UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Data-driven Health Assessment of Levee Systems for Flood Resiliency

Permalink

https://escholarship.org/uc/item/1tj068vz

Author

Russo, Brittany Monterey

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Data-driven Health Assessment of Levee Systems for Flood Resiliency

By

Brittany Monterey Russo

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Civil and Environmental Engineering

in the

Graduation Division

of the

University of California, Berkeley

Committee in charge:

Professor Adda Athanasopoulos-Zekkos, Chair
Professor Kenichi Soga
Professor Iryna Dronova

Summer 2024

Data-driven Health Assessment of Levee Systems for Flood Resiliency

© Copyright 2024
Brittany M. Russo
All rights reserved

Abstract

Data-driven Health Assessment of Levee Systems for Flood Resiliency

by

Brittany Monterey Russo

Doctor of Philosophy in Engineering - Civil and Environmental Engineering

University of California, Berkeley

Professor Adda Athanasopoulos-Zekkos, Chair

Levees are complex earthen structures used for flood control management. With over 24,000 miles of levees in the United States, the performance of these structures is crucial as they put over 23 million people at risk if they fail. The American Society of Civil Engineers rates levees based on their current performance, consistently assigning them a grade of "D," indicating significant deterioration and a strong risk for failure.

One major contributing factor to the poor performance of levees is the current inspection practices in the United States, which primarily rely on visual assessments of the surface with sparse, discrete instrumentation providing a data point in the subsurface at the point of installation. Another challenge in inspection is the lack of standardization among different organizations, leading to weak contacts in adjacent levee segments. Due to the spatial extent of levees, the temporal frequency of inspection is not sufficient to ensure the health of these systems. Given these challenges, there is a need for improved levee inspection practices to assess the surface and subsurface health.

This resiliency monitoring framework for levee health maintenance leverages a combination of non-invasive and continuous UAV-enabled sensors and geophysical techniques to incorporate surface and subsurface health information. Data were collected along three stretches on the Grand Island levee system in the Sacramento-San Joaquin Delta during the 2022 dry season and the 2023 wet season. UAV-enabled optical imagery, thermal imagery, and LIDAR were collected at all three sites, along with two geophysical techniques: the Multichannel Analysis of Surface Waves (MASW) to capture the shear wave velocity distribution of the subsurface and Electromagnetic Induction (EMI) to capture the electrical resistivity distribution of the subsurface.

This framework utilizes remote sensing change detection techniques on the surface data to understand how the health of the system changes over time using image differencing and classification methods. Change detection was applied to the subsurface data, along with unsupervised clustering and anomaly detection algorithms, to identify potentially problematic zones in the subsurface that could be attributed to internal failure mechanisms. The framework fuses the surface and subsurface health information into hazard maps, revealing increased areas of hazard on the levee. This information justifies locations for more detailed investigation and aids stakeholders in decision-making for maintenance and future rehabilitation.

To Opa and Nonno

Table of Contents

Abstract 1
Table of Figures v
Table of Tables xiii
List of Symbols xv

1. Introduction 1
1.1 Research Motivation 1
1.2 Research Objectives 5
1.3 Limitations 5
1.4 Dissertation Structure 6

2. Laboratory Testing 7
2.1 Laboratory Testing Introduction 7
2.2 Literature Review 7
2.3 Experimental Testing Design 11
2.3.1 Materials 11
2.3.2 Sample Preparation 16
2.4 Soil Cell Design 17
2.4.1 Bender Elements 18
2.4.2 Electrodes 24
2.5 Laboratory Investigation Results 25
2.5.1 Shear Wave Velocity Results 27
2.5.2 Electrical Resistivity Results 30
2.5.3 Validation of Laboratory Shear Wave Velocity 32
2.5.4 Validation of Laboratory Electrical Resistivity 33
2.6 The Importance of Geotechnical Conditions 36
2.7 Summary of Laboratory Results for Field Profile Interpretation 36

3. Field Data Collection and Remote Sensing Analysis 38
3.1 Geophysical and Remote Sensing Literature Review 38
3.1.1 Geophysical Methods for Subsurface Investigation 38
3.1.2 Remote Sensing for Surface Investigation 42
3.1.3 Imaging the Surface and Subsurface 45
3.2 Field Testing Locations 46
3.2.1 Grand Island, California 48
3.2.2 Bacon Island, California 51
3.3 Geophysical Data Collection 53
3.3.1 Multichannel Analysis of Surface Waves 53
3.3.2 Electromagnetic Induction 64
3.3.3 Electrical Resistivity Tomography 73
3.3.4 Comparison of Laboratory Electrical Resistivity to Field Electrical Resistivity 75
3.4 Unmanned Aerial Vehicle Mounted Sensors 77
3.4.1 Structure-From-Motion 79

3.4.2 Red, Green, Blue Imagery (Optical Imagery) 80
3.4.3 Thermal Imagery 82
3.4.4 LIDAR 83
3.5 Change Detection 85
3.5.1 Change Detection of the Subsurface 86
3.5.2 Change Detection of the Surface 88

4. Machine Learning Analysis 106
4.1 Machine Learning Literature Review 106
4.1.1 Machine Learning in Geotechnical Engineering 106
4.1.2 Machine Learning and Geophysics 108
4.1.3 Machine Learning for Levees 110
4.2 Machine Learning Introduction 111
4.3 Feature Engineering 113
4.3.1 Standardization and Normalization 113
4.3.2 Feature Selection 114
4.3.3 Resampling 114
4.4 Unsupervised Machine Learning 115
4.4.1 Clustering Algorithms 116
4.4.2 Clustering Model Parameters 121
4.4.3 Anomaly Detection 126
4.4.4 Anomaly Detection Model Parameters 129
4.5 Supervised Machine Learning 132
4.5.1 Supervised Algorithms 132
4.5.2 Hyperparameter Optimization 135
4.5.3 Validation Methods 137
4.6 Unsupervised Machine Learning Analysis on Grand Island 138
4.6.1 Unsupervised Feature Engineering 138
4.6.2 Unsupervised Clustering Results 141
4.6.3 Potentially Problematic Zones Using Clusters 152
4.7 Anomaly Detection Results 153
4.7.1 Factors Affecting the Contamination Factor 154
4.7.2 Anomaly Detection Analysis 157
4.8 Combination of Anomaly Detection and Clustering 164
4.8.1 Advantages and Disadvantages of Clustering 164
4.8.2 Advantages and Disadvantages of Anomaly Detection 165
4.8.3 Framework for Understanding Subsurface Levee Health 165
4.8.4 Subsurface Validation at Grand Island 170
4.9 Supervised Machine Learning Analysis on Bacon Island 170
4.9.1 Data and Feature Engineering 172
4.9.2 Analysis #1 - Predicting Individual SBTn 178
4.9.3 Predicting SBTn of a Site 187
4.9.4 Feature Selection for Algorithm Performance 191
4.9.5 Supervised Machine Learning Results Discussion 193
4.10 Supervised Compared with Unsupervised 194
4.10.1 Unsupervised Machine Learning Results for Bacon Island 194

4.10.2 Subsurface Anomalous Zones Compared to SBTn Distribution 198
4.10.3 Unsupervised Results Compared to Supervised Results (Using all Features) 202
4.10.4 Unsupervised and Supervised Machine Learning for Levee Subsurface Health Investigation 203

5. Data Fusion and Hazard Maps 205
5.1 Robust Levee Health Framework 205
5.2 Projection of Subsurface Data to the Surface 205
5.2.1 Change Detection Analysis Projection 205
5.2.2 Potentially Problematic Zones Projection 206
5.3 Model Layers 208
5.3.1 Surface Layers 208
5.3.2 Subsurface Layers 209
5.4 Model Weights 211
5.4.1 Surface Weights 211
5.4.2 Subsurface Weights 216
5.5 Building the Model 219
5.6 Comparison of Hazard Maps 222
5.6.1 Comparison Between Clustering Techniques 222
5.6.2 Comparison Between Compilation Datasets 225
5.7 Hazard Model Discussion and Conclusion 227

6. Conclusions and Recommendations for Future Research 230
6.1 Major Contributions of this Research 230
6.1.1 Development of Library for Shear Wave Velocity and Electrical Resistivity 230
6.1.2 Field Data Collection 231
6.1.3 Change Detection Analysis of Surface and Subsurface Data 231
6.1.4 Machine Learning Analysis of Geophysical Data of Levee Subsurface 231
6.1.5 Hazard Map 232
6.2 Recommendations for Future Research 233
References 235

Appendix A – Laboratory Data 250
Appendix B – Geotechnical Borings and CPTs 280
Appendix C – Shear Wave Velocity and Electrical Resistivity Profiles 315
Appendix D – Remote Sensing Profiles 374
Appendix E – Change Detection Profiles 420
Appendix F – Machine Learning Model Parameters 468
Appendix G – Clustering Profiles ..... -- 473
Appendix H – Anomaly Detection 554
Appendix I – Clustering and Anomaly Detection Fusion 701
Appendix J – Supervised Machine Learning 794
Appendix K – ModelBuilder Models 798
Appendix L – Hazard Maps and Comparisons 809

Table of Figures

Figure 1.1 Typical Structure of a Levee 1

Figure 1.2 Typical Levee Failure Mechanisms (ASDSO, 2020) 3

Figure 2.1 Gradation of curves of soils tested in the laboratory (a) quartz grained soils, (b) Topatopa soils, and (c) New Zealand soils 13

Figure 2.2 Materials tested (a) Crushed angular soil, (b) C109 Sand, (c) Nevada Sand, (d) C109-Nevada Mixture, (e) Topatopa Sand with organics, (f) Topatopa Sand with no-organics, (g) New Zealand - Fine, (h) New Zealand - Coarse, and (i) New Zealand Well Graded..... 15

Figure 2.3 Soil testing cell design schematic showing the (a) side view and the (b) plan view.. 18

Figure 2.4 Fabrication of bender elements (a) soldering wire to the front of the bender element, (b) soldering wire to the back of the bender element, (c) silver paint coating, and (d) epoxy and polyurethane coating 19

Figure 2.5 Laboratory setup showing the soil testing cell, dial gauge, oscilloscope, waveform generator, and LCR meter 20

Figure 2.6 Resonant frequency plot ranging from 1 kHz to 25 kHz..... 21

Figure 2.7 First arrival of received wave compared to input wave..... 22

Figure 2.8 Cross-correlation results showing how maximum peak corresponds to first arrival of the received wave compared to the input wave 23

Figure 2.9 Frequency sweep results of resistance with respect to frequency for three NaCl solutions showing optical frequency range between 100 Hz and 10,000 Hz 24

Figure 2.10 Geometric factor trendline to account for changes in sample height during testing 25

Figure 2.11 Shear wave velocity results between Nevada Sand and the C109-Nevada Mixture at (a) w=1%, Dr=30%, (b) w=5%, Dr30%, (c) w=10%, Dr=30%, (d) w=20%, Dr=30%, (e) w=1%, Dr=50%, (f) w=5%, Dr=50%, (g) w=10%, Dr=50%, (h) w=20%, Dr=50%, (i) w=1%, Dr=70%, (j) w=5%, Dr=70%, (k) w=10%, Dr=70%, (l) w=20%, Dr=70% 26

Figure 2.12 Showing the results of the shear wave velocity laboratory testing for each geotechnical condition 28

Figure 2.13 Showing the results of the electrical resistivity laboratory testing for each geotechnical condition 31

Figure 2.14 Ottawa C109 Sand results compared to Robertson et al. 1995 for validation of testing procedure (a) w = 0%, (b) w = 1%, (c) w = 5%, (d) w = 10%, and (e) w = 20%..... 32

Figure 2.15 Literature data with specified water contents 33

Figure 2.16 Compilation of electrical resistivity data from literature..... 34

Figure 2.17 Comparison of electrical resistivity results with Pandey et al. (2015) with (a) Topatopa without organics, (b) fine New Zealand Sand, and (c) coarse New Zealand Sand..... 35

Figure 3.1 Location of the Sacramento-San Joaquin Delta 47
Figure 3.2 Data collection sites along the Grand Island levees 49
Figure 3.3 Locations of field testing sites on Bacon Island 51
Figure 3.4 MASW setup with (a) geophones, (b) hammer, (c) Geode and (d) field laptop..... 56
Figure 3.5 Field photos of active MASW surveys for (a) crest survey in the dry season at Site 1 and (b) crest survey in the wet season at Site 2 59
Figure 3.6 MASW geophone and shot array geometry. 60
Figure 3.7 MASW survey sites on Grand Island - Site 1 61
Figure 3.8 MASW survey sites on Grand Island - Site 2 61
Figure 3.9 MASW survey sites on Grand Island - Site 3 62
Figure 3.10 Sample MASW Data showing (a) wave arrival time series, (b) dispersion curves, and (c) 2D shear wave velocity profile. 64
Figure 3.11 EMI sensor showing the eddy currents and the primary and secondary electromagnetic fields 65
Figure 3.12 GEM-2 sensor connected to Reach RS2 GPS and Trimble Nomad data logger 68
Figure 3.13 EMI surveys on Grand Island for (a) Site 1, (b) Site 2, and (c) Site 3 70
Figure 3.14 Final electrical resistivity inversion profile 73
Figure 3.15 Electrical Resistivity Tomography electrode configurations 74
Figure 3.16 Berkeley Hills precipitation and data collection (modified from NOAA, 2023). 76
Figure 3.17 Field and laboratory comparison of electrical resistivity 76
Figure 3.18 DJI Matrice 300 with (a) equipped with the thermal camera and (b) equipped with the LIDAR sensor. 78
Figure 3.19 Structure-from-Motion workflow for 3D model construction (Carrivick, Smith, & Quincey, 2016) 79
Figure 3.20 Electromagnetic spectrum showing wavelengths of the remote sensing sensors..... 80
Figure 3.21 Examples of Grand Island Site 2 showing the (a) 3D textured mesh and (b) 3D thermal signature. 82
Figure 3.22 3D LIDAR point cloud of Site 2 on Bacon Island. 85
Figure 3.23 Comparison of (a) average temperature, (b) precipitation and (c) wind speed in the Sacramento-San Joaquin Delta during the months of data collection. 86
Figure 3.24 MASW subsurface profile showing change detection results with the grey points representing no meaningful change, the green points representing areas with above 10% difference, and the red points representing areas with below -10% difference. 87
Figure 3.25 Change Detection Workflow (Lu, Mausel, Brondizio, & Moran, 2004). 88
Figure 3.26 Classification of Site 2 at Grand Island during the Dry Season showing (a) the optical imagery and (b) the classification results. 95

Figure 3.27 Classification of Site 2 at Grand Island during the Wet Season showing (a) the optical imagery and (b) the classification results. 96

Figure 3.28 Classification Change Detection Results for Site 2 on Grand Island showing (a) where no change occurred, (b) where vegetation was classified as dead in the dry season to green in the wet season, and (c) where vegetation was classified as green in the dry season and dead in the wet season. 97

Figure 3.29 Reflectance of vegetation, bare soil, and water with respect to the red, green, and blue optical bands (Modified from Wilkie and Finn, 1996). 98

Figure 3.30 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season, (b) optical imagery from the wet season, and (c) difference in the red band. 99

Figure 3.31 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season, (b) optical imagery from the wet season, and (c) difference in the green band. 100

Figure 3.32 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season, (b) optical imagery from the wet season, and (c) the difference in the Normalized Green-Red Difference Ratio between the dry and wet season. 101

Figure 3.33 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season with bare earth locations and (b) optical imagery from the wet season with bare earth locations. 102

Figure 3.34 Example from Site 2 at Grand Island showing (a) the temperature difference between the dry and wet seasons and (b) the difference in the Normalized Relative Canopy Temperature between the dry and wet seasons. 103

Figure 4.1 Supervised Machine Learning Workflow. 112

Figure 4.2 Unsupervised Machine Learning Workflow. 112

Figure 4.3 Resampling techniques for imbalanced datasets (a) up-sampling and (b) down-sampling. 115

Figure 4.4 K-Means Algorithm (1) Randomly initiate centroids (2) assign points to their closest centroid (3) calculate the mean of the cluster (4) update new centroid location to mean location. 116

Figure 4.5 Agglomerative Hierarchical Clustering Algorithm (1) initial dataset (2) finds the two closest points (3) merges the two closest points into one cluster until one cluster remains. 118

Figure 4.6 Example of possible dendrograms from Hierarchical Clustering (a) shows the points in the dataset (b) assigned “labels” to each point in the dataset for dendrogram viewing (c) example of a symmetric dendrogram (d) example of an asymmetric dendrogram where the blue line represents three clusters and the orange line represents four clusters. 119

Figure 4.7 DBSCAN Algorithm (a) initial dataset (b) applying circles of radius, ε, around each point (c) the black points have no points in their radius and are classified as outliers (d) the colored points are core points and have the minimum number of points within their radius (e) the colored points are border points and (f) showing the clustered data points as red and outliers as black. 120

Figure 4.8 Elbow Rule for determining k for K-Means and Hierarchical clustering Analysis – optimal number of clusters is four. 122

Figure 4.9 Elbow Rule for determining ε for DBSCAN – optimal choice for ε is 1. 122

Figure 4.10 Silhouette Rule for determining k for K-Means and Hierarchical Clustering – optimal was chosen as two clusters. 123

Figure 4.11 Gap Statistics using the Gap Value. Eight clusters maximize the Gap Value. 124

Figure 4.12 Gap Statistics using the Diff Value. The smallest k with a positive diff value is three clusters. 125

Figure 4.13 Gap Statistics using the Gap* Value. Three clusters maximize Gap*. 126

Figure 4.14 Isolation Forest concept showing anomaly and tree leaves. 127

Figure 4.15 Local Outlier Factor showing the reachability density of a point (O). 128

Figure 4.16 Isolation Forest anomaly detection using contamination factor as "auto" showing (a) the shear wave velocity distribution, (b) the electrical resistivity distribution, and (c) the spatial distribution of anomalies. 130

Figure 4.17 OCSVM anomaly classification results (a). classification results using γ = "scale" (b). results from (a) overlain on the MASW profile (c). Classification results using γ = "auto" (d). Results from (c) overlain on the MASW profile. 131

Figure 4.18 Supervised Machine Learning steps. 132

Figure 4.19 K-Nearest Neighbors Algorithm (a) the dataspace, (b) polygons of closeness around each data point, (c) nonlinear decision boundary around the two classes, (d) new data point is introduced and showing the polygon it falls into, (e) classifying the new datapoint based on the closest point. 133

Figure 4.20 Random Forest depiction with multiple trees and taking the majority vote for the final classification. 134

Figure 4.21 Support Vector Machine (a) hyperplane with a small margin (b) hyperplane with a larger margin. 135

Figure 4.22 5-Fold Cross Validation (a) training dataset is split into five folds and parameters are iterated (b) plot showing the parameter vs. accuracy and picking the parameter with the best accuracy. 136

Figure 4.23 Uniform grid of points for machine learning. 138

Figure 4.24 K-Means Clustering results (a) using the raw data, (b) using the standardized data, (c) using the normalized data. 142

Figure 4.25 DBSCAN results (a) using the raw data, (b) using the standardized data, (c) using the normalized data. 142

Figure 4.26 Exploration of data (a) showing the data as a function of depth below ground surface, (b) showing data by x-coordinate of profile, and (c) K-Means Clustering results using 7 clusters. 143

Figure 4.27 Examples of profiles where (a) the data was not clustered, (b) the data was clustered based upon Vs, (c) the data was clustered based upon ρ, and (d) the data was clustered by Vs and ρ..... 145

Figure 4.28 Clustering labels as a function of shear wave velocity and electrical resistivity.... 147

Figure 4.29 Subsurface profiles from the dry season on the crest at Site 2 showing (a) the shear wave velocity distribution and (b) the electrical resistivity distribution..... 148

Figure 4.30 Subsurface profiles from the dry season on the crest at Site 2 showing (a) clustering profile from the Crest compilation dataset, (b) clustering profiles from the Dry compilation dataset, (c) clustering profile from the Dry Site 2 compilation dataset, (d) clustering profile from the Site 2 compilation dataset, and (e) clustering profile when clustering on the individual profile. 149

Figure 4.31 Number of clusters with respect to the normalized shear wave velocity and electrical resistivity for all profiles in a compilation dataset clustered with (a) K-Means Clustering and (b) Hierarchical Clustering. 150

Figure 4.32 Interpretability of potentially problematic zones for clustering showing (a) shear wave velocity subsurface profile, (b) electrical resistivity subsurface, (c) clustering with respect to shear wave velocity and electrical resistivity, and (d) clustering distribution in the subsurface. 153

Figure 4.33 Effect of shear wave velocity and electrical resistivity on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation..... 154

Figure 4.34 Effect of site on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation. 155

Figure 4.35 Effect of season on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation. 156

Figure 4.36 Effect of location on the levee on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation..... 157

Figure 4.37 Effect of contamination factor on performance for (a) fitted on compilation dataset and predicted on individual profiles for Isolation Forest, (b) fitted and predicted on compilation dataset for Isolation Forest, (c) fitted on compilation dataset and predicted on individual profiles for One-Class Support Vector Machines, and (d) fitted and predicted on compilation dataset for One-Class Support Vector Machines. 160

Figure 4.38 Effect of shear wave velocity and electrical resistivity on performance for (a) fitted on compilation dataset and predicted on individual profiles for Isolation Forest, (b) fitted and predicted on compilation dataset for Isolation Forest, (c) fitted on compilation dataset and predicted on individual profiles for One-class Support Vector Machines, and (d) fitted and predicted on compilation dataset for One-class Support Vector Machines. 163

Figure 4.39 Clustering example (a) with respect to shear wave velocity and electrical resistivity and (b) subsurface profile showing the spatial distribution of clusters 164

Figure 4.40 Subsurface profile showing anomaly detection results as anomalous zones..... 165

Figure 4.41 Example of thresholds for determining if an anomalous point is spatially significant when (a) N = 4 and M = 1, (b) N = 4 and M = 3, (c) N = 8 and M = 4, and (d) N = 8 and M = 6 166

Figure 4.42 Subsurface profile combining the clustering results with the anomaly detection results showing clusters in the anomalous zones. 167

Figure 4.43 Subsurface profile showing the anomalous zones as a function of the relative change in shear wave velocity and electrical resistivity. 169

Figure 4.44 Subsurface profile showing the anomalous zones as a function of geotechnical conditions controlled for in the laboratory investigation. 170

Figure 4.45 Longitudinal profile of the Northwest Site on Bacon Island. 172

Figure 4.46 Longitudinal profile of the Northcentral Site on Bacon Island. 173

Figure 4.47 Longitudinal profile of the South Site on Bacon Island. 173

Figure 4.48 Cross-sectional profile of the Northwest Site on Bacon Island. 174

Figure 4.49 Cross-sectional profile of the Northcentral Site on Bacon Island. 175

Figure 4.50 Cross-sectional profile of the South Site on Bacon Island. 176

Figure 4.51 CPT Profiles at Bacon Island for (a) Northcentral site, (b) Northwest site, and (c) South site. 177

Figure 4.52 Violin plots showing the distribution at each site of (a) shear wave velocity and (b) electrical resistivity. 180

Figure 4.53 Correctly Predicted data (black circles) and incorrectly predicted data (red x's) as a function of shear wave velocity and electrical resistivity for (a) K-Nearest Neighbors, (b) Random Forest, and (c) Support Vector Machines. 181

Figure 4.54 Normalized Classification Performance of each SBTn corresponding to abundance of SBTn in the Training Dataset. 182

Figure 4.55 Normalized Classification Performance of each SBTn of the baseline analysis (Analysis #1), stratified randomization (Analysis #2), and up/down sampling (Analysis #3) to look at the effect on class imbalance. Areas shaded in green represent low class representation. 183

Figure 4.56 Showing the effect of scoring metric on overfitting with respect to the number of neighbors for (a) the baseline analysis and (b) the sampled analysis using 500 samples per class. 185

Figure 4.57 Showing the effect of scoring metric on overfitting with respect to the number of samples per class for (a) when k = 2 and (b) the k = 4. 186

Figure 4.58 Distribution of (a) shear wave velocity and (b) electrical resistivity in each dataframe for predicting SBTn at a site. 188

Figure 4.59 Spatial distribution of correct (grey) and incorrect (red) SBTn classification using (a) K-Nearest Neighbors, (b) Random Forest, (c) Support Vector Machines, and (d) a combination of the three algorithms. 190

Figure 4.60 Spatial distribution of correct and incorrect classification looking at all three algorithms for (a) the Northwest site and (b) the South site. 191

Figure 4.61 The distribution of (a) shear wave velocity and (b) electrical resistivity at the Northcentral Site for data that was correctly and incorrectly classified by the three algorithms (SVM, RF, KNN). 191

Figure 4.62 Feature Importance from Random Forest. 192

Figure 4.63 Overfitting Comparison between the Baseline Analyses (Analyses #5-7) and the Feature Importance Analyses for Predicting Site SBTn (Analyses #8-10). 193

Figure 4.64 Subsurface profiles for the Northwest Site showing potentially problematic zones using (a) K-Means Clustering and (b) Hierarchical Clustering. 195

Figure 4.65 Subsurface profiles for the Northcentral Site showing potentially problematic zones using (a) K-Means Clustering and (b) Hierarchical Clustering. 196

Figure 4.66 Subsurface profiles for the South Site showing potentially problematic zones using (a) K-Means Clustering and (b) Hierarchical Clustering. 197

Figure 4.67 Distribution of SBTn for Northwest Site for potentially problematic zones (a) coarser particles (higher WCs), more well graded (finer), (b) rounder particles, coarser particles (lower WCs), (c) high Vs, similar resistivity, (d) low Vs, similar resistivity, and (e) normal – not at a potentially problematic zone. 199

Figure 4.68 Distribution of SBTn for Northcentral Site for potentially problematic zones (a) coarser particles (higher WCs), more well graded (finer), (b) finer particles (higher WCs), more poorly graded (coarser), (c) high Vs, similar resistivity, (d) low Vs, similar resistivity, and (e) normal – not at a potentially problematic zone. 200

Figure 4.69 Distribution of SBTn for South Site for potentially problematic zones (a) coarser particles (higher WCs), more well graded (finer), (b) low Vs, similar resistivity, and (c) normal – not at a potentially problematic zone. 201

Figure 5.1 Projection of subsurface change detection to the surface. 206

Figure 5.2 Projection of subsurface machine learning analysis to the surface showing (a) the subsurface profile, (b) the binary classification with respect to the x-coordinate for all labels, and (c) the projected data at the surface. 207

Figure 5.3 Thermal change for Site 1 showing (a) distribution of thermal differences, (b) 1 standard deviation of change, and (b) 2 standard deviation of change. 213

Figure 5.4 Thermal change for Site 2 showing (a) distribution of thermal differences, (b) 1 standard deviation of change, and (b) 2 standard deviation of change. 214

Figure 5.5 Thermal change for Site 3 showing (a) distribution of thermal differences, (b) 1 standard deviation of change, and (b) 2 standard deviation of change. 215

Figure 5.6 Hazard map example with (a) shadows and (b) without shadows for Site 2 on Grand Island. 216

Figure 5.7 Model processing steps for calculating the hazard weight. 220

Figure 5.8 Hazard Map of Site 1 on Grand Island showing locations of potential hazards. 222

Figure 5.9 Comparison between the hazard maps for the Crest and Toe Compilation datasets at Site 2 showing (a) the distribution of hazard weights and (b) spatial difference in hazard weights. 223

Figure 5.10 Variability in Count for Hazard Weights between Compilation Profiles for Site 1 showing (a) Hierarchical Clustering and (b) K-Means Clustering. 225

Figure 5.11 Variability in Count for Hazard Weights between Compilation Profiles for Site 1 showing (a) Hierarchical Clustering and (b) K-Means Clustering. 226

Figure 5.12 Variability in Count for Hazard Weights between Compilation Profiles for Site 1 showing (a) Hierarchical Clustering and (b) K-Means Clustering. 226

Table of Tables

Table 1.1 Typical Levee Failure Mechanisms 2

Table 2.1 Soil Properties and Classification 15
Table 2.2 Specific Surface and Capillary Force 29
Table 2.3 Summary Results of Laboratory Shear Wave Velocity and Electrical Resistivity 37

Table 3.1 Soil Behavior Index (Robertson & Cabal, 2014) 53
Table 3.2 NEHRP Site Classification (Building Seismic Safety Council, 2004) 56
Table 3.3 MASW Acquisition Software Parameters 58
Table 3.4 Coil Configuration Setup Parameters 71
Table 3.5 Iterations of EMI Inversions 72
Table 3.6 Number of Images Collected per Survey 81
Table 3.7 Ellipsoid – GEOID18 Height Conversion 84
Table 3.8 Site 1 Dry Season Classification Accuracy 90
Table 3.9 Site 2 Dry Season Classification Accuracy 90
Table 3.10 Site 3 Dry Season Classification Accuracy 91
Table 3.11 Site 1 Wet Season Classification Accuracy 91
Table 3.12 Site 2 Wet Season Classification Accuracy 91
Table 3.13 Site 3 Wet Season Classification Accuracy 92
Table 3.14 Site 1 Classification Change Detection 93
Table 3.15 Site 2 Classification Change Detection 93
Table 3.16 Site 3 Classification Change Detection 93

Table 4.1 Feature Engineering Methods 113
Table 4.2 Disadvantages to K-Means Clustering 117
Table 4.3 Hyperparameters for each Algorithm for Optimization 136
Table 4.4 Compilation Datasets 139
Table 4.5 Optimal Hyperparameters for Feature Engineering Methods 141
Table 4.6 Percentages of Clustering Label per Compilation Dataset 145
Table 4.7 Difference in Clusters between K-Means and Hierarchical Clustering 151
Table 4.8 Percentage of Profiles Classified in Each Performance Category 158
Table 4.9 Performance of Predicting on Individual and Compilation Datasets 158
Table 4.10 Effect of Site on Anomaly Detection Performance 159
Table 4.11 Effect of Profile Location on Anomaly Detection Performance 161
Table 4.12 Effect of Season on Anomaly Detection Performance 162
Table 4.13 Shear Wave Velocity and Electrical Resistivity Change for Case Study 1 168
Table 4.14 Shear Wave Velocity and Electrical Resistivity Change for Case Study 2 and Case Study 3 168
Table 4.15 Supervised Machine Learning Analyses Accuracy Results 171
Table 4.16 Summary of Normalized Correct Percentages per Algorithm per Site 179
Table 4.17 Summary of SBTn performance for each site 179

Table 4.18 Percent Difference between Training and Testing Datasets for Predicting Each Site 188
Table 4.19 Optimal Hyperparameters for Bacon Island 194
Table 4.20 Percent Correctly Predicted SBTn per Potentially Problematic Zone for the Northwest Site for all Features 202
Table 4.21 Percent Correctly Predicted SBTn per Potentially Problematic Zone for the Northcentral Site for all Features 203
Table 4.22 Percent Correctly Predicted SBTn per Potentially Problematic Zone for the South Site for all Features 203

Table 5.1 Summary of Surface Layers Weights 211
Table 5.2 Machine Learning Layers Base Weights 217
Table 5.3 Machine Learning Analysis Profile Weights 217
Table 5.4 Change Detection Profile Weights 218
Table 5.5 Change Detection MASW/EMI Weights 219

List of Symbols

Symbol	Description	Units
Vs	Shear wave velocity	Meter/second
ρ	Electrical resistivity	Ωm (ohm-meter)
t	Time	Seconds
Cc	Coefficient of curvature	-
Cu	Coefficient of uniformity	-
γ_{max}	Maximum unit weight	Kilonewton/meter^3
γ_{min}	Minimum unit weight	Kilonewton/meter^3
D50	The average particle size of the soil	Millimeter
w	Water content	%
wm	Weight of moist soil	Grams
wd	Weight of dry soil	Grams
Dr	Relative Density	%
γ_d	Dry unit weight	Kilonewton/meter^3
γ_{moist}	Moist unit weight	Kilonewton/meter^3
d_{moist}	Density of the moist soil	Grams/centimeter^3
g	Gravitational acceleration	Meter/second^2
M	Mass of the soil	Grams
V	Volume of soil	Meter^3
e_{max}	Maximum void ratio	-
e_{min}	Minimum void ratio	-
L	Tip-to-tip distance between bender elements	Meter
Δt	Travel time difference between bender elements	Second
$\varphi_{fg}(t)$	Cross-correlation in the time domain	-
f(t)	Signal 1 in the time domain	-
g(t)	Signal 2 in the time domain	-
τ	Time shift in the time domain	Second
F(f)	Signal 1 in the frequency domain	-
G(f)	Signal 2 in the frequency domain	-
k	Geometric factor	Meters
R	Resistance	Ω (ohm)
h	Height of soil	Inches
Ss	Specific surface	Meter^2/grams
Pc	Capillary pressure	Kilopascal
dw	Density of water	Gram/centimeter^3
Gs	The specific gravity of the soil	-
Ts	Interfacial tension	Newton/meter
dm	Mineral density	Kilogram/meter^3
e	Void ratio	-
K	Hydraulic conductivity	Centimeter/second
C	Empirical constant	Meter/second
D_{10}	The grain size of the soil where 10% of the soil is finer	Millimeter
μ	Soil permeability	Centimeter/second

p	The porosity of the soil	-
Qt	Cone tip resistance	-
FR	Friction ratio	%
qt	Total cone resistance	Megapascal
σ_{vo}	Total vertical stress	Kilopascal
σ'_{vo}	Effective vertical stress	Kilopascal
fs	Sleeve friction	Megapascal
Ic	Soil behavior index type	-
SBTn	Normalized soil behavior index	-
Qtn	Normalized cone resistance	-
n	Stress exponent for CPT	-
pa	Atmospheric pressure	Kilopascal
α	Shear modulus	Kilopascal
Vs30	Shear wave velocity in the upper 30 m	Meter/second
Z_{max}	Maximum depth of investigation	Meter
λ	MASW array length	Meter
E	Electric field	Newton/Coulomb
B	Magnetic field	Tesla
ξ	Magnetic permeability	Henries/meter
ψ	Electrical conductivity	Siemens/meter
η	Dielectric permittivity	Farads/meter
δ	Skin depth	Meters
ω	Angular frequency	Radians
f	Frequency	Hertz
π	Pi (3.41)	-
ς	Radial distance at the center of the transmitter coil	Meters
φ	Phase shift	Radians/degrees
i	Current	Amperes
Λ	Voltage	Volts
l	Length	Meter
A	Cross-sectional area	$Meters^2$
a	Spacing between electrodes for Wenner Array	Meters
N	Distance of potential electrodes from center of array for Schlumberger array	Meters
J	Distance of current electrodes from center of array for Schlumberger array	Meters
n	Spacing between current and potential electrodes for dipole-dipole array	Meters
sf	Scale factor	-
fl	Focal length	Millimeters
fh	Flying height	Meters
T_{rad}	Radiant temperature	°Celsius/°Fahrenheit
T_{kin}	Kinetic Temperature	°Celsius/°Fahrenheit
ϵ	Emissivity	-
c	Speed of light	Meters/second
dist	Distance between LIDAR sensor and object	Meters

Gr	Green band digital number	-
Rd	Red band digital number	-
T	Surface temperature	°Celsius/°Fahrenheit
T_{min}	Minimum surface temperature	°Celsius/°Fahrenheit
T_{max}	Maximum surface temperature	°Celsius/°Fahrenheit
s	Shear strength	Kilopascals
c’	Cohesion	Kilopascals
C'_R	Apparent cohesion	Kilopascals
σ	Normal vertical stress	Kilopascals
u	Pore-water pressure	Kilopascals
ϕ'	Effective friction angle	Degrees
$\overline{X}$	Standardized data	-
X	Raw data	-
std	Standard deviation of data	-
X_{min}	Minimum value in the dataset	-
X_{max}	Maximum value in the dataset	-
E	Euclidean distance	Distance
S(i)	Silhouette coefficient	-
W_k	Within cluster dispersion	-
γ	Kernel coefficient for OCSVM	-
ν	Contamination factor for OCSVM	-

Acknowledgments

First, I would like to express my gratitude to Dr. Adda Athanasopoulos-Zekkos, my Ph.D. advisor. Your mentoring and expertise have truly guided me through my time at the University of California, Berkeley. I am honored to be your first Ph.D. student at Berkeley, and you have taught me how to be a good mentor to graduate students. Your advice and feedback on my research and presentations have truly been instrumental in my success as a graduate student.

I would also like to thank my Dissertation Committee, Dr. Kenichi Soga and Dr. Iryna Dronova, for their guidance and feedback on my project. Additionally, I would like to thank my Qualifying Exam Committee, Dr. Kenichi Soga, Dr. John Radke, and Dr. Marta Gonzalez for their expertise and insight into the different components of my research. I would like to thank all the faculty at UC Berkeley and the University of Waterloo who have written letters of recommendation for me: Robert Kayen, Jonathon Bray, Tony Endres, Bill Annable, and Rohan Jayasundera. I would not be where I am today without all your support and encouragement.

I am deeply grateful to Robert Kayen for his mentorship through my time at UC Berkeley. It started in the master's program while he encouraged me to talk about Stereonets in more detail than the class entailed and continued into the four terms where I was his GSI. Thank you for giving my ideas merit and letting me experiment with the Engineering Geology labs and giving me an active role in teaching. Thank you for always having an open door, chocolate, and a welcoming presence. The experience and tools I learned from you make me a better teacher and mentor.

I would also like to thank Dr. Michael Riemer and Dr. Jongchan Kim for their support and guidance in the laboratory. Additionally, I would like to acknowledge all the graduate students who joined me at Grand Island during sun, rain, and hail to collect field data on levees: Parker Blunts, Drew Gomberg, Jhih-Rou Huang, Weibing Gong, Gabriela Paredes, Satuk Sari, Fernanda Corvillon, Ashely Woodward, Leslie Ramirez, and Zi-Xiang Gng. My field work for my Ph.D. would not have been possible without Reclamation District 3 and MBK Engineers who provided me with access to my field sites and additional subsurface information. I would like to acknowledge the United States Society on Dams for awarding my scholarship funds in 2022 to perform field work and GreenValley International for their guidance on LIDAR data processing.

I would not be here today without all the high school teachers and professors who recognized my potential even before I did. Their belief in me even when I struggled through classes, homework, and exams proved to me that I could do it.

This degree would not have been possible without my friends and family as my support system. To my friends in the Bay Area, Paige, Andrew, Parker, Joel, Margaret, Jon, Sydney, Leslie, Jaewon and Sarah, who helped keep my spirits up and got me away from the computer. And to my friends in Canada, Morgan, Taskia, and Mike, who always gave me support, even from another country. I would not have finished without my family. My siblings, Kayla, Anastasia, Alyssa, Geoffery, and Courtney have been with me throughout this entire journey and their unwavering support and numerous phone calls have been instrumental in the success of this degree. I cannot express how

grateful I am to have amazing parents who always supported me and encouraged me to move to Canada for undergrad and to move to Berkeley for graduate school. They have been my rock and greatest supporters. I also want to thank my grandma for always having unwavering optimism and believing in my success. Thank you for the weekly FaceTime chats and always putting a smile on my face.

Lastly, to my dog Hades, you have truly been the light during my Ph.D. The early morning cuddles, sitting on my lap during lectures, and curling up in the corner of my desk have brought great comfort to me and I am glad you came into my life when you did.

Chapter 1. Introduction

1.1 Research Motivation

Levees are earthen embankments used for flood control management. They are typically found along riverbanks to prevent flooding and can span many kilometers in length. When engineered, they are designed to withstand high-water storm events while considering the land use behind the levee. Many levees are not engineered. They can be built by landowners using native materials to protect their land or they can be made naturally over time through multiple flooding events where sediment is deposited along riverbanks where the velocity of water is lower. Often, engineered levees are built on or along naturally occurring levees, adding an additional level of complexity to the system. Additionally, levees occur in many environments where flooding is problematic including along rivers, in deltas, and along lake shores (USSD, 2016). Even the loading conditions of levees are variable. Most levees are only loaded during a storm event but levees in deltaic environments are loaded frequently and act as dams. Due to the nature of how they were built as well as the conditions they exist in, no two levees are the same. Levees come in all shapes and sizes, but the typical geometry of an earthen levee is depicted in Figure 1.1.

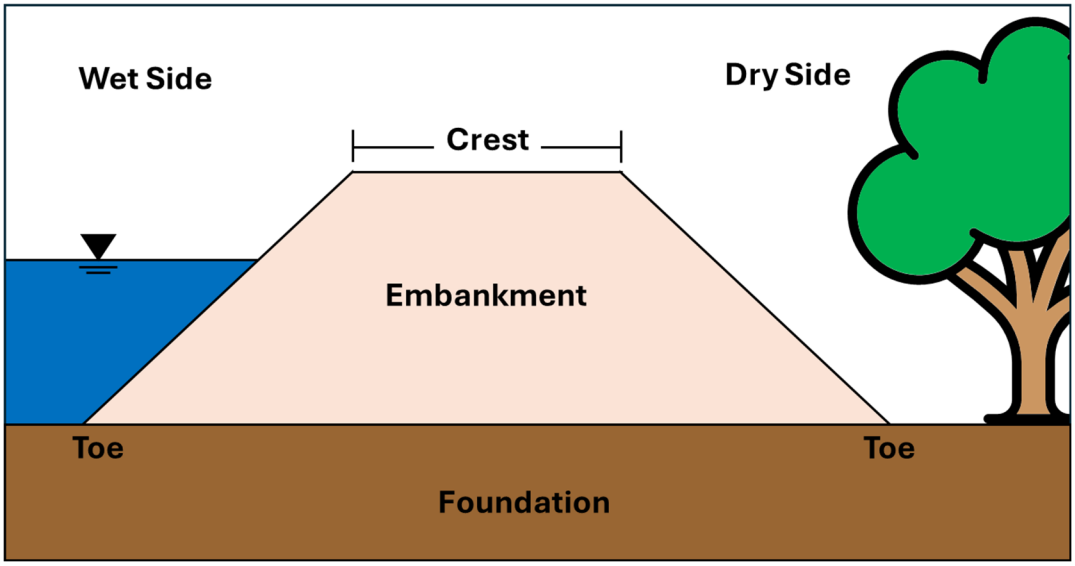


Figure 1.1 Typical Structure of a Levee

The primary structure of the levee is the earthen embankment, typically formed from compacted clay or the native soil in the region and can differ from the foundation material. The height of the structure is designed to withstand high flooding events where the embankment crest is the highest point on the levee. The width of the crest can be defined by a stability analysis or anthropogenic uses such as roads or pedestrian paths. The earthen embankment has two slopes, and the steepness is dependent on many factors including slope stability, maintenance requirements, and land use needs. The slopes and crest of levees may be covered by vegetation, pavement, or gravel, depending on the desired secondary function of the levee. The embankment also has two toes. The toes of a levee are where the bottom of the embankment slopes meet the foundation. Every levee has two sides, the wet side and the dry side. The wet side of a levee is the side with water during a high-water event while the dry side is the area being protected by the levee.

Because levees are complex earthen structures, they have many failure mechanisms that put people and structures at risk. These failures can be induced during high water events where an additional water load is acting on the system or from dry-weather failures that are unrelated to external hazards, such as storms and earthquakes. Typical levee failure mechanisms are listed in Table 1.1 and shown in Figure 1.2.

Table 1.1 Typical Levee Failure Mechanisms

Failure Mechanism	Description
Overtopping	Water exceeds the height of the levee crest and causes erosion on the crest or the dry toe
Through seepage	Water moves through sandy materials in the embankment and can internally erode soil
Under seepage	Water moves through sandy material below the embankment and can create sand boils
Surface erosion	Removal of material from wind or rain
Bank erosion	Removal of material from flowing water
Slope failure	Dry or wet slope collapses
Liquefaction	From strong shaking during earthquakes
Settlement	The weight of the structure causes displacement in the foundation material
Surface cracking	Shrinking of soil as it dries
Vegetation encroachment	Roots can cause preferred conduits for fluid movement
Animal burrowing	Cavities created in the embankment from animals
Anthropogenic hazards	Shipwrecks, car wrecks, explosions, etc.

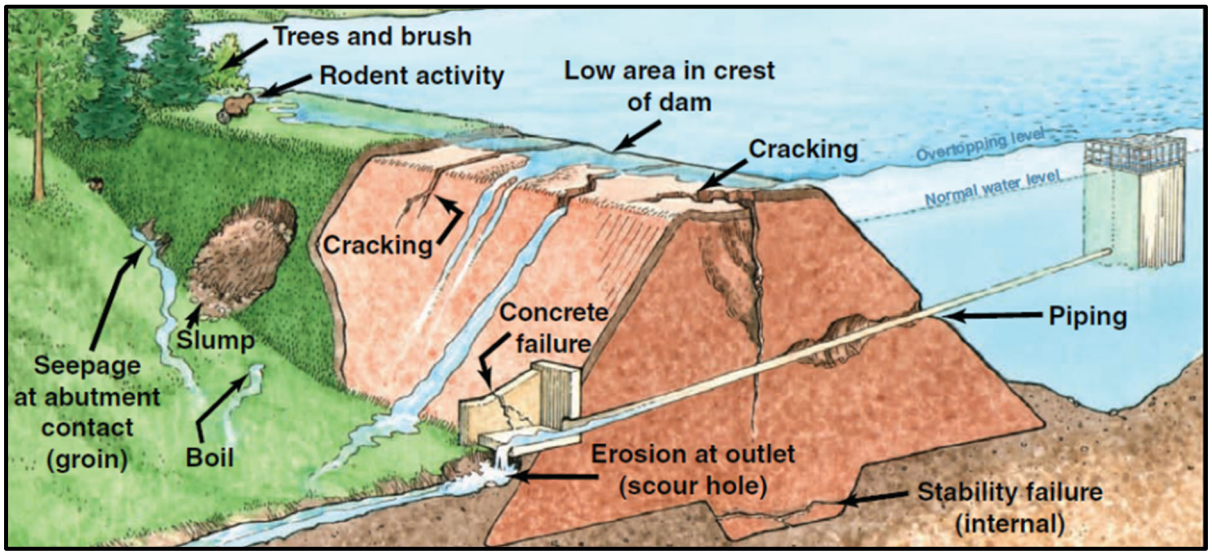


Figure 1.2 Typical Levee Failure Mechanisms (ASDSO, 2020)

With levees, some failure mechanisms are more likely to prove fatal to the system and cause major economic damage. Overtopping of levees is typically the most concerning during a high-water event as the force of the water erodes the crest and potentially the dry toe of the levee causing a breach in the structure, effectively flooding the land that was being protected.

Due to the aging infrastructure, poor inspection, and maintenance practices, as well as the multitude of failure mechanisms, the ASCE has repeatedly given levees a grade of D in their infrastructure report card since 2009 (ASCE, 2021). This grade indicates that these structures have significant deterioration and a strong risk for failure. With these degrading structures, better maintenance and inspection practices are needed to gain infrastructure resiliency against the increasing demand put on these systems due to climate change.

Current levee inspection practices in the United States are not sufficient to ensure the health of these systems. There are many problems regarding inspection including there being no standard among different organizations, inspection is primarily performed visually, and the temporal frequency is not enough to ensure the health of these systems.

Currently, in the United States, levees are owned and operated by the USACE, levees are owned by the USACE but maintained by another entity, levees are owned federally but not by the USACE, or levees are not federally owned (USSD, 2016). Approximately 70% of the levees owned by the USACE are operated by other organizations such as municipalities, levee districts, water management districts, and tribal nations (ASCE, 2021). These organizations are responsible for determining the frequency of inspection as well as inspection methods. Rarely is there collaboration between these entities in creating standards for operation resulting in inconsistencies in levee design, an increased risk of undetected failure mechanisms, and no coordination in emergency response and preparedness.

Most levees in the United States are inspected visually. These visual reconnaissance efforts involve a field technician on site looking for surface signs of failure including vegetation, cracking or scarps, depressions in the crest, sinkholes, or bulging at the toe. Surface signs of failure are important to document, but since levees are complex earthen structures, there are many internal

failure mechanisms that are not visible from the surface until it is too late. Some instrumentation has been used by organizations to monitor the internal health of levees but many of these methods only measure data at a discrete point in the subsurface (USSD, 2016). Examples of these instruments include piezometers to measure pore pressures, soil moisture sensors, settlement cells and extensometers to measure vertical displacement, inclinometers to measure lateral displacement, and thermistors or thermocouples to measure subsurface temperature (USSD, 2016).

Including these instruments in a levee structure while it is being built can be useful for collecting important subsurface data but installing these types of instruments into an existing levee has its challenges such as an increase in the risk of failure from hydraulic fracturing (USSD, 2016).

Another major problem with levee inspection is the temporal frequency at which it is performed. Temporal frequency for levee inspection is currently not enough to ensure the health of these systems. For levees owned by the USACE, there are two types of inspection for levees, with differing temporal frequencies. The first type of inspection is called the Routine Inspection and occurs annually and mainly consists of visual inspections. The second type of inspection is called the Periodic Inspection and occurs every five years. With this type of inspection, a more rigorous inspection is performed to understand the condition of the levee focusing on structural stability, safety, and operational adequacy (USACE, 2024). Having levee inspections annually to look for surface signs of failure or every five years for a more rigorous understanding of structural health leads to an increased risk of not detecting failure mechanisms, especially internally.

According to the National Levee Database, there are over 24 thousand miles of known levees in the United States that are putting 23 million people at risk. There are 7 million buildings and 5 million acres of farmland being protected by levees with $2 trillion of property value at risk (USACE & FEMA, 2024). The average age of levees in the United States is currently 60 years.

Given the challenges of having a variety of levee geometries, loading conditions, a multitude of failure mechanisms, no standard in inspection and maintenance, no subsurface health information, and a sparse temporal frequency of inspection, this research aims to create a robust levee health assessment framework. The goal of this research is to systematically analyze levees at the system level to provide valuable insights to stakeholders for informed decision-making. This will be achieved through integrating various sensing techniques for comprehensive imaging of both the surface and subsurface of levees. This framework aims to standardize inspection procedures and to increase inspection efficiency to allow for a greater temporal assessment of levee conditions.

Imaging the surface was performed using UAV-enabled optical imagery to classify land use and thermal imagery to understand temperature variations of the surface. UAV-mounted LIDAR was used to create detailed 3D models of the levee geometry and to create Digital Elevation Models. Leveraging change detection techniques on these datasets can show changes in vegetation, surface moisture content, slope instability, and toe bulging.

The subsurface will be imaged using two geophysical methods, the Multichannel Analysis of Surface Waves (MASW) and Electromagnetic Induction (EMI). MASW is a common surface wave seismic method used in geotechnical engineering to map variations in shear wave velocity (Vs). Vs is an important parameter as it is directly related to soil stiffness and provides information on changes in lithology and the location of structures. EMI uses electromagnetic fields which interact with conductive bodies in the subsurface to infer the apparent electrical resistivity (ρ). The ρ is useful for understanding levee conditions as it is a proxy for moisture content and clay content.

Change detection techniques can be leveraged on the subsurface data to understand how the subsurface is changing over time while machine learning analyses can be leveraged to identify locations in the subsurface that could lead to potential internal failure mechanisms.

The surface and subsurface imaging sensors used are non-invasive which is important for understanding levee conditions compared to invasive exploration methods and sensors such as drilling which can damage the existing structure. These sensors can also be scaled up to image spatially extensive systems like levees. Because of this, the temporal frequency of inspection can be increased, and it becomes more efficient to image large extents of area.

1.2 Research Objectives

The focus of this research was to develop a robust levee health assessment framework to aid stakeholders in understanding the conditions of levees at a system level for decision-making about maintenance and future rehabilitation. To achieve this, this research has five objectives:

1. Create a comprehensive library of Vs and ρ for a variety of geotechnical conditions in the laboratory to aid in the interpretation of subsurface conditions from unsupervised machine learning analyses.
2. Use a variety of non-invasive and spatially distributed sensing techniques such as UAV-mounted LIDAR, optical imagery, thermal imagery, MASW, and EMI to image the surface and subsurface levee conditions during the dry and wet seasons.
3. Explore the effectiveness of various feature engineering techniques and various supervised and unsupervised machine learning algorithms for analyzing levee subsurface data.
4. Investigate the correlation between change detection techniques between surface and subsurface levee data and different surface and internal failure mechanisms.
5. Create a semi-automated framework that fuses change detection surface analyses with subsurface change detection analyses and machine learning analyses to produce a model of the levee segment showing areas of increased vulnerability.

This robust levee health assessment framework produces a model that fuses surface and subsurface health information to understand locations of increased hazard along a levee. The importance of this framework is that it creates a standardization for inspection amongst different organizations and it increases the efficiency of inspection while allowing for more inspections in a year's timeframe. The outcome of this research is a hazard map that justifies locations where more detailed subsurface investigations might be needed, and it can aid stakeholders in decision-making regarding maintenance and future rehabilitation.

1.3 Limitations

Collecting data on levees provided some challenges. With over 24 thousand miles of known levees, they are under a variety of loading conditions and are composed of a variety of subsurface materials. This data collected for the levee framework was collected at one levee system that was continuously loaded, with known seepage problems, and composed predominantly of sandy soils. The dataset is not representative of all levees in the United States and is not representative of all possible failure mechanisms. Another limitation of the study was that ground truth data was not

available for the sites where the levee framework was developed. The sites were chosen as boreholes and CPTs were scheduled, but permits were not approved before this research was completed. This provided challenges in validating the subsurface change detection and machine learning results. The final limitation of this study was that the laboratory study only analyzed geotechnical conditions associated with non-cohesive soils. As a result, the laboratory labels discussed for machine learning interpretation should only be applied to subsurface conditions composed predominantly of sandy soils.

1.4 Dissertation Structure

This dissertation is composed of five chapters. At the beginning of each chapter will be a literature review regarding the topics discussed in that chapter. A summary for each chapter is as follows:

Chapter 2 describes the laboratory investigation to measure Vs and ρ under a variety of geotechnical conditions using bender elements and electrodes. The development of the testing cell is discussed along with the results from the analysis.

Chapter 3 discusses the collection of the field data including the theory behind the methods, the equipment used, and the location of the field sites. This chapter also presents the surface and subsurface change detection results between the dry and wet seasons.

Chapter 4 introduces machine learning methods, feature engineering techniques, and a variety of supervised and unsupervised machine learning algorithms. This section describes the supervised machine learning analysis for Bacon Island using the geophysical data discussed in Chapter 3. This chapter also describes the unsupervised machine learning framework for fusing clustering and anomaly detection techniques for the identification of potentially problematic zones in the subsurface. This analysis uses the remote sensing data and geophysical data discussed in Chapter 3. This section also discusses how the laboratory results discussed in Chapter 2 are fused with the unsupervised machine learning analysis.

Chapter 5 presents the outcome of the levee health assessment framework. This chapter describes how the surface and subsurface data were fused from the analyses discussed in Chapter 2 and Chapter 3 as well as describing the workflow to provide a semi-automatic model to produce a hazard map to understand the conditions along the site. Also presented is a comparison of hazard maps that were created using different machine learning parameters.

Chapter 6 summarizes the major conclusions and findings of this research. This chapter also outlines the recommendations for future research.

Chapter 2. Laboratory Testing

2.1 Laboratory Testing Introduction

Continuous and non-invasive field sensing techniques to measure shear wave velocity (*Vs*) and electrical resistivity (ρ) are incredibly effective at imaging large spans of the subsurface but interpretation can be ambiguous. Subsurface *Vs* and ρ can be affected by numerous geologic and environmental factors such as soil lithology, water content, and organics. To assist in the interpretation of field profiles, *Vs* and ρ were measured in a laboratory setting where a range of geologic and environmental factors were controlled. The results of the laboratory testing were applied to the unsupervised machine learning results in Section 4.8. Portions of the results presented in this chapter have previously been published in the following conference proceedings and are in preparation for journal publication:

Brittany M. Russo, Adda Athanasopoulos-Zekkos, Jongchan Kim, "Effect of Grain Size of Granular Soils on Shear Wave Velocity and Electrical Resistivity for Levee Health Monitoring". In the Proceedings of *Geo-Congress 2023*. (2023).

Brittany M. Russo, Adda Athanasopoulos-Zekkos, Jongchan Kim, "Influence of Geologic and Environmental Conditions of Soil on Electrical Resistivity." [In Preparation]

2.2 Literature Review

Laboratory measurements of *Vs* have been extensively researched as it is a popular way to understand soil stiffness since it is directly related to the small strain stiffness. Some of the earliest work of measuring *Vs* in the laboratory included the use of resonant columns which operate by initiating a sinusoidal torque that causes waves to propagate through a cylindrical soil specimen (Drnevich, 1978). One of the most influential studies using the resonant column to measure the *Vs* of soil was conducted by Hardin and Richart (1963). This study measured the *Vs* in controlled experiments on Ottawa sand under varying confining stress with respect to moisture content, void

ratio (e), and grain characteristics. The findings of this work showed that grain shape affected *Vs* while grain size and gradation had little effect. This research also showed that *Vs* decreased linearly with void ratio. The major contribution of this work was the relationship between *Vs* and the vertical effective stress (σ_o'):

$$V_s = (me)\sigma_o'^{0.25} \quad (2.1)$$

Where m is a function of the material.

The development of bender elements made measuring *Vs* in the laboratory easier. The study conducted by Shirley and Hampton (1978) developed shear wave transducers using bender elements to measure the *Vs* of soil, in situ. This study highlighted how easy it was to measure *Vs* in situ using bender elements, which caused a decrease in the use of resonant columns to measure *Vs* in the laboratory. This was further cemented when Bates (1989) created a custom triaxial cell with bender elements installed in the top and bottom cap to measure *Vs*. This study eliminated the need for resonant columns as the *Vs* of soil could now be measured while performing routine triaxial tests. A study conducted by Brignoli, Gotti, and Stoke (1996) compared bender elements, shear plate transducers, and resonant columns for measuring *Vs* in the laboratory. They found that bender elements were better for measuring *Vs* at a small strain range compared to resonant columns and shear plate transducers are better for measuring *Vs* of undisturbed soil samples and for measuring *Vs* of soils with large aggregate.

While bender elements became the popular choice for measuring *Vs* in the laboratory, there were few standards with respect to experimental procedure. Viggiani and Atkinson (1995) were some of the first pioneers who explored best practices for measuring *Vs* using bender elements. That study used bender elements on kaolin clay to determine the appropriate travel time length through the soil and determination of wave travel time. Through their work, the travel time length was standardized to be the tip-to-tip distance between the two bender elements. They also suggested using the cross-correlation or cross-power spectrum for travel time arrival. In comparison, the results of the study conducted by Jovicic, Coop, and Simic (1996) showed that the travel time measurement of the received wave can be directly measured from the oscilloscope trace instead of using the Fourier Transform with the caveat that soils should be tested several times to increase the confidence in the data. Traditionally, square waves were used as the input waves for bender element testing. The study performed by Leong, Yeo, and Rahardjo (2005) challenged that by comparing the *Vs* results of soil samples when using square waves and sinusoidal waves. Their study showed that the highest quality receiver signal was produced from the sinusoidal waves instead of the traditional square waves.

As standards were being created for the measurement of *Vs* using bender elements, this methodology was starting to be used extensively to measure the physical properties of soils. Research conducted by Salgado, Bandini, and Karim (2000) was interested in how *Vs* changed with fines content, varying relative densities, and effective vertical stress while Lee, Suh, Yoon, and Yun (2017) were interested in how *Vs* was affected by void ratio, effective vertical stress, and the shape of soil grains. The study performed by De Alba et al. (1984) expanded on Hardin and Richart's work by looking at how *Vs* varied with different grain sizes and shape characteristics for

sands. Their work summarized how *Vs* was a function of the microstructure of the sand and was heavily dependent on how grains were oriented and distributed.

The Roberston (1995) study used bender elements to measure the *Vs* of Ottawa sands under varying densities and effective confining stresses under undrained and drained triaxial compression loading. The results of this study were limited to clean and uncemented sands but showed that it was possible to use *Vs*, vertical effective stress, and the lateral earth pressure coefficient to estimate the contractive and dilative behavior of soils.

Research conducted by Park and Lee (2014) was interested in the physical characteristics of sand-silt mixtures during freezing and thawing conditions. This study used bender elements to measure the *Vs* of the soil while temperatures ranged between 10°C and 20°C. The soil was prepared to saturations that ranged from 5% to 100% and found that while the degree of freezing of the soil increased, the *Vs* also increased, but once the soil was frozen, the velocities remained constant.

More recently, research using *Vs* in the laboratory for liquefaction applications has become more popular. Yunmin, Han, and Ren-peng (2005) measured *Vs* using bender elements to develop a simplified procedure for evaluating liquefaction resistance while Cabalar, Demir, and Khalaf (2021) measured *Vs* in the lab to determine G_{max} and the Cyclic Resistance Ratio (CRR) for soil samples of different particle sizes, particle shapes, and fine contents to understand their effect on liquefaction resistance. Work has been done to incorporate bender elements to measure *Vs* during centrifuge testing. Fu, Liu, and Zeng (2009) developed a technique to measure the *Vs* of soils during spin-up, spin-down, and in-flight stages in centrifuge testing for dry and saturated samples while work performed by El-Sekelly, Tessari, and Abdoun (2014) summarized the caution required to use bender elements in noisy environments, like centrifuge testing.

While work measuring *Vs* in the laboratory has been extensively studied, measuring ρ in the laboratory is not as prevalent. Compared to the standards to measure *Vs* in the laboratory, there are no standards to measure ρ. The closest standardization is derived from work done by Cho, Lee, and Santamarina (2004) measured the electromagnetic properties of soils using a custom-made needle-size probe and explored important parameters that affect the ρ readings, including the probe shape, the probe material, voltage effects, and the operating frequency of the current. This study concluded that many factors should be considered when designing a system to measure ρ.

A study performed by Pandey, Shukla, and Habibi (2015) explored some of these factors in more detail. These authors measured the ρ of a natural sandy soil from Perth, Australia while controlling for water content, relative density, the type of water used in the soil, as well as the input voltage type, frequency, and electrode material. The authors found that the frequency, voltage, and electrode material were inconsequential to testing. Instead, the authors found that the ρ measurements were significantly affected by the soil properties. Regardless of the type of water used, as the water content increased, the ρ decreased rapidly, but the rate of decrease was dependent on the water type. Relative density was also found to have a linear relationship with ρ but only at low water contents.

Measurements of ρ in the lab are primarily used to understand varying soil properties. The most common soil property measured is water content. A study performed by Cardoso and Dias (2017)

was interested in measuring the ρ of kaolin clay at different void ratios and different water potentials. Their study found that at low degrees of saturation, there is a strong rate of increase in ρ and ρ can be used to estimate total suction. Bhatt and Jain (2014) measured ρ on sands under varying water contents. Their study showed that at low water contents, the soil was not conductive enough to measure ρ. They also found that a non-linear regression model provided the best correlation between the decreasing ρ and increasing water content. A study performed by Kibria and Hossain (2012) looked at how the ρ of highly plastic clays was affected by a variety of factors including water content. The authors showed that as the surface area of the clay particles increased, the ρ also increased and was the most important factor influencing soil ρ.

Other factors that researchers have studied include the density and fines content of soils. Research conducted by Seladji et al. (2010) wanted to understand the physical relationship between ρ, soil porosity, and water content for agricultural and forest soils by altering the density of the soil. The researchers found that regardless of soil texture, the ρ was sensitive to soil density, and they recommended that using ρ for soil compaction should be conducted in the dry density state. Other work by Mostafa, Anwar, and Radwan (2018) was interested in how soil compaction and fines content affected the ρ of calcareous soils. The fines content had an inverse relationship with ρ. For compaction, the side of optimum moisture content was deemed to be important. On the wet side of optimum, results in ρ decreased with increasing compaction while on the dry side of optimum, results in ρ increased with increasing compaction.

The effect of weathering on ρ has also been studied in the laboratory by Son, Oh, and Lee (2010). The authors wanted to correlate ρ with chemical weathering. They found there was no clear relationship with weathering, but at low water contents, there was a linear relationship, showing that water content dominated ρ readings and not the degree of weathering. Tang et al. (2018) showed that ρ can be used to characterize how clays dry out due to evaporation, volumetric shrinkage, and the formation of cracks.

While most of the laboratory measurements of the ρ of soils are of a single measurement, work by An et al. (2020) developed a method with high spatial resolution to measure the ρ throughout the entire sample. This study aimed to understand how the water content of the soil changed with depth due to evaporation. The authors were able to successfully measure the temporal and spatial variations of soil water content during evaporation and showed that water content decreased exponentially with increasing ρ.

There has been little work in correlating laboratory ρ measurements with field ρ measurements recorded by geophysical sensors. Work done by Bloss and Bedrosian (2015) explored creating these correlations for cored boreholes and surface outcrops at Fort Irwin National Training Center in Southern California. They experimented with both soil and rock samples of a variety of lithological units including plutonic rocks, metamorphic rocks, lava flows, consolidated sedimentary rocks, and unconsolidated sedimentary deposits. The field measurements were collected using a time-domain electromagnetic geophysical sensor, which was used near most of the boreholes. The study showed that the laboratory ρ measurements of the rock and soil samples had an excellent correlation with the time-domain electromagnetic geophysical data, proving that

ρ measured in the laboratory using direct current was a reasonable estimate of the ρ measured in the field using geophysical sensors.

Measuring *Vs* and ρ in tandem has not been extensively researched. There has been some work conducted by Yun et al. (2006) who were interested in measuring the physical properties of sediments that were recovered from pressure coring devices. The authors designed a chamber that could measure the *Vs* of the soil, the undrained shear strength, as well as the electrical conductivity.

Other work by Kang and Lee (2015) looked at how *Vs* and ρ were affected by silt content and temperature. The soil sample was a sandy silt mixture, and the authors systematically changed the silt content with different experiments. For each experiment, the sample experienced temperatures that ranged from -13.5°C to 20°C. The *Vs* was measured by bender elements that were applied to the sides of the soil sample while the ρ was measured using four stainless steel electrodes installed at the corners of the soil cell. The results of this study showed how *Vs* and ρ were affected by temperature under the same conditions. After one cycle of freezing, the *Vs* decreased while the ρ increased due to the change in soil fabric. This study also showed how the smallest void ratio of the soil produced the lowest *Vs* and ρ.

While research has been conducted to understand how certain physical and chemical properties of soils affect *Vs* and ρ independently, there is a gap in understanding how *Vs* and ρ are affected in tandem. This chapter describes a systematic approach to measuring the *Vs* and ρ of soils under the same conditions by controlling for geotechnical conditions in the laboratory. These conditions include particle grain size, particle angularity, soil gradation, organic content, water content, relative density, and vertical stress.

2.3 Experimental Testing Design

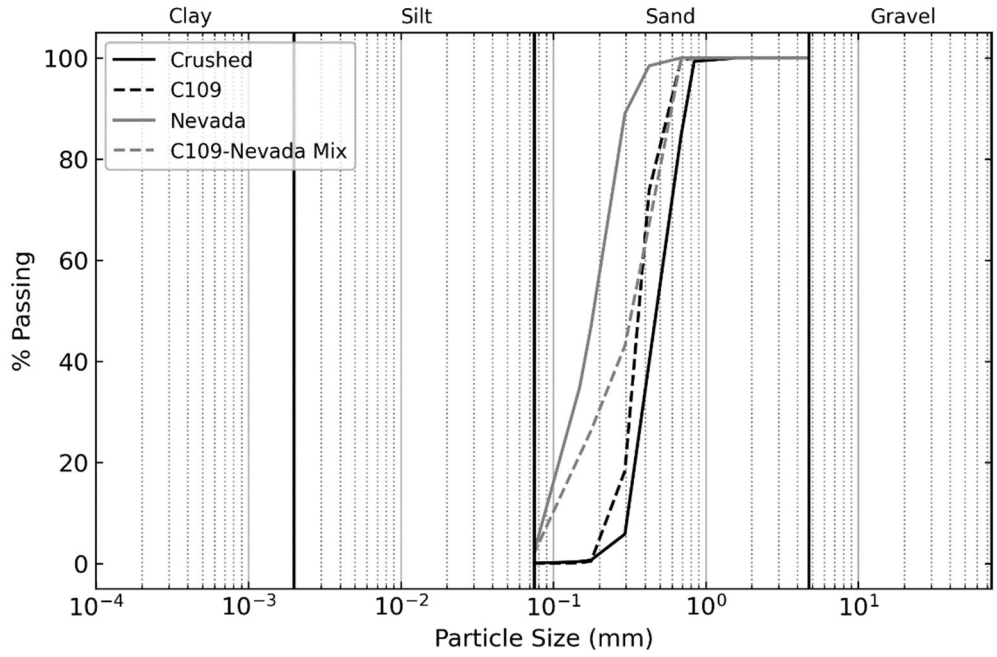
2.3.1 Materials

The goal of the laboratory testing was to determine how certain geologic and environmental conditions affect the *Vs* and ρ of soil. To accomplish this, various geotechnical conditions were controlled for, including particle grain size, particle angularity, soil gradation, the addition of organic material in the soil, water content, and relative density.

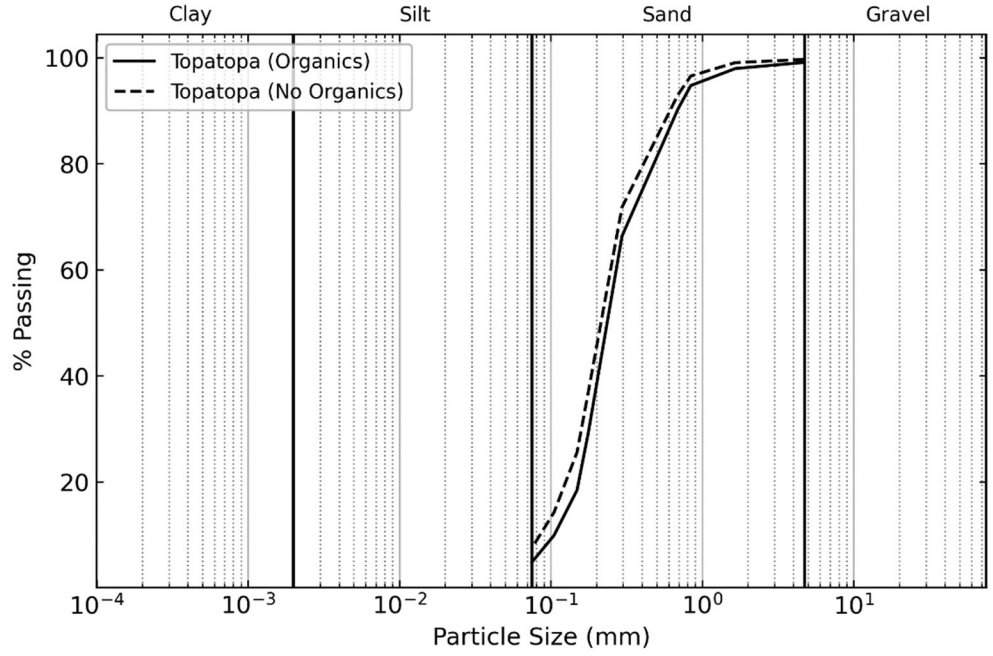
Nine non-cohesive granular soils were used to explore how the various geologic and environmental conditions affect *Vs* and ρ. Table 2.1 lists the soil names, USCS classification, Cc, Cu, minimum dry unit weight ($(\gamma_d)_{min}$), and maximum dry unit weight ($(\gamma_d)_{max}$). Figure 2.1 displays the gradation curves of each soil and Figure 2.2 shows images of a sample of each soil.

A naturally occurring soil sample from the Topatopa Mountain Range in Southern California was tested to look at the effect of organics on ρ and *Vs*. A representative sample of soil was taken from the slope of a mountain, taking care to preserve the naturally occurring organic material. This soil was tested, as is, with the organics. Afterwards, the large pieces of organics were sieved out and then the remaining organic material was burned off per ASTM 2974. Once all the organics were burned off, the soil was retested to look at the effects of organics. The soil had approximately 1% organics by weight.

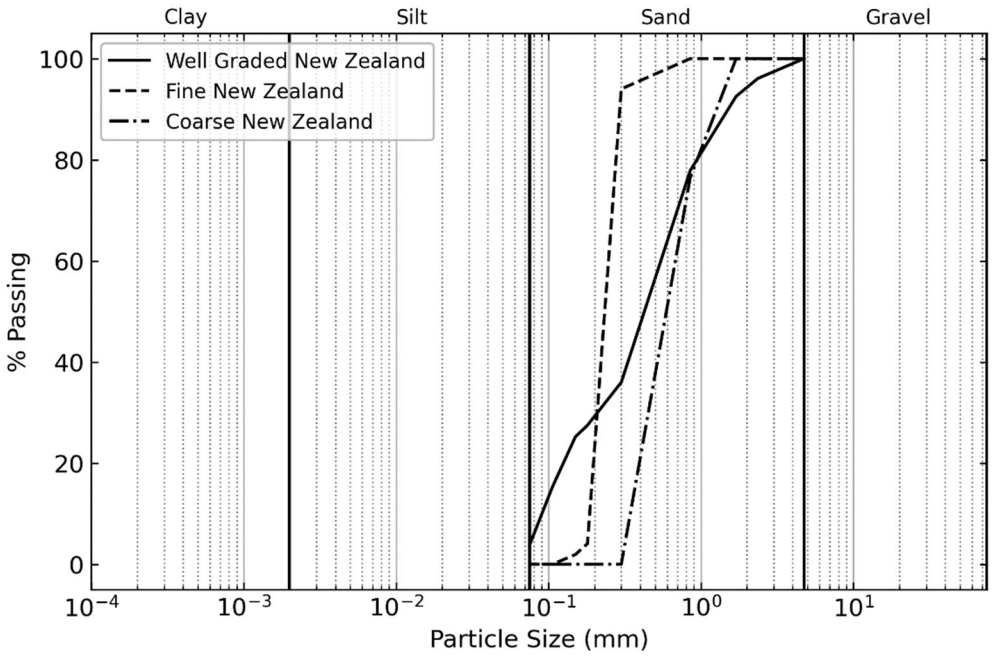
To study how the angularity of granular soils affects ρ and Vs, a silica-rich rock was crushed, resulting in angular particles with similar gradation to C109 sand, a rounded commercial silica sand. Both soils were classified as poorly graded with less than 12% fines content.



(a)



(b)



(c)

Figure 2.1 Gradation of curves of soils tested in the laboratory (a) quartz grained soils, (b) Topatopa soils, and (c) New Zealand soils



(a) (b)



(c) (d)



(e) (f)



(g) (h)



(i)

Figure 2.2 Materials tested (a) Crushed angular soil, (b) C109 Sand, (c) Nevada Sand, (d) C109-Nevada Mixture, (e) Topatopa Sand with organics, (f) Topatopa Sand with no-organics, (g) New Zealand - Fine, (h) New Zealand - Coarse, and (i) New Zealand Well Graded

Table 2.1 Soil Properties and Classification

Name	USCS Classification	Cc	Cu	γ_{max}	γ_{min}
Angular	SP	2.5	0.8	15.6	13.2
C109	SP	1.6	0.9	17.3	14.8
Nevada	SP	1.1	2	17.2	14.4
C109-Nevada Mixture	SP	1.0	3.9	18.1	15.1
Topatopa – Organic	SP	1.2	2.5	15.7	12.5
Topatopa – Non-organic	SP-SM	1.4	2.9	16.5	13.0
New Zealand Fine	SP	0.9	1.3	15.8	11.2
New Zealand Coarse	SP	0.9	2.1	16.7	13.0
New Zealand Well Graded	SW	1.0	6.1	17.5	13.3

C109 Sand and Nevada Sand were compared to study the effect of grain size on ρ and *Vs*. Both sands are poorly graded, silica sands, with rounded grain particles. The gradation curves between the two soils are similar where C109 Sand has a Cu of 1.60 with a D_{50} of 0.36 mm while Nevada Sand has a Cu of 2 and a D_{50} of 0.18 mm.

To study the effect of soil gradation on ρ and *Vs*, C109 Sand and Nevada Sand were mixed to create a silica sand with rounded particles that ranged in particle sizes from the largest grains of C109 Sand to the finest grain size in Nevada Sand. Multiple iterations of this sand were done to create the largest gradation possible with the grain particles available. To compare gradation, Nevada Sand and C109 were both used, separately. Even though the C109-Nevada Sand mixture was classified as poorly graded, it had a Cu of 3.94 which was much larger compared to C109 Sand (Cu of 0.9) and Nevada Sand (Cu of 2.0).

To confirm the effects of gradation and particle size, a natural soil from New Zealand was engineered to have certain gradations. The New Zealand soil was first separated into different grain sizes through sieving. The different weights of each soil size were appropriately combined to achieve the desired gradation of a Cu of 6.0 and a Cc of 1.0. This was tested against a fine grained poorly graded soil of the same material with a Cu of 1.3 and a Cc of 0.9 as well as a coarse grained poorly graded soil of the same material with a Cu of 2.1 and a Cc of 0.9.

The poorly graded fine and coarse New Zealand soils were also compared to confirm the effect of soil particle size on *Vs* and ρ.

2.3.2 Sample Preparation

The first step in the soil preparation was to oven-dry the material for a minimum of 24 hours at 105 ± 5°C to ensure there was no moisture in the soil sample. Each soil was tested at a variety of water contents including 1%, 5%, 10%, and 20%. For each water content sample, the soil was tested at three relative densities, 30%, 50%, and 70%.

To build the sample for testing, the weight of the moist sample was first calculated using Equation 2.2 using the target water content:

$$w = \frac{Wm - Wd}{Wm} \times 100 \tag{2.2}$$

where w is the water content in percent, w_m is the weight of the moist soil (g), and w_d is the weight of the dry soil (g). Once the weight of the moist soil was calculated, distilled water was added to the soil until the target weight was achieved. The soil and water were thoroughly mixed until the water was evenly distributed among the soil particles. Deionized water was used to reduce the impacts of variations in water chemistry. The ρ is highly sensitive to changes in water chemistry and differences in tap water could affect the results.

Once the target water content was achieved, the target relative density (Dr) was used to calculate the mass needed for the sample given the volume of the soil cell using Equation 2.3:

$$Dr = \frac{\frac{1}{(\gamma_d)_{min}} - \frac{1}{\gamma_d}}{\frac{1}{(\gamma_d)_{min}} - \frac{1}{(\gamma_d)_{max}}} \tag{2.3}$$

Where γ_d is the dry unit weight of the soil (kN/m^3). The tests to determine the maximum and minimum unit weights were conducted following the Japanese Standard JIS A 1224:2009. This standard was preferred because the testing required only a small sample of the soil (Mijic, Bray, Riemer, Cubrinovski, & Rees, 2021).

The γ_d can be converted into the moist unit weight (γ_{moist}) through Equation 2.4.

$$\gamma_{moist} = \gamma_d(1 + w) \tag{2.4}$$

With a calculated unit weight, the moist density (d_{moist}) of the soil was calculated which can be used to calculate the mass needed for testing, as shown in Equations 2.5 and 2.6, respectively.

$$d_{moist} = \frac{\gamma_{moist}}{g} \tag{2.5}$$

$$M = d_{moist}V \tag{2.6}$$

where g is the gravitational acceleration constant at 9.81 m/s^2, M is the mass of the soil (g), and V is the volume of the soil cell. For each test, the volume of the soil cell remained constant at 174 cm^3.

The soil was then divided into three equal lifts. Each lift was placed in the soil cell and compacted to the desired relative density using the tamping technique. This was repeated for the remaining two lifts.

After testing, the water content of the sample was measured following ASTM standard D2216-19.

For each test, the ρ and *Vs* of the soil were measured at seven vertical stresses during loading: 0 kPa, 5 kPa, 10 kPa, 25 kPa, 50 kPa, 75 kPa, and 100 kPa. The cell was placed on a triaxial load cell and weights were placed to reach the desired vertical stress. To ensure no electrical crosstalk during measuring, the connections to the LCR meter were disconnected when not measuring ρ and the connections to the waveform generator and oscilloscope were disconnected when not measuring *Vs*.

2.4 Soil Cell Design

A testing cell was designed to hold soil samples of granular materials. The cell was designed so that *Vs* and ρ could be measured under the same conditions. The body of the cell was 0.6 cm thick acrylic tube cut to 12.7 cm length with a 7.6 cm inside diameter. The top and bottom caps were a 7.6 cm diameter acrylic rod cut to 5.1 cm and 2.5 cm length, respectively. The bottom cap had an O-ring groove cut along the outside of the rod to fit a 3/32 fractional width O-ring to prevent leakage during testing. A schematic of the cell design is shown in Figure 2.3.

Installed in the top and bottom caps were two bender elements to measure *Vs* of the soil and four stainless steel rods which act as electrodes to measure the ρ of the soil. The bender elements were installed in the middle of the top and bottom cap to reduce noise and are discussed more in Section 2.4.1. Two electrodes were installed in the top cap and two in the bottom cap with a 1.5 cm offset from the bender elements. The spacing between the two electrodes was 0.9 cm. More detailed information about the electrodes is discussed in Section 2.4.2.

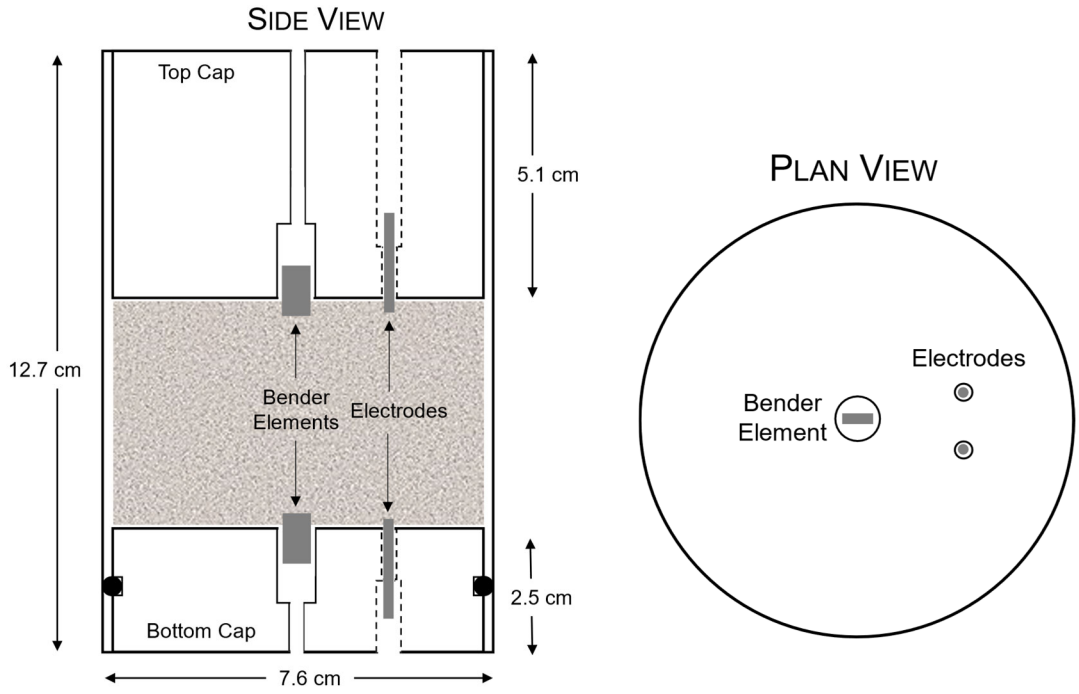


Figure 2.3 Soil testing cell design schematic showing the (a) side view and the (b) plan view

2.4.1 Bender Elements

Bender elements are used to measure the *Vs* of soils in the laboratory setting. They consist of two piezoelectric transducers. A single bender element has two outer piezoelectric ceramic plates with a center metal plate (Montoya, et al., 2012). The bender element acts as a cantilever beam where one end is anchored and the other end is embedded within the soil (Lee & Santamarina, 2005). A voltage is applied to one bender element which excites the sensor and causes it to oscillate. The bending of the piezoelectric material generates a wave that travels through the soil specimen. On the other side of the soil specimen is a second bender element. When the wave from the first bender element reaches the second bender element, it causes the second bender element to oscillate. This bending motion in the second bender element generates a voltage that can be measured as a waveform using an oscilloscope (Viggiani & Atkinson, 1995).

The bender elements used in testing were made from T220-A4BR-2513YE piezoelectric material from Piezo.com. The piezoelectric material was cut to 8 mm by 4 mm by 0.5 mm. Coaxial cables were soldered to the piezoelectrical material in a parallel circuit. In a parallel connection, the inner wire of the coaxial cable is attached to the metal shim while the outer wire of the coaxial cable was connected to either side of the piezoelectric ceramic for grounding. A parallel-type wiring was used because both piezoelectric ceramic pieces of the bender element are polarized in the same direction which allowed for double the displacement compared to a series-type bender element (Lee & Santamarina, 2005). After the wiring was soldered to the bender elements, layers of epoxy and M-coat A polyurethane were applied for waterproofing. The waterproofing of the bender

elements was necessary to protect the sensor from pore fluids during testing, as well as to protect the sensors from any abrasion due to the surrounding soil (Montoya, et al., 2012). The epoxy was checked before each test to ensure no damage has occurred and recoating is applied if necessary. To improve the grounding of the bender element, a single layer of conductive silver paint was applied. Figure 2.4 shows different stages of the bender element fabrication process. Each sensor was checked for short-circuiting throughout the fabrication process using a multimeter. After the bender elements were completed, each sensor was attached to a function generator at 20 Hz to check that the sensor was working properly. They were installed into the top and bottom cap using epoxy, extending 3 mm into the soil.

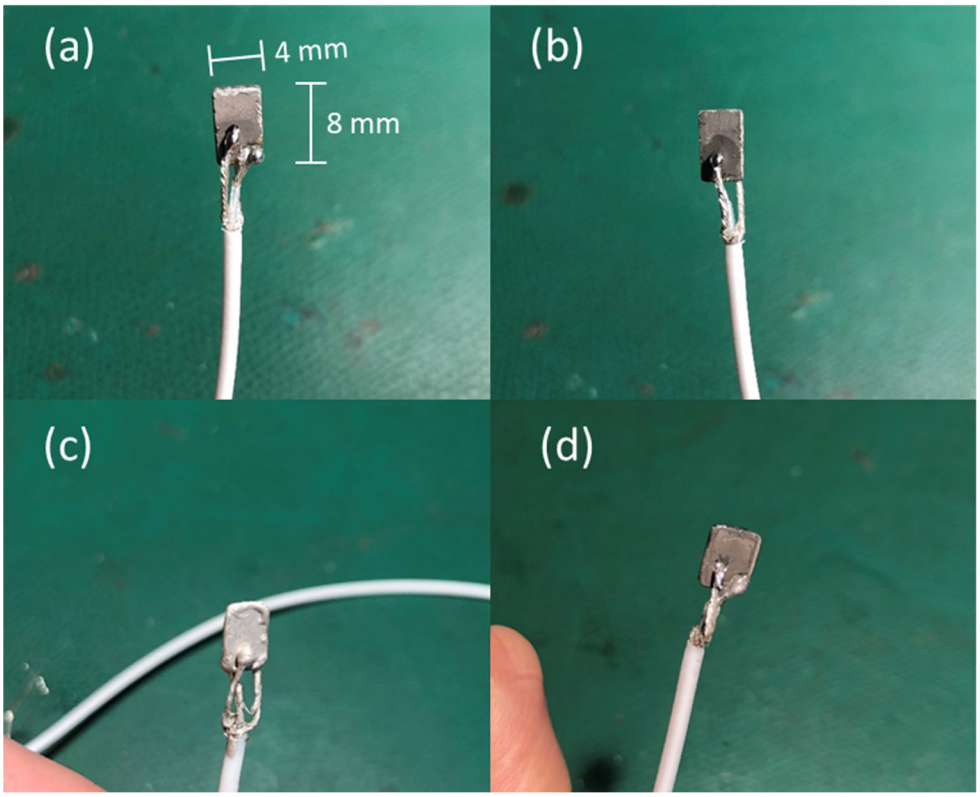


Figure 2.4 Fabrication of bender elements (a) soldering wire to the front of the bender element, (b) soldering wire to the back of the bender element, (c) silver paint coating, and (d) epoxy and polyurethane coating

The input wave for testing was generated by an Agilent 33500B Series waveform generator which was connected to the bender element in the bottom cap of the soil cell. A voltage of 10 V peak-to-peak was applied using a sine wave. The frequency of the sine wave for each test was dependent upon the resonant frequency of the soil. The input and received waves were recorded using a Tektronix MSO 2012 Mixed Signal Oscilloscope. The received signal was averaged with 512 points to reduce noise in the data. The data was then saved to a computer for further processing. Figure 2.5. shows the laboratory setup with the soil cell, oscilloscope, and waveform generator.

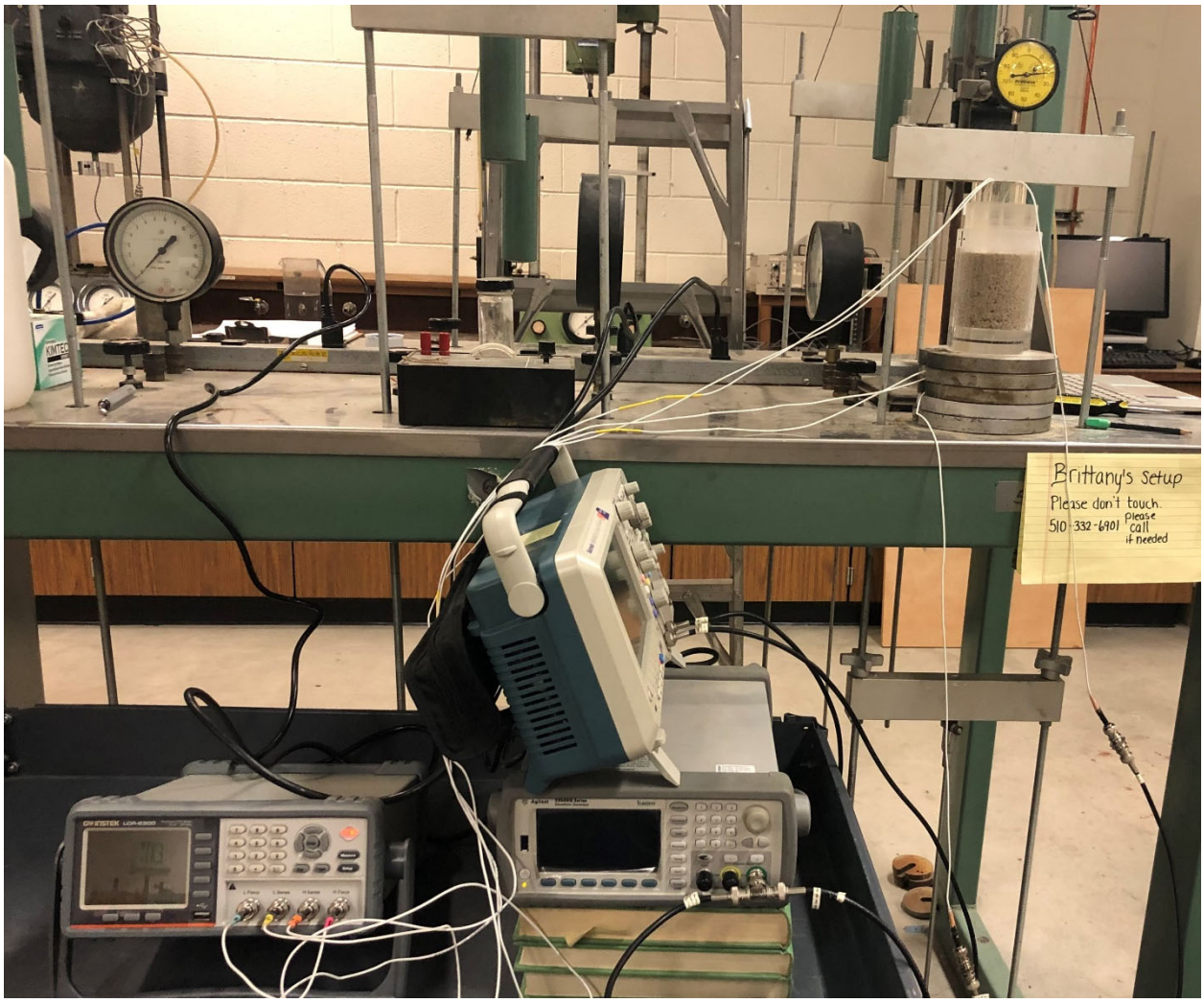


Figure 2.5 Laboratory setup showing the soil testing cell, dial gauge, oscilloscope, waveform generator, and LCR meter

The frequency used for testing was the resonant frequency of the bender element in each soil. Each soil was tested for the resonant frequency before analysis as the soil properties can affect the resonant frequency (Lee & Santamarina, 2005). To be consistent through testing, the resonant frequency was found using the soil with a water content of 1%, a relative density of 30%, and no additional load on the top cap (total stress acting on the soil was 0.65 kPa). Each test was performed using the resonant frequency as it results in the largest amount of energy transmitted leading to the largest signal received (Montoya, et al., 2012). The resonant frequency was found using a sine sweep while changing the frequency between 1 kHz and 25 kHz. Then, the peak magnitude of the received wave from each frequency was found and the frequency with the largest magnitude was determined to be the resonant frequency. The waveforms used in the resonant frequency testing for an SP soil with e_{max} of 0.96 and e_{min} of 0.64 is shown in Figure 2.6. This soil had a resonant frequency of 4 kHz.

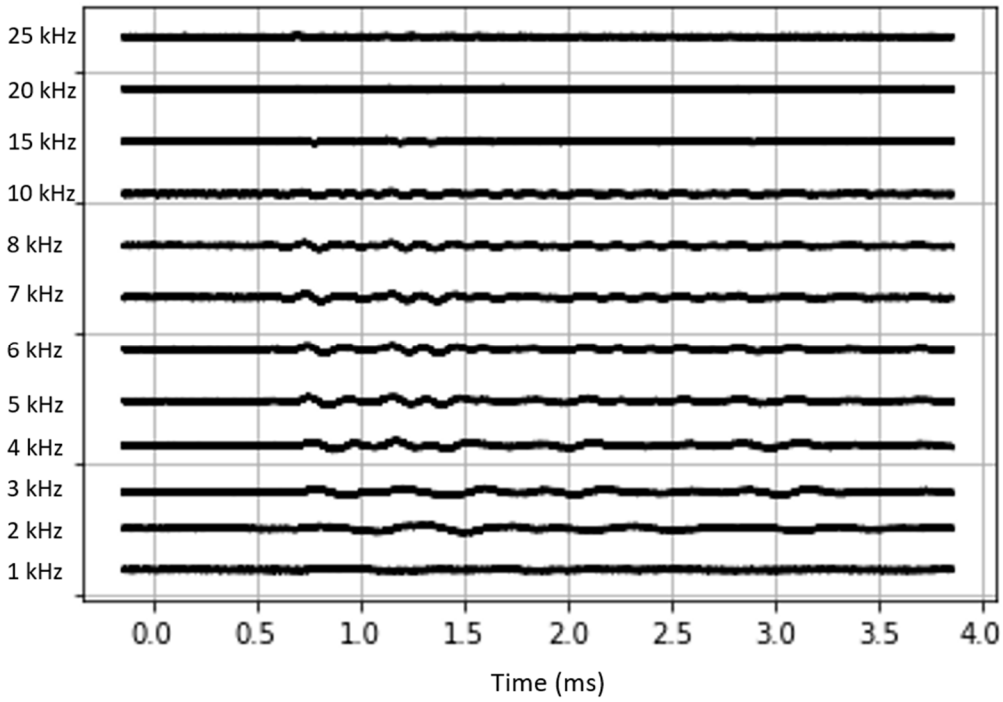


Figure 2.6 Resonant frequency plot ranging from 1 kHz to 25 kHz

The literature outlines many ways to measure the *Vs* of the soil in a laboratory setting. Each method results in the difference in travel time between the input wave and the received wave. Knowing this change in travel time, the *Vs* of the soil can be calculated using Equation 2.7 (Viggiani & Atkinson, 1995).

$$V_s = \frac{L}{\Delta t} \tag{2.7}$$

Where L is the tip-to-tip distance between the two bender elements in meters (m) and Δt is the travel time difference between the input and received waves in seconds (s).

The most common way to measure Δt is by picking the first arrival of the received wave in the time domain as shown in Figure 2.7. This method assumes that the received wave will retain the shape of the input wave through the soil (Yamashita, et al., 2009). By knowing the time of the arrival of the input wave and the first arrival of the received wave, the difference in the arrivals is the amount of time the shear wave took to travel through the soil. This method can also be done by picking the first peak-to-peak of the input and received wave or the first trough-to-trough. For testing, the first arrivals were used and recorded.

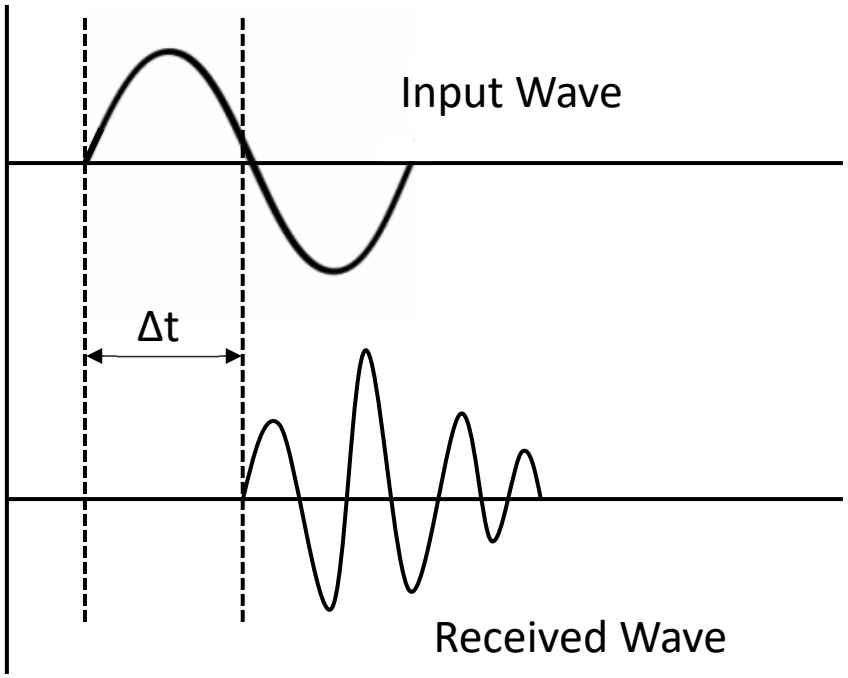


Figure 2.7 First arrival of received wave compared to input wave

A second method to determine the arrival is through cross-correlation between the input and received wave. The cross-correlation is used when comparing two signals to determine the similarity between them. When the cross-correlation is high in value, the two waves are very similar. The cross-correlation can be calculated using Equation 2.8:

$$\varphi_{fg}(t) = \int_{-\infty}^{\infty} f(t)g(t+\tau)d\tau \tag{2.8}$$

Where $\varphi_{fg}(t)$ is the cross-correlation in the time domain, f(t) is one signal in the time domain, g(t) is the second signal in the time domain, and τ is the time shift between the two signals. Through the cross-correlation theorem, the process of calculating the cross-correlation is simplified in the frequency domain. This can be done by first finding the Fast Fourier Transform (FFT) of each signal as shown in Equation 2.9:

$$F(f) = FFT\big(f(t)\big) \text{ and } G(f) = FFT\big(g(t)\big) \tag{2.9}$$

Where F(f) is the first signal in the frequency domain and G(f) is the second signal in the frequency domain. Then, the cross-power spectrum can be calculated using Equation 2.10.

$$\emptyset_{FG} = F^{*}(f)G(f) \tag{2.10}$$

Where the asterisk (*) denotes the complex conjugate. The cross-correlation of the signals can be found by bringing the cross-power spectrum back into the time domain through an Inverse FFT as shown in Equation 2.11.

$$\varphi_{fg}(t) = IFFT(\emptyset_{FG}) \tag{2.11}$$

To determine Δt from the cross-correlation, the time that corresponds to the maximum amplitude of the signal is considered to be the difference in travel time between the input and received waves (Yamashita, et al., 2009; Viggiani & Atkinson, 1995; Lee & Santamarina, 2005; Ogino, Kawaguchi, Yamashita, & Kawajiri, 2015; Airey & Mohsin, 2013; Arulnathan, Boulanger, & Riemer, 1998). Figure 2.8 shows the input wave, received wave, and the normalized cross-correlation.

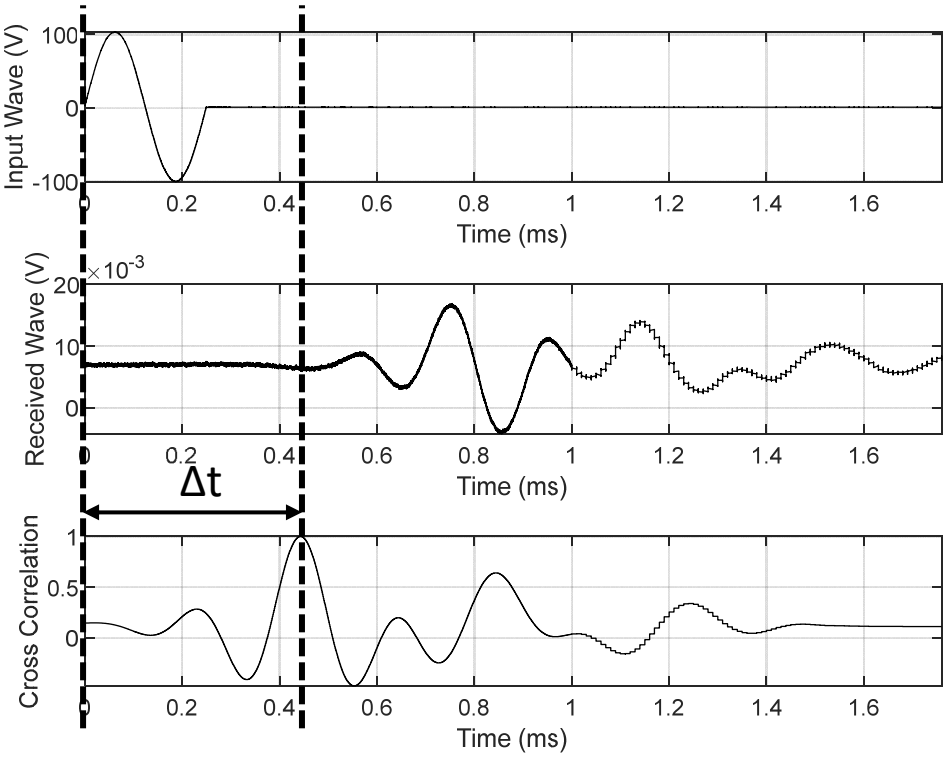


Figure 2.8 Cross-correlation results showing how maximum peak corresponds to first arrival of the received wave compared to the input wave

It was found that the cross-correlation method was unreliable with sandy soils, which confirms the statement made by (Ogino, Kawaguchi, Yamashita, & Kawajiri, 2015). The peak magnitude of the cross-correlation rarely matched the travel time difference between the first arrival of the input and received wave.

Since cross-correlation was found to be unreliable for sandy soils, the first arrival was manually chosen in the time domain. To achieve the most accurate *Vs*, the tip-to-tip distance between the bender elements was calculated at each measurement. A dial gauge with millimeter accuracy was used to measure the vertical difference due to an applied load. This vertical difference was subtracted from the initial tip-to-tip distance to calculate the actual distance between the bender elements.

2.4.2 Electrodes

The electrode probes used in testing were 304 stainless-steel rods with a 2 mm diameter. The rod was cut into 15.9 mm long pieces and sanded on the ends to ensure a smooth surface. The stainless-steel rods were attached to a GW Instek 6300 LCR meter, shown in Figure 2.5, to measure the resistance of the soil. The ρ of the soil can be measured using Equation 2.12:

$$\rho = kR \tag{2.12}$$

where k is the geometric factor with units of meters (m) and R is the resistance in Ω measured by the LCR meter. The geometric factor is determined by using multiple NaCl solutions with known ρ and measuring the resistances (Yun, Narsilio, Santamarina, & Ruppel, 2006; Cho, Lee, & Santamarina, 2004; Won, Park, Choo, & Burns, 2019; Merritt, et al., 2016).

Three different calibration solutions were used to determine the geometric factor, 84 µS/cm (119 Ωm), 1,413 µS/m (7 Ωm), and 12,880 µS/m (0.8 Ωm). For each calibration solution, a frequency sweep test was performed with the top cap at different heights above the bottom cap. The height of the top cap was changed to determine if that affects the geometric factor since testing will require the addition of loads to the top cap that could change the distance between the electrodes. The frequency sweep test was used to determine when electrical resonance starts to pick the optimal frequency for testing (Cho, Lee, & Santamarina, 2004; Kang & Lee, 2015; Son, Oh, & Lee, 2010).

It is important to know at what frequency electrical resonance begins as it makes the ρ readings unstable (Cho, Lee, & Santamarina, 2004). The frequency at which electrical resonance begins is a function of the geometry of the electrode as well as the cables that connect the electrode to the measurement system (Cho, Lee, & Santamarina, 2004). To determine when electrical resonance begins for the system designed for testing, 15 frequencies were used ranging from 10 Hz to 300,000 Hz. The results of the frequency sweep are plotted in Figure 2.9.

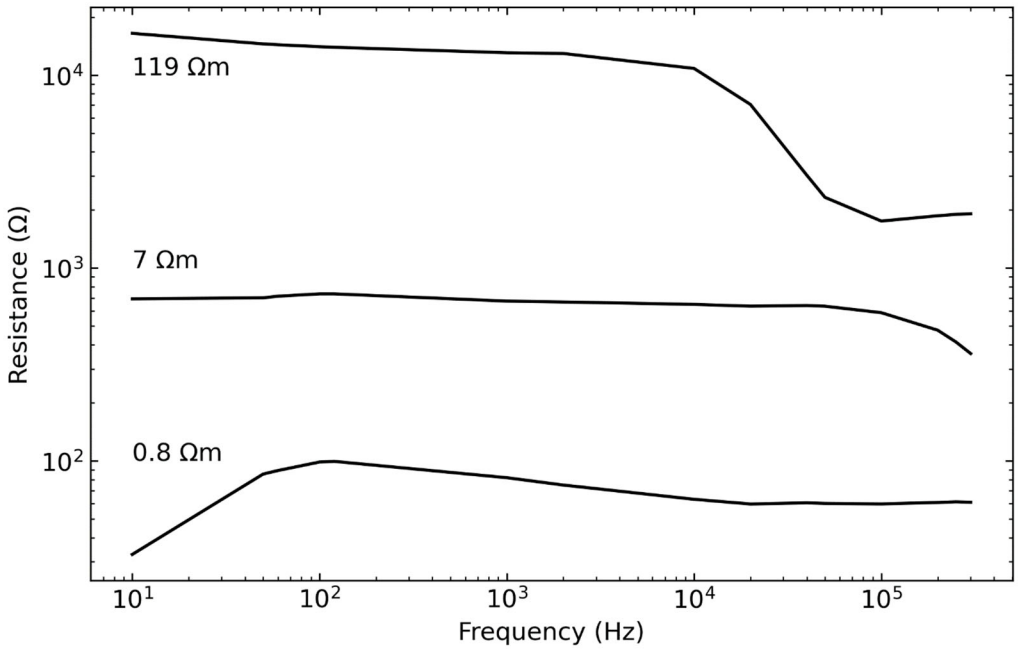


Figure 2.9 Frequency sweep results of resistance with respect to frequency for three NaCl solutions showing optical frequency range between 100 Hz and 10,000 Hz

Electrical resonance begins around 10,000 Hz as shown by the steep decrease in resistance in the solution with a ρ of 119 Ωm. Any frequencies above 10,000 Hz were deemed unsuitable for testing. Frequencies below 100 Hz were also unstable as shown in the solution with a ρ of 0.8 Ωm where the resistance is not constant. The frequency chosen for testing was 1,000 Hz, which lies below the electrical resonance range.

For the tests that were conducted at 1,000 Hz, the optimal frequency for testing, the geometric factor was calculated using Equation 2.9 for different heights of the top cap above the bottom cap. The geometric factor was calculated for heights of 2.25 in, 2.5 in, 2.75 in, and 3 in above the bottom cap to see how the geometric factor changes. The results are plotted in Figure 2.10.

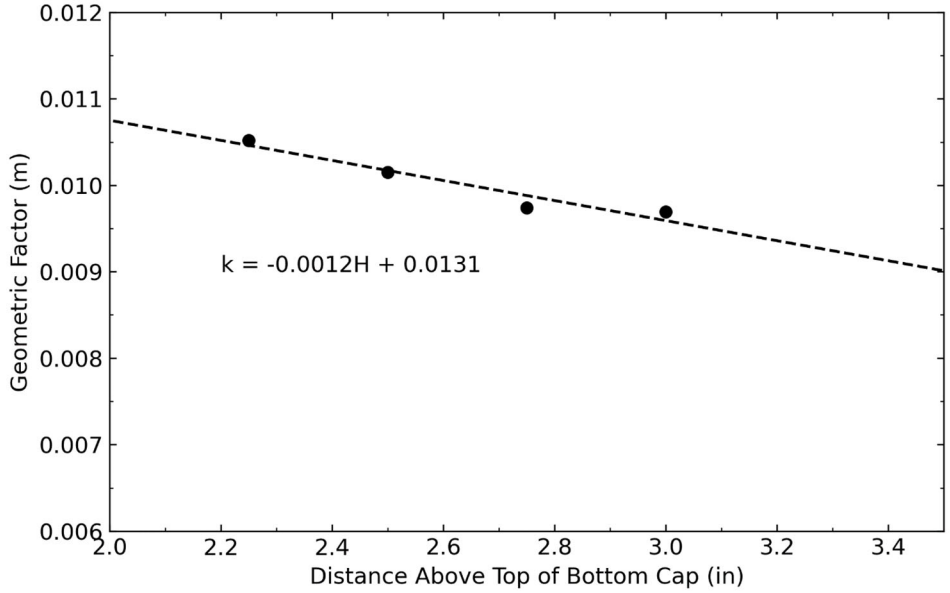


Figure 2.10 Geometric factor trendline to account for changes in sample height during testing

There is some variation in the geometric factor with height. As the distance between the top and bottom cap increases, the geometric factor decreases. Due to this slight variation in the geometric factor, for testing, the geometric factor was calculated for each load to account for changes in the soil cell dimensions through a regression line shown by Equation 2.13:

$$k = -0.012h + 0.0131 \quad (2.13)$$

where h is the height of the soil in inches and k is measured in meters.

2.5 Laboratory Investigation Results

The purpose of the laboratory testing is to understand how certain geologic and environmental conditions affect *Vs* and ρ to better interpret the MASW and EMI subsurface profiles. The geotechnical conditions analyzed included particle grain size, particle angularity, presence of organics, and soil gradation. The measured *Vs* and ρ values for all tests are listed in Appendix A.

For each geotechnical condition comparison, the soils used to look at the effects were plotted on the same plot with respect to *Vs* or ρ as a function of vertical stress for each combination of water content and relative density. Figure 2.11. shows an example of how the data was plotted for the gradation analysis for *Vs* using the Nevada Sand, represented as a black square, and the C109-Nevada Sand mixture represented as a white circle. The figures for the other comparisons can be found in Appendix A.

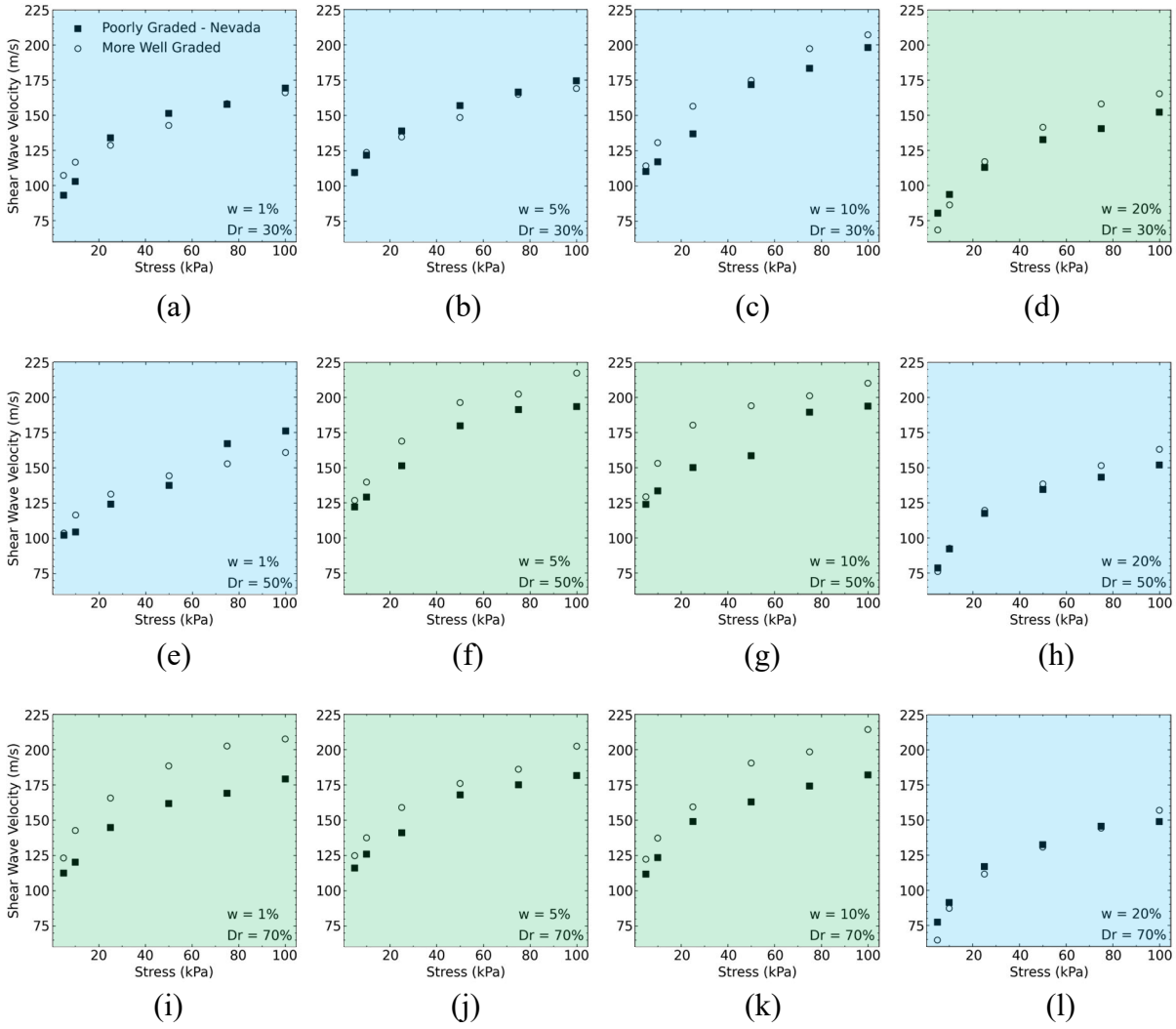


Figure 2.11 Shear wave velocity results between Nevada Sand and the C109-Nevada Mixture at (a) w=1%, Dr=30%, (b) w=5%, Dr30%, (c) w=10%, Dr=30%, (d) w=20%, Dr=30%, (e) w=1%, Dr=50%, (f) w=5%, Dr=50%, (g) w=10%, Dr=50%, (h) w=20%, Dr=50%, (i) w=1%, Dr=70%, (j) w=5%, Dr=70%, (k) w=10%, Dr=70%, (l) w=20%, Dr=70%

For these figures, as the plots move horizontally, the water content of the test increases. As the plots move down vertically, the Dr of the test increases. Each subplot has two markers, each representing the two soils being compared for that condition. To understand how these soils are affecting the *Vs* and ρ, one soil was assigned as Soil A and the other was assigned Soil B. Each

subplot in Figure 2.11 was analyzed to determine if Soil A had higher values, if Soil B had higher values, or if there was little difference between Soil A and Soil B. Once this was determined, the subplot was assigned a color to match the result. If the subplot showed that Soil A had higher values, the subplot was represented as red. If the subplot showed that Soil B had higher values, it was represented as green, and if the soils had similar values, then the subplot was represented as blue. The majority vote between the subplots results became the assignment for the condition.

In Figure 2.11, tests (a), (b), (c), (e), (h), and (l) show little difference in *Vs* between Nevada Sand and the C109-Nevada Sand Mixture. Since these tests had little change, they were assigned blue. For tests (d), (f), (g), (i), (j), and (k), the C109-Nevada Sand Mixture (Soil B) had higher *Vs* than Nevada Sand (Soil A). As a result, the tests were assigned green. Since six tests were assigned blue and six tests were assigned green, there was no majority vote. As a result, these tests showed that a more well graded soil would have higher *Vs,* but it could be difficult to distinguish in data.

2.5.1 Shear Wave Velocity Results

The results of the *Vs* laboratory testing are shown in tabular format in Figure 2.12. Each test for each soil was given a color as described in Section 2.5 which represents which soil had a higher *Vs*. Analyzing the results for each test allows for the understanding of how the geologic or environmental condition affects the *Vs* overall, but also shows if water content and relative density also affect the *Vs*. The final classification is shown in the right-most column as either one or two colors depending on the majority vote.

Particle Angularity: the particle angularity testing showed that overall, the soil with the rounded particles had higher *Vs* than the soil with the angular particles. 75% of the tests showed rounded particles with higher *Vs*, 17% of the tests showed similar *Vs* between the two soils, and 8% of the tests showed angular particles had higher *Vs*. The majority vote was that rounded particle sands had higher *Vs* than angular particle sands.

This result comes with the caveat that the difference in *Vs* between sands of different particle angularity is more pronounced at lower water contents. At higher water contents, it is still possible to distinguish between sands with different particle angularity using *Vs*, but it becomes more difficult.

Particle Grain Size: Four soils were used to test the effect of particle grain size on *Vs*. The first pair of soils were the quartz based C109 Sand and Nevada Sand. These soils showed that 33% of the tests resulted in coarser particle sands having higher *Vs* while 42% of the tests showed that particle grain size had very little to no effect on the *Vs*. Because these percentages were similar, the final classification for this pair of soils was that coarser particle sands tended to have higher *Vs* at lower water contents, and at higher water contents (above 5%), it is hard to distinguish between particle size.

The second pair of soils used to test particle grain size was the New Zealand fine and coarse sands. The tests conducted with these soils showed 33% of the tests not being able to distinguish between the soils, 33% of the tests showed coarser sands had higher *Vs*, and 33% of the tests showed that finer sands had higher *Vs*. Since there was an equal distribution, the final classification was determined as “hard to distinguish” between grain size using *Vs*. This classification matched the classification chosen for the quartz-based sands.

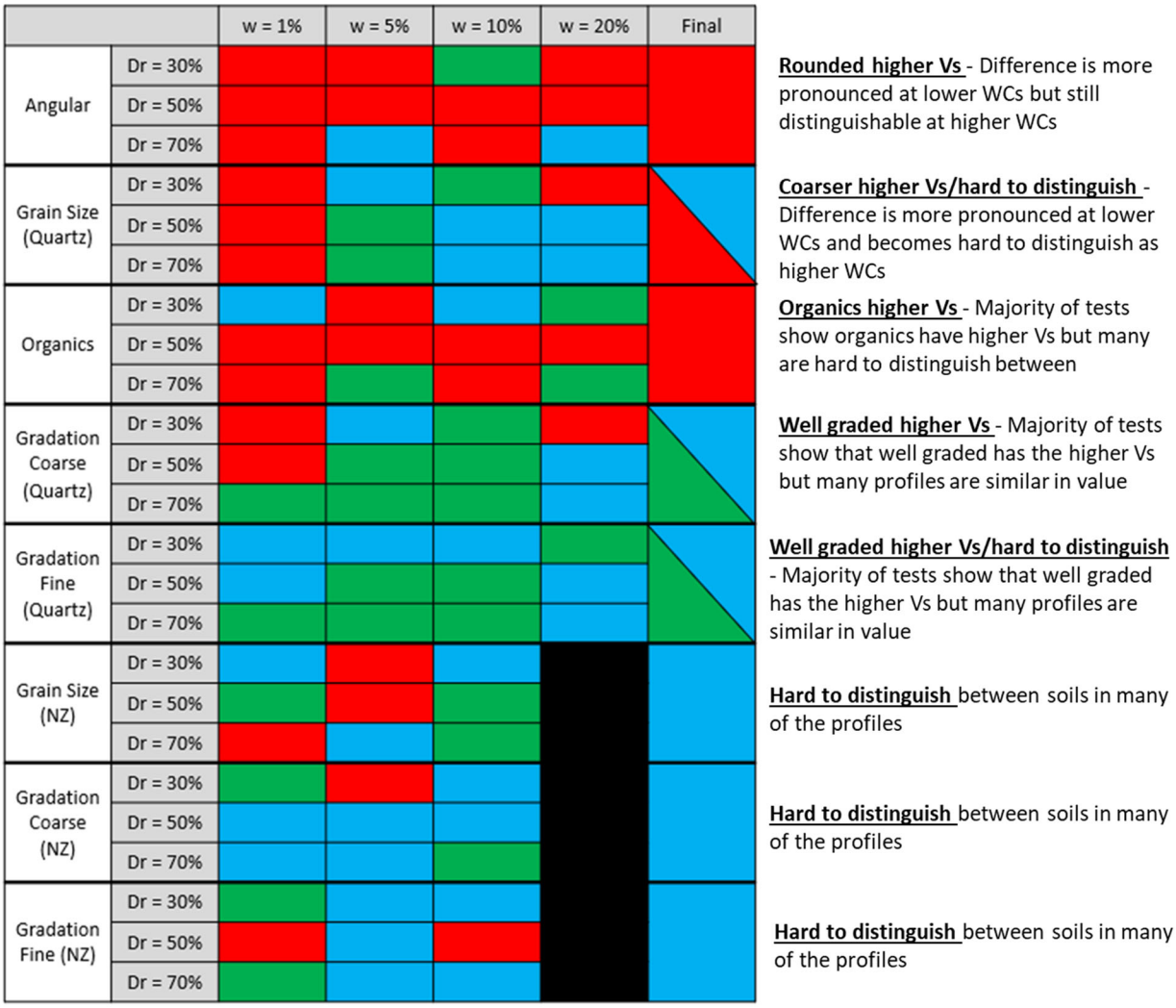


Figure 2.12 Showing the results of the shear wave velocity laboratory testing for each geotechnical condition

In unsaturated conditions, as water content decreases, a higher capillary force (Pc) is expected and can cause higher *Vs* in soils with finer grain particles (Kim, Won, & Park, 2021). To understand if the grain size or capillary pressure is causing these changes in *Vs*, the specific surface (Ss) and Pc were calculated for Nevada Sand, C109 Sand, as well as the New Zealand fine and coarse sands using Equations 2.14 and 2.15 defined by (Santamarina, Klein, Wang, & Prencke, 2002):

$$Ss = \frac{3(C_u+7)}{4 d_w G_s D_{50}} \tag{2.14}$$

$$Pc = 2\frac{T_s S_s d_m}{e} \tag{2.15}$$

Where Ss is measured in m^2/g, d_w is the density of water, G_s is the specific gravity of the soil, Ts is the interfacial tension (N/m) and assumed to be 0.072 N/m, d_m is the mineral density (kg/m^3) and assumed to be 2650 kg/m^3 for all soils, and e is the void ratio. The Ss and Pc were calculated for the tests conducted at a w = 1%, as this water content could be low enough for Pc to affect the *Vs*. The results are summarized in Table 2.2.

Table 2.2 Specific Surface and Capillary Force

Soil	Ss (m^2/g)	Pc (kPa)		
		Dr = 30%	Dr = 50%	Dr = 70%
Nevada Sand	0.007	7.6	8.1	9.0
C109 Sand	0.014	3.8	4.1	4.4
New Zealand – Fine	0.01	3.5	4.0	4.6
New Zealand – Course	0.004	1.9	2.1	2.4

For C109 Sand and Nevada Sand as well as the two New Zealand soils, the finer particle soil exhibited lower Pc compared to their coarser grained counterparts. This is to be expected but since the difference in Pc between the finer particle soil and their respective coarser grained soil is negligible, it can be concluded that the change in *Vs* between the soils is attributed to particle size rather than Pc.

Soil Gradation (fine): Gradation was tested using the quartz-based sands as well as the New Zealand sands using the finer particle sands. The Nevada Sand was compared to the C109-Nevada Sand mixture and showed that in 50% of the tests, it was hard to distinguish between the two soils using *Vs* and the other 50% of the tests showed that the C109-Nevada Sand Mixture had higher *Vs* than the Nevada Sand. This pair of soils were classified as both hard to distinguish and well graded soil had higher *Vs*.

The other pair of soils used for testing soil gradation was the New Zealand well graded soil against the New Zealand finer particle soil. 22% of the tests showed that the well graded soil had higher *Vs*, 22% of the tests showed that the finer poorly graded soil had higher *Vs*, and 56% of the tests showed that it was hard to distinguish between the two soils. The final classification was harder to distinguish between soil gradation when using *Vs*.

Soil Gradation (coarse): Gradation was also tested by comparing the well graded soils to the coarser particle soil counterpart. The quartz-based sands showed that for 50% of the tests, the well graded soil had higher *Vs* but many of these were close in value. 25% of the tests showed the poorly graded soil had higher *Vs* while for the remaining 25% of the tests, it was too hard to distinguish between the soils. The final classification for this soil combination was well graded soils have higher *Vs* and it is hard to distinguish. This classification was selected even though 50% of the tests involved well-graded soils with higher *Vs*. This decision was made because, among those tests within the 50% category, approximately half of them displayed *Vs* that were exceedingly similar, making them challenging to differentiate.

The second gradation testing using coarse soils was with the New Zealand soils. 67% of these tests were classified as hard to distinguish between when using *Vs*. 22% of the tests showed the well graded soil having higher *Vs* and 11% of the tests had the poorly graded soil having higher *Vs*. Like the other gradation comparisons, this set of tests was also classified as hard to distinguish.

Organics: Overall, with 58% of the tests, soils with organic material such as leaves and twigs, resulted in having higher *Vs* than soils without organics. 25% of the tests showed the organics had

higher *Vs* and 17% of the tests showed no difference in *Vs*. The final classification was that soils with organics have higher *Vs* than soils without organics.

Even though most of the tests showed clear trends with organic material having higher *Vs* than soils without organics, the difference in *Vs* between the two soils was not significant. This is most likely because of the low organic content in the soil, which was measured to be approximately 1% of the total weight.

2.5.2 Electrical Resistivity Results

Like the *Vs* analysis for different geologic and environmental conditions, the results from the ρ laboratory testing were analyzed in a tabular format using the same color scheme as with *Vs*. The results for ρ are shown in Figure 2.13.

Particle Angularity: 75% of the tests for looking at the effect of particle angularity on ρ resulted in the rounded particle soil having higher ρ. 17% of the tests were too hard to distinguish between the soils and 8% of the tests showed the angular particles had higher ρ. The final classification was that soils with rounded particles have higher ρ than soils with angular particles.

Particle Grain Size: The quartz-based sands used to test the effect of particle grain size on ρ showed that 75% of the tests resulted in the coarser particle grain size having higher ρ than the finer particle grain size. 8% of the tests showed the opposite with the finer particle grain size sand with higher ρ and in the remaining 17% of the tests, it was too hard to distinguish between particle grain size using ρ. The final classification result was that soils with coarser particle grain size will have higher ρ than soils with finer particle grain size.

The New Zealand soils that tested grain size resulted in the same final classification as the quartz-based grain size tests with 78% of the tests showing higher ρ in the soil with the coarser grain size. For these soils, at higher water contents, the difference in ρ between the soils decreased significantly. This is because ρ is sensitive to clay content and water content. As the water content increased, the ρ significantly decreased and dominated the reading.

Soil Gradation (fine): The New Zealand and C109-Nevada Sand Mixture (against the fine poorly graded sands) both showed that most tests resulted in the well graded soil having higher ρ than the fine poorly graded soil. The quartz-based sands had 75% of the tests showing the well graded soil had higher ρ than the poorly graded soil. Similarly, 67% of the tests using the New Zealand soils also showed the well graded soil having higher ρ than the poorly graded soil.

As the water content increased for the New Zealand tests, it became difficult to distinguish between poorly graded and well graded soil using ρ because the water content was dominating the reading.

Soil Gradation (coarse): The gradation tests using the coarser grained poorly graded soil had different results when compared to the gradation tests when using the finer grained poorly graded soil. The tests using the quartz-based soils showed that 75% of the tests resulted in the poorly graded soil having higher ρ compared to the well graded soil. At a water content of 20% though, it became hard to distinguish between these soils for two out of the three tests.

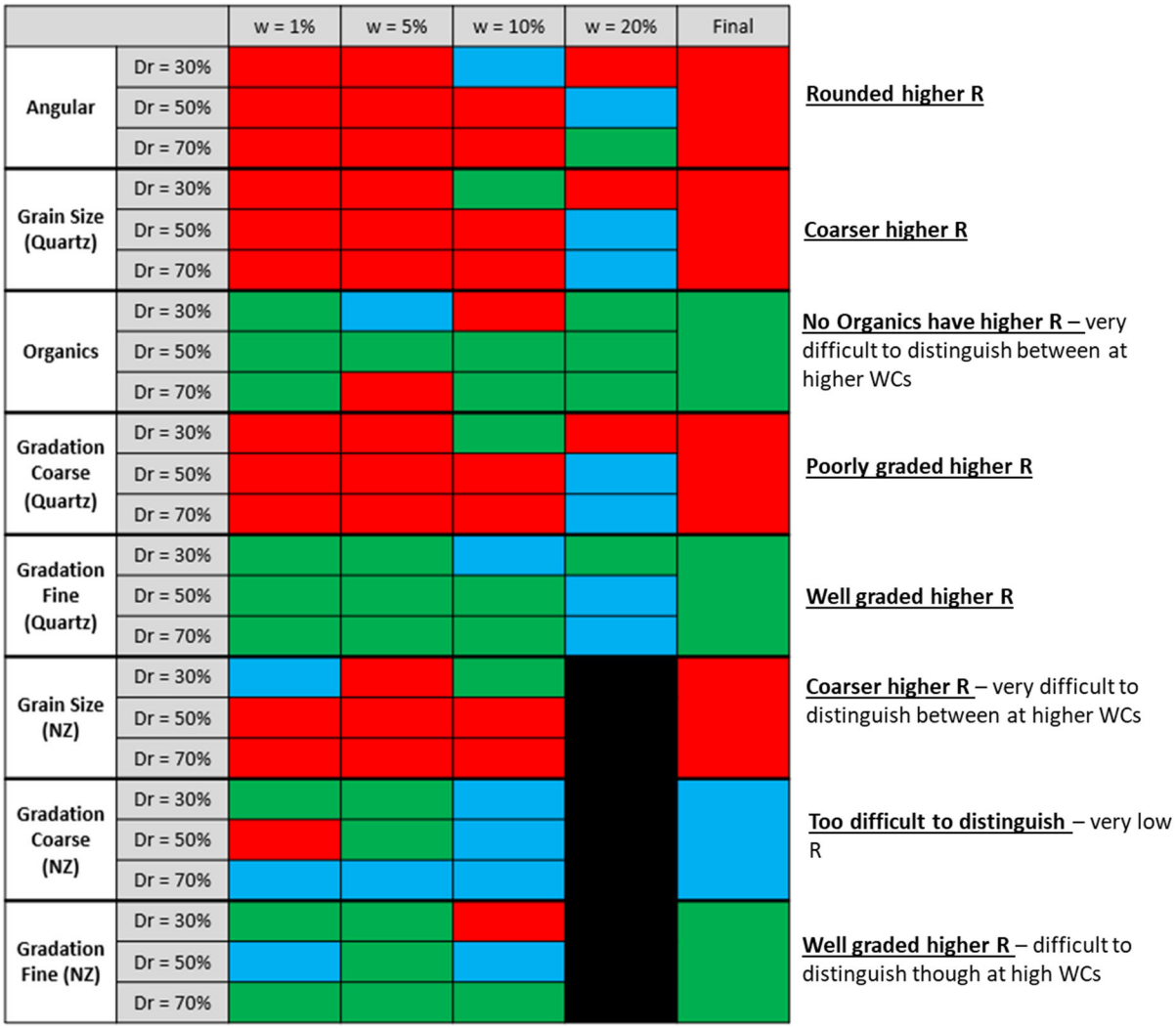


Figure 2.13 Showing the results of the electrical resistivity laboratory testing for each geotechnical condition

In comparison, the New Zealand soils did not show clear results. 11% of the tests showed the poorly graded soil had higher ρ while 33% of the tests showed the opposite. For 56% of the tests, it was too hard to distinguish between the soils using ρ.

Organics: The final classification for organics was that soils without organics have higher ρ than soils with organics. This was found true at all water contents but became more difficult to distinguish between at higher water contents where the water content starts to dominate the ρ readings. It was found that 75% of the tests showed the soil without organics had higher ρ, while 17% of the tests showed the opposite, and 8% of the tests it was too hard to distinguish between the soils.

2.5.3 Validation of Laboratory Shear Wave Velocity

The experimental *Vs* testing setup was validated by comparing the results tested on the Ottawa C109 Sand with the results from Roberston et al. (1995). The Robertson et al. (1995) study used Ottawa C109 Sand which was prepared with moist tamping or water pluviation, and the water content of the samples was not stated. The data from Robertson et al.'s testing used for comparison included both the moist tamping and water-pluviated data with vertical stresses between 50 and 100 kPa. The comparison results are shown in Figure 2.14.

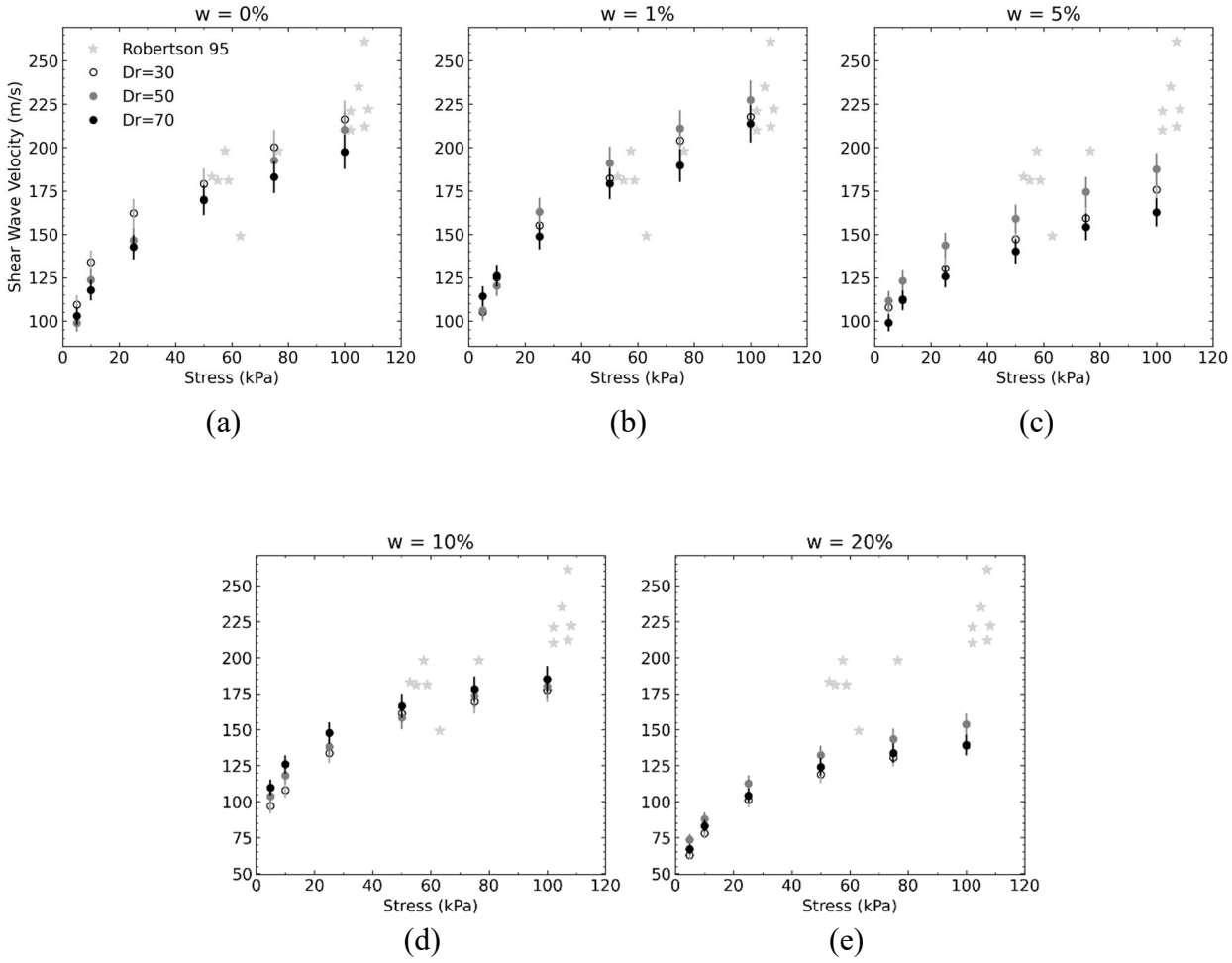


Figure 2.14 Ottawa C109 Sand results compared to Robertson et al. 1995 for validation of testing procedure (a) w = 0%, (b) w = 1%, (c) w = 5%, (d) w = 10%, and (e) w = 20%

Since the water content and relative density of the Robertson et al. data were not specified in the paper, the data was compared to all water contents and relative densities tested in this laboratory setup. The data represented as circles were part of this study while the grey starts were data from the Robertson et al study. When the water content in this study was low at 0% and 1%, the data was similar to the data from Robertson et al. and followed the same trajectory, validating the experimental testing procedure used in this study to measure *Vs*.

As water content increases to 5%, 10%, and 20%, the measured *Vs* in this study is less than the Robertson et al. data. This trend is more noticeable at higher vertical stresses. At lower vertical stresses, the 5% and 10% water content data were similar to the Robertson et al. data.

2.5.4 Validation of Laboratory Electrical Resistivity

To validate the ρ results, the values were compared to values found in the literature. Figure 2.16 shows a compilation of ρ values from 11 sources across a range of soil types from gravels to clays. This shows the data independent of water content.

The source of the values is on the y-axis with the ρ on the x-axis. The bars represent the range of values and the color of the bar represents the soil type. The data from literature shows clays and loams have low ρ while sands and silts have higher ρ. The data collected in this study falls within the same range as other sandy soils. Some of these measurements in literature also state the water content they were collected at. The data collected in this study was also compared with data associated with different water contents in Figure 2.15, since water contents significantly impacts the ρ.

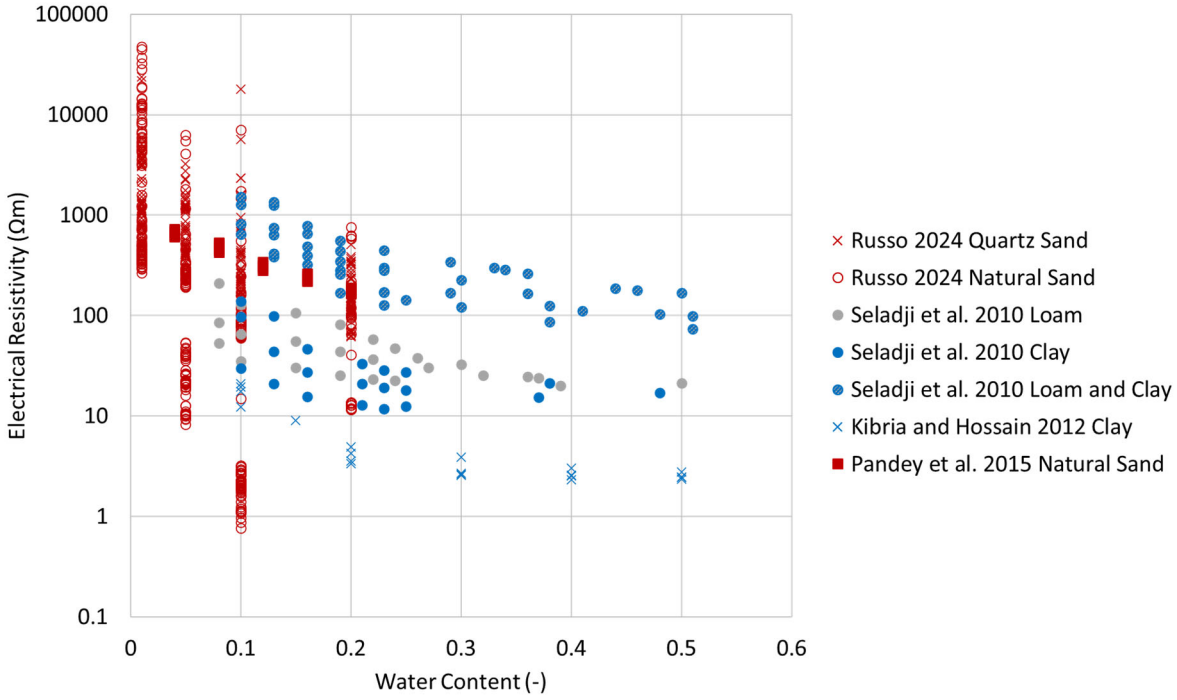


Figure 2.15 Literature data with specified water contents

The ρ is on the y-axis with the water content on the x-axis. The different studies are shown by different marker symbols and the soil type is represented by different colors. The data collected in this study follows the same trends as the data collected from other studies where there is a quick decay in ρ as water content increases.

The study that most closely resembled this study was the Pandey et al. (2015) study. As a result, the data from this study was compared individually with the Pandey et al. (2015) study

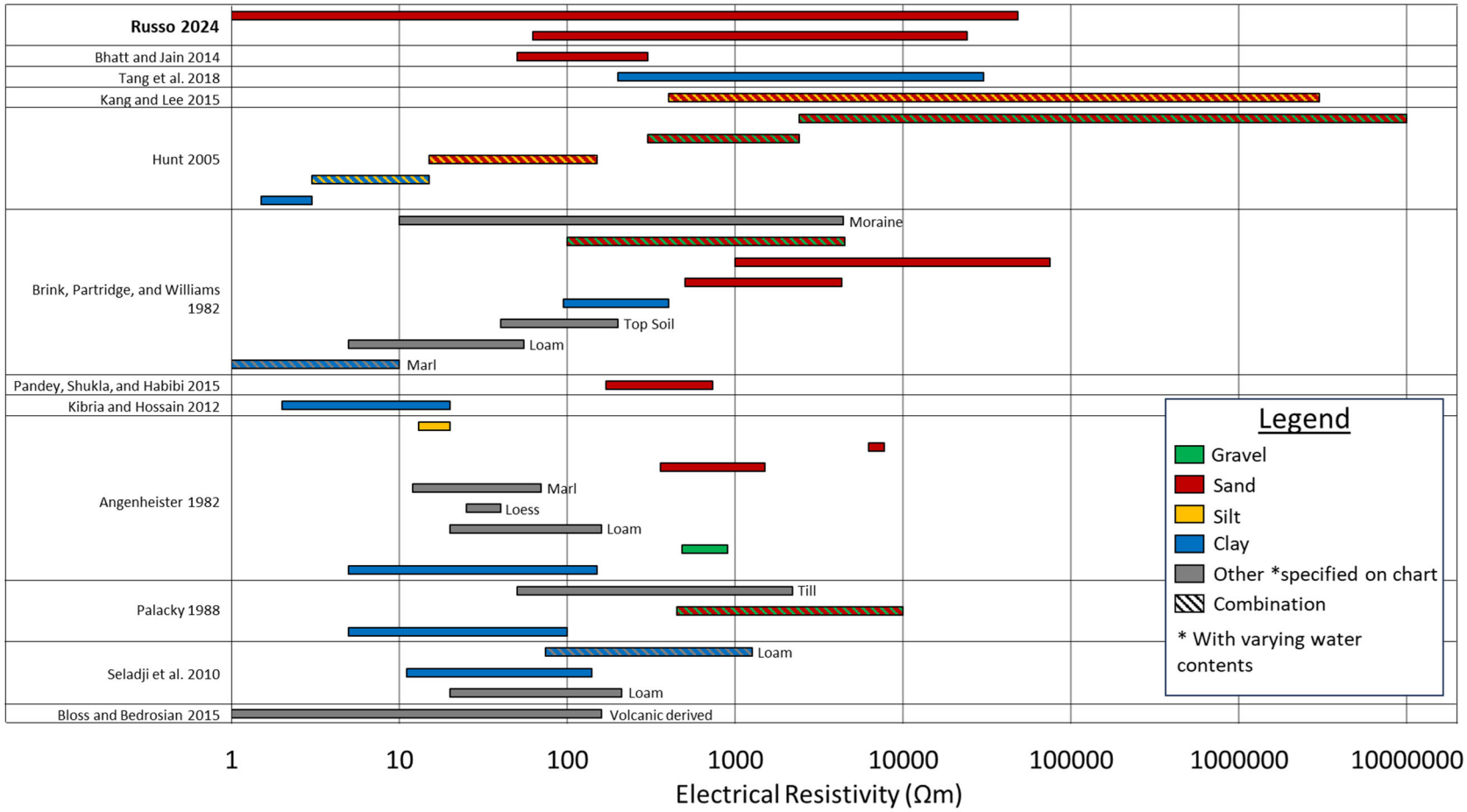


Figure 2.16 Compilation of electrical resistivity data from literature

which used a poorly graded sandy soil with distilled water, with water contents measured at 4%, 8%, 12%, 16%, and 20% and relative densities ranging from 0% to 100%. The Pandey et al. (2015) study also used stainless-steel electrodes, similar to this study. The comparison between the results from this study and the Pandey et al. (2015) study is shown in Figure 2.17.

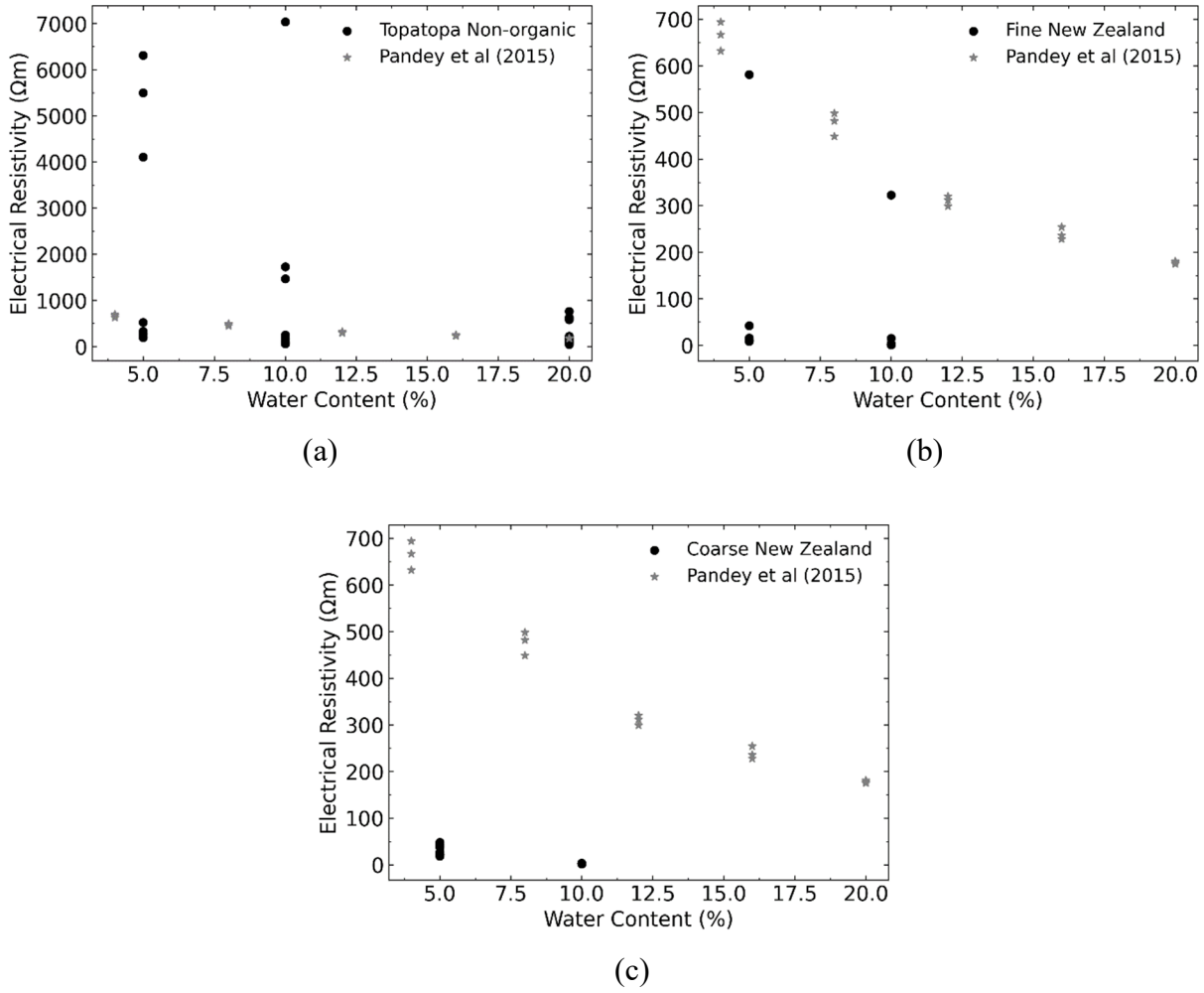


Figure 2.17 Comparison of electrical resistivity results with Pandey et al. (2015) with (a) Topatopa without organics, (b) fine New Zealand Sand, and (c) coarse New Zealand Sand

Because the chemical composition of soils affects ρ, the data from Pandey et al. was plotted against the Topatopa non-organic soil, the coarse New Zealand sand, and the fine New Zealand sand. The data from this study is shown as black circles while the data from the Pandey et al. study is represented as grey stars. Looking at the results, the New Zealand soil consistently had smaller ρ compared to the sand from Pandey et al. This could be attributed to differing chemical compositions. The Topatopa non-organic sand falls within a similar range of the Pandey et al. data, validating the ρ laboratory experiment. The results from the Topatopa data and Pandey et al. show the same trends where the ρ decreases with increasing water content, where there is a higher decrease in ρ at lower water contents, and then a slower decrease at higher water contents.

2.6 The Importance of Geotechnical Conditions

The geotechnical conditions controlled for in the laboratory were chosen as they could contribute to internal failure mechanisms such as seepage. The Hazen Method, shown in Equation 2.16, is a relationship derived between particle grain size and hydraulic conductivity showing that soils with coarser particles are more likely to contribute to seepage than soils with finer particles (Hazen, 1911).

$$K = C(d_{10})^2 \tag{2.16}$$

Where K is the hydraulic conductivity, C is an empirical constant, typically assumed to be 1, and d_{10} is the grain size where 10% of the soil is finer. As the size of the particle increases, there is an increase in hydraulic conductivity which could lead to more seepage. Poorly graded soil compared to well graded soil is also more likely to lead to seepage which is supported by Equation 2.17 (Kozeny, 1927):

$$\mu = \frac{Cd^2p^3}{(1-p)^2} \tag{2.17}$$

where μ is the soil permeability, d is related to particle grain size, and p is the porosity of the soil. Well graded soils have multiple grain sizes resulting in a lower porosity since smaller grain sizes can fit between larger grain sizes, reducing the void space in the soil matrix. As a result, soils that are poorly graded, which have uniform grain sizes and more void space, will have a higher porosity which leads to an increase in the permeability of the soil. With higher permeability, seepage is more likely to occur as it allows for a greater flow of fluids through the soil.

Particle angularity is another geotechnical condition that affects the porosity of the soil which could contribute to seepage because more angular particles increase the void ratio of the soil (Shin & Santamarina, 2013). Equation 2.18 shows how porosity is a function of the void ratio of the soil:

$$p = \frac{e}{1+e} \tag{2.18}$$

where e is the void ratio. As the void ratio of the soil increases, the porosity of the soil increases. As a result, if the soil had more angular particles, it would have a higher porosity than soil with more rounded particles.

The amount of organic matter in the soil was also of interest as organics, especially vegetation, could cause preferred conduits for water movement through large tree trunks. Vegetation can also cause stabilization of slopes against erosion and banks against water currents during storms.

2.7 Summary of Laboratory Results for Field Profile Interpretation

The goal of the laboratory testing is to understand how *Vs* and *ρ* changed based on a variety of geologic and environmental conditions for interpreting field profiles to help detect potential problematic zones in the subsurface that could potentially lead to failure. To achieve this, the results from the laboratory testing are summarized in Table 2.3. focusing on how *Vs* and *ρ* change in tandem.

Table 2.3 Summary Results of Laboratory Shear Wave Velocity and Electrical Resistivity

ID	Change in Shear Wave Velocity	Change in Resistivity	Potential Factor (s)
1	No Change	Increased	Coarser particles (higher WCs), Well graded soils (finer soils)
2	No Change	Decreased	Finer particles (higher WCs), Poorly graded soils (coarser soils)
3	Increased	Increased	Rounded particles, Coarser particles (at low WCs)
4	Increased	Decreased	Soils with organics, well graded soils (coarser soils)
5	Increased	No Change	No test results
6	Decreased	Increased	Soils without organics, poorly graded soils (coarser soils)
7	Decreased	Decreased	Angular particles, Finer particles (at low WCs)
8	Decreased	No Change	No test results

For example, when the MASW profile shows very little change in *Vs* at a certain location, but the EMI profile shows an increase in ρ compared to the rest of the profile, it could potentially be attributed to coarser particle soils in that location or a more well graded soil at that location. By connecting locations in the profile with potential geotechnical conditions, the field profiles become more interpretable for understanding subsurface conditions.

Chapter 3. Field Data Collection and Remote Sensing Analysis

3.1 Geophysical and Remote Sensing Literature Review

With current levee inspection practices being predominantly visual through human observation, non-invasive sensing techniques can be leveraged to identify surface and subsurface conditions. Geophysical methods are non-invasive methods that can image the subsurface through seismic waves or electrical currents to understand lithological variation and subsurface structures. Remote sensing uses active and passive sensors which are non-invasive tools that have been found to effectively detect surface conditions that could lead to failure and certain sensors have been found effective at imaging near-surface failure mechanisms, such as animal burrows.

3.1.1 Geophysical Methods for Subsurface Investigation

Understanding the subsurface variability is an important component of analyzing levee health. There are many techniques available to image the subsurface, including invasive methods like downhole seismic tests, CPTs, boreholes, inclinometers, and piezometers. A newer method of using fiber optics has been gaining popularity for imaging the subsurface. Once the fiber is installed in the levee, it can non-invasively provide subsurface information. Wang et al. (2016) used Distributed temperature sensing (DTS) fiber optics in embankments to monitor the seepage velocity. The experiment utilized a flume testing platform with scaled-down earthen embankment dams. The velocity of water was controlled in the experiment to look at the velocity effect on seepage. These laboratory experiments proved to be successful in using fiber optics to monitor seepage. The testing had good precision and showed the promise for real-time seepage monitoring of earthen dams.

DTS fiber optics has also been used for the detection of backward erosion piping in levees. A large-scale experimental levee was built in the Netherlands for this research by Bersan, Koelewijn, and Simonini (2018). The model levee was built with DTS fiber optics, pore pressure transducers, liquid-level sensors, and flowmeters. The fiber optic cables were installed during the construction

of the levee and were placed along the boundary between the bottom sand layer of the levee and the foundation as well as on the downstream slope. Water was introduced in a controlled manner in the study until the levee collapsed. The results of this work showed that the optimal location for the DTS fiber optic cables with respect to monitoring backward erosion piping was close to the landslide toe of the levee.

While studies have shown the promise of using fiber optics for subsurface levee investigations, it is not always the most practical solution. Fiber optics must be installed in the levee structure, which is problematic for existing levees. A study performed by Sekula et al. (2017) created a system for monitoring the changes within the core of an experimental earthen levee in Czernichow, Poland. The experimental levee was at a 1:1 scale and equipped with fiber optics, earth pressure sensors, pore water pressure sensors, and thermistors. The study aimed to use the combination of these sensors to detect filtration and erosion processes. The conclusion made by the authors was that this system works well for newly built or renovated levees. For levees that already exist, this monitoring system is advised against since it would be costly to install the invasive sensors and could damage the pre-existing structure. With the majority of the levees in the United States, already built, the use of fiber optics is not the most efficient method for imaging their subsurface.

Instead, using fully non-invasive methods such as geophysical sensors can be leveraged to gain continuous measurements of the subsurface. There are a variety of geophysical sensors available that utilize seismic waves, direct electric currents, electromagnetic waves, and magnetic fields to characterize subsurface variability. Seismic methods are popular geophysical sensors as they can be used to map lithology changes, structures, and soil stiffness. A study conducted by Osazuwa and Chinedu (2008) used the seismic refraction technique to detect highly permeable zones within an earthen dam. It was believed that these permeable zones dominate the flow regime of the system and influence seepage, as a result, it is important to be able to locate these zones. The seismic refraction surveys were conducted parallel and perpendicular to the axis of the dam. The seismic refraction data identified five thick deposits of sediment and the location of the basement rock. The data was also effective at detecting highly permeable subsurface channels that could cause seepage. When compared to ground truth borehole logs, the geophysical data correlated well.

Research performed by Woolery (2018) also showed the effectiveness of using seismic geophysical methods for the identification of subsurface structures. SH-wave seismic reflection methods were used at two earthen dams in this study to image and locate internal boundaries in the embankment. The seismic geophysical data was able to identify the bedrock foundation as well as colluvium wedges and alluvium deposits within the dam. These methods work well as different geologic units will exhibit different seismic velocities, allowing for the detection of these structures.

With the ease of using seismic methods for levee inspection, there are also some challenges as identified by Ivanov et al. (2017). These authors used Multichannel Analysis of Surface Waves (MASW), a surface waves seismic technique, on levees in Southern Texas. Traditionally, MASW uses Rayleigh waves, but the authors used Love Waves to image the levees instead. The contribution of this work was that the geometry of the levee can cause an overestimation of Rayleigh wave velocities which can affect the measured shear wave velocity (*Vs*).

Where seismic wave techniques are useful for identifying structures and lithology changes, electrical and electromagnetic geophysical sensors have proven to be useful imaging techniques for levees as the electrical resistivity (ρ) of the subsurface is significantly impacted by water content, making them ideal for identifying seepage zones. For example, Raji and Adedoyin (2020) used Electrical Resistivity Tomography (ERT) with a dipole-dipole array to detect structural anomalies in Unilorin Dam, an earthen dam in Nigeria. The authors used 2D ρ mapping in combination with geologic mapping of the surface and subsurface of the dam. The authors found erosional channels within the rock and cavities using ρ data. These structures can act as seepage conduits for the dam.

The electrical and electromagnetic geophysical methods are also useful tools for detecting subsurface layers with contrasting ρ as shown by Asch et al. (2008). The authors used a combination of ERT, Electromagnetic Induction (EMI), and Capacitively-Coupled Resistivity (CCR) to map the thickness of sand lenses and depth to clay in the American River levee system in Sacramento. There was a need to map the sand lenses as there were concerns of these lenses eroding and leading to levee failure. Overall, the study was successful at detecting the sand lenses due to the high contrast in ρ between sand and clay. These methods were also successful at determining the depth of the clay zones, ranging from 10 ft to 60 ft. This study highlighted a disadvantage of using EMI for this application, though. In highly resistive terrains, the environment was too noisy and caused the EMI inversions to be unsuccessful. Depending on site geology, certain geophysical methods are not appropriate. As a result, it is always important to perform a preliminary site investigation and to use a combination of techniques.

Other work performed by Stepler, Witten, and Slatt (2004) further highlights the effectiveness of using these methods for delineating structures in the subsurface. This study used a combination of EMI and Ground Penetrating Radar (GPR) on a levee to characterize the lateral and basal geologic boundaries between sandstones and shales. These methods were successful in imaging the lateral boundaries but had difficulty in imaging the basal boundaries since EMI and GPR are not the best methods for deep investigations. This study also highlighted the advantage of efficiency for EMI that was not present with GPR since EMI does not have to touch the ground and is unaffected by vegetation density (if it can be walked through).

Smaller structures than geologic boundaries have also been imaged using these methods. A study performed by Tanajewski and Bakula (2016) wanted to locate animal burrows and tunnels in a levee. The study leveraged GPR with an integrated Global Positioning System (GPS) to locate these structures. As data was being collected using GPR, the GPS signals were sent to the GPR controller to record the position of each data point. Due to the GPS signal being unable to update as quickly as the GPR system, some GPR traces had the wrong coordinate and affected the calculation of the total length of the geophysical profile. This study also found that GPR was able to detect animal burrows but with only approximate shapes and depths. Even though the detection was approximate, it will still allow for more targeted maintenance in these areas to have a better understanding of the size and shape of these structures.

A study performed by Tresoldi et al. (2018) used a combination of EMI and GPR to develop a framework called Tech-Levee-Watch. This framework combined vehicle-mounted EMI and GPR

sensors to create a system that can diagnose, monitor, and alert to vulnerabilities in levees. This study was able to successfully record the ρ of the levee every 24 hours. The data was sent through an internet connection and used to create maps of the subsurface showing the water content distribution. Daily maps like this can be used to understand water flow through the levee which can help give early warning towards failure.

While there are many benefits to using electrical current and electromagnetic sensors, they also have some limitations as shown by Weller et al. (2014). This research developed a monitoring system for levees on the Red River in Vietnam by combining a permanent ERT array (half-Wenner configuration), frequency domain reflectometry sensors, tensiometers, and observation wells installed with pressure sensors. By combining these instruments, the goal was to monitor the natural variations of the water level on the wet side of the levee. This system was deployed in natural conditions for five years to monitor water levels. The first issue with the system was that the frequency domain reflectometry sensors had a high failure rate due to the subtropical region. The second issue was that the seasonal variation in precipitation between the dry periods and monsoon periods caused high ρ changes in the near-surface. This was problematic as these zones make it difficult to image deeper seepage zones in the levee.

Another major limitation of these sensors was described by Lin et al. (2014). These authors used ERT with a Wenner configuration to detect leakage in an earthen dam. The arrays were installed on the left abutment, dam crest, and the downstream shell. While the study was successful in characterizing hydrophysical properties under homogeneous geological conditions, it became challenging to characterize these properties when there were heterogeneous geologic conditions. Often, natural depositions of soil are not homogeneous, which could cause uncertainty in the interpretation of these types of datasets.

While using just seismic waves or just electrical currents and electromagnetic waves can give important information about subsurface health, it is more beneficial to use these geophysical sensors in tandem. One of the earliest research projects conducted for levee applications using these sensors in tandem was by Inazaki and Sakamoto (2005). This study used MASW, CCR, EMI, seismic reflection, and ERT to assess the failure vulnerability of levees during floods and earthquakes. The CCR and EMI datasets were successful in detecting anomalously high and low ρ zones in the levee body and the ρ was used to estimate permeability. Seismic reflection showed the stratigraphy of the levee along with the foundation soils while MASW imaged high and low *Vs* zones in the levee which were then used to estimate soil stiffness.

This work was further developed by Inazaki, Aoike, and Kaneko (2016). In this study, the authors integrated GPR, global navigation satellite system (GNSS), ERT, and high-resolution surface wave surveys to understand seepage in a levee along the Kokai River in Japan. This study found that by combining the geophysical results from these sensors, a 3D model can be made showing subsurface structures. The high ρ and high *Vs* zones found by these methods resulted in the detection of an irregularly shaped permeable layer.

Research conducted by Rahimi et al. (2019) also showed the effectiveness of using geophysical sensors in tandem for detecting internal structures. ERT, MASW, and CCR were combined to

image the earthen dam, Kinion Lake Dam. This dam had historical seepage and internal erosion issues. The surveys were performed on the crest and downstream toe of the dam. The study showed that MASW was successful at finding depth to bedrock but could not find locations of seepage. By also using CCR and ERT, the ρ found seepage paths in the dam by comparing seasonal data and detected highly fractured saturated limestone. This study showed the importance of using both seismic techniques and electrical current/electromagnetic techniques, as certain subsurface structures are highlighted differently.

Research has also shown the effectiveness of jointly interpreting the results of *Vs* and ρ from geophysical sensors. A study conducted by Poomvises et al. (2020) used ERT, MASW, and GPR data collected along the centerline of an earthen dam in Thailand to understand the spatial risk of the dam. These geophysical methods were chosen as they have different depth penetrations, with ERT penetrating to the foundation, MASW for the middle of the dam, and GPR focusing on the near-surface. To analyze the risk, the authors created a 5x5 matrix over the area of the dam where the X axis of the matrix represented the distance along the dam and the Y axis of the matrix represented the depth. This study showed how a spatially defined risk matrix can be used for understanding subsurface health. Where anomalies were found from each geophysical method, they were plotted into the corresponding matrix cell to show the spatial risk of the system.

Another approach to jointly interpret the results of *Vs* and ρ was to create a cross plot that was used to identify soil type. Work conducted by Hayashi et al. (2008) used MASW and CCR to obtain *Vs* and ρ measurements of the subsurface of a levee to estimate soil conditions. The surveys were performed on the crest and toe of the levee. The authors developed a cross plot where ρ was used to distinguish between clays and sands while *Vs* was used to distinguish between soft and stiff soil. Measurements that plotted in high ρ and low *Vs* were deemed to be dangerous as they represented loose sandy materials while measurements that plotted in low ρ and high *Vs* were deemed to be safe as they represented compacted clayey materials.

3.1.2 Remote Sensing for Surface Investigation

Surface variability is an important component of levee health. With current visual inspection practices for levees, a more robust framework is needed to understand surface conditions. Remote sensing techniques are non-invasive sensors with the ability to image large areas of the Earth's surface. The use of these methods can standardize inspection among different organizations by removing human bias from inspection.

There are many remote sensing techniques to choose from for imaging the surface, including satellite imagery, drone imagery, and Light Detection and Ranging (LIDAR) sensors. Landsat imagery and radar cover large spatial areas and are useful for change detection analyses as shown by Nguyen et al. (2018). The authors performed a change detection analysis from 2009 to 2014 using Interferometric Synthetic Aperture Radar (InSAR) data from the TerraSAR-X satellite. The study aimed to monitor settlement over time of levees in New Orleans. The settlement determined using the InSAR data had good agreement with the settlement observed using GPS and extensometers.

A study carried out by Ozer et al. (2019) also found that InSAR data from the TerraSAR-X satellite was a useful tool for measuring swelling and shrinking deformation over weeks on levees. This study showed that InSAR images with meter-level resolution were high enough for levee applications. One of the other benefits of this method described by the authors was the applicability of this method during the night and that it can operate in all weather conditions, which is important for monitoring levees during hurricanes where UAV-based imagery would be not applicable due to high wind and rain conditions.

Research conducted by Sivasankar et al. (2019) also praised InSAR data for its ability to operate in all weather conditions. This study used the C-band dual polarized (VV and VH) Sentinel-1 InSAR data to monitor for breach locations along levees in the Solengi River in Assam. This study also intended to use the InSAR data to characterize breach length and the area affected due to inundation. The study found that the InSAR had high sensitivity to the dielectric constant which is important when monitoring water-related failures such as breaching. This study also showed that VH polarized InSAR data was more suitable for monitoring breaching as areas with high soil moisture content were more prominent than when using the VV polarized InSAR data. The last conclusion of this study was that InSAR data was not suitable for breach monitoring during high water as it could not effectively monitor seepage within the levee.

While InSAR has proven useful for levee monitoring, a study by Maltese et al. (2021) used InSAR data from the Sentinel-1 satellite and thermal imagery from the Landsat 8 satellite. The objective of this study was to monitor displacements of the Magazzolo Reservoir behind the Castello Dam in Sicily, Italy. One of the major disadvantages of the satellite imagery data was that a clear sky was needed for useful images. Also, the satellite imagery had a ground resolution of 14 m and the authors felt that was not high enough for the application. In comparison, the InSAR data had high enough accuracy that when passed through a K-Means clustering algorithm, the water surface extent behind the dam was successfully estimated.

Dardanelli et al. (2014) also tried to monitor the reservoir water surface of an earthen dam using satellite imagery. This study used imagery from the Landsat 7 satellite in conjunction with permanent GNSS stations. In contrast to the results from Maltese et al. (2021), this study found that applying the K-Means clustering algorithms to the satellite imagery was sufficient to model the reservoir area with good accuracy with the GNSS data validating the results.

A major disadvantage of using satellite data for surface inspection is the inability to control when data is collected. This can be problematic if inspection needs to occur immediately after a natural hazard to ensure the health of the system and the satellite is not due to revisit the location immediately. A solution to this problem is to use UAV-based data instead, as shown by the study conducted by Bakula et al. (2016). This study utilized a UAV Hawk Moth platform equipped with LIDAR and a digital camera to create Digital Elevation Models (DEM), orthoimages, and point clouds. The benefits of these methods as described by the authors include the continuity in data, non-invasive, mobile, the ability to collect data in areas with access issues, and the penetration of vegetation cover.

The work carried out by Yu et al. (2018) also showed the benefits of using UAV-enabled sensors. This study used a Phantom 4 Pro drone equipped with an MP CMOS Sony camera to monitor the construction of a levee on the Xiangxi River. The authors had 30 minutes of flight time with this drone and payload. The authors highlighted how the use of UAVs improves monitoring efficiency as it shortens inspection time and how time intervals between UAV flights are adjustable, unlike with satellite revisit times.

In contrast to satellite data, UAV data tends to have higher resolution and accuracy. Wang, Wang, and Zhang (2020) used UAV-enabled LIDAR for their study because it gave high-resolution data, which was important for their objective of obtaining geometric parameters of levees to quickly assess stability from overtopping threats. The authors needed high-resolution data to obtain crest height, crest width, and the slopes of a levee in the Pearl River Delta in Southern China. A point cloud was generated from the LIDAR acquisition which was processed into a DEM to get the geometric parameters. Because LIDAR data was high resolution, large-scale examinations of the levee were possible and allowed for the automatic detection of the levee's geometry for stability assessment.

A study by Choung (2014) also used UAV-enabled LIDAR and multispectral images due to the high resolution. The multispectral camera contained the red, blue, green, and near-infrared (NIR) bands at a resolution of 25 cm. For levee applications, high-resolution imagery is needed to detect certain surface features such as small soil patches or eroded areas that could become problematic as they grow. As a result, early detection of these areas is important, so high-resolution data was needed. Because of the high-resolution multispectral data, surface features such as asphalt roads, gravel roads, soil roads or soil patches, concrete blocks, vegetation, and eroded areas were detectable using K-Means and ISODATA clustering techniques.

With all the benefits of UAV data, there are some disadvantages as well, primarily drone battery power as experienced by Zhao et al. (2021). This research was to develop a dam health monitoring model using a DJI M210 drone equipped with a Zenmuse X5S camera. The authors found that the battery time limited their monitoring capabilities. UAV battery time is dependent on many factors including the size of the payload, wind speed, as well as flight speed, and many others. Unlike satellite data, ground control points (GCPs) are required to accurately georeference models created from UAV data. This study placed GCPs on the downstream surface of the dam, dam crest, and edge of the dam to understand where GCPs were most effective. The authors discovered that using GCPs only on the dam crest did not provide a precise model due to the flat terrain. The GCPs located on the slopes of the earthen dam provided a more precise model. When performing any UAV flight, it is important to spread out the GCPs, and taking care not to place them linearly.

Another UAV option for monitoring levees is through NASA's Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR) system. This is an airplane-mounted InSAR system that can be deployed multiple times in the same area for change detection analyses. The sensor uses L-band InSAR which results in sub-meter ground resolution and the ability to penetrate vegetation, making it a popular option for levee applications. The work performed by Aanstoos et al. (2010) used UAVSAR data in combination with InSAR data from the TerraSAR-X satellite to create a methodology to automatically detect slides on levees. The authors leveraged Stepwise Linear

Discriminant Analysis for dimensionality reduction and Gray Level Co-occurrence Matrix, and K-Means clustering to create a binary classification of slide versus healthy levee. Training data included locations of existing sand boils, slough slides, locations of historically frequent slides or boils, and in situ measurements of soil properties. This study was successful in splitting the levee segment into slide locations and healthy locations. The healthy locations were characteristic of uniform soil properties and vegetation while the slide locations produced anomalous soil properties.

This work was expanded on by Dabbiru, Aanstoos, and Younan (2016), who also used UAVSAR on the levees along the Mississippi River to develop a machine learning framework to detect slump slides. Unlike the Aanstoos et al. (2010) study, this study leveraged RX Anomaly Detector as it does not require training data and looks for spectrally anomalous data. This study was effective at using anomaly detection to classify slump slides through surface roughness determined from the InSAR data. The anomalous locations along the levee segments were where known current slides, repaired slides, and areas were slides developed after the study. This type of analysis is useful for stability assessment as it can detect current and future stability issues.

UAVSAR data was also used by Nobrega et al. (2013) on the Mississippi River levee system to investigate multiple failure mechanisms. The goal of the study was to use InSAR data to detect soil moisture, soil electrical conductivity, sand boils, landslides, and vegetation. While most other research focuses on detecting one failure mechanism from a dataset, this research aimed at detecting multiple mechanisms from the same dataset. Backpropagation Neural Networks were used to classify soil conductivity to estimate soil moisture of a levee segment by using radar backscattering coefficients as inputs since these coefficients are a function of moisture, surface roughness, and radar incidence angles. Supervised machine learning techniques such as Support Vector Machines (SVM) were applied to the data to detect landslides using a binary classification and performed with an accuracy of 99.6% while unsupervised analyses using RX Anomaly Detection were applied to the dataset to detect sand boils on a clayey levee. Object-based Image Analysis and Principal Component Analysis successfully detected high moisture areas on the toe of the levee. This study shows the feasibility of using different techniques on the same dataset for detecting multiple failure mechanisms.

3.1.3 Imaging the Surface and Subsurface

Since levees are geologic structures, failure mechanisms can exist in the subsurface or surface of the structure. As a result, understanding the condition of the surface and subsurface is an integral component for understanding the health of the system. Research conducted by Zumr et al. (2020) used geophysics and remote sensing to detect internal structures in an earthen dam. The sensors to image the surface included a DJI Phantom 4 Pro equipped with an FC330 RBG camera and a Flir Tau2 radiometric infrared camera and a Tarot X hex-copter equipped with an Optris Micro-Epsilon TIM 450 infrared camera. The geophysical method used in this study was ERT using a Wenner-Schlumberger array. All structures identified in this dam were water related. The ERT data was able to detect traces of the original stream location beneath the dam while the remote sensing platforms could detect saturated zones on the dam structure by using a combination of optical

imagery and thermal imagery. While no other structures were identified with these datasets, this study did show the effectiveness of using the data in combination.

Another study that used geophysical sensors and remote sensing sensors in combination was by Borgatti et al (2017). This study used GPR, ERT (dipole-dipole array), EMI, and thermal imagery to try and detect European Badger burrows on sandy and silty levees along the Panaro River in Northern Italy. These sensors were chosen as cavities created by animals can have significantly different ρ and thermal temperatures compared to the surrounding soil. To detect these structures with thermal cameras, photos were captured throughout the day as it made changes in air temperature easier to capture. The thermal images did detect the burrows and showed that the cavities had lower temperatures during the day with respect to the surrounding soil. During the night, the burrows showed warmer temperatures than the surrounding soil. The geophysical sensors were also effective at locating the burrows. GPR was found to be useful when detecting the top of the burrows while ERT and EMI were useful for estimating the volume of the burrow.

While detecting one structure is easy, being able to detect multiple failure mechanisms is important as levees are complex systems. Mooney et al. (2014) developed a study using remote sensing and geophysics to detect sand boils and crest depressions. The study used an experimental levee in Booneschans, Netherlands which was forced to fail through internal erosion. The levee was made from a highly plastic clay and sat on a poorly graded sand foundation. The sensors used included self-potential, passive seismic methods, and terrestrial LIDAR. This study showed that the geophysical methods could detect sand boils on the toe while the LIDAR measured 15 cm of deformation on the crest, highlighting the strengths of surface and subsurface imaging techniques.

One of the earliest works that combined surface and subsurface imaging was conducted by Bishop, Dunbar, and Peyman-Dove (2003). In this study, the authors used airborne LIDAR and EMI to classify levee segments into poor condition and good condition. LIDAR was used to define levee geometry, vegetation encroachment, and erosion, while EMI was used to obtain subsurface ρ. In conjunction with this data, previous data on construction and performance history, visual inspection of levee surface, material type, presence of man-made intrusions, geology, and proximity to burrow areas were also considered. These components were scored individually, where the score represented how that component negatively or positively affected the health of the levee. The scale of scores was then redefined to range from poor condition to good condition. This framework is a first step in integrating the results from remote sensing techniques and geophysical techniques for understanding the health of the overall system but fails to include information about how levee health changes with time. Understanding how conditions are progressively changing will allow for the early detection of failure mechanisms.

3.2 Field Testing Locations

The field data was collected on levees from two islands in the Sacramento-San Joaquin Delta as shown in Figure 3.1. The Delta was historically a freshwater marsh that was reclaimed in the late 1800s for agricultural purposes (Ingebrisen, Ikehara, Galloway, & Jones, 2000). To keep the

islands in the Delta from flooding, farmers used dredged materials from the waterways to build 1,300 miles of levees.

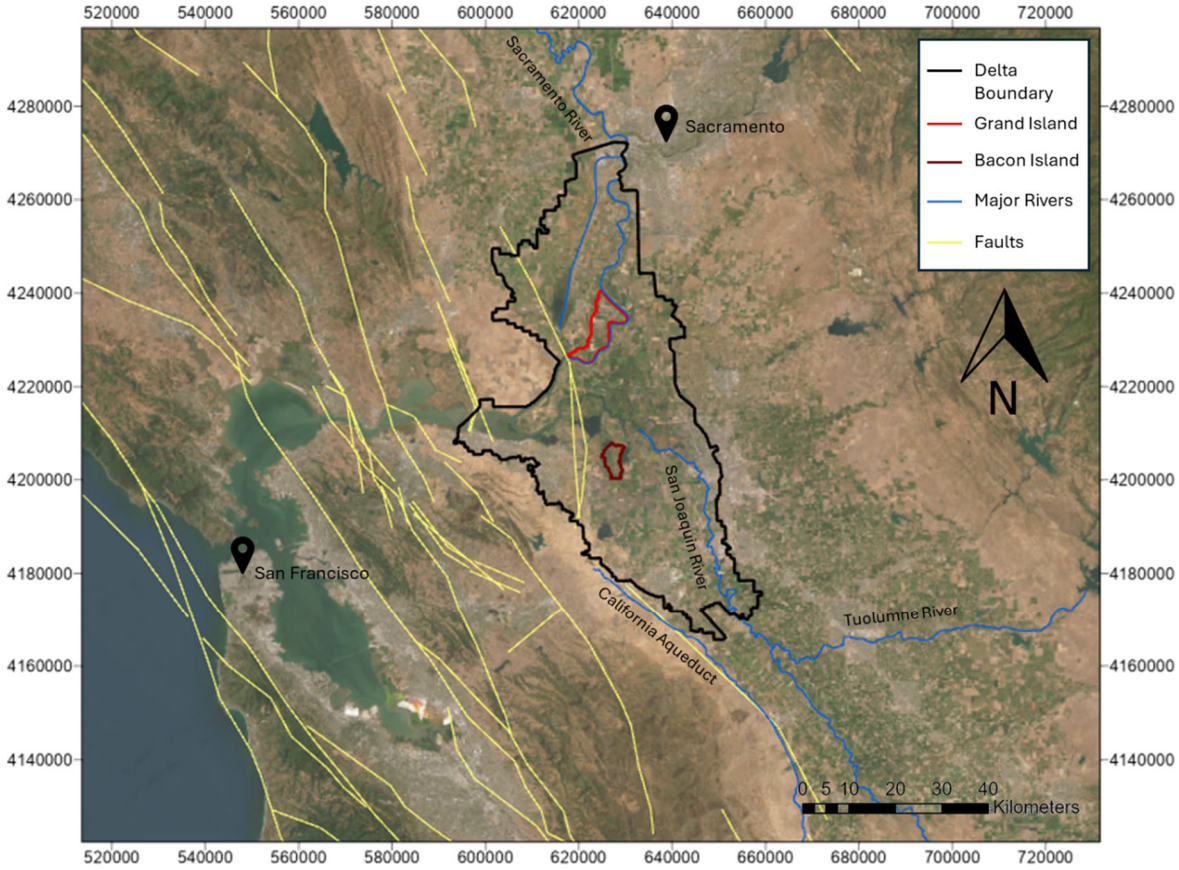


Figure 3.1 Location of the Sacramento-San Joaquin Delta

The geology of the Sacramento-San Joaquin Delta is composed of sedimentary deposits that range from 5 to 10 km thick (Atwater, 1982). Most of these sediments were formed in marine environments roughly 175 million to 25 million years ago. Alluvial fan deposits are found in the Delta along the tributaries due to historic glaciers in the Sierra Nevada Mountain Range that were washed down in meltwater streams. Roughly 7,000 years ago during an interglacial period, the sea levels rose and invaded the Central Valley in California allowing for the deposition of soft peat.

After the conversion to agricultural land, the peat in the Delta started undergoing significant decomposition and oxidation due to exposure to air causing land subsidence (Kraus & Sneed, 2018). It is estimated that the land in the Delta subsides approximately 3 to 8 centimeters per year, causing most of the islands to be upwards of 6 meters below sea level (Ingebrisen, Ikehara, Galloway, & Jones, 2000).

With islands below sea level, the levees in the Sacramento-San Joaquin Delta are vital flood protection structures as they are continuously holding back water to keep the islands from flooding. It is imperative that these structures do not fail to protect agriculture and fresh drinking water.

In addition to other threats, the levees in the Sacramento-San Joaquin Delta are also seismically vulnerable. The Delta Risk Management Study (DRMS) Phase 1 from 2009 explored the impact of seismic hazards due to the San Francisco Bay Area's high seismicity on the levee system in the Sacramento-San Joaquin Delta. The study showed that given a 6.7 Mw earthquake on the Hayward, Calaveras, or San Andreas Fault in the Bay Area, which has a 62% probability of occurring between 2003 and 2032, there is a 40% probability of 27 or more islands in the Delta flooding simultaneously (California Department of Water Resources, 2009). According to the study, if 20 of the islands were to flood due to a seismic event, the export of fresh drinking water to two-thirds of Californians would be interrupted for 1.5 years with water supply losses of 8 million acre-feet. It would take 3 years and upwards of 2.3 billion dollars to repair the levee system and all that was destroyed due to the failures.

Due to the fact that the Sacramento-San Joaquin Delta is connected to the Pacific Ocean and is subjected to changes in tides, climate impact on sea level rise is also a concern to these levees. High water is a concern to levee health as it leads to potential failure mechanisms such as overtopping and seepage. The DRMS expects 140 levee failures in the Sacramento-San Joaquin Delta in the next 100 years because of increasing high water risks (California Department of Water Resources, 2009).

Due to the age and deterioration of the levees in the Sacramento-San Joaquin Delta and the important role the levees play in flood protection and fresh water supply for the state, as well as the seismic and climate impacts on these systems, the Sacramento-San Joaquin Delta was chosen as the primary location for field deployments. Field deployments in the Sacramento-San Joaquin Delta occurred at two islands, Grand Island and Bacon Island.

3.2.1 Grand Island, California

The first island chosen for field deployment was Grand Island which is located approximately 50 kilometers south of Sacramento. The island is boarded by two rivers, Steamboat Slough to the west and the Sacramento River to the north, south, and east. This island was chosen because there are approximately 47 kilometers of continuous levees surrounding the island which are being continuously loaded by Steamboat Slough and the Sacramento River (USACE & FEMA, 2024). The Grand Island levees protect 67 square kilometers of land, a population of 714 people, 313 buildings, and a property value upwards of $74.5 million (USACE & FEMA, 2024). If the levees were to fail, two state highways would be inaccessible including Highway 160 which overlays the levee along the Sacramento River, and Highway 220 which crosses through the center of the island connecting Highway 160 with the J-Mack Ferry, providing access to Ryer Island.

The levees on Grand Island are operated and maintained by Reclamation District 0003 with MBK Engineers as the district engineers. The failure modes of concern for these levees include overtopping, embankment erosion, embankment seepage, and foundation seepage (Stillwater Sciences and MBK Engineers, 2020). Being approximately 105 kilometers from the San Francisco Bay Area, there is an additional seismic risk associated with these levees. If the levees were to breach, the water depth on Grand Island could be upwards of 8 meters above sea level (USACE & FEMA, 2024).

Three site locations were chosen along the Grand Island Levee system for field deployment, as shown in Figure 3.2.

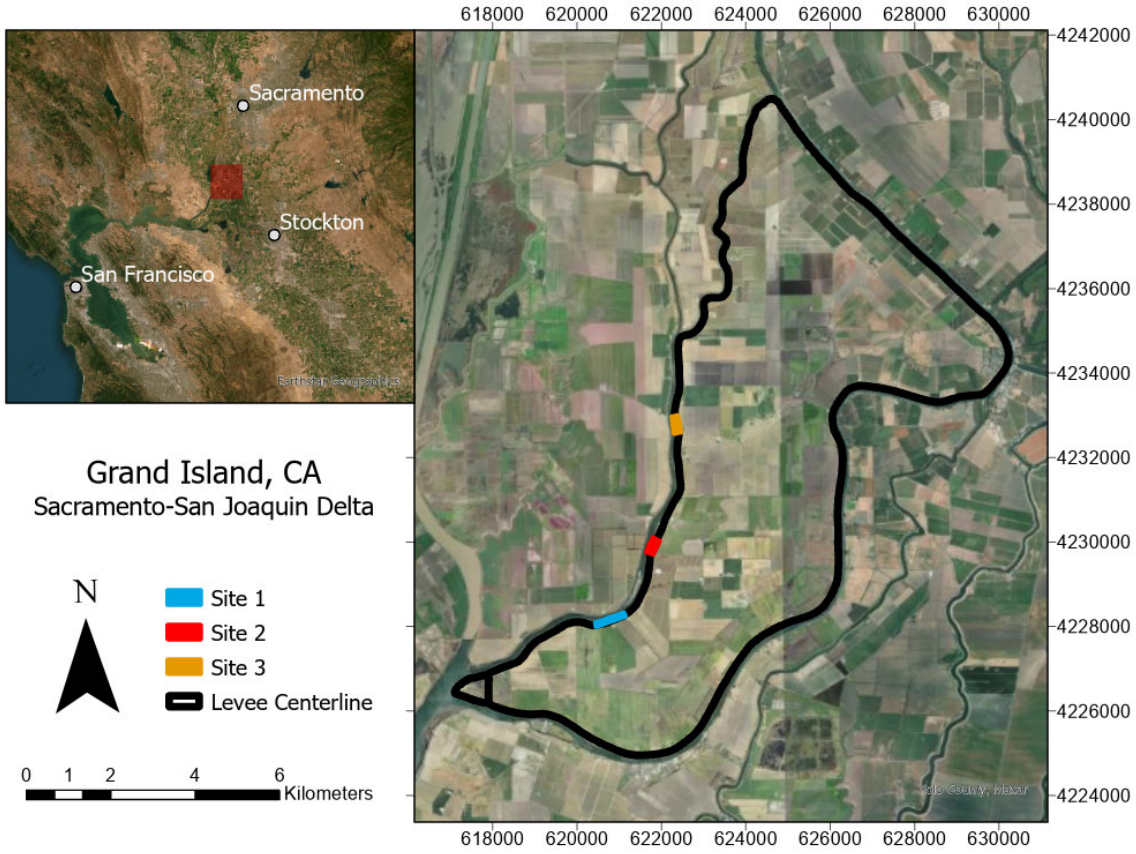


Figure 3.2 Data collection sites along the Grand Island levees

These sites were chosen in collaboration with MBK Engineers, the engineering firm in charge of maintaining these levees. These sites were chosen because they have existing seepage problems and MBK Engineers had planned detailed geotechnical subsurface explorations using Standard Penetration Test (SPT) and Cone Penetration Test (CPT) along the centerline of the levee for three segments of the levee. All the sites are located on the western side of the island, bordered by the Steamboat Slough waterway on the wet side. The dry side of the levee at each of these sites is bordered by private agricultural land.

Site 1 is the southernmost site and is the longest of the three sites located between stations 460+00 and 474+00 adjacent to Steamboat Slough. The crest elevation at this site is approximately 5 m from the National Geodetic Vertical Datum (NGVD29) with crest widths that vary between 8 to 9 m along the length of the site. The slopes of the levees also vary. The dry slope is the most gradual at 4H:1V and the steepest at 5H:1V while the wet slope tends to have shallower slopes ranging from 2H:1V to 4H:1V (Hultgren-Tillis Engineers, 2021).

Site 2 is located between stations 394+00 to 406+00 adjacent to Steamboat Slough on the west side of Grand Island. The crest elevation ranges from 6 to 7 m using the NGVD29 at this site with crest widths that range from 7.5 to 10.5 m. The dry slopes at this site are steeper than the wet slopes

with inclines ranging from 4H:1V to 4.5H:1V. In comparison, the wet slopes range from 2H:1V to 4H:1V along this site (Hultgren-Tillis Engineers, 2021).

Site 3 is the northernmost site located on Grand Island. It is located between stations 292+00 and 298+00, making it the shortest of the three sites. The crest elevation for this site ranges between 6 m and 7 m using the NGVD29 while the width of the crest ranges from 7.5 m to 9 m. Like with Site 1 and Site 2, the dry slope tends to be steeper than the wet slope with inclines ranging from 3.5H:1V to 4H:1V while the wet slopes incline between 2H:1V to 4H:1V (Hultgren-Tillis Engineers, 2021).

Existing Geotechnical Information for Grand Island

There have been two previous geotechnical explorations of the levees on Grand Island near the sites of interest. The first geotechnical exploration was performed by Harlan Tait Associates for the USACE in 1990 and can be found in Appendix B (Harlan Tait Associates, 1991). Nine borings were drilled on the levees adjacent to Steamboat Slough along the crest and the dry toe and were used to help understand the subsurface conditions at these sites to design the field surveys.

Four of the nine boreholes were drilled along the levee crest. These boreholes were drilled to a depth that ranged between 9 m below ground level to upwards of 21 m below ground level. The logs show that the top 3 to 6 m of the levee is sand with a fines content less than 12%. Below this sand layer is typically a silt layer, underlain by a clay layer for the southern boreholes. The northern boreholes show another sand layer below the silt followed by either an organic clay layer or another silt layer.

The remaining five boreholes from the 1991 geotechnical exploration were drilled on the dry toe. These boreholes extend approximately 7.5 to 10.5 m below ground level. The northernmost borehole shows a thin silt layer followed by almost 9 m of sand underlain by an organic clay layer until the end of the borehole. The southernmost borehole shows sand near the surface underlain by a 6 m thick clay layer. The three boreholes in the middle of the exploration area show alternating layers of silts and clays underlain by thick organic clay layers. One of these boreholes shows thin beds of peat which is not observed in any of the other boreholes.

The second geotechnical exploration was done in 2020 and 2021 by Hultgren-Tillis Engineers for MBK Engineers and can be found in Appendix B (Hultgren-Tillis Engineers, 2021). In December 2020, 11 test pits were excavated ranging from 1.5 to 3 m below the ground surface. A rubber-tired backhoe was used for the excavation.

In April 2021, 5 boreholes were drilled by Hultgren-Tillis Engineers along the dry toe at each of the three sites. Drilling was done by rotary wash and sampling was done using 2.5 inch outside diameter split barrel samplers and Shelby tube samplers. A 140-pound hammer was used for sampling and SPT-N values were recorded, correcting only for hammer energy.

The borehole and test pit excavations showed that Site 1 was composed of very loose to loose sand that was 3 to 4 m in depth. The fines content of this sand was determined to be 4 to 9%. Below the sand was a soft to firm silt layer with a thickness that ranged from 8.2 to 8.8 m. The Liquid Limit (LL) was measured to be between 88 to 90 and the Plasticity Index (PI) was measured to be

between 33 to 35. A soft to very stiff clay layer was found beneath the silt with a thickness of 2.4 m which overlays a medium dense to dense sand layer.

The boreholes and test pits at Site 2 showed a soft sandy silt layer near the surface that ranged in thickness from 2.4 to 2.6 m and with a fines content of 57%. This sandy silt exhibited plasticity behaviors with a measured LL of 45 and PI of 13. Below this layer was a medium-dense sand layer that was 5.8 to 5.9 m thick with a fines content of 5 to 6%. This overlayed a soft to very stiff elastic silt that was 6 to 6.4 m thick. The LL was measured to be 56 while the PI was 20. Below this layer until the end of the borehole was a medium-dense sand.

At Site 3, the geotechnical exploration showed a very soft to stiff elastic silt near the surface that was 11 thick. The LL was measured to be 65 with a PI of 23. At 8.5 m below ground level in the elastic silt layer, a five-foot-thick layer of organic silt was observed. Below the elastic silt was a very stiff to hard lean clay layer that was 2.7 m thick. Underlaying that layer until the end of the borehole was a medium-dense silty sand.

3.2.2 Bacon Island, California

An additional dataset was provided by researchers at Lawrence Berkeley Laboratory for Bacon Island in the Sacramento-San Joaquin Delta. This is a privately owned island with approximately 23 km of continuous levee surrounding the island and continuously holding back water. The data being used for this analysis was collected at three locations along the toe and crest of the levee surrounding the island, with two sites on the north end of the island and the third site at the south end of the island, as shown in Figure 3.3.

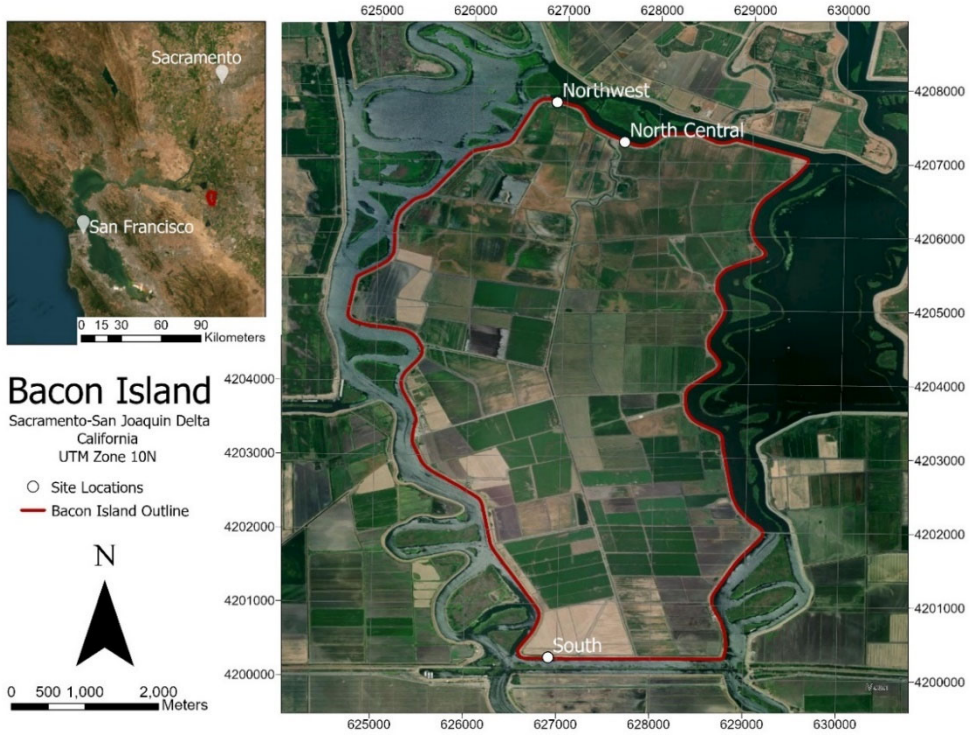


Figure 3.3 Locations of field testing sites on Bacon Island

Unlike Grand Island, peat can be found almost everywhere in the subsurface at Bacon Island. The peat thickness varies drastically, being anywhere from 2 m to 15 m in thickness.

Existing Geotechnical Information for Bacon Island

Historic borehole and CPT data exist for the three sites chosen for analysis at Bacon Island. The Northwest Site has one CPT found on the crest of the levee which shows that within the top meter, the soil is stiff sand and silty sand. The subsurface then starts to transition into clay and silty clay until approximately 4 m depth. Between 4 m and 10.5 m, the subsurface is mostly organic soil with some clay layers. Underlying the peat is more silty clay and silty sand with sand starting at 12 m until the end of the CPT at 14.5 m.

The Northcentral site has the most boreholes and CPT information with two CPTs and one borehole on the crest of the levee and one CPT on the toe of the levee. The toe of the levee consisted of a very stiff fine-grained material within the top 0.6 m of the surface. Underlying this top material was clay until the end of the CPT at a depth of 8.2 m. The crest information shows stiff sands, silty sands, and silty clays for the top 2 m below the levee crest. Underlying this is alternating layers of clay and organic soil for 7 m. From 11 m in depth to 18 m in depth is sand followed by 2 m of clay and an additional 11 m of sand. From 31 m until the end of the CPT at 38 m is clay.

The CPT data was processed using the CPeT-IT software following the Roberston (1990) procedure to obtain the normalized soil behavior type (SBTn) for the supervised machine learning analysis, discussed in Chapter 4. The first step in processing the data was to digitize the CPT records. The original CPT records can be found in Appendix B.

The SBTn was used to remove the effect that overburden stress has on soil classification when using CPT (Robertson P. K., 1990). This required normalizing the cone tip resistance (Q_t) and Friction Ratio (FR), described in Equation 3.1 and 3.2, respectively.

$$Qt = \frac{(q_t - \sigma_{vo})}{\sigma'_{vo}} \tag{3.1}$$

Where q_t is the total cone resistance in MPa, σ_{vo} is the total vertical stress in kPa, and σ'_{vo} is the effective vertical stress in kPa.

$$FR = \frac{fs}{q_t - \sigma_{vo}} \tag{3.2.}$$

Where fs is the sleeve friction in MPa.

The Q_t and FR are then used to calculate the soil behavior type index (I_c) which can be used to determine the soil behavior type without plotting on the SBTn chart.

$$I_c = [(3.47 - \log Q_t)^2 + (\log FR + 1.22)^2]^{0.5} \tag{3.3}$$

Using the normalized method, Q_t must be upgraded to Q_{tn} by using a stress exponent, n, which is a function of I_c, resulting in iterations needing to be performed to converge on I_c and Q_{tn}. Q_{tn} can be upgraded through Equation 3.4 using Equation 3.5 to calculate n based upon I_c.

$$Q_{tn} = \frac{q_t - \sigma_v}{p_a}\left(\frac{p_a}{\sigma'_{vo}}\right)^n \tag{3.4}$$

$$n = 0.381(I_c) + 0.05\left(\frac{\sigma'_{vo}}{p_a}\right) - 0.15 \quad (3.5)$$

where n must be less than or equal to one and p_a is the atmospheric pressure.

The SBTn was assigned to each CPT data point based upon the calculated I_c value using Table 3.1.

Table 3.1 Soil Behavior Index (Robertson & Cabal, 2014)

SBTn Zone	I_c Range	Soil Behavior Type
1	-	Sensitive, fine grained
2	> 3.6	Organic soils, clay
3	2.95 – 3.60	Silty clay to clay
4	2.60 – 2.95	Clayey silt to silty clay
5	2.05 – 2.60	Silty sand to sandy silt
6	1.31 – 2.05	Clean sand to silty sand
7	< 1.31	Gravelly sand to dense sand
8	-	Very stiff sand to clayey sand
9	-	Very stiff, fine grained

3.3 Geophysical Data Collection

3.3.1 Multichannel Analysis of Surface Waves

MASW was performed at Grand Island as well as Bacon Island to collect *Vs* of the subsurface of levees. At Grand Island, there were three MASW survey profiles at Site 1, three MASW survey profiles at Site 2, and two MASW survey profiles at Site 3. Data was collected along the same profiles during both the dry and wet seasons. For Bacon Island, continuous MASW surveys were performed along the levees protecting the north end of the island as well as the south end of the island. All geophysical profiles can be found in Appendix C.

3.3.1.1 Seismic Waves Background

There are two broad categories of seismic waves. The first are body waves which can travel through a medium with depth. These waves can be divided into two types. The first are primary waves, or p-waves. P-waves exhibit compressional movement with volumetric and shear deformations. The second type of body wave are secondary waves, or s-waves. An s-wave exhibits shearing deformation but no volumetric deformation.

In comparison to body waves, there are also surface waves. Surface waves are confined to the surface, unlike body waves which can travel through the entire medium. Surface waves are produced when p-waves and s-waves interact at the surface where the shear stresses vanish (Lay & Wallace, 1995). There are two major types of surface waves. The first are Rayleigh Waves which exhibit a retrograde elliptical particle motion and are confined to propagate along the surface. The second type of surface waves are Love Waves which are confined to the surface by a layer below and reverberate in a transverse motion to the direction of wave propagation.

Both body and surface waves are integral for geophysical explorations of the subsurface. Many types of seismic geophysical methods require an impact from an active source to generate seismic waves. These impacts are approximately 67% surface waves, 26% s-waves, and 7% p-waves (Miller & Pursey, 1955). After the impact, the waves spread away from the source. Body waves spread from the source in a spherical fashion and the decay of the amplitude of these waves is $1/r^2$ where r is the distance from the source. In contrast, surface wave amplitude falls off at 1/r and decays rapidly with depth. Due to these relationships, surface waves have greater energy near the surface compared to body waves (Lay & Wallace, 1995).

This becomes important when deciding which geophysical method is necessary for the outcome of the investigation.

3.3.1.2 Surface Wave Geophysical Methods

Surface wave seismic geophysical methods, such as MASW and Spectral Analysis of Surface Waves (SASW), which utilize surface waves, are a better option for geotechnical explorations as surface waves have lower frequencies, on the order of 1 to 30 Hz, which is needed for near-surface explorations (Park, Miller, Xia, & Ivanov, 2007).

SASW looks at the time delay between receivers at various frequencies, using two receivers, an impulse source, and a spectral analyzer. MASW uses multiple geophones, typically 12, 24, or 48 to record the ground movement from seismic waves. MASW measures the Rayleigh Waves from an impact source which are dispersive in a heterogeneous medium, such as a naturally occurring subsurface (Everett, 2013). This dispersive nature of the wave allows for the wave to be decomposed into its fundamental frequencies through the Fourier Transform. At each fundamental frequency, there is a characteristic propagation velocity called the phase velocity which can be used to construct a dispersion curve which can then be used to construct a *Vs* profile of the subsurface (Park, Miller, & Xia, 1999).

The general procedure for conducting an MASW survey outlined by Park, Miller, and Xia (1999) is broken into three steps. The first is field data collection to acquire multichannel records. The second step is data processing to extract dispersion curves. And the final step is back calculating to create a 1D *Vs* profile. If a 2D *Vs* profile is the desired outcome such in this study, a subsequent step to the final step is to interpolate between the 1D *Vs* profiles.

3.3.1.3 Comparison of MASW with Other Geophysical Methods

Body wave geophysical methods include Seismic Reflection and Seismic Refraction. Seismic Reflection utilizes subcritical reflections of p-waves at geologic interfaces in the subsurface to differentiate between different lithological units (Burger, Sheehan, & Jones, 2006). Seismic Refraction, on the other hand, looks at the first arrival of p-waves after they refract from surfaces in the subsurface to map structures, stratigraphy, and other geologic structures of interest (Burger, Sheehan, & Jones, 2006). While these methods are popular for hydrocarbon exploration because the wavelets have higher frequencies (typically higher than 50 Hz), they are more useful for deeper investigations instead of for shallower investigations like surface wave geophysical methods (Park, Miller, Xia, & Ivanov, 2007).

For levee health assessment, it was decided that MASW was more appropriate than SASW as it is less time-consuming. Because the frequency is dependent on receiver spacing for SASW, the receivers must be reconfigured many times to sample the desired frequency range which requires a new shot with each reconfiguration (Everett, 2013). In comparison, MASW uses a linear array of geophones with constant spacing, so the array layout does not need to be changed each time, allowing for a quicker and more economical survey which is important when trying to characterize long stretches of ground, such as with levees. MASW also has a better Signal-to-Noise Ratio (S/N) which makes it preferable to SASW. With SASW, the noise can change unpredictably due to shot and receiver locations while with MASW, noise sources can be identified among different channels, effectively increasing the S/N. With a higher S/N, there will be higher accuracy in the dispersion curves which will result in a more accurate *Vs* profile (Everett, 2013).

Other geophysical methods that are used in geotechnical engineering for characterizing the *Vs* in the subsurface include down-hole and cross-hole seismic surveys. MASW is preferred for levee health assessment in comparison to these methods because it is non-destructive and non-invasive. For down-hole and cross-hole surveys, drilling through the subsurface is required.

3.3.1.4 MASW Benefits in Geotechnical Engineering

The dispersion curve is critical for creating a *Vs* profile of the subsurface when using MASW. It is created by plotting the phase velocity as a function of frequency. This phase velocity is a function of the elastic properties of the subsurface, which allows for the inference of near-surface elastic properties which are important for understanding the stiffness of the subsurface (Park, Miller, & Xia, 1998). The relationship between the shear modulus (α) and *Vs* is well known, as shown in Equation 3.6 (Park C. B., 2013):

$$\alpha = d_m V_s^2 \tag{3.6}$$

Where α is measured in kPa and d_m is the density of the material (g/m^3). This is important because *Vs* can be used to distinguish between structures with differing elastic mediums such as changes in soil layers, voids, structures, and the soil/bedrock interface.

Because stiffness is important in geotechnical engineering, MASW has many uses for geotechnical analysis such as assessment of bearing capacity and understanding ground behavior under continuous vibrations (Park C. B., 2013). MASW is most often used for ground amplification and liquefaction analysis, though. Both the International Building Code (IBC) and the National Earthquake Hazards Reduction Program (NEHRP) use *Vs* in the upper 30 m (V_{s30}) to classify the site class. As an example, Table 3.2 lists the NEHRP site class, with the corresponding V_{s30}, and the geologic unit description.

Table 3.2 NEHRP Site Classification (Building Seismic Safety Council, 2004)

Site Class	V_{s30}	Description
A	$V_{s30} > 5000$ ft/s	Hard Rock
B	2500 ft/s < V_{s30} < 5000 ft/s	Rock
C	1200 ft/s < V_{s30} < 2500 ft/s	Very dense soil and soft rock
D	600 ft/s < V_{s30} < 1200 ft/s	Stiff soil
E	$V_{s30} < 600$ ft/s	Soil
F	$V_{s30} < 600$ ft/s	Soil requiring site-specific evaluation

3.3.1.5 MASW Field Data Collection for Grand Island

The first step in performing an MASW survey is to collect the field data. Many considerations need to be accounted for when designing an MASW survey for the application desired, including the equipment needed, geophone frequency, geophone layout, and the desired outcome of the data collection.

Equipment

There is a lot of equipment that is needed to perform an MASW survey, which is shown in Figure 3.4 **MASW setup with (a) geophones, (b) hammer, (c) Geode and (d) field laptop**. One of the most important pieces of equipment needed is a seismograph that can convert analog signals from multiple receivers to a digital signal that can be used by a computer. The seismograph used for the surveys conducted in the Sacramento-San Joaquin Delta was the Geode from Geometrics. The Geode is a small and lightweight seismograph that is ideal for carrying up and down levee slopes. It is used for MASW surveys as it has different channels that can do the analog-to-digital conversion for each receiver separately. The Geode used for analysis had 24 channels, which allows for 24 geophones.

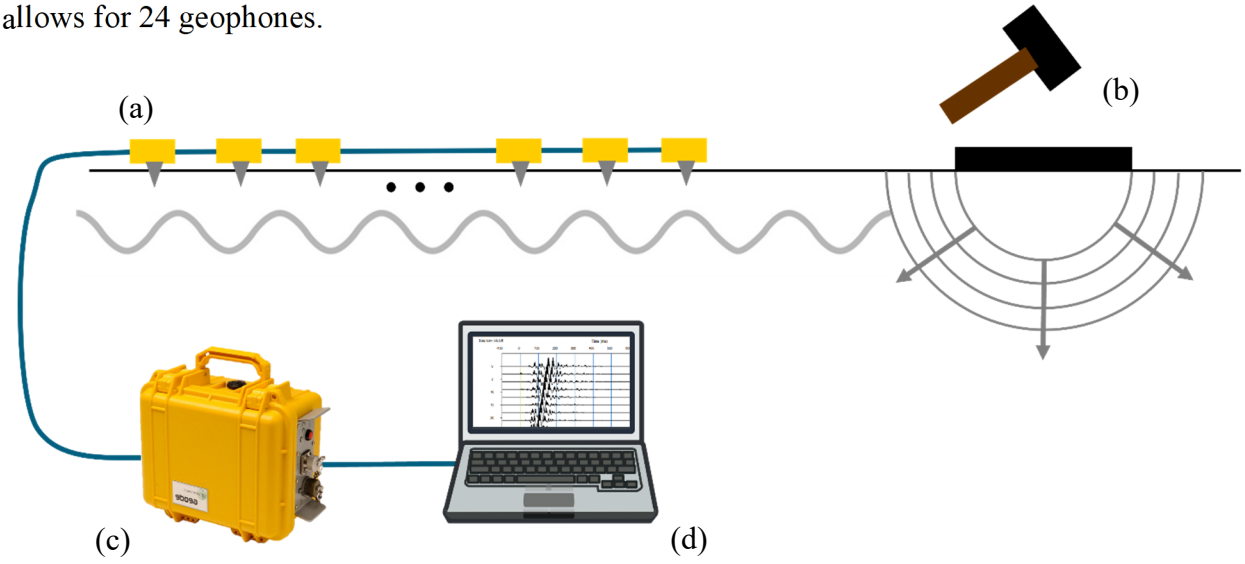


Figure 3.4 MASW setup with (a) geophones, (b) hammer, (c) Geode and (d) field laptop

Geophones are also integral to an MASW survey. They are a small device that sits either on the ground surface or is inserted into the subsurface using a small spike. Their role in the survey is to measure the ground motion of the wave energy. This is done by having a weight attached to a cylindrical wire surrounded by a magnet. As the ground moves, the mass moves proportional to the ground movement causing relative movement between the coil and magnet, generating a voltage. This voltage is the signal measured by the Geode and is proportional to the ground movement (Burger, Sheehan, & Jones, 2006).

The natural frequency of the geophone is also an important characteristic to consider when designing an MASW survey. Typically, a 4.5 hertz (Hz) geophone is used for an MASW survey. The benefit of having low-frequency geophones is that they more accurately record lower frequencies which are needed for deeper investigation depths (Park, Miller, & Xia, Multichannel analysis of surface waves, 1999).

Other equipment necessary for the MASW survey includes a geophone spread cable to connect the geophones to the Geode and a laptop to record the data. A sledgehammer is needed as well to create seismic waves in the ground when it hits a plastic strike plate on the ground. A hammer switch is taped to the hammer and connected to the Geode. The purpose of this switch is to tell the data acquisition program when to start recording. A trigger extension cord may also be needed when performing a 2D MASW survey. It connects to the hammer switch and Geode, to allow the hammer to be carried the entire length of the survey. To power the Geode, a 12-volt (V) battery is needed. And finally, a tape measure is needed to accurately place the geophones and strike plate along the length of the survey.

Geophone Spacing

MASW survey design is critical to achieve the desired outcome of the data collection. One of the many considerations of survey design is the geophone spacing. Geophone spacing is integral for the maximum depth of investigation (Z_{max}). Using Equation 3.7. from Park, Miller, & Xia (1999), Z_{max} can be estimated knowing the length of the array (λ):

$$Z_{max} = 0.5\lambda \tag{3.7}$$

For this study, it was decided that imaging the near-surface is more important than depth as most of the failure mechanisms occur in the top 9 m and it will result in higher resolution near the surface. To be able to image to a depth of 9 m, the total array length needs to be doubled the desired depth, requiring the total array length to be at least 18 m long. With 24 geophones, the spacing between geophones must be at least 0.8 m to achieve a total length of 18 m. To accommodate the strike plate, the geophone spacing was determined to be 0.9 m, giving a total array length of 22 m, meaning the depth of investigation was 11 m. The majority of the MASW surveys on Grand Island were conducted using a 0.9 m geophone spacing. One survey on the crest of the levee at Site 1 was conducted using a 1.5 m spacing instead of the 0.9 m spacing in the same location for comparison.

Active Survey vs. Passive Survey

Another design consideration for an MASW survey is whether to do an active or passive survey. The traditional MASW survey used in geotechnical engineering is an active survey. An active

MASW survey utilizes an anthropogenically made impact source to generate shear waves (Park, Miller, Xia, & Ivanov, 2007). Impact sources can include sledgehammers, weights attached to the back of vehicles, and shaker vehicles. For this type of survey, the geophones are placed linearly.

In comparison, a passive survey uses natural or cultural sources such as wind, waves, and traffic to generate surface waves (Park, Miller, Xia, & Ivanov, 2007). For this type of survey, the geophones are typically placed in a circular layout and require at least 24 geophones. This type of survey has become more popular because it allows for greater investigation depth compared to an active MASW survey, but it requires more time to collect data.

The MASW surveys used for this study were all active surveys, as shown in Figure 3.5. The geophones were placed in a linear array and a 10-pound (lb) sledgehammer was used to generate surface waves. An active survey was chosen because a deep investigation was not needed, ruling out the use of a passive survey, and the framework requires a quick and efficient method for collecting subsurface data.

Figure 3.5(a) was taken during the dry season on the crest of the levee at Site 1. This image shows Geode, as well as the linear array of geophones. Figure 3.5(b) was taken during the wet season on the crest of the levee at Site 2. This image depicts the use of the sledgehammer when performing an active MASW survey.

Field Data Procedure

The field procedure for data collection was designed to produce a 2D MASW profile. A 2D MASW survey takes longer than a 1D survey as the shot location must change multiple times. To start, the geophones are placed linearly at the desired spacing to reach the maximum depth desired. Then all the cables were connected, and the data acquisition software was loaded on the computer. The data acquisition software used for this project was Geometrics' SeisImager software. Table 3.3 below lists all the software inputs required for data acquisition.

Table 3.3 MASW Acquisition Software Parameters

Sample Interval	0.125
Record Length	1 second
Delay	-0.1 second
Stack	5
Stack Type	Auto Stack
Gain	Low
Trigger Holdoff	0.2 5

The first shot location for the survey is typically recommended to be between 10% to 20% of the total array length from the first geophone. With an array length of 22 m, it was decided that the first shot location would be 2 m from the first geophone. An example of the survey geometry and shot locations is shown in Figure 3.6.



(a)



(b)

Figure 3.5 Field photos of active MASW surveys for (a) crest survey in the dry season at Site 1 and (b) crest survey in the wet season at Site 2

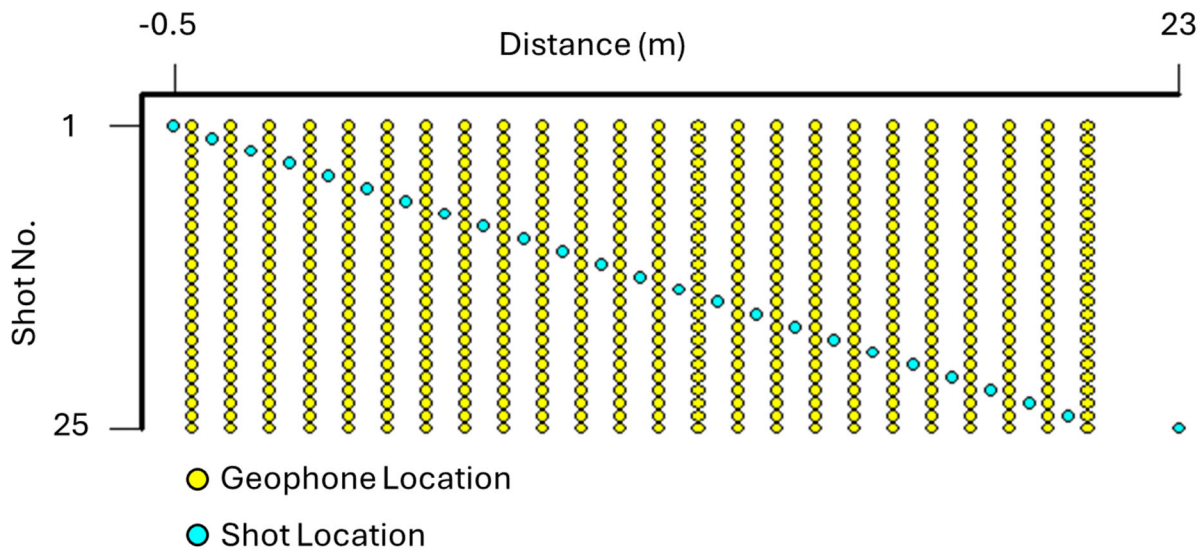


Figure 3.6 MASW geophone and shot array geometry.

The x-axis in the figure dictates the distance along the survey while the y-axis shows the shot numbers which correspond to each file recorded. Each movement of the shot location results in a new file recorded and a new shot number. The yellow dots represent the geophones while the blue dots represent the shot locations.

A fixed-array survey was performed during data collection. A fixed-array survey is where the location of the geophones does not change but the shot location gradually moves down the geophone array. The first shot location is dictated by Shot No. 25 in the figure. After the first shot location was completed, the shot moved to in between the first and second geophone. Then, the shot location moved every 3 ft, resulting in shots being taken in between each set of geophones. The final shot location was 1.5 ft away from the last geophone, as shown in Figure 3.6 by Shot No. 1.

MASW Survey Locations on Grand Island

Each site at Grand Island had at least two MASW surveys, one on the wet toe and one on the crest. For Site 1, three MASW surveys per season were conducted, with one being on the wet toe and two on the crest as shown in Figure 3.7.

Two crest surveys were conducted at the same location at Site 1 but with different geophone spacings to gather more information with depth. In Figure 3.7, the white survey on the crest was performed using a geophone spacing of 0.9 m while the red survey was performed using a geophone spacing of 1.5 m.

Site 2 on Grand Island had three MASW surveys per season. This site had one survey on the crest, one survey on the wet toe, and one survey on the dry toe as shown in Figure 3.8.

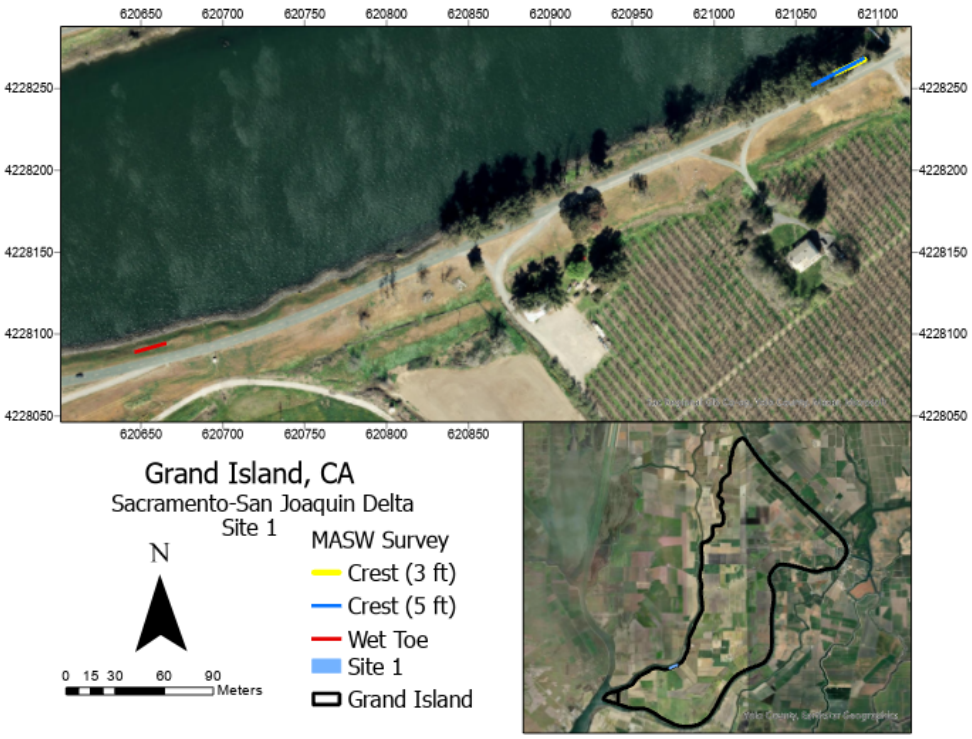


Figure 3.7 MASW survey sites on Grand Island - Site 1

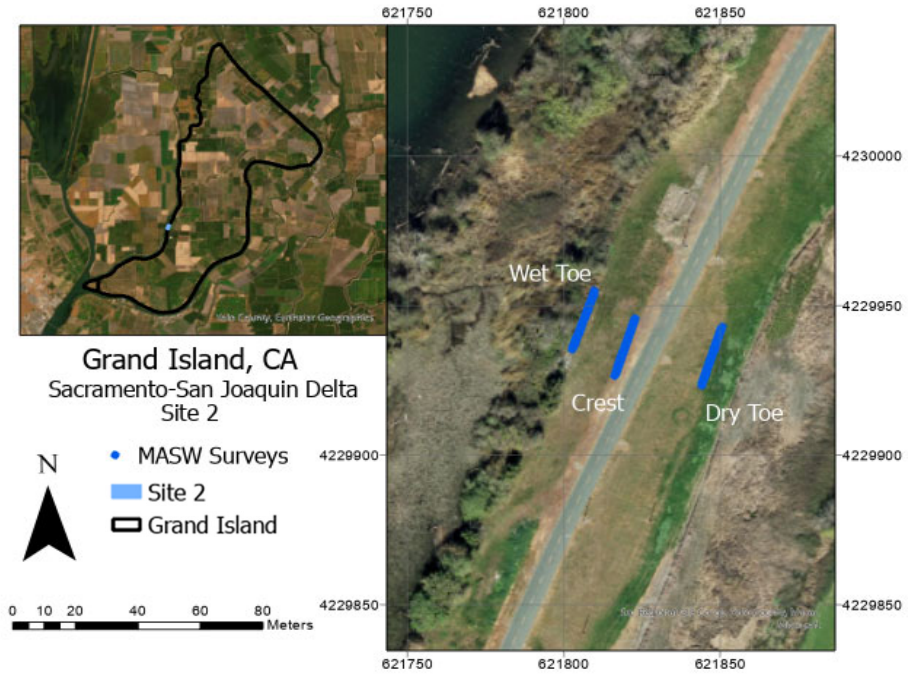


Figure 3.8 MASW survey sites on Grand Island - Site 2

Site 2 is the only site where a MASW survey was performed on the dry toe. By obtaining *Vs* on the dry toe, comparisons can be made with the wet toe and crest to look for differences. Having dry toe data can be used to see if problematic issues such as seepage are more visible in the data than compared to the crest and wet toe data.

Site 3 on Grand Island had two MASW surveys as it was the smallest site and the most vegetated site with the steepest slopes, making many of the slopes difficult to survey on. One survey was performed on the wet toe and one survey was performed on the crest, as shown in Figure 3.9.



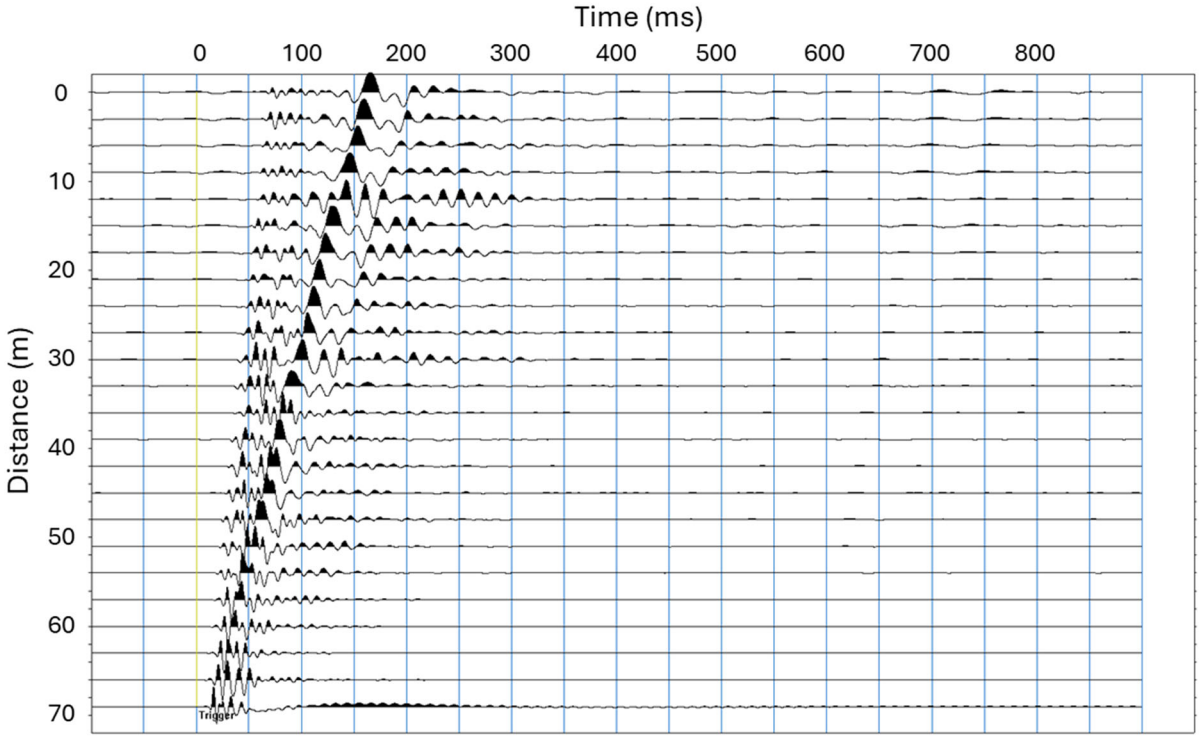
Figure 3.9 MASW survey sites on Grand Island - Site 3

4.2.1.6 MASW Field Data Collection for Bacon Island

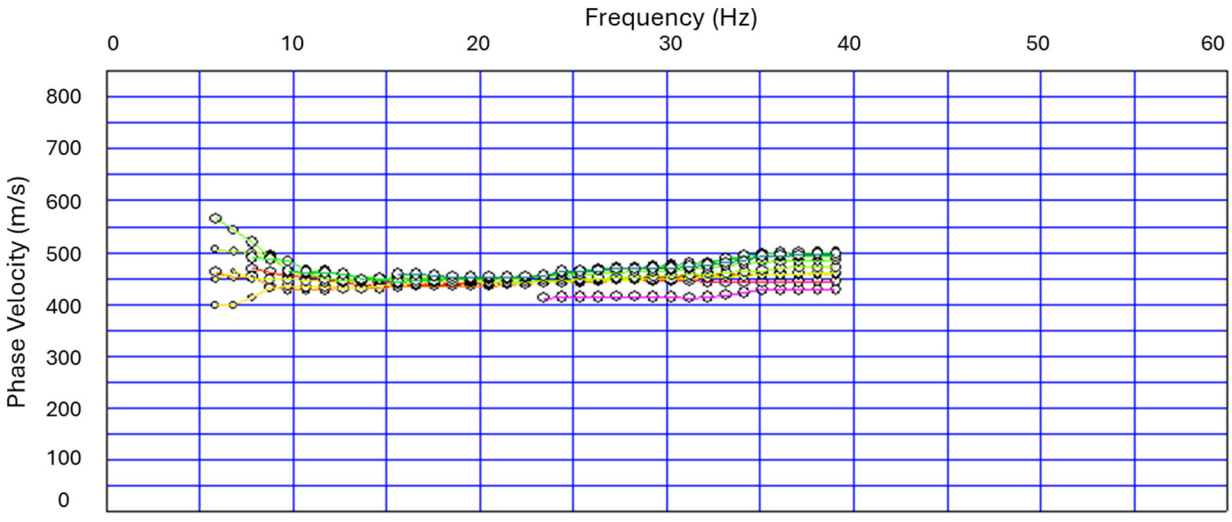
MASW data was collected on Bacon Island by researchers at the Lawrence Berkeley Laboratory. Three MASW surveys on the island are being used, with two surveys on the north end of the island and one survey on the south end of the island, as shown in Figure M.5. The data was collected along the crest of the levee using the roll-along method. These surveys used 48 geophones with a constant spacing of 2 m on a land streamer. The source location was 9 m from the start of the first geophone and a truck with a weight was used to create the seismic source. After each shot location, the array was moved 10 m, keeping a source-receiver distance of 9 m.

3.3.1.6 MASW Data Processing

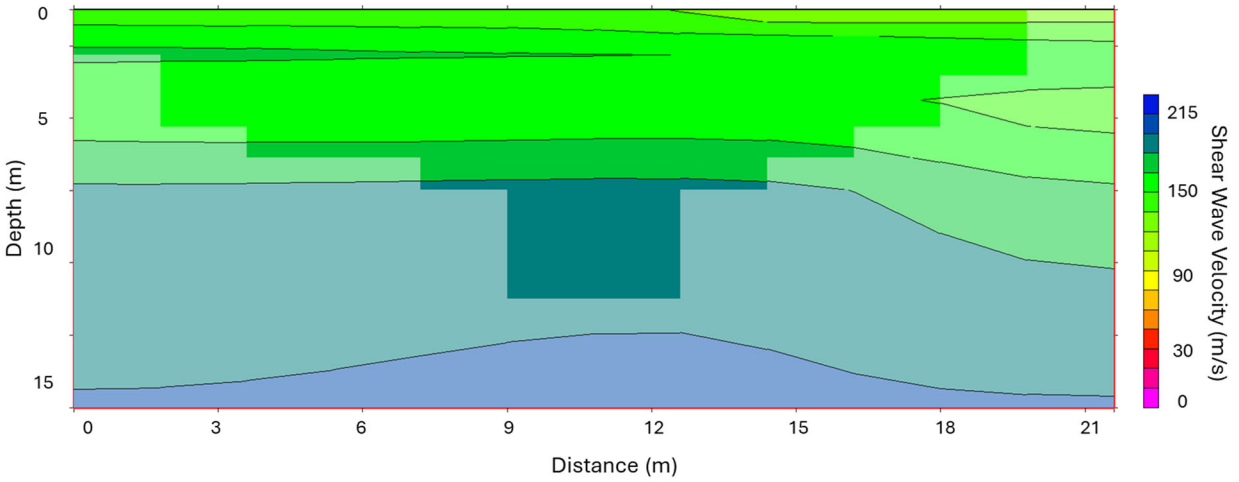
The Grand Island and Bacon Island MASW surveys were processed using the SeisImager/SW software provided by Geometrics. Starting with the Surface Wave Analysis Wizard, the "Active Source 2D MASW" option was chosen, and the first step is to upload the raw files collected by the Geode during data acquisition and check wave arrivals, shown in Figure 3.10(a), and that the geometry is correct. With the Grand Island surveys, the geometry during field acquisition was geometrically referenced for each shot, resulting in no changes needed for geometry. The Bacon Island dataset required manual input of geometry as each shot location was entered as the same.



(a)



(b)



(c)

Figure 3.10 Sample MASW Data showing (a) wave arrival time series, (b) dispersion curves, and (c) 2D shear wave velocity profile.

Once the geometry had been verified, the next processing step was to calculate the common mid-point cross-correlations (CMPCC). The CMPCC are cross-correlations of the wave traces that have mid-point locations in common. When performing an MASW survey, the subsurface structures are estimated at the mid-point of the survey spread. As a result, to increase the lateral resolution of the survey, CMPCCs are used (Hayashi & Suzuki, 2004).

The next step in processing was to calculate phase velocity and generate dispersion curves, shown in Figure 3.10(b). Once the dispersion curves are generated, the *Vs* profile of the subsurface can be generated through inversion to find the *Vs* model that most accurately fits the calculated dispersion curves.

Once the inversion is completed, the 2D *Vs* profile with depth is displayed, as shown in Figure 3.10(c). This data was downloaded and brought into Python where the data was georeferenced with x- and y- UTM Zone 10N coordinates for Grand Island and Bacon Island. The MASW data for Grand Island was also assigned elevation values from the LIDAR surveys. Bacon Island MASW data was assigned elevation based on assumptions made using CPTs and elevations provided with the ERT data.

3.3.2 Electromagnetic Induction

EMI is a geophysical technique that non-invasively measures the apparent conductivity of the subsurface and responds to material characteristics such as fluid composition, clay content, and porosity of the soil (Everett, 2013). This method was based on Maxwell's Equations, which combines Gauss's Law, Ampere's Law, and Faraday's Law to produce a set of coupled partial derivatives for electrical and magnetic fields as shown in Equations 3.8 and 3.9, respectively (Burger, Sheehan, & Jones, 2006).

$$\nabla^2 \mathbf{E} = \xi\psi \frac{\partial \mathbf{E}}{\partial t} + \xi\eta \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{3.8}$$

Where **E** is the electric field, ξ is the magnetic permeability, ψ is the electrical conductivity, and η is the dielectric permittivity.

$$\nabla^2\mathbf{B} = \xi\psi\frac{\partial\mathbf{B}}{\partial t} + \xi\eta\frac{\partial^2\mathbf{B}}{\partial t^2} \tag{3.9}$$

and **B** is the magnetic field.

Each equation has two terms. The $\xi\psi\frac{\partial}{\partial t}$ term relates to the conduction of currents and when this term dominates, then the electromagnetic (EM) field will diffuse into the subsurface. The process of the EM field diffusing into the subsurface is the driving mechanism behind frequency and time domain EM (FDEM and TDEM, respectively). In comparison, the $\xi\psi\frac{\partial^2}{\partial t^2}$ term relates to the displacement of currents in the subsurface, and when this term dominates, the EM field will propagate like a wave, which is the mechanism used in GPR (Burger, Sheehan, & Jones, 2006).

EMI methods are conducted using two coils, a transmitter coil and a receiver coil. First, an electrical current is generated, and it flows through the transmitter coil, creating a primary EM field, shown by the blue lines in Figure 3.11. When this primary EM field interacts with a conductive body in the subsurface, Eddy currents are generated which loop around the conductive body, as depicted by orange lines in Figure 3.11. These eddy currents generate a secondary EM field, depicted by green lines in Figure 3.11. The secondary EM field depends on the characteristics of the conductive body's geometry, position, and conductivity (Everett, 2013).

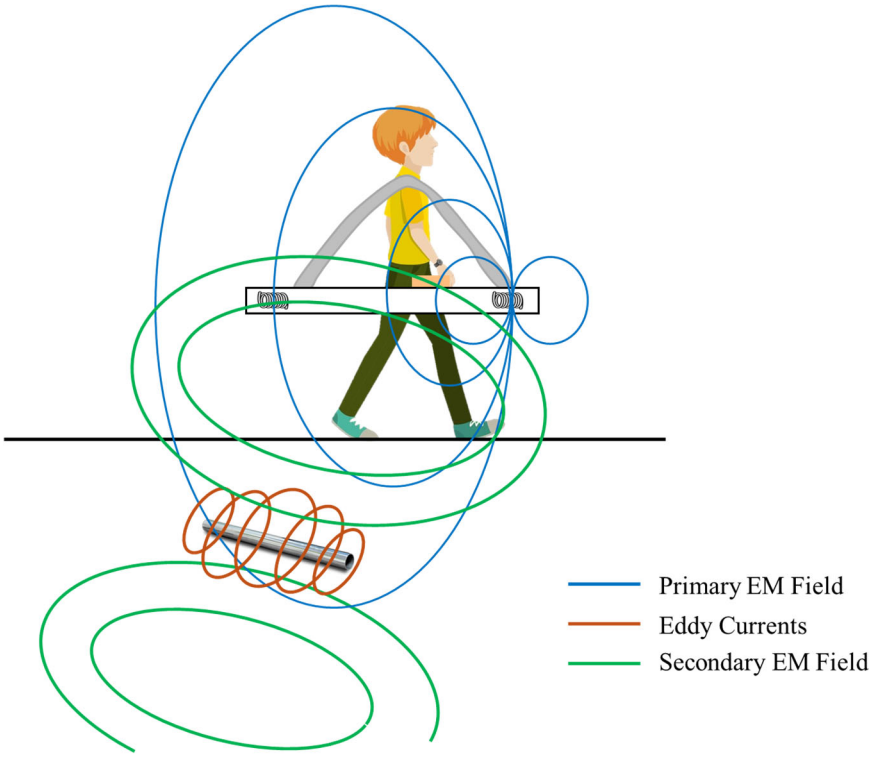


Figure 3.11 EMI sensor showing the eddy currents and the primary and secondary electromagnetic fields

Because the primary EM field is propagating in all directions, the primary EM field is going to interact with the secondary EM field. This causes the receiver coil to measure both the primary and secondary EM field, even though only the secondary field is the one of importance. Since the primary EM field is known, it can be separated from the secondary EM field after it has been measured by the receiver coil. Isolating the secondary EM field is important for gaining accurate conditions of the subsurface.

When determining the depth of investigation for EMI surveys, the skin depth needs to be calculated. The skin depth (δ) is the depth where the EM signal attenuates to approximately 37% of the incident amplitude. Skin depth can be calculated using Equation 3.10 (Everett, 2013):

$$\delta = \sqrt{\frac{2}{\xi\psi\omega}} \tag{3.10}$$

Where ω is the angular frequency measured in radians:

$$\omega = 2\pi f \tag{3.11}$$

and f is the frequency of the EM wave (Hz) and π is approximated at 3.14.

Equation 3.10 shows that the skin depth is dependent on the electrical conductivity of the subsurface. When the subsurface has a higher electrical conductivity, the skin depth is going to be shallower. Similarly, if Equation 3.11 is substituted into Equation 3.10, the skin depth is also a function of frequency. With higher frequencies, the depth of penetration is also going to be shallower. The reason why electrical conductivity and frequency result in shallower depths of penetration is due to the efficiency of converting EM energy into kinetic energy. A higher electrical conductivity is going to result in higher efficiency of converting EM energy into kinetic energy which is going to result in a shallower skin depth (Everett, 2013).

3.3.2.1 Time Domain EM vs Frequency Domain EM

TDEM and FDEM both utilize diffusing EM signals in the subsurface to map electrical conductivities. A TDEM system starts with a current through the transmitter coil that is rapidly shut off. By rapidly shutting off the current, a transient primary EM field will be generated that is in phase with the current. This is important because it will generate a decaying eddy current due to resistive heat dissipation. The rate of decay of the eddy currents is determined by the electrical conductivity of the subsurface. This decaying eddy current will then produce a weak secondary EM field when the primary EM field is no longer present (Everett, 2013). This procedure allows for the recording of the secondary EM field without the primary EM field also being measured.

The FDEM method is dependent upon the coil spacing and the frequency used. These systems work by transmitting harmonic waves at known frequencies to generate a primary EM field with the form (Everett, 2013):

$$\mathbf{B}^{\mathrm{P}} = \mathbf{B}_0(\varsigma)\sin(\omega t) \tag{3.12}$$

Where ς is the radial distance at the center of the transmitter coil and t is time. Eddy currents are generated, which in turn generate the secondary EM field, $\mathbf{B}^{\mathrm{S}}$, which takes the form (Everett, 2013):

$$\mathbf{B}^S = \mathbf{B}_1 \sin(\omega t + \varphi) \quad (3.13)$$

Where φ is the phase shift between the primary and secondary EM field. The phase shift is important because it is a function of how resistive or conductive the subsurface is. The higher the electrical conductivity, the higher the phase shift. If the phase shift is 90°, then the subsurface is perfectly conducting (Everett, 2013).

With the FDEM method, the receiver coil measures both the primary and secondary EM waves. As a result, the total EM field measured is a combination of the primary and secondary EM waves:

$$\mathbf{B}^T = \mathbf{B}^P + \mathbf{B}^S \quad (3.14)$$

By decomposing $\mathbf{B}^T$ into two orthogonal components, the signal can be broken into a real (Re) response or in-phase signal, and a quadrature (Qd) response, or out-of-phase signal. The equations for the real and quadrature signals are expressed in Equations 3.15 and 3.16, respectively (Everett, 2013):

$$\mathrm{Re} = \mathbf{B}_1 \cos\varphi \quad (3.15)$$

$$\mathrm{Qd} = \mathbf{B}_1 \sin\varphi \quad (3.16)$$

The Re and Qd signals are important as they are the outputs of the FDEM instruments and are used to convert to the apparent conductivity of the subsurface. The apparent conductivity is the electrical conductivity of the half-space that is equivalent to the Qd response (Everett, 2013). As the Re component of the signal increases, the Qd signal will decrease, and the apparent conductivity of the subsurface increases (Burger, Sheehan, & Jones, 2006). The Re and Qd signal is measured at the midpoint between the transmitter and receiver coil.

3.3.2.2 EMI Field Data Collection

The EMI sensor used for surveying was a GEM-2 handheld multifrequency ski from Geophex. It is a broadband electromagnetic sensor that can operate at up to 10 frequencies in the range of 30 Hz to 93 kHz. This sensor was chosen because it is lightweight and allows for quick surveying of large areas. Figure 3.12 shows the setup and how portable the system is. Additionally, because the sensor can operate at multiple frequencies, information with depth can be obtained with one survey.

The more frequencies being operated at once by the sensor, the lower the S/N. Even though the sensor can operate at 10 frequencies at once, the operator suggests not to use more than five frequencies as the S/N significantly decreases. Four frequencies were used in this study: 1530 Hz, 5310 Hz, 18330 Hz, and 63030 Hz. The sensor outputs the Re and Qd components of the signal in parts per million (ppm).

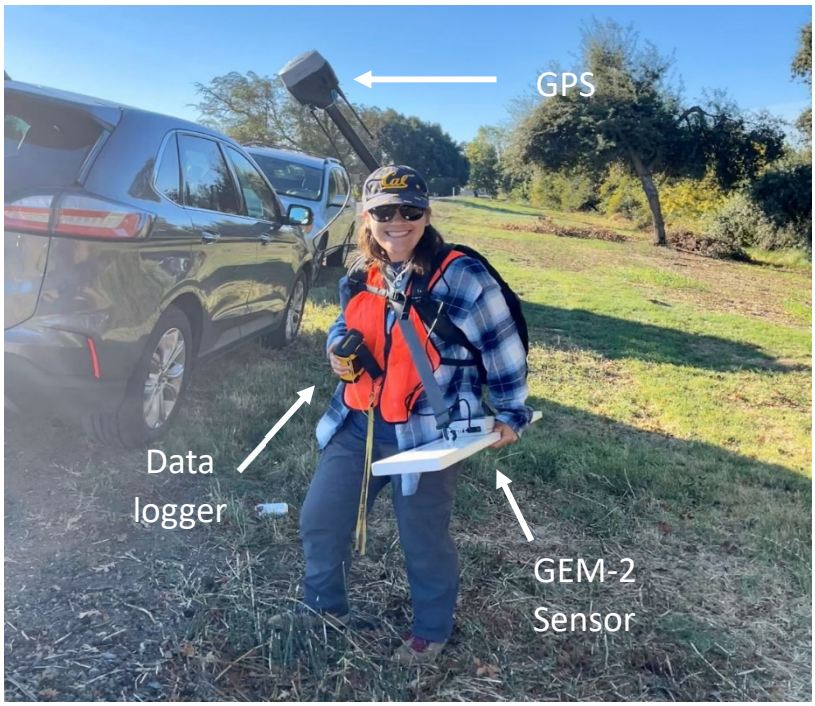


Figure 3.12 GEM-2 sensor connected to Reach RS2 GPS and Trimble Nomad data logger

The GEM-2 sensor is connected via Bluetooth to a Trimble Nomad datalogger that operates the WinGEM2 software for data collection. The GEM-2 sensor is also connected to a GPS to collect precise coordinates every time a measurement is collected. The GPS used for surveying was the Reach RS2 system with multi-band RTK GNSS receivers. This system included a base GPS which was set on a tripod with minimal sky coverage. The base station was set up first and was allowed to average for five minutes before connecting the rover GPS. The rover GPS was attached to a backpack via a pole and connected to the GEM-2 sensor, as shown in Figure 3.12. The GPS signals were checked before surveying at each location.

3.3.2.3 EMI Survey Locations on Grand Island

Because the GEM-2 sensor is non-invasive and lightweight, large amounts of area can be surveyed in a relatively short amount of time. As a result, the sites on Grand Island were surveyed densely compared to the MASW profiles. Figure 3.13 shows the change detection EMI survey lines for Site 1, Site 2, and Site 3 that were used for analysis.

For Site 1, there were two surveys parallel to the levee along the crest on either side of the road starting at the east end of the site until the west end of the site. On the dry side of the levee, a survey parallel to the levee centerline was conducted on the toe as well as on the slope halfway between the crest and toe. On the wet side of the levee, there were two surveys along the toe. One survey was on the west side of the site and stopped at the eucalyptus grove. The second survey was on the east side of the site through the eucalyptus grove to the eastern end of the site. The last

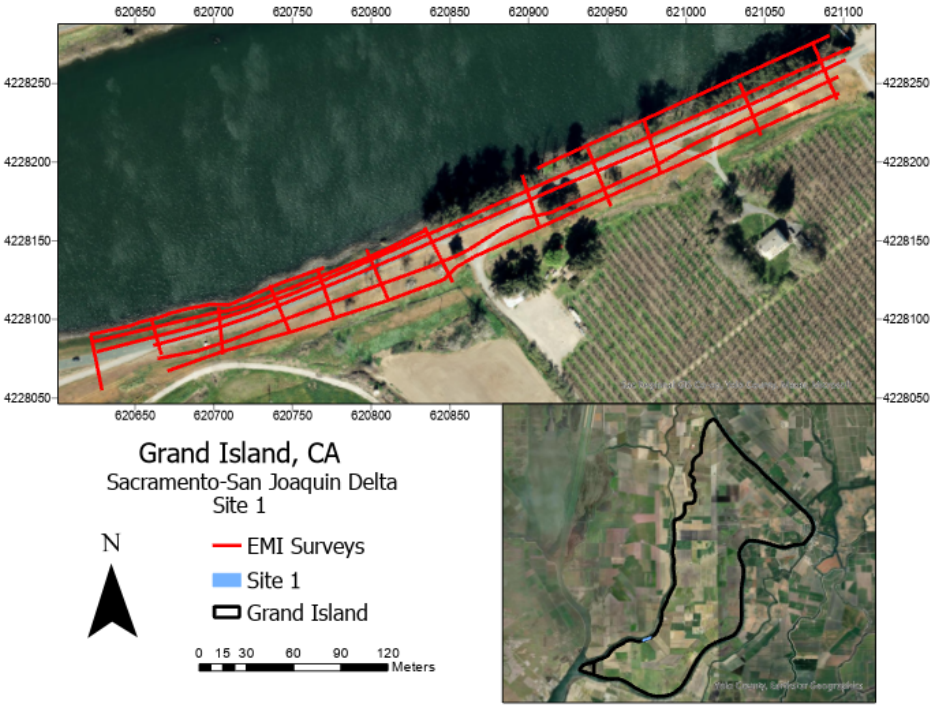
longitudinal survey was on the wet slope midway down and started on the west side of the site and ended at the eucalyptus grove. Twelve cross-sectional surveys span the entire length of the site that crossed the wet toe, crest, and dry toe.

At Site 2, the dry season deployment resulted in two longitudinal crest surveys on either side of the road that span the entire length of the site. There was a longitudinal survey on the wet toe, dry toe, as well as halfway up the dry slope. Along the length of the site spans seven cross-sectional surveys that span from the trees on the wet to the dry toe.

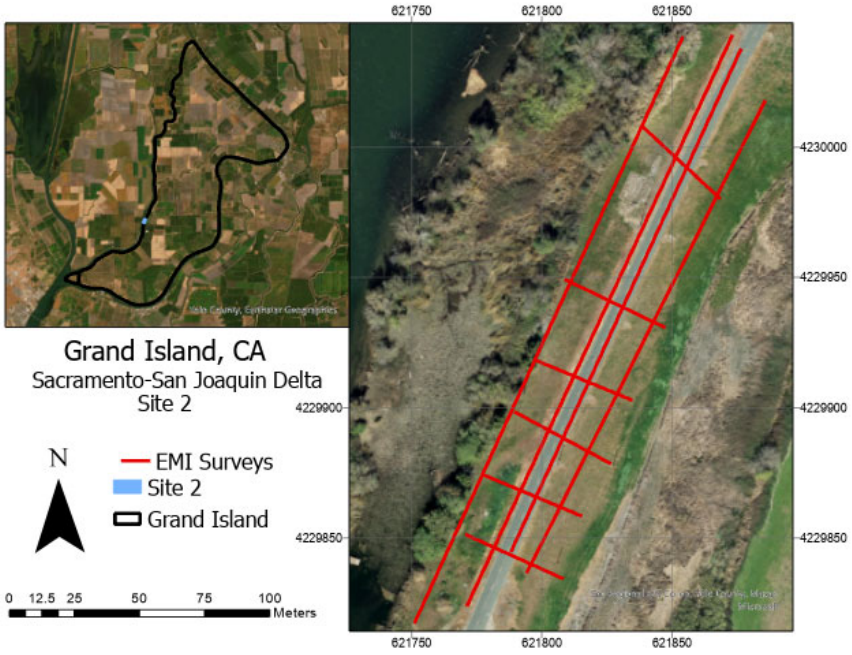
During the wet season deployment, the vegetation was thick on the slopes and made some locations impossible to survey. All the surveys from the dry season were repeated during the wet season except for the dry toe and Cross Section 6.

At Site 3 during the dry season, multiple longitudinal surveys spanned the entire length of the profile on the wet toe, the crest on either side of the road, the dry toe, and halfway up the dry slope. There was a total of six cross-sectional surveys that started at the wet toe and ended at the dry toe.

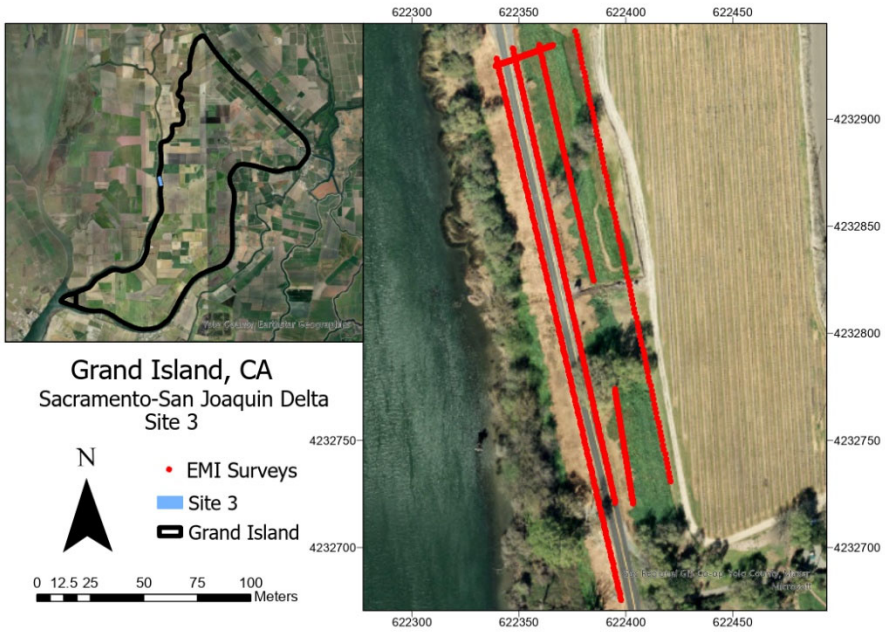
Like with Site 2, the vegetation was so densely packed on the slopes that some survey locations from the dry season were impossible to traverse. These sites include five of the six cross sections and some of the dry mid-slope survey. Data collection was attempted on the wet toe but due to backtracking and multiple issues during data collection, this survey was not used for change detection due to the poor quality of the data.



(a)



(b)



(c)

Figure 3.13 EMI surveys on Grand Island for (a) Site 1, (b) Site 2, and (c) Site 3

3.3.2.4 EMI Processing

The EMI data is first processed by inputting the raw data into ArcGIS Pro. The elevation data from the LIDAR data was spatially joined to the raw EMI data. The raw EMI data with associated elevations were exported for inversion processing.

To process the EMI data to create 2D ρ profiles, Aarhaus Workbench was used. The first step in processing the data is the coil configuration. This includes defining the frequencies used, the transmitter-receiver coil separation, as well as the transmitter and receiver orientation. Other variables that are defined are the sensor height above the ground surface and the coordinate system. Table 3.4 summarizes the inputs for the coil configuration setup.

Table 3.4 Coil Configuration Setup Parameters

Name	Parameter
Number of Coil Configurations	4
Frequencies (Hz)	1530, 5310, 18330, 63030
Coil Separation (m)	1.66
Transmitter Orientation	Z – vertical magnetic dipole
Receiver Orientation	Z – vertical magnetic dipole
Sensor Height (m)	1
Coordinate System	WGS 84 UTM Zone 10N

After the coil configuration was complete, the next step was to map the information in the input file to the variables in the program. Important parameters to select include elevation, x-coordinate, y-coordinate, powerline, and the Qd and Re data for each frequency.

After the data was inputted into the program, the data was initially processed before performing the inversion. This includes deciding when to remove negative data. Negative data can be removed before processing, after processing, or not removed. How negative data is removed was iterated on while processing the data. A polynomial interpolation order can also be specified. This forces the data to use the running mean or fit the data to a polynomial with an order that is specified. The polynomial interpolation order was set to zero to use the running mean. The sounding distance also needs to be specified and is the distance between the soundings. It is recommended to use an input that is greater than the sounding distance of the raw data. The sounding distance for this study was set to 0.5 m.

The data can also be automatically smoothed to remove noise by setting the running mean width for a running mean filter for each frequency channel. A smaller running mean width is useful for preserving geologic features in the subsurface while a larger running mean width smears out geologic features but suppresses noise. Since choosing the proper running mean width can be difficult, this parameter is iterated on to determine the model that represents the profile best while minimizing the total residual. The final input for the initial processing is the GPS shift along the profile. This GPS shift is the relative position of the frequency channels along the profile relative to the GPS receiver. For this data, the GPS shift was set to 0 meters for all channels.

After the initial processing of the data, the next step was to perform the inversion to create 2D models. The Aarhaus Workbench software leverages the Levenberg-Marquadt Algorithm (or the

Damped Least Squares method). This is an iterative procedure used to solve non-linear least squares problems (Levenberg, 1944). This algorithm is a combination of gradient descent and Gauss-Newton (Gavin, 2022). Gradient descent chooses values to minimize the cost function in the steepest-descent direction while Gauss-Newton assumes that the function is quadratic and finds the minimums by finding the x-intercepts. The Levenberg-Marquadt Algorithm iteratively chooses gradient descent or Gauss-Newton through λ, a damping parameter that smooths the graph. When λ is small, the parameters are close to the optimal value and Gauss-Newton is chosen. Conversely, when λ is large, the parameters are far from the optimal value, and gradient descent is chosen (Gavin, 2022).

When performing the inversion, a laterally constrained inversion (LCI) or a spatially constrained inversion (SCI) can be performed. An LCI model uses 1D models that are laterally constrained to produce continuous 2D ρ models. The SCI model uses 1D models that are constrained to produce 3D models of ρ profiles.

An LCI model was used to produce 2D ρ profiles of the subsurface. The inversions used smooth models which discretizes the ρ model using multiple layers that have fixed layer boundaries. For each inversion, the Qd and Re channels are chosen to be included in the processing. Removing a Re channel can help build a better profile by reducing noise. Choosing which channels to include in the analysis was iterated on to minimize the total residual while still producing data with good information at depth.

An example of how iterations were performed on a survey is shown in Table 3.5 with the final inversion shown in Figure 3.14.

Table 3.5 Iterations of EMI Inversions

Iteration	Running Mean	Negatives	Channels	Residuals	Notes
1	0.5 for all	Keep them	All	2.23	Noisy but okay profile
2	0.5 for all	Remove before	All	2.24	Like iteration 1 but noisier
3	0.5 for all	Remove after	All	2.20	Like iteration 1
4	2 for ch. 1 and 2, 0.5 for rest	Remove after	All	1.78	Looks like iteration 3
5	2 for ch. 1 and 2, 0.5 for rest	Remove after	No Re 1	1.67	Looks like iteration 4

In this example, iteration 1 used a running mean width of 0.5 m for all channels, the negatives were kept, and all frequency channels were used. The final total residual was 2.23 and the profile was noisy but acceptable. The next two iterations kept the running mean width and the channels the same. How negative data are used changed in iterations 2 and 3. In iteration 2, the negative data was removed before processing resulting in a higher residual and creating a noisier profile. Iteration 3 removed negative data after processing and reduced the residual from 2.23 to 2.20 without significantly affecting the quality of the profile. Removing negatives after was used for the remaining iterations as it had a lower residual compared to iteration 1.

To further reduce the residual, the running mean width of channels 1 and 2 was changed to 2 m and 1 m for channels 3 and 5. The higher the running mean width, the smoother the data. Channels 1 and 2 were noisy so a higher running mean width was applied to reduce the noise in the profile. The resulting residual was 1.78 which was lower than iteration 3 without reducing the quality of the profile.

Keeping the new running mean widths, the channels included were iterated to try to further reduce the residual. Channel 1 was the noisiest and significantly affected how well the model fits the data. To help the data fit the model better, the Re data for channel 1 was removed from the analysis. This resulted in a residual of 1.67 and the profile looked like iteration 4. Since this iteration had the smallest residual while maintaining the quality of data in the profile, it was chosen as the final iteration. The 2D profile for iteration 5 is shown in Figure 3.14.

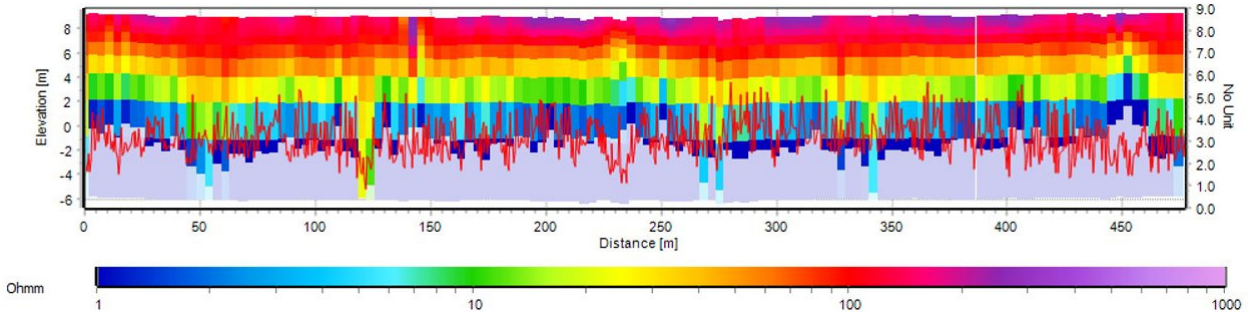


Figure 3.14 Final electrical resistivity inversion profile

This figure shows the distance of the survey profile on the x-axis and the elevation on the left y-axis. The ρ is represented as 1D bars that are stitched together and colored based on ρ value. The right y-axis represents the residual, which is shown as a red line along the profile. Looking at where the residual is high along the profile can be used to determine if a channel is not fitting the model well and can help to determine if it should be removed from the inversion.

3.3.3 Electrical Resistivity Tomography

ERT is a geophysical method that leverages direct current to determine the ρ of the subsurface through the potential difference between electrodes. Subsurface structures and lithology can be identified through different resistances to current flow characterized by variations in potential differences (Burger, Sheehan, & Jones, 2006). Since ERT uses direct current, it follows Ohm's Law:

$$\mathrm{i} = \frac{\Lambda}{\mathrm{R}} \tag{3.17}$$

Where i is current, Λ is voltage, and R is resistance. For soil, resistance is not just a function of current and voltage though, but also of the dimensions of the material the current flows through such as the length (l) and cross-sectional area (A). The dimensions of the material can be used to relate resistance to the ρ of the material through Equation 3.18:

$$\rho = \mathrm{R}\frac{\mathrm{A}}{\mathrm{l}} \tag{3.18}$$

To measure the ρ of the subsurface, a series of electrodes are placed in the ground where some electrodes act as source electrodes and some act as sink electrodes. The source electrodes are positive current electrodes where current is introduced into the subsurface. The current radiates outward equally in all directions through the subsurface creating equipotential surfaces. The sink electrodes are negative current electrodes that are used to measure the potential difference in the subsurface. The potential difference at a point P from the source and sink electrodes can be measured using Equation 3.19:

$$\Lambda_p = \frac{\rho I}{2\pi}\left(\frac{1}{r_{AP}} - \frac{1}{r_{BP}}\right) \tag{3.19}$$

Where r_{AP} is the distance from point P to the source electrode and r_{BP} is the distance from point P to the sink electrode.

When performing an ERT survey, multiple electrode configurations should be considered including the Wenner array, Schlumberger array, and dipole-dipole array. The Wenner array has equal spacing (a) between all electrodes regardless of being a potential or current electrode, as shown in Figure 3.15. The current electrodes are at the end of the survey with the potential electrodes between the two current electrodes (Everett, 2013).

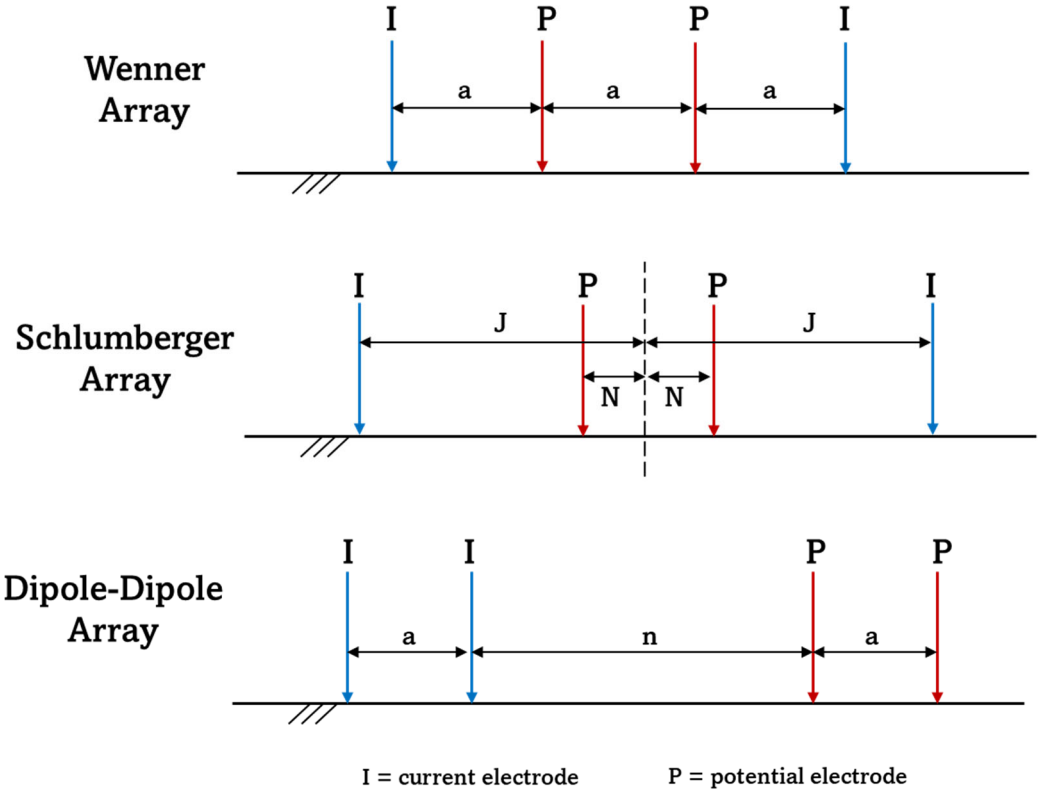


Figure 3.15 Electrical Resistivity Tomography electrode configurations

The Schlumberger array is like the Wenner array with the current electrodes at the end of the survey and with two potential electrodes between the current electrodes. The difference between a Schlumberger configuration and the Wenner configuration is the electrode spacing. The current

electrodes are placed at a distance J from the center of the array while the potential electrodes are placed at a distance N from the center of the array. The spacing between the current and electrodes should satisfy 2J > 5N.

The third electrode configuration is the dipole-dipole array. With this array configuration, the spacing between the current electrodes and the potential electrodes are both a. The two potential electrodes are on one side of the array with the pair of current electrodes on the opposite side of the array with a spacing of n as shown in Figure 3.15.

3.3.3.1 Bacon Island ERT Field Data Collection

ERT data was collected at 3 sites on Bacon Island in the Sacramento-San Joaquin Delta as shown in Figure 3.3. Data was collected at two sites on the northern section of the island, Northwest and Northcentral, as well as one site on the southern levee. All ERT surveys were performed parallel to the levee crest and were on the dry toe of the levee to ensure good soil and electrode interaction.

The ERT data was collected using a dipole-dipole configuration (Uhlemann, et al., 2021). The dipole spacing between both the current and potential electrodes (a) was 2, 4, 6, 8, and 10 m while the spacing between the current and potential electrodes (n) was 1 to 8a.

3.3.3.2 ERT Processing

The ERT data was processed using ResIPy, an open-source Python GUI for ρ data processing through inversion modeling (Blanchy, Saneiyan, Boyd, & Binley, 2020). The first step is to input the raw data into the program and add elevation data if available. The next step would be pre-processing where the raw data can be filtered before inversion. Before the inversion model can run, a mesh needs to be generated over the subsurface. For the ERT data processing, a fine triangular mesh was used. The final step was to start the inversion.

ResIPy uses an iterative weighted least squares inversion model with regularization options such as time-lapse data analysis. This model aims to minimize the misfit between a predicted response and the ρ from the electrodes to find the best geologic parameters that are like the observed measurements (Blanchy, Saneiyan, Boyd, & Binley, 2020).

3.3.4 Comparison of Laboratory Electrical Resistivity to Field Electrical Resistivity

Since the results from the laboratory testing are being applied to data collected in the field, it is important to understand how ρ measured in the field compares to ρ measured in the laboratory. This was done by performing two short EMI surveys using the GEM-2 sensor at a site located in Berkeley Hills. The first data collection took place on November 19th, 2023, after approximately 26 mm of precipitation in three days. The second data collection occurred on December 03, 2023, after approximately 3 mm of precipitation over two days, as shown in Figure 3.16. (NOAA, 2023). On both days, soil samples were taken from approximately 0.3 m below ground level for laboratory testing and placed in airtight Ziploc bags until testing.

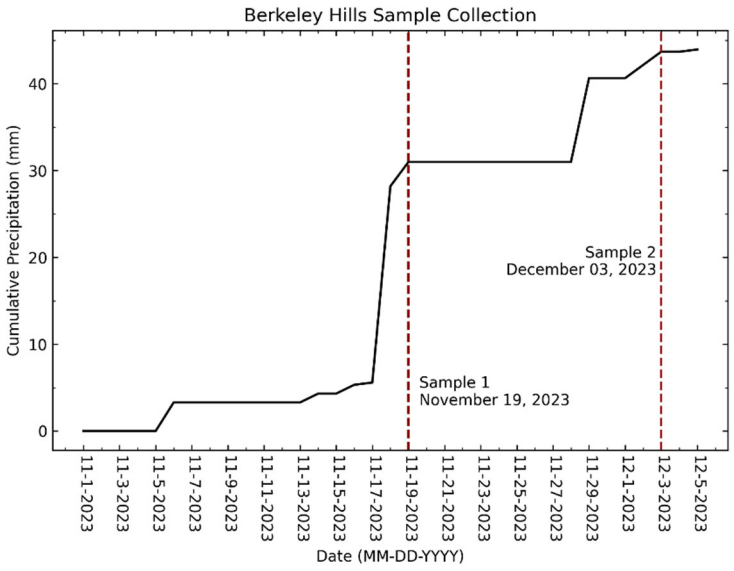


Figure 3.16 Berkeley Hills precipitation and data collection (modified from NOAA, 2023).

The field data was inverted using the processing described in Section 3.3.2. The laboratory testing was performed using the same setup described in Chapter 2. The water contents of both samples were measured. Sample 1, which was collected on November 19th, had a moisture content of 14%, and Sample 2, which was collected on December 3rd, had a moisture content of 10%. The laboratory results were compared to the upper 1 m of the field profile, as shown in Figure 3.17.

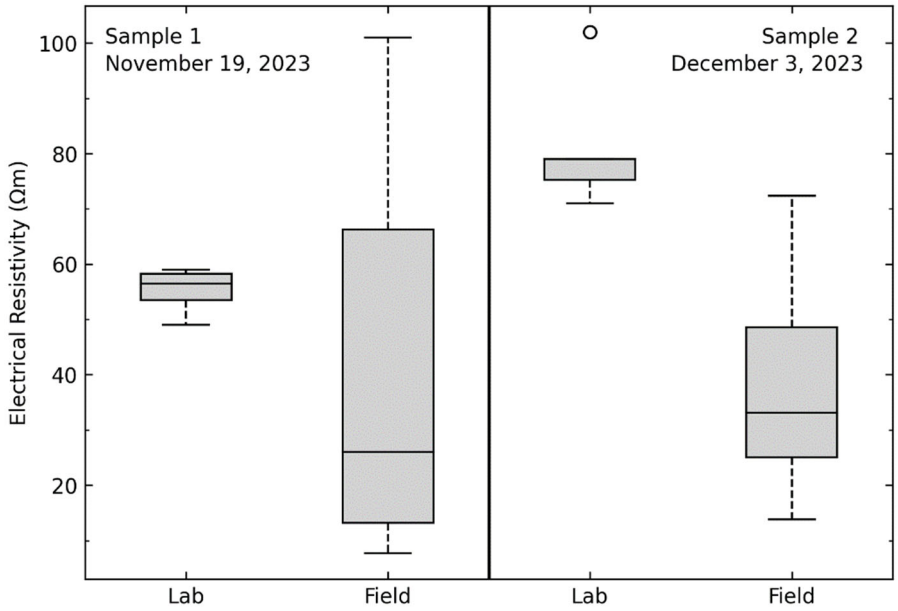


Figure 3.17 Field and laboratory comparison of electrical resistivity

During the November 19th data collection, the upper 1 m field data showed ρ that ranged from 8 Ωm to 101 Ωm with a median of 26 Ωm. The laboratory ρ ranged from 49 Ωm to 59 Ωm with a median of 57 Ωm. The laboratory matched closely with the field data as it fell under the range shown in the field. Other factors such as vegetation and geologic units below where the soil sample was taken could affect the measurements recorded in the field.

During the December 3rd data collection, the upper 1 m field data showed ρ that ranged from 14 Ωm to 72 Ωm with a median of 33 Ωm. The laboratory ρ ranged from 71 Ωm to 102 Ωm with a median of 79 Ωm. While the field measurements were lower in ρ compared to the laboratory ρ, the laboratory measurements were similar. It is expected for the sample collected in December to have higher ρ than the sample in November because the sample in December had a lower moisture content. With less water in the sample, the soil was more resistive, resulting in higher ρ.

One reason for the field measurements having lower values than the laboratory measurements is that the field measurements are averaging in the soil below and above. With free-standing water on the ground surface due to the rain from the night before, the ρ collected by the EMI ski for near-surface measurements would be slightly lower than the actual ρ of the soil.

3.4 Unmanned Aerial Vehicle Mounted Sensors

The use of remote sensing for infrastructure inspection and maintenance has become very popular in civil engineering as UAVs and satellite imagery allow for more efficient inspection, a reduction in human bias, and increased feasibility of inspecting locations that are difficult for someone to get to or are dangerous.

Because levees are spatially distributed systems, the use of UAV-enabled sensors to image the surface can significantly increase the temporal frequency of inspection and increase the efficiency of surface inspection while reducing human bias. Unlike satellite imagery, UAV imagery allows for rapid deployment which is important to collect data about failure mechanisms. Satellite imagery cannot be collected on demand and the user must wait for the satellite to revisit the location of interest. Some satellites have low temporal revisit times such as the Landsat Satellites. These satellites will revisit a point on the Earth every 16 days (Markham, 2013). This could become problematic if a potential failure mechanism has been identified and the site needs to be imaged immediately. Satellite imagery also typically has lower spatial resolution compared to UAV-enabled sensors. The Landsat 8 satellite has a spatial resolution of 30 m for the red, blue, and green bands while a UAV-enabled camera can have centimeter-level resolution (Markham, 2013). As a result, UAV-enabled sensors were chosen as optimal for levee inspection as the inspector can choose when to collect data and they have higher spatial resolution.

The drone used to collect data was a DJI Matrice 300 with three UAV-enabled sensors to image the surface of the levees in Grand Island, including a Red, Blue, and Green (RGB) camera, a thermal camera, and a LIDAR sensor, shown in Figure 3.18.



(a)



(b)

Figure 3.18 DJI Matrice 300 with (a) equipped with the thermal camera and (b) equipped with the LIDAR sensor.

3.4.1 Structure-From-Motion

Using Structure-From-Motion (SfM), hundreds of images collected from UAVs were stitched together to create detailed images and 3D models of each site. Figure 3.19 describes the workflow for the SfM procedure.

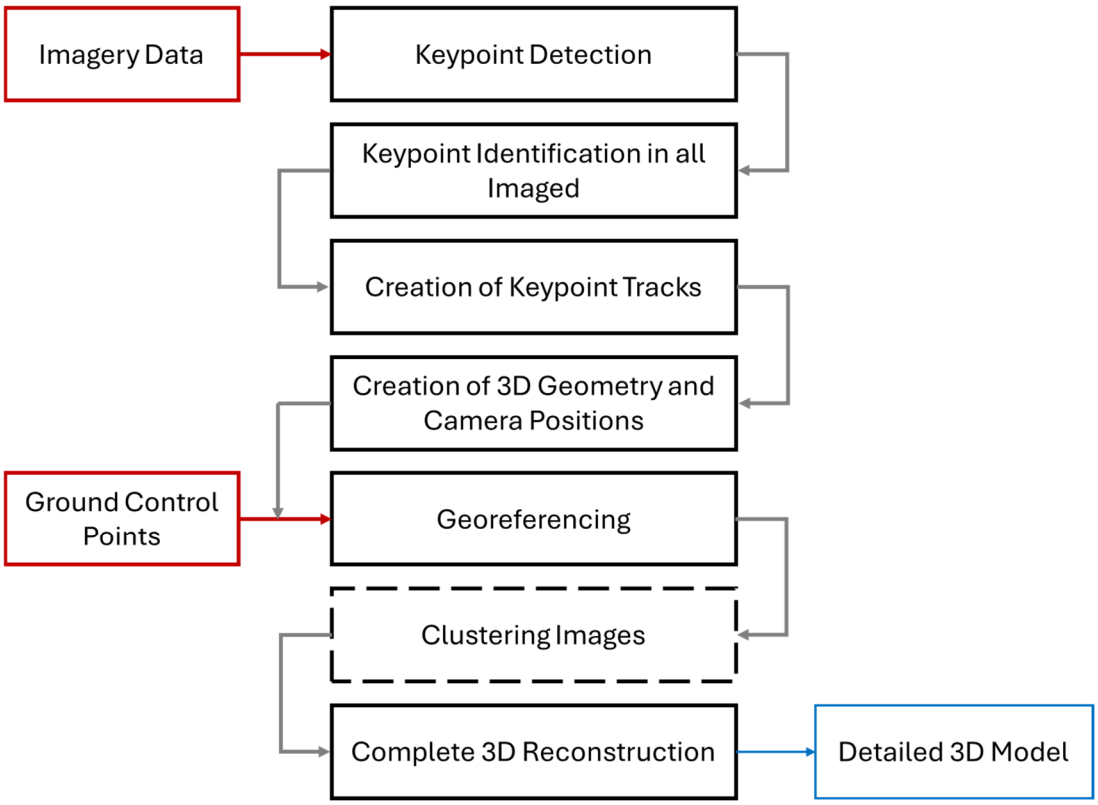


Figure 3.19 Structure-from-Motion workflow for 3D model construction (Carrivick, Smith, & Quincey, 2016)

The first step in the SfM workflow is to input the imagery dataset and keypoints are detected. Keypoints are points identified in numerous photos and act as common points. Once the keypoints are chosen, all the images are searched for the keypoints and are identified. For a keypoint to be useful in the algorithm, a minimum of three images requires that keypoint to be present. The keypoints that do not have a minimum of three corresponding partners in other images are filtered out. Tracks and links between images are created based on partnered keypoints. Tracks without at least two keypoints are filtered out. These tracks are used to connect images among the entire dataset (Carrivick, Smith, & Quincey, 2016).

Once the tracks between all the images are created, the 3D geometry and camera positions are reconstructed. The product of this step is a sparse point cloud and relative camera locations in an arbitrary coordinate system. To perform temporal comparisons of the models, Ground Control Points (GCPs) are inputted into the model with known x- and y-coordinates and elevations (Z) in

a global coordinate system. A survey should include a minimum of three GCPs and should avoid being placed in a linear alignment while being widely distributed amongst the site (Carrivick, Smith, & Quincey, 2016).

Often the imagery datasets to create SfM models are large, containing hundreds of photos. As the size of the dataset increases, the processing time increases. To make the process more efficient for large datasets, an optional step is to cluster the images. This step splits the model into smaller sections to reduce the amount of RAM needed to complete the modeling. The final step was applying a Multi-View Stereo algorithm to create a dense point cloud showing the 3D reconstruction of the site's geometry (Carrivick, Smith, & Quincey, 2016).

3.4.2 Red, Green, Blue Imagery (Optical Imagery)

Digital imagery is an important remote sensing tool for representing an area of interest. A digital image is a 2D array of pixels. Each pixel is assigned a number that represents brightness. The dimensionality of an image can be increased by adding layers, called bands, which represent the same area but contain different images (Liu & Mason, 2016).

Each band of an image represents a range of wavelengths on the electromagnetic spectrum. The electromagnetic spectrum ranges from gamma rays at short wavelengths to radio waves at long wavelengths. The spectrum showing the location of each remote sensing sensor is represented in Figure 3.20.

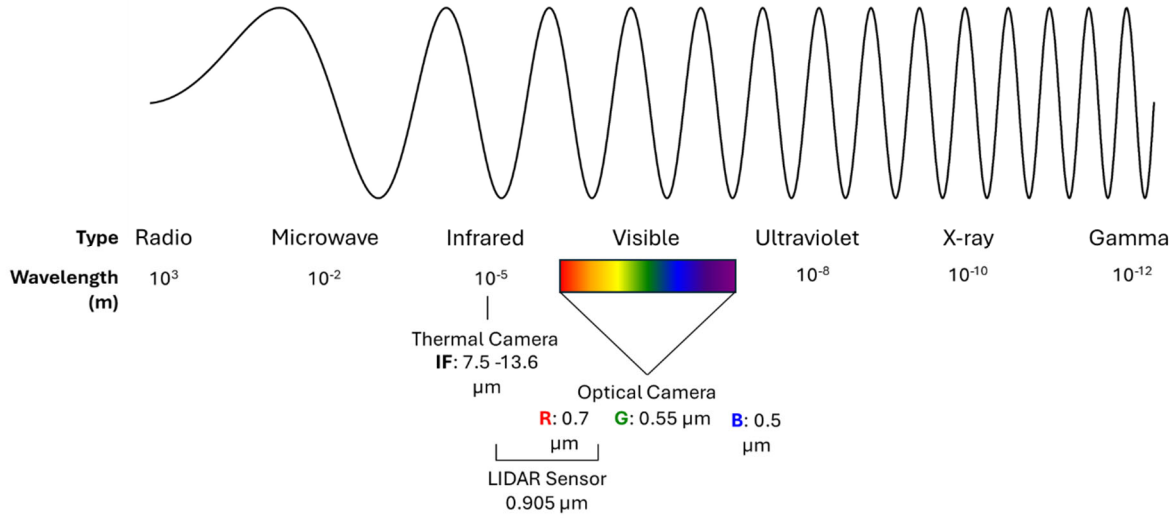


Figure 3.20 Electromagnetic spectrum showing wavelengths of the remote sensing sensors

The scale and spatial resolution of the images when using UAV cameras are dependent upon the focal length (fl) of the sensor as well as the flying height (fh) of the UAV. The ratio of the fl and fh is the scale factor (sf). As the sf increases, the ground coverage also increases, resulting in lower spatial resolution. The sf of the photo can be described by (Bakker, et al., 2001):

$$sf = \frac{fh}{fl} \tag{3.20}$$

Changing the focal length of the camera or the height of the UAV will affect the scale of the image (Bakker, et al., 2001). Increasing the height of the UAV or decreasing the lens's focal length while keeping the other constant will increase the scale of the image.

3.4.2.1 Optical Imagery Survey Acquisition

The optical camera used for this analysis was the DJI Zenmuse XT2 thermal camera which has both optical and thermal sensors. The optical sensor is a 1/1.7" complementary metal oxide semiconductor (CMOS) image sensor with 12 million effective pixels. The lens focus is 8 mm with a field of view of 57.12° x 42.44° (DJI, 2018). This camera was used to capture hundreds of digital images of levee surfaces using the red (0.4-0.5 μm), green (0.5-0.6 μm), and blue (0.6-0.7 μm) wavelengths.

Optical imagery was collected at the three levee sites on Grand Island twice. The first deployment was during September 2022 to image the dry season and the second deployment was during February 2023 to image the wet season. For both deployments, the drone flew at 48 m above ground surface using a side overlap of 80% and a frontal overlap of 80%. The number of photos collected for each survey is listed in Table 3.6.

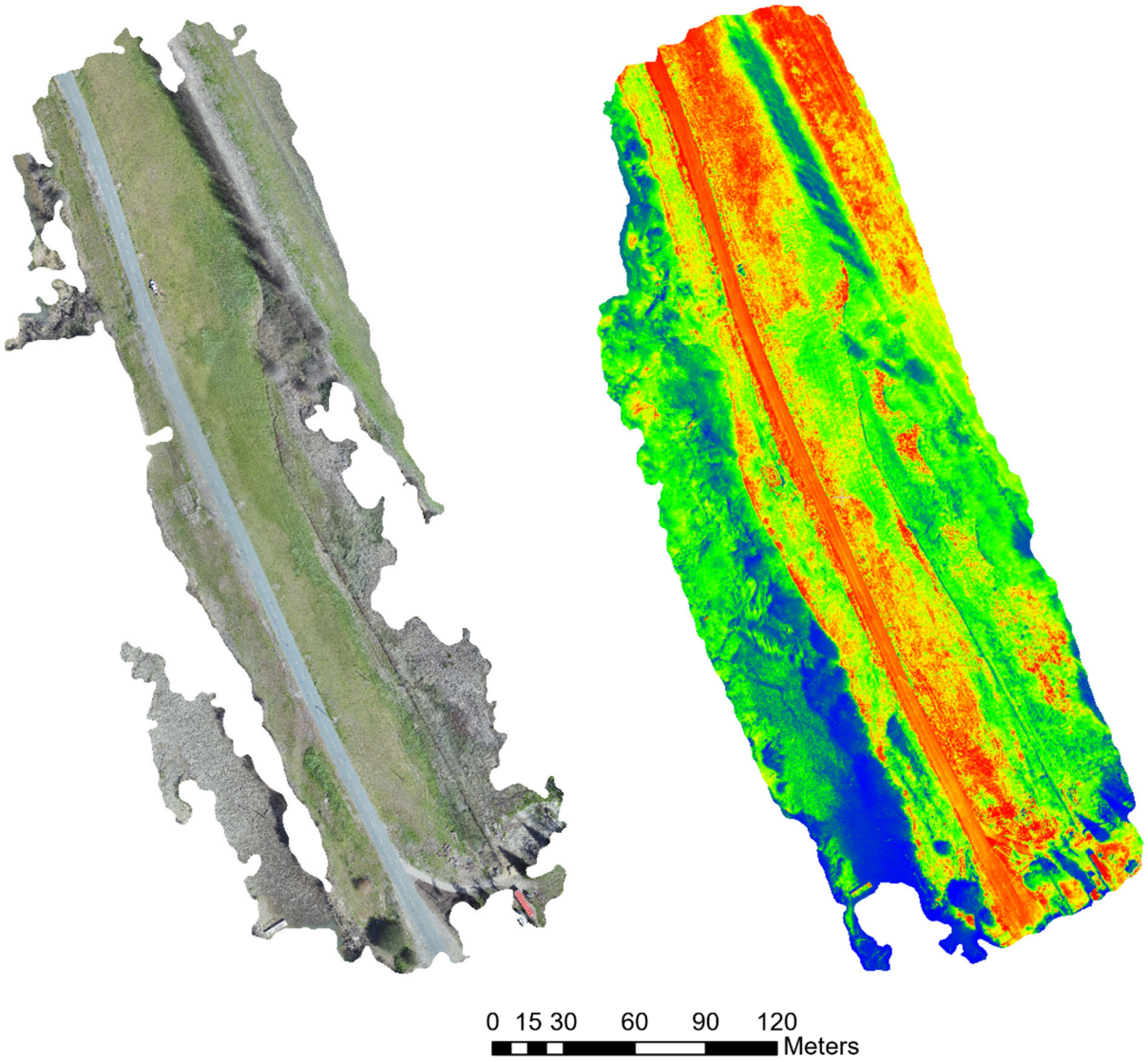
Table 3.6 Number of Images Collected per Survey

	Number of Images Collected	
Location	**September Deployment**	**February Deployment**
Site 1	851	269*
Site 2	701	606
Site 3	203	462

* high wind conditions

The collection of images during the February deployment was challenging due to high wind conditions. As a result, the number of photos collected during the February deployment for Site 1 and Site 2 was significantly smaller compared to the September deployment. The UAV had to be landed early when the wind speed became too dangerous to fly in.

The images collected were processed using Pix4D with the "3D Model" template for SfM. During the data processing stage, four GCPs were included in the model for each site and deployment to increase location accuracy. Processing the data resulted in 3D textured meshes of each survey, a point cloud of each survey, and high-resolution orthomosaics of each survey. An example of a 3D textured mesh for Site 2 on Grand Island is shown in Figure 3.21(a). Images of the meshes, orthomosaics, and point clouds for each survey can be found in Appendix D.



(a) (b)

Figure 3.21 Examples of Grand Island Site 2 showing the (a) 3D textured mesh and (b) 3D thermal signature.

3.4.3 Thermal Imagery

Thermal imagery is a branch of remote sensing that specializes in the collecting and interpretation of infrared wavelengths to understand the thermal properties of the landscape. Where RGB imagery uses reflected wavelengths to determine properties of the ground surface, thermal imagery measures the radiation emitted by objects to estimate radiant temperature (T_{rad}) (Prakash, 2000).

The T_{rad} measured by the sensor is a function of the kinetic temperature (T_{kin}) and the emissivity (ϵ) as shown in Equation 3.21 (Bakker, et al., 2001):

$$T_{rad} = \epsilon^{\frac{1}{4}} T_{kin} \quad (3.21)$$

The kinetic temperature is generated by the vibration of the molecules in the object and is dependent upon many factors including solar heating, cloud cover, altitude, thermal conductivity, and heat capacity. Kinetic temperature is measured in terms of Kelvin (K), Celsius (C), or Fahrenheit (F) (Prakash, 2000).

In contrast, the emissivity of an object is a function of the object's composition and geometric configuration and is the ability to radiate energy compared to a black body. A black body is a theoretical object that absorbs and emits all energy at all wavelengths. Emissivity is described as a ratio between 0 and 1, where 0 represents an object that is a perfect reflector and 1 represents the black body. Bare soil and vegetation typically have emissivity ranges close to 0.9 (Prakash, 2000).

3.4.3.1 Thermal Imagery Survey Acquisition

The thermal camera used to survey the levees was the DJI Zenmuse XT2 thermal camera. The thermal sensor operated at spectral bands between 7.5 to 13.6 μm with a pixel pitch of 17 μm and a full frame rate of 30 Hz. The Noise Equivalent Temperature Difference (NETD) was <50 millikelvin (mk). The NETD describes the thermal sensitivity of the sensor. With a smaller NETD value, the sensor can better register small temperature differences. According to the manufacturer, a NETD value of < 50 mk is considered "Good" for the quality of thermal temperature detection.

Like with the optical imagery, the thermal imagery was collected at the three Grand Island levee sites with two deployments, the first during the dry season and the second during the wet season. All the flights were flown at 48 m above ground level with 80% frontal and side image overlap. The number of thermal images collected is the same as the number of optical images collected. Refer to Table 3.6 for the number of photos collected per survey.

The images were processed using Pix4D with the "Thermal Camera" template for SfM. Like processing the optical imagery, four GCPs were used to increase global accuracy creating 3D texture meshes and high-resolution orthomosaics of each survey. Figure 3.21(b) shows an example of the thermal 3D textured mesh for Site 2 on Grand Island. Images of the 3D models and orthomosaics for Site 1 and Site 3 can be found in Appendix D.

3.4.4 LIDAR

Unlike optical and thermal remote sensing, LIDAR is an active remote sensing technique. Optical and thermal sensors record natural wavelengths that are reflected or emitted from a source other than the sensor. With LIDAR, the LIDAR sensor actively emits a pulse of light (beam) that is typically in the optical or infrared wavelength of the electromagnetic spectrum. That beam travels in all directions, hits an object, and is reflected. The light then travels back to the sensor where the two-way travel time of the beam is recorded. Knowing the two-way travel time (t) and speed of light (c), the distance (dist) to the object from the sensor can be calculated through Equation 3.22 (McManamon, 2019):

$$dist = \frac{ct}{2} \quad (3.22)$$

One of the many benefits of LIDAR is the ability to "penetrate" through vegetation which is necessary if trying to create detailed digital elevation models (DEMs). DEMs are models that represent the elevation and structure of the ground surface without vegetation and buildings. These models can be created in areas with vegetation with gaps in the canopy or foliage. When gaps are present, the LIDAR sensor beams will travel through those gaps, resulting in ground information below the vegetation (Carter, et al., 2012).

LIDAR returns are important when creating a DEM. A return is the number of signals received per pulse. For example, the first return measures the first object measured, such as the top of the canopy of a tree. Progressively later returns measure objects that are reflected after the first return. If the first return is the tree canopy, the second return can be vegetation at the ground surface, and the last return can be the ground surface. To create DEMs, the last return is used (Raber & Cannistra, 2005).

The data produced from a LIDAR sensor is a point cloud and is dependent upon the pulse rate. The pulse rate, or the number of light pulses emitted per second, of a sensor is important for resolution. If a sensor has a high pulse rate, then the point cloud is going to be dense resulting in higher-resolution elevation points and a higher-resolution DEM (Carter, et al., 2012).

One of the most important features of a LIDAR sensor is the GPS and Inertial Measuring Units (IMU). Without the GPS and IMU, none of the points would be georeferenced or contain altitude information. A base station GPS is placed on the ground in an area with minimum sky impedance and provides the GPS on the LIDAR with a position correction factor (Carter, et al., 2012).

3.4.4.1 LIDAR Survey Acquisition

The LIDAR sensor used was the LiAir V70 from GreenValley International. The laser scanner is a Livox AVIA that operates at 905 nm with an accuracy of +/- 5 cm. The scan rate of the sensor is 240,000 points per second. The data was processed using the GreenValley LiGeoreference software. Because the survey sites were spatially extensive, the sites were divided into multiple raw point clouds by the sensor. In the LiGeoreference software, boresight corrections were applied to reduce the error between overlapping segments. Trajectory adjustments, elevation adjustments, and removal of outlier points were also applied to the dataset before the point clouds were merged into one. The LIDAR sensor defaults to measuring the data with respect to the WGS84 ellipsoid. Each point cloud was converted to GEOID18 Height by applying a linear transformation to each site, as listed in Table 3.7 based upon NOAA's conversion (Shirley & Hampton, 1978).

Table 3.7 Ellipsoid – GEOID18 Height Conversion

Site	Latitude	Longitude	Geoid Height
1	38.193427	121.619377	-32.020
2	38.209403	121.608318	-31.951
3	38.235563	121.601945	-31.868

DEMs and Digital Surface Models (DSMs) were created using the point clouds from the LIDAR surveys. These models were created using ArcGIS Pro. The DSMs were created using the average elevation data (the first return) and the DEMs were created using the minimum elevation data (the

last return). Figure 3.22 shows an example of the 3D point cloud of Site 2 on Grand Island. The LIDAR point clouds and DEMs for the other sites can be found in Appendix D.



Figure 3.22 3D LIDAR point cloud of Site 2 on Bacon Island.

3.5 Change Detection

Levees are complex earthen structures, and their health will vary depending on seasonal conditions, such as increased temperature or rainfall. These factors are typically experienced during different seasons, so understanding how the system changes with seasonal effects is imperative. Surface and subsurface data was collected on the levees on Grand Island twice. A comparison between the average temperature, precipitation, and wind speed between data collection is shown in Figure 3.23.

The first data collection was in September 2022 during the dry season where the month's average temperature ranged from 61°F to 93°F with an average of 77°F with an average wind speed of 12 kph with a total of 1.5 mm of precipitation for the entire month. The second data collection was in February 2023 during the wet season where the month's average temperature ranged from a low of 38°F to a high of 59°F and an average of 47°F. The average wind speed in February was 12 kph and the precipitation totaled 42 mm over the entirety of the month.

Change detection was analyzed at the subsurface level looking at how Vs changed between the seasons and how ρ changed between the seasons. The change detection of the surface was also analyzed by looking at changes in the optical properties, changes in the thermal properties, changes in surface classification, and changes in elevation, slope, and aspect.

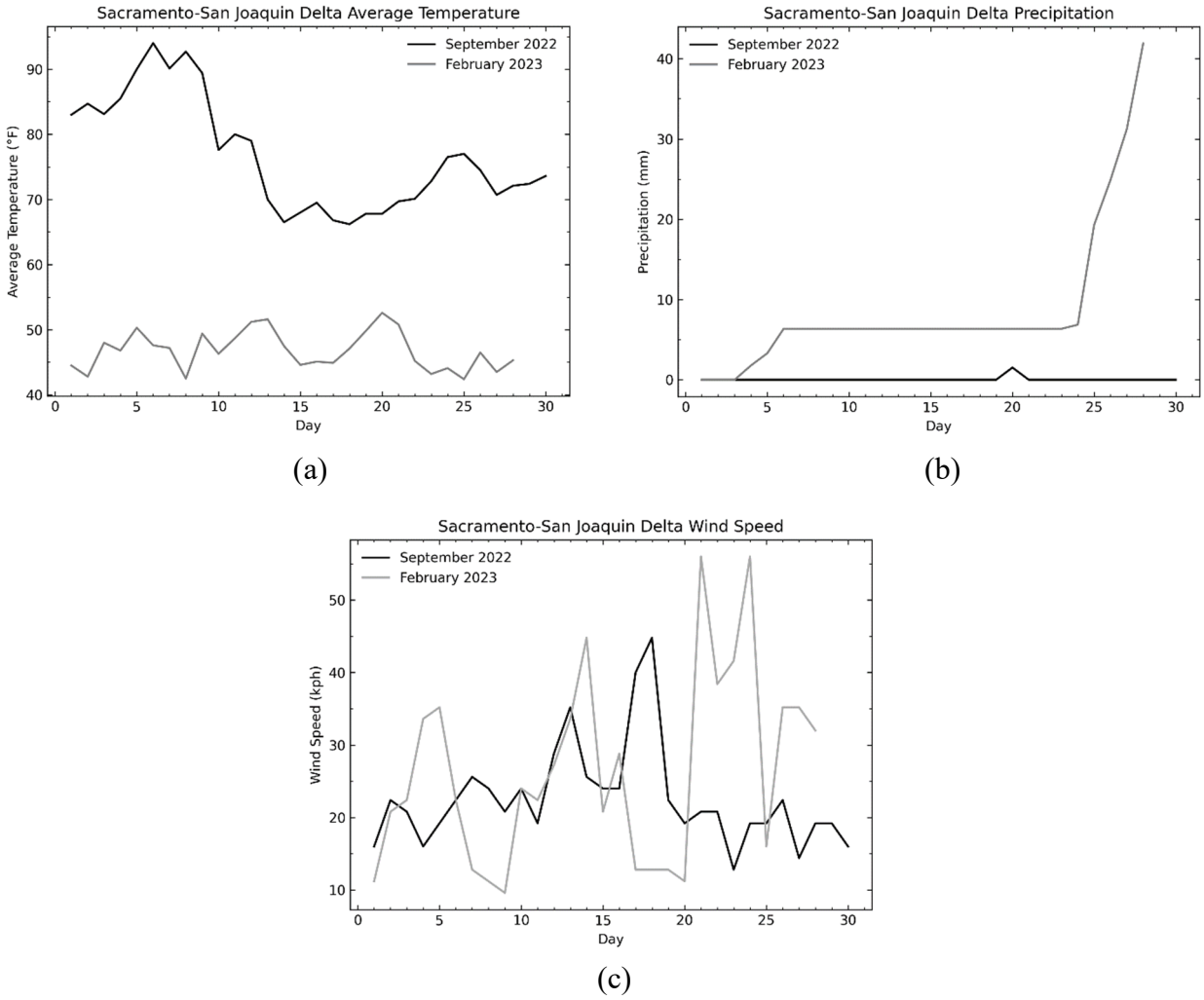


Figure 3.23 Comparison of (a) average temperature, (b) precipitation and (c) wind speed in the Sacramento-San Joaquin Delta during the months of data collection.

3.5.1 Change Detection of the Subsurface

3.5.1.1 Dataset Creation

The first step in analyzing change detection of the subsurface information is to create a dataset that contains the dry season data and the wet season data for a single survey. This was done for all the MASW surveys and all the EMI surveys. Due to dense vegetation, some EMI surveys were not repeated in February and those surveys were not included in the change detection analysis.

To create the dataset, a grid of user-defined points was created with known elevations, x-coordinates, and y-coordinates. The lateral limits and elevational limits were defined by where the wet and dry surveys overlapped. Data that did not overlap between the seasons were excluded. For

the EMI data, before the elevational limits were determined, a skin depth cutoff was applied to remove data at the depth that was not sensitive to the inverse modeling.

With a grid of known points created, the *Vs* or ρ for both seasons was interpreted to each point using the nearest-neighbors algorithm. This resulted in a dataset where each point had a *Vs* or ρ for both seasons which allowed for direct comparison between profiles. Refer to Appendix E for all the MASW and EMI profiles used in the change detection analysis.

3.5.1.2 Analyzing Change in the Subsurface

To look at how the subsurface changed between seasons, the percent change in ρ or *Vs* was calculated for each point in the dataset. When calculating the percent difference, values were positive when data in the dry season had higher values than in the wet season, or negative when data in the dry season had lower values in the dry season. It was important to keep the sign in the analysis as it provides valuable insights into the seasonal variations of the subsurface. This approach goes beyond merely exploring the overall change and enables the understanding of how the subsurface changes throughout different seasons.

For both *Vs* and ρ, it was decided that if the percent change was within 10% between the two seasons, then this signified that there was no meaningful change. With the measurement and processing of geophysical data, there is inherent uncertainty in the values due to inversion assumptions. Inversions produce a non-uniqueness solution to the velocity model. Using a threshold of 10% change accounts for some of the uncertainty in processing.

The percent change that ranged between -10% and 10% were grouped to create a category of no significant change. Anything above a +/- 10% difference indicated a change in the subsurface. This was split into two categories, less than -10% and greater than 10%, to represent seasonal changes. These categories can be represented spatially, as shown in Figure 3.24.

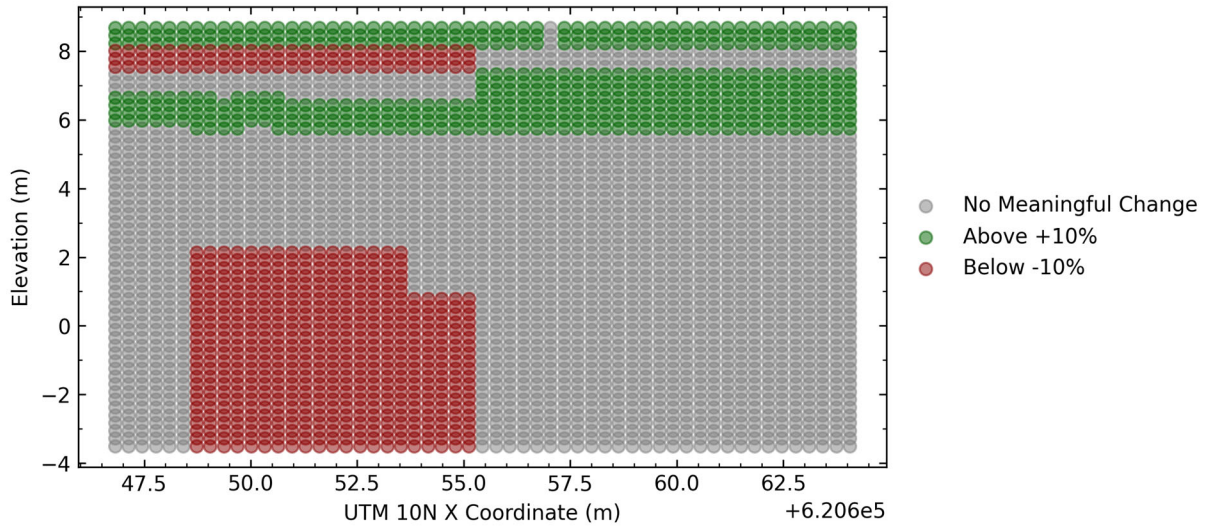


Figure 3.24 MASW subsurface profile showing change detection results with the grey points representing no meaningful change, the green points representing areas with above 10% difference, and the red points representing areas with below -10% difference.

Figure 3.24. shows a change detection subsurface profile of a MASW survey with respect to elevation and the x-coordinate. The grey points represent the areas where there is no significant change. The green points represent areas where the percent change was greater than 10%. These are areas where the dry season had a significant increase in *Vs* compared to the wet season. The red points represent where the percent change was less than -10%. In these areas, the dry season had a significant decrease in *Vs* compared to the wet season. For all subsurface change detection profiles, please refer to Appendix E.

3.5.2 Change Detection of the Surface

The change detection analysis for the surface of the levees was performed on three datasets, the optical imagery, the thermal imagery, and the LIDAR point cloud for each site on Grand Island. The change detection analysis aimed to understand the seasonal effects on the levee system on surface conditions. The change detection workflow used in this analysis was modeled after the workflow proposed by Lu, Mausel, Brondizio, & Moran, (2004) and is shown in Figure 3.25.

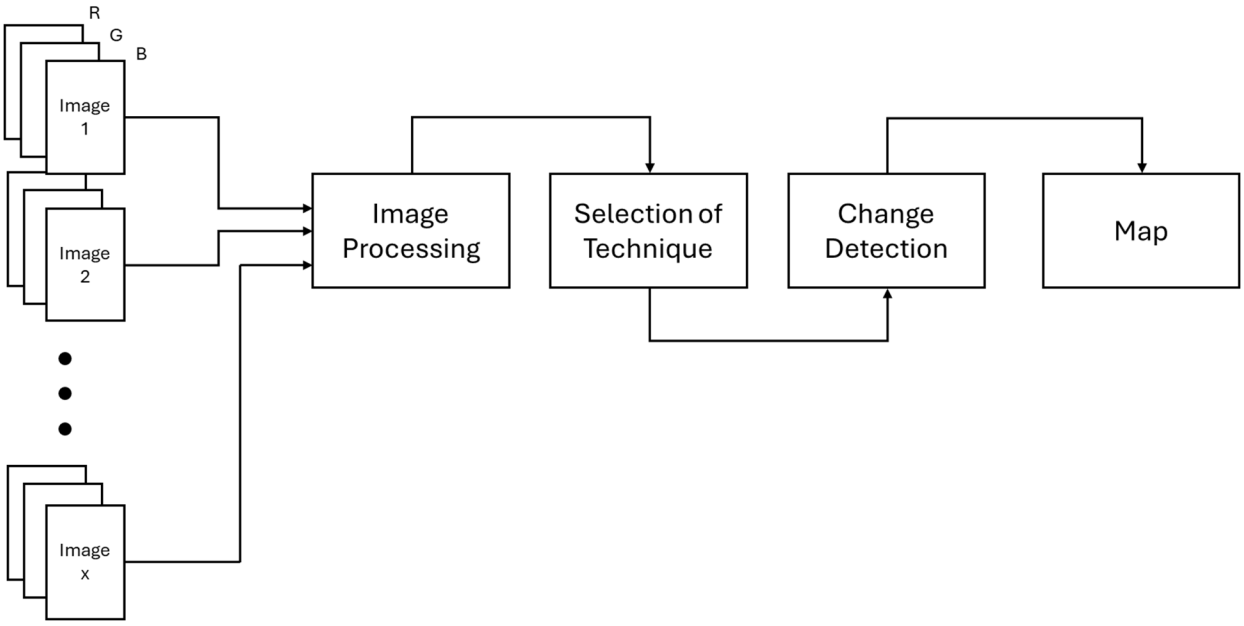


Figure 3.25 Change Detection Workflow (Lu, Mausel, Brondizio, & Moran, 2004).

The first step in the workflow was image processing. This included creating orthomosaics as well as radiometric corrections and geometric corrections. Geometric corrections are needed when images of the same geographic location do not line up correctly (Ban & Yousif, 2016). This is important with change detection as changes in location will significantly skew the results. Ways to remedy this were through ground control points, as discussed in Section 3.4.1.

Radiometric corrections need to be applied when images of the same objects with no change have different brightness values (Ban & Yousif, 2016). Causes for changes in brightness values could be seasonal, meteorological conditions, and changes in solar angles. To check if radiometric corrections were needed, the histograms of each spectral band can be plotted. If the darkest pixel

in the band does not have a brightness of zero, then radiometric corrections need to be applied. This can be done by subtracting the minimum value from each band or through image-to-image normalization, which is a linear model that calibrates the subsequent images to the baseline image (Ban & Yousif, 2016). The histograms of each red, green, and blue band for each optical image at Grand Island can be found in Appendix D. Each band started at zero, indicating that no radiometric corrections were necessary.

The next step in the workflow was to select the change detection technique. There are two major categories of techniques, the first is pixel-based change detection and the second is object-based change detection. Pixel-based change detection compares individual pixels while object-based change detection creates objects through image segmentation where an object is a group of pixels (Walter, 2004). Object-based change detection compares the mean brightness of each object, as well as shape and size.

This study used pixel-based change detection instead of object-based change detection. With pixel-based change detection, multiple techniques can be used including algebraic methods, transformation methods such as PCA, Tasselled Cap, Gramm-Schmidt, and Chi-square, as well as classification methods. This analysis focused on algebraic methods and classification methods (Lu, Mausel, Brondizio, & Moran, 2004).

Algebraic methods were applied to both optical and thermal imagery using image differencing and index differencing. Equation 3.23 was applied to optical and thermal imagery in ArcGIS Pro using the raster calculator to calculate the difference in pixel brightness and temperature between the dry and wet seasons.

$$\text{Difference} = \text{Image}_{\text{Dry}} - \text{Image}_{\text{Wet}} \tag{3.23}$$

Where $\text{Image}_{\text{Dry}}$ represents the thermal, optical, or index related to the dry season and $\text{Imagey}_{\text{Wet}}$ represents the thermal, optical, or index of the same location related to the wet season.

The final steps of the workflow are to perform the change detection analysis using the chosen technique and then to create maps showing locations where significant change has occurred.

3.5.2.1 Classification Change Detection

Classification change detection first classifies both images and then uses supervised machine learning or unsupervised machine learning to classify pixels. Once both images are classified, the classification is compared to see what changed. Classification is a popular method for change detection as effects from atmospheric conditions and environmental conditions have less of an impact on the change detection results (Asokan & Anitha, 2019). One of the major downfalls of using classification for change detection though is that additional bias from classifying the images can affect the change detection results.

The classification change detection technique used was the supervised method, Random Forest in ArcGIS Pro. When performing the analysis, the maximum number of trees was set to 100 with a maximum tree depth of 30 and a maximum number of samples per class at 1000. Before the data was passed through the Random Forest algorithm, all the optical images were resampled to match the spatial resolution of the thermal images at the site.

Since Random Forest is a supervised machine learning algorithm, training samples were needed to classify each image. A single schema was created for the training samples with the different classes chosen based on known land covers at the sites. The classes included water, trees, urban, dead grass, green grass, and shadows. Each optical image used the same training schema, but training samples were created for each image at each site to increase the accuracy of the classification results. Polygons were manually created for each training set using visuals from the images and field notes to determine each class. Great care went into ensuring that a minimum of 10N (with N being the number of bands) training pixels per class were created. The red, green, and blue bands were used for classifying the optical images.

Accuracy Assessment of Classification

Each image was classified separately and then the accuracy of each classification was analyzed. To analyze the accuracy of each classified image, points were randomly created at each site. The land use for each point was manually labelled based on the optical imagery for that season. Once the accuracy points were labelled, the labelled land use was compared to the land use classified by the algorithm. Tables 3.8 through 3.13 summarize the accuracies of each land use for each site on Grand Island for each season.

Table 3.8 Site 1 Dry Season Classification Accuracy

		Reference Data						
		Water	Trees	Urban	Dead Grass	Green Grass	Shadows	User's Accuracy
Classified Data	Water	45	0	3	0	1	3	0.87
	Trees	3	51	2	29	24	18	0.40
	Urban	0	3	66	37	13	5	0.53
	Dead Grass	1	2	9	84	8	0	0.81
	Green Grass	4	8	2	4	39	0	0.68
	Shadows	1	2	3	1	0	29	0.81
	Producer's Accuracy	0.83	0.77	0.78	0.54	0.46	0.53	*Overall Accuracy: 0.63*

Table 3.9 Site 2 Dry Season Classification Accuracy

		Reference Data						
		Water	Trees	Urban	Dead Grass	Green Grass	Shadows	User's Accuracy
Classified Data	Water	6	2	0	0	2	0	0.60
	Trees	0	47	0	15	24	6	0.51
	Urban	0	0	16	3	0	0	0.84
	Dead Grass	0	2	2	99	0	1	0.95
	Green Grass	0	15	0	9	23	0	0.49
	Shadows	0	4	0	1	1	24	0.80
	Producer's Accuracy	1.00	0.67	0.89	0.78	0.46	0.77	*Overall Accuracy: 0.71*

Table 3.10 Site 3 Dry Season Classification Accuracy

		Reference Data						
		Water	Trees	Urban	Dead Grass	Green Grass	Shadows	User's Accuracy
Classified Data	Water	9	0	2	0	0	2	0.69
	Trees	1	62	1	14	16	22	0.53
	Urban	1	4	18	7	1	2	0.55
	Dead Grass	0	7	2	41	1	2	0.77
	Green Grass	0	7	0	1	2	0	0.20
	Shadows	7	4	1	0	0	20	0.63
	Producer's Accuracy	0.50	0.74	0.75	0.65	0.10	0.42	*Overall Accuracy: 0.59*

Table 3.11 Site 1 Wet Season Classification Accuracy

		Reference Data						
		Water	Trees	Urban	Dead Grass	Green Grass	Shadows	User's Accuracy
Classified Data	Water	52	1	0	10	3	0	0.79
	Trees	1	27	2	33	27	0	0.30
	Urban	6	2	67	28	0	0	0.65
	Dead Grass	1	8	6	96	12	0	0.78
	Green Grass	0	4	0	9	61	0	0.82
	Shadows	14	7	3	17	3	0	0
	Producer's Accuracy	0.70	0.55	0.86	0.50	0.58	0	*Overall Accuracy: 0.61*

Table 3.12 Site 2 Wet Season Classification Accuracy

		Reference Data						
		Water	Trees	Urban	Dead Grass	Green Grass	Shadows	User's Accuracy
Classified Data	Water	8	0	0	2	2	0	0.67
	Trees	1	9	0	22	23	4	0.15
	Urban	0	0	16	4	0	0	0.80
	Dead Grass	0	0	7	92	4	7	0.84
	Green Grass	0	2	0	5	73	0	0.91
	Shadows	0	0	0	4	0	14	0.78
	Producer's Accuracy	0.89	0.82	0.70	0.71	0.72	0.56	*Overall Accuracy: 0.71*

Table 3.13 Site 3 Wet Season Classification Accuracy

		Reference Data						
		Water	Trees	Urban	Dead Grass	Green Grass	Shadows	User's Accuracy
Classified Data	Water	11	1	0	0	5	0	0.65
	Trees	0	0	0	0	0	0	0.00
	Urban	2	0	19	16	1	0	0.50
	Dead Grass	2	0	3	91	9	4	0.83
	Green Grass	2	4	0	7	47	1	0.77
	Shadows	0	0	2	6	1	16	0.64
	Producer's Accuracy	0.65	0.00	0.79	0.76	0.75	0.76	*Overall Accuracy: 0.74*

Overall, the classification performed on Site 3 during the wet season had the highest accuracy of 74% while the classification on Site 3 during the dry season had the lowest accuracy of 59%. Overall accuracy was affected mostly by the vegetation land use classes. At Site 1, Site 2, and Site 3 during the dry season, the green grass land use class was the most incorrectly classified land use class. Due to the similar optical properties of lush trees and grass, tree pixels were incorrectly classified as green grass the most.

The shadows land use class was also incorrectly classified often as green grass or trees. The shadows that were present at the site were typically caused by large trees. As a result, many locations with shadows were associated with trees and were cast over grass, causing the algorithms to struggle to differentiate between the shadow and vegetation. This was decided not to be problematic as the change detection analysis did not differentiate between types of vegetation (i.e. green grass versus trees) and instead identified areas were vegetation went from healthy to dead.

The water and urban land use classes were typically classified as the best with the fewest misclassified pixels because they were the most optically different at the sites. When the urban land class was misclassified, it was typically misclassified as dead grass. At many of the sites, the misclassification of urban for dead grass happened along the sides of roads where there was a transition between asphalt/gravel to dead grass.

Change Detection

Once each site during each season was classified, the change detection between the two seasons was performed. A categorical change detection analysis was performed on the classification data using the categorical change detection tool in ArcGIS Pro to see where major changes occurred in land cover. The results of the change detection were exported and formatted into tables to understand which classes stayed the same and which were classified as something else, as shown in Tables 3.14, 3.15, and 3.16.

Table 3.14 Site 1 Classification Change Detection

Class	% Classified in the Dry Season	% Classified in the Wet Season	% Changed from Dry to Wet Season	% Same from Dry to Wet Season
Water	10	13	21	79
Trees	25	18	74	26
Urban	25	21	57	43
Dead Grass	21	25	61	39
Green Grass	11	15	80	20
Shadows	7	9	75	25

Table 3.15 Site 2 Classification Change Detection

Class	% Classified in the Dry Season	% Classified in the Wet Season	% Changed from Dry to Wet Season	% Same from Dry to Wet Season
Water	3	4	52	48
Trees	31	20	82	18
Urban	6	7	28	72
Dead Grass	35	37	78	22
Green Grass	16	27	83	17
Shadows	10	6	82	18

Table 3.16 Site 3 Classification Change Detection

Class	% Classified in the Dry Season	% Classified in the Wet Season	% Changed from Dry to Wet Season	% Same from Dry to Wet Season
Water	5	7	45	55
Trees	46	0	100	0
Urban	13	15	54	46
Dead Grass	21	44	62	38
Green Grass	2	24	81	19
Shadows	13	10	92	8

The classification change detection for Site 1 showed that 38% of all the pixels did not change between seasons. In the dry season, trees and urban were classified the most with 25% of the pixels with shadows the least at 7% of pixels. In the wet season though, urban was still classified highly with 21% of the pixels but dead grass increased to have the highest number of pixels at 25% while trees were at 18% in the wet season. The greatest increase of change between the dry and wet seasons was green grass at 80% and shadows at 75%. Water stayed the same the most between the two seasons with 79% of the pixels staying the same between the dry and wet seasons.

For Site 2, overall, 23% of the pixels stayed the same between the dry and wet season. Water and urban land covers had the fewest pixels at 3% and 6% in the dry season with 4% and 7% in the wet season. In the dry season, the dead grass and trees contained the most pixels at 35% and 31%, respectively. In the wet season, the classes with the most pixels were dead grass and green grass at 37% and 27%, respectively. For Site 2, all the classes had significant changes in classification

between the dry and wet seasons with urban and water being the least at 28% and 52%. Trees, dead grass, green grass, and shadows had 78-83% of their pixels change between seasons.

Site 3 had 18% of the overall pixels stay the same between the dry and wet seasons. In the dry season, only 2% of pixels were classified as green grass and 5% of pixels were classified as water. The highest land cover in the dry season was trees with 46% of pixels followed by dead grass at 21% of pixels. With the change to the wet season, no pixels were classified as trees anymore, but the number of green grass pixels had significantly increased to 24% of the pixels. Most of the water and urban pixels remained the same while very few of the shadows and green grass pixels stayed the same.

In terms of levee health, the most significant classification changes are those concerning vegetation. Vegetation on levees can both be beneficial and harmful to the system. The benefits of vegetation include slope stability and soil reinforcement from dense root systems as well as the reduction of impact on slope hydrology (Gray & Megahan, 1981; Stokes, et al., 2008). The root systems of grasses and trees protect surface soil from erosion by binding the soil together through an apparent cohesion, increasing the overall strength of the soil, through a modified Mohr-Coulomb Failure Criterion shown in Equation 3.24 (Gray & Leiser, 1982):

$$s = c' + c'_R + (\sigma - u)\tan\phi' \tag{3.24}$$

where s is the shear strength, c' is the cohesion, c'_R is the apparent cohesion from the root system, σ is the normal stress, u is the pore-water pressure, and ϕ' is the effective friction angle. This additional cohesion causes the vegetation to act as a protective layer for the soil.

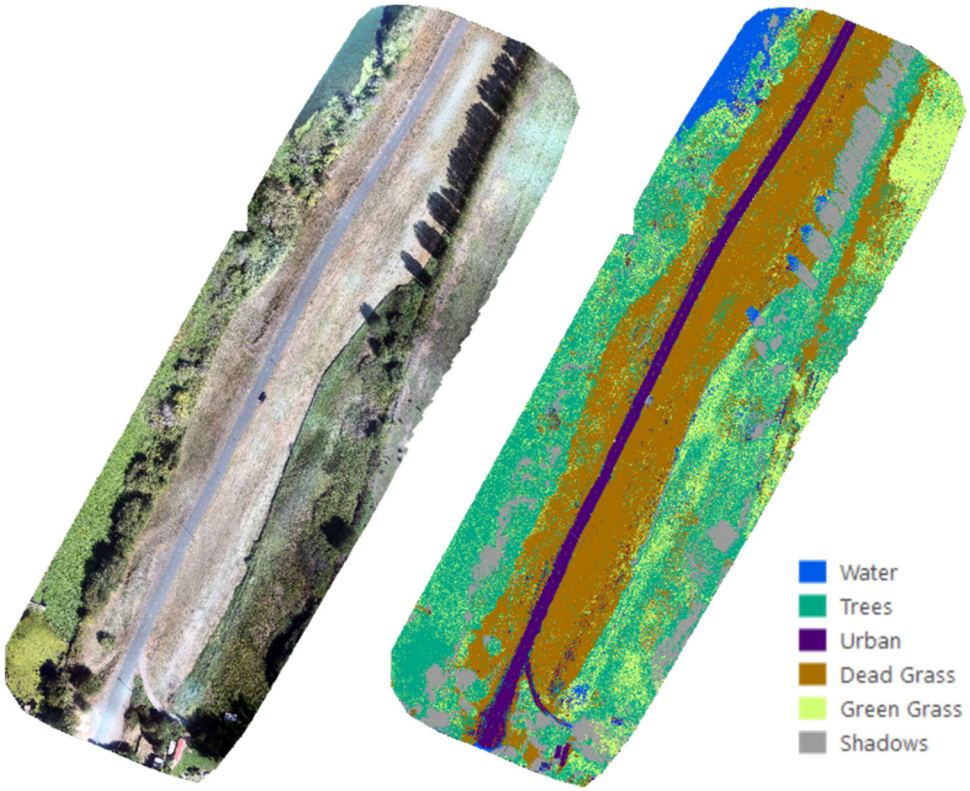
The vegetated protective layer is also crucial for reducing the harmful effects of slope hydrology. Without vegetation during rainfall, the rain will penetrate the soil and increase the pore water pressure in the subsurface pushing the slope towards failure as well as actively eroding the soil sitting on top of the slope (Norris, et al., 2008). Vegetation on the slopes reduces the kinetic energy of falling rain which in return reduces the overland velocity resulting in less surface erosion. As rain falls, canopy covers can intercept rain causing rain to collect on leaves instead of falling to the ground. This results in less infiltration since the water collected on the canopy evaporates instead, leading to decreased moisture content in the soil (Stokes, et al., 2008).

The levees in the Sacramento-San Joaquin Delta are continuously loaded by water. The wet toes and slopes of these levees are continuously being impacted by water currents. Shrubs and trees have been found to act as barriers against the erosive potential of water currents (USACE, 2019). This becomes highly beneficial for levees like those in the Sacramento-San Joaquin Delta that are constantly being eroded by water currents and wave action.

While vegetation can be highly beneficial to the strength of a levee system, it can also pose some challenges and negative impacts on health as well. Roots were described previously as increasing the strength of the soil, but they can also lead to piping. With extensive root systems that have decayed, conduits for preferred water flow develop through the internal structure of the levee, which could lead to soil being eroded internally. Vegetation can also harm levees during heavy rainstorms or hurricanes. Large gusts of winds can overturn trees causing structural damage in the levee in the form of holes. If the tree is large enough, it can cause water to breach the structure during high water situations, endangering the people living on the other side of the levee (USACE, 2019).

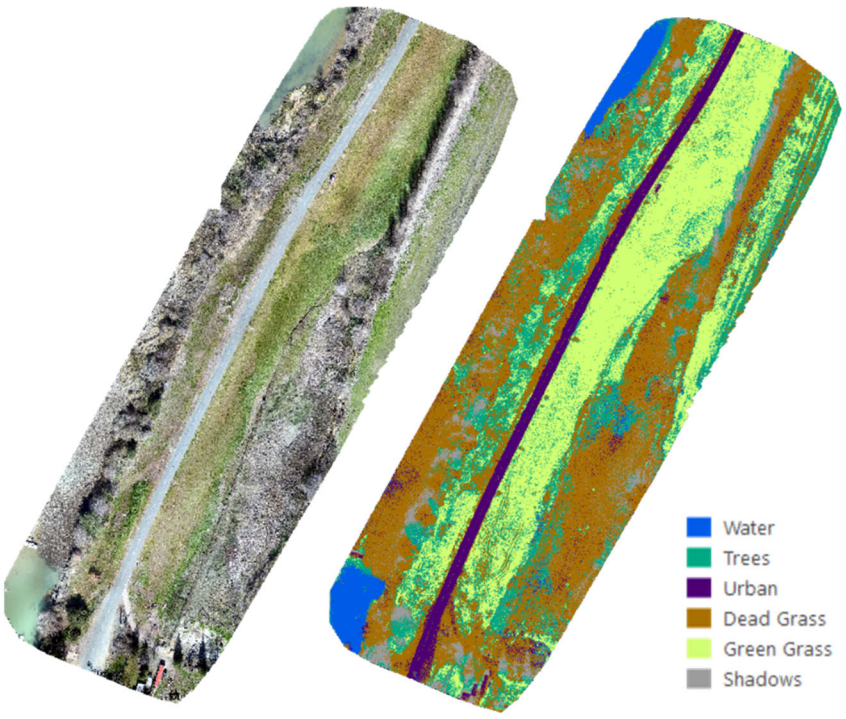
The influence of vegetation not only poses challenges to the stability of infrastructure but also more importantly impacts various aspects of operational efficiency and emergency response. The USACE requires a vegetation-free zone around levees to ensure access to the structure for personnel and equipment for inspection, maintenance, and monitoring (USACE, 2019). With current levee inspection consisting of mostly visual practices, having a clear view of the condition of the levee is imperative for locating developing issues. Having this vegetation-free zone is also important for flood-fighting. The quicker emergency personnel can repair a breached section of levee during a flooding situation, the fewer lives lost, and the less damage caused.

Because vegetation is important both structurally and operationally to levees, the classification change detection focused on how vegetation changed between the dry and wet seasons. First, the classification looked at where no change occurred among all classes, where vegetation went from a dead state in the dry season to a green state in the wet season, and where vegetation went from a green state in the dry season to a dead state in the wet season. An example of the classification results for Site 2 at Grand Island is shown in Figure 3.26 showing the dry season, Figure 3.27 showing the classification for the wet season, and Figure 3.28 showing the change detection results. For the classification results for the other sites, please refer to Appendix E.



(a) (b)

Figure 3.26 Classification of Site 2 at Grand Island during the Dry Season showing (a) the optical imagery and (b) the classification results.



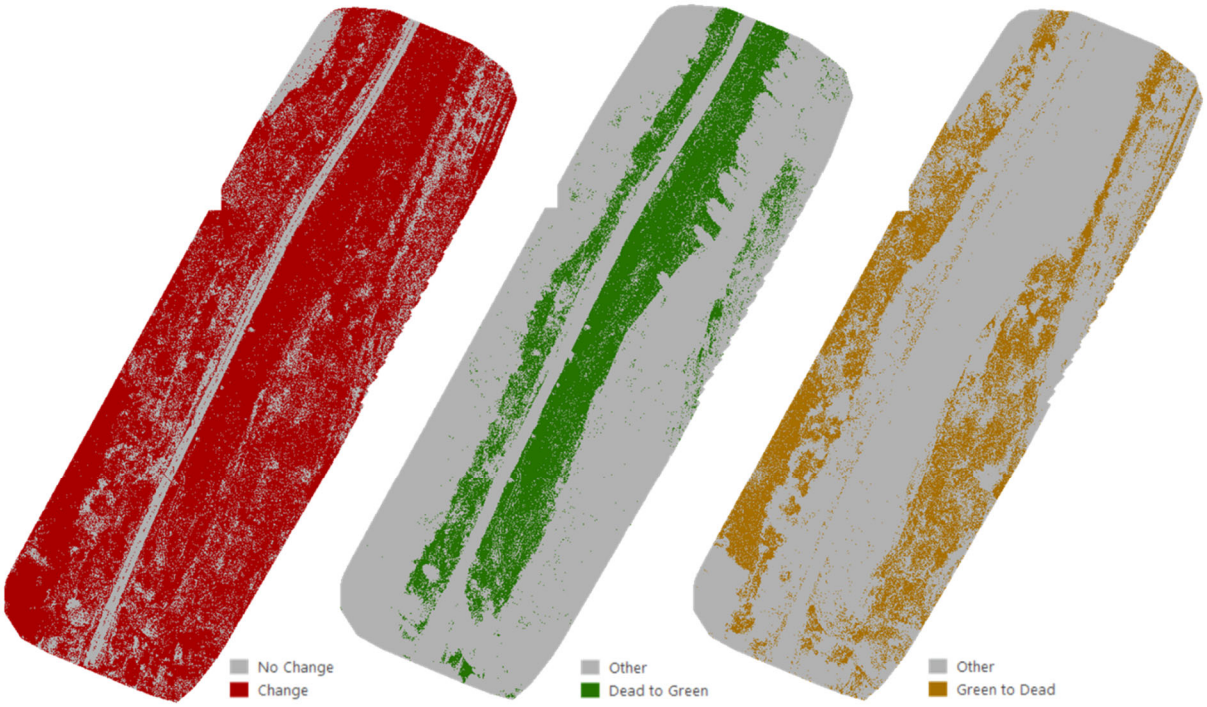
(a) (b)

Figure 3.27 Classification of Site 2 at Grand Island during the Wet Season showing (a) the optical imagery and (b) the classification results.

In the dry season shown in Figure 3.26, the slopes of the levee were dry and were correctly classified as dead grass while trees and green grass can be seen in the optical imagery on either toe of the levee and were correctly classified. In the wet season shown in Figure 3.27, the classification was almost opposite where green grass was visible on the slopes of the levee in the optical imagery and correctly classified with dead grass and trees on either toe of the levee.

The classification results from Figure 3.26(b) and Figure 3.27(b) were compared to spatially determine where no change occurred as shown in Figure 3.28(a). Pixels where the classification changed are shown in red while pixels where the land cover did not change are shown in grey. Most of the pixels that showed no change between the seasons were the urban areas such as the road along the centerline of the levee as well as the water at the top left corner of the image

More importantly, looking at where vegetation changed between seasons is shown in Figure 3.28(b) and Figure 3.28(c). Figure 3.28(b) highlights green pixels where the vegetation was classified as dead in the dry season and reclassified to green in the wet season. The green vegetation includes pixels that were classified as green grass or trees. This change occurred predominantly on both slopes of the levee. This indicates that during the dry season, the levee is more accessible to personnel for maintenance, inspection, and flood-fighting, but less protected from erosion and will have reduced slope stability. During the dry season though, the effect of rainfall on the system is not significant enough to cause concern for stability due to the minimal rainfall.



(a) (b) (c)

Figure 3.28 Classification Change Detection Results for Site 2 on Grand Island showing (a) where no change occurred, (b) where vegetation was classified as dead in the dry season to green in the wet season, and (c) where vegetation was classified as green in the dry season and dead in the wet season.

Figure 3.28(c) shows pixels in brown where the vegetation was classified as green in the dry season and reclassified as dead in the wet season. Like with the previous analysis, the green includes green grass and trees. This change occurred predominantly at both the wet and dry toe of the levee indicating that there is better access to the toes during the wet season but less access to the slopes as they are covered with lush vegetation. With the lush vegetation comes better protection from heavy rains that can occur during the wet season, but access becomes nearly impossible which is problematic for inspection and emergency response.

3.5.2.2 Optical Change Detection

The optical change detection compared the red, green, and blue bands between the dry season and the wet season for the three sites at Grand Island. The techniques used for these orthomosaics were RBG differencing and index differencing. For all optical change detection results, please refer to Appendix E.

3.5.2.2.1 RGB Differencing

RGB differencing subtracts pixels from the dry and wet seasons as shown in Equation 3.23 by comparing each optical band separately. The red band, green band, and blue band of each optical image were independently differenced to look for changes in the surface condition. Large changes in the red band can be used to determine vegetation health as chlorophyll from plants is strongly absorbed by red light, as shown in Figure 3.29 (Wilkie & Finn, 1996).

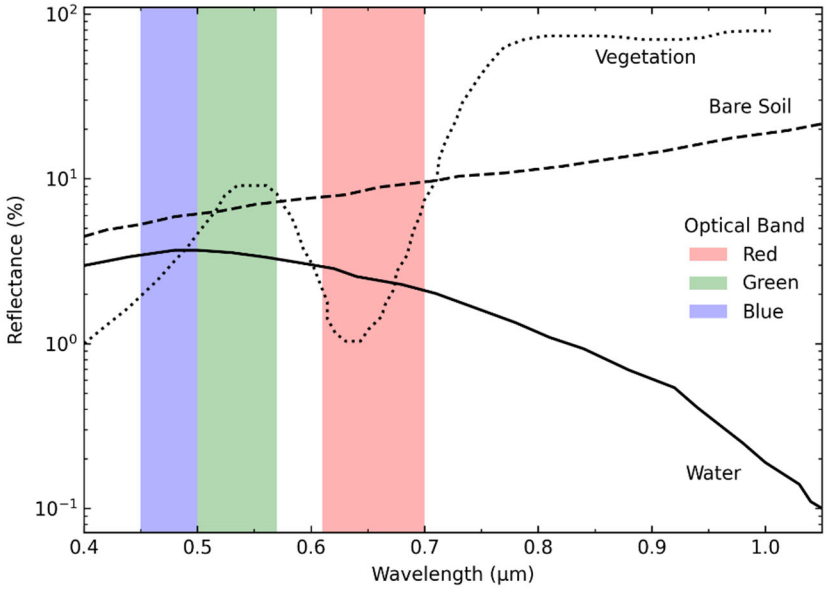
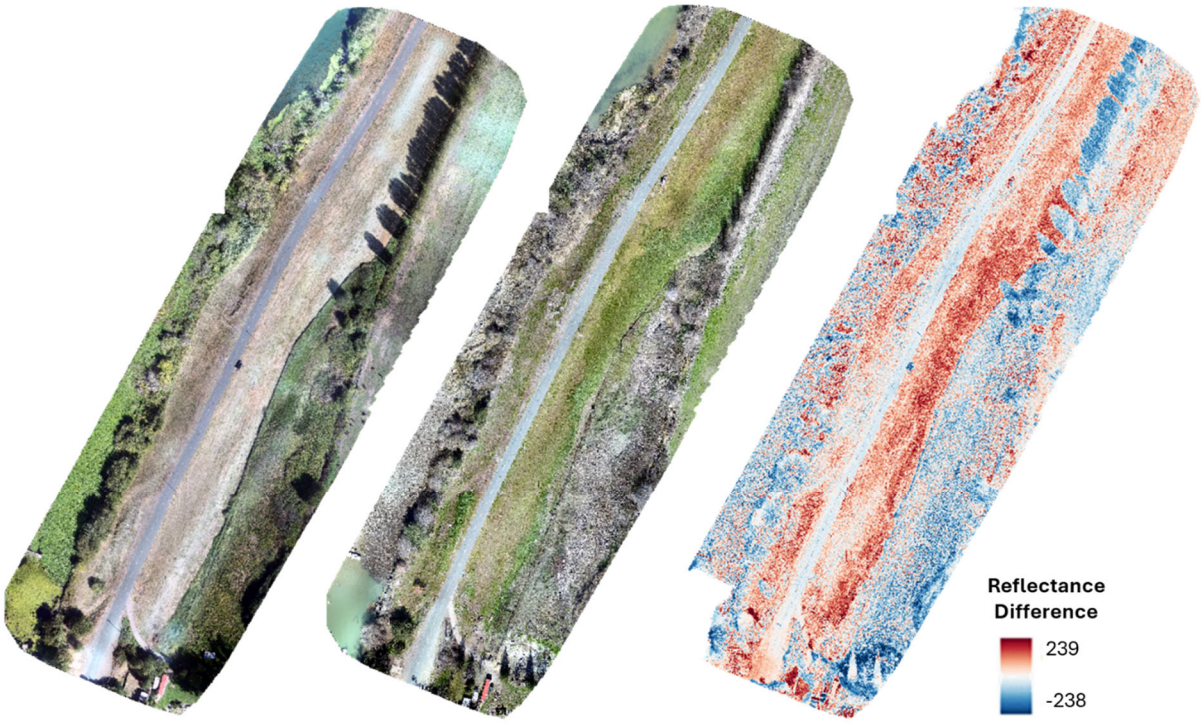


Figure 3.29 Reflectance of vegetation, bare soil, and water with respect to the red, green, and blue optical bands (Modified from Wilkie and Finn, 1996).

With a large positive change, the red band has higher reflectance during the dry season than the wet season, resulting in less lush vegetation in the dry season. With a large negative change, the red band had higher reflectance during the wet season and had less vegetation compared to the dry season. The closer the reflectance change is to zero, the less change there was in the red band between the dry and wet seasons. An example of Site 2 looking at the difference in the red band is shown in Figure 3.30.

The areas shown in red in Figure 3.30(c) had a high positive change in the red band between the wet and dry seasons. It can be seen from Figure 3.30(a) and Figure 3.30(b) that these were areas with little dead vegetation in the dry season and lush vegetation in the wet season. These areas typically occurred on the slopes of the levee. The areas in blue in Figure 3.30(c) had a high negative change in the red band and represent areas with lush vegetation in the dry season and little vegetation in the wet season. This predominantly occurred on either toe of the levee where trees and farmland were lush in the dry season and died in the wet season.

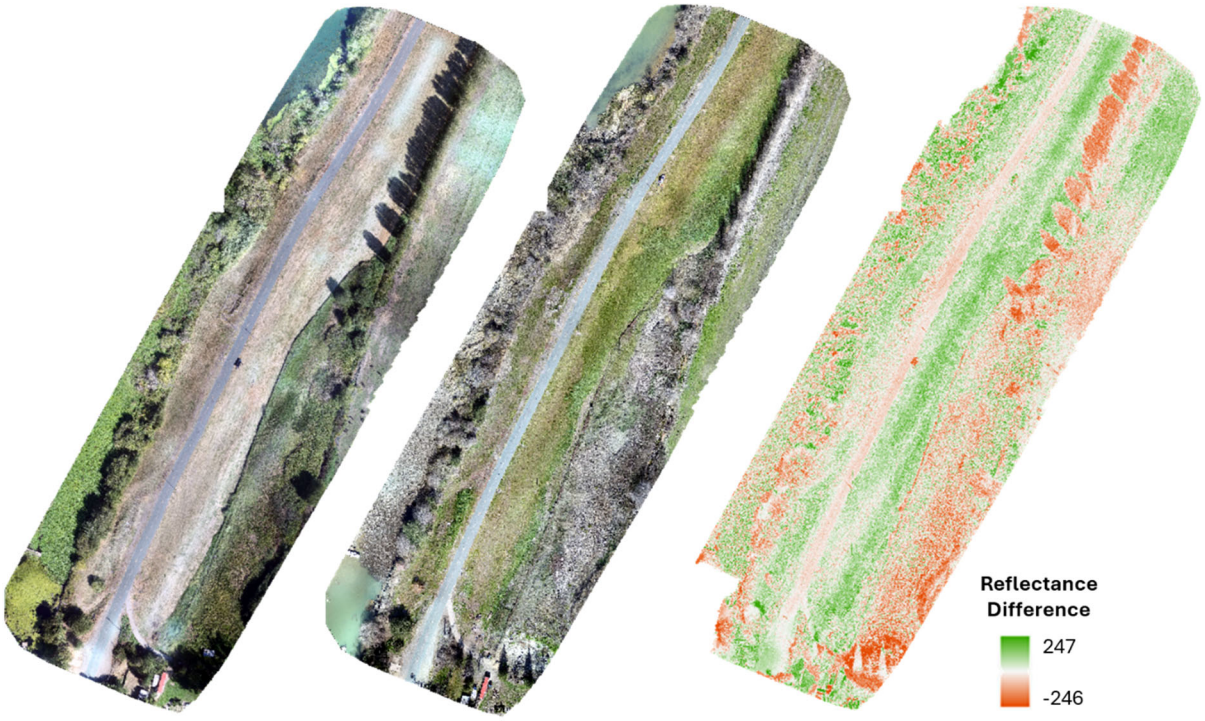


(a) (b) (c)

Figure 3.30 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season, (b) optical imagery from the wet season, and (c) difference in the red band.

In comparison to the red band, the green band can also be used to look at the change in healthy vegetation. While the chlorophyll in plants absorbs the light in the red region, it reflects the light in the green light, as shown in Figure 3.29. With a large positive change, the green band had higher reflectance during the dry season due to more lush vegetation compared to the wet season. With a large negative change, the green band has higher reflectance in the wet season due to lush vegetation. As the reflectance converges to zero, there was less change in the green band between the dry and wet seasons. An example from Site 2 looking at the difference in the green band is shown in Figure 3.31.

The blue band in optical images reflects deep, clear water, making it useful in detecting water bodies. As the water starts to suspend soils though, the reflectance of the water will change according to the nature of the suspended particles. In the infrared wavelengths, clear water bodies appear black as these wavelengths are fully absorbed by the water features. Without the infrared wavelengths, it can be difficult to distinguish water bodies from other features (Wilkie & Finn, 1996).



(a) (b) (c)

Figure 3.31 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season, (b) optical imagery from the wet season, and (c) difference in the green band.

3.5.2.2.2 RGB Index Differencing

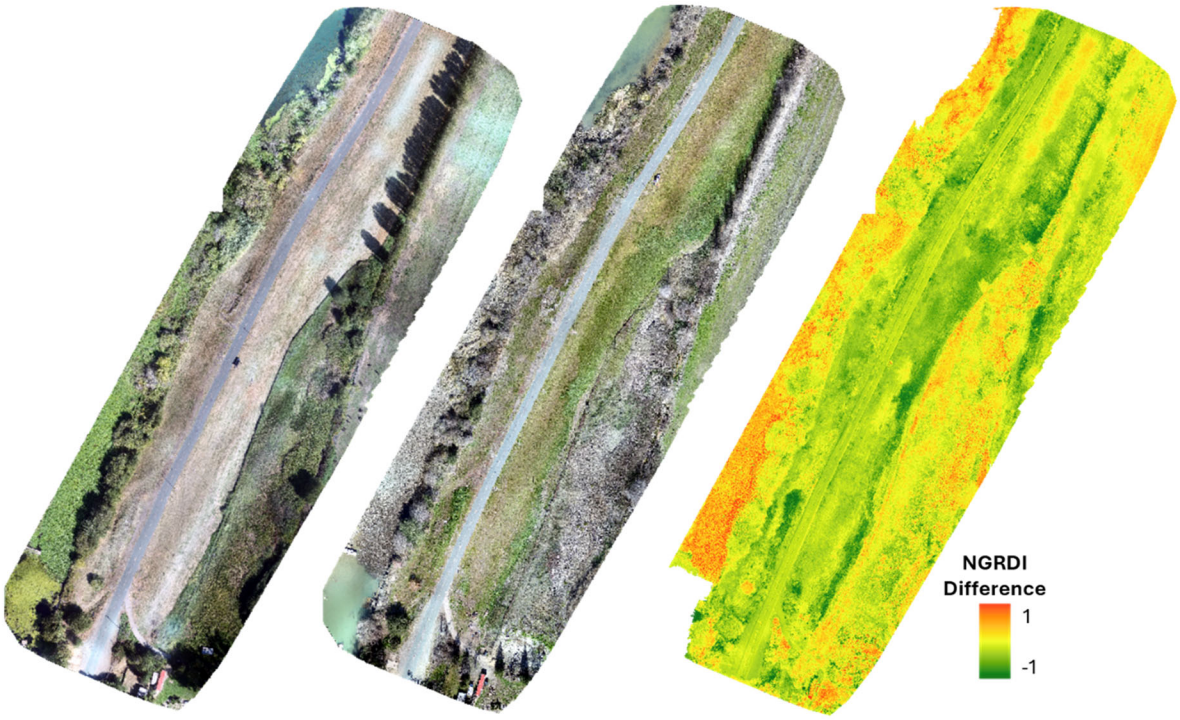
While some bands individually may not give much information about land use, combining multiple bands can be more effective at highlighting surface features as certain land covers absorb and reflect light differently at different wavelengths. Different indices in remote sensing use a combination of two or more bands to detect certain land covers, such as vegetation or water. Change detection can leverage these indices by looking at where an index changes over time. This process is like the RGB differencing, but the index is first calculated and then compared to the same index from the other season.

With only the red, green, and blue bands available to use, the choices of indices are limited for the analysis. Many vegetation indices require the near-infrared band while many water and bare earth indices require the short-wave infrared band. The Normalized Green-Red Difference Index is a vegetation index that only requires the red and green bands, as shown in Equation 3.25 (Zakaluk & Sri Ranjan, 2008):

$$\text{NGRDI} = \frac{\text{Gr}-\text{Rd}}{\text{Gr}+\text{Rd}} \tag{3.25}$$

Where Gr is the digital number of the green band and Rd is the digital number of the red band. This equation was applied to the optical images from the dry season and the optical images from the wet season. Then, using Equation 3.23, the wet and dry NRGDI were differenced to look at where vegetation changed.

The NGRDI ranges between 0 and 1 where 0 represents little vegetation and 1 represents lush vegetation. When subtracted, the NGRDI difference can range from -1 to 1 where -1 indicates that the dry season had little vegetation and the wet season had a lot of vegetation. With a difference close to 1, this indicates that the dry season had a lot of vegetation and the wet season had little vegetation. An example of this index at Site 2 is shown in Figure 3.32.



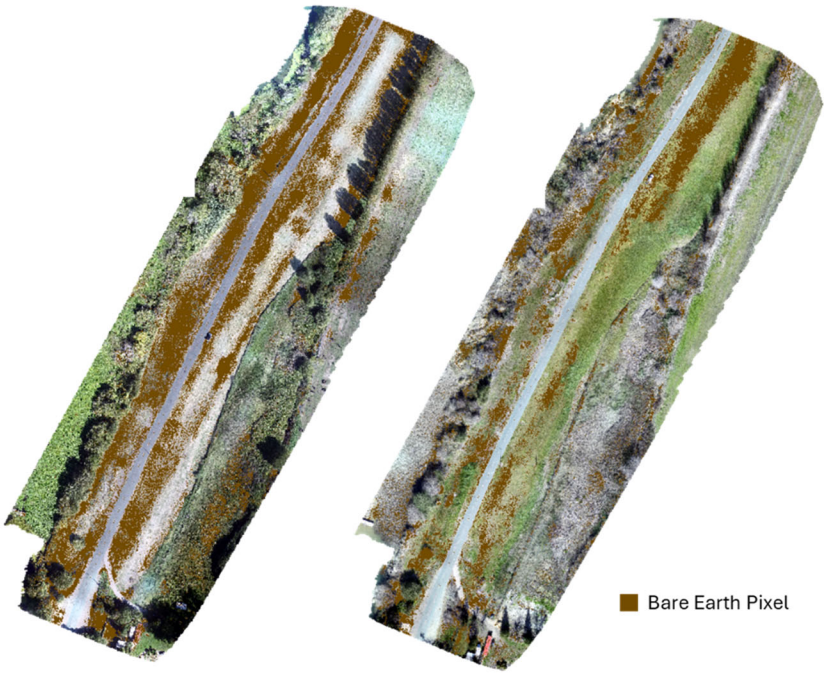
(a) (b) (c)

Figure 3.32 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season, (b) optical imagery from the wet season, and (c) the difference in the Normalized Green-Red Difference Ratio between the dry and wet season.

The green areas in Figure 3.32(c) show where the dry season had little vegetation, but the wet season had a lot of vegetation. This occurred predominantly on the slopes of the levees. The areas in red indicate areas where the vegetation was healthy in the dry season but dead in the wet season. This occurred predominantly on the toes where the trees and farmland were located.

While the blue band by itself is not a useful indicator for any one land cover, in conjunction with the green and red bands, it can be potentially used to identify bare soil and rock. From Figure 3.29, bare soil and rock increase reflectance with increasing wavelength. If an area is shown to increase from the blue to the green band and then the green to red band, that location could be considered bare earth (Wilkie & Finn, 1996). To pull out potential bare earth land cover from the optical bands, the first step was to find where the green band had higher reflectance than the blue band and where the red band had higher reflectance than the green band. Those two raster layers were compared and where the green band was higher than the blue band and where the red band was

higher than the green band, the overlapping areas were considered bare earth, as shown in Figure 3.33.



(a) (b)

Figure 3.33 Example from Site 2 at Grand Island showing (a) optical imagery from the dry season with bare earth locations and (b) optical imagery from the wet season with bare earth locations.

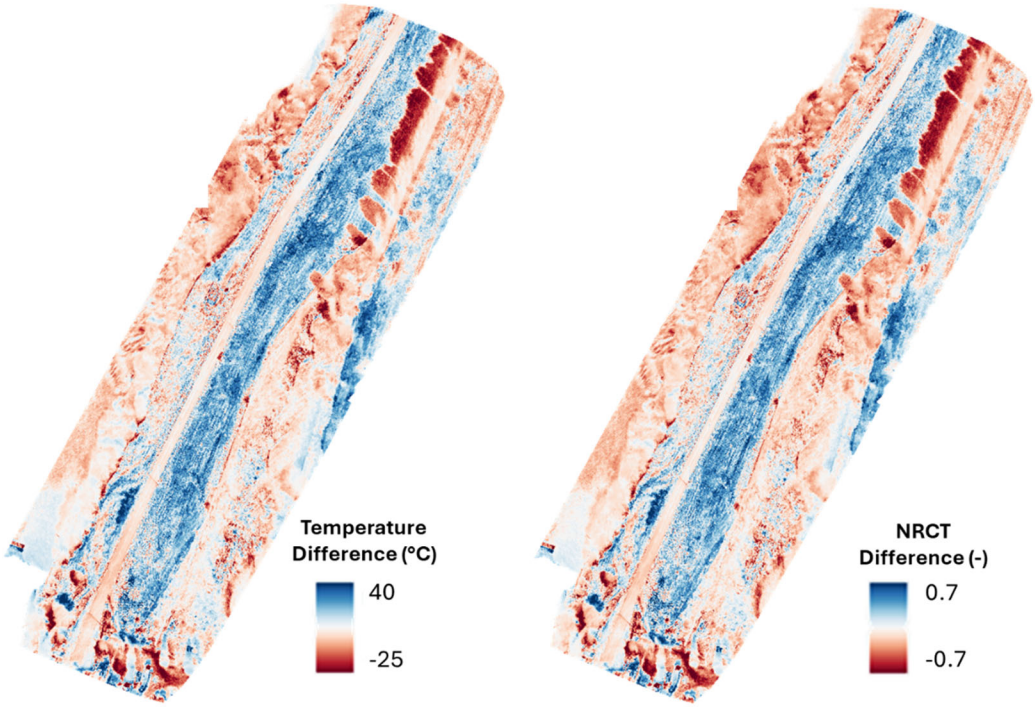
This method was not entirely accurate as many locations that were being classified as bare earth were areas with little or dead vegetation. This method does give an indicator of potentially bare earth locations or locations where the vegetation is much thinner than around it. To increase the accuracy of finding bare earth locations, short-wave infrared bands are required.

3.5.2.3 Thermal Change Detection

Like the RBG imagery, change detection can be performed on the thermal imagery using differencing of the pixels as well as differencing of indices between the dry and wet seasons. Changes in thermal temperature on levees can be attributed to differences in vegetation, soil moisture content, and the presence of near-surface animal burrows (Borgatti, et al., Detection and characterization of animal burrows within river embankments by means of coupled remote sensing and geophysical techniques: Lessons from River Panaro (northern Italy), 2017).

3.5.2.3.1 Thermal Differencing

The thermal images collected contain one band. As a result, when performing a differencing analysis for change detection, only one band is used, unlike the optical images which had three bands. Each pixel in the thermal images contained a value that represented the absolute temperature in °C. Equation 3.23 can be applied to the thermal images from the dry and wet seasons to see how the temperature changed, as shown in Figure 3.34.



(a) (b)

Figure 3.34 Example from Site 2 at Grand Island showing (a) the temperature difference between the dry and wet seasons and (b) the difference in the Normalized Relative Canopy Temperature between the dry and wet seasons.

Figure 3.34 is an example showing Site 2 on Grand Island. In Figure 3.34(a), the blue indicates higher temperatures in the dry season compared to the wet season while the red indicates higher temperatures in the wet season compared to the dry season. The dry season tended to have higher temperatures on the slopes of the levees. This is most likely due to the dead grasses on the slope in the dry season which absorb heat more than lush vegetation—the wet season tended to have higher temperatures on the toes of the levees. Like with the dry season, the toes in the wet season were where dead vegetation was located while the slopes had healthy vegetation.

3.5.2.3.2 Thermal Index Differencing

Index differencing can also be used with thermal images. The Normalized Relative Canopy Temperature (NRCT) is a vegetation index that ranges between 0 and 1 and uses temperature

instead of the blue, green, red, or near-infrared bands. It can be calculated using the minimum and maximum temperature, as shown in Equation 3.26 (Elsayed, et al., 2021):

$$\mathrm{NRCT} = \frac{\mathrm{T}-\mathrm{T_{min}}}{\mathrm{T_{max}}-\mathrm{T_{min}}} \tag{3.26}$$

Where T is the surface temperature, T_{min} is the minimum temperature, and T_{max} is the maximum temperature. This index measures the canopy temperature of lush vegetation which is a function of evaporation and transpiration (Elsayed, et al., 2021). The NRCT had higher values in locations where vegetation was dead, indicating less moisture in these locations, which is supported by the findings of Elsayed, et al. (2021). This method can be used to look at the amount of moisture and locations where lush vegetation occurs.

Change detection using this index was done by calculating the NRCT during the dry season and the wet season. Then, Equation 3.23 was applied to both indices to look for differences between the seasons. An example of the results is shown in Figure 3.29(b) The blue locations represent areas in the dry season with higher NRCT while the red locations represent areas in the wet season with higher NRCT. The blue locations are predominantly on the slope of the levee where the dry season had little moisture indicated by the dead grass while the rest of the structure in the dry season had more moisture indicated by the lush vegetation. The red locations are predominantly on the toes of the levee which had dead vegetation in the wet season as indicated by dead vegetation. During the wet season, other areas were covered with lush vegetation, indicating more moisture, resulting in lower NRCT.

3.5.2.4 LIDAR Change Detection

The detailed point clouds produced from the LIDAR surveys were also used for change detection analyses. LIDAR point clouds are useful for creating detailed 3D geometries of a survey. As a result, the LIDAR datasets were used to look at changes in the elevation, slope, and aspect of the levee. Elevation changes can be useful for looking at depressions in the crest or bulges in the toe. The change in slope and change in aspect are useful for looking for signs of slope failure.

The first step in the change detection analysis when using the LIDAR data was to create DEMs from all the point clouds. This process was done in ArcGIS Pro by first filtering the point cloud to use only the last returns and then feeding that point into the 'LAS Dataset to Raster" tool. Once a DEM was calculated, they were clipped to the geometry of the levee, and trees and structures were clipped out.

To calculate the elevation change, the DEMs from the dry and wet seasons were compared by finding the percent change between the two. If the percent change was within 10%, it was considered there was no significant change to account for uncertainty in the sensor.

To calculate the change in slope, the 'Slope' tool was used in ArcGIS Pro. This tool uses a DEM to calculate the slope and then returns a raster that ranges from 0° to 90°. Lastly, to look at the change in aspect, which is the azimuthal direction the slope is dipping measured from North, the DEMs were inputted into the 'Aspect' tool in ArcGIS Pro. This returns raster layers where pixels range from 000° to 359° where 000° represents North, 090° represents East, 180° represents South, and 270° represents West.

3.5.2.5 Change Detection Discussion

Performing change detection analyses on these different datasets allows for the detection of multiple failure mechanisms to look at levees as a system. Classification change detection allows for the detection between different land uses while thermal change detection can be used to understand where temperature changes at the site which could be evidence of developing failure mechanisms such as animal burrowing and seepage. The LIDAR change detection is a useful tool to understand how the shape of the structure changes over time. It can be used to identify depressions in the crest, bulging at the toe, and stability issues on the slopes. These change detection results will be split into different surface layers showing how the levee as a system is changing between the dry and wet season. These changes will be a component used to evaluate the potential hazard of the structure in conjunction with the subsurface change detection results and an unsupervised machine learning analysis of the subsurface data, which is discussed in Chapter 5. The process of combining the surface change detection layers with the subsurface change detection and machine learning analysis is discussed in Chapter 5.

A limitation of this study is using a three-band optical camera instead of a multispectral camera. With only the red, green, and blue bands, the analysis is limited in what can be analyzed using index differencing. Many indices, including vegetation indices, bare earth, and water indices, require the near-infrared or short-wave infrared bands to be effective for analysis. Some indices have been developed for bare earth and vegetation using only the red, green, and blue bands, but they are not as effective or accurate as when the infrared bands are used.

Even without the near-infrared band, it was found that using the vegetation index was more accurate than looking at the change detection of each band independently. Independently, the red and green bands can be used to look at changes in vegetation, but where there are no significant changes, it can be difficult to interpret the results. Looking at changes in the blue band individually provided no meaningful results and should not be included in the analysis.

The thermal change detection results showed that differing the thermal images produces the same results as first calculating the NRCT and then differing it. There are very minor differences between the two analyses, but the trends are the same. As a result, the use of either analysis is suitable but not both. Using both differences will result in double counting temperature and moisture.

Chapter 4. Machine Learning Analysis

Portions of the results presented in this chapter have previously been published in the following conference proceedings and are in preparation for journal publications:

Brittany M. Russo and Adda Athanasopoulos-Zekkos, "Exploration of Feature Engineering Techniques and Unsupervised Machine Learning Clustering Algorithms for Geophysical Data on Levees". In the Proceedings of *Geo-Congress 2024*. (2024).

Brittany M. Russo, Adda Athanasopoulos-Zekkos, Craig Ulrich, and Sebastian Uhlemann, "Levee Soil Type Classification using Shear Wave Velocity and Electrical Resistivity Measurements". In the Proceedings of *USSD 2024 Annual Conference*. (2024).

Brittany M. Russo and Adda Athanasopoulos-Zekkos, "Unsupervised clustering analysis of crest and toe surveys on levees using geophysical methods." [In Preparation]

Brittany M. Russo and Adda Athanasopoulos-Zekkos, "Combining Clustering and Anomaly Detection Methods for Detection of Problematic Zones in Levee Subsurface." [In Preparation]

4.1 Machine Learning Literature Review

4.1.1 Machine Learning in Geotechnical Engineering

Machine learning has become a popular tool in the last couple of decades for analyzing large amounts of data as it can learn trends and properties from previous datasets. This is a universal tool that has been used for social media, fraud detection, self-driving cars, and speech recognition. As more data emerges in the geotechnical engineering field, the more beneficial machine learning has become for understanding soil properties at a greater spatial scale, structural vulnerability, and hazard susceptibility. Common data inputs for geotechnical engineering to determine soil properties are borehole SPT N-values and laboratory-measured properties. Puri, Prasad, and Jain (2018) used 1,053 SPT N-values as inputs in linear regression, Artificial Neural Networks (ANN), SVM, Random Forest (RF), and M5 Tree to predict the friction angle and cohesion of soil. The authors also measured the liquid limit and void ratio in the laboratory and applied the same

algorithms to try and estimate Cc. This study showed how laboratory-based measurements and field measurements can be used to accurately estimate important design parameters.

There has been work where the SPT N-values were predicted using K-Nearest Neighbors (KNN) (Galupino & Dungca, 2022). The training data was developed from 24 borehole logs drilled in Valenzuela. The borehole logs were discretized into 1,912 points, where each point contained latitude, longitude, depth, and SPT blow count. The study aimed to predict SPT N-values in the City of Valenzuela in the Philippines. This approach worked well for this city, but the data was biased to stiff soils as most of the data had blow counts of 50. Because the latitude and longitude were included in the training data, the model created would not be able to generalize to locations outside of Valenzuela.

Another geotechnical application machine learning has been applied to was assessing liquefaction potential (Goh & Goh, 2007). The dataset included CPT data including cone tip resistance, sleeve friction ratio, effective stress, total stress, as well as maximum horizontal ground surface acceleration, and earthquake moment magnitude. Within the dataset, there were 133 cases of liquefaction and 93 cases of no liquefaction. The data was normalized, and two-thirds of the data was used as training data while the remaining one-third of the data was used as testing data. SVM was used to predict liquefaction with 97% accuracy.

These algorithms can also be used to create an early warning system for geotechnical engineering problems. A study conducted by Chou and Thedja (2016) successfully used SVM to build a model to assess soil liquefaction with 95% accuracy, slope collapse with 95% accuracy, groutability of sandy silt soil with 95% accuracy, and monitoring seismic hazards in coal mines with 94% accuracy.

Machine learning has been applied to pile foundation design. Research carried out by Kardani et al. (2020) used a variety of supervised machine learning algorithms to estimate the bearing capacity of piles. The data was derived from 59 load test records of driven piles in sand. The training dataset included information about the soil shear resistance angle at the shaft of the pile, the soil shear resistance angle at the tip of the pile, the length of the pile, the cross-sectional area of the pile, and the effective stress at the tip of the pile. The data was normalized and then randomly split into 80% training data and 20% testing data. With machine learning, the algorithms and features used to train the algorithm all affect the result. It is important when designing a machine learning analysis to assess how different algorithms perform since all algorithms have inherent assumptions that may interact with the input data differently. This study showed that KNN performed the worst with an R-squared value of 0.731 while Extreme Gradient Boosting performed the best with an R-squared value of 0.9615. Effective stress was also found to be the most influential variable when predicting bearing capacity.

A study performed by Rabby, Hossain, and Abedin (2022) also evaluated three supervised machine learning algorithms, KNN, RF, and Extreme Gradient Boosting, but for landslide susceptibility mapping in Rangamati District, Bangladesh. The input dataset was a landslide inventory dataset containing information regarding the location of landslides, size of landslides, volume of mass moved, number of casualties, the type of landslide, as well as geologic, topographic, hydrologic,

and anthropogenic factors. When examining feature importance, the dip of the slope was the most important factor followed by the distance from the road network and the topographic wetness index. This shows that the slope combined with the road cut and water were the primary causes of landslides. When comparing algorithms, the Extreme Gradient Boosting model performed the best. Performance is not the only metric that should be examined, though. This study showed that Extreme Gradient Boosting overestimated the landslide susceptibility and could not distinguish between high and low susceptibility.

A significant issue with geotechnical data for machine learning analyses is not having enough data to develop an unbiased model. A way to create a larger dataset can be to create synthetic data. For geotechnical engineering, this can be done with the Discrete Element Method and the Finite Element Method (Pirnia, Duhaime, & Manashti, 2018). The Discrete Element Method can be used to generate synthetic photos of soil showing the particle grain size distribution while the Finite Element Method can be used to generate pore pressure time series of a dam with increasing water levels upstream. These datasets can be used as training data for supervised machine learning algorithms like ANN. When using only synthetic images of particle distribution, there was bias in the model since the photos were not fully realistic. The synthetic images did not accurately represent soil angularity or texture and were unable to generalize to real images of soil. This study showed how important it was to have some real images of soil in the training dataset.

While the dataset is the most important component of any machine learning analysis, hyperparameters can significantly impact performance. By not optimizing the hyperparameters, it could cause the model to overfit or underfit the training data, resulting in poor generalization to new data and poor performance. A study performed by Contreras et al. (2021) used RF to understand the sensitivity of runoff forecasting models in the Tomebamba catchment in Ecuador. The input data included precipitation and discharge. This study performed a sensitivity analysis by running the RF algorithm first with the default parameters and then with optimized parameters determined from 3-fold cross-validation and grid-search. The RF hyperparameters included in the study were (from the Scikit Learn library) criterion, max depth, max features, max leaf nodes, max samples, min samples leaf, min samples split, and n estimators. The study found that the model that used the default hyperparameters performed worse than the model that used the optimized hyperparameters. The study also showed that the number of estimators used (number of trees), improved the performance the most compared to the other hyperparameters. As a result, when using RF, the number of estimators should always be optimized instead of using the default value.

4.1.2 Machine Learning and Geophysics

Geophysical data is easy to collect, efficient, non-invasive, and continuous. If data is collected as 2D profiles, large datasets needed for a machine learning analysis can be created easily. If machine learning models for understanding the subsurface conditions can be created using geophysical data as inputs, it will reduce the need for extensive invasive data such as boreholes and CPTs. A study performed by de Mello et al. (2022) used gamma-ray emission from uranium, thorium, and potassium, as well as magnetic susceptibility and Electromagnetic Induction (EMI) as inputs to the Cubist model, SVM, RF, and the Linear Model. The study aimed to compare the supervised

machine learning models to predict soil properties. By combining the geophysical data, soil properties were efficiently predicted.

Subsurface lithology is typically determined through invasive techniques such as boreholes and CPTs, as a result, predicting lithology using machine learning is popular. Predicting lithology is possible through a combination of geophysical techniques as geophysical measurements are directly related to subsurface properties. Certain geophysical methods are sensitive to subsurface properties such as shear wave velocity (*Vs*) with soil stiffness and electrical resistivity (ρ) with soil moisture. Research carried out by Bressan et al. (2020) tried to predict subsurface lithology based on a variety of geophysical data. Offshore wells were used to collect gamma-ray attenuation bulk density, p-wave velocities (*Vp*), magnetic susceptibility, reflectance spectrophotometry, and shock remanent magnetization. The data was randomly split into 70% training data, 20% test data, and 10% validation data for Decision Tree, RF, SVM, and Multilayer Perceptron. This study showed that geophysical data was beneficial for predicting lithology where RF performed best after cross-validation and Multilayer Perceptron performed best before cross-validation.

Lithology was also predicted by Lim et al. (2020) by using a combination of geophysical data with laboratory-measured data. Seismic downhole logging was used to measure the p-wave and s-wave velocities of the subsurface while the moisture content, liquid limit, plastic limit, and specific gravity of the soils were measured in the laboratory. This data was used as inputs for Decision Tree, ANN, and Bayesian Networks. With a small dataset of 151 points, the authors determined this was insufficient to create a model to classify soil type. The Bayesian Network model performed the best with a training accuracy of 100% but a testing accuracy of 34%, indicating that the model was overfitting the training data. Having enough data to avoid overfitting is crucial for machine learning analyses.

When looking only at near-surface lithology to infer 3D geology, Reading et al. (2015) used a combination of geophysical data and remote sensing data as input data for RF. The geophysical data consisted of total magnetic intensity and radiometrics while the remote sensing data consisted of digital elevation models and Landsat imagery. This study showed that 3D geological models can be created by using RF to constrain the geometry of the major lithological units in the subsurface.

A study by Marzan et al. (2021) leveraged *Vp* from well logs and electrical resistivity from ERT. These authors explained their method for feature engineering their geophysical data. It consisted of creating a common grid between the surveys and interpolating velocities and resistivity to each point. A disadvantage of this study was poor coverage of *Vp* near the surface. To perform a machine learning analysis, every point in the dataset is required to have a *Vp* and ρ value. To fill this gap in the dataset, unsupervised Expectation-maximation clustering was applied to estimate the *Vp* at the surface. Then, the dataset was passed through supervised Linear Discriminant Analysis for predicting lithology to create 3D models that were consistent with mapped geological units.

When manually developing subsurface lithological profiles, each point in the profile is not analyzed separately. Instead, a given area is considered when understanding lithological variation. With traditional machine learning, a single point in the profile is analyzed. Work carried out by

Wrona et al. (2018) wanted to use supervised machine learning to create a subsurface profile by analyzing an area instead of each point individually. The authors used seismic reflection data to extract features for inputs into Adaboost, Decision Tree, Gradient Boosting, KNN, Logistic Regression, Neural Networks, RF, stochastic Gradient Descent, and SVM. In parallel, the authors also manually classified the subsurface. The machine learning models were altered to look at a window around each point to mimic manually classifying the data point. By doing this, the supervised models performed well in classifying lithology with accuracies of 98% compared to the manual classification.

Geophysical data can be used to classify more than subsurface lithology. A study conducted by Uhlemann et al. (2022) wanted to use regression analyses to predict the ρ variation across a watershed. The dataset included nuclear magnetic geophysical borehole data, airborne EMI data, Electrical Resistivity Tomography (ERT) data, LIDAR data, and hyperspectral images. The LIDAR data was used to create a digital terrain model to calculate geometric properties of the basin including slope and aspect while the hyperspectral data was used as the input data. The supervised machine learning techniques classified the hyperspectral data into meadow plants, shrubs, deciduous forests, evergreen forests, and barren land. The regression analysis was able to explain geologic properties on a small scale but was unable to be generalized to a different location. The one regression model was unable to explain the measured data well. This analysis shows that regression is not always the solution to classification problems. With high variation in the data, more sophisticated machine learning models are required to understand the complex relationships within the dataset for prediction.

4.1.3 Machine Learning for Levees

With the ability of machine learning to process large datasets and predict properties of novel data, and it is a useful tool for levee applications where there is limited data. The use of machine learning for levee and earthen embankment applications is relatively new when not using GIS applications to process the data. A study performed by Kuchi et al. (2020) used Viola-Jones, Single Shot Multibox Detector, SVM, and gradient boosting for crack detection on levees. The dataset contained 12,800 optical images collected with UAVs of levees in New Orleans. All the images were altered so the cracks were the most prominent feature in the image. Because there were so few images with cracks, synthetic images with cracks were created to reduce the bias towards class imbalance. This study showed how drone imagery can have issues as inputs into algorithms if there are slight changes in lighting conditions and photo angle. Also, each photo in the dataset had to be manually annotated to be used in supervised machine learning algorithms, which is time-consuming. This study also showed that the Viola-Jones algorithm was not useful for dam and levee applications because it works best by detecting the highest contrast feature on the surface, which are typically not levee failure mechanisms.

Other research looked at detecting cracks in levees using passive seismic geophysical data. A study carried out by Belcher, Camp, and Krzhizhanovskaya (2015) built an experimental levee in the laboratory and forced it to fail through cracking. Geophones were placed on the levee to record time series data during the process. The authors applied Principal Component Analysis to the time

series to reduce the dimensionality of the data and then applied a variety of unsupervised machine learning algorithms, K-Means clustering, Hierarchical clustering, Gaussian Mixture Model, Partitioning Around Medoids, and Fuzzy C-Means Clustering to understand natural patterns in the data. The results were able to partition the data into crack events, normal events, and noise events, which can be useful for real-time monitoring.

This work was further developed by Fisher, Camp, and Krzhizhanovskaya (2016) and Fisher et al. (2017) with the same passive seismic dataset for detecting cracks in levees. In this study, the authors used one unsupervised method, One-Class Support Vector Machines (OCSVM), and two supervised methods, SVM and Multivariate Gaussian, to predict anomalies associated with cracking in the dataset. With this analysis, features including zero-cross, centroid, spread, root mean square, 85% rolloff frequency, flatness, kurtosis, irregularity, and skewness, were extracted from the time series to be used as inputs into the models. SVM predicted anomalies with 97% accuracy while OCSVM split the data into binary classifications of anomalies and normal. The authors found that OCSVM worked well for levee applications since the algorithm does not need all possible non-normal observations to be in the dataset. For levee failure mechanisms, creating a dataset that contains all types of failure mechanisms is extremely difficult, so using a technique like OCSVM is very useful for detecting all types of levee failure mechanisms.

While cracks are an important failure mechanism, many other subsurface levee failure mechanisms should be considered when performing a machine learning analysis. By developing models that only analyze one failure mechanism, the algorithms will be biased and unable to analyze the health of levees at the system level.

4.2 Machine Learning Introduction

Machine learning has become very popular in research as it is a way to analyze large amounts of data with high dimensionality. Large datasets are inputted into an algorithm and the algorithm learns patterns from the data. The goal of machine learning is to approximate these patterns to understand the process or to make predictions about future data (Alpaydin, 2010). Since machine learning excels at analyzing large amounts of data, it is a useful tool to evaluate data collected on levees. Due to levees being spatially extensive systems, large amounts of data of the surface and subsurface of these structures need to be collected to assess the system. Manually analyzing kilometers worth of surface and subsurface data becomes increasingly difficult and time-consuming, potentially resulting in the inability to identify trends in the data. Since machine learning works best with large datasets such as kilometers worth of levee data, it is a useful tool for identifying trends in the data spatially to assess the conditions of levees.

While machine learning is an effective tool for analyzing large amounts of data, not all machine learning algorithms are useful for every problem. Within machine learning, two main categories can be used to analyze data: supervised machine learning and unsupervised machine learning. Figure 4.1 illustrates how supervised learning works.

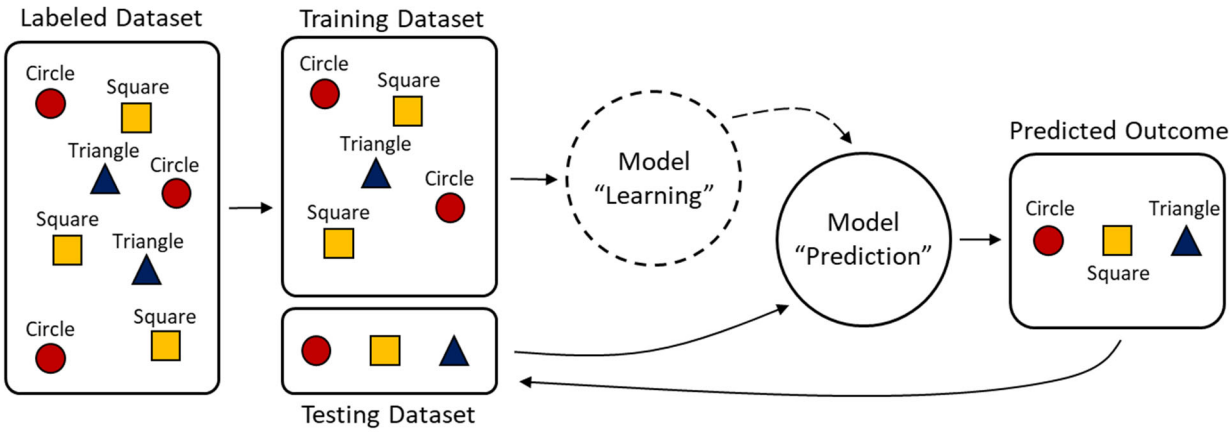


Figure 4.1 Supervised Machine Learning Workflow.

First, the dataset is split into training and testing datasets where each instance has a known label. The supervised algorithm learns the "right" answer from the training dataset. The testing dataset is then inputted into the model and the label is predicted and can be used as a measure of how well the algorithm learned the patterns in the dataset (Alpaydin, 2010). Uses of supervised machine learning include classification and regression problems.

The second type of machine learning is unsupervised machine learning. Unlike supervised machine learning, datasets in unsupervised are not labeled and the "right" answer is not known. The algorithm is not trained to predict an outcome as shown in Figure 4.2, but the goal of unsupervised learning is to look for natural clustering or patterns within the data by looking for similarities or differences in the data (Alpaydin, 2010).

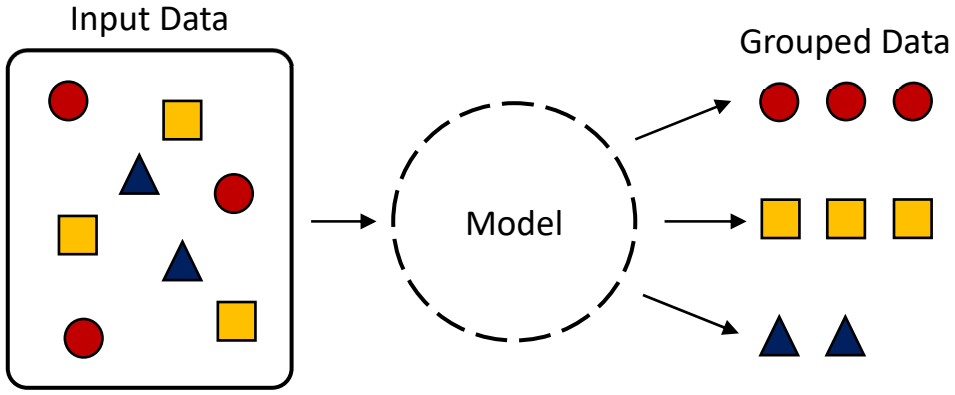


Figure 4.2 Unsupervised Machine Learning Workflow.

When performing a machine learning analysis, many factors are important to consider, such as feature engineering methods, choosing an appropriate algorithm, and choosing which model parameters to optimize. Because there are no universal selection criteria for these factors, multiple algorithms and feature engineering methods should be explored for any analysis to understand how these affect the problem.

4.3 Feature Engineering

Feature engineering is an important component of any machine learning analysis, regardless of it being supervised or unsupervised. Feature engineering is the process of transforming data to make it a useful input into the machine learning algorithm. Data needs to be transformed because machine learning algorithms either make assumptions or use metrics that are greatly affected by the structure of the dataset. Important considerations when choosing a feature engineering method include the scale, magnitude, and distribution of each variable (Zheng & Casari, 2018). Table 4.1 lists some common feature engineering methods that can be applied to a dataset.

Table 4.1 Feature Engineering Methods

	Method	Description
1	Standardization	Scaling features to have a mean of zero and a standard deviation of 1
2	Normalization	Scaling features to be between 0 and 1
3	Feature Selection	Choosing appropriate features to use in the analysis
4	Dropping Data	Removing missing data or removing features
5	Resampling	Up-sampling or down-sampling imbalanced datasets
6	Interaction Features	Combining features to include the interaction of the features

Standardization, normalization, feature selection, and resampling are discussed in more detail in the following sections. Dropping data is done when data points are missing. Dropping data could include removing an entire observation (i.e. row in a dataframe) where a data point is missing or dropping an entire variable (i.e. column in a dataframe). The dataset should always be checked for missing data before performing a machine learning analysis as algorithms handle missing data in different ways, like dropping the row or inputting a zero as a placeholder.

Interaction features combine features to examine how the interaction between features affects the analysis. The interaction features could be created by multiplying/dividing features, adding/subtracting features, or raising a feature to a power.

4.3.1 Standardization and Normalization

Scaling features are important when working with machine learning models because one feature could be given more weight in the model depending on the assumptions of the model. For example, algorithms that include a distance metric, such as K-Means Clustering, are more greatly affected by features with larger scales. By putting all features on the same scale, each feature will contribute equally to the distance, resulting in a less biased analysis (Xu & Wunsch II, 2009). Two of the most common methods for scaling features include standardization and normalization. Standardization is defined by Equation 4.1 (Zheng & Casari, 2018).

$$\overline{X} = \frac{X - \text{mean}}{\text{std}} \tag{4.1}$$

Where $\overline{X}$ is the standardized data or the z-score, X is the raw data, and std is the standard deviation of the raw data. By performing this operation, the standardized feature has a mean of zero and a

standard deviation of one. This is useful when comparing multiple features that are measured in different units or have large differences in scale (Xu & Wunsch II, 2009).

In contrast, normalization is another popular method to rescale features. Unlike standardization where the data is manipulated to be at a certain mean and standard deviation, normalization forces the data to be within a range of zero and one. This is done by comparing each data point to the minimum (X_{min}) and maximum value (X_{max}) of the feature (Zheng & Casari, 2018).

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (4.2)$$

Where X' is the normalized data. Normalizing the data either compresses or stretches each data point, making it inefficient when outliers are present. As a result, standardization is the preferred method for scaling features, but each method should be considered regardless to see how the analysis is affected.

4.3.2 Feature Selection

Feature selection is the process of choosing which variables to include in the machine learning analysis. When the dimensionality of the dataset is high, it can be computationally expensive to use all variables. Since not all variables contribute to the algorithm equally, choosing features to use in the analysis can provide benefits such as improved data visualization, data interpretation, a reduction in training times, a reduction in the storage requirement, and improved prediction performance (Guyon & Elisseeff, 2003). There are many feature selection methods to choose from including variable ranking and forward/backward selection.

In variable ranking, a scoring function computes a score for each variable using the training data. The higher the score, the more valuable the variable is to the prediction analysis. The scores can be sorted in decreasing order, ordering the relevance of each variable. This method is computationally efficient and robust against overfitting (Guyon & Elisseeff, 2003).

Given k number of features, backward selection will train the model first using all k features. Then k models will be generated using k-1 features. Out of the k new models, the best model is chosen, removing the feature that resulted in the best model. This process continues, each time dropping the feature that results in the best model when it is removed. This is repeated until a stopping condition is reached. Forward selection is the same as backward selection but instead of dropping features, features are added with each iteration.

Redundant features should also be considered in feature selection. Having redundant features can place too much emphasis on one variable which can skew the machine learning results. A method that can be used to remove redundant features is to compute the variance for each feature using the training data and remove the features with low variance.

4.3.3 Resampling

Resampling is a technique used on an imbalanced dataset. An imbalanced dataset contains significantly more observations of a predictor. In terms of levees, an imbalanced dataset would

typically contain more observations of "healthy" levees than "unhealthy" levees. Understanding if the dataset is imbalanced is important as the algorithm will be biased toward most observations.

There are two types of resampling, down-sampling and up-sampling. The difference between the two is illustrated in Figure 4.3.

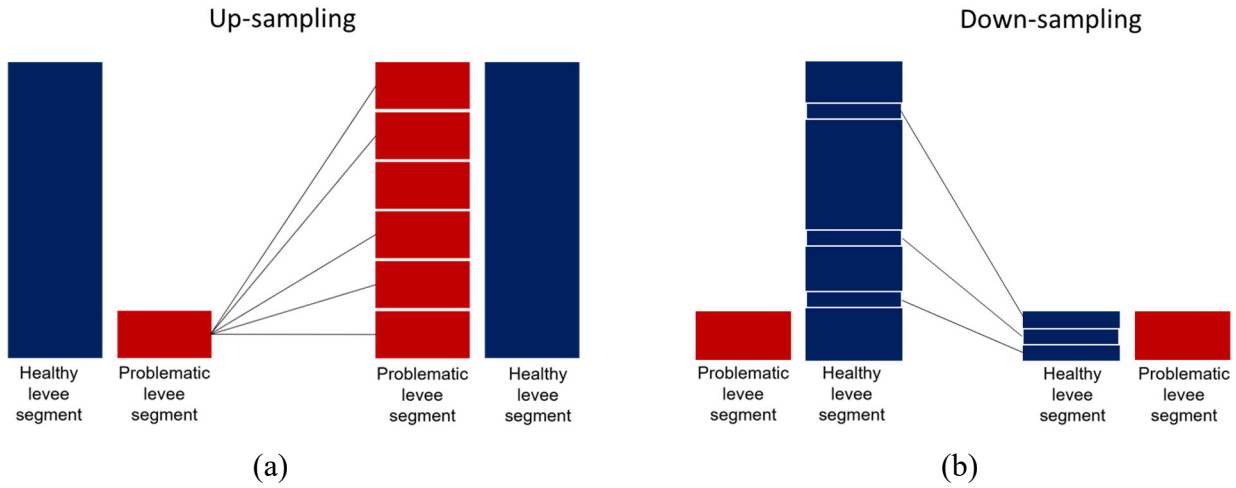


Figure 4.3 Resampling techniques for imbalanced datasets (a) up-sampling and (b) down-sampling.

Down-sampling includes throwing out some of the majority class data to have the same number of instances of the minority class. All the minority class observations are used. The majority class data is chosen based upon a probability of $1/\alpha$ where α is the ratio of majority to minority class. This method is not preferred with small datasets as data must be removed to create a balanced dataset. Instead, up-sampling can be used.

Up-sampling includes every instance of the majority class. To create a balanced dataset, the minority class is repeated several times until the number of observations is equal to the majority class. This method is typically preferred, especially with smaller datasets, as no data is removed.

4.4 Unsupervised Machine Learning

Unsupervised machine learning algorithms are very useful for learning subsurface information about levees since labels are not always available for subsurface information. Without invasive methods such as boreholes and CPTs, knowing the soil condition is very difficult, resulting in few to no labels for subsurface data. As a result, being able to analyze large amounts of data without labels becomes extremely valuable. Unsupervised machine learning techniques were leveraged in this study to identify potentially problematic zones in the subsurface using Vs and ρ as inputs for the Grand Island dataset. Multiple unsupervised machine learning algorithms were used to understand the effectiveness of the algorithms for levee health assessment.

4.4.1 Clustering Algorithms

Three unsupervised clustering techniques were used in exploring levee subsurface health data to determine if natural groupings of data exist. Exploring the results of different algorithms could lead to differences in the clustering of the data. As a result, it is important to consider multiple clustering algorithms. The algorithms explored in this study include K-Means clustering, Hierarchical clustering, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN).

4.4.1.1 K-Means Clustering

K-Means clustering is a distance-based method that groups data points using the closest distance to a centroid. First, centroids are randomly initiated. The number of centroids is user-specified and is equal to the number of clusters. Once the centroids are initiated, the distance between each point and each centroid is calculated. Then, each point is assigned to its closest centroid. For each cluster, the mean is found and the centroid shifts to the mean. This process is repeated until the mean and centroid converge as shown in Figure 4.4 (Lloyd, 1982; MacQueen, 1967).

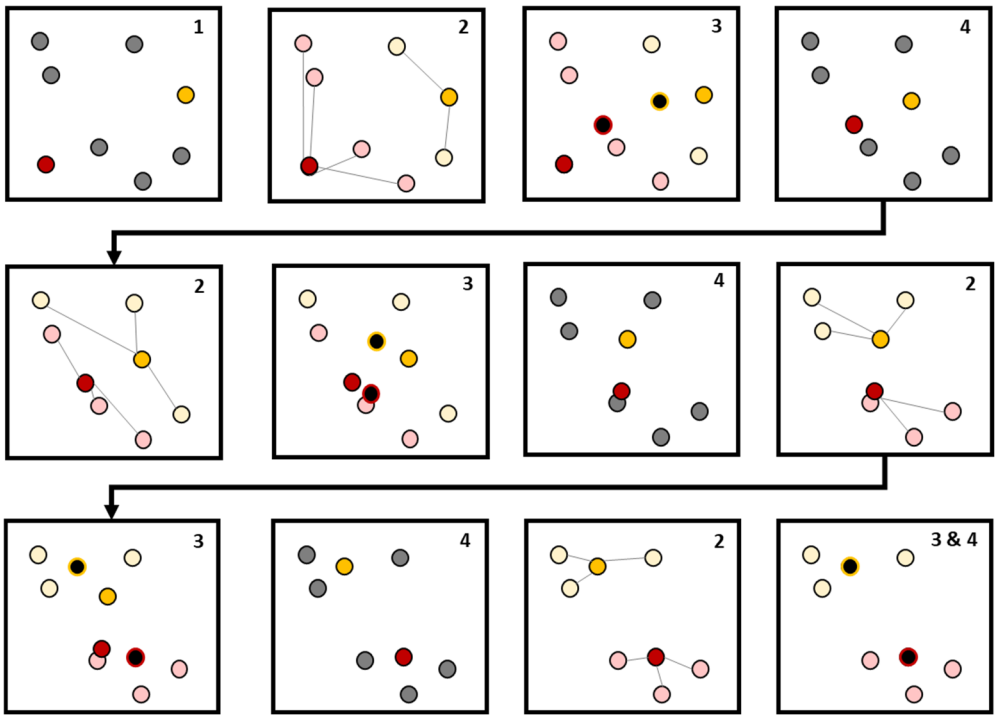


Figure 4.4 K-Means Algorithm (1) Randomly initiate centroids (2) assign points to their closest centroid (3) calculate the mean of the cluster (4) update new centroid location to mean location.

Because K-Means is a distance-based method, the distance metric used for analysis is important. Typically, the Euclidean Distance (E) is used as the length of a line segment between two points in space and is calculated through Equation 4.3.

$$E = \sqrt{(X_c - X_p)^2 + (Y_c - X_p)^2} \quad (4.3)$$

where X_c and Y_c are the x- and y- coordinates of the centroid and X_p and Y_p are the x- and y-coordinates of the data point in the dataset. Using the Euclidean Distance as the distance metric causes the clusters to form a hyperspherical shape. As a result, a K-Means clustering analysis works best when the data forms compact and hyperspherical clusters, naturally. Other distance metrics that can be used include the Manhattan Distance, Minkowski Distance, and the Hamming Distance (Xu & Wunsch II, 2009).

One of the major benefits of the K-Means algorithm is computational efficiency. Because simple distance metrics like the Euclidean Distance are used between each point in the data set, computation time is minimal. This allows for large datasets to be analyzed efficiently.

Since K-Means is a relatively simple technique that can be applied to a range of large datasets with minimal computation time, there are disadvantages to these methods as outlined in Table 4.2.

Table 4.2 Disadvantages to K-Means Clustering

	Disadvantage
1	Random centroid initialization could affect clustering
2	No universal method for determining the number of clusters
3	Sensitive to outliers and noise
4	Not compatible with categorical variables

Depending on where the centroid is first randomly located, different clusters could be created. This is problematic as running the analysis multiple times on the same dataset with different initial centroids could result in different clusters. It is important when performing a K-Means analysis to start with reasonable initial cluster locations, if possible. The sensitivity of clustering results can be examined by running the algorithm multiple times with different initial centroid locations and seeing how vastly the clusters change with each run.

One of the major hyperparameters for the K-Means algorithm is the number of clusters (k). Because the results are severely dependent on the input k, choosing the right k for analysis is essential. Unfortunately, there is no universal method for choosing k. When choosing k, multiple methods should be considered and are described in Section 4.4.2.

With K-Means, every data point in the dataset will be assigned to a cluster. This creates issues when there are outliers in the dataset. Even though an outlier may be far from all centroids and other data points, it will be forced into the nearest cluster. Because the algorithm calculates the mean in every iteration, outliers are going to significantly affect cluster locations.

Finally, since K-Means uses distance to sort data points into clusters, the algorithm is not compatible with categorical variables. Only numeric variables can be sorted into clusters.

4.4.1.2 Hierarchical Clustering

Hierarchical clustering is like K-Means clustering as it is also a distance-based method. There are two types of Hierarchical clustering. The first is called agglomerative clustering. In this method, each data point starts as its own cluster. Then, the two closest clusters are combined into one. Then, the next two closest clusters are combined. This process of combining the two closest clusters continues until all clusters have been merged into one cluster as shown in Figure 4.5.

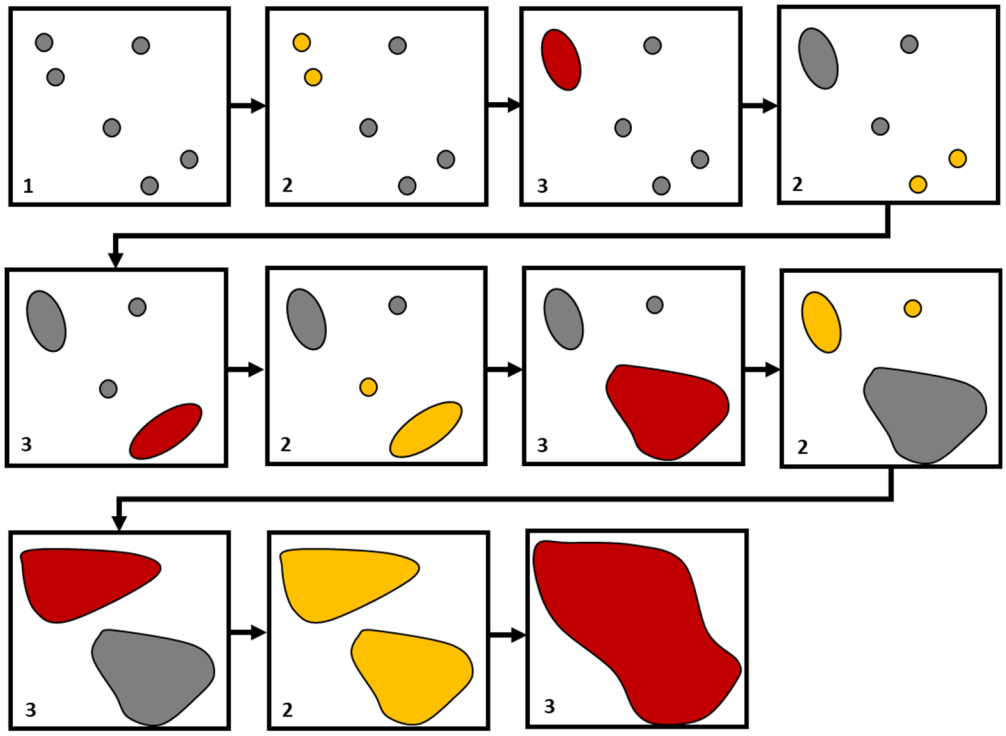


Figure 4.5 Agglomerative Hierarchical Clustering Algorithm (1) initial dataset (2) finds the two closest points (3) merges the two closest points into one cluster until one cluster remains.

The second type of Hierarchical clustering is divisive clustering which is the opposite of agglomerative clustering. In divisive, it starts with one cluster and then splits into smaller clusters until each cluster is a single data point. Because divisive is computationally expensive, it is not commonly used. Instead, agglomerative clustering is used more often (Xu & Wunsch II, 2009).

Since Hierarchical clustering is a distance-based metric, the choice of calculating distance will affect the clustering results, like with K-Means clustering. In the Hierarchical clustering method, the distance metric is called linkage and represents the distance between two clusters. There are many types of linkage, starting with the single linkage. Single linkage is where the distance is determined by the two closest points. As a result, a single linkage is optimal for clusters that are far apart (Sneath, 1957; Jain & Dubes, 1988; Everitt, Landau, & Leese, 2001).

Another type of linkage is complete linkage. This type of linkage is good for small and compact clusters as it uses the distance defined as the farthest distance between two points in different clusters. (Jain & Dubes, 1988; Evertitt, Landau, & Leese, 2001).

The third type of linkage is the average linkage. This distance is found by finding the distance between each point between the two clusters. Then the average of these distances is used to determine clusters. Similarly, the centroid method first finds the mean of each cluster and uses the distance between each cluster centroid to merge clusters (Sokal & Michener, 1958; Jain & Dubes, 1988; Everitt, Landau, & Leese, 2001).

A final type of linkage is called Ward's Method. This method aims at minimizing the growth of the sum of squares as clusters merge and is best suited for numerical variables (Ward, 1963; Jain & Dubes, 1988; Everitt, Landau, & Leese, 2001).

The output of Hierarchical Clustering is a dendrogram. A dendrogram is a graphical representation showing the optimal allocation of points to clusters using a tree-like structure (Xu & Wunsch II, 2009). An example of dendrograms is shown in Figure 4.6.

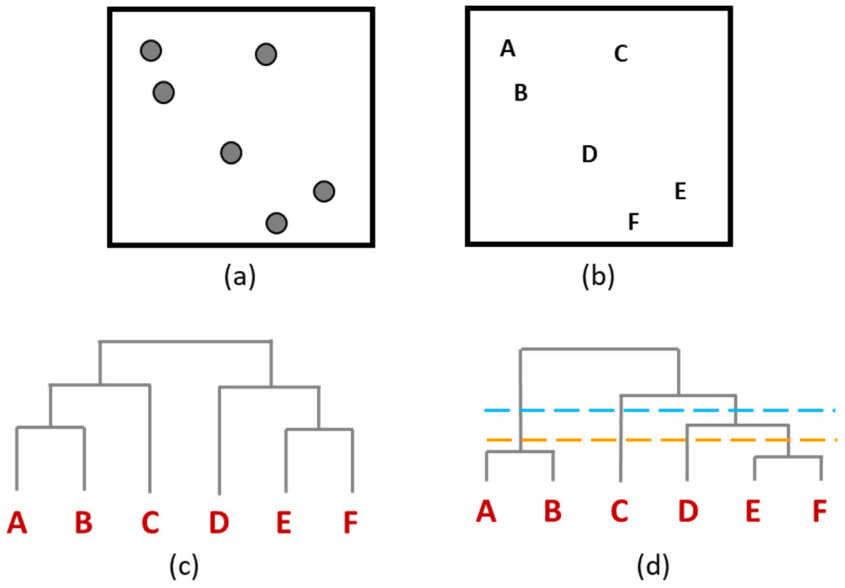


Figure 4.6 Example of possible dendrograms from Hierarchical Clustering (a) shows the points in the dataset (b) assigned "labels" to each point in the dataset for dendrogram viewing (c) example of a symmetric dendrogram (d) example of an asymmetric dendrogram where the blue line represents three clusters and the orange line represents four clusters.

Figure 4.6(a) shows the dataset in a 2-dimensional space as points while Figure 4.6(b) shows the same dataset, but each point has been given a unique label to differentiate between points. Figures 4.6(c) and 4.6(d) are examples of possible dendrograms. Each point in the dataset is represented as a leaf node. When two nodes fulfill the linkage requirement and are merged as a cluster, then data points are connected through branches. When the branches are short, the distance between the points is small. When the branches are long, the distance between a pair of points is large.

The dendrograms can be terminated at different heights to represent the number of desired clusters. For example, if the dendrogram in Figure 4.6(d) is cut at the blue dashed line, then the dataset will be split into three clusters. The first cluster would contain points A and B, the second cluster only contains point C, and the last cluster will contain points D, E, and F. The lower the termination on the dendrogram results in more clusters. If the dendrogram in Figure 4.6(d) is instead cut at the

dashed orange line, then the data will be split into four clusters. The first cluster would contain points A and B, the second cluster would contain point C, the third cluster would contain point D, and the last cluster would contain points E and F.

4.4.1.3 DBSCAN

Unlike the previous methods, DBSCAN is not a distance-based algorithm and does not need to be told the number of clusters. Instead, this method assumes that high-density regions surrounded by low-density regions are clusters. Because this method uses point density to find clusters, the clusters can be arbitrary shapes (i.e. not spherical) and can be of any size. Figure 4.7 graphically shows the components of the DBSCAN algorithm (Ester, Kriegel, Sander, & Xiaowei, 1996).

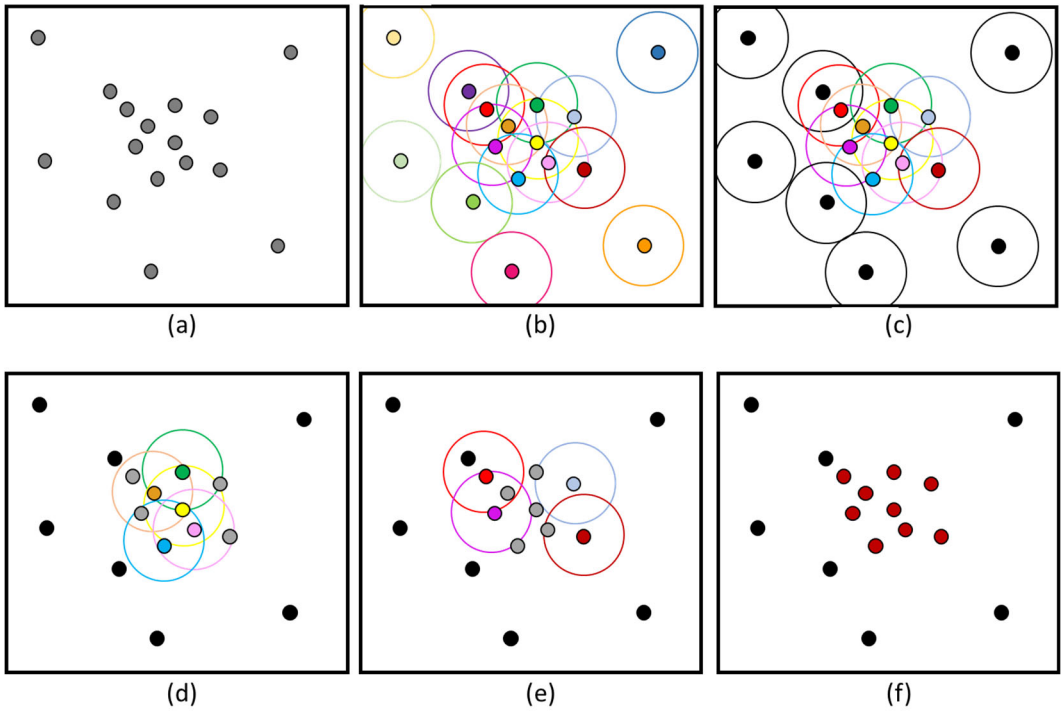


Figure 4.7 DBSCAN Algorithm (a) initial dataset (b) applying circles of radius, ε, around each point (c) the black points have no points in their radius and are classified as outliers (d) the colored points are core points and have the minimum number of points within their radius (e) the colored points are border points and (f) showing the clustered data points as red and outliers as black.

This method works by first drawing a circle, or neighborhood, of radius, ε, around each point and counting the number of points in each neighborhood. If a point contains the user-defined minimum number of points *(minPts*) within its neighborhood, then the point is a core point. If there are two core points within their neighborhood, they are part of the same cluster. If a point has at least one core point within the neighborhood but does not satisfy *minPts*, then the point is a border point, but still within the cluster. Finally, if a point does not contain a core point and does not satisfy

minPts, then the point is classified as an outlier and is considered noise. This method creates clusters by grouping all the points within its neighborhood.

4.4.2 Clustering Model Parameters

Many machine learning algorithms are sensitive to input parameters. For K-Means and Hierarchical clustering, the most important parameter that will dictate the clustering results is *k*. In contrast, DBSCAN is sensitive to *ε* and *minPts*. Choosing these parameters is important and there are no universal methods for choosing these parameters. This section describes the methods that were used to determine input parameters for the unsupervised machine learning algorithms.

4.4.2.1 Elbow Rule

The Elbow Rule can be applied to K-Means and Hierarchical clustering for determining the optimal *k* for analysis, but it can also be used with DBSCAN for determining the optimal *ε*. To find the optimal parameter for the Elbow Rule, regardless of determining *k* or *ε*, the goal is to choose the maximum curvature in the bend of the graph.

Determining *k*

When using the Elbow Rule to determine *k* for K-Means or Hierarchical clustering, the goal is to determine the within-cluster sum of squares (WCSS) for the range of *k* of interest. The first step in doing this is to find the Euclidean distance between every point in the cluster and the centroid, using Equation 4.3. To calculate WCSS, the distances are summed and divided by the number of points in the cluster using Equation 4.4 (Thorndike, 1953).

$$\mathrm{WCSS} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{E_i} \tag{4.4}$$

Where N is the total number of points in a cluster. An example of an Elbow Rule graph for determining *k* is shown in Figure 4.8.

In this figure, WCSS is on the y-axis with the number of clusters on the x-axis. For this dataset, the maximum curvature in the plot occurs at four clusters. As a result, four clusters should be considered for K-Means and Hierarchical clustering.

Determining ε

Determining *ε* using the Elbow Rule is like determining *k,* except WCSS is not used. Instead, the average distance between each point to its nearest neighbor is plotted in ascending order as shown in Figure 4.9.

A small nearest-neighbor difference indicates points within clusters since the points are close to each other. As the distance starts to increase, the points become more isolated and are considered noise. The goal is to choose an *ε* that separates points in clusters with points considered as noise (Hahsler, Piekenbrock, & Doran, 2019).

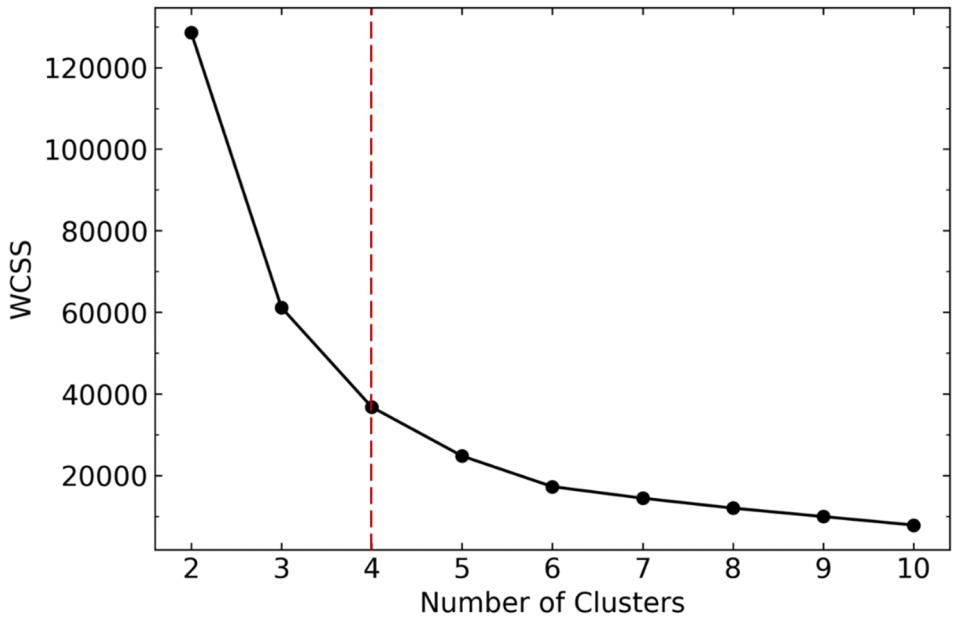


Figure 4.8 Elbow Rule for determining k for K-Means and Hierarchical clustering Analysis – optimal number of clusters is four.

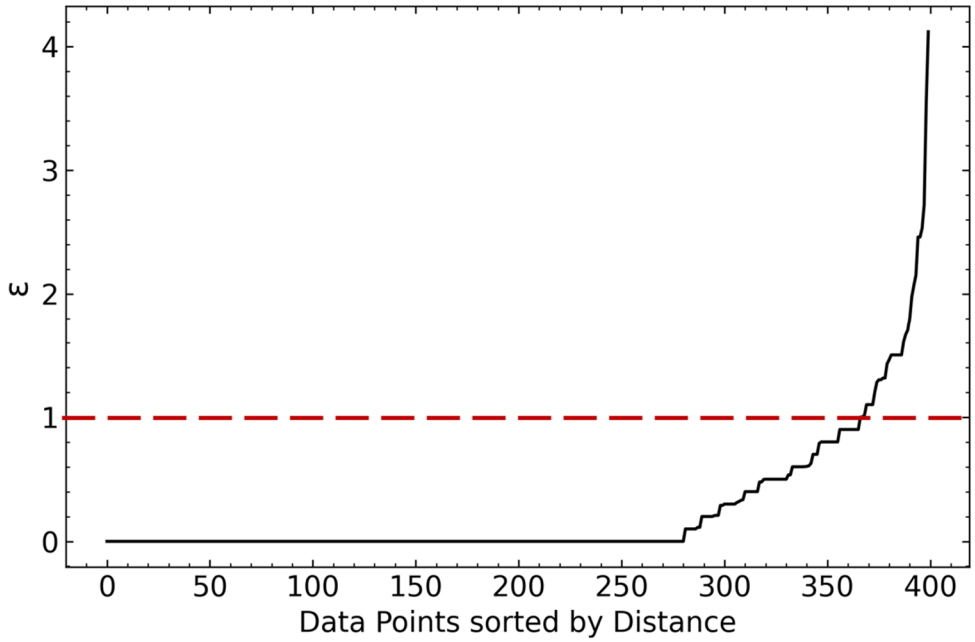


Figure 4.9 Elbow Rule for determining ε for DBSCAN – optimal choice for ε is 1.

4.4.2.2 Silhouette Rule

Another method to determine *k* for K-Means and Hierarchical clustering is through the Silhouette Rule. This method calculates a Silhouette coefficient that ranges between -1 and 1 and measures how similar data points are within a cluster to data points in other clusters. Calculating the coefficient can be done using Equation 4.5.

$$S(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \tag{4.5}$$

Where i represents a data point, a(i) is the average distance between i and all the other data points in the cluster of which i is a part while b(i) is the average distance to the closest cluster of which i is not a part of.

When the coefficient is 1, this is optimal as it represents a compact cluster far away from other clusters. As a result, the goal is to choose the *k* that maximizes the Silhouette Coefficient, as shown in Figure 4.10.

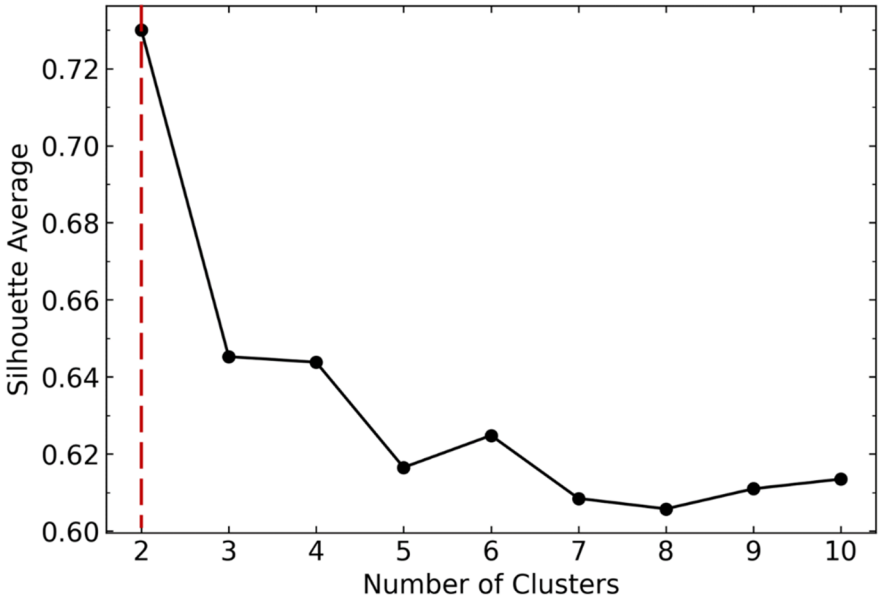


Figure 4.10 Silhouette Rule for determining k for K-Means and Hierarchical Clustering – optimal was chosen as two clusters.

In this example, the Silhouette Coefficient is at a maximum with two clusters. This example, though, shows a monotonically decreasing Silhouette Coefficient. In cases like this, the Silhouette Rule is not the best choice of method for determining *k*.

4.4.2.3 Gap Statistics

The final method for determining *k* is Gap Statistics. This method aims at comparing a reference null distribution with the within-cluster dispersion (W_k) for determining the optimal *k* (Tibshirani,

Walther, & Hastie, 2001). With this method, three values need to be considered when choosing *k*: the Gap Value, the Diff Value, and the Gap* Value.

Gap Value

The Gap Value is the $\log(W_k)$ standardized with the expectation of $\log(W_k)$ under a null reference distribution of the data and is defined as:

$$\mathrm{Gap}_n(k) = E_n^*\{\log(W_k)\} - \log(W_k) \tag{4.6}$$

$$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} D_r \tag{4.7}$$

$$D_r = \sum d_{ii'} \tag{4.8}$$

Where E* is the expectation from the reference distribution of a sample size n and calculated through Equation 4.9 where r is the cluster and $d_{ii'}$ is the squared Euclidean distance between point i and point i'.

$$E_n^* \log(W_k) = \frac{1}{B} \log(\prod W_{kb}^*) \tag{4.9}$$

Where B is the average number of reference datasets of $\log(W_k)$ computed from Monte Carlo samples from the reference distribution. The choice of optimal *k* for analysis is chosen by the one that optimizes $\mathrm{Gap}_n(k)$. An example of the figure produced from this method and the chosen value of *k* is shown in Figure 4.11.

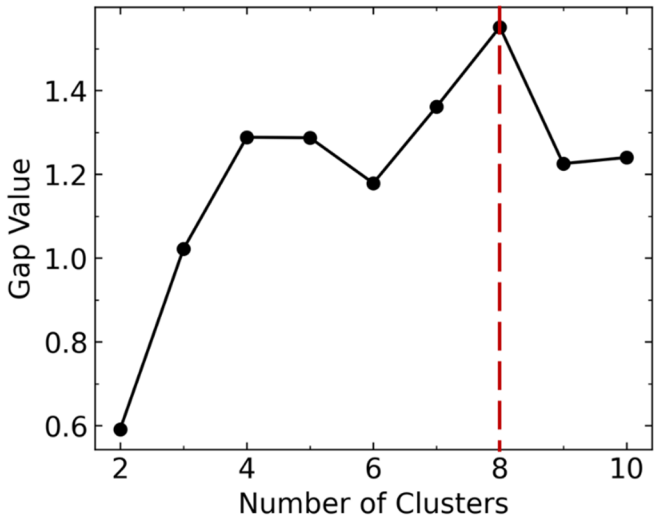


Figure 4.11 Gap Statistics using the Gap Value. Eight clusters maximize the Gap Value.

With the number of clusters on the x-axis and the Gap Value on the y-axis, the optimal number of clusters is eight because eight clusters maximize the Gap Value. When the Gap Value monotonically increases or decreases, then this method is not appropriate to use.

Diff Value

The Diff Value was introduced because the Gap Value may potentially be monotonically increasing or decreasing. This value compares the Gap Value of one value of *k* with the Gap Value of the next highest value of *k* as shown in Equation 4.10.

$$\mathrm{Diff} = \mathrm{Gap(k)} - \mathrm{Gap(k+1)} + \mathrm{s(k+1)} \tag{4.10}$$

$$\mathrm{s} = \sqrt{\left(1 + \frac{1}{\mathrm{B}}\right)}\,\mathrm{std(k)} \tag{4.11}$$

Where s is the simulation error calculated from the standard deviation of the Monte Carlo Samples (Tibshirani, Walther, & Hastie, 2001).

When choosing *k* using this method, the smallest *k* where the Diff Value is first positive should be chosen as shown in the example in Figure 4.12.

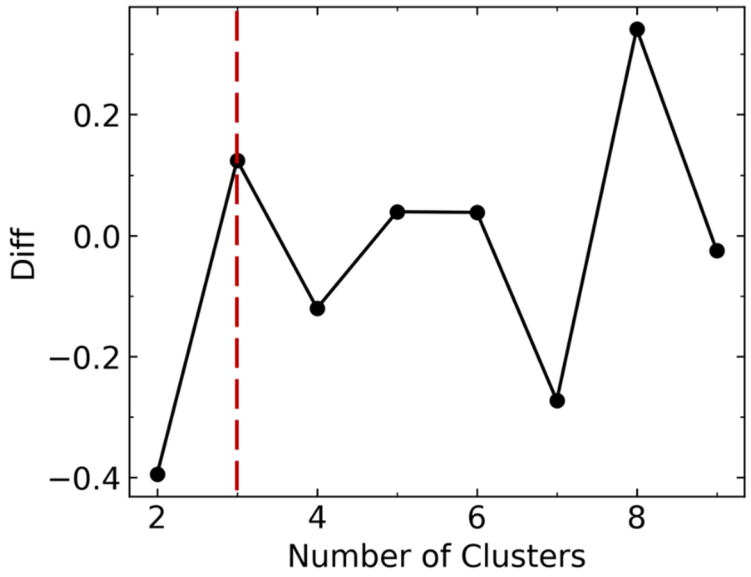


Figure 4.12 Gap Statistics using the Diff Value. The smallest k with a positive diff value is three clusters.

In this scenario, the smallest *k* analyzed was two clusters but resulted in a Diff Value of -0.4. The next smallest *k* of three clusters resulted in a Diff Value of 0.1. Because three clusters have the smallest *k* with a positive Diff Value, three clusters should be chosen for the analysis.

Gap* Value

The final parameter in Gap Statistics that can be used to help determine *k* is the Gap* Value. Like with the other methods for determining *k*, the Gap Value does not always work. Dudoit and Fridlyand (2002) have found that the Gap Value can overestimate *k* while Yin et al. (2008) showed that the Gap Value does not work well when the dataset contains clusters of different densities. As

a result, (Mohajer, Englmeier, & Schmid (2010) derived a new parameter, Gap*, based upon the Gap Value but removed the logarithmic functions, as shown in Equation 4.12.

$$\mathrm{Gap}_{\mathrm{n}}^{*}(\mathrm{k}) = \mathrm{E}_{\mathrm{n}}^{*}(\mathrm{W}_{\mathrm{k}}^{*}) - \mathrm{W}_{\mathrm{k}} \tag{4.12}$$

$$\mathrm{E}_{\mathrm{n}}^{*}(\mathrm{W}_{\mathrm{n}}^{*}) = \frac{1}{\mathrm{B}} \log\left(\prod W_{kb}^{*}\right) \tag{4.13}$$

When the Gap* is plotted with respect to *k*, as shown in Figure 4.13, the value chosen is the one that maximizes Gap*. When Gap* monotonically increases or decreases, then it is not the most appropriate method for determining *k*.

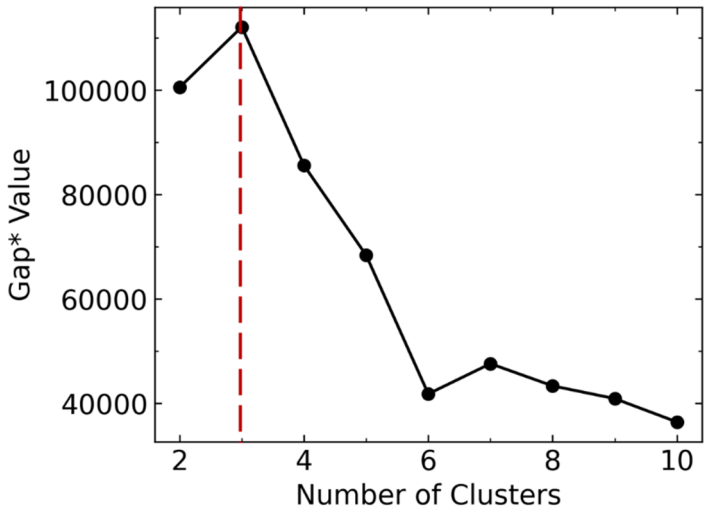


Figure 4.13 Gap Statistics using the Gap* Value. Three clusters maximize Gap*.

In this example, three clusters maximize Gap* and should be considered for the clustering analyses.

4.4.2.4 Determining *minPts*

The final parameter needed for DBSCAN is *minPts*. When determining *minPts*, Equation 4.14 is used (Ester, Kriegel, Sander, & Xiaowei, 1996; Sander, Ester, Kriegel, & Xiaowei, 1998).

$$\mathrm{minPts} = 2 \times \mathrm{dim} \tag{4.14}$$

Where dim is the dimension of the dataset. When working with two variables, the recommended value for *minPts* is four. Because *minPts* is dependent on the dimensionality of the dataset, as more variables are being introduced, the higher the *minPts* are needed. Typically, noisier datasets require a larger *minPts* (Ester, Kriegel, Sander, & Xiaowei, 1996).

4.4.3 Anomaly Detection

Anomaly Detection is a type of unsupervised machine learning that identifies data that is statistically different, or anomalous, from most of the data. These methods can work with training

data by learning on a dataset and then predicting on novel data or without training data by learning from the input data and determining which points are statistically different. Anomaly detection is unique as it is a binary classification, identifying points as either “normal” or “anomaly”.

4.4.3.1 Isolation Forest

Isolation Forest (IF) is an unsupervised Anomaly Detection algorithm that is a tree-based approach for identifying anomalies in the dataset. IF works on the concept that anomalous data points are far from normal data points. This method randomly initializes isolation trees to isolate the anomalous points with nodes that represent the internal features. These nodes are then split using branches until the feature is terminal. Once a node is no longer able to split further, it is considered a leaf. The more data is split, the longer the path in the tree (Lesouple, Baudoin, Spigai, & Tourneret, 2021).

In the tree, anomalous points will have shorter paths or tend to be higher in the tree. If data is considered “normal”, then that feature should be able to be split multiple times in the tree. If a data point is considered “anomalous” then it should not be able to be split multiple times. A visual concept of Isolation Forest is shown in Figure 4.14.

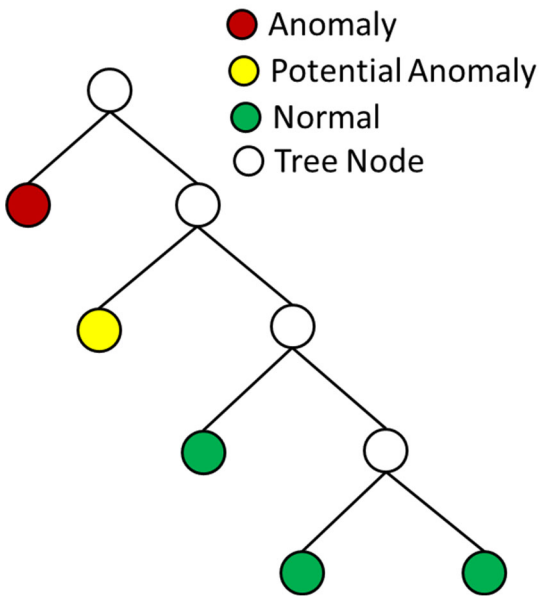


Figure 4.14 Isolation Forest concept showing anomaly and tree leaves.

This diagram shows an isolation tree where nodes are represented as white circles, branches as black lines, and leaves as red, yellow, or green circles. The green leaves are deeper in the tree, indicating a long path. As a result, the green leaves would be classified as “normal”. The red and yellow leaves have much shorter paths as they are higher in the tree. As a result, the red would be considered “anomalous” and the yellow would be considered potentially “anomalous” depending on the threshold set.

4.4.3.2 Local Outlier Factor

Local Outlier Factor (LOF) is a density-based anomaly detection algorithm that only considers the neighborhood of points around the point being analyzed. This method assumes that anomalous data points have a lower density of points compared to their neighbors.

The LOF value calculated by the algorithm is used to determine whether a point (O) is anomalous or normal. To do this, the Reachability Density (RD) first needs to be calculated using Equation 4.15 to determine if the object (P) is close. If the object is too close, such as object P1 in Figure 4.15, a distance controlled by a parameter, k, becomes the new distance to reduce statistical fluctuations of close neighbors (Breunig, Kriegel, Ng, & Sander, 2000).

$$\mathrm{RD} = \max(\mathrm{k-distance(O)}, \mathrm{d(P,O)} \tag{4.15}$$

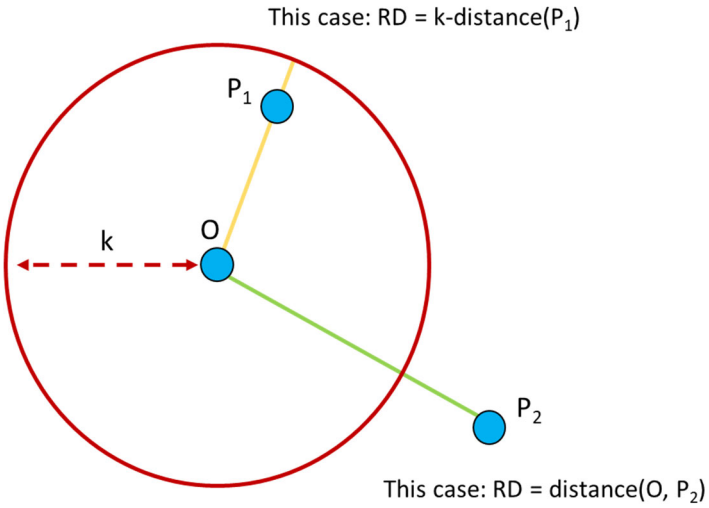


Figure 4.15 Local Outlier Factor showing the reachability density of a point (O).

Once the RD is known, the Local Reachability Density (LRD) can be calculated using Equation 4.16 to determine how far a data point is from the nearest grouping of points (Breunig, Kriegel, Ng, & Sander, 2000).

$$\mathrm{LRD} = \frac{1}{\frac{\sum_{O \in N_k(P)} \mathrm{RD(O,P)}}{|N_k(P)|}} \tag{4.16}$$

Where $N_k(P)$ represents the density of neighbors of point P. This equation shows that the higher the RD, there is going to be a smaller density of points around point O which results in the nearest cluster of points being far.

The LRD can then be used to calculate the LOF to determine if a point is anomalous or not through Equation 4.17 (Breunig, Kriegel, Ng, & Sander, 2000).

$$\mathrm{LOF} = \frac{\sum_{O \in N_k(P)} \frac{\mathrm{LRD_k(O)}}{\mathrm{LRD_k(P)}}}{|N_k(P)|} \tag{4.17}$$

The LOF of P is going to be higher when the LRD of P is low and when the LRD of P's nearest neighbors are high.

4.4.3.3 One-Class Support Vector Machines

One-Class Support Vector Machines (OCSVM) is an unsupervised binary classification system to minimize a hypersphere around a dataset in a feature space using a kernel function. The points that fall within the hypersphere are considered the normal data points and the points that fall outside the hypersphere are considered anomalies. The hypersphere boundary is determined by the contamination factor which describes the percentage of normal to anomalous points in the dataset (Scholkopf, Platt, Shawe-Taylor, Smola, & Williamson, 2001).

4.4.4 Anomaly Detection Model Parameters

The most important model parameter for anomaly detection is the contamination factor which describes the percentage of anomalous to normal points in the dataset. This parameter is needed for IF, LOF, and OCSVM, though in OCSVM, the parameter is called ν (nu). The other parameter needed is γ for OCSVM, which is the kernel coefficient for creating the hypersphere.

4.4.4.1 Contamination Factor

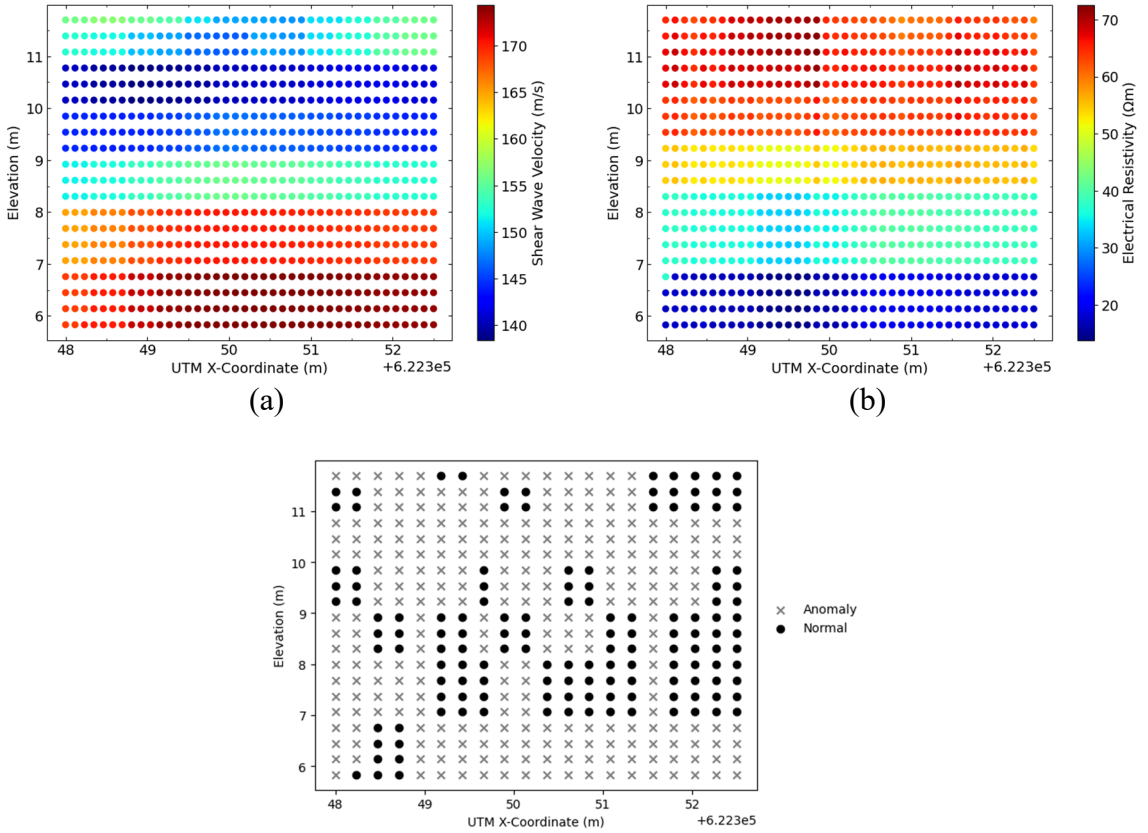
IF, LOF, and OCSVM are sensitive to the contamination factor, which is a value between 0 and 1 as it describes the percentage of anomalies in the dataset. For this study, the contamination factor was chosen to be the same for all anomaly detection algorithms based on the assumption that the dataset has a single percentage of anomalies in the dataset.

To determine the contamination factor, the dataset was first inputted into IF with the contamination function as "auto" in the SKLearn IsolationForest function in Python (Pedregosa, et al., 2011). The results showing which data points were considered anomalies were plotted with respect to the x-coordinate and elevation. It was concluded that the number of anomalous data points was too high, as shown in Figure 4.16.

The data points represented as black circles were determined to be "normal" while the data points represented as grey x's were classified as "anomalies." In this profile, 65% of the data was classified as anomalous. When compared to the *Vs* and ρ profiles in Figure 4.16(a) and Figure 4.16(b), respectively, the anomalously classified data points do not match areas of higher or lower *Vs* and/or ρ. To obtain a more appropriate estimate of the percentage of outliers in the dataset, the score describing how anomalous a data point is, was calculated using the SKLearn score_samples function for IF (Pedregosa, et al., 2011).

This score is constant regardless of the contamination factor. As a result, the score was used to determine a better estimate of the contamination factor by calculating one standard deviation from the mean of the scores. The more anomalous a point is the lower the score. Taking the number of data points that fell below the one standard deviation threshold resulted in a better estimate of the

percentage of outliers (Liu, Ting, & Zhou, 2008). This value was then used as the contamination input for IF, LOF, and OCSVM.



(a) (b)

(c)

Figure 4.16 Isolation Forest anomaly detection using contamination factor as "auto" showing (a) the shear wave velocity distribution, (b) the electrical resistivity distribution, and (c) the spatial distribution of anomalies.

4.4.4.2 Gamma (γ)

The final parameter needed for anomaly detection was γ which is the kernel coefficient required to draw the hypersphere for OCSVM. In the SKLearn library in Python, there are two options for γ, scale and auto, which calculate γ differently. To determine which to use for analysis, ν was kept constant at 0.15 (15%) and the algorithm was run on the same dataset twice, once using γ = "scale" and then again using γ = "auto" where "scale" and "auto" are defined by Equations 4.18 and 4.19, respectively (Pedregosa, et al., 2011).

$$\text{Scale} = \frac{1}{n_{features} \times \text{variance}} \tag{4.18}$$

$$\text{Auto} = \frac{1}{n_{features}} \tag{4.19}$$

Choosing between "scale" and "auto" was determined experimentally by looking at how changing γ affected the anomaly detection results. When γ was set to 'auto', more data was classified as anomalies and did not match the profiles as shown in Figure 4.17. When using "auto" most of the data points were classified as anomalies.

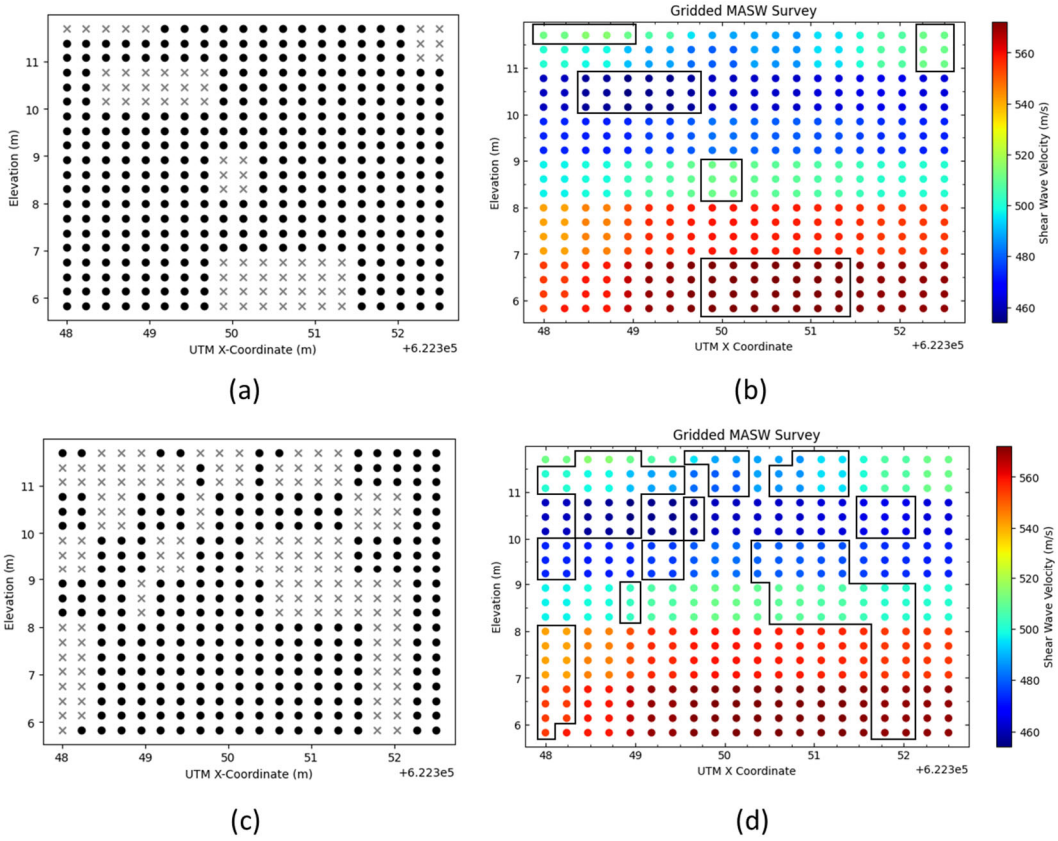


Figure 4.17 OCSVM anomaly classification results (a). classification results using γ = "scale" (b). results from (a) overlain on the MASW profile (c). Classification results using γ = "auto" (d). Results from (c) overlain on the MASW profile.

MASW data on the crest of Site 3 from the dry season was used as the input into the algorithm and the distribution of *Vs* was plotted spatially. Figure 4.17(a) shows the spatial distribution of anomalies when γ = "scale" and Figure 4.17(b) shows the *Vs* distribution with the anomalous areas overlain. Figure 4.17(c) shows the spatial distribution of anomalies when γ = "auto" with Figure 4.17(d) showing the *Vs* distribution with the anomalous areas overlain.

When using "scale" for γ, fewer sections of the profile were classified as anomalous and the sections that were classified correspond to a visual outlier in the profile. In comparison, when γ was set to "auto", more sections of the profile were classified as anomalous with many of these sections not corresponding to anomalous zones shown in the *Vs* profiles. Since using "scale" for γ classified the profile best, it was determined that for the analysis, γ would be set to "scale".

4.5 Supervised Machine Learning

Supervised machine learning is a tool used for classification, prediction, and regression. An algorithm is trained on known labels to create a model that can be applied to other data. This can be beneficial for levees where it is impossible to have ground truth data over the entire extent of the system. Creating a model using MASW and EMI can be extended to other extents of levees without ground truth, resulting in a powerful subsurface inspection framework. Ground truth data for levees include CPTs and borehole logs that describe the lithology of the subsurface. This ground truth data acts as labels the model will train on. The process used for performing a supervised machine learning analysis is shown in Figure 4.18.

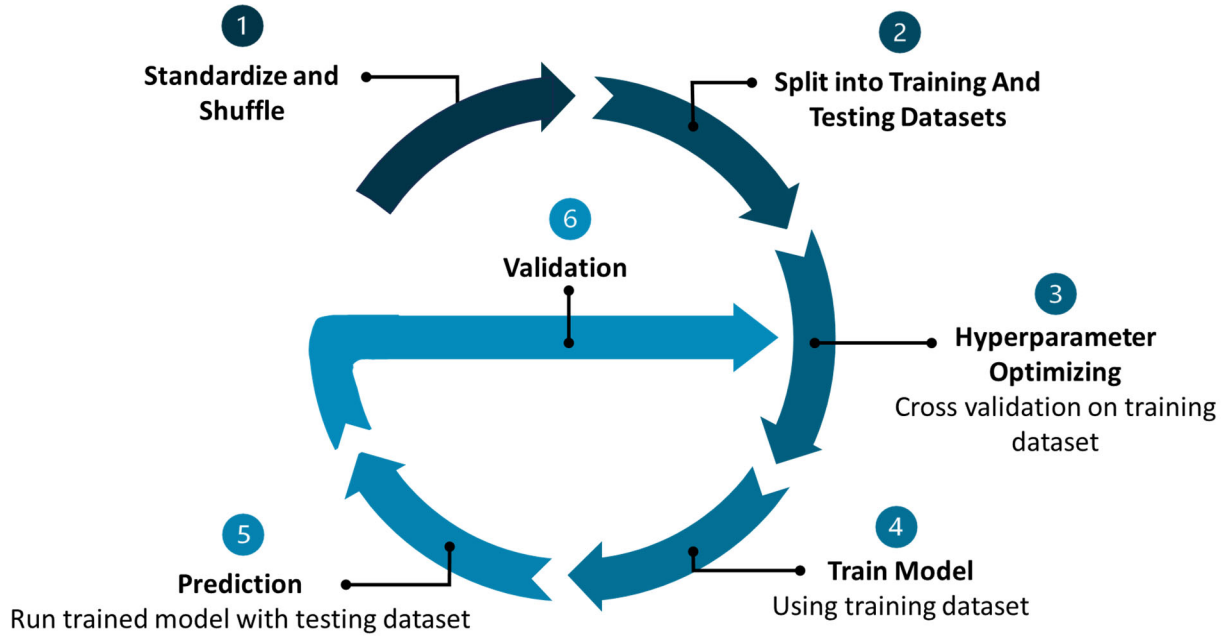


Figure 4.18 Supervised Machine Learning steps.

First, the dataset is standardized and then shuffled to ensure that no bias from sample preparation affects the classification prediction. Then, the dataset is split into a training dataset that is used to train the algorithm and a testing dataset that is used to test how well the algorithm can predict the classification on data it has never seen before.

The training dataset is used for cross-validation and a grid search to find the optimal hyperparameters for the model. The model is then trained using the optimal hyperparameters and training data. The testing data is then inserted into the trained model. The result of this prediction is compared to the actual results in the validation step. Based on how good the validation is, the hyperparameters may have to be turned and the process repeated.

4.5.1 Supervised Algorithms

Multiple supervised machine learning algorithms were performed on the Bacon Island geophysical dataset since ground truth data existed in the form of CPT data. These algorithms include K-Nearest Neighbors, Random Forest, and Support Vector Machines. The purpose of this study was

to understand the effectiveness of using supervised machine learning for predicting subsurface lithology from *Vs* and ρ.

4.5.1.1 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is a simple supervised machine learning algorithm that divides the dataspace into different classifications by finding the closest data point to itself and creating Voronoi Diagrams. Figure 4.19 shows a dataset with two classes, red and blue. Polygons are created around each point indicating zones of proximity, typically through the Euclidean Distance. When classifying or predicting a new data point, the polygon it falls into will control the classification. This method works by assuming that similar objects are in proximity. The number of neighbors, *K*, controls how smooth the decision boundary is. When there are more neighbors (i.e., larger value of *K*), the decision boundary will be smoother. This algorithm can learn complex decision boundaries and can separate data that is not linearly separable (Mitchell, 1997).

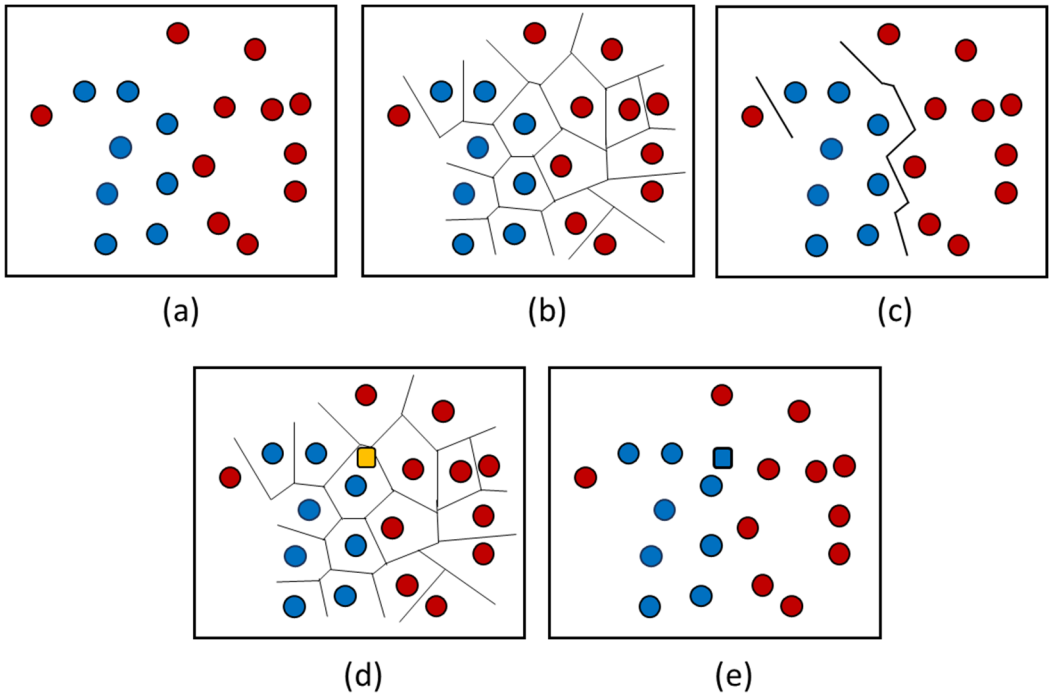


Figure 4.19 K-Nearest Neighbors Algorithm (a) the dataspace, (b) polygons of closeness around each data point, (c) nonlinear decision boundary around the two classes, (d) new data point is introduced and showing the polygon it falls into, (e) classifying the new datapoint based on the closest point.

Overfitting of a supervised machine learning algorithm needs to be considered to classify new data. For KNN, a larger value of *K* should be chosen to prevent overfitting. Choosing this hyperparameter is discussed more in Section 4.5.2.

4.5.1.2 Random Forest

Random Forest (RF) is a supervised technique that builds multiple estimators, or decision trees, to classify and predict unseen data. A decision tree is a hierarchical predictive model that consists of nodes and branches. The node at the top of the tree is the root node which then branches out to internal nodes that test a variable and then splits. The variables that are not split on are leaf nodes and are the outcome of the test, representing a class label. Each internal node is split until only leaf nodes remain (Rokach & Maimon, 2015).

The complexity of the tree is dependent on the tree size. As the tree becomes larger, the more complex the tree becomes and the higher the risk of overfitting. The size of the tree depends on many factors including the minimum number of leaves and the tree depth (Rokach & Maimon, 2015).

A Random Forest is built by first randomly subsampling the training dataset using replacement. Each subsample then builds a decision tree using a random sample of features and splits on those features. This process is repeated until the desired number of decision trees are made. Prediction is done by taking the majority vote amongst the trees which becomes the final classification (Rokach & Maimon, 2015). This process is illustrated in Figure 4.20.

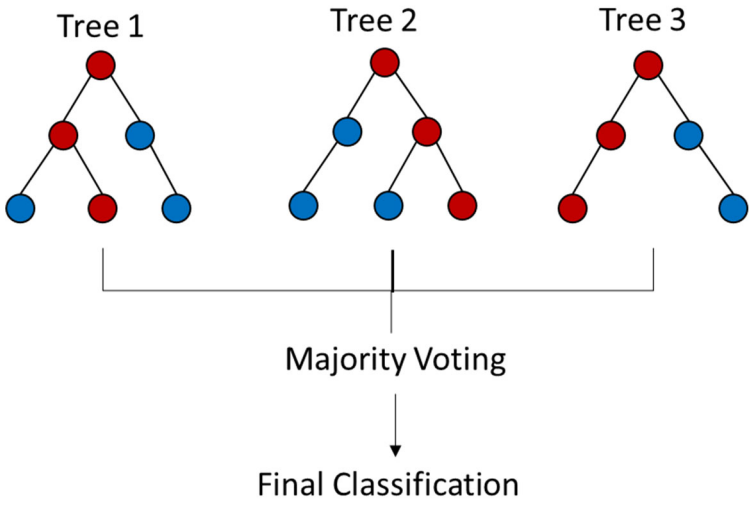


Figure 4.20 Random Forest depiction with multiple trees and taking the majority vote for the final classification.

Important features for classification in Random Forest are easy to identify as they occur in many trees, are split on often, and occur higher in the trees. This becomes important when trying to classify how important certain features are for classification.

Because RF uses a large ensemble of trees, overfitting is not possible, unlike when using a single Decision Tree. As a result, no additional regularization methods need to be employed when using RF (Breiman, 2001).

4.5.1.3 Support Vector Machines

The final algorithm used was Support Vector Machines (SVM). The goal of this algorithm is to find a hyperplane that maximizes the margin between two classes. This is done by first drawing a hyperplane at a random angle and determining the amount of space on either side of the line that separates the two classes. Then, the angle of the hyperplane is altered slightly, and the new margin is determined. This is repeated until the maximum margin between the two classes is found. An example of two hyperplanes with different margins is shown in Figure 4.21.

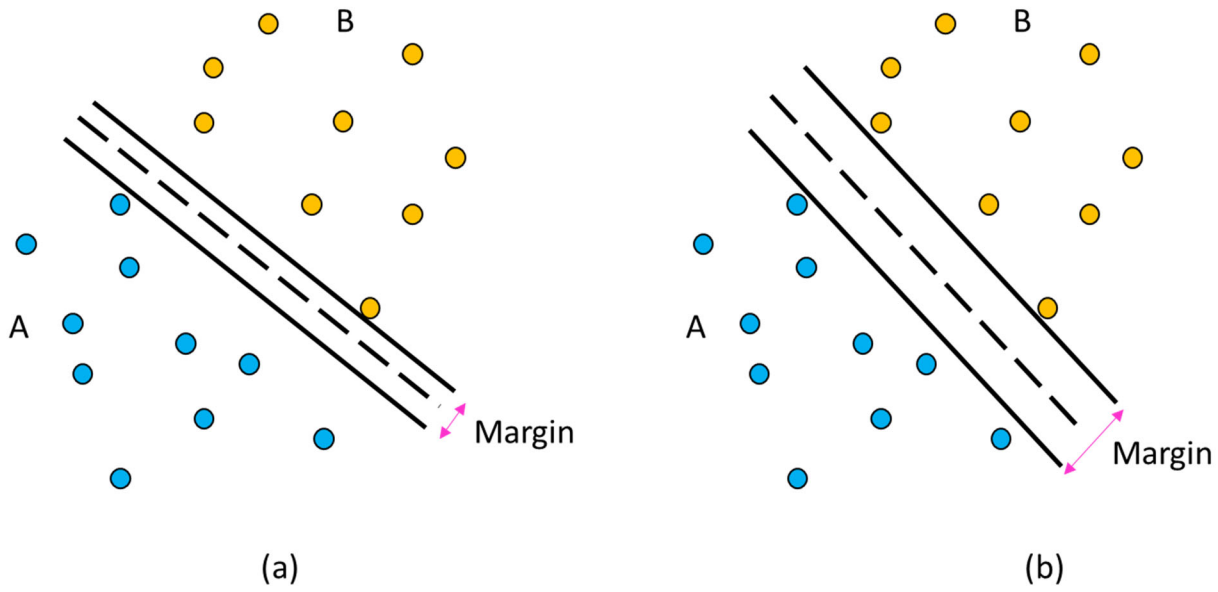


Figure 4.21 Support Vector Machine (a) hyperplane with a small margin (b) hyperplane with a larger margin.

Figure 4.21 shows a binary dataset with a blue and yellow class. In Figure 4.21(a), the margin drawn separating the two classes is relatively small. In comparison, Figure 4.21(b) shows the hyperplane drawn at a slightly different angle but creates a larger margin between the two classes. In this example, the hyperplane in Figure 4.21(b) would be chosen as the optimal decision boundary for SVM.

There are two hyperparameters of importance for SVM. The first is the regularization parameter (C) and it specifies the amount of training error. With a smaller C, the margin will be larger resulting in more training error. A larger C results in a smaller margin but classifies the data better. The second hyperparameter is γ and controls the curvature of the decision boundary when using the Gaussian RBF Kernel. If a linear or polynomial kernel is used instead, then γ is not needed. A larger γ will create a more complex decision boundary (more curvature) which could cause the model to overfit the data.

4.5.2 Hyperparameter Optimization

Hyperparameter optimization is important when building the supervised machine model to create a model that will not overfit or underfit the model and still provide high accuracy. Different

machine learning models will have different hyperparameters to tune. Table 4.3 lists the parameters that need to be optimized for their given model.

Table 4.3 Hyperparameters for each Algorithm for Optimization

Model	Parameter
K-Nearest Neighbor	K
Random Forest	Max tree depth, number of trees, minimum leaf samples
Support Vector Machines	C, γ

To optimize hyperparameters, cross-validation is used. Cross-validation is when the training data is split into a specified number of folds, or sections. For example, Figure 4.22 shows a 5-fold cross-validation. One of the folds is placed through the model with a given set of parameters and the accuracy is calculated. This is repeated for each fold and then the average of the accuracy is calculated. This process is iterated using different model parameters.

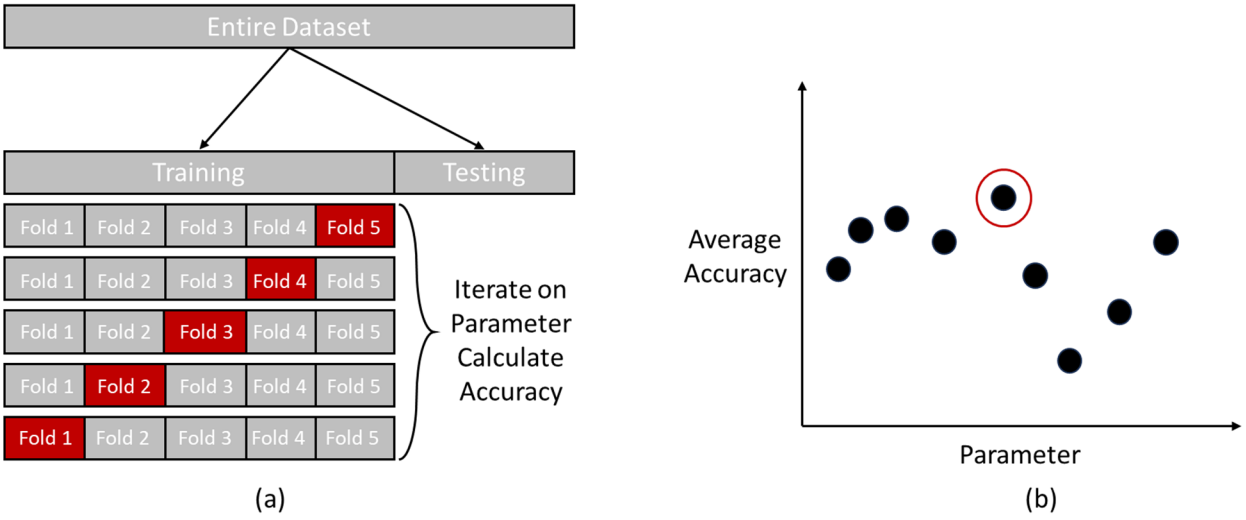


Figure 4.22 5-Fold Cross Validation (a) training dataset is split into five folds and parameters are iterated (b) plot showing the parameter vs. accuracy and picking the parameter with the best accuracy.

The accuracy measure can be plotted to determine the optimal parameter. When optimizing multiple parameters, a grid search is performed to find the combination of parameters that results in the best accuracy. The accuracy measures that can be used are described in more detail in Section 4.5.3.

4.5.3 Validation Methods

Many validation methods can be used to determine the accuracy of the model. One of the simplest is the Root Mean Square Error (RMSE). This measures how far the observation is from the regression line and is calculated through Equation 4.20:

$$\text{RMSE} = \sqrt{\sum \frac{|y_{\text{predict}} - y_{\text{true}}|^2}{N}} \tag{4.20}$$

Where $y_{predict}$ is the observation from the model, y_{true} is the actual observation, and N is the total number of observations. Because RMSE describes errors, it should be minimized when used as a validation metric.

Another validation metric is accuracy. Accuracy is commonly used as it describes the number of correctly classified observations in the entire dataset. This is calculated through Equation 4.21.

$$\text{Accuracy} = \frac{TN+TP}{N} \tag{4.21}$$

Where TN represents the number of points classified as True Negative, TP represents data points that are True Positive, and N is the total number of data points. TN is when the actual data point is negative and is classified as negative while TP is when the data point is actually positive and classified as positive. When using accuracy, it should be maximized for tuning hyperparameters. When the dataset is highly imbalanced, then accuracy may not be the best measure to use for validation.

When accuracy is not the best measure, precision, recall, and F-scores can be used. They can be calculated using Equation 4.22, 4.23, and 4.24, respectively:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4.22}$$

$$\text{Recall} = \frac{TN}{TP+FN} \tag{4.23}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}} \tag{4.24}$$

Where FP is False Positive and FN is False Negative. False Positive is when an observation is actually negative and classified as positive while False Negative is when an observation is actually positive and classified as negative.

Precision describes the quality of positive observations in the prediction by comparing the True Positive observations with all the positive classified observations. Recall compares the number of True Positives to all the actual positive observations. For both precision and recall, the higher the value, the better.

The F-scores are important when both precision and recall need to be represented at once. The higher the F-score, the better, as it indicates a high precision and recall.

4.6 Unsupervised Machine Learning Analysis on Grand Island

The unsupervised machine learning analysis was performed on the 2D *Vs* and ρ profiles collected at Grand Island for both seasons. Unsupervised machine learning was used on Grand Island since ground truth in the form of CPTs and boreholes was unavailable. Without ground truth, the dataset did not contain labels, as a result, supervised machine learning was not a suitable tool for analyzing this dataset. Unsupervised machine learning though, does not require labels since it learns patterns from the data. The unsupervised analysis first consisted of feature engineering the data and then applying clustering and anomaly detection algorithms to the datasets.

4.6.1 Unsupervised Feature Engineering

The first step in the unsupervised machine learning analysis was feature engineering. The dataset contained 2D *Vs* and ρ profiles from Bacon Island that were a function of elevation and the UTM 10N x and y coordinates. Due to the *Vs* and ρ profiles having the potential of having different spatial resolutions and spatial overlap, the first step was to create a uniform grid of points that represented the subsurface. An example of one of these grids is shown in Figure 4.23.

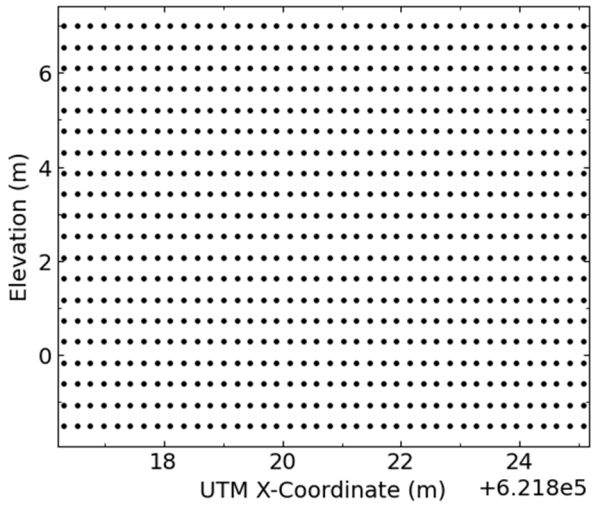


Figure 4.23 Uniform grid of points for machine learning.

In this grid, every point has a known x-coordinate, a known y-coordinate, and a known elevation. The elevation for the profiles was found by spatially joining the DEM from the LIDAR data to the MASW and EMI data.

The lateral and vertical extents of the grid were chosen based on where the MASW and EMI or ERT profiles overlapped, considering the skin depth of the EMI profile where applicable. The spatial resolution of the grid was determined by the profile with the lower spatial resolution. By creating a uniform grid of points, it allowed for direct comparison between the two profiles. For the unsupervised machine learning analysis, the *Vs* and ρ were interpolated to each point in the grid through nearest neighbors.

This process of creating a uniform grid of points was completed for each profile during the dry and wet seasons, individually. This resulted in a total of sixteen profiles that represented the crest and toe profiles from the three sites on Grand Island.

Another feature engineering technique applied to the dataset was the creation of the input datasets for the algorithms. A major component of this analysis was to look at how different profiles in the dataset affected clustering and anomaly detection results. To understand this effect, thirteen datasets were created, each containing different Vs/ρ gridded profiles depending on the aim of the dataset. Throughout this analysis, these datasets are called "Compilation Datasets." Table 4.4 summarizes each compilation dataset including a description of what each dataset represents, and the profiles included.

The first letter in the profile name indicates the season in which the data was collected, "S" for the dry season and "F" for the wet season. The next one to two letters represent where on the levee the data was collected, "c" was collected on the crest, "dt" was collected on the dry toe, and "wt" was collected on the wet toe. The name also includes the site number to indicate the site at which the data was collected.

For each compilation dataset, the Vs and ρ were then standardized using Equation 4.1 to reduce bias created by having the Vs and ρ at different scales.

Table 4.4 Compilation Datasets

	Dataset Name	Description	Size of Dataset	Included Profiles
1	Crest	Includes all the crest profiles regardless of site and season	7500	Fc_site2 Fc_site3 Fc3_site1 Fc5_site1 Sc_site2 Sc_site3 Sc3_site1 Sc5_site1
2	Toe	Includes all the toe profiles regardless of site and season	4225	Fdt_site2 Fwt_site1 Fwt_site2 Fwt_site3 Sdt_site2 Swt_site1 Swt_site2 Swt_site3
3	Dry	Includes all the profiles collected in the dry season regardless of site and crest/toe	5225	Sc_site2 Sc_site3 Sc3_site1 Sc5_site1 Sdt_site2 Swt_site1 Swt_site2

				Swt_site3
4	Wet	Includes all the profiles collected in the wet season regardless of the site and crest/toe	6500	Fc_site2 Fc_site3 Fc3_site1 Fc5_site1 Fdt_site2 Fwt_site1 Fwt_site2 Fwt_site3
5	Dry Site 1	Includes all the profiles collected in the dry season at site 1 regardless of crest/toe	2400	Sc3_site1 Sc5_site2 Swt_site1
6	Dry Site 2	Includes all the profiles collected in the dry season at site 2 regardless of crest/toe	1800	Sc_site2 Sdt_site2 Swt_site2
7	Dry Site 3	Includes all the profiles collected in the dry season at site 3 regardless of crest/toe	1025	Sc_site3 Swt_site3
8	Wet Site 1	Includes all the profiles collected in the wet season at site 1 regardless of crest/toe	3100	Fc3_site1 Fc5_site1 Fwt_site1
9	Wet Site 2	Includes all the profiles collected in the wet season at site 2 regardless of crest/toe	2000	Fc_site2 Fdt_site2 Fwt_site2
10	Wet Site 3	Includes all the profiles collected in the wet season at site 3 regardless of crest/toe	1400	Fc_site3 Fwt_site3
11	Site 1	Includes all the profiles collected at site 1 regardless of season and crest/toe	5500	Fc3_site1 Fc5_site1 Fwt_site1 Sc3_site1 Sc5_site1 Swt_site1
12	Site 2	Includes all the profiles collected at site 2 regardless of season and crest/toe	3800	Fc_site2 Fdt_site2 Fwt_site2 Sc_site2 Sdt_site2 Swt_site2
13	Site 3	Includes all the profiles collected at site 3 regardless of season and crest/toe	2425	Fc_site3 Fwt_site3 Sc_site3 Swt_site3

4.6.2 Unsupervised Clustering Results

This section will explore how feature engineering and hyperparameter optimization affect clustering results as well as the effectiveness of different clustering techniques for understanding the subsurface distribution of *Vs* and ρ. The factors affecting clustering and the number of clusters in a profile for compilation datasets were also analyzed. Lastly, this section describes how compilation datasets affect the global and local clusters in each profile. For a full list of model parameters used in the clustering analysis, refer to Appendix F.

4.6.2.1 Feature Engineering and Hyperparameter Effect on Clustering Techniques

Unsupervised clustering techniques are sensitive to model parameters as well as to the feature engineering techniques used on the dataset. Featuring engineering that will affect the analysis results include standardizing or normalizing the dataset while hyperparameters include *k* for K-Means and Hierarchical clustering and *minPts* and ε for DBSCAN. To understand the effect of these hyperparameters and featuring engineering techniques on clustering results, the profile collected in the dry season on the crest of the levee at Site 3 (sc_site3) will be used for a detailed analysis.

The Elbow Rule and Silhouette Rule were applied to this profile to determine the optimal *k* when the data was not manipulated, when *Vs* and ρ were standardized independently, and when the *Vs* and ρ were normalized independently. The Elbow Rule for DBSCAN was also applied to this profile to determine the ε for the three feature engineering techniques while *minPts* was chosen as four since the input dataset is 2-dimensional. The results for the optimal hyperparameters are shown in Table 4.5.

Table 4.5 Optimal Hyperparameters for Feature Engineering Methods

	Number of Clusters (k)		
Method	**Raw**	**Standardized**	**Normalized**
Elbow Rule	4	5	5
Silhouette Rule	2	2	2
DBSCAN – Elbow Rule	$\varepsilon = 1$	$\varepsilon = 0.05$	$\varepsilon = 0.01$
k used for Analysis	**4**	**5**	**5**

The Silhouette Rule showed the optimal number of clusters to be two for all feature engineering methods. When looking at the Silhouette graph, shown in Figure 4.12 in Section 4.4.2, the Silhouette Average was monotonically decreasing. When the Silhouette Average monotonically decreases, it should not be used to determine the optimal *k*. As a result, for this analysis, the *k* was chosen solely on the *k* from the Elbow Rule.

These optimal values of *k* were used in the K-Means clustering algorithm for all three feature engineering methods. The raw dataset resulted in four optimal clusters while the standardized and normalized datasets both had five optimal clusters. The results of the clustering are shown in Figure 4.24.

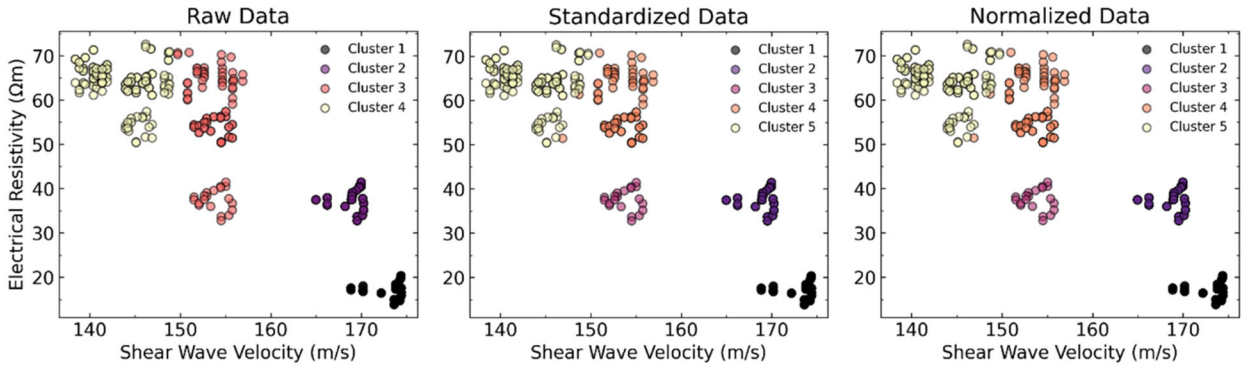
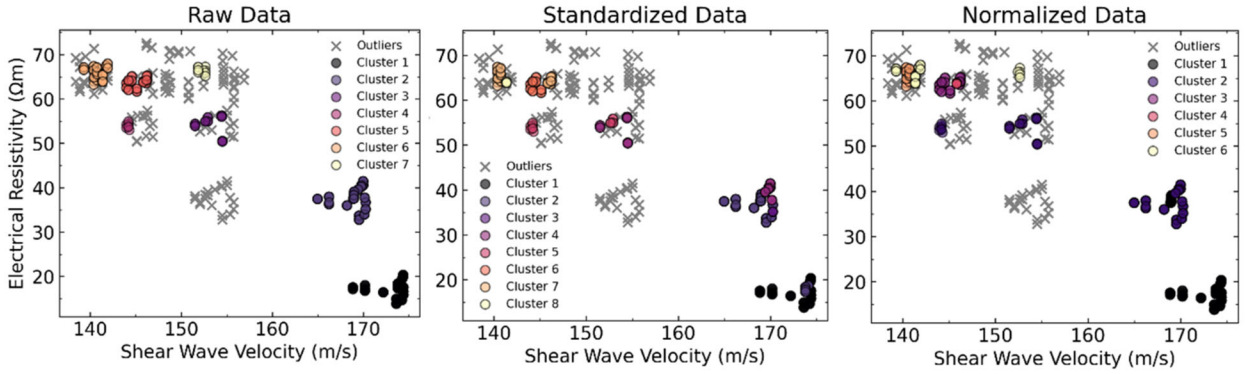


Figure 4.24 K-Means Clustering results (a) using the raw data, (b) using the standardized data, (c) using the normalized data.

Standardizing and normalizing the data added an extra cluster compared to the raw dataset. This is evident in the results shown in Figure 4.24 where the same clusters between the normalized and standardized data emerge with respect to Vs and ρ. The clusters for both the standardized and normalized datasets are: (1) Vs = 175 m/s, ρ = 25 Ωm, (2) Vs = 170 m/s, ρ = 35 Ωm, (3) Vs = 155 m/s, ρ = 35 Ωm, (4) Vs = 155 m/s, ρ = 60 Ωm, and (5) Vs = 145 m/s, ρ = 60 Ωm. Since the raw dataset only had four clusters, the group of data centered around a Vs of 155 m/s and a ρ of 35 Ωm did not have its own cluster but was grouped with the cluster that represented Vs of 155 m/s and ρ of 60 Ωm.

When using the optimized parameters for clustering using DBSCAN, the data was clustered into six to eight clusters based on the feature engineering technique applied to the dataset. In addition to the clusters, DBSCAN also classified data points as anomalies if they were found to be statistically different from most of the data, as shown in Figure 4.25.



(a) (b) (c)

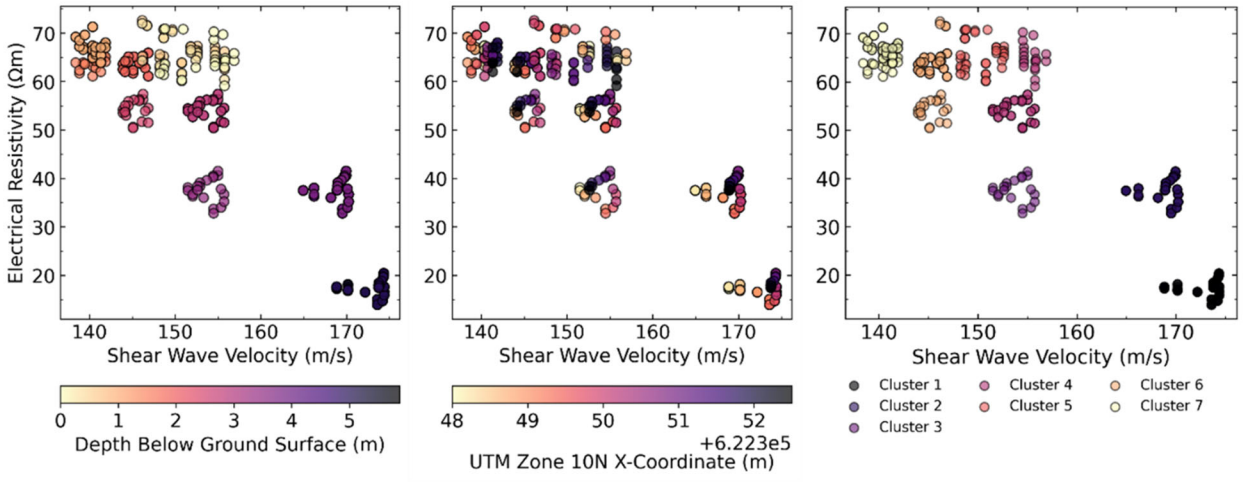
Figure 4.25 DBSCAN results (a) using the raw data, (b) using the standardized data, (c) using the normalized data.

Each feature engineering technique for DBSCAN resulted in a different number of clusters as well as a different number of data points classified as anomalies. When using the raw dataset, DBSCAN

used seven clusters and classified 136 data points as anomalies. Using the standardized dataset resulted in the highest number of clusters with eight clusters and the lowest number of outliers, with 109 outliers. When the normalized dataset was used, DBSCAN found the minimum number of clusters of six between the three datasets and the highest number of anomalies with 179 data points. Most of the anomalies were classified at high ρ and low Vs.

The clustering in DBCAN was similar to K-Means with the same clusters at high Vs and low ρ. At low Vs and high ρ though, DBSCAN broke the data into more distinct clusters compared to K-Means. The cluster at a Vs of 155 m/s and ρ of 35 Ωm that appeared when using K-Means on the normalized and standardized dataset was fully classified as anomalies when using DBSCAN.

Both DBSCAN and K-Means produced clear clustering results with similar clusters. To understand what physically in the subsurface could be causing these clusters, the depth below the ground surface and the x-coordinate of the profile were analyzed and are shown in Figure 4.26.



(a) (b) (c)

Figure 4.26 Exploration of data (a) showing the data as a function of depth below ground surface, (b) showing data by x-coordinate of profile, and (c) K-Means Clustering results using 7 clusters.

Figures 4.26(a) and 4.26(b) show the data as a function of Vs and ρ but also a function of the depth below the ground surface and the x-coordinate of the profile, respectively. The colors in both figures represent depth or the x-coordinate. Examining the data as a function of the x-coordinate in Figure 4.26(b) shows no natural groupings in the data. The x-coordinates with respect to Vs and ρ are scattered throughout the dataset.

When looking at the depth below the ground surface though, clear natural groupings in the data emerge. These clusters are like the natural groupings found in K-Means and DBSCAN showing similar ranges of Vs and ρ. When analyzing the color groupings in Figure 4.26(a), seven to eight natural groupings based upon depth emerge, which is more than the optimal k chosen by the Elbow Rule but similar to the number of clusters found while using DBSCAN.

To see if K-Means can more accurately represent the natural groupings in the data based on depth, seven clusters were used in the K-Means algorithm. The clustering results when using seven clusters are shown in Figure 4.26(c). When Figure 4.26(a) is compared to Figure 4.26(c), they both show similar clusters at the same ranges of *Vs* and ρ. As a result, using the Elbow Rule may underestimate the optimal *k*. In comparison, using the number of clusters found in DBSCAN as inputs into the K-Means and Hierarchical clustering algorithms may produce better results, as it more closely represents the natural grouping of the data.

Overall, this analysis showed that data should always be standardized or normalized to produce the best clustering results. With this dataset, the standardized data performed better than the normalized data as it had fewer classified anomalies, and the clustering in DBSCAN more closely resembled the natural groupings with respect to depth. For the application of levee subsurface health though, it was decided that DBSCAN was not the optimal algorithm to use as it overestimates the number of anomalous data points in the subsurface. As a result, further analysis for levee health using clustering was done with K-Means and Hierarchical clustering.

4.6.2.2 Factors Affecting Clustering

All profiles, including those clustered individually and clustered in a compilation dataset, were analyzed spatially and with respect to *Vs* and ρ. All clustering profiles can be found in Appendix G.

While analyzing each profile with respect to *Vs* and ρ, it became clear that some profiles were being clustered based on *Vs* and ρ, some profiles showed little to no clustering, and some profiles were being clustered solely on *Vs* or ρ. Examples of these profiles are shown in Figure 4.27.

Figure 4.27(a) shows an example of a profile where the data was not clustered at all by *Vs* and ρ. Figure 4.27(b) shows an example of a profile where the clustering was dominated by *Vs* shown by the defined vertical boundaries of the clusters while Figure 4.27(c) shows an example of a profile where the clustering was dominated by ρ shown by the defined horizontal boundaries of the clusters. The final type of profile is where *Vs* and ρ both contribute to clustering which is evident by clusters that have defined vertical and horizontal boundaries, as shown in Figure 4.27(d).

All the profiles that were clustered individually and clustered as part of a compilation dataset were given a label ranging from zero to three to represent how the profile clustered. The profiles given a label of (0) showed none or very little clustering, (1) represented profiles that were equally clustered by *Vs* and ρ, (2) represented profiles where *Vs* dominated clustering, and (3) were given to profiles where the ρ dominated clustering.

With a total of 80 profiles examined, 6.3% of them had little to no clustering. Profiles that were clustered equally using *Vs* and ρ comprised 50% of the total profiles while 32.5% of profiles were clustered on *Vs* and the remaining 11.3% of profiles were clustered on ρ.

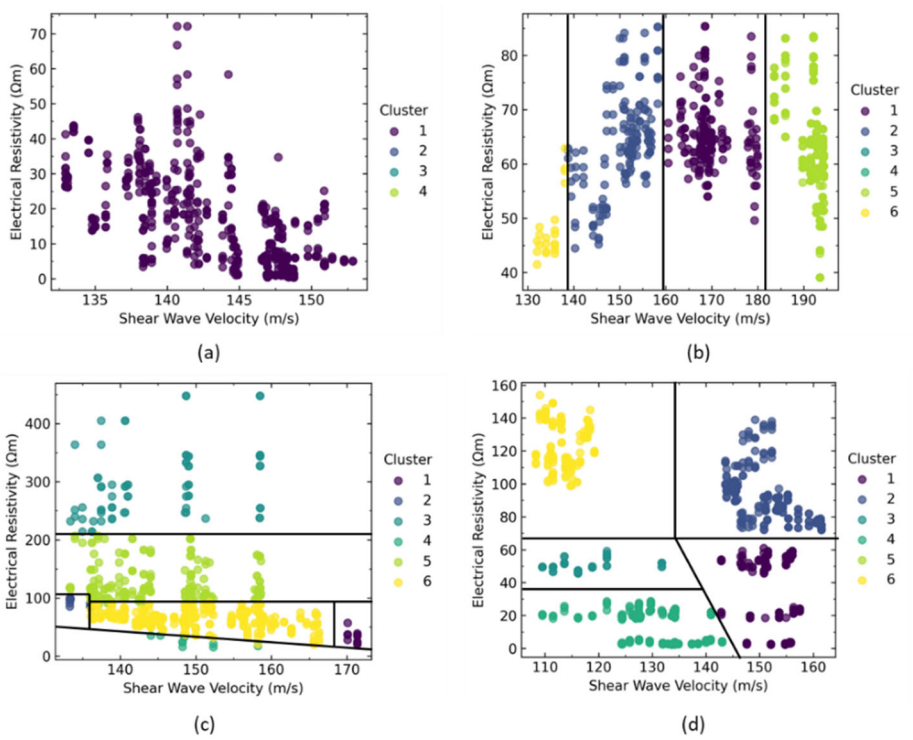


Figure 4.27 Examples of profiles where (a) the data was not clustered, (b) the data was clustered based upon Vs, (c) the data was clustered based upon ρ, and (d) the data was clustered by Vs and ρ.

To analyze how many profiles in each compilation dataset were clustered in a certain way, the percentage of each profile label in each compilation dataset was calculated and displayed in Table 4.6.

Table 4.6 Percentages of Clustering Label per Compilation Dataset

	Label			
Compilation	**0**	**1**	**2**	**3**
Individual	0.00	93.75	6.25	0.00
Crest	0.00	37.50	62.50	0.00
Toe	0.00	12.50	62.50	25.00
Dry	25.00	37.50	37.50	0.00
Wet	12.50	0.00	50.00	37.50
Dry Site 1	0.00	66.67	33.33	0.00
Dry Site 2	33.33	0.00	33.33	33.33
Dry Site 3	0.00	50.00	0.00	50.00
Wet Site 1	0.00	66.66	33.33	0.00
Wet Site 2	0.00	66.66	33.33	0.00
Wet Site 3	0.00	50.00	50.00	0.00
Site 1	0.00	83.33	16.67	0.00
Site 2	16.67	33.33	33.33	16.67
Site 3	0.00	75.00	0.00	25.00

When profiles were clustered individually, 93.75% of the profiles were clustered equally based upon *Vs* and ρ while the remaining 6.25% of profiles were clustered based solely on *Vs*. No profiles that were clustered individually had no clustering or clustering based solely on ρ.

Overall, most of the compilation datasets had no profiles that exhibited no clustering. The Dry Site 2 compilation dataset had the most at 33.33% of its profiles with the Dry dataset following at 25% of its profiles. Clustering based solely on ρ had the second lowest number of compilation datasets exhibiting profiles with this clustering. The Dry Site 3 compilation dataset had the most ρ clustered profiles at 50% followed by the Wet compilation dataset with 37.5% of the profiles.

Most of the compilation datasets showed that most of their profiles were clustered equally based on *Vs* and ρ. This trend is most pronounced in the Site 1 and Site 3 compilation datasets with 83.33% and 75% of their profiles being labeled as (1), respectively. Most of the compilation datasets had 33.33% to 66.67% of their profiles dominated by both *Vs* and ρ. Only the Wet compilation dataset and the Dry Site 2 compilation dataset had no profiles that were clustered on *Vs* and ρ.

It was also common for most of the profiles in a compilation dataset to be dominated solely on *Vs* with the Crest and Toe compilation datasets having the highest number of profiles at 62.50%. Most of the compilation datasets have approximately 33% of their profiles being dominated solely on *Vs*. The Dry Site 3 and the Site 3 compilation datasets were the only datasets to not contain any profiles that were dominated by *Vs*.

With these clustering mechanisms appearing throughout the profiles, it was important to understand how certain factors were attributed to how these profiles were clustering. Since the clustering was a function of *Vs* and ρ, it was hypothesized that the range of *Vs* and ρ of an individual profile compared to the compilation dataset it was a part of, was affecting how the data was being clustered. This hypothesis was supported by having 93.75% of the individually clustered profiles being a function of both *Vs* and ρ.

To analyze this, the maximum and minimum *Vs* and ρ of each compilation dataset was determined and the difference in the maximum and minimum *Vs* and ρ were used to normalize the individual profile to the compilation dataset. Normalization was important to allow for an unbiased comparison with other profiles. All the profiles in the compilation datasets were plotted with respect to the normalized *Vs* and ρ ranges, as shown in Figure 4.28, to see if there was a trend with how the profile was clustered.

How the profile was clustered is represented by shapes and colors. For profiles that showed no clustering, they are represented by purple circles. The profiles that exhibited equal clustering on *Vs* and ρ are denoted by blue squares. The green stars show the profiles where *Vs* dominated clustering and the yellow triangles show profiles that were dominated by ρ.

Profiles that exhibited little to no clustering have low normalized *Vs* and ρ ranges. All data is below a normalized range of 0.5 for both *Vs* and ρ. Profiles that were equally dominated by *Vs* and ρ tended to plot along the 45° line where the normalized *Vs* range was equal to the normalized ρ range.

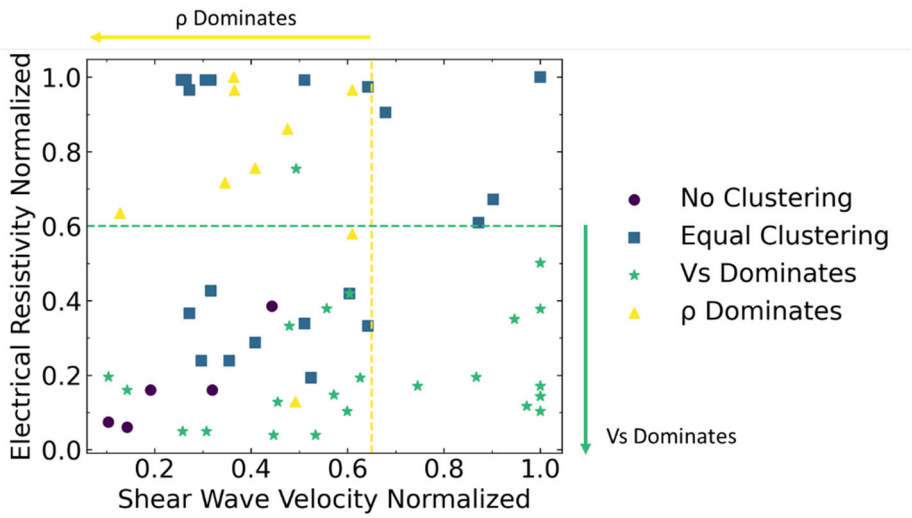


Figure 4.28 Clustering labels as a function of shear wave velocity and electrical resistivity.

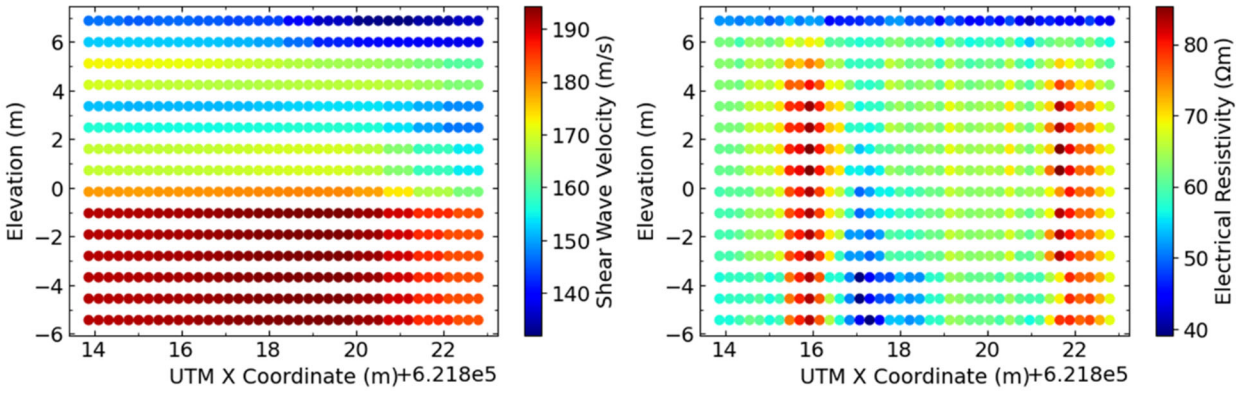
The profiles that were dominated by *Vs* were found at all ranges of the normalized *Vs* but typically less than a normalized ρ range of 0.6. Similarly, profiles that were dominated by ρ tended to plot at all ranges of normalized ρ but less than 0.63 of normalized *Vs*.

These results show that the compilation datasets will affect how data is clustered based on the range of *Vs* and ρ of the entire dataset compared to the individual profile. When the individual profile has a lower range compared to the dataset, the less likely the profile will be clustered. In comparison, when the profile has a larger range of *Vs* or ρ, the more likely the profile will be clustered based upon *Vs* and/or ρ. This is important if trying to target global or local effects in the dataset.

4.6.2.3 Global and Local Effects

While exploring how the range of *Vs* and ρ affected how the profile was being clustered, it was also apparent that the range of *Vs* and ρ of the compilation dataset affected what information was present in the clustering profile. For example, Figure 4.29 shows the *Vs* and ρ subsurface profiles collected on the crest of the levee in the dry season at Site 2.

The *Vs* profile shows lower *Vs* near the surface and higher *Vs* at depth in the profile. The ρ profile shows a relatively uniform ρ for the entirety of the profile except at an x-coordinate of 621816 m and 621822 m. At these two locations, there is an increase in the ρ for the entire depth of the profile. When these profiles were clustered individually compared to being a part of a compilation dataset, the resulting clustering was significantly different as shown in Figure 4.30.



(a) (b)

Figure 4.29 Subsurface profiles from the dry season on the crest at Site 2 showing (a) the shear wave velocity distribution and (b) the electrical resistivity distribution.

Figures 4.30(a)-(d) show clustering profiles from different compilation datasets. Each of these profiles shows the global effects in the dataset. These global effects are highlighting areas in the subsurface with significant *Vs* and/or ρ in the compilation datasets. In the case of the crest profile in the dry season on Site 2, this is the lower *Vs* zone near the subsurface. The high ρ zones in the subsurface are drowned out in the clustering because the ρ range for this profile is significantly lower than the ρ ranges for the compilation datasets.

In comparison, Figure 4.30(e) shows the cluster distribution in the subsurface when the profile was clustered individually. Since this profile was not part of a compilation dataset and did not have other profiles with higher ranges of *Vs* or ρ, the local effects at this site were clustered. The clustering shows areas with higher and low *Vs* as well as the high and low ρ. By clustering on the individual profile instead of a compilation dataset, the high ρ zones were captured in the clustering.

This concept of global and local effects in the clustering profiles is important for levee stakeholders. A stakeholder must decide if they are interested in some of the smaller changes in *Vs* and/or ρ when trying to understand subsurface levee health. If they decide that local effects are significant to levee health, it is recommended to collect geophysical data and cluster on profiles individually. If stakeholders decide that global effects with more significant zones of *Vs* and/or ρ are more important for levee health, then fitting the data in a compilation dataset and predicting on the profile is recommended for clustering.

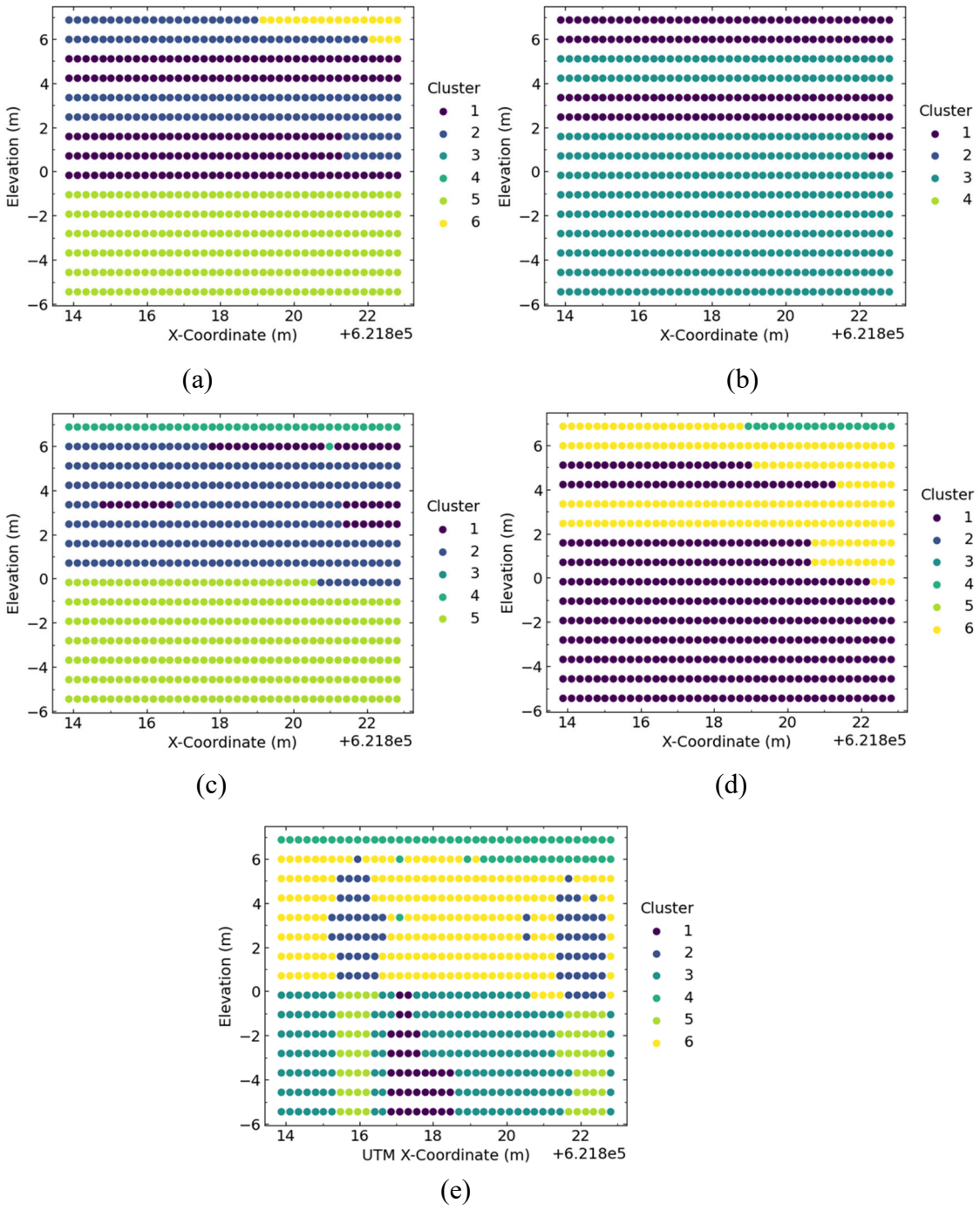


Figure 4.30 Subsurface profiles from the dry season on the crest at Site 2 showing (a) clustering profile from the Crest compilation dataset, (b) clustering profiles from the Dry compilation dataset, (c) clustering profile from the Dry Site 2 compilation dataset, (d) clustering profile from the Site 2 compilation dataset, and (e) clustering profile when clustering on the individual profile.

4.6.2.4 Factors Affecting the Number of Clusters per Profile

To perform a K-Means and Hierarchical clustering analysis, the number of clusters must be specified as a hyperparameter in the model. When performing clustering on an individual profile, the number of clusters present will always be the same as the number of clusters specified in the model. When performing clustering on a compilation dataset, the individual profiles in the compilation dataset may not utilize all the clusters specified for the model. As these models are a function of *Vs* and ρ, it was hypothesized that the number of clusters used for a profile in a compilation dataset was directly related to the profile's *Vs* and ρ range compared to the compilation datasets *Vs* and ρ range.

To test this hypothesis, each profile clustered in a compilation dataset was plotted as a function of their normalized *Vs* and ρ range for both K-Means and Hierarchical with respect to the profile's normalized number of clusters. The results of this analysis are shown in Figure 4.31.

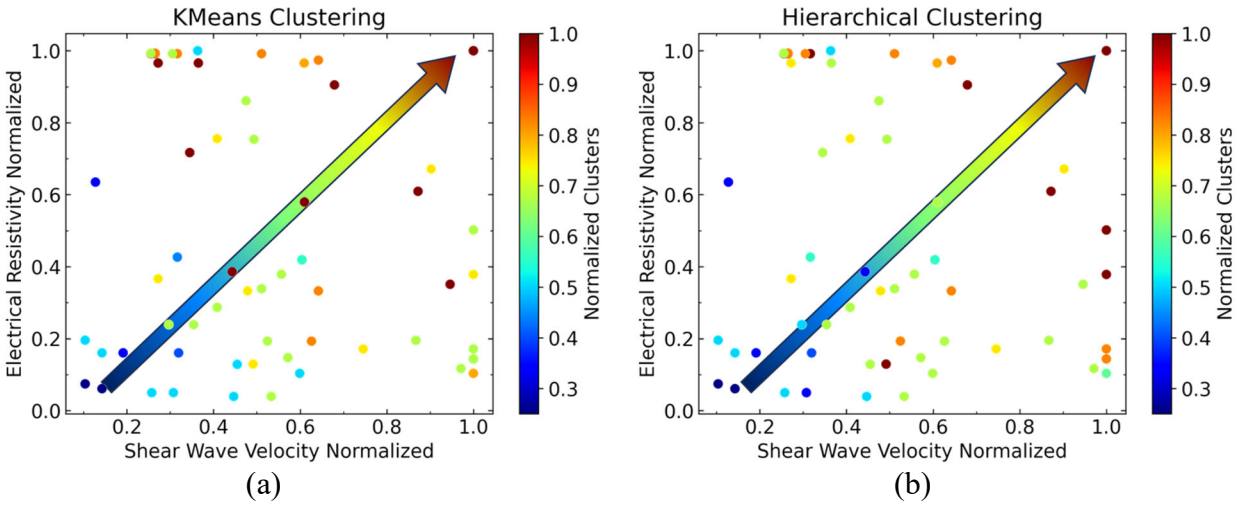


Figure 4.31 Number of clusters with respect to the normalized shear wave velocity and electrical resistivity for all profiles in a compilation dataset clustered with (a) K-Means Clustering and (b) Hierarchical Clustering.

The colors in these figures represent the normalized cluster per profile. The normalized clusters were found by normalizing the number of clusters used in a profile compared to the total number of clusters used in the compilation dataset. This allows for the comparison of profiles between compilation datasets. The cooler colors represent profiles that had significantly fewer clusters compared to the total number of clusters in the compilation dataset while the warmer colors represent profiles that had a similar number of clusters to the total number of clusters in the compilation dataset. When looking at these profiles with respect to the normalized *Vs* and ρ ranges, a trend emerges for both K-Means and Hierarchical clustering where profiles with fewer clusters compared to the total number of clusters are plotted at lower ranges (less than 0.5) of normalized *Vs* and ρ. Profiles that had a similar number of clusters to the compilation dataset were found at normalized ranges of *Vs* and ρ greater than 0.5.

The type of clustering algorithm could also affect the number of clusters used in each profile in a compilation dataset. To analyze this, the number of clusters for each profile was normalized by the

total number of clusters used for the compilation dataset. This was done for all compilation datasets and K-Means and Hierarchical clustering. To understand the difference between the number of clusters used for K-Means and Hierarchical clustering, the normalized clusters for each profile in each compilation dataset were subtracted between K-Means and Hierarchical. The results are shown in Table 4.7.

Table 4.7 Difference in Clusters between K-Means and Hierarchical Clustering

	Difference in Normalized Clusters			
Profile	**Crest/Toe**	**Dry/Wet**	**Dry/Wet - Site**	**Site**
Fc_site2	-0.17	0.67	0.00	0.00
Fc_site3	0.00	0.33	0.00	0.00
Fc3_site1	0.17	0.33	0.00	-0.17
Fc5_site1	0.00	0.33	0.00	0.17
Fdt_site2	0.00	0.00	0.00	0.00
Fwt_site1	0.00	0.00	0.00	0.00
Fwt_site2	0.00	0.00	-0.33	-0.17
Fwt_site3	-0.25	0.00	0.00	0.00
Sc_site2	0.00	0.00	0.20	-0.17
Sc_site3	0.17	0.00	0.00	-0.11
Sc3_site1	-0.17	0.00	0.00	0.00
Sc5_site1	-0.17	0.00	0.00	0.00
Sdt_site2	0.00	0.00	0.00	0.00
Swt_site1	-0.25	0.00	0.00	-0.17
Swt_site2	0.25	0.00	0.00	0.33
Swt_site3	0.00	0.00	0.00	0.00

If the difference in the normalized clusters is 0, then there is no difference in the number of clusters for that profile between K-Means and Hierarchical clustering. The closer the value is to +1 (shown in green in Table 4.7), the more clusters K-Means had for that profile compared to Hierarchical clustering while when the value is close to -1 (shown in red in Table 4.7), the more clusters Hierarchical had compared to K-Means clustering.

For 67% of the profiles, K-Means clustering and Hierarchical clustering produced the same number of clusters in a profile in the compilation dataset. Out of the remaining profiles that did not have the same number of clusters, 17% of them showed that Hierarchical clustering resulted in more clusters in the profile while the remaining 16% of the profiles had K-Means clustering using more clusters in the profile.

The Dry compilation dataset had no difference in the number of clusters between the two algorithms while the crest profiles in the Wet compilation dataset had more clusters when using K-Means compared to Hierarchical. The Crest compilation dataset was equally split between having profiles that had the same number of clusters, profiles that had more clusters when using K-Means, and profiles that had more clusters when using Hierarchical. The Toe compilation dataset though, showed that most of the profiles had no difference in clusters. The Dry and Wet (site specific) compilation datasets had a good agreement between K-Means and Hierarchical with only two out of 16 profiles showing a difference. The wet toe profile at Site 2 during the wet season (Fwt_site2) had more clusters when using Hierarchical and the crest profile at Site 2 during the dry season (Sc_site2) had more clusters when using K-Means clustering. The Site 1 and Site 2

compilation datasets had similar results where three out of the six profiles had no difference in clustering whereas one profile had more clusters when using K-Means and two profiles had more clusters when using Hierarchical clustering. The Site 3 compilation dataset had all but one profile with no change in clusters and the remaining one profile had Hierarchical have more clusters.

When the normalized difference in Table 4.6 was summed across all profiles for K-Means and Hierarchical clustering, K-Means clustering had a total of 2.95 while Hierarchical clustering had a total of -2.11. Since K-Means had a higher absolute total, this shows that K-Means tended to have more clusters overall compared to Hierarchical clustering. Even though K-Means had a slightly higher number of clusters, overall, there was very little difference in the number of clusters between clustering algorithms.

4.6.3 Potentially Problematic Zones Using Clusters

This levee health framework aims to be able to characterize subsurface conditions using geophysical data. Clustering is a method that was utilized with the hope of being able to determine areas in the subsurface that may be associated with potential failure mechanisms such as seepage. The clustering methods excelled at splitting the data into zones based upon *Vs* and ρ but because ground truth data was not available for Grand Island, determining which clusters were associated with potentially problematic zones was difficult. Certain clusters could be attributed to soil stiffness based on *Vs* or sandy/clayey material based on ρ (Hayashi & Suzuki, 2004). This difficulty in determining potentially problematic zones increases as the number of clusters in the profile increases. Figure 4.32 shows an example of a subsurface profile and the clustering results where nine clusters were determined to be optimal.

Figure 4.32(c) shows clear clustering based on *Vs* and ρ, splitting the data into nine clusters. When these clusters are analyzed spatially in the subsurface as shown in Figure 4.32(d), the clusters are in multiple locations in the profile, breaking the subsurface into many locations. Breaking the profile into many zones increases the difficulty in understanding which clusters are associated with potentially problematic zones. Using a high number of clusters to detect potentially problematic zones in the levee subsurface is not always beneficial as it reduces the interpretability of the profile. For example, in Figure 4.32(c), there is a cluster that represents *Vs* of 150 to 175 m/s and ρ of 0 to 20 Ωm and a cluster that represents *Vs* of 150 to 180 m/s with a ρ of 20 to 50 Ωm. These clusters represent similar ranges of *Vs* and slightly different ρ. Based on the similar ranges of *Vs* and ρ, it is difficult to determine if either or both clusters could be problematic.

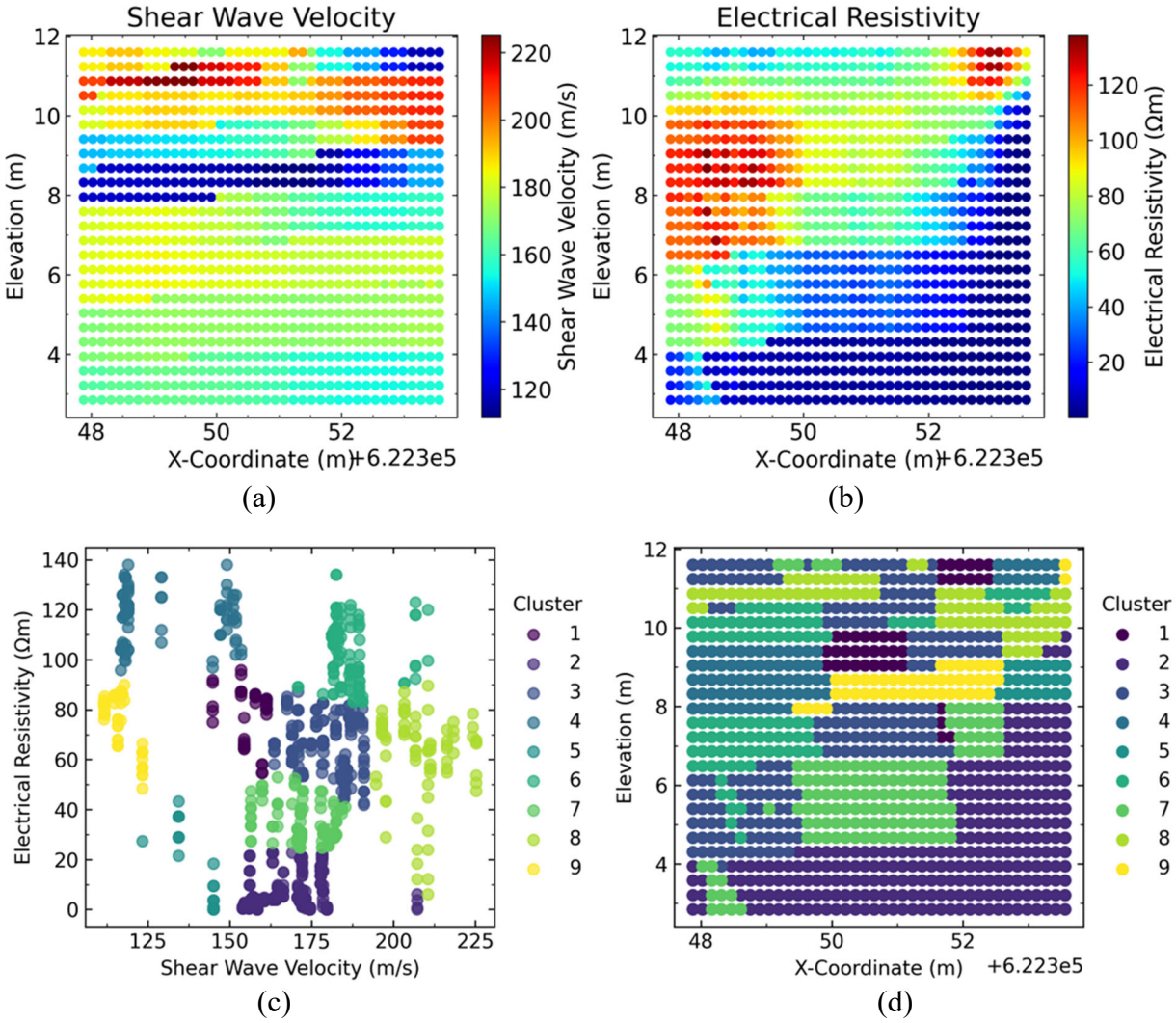


Figure 4.32 Interpretability of potentially problematic zones for clustering showing (a) shear wave velocity subsurface profile, (b) electrical resistivity subsurface, (c) clustering with respect to shear wave velocity and electrical resistivity, and (d) clustering distribution in the subsurface.

4.7 Anomaly Detection Results

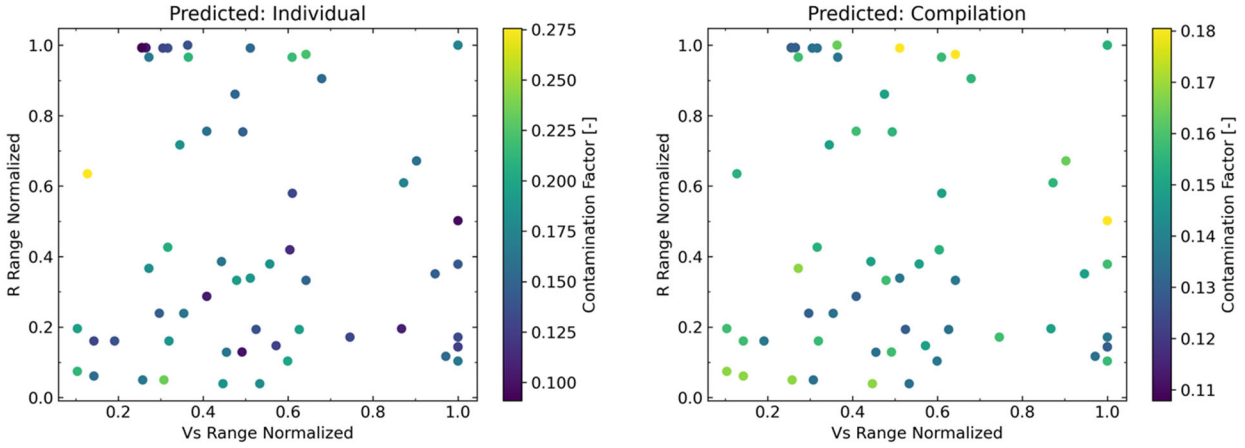
Anomaly detection was performed on the Grand Island dataset using the subsurface *Vs* and ρ data to locate potentially problematic zones. The analysis was performed on the thirteen compilation datasets to see how location on the levee, season, and site affected the detection of anomalies within the subsurface. Anomaly detection was conducted by fitting the algorithm on individual profiles and predicting anomalies in each profile, by fitting the algorithm with the entirety of each compilation dataset and predicting on the entire compilation dataset, and by fitting on the compilation dataset and predicting each profile individually. For IF, the number of estimators used

was 50. The LOF algorithm was built with 20 nearest neighbors with the Minkowski metric and a p of 2 which results in the Euclidean distance being utilized by the algorithm. For OCSVM, γ was set to 'scale'. For a full list of the model parameters used for each profile, refer to Appendix F.

4.7.1 Factors Affecting the Contamination Factor

Since the anomaly detection algorithms are sensitive to the contamination factor, multiple factors were analyzed to see if they had an impact on the contamination factor including *Vs* and ρ, site, season, and location on the levee. Fitting on the compilation datasets and predicting the individual profiles resulted in higher contamination factors ranging from 0.100 to 0.275 compared to predicting on the compilation datasets which resulted in ranges of 0.11 to 0.18.

To look at how the *Vs* and ρ affected the contamination factor, the *Vs* and ρ were first normalized to the compilation dataset and then plotted as shown in Figure 4.33.

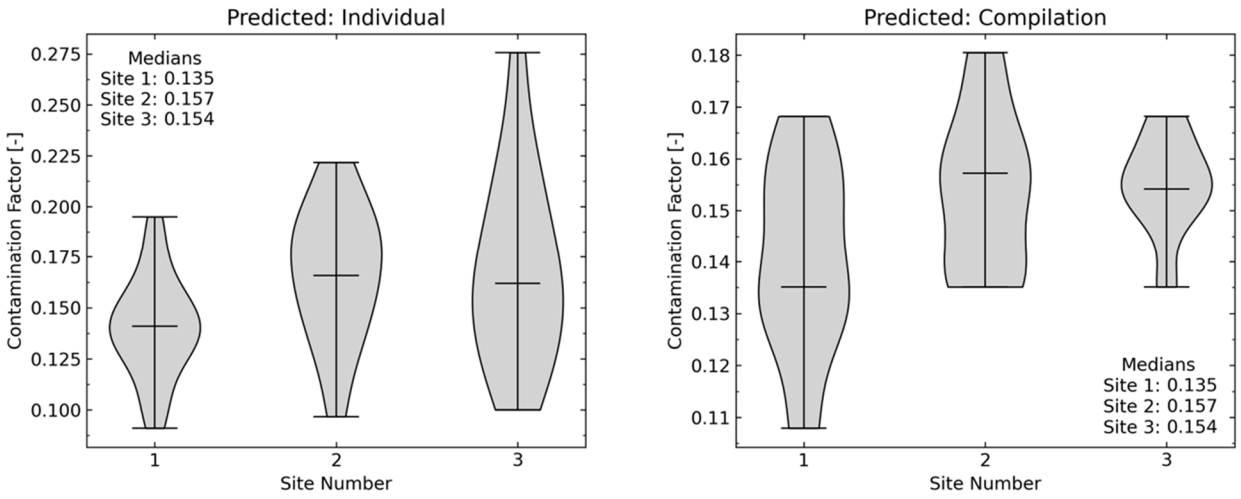


(a) (b)

Figure 4.33 Effect of shear wave velocity and electrical resistivity on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation.

The two figures represent whether the algorithms predicted the profiles individually or if they predicted the profiles as part of the compilation dataset. The different colors represent the range of contamination factors where yellow represents higher contamination factors and purple represents the lower contamination factors. From both figures, there was no clear trend between the contamination factor and the *Vs* and ρ.

Another factor that was analyzed was whether the site number along the levee resulted in higher or lower contamination factors. In this case, if Site 1 had higher contamination factors compared to Site 2 or Site 3, there could be a physical reason behind this difference. To look at this trend, violin plots were created and shown in Figure 4.34 for each site showing the distribution of the contamination factor at that site.



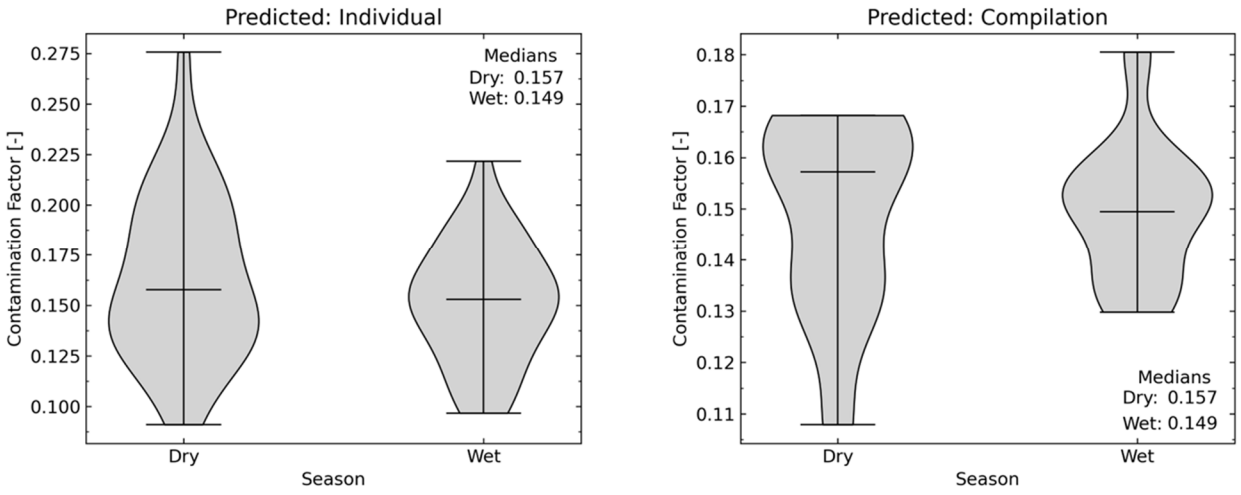
(a) (b)

Figure 4.34 Effect of site on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation.

Figure 4.34(a) shows the results for when the data was fitted on a compilation dataset but predicted on the individual profiles. There is little difference in the median between the different sites with Site 1 having a median of 0.135, Site 2 having a median of 0.157, and Site 3 having a median of 0.154. The distributions between Site 1 and Site 2 are similar where they taper off near the bottom and top of the distribution with most of the data near the mean. The distribution of Site 3 has most of the data around the mean but tapers off at much higher contamination factors than Site 1 and Site 2.

Figure 4.34(b) shows the results for when the data was fitted on a compilation dataset and predicted on the entire compilation dataset. In this case, Site 2 and Site 3 had little difference in the median with 0.157 and 0.154, respectively. Site 1 had a slightly lower median of 0.135 and a much larger range in contamination factor from 0.11 to 0.17. Site 2 and Site 3 had minimum contamination factors of 0.135. Based on the results from predicting the individual profile and the compilation dataset, there was no clear trend between the site number and the contamination factor.

The season when the data was collected was also analyzed to see if it impacted the contamination factor. The distribution of contamination factors for each season was plotted as violin plots in Figure 4.35.



(a) (b)

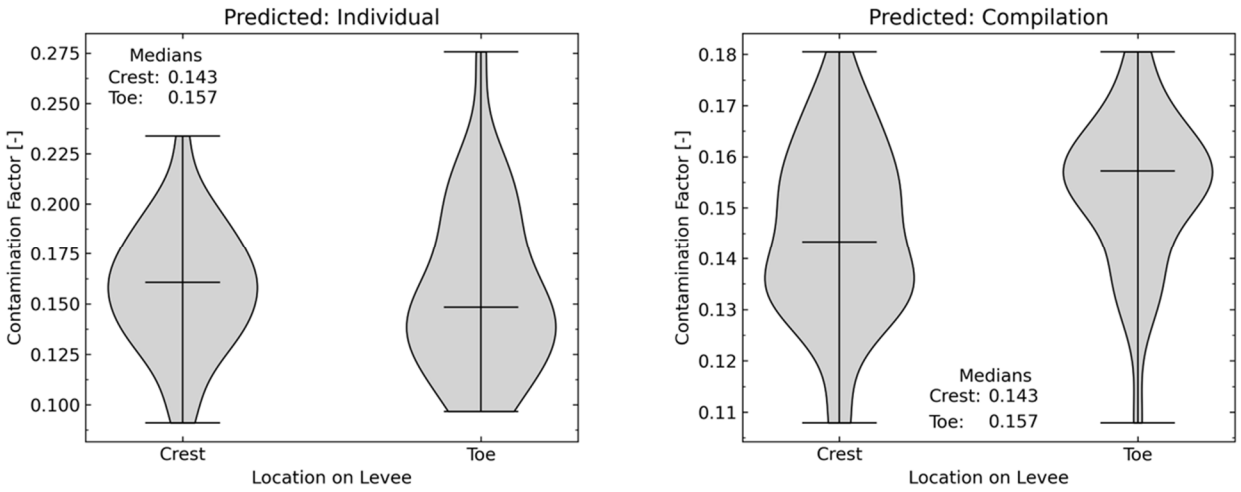
Figure 4.35 Effect of season on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation.

Figure 4.35(a) shows the results for when the data was fitted on a compilation dataset but predicted as individual profiles. The distribution of contamination factor shows similar shapes where most of the data falls around the median of 0.157 and 0.149 for the dry and wet seasons, respectively. The distributions both taper as the contamination decreases towards 0.100 and as the contamination increases above the median. The dry season has slightly higher contamination factors upwards of 0.275 while the wet season maxes out at 0.225. Overall, there is no discernable trend.

Figure 4.35(b) shows the results for when the data was fitted and predicted on the compilation datasets. The distributions of contamination factors are widely different between the dry and wet seasons. The dry season has a median of 0.157 with most of the data centered around it. The distribution maxes out around 0.17 but the minimum extends much farther from the median and thins out to 0.11. In comparison, the wet season has a median of 0.149 with most of the data centered around the median. The distribution of contamination factors is much higher at 0.18 but significantly thinner with a minimum of 0.13. These results show that the dry season had more variation in the contamination factor compared to the wet season.

The final factor analyzed to see the impact of the contamination factor was the location of the profile on the levee. For this analysis, the distribution of contamination factors was plotted in Figure 4.39 for profiles collected on the crest of the levees and profiles collected on the toe of the levees.

Figure 4.36(a) shows the results for when the algorithm fitted on the compilation dataset and predicted on the individual profiles. The distribution of the contamination factors for the crest and toe profiles were very similar with medians of 0.143 and 0.157 for the crest and toe, respectively. Both distributions had most of the data centered around the median and tapered off as the contamination factor increased.



(a) (b)

Figure 4.36 Effect of location on the levee on the contamination factor for (a) profiles that were fitted on a compilation and predicted individually and (b) profiles that were fitted on a compilation and predicted as a compilation.

For the results from fitting and predicting on the compilation dataset, the distributions of contamination factors are similar. Most of the data is centered around the median and tapers off as the contamination factor increases and decreases. The distribution near the minimum is thinner than the distribution near the maximum. There is no significant difference in the distribution of contamination factors between the crest and toe profiles.

4.7.2 Anomaly Detection Analysis

The anomaly detection algorithms classified each point using a binary classification scheme. Each data point was given either a 0, which represents a normal data point, or a 1, which represents an anomalous data point. All anomaly detection profiles can be found in Appendix H.

Each profile was fit on the individual profile and predicted on the individual profile, fit on compilation and predicted compilation, and fit compilation and predicted individual. These profiles were compared to the *Vs* and ρ profiles to determine how well the anomaly detection algorithms performed. Each profile was given one of four numerical ratings ranging from 0 to 3 depending on performance:

- Excellent (0) – performed well with no noise
- Good (1) – performed well but with some noise
- Fair (2) – very noisy but anomalies are slightly recognizable
- Poor (3) – very noisy where anomalies are not recognizable

To understand which algorithm performed the best, the assigned performance values assigned to each profile were normalized, where closer to 1 represented poor performance and closer to 0 represented perfect performance. The IF algorithm performed the best with a normalized score of 0.18 with OCSVM performing the second best with a normalized score of 0.25 and LOF

performing the worst with a normalized score of 0.85. Table 4.8 summarizes the percentage of profiles classified into each performance category for each anomaly detection algorithm.

Table 4.8 Percentage of Profiles Classified in Each Performance Category

	Algorithm		
Performance	**IF**	**LOF**	**OCSVM**
Excellent	67	1	47
Good	20	8	37
Fair	5	24	11
Poor	8	67	6

For IF, 88% of the profiles were classified as excellent or good while OCSVM had 83% of the profiles classified as excellent or good and LOF had only 9% of the profiles classified as excellent or good. LOF was mostly classified as fair and poor with 91% of the profiles not performing well.

4.7.2.1 Factors Affecting Anomaly Detection Performance

Predicting on Individual Profiles or Compilation Datasets

Many factors could affect the performance of the anomaly detection results including predicting on the individual or compilation datasets, the contamination factor, site, location on the levee, season, and the Vs and ρ range of the profiles. To understand how predicting on individual profiles or compilation datasets affects performance, the performance was normalized among all sixteen subsurface profiles for each combination of predicting on individual/compilation for IF and OCSVM. The results are shown in Table 4.9 where a value closer to 0 represents perfect performance and a value closer to 1 represents poor performance.

Table 4.9 Performance of Predicting on Individual and Compilation Datasets

		Algorithm	
Fit on:	**Predicted on:**	**IF**	**OCSVM**
Individual	Individual	0.05	0.20
Crest / Toe	Crest / Toe	0.22	0.28
Crest / Toe	Individual	0.08	0.09
Dry / Wet	Dry / Wet	0.31	0.27
Dry / Wet	Individual	0.05	0.09
Dry / Wet (Site)	Dry / Wet (Site)	0.17	0.19
Dry / Wet (Site)	Individual	0.06	0.11
Site	Site	0.22	0.30
Site	Individual	0.03	0.17

Regardless of the anomaly detection algorithm and the compilation being fitted, fitting on the compilation dataset and predicting on the individual profiles performed better than when fitting on the compilation dataset and predicting on the compilation dataset. When predicting on the individual profiles, more local anomalies were detected but when predicting on the compilation profiles only global anomalies were detected. A higher number of profiles that were predicted

individually performed better than their predicted compilation counterpart. As a result, predicting the individual profile results in more accurate anomaly detection results for local anomalies.

The performance results of each profile were also used to determine if one compilation type or one profile performed significantly better or worse compared to the others. The results in Table 4.8 showed that no one compilation dataset performed significantly better for both algorithms. When looking at profile performance, no profile performed significantly better or worse than other profiles.

Contamination Factor

Since the contamination factor is the most important model parameter, it would be expected that the contamination factor would affect the performance of the anomaly detection results. To understand if there is a relationship between the contamination factor and the performance, the distribution of the contamination factor was plotted in Figure 4.37 for IF and OCSVM as well as for predicted individual profiles versus compilation profiles. LOF was not considered for this analysis as over 90% of the data performed poorly. The corresponding figures showing the effect of the contamination factor on performance for LOF can be found in Appendix H.

Figures 4.37(a) and 4.37(b) show the results from IF while Figures 4.37(c) and 4.37(d) show the results from OCSVM. Overall, there is a trend where the profiles with excellent and good performance had a higher range of contamination factors than the profiles with fair and poor performance. Among all the algorithms and prediction methods, there is very little difference in the median of contamination factors between the different rating classes.

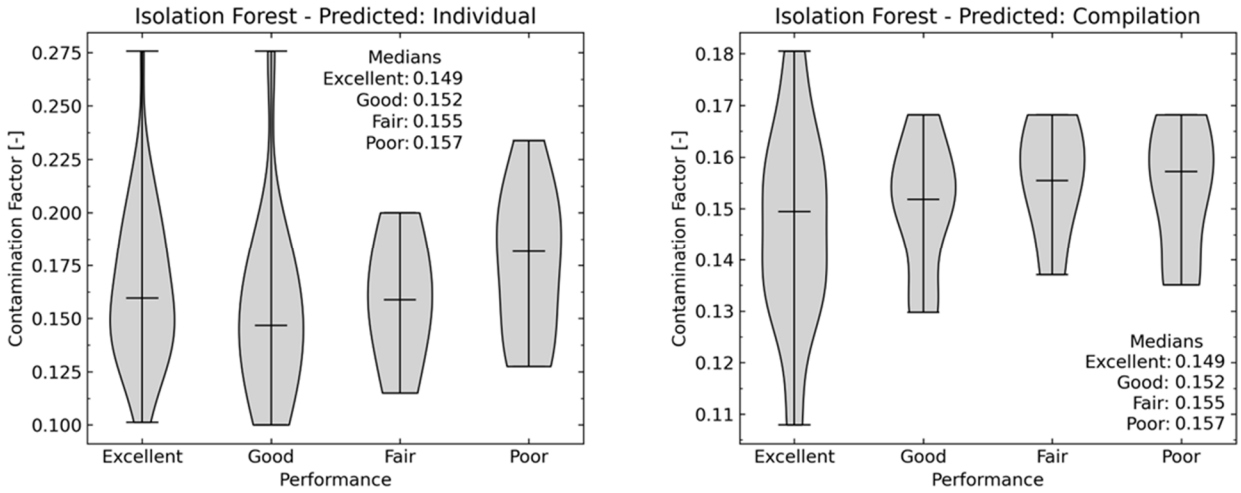
Another main difference between the excellent and good classes compared to the fair and poor classes is the shape of the distributions. In Figure 4.37, the distributions of the fair and poor classes were relatively uniform with insignificant tapering at either end. In comparison, the excellent and good classes had most of the data centered around the mean with significant tapering at the minimum and maximum contamination factors.

Site Number

Different segments of levees could potentially perform differently. To understand how this affected the anomaly detection performance, the performance of each profile in each site was summed separately for predicting the individual profiles and predicting the compilation dataset as shown in Table 4.10. The lower the score, the better the performance as more profiles were given an excellent or good rating. Like with the contamination factor analysis on performance, LOF was not included in the analysis as most of the profiles performed poorly.

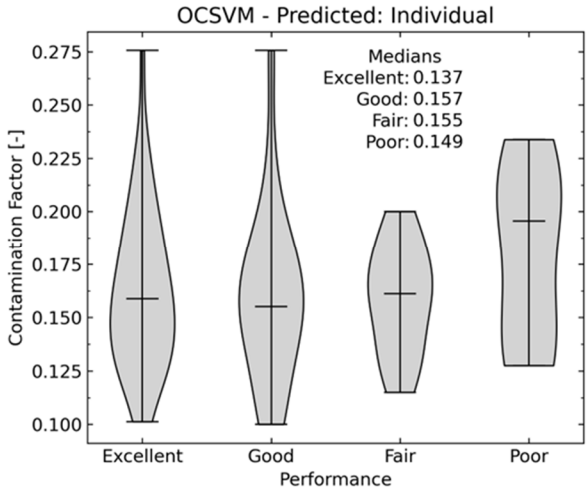
Table 4.10 Effect of Site on Anomaly Detection Performance

	Isolation Forest		One-Class Support Vector Machines	
Site	Individual	Compilation	Individual	Compilation
1	4	9	8	18
2	5	23	12	20
3	4	27	10	28

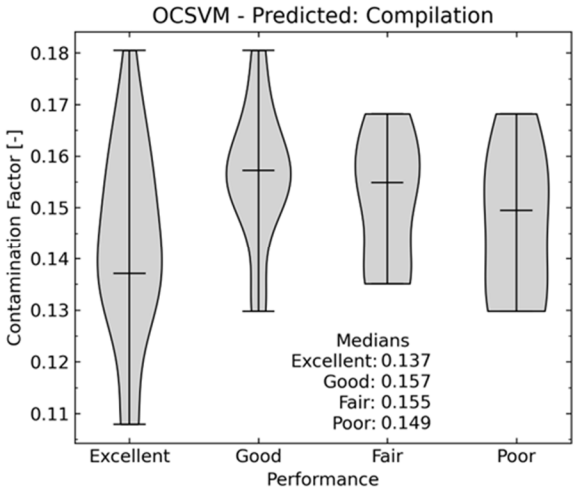


(a)

(b)



(c)



(d)

Figure 4.37 Effect of contamination factor on performance for (a) fitted on compilation dataset and predicted on individual profiles for Isolation Forest, (b) fitted and predicted on compilation dataset for Isolation Forest, (c) fitted on compilation dataset and predicted on individual profiles for One-Class Support Vector Machines, and (d) fitted and predicted on compilation dataset for One-Class Support Vector Machines.

Overall, Site 1 performed the best regardless of predicting on individual or compilation and regardless of the anomaly detection algorithm. IF for Site 1 produced a total performance score of 4 for predicting on individual and a score of 9 for predicting on compilation datasets. These scores are lower than Site 2 and lower than Site 3 for compilation. OCSVM for Site 1 produced a total performance score of 8 for predicting on individual and 18 for predicting on compilation which were lower than both Site 2 and Site 3 for individual and compilation. This could be due to Site 1 having the highest variety of Vs and ρ amongst the three sites.

Site 2 performed better than Site 3 when predicting the compilation dataset for both IF and OCSVM. Otherwise, Site 3 performed better than Site 2 when predicting the individual dataset for both IF and OCSVM.

Profile Location on Levee (Crest/Toe)

To explore whether the location of the profile on the levee had any effect on performance, the performances for all the profiles on the crests and toes were summed separately for IF and OCSVM as well as if predicted as an individual profile or as a compilation dataset. The results are shown in Table 4.11. The closer the value to 0, the better the performance. Due to LOF being classified as mostly fair and poor performance, it was excluded from this analysis.

Table 4.11 Effect of Profile Location on Anomaly Detection Performance

	Isolation Forest		One-Class Support Vector Machines	
Location	**Individual**	**Compilation**	**Individual**	**Compilation**
Crest	2	26	12	34
Toe	12	33	18	32

When predicting the individual profiles, the profiles surveyed on the crests of the levees performed better than the profiles collected on the toes. The summed performance for crest surveys using IF when predicting individual profiles was 2 while the toe was 12. OCSVM had a summed performance of 12 for crest surveys while the toe had a total performance of 18. The crest profiles should be more uniform in subsurface properties than the toe profiles and since predicting on the individual profiles is better at finding local anomalies, finding anomalies in a more uniform subsurface is easier than in a heterogenous subsurface.

When predicting the compilation dataset, there was no clear trend on whether crest surveys or toe surveys performed better. The summed performance for crest profiles when using IF was 26 while the toe surveys was 33 suggesting that the crest surveys perform better than the toe surveys. OCSVM contradicted this finding when the crest surveys had a summed performance of 34 and the toe surveys had a summed performance of 32.

Season

The season when the data was collected was also analyzed to see if it affected the anomaly detection performance. This was done by summing the performance values for each profile in the wet season for IF and OCSVM separately, as well as for predicting individual profiles and

compilation profiles. The same process was repeated for the dry season profiles. The results are summarized in Table 4.12. LOF was removed from this analysis since most of the profiles were classified as fair and poor.

Table 4.12 Effect of Season on Anomaly Detection Performance

	Isolation Forest		One-Class Support Vector Machines	
Season	**Individual**	**Compilation**	**Individual**	**Compilation**
Dry	4	41	13	36
Wet	10	18	17	30

When predicting the individual profiles, the dry season performed better than the wet season regardless of the algorithm. For IR, the summed performance was 4 for the dry season while the wet season had a summed performance of 10. OCSVM had a summed performance of 13 for the dry season while the wet season had a summed performance of 17. Similar to the crest and toe results, the dry season is expected to have more uniform subsurface ρ since there is little additional moisture from rainfall, making the detection of anomalies easier in a more homogeneous subsurface.

This trend does not hold when the profiles were predicted using the compilation dataset. Instead, the trend is the opposite where the wet season performed better than the dry season regardless of the algorithm. The summed performance of the dry season for IF was 41 while the wet season was 18. Similarly, the summed performance for the dry season for OCSVM was 36 while the wet season was 30.

Shear Wave Velocity and Electrical Resistivity

The final factors analyzed to see the effect on the anomaly detection performance were *Vs* and ρ. To look at how this affected performance, the *Vs* and ρ were first normalized to the range of *Vs* and ρ associated with the compilation dataset they were fitted with. The results are plotted in Figure 4.38 showing the distribution of performance ranking amongst the normalized *Vs* and ρ range for IF and OCSVM. LOF was not included in the analysis as 91% of the profiles were classified as fair or poor performance and showed no trend. The corresponding figures for LOF showing the effect of *Vs* and ρ on anomaly detection performance can be found in Appendix H.

For predicting individual profiles for both IF and OCSVM, there was no discernable trend that appears concerning performance, *Vs*, and ρ. Most of the data is classified as either excellent or good and no trend appears showing clear clusters of a performance in a certain location on the plots.

In comparison, when predicting using the compilation datasets, a trend appears between well-performing profiles and poorly performing profiles. The profiles that were classified as poor tend to plot below 0.6 normalized *Vs* and below 0.5 normalized ρ. The profiles that were classified as fair, plot at all ranges of normalized *Vs* but never surpassed 0.7 normalized ρ. The profiles that were classified as excellent and good plot throughout the total range of normalized *Vs* and ρ but tend to be less dense with normalized *Vs* below 0.3.

In summary, profiles with higher normalized ranges of *Vs* and ρ are more likely to perform better than profiles with lower normalized *Vs* and ρ. Profiles that do tend to have lower normalized *Vs* and lower ρ have the potential to perform poorly but can perform well.

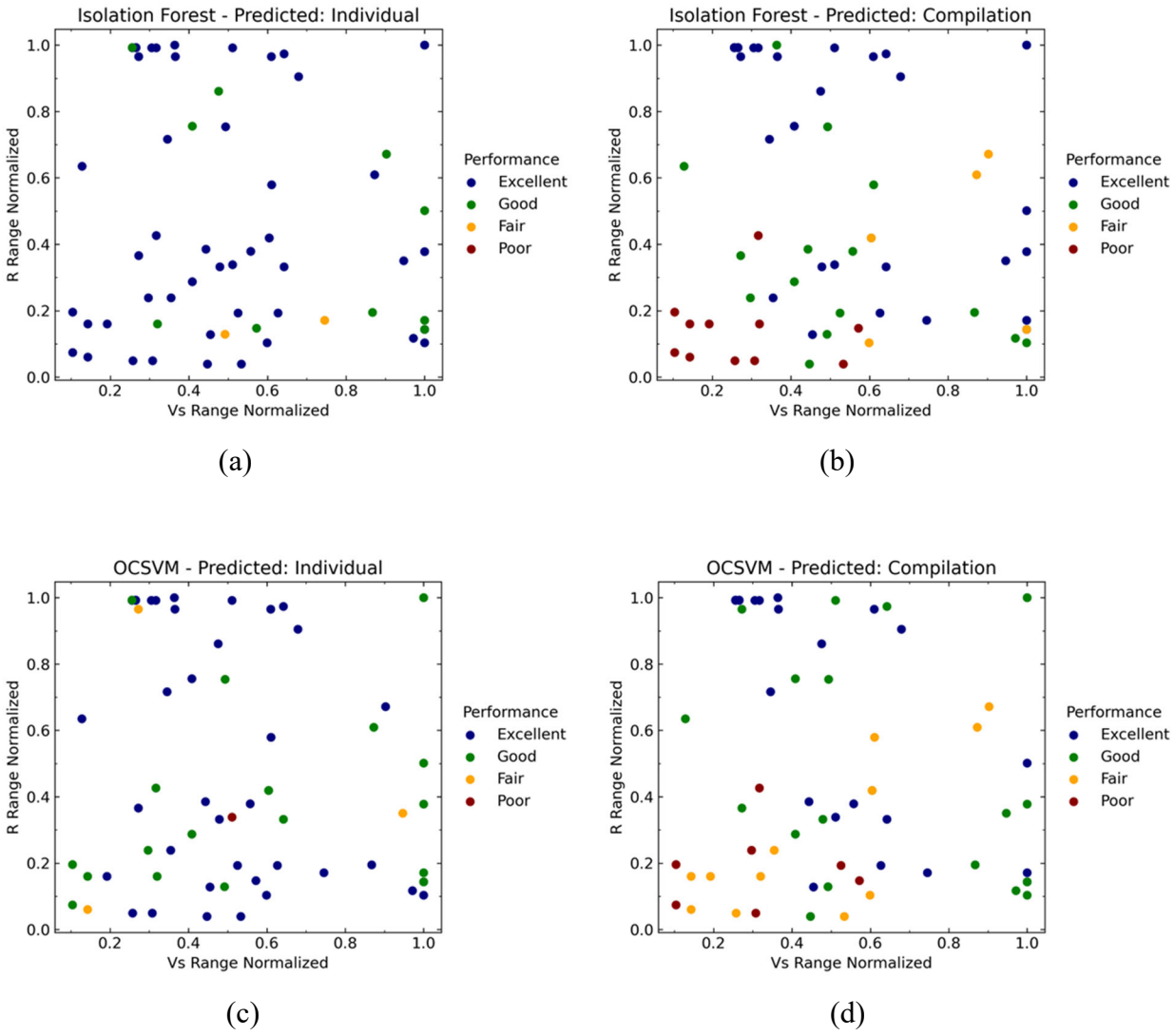


Figure 4.38 Effect of shear wave velocity and electrical resistivity on performance for (a) fitted on compilation dataset and predicted on individual profiles for Isolation Forest, (b) fitted and predicted on compilation dataset for Isolation Forest, (c) fitted on compilation dataset and predicted on individual profiles for One-class Support Vector Machines, and (d) fitted and predicted on compilation dataset for One-class Support Vector Machines.

4.8 Combination of Anomaly Detection and Clustering

Unsupervised anomaly detection and clustering techniques were applied to the subsurface data at multiple locations along the levees during two seasons. After analyzing the results from the clustering and anomaly detection analyses on the subsurface data, it was determined that no one unsupervised method was independently suitable for the application of understanding levee subsurface conditions.

4.8.1 Advantages and Disadvantages of Clustering

For the application of levee subsurface health investigation, unsupervised clustering had some advantages and disadvantages. One of the advantages of clustering was that it excelled at clustering the data based on *Vs* and ρ as shown in Figure 4.39.

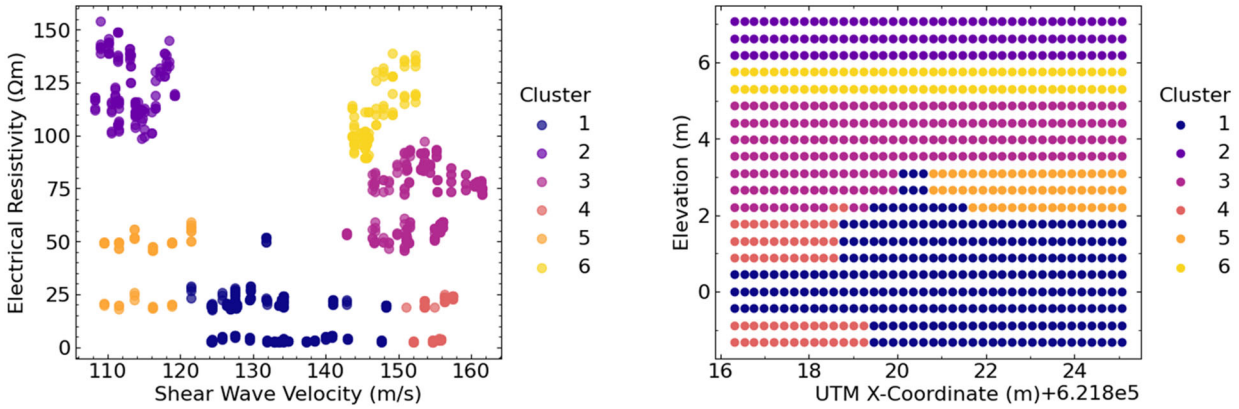


Figure 4.39 Clustering example (a) with respect to shear wave velocity and electrical resistivity and (b) subsurface profile showing the spatial distribution of clusters

In Figure 4.39(a), the clusters are represented by different color data points as a function of *Vs* and ρ. In this example, six clusters are present: (1) *Vs* = 110 m/s and ρ = 125 Ωm, (2) *Vs* = 110 m/s and ρ = 30 Ωm, (3) *Vs* = 135 m/s and ρ = 25 Ωm, (4) *Vs* = 150 m/s and ρ = 100 Ωm, (5) *Vs* = 155 m/s and ρ = 75 Ωm, and (6) *Vs* = 155 m/s and ρ = 20 Ωm.

The clusters can also be analyzed spatially in the subsurface, as shown in Figure 4.39(b). Understanding where clusters are located spatially is useful for understanding subsurface health as it can show areas with anomalously high or low *Vs* and ρ. In the profile shown in Figure 4.39(b), Cluster 1 is mostly found at depth while Cluster 2 is near the top of the profile. Cluster 6 is found between 5 to 6 m in elevation along the entire length of the profile. Cluster 3 is just below Cluster 6 and extends to 3 to 2 m in elevation. Cluster 4 is found around an x-coordinate of 621817 m at an elevation of 1 m and -1 m. Cluster 5 is found between 621820 m and 621825 m at an elevation of 3 m.

While clustering was beneficial in finding the natural groupings of the data based on *Vs* and ρ, it was unable to determine potentially problematic zones. Without significant experience with *Vs* and ρ measurements, stakeholders would be unable to identify potentially problematic zones using *Vs* and ρ alone.

4.8.2 Advantages and Disadvantages of Anomaly Detection

While unsupervised clustering techniques were unable to determine potentially problematic zones in the subsurface, unsupervised anomaly detection could. The purpose of anomaly detection is to classify data points as normal or anomalous. As a result, groupings of anomalous data points in the subsurface can be considered a potentially problematic zone as shown in Figure 4.40.

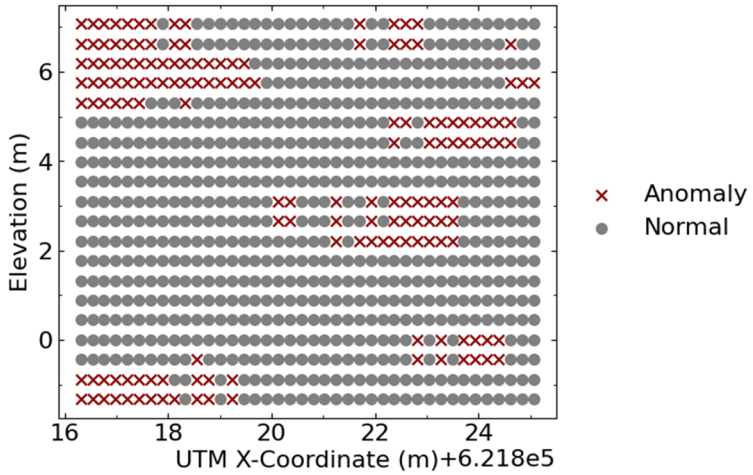


Figure 4.40 Subsurface profile showing anomaly detection results as anomalous zones.

This shows a subsurface profile with the binary anomaly detection classification. Data points shown as grey circles were considered "normal" while data points represented as red x's were classified as "anomalies." From Figure 4.40, clear zones of anomalous data points are exhibited throughout the profile. With clear anomalous zones, stakeholders can easily identify areas that could be problematic to levee health.

One of the major disadvantages of anomaly detection for levee subsurface health is the inability to determine the *Vs* and ρ for one anomalous zone. Due to the binary classification, pulling out the *Vs* and ρ for one zone only is difficult. To understand what is potentially problematic about the area, understanding the trends of the *Vs* and ρ is critical.

4.8.3 Framework for Understanding Subsurface Levee Health

Both clustering and anomaly detection had many advantages, but their disadvantages made it difficult to use them independently for levee subsurface health investigation. Clustering techniques subdivided the subsurface into zones of similar *Vs* and ρ while anomaly detection made it difficult to pull out *Vs* and ρ for one anomalous zone because of the binary classification. However, anomaly detection could detect potentially anomalous zones in the subsurface while it was difficult to interpret which clusters would be considered potentially problematic. Since these two unsupervised machine learning methods complement each other, the clustering and anomaly detection results were fused into one profile to understand the health of the subsurface.

The first step in fusing the data was to remove noise in the anomalous profiles. It was assumed that within the subsurface, a potentially anomalous zone would not occur if there was a single anomalous data point. To remove these single data points, a threshold was set to determine if the location was spatially significant. This threshold was set experimentally using a nearest-neighbors methodology.

Let N denote the total number of data points (either normal or anomalous) surrounding an anomalous data point and let M represent the number of data points within N that are classified as anomalous. If the anomalous data point had less than M points surrounding it, then the anomalous data point was reclassified as normal. For this analysis, the values of N ranged from 2 to 8 in increments of 2 and the values of M ranged from 2 to 7 in increments of 1. Every combination of N and M was analyzed to determine the optimal set of parameters. A diagram showing examples of different combinations of N and M is shown in Figure 4.41.

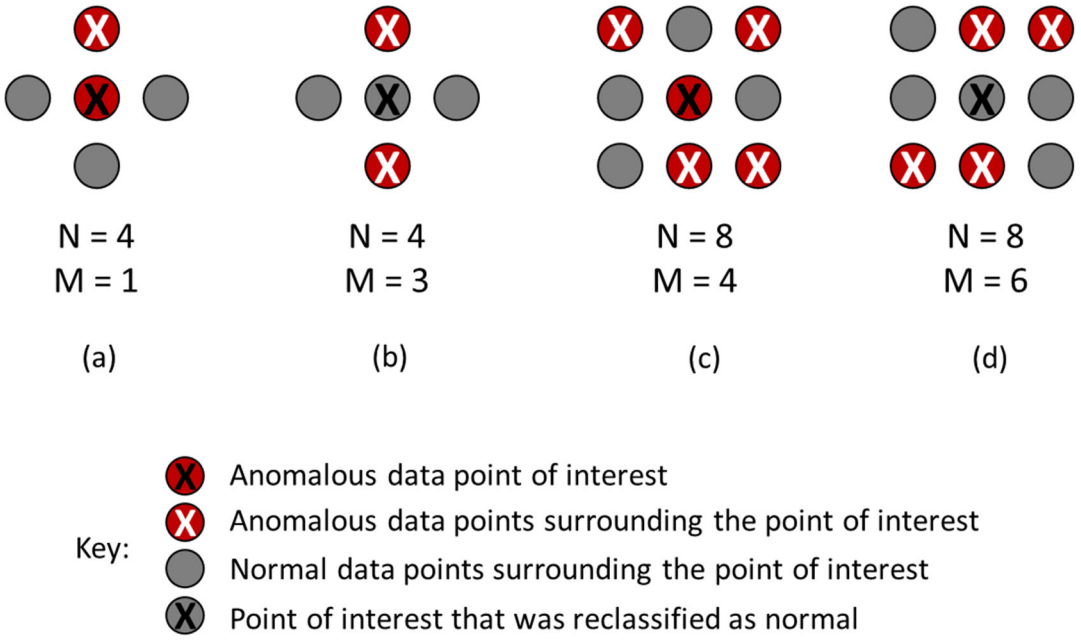


Figure 4.41 Example of thresholds for determining if an anomalous point is spatially significant when (a) N = 4 and M = 1, (b) N = 4 and M = 3, (c) N = 8 and M = 4, and (d) N = 8 and M = 6

In these diagrams, the anomalous point that is being explored to determine if it is spatially significant is denoted as a red circle with a black X and if it was reclassified as normal, it is denoted as a grey circle with a black X. The red circles with the white X represent anomalous points surrounding the point of interest and the grey circles represent normal data points surrounding the point of interest. In example (a), N = 4 which is shown by four points surrounding the point of interest. When M is set to 1, then the anomalous point of interest would remain classified as anomalous since there is at least one other anomalous data point surrounding it. The example shown in Figure 4.41(b) had an N of 4, which is represented by the four points surrounding the point of interest, and an M of 3. Since only two anomalous points are around the point of interest, the point of interest will be reclassified as normal. In example (c), N was set to 8, which is shown

by the eight points surrounding the point of interest. The M value was set to 4. In the example, since there are four anomalous points within the eight surrounding points, the point of interest will remain classified as anomalous. In example (d), N was set to 8 and M was set to 6. Since there are fewer than six anomalous data points surrounding the point of interest, it was reclassified as normal.

After experimentally analyzing every combination of N and M, the optimal set of parameters was set to N = 8 neighbors and M = 2 anomalies. The figures that accompanied this analysis can be found in Appendix I.

Once the noise filtering threshold was applied to the anomaly detection profile, the next step in the framework was to spatially join the clusters from the clustering analysis to the potentially anomalous zones from the anomaly detection analysis. An example of this spatially joined profile is shown in Figure 4.42.

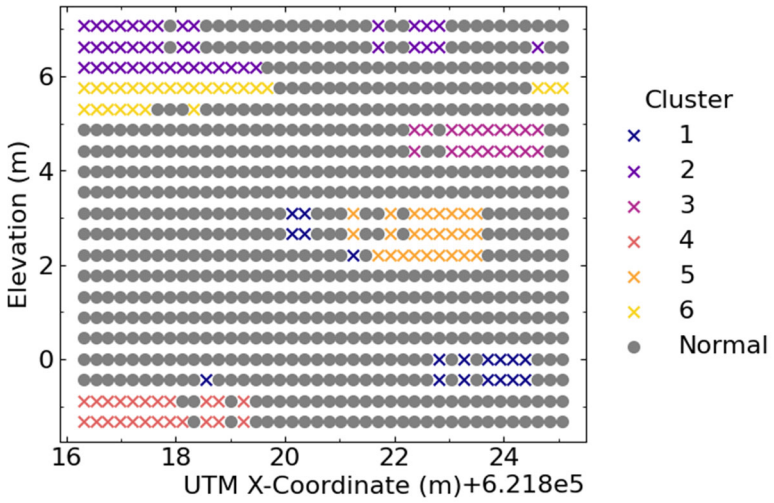


Figure 4.42 Subsurface profile combining the clustering results with the anomaly detection results showing clusters in the anomalous zones.

In this subsurface profile, the grey circles represent the normal data points, the x's represent potentially problematic locations, and the colors of the x's represent the clusters. This profile is powerful as it now shows the anomalous zones, but it also shows which clusters are associated with each zone. While it is important to understand where the clusters are located spatially in the subsurface, it is more beneficial to understand the *Vs* and ρ in the subsurface. This can be done by pulling the *Vs* and ρ out from each cluster and comparing it to the entire profile to understand how *Vs* and ρ change in an anomalous zone.

For the clusters that corresponded to an anomalous zone, the mean (centroid) of the *Vs* and ρ was found. It was then compared to the mean of the *Vs* and ρ of the entire profile. New labels were given to the anomalous zones depending on whether the cluster mean of the *Vs* and ρ increased, decreased, or had no significant change compared to the rest of the profile. The threshold for significant change was determined experimentally by comparing the +/- 1 standard deviation from the mean, 1% change, 5% change, 10% change, and 25% change.

Three case studies were conducted on different profiles to understand how these different levels of change affected the subsurface profile. The first case study was conducted on the crest survey at Site 2 on Grand Island during the wet season where anomaly detection fit and predicted on the profile individually. For this site, the 1 standard deviation, 1%, 5%, 10%, and 25% change were calculated for *Vs* and ρ and are shown in Table 4.13.

Table 4.13 Shear Wave Velocity and Electrical Resistivity Change for Case Study 1

Threshold	Shear Wave Velocity (m/s)	Electrical Resistivity (Ωm)
1 Standard Deviation	15.9	43.77
1% Change	1.37	0.56
5% Change	6.84	2.82
10% Change	13.68	5.64
25% Change	34.20	14.09

For this case study, the 10% change in *Vs* is close to the 1 standard deviation in *Vs*. No change percentage was close to 1 the standard deviation for ρ. The histograms showing the distribution of ρ and *Vs* for each change threshold can be found in Appendix I. For each of these thresholds, the spatial profile was plotted showing the anomalous zones which were colored based on the change in *Vs* and ρ based on the change threshold and can be found in Appendix I.

Each change threshold subsurface figure was compared to the *Vs* and ρ distribution of the subsurface to determine which threshold most accurately represented the change in the subsurface. After examination, the 1 standard deviation threshold most accurately represented the change for Case Study 1.

This process was repeated for Case Study 2 and Case Study 3. Case Study 2 was conducted on the wet toe profile at Site 1 on Grand Island during the dry season. This profile was from the Toe compilation profile. Case Study 3 was conducted on the crest profile at Site 2 on Grand Island during the dry season from the Site 2 compilation dataset. The 1 standard deviation, 1%, 5%, 10%, and 25% change for *Vs* and ρ for Case Study 2 and Case Study 3 are summarized in Table 4.14.

Table 4.14 Shear Wave Velocity and Electrical Resistivity Change for Case Study 2 and Case Study 3

	Shear Wave Velocity (m/s)	Electrical Resistivity (Ωm)
Threshold	**Case Study 2**	
1 Standard Deviation	15.9	43.77
1% Change	1.34	0.63
5% Change	6.71	3.16
10% Change	13.41	6.33
25% Change	33.54	15.82
	Case Study 3	
1 Standard Deviation	18.18	8.98
1% Change	1.72	0.64
5% Change	8.62	3.19
10% Change	17.23	6.38
25% Change	43.08	15.96

The histograms showing the change thresholds and the change threshold subsurface figures for Case Study 2 and Case Study 3 can be found in Appendix I. The results from these case studies agree with Case Study 1 where the +/- 1 standard deviation from the mean, models the significant change in the subsurface the best.

With the change threshold set to +/- 1 standard deviation from the mean of the entire subsurface profile, the anomalous zones were given one of three labels: (1) increase in *Vs* or ρ when the mean of the cluster was greater than +/ standard deviation, (2) decrease in *Vs* or ρ when the mean of the cluster was less than -1 standard deviation, and (3) similar or no significant change when the mean of the *Vs* or ρ fell within +/- 1 standard deviation from the mean. An example of the subsurface profile showing the change in *Vs* and ρ is shown in Figure 4.43.

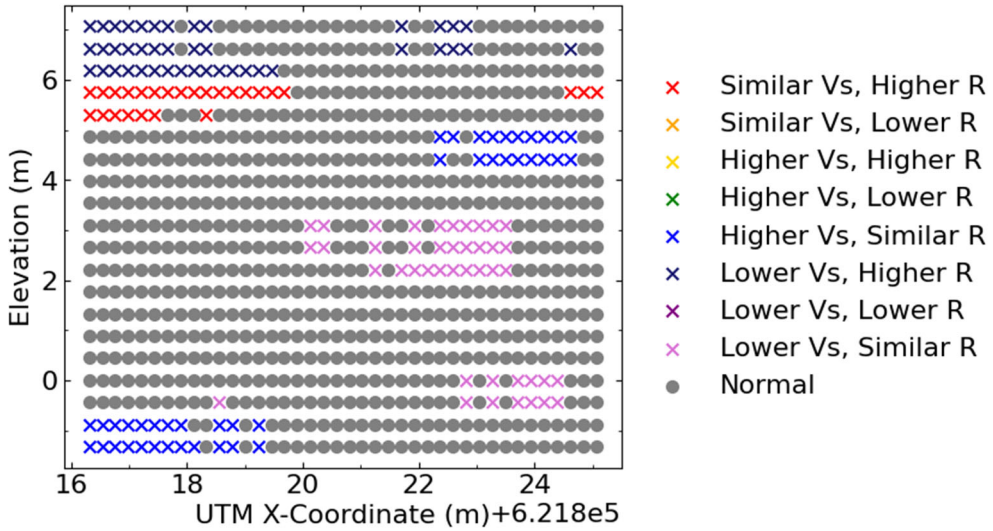


Figure 4.43 Subsurface profile showing the anomalous zones as a function of the relative change in shear wave velocity and electrical resistivity.

In this subsurface profile, the grey circles represent normal data points and the x's represent anomalous zones. The colors associated with the anomalous zones show all the combinations of increasing, decreasing, or no significant change between *Vs* and ρ. For understanding levee subsurface health, relative changes in *Vs* and ρ are important but can be difficult to interpret. To increase the interpretability of the profile, it is important to understand what geotechnical conditions in the subsurface cause similar trends in *Vs* and ρ.

The results from the laboratory analysis that were discussed in Section 2.7 were applied to the subsurface profile. The results are summarized in Table 2.3 showing how different geotechnical conditions increased/decreased/had no significant change in *Vs* and ρ in tandem. These laboratory results were then applied to the subsurface profile by matching the changes in *Vs* and ρ in the subsurface with the changes in *Vs* and ρ in Table 2.3. The subsurface profile shown in Figure 4.46 was changed to include the geotechnical condition labels instead of the relative changes in *Vs* and ρ. The updated subsurface profile is shown in Figure 4.44.

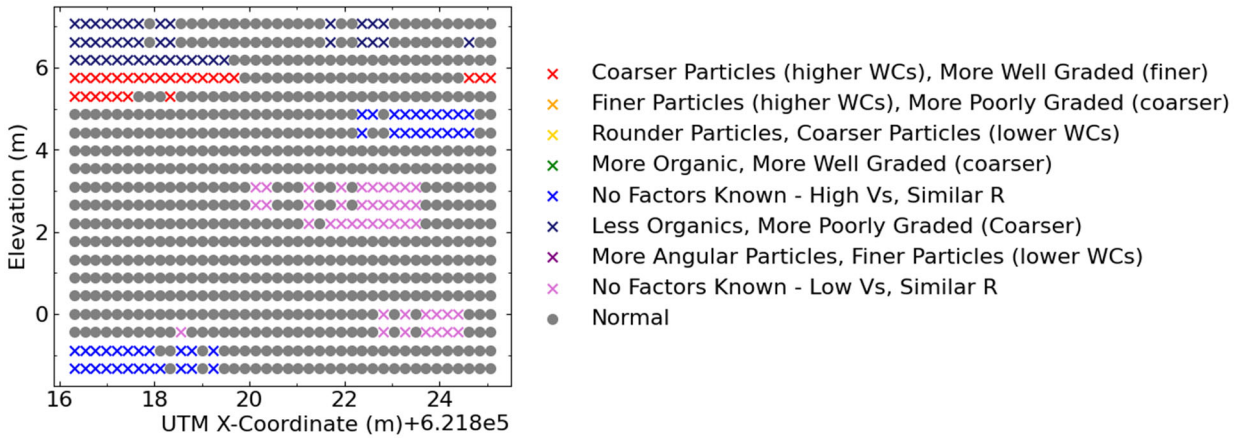


Figure 4.44 Subsurface profile showing the anomalous zones as a function of geotechnical conditions controlled for in the laboratory investigation.

This subsurface profile now shows the normal data as grey circles, the potentially problematic zones as x's, and the colors represent the geotechnical conditions that may exist at that location in the subsurface. These zones are now clearly delineated in the subsurface with understandable labels that are directly related to internal failure mechanisms, such as seepage. This profile is powerful as it fuses clustering results with anomaly detection results to create a profile that is easily interpretable to anyone who needs to make decisions about levee health. For all potentially problematic zone profiles at Grand Island, please refer to Appendix I.

4.8.4 Subsurface Validation at Grand Island

Subsurface validation is important for understanding how well the framework performs. At the time of planning the project, boreholes and CPTs were planned along the crest of the levee for all three sites at Grand Island to provide ground truth for validation of the method. Unfortunately, permits were never approved for the invasive testing, and ground truth data was never obtained. Future work for this framework includes performing this analysis in another location where ground truth data is readily available at a high spatial resolution to validate this method. This framework was applied to the Bacon Island dataset for validation and is discussed in more detail in Section 4.10.

4.9 Supervised Machine Learning Analysis on Bacon Island

Supervised machine learning was performed on the Bacon Island dataset to predict individual SBTn and to predict the SBTn for a given site using geophysical data. Multiple analyses were performed on this dataset to look at the effect of feature engineering, including feature selection, methods to reduce overfitting, and methods to reduce the effect of class imbalance. Table 4.15 summarizes each analysis results with the features used, the data being classified, and other analyses applied.

Table 4.15 Supervised Machine Learning Analyses Accuracy Results

Analysis	Classification	Features	Other	Training Dataset Size	KNN		RF		SVM	
					Training Accuracy	Testing Accuracy	Training Accuracy	Testing Accuracy	Training Accuracy	Testing Accuracy
1	Random 30%	All		3640	81	67	99	94	82	75
2	Random 30%	All	Stratified Randomization	3640	83	68	99	94	82	73
3	Random 30%	All	Sampling	4500	96	56	99	89	100	59
4	Random 30%	Vs and R		3640	68	57	88	66	74	58
5	Northwest	All		3200	87	21	99	24	82	21
6	Northcentral	All		3200	72	44	100	43	78	42
7	South	All		4000	75	18	99	30	86	9
8	Northwest	Vs and R		3200	71	22	90	22	76	23
9	Northcentral	Vs and R		3200	60	35	88	35	63	45
10	South	Vs and R		4000	69	23	87	23	75	43

4.9.1 Data and Feature Engineering

The data used in this analysis included the 2D *Vs* profiles from the MASW surveys and the 2D ρ profiles from the ERT surveys at the Northwestern Site, the Northcentral Site, and the South Site at Bacon Island. This analysis also leveraged the SBTn from the CPT data at these sites on Bacon Island to understand lithological variation in the subsurface.

The first step in feature engineering for this dataset was to create a uniform grid of points with known elevation and UTM 10N x-coordinates that overlapped with the MASW, ERT, and CPT data. Figures 4.45 through 4.47 shows longitudinal profiles of each site and how the MASW survey, ERT survey, and CPT overlap while Figures 4.48 through 4.50 shows the cross-sectional profiles of each site with locations of the CPT data.

The first step in creating the uniform grid was to determine where the geophysical surveys overlapped with the CPT data. The x-coordinates and y-coordinates that fell within the lateral extents of the geophysical surveys and the elevation extents of the CPT data were used to create the grid. The MASW, ERT, and CPT data were interpolated to each point in the grid to create the dataset using nearest neighbors. Figure 4.51 shows the interpolated CPT profiles for the three Bacon Island sites.

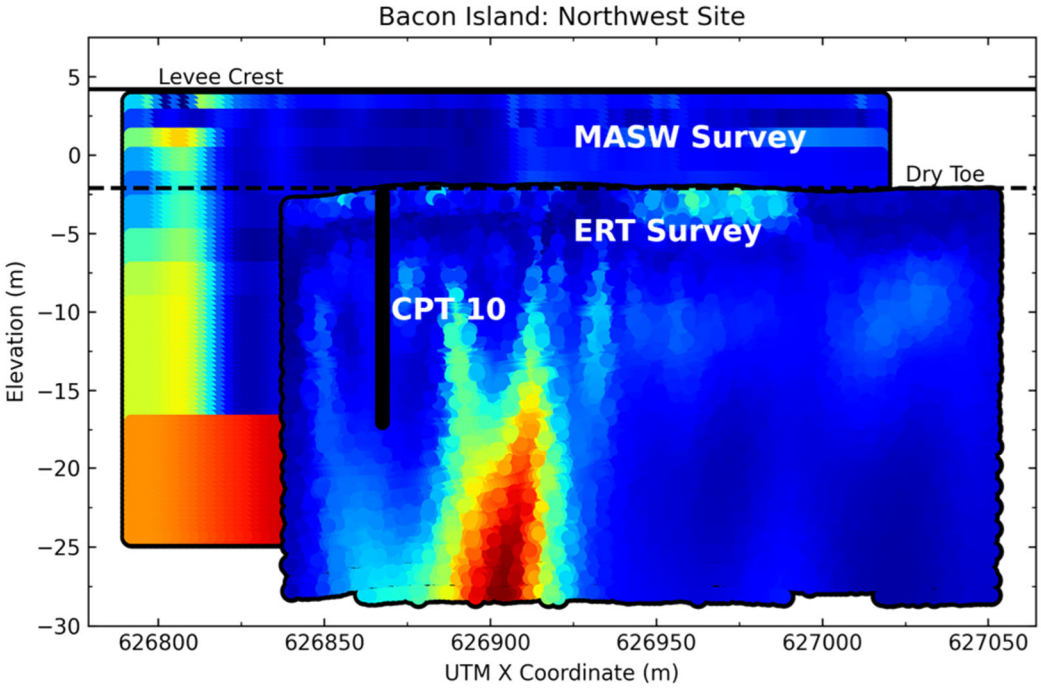


Figure 4.45 Longitudinal profile of the Northwest Site on Bacon Island.

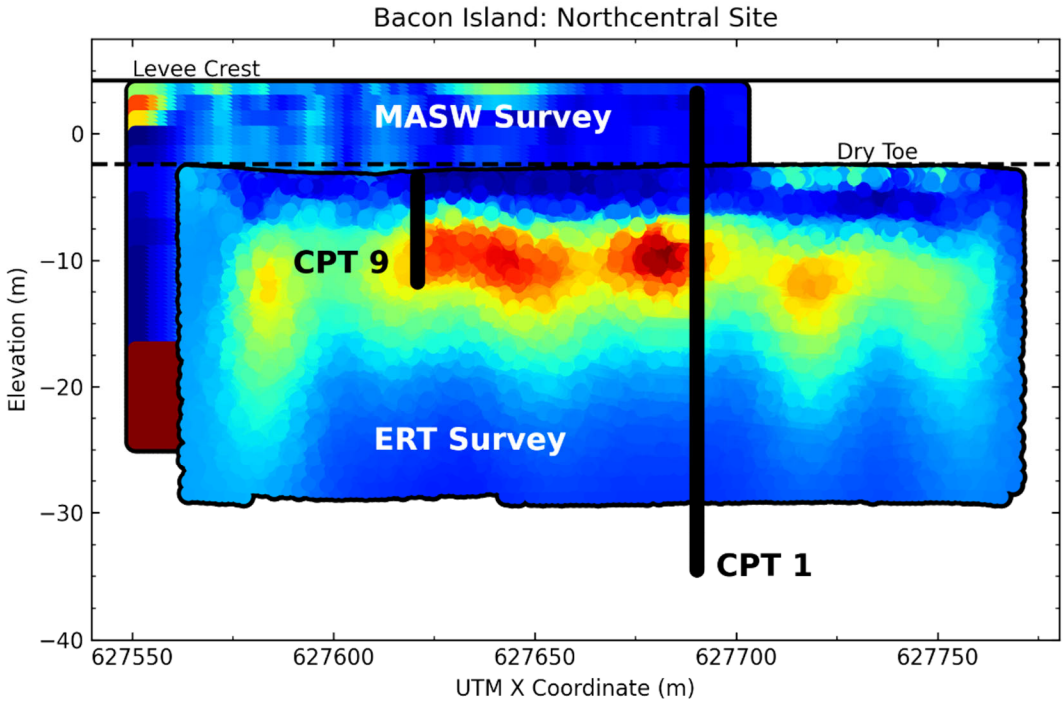


Figure 4.46 Longitudinal profile of the Northcentral Site on Bacon Island.

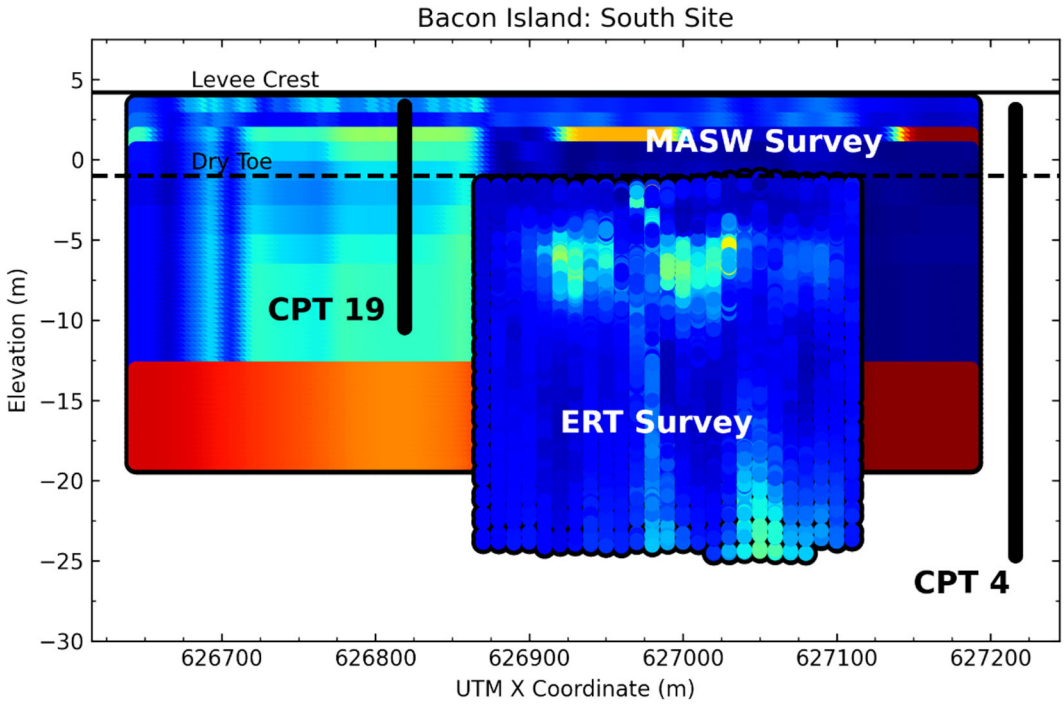


Figure 4.47 Longitudinal profile of the South Site on Bacon Island.

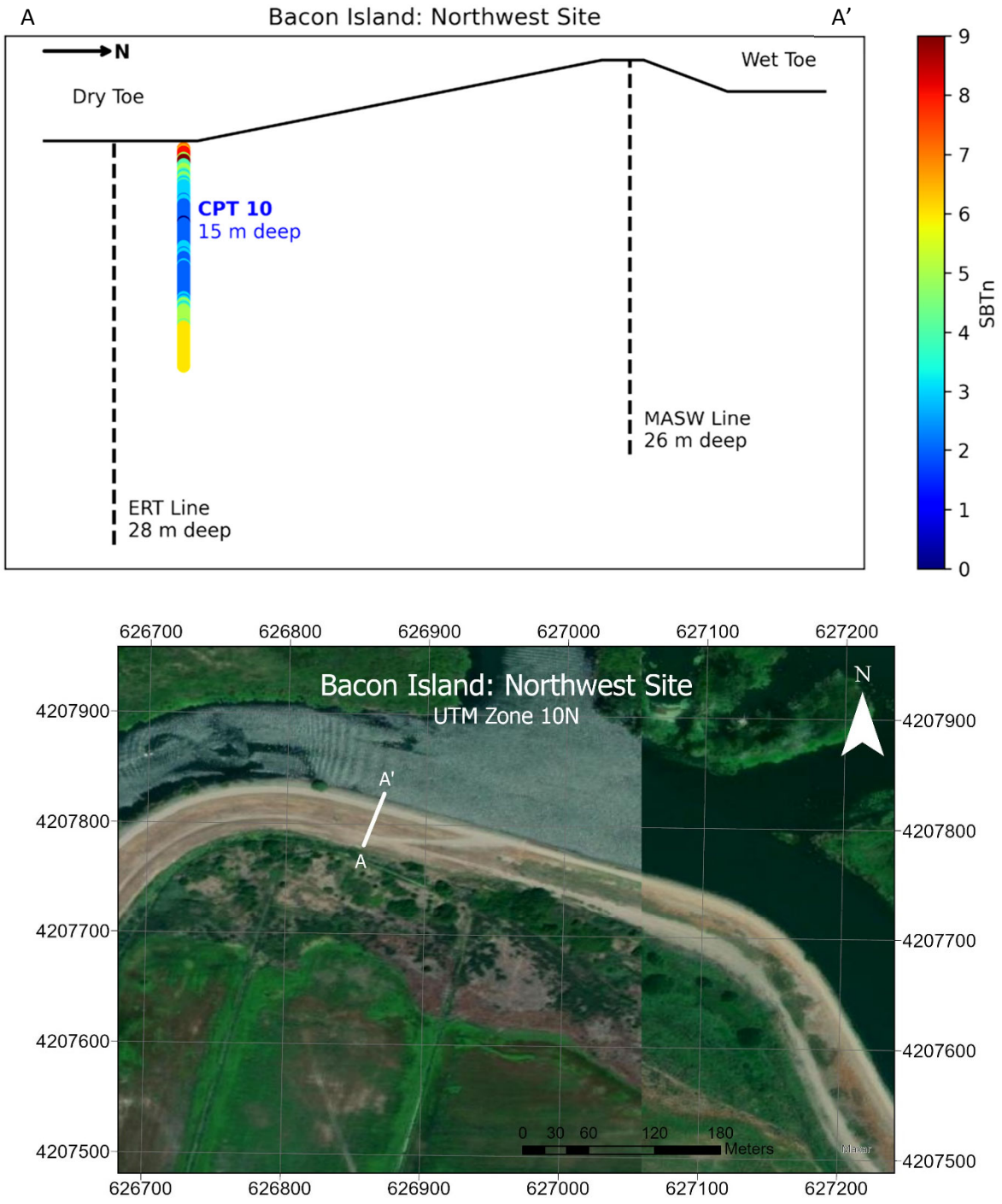


Figure 4.48 Cross-sectional profile of the Northwest Site on Bacon Island.

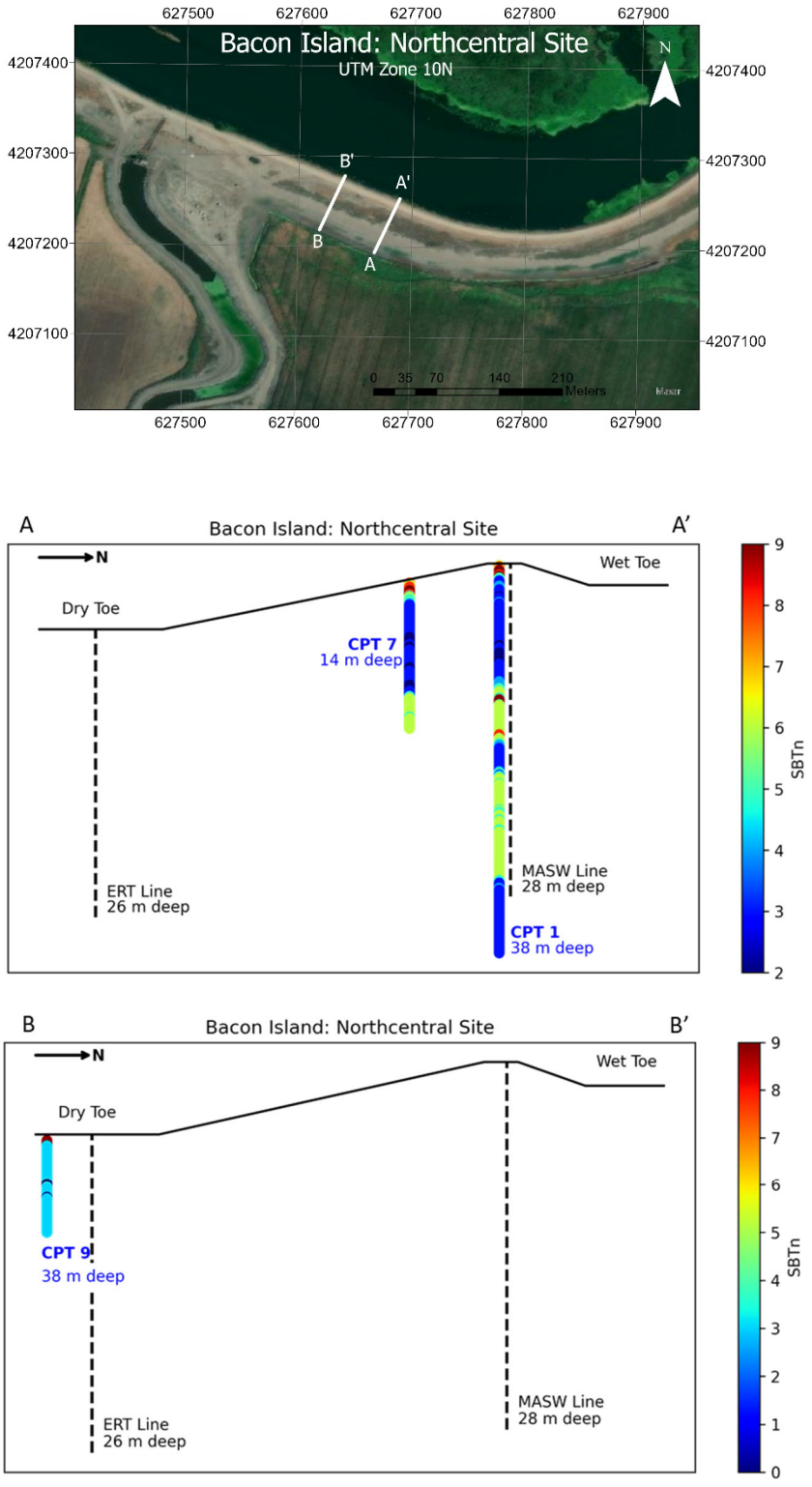


Figure 4.49 Cross-sectional profile of the Northcentral Site on Bacon Island.

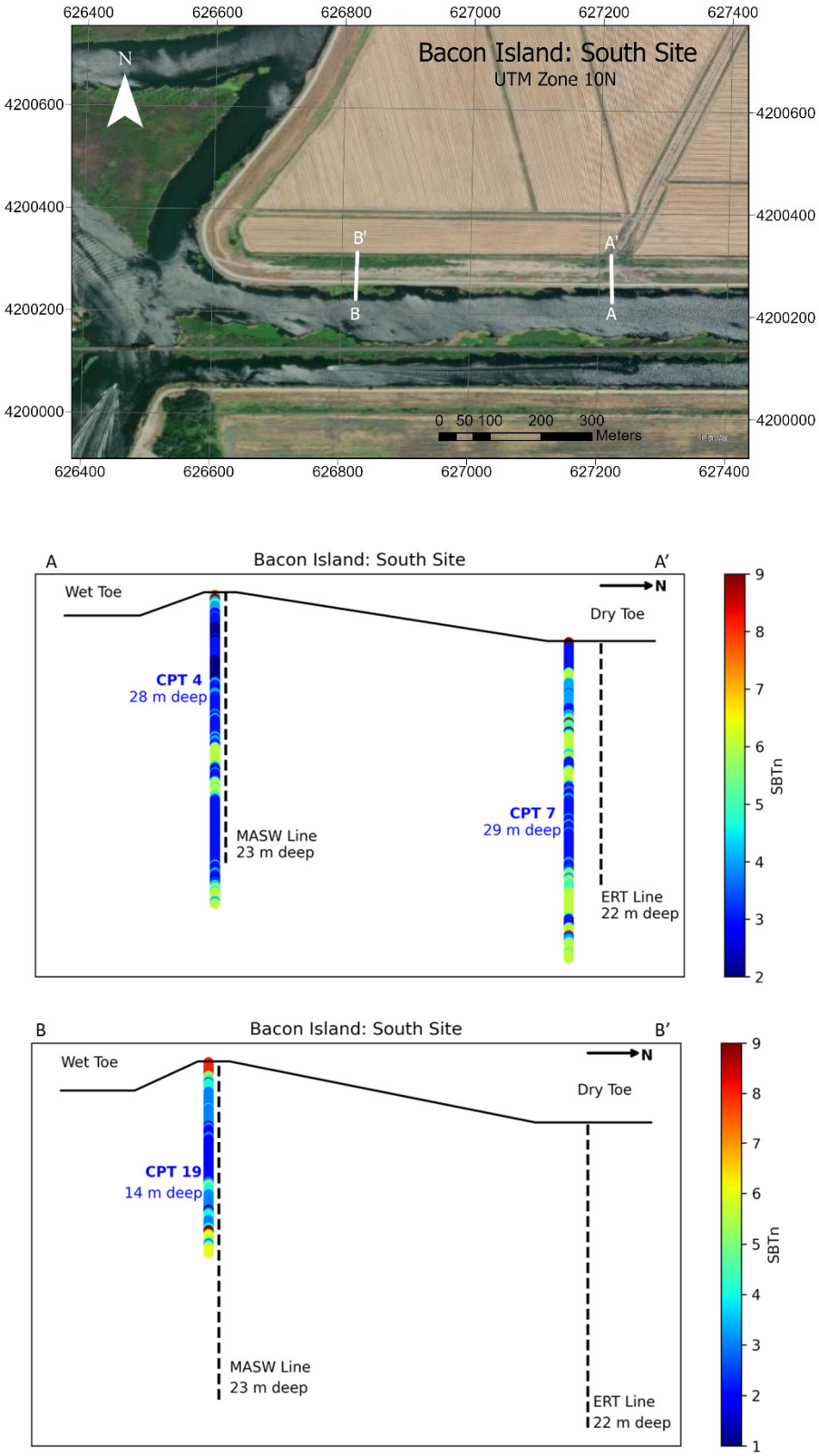
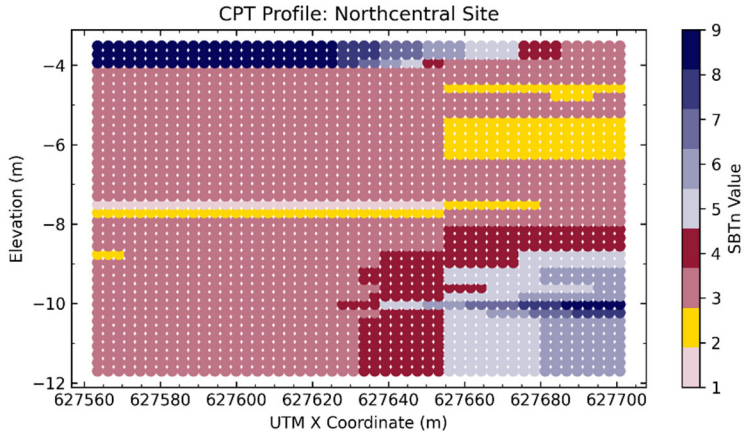
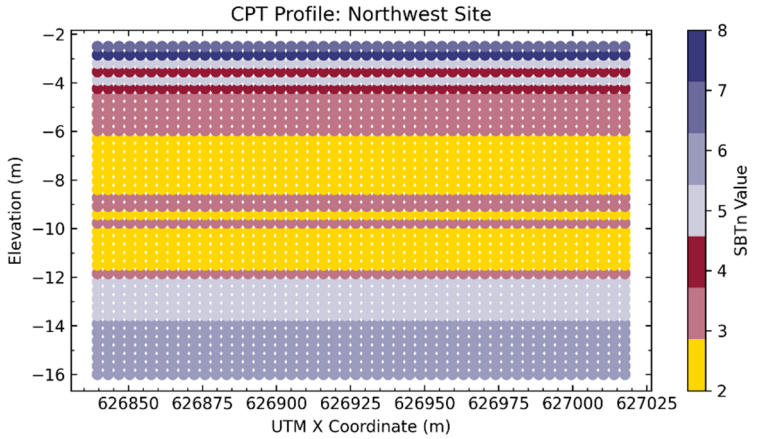


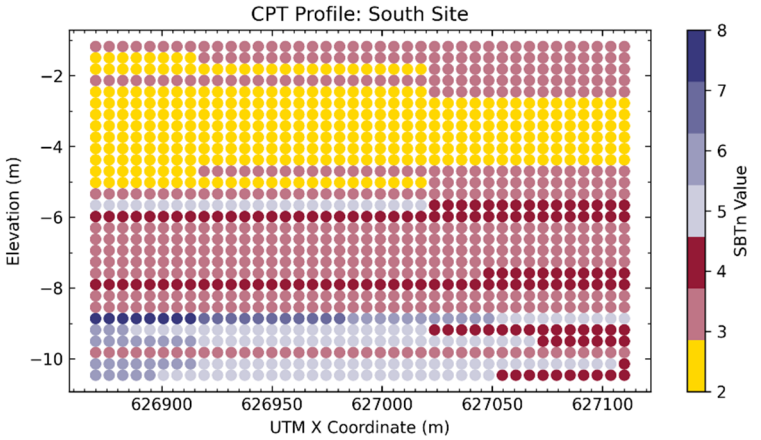
Figure 4.50 Cross-sectional profile of the South Site on Bacon Island.



(a)



(b)



(c)

Figure 4.51 CPT Profiles at Bacon Island for (a) Northcentral site, (b) Northwest site, and (c) South site.

Interpolating the data to a uniform grid allows for the data to be used in the machine learning algorithms as every point will have a *Vs*, ρ, SBTn, elevation, and x-coordinate. The *Vs* and ρ values were then standardized, separately. One dataframe including all three sites was initially created. The features included in the dataset were the site number, x-coordinate, elevation, standardized *Vs*, and standardized ρ.

4.9.2 Analysis #1 - Predicting Individual SBTn

The three supervised machine learning models classified the individual SBTn values based upon *Vs*, ρ, elevation, x-coordinate, and site number with varying success shown as Analysis #1 in Table 4.15. Cross-validation for KNN determined that three neighbors were optimal but resulted in the poorest performing algorithm. This model had good training accuracy but poor testing accuracy, causing the algorithm to overfit the training data. For SVM, cross-validation and grid search were used to determine optimal parameters of C, at 100, and γ, at 0.01. With these parameters, the training accuracy was similar to KNN but with a higher testing accuracy, resulting in the model not overfitting the training data. The best-performing model for classifying individual SBTn values was RF. With RF, grid search and cross-validation determined that the maximum depth of the trees should be 20, the number of tree estimators should be 201, with a minimum leaf sample of 2, and the minimum number of split samples should be 2. Using these optimized parameters, the model had a 99% training and 94% testing accuracy. RF had the best performance and did not overfit the training data.

4.9.2.1 Performance at the Site Level

To predict individual SBTn, all the features were used in the algorithm. This dataset was randomly shuffled and a random 30% of the data was removed to be the testing dataset while the remainder 70% of the data was the training dataset. The 70% training dataset was first inputted into the machine learning algorithms to train a model. The 30% testing dataset was then put through the trained models to predict the SBTn.

To understand the performance at the site level, classified data was analyzed separately for each site on Bacon Island. For each site and each algorithm (RF, KNN, SVM), the classified SBTn value from the algorithm was compared to the "true" SBTn value from the dataset. For each value of SBTn (1-9), the percentage of correct SBTn was calculated.

To look at how each algorithm performed for each site, the percentages representing the correct classification for each algorithm at a site were summed and normalized for the entire range of SBTn, resulting in a value between 0 and 1. When the value was closer to 0, the algorithm generally classified the SBTn values poorly, while a value closer to 1 indicated that the algorithm was able to classify the SBTn values well. The results are summarized in Table 4.16.

Table 4.16 Summary of Normalized Correct Percentages per Algorithm per Site

Algorithm	Site		
	Northwestern	Northcentral	South
KNN	0.42	0.38	0.34
RF	0.94	0.60	0.81
SVM	0.52	0.51	0.41

This analysis showed that KNN performed the poorest out of the three algorithms for each site followed by SVM for each site. RF performed the best for all three sites with significantly higher normalized correct percentages than KNN and SVM, resulting in RF predicting more SBTn values correctly for each site compared to the other algorithms. These results are expected when compared to the total accuracy of each algorithm from Analysis 1 in Table 4.15, which shows the same overall trend.

At the site level, the number of individual SBTn values being correctly classified can also be analyzed. This was done by averaging and normalizing the correct percentages of the three algorithms for each site for each SBTn value. For this analysis, SBTn 1 and 9 were not included as they were not present at each site. The results are summarized in Table 4.17.

Table 4.17 Summary of SBTn performance for each site

SBTn	Northwestern	Northcentral	South	Site that Performed the Best
2	0.95	0.88	0.84	Northwest
3	0.55	0.71	0.75	South
4	0.37	0.54	0.60	South
5	0.73	0.72	0.69	Northwest
6	0.94	0.86	0.78	Northwest
7	0.48	0.48	0.29	Northwest or Northcentral
8	0.38	0.38	0.21	Northwest or Northcentral

Like the previous analysis, the closer the value is to 1, the better the site performed overall in classifying the SBTn value. In terms of which SBTn values were correctly classified most often, SBTn 2, 5, and 6 were classified the best with SBTn 7 and 8 being classified the worst.

When looking at which site performed best overall, the normalized value of each site for each SBTn was compared. The site with the highest normalized value for a given SBTn value was deemed the best site for that SBTn value, as summarized in Table 4.17. The site that resulted in classifying the most SBTn values correctly was the Northwest Site, containing 43-71% of the SBTn values. The next best site was the South Site which classified 29% of the SBTn values as the best. The worst site was the Northcentral Site which classified 0-29% of the SBTn values as the best.

To understand what caused the Northwest Site to predict the best and the Northcentral Site to predict the worst, the Vs and ρ of each site were visualized using violin plots shown in Figure 4.52.

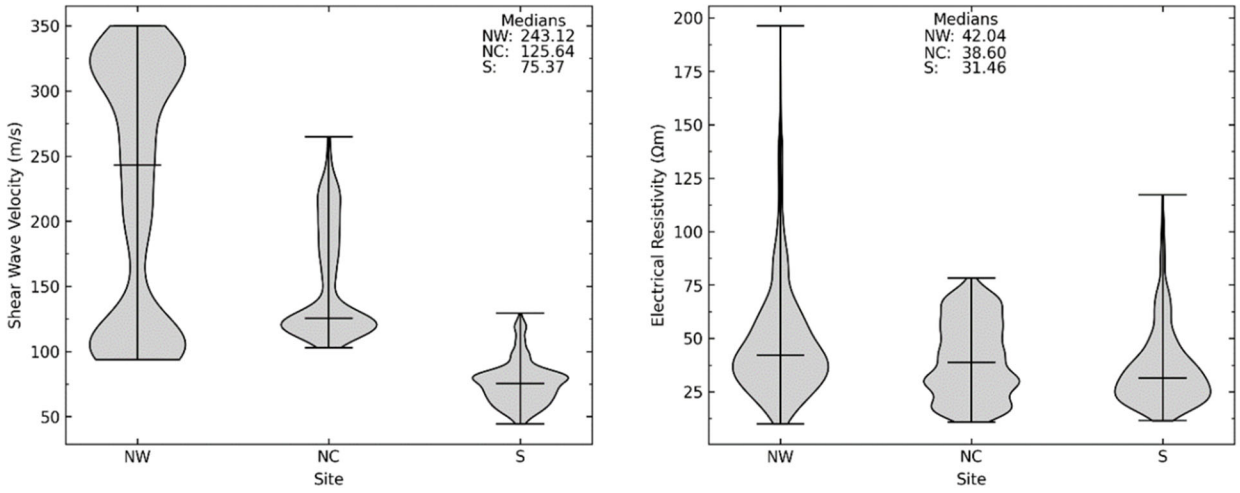


Figure 4.52 Violin plots showing the distribution at each site of (a) shear wave velocity and (b) electrical resistivity.

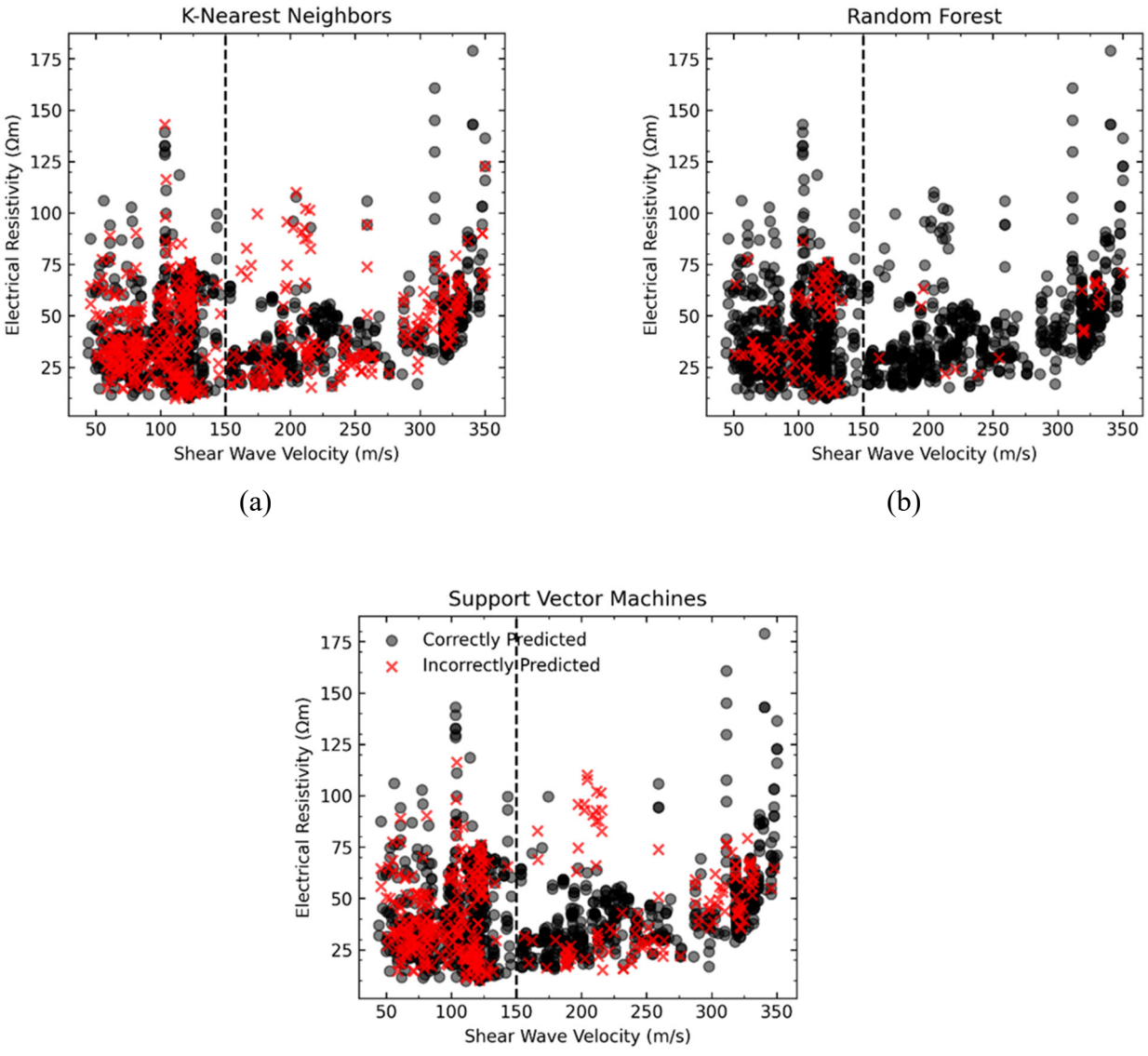
When comparing the *Vs* and ρ of each site, the South Site has the lowest *Vs* and the smallest range with bounds of 50 to 125 m/s and a median of approximately 75 m/s. The South Site also exhibited a unimodal distribution of *Vs*.

The Northcentral Site has the next lowest *Vs* median of 126 m/s with a wider range of *Vs* bounds of 100 to 275 m/s. The data for the Northcentral Site was heavily skewed around the median of the range with a long thin tail of data at higher *Vs*.

The best-performing site, the Northwest Site, showed a bimodal distribution in *Vs* where the data congregates near the upper and lower bounds of the range. This site has a median of 243 m/s with a lower bound of 100 m/s and an upper bound of 350 m/s. The subsurface profile for this site, which can be found in Appendix C, shows a lower *Vs* zone near the surface but also between x-coordinates of 626815 m to 626905 m with stiffer soils below -5 m below ground surface on either side of the lower *Vs* zone. This could be due to remediation of the levee where softer soils in the subsurface were replaced with stiffer soils.

When looking at the ρ distributions, each site has a similar unimodal distribution. The South Site had a median of 31 Ωm, the Northcentral Site had a slightly higher median of 39 Ωm, and the Northwest Site had a median of 42 Ωm. The Northwest and South sites have similar shapes in distribution with most of the data falling around the median and then tapering off towards their respective upper bounds. The Northcentral Site was slightly different where the distribution of the data was relatively consistent among the ρ range for that site.

To further explore how *Vs* and ρ were affecting the classification results, the correct and incorrect predictions with respect to *Vs* and ρ were plotted and visualized in Figure 4.56. Each plot in Figure 4.53 represents a different supervised machine learning algorithm. The data was plotted with respect to *Vs* and ρ with correctly classified points represented as black circles and incorrectly classified points represented by red x's. At 150 m/s, there is a clear demarcation in RF and SVM which separates data that was correctly classified from data that was incorrectly classified. This demarcation is still visible for KNN but not as prevalent as with the other algorithms.



(a)

(b)

(c)

Figure 4.53 Correctly Predicted data (black circles) and incorrectly predicted data (red x's) as a function of shear wave velocity and electrical resistivity for (a) K-Nearest Neighbors, (b) Random Forest, and (c) Support Vector Machines.

With the KNN model, 64% of the incorrectly classified data fell below a *Vs* of 150 m/s. In comparison, SVM had 67% of the incorrectly classified data fall below this 150 m/s demarcation while RF had 82% of the incorrectly classified data fall below the demarcation. As the overall accuracy of the model increased, the fewer incorrectly classified data points fell above the 150 m/s demarcation. With most of the incorrectly predicted data falling below the 150 m/s demarcation line, this suggests that subsurface profiles with stiffer materials can be modeled more accurately. With only 35% of the training dataset representing stiffer soils, class imbalance is not a cause for the enhanced classification performance of stiffer materials.

4.9.2.2 Performance of Each SBTn Value

Using the results from Analysis #1, the classification results were analyzed to look at which SBTn value was predicted the best and which was predicted the worst, regardless of the site. For each SBTn, the total percentage of correctly classified data was calculated. Then, to look at the total performance of each SBTn value, the percentages of each algorithm for each SBTn were summed and normalized to result in values between 0 and 1. Where a value of 0 resulted in poor performance for that SBTn and 1 represented good performance for that SBTn. The results are shown in Figure 4.54.

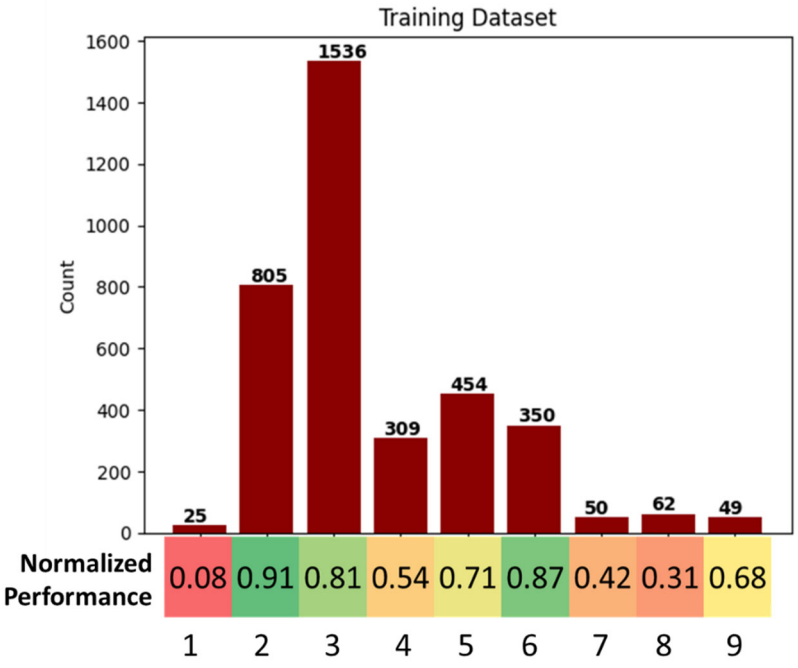


Figure 4.54 Normalized Classification Performance of each SBTn corresponding to abundance of SBTn in the Training Dataset.

Out of all the SBTn values, SBTn 2, 3, and 6 were classified the best with SBTn 1, 7, and 8 being classified the worst. To understand why certain SBTn were classified better than others, a histogram of the occurrence of each SBTn in the training dataset was created and compared to the normalized classification, as shown in Figure 4.57. Where the normalized performance was the highest, there were typically more SBTn in the training dataset while where there were fewer SBTn in the training dataset, the normalized performances were poor. For example, for SBTn 1, there were only 25 counts in the training dataset of 3640 total data points, resulting in the lowest normalized performance of 0.08. In contrast, SBTn 2 had 805 counts in the training dataset with a normalized performance of 0.91 and SBTn 3 had the highest count at 1536 with the third highest normalized performance of 0.81. This trend of SBTn that appear more frequently in the training dataset having higher normalized performances compared to their counterparts with lower frequency and poor normalized performance suggests that class imbalance is affecting the accuracy of classification.

Multiple methods were used to reduce the effect of class imbalance on the ability of the model to accurately classify the SBTn. These methods include stratified randomization (Analysis #2), up-

sampling and down-sampling (Analysis #3), and using different scoring metrics to determine how well the model performed.

Analysis #2 – Stratified Randomization

When using stratified randomization, the dataset is split into subgroups that represent class imbalance. In this analysis, the subgroups represented each of the SBTn values. Within each subgroup, a random sample of data points was selected based on the size of the subgroup. This was done using Python's 'train_test_split' function from the SKLearn Learn library (Pedregosa, et al., 2011).

The data was first randomly shuffled and then 30% of the data was used as the testing dataset and the remaining 70% was the training dataset. Stratified randomization was applied to the SBTn data to help reduce the effect of class imbalance.

To understand how stratified randomization affected the model prediction, the normalized correct performance described in Section 4.9.2.1 was applied to the classified results. Figure 4.55 summarizes how the previous baseline analysis (Analysis #1) performed against the stratified randomization analysis and the sampling analysis.

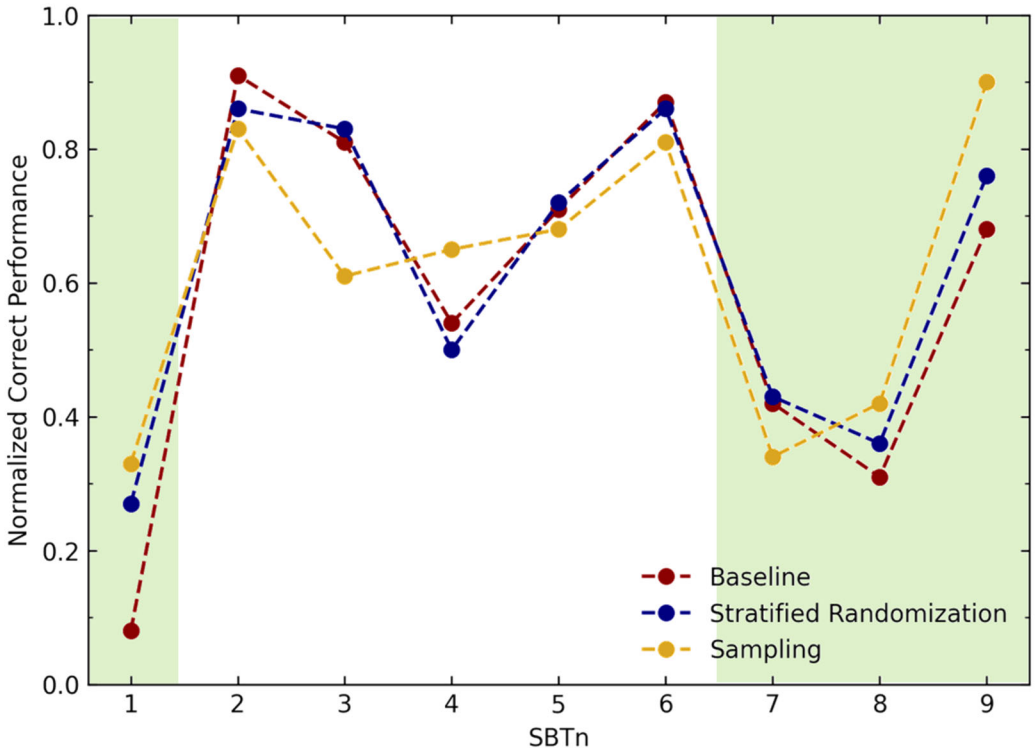


Figure 4.55 Normalized Classification Performance of each SBTn of the baseline analysis (Analysis #1), stratified randomization (Analysis #2), and up/down sampling (Analysis #3) to look at the effect on class imbalance. Areas shaded in green represent low class representation.

If the stratified randomization reduced the effect of class imbalance on classification, the normalized correct performance value would be closer to 1 compared to the baseline value. When compared to the baseline value, stratified randomization had slightly higher values for the classes that are not represented well, such as SBTn 1, 7, 8, and 9. For SBTn values that are more represented in the training dataset, the normalized correct performance value had very little difference from the baseline normalized correct performance. These results show that stratified randomization helped with class imbalance slightly but not significantly.

Analysis #3 – Sampling

Another method that was used to help reduce the effect of class imbalance was sampling. With sampling all classes are forced to have the same number of occurrences in the training dataset. If a class had more occurrences than the specified number of samples, down-sampling was used. A random number of data points for that class was removed from the training set to match the specified number of samples. Alternatively, if the class had fewer occurrences than the specified number of samples, up-sampling was used. This entails randomly repeating existing data in the class until the specified number of samples was met.

For this analysis, each SBTn value was forced to have 500 occurrences in the training dataset. SBTn 1, 4, 5, 6, 7,8, and 9 used up-sampling to reach 500 occurrences while SBTn 2 and 3 used down-sampling to reach 500 occurrences. With the new dataset, a random 30% was used for the testing dataset and the remaining 70% was used for the training dataset.

To understand how sampling affected the normalized correct performance of each SBTn value, it was compared to the baseline normalized correct performance, as shown in Figure 4.58. For the SBTn values that were not represented well in the original dataset, up-sampling resulted in higher normalized correct performances except for SBTn 7 which was less than the baseline. Where the SBTn values were well represented in the training dataset, sampling performed worse than the baseline except for SBTn 4, where it performed better. Like with stratified randomization, sampling helped slightly but not significantly when dealing with the effect of class imbalance.

A major weakness of up-sampling is memorizing data when repeated significantly enough, which results in poor classification results. When significantly up-sampling data, such as with SBTn 1 where there were 25 occurrences in the baseline dataset and copied 20 times to reach 500 occurrences, the algorithm starts memorizing the few data points that are repeated multiple times. A similar issue occurs with down-sampling. If the class needs to remove a large number of occurrences, such as with SBTn 3, this could result in significant information loss and the algorithm may not be able to generalize and predict new instances correctly.

4.9.2.3 Overfitting

The scoring metric used to determine the performance of the model can also be influenced by class imbalance and is used to determine if the model is overfitting. The results reported in Table 4.15 were all measures of accuracy. Other scoring metrics such as the F1 score were analyzed to determine which scoring metric produced the least overfitting. Overfitting of the data occurs when the training dataset has a significantly higher success rate compared to the testing success rate.

This occurs when the model learns the nuances of the training data and cannot generalize to new data, such as the testing data. When class imbalance is present in the dataset, this can contribute to overfitting.

To look at which scoring metric was best for the dataset, an analysis was performed using the KNN model with the baseline analysis (Analysis #1) to see how different scoring metrics changed with increasing neighbors. The scoring metrics used were accuracy, F1 score with macro averaging, F1 score with micro averaging, and F1 score with weighted averaging. The results are shown in Figure 4.56.

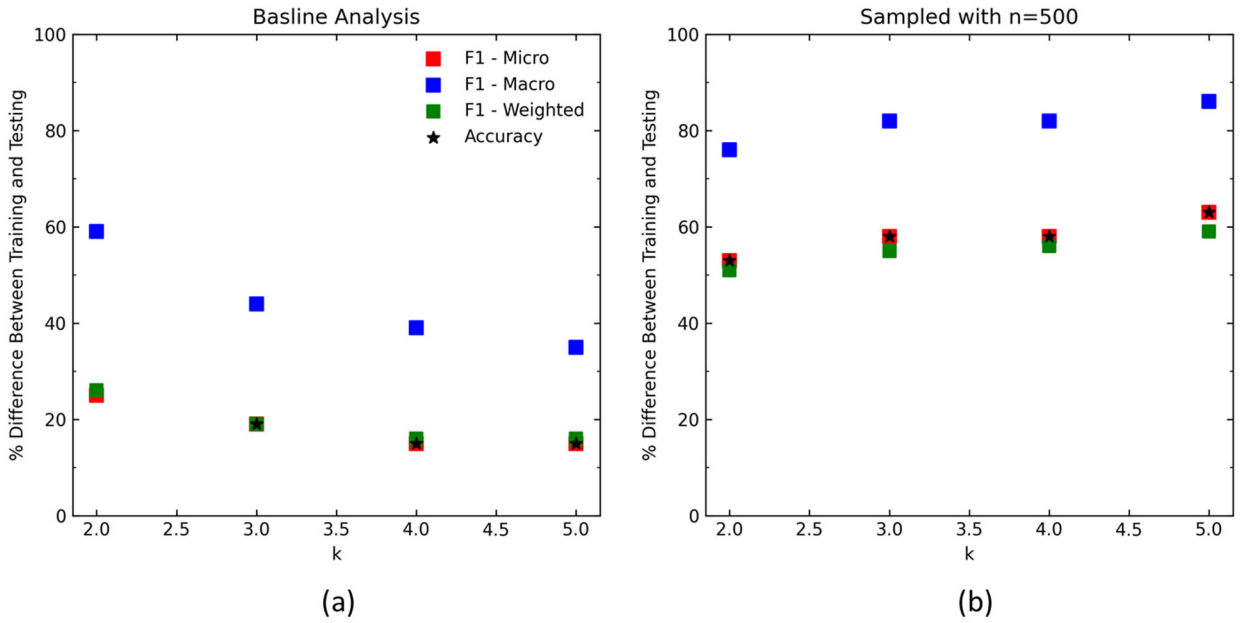


Figure 4.56 Showing the effect of scoring metric on overfitting with respect to the number of neighbors for (a) the baseline analysis and (b) the sampled analysis using 500 samples per class.

Figure 4.56(a) shows how scoring metrics affected overfitting concerning *K* for the baseline analysis. With a lower percent difference between the training and testing datasets, there is less overfitting in the model. From this analysis, the F1 score using macro averaging had the highest amount of overfitting among all the scoring metrics ranging from 35% to 59% difference between the training and testing datasets.

Accuracy and F1 score using the micro averaging had the same percent difference between the training and testing datasets for each value of *K*. These two metrics also showed the best results concerning overfitting with a range of 15% to 25% difference. The F1 score using the weighted average had slightly more overfitting compared to Accuracy and the F1 score with micro with a range of 16% to 26%.

This same analysis was done using the sampled dataset with 500 samples per class, as shown in Figure 4.56(b). Like with the baseline analysis, the F1 score with the macro averaging performed the worst with the highest range of percent difference at 76% to 86%. The best-performing scoring metric against overfitting was the F1 score with a weighted average with a range of 51% to 59%.

Like with the baseline analysis, accuracy and F1 score with the micro average performed the same for each *K*. Unlike the baseline average though, accuracy and F1 score with the micro average performed slightly worse than the F1 score with the weighted average with a range of 53% to 63%.

This analysis also showed that as the number of neighbors increased, the overfitting decreased for all scoring metrics. This is expected as *K* acts as a smoothing parameter for KNN. With an increase in *k*, there is more smoothing in the decision boundary for classification which should result in less overfitting. This trend was not observed with the sampled dataset when using 500 samples per class. Instead, as *K* increased, the overfitting in all scoring metrics increased slightly. This could be due to data being repeated multiple times to reach 500 samples causing the algorithm to memorize the data instead of learning trends.

The effect of overfitting on the number of samples for each class used in the sampling database was also analyzed. The results of this analysis are shown in Figure 4.57.

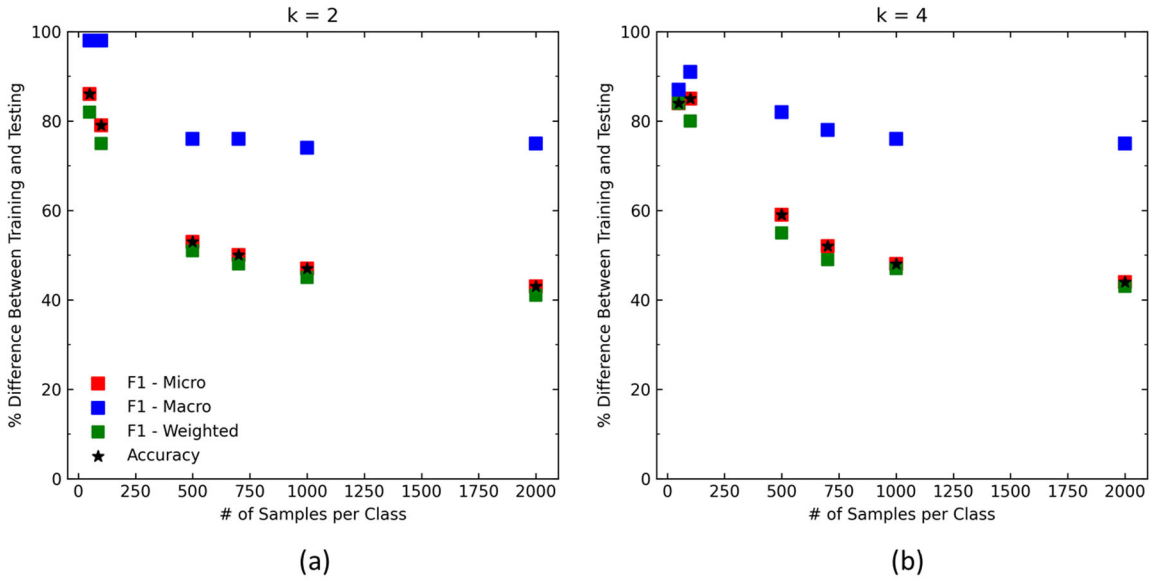


Figure 4.57 Showing the effect of scoring metric on overfitting with respect to the number of samples per class for (a) when k = 2 and (b) the k = 4.

Figure 4.57(a) shows the results with the optimal number of neighbors determined through cross-validation for KNN. As the number of samples per class increases, the degree of overfitting decreases for all scoring metrics. With a few samples per class, such as 50 and 100 samples, there was significant overfitting, but as the number of samples started to become large, such as 1000 and 2000 samples, the degree of overfitting started to flatten out. Like with the previous analysis looking at how *K* affects overfitting, the F1 score with the macro averaging results in the worst overfitting with a range of 75% to 98%. F1 score with the weighted average performs the best with a percent difference range between 41% to 82%. Accuracy and F1 score with the micro average performed equally well and slightly worse than F1 score with the weighted average with a range of 43% to 86%.

This analysis was repeated using $K = 4$, which was not optimal for the KNN analysis, to see if the trends described previously still hold. The same trends did exist where a low number of samples

resulted in large degrees of overfitting and decreased overfitting as the number of samples increased until it flattened out. The same trends with the scoring metrics also held. F1 score with macro averaging performed the worst with a percent difference between 75% and 87%. F1 score with the weighted average performed the best with a range of 43% and 84% while accuracy and F1 score with micro average performed the same and slightly worse than F1 score with weighted average with a range of 44% to 84%.

When using sampling to reduce the effect of overfitting, there is a trade-off between the degree of overfitting and computational effort. As the number of samples per class increases, the dataset becomes larger, resulting in less overfitting but the analysis becomes more computationally demanding. The trade-off between computational effort and the degree of overfitting must be addressed when using sampling to reduce the effect of class imbalance.

This analysis also showed that overall, the F1 score with the macro average performed the worst regardless of K and the sampling techniques by significant amounts. The macro averaging finds the unweighted mean of each class and does not take into class imbalance, causing it to perform the worst with an imbalanced dataset.

F1 score with the micro and weighted average, as well as accuracy, had similar performance in terms of overfitting. As a result, since the baseline analysis, stratified randomization, and sampling analysis were performed using the accuracy metric, the subsequent analyses were also scored using accuracy to be consistent.

4.9.3 Predicting SBTn of a Site

While predicting individual SBTn is important, predicting the SBTn of an entire section of a levee is more important. In practice, once the model is created, it is more beneficial for a stakeholder to predict the subsurface lithology of a novel site instead of predicting the subsurface lithology of one point in the subsurface. As a result, it is important to understand how well these supervised machine learning models perform when predicting the subsurface lithology of an entire site.

This analysis was done by creating three dataframes using all the features: x-coordinate, elevation, Vs, and ρ. In these analyses, a combination of two sites was used as the training data for the model instead of a random 70% of the data, and the third site was used as the testing data instead of a random 30% of the data. The SBTn at the South Site on Bacon Island was predicted based on the Northwest and Northcentral sites (Analysis #7), while the Northwest Site was predicted based on a model trained by the South and Northcentral sites (Analysis #5), and the Northcentral Site was predicted from the South and Northwest sites (Analysis #6).

There was little success in predicting a site based on the other two sites as summarized in Table 4.15. When trying to predict the SBTn at the Northwest Site, the training accuracy ranged from 82% to 99% among the three algorithms but the testing accuracy ranged from 21% to 24%. When predicting the Northcentral Site, the training accuracy ranged from 72% to 100% with testing accuracies of 42% to 44%. Finally, when predicting the South Site, the training accuracies ranged from 75% to 99% with testing accuracies of 9% to 30%.

The percent difference between the training and testing accuracy for each algorithm and each site is represented in Table 4.18 with the average of the percent difference among the three algorithms for each site.

Table 4.18 Percent Difference between Training and Testing Datasets for Predicting Each Site

	Percent Difference (%)			
Site	**KNN**	**RF**	**SVM**	**Average**
Northwest	122	122	118	121
Northcentral	48	80	60	63
South	123	107	162	131

The South Site had the highest average percent difference between the training and testing datasets with 131% followed by the Northwest Site with an average percent difference of 121%. The Northcentral Site had the smallest average percent difference at 63%, indicating that using the South and Northwest sites as training data created a model that was better at generalizing to a new site. In the case of all three sites, there was overfitting on the training data.

The Northcentral Site has a percent difference of almost half the next lowest site. To understand why the Northcentral Site classified SBTn better compared to the other two sites, violin plots showing the distribution of *Vs* and ρ in the training dataset were created and are shown in Figure 4.58.

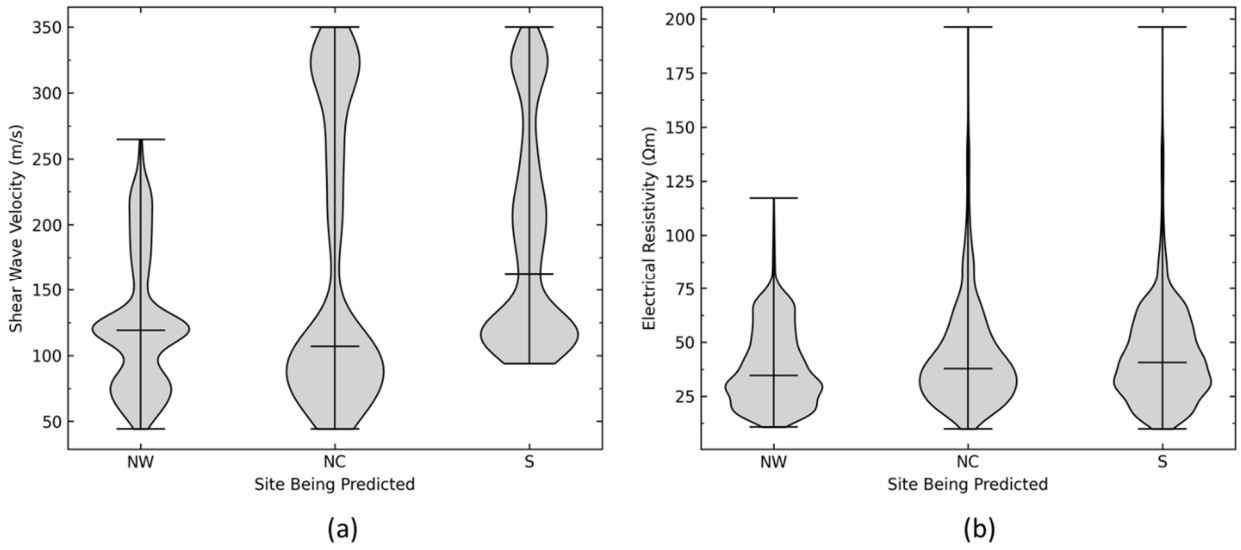


Figure 4.58 Distribution of (a) shear wave velocity and (b) electrical resistivity in each dataframe for predicting SBTn at a site.

The training dataset for classifying the Northcentral (NC) Site had a larger range of *Vs* compared to the other two training datasets ranging from 44 m/s to 350 m/s with a median of 107 m/s. The training dataset for the Northwest (NW) Site had *Vs* that ranged from 44 m/s to 265 m/s with a median of 119 m/s while the South (S) Site had *Vs* that ranged from 94 m/s to 350 m/s with a median of 163 m/s.

For the ρ ranges, the Northwest Site had a mean of 35 Ωm and spanned 11 Ωm to 117 Ωm. The Northcentral Site had ρ that ranged from 10 Ωm to 196 Ωm with a median of 38 Ωm while the South Site was like the Northcentral Site with the same range but a slightly higher median of 41 Ωm.

One potential possibility for the Northcentral Site having better classification than the other sites is due to the large range in *Vs* and ρ. The training dataset for the Northwest Site does not have data above 265 m/s 117 Ωm while the South Site does not have training data below 94 m/s. Since the Northwest training data spans the entirety of the available *Vs* and ρ data, this could help the model to generalize better to a new site.

Another possibility that could affect how well a model performs is the amount of training data used to train the model. Typically, models with more data perform better than models trained on less data. The Northwest classification was trained on 3,200 data points, the Northcentral was trained on 3,200 data points, and the South Site was trained on 4,000 data points. Since the Northcentral Site is the same as the Northwest Site, the amount of data did not affect the results.

4.9.3.1 Spatial Accuracy of SBTn

Since an entire site was classified instead of just a random 30%, the entirety of the site profile can be analyzed to understand spatially where data was being classified correctly and where data was being classified incorrectly by comparing the true SBTn with the predicted SBTn. Each profile can be analyzed separately for each algorithm (KNN, RF, SVM), for example as shown in Figure 4.59 for the Northcentral Site.

Figures 4.59(a), (b), and (c) show a correct classification as a grey circle and an incorrect classification as a red circle for KNN, RF, and SVM, respectively. Elevation is on the y-axis and x-coordinate is on the x-axis. Using this method, the prediction results can be analyzed spatially. The three profiles can be combined as shown in Figure 4.59(d) Since the location of points in the profile is consistent among the different algorithms, the number of algorithms that classify a point incorrectly can be summed. Each point has the possibility of having 0, indicating that SVM, RF, and KNN, all classified the point correctly, or up to 3 which indicates that SVM, RF, and KNN all incorrectly classified the point.

This was done for the Northwest and South sites as well, as shown in Figure 4.60. For the profiles showing the individual performance of SVM, RF, and KNN for the Northwest and South sites, please refer to Appendix J. The Northwest and South sites were mostly classified incorrectly as expected based on the training accuracy among the three algorithms. Because most of the profile was classified incorrectly, there were no trends of where data was being misclassified, spatially.

In comparison, the Northcentral Site, which had the best training accuracy, showed the near-surface was one location that had the poorest classification. Within the top 1 meter, most of the data points were misclassified by two to three algorithms. Another location that was significantly misclassified was bounded by 627630 m to 627700 m and below -8 in elevation until the terminal depth of the survey. The Northcentral Site also had a section from an elevation of about -5 m to just below -6 m that spanned almost the entirety of the site and had perfect classification by all three algorithms. From -6 m elevation to the terminal depth spanning 627560 m to 627630 m also

had good classification results. For this zone, SVM and RF correctly classified the data points while KNN did not.

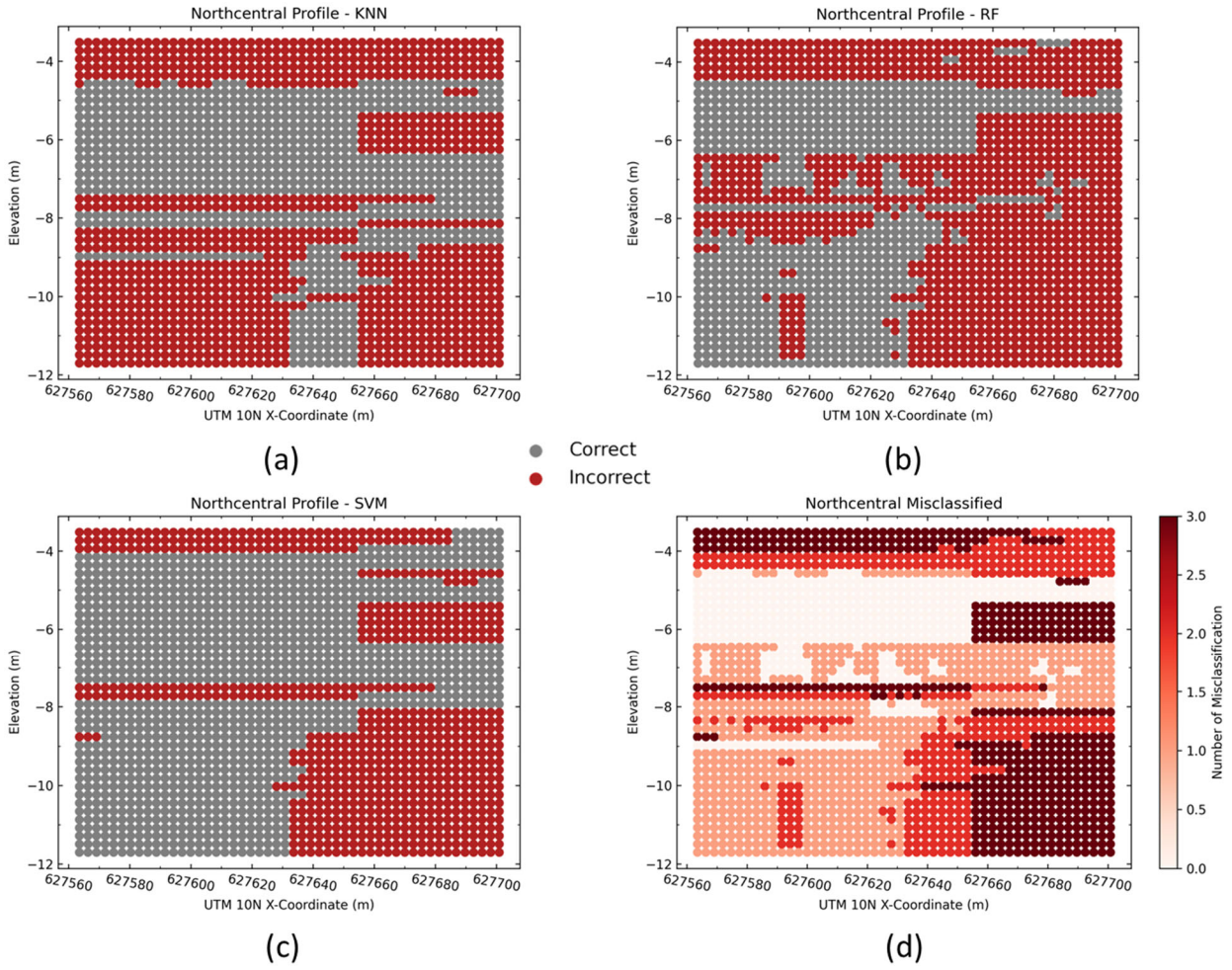


Figure 4.59 Spatial distribution of correct (grey) and incorrect (red) SBTn classification using (a) K-Nearest Neighbors, (b) Random Forest, (c) Support Vector Machines, and (d) a combination of the three algorithms.

The distribution of *Vs* and ρ for the correctly and incorrectly classified data for the Northcentral Site was pulled out to understand if *Vs* and ρ were influencing classification. The results are shown in Figure 4.61.

Data points that were correctly classified showed a *Vs* distribution that ranged from 103 m/s to 265 m/s with a median of 172 m/s. The distribution was bimodal showing that most of the data was around a *Vs* of 120 m/s and 210 m/s. The ρ distribution was unimodal and ranged from 13 Ωm to 70 Ωm with a median of 36 Ωm. In comparison, the incorrect data showed a *Vs* distribution that was unimodal mostly around 122 m/s while the ρ distribution was bimodal with data centered around 15 Ωm and 65 Ωm. The locations that were classified incorrectly typically occurred in areas with softer soils.

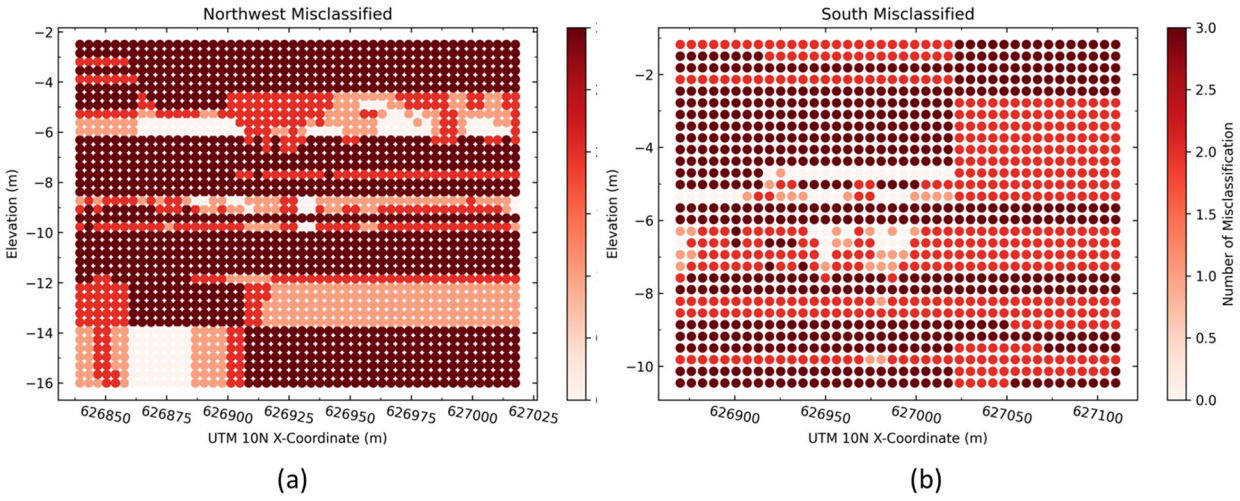


Figure 4.60 Spatial distribution of correct and incorrect classification looking at all three algorithms for (a) the Northwest site and (b) the South site.

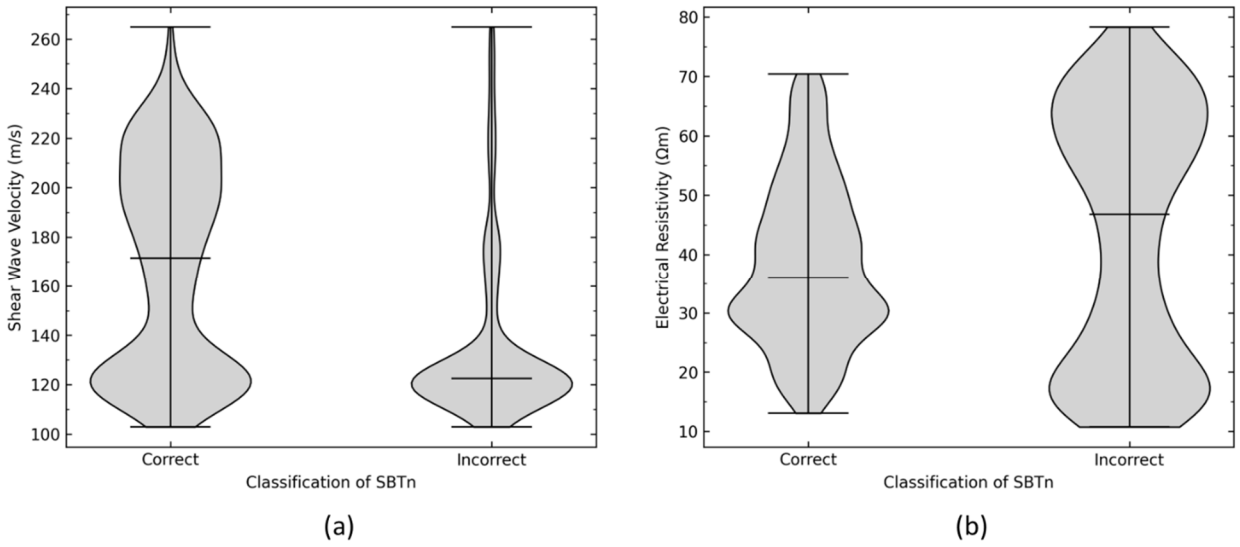


Figure 4.61 The distribution of (a) shear wave velocity and (b) electrical resistivity at the Northcentral Site for data that was correctly and incorrectly classified by the three algorithms (SVM, RF, KNN).

4.9.4 Feature Selection for Algorithm Performance

Feature selection is an important feature engineering consideration when performing a machine learning analysis. Including or removing certain variables can cause model improvement or model interpretability and should be considered when designing an analysis. One way to determine the importance of a feature is through feature importance which generates a score for each input model that represents how impactful that feature was to the model. When a feature has a higher score, it

will have a greater impact on the model. The scores of the input variables for the previous analyses were calculated as a Gini Importance using Random Forest as shown in Figure 4.62.

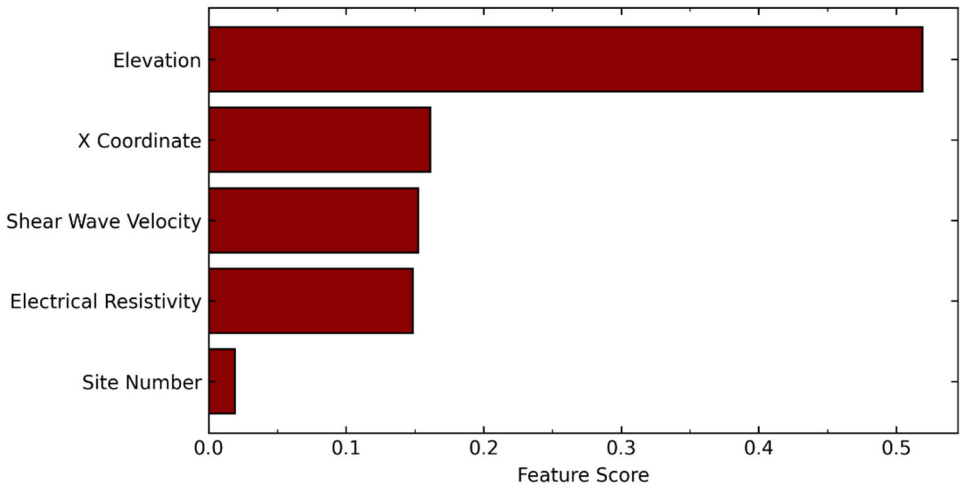


Figure 4.62 Feature Importance from Random Forest.

Elevation had the highest impact on the model with a Gini Importance of 0.5 with the site number having the lowest impact on the model with a Gini Importance of 0.02. The x-coordinate, *Vs*, and ρ had a similar impact on the model with a Gini Importance of approximately 0.15. Elevation and x-coordinate were included in the analysis originally to represent the spatial distribution of the data. The inclusion of these variables could affect how well the model can generalize to a new location, as a new location might not have overlapping elevations or x-coordinates.

The first analysis conducted was looking at how removing elevation, x-coordinate, and site number from the dataset affected the individual classification of SBTn values. Like with the baseline analysis (Analysis #1 in Table 4.15), the dataset was randomly shuffled with 70% of the data used as the training dataset and the remaining 30% as the testing dataset. The results are shown as Analysis #4 in Table 4.15. When classifying individual SBTn values, removing the elevation and x-coordinate caused worse accuracy than when including them in the model. Also, removing elevation and x-coordinate resulted in worse overfitting with an 18%, 29%, and 24% difference between the training and testing datasets for KNN, RF, and SVM, respectively. In comparison, the baseline analysis had a 19%, 5%, and 10% difference between the training and testing datasets for KNN, RF, and SVM, respectively.

The analysis looking at the model's ability to generalize to a new location from a model trained on the other two models was repeated without the x-coordinate and elevation, as Analysis #8 for predicting the Northwest Site, Analysis #9 for predicting the Northcentral Site, and Analysis #10 for predicting the South Site in Table 4.15. The percent difference between the training and testing datasets was calculated for each supervised algorithm for each site being classified. The results are shown in Figure 4.63.

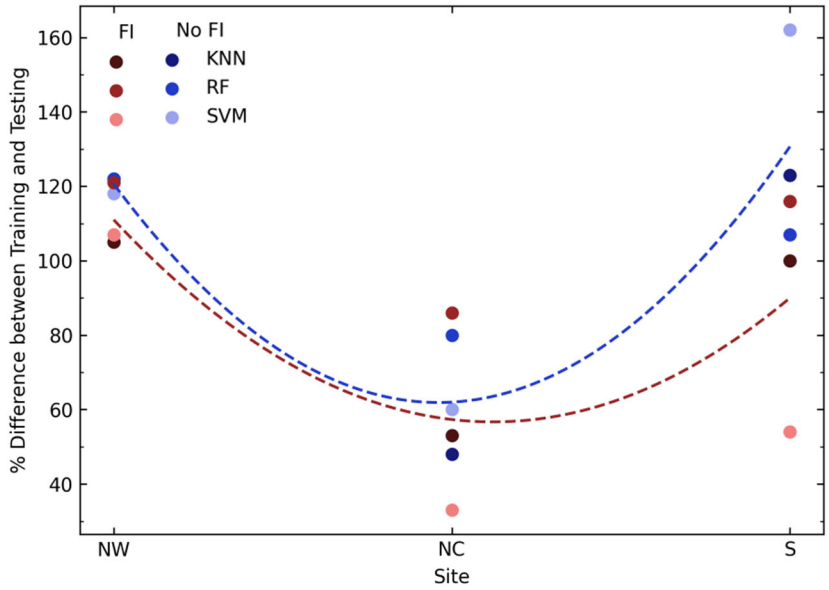


Figure 4.63 Overfitting Comparison between the Baseline Analyses (Analyses #5-7) and the Feature Importance Analyses for Predicting Site SBTn (Analyses #8-10).

The red data points represent the overfitting in the feature importance analyses while the blue data points represent the overfitting from the baseline analyses where all variables were included. From the results, for all sites, the feature selection produced less overfitting in the model, as shown by the red and blue trendlines through the data. The Northcentral Site and the Northwest Site had relatively little difference in overfitting between the two analyses, but the South Site showed significant improvement. Removing the x-coordinate and elevation when classifying new sites increased the model's ability to generalize to new data since the model was not heavily dependent on elevation and x-coordinates that were not present in the dataset.

4.9.5 Supervised Machine Learning Results Discussion

The supervised machine learning analyses performed on the Bacon Island dataset provided useful insight into spatial data analyses. If the goal of the analysis was to classify individual data without spatial consideration, including the x-coordinate and elevation in the analysis led to better results. If trying to generalize to a different location, the x-coordinate and elevation should not be included in the analysis as the model was heavily impacted by these features and will not be able to generalize to a new location without an overlapping x-coordinate and elevation.

With the current Bacon Island dataset, all algorithms struggled to not overfit the training data when generalizing to a new site, regardless of what methods were applied to reduce overfitting. This was probably due to the lack of sites in the training dataset when training the model. Currently, only two sites are used to train a model to generalize to new sites. Adding more sites with widely varying Vs and ρ would help with accuracy and reduce overfitting for new sites.

A limitation of this analysis was the lack of CPT data at Bacon Island. Having more CPT data that are close in spatial proximity will reduce uncertainty in SBTn in the subsurface. This analysis used

1 to 2 CPTs per site which resulted in high uncertainty in the subsurface and could have been a reason for overfitting and poor accuracy results. More work needs to be done to create a more comprehensive dataset for generalizing to new sites and reducing uncertainty in SBTn values.

Even with these challenges and limitations, this research has shown that using supervised machine learning for levee health analysis is feasible and could be incorporated into a resiliency framework for levee health assessment.

4.10 Supervised Compared with Unsupervised

In this section, the supervised machine learning results will be compared to the unsupervised machine learning framework for the dataset collected on the Bacon Island levee system. The identified potentially problematic zones from the unsupervised analysis will be compared to the distribution of SBTn from the CPT data along with how it correlates to the incorrectly predicted SBTn in the subsurface.

4.10.1 Unsupervised Machine Learning Results for Bacon Island

The unsupervised machine learning analysis utilized the fused anomaly and clustering framework to detect potentially problematic zones in the subsurface with integrated laboratory labels. Both K-Means and Hierarchical clustering were applied to the three sites along the Bacon Island levee system. The Elbow Rule, Silhouette Rule, and Gap Statistics were applied to each section of the levee to determine the optimal number k to use for the clustering algorithms. The same method described in Section 4.4.4 was applied to this dataset to determine the contamination factor for IF, LOF, and OCSVM for anomaly detection. The optimal hyperparameters for the three sites for clustering and anomaly detection are summarized in Table 4.19.

Table 4.19 Optimal Hyperparameters for Bacon Island

Survey	k	Contamination Factor
Northwest	6	0.134
Northcentral	6	0.155
South	3	0.156

The clustering and anomaly detection profiles for the three sites can be found in Appendix G and Appendix H, respectively. The anomaly detection profiles show that LOF did not perform well compared to IF and OCSVM, like the Grand Island dataset. The data that was classified as anomalies using LOF were random compared to IF and OCSVM, which showed similar results. Due to the poor performance of LOF, it was not included in the fused clustering and anomaly framework for Bacon Island.

There was a slight variation in the clusters found compared to K-Means and Hierarchical clustering. Even with the differences, the same relative zones of Vs and ρ were found between the two algorithms. Both K-Means and Hierarchical clustering were used in the fusion framework separately, to create potentially problematic zone profiles. These profiles for the three Bacon Island sites are shown in Figures 4.64, 4.65, and 4.66.

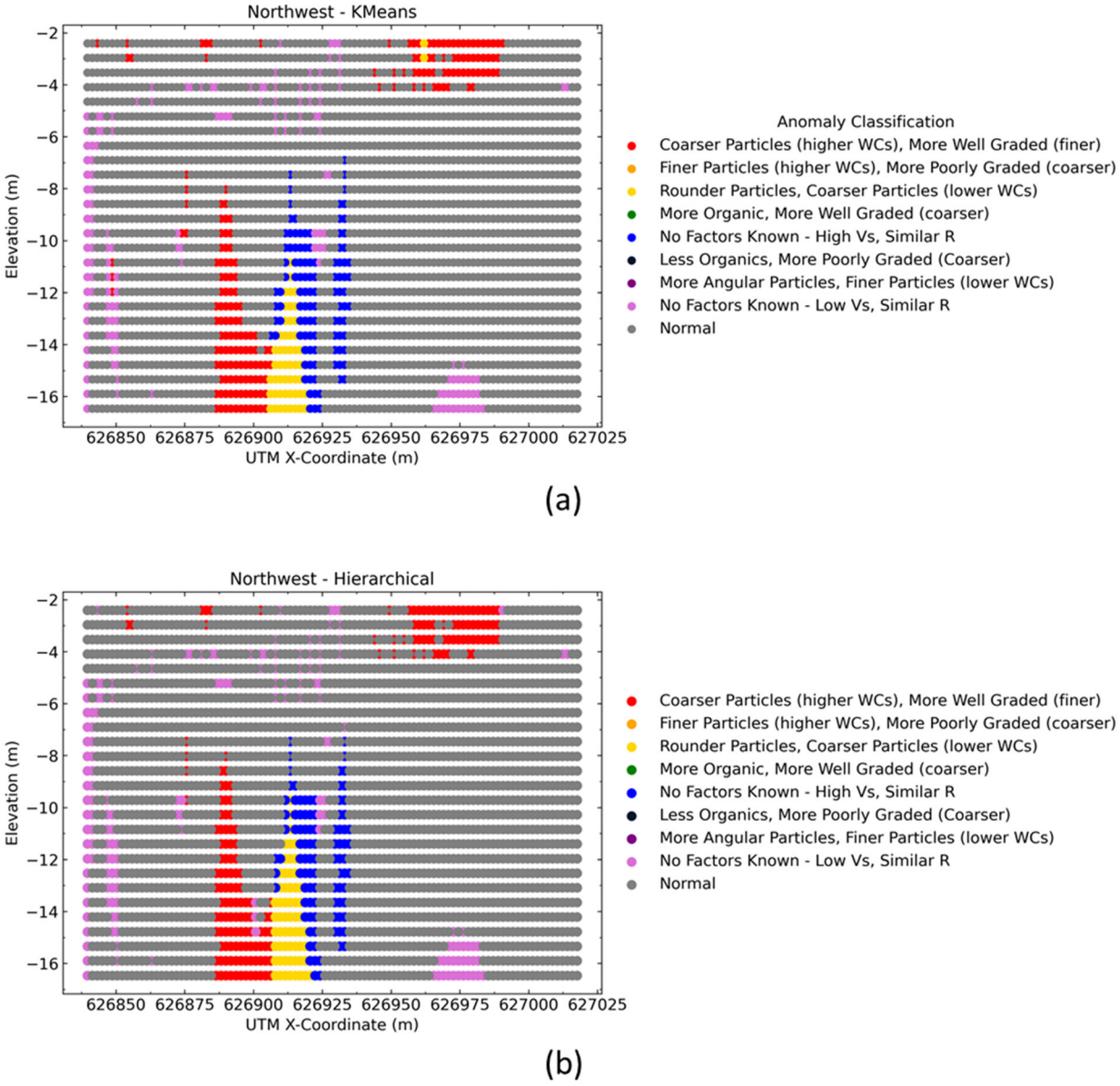


Figure 4.64 Subsurface profiles for the Northwest Site showing potentially problematic zones using (a) K-Means Clustering and (b) Hierarchical Clustering.

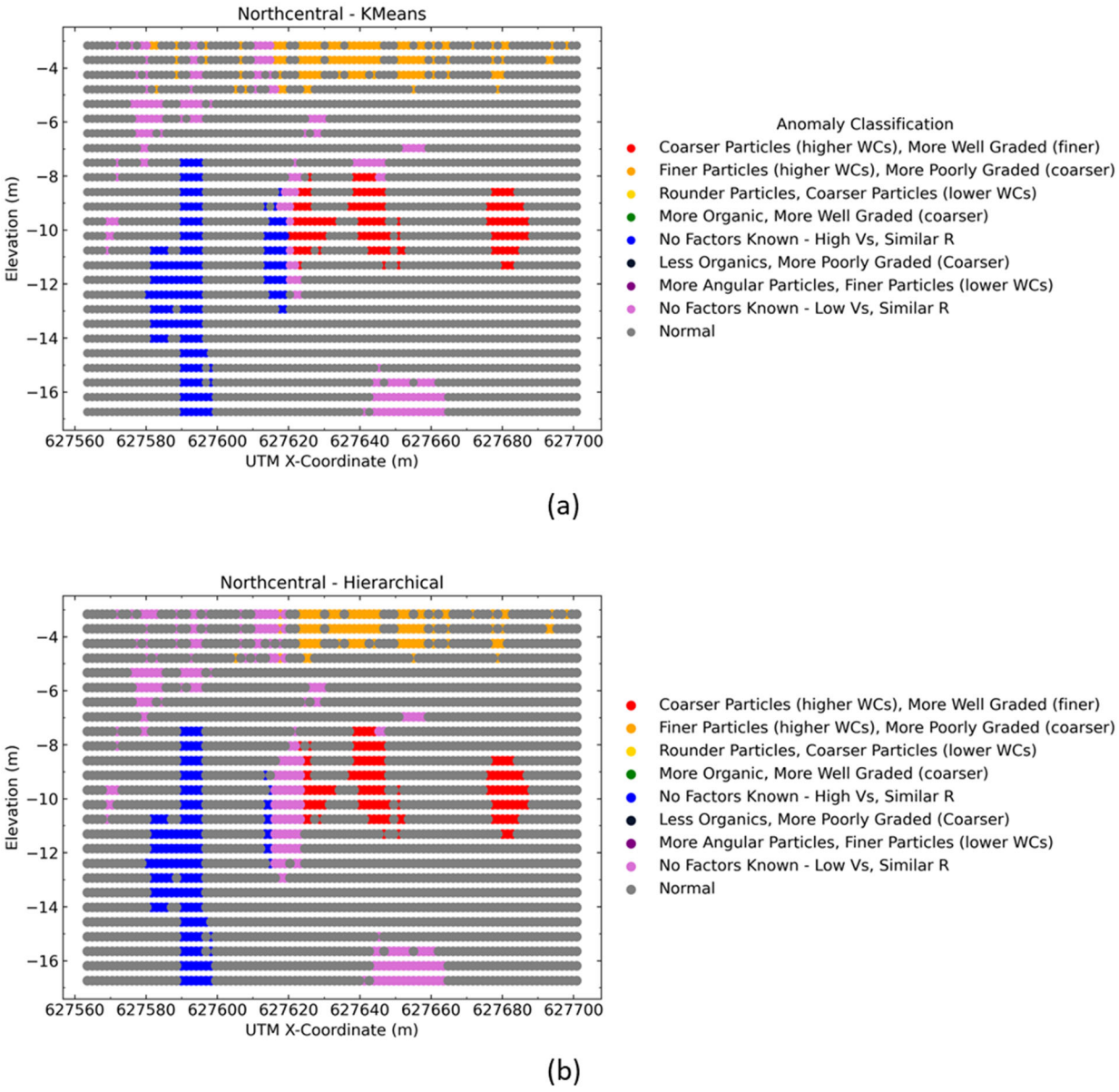


Figure 4.65 Subsurface profiles for the Northcentral Site showing potentially problematic zones using (a) K-Means Clustering and (b) Hierarchical Clustering.

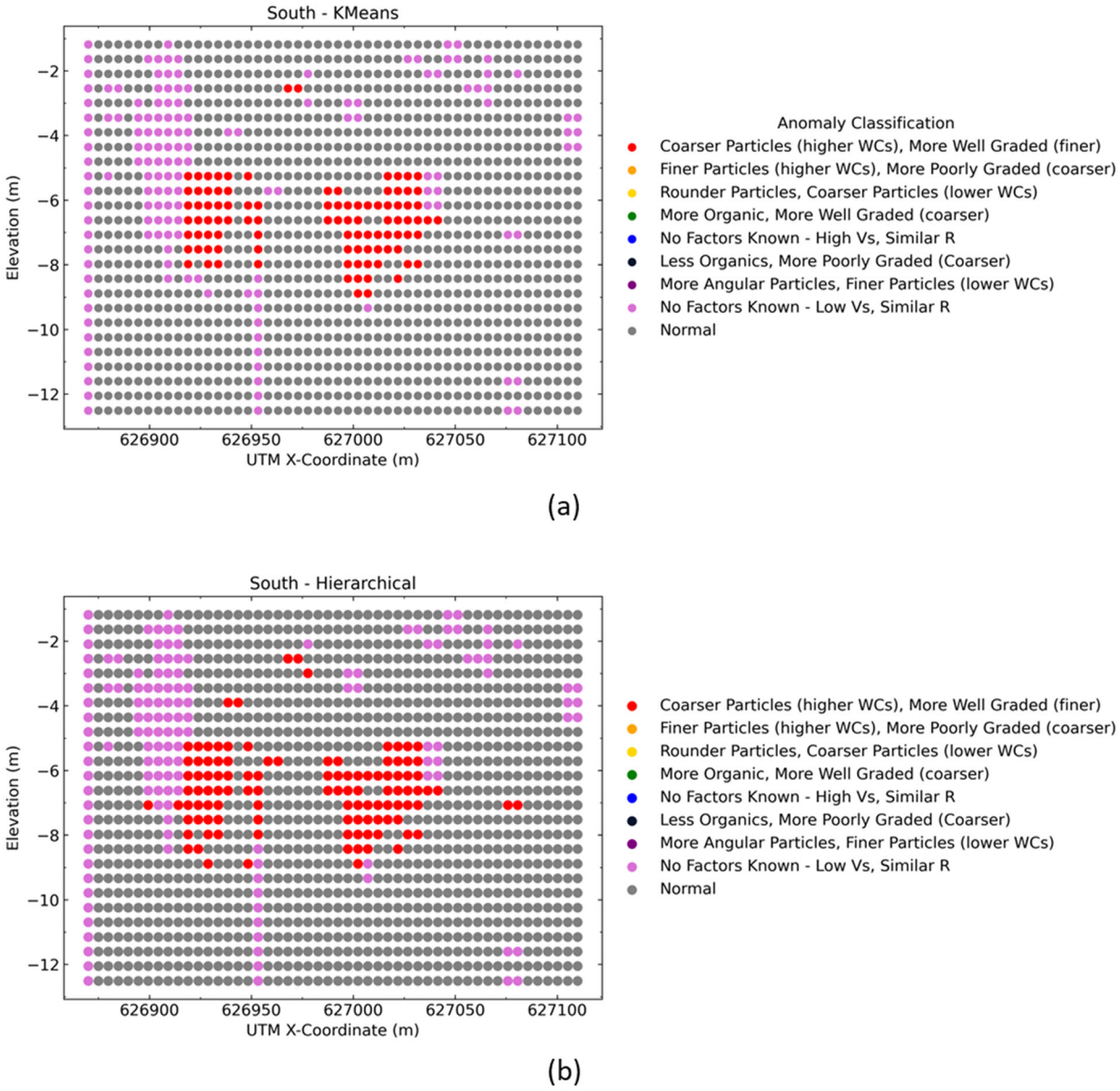


Figure 4.66 Subsurface profiles for the South Site showing potentially problematic zones using (a) K-Means Clustering and (b) Hierarchical Clustering.

When comparing the K-Means and Hierarchical clustering potentially problematic zone profiles for each site, there was relatively good agreement between the profiles. Between each profile for each site, there were slight differences in the boundaries between clusters. Even with the differences in boundaries, the geotechnical conditions labels were the same between K-Means and Hierarchical clustering.

4.10.2 Subsurface Anomalous Zones Compared to SBTn Distribution

The potentially problematic zones that were identified using clustering and anomaly detection were compared with the spatial distribution of the SBTn to understand what subsurface conditions may exist at these locations. The distribution of SBTn for each potentially problematic zone label for the Northwest Site, the Northcentral Site, and the South Site, are shown in Figure 4.67, Figure 4.68, and Figure 4.69, respectively.

The Northwest Site had four different geotechnical condition labels for potentially problematic zones. The areas of the profile that were considered "normal" in the clustering and anomaly detection analysis showed that most of the profile had SBTn values of 2 and 3 which represent clays to clayey silts (refer to Table 3.1.), shown in Figure 4.70(e). The potentially problematic zones that were classified as areas with potentially coarser particles (at higher water contents) or a more well graded soil (at finer particles) were mostly composed of SBTn values of 5 and 6 which were silty sands and clean sands with some 7 and 8 which represent dense sands to gravelly sands as well as very stiff sands. While these SBTn values have no information about stiffness, the sands in these zones have coarser particles compared to the silty clays in the majority of the profile.

The problematic zones in the Northwest Site that were classified as rounded particles or coarser particles (at lower water contents) overlapped with mostly SBTn of 5 and 6, as shown in Figure 4.67(b). The laboratory label agrees with the geotechnical conditions found using CPTs as the silty sand to sand will have a larger grain size than silty clays and clays that are found in most of the profile.

The final two potentially problematic zones showed high Vs and similar ρ as well as low Vs and similar ρ. The distributions of SBTn for both potentially problematic zones were similar with almost equal distributions of 2 and 3 which represent organic soil to silty clay and 5 and 6 which represent sandy silt to clean sand, as shown in Figure 4.67(c) and Figure 4.68(d), respectively. No geotechnical conditions controlled in the laboratory showed these changes in Vs and ρ, but more clayey and sandy soils could be the cause for these labels.

The Northcentral Site had four different geotechnical conditions labels for potentially problematic zones. Figure 4.68(e) shows the distribution of SBTn in the areas of the profile that were determined to be "normal" from the clustering and anomaly detection analysis with SBTn 3 being the most prevalent. This indicates most of the profile is silty clay or clay. The potentially problematic zones that were classified as areas with potentially coarser particles (at higher water contents) or a more well graded soil (at finer particles) were mostly composed of SBTn 3 and 4, with some 5 and 6. This indicates that these zones are mostly clays to clayey silts with sandy silt to clean sands. The coarser-grained sands compared to the clay in the "normal" component of the profile could be causing the changes in Vs and ρ that are represented in the clustering and anomaly detection analysis.

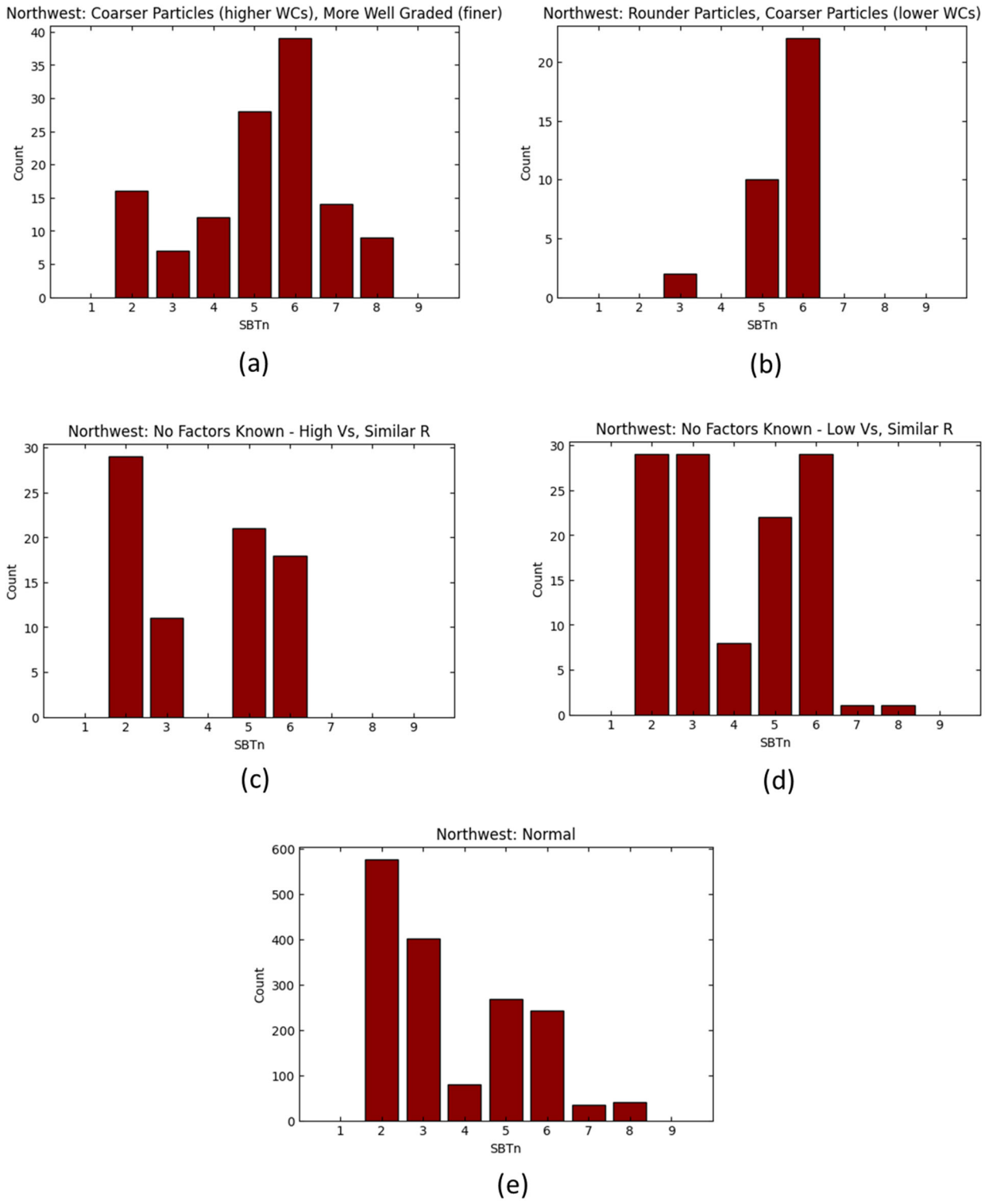


Figure 4.67 Distribution of SBTn for Northwest Site for potentially problematic zones (a) coarser particles (higher WCs), more well graded (finer), (b) rounder particles, coarser particles (lower WCs), (c) high Vs, similar resistivity, (d) low Vs, similar resistivity, and (e) normal – not at a potentially problematic zone.

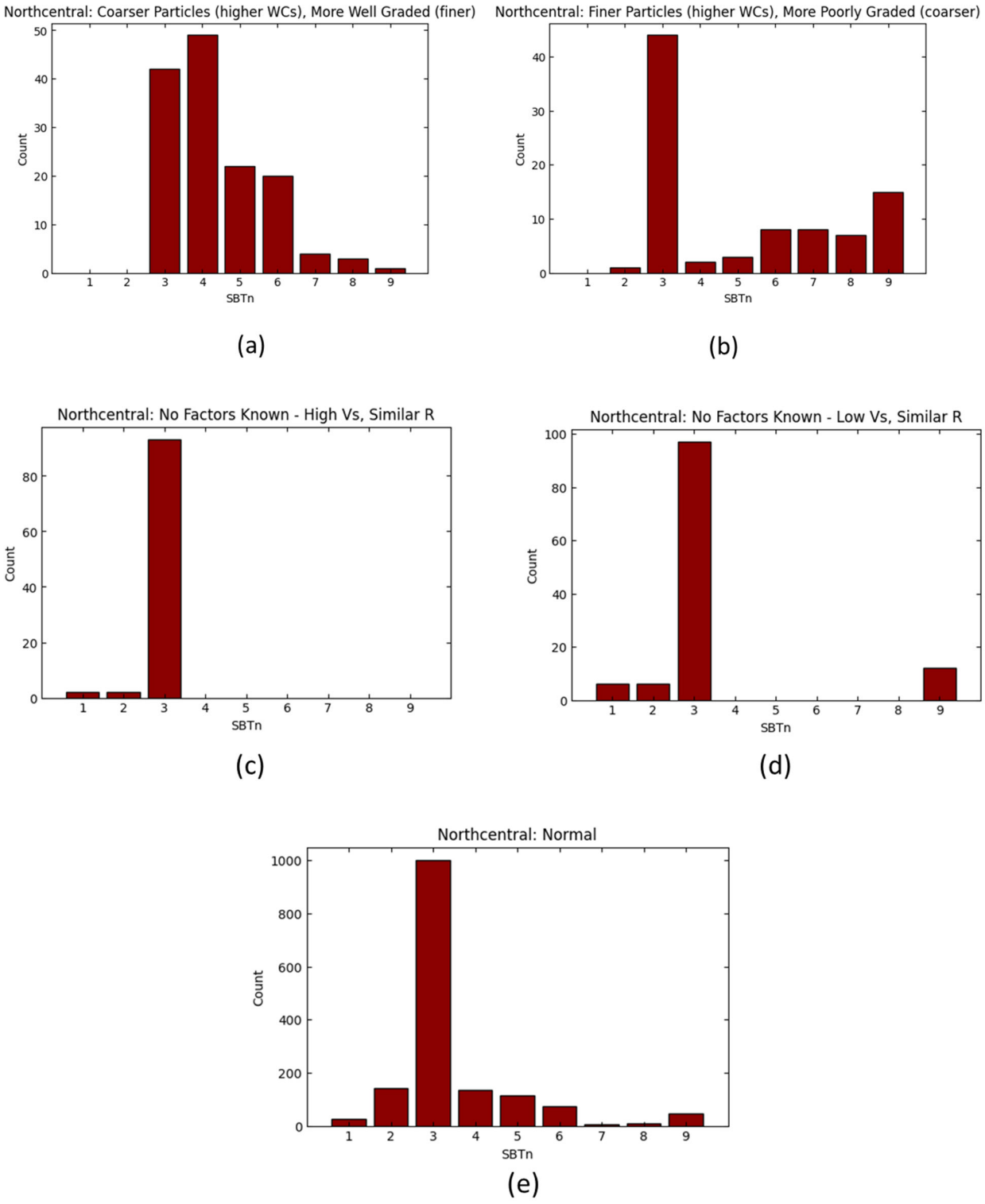


Figure 4.68 Distribution of SBTn for Northcentral Site for potentially problematic zones (a) coarser particles (higher WCs), more well graded (finer), (b) finer particles (higher WCs), more poorly graded (coarser), (c) high Vs, similar resistivity, (d) low Vs, similar resistivity, and (e) normal – not at a potentially problematic zone.

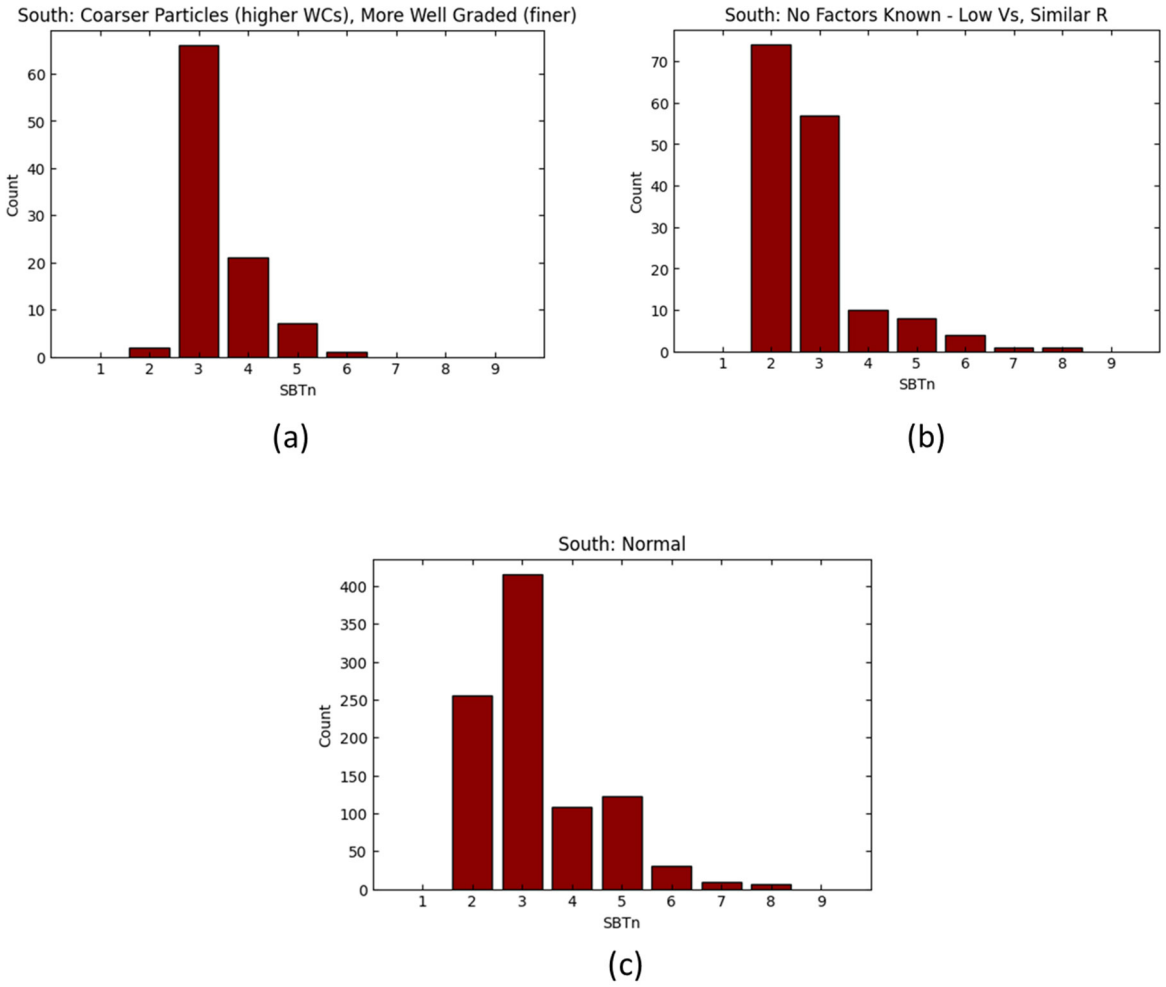


Figure 4.69 Distribution of SBTn for South Site for potentially problematic zones (a) coarser particles (higher WCs), more well graded (finer), (b) low Vs, similar resistivity, and (c) normal – not at a potentially problematic zone.

The potentially problematic zones in the Northcentral Site that were classified as finer particles (at higher water contents) and more poorly graded (for coarser particles) were mostly composed of SBTn 3, silty clay to clay. There were trace amounts of organic soils and silty clay to very stiff sands and fine-grained soils. These subsurface conditions are mostly finer particles that align with the potentially problematic zones label.

The remaining problematic zones were classified as high *Vs* and similar ρ as well as low *Vs* and similar ρ, shown in Figures 4.68(c) and 4.68(d), respectively. Both zones were dominated by SBTn 3, silty clays to clays with very little presence of other soils. Clayey soils could be what is causing these changes in *Vs* and ρ as clays were not measured in the laboratory.

The South Site had two geotechnical conditions labels for potentially problematic zones. The distribution of SBTn in most of the profile is shown as 2 and 3, representing organic soils to clay with some 4 and 5, silty clay to silty sand. The zones that were considered coarser particles (at

higher water contents) or more well graded (at finer particles) were dominated by SBTn of 3 and 4, silty clays to silty sands. These results validate the label assigned to the potentially problematic zone.

The final potentially problematic zones were classified as low *Vs* and similar ρ. These zones were dominated by 2 and 3, organic soils to silty clays. This confirms what the other sites showed where these potentially problematic zones were also dominated by clays.

The results from the sites along the Bacon Island levee system confirm the results from the clustering and anomaly detection analysis. Where most of the profiles were silty clays or clays, the areas that were different tended to be classified as anomalous with laboratory labels that described the change.

4.10.3 Unsupervised Results Compared to Supervised Results (Using all Features)

The clustering and anomaly detection analysis resulted in a subsurface profile showing potentially problematic zones while the supervised machine learning analysis resulted in a subsurface profile showing where SBTn was correctly or incorrectly predicted. This analysis compares the two profiles to understand if areas that are considered potentially problematic correctly classify SBTn more than areas that were not potentially problematic. The supervised analysis used was predicting SBTn at each site when using all features. This analysis was used as it had the best accuracy compared to predicting SBTn at each site using only *Vs* and ρ.

The analysis of comparing the potentially problematic zones with the supervised prediction profiles was done by finding the percentage of correctly classified SBTn values within each potentially problematic zone and the "normal" zone for each site at Bacon Island. The results for the Northwest Site are summarized in Table 4.20.

Table 4.20 Percent Correctly Predicted SBTn per Potentially Problematic Zone for the Northwest Site for all Features

	% Correctly Predicted in Problematic Zone		
	KNN	**RF**	**SVM**
Coarser/Well Graded	23	6	33
Rounder/Coarser	15	6	29
High Vs, Similar ρ	30	14	30
Low Vs, Similar ρ	17	29	13
Normal	21	26	20

For the Northwest Site, SVM had the highest number of correctly predicted SBTn values for potentially problematic zones that were labeled as coarser particles (at higher water contents), more well graded (at finer particles) and rounder particles, coarser particles (at lower water contents). In comparison, RF showed minimal correctly predicted SBTn values, except for the zones classified as low *Vs* and similar ρ. In terms of comparison between correctly predicted SBTn of potentially problematic zones with the "normal" portion of the profile, the percentage of correctly predicted data in the "normal" profile is similar to the percentage of correctly predicted data in the potentially problematic zones. For the Northwest Site, the potentially problematic zones did not affect whether SBTn was correctly classified or not.

The results for the Northcentral Site showing the percentage of correctly predicted SBTn for each potentially problematic zone and the "normal" section of the profile are shown in Table 4.21.

Table 4.21 Percent Correctly Predicted SBTn per Potentially Problematic Zone for the Northcentral Site for all Features

	% Correctly Predicted in Problematic Zone		
	KNN	**RF**	**SVM**
Coarser/Well Graded	38	26	30
Finer/Poorly Graded	16	22	50
High Vs, Similar ρ	10	69	96
Low Vs, Similar ρ	56	60	80
Normal	47	42	64

For the Northcentral Site, SVM showed high percentages of correctly predicted SBTn for the potentially problematic zones. RF also showed high percentages of correctly predicted data for zones that were labeled as high *Vs*, similar *ρ,* and low *Vs*, similar *ρ*. KNN had high percentages of correctly predicted data for zones that were labeled as coarser particles (at higher water contents), more well graded (at finer particles), and low Vs, similar ρ. The percentage of correctly predicted data in the "normal" part of the profile was similar to the percentages of correctly predicted data for the potentially problematic zones.

The results for the South Site showing the percentage of correctly predicted SBTn for each potentially problematic zone and the "normal" section of the profile are shown in Table 4.22.

Table 4.22 Percent Correctly Predicted SBTn per Potentially Problematic Zone for the South Site for all Features

	% Correctly Predicted in Problematic Zone		
	KNN	**RF**	**SVM**
Coarser/Well Graded	11	56	7
Low Vs, Similar ρ	14	23	5
Normal	19	29	9

The South Site showed that RF had the highest percentage of correctly predicted SBTn for each potentially problematic zone compared to KNN and SVM. SVM had the lowest percentage of correctly predicted SBTn. The percentage of correctly predicted SBTn in the "normal" component of the profile was like the percentage of correctly predicted SBTn for each potentially problematic zone, for each supervised machine learning algorithm. All three sites at Bacon Island had the same result where there was no significant difference in how many data points were correctly classified between a potentially problematic zone and the "normal" portion of the profile.

4.10.4 Unsupervised and Supervised Machine Learning for Levee Subsurface Health Investigation

Unsupervised and supervised machine learning have their advantages and disadvantages for levee health assessment. The unsupervised machine learning analysis does not need labels, or the right

answer, to create a good model. It has the ability to use only non-invasive geophysics techniques to identify potentially problematic zones in the subsurface using cluster and anomaly detection techniques. However, there is uncertainty in these zones regarding detailed knowledge of what is causing these zones and if failure mechanisms are currently found in the field.

Supervised machine learning analysis though can use the non-invasive geophysical data to train the model to predict ground truth data. The ground truth data provides more certainty in understanding the subsurface conditions but is more difficult to collect. Many stakeholders are hesitant to collect ground truth data as invasive methods like drilling and CPTs are typically needed. This is problematic as a good, supervised machine learning model requires a large dataset. This would require thousands of stretches of levees to have CPT or borehole data for a range of levee subsurface conditions. If a dataset can be created with high enough spatial resolution, the model could be applied to stretches of levees where ground truth data does not exist or is difficult to collect. This will allow stakeholders to make more informed decisions regarding levee health by understanding the subsurface stratigraphy.

Having the Bacon Island dataset that had both Vs and ρ data along with CPT data allowed for the validation of the unsupervised machine learning framework. By comparing the laboratory labels with their respective potentially problematic zones with the SBTn values at the same area in the subsurface, it can be determined if the unsupervised methodology is useful. After comparing the Bacon Island unsupervised results with the supervised results for the three sites, it was determined that the unsupervised framework was reasonable for understanding subsurface health without the need for ground truth data. The SBTn values matched locations in the unsupervised framework that had similar changes in characteristics, such as changes in soil grain size. While there are still laboratory labels in the unsupervised framework that were not validated using the SBTn data, such as organics and gradations, having the grain size validation shows the potential of this framework and it’s applicability to understanding subsurface levee conditions.

Chapter 5. Data Fusion and Hazard Maps

5.1 Robust Levee Health Framework

Levee inspection practices currently have no information about subsurface health, but it does contain surface information. To understand the health of the system, it is imperative to analyze the surface and subsurface health in tandem. To do this, the surface change detection analyses must be fused with the subsurface geophysical change detection and machine learning analyses. One major complication of fusing surface and subsurface data is the spatial inconsistency between the two datasets. The surface data is two-dimensional in the x- and y-direction while the subsurface data is two-dimensional in the x- and z-direction. This study projected the subsurface data to the surface to address the issues of spatial inconsistency between the surface and subsurface data.

With the subsurface analyses projected to the surface, the next step was to decide the significance of the surface and subsurface data layers in the analysis and to understand what information each layer conveyed regarding levee health. Once the layers are finalized, negative weights are assigned to each surface and subsurface layer to represent the hazard they contribute to the system.

The final health of the system is modeled by spatially adding the weights of the layers to create a hazard map of the levee segment. Increased weights on the hazard map represent areas subjected to increased hazard based on the surface and subsurface analyses. These hazard maps can justify more invasive investigations and maintenance.

5.2 Projection of Subsurface Data to the Surface

5.2.1 Change Detection Analysis Projection

For the final data fusion framework, the subsurface change detection had to be projected to the surface to integrate with the surface data. This was done by looking at how the percent change changed along the length of the survey line for each category of change: no meaningful change, greater than 10% change, and less than -10% change. For each unique combination of the x- and y-coordinate, the percent change of one category of change was averaged over the entire range of elevation for that x- and y-coordinate. That averaged percent change was then attributed to that

specific x- and y-coordinate as the "projected" change at the surface. This was repeated for all the change categories. An example of this is shown in Figure 5.1.

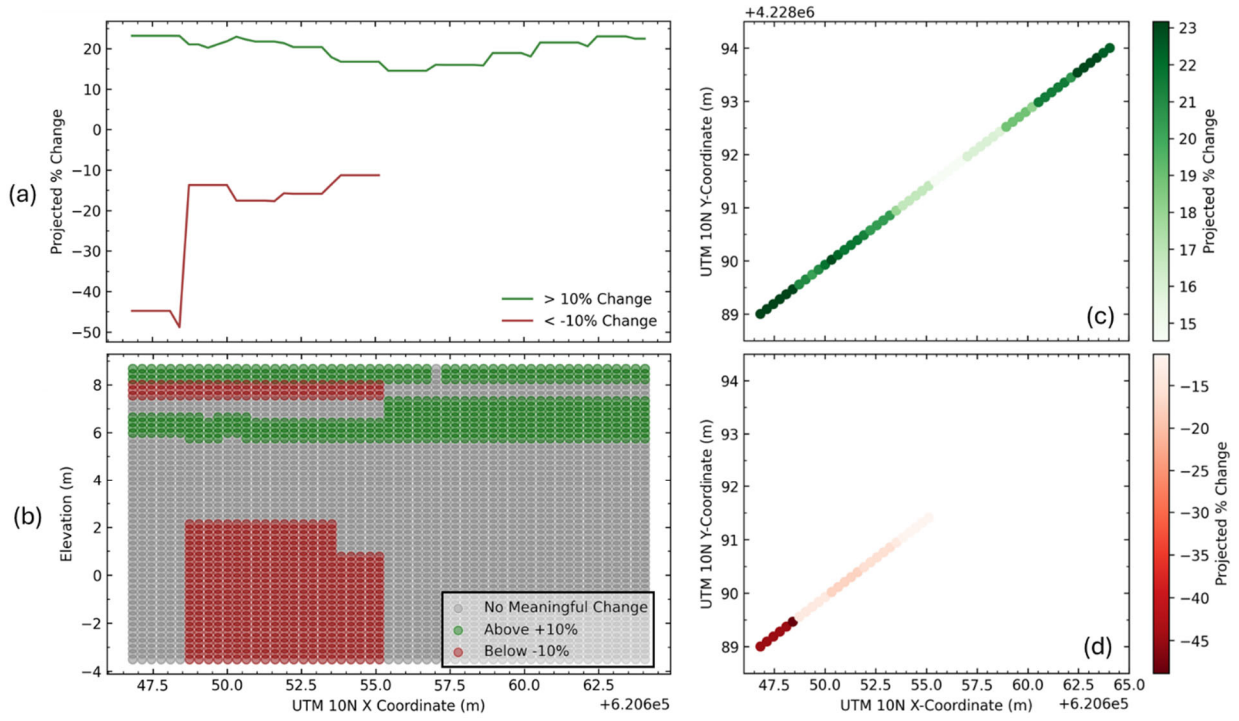


Figure 5.1 Projection of subsurface change detection to the surface.

Figure 5.1(a) shows the "projected" percent change over the length of the survey line for both categories of significant change while Figure 5.1(b) shows the spatial distribution of the change categories. Figure 5.1(c) and Figure 5.1(d) show what the data looks like when projected to the surface and the darker the color along the line, the higher the percent change at depth below that point. Figure 5.1(c) shows the percent change above 10% and shows that at either end of the survey, there are higher changes between the wet and dry seasons compared to the middle of the survey. Figure 5.1(d) is the percent change below -10%. The projection ends halfway along the survey line because after an x-coordinate of 620655 m, there are no areas at depth that represent less than -10% change, which is also shown in Figure 5.1(a). When there was no category at depth, there was no projected data point at the surface.

5.2.2 Potentially Problematic Zones Projection

The machine learning results that fused clustering and anomaly detection to identify areas that were potentially problematic in the subsurface had to be projected to the surface. For each subsurface profile, the different laboratory-labeled zones were projected to the surface, separately. This was done by looking at each combination of the x- and y-coordinate and determining if a potentially problematic zone existed in the subsurface. If the potentially problematic zone existed in the profile for that x- and y-coordinate, then that coordinate pair was given a value of 1. If that zone did not exist in the profile, the coordinate pair was given a value of 0. An example of this is shown in Figure 5.2.

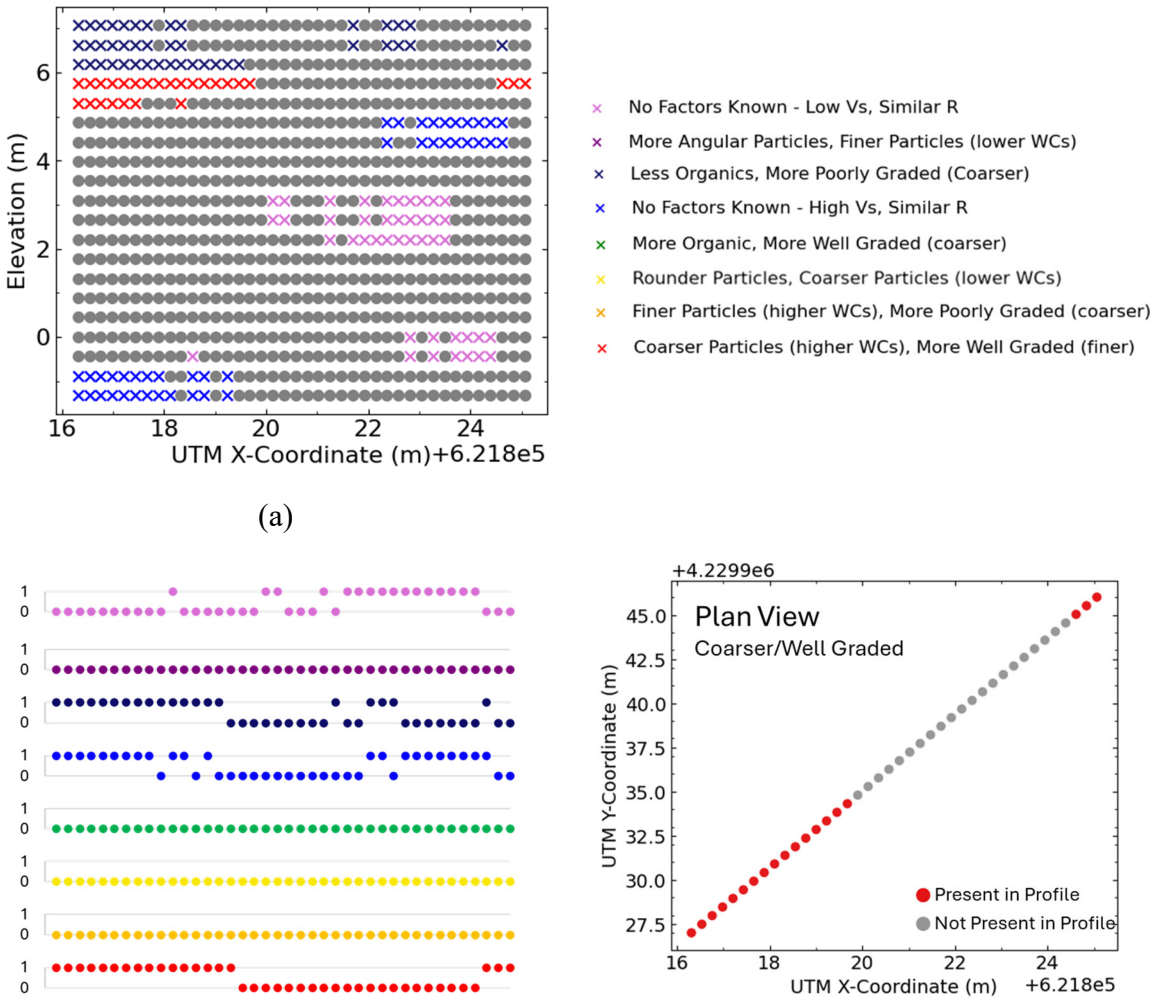


Figure 5.2 Projection of subsurface machine learning analysis to the surface showing (a) the subsurface profile, (b) the binary classification with respect to the x-coordinate for all labels, and (c) the projected data at the surface.

In this profile, four different potentially problematic labels exist: (1) coarser particles (at higher water contents), more well graded (with finer particles), (2) no factors known, high *Vs* and similar ρ, (3) less organics, more poorly graded (with coarser particles), and (4) no factors known, low *Vs* and similar ρ. For each of these, their distribution of 0 and 1 are shown in Figure 5.2(b) For the four labels that do not exist in this profile, all x- and y-coordinate pairs were given a value of 0, indicating that they do not exist as shown in Figure 5.2(b).

Each one of the potentially problematic zone labels are then represented on the surface by plotting points at each x- and y-coordinate pair and coloring the points by whether the zone existed in the subsurface. An example of the plan view is shown in Figure 5.2(c) where the line is plotted with respect to the x-coordinate and the y-coordinate showing the results for the zones labeled as coarser particles (with higher water contents), more well graded (for finer particles). In this plan view of

the surface, the red points represent that the potentially problematic zone existed in the profile at depth while the grey points represent that the potentially problematic zone does not exist in the profile at depth.

5.3 Model Layers

The model layers for the final hazard model are spatial layers representing different levee hazards and failure modes. These layers represent either surface features or subsurface features. The surface layers were created by analyzing the change in the UAV collected data between seasons while the subsurface layers show the projected change detection analyses and projected machine learning analyses using the geophysical data. This section explains the different surface and subsurface layers that were analyzed and whether they were included in the final hazard model.

5.3.1 Surface Layers

The surface layers are a result of the change detection techniques performed on the UAV-derived data to explore how different transformations of the data highlight surface features. The classification change detection resulted in three layers called NoChange, GreenDead, and DeadGreen. The thermal change detection resulted in two layers called ThermalDiff and NRCT. Five layers were explored from the optical change detection, NormGR, Band1, Band2, Band3, and AddBand. The LIDAR change detection had three layers, Aspect, Slope, and ElevationDiff. A description of these thirteen surface layers and whether they were included in the hazard model is described below. The layers can be found in Appendix E.

- NoChange: this is a layer derived from the classification change detection analysis. It shows pixels that had a change in classification between seasons and pixels that had no change in classification between seasons. This layer was not selected for the analysis as it does not have enough detail about how pixels changed or the degree of hazard at the site.
- GreenDead: this is a layer derived from the classification change detection analysis. It shows pixels that were classified as vegetation in the dry season and then were classified as dead vegetation in the wet season. This layer was used in the analysis as it represents an increase in vegetation in the dry season which can be an indicator of seepage.
- DeadGreen: this is a layer derived from the classification change detection analysis. This layer shows pixels that were classified as vegetation in the wet season but not in the dry season. This layer was used in the analysis as it shows an increase in vegetation in the wet season which could be problematic for access to levee slopes, could create preferential conduits for water flow, but can also help prevent surface erosion and bank reinforcement.
- ThermalDiff: this is the difference in thermal temperature between the dry and wet seasons, showing positive and negative changes to preserve seasonal effects. This layer shows differences in everything that could cause a change in temperature including shadows, vegetation, and bare soil. This layer was kept for the model as it can give an indication of changes in bare soil, near surface cavities, vegetation, and urban structures.
- NRCT: this layer shows the difference in the Normalized Relative Canopy Temperature between the dry and wet seasons. This layer shows the difference between vegetation.

Since this layer only shows vegetation and other layers already show these trends, it was not included in the final hazard model.

- NormGR: this layer shows the difference in the Normalized Green-Red difference index. This is a vegetation index that shows how vegetation has changed over time. This layer was not included in the analysis as it is not a common vegetation index when using optical imagery. This one was initially explored since the NIR band was not available. Other layers were deemed to highlight vegetation better.
- Band1: this layer shows the difference in the red band brightness values. With difficulty, this difference can show changes in healthy vegetation because vegetation reflects red light. As a result, low reflectance can indicate healthy vegetation. Due to the difficulty of looking at vegetation with only the red band though, it was not used in the hazard model.
- Band2: this layer shows the difference in the green band brightness values between the dry and wet seasons. Because vegetation reflects green light, a high green reflectance can be used to identify vegetation. It was decided not to use this layer as other layers modeled vegetation better.
- Band3: this layer shows the difference in the blue band brightness values between the dry and wet seasons. Without the NIR band, the change in the blue band is meaningless and does not highlight any significant change. As a result, it was not used in the analysis.
- AddBand: this layer was created by adding the differences of the red, green, and blue bands to make a composite difference band. After analyzing this layer, it was determined that the layer does not provide meaningful data about changes in surface features and was not included in the final model.
- Aspect: this layer shows the changes in the aspect between the dry and wet seasons. The change in aspect would be useful for landslide detection. Changes in the riprap along the wet toe of the levee dominated this layer and would have created an anomalously high hazard at the toe. As a result, this layer was not included in the final model.
- Slope: this layer shows the difference in slope between the seasons. This layer is useful for identifying landslides on the slopes of the levees. As a result, it was included in the final hazard model.
- ElevationDiff: this layer represents the change in elevation in the levee surface between the seasons. Significant elevation changes can be a result of toe bulging and crest depressions. As a result, it was included in the final hazard model.

Of the thirteen surface layers that were explored using change detection techniques, five were selected to be included in the final model. These five were selected as they highlighted different failure mechanisms and did not significantly overlap failure mechanisms with other layers.

5.3.2 Subsurface Layers

The subsurface layers are the projected layers from the geophysical change detection analysis as well as the machine learning analysis. The geophysical change detection analysis resulted in six subsurface layers, EM_NoChange, EM_NegativeChange, EM_PositiveChange, MASW_NoChange, MASW_NegativeChange, and MASW_PositiveChange. All of these layers were included in the final hazard models.

- EM_NoChange: this layer shows the percentage of profile depth for each x- and y-coordinate pair with no significant change for the EMI surveys.

- EM_NegativeChange: this layer shows the amount (average percentage) of significant change in the EMI profiles for each x- and y-coordinate pair when the wet season had higher ρ.
- EM_PositiveChange: this layer shows the amount (average percentage) of significant change in the EMI profiles for each x- and y-coordinate pair when the dry season had higher ρ.
- MASW_NoChange: this layer shows the percentage of profile depth for each x- and y-coordinate pair with no significant change for the MASW surveys.
- MASW_NegativeChange: this layer shows the amount (average percentage) of significant change in the MASW profiles for each x- and y-coordinate pair when the wet season had higher *Vs*.
- MASW_PositiveChange: this layer shows the amount (average percentage) of significant change in the MASW profiles for each x- and y-coordinate pair when the dry season had higher *Vs*.

The subsurface layers that show the machine learning analysis contained eight layers for the dry season and eight layers for the wet season. For each season, a separate projected layer was used to represent each of the potentially problematic zone labels from the laboratory investigation, resulting in a total of sixteen layers.

- Coarse/Well Graded: this layer represents areas in the subsurface where potentially problematic zones were identified as potentially having coarser particles (at higher water contents) or more well-graded soil (with finer particles). Coarser particles could contribute to seepage, as shown by the Hazen Method where coarser particles result in higher hydraulic conductivity.
- Fine/Poorly Graded: this layer represents areas in the subsurface where potentially problematic zones were identified as potentially having finer grain size (at higher water contents) or more poorly graded soils (with coarser particles). The poorly graded soils could contribute to seepage since they have more void space in the soil matrix which results in an increased soil permeability.
- Round/Coarse: this layer represents areas in the subsurface where potentially problematic zones were identified as potentially having rounder particles or coarser particles (at lower water contents). Coarser particles are more likely to contribute to seepage than finer particles through the Hazen Method.
- Organic/Well Graded: this layer represents areas in the subsurface where potentially problematic zones were identified as potentially having more organic matter in the soil or more well graded soil (for coarser particles). Having organics can either contribute to seepage by providing preferred conduits for water movement but can also provide stabilization against slope and bank erosion.
- Less Organics/Poorly Graded: this layer represents areas in the subsurface where potentially problematic zones were identified as potentially having less organics or more poorly graded soil (with coarser particles). Poorly graded soil has more void space in the soil matrix which results in a higher soil permeability and is more likely to contribute to seepage.
- Angular/Finer: this layer represents areas in the subsurface where potentially problematic zones were identified as potentially having more angular particles or finer particles (at

lower water contents). Soils with angular particles contribute to seepage as they tend to have higher void ratios which lead to higher porosity.

- High *Vs*/Similar ρ: this layer represents areas in the subsurface where potentially problematic zones were identified as having higher *Vs* compared to the profile with very little change in the ρ. No factors controlled for in the laboratory investigation showed these trends in *Vs* and ρ.
- Low *Vs*/Similar ρ: this layer represents areas in the subsurface where potentially problematic zones were identified as having lower *Vs* compared to the profile with very little change in the ρ. No factors controlled for in the laboratory investigation showed these trends in *Vs* and ρ.

The machine learning subsurface layers are a function of the clustering method used, either K-Means or Hierarchical clustering as well as the compilation dataset used to fit and predict the data. As a result, the change detection layers for the surface and subsurface are constant for all models but different models are needed to represent the different clustering algorithms and compilation datasets.

5.4 Model Weights

The hazard model is created by spatially summing weights associated with the surface and subsurface layers. Since the model represents hazard, only negative weights were assigned to layers with higher negative weights associated with more hazard. The range of weights given for a given layer was between 0 and -1. This section describes how weights were assigned to each layer and the weight given.

5.4.1 Surface Weights

The surface layers were given weights based on the surface features they highlight. Table 5.1 summarizes the surface layers with the potential levee hazard, the weight assigned, and the threshold of change applied, if applicable.

Table 5.1 Summary of Surface Layers Weights

Name	Hazard	Threshold	Weight
GreenDead	Seepage	N/A	-1
DeadGreen	Preferred conduit for water movement, access to slopes	N/A	-0.5
ThermalDiff	Vegetation, animal burrows, bare soil/erosion	+/- 1 STD	-0.75
Slope	Slope stability	+/- 1 STD	-1
ElevationDiff	Toe bulging, crest depressions	10%	-1

For the GreenDead and DeadGreen layers, no threshold of change was needed as the change was already represented by pulling out pixels that had a different classification between seasons. The pixels that represented a change in the GreenDead layer were given a weight of -1 as they showed an increase in vegetation in the dry season compared to the wet season, which indicates the potential for seepage. All other pixels were given a weight of 0 to represent no hazard.

The pixels in the DeadGreen layer that represented change were given a weight of -0.5 while all other pixels were given a weight of 0. This layer showed an increase in vegetation in the wet season compared to the dry season. This is problematic for levee health as vegetation on the slopes and toes can impede access to the levee surface during inspection. Dense vegetation can also cause preferred conduits for water movement through the structure. This layer was given a -0.5 instead of a -1 since vegetation can also benefit levee health, resulting in less of a hazard potential. A dense root system can help prevent surface erosion and bank erosion.

For the ThermalDiff and Slope layers, a threshold was needed to determine the change between the seasons. Work performed by Behrens (1998) stated that areas with little to no change are located near the mean of the distribution. The farther from the mean the change values are, the greater the amount of change. The change away from the mean is typically taken as standard deviations away from the mean. Once the appropriate standard deviation is identified, then the pixels that are within the tail of that standard deviation can be highlighted in the image to show areas with change.

Figure 5.3, Figure 5.4, and Figure 5.5 shows an example of looking at the 1 standard deviation and 2 standard deviations away from the mean for thermal change detection on Grand Island for Site 1, Site 2, and Site 3, respectively.

For these figures, the locations shown in blue are areas that fall within the standard deviation tail of interest. All three sites show significantly more change using a 1 standard deviation change threshold than a 2 standard deviation change threshold. The standard deviation that represented the actual change more accurately was the 1 standard deviation threshold. This was based on site observations during the dry and wet seasons and noting the drastic change in vegetation which is modeled best by the 1 standard deviation threshold.

The pixels in the 1 standard deviation tail for the ThermalDiff layer were given a weight of -0.75 instead of -1 because the changes could be attributed to changes in bare soil or near surface cavities, but it mostly shows changes in vegetation. Since other surface layers also show changes in vegetation, this layer's weight was reduced slightly to not overly penalize vegetation in the model. All other pixels that were not in the 1 standard deviation tail were given a weight of 0 to represent no hazard.

When creating the hazard maps, it was noticed that locations with shadows had anomalously high hazard as shown in Figure 5.6. The area circled in black in both photos is an area where shadows were recorded in the dry season but not in the wet season.

Since shadows have a large temperature change if not recorded in both the dry and wet season images, those areas were found to have significant thermal changes. In Figure 5.6(a), the shadows were included in creating the hazard map that area showed significant hazard. In comparison, Figure 5.6(b) shows the hazard for the same area with the same data except the shadows were removed from the analysis. This area now has significantly less hazard than when the shadows were included. As a result, it was important to remove the effects of shadows from the thermal change detection analysis. This was done by using the classification results from the dry and wet seasons to determine the locations of shadows. Wherever the location of shadows overlapped with significant thermal change, the thermal change locations were removed. This allowed for a more accurate representation of the hazard from the thermal change detection while not having shadows dominate the hazard at the site.

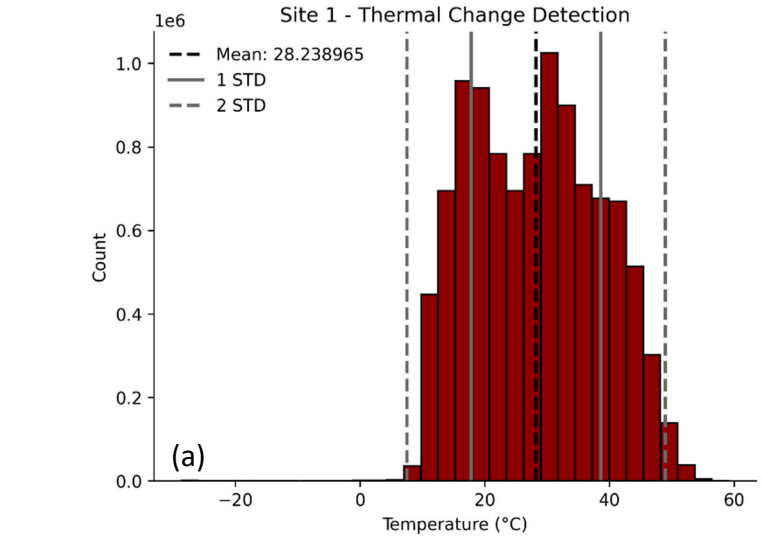






Figure 5.3 Thermal change for Site 1 showing (a) distribution of thermal differences, (b) 1 standard deviation of change, and (b) 2 standard deviation of change.

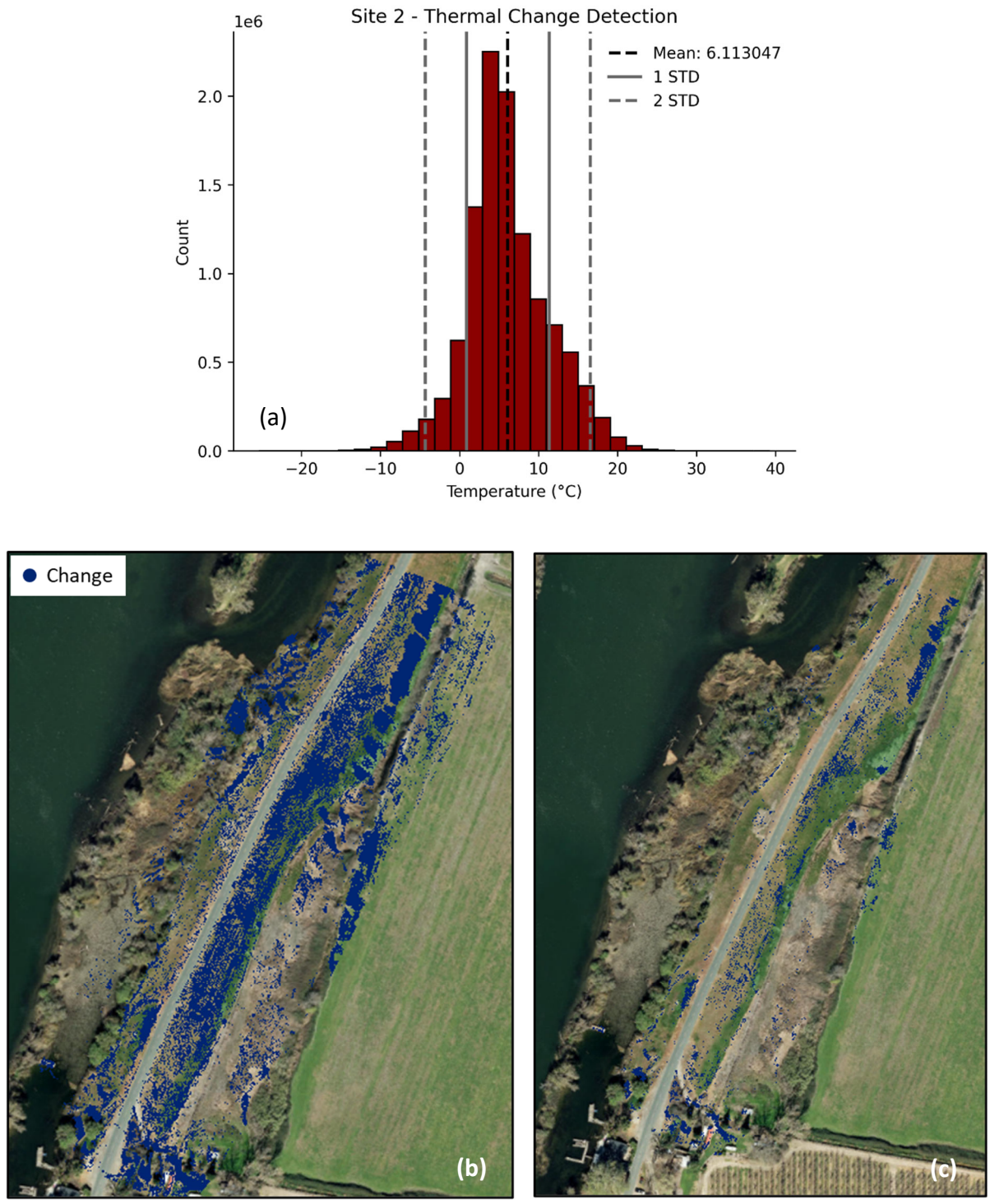
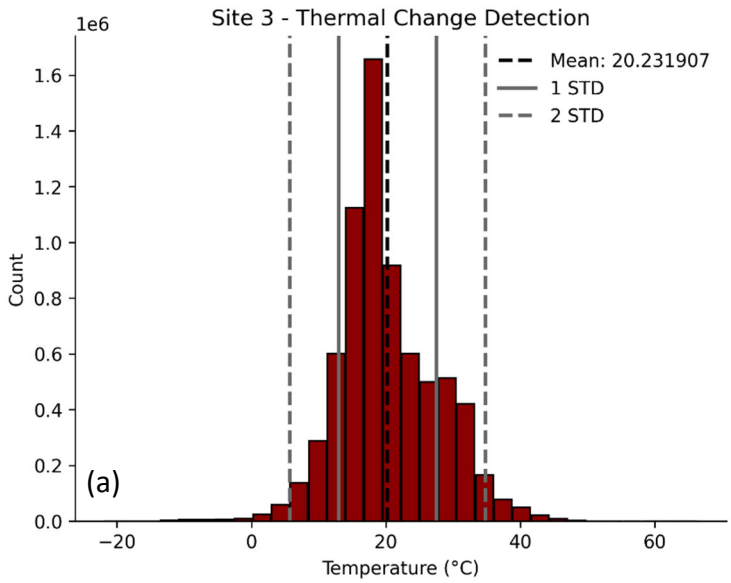


Figure 5.4 Thermal change for Site 2 showing (a) distribution of thermal differences, (b) 1 standard deviation of change, and (b) 2 standard deviation of change.



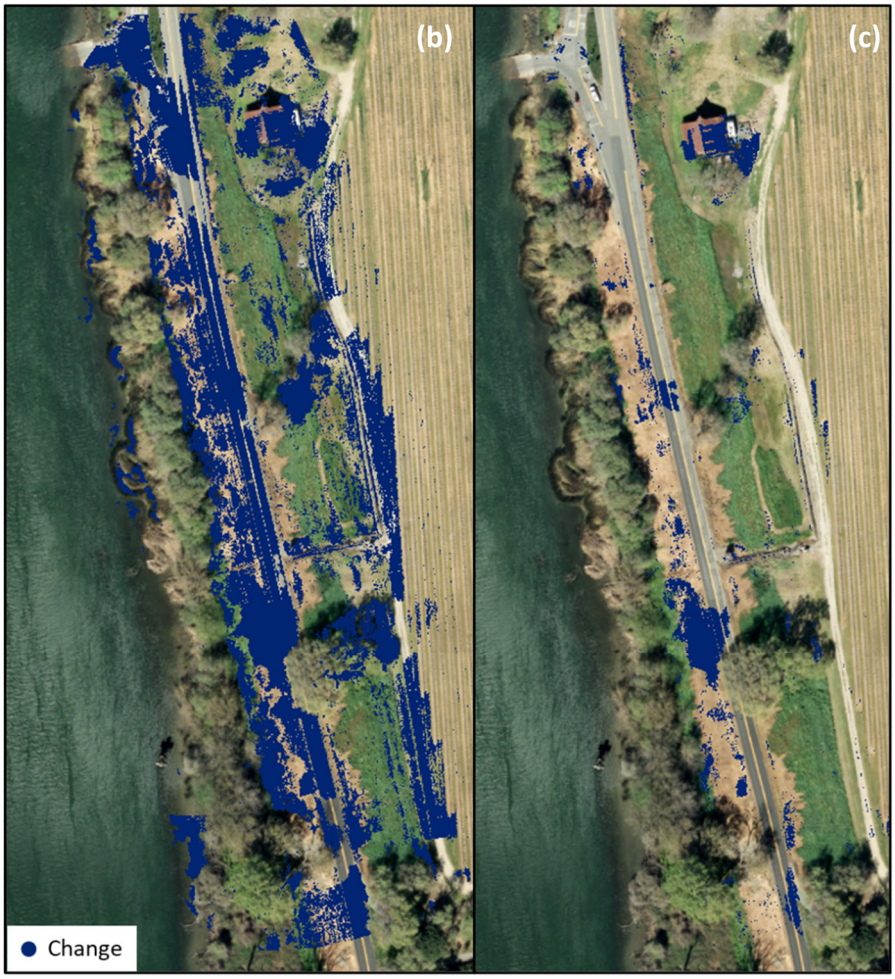


Figure 5.5 Thermal change for Site 3 showing (a) distribution of thermal differences, (b) 1 standard deviation of change, and (b) 2 standard deviation of change.

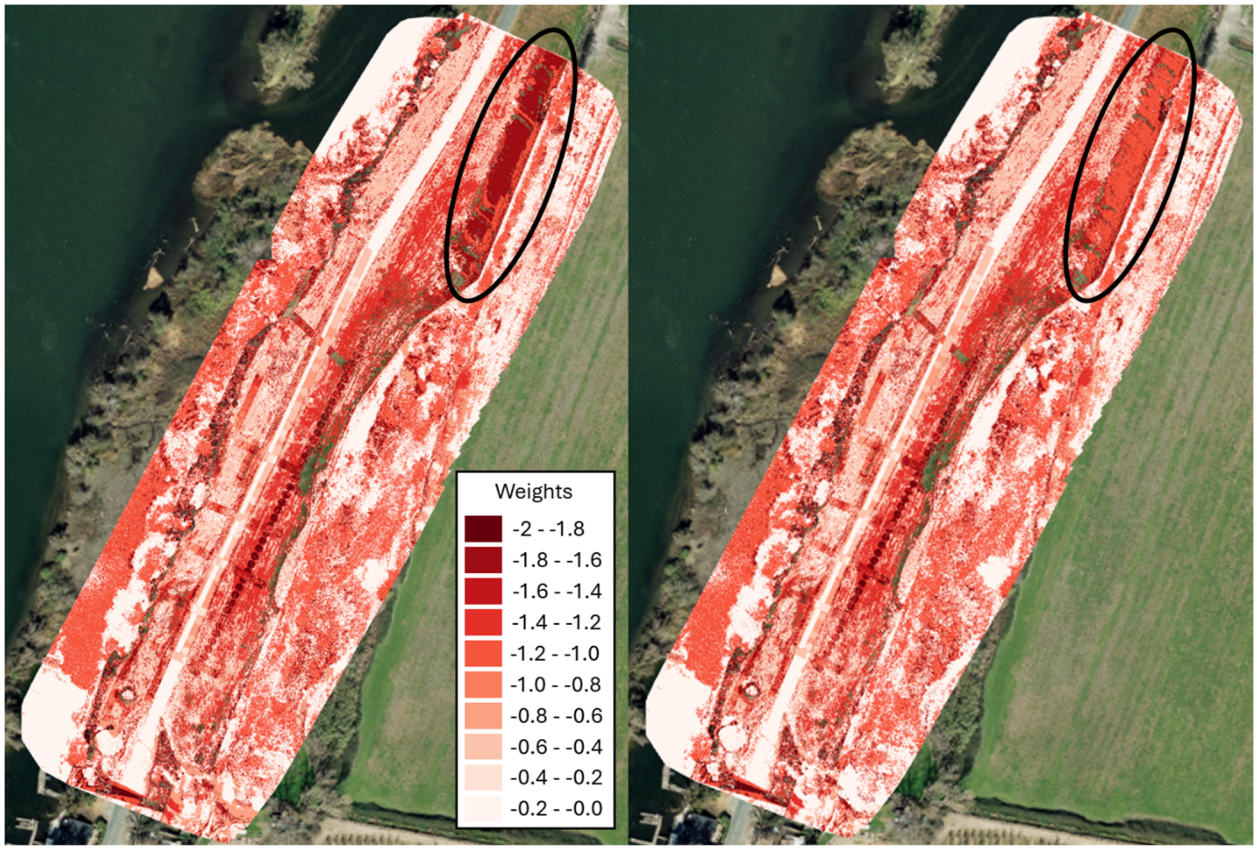


Figure 5.6 Hazard map example with (a) shadows and (b) without shadows for Site 2 on Grand Island.

The Slope layer also used a 1 standard deviation change threshold based on the results from the thermal change detection analysis. Since this layer highlights slope movement best, the pixels within the 1 standard deviation tails were given a weight of -1. All pixels that were not in the 1 standard deviation tails were given a weight of 0.

The final surface layer was ElevationDiff and a threshold of 10% was assumed to be suitable to look for changes in the elevation of the levee structure. For every pixel in the DEM, the percent change was calculated. The pixels that had a percent change above 10% were given a weight of -1 as it could represent toe bulging or crest depressions. All other pixels were assigned a weight of 0 to represent no hazard.

5.4.2 Subsurface Weights

The subsurface layers could not be given a weight solely upon the hazard they represent, unlike the surface layers. Because the subsurface layers are projected to the surface, it was important to include depth information when assigning layer weights.

5.4.2.1 Machine Learning Layers

For the layers that show the potentially problematic zones, each geotechnical condition label associated with factors controlled for in the laboratory were given a base weight. Table 5.2 lists the layers along with their associated base weight. These base weights were applied to each machine learning layer where it was present at depth for each x- and y-coordinate.

Table 5.2 Machine Learning Layers Base Weights

Layer Name	Base Weight
Coarse/Well Graded	1
Fine/Poorly Graded	1
Round/Coarse	1
Organic/Well Graded	0.5
High Vs/Similar ρ	1
Less Organics/Poorly Graded	1
Angular/Fine	1
Low Vs/Similar ρ	1

All layers were given a base weight of 1 except for the Organic/Well Graded layer. This layer was defined by potentially problematic zones that could be attributed to organics or well graded soil. Since well graded soils are less likely to contribute to seepage compared to poorly graded soils and organics can be problematic in terms of seepage but also beneficial in terms of erosion, the layer was given a weight of 0.5.

The High *Vs*/Similar ρ and Low *Vs*/Similar ρ layers were penalized with a weight of 1 since there are no known geotechnical conditions from the laboratory experiment that showed these trends. With no known geotechnical conditions, it is difficult to determine if these zones are hazardous to the system and should be penalized to represent that uncertainty.

All other layers were given a weight of 1 as at least one of the geotechnical conditions in the layer name is associated with the potential for seepage, as discussed in Section 2.7.

To account for the depth information, each point at the surface for each layer was also assigned a profile weight. This profile weight represents the percentage of the profile for each x- and y-coordinate pair where the potentially problematic zone was present. This weight was assigned to each x- and y-coordinate at the surface. The profile weights are shown in Table 5.3.

Table 5.3 Machine Learning Analysis Profile Weights

Percentage of Profile where Zone is Present	Profile Weight
1 – 19	0.05
20 – 39	0.25
40 – 59	0.5
60 – 79	0.75
80 – 100	1

The purpose of the profile weights was to penalize potentially problematic zones that are more prevalent at depth. Very narrow areas in the subsurface should not be penalized as more extensive areas in the subsurface. As a result, potentially problematic zones in the profile that are extensive were penalized more with a higher weight than the thinner areas.

The final weight for the model was a function of the profile weight and the base weight. For each point that was projected to the surface, for each geotechnical condition layer, a profile weight and base weight were assigned. For each point, the final weight can be calculated using Equation 5.1:

$$\text{Weight} = -1(\text{Profile Weight} \times \text{Base Weight}) \quad (5.1)$$

By multiplying the profile weight from Table 5.3 with the base weight from Table 5.2, it accounted for the depth information as well as the hazard type, resulting in weights that can range from 0 to -1. The weight was multiplied by -1 to be consistent with the negative weights assigned to the surface layers. This process was completed in Python before being integrated into the Hazard models in ArcGIS Pro.

5.4.2.2 Change Detection Layers

For the subsurface change detection layers, the weights need to account for the percent of depth where change is, as well as the amount of change that has occurred between the two seasons. The depth information was represented by a profile weight that was attributed to the percentage of the profile for that x- and y-coordinate that had no significant change, as described in Table 5.4.

Table 5.4 Change Detection Profile Weights

Percentage of Profile with No Significant Change	Profile Weight
80 – 100	0.05
60 – 79	0.25
40 – 59	0.5
20 – 39	0.75
0 – 19	1

With these profile weights, areas in the profile with less significant change will not be penalized as harshly as areas in the profile where more significant change occurs. These profile weights are assigned to each projected point at the surface.

With change detection, it was also important to consider the amount of change that occurred. The projected subsurface change detection data contains information about the average percent change in the profile for that x- and y-coordinate. Coordinates with more change should be penalized more than areas with less change. This was done by having weights that represented the amount of change in the subsurface. There are different weights for the projected MASW data compared to the EMI projected data, as shown in Table 5.5.

Table 5.5 Change Detection MASW/EMI Weights

MASW Surveys		EMI Surveys	
Percent Change	**MASW Weight**	**Percent Change**	**EMI Weight**
11 – 29	0.05	10 – 49	0.05
30 – 49	0.25	50 – 99	0.25
50 – 69	0.5	100 – 499	0.5
70 – 89	0.75	500 – 999	0.75
>= 90	1	>= 500	1

Areas that exhibit higher percent changes between the seasons are more heavily penalized than areas with lower percent change. The percent change categories for the EMI data were determined by finding groups of percent change that had roughly the same number of data points in each.

The final weight for the change detection layers was a function of the profile weight from Table 5.4 and the percent change weight from Table 5.5. For each point that was projected to the surface for the MASW and EMI profiles, a profile weight and percent change weight was assigned. For each point, the final weight was calculated using Equation 5.2:

$$\text{Weight} = -1(\text{Profile Weight} \times [\text{MASW Weight or EMI Weight}]) \quad (5.2)$$

With this equation, the profile weight of a point was multiplied by the corresponding MASW weight, if the point is part of an MASW survey, or it was multiplied by the corresponding EMI weight, if the point was part of an EMI survey. The weight was then multiplied by -1 to be consistent with the negative weights assigned to the surface layers. This process of assigning weights to the MASW and EMI change detection profiles was completed using ModelBuilder in ArcGIS Pro and the model layouts can be found in Appendix K.

5.5 Building the Model

The final hazard model is a compilation of the surface and subsurface layers. Part of the goal of the levee health framework is to increase the frequency of inspection. If inspection occurs often, building the model needs to be as automated as possible to support efficient inspection. To create an automated model, ModelBuilder in ArcGIS Pro was leveraged. This model had two main components: change detection analyses and the machine learning analysis.

The model was split into two components to allow for an analysis of how different clustering algorithms and compilation datasets affect the hazard model results. The change detection component of the model was consistent between all the hazard models while the machine learning results changed and affected the results. The change detection model was created separately from the machine learning model to increase the efficiency of the model run time as it only had to be run once. This was important as multiple hazard maps were going to be made to represent the different clustering techniques and compilation datasets and the change detection analysis would not change between these maps. The ModelBuilder change detection model can be found in Appendix K. The flowchart of the model process for calculating the final hazard weights is shown in Figure 5.7.

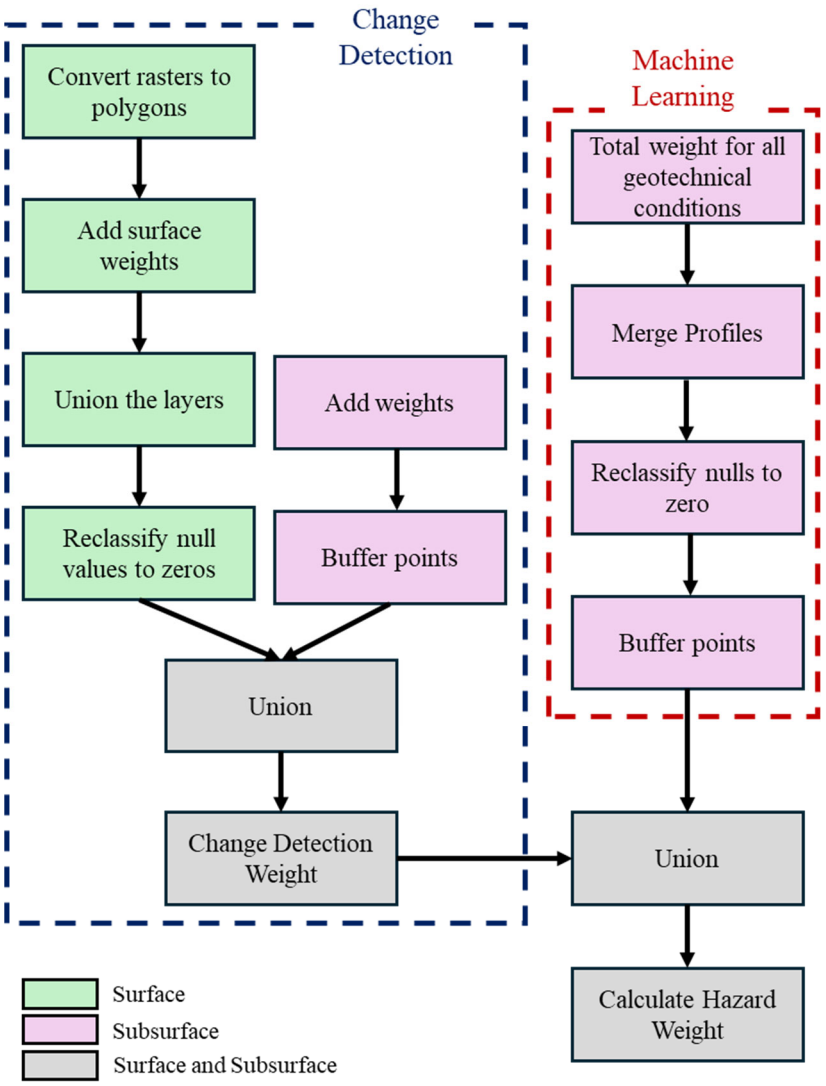


Figure 5.7 Model processing steps for calculating the hazard weight.

The change detection model processes both the surface and subsurface change detection results. For the surface layers, the first step was to convert the raster files to polygons. This was done through ArcGIS Pro's "Raster to Polygon" tool. This was a necessary step in the process as the subsurface data was represented as points and to spatially sum the weights between the layers, the surface and subsurface layers needed to be the same file type.

Once the surface layers were converted to polygons, the next step in building the model was to add the weights to each layer. This was done by creating a new column in the layer's attribute table and populating it with the appropriate weight using if statements. Before the weight was assigned to the thermal change detection layer, the areas that overlapped with shadows were removed since they created artificial hazard locations in the thermal data.

Once the weights were assigned to each surface layer, the polygons making up each layer were combined using the "Union" tool. This tool was the most beneficial for this analysis as it retains

the original geometry of each polygon and creates new polygons where overlap occurred. One of the disadvantages of this method was that polygons that do not overlap between layers create null values in the attribute tables. Null values in the attribute table prevent the layers from spatially summing correctly. To account for this, the model goes into each column in the attribute table and reclassifies the null values to zeros. When that was completed, the weights of the surface layers were totaled to calculate a surface change detection weight.

The change detection model also incorporates the subsurface change detection layers. The subsurface change detection layers are incorporated into the change detection model in parallel with the surface layers. To account for uncertainty in the area directly surrounding each profile, the points that make up the subsurface layers are buffered by 2 m using the "Buffer" tool in ArcGIS Pro. This value was chosen as anything less than this would be drowned out by the surface weights but still high resolution enough to understand problematic areas in the subsurface.

After the subsurface layers were buffered, those results were incorporated into the surface results using the "Union" tool. Now that all the change detection layers are spatially combined, the final change detection weight was calculated. The model was used at Site 1, Site 2, and Site 3 on Grand Island to calculate the change detection weight for each levee segment. The change detection weight for each levee segment will remain constant in all the hazard maps, regardless of the clustering technique or compilation dataset.

To account for different clustering techniques and compilation datasets, a second model was created in ModelBuilder to combine the machine learning layers with the change detection results. The machine learning model integrated with the change detection results can be found in Appendix K.

The first step in building this model was to calculate the total weight for all the geotechnical conditions for a specific area. This was done by summing the weights for each geotechnical condition at the same x- and y-coordinate. This process was completed for the dry and wet profiles.

Once the total weight was calculated for each profile, the next step was to merge the profiles in the wet season into one layer. This was also completed simultaneously for the profiles in the dry season. Merging sites caused null values for points that were not spatially in the same location. As a result, the model reclassified all null values to zero. For both the dry and wet seasons, the projected points were buffered by 2 m to avoid these profiles being overshadowed by the surface layers.

The next step was to union the dry season profiles with the wet season profiles as well as with the change detection results. Once the layers were fused, the final step was to add a new column to the attribute table to calculate the final hazard weight, which was the spatial sum of all the layers. The resulting potential hazardous zones were visualized using a red scale where darker reds indicated more potentially problematic areas while white represented no issues, as shown in Figure 5.8.

Figure 5.8 shows a hazard map of Site 1 on Grand Island created using Hierarchical Clustering by fitting the individual profiles in the anomaly detection analysis. From this map there are areas highlighted in darker reds and areas highlighted in lighter reds. The lighter red areas are located along the water, at the western end of the site, at the eastern end of the site, and along the dry toe. The areas of dark red are located mostly on the dry slope of the levee near the west end of the site.

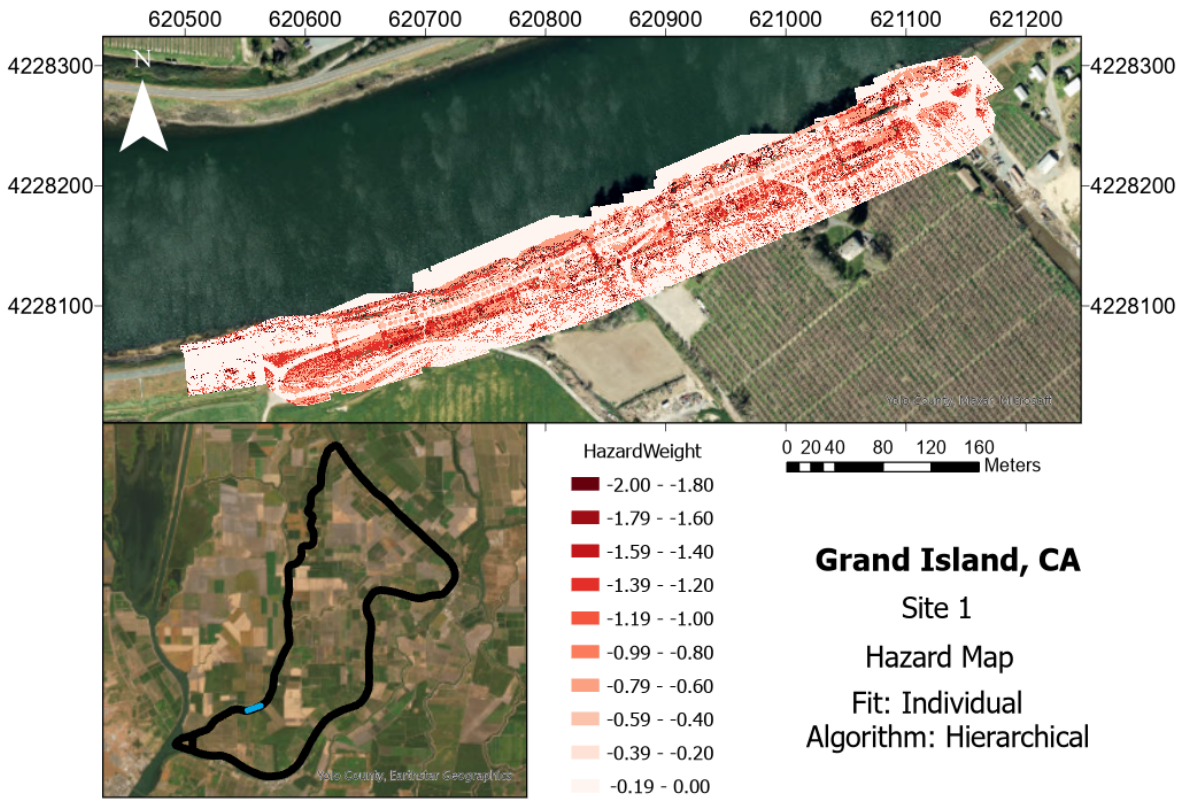


Figure 5.8 Hazard Map of Site 1 on Grand Island showing locations of potential hazards.

The areas highlighted in the dark red represent areas that could be potentially problematic, either at the surface or in the subsurface of the levee. Using this information, stakeholders can make informed decisions about more targeted inspection practices, such as drilling boreholes to better understand the subsurface health or sending a crew to the site to inspect the surface for seepage problems or vegetation encroachment.

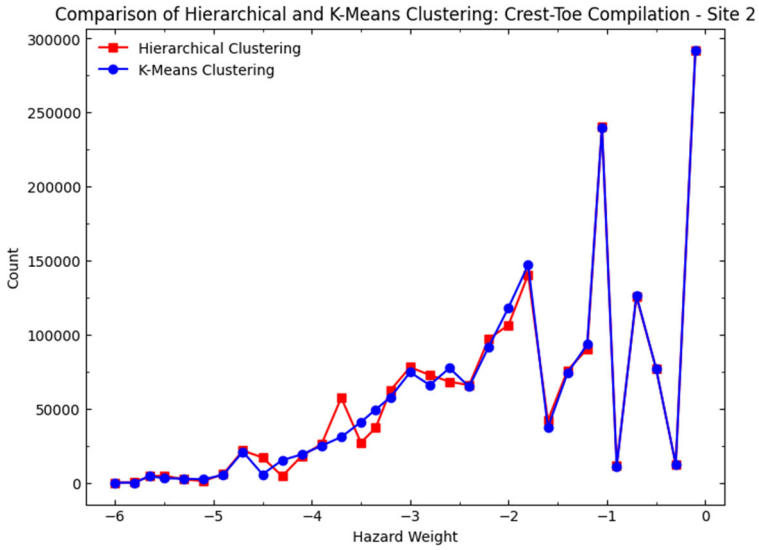
5.6 Comparison of Hazard Maps

Multiple hazard maps were created to see how different clustering techniques and different compilation datasets affected the overall map. This analysis was performed for each site on Grand Island. When creating these maps, the change detection analyses for the surface and subsurface remained constant. The parameter that changed in the creation of the hazard maps was the machine learning analysis layers. For all hazard maps, please refer to Appendix L.

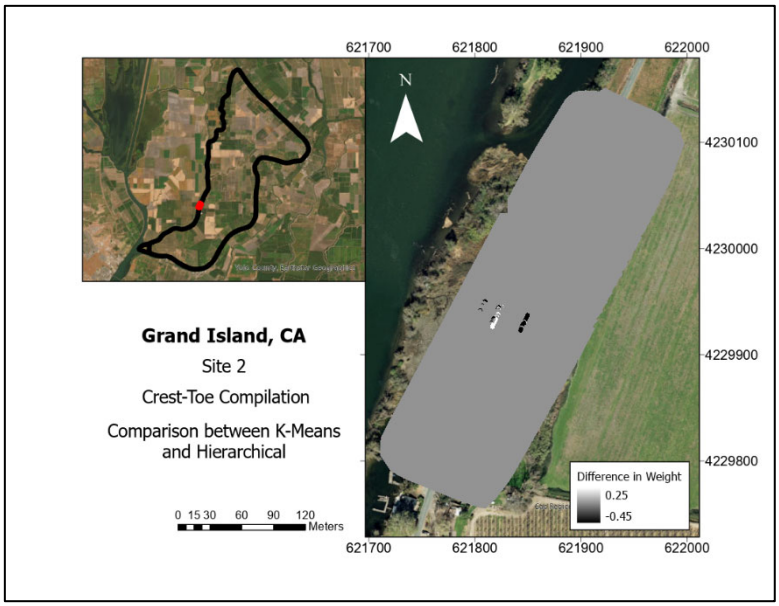
5.6.1 Comparison Between Clustering Techniques

To look at how the clustering techniques affected the hazard maps, the maps with the same compilation dataset were compared. The distribution of the number of occurrences of each hazard weight was plotted for each compilation dataset at each site. For example, the comparison between

K-Means clustering and Hierarchical clustering for the Crest and Toe Compilation for Site 2 is shown in Figure 5.9. For all other distributions, please refer to Appendix L.



(a)



(b)

Figure 5.9 Comparison between the hazard maps for the Crest and Toe Compilation datasets at Site 2 showing (a) the distribution of hazard weights and (b) spatial difference in hazard weights.

In Figure 5.9(a), the distribution of hazard weights for Hierarchical clustering is shown with a red line and square markers while the distribution of hazard weights for K-Means Clustering is shown with a blue line and circle markers. For this compilation dataset, there was good agreement in

hazard weights between K-Means and Hierarchical clustering, especially for weights between -1 and 0 as well as -6 and -5. Most of the variability in the weights occurs between -5.5 and -2, but the variability was not significant.

Overall, there was little variability in the distribution of the hazard weights between the K-Means and Hierarchical clustering maps for Site 1 and Site 2. For each compilation dataset, the distributions have relatively the same shape between K-Means and Hierarchical Clustering with most of the variability occurring between weights of -4 to -2.

For Site 3, there was significantly less agreement in the weight distributions between the K-Means and Hierarchical hazard maps. Some compilation datasets have good agreement in the distributions like the Site Compilation dataset and the Dry and Wet Compilation datasets. Other maps, such as the hazard maps created from fitting and predicting on the individual profiles show very little agreement. For this comparison, the peak of the weight distribution for Hierarchical Clustering occurred around a weight of -2 while the peak for K-Mean clustering was around a weight of -3. This could be due to the high spatial variability in the individual profiles that could cause differences in how the data is clustered between different methods.

The comparison between the K-Means and Hierarchical clustering results on the hazard maps was also analyzed spatially. This was done by comparing the hazard maps from the same compilation dataset but with different clustering techniques, converting the polygon layers to raster layers, and then subtracting the weights at each pixel. This was done for each compilation dataset at all sites on Grand Island. Figure 5.9(b) shows an example of the levee system at Site 2 on Grand Island with these results. Refer to Appendix L for all the maps showing the change in hazard weights spatially on Grand Island.

For the spatial map in Figure 5.9(b), the grey area represents areas where the hazard weight did not change. Most of the profile shows no change in the hazard weights between the two clustering techniques, which was to be expected as the clustering techniques were only applied to the areas where the MASW and EMI surveys overlapped. At Site 2, the surveys were located on the wet toe, dry toe, and crest. From the figure, the crest survey shows that K-Means clustering had higher negative weights compared to Hierarchical clustering while the dry toe survey showed the opposite. Most of the wet toe showed no change in the hazard weight between the two clustering techniques. Clustering on the individual profiles had the least amount of change between the techniques for Site 2 with almost no change in hazard weight at any of the profiles. The Dry and Wet Compilation datasets had no change on the crest and wet toe but higher K-Mean weights on the dry toe with a difference of 0.2 in weight. The Dry and Wet (Site) Compilation datasets also showed very little change between K-Means and Hierarchical clustering with all change occurring on the wet toe but with a difference of 0.25 in weight.

For Site 1, the difference in the hazard weight between K-Means and Hierarchical clustering spatially is dependent on the compilation datasets. Some datasets, such as the individual profiles and Crest and Toe compilation datasets, had very little difference in hazard weights. The other compilation profiles show one survey having little change but the other survey location having upwards of 0.95 change in the hazard weight between K-Means and Hierarchical clustering.

Site 3 at Grand Island was also variable in how the weights compared between K-Means and Hierarchical clustering. Overall, there as little difference in weight, averaging around 0.35

difference. The hazard maps created when clustering on the individual profiles showed the highest change though at a hazard weight difference of 1.45 on the crest and wet toe surveys.

Overall, there was little difference in the hazard weights between K-Means and Hierarchical clustering for all sites and profiles at Grand Island. As a result, using K-Means or Hierarchical clustering in the framework will not significantly affect the health analysis of the levee system.

5.6.2 Comparison Between Compilation Datasets

The analysis of the compilation datasets is important as it shows how combining different profiles for clustering and anomaly detection affects the ability to understand the health of the subsurface. A list of the compilation datasets, as well as which profiles are included, is shown in Table 4.2. To analyze the effect of the compilation datasets on the hazard weights, the distribution of the number of occurrences for each hazard weight for each compilation dataset was plotted. Refer to Appendix L for the figures that show how all the compilation datasets are plotted with respect to one another. Figures 5.10, 5.11, and 5.12, show the mean and variability in the distribution for each weight amongst the compilation datasets for Site 1, Site 2, and Site 3, respectively.

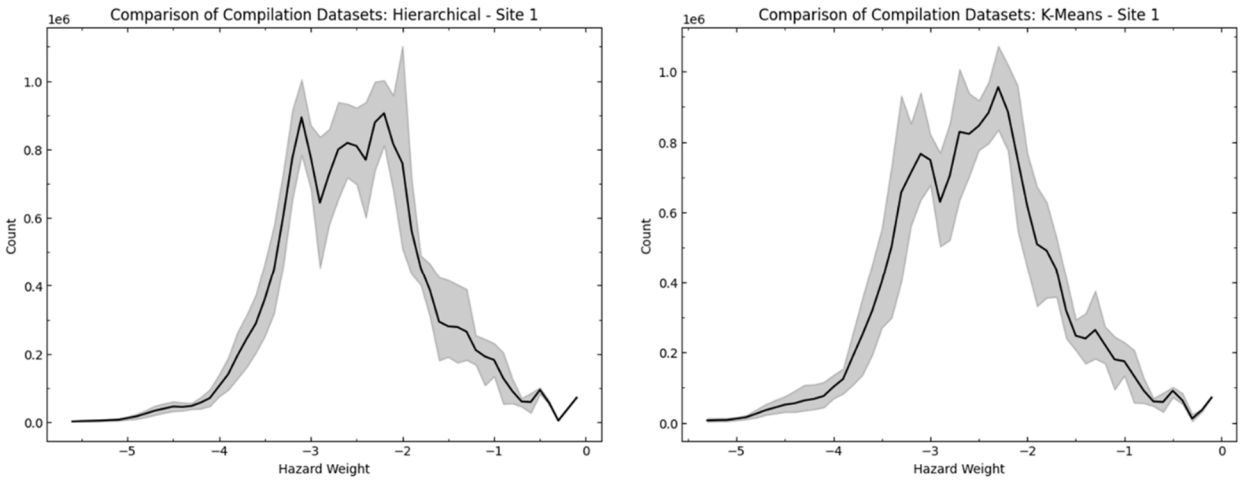


Figure 5.10 Variability in Count for Hazard Weights between Compilation Profiles for Site 1 showing (a) Hierarchical Clustering and (b) K-Means Clustering.

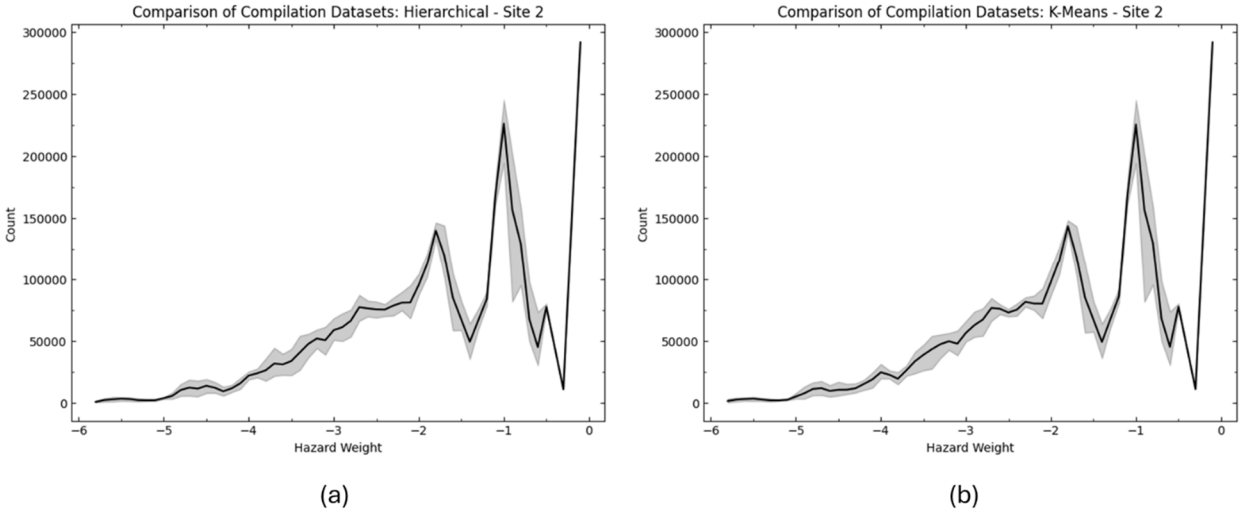


Figure 5.11 Variability in Count for Hazard Weights between Compilation Profiles for Site 1 showing (a) Hierarchical Clustering and (b) K-Means Clustering.

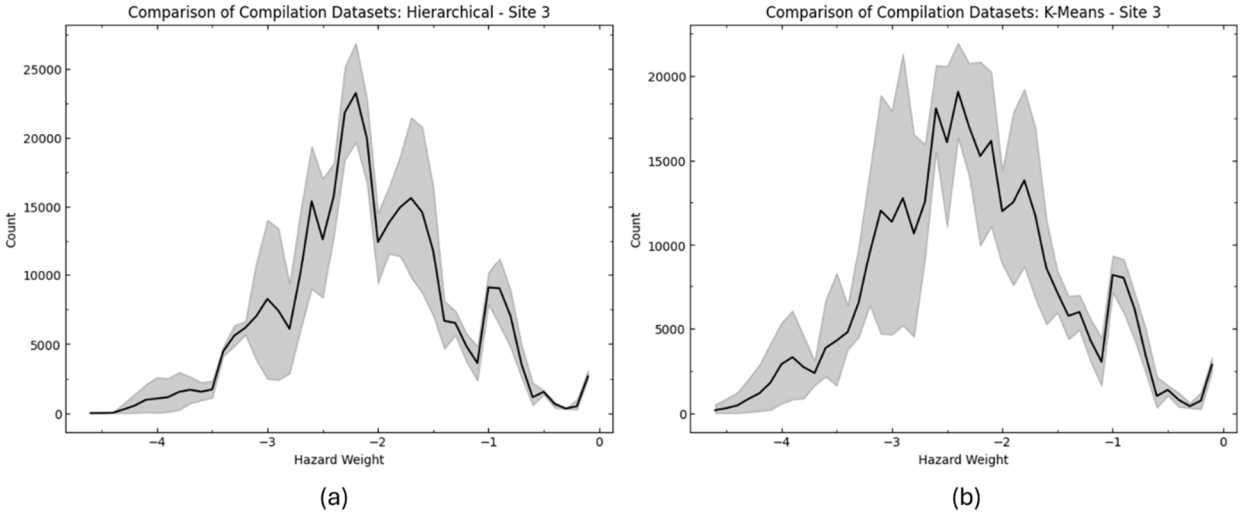


Figure 5.12 Variability in Count for Hazard Weights between Compilation Profiles for Site 1 showing (a) Hierarchical Clustering and (b) K-Means Clustering.

Figure 5.10 shows the distribution of the occurrences of each weight for Site 1. The black line represents the mean occurrence between the different compilation datasets and the grey areas show the variability in occurrence. At Site 1 for Hierarchical and K-Means clustering, all the compilation datasets show the same distribution with most of the levee section having a weight that lies between -2 and -4. Most of the variability in the hazard weights was at the peak of the occurrences between weights of -1 and -4.

Site 2, represented in Figure 5.11, shows good agreement in the hazard weight distribution between the compilation datasets for both K-Means and Hierarchical clustering. There was little variation in the occurrence of each hazard weight with most of the variation around the -0.5 weight and between -2.5 and -4. The compilation datasets show the same distribution where most of the data is found at low hazard weights.

Site 3, shown in Figure 5.12, shows the most variability in the hazard weights between the compilation datasets. For both K-Means and Hierarchical clustering, all of the compilation datasets show the same distribution in hazard weight occurrences with most of the data falling between weights of -1 to -3. At low weights between -1 and 0, there was little variation between compilation datasets. For Hierarchical clustering, the Crest and Toe Compilation dataset was similar to the Dry and Wet Compilation dataset with peaks around a hazard weight of -2.5. The Site Compilation dataset, Dry and Wet (Site) Compilation dataset, and the individual profiles had similar distributions that peaked around a hazard weight of -2. For the K-Means analysis, the Crest and Toe Compilation was similar to the Dry and Wet Compilation as well as the individual profiles with means close to -2.5. The Dry and Wet (Site) Compilation dataset was similar to the Site Compilation dataset, but the mean was shifted towards -2.

To understand how the compilation datasets affected the hazard maps spatially, the weights from each compilation dataset were compared to the weights created when clustering and detecting anomalies on individual profiles. The maps that show these results for all compilation datasets at each site on Grand Island are shown in Appendix L.

For Site 1, the change between each compilation dataset and the individual profiles was typically around a difference in weight of 1 with the compilation dataset (not the individual profiles) having the larger negative weight. For all combinations, both the crest and wet toe profiles had weight changes. For Site 2, the change between the compilation datasets and the individual profiles was around a difference in weight of 1.4 with the compilation dataset having the larger negative weight. For all compilation datasets, the crest, wet toe, and dry toe profiles showed weight changes. For Site 3, the change between the compilation datasets and the individual profiles was a difference in a weight of 1.3 with some showing the compilation dataset having the larger negative weight and others showing the individual profiles having the larger negative weight.

This analysis shows that the compilation datasets do affect the hazard maps. Having different profiles in the clustering and anomaly detection dataset will affect the overall hazard weight along the levee as it will affect the clustering and anomaly detection results (discussed in more detail in Section 4.8. The stakeholder needs to decide if they want to detect subsurface anomalies in small stretches of levees or if they want to detect subsurface anomalies for the entire system. Detecting subsurface anomalies for the entire system will require a compilation dataset and may be more representative of more distinct failure mechanisms and ignores small changes in the subsurface.

5.7 Hazard Model Discussion and Conclusion

Levees are complex earthen structures that can be naturally occurring or man-made. Levees are found in many different conditions, including fluvial, deltaic, and lacustrine environments. Due to this, levee conditions vary widely in terms of height, width, loading conditions, soil composition, and surface conditions. These factors are a result of poor inspection and maintenance practices of

levees. With mostly visual inspection that occurs annually and few instruments to measure the subsurface, levees in the United States have significant deterioration that results in a strong risk for failure.

This research created a robust levee health framework to address the challenges of inspecting and analyzing the surface and subsurface of levees by leveraging a combination of non-invasive sensing techniques. UAV-enabled sensors were used to map the surface conditions using optical properties, surface temperatures, and LIDAR for detailed elevation models. The surface conditions were analyzed using a variety of change detection techniques to understand changes over time that could be attributed to failure mechanisms such as seepage, crest depression, and slope failure.

The subsurface conditions were mapped using geophysical techniques such as MASW to develop the *Vs* distribution of the subsurface and EMI to understand the ρ of the subsurface. The subsurface was analyzed using change detection methods to understand changes over time. This framework also leveraged machine learning to further enhance the understanding of subsurface variations. Unsupervised machine learning clustering techniques were fused with anomaly detection techniques to identify potentially problematic zones that could be attributed to internal failure mechanisms, such as seepage.

To improve the interpretability of these zones, a systematic laboratory investigation was conducted to create a library of *Vs* and ρ values under the same conditions for various geotechnical conditions. The results from the laboratory investigation were used to understand which geotechnical conditions increased, decreased, or did not affect *Vs* and ρ, in tandem. These results were then mapped to the potentially problematic zones from the machine learning analysis to give stakeholders a better understanding of what was happening in the subsurface.

While performing the machine learning analysis, a variety of clustering techniques were used to understand their effect on the analysis. It was determined that there was very little difference in the results and either K-Means or Hierarchical clustering are suitable for this framework.

The datasets included in the clustering and anomaly detection algorithms were also varied to understand how that affects the results. When performing the unsupervised analysis on individual profiles, local anomalies and clusters are detected. These represent variations within the profile of interest and are important for looking for small changes in the profile. Every profile will have anomalies detected. When grouping profiles to fit the unsupervised models, global anomalies and clusters were detected. Because of the larger ranges in *Vs* and ρ, these potentially problematic zones were significantly different than the surrounding areas and could detect problems more easily.

The surface and subsurface conditions were fused to create a hazard map showing the potential hazard at a site. This was done by projecting the subsurface change detection and potentially problematic zone analyses to the surface. Then, the surface layers and subsurface layers were assigned weights that were representative of the hazard associated with them. Once all layers were assigned weights, the weights were spatially summed to create the final hazard map.

These hazard maps integrate surface and subsurface conditions to understand the levee as a system. This framework creates a standardization among organizations for inspection by defining tools for data collection and analysis. Because this framework uses spatially extensive non-invasive sensors, it increases the efficiency of inspection which allows for a higher temporal frequency of inspection.

And since the framework produces a hazard map that shows areas of increased hazard, it justifies locations that require more detailed subsurface investigation and can aid stakeholders in decision-making about maintenance and future rehabilitation.

Chapter 6. Conclusions and Recommendations for Future Research

6.1 Major Contributions of this Research

The major findings and conclusions of this research are described in the following sections for the laboratory investigation on measuring shear wave velocity (*Vs*) and electrical resistivity (ρ), the change detection analysis of the surface and subsurface data, the supervised and unsupervised machine learning analysis of the geophysical data, and the final levee hazard maps.

6.1.1 Development of Library for Shear Wave Velocity and Electrical Resistivity

- A library of *Vs* and ρ data was created under the same conditions for soils with different particle grain sizes, different particle angularities, different gradations, varying organic material, as well as water content, relative density, and vertical stress.
- Soils with rounded particles had higher *Vs* and ρ than soils with angular particles. This relationship with *Vs* is more pronounced at lower water contents. As water content increases, it becomes more difficult to distinguish between particle angularity using *Vs*.
- At higher water contents, soils with coarser particles tended to show similar *Vs* but an increase in ρ compared to finer-grained soil particles. At low water contents though, soils with coarser particles showed higher *Vs* and ρ compared to finer-grained soil particles.
- A soil that was well-graded had similar *Vs* and higher ρ compared to a poorly graded soil that had fine-grained soil particles.
- A soil that was well-graded had higher *Vs* and lower ρ compared to a poorly graded soil that had coarse-grained soil particles.
- Soils that had more organic material exhibited higher *Vs* and lower ρ than the soil with no organic material.

6.1.2 Field Data Collection

- Collecting data in the wet season was challenging due to seasonal changes in vegetation and weather. Levees located in areas with seasons that have a lot of rain can result in heavy vegetation on the slopes of the levees. When the slopes of the levees are densely packed with vegetation, it can be impossible to collect the subsurface data as access to slopes is required.
- When comparing the laboratory measured ρ with the field measured ρ of the same soil samples, there was good agreement in measured ρ between the two measurement techniques. The laboratory measurements of ρ fell within the range of field measurements of ρ, indicating that laboratory measurements were representative of field measurements.

6.1.3 Change Detection Analysis of Surface and Subsurface Data

- The water and urban pixels had the highest rate of being classified as the same land use between the two seasons while the trees, dead grass, and green grass had the lowest rate of being classified as the same land use between the two seasons.
- The dry season had dead vegetation on the slopes of the levees but were highly vegetated on the toes of the levees. During the wet season, the slopes were highly vegetated while the toes had dead vegetation due to the leaves falling off the trees during winter.

6.1.4 Machine Learning Analysis of Geophysical Data of Levee Subsurface

- The clustering in DBSCAN more closely reflected the natural groupings of data compared to K-Means or Hierarchical clustering but, DBSCAN overestimates the anomalous data in the subsurface making it not the optimal algorithm for a resiliency monitoring framework.
- When clustering the subsurface profiles, certain profiles were clustered only on *Vs*, clustered only on ρ, not clustered at all, or were clustered equally based on *Vs* and ρ. This was due to the range of *Vs* and ρ of a profile compared to the range of *Vs* and ρ of the compilation dataset. If the profile had a low range of *Vs* and ρ compared to the dataset, the profile exhibited no clustering. At low ranges of ρ only, the profile was clustered solely on *Vs* while at low ranges of *Vs* only, the profile was clustered solely on ρ. At medium and high ranges of both *Vs* and ρ, the profile was clustered equally on *Vs* and ρ.
- Clustering and anomaly detection on individual profiles show local effects in the results while clustering and anomaly detection on compilation datasets show global effects. Local effects are important when looking for small-scale changes in the dataset while global effects are important when looking for significant changes in the levee.
- The number of clusters used in a profile compared to the number of clusters used for the analysis is dependent on the range of *Vs* and ρ for the profile compared to the compilation dataset. Profiles that used fewer clusters compared to the number of clusters used for the analysis had lower ranges of *Vs* and ρ compared to the compilation dataset. The profiles that had a similar number of clusters as the optimal clusters had higher ranges of *Vs* and ρ.
- The factor affecting the anomaly detection contamination factor was how the data was fitted in the machine learning algorithm. If the algorithm was fitted on the individual profiles, it resulted in higher contamination factors but if the algorithm was fitted on the compilation dataset, it resulted in lower contamination factors. Other factors such as *Vs*, ρ, site location, season, and location on levee had very little effect on the contamination factor.

- The clustering analysis was beneficial in finding the natural groupings in the data based on *Vs* and ρ but it was difficult to determine which cluster should be considered potentially problematic. Anomaly detection could determine potentially problematic zones by grouping anomalies but because it was a binary classification, it was difficult to pull out *Vs* and ρ for one zone. By combining the spatial cluster and anomaly detection profiles, potentially problematic zones in the subsurface are identified along with the *Vs* and ρ in each zone.
- When predicting lithology from CPT data using geophysical data and supervised machine learning analyses, Random Forest performed the best with 94% accuracy with the Northwest Site on Bacon Island performing the best due to the large range in *Vs* and ρ and stiffer soils present at the site.
- Since the Bacon Island dataset contained only three sites, the supervised machine learning analysis was unable to generalize to a new site when predicting the lithology of an entire site due to the lack of training data points and that the models relied heavily on the x-coordinate and elevation.
- When the supervised machine learning analysis was compared with the unsupervised machine learning analysis on Bacon Island, the unsupervised clustering and anomaly detection framework identified areas that were significantly different from the majority of the profile with respect to lithology.

6.1.5 Hazard Map

- The surface layers used in the analysis included the GreenDead as an indicator for seepage, the DeadGreen layer to represent slope access and surface erosion, ThermalDiff as temperature can be used to identify changes in bare soil, near surface cavities, vegetation, and urban structures, Slope to help identify landslides, and ElevationDiff for toe bulging and crest depressions. The surface layers not included in the analysis were NoChange as it shows very little about hazards at the site, NRCT as it only shows vegetation, NormGR as it shows vegetation but without the use of the NIR band, Band1 since it only uses the red band to understand changes in vegetation, Band2 since using a combination of bands represents changes in vegetation better, Band3 as using the blue band shows no significant change for any land use, AddBand as it does not provide any meaningful information about changes in surface features and Aspect because riprap along the toes of levees would have caused anonymously high hazards.
- The distribution of hazard weights between K-Means Clustering and Hierarchical Clustering showed little variability for all sites and profiles on Grand Island. As a result, either clustering algorithm was deemed suitable for this framework as they produce similar results.
- The profiles used to predict anomaly detection and cluster profiles had a significant effect on the hazard weights in the final map. Depending on the profiles used to detect anomalies or to detect clusters, local or global effects in the data will dictate the final hazard weights. Analyses performed using the data that was predicted and fitted on individual profiles showed more local problematic areas which caused an increase in the hazard weight. In comparison, analyses performed using multiple profiles to predict and fit on the profiles resulted in identifying global anomalies along the levee, resulting in lower hazard weights in the maps. If small changes in the subsurface are of interest, the machine learning

analyses should predict and fit on individual subsurface profiles. If large changes that represent more drastic failure mechanisms are of interest, the machine learning analyses should predict on a compilation dataset and fit on each subsurface profile.

- Hazard maps of levee segments were able to show areas of increased hazard using both surface and subsurface data to understand the health of the system. This framework will increase the efficiency of inspection and will help increase the temporal frequency of inspection. It allows for the justification of more detailed investigations and can aid stakeholders in decision-making about maintenance and future rehabilitation.

6.2 Recommendations for Future Research

The framework developed in this study was successful in integrating surface and subsurface levee information for a better understanding of the levee system. Multiple research areas could be explored further that would improve the framework including:

- Collecting surface and subsurface data more frequently. This study analyzed data that were collected during one dry season and the following wet season. Further investigation of collecting data during another season and comparing it to previous data will further develop the framework for increased temporal frequency of data collection/inspection.
- The current framework collects MASW data in finite surveys along sections of levees. Further work should include scaling up the MASW surveys using a Land Streamer to increase the area of the levee surveyed. By scaling up the MASW surveys, 3D models of the *Vs* distribution can be generated. When compared to 3D models of the ρ distribution of the subsurface, a better understanding of subsurface conditions can be analyzed. Having greater spatial distribution of the geophysical data will create better hazard maps as buffering of points will not be needed and there will be no large gaps in subsurface conditions where surface conditions exist.
- This study performed a supervised machine learning analysis on a small subset of levee subsurface data from Bacon Island. Future research should expand this dataset to include sections of levees from other sites under varying conditions. This would help to reduce the effect of class imbalance and create a more robust supervised model that could be applied to novel sections of levees. Another supervised analysis of the subsurface should be conducted where the model is predicting internal failure mechanisms such as animal burrows or seepage.
- This framework utilized unsupervised machine learning techniques for understanding subsurface levee conditions because ground truth data in the form of CPTs and boreholes were not available at the project site. More research should be conducted in areas where CPTs and boreholes are readily available on levees to act as validation for the unsupervised machine learning framework developed in this study.
- A major limitation of this study is the lack of machine learning analysis on the surface data to classify surface failure mechanisms. To perform a study of this nature, an extensive dataset would need to be created consisting of many instances of every single failure mechanism for every possible levee condition. For example, the dataset would need to include instances of crest depressions for levees that are constantly loaded and rarely loaded, as well as for levees that are predominantly clay and levees that are predominantly sandy, as well as for fluvial levees, deltaic levees, lacustrine levees, and all combination of

these factors, and all other failure mechanisms. A dataset of this nature would be difficult to create but would be necessary to classify failure mechanisms of levees at a system levee.

- While performing the surface change detection analysis, a major limitation was the lack of the near-infrared band. Without this band, many change detection techniques were unavailable. Many useful remote sensing indices rely on the near-infrared band to highlight land use features such as NDVI for vegetation or NDWI for water bodies. Future research should include using a multispectral camera to capture the near-infrared wavelengths.
- Further research should be conducted to understand how large and how deep near-surface failure mechanisms need to be to be visible in thermal imagery. In this study, the thermal imagery was dominated by changes in shadows and vegetation. Previous research has used thermal imagery to image badger burrows in levees (Borgatti, et al., 2017) where the burrows were in the range of 2 m by 5 m and near the surface. More research is needed to understand the limitations of using thermal imagery for near-surface failure mechanisms.
- This framework produced hazard maps to aid stakeholders in decision-making regarding inspection and maintenance. Future recommendations include using digital twins and 3D maps of the levees with the ability to toggle on and off different layers, such as the change detection layers as well as the machine learning layers.
- The final recommendation is to integrate social sciences in the hazard assessment of levees. Studies have shown that socially vulnerable communities behind these structures are disproportionally affected by levee failures due to the lack of resources to prepare and recover from flooding events as well as social impediments such as lack of political power and inadequate access to information about the risk of these systems (Vahedifard, Azhar, & Brown, 2023). Since these structures are built to protect communities from flooding, the communities behind these structures should be included in the hazard assessment to help stakeholders make the best informed decision possible.

References

Aanstoos, J. V., Hasan, K., O'Hara, C. G., Prasad, S., Dabbiru, L., Mahrooghy, M., . . . Shrestha, B. (2010). Use of Remote Sensing to Screen Earthen Levees. *2010 IEEE 39th Applied Imagery Pattern Recognition Workshop*, (pp. 1-6). Washington, DC.

AGS Aarhus GeoSoftware. (2024). *Aarhus Workbench*. Retrieved from AGS Aarhus GeoSoftware: https://www.aarhusgeosoftware.dk/aarhus-workbench

Airey, D., & Mohsin, A. K. (2013). Evaulation of Shear Wave Velocity from Bender Elements Using Cross-correlation. *Geotechnical Testing Journal, 36*(4), 506-514.

Alpaydin, E. (2010). *Introduction to Machine Learning: 2nd Edidtion.* Cambridge: The MIT Press.

An, N., Tang, C.-S., Cheng, Q., Wang, D.-Y., & Shi, B. (2020). Laboratory characterization of sandy soil water content during drying process using electrical resistivity/resistance method (ERM). *Bulletin of Engineering Geology and the Environment, 79*, 4411-4427.

Angenheister, G. (1982). Physical properties of rocks. *Landolt-Bornstein, 1b*.

Arulnathan, R., Boulanger, R. W., & Riemer, M. F. (1998). Analysis of Bender Element Tests. *Geotechnical Testing Journal, 21*(2), 120-131.

ASCE. (2021). *Report Card for America's Infrastructure.* ASCE. Retrieved from https://infrastructurereportcard.org/cat-item/levees-infrastructure/

Asch, T. H., Deszcz-Pan, M., Burton, B. L., & Ball, L. B. (2008). *Geophysical characterization of the American River levees, Sacramento, California, using electromagnetics, capacitively coupled resistivity, and DC resistivity.* Virginia: USGS.

ASDSO. (2020). *Lesson Learned: Intervention can stop or minimize consequences of a dam failure. Warning signs should not be ignored.* Retrieved May 21, 2024, from Association of State Dam Safety of Officials: https://damfailures.org/lessons-learned/intervention-can-stop-or-minimize-consequences-of-a-dam-failure-warning-signs-should-not-be-ignored/

Asokan, A., & Anitha, J. (2019). Change detection techniques for remote sensing applications: a survey. *Earth Science Infomatics, 12*, 143-160.

ASTM. (2019). *ASTM D2216-19: Standard Test Methods for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass.* ASTM.

Atwater, B. F. (1982). *Geologic Maps of the Sacramento-San Joaquin Delta, California.* United States Geological Survey.

Bakker, W. H., Feringa, W., Gieske, B. G., Grabmaier, K. A., Hecker, C. A., Horn, J. A., . . . Woldai, T. (2001). *Principles of Remote Sensing.* Enschede: The International Institute for Geo-Information Science and Earth Observation.

Bakula, K., Ostrowski, W., Szender, M., Plutecki, W., Salach, A., & Gorski, K. (2016). Possibilities for using LiDar and photogrammetric data obtained with an unmanned aerial vehicle for levee monitoring. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 41-B1*.

Ban, Y., & Yousif, O. (2016). Chapter 2: Change Detection Techniques: A Review. In B. Yifang, *Multitemporal Remote Sensing: Methods and Applications* (Vol. 20, pp. 19-43). Springer.

Bates, C. R. (1989). Dynamic soil property measurements during triaxial testing. *Geotechnique, 39*(4), 721-726.

Belcher, W., Camp, T., & Krzhizhanovskaya, V. V. (2015). Detecting Erosion Events in Earth Dam and Levee Passive Seismic Data with Clustering. *2015 IEEE 14th International Conference on Machine Learning and Applications* (pp. 903-910). IEEE.

Bersan, S., Koelewijn, A. R., & Simonini, P. (2018). Effectiveness of distributed temperature measurements for early detection of piping in river embankments. *Hydrology and Earth System Sciences, 22*, 1491-1508.

Bhatt, S., & Jain, P. K. (2014). Correlation between electrical resistivity and water content of sand - a statistical approach. *American International Journal of Research in Science, Technology, Engineering & Mathematics, 6*(2), 115-121.

Bishop, M., Dunbar, J., & Peyman-Dove, L. (2003). Integration of remote sensing (LIDAR, electromagnetic conductivity) and geologic data toward the condition assessment of levee systems. *SPIE*, *4886*, pp. 400-407.

Blanchy, G., Saneiyan, S., Boyd, J., & Binley, A. (2020). ResIPy, an intuitive open source software for complex geoelectrical inversion/modeling. *Computers & Geosciences, 137*, 104423.

Bloss, B. R., & Bedrosian, P. A. (2015). Laboratory Electrical Resistivity Analysis of Geologic Samples from Fort Irwin, California. In D. C. Buesch, *Geology and Geophysics Applied to Groundwater Hydrology at Fort Irwin.* USGS.

Borgatti, L., Forte, E., Mocnik, A., Zambrini, R., Cervi, F., Martinucci, D., . . . Zamariolo, A. (2017). Detection and characterization of animal burrows within river embankments by means of coupled remote sensing and geophysical techniques: Lessons from River Panaro (northern Italy). *Engineering Geoloy, 226*, 277-289.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5-32.

Bressan, T. S., de Souza, M. K., Girelli, T. J., & Junior, F. C. (2020). Evaluation of machine learning methods for lithology classification using geophysical data. *Computers and Geoscience, 139*, 104475.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying Density-Based Local Outliers. *ACM SIGMOD International Conference on Management of Data. 29*, pp. 93-104. Dallas: ACM SIGMOD Record.

Brignoli, E. G., Gotti, M., & Stokoe, II, K. H. (1996). Measurement of Shear Waves in Laboratory Specimens by Means of Piezoelectric Transducers. *Geotechnical Testing Journal, 19*(4), 384-397.

Brink, A. B., Partridge, T. C., & Williams, A. B. (1982). *Soil Survey for Engineering.* Clarendon Press.

Building Seismic Safety Council. (2004). *NEHRP RECOMMENDED PROVISISIONS FOR SEISMIC REGULATIONS FOR NEW BUILDINGS AND OTHER STRUCTURES.* Washington D.C.: Building Seismic Safety Council.

Burger, H. R., Sheehan, A. F., & Jones, C. H. (2006). *Introduction to Applied Geophysics.* New York, N.Y., USA: W. W. Norton and Company Inc.

Cabalar, A. F., Demir, S., & Khalaf, M. M. (2021). Liquefaction Resistance of Different Size/Shape Sand-Clay Mixtures using a Pair of Bender Element-Mounted Molds. *Journal of Testing and Evaluation, 49*(1), 509-524.

California Department of Water Resources. (2009). *Delta Risk Management Study: Phase 1.* California Department of Water Resources.

Cardoso, R., & Dias, A. S. (2017). Study of the electrical resistivity of compacted kaolin based on water potential. *Engineering Geology, 226*, 1-11.

Carrivick, J. L., Smith, M. W., & Quincey, D. J. (2016). *Structure From Motion in the Geosciences.* West Sussex: John Wiley & Sons, Ltd.

Carter, J., Schmid, K., Waters, K., Betzhold, L., Hadley, B., Mataosky, R., & Halleran, J. (2012). *Lidar 101: An Introduction to Lidar Technology, Data, and Applications.* Charleston: NOAA.

Cho, G. C., Lee, J.-S., & Santamarina, J. C. (2004). Spatial Variability in Soils: High Resolution Assessment with Electrical Needle Probe. *Journal of Geotechnical and Geoenvironmental Engineering*, 843-850.

Chou, J.-S., & Thedja, J. P. (2016). Metaheuristic optimization within machine learning-based classification system for early warnings related to geotechnical problems. *Automation in Construction, 68*, 65-80.

Choung, Y. (2014). Mapping Levees using LIDAR Data and Multispectral Orthoimages in the Nakdong River Basins, South Korea. *Remote Sensing, 6*, 8696-8717.

Contreras, P., Orellana-Alvear, J., Munoz, P., Bendix, J., & Celleri, R. (2021). Influence of Random Forest Hyperparameterization on Short-Term Runoff Forecasting in an Andean Mountain Catchment. *Atmosphere, 12*, 238.

Dabbiru, L., Aanstoos, J. V., & Younan, N. H. (2016). Earthen levee slide detection via automated analysis of synthetic aperature radar imagery. *Landslides, 13*, 643-652.

Dardanelli, G., La Loggia , G., Perfetti, N., Capodici, F., Puccio, L., & Maltese, A. (2014). Monitoring displacements of an earthen dam using GNSS and remote sensing. *Proceedings of SPIE 9239, Remote Sensing for Agriculture, Ecosystems, and Hydrology XVI* (p. 923928). Amsterdam: SPIE.

De Alba, P., Baldwin, K., Janoo, V., Roe, G., & Celikkol, B. (1984). Elastic-Wave Velocities and Liquefaction Potential. *Geotechnical Testing Journal, 7*(2), 77-87.

de Mello, D. C., Veloso, G. V., de Lana, M. G., de Oliveira Mello, F. A., Poppiel, R. R., Caberro, D. R., . . . Dematte, J. (2022). A new methodological framework for geophysical sensor combinations associated with machine learning algorithms to understand soil attributes. *Geoscientific Model Development, 15*, 1219-1246.

DJI. (2018). *Zenmuse XT2 User Manual.* Retrieved June 2023, from DJI: https://www.dji.com/zenmuse-xt2

Drnevich, V. (1978). *Resonant Column Testing - Problems and Solutions.* ASTM STP 654.

Dudoit, S., & Fridlyand, J. (2002). A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biology, 3*, research00361.

Elsayed, S., El-Hendawy, S., Khadr, M., Elsherbiny, O., Al-Suhaibani, N., Alotaibi, M., . . . Darwish, W. (2021). Combining Thermal and RGB Imaging Indices with Multivariate and Data-Driven Modeling to Estimate the Growth, Water Status, and Yield of Potato under Different Drip Irrigation Regimes. *Remote Sensing, 13*.

El-Sekelly, W., Tessari, A., & Abdoun, T. (2014). Shear Wave Velocity Measurement in the Centrifuge using Bender Elements. *Geotechnical Testing Journal, 37*(4), 689-704.

Ester, M., Kriegel, H. P., Sander, J., & Xiaowei, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise.

Everett, M. E. (2013). *Near-Surface Applied Geophyics.* Cambridge: Cmabridge University Press.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster Analysis.* London, UK: Arnold.

Fisher, W. D., Camp, T. K., & Krzhizhanovskaya, V. V. (2016). Crack Dection in Earth Dam and Levee Passive Seismic Data using Support Vector Machines. *Procedia Computer Science, 80*, 577-586.

Fisher, W. D., Camp, T. K., & Krzhizhanovskaya, V. V. (2017). Anomaly detection in earth dam and levee passive seismic data using support vector machines and automatic feature selection. *Journal of Computational Science, 20*, 143-153.

Fu, L., Liu, G., & Zeng, X. (2009). Evaluation of Shear Wave Velocity Based Soil Liquefaction Resistance Criteria by Centrifuge Tests. *Geotechnical Testing Journal, 32*(1), 39-44.

Galupino, J., & Dungca, J. (2022). Development of a k-Nearest Neighbor (kNN) Machine Learning Model to Estimate the SPT N-Values of Valenzuela City, Philippines. *The 9th AUN/SEED-Net Regional Conference on Natural Disaster* (p. 012021). IOP Publishing.

Gavin, H. P. (2022). *The Levenberg-Marquardt Algorithm for Nonlinear Least Squares Curve-fitting Problems.* Duke University. Retrieved from https://people.duke.edu/~hpgavin/ExperimentalSystems/lm.pdf

Geometrics. (2024). *Seismographs*. Retrieved from Geometrics: https://www.geometrics.com/seismographs/

Geophex Ltd. (2024). *Hand-held Gem-2 Ski*. Retrieved from Geophex Ltd.: https://geophex.com/hand-held-gem-2-ski/

Goh, A. T., & Goh, S. H. (2007). Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data. *Computers and Geotechnics, 34*, 410-421.

Gray, D. H., & Leiser, A. T. (1982). *Biotechnical slope protection and erosion control.* New York: Van Nostrand Reinhold Co.

Gray, D. H., & Megahan, W. F. (1981). *Forest Vegetation Removal and Slope Stability in the Idaho Batholith.* United States Department of Agriculture.

GreenValley International. (2024). *GreenValley International*. Retrieved from GreenValley International: https://www.greenvalleyintl.com/

Guyon, I., & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research, 3*, 1157-1182.

Hahsler, M., Piekenbrock, M., & Doran, D. (2019). dbscan: Fast Density-Based Clustering with R. *Journal of Statistical Software, 91*(1).

Hardin, B. O., & Richart, F. E. (1963). Elastic Wave Velocities in Granular Soils. *Journal of the Soil Mechanics and Foundations Division, 89*(1), 33-65.

Harlan Tait Associates. (1991). *Report on the Preliminary Geotechnical Assessment of the Levees in the Southern Portion of the Lower Sacramento River Area, Phase iV, Sacramento River Flood Control Sytem Evaluation.*

Hayashi, K., & Suzuki, H. (2004). CMP cross-correlation analysis of multi-channel surface-wave data. *Exploration Geophysics, 35*, 7-13.

Hazen, A. (1911). Discussion on Dams on Sand Foundations. *Transactions of the American Society of Civil Engineers, 73*, 199-203.

Hultgren-Tillis Engineers. (2021). *Draft Preliminary Geotechnical Investigation.* Concord: Hultgren-Tillis Engineers.

Hunt, R. E. (2005). Geotechnical Engineering investigation. In *Hand book second edition* (pp. 56-63).

Inazaki, T., & Sakamoto, T. (2005). *Geotechnical Characterization of Levee By Intergrated Geophysical Surveying.*

Inazaki, T., Aoike, K., & Kaneko, M. (2016). Detailed Geophysical Imaging of the Shallow Surfaces at an Underseepage Site Behind a Levee. *SAGEEP*, (pp. 90-95). Denver.

Ingebrisen, S., Ikehara, M. E., Galloway, D. L., & Jones, D. R. (2000). *Sacramento-San Joaquin Delta: The sinking heart of the state.* Sacramento: United States Geological Survey.

Ivanov, J., Miller, R. D., Feigenbaum, D., Morton, S. L., Peterie, S. L., & Dunbar, J. B. (2017). Revisiting Levees in Southern Texas using Love-wave Multichannel Analysis of Surface Waves with the High-resolution Linear Radon Transform. *Interpretation*, T287-T298.

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data.* Englewood Cliffs, NJ: Prentice Hall.

Jovicic, V., Coop, M. R., & Simic, M. (1996). Objective criteria for determining Gmax from bender element tests. *Geotechnique, 46*(2), 357-362.

JSA. (2022). *JIS A 1224:2009- Test Method for Minimum and Maximum Densities of Sands.* Japanese Standards Association.

Kang, M., & Lee, J. S. (2015). Evaluation of the freezing-thawing effects in sand-silt mixtures using elastic waves and electrical resistivity. *Cold Regions Science and Technology, 113*, 1-11.

Kardani, N., Zhou, A., Nazem, M., & Shen, S.-L. (2020). Estimation of Bearing Capacity of Piles in Cohesionless Soil using Optimised Machine Learning Approaches. *Geotechnical and Geological Engineering, 38*, 2271-2291.

Kibria, G., & Hossain, M. S. (2012). Investigation of Geotechnical Parameters Affecting Electrical Resistivity of Compacted Clays. *Journal of Geotechnical and Geoenvironmental Engineering, 138*(12), 1520-1529.

Kim, J., Won, J., & Park, J. (2021). Effects of water saturation and distribution on small-strain stiffness. *Journal of Applied Geophysics, 186*, 104278.

Kozeny, J. (1927). Ueber kapillare Leitung des Wassers im Boden. *Sitzungsber Akad, 136*(2a), 271-306.

Kraus, T., & Sneed, M. (2018, December 6). *Subsidence in the Sacramento-San Joaquin Delta*. Retrieved March 7, 2024, from United States Geological Survey: https://www.usgs.gov/centers/land-subsidence-in-california/science/subsidence-sacramento-san-joaquin-delta

Kuchi, A., Hoque, M. T., Abdelfuerfi, M., & Flanagin, M. C. (2020). Levee-crack detectiom satellite or drone imagery using machine learning aproaches. *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium* (pp. 976-979). Waikoloa, HI: IEEE.

Lay, T., & Wallace, T. (1995). *Modern Global Seismology.* Academic Press.

Lee, C., Suh, H. S., Yoon, B., & Yun, T. S. (2017). Particle shape effect on thermal conductivity and shear wave velocity in sands. *Act Geotechnica, 12*, 615-625.

Lee, J.-S., & Santamarina, J. C. (2005). Bender Elements: Performance and Signal Interpretation. *Journal of Geotechnical and Geoenvironmental Engineering, 131*(9), 1063-1070.

Leong, E. C., Yeo, S. H., & Rahardjo, H. (2005). Measuring Shear Wave Velocity using Bender Elements. *Geotechnical Testing Journal, 28*(5), 488-498.

Lesouple, J., Baudoin, C., Spigai, M., & Tourneret, J.-Y. (2021). Generalized isolation forest for anomaly detection. *Pattern Recognition Letters, 149*, 109-119.

Levenberg, K. (1944). A Method for the Solution of Certain Non-linear Problems in Least Squares. *Quarterly of Applied Mathematics, 2*, 164-168.

Lim, C. S., Mohamad, E. T., Motahari, M. R., Armaghani, D. J., & Saad, R. (2020). Machine Learning Classifiers for Modeling Soil Characteristics by Geophysics Investigations: A Comparative Study. *Applied Sciences, 10*, 5734.

Lin, C.-P., Hung, Y.-C., Wu, P.-L., & Yu, Z.-H. (2014). Performance of 2-D ERT in Investigation of Abnormal Seepage: A Case Study at the Hsin-Shan Earth Dam in Taiwan. *Journal of Environmental and Engineering Geophysics, 19*(2), 101-112.

Liu, F., Ting, K., & Zhou, Z.-H. (2008). Isolation Forest. *Eighth IEEE International Conference on Data Mining* (pp. 413-422). Pisa: IEEE.

Liu, J. G., & Mason, P. J. (2016). *Image Processing and GIS for Remote Sensing: Techniques and Applications.* West Sussex: John Wily & Sons Ltd.

Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory, 28*(2), 129-137.

Lu, D., Mausel, P., Brondizio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*(12), 2365-2401.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Fifth Berkeley Symposium*, (pp. 281-297).

Maltese, A., Pipitone, C., Dardanelli, G., Capodici, F., & Muller, J.-P. (2021). Toward a Comprehensive Dam Monitoring: On-Site and Remote-Retrieved Forcing Factors and Resulting Displacements (GNSS and PS-InSAR). *Remote Sensing, 13*, 1543.

Markham, B. (2013). *LANDSAT 8*. Retrieved June 2023, from Landsat Science: https://landsat.gsfc.nasa.gov/satellites/landsat-8/

Marzan, I., Marti, D., Lobo, A., Alcalde, J., Ruiz, M., Alvarez-Marron, J., & Carbonell, R. (2021). Joint interpretation of geophysical data: Applying machine learning to the modeling of an evaporitic sequence in Villar de Canas (Spain). *Engineering Geology, 288*, 106126.

McManamon, P. (2019). *LiDAR Technologies and Systems.* Bellingham: SPIE Press.

Merritt, A. J., Chambers, J. E., Wilkinson, P. B., West, L. J., Murphy, W., Gunn, D., & Uhlemann, S. (2016). Measurement and modelling of moisture-electrical resistivity relationship of fine-grained unsaturated soils and electrical anisotropy. *Journal of Applied Geophysics, 124*, 155-165.

Mijic, Z., Bray, J. D., Riemer, M. F., Cubrinovski, M., & Rees, S. D. (2021). Test method for minimum and maximum densities of small quantities of soil. *Soils and Foundations, 61*, 533-540.

Miller, G. F., & Pursey, H. (1955). On the partition of energy between elastic waves in a semi-infinite solid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 233*(1192), 55-69.

Mitchell, T. M. (1997). *Machine Learning.* McGraw-Hill.

Mohajer, M., Englmeier, K.-H., & Schmid, V. J. (2010). *A comparison of Gap statistic definitions with and with-out logarithm function.* Munich: University of Munich.

Montoya, B. M., Gerhard, R., DeJong, J. T., Wilson, D. W., Weil, M. H., Martinez, B. C., & Pederson, L. (2012). Fabrication, Operation, and Health Monitoring of Bender Elements for Aggressive Environments. *Geotechnical Testing Journal, 35*(5), 728-742.

Mooney, M. A., Parekh, M. L., Lowry, B., Rittgers, J., Grasmick, J., Koelewijin, A. R., . . . Zhou, W. (2014). Design and Implementation of Geophysical Monitoring and Remote Sensing during a Full Scale Embankment Internal Erosion Test. *Geo-Congress 2014* (pp. 202-211). ASCE.

Mostafa, M., Anwar, M. B., & Radwan, A. (2018). Application of electrical resistivity measurements as quality control test for calcareous soil. *Housing and Building National Research Center Journal, 14*, 379-384.

Nguyen, D. C., Bennett, V., Jones, C., Abdoun, T., Zeghal, M., & O'Meara, K. (2018). New Orleans Levee Monitoring via Remote Sensing and in Situ Instrumentation. *IFCEE 2018* (pp. 516-525). Orlando, FL: ASCE.

NOAA. (2023). *Daily Summaries Station Detailes (Albany 1.7 E, CA US)*. Retrieved 12 07, 2023, from National Centers for Environmental Information: https://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/stations/GHCND:US1CAAL0025/detail

Nobrega, R. A., Anastoos, J., Gokaraju, B., Mahrooghy, M., Dabirru, L., & O'Hara, C. G. (2013). Mapping Weaknesses in the Mississippi River Levee System using Multi-Temporal UAVSAR Data. *Revista Brasileira de Cartografia, 65*(4), 681-694.

Norris, J. E., Greenwood, J. R., Achim, A., Gardiner, B. A., Nicoll, B. C., Cammeraat, E., & Michovski, S. B. (2008). Hazard Assessment of Vegetated Slopes. In J. E. Norris, A. Stokes, S. B. Michovski, E. Cammeraat, R. Beek, B. C. Nicoll, & A. Achim, *Slope Stability and Erosion Control: Ecotechnicalogical Solutions* (pp. 119-166). Springer Dordrecht.

Ogino, T., Kawaguchi, T., Yamashita, S., & Kawajiri, S. (2015). Measurement deviations for shear wave velocity of bender element test using time domain, cross-correlation, and frequency domain approaches. *Soils and Foundations, 55*(2), 329-342.

Osazuwa, I. B., & Chinedu, A. D. (2008). Seismic refraction tomorgraphy imaging of high-permeabilty zones beneath an earthen dam, in Zaria area, Nigeria. *Journal of Applied Geophysics, 66*, 44-58.

Ozer, I. E., Rikkert, S. J., van Leijen, F. J., Jonkman, S. N., & Hanssen, R. F. (2019). Sub-seasonal Levee Deformation Observed using Satellite Radar Interferometry to Enhance Flood Protection. *Scientific Reports, 9*, 2646.

Palacky, G. J. (1988). *3. Resistivity Characteristics of Geologic Targets.* Tulsa, OK: Society of Exploration Geophysicists.

Pandey, L. M., Shukla, S. K., & Habibi, D. (2015). Electrical resistivity of sandy soil. *Geotechnique Letters, 5*, 178-185.

Park, C. B. (2013). MASW for geotechnical site investigation. *The Leading Edge*, 656-662.

Park, C. B., Miller, R. D., & Xia, J. (1998). Ground roll as a tool to image near-surface anomaly. *68th Annual International Meeting Society of Exploration Geophysics, Expanded Abstracts*, (pp. 1377-1380).

Park, C. B., Miller, R. D., & Xia, J. (1999). Multichannel analysis of surface waves. *Geophysics, 64*(3), 800-808.

Park, C. B., Miller, R. D., Xia, J., & Ivanov, J. (2007). Multichannel analysis of surface waves (MASW) - active and passive methods. *The Leading Edge*, 60-64.

Park, J.-H., & Lee, J.-S. (2014). Characteristics of elastic waves in sand-silt mixtures due to freezing. *Cold Regions Science and Technology, 99*, 1-11.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Piezo.com. (2024). *Piezoelectric Actuators and Motors*. Retrieved from Piezo.com: https://piezo.com/collections/piezoelectric-actuators-motors?_=pf&pf_t_quantity=Quantity__1

Pirnia, P., Duhaime, F., & Manashti, J. (2018). Machine learning algorithms for applications in geotechnical engineering. *GeoEdmonton 2018.* Edmonton.

Poomvises, N., Pakdeerod, P., Kongsuk, A., Sintuboon, N., Sawasdipong, B., Khardzir Hussain, M., & Harith, Z. Z. (2020). The Geophysical Risk Matrix: An Innovative Risk Analysis Tool for Earth-Rock Dam Failures Inspection. *ICDSME 2019* (pp. 175-183). Springer.

Prakash, A. (2000). Thermal Remote Sensing: Concepts, Issues, and Applications. *International Archives of Photogrammetry and Remote Sensing, 33*(Part V1), 239-243.

Puri, N., Prasad, H. D., & Jain, A. (2018). Prediction of Geotechnical Parameters using Machine Learning Techniques. *Procedia Computer Science, 125*, 509-517.

Rabby, Y. W., Hossain, M. B., & Abedin, J. (2022). Landslide susceptibility mapping in three Upazilas of Rangamati hill district Bangladesh: application and comparison of GIS-based machine learning methods. *Geocarto International, 37*(12), 3371-3396.

Raber, B. R., & Cannistra, J. (2005). *LIDAR Guidebook: Concepts, Project Design, and Practical Applications.* Park Ridge: Urban and Regional Information Systems Association.

Rahimi, S., Moody, T., Wood, C., Kouchaki, B. M., Barry, M., Tran, K., & King, C. (2019). Mapping Subsurface Conditions and Detecting Seepage Channels for an Embankmnet Dam using Geophysical Methods: a Case Study of the Kinion Lake Dam. *Journal of Environmental and Engineering Geophysics, 24*(3), 373-386.

Raji, W. O., & Adedoyin, A. D. (2020). Dam safety assessment using 2D electrical resistivity geophysical survey and geological mapping. *Jounral of King Saud University - Science, 32*, 1123-1129.

Reading, A. M., Cracknell, M. J., Bombardieri, D. J., & Chalke, T. (2015). Combining Machine Learning and Geophysical Inversion for Applied Geophysics. *ASEG-PESA 2015.* Perth, Australia.

Robertson, P. K. (1990). Soil classification using the cone penetration test. *Canadian Geotechnical Journal, 27*, 151-158.

Robertson, P. K., & Cabal, K. L. (2014). *Guide to Cone Penetration Testing for Geotechnical Engineering.* Signal Hill, CA: Gregg Drilling and Testing, Inc.

Robertson, P. K., Sasitharan, S., Cunning, J. C., & Sego, D. C. (1995). Shear-Wave Velocity to Evaluate In-Situ State of Ottawa Sand. *Journal of Geotechnical Engineering, 121*(3), 262-273.

Rokach, L., & Maimon, O. (2015). *Data Mining with Decision Trees: Theory and Applications, 2nd Edition.* Singapore: World Scientific Publishing Co. Pte. Ltd.

Russo, B. M., & Athanasopoulos-Zekkos, A. (2024). Exploration of Feature Engineering Techniques and Unsupervised Machine Learning Clustering Algorithms for Geophysical Data on Levees. *Geo-Congress 2024* (pp. 454-463). Vancouver: ASCE.

Russo, B. M., Athanasopoulos-Zekkos, A., & Kim, J. (2023). Effect of Grain Size on Granular Soils on Shear Wave Velocity and Electrical Resistivity for Levee Health Monitoring. *Geo-Congress 2023* (pp. 415-423). Los Angeles: ASCE.

Russo, B. M., Athanasopoulos-Zekkos, A., Ulrich, C., & Uhlemann, S. (2024). Levee Soil Type Classification using Shear Wave Velocity and Electrical Resistivity Measurements. *USSD 2024 Annual Conference.* Seattle: USSD.

Salgado, R., Bandini, P., & Karim, A. (2000). Shear Strength and Stiffness of Silty Sand. *Journal of Geotechnical and Geoenvironmental Engineering, 126*(5), 451-462.

Sander, J., Ester, M., Kriegel, H.-P., & Xu, X. (1998). Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and its Applications. *Data Mining and Knowledge Discovery, 2*, 169-194.

Santamarina, J., Klein, K., Wang, Y., & Prencke, E. (2002). Specific surfaceL determination and relevance. *Canadian Geotechnical Journal, 39*, 233-241.

Scholkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the Support of a High-Dimensional Distribution. *Neural Computations, 13*, 1443-1471.

Sekula, K., Borecka, A., Kessler, D., & Majerski, P. (2017). Smart Levee in Poland. Full-Scale Monitoring Experimental Study of Levees by Different Methods. *Computer Science, 14*(4), 357-384.

Seladji, S., Cosenza, P., Tabbagh, A., Ranger, J., & Richard, G. (2010). The effect of compaction on soil electrical resistivity: a laboratory investigation. *European Journal of Soil Science, 61*, 1043-1055.

Shin, H., & Santamarina, J. C. (2013). Role of Particle Angularity on the Mechanical Behavior of Granular Mixtures. *Journal of Geotechnical and Geoenvironmental Engineering*, 353-355.

Shirley, D. J., & Hampton, L. D. (1978). Shear-wave measurements in laboratory sediments. *The Journal of Acoustical Society of America, 63*, 607-613.

Sivasankar, T., Das, R., Borah, S. B., & Raju, P. L. (2019). Insight to the Potentials of Sentintel-1 SAR Data for Embankment Breach Assessment. *Proceedings of International Conference on Remote Sensing for Disaster Management* (pp. 33-41). Springer .

Sneath, P. H. (1957). The Application of Computers to Taxonomy. *The Jounral of General Microbiology, 17*, 201-226.

Sokal, R. R., & Michener, C. D. (1958). A Statistical Methods for Evaluating Relationships. *University of Kansas Bulletin, 38*, 1409-1448.

Son, Y., Oh, M., & Lee, S. (2010). Estimation of soil weathering degree using electrical resistivity. *Environmental Earth Science, 59*, 1319-1326.

Stepler, R. P., Witten, A. J., & Slatt, R. M. (2004). Three-dimensional imaging of a deep marine channel- levee/overbank sandstone behind outcrop with EMI and GPR. *The Leading Edge*, 974-976.

Stillwater Sciences and MBK Engineers. (2020). *Initial Study/Mitigated Negative Declaration Grand Island Levee Erosion Repair Project, Sacramento County, CA.* Berkeley: Stillwater Sciences.

Stokes, A., Norris, J. E., van Beek, L., Bogaard, T., Cammeraat, E., Mickovski, S. B., . . . Fourcaud, T. (2008). How Vegetation Reinforces Soil on Slopes. In J. E. Norris, A. Stokes, S. B. Mickovski, E. Cammeraat, R. Beek, B. C. Nicoll, & A. Achim, *Slope Stability and Erosion Control: Ecotechnological Solutions* (pp. 65-118). Springer Dordrecht.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction 2nd Edition.* Cambridge: The MIT Press.

Tanajewski, D., & Bakula, M. (2016). Application of Ground Penetrating Radar Surveys and GPS Surveys for Monitoring the Condition of Levees and Dykes. *Acta Geophysica, 64*(4), 1093-1111.

Tang, C.-S., Wang, D.-Y., Zhu, C., Zhou, Q.-Y., Xu, S.-K., & Shi, B. (2018). Characterizing drying-induced clayey soil desiccation cracking process using electrical resistivity method. *Applied Clay Science, 152*, 101-112.

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika, 18*, 267-276.

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society. Series B., 63*(2), 411-423.

Tresoldi, G., Arosio, D., Hojat, A., Longoni, L., Papini, M., & Zanzi, L. (2018). Tech-Levee-Watch: experimenting an integrated geophysical system for stability assessment of levees. *Rendiconti Online Societa Geologica Italiana, 46*, 38-43.

Uhlemann, S., Dafflon, B., Wainwright, H. M., Williams, K. H., Minsley, B., Zamudio, K., . . . Hubbard, S. (2022). Surface parameters and bedrock properties covary across a mountainous watershed: Insights from machine learning and geophysics. *Science Advances, 8*.

Uhlemann, S., Ulrich, C., Gniazdowska, A., Craig, M., Nico, P., & Brandenberg, S. (2021). Geophyiscal Data - Next Generation Multi-Hazard Levee Assessment. DesignSafe-CI. Retrieved from https://doi.org/10.17603/ds2-fp5z-1a55

USACE & FEMA. (2024). Retrieved March 6, 2024, from National Levee Database: https://levees.sec.usace.army.mil/#:~:text=The%20NLD%20currently%20includes%20information,the%20NLD%20and%20its%20mission.

USACE. (2019). *Guidelines for Landscape Planning and Vegetation Management at Levees, Floodwalls, Embankment Dams, and Appurtenant Structures.* USACE. Retrieved from US Army Corps of Engineers.

USACE. (2024). *Levees and Public Safety*. Retrieved March 6, 2024, from US Army Corps of Engineers: https://www.nwp.usace.army.mil/Missions/Flood-Risk-Management/Levees/Safety/#:~:text=Routine%20Inspections%2C%20also%20called%20annual,conducted%20on%20an%20annual%20basis.

USSD. (2016). *Monitoring Levees.* Denver: USSD.

Vahedifard, F., Azhar, M., & Brown, D. C. (2023). Overrepresentation of Historically Underserved and Socially Vulnerable Communities Behind Levees in the United States. *Earth's Future, 11*, e2023EF003619.

Viggiani, G., & Atkinson, J. H. (1995). Interpretation of bender element tests. *Geotechnique, 45*(1), 149-154.

Walter, V. (2004). Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing, 58*, 225-238.

Wang, C., Chen, J., Wang, J., & Chen, J. (2016). Flume testing of seepage velocity monitoring using optic fiber distributed temperature sensing for embankments. *Sensor Review, 36*(2), 120-129.

Wang, X., Wang, L., & Zhang, T. (2020). Geometry-based assessment of levee stability and overtopping using airborne LiDar altimetry: A case study in the Pearl River Delta, Southern China. *Water, 12*(403).

Ward Jr., J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association, 58*, 236-244.

Weather Underground. (2023). *Stockton, CA Weather History*. Retrieved 12 11, 2023, from Weather Underground: https://www.wunderground.com/history/monthly/us/ca/stockton/KSCK/date/2023-2

Weller, A., Lewis, R., Canh, T., Moller, M., & Scholz, B. (2014). Geotechnical and Geophysical Long-term Monitoring at a Levee of Red River in Vietnam. *Journal of Environmental and Engineering Geophysics, 19*(3), 183-192.

Wilkie, D. S., & Finn, J. T. (1996). *Remote Sensing Imagery for Natural Resources Monitoring.* Columbia University Press.

Won, J., Park, J., Choo, H., & Burns, S. (2019). Estimation of saturated hydraulic conductivity of coarse-grained soils using particle shape and electrical resistivity. *Journal of Applied Geophysics, 167*, 19-25.

Woolery, E. W. (2018). SH-Mode Seismic Reflection Imaging of Earthen Dams. *Engineering, 4*, 694-701.

Wrona, T., Pan, I., Gawthorpe, R. L., & Fossen, H. (2018). Seismic facies analysis using machine learning. *Geophysics, 83*(5), 83-95.

Xu, R., & Wunsch II, R. C. (2009). *Clustering.* Hoboken: John Wiley & Sons, Inc.

Yamashita, S., Kawaguchi, T., Nakata, Y., Mikami, T., Fujiwara, T., & Shibuya, S. (2009). Interpretation of International Parallel Test on the Measurement of Gmax using Bender Elements. *Soils and Foundations, 49*(4), 631-650.

Yin, Z., Zhou, X., Bakal, C., Li, F., Sun, Y., Perrimon, N., & Wong, S. T. (2008). Using iterative cluster merging with improved gap statistics to perform online phenotype discovery in the context of high-throughput RNAi screens. *BMC Bioinformatics, 9*, 264.

Yu, J., Gan, Z., Zhong, L., & Deng, L. (2018). Research and practice of UAV remote sensing in the monitoring and management of construction projects in riparian areas. *The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3*, (pp. 2161-2165). Beijing.

Yun, T. S., Narsilio, G. A., Santamarina, J. C., & Ruppel, C. (2006). Instrumented pressure testing chamber for charaacterizing sediment cores recovered at in situ hydrostatic pressure. *Marine Geology, 229*, 285-293.

Yunmin, C., Han, K., & Ren-peng, C. (2005). Correlation of shear wave velocity with liquefaction resistance based on laboratory tests. *Soil Dynamics and Earthquake Engineering, 25*, 461-469.

Zakaluk, R., & Sri Ranjan, R. (2008). Predicting the leaf water potential of potato plants using RGB reflectance. *Canadian Biosystems Engineering, 50*, 7.1-7.12.

Zhao, S., Kang, F., Li, J., & Ma, C. (2021). Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction. *Automation in Construction, 130*, 103832.

Zheng, A., & Casari, A. (2018). *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists.* Sebastopol, CA, USA: O'Reilly Media Inc.

Zhu, X., & Goldberg, A. B. (2009). *Introduction to Semi-Supervised Learning.* Morgan & Claypool.

Zumr, D., David, V., Jerabek, J., Noreika, N., & Krasa, J. (2020). Monitoring of the soil moisture regime of an earth-filled dam by means of electrical resistance tomography, close range photogrammetry, and thermal imaging. *Environmental Earth Sciences, 79*(299).

Appendix A – Laboratory Data

Table A.1. Laboratory Data

Soil	Vertical Stress (kPa)	Dr (%)	WC (%)	Vs (m/s)	ρ (Ωm)
Angular	4.949112	30	1	84.99	5865.05
Angular	10.0043	30	1	98.53	1307.65
Angular	25.06229	30	1	114.32	1247.85
Angular	50.01555	30	1	134.6	1179.53
Angular	74.99031	30	1	145.43	1133.43
Angular	99.92205	30	1	152.06	1100.86
Angular	4.949112	50	1	95.21	1131.31
Angular	10.0043	50	1	103.25	1104.02
Angular	25.06229	50	1	118.19	1081.49
Angular	50.01555	50	1	133.68	1064.76
Angular	74.99031	50	1	144.48	1054.05
Angular	99.92205	50	1	151.9	1036.96
Angular	4.949112	70	1	83.9	812.95
Angular	10.0043	70	1	90.48	797.28
Angular	25.06229	70	1	112.44	771.22
Angular	50.01555	70	1	123.91	765.71
Angular	74.99031	70	1	134.72	750.37
Angular	99.92205	70	1	143.68	738.57
Angular	4.949112	30	5	86.24	483.12
Angular	10.0043	30	5	102.56	417.11
Angular	25.06229	30	5	116.85	406.39
Angular	50.01555	30	5	132.93	403.57
Angular	74.99031	30	5	144.05	401.77
Angular	99.92205	30	5	149.15	401.22
Angular	4.949112	50	5	85.68	1200.26
Angular	10.0043	50	5	97.77	698.8
Angular	25.06229	50	5	119.33	540.87
Angular	50.01555	50	5	132.88	495.58
Angular	74.99031	50	5	142.1	476.08
Angular	99.92205	50	5	151.96	428.55
Angular	4.949112	70	5	96.93	262.91
Angular	10.0043	70	5	111.14	392.51
Angular	25.06229	70	5	132.36	257.7
Angular	50.01555	70	5	147.14	257.3
Angular	74.99031	70	5	157.97	257.18
Angular	99.92205	70	5	167.79	256.97
Angular	4.949112	30	10	94.62	164.87
Angular	10.0043	30	10	110.31	159.79
Angular	25.06229	30	10	129.6	156.85
Angular	50.01555	30	10	145	155.07

Angular	74.99031	30	10	153.82	154.43
Angular	99.92205	30	10	164.66	169.15
Angular	4.949112	50	10	95.1	183.28
Angular	10.0043	50	10	113.75	173.84
Angular	25.06229	50	10	128.48	172.63
Angular	50.01555	50	10	142.44	174.45
Angular	74.99031	50	10	153.16	173.93
Angular	99.92205	50	10	162.34	173.19
Angular	4.949112	70	10	85.99	263.34
Angular	10.0043	70	10	96.37	262.91
Angular	25.06229	70	10	110	261.85
Angular	50.01555	70	10	127.64	226.17
Angular	74.99031	70	10	139.65	223.44
Angular	99.92205	70	10	147.86	219.53
Angular	4.949112	30	20	94.9	323.46
Angular	10.0043	30	20	106.29	323.64
Angular	25.06229	30	20	123.12	322.52
Angular	50.01555	30	20	135.83	294.98
Angular	74.99031	30	20	144.21	270.41
Angular	99.92205	30	20	151.52	253.3
Angular	4.949112	50	20	82.23	152.53
Angular	10.0043	50	20	90.31	151.19
Angular	25.06229	50	20	105.08	149.4
Angular	50.01555	50	20	117.12	147.88
Angular	74.99031	50	20	127.19	146.86
Angular	99.92205	50	20	135.09	144.19
Angular	4.949112	70	20	79.53	509.33
Angular	10.0043	70	20	93.15	387.9
Angular	25.06229	70	20	113.27	368.81
Angular	50.01555	70	20	133.88	345.55
Angular	74.99031	70	20	144	342.04
Angular	99.92205	70	20	153.54	333.62
C109	4.949112	30	1	105.28	5138.75
C109	10.0043	30	1	125.16	4645.45
C109	25.06229	30	1	155.19	4377.87
C109	50.01555	30	1	182.17	4223.28
C109	74.99031	30	1	204.04	4117.95
C109	99.92205	30	1	217.74	4023.27
C109	4.949112	50	1	106.16	23923.33
C109	10.0043	50	1	120.38	21918.43
C109	25.06229	50	1	163.01	4102.61
C109	50.01555	50	1	190.96	3828.12
C109	74.99031	50	1	210.99	3681.7

C109	99.92205	50	1	227.43	3516.68
C109	4.949112	70	1	114.23	3489.24
C109	10.0043	70	1	126.1	3268.52
C109	25.06229	70	1	148.74	3233.94
C109	50.01555	70	1	179.21	3138.29
C109	74.99031	70	1	189.7	3074.83
C109	99.92205	70	1	213.69	3047.63
C109	4.949112	30	5	107.92	2245.11
C109	10.0043	30	5	112.53	1761.46
C109	25.06229	30	5	130.37	1593.29
C109	50.01555	30	5	147.18	1505.96
C109	74.99031	30	5	159.35	1432.18
C109	99.92205	30	5	175.72	1405.16
C109	4.949112	50	5	111.74	1213.15
C109	10.0043	50	5	123.14	1191.47
C109	25.06229	50	5	143.72	1174.9
C109	50.01555	50	5	158.96	1147.32
C109	74.99031	50	5	174.37	1110.79
C109	99.92205	50	5	187.47	1109.09
C109	4.949112	70	5	99.05	1275.53
C109	10.0043	70	5	111.99	1275.29
C109	25.06229	70	5	125.69	1273.93
C109	50.01555	70	5	140.22	1261.67
C109	74.99031	70	5	154.25	1255.54
C109	99.92205	70	5	162.7	1238.94
C109	4.949112	30	10	96.84	780.44
C109	10.0043	30	10	107.98	749.51
C109	25.06229	30	10	133.77	731.27
C109	50.01555	30	10	161.11	716.11
C109	74.99031	30	10	169.41	709.82
C109	99.92205	30	10	177.65	690.53
C109	4.949112	50	10	103.77	2330.65
C109	10.0043	50	10	118.13	948.3
C109	25.06229	50	10	138.15	818.8
C109	50.01555	50	10	158.27	756.52
C109	74.99031	50	10	173.48	731.29
C109	99.92205	50	10	180.11	681.48
C109	4.949112	70	10	109.8	1592.85
C109	10.0043	70	10	126	1583.48
C109	25.06229	70	10	147.65	1574.9
C109	50.01555	70	10	166.41	1576.4
C109	74.99031	70	10	178.16	1577.03
C109	99.92205	70	10	185.12	1574.88

C109	4.949112	30	20	62.8	61.8
C109	10.0043	30	20	78	63.76
C109	25.06229	30	20	101.05	67.06
C109	50.01555	30	20	118.87	68.88
C109	74.99031	30	20	130.75	70.12
C109	99.92205	30	20	138.98	71.07
C109	4.949112	50	20	73.52	96.77
C109	10.0043	50	20	87.96	100.73
C109	25.06229	50	20	112.67	103.15
C109	50.01555	50	20	132.35	106.58
C109	74.99031	50	20	143.52	116.45
C109	99.92205	50	20	153.68	115.28
C109	4.949112	70	20	67.08	166.27
C109	10.0043	70	20	83.19	166.24
C109	25.06229	70	20	104.28	170.13
C109	50.01555	70	20	124.13	171.58
C109	74.99031	70	20	133.74	174.99
C109	99.92205	70	20	139.42	178.82
Topatopa-Organics	4.949112	30	1	91.8	12448.51
Topatopa-Organics	10.0043	30	1	110	8393.58
Topatopa-Organics	25.06229	30	1	136.91	6900.07
Topatopa-Organics	50.01555	30	1	168.56	6035.76
Topatopa-Organics	74.99031	30	1	195.65	5385.28
Topatopa-Organics	99.92205	30	1	207.75	5064.92
Topatopa-Organics	4.949112	50	1	96.87	6553.36
Topatopa-Organics	10.0043	50	1	114.03	5835.49
Topatopa-Organics	25.06229	50	1	141.21	5281.58
Topatopa-Organics	50.01555	50	1	167.96	4789.22
Topatopa-Organics	74.99031	50	1	192.98	4480.73
Topatopa-Organics	99.92205	50	1	203.49	4254.36
Topatopa-Organics	4.949112	70	1	106.46	28417.08
Topatopa-Organics	10.0043	70	1	125.73	6481.65
Topatopa-Organics	25.06229	70	1	145.32	5488.07
Topatopa-Organics	50.01555	70	1	156.9	4856.79
Topatopa-Organics	74.99031	70	1	174.51	4511.24
Topatopa-Organics	99.92205	70	1	188.63	4239.59
Topatopa-Organics	4.949112	30	5	94.02	1811.72
Topatopa-Organics	10.0043	30	5	123.88	463.62
Topatopa-Organics	25.06229	30	5	132.75	332.93
Topatopa-Organics	50.01555	30	5	148.83	296.07
Topatopa-Organics	74.99031	30	5	160.05	266.37
Topatopa-Organics	99.92205	30	5	181.33	254.51
Topatopa-Organics	4.949112	50	5	119.89	237.49

Topatopa-Organics	10.0043	50	5	133.5	233.15
Topatopa-Organics	25.06229	50	5	151.37	232.44
Topatopa-Organics	50.01555	50	5	164.12	229.34
Topatopa-Organics	74.99031	50	5	184.48	225.89
Topatopa-Organics	99.92205	50	5	191.76	222.39
Topatopa-Organics	4.949112	70	5	96.99	1165.32
Topatopa-Organics	10.0043	70	5	107.1	642.47
Topatopa-Organics	25.06229	70	5	119.03	346.29
Topatopa-Organics	50.01555	70	5	135.79	298.7
Topatopa-Organics	74.99031	70	5	152.3	269.57
Topatopa-Organics	99.92205	70	5	161.94	250.82
Topatopa-Organics	4.949112	30	10	92.54	122.55
Topatopa-Organics	10.0043	30	10	108.18	86.25
Topatopa-Organics	25.06229	30	10	124.65	84.31
Topatopa-Organics	50.01555	30	10	146.22	79.61
Topatopa-Organics	74.99031	30	10	170.65	74.6
Topatopa-Organics	99.92205	30	10	186.9	71.38
Topatopa-Organics	4.949112	50	10	107.36	555.11
Topatopa-Organics	10.0043	50	10	118.27	103.31
Topatopa-Organics	25.06229	50	10	146.56	87.42
Topatopa-Organics	50.01555	50	10	153.86	83.8
Topatopa-Organics	74.99031	50	10	176.56	80.55
Topatopa-Organics	99.92205	50	10	180.17	77.99
Topatopa-Organics	4.949112	70	10	112.43	65.42
Topatopa-Organics	10.0043	70	10	129.43	64.77
Topatopa-Organics	25.06229	70	10	144.32	64.41
Topatopa-Organics	50.01555	70	10	160.71	63.67
Topatopa-Organics	74.99031	70	10	169.78	62.87
Topatopa-Organics	99.92205	70	10	185.35	61.86
Topatopa-Organics	4.949112	30	20	67	13.43
Topatopa-Organics	10.0043	30	20	77.27	13.5
Topatopa-Organics	25.06229	30	20	89.14	13.56
Topatopa-Organics	50.01555	30	20	99.68	13.64
Topatopa-Organics	74.99031	30	20	108.56	13.6
Topatopa-Organics	99.92205	30	20	116.54	13.39
Topatopa-Organics	4.949112	50	20	68.24	11.64
Topatopa-Organics	10.0043	50	20	75.11	11.66
Topatopa-Organics	25.06229	50	20	83.66	11.73
Topatopa-Organics	50.01555	50	20	93.37	11.83
Topatopa-Organics	74.99031	50	20	101.86	11.87
Topatopa-Organics	99.92205	50	20	107.82	11.91
Topatopa-Organics	4.949112	70	20	62.04	12.85
Topatopa-Organics	10.0043	70	20	67.79	12.87

Topatopa-Organics	25.06229	70	20	78.1	12.86
Topatopa-Organics	50.01555	70	20	89.6	12.88
Topatopa-Organics	74.99031	70	20	98.25	12.87
Topatopa-Organics	99.92205	70	20	106.6	12.87
Topatopa-No Organics	4.949112	30	1	84.37	37235.46
Topatopa-No Organics	10.0043	30	1	100.53	18619.86
Topatopa-No Organics	25.06229	30	1	130.57	14406.58
Topatopa-No Organics	50.01555	30	1	179.32	12975.24
Topatopa-No Organics	74.99031	30	1	188.45	12240.39
Topatopa-No Organics	99.92205	30	1	207.63	11638.72
Topatopa-No Organics	4.949112	50	1	74.32	44590.26
Topatopa-No Organics	10.0043	50	1	94.69	19110.01
Topatopa-No Organics	25.06229	50	1	119.1	14643.08
Topatopa-No Organics	50.01555	50	1	141.47	12769.53
Topatopa-No Organics	74.99031	50	1	163.04	11860.9
Topatopa-No Organics	99.92205	50	1	181.38	11111.79
Topatopa-No Organics	4.949112	70	1	64.49	47838.93
Topatopa-No Organics	10.0043	70	1	82.82	32458.22
Topatopa-No Organics	25.06229	70	1	107.31	10290.43
Topatopa-No Organics	50.01555	70	1	133.45	8949.05
Topatopa-No Organics	74.99031	70	1	144.75	8361.46
Topatopa-No Organics	99.92205	70	1	165.65	7912.66
Topatopa-No Organics	4.949112	30	5	89.89	6312.14
Topatopa-No Organics	10.0043	30	5	104.23	5494.77
Topatopa-No Organics	25.06229	30	5	125.55	289.94
Topatopa-No Organics	50.01555	30	5	141.29	286.46
Topatopa-No Organics	74.99031	30	5	153.19	282.59
Topatopa-No Organics	99.92205	30	5	160.61	280.46
Topatopa-No Organics	4.949112	50	5	92.16	4102.3
Topatopa-No Organics	10.0043	50	5	100.22	520.74
Topatopa-No Organics	25.06229	50	5	118	327.51
Topatopa-No Organics	50.01555	50	5	137.87	279.84
Topatopa-No Organics	74.99031	50	5	155.71	261.16
Topatopa-No Organics	99.92205	50	5	167.77	250.5
Topatopa-No Organics	4.949112	70	5	110.1	212.08
Topatopa-No Organics	10.0043	70	5	121.06	205.45
Topatopa-No Organics	25.06229	70	5	137.97	203.04
Topatopa-No Organics	50.01555	70	5	156.6	198.7
Topatopa-No Organics	74.99031	70	5	165.31	195.77
Topatopa-No Organics	99.92205	70	5	178.03	193.32
Topatopa-No Organics	4.949112	30	10	98.44	68.93
Topatopa-No Organics	10.0043	30	10	110.34	63.23
Topatopa-No Organics	25.06229	30	10	119.81	61.77

Topatopa-No Organics	50.01555	30	10	145.6	61.27
Topatopa-No Organics	74.99031	30	10	154.34	60.72
Topatopa-No Organics	99.92205	30	10	169.68	59.99
Topatopa-No Organics	4.949112	50	10	101.58	7034.26
Topatopa-No Organics	10.0043	50	10	105.64	1729.58
Topatopa-No Organics	25.06229	50	10	122.73	172.04
Topatopa-No Organics	50.01555	50	10	149.91	116.3
Topatopa-No Organics	74.99031	50	10	155.34	102.18
Topatopa-No Organics	99.92205	50	10	169.89	91.07
Topatopa-No Organics	4.949112	70	10	114.57	1470.07
Topatopa-No Organics	10.0043	70	10	123.19	246.62
Topatopa-No Organics	25.06229	70	10	139	160.63
Topatopa-No Organics	50.01555	70	10	143.51	106.87
Topatopa-No Organics	74.99031	70	10	167.32	89.25
Topatopa-No Organics	99.92205	70	10	195.86	78.73
Topatopa-No Organics	4.949112	30	20	79.65	165.06
Topatopa-No Organics	10.0043	30	20	87.96	153.68
Topatopa-No Organics	25.06229	30	20	103.32	101.1
Topatopa-No Organics	50.01555	30	20	116.03	83.22
Topatopa-No Organics	74.99031	30	20	127.45	65.68
Topatopa-No Organics	99.92205	30	20	134.81	97.74
Topatopa-No Organics	4.949112	50	20	68.42	219.96
Topatopa-No Organics	10.0043	50	20	74.45	180.11
Topatopa-No Organics	25.06229	50	20	86.52	165.45
Topatopa-No Organics	50.01555	50	20	98.94	119.83
Topatopa-No Organics	74.99031	50	20	107.99	105.36
Topatopa-No Organics	99.92205	50	20	115.18	94.77
Topatopa-No Organics	4.949112	70	20	73.94	764.59
Topatopa-No Organics	10.0043	70	20	83.82	621.37
Topatopa-No Organics	25.06229	70	20	97.29	582.08
Topatopa-No Organics	50.01555	70	20	111.94	629.08
Topatopa-No Organics	74.99031	70	20	120.61	75.02
Topatopa-No Organics	99.92205	70	20	129.22	40.91
Nevada	4.949112	30	1	92.96	1652.25
Nevada	10.0043	30	1	102.77	1135.74
Nevada	25.06229	30	1	133.87	941.66
Nevada	50.01555	30	1	151.29	859.99
Nevada	74.99031	30	1	157.66	798.26
Nevada	99.92205	30	1	169.2	758.01
Nevada	4.949112	50	1	102.17	663.98
Nevada	10.0043	50	1	104.36	613.52
Nevada	25.06229	50	1	124.12	575.76
Nevada	50.01555	50	1	137.42	546.3

Nevada	74.99031	50	1	166.91	539.07
Nevada	99.92205	50	1	175.96	529.77
Nevada	4.949112	70	1	112.38	529.91
Nevada	10.0043	70	1	120.07	504.33
Nevada	25.06229	70	1	144.66	497.62
Nevada	50.01555	70	1	161.68	489.44
Nevada	74.99031	70	1	168.94	485.38
Nevada	99.92205	70	1	179.1	482.19
Nevada	4.949112	30	5	109.45	3226.54
Nevada	10.0043	30	5	121.75	745.46
Nevada	25.06229	30	5	138.78	412.03
Nevada	50.01555	30	5	156.85	373.35
Nevada	74.99031	30	5	166.44	352.5
Nevada	99.92205	30	5	174.4	337.43
Nevada	4.949112	50	5	121.98	296.15
Nevada	10.0043	50	5	128.91	263.43
Nevada	25.06229	50	5	151.19	259.87
Nevada	50.01555	50	5	179.66	257.58
Nevada	74.99031	50	5	191.21	255.23
Nevada	99.92205	50	5	193.37	252.75
Nevada	4.949112	70	5	115.99	1094.78
Nevada	10.0043	70	5	125.76	289.6
Nevada	25.06229	70	5	140.93	253.76
Nevada	50.01555	70	5	167.84	240.43
Nevada	74.99031	70	5	174.95	231.23
Nevada	99.92205	70	5	181.6	227.52
Nevada	4.949112	30	10	110.17	18059.15
Nevada	10.0043	30	10	116.96	1376.17
Nevada	25.06229	30	10	136.78	286.8
Nevada	50.01555	30	10	171.73	234.71
Nevada	74.99031	30	10	183.29	204.77
Nevada	99.92205	30	10	197.94	193.85
Nevada	4.949112	50	10	123.84	5683.05
Nevada	10.0043	50	10	133.39	356.55
Nevada	25.06229	50	10	149.95	294.02
Nevada	50.01555	50	10	158.45	270.69
Nevada	74.99031	50	10	189.34	244.74
Nevada	99.92205	50	10	193.68	237.71
Nevada	4.949112	70	10	111.71	242.49
Nevada	10.0043	70	10	123.44	235.34
Nevada	25.06229	70	10	148.93	227.98
Nevada	50.01555	70	10	162.87	221.42
Nevada	74.99031	70	10	174.18	217.01

Nevada	99.92205	70	10	181.95	211.35
Nevada	4.949112	30	20	80.2	120.99
Nevada	10.0043	30	20	93.57	120.71
Nevada	25.06229	30	20	112.87	121.76
Nevada	50.01555	30	20	132.51	122.9
Nevada	74.99031	30	20	140.36	123.43
Nevada	99.92205	30	20	152.17	124.16
Nevada	4.949112	50	20	78.48	107.42
Nevada	10.0043	50	20	92.08	107.43
Nevada	25.06229	50	20	117.41	107.24
Nevada	50.01555	50	20	134.38	107.37
Nevada	74.99031	50	20	143.14	107.61
Nevada	99.92205	50	20	151.84	107.88
Nevada	4.949112	70	20	77.21	63.09
Nevada	10.0043	70	20	91.25	135.44
Nevada	25.06229	70	20	116.85	138.07
Nevada	50.01555	70	20	132.46	140.16
Nevada	74.99031	70	20	145.55	142.26
Nevada	99.92205	70	20	148.78	146.55
C109-Nevada	4.949112	30	1	107.1	3487.12
C109-Nevada	10.0043	30	1	116.47	1314.4
C109-Nevada	25.06229	30	1	128.55	1161.65
C109-Nevada	50.01555	30	1	142.66	1094.08
C109-Nevada	74.99031	30	1	158.19	1050.32
C109-Nevada	99.92205	30	1	165.98	1023.71
C109-Nevada	4.949112	50	1	103.58	2132.23
C109-Nevada	10.0043	50	1	116.4	1520.04
C109-Nevada	25.06229	50	1	131.19	1412.13
C109-Nevada	50.01555	50	1	144.23	1360.94
C109-Nevada	74.99031	50	1	152.71	1310.63
C109-Nevada	99.92205	50	1	160.71	1268.84
C109-Nevada	4.949112	70	1	123.09	2292.15
C109-Nevada	10.0043	70	1	142.57	1120.97
C109-Nevada	25.06229	70	1	165.57	1106.59
C109-Nevada	50.01555	70	1	188.35	1078.24
C109-Nevada	74.99031	70	1	202.41	1058.96
C109-Nevada	99.92205	70	1	207.4	1036.4
C109-Nevada	4.949112	30	5	109.25	2305.29
C109-Nevada	10.0043	30	5	123.59	666.64
C109-Nevada	25.06229	30	5	134.6	574.98
C109-Nevada	50.01555	30	5	148.36	514.36
C109-Nevada	74.99031	30	5	164.84	485.22
C109-Nevada	99.92205	30	5	168.93	473.77

C109-Nevada	4.949112	50	5	126.6	2742.15
C109-Nevada	10.0043	50	5	139.61	1407.27
C109-Nevada	25.06229	50	5	168.73	1051.42
C109-Nevada	50.01555	50	5	196.23	871.31
C109-Nevada	74.99031	50	5	202.23	787.43
C109-Nevada	99.92205	50	5	217.19	712.82
C109-Nevada	4.949112	70	5	124.71	753.13
C109-Nevada	10.0043	70	5	137.38	465.3
C109-Nevada	25.06229	70	5	158.86	449.31
C109-Nevada	50.01555	70	5	175.91	433.63
C109-Nevada	74.99031	70	5	185.96	423.22
C109-Nevada	99.92205	70	5	202.29	408.93
C109-Nevada	4.949112	30	10	114.09	173.99
C109-Nevada	10.0043	30	10	130.57	171.84
C109-Nevada	25.06229	30	10	156.32	171.17
C109-Nevada	50.01555	30	10	174.67	170.29
C109-Nevada	74.99031	30	10	197.14	169.84
C109-Nevada	99.92205	30	10	206.99	169.61
C109-Nevada	4.949112	50	10	129.19	474.38
C109-Nevada	10.0043	50	10	152.98	425.57
C109-Nevada	25.06229	50	10	180.17	429.1
C109-Nevada	50.01555	50	10	193.93	420.17
C109-Nevada	74.99031	50	10	201.02	418.86
C109-Nevada	99.92205	50	10	210	416.96
C109-Nevada	4.949112	70	10	122.22	2337.98
C109-Nevada	10.0043	70	10	137.16	448.88
C109-Nevada	25.06229	70	10	159.3	421.35
C109-Nevada	50.01555	70	10	190.36	405.38
C109-Nevada	74.99031	70	10	198.3	386.66
C109-Nevada	99.92205	70	10	214.15	351.49
C109-Nevada	4.949112	30	20	68.28	229.8
C109-Nevada	10.0043	30	20	86.1	232.25
C109-Nevada	25.06229	30	20	116.89	233.54
C109-Nevada	50.01555	30	20	141.33	234.85
C109-Nevada	74.99031	30	20	157.96	236.01
C109-Nevada	99.92205	30	20	165.21	236.65
C109-Nevada	4.949112	50	20	75.98	124.73
C109-Nevada	10.0043	50	20	92.53	127.61
C109-Nevada	25.06229	50	20	119.51	123.21
C109-Nevada	50.01555	50	20	138.33	127.33
C109-Nevada	74.99031	50	20	151.35	124.41
C109-Nevada	99.92205	50	20	162.95	123.79
C109-Nevada	4.949112	70	20	64.47	116.71

C109-Nevada	10.0043	70	20	87.08	114.9
C109-Nevada	25.06229	70	20	111.47	114.19
C109-Nevada	50.01555	70	20	130.66	113.56
C109-Nevada	74.99031	70	20	144.16	111.2
C109-Nevada	99.92205	70	20	156.87	112.13
New Zealand-Well Graded	4.949112	30	1	119.37	2143.25
New Zealand-Well Graded	10.0043	30	1	127.3	1582.6
New Zealand-Well Graded	25.06229	30	1	147.24	1237.78
New Zealand-Well Graded	50.01555	30	1	181.55	904.86
New Zealand-Well Graded	74.99031	30	1	198.54	797.37
New Zealand-Well Graded	99.92205	30	1	213.63	731.86
New Zealand-Well Graded	4.949112	50	1	98.74	1233.8
New Zealand-Well Graded	10.0043	50	1	104.33	600.18
New Zealand-Well Graded	25.06229	50	1	126.17	448.23
New Zealand-Well Graded	50.01555	50	1	146.85	356.41
New Zealand-Well Graded	74.99031	50	1	164.17	305.33
New Zealand-Well Graded	99.92205	50	1	187.72	265.61
New Zealand-Well Graded	4.949112	70	1	95.04	3272.64
New Zealand-Well Graded	10.0043	70	1	98.67	3131.53
New Zealand-Well Graded	25.06229	70	1	127.01	726.98
New Zealand-Well Graded	50.01555	70	1	157.23	439.88
New Zealand-Well Graded	74.99031	70	1	172.12	367.54
New Zealand-Well Graded	99.92205	70	1	190.24	325.67
New Zealand-Well Graded	4.949112	30	5	89.36	265.45
New Zealand-Well Graded	10.0043	30	5	109.22	53.78
New Zealand-Well Graded	25.06229	30	5	121.22	53.48
New Zealand-Well Graded	50.01555	30	5	139.52	46.49
New Zealand-Well Graded	74.99031	30	5	157.45	40.96
New Zealand-Well Graded	99.92205	30	5	170.05	35.86
New Zealand-Well Graded	4.949112	50	5	102.36	192.42
New Zealand-Well Graded	10.0043	50	5	105.59	19.4
New Zealand-Well Graded	25.06229	50	5	129.96	25.75
New Zealand-Well Graded	50.01555	50	5	152.79	25.39
New Zealand-Well Graded	74.99031	50	5	160.71	26.14
New Zealand-Well Graded	99.92205	50	5	168.72	25.67
New Zealand-Well Graded	4.949112	70	5	136.27	40.45
New Zealand-Well Graded	10.0043	70	5	142.67	28.56
New Zealand-Well Graded	25.06229	70	5	144.93	25.68
New Zealand-Well Graded	50.01555	70	5	160.25	21.64
New Zealand-Well Graded	74.99031	70	5	170.16	20.74
New Zealand-Well Graded	99.92205	70	5	185.61	19.51
New Zealand-Well Graded	4.949112	30	10	86.19	0.77
New Zealand-Well Graded	10.0043	30	10	101.84	0.87

New Zealand-Well Graded	25.06229	30	10	120.49	0.97
New Zealand-Well Graded	50.01555	30	10	142.97	1.07
New Zealand-Well Graded	74.99031	30	10	153.95	1.12
New Zealand-Well Graded	99.92205	30	10	167.59	1.12
New Zealand-Well Graded	4.949112	50	10	92.59	2.26
New Zealand-Well Graded	10.0043	50	10	95.13	2.28
New Zealand-Well Graded	25.06229	50	10	111.64	2.29
New Zealand-Well Graded	50.01555	50	10	124.96	2.31
New Zealand-Well Graded	74.99031	50	10	138.94	2.29
New Zealand-Well Graded	99.92205	50	10	141.86	2.26
New Zealand-Well Graded	4.949112	70	10	89.89	2.21
New Zealand-Well Graded	10.0043	70	10	103.35	2.06
New Zealand-Well Graded	25.06229	70	10	120.53	2.03
New Zealand-Well Graded	50.01555	70	10	130.45	2.03
New Zealand-Well Graded	74.99031	70	10	146.9	2
New Zealand-Well Graded	99.92205	70	10	150.81	1.95
New Zealand-Coarse	4.949112	30	1	84.16	3595.02
New Zealand-Coarse	10.0043	30	1	111.15	872.06
New Zealand-Coarse	25.06229	30	1	141.24	506.68
New Zealand-Coarse	50.01555	30	1	173.55	393.29
New Zealand-Coarse	74.99031	30	1	192.74	329.35
New Zealand-Coarse	99.92205	30	1	199.43	289.35
New Zealand-Coarse	10.0043	50	1	116.75	759.58
New Zealand-Coarse	25.06229	50	1	129.55	598.34
New Zealand-Coarse	50.01555	50	1	143.95	482.63
New Zealand-Coarse	74.99031	50	1	161.44	408.8
New Zealand-Coarse	99.92205	50	1	184.18	355.39
New Zealand-Coarse	4.949112	70	1	98.44	721.37
New Zealand-Coarse	10.0043	70	1	117.34	621.85
New Zealand-Coarse	25.06229	70	1	135.81	552
New Zealand-Coarse	50.01555	70	1	154.35	495.57
New Zealand-Coarse	74.99031	70	1	156.23	449.95
New Zealand-Coarse	99.92205	70	1	167.16	405.6
New Zealand-Coarse	4.949112	30	5	111.14	47.82
New Zealand-Coarse	10.0043	30	5	134.1	43.8
New Zealand-Coarse	25.06229	30	5	145.23	42.05
New Zealand-Coarse	50.01555	30	5	167.42	40.46
New Zealand-Coarse	74.99031	30	5	186.53	39.06
New Zealand-Coarse	99.92205	30	5	198.74	37.17
New Zealand-Coarse	4.949112	50	5	102.72	23.07
New Zealand-Coarse	10.0043	50	5	125.8	22.47
New Zealand-Coarse	25.06229	50	5	138.48	22.32
New Zealand-Coarse	50.01555	50	5	152.44	21.97

New Zealand-Coarse	74.99031	50	5	163.76	21.69
New Zealand-Coarse	99.92205	50	5	180.17	21.47
New Zealand-Coarse	4.949112	70	5	107.27	26.51
New Zealand-Coarse	10.0043	70	5	119.74	18.43
New Zealand-Coarse	25.06229	70	5	131.69	21.67
New Zealand-Coarse	50.01555	70	5	160.3	21.33
New Zealand-Coarse	74.99031	70	5	171.84	21.04
New Zealand-Coarse	99.92205	70	5	187.51	20.7
New Zealand-Coarse	4.949112	30	10	84.11	2.68
New Zealand-Coarse	10.0043	30	10	104.09	2.68
New Zealand-Coarse	25.06229	30	10	122.41	2.68
New Zealand-Coarse	50.01555	30	10	138.39	2.59
New Zealand-Coarse	74.99031	30	10	156.26	2.68
New Zealand-Coarse	99.92205	30	10	172.84	2.68
New Zealand-Coarse	4.949112	50	10	73.43	2.15
New Zealand-Coarse	10.0043	50	10	91.75	2.59
New Zealand-Coarse	25.06229	50	10	104.31	2.72
New Zealand-Coarse	50.01555	50	10	127.28	2.86
New Zealand-Coarse	74.99031	50	10	145.5	2.92
New Zealand-Coarse	99.92205	50	10	158.15	2.95
New Zealand-Coarse	4.949112	70	10	78.75	1.59
New Zealand-Coarse	10.0043	70	10	92.64	1.74
New Zealand-Coarse	25.06229	70	10	108.66	1.84
New Zealand-Coarse	50.01555	70	10	121.3	1.89
New Zealand-Coarse	74.99031	70	10	137.48	2
New Zealand-Coarse	99.92205	70	10	150.24	2.1
New Zealand-Fine	4.949112	30	1	109.9	521.03
New Zealand-Fine	10.0043	30	1	123.54	504.38
New Zealand-Fine	25.06229	30	1	140.31	464.71
New Zealand-Fine	50.01555	30	1	174.12	408.34
New Zealand-Fine	74.99031	30	1	187.2	369.02
New Zealand-Fine	99.92205	30	1	208.13	344.1
New Zealand-Fine	4.949112	50	1	102.66	497.66
New Zealand-Fine	10.0043	50	1	115.38	450.99
New Zealand-Fine	25.06229	50	1	139.69	411.93
New Zealand-Fine	50.01555	50	1	159.8	355.19
New Zealand-Fine	74.99031	50	1	171.09	327.37
New Zealand-Fine	99.92205	50	1	199.04	294.15
New Zealand-Fine	4.949112	70	1	96.93	424.87
New Zealand-Fine	10.0043	70	1	106.36	398.5
New Zealand-Fine	25.06229	70	1	115.63	379.02
New Zealand-Fine	50.01555	70	1	139.65	351.47
New Zealand-Fine	74.99031	70	1	155.39	330.42

New Zealand-Fine	99.92205	70	1	176.03	310.68
New Zealand-Fine	4.949112	30	5	106.06	581.05
New Zealand-Fine	10.0043	30	5	118.41	10.3
New Zealand-Fine	25.06229	30	5	126.43	9.33
New Zealand-Fine	50.01555	30	5	144.98	11.19
New Zealand-Fine	74.99031	30	5	153.95	12.74
New Zealand-Fine	99.92205	30	5	162.44	12.61
New Zealand-Fine	4.949112	50	5	109.73	8.22
New Zealand-Fine	10.0043	50	5	125.43	10.43
New Zealand-Fine	25.06229	50	5	131.74	10.61
New Zealand-Fine	50.01555	50	5	149.15	10.4
New Zealand-Fine	74.99031	50	5	160.77	10.34
New Zealand-Fine	99.92205	50	5	173.37	9.96
New Zealand-Fine	4.949112	70	5	131.35	42.33
New Zealand-Fine	10.0043	70	5	156.78	16.08
New Zealand-Fine	25.06229	70	5	160.2	9.52
New Zealand-Fine	50.01555	70	5	162.91	10.14
New Zealand-Fine	74.99031	70	5	169.36	10.28
New Zealand-Fine	99.92205	70	5	180.34	10.62
New Zealand-Fine	4.949112	30	10	98.4	2.2
New Zealand-Fine	10.0043	30	10	108.18	2.18
New Zealand-Fine	25.06229	30	10	131.53	2.7
New Zealand-Fine	50.01555	30	10	136.45	3.12
New Zealand-Fine	74.99031	30	10	148	3.22
New Zealand-Fine	99.92205	30	10	171.24	3.2
New Zealand-Fine	4.949112	50	10	102.46	322.85
New Zealand-Fine	10.0043	50	10	107.38	15.02
New Zealand-Fine	25.06229	50	10	120.87	1.41
New Zealand-Fine	50.01555	50	10	138.91	1.63
New Zealand-Fine	74.99031	50	10	151.05	1.78
New Zealand-Fine	99.92205	50	10	166.16	1.88
New Zealand-Fine	4.949112	70	10	95.12	1.92
New Zealand-Fine	10.0043	70	10	107.65	1.14
New Zealand-Fine	25.06229	70	10	115.87	1.11
New Zealand-Fine	50.01555	70	10	133.78	1.12
New Zealand-Fine	74.99031	70	10	142.93	1.2
New Zealand-Fine	99.92205	70	10	157.16	1.25

Comparison Figures

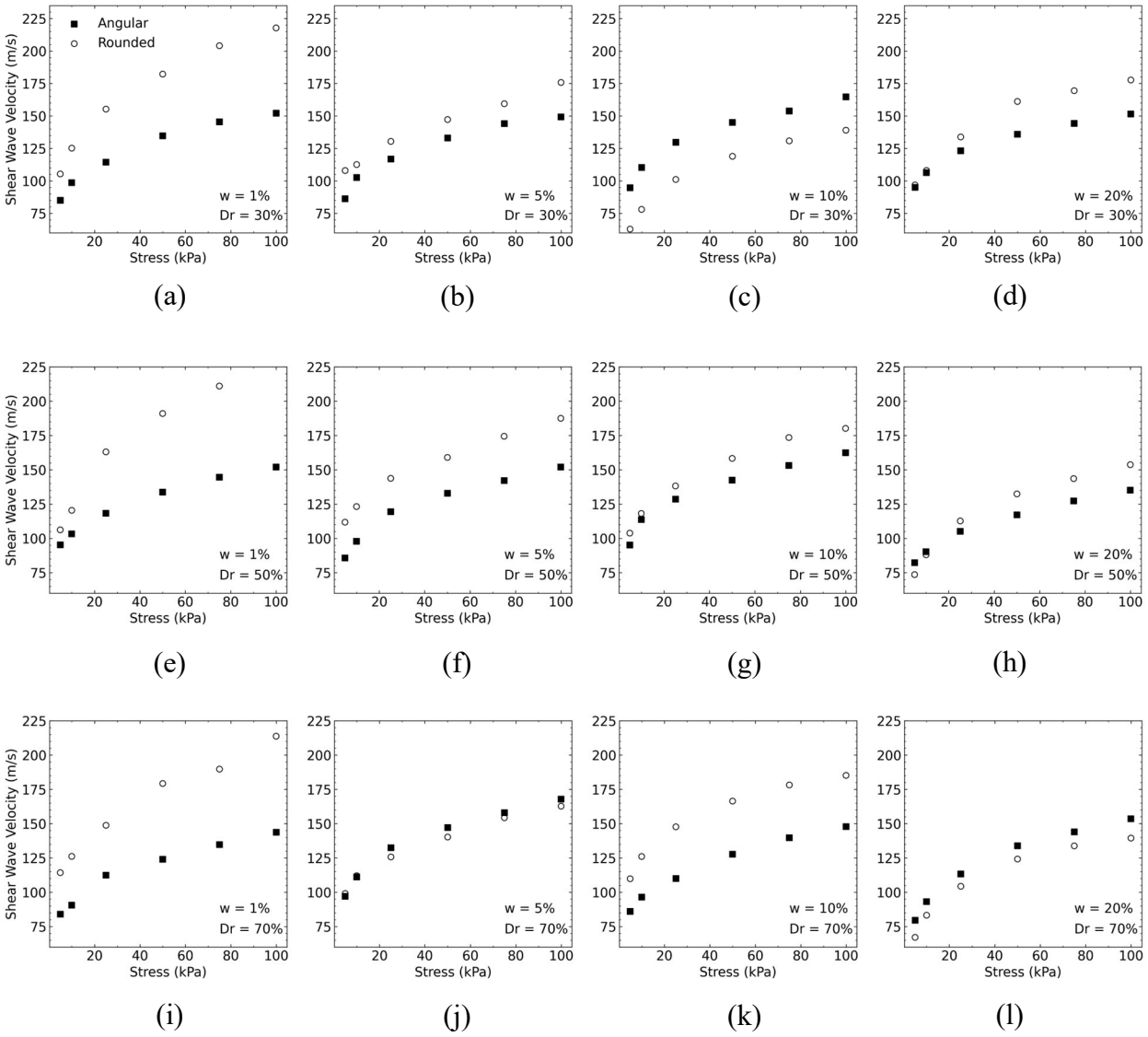


Figure A.1. Particle angularity comparison for shear wave velocity between the crushed angular soil and C109 sand for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

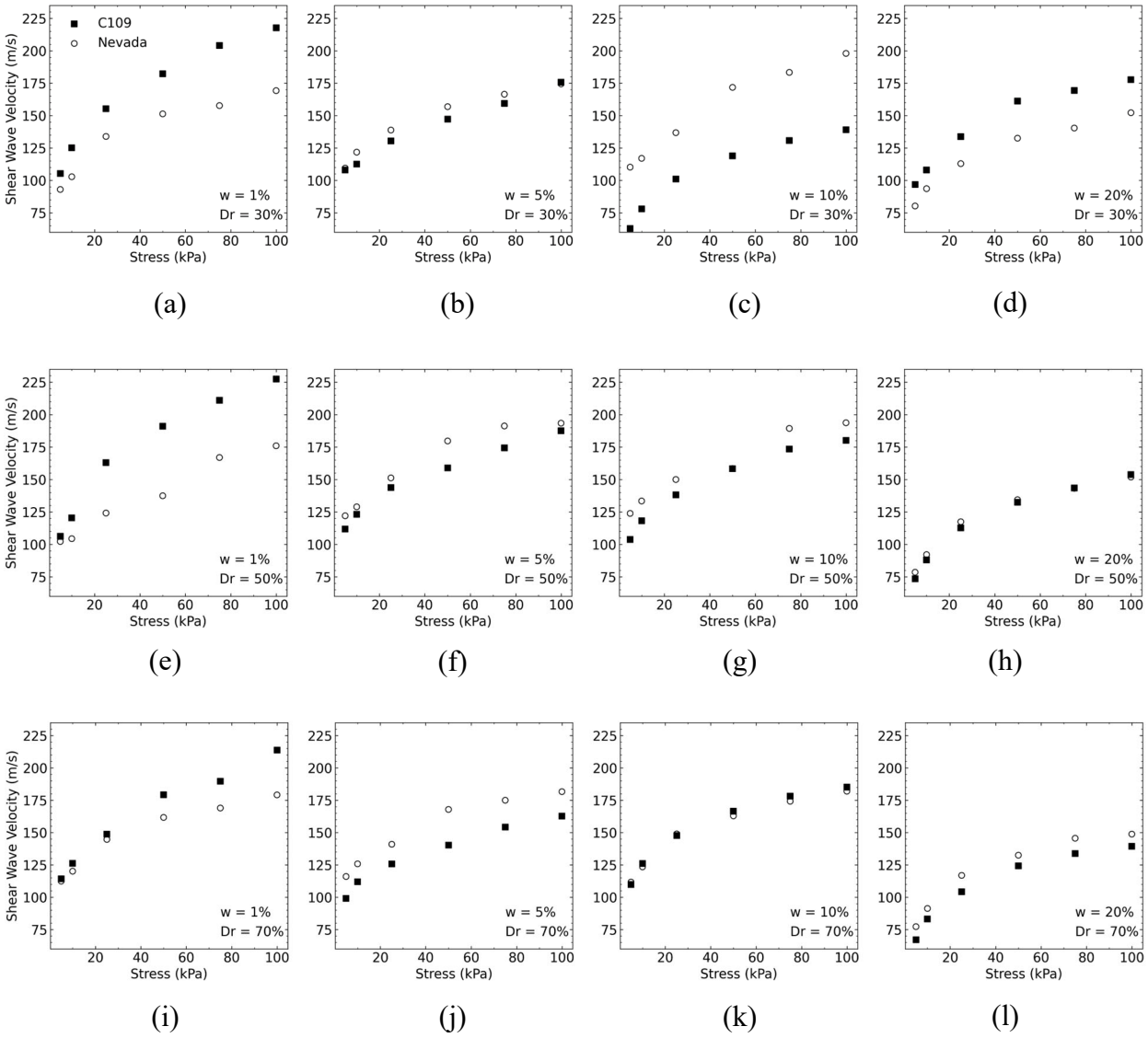


Figure A.2. Particle grain size comparison for shear wave velocity between C109 sand and Nevada sand for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

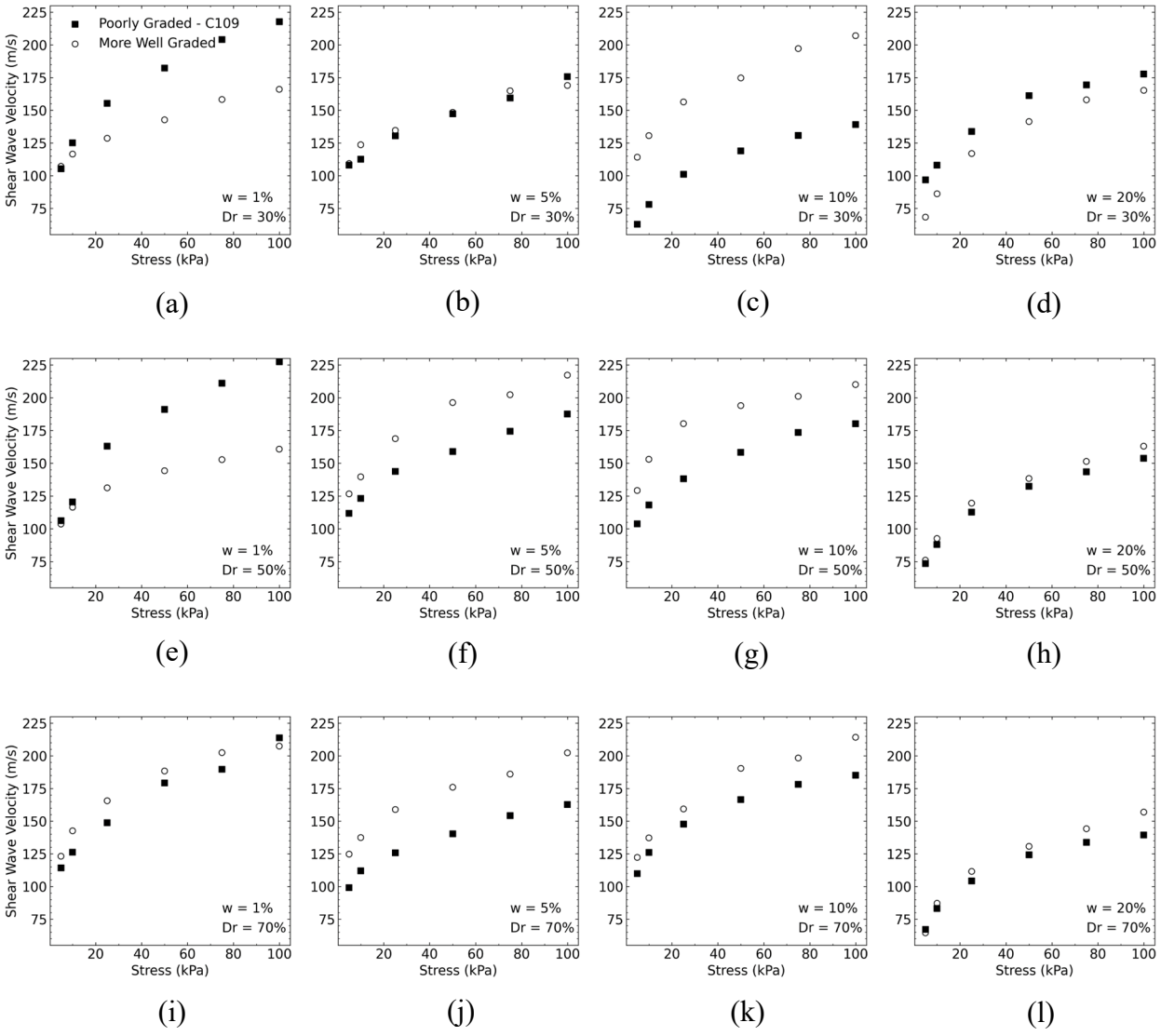


Figure A.3. Soil gradation comparison for shear wave velocity between C109 sand and the C109-Nevada mixture for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

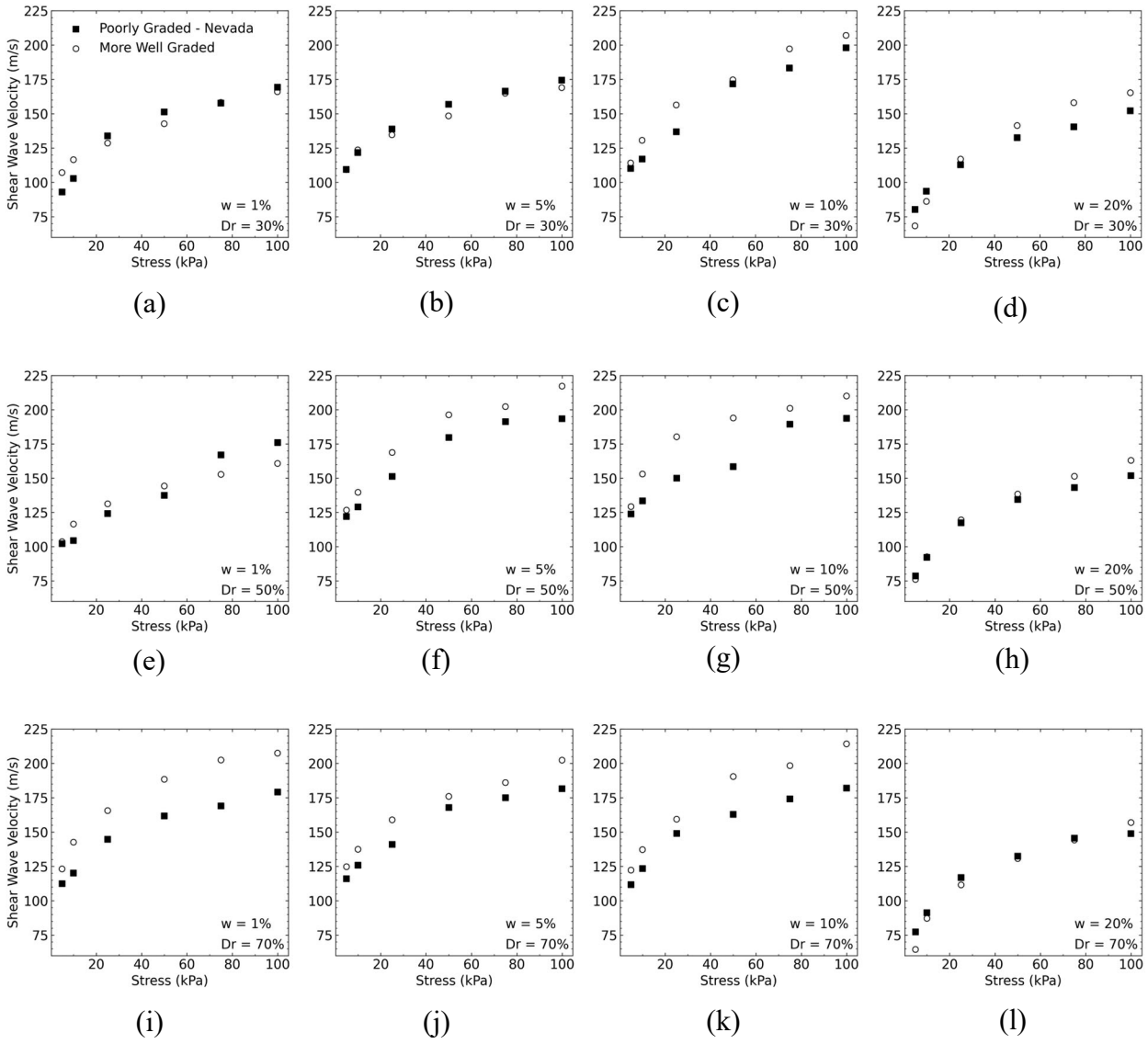


Figure A.4. Soil gradation comparison for shear wave velocity between Nevada sand and the C109-Nevada mixture for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

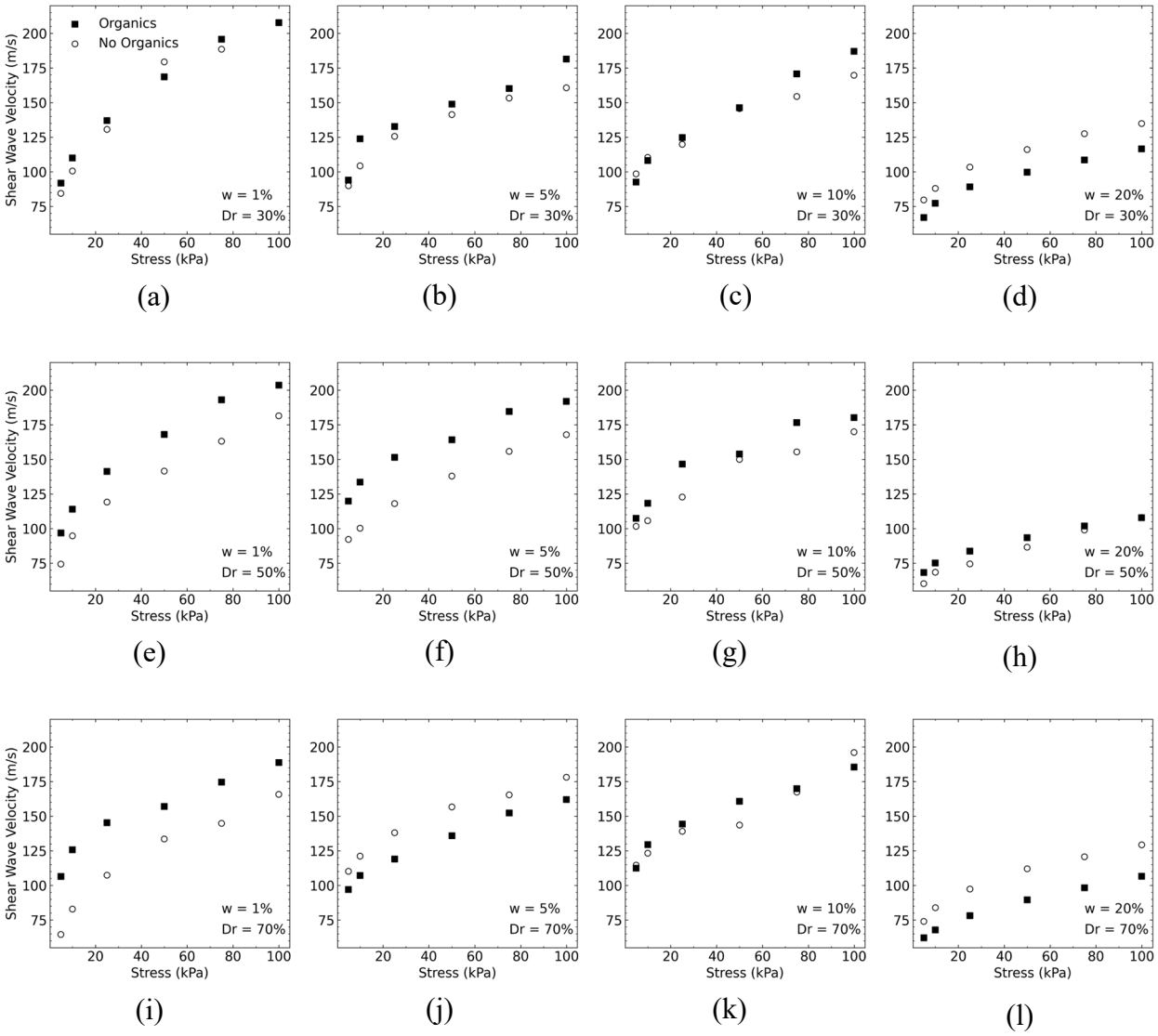


Figure A.5. Organic matter comparison for shear wave velocity between the Topatopa soils for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

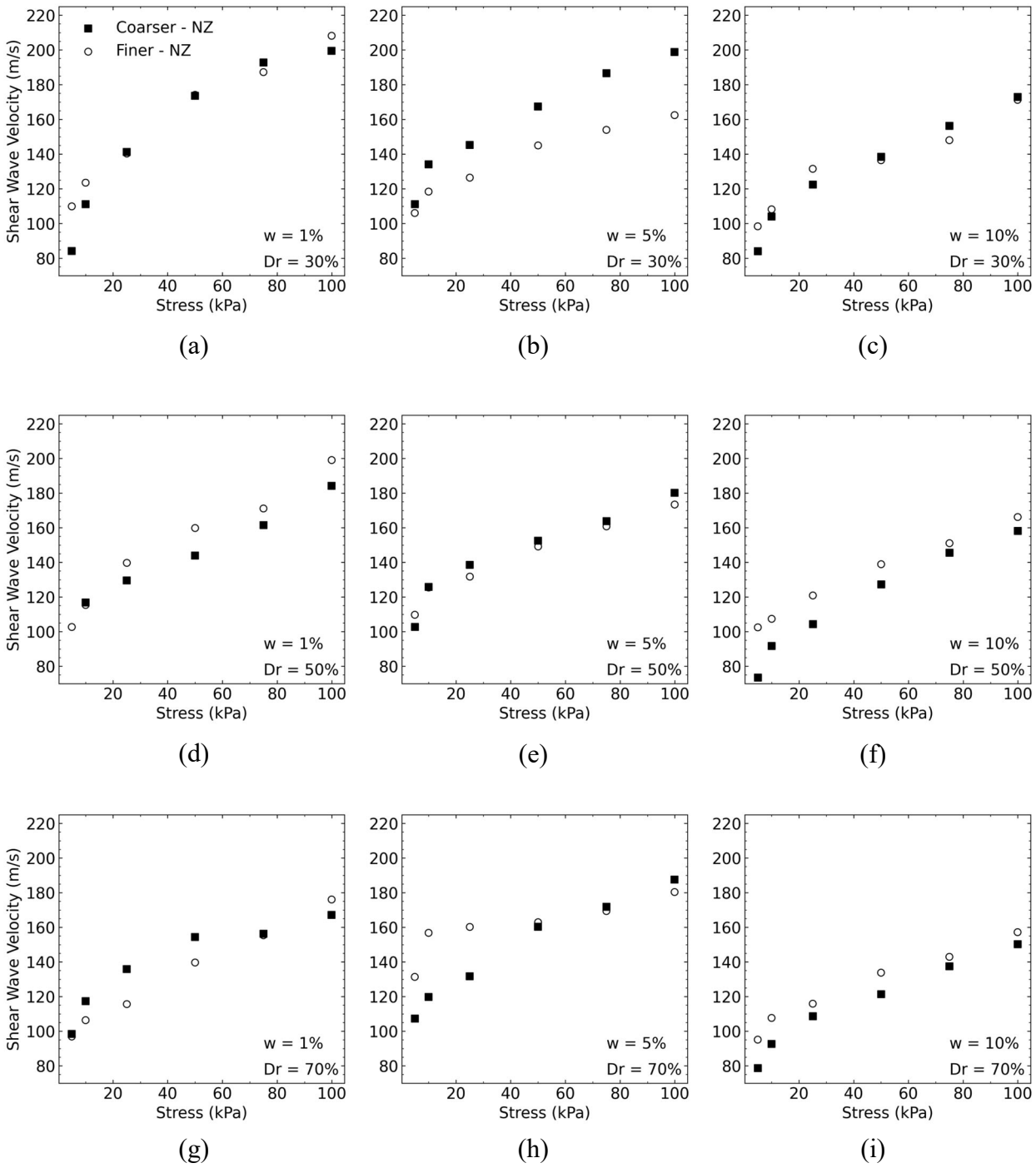


Figure A.6. Particle grain size comparison for shear wave velocity between the coarse and fine New Zealand soils (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 1% and Dr = 50%, (e) w = 5% and Dr = 50%, (f) w = 10% and Dr = 50%, (g) w = 1% and Dr = 70%, (h) w = 5% and Dr = 70%, and (i) w = 10% and Dr = 70%

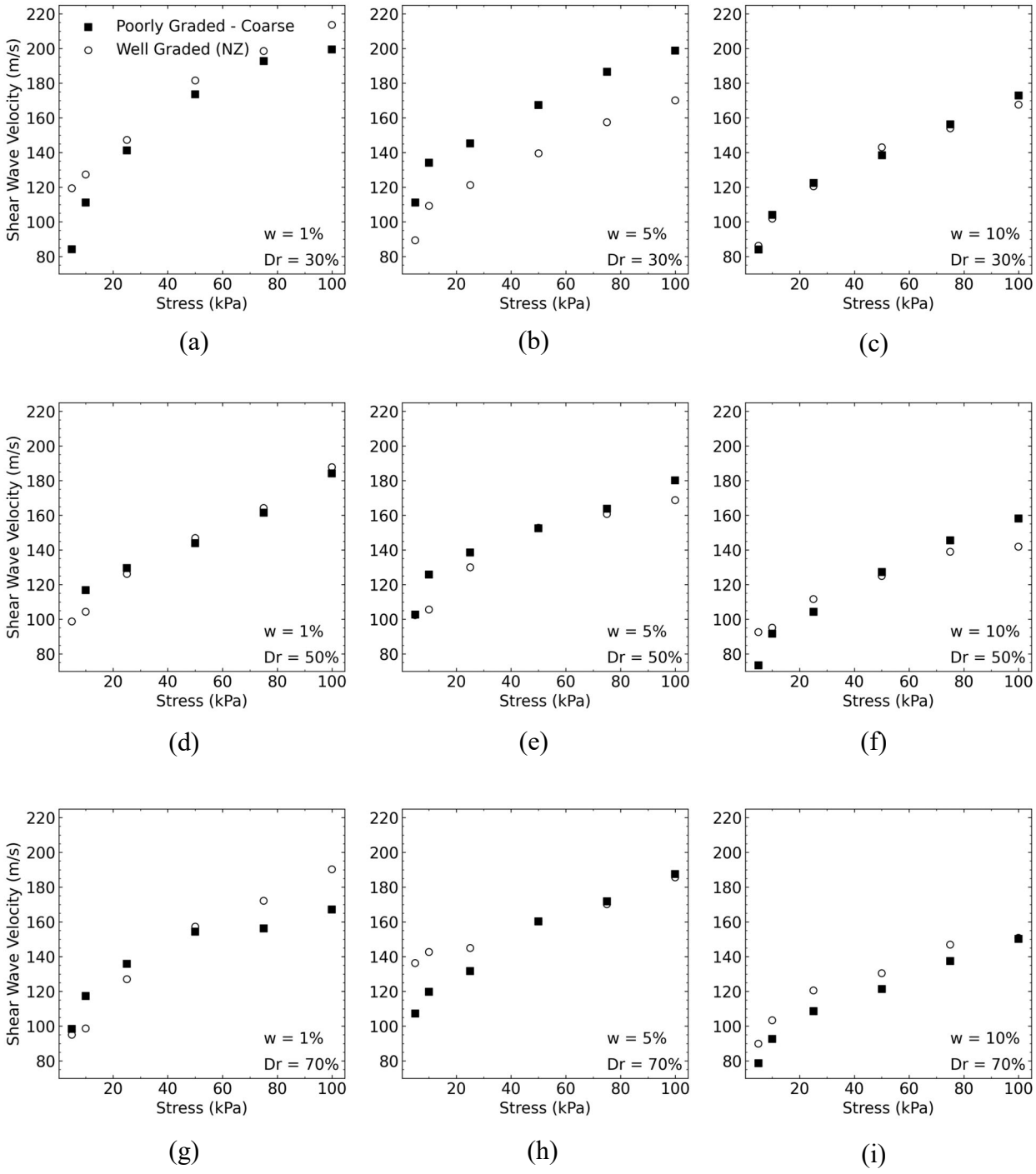


Figure A.7. Soil gradation comparison for shear wave velocity for the New Zealand soils using the coarse poorly graded soil for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 1% and Dr = 50%, (e) w = 5% and Dr = 50%, (f) w = 10% and Dr = 50%, (g) w = 1% and Dr = 70%, (h) w = 5% and Dr = 70%, and (i) w = 10% and Dr = 70%

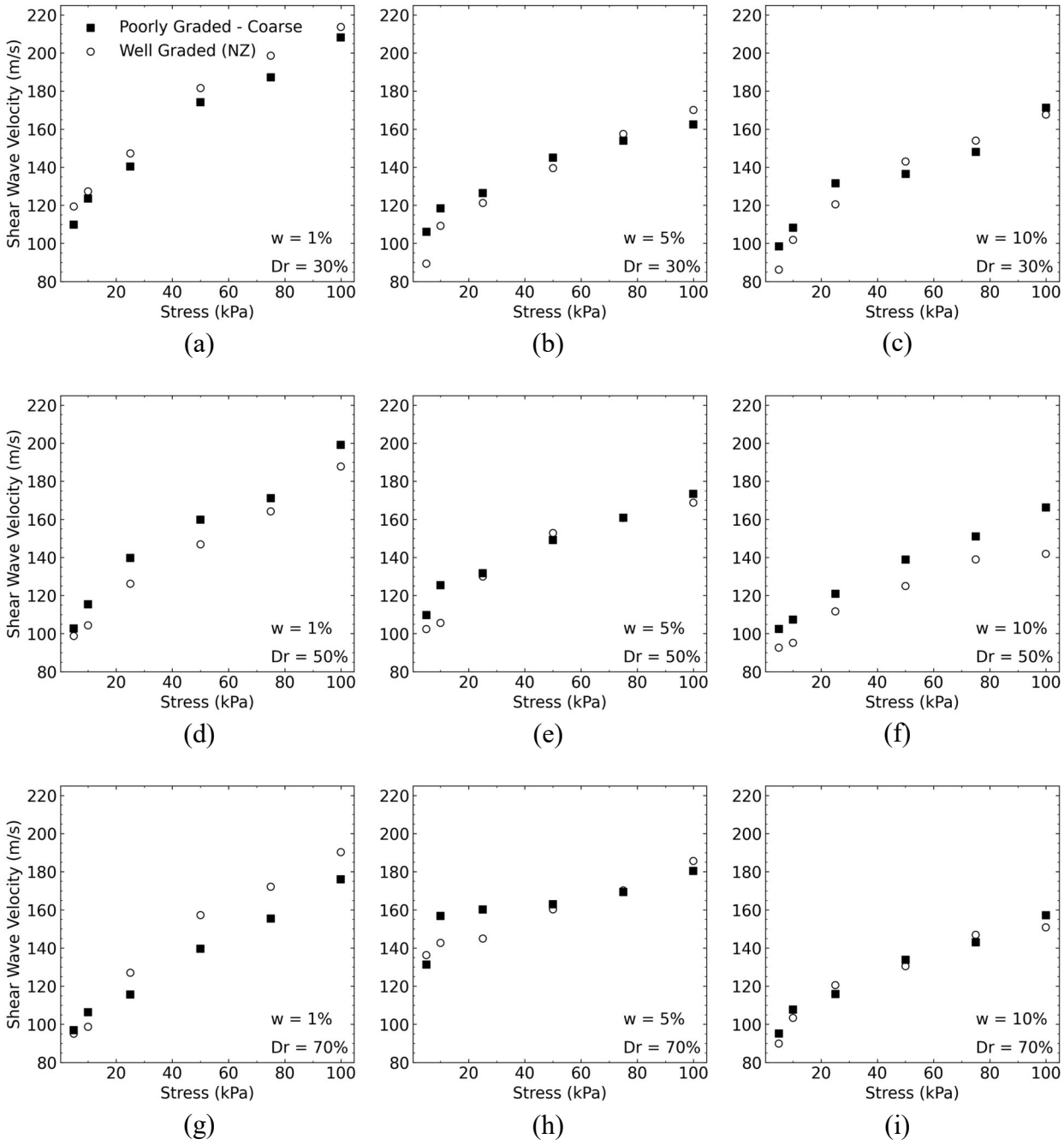


Figure A.7. Soil gradation comparison for shear wave velocity for the New Zealand soils using the fine poorly graded soil for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 1% and Dr = 50%, (e) w = 5% and Dr = 50%, (f) w = 10% and Dr = 50%, (g) w = 1% and Dr = 70%, (h) w = 5% and Dr = 70%, and (i) w = 10% and Dr = 70%

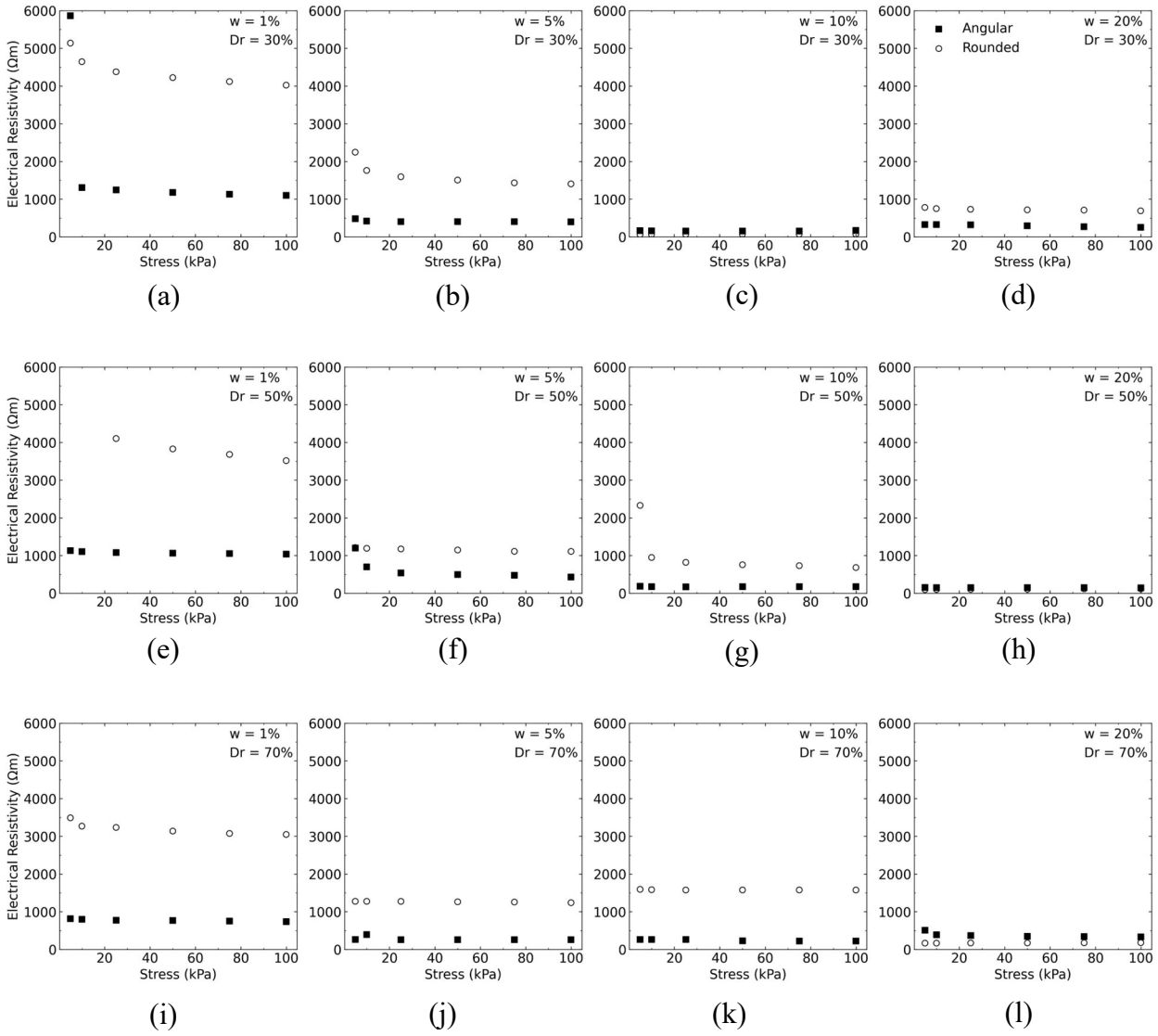


Figure A.9. Particle angularity comparison for electrical resistivity between the crushed angular soil and the C109 sand for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

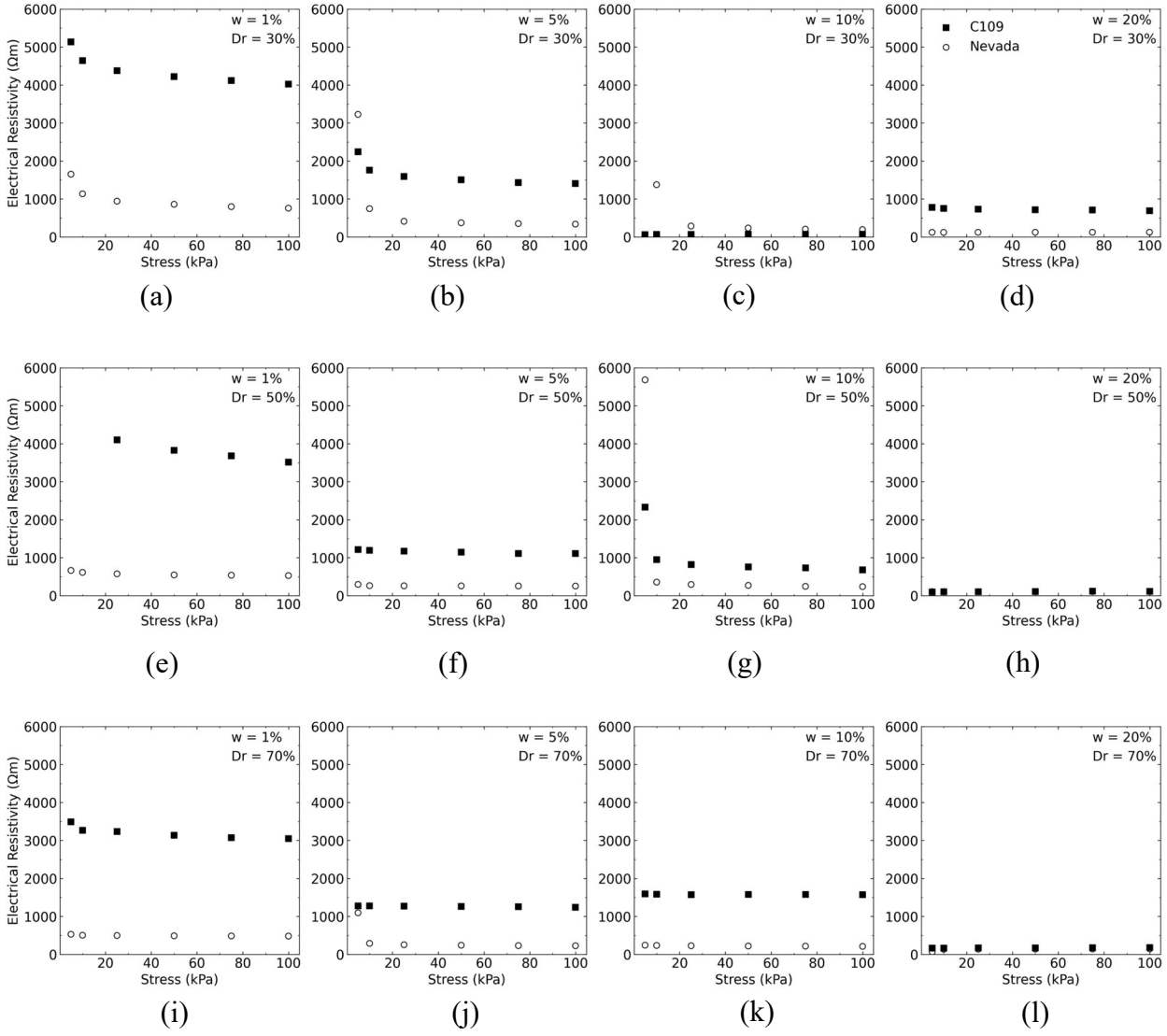


Figure A.10. Particle grain size comparison for electrical resistivity between C109 sand and Nevada sand for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

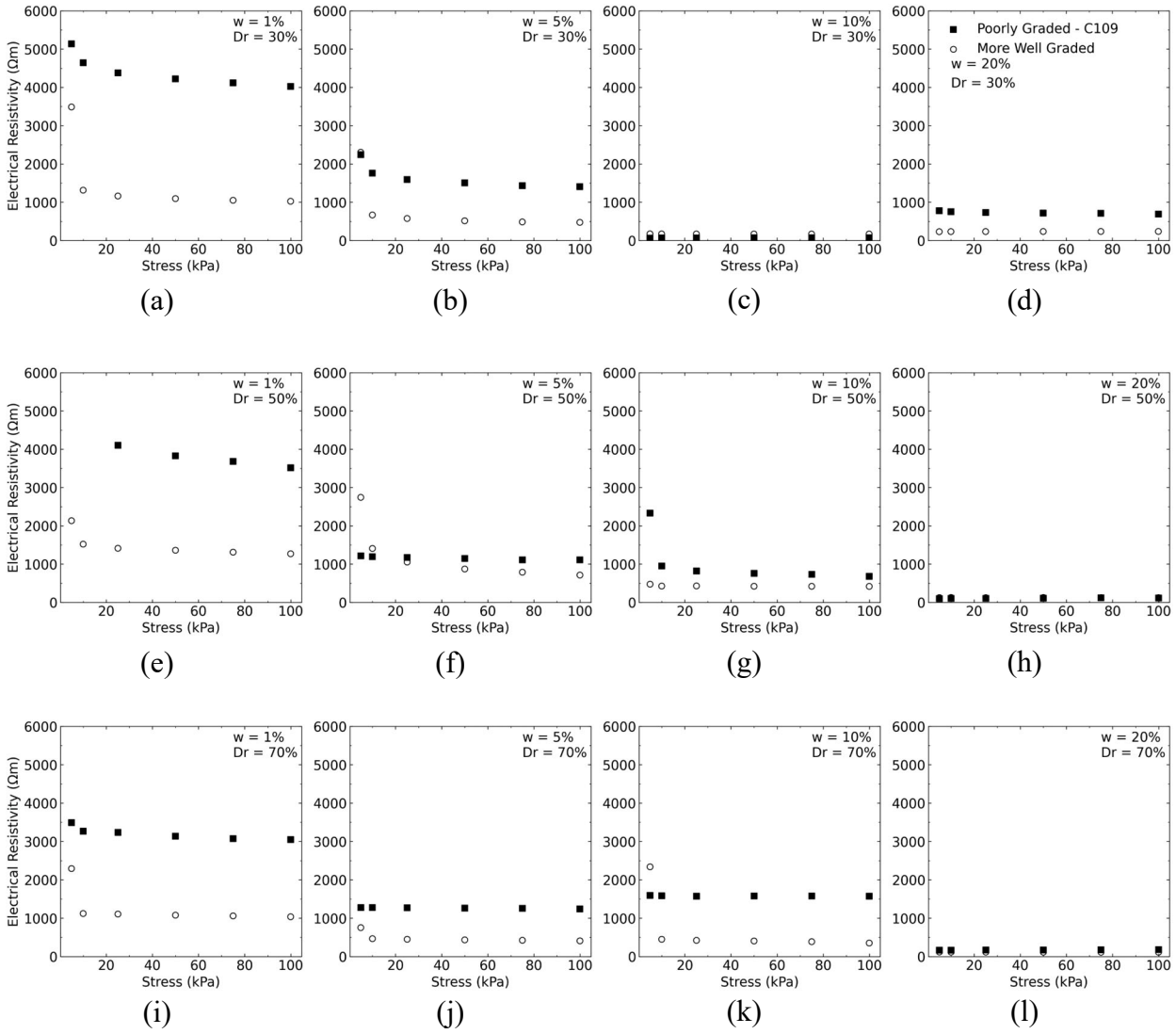


Figure A.11. Soil gradation comparison for electrical resistivity between C109 sand and the C109-Nevada mixture for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

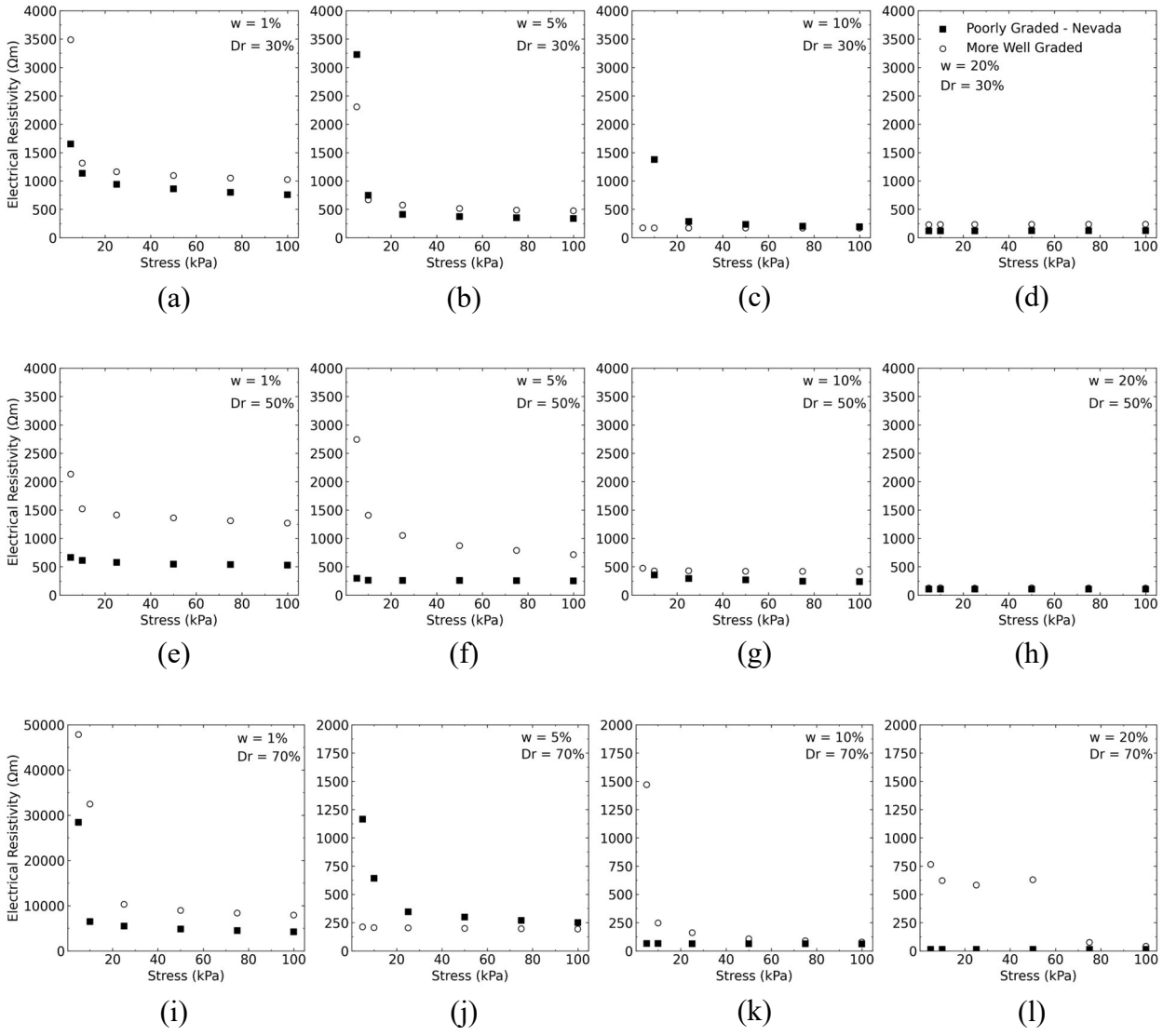


Figure A.12. Soil gradation comparison for electrical resistivity between C109 sand and the C109-Nevada mixture for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

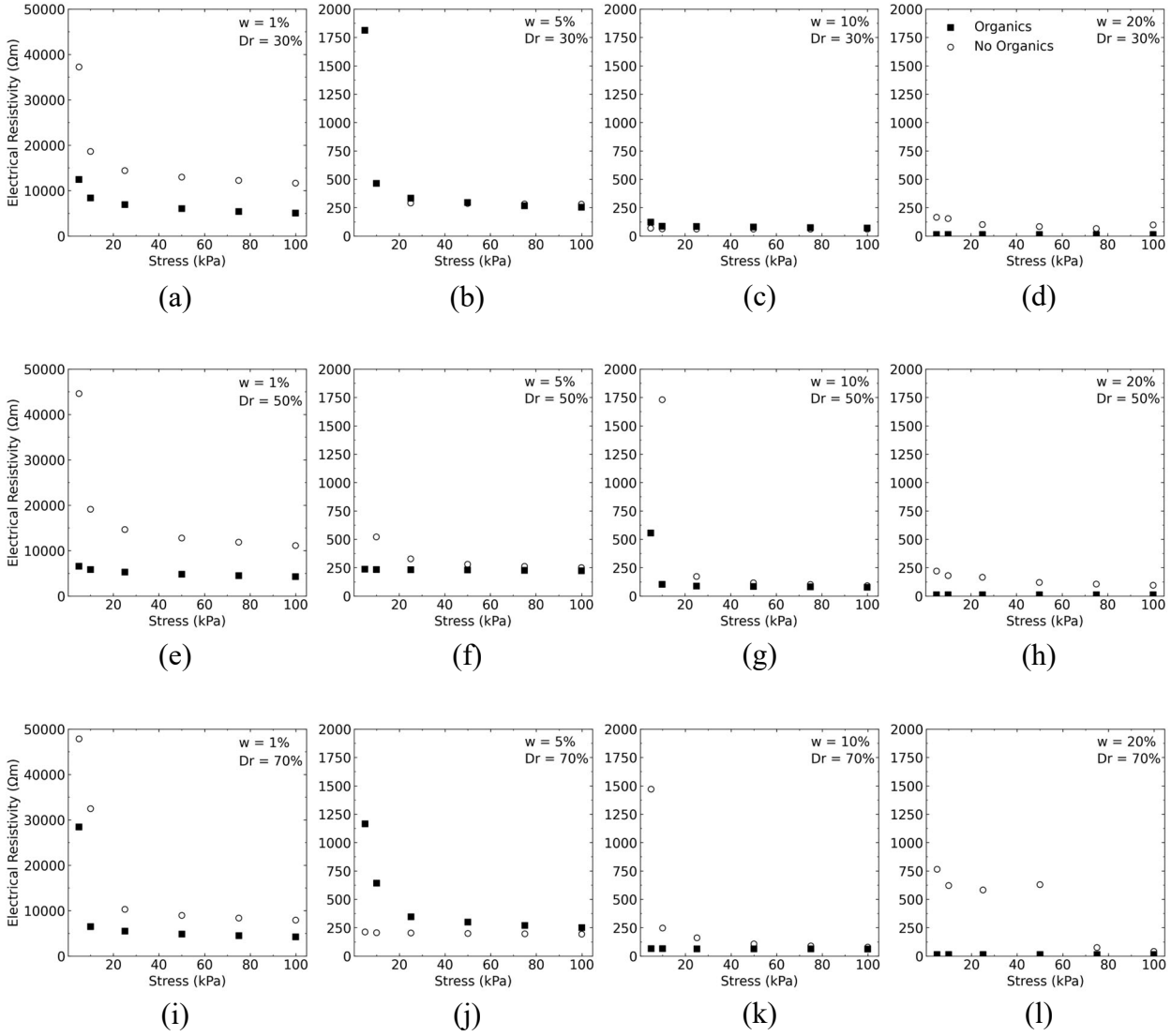


Figure A.13. Organic matter comparison for electrical resistivity between the Topatopa soils for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 20% and Dr = 30%, (e) w = 1% and Dr = 50%, (f) w = 5% and Dr = 50%, (g) w = 5% and Dr = 50%, (h) w = 20% and Dr = 50%, (i) w = 1% and Dr = 70%, (j) w = 5% and Dr = 70%, (k) w = 10% and Dr =70% and (l) w = 20% and Dr = 70%

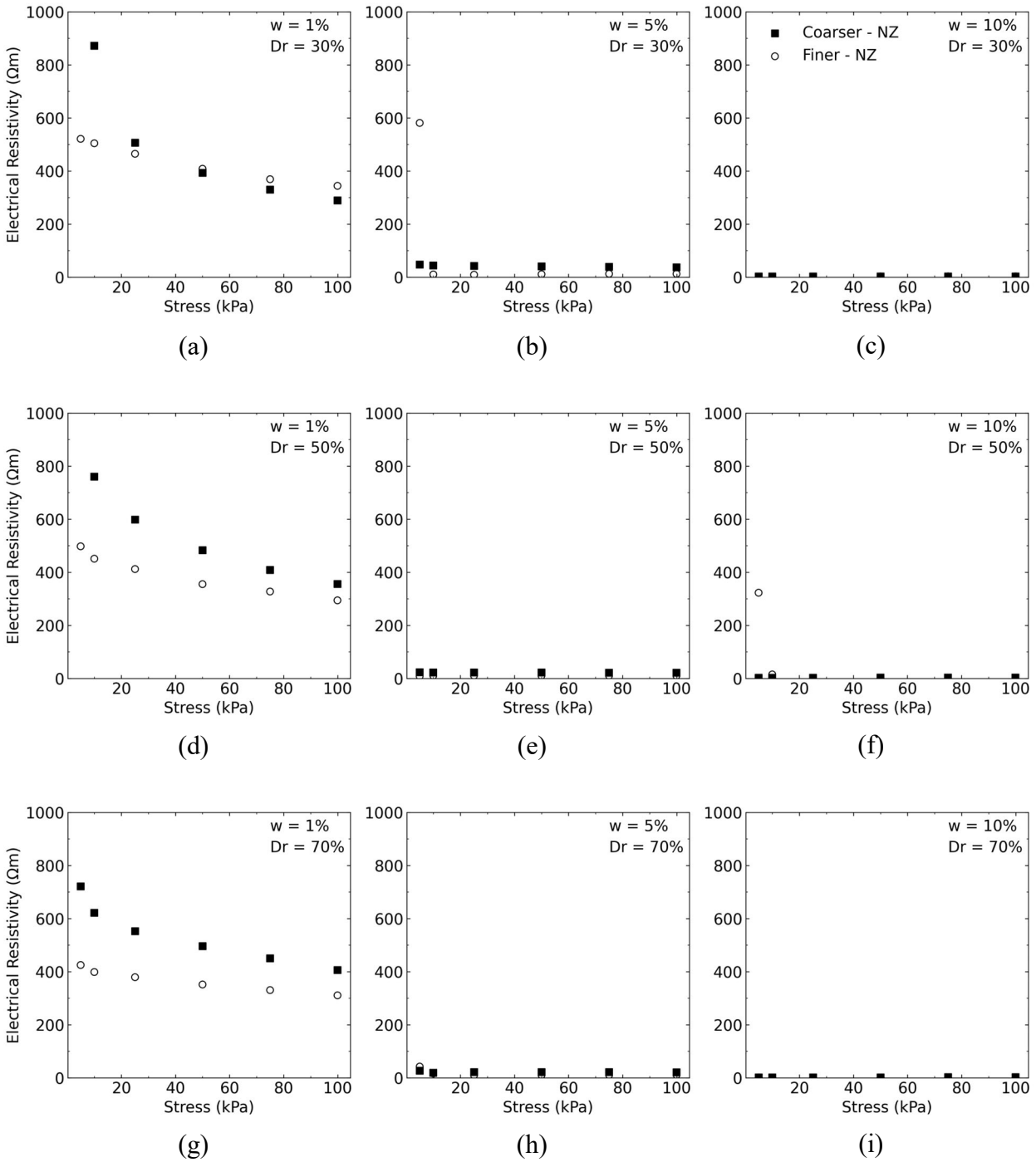


Figure A.14. Particle grain size comparison for electrical resistivity between the coarse and fine New Zealand soils (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 1% and Dr = 50%, (e) w = 5% and Dr = 50%, (f) w = 10% and Dr = 50%, (g) w = 1% and Dr = 70%, (h) w = 5% and Dr = 70%, and (i) w = 10% and Dr = 70%

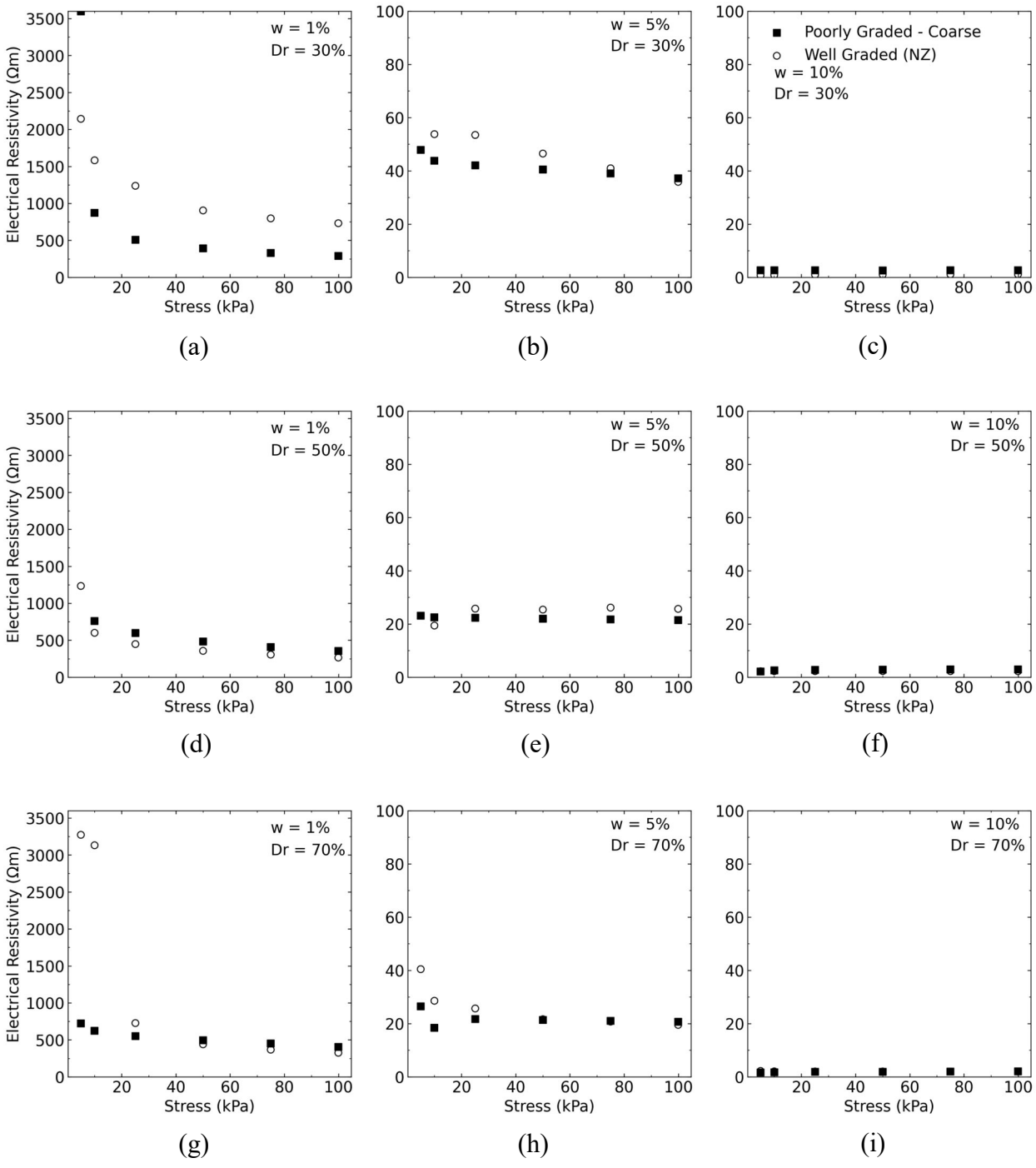


Figure A.15. Soil gradation comparison for electrical resistivity for the New Zealand soils using the coarse poorly graded soil for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 1% and Dr = 50%, (e) w = 5% and Dr = 50%, (f) w = 10% and Dr = 50%, (g) w = 1% and Dr = 70%, (h) w = 5% and Dr = 70%, and (i) w = 10% and Dr = 70%

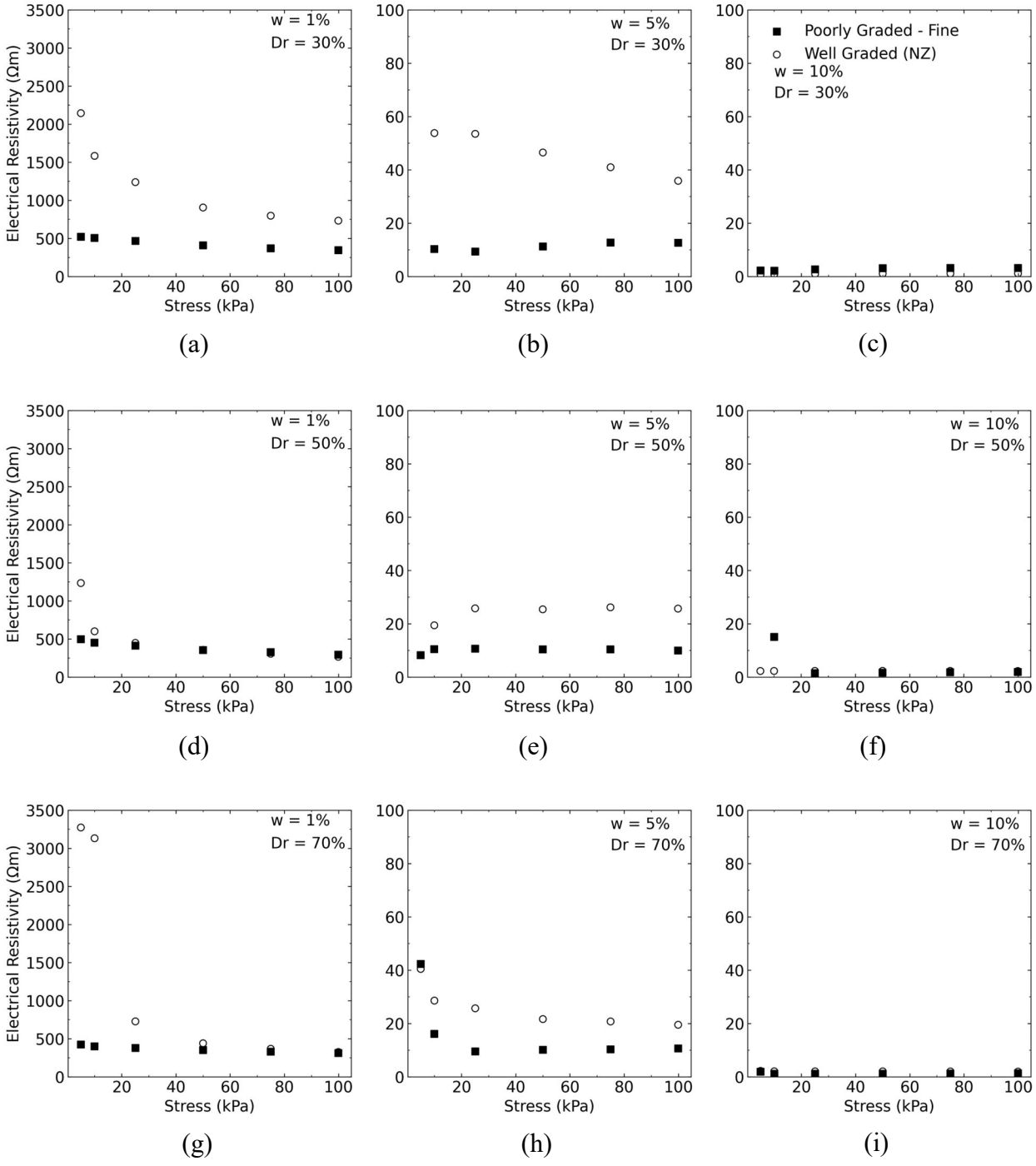


Figure A.16. Soil gradation comparison for electrical resistivity for the New Zealand soils using the fine poorly graded soil for (a) w = 1% and Dr = 30%, (b) w = 5% and Dr = 30%, (c) w = 10% and Dr = 30%, (d) w = 1% and Dr = 50%, (e) w = 5% and Dr = 50%, (f) w = 10% and Dr = 50%, (g) w = 1% and Dr = 70%, (h) w = 5% and Dr = 70%, and (i) w = 10% and Dr = 70%

Appendix B – Geotechnical Borings and CPTs

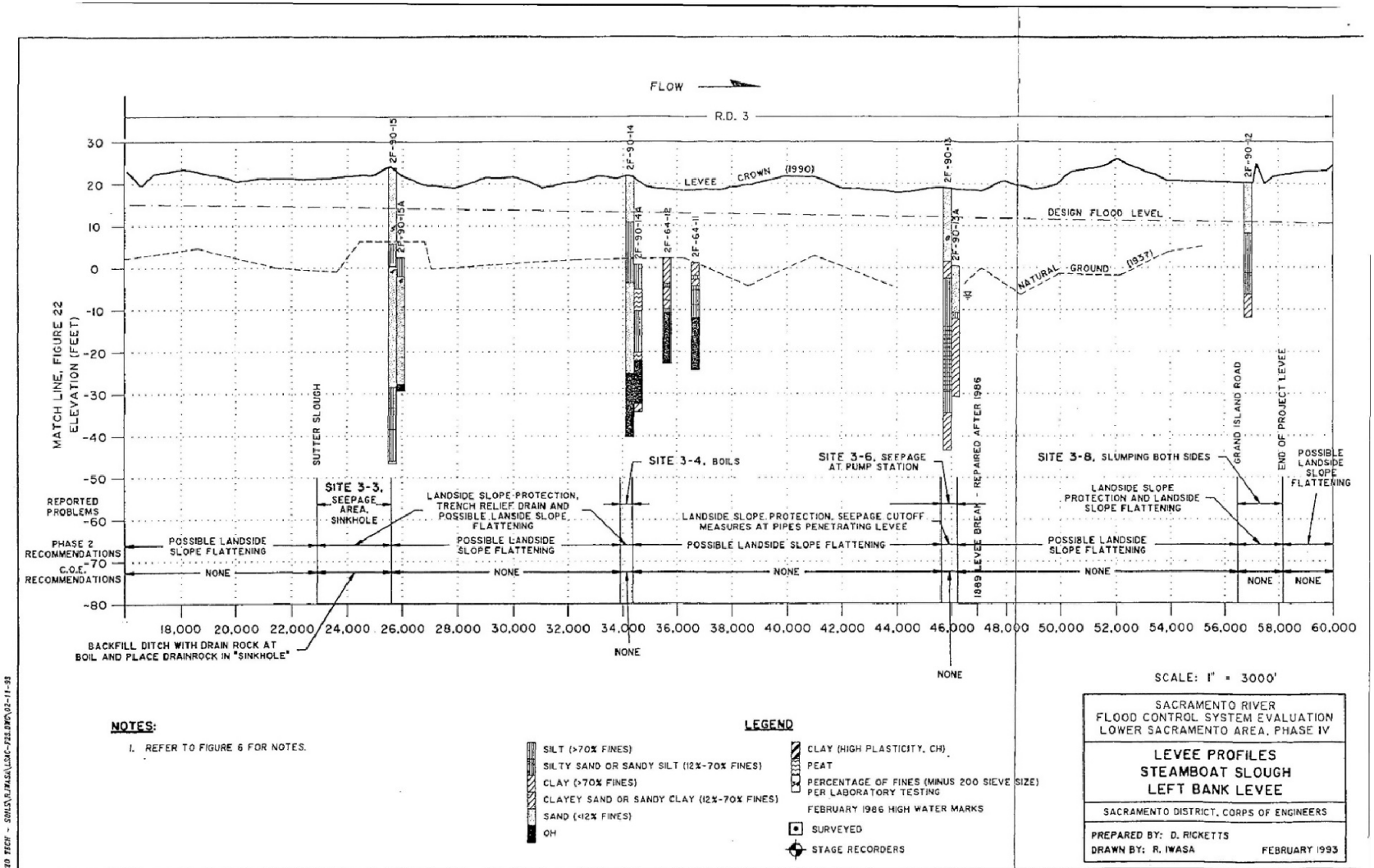

FLOW
R.D. 3
LEVEE CROWN (1990)
DESIGN FLOOD LEVEL
NATURAL GROUND (1937)
MATCH LINE, FIGURE 22
ELEVATION (FEET)
SUTTER SLOUGH
SITE 3-3, SEEPAGE AREA, SINKHOLE
SITE 3-4, BOILS
SITE 3-6, SEEPAGE AT PUMP STATION
SITE 3-8, SLUMPING BOTH SIDES
1986 LEVEE BREAK - REPAIRED AFTER 1986
GRAND ISLAND ROAD
END OF PROJECT LEVEE
POSSIBLE LANDSIDE SLOPE FLATTENING
LANDSIDE SLOPE PROTECTION, TRENCH RELIEF DRAIN AND POSSIBLE LANDSIDE SLOPE FLATTENING
LANDSIDE SLOPE PROTECTION, SEEPAGE CUTOFF MEASURES AT PIPES PENETRATING LEVEE
LANDSIDE SLOPE PROTECTION AND LANDSIDE SLOPE FLATTENING
REPORTED PROBLEMS
PHASE 2 RECOMMENDATIONS
C.O.E. RECOMMENDATIONS
NONE
BACKFILL DITCH WITH DRAIN ROCK AT BOIL AND PLACE DRAINROCK IN "SINKHOLE"
SCALE: 1" = 3000'
NOTES:
1. REFER TO FIGURE 6 FOR NOTES.
LEGEND
SILT (>70% FINES)
SILTY SAND OR SANDY SILT (12%-70% FINES)
CLAY (>70% FINES)
CLAYEY SAND OR SANDY CLAY (12%-70% FINES)
SAND (<12% FINES)
OH
CLAY (HIGH PLASTICITY, CH)
PEAT
PERCENTAGE OF FINES (MINUS 200 SIEVE SIZE) PER LABORATORY TESTING
FEBRUARY 1986 HIGH WATER MARKS
SURVEYED
STAGE RECORDERS
SACRAMENTO RIVER
FLOOD CONTROL SYSTEM EVALUATION
LOWER SACRAMENTO AREA, PHASE IV
LEVEE PROFILES
STEAMBOAT SLOUGH
LEFT BANK LEVEE
SACRAMENTO DISTRICT, CORPS OF ENGINEERS
PREPARED BY: D. RICKETTS
DRAWN BY: R. IWASA
FEBRUARY 1993

FIGURE 23

Date : 4/28/2021
Drilling Method : Rotary Wash
Elevation (Feet) : 1
Latitude : 38.19296
Longitude : -121.61982

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Topsoil, brown, moist, with rootlets, (12-inches)					
	M	2		SP-SM	▽	Poorly-Graded Sand with Silt (SP-SM), gray, medium grained sand, wet, very loose to loose, plant rootlets			48 55	66	Sieve -200=9
5	M	4									
10	S	4		OH		Organic Silt (OH), gray, wet, soft, low plasticity, no dilatancy, fibrous/woodey debris					
						Organic Silt (OH), gray, wet, soft to firm, medium plasticity, no dilatancy, fibrous, woody, peaty debris					
15	M	2					0.25 0.20	0.5	119	39	LL=88 PI=35
20	M	1					0.18	0.5	145	32	
25	M	4		OH			0.23 0.23	0.5 0.5			
30	M	2				Gray to dark brown	0.25 0.25	0.8 0.8			
						minor fibrous debris					
35	M	2					0.35 0.20	0.5 0.5			
40	M	4		CL		Lean Clay (CL), dark gray, wet, soft to firm, low plasticity, no dilatancy, medium toughness	0.25	0.5			

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 1
(Page 1 of 2)

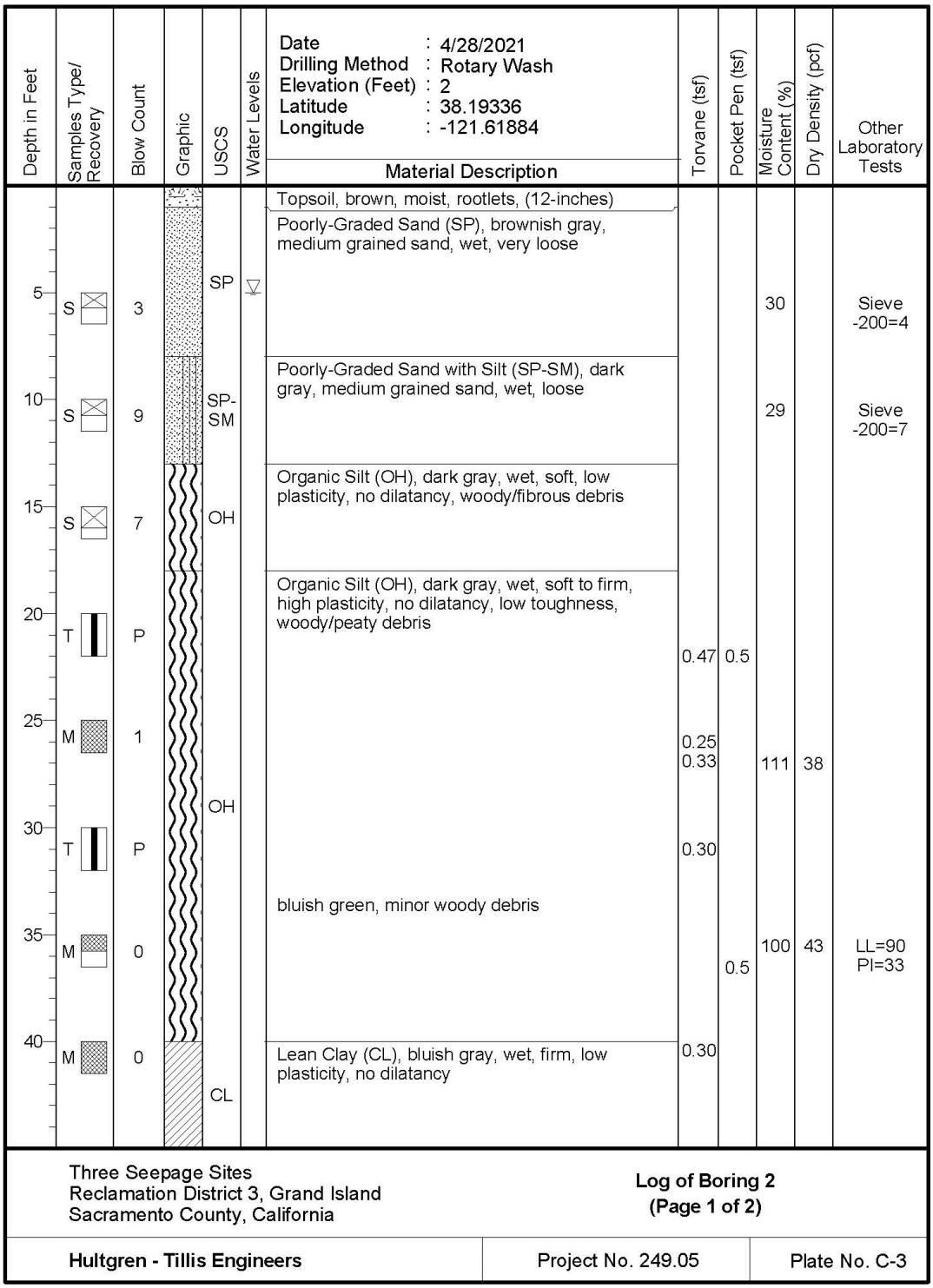
Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-1

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Date : 4/28/2021 Drilling Method : Rotary Wash Elevation (Feet) : 1 Latitude : 38.19296 Longitude : -121.61982					
	M	32		CL		Very stiff	0.60	3.3 3.5			
50	M	34		SP		Poorly-Graded Sand (SP), dark gray, medium grained sand, wet, dense					
						Bottom of boring at 51.5 feet Groundwater encountered at 2.5 feet					

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 1
(Page 2 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-2



Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 2
(Page 1 of 2)

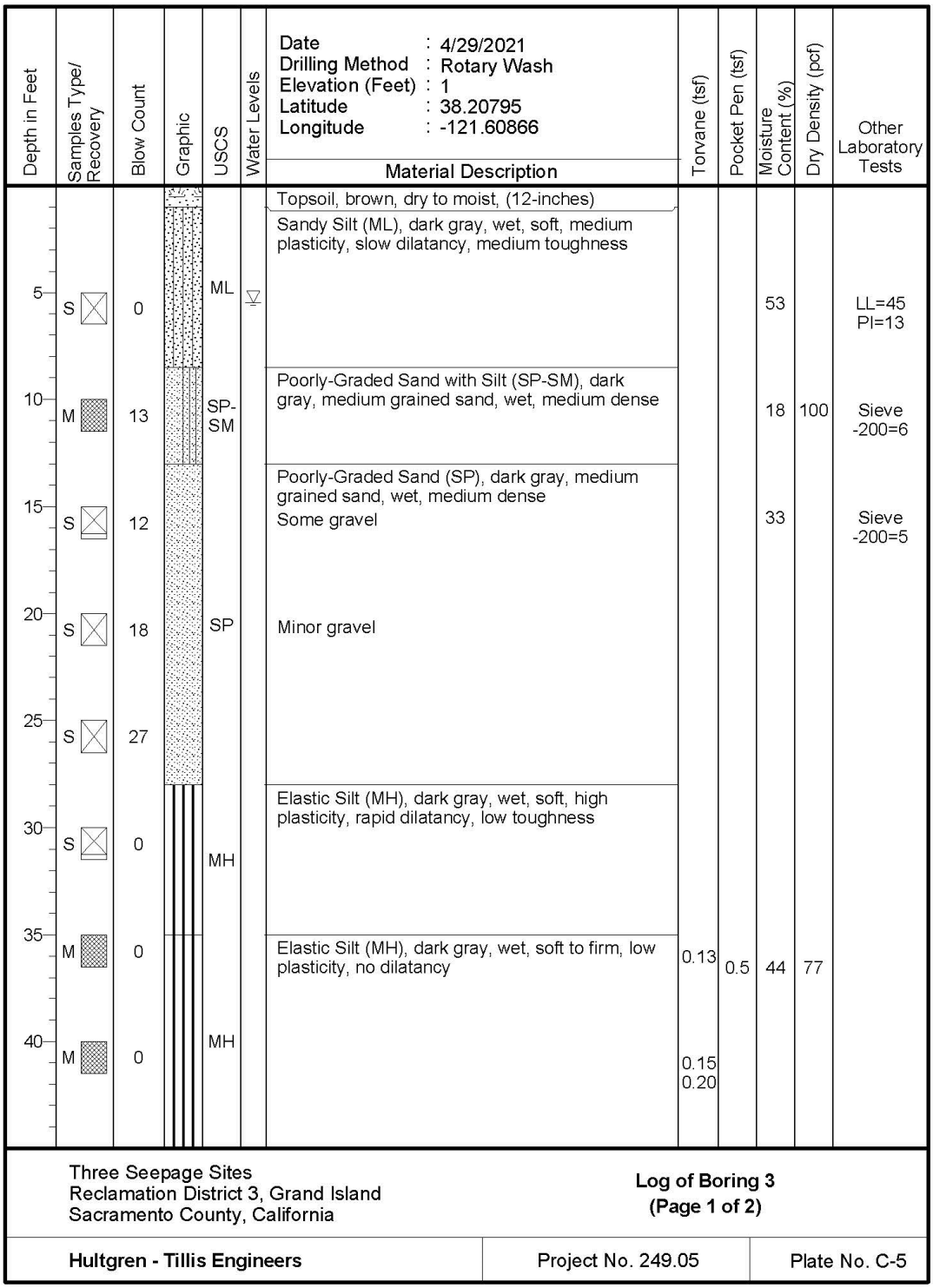
Hultgren - Tillis Engineers	Project No. 249.05	Plate No. C-3

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Date : 4/28/2021 Drilling Method : Rotary Wash Elevation (Feet) : 2 Latitude : 38.19336 Longitude : -121.61884					
	M	25		CL		Very stiff	0.55	3.5 3.5	26	98	
50	S	50		SM		Silty Sand (SM), gray, medium grained sand, wet, dense					
						Bottom of boring at 51.5 feet Groundwater encountered at 5 feet					

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 2
(Page 2 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-4



Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 3
(Page 1 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-5

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Date : 4/29/2021 Drilling Method : Rotary Wash Elevation (Feet) : 1 Latitude : 38.20795 Longitude : -121.60866					
	M	14		MH		Stiff	2.00	1.5 2.0			
50	M	18		SP-SM		Poorly-Graded Sand with Silt (SP-SM), dark gray, wet, medium dense					Sieve -200=9
						Bottom of boring at 51.5 feet Groundwater encountered at 5.5 feet					

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 3
(Page 2 of 2)

Hultgren - Tillis Engineers	Project No. 249.05	Plate No. C-6

Date : 4/29/2021
Drilling Method : Rotary Wash
Elevation (Feet) : 4
Latitude : 38.20835
Longitude : -121.60843

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Topsoil, brown, dry, (30-inches)					
5	S	0		ML	▽	Sandy Silt (ML), dark gray, medium grained sand, wet, soft					Sieve -200=57
10	S	4		SP-SM		Poorly-Graded Sand with Silt and Gravel (SP-SM), dark gray, medium grained sand, wet, loose					
15	S	13				Poorly-Graded Sand with Silt (SP-SM), dark gray, medium grained sand, wet, medium dense					
20	S	15		SP-SM					24		Sieve -200=5
25	S	15									
30	S	15		MH		Elastic Silt (MH), dark gray, wet, soft, high plasticity, rapid dilatancy					
35	T	P					0.15				
						Elastic Silt (MH), dark gray, wet, very stiff, low plasticity, no dilatancy					
40	M	0		MH			0.30		46	77	LL=56 PI=20

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 4
(Page 1 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-7

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Date : 4/29/2021 Drilling Method : Rotary Wash Elevation (Feet) : 4 Latitude : 38.20835 Longitude : -121.60843					
	M	27		MH				3.5 2.5	24	105	
50	S	24		SM		Silty Sand (SM), dark gray, fine grained sand, wet, medium dense					
						Bottom of boring at 51.5 feet Groundwater encountered at 5 feet					

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 4
(Page 2 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-8

Date : 4/30/2021
Drilling Method : Rotary Wash
Elevation (Feet) : 5
Latitude : 38.23594
Longitude : -121.60162

Depth in Feet	Samples Type/ Recovery	Blow Count	USCS	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
				Topsoil, brown, dry, (24-inches)					
				Elastic Silt (MH), gray, wet, soft, rapid dilatancy					
5	M	0		Very soft	0.15 0.05		45	76	
10	M	0			0.20 0.13		69	58	LL=65 PI=23
15	M	0	MH						
20	M	0							
	M	0			0.05		48		
				With sand					
25	M	0		With sand	0.25				
				Organic Silt (OH), gray, wet, soft, high plasticity, no dilatancy, organic/woody debris					
30	M	0	OH		0.35		78	52	LL=67 PI=26
				Elastic Silt (MH), gray, wet, stiff, high plasticity, no dilatancy					
35	T	P	MH		0.70 1.00	1.3			
				Lean Clay (CL), dark gray, wet, very stiff, low plasticity, no dilatancy					
40	M	28	CL			3.0 3.8			

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 5
(Page 1 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-9

Depth in Feet	Samples Type/ Recovery	Blow Count	Graphic	USCS	Water Levels	Material Description	Torvane (tsf)	Pocket Pen (tsf)	Moisture Content (%)	Dry Density (pcf)	Other Laboratory Tests
						Date : 4/30/2021 Drilling Method : Rotary Wash Elevation (Feet) : 5 Latitude : 38.23594 Longitude : -121.60162					
	M	46		CL		Hard		4.5+ 4.5+			
50	S	27		SM		Silty Sand (SM), bluish gray, fine grained sand, wet, medium dense					
						Bottom of boring at 51.5 feet Groundwater encountered at 2 feet					

Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

Log of Boring 5
(Page 2 of 2)

Hultgren - Tillis Engineers | Project No. 249.05 | Plate No. C-10

MAJOR DIVISIONS				GROUP NAMES
COARSE GRAINED SOILS MORE THAN 50% RETAINED ON NO. 200 SIEVE	GRAVELS MORE THAN 50% OF COARSE FRACTION IS RETAINED ON NO. 4 SIEVE	CLEAN GRAVELS WITH LESS THAN 5% FINES	GW	WELL GRADED GRAVEL
			GP	POORLY GRADED GRAVEL
		GRAVELS WITH OVER 12% FINES	GM	SILTY GRAVEL
			GC	CLAYEY GRAVEL
	SANDS 50% OR MORE OF COARSE FRACTION PASSES NO. 4 SIEVE	CLEAN SANDS WITH LESS THAN 5% FINES	SW	WELL GRADED SAND
			SP	POORLY GRADED SAND
		SANDS WITH OVER 12% FINES	SM	SILTY SAND
			SC	CLAYEY SAND
FINE GRAINED SOILS 50% OR MORE PASSES NO. 200 SIEVE	SILTS AND CLAYS LIQUID LIMIT LESS THAN 50		ML	SILT
			CL	LEAN CLAY
			OL	ORGANIC CLAY, ORGANIC SILT
	SILTS AND CLAYS LIQUID LIMIT 50 OR MORE		MH	ELASTIC SILT
			CH	FAT CLAY
			OH	ORGANIC CLAY, ORGANIC SILT
HIGHLY ORGANIC SOILS			Pt	PEAT

UNIFIED SOIL CLASSIFICATION SYSTEM- ASTM D 2487

S - SPT

M - 2.5 inch

C - 3.0 inch

T - Shelby Tube

B - Bag

- Water Level at Time of Drilling
- Water Level after Drilling (with date measured)

Consol - Consolidation
Gs - Specific Gravity
LL - Liquid Limit (%)
PI - Plasticity Index (%)
TxUU - Shear Strength (psf) - Unconsolidated Undrained Triaxial Shear
TxCU - Shear Strength (psf) - Consolidated Undrained Triaxial Shear
UC - Compressive Strength (psf) - Unconfined Compression

P - Push
Perm - Permeability
Sieve - Particle Size Analysis
VS - Laboratory Vane Shear (psf)
-200 - % Passing No. 200 Sieve

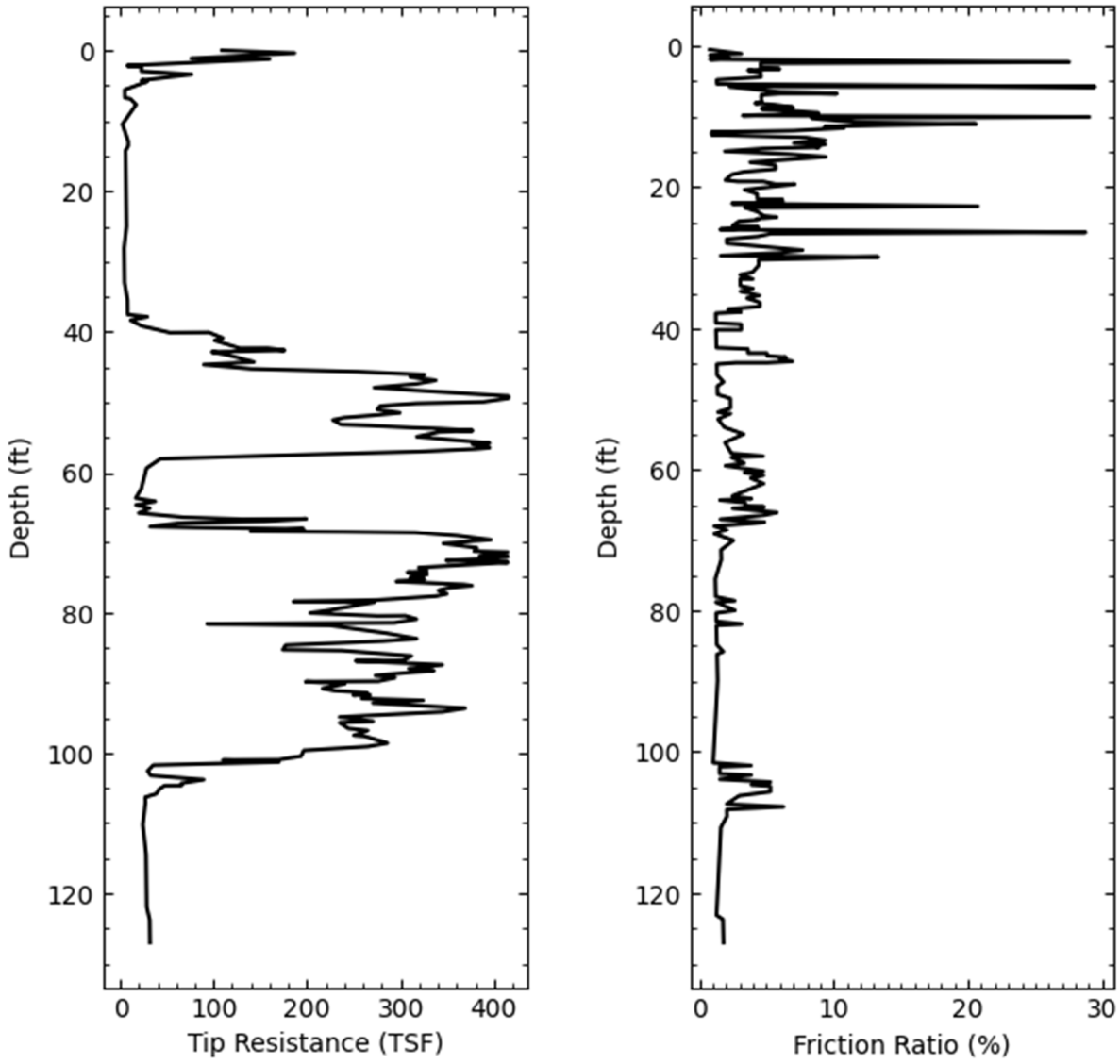
KEY TO TEST DATA

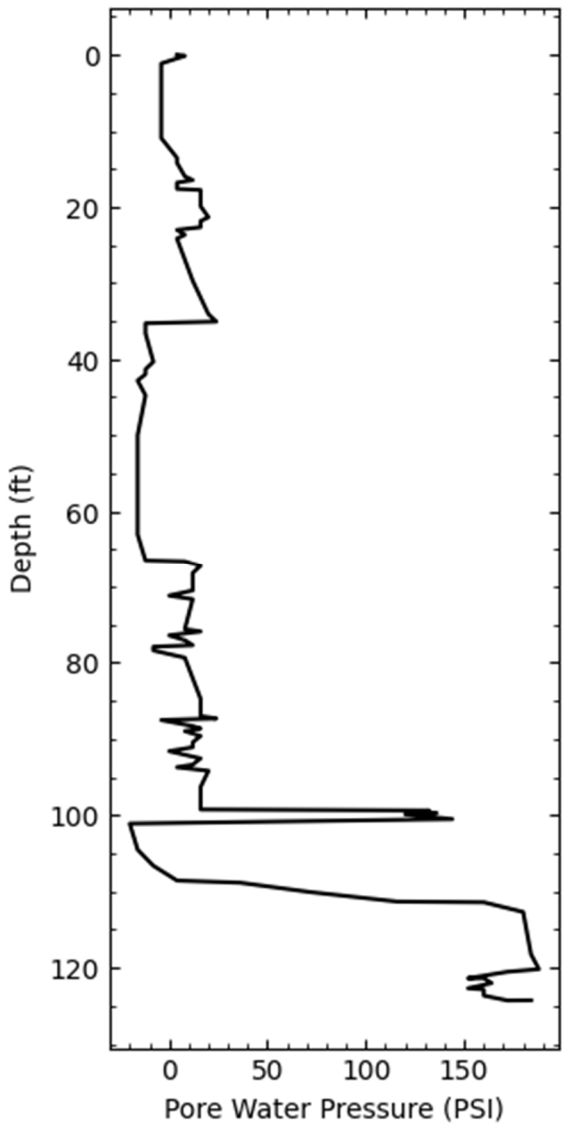
Three Seepage Sites
Reclamation District 3, Grand Island
Sacramento County, California

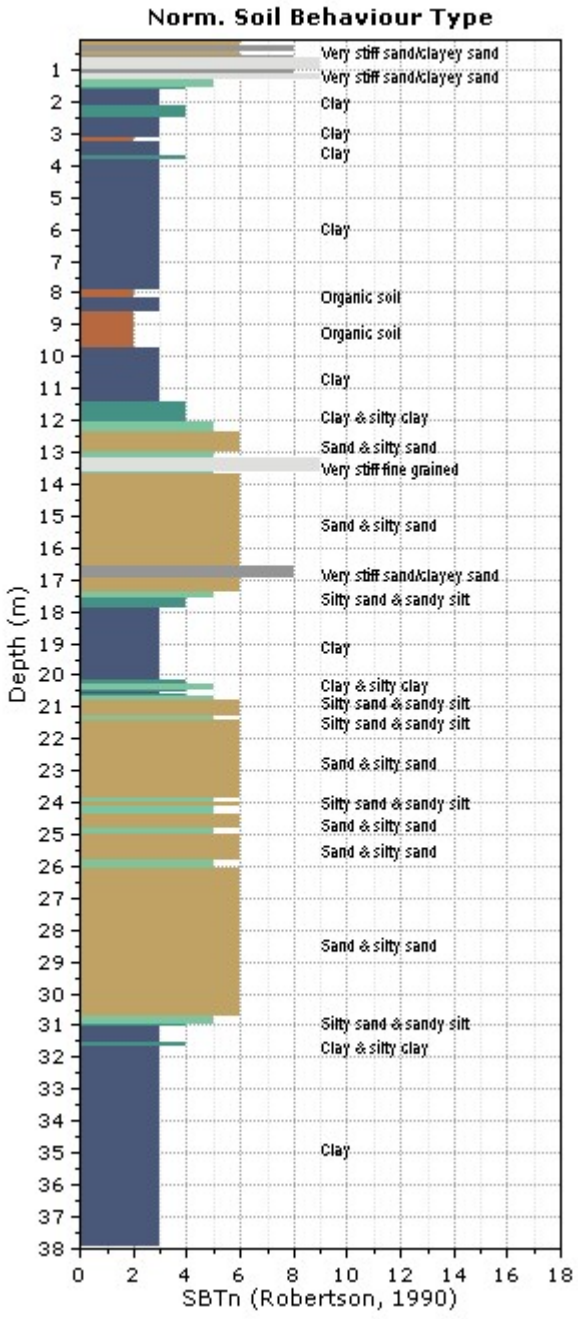
Soil Classification Chart

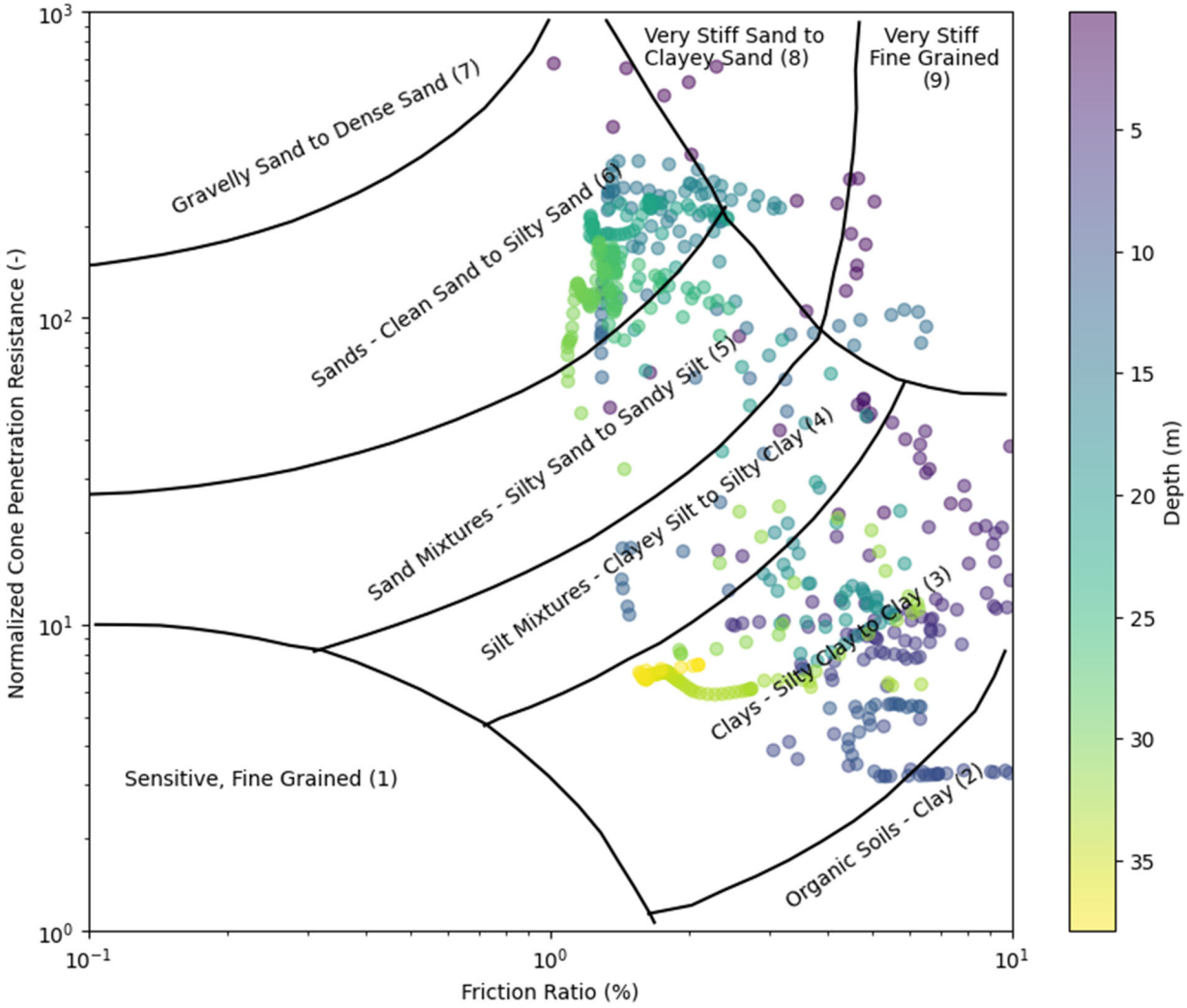
Hultgren - Tillis Engineers	Project No. 249.05	Plate No. C-11

BI-CPT-01-1

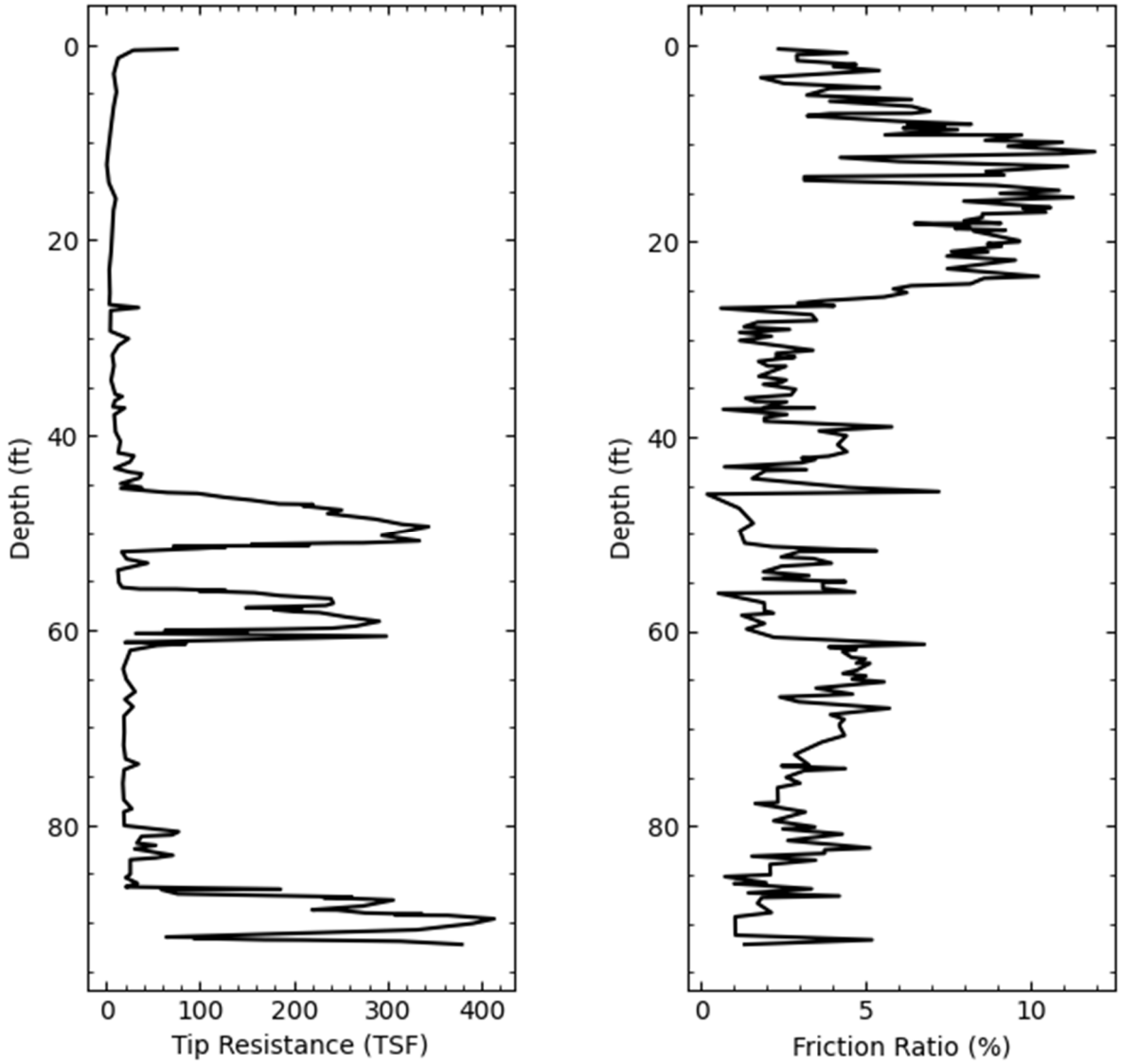

Depth (ft)
0
20
40
60
80
100
120
0
100
200
300
400
Tip Resistance (TSF)
Depth (ft)
0
20
40
60
80
100
120
0
10
20
30
Friction Ratio (%)

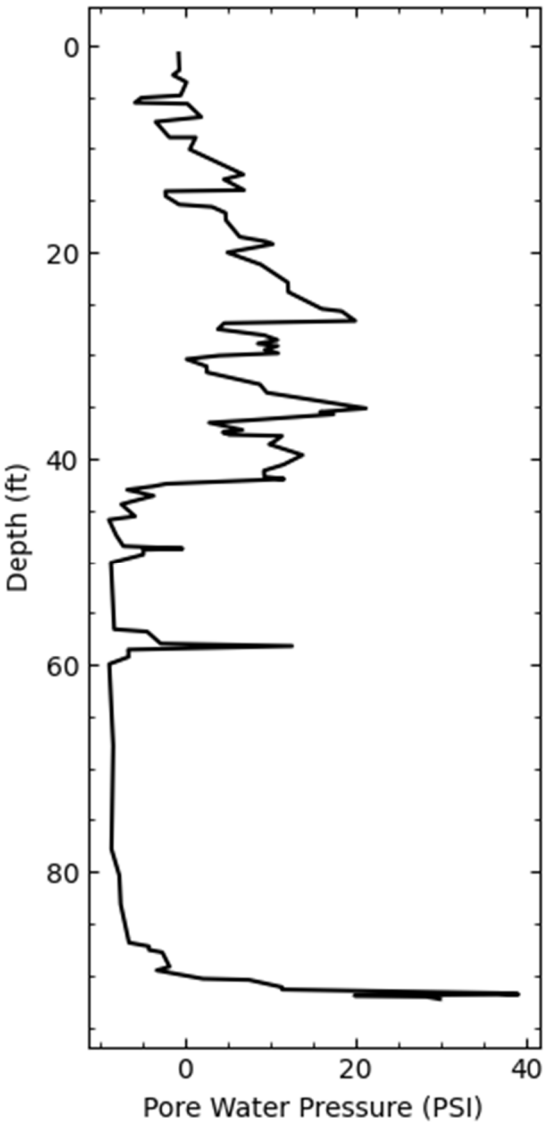

Depth (ft)
0
20
40
60
80
100
120
0
50
100
150
Pore Water Pressure (PSI)

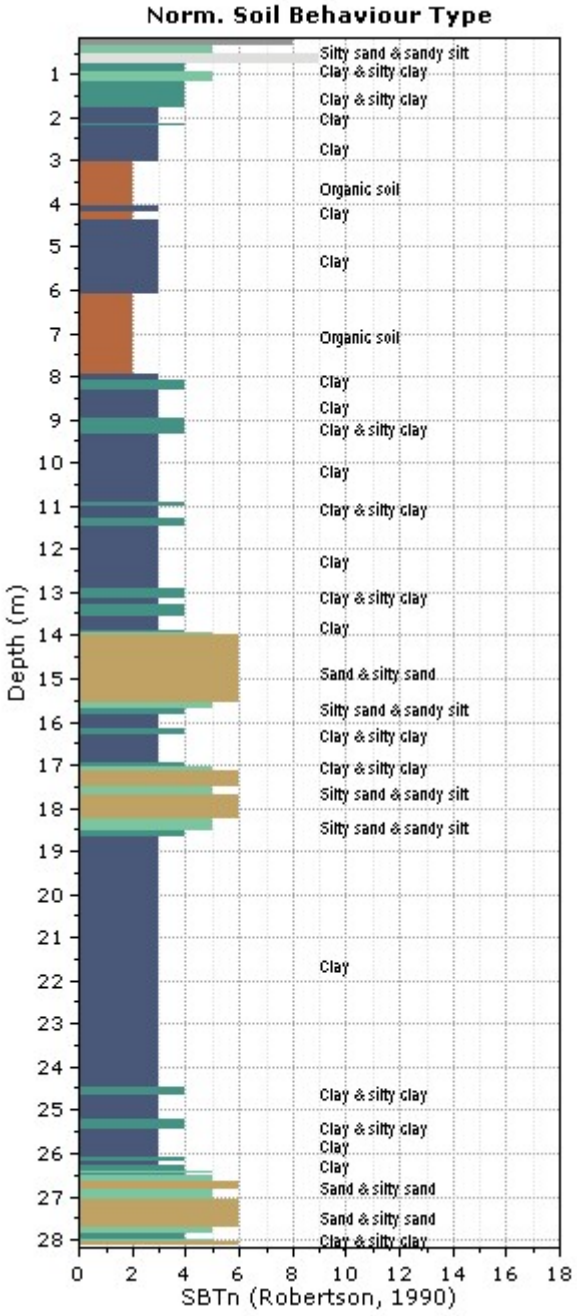

Norm. Soil Behaviour Type
Depth (m)
Very stiff sand/clayey sand
Very stiff sand/clayey sand
Clay
Clay
Clay
Clay
Organic soil
Organic soil
Clay
Clay & silty clay
Sand & silty sand
Very stiff fine grained
Sand & silty sand
Very stiff sand/clayey sand
Silty sand & sandy silt
Clay
Clay & silty clay
Silty sand & sandy silt
Silty sand & sandy silt
Sand & silty sand
Silty sand & sandy silt
Sand & silty sand
Sand & silty sand
Sand & silty sand
Silty sand & sandy silt
Clay & silty clay
Clay
0
2
4
6
8
10
12
14
16
18
SBTn (Robertson, 1990)

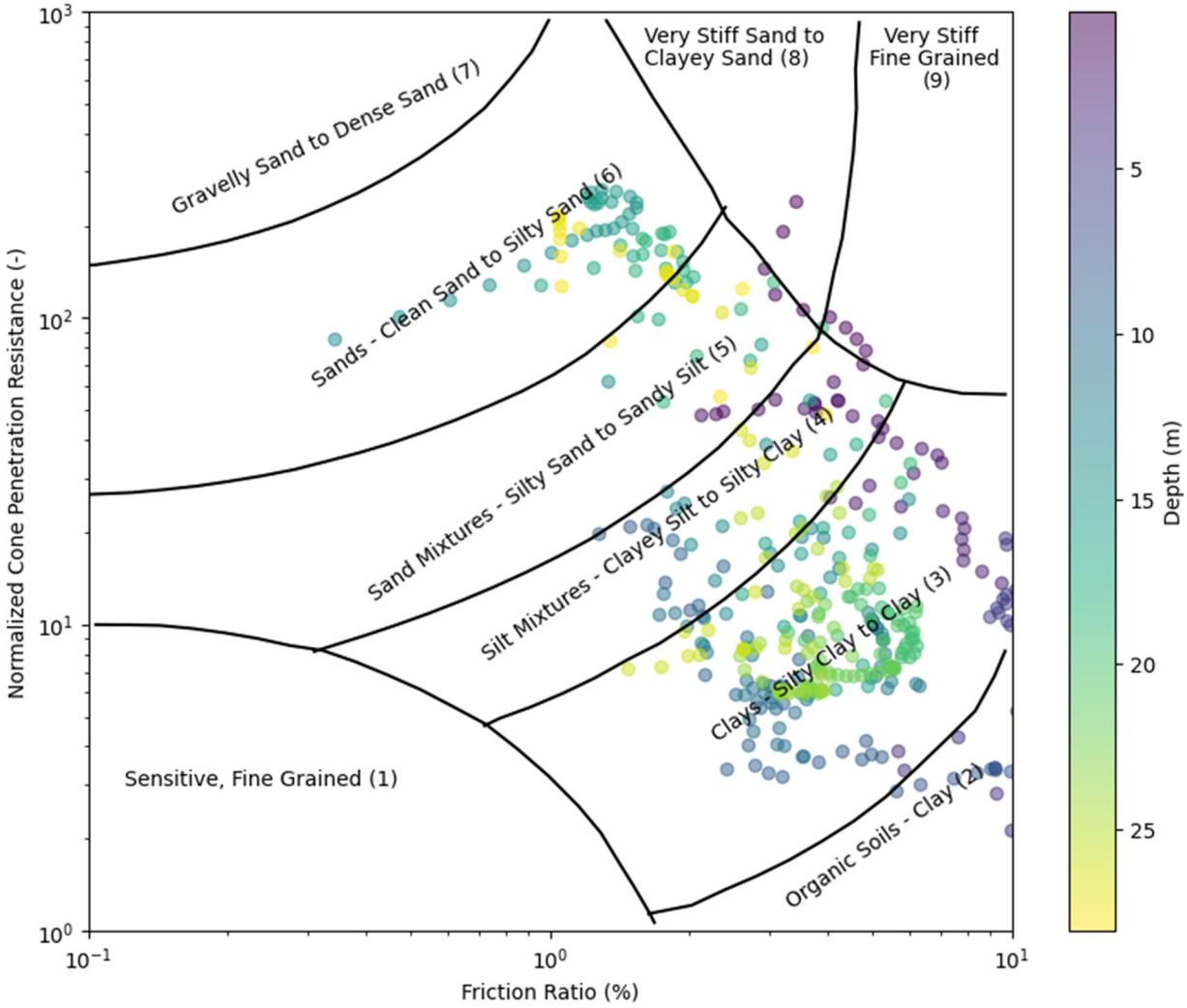

Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)
Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)

BI-CPT-01-4

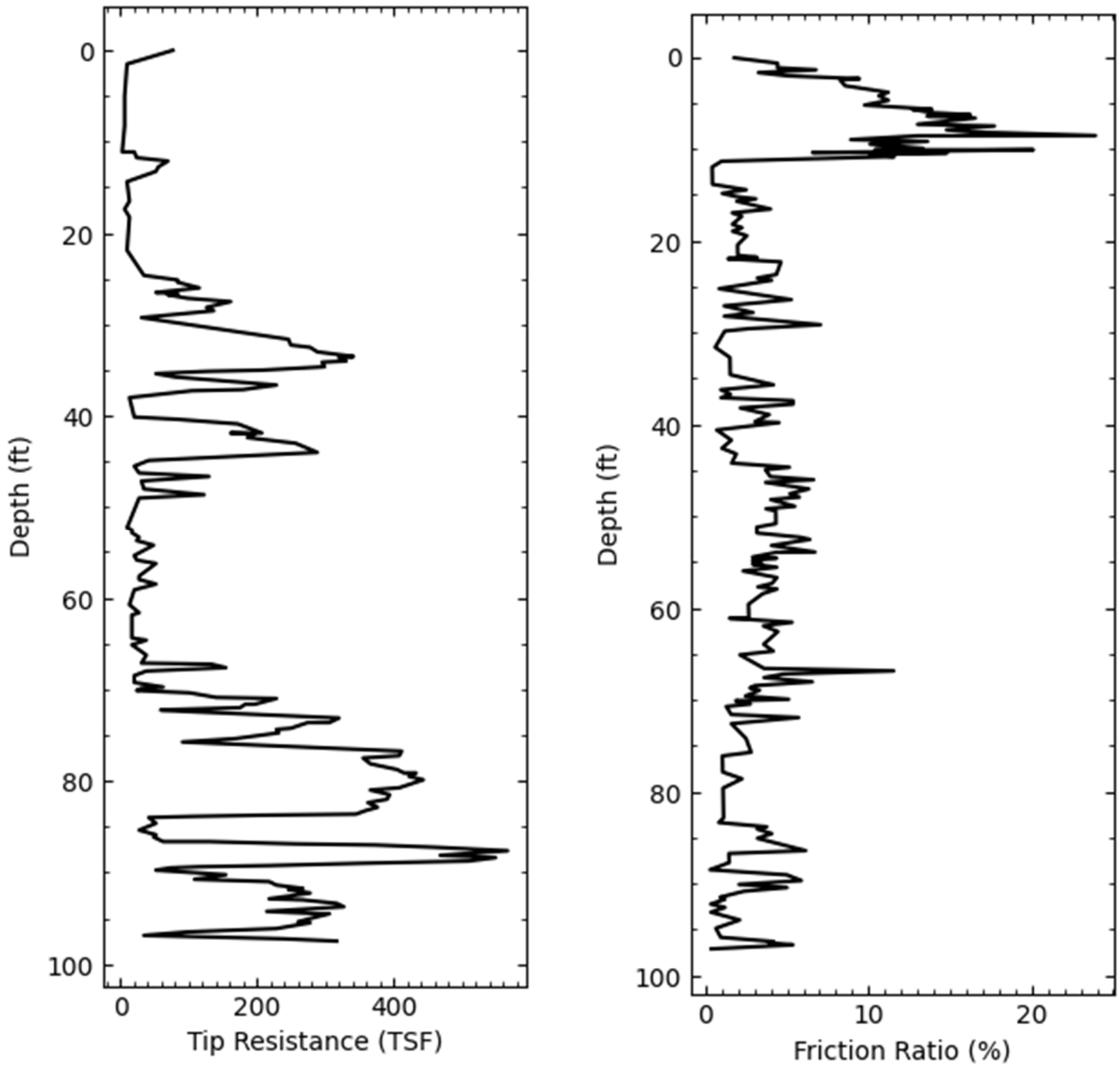

Depth (ft)
0
20
40
60
80
0
100
200
300
400
Tip Resistance (TSF)
Depth (ft)
0
20
40
60
80
0
5
10
Friction Ratio (%)

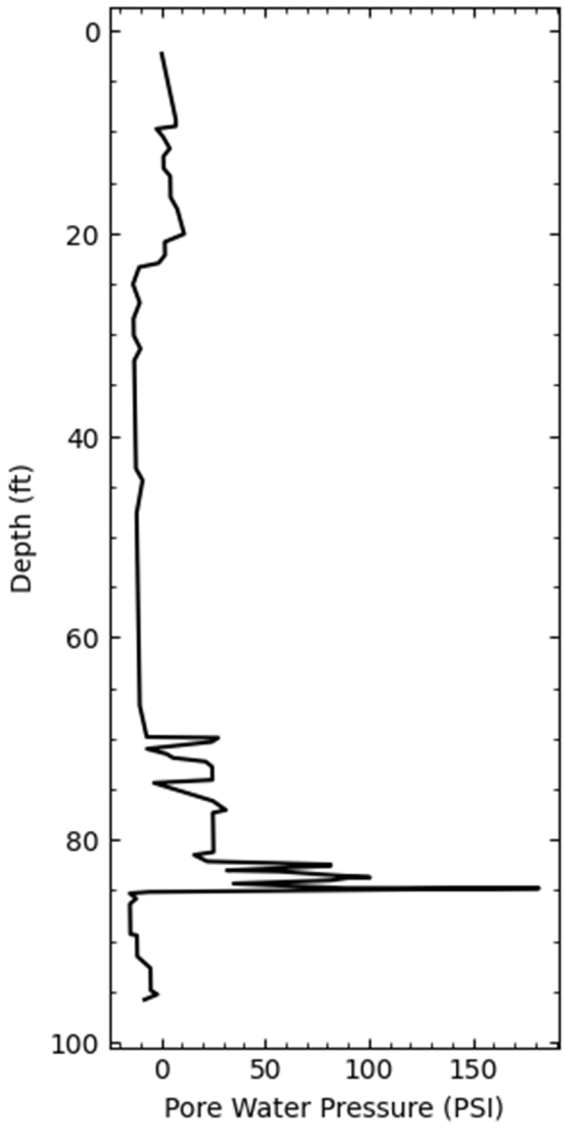

Depth (ft)
0
20
40
60
80
0
20
40
Pore Water Pressure (PSI)

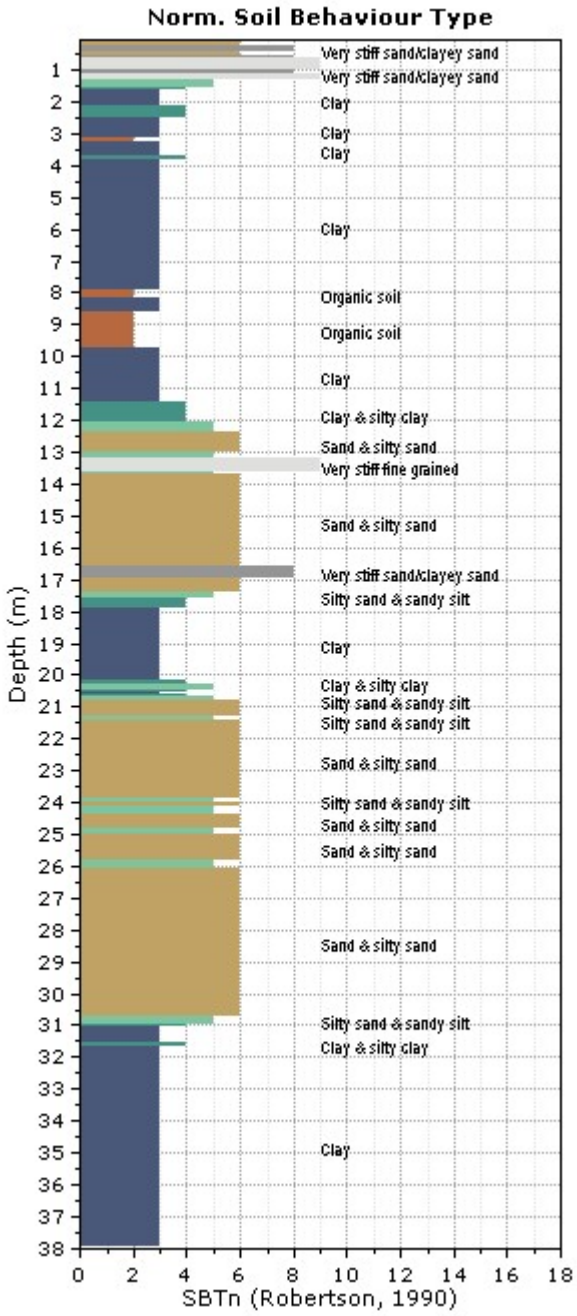

Norm. Soil Behaviour Type
Depth (m)
Silty sand & sandy silt
Clay & silty clay
Clay & silty clay
Clay
Clay
Organic soil
Clay
Clay
Organic soil
Clay
Clay
Clay & silty clay
Clay
Clay & silty clay
Clay
Clay & silty clay
Clay
Sand & silty sand
Silty sand & sandy silt
Clay & silty clay
Clay & silty clay
Silty sand & sandy silt
Silty sand & sandy silt
Clay
Clay & silty clay
Clay & silty clay
Clay
Clay
Sand & silty sand
Sand & silty sand
Clay & silty clay
0
2
4
6
8
10
12
14
16
18
SBTn (Robertson, 1990)

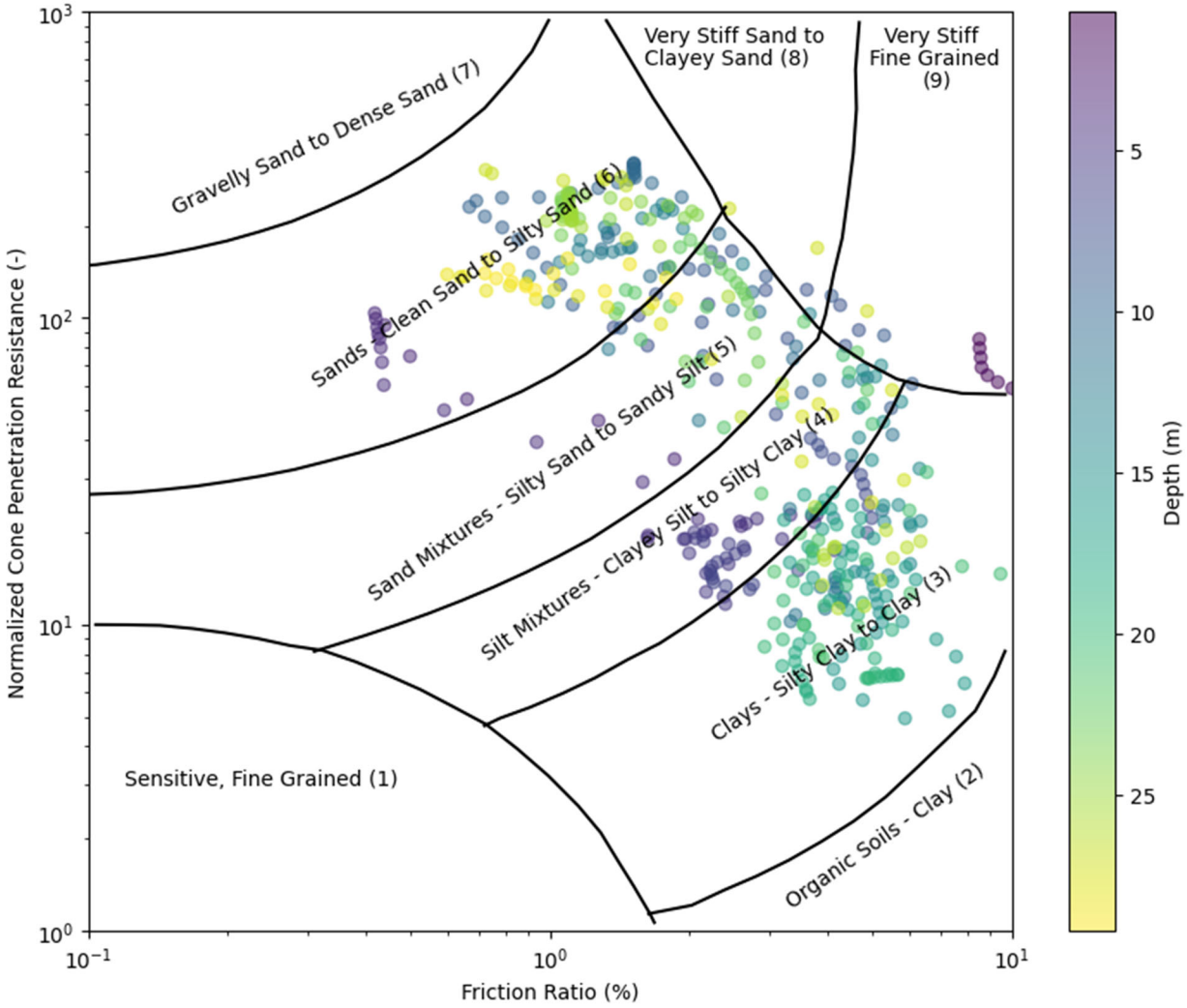

Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)
Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)

BI-CPT-01-7

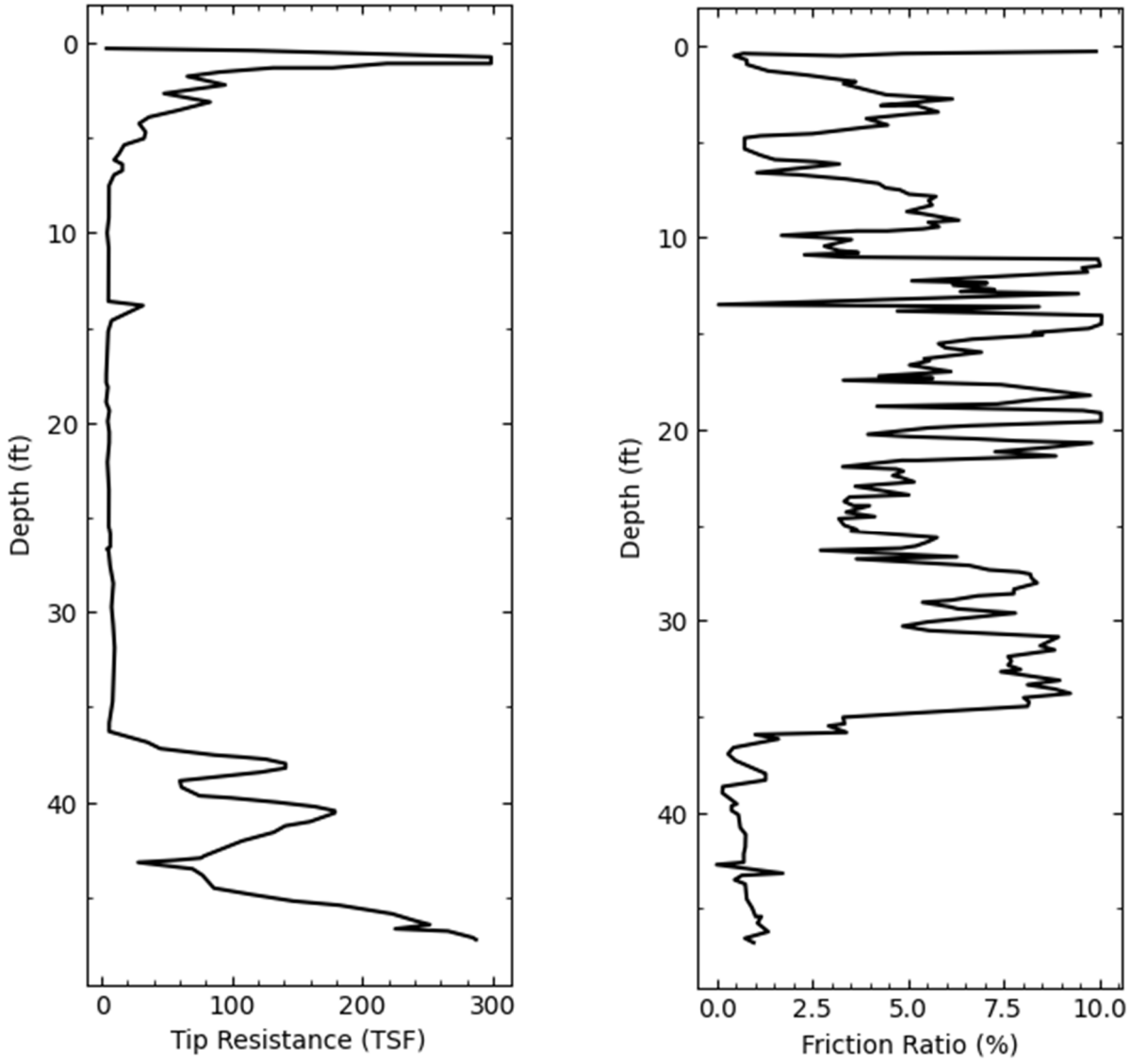

Depth (ft)
0
20
40
60
80
100
0
200
400
Tip Resistance (TSF)
Depth (ft)
0
20
40
60
80
100
0
10
20
Friction Ratio (%)

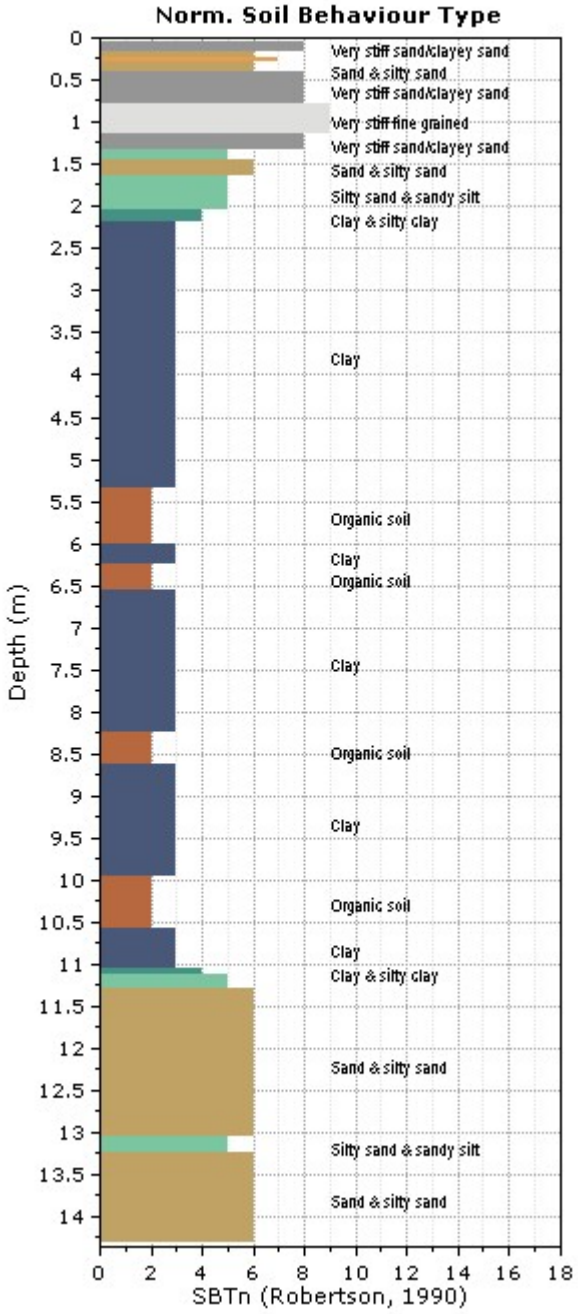

Depth (ft)
0
20
40
60
80
100
0
50
100
150
Pore Water Pressure (PSI)

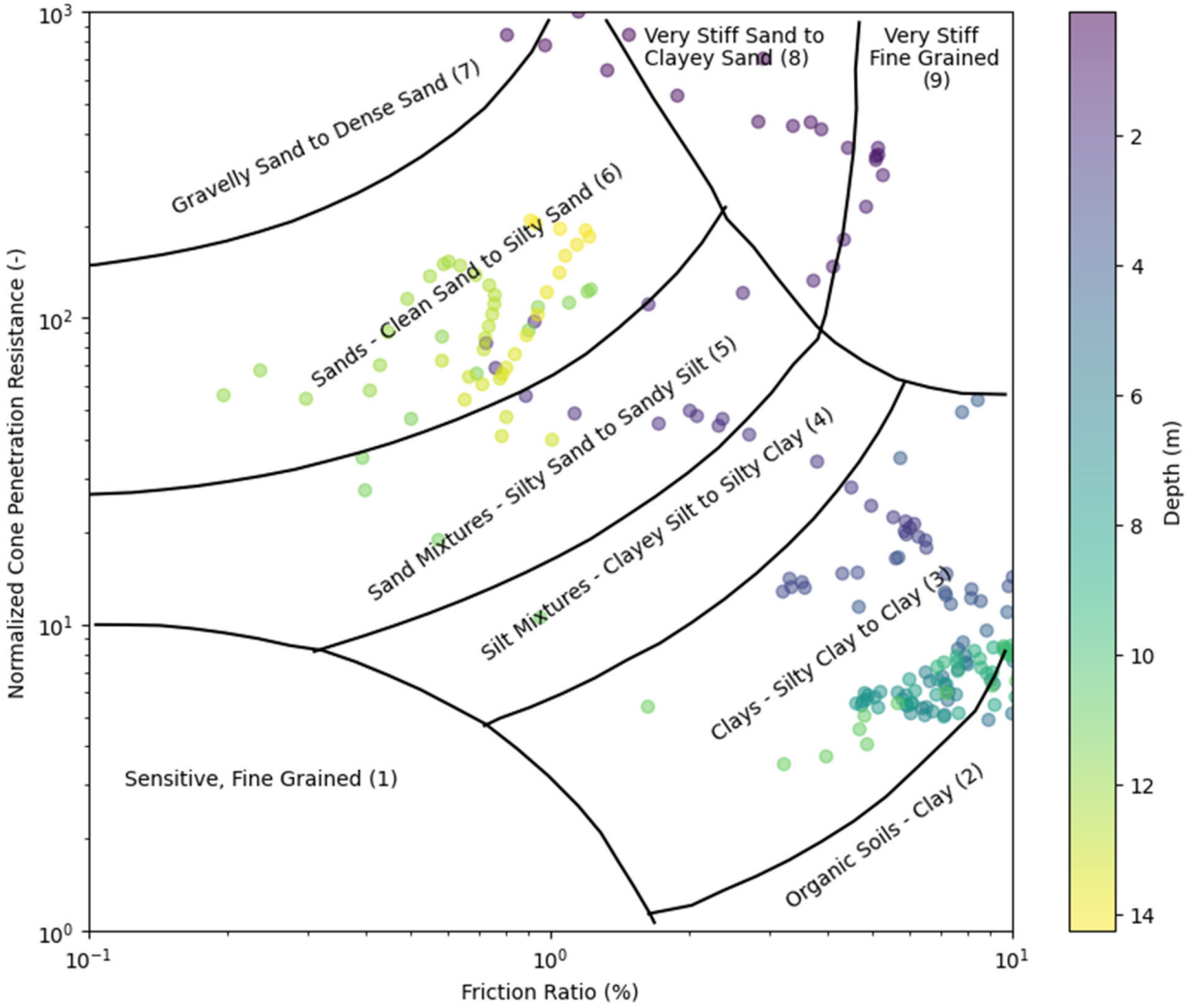

Norm. Soil Behaviour Type
Depth (m)
Very stiff sand/clayey sand
Very stiff sand/clayey sand
Clay
Clay
Clay
Clay
Organic soil
Organic soil
Clay
Clay & silty clay
Sand & silty sand
Very stiff fine grained
Sand & silty sand
Very stiff sand/clayey sand
Silty sand & sandy silt
Clay
Clay & silty clay
Silty sand & sandy silt
Silty sand & sandy silt
Sand & silty sand
Silty sand & sandy silt
Sand & silty sand
Sand & silty sand
Sand & silty sand
Silty sand & sandy silt
Clay & silty clay
Clay
0
2
4
6
8
10
12
14
16
18
SBTn (Robertson, 1990)


Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)
Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)

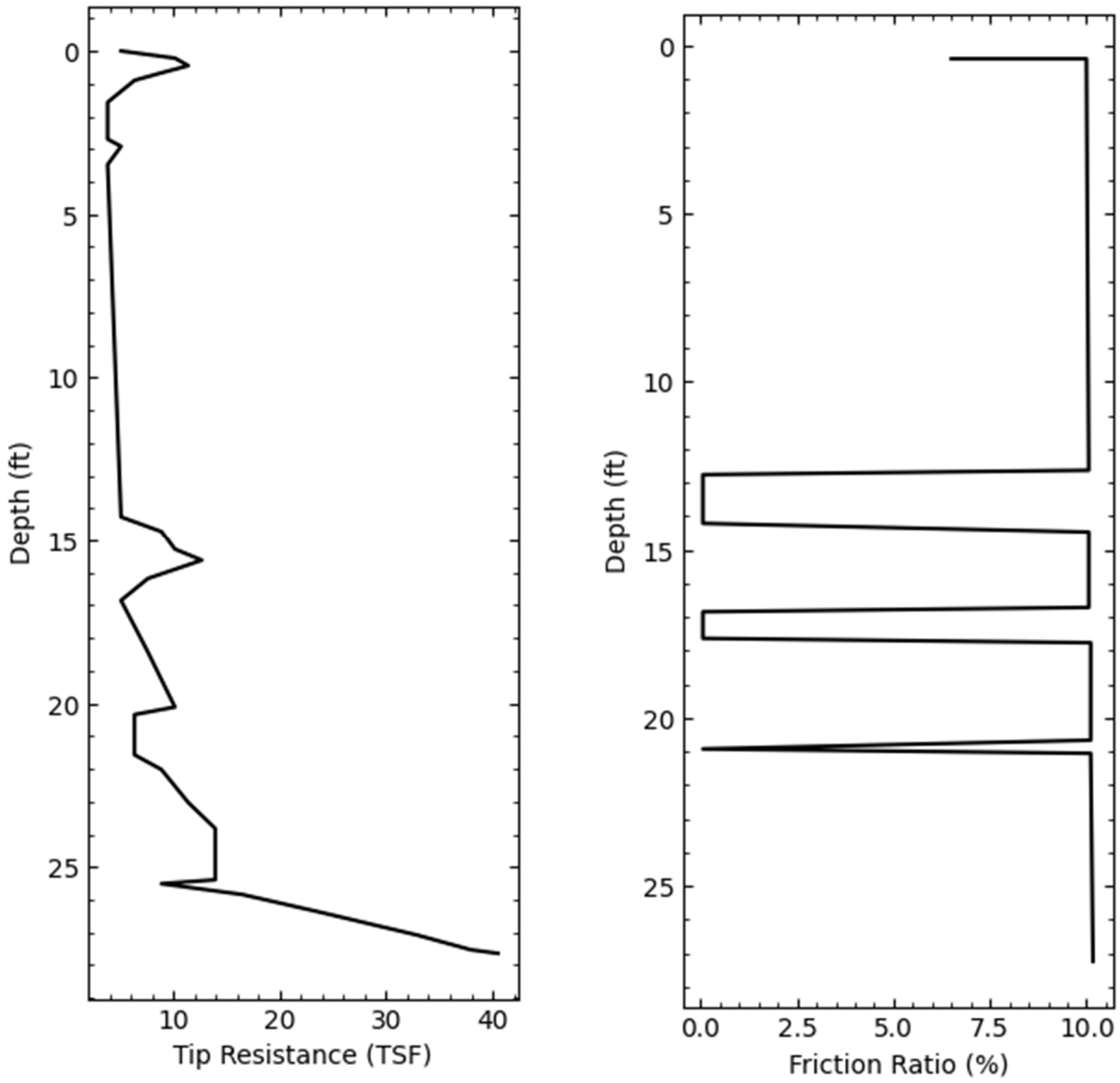
DWP-BI-BA8-CPT

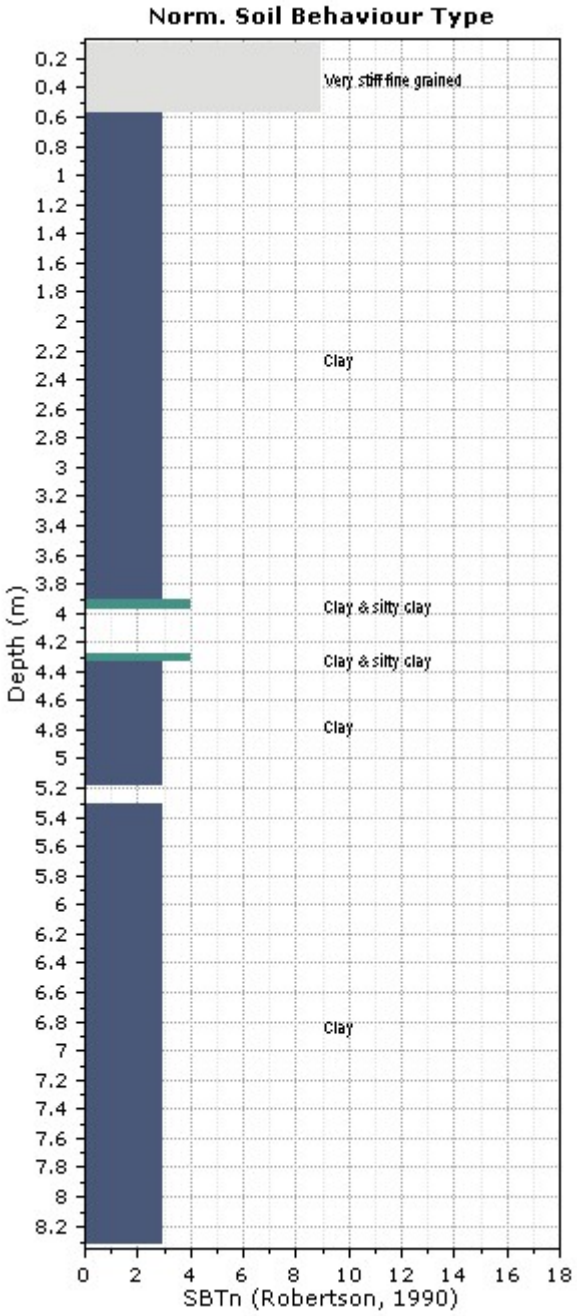


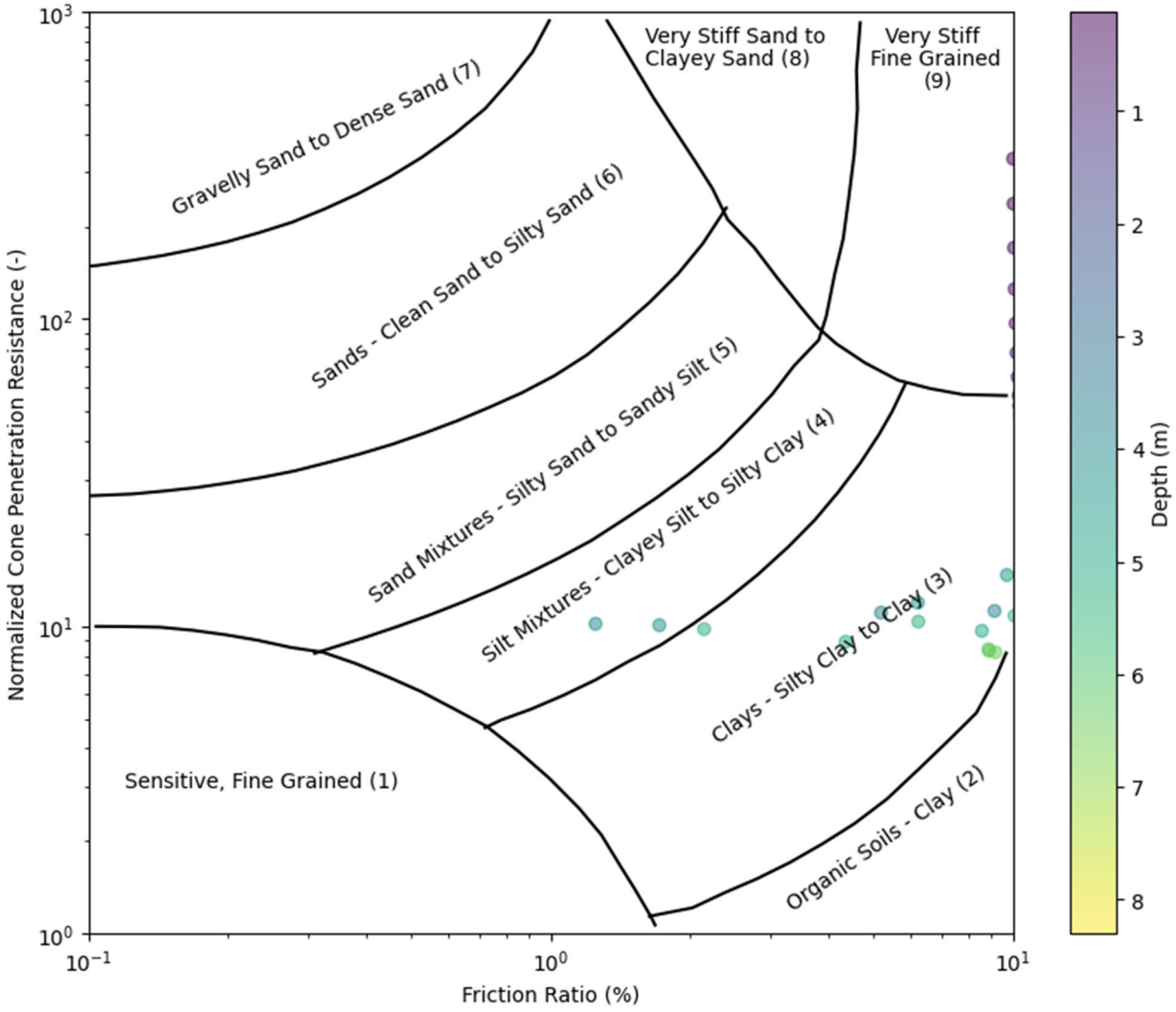

Norm. Soil Behaviour Type
Depth (m)
SBTn (Robertson, 1990)
Very stiff sand/clayey sand
Sand & silty sand
Very stiff sand/clayey sand
Very stiff fine grained
Very stiff sand/clayey sand
Sand & silty sand
Silty sand & sandy silt
Clay & silty clay
Clay
Organic soil
Clay
Organic soil
Clay
Organic soil
Clay
Organic soil
Clay
Clay & silty clay
Sand & silty sand
Silty sand & sandy silt
Sand & silty sand


Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)
Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)

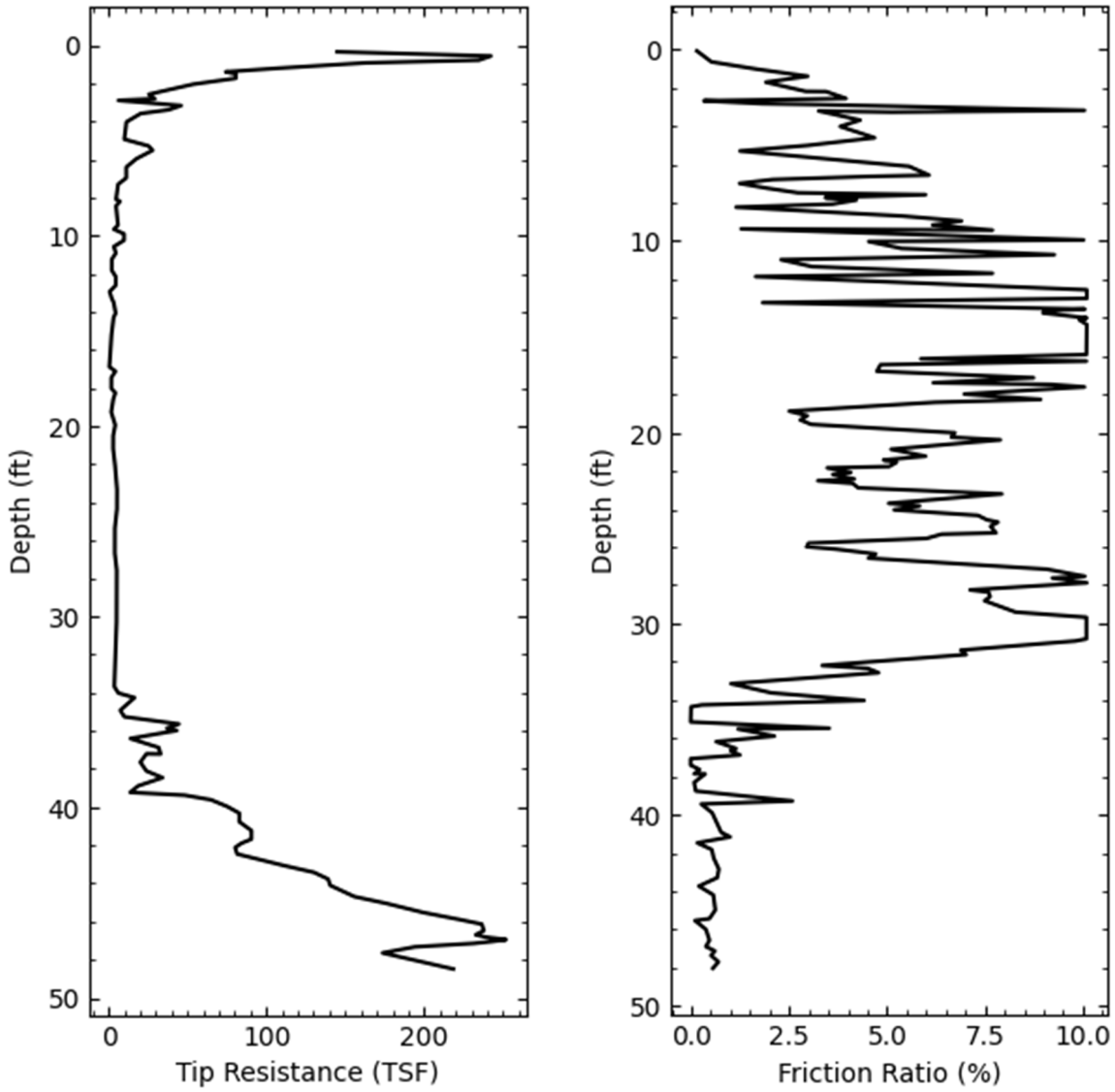
DWP-BI-BA9-CPT

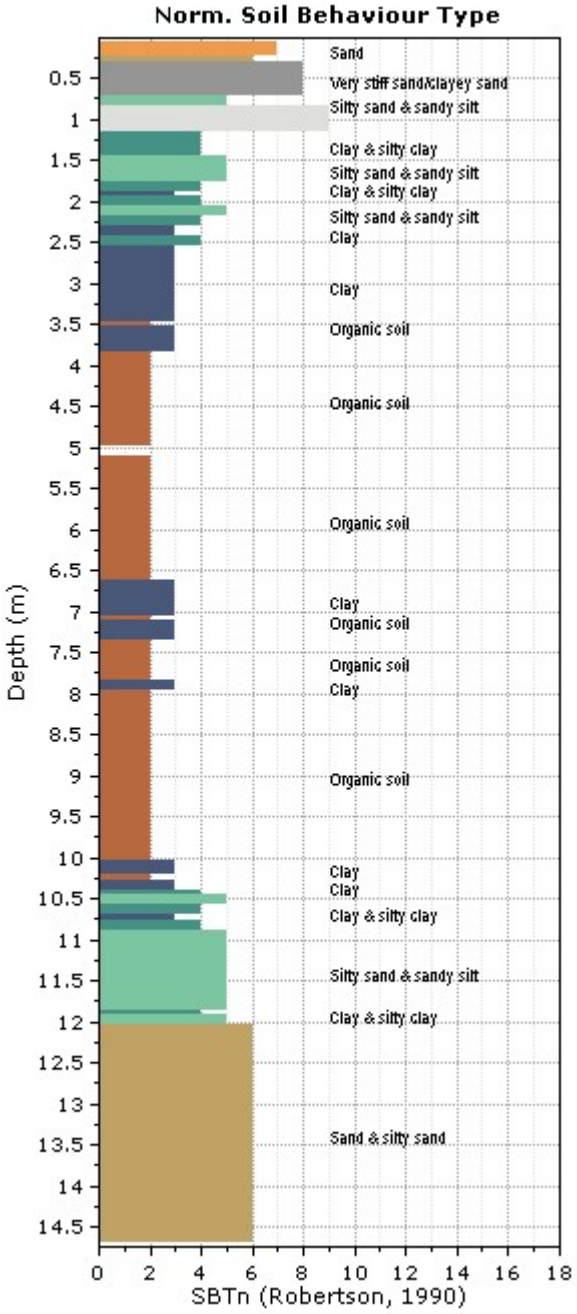


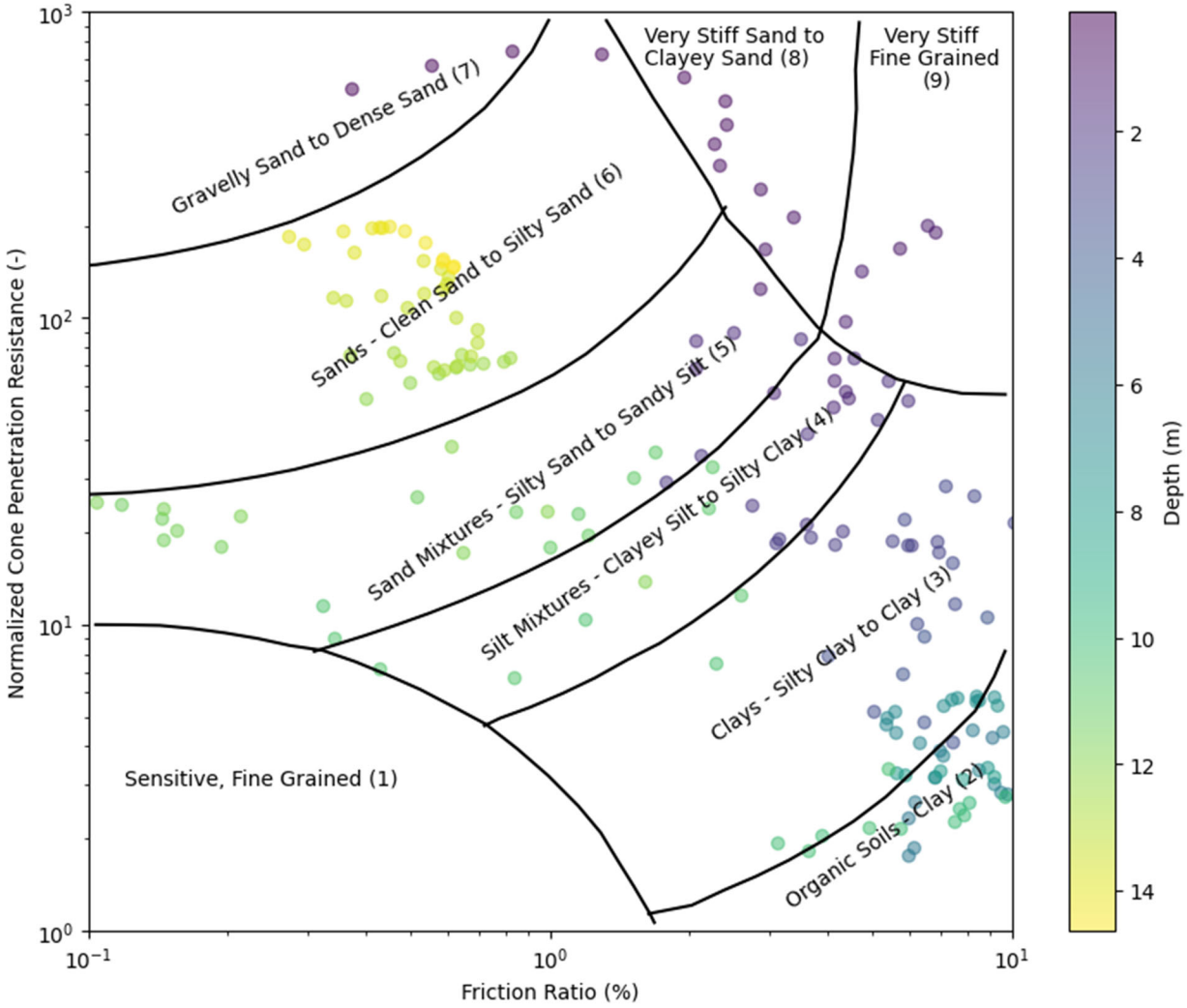

Norm. Soil Behaviour Type
Very stiff fine grained
Clay
Clay & silty clay
Clay & silty clay
Clay
Clay
Depth (m)
SBTn (Robertson, 1990)


Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)
Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)

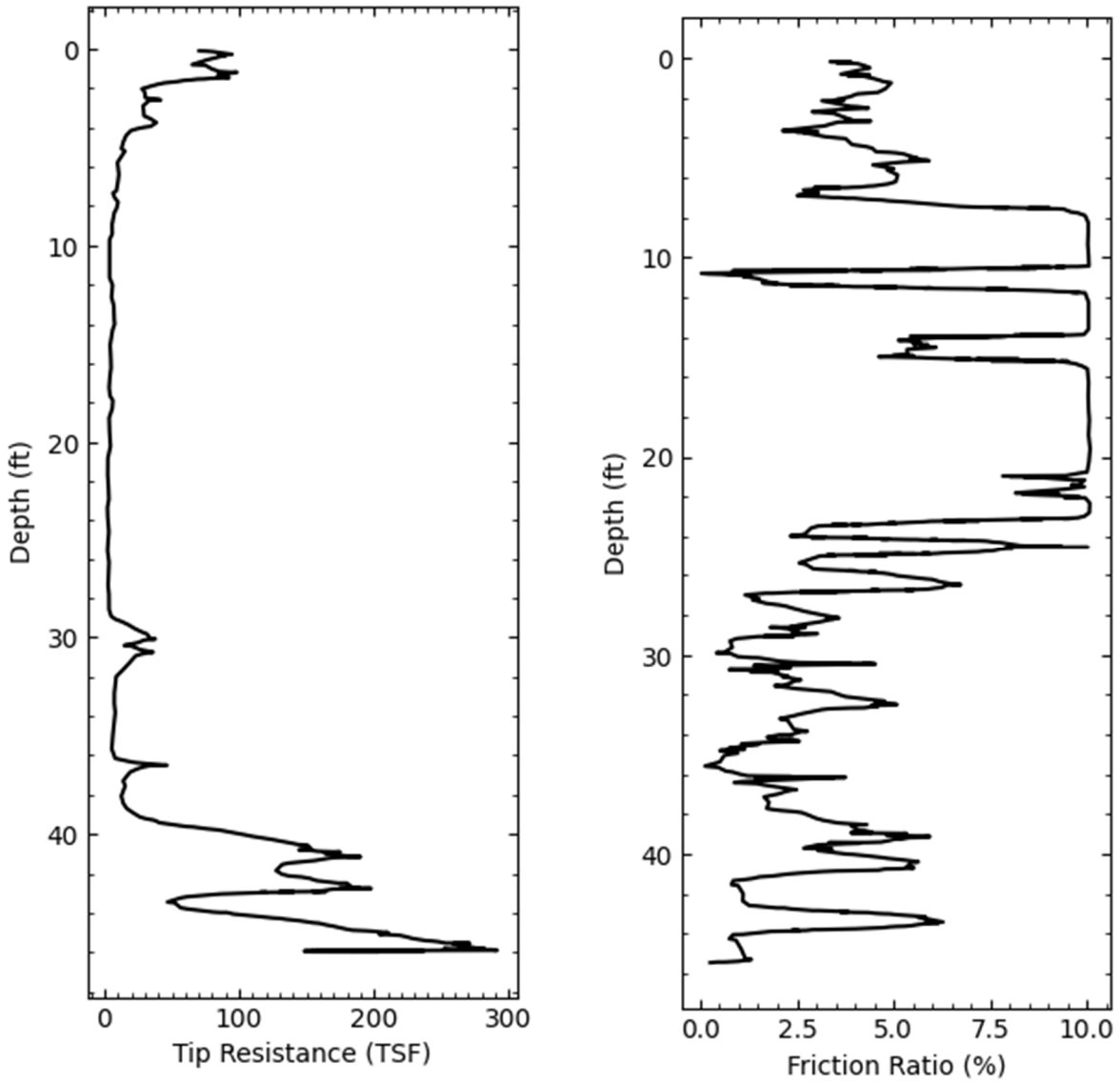
DWP-BI-BA10-CPT

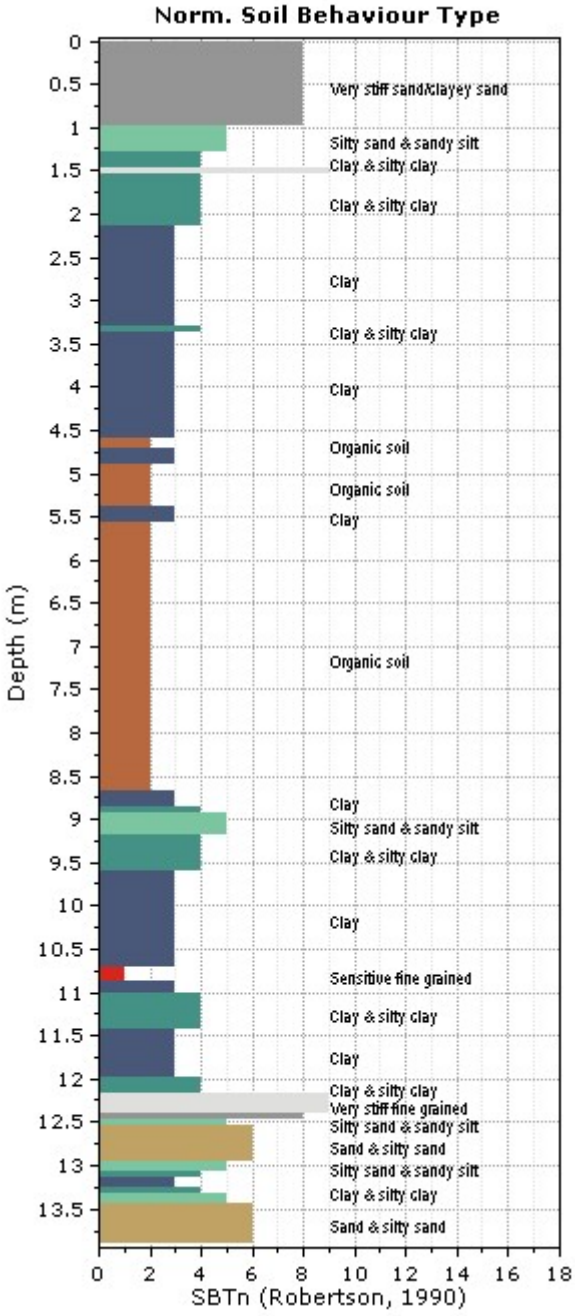


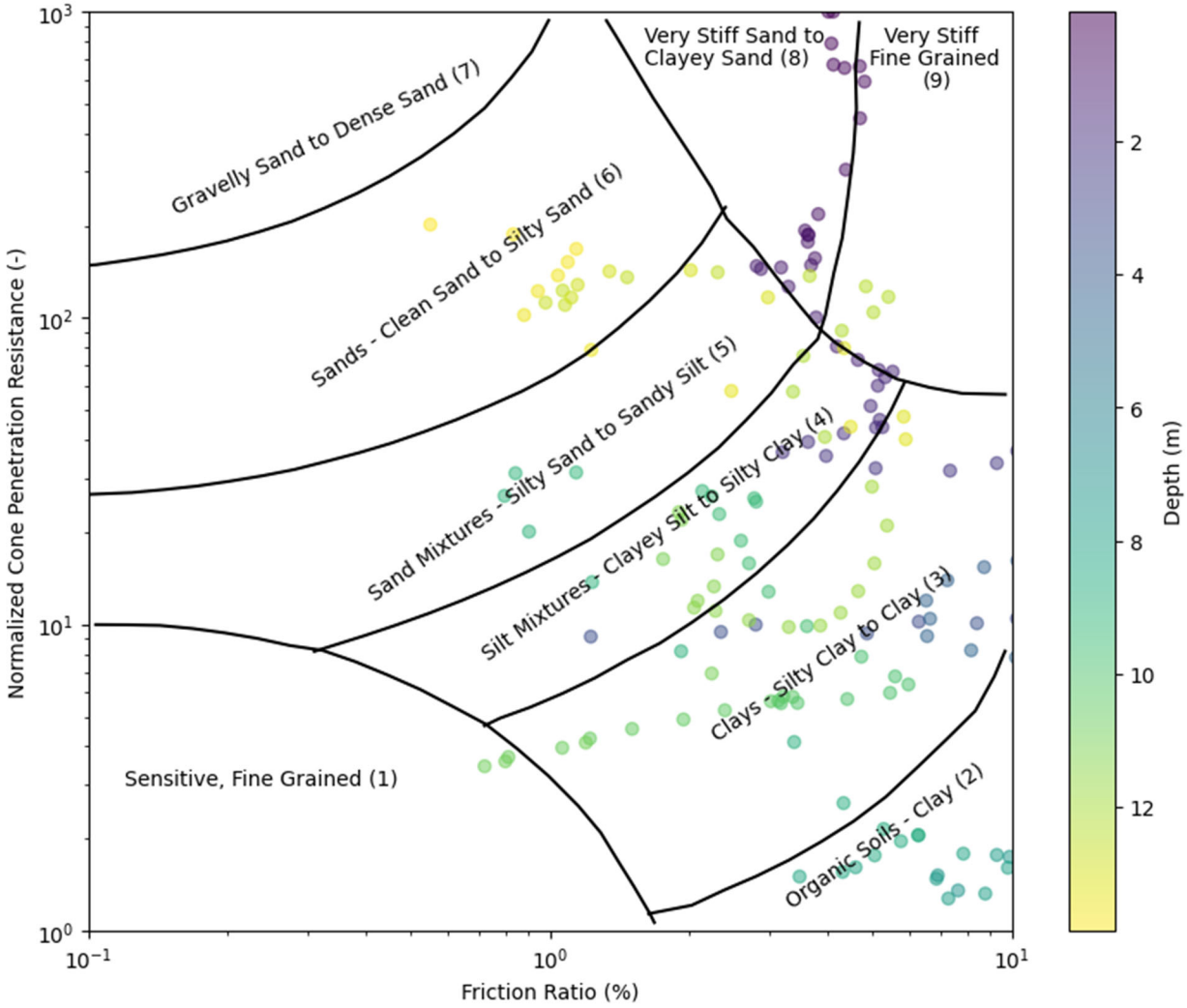

Norm. Soil Behaviour Type
Depth (m)
SBTn (Robertson, 1990)
Sand
Very stiff sand/clayey sand
Silty sand & sandy silt
Clay & silty clay
Silty sand & sandy silt
Clay & silty clay
Silty sand & sandy silt
Clay
Clay
Organic soil
Organic soil
Organic soil
Clay
Organic soil
Organic soil
Clay
Organic soil
Clay
Clay
Clay & silty clay
Silty sand & sandy silt
Clay & silty clay
Sand & silty sand


Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)
Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)
10^3
10^2
10^1
10^0
10^{-1}
10^0
10^1
2
4
6
8
10
12
14

DWP-BI-BA19-CPT



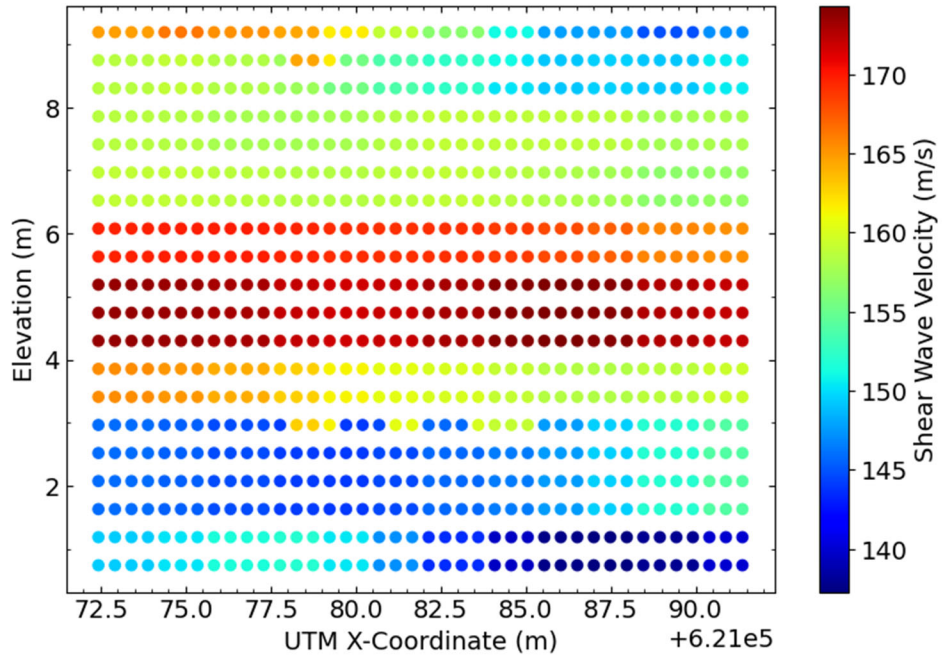

Norm. Soil Behaviour Type
Depth (m)
SBTn (Robertson, 1990)
Very stiff sand/clayey sand
Silty sand & sandy silt
Clay & silty clay
Clay & silty clay
Clay
Clay & silty clay
Clay
Organic soil
Organic soil
Clay
Organic soil
Clay
Silty sand & sandy silt
Clay & silty clay
Clay
Sensitive fine grained
Clay & silty clay
Clay
Clay & silty clay
Very stiff fine grained
Silty sand & sandy silt
Sand & silty sand
Silty sand & sandy silt
Clay & silty clay
Sand & silty sand

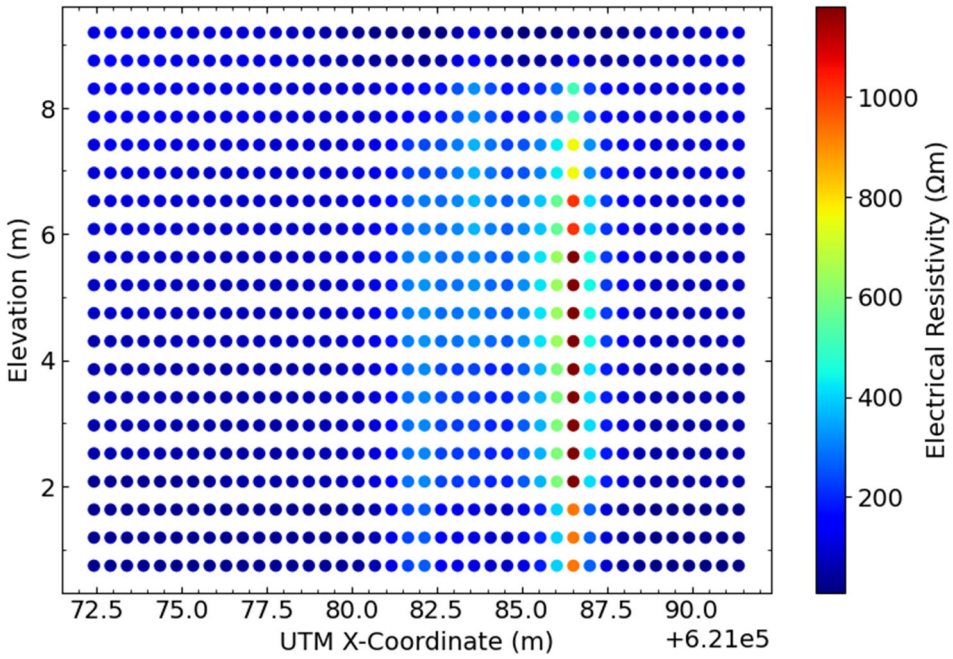

Gravelly Sand to Dense Sand (7)
Very Stiff Sand to Clayey Sand (8)
Very Stiff Fine Grained (9)
Sands - Clean Sand to Silty Sand (6)
Sand Mixtures - Silty Sand to Sandy Silt (5)
Silt Mixtures - Clayey Silt to Silty Clay (4)
Clays - Silty Clay to Clay (3)
Sensitive, Fine Grained (1)
Organic Soils - Clay (2)
Normalized Cone Penetration Resistance (-)
Friction Ratio (%)
Depth (m)
10^3
10^2
10^1
10^0
10^{-1}
10^0
10^1
2
4
6
8
10
12

Appendix C – Shear Wave Velocity and Electrical Resistivity Profiles

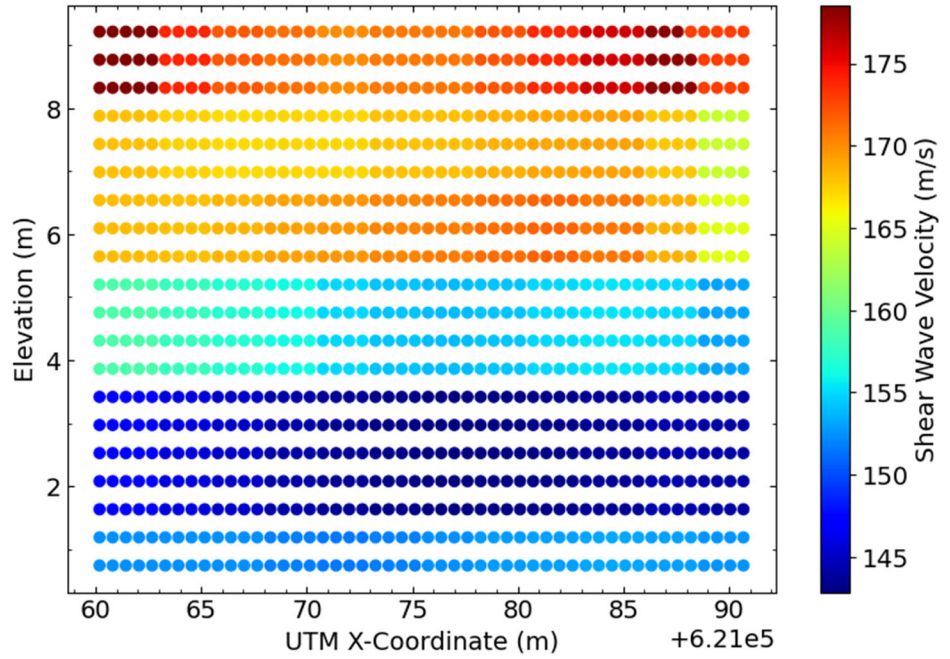
Grand Island – MASW and ERT Profiles for Machine Learning

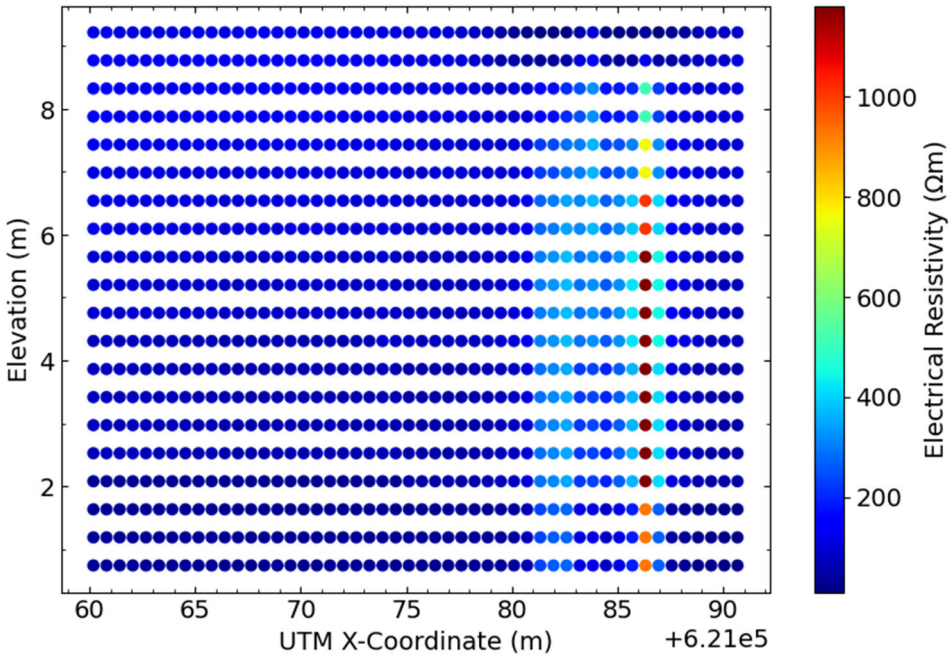
Site 1
Profile: sc3_site1



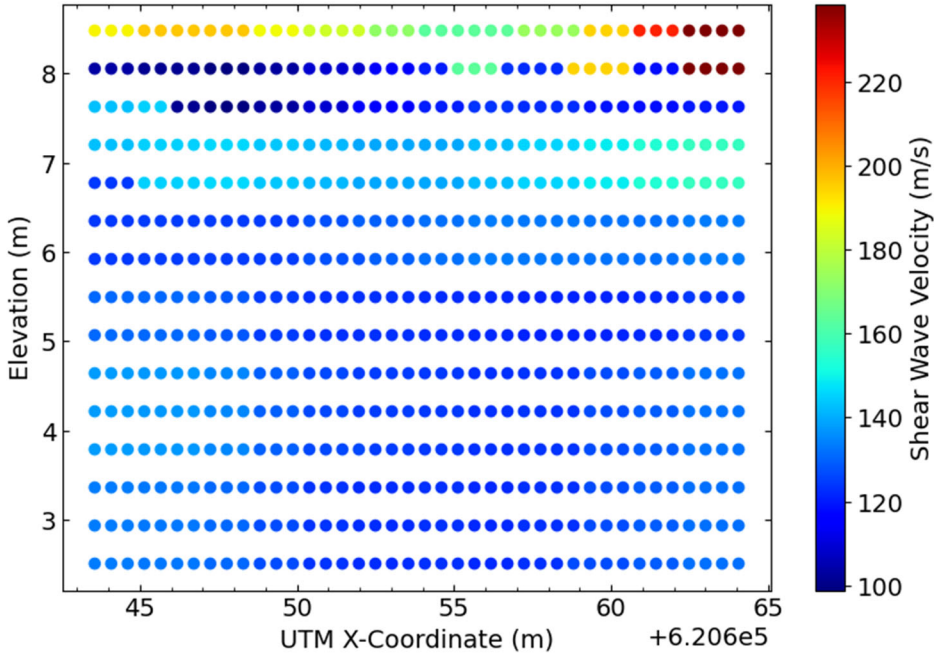


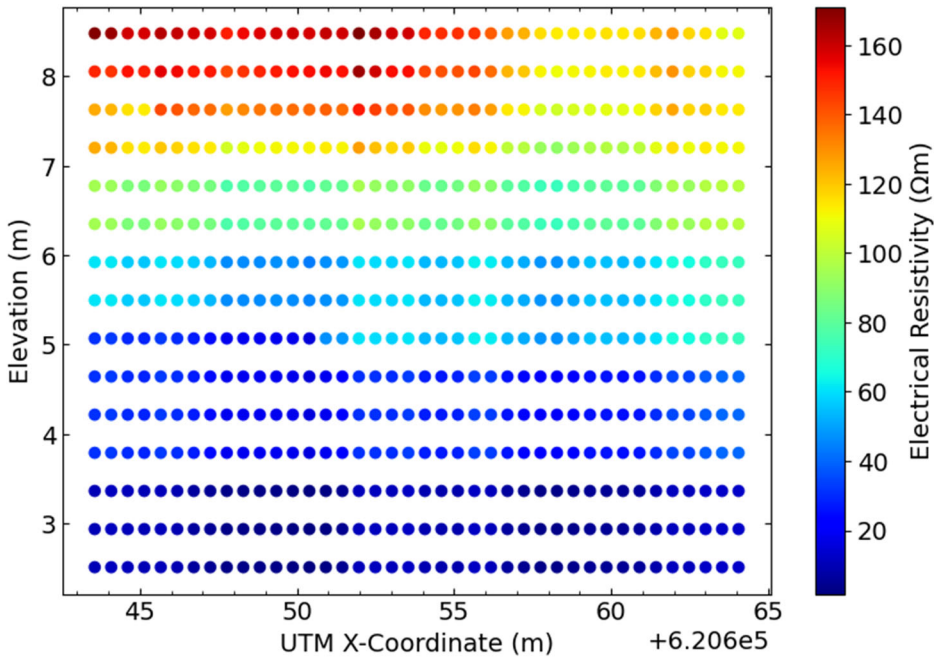
Grand Island
Site 1
Profile: sc5_site1



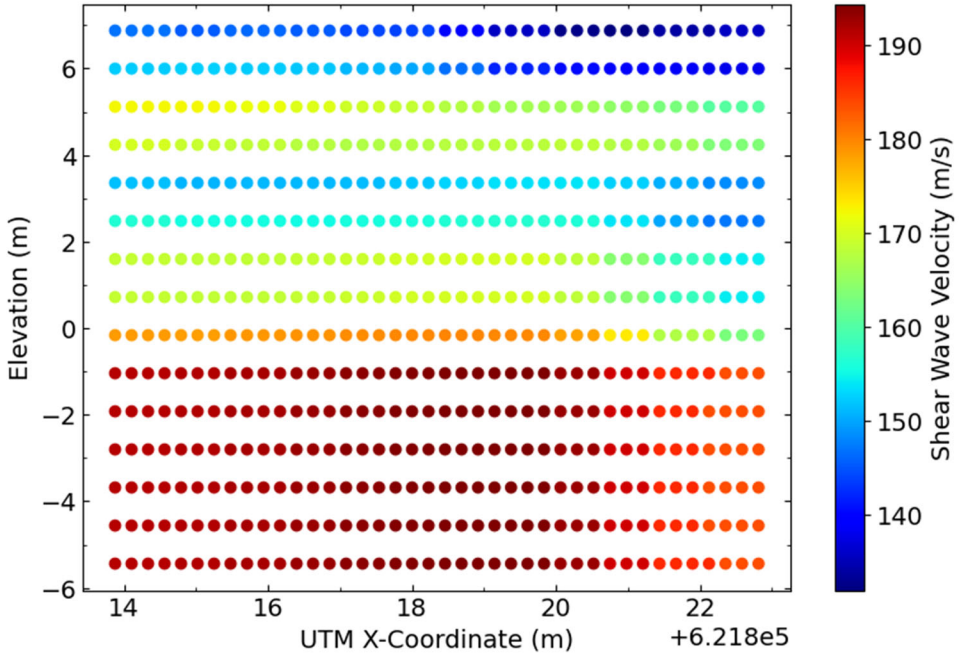


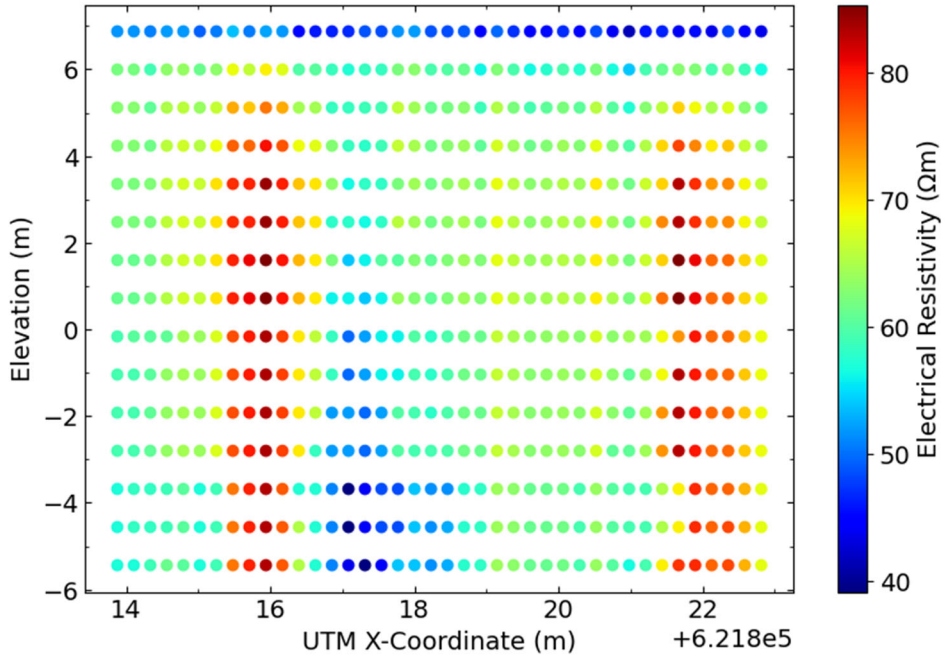
Grand Island
Site 1
Profile: swt_site1



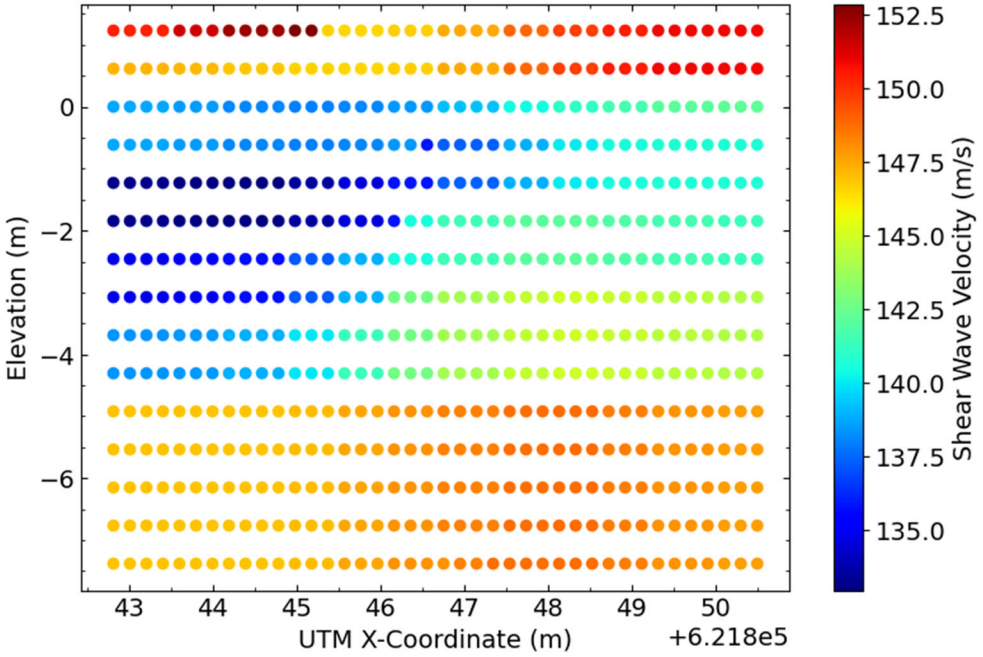


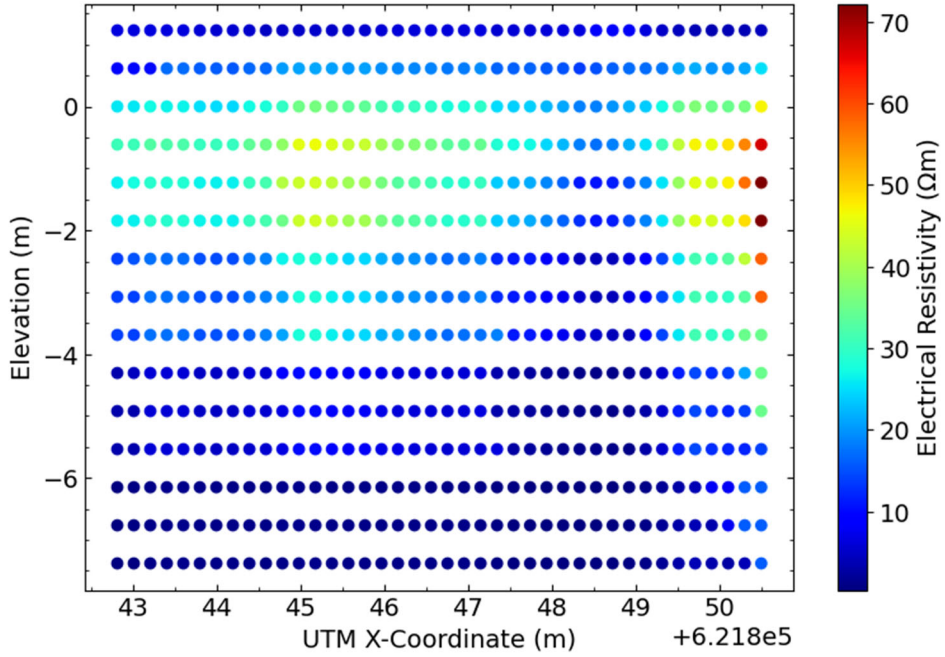
Grand Island
Site 2
Profile: sc_site2



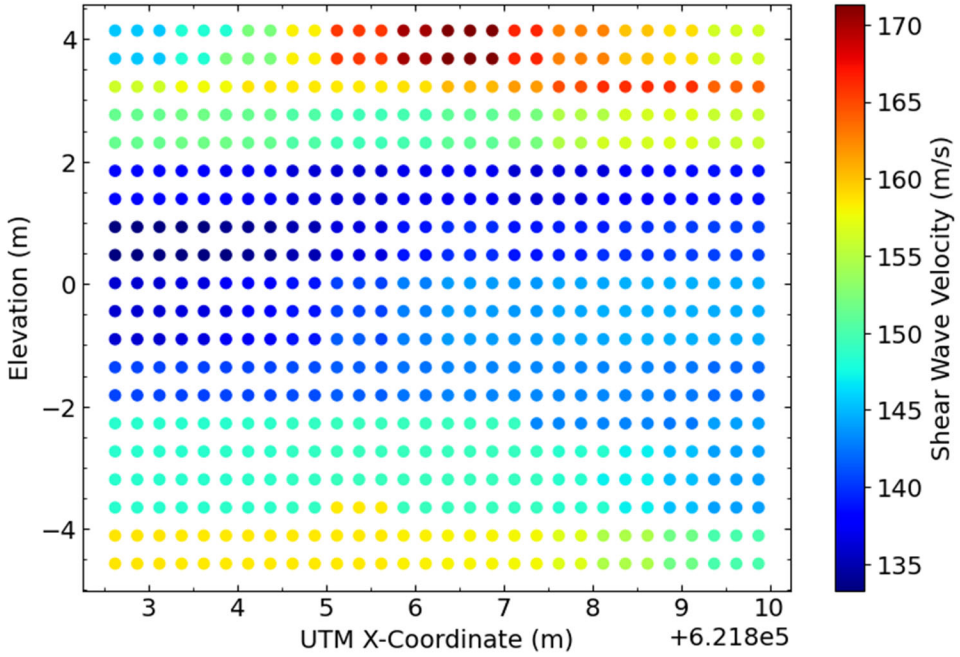


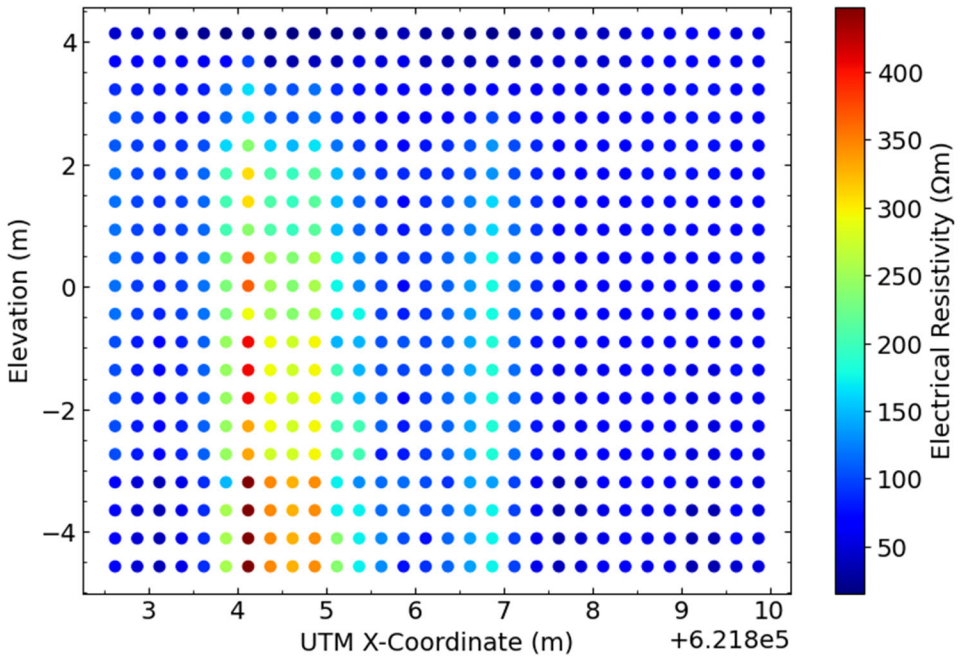
Grand Island
Site 2
Profile: sdt_site2



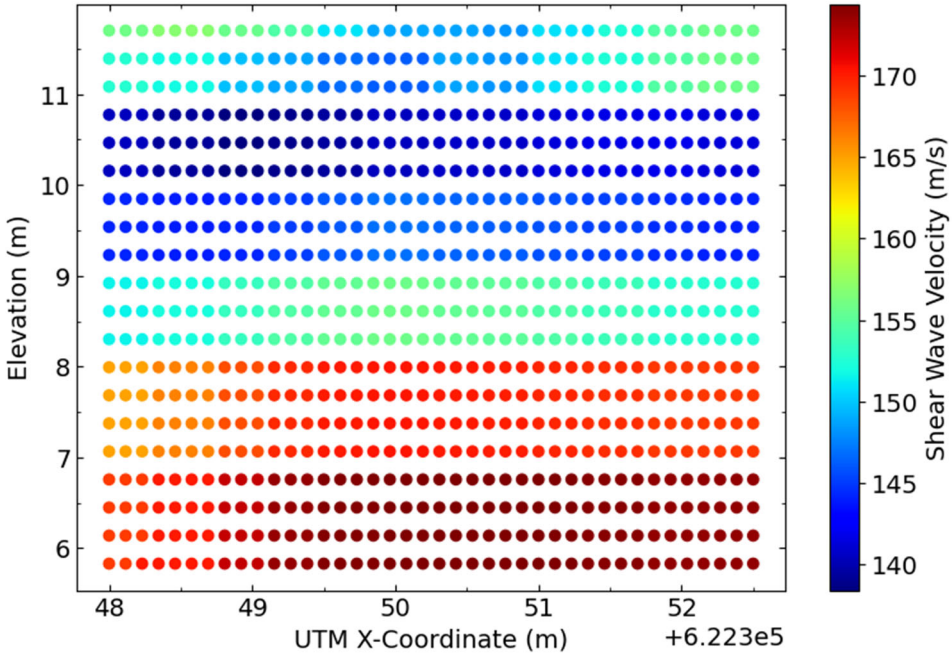


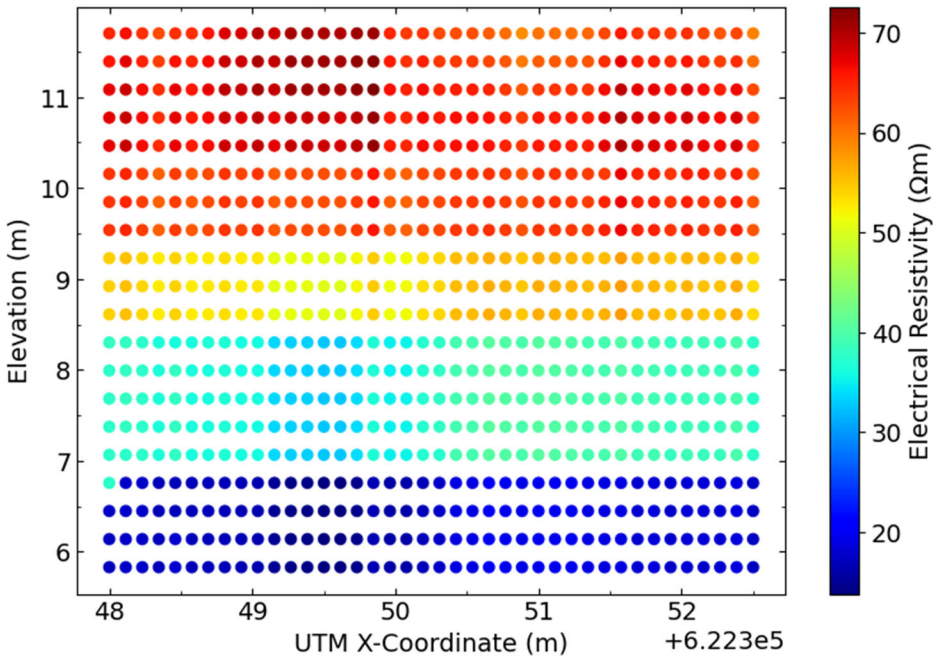
Grand Island
Site 2
Profile: swt_site2



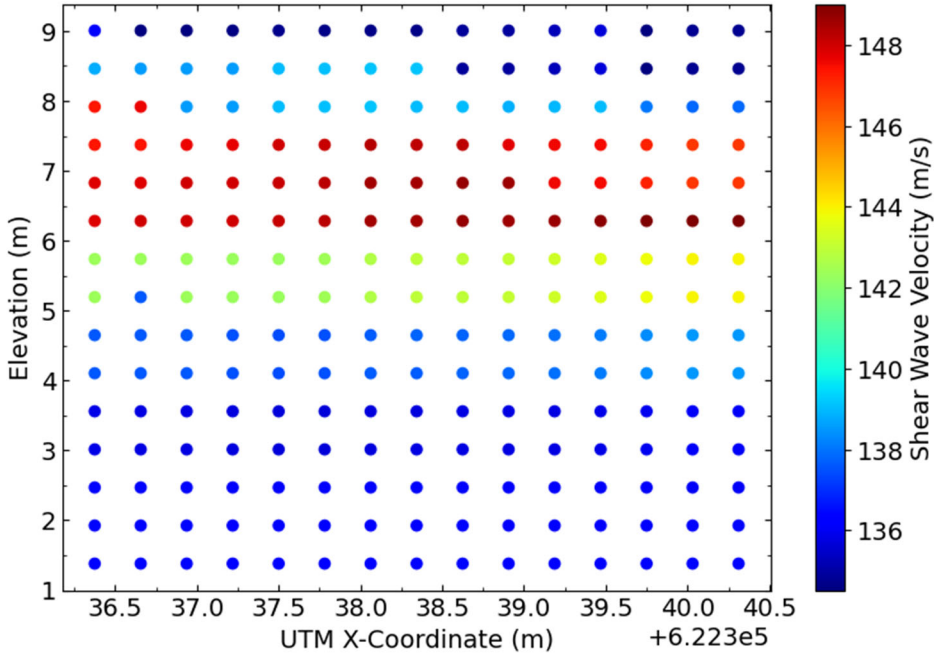


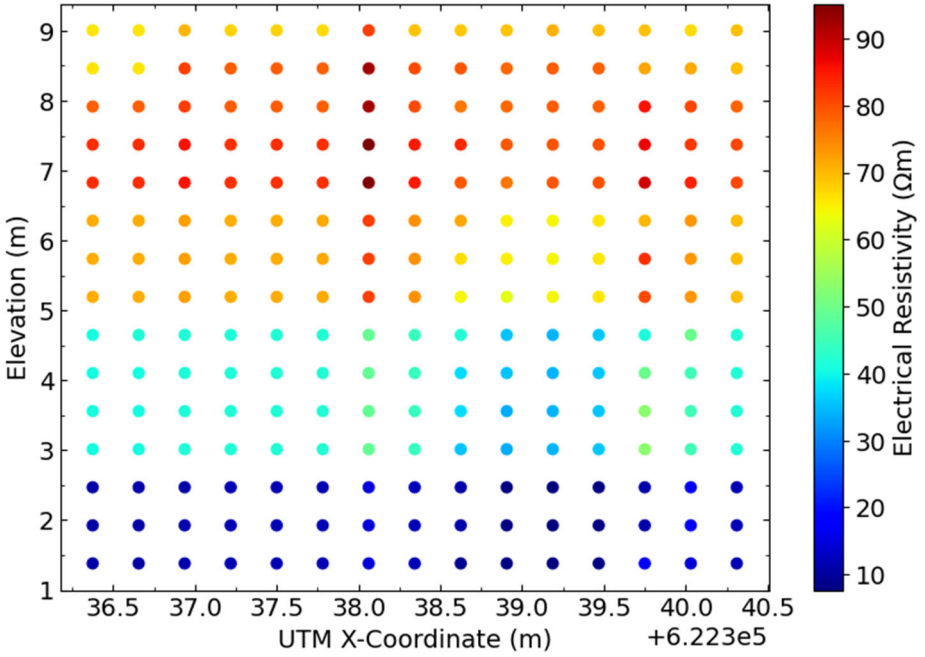
Grand Island
Site 3
Profile: sc_site3



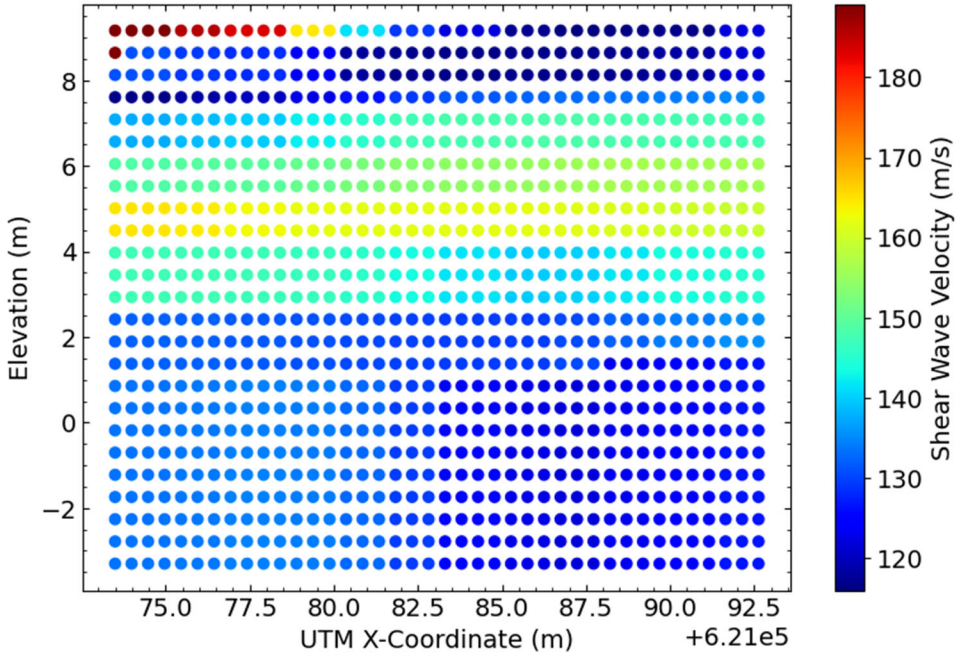


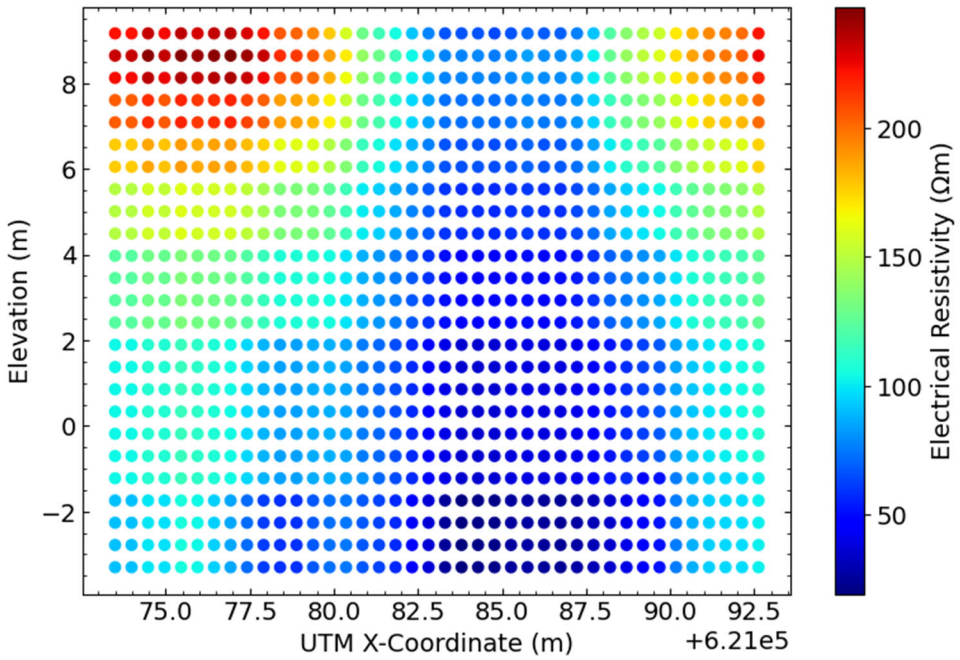
Grand Island
Site 3
Profile: swt_site3



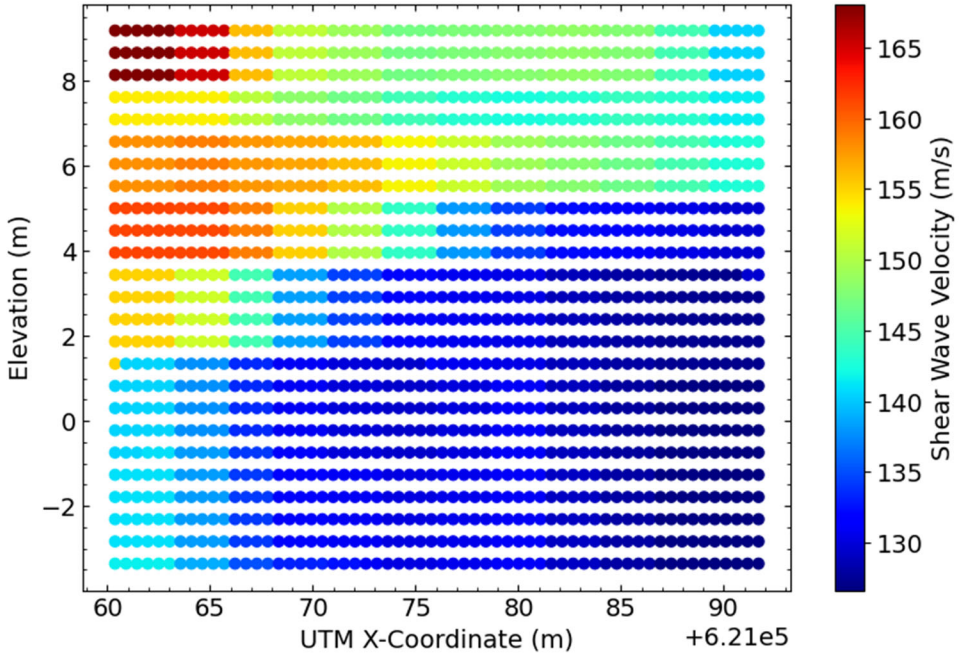


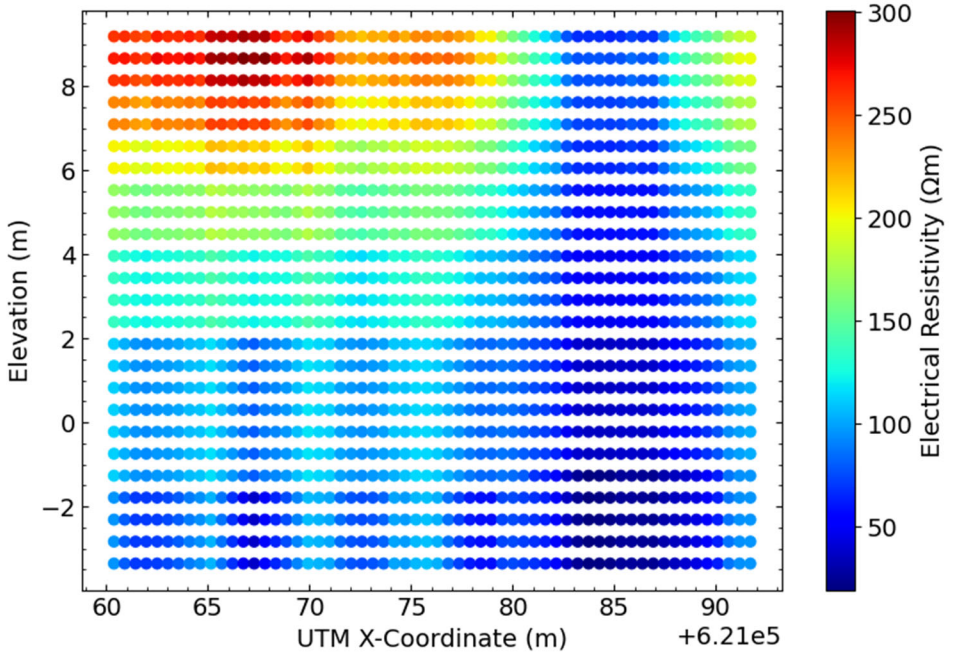
Grand Island
Site 1
Profile: fc3_site1



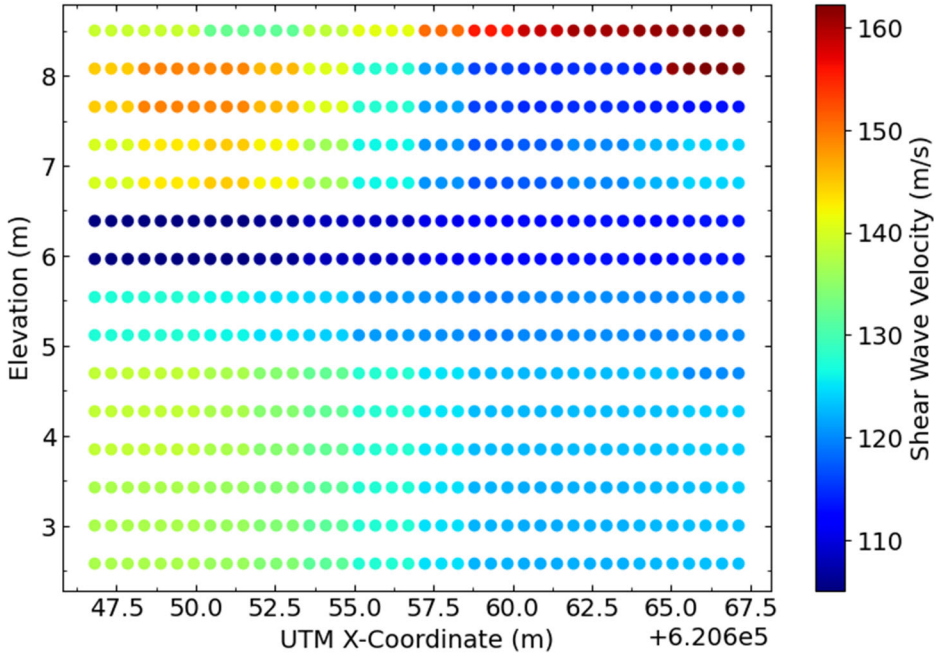


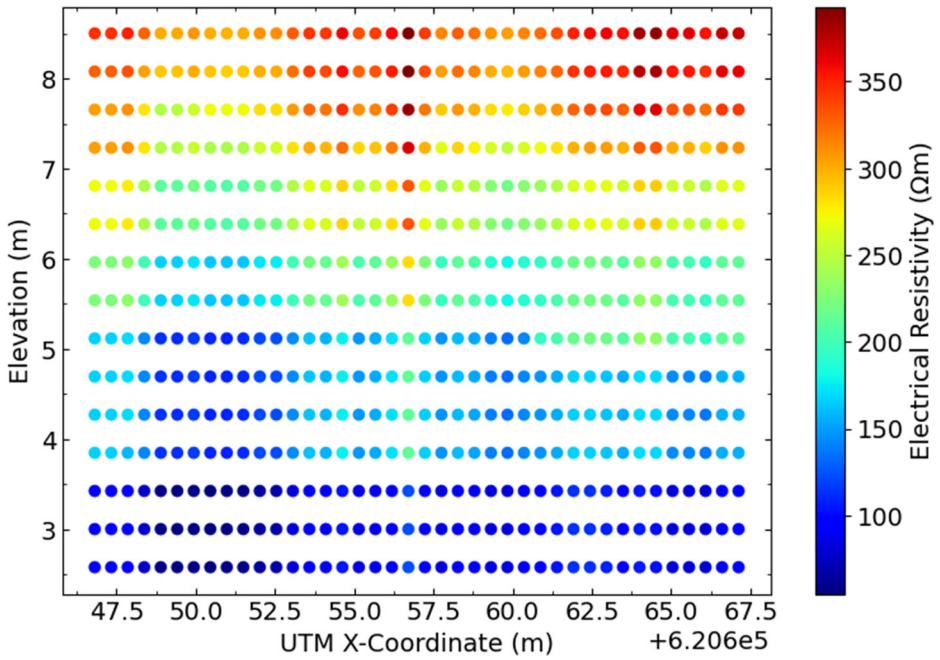
Grand Island
Site 1
Profile: fc5_site1



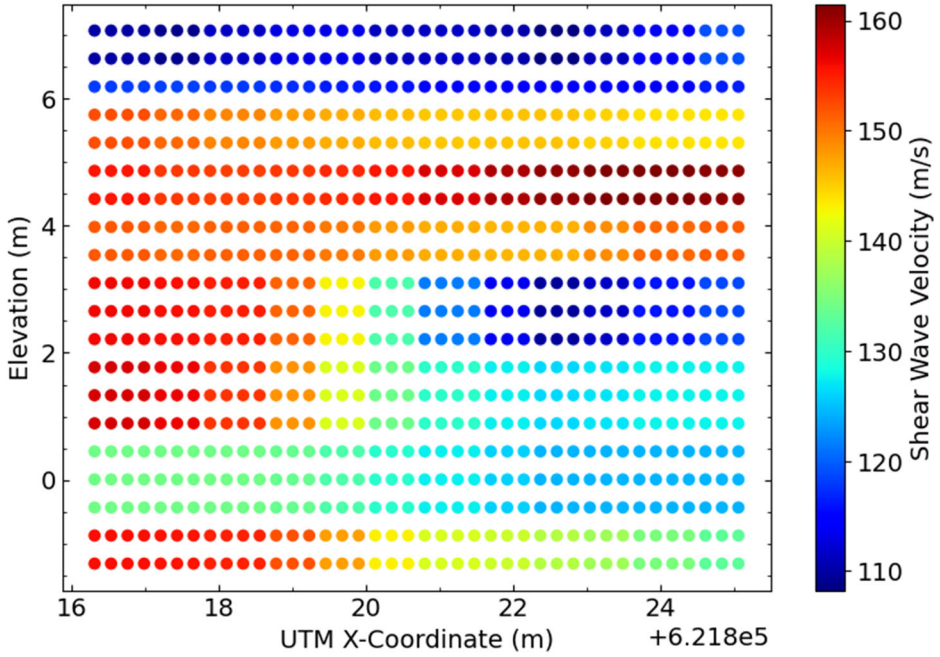


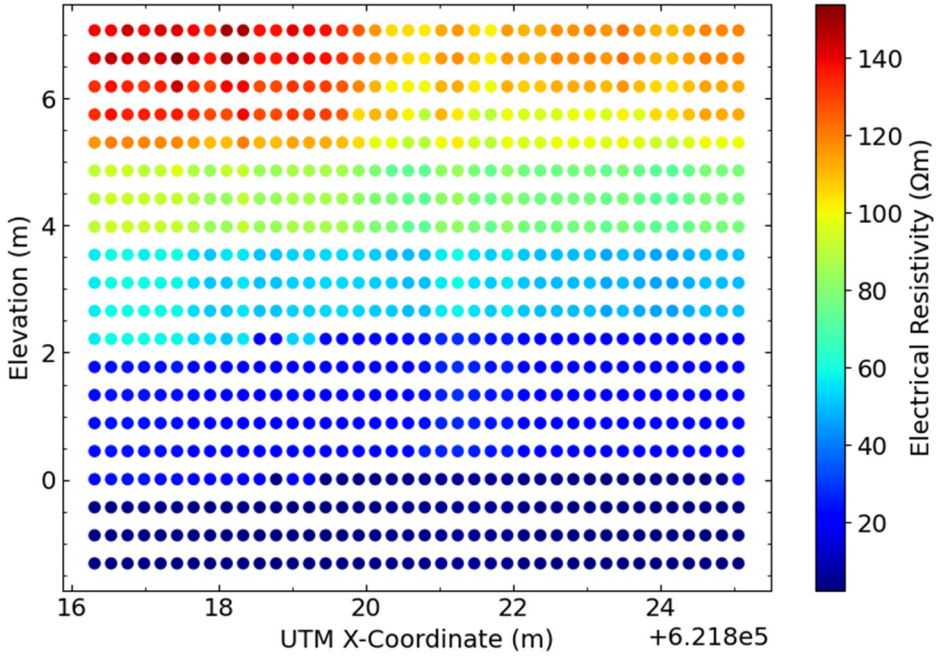
Grand Island
Site 1
Profile: fwt_site1



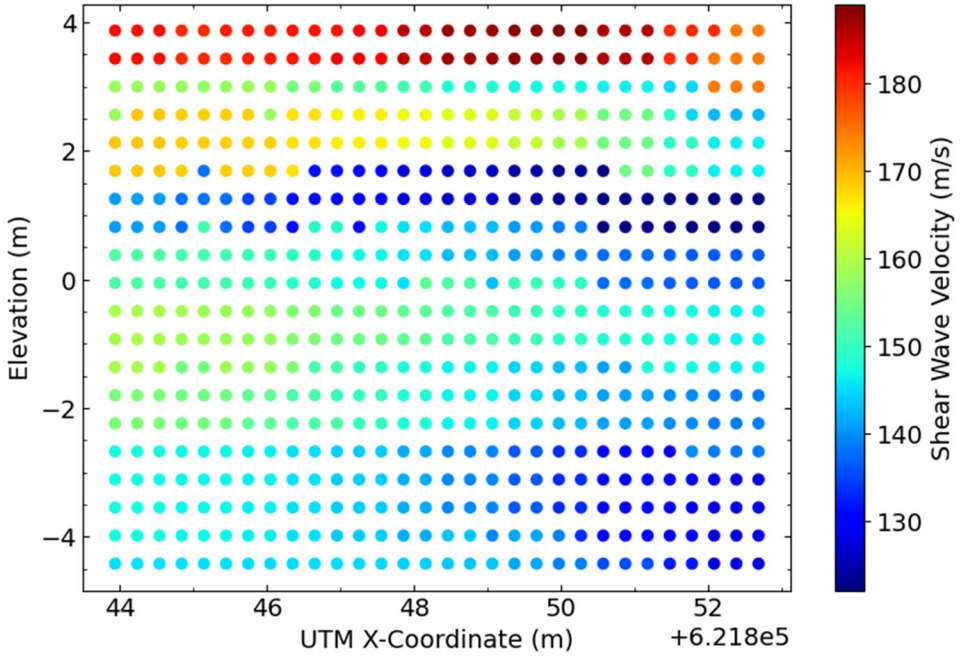


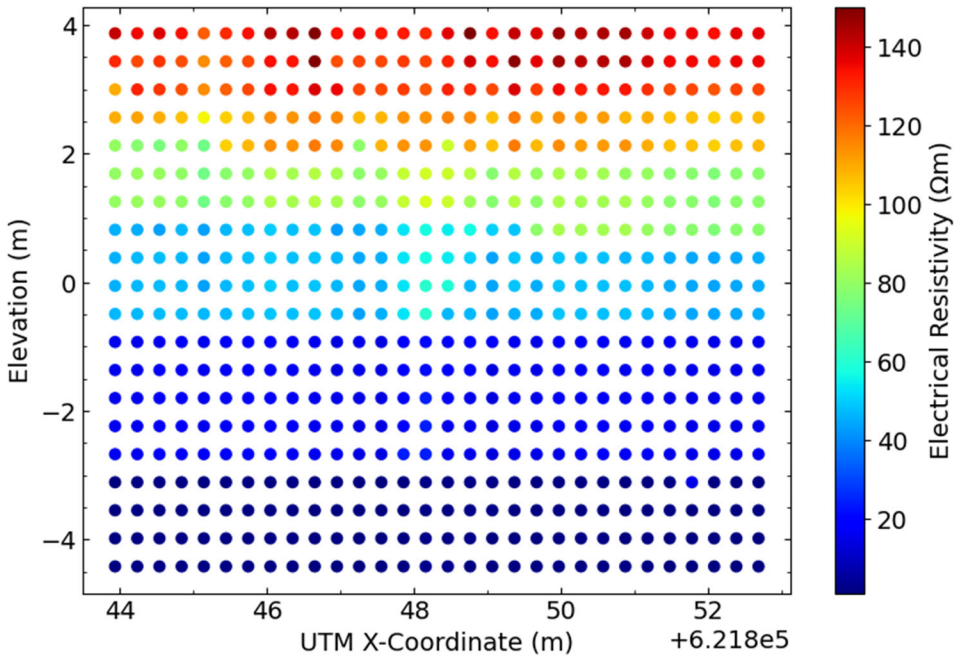
Grand Island
Site 2
Profile: fc_site2



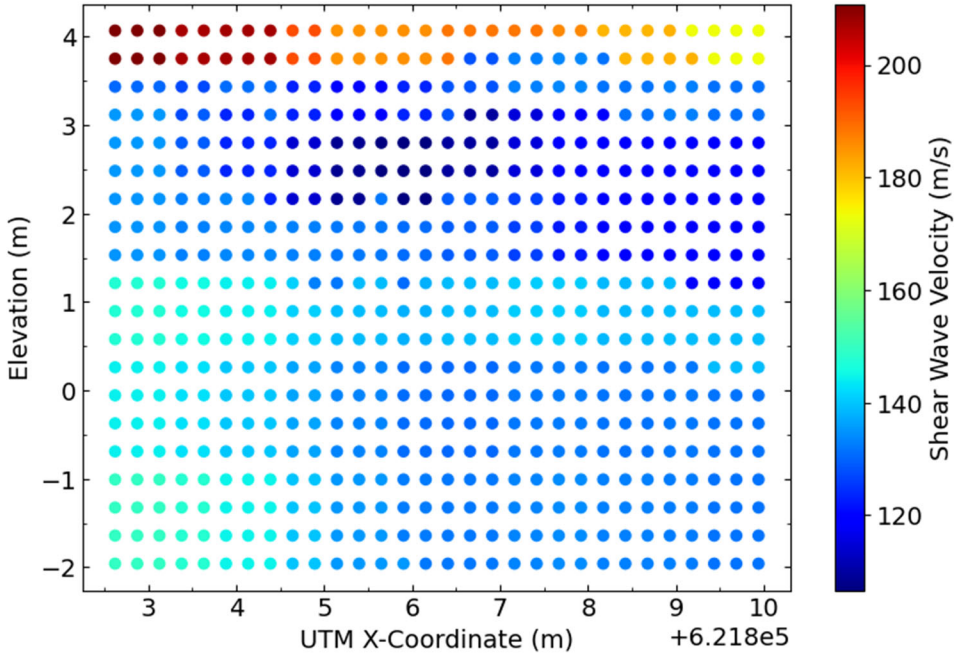


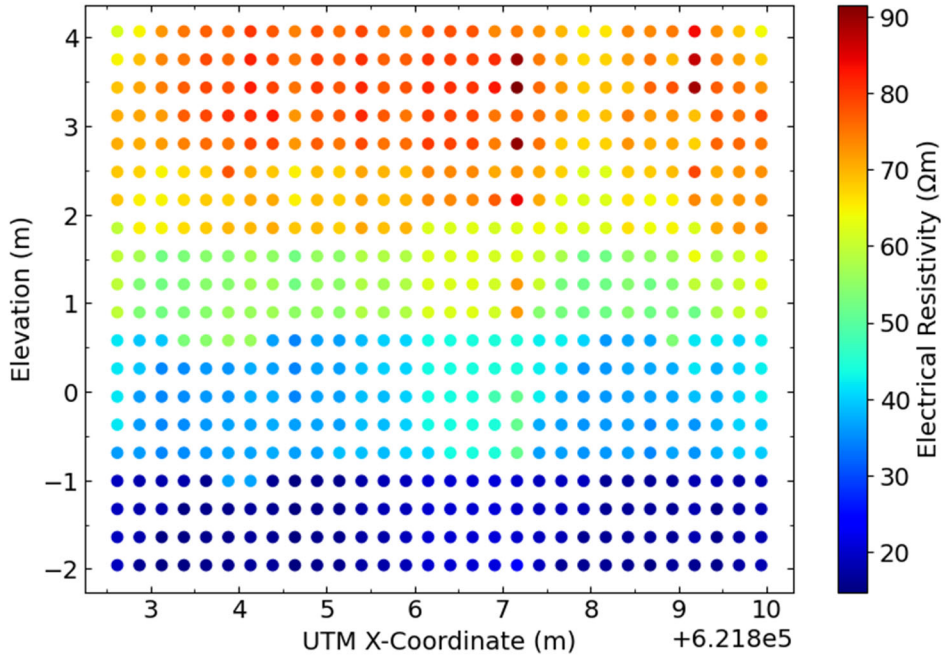
Grand Island
Site 2
Profile: fdt_site2



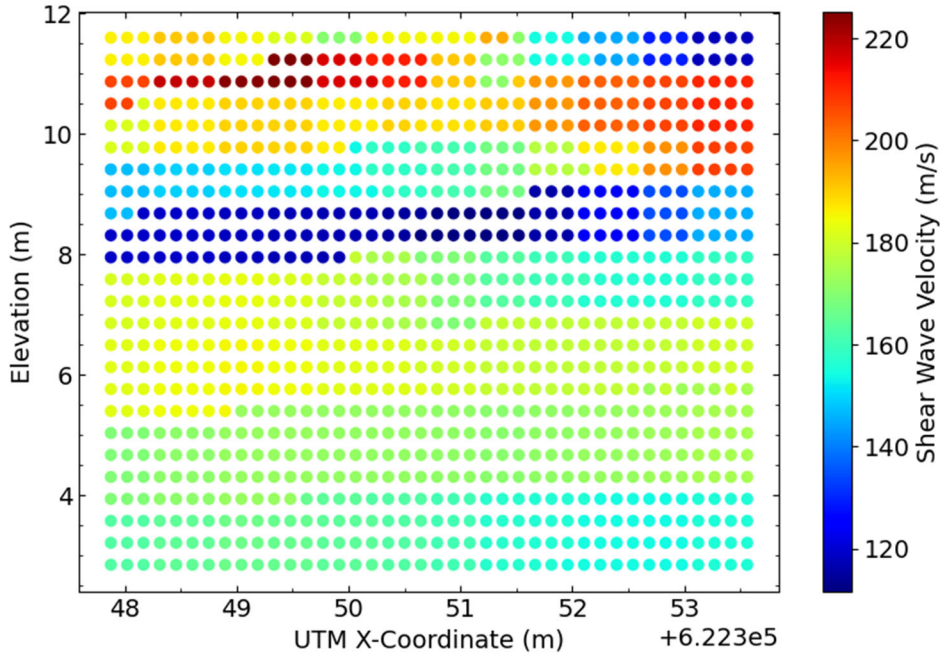


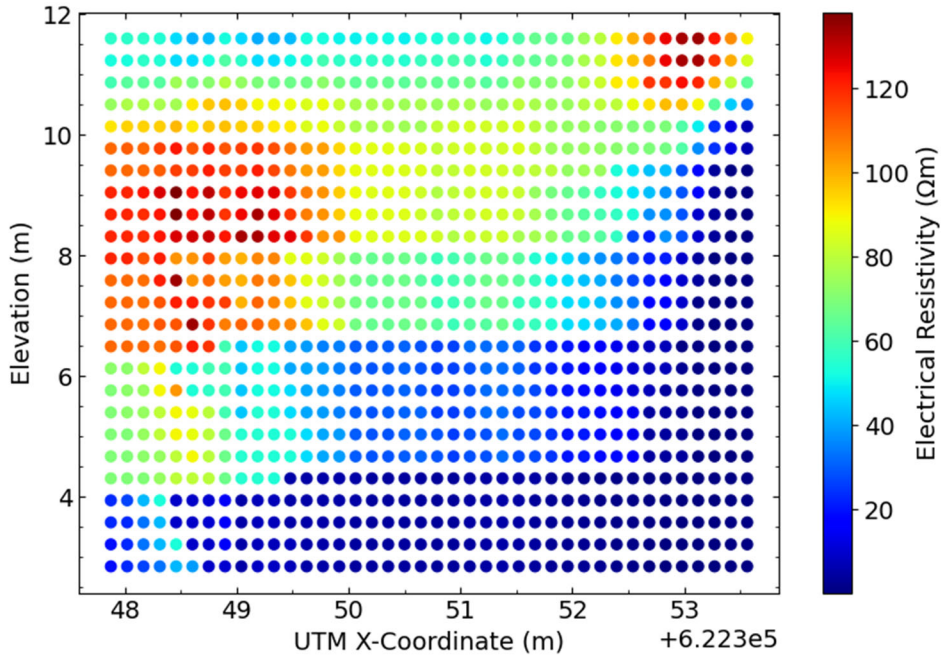
Grand Island
Site 2
Profile: fwt_site2



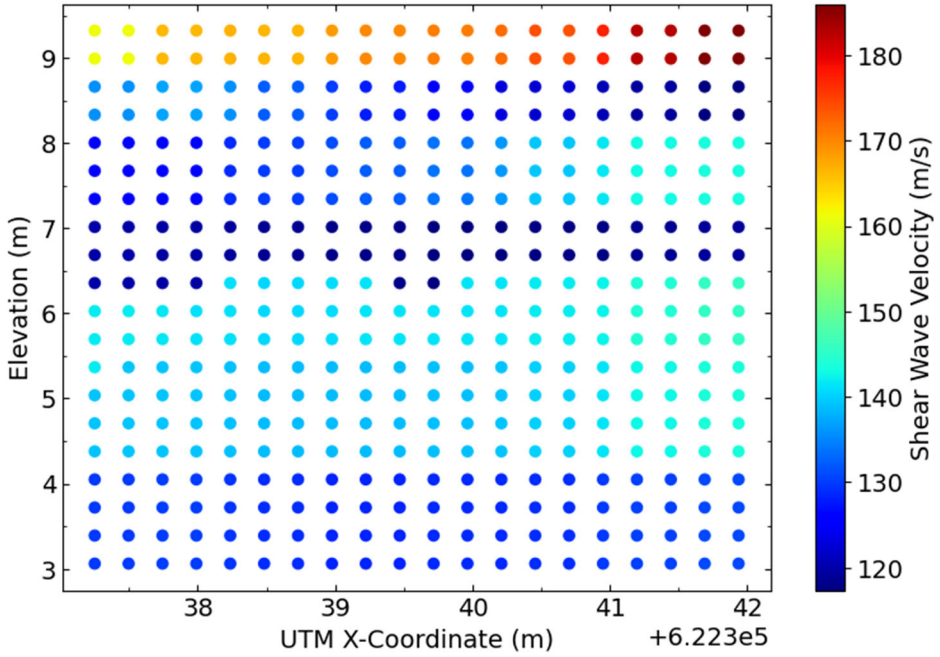


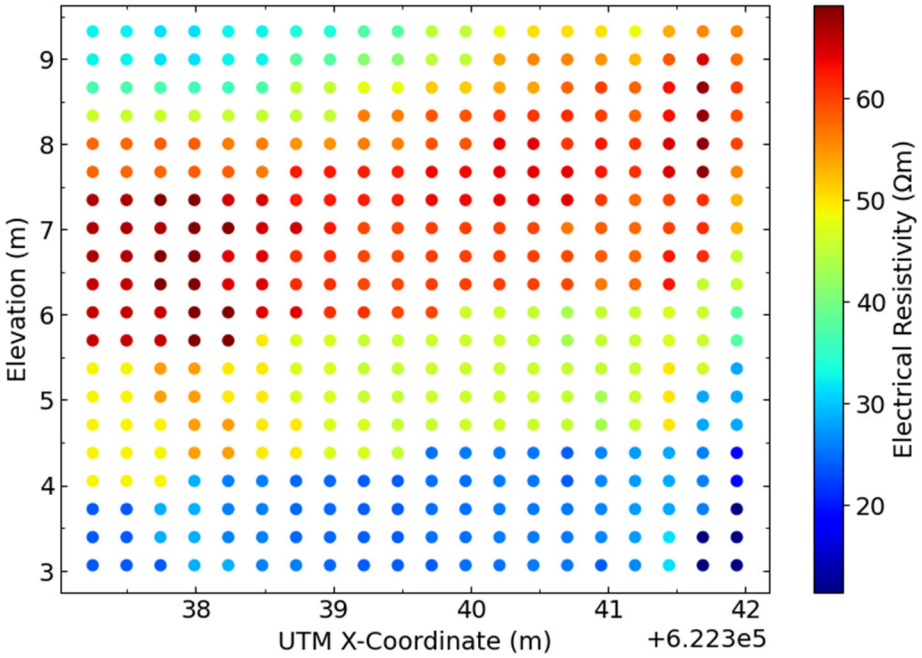
Grand Island
Site 3
Profile: fc_site3





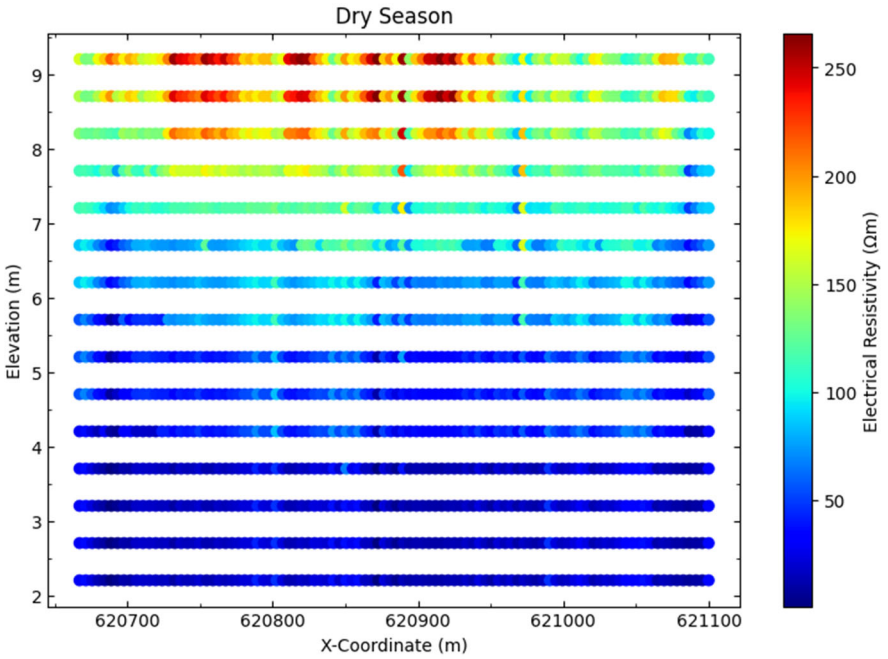
Grand Island
Site 3
Profile: fwt_site3

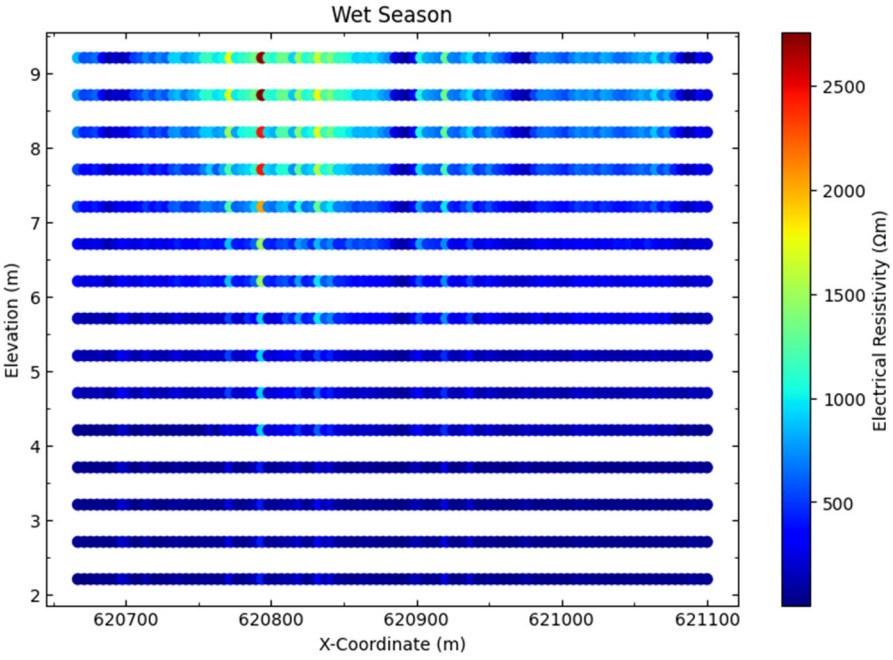




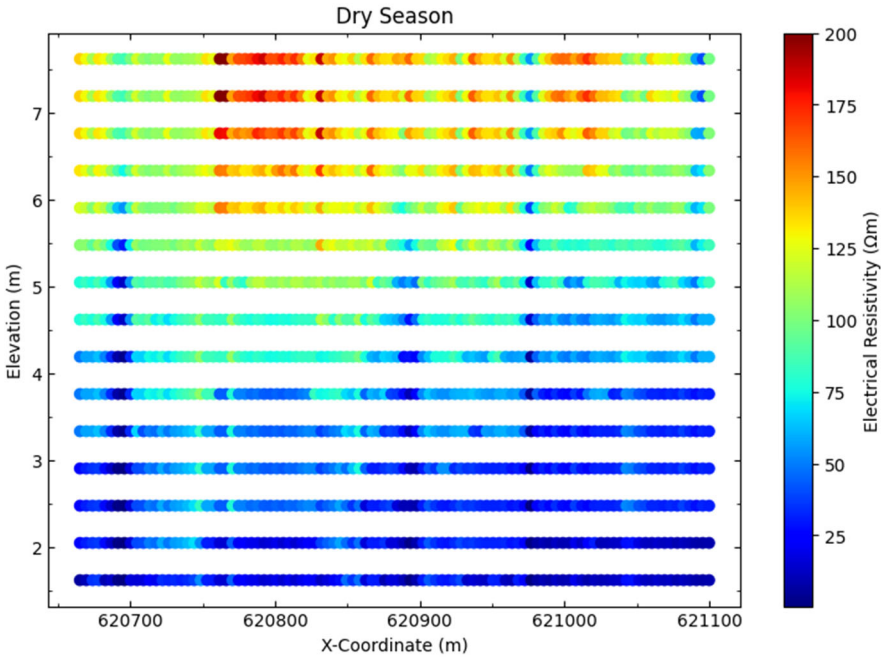
Grand Island – EMI Profiles for Comparisons

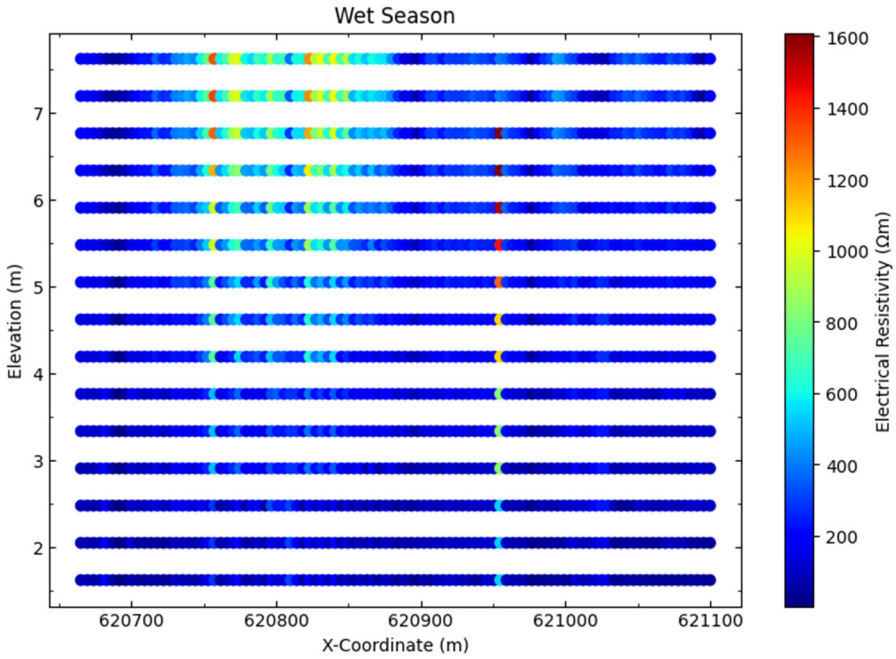
Grand Island
Site 1
Dry Crest



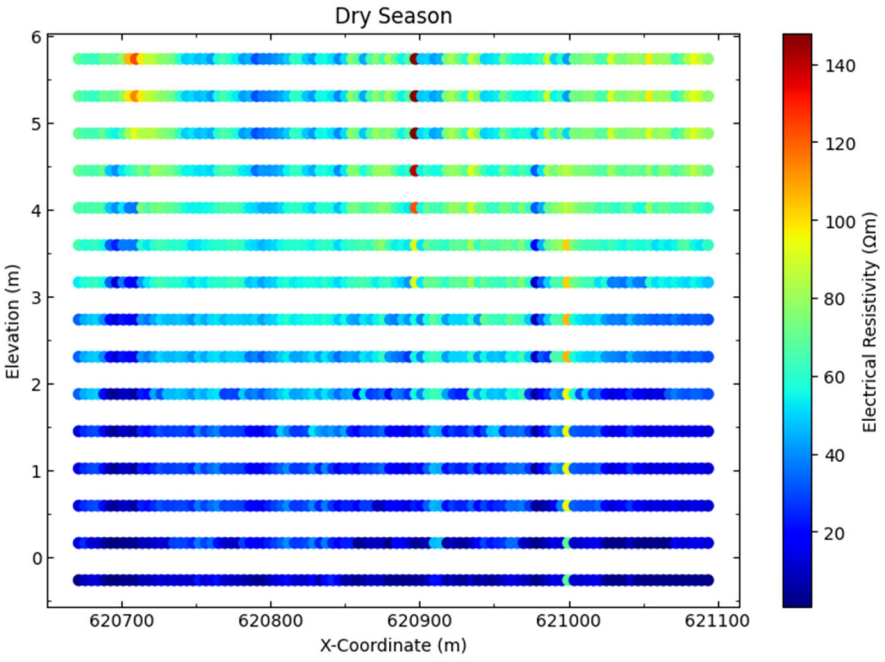


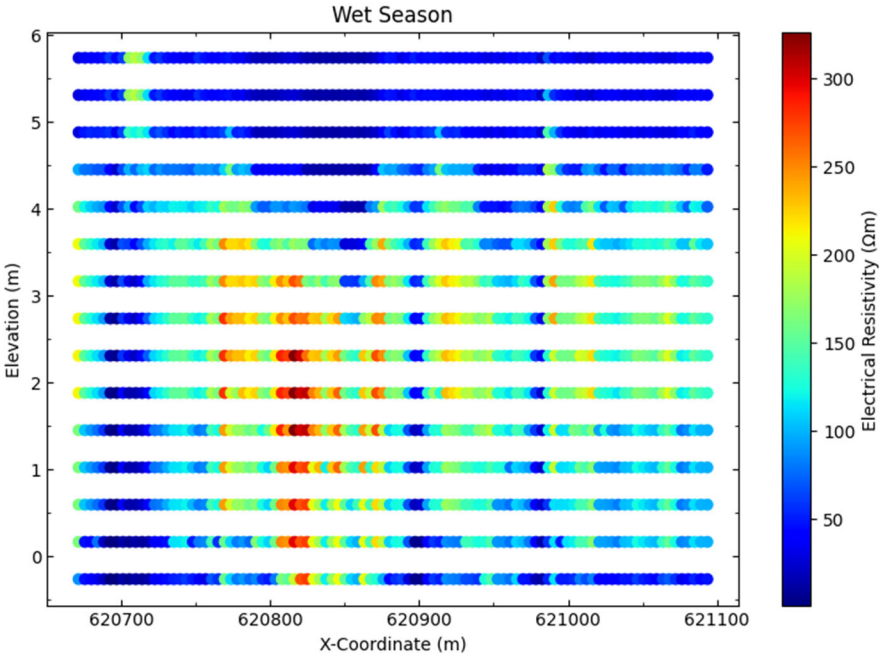
Grand Island
Site 1
Dry Mid-Slope



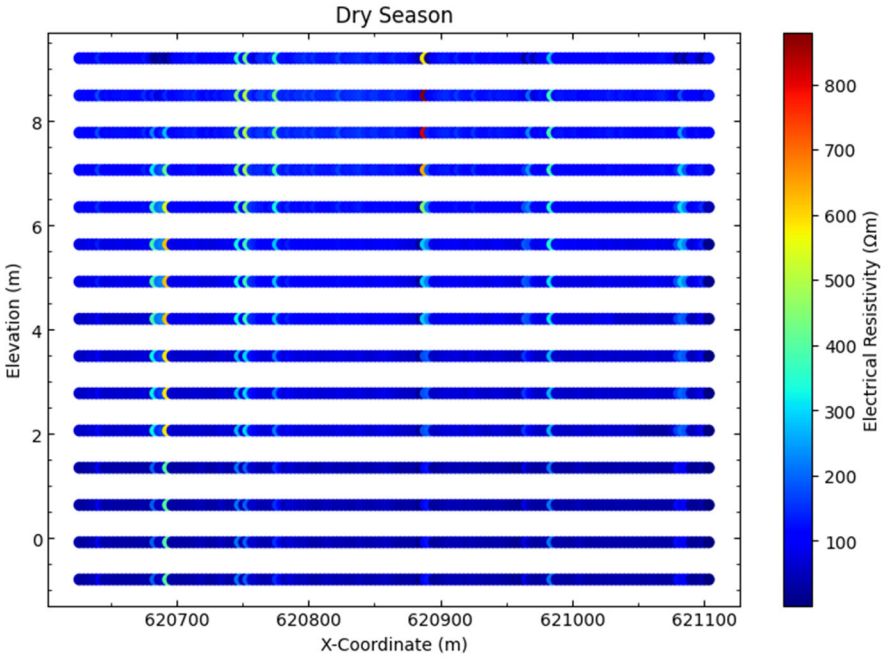


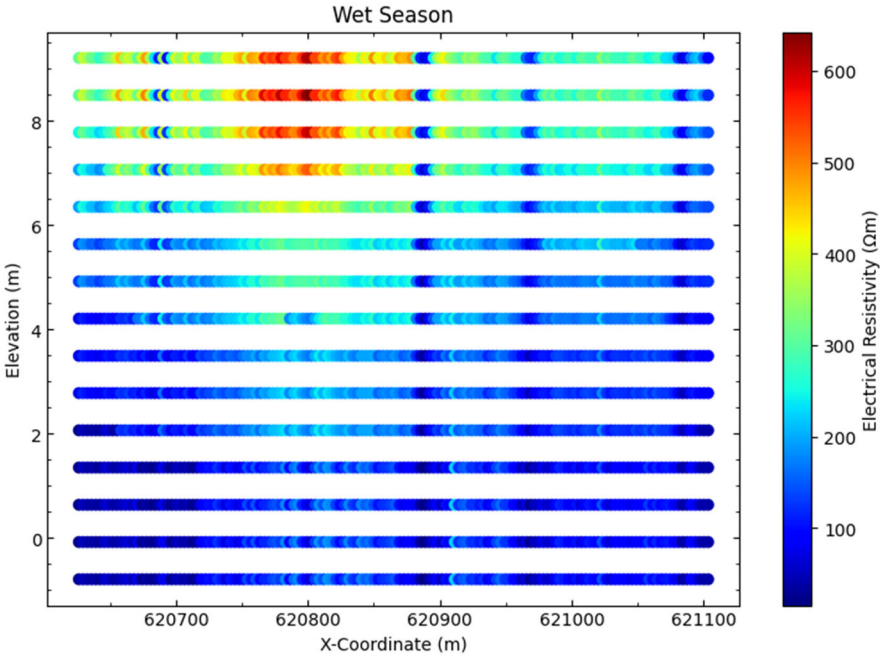
Grand Island
Site 1
Dry Toe



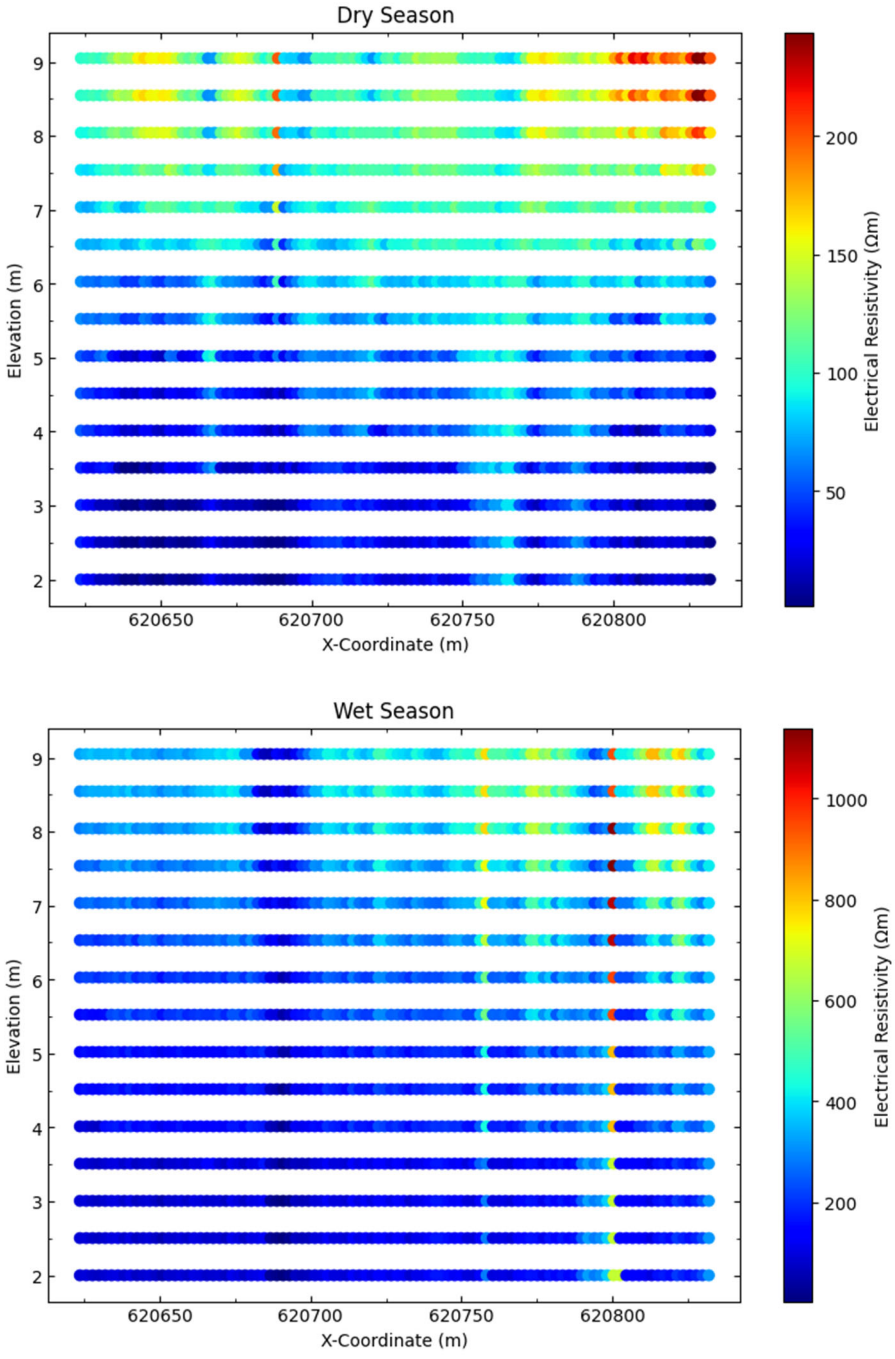


Grand Island
Site 1
Wet Crest

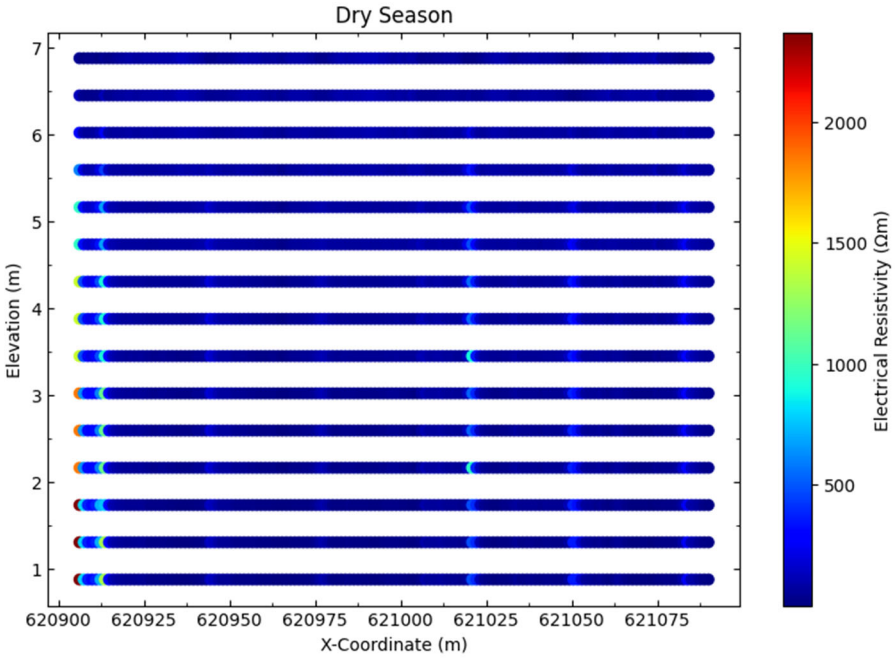


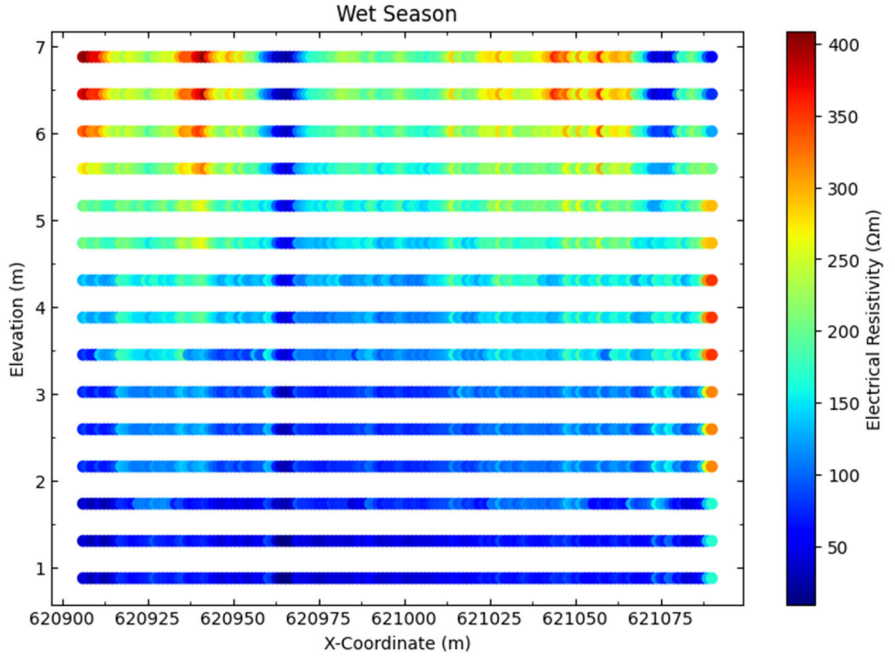


Grand Island
Site 1
Wet Mid-Slope

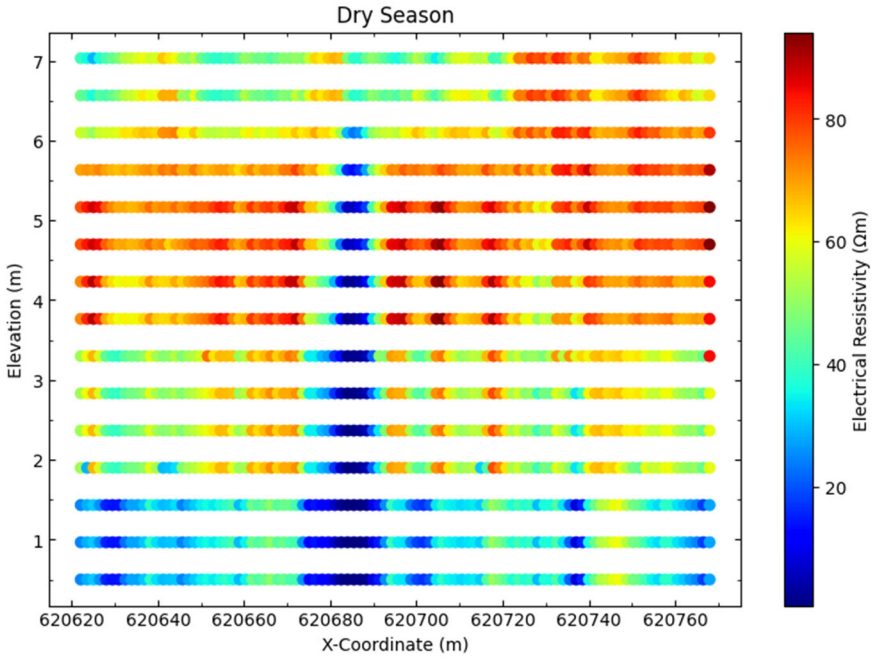


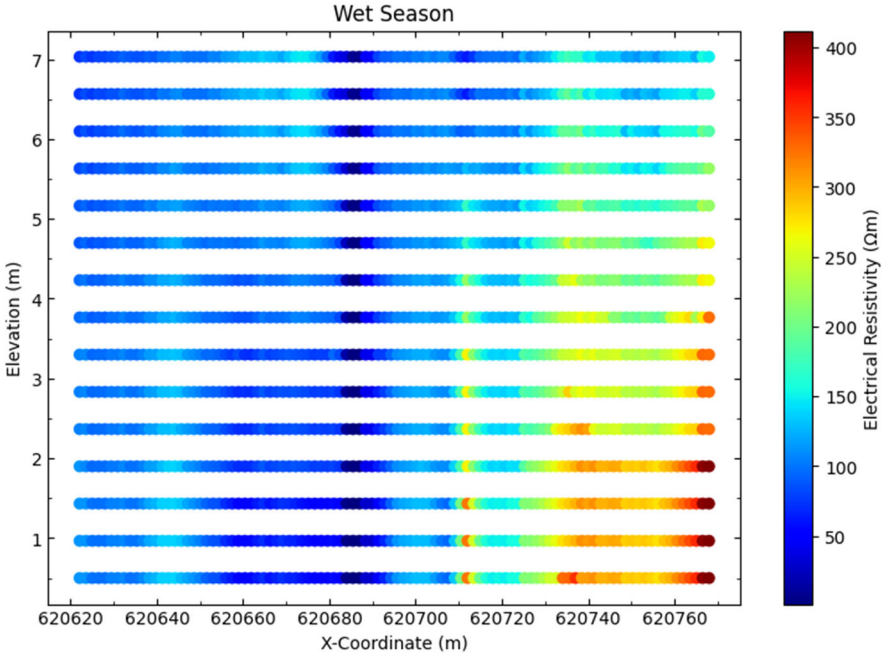
Grand Island
Site 1
Wet Toe - East



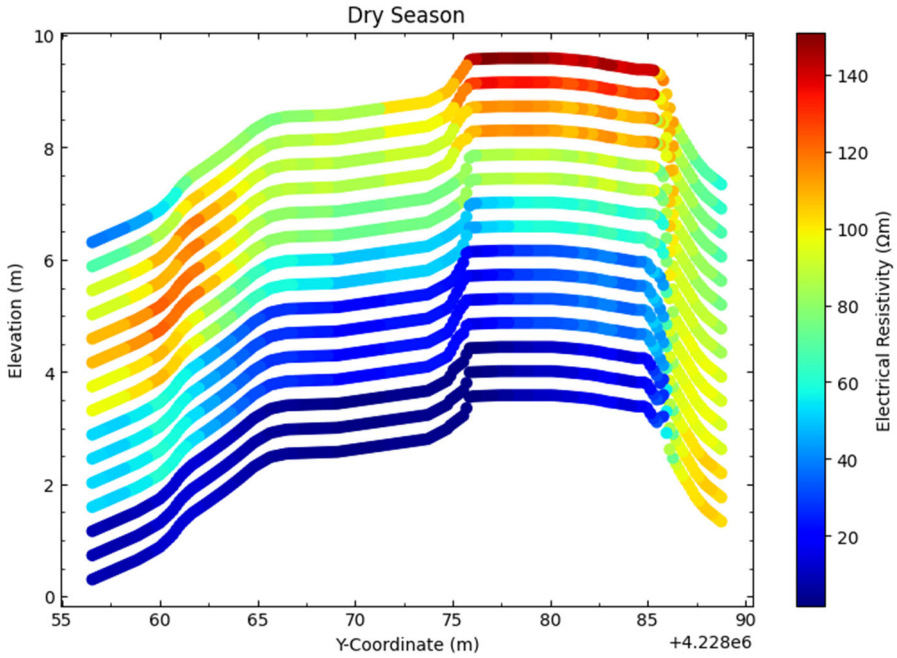


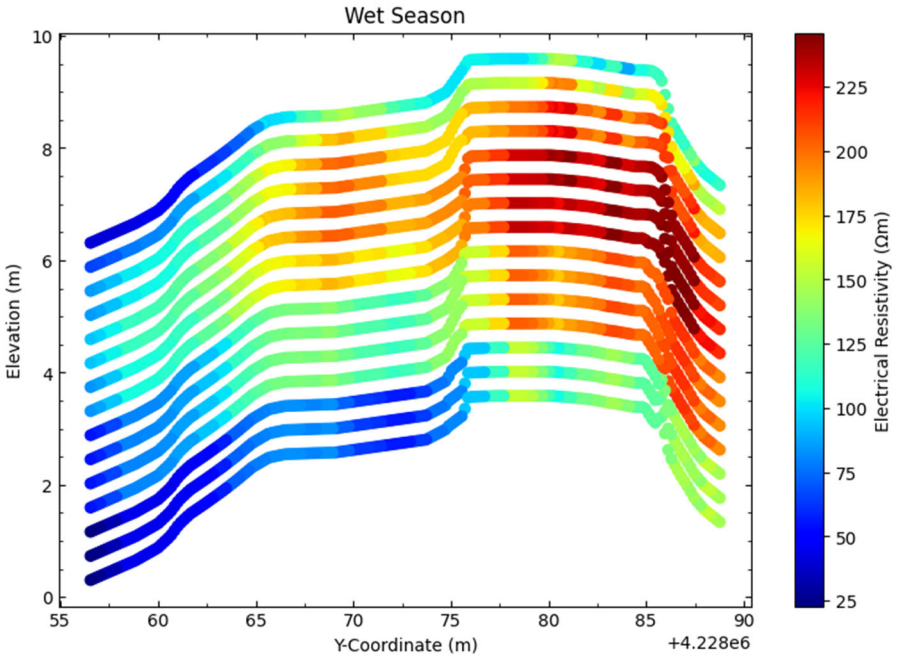
Grand Island
Site 1
Wet Toe – West



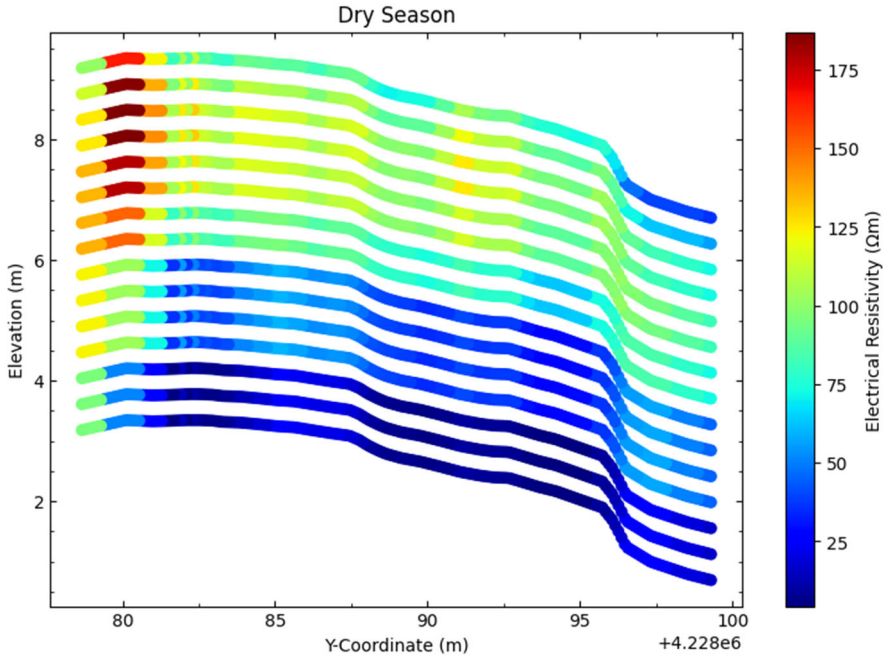


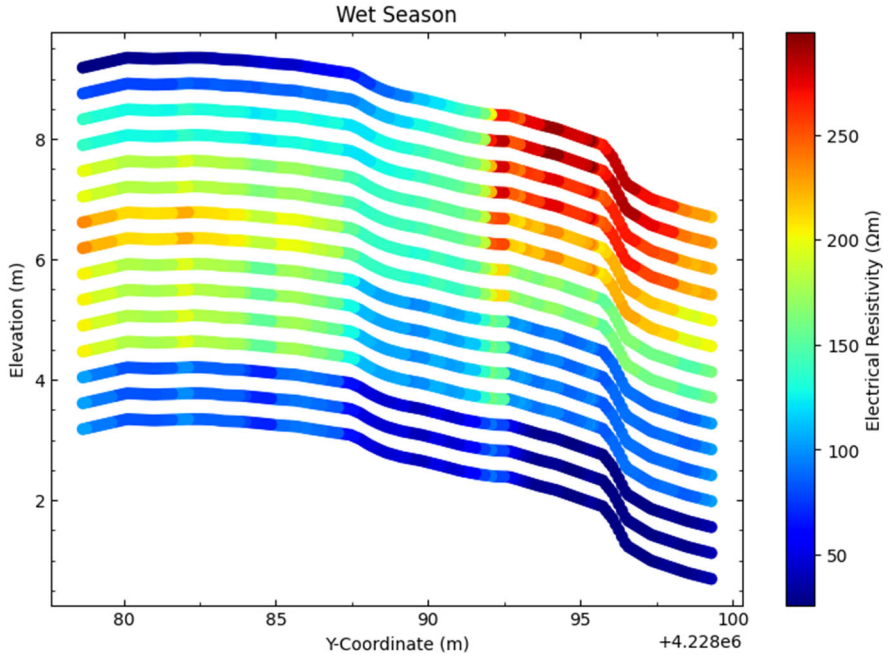
Grand Island
Site 1
Cross Section 1



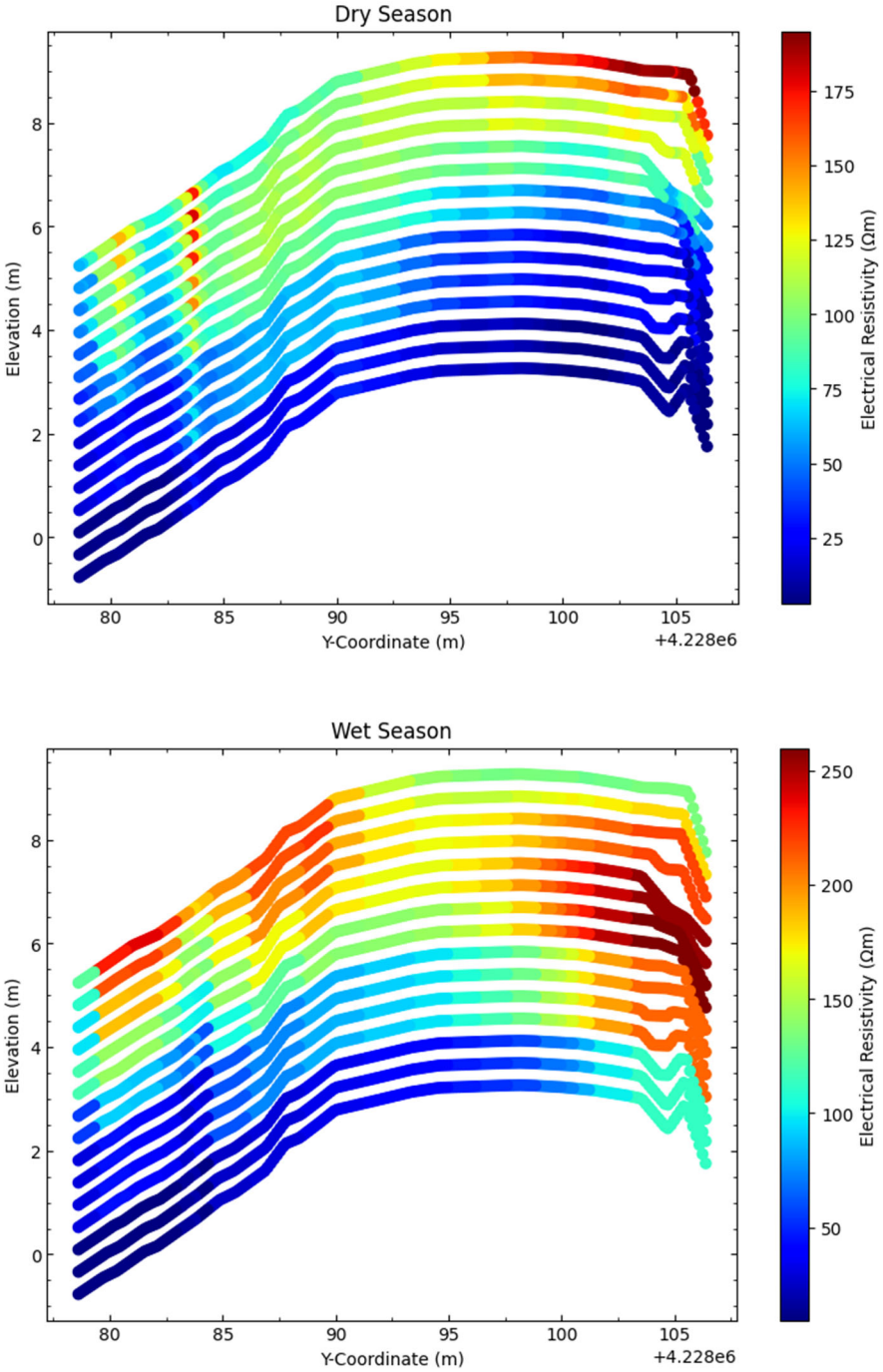


Grand Island
Site 1
Cross Section 2

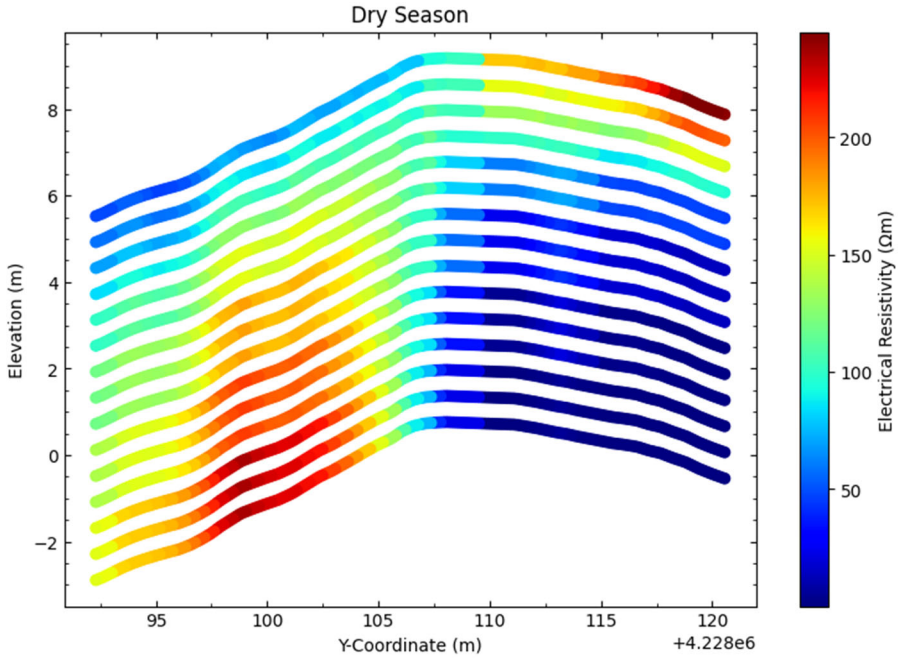


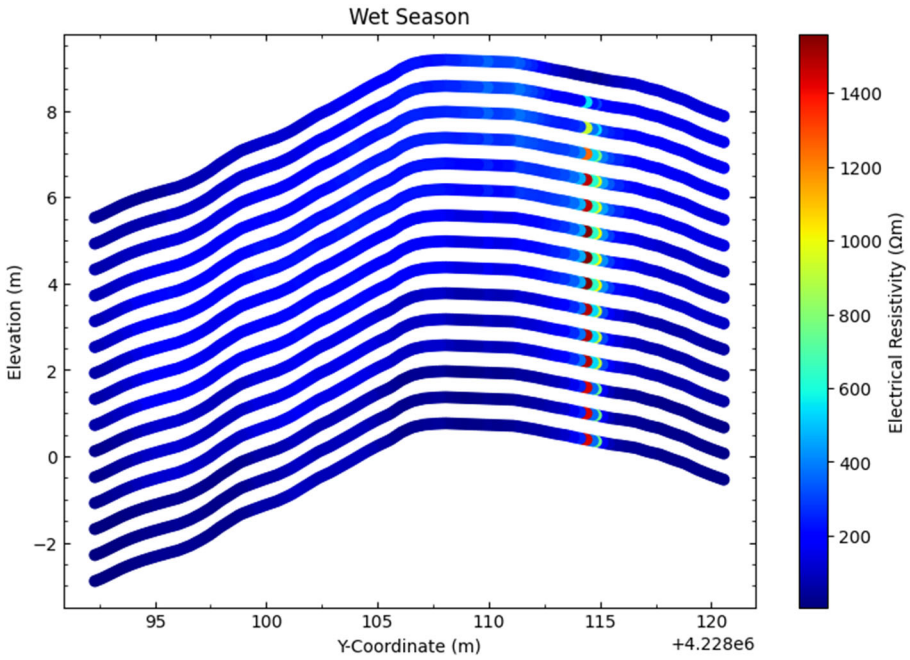


Grand Island
Site 1
Cross Section 3

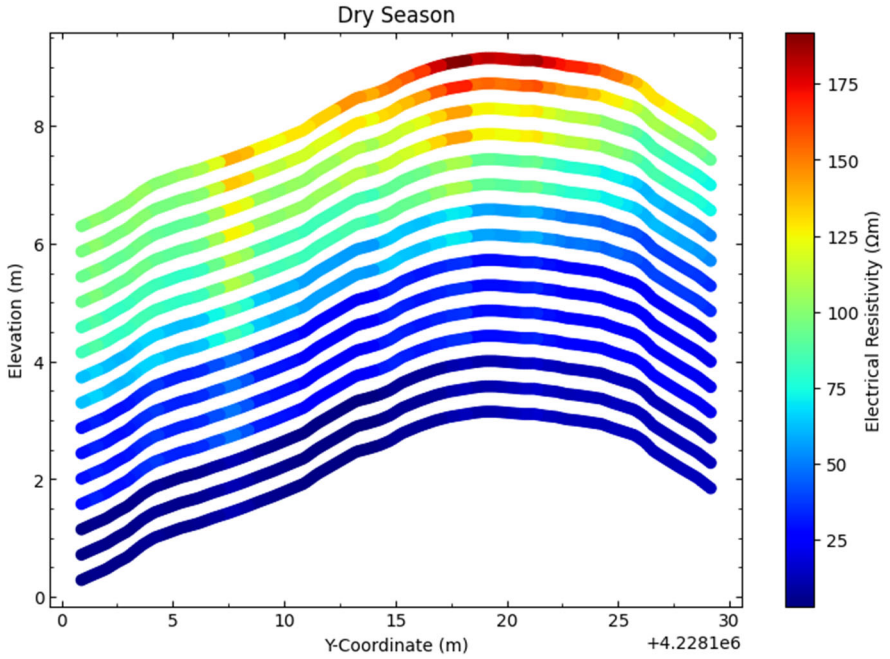


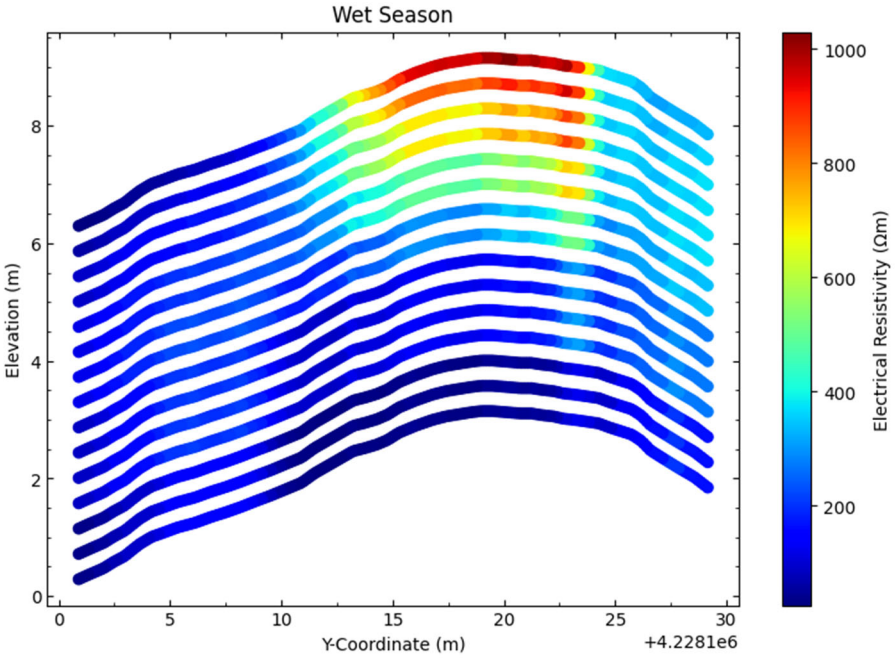
Grand Island
Site 1
Cross Section 4



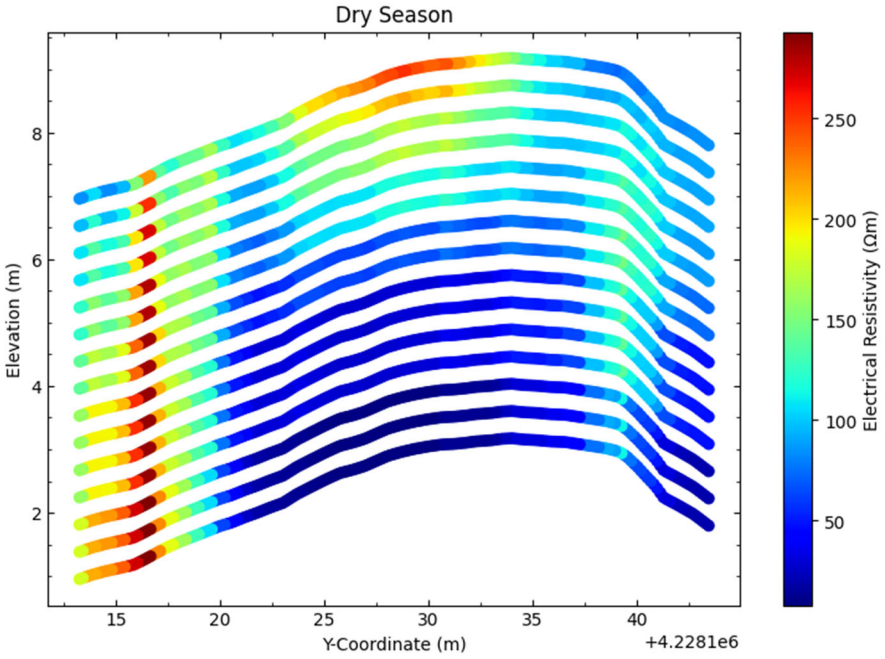


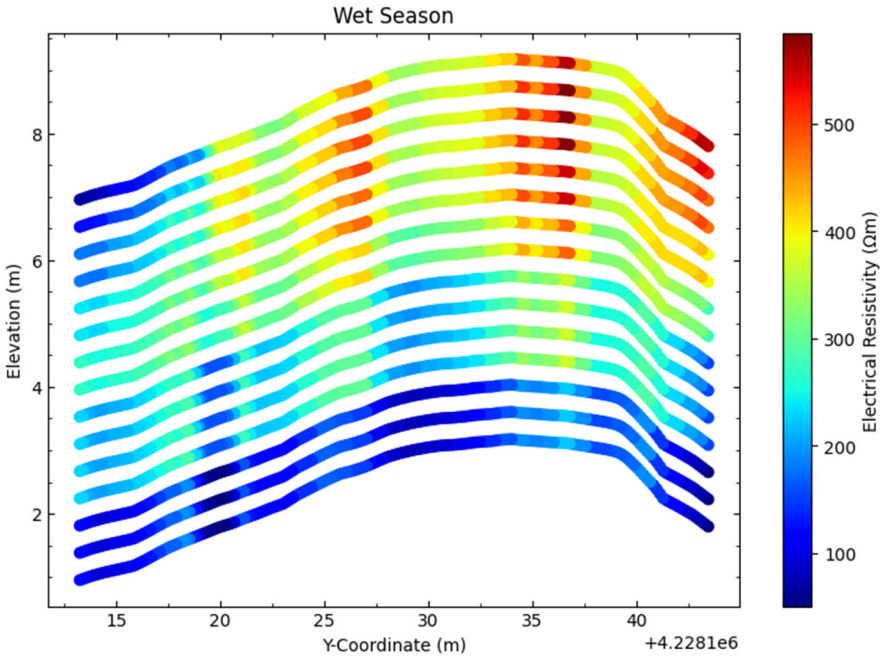
Grand Island
Site 1
Cross Section 5



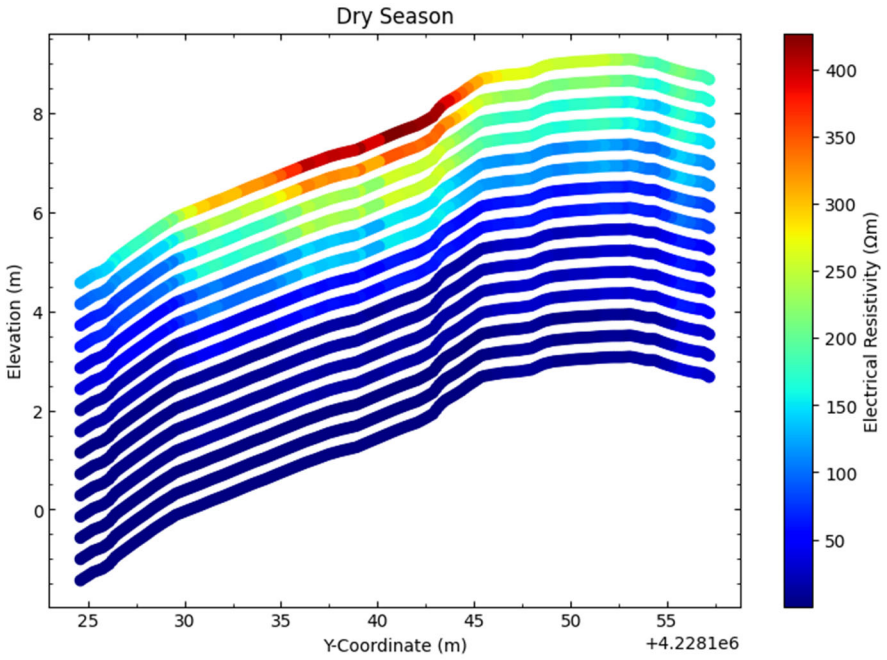


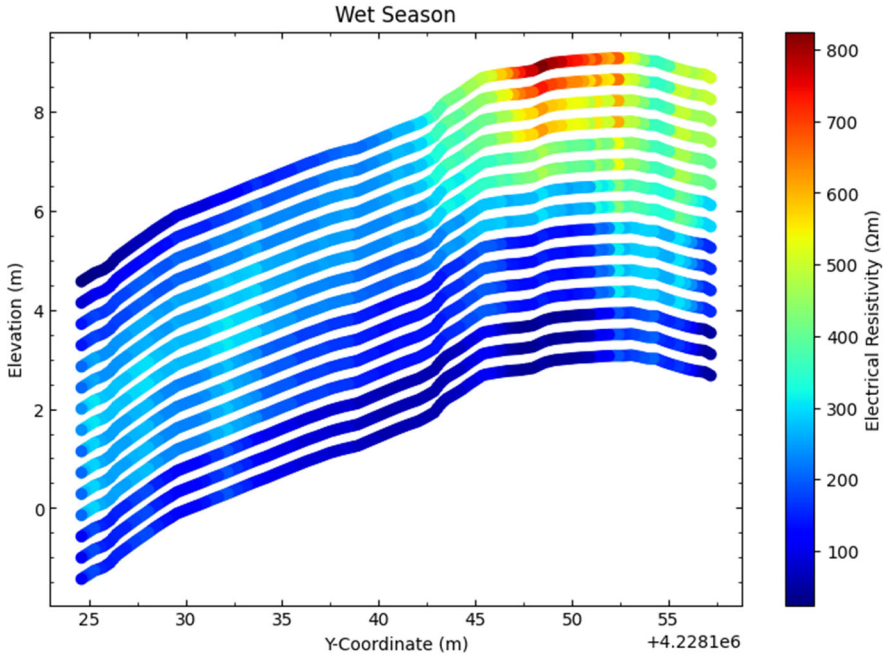
Grand Island
Site 1
Cross Section 6



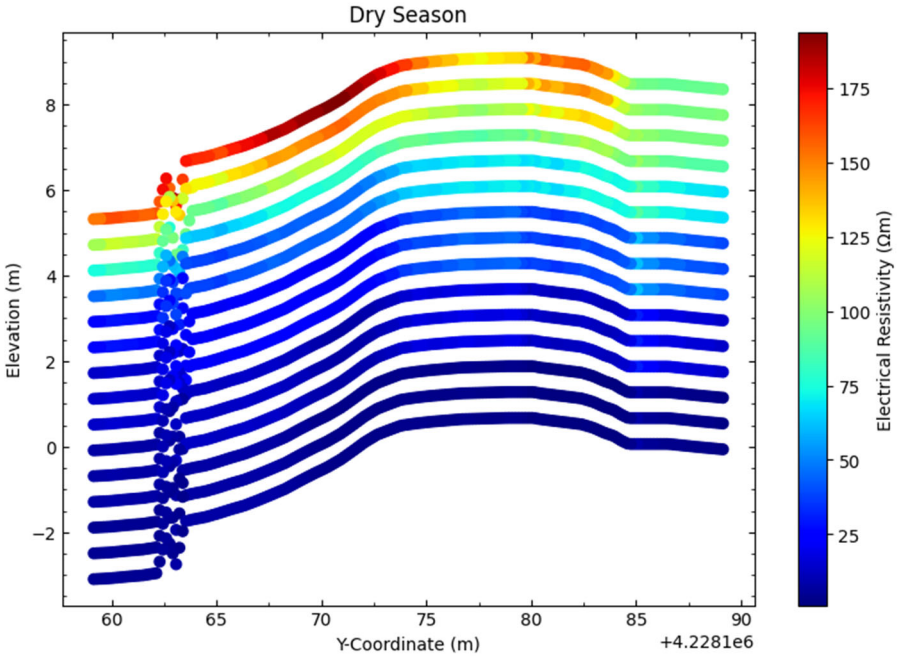


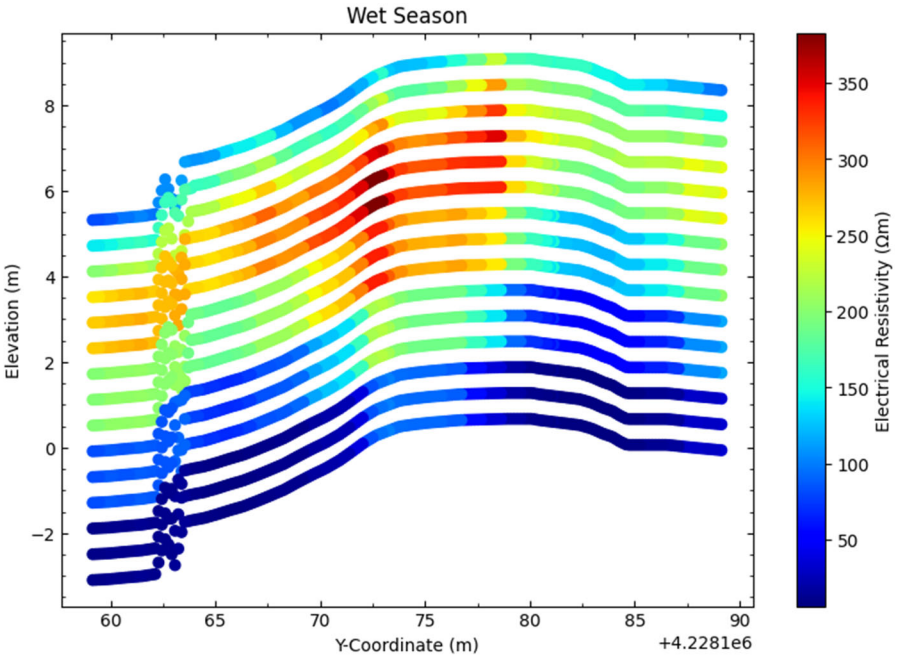
Grand Island
Site 1
Cross Section 7



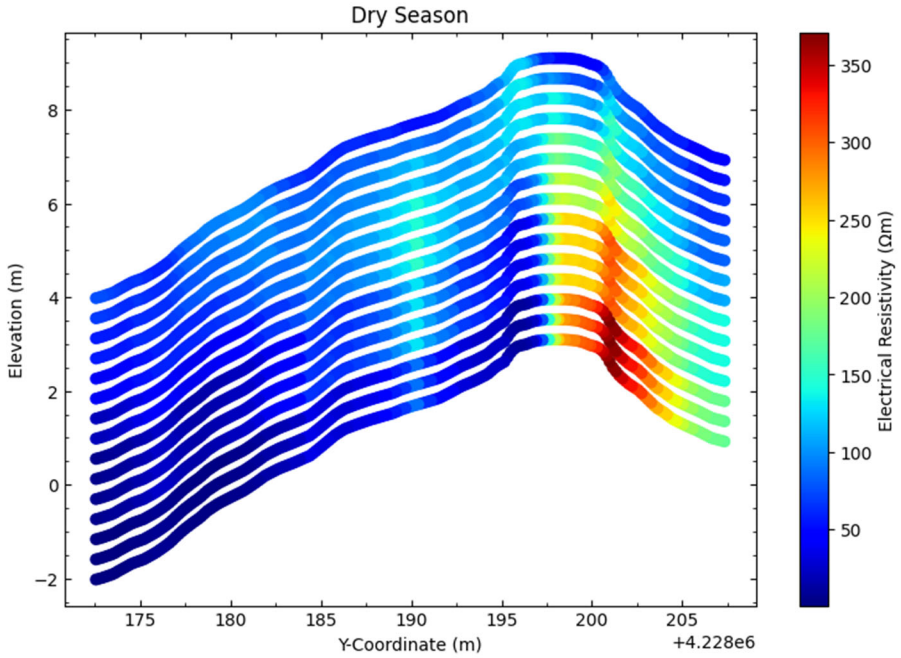


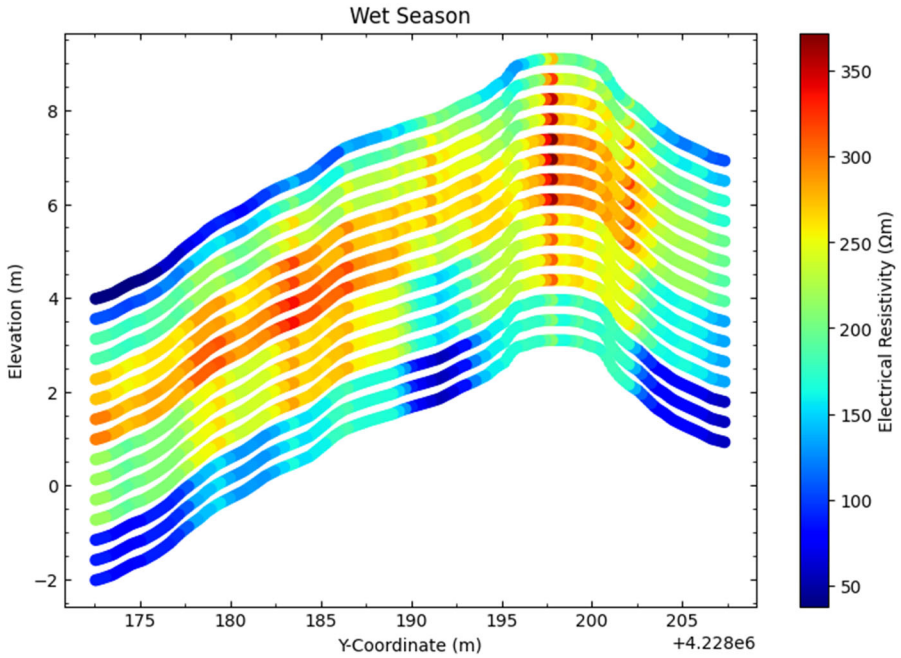
Grand Island
Site 1
Cross Section 8



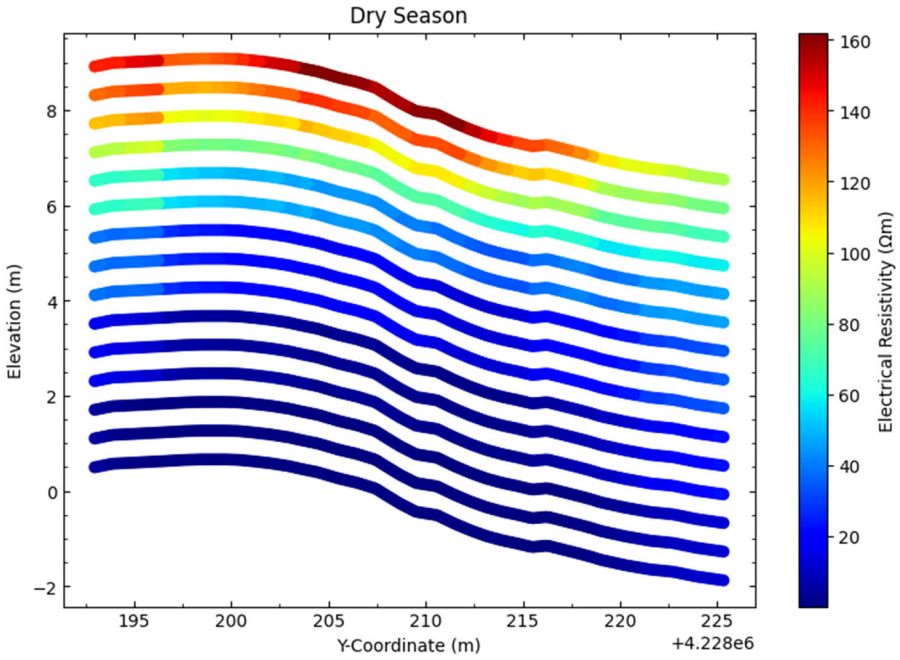


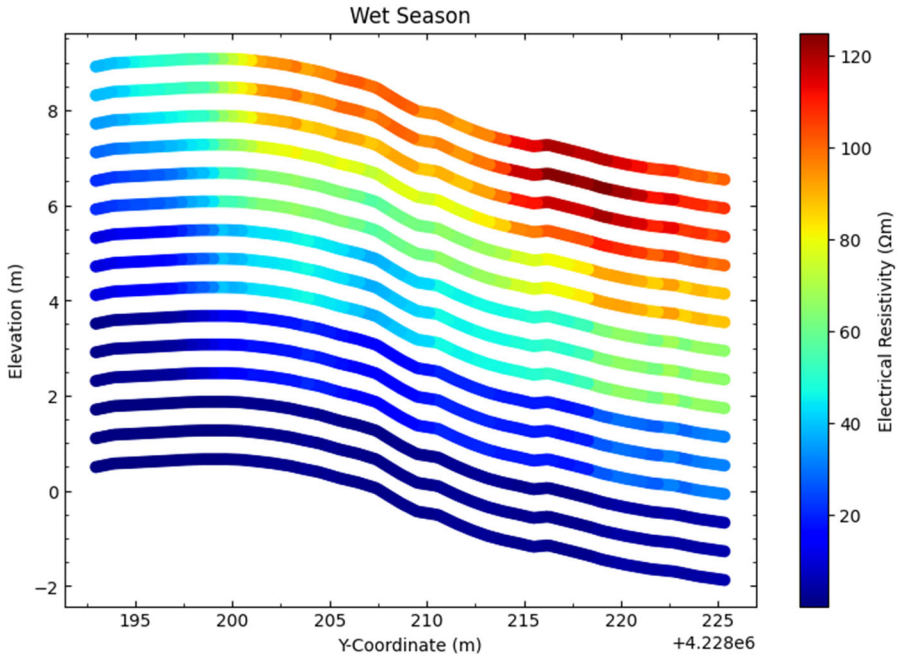
Grand Island
Site 1
Cross Section 9



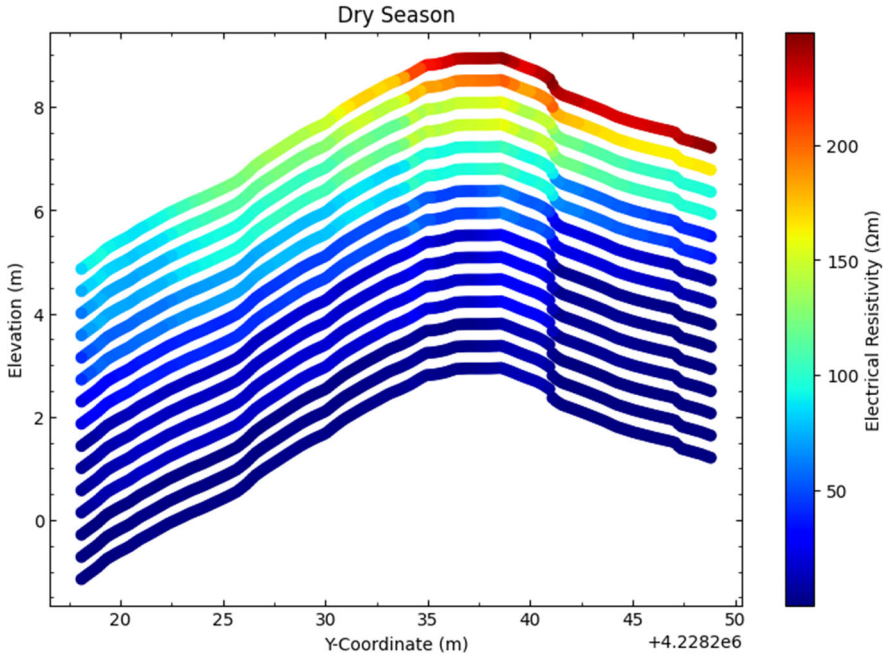


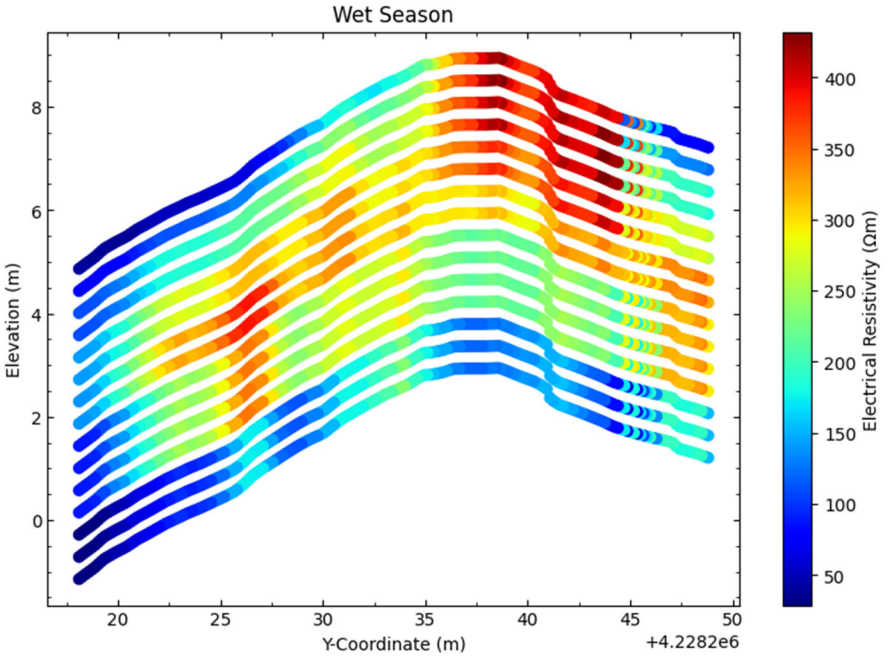
Grand Island
Site 1
Cross Section 10



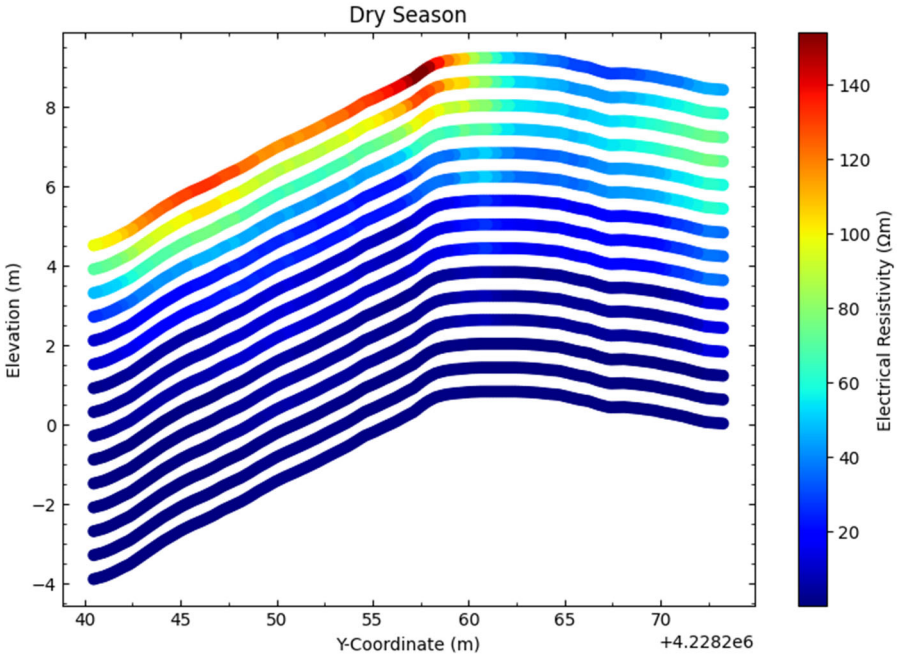


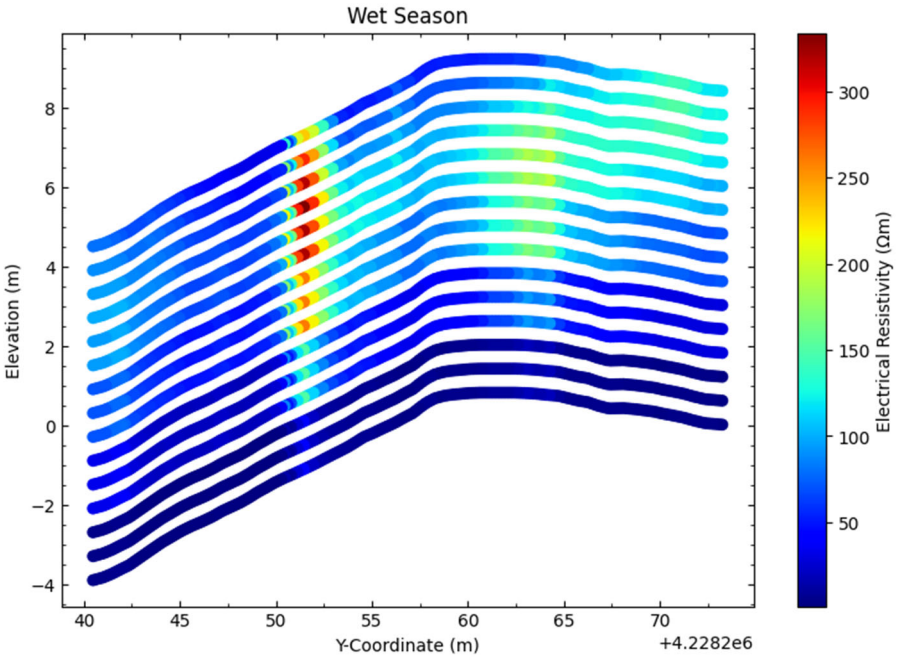
Grand Island
Site 1
Cross Section 11



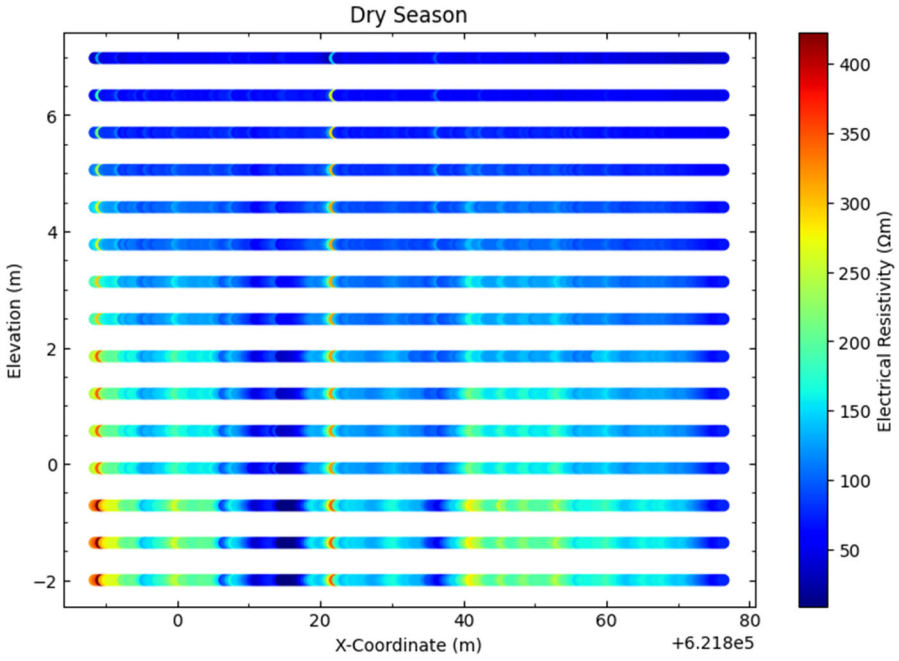


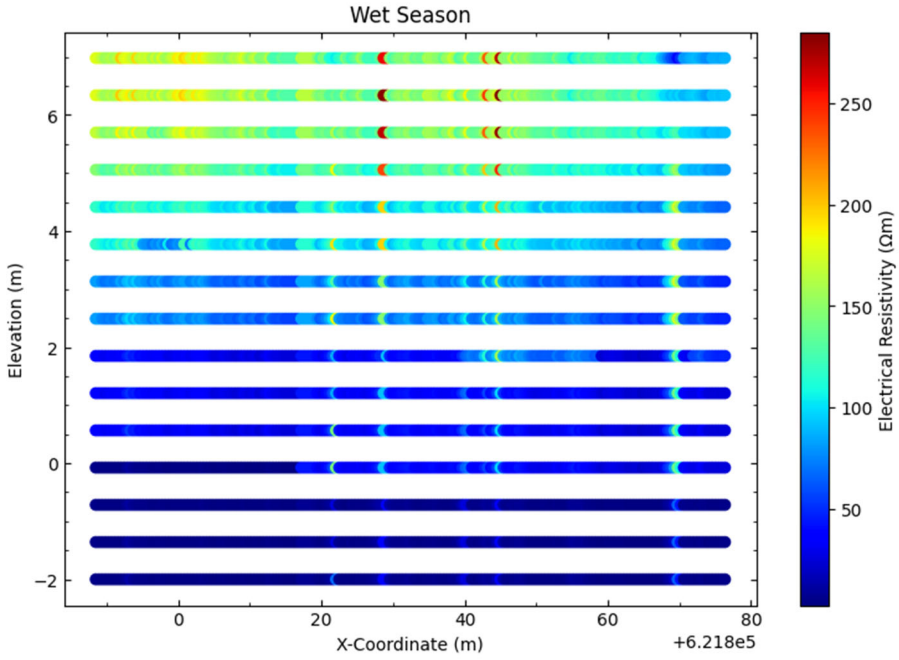
Grand Island
Site 1
Cross Section 12



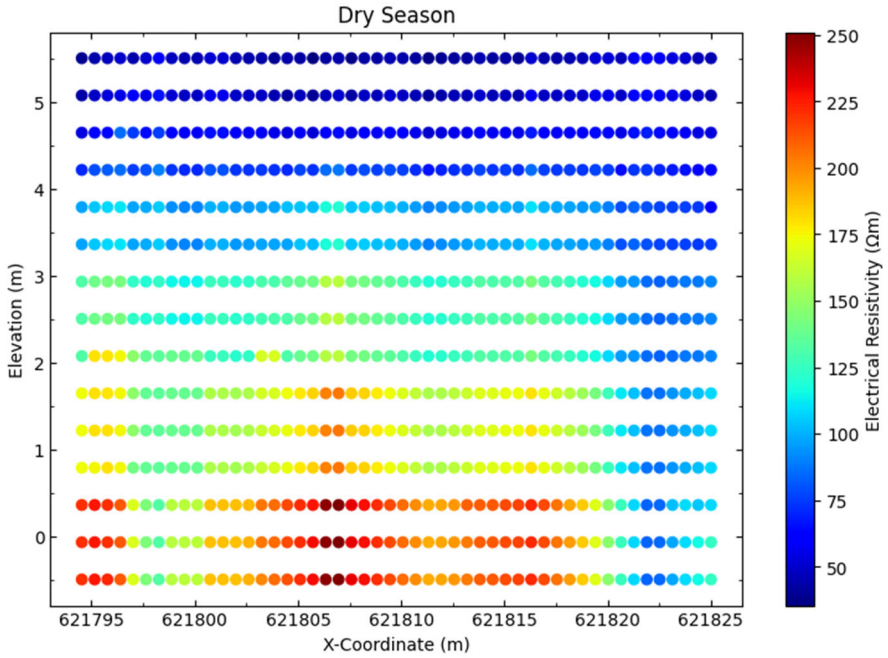


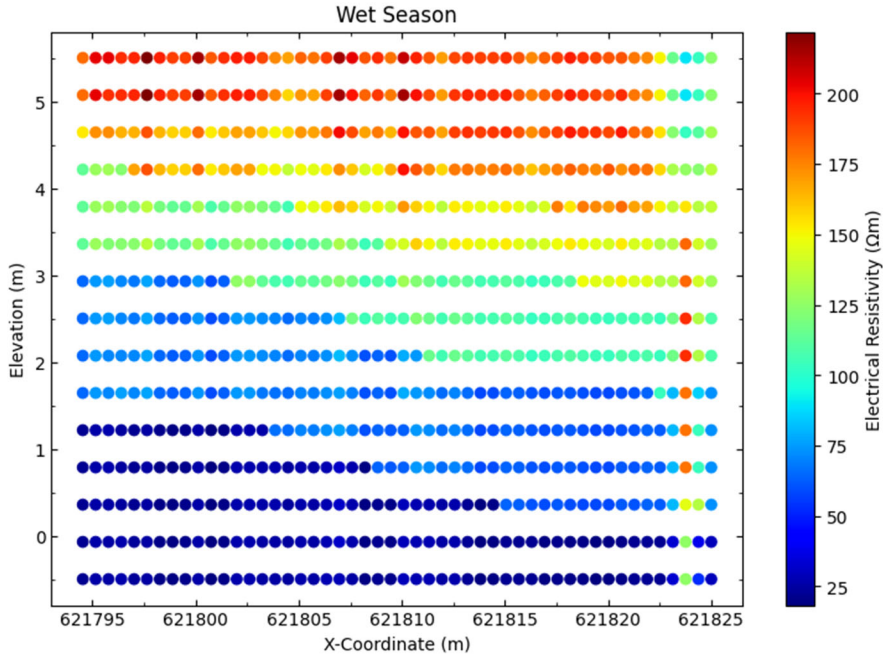
Grand Island
Site 2
Dry Crest



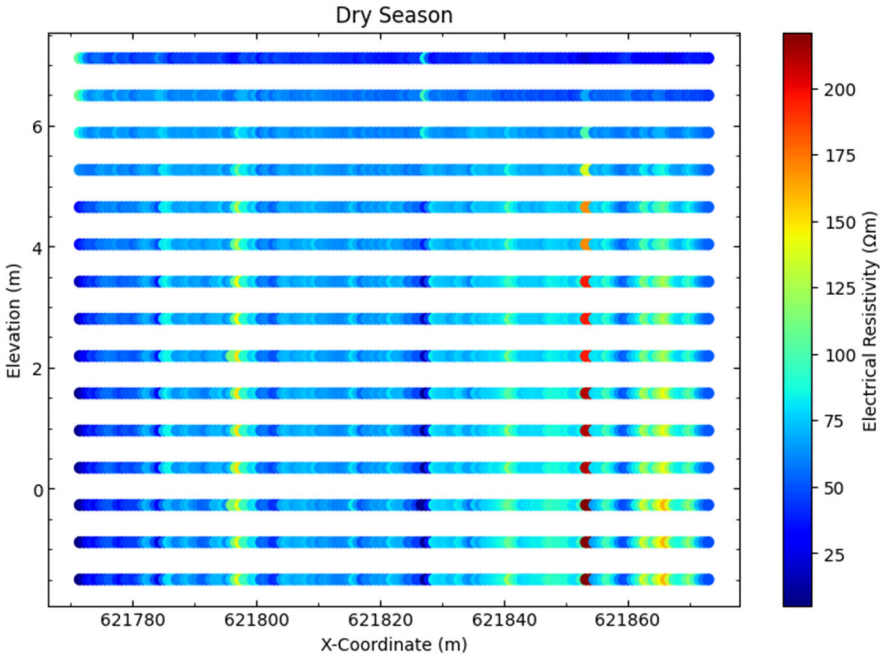


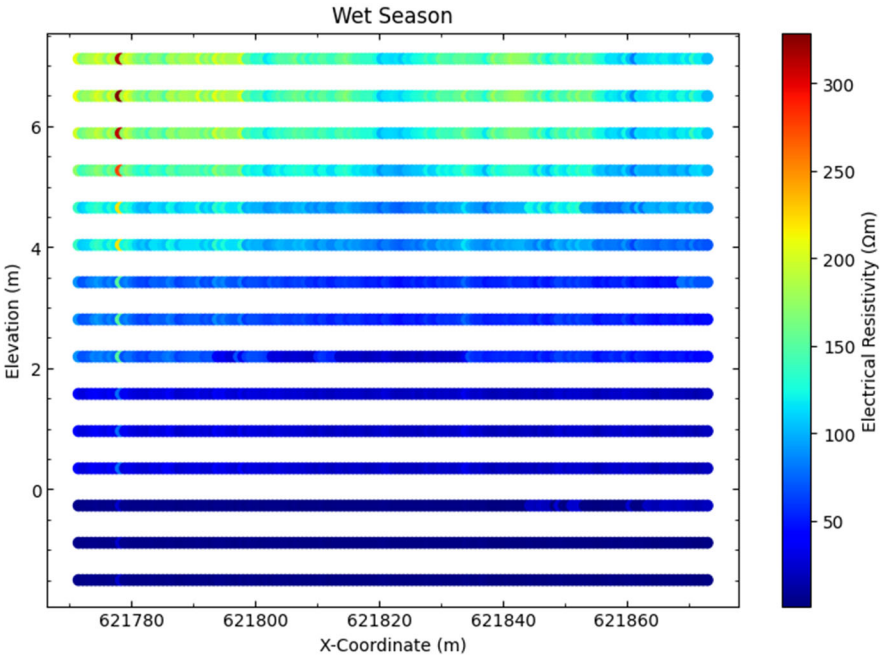
Grand Island
Site 2
Dry Mid-Slope



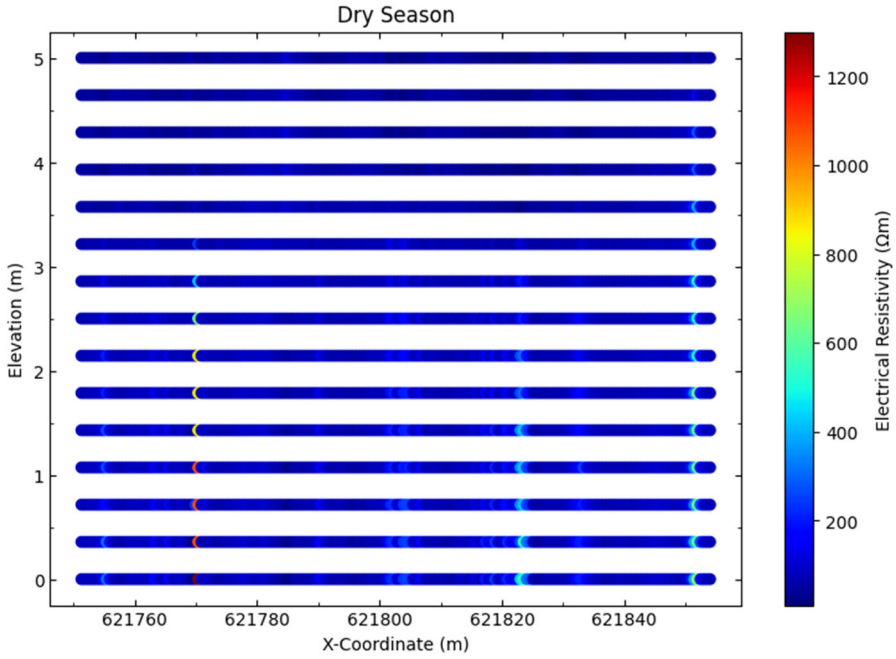


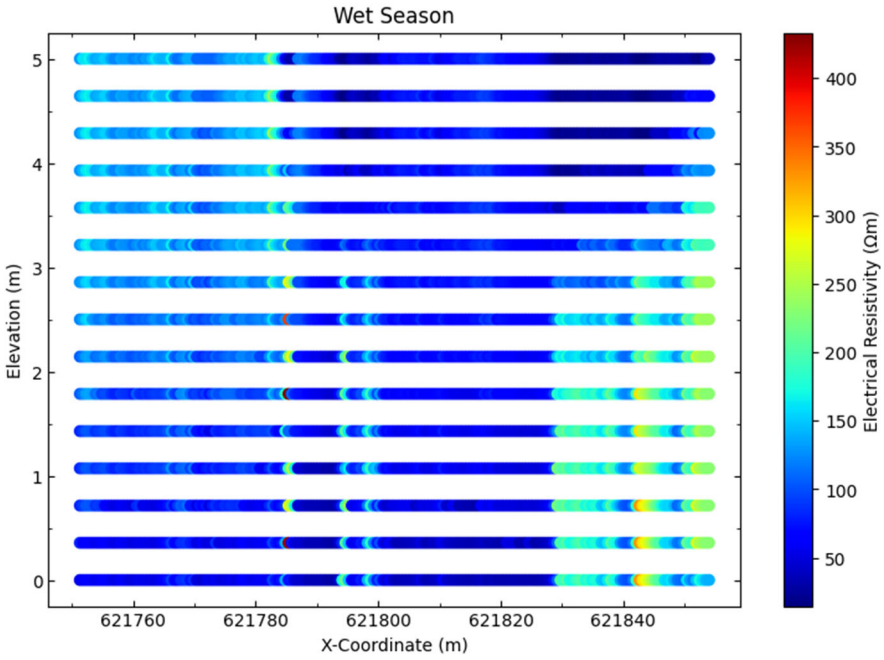
Grand Island
Site 2
Wet Crest



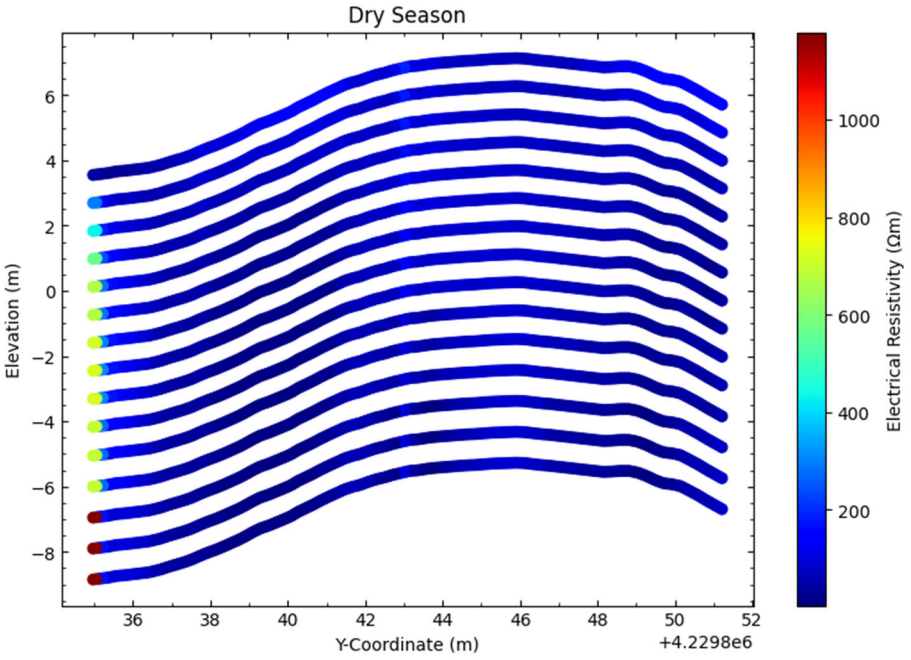


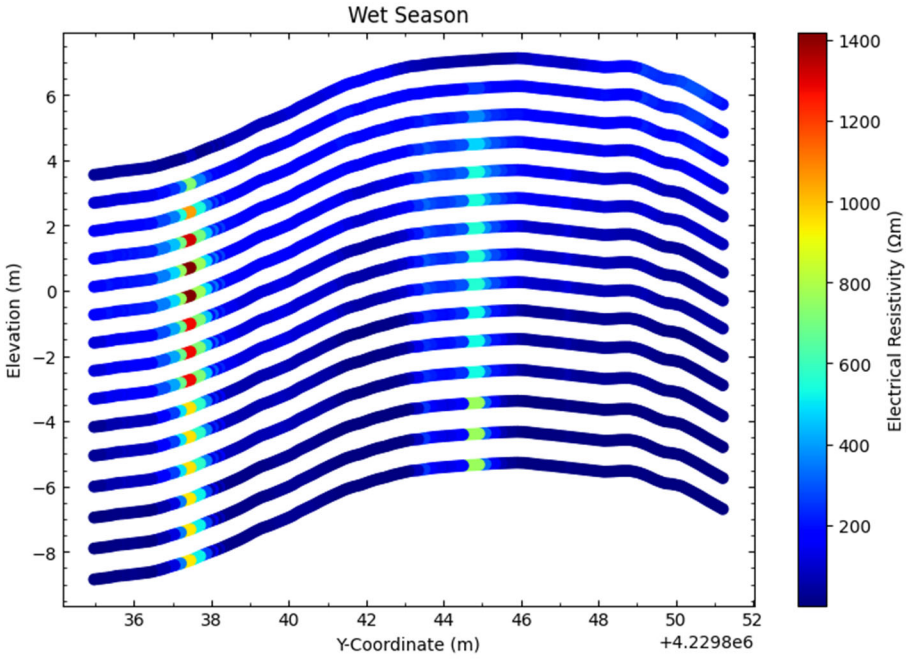
Grand Island
Site 2
Wet Toe



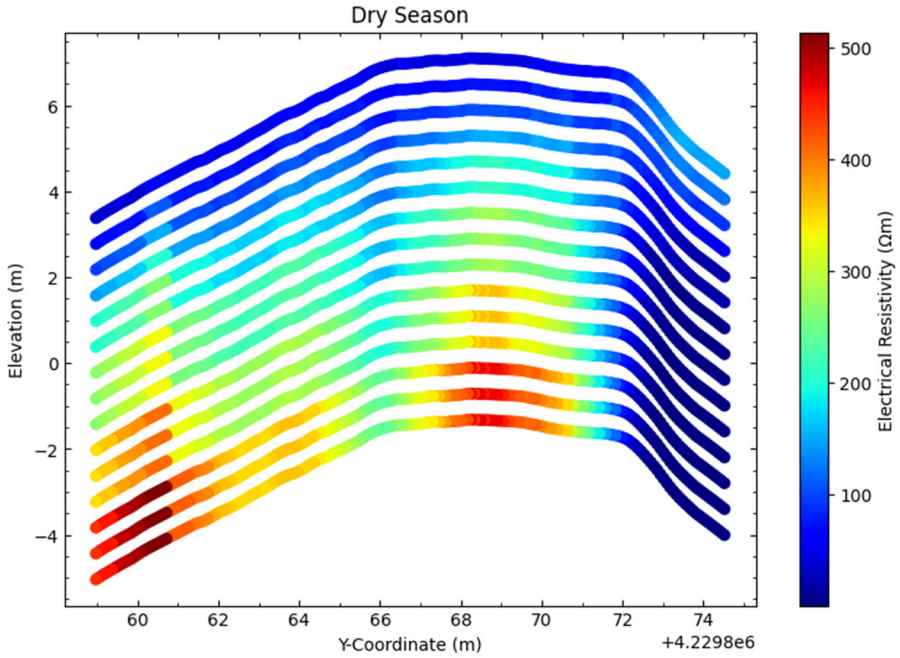


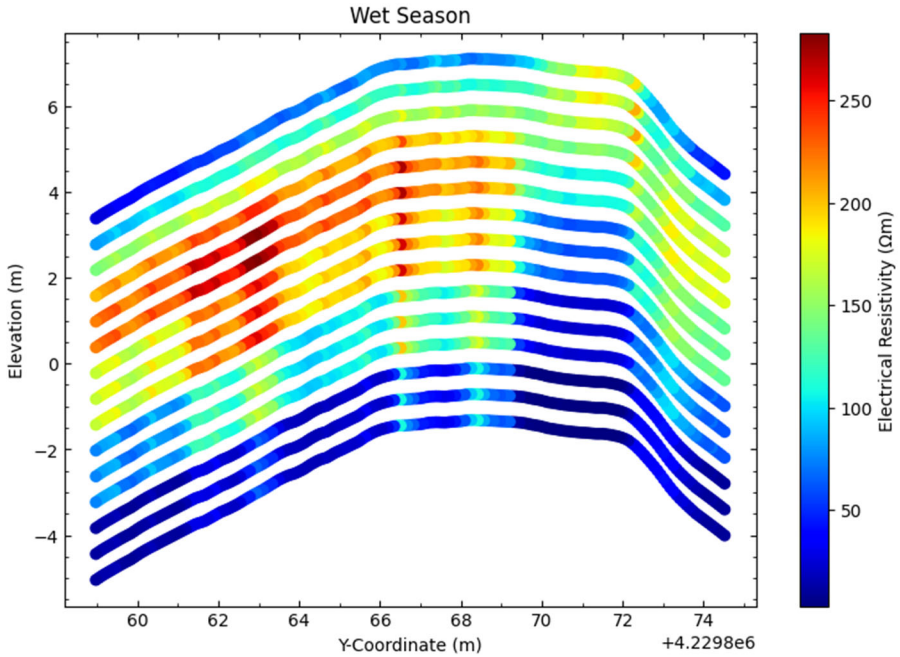
Grand Island
Site 2
Cross Section 1



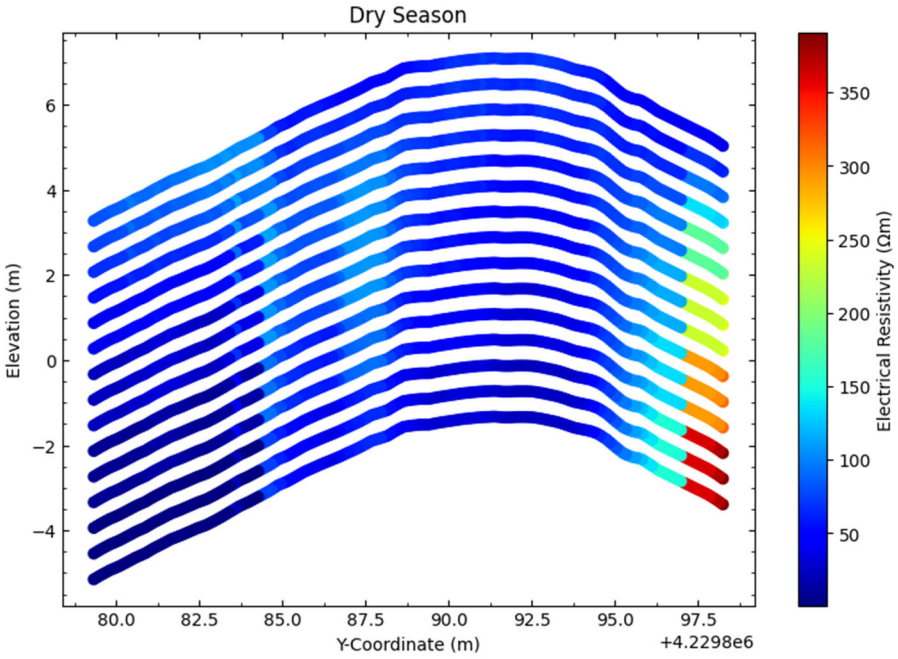


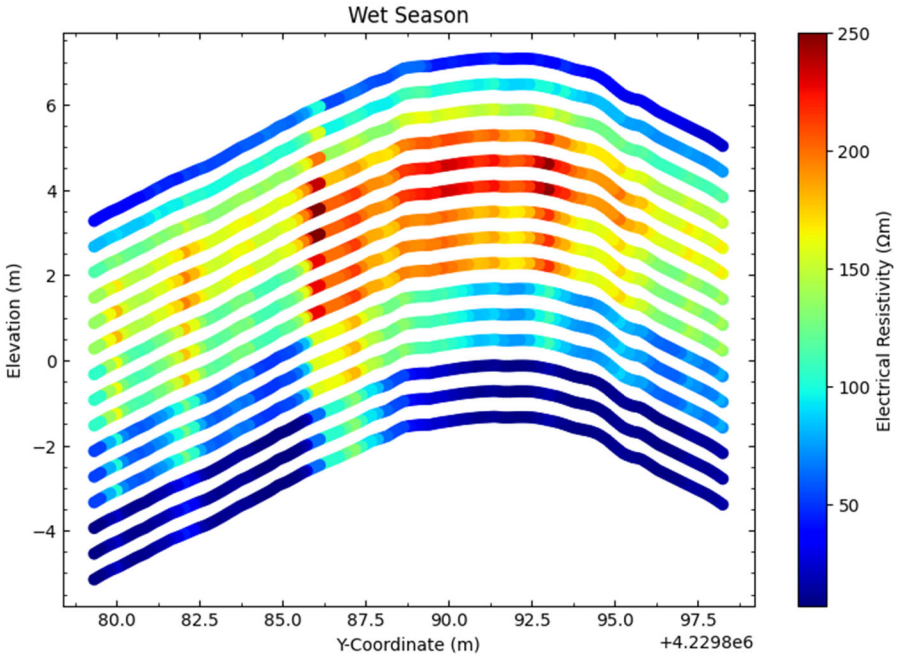
Grand Island
Site 2
Cross Section 2



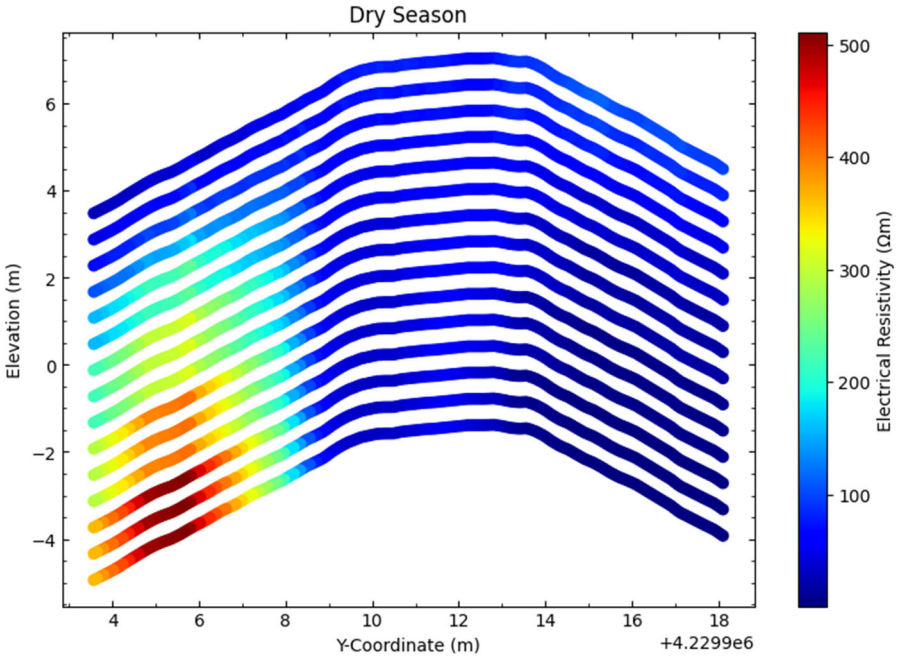


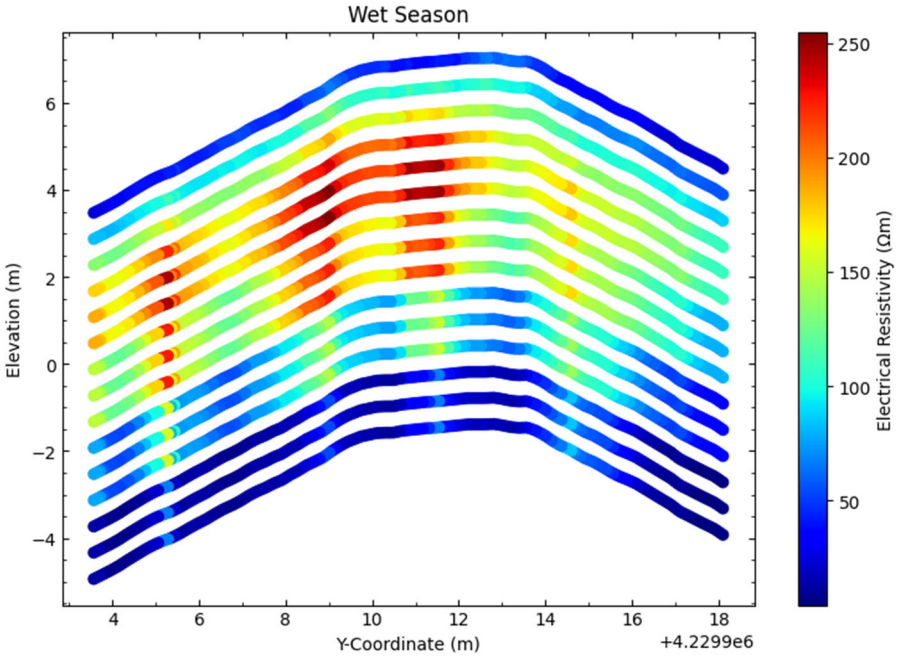
Grand Island
Site 2
Cross Section 3



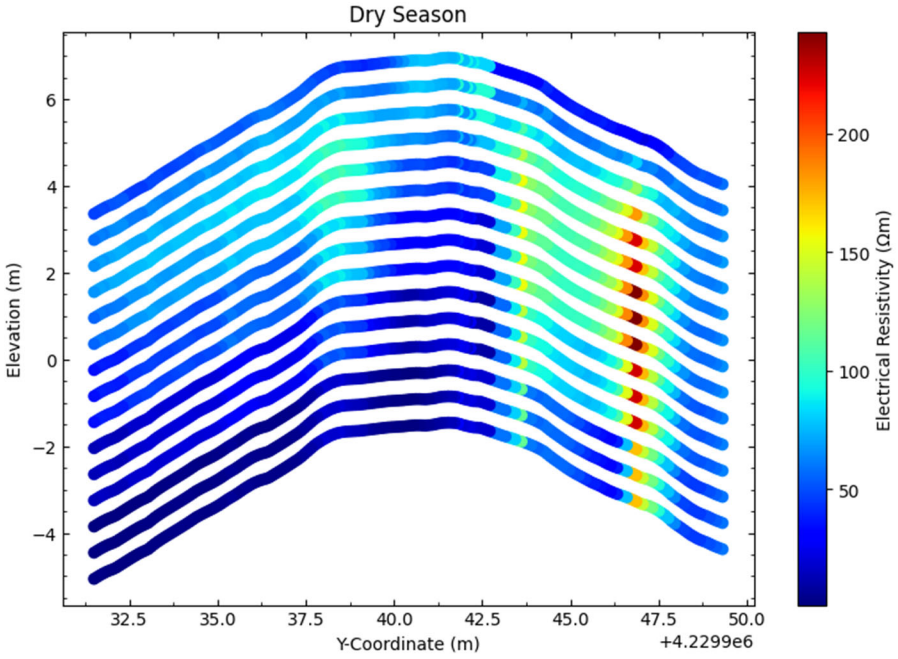


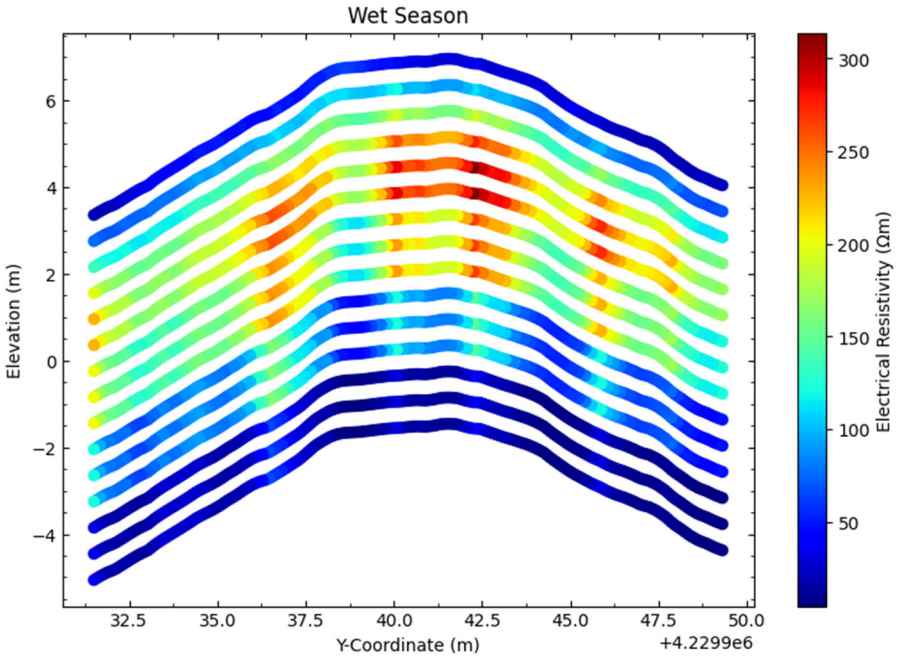
Grand Island
Site 2
Cross Section 4



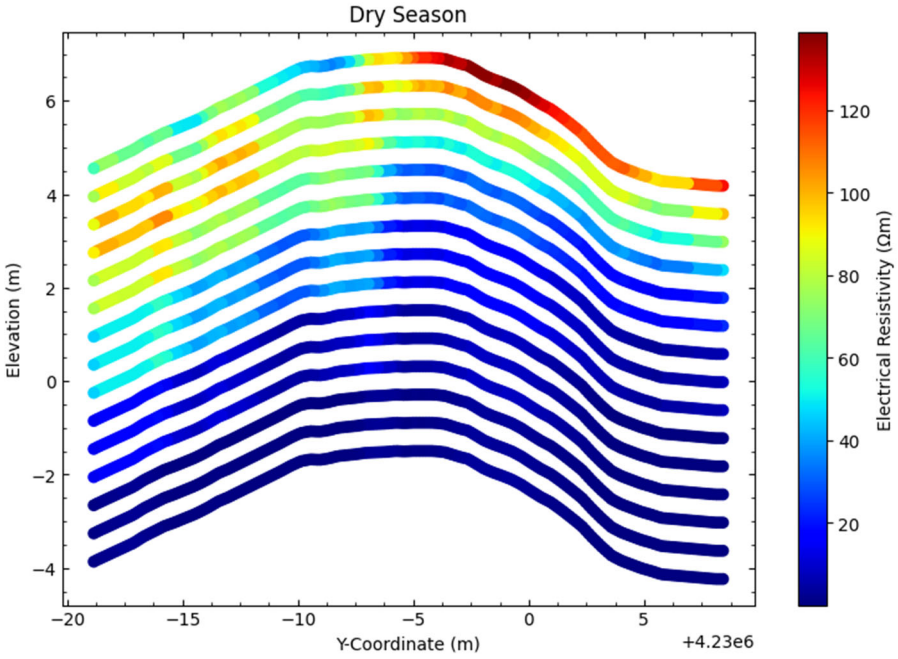


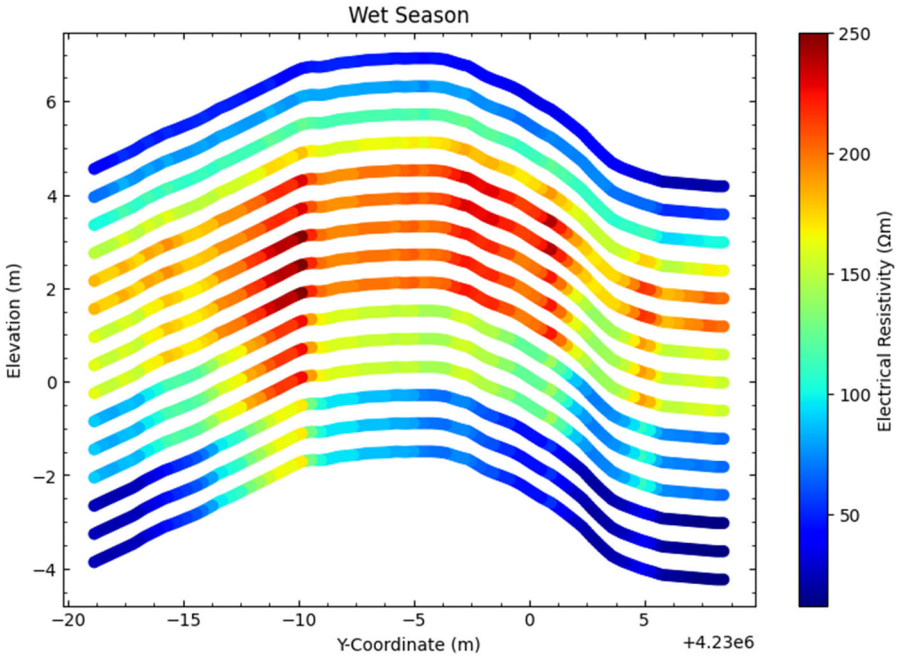
Grand Island
Site 2
Cross Section 5



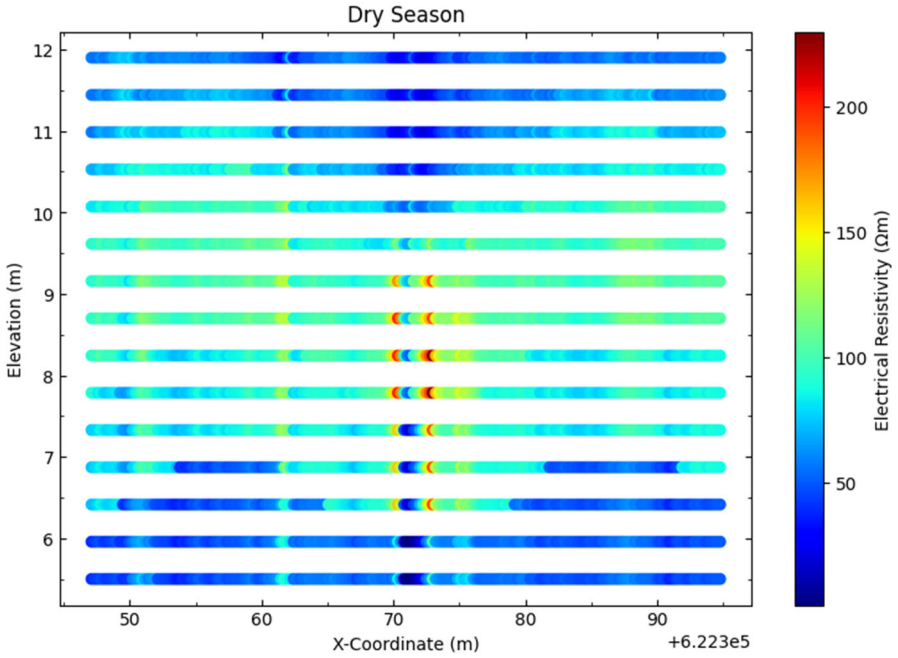


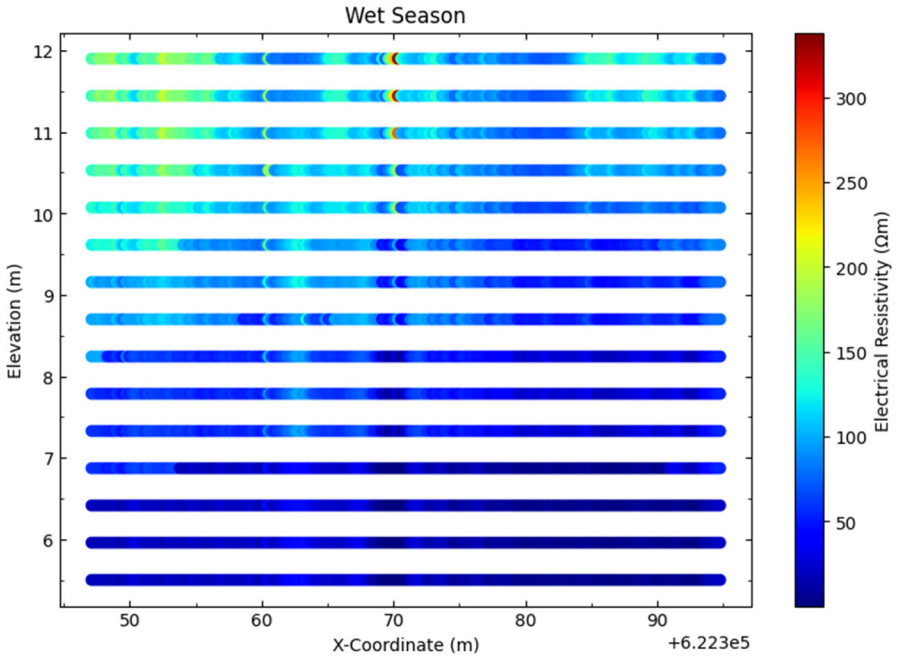
Grand Island
Site 2
Cross Section 7



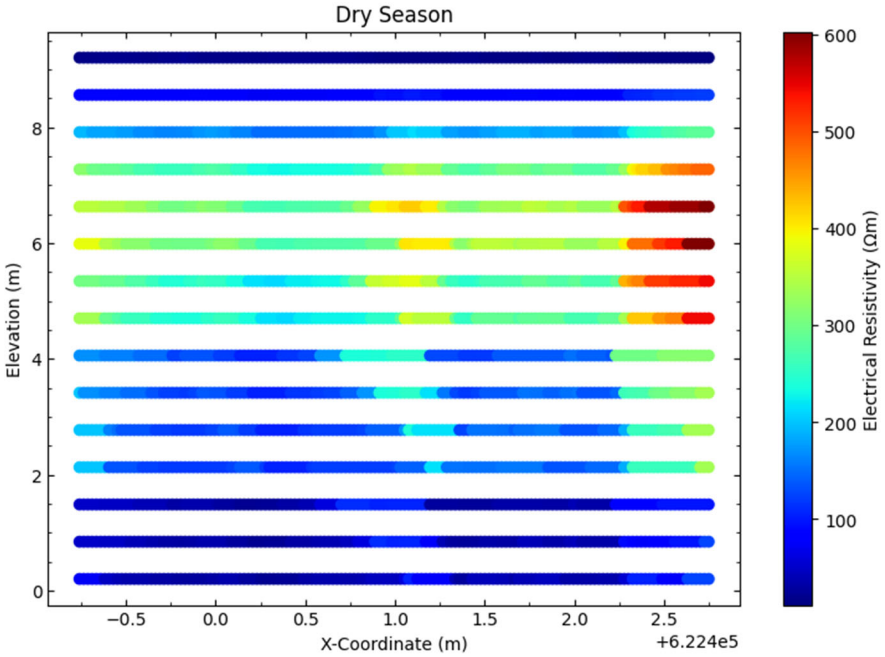


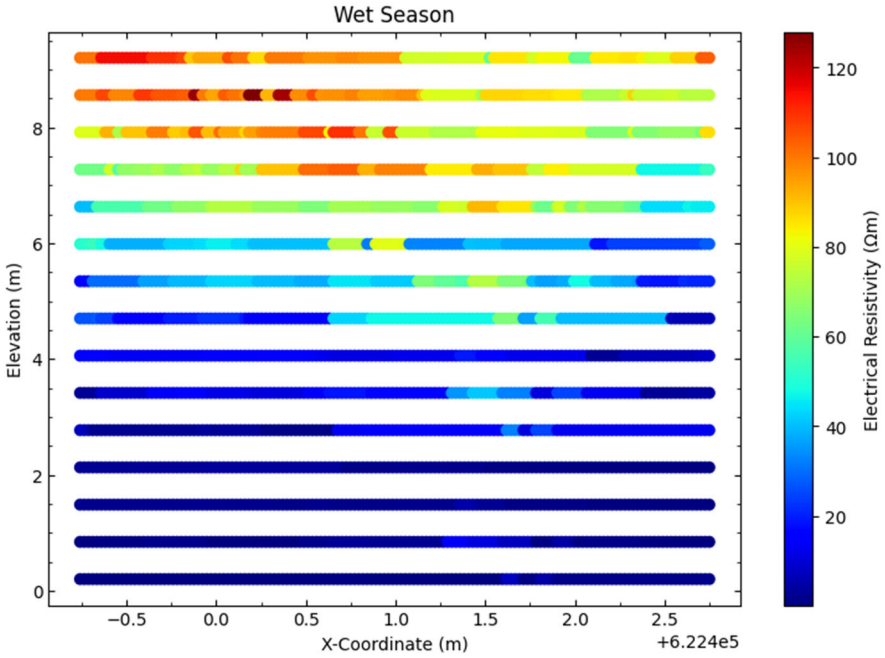
Grand Island
Site 3
Dry Crest



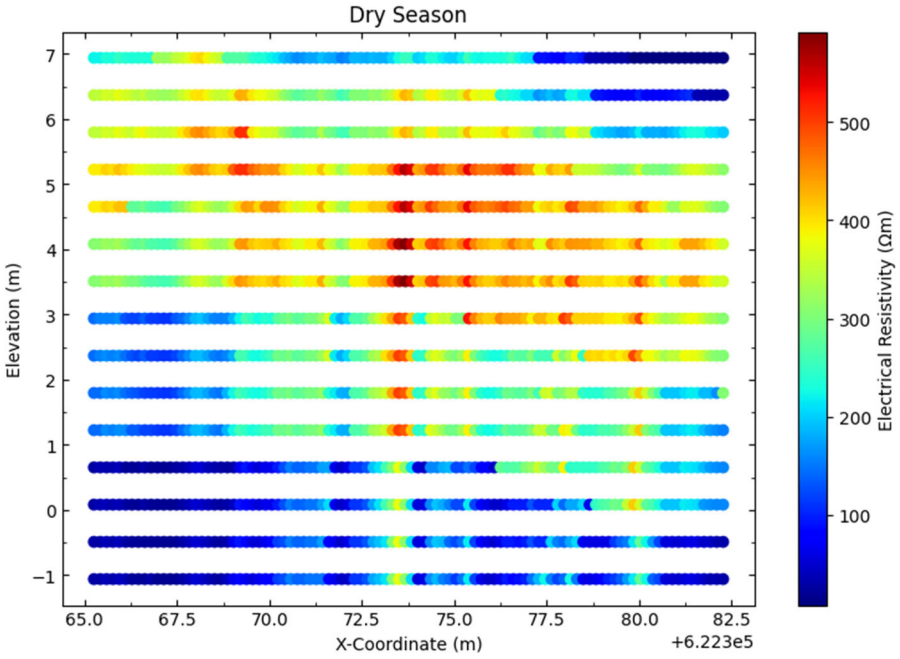


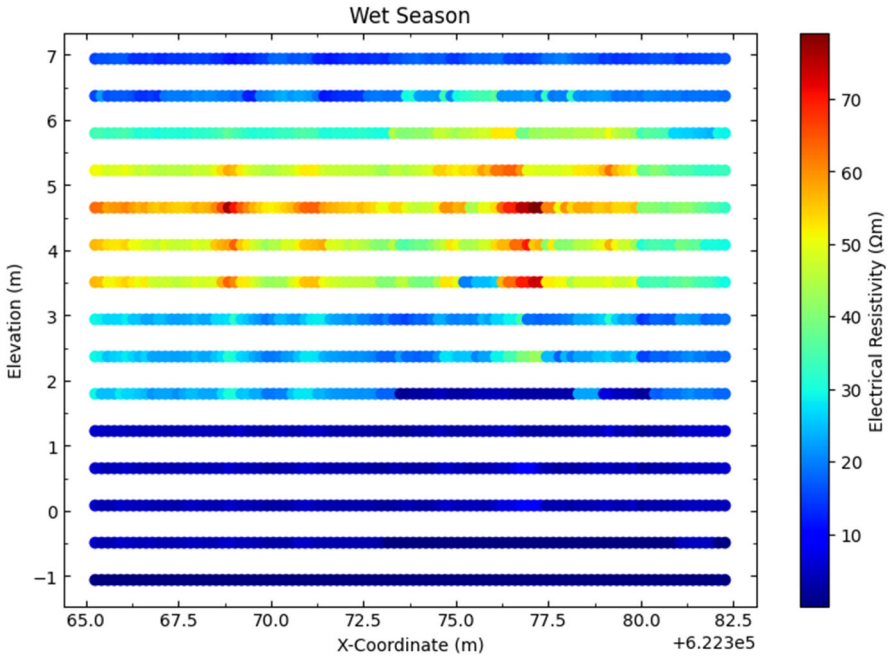
Grand Island
Site 3
Dry Mid-Slope 1



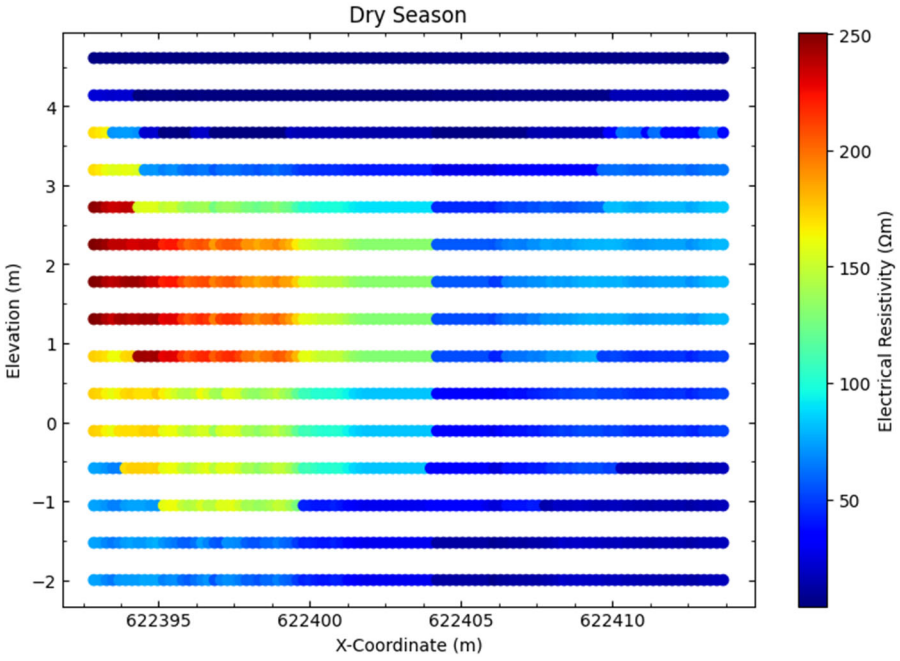


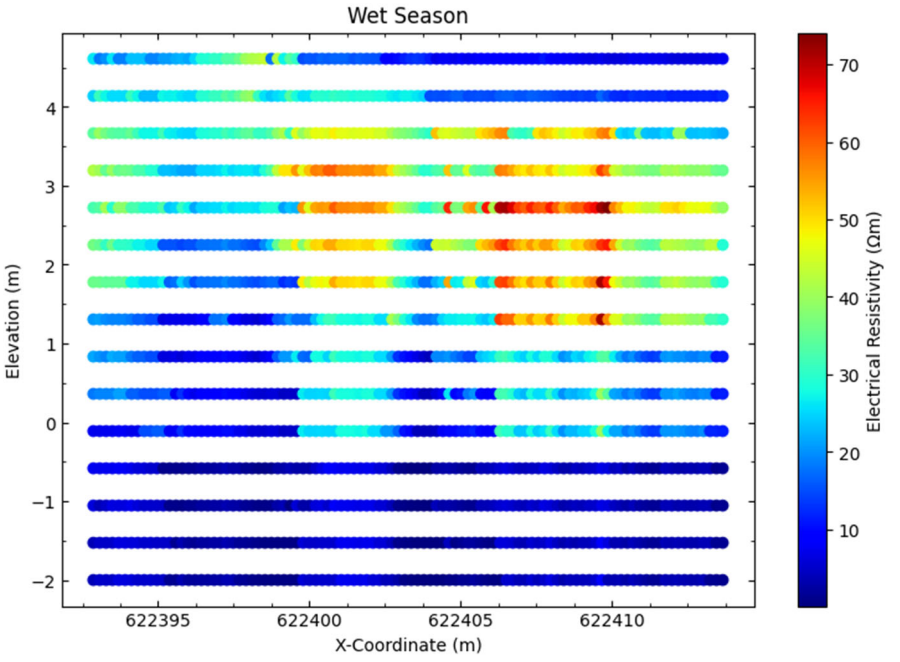
Grand Island
Site 3
Dry Mid-Slope 2



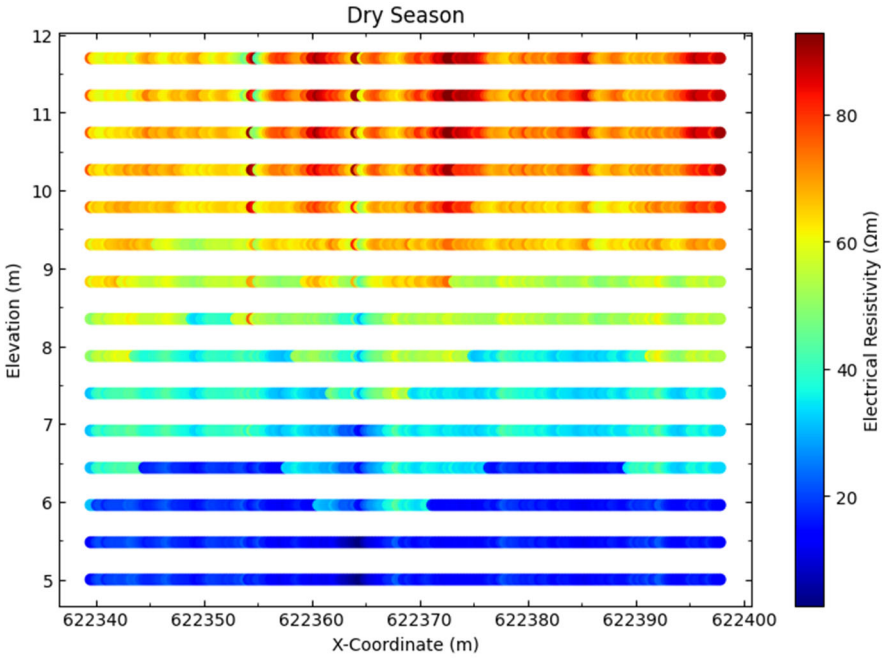


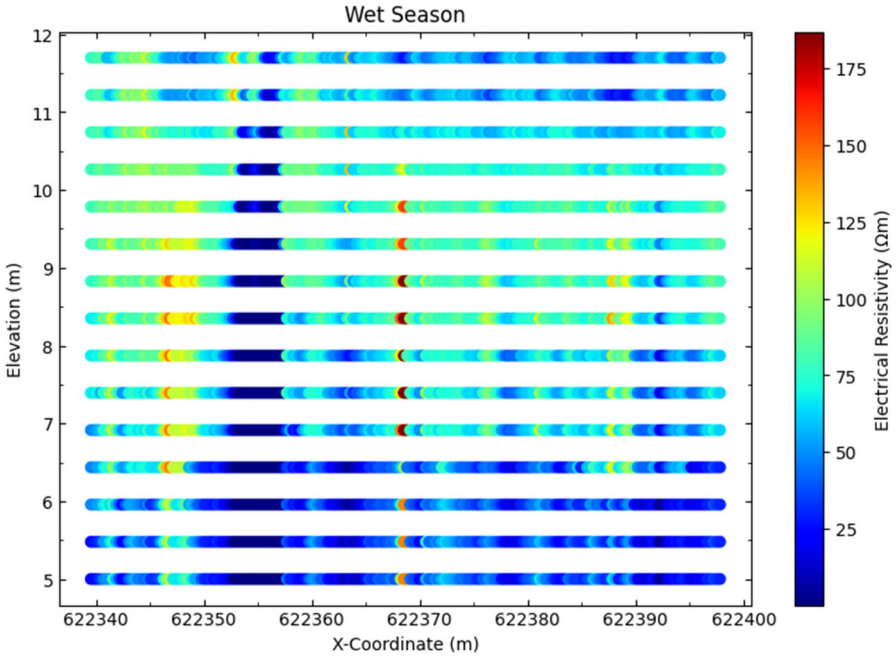
Grand Island
Site 3
Dry Toe



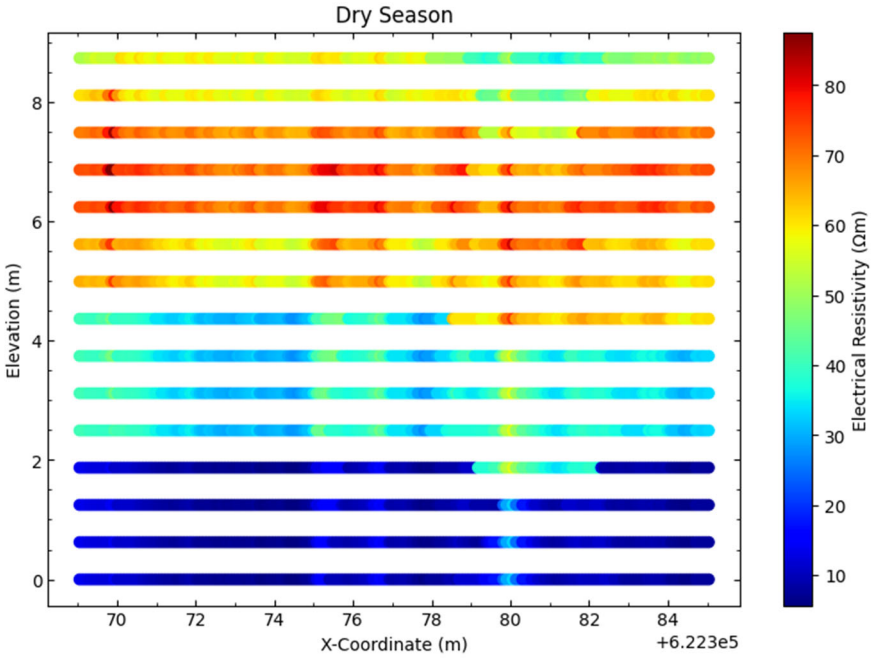


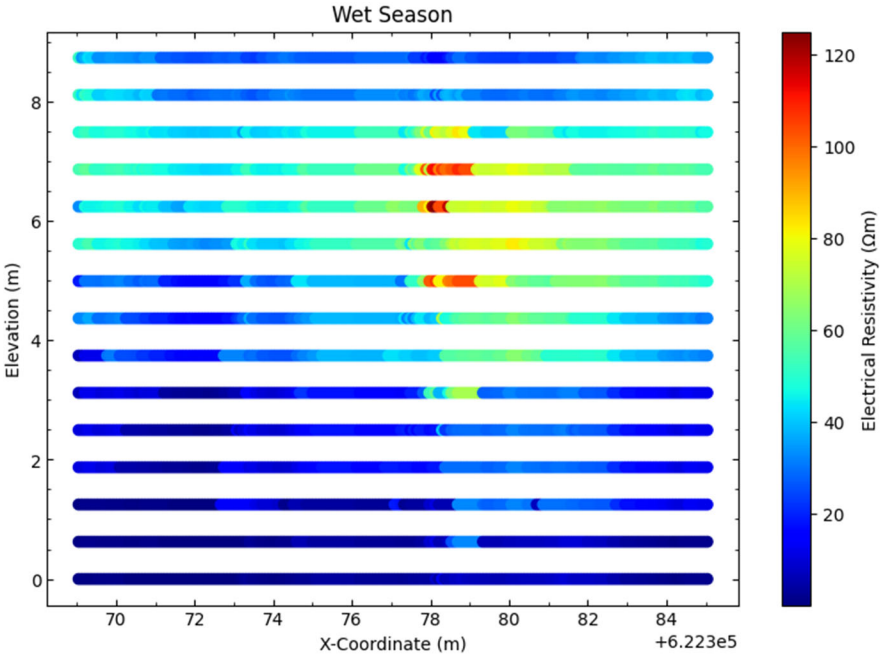
Grand Island
Site 3
Wet Crest



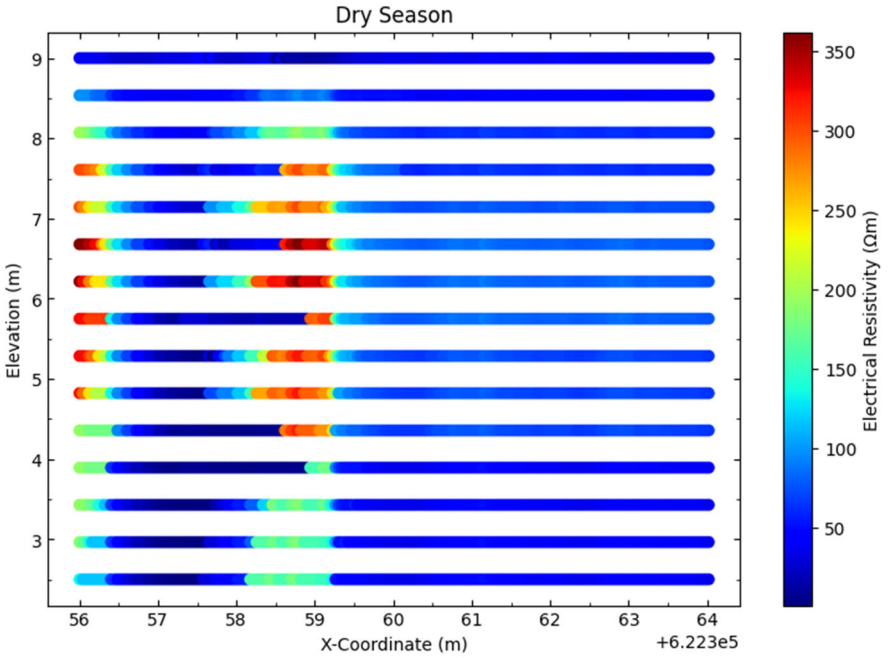


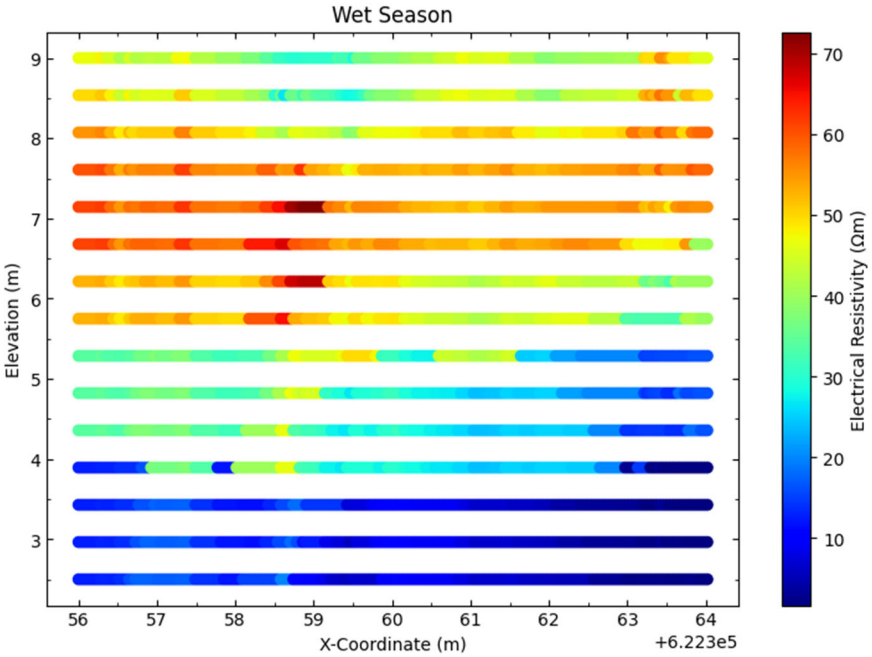
Grand Island
Site 3
Wet Toe 1



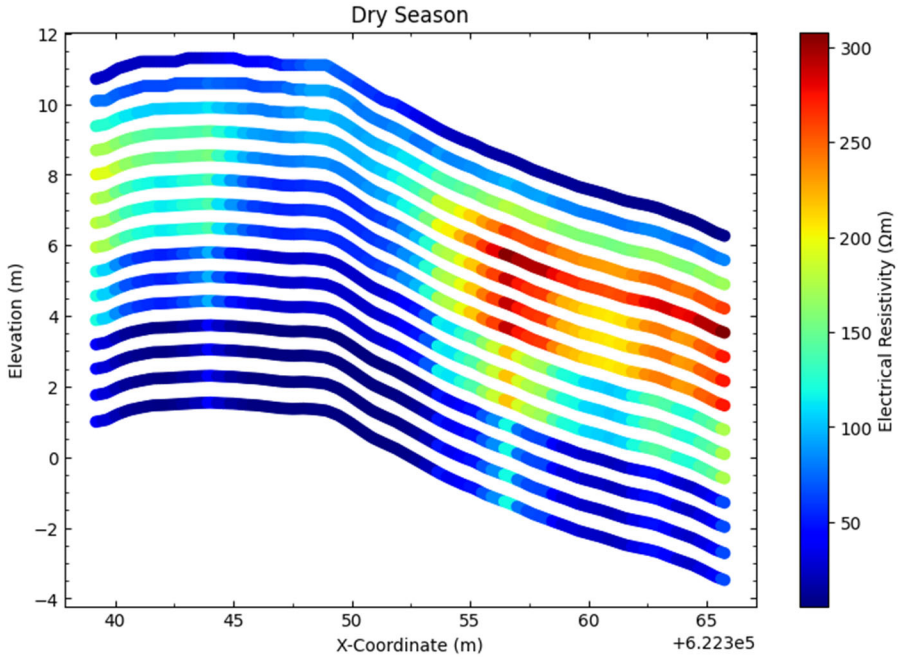


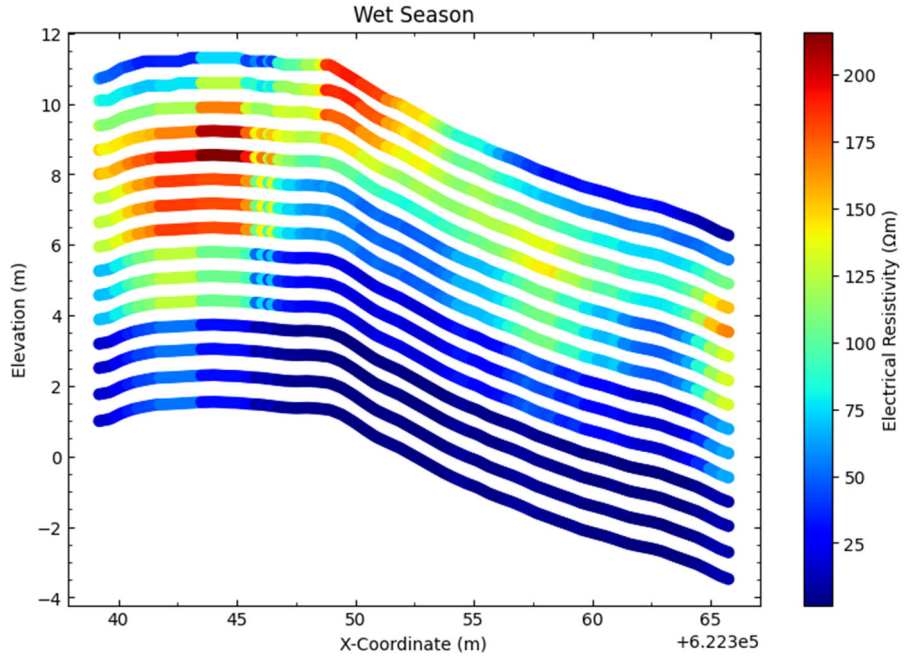
Grand Island
Site 3
Wet Toe 2





Grand Island
Site 3
Cross Section 6





Skin Depth Estimates

Table C.1. Site 1 Skin Depth Estimates for the EMI Surveys

Survey Name	September Residual	September Skin Depth	February Residual	February Skin Depth
Cross Section 1	1.72	2 m (wet toe), 5 m (crest and dry toe)	2.28	12 m
Cross Section 2	1.92	2 m (wet toe), 6 m (crest and dry toe)	2.71	11 m
Cross Section 3	1.92	6 m	2.59	8 m dry toe, 10 m, crest and wet toe
Cross Section 4	1.51	8 m (wet toe and crest), 4 m (dry toe)	2.43	12 m
Cross Section 5	1.95	8 m (wet toe), 6 m (dry toe)	2.38	10 m (dry toe and crest), 14 m (wet toe)
Cross Section 6	1.80	8 m (wet toe and crest), 4 m (dry toe)	3.15	10 m
Cross Section 7	2.14	5 m	2.92	11 m
Cross Section 8	1.59	8 m	2.95	12 m
Cross Section 9	1.74	3 m (wet toe and crest), 6 m (dry toe)	3.03	12 m
Cross Section 10	1.20	7 m (wet toe and crest), 2 m (dry toe)	2.31	9 m
Cross Section 11	2.05	4 m (wet toe), 5 m (crest and dry toe)	2.79	10 m elsewhere, 14 m (13 meters to 23 meters)
Cross Section 12	1.53	7 m	2.43	11 m
Dry Crest	1.67	7 m	2.52	11 m
Dry Mid	1.67	6 m	2.04	10 m
Dry Toe 1	1.55	6 m	2.41	8 m
Dry Toe 2	1.41	8 m	-	-
Wet Crest	1.92	10 m	1.77	14 m
Wet Mid 1	1.58	8 m	1.87	12 m
Wet Mid 2	1.38	8 m	-	-
Wet Toe1	1.29	8 m	0.65	3 m (until 35 m), 6 m (until the end)
Wet Toe2	-	-	0.88	14 m
Wet Toe 3	-	-	1.85	11 m

Table C.2. Site 2 Skin Depth Estimates for the EMI Surveys

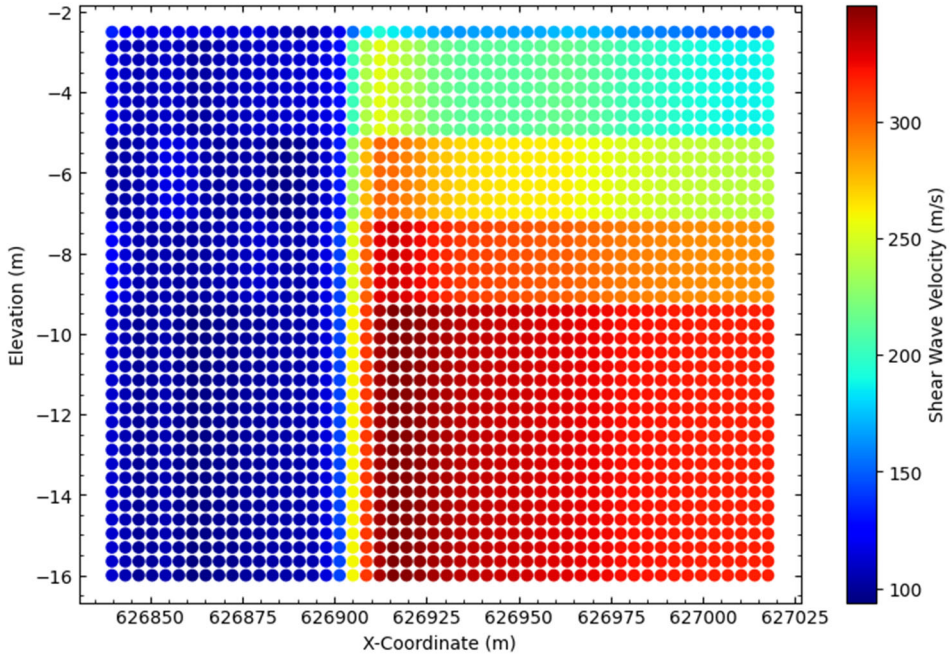
Survey Name	September Residual	September Skin Depth	February Residual	February Skin Depth
Cross Section 1	1.60	10 m	3.24	10 m
Cross Section 2	2.56	7 m	3.01	9 m
Cross Section 3	1.89	9 m	3.01	9 m
Cross Section 4	1.60	Dry toe (4 m), 10 m (crest and wet toe)	2.90	8 m
Cross Section 5	1.70	8 m	2.85	8 m
Cross Section 6	1.47	10 m	3.57	12 m
Cross Section 7	2.18	8 m	2.43	8 – 12 m
Cross Section 8	1.74	8 m (wet toe), 14 m (dry toe and crest)	-	-
Dry Crest 1	1.21	Mostly 2 m, few locations upwards of 14 m	2.68	11 m
Dry Crest 2	-	-	2.72	11 m
Dry Mid	0.85	Mostly 2 m, few locations upwards of 14 m	-	-
Dry Toe 1	1.21	2 - 10 m	2.59	11 m
Wet Crest	1.64	14 m	2.03	11 m
Wet Toe	1.24	6 - 10 m	2.24	8 m

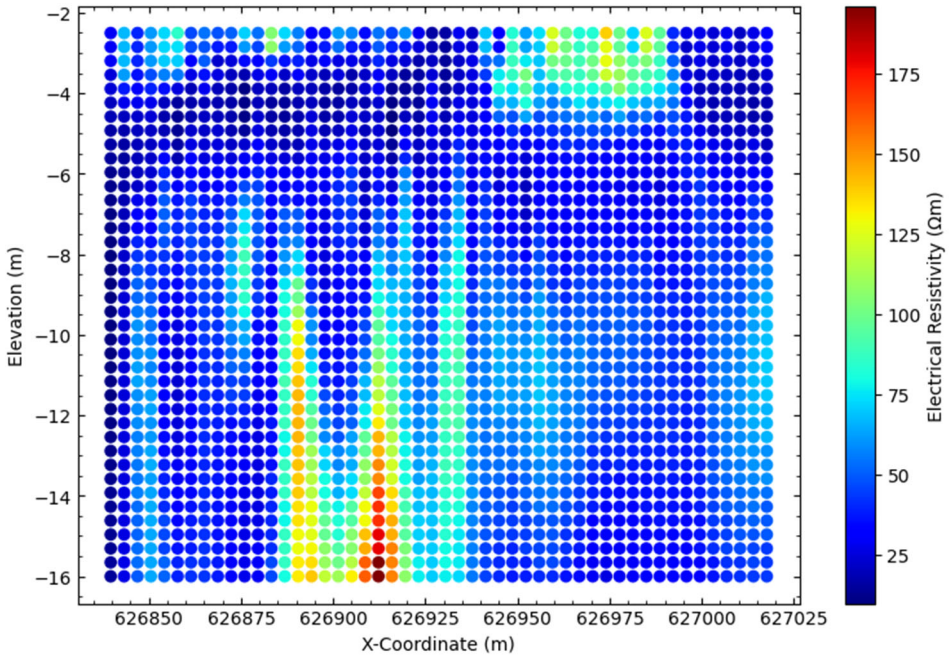
Table C.3. Site 3 Skin Depth Estimates for the EMI Surveys

Survey Name	September Residual	September Skin Depth	February Residual	February Skin Depth
Cross Section 1	1.52	8 m (dry toe), 10 m (crest and wet toe)	-	-
Cross Section 2	2.14	8 m	-	-
Cross Section 3	2.20	8 m	-	-
Cross Section 4	2.41	10 m	-	-
Cross Section 5	2.25	10 m	-	-
Cross Section 6	2.32	8 m	2.57	11 m
Dry Crest	1.53	10 m	2.59	8 m
Dry Mid 1	2.70	10 m	1.92	9 m
Dry Mid 2	2.32	10 m	-	-
Dry Mid 3	2.79	10 m	-	-
Dry Toe 1	1.91	10 m	1.81	9 m
Dry Toe 2	1.32	9 m	2.17	8 m
Wet Crest	1.51	8 m	2.57	11 m, 4 m (short distance)
Wet Toe	1.56	12 m	1.89	8 m
Wet Toe 1	-	-	1.75	10 m
Wet Toe 2	-	-	1.91	9 m
Wet Toe 3	-	-	1.96	8 m

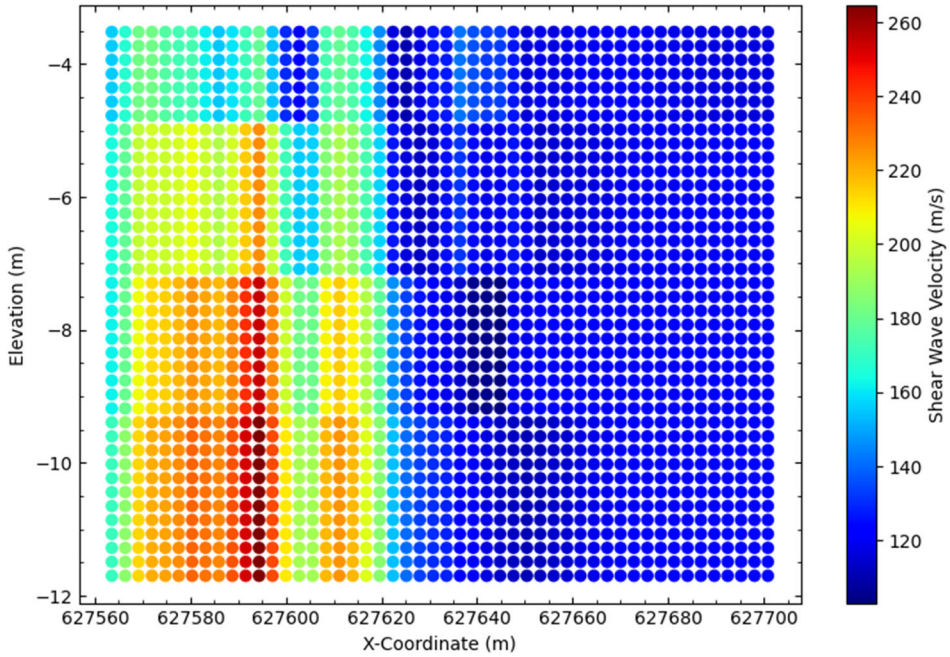
Bacon Island Geophysical Profiles

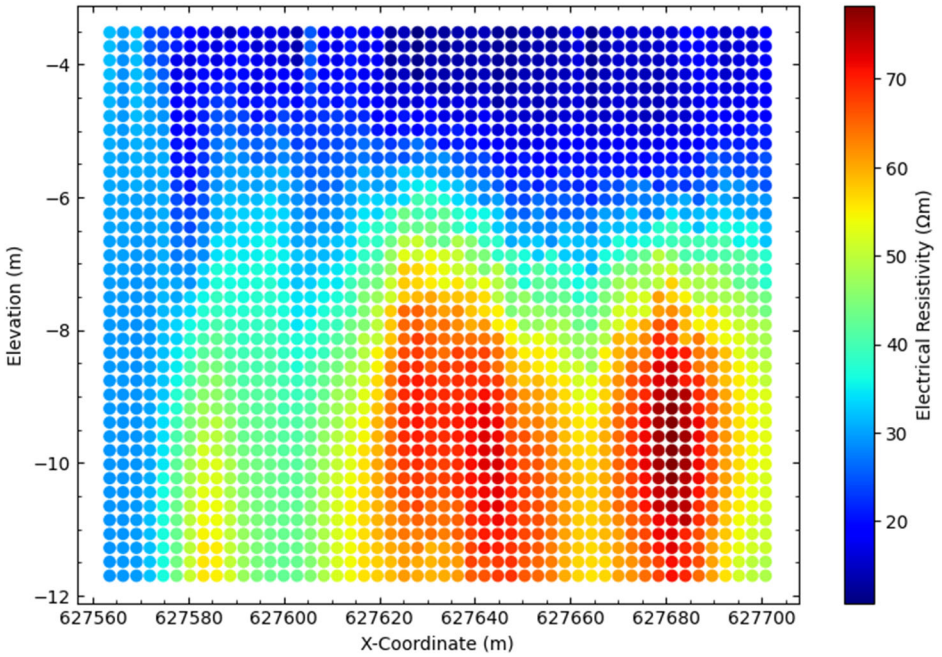
Bacon Island
Northwest Site



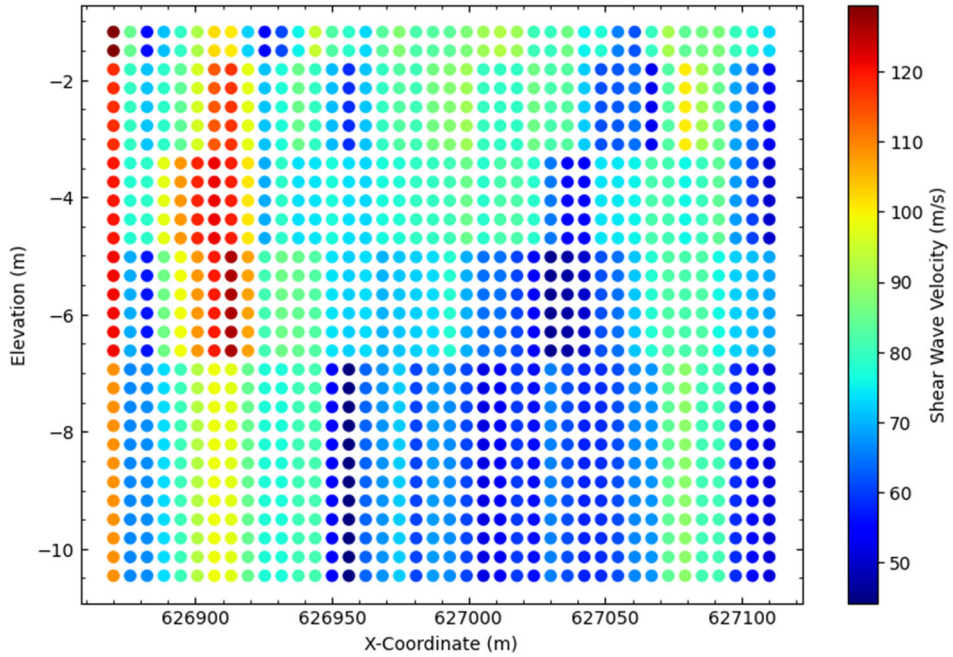


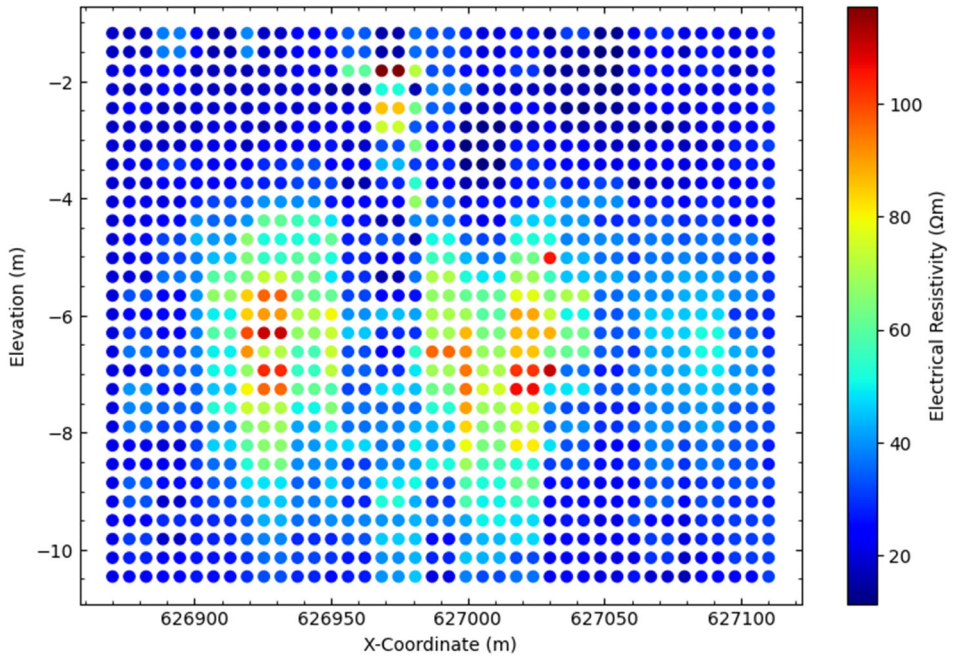
Bacon Island
Northcentral Site





Bacon Island
South Site





Appendix D – Remote Sensing Profiles

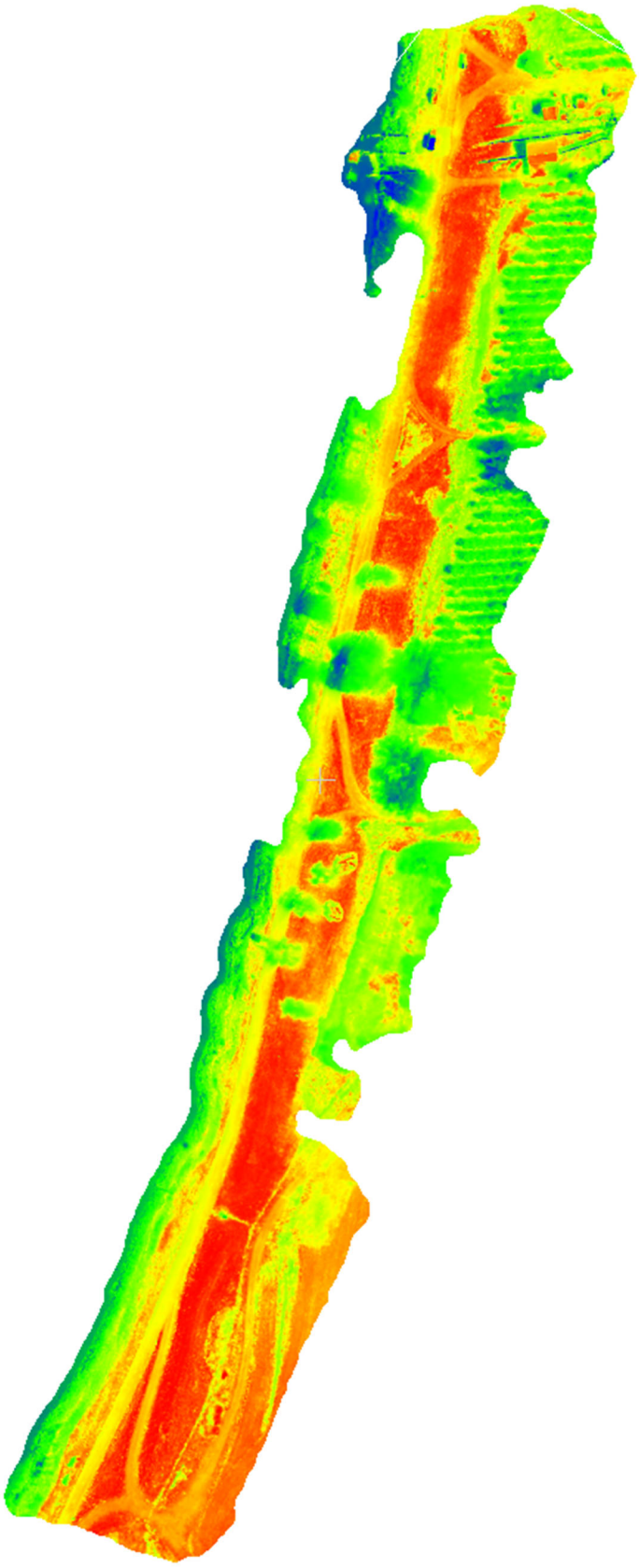
Grand Island
Site 1
3D Textured Mesh - Optical
Dry Season



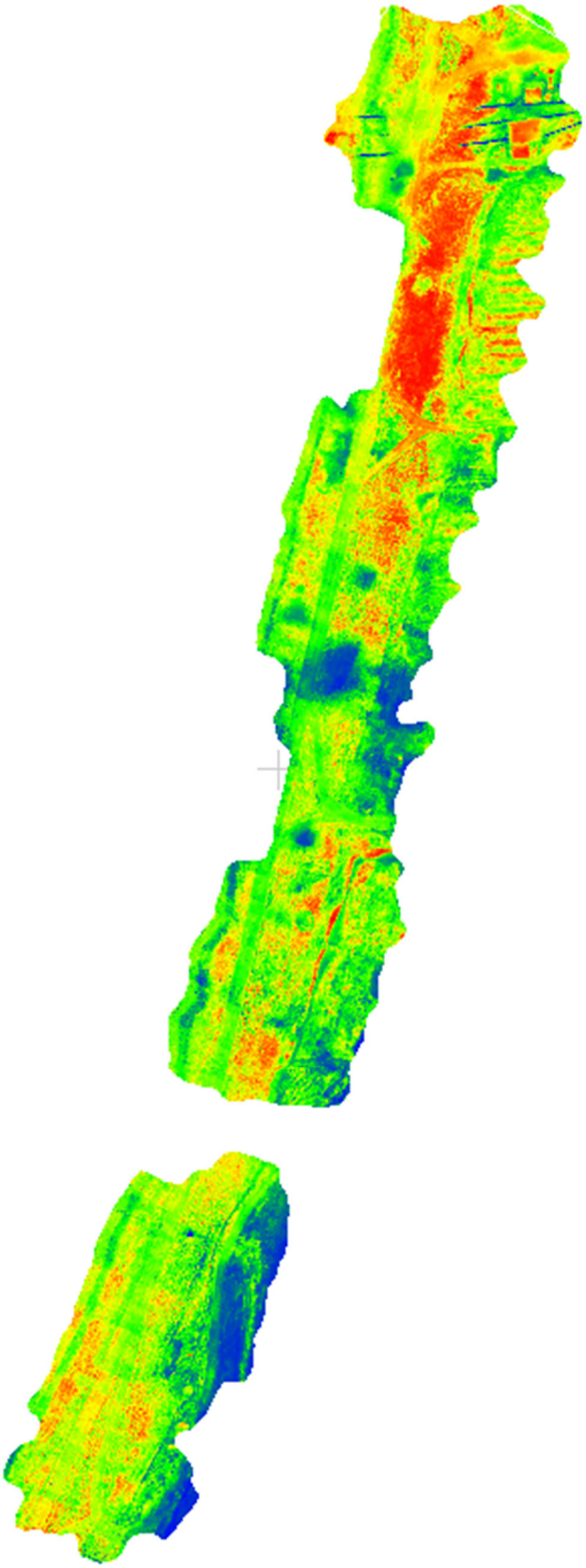
Grand Island
Site 1
3D Textured Mesh - Optical
Wet Season



Grand Island
Site 1
3D Textured Mesh – Thermal
Dry Season



Grand Island
Site 1
3D Textured Mesh – Thermal
Wet Season



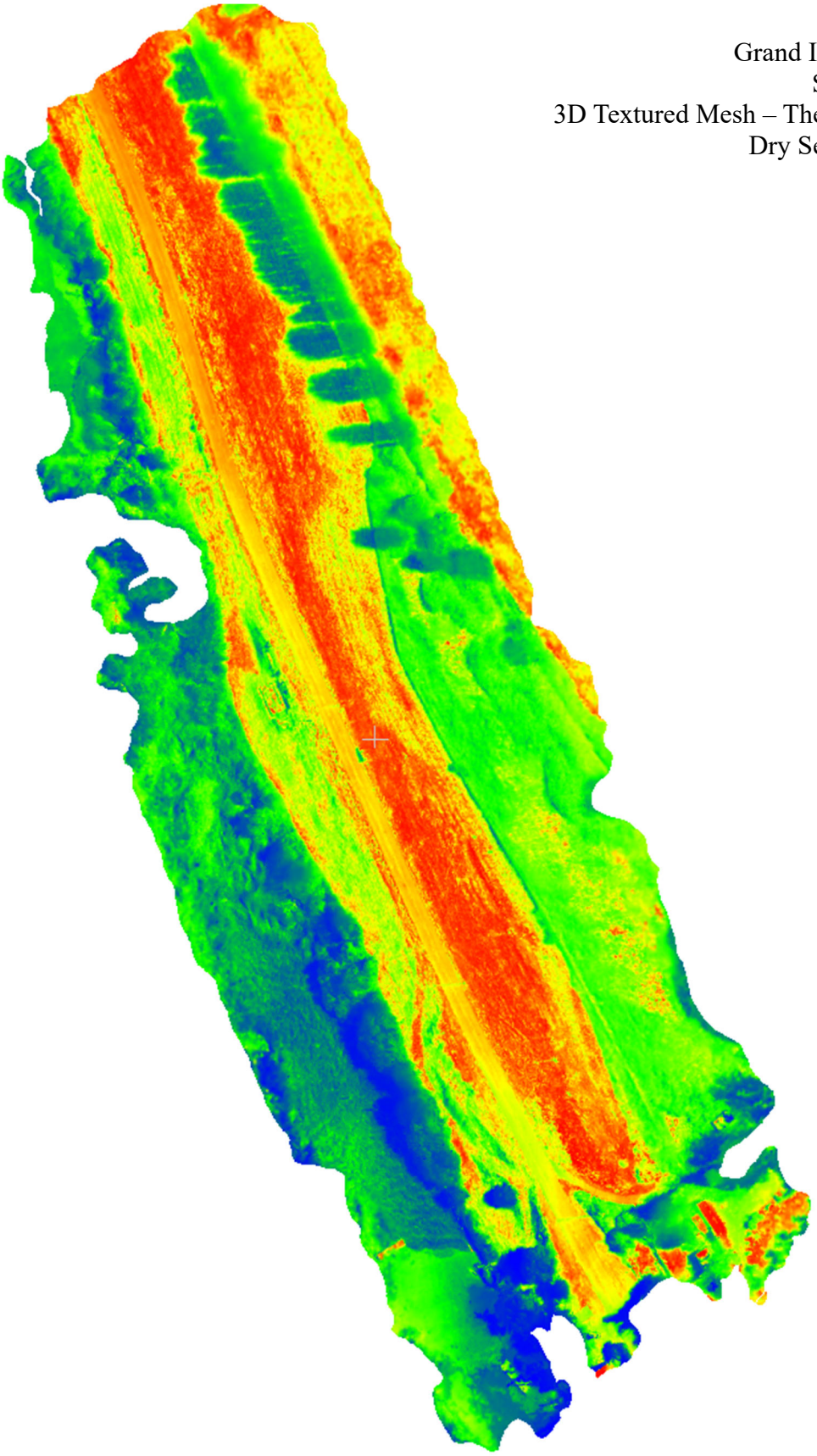
Grand Island
Site 2
3D Textured Mesh – Optical
Dry Season



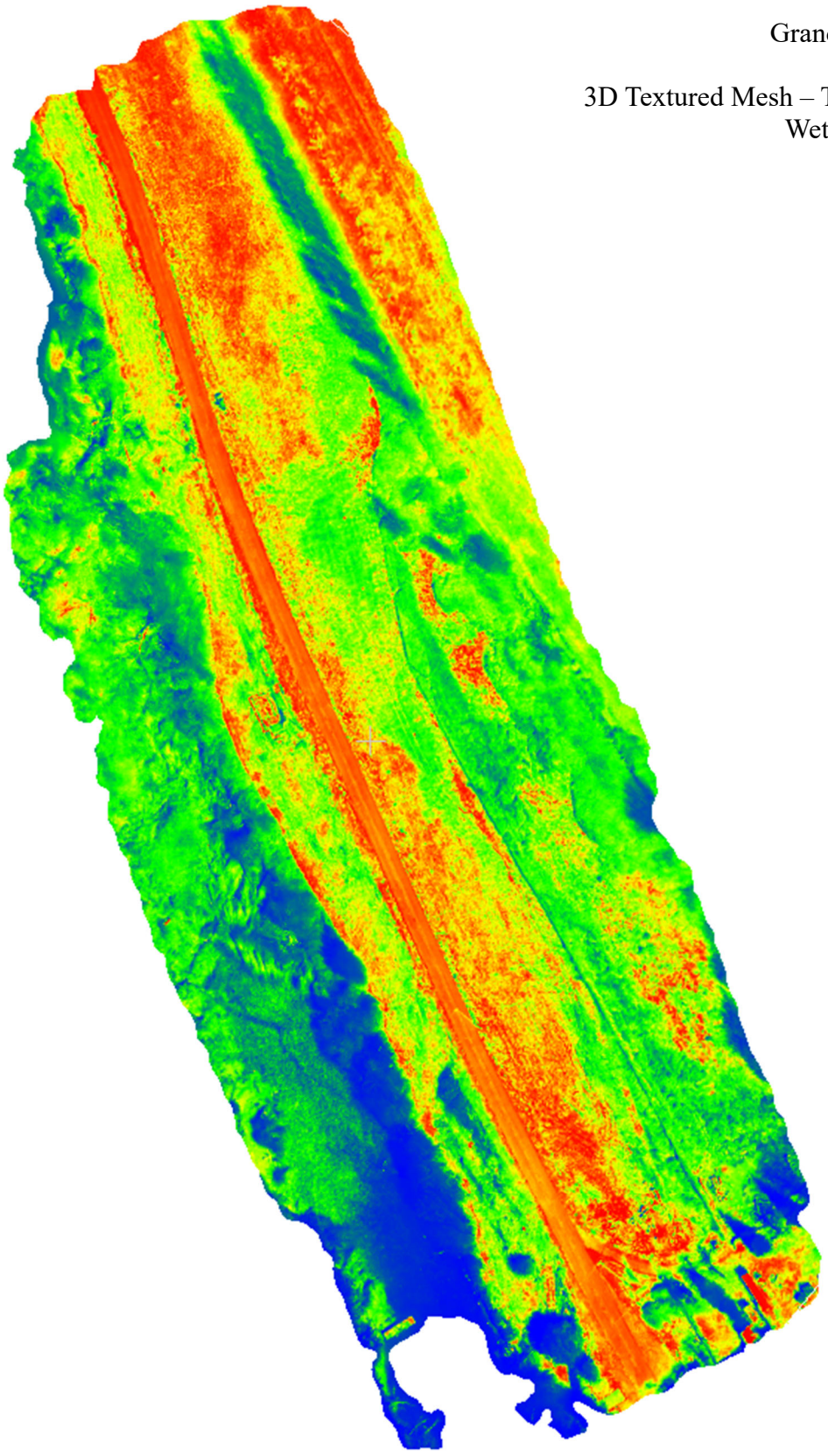
Grand Island
Site 2
3D Textured Mesh – Optical
Wet Season



Grand Island
Site 2
3D Textured Mesh – Thermal
Dry Season



Grand Island
Site 2
3D Textured Mesh – Thermal
Wet Season



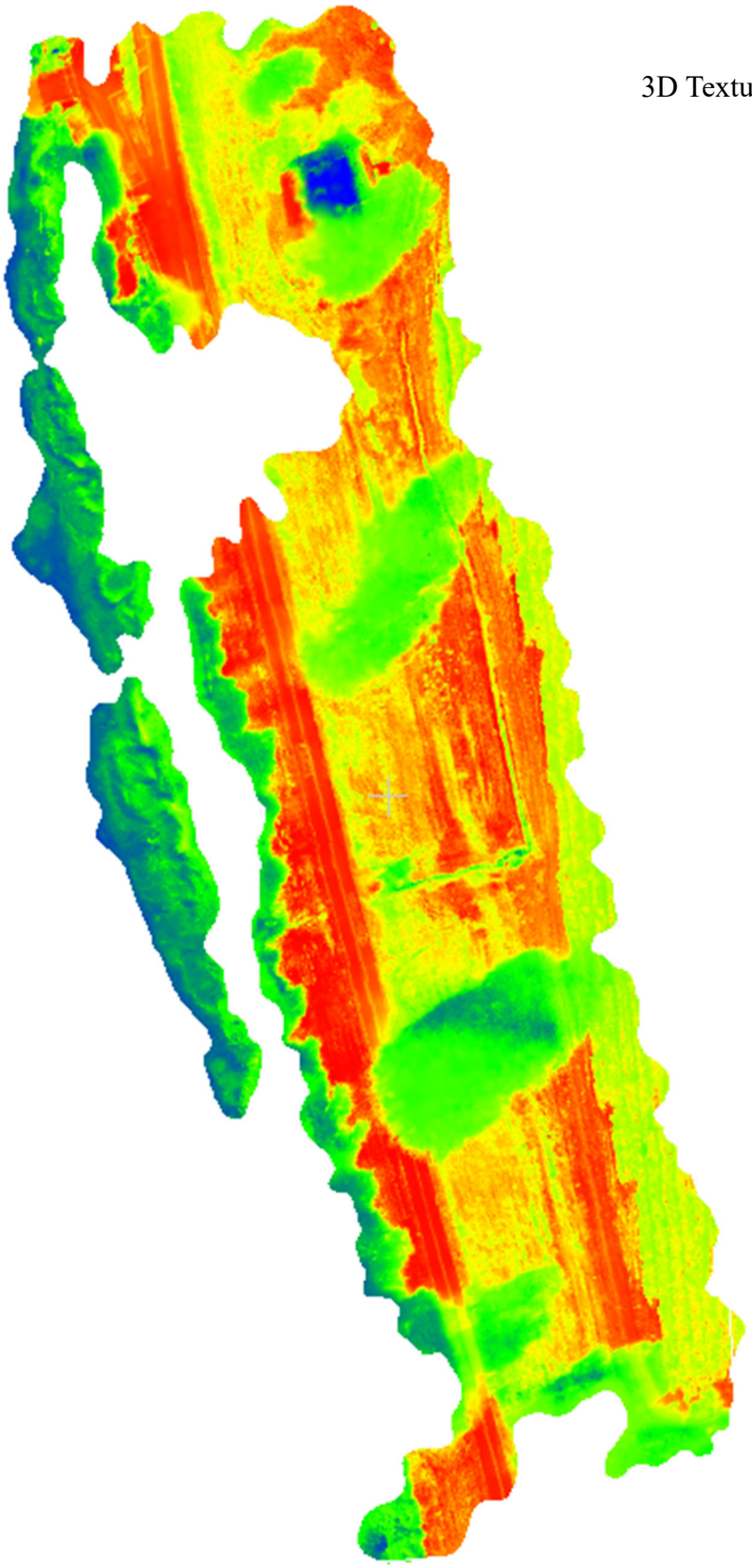


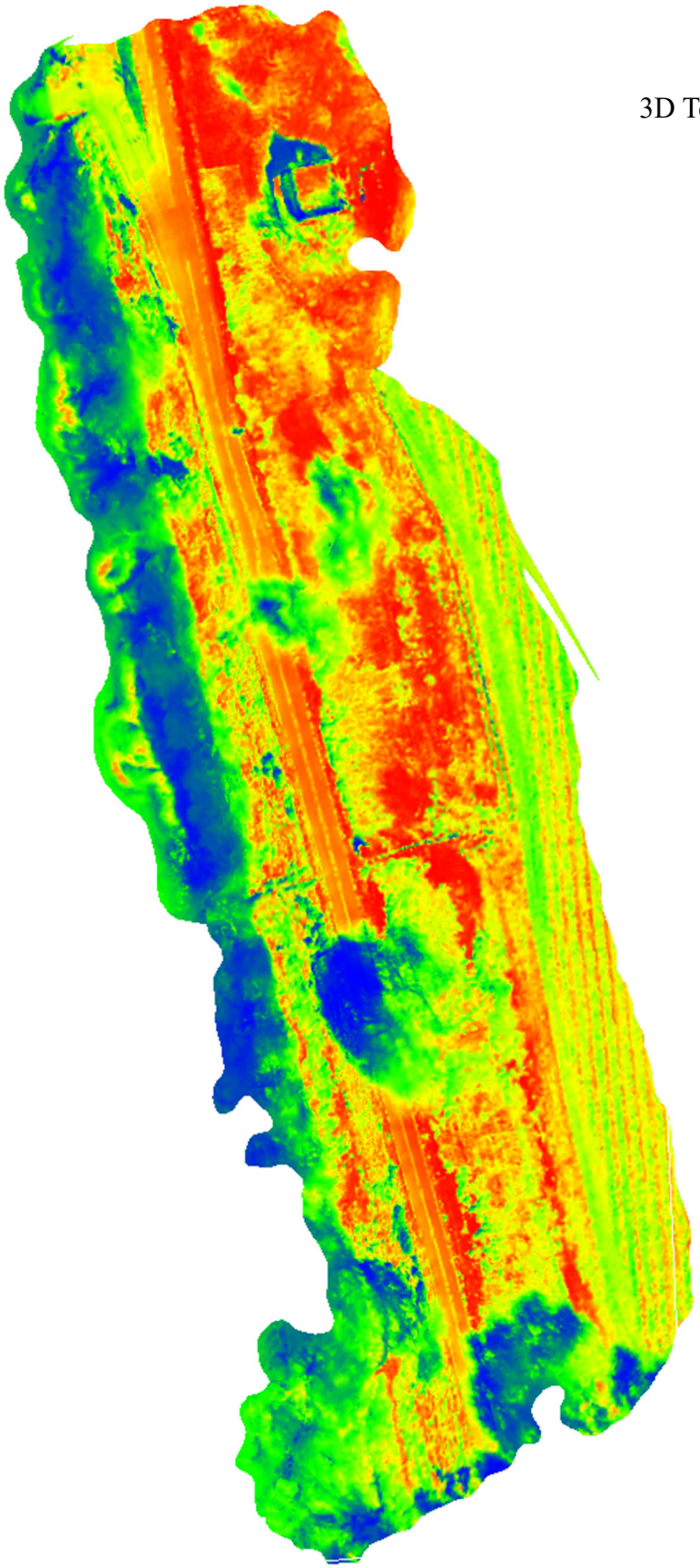
Grand Island
Site 3
3D Textured Mesh – Optical
Dry Season



Grand Island
Site 3
3D Textured Mesh – Optical
Wet Season

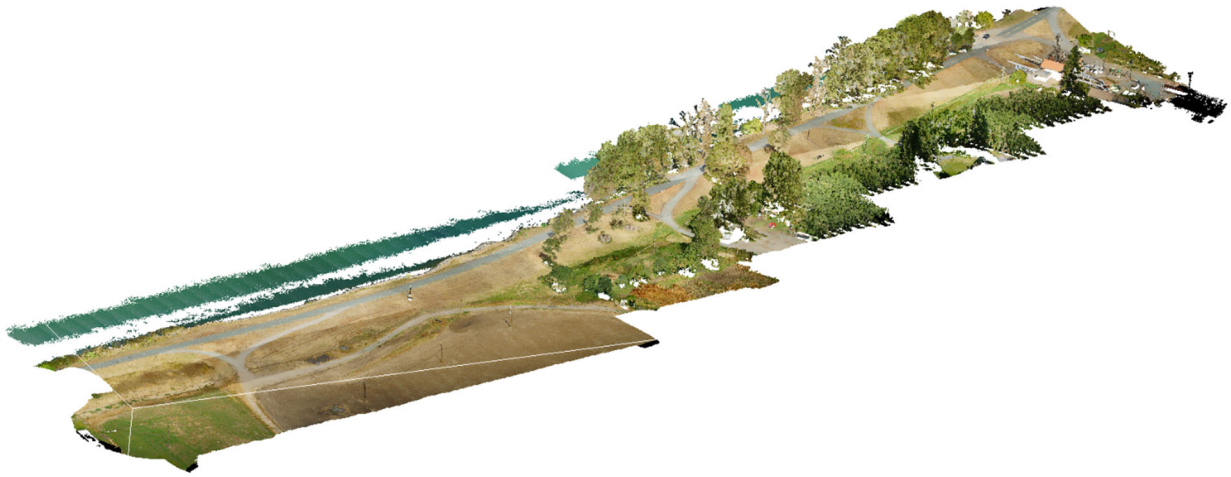
Grand Island
Site 3
3D Textured Mesh – Thermal
Dry Season



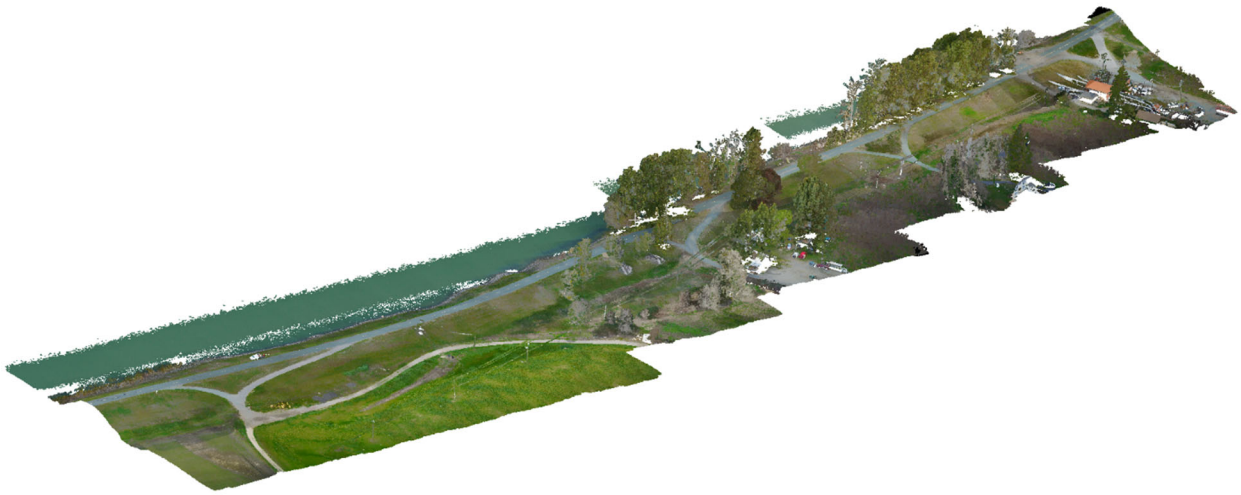


Grand Island
Site 3
3D Textured Mesh – Thermal
Wet Season

Grand Island
Site 1
3D Point Cloud – LIDAR
Dry Season



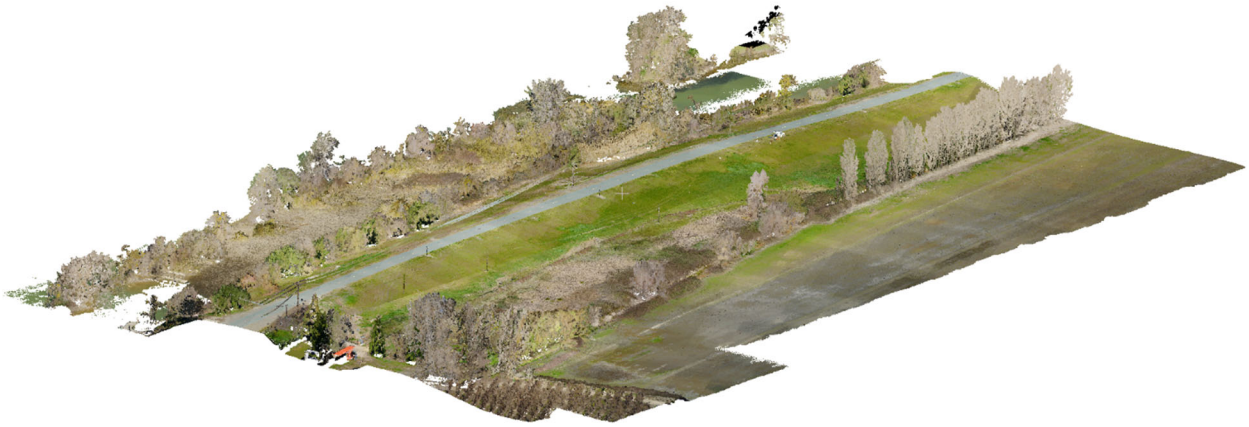
Grand Island
Site 1
3D Point Cloud – LIDAR
Wet Season



Grand Island
Site 2
3D Point Cloud – LIDAR
Dry Season



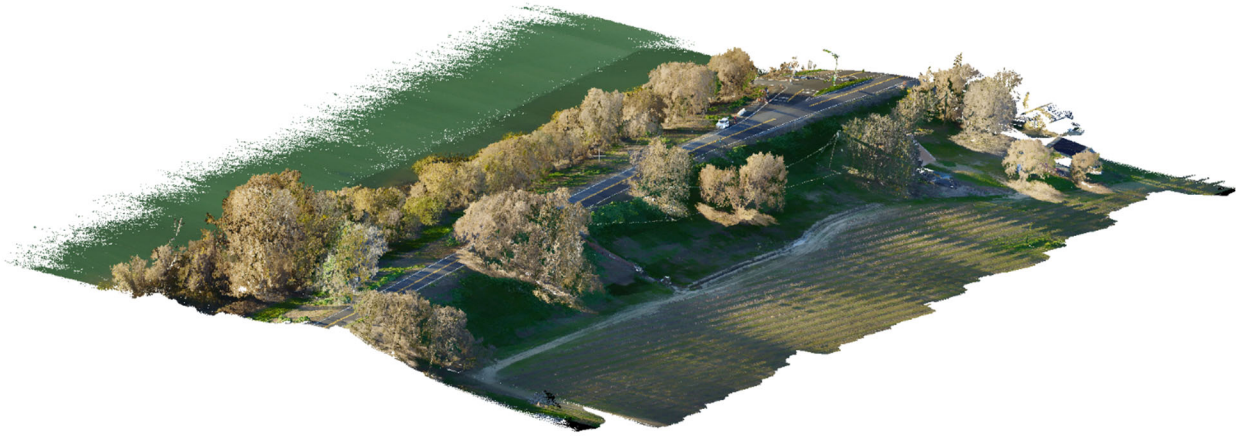
Grand Island
Site 2
3D Point Cloud – LIDAR
Wet Season



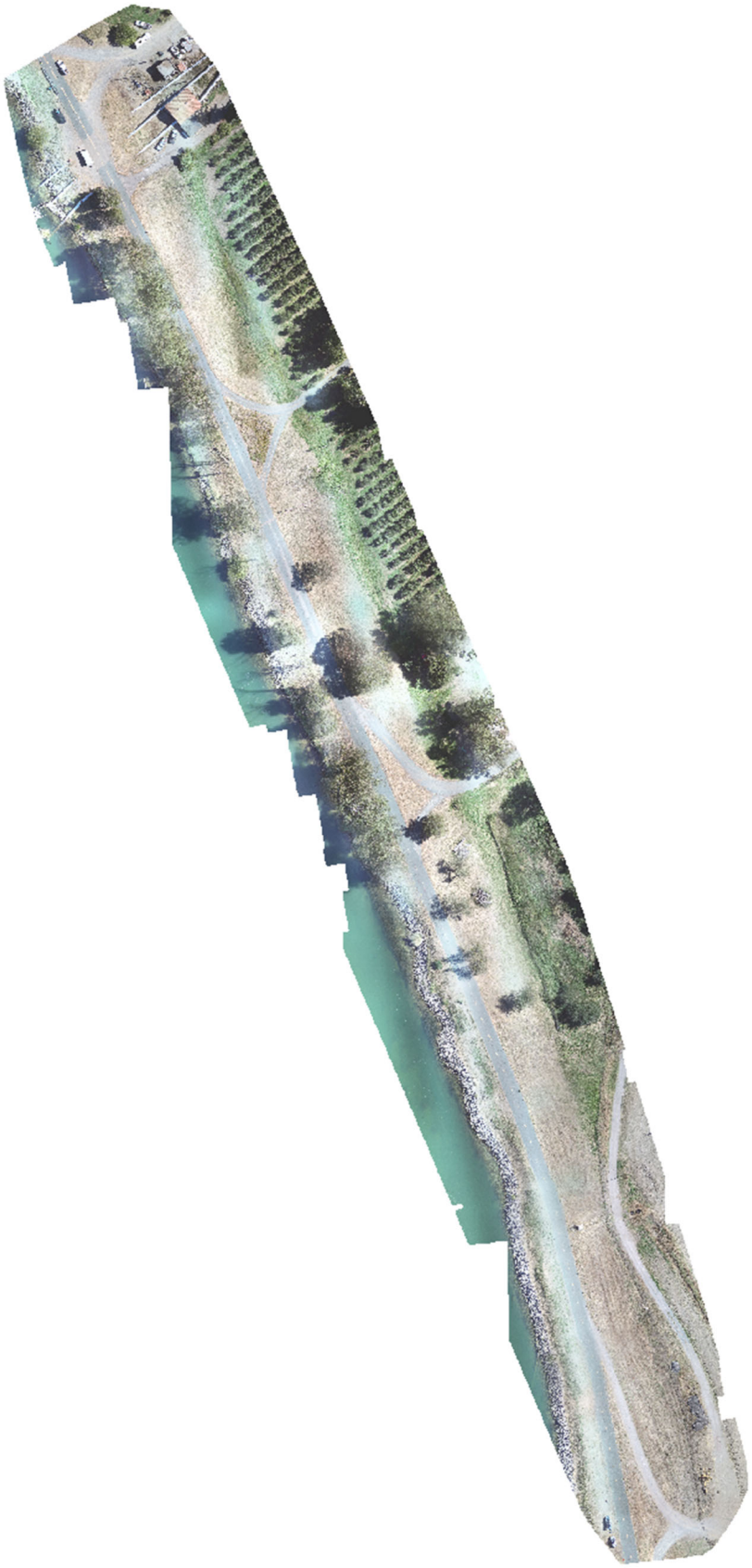
Grand Island
Site 3
3D Point Cloud – LIDAR
Dry Season



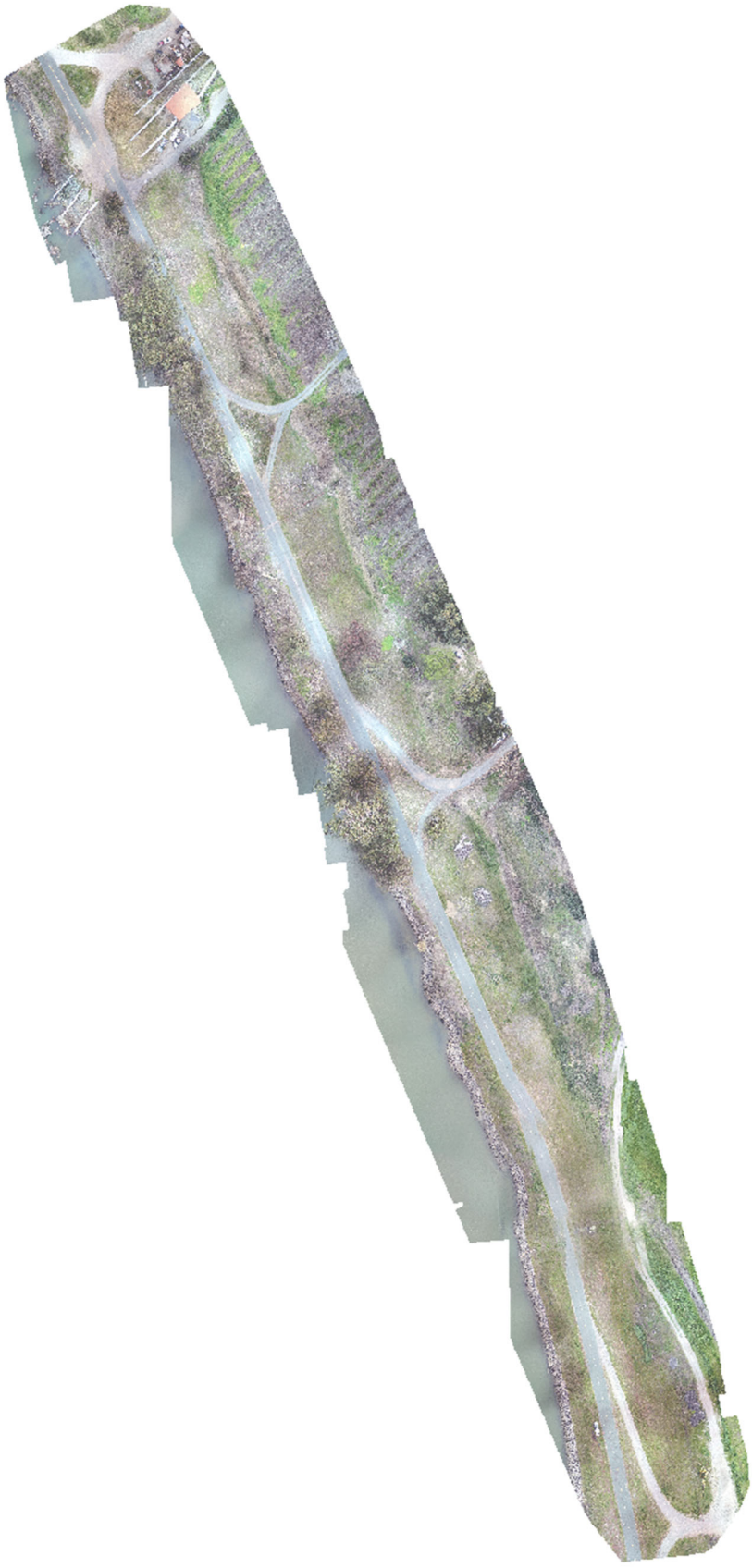
Grand Island
Site 3
3D Point Cloud – LIDAR
Wet Season



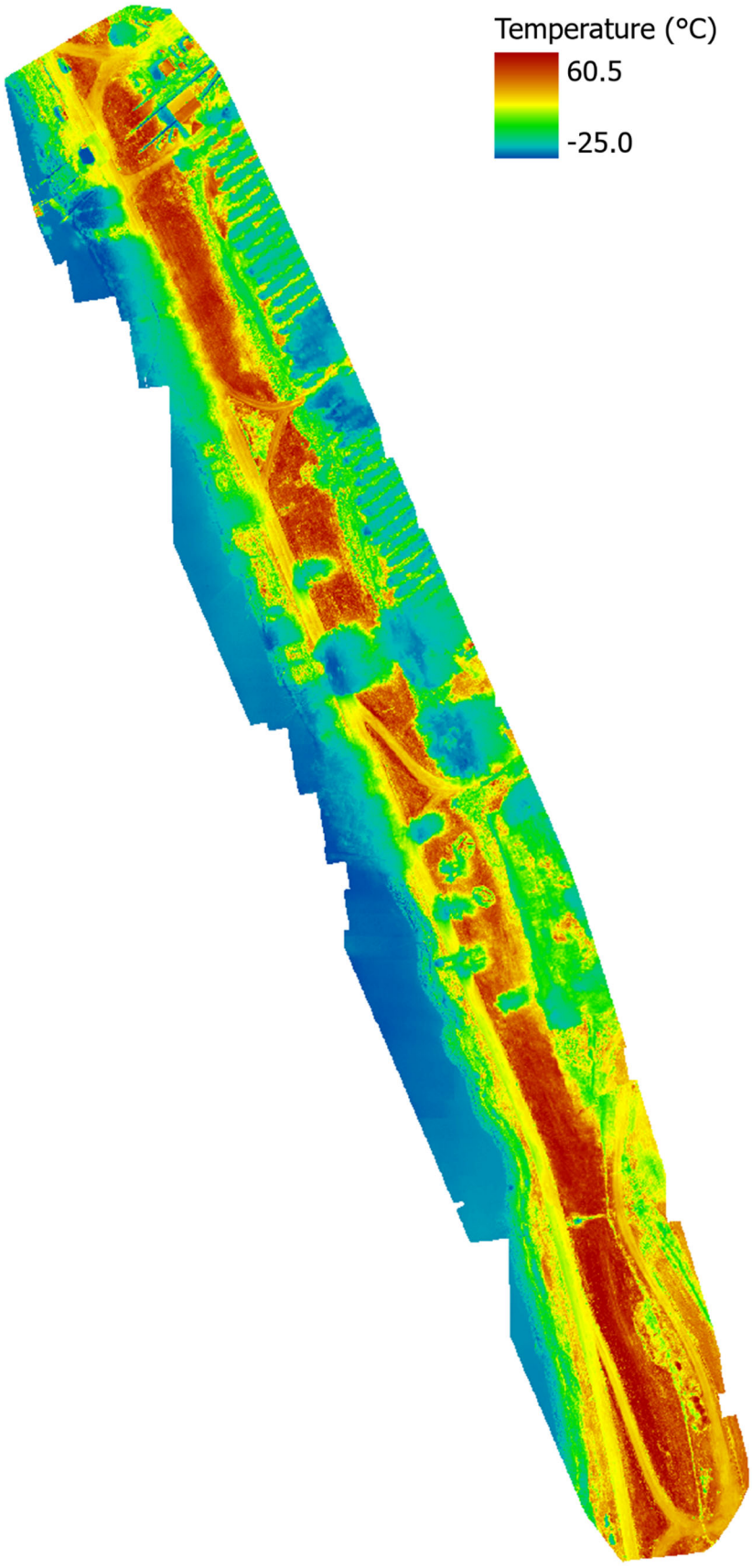
Grand Island
Site 1
Orthomosaic – Optical
Dry Season



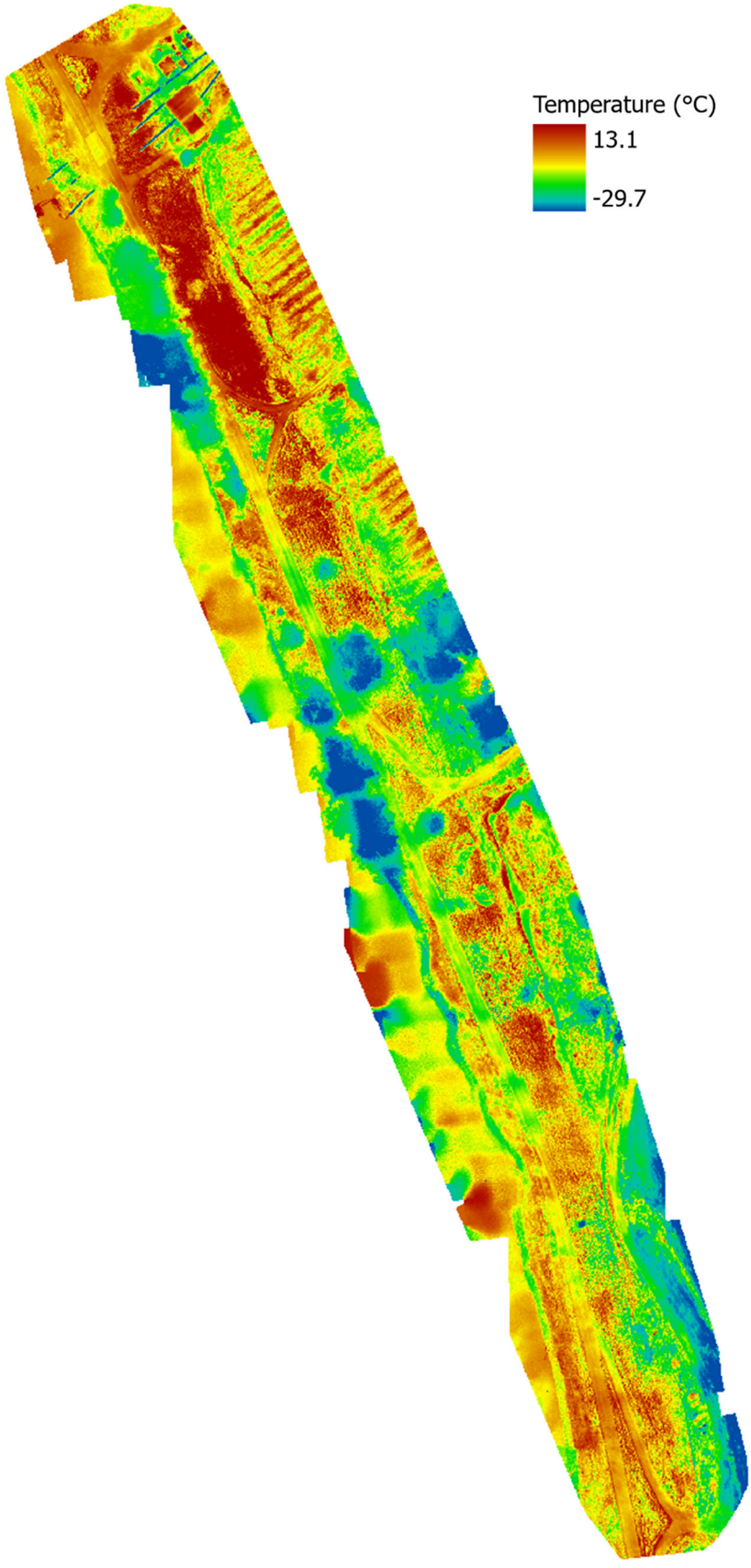
Grand Island
Site 1
Orthomosaic – Optical
Wet Season



Grand Island
Site 1
Orthomosaic – Thermal
Dry Season



Grand Island
Site 1
Orthomosaic – Thermal
Wet Season



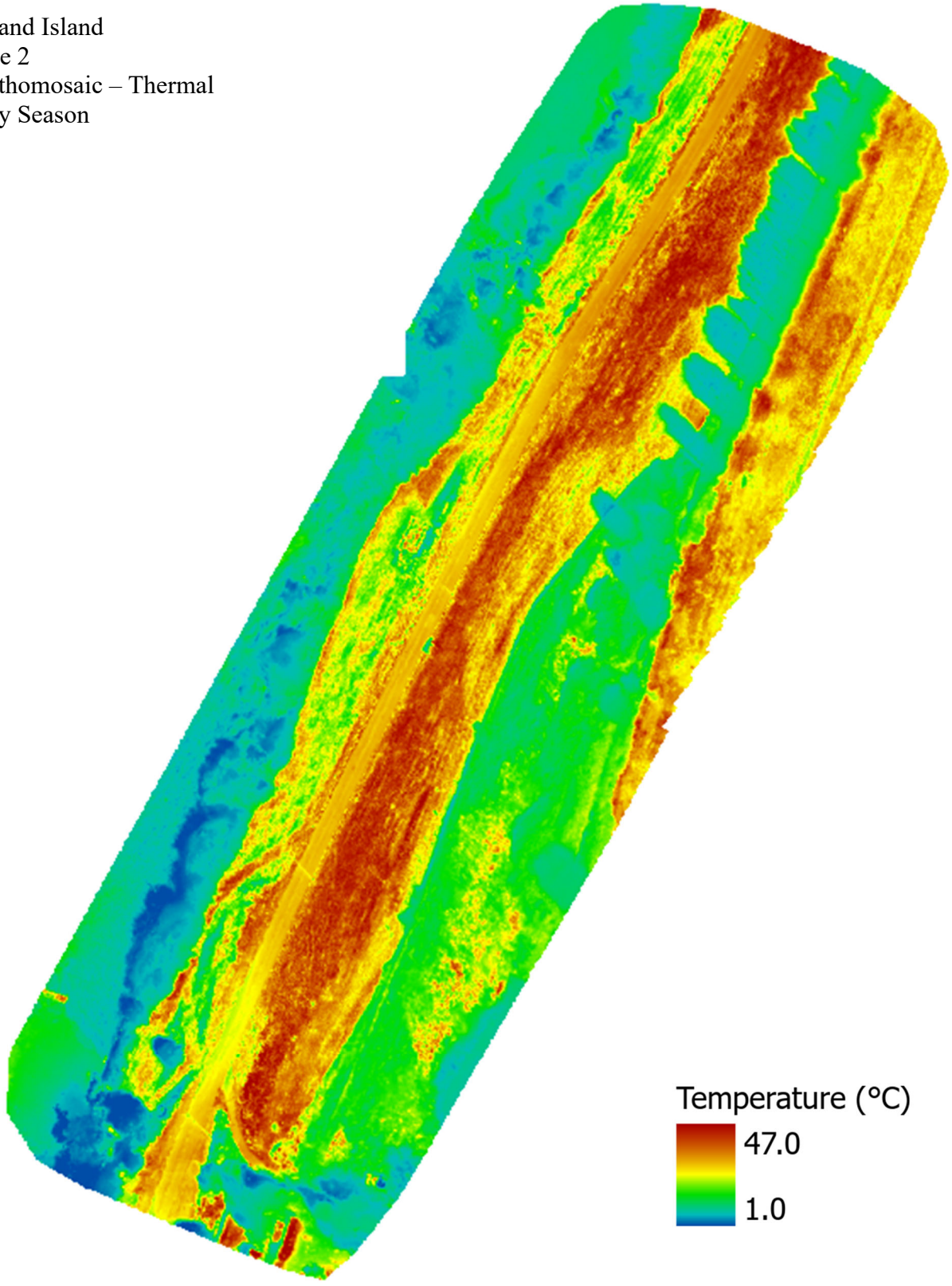
Grand Island
Site 2
Orthomosaic – Optical
Dry Season



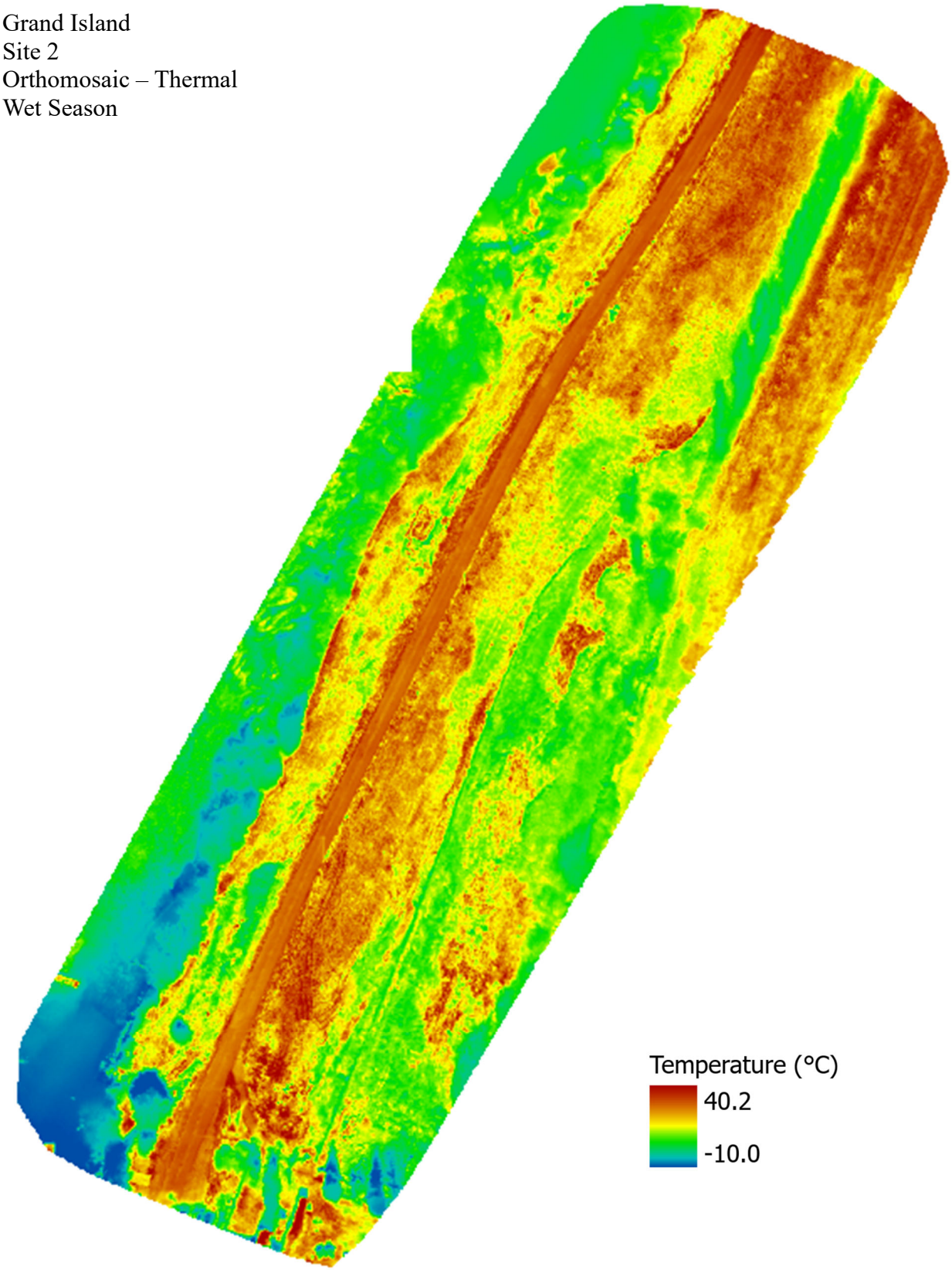
Grand Island
Site 2
Orthomosaic – Optical Wet Season



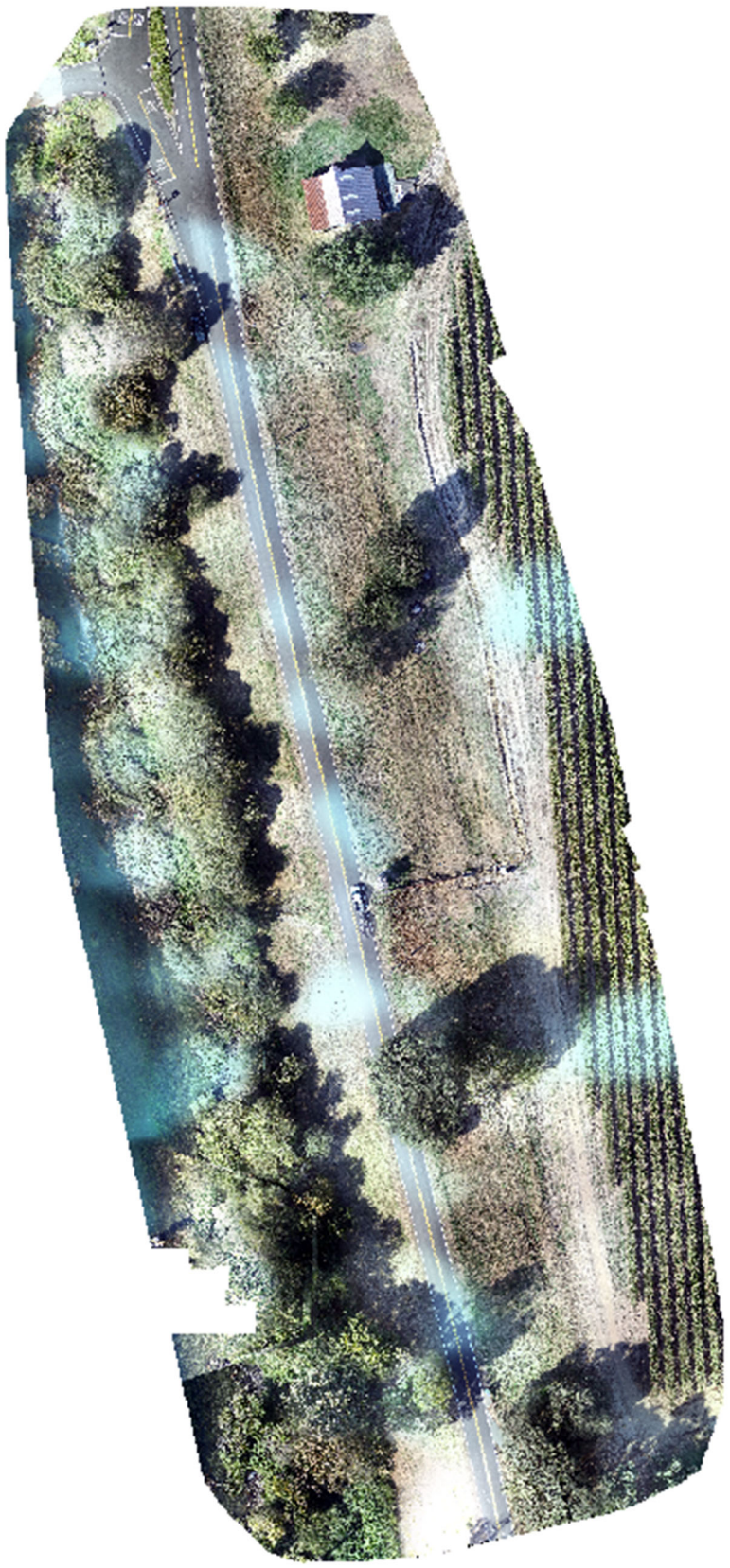
Grand Island
Site 2
Orthomosaic – Thermal
Dry Season



Grand Island
Site 2
Orthomosaic – Thermal
Wet Season



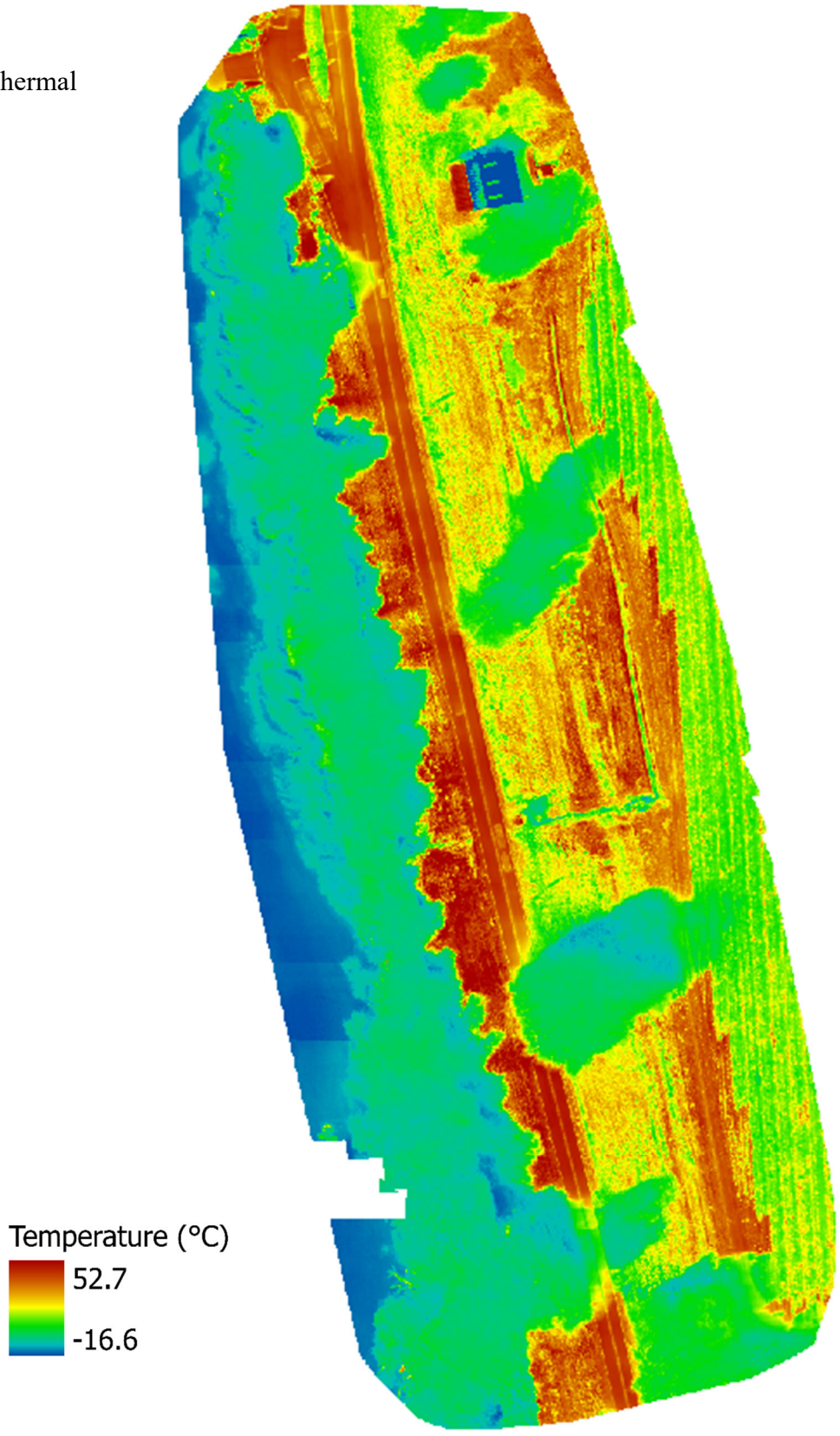
Grand Island
Site 3
Orthomosaic – Optical
Dry Season



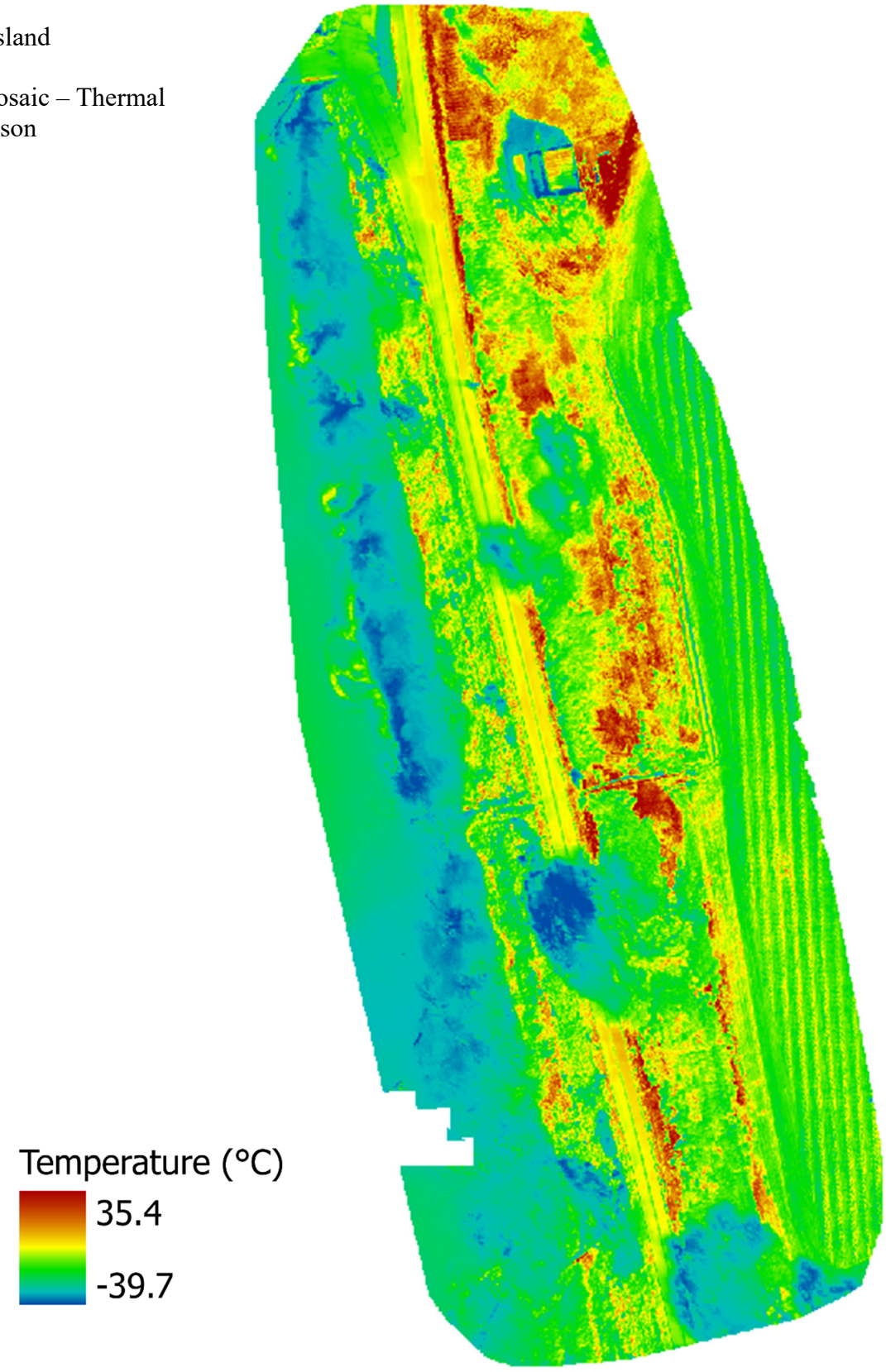
Grand Island
Site 3
Orthomosaic – Optical
Wet Season



Grand Island
Site 3
Orthomosaic – Thermal
Dry Season



Grand Island
Site 3
Orthomosaic – Thermal
Wet Season



Grand Island
Site 1
Digital Elevation Model
Dry Season



Grand Island
Site 1
Digital Elevation Model
Wet Season



Grand Island
Site 2
Digital Elevation Model
Dry Season



Grand Island
Site 2
Digital Elevation Model
Wet Season



Grand Island
Site 3
Digital Elevation Model
Dry Season

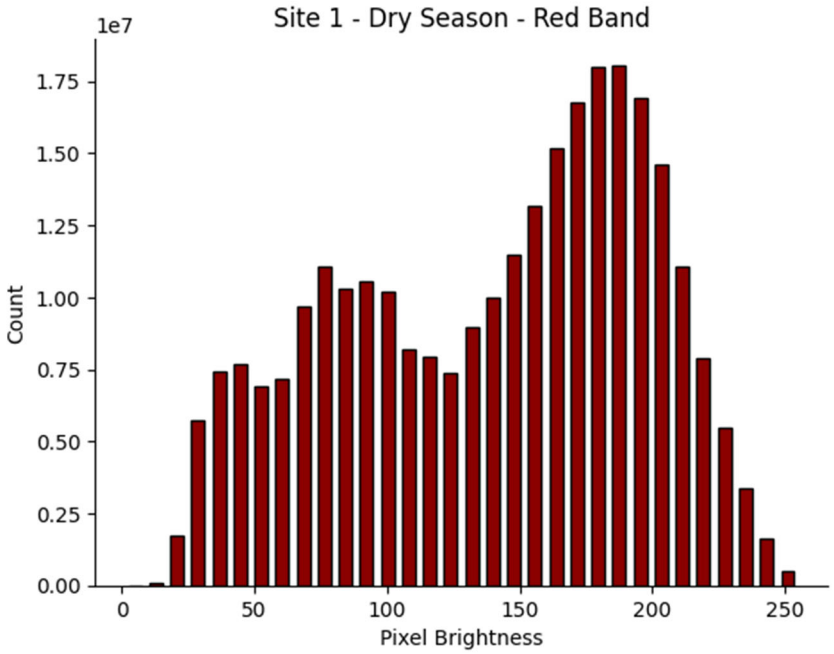


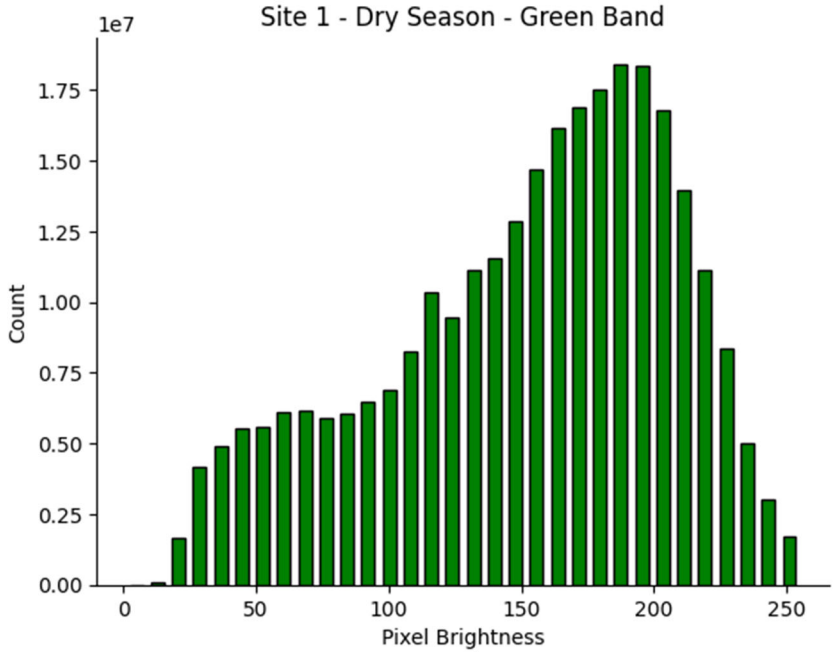
Grand Island
Site 3
Digital Elevation Model
Wet Season

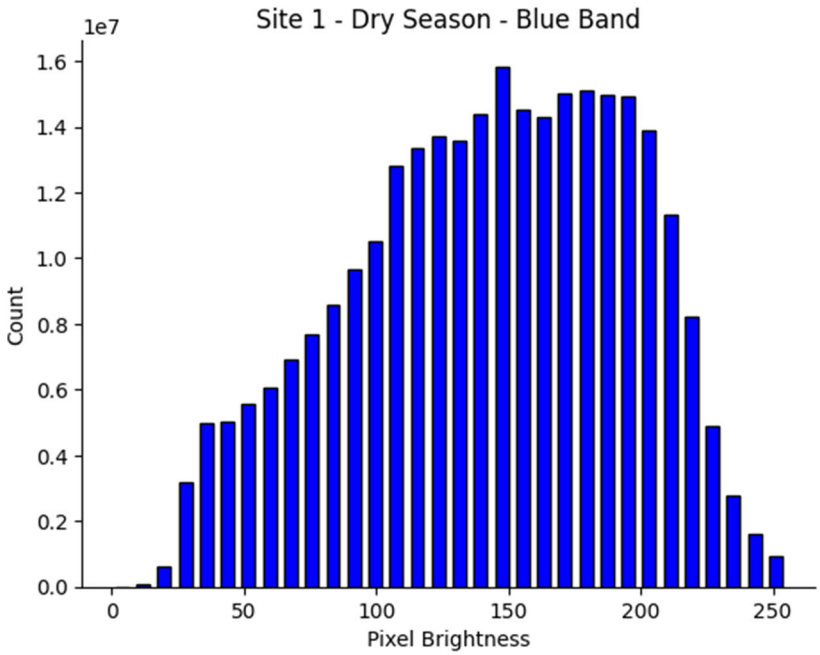


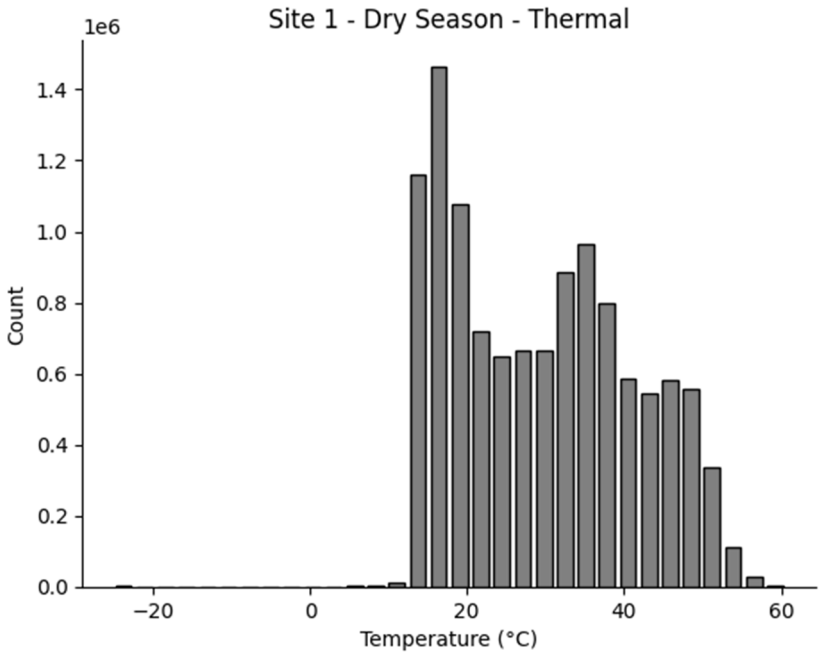
Histograms

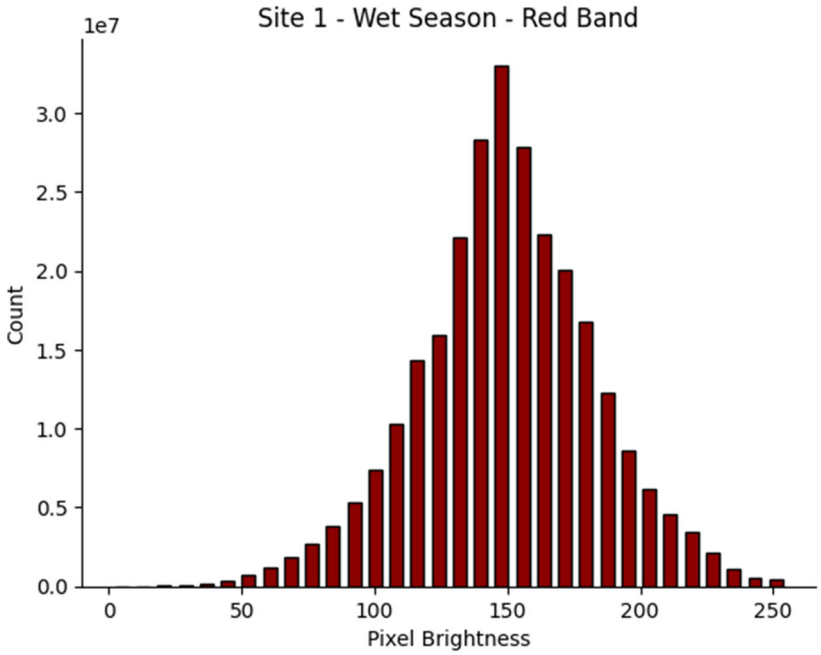
Grand Island
Site 1

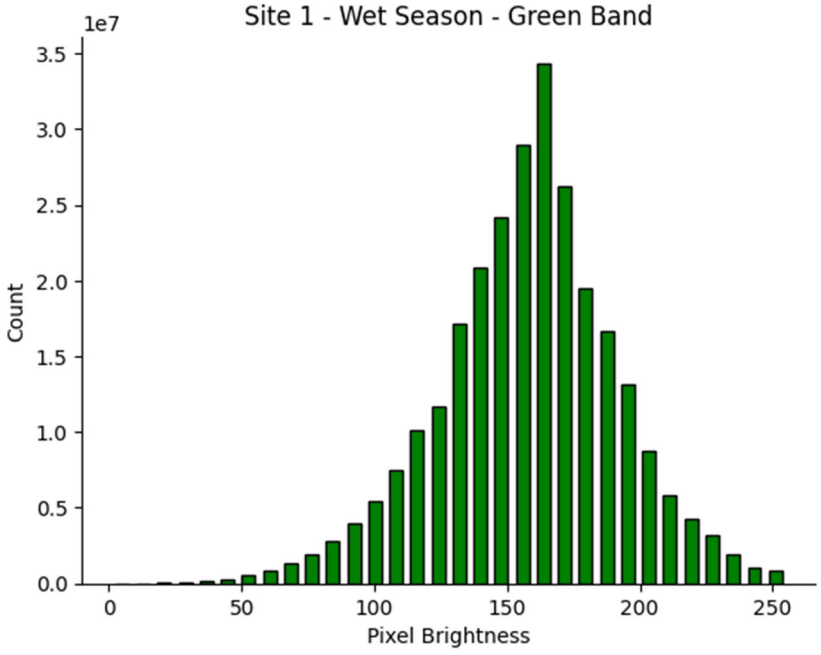


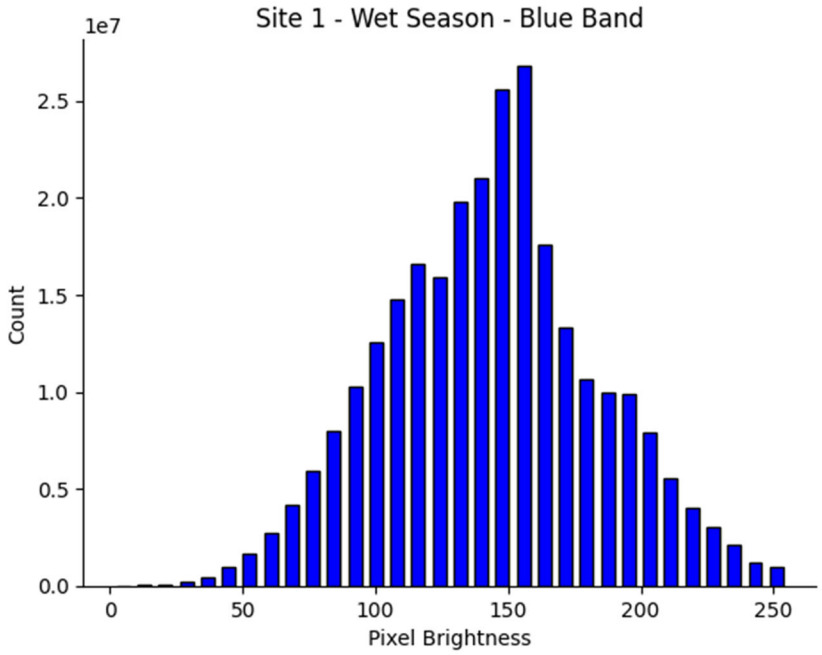


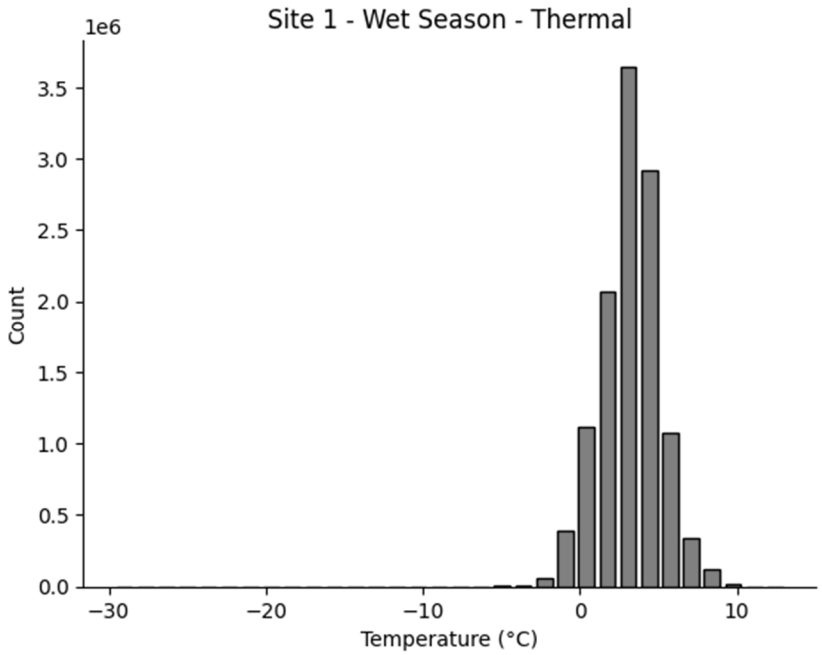

Site 1 - Dry Season - Blue Band
1e7
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
0
50
100
150
200
250
Pixel Brightness


Site 1 - Dry Season - Thermal
1e6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
−20
0
20
40
60
Temperature (°C)

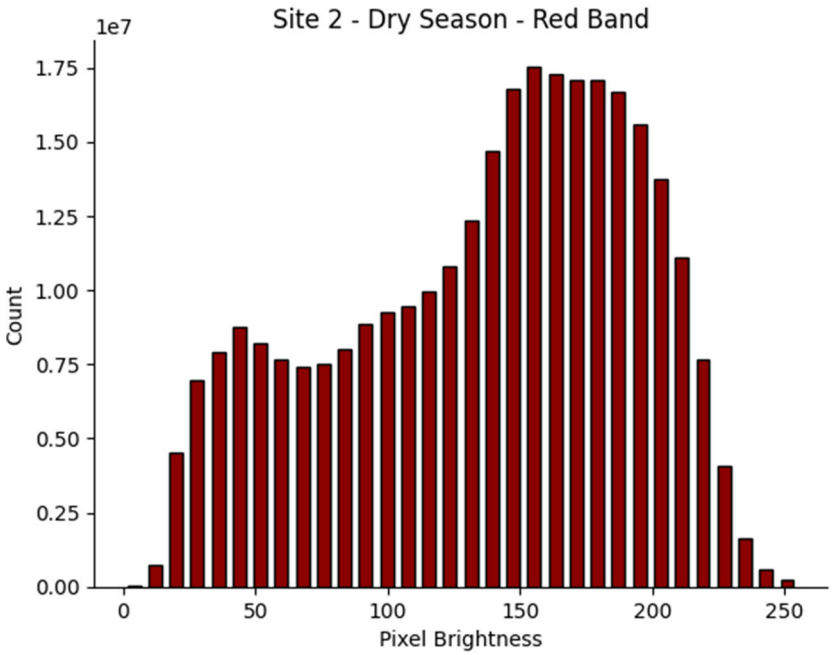

Site 1 - Wet Season - Red Band
1e7
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
0
50
100
150
200
250
Pixel Brightness

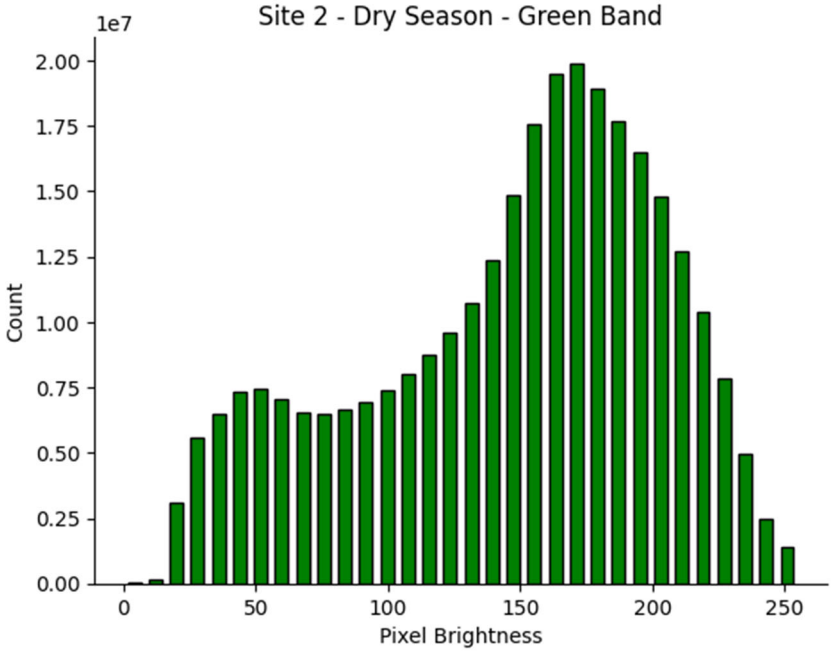

Site 1 - Wet Season - Green Band
1e7
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
0
50
100
150
200
250
Pixel Brightness

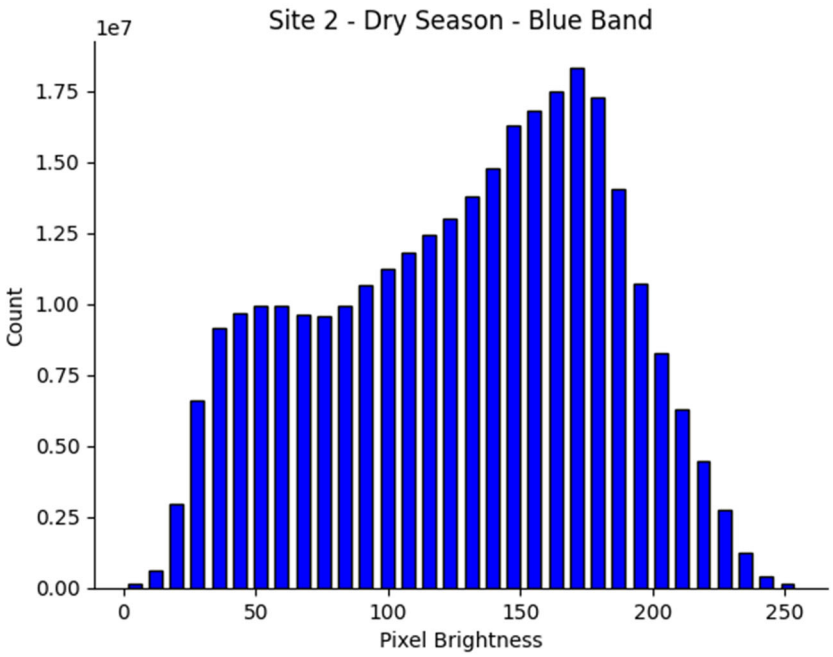

Site 1 - Wet Season - Blue Band
1e7
2.5
2.0
1.5
1.0
0.5
0.0
Count
0
50
100
150
200
250
Pixel Brightness

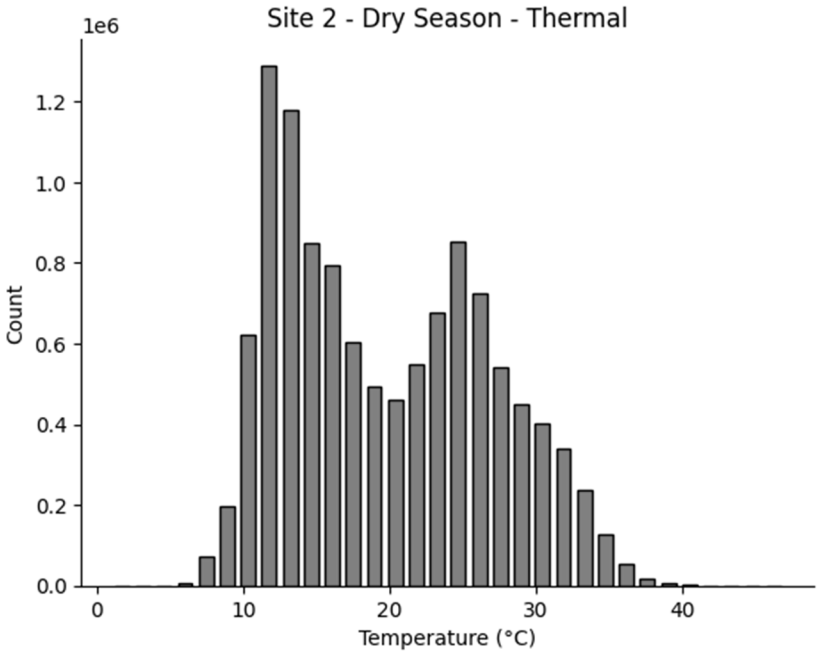

Site 1 - Wet Season - Thermal
1e6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
−30
−20
−10
0
10
Temperature (°C)

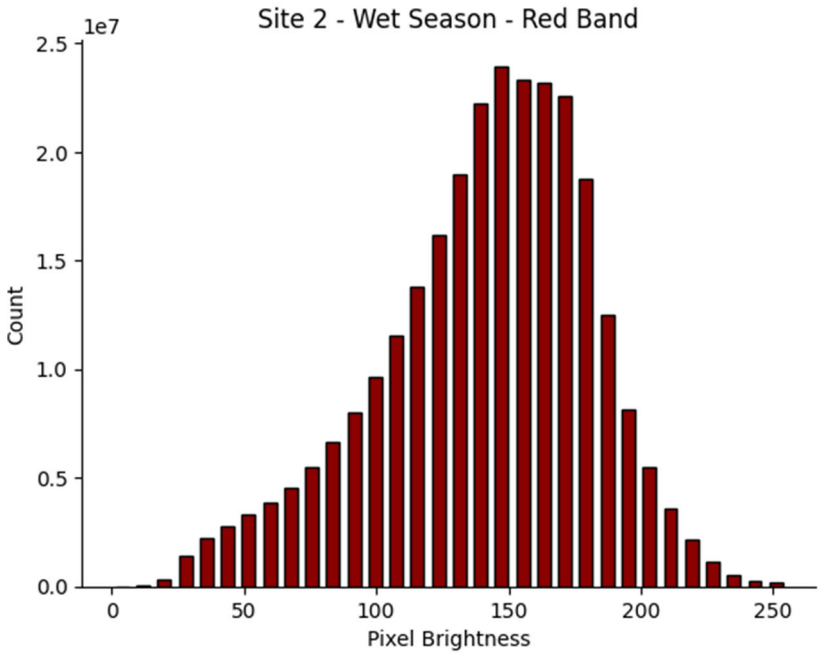
Grand Island
Site 2

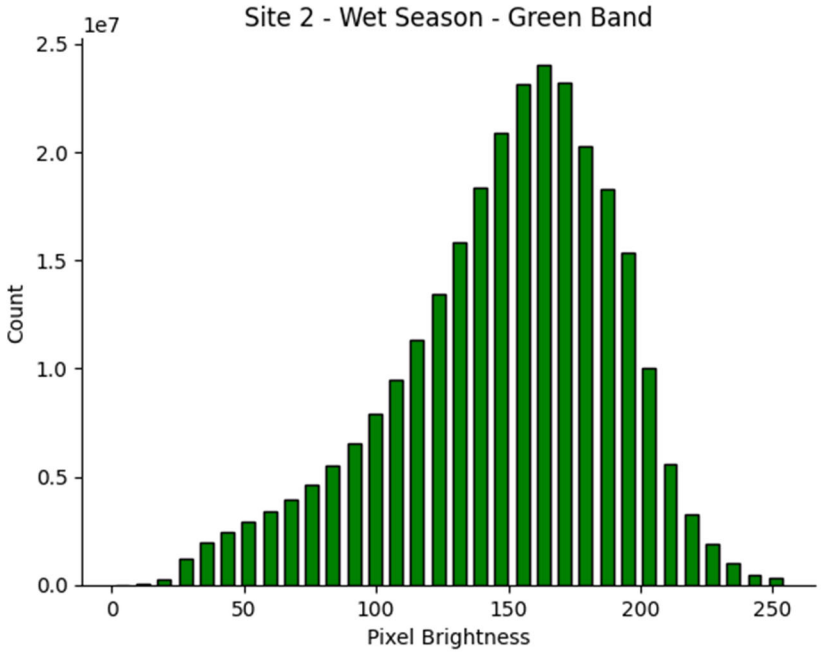


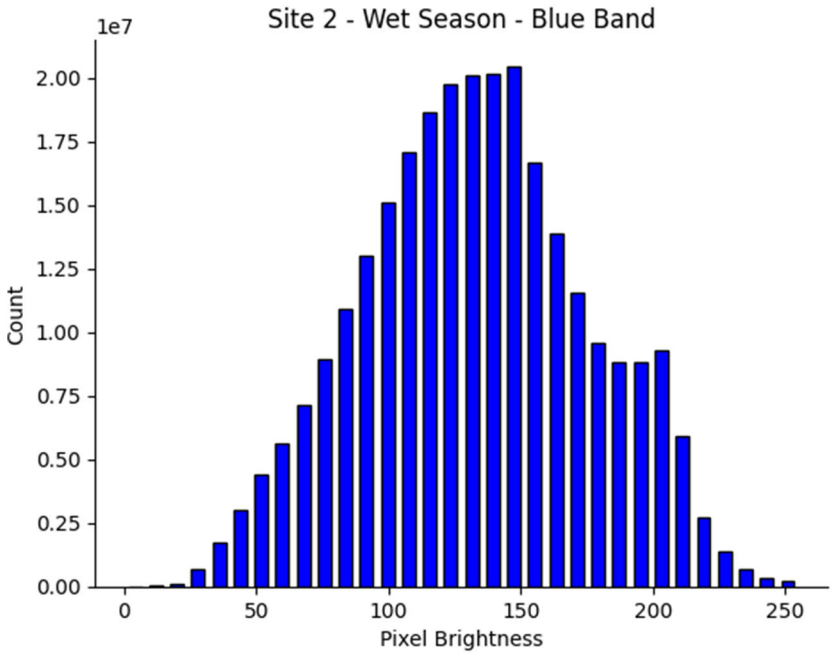


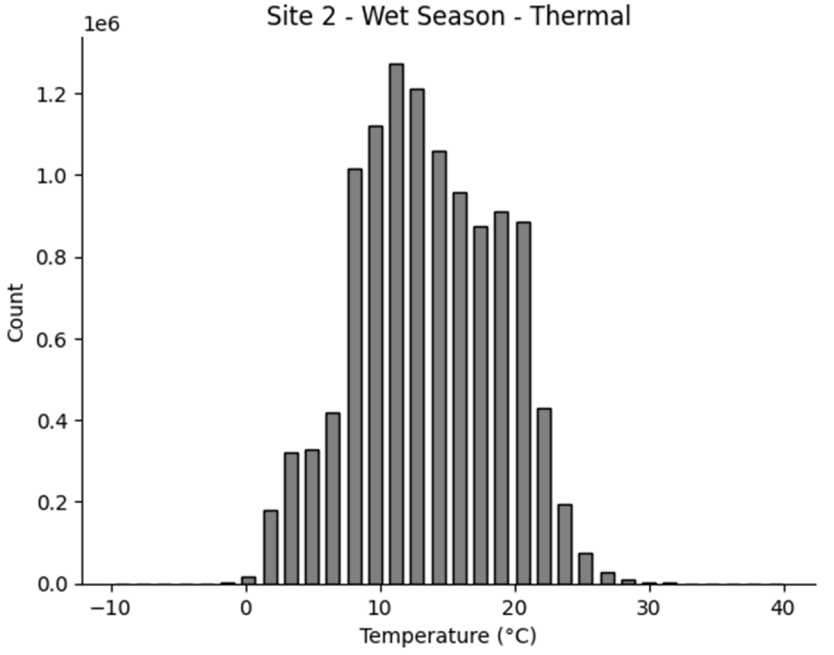

Site 2 - Dry Season - Blue Band
1e7
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Count
0
50
100
150
200
250
Pixel Brightness


Site 2 - Dry Season - Thermal
1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
0
10
20
30
40
Temperature (°C)

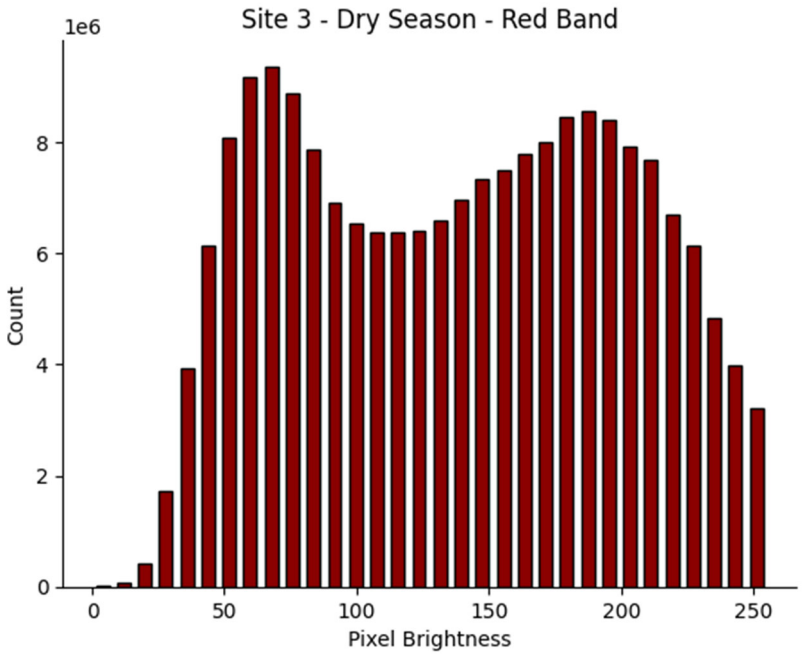

Site 2 - Wet Season - Red Band
1e7
2.5
2.0
1.5
1.0
0.5
0.0
Count
0
50
100
150
200
250
Pixel Brightness

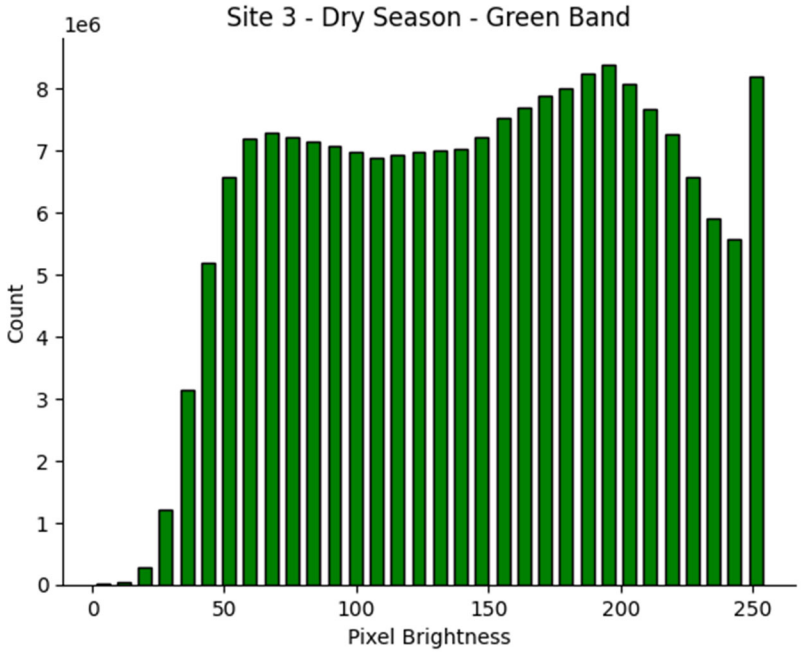

Site 2 - Wet Season - Green Band
1e7
2.5
2.0
1.5
1.0
0.5
0.0
Count
0
50
100
150
200
250
Pixel Brightness

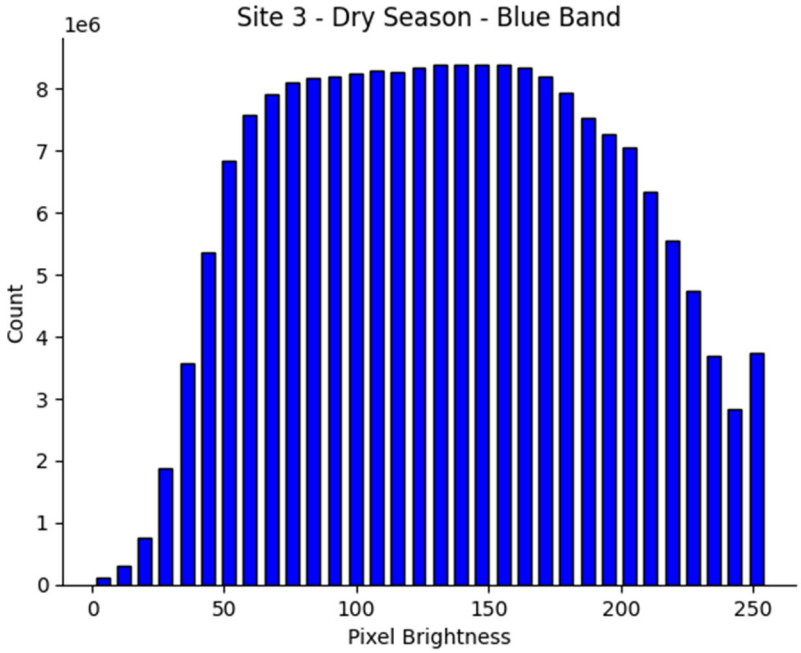

Site 2 - Wet Season - Blue Band
1e7
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Count
0
50
100
150
200
250
Pixel Brightness

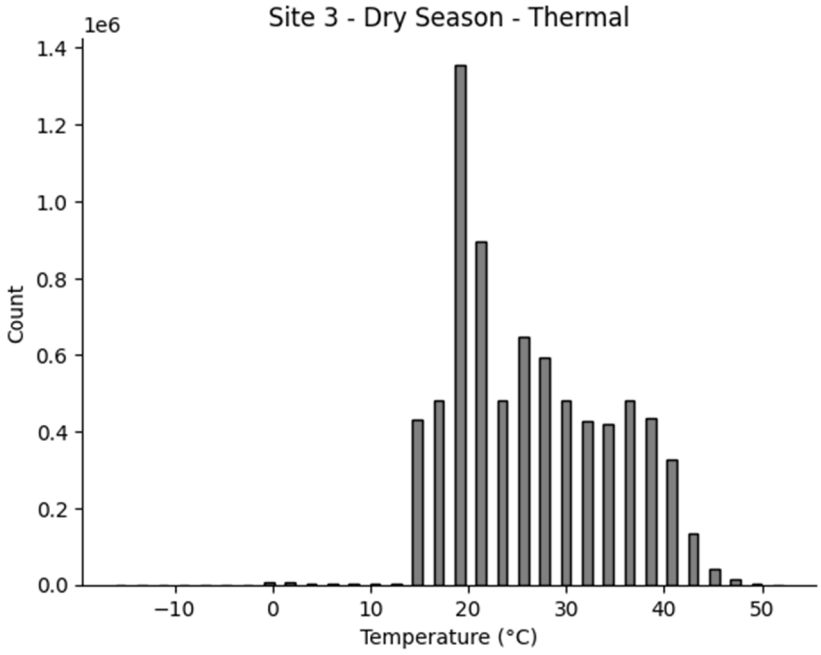

Site 2 - Wet Season - Thermal
1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
−10
0
10
20
30
40
Temperature (°C)

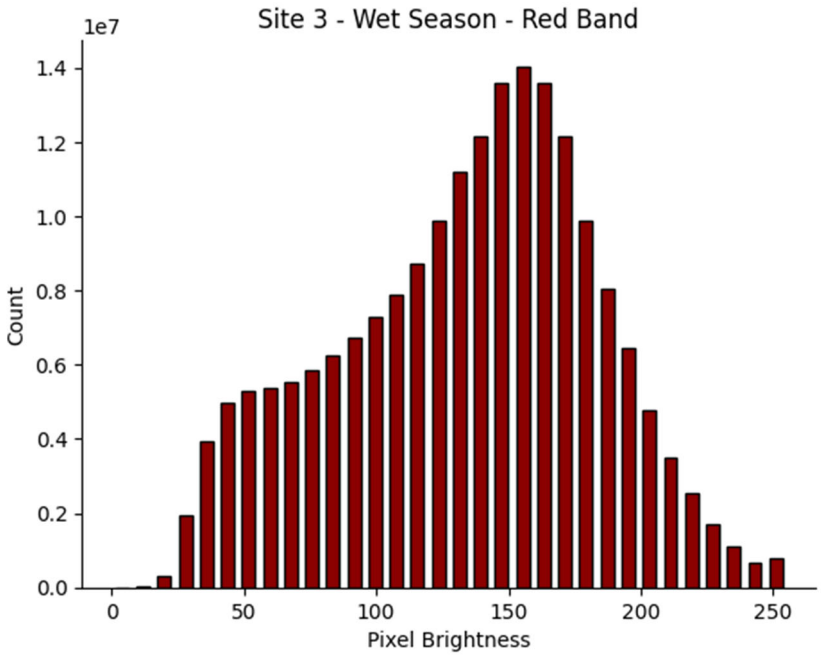
Grand Island
Site 3

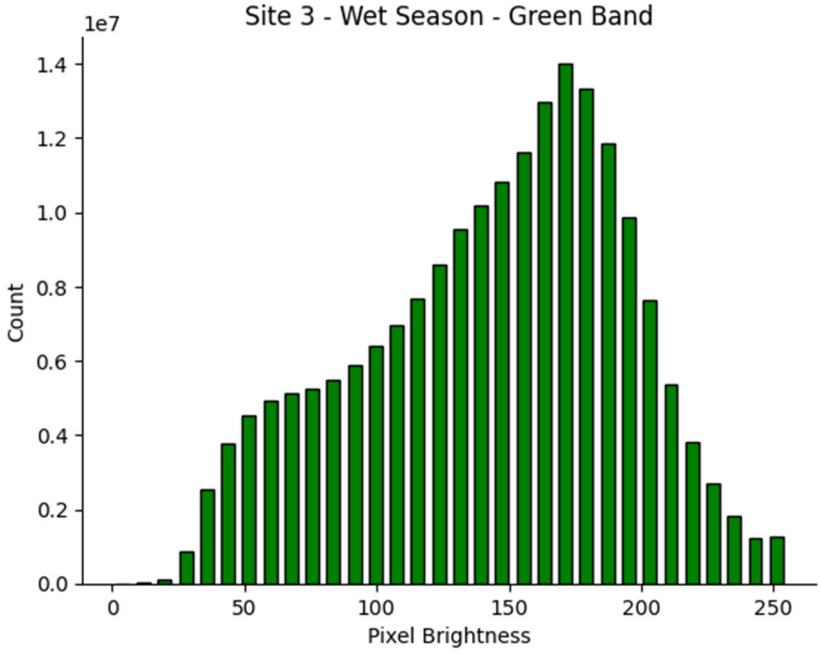


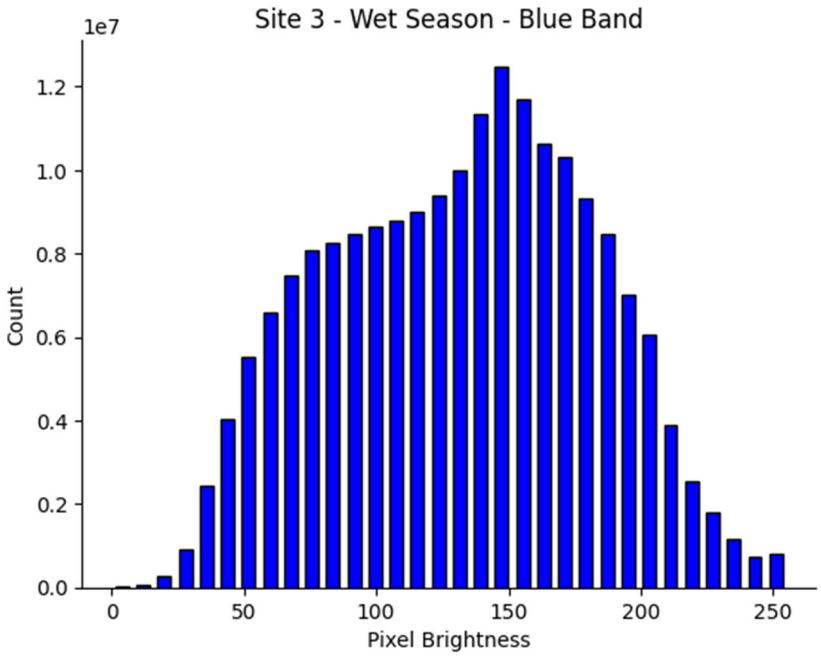


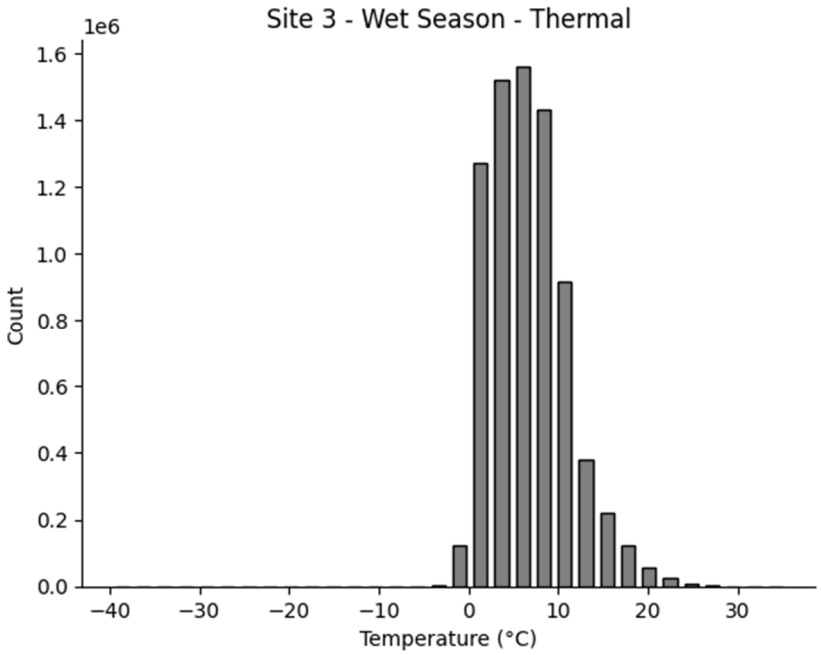

Site 3 - Dry Season - Blue Band
1e6
Count
Pixel Brightness
0
1
2
3
4
5
6
7
8
0
50
100
150
200
250


Site 3 - Dry Season - Thermal
1e6
Count
Temperature (°C)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
−10
0
10
20
30
40
50


Site 3 - Wet Season - Red Band
1e7
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
0
50
100
150
200
250
Pixel Brightness


Site 3 - Wet Season - Green Band
1e7
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
0
50
100
150
200
250
Pixel Brightness

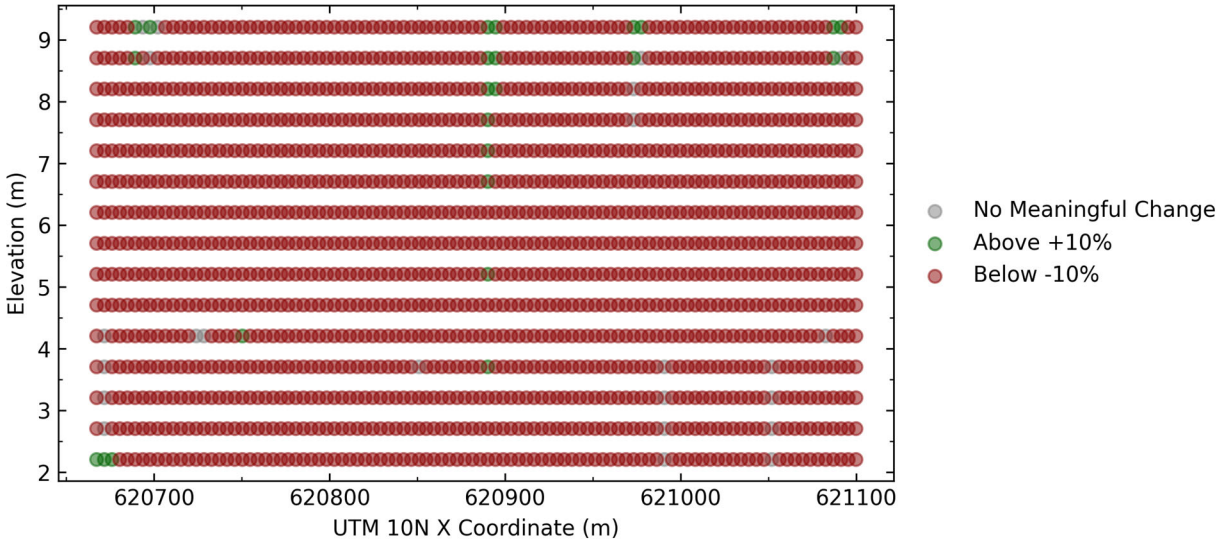

Site 3 - Wet Season - Blue Band
1e7
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
0
50
100
150
200
250
Pixel Brightness


Site 3 - Wet Season - Thermal
1e6
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
−40
−30
−20
−10
0
10
20
30
Temperature (°C)

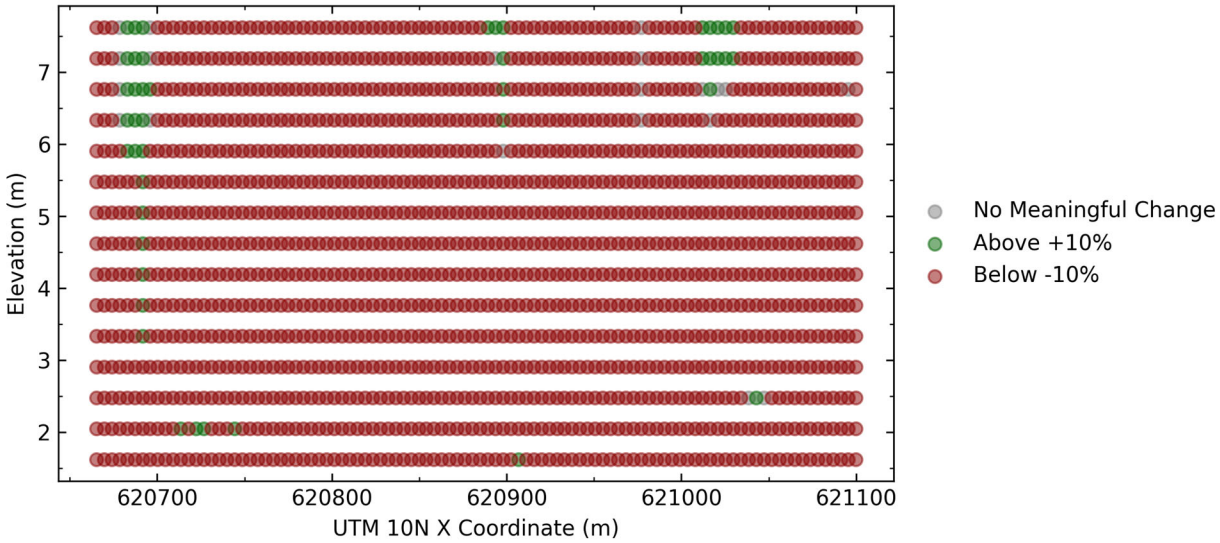
Appendix E – Change Detection Profiles

Subsurface EMI Change Detection Profiles

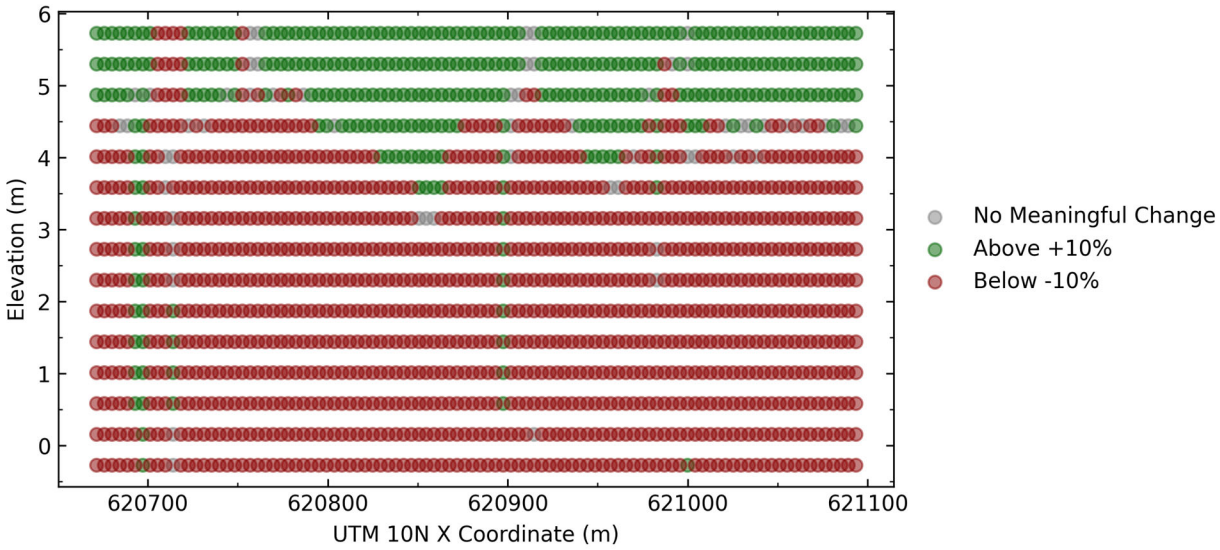
Grand Island
Site 1
EMI - Dry Crest



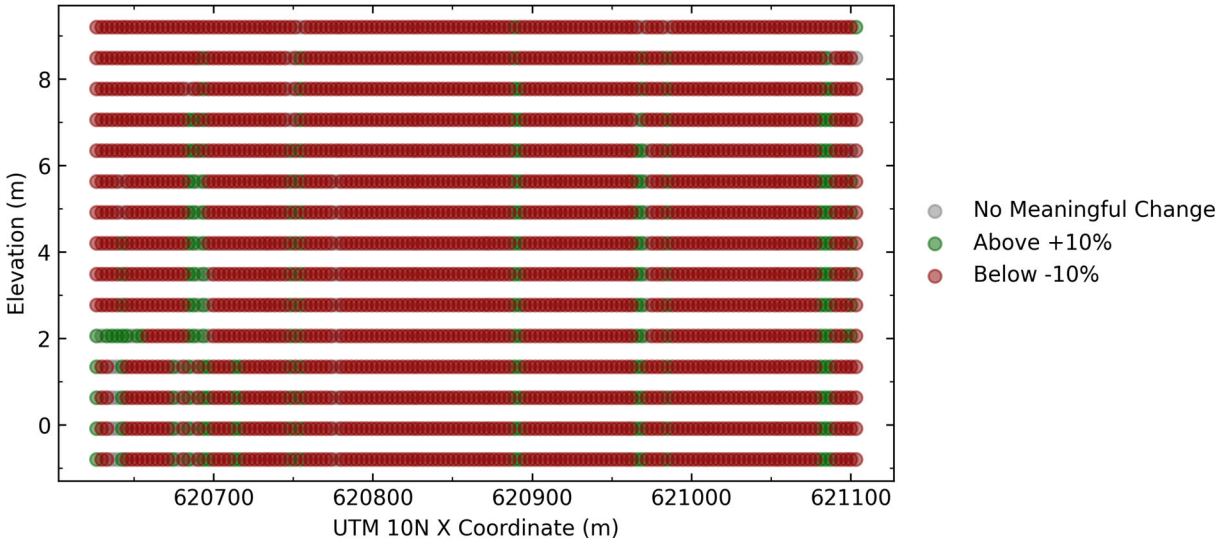
Grand Island
Site 1
EMI - Dry Mid-Slope



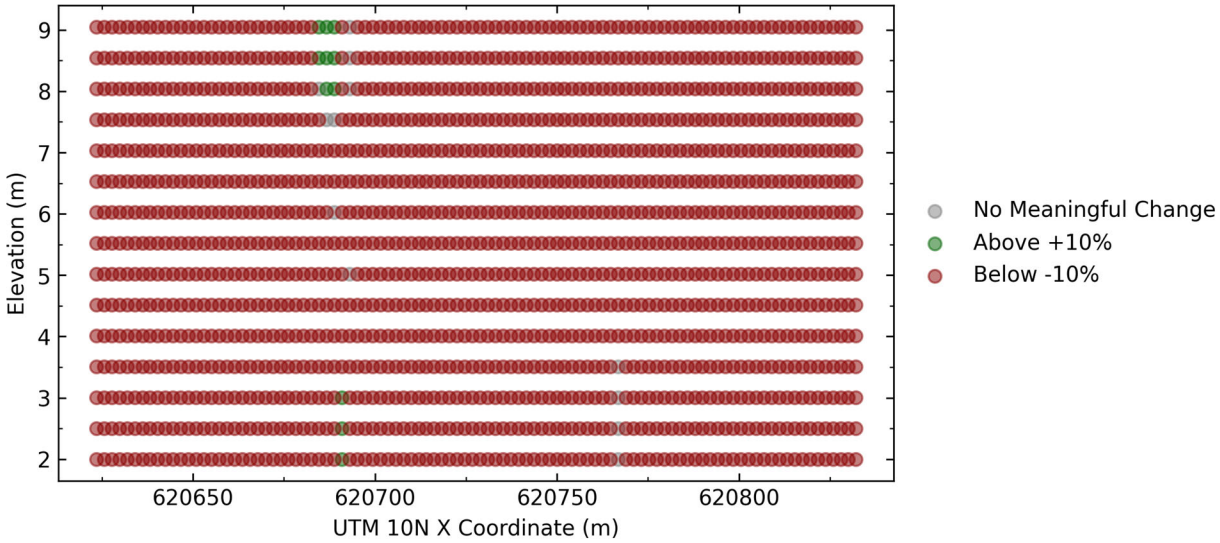
Grand Island
Site 1
EMI - Dry Toe



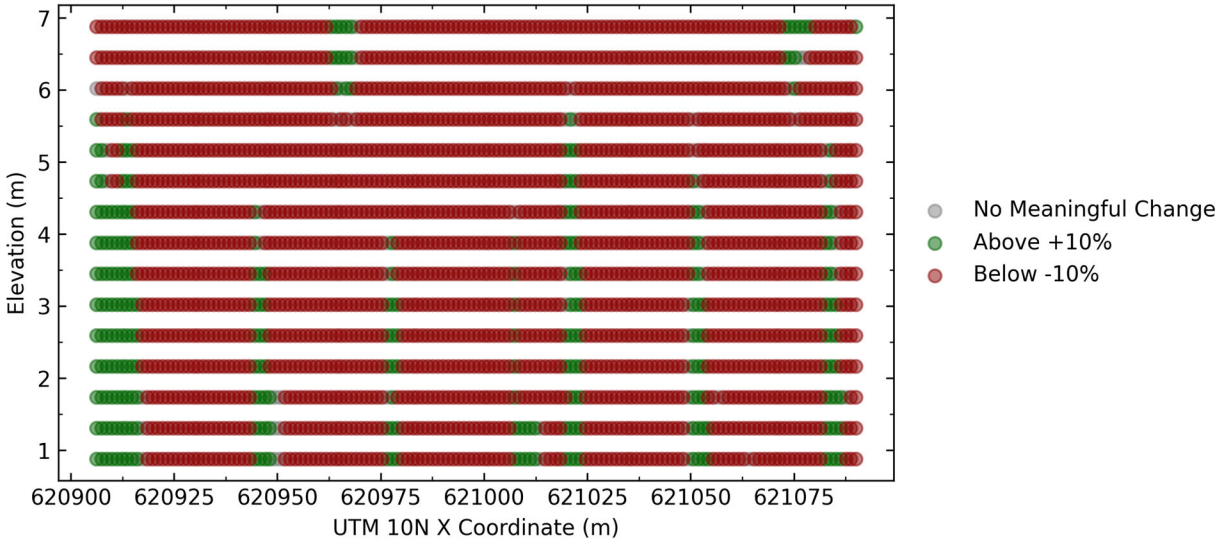
Grand Island
Site 1
EMI – Wet Crest



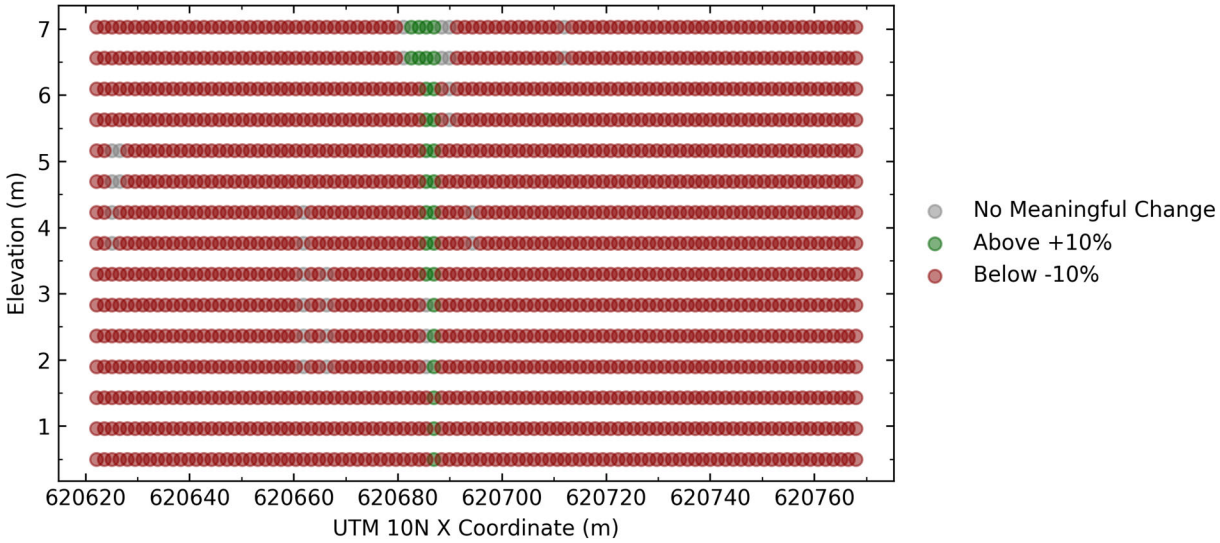
Grand Island
Site 1
EMI – Wet Mid-Slope



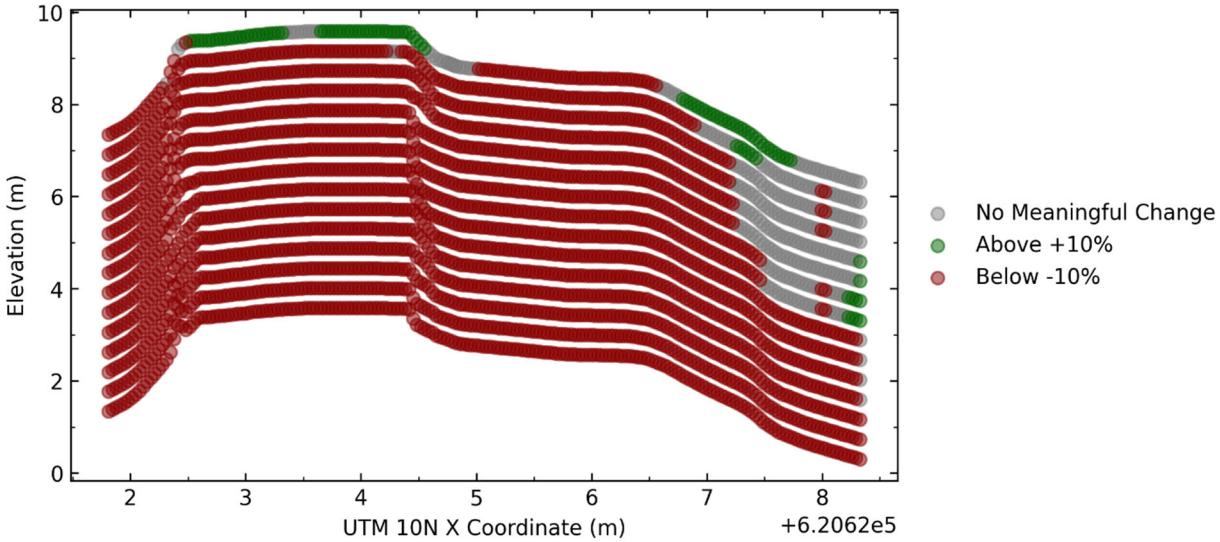
Grand Island
Site 1
EMI – Wet Toe East



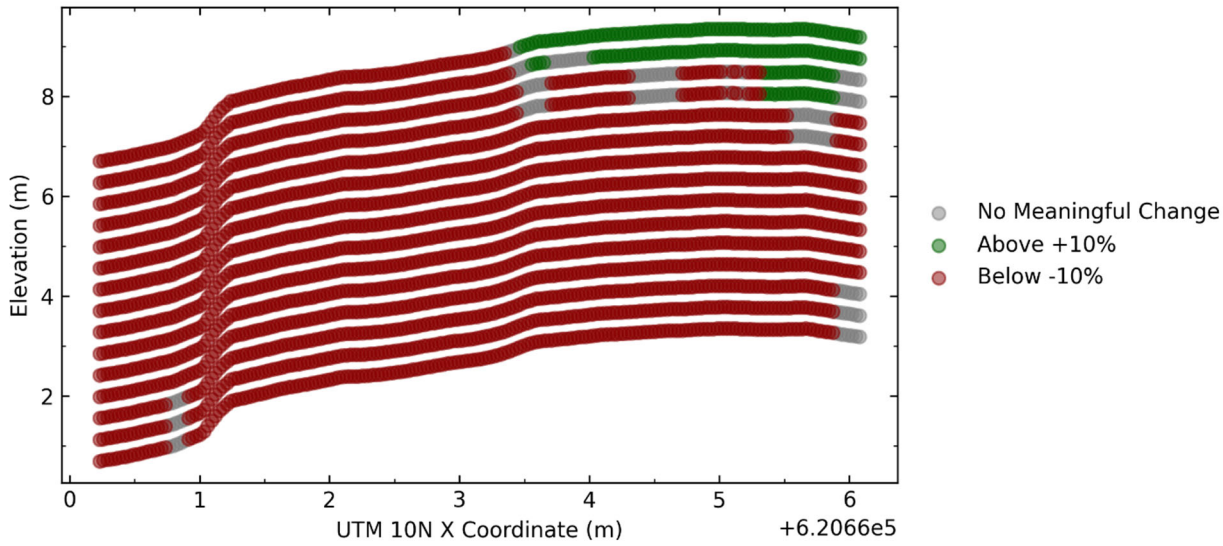
Grand Island
Site 1
EMI – Wet Toe West



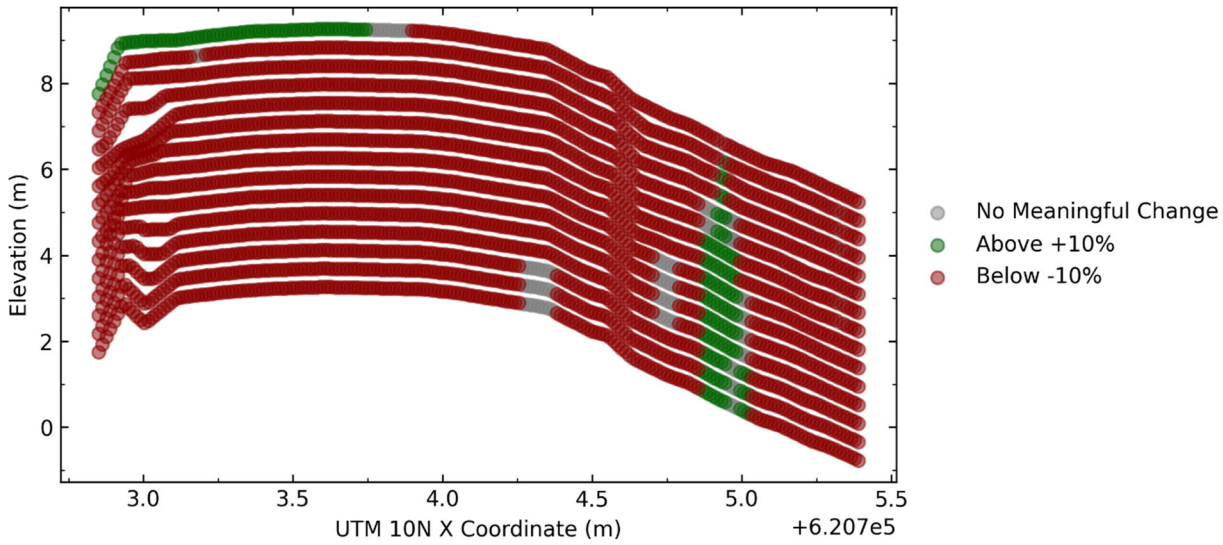
Grand Island
Site 1
EMI – Cross Section 1



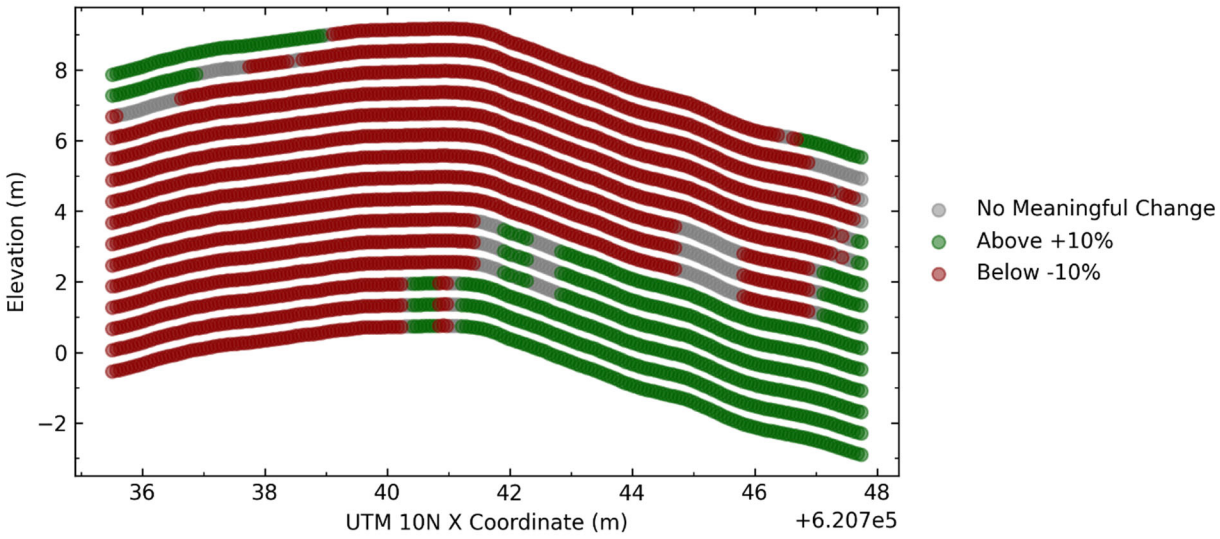
Grand Island
Site 1
EMI – Cross Section 2



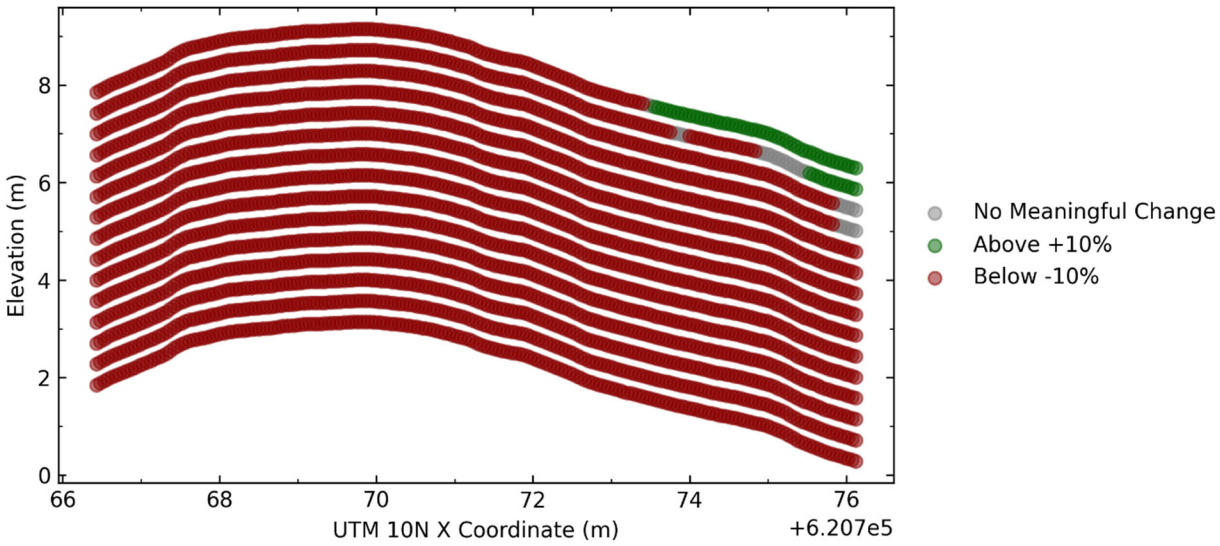
Grand Island
Site 1
EMI – Cross Section 3



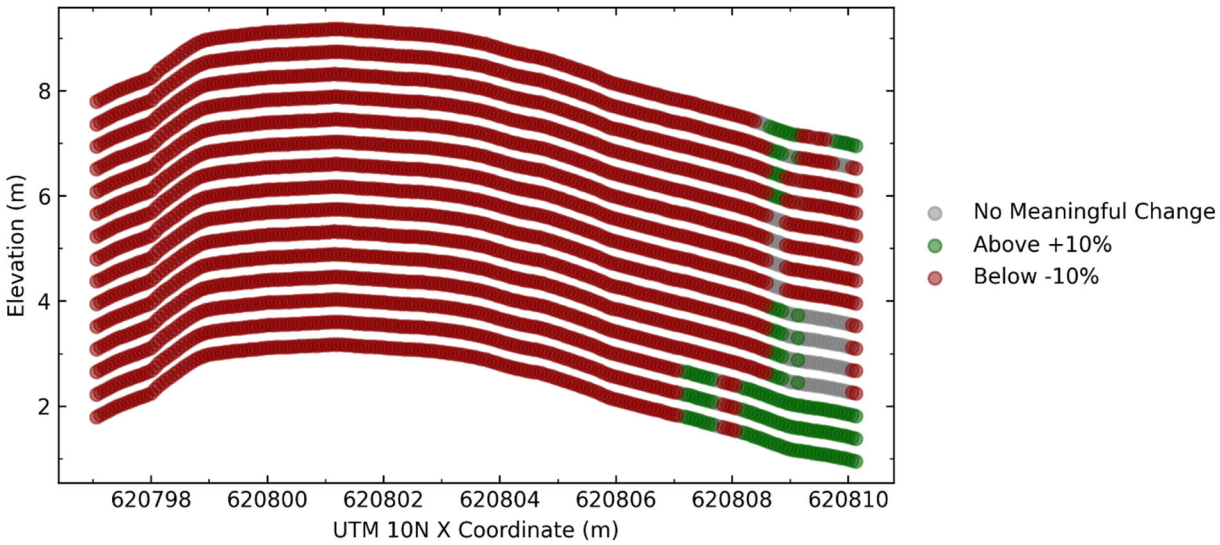
Grand Island
Site 1
EMI – Cross Section 4



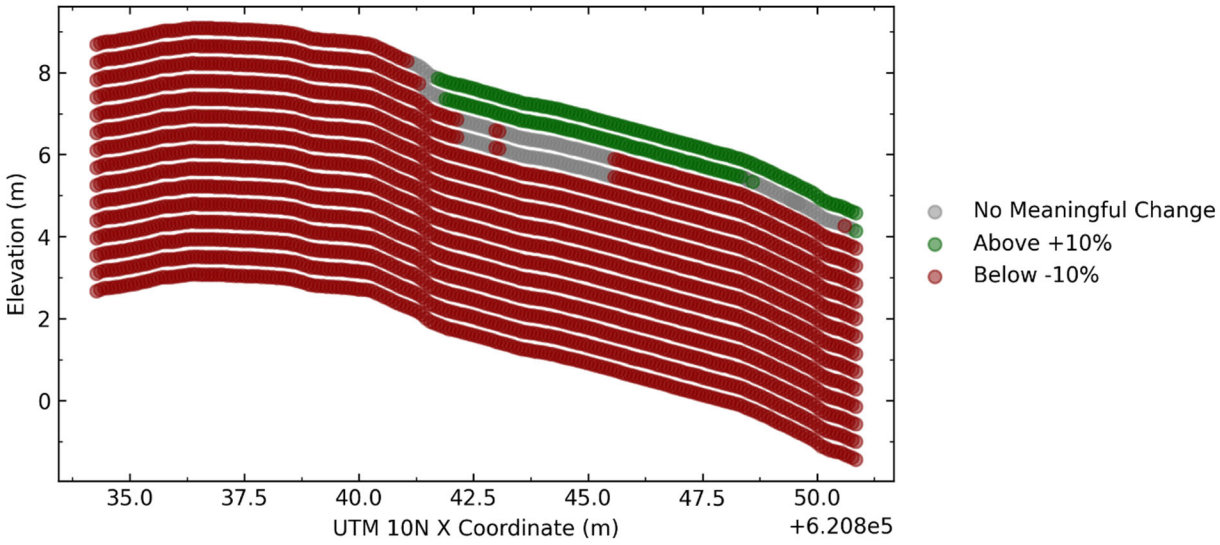
Grand Island
Site 1
EMI – Cross Section 5



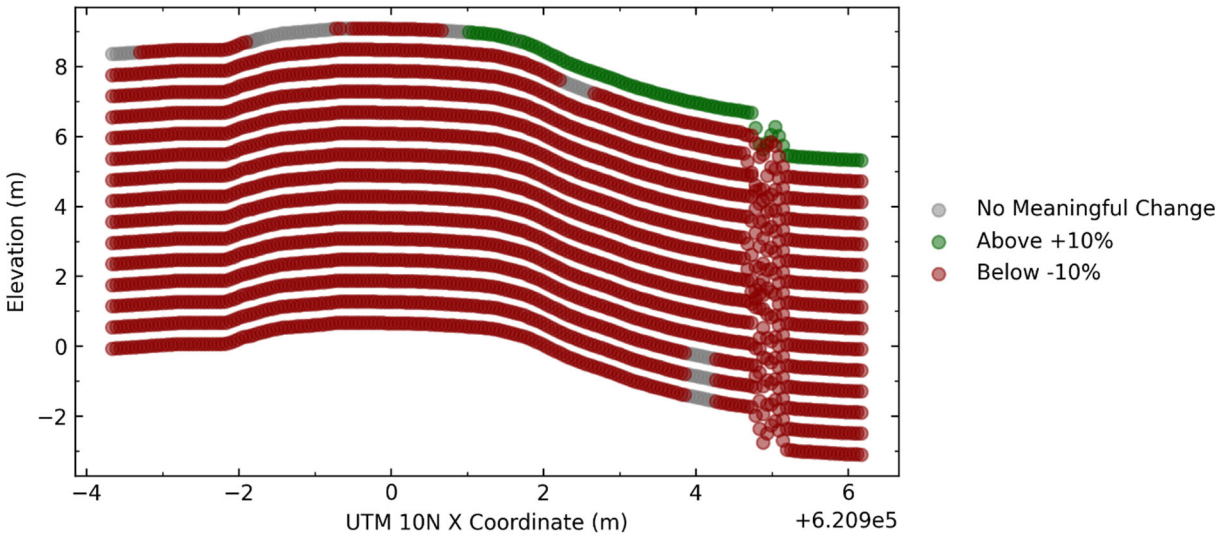
Grand Island
Site 1
EMI – Cross Section 6



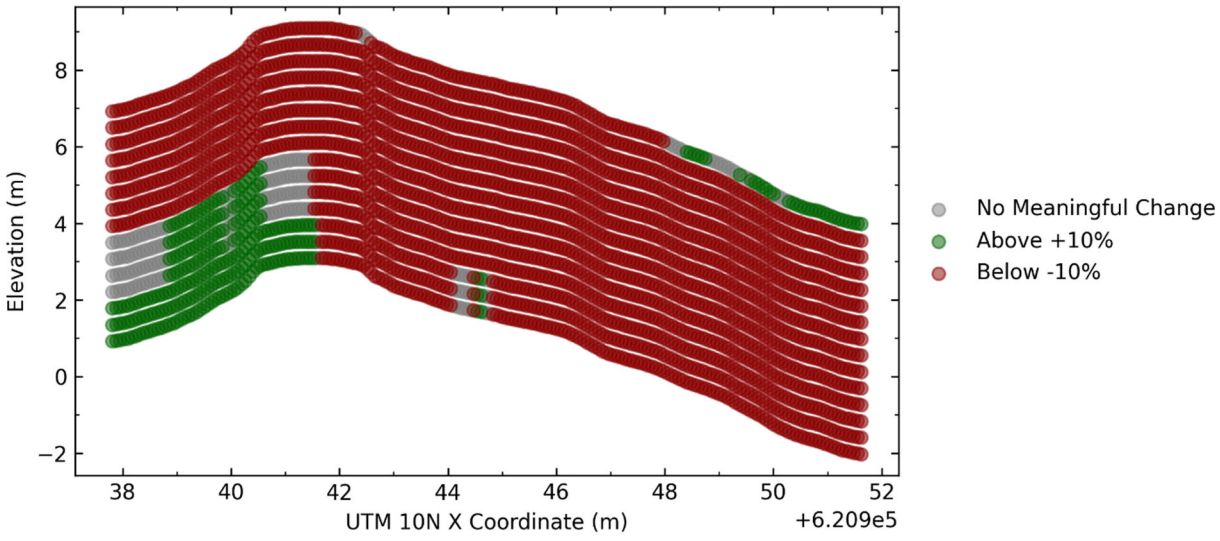
Grand Island
Site 1
EMI – Cross Section 7



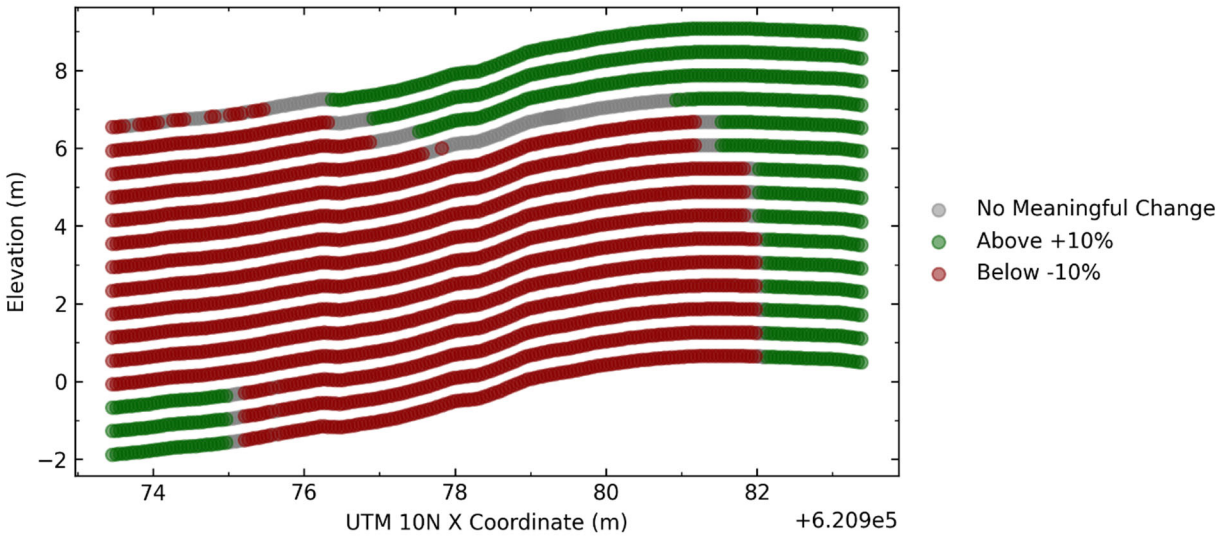
Grand Island
Site 1
EMI – Cross Section 8



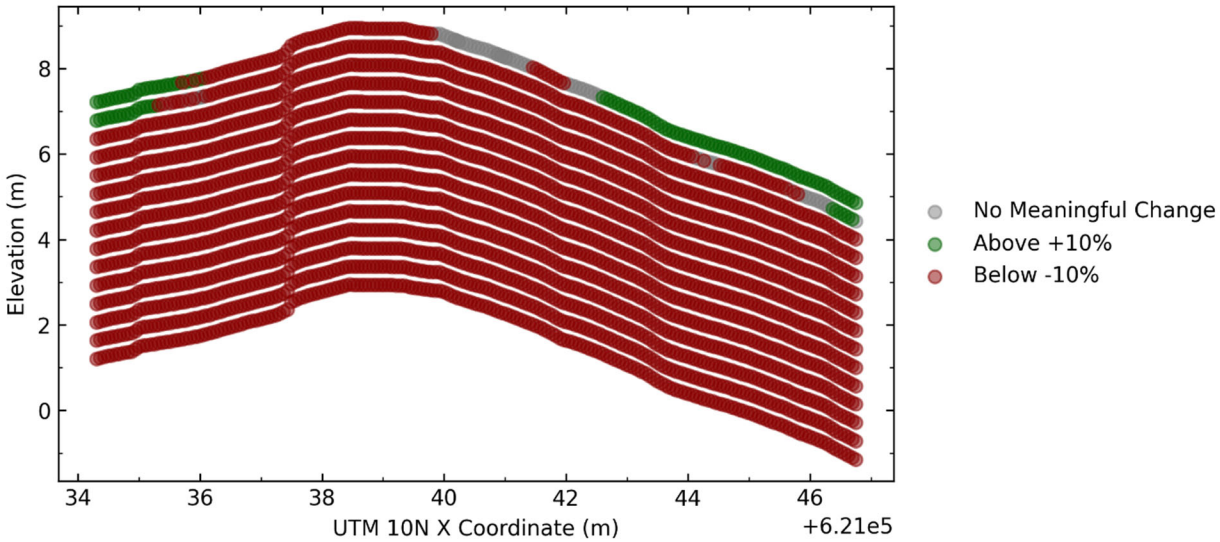
Grand Island
Site 1
EMI – Cross Section 9



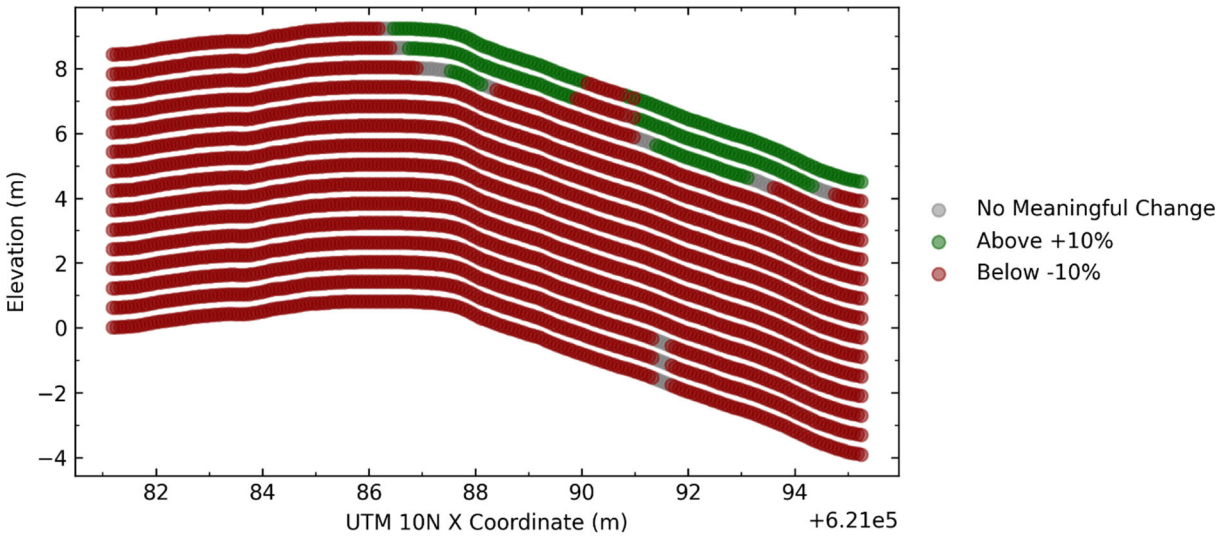
Grand Island
Site 1
EMI – Cross Section 10



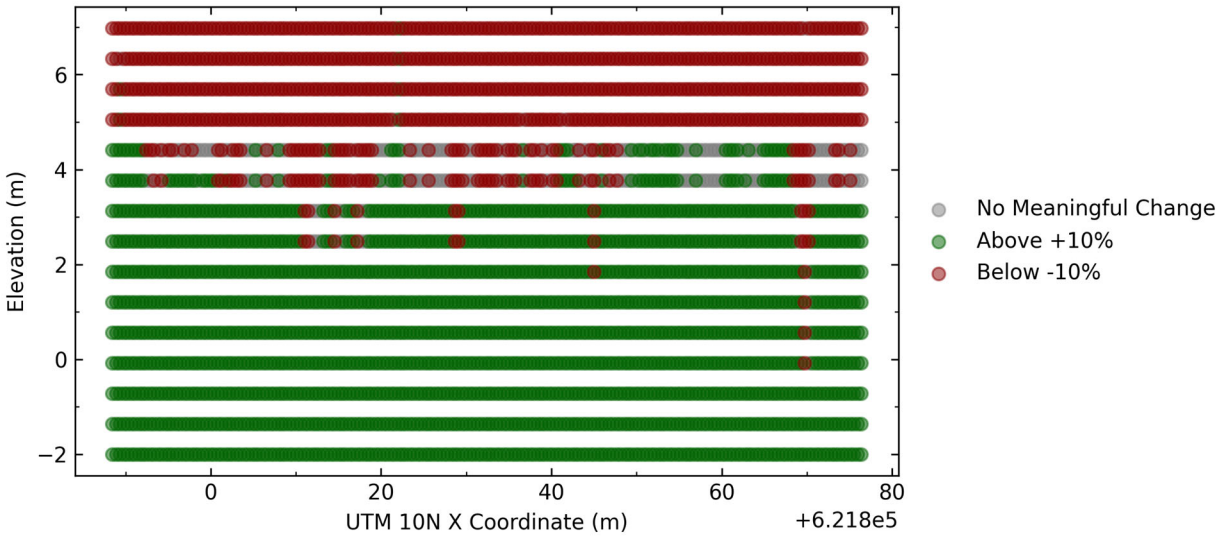
Grand Island
Site 1
EMI – Cross Section 11



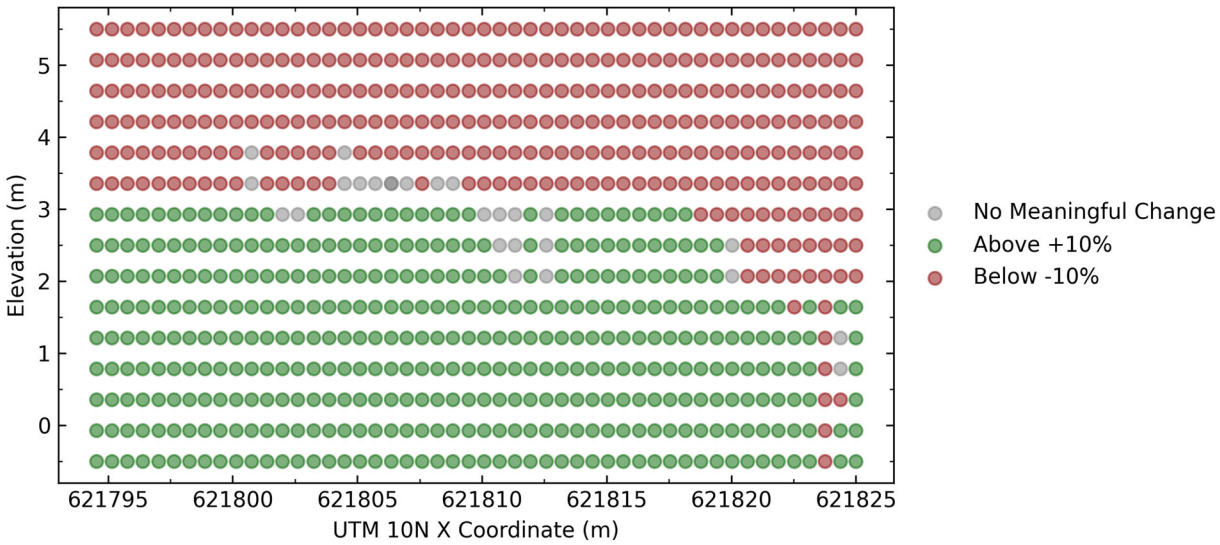
Grand Island
Site 1
EMI – Cross Section 12



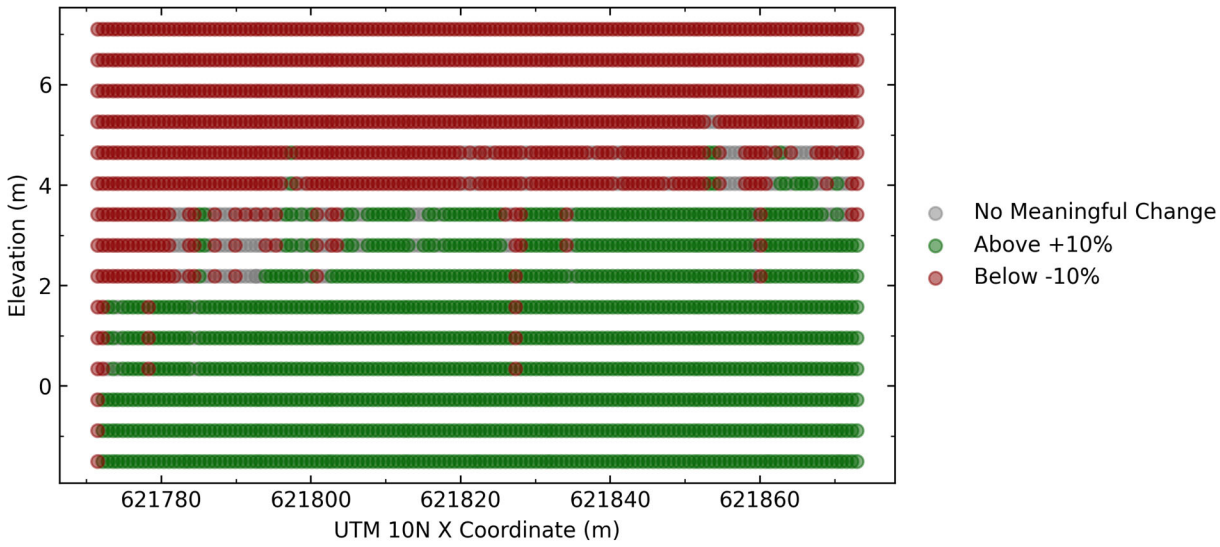
Grand Island
Site 2
EMI – Dry Crest



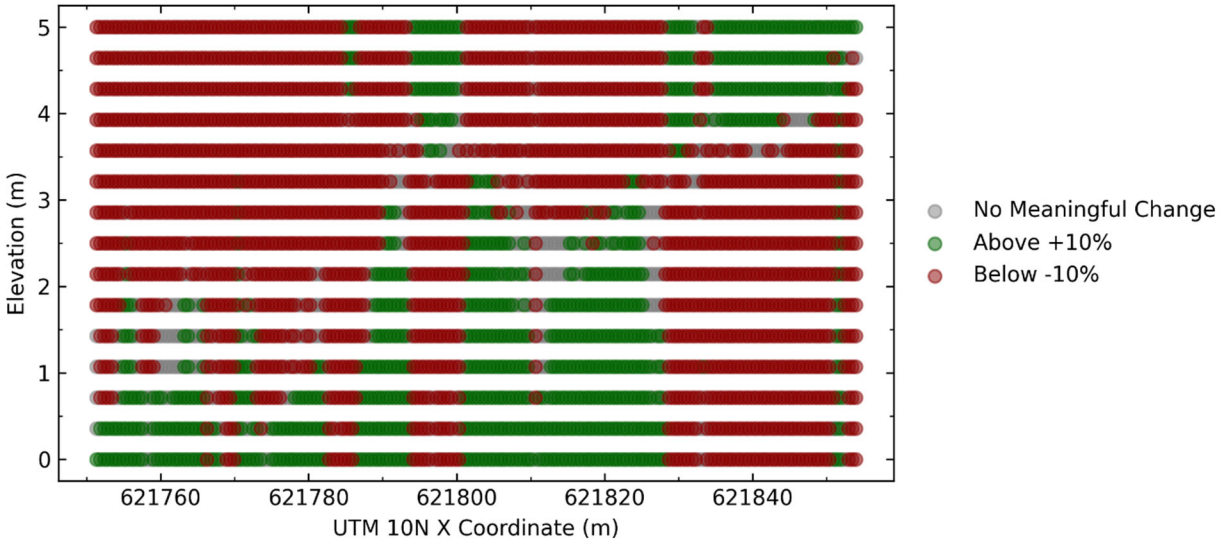
Grand Island
Site 2
EMI – Dry Mid-Slope



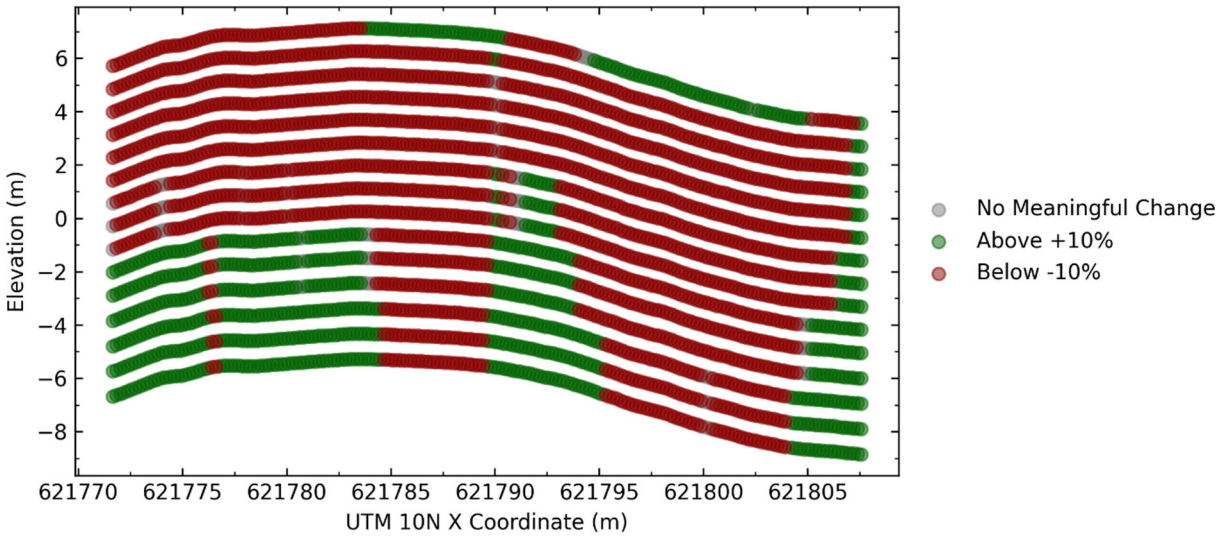
Grand Island
Site 2
EMI – Wet Crest



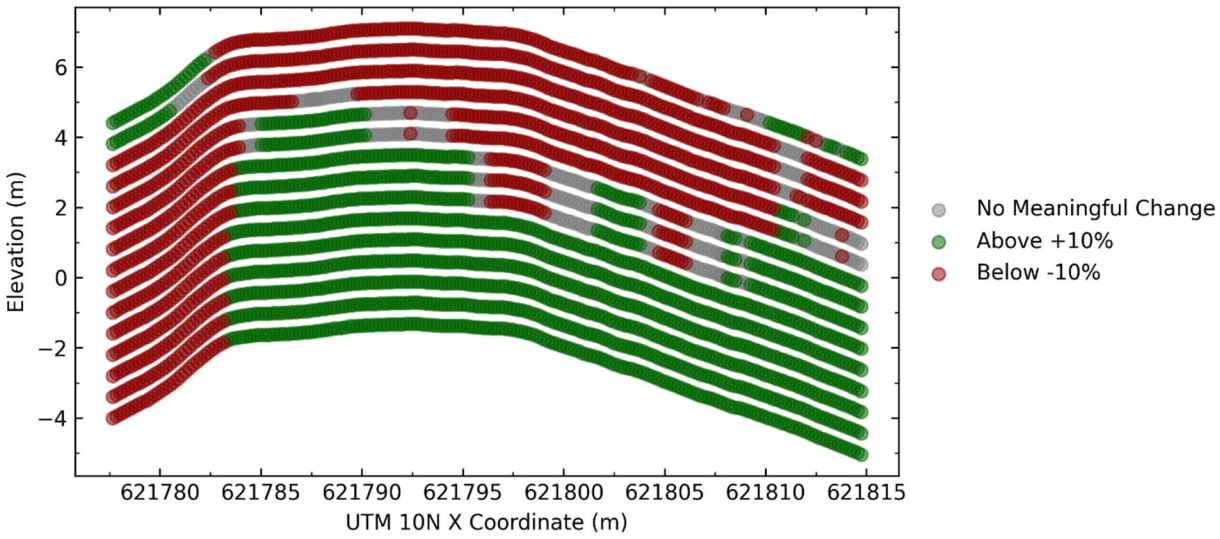
Grand Island
Site 2
EMI – Wet Toe



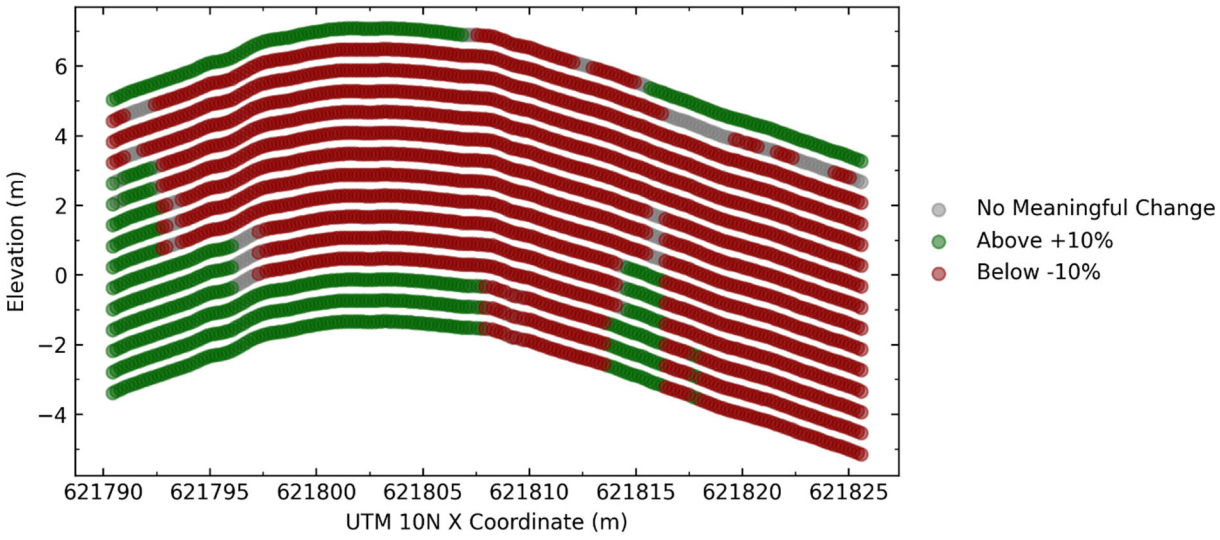
Grand Island
Site 2
EMI – Cross Section 1



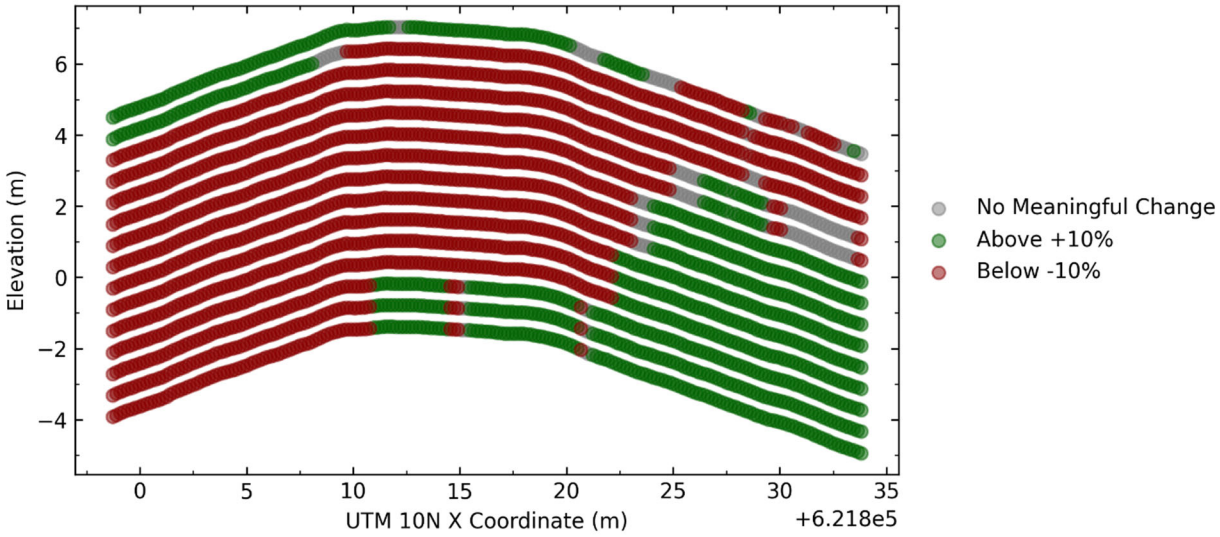
Grand Island
Site 2
EMI – Cross Section 2



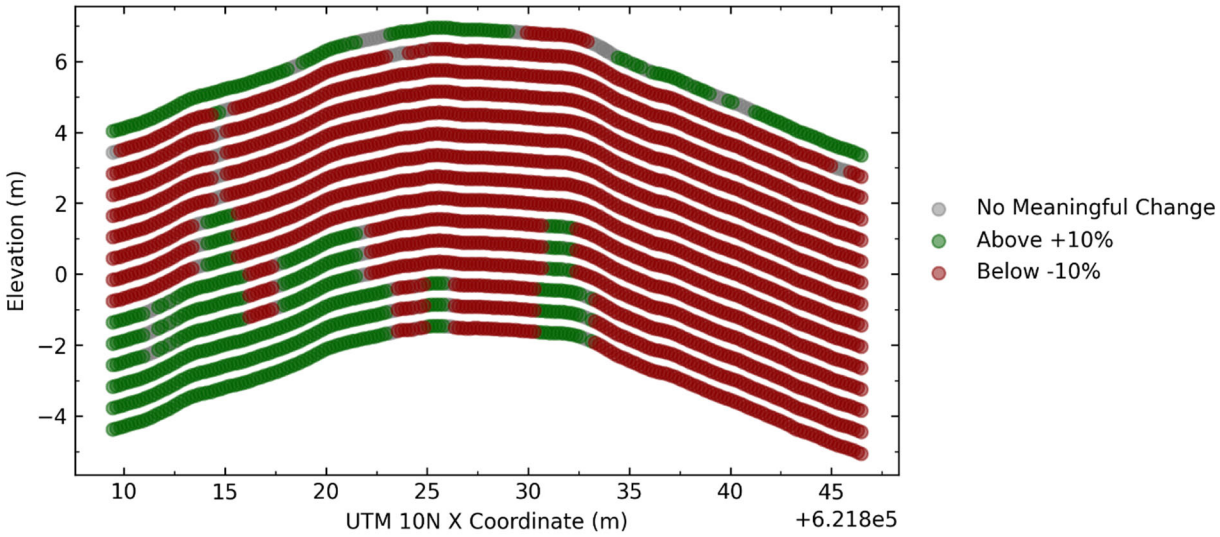
Grand Island
Site 2
EMI – Cross Section 3



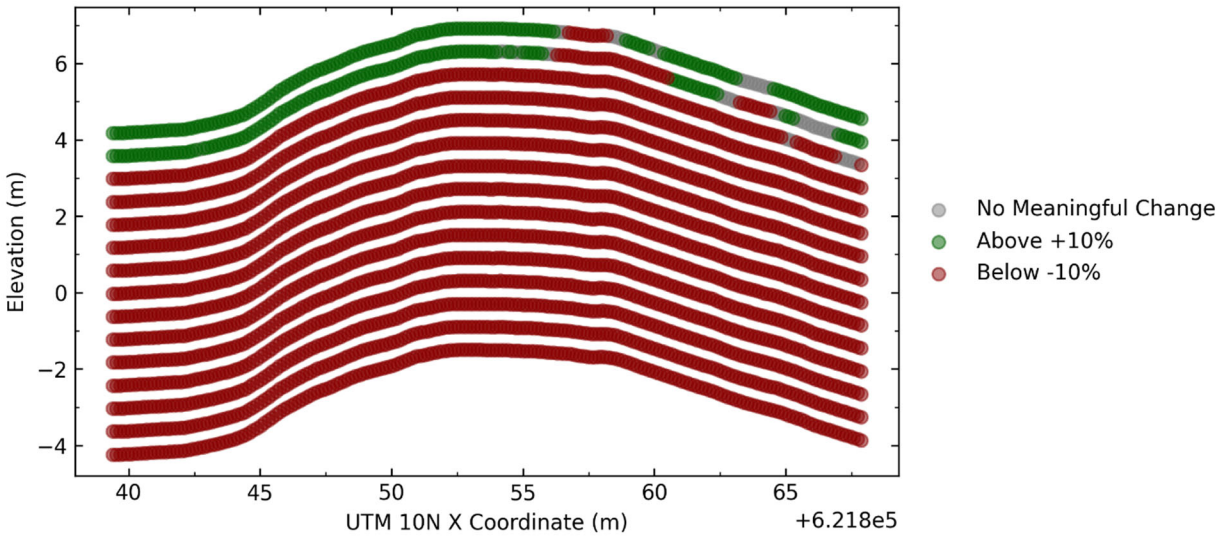
Grand Island
Site 2
EMI – Cross Section 4



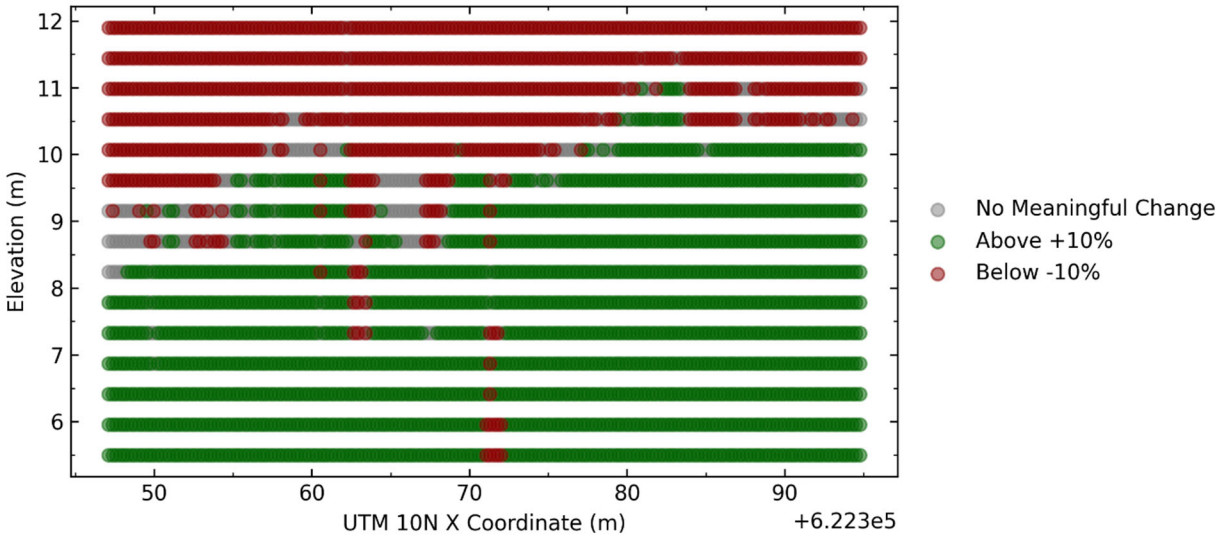
Grand Island
Site 2
EMI – Cross Section 5



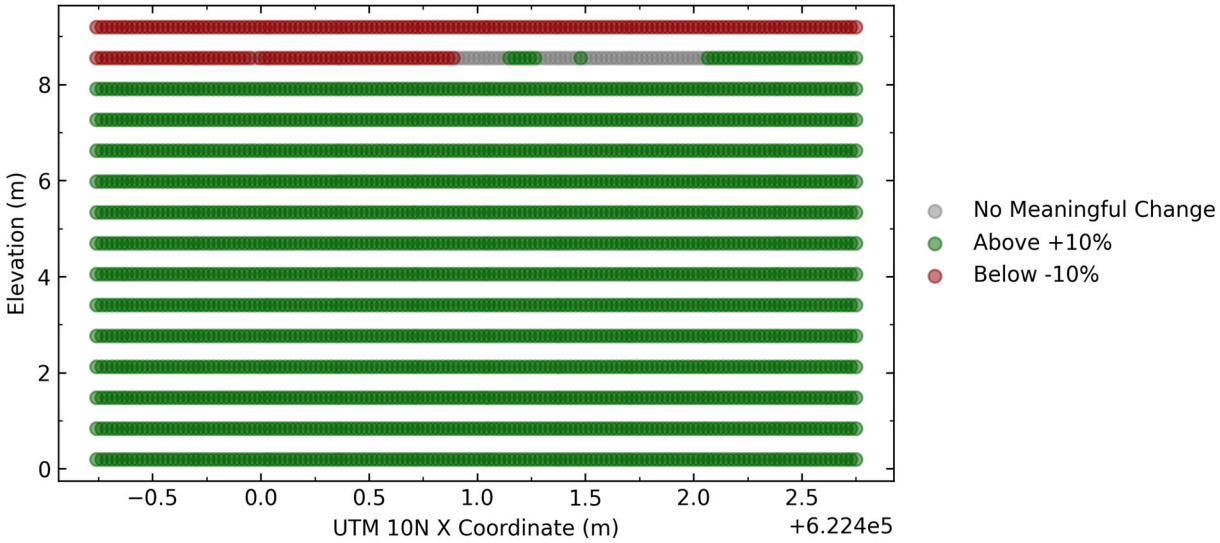
Grand Island
Site 2
EMI – Cross Section 7



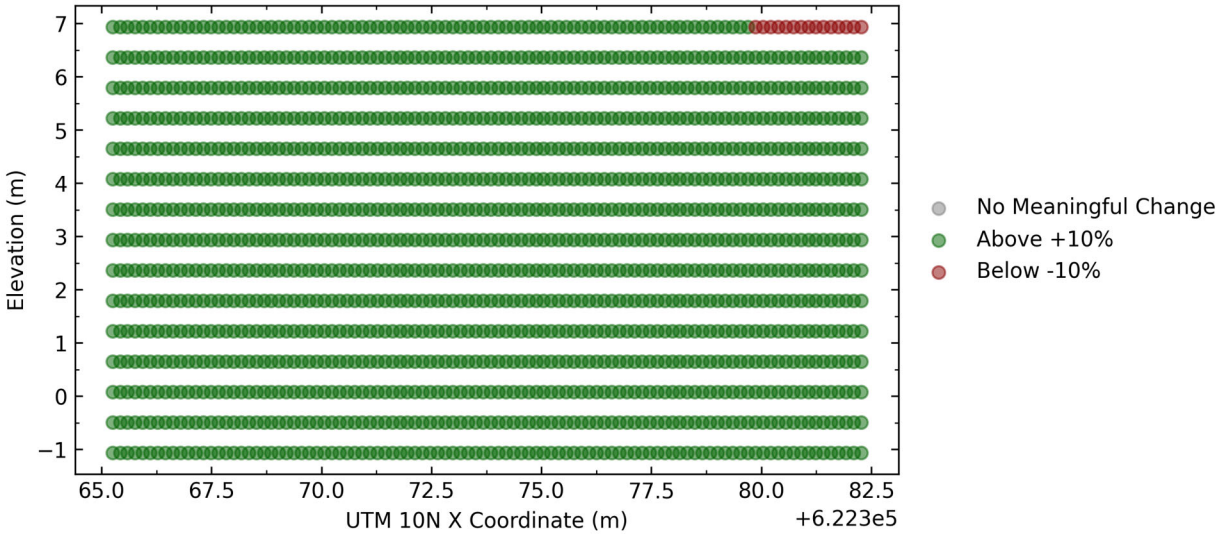
Grand Island
Site 3
EMI – Dry Crest



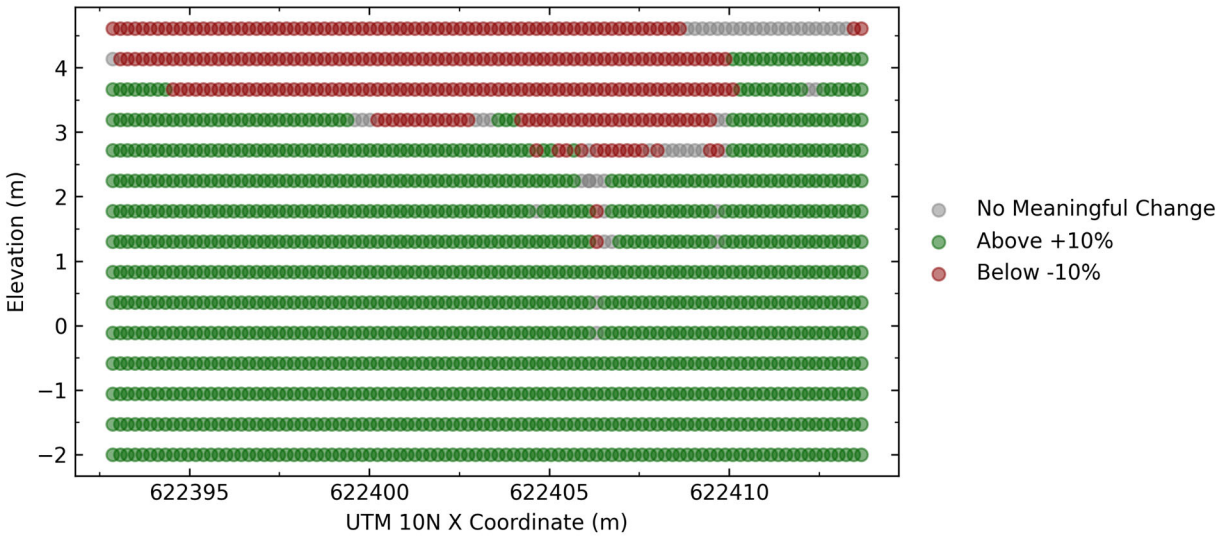
Grand Island
Site 3
EMI – Dry Mid-Slope 1



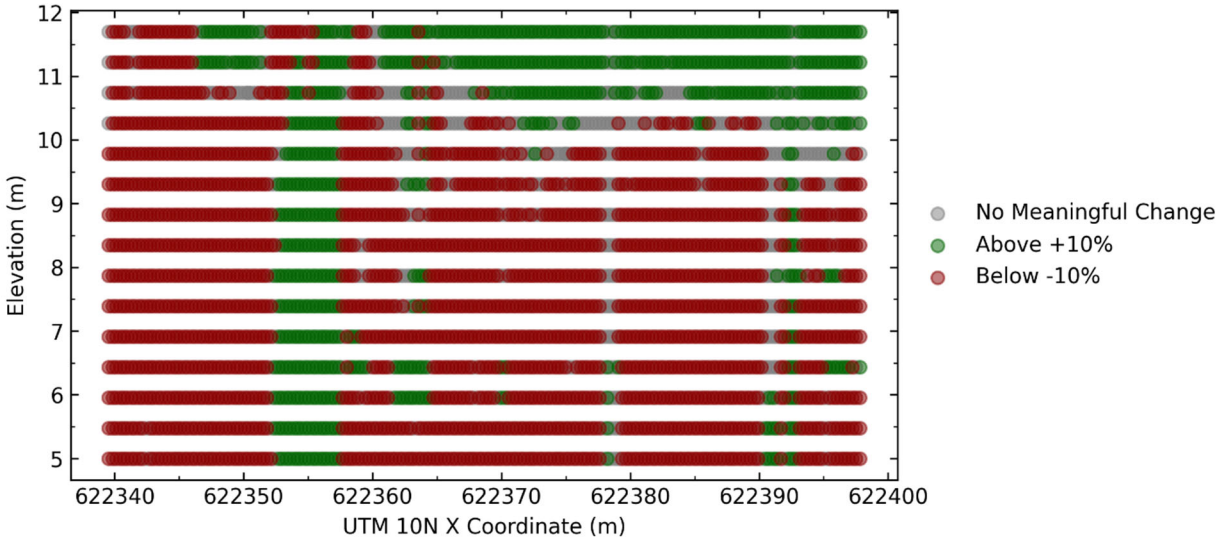
Grand Island
Site 3
EMI – Dry Mid-Slope 2



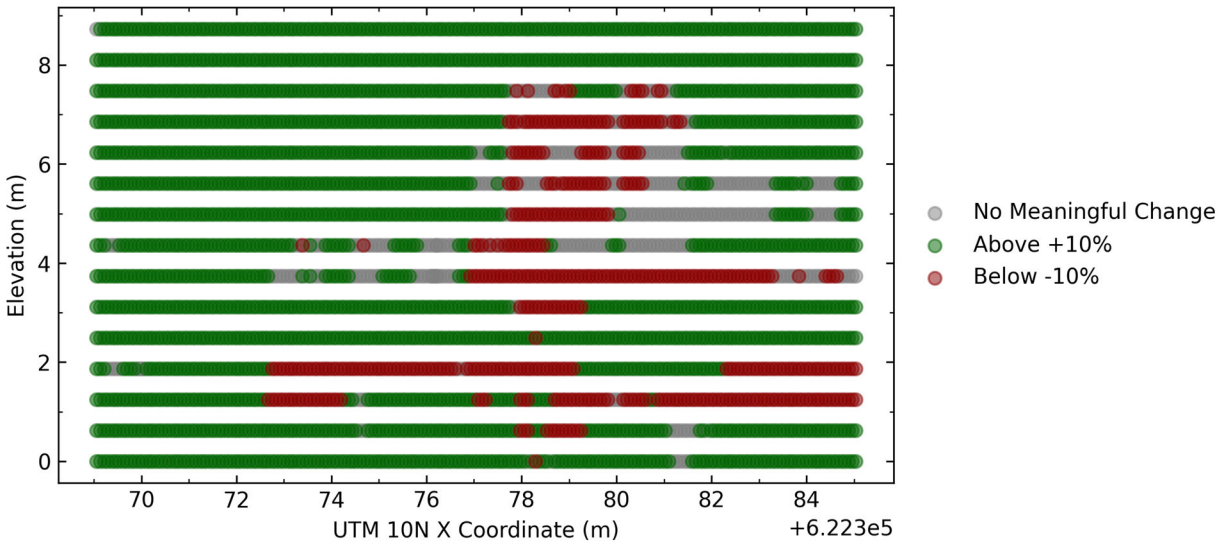
Grand Island
Site 3
EMI – Dry Toe



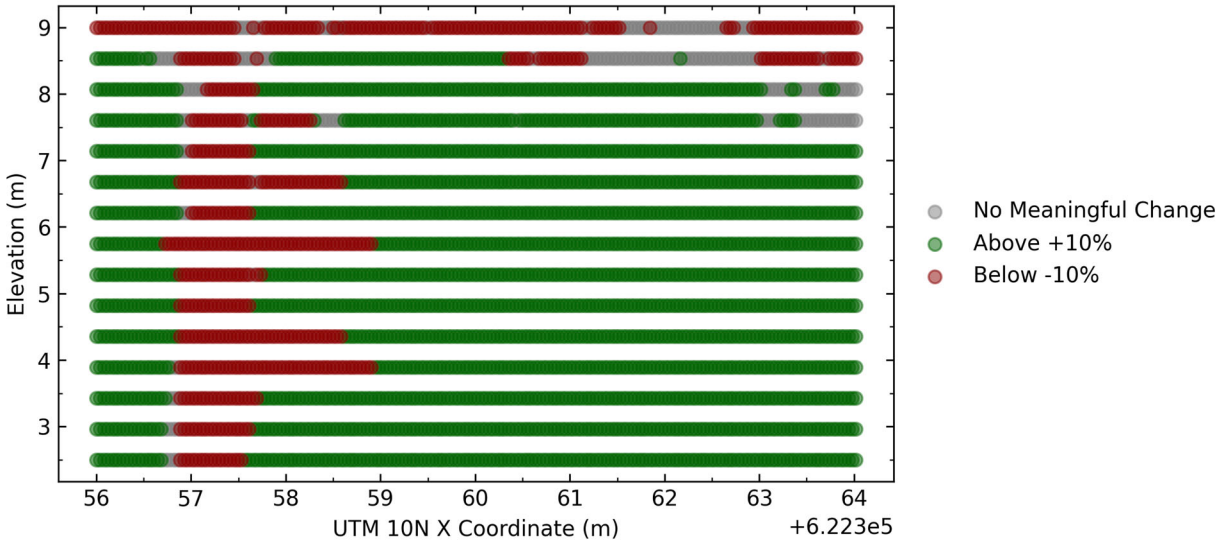
Grand Island
Site 3
EMI – Wet Crest



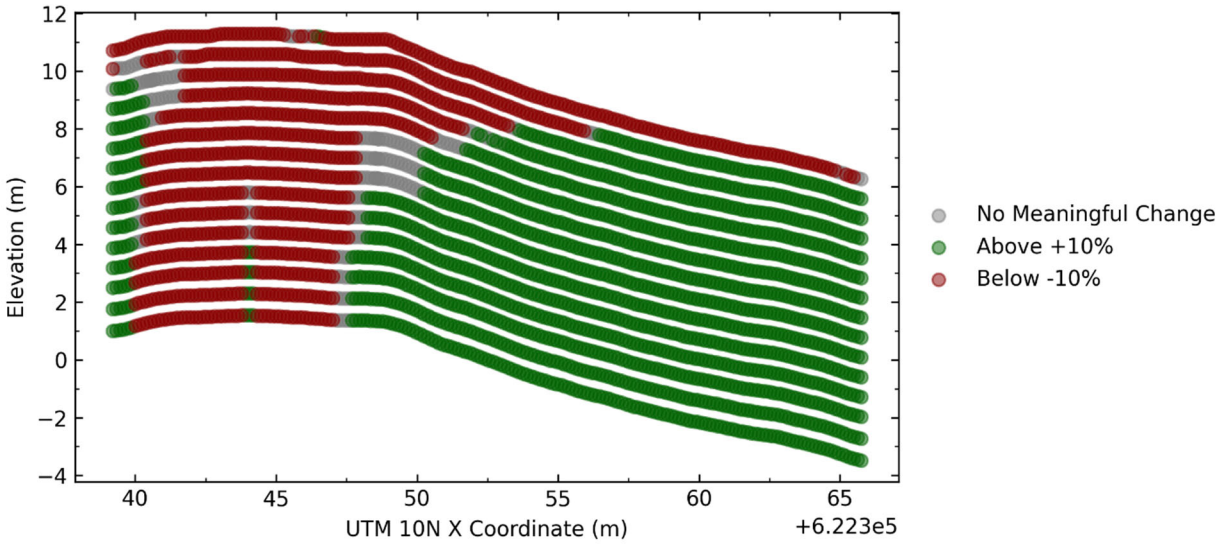
Grand Island
Site 3
EMI – Wet Toe 1



Grand Island
Site 3
EMI – Wet Toe 2

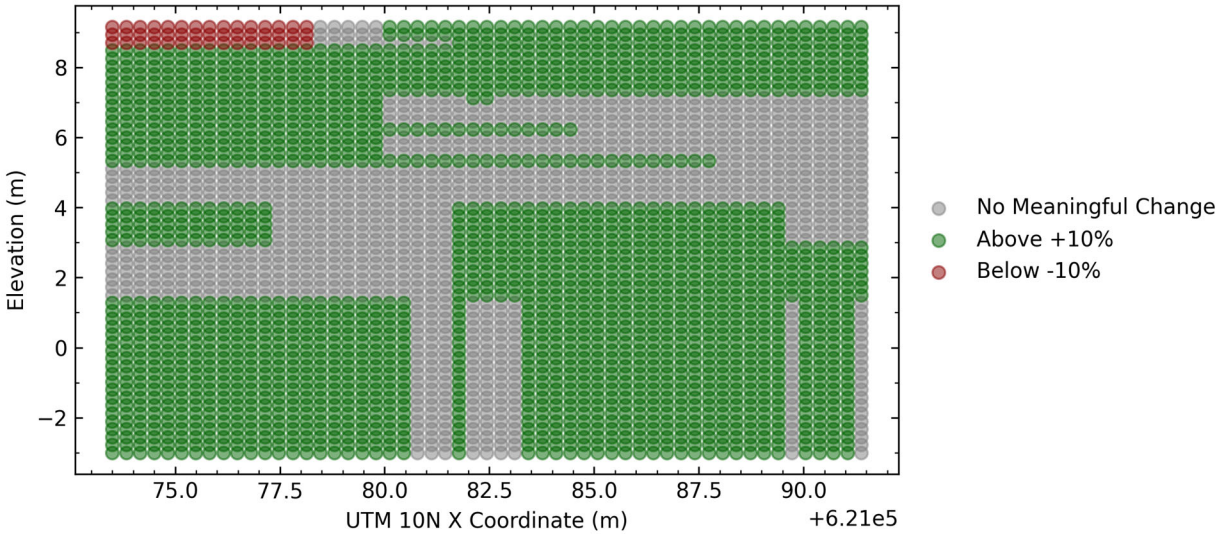


Grand Island
Site 3
EMI – Cross Section 6

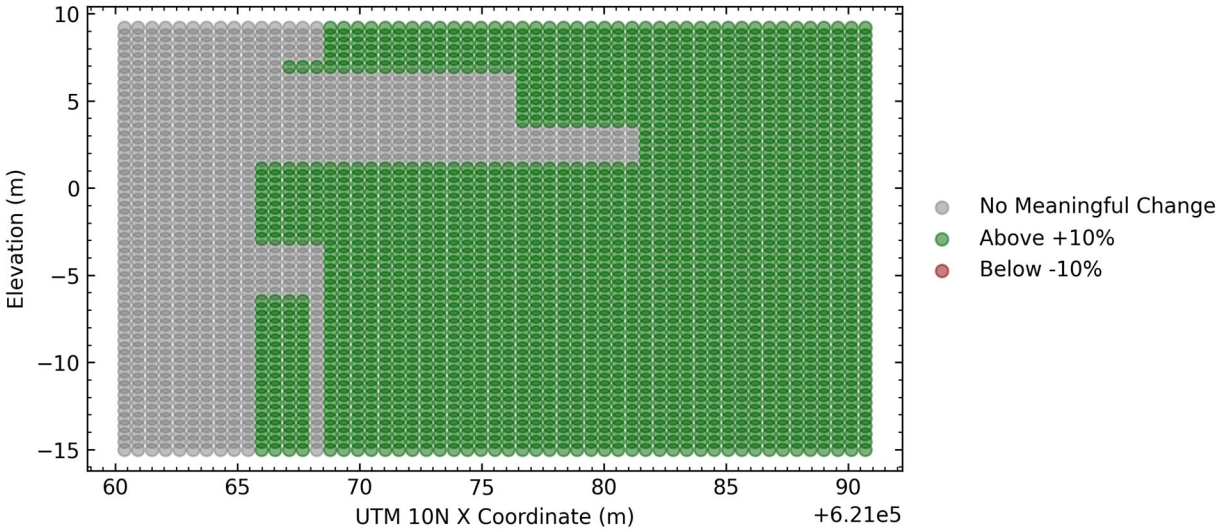


Subsurface MASW Change Detection Profiles

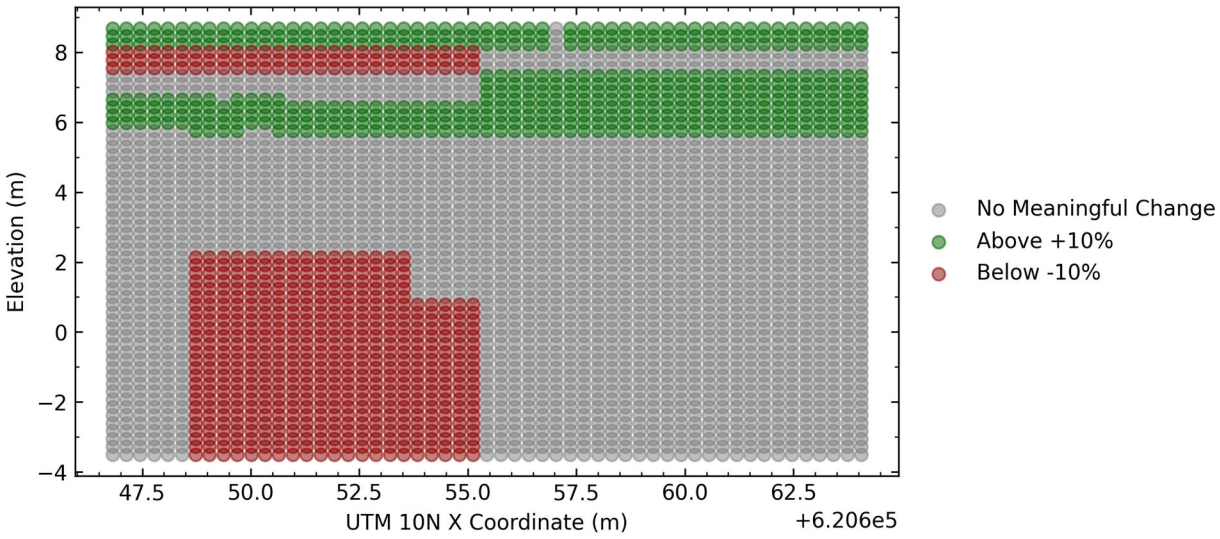
Grand Island
Site 1
MASW – Crest 3 ft Spacing



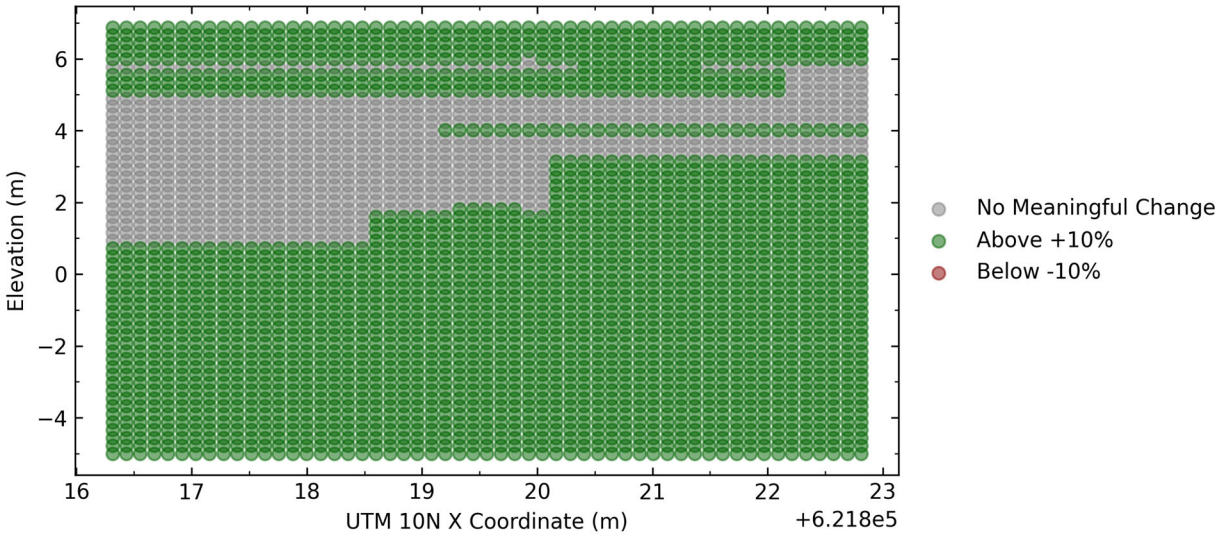
Grand Island
Site 1
MASW – Crest 5 ft Spacing



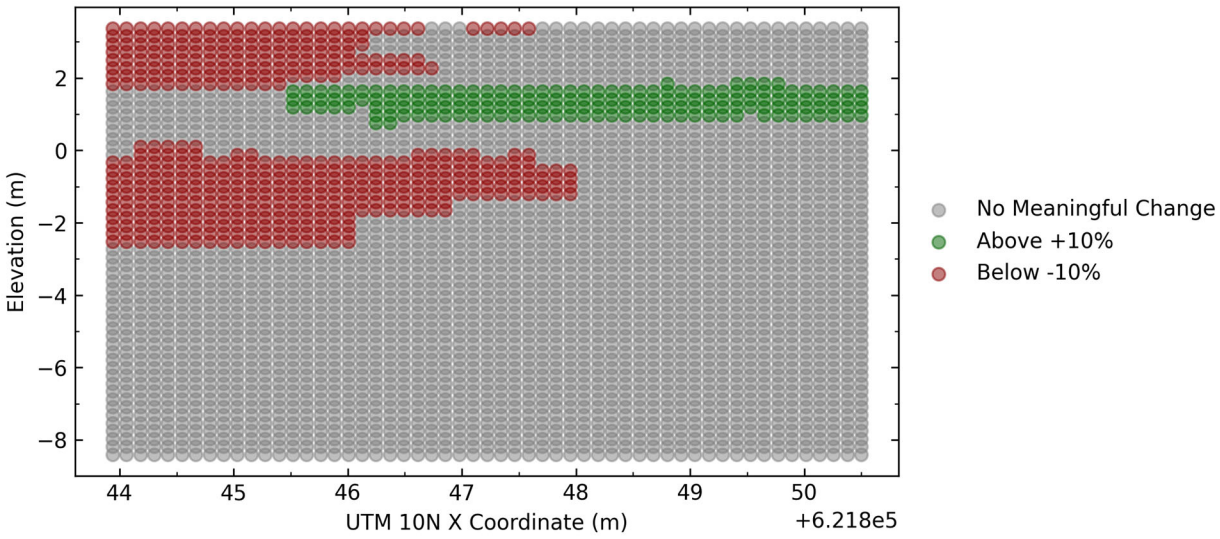
Grand Island
Site 1
MASW – Wet Toe



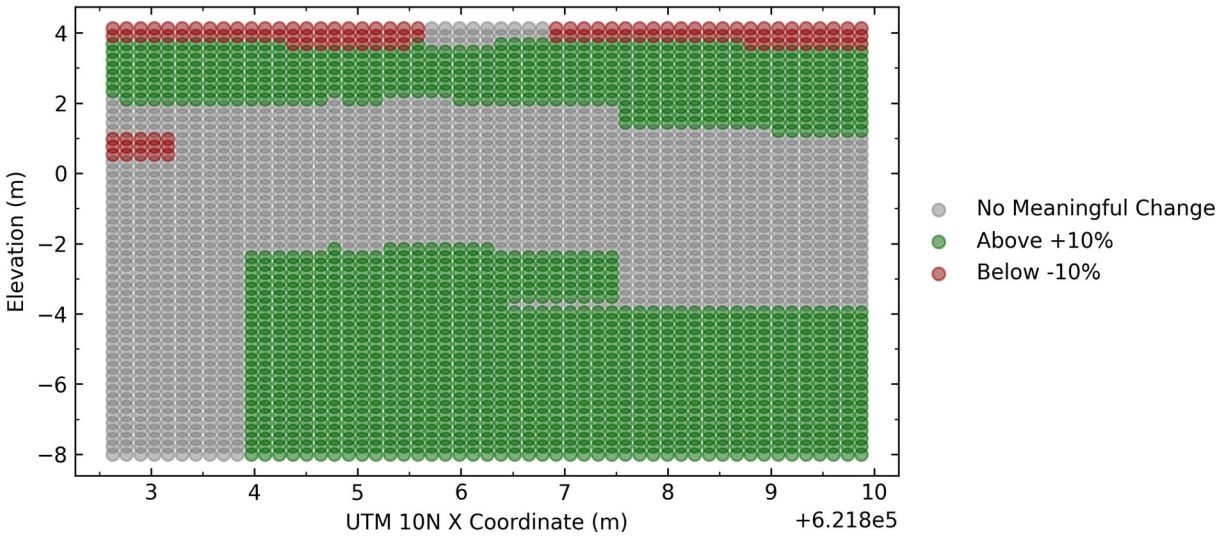
Grand Island
Site 2
MASW – Crest



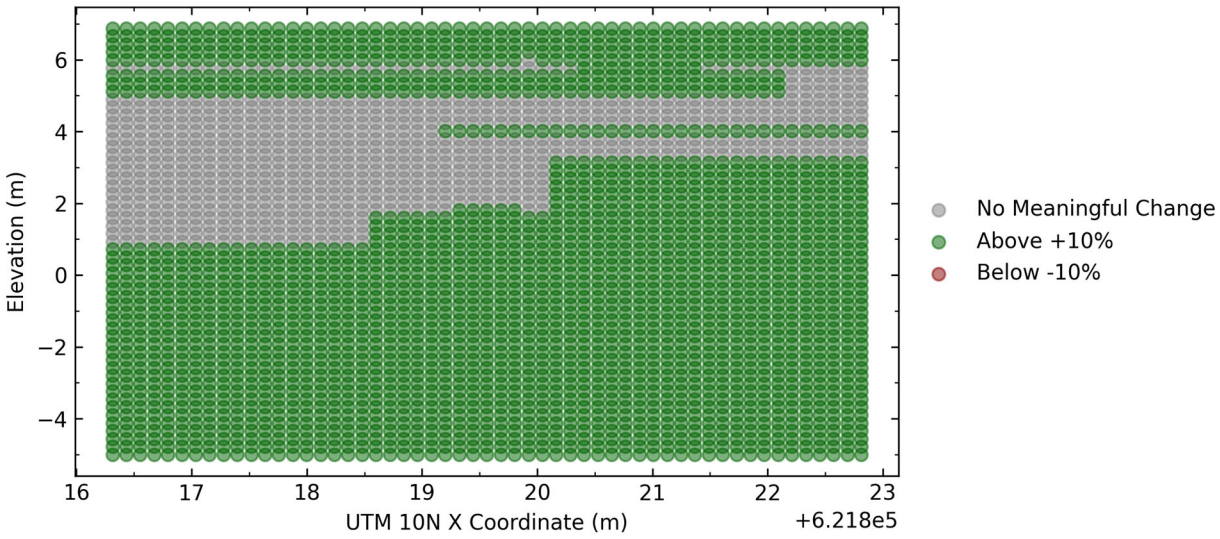
Grand Island
Site 2
MASW – Dry Toe



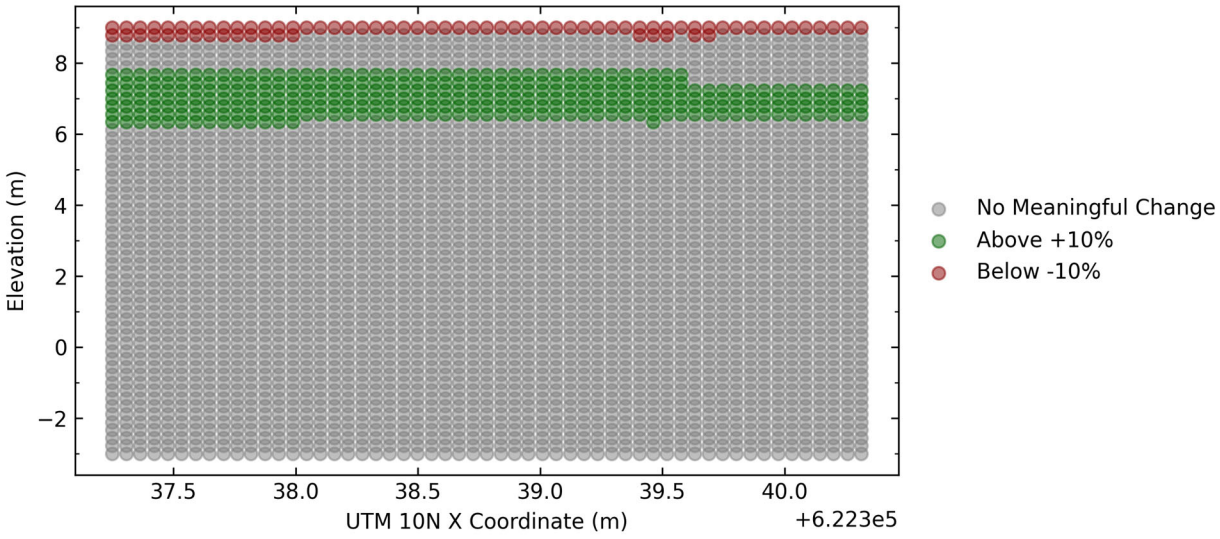
Grand Island
Site 2
MASW – Wet Toe



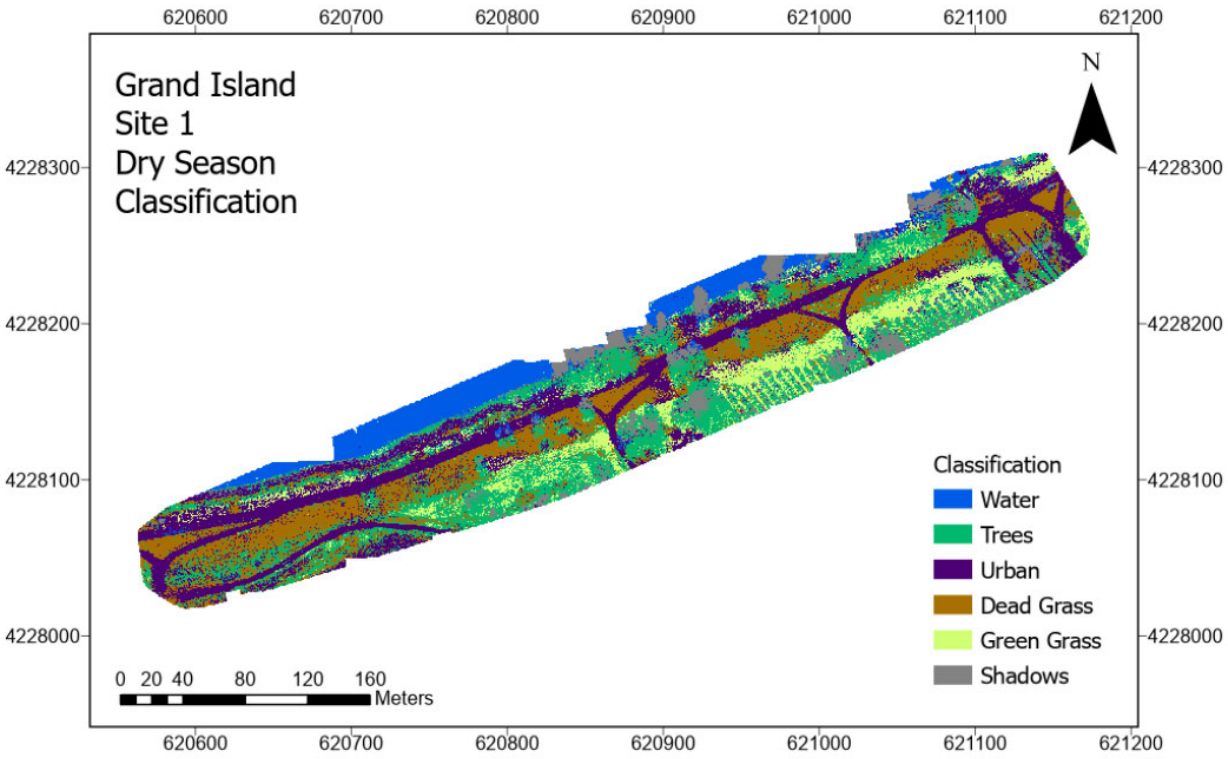
Grand Island
Site 3
MASW – Crest

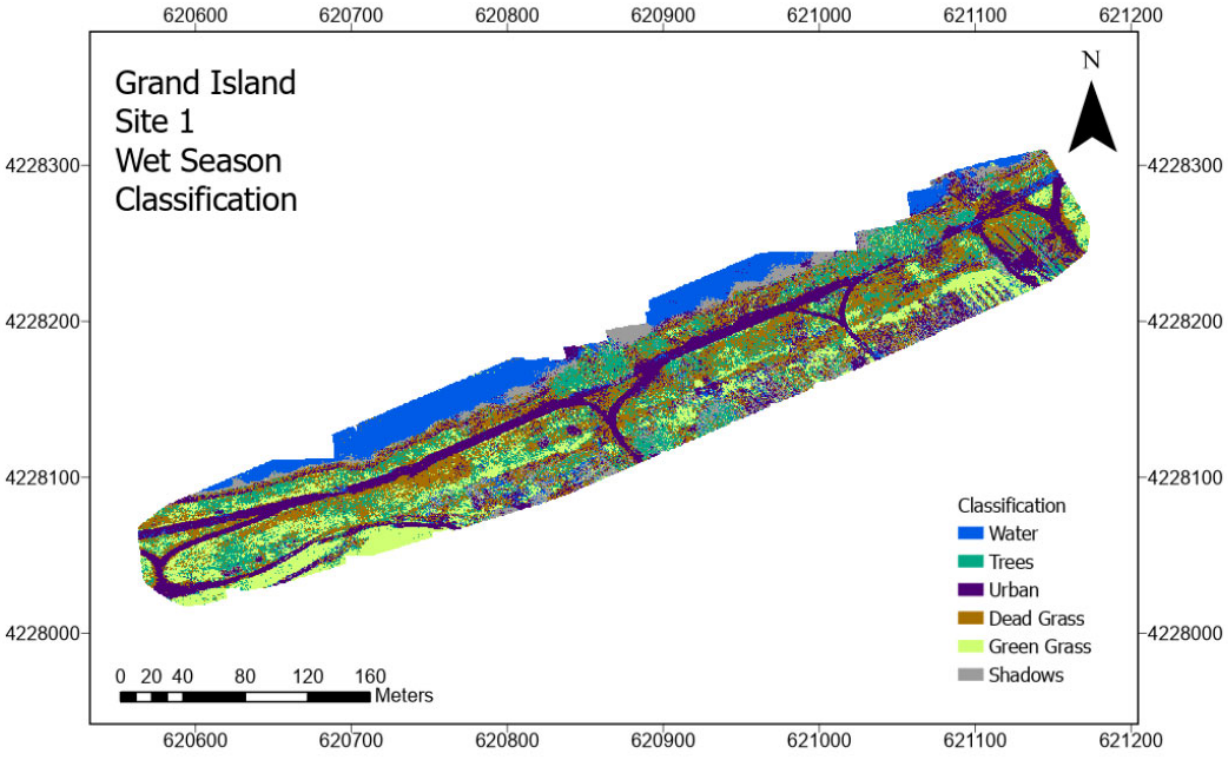


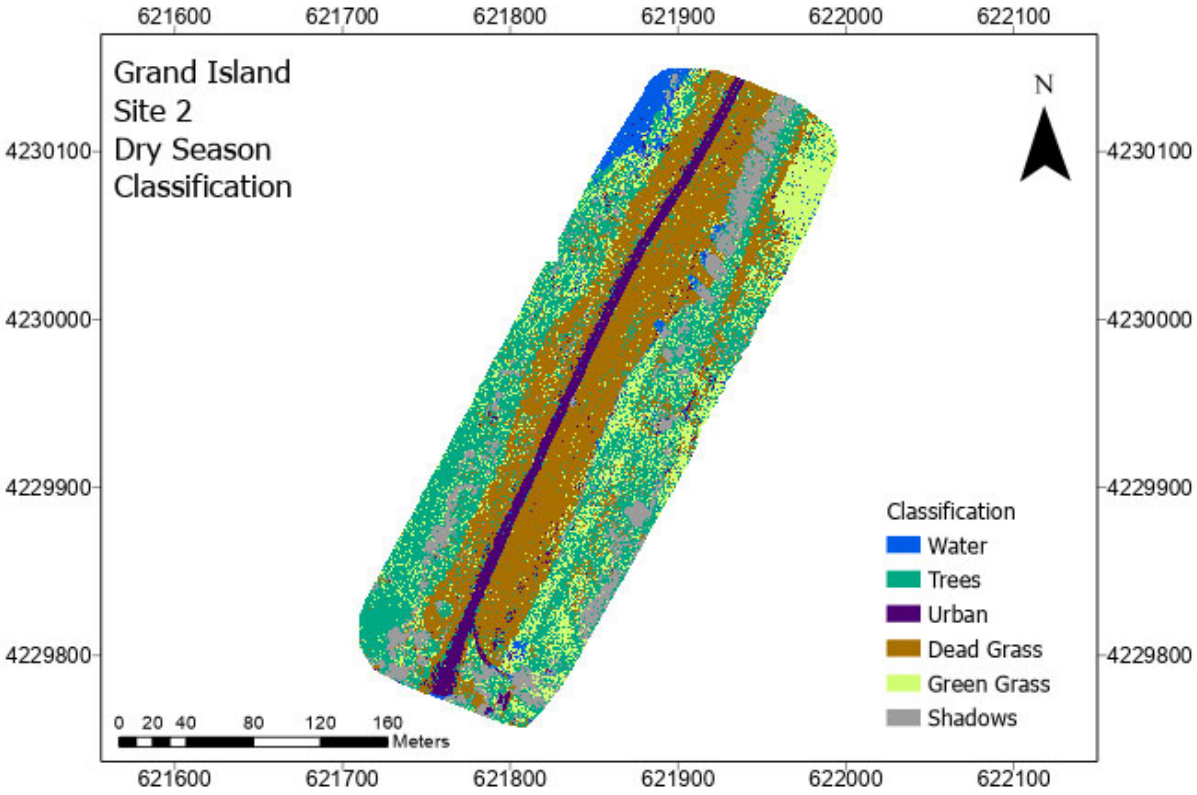
Grand Island
Site 3
MASW – Wet Toe

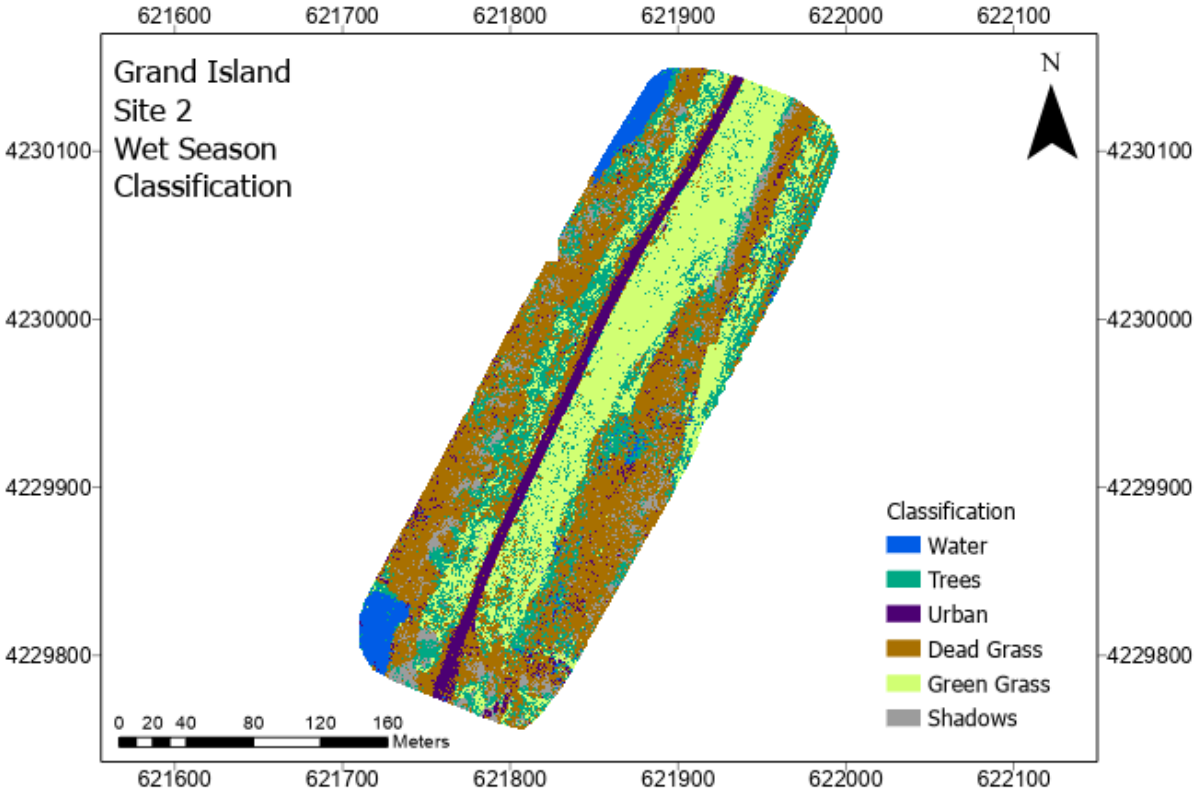


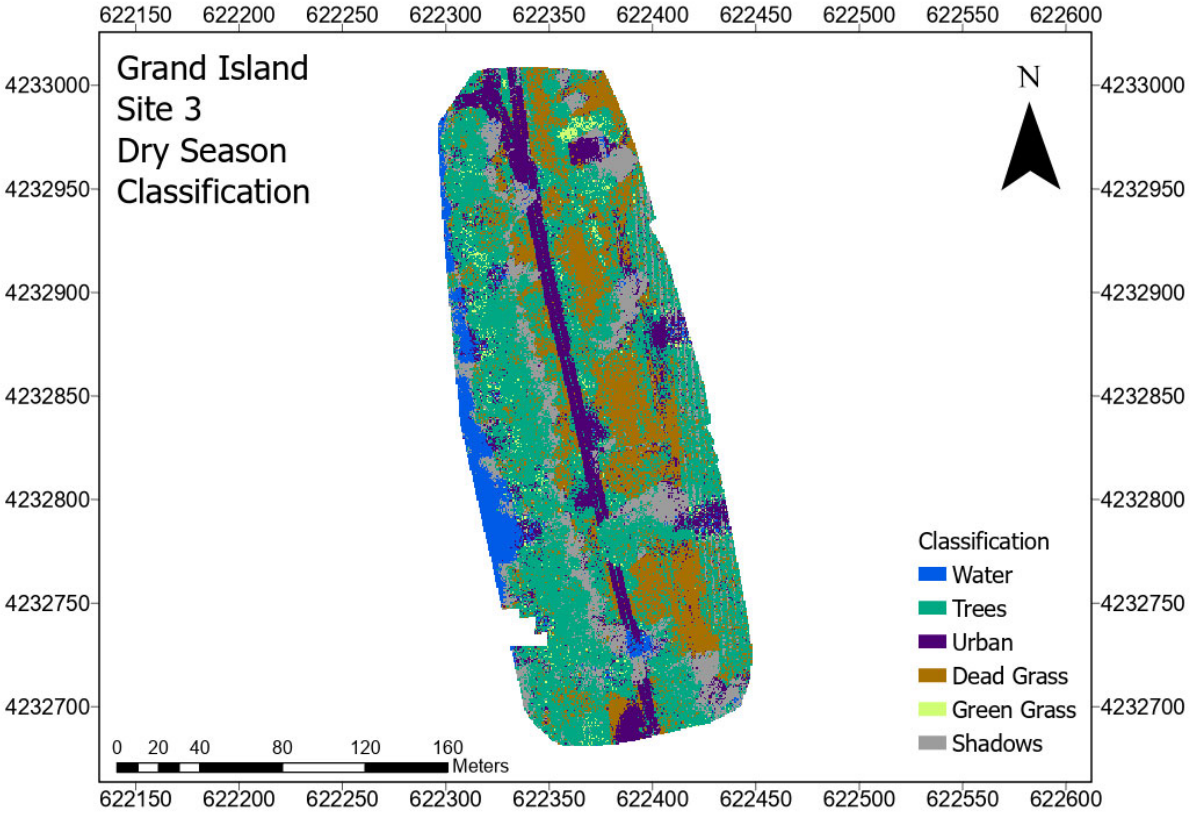
Surface Classification Results

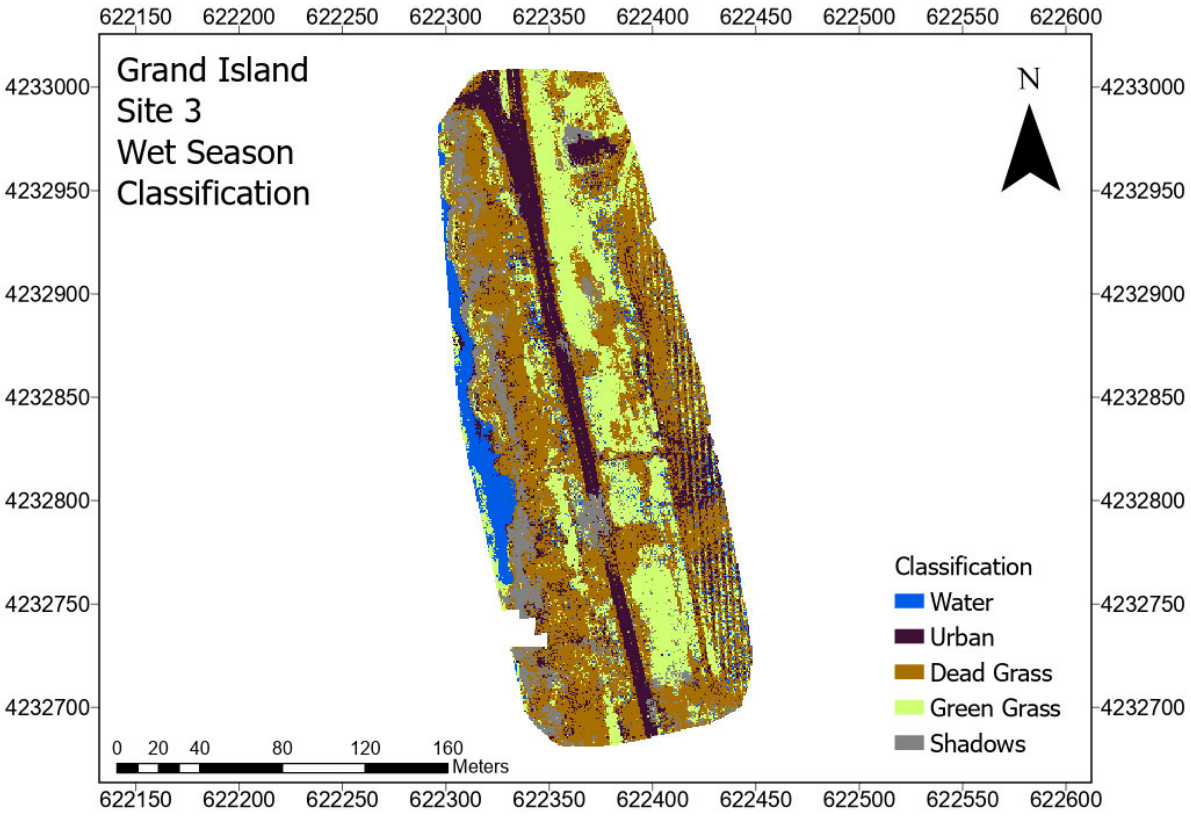




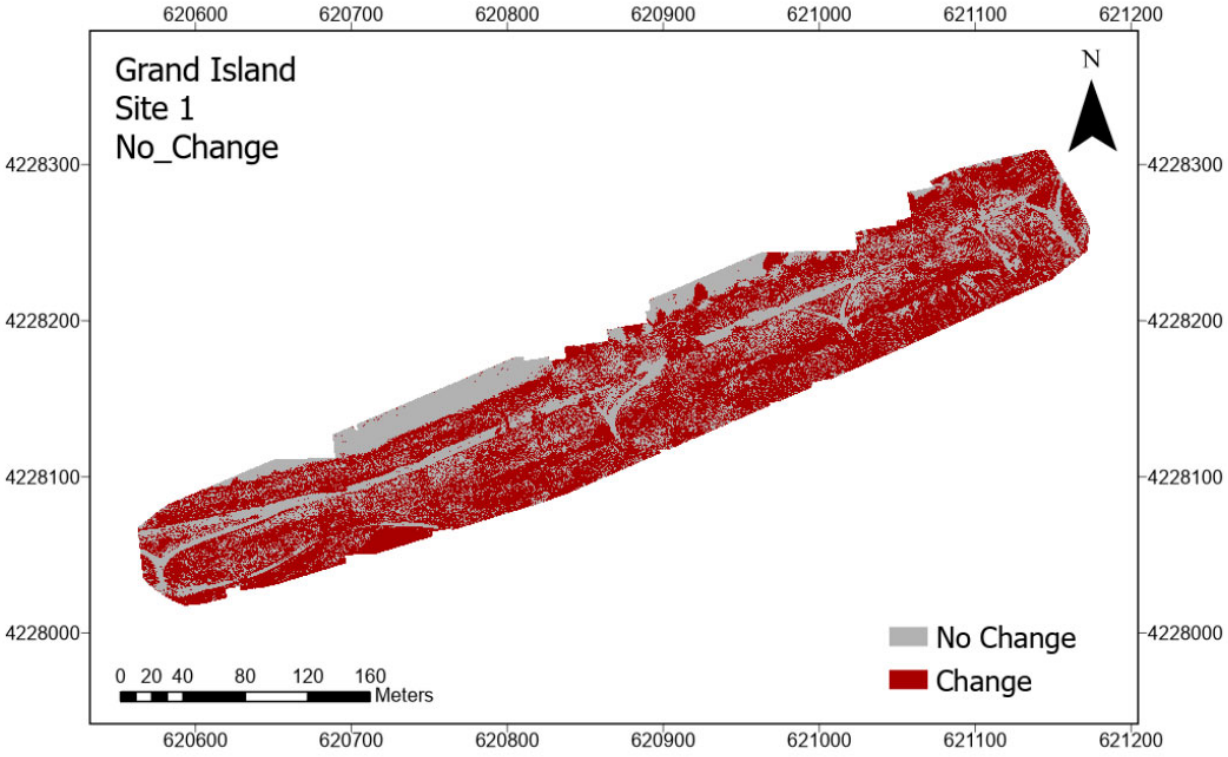

Grand Island
Site 2
Dry Season
Classification
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Classification
Water
Trees
Urban
Dead Grass
Green Grass
Shadows
0 20 40 80 120 160
Meters

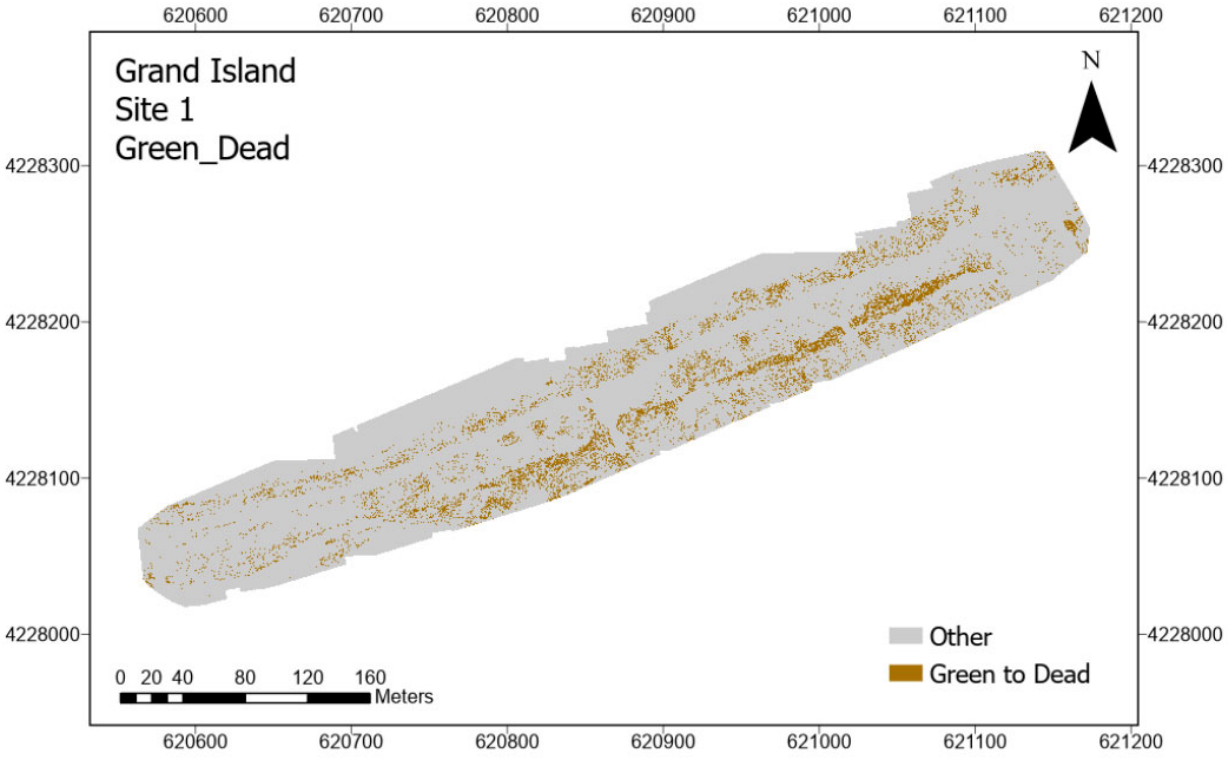

Grand Island
Site 2
Wet Season
Classification
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Classification
Water
Trees
Urban
Dead Grass
Green Grass
Shadows
0 20 40 80 120 160
Meters

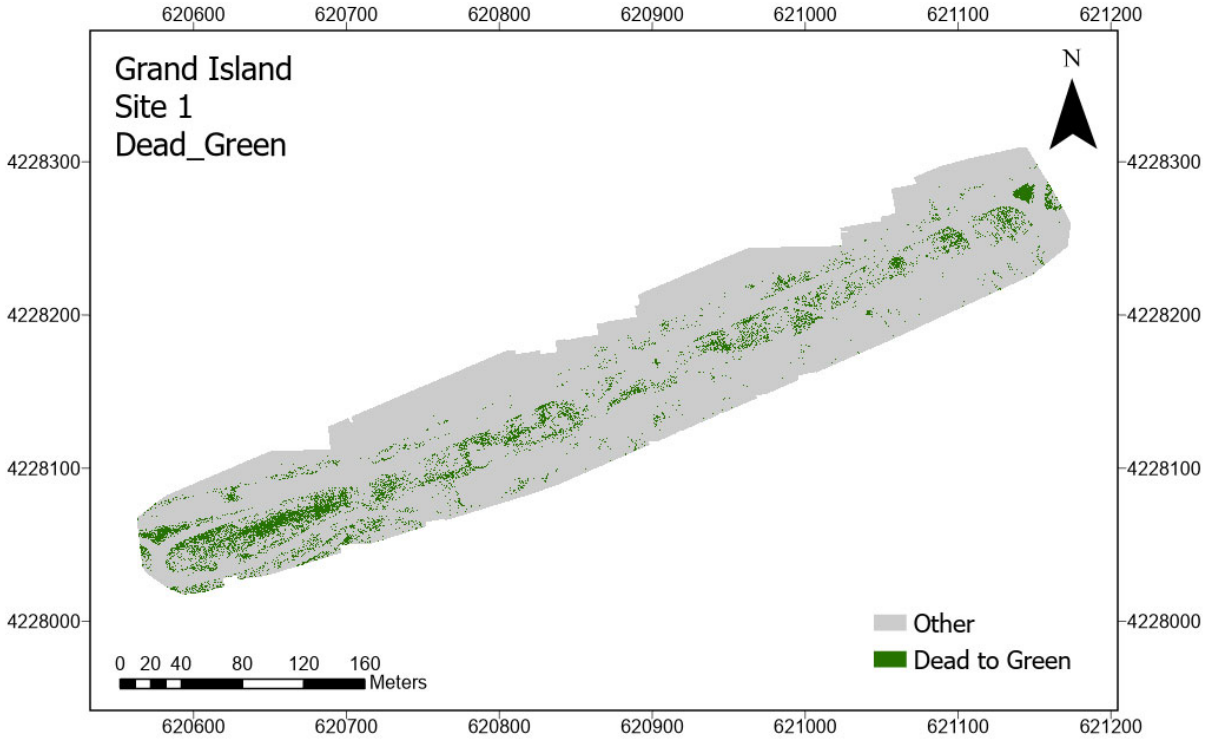

Grand Island
Site 3
Dry Season
Classification
N
622150 622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700
Classification
Water
Trees
Urban
Dead Grass
Green Grass
Shadows
0 20 40 80 120 160
Meters

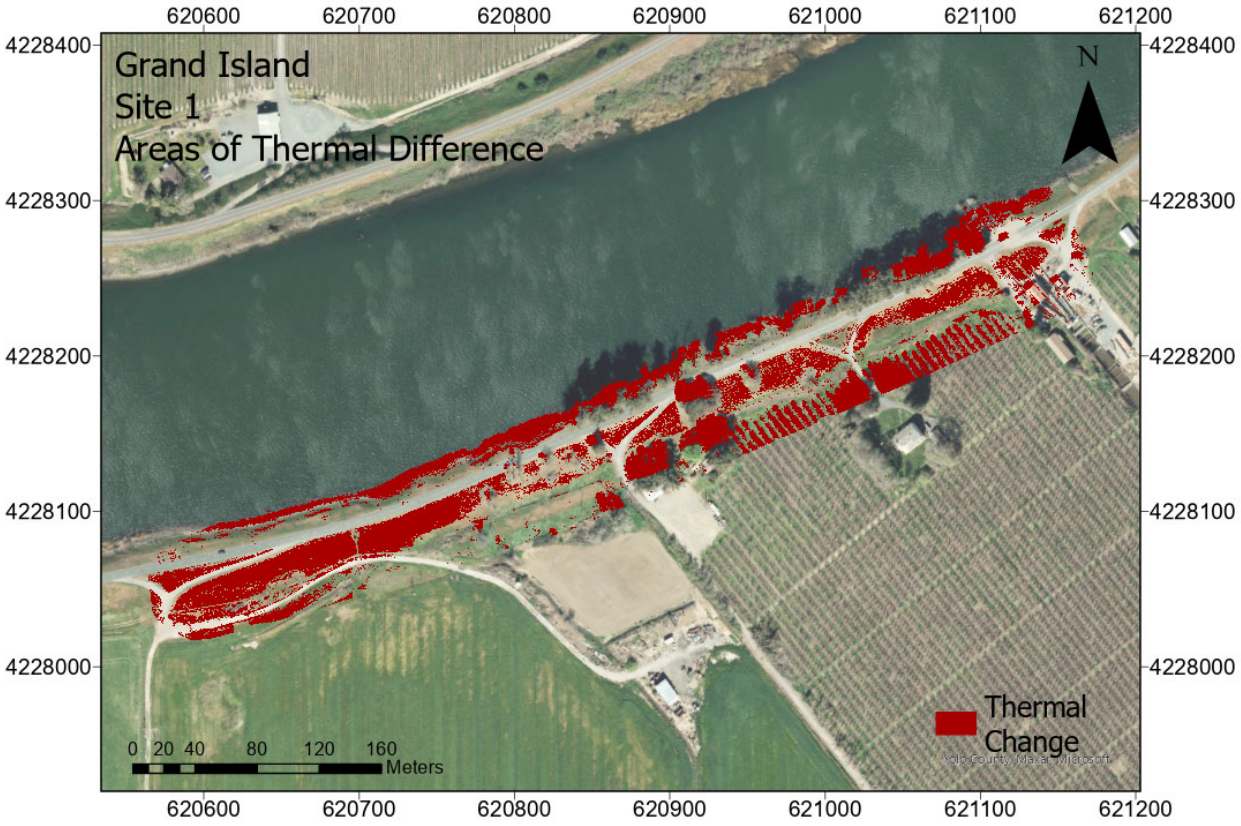

Grand Island
Site 3
Wet Season
Classification
N
622150 622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700
Classification
Water
Urban
Dead Grass
Green Grass
Shadows
0 20 40 80 120 160
Meters

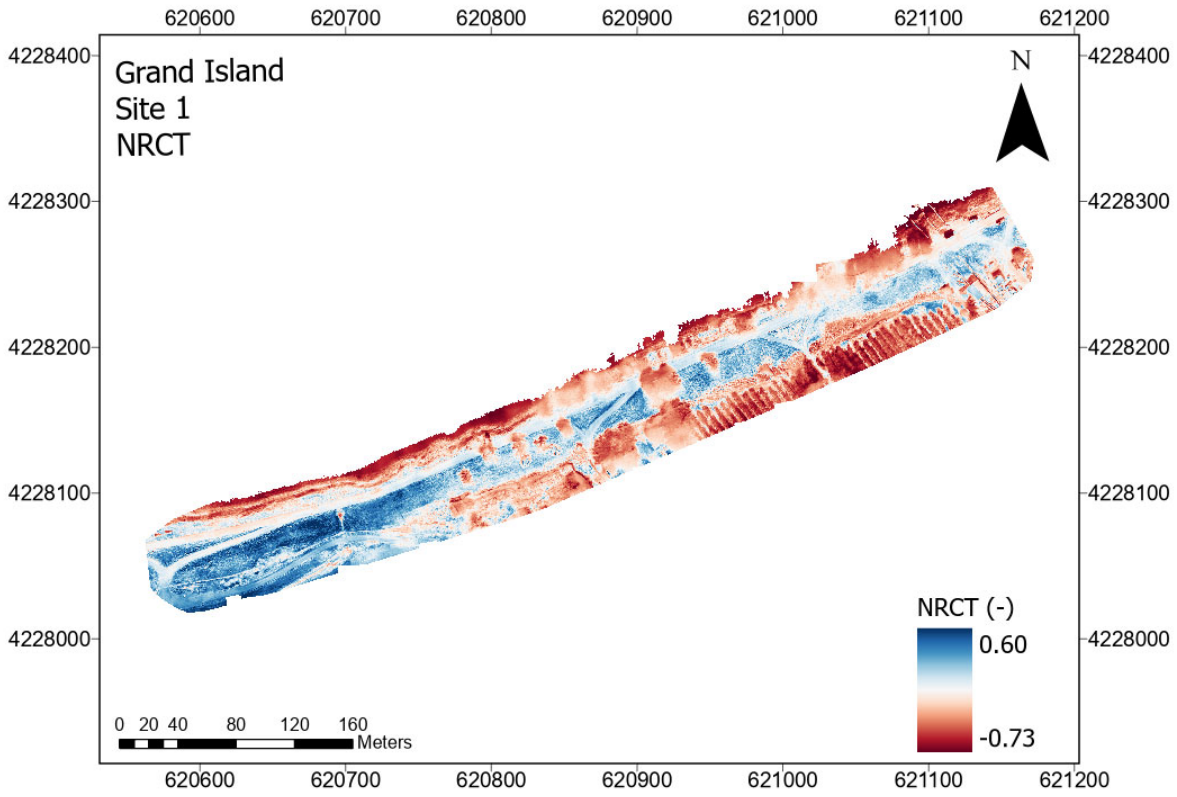
Surface Change Detection Results – Site 1

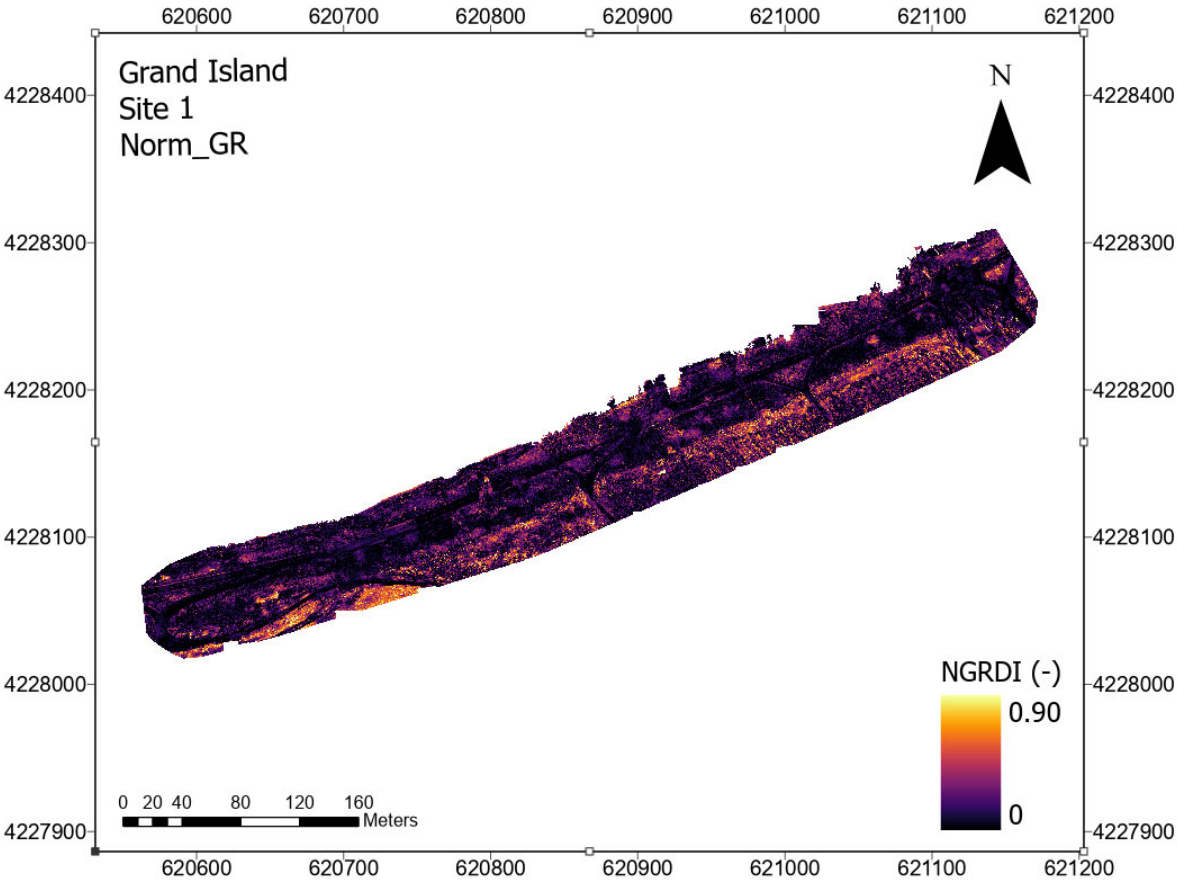


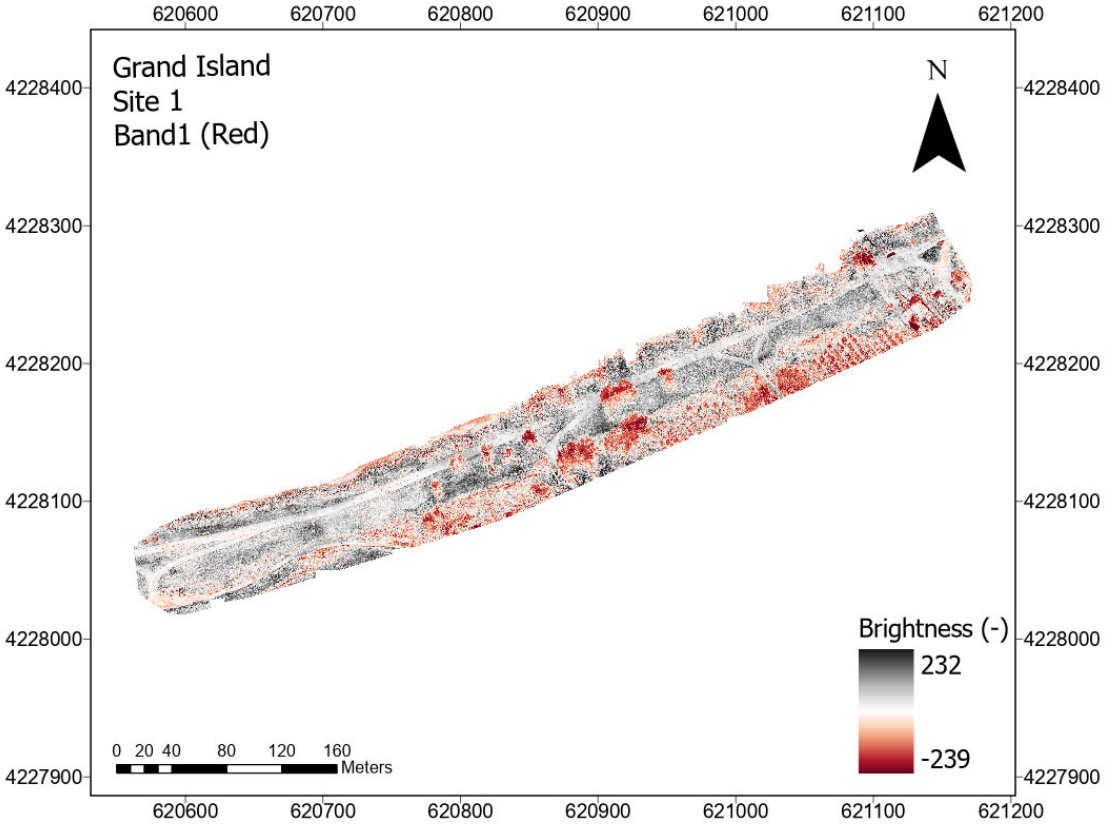


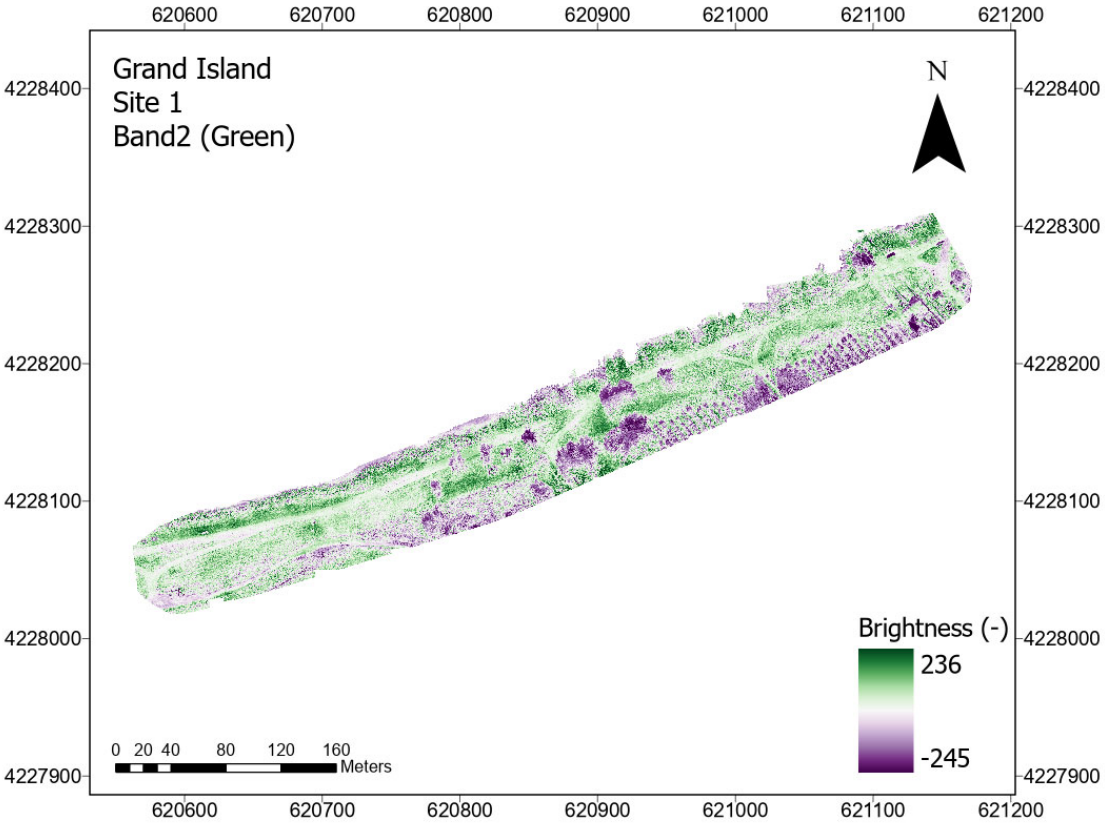

Grand Island
Site 1
Dead_Green
N
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
0 20 40 80 120 160
Meters
Other
Dead to Green

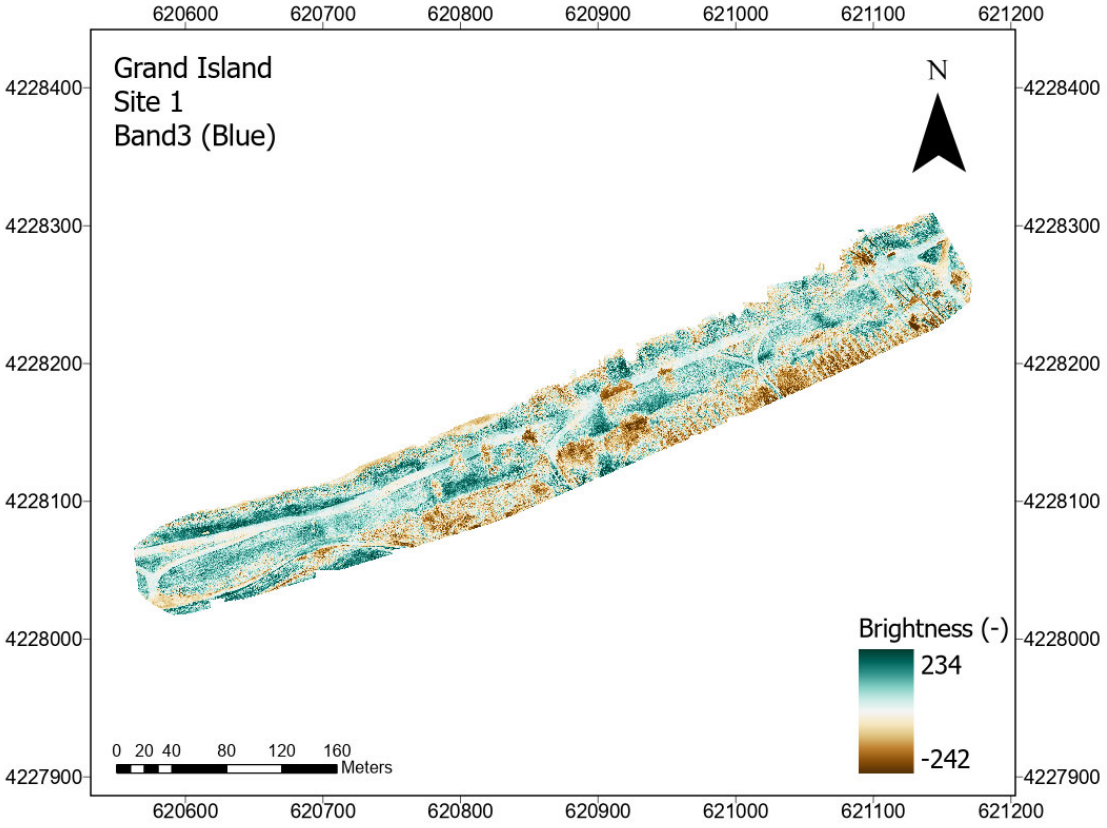

Grand Island
Site 1
Areas of Thermal Difference
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
0 20 40 80 120 160
Meters
Thermal
Change

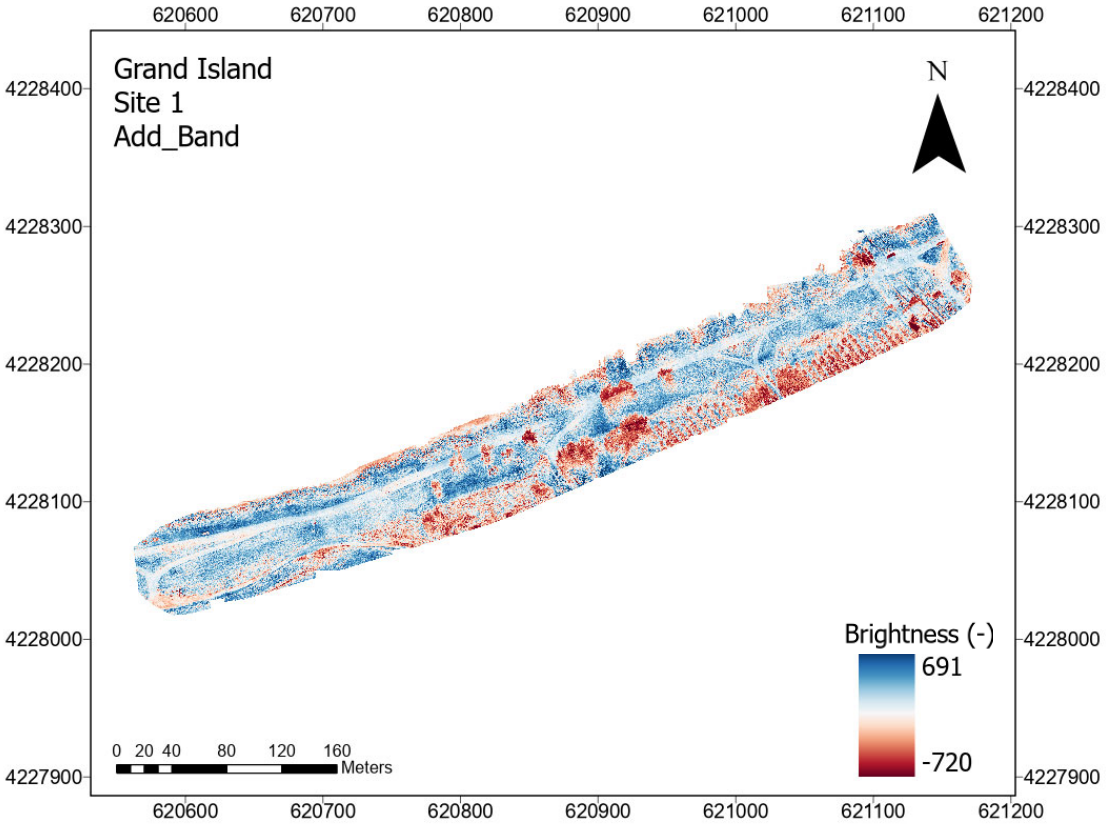

Grand Island
Site 1
NRCT
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
NRCT (-)
0.60
-0.73
0 20 40 80 120 160
Meters

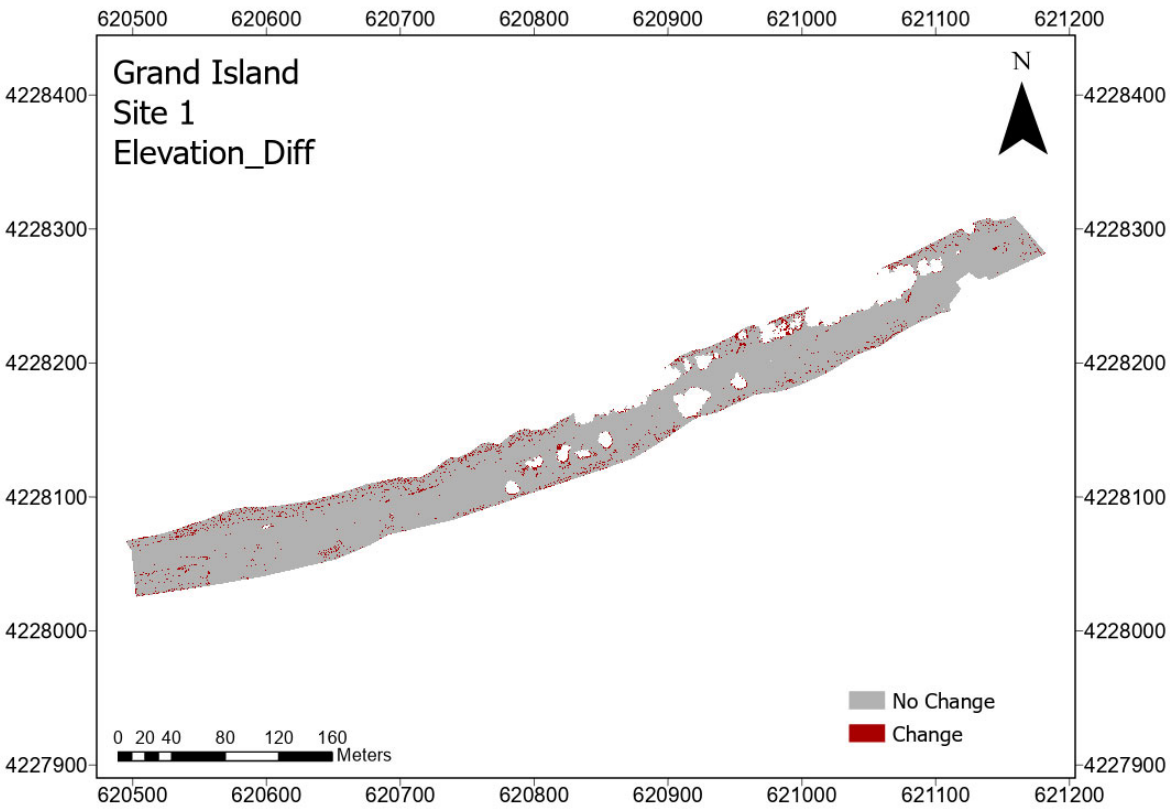

Grand Island
Site 1
Norm_GR
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
NGRDI (-)
0.90
0
0 20 40 80 120 160
Meters

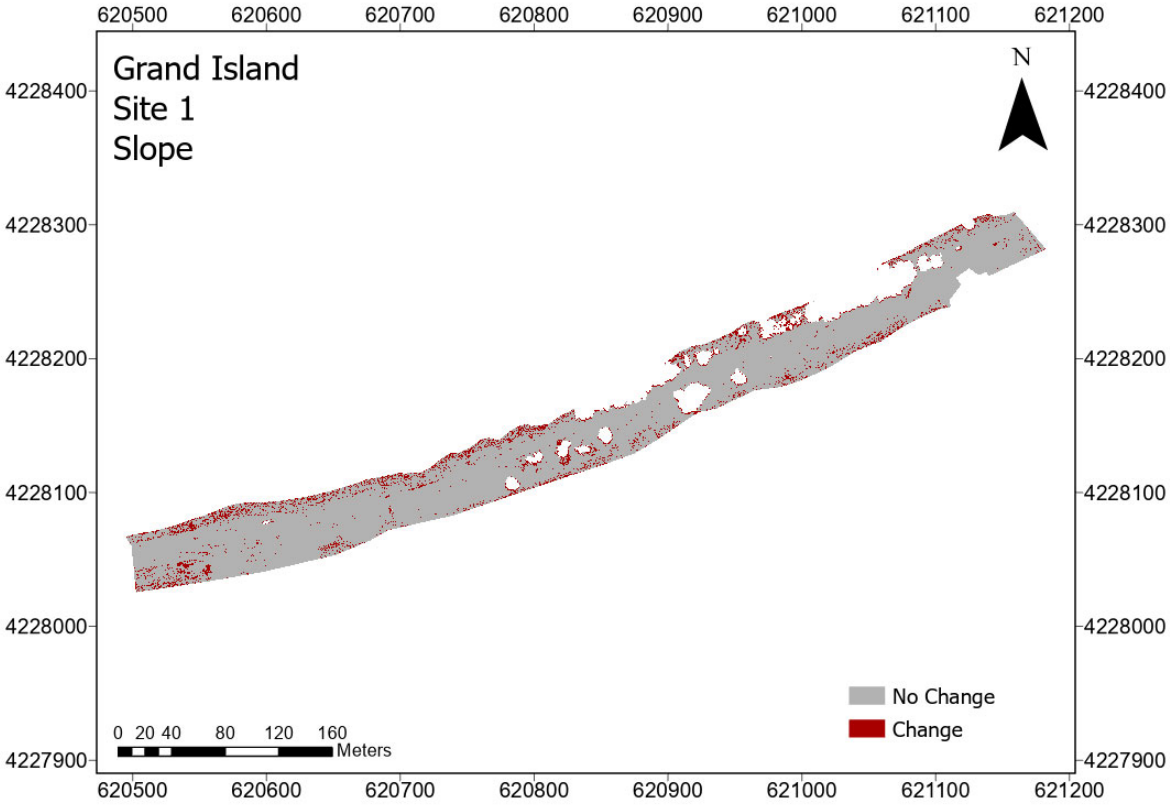

Grand Island
Site 1
Band1 (Red)
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
Brightness (-)
232
-239
0 20 40 80 120 160
Meters

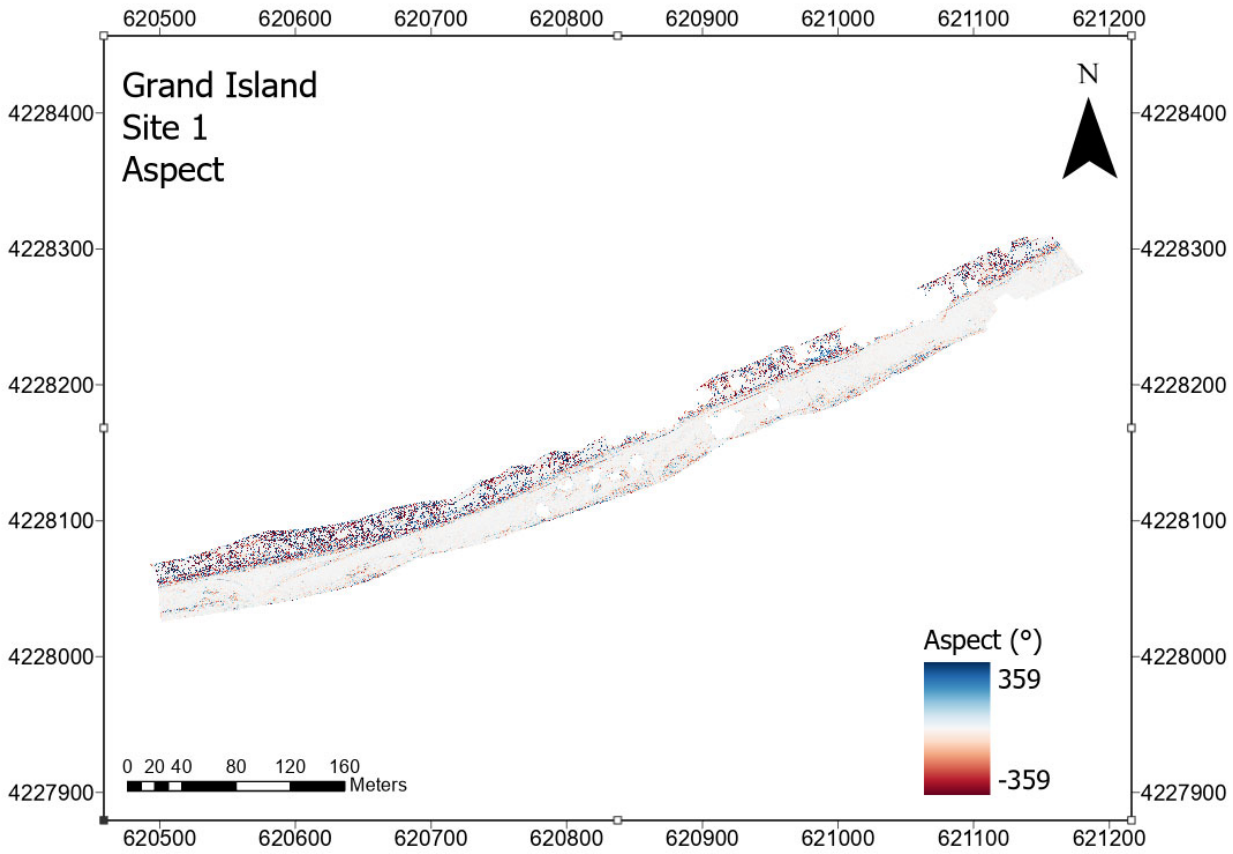

Grand Island
Site 1
Band2 (Green)
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
Brightness (-)
236
-245
0 20 40 80 120 160
Meters


Grand Island
Site 1
Band3 (Blue)
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
Brightness (-)
234
-242
0 20 40 80 120 160
Meters

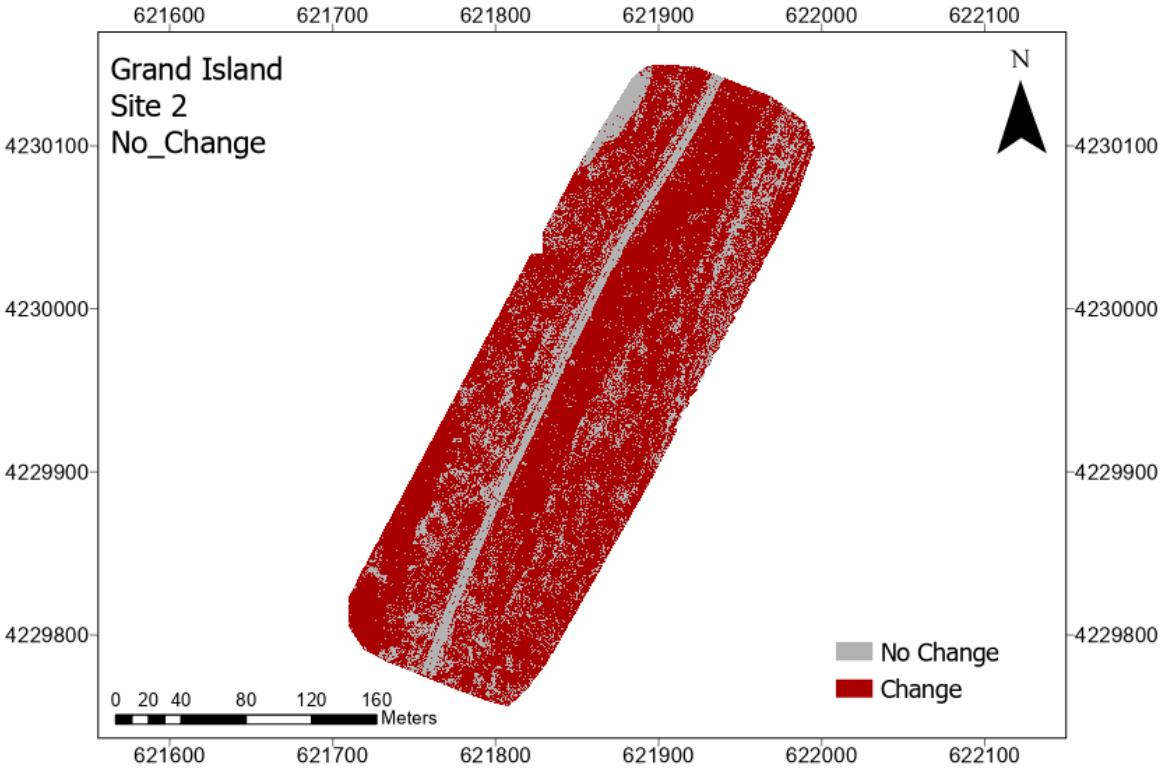

Grand Island
Site 1
Add_Band
N
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
Brightness (-)
691
-720
0 20 40 80 120 160
Meters

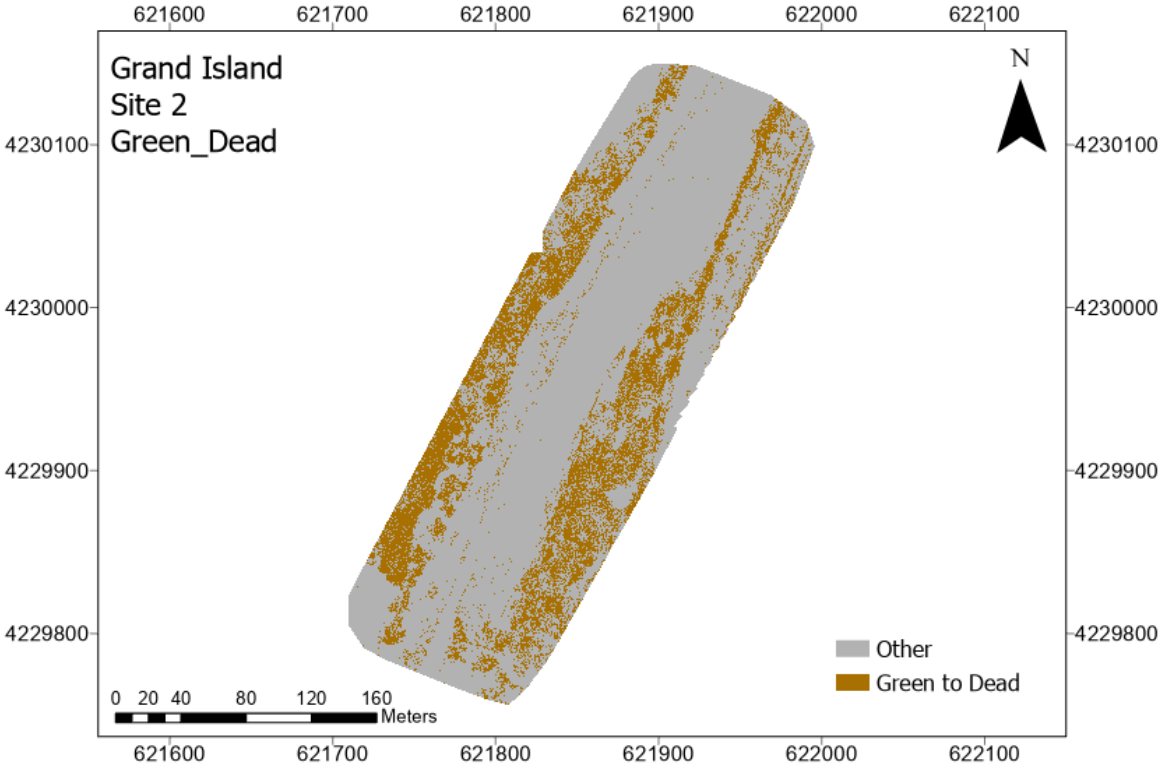

Grand Island
Site 1
Elevation_Diff
N
620500
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
No Change
Change
0 20 40 80 120 160
Meters

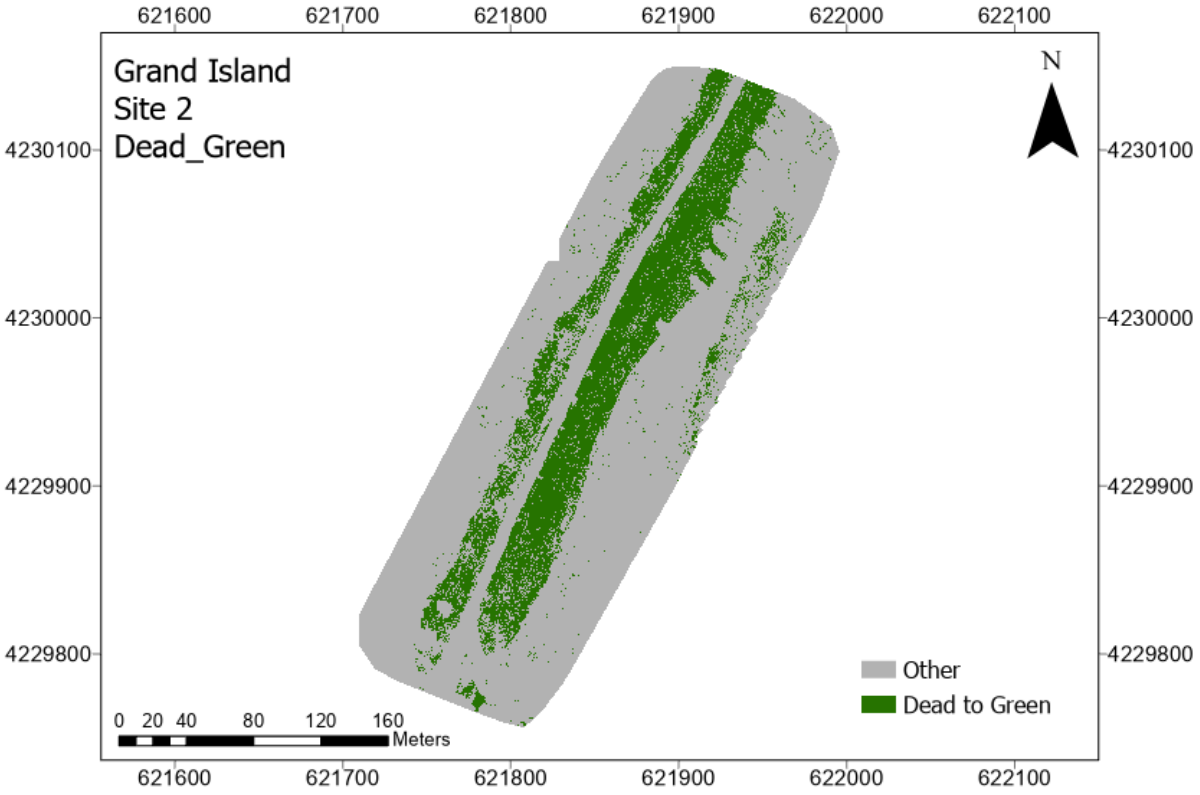

Grand Island
Site 1
Slope
N
620500
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
No Change
Change
0 20 40 80 120 160
Meters

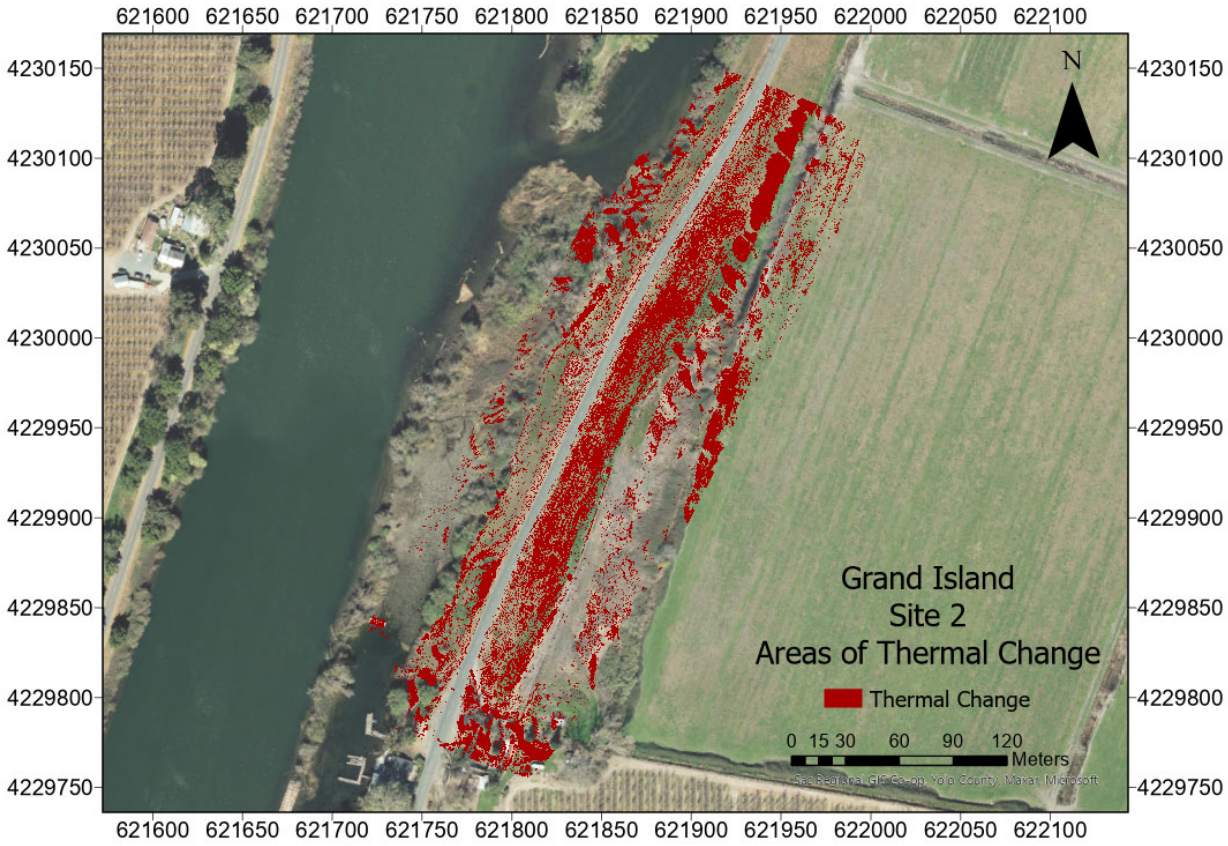

Grand Island
Site 1
Aspect
N
620500
620600
620700
620800
620900
621000
621100
621200
4228400
4228300
4228200
4228100
4228000
4227900
Aspect (°)
359
-359
0 20 40 80 120 160
Meters

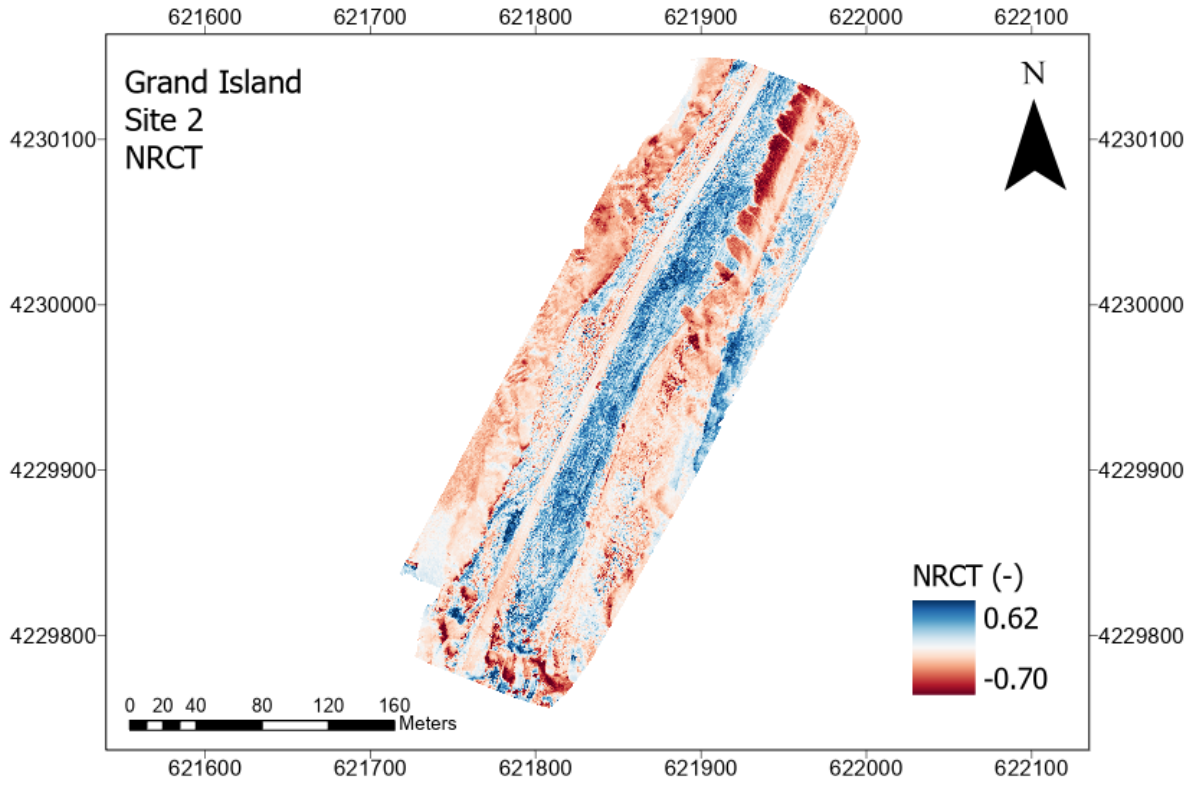
Surface Change Detection Results – Site 2

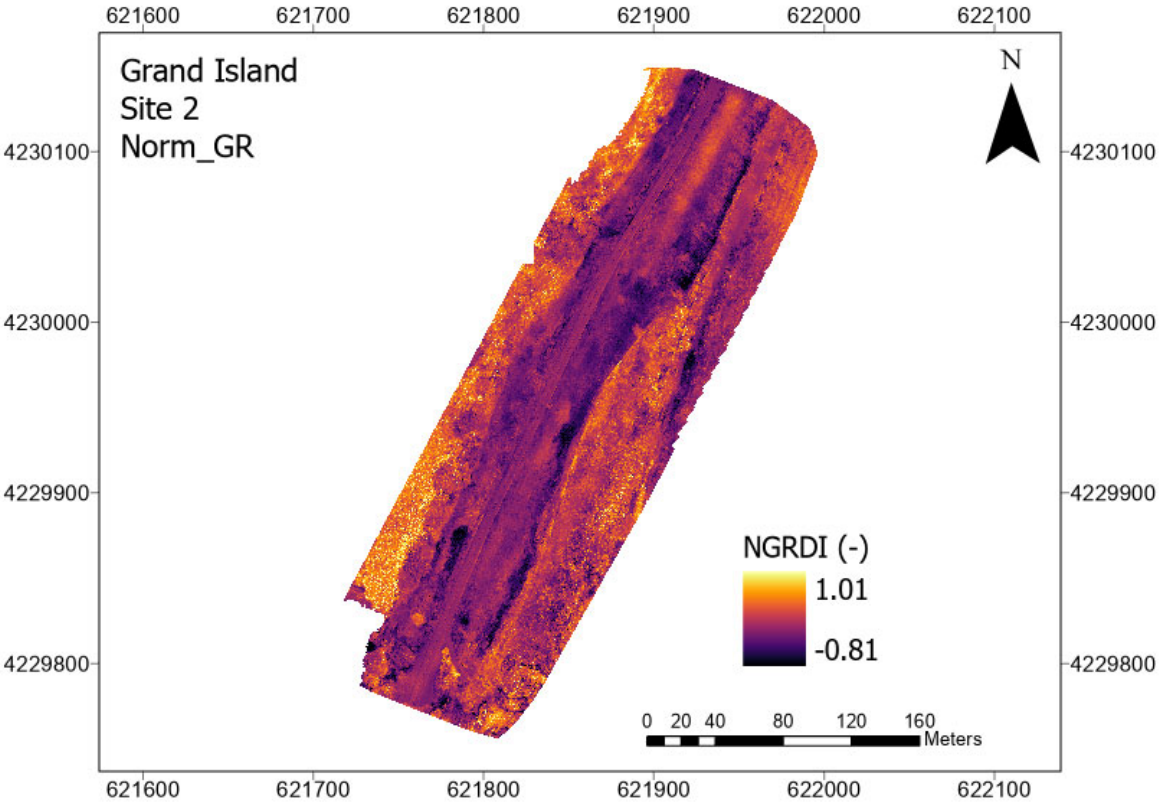


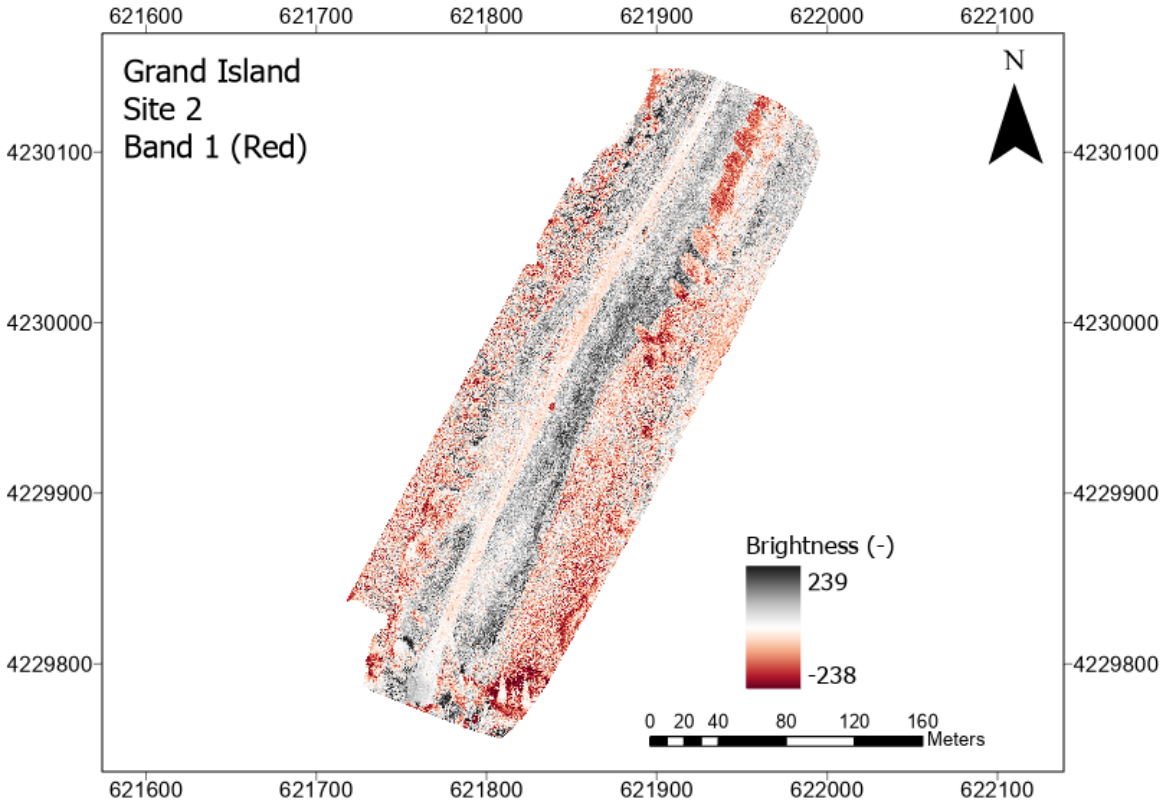


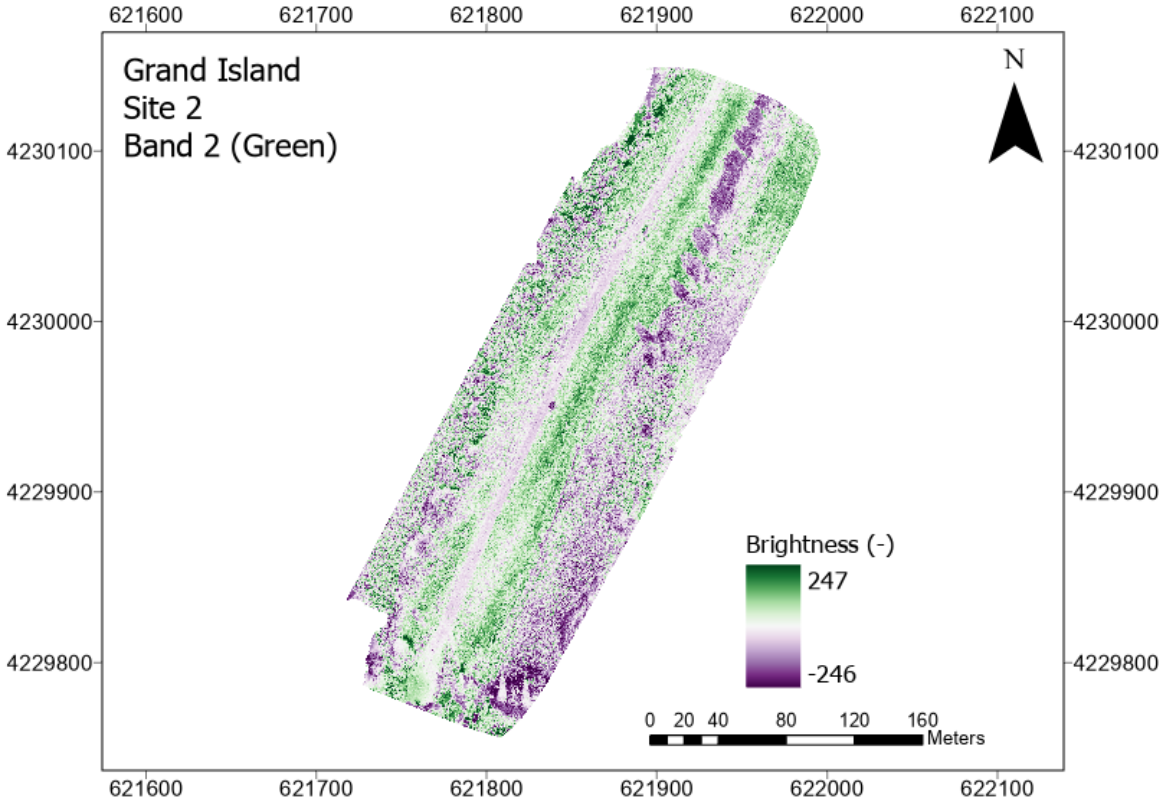

Grand Island
Site 2
Dead_Green
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Other
Dead to Green
0 20 40 80 120 160
Meters

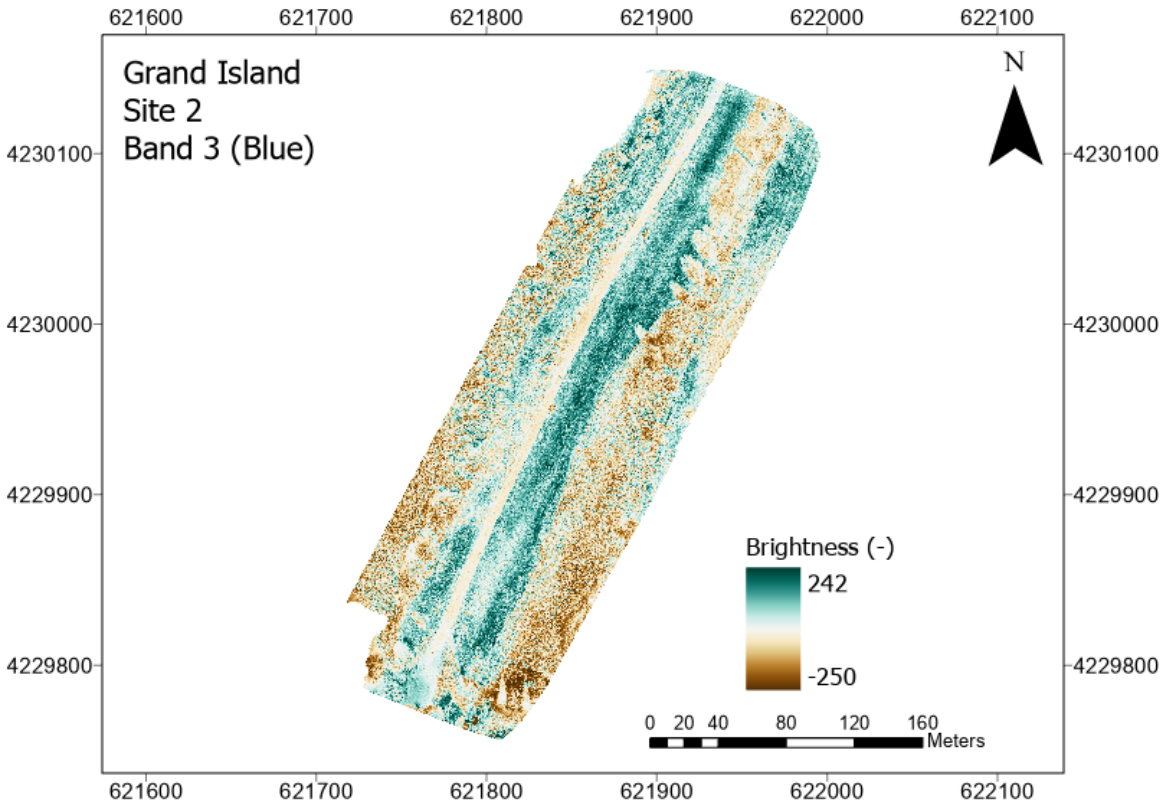

Grand Island
Site 2
Areas of Thermal Change
Thermal Change
N
621600
621650
621700
621750
621800
621850
621900
621950
622000
622050
622100
4230150
4230100
4230050
4230000
4229950
4229900
4229850
4229800
4229750
0 15 30 60 90 120
Meters

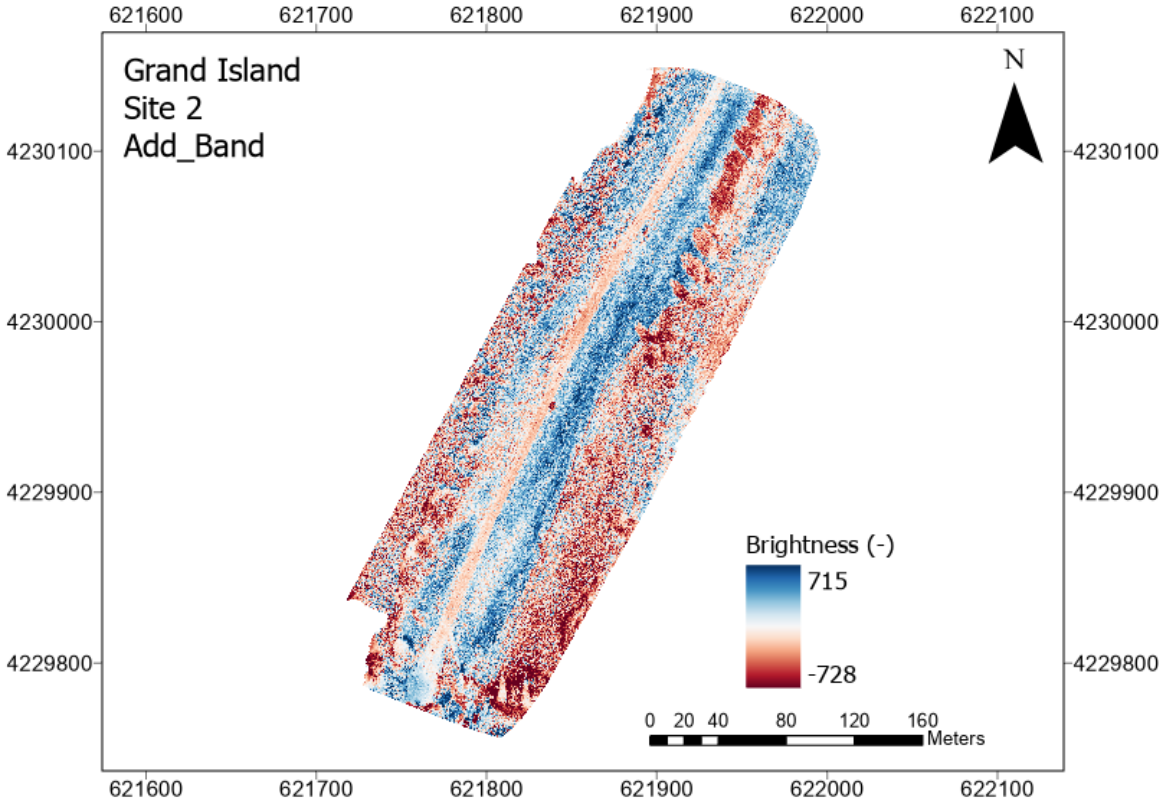

Grand Island
Site 2
NRCT
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
NRCT (-)
0.62
-0.70
0 20 40 80 120 160
Meters

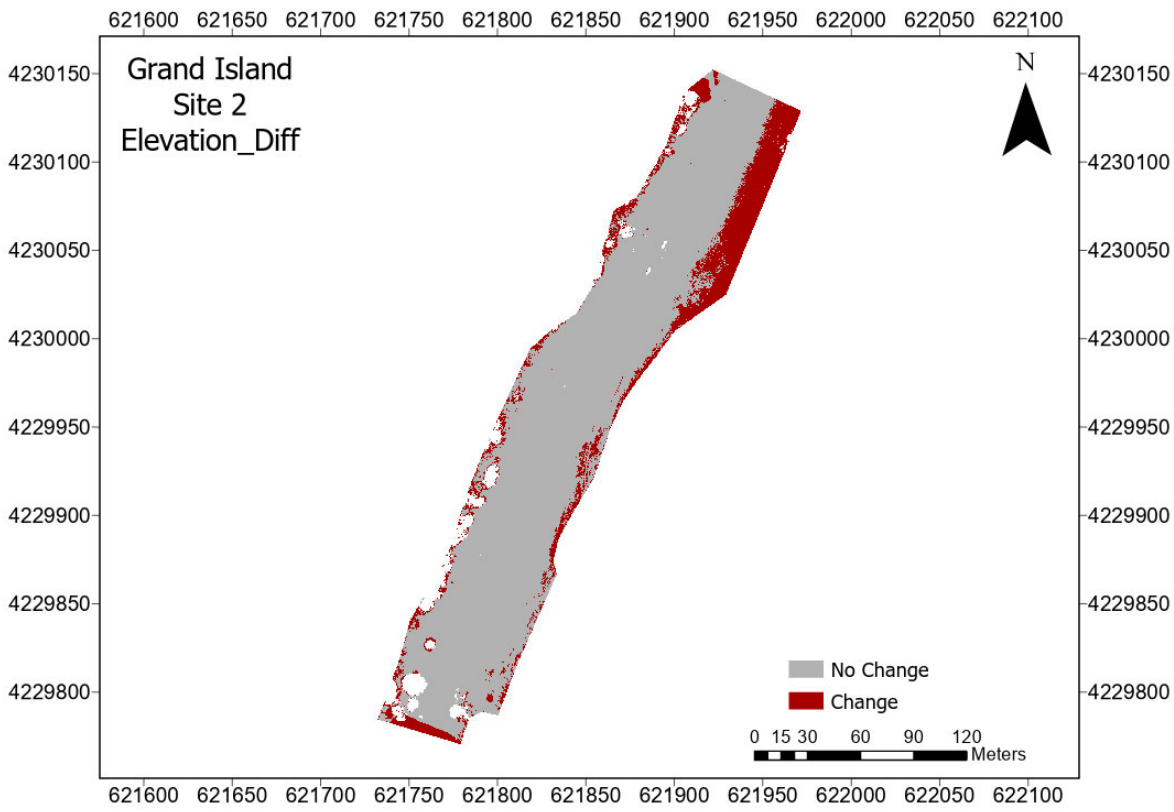

Grand Island
Site 2
Norm_GR
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
NGRDI (-)
1.01
-0.81
0 20 40 80 120 160
Meters

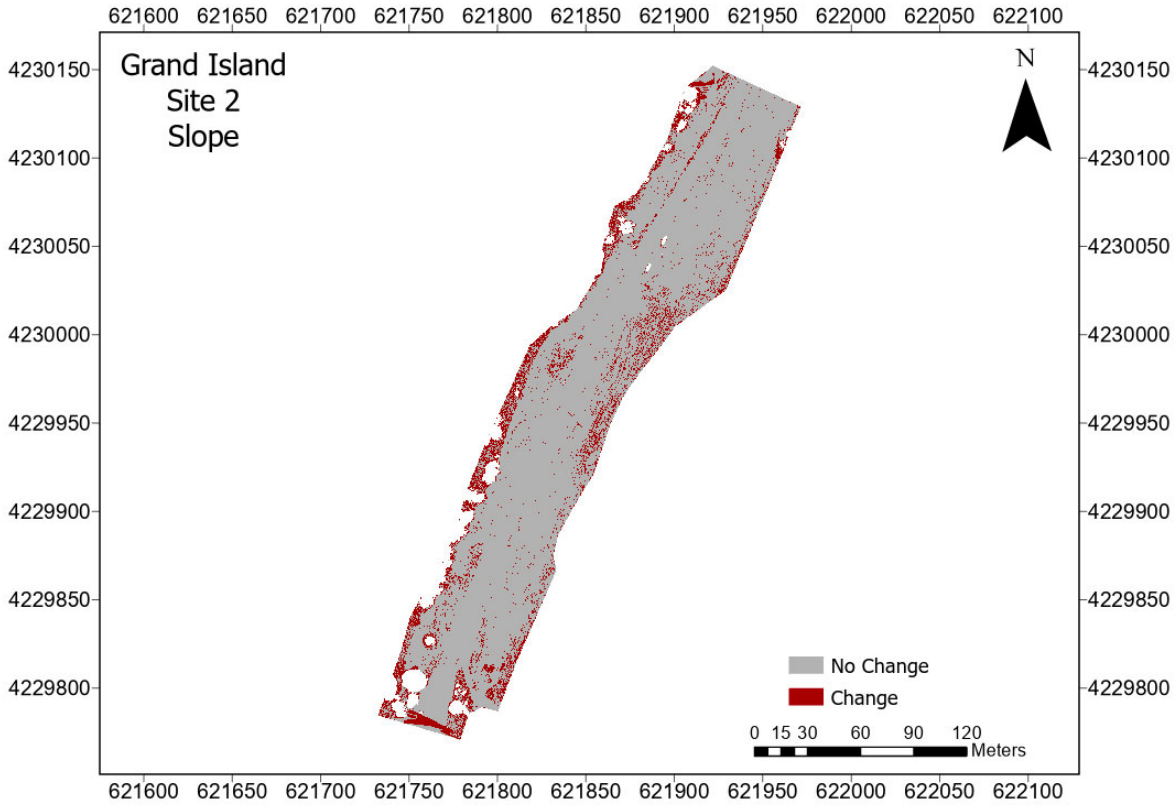

Grand Island
Site 2
Band 1 (Red)
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Brightness (-)
239
-238
0 20 40 80 120 160
Meters

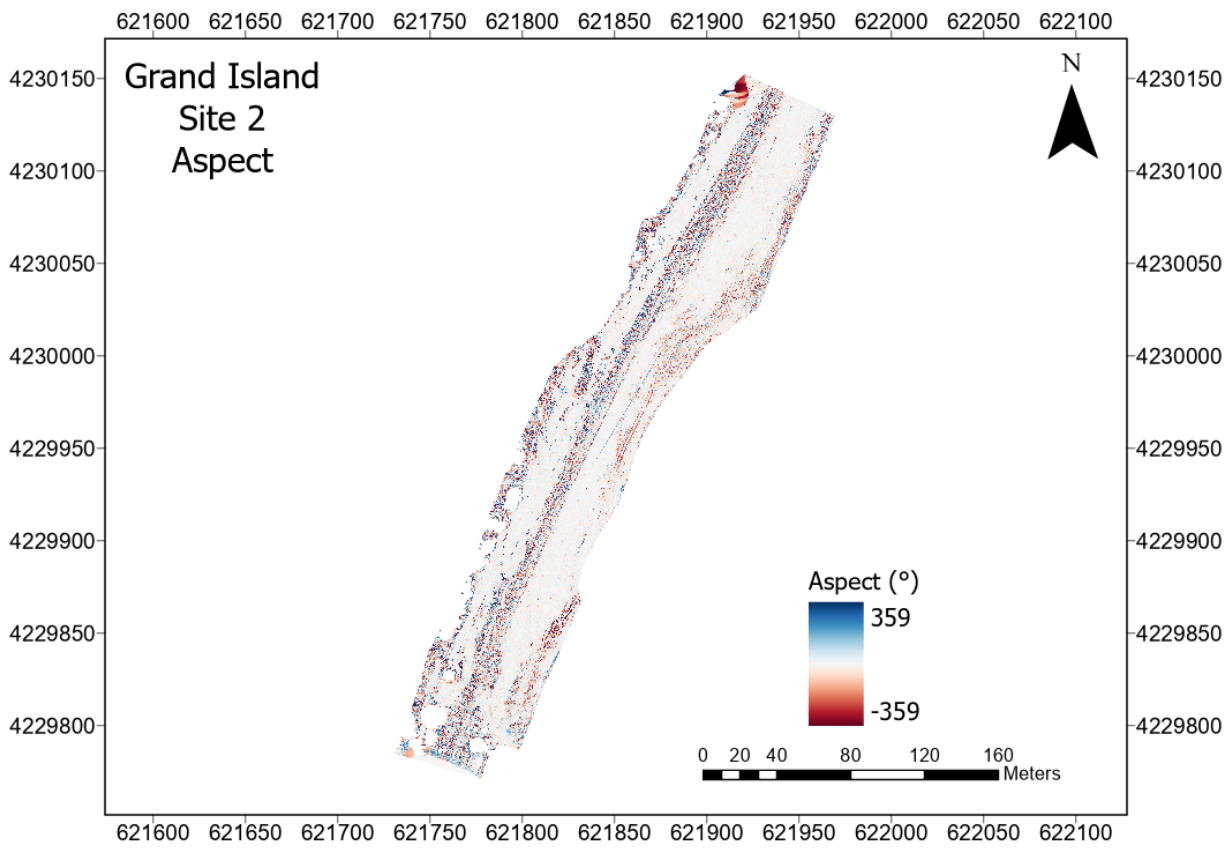

Grand Island
Site 2
Band 2 (Green)
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Brightness (-)
247
-246
0 20 40 80 120 160
Meters


Grand Island
Site 2
Band 3 (Blue)
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Brightness (-)
242
-250
0 20 40 80 120 160
Meters

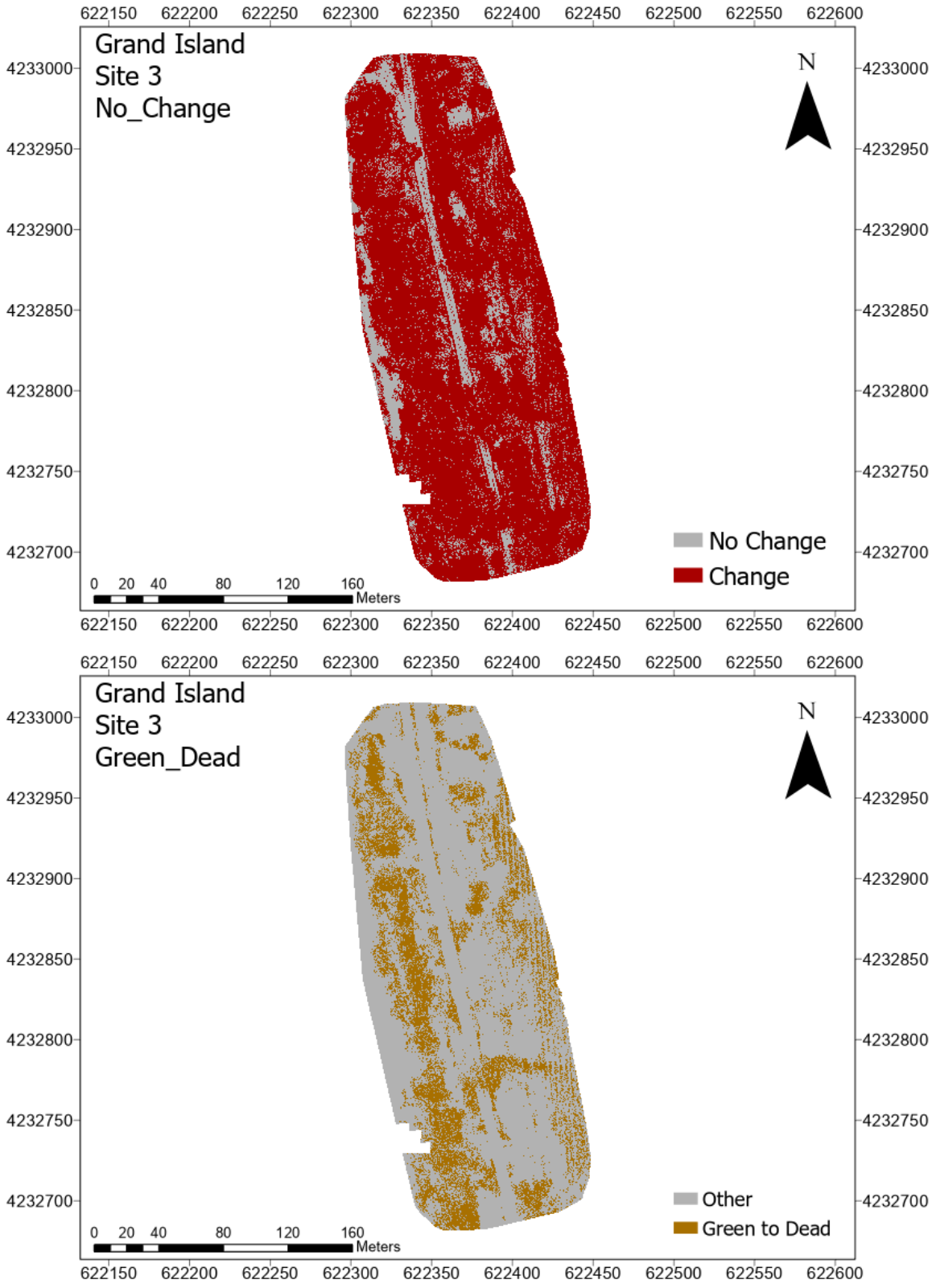

Grand Island
Site 2
Add_Band
N
621600
621700
621800
621900
622000
622100
4230100
4230000
4229900
4229800
Brightness (-)
715
-728
0 20 40 80 120 160
Meters

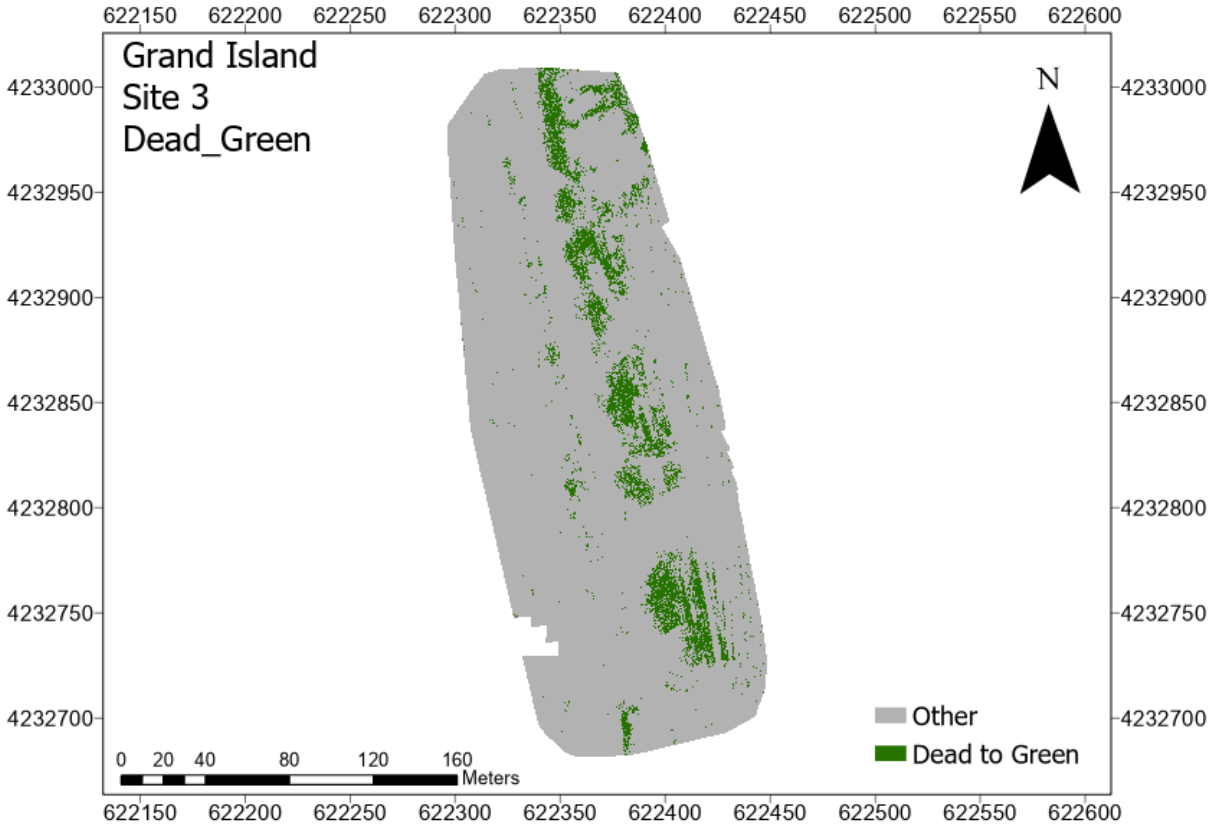

Grand Island
Site 2
Elevation_Diff
N
621600 621650 621700 621750 621800 621850 621900 621950 622000 622050 622100
4230150 4230100 4230050 4230000 4229950 4229900 4229850 4229800
No Change
Change
0 15 30 60 90 120
Meters

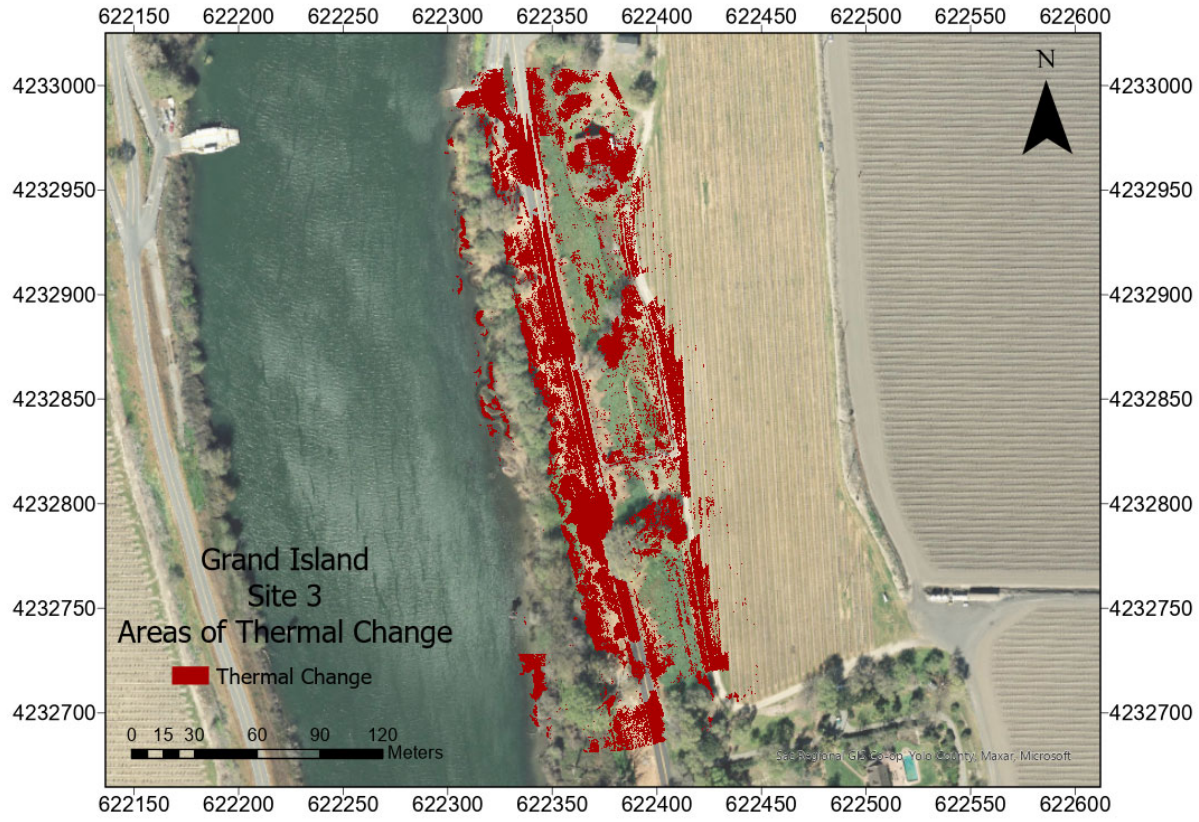

Grand Island
Site 2
Slope
N
621600 621650 621700 621750 621800 621850 621900 621950 622000 622050 622100
4230150 4230100 4230050 4230000 4229950 4229900 4229850 4229800
No Change
Change
0 15 30 60 90 120
Meters

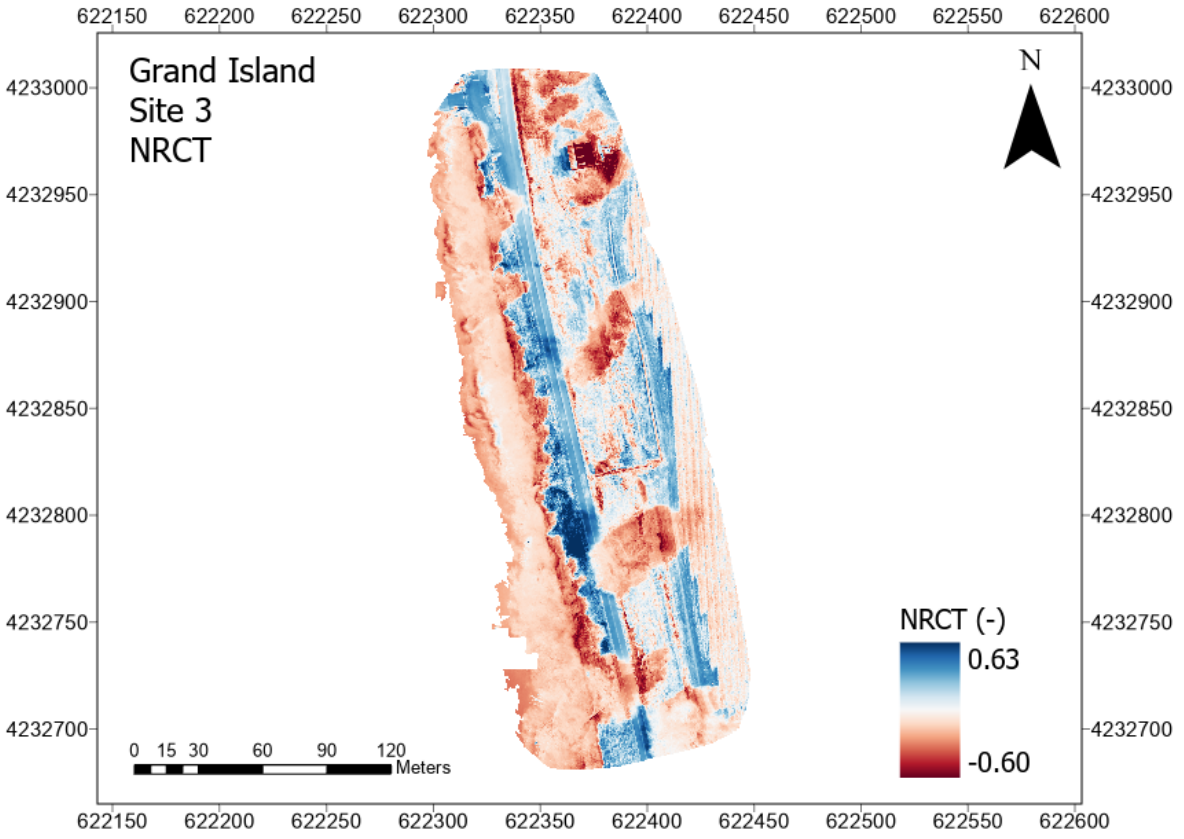

Grand Island
Site 2
Aspect
N
Aspect (°)
359
-359
0 20 40 80 120 160
Meters
621600 621650 621700 621750 621800 621850 621900 621950 622000 622050 622100
4230150 4230100 4230050 4230000 4229950 4229900 4229850 4229800

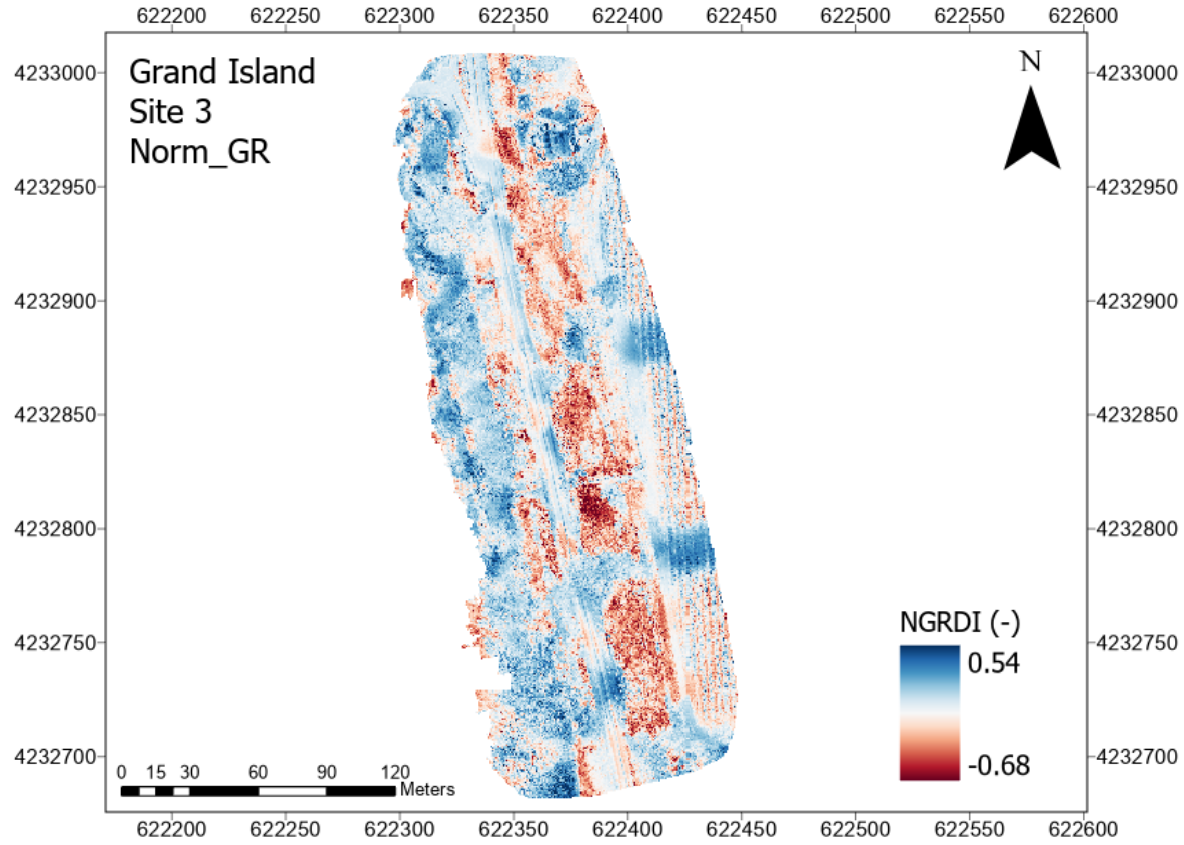
Surface Change Detection Results – Site 3

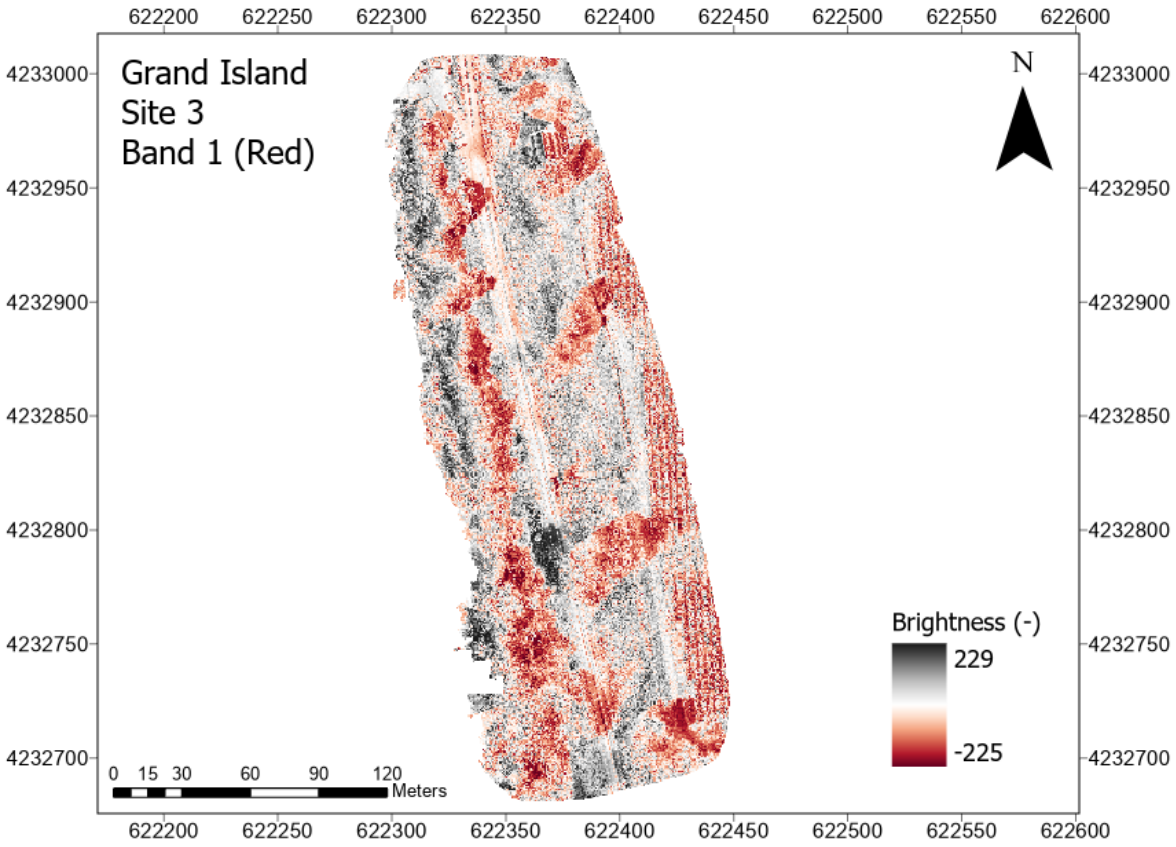


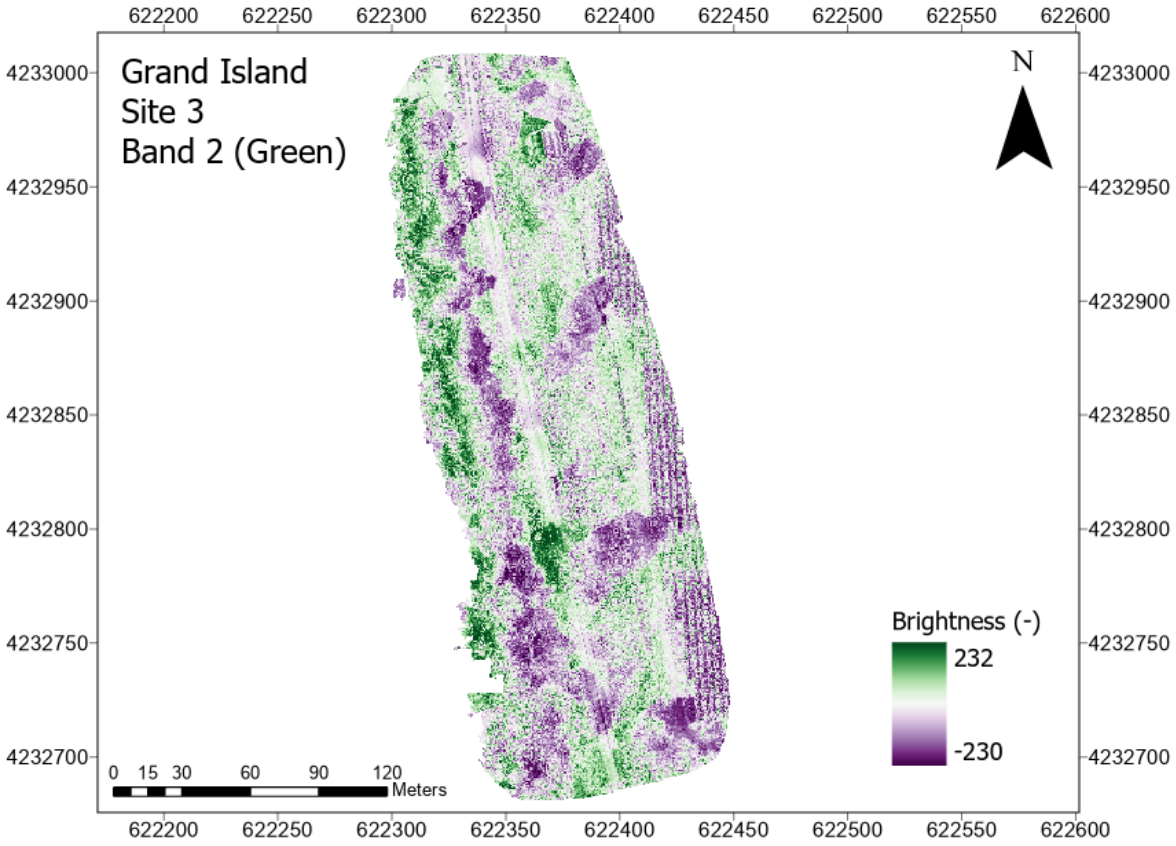

Grand Island
Site 3
Dead_Green
N
622150 622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700
0 20 40 80 120 160
Meters
Other
Dead to Green

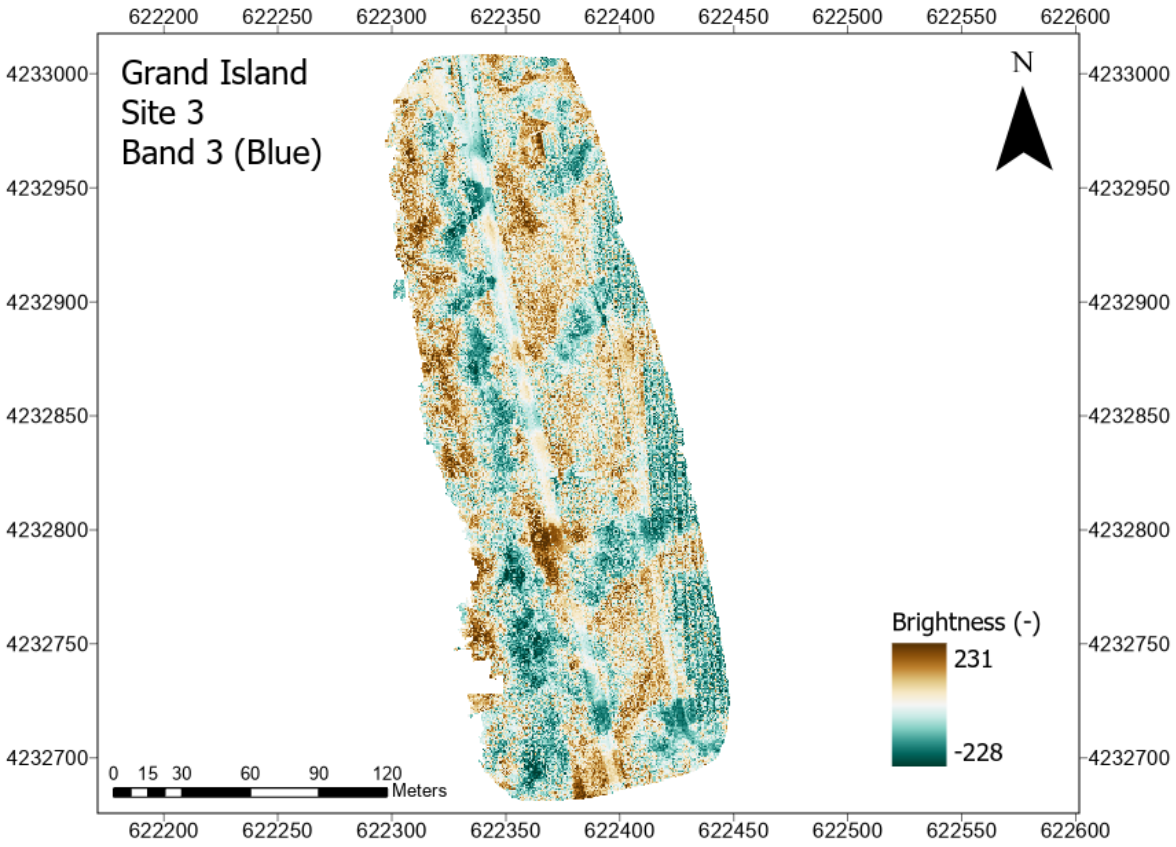

Grand Island
Site 3
Areas of Thermal Change
Thermal Change
N
622150 622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700
0 15 30 60 90 120
Meters

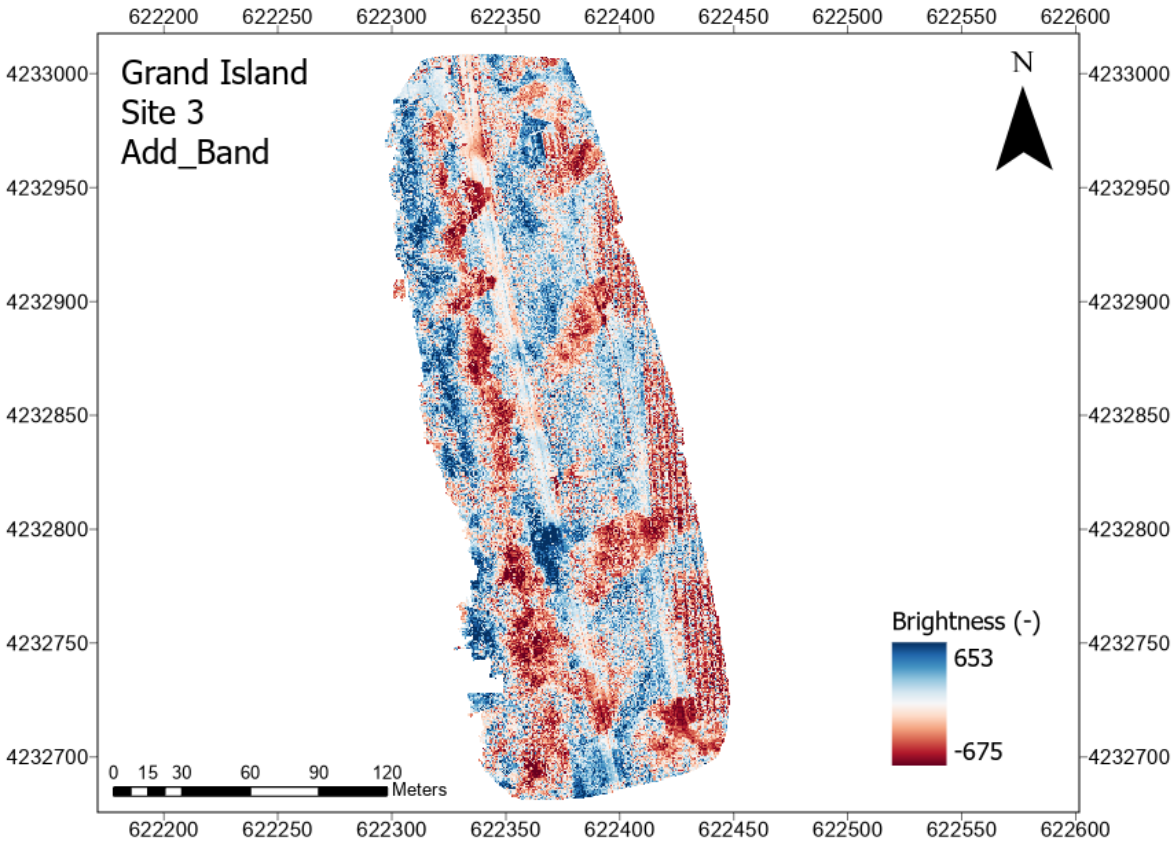

Grand Island
Site 3
NRCT
N
NRCT (-)
0.63
-0.60
0 15 30 60 90 120
Meters
622150 622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700

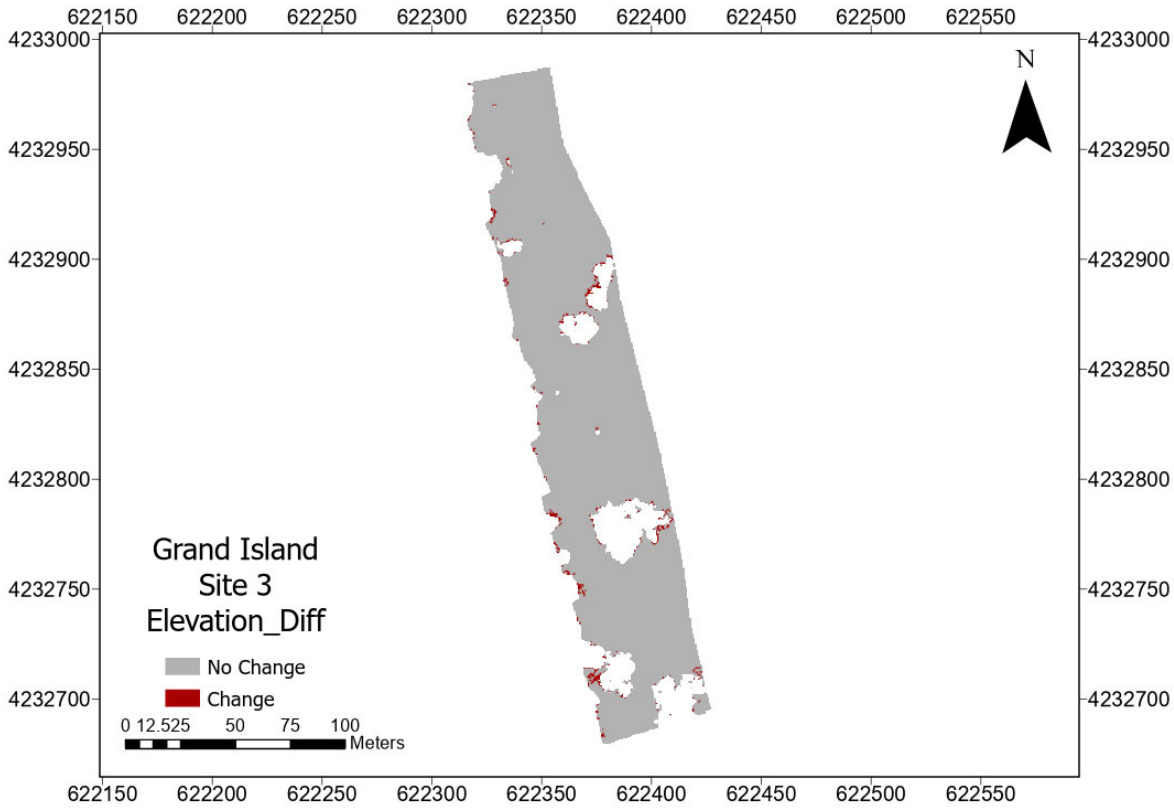

Grand Island
Site 3
Norm_GR
N
NGRDI (-)
0.54
-0.68
0 15 30 60 90 120
Meters
622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700

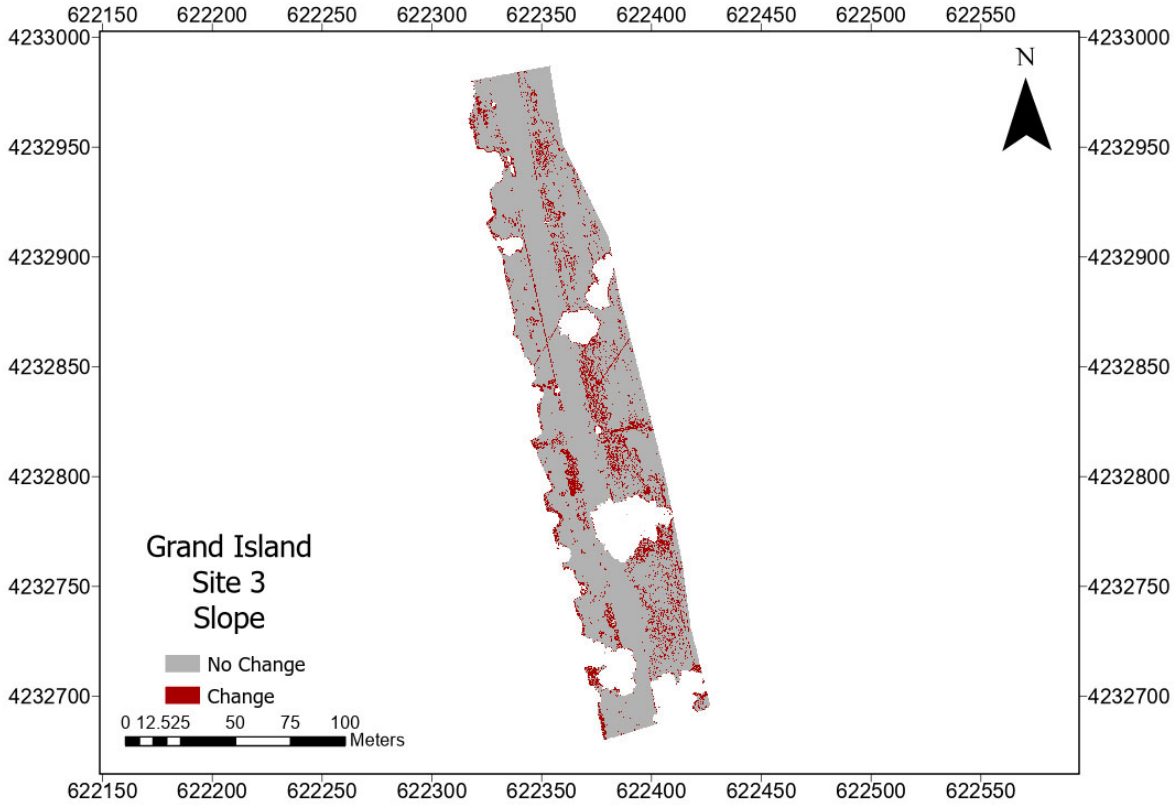

Grand Island
Site 3
Band 1 (Red)
N
Brightness (-)
229
-225
0 15 30 60 90 120
Meters
622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700

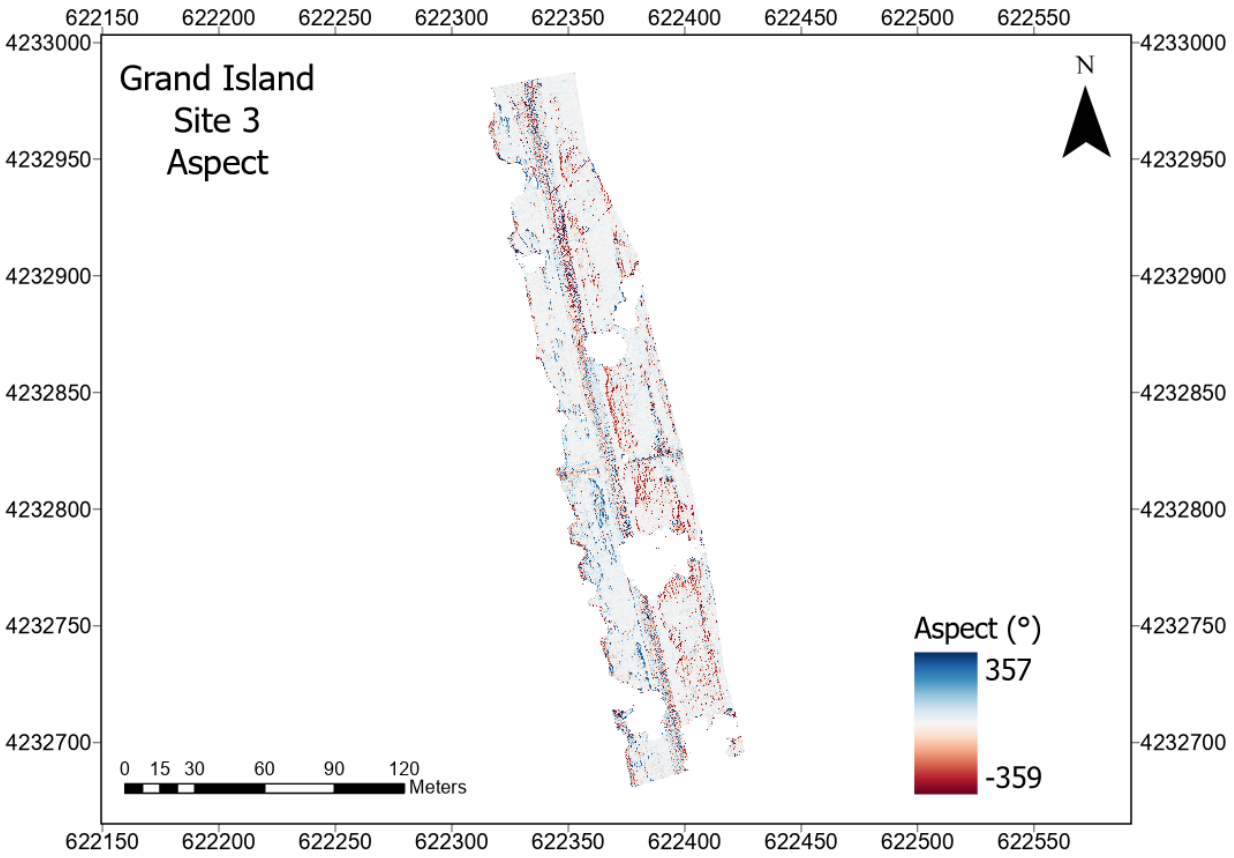

Grand Island
Site 3
Band 2 (Green)
N
Brightness (-)
232
-230
0 15 30 60 90 120
Meters
622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700


Grand Island
Site 3
Band 3 (Blue)
N
Brightness (-)
231
-228
0 15 30 60 90 120
Meters
622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700


Grand Island
Site 3
Add_Band
N
Brightness (-)
653
-675
0 15 30 60 90 120
Meters
622200 622250 622300 622350 622400 622450 622500 622550 622600
4233000 4232950 4232900 4232850 4232800 4232750 4232700


622150
622200
622250
622300
622350
622400
622450
622500
622550
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island
Site 3
Elevation_Diff
No Change
Change
0 12.525 50 75 100
Meters


622150
622200
622250
622300
622350
622400
622450
622500
622550
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island
Site 3
Slope
No Change
Change
0 12.525 50 75 100
Meters


Grand Island
Site 3
Aspect
N
622150
622200
622250
622300
622350
622400
622450
622500
622550
4233000
4232950
4232900
4232850
4232800
4232750
4232700
0 15 30 60 90 120
Meters
Aspect (°)
357
-359

Appendix F – Machine Learning Model Parameters

Table F.1 - Anomaly Detection Model Parameters

			IF	Contamination	LOF			OCSVM
ID	Fit	Predict	n_estimators	Vs and R	n_neighbors	metric	p	gamma
1	fc_site2	fc_site2	50	0.1375	20	minkowski	2	scale
2	fc_site3	fc_site3	50	0.209	20	minkowski	2	scale
3	fc3_site1	fc3_site1	50	0.148	20	minkowski	2	scale
4	fc5_site1	fc5_site1	50	0.1207	20	minkowski	2	scale
5	fdt_site2	fdt_site2	50	0.16	20	minkowski	2	scale
6	fwt_site1	fwt_site1	50	0.155	20	minkowski	2	scale
7	fwt_site2	fwt_site2	50	0.2017	20	minkowski	2	scale
8	fwt_site3	fwt_site3	50	0.165	20	minkowski	2	scale
9	sc_site2	sc_site2	50	0.1517	20	minkowski	2	scale
10	sc_site3	sc_site3	50	0.16625	20	minkowski	2	scale
11	sc3_site1	sc3_site1	50	0.13375	20	minkowski	2	scale
12	sc5_site1	sc5_site1	50	0.149	20	minkowski	2	scale
13	sdt_site2	sdt_site2	50	0.1767	20	minkowski	2	scale
14	swt_site1	swt_site1	50	0.1367	20	minkowski	2	scale
15	swt_site2	swt_site2	50	0.1683	20	minkowski	2	scale
16	swt_site3	swt_site3	50	0.12	20	minkowski	2	scale
17	all_toe	all_toe	50	0.1583	20	minkowski	2	scale
18	all_toe	fdt_site2	50	0.1933	20	minkowski	2	scale
19	all_toe	fwt_site1	50	0.16	20	minkowski	2	scale
20	all_toe	fwt_site2	50	0.1317	20	minkowski	2	scale
21	all_toe	fwt_site3	50	0.1	20	minkowski	2	scale
22	all_toe	sdt_site2	50	0.13	20	minkowski	2	scale
23	all_toe	swt_site1	50	0.1433	20	minkowski	2	scale
24	all_toe	swt_site2	50	0.1667	20	minkowski	2	scale
25	all_toe	swt_site3	50	0.1956	20	minkowski	2	scale
26	site2	site2	50	0.1371	20	minkowski	2	scale
27	site2	fc_site2	50	0.18875	20	minkowski	2	scale
28	site2	fdt_site2	50	0.1517	20	minkowski	2	scale
29	site2	fwt_site2	50	0.1333	20	minkowski	2	scale
30	site2	sc_site2	50	0.2	20	minkowski	2	scale
31	site2	sdt_site2	50	0.1383	20	minkowski	2	scale
32	site2	swt_site2	50	0.2117	20	minkowski	2	scale
33	all crest	all crest	50	0.1351	20	minkowski	2	scale
34	all crest	fc_site2	50	0.1625	20	minkowski	2	scale
35	all crest	fc_site3	50	0.154	20	minkowski	2	scale
36	all crest	fc3_site1	50	0.195	20	minkowski	2	scale
37	all crest	fc5_site1	50	0.1633	20	minkowski	2	scale
38	all crest	sc_site2	50	0.1817	20	minkowski	2	scale
39	all crest	sc_site3	50	0.2338	20	minkowski	2	scale
40	all crest	sc3_site1	50	0.1363	20	minkowski	2	scale
41	all crest	sc5_site1	50	0.131	20	minkowski	2	scale
42	all_dry	all_dry	50	0.1682	20	minkowski	2	scale
43	all_dry	sc_site2	50	0.1867	20	minkowski	2	scale
44	all_dry	sc_site3	50	0.165	20	minkowski	2	scale
45	all_dry	sc3_site1	50	0.13	20	minkowski	2	scale
46	all_dry	sc5_site1	50	0.142	20	minkowski	2	scale
47	all_dry	sdt_site2	50	0.1567	20	minkowski	2	scale

48	all_dry	swt_site1	50	0.145	20	minkowski	2	scale
49	all_dry	swt_site2	50	0.1833	20	minkowski	2	scale
50	all_dry	swt_site3	50	0.2133	20	minkowski	2	scale
51	all_wet	all_wet	50	0.1495	20	minkowski	2	scale
52	all_wet	fc_site2	50	0.1625	20	minkowski	2	scale
53	all_wet	fc_site3	50	0.147	20	minkowski	2	scale
54	all_wet	fc3_site1	50	0.131	20	minkowski	2	scale
55	all_wet	fc5_site1	50	0.1847	20	minkowski	2	scale
56	all_wet	fdt_site2	50	0.185	20	minkowski	2	scale
57	all_wet	fwt_site1	50	0.1517	20	minkowski	2	scale
58	all_wet	fwt_site2	50	0.1083	20	minkowski	2	scale
59	all_wet	fwt_site3	50	0.1275	20	minkowski	2	scale
60	site1_dry	site1_dry	50	0.1079	20	minkowski	2	scale
61	site1_dry	sc3_site1	50	0.1263	20	minkowski	2	scale
62	site1_dry	sc5_site1	50	0.112	20	minkowski	2	scale
63	site1_dry	swt_site1	50	0.14	20	minkowski	2	scale
64	site1_wet	site1_wet	50	0.1555	20	minkowski	2	scale
65	site1_wet	fc3_site1	50	0.175	20	minkowski	2	scale
66	site1_wet	fc5_site1	50	0.1473	20	minkowski	2	scale
67	site1_wet	fwt_site1	50	0.155	20	minkowski	2	scale
68	site2_dry	site2_dry	50	0.1572	20	minkowski	2	scale
69	site2_dry	sc_site2	50	0.165	20	minkowski	2	scale
70	site2_dry	sdt_site2	50	0.1883	20	minkowski	2	scale
71	site2_dry	swt_site2	50	0.2167	20	minkowski	2	scale
72	site2_wet	site2_wet	50	0.1805	20	minkowski	2	scale
73	site2_wet	fc_site2	50	0.155	20	minkowski	2	scale
74	site2_wet	fdt_site2	50	0.2217	20	minkowski	2	scale
75	site2_wet	fwt_site2	50	0.0967	20	minkowski	2	scale
76	site3_dry	site3_dry	50	0.1639	20	minkowski	2	scale
77	site3_dry	sc_site3	50	0.1588	20	minkowski	2	scale
78	site3_dry	swt_site3	50	0.1378	20	minkowski	2	scale
79	site3_wet	site3_wet	50	0.1542	20	minkowski	2	scale
80	site3_wet	fc_site3	50	0.17	20	minkowski	2	scale
81	site3_wet	fwt_site3	50	0.12	20	minkowski	2	scale
82	site1	site1	50	0.1298	20	minkowski	2	scale
83	site1	fc3_site1	50	0.138	20	minkowski	2	scale
84	site1	fc5_site1	50	0.146	20	minkowski	2	scale
85	site1	fwt_site1	50	0.105	20	minkowski	2	scale
86	site1	sc3_site1	50	0.10125	20	minkowski	2	scale
87	site1	sc5_site1	50	0.091	20	minkowski	2	scale
88	site1	swt_site1	50	0.1267	20	minkowski	2	scale
89	site3	site3	50	0.1542	20	minkowski	2	scale
90	site3	fc_site3	50	0.175	20	minkowski	2	scale
91	site3	fwt_site3	50	0.115	20	minkowski	2	scale
92	site3	sc_site3	50	0.2063	20	minkowski	2	scale
93	site3	swt_site3	50	0.2756	20	minkowski	2	scale

Table F.2 - K-Means and Hierarchical Clustering Model Parameters

		K-Means and Hierarchical					
	Survey	Elbow - k	Silhouette - k	Gap - k	Diff - k	Gap* - k	Average k
1	fc_site2	5	3	15	4	3	6
2	fc_site3	7	3	20	3	3	7
3	fc3_site1	6	18	17	3	2	9
4	fc5_site1	7	2	18	3	2	6
5	fdt_site2	6	2	16	3	2	6
6	fwt_site1	6	3	3	3	3	4
7	fwt_site2	5	2	19	2	4	6
8	fwt_site3	6	11	17	5	3	8
9	sc_site2	7	7	10	3	3	6
10	sc_site3	6	2	16	2	2	6
11	sc3_site1	6	16	8	2	2	7
12	sc5_site1	6	6	6	6	3	5
13	sdt_site2	7	2	2	2	2	3
14	swt_site1	6	2	19	4	3	7
15	swt_site2	6	2	9	3	2	4
16	swt_site3	5	6	10	3	2	5
17	all_crest	6	4	16	3	2	6
18	all_toe	7	2	4	4	2	4
19	all_dry	6	4	7	3	2	4
20	all_wet	6	3	3	3	2	3
21	site1_dry	6	6	16	2	2	6
22	site1_wet	5	3	3	3	2	3
23	site2_dry	5	3	8	6	3	5
24	site2_wet	6	3	14	3	3	6
25	site3_dry	6	13	18	2	2	8
26	site3_wet	5	18	18	3	2	9
27	site1	6	4	16	2	2	6
28	site2	6	3	18	3	2	6
29	site3	6	19	16	3	2	9

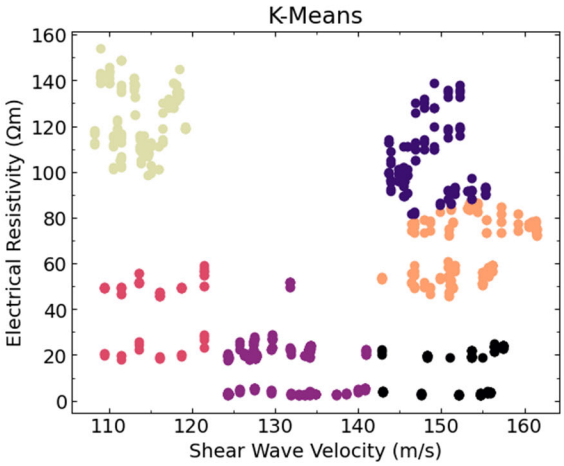
Table F.3 - DBSCAN Clustering Model Parameters

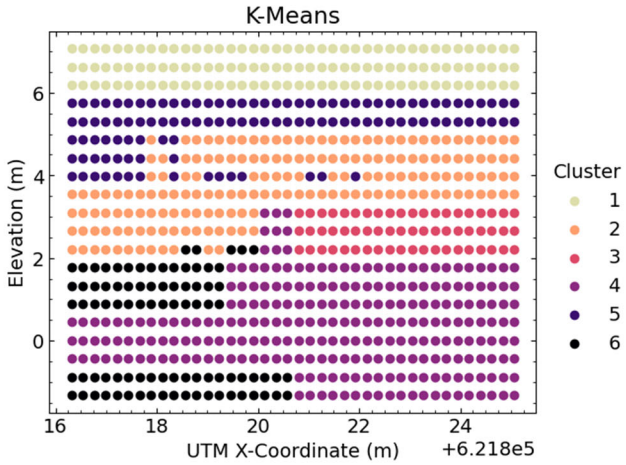
		DBSCAN					
	Survey	minPts	ε	Num clusters	Num outliers	Num points	% outliers
1	fc_site2	4	0.025	~	318	800	40
2	fc_site3	4	0.08	~	190	1000	19
3	fc3_site1	4	0.09	~	212	1000	21
4	fc5_site1	4	0.04	~	460	1500	31
5	fdt_site2	4	0.04	~	242	600	40
6	fwt_site1	4	0.09	~	219	600	37
7	fwt_site2	4	0.1	~	70	600	12
8	fwt_site3	4	0.1	~	81	400	20
9	sc_site2	4	0.1	~	168	600	28
10	sc_site3	4	0.03	~	232	800	29
11	sc3_site1	4	0.1	~	147	800	18
12	sc5_site1	4	0.1	~	58	1000	6
13	sdt_site2	4	0.16	~	30	600	5
14	swt_site1	4	0.07	~	146	600	24
15	swt_site2	4	0.08	~	193	600	32
16	swt_site3	4	0.08	~	61	225	27

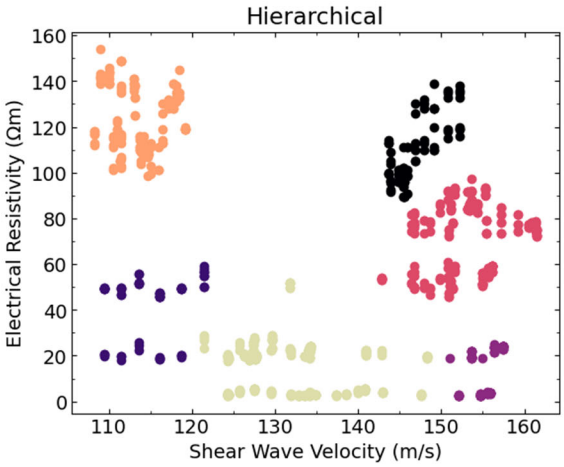
Appendix G – Clustering Profiles

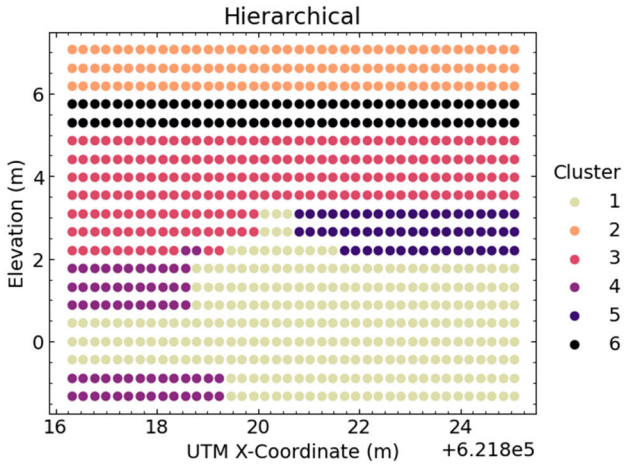
Fit: Individual Profiles

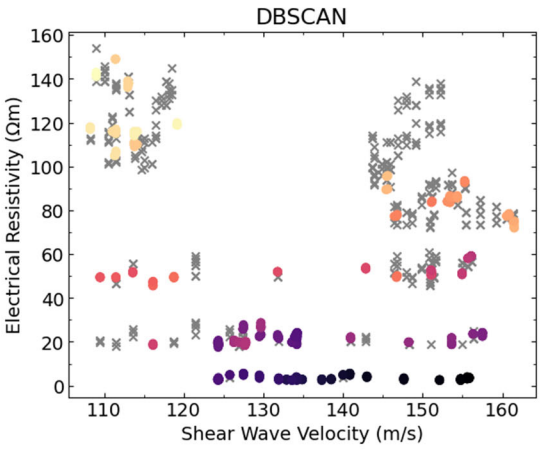
Profile: fc_site2

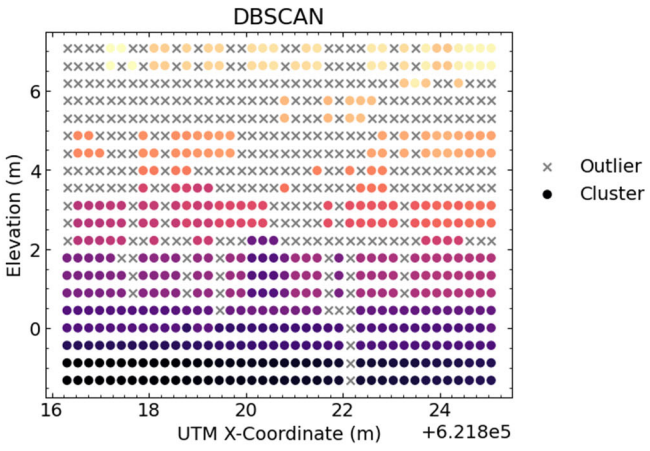






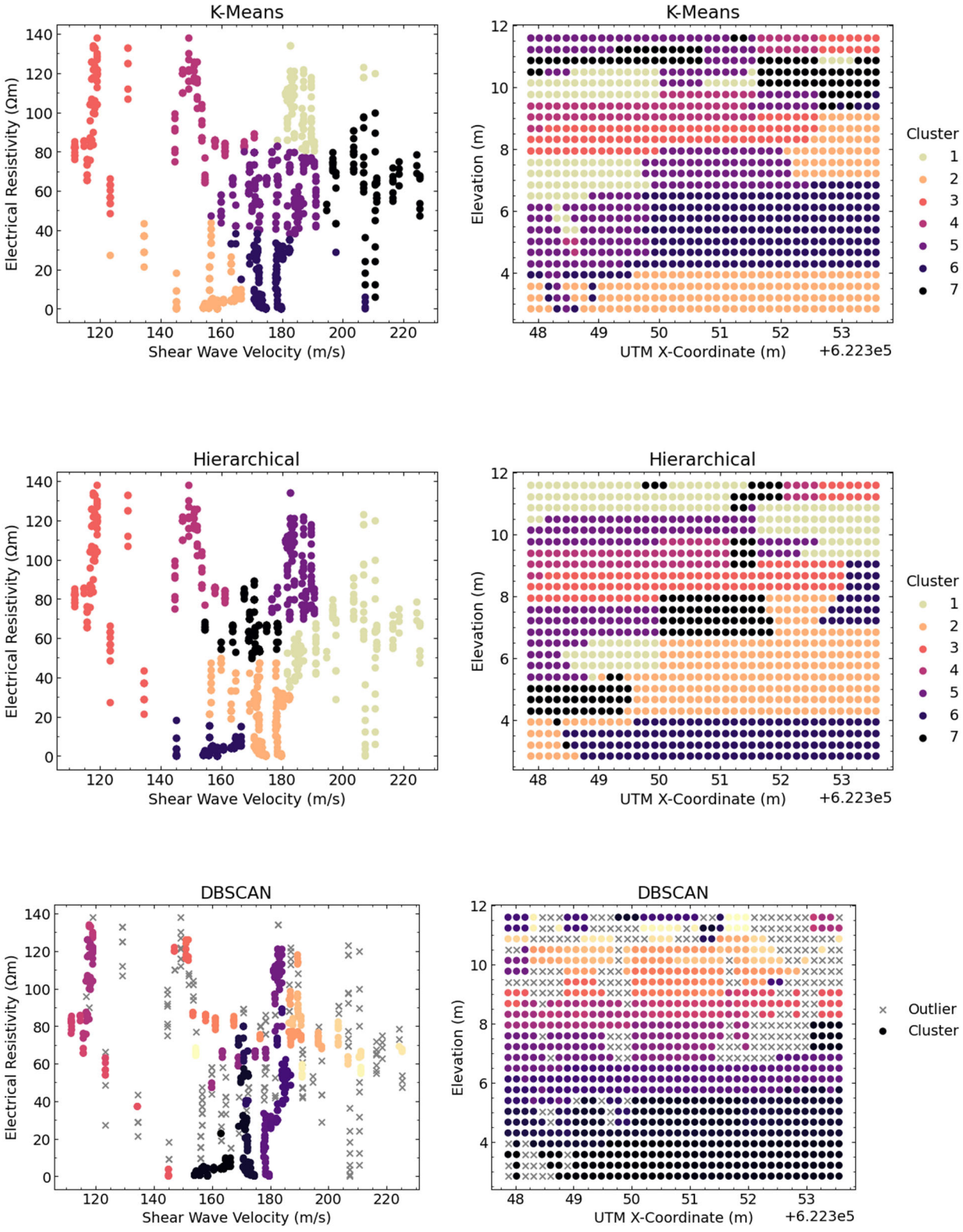






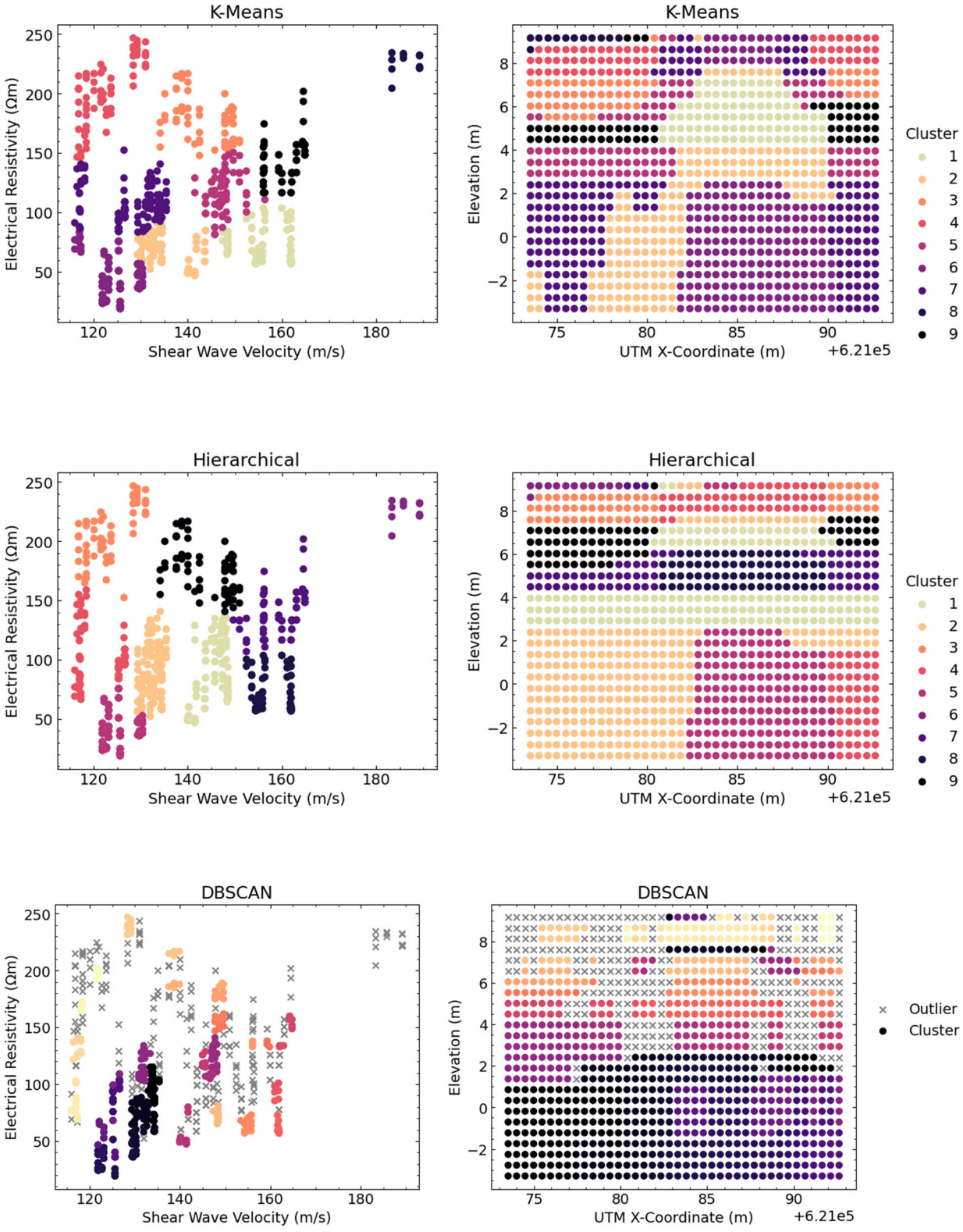
Fit: Individual Profiles

Profile: fc_site3



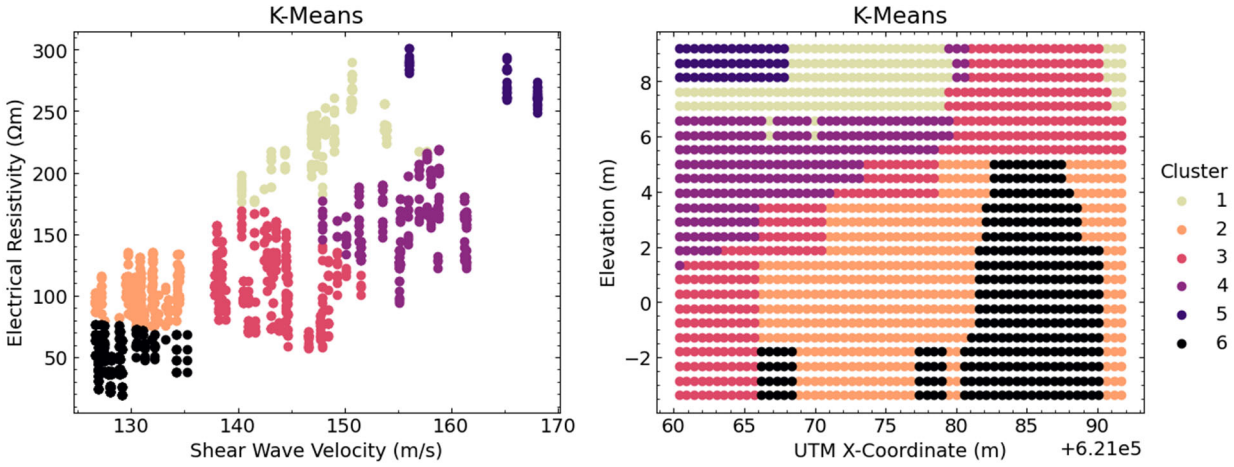
Fit: Individual Profiles

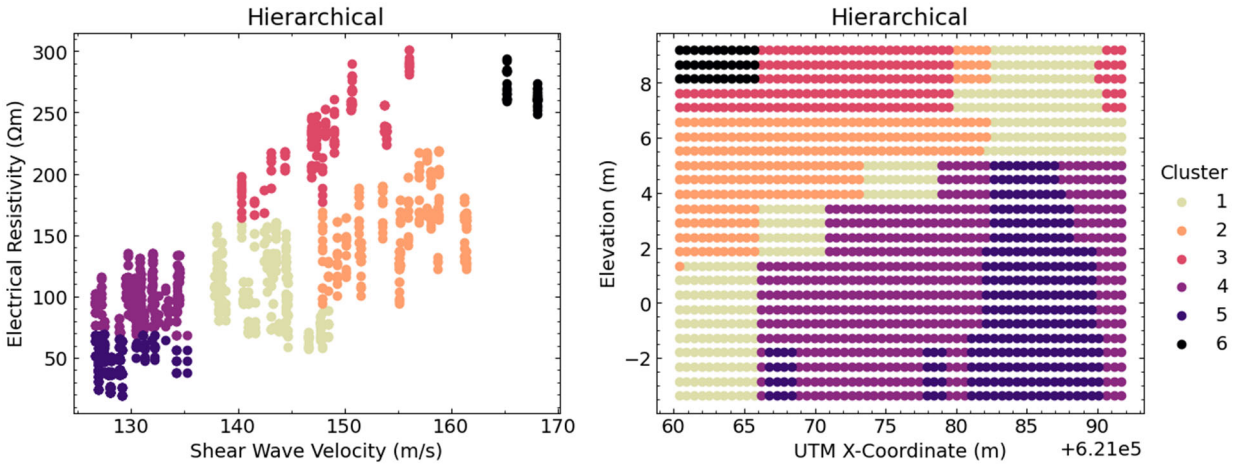
Profile: fc3_site1

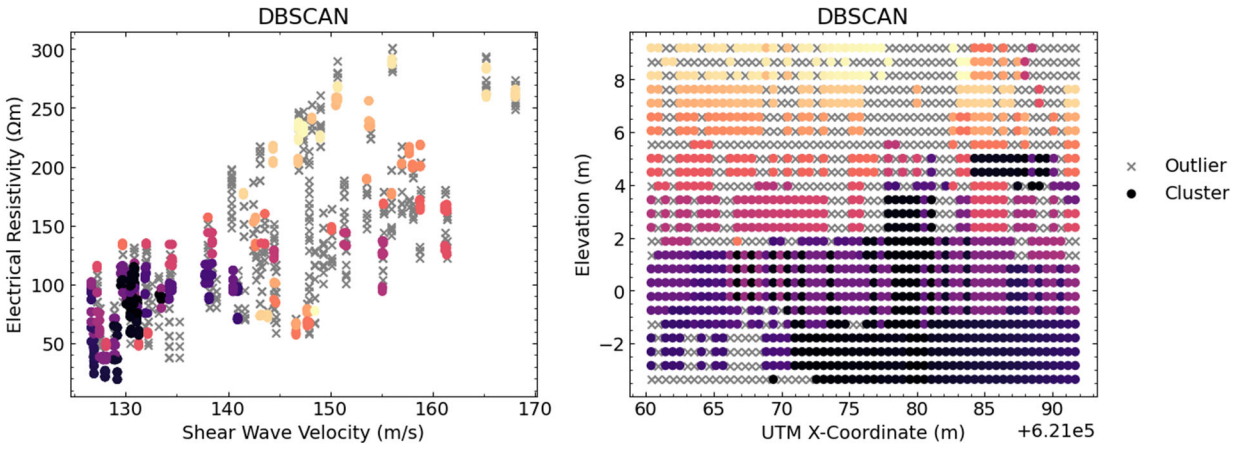


Fit: Individual Profiles

Profile: fc5_site1

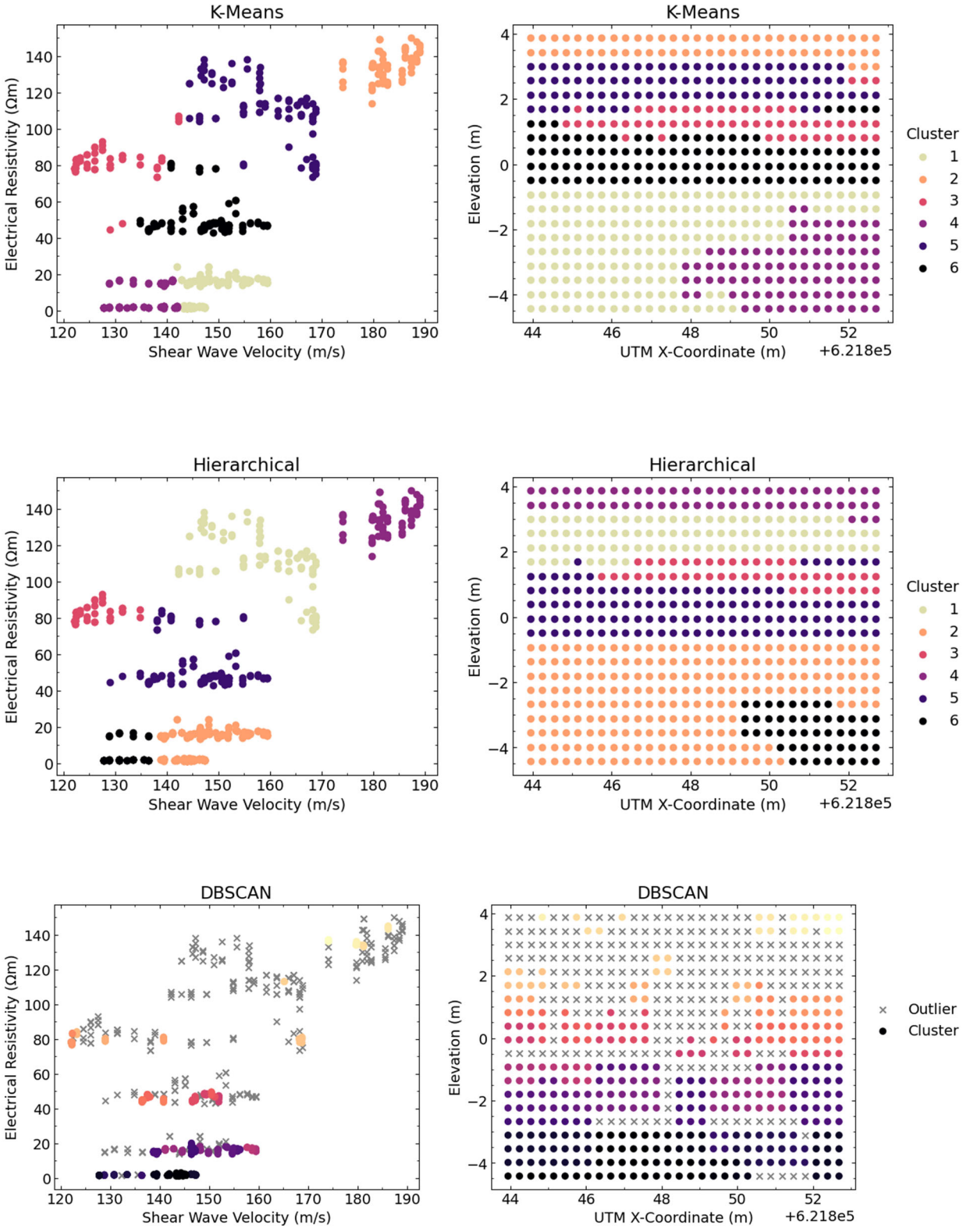






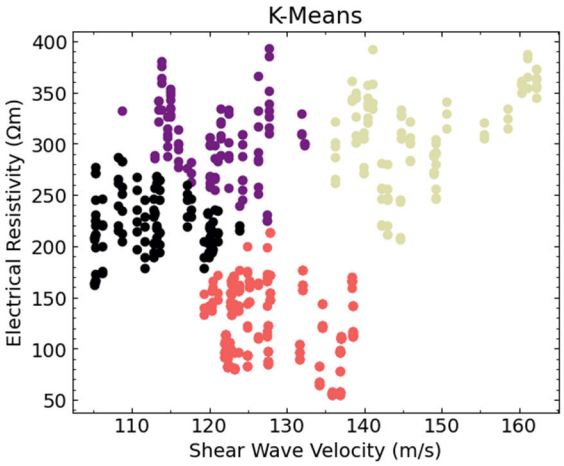
Fit: Individual Profiles

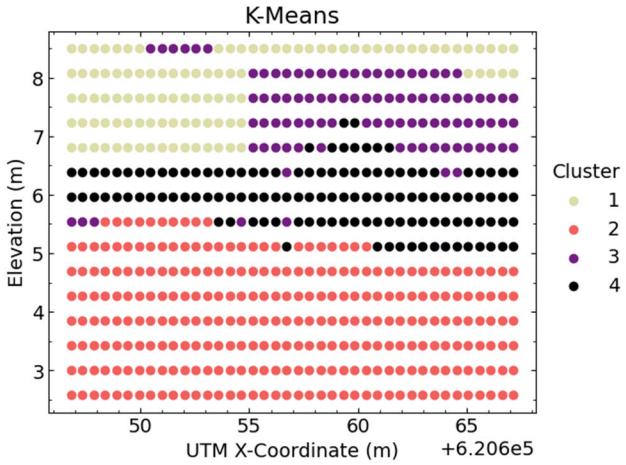
Profile: fdt_site2

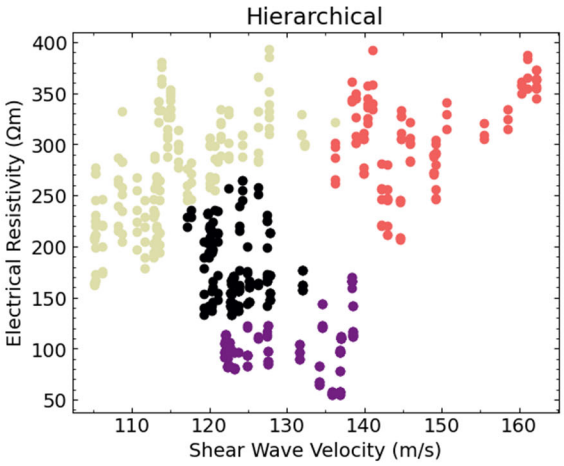


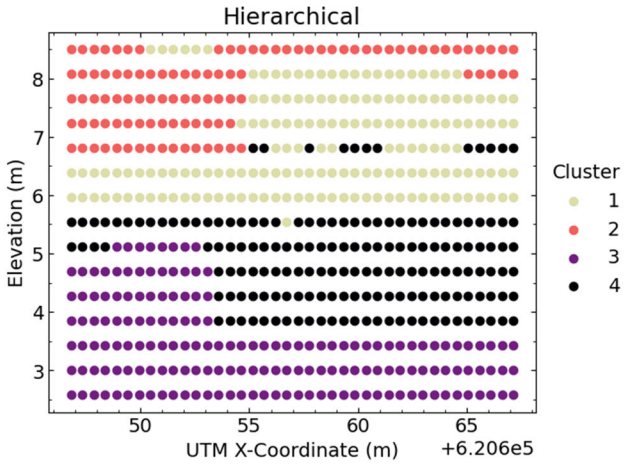
Fit: Individual Profiles

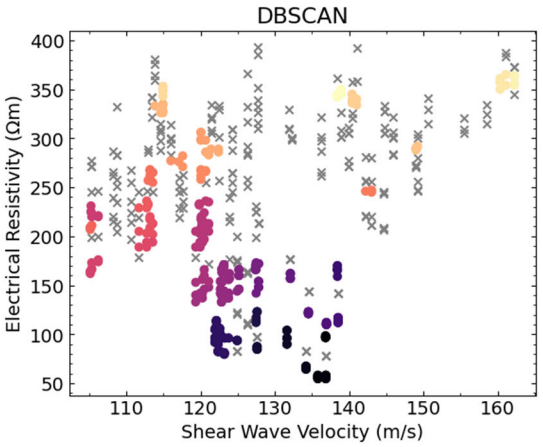
Profile: fwt_site1

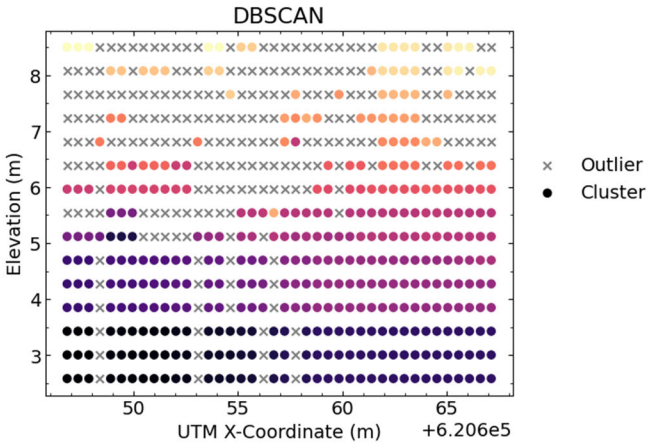






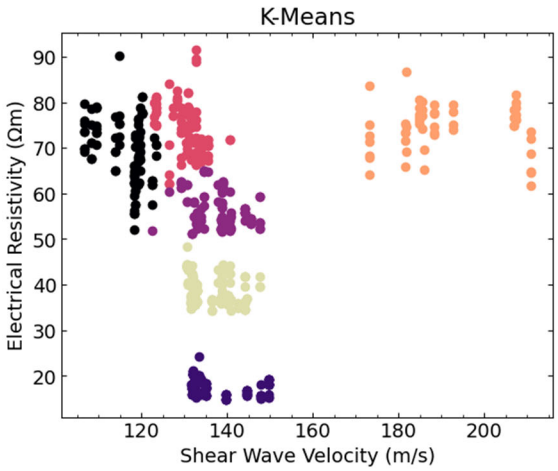


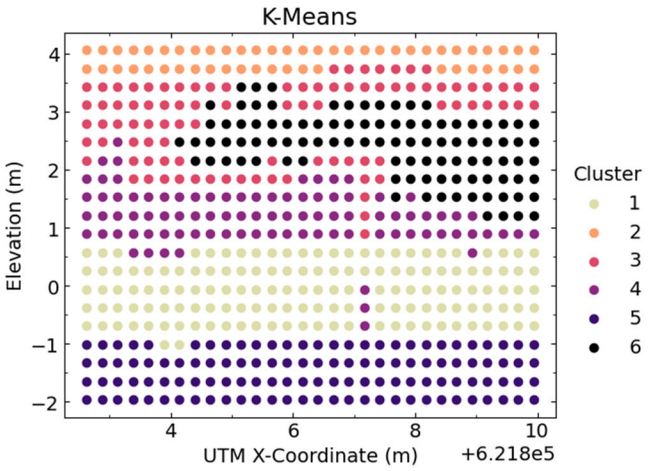


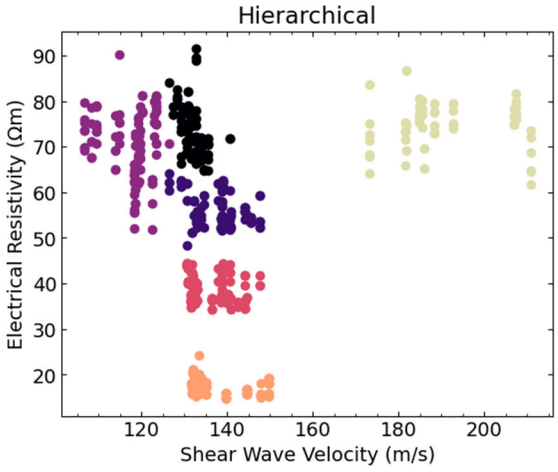


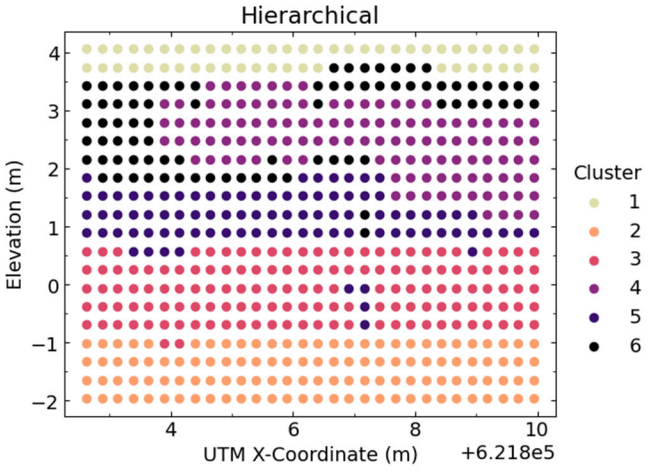
Fit: Individual Profiles

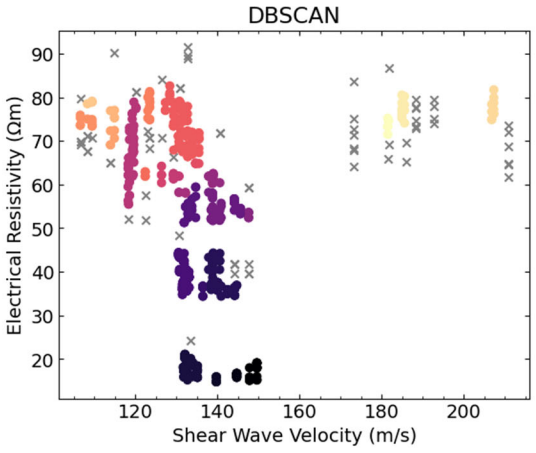
Profile: fwt_site2

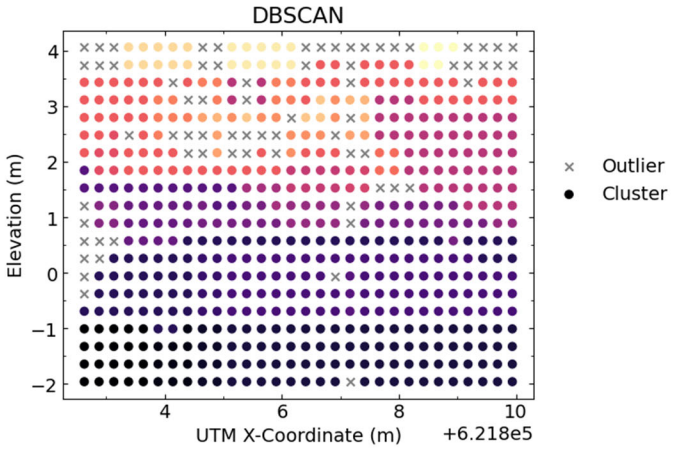






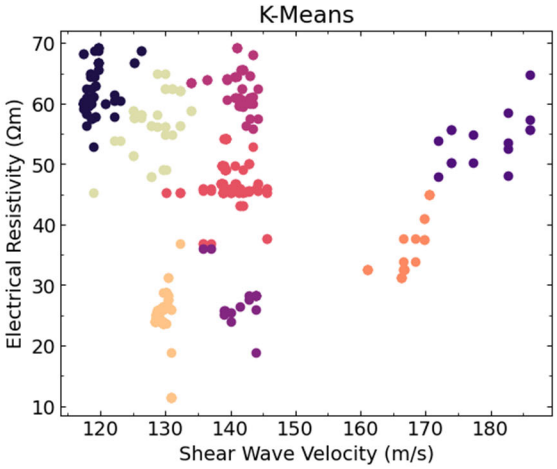


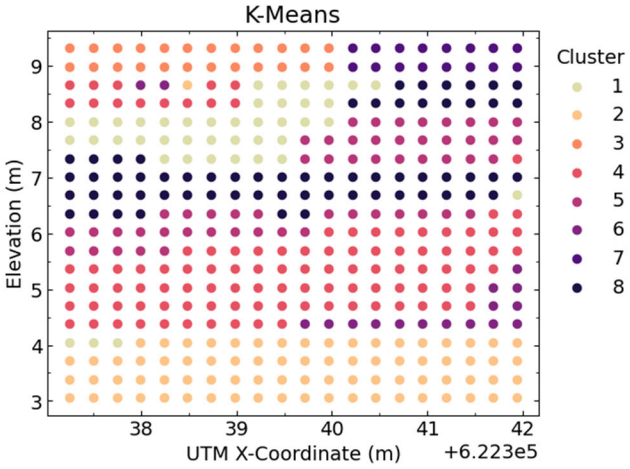


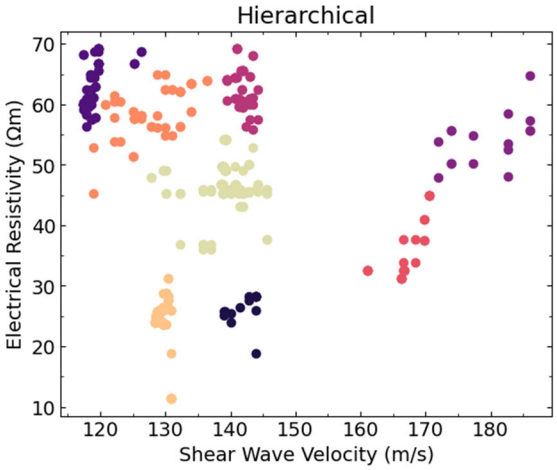


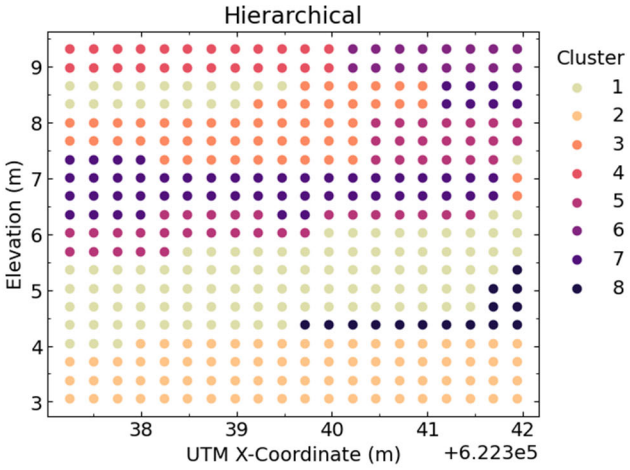
Fit: Individual Profiles

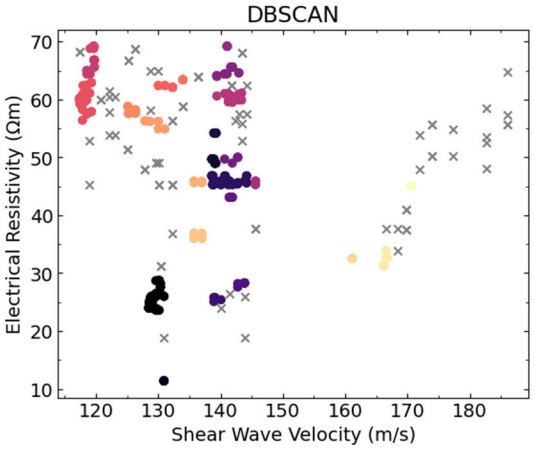
Profile: fwt_site3

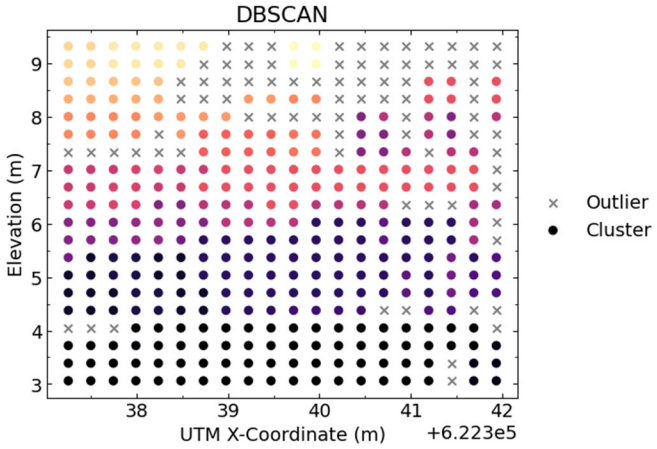






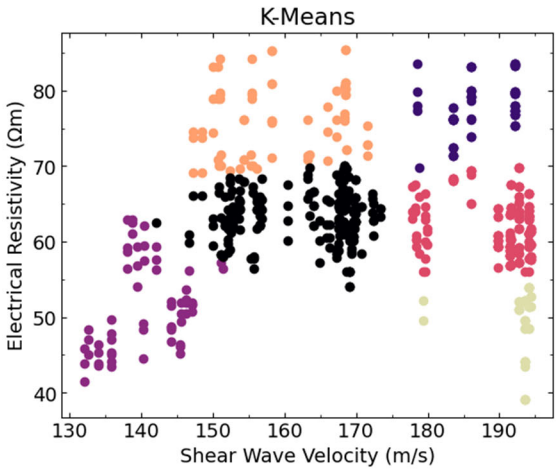


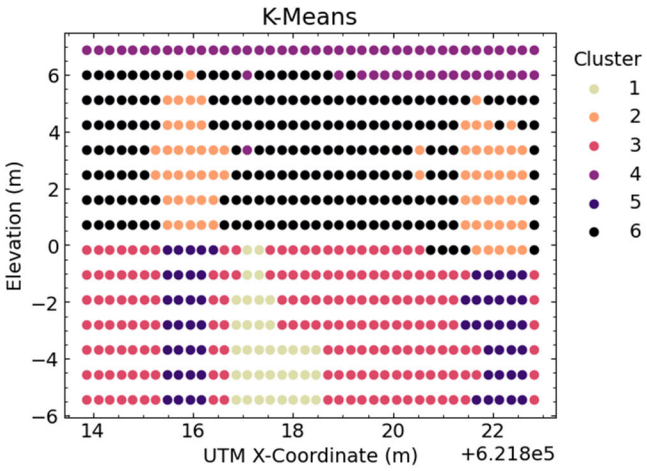


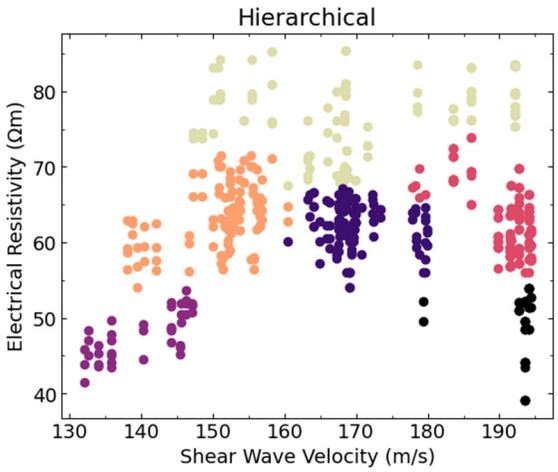


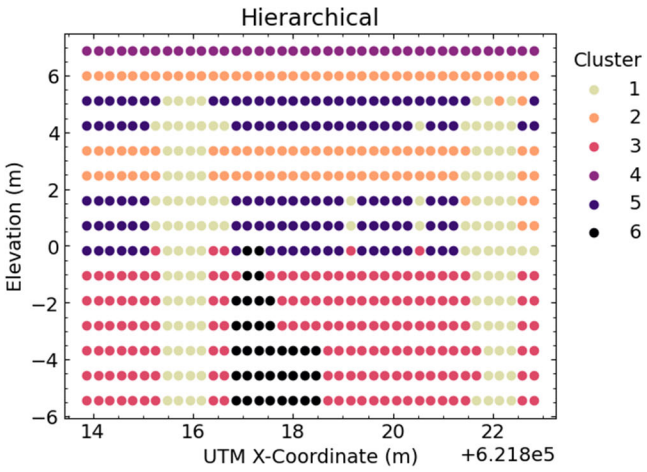
Fit: Individual Profiles

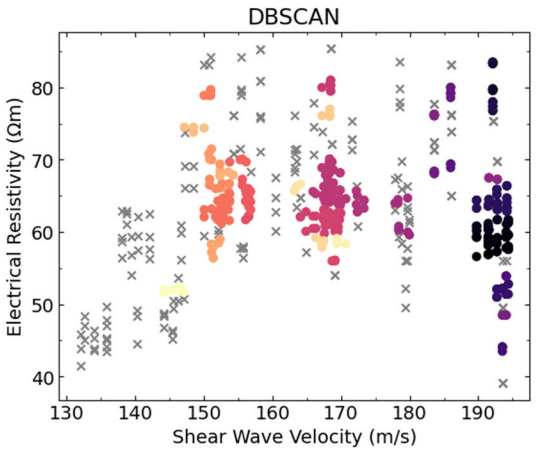
Profile: sc_site2

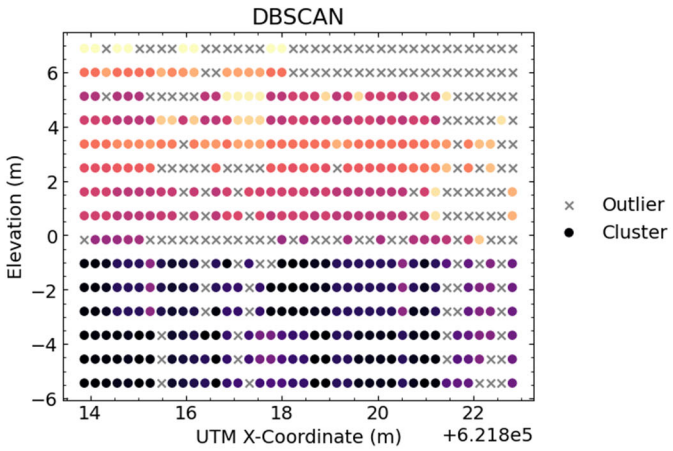






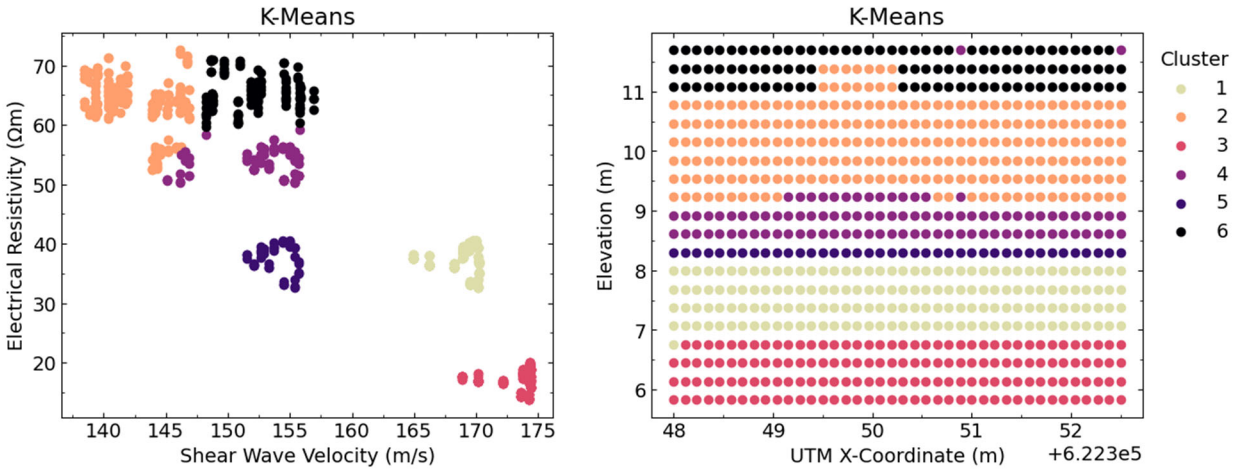


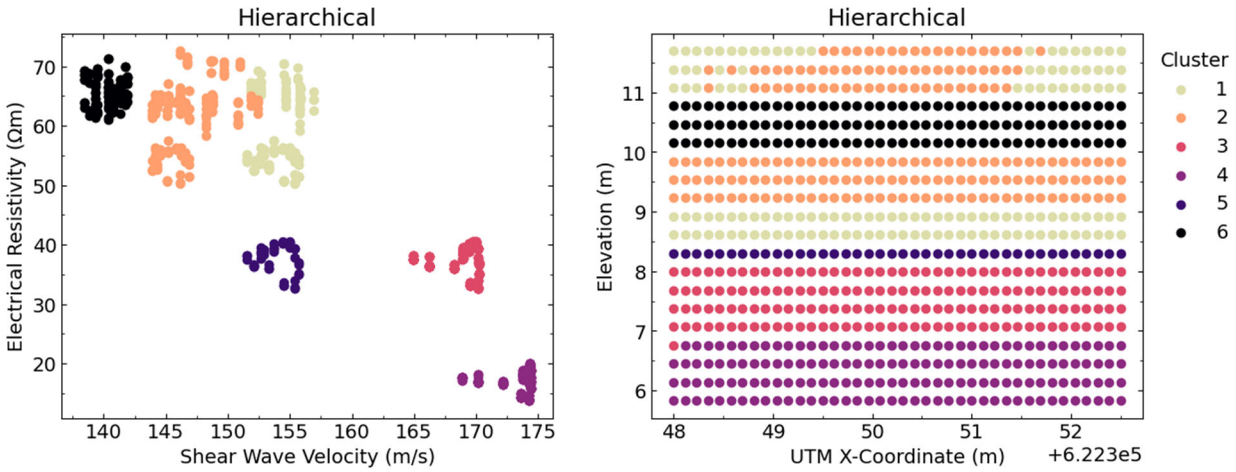


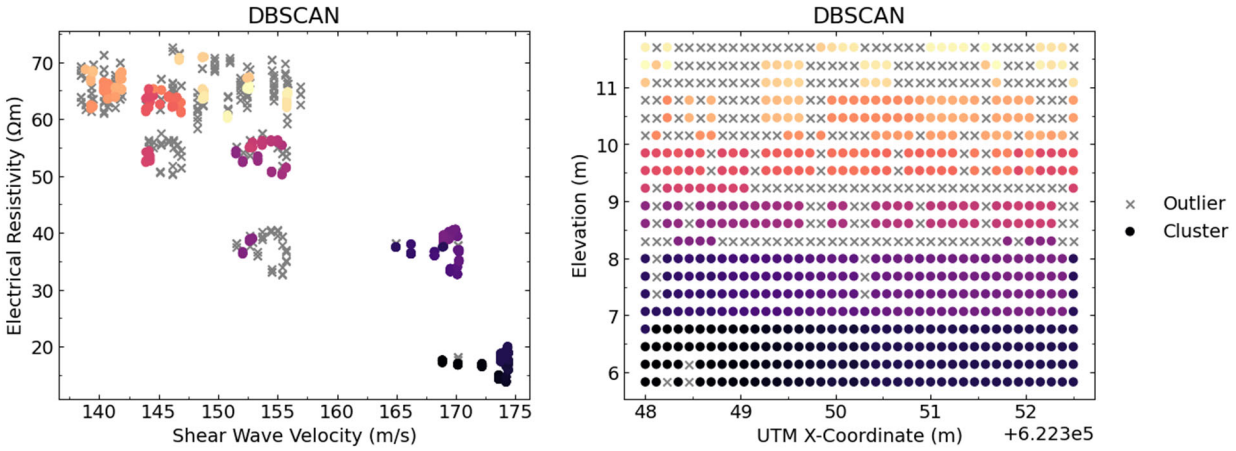


Fit: Individual Profiles

Profile: sc_site3

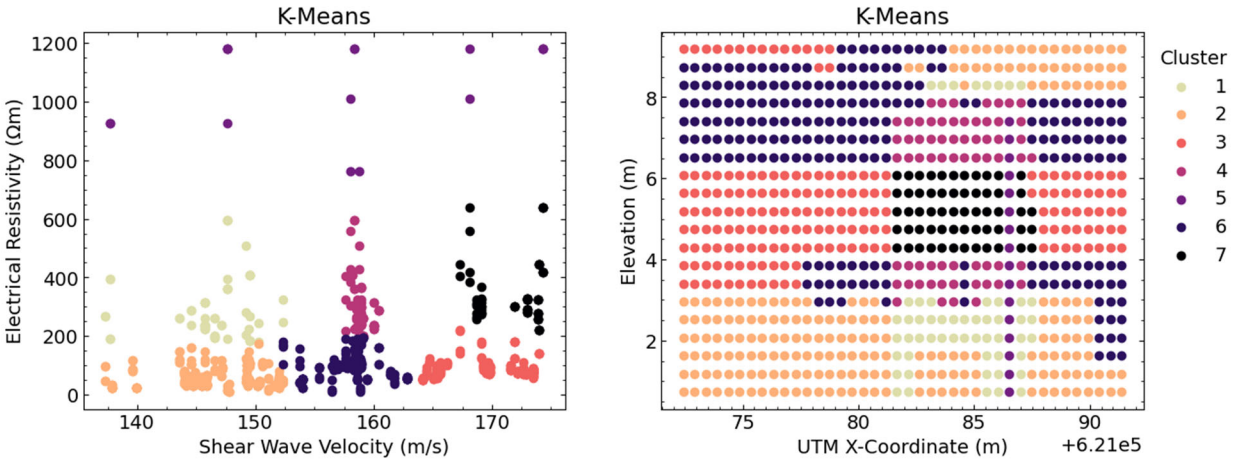


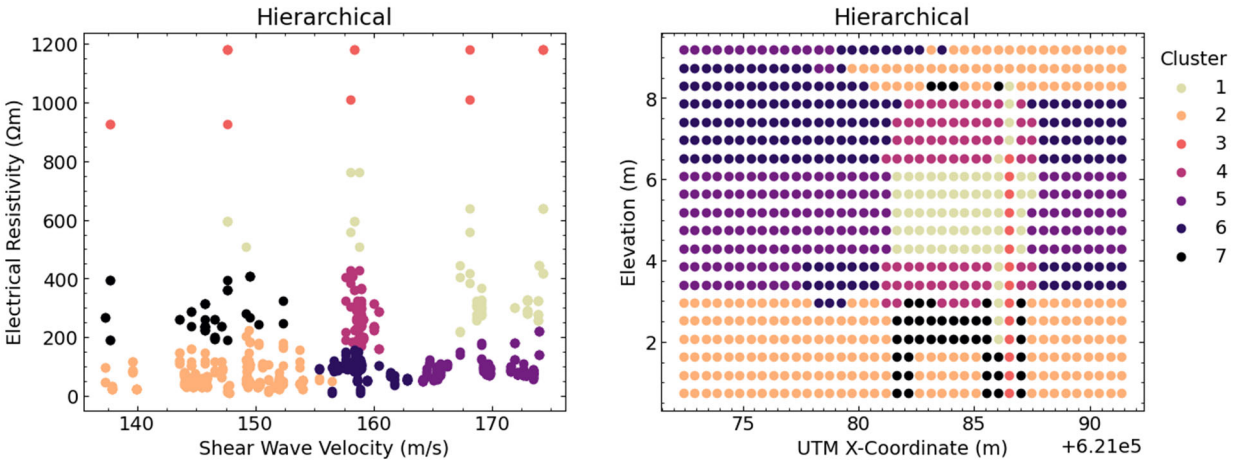


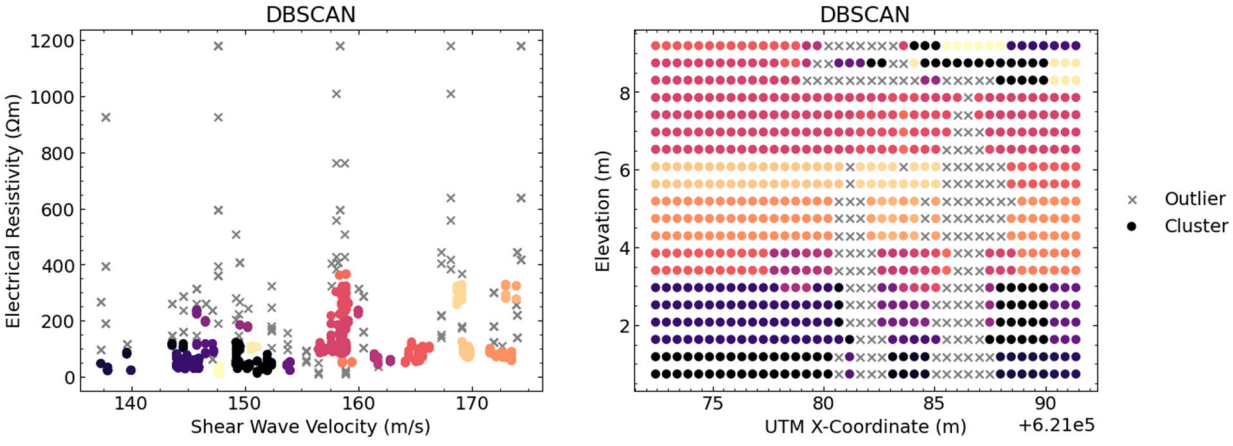


Fit: Individual Profiles

Profile: sc3_site1

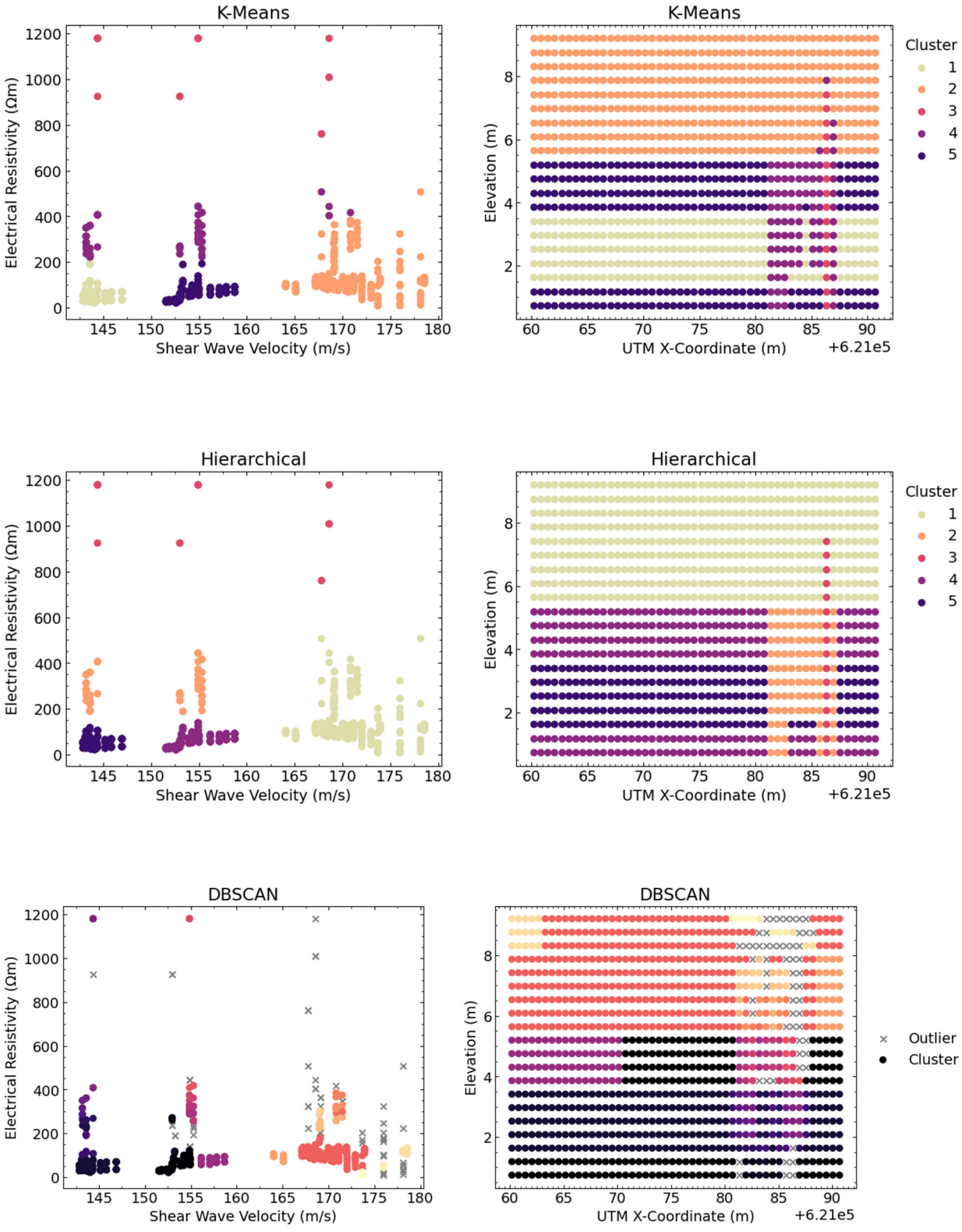






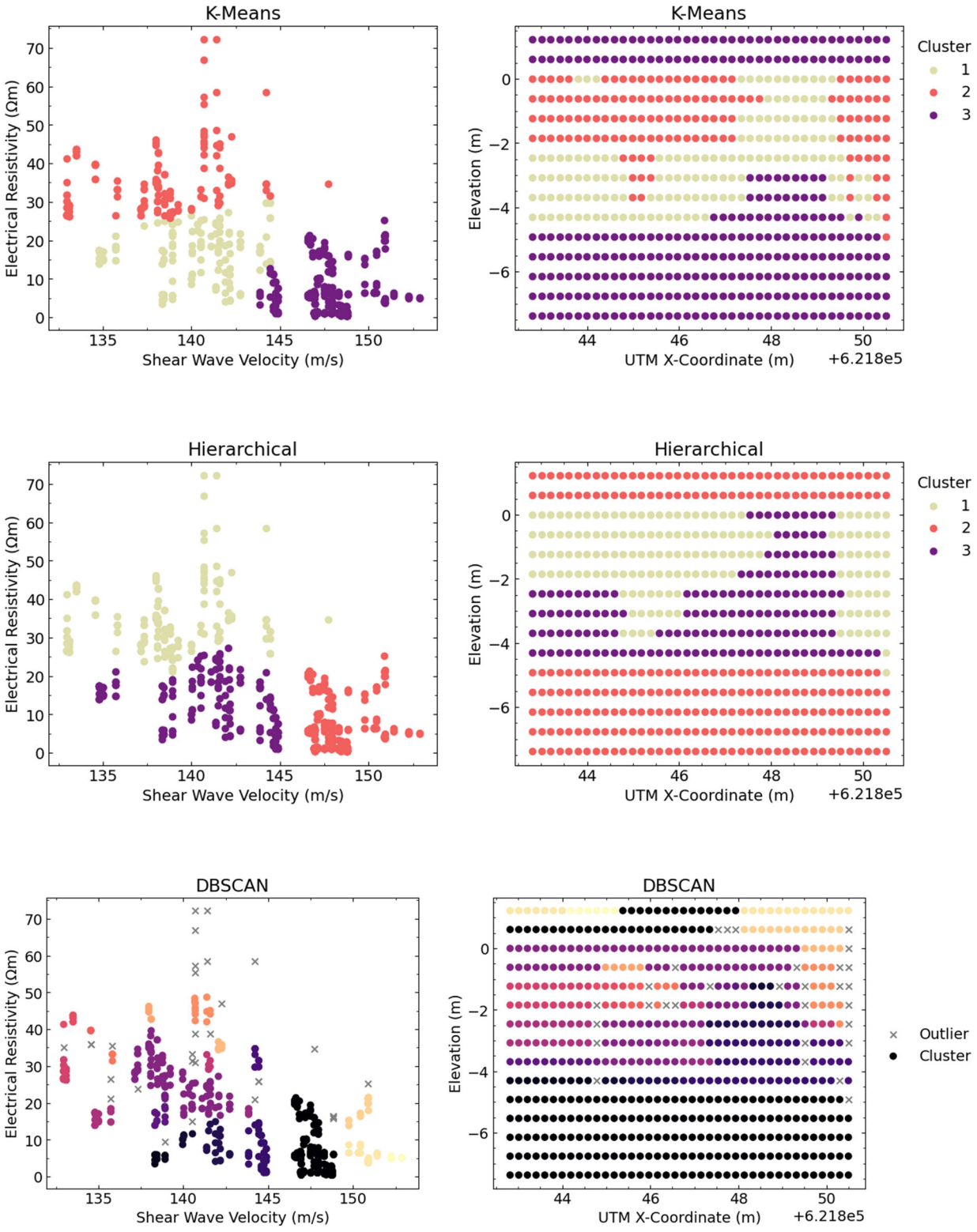
Fit: Individual Profiles

Profile: sc5_site1



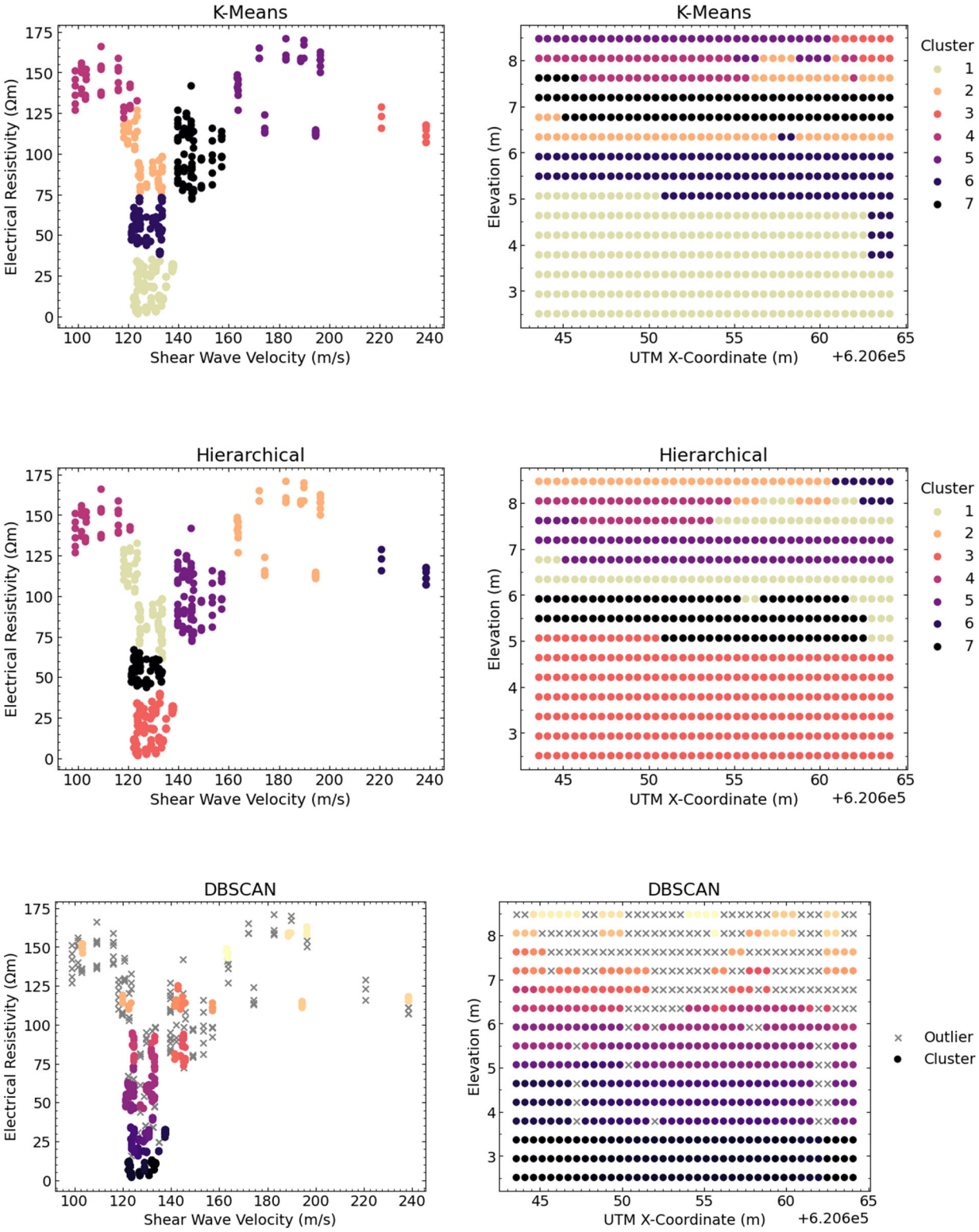
Fit: Individual Profiles

Profile: sdt_site2



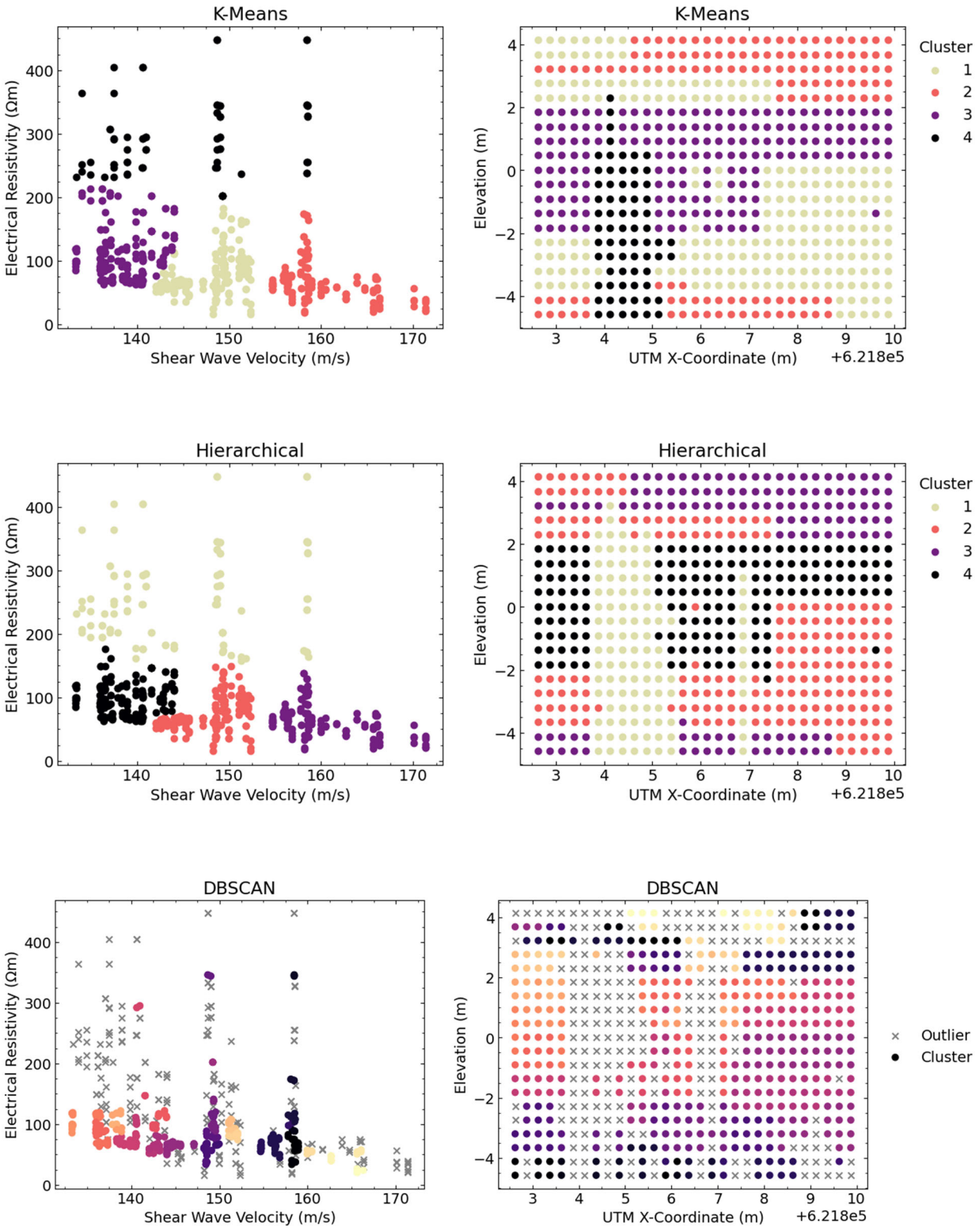
Fit: Individual Profiles

Profile: swt_site1



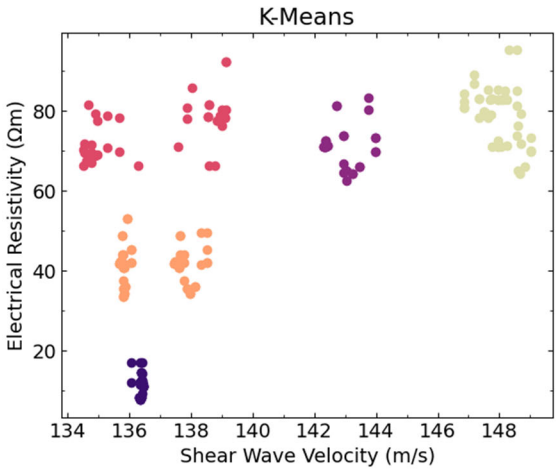
Fit: Individual Profiles

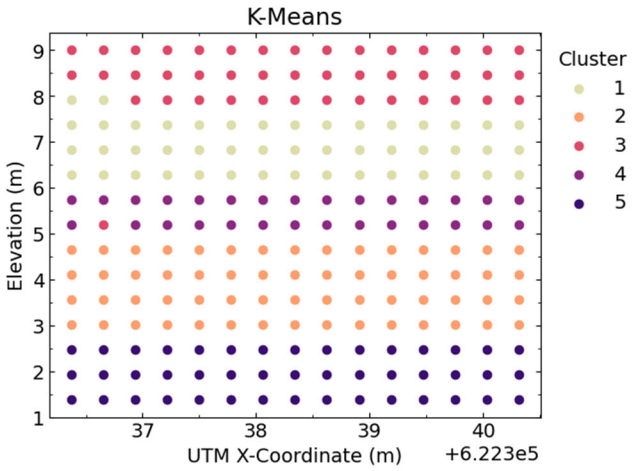
Profile: swt_site2

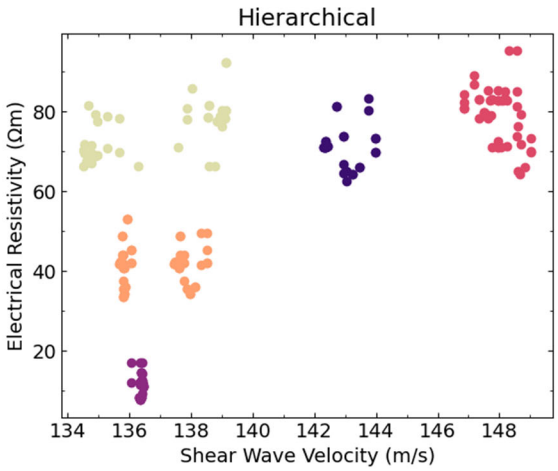


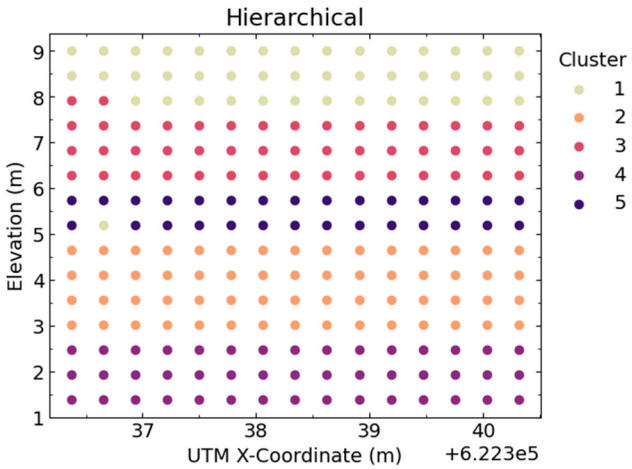
Fit: Individual Profiles

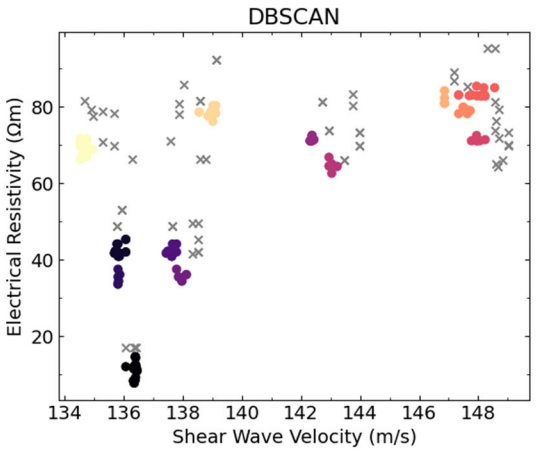
Profile: swt_site3

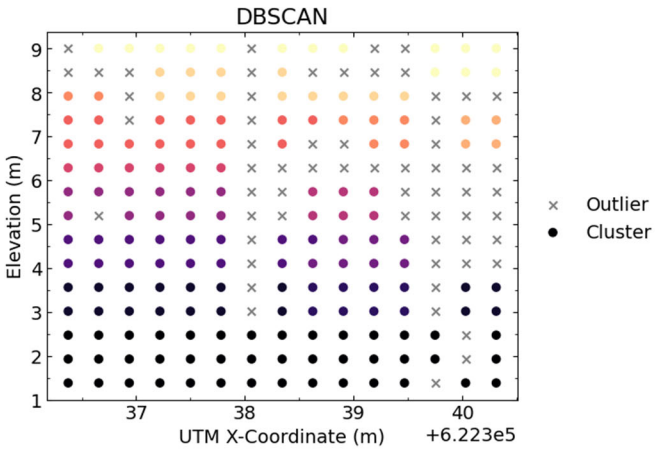






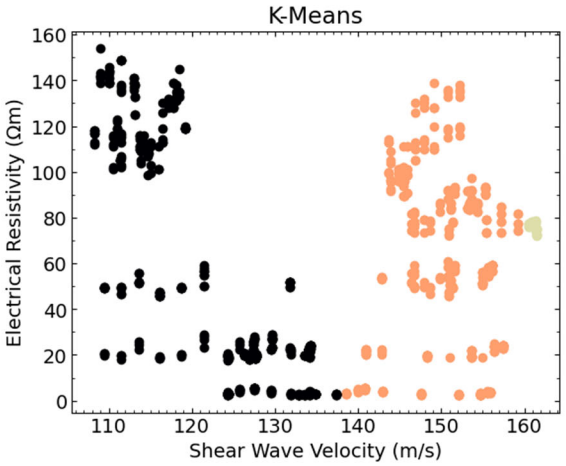


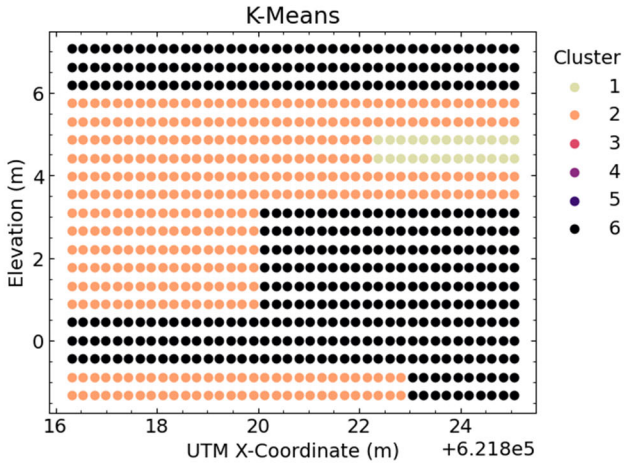


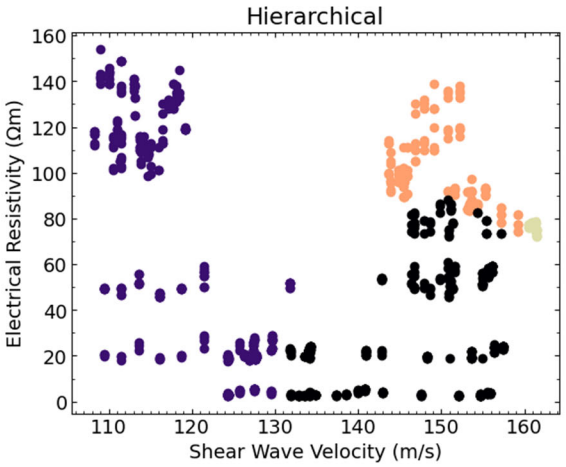


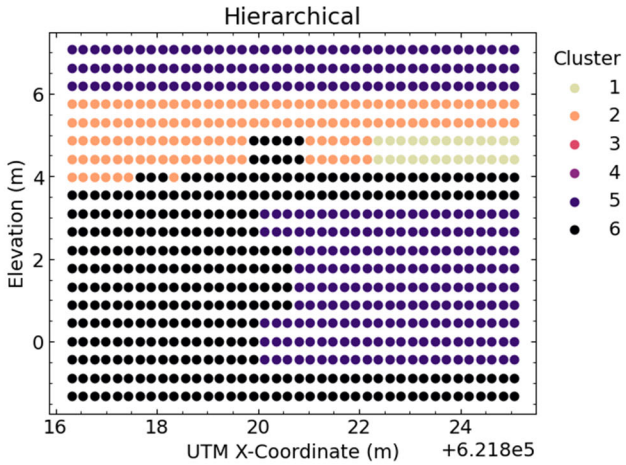
Fit: All_Crest

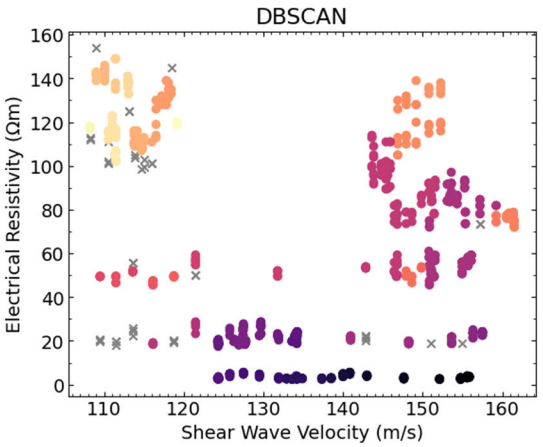
Profile: fc_site2

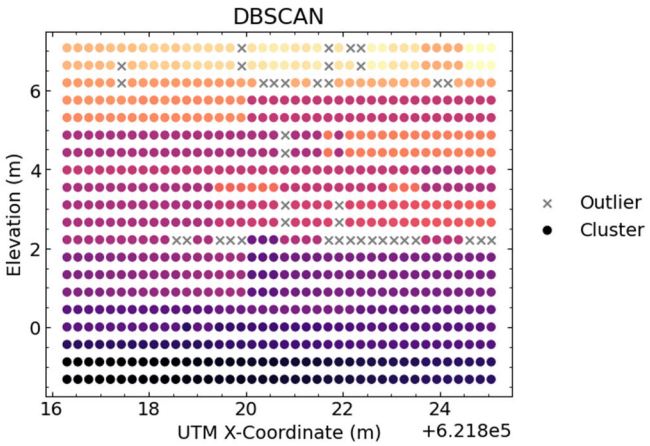






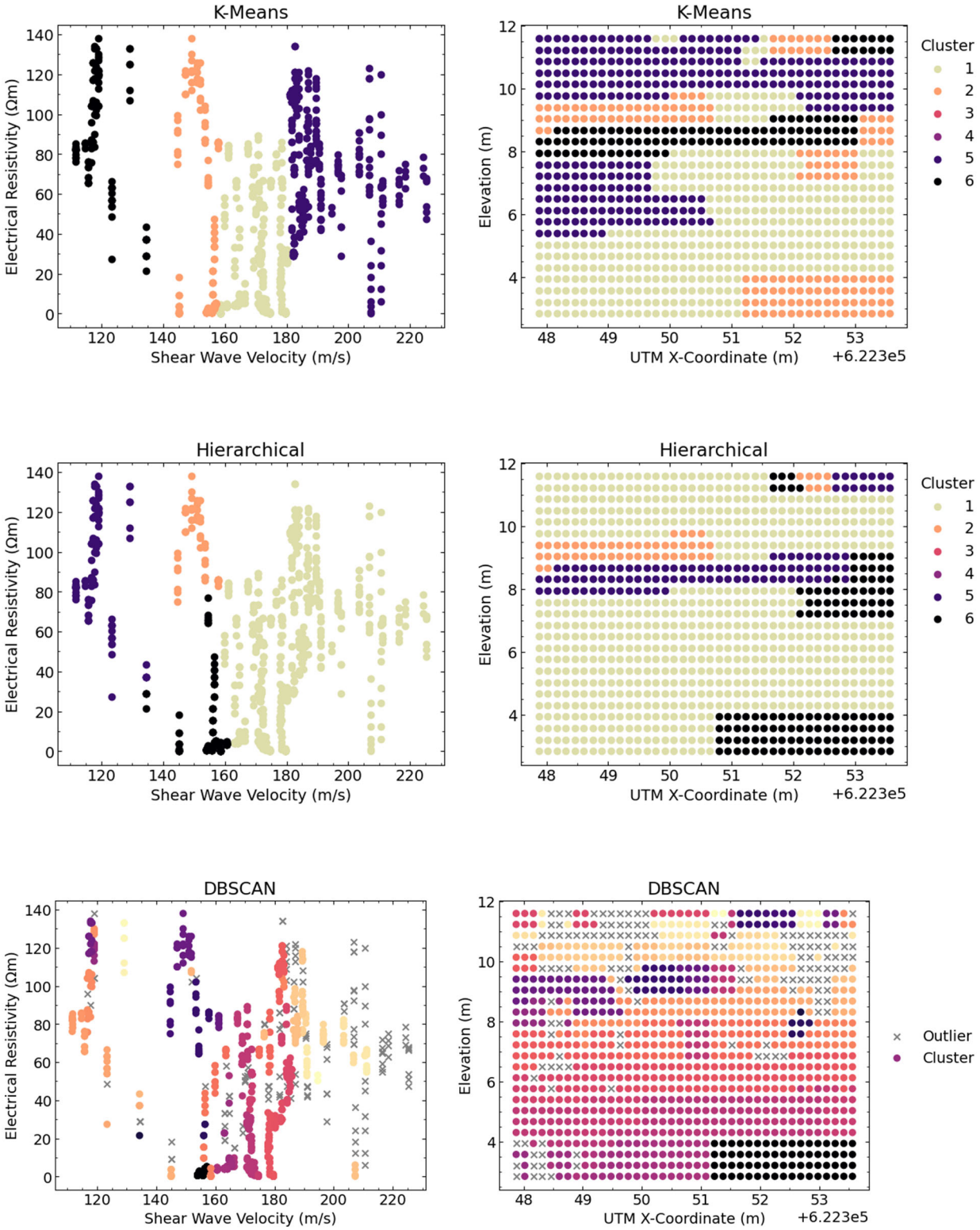






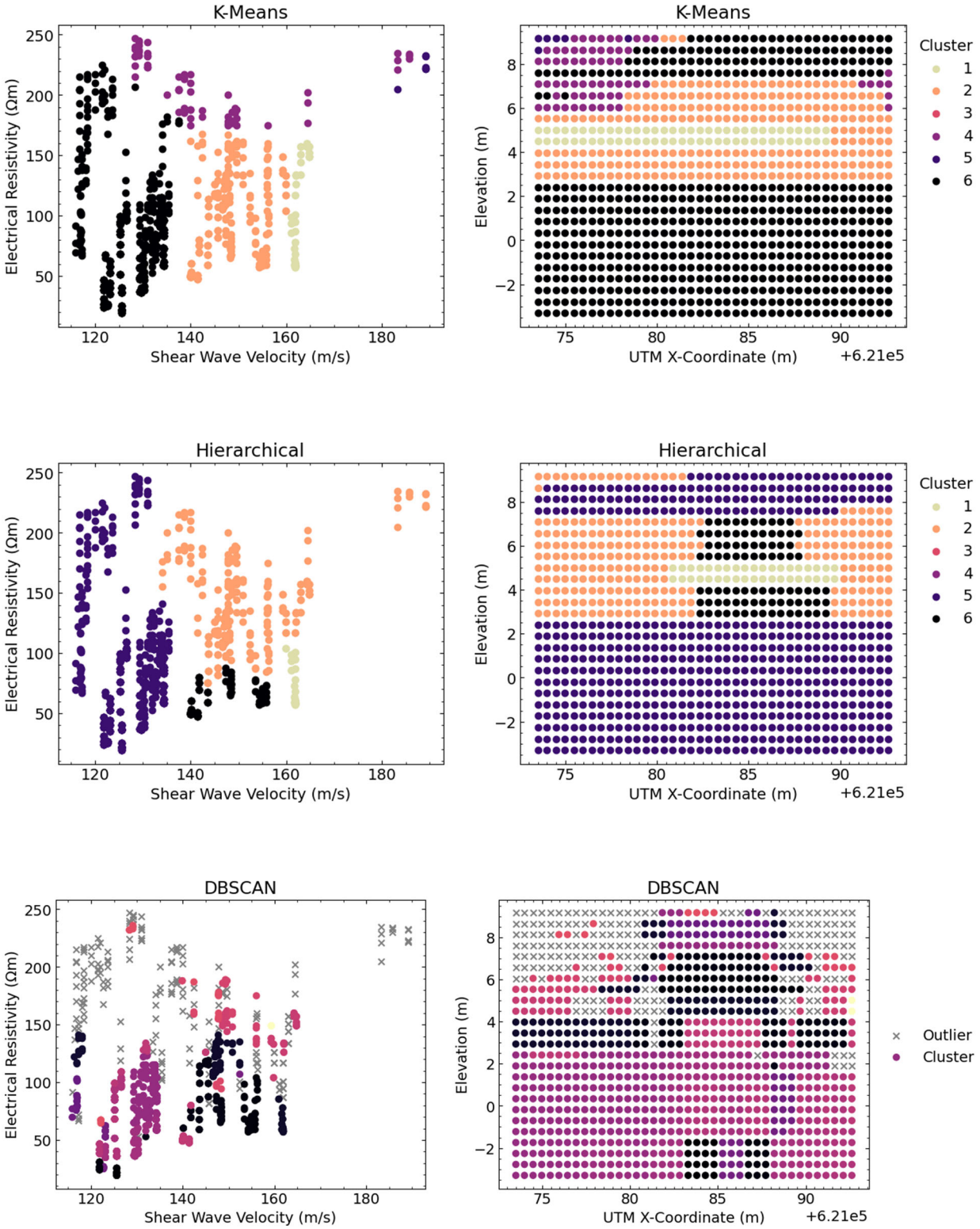
Fit: All_Crest

Profile: fc_site3



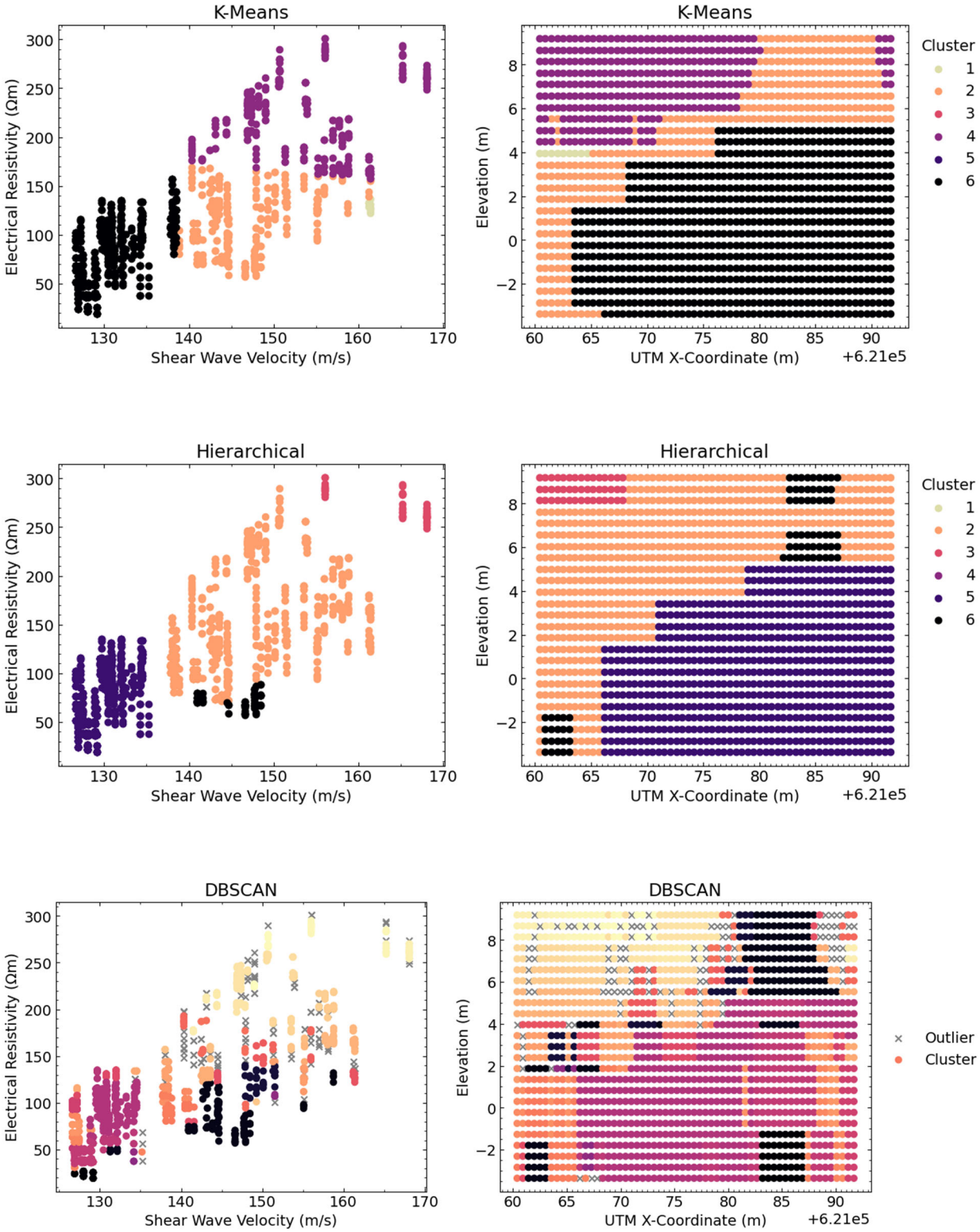
Fit: All_Crest

Profile: fc3_site1



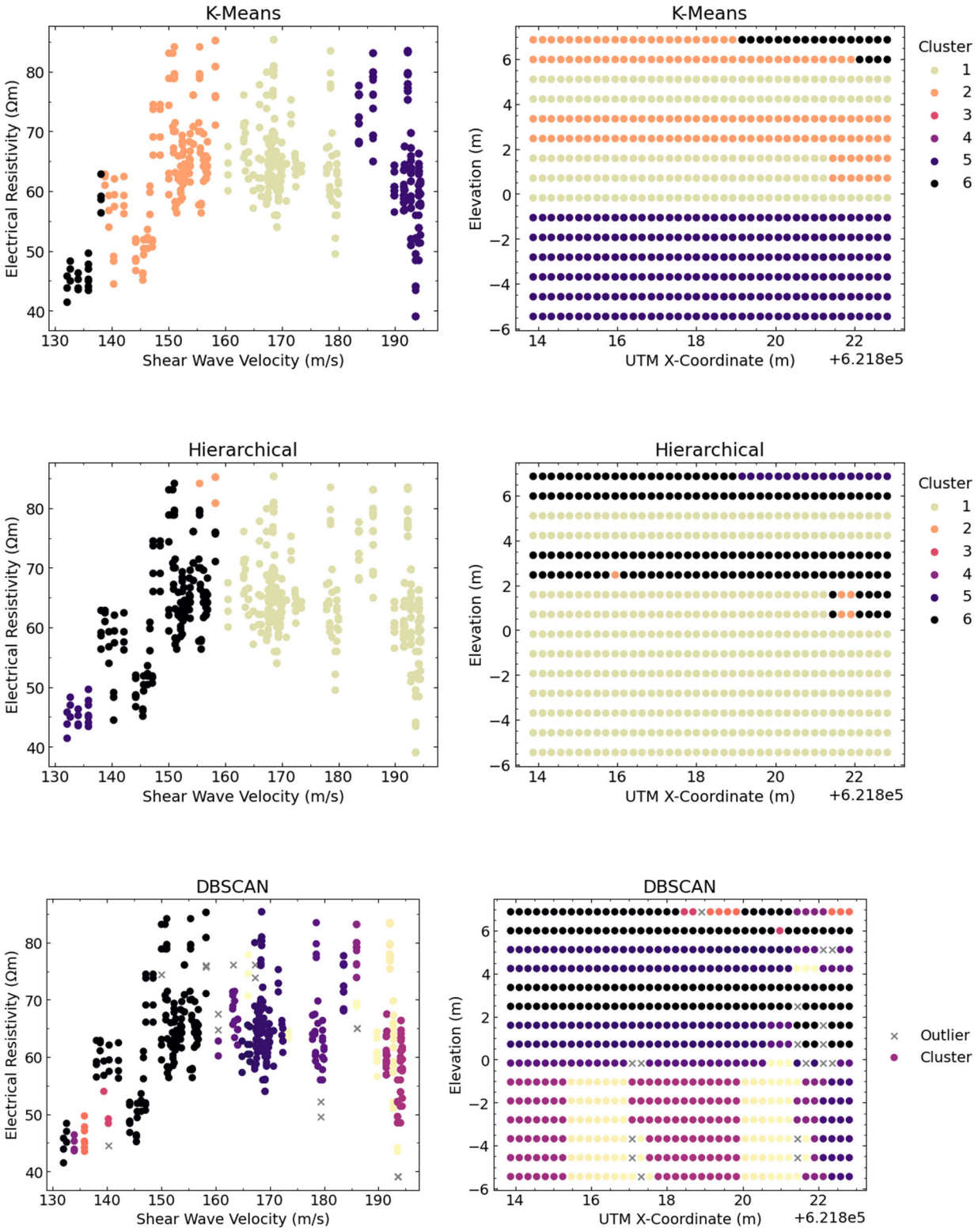
Fit: All_Crest

Profile: fc5_site1



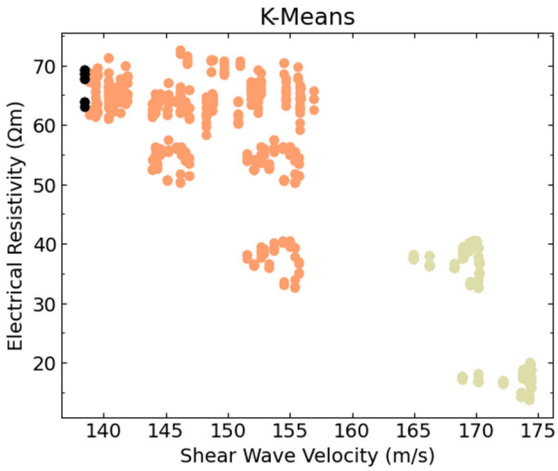
Fit: All_Crest

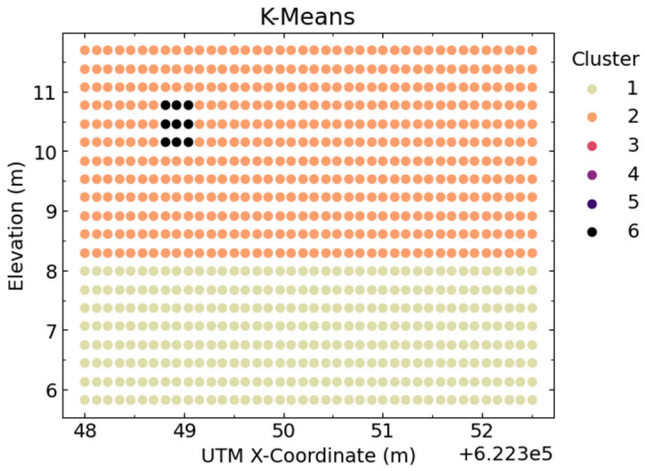
Profile: sc_site2

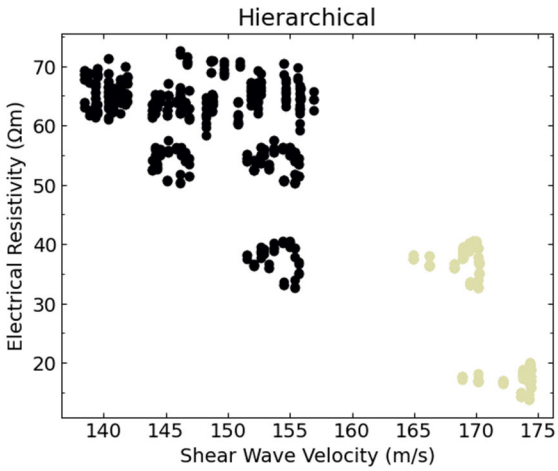


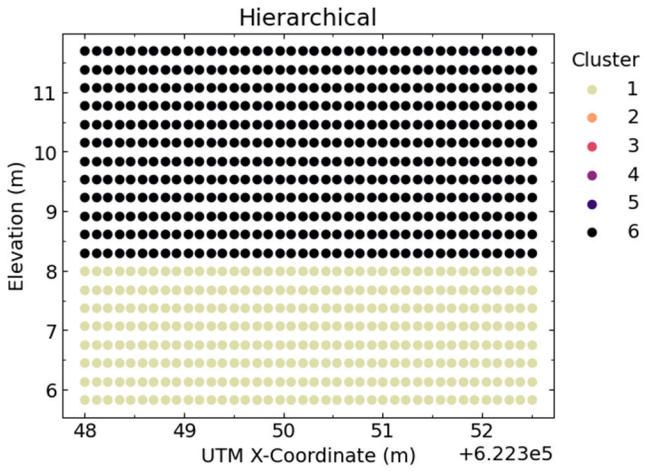
Fit: All_Crest

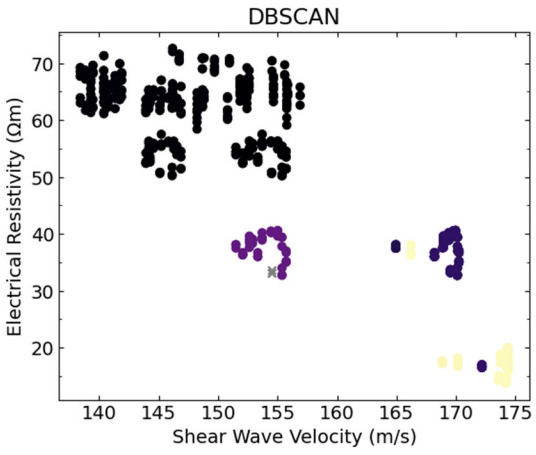
Profile: sc_site3

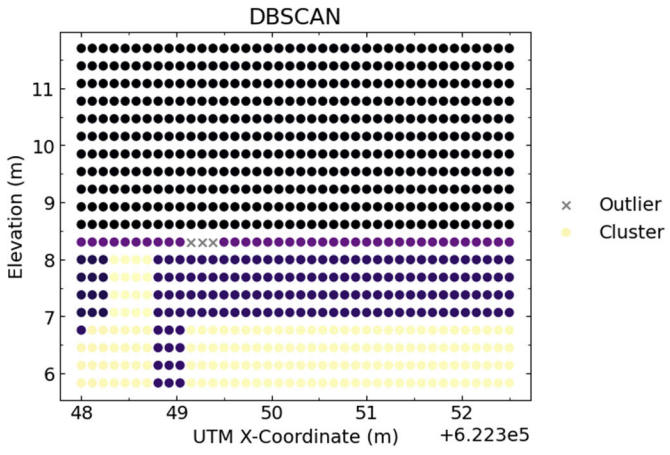






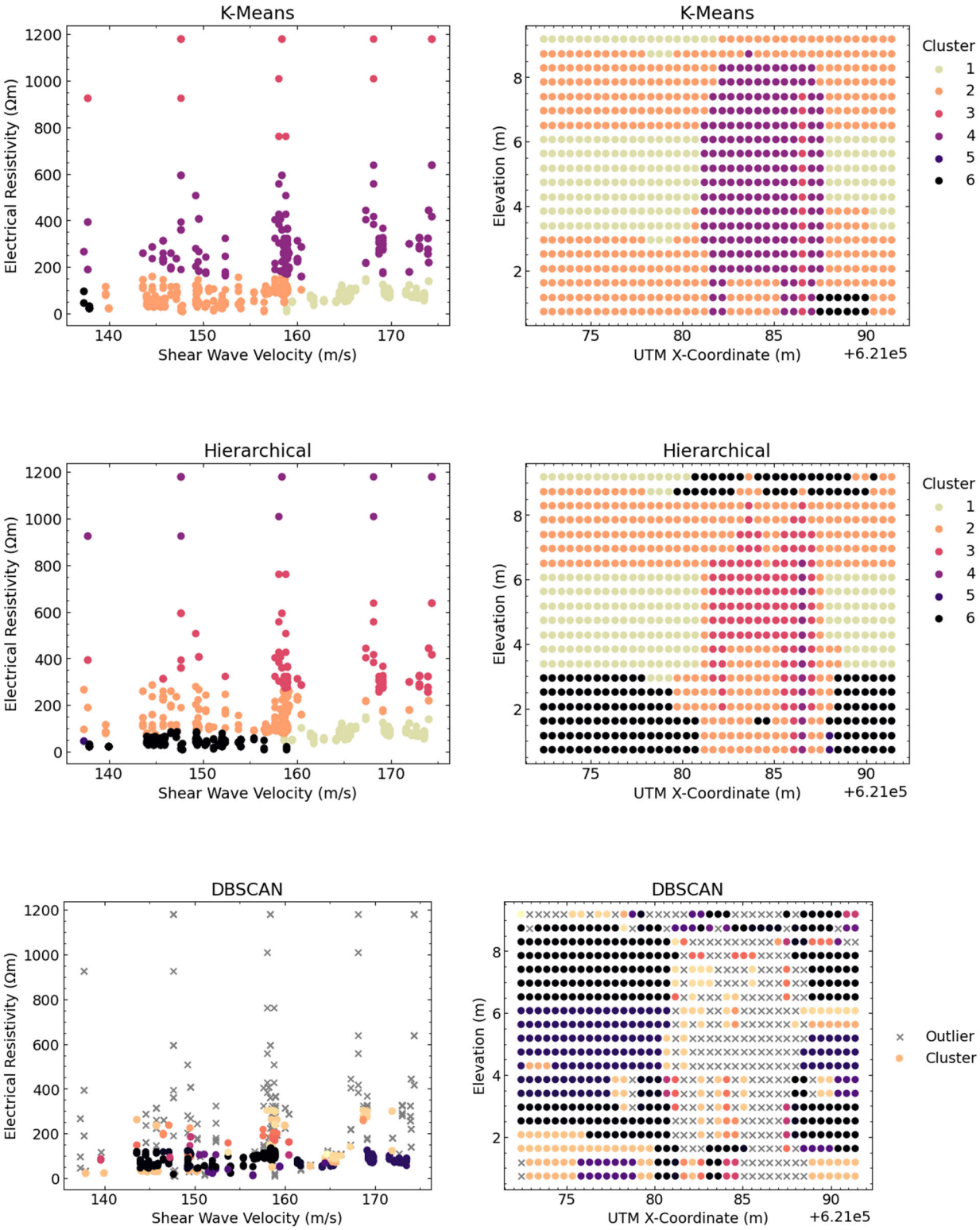






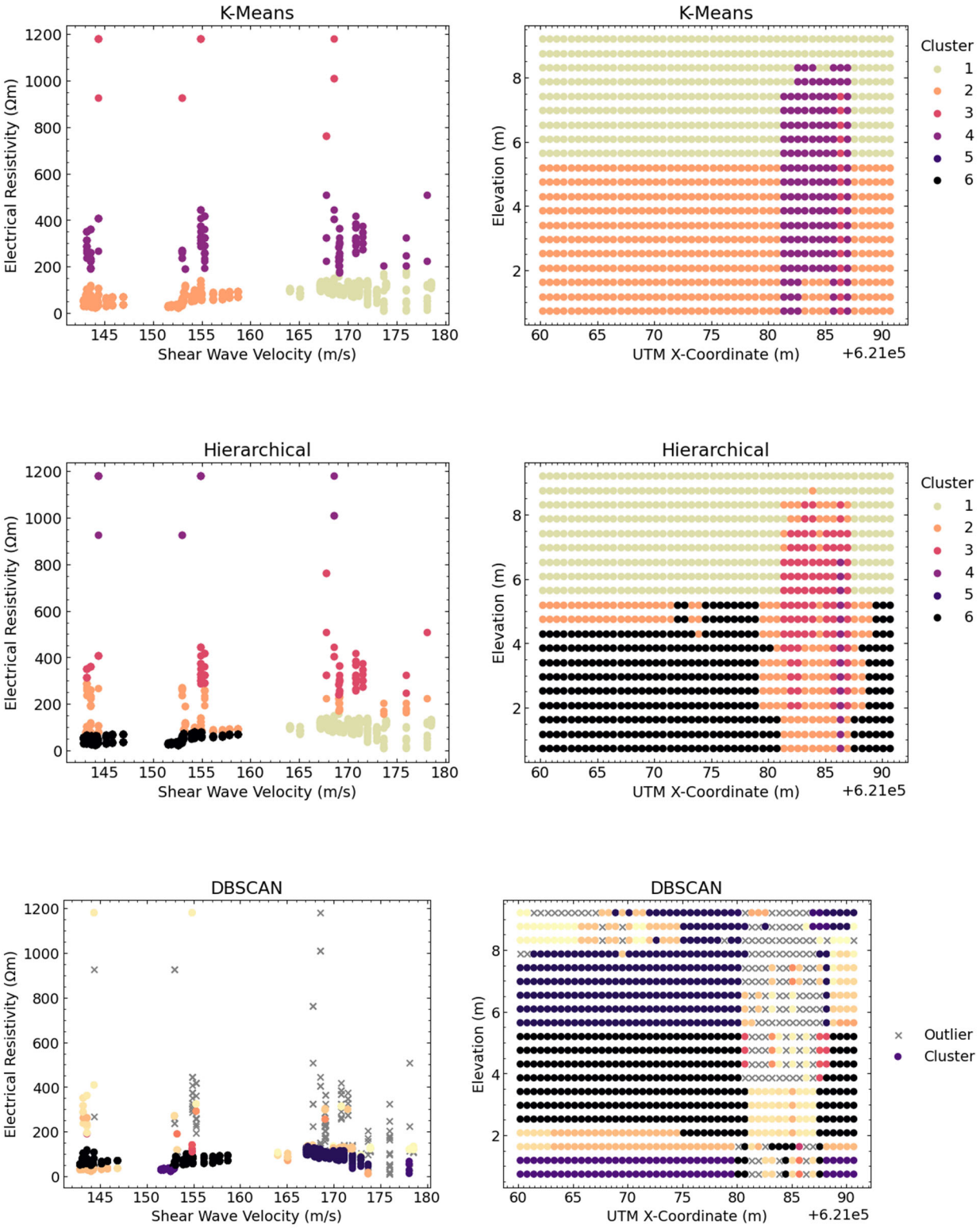
Fit: All_Crest

Profile: sc3_site1



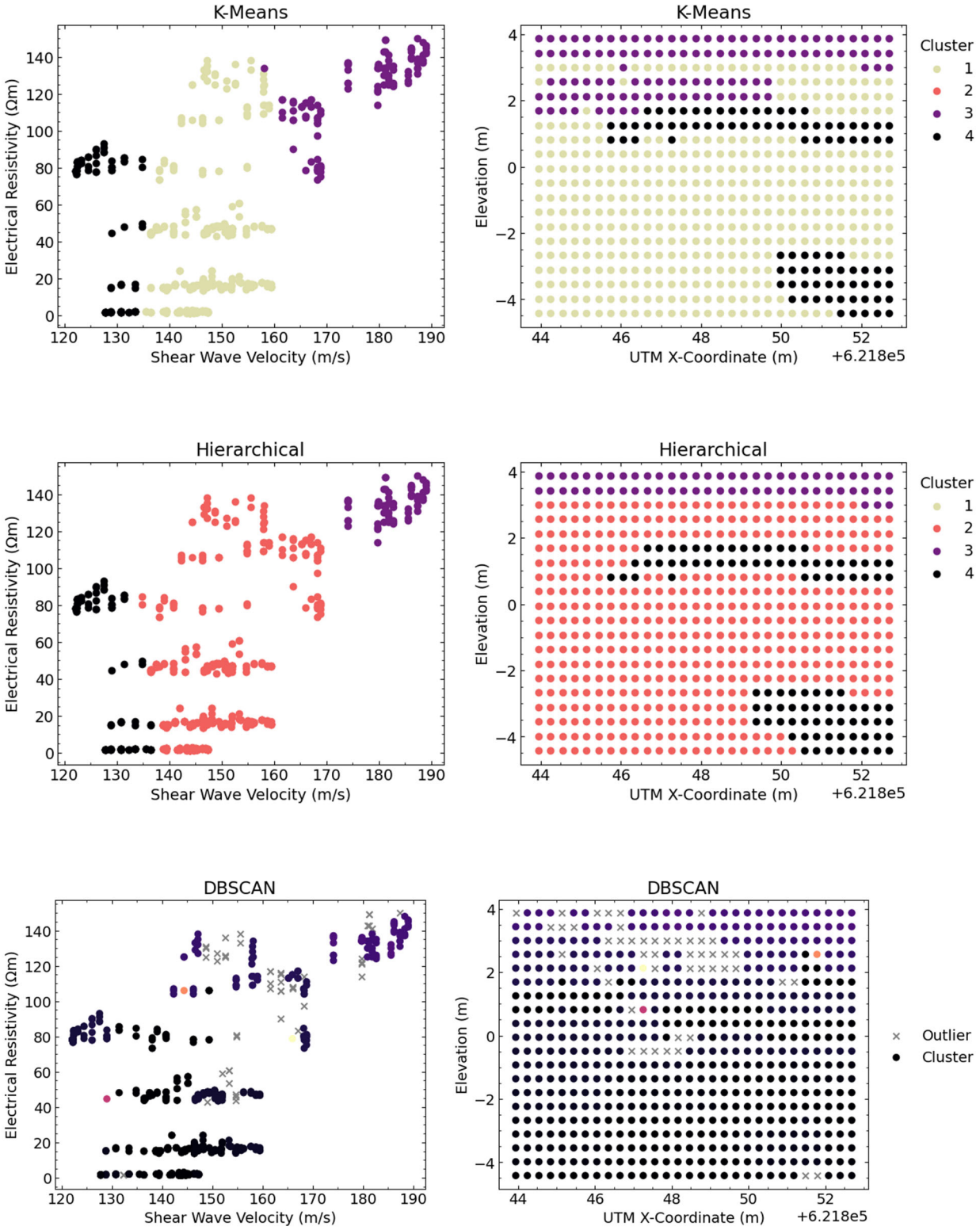
Fit: All_Crest

Profile: sc5_site1



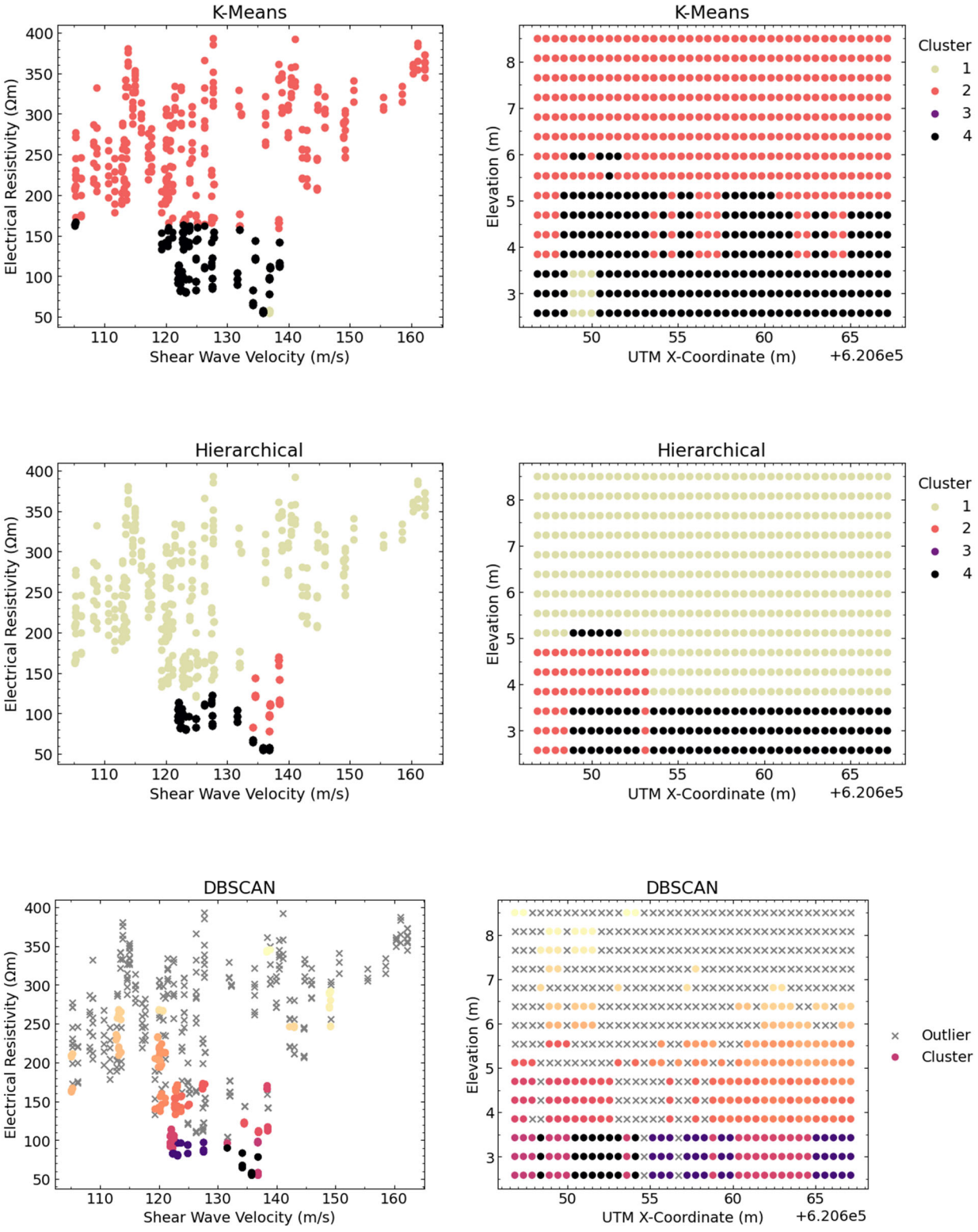
Fit: All_Toe

Profile: fdt_site2



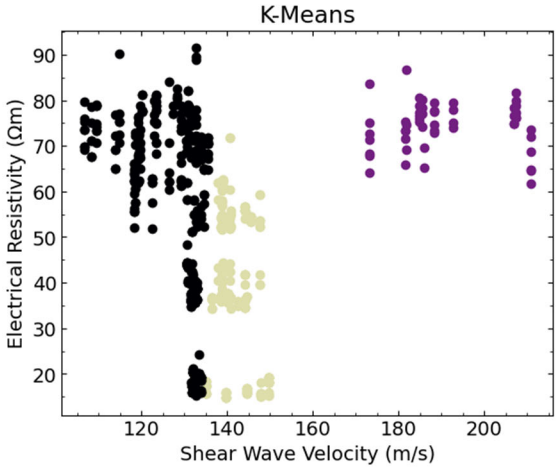
Fit: All_Toe

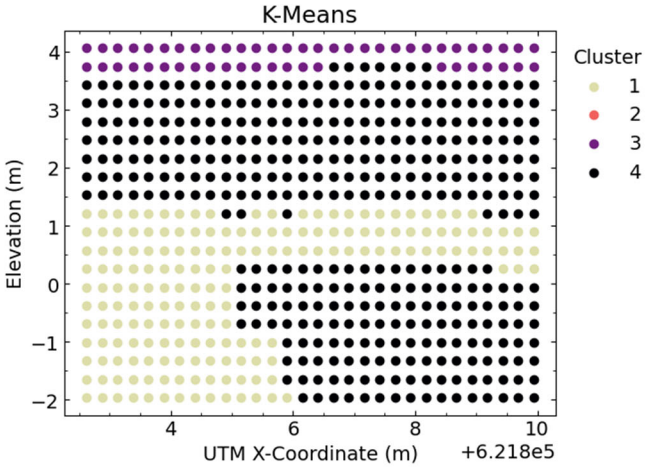
Profile: fwt_site1

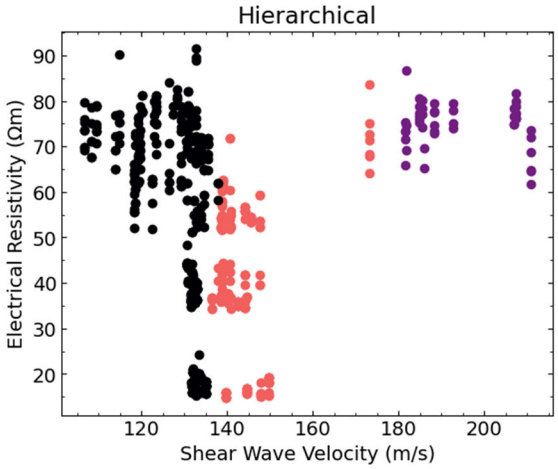


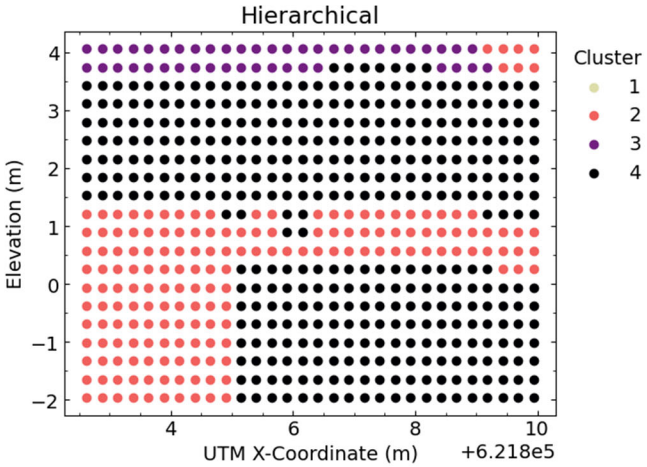
Fit: All_Toe

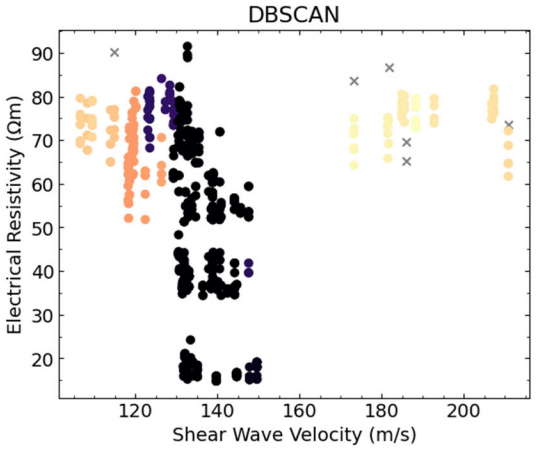
Profile: fwt_site2

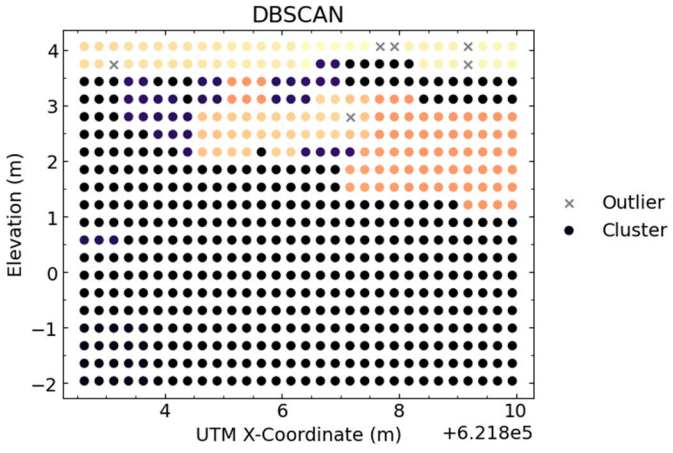






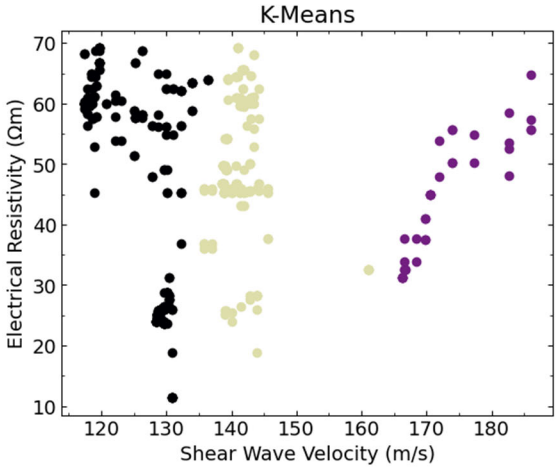


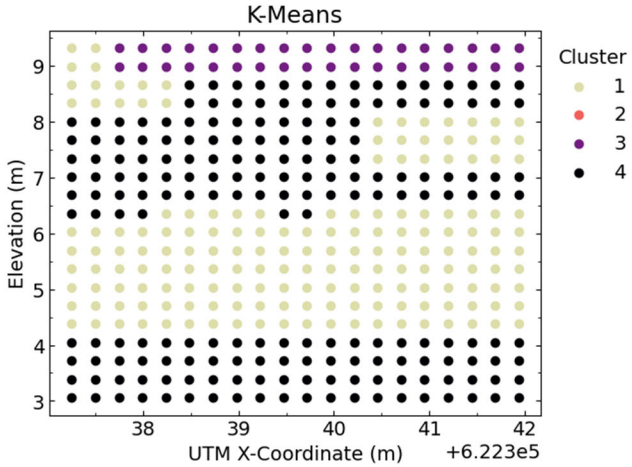


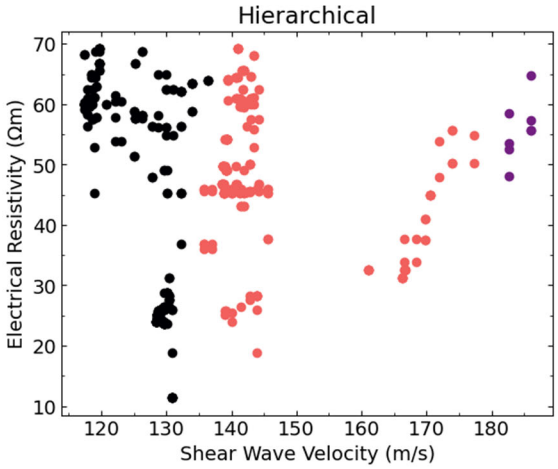


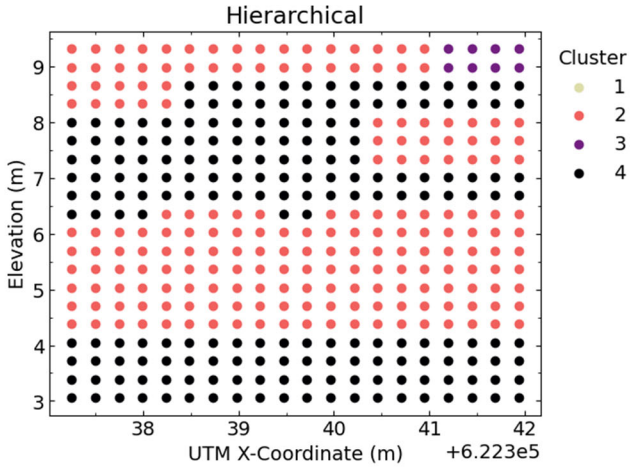
Fit: All_Toe

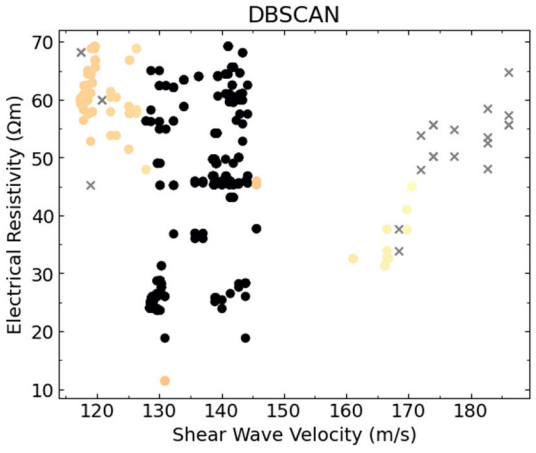
Profile: fwt_site3

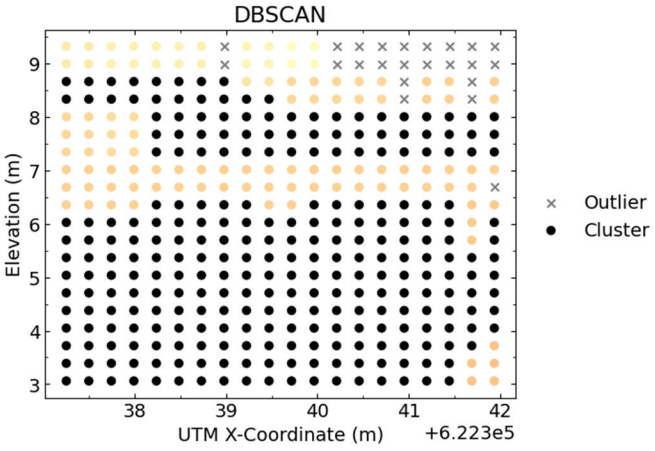






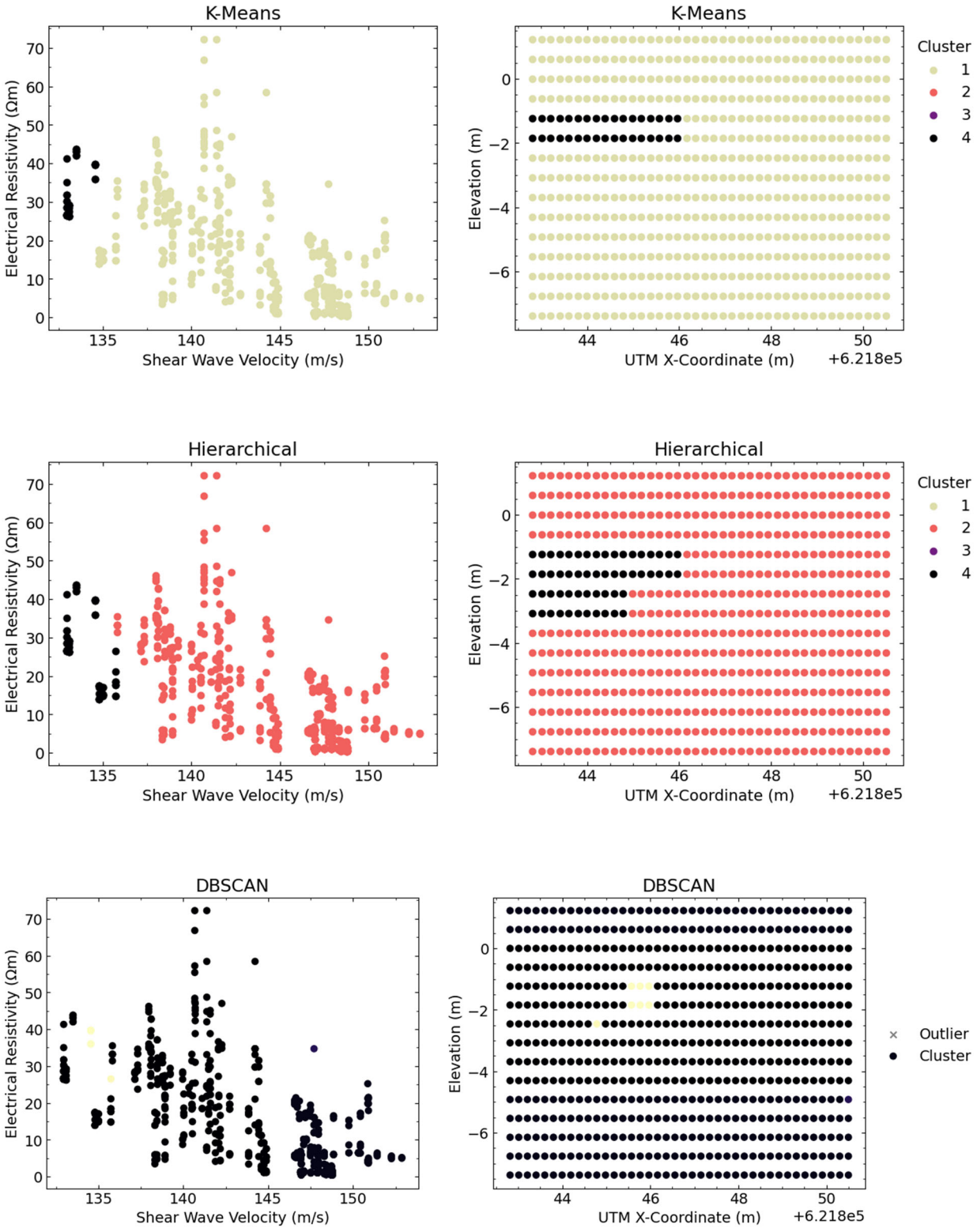






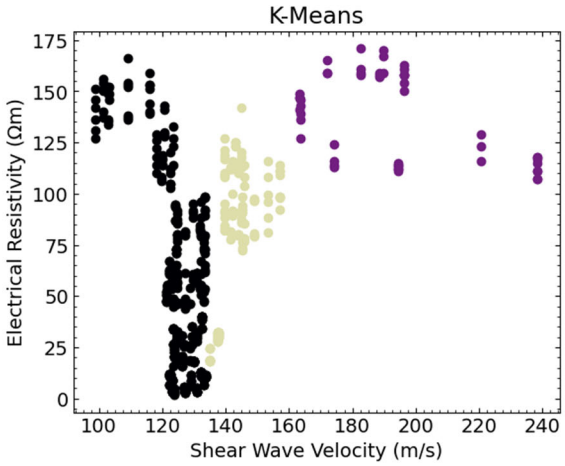
Fit: All_Toe

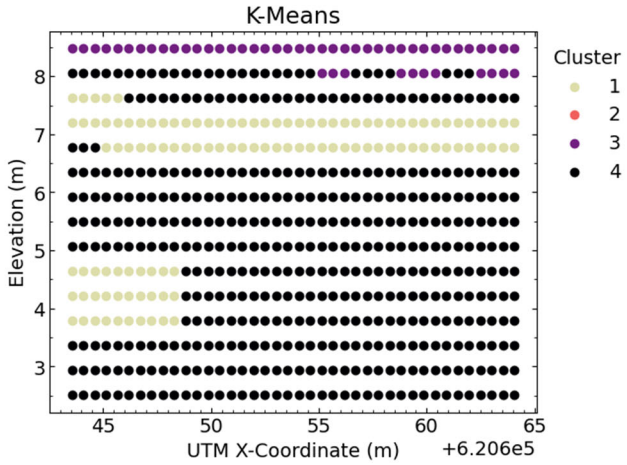
Profile: sdt_site2

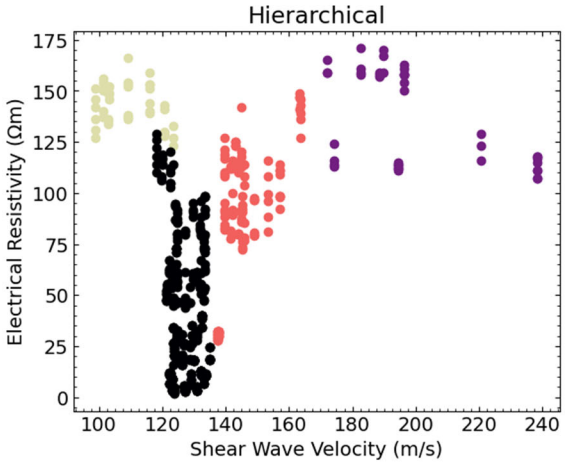


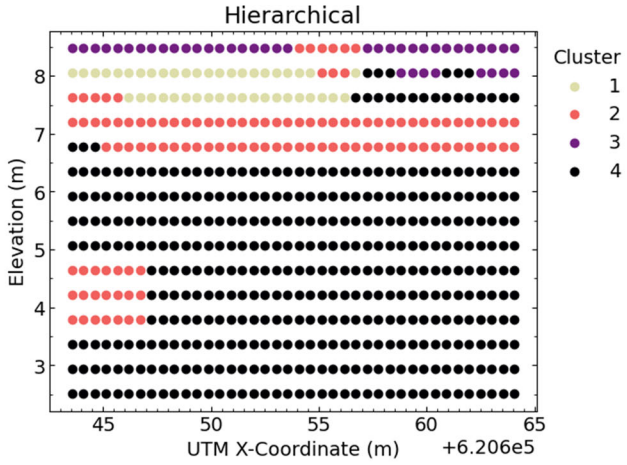
Fit: All_Toe

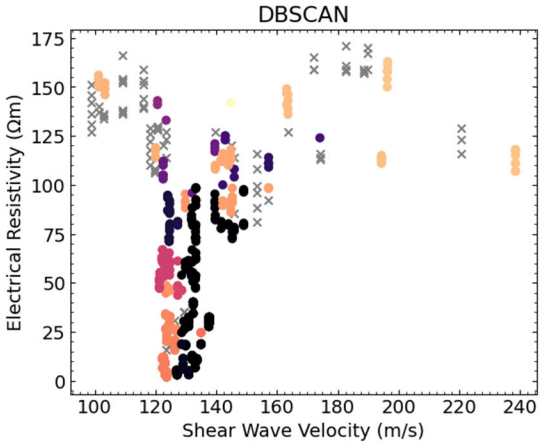
Profile: swt_site1

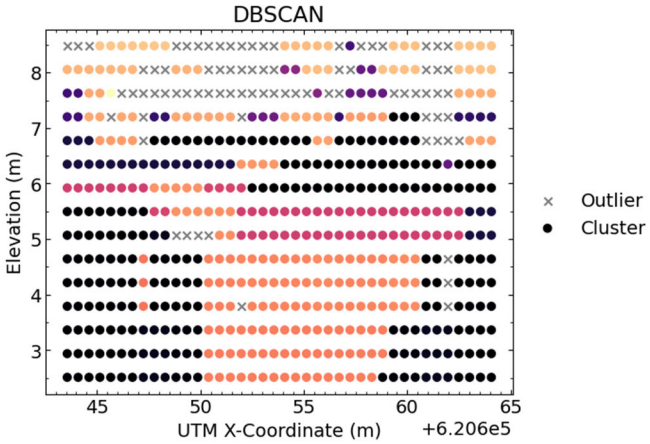






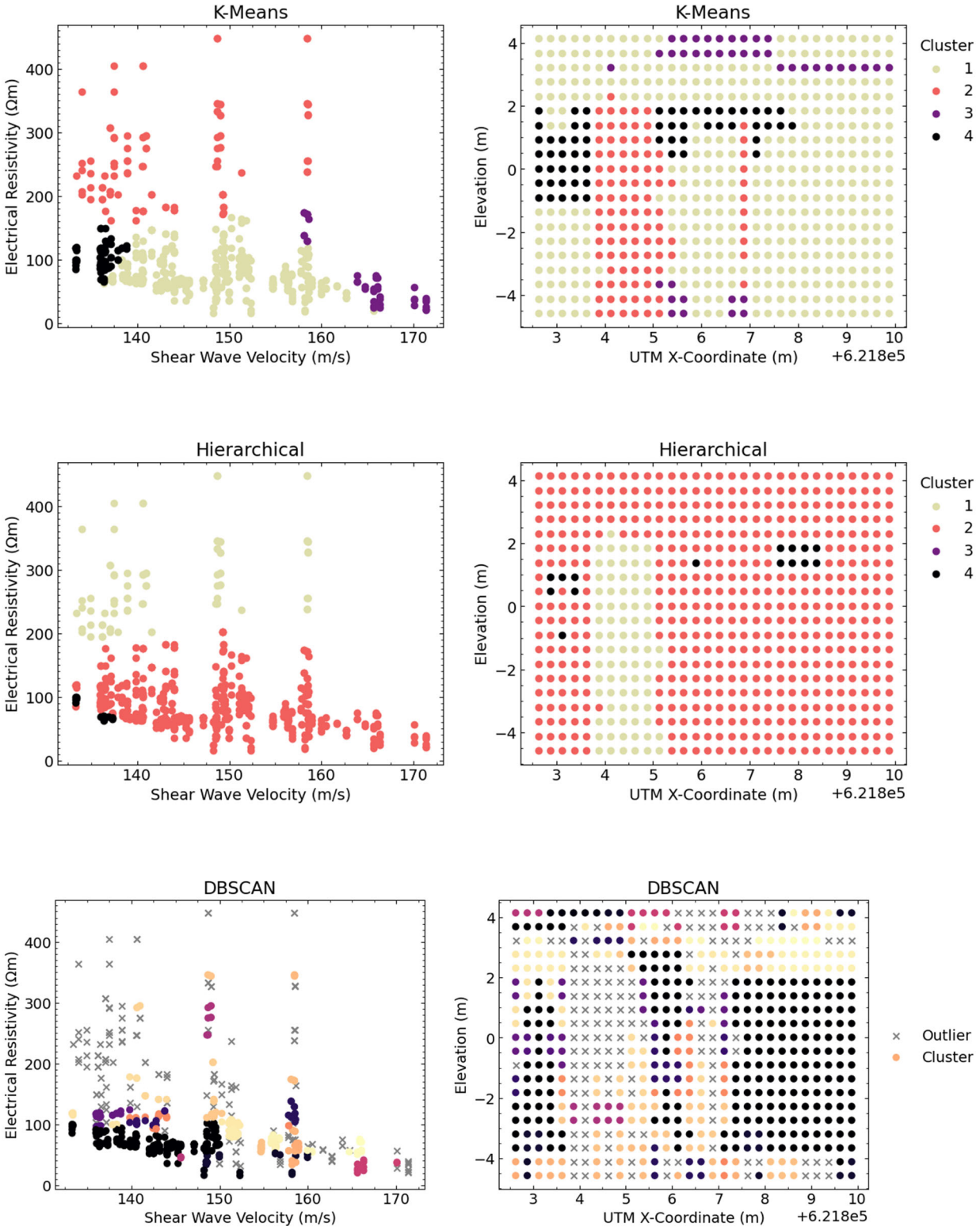






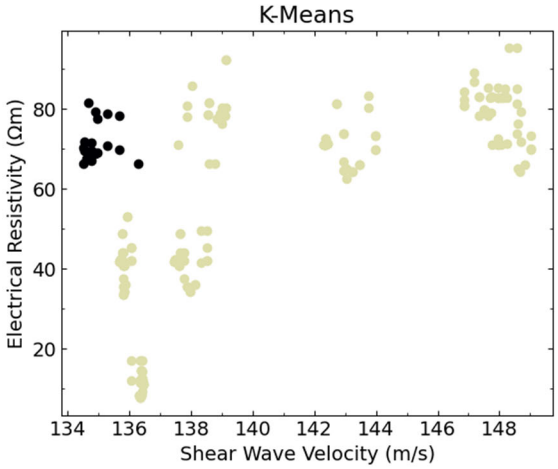
Fit: All_Toe

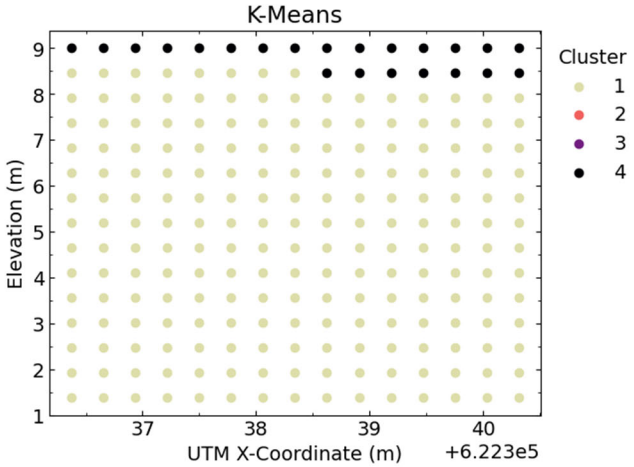
Profile: swt_site2

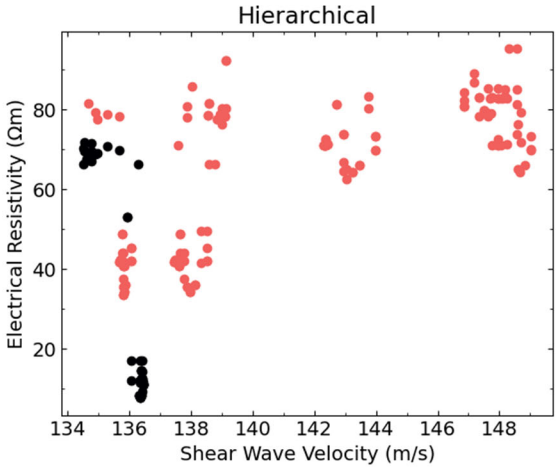


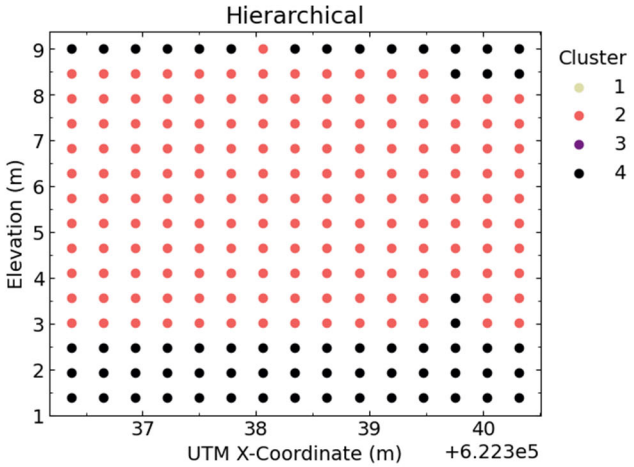
Fit: All_Toe

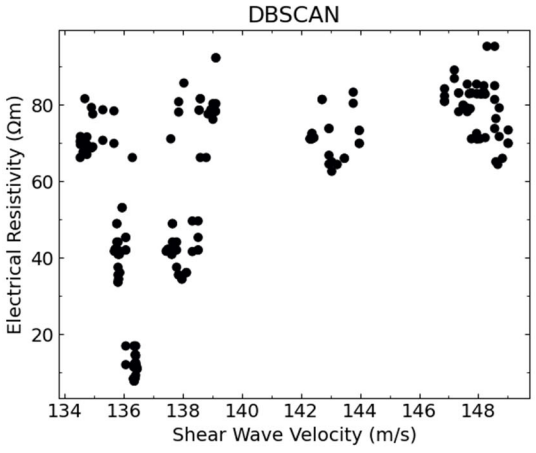
Profile: swt_site3

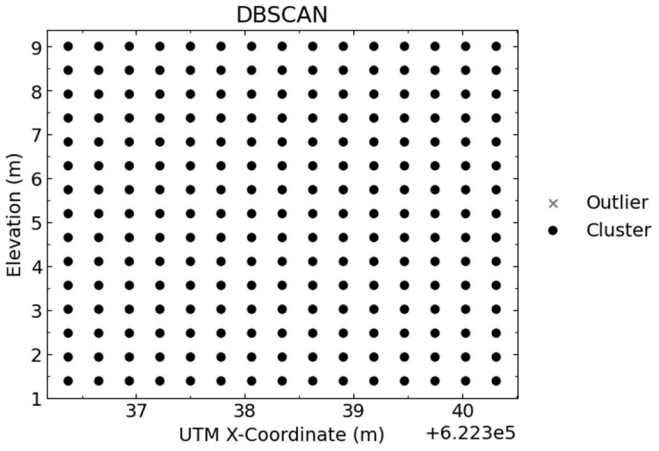






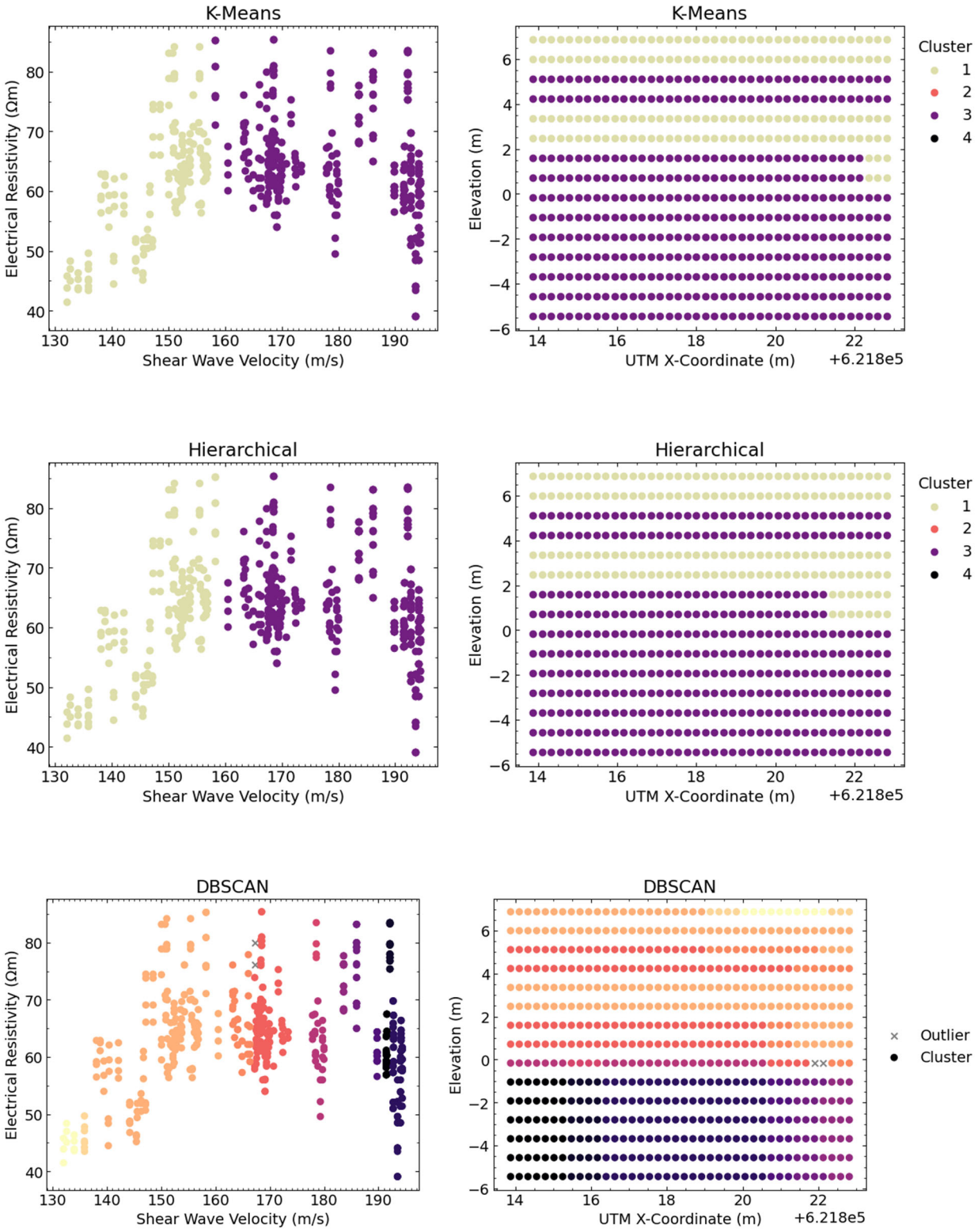






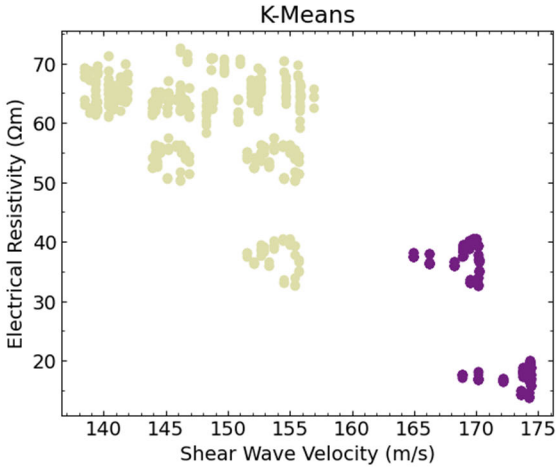
Fit: All_Dry

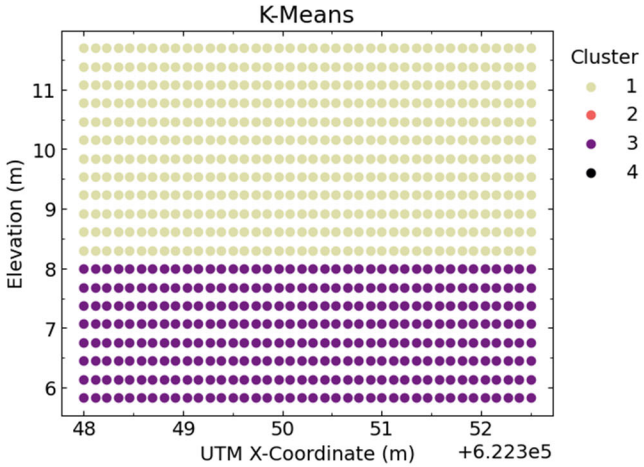
Profile: sc_site2

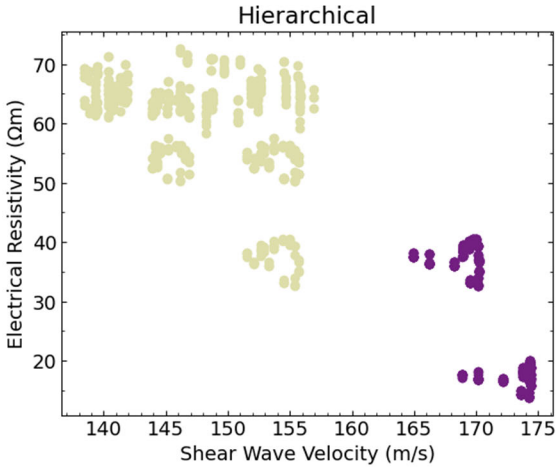


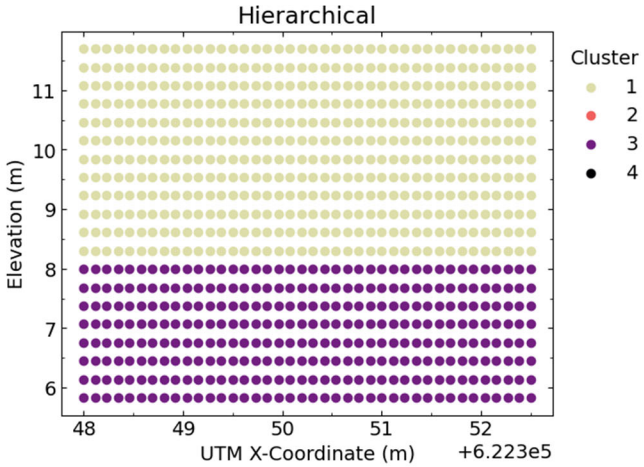
Fit: All_Dry

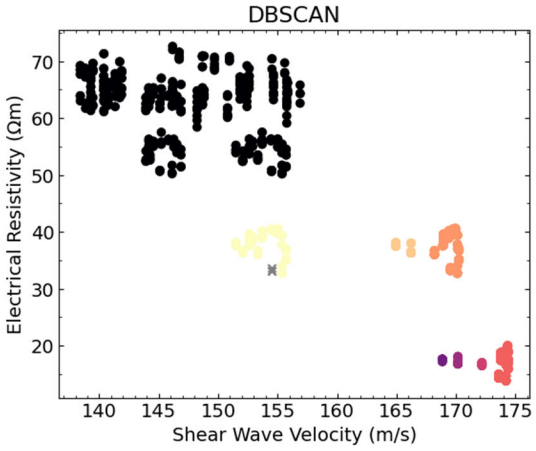
Profile: sc_site3

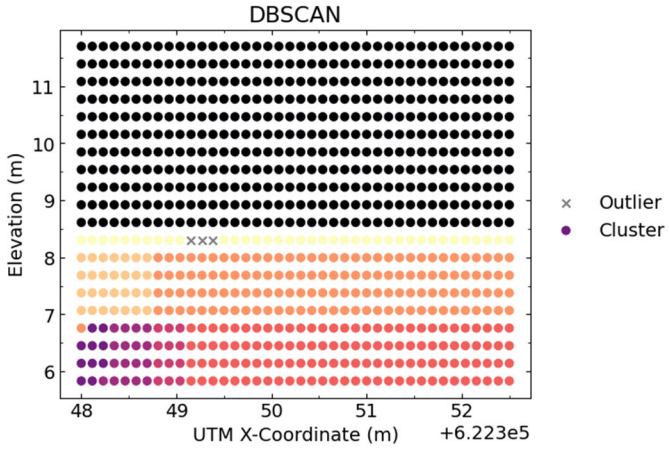






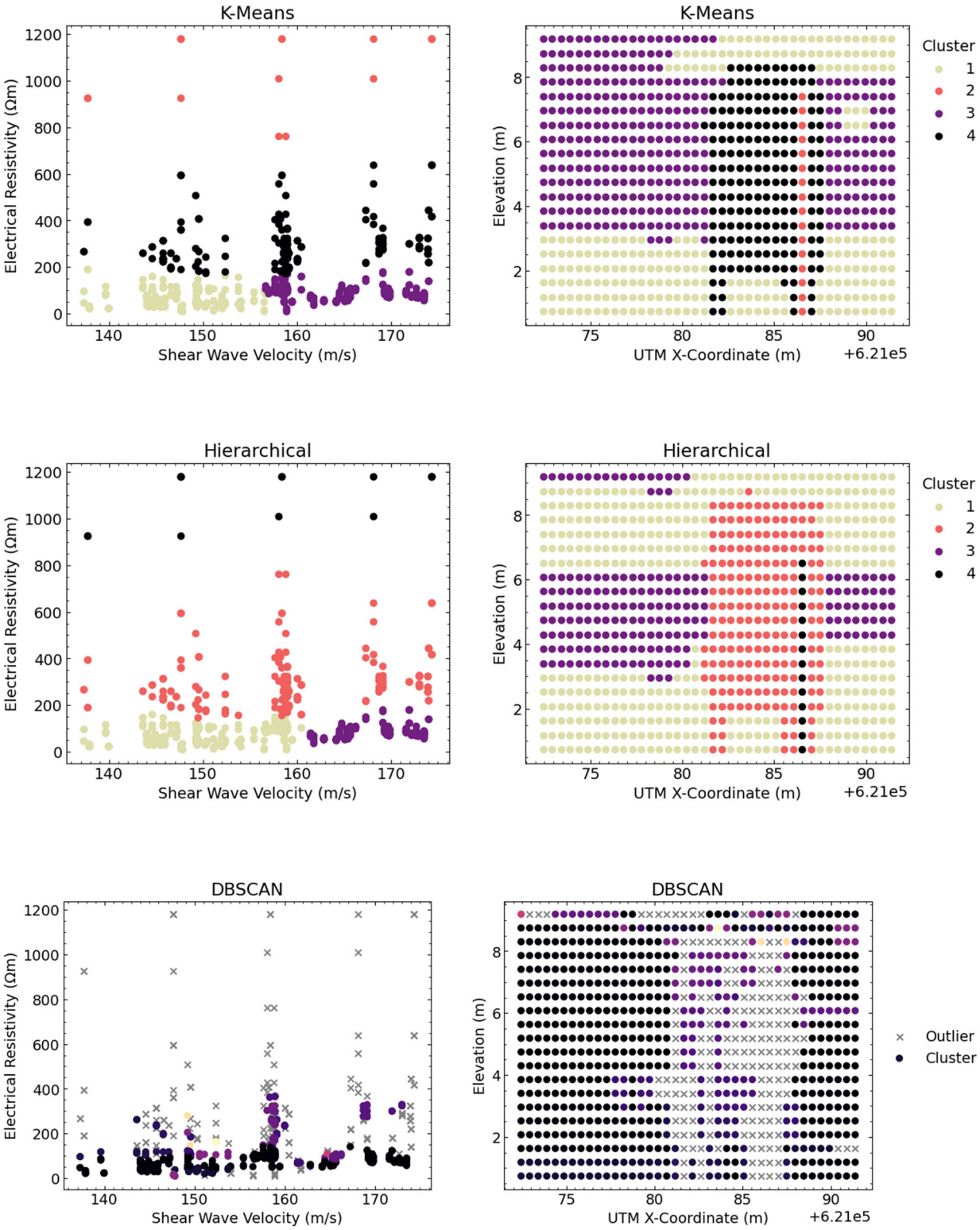






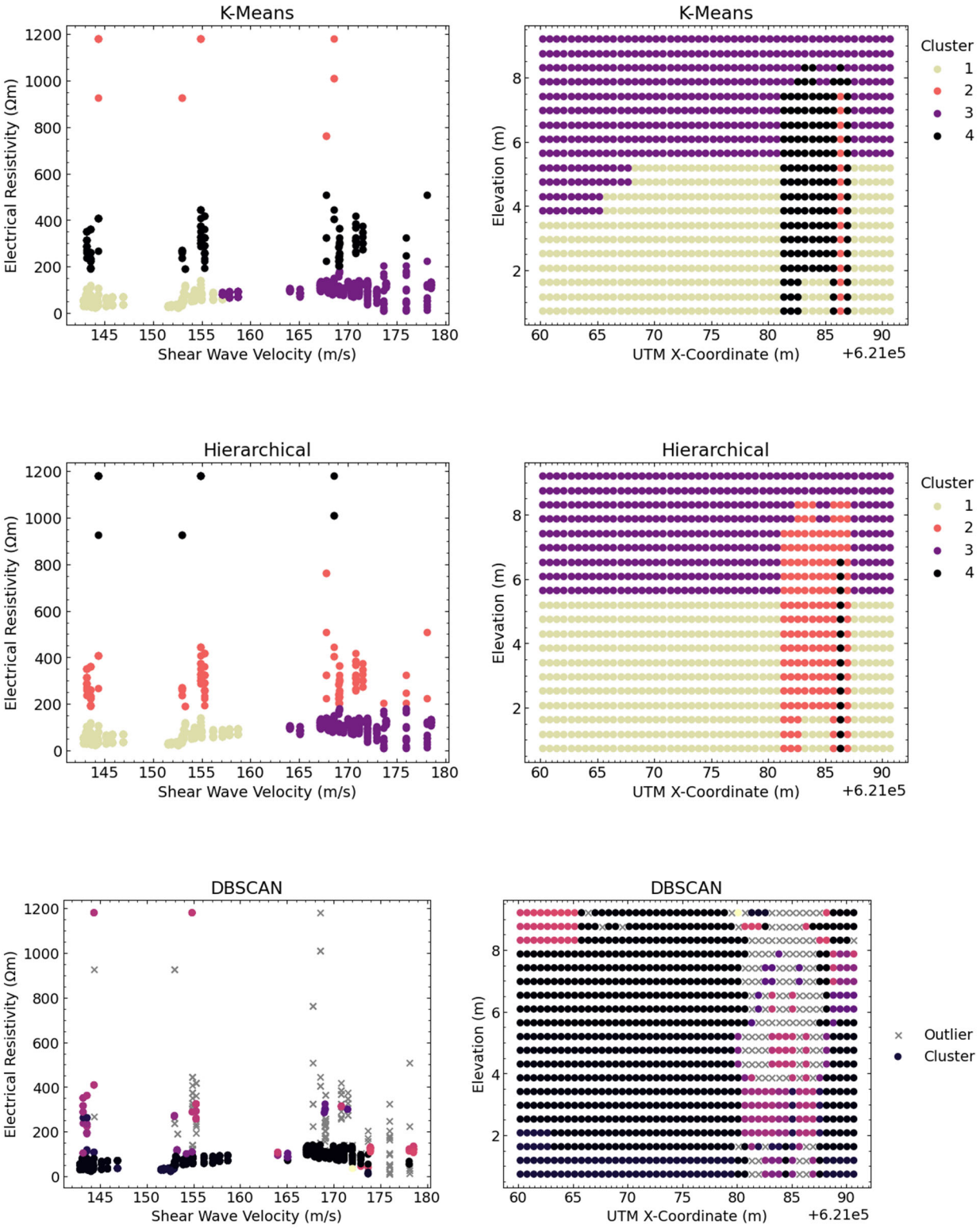
Fit: All_Dry

Profile: sc3_site1



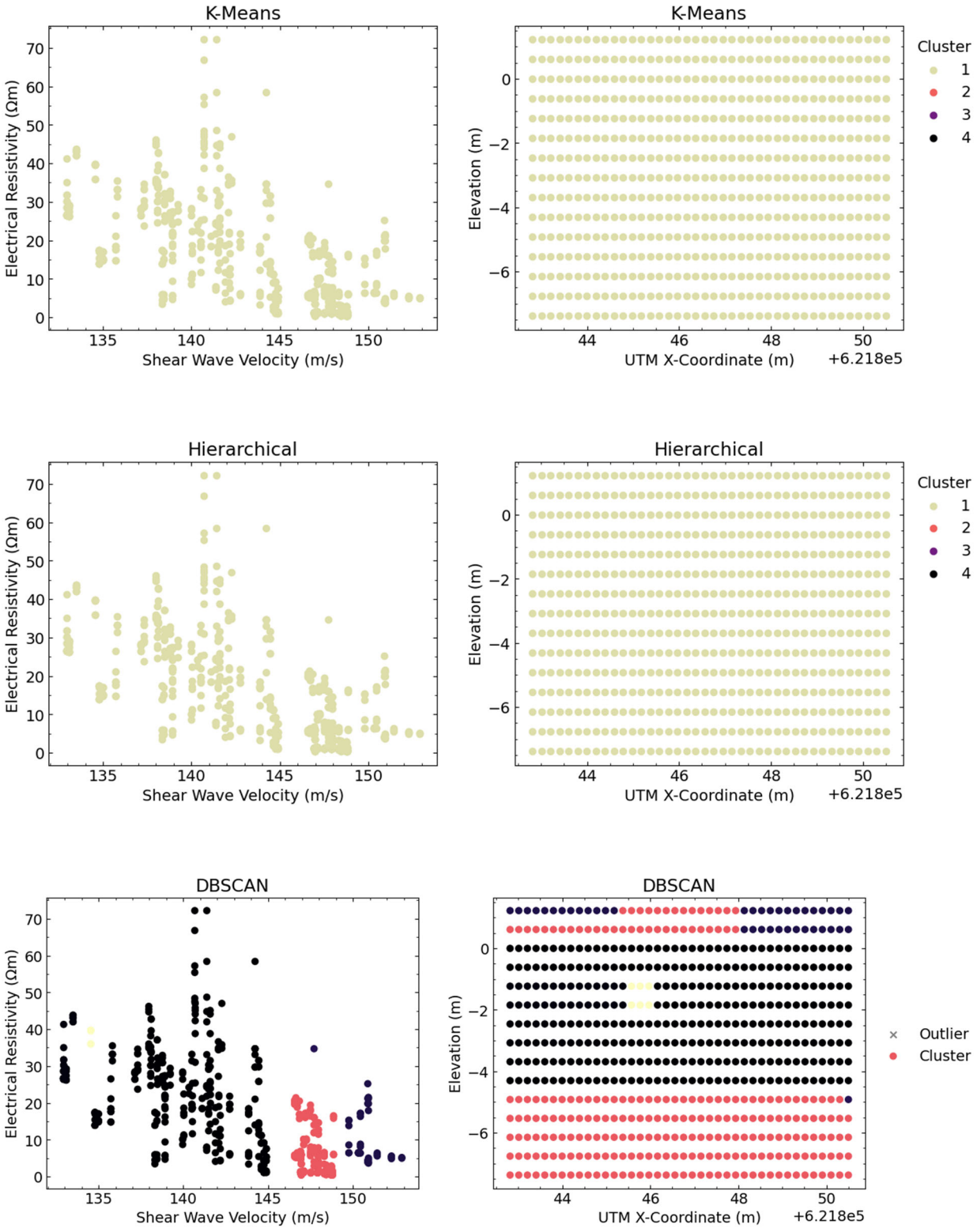
Fit: All_Dry

Profile: sc5_site1



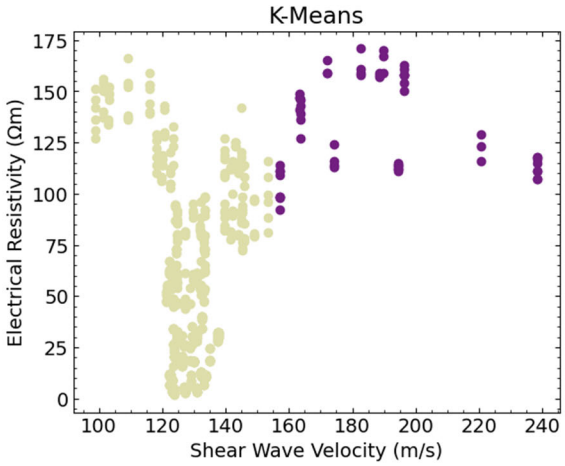
Fit: All_Dry

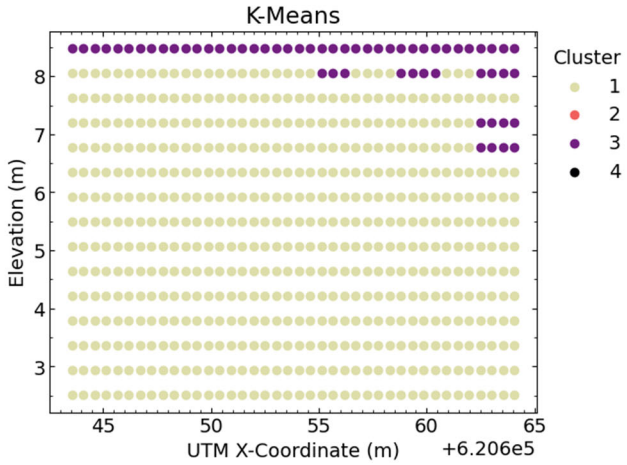
Profile: sdt_site2

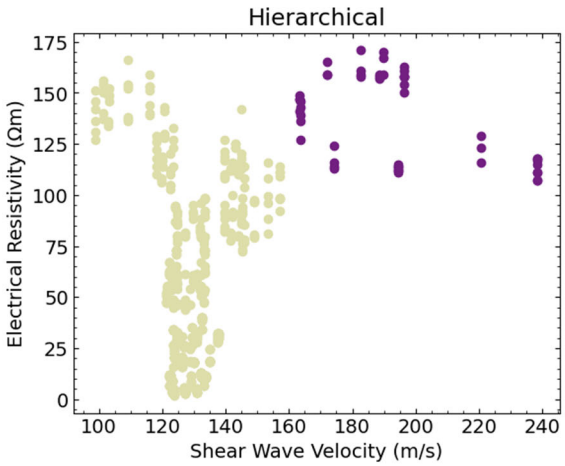


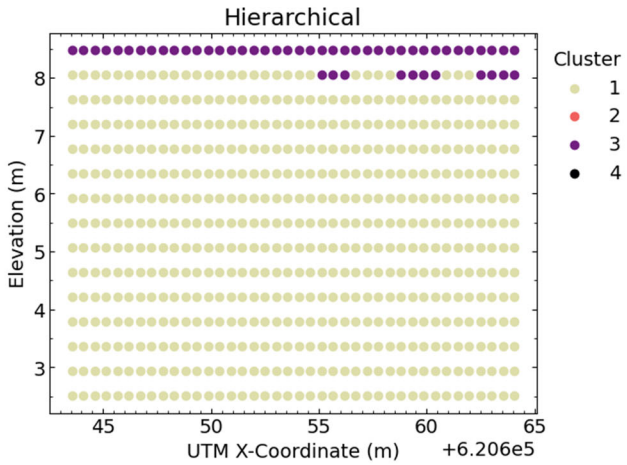
Fit: All_Dry

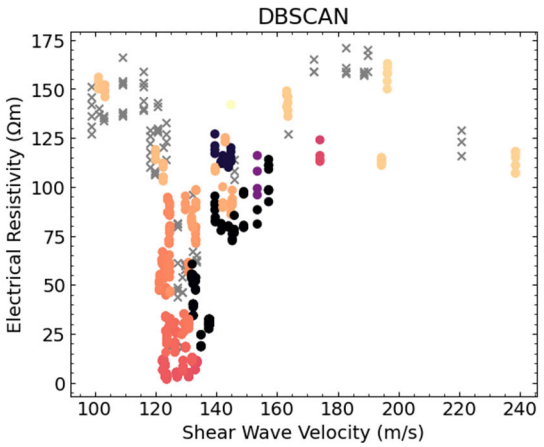
Profile: swt_site1

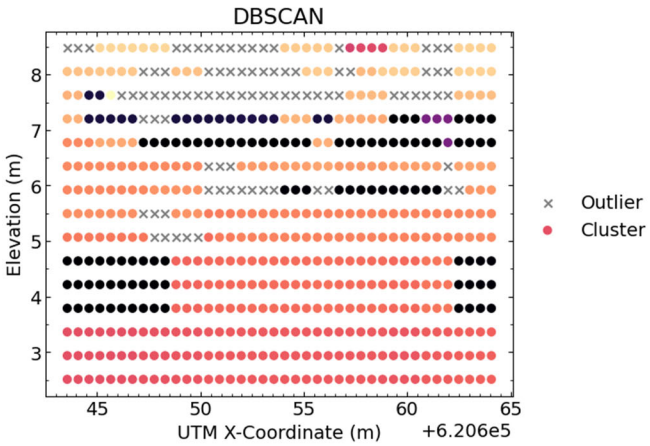






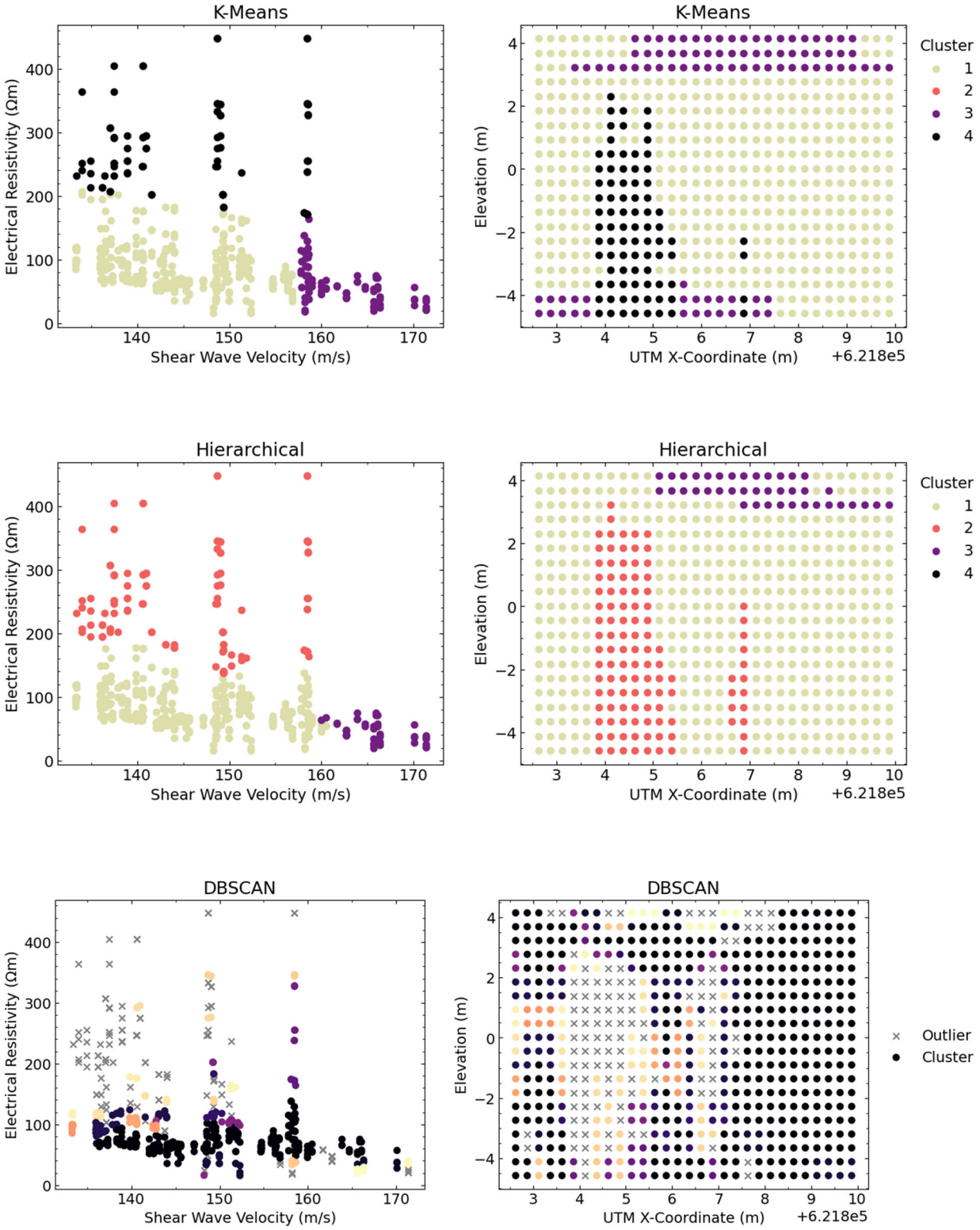






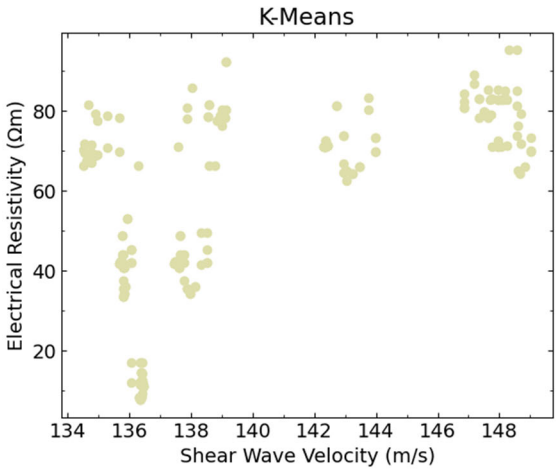
Fit: All_Dry

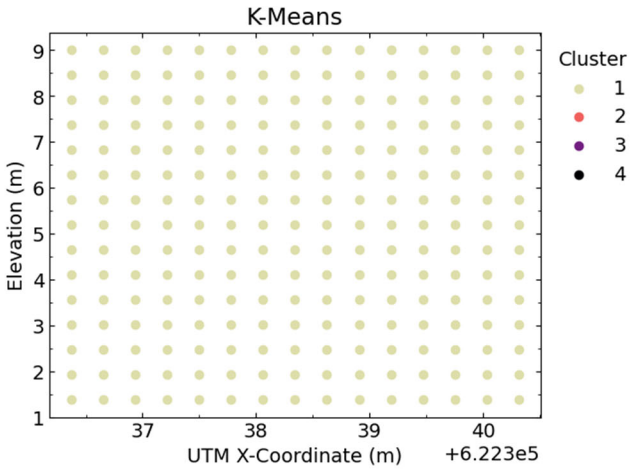
Profile: swt_site2

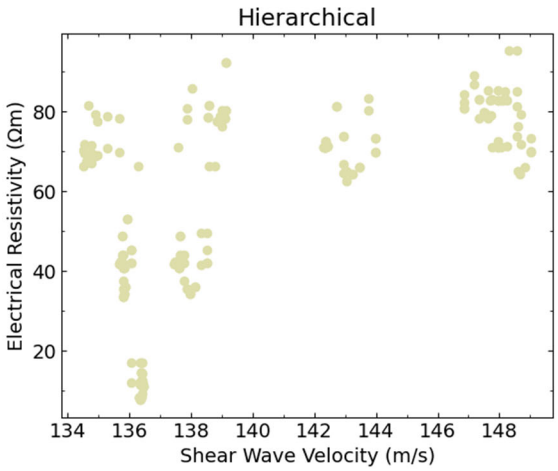


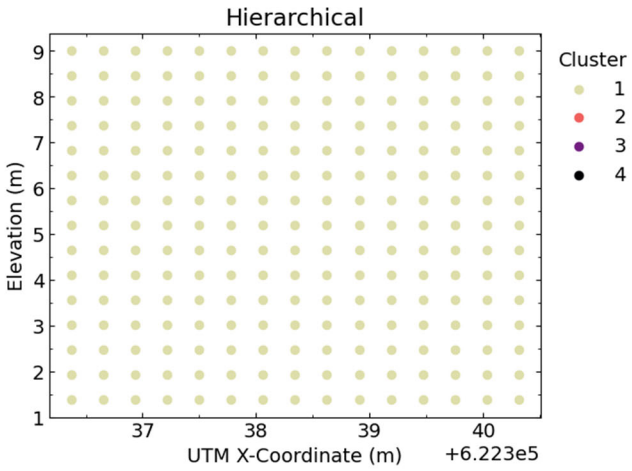
Fit: All_Dry

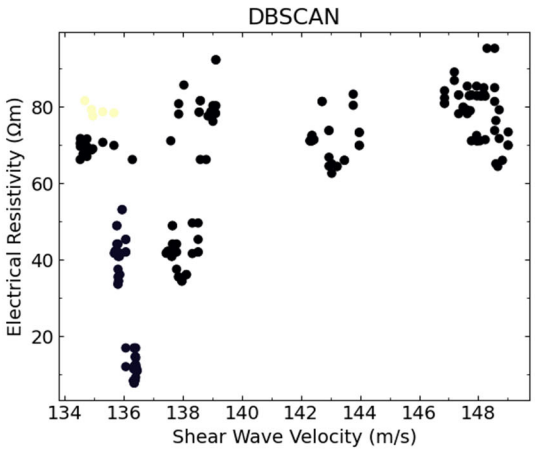
Profile: swt_site3

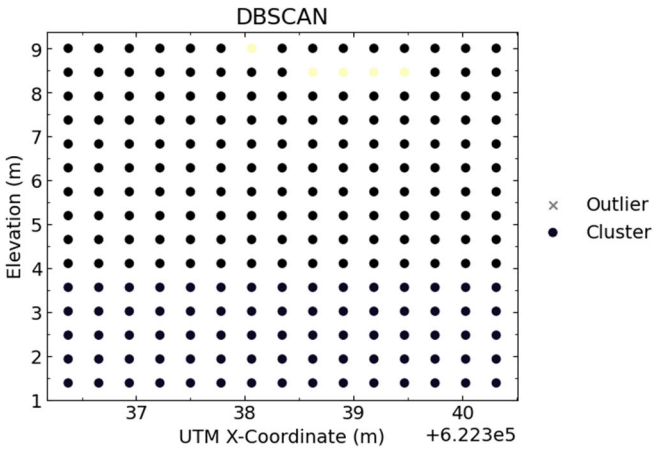






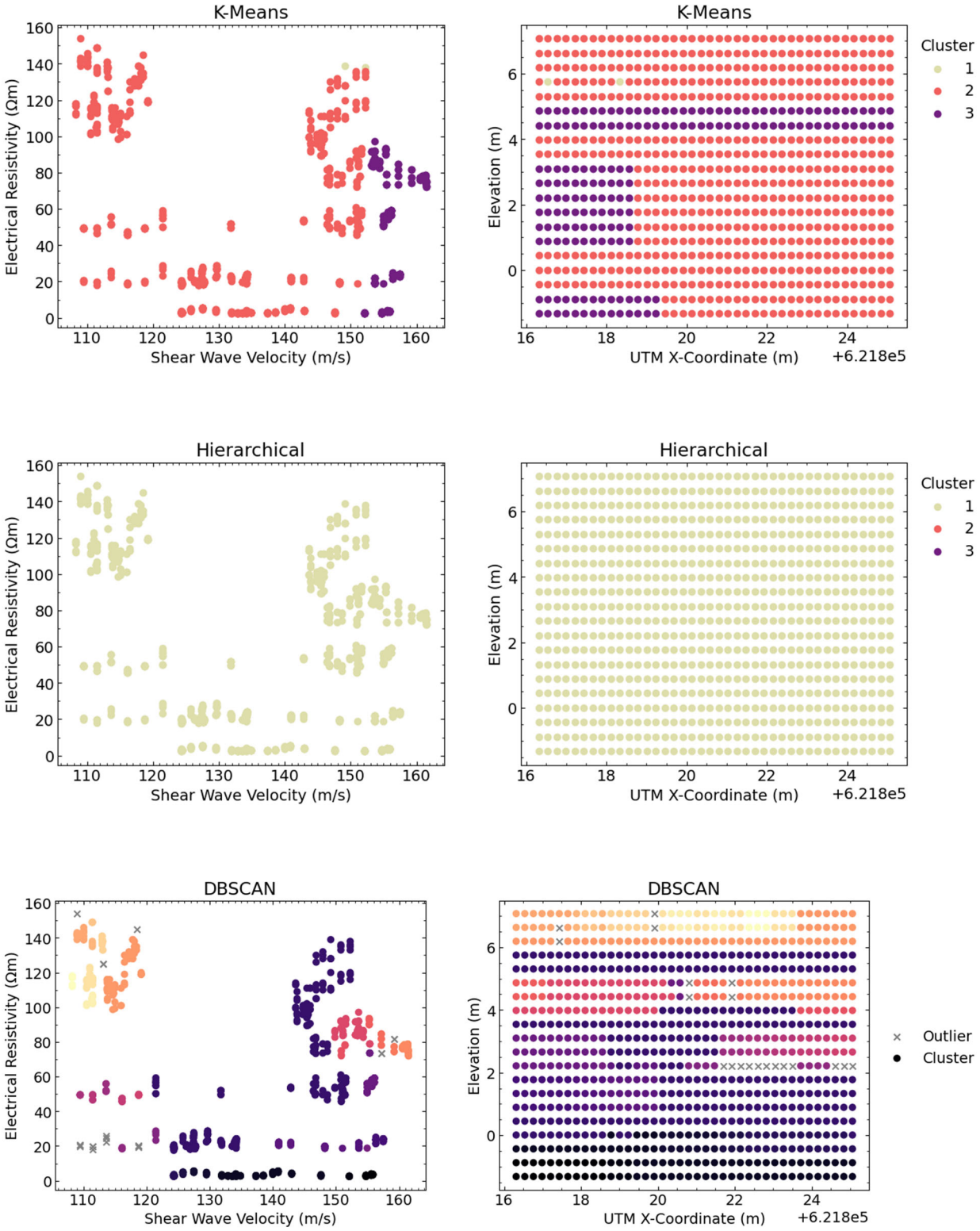






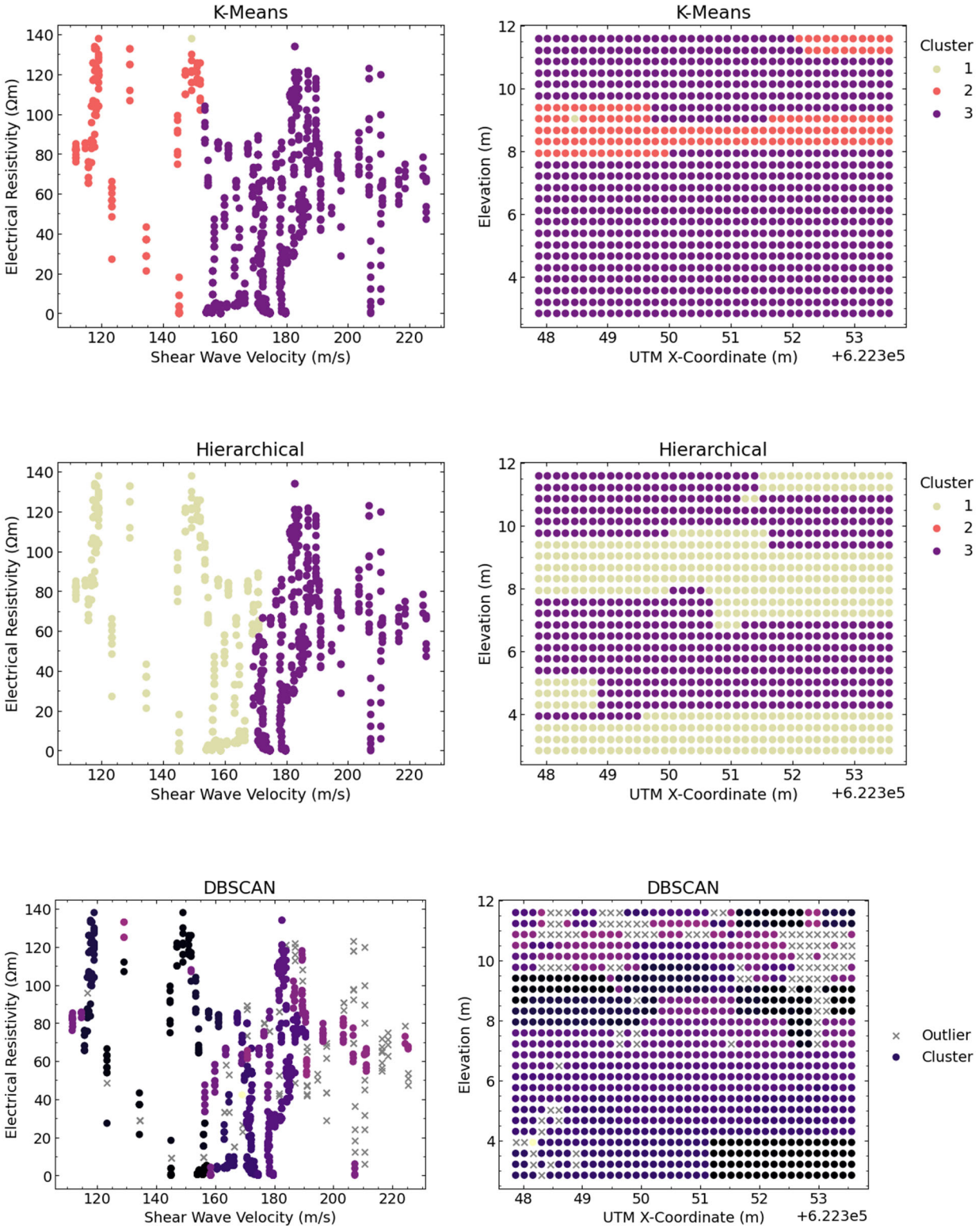
Fit: All_Wet

Profile: fc_site2



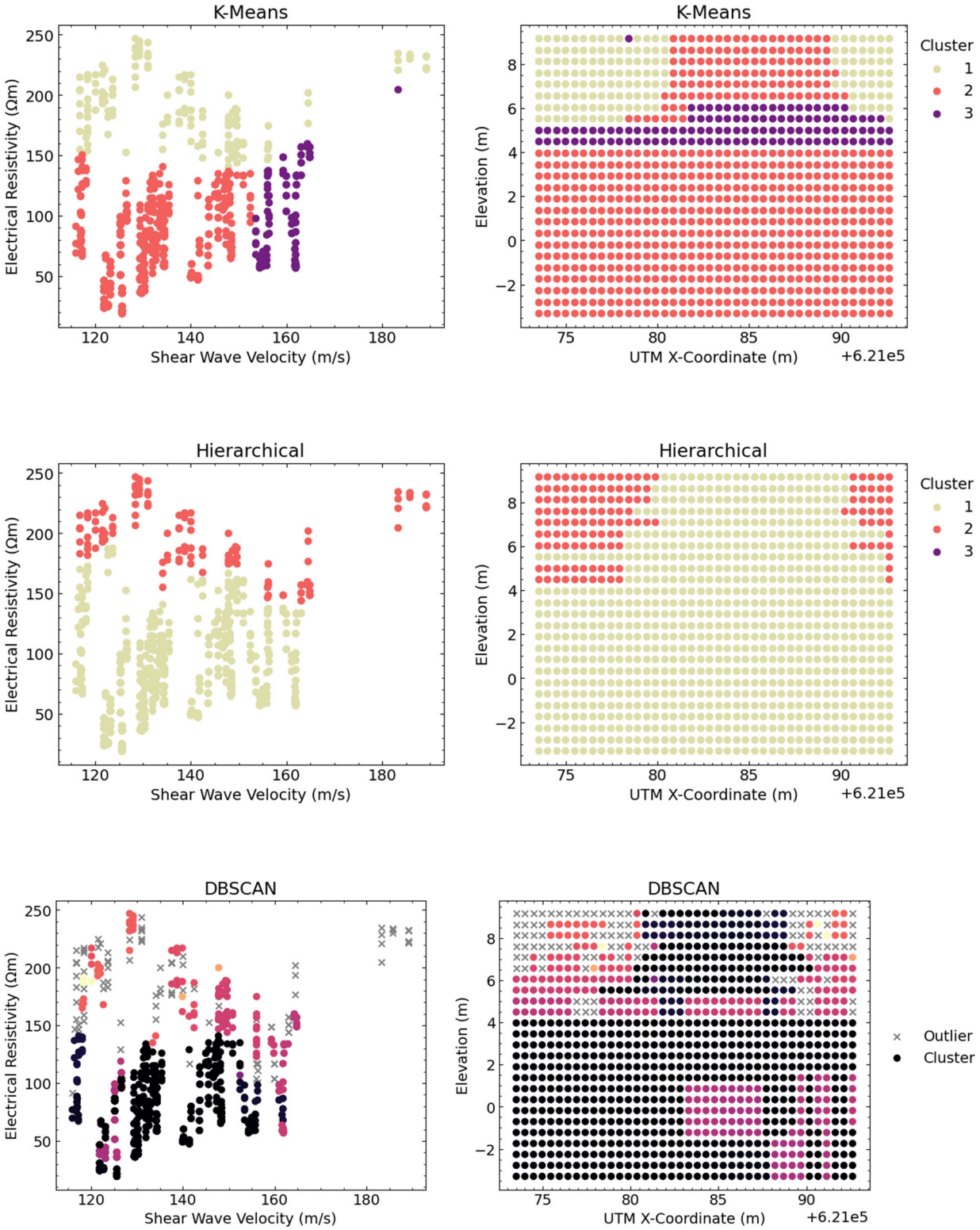
Fit: All_Wet

Profile: fc_site3



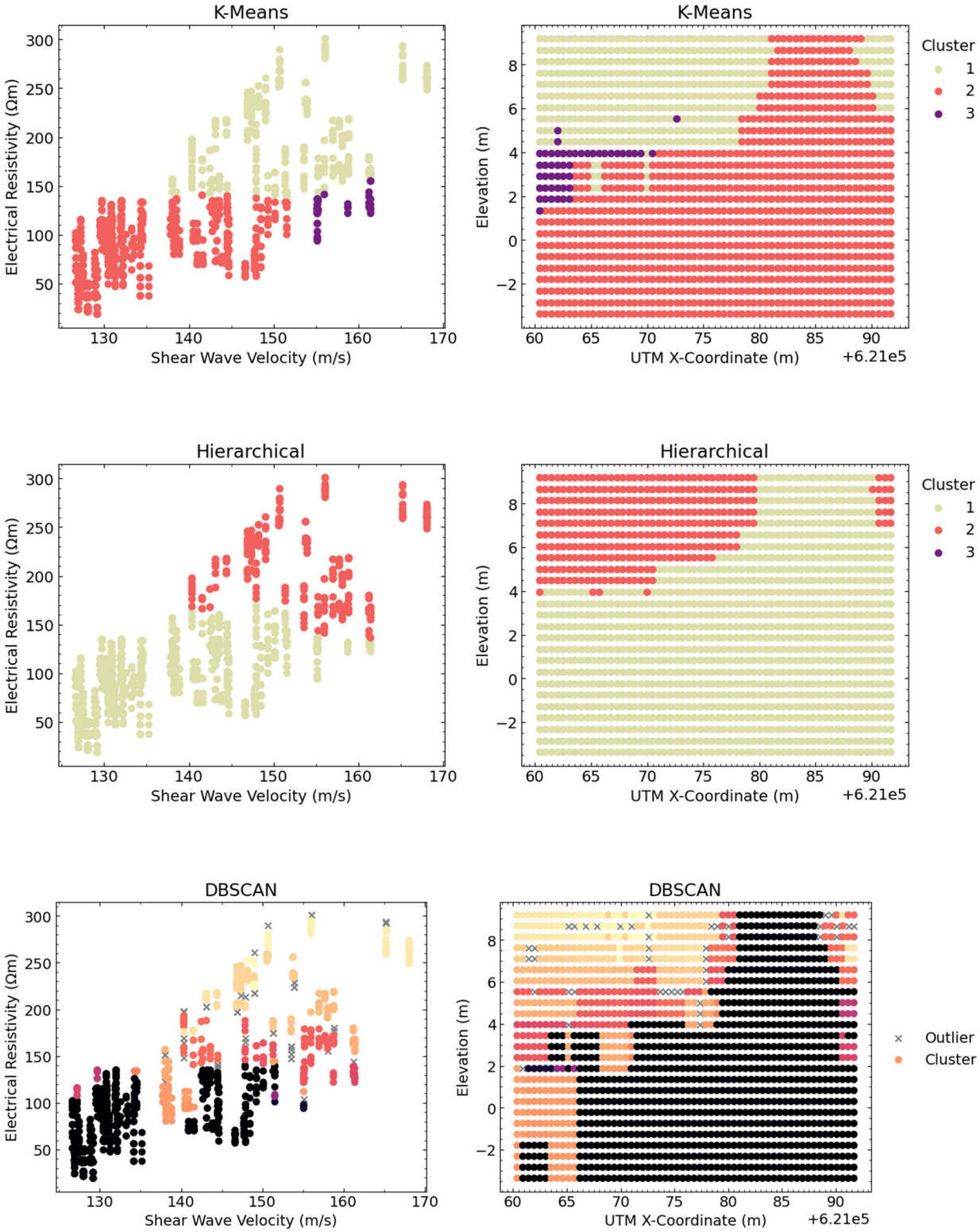
Fit: All_Wet

Profile: fc3_site1



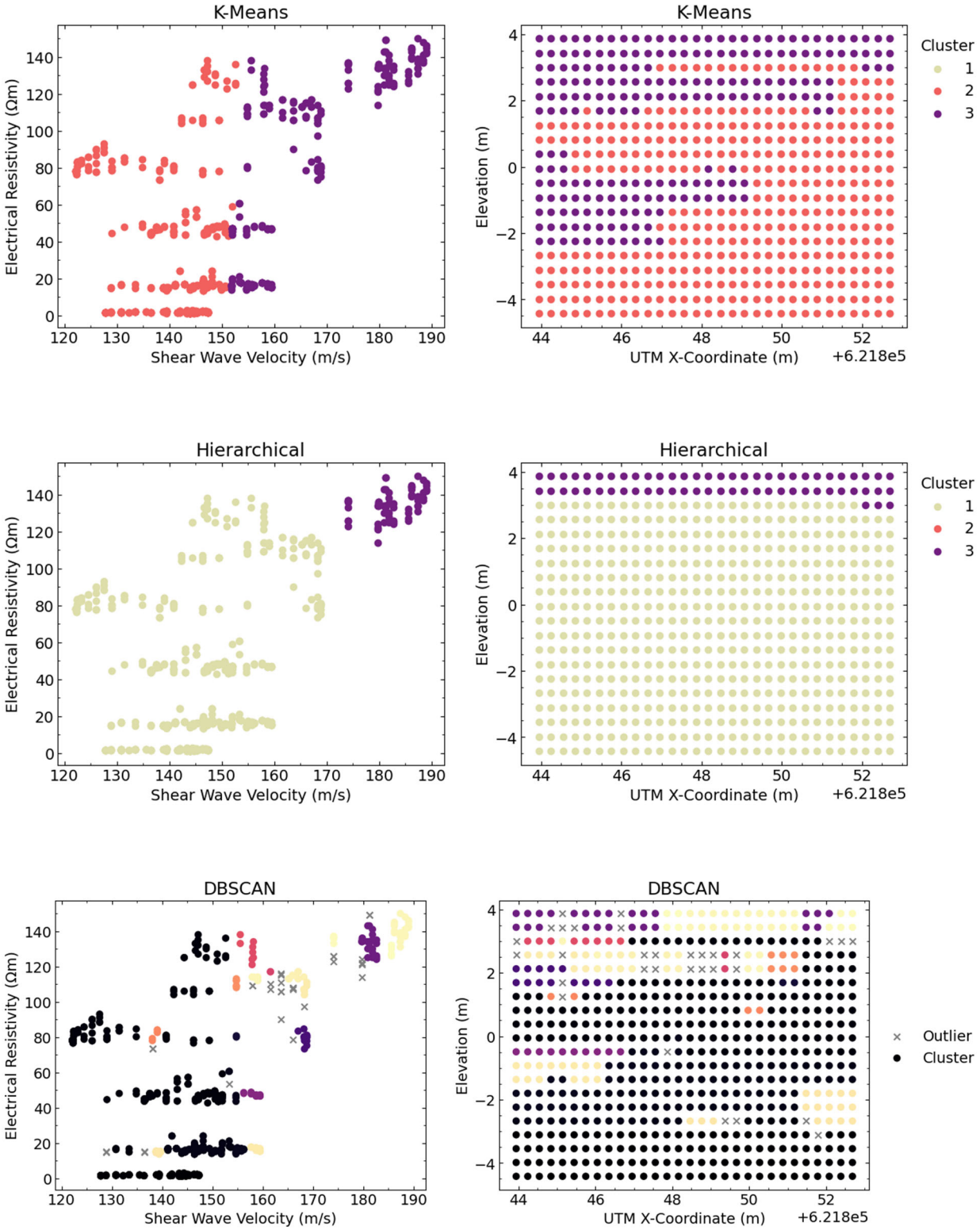
Fit: All_Wet

Profile: fc5_site1



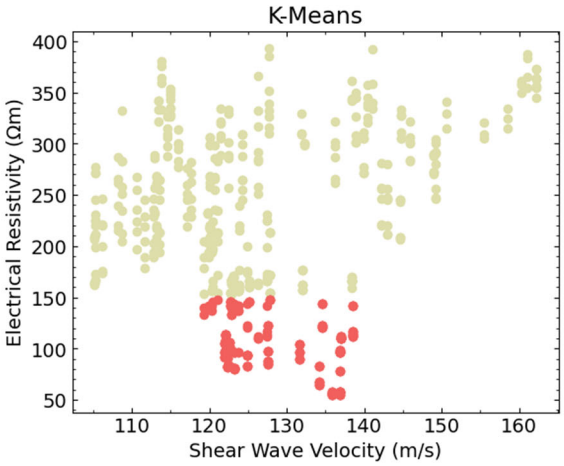
Fit: All_Wet

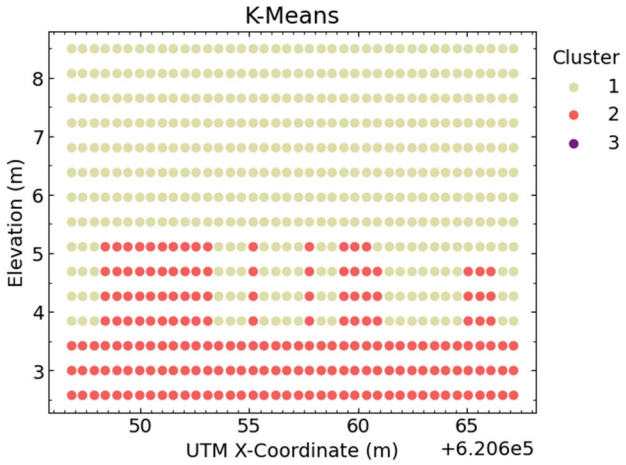
Profile: fdt_site2

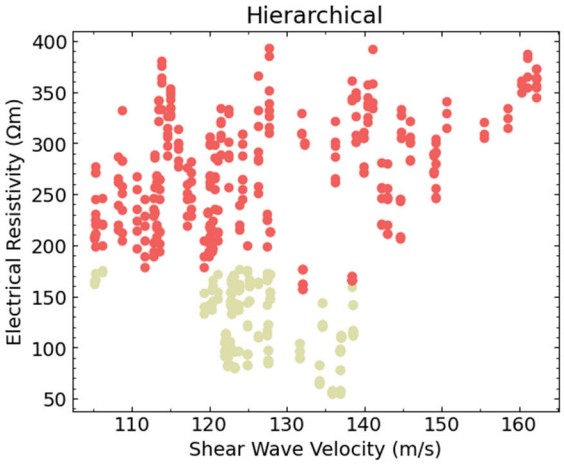


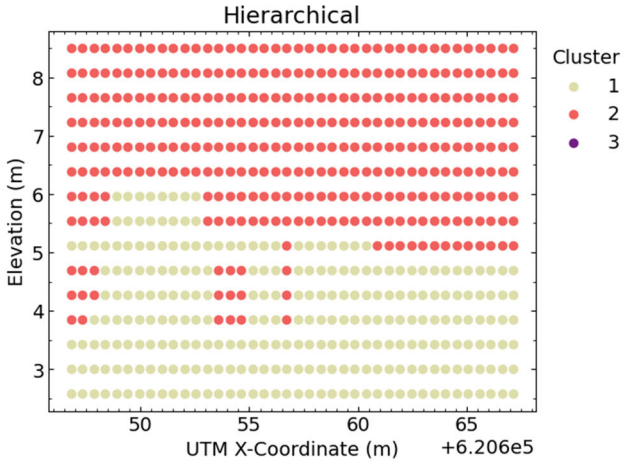
Fit: All_Wet

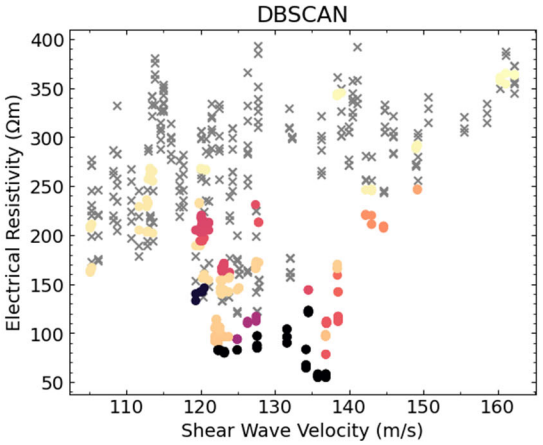
Profile: fwt_site1

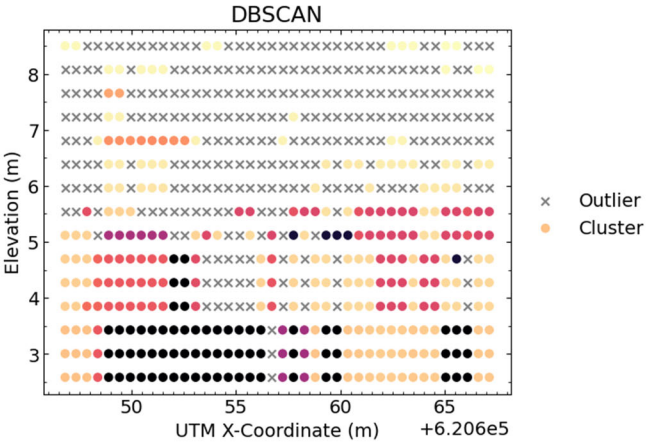






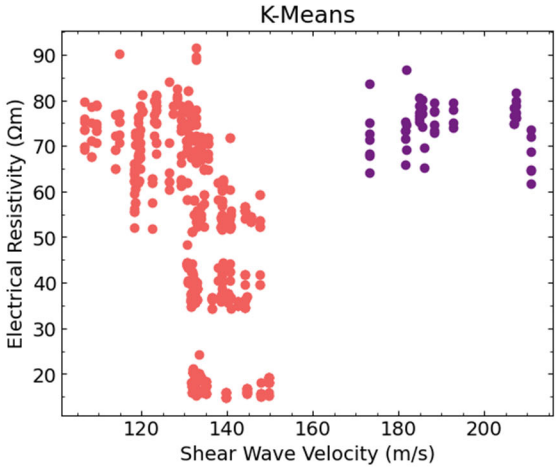


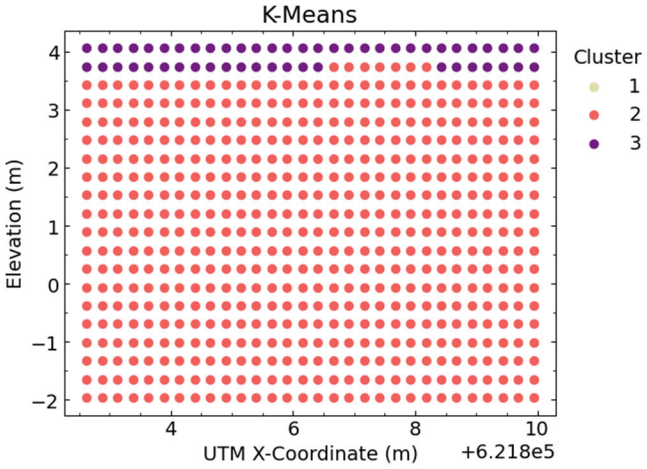


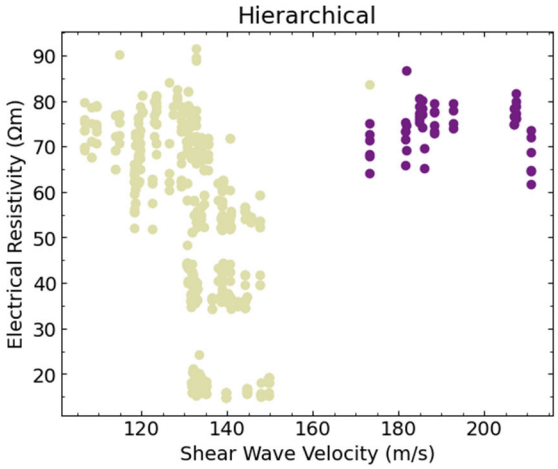


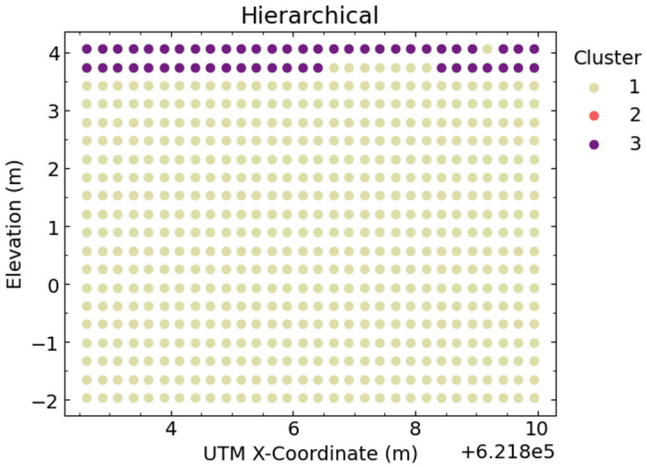
Fit: All_Wet

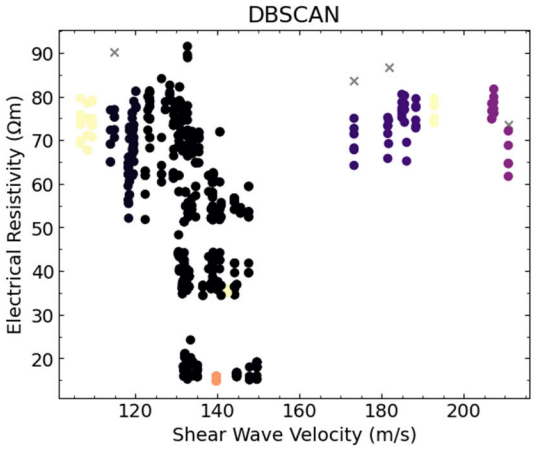
Profile: fwt_site2

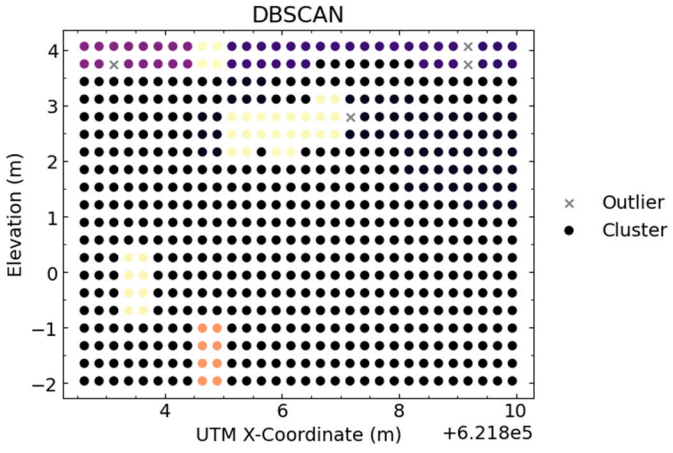






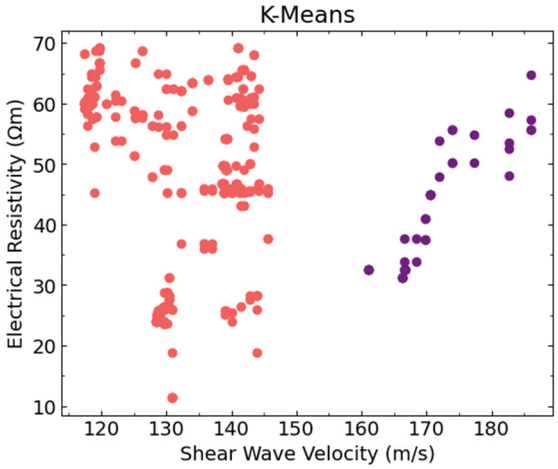


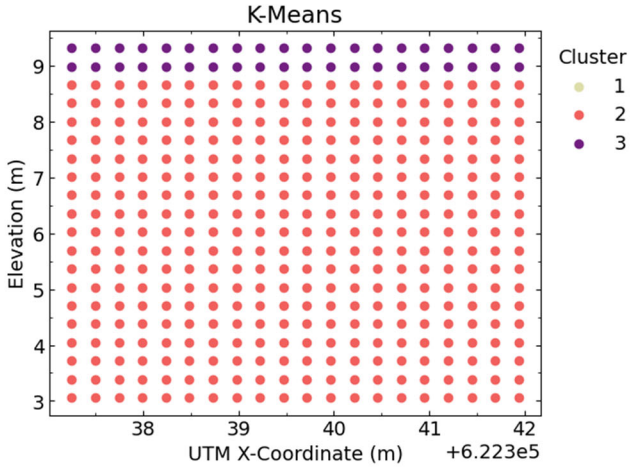


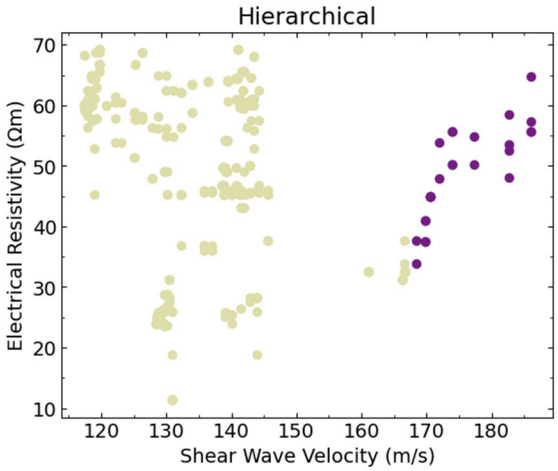


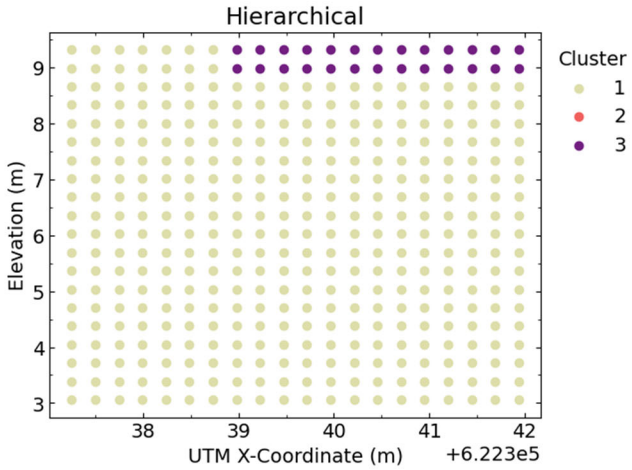
Fit: All_Wet

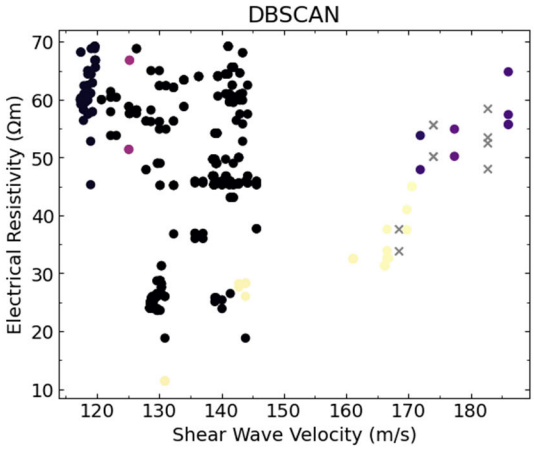
Profile: fwt_site3

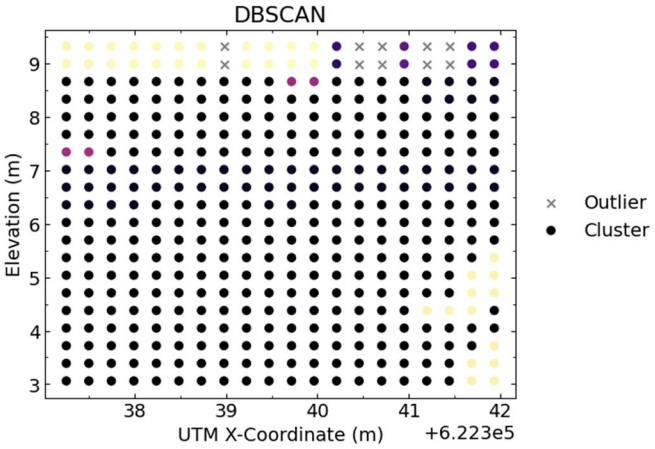






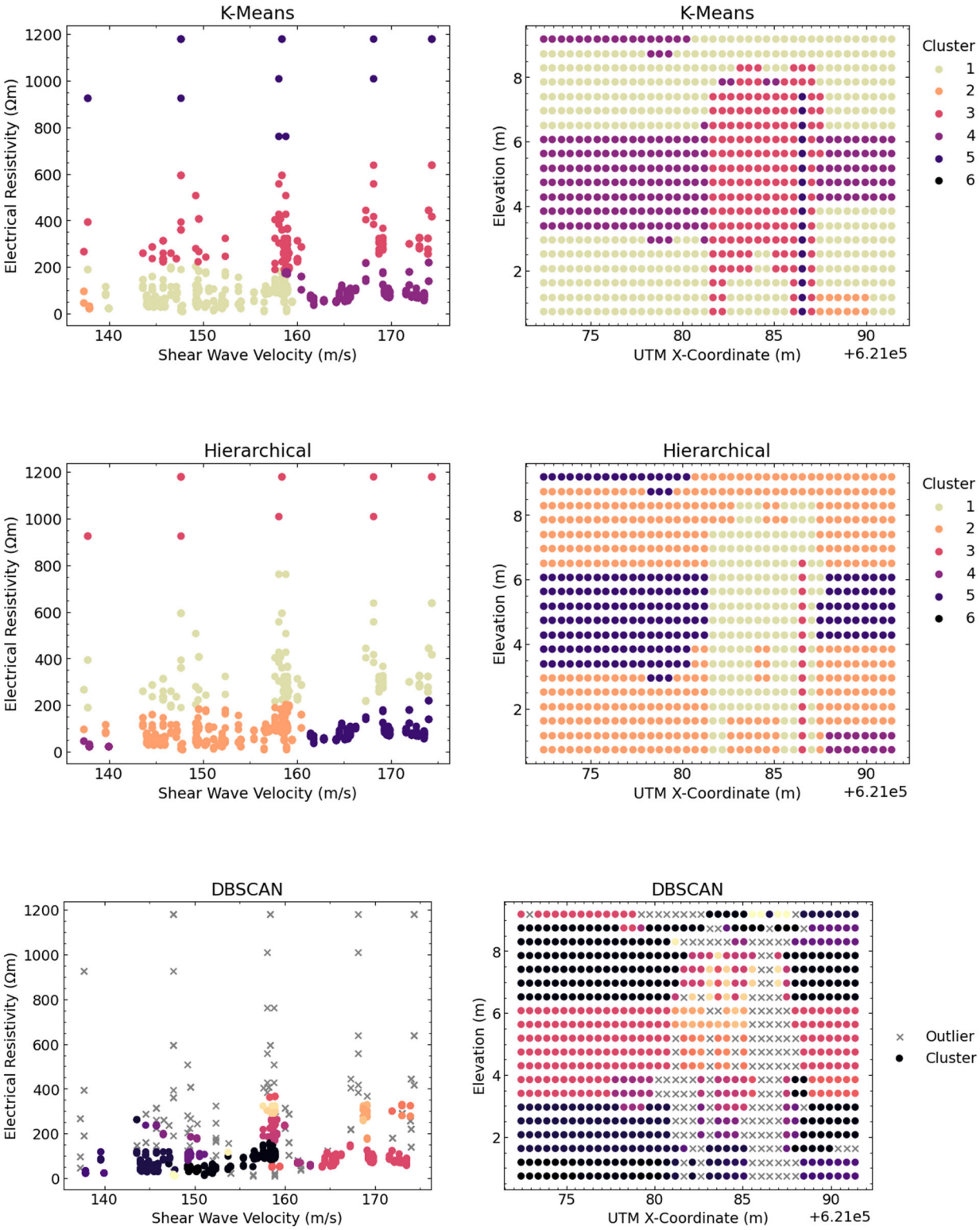






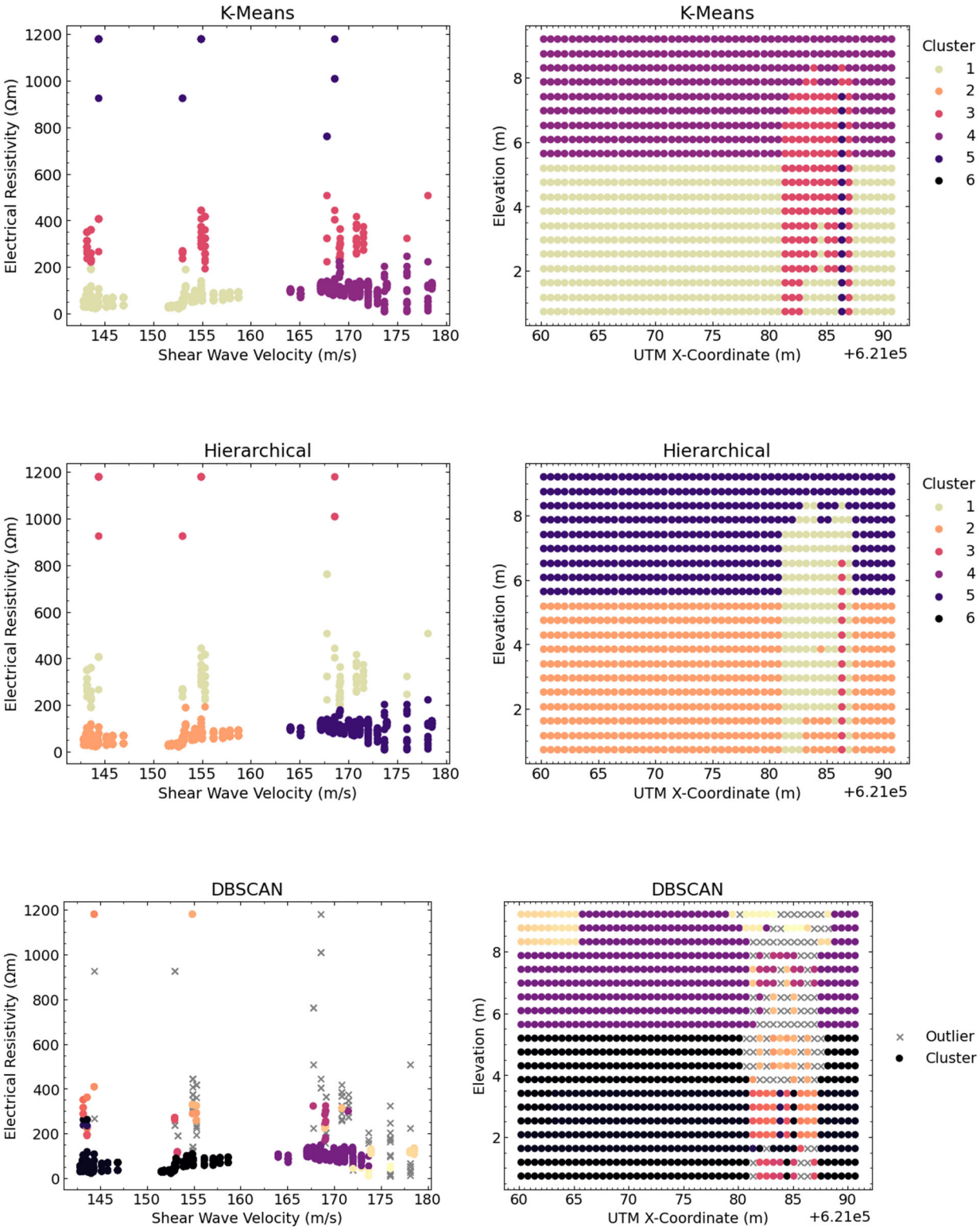
Fit: Dry_Site1

Profile: sc3_site1



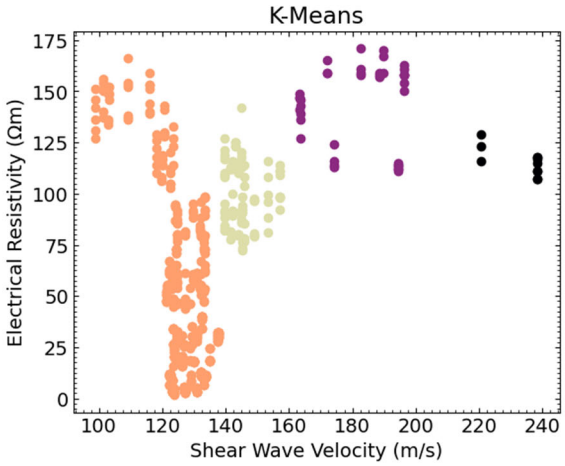
Fit: Dry_Site1

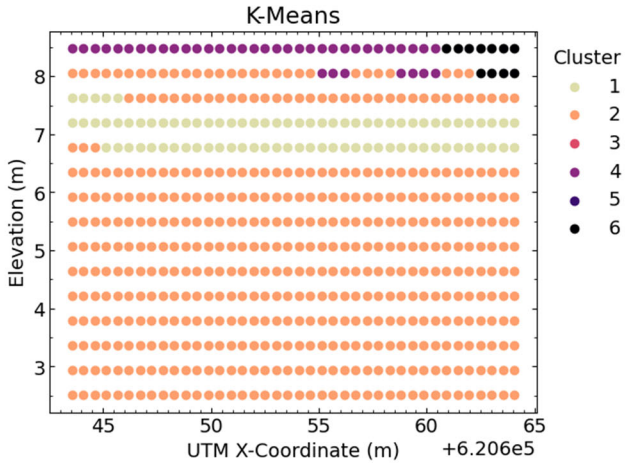
Profile: sc5_site1

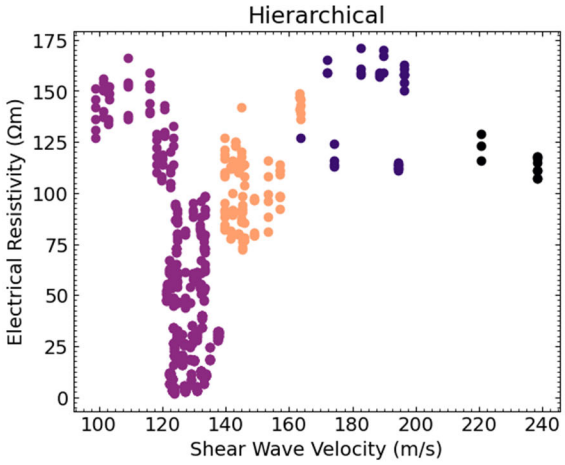


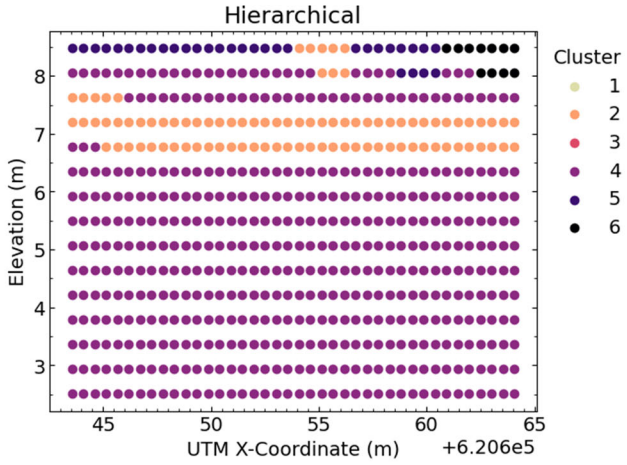
Fit: Dry_Site1

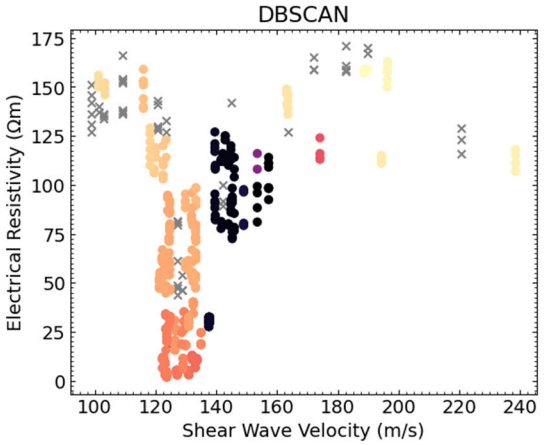
Profile: swt_site1

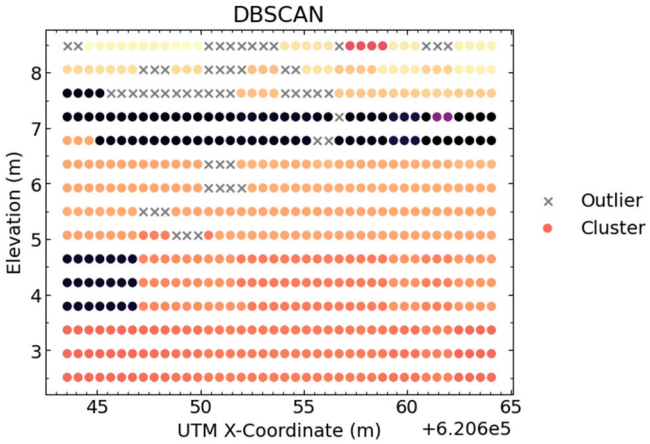






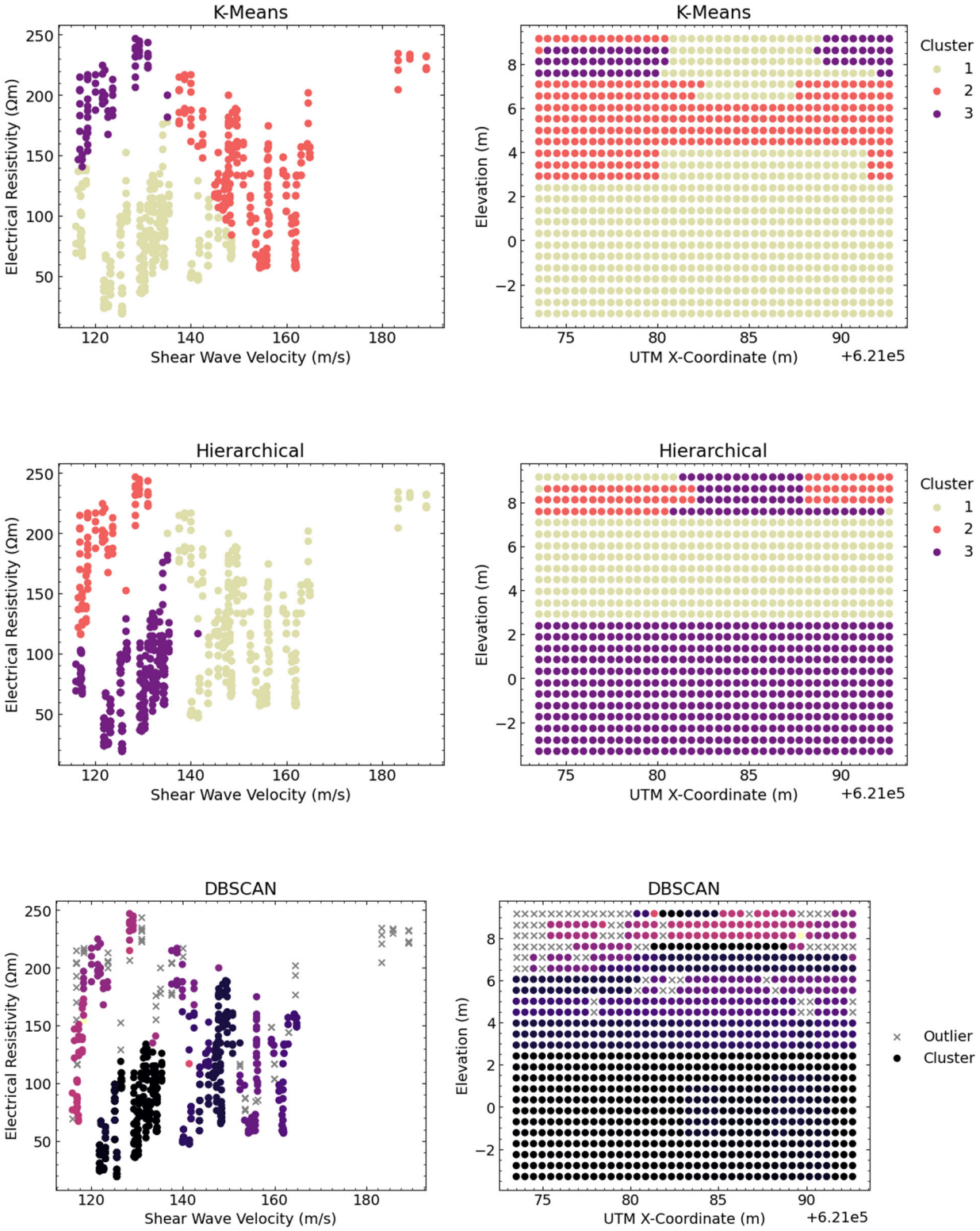






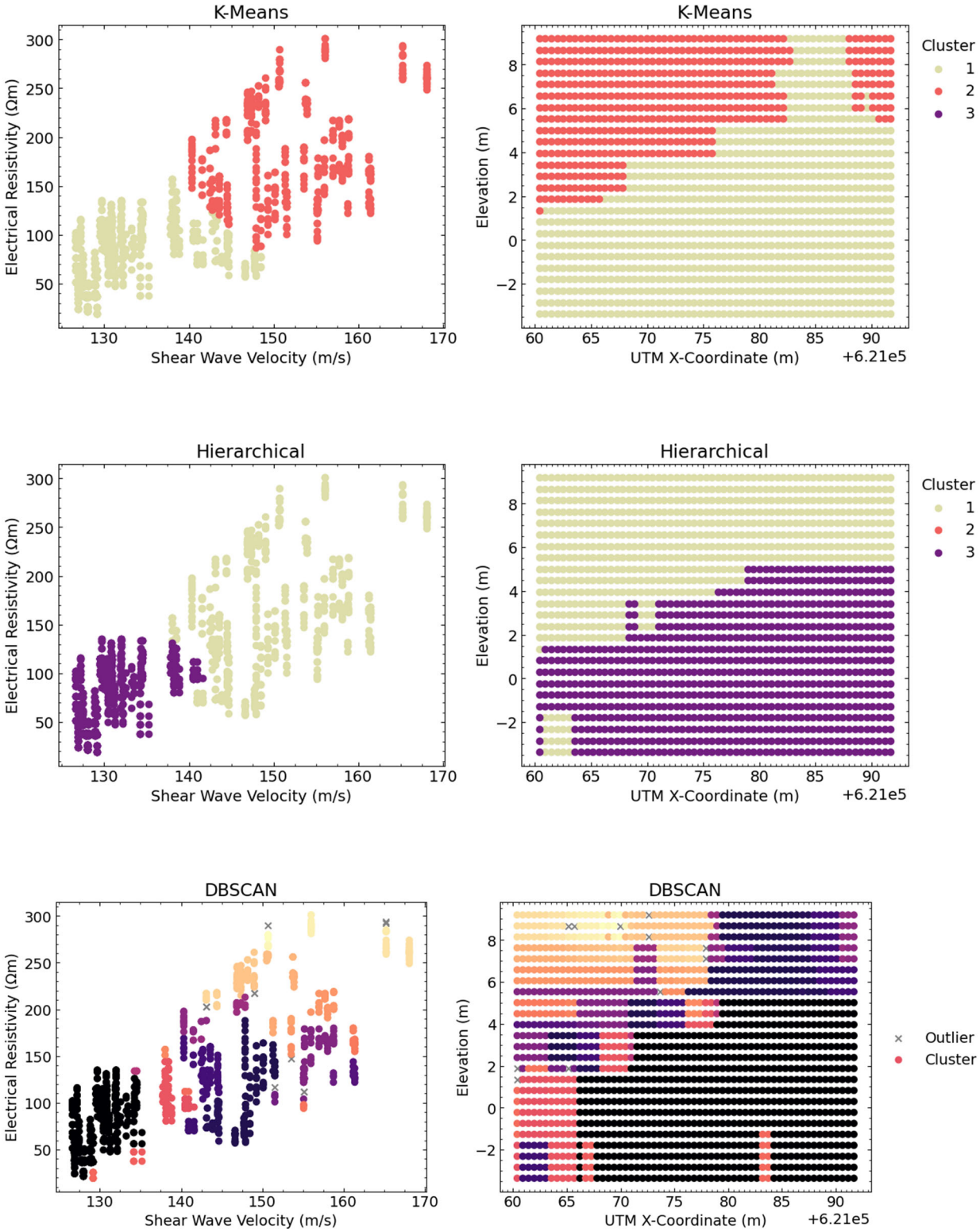
Fit: Wet_Site1

Profile: fc3_site1



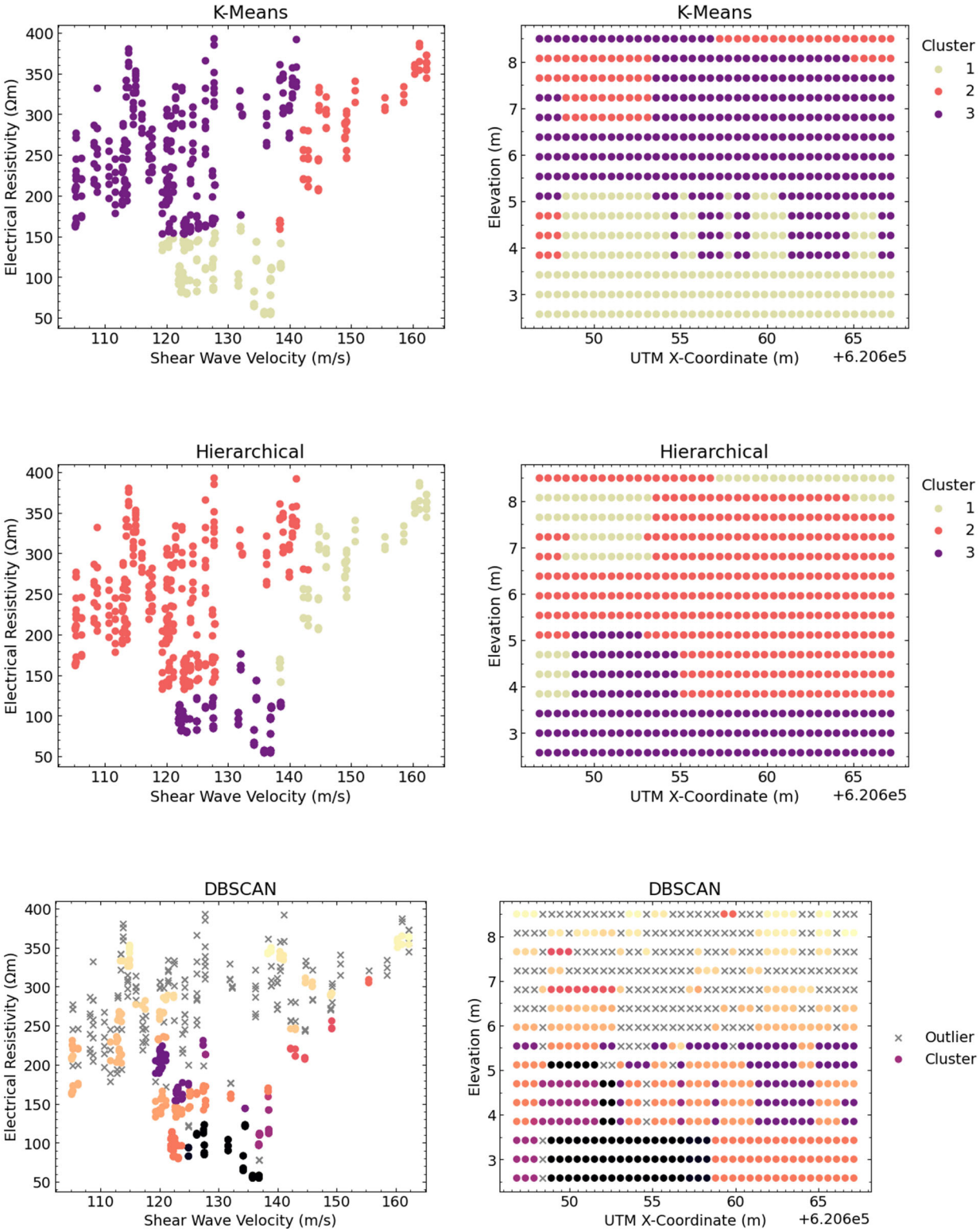
Fit: Wet_Site1

Profile: fc5_site1



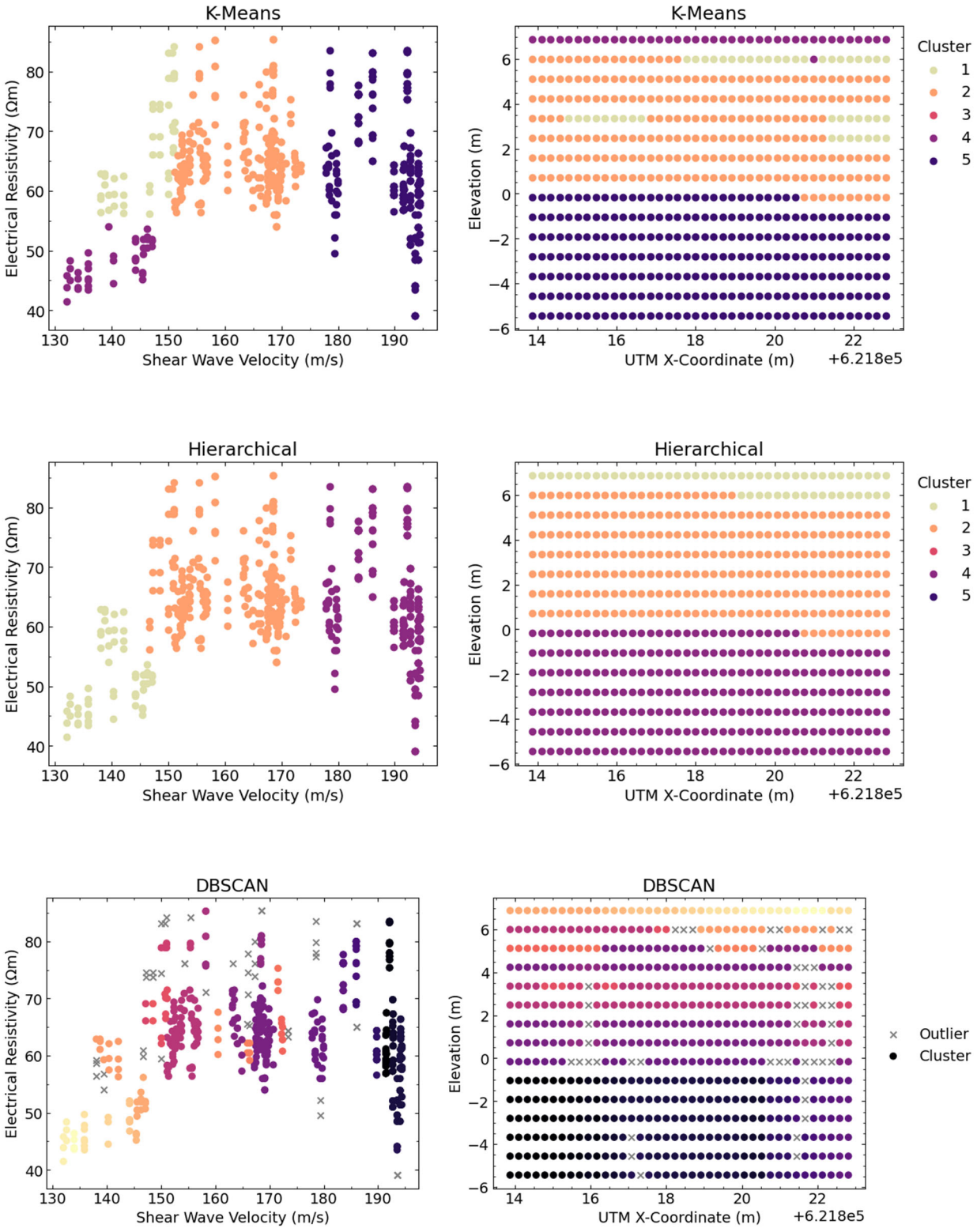
Fit: Wet_Site1

Profile: fwt_site1



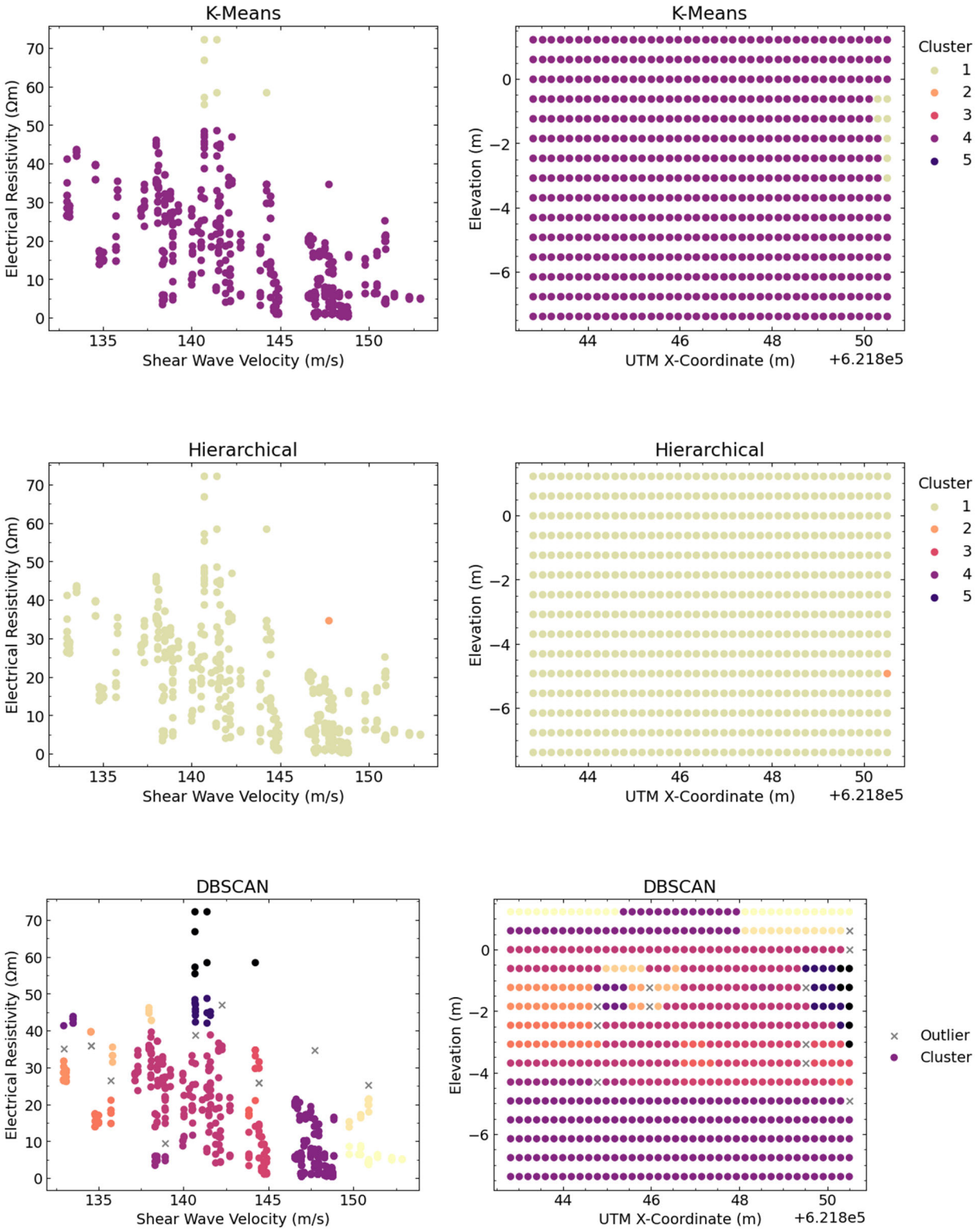
Fit: Dry_Site2

Profile: sc_site2



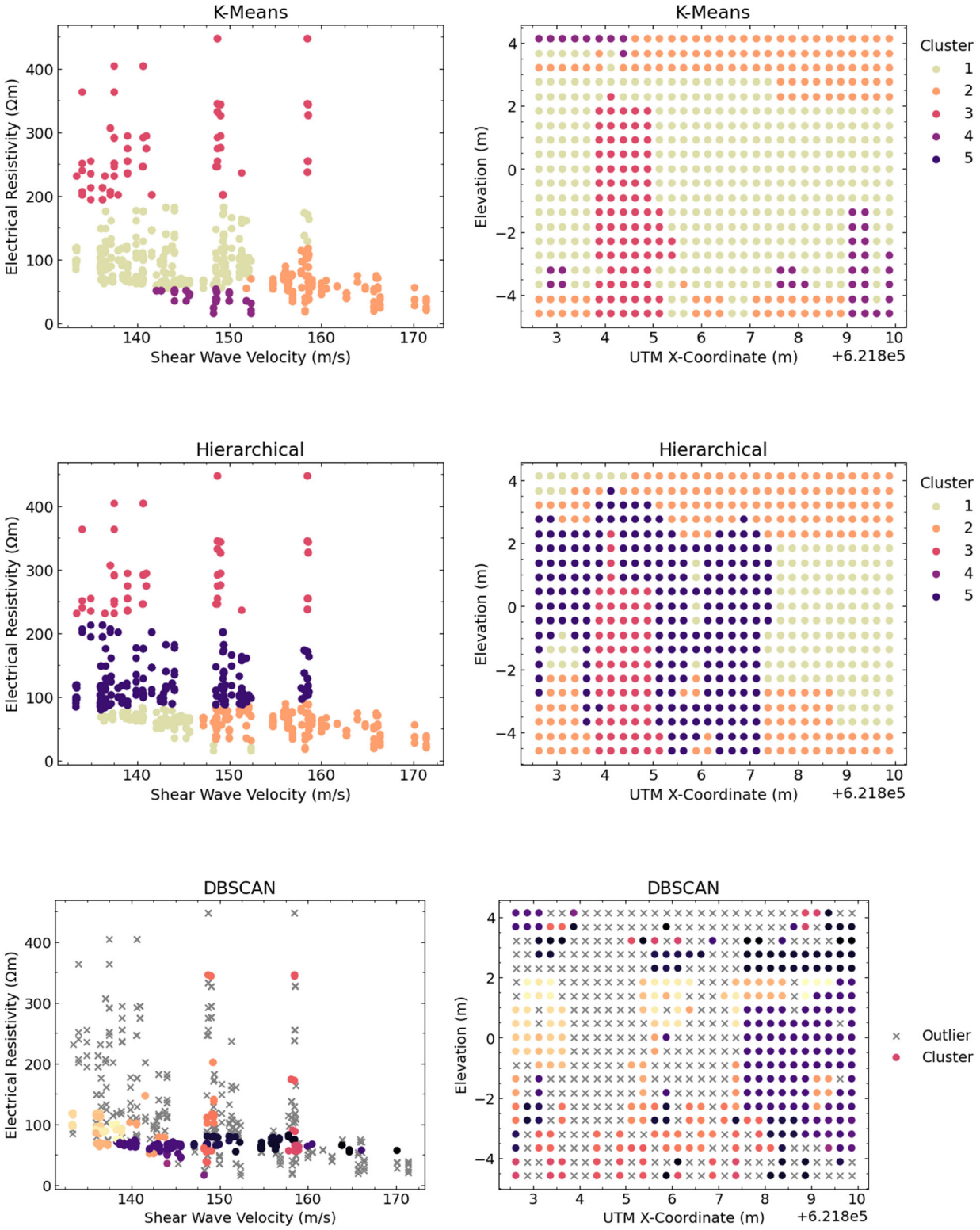
Fit: Dry_Site2

Profile: sdt_site2



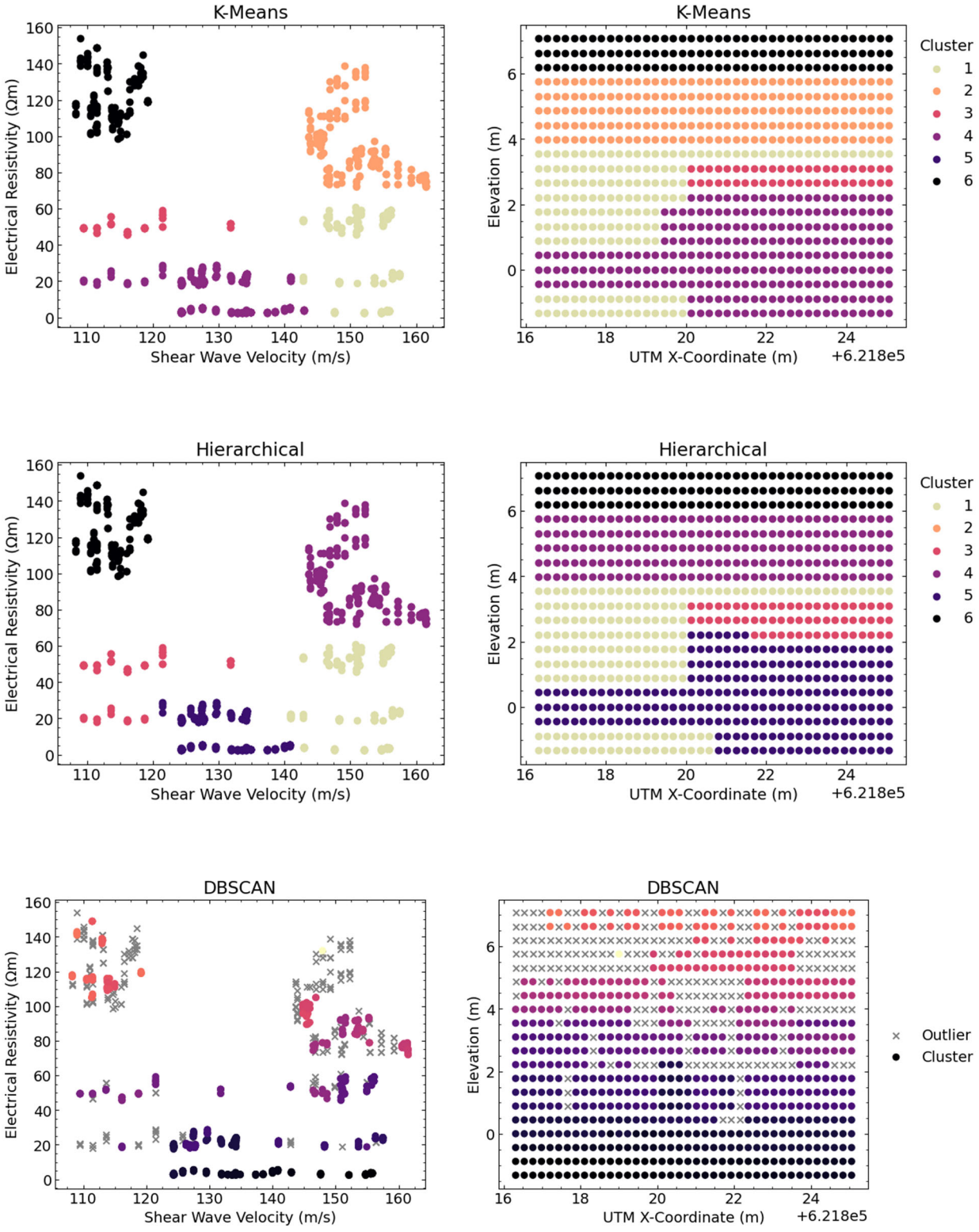
Fit: Dry_Site2

Profile: swt_site2



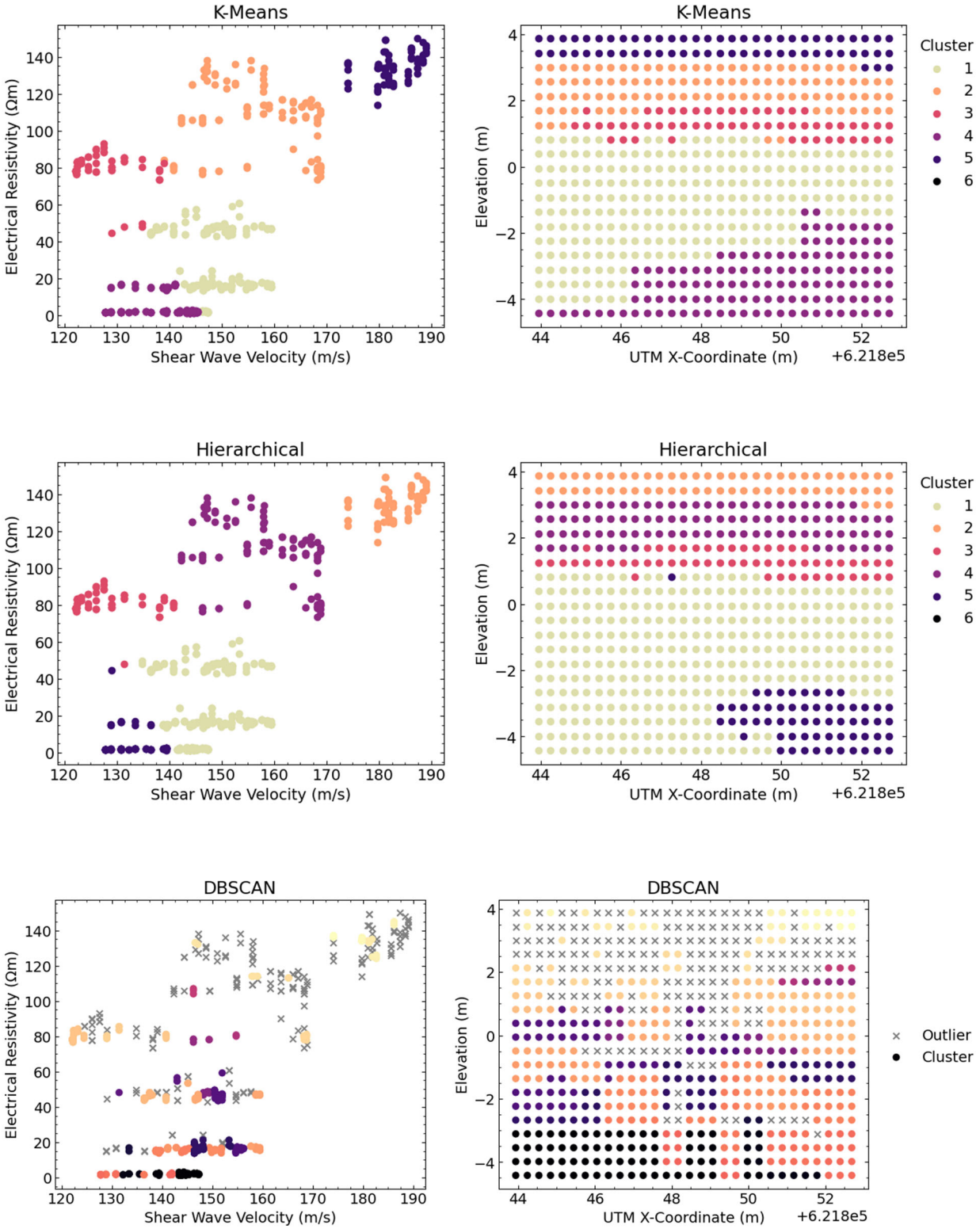
Fit: Wet_Site2

Profile: fc_site2



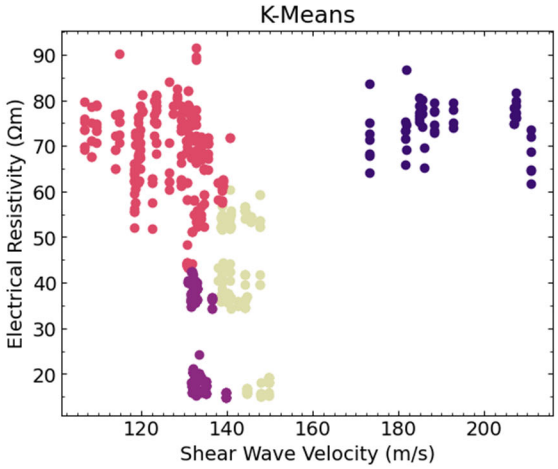
Fit: Wet_Site2

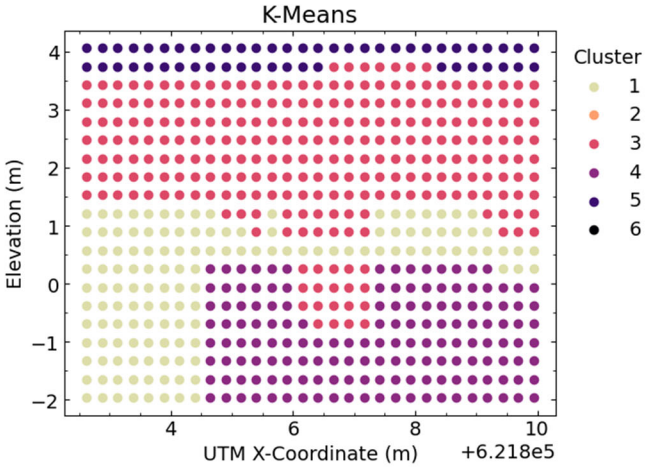
Profile: fdt_site2

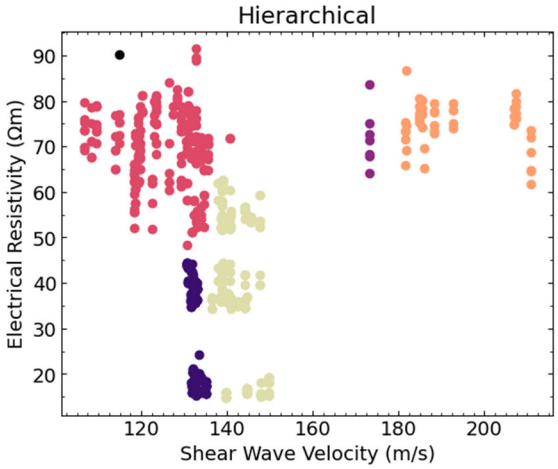


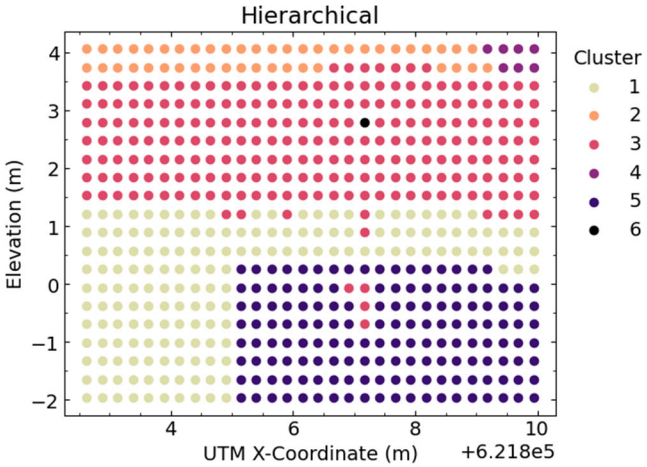
Fit: Wet_Site2

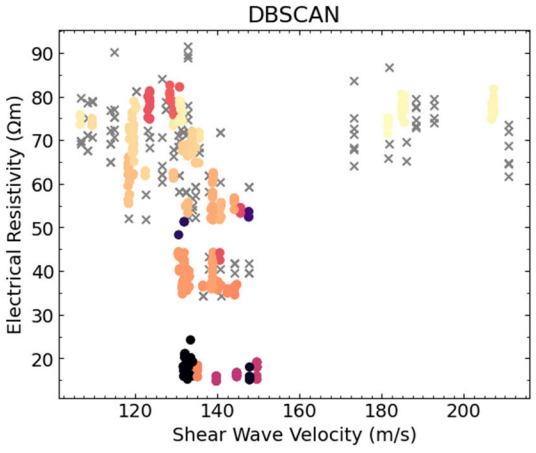
Profile: fwt_site2

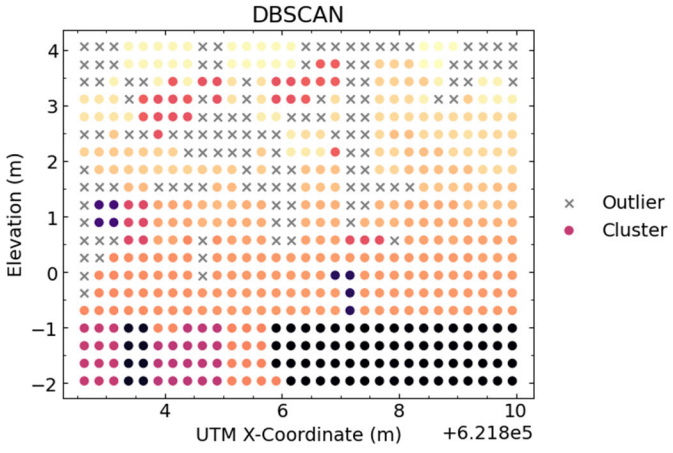






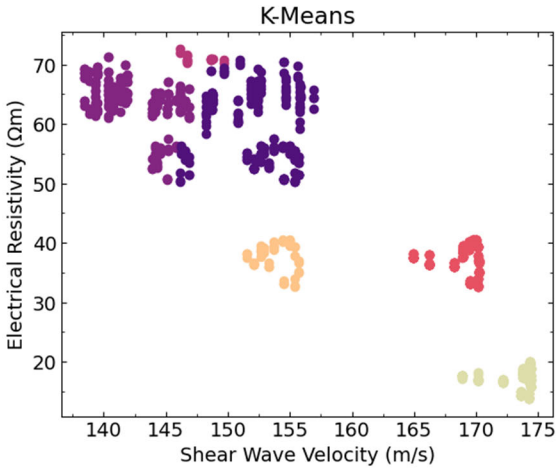


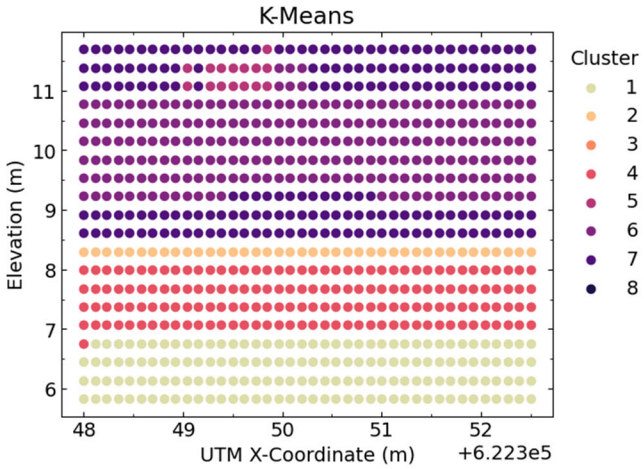


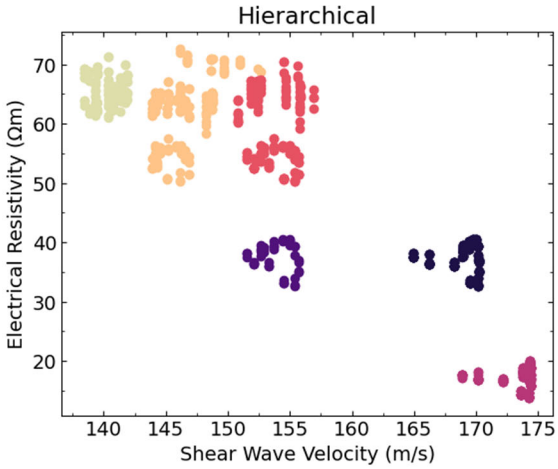


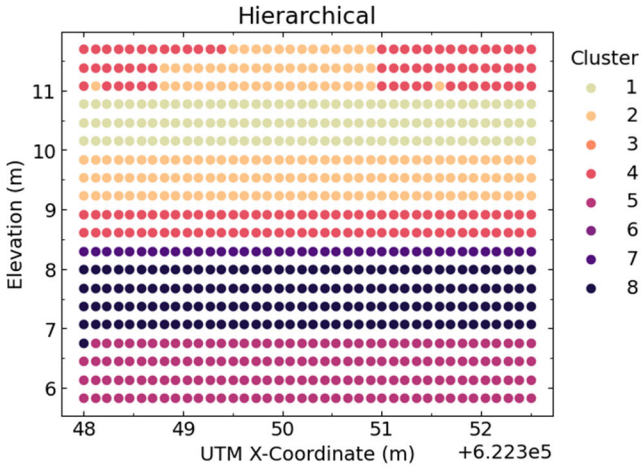
Fit: Dry_Site3

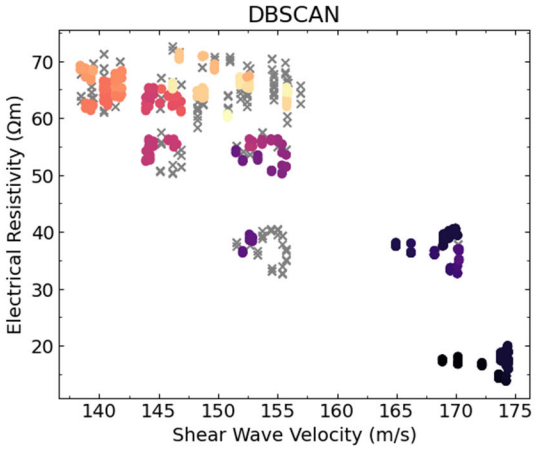
Profile: sc_site3

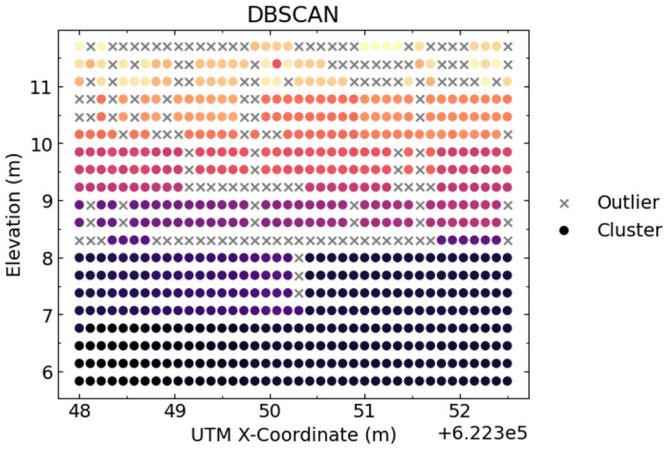






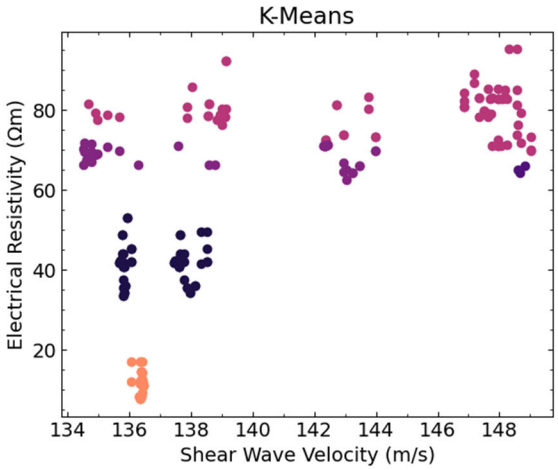


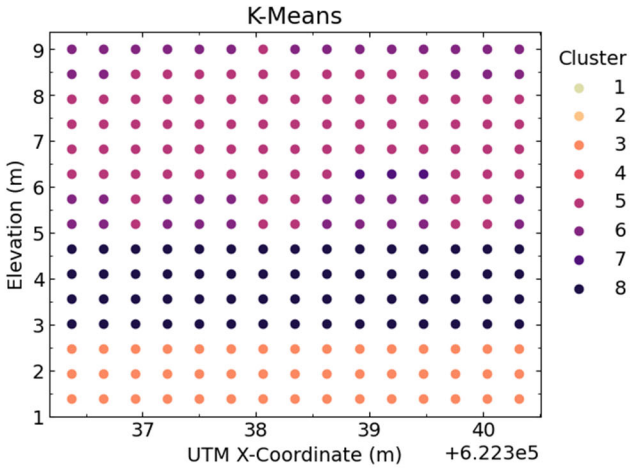


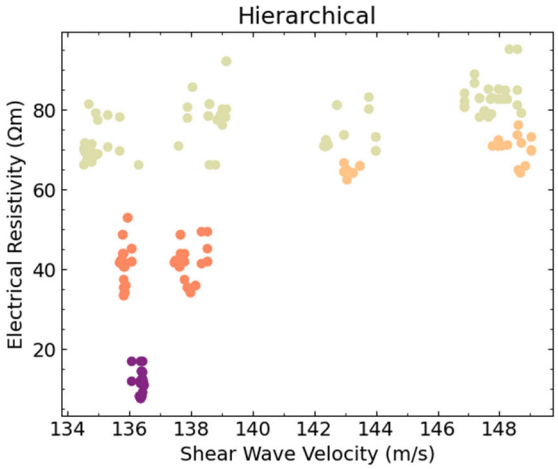


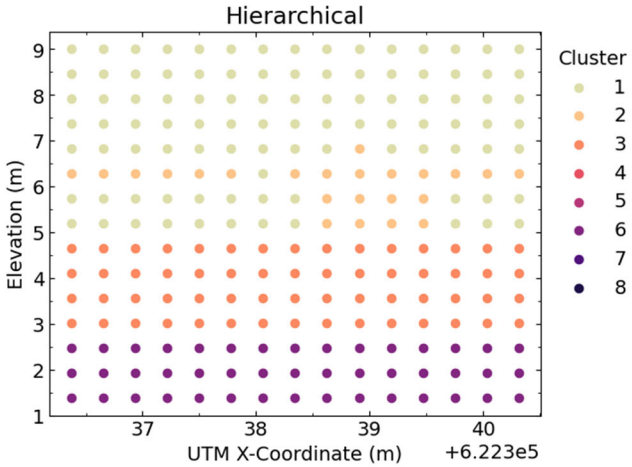
Fit: Dry_Site3

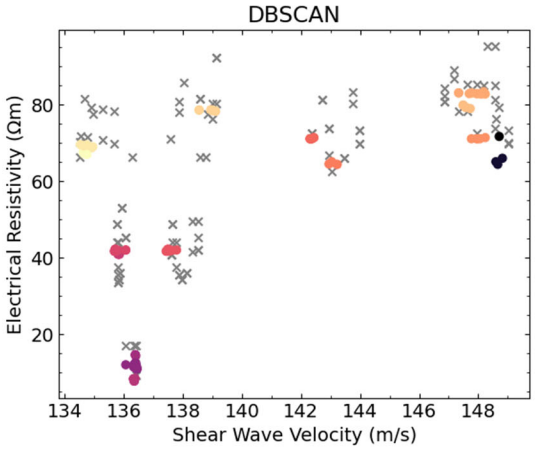
Profile: swt_site3

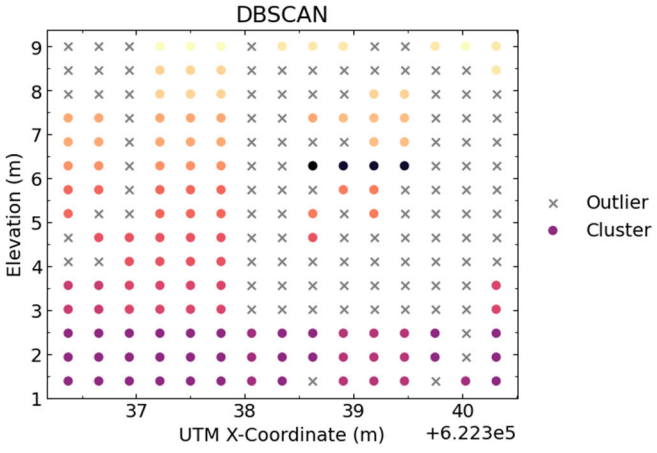






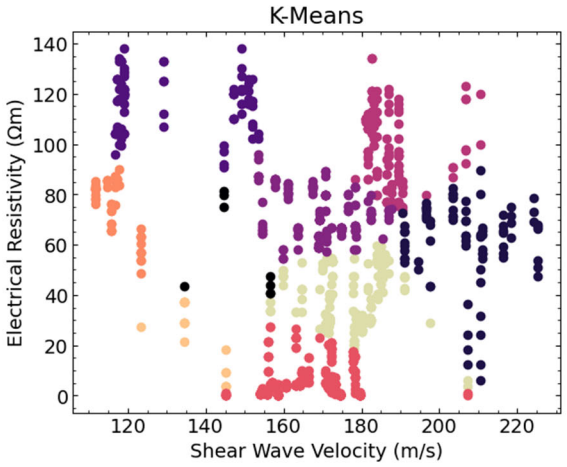


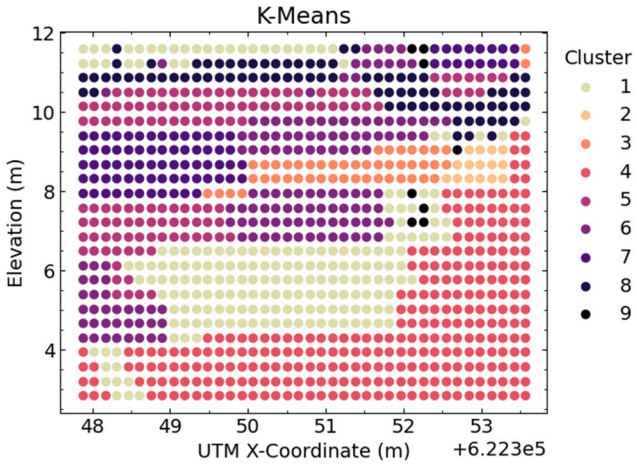


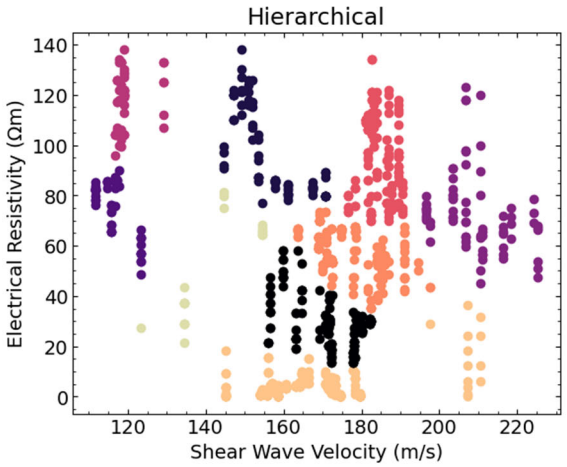


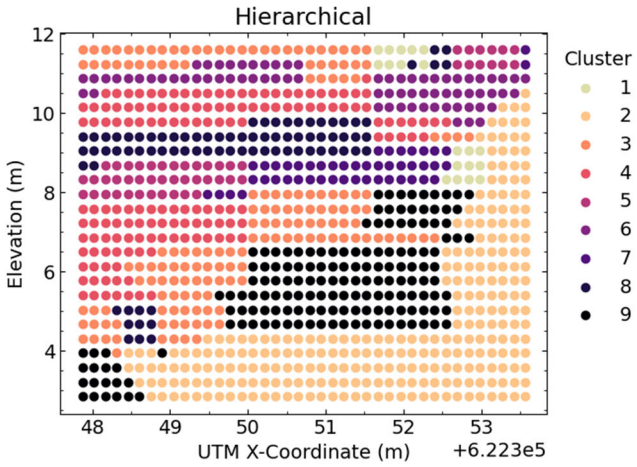
Fit: Wet_Site3

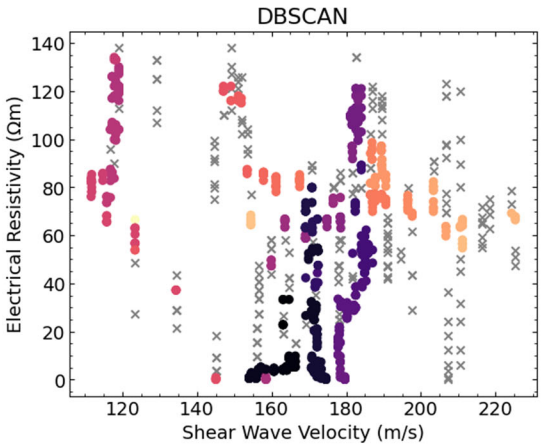
Profile: fc_site3

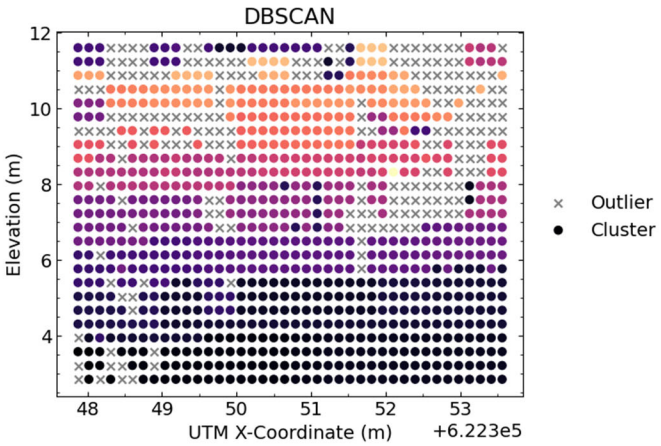






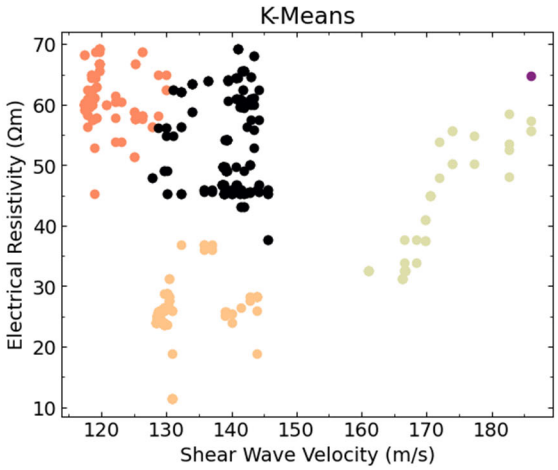


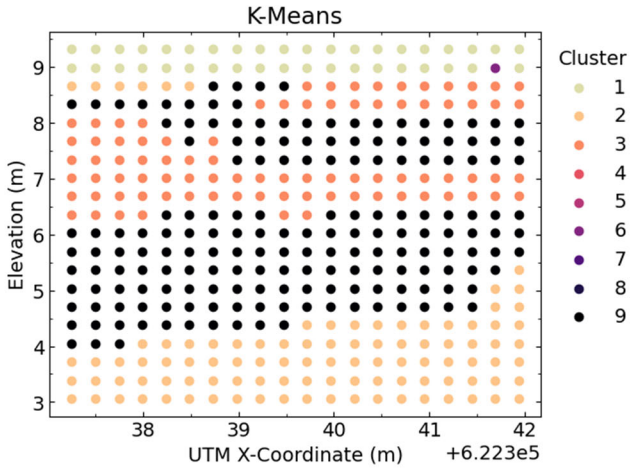


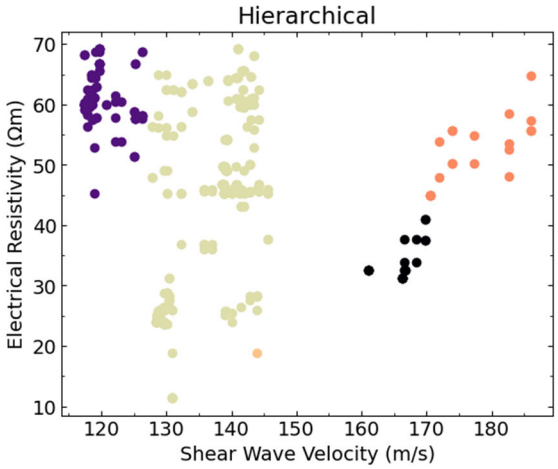


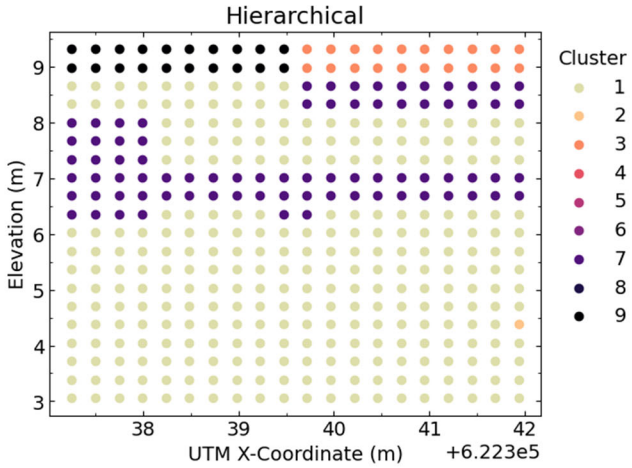
Fit: Wet_Site3

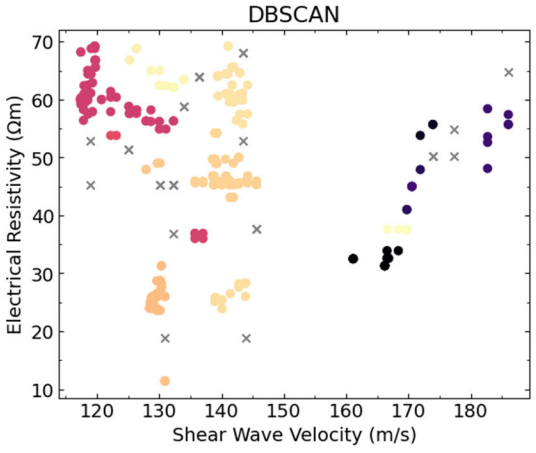
Profile: fwt_site3

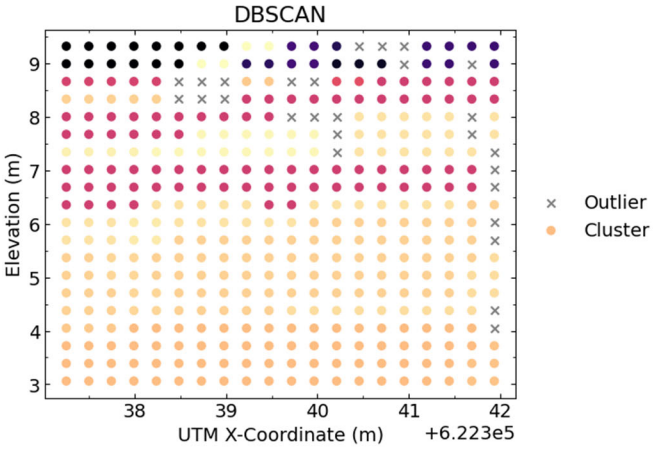






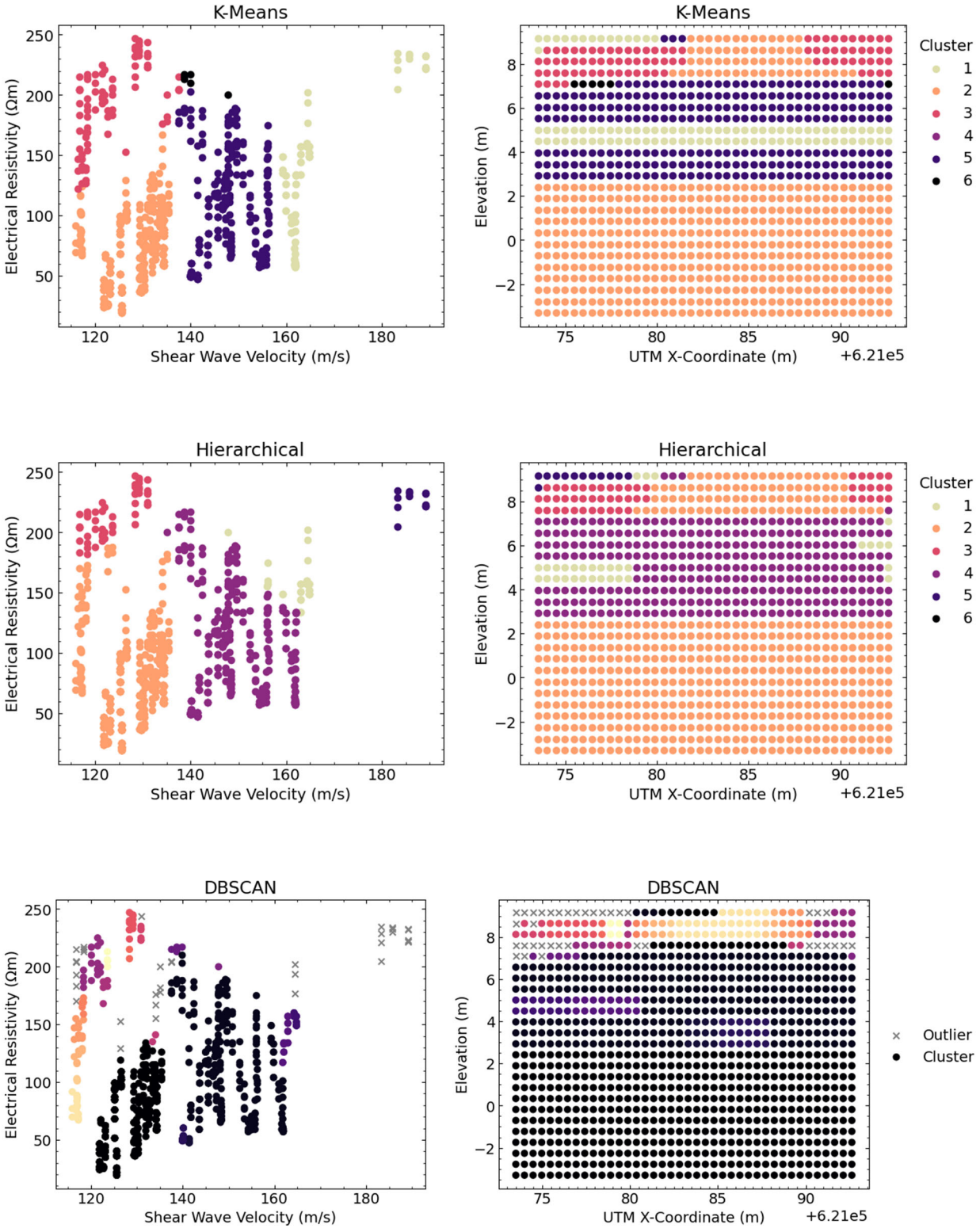






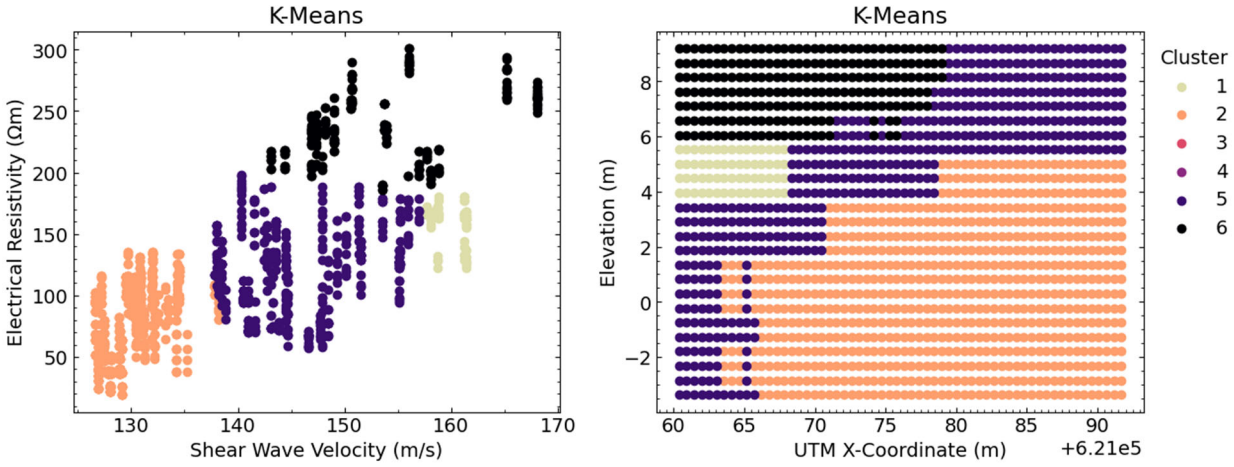
Fit: Site1

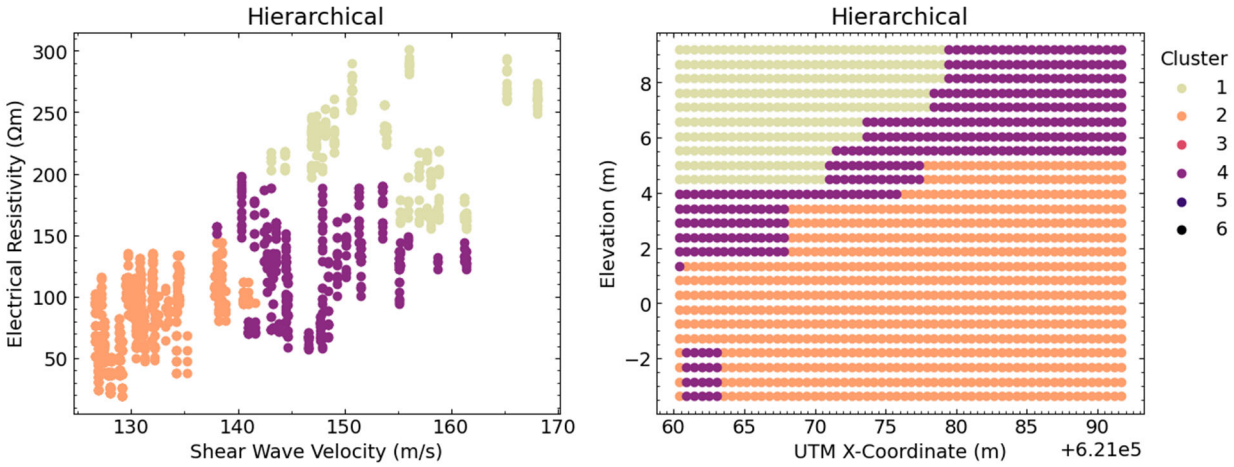
Profile: fc3_site1

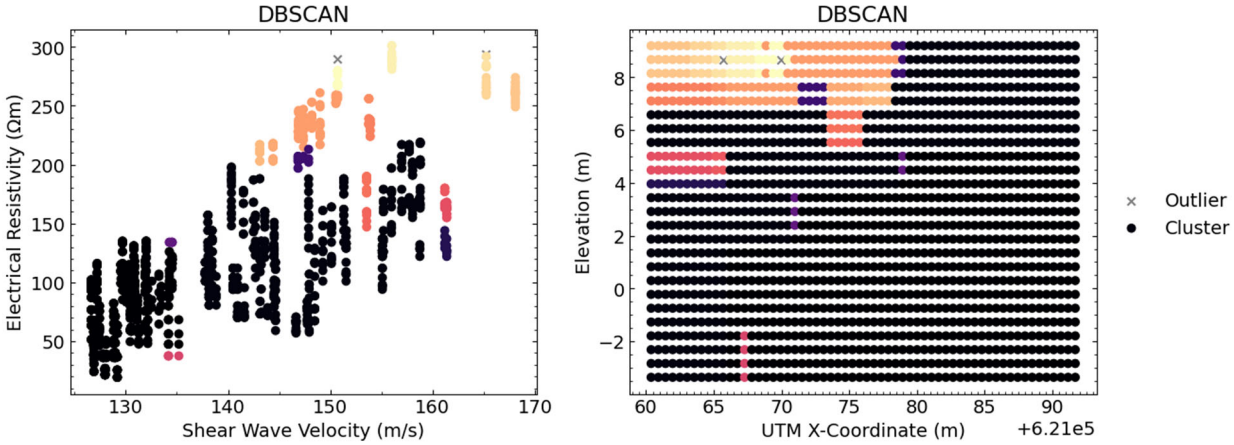


Fit: Site1

Profile: fc5_site1

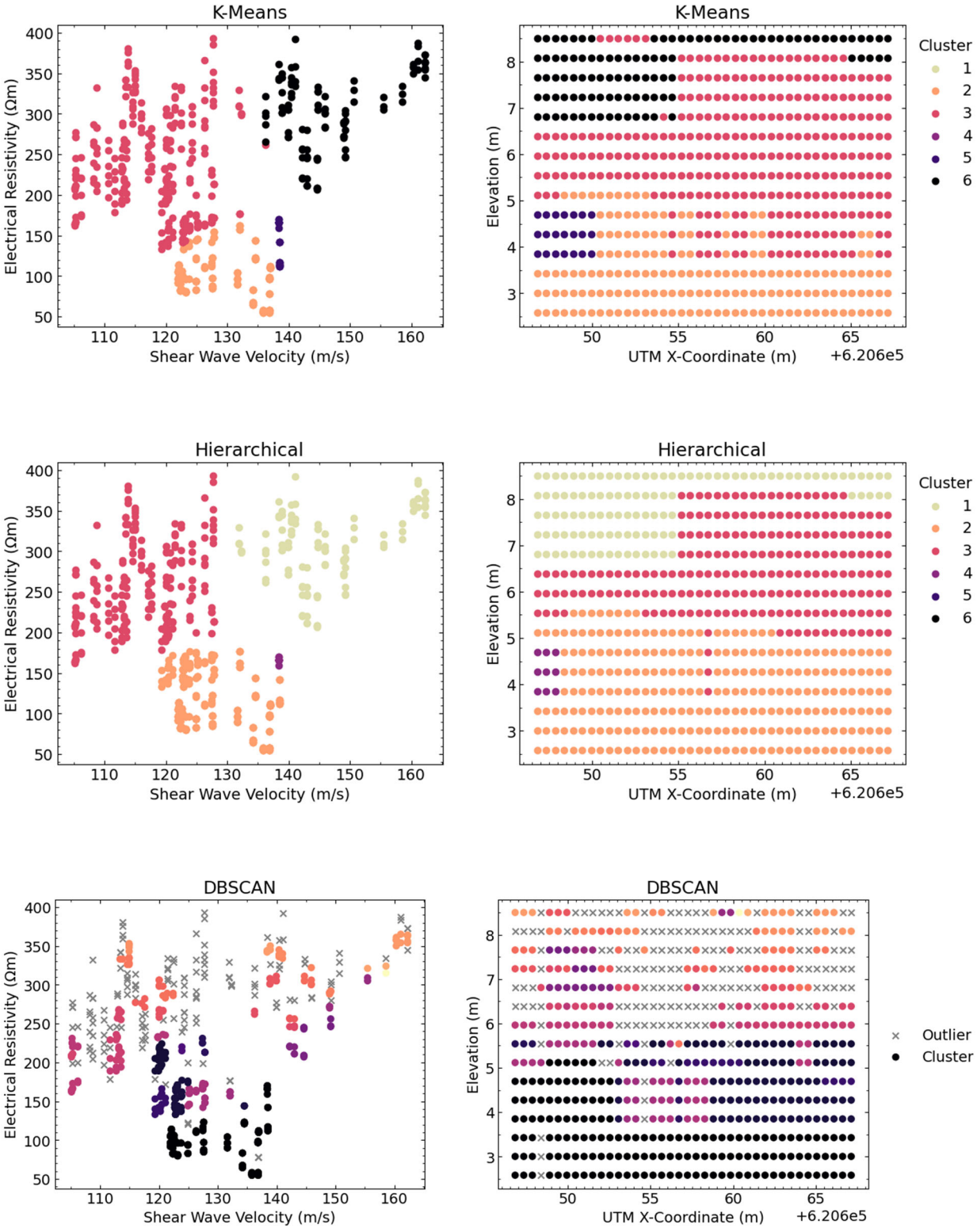






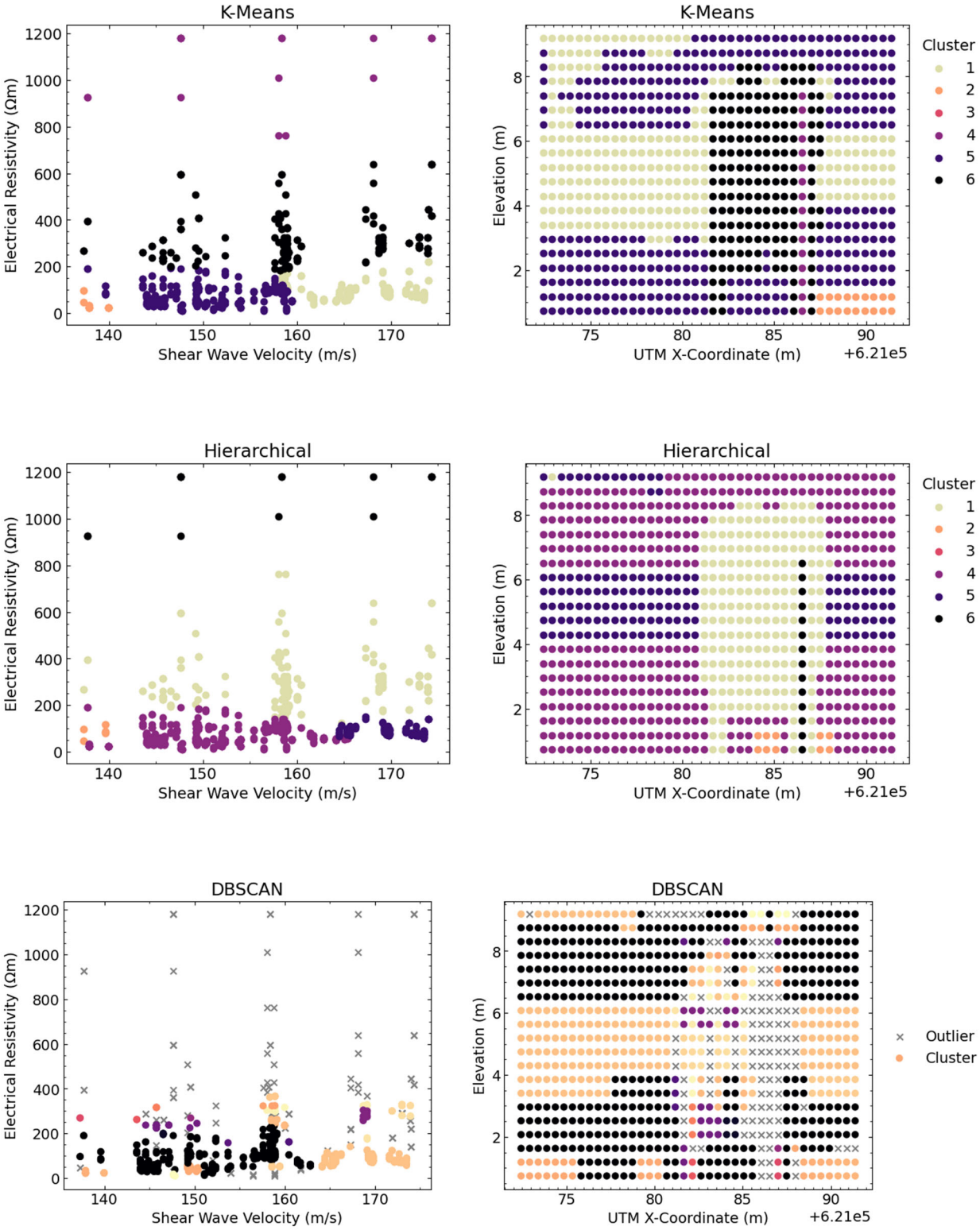
Fit: Site1

Profile: fwt_site1



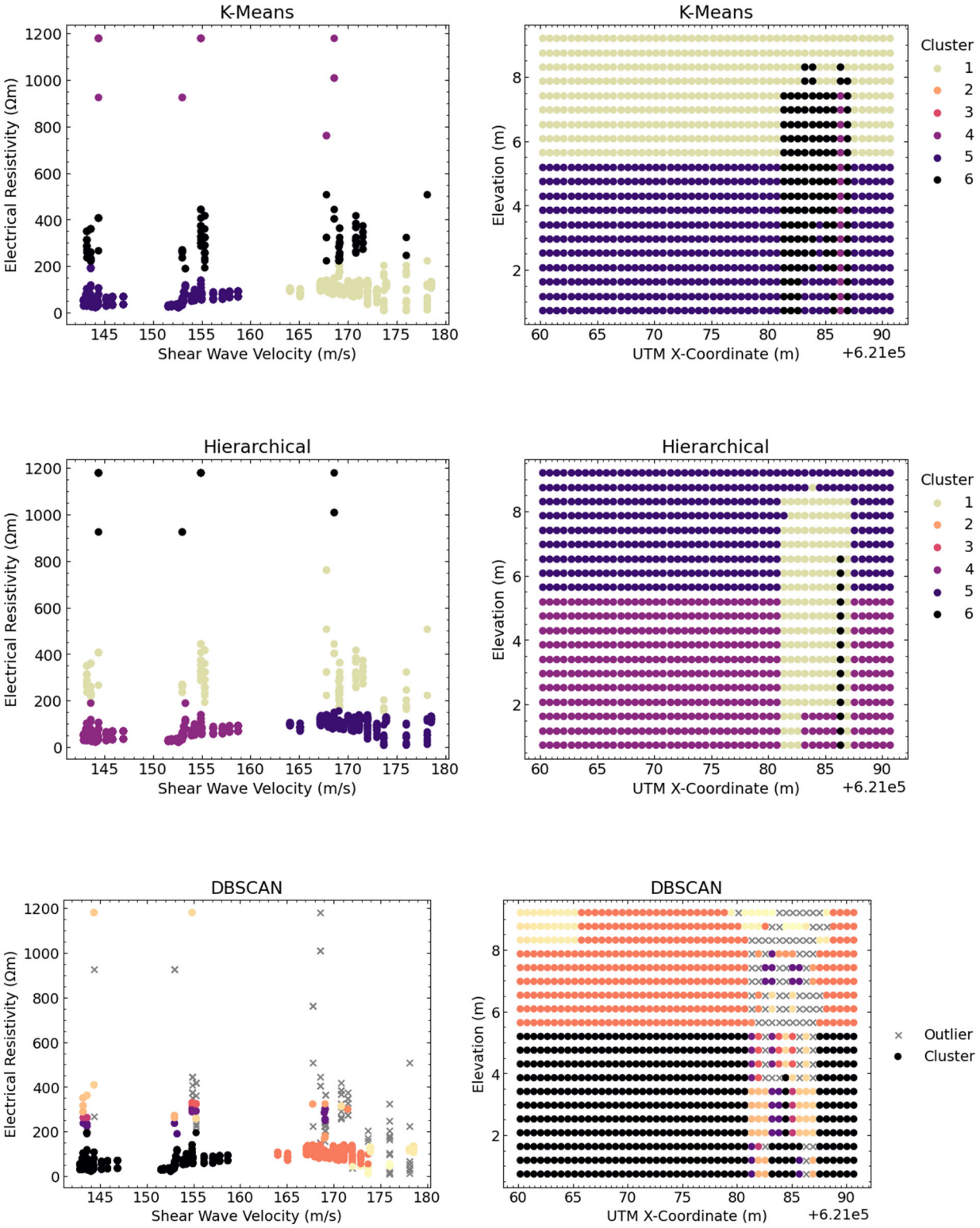
Fit: Site1

Profile: sc3_site1



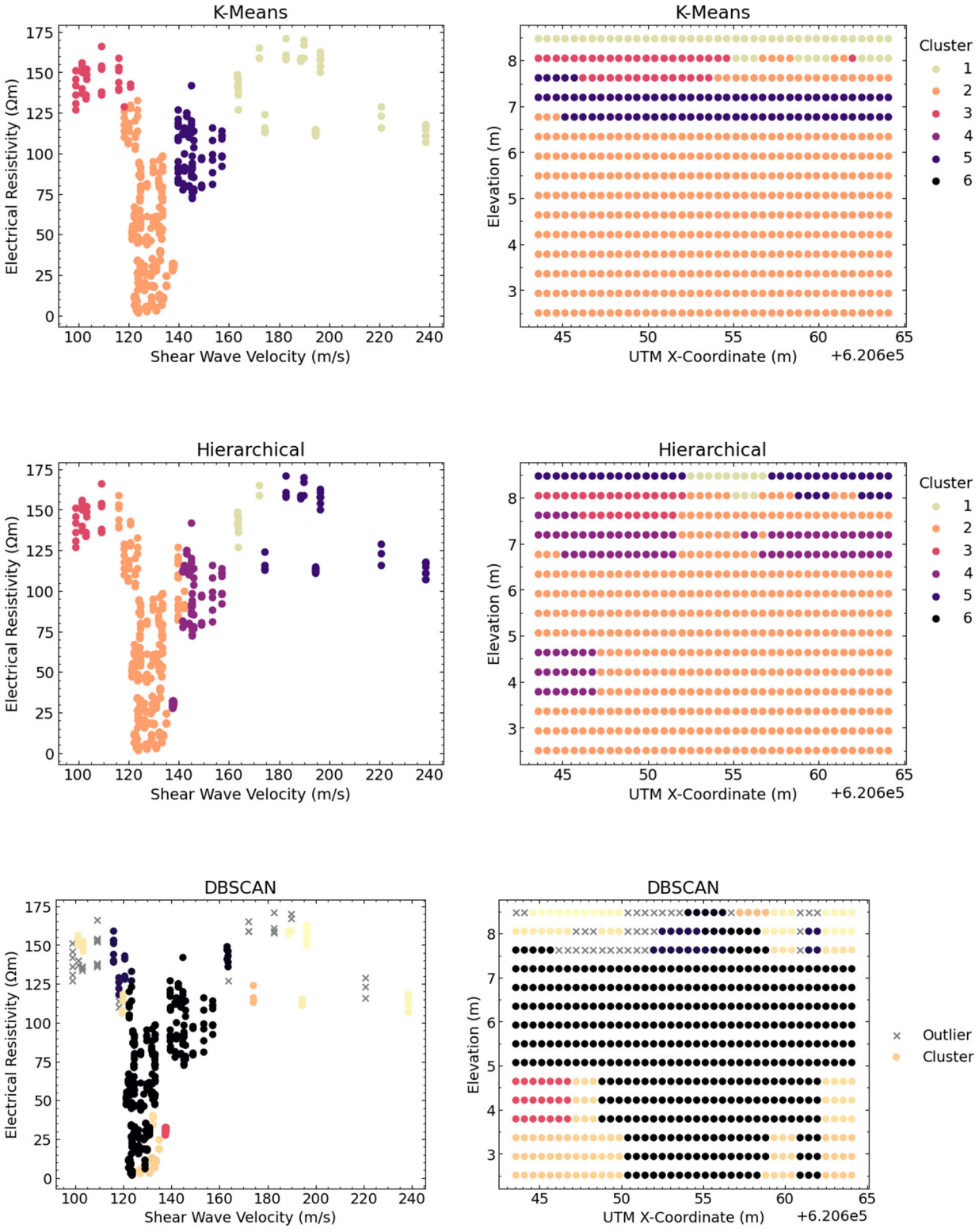
Fit: Site1

Profile: sc5_site1



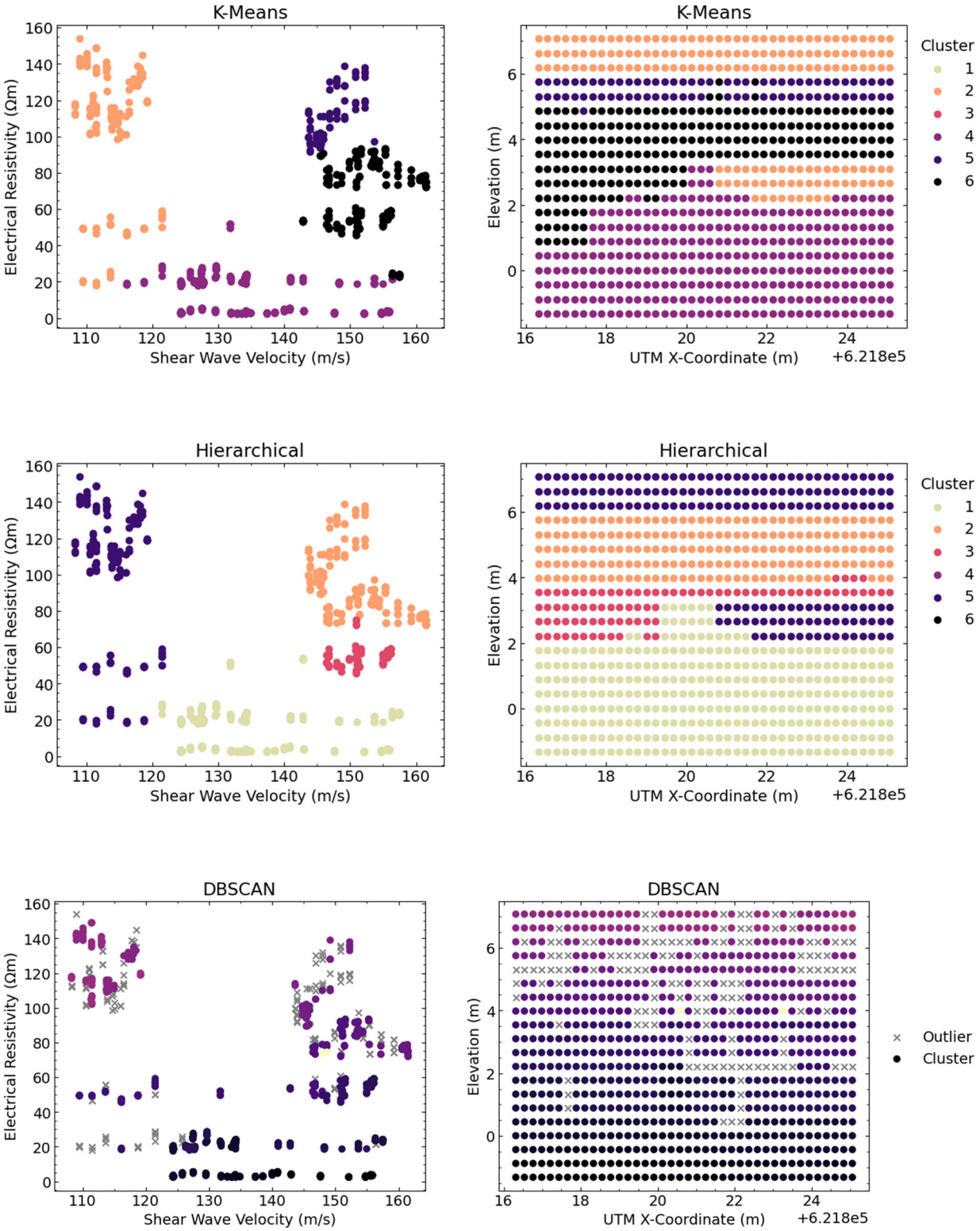
Fit: Site1

Profile: swt_site1



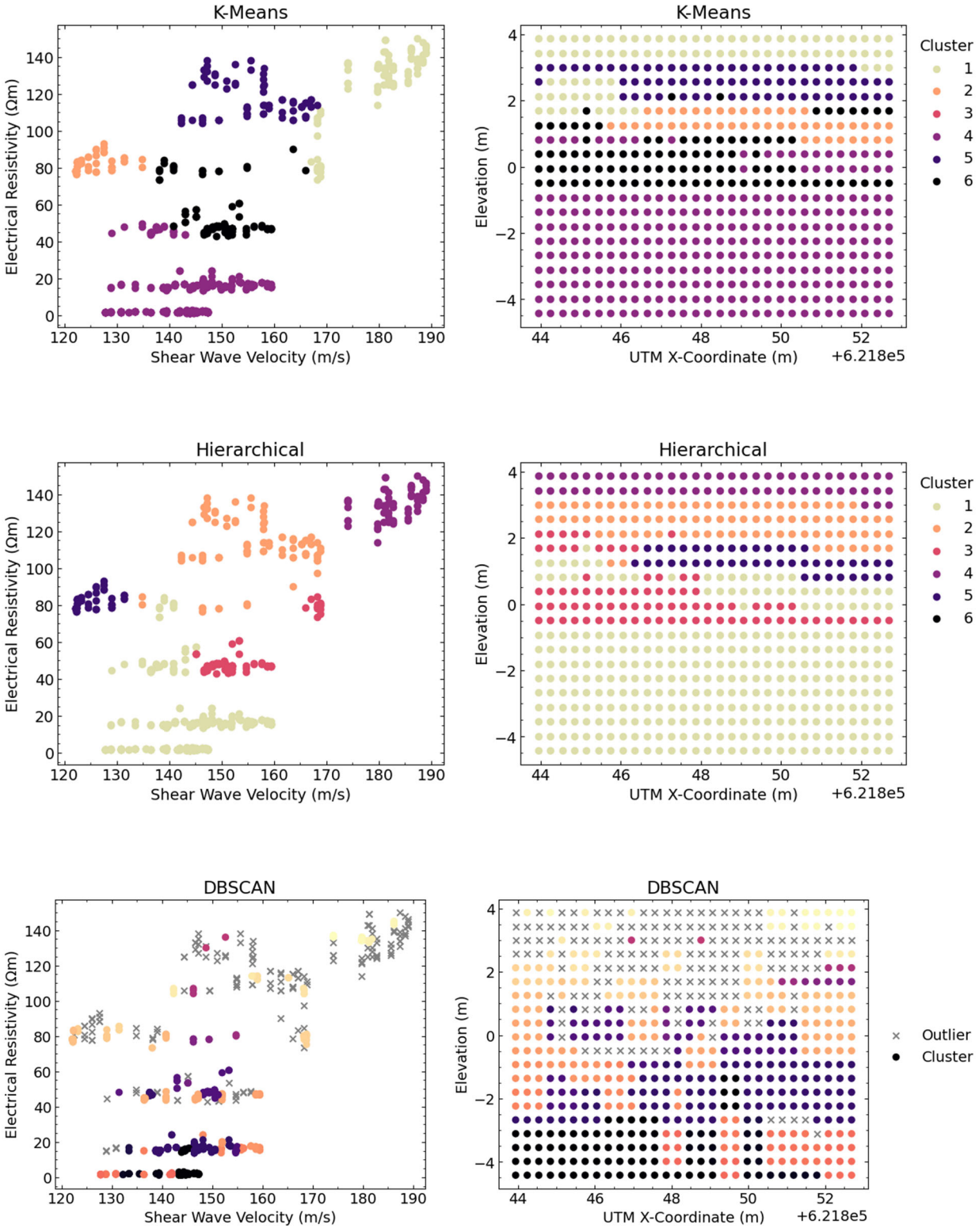
Fit: Site2

Profile: fc_site2



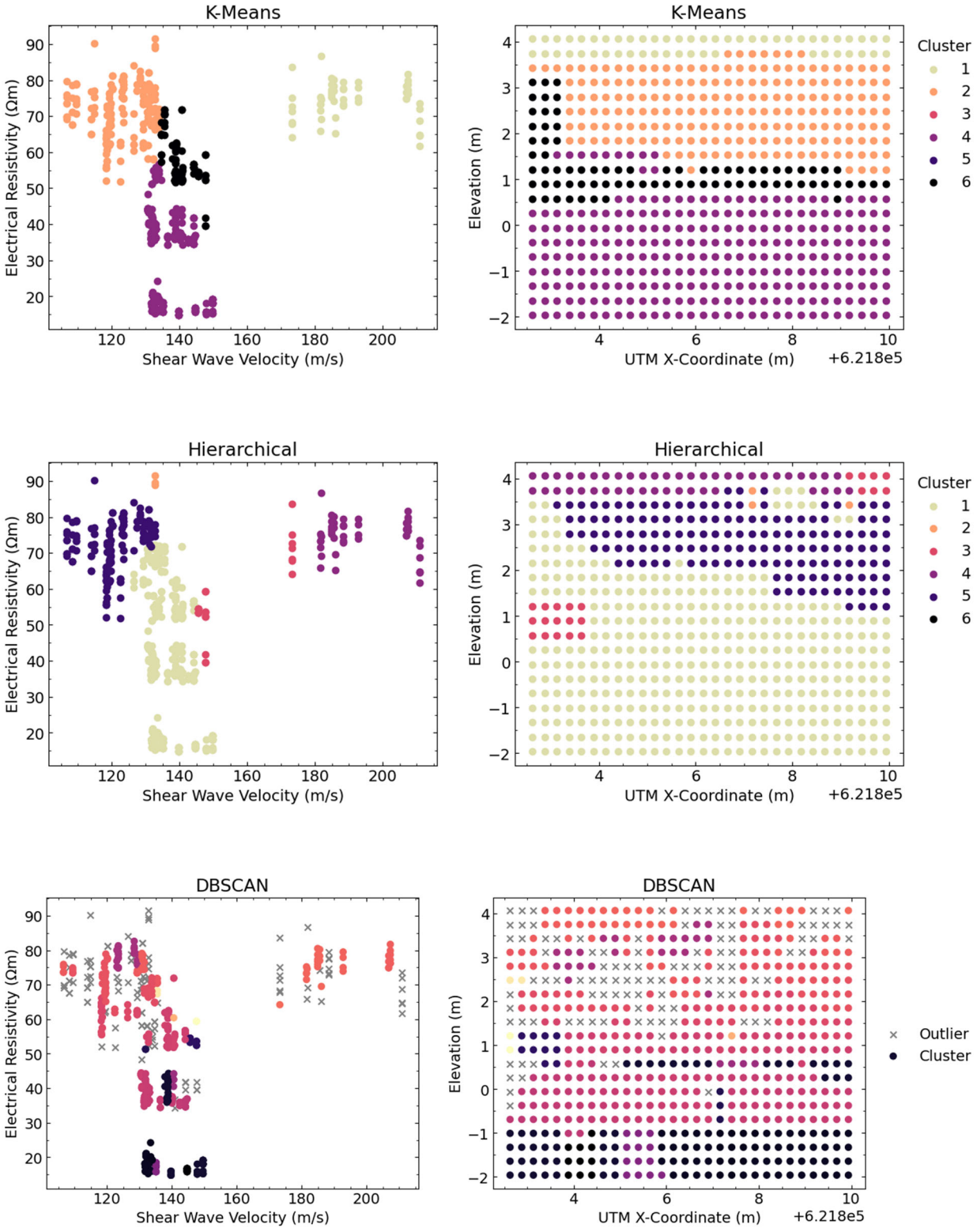
Fit: Site2

Profile: fdt_site2



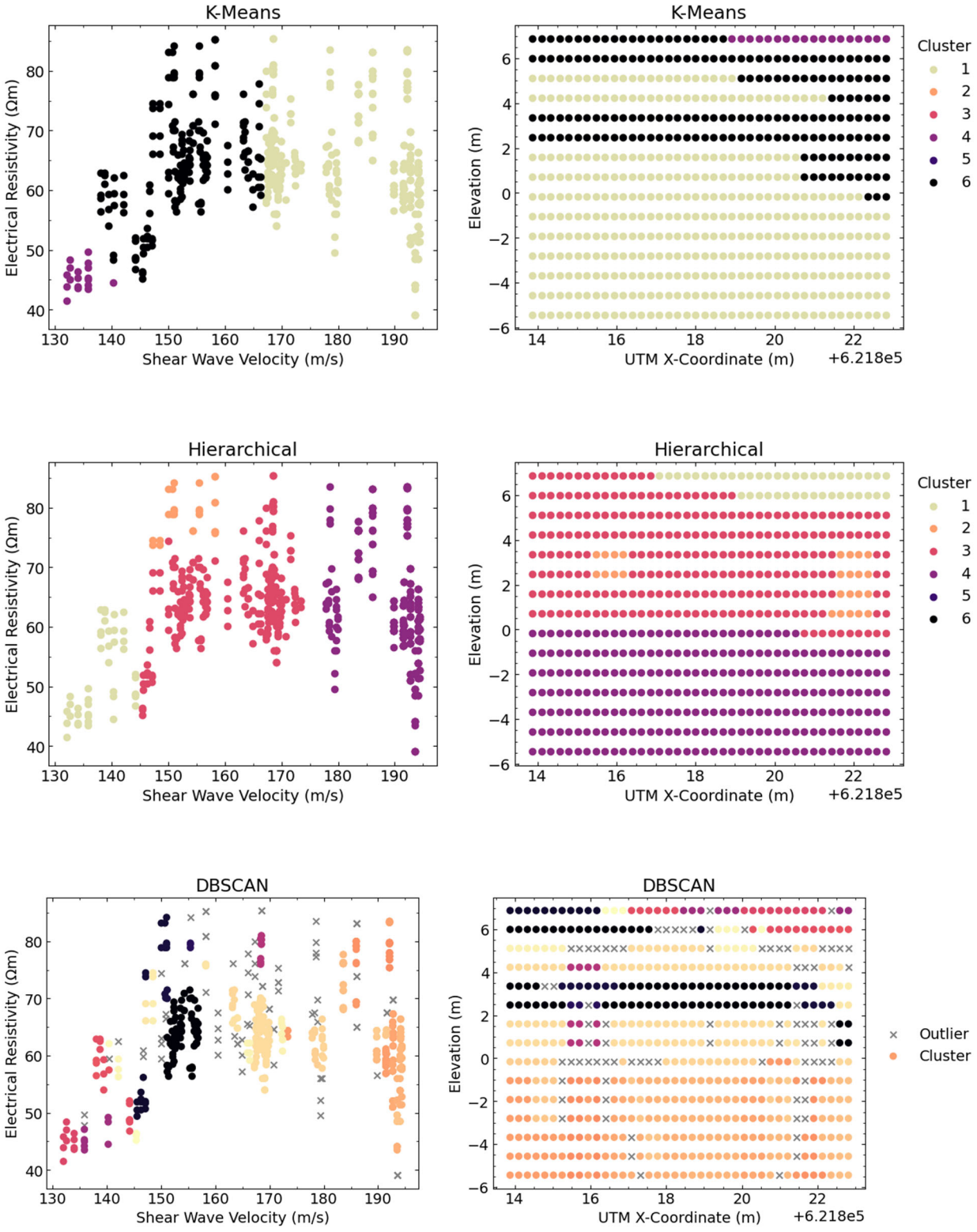
Fit: Site2

Profile: fwt_site2



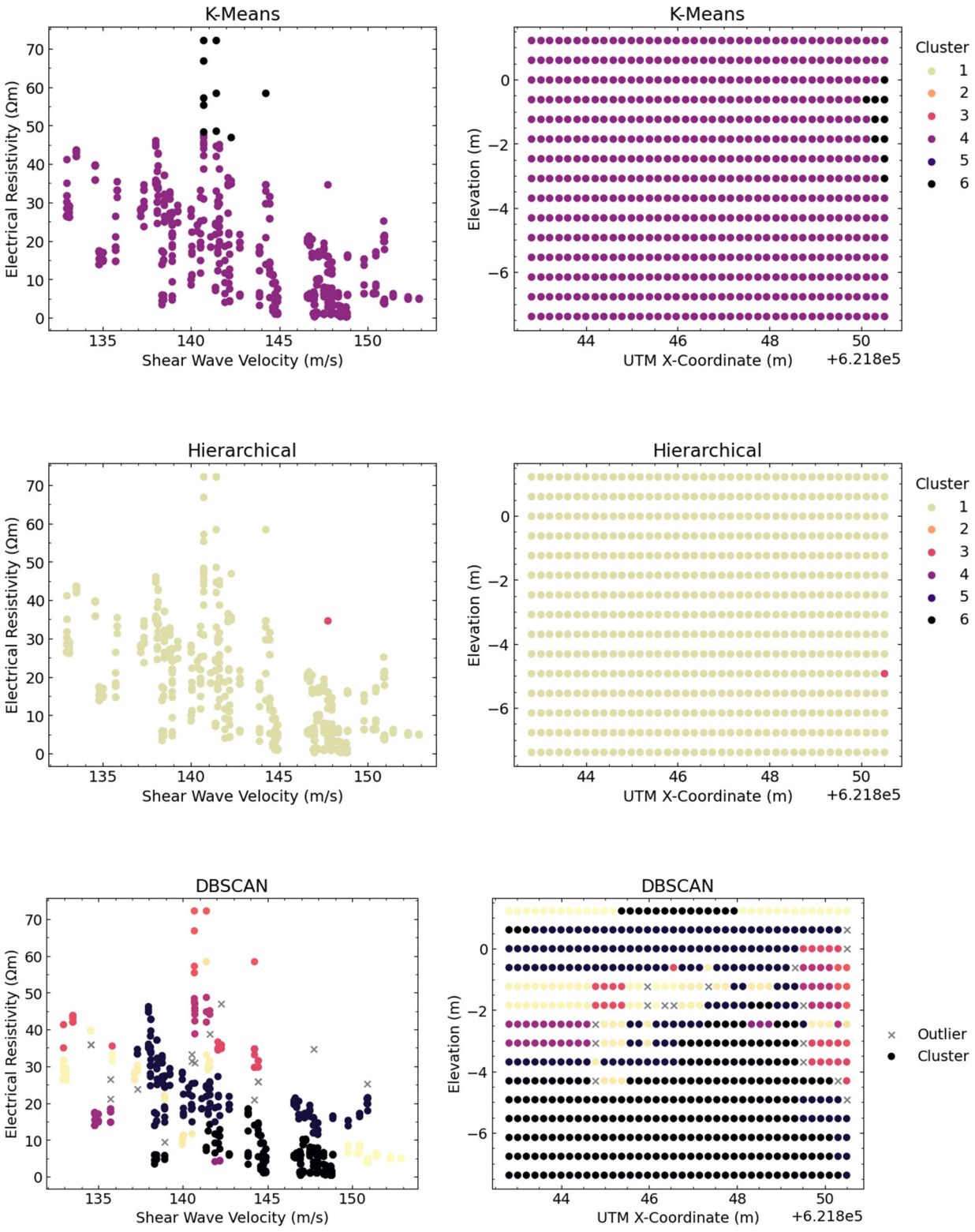
Fit: Site2

Profile: sc_site2



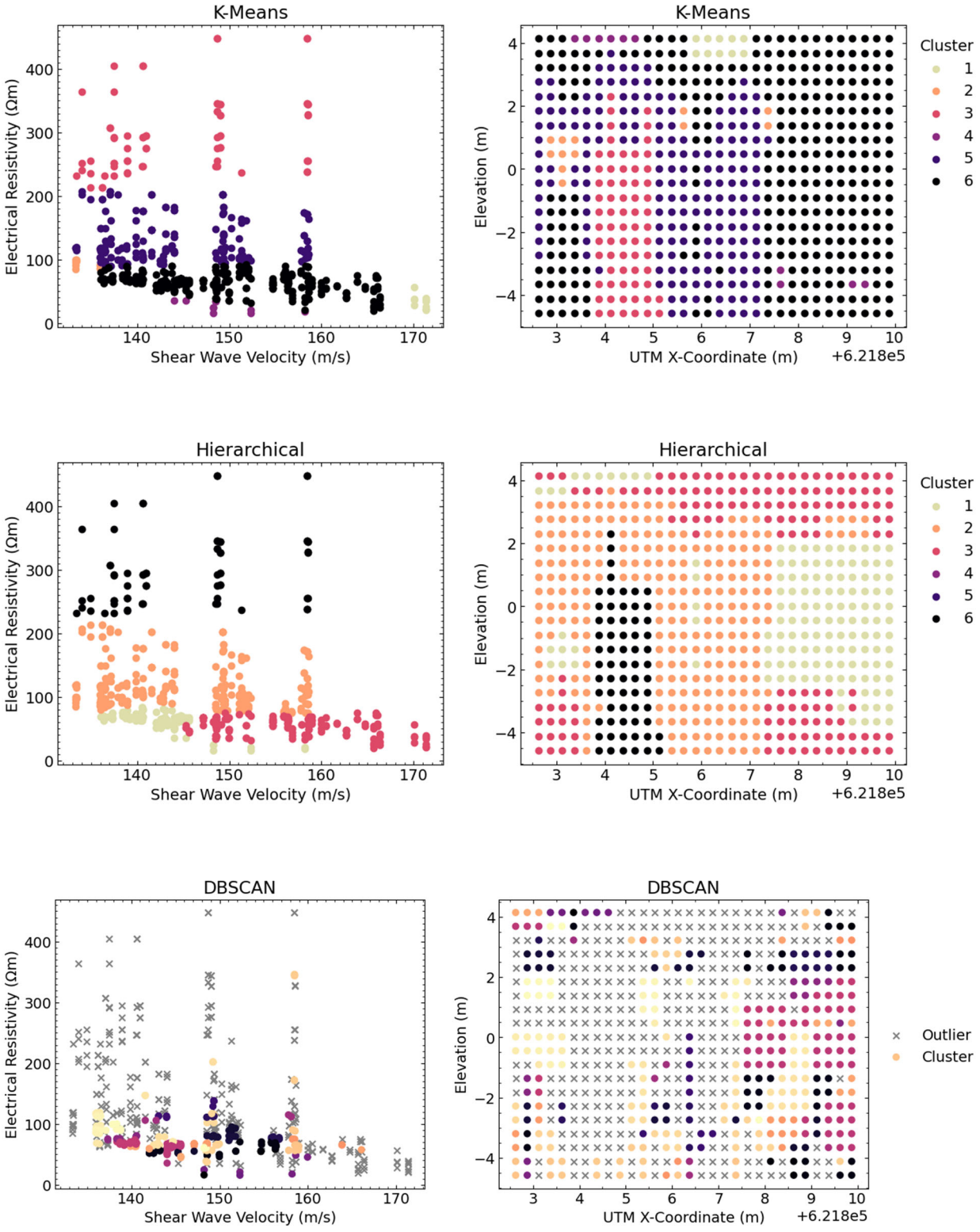
Fit: Site2

Profile: sdt_site2



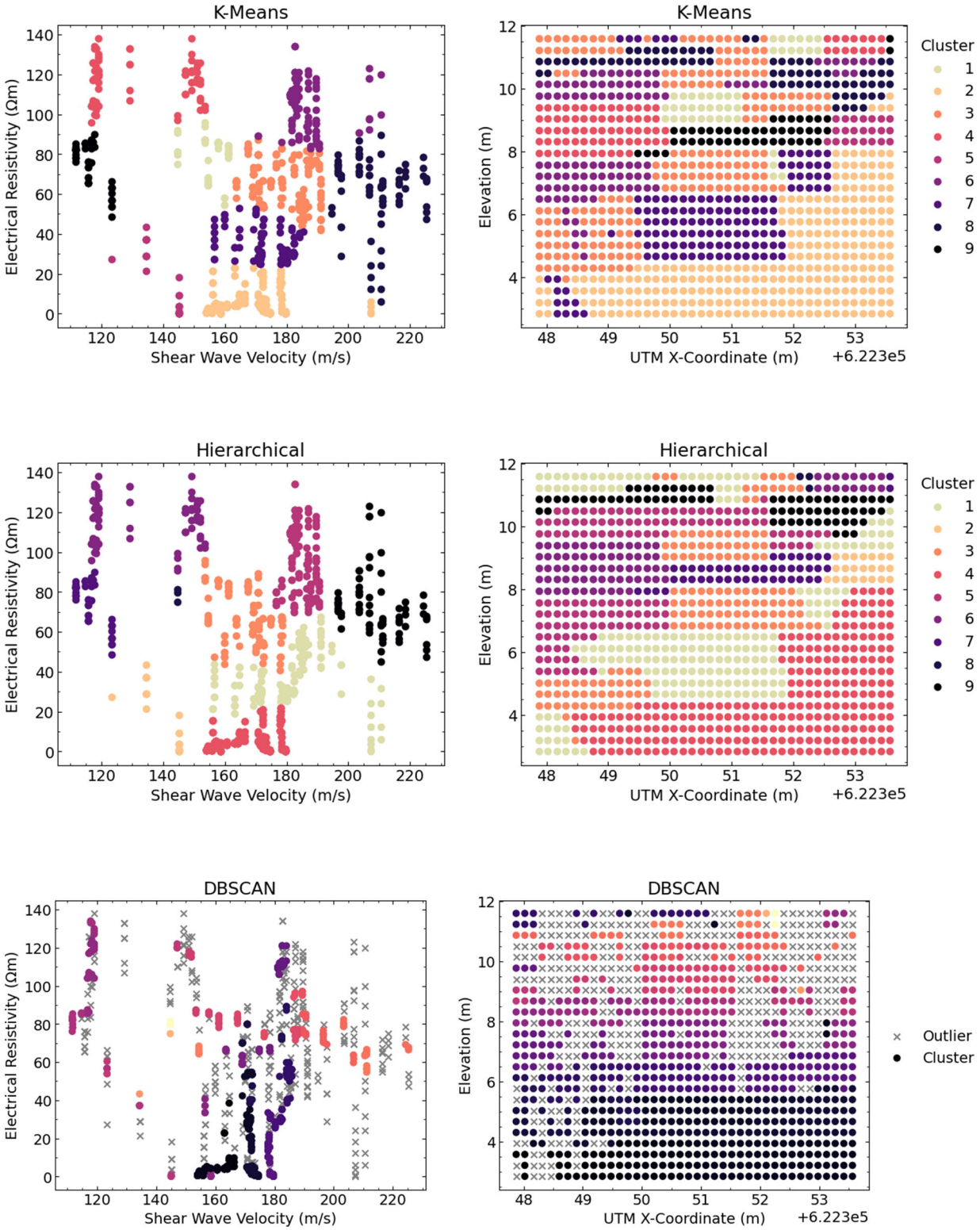
Fit: Site2

Profile: swt_site2



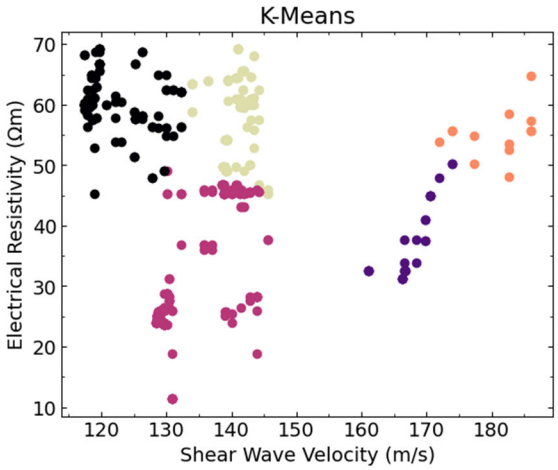
Fit: Site3

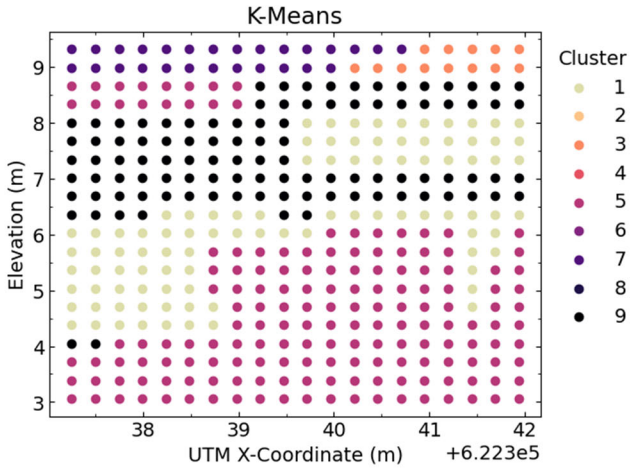
Profile: fc_site3

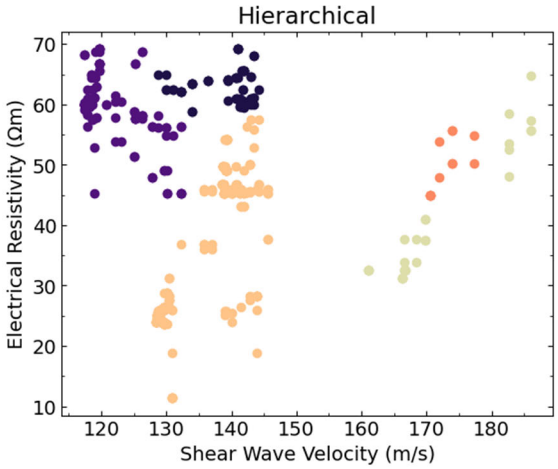


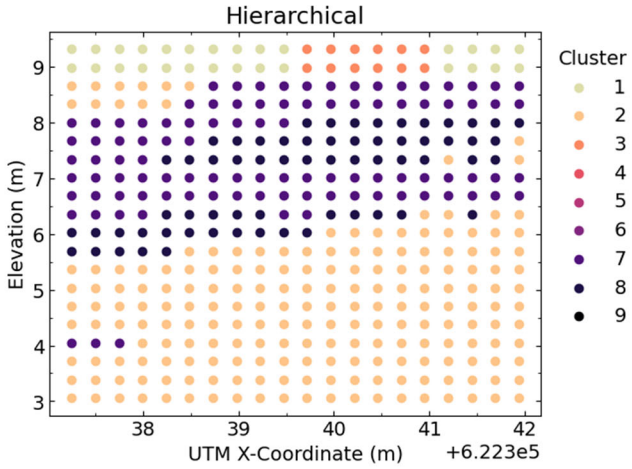
Fit: Site3

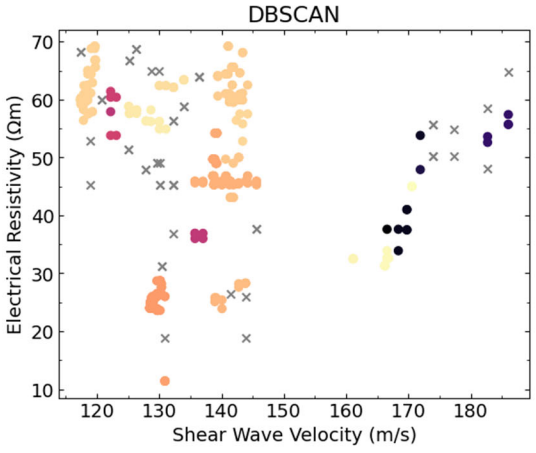
Profile: fwt_site3

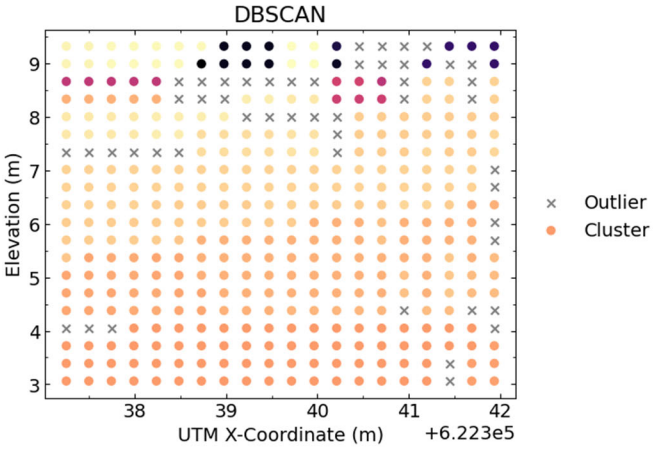






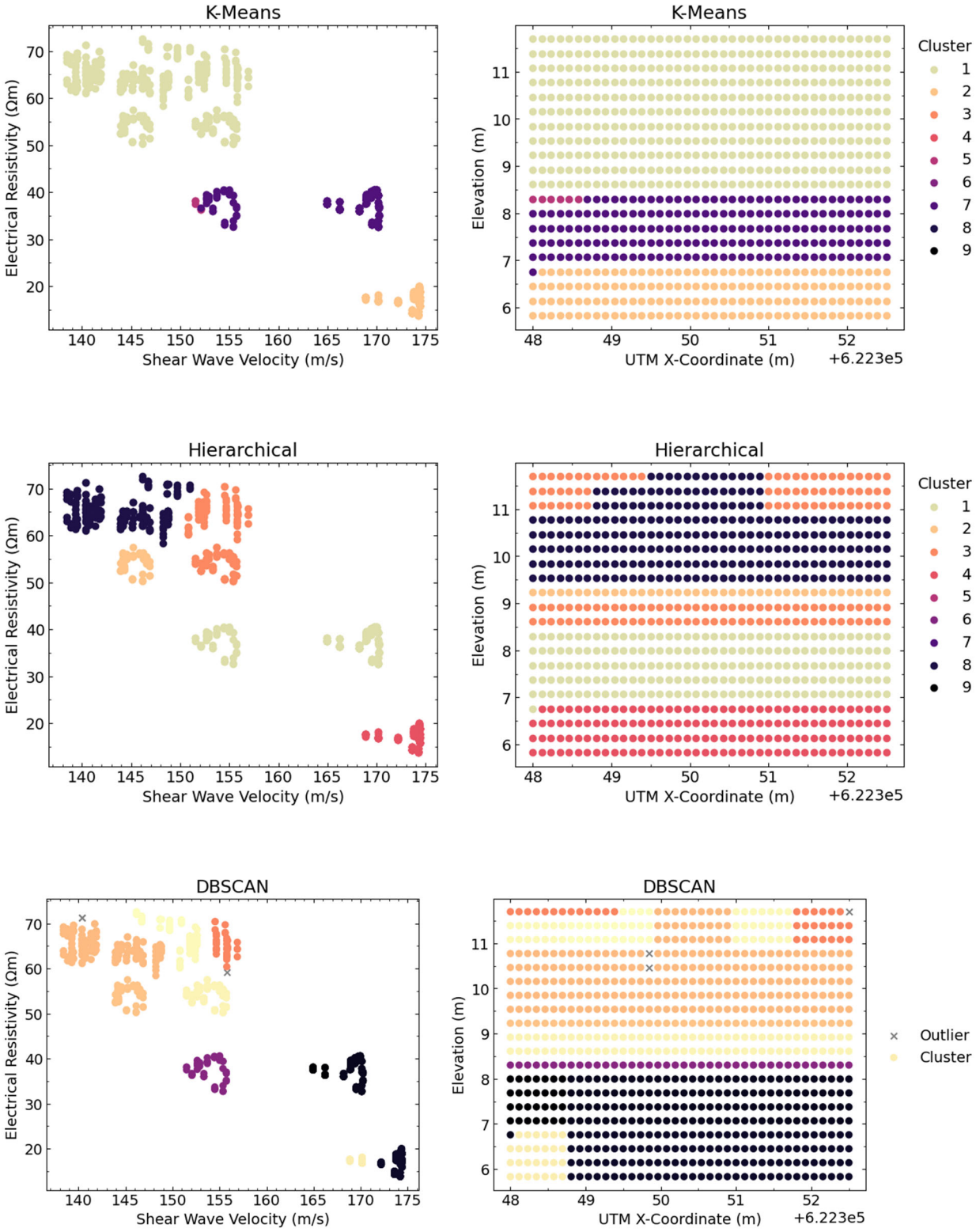






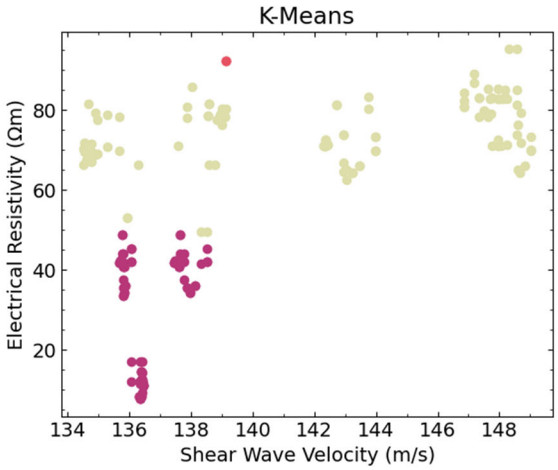
Fit: Site3

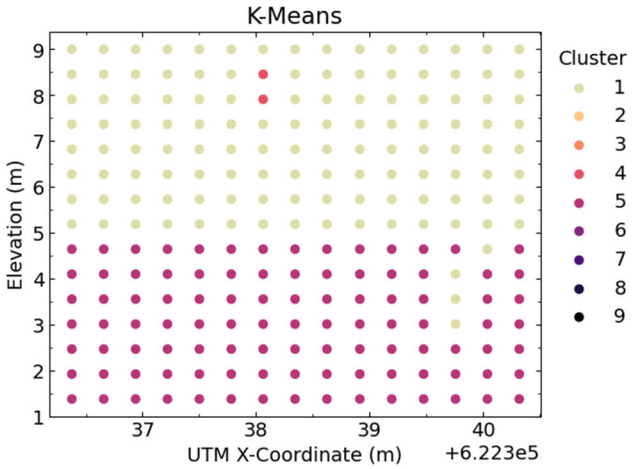
Profile: sc_site3

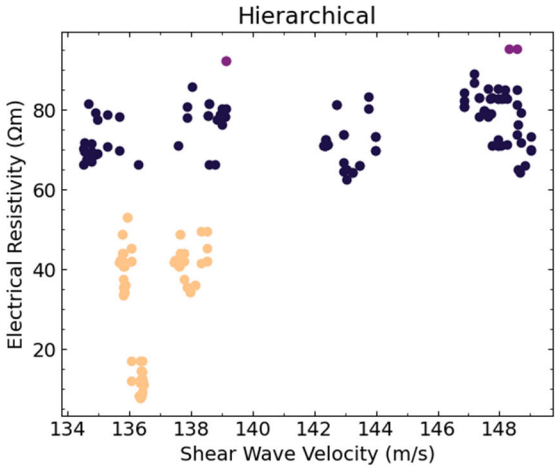


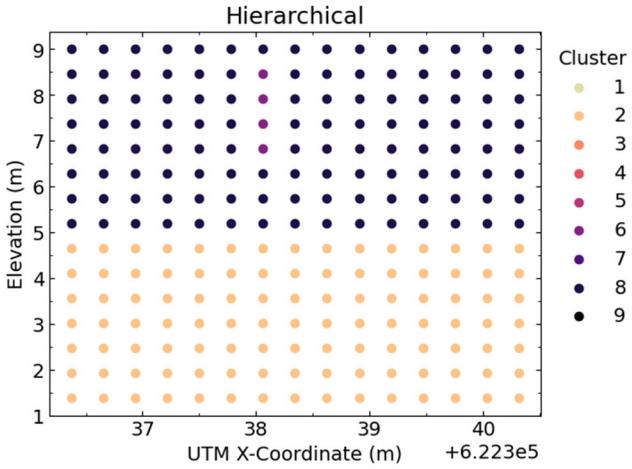
Fit: Site3

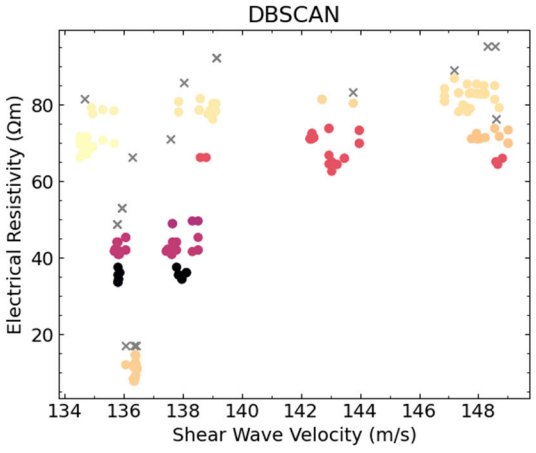
Profile: swt_site3

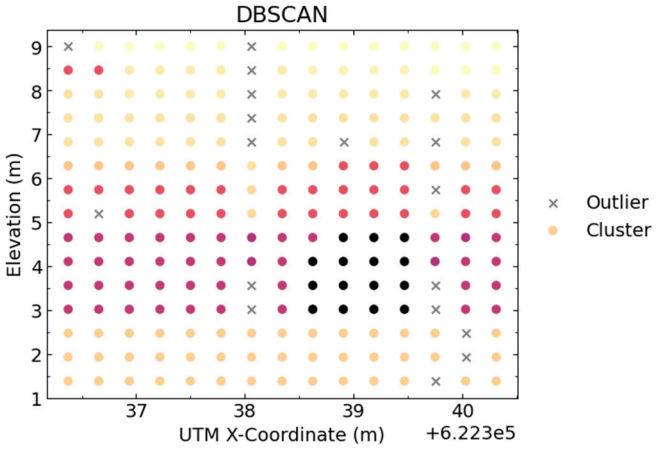








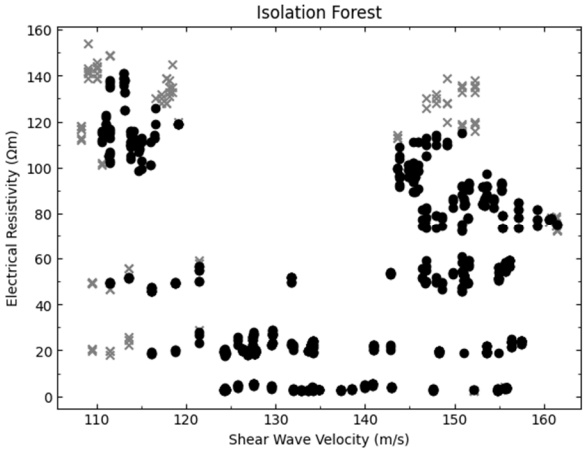


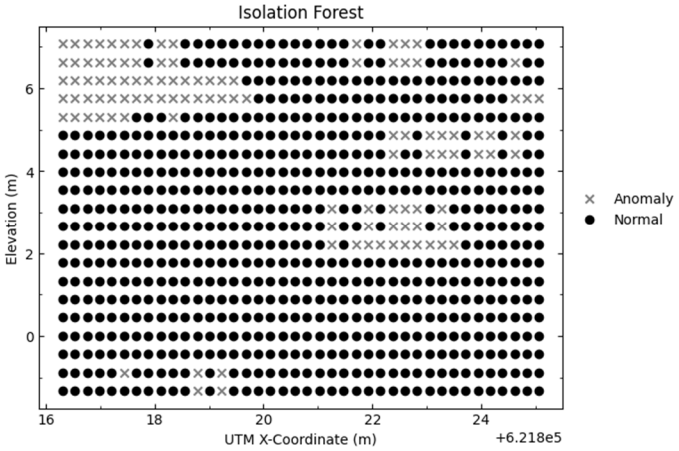


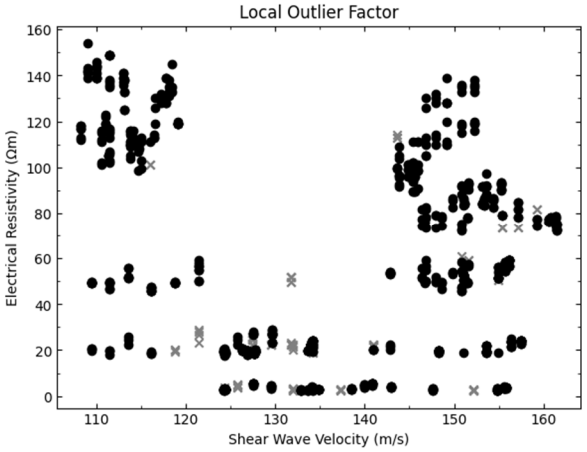
Appendix H – Anomaly Detection

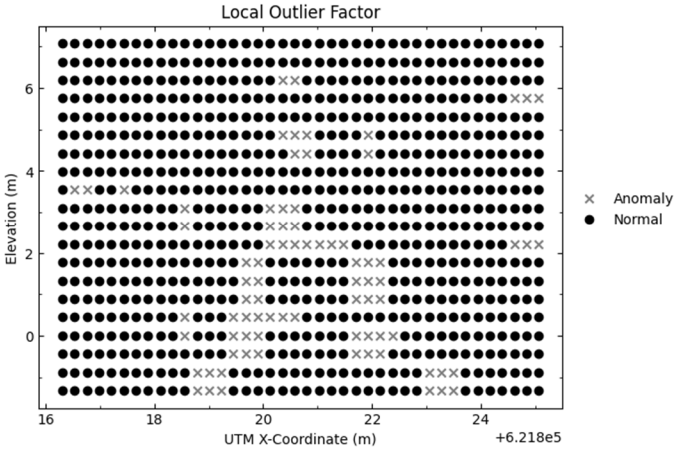
Fit: Individual Profiles
Predicted: Individual Profiles

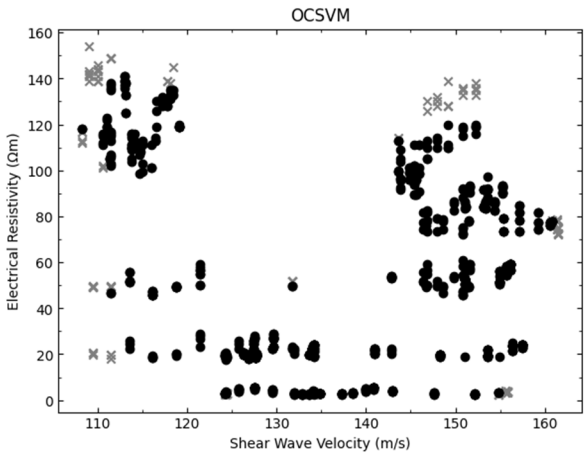
Profile: fc_site2

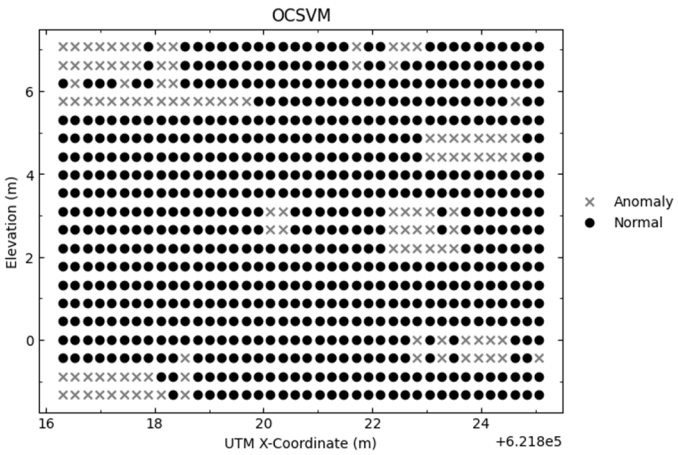




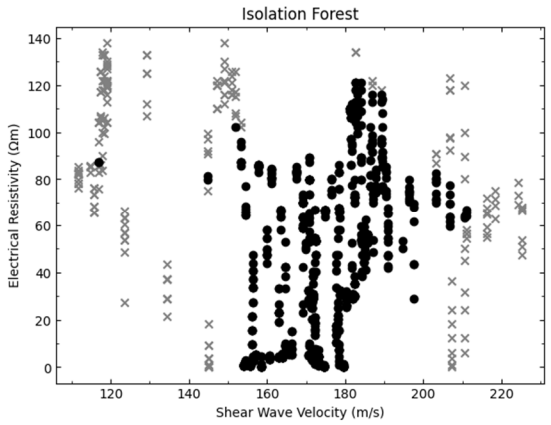


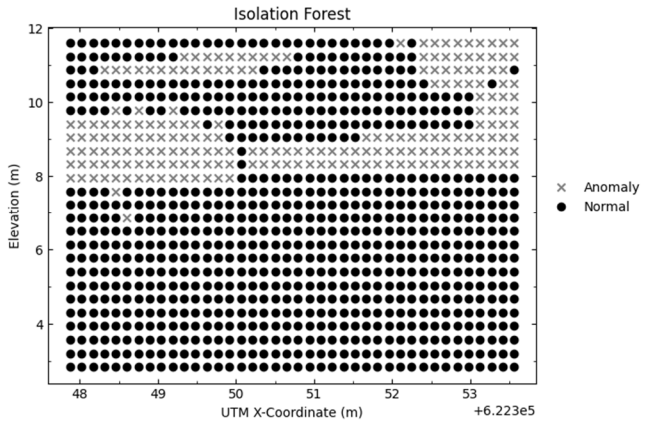


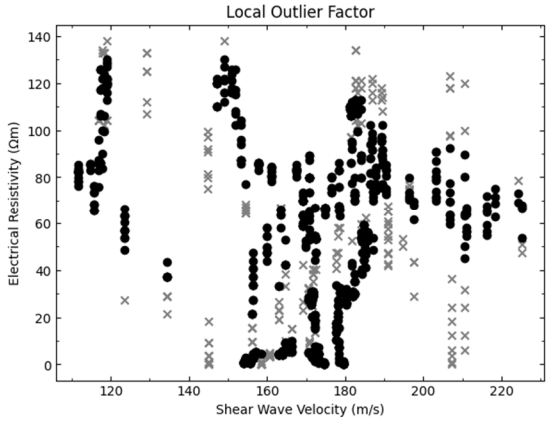


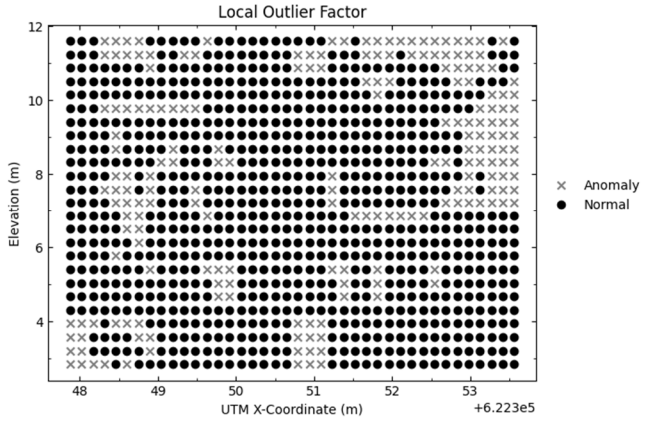


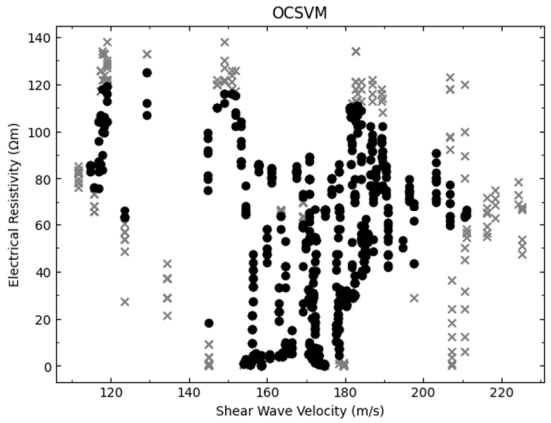
Profile: fc_site3

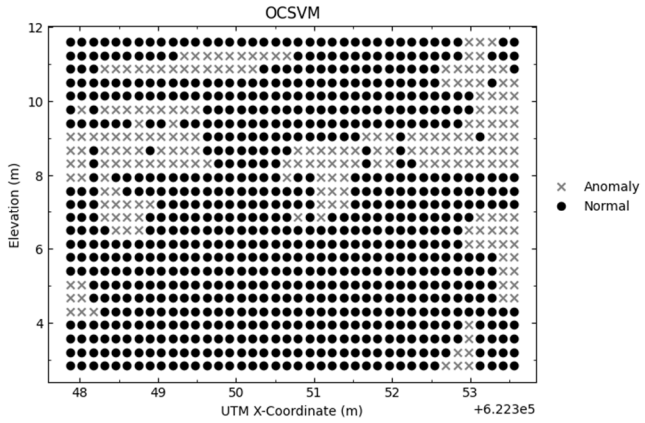




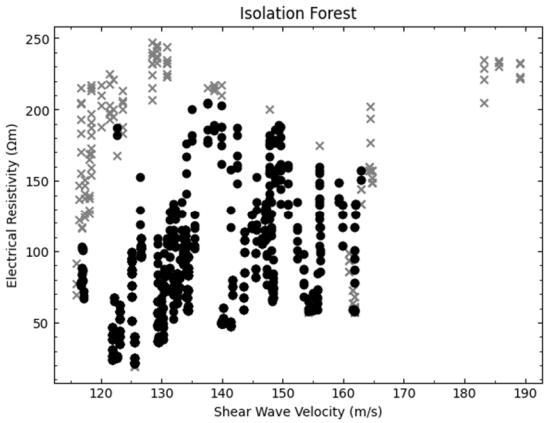


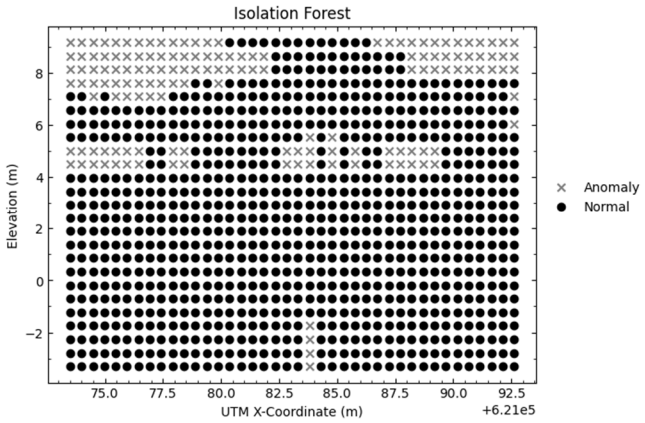


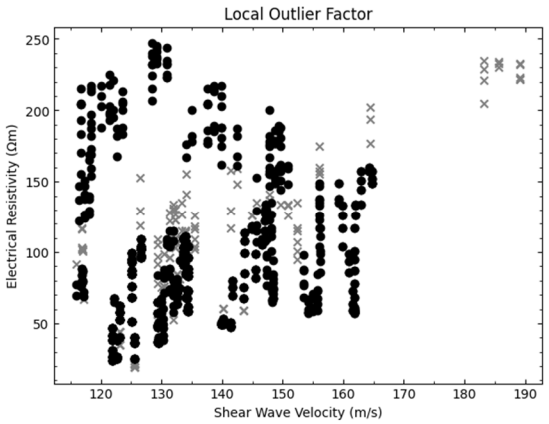


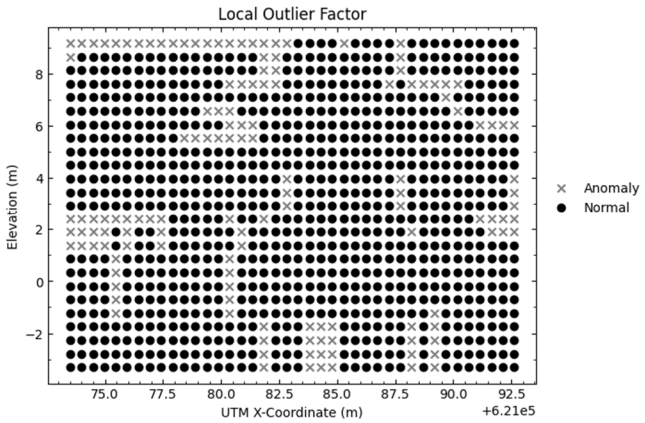


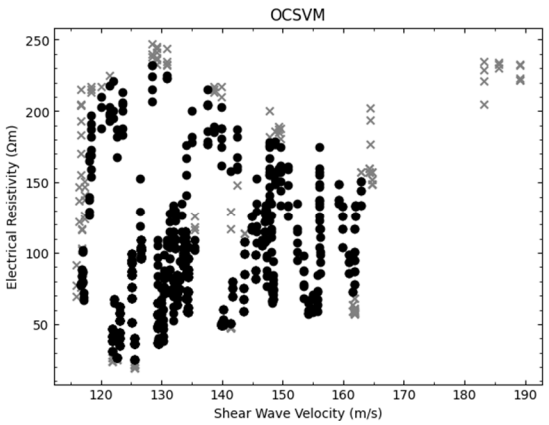
Profile: fc3_site1

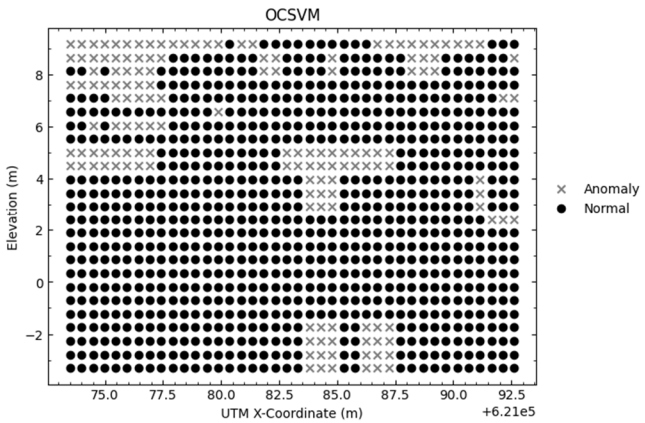




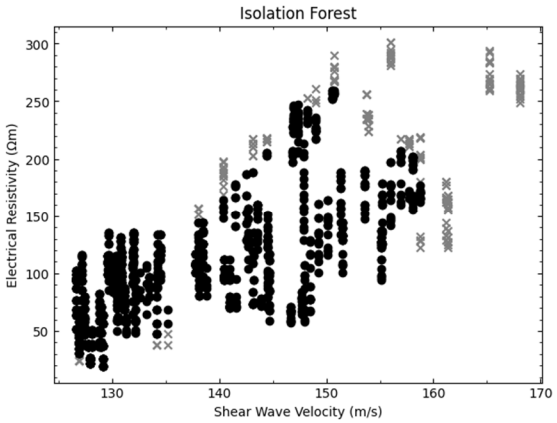


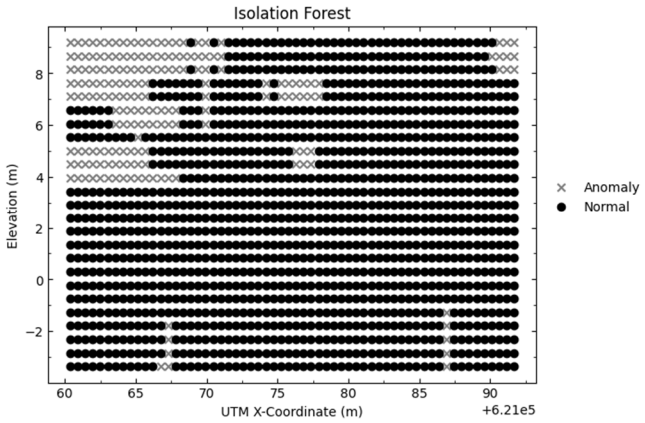


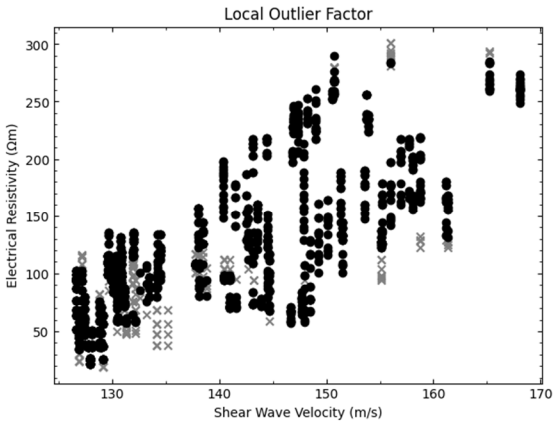


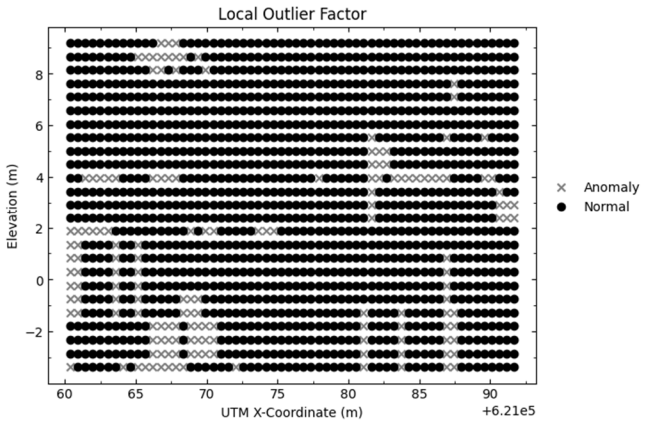


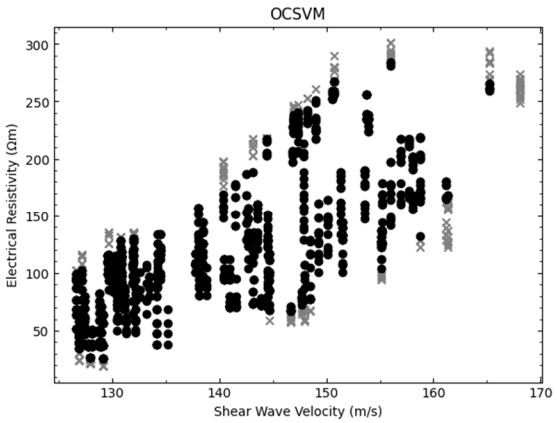
Profile: fc5_site1

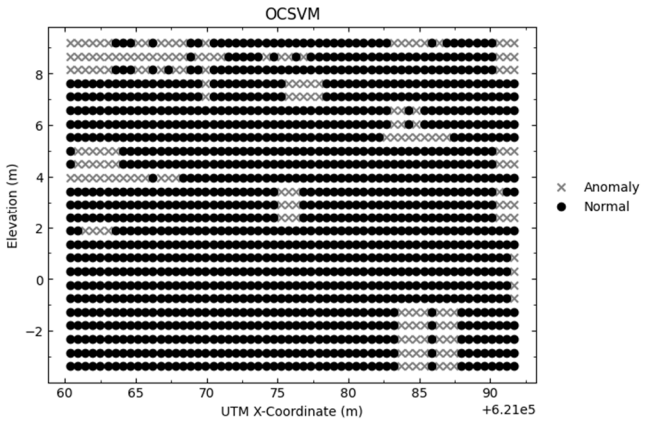




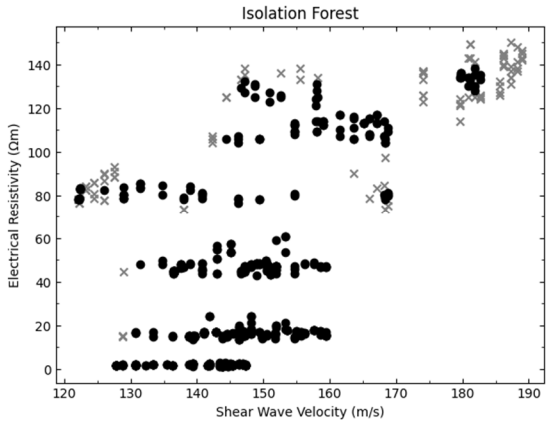


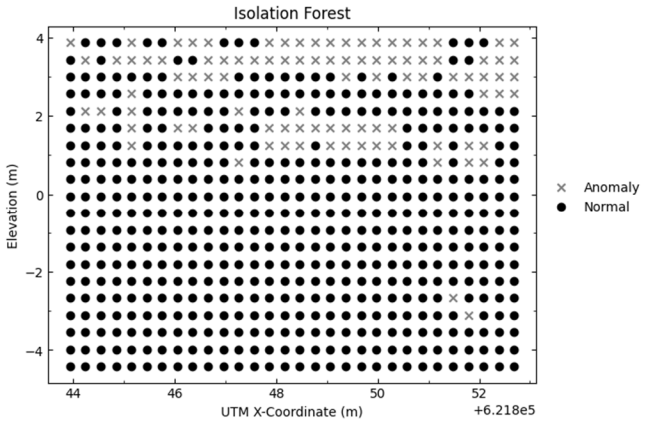


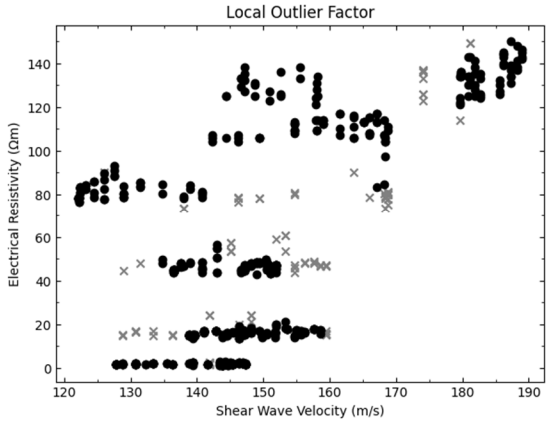


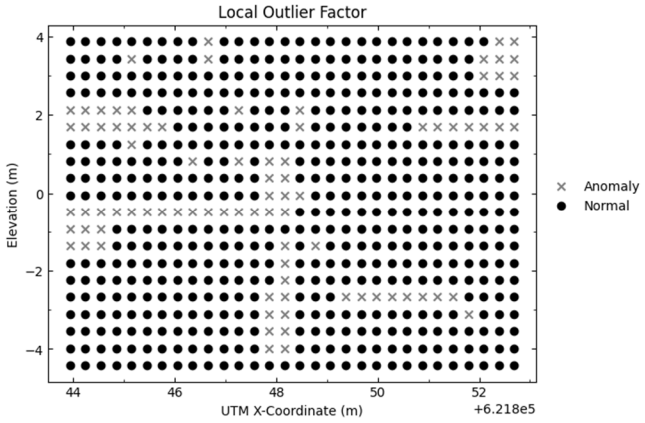


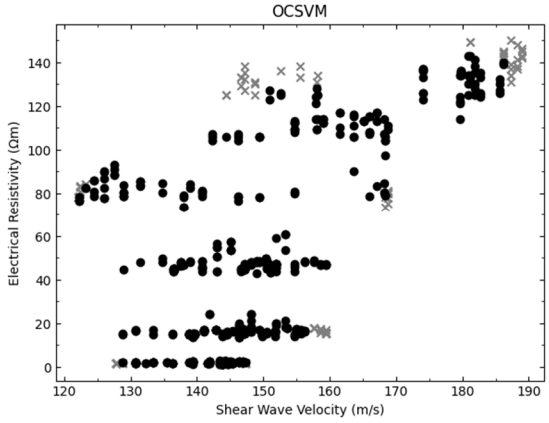
Profile: fdt_site2

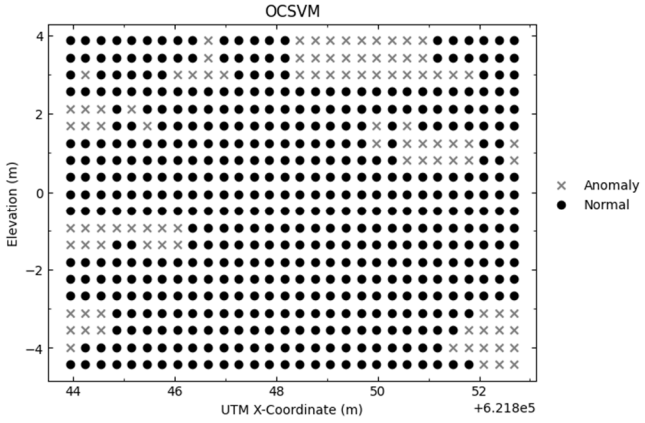




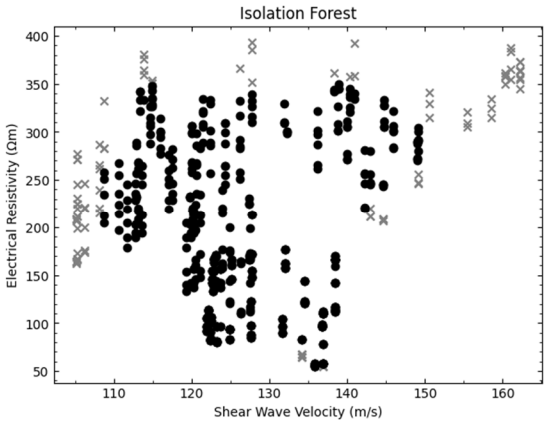


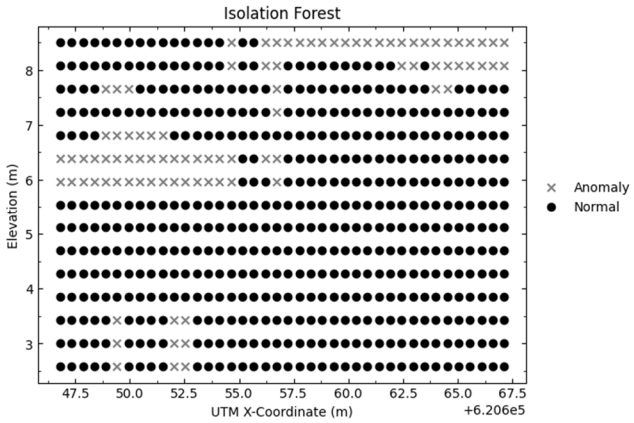


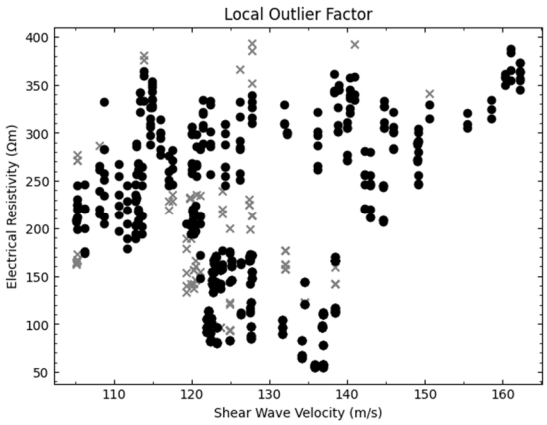


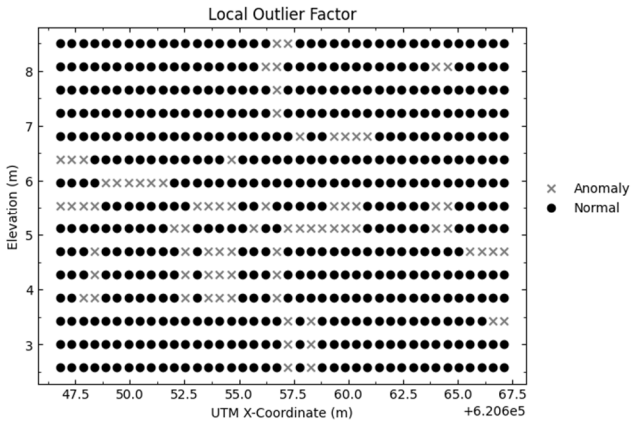


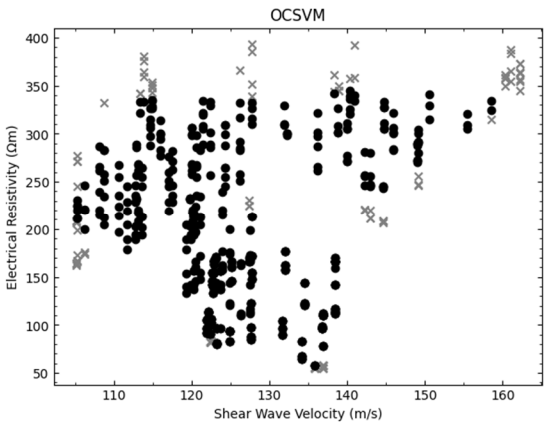
Profile: fwt_site1

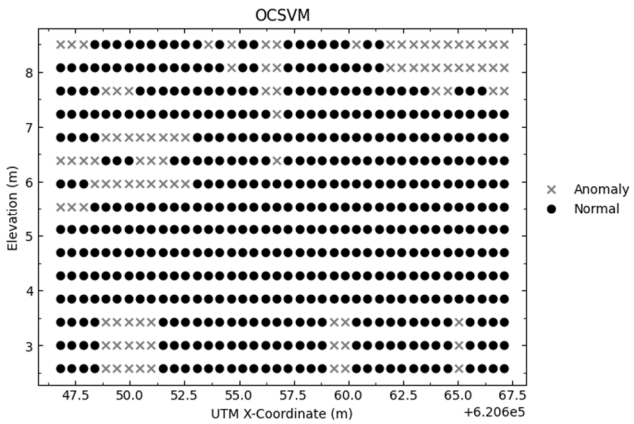




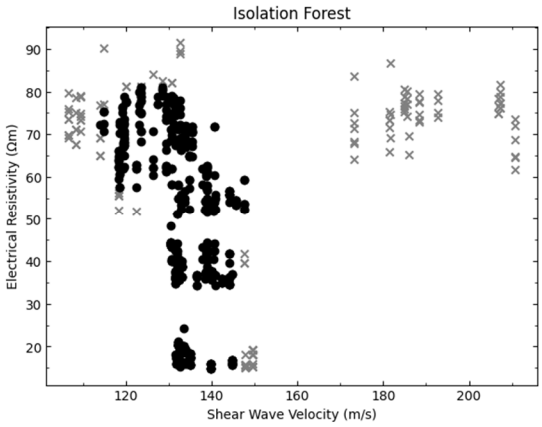


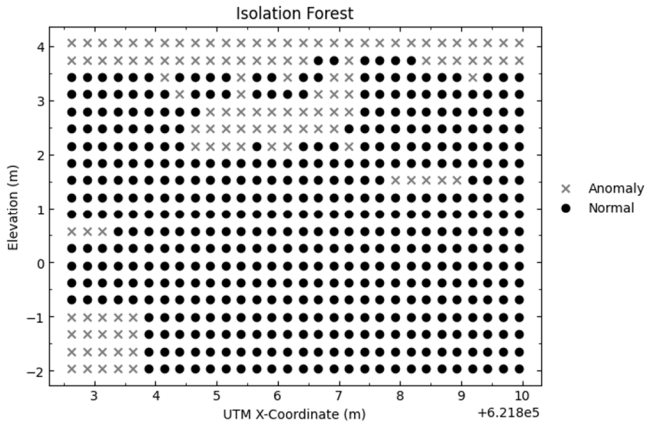


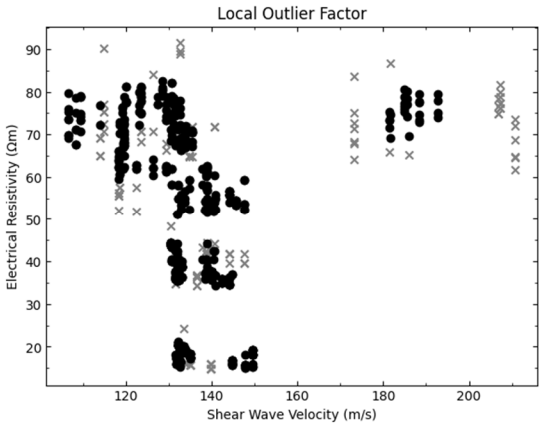


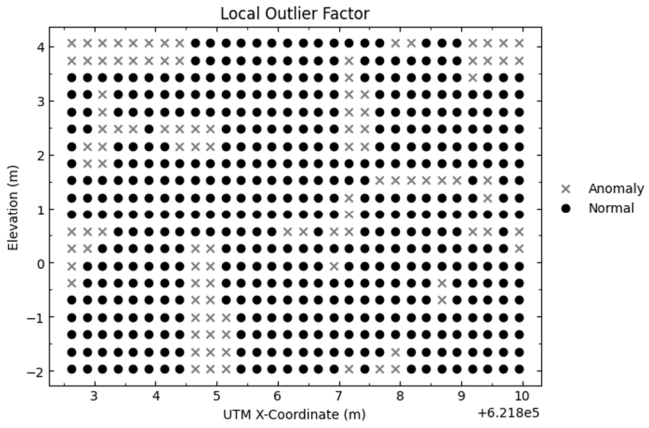


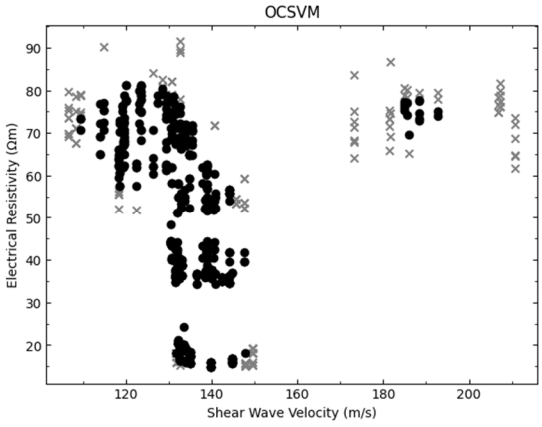
Profile: fwt_site2

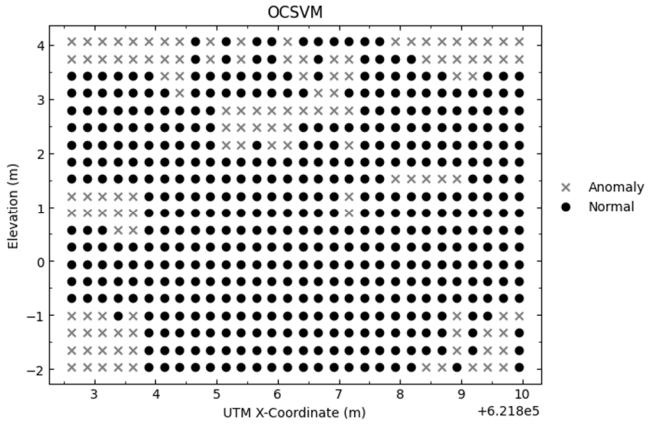




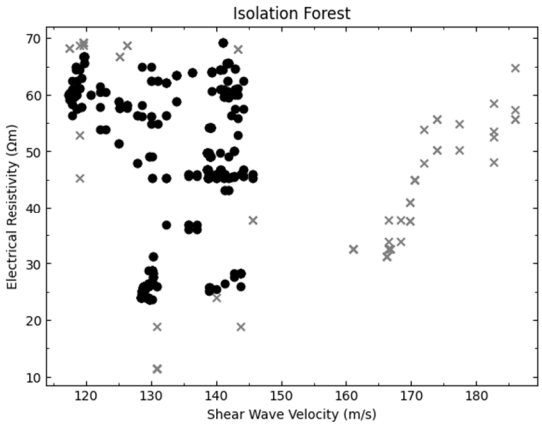


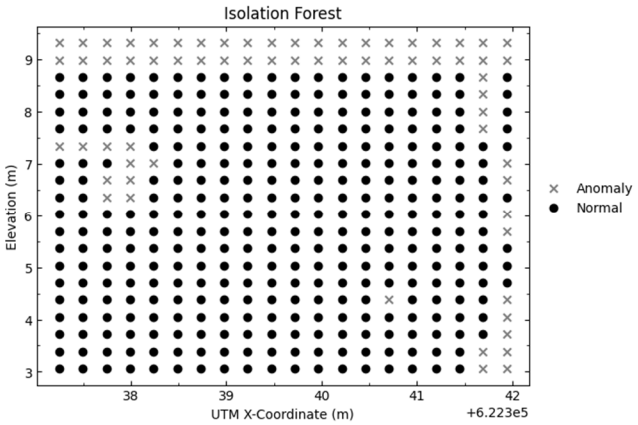


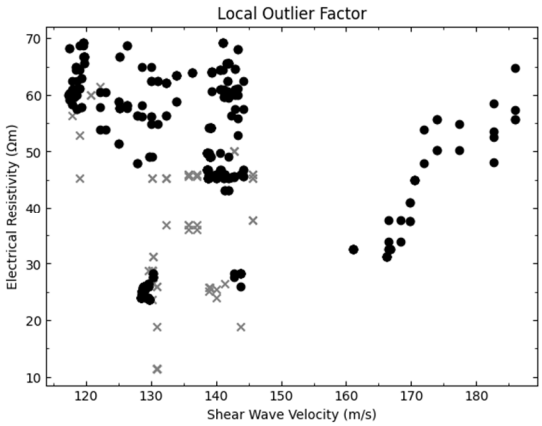


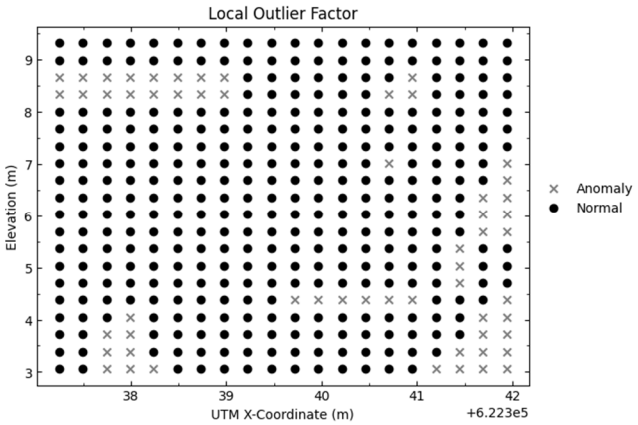


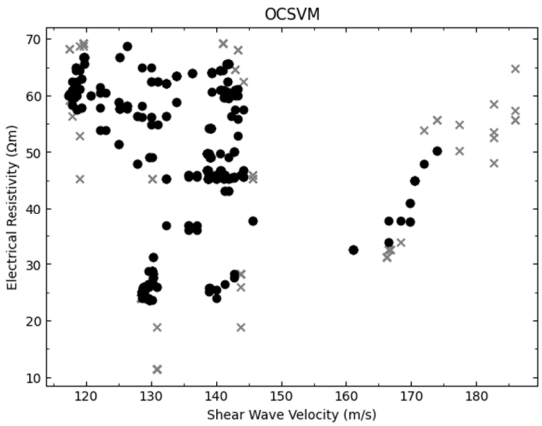
Profile: fwt_site3

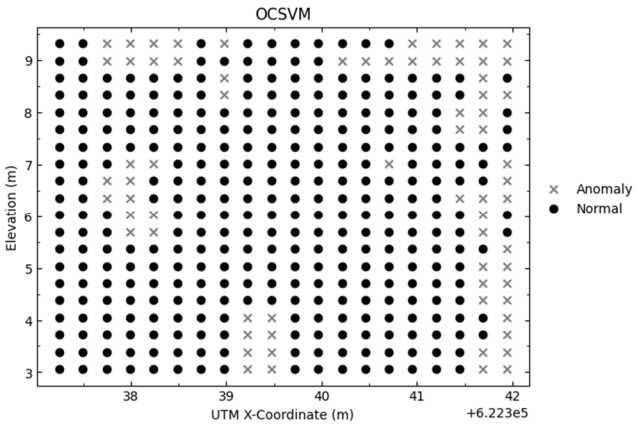




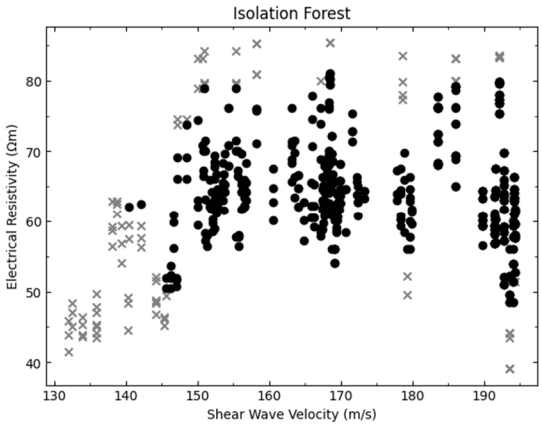


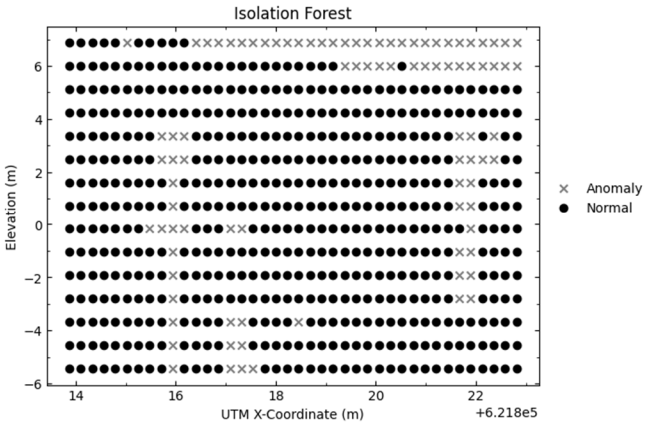


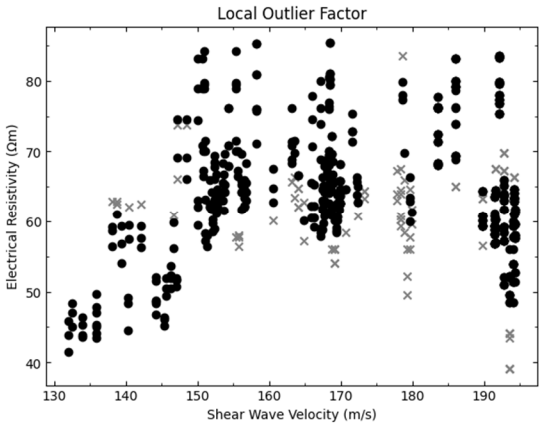


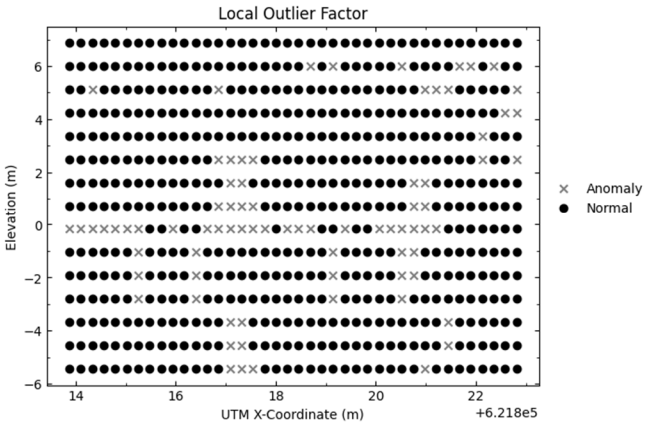


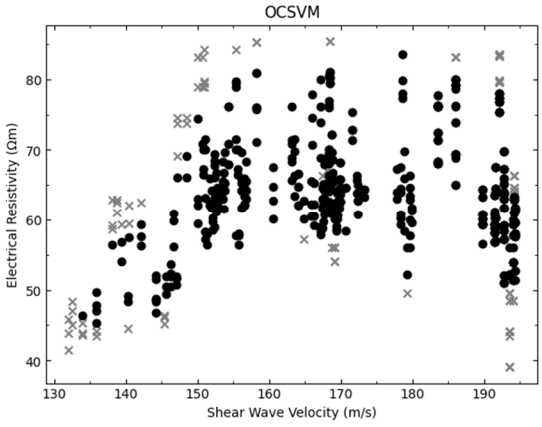
Profile: sc_site2

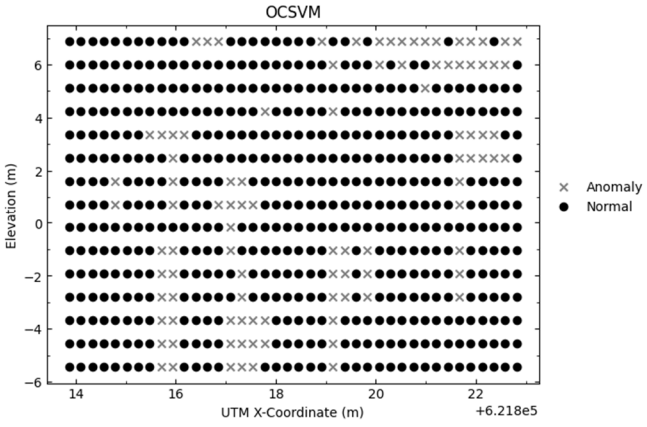




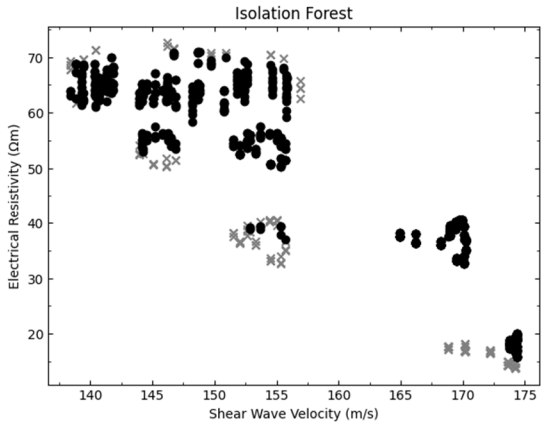


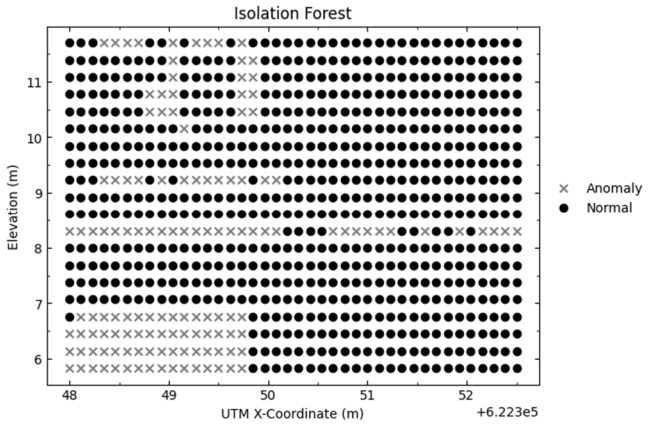


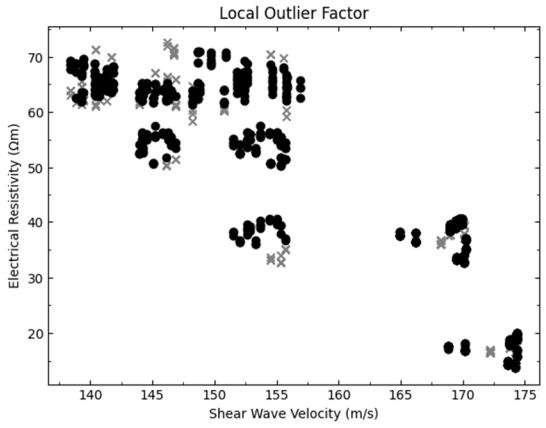


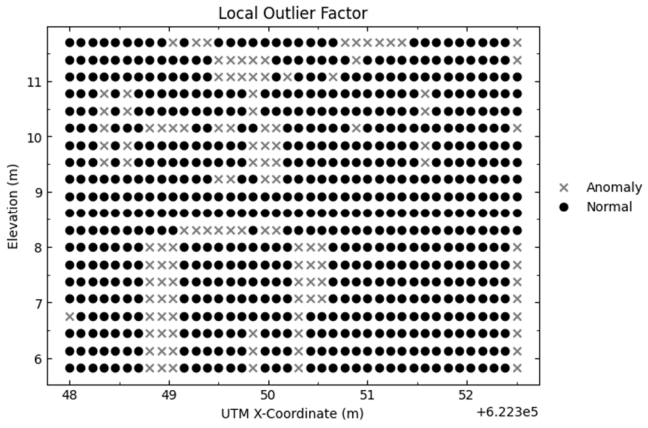


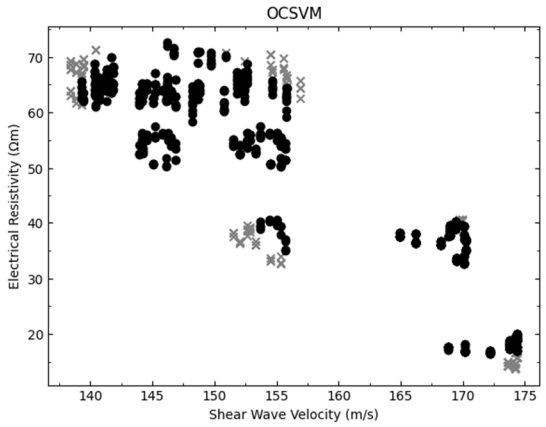
Profile: sc_site3

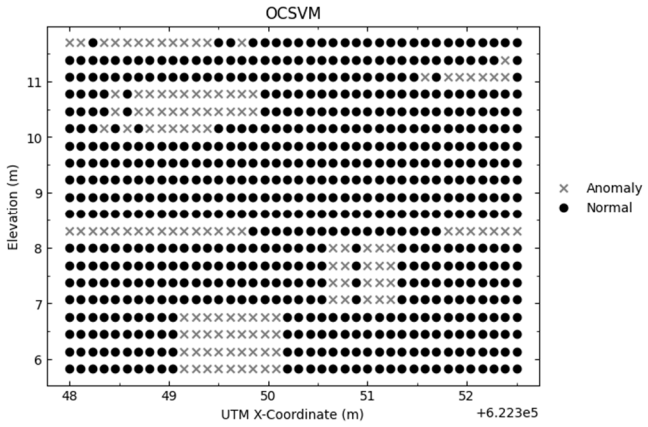




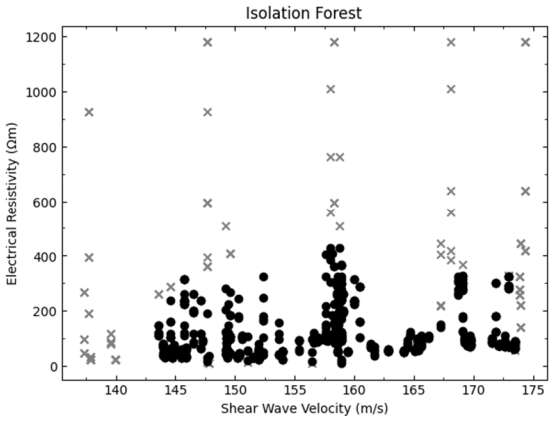


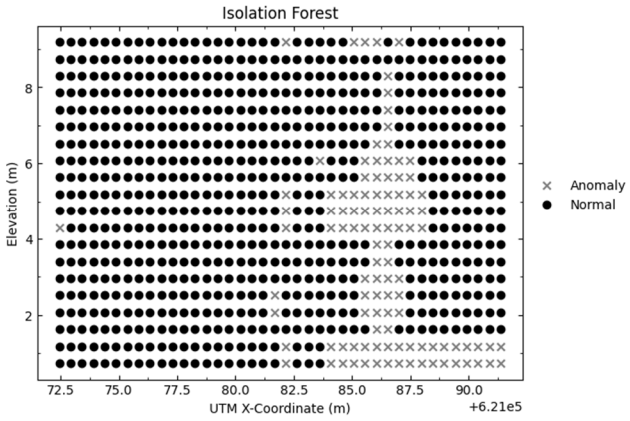


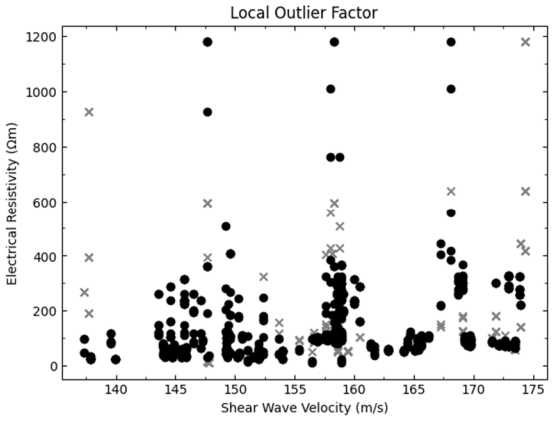


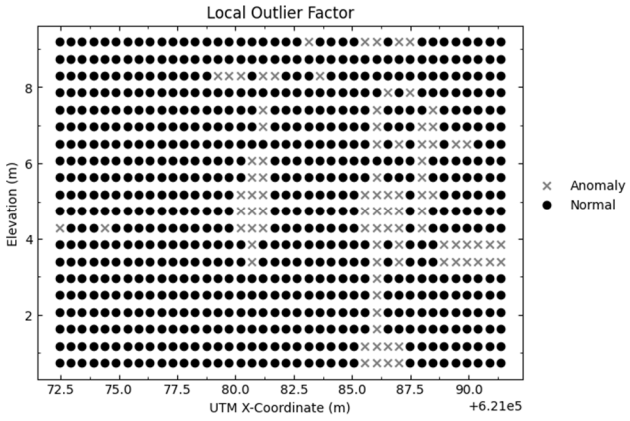


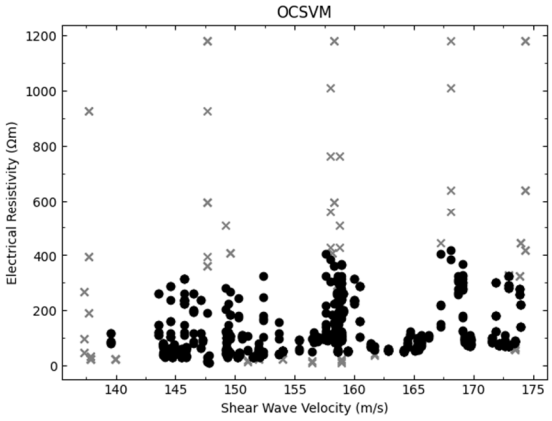
Profile: sc3_site1

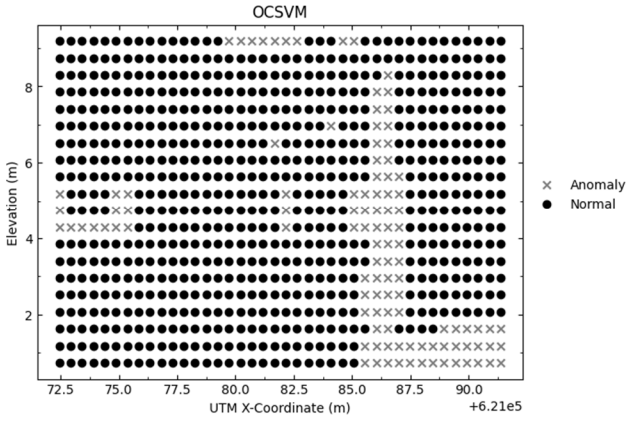




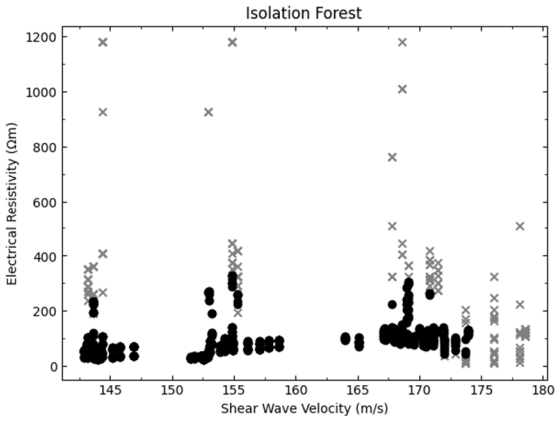


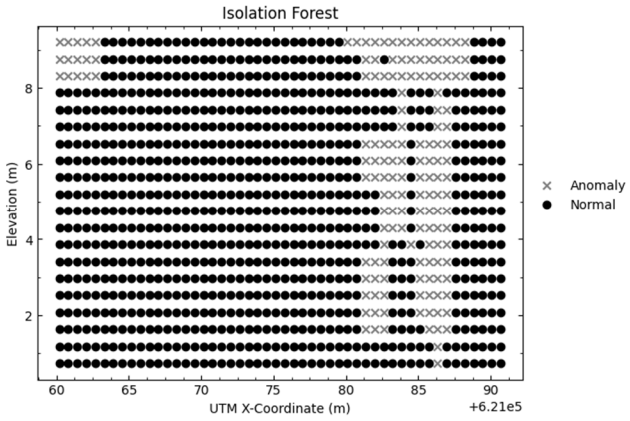


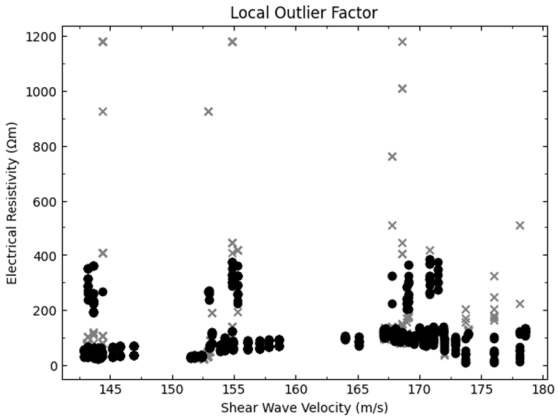


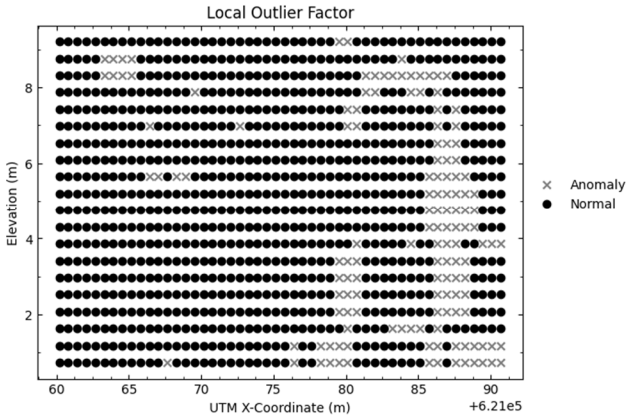


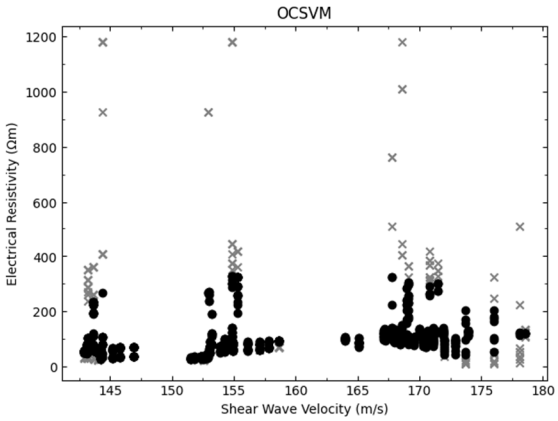
Profile: sc5_site1

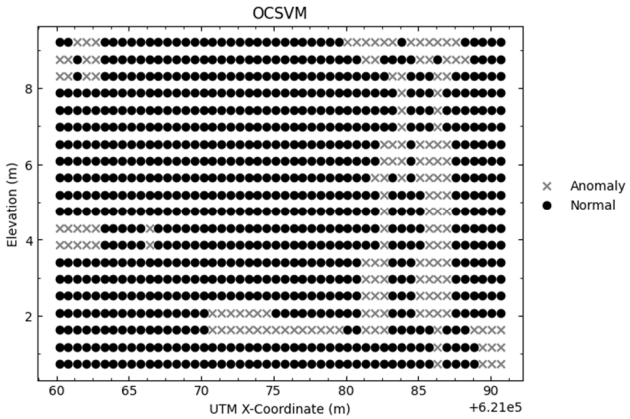




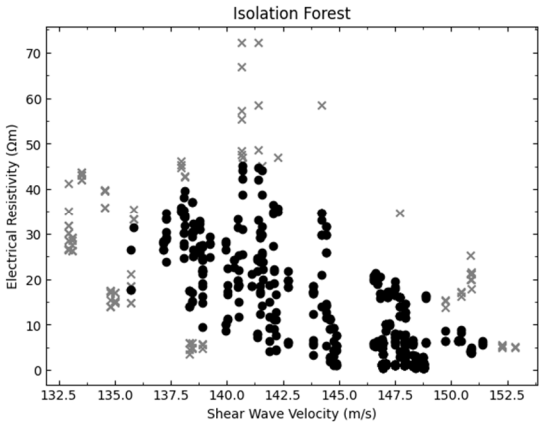


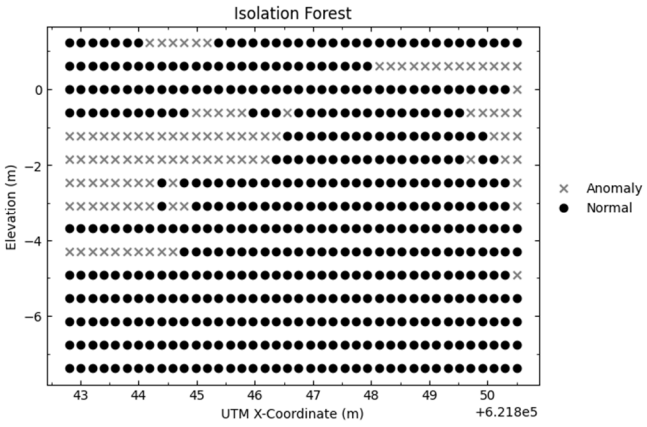


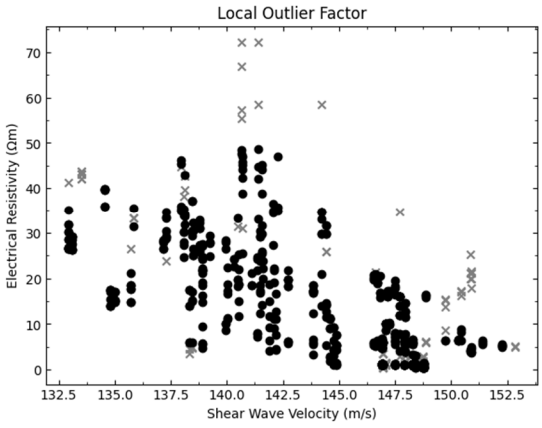


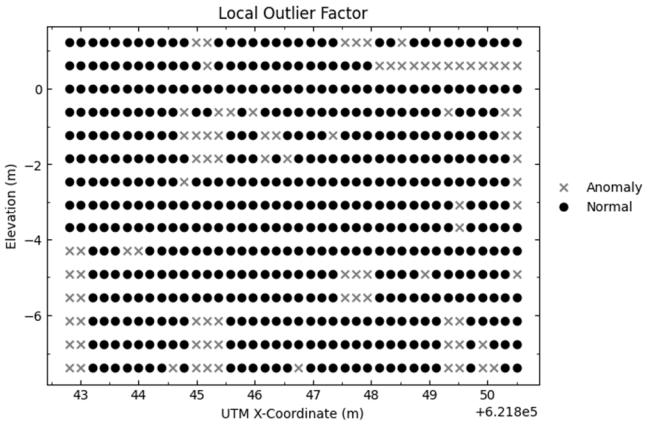


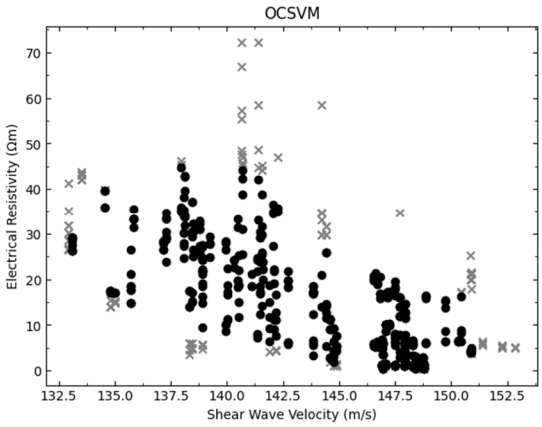
Profile: sdt_site2

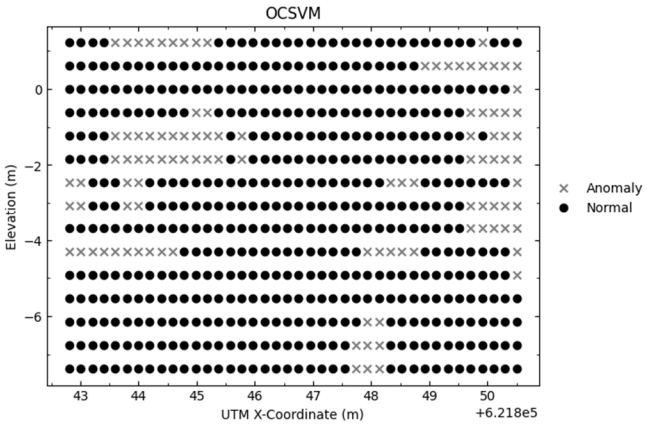




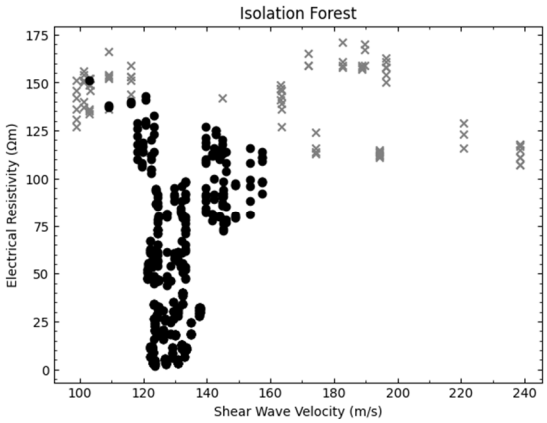


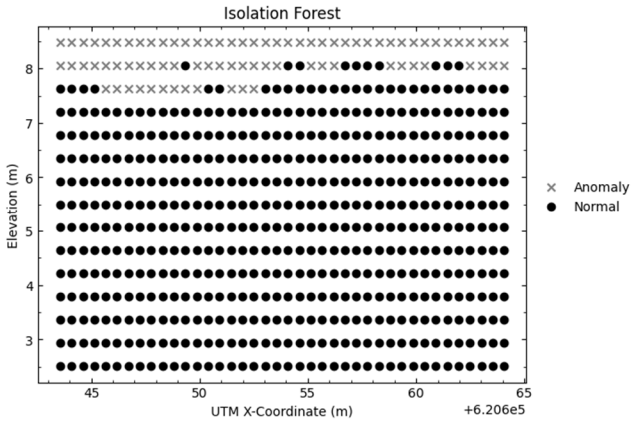


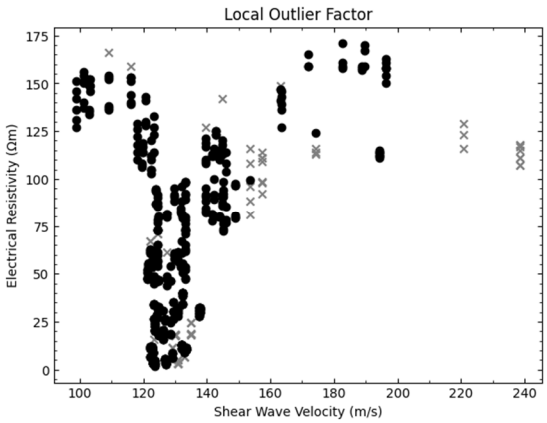


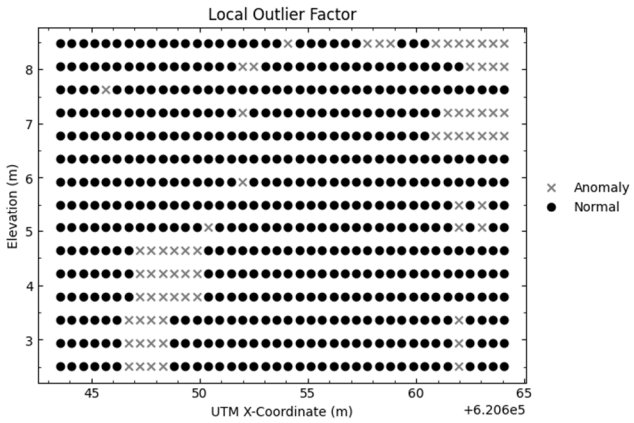


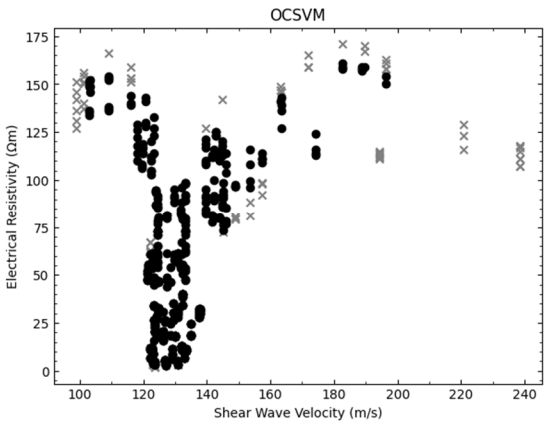
Profile: swt_site1

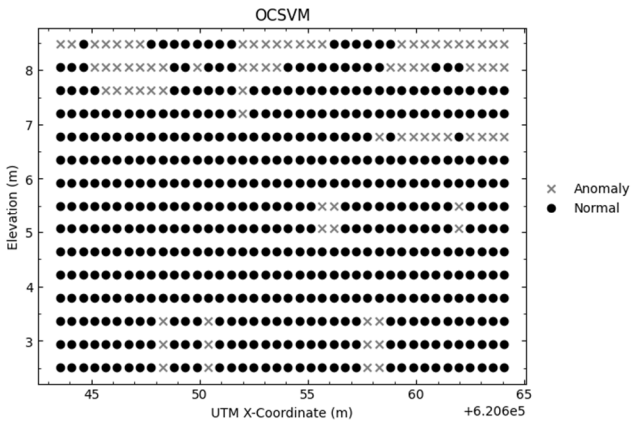




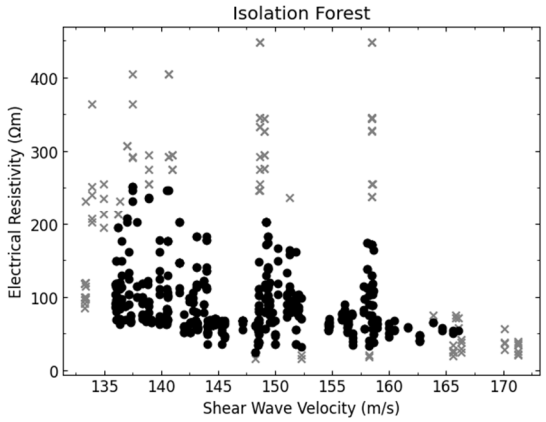


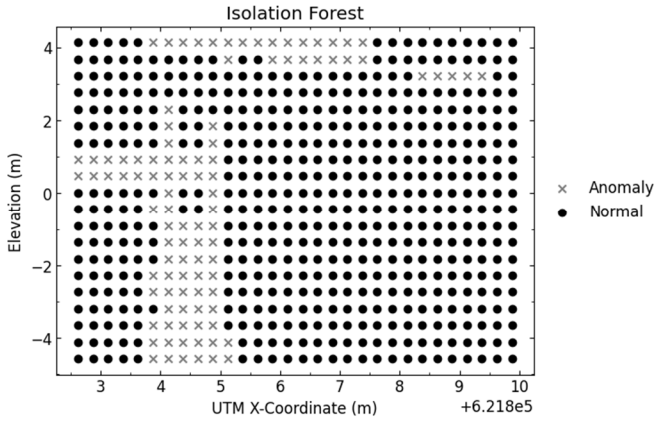


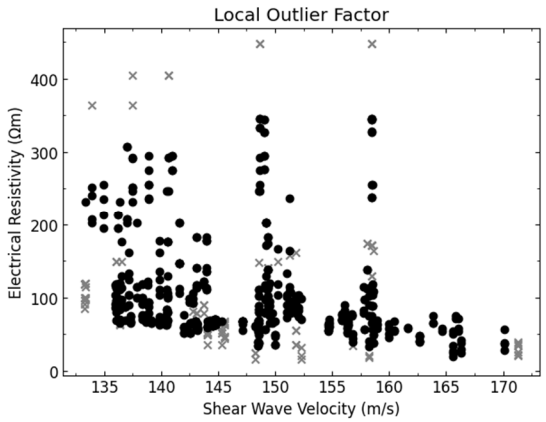


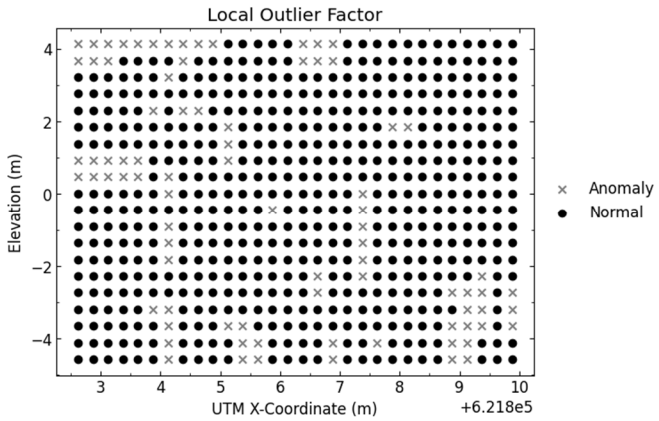


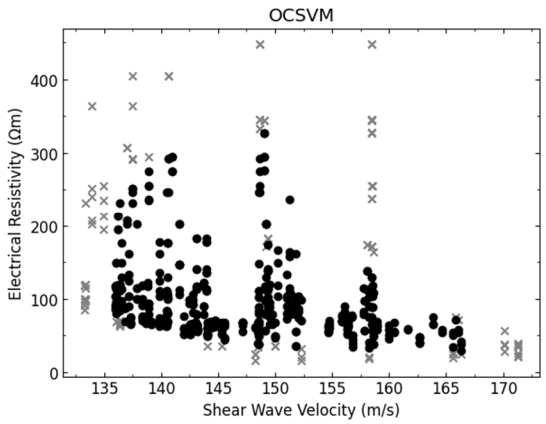
Profile: swt_site2

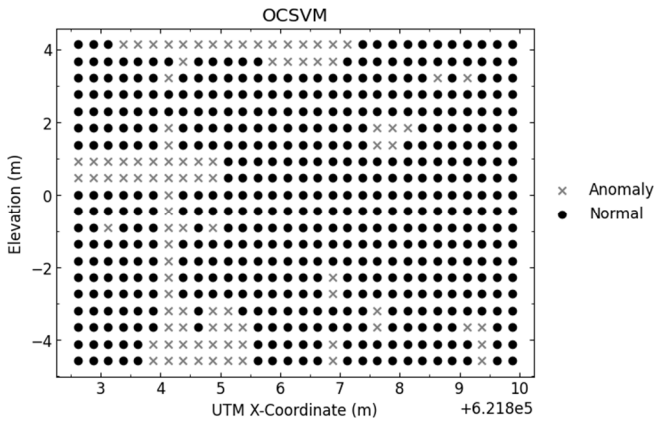




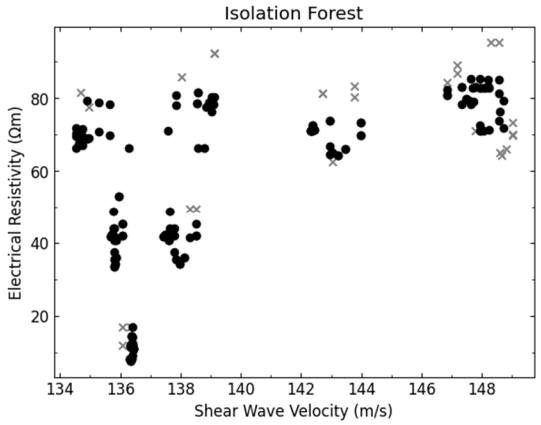


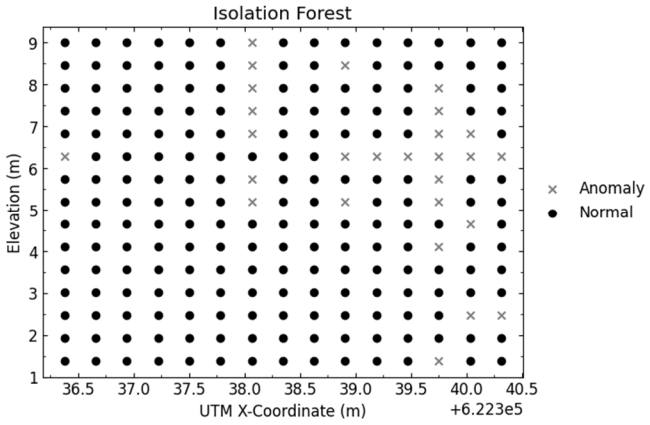


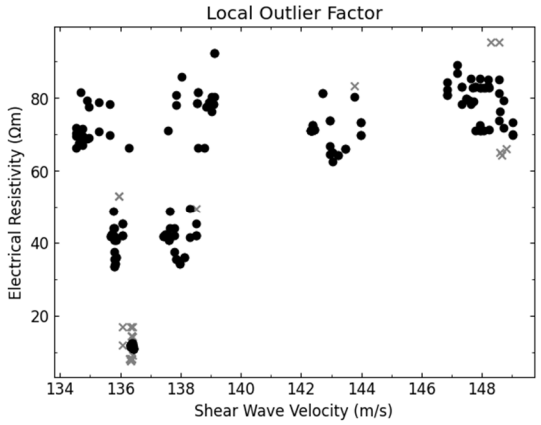


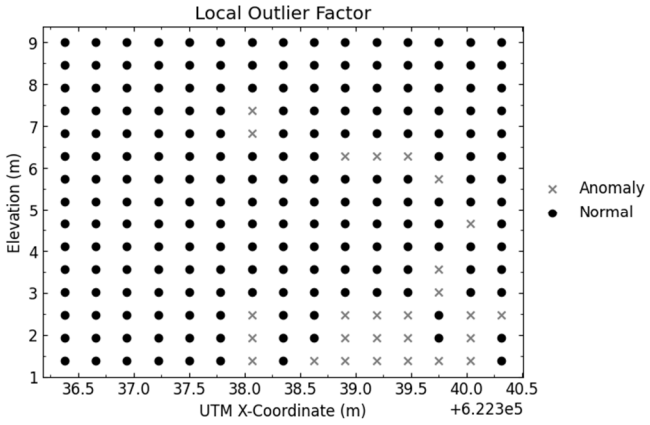


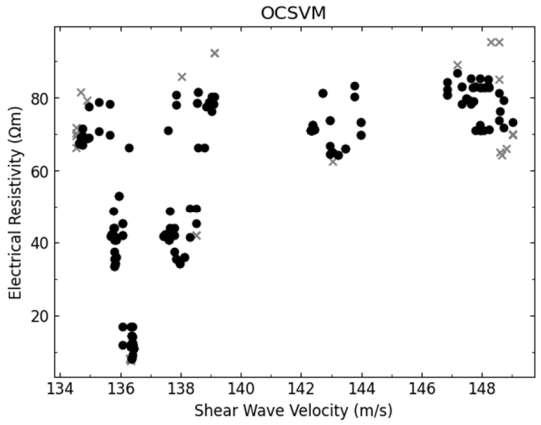
Profile: swt_site3

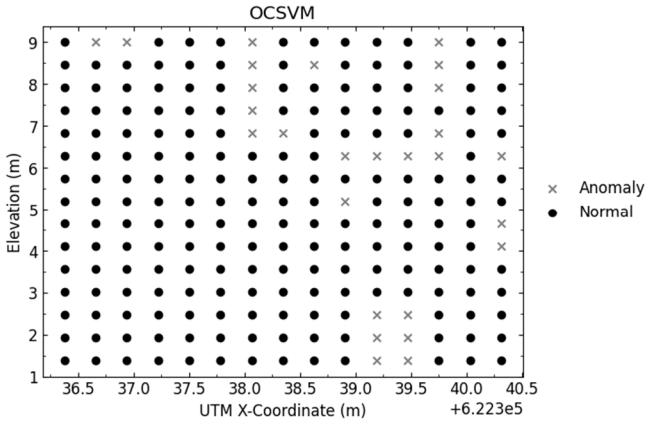






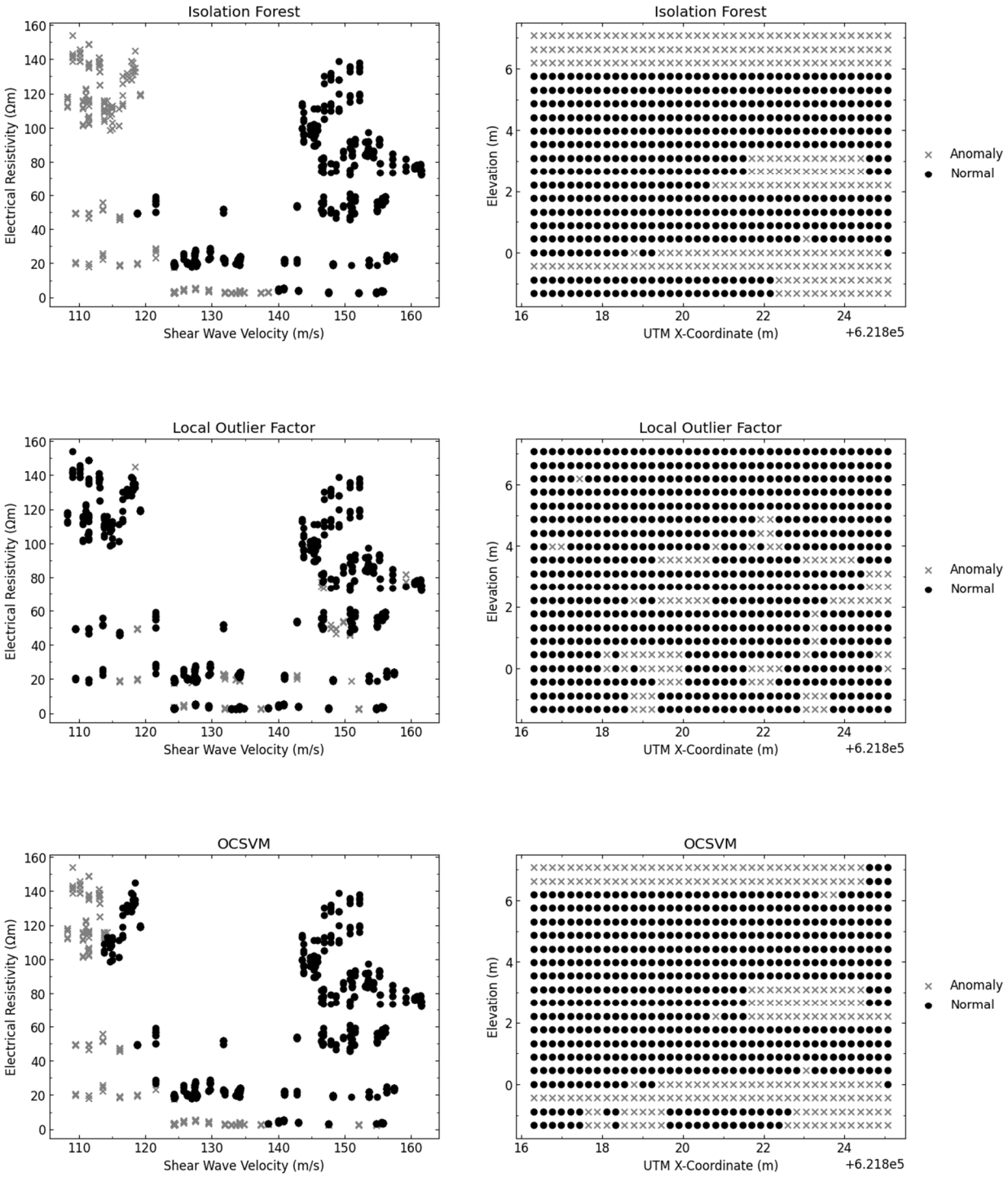




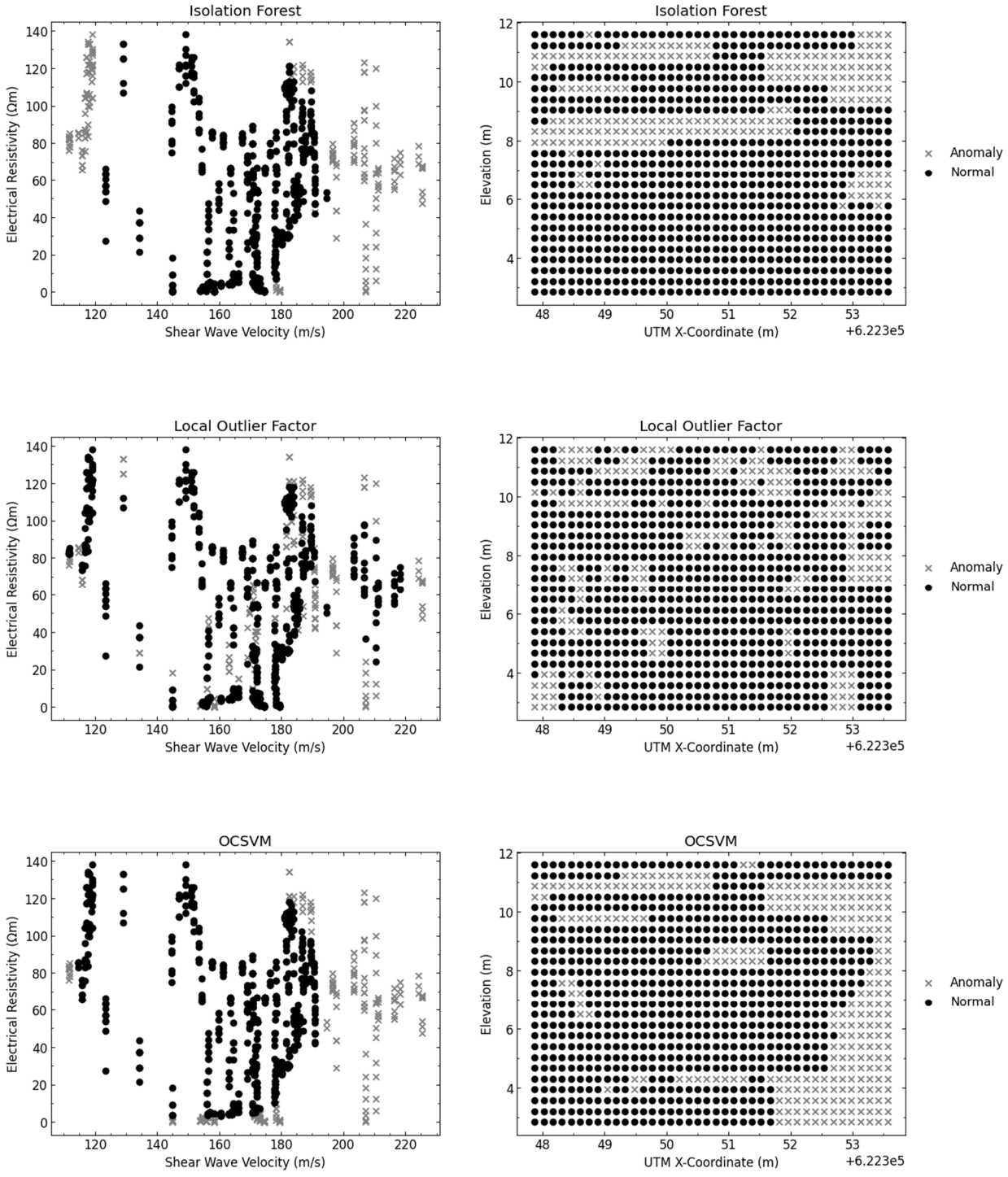


Fit: Crest Compilation Dataframe
Predicted: Crest Compilation Dataframe

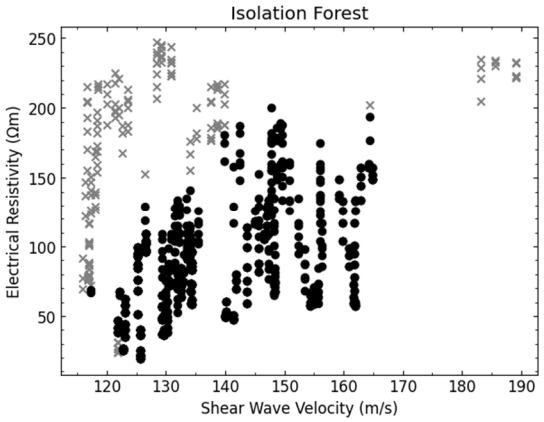
Profile: fc_site2

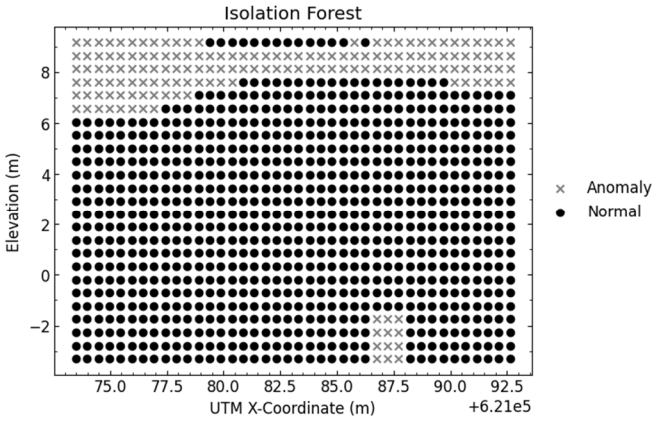


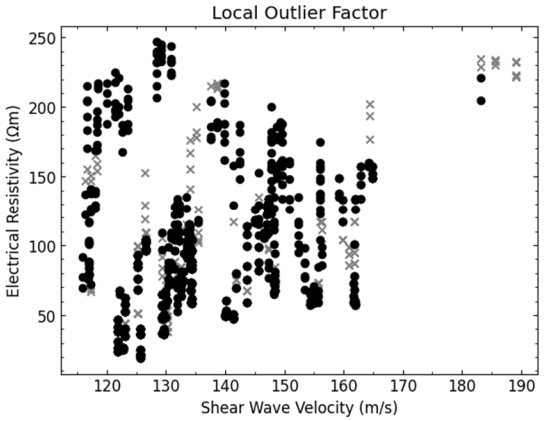
Profile: fc_site3

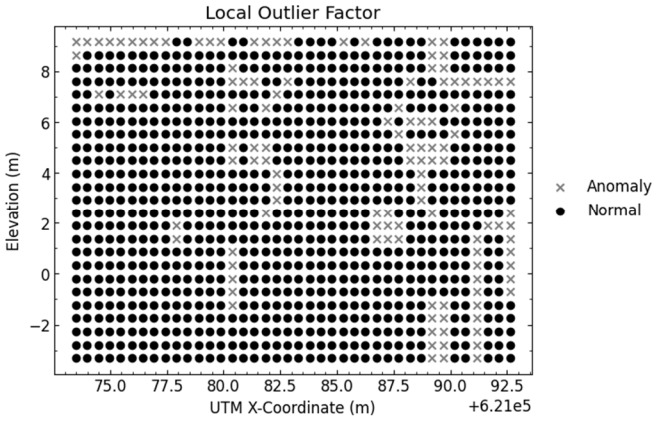


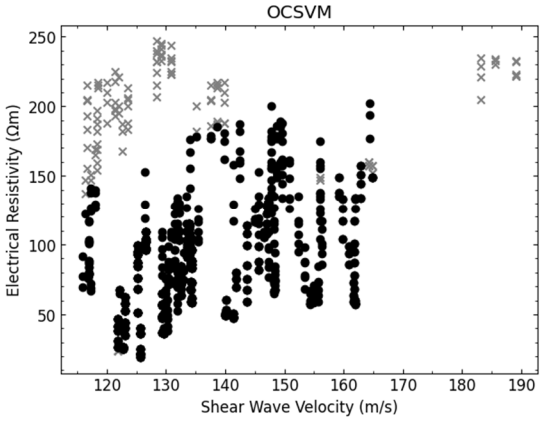
Profile: fc3_site1

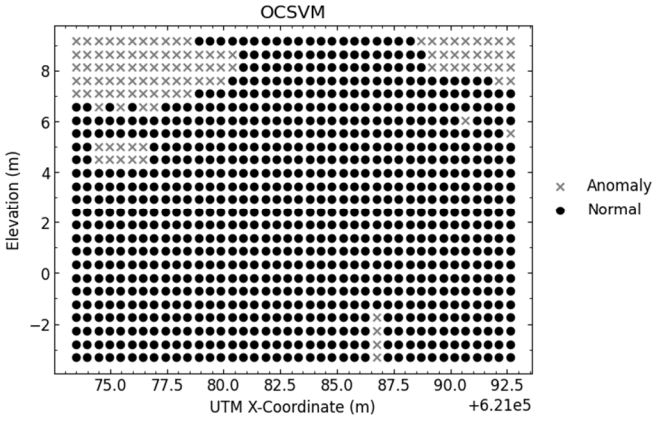




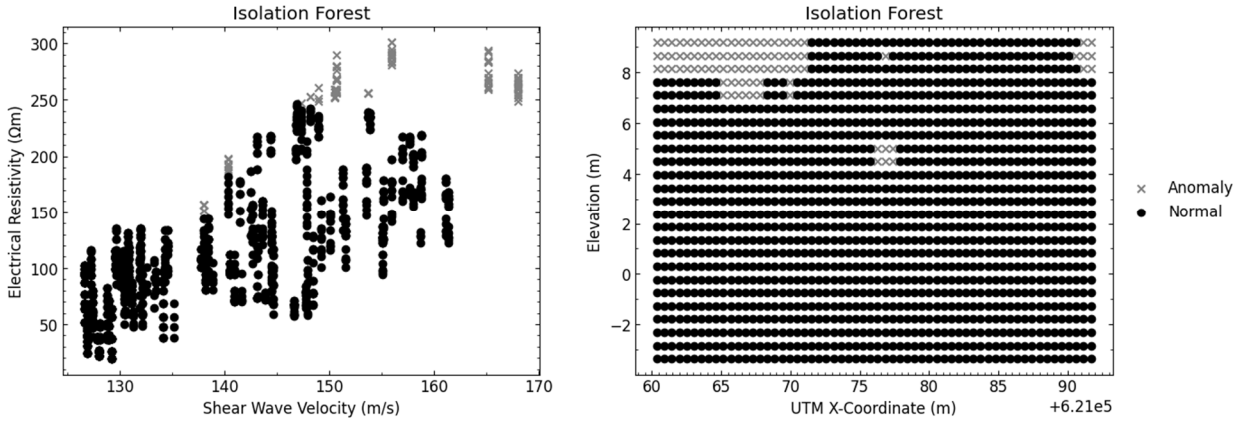


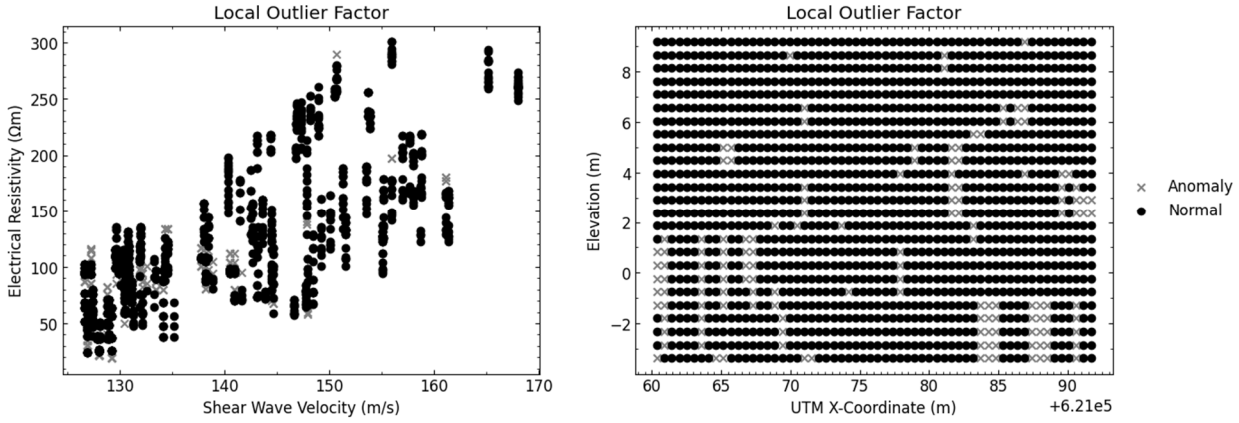


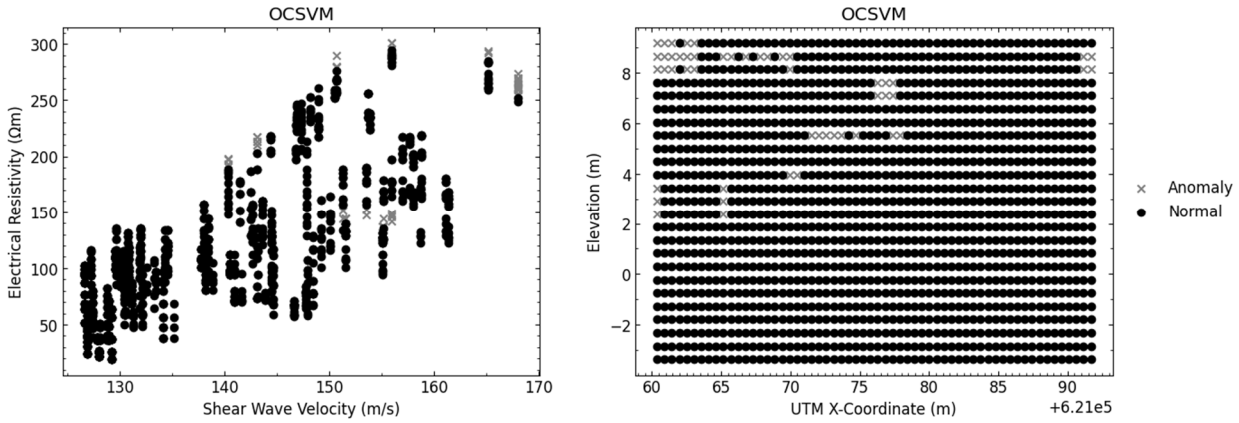




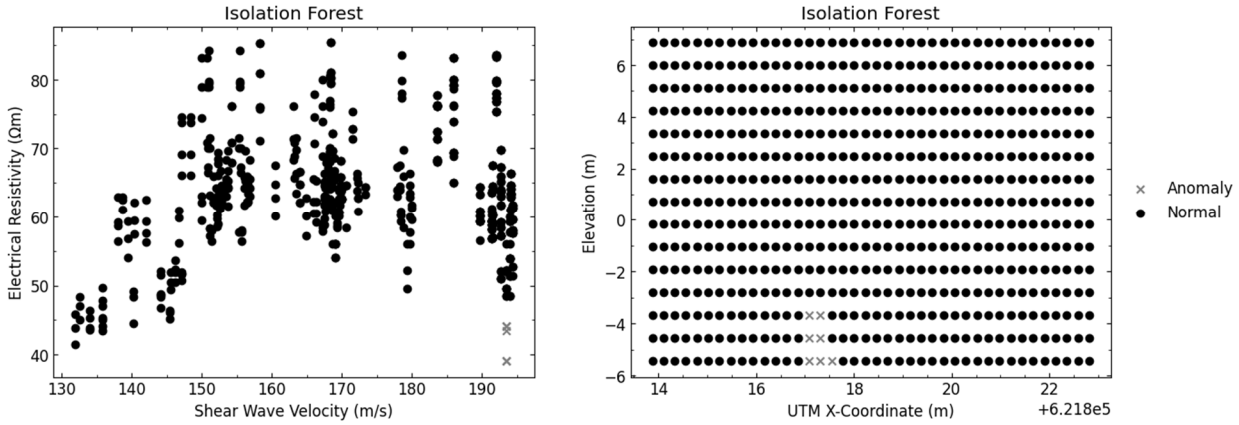
Profile: fc5_site2

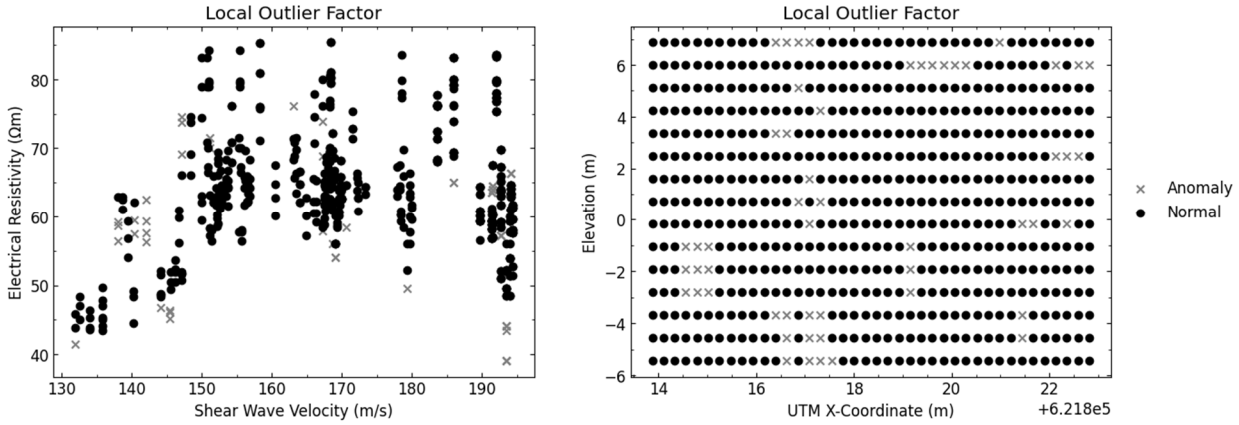


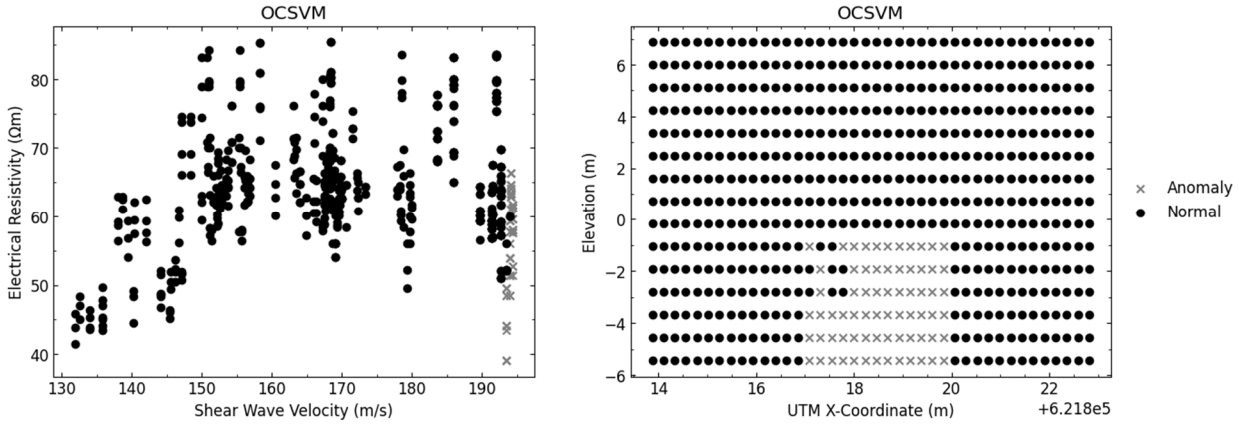




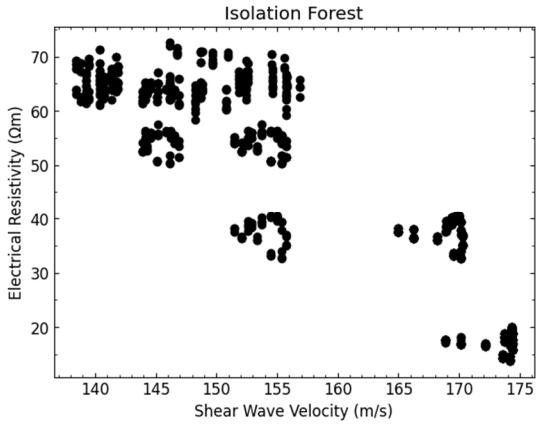
Profile: sc_site2

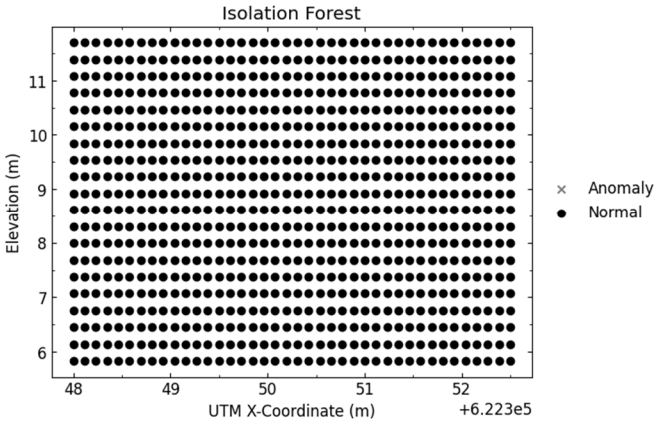


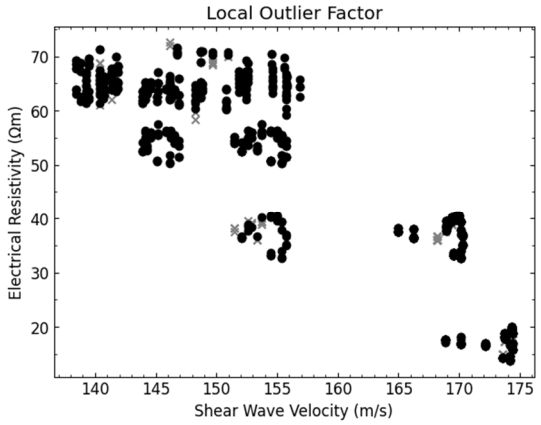


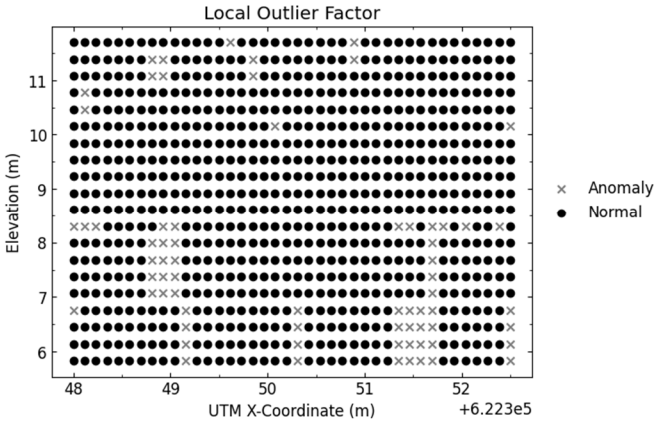


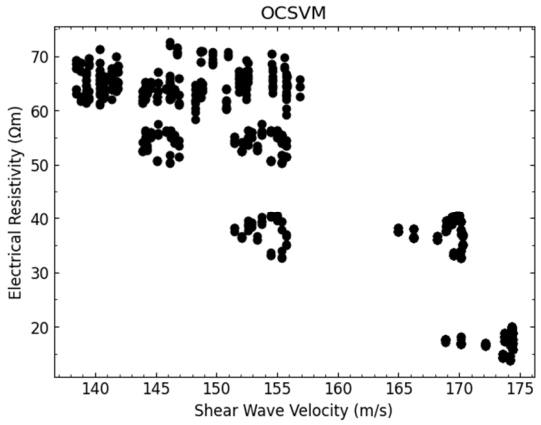
Profile: sc_site3

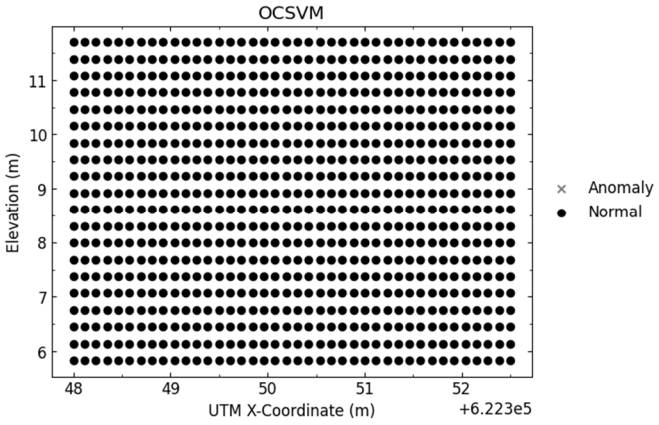




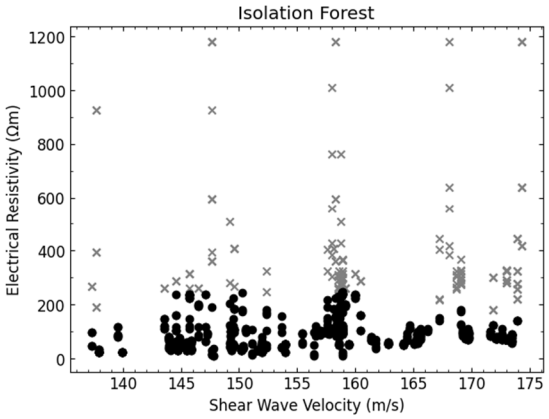


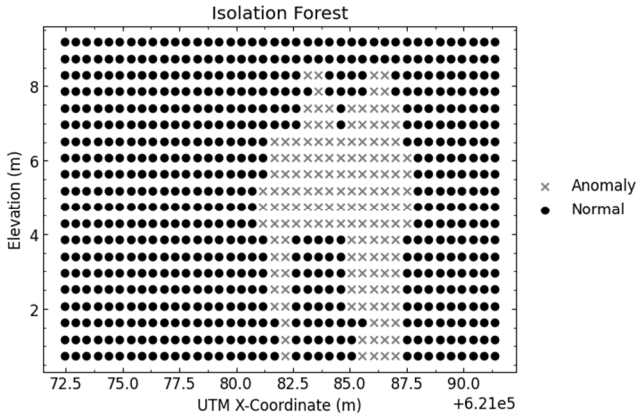


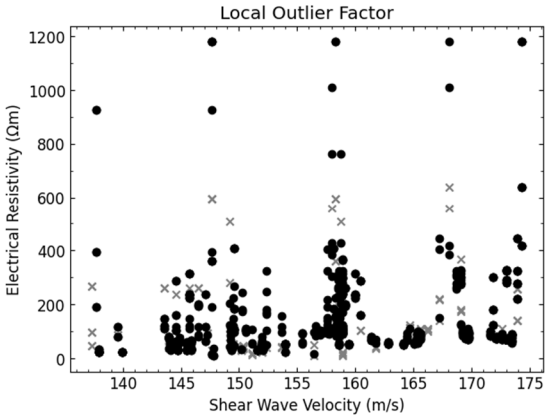


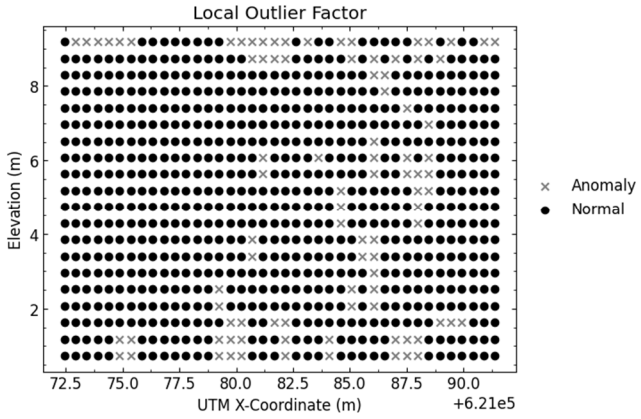


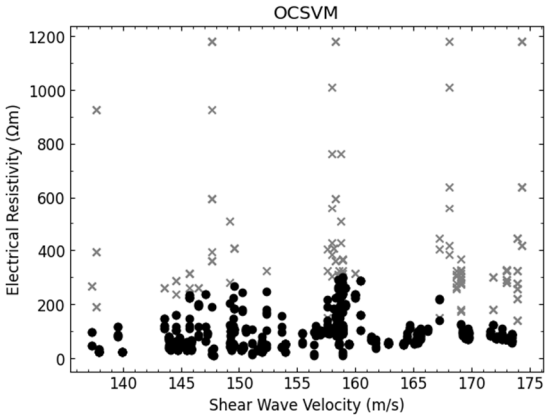
Profile: sc3_site1

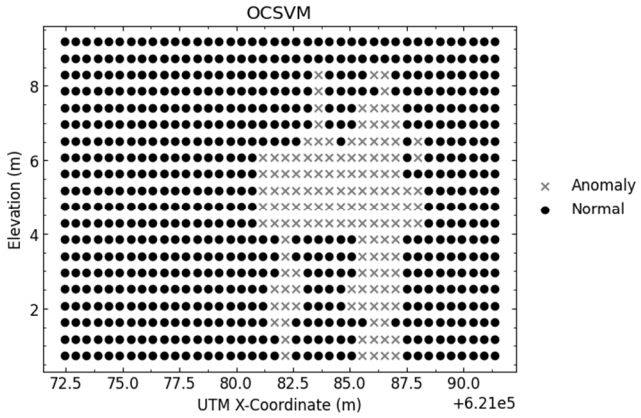




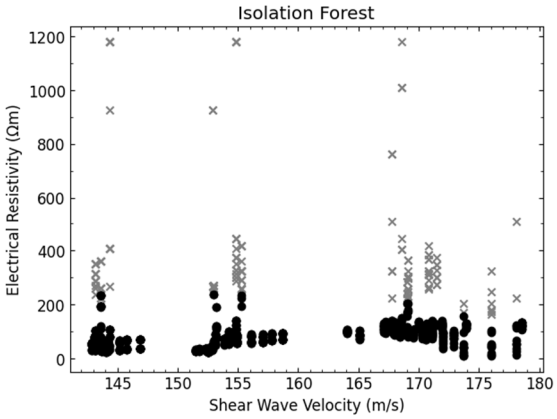


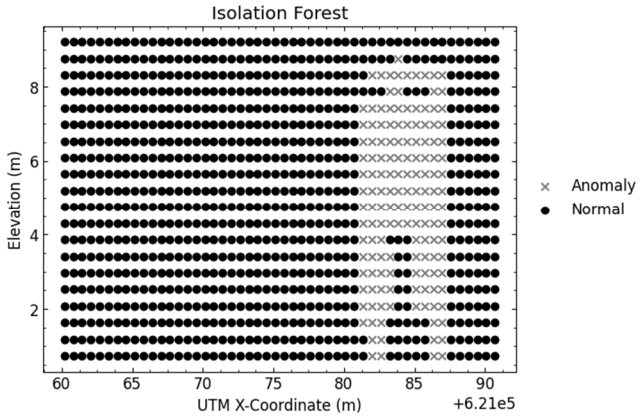


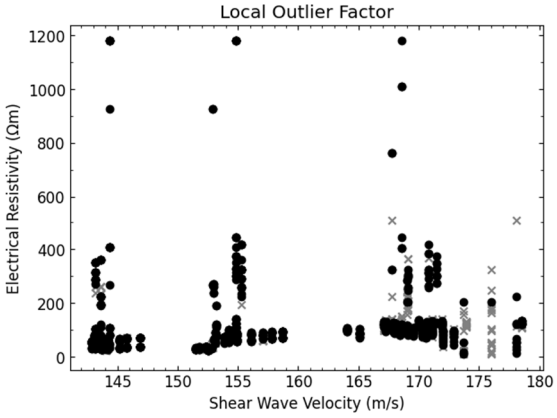


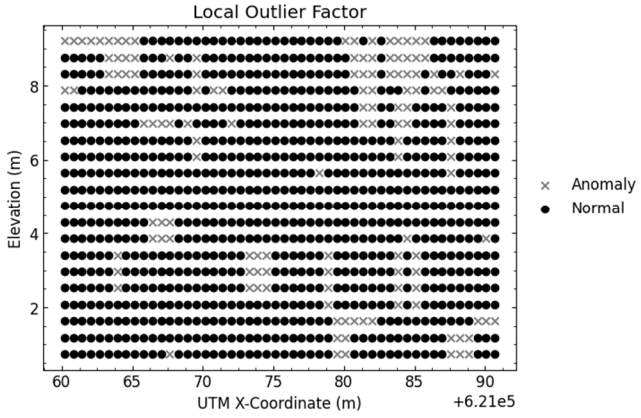


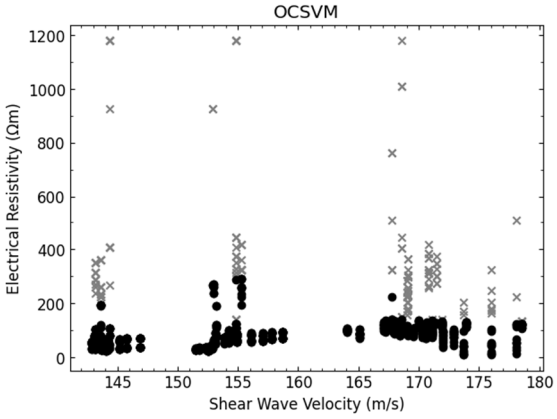
Profile: sc5_site1

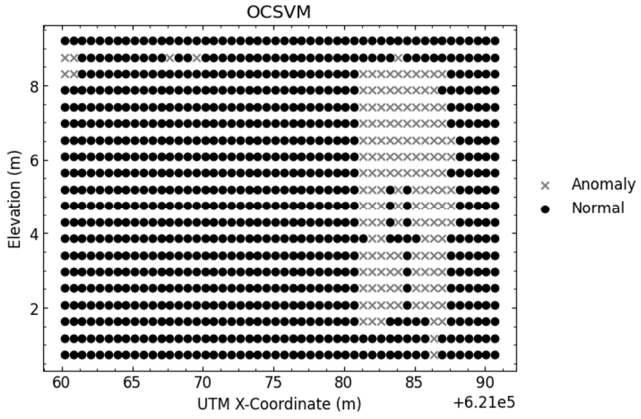






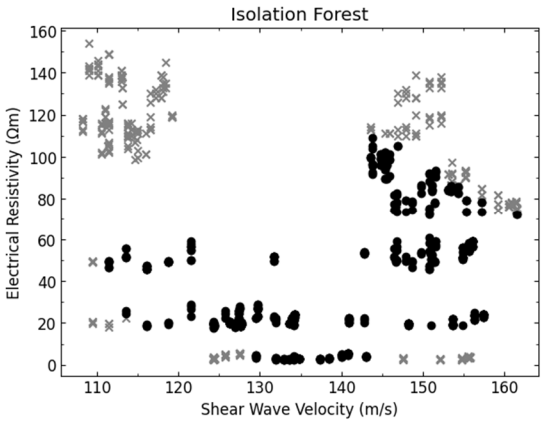


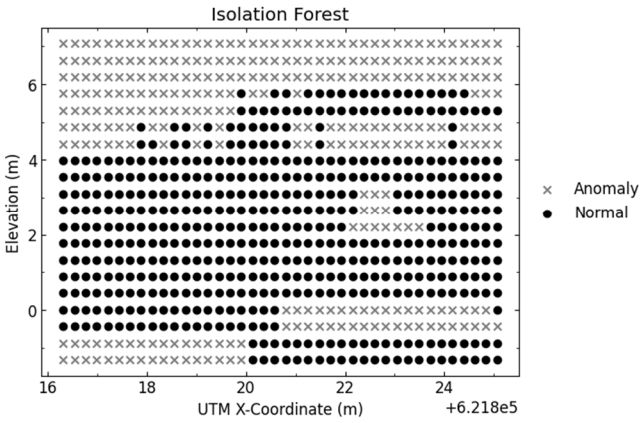


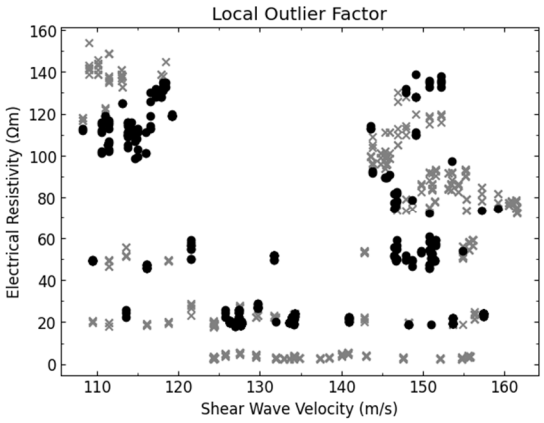


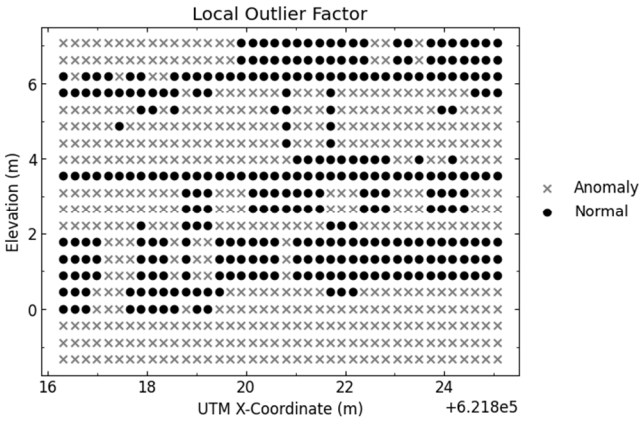
Fit: Crest Compilation Dataframe
Predicted: Individual Profile

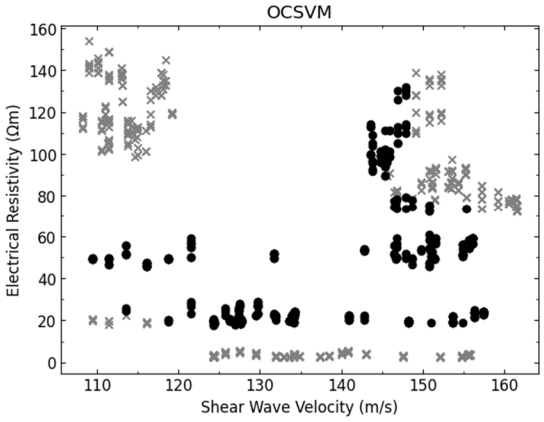
Profile: fc_site2

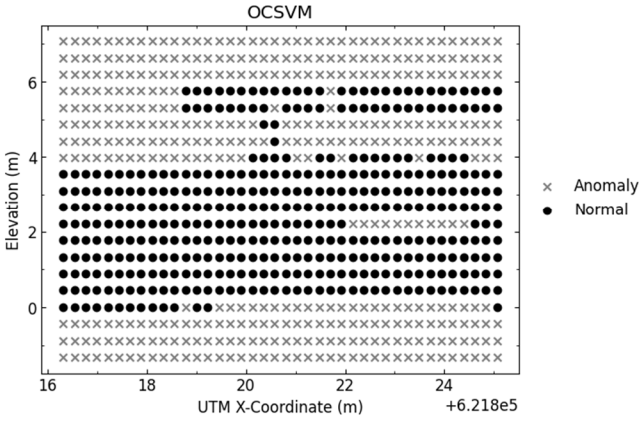




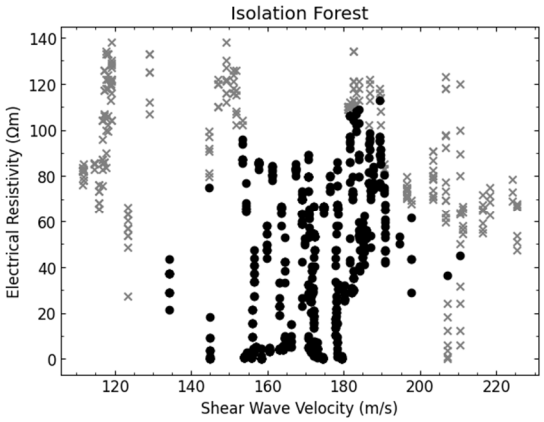


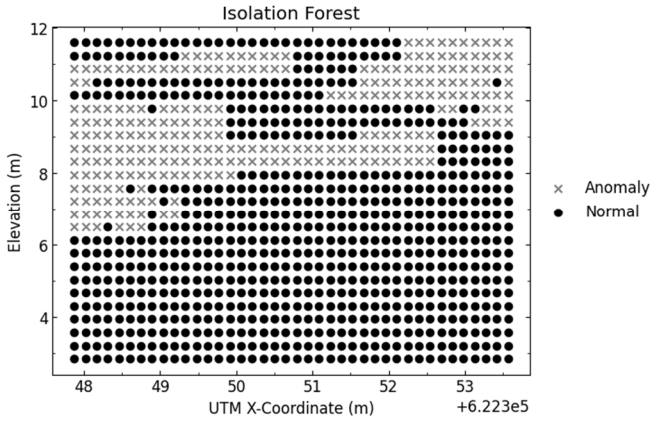


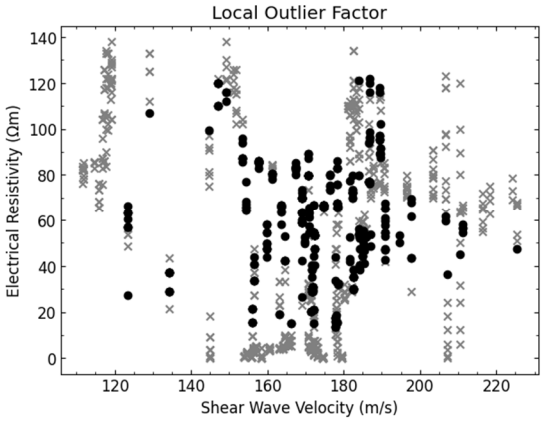


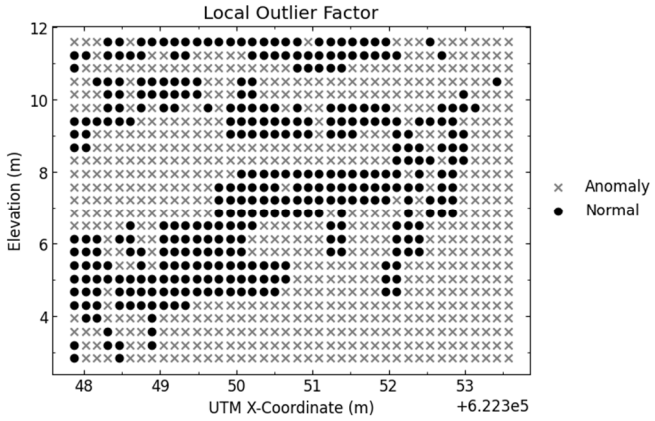


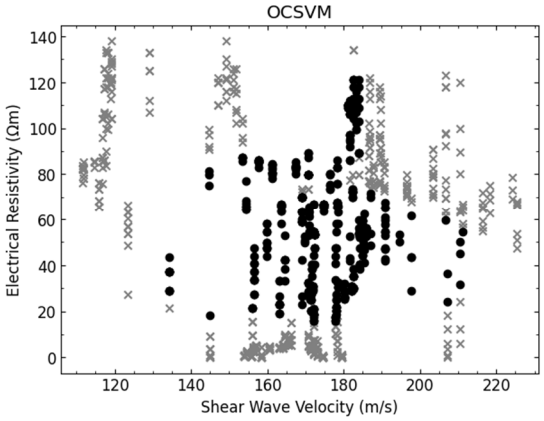
Profile: fc_site3

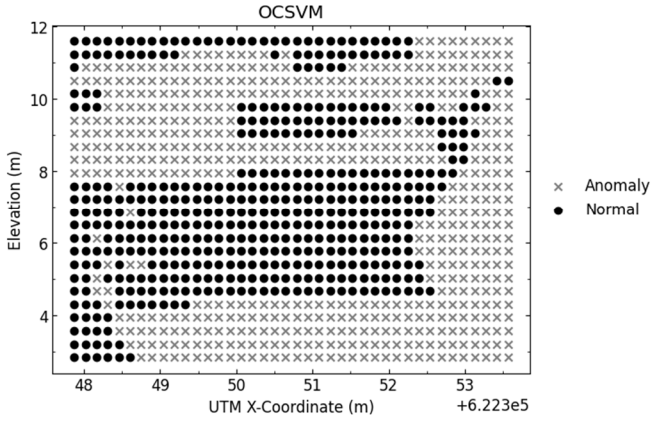




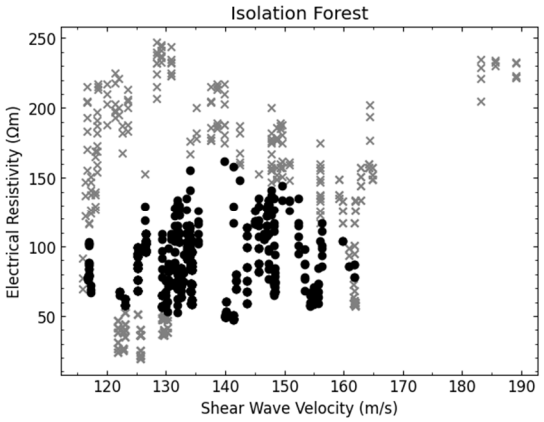


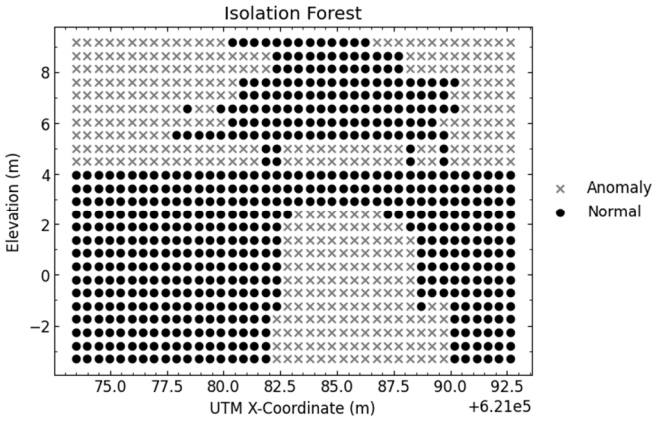


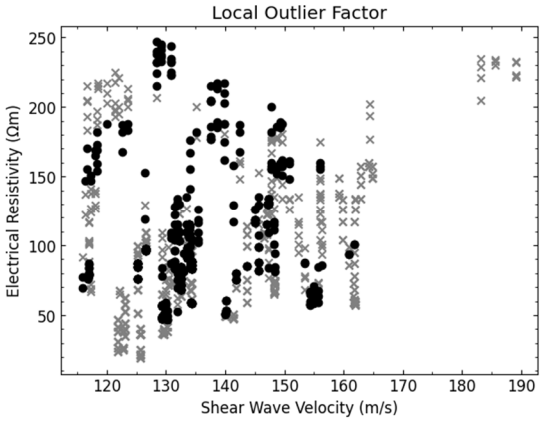


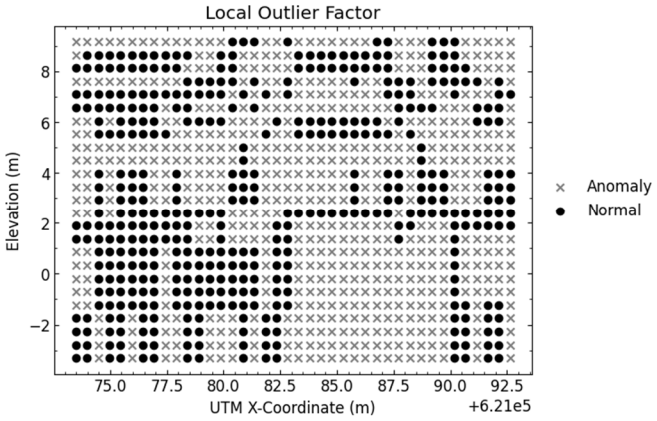


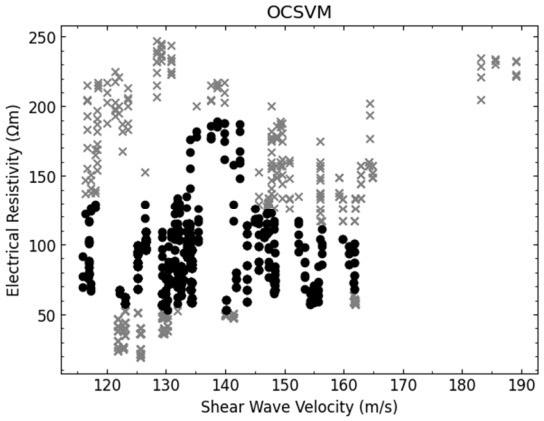
Profile: fc3_site1

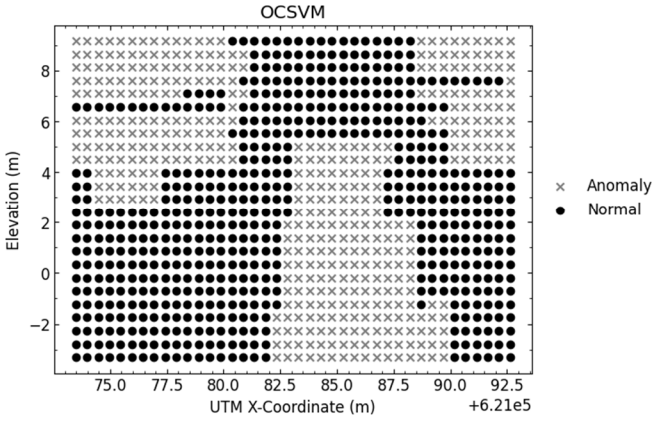




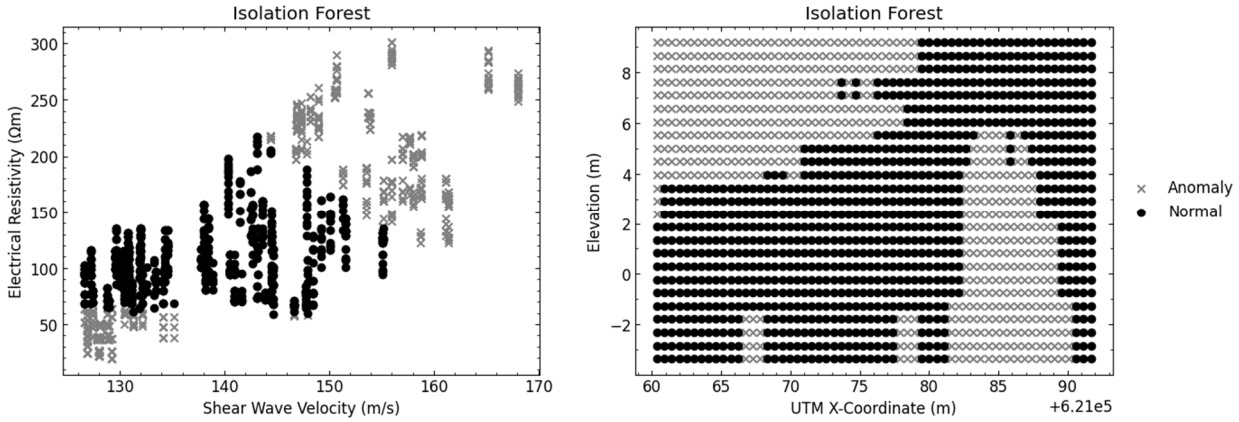


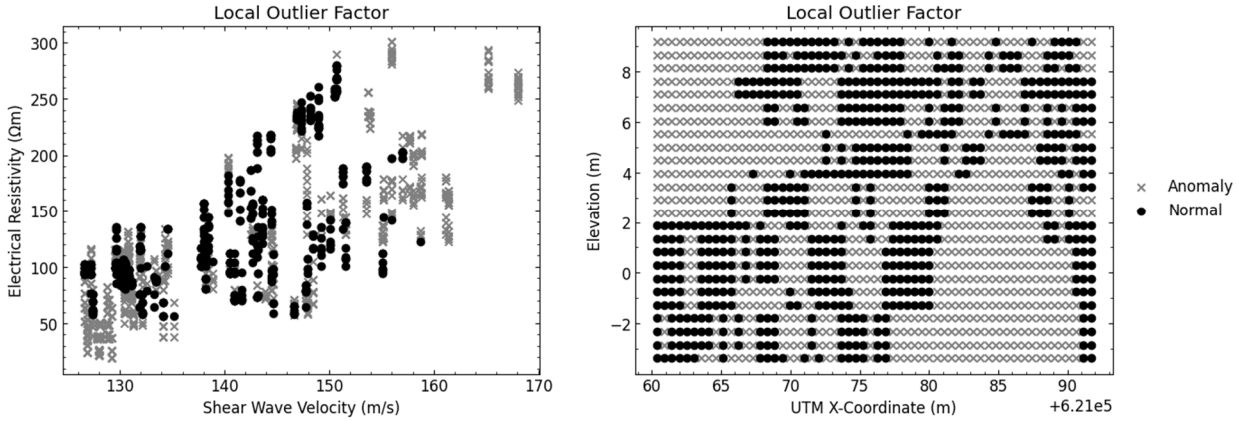


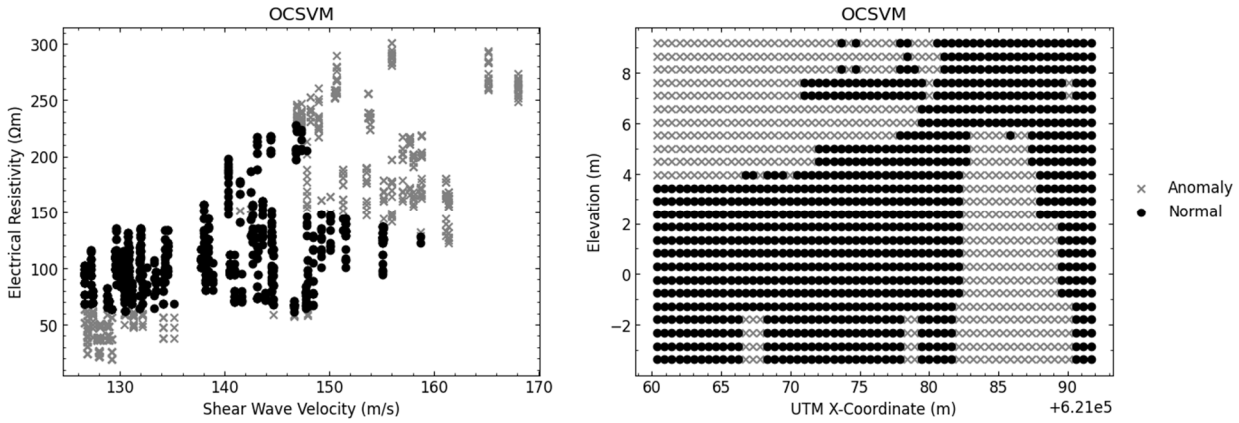




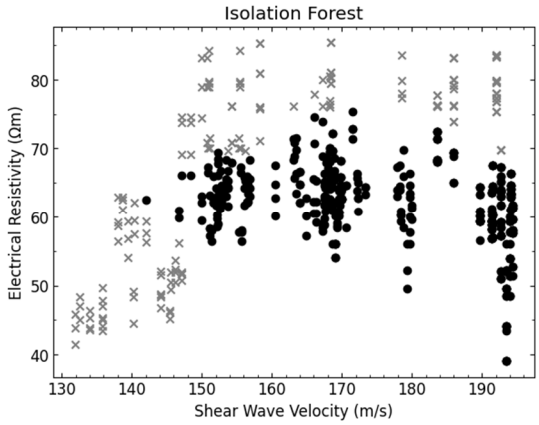
Profile: fc5_site2

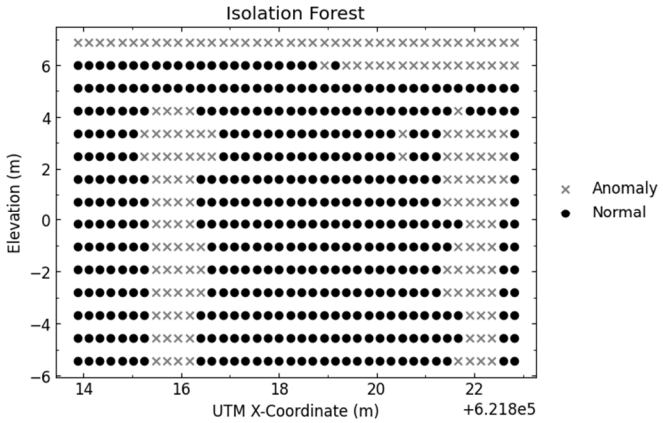


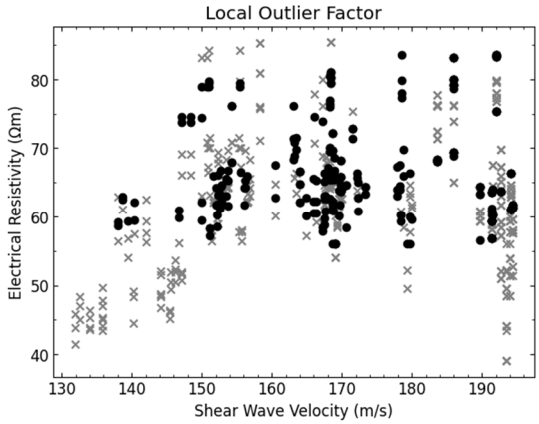


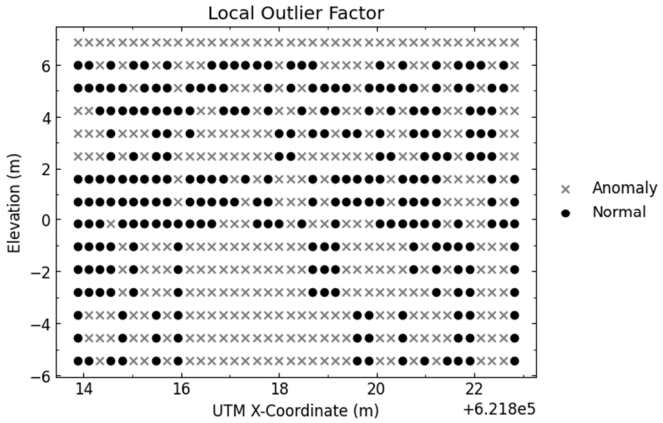


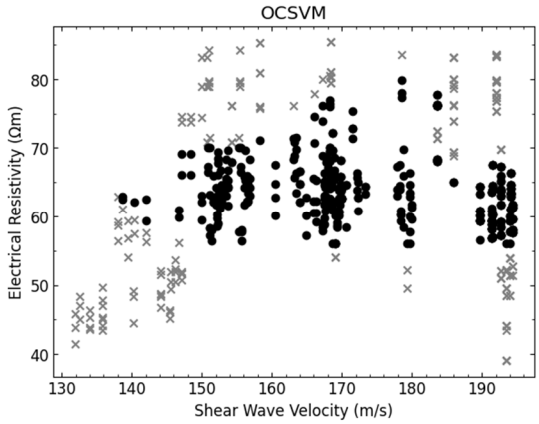
Profile: sc_site2

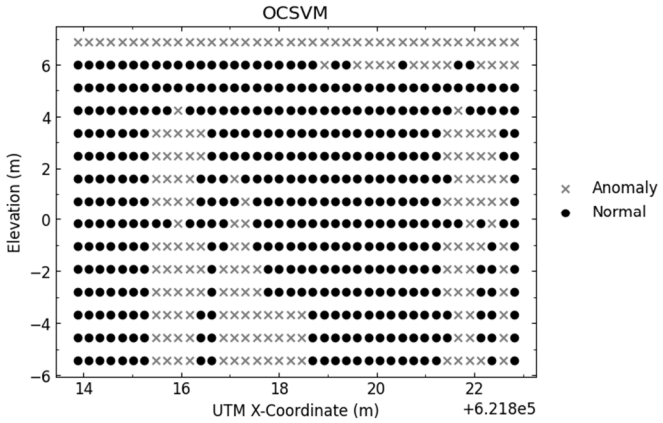




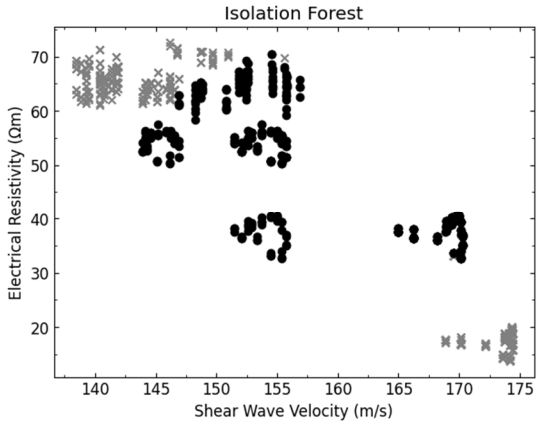


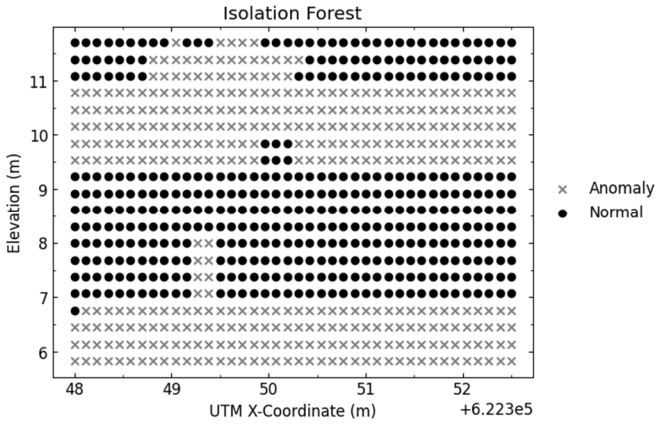


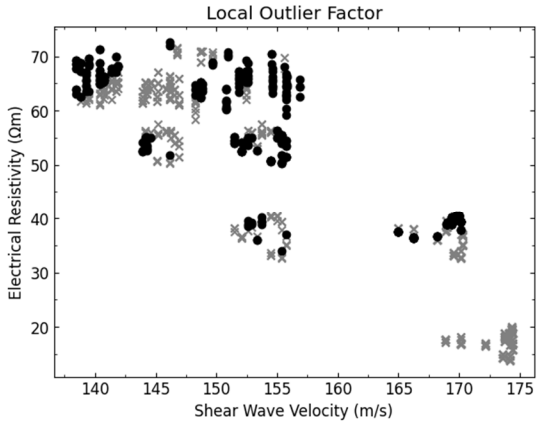


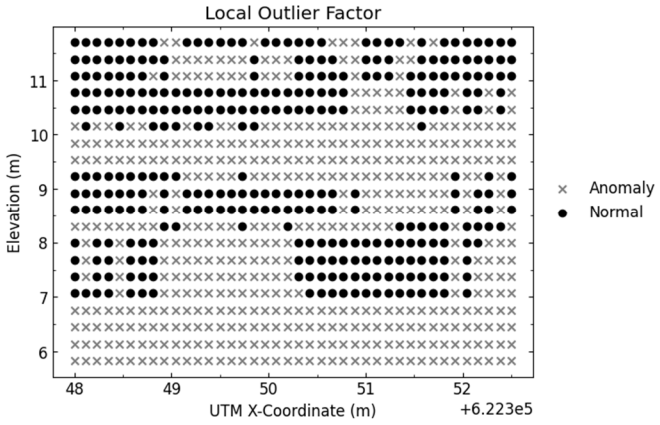


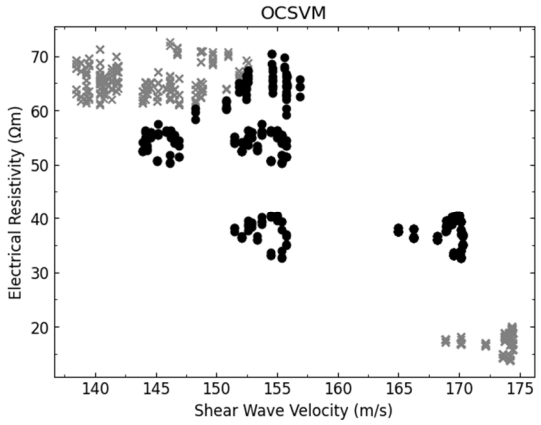
Profile: sc_site3

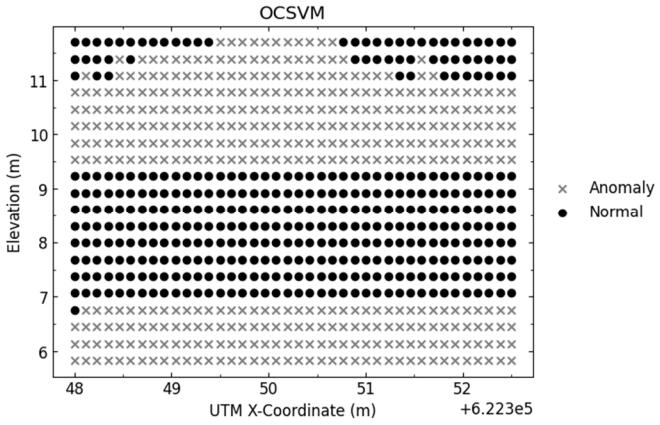




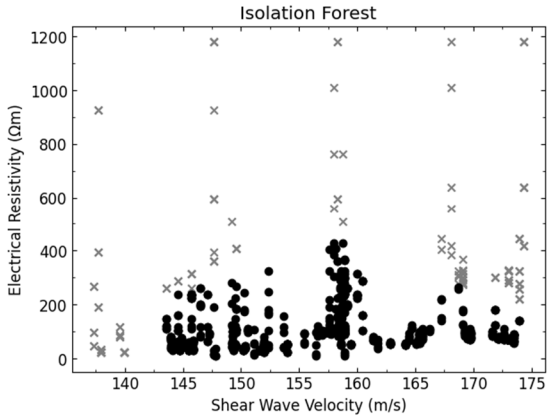


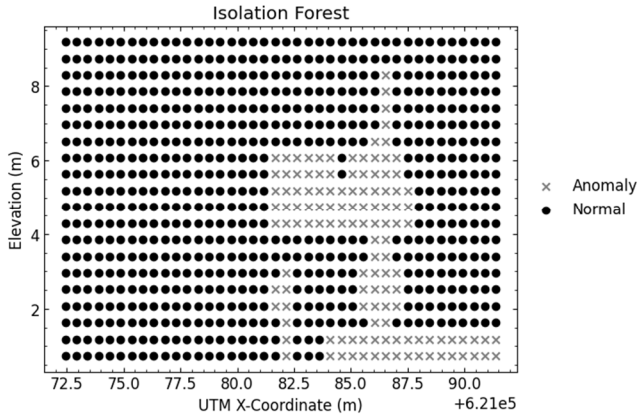


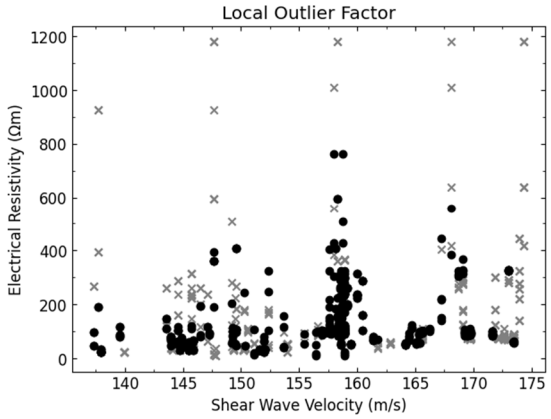


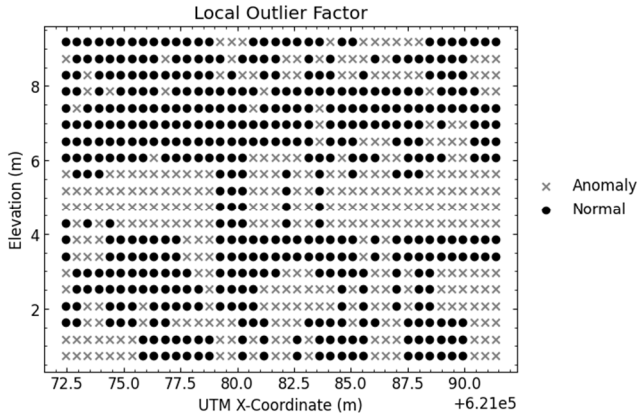


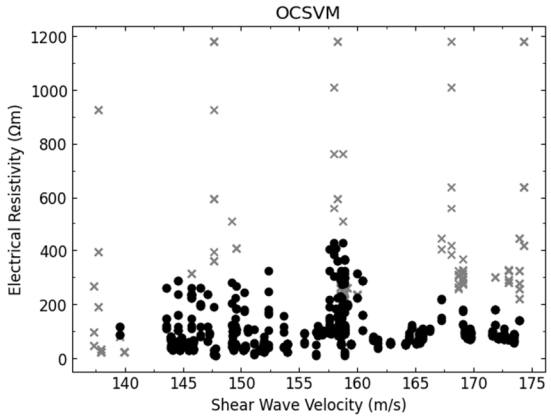
Profile: sc3_site1

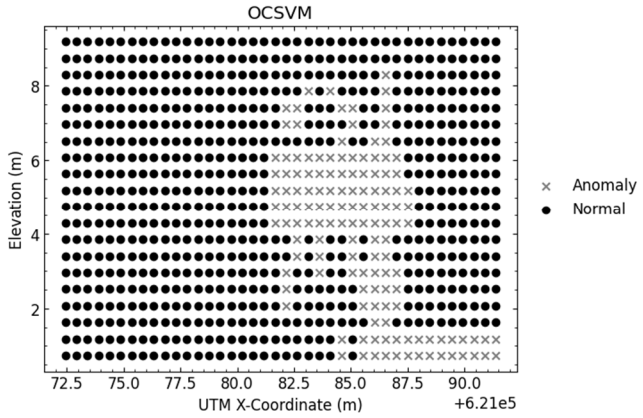




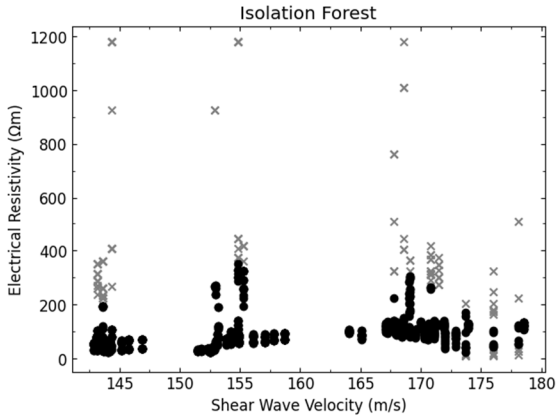


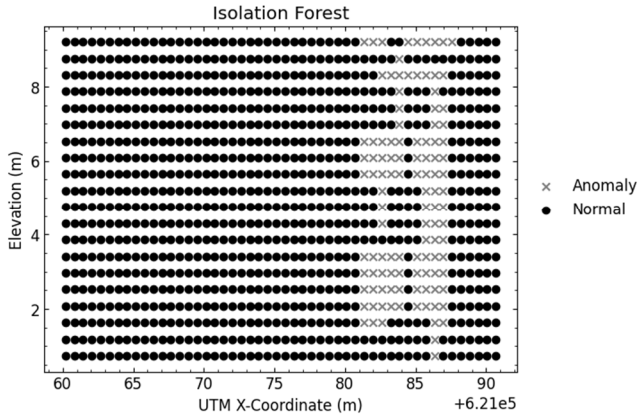


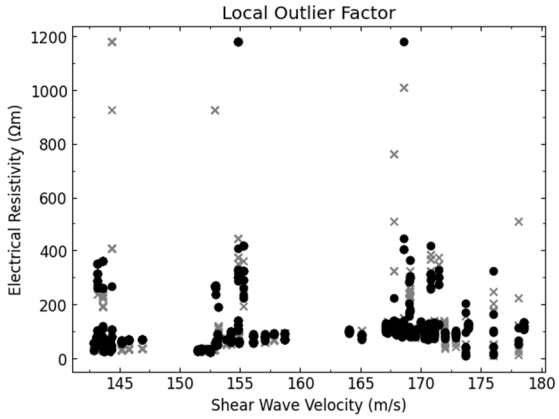


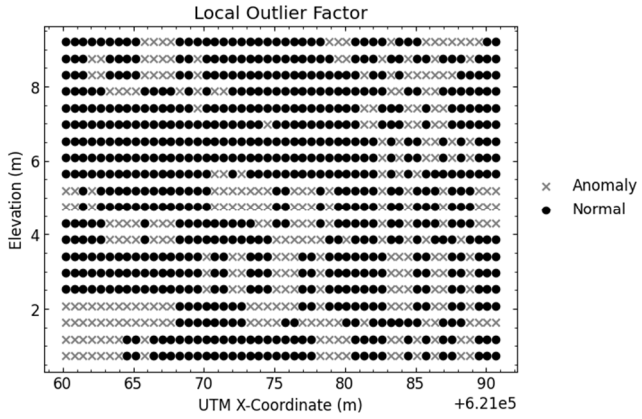


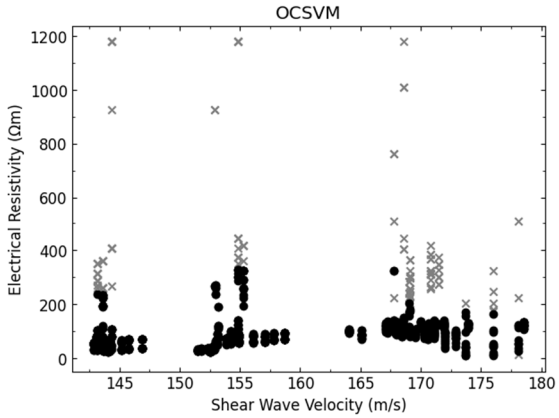
Profile: sc5_site1

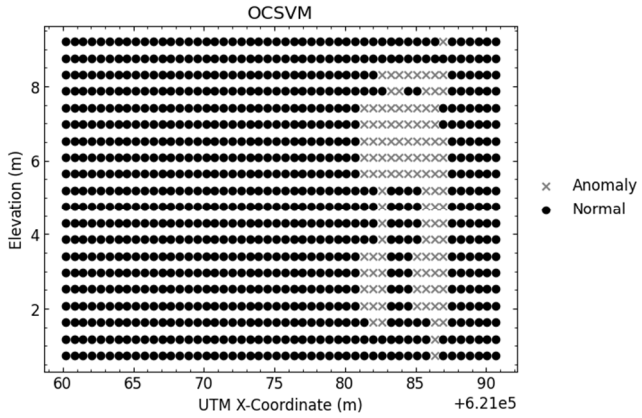






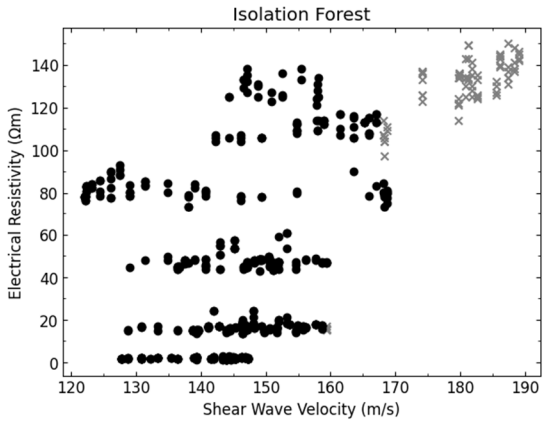


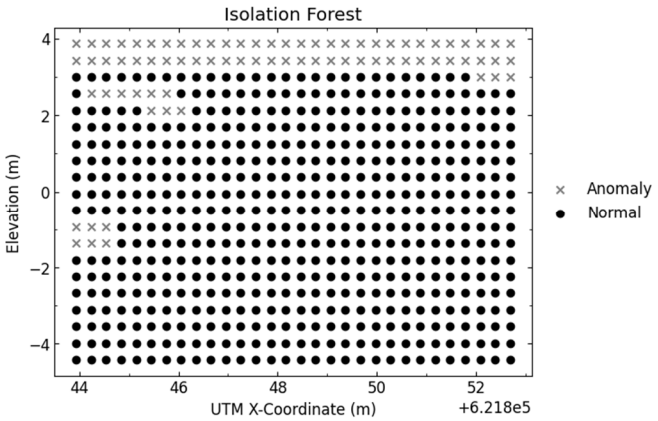


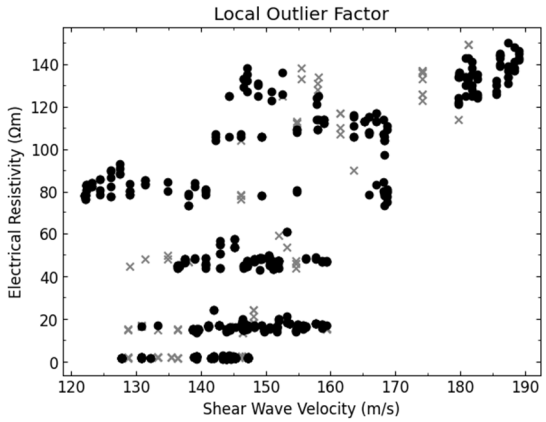


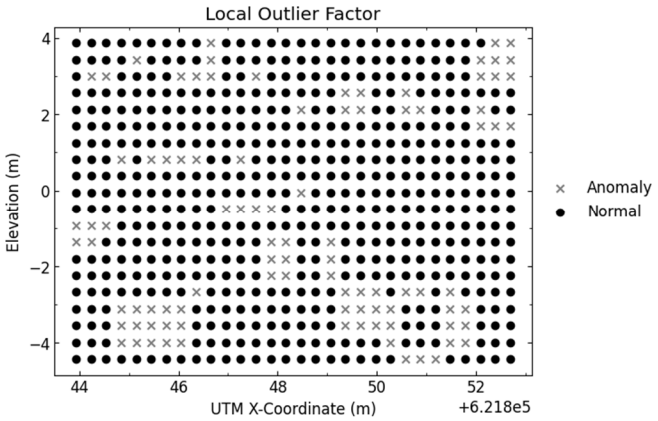
Fit: Toe Compilation Dataframe
Predicted: Toe Compilation Dataframe

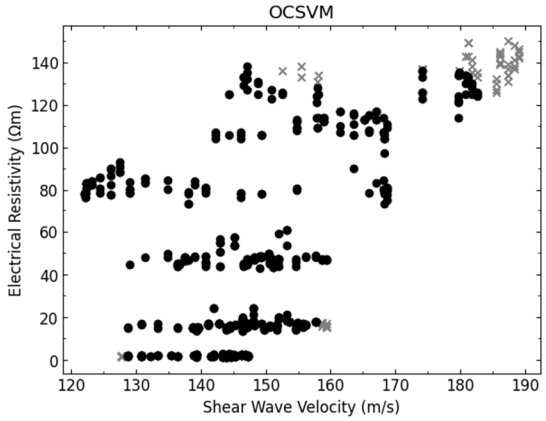
Profile: fdt_site2

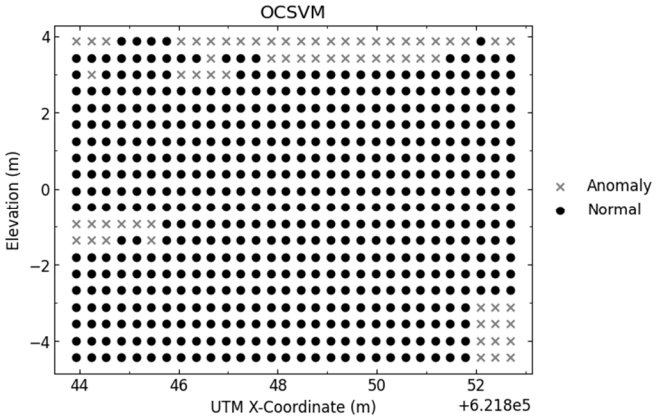




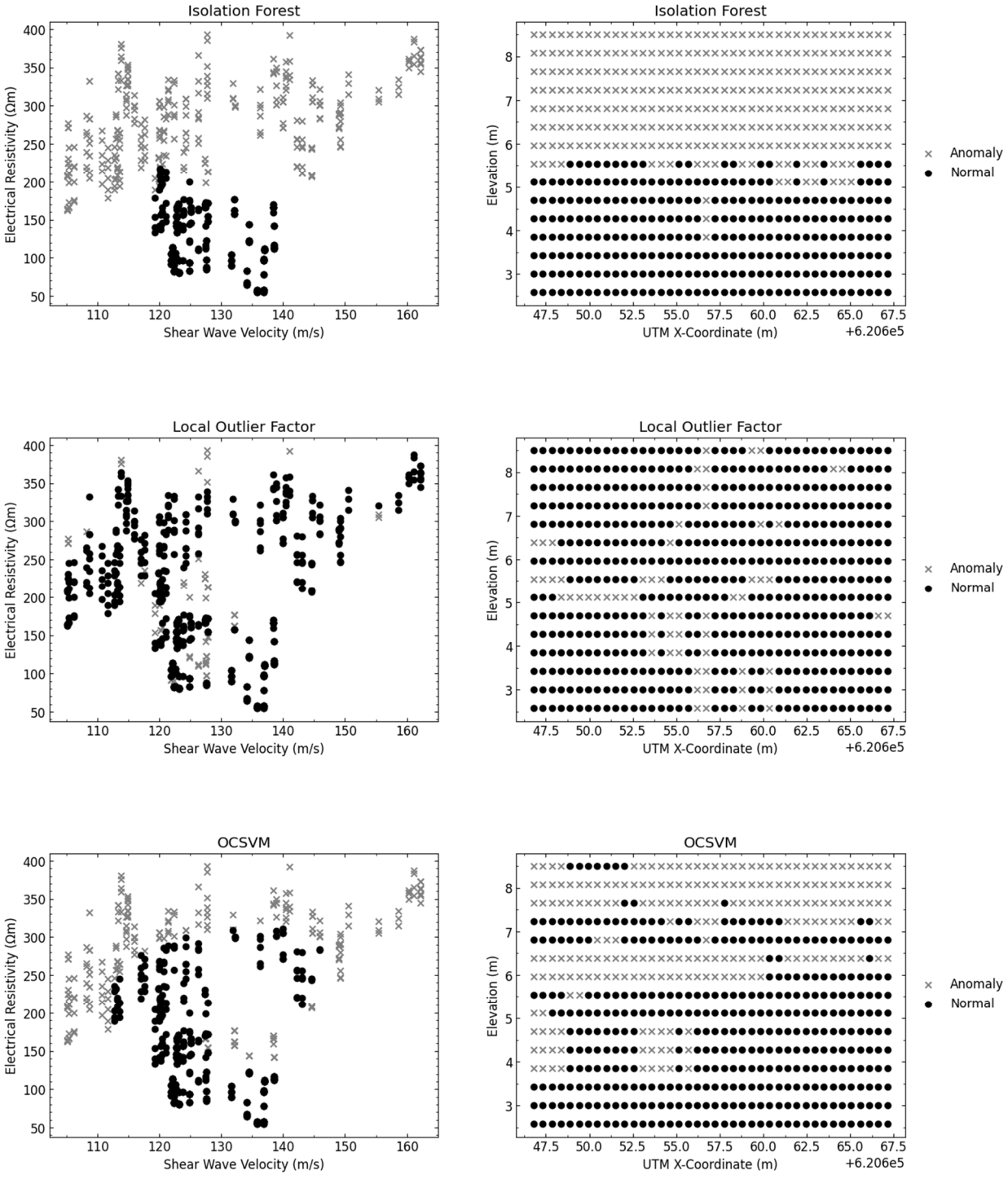




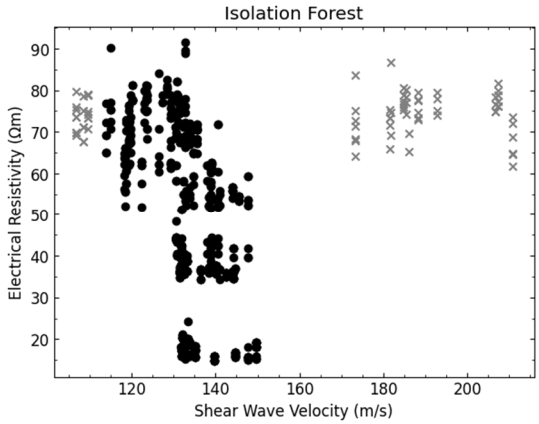


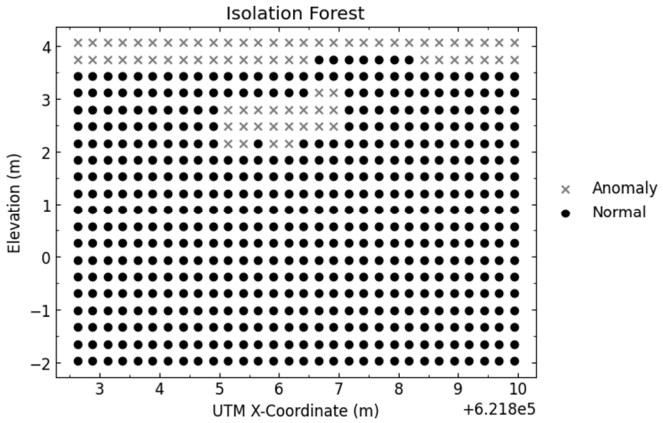


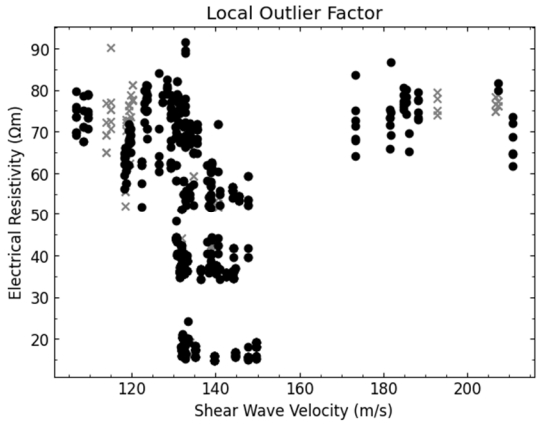
Profile: fwt_site1

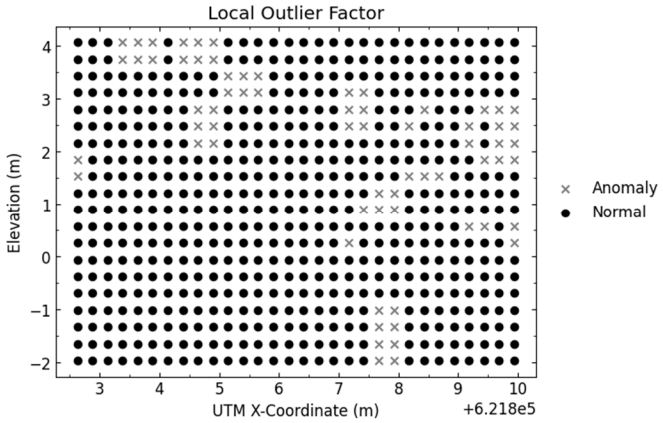


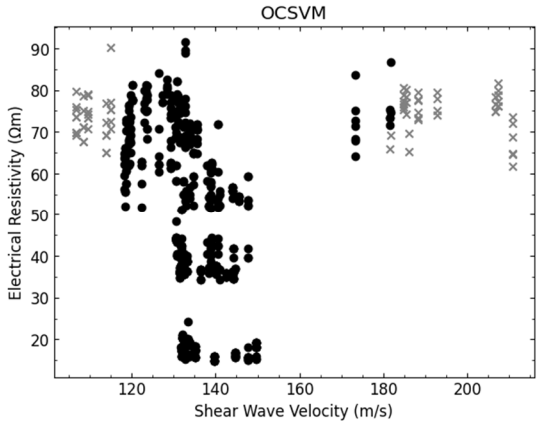
Profile: fwt_site2

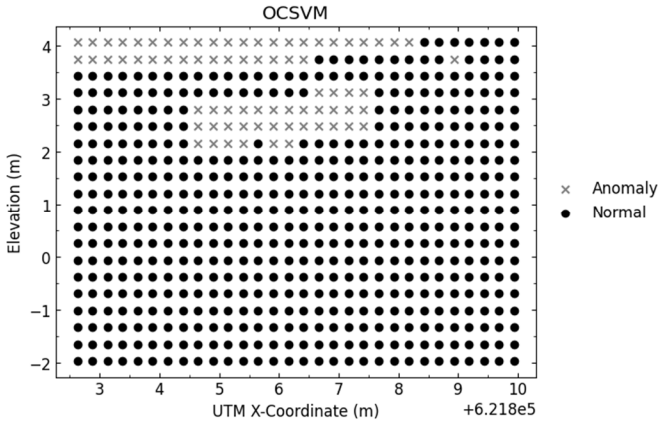




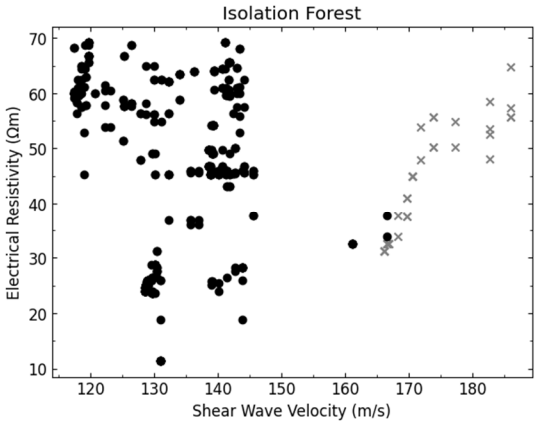


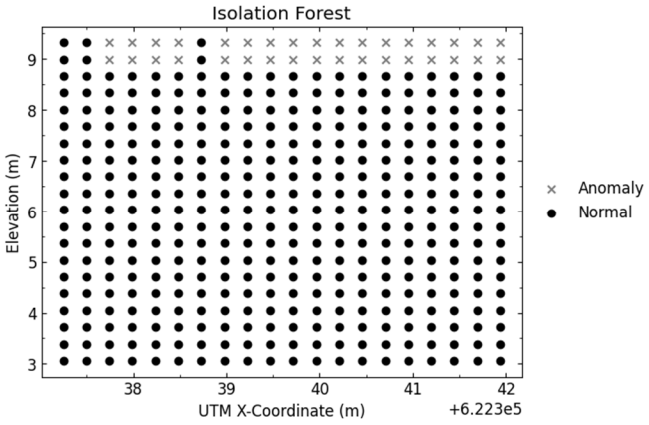


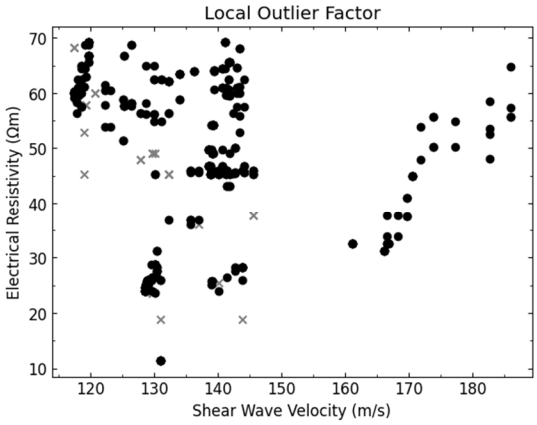


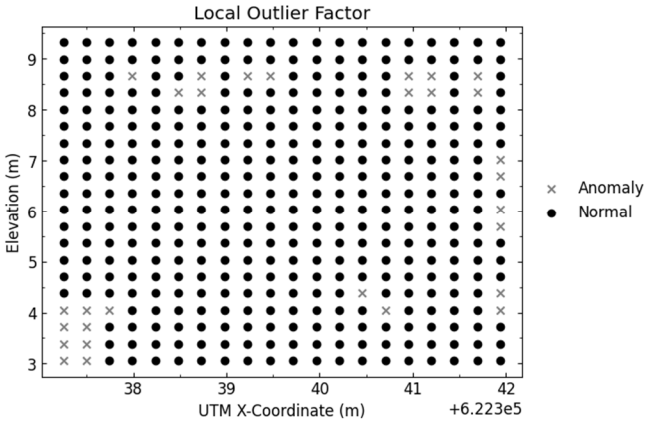


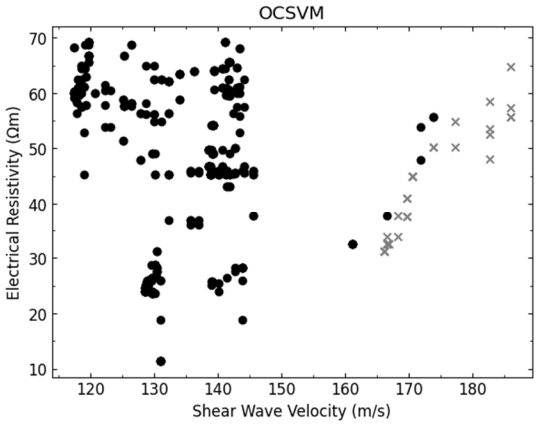
Profile: fwt_site3

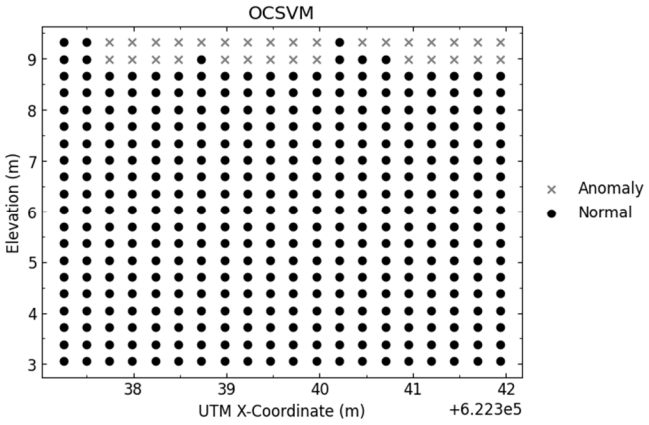




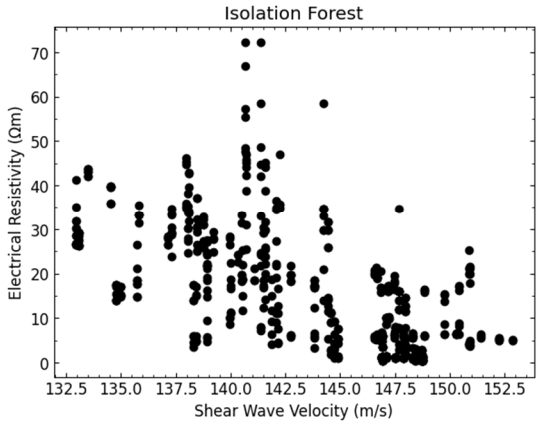


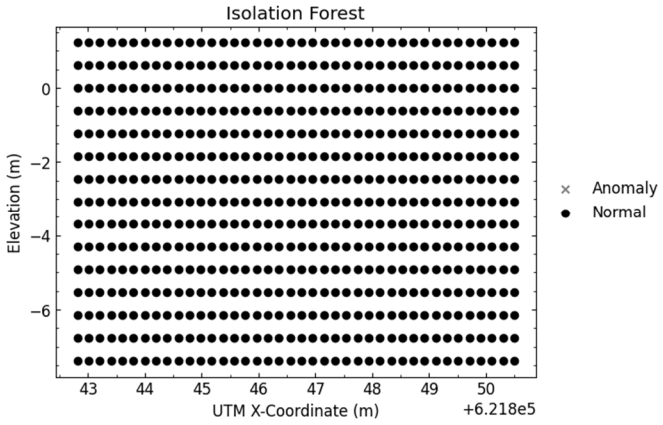


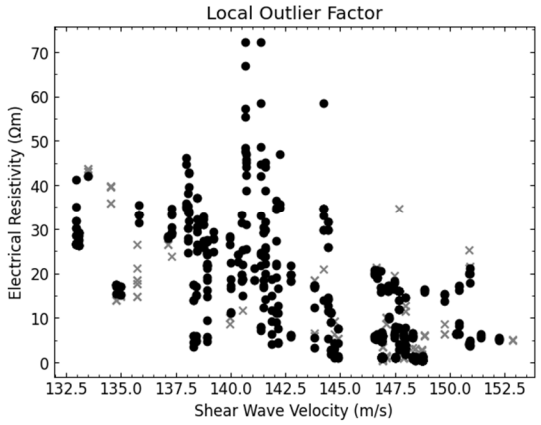


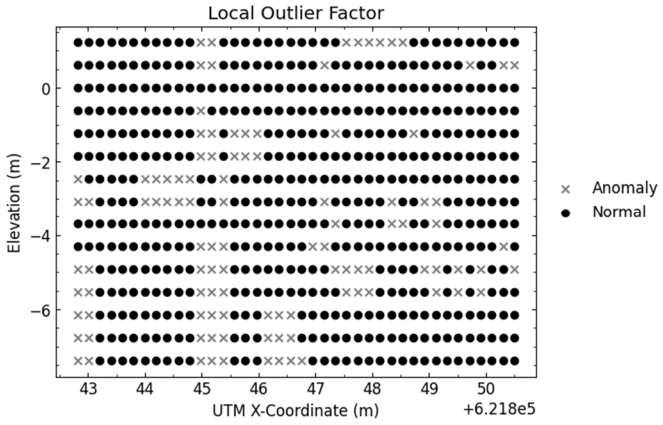


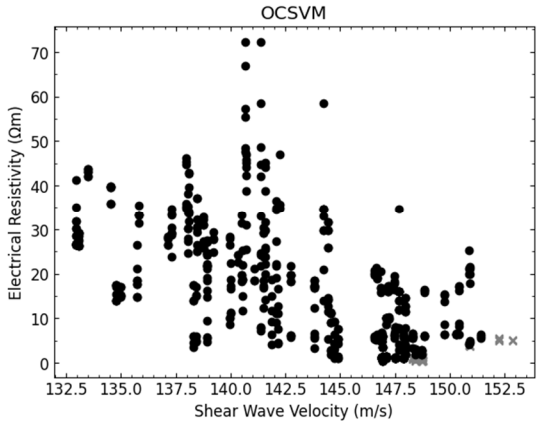
Profile: sdt_site2

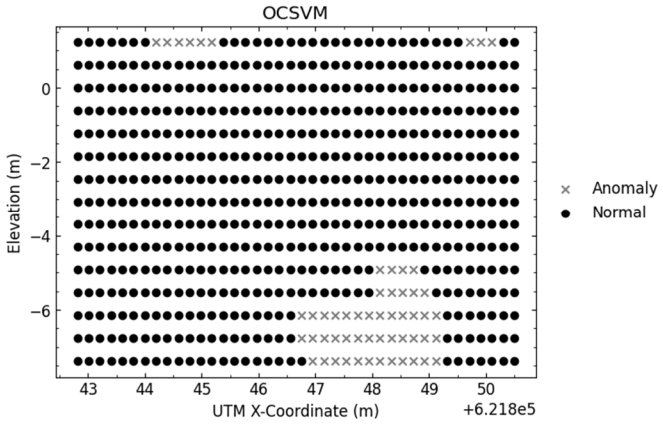




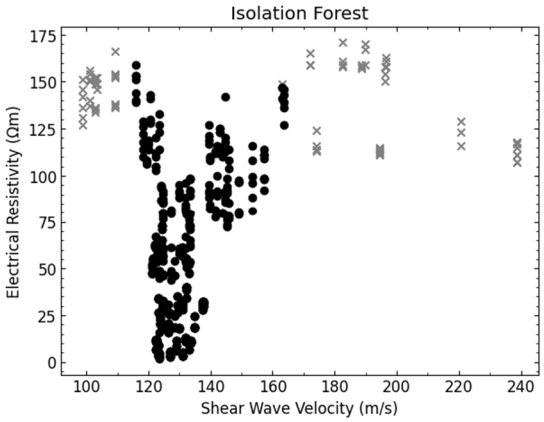


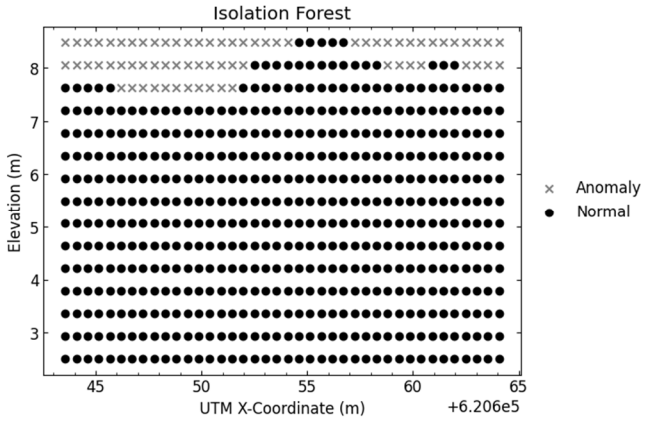


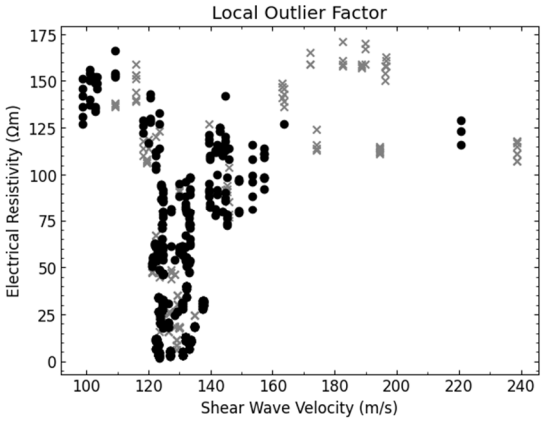


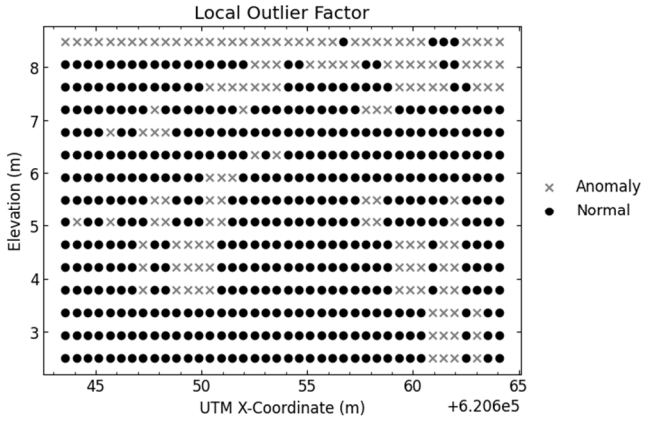


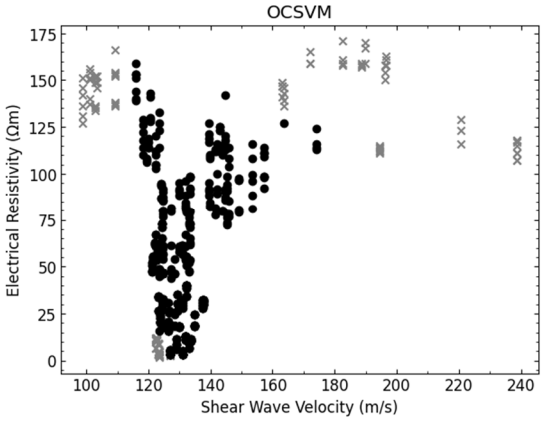
Profile: swt_site1

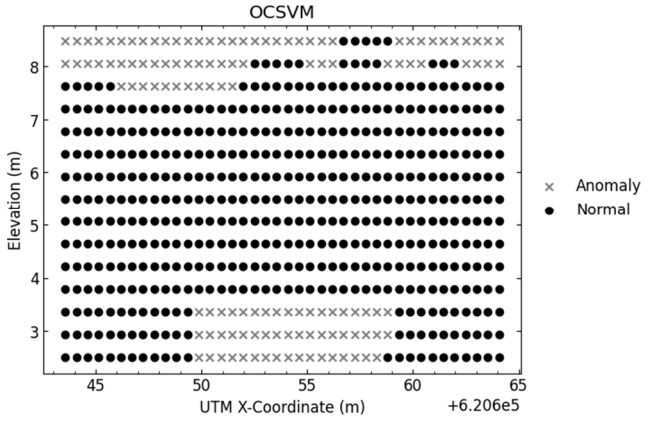




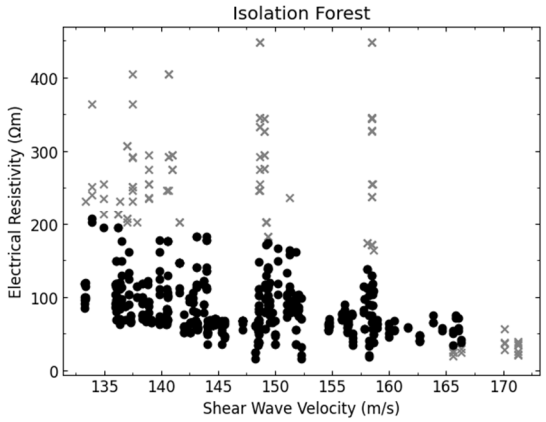


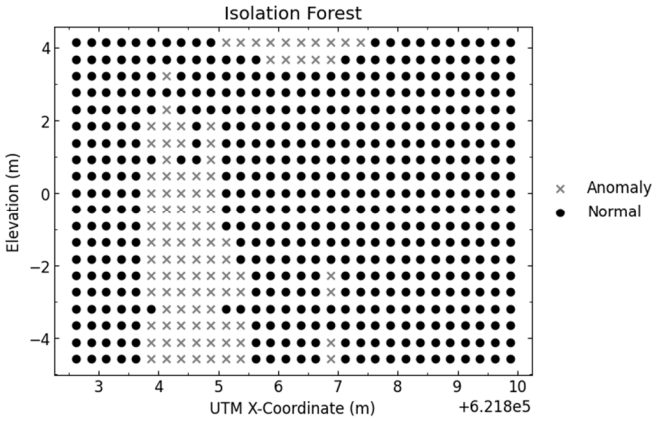


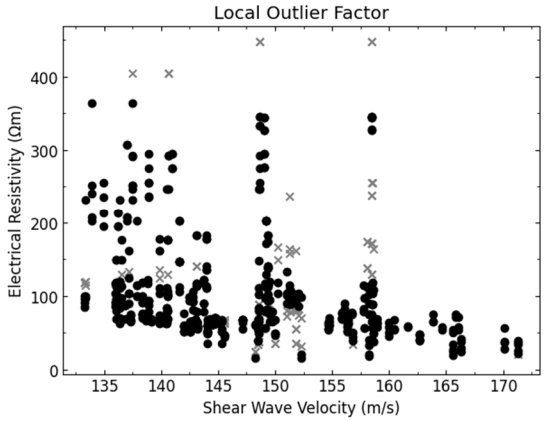


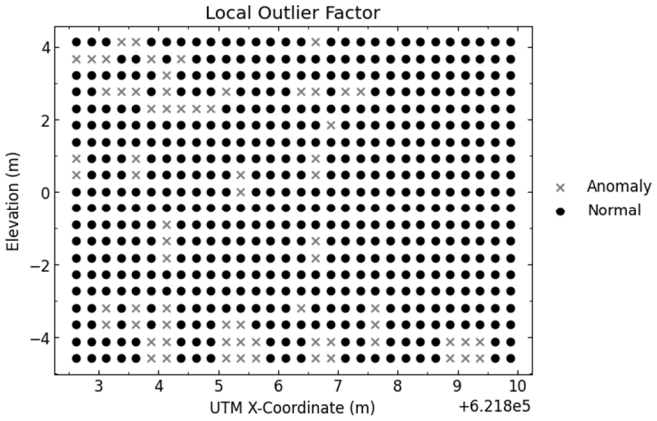


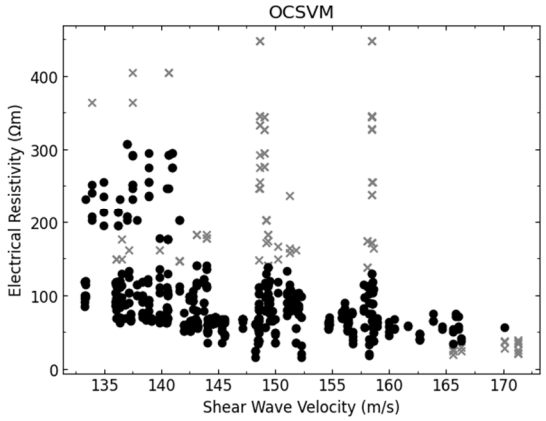
Profile: swt_site2

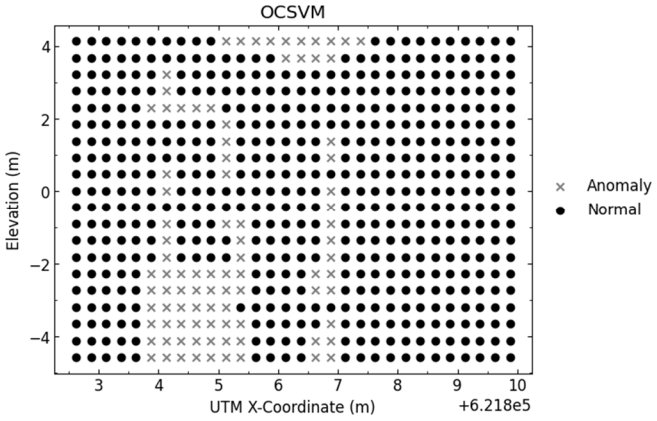




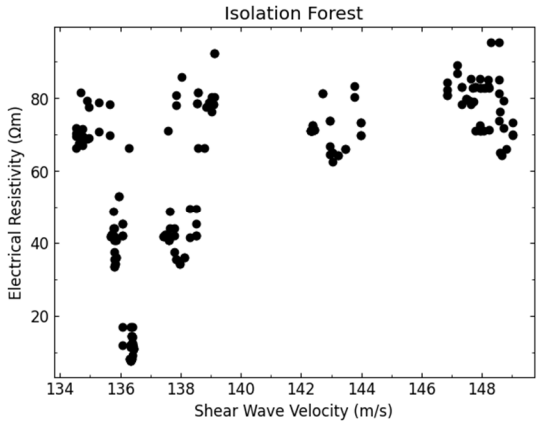


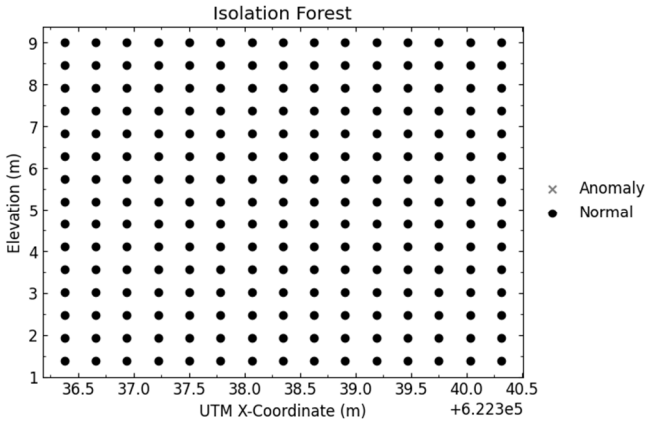


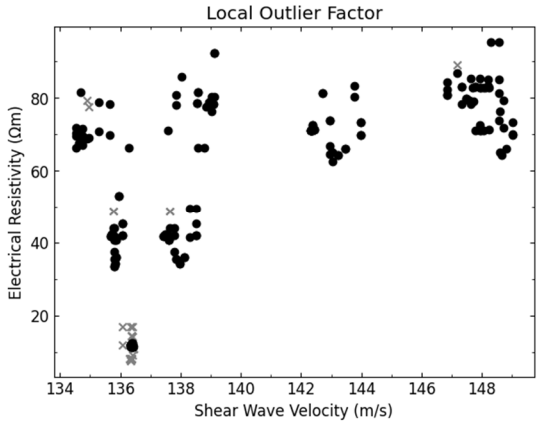


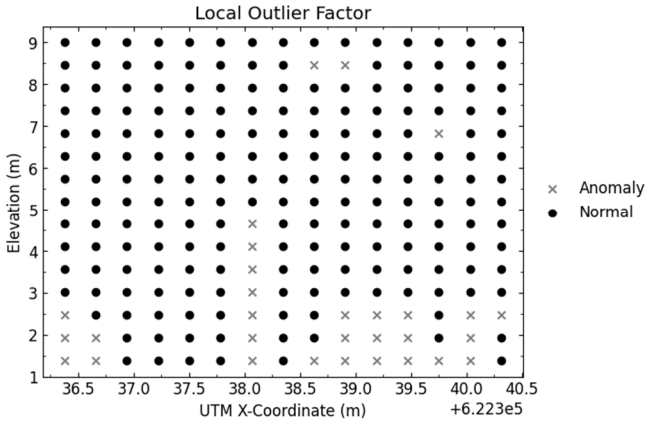


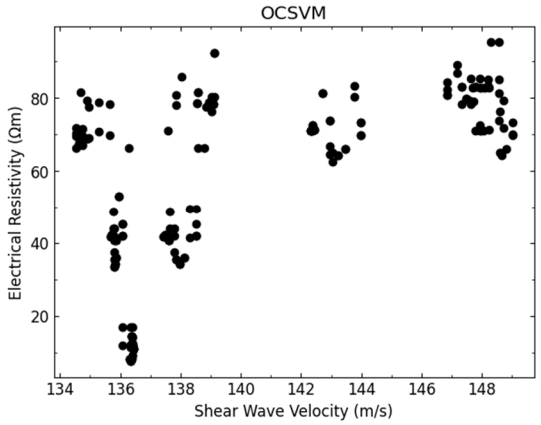
Profile: swt_site3

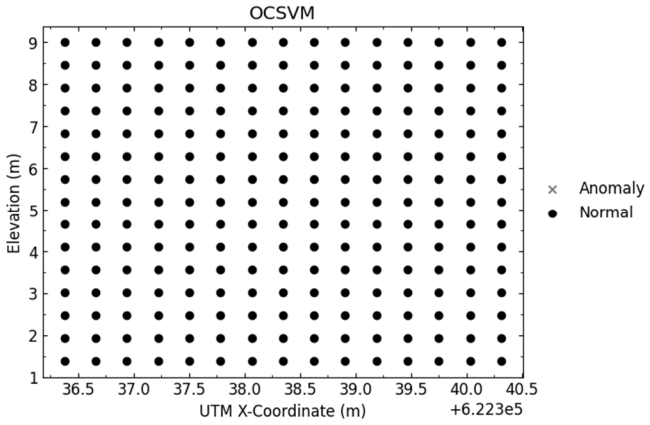






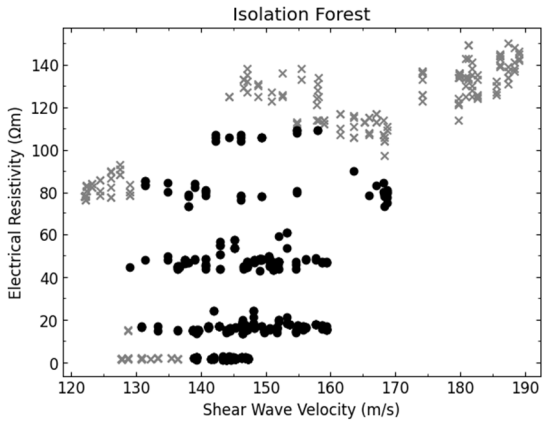


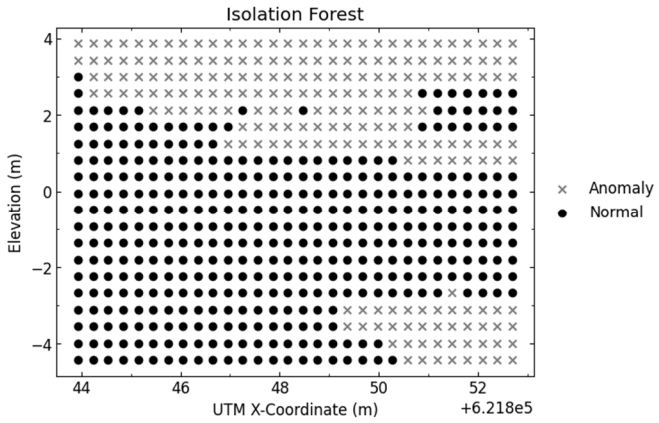


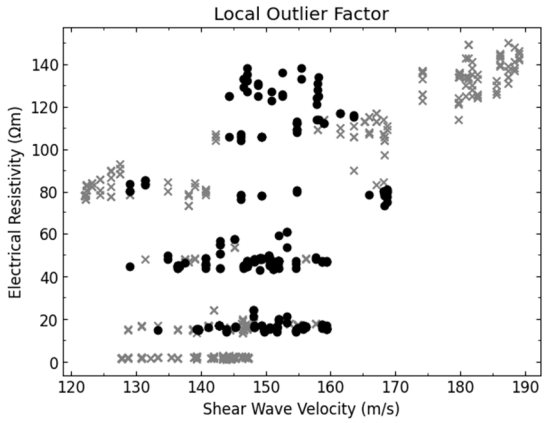


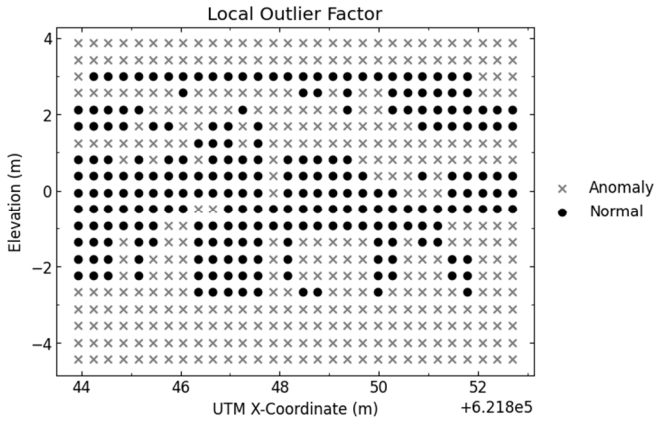
Fit: Toe Compilation Dataframe
Predicted: Individual Profiles

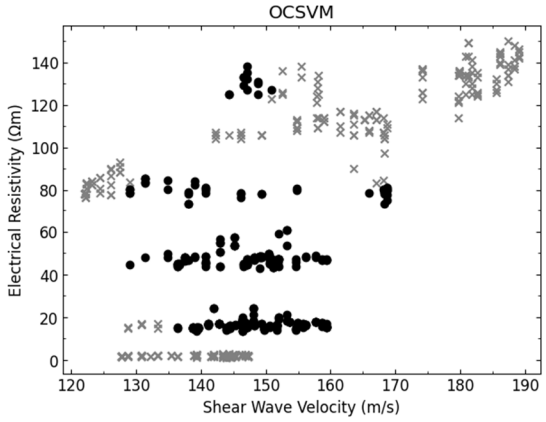
Profile: fdt_site2

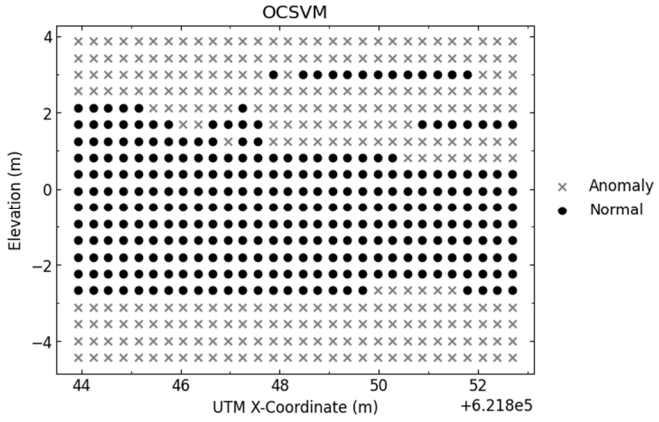




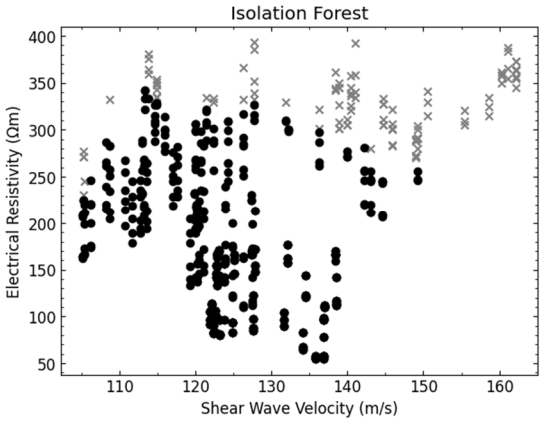


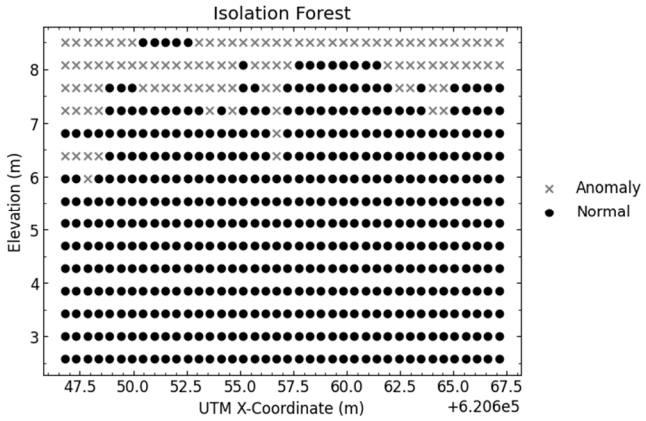


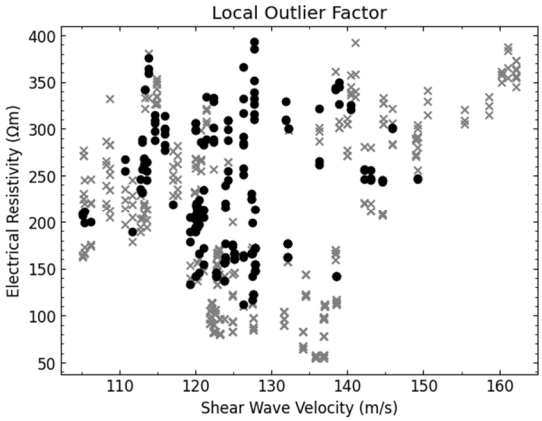


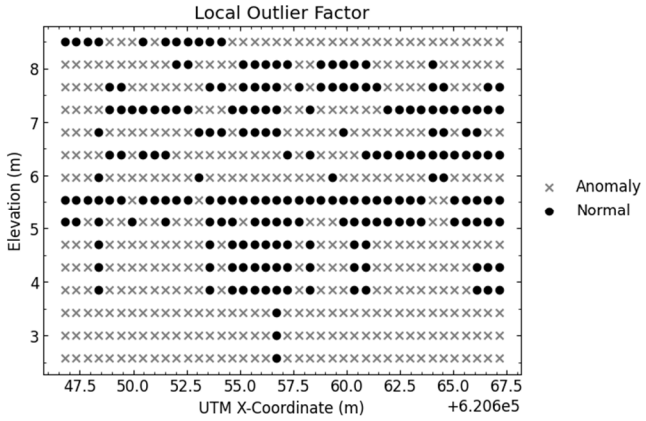


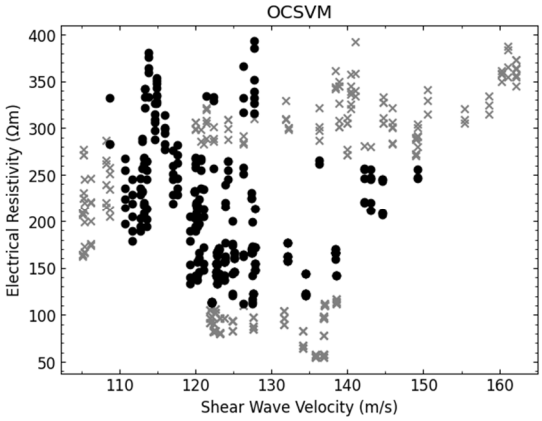
Profile: fwt_site1

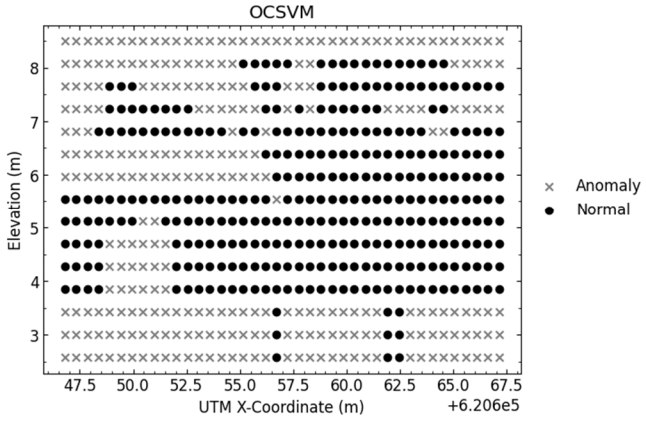




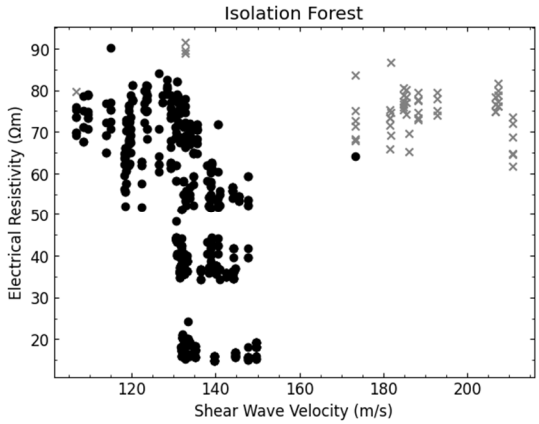


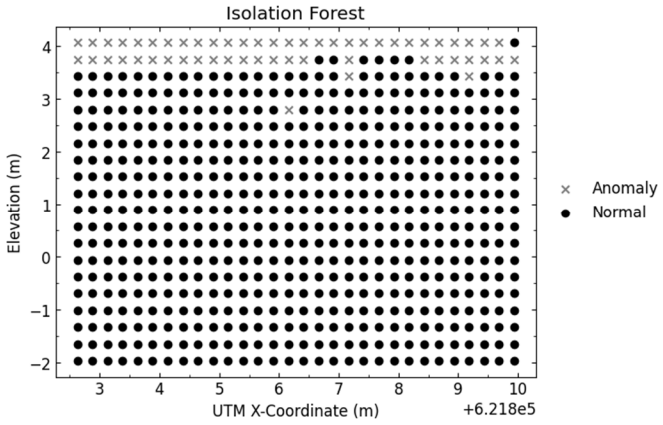


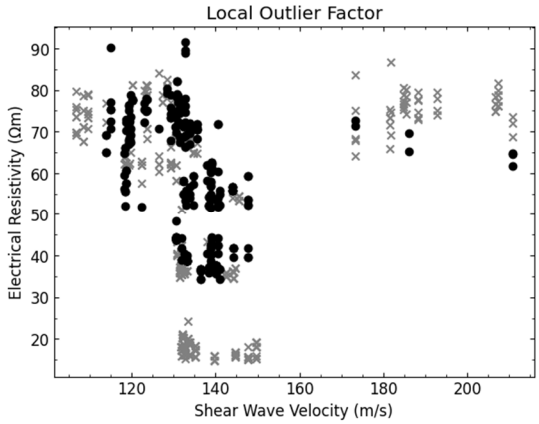


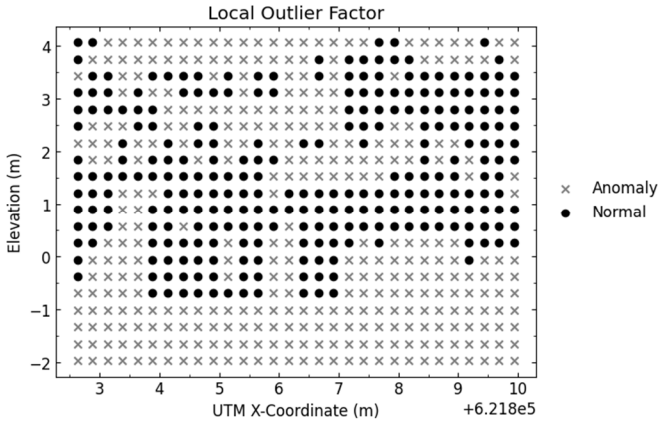


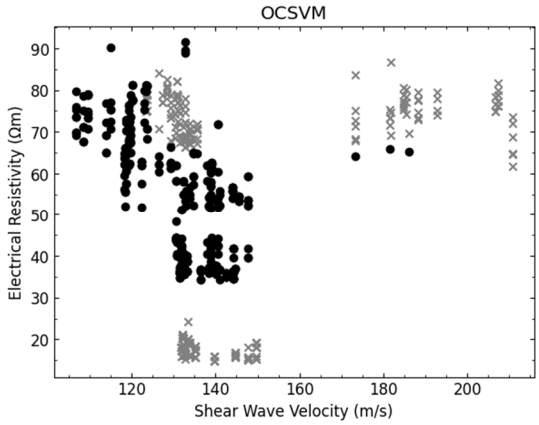
Profile: fwt_site2

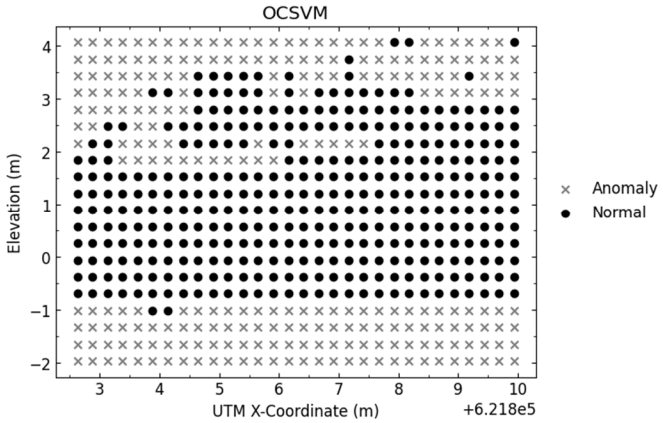




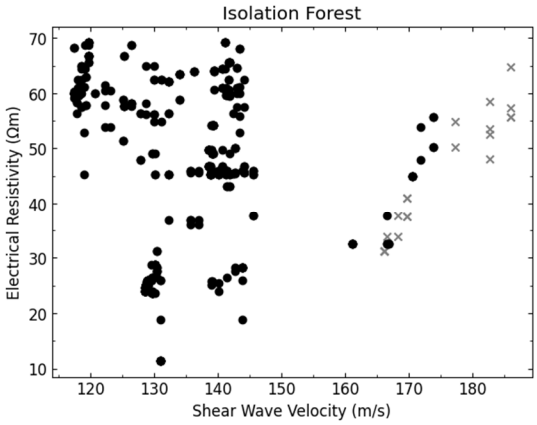


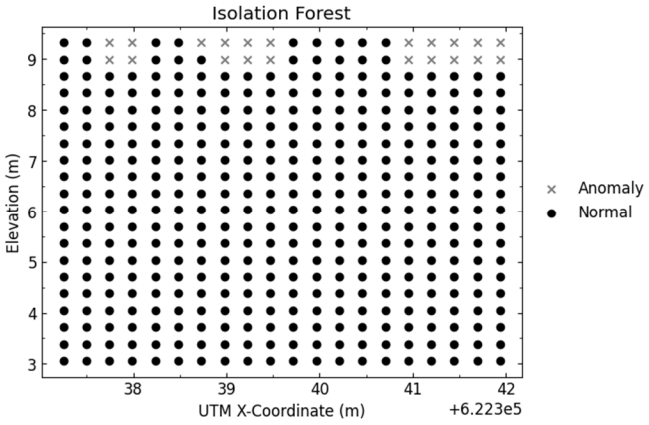


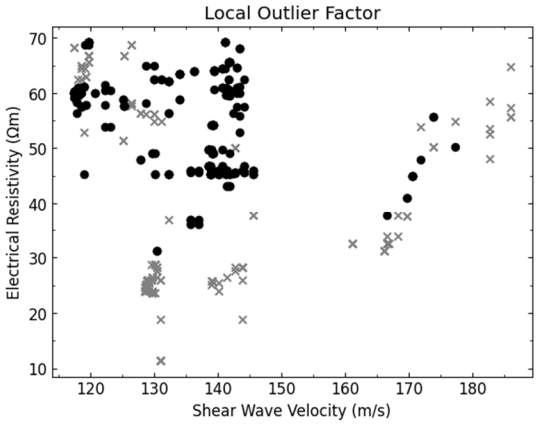


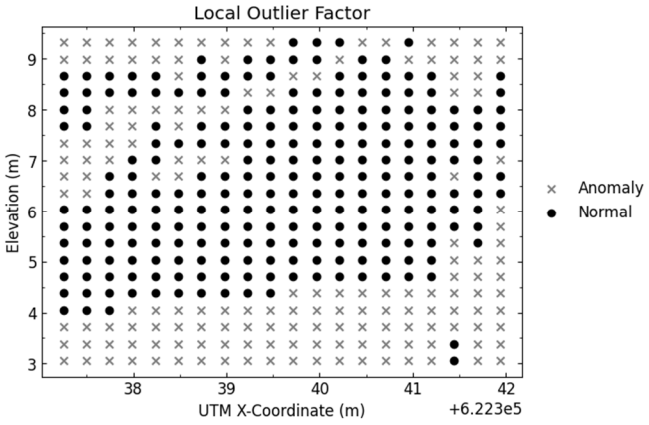


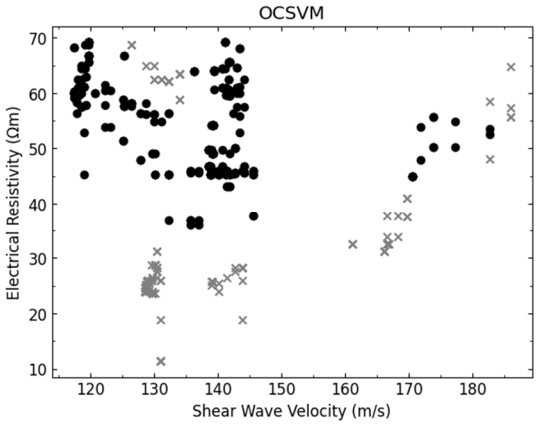
Profile: fwt_site3

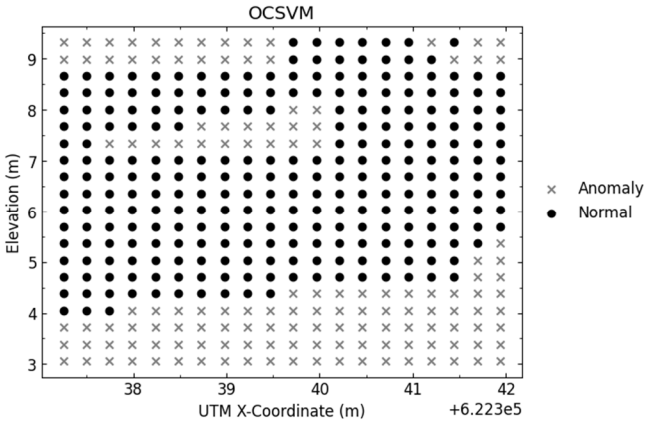




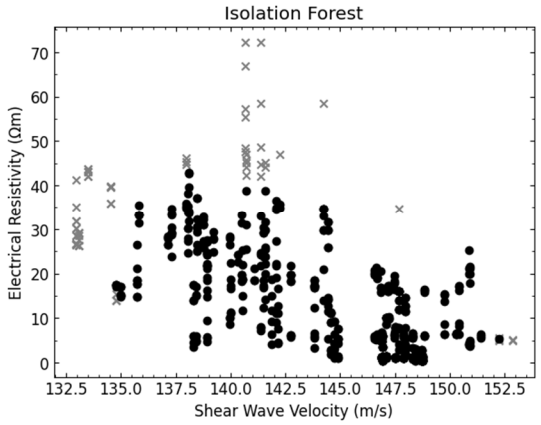


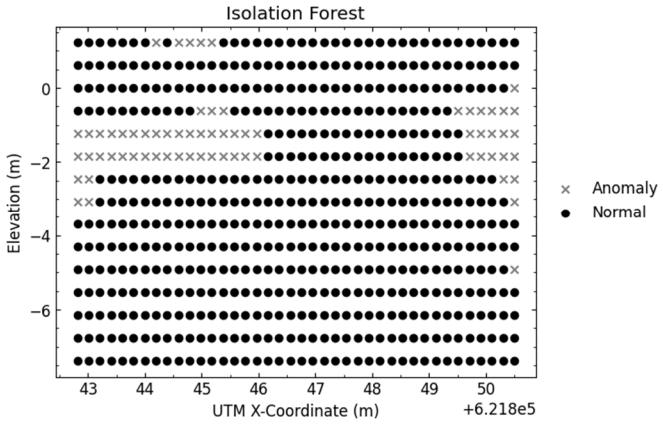


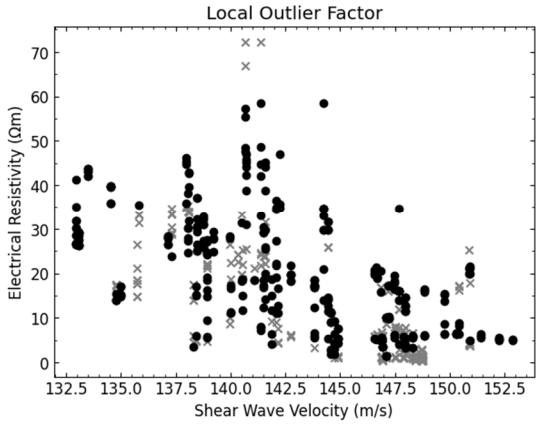


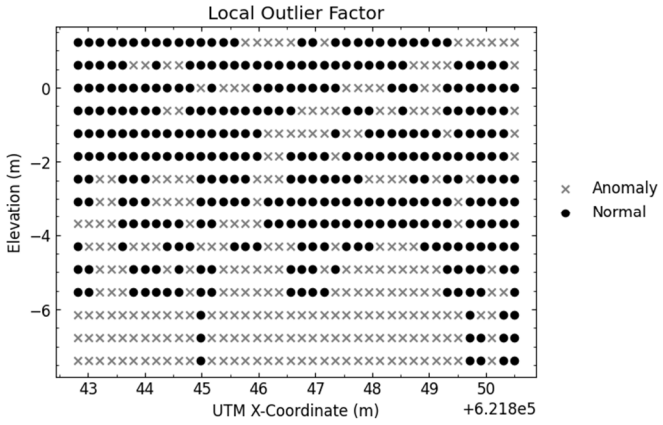


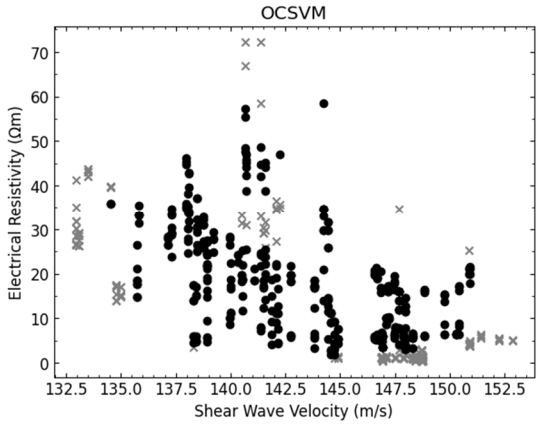
Profile: sdt_site2

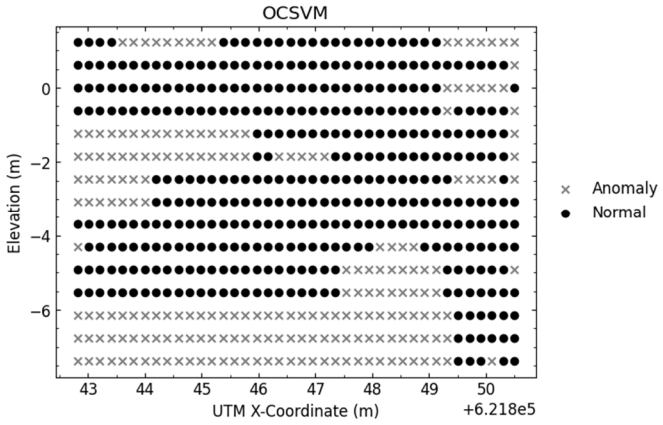




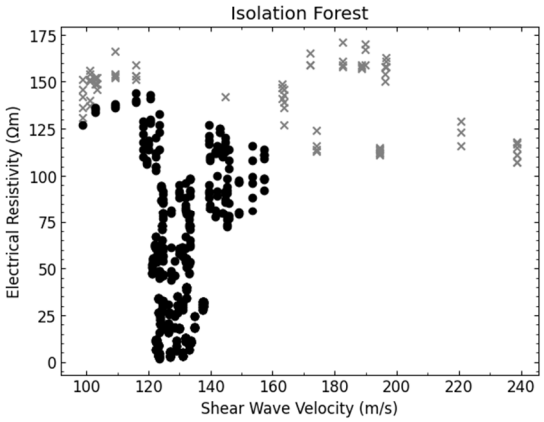


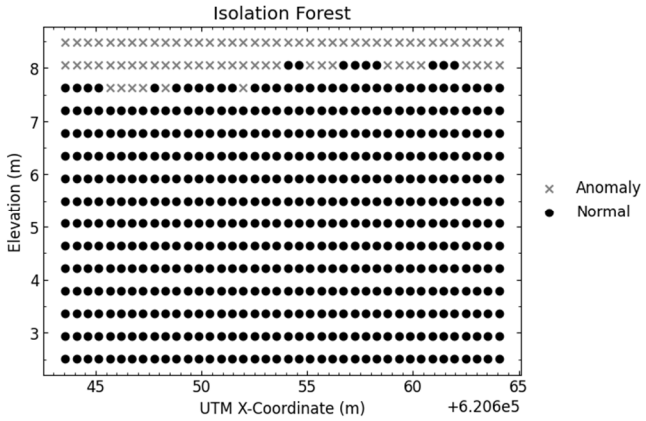


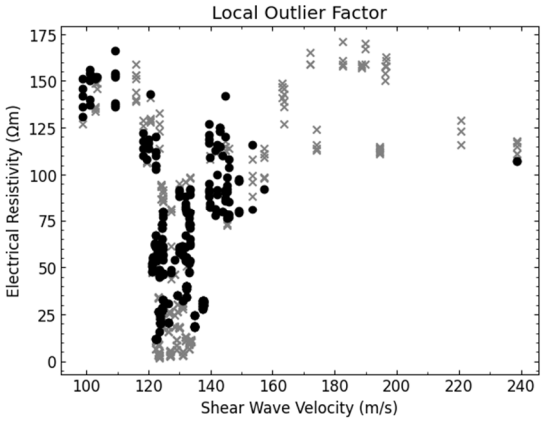


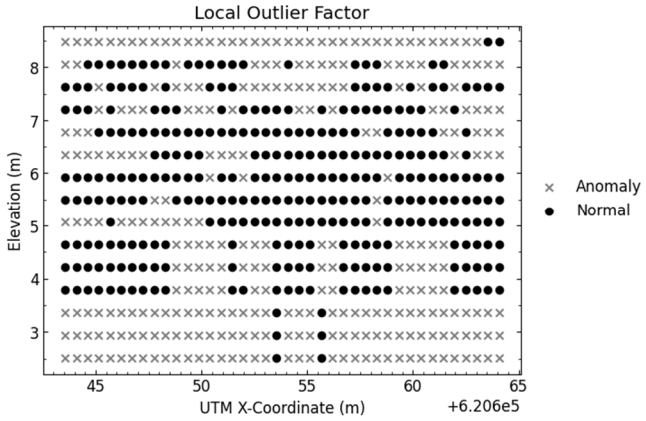


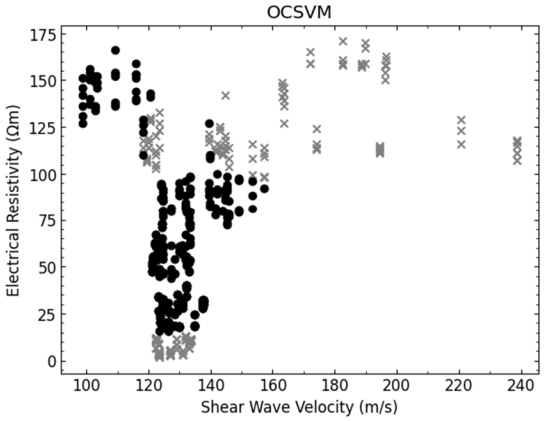
Profile: swt_site1

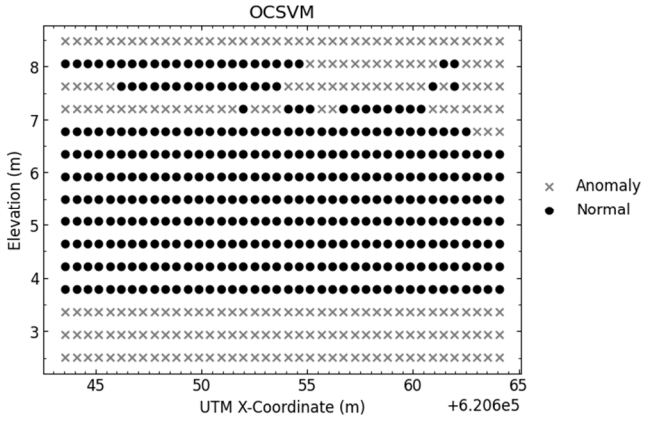




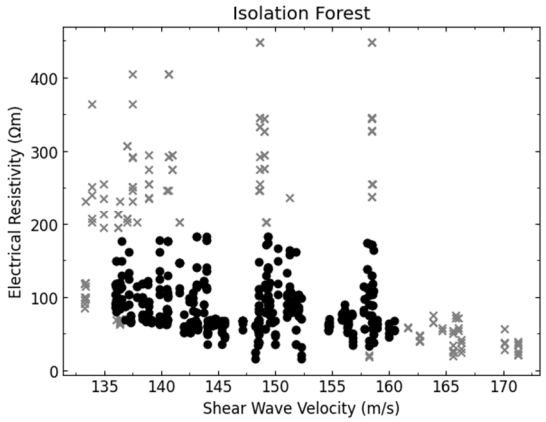


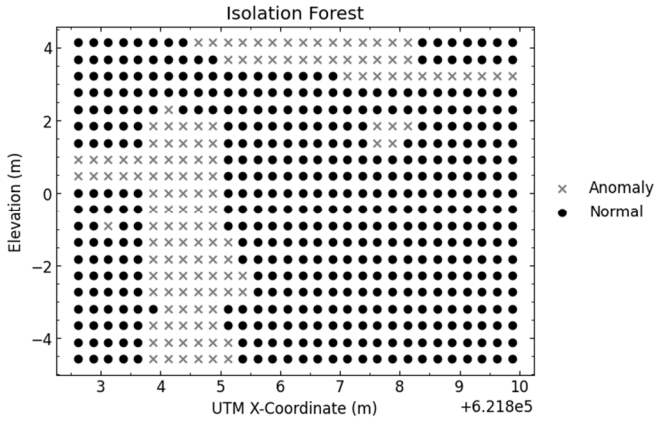


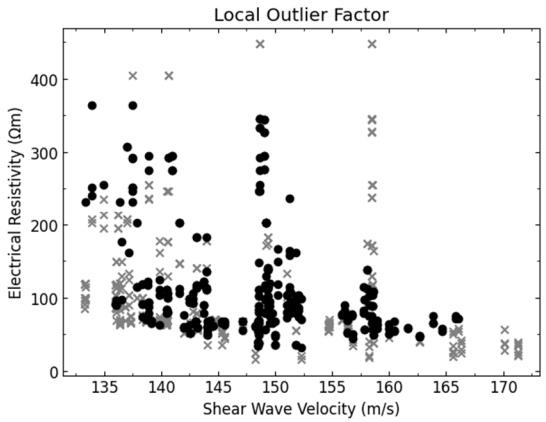


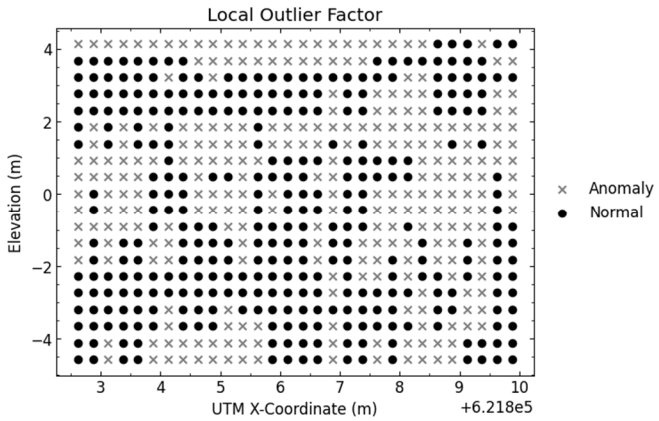


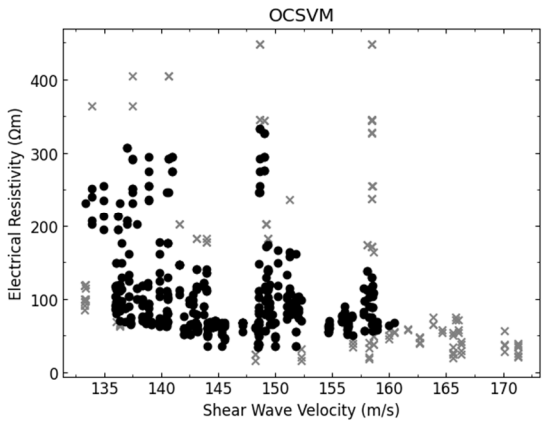
Profile: swt_site2

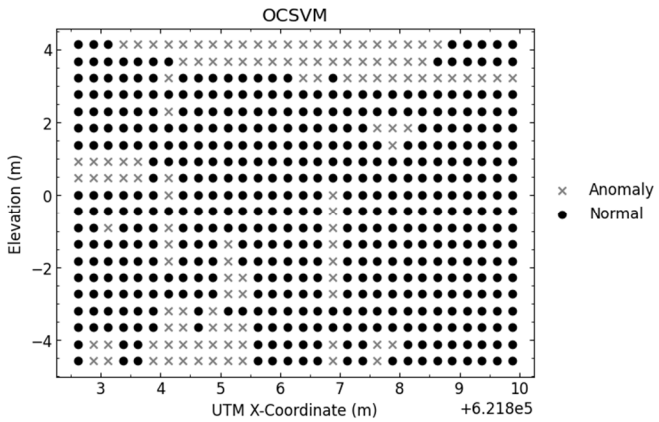




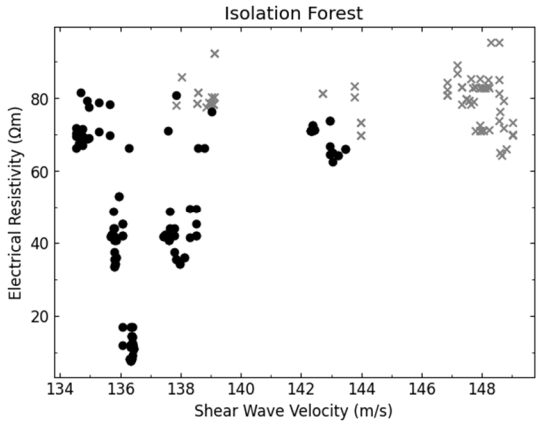


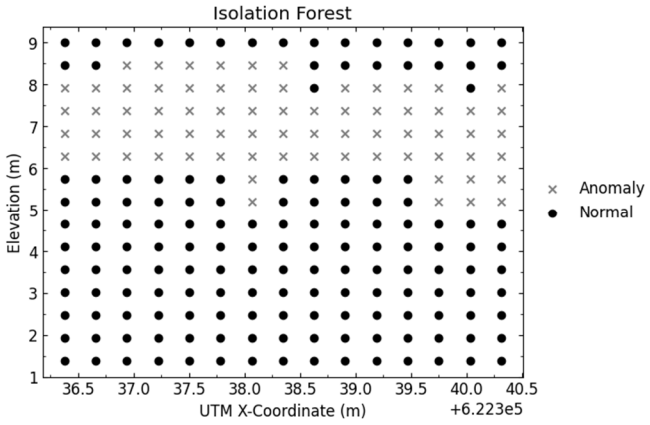


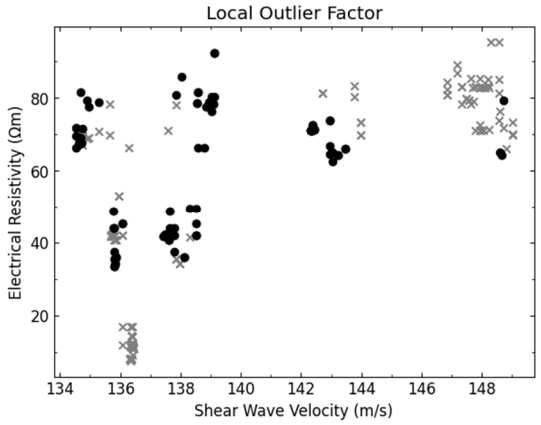


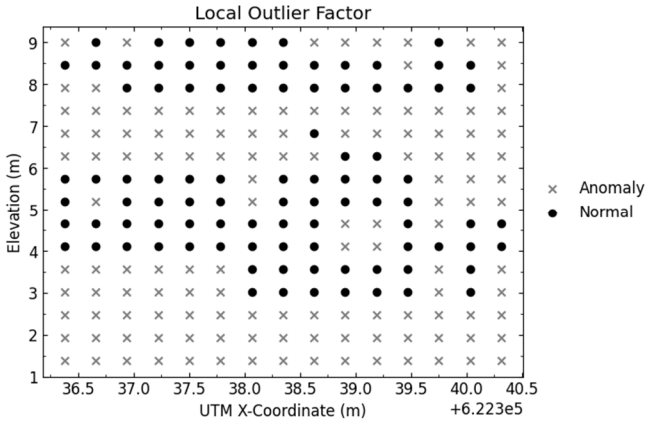


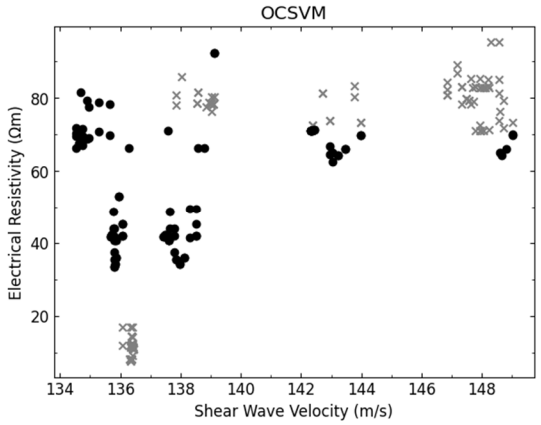
Profile: swt_site3

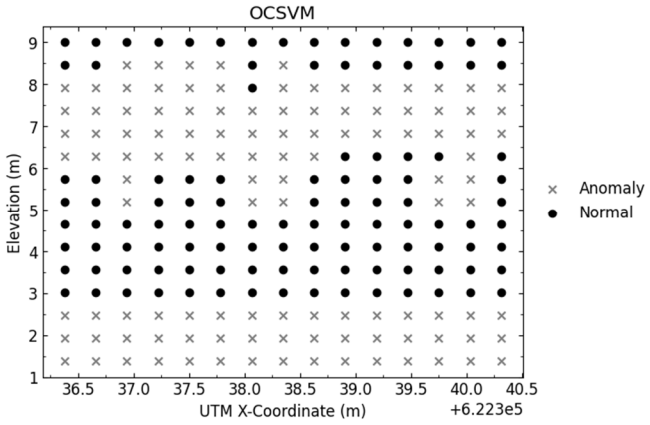






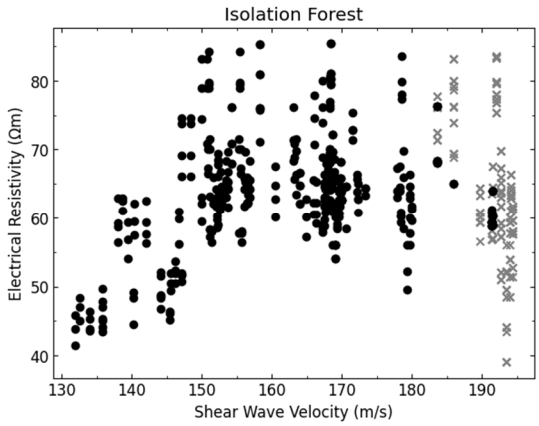


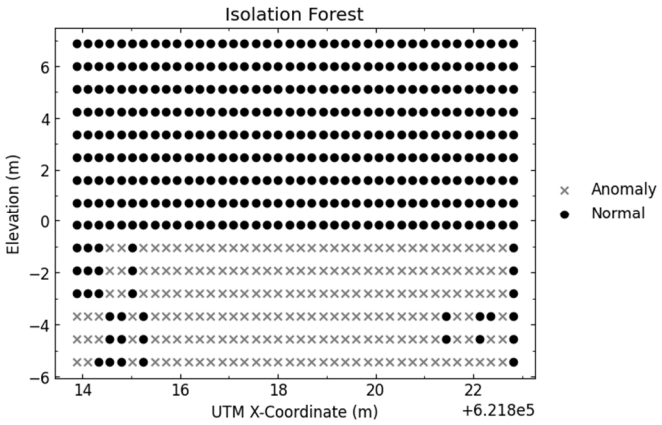


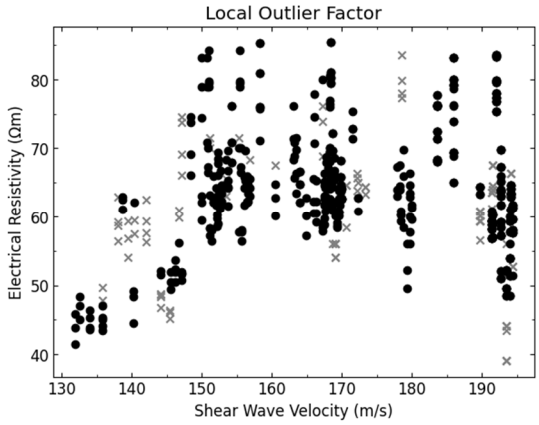


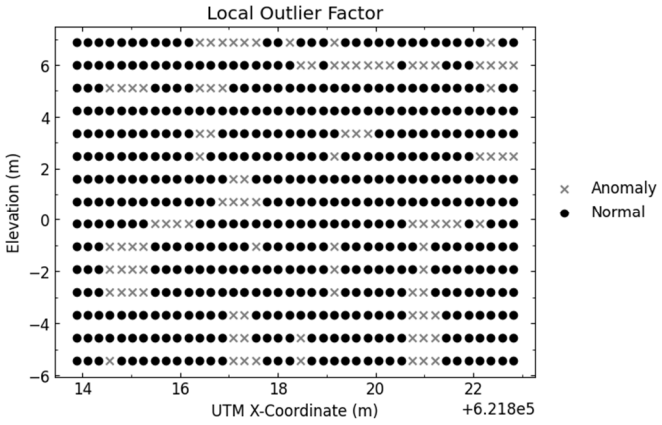
Fit: Dry Season Compilation Dataframe
Predicted: Dry Season Compilation Dataframe

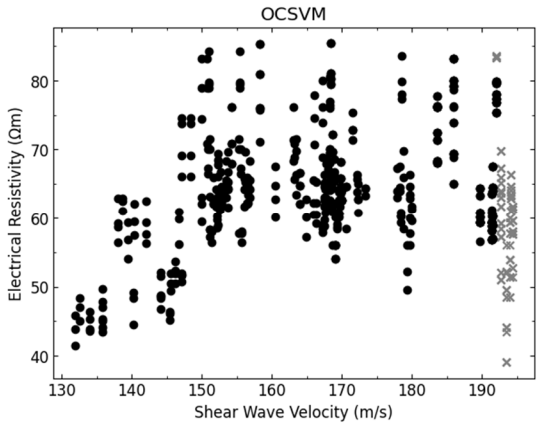
Profile: sc_site2

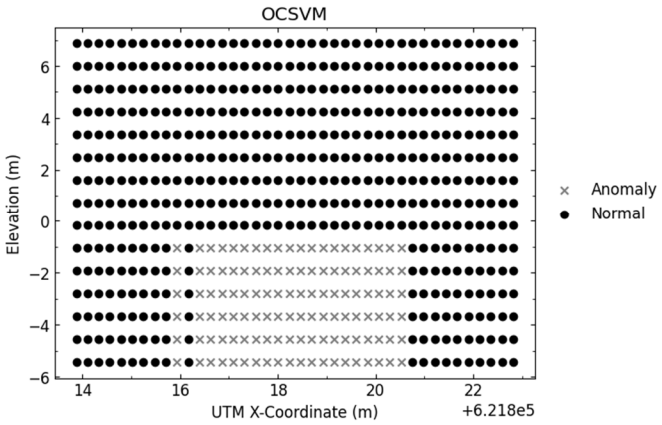




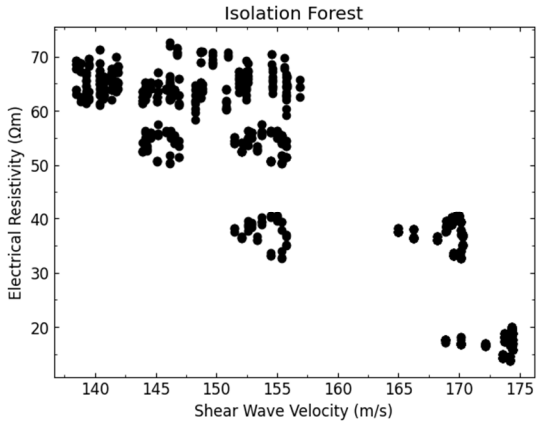


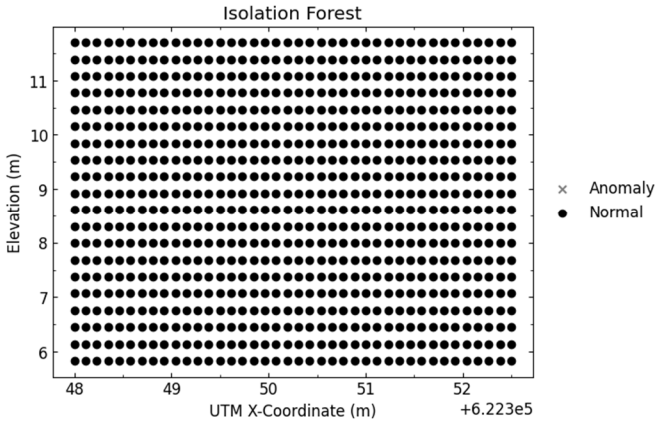


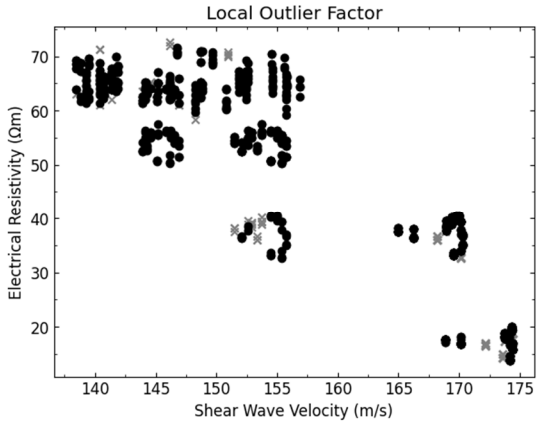


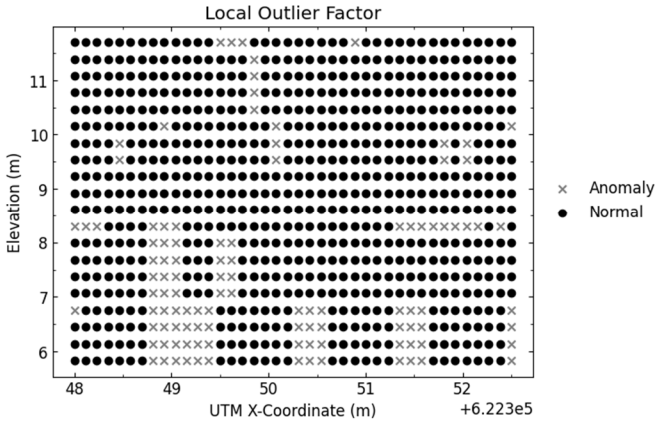


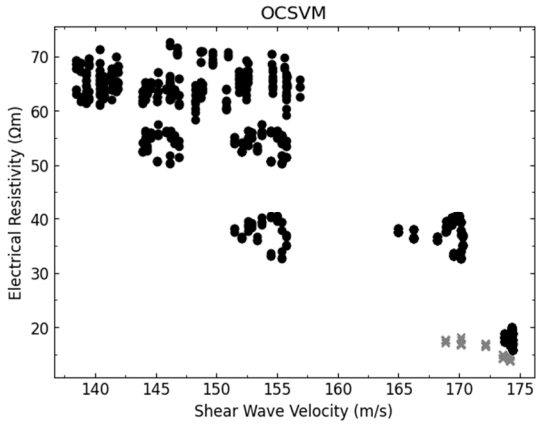
Profile: sc_site3

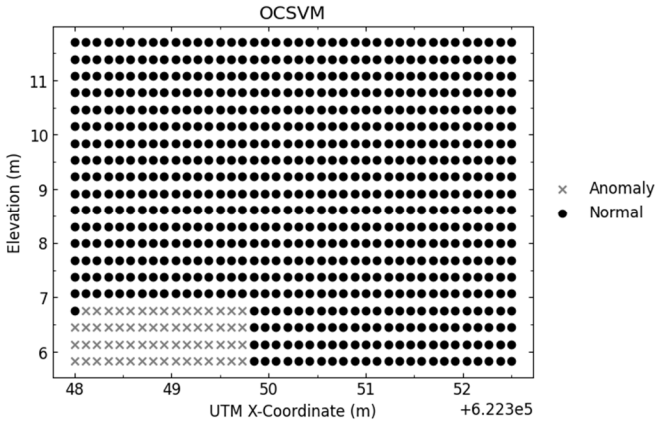




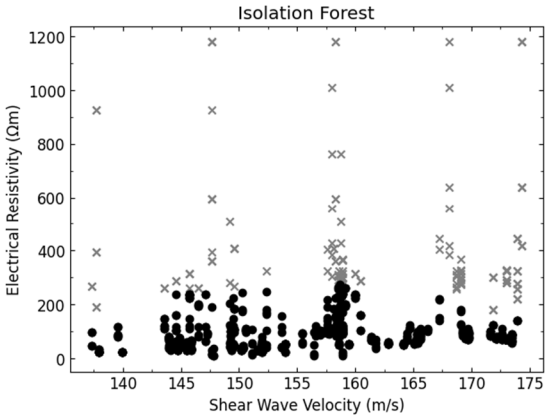


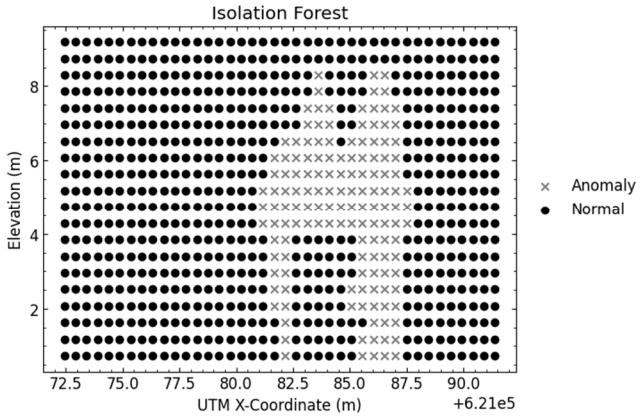


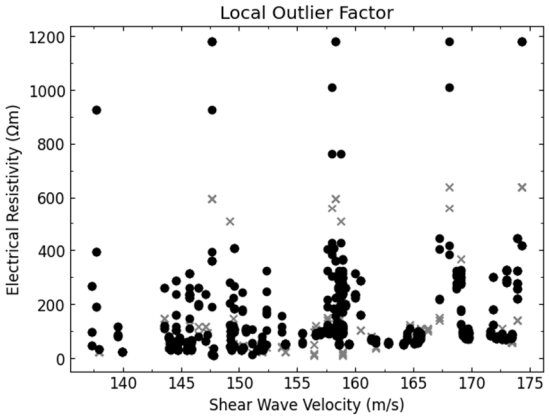


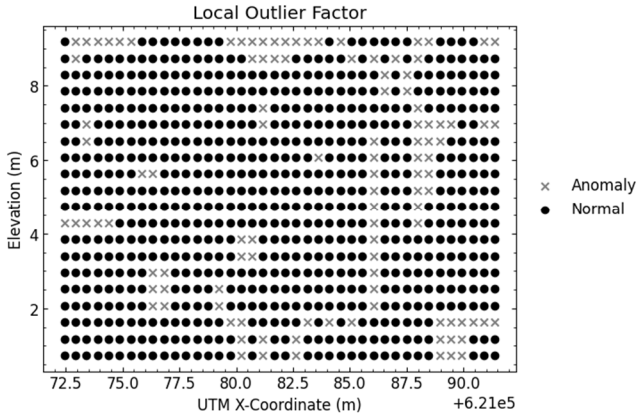


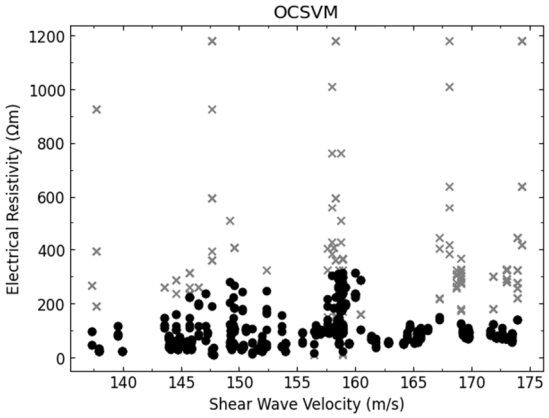
Profile: sc3_site1

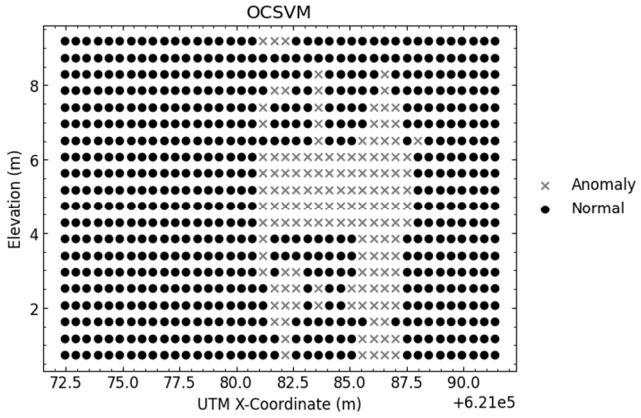




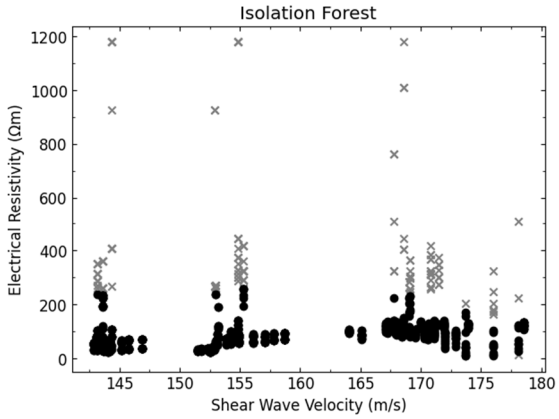


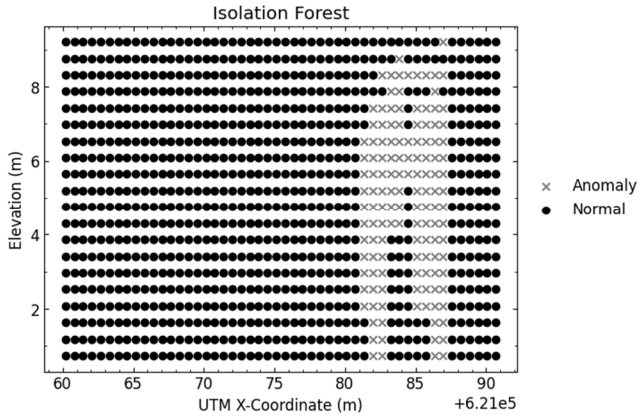


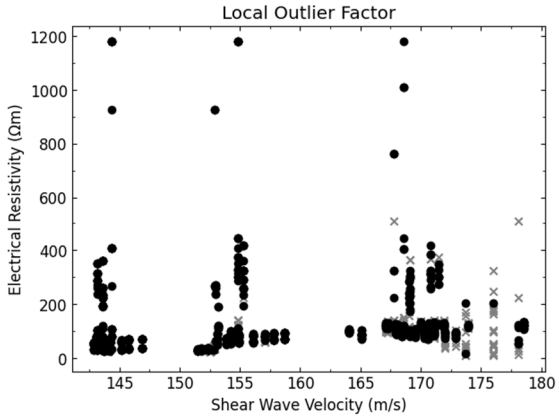


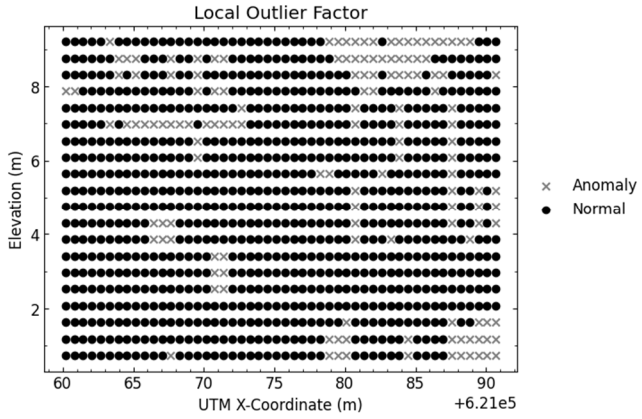


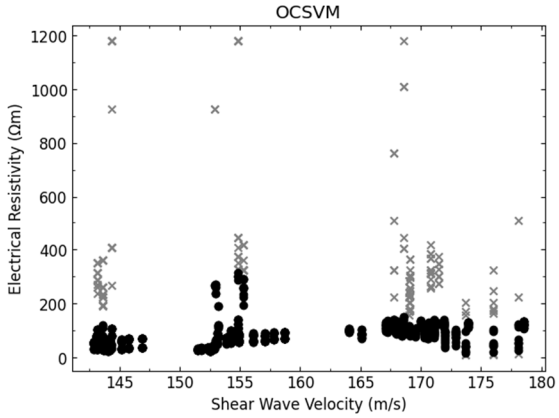
Profile: sc5_site1

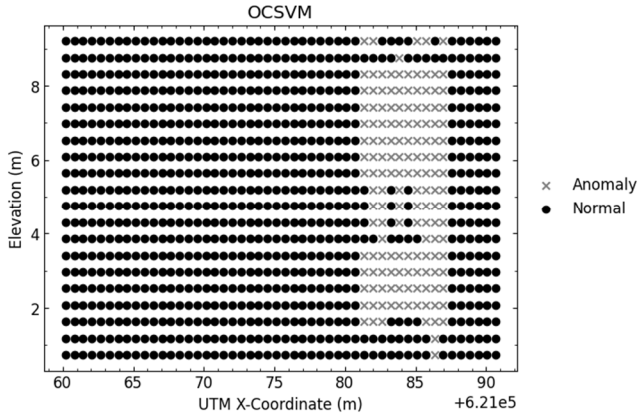




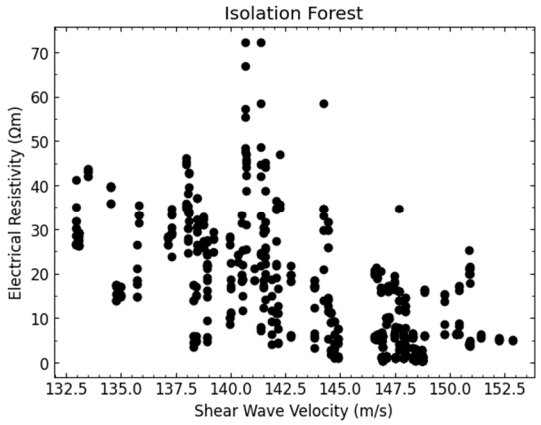


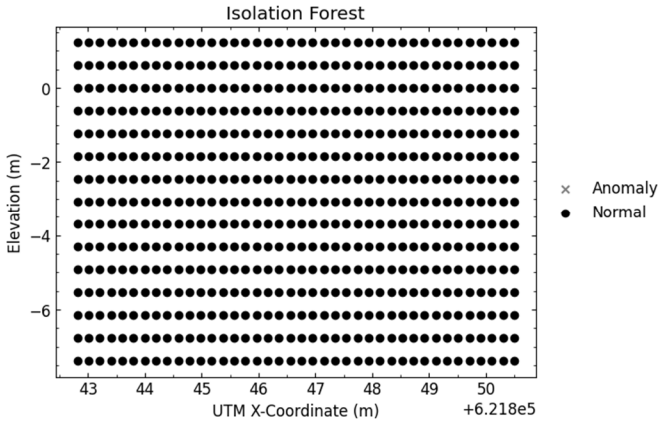


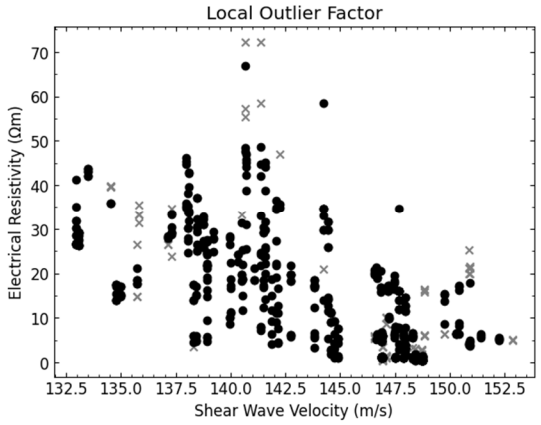


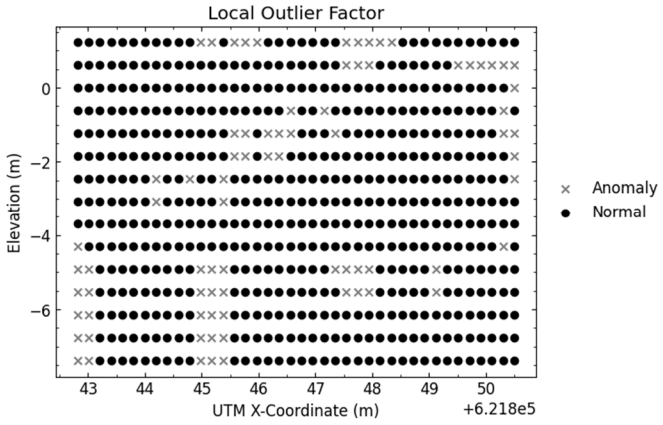


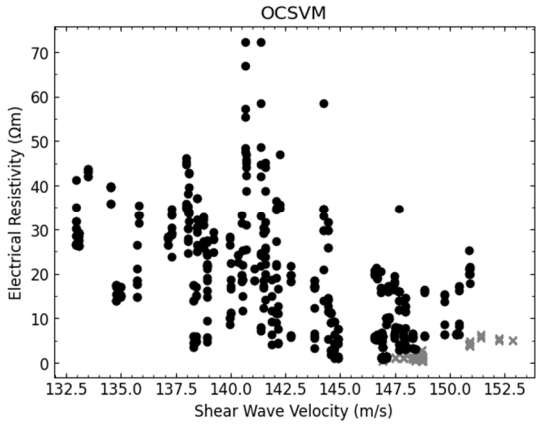
Profile: sdt_site2

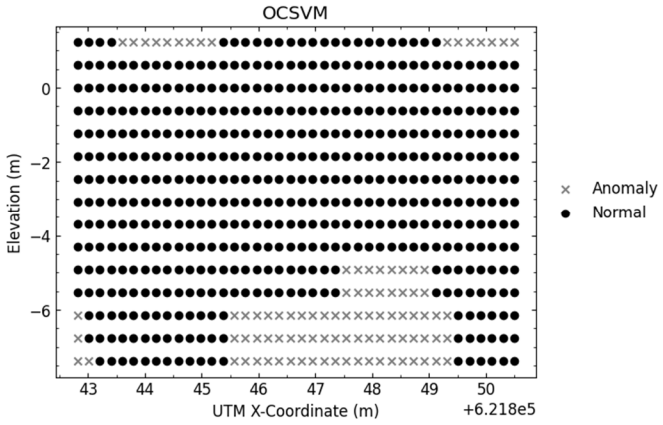




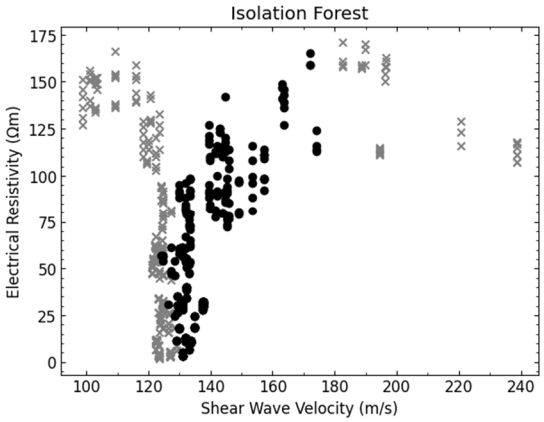


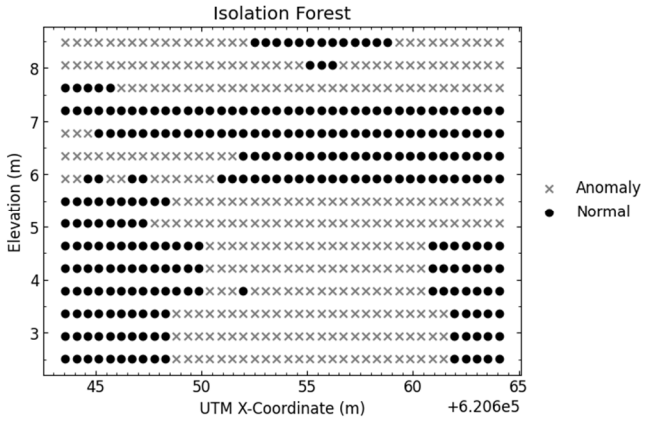


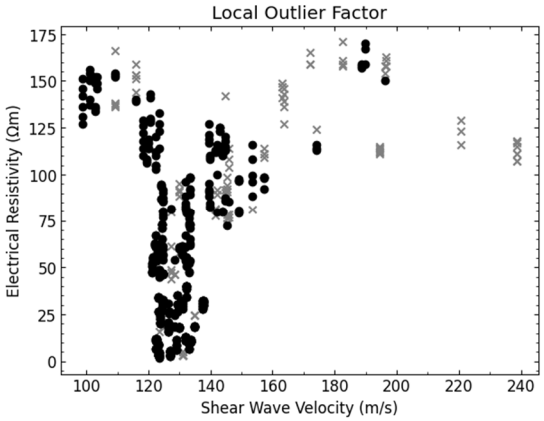


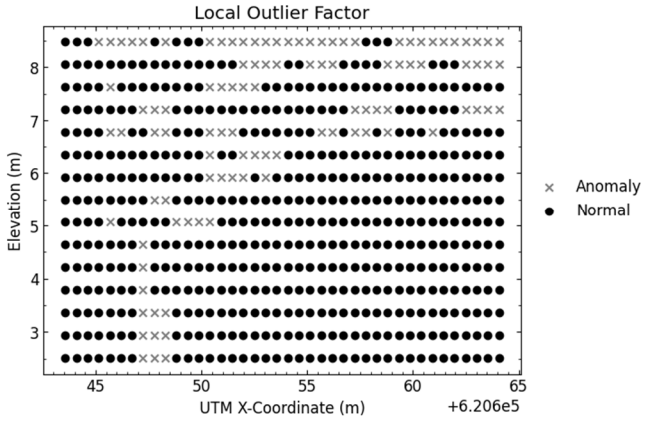


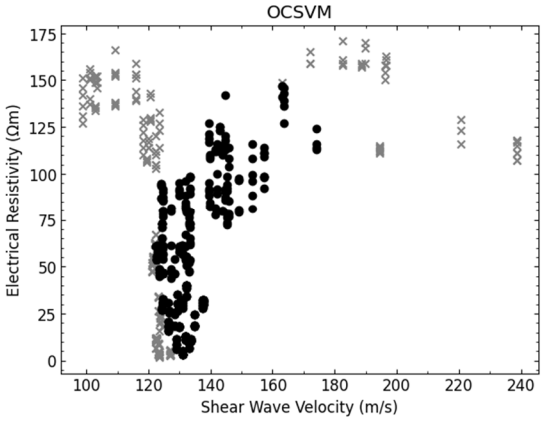
Profile: swt_site1

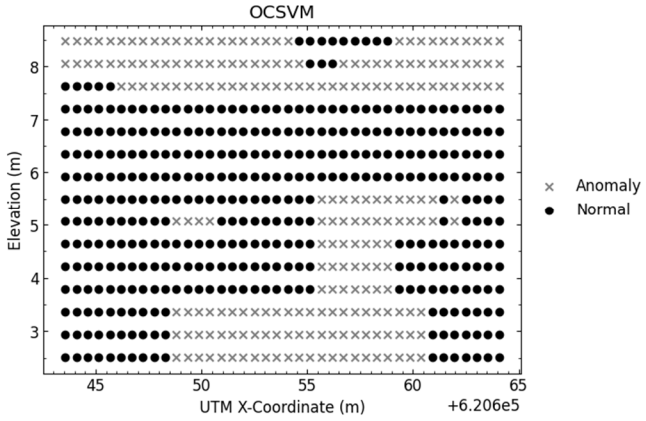




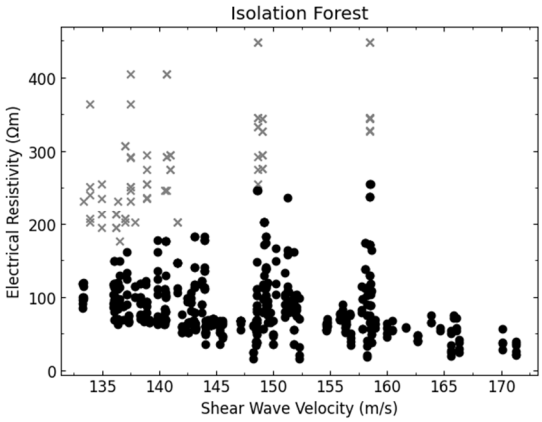


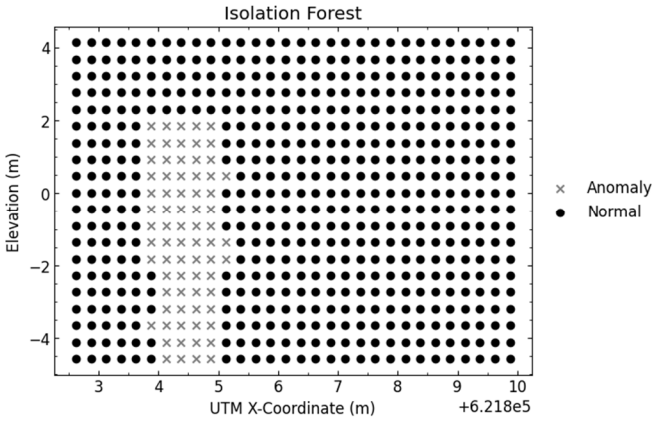


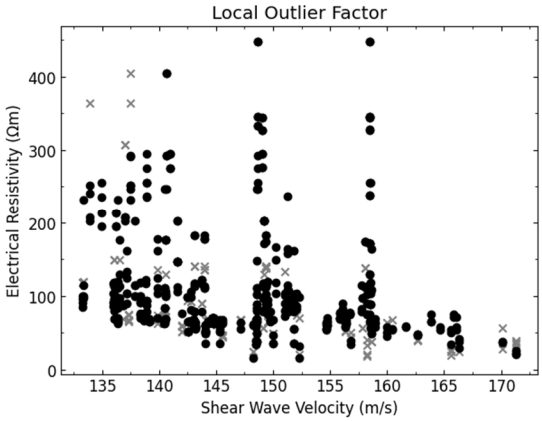


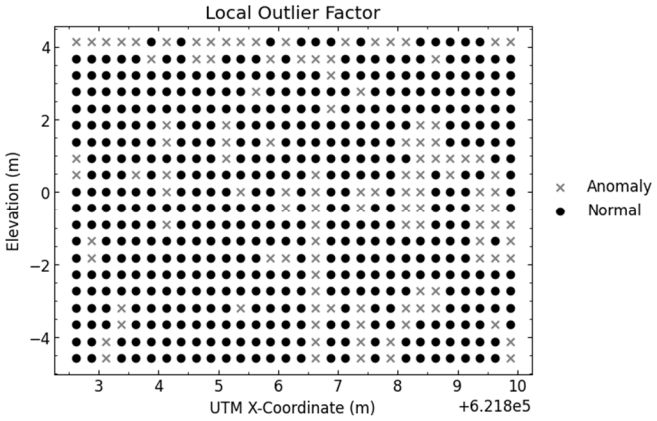


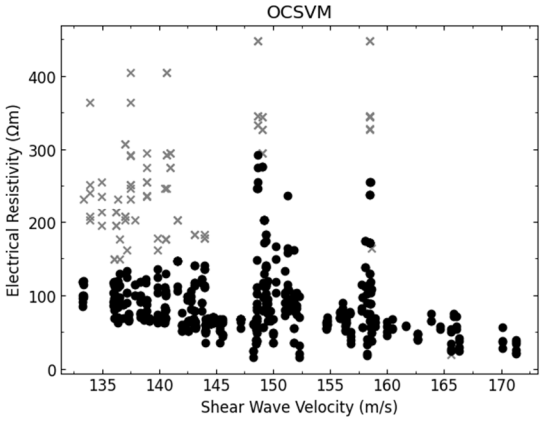
Profile: swt_site2

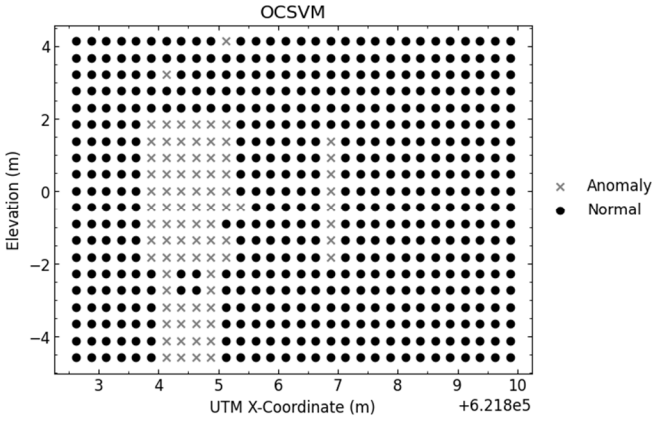




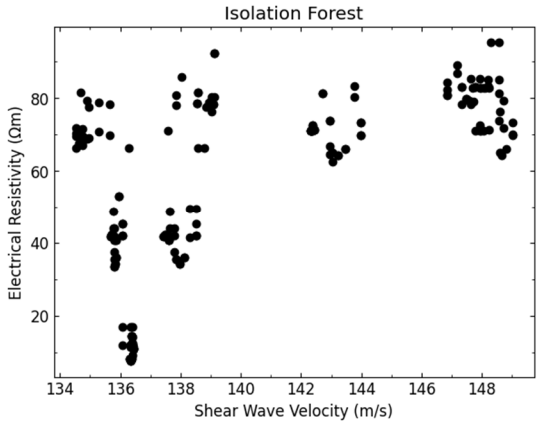


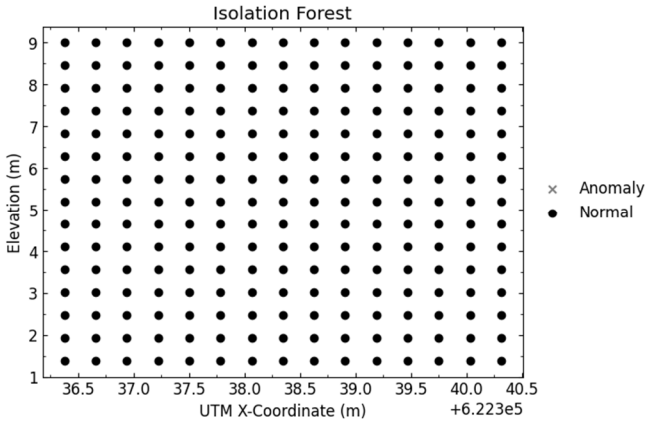


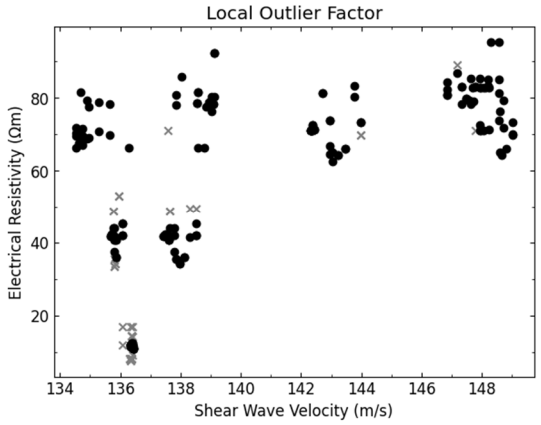


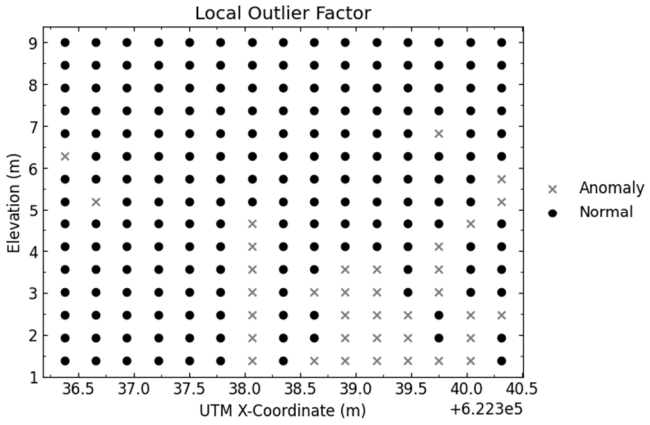


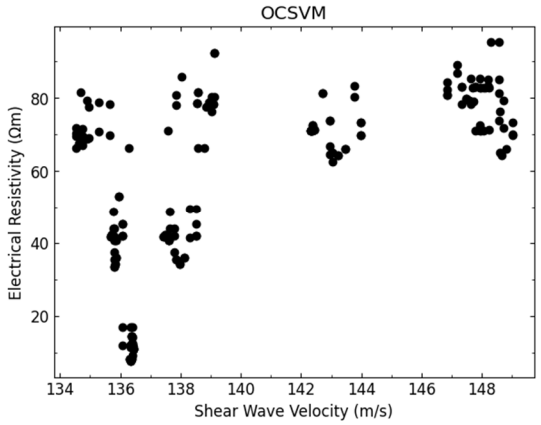
Profile: swt_site3

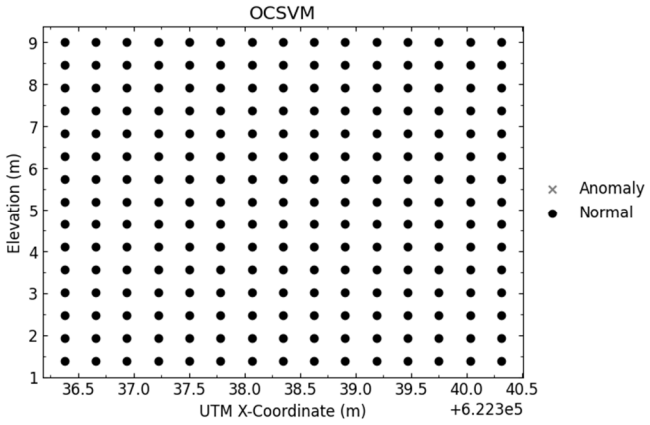






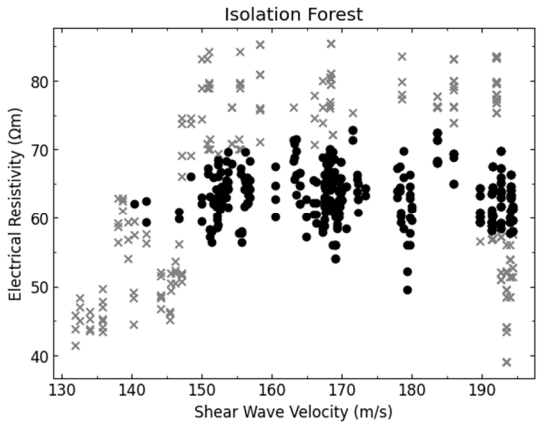


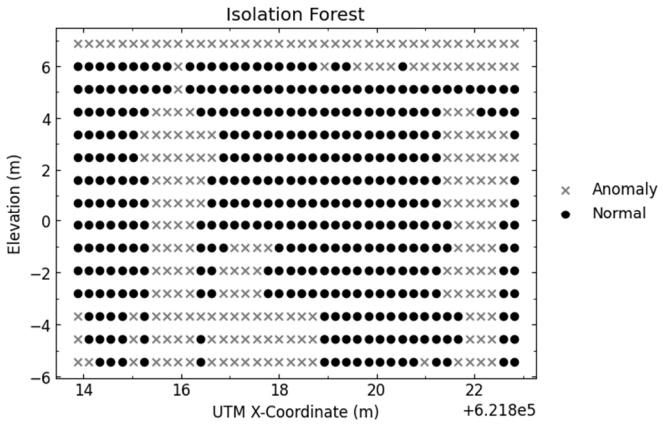


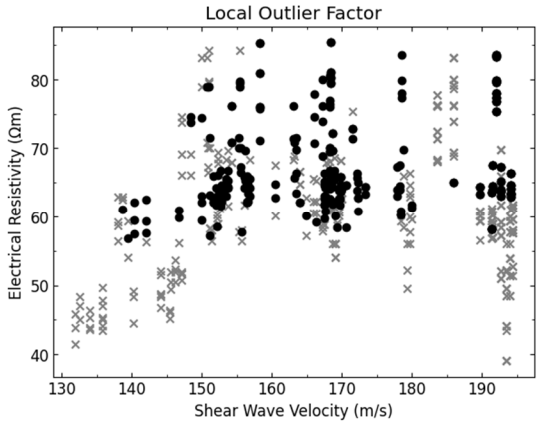


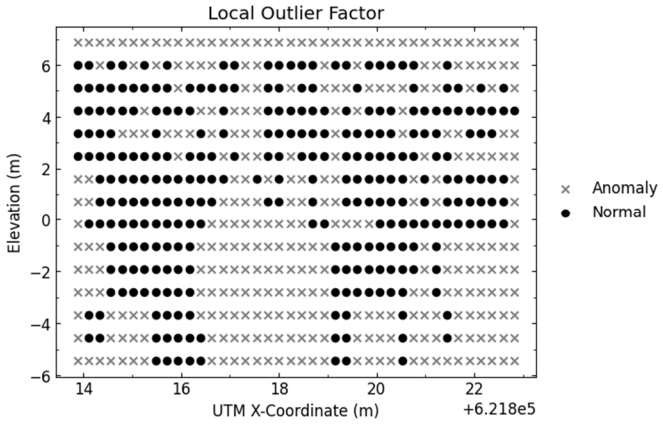
Fit: Dry Season Compilation Dataframe
Predicted: Individual Profiles

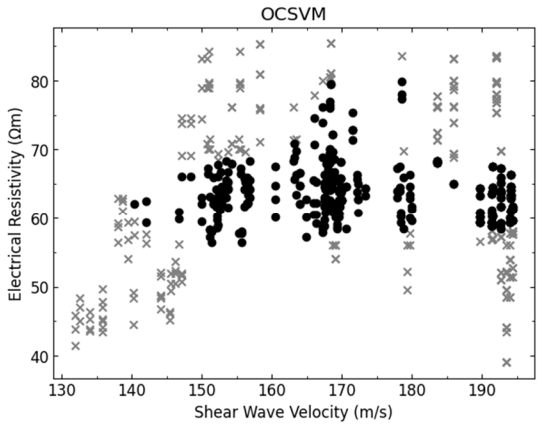
Profile: sc_site2

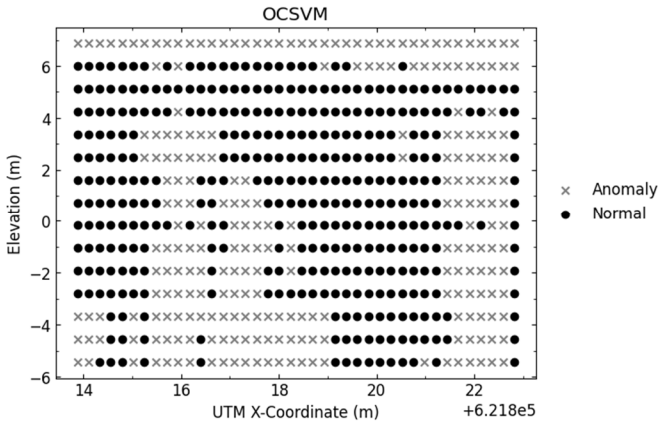




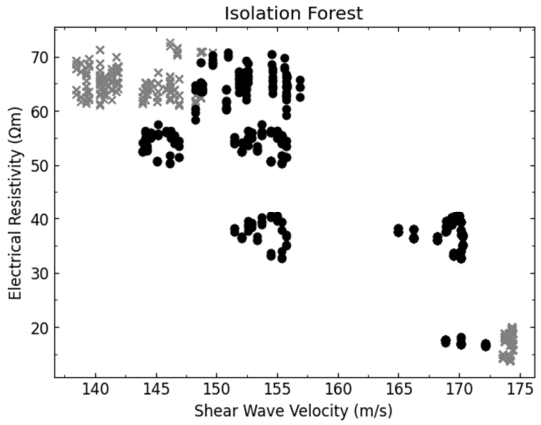


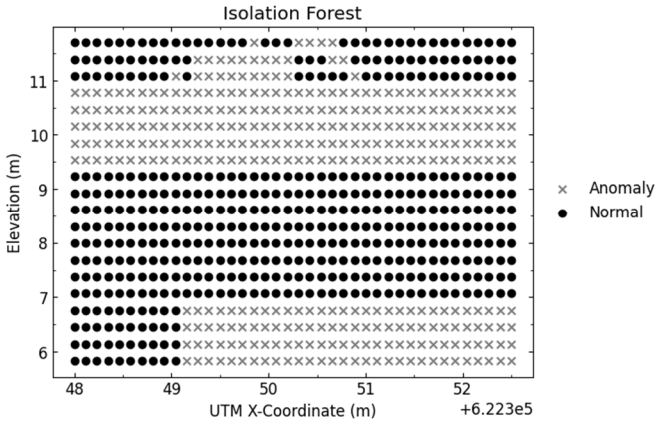


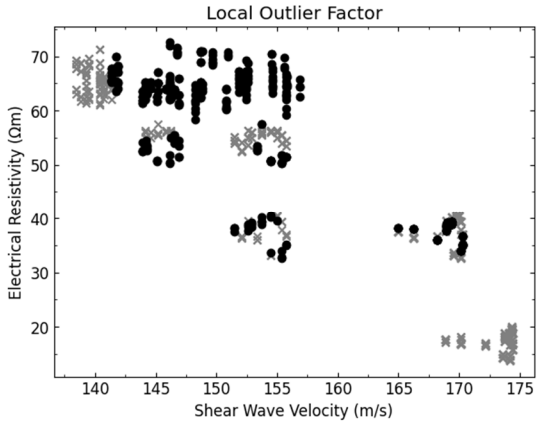


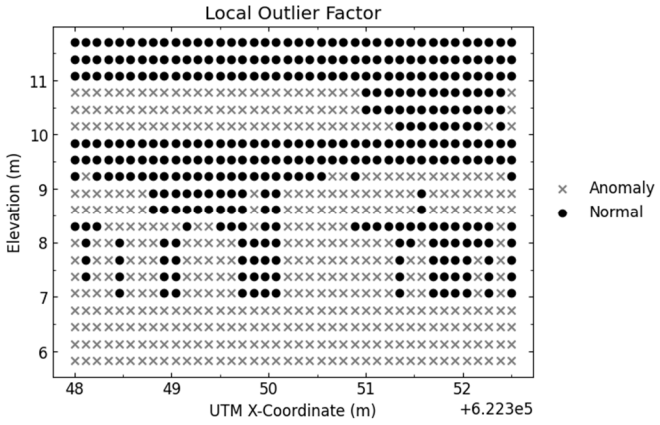


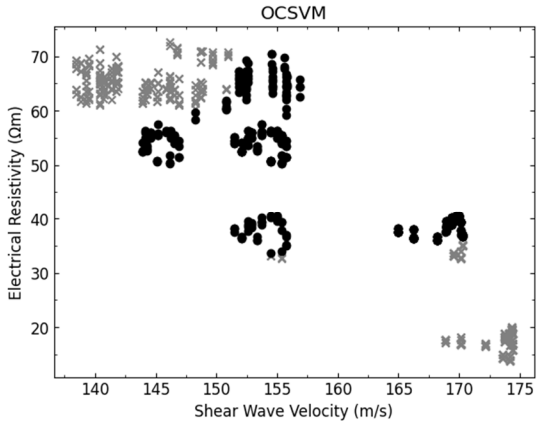
Profile: sc_site3

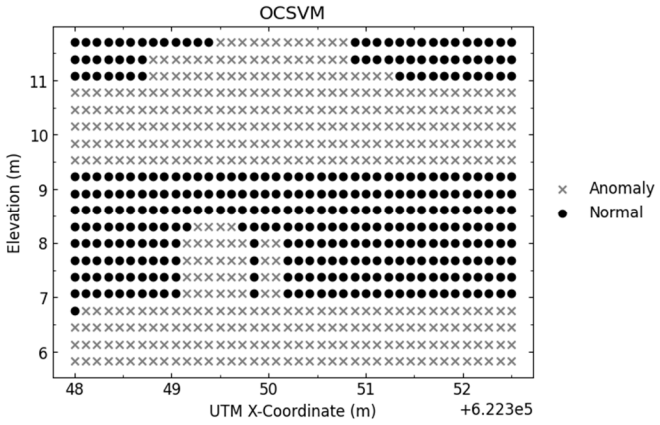




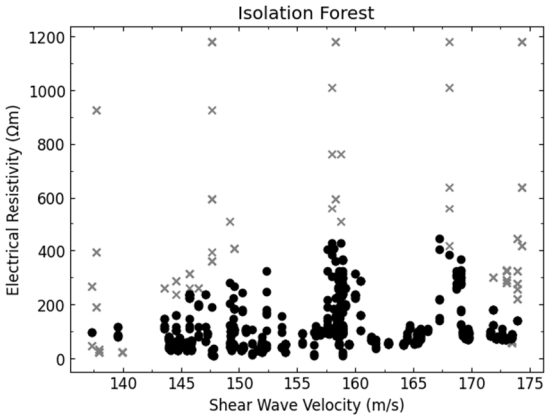


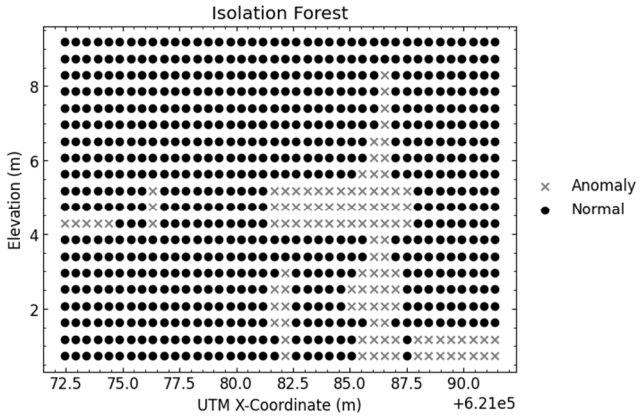


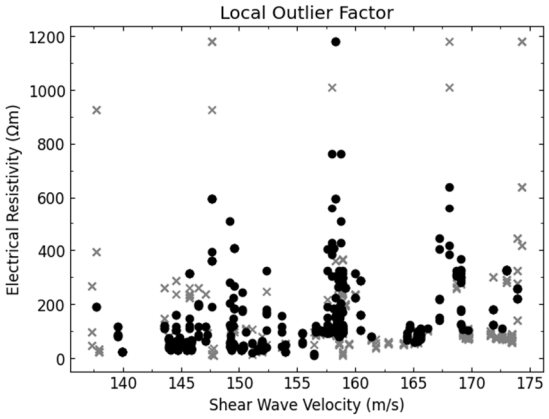


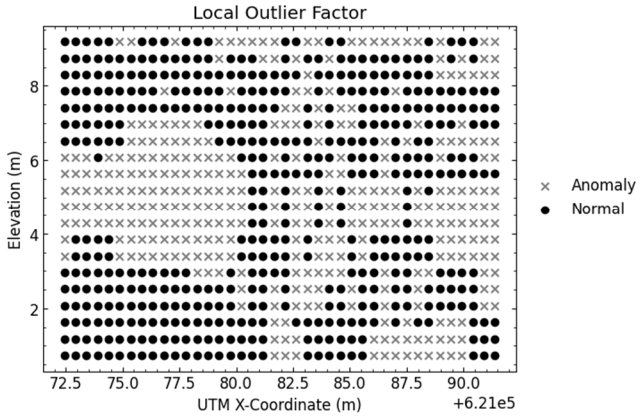


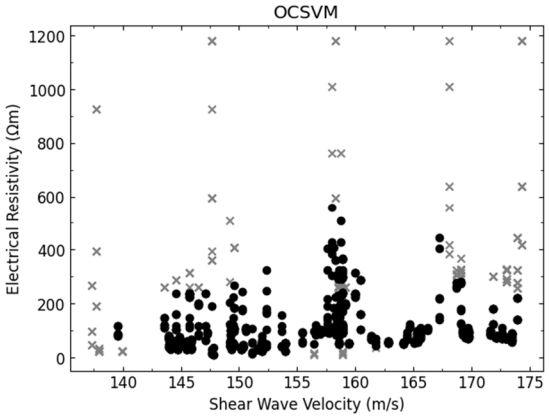
Profile: sc3_site1

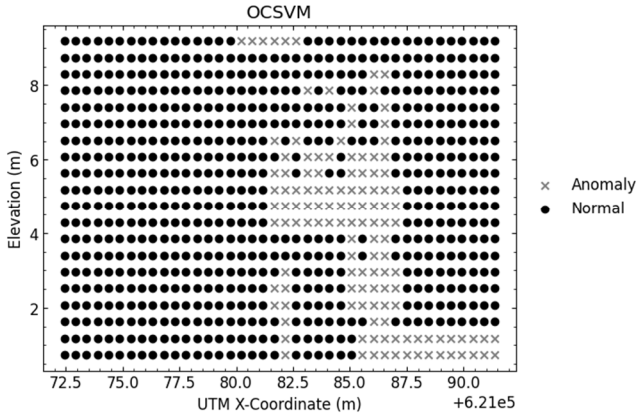




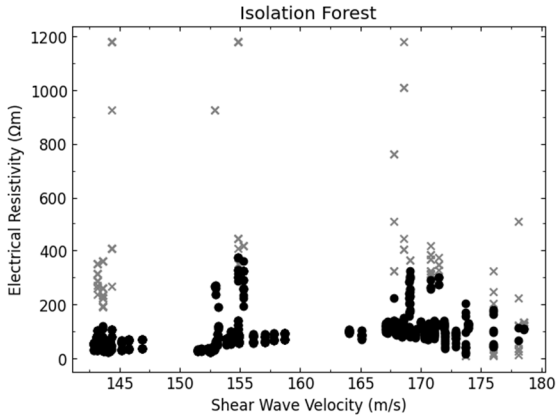


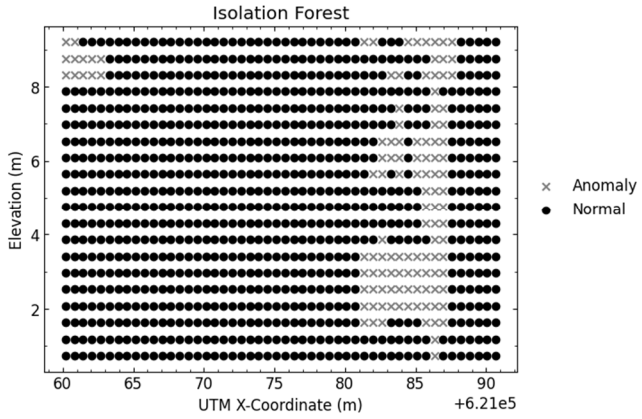


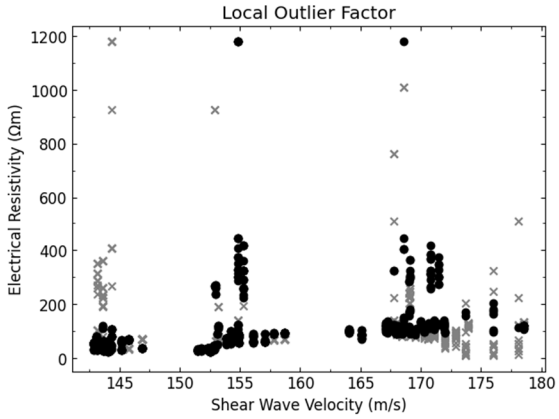


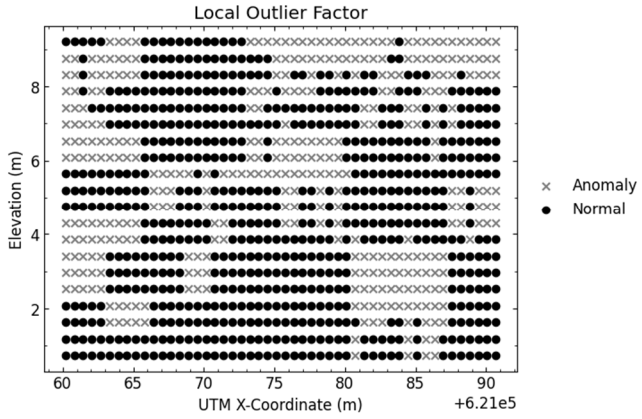


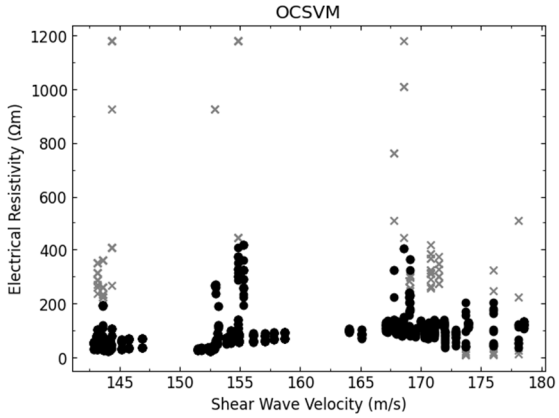
Profile: sc5_site1

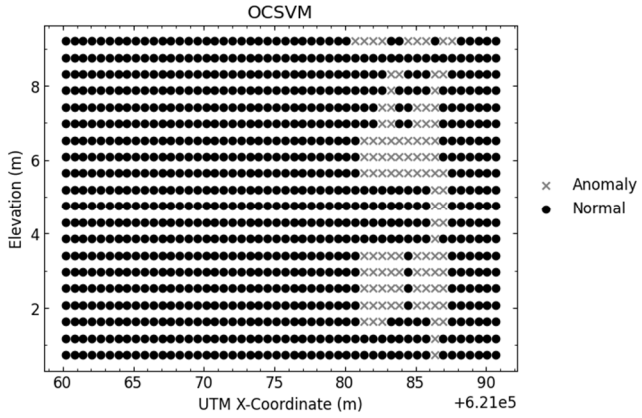




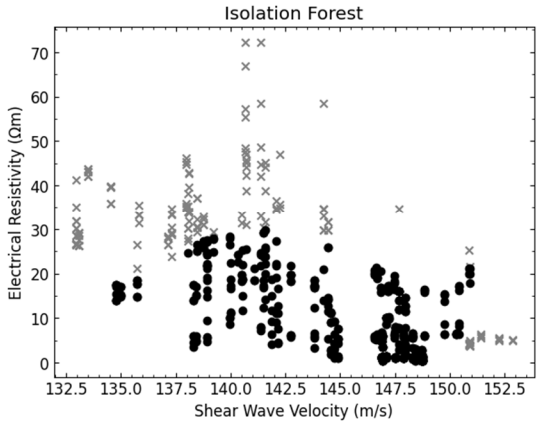


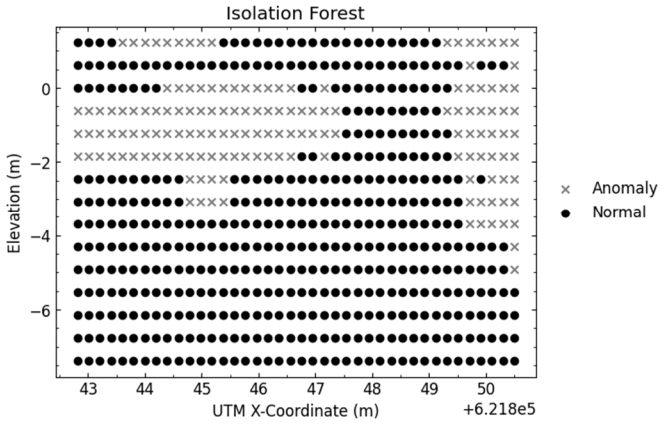


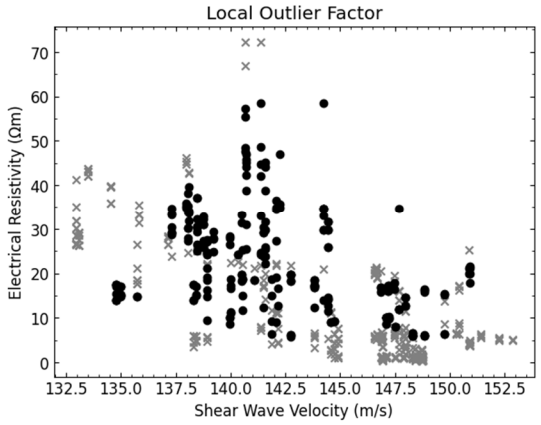


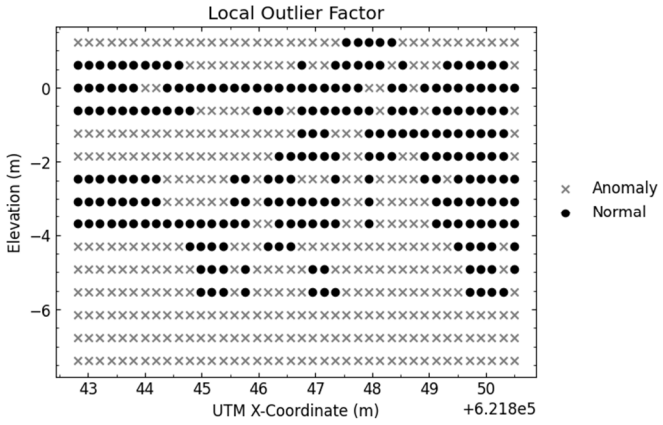


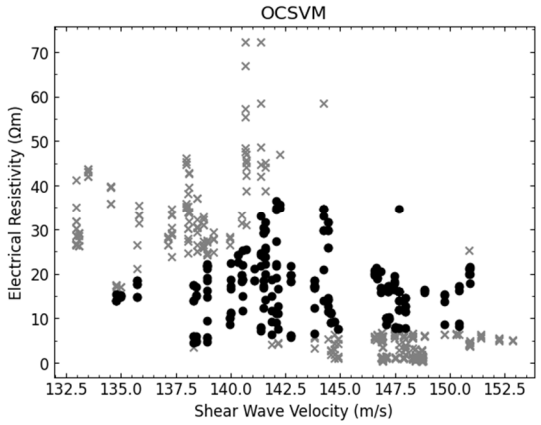
Profile: sdt_site2

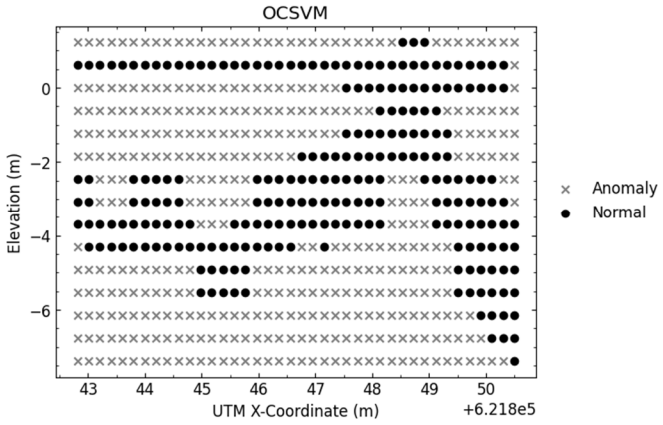




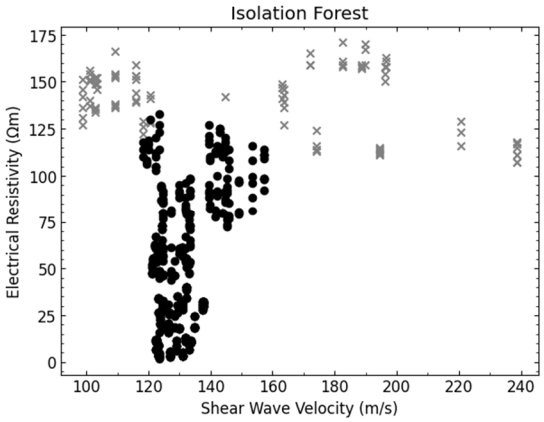


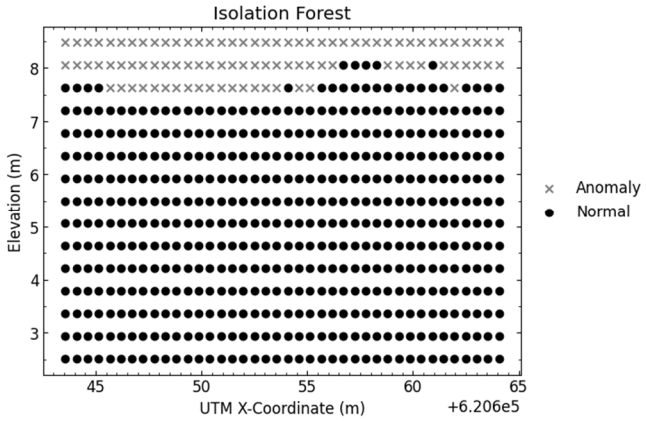


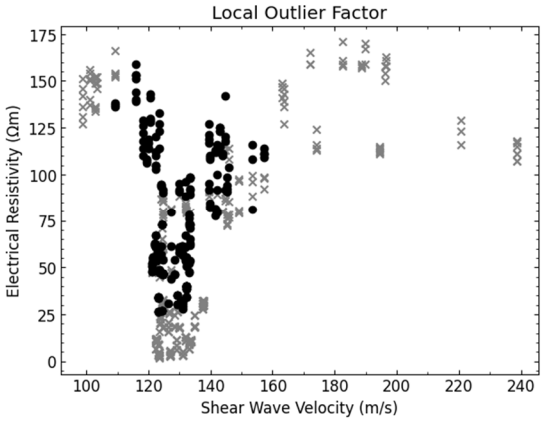


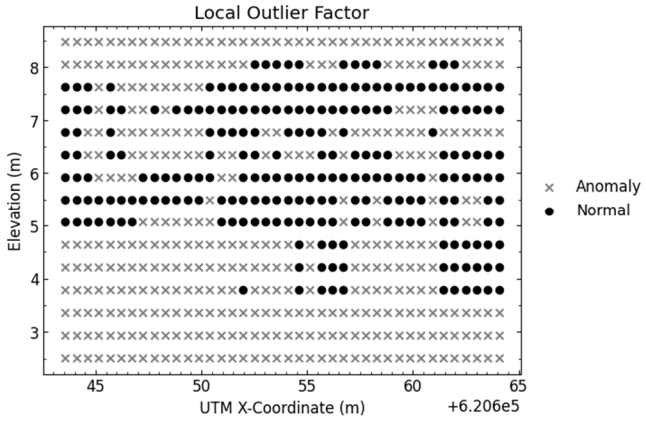


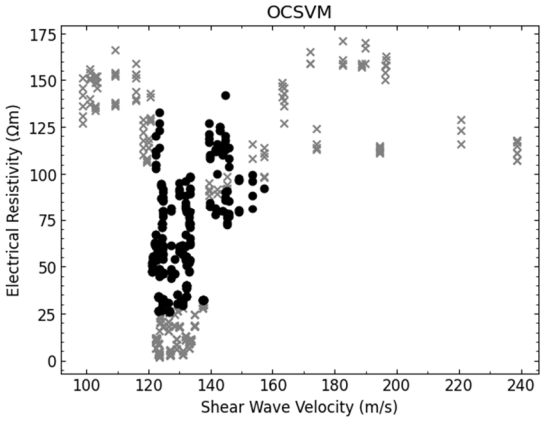
Profile: swt_site1

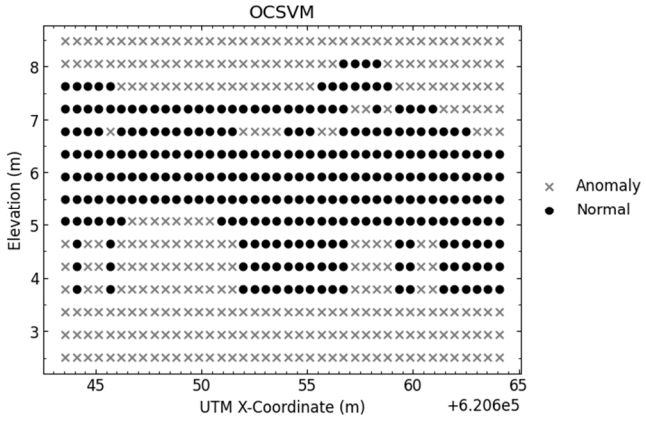




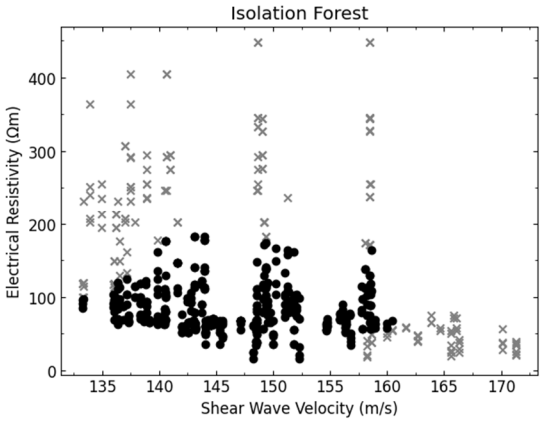


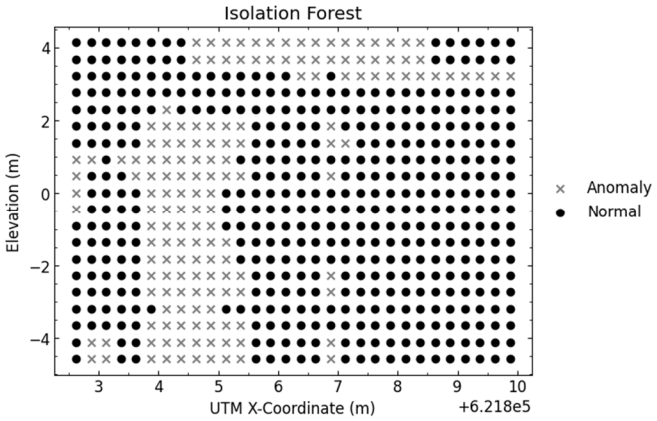


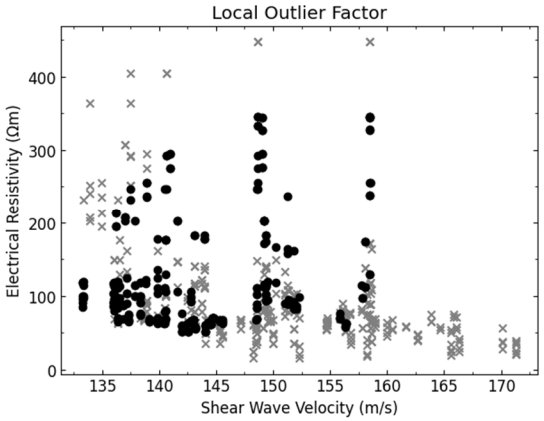


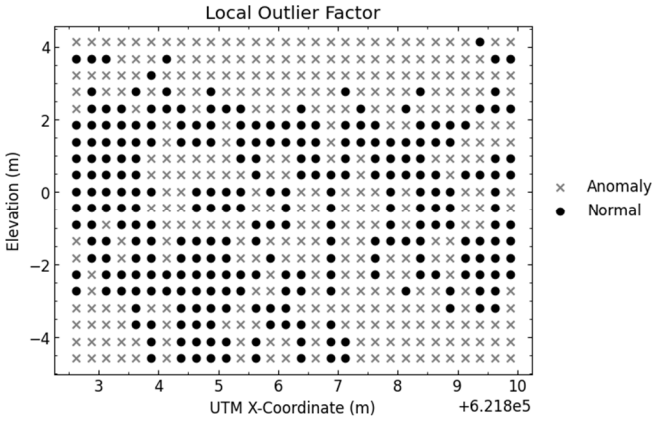


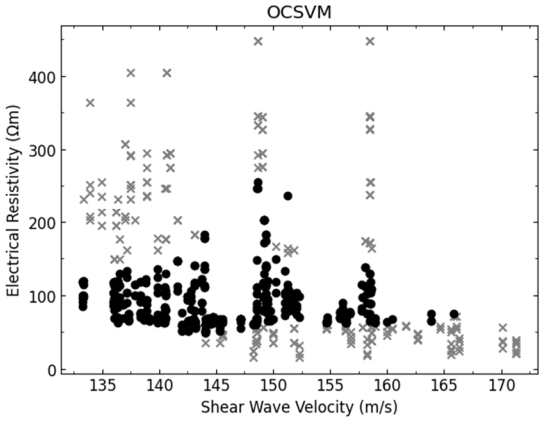
Profile: swt_site2

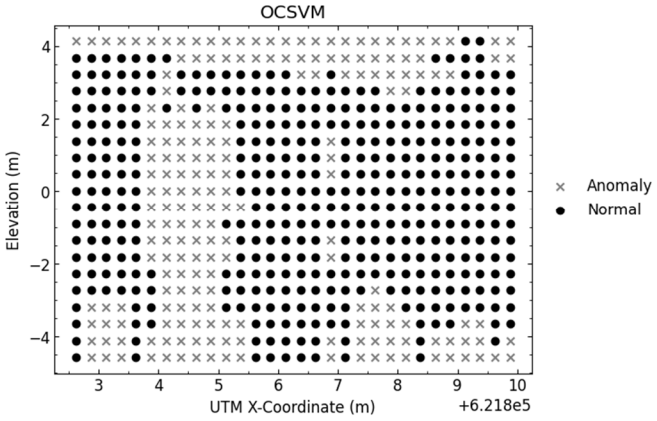




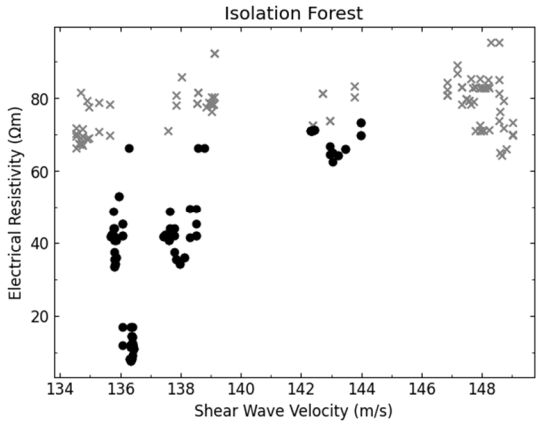


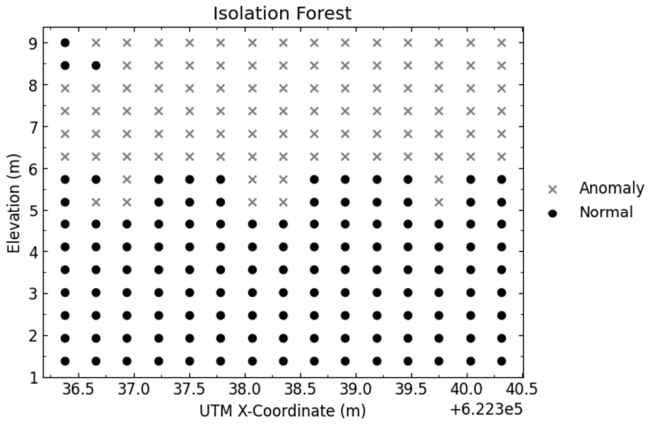


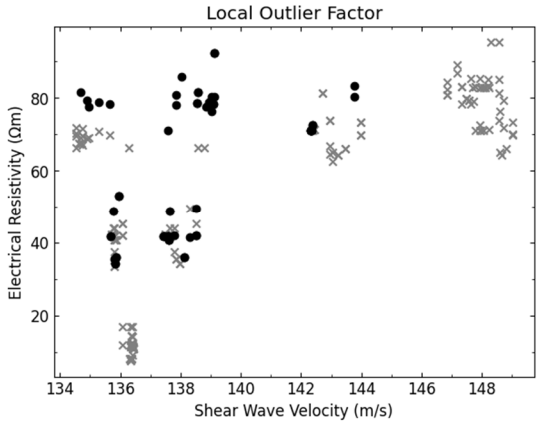


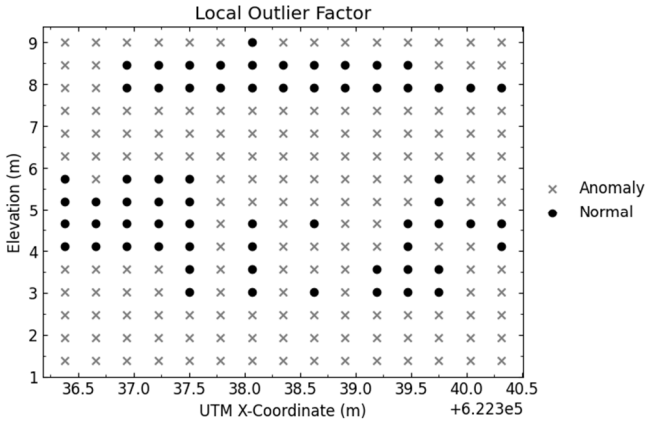


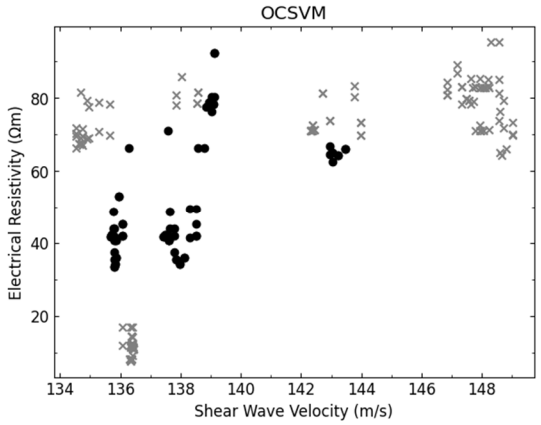
Profile: swt_site3

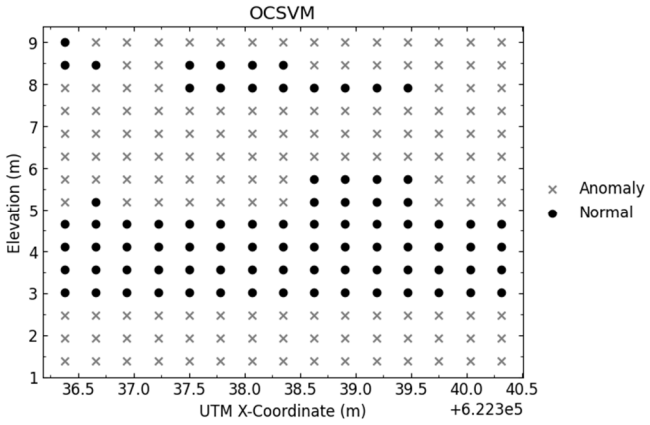






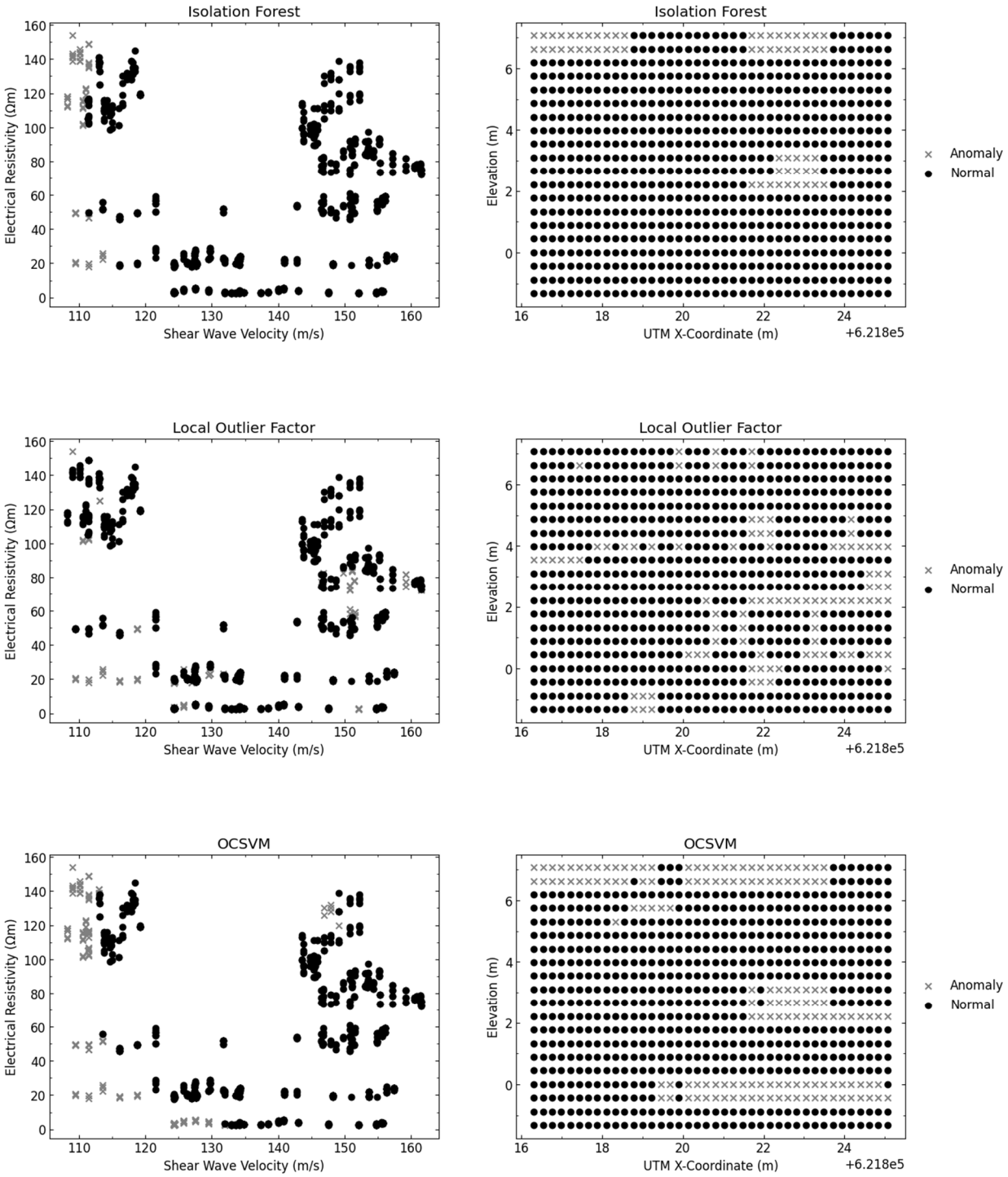




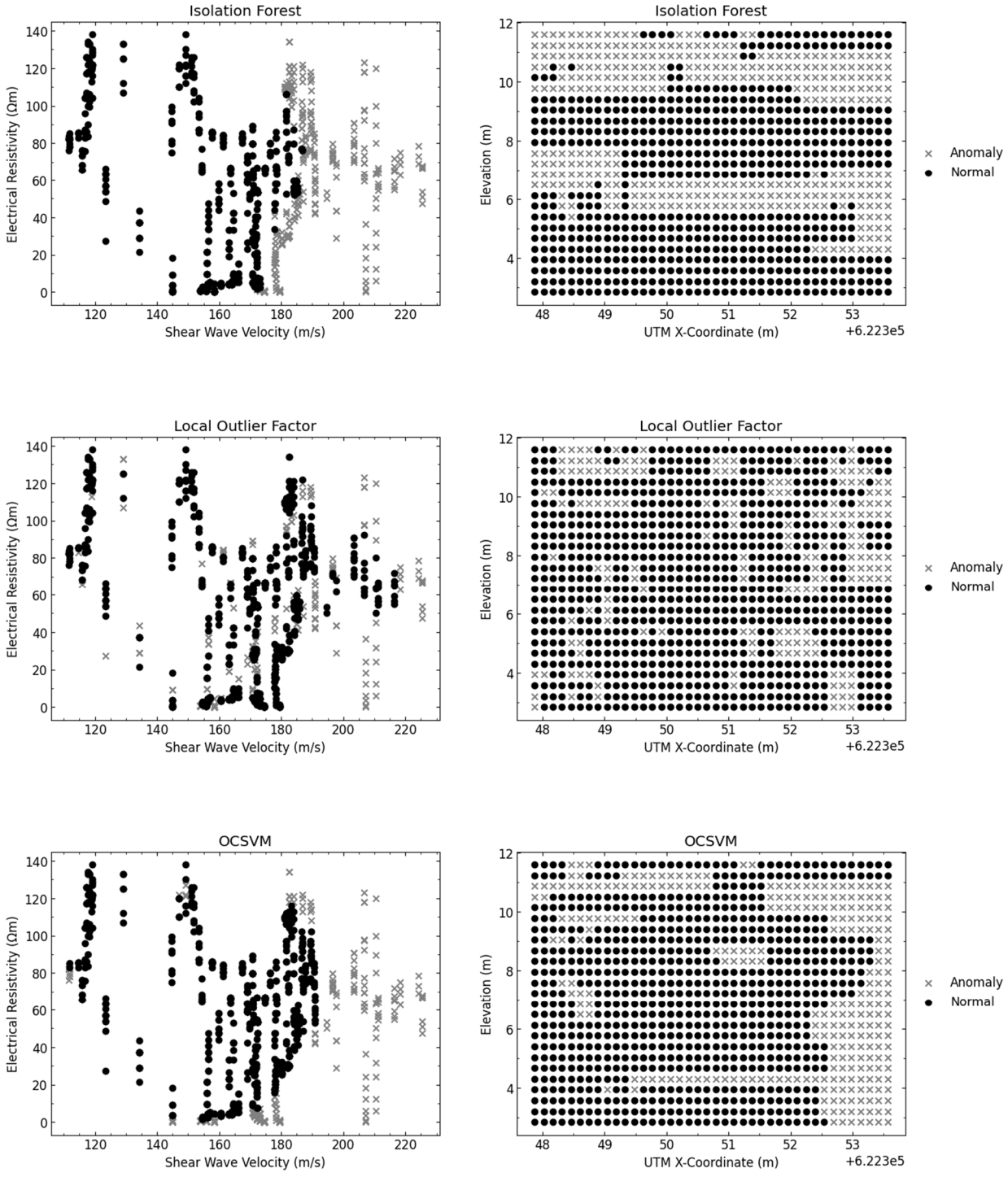


Fit: Wet Season Compilation Dataframe
Predicted: Wet Season Compilation Dataframe

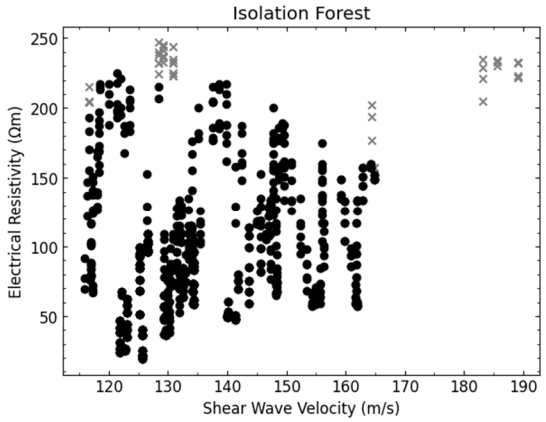
Profile: fc_site2

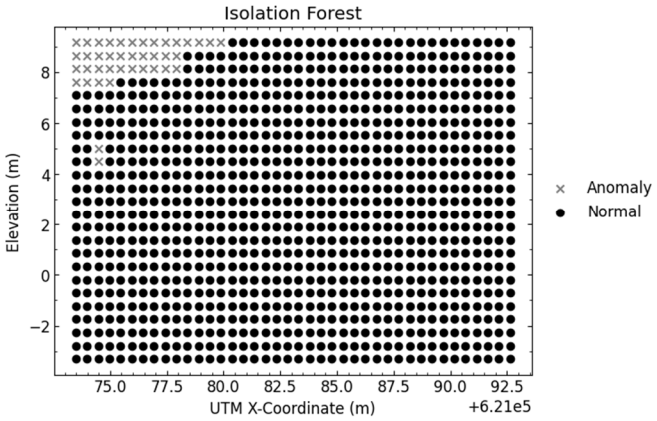


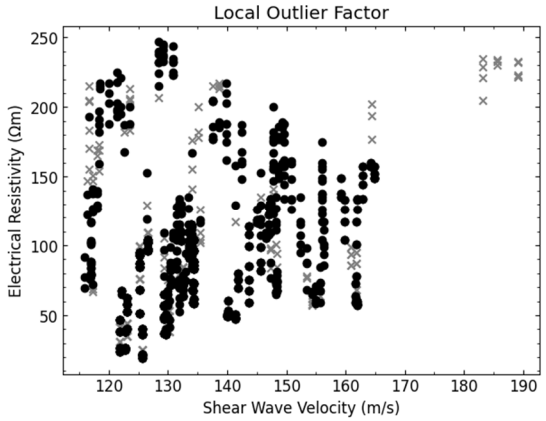
Profile: fc_site3

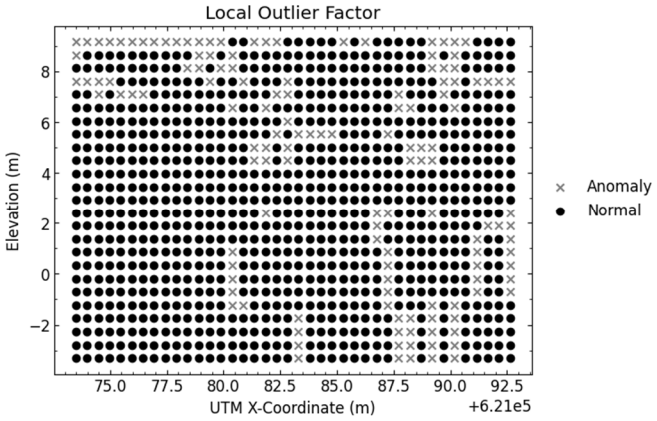


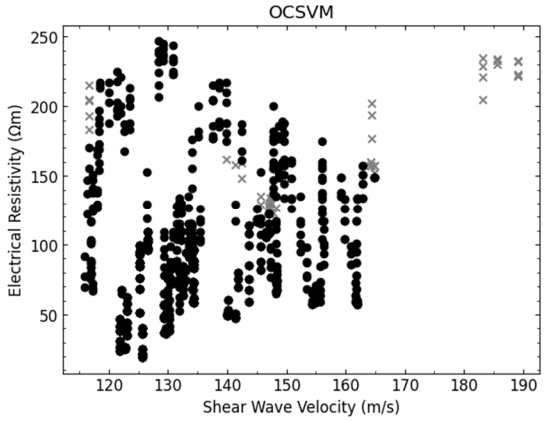
Profile: fc3_site1

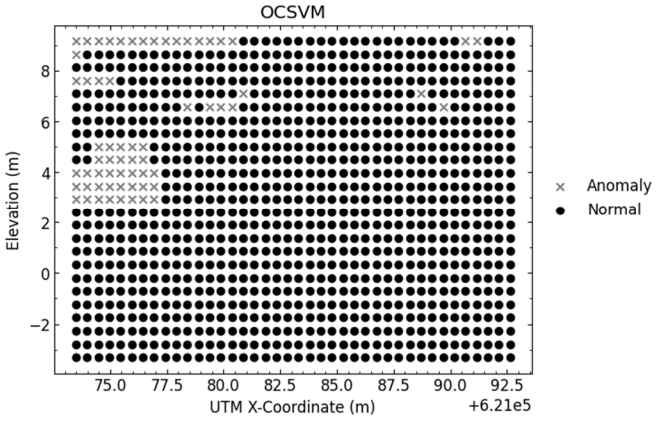




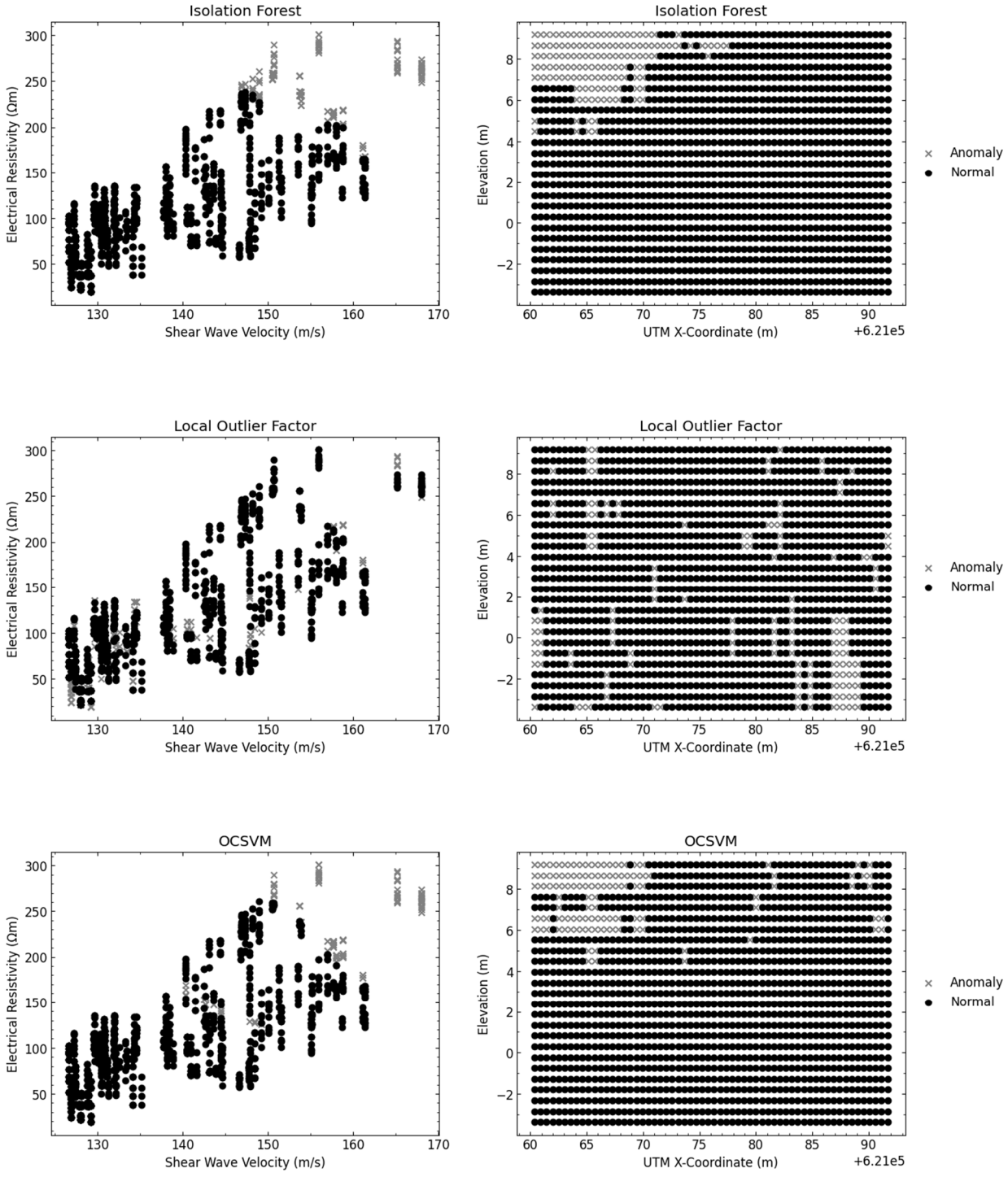




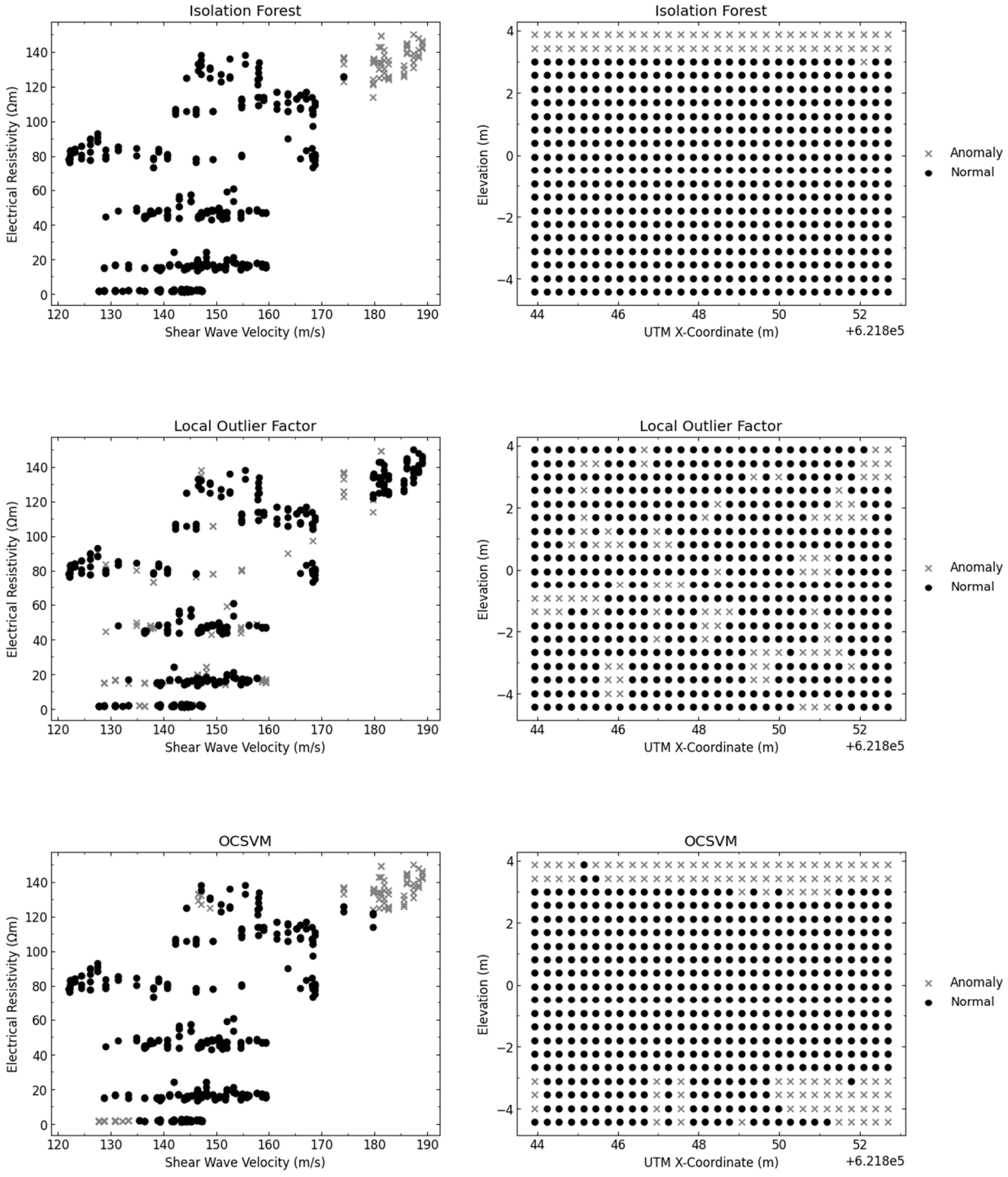




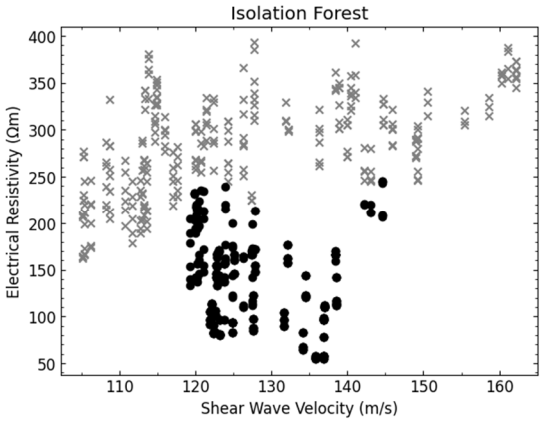
Profile: fc5_site1

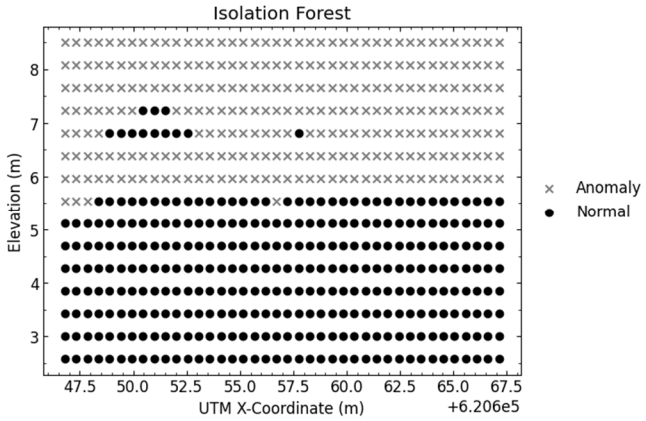


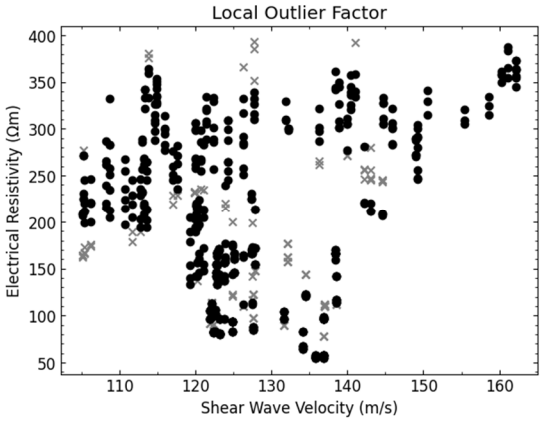
Profile: fdt_site2

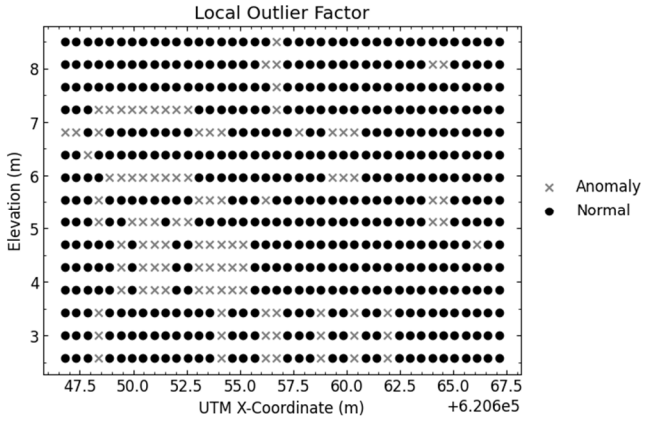


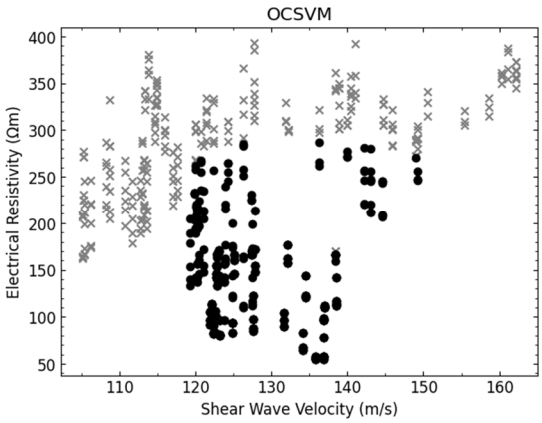
Profile: fwt_site1

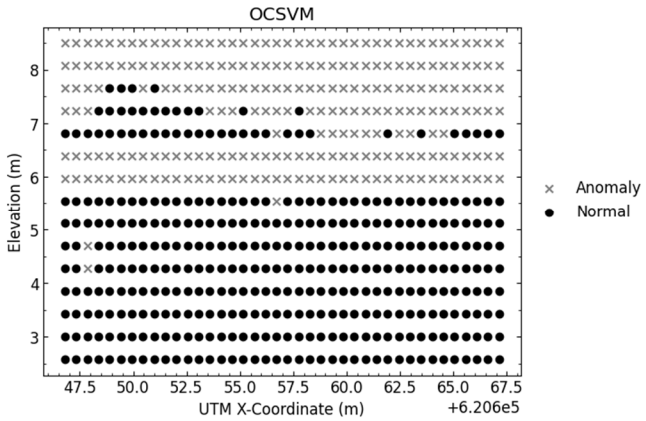




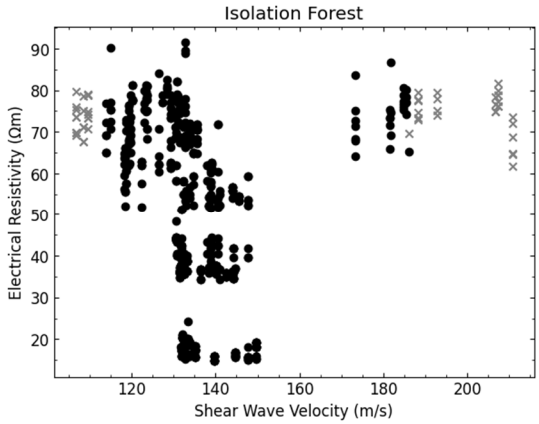


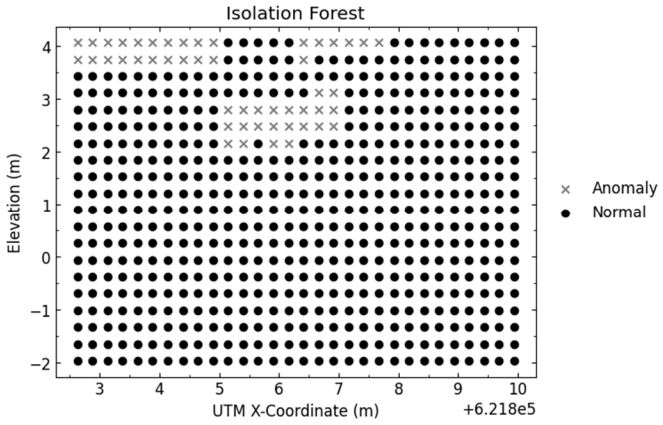


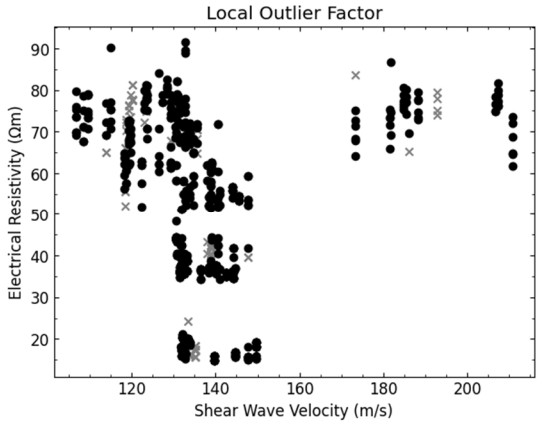


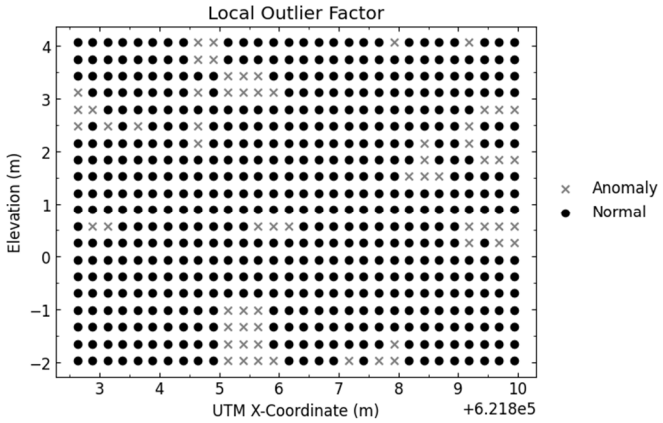


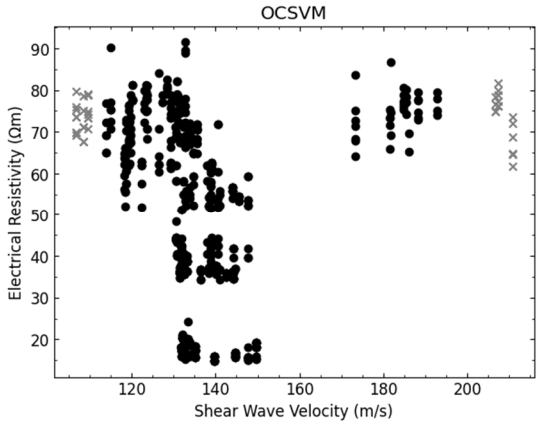
Profile: fwt_site2

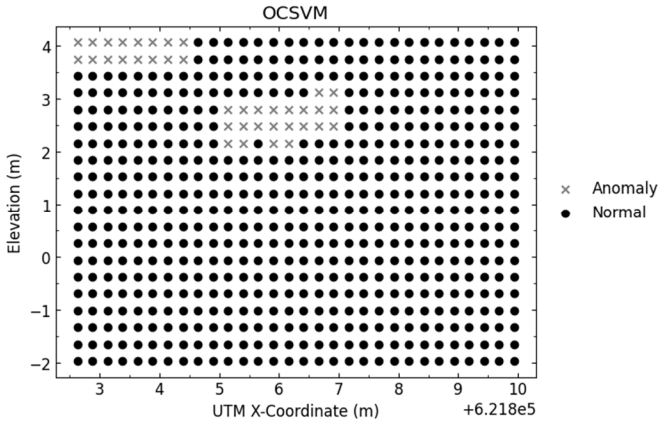




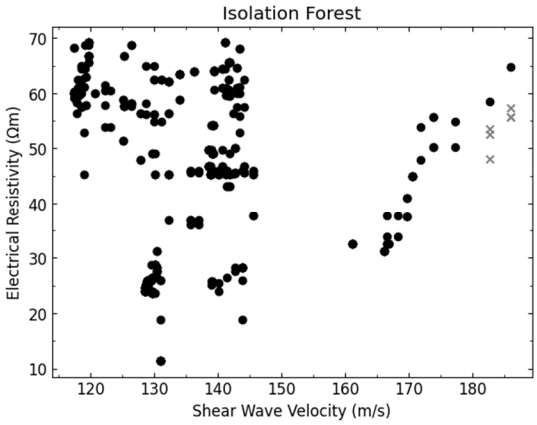


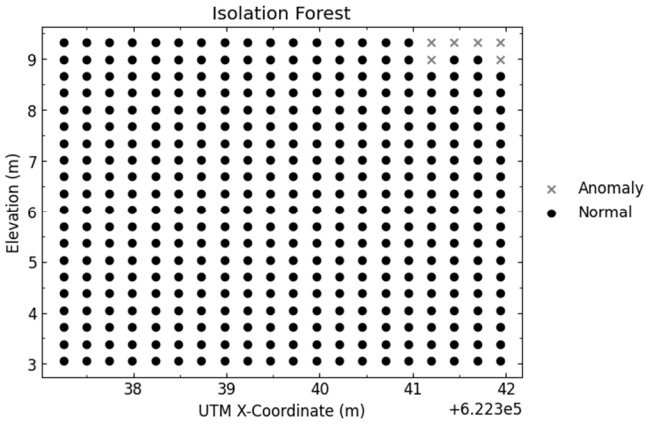


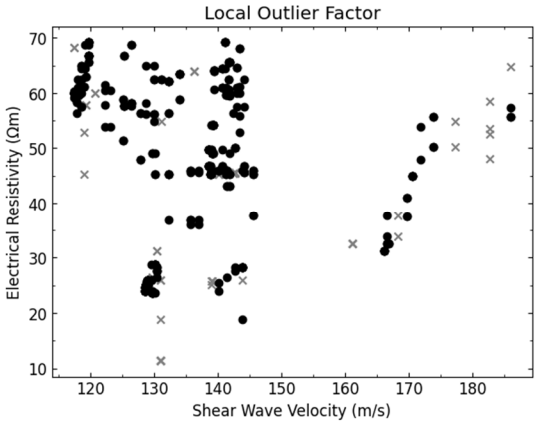


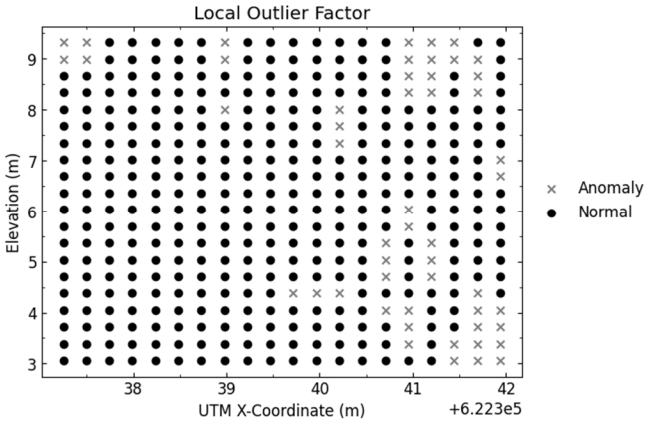


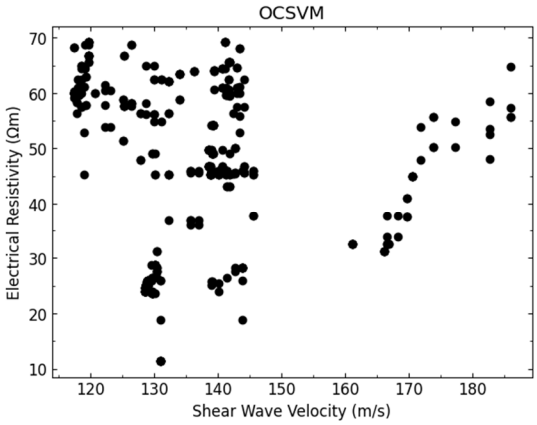
Profile: fwt_site3

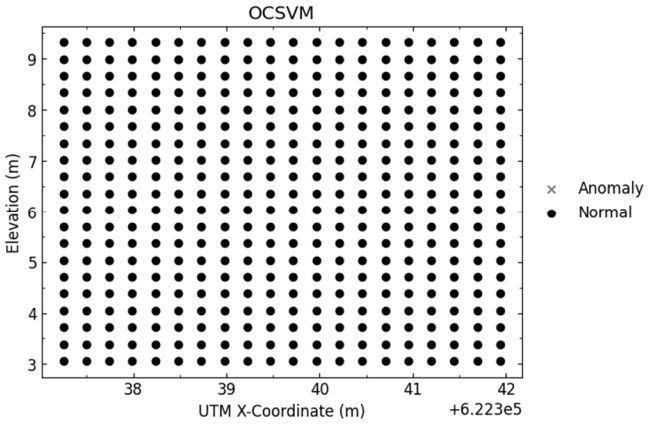






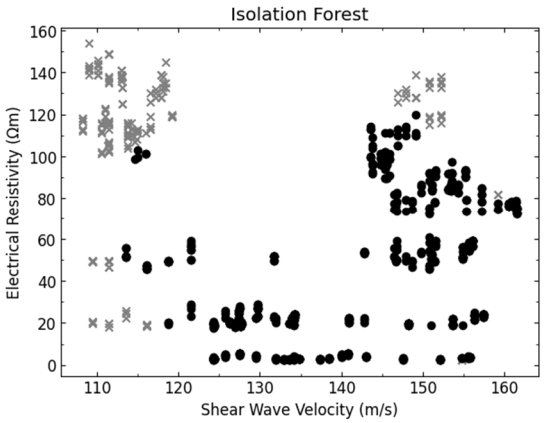


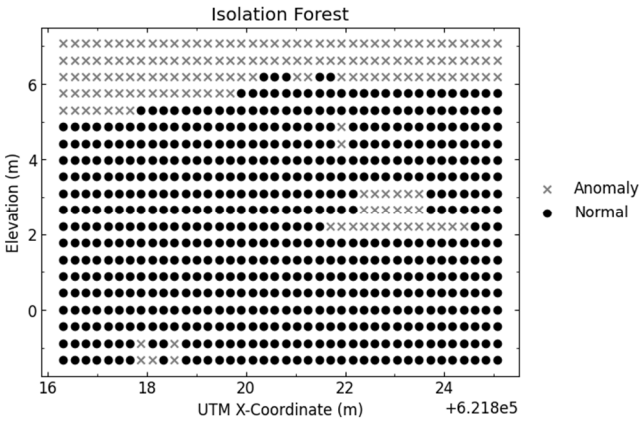


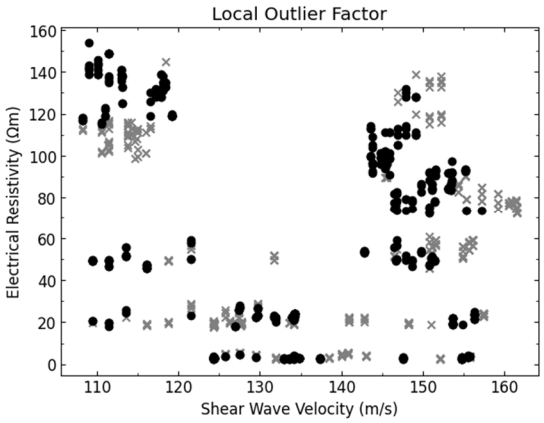


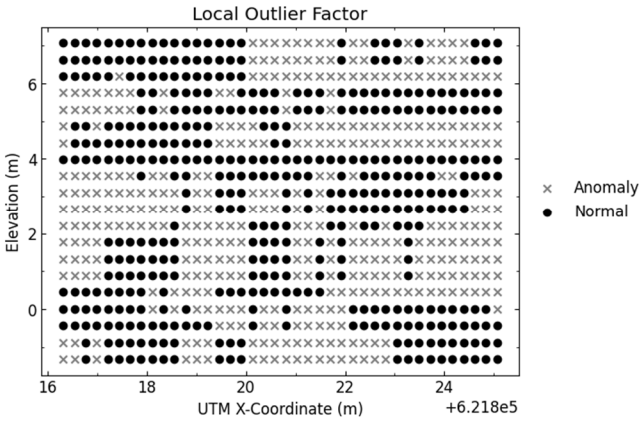
Fit: Wet Season Compilation Dataframe
Predicted: Individual Profiles

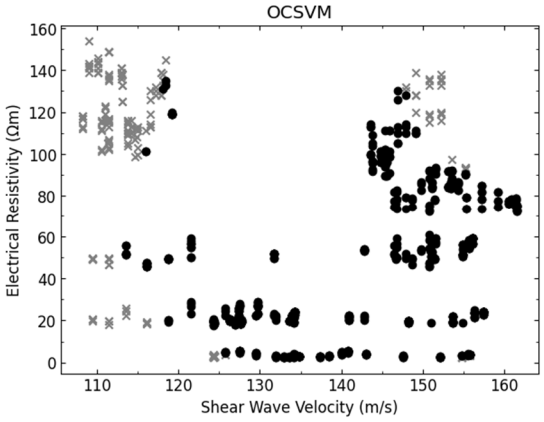
Profile: fc_site2

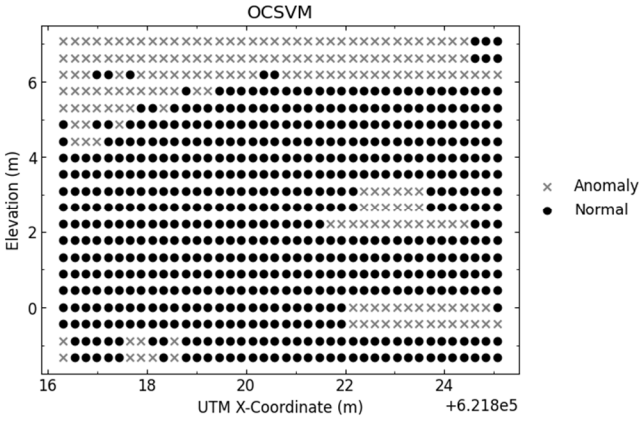




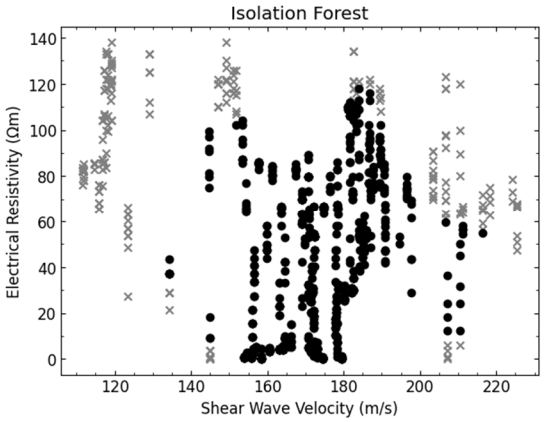


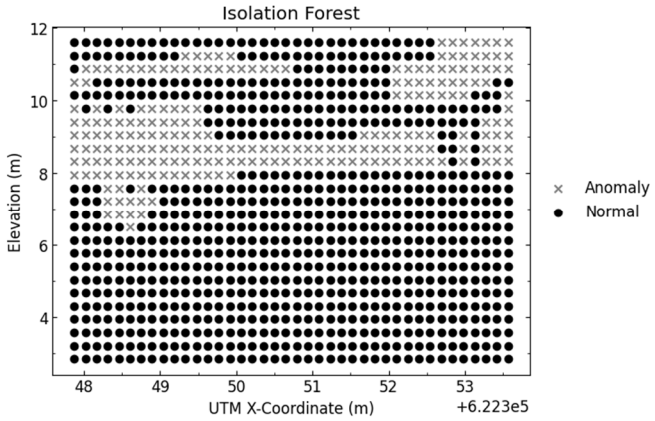


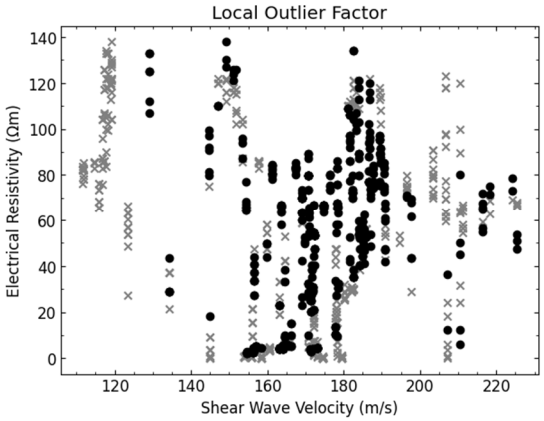


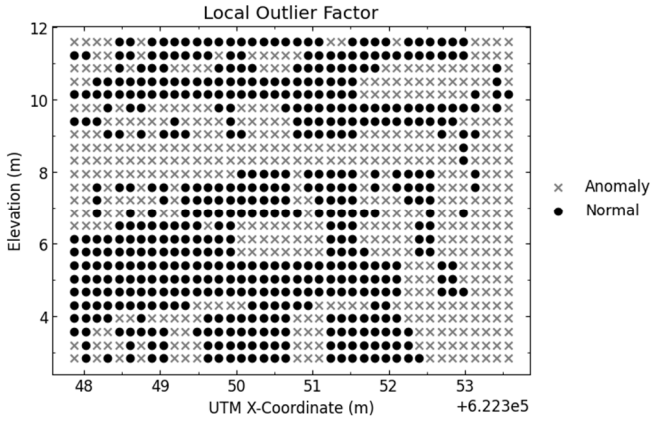


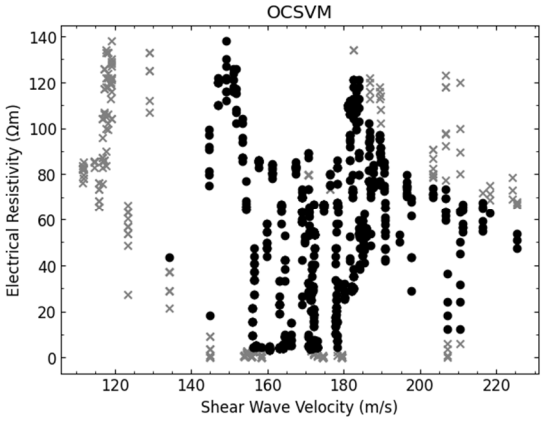
Profile: fc_site3

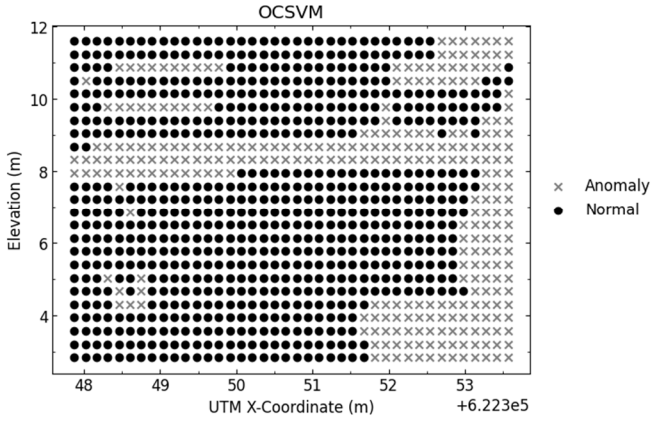




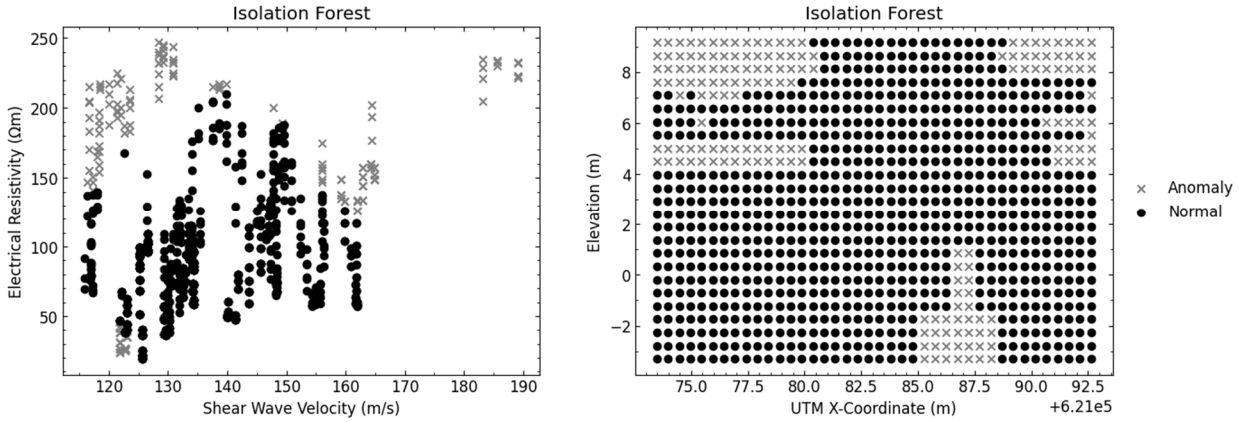


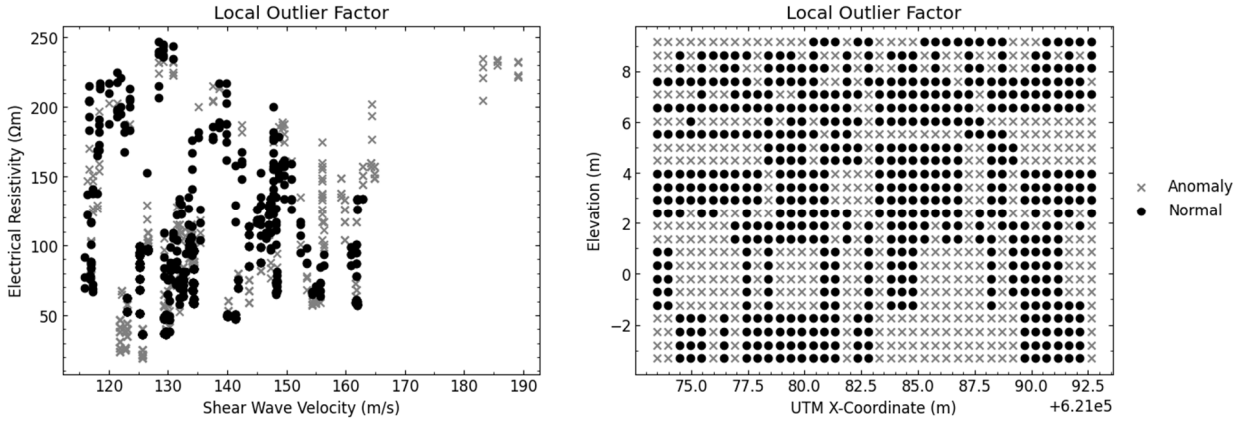


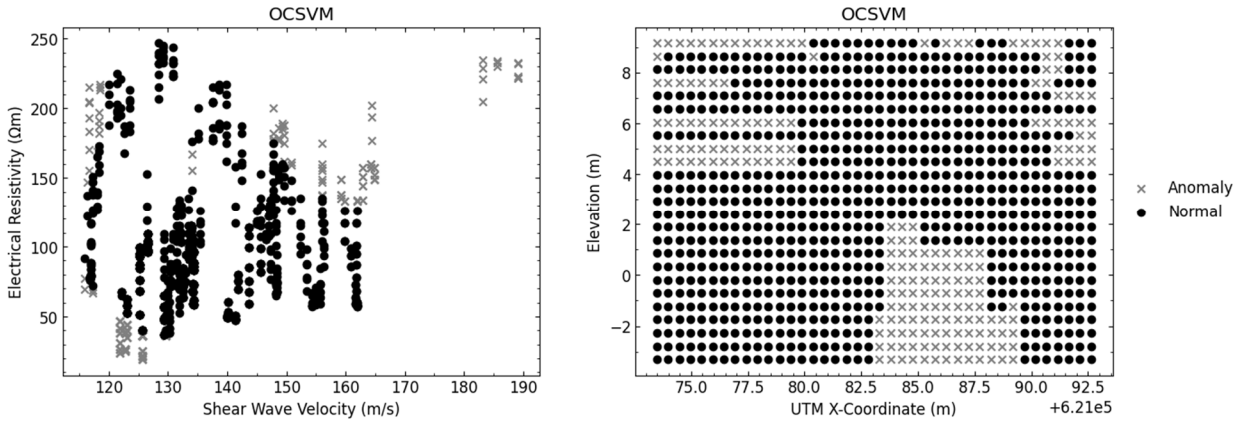




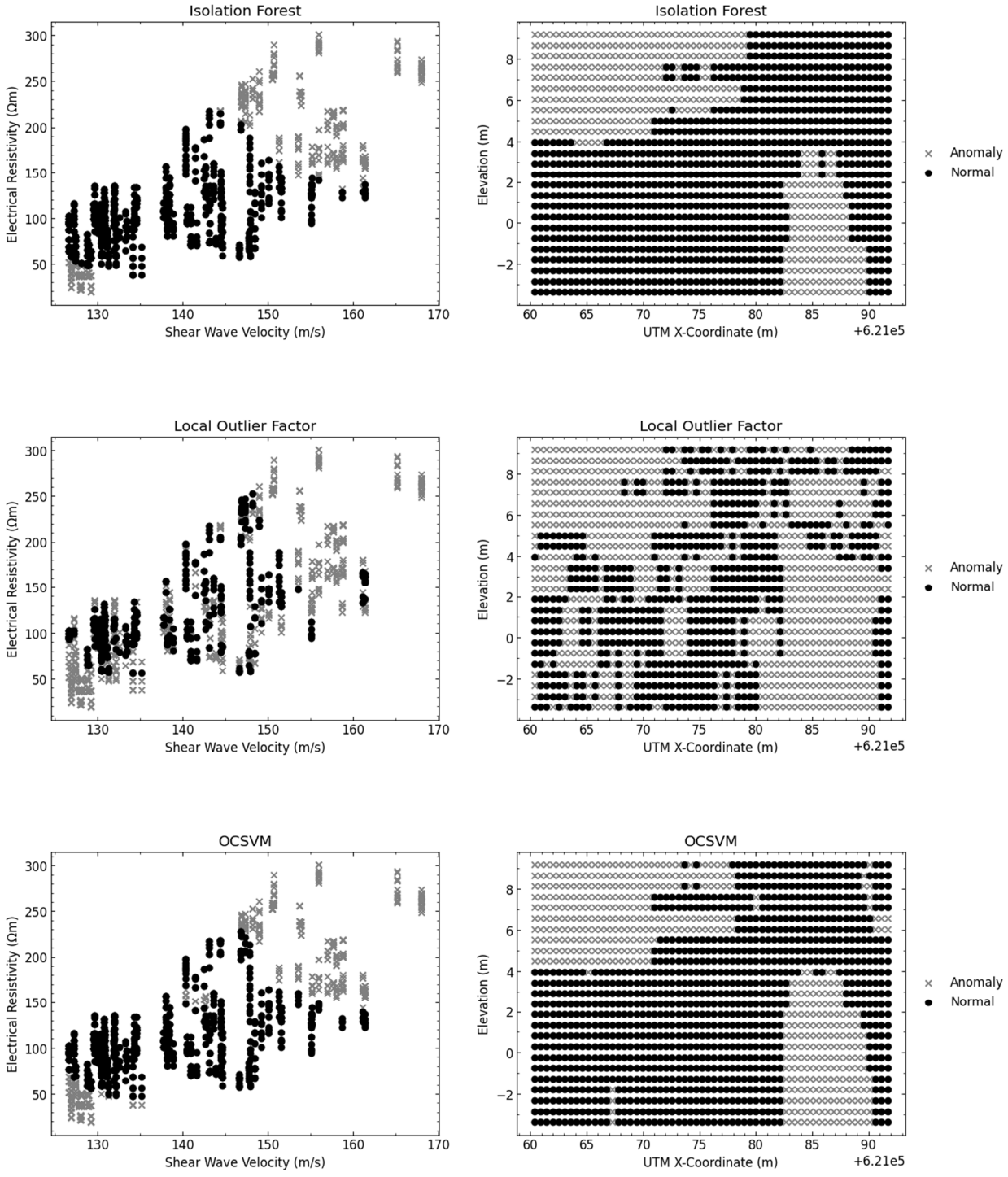
Profile: fc3_site1



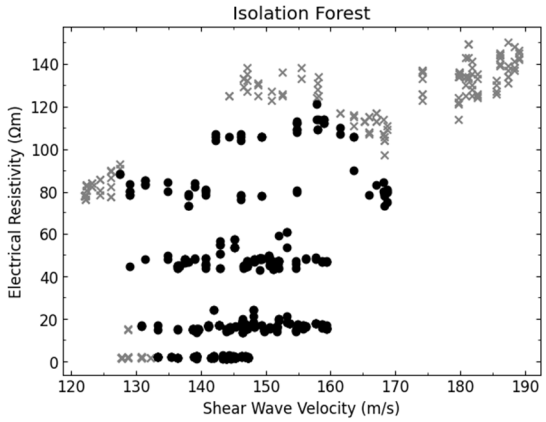


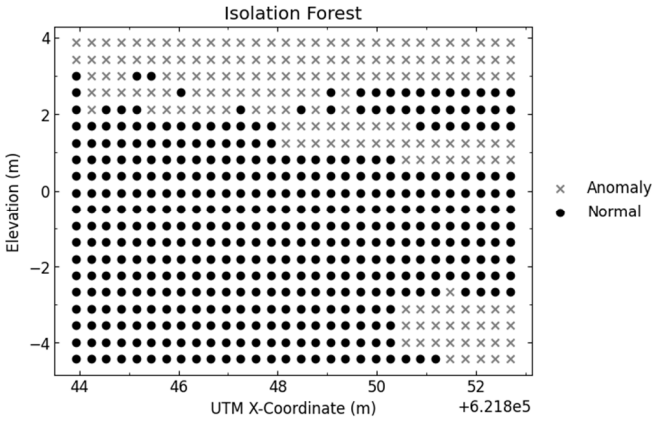


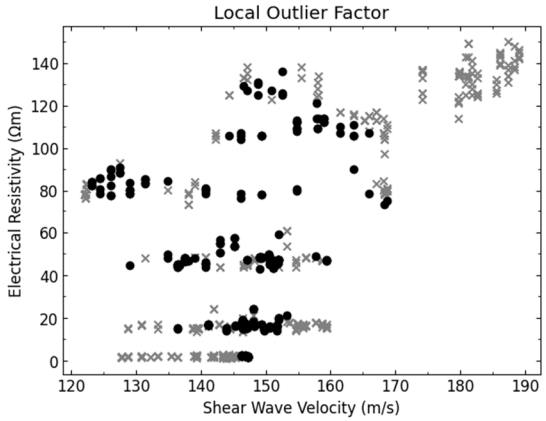
Profile: fc5_site1

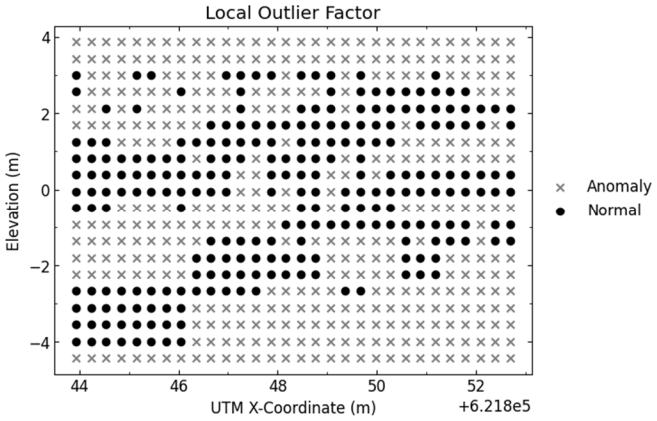


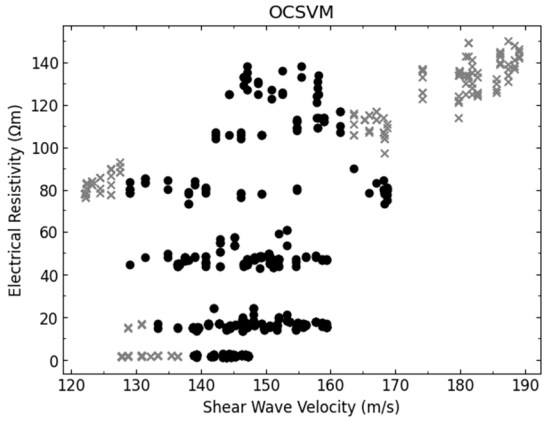
Profile: fdt_site2

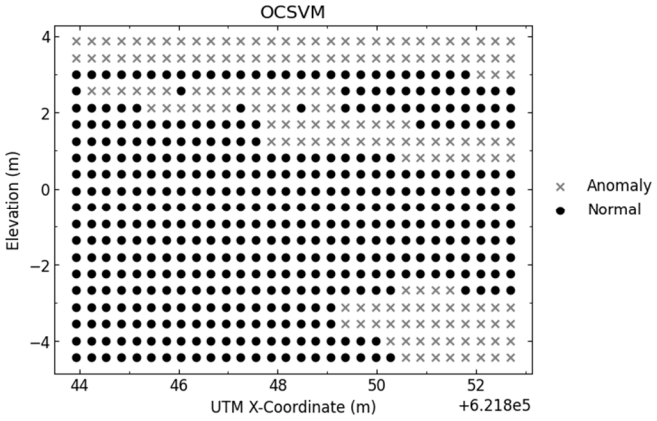




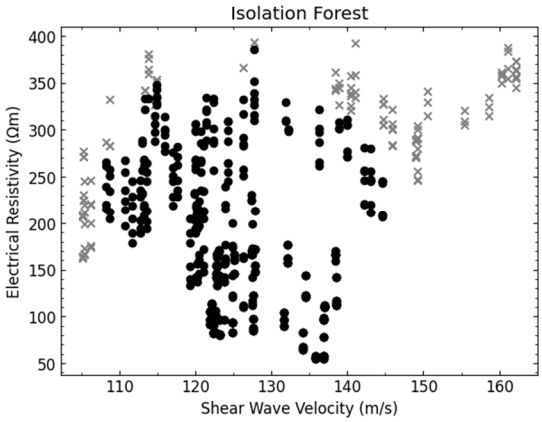


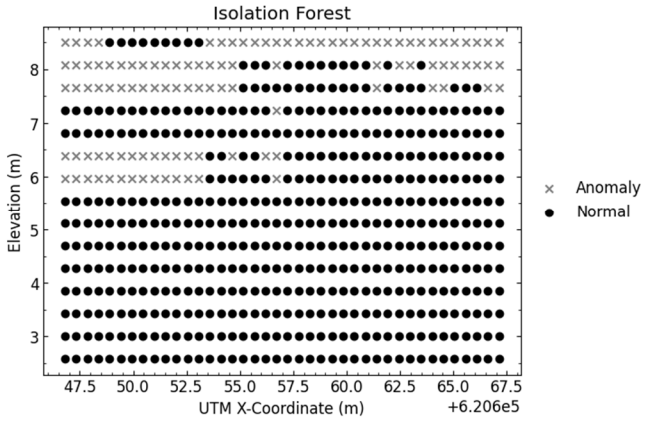


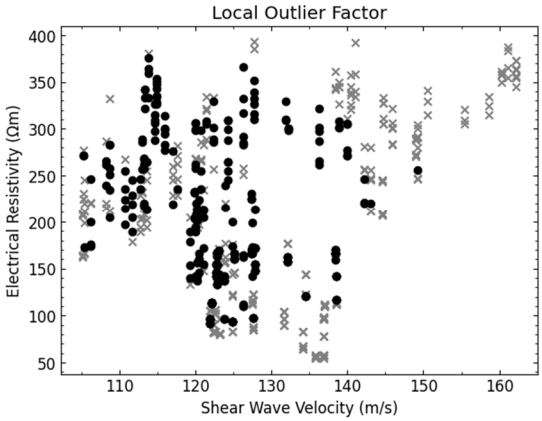


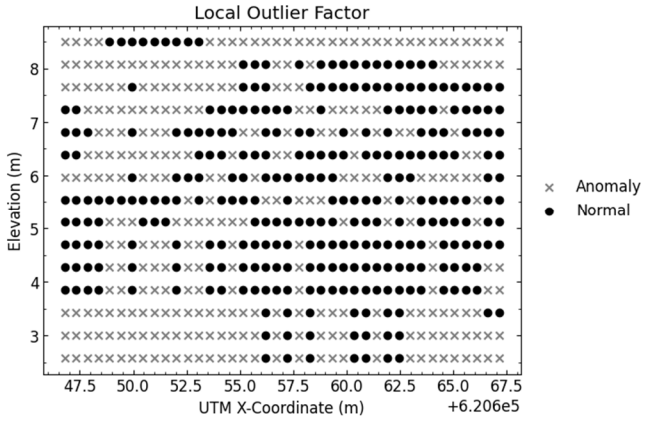


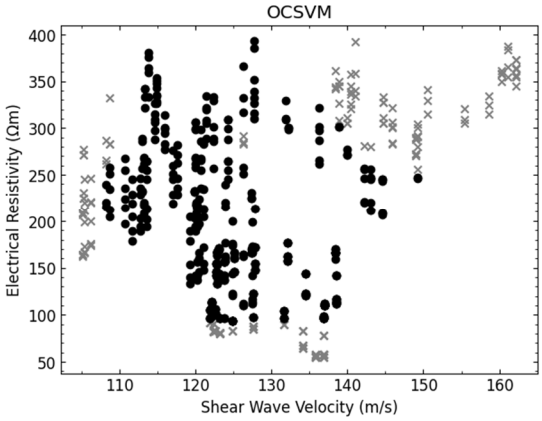
Profile: fwt_site1

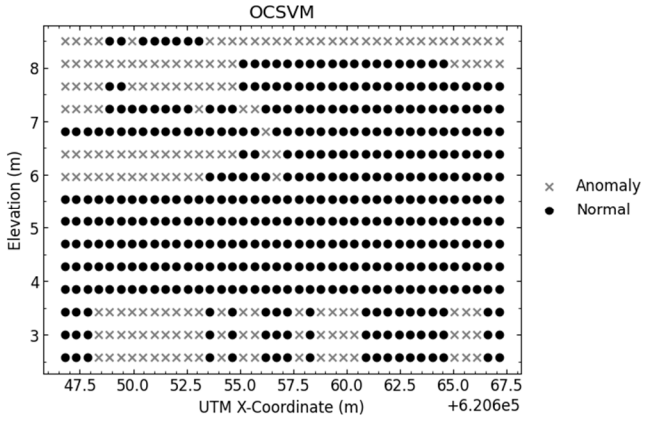




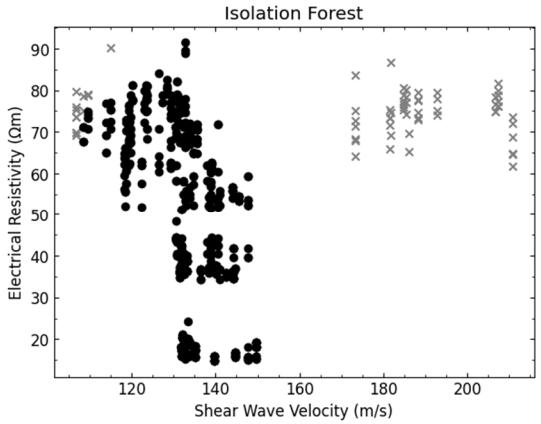


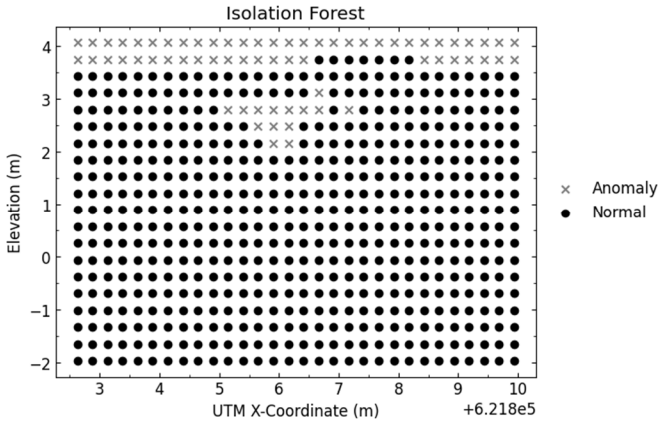


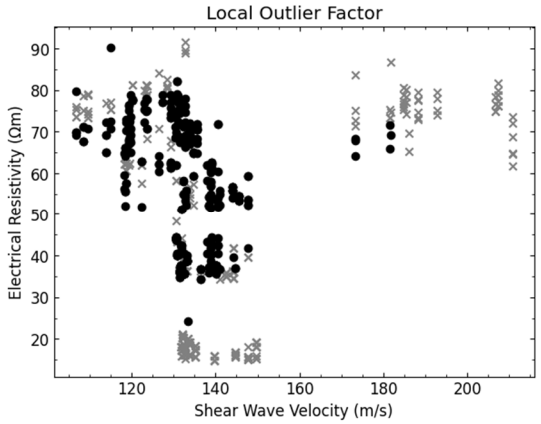


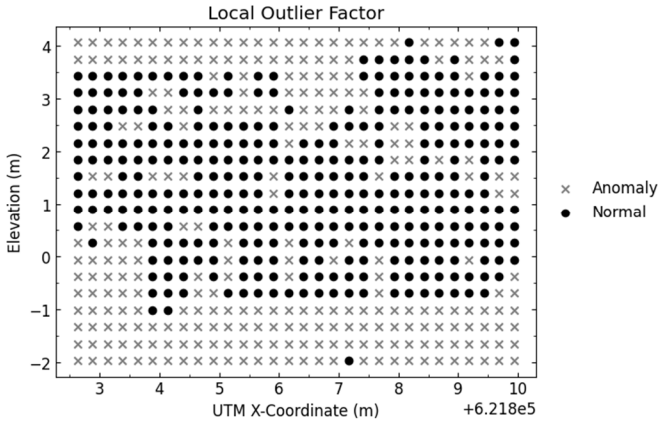


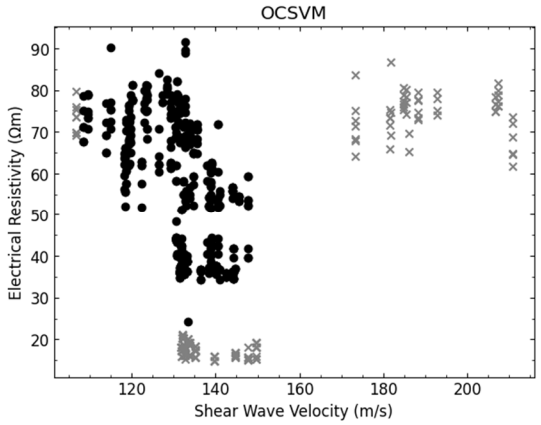
Profile: fwt_site2

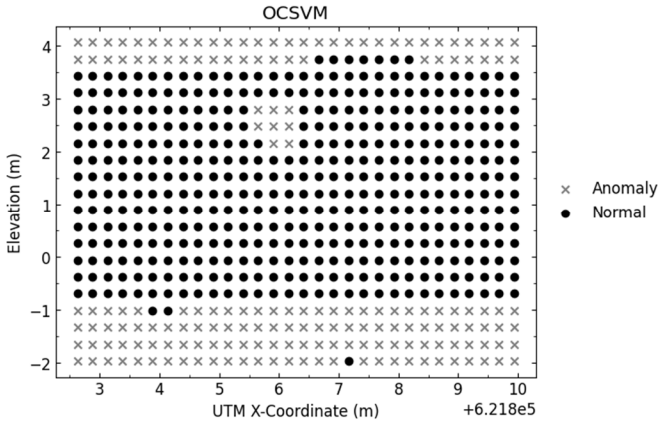




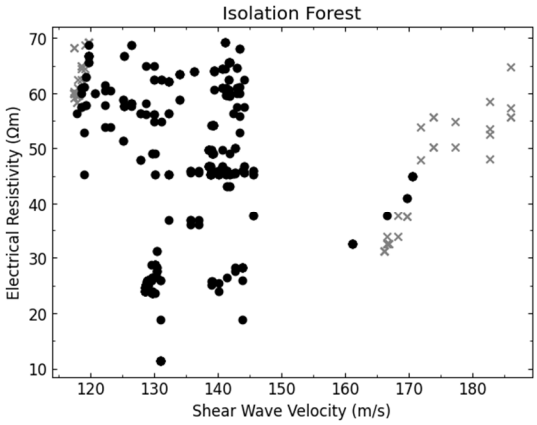


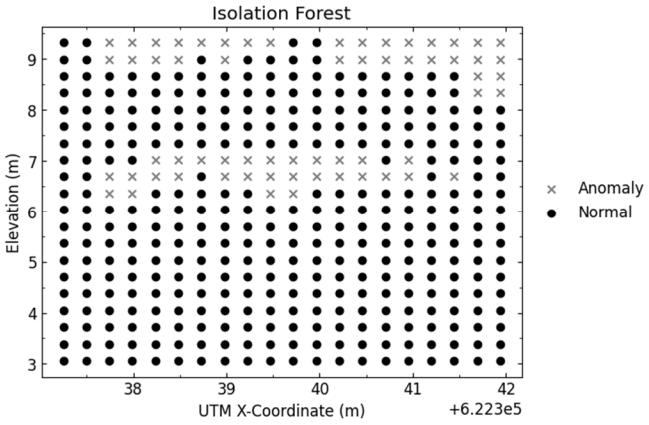


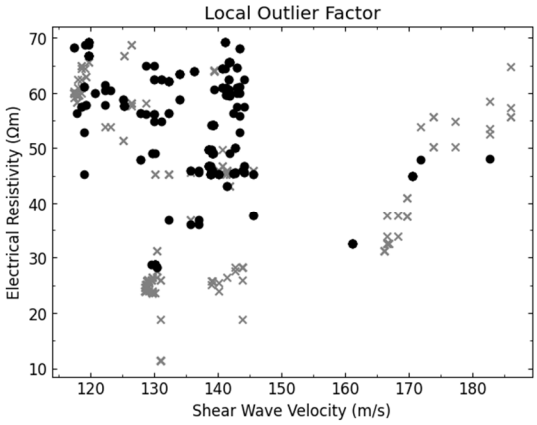


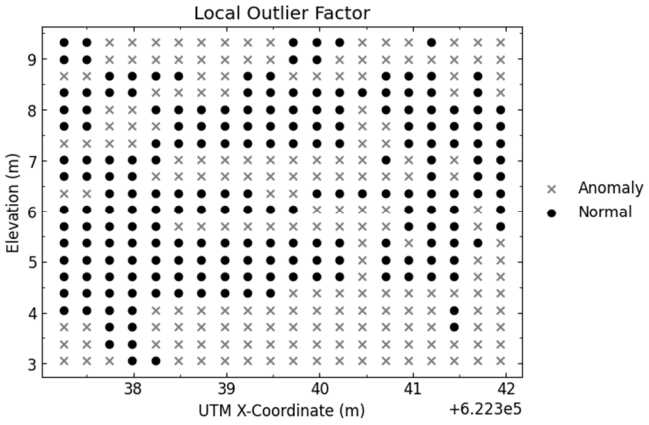


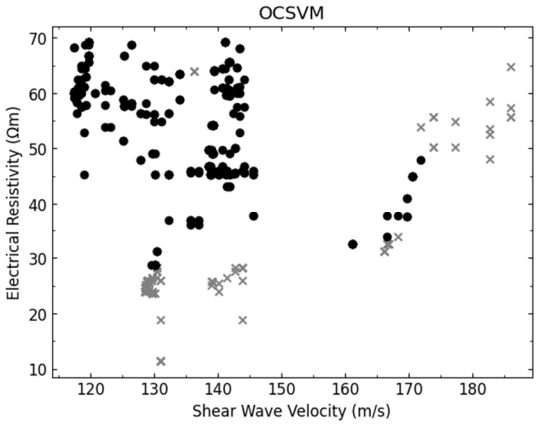
Profile: fwt_site3

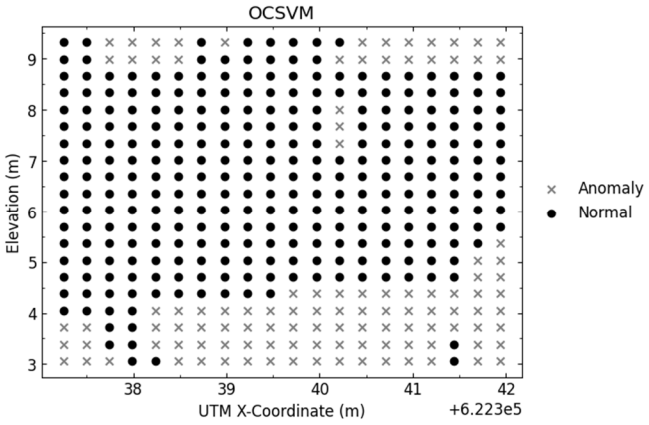






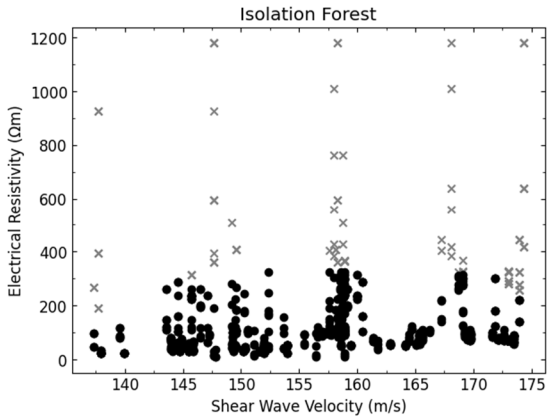


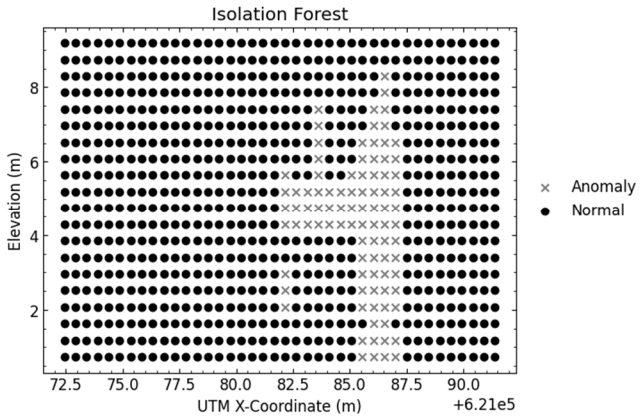


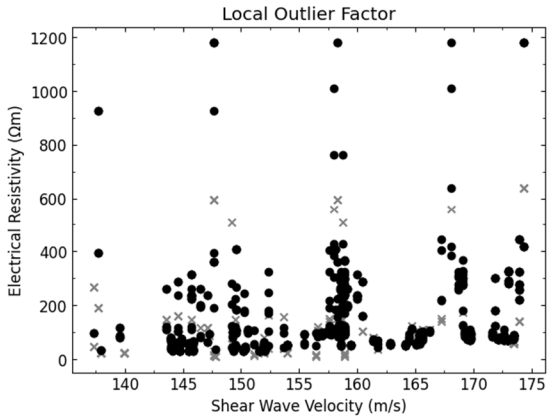


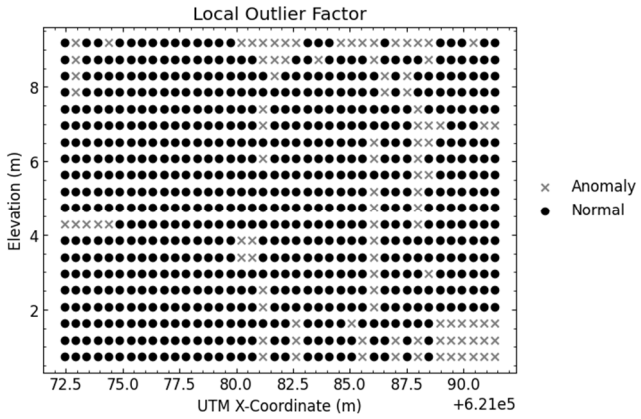
Fit: Dry Season Site 1 Compilation Dataframe
Predicted: Dry Season Site 1 Compilation Dataframe

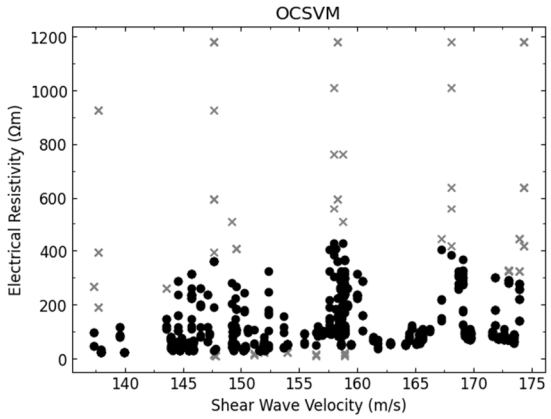
Profile: sc3_site1

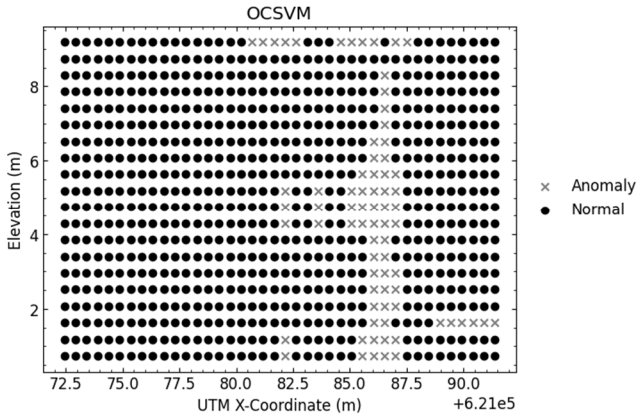




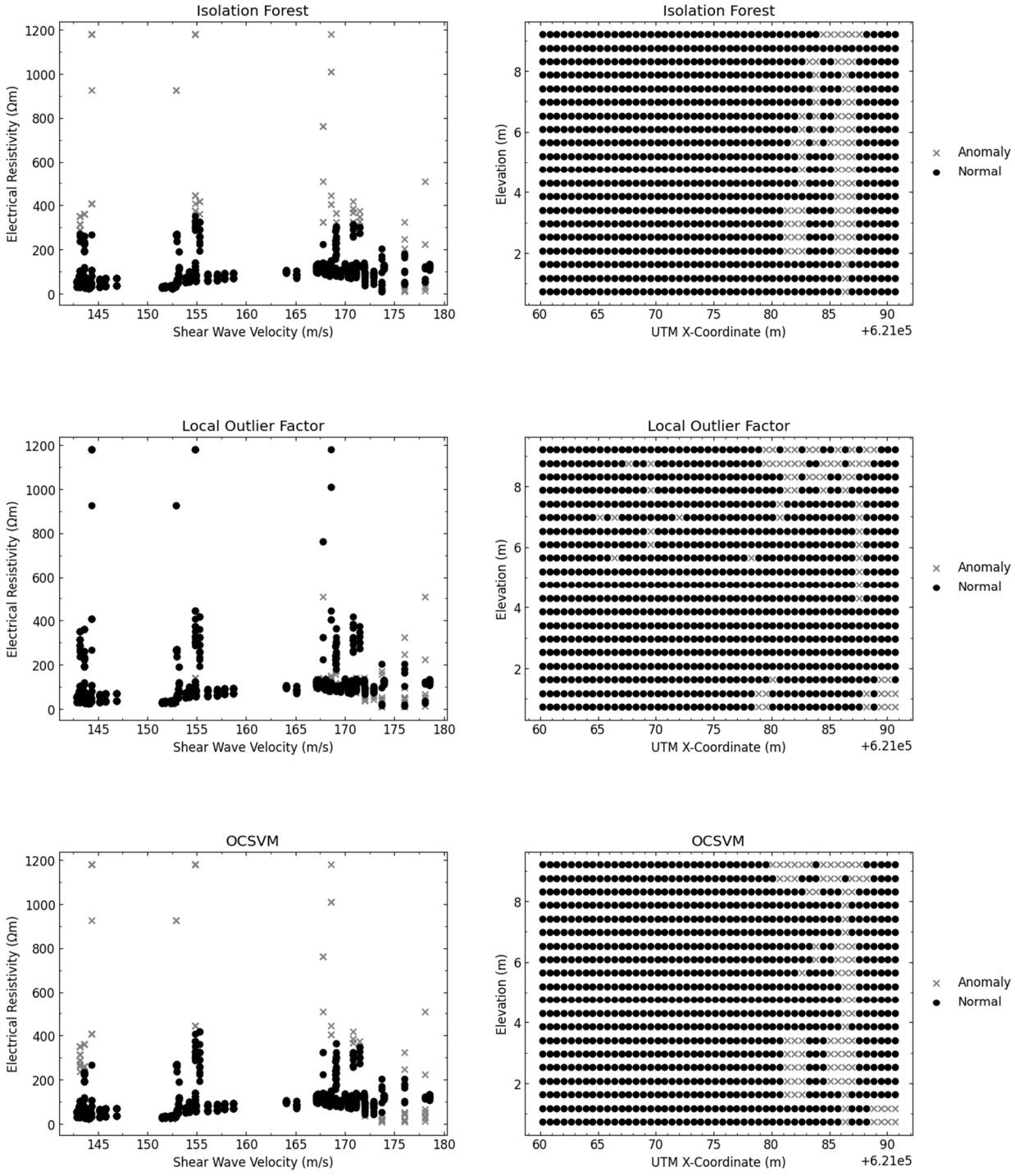




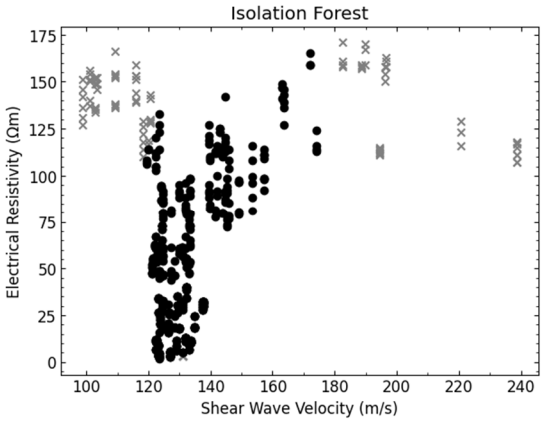


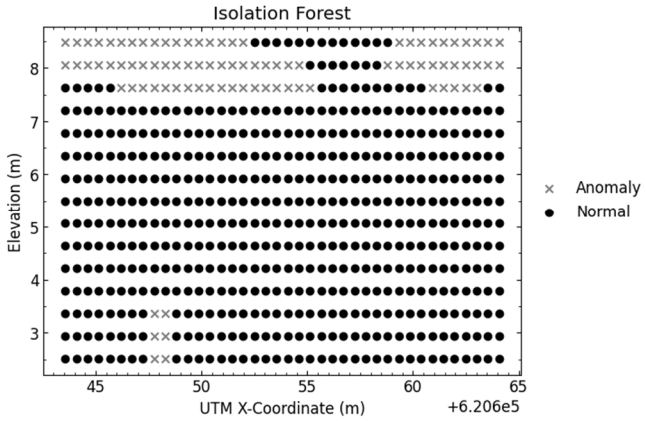


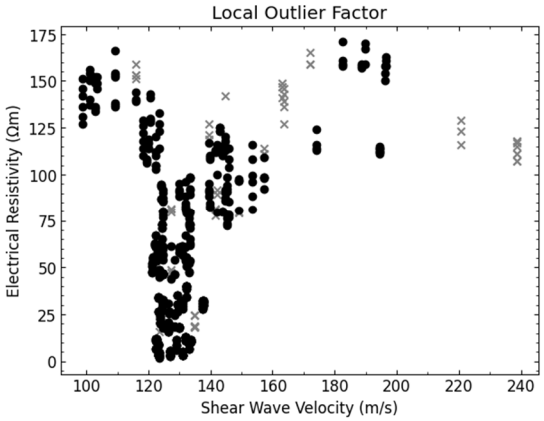
Profile: sc5_site1

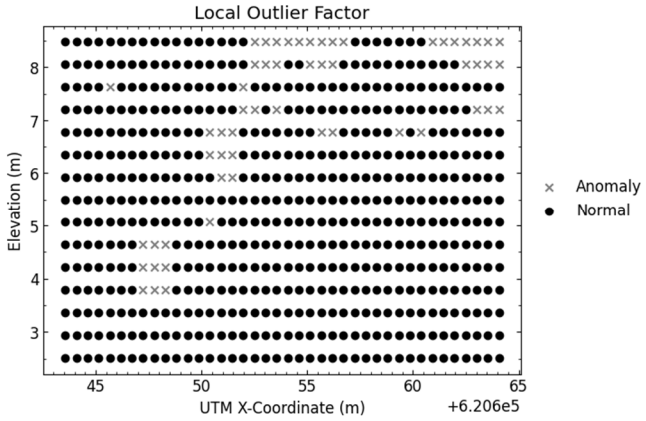


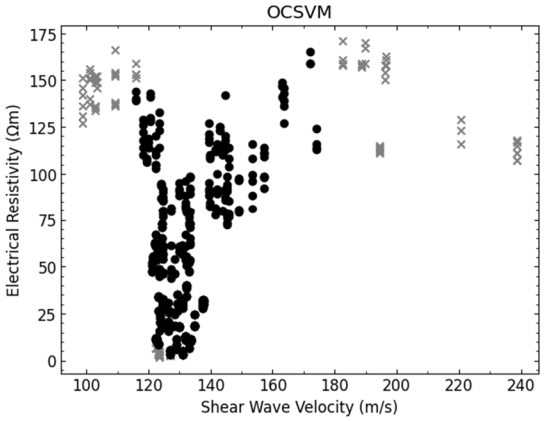
Profile: swt_site1

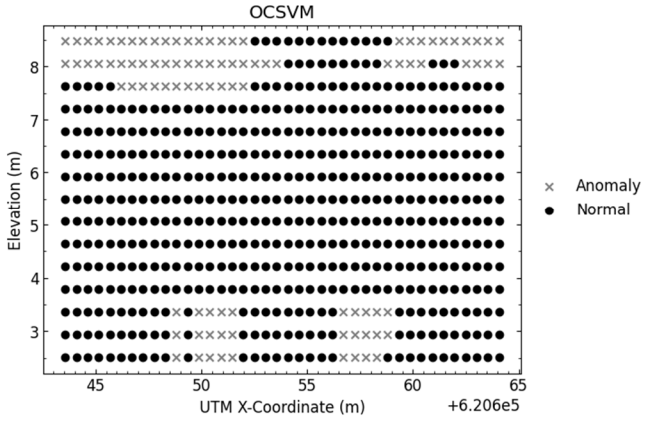






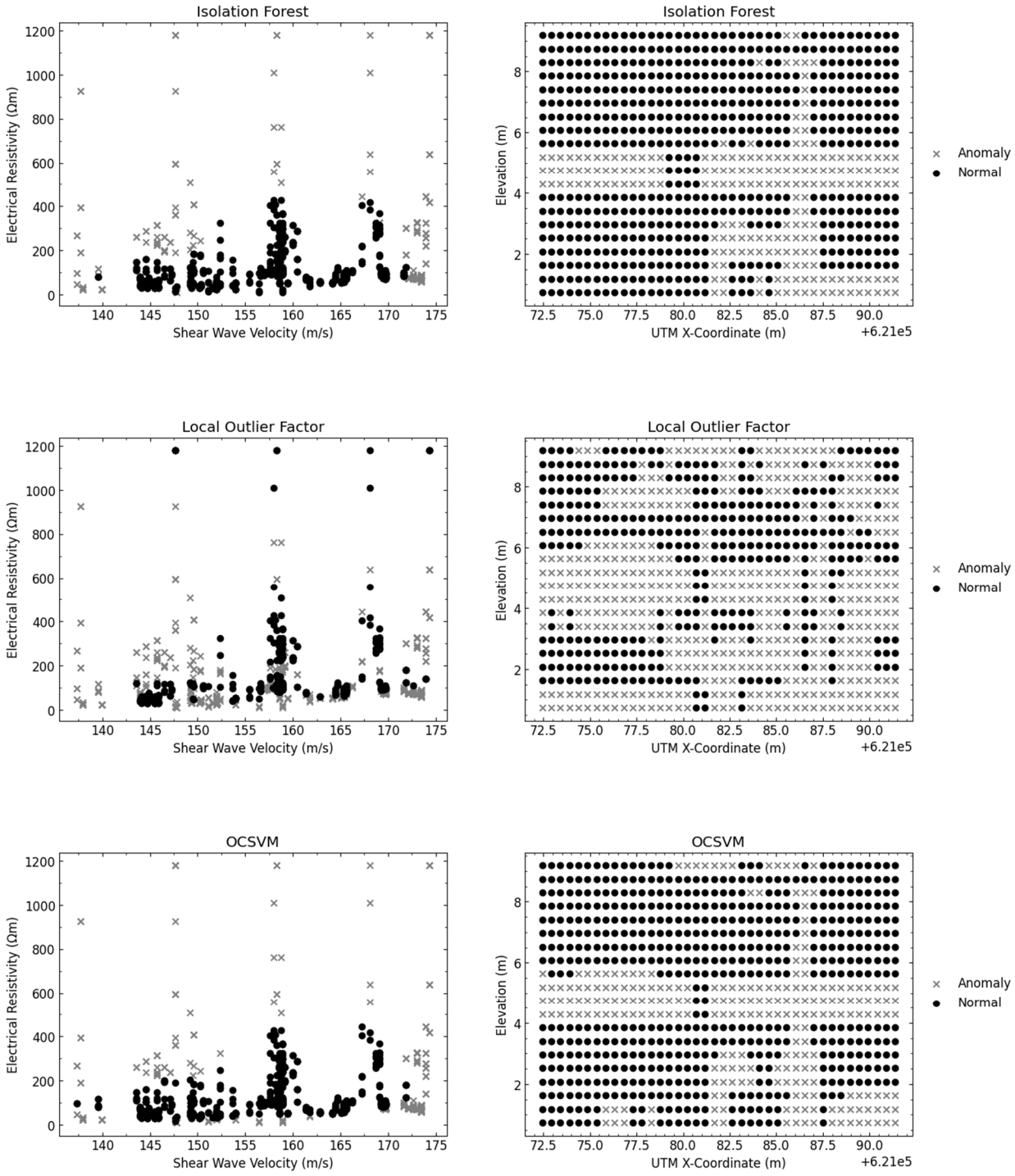




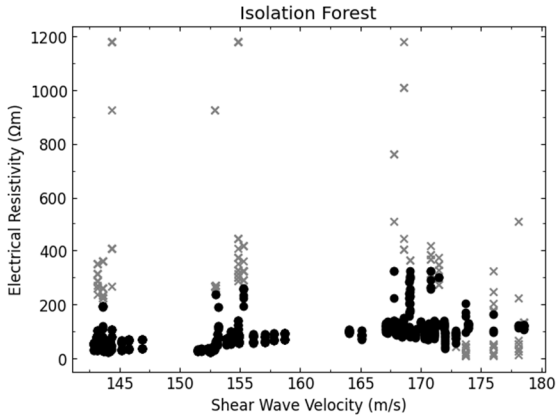


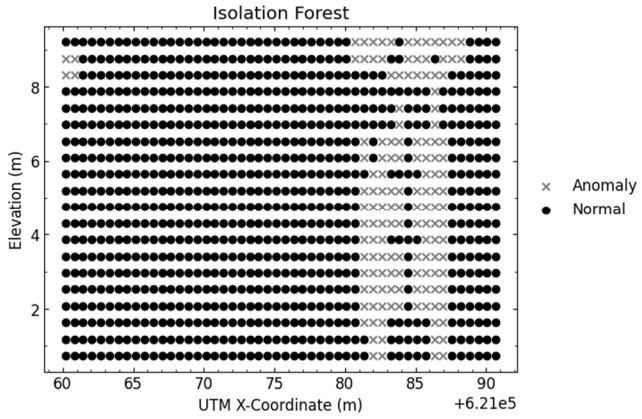
Fit: Dry Season Site 1 Compilation Dataframe
Predicted: Individual Profiles

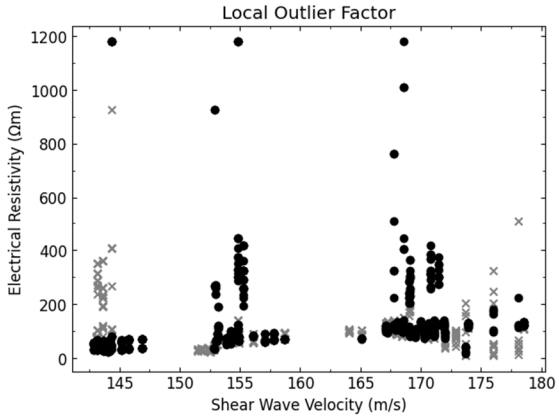
Profile: sc3_site1

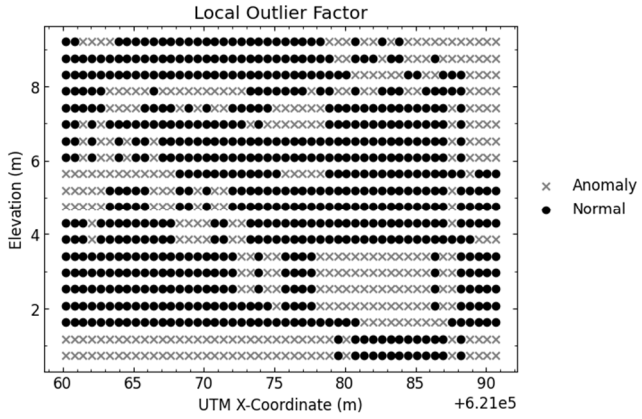


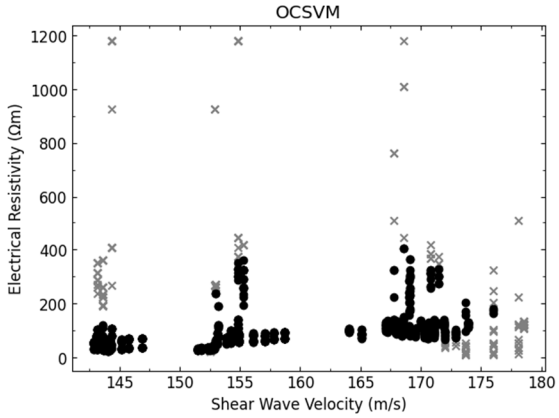
Profile: sc5_site1

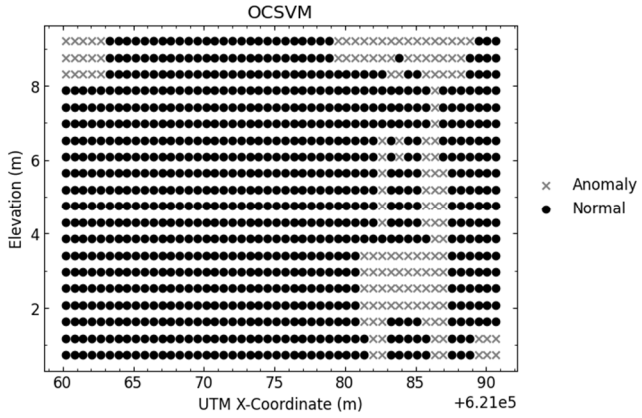




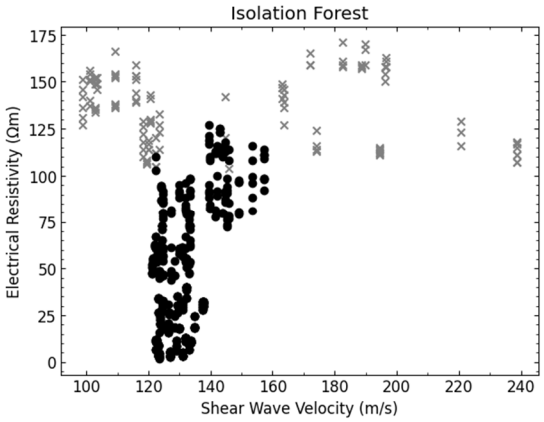


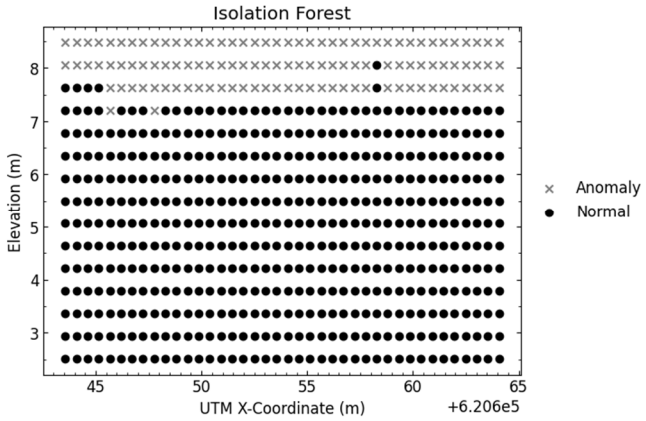


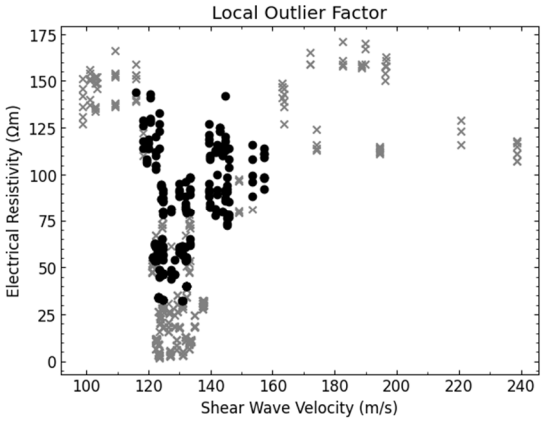


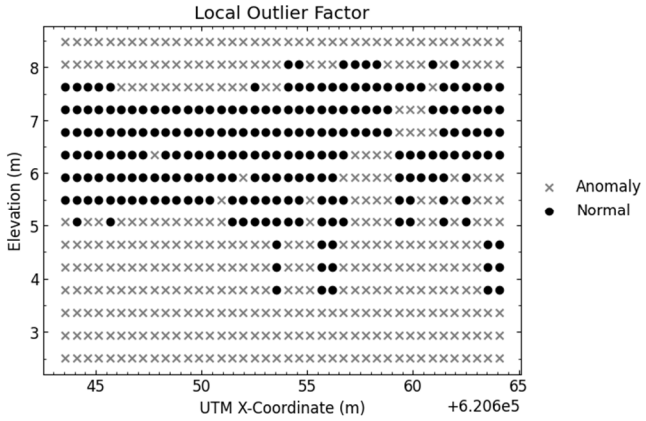


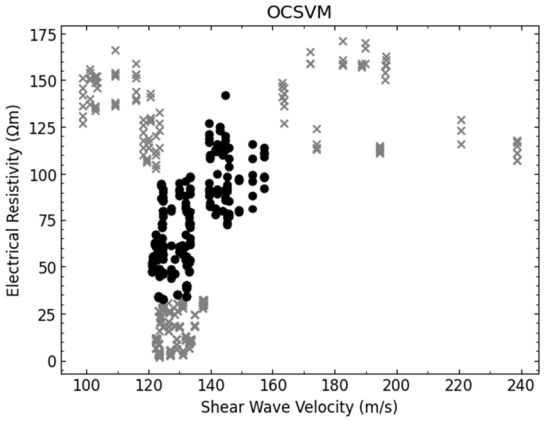
Profile: swt_site1

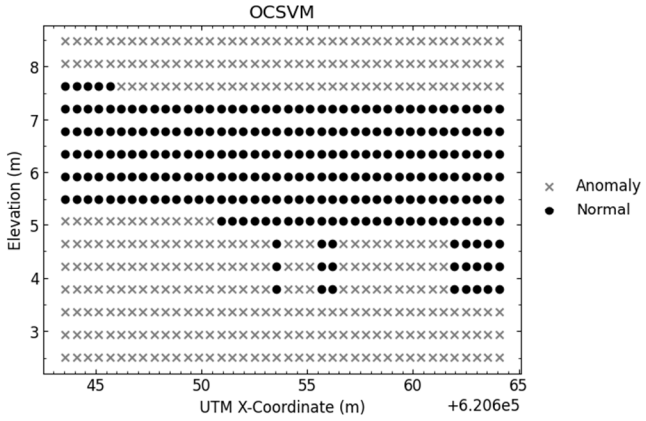






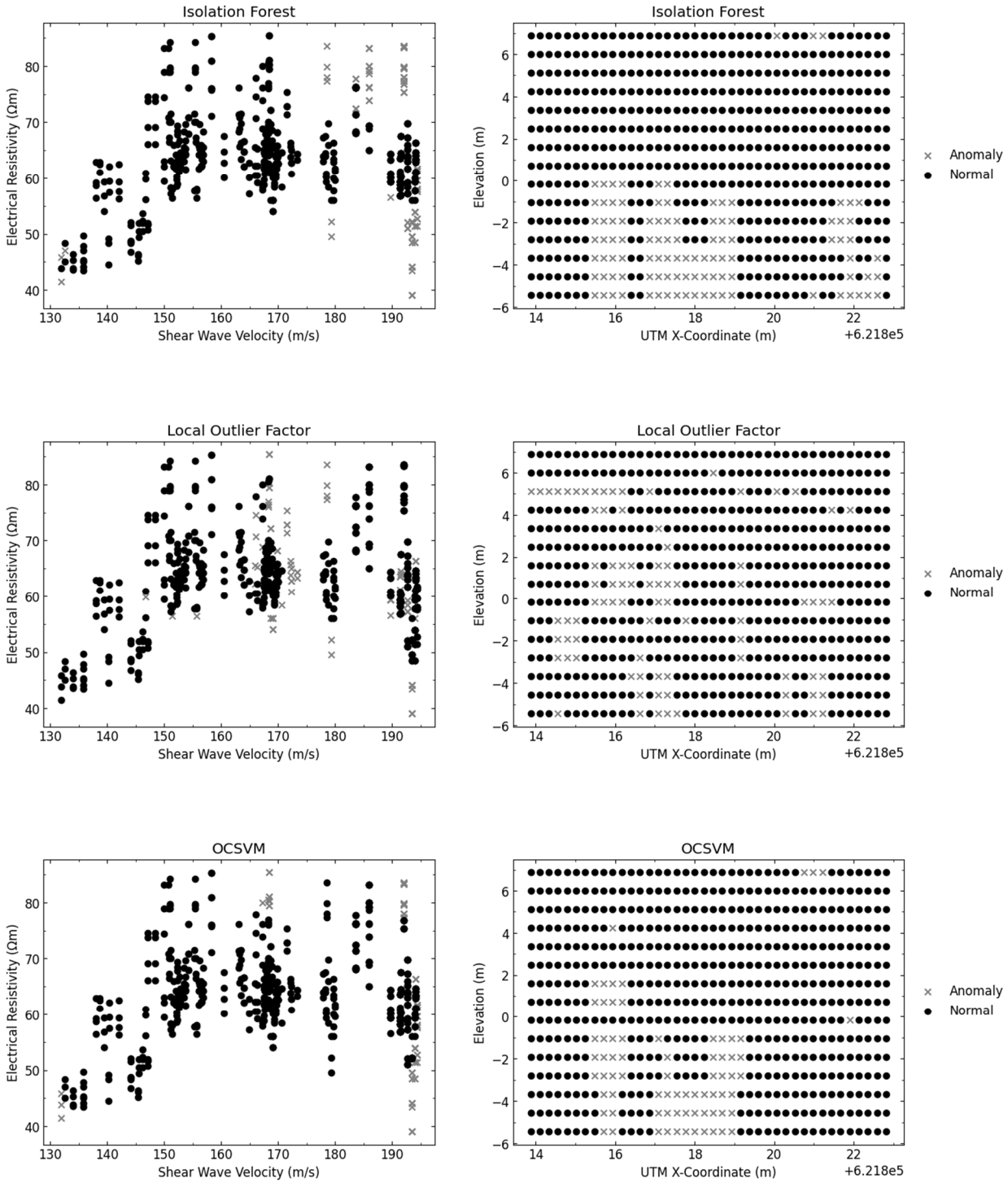




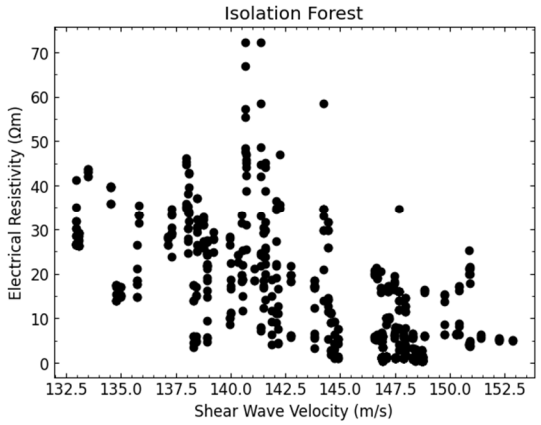


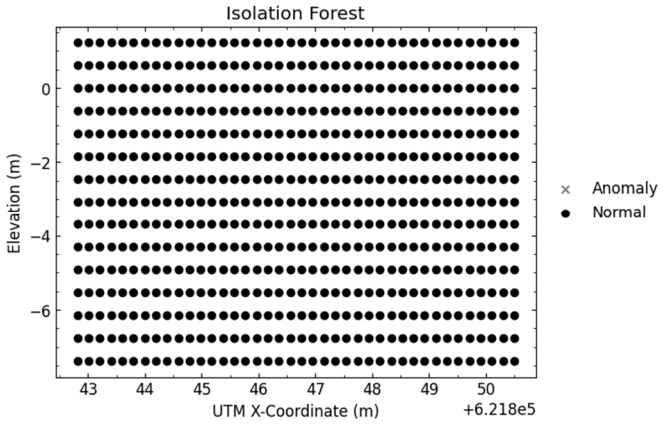
Fit: Dry Season Site 2 Compilation Dataframe
Predicted: Dry Season Site 2 Compilation Dataframe

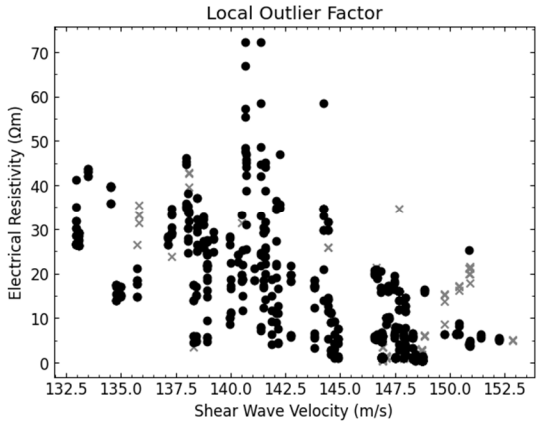
Profile: sc_site2

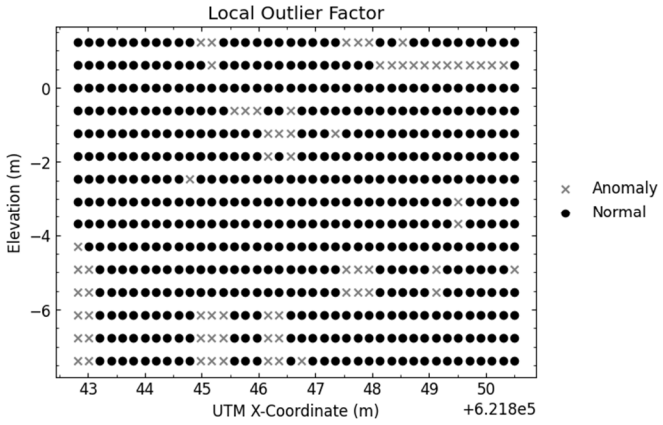


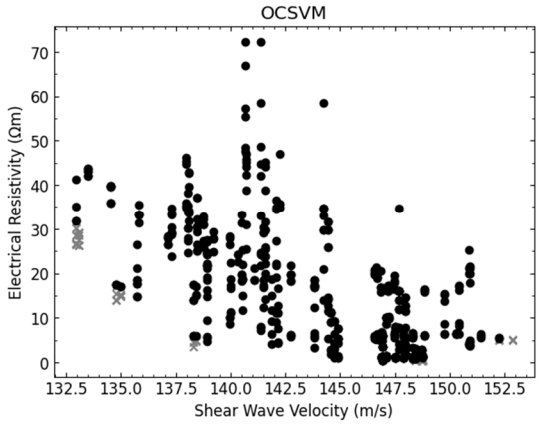
Profile: sdt_site2

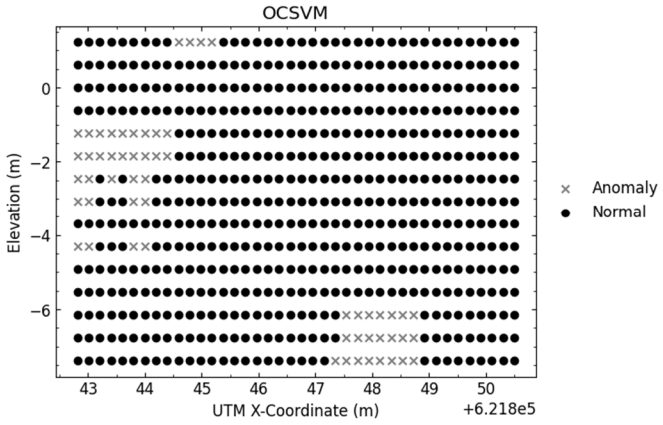




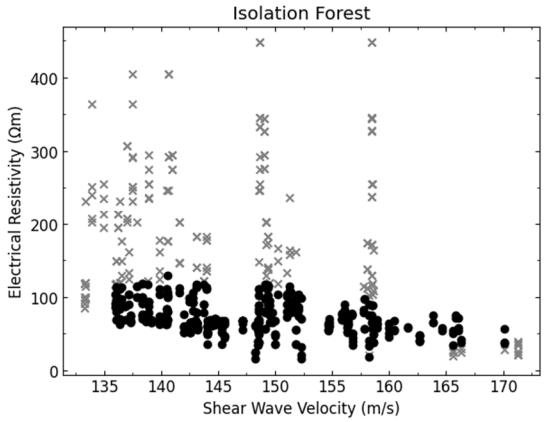


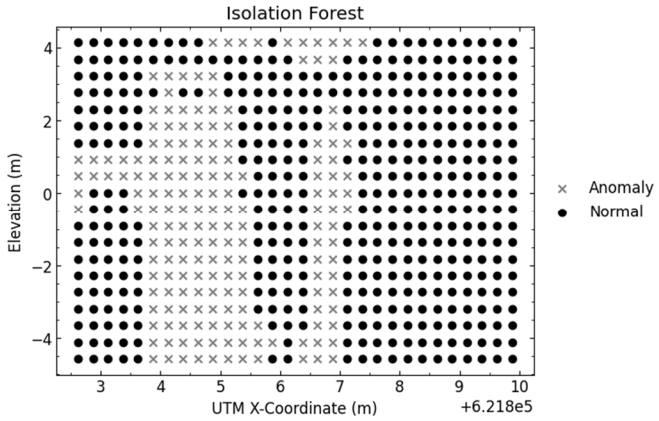


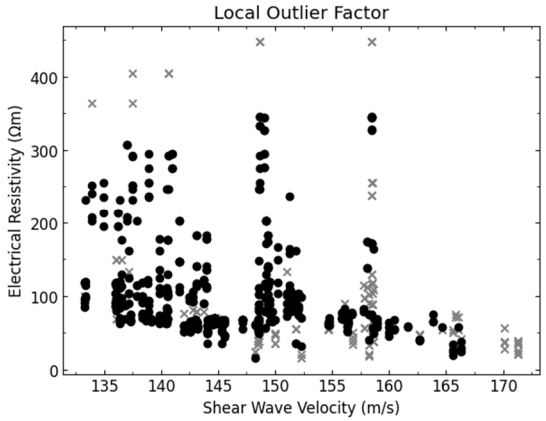


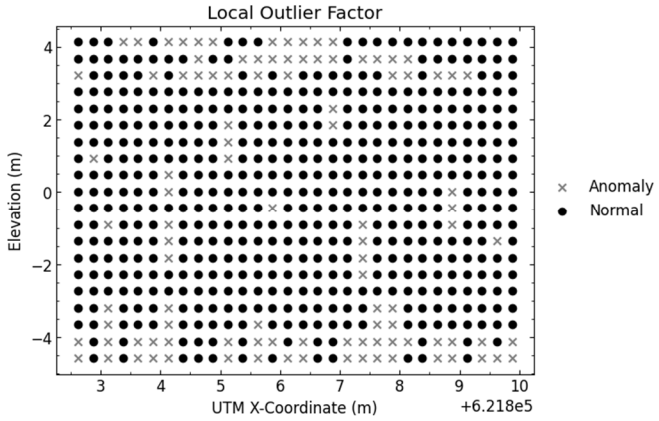


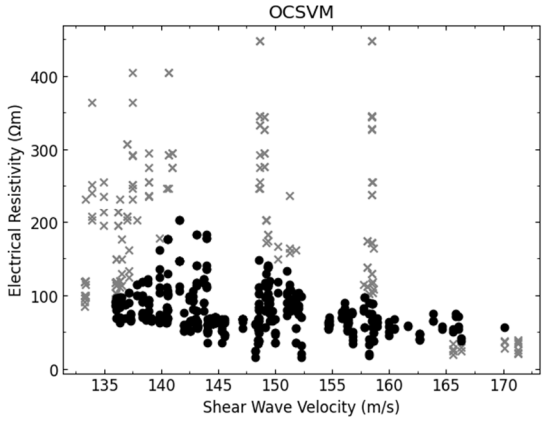
Profile: swt_site2

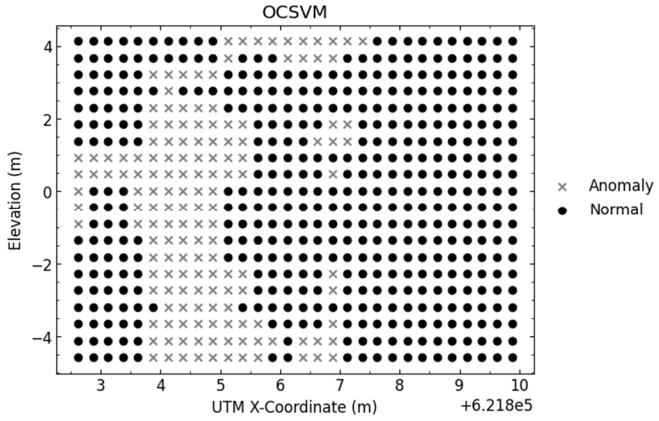






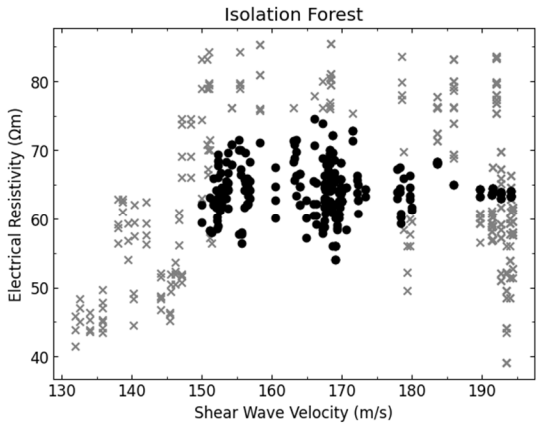


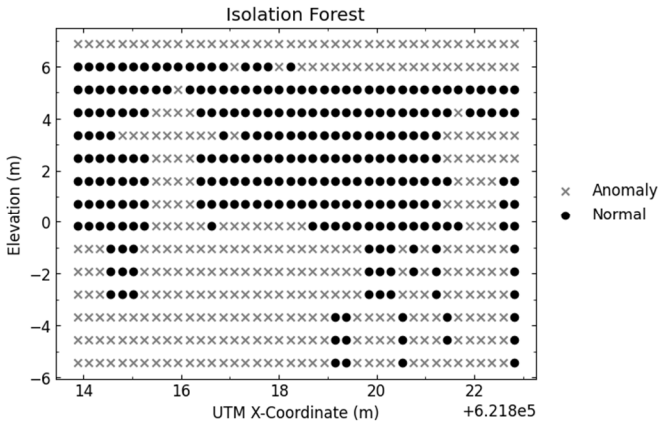


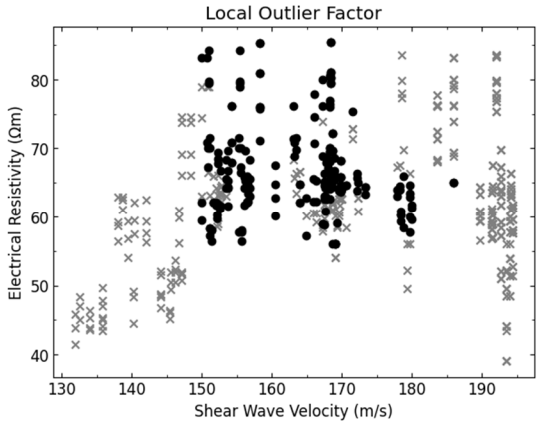


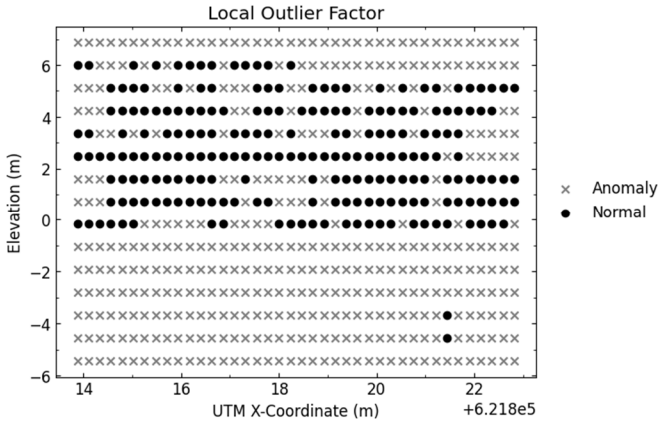
Fit: Dry Season Site 2 Compilation Dataframe
Predicted: Individual Profiles

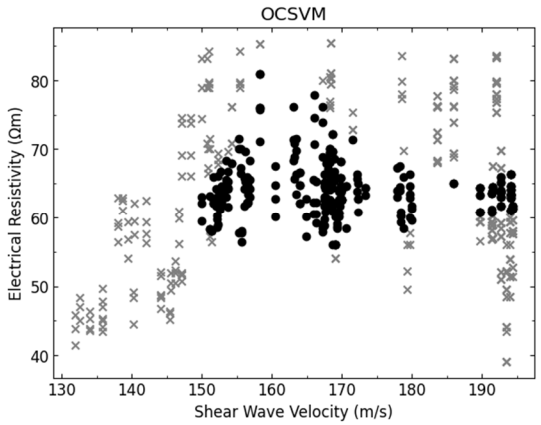
Profile: sc_site2

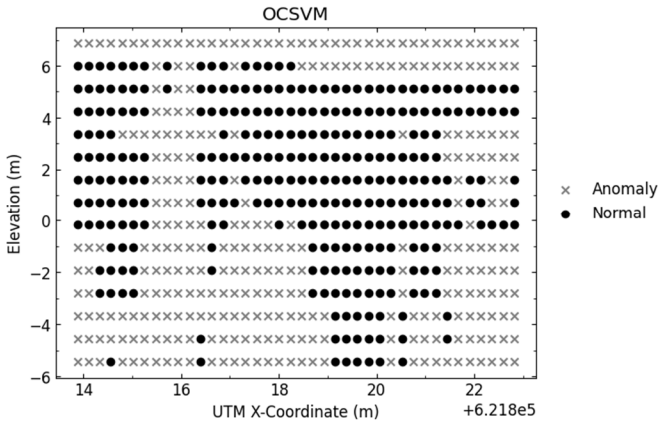




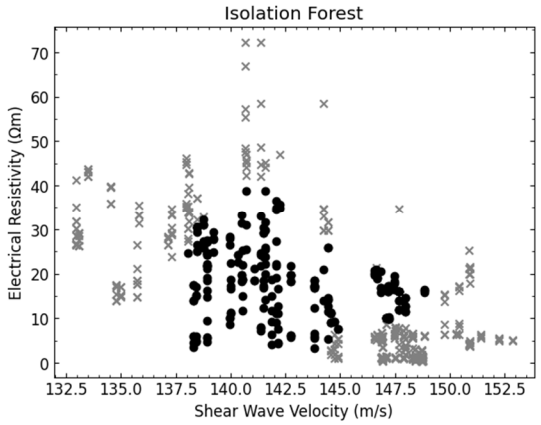


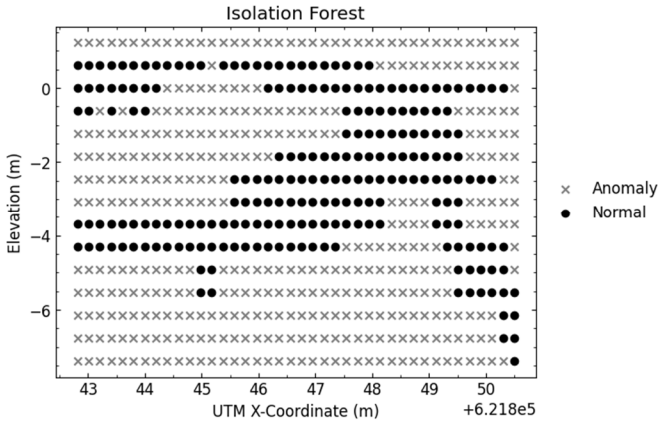


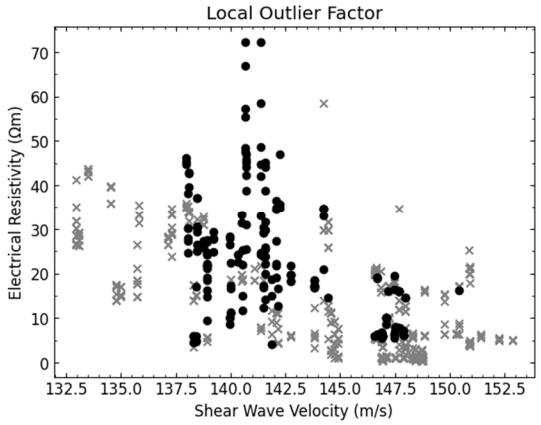


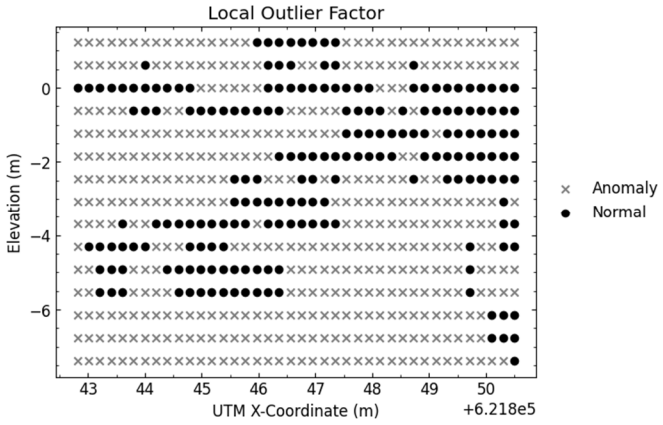


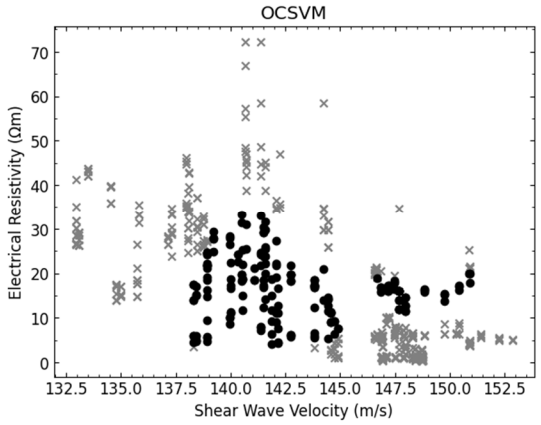
Profile: sdt_site2

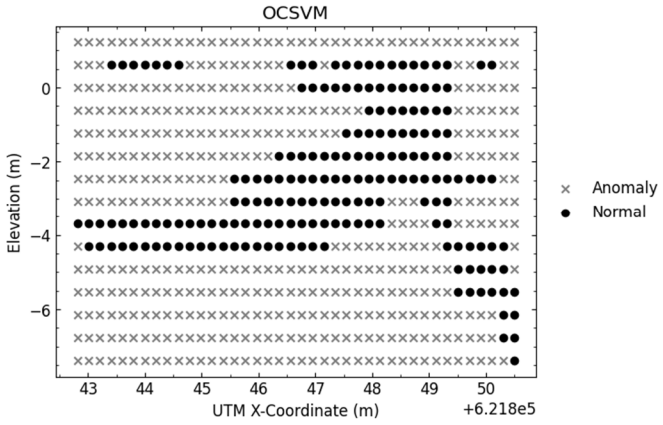




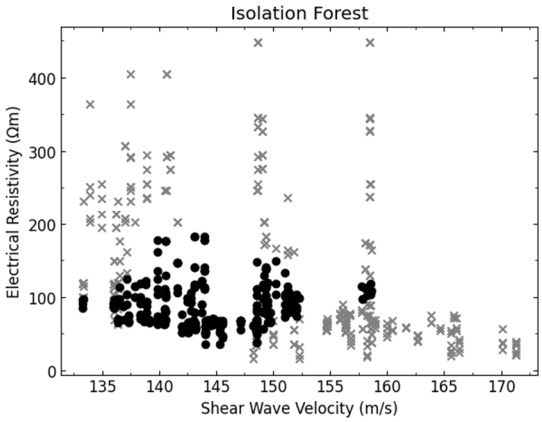


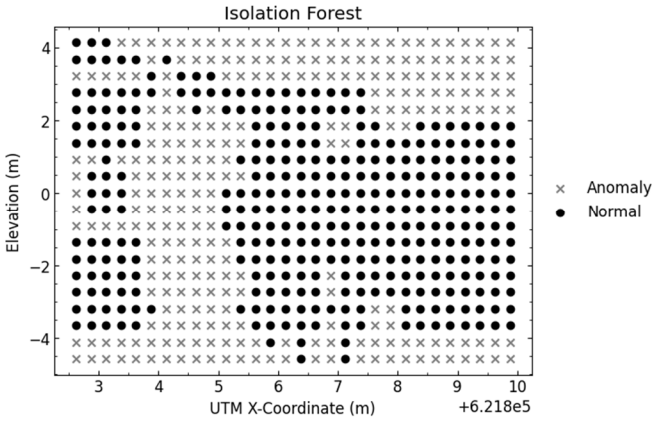


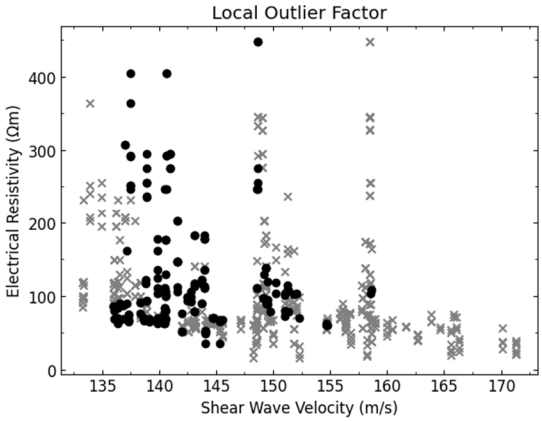


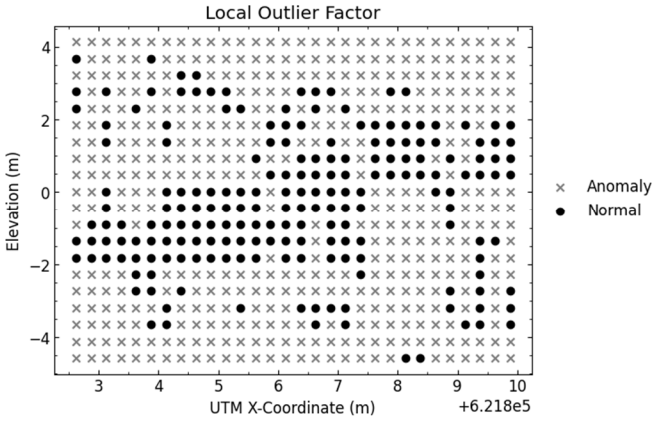


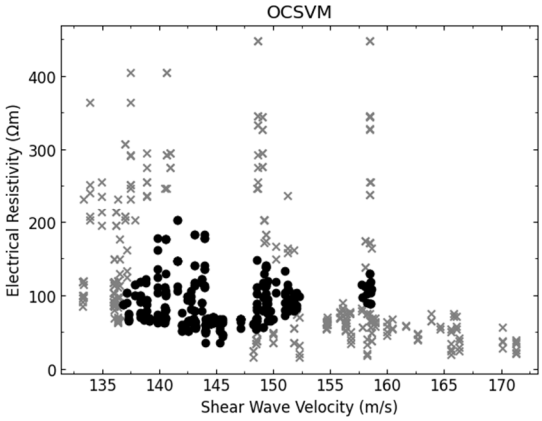
Profile: swt_site2

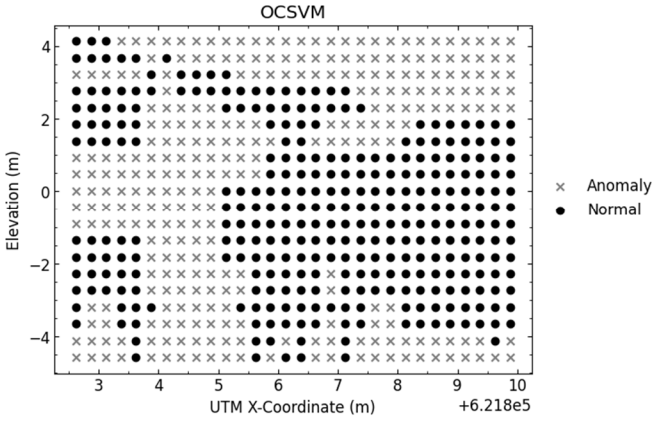






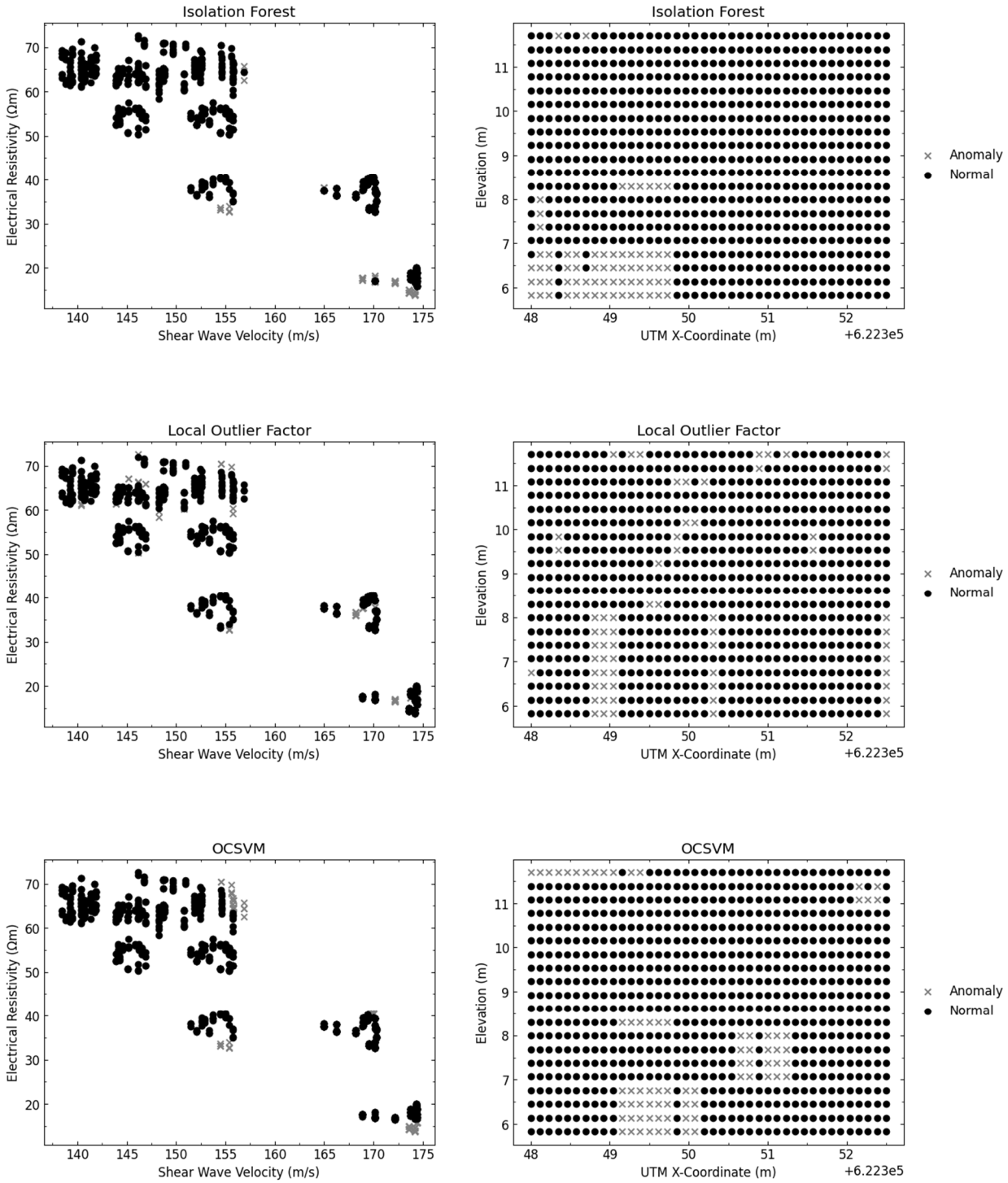




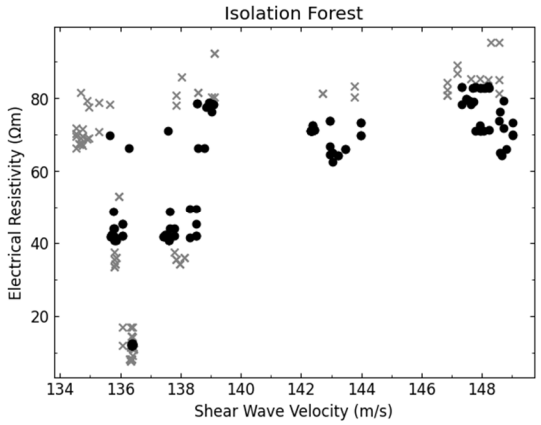


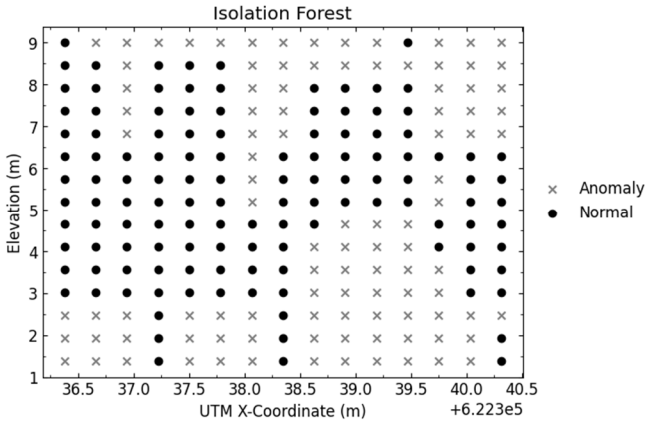
Fit: Dry Season Site 3 Compilation Dataframe
Predicted: Dry Season Site 3 Compilation Dataframe

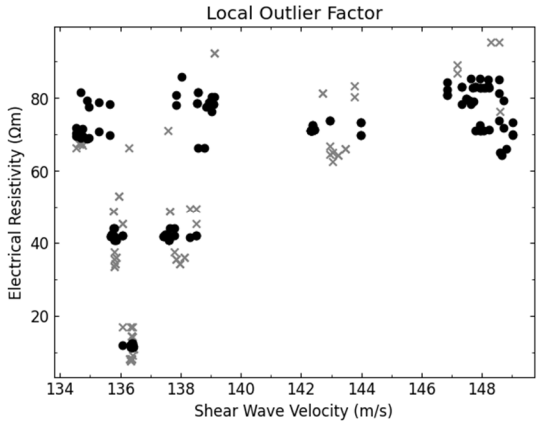
Profile: sc_site3

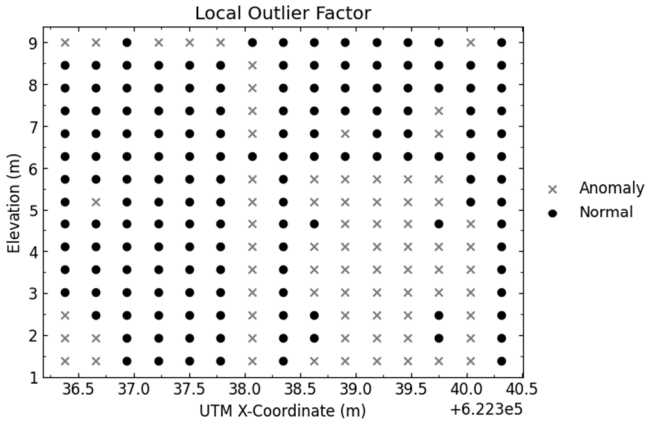


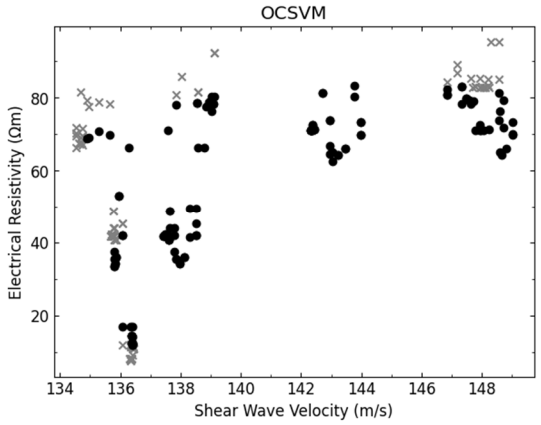
Profile: swt_site3

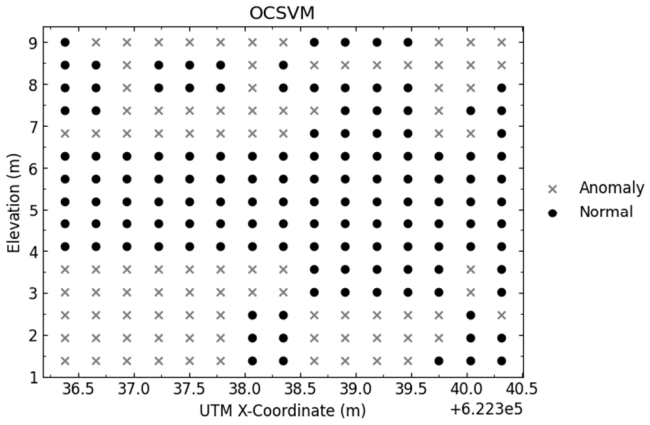






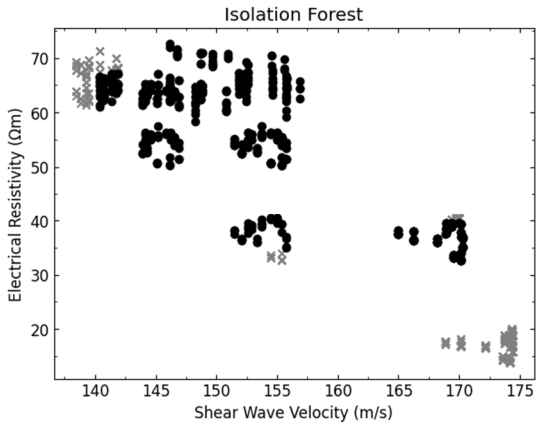


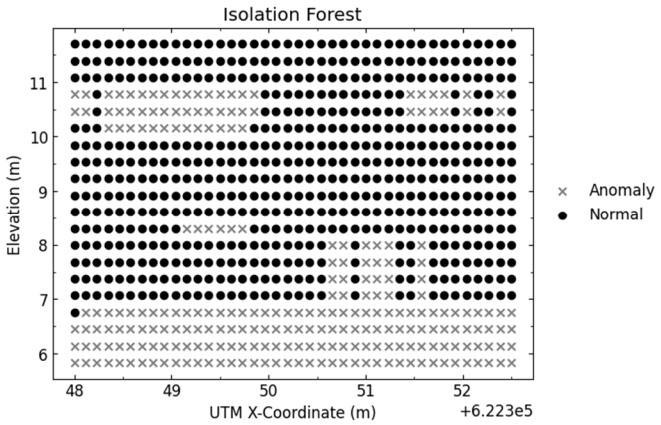


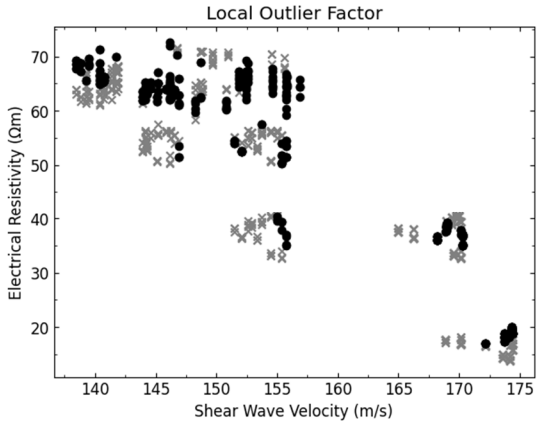


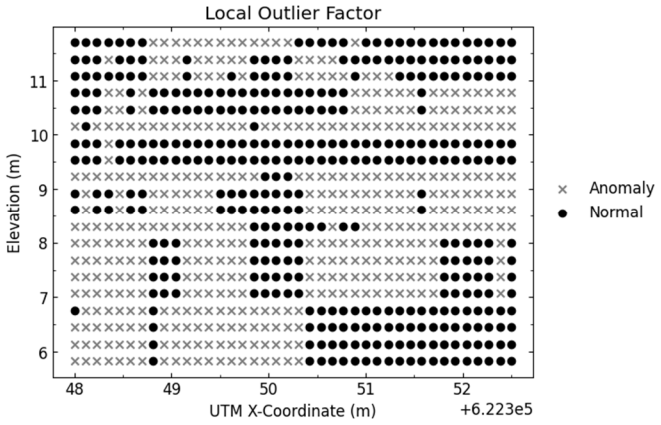
Fit: Dry Season Site 3 Compilation Dataframe
Predicted: Individual Profiles

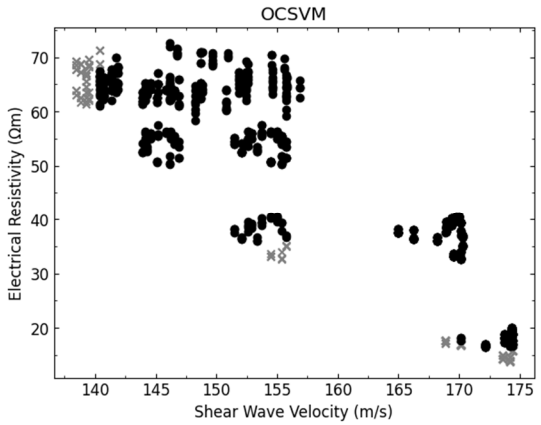
Profile: sc_site3

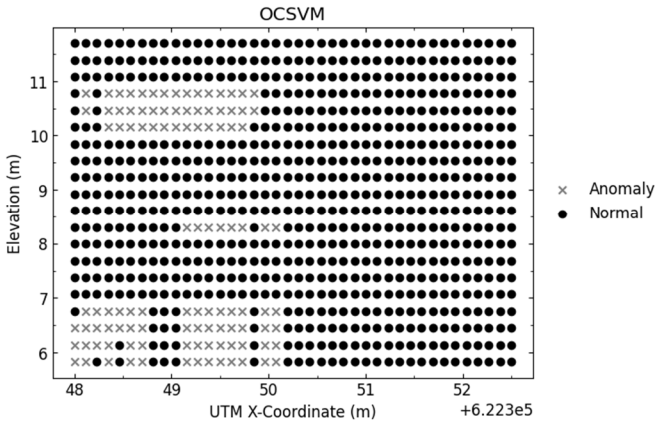




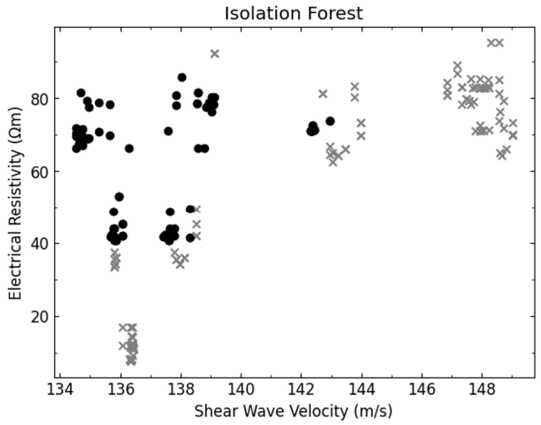


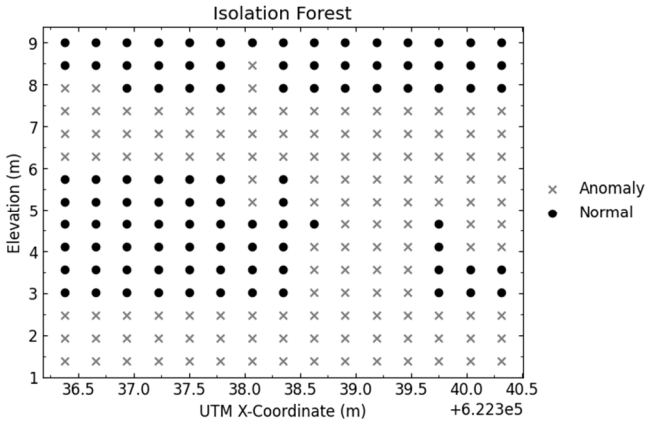


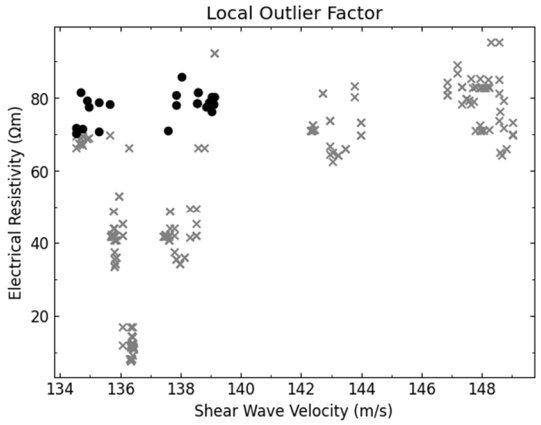


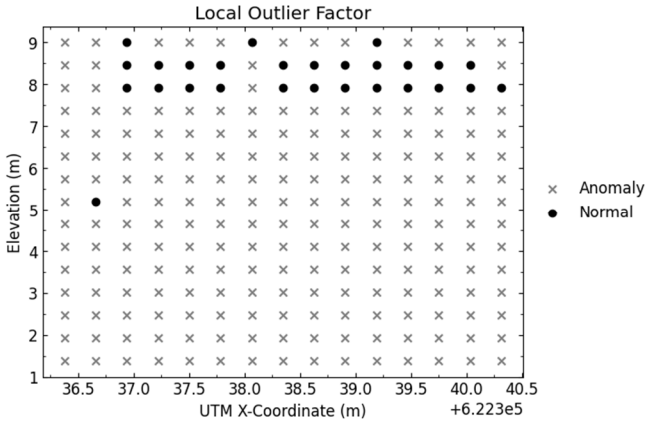


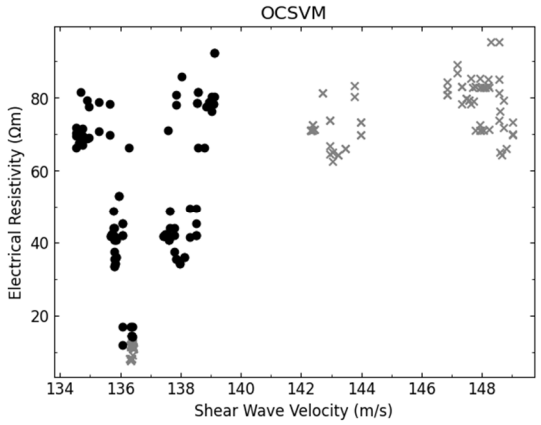
Profile: swt_site3

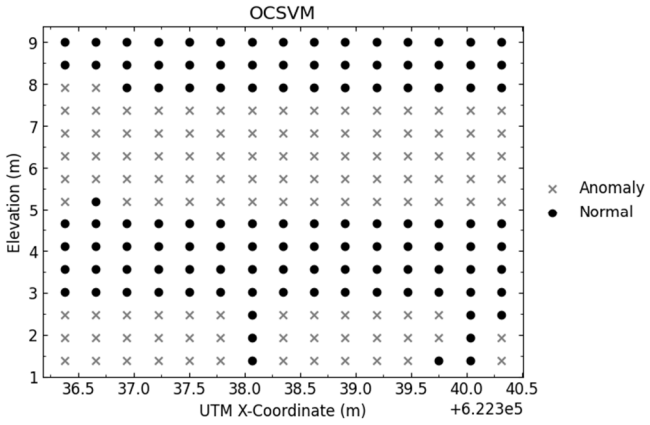






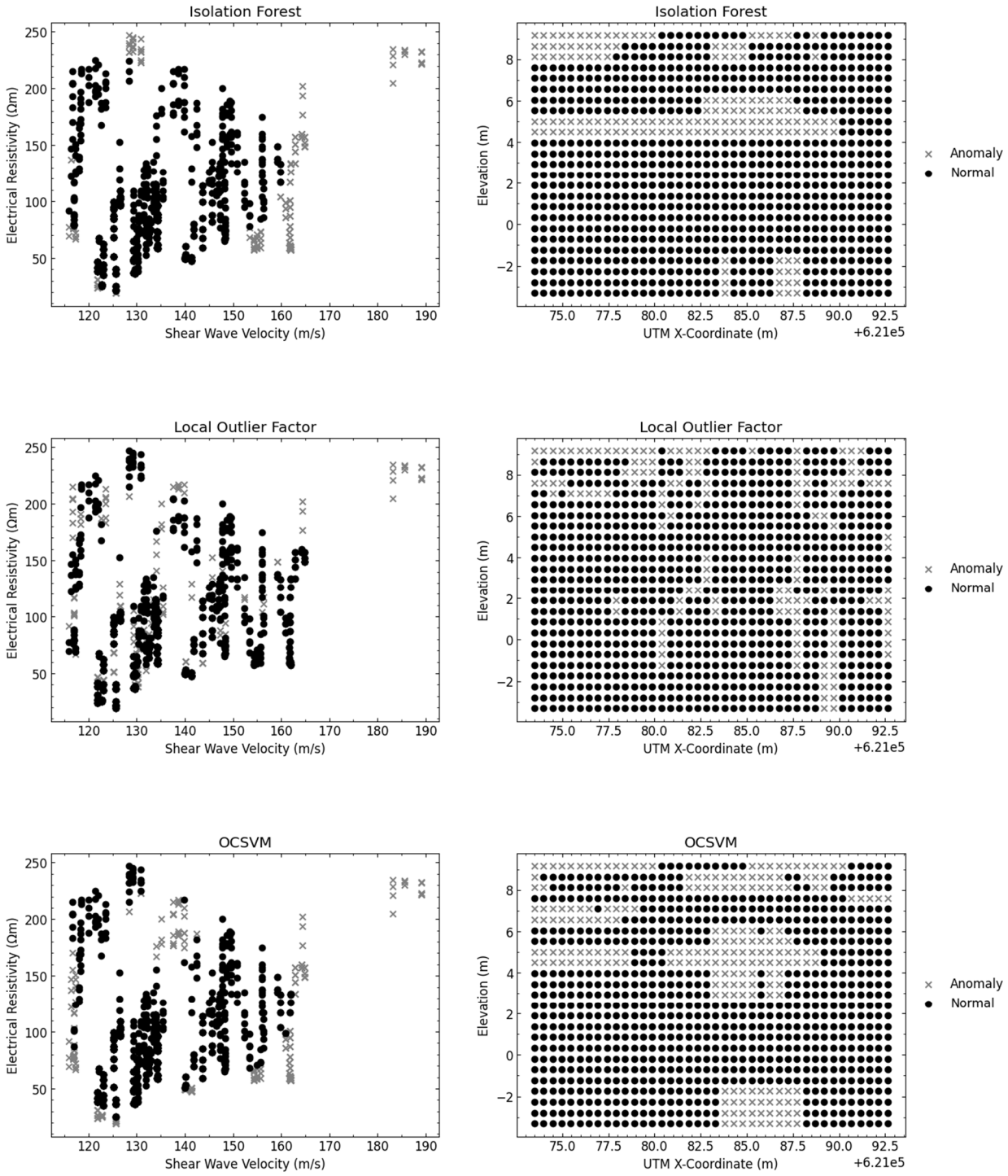




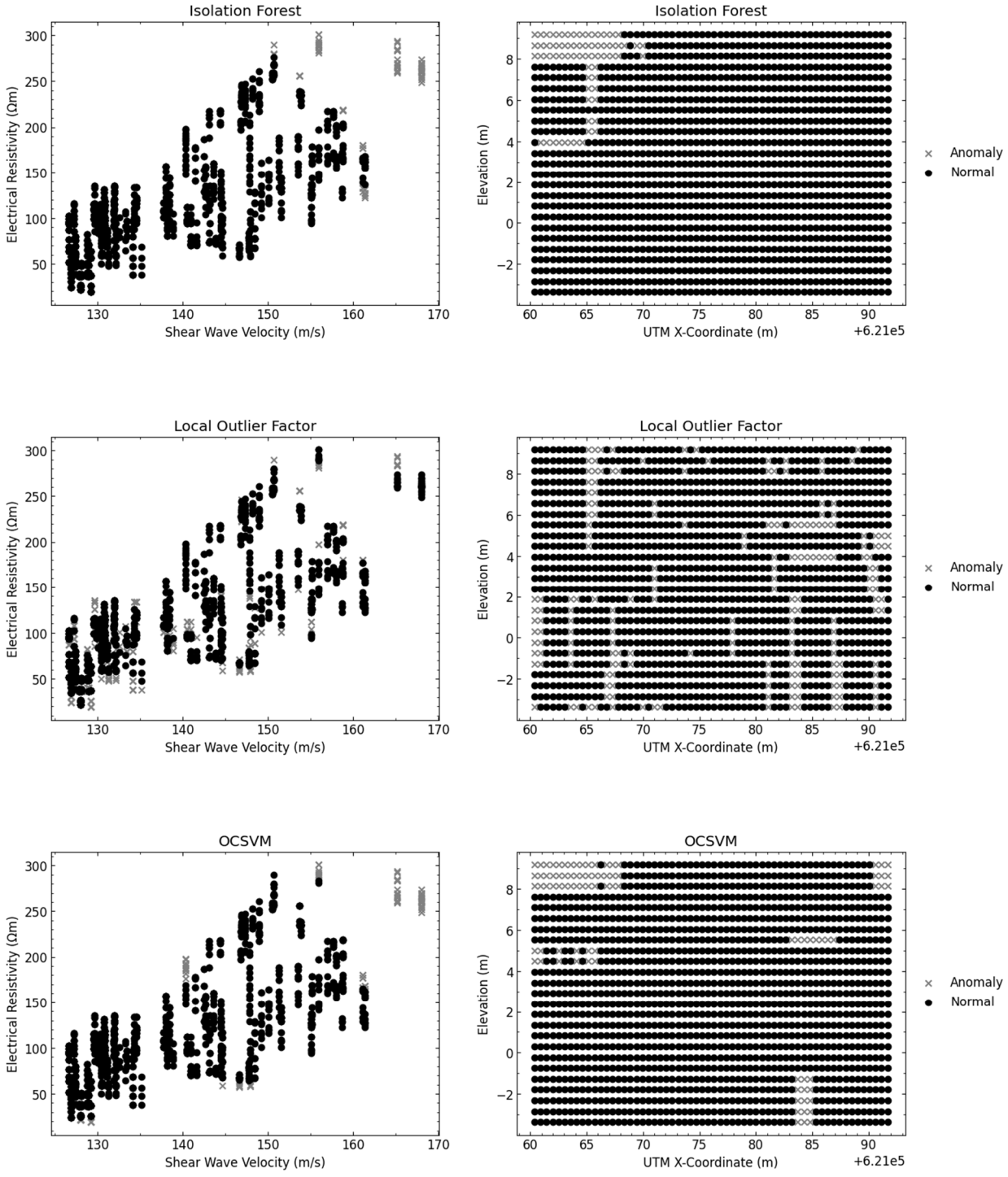


Fit: Wet Season Site 1 Compilation Dataframe
Predicted: Wet Season Site 1 Compilation Dataframe

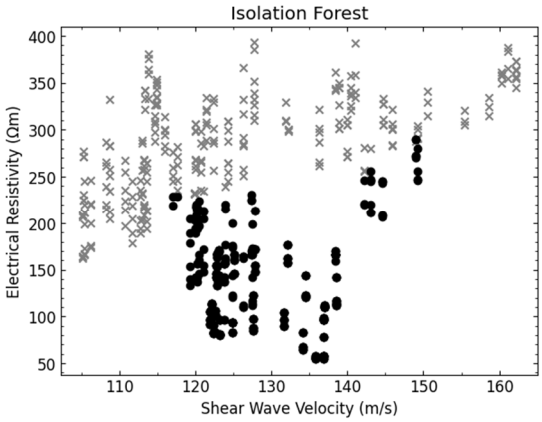
Profile: fc3_site1

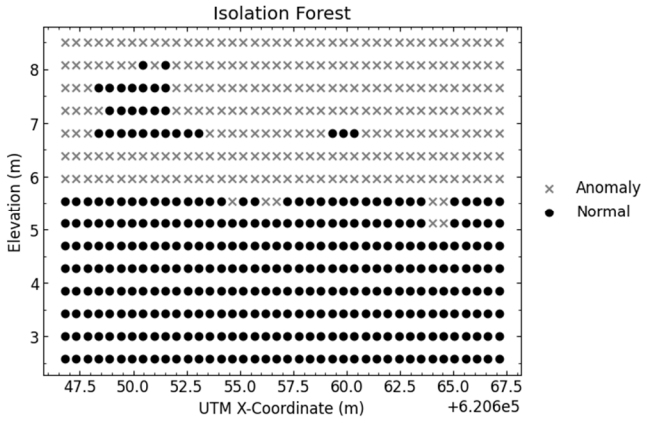


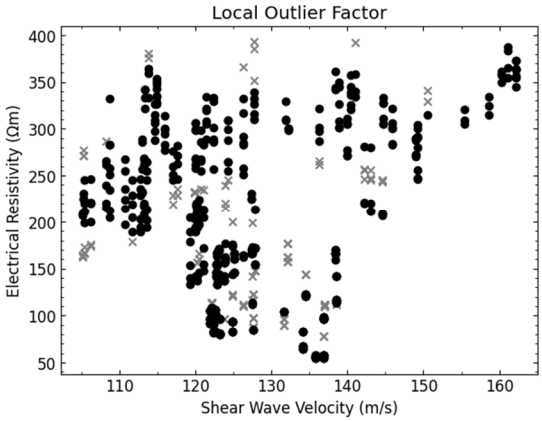
Profile: fc5_site1

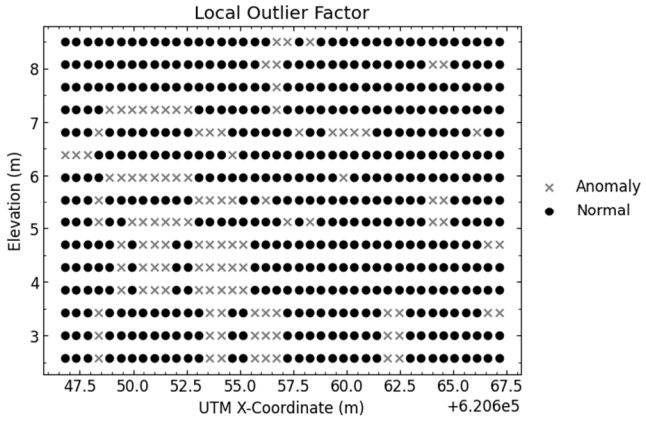


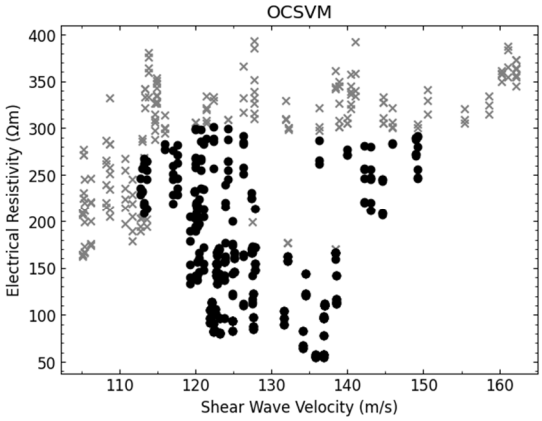
Profile: fwt_site1

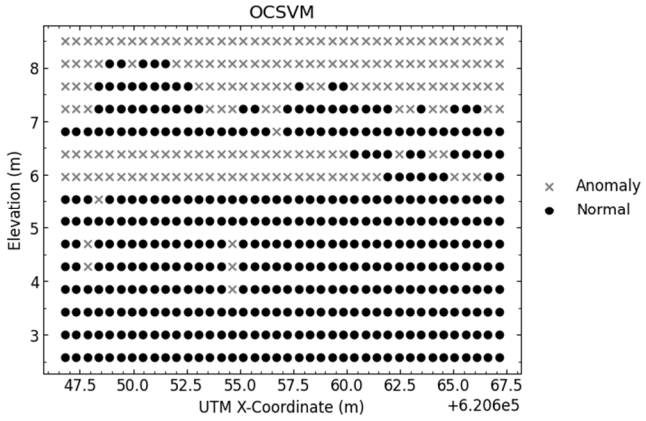






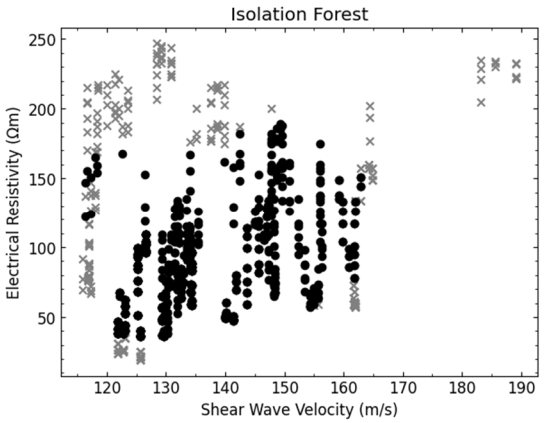


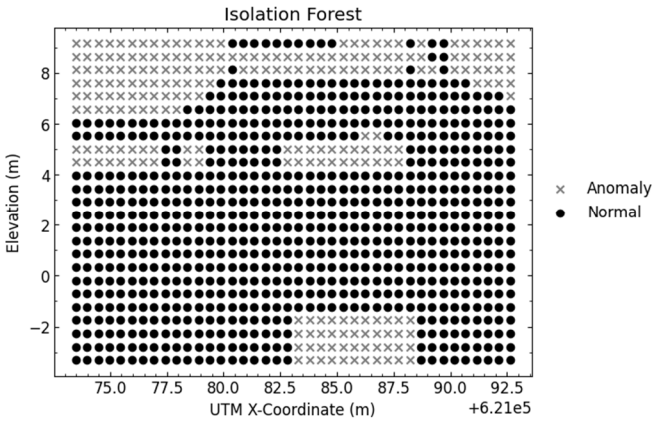


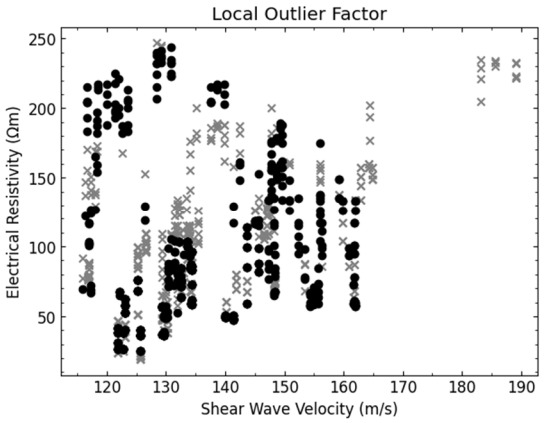


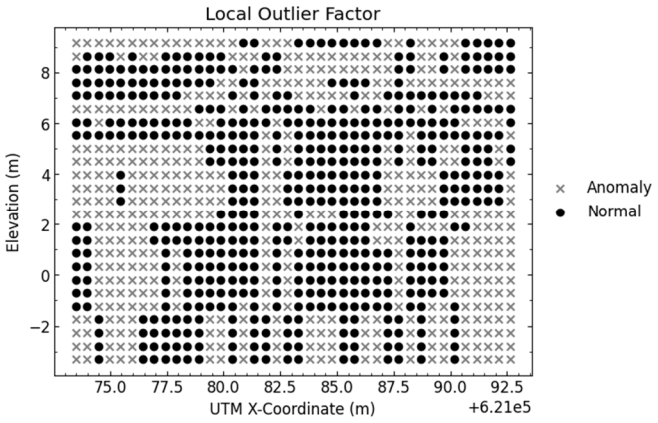
Fit: Wet Season Site 1 Compilation Dataframe
Predicted: Individual Profiles

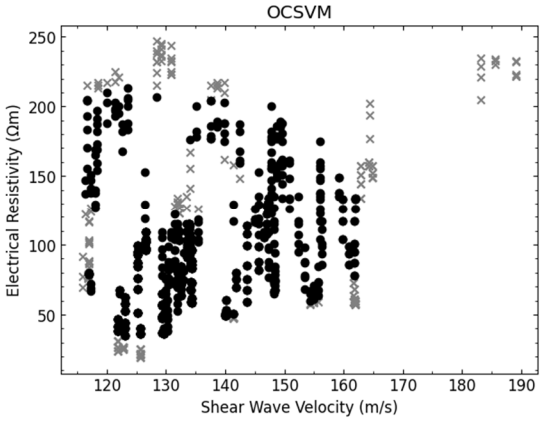
Profile: fc3_site1

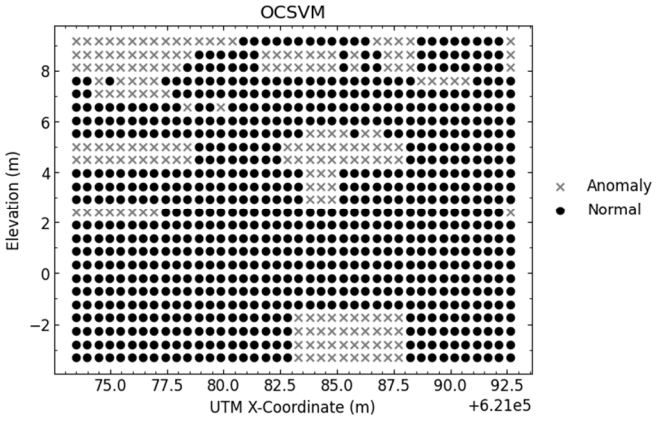




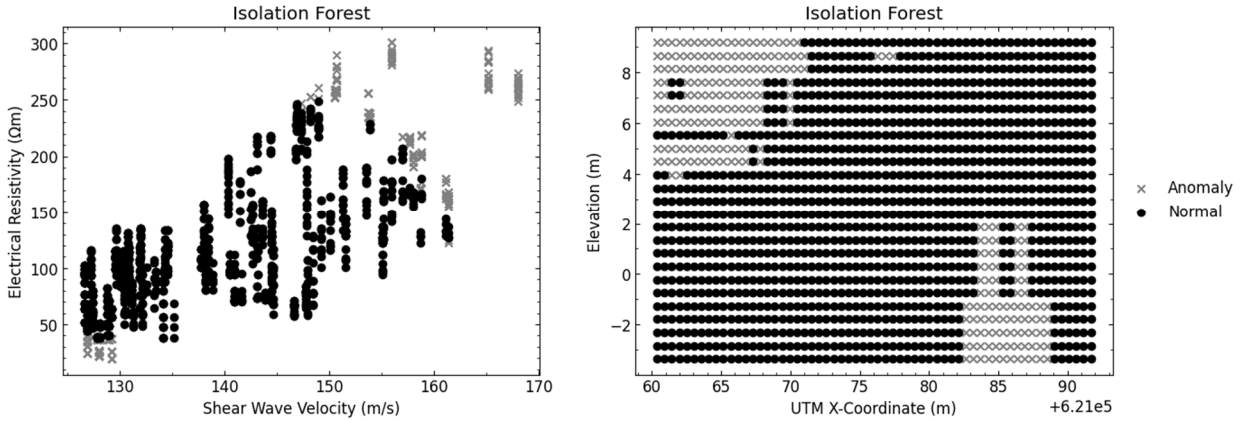


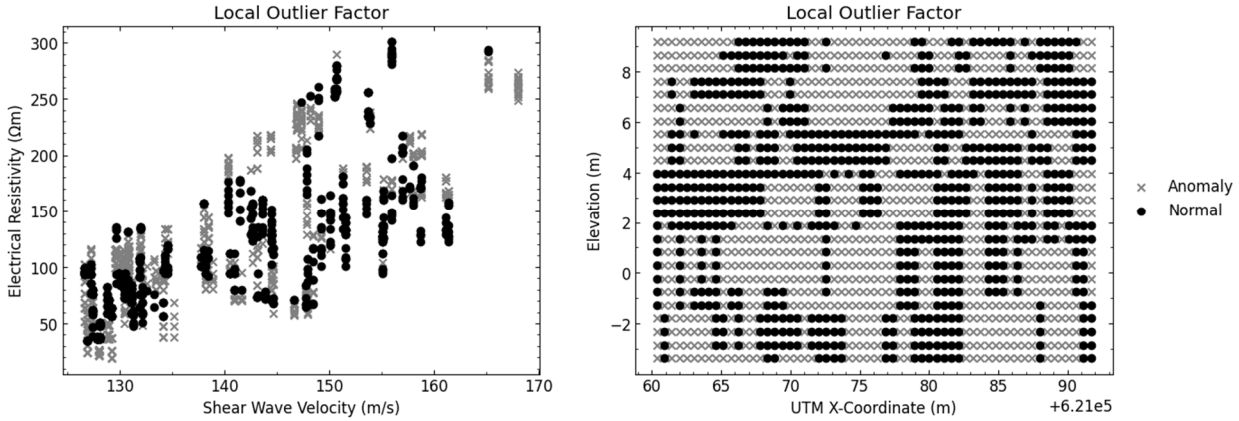


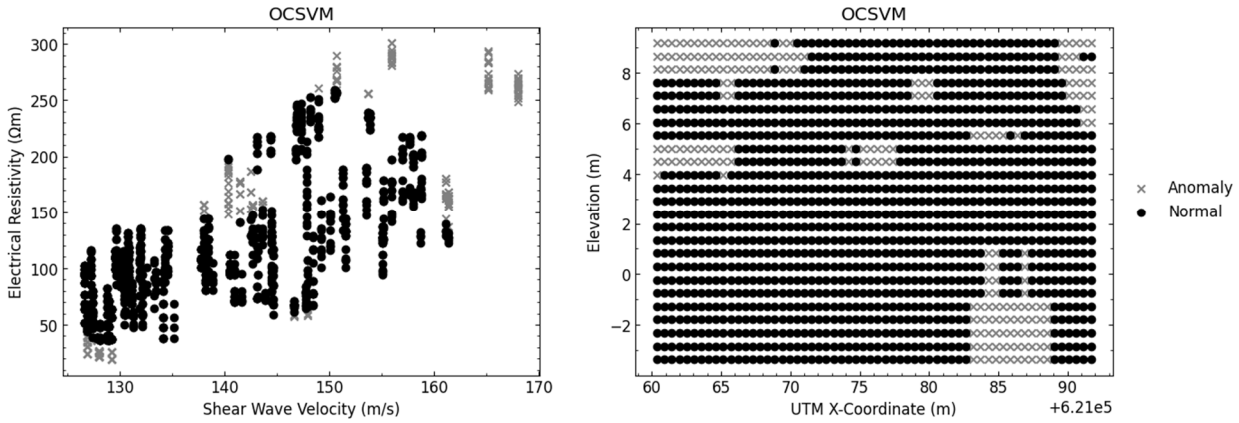




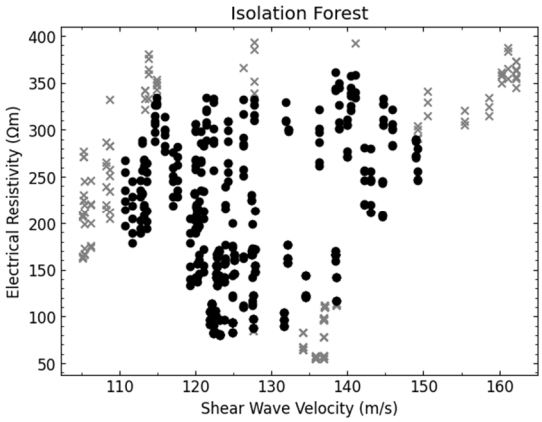
Profile: fc5_site1

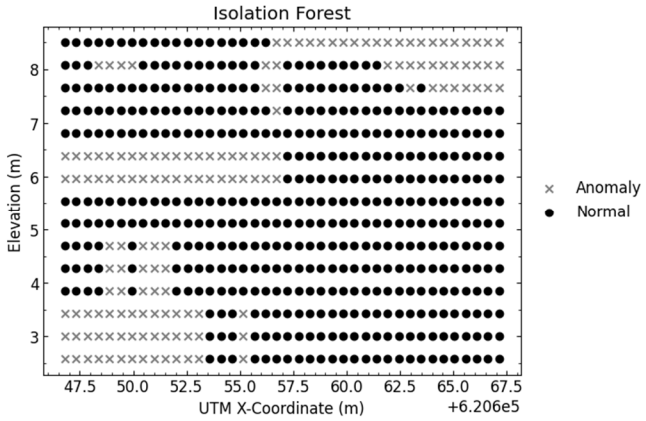


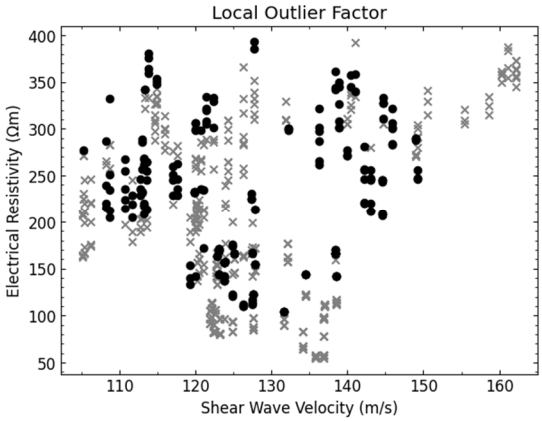


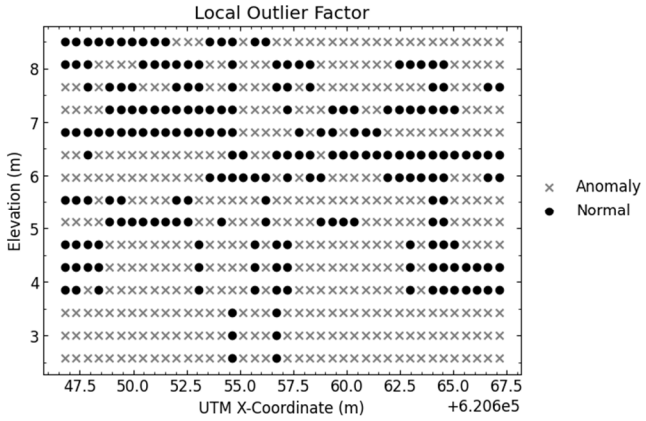


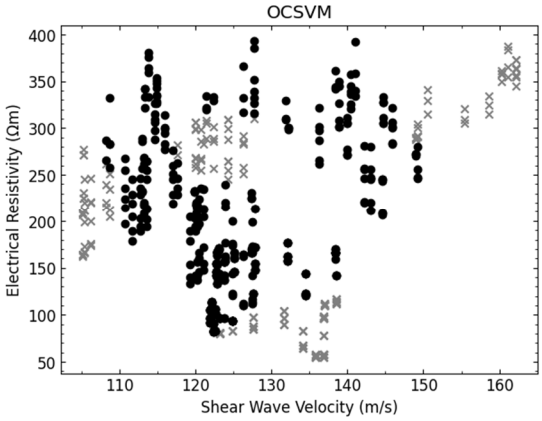
Profile: fwt_site1

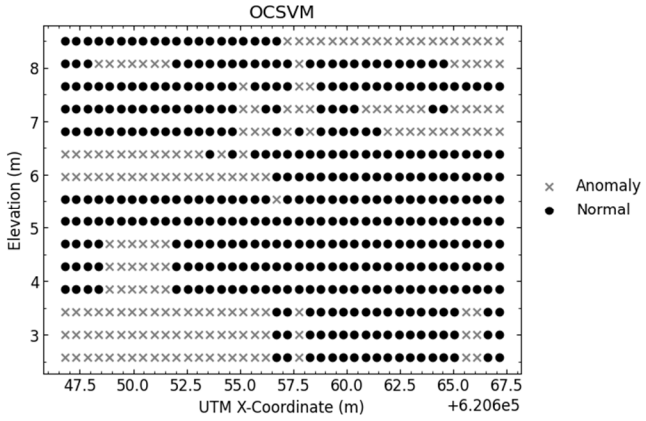






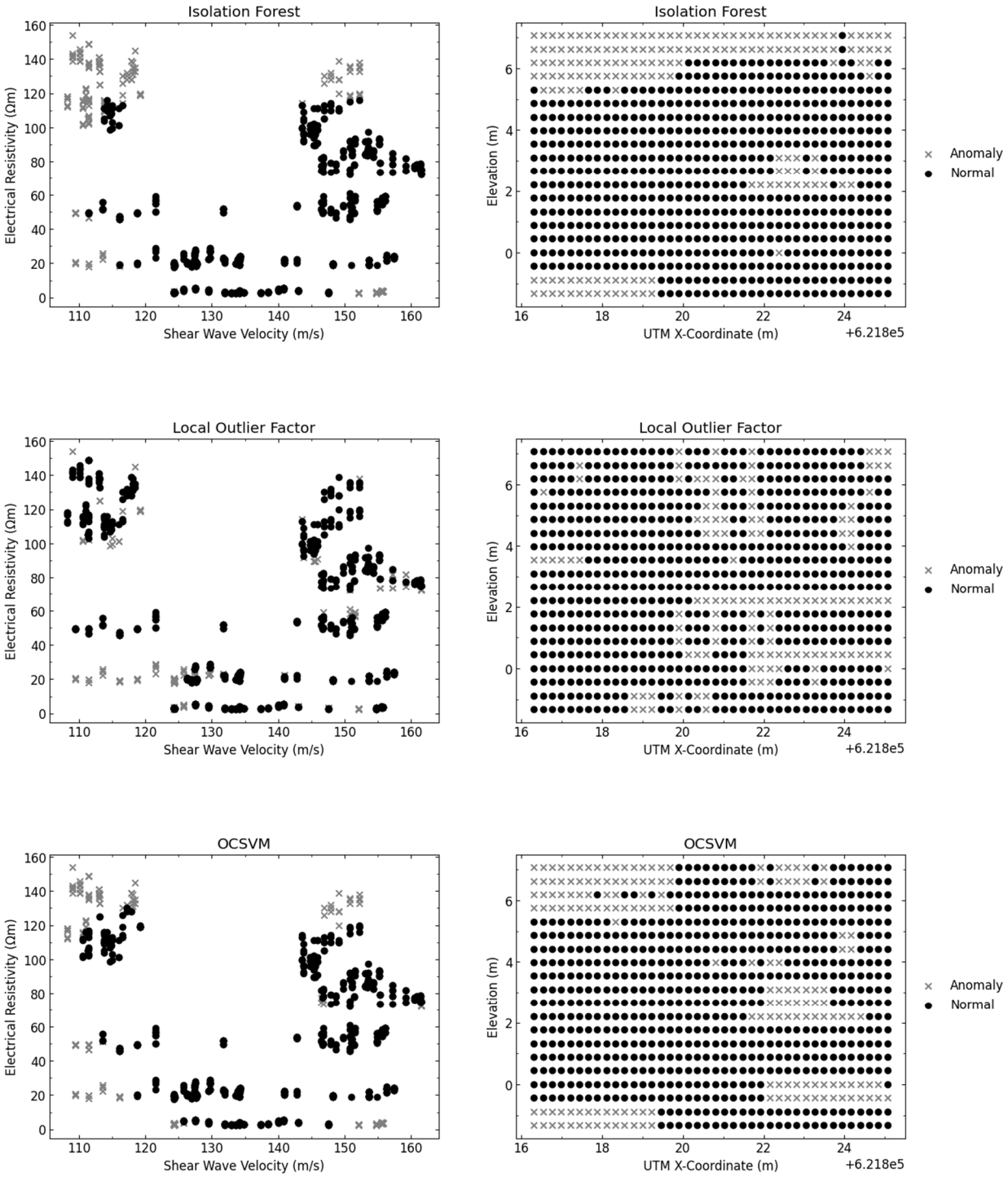




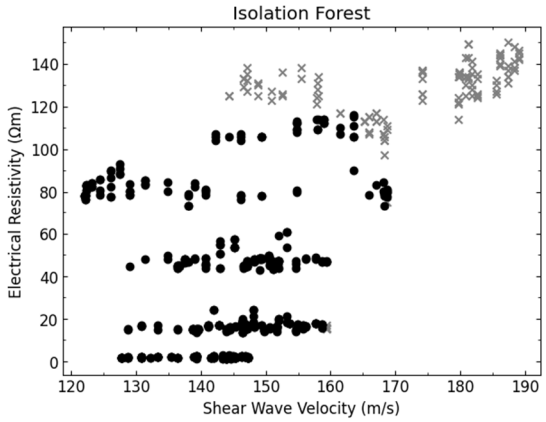


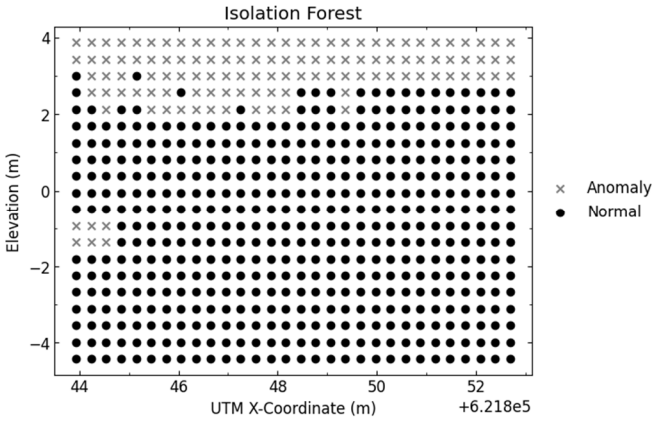
Fit: Wet Season Site 2 Compilation Dataframe
Predicted: Wet Season Site 2 Compilation Dataframe

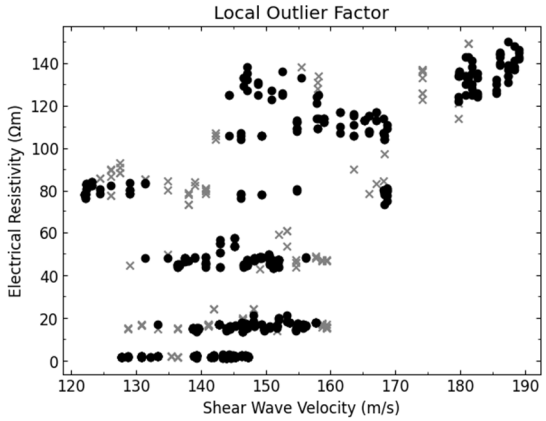
Profile: fc_site2

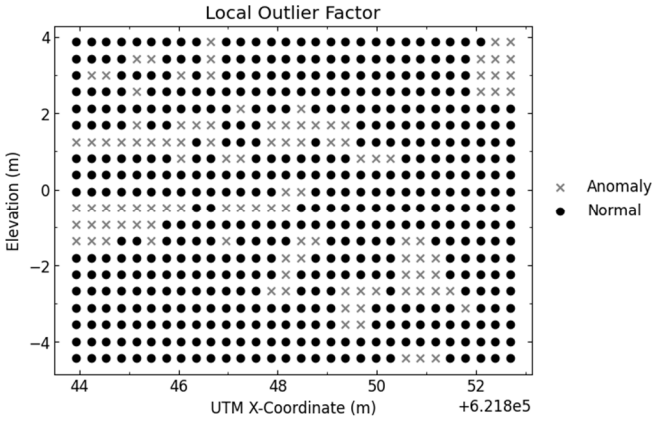


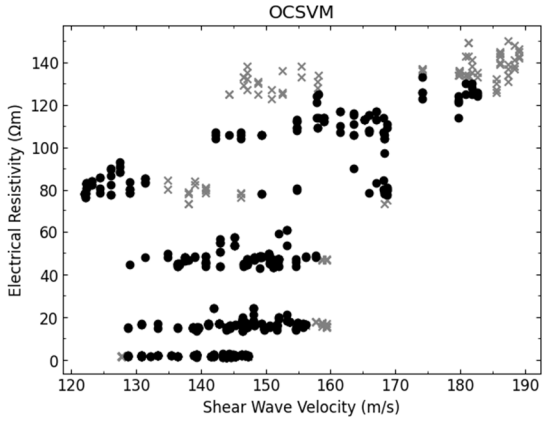
Profile: fdt_site2

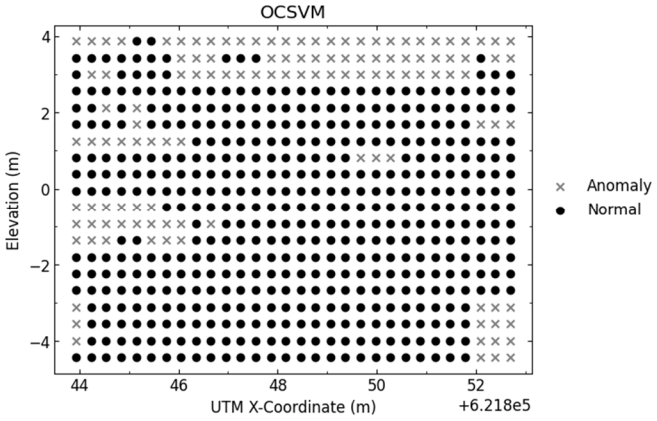




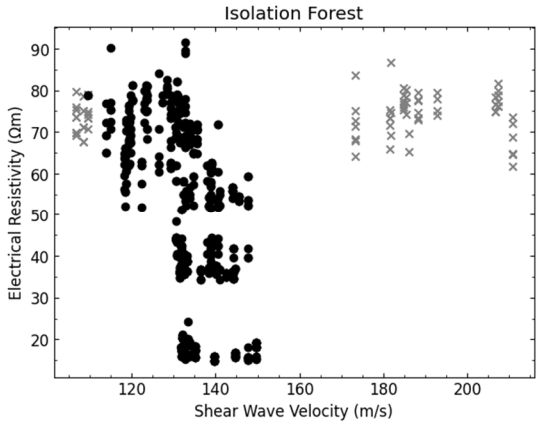


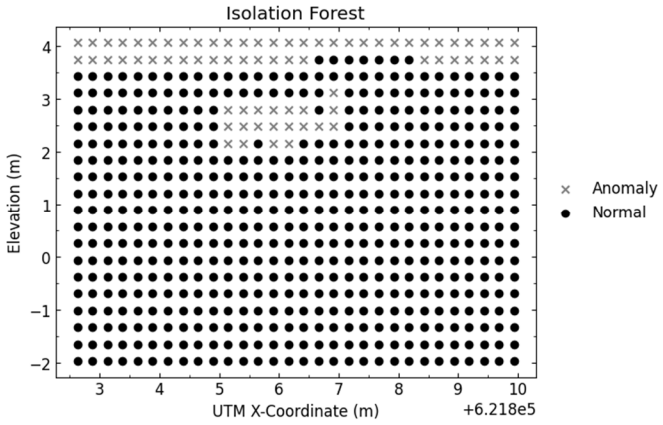


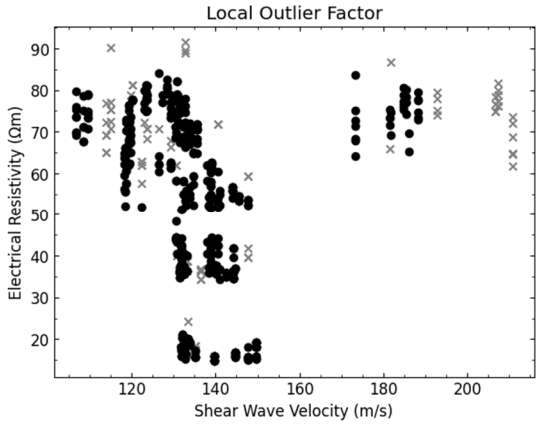


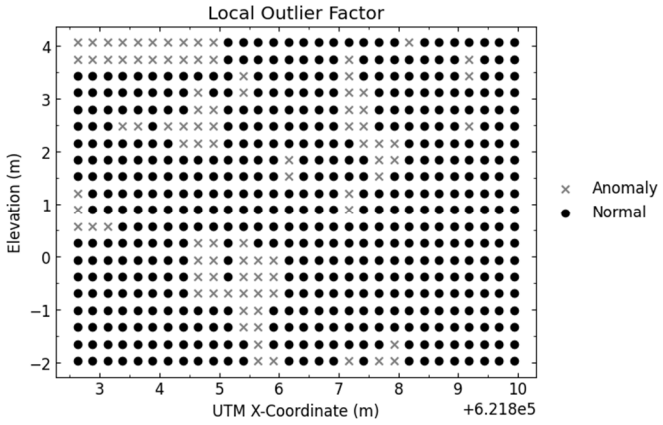


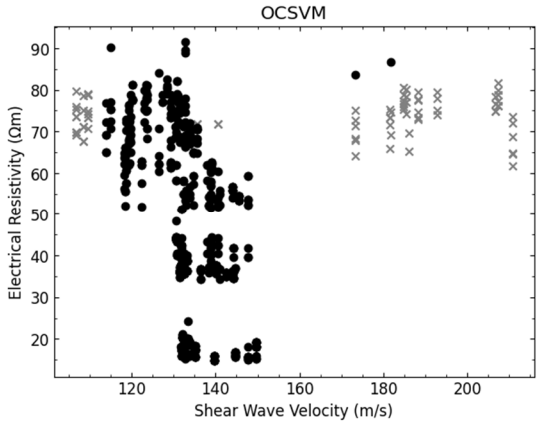
Profile: fwt_site2

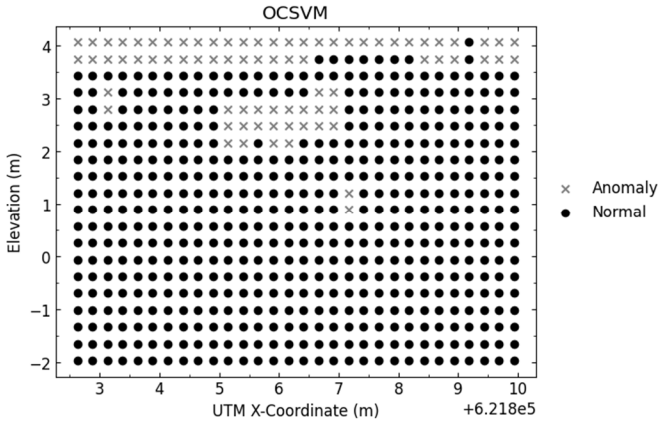






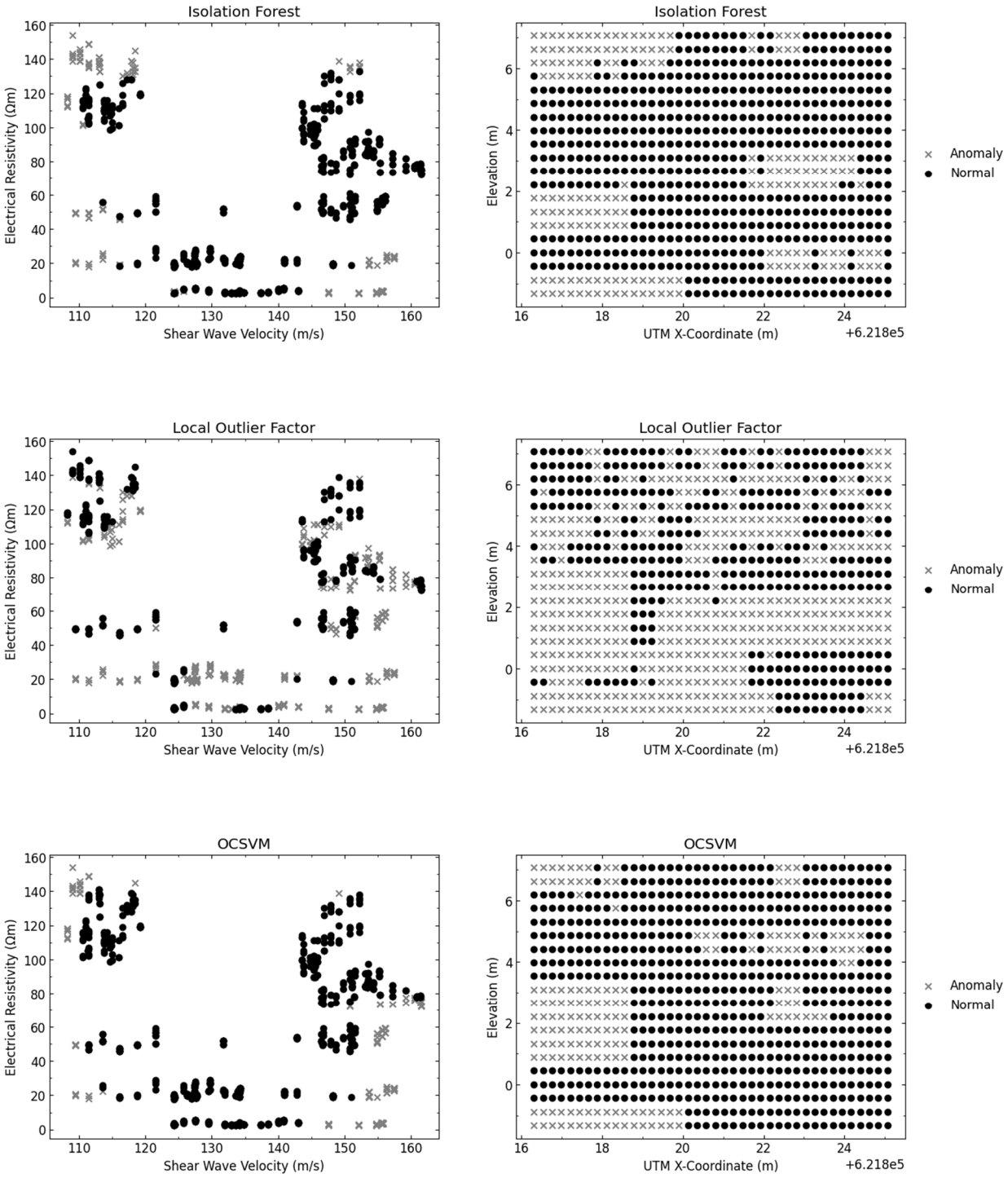




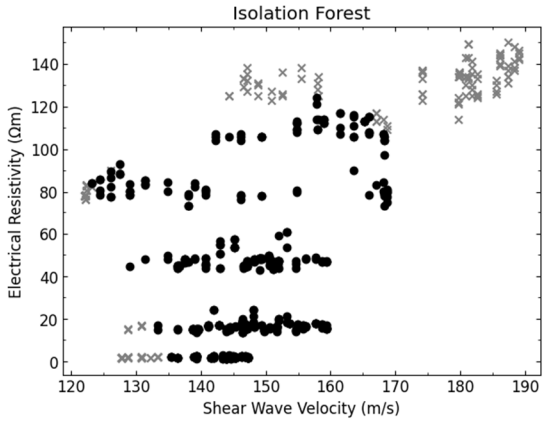


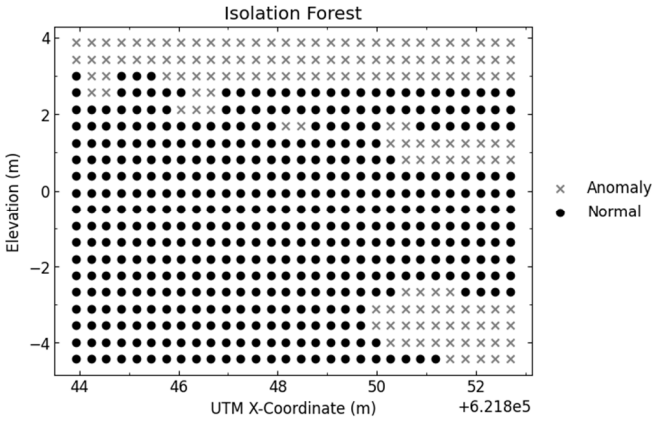
Fit: Wet Season Site 2 Compilation Dataframe
Predicted: Individual Profiles

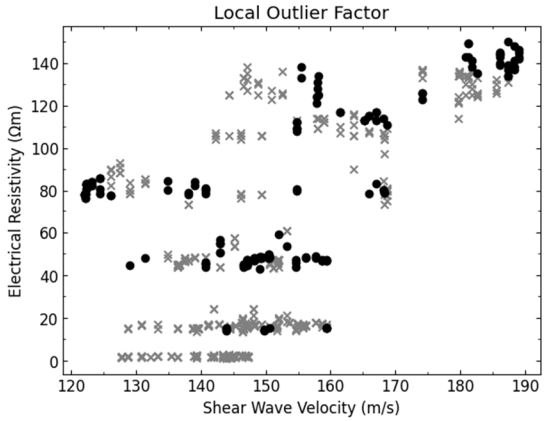
Profile: fc_site2

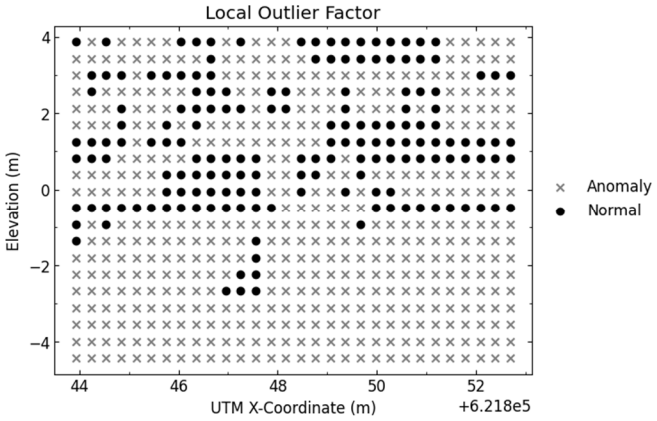


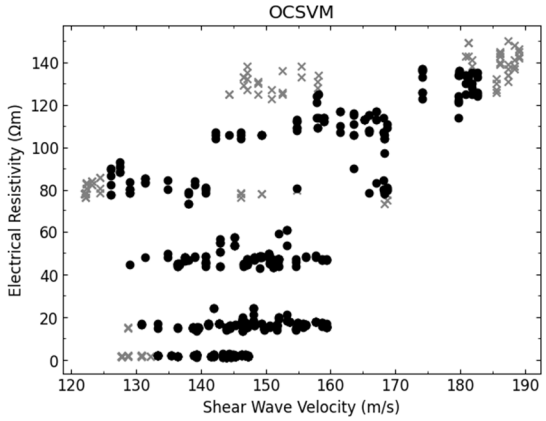
Profile: fdt_site2

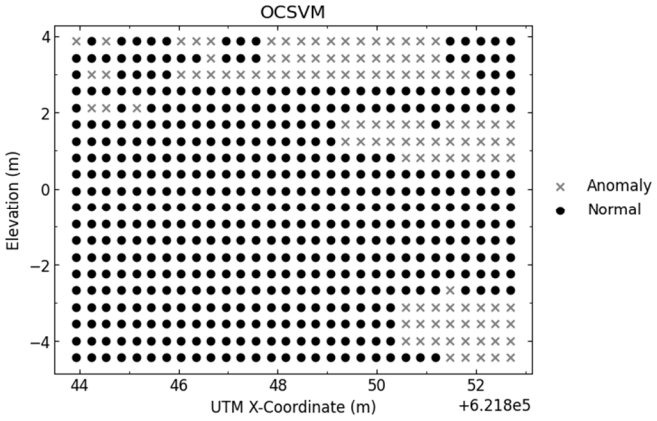




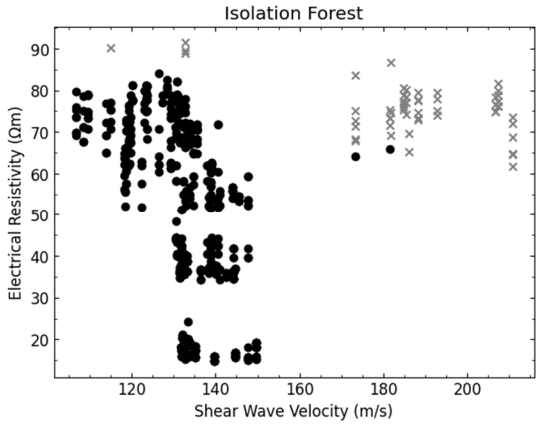


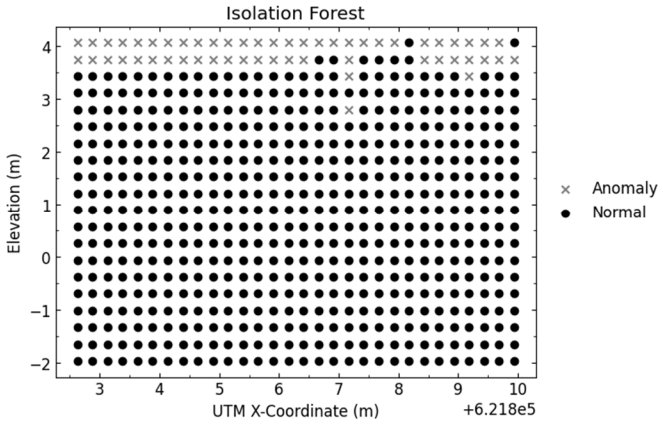


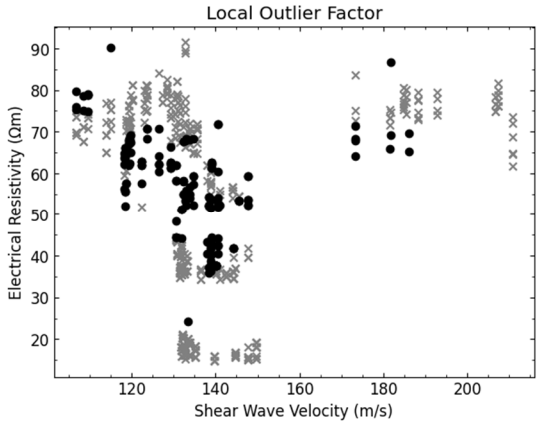


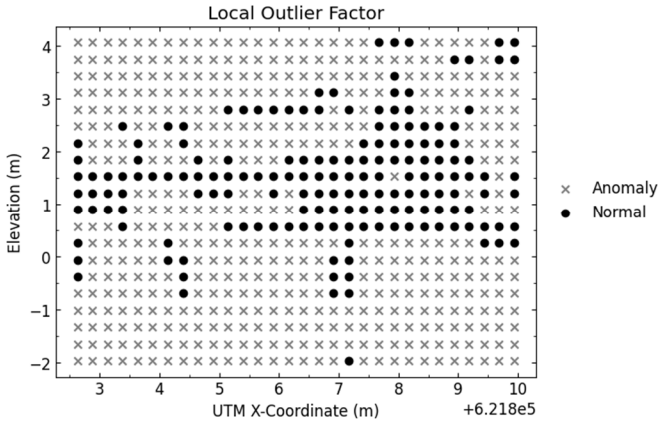


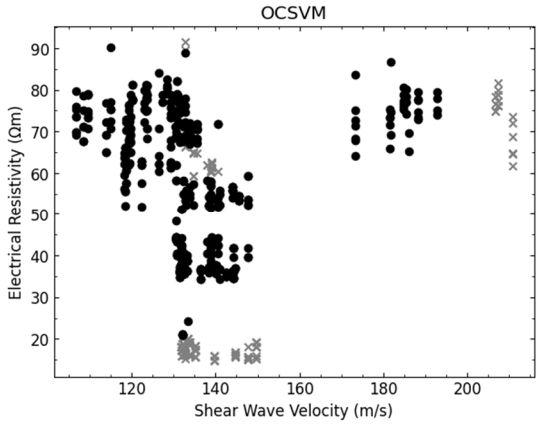
Profile: fwt_site2

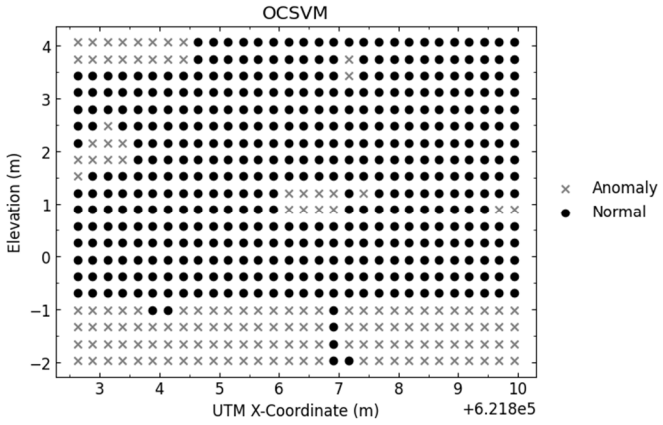






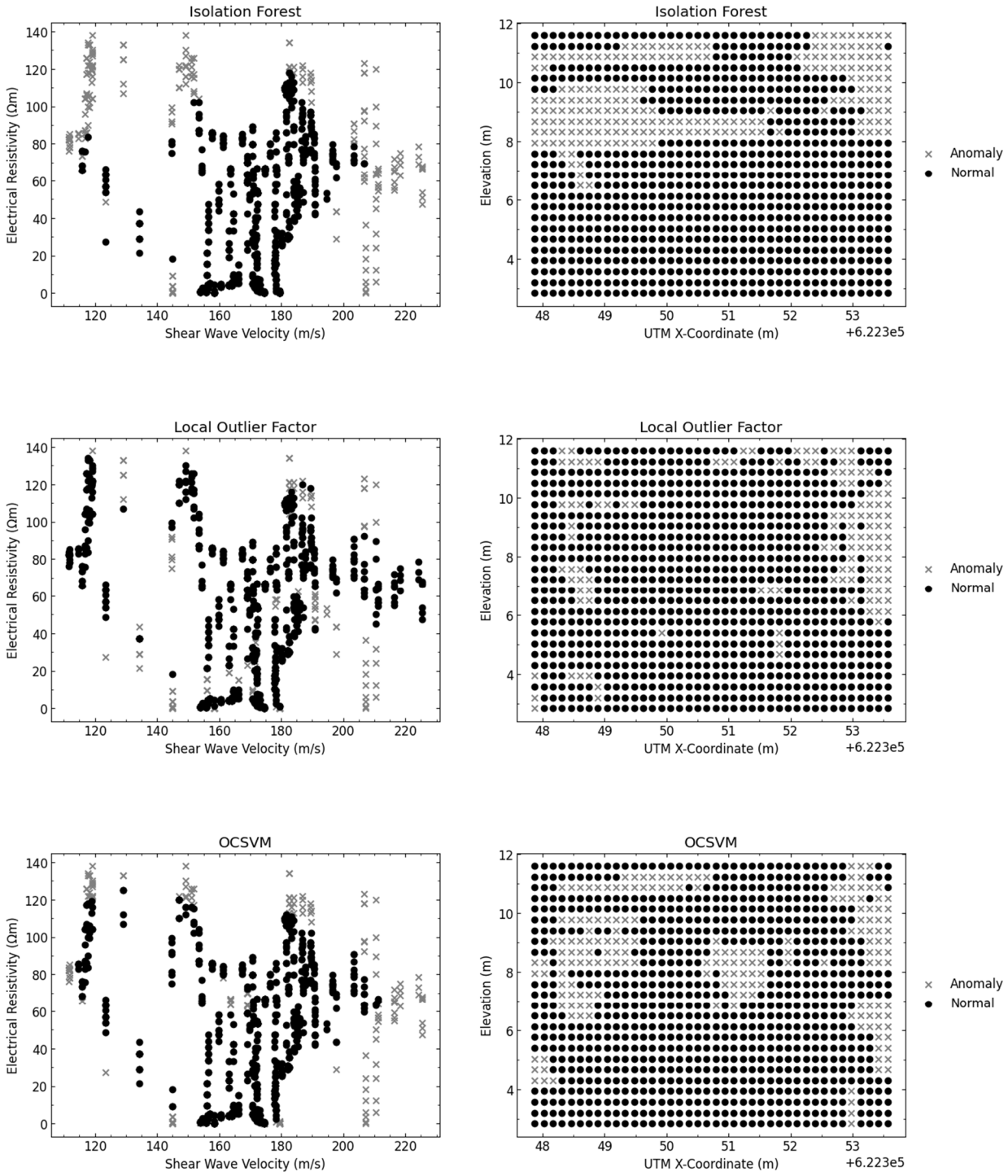




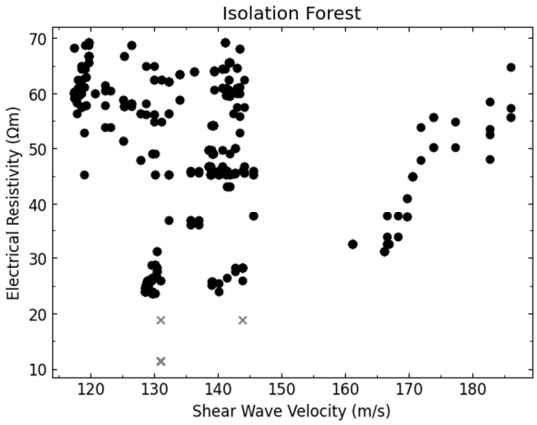


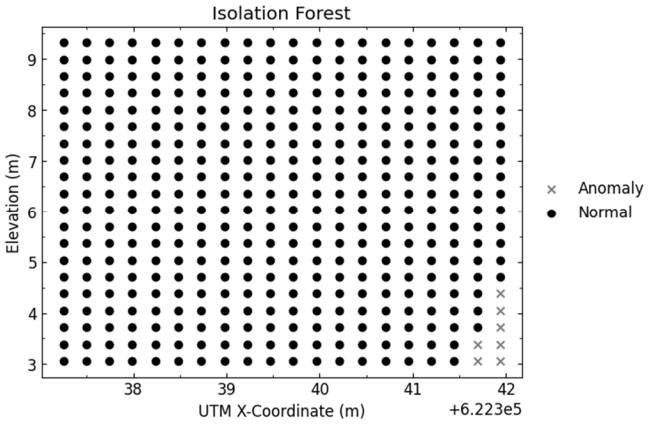
Fit: Wet Season Site 3 Compilation Dataframe
Predicted: Wet Season Site 3 Compilation Dataframe

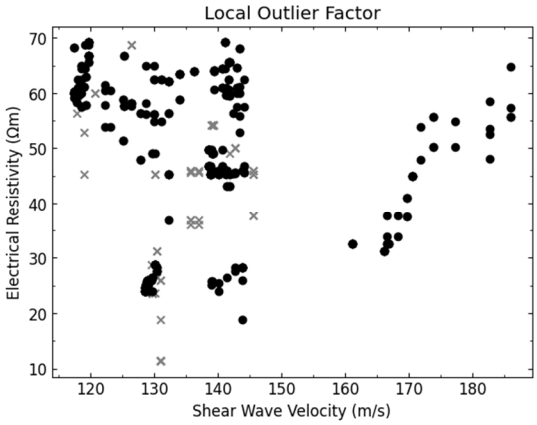
Profile: fc_site3

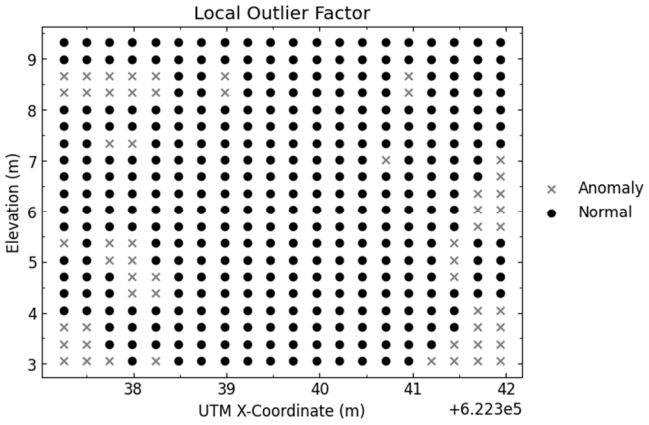


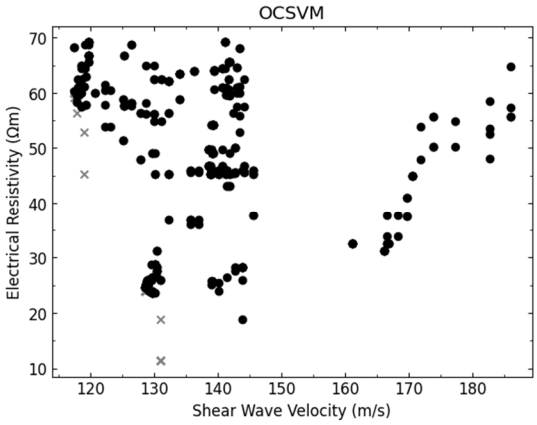
Profile: fwt_site3

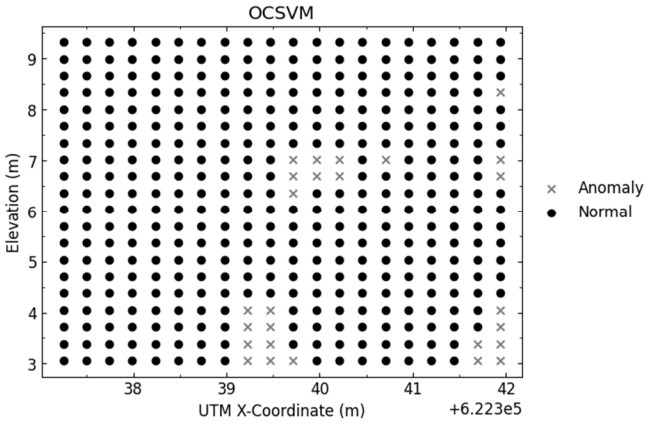






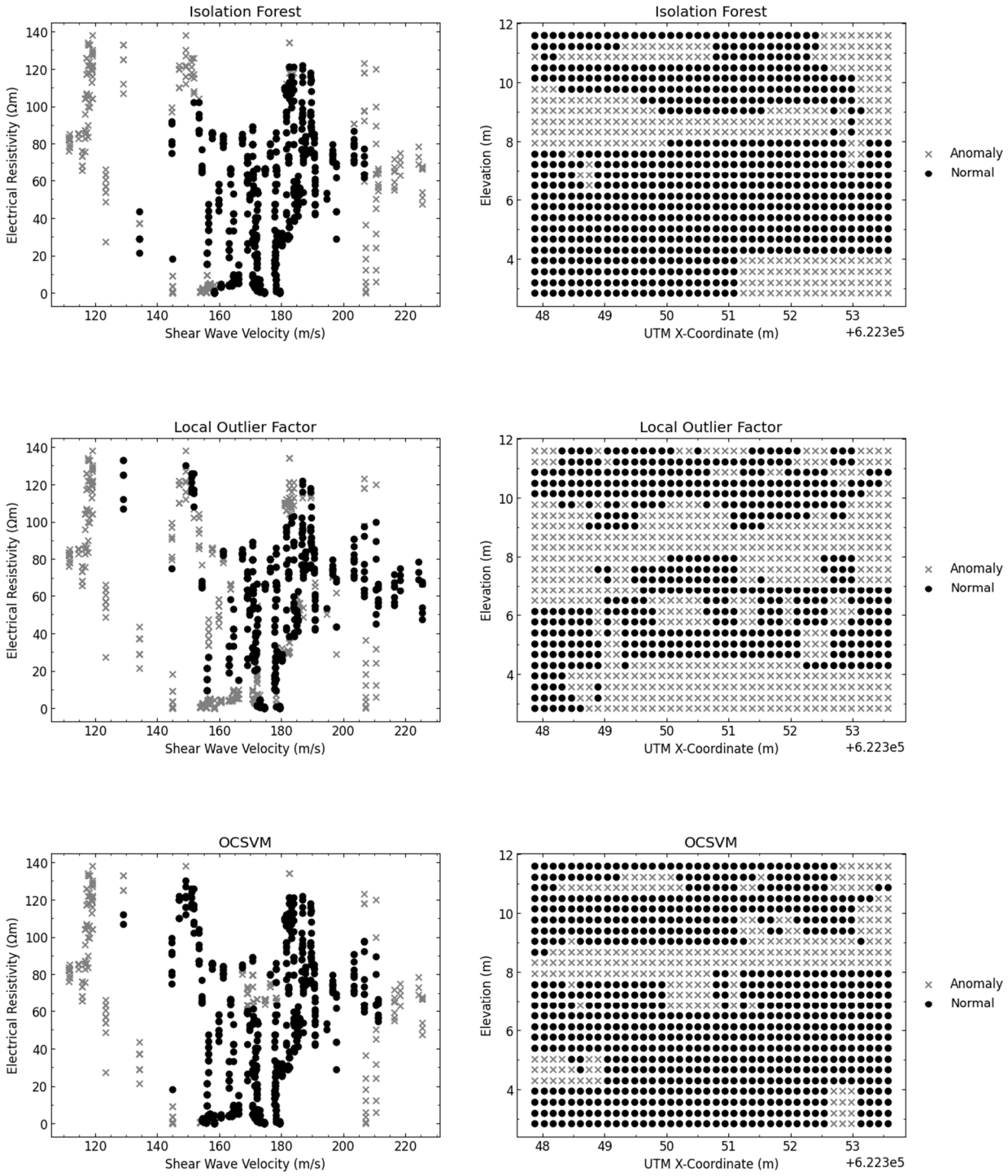




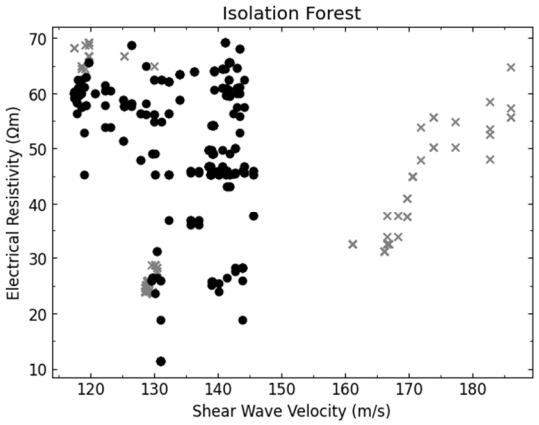


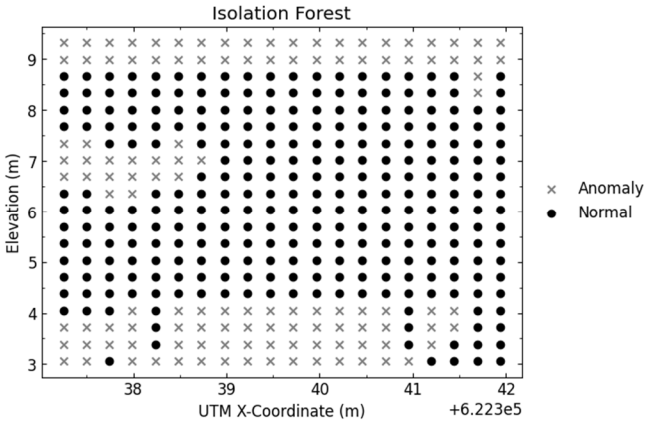
Fit: Wet Season Site 3 Compilation Dataframe
Predicted: Individual Profiles

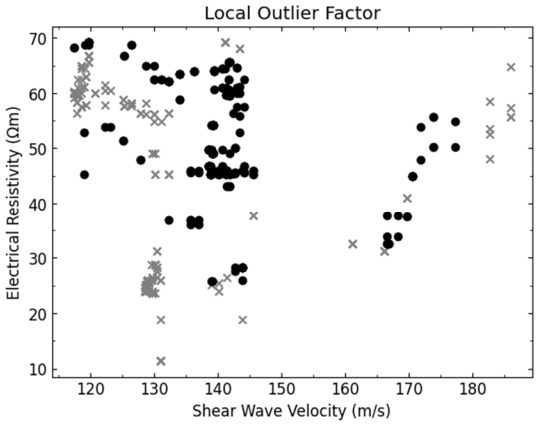
Profile: fc_site3

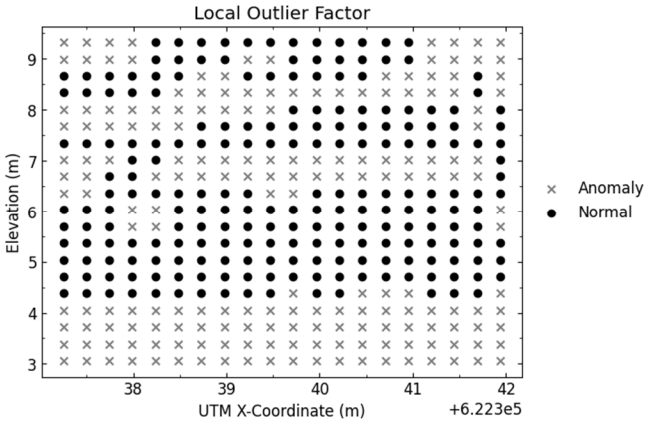


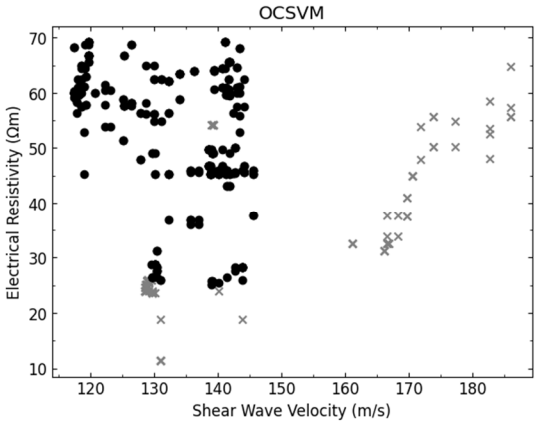
Profile: fwt_site3

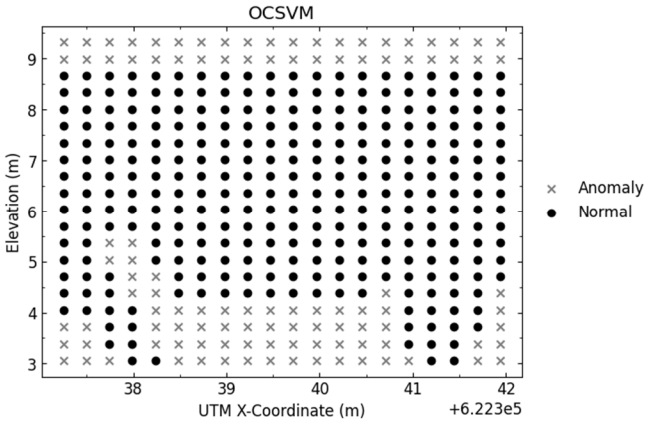






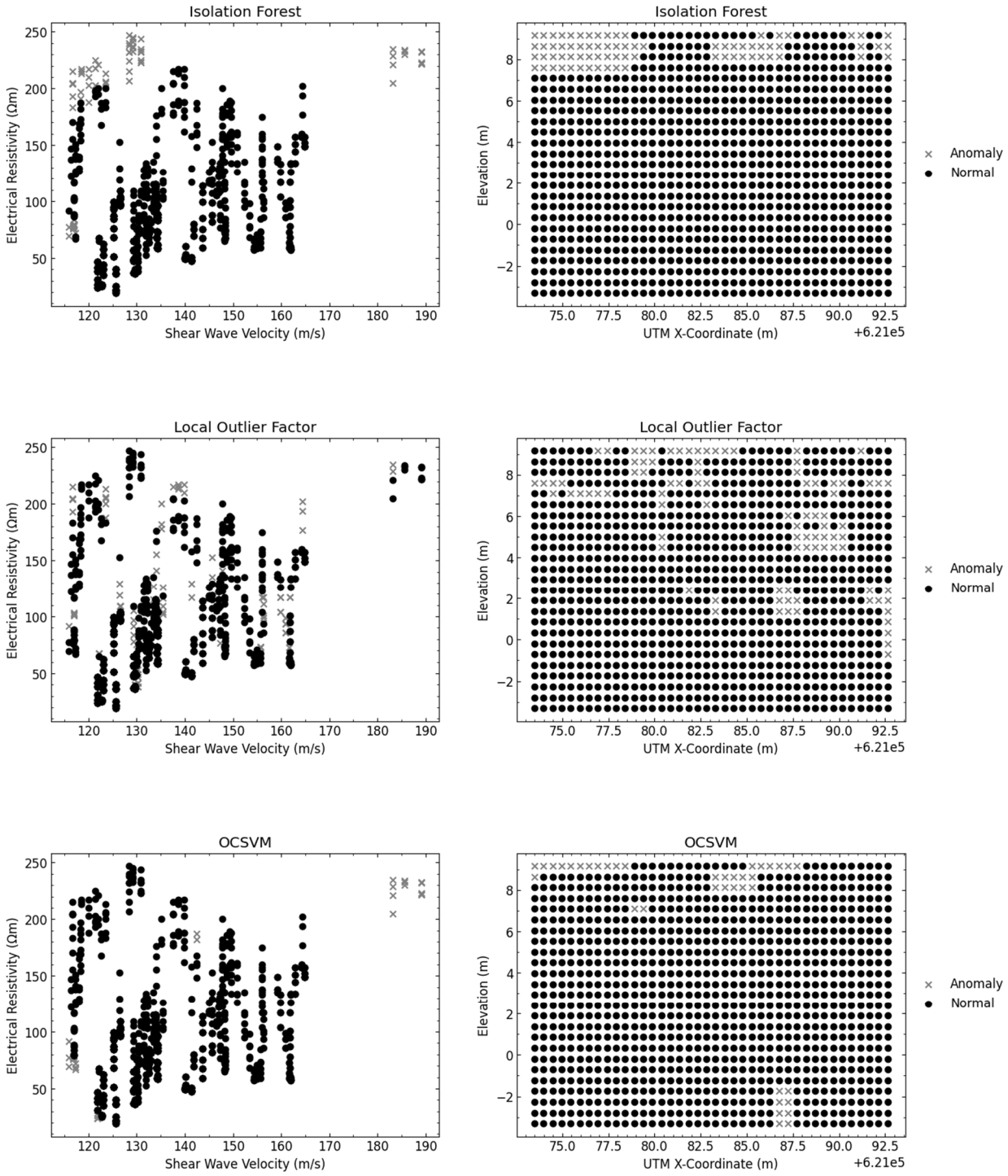




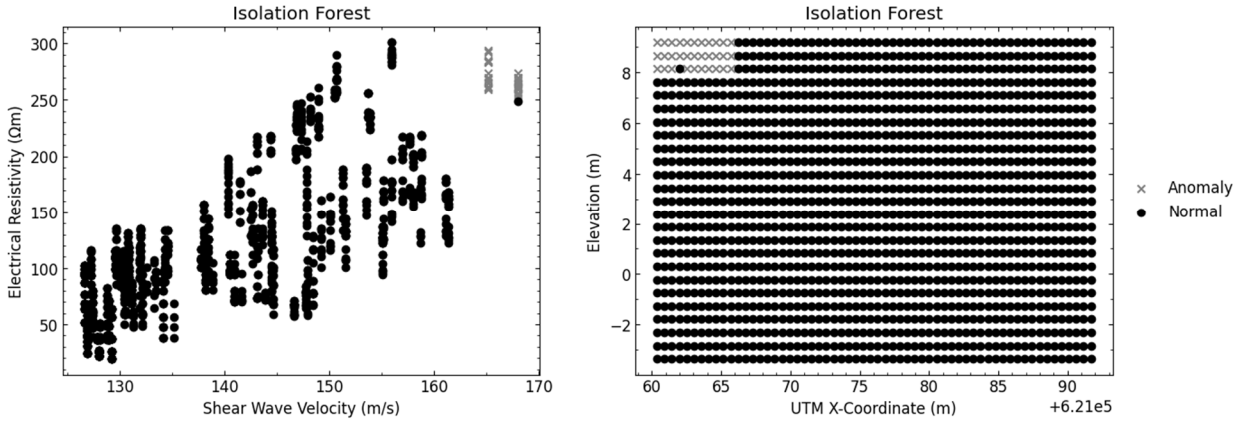


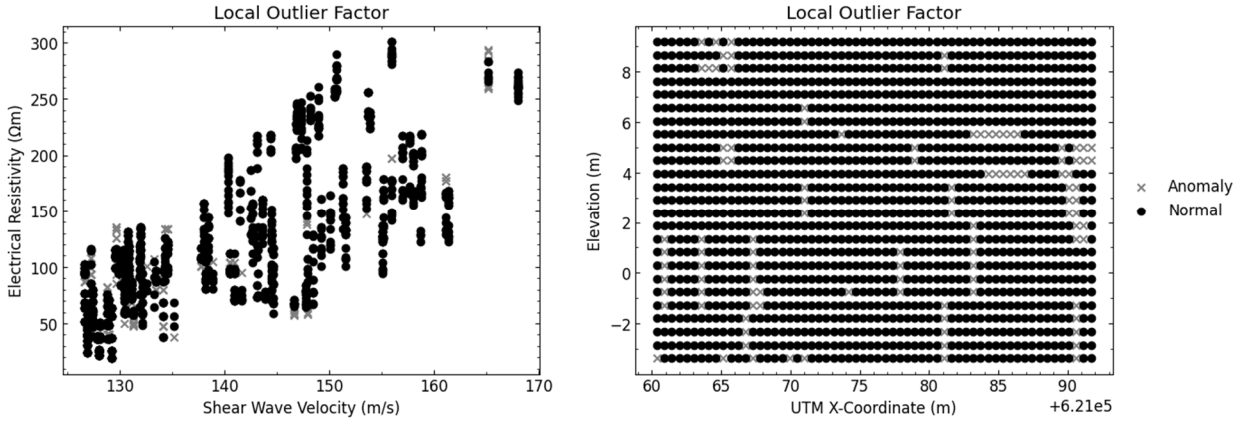
Fit: Site 1 Compilation Dataframe
Predicted: Site 1 Compilation Dataframe

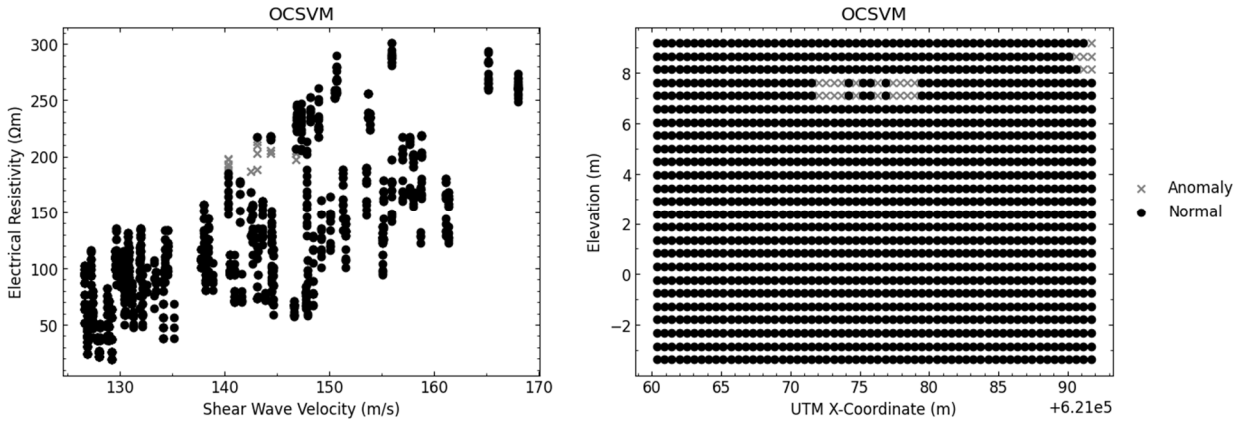
Profile: fc3_site1



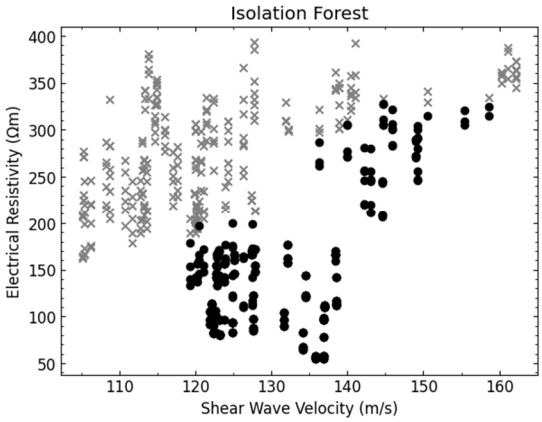
Profile: fc5_site1

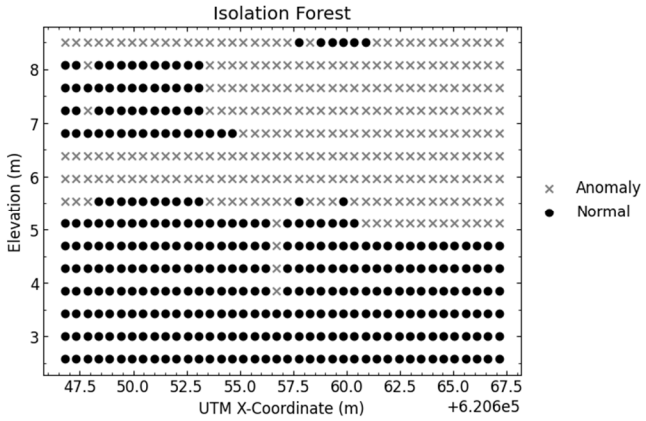


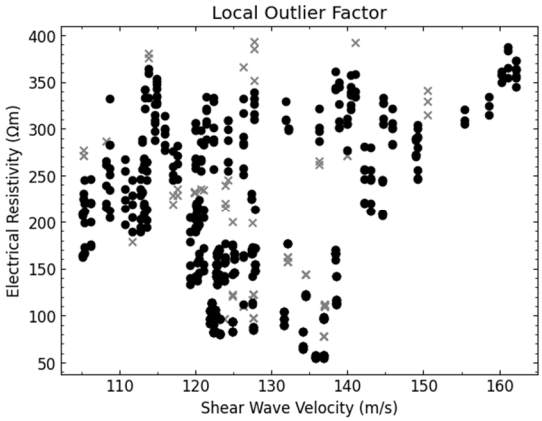


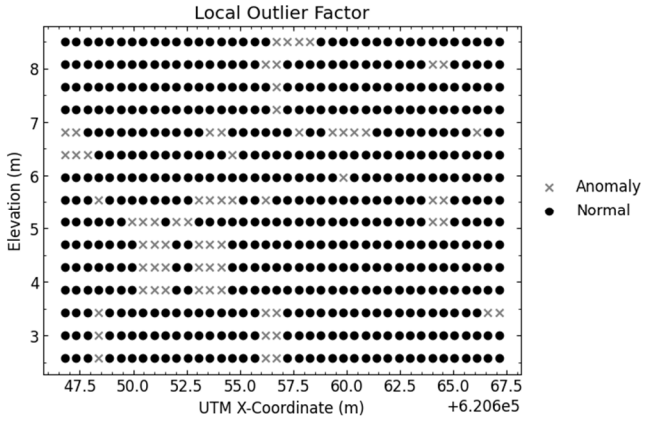


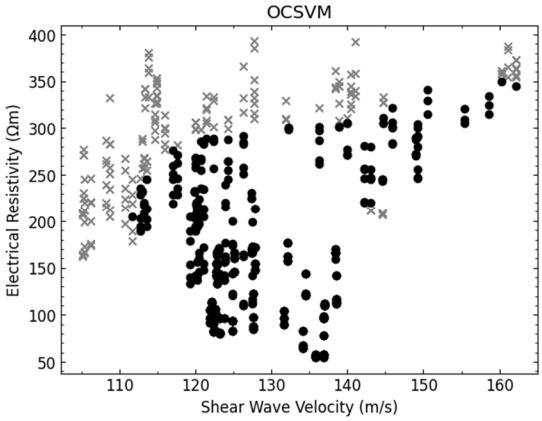
Profile: fwt_site1

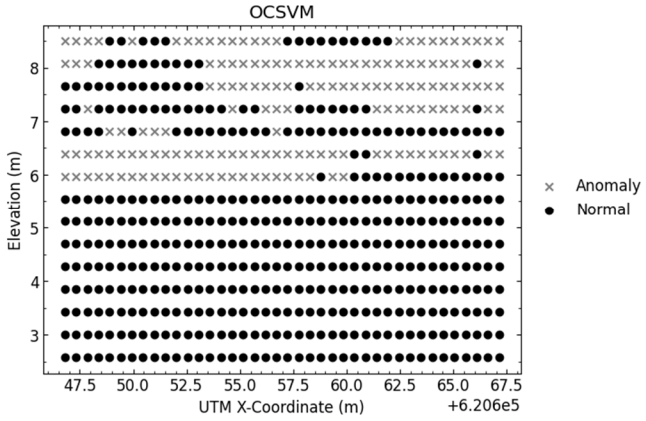




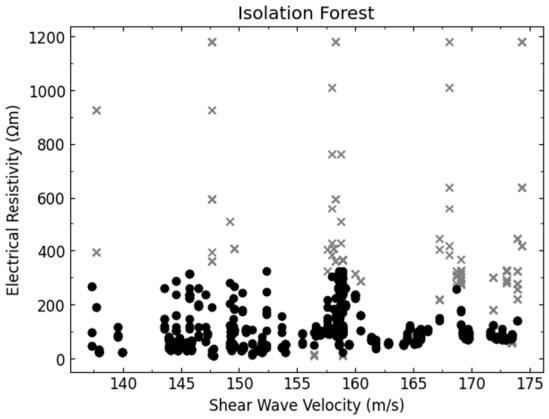


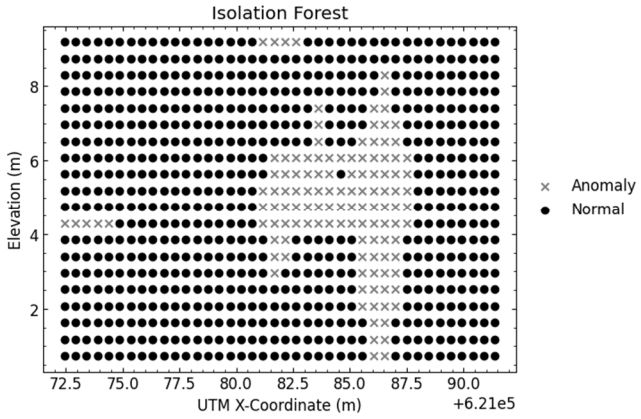


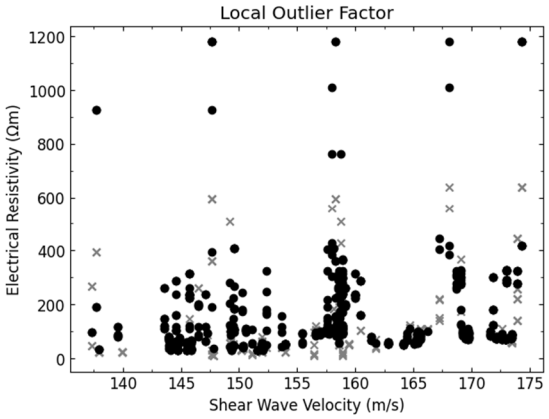


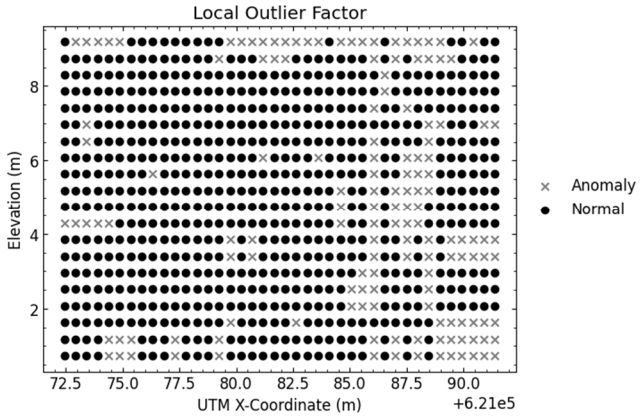


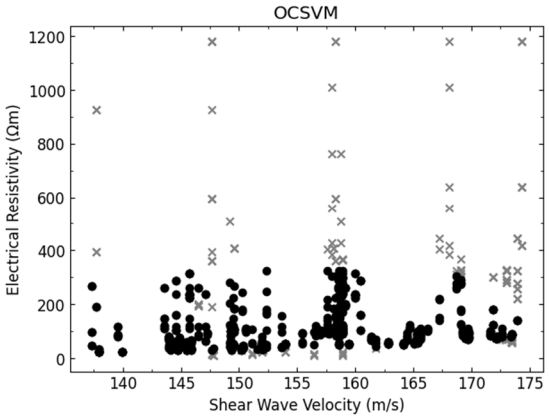
Profile: sc3_site1

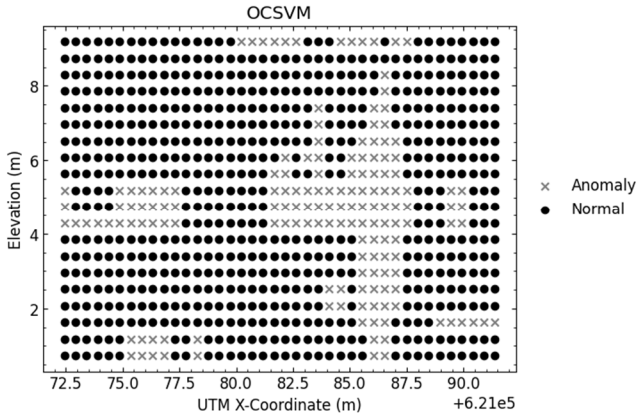




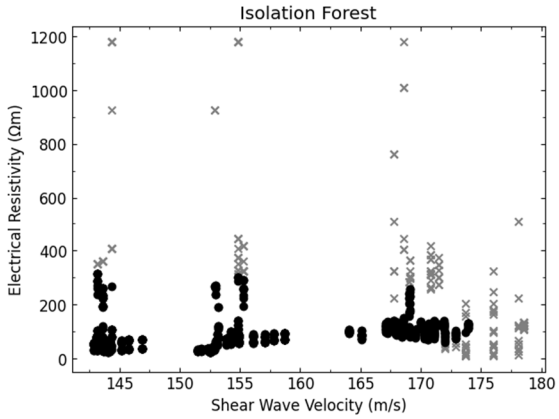


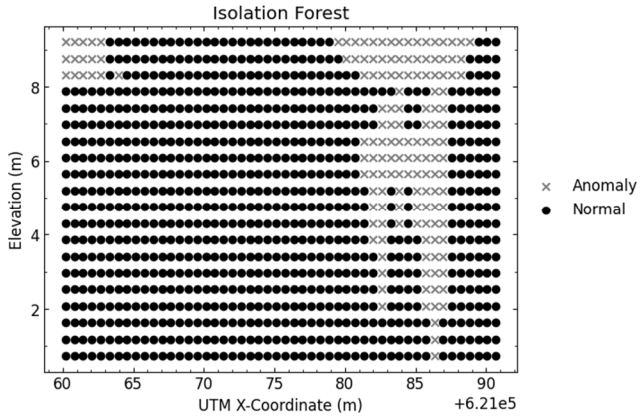


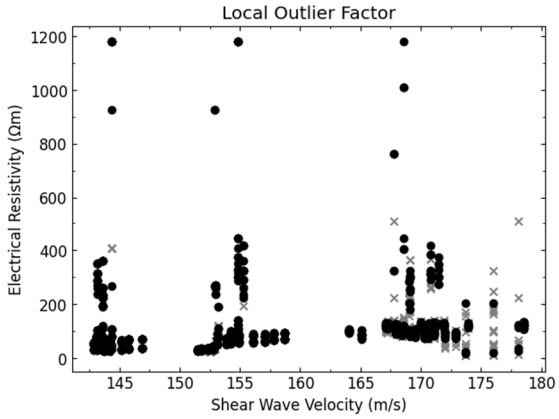


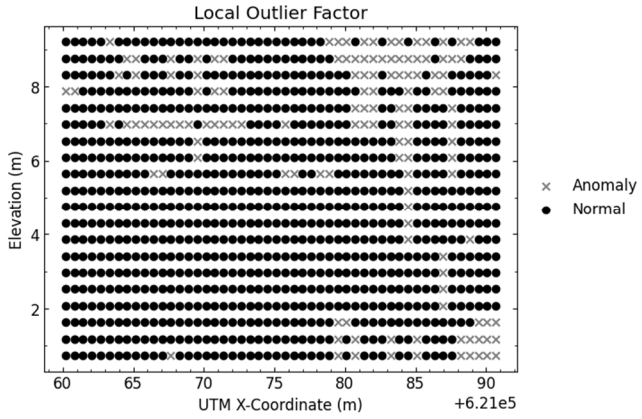


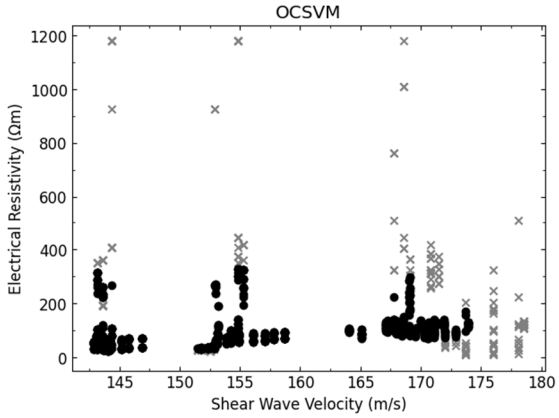
Profile: sc5_site1

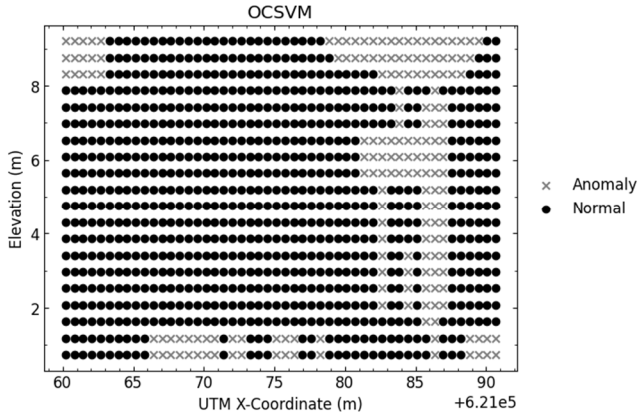




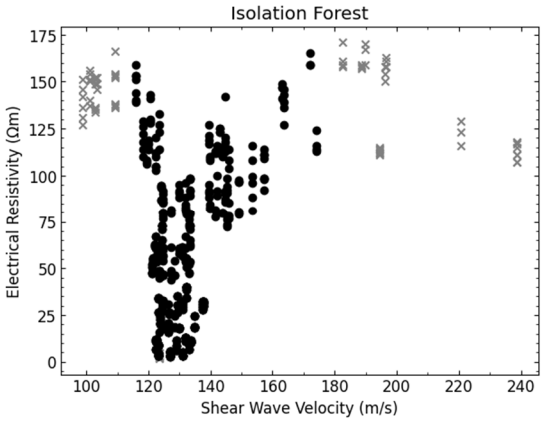


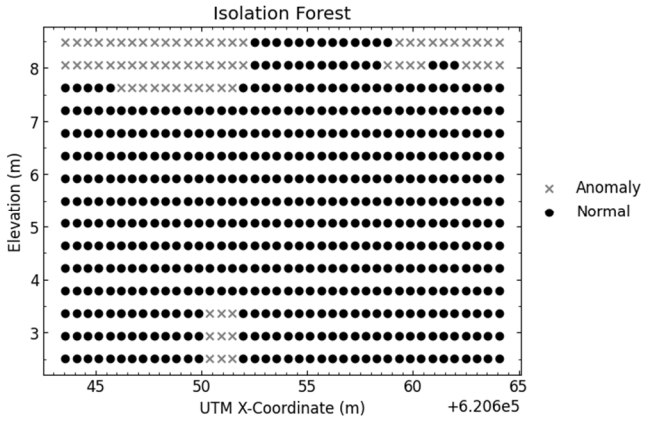


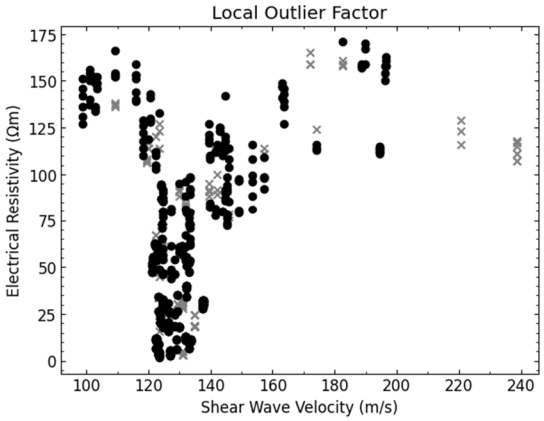


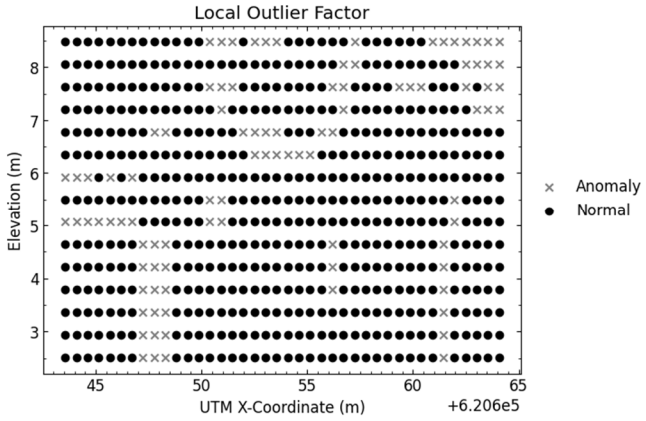


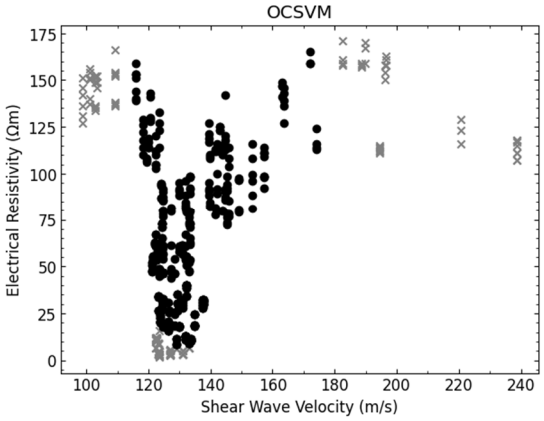
Profile: swt_site1

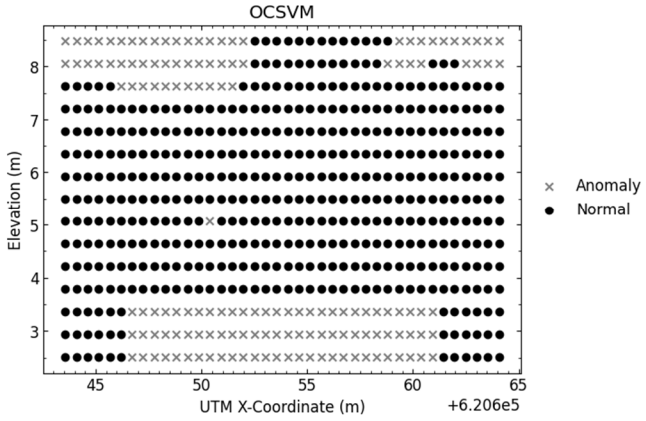






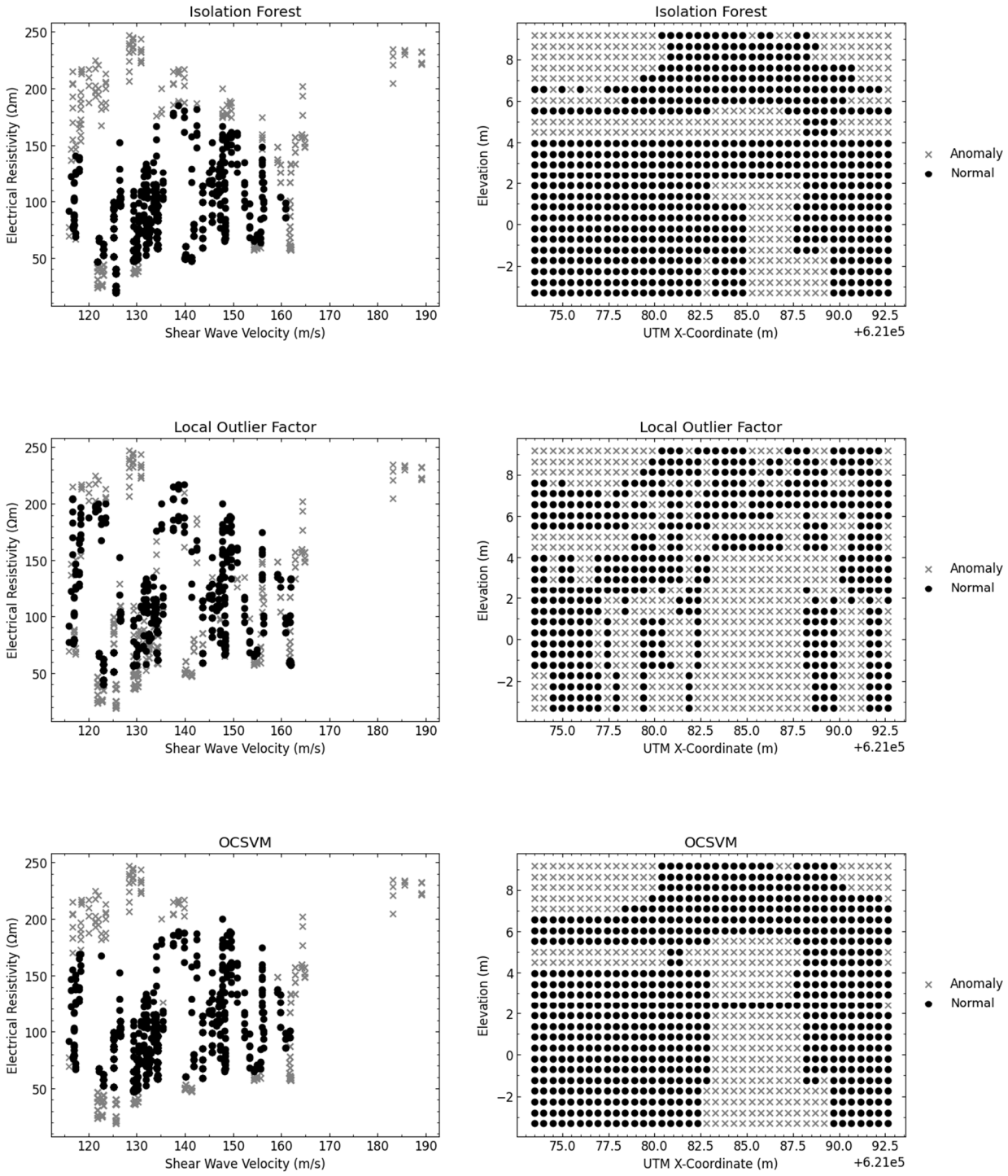




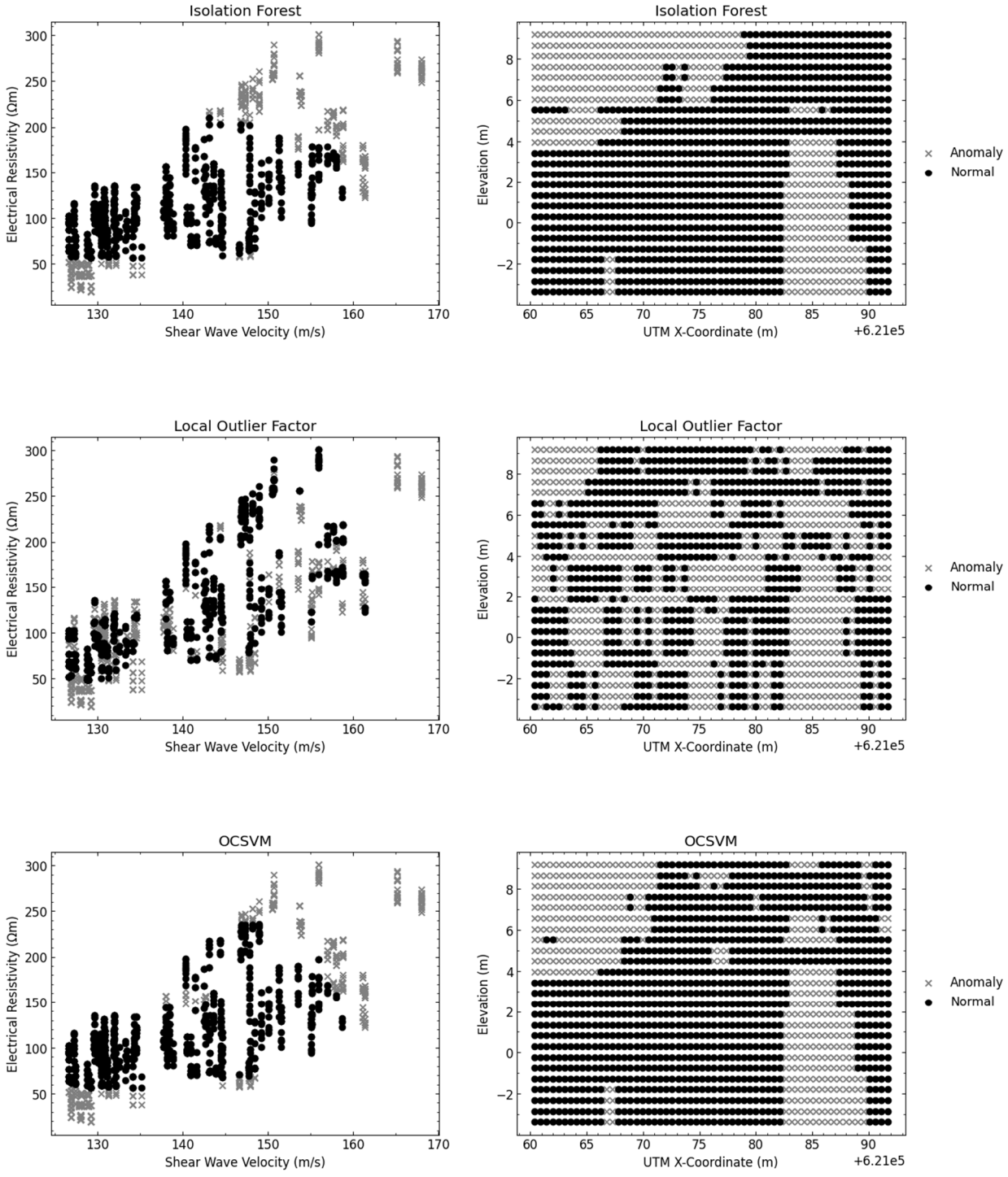


Fit: Site 1 Compilation Dataframe
Predicted: Individual Profiles

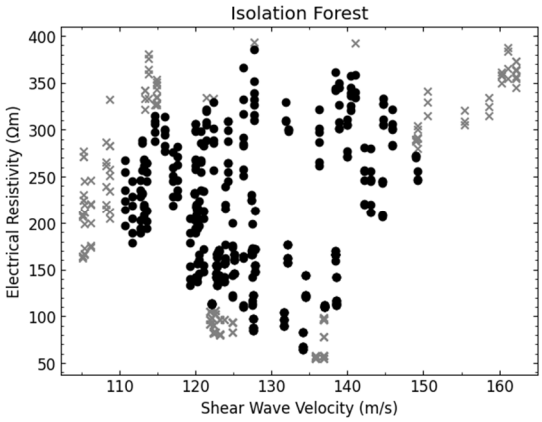
Profile: fc3_site1

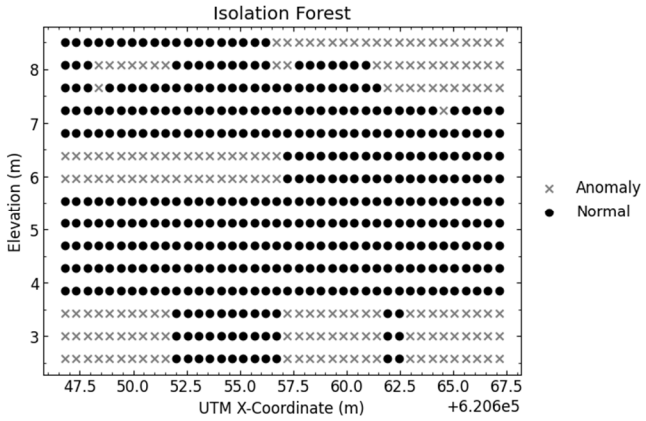


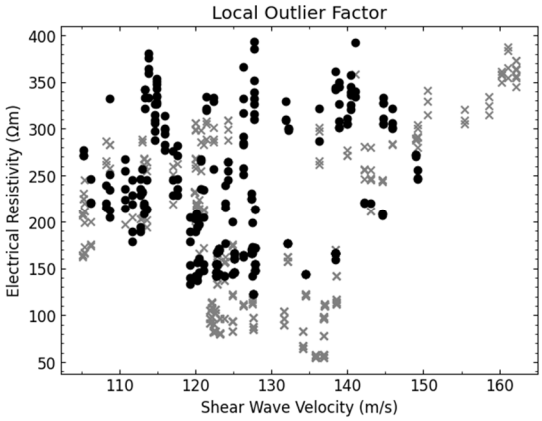
Profile: fc5_site1

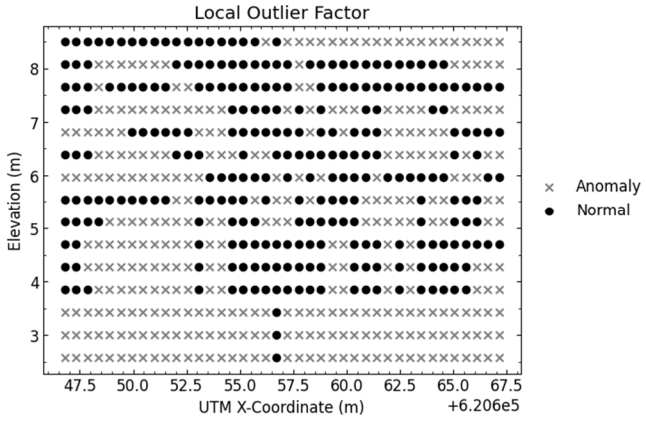


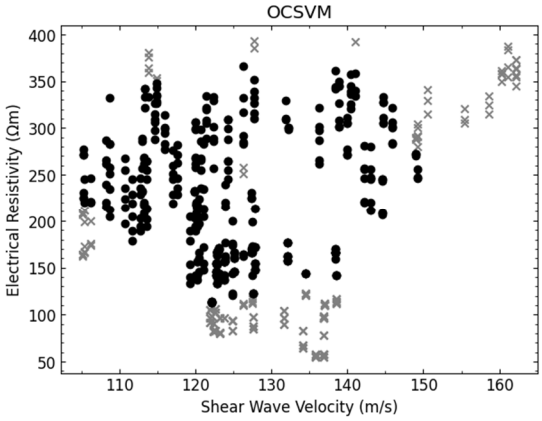
Profile: fwt_site1

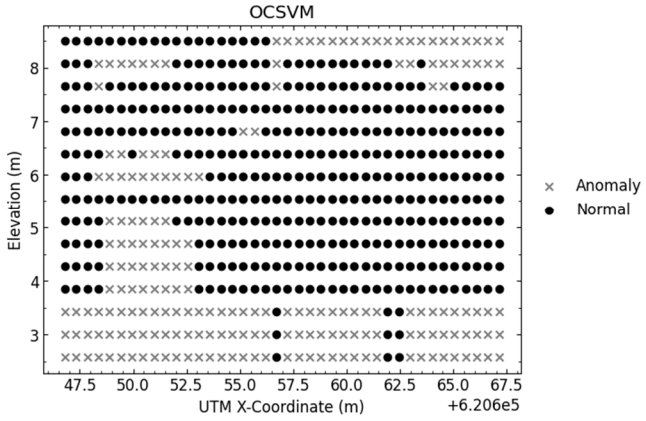




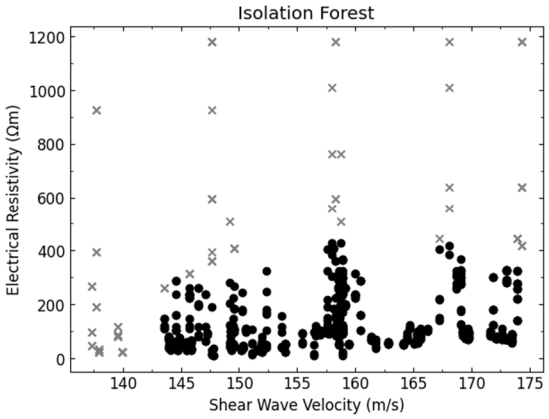


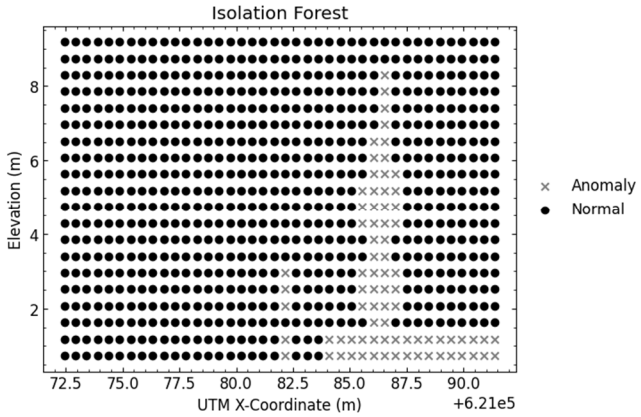


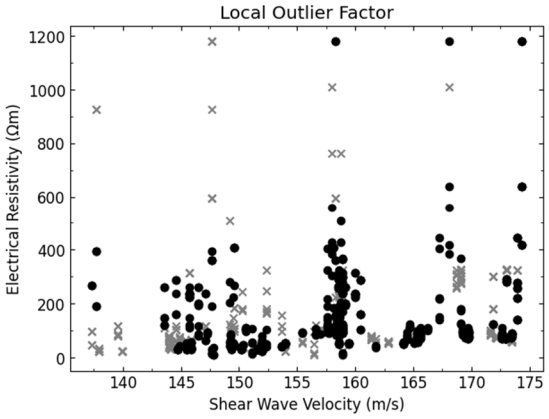


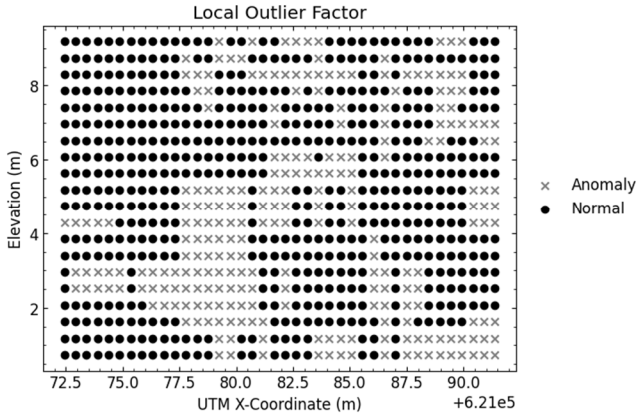


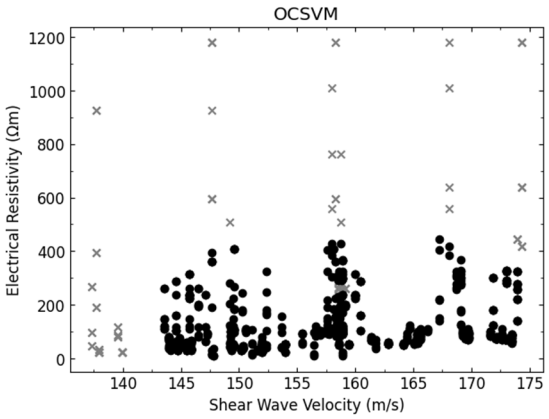
Profile: sc3_site1

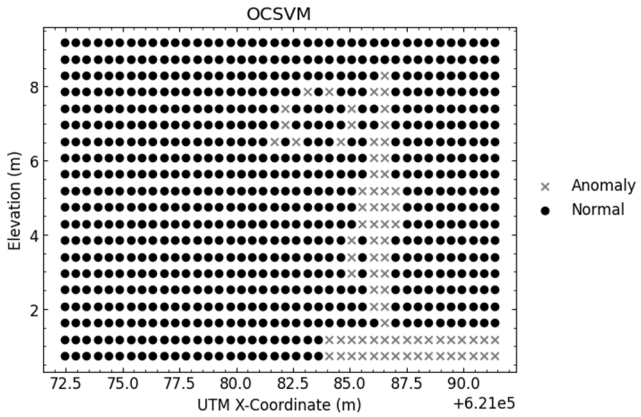




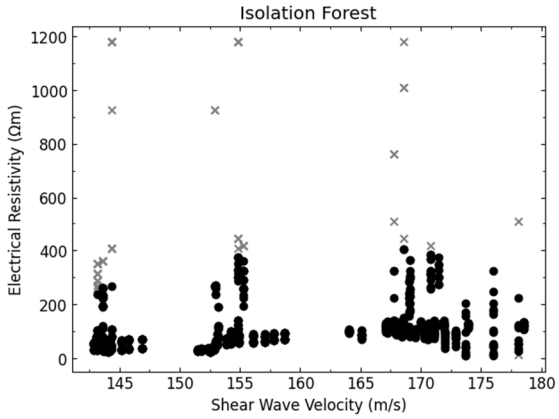


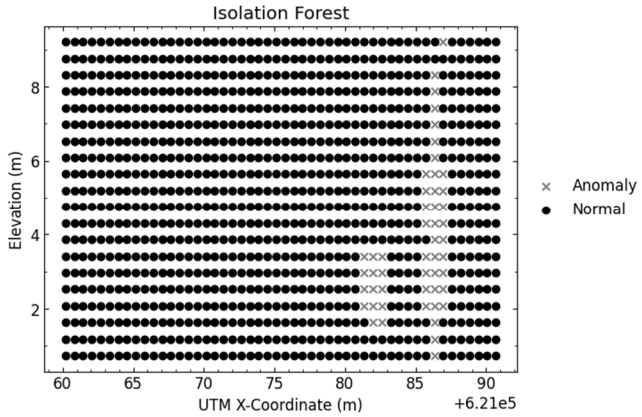


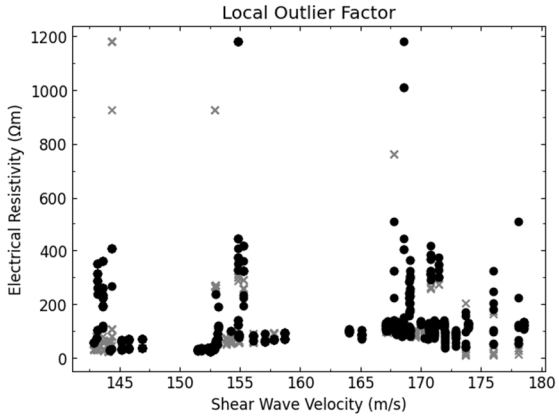


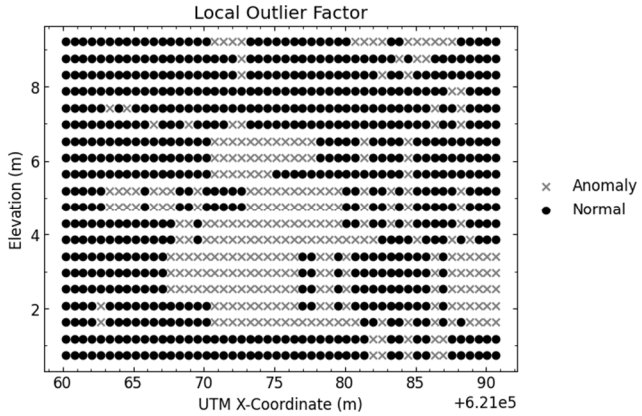


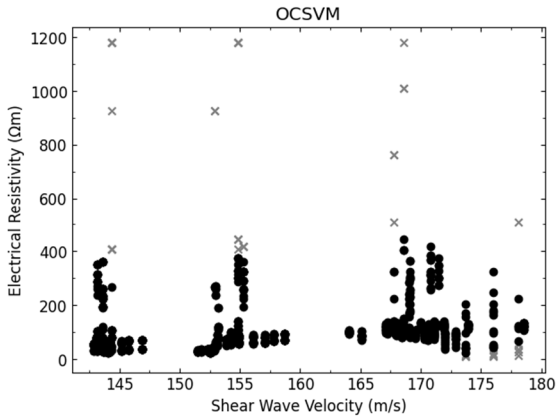
Profile: sc5_site1

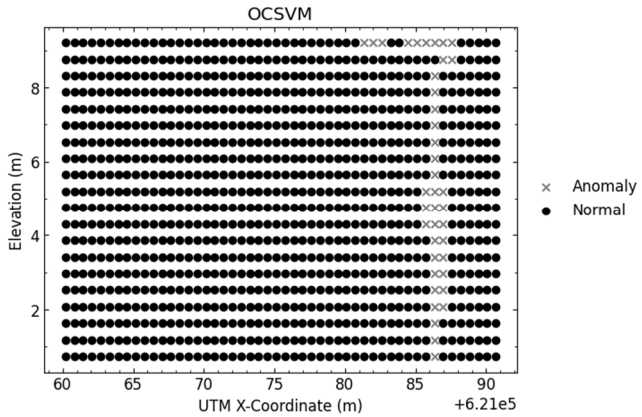




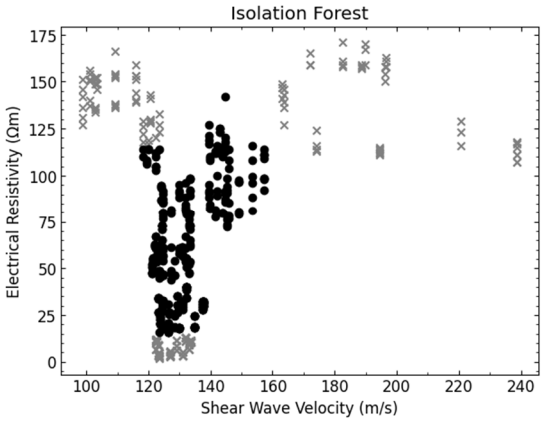


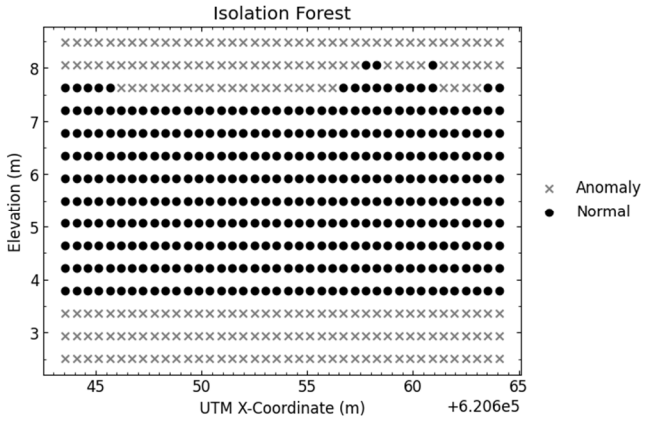


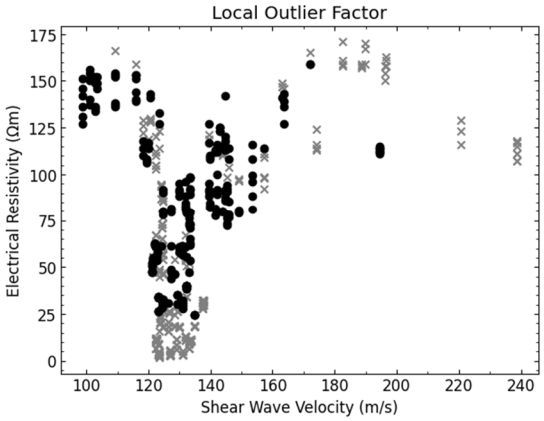


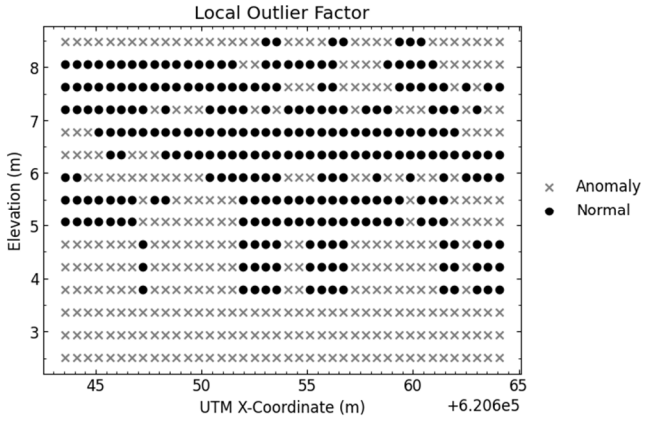


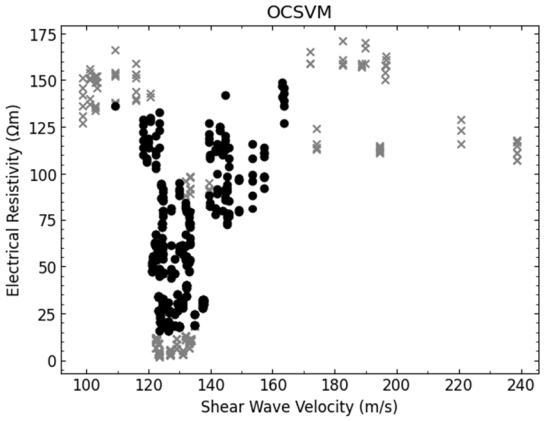
Profile: swt_site1

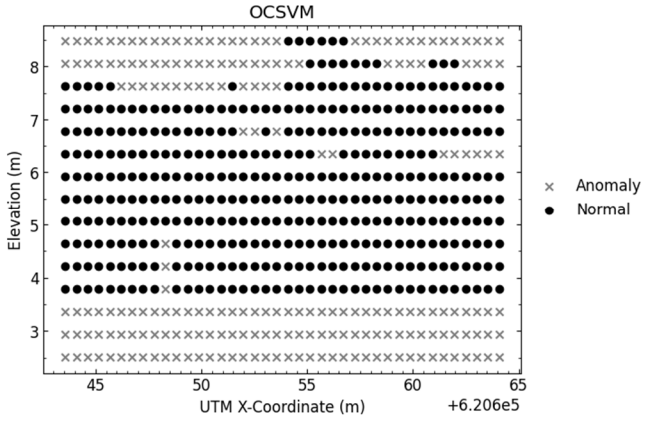






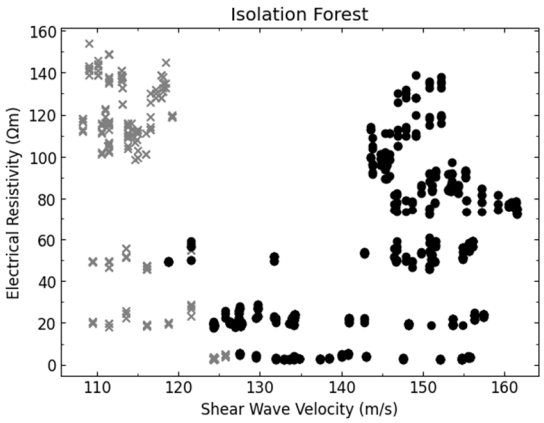


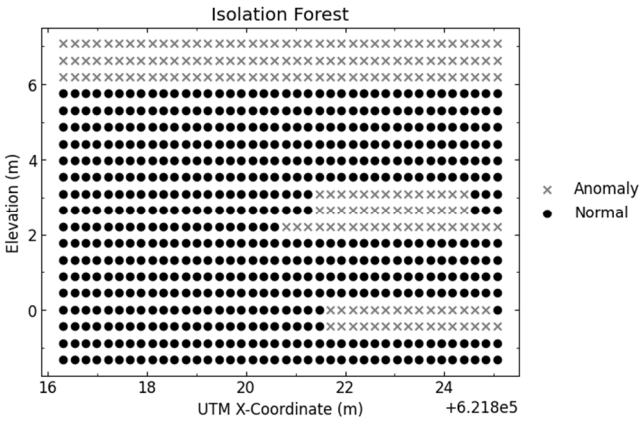


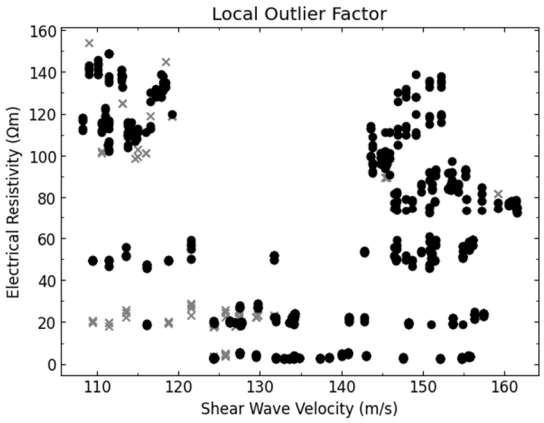


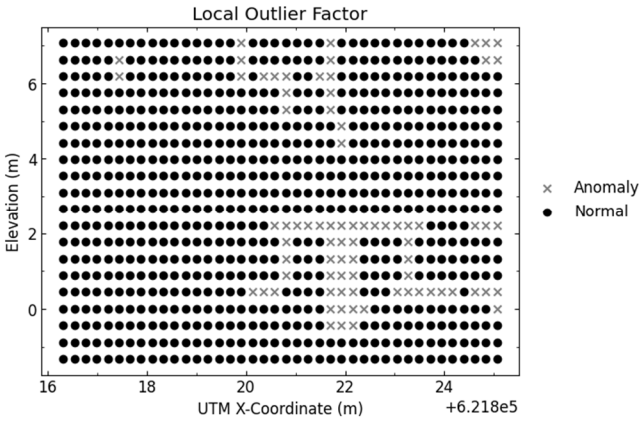
Fit: Site 2 Compilation Dataframe
Predicted: Site 2 Compilation Dataframe

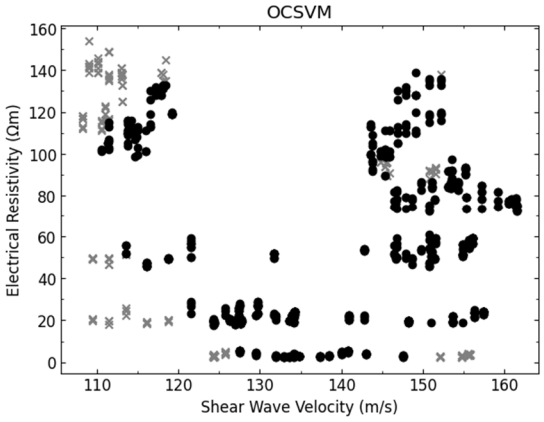
Profile: fc_site2

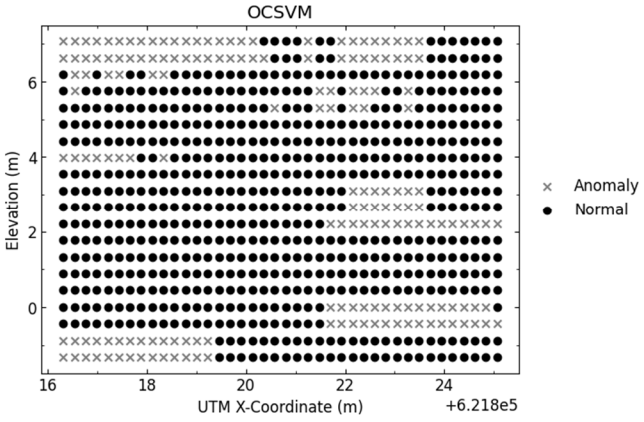




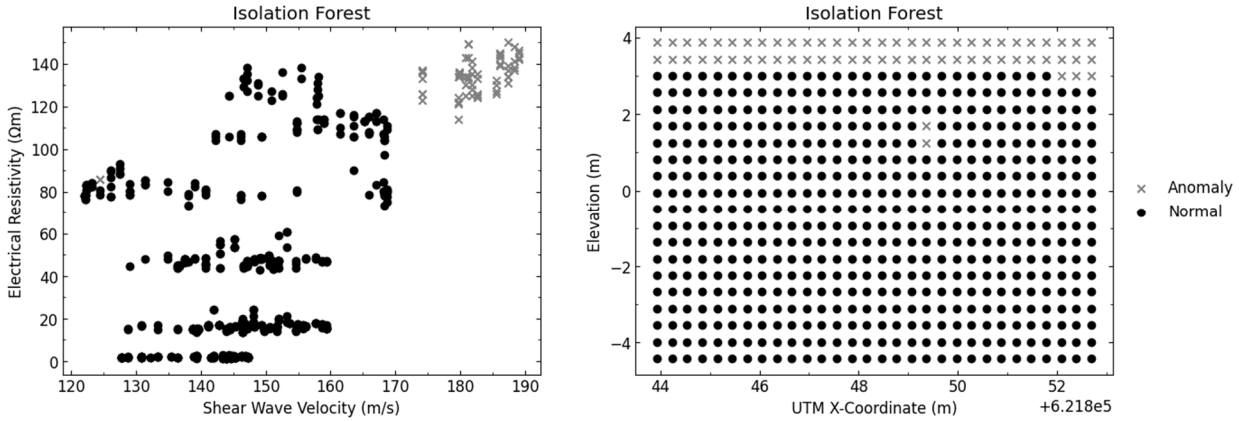


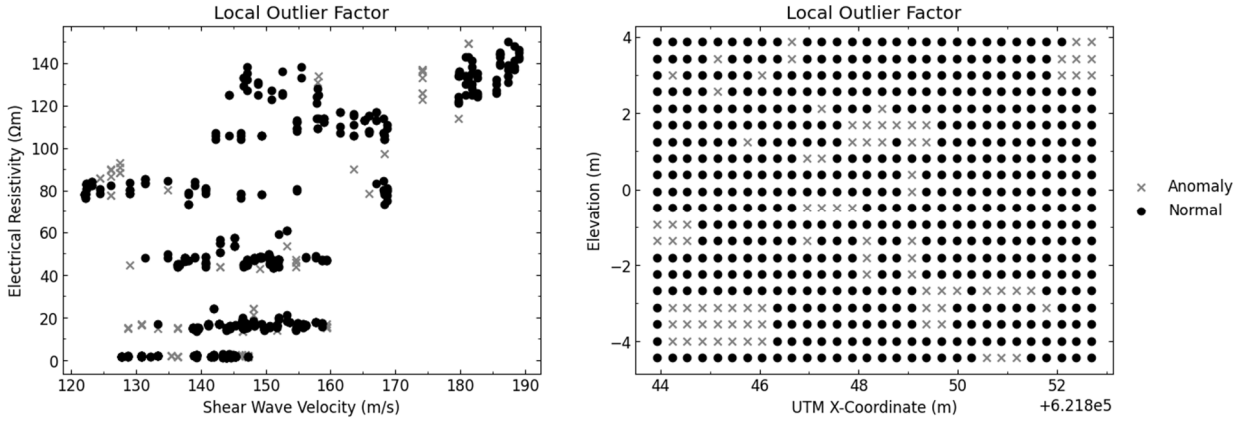


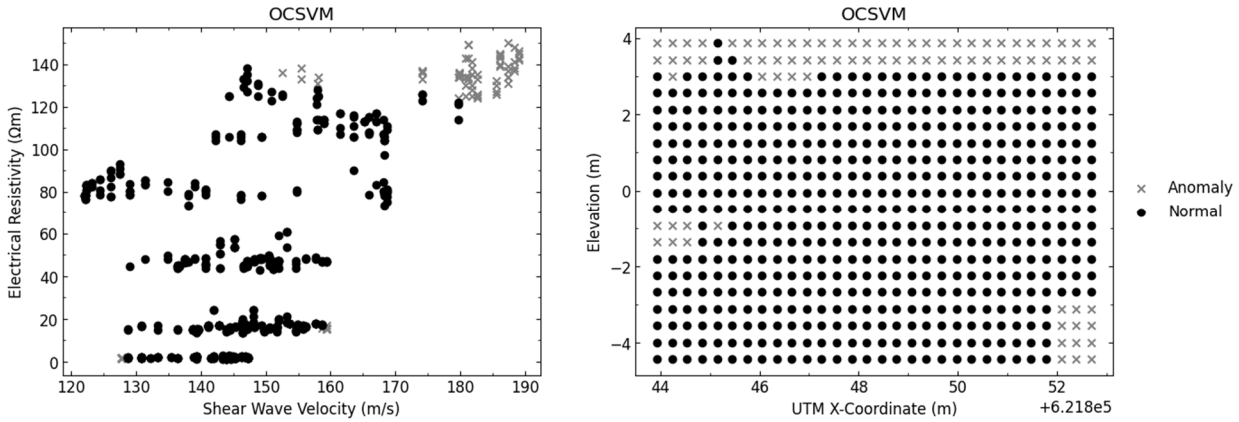




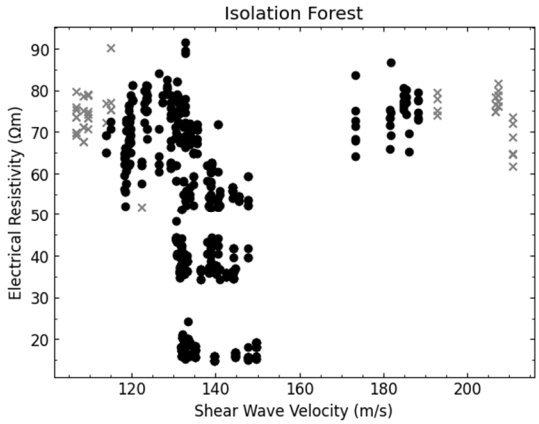
Profile: fdt_site2

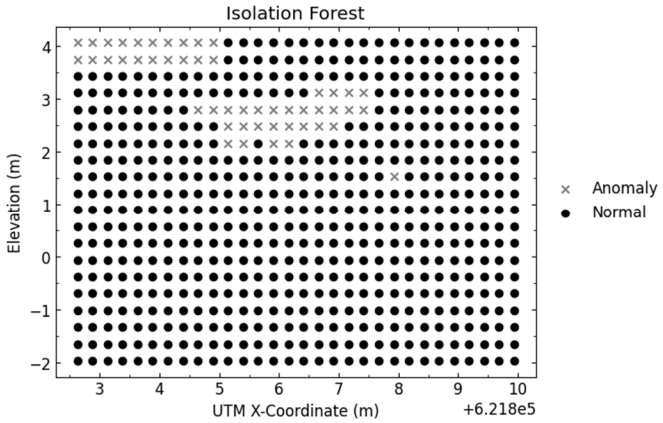


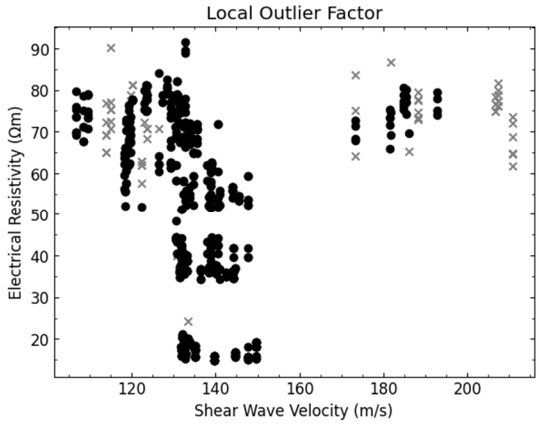


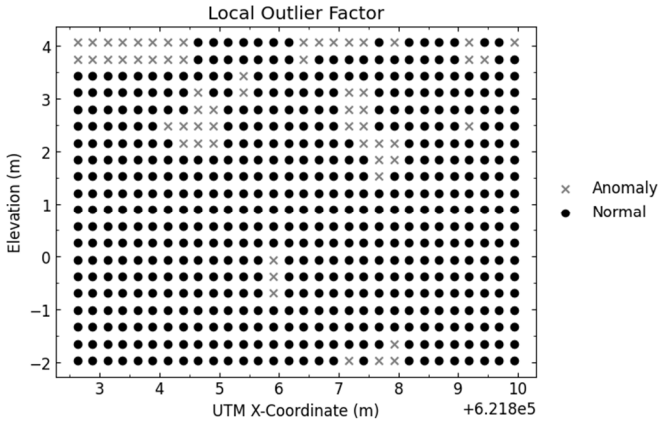


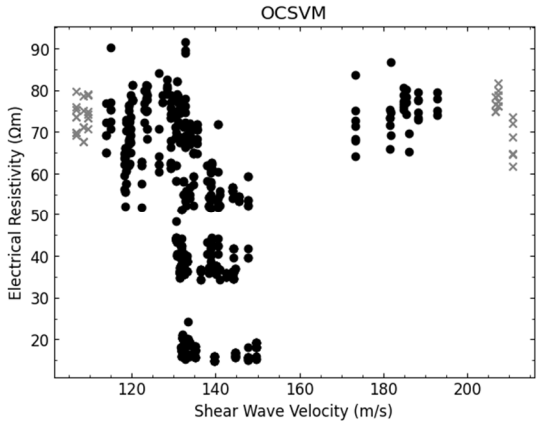
Profile: fwt_site2

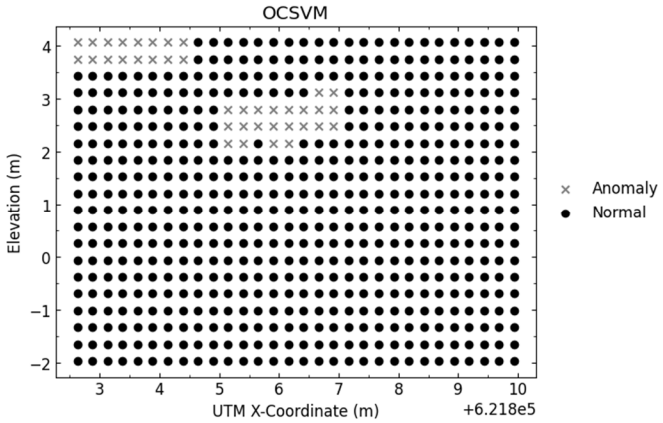




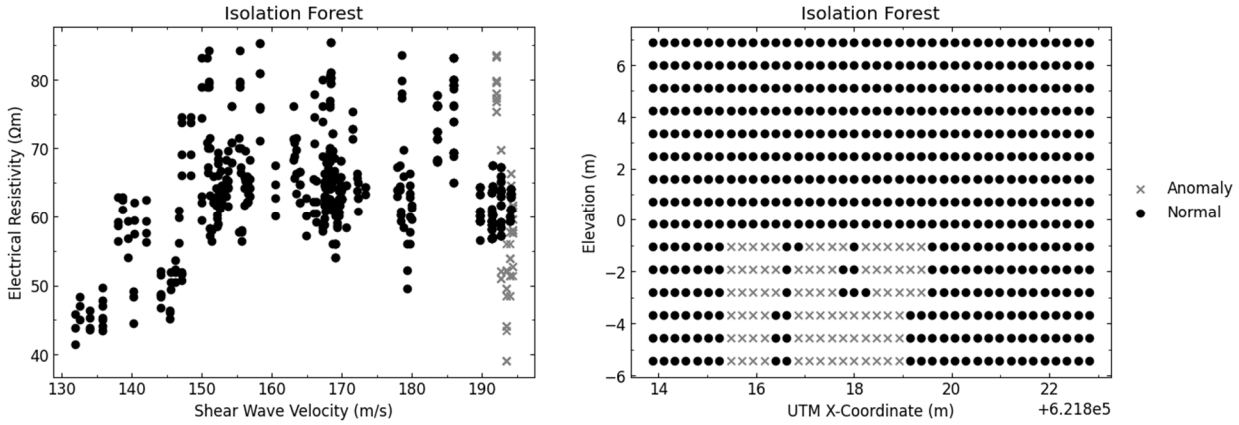


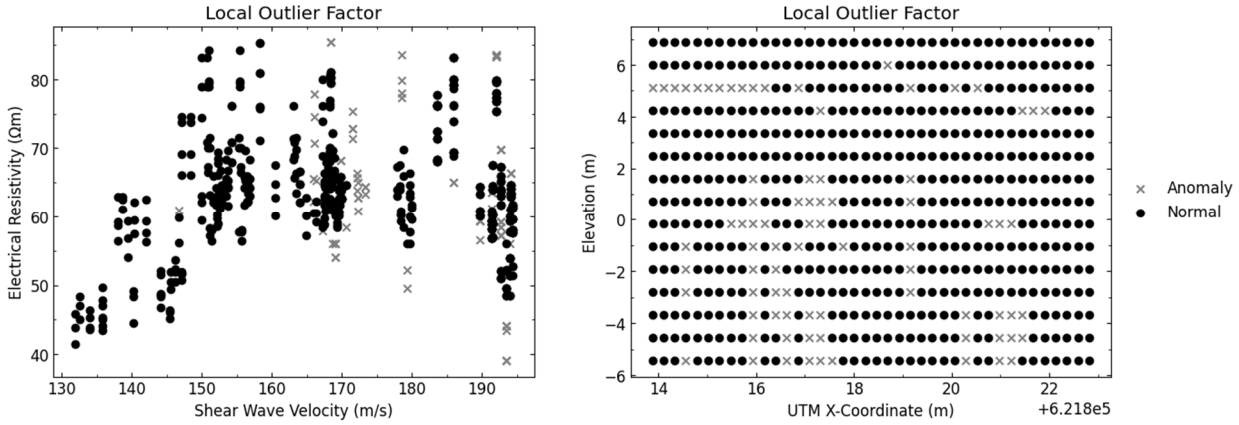


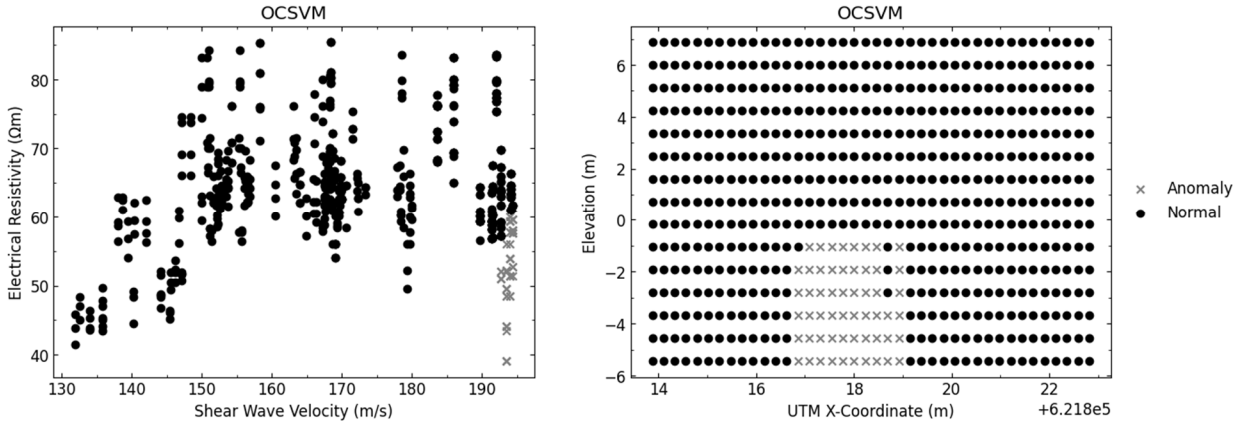




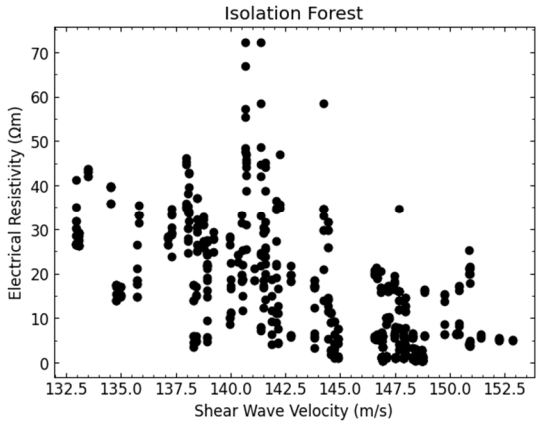
Profile: sc_site2

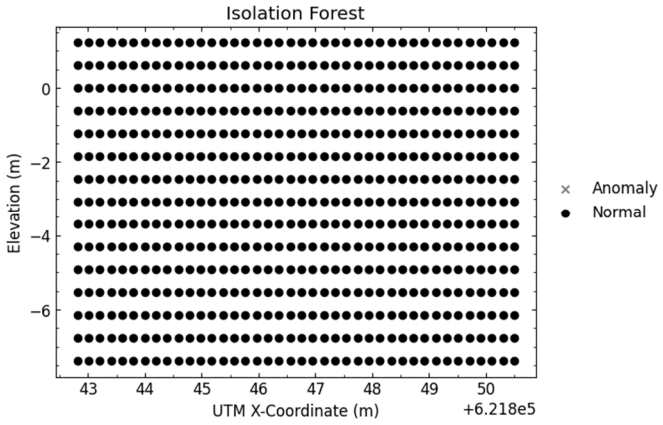


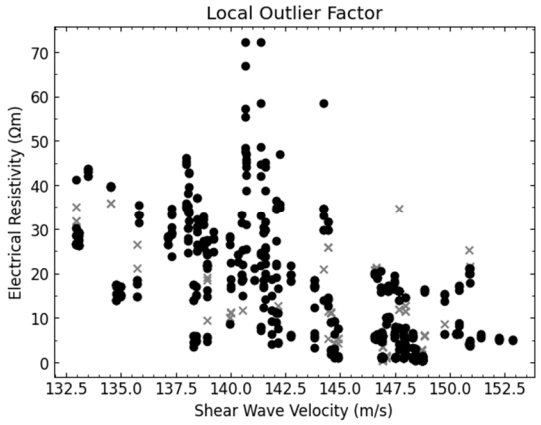


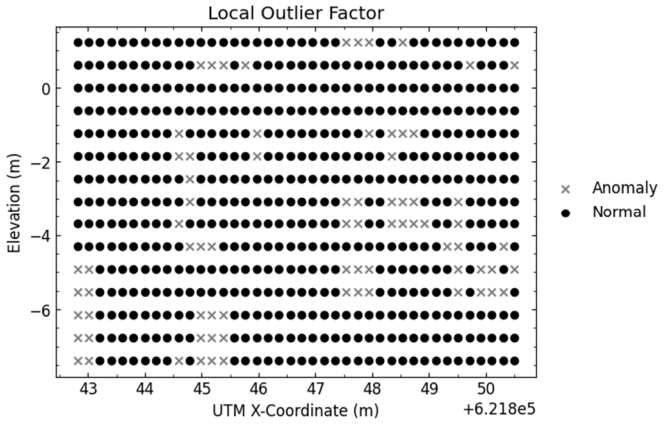


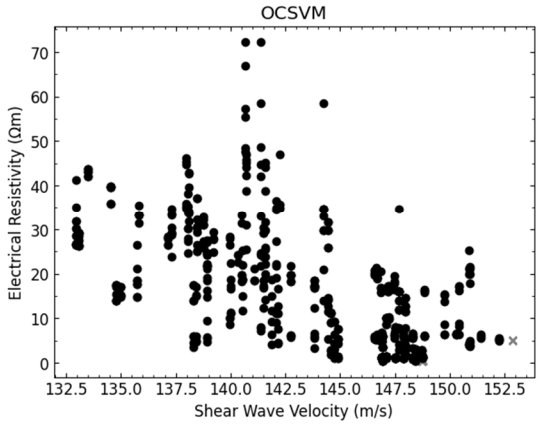
Profile: sdt_site2

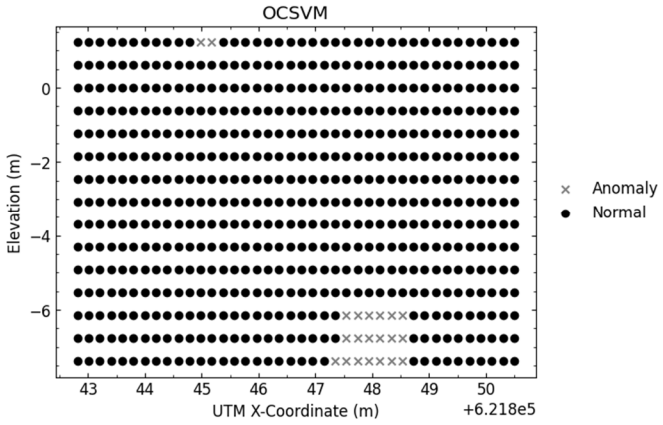




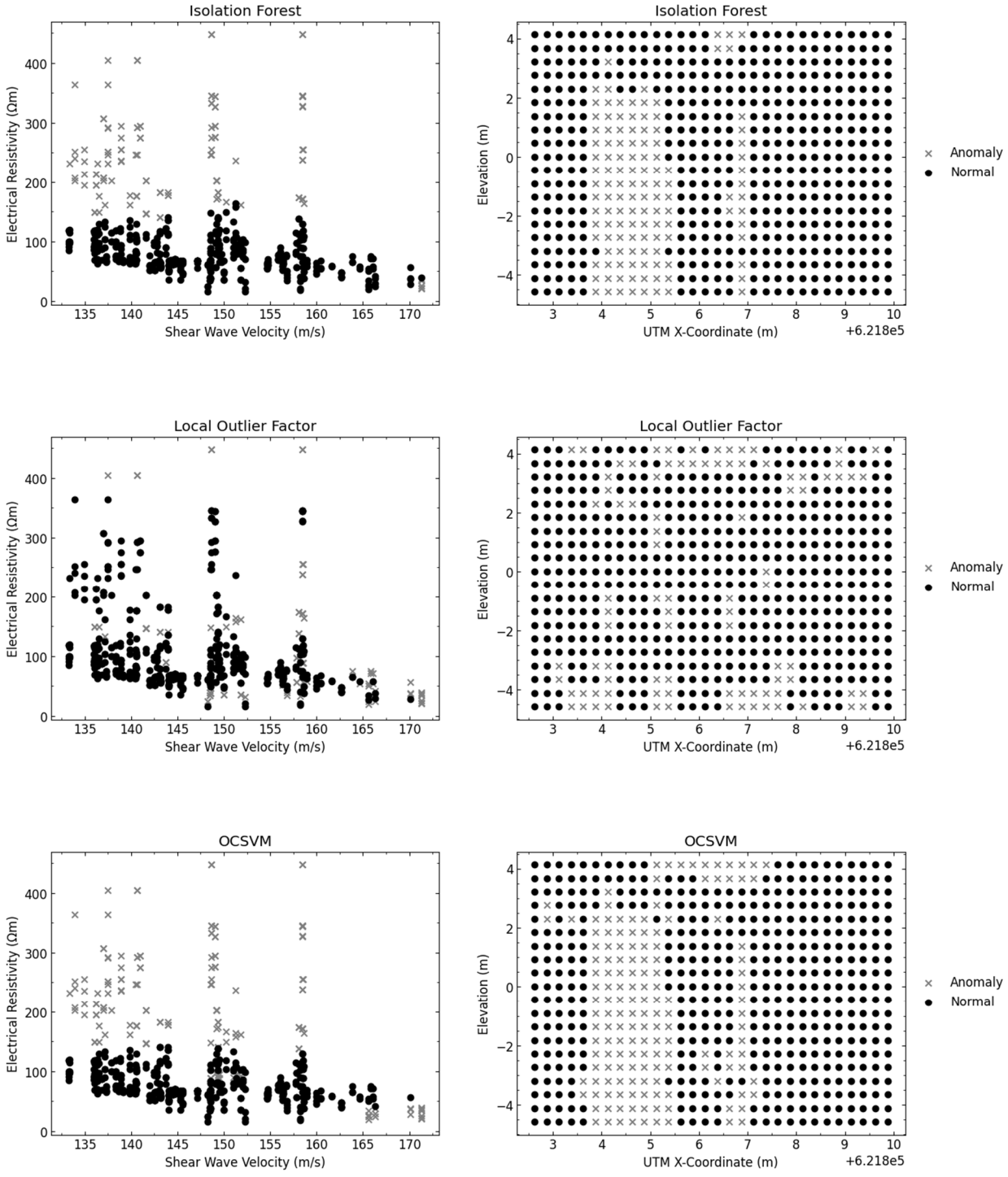






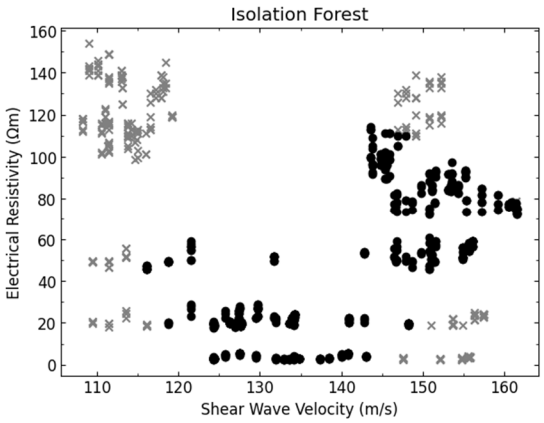


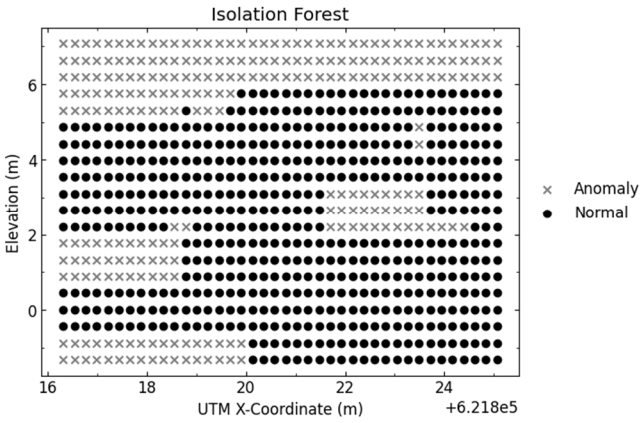
Profile: swt_site2

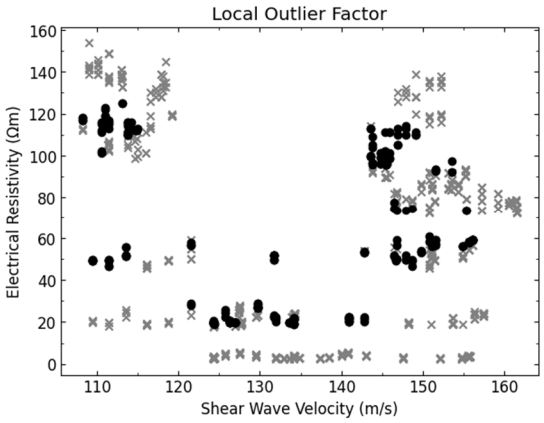


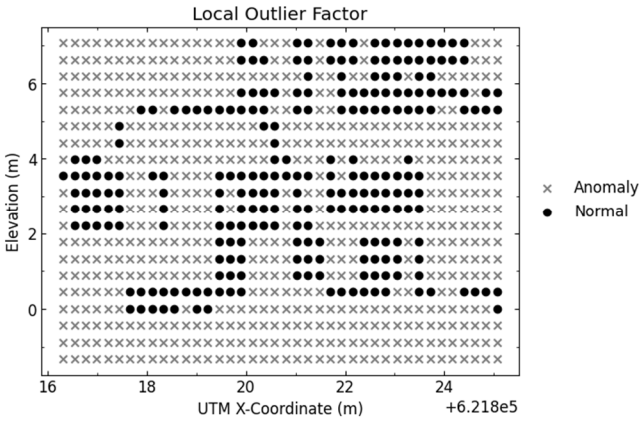
Fit: Site 2 Compilation Dataframe
Predicted: Individual Profiles

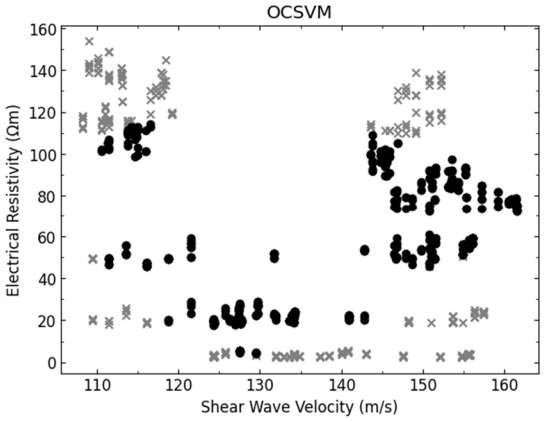
Profile: fc_site2

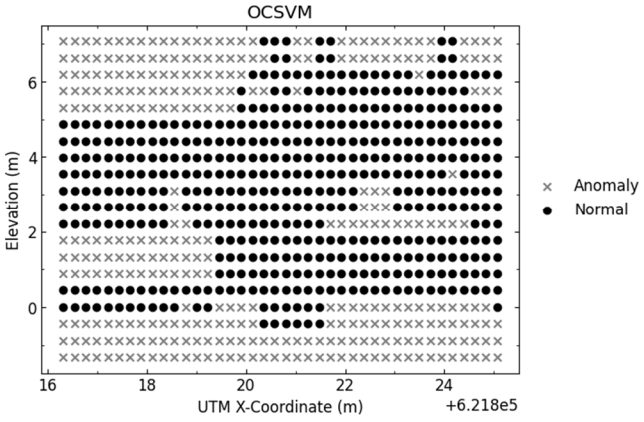




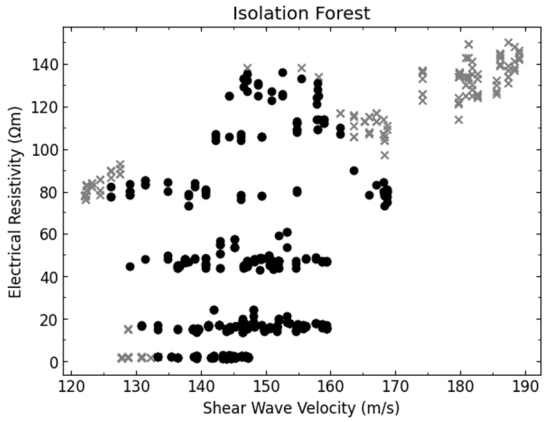


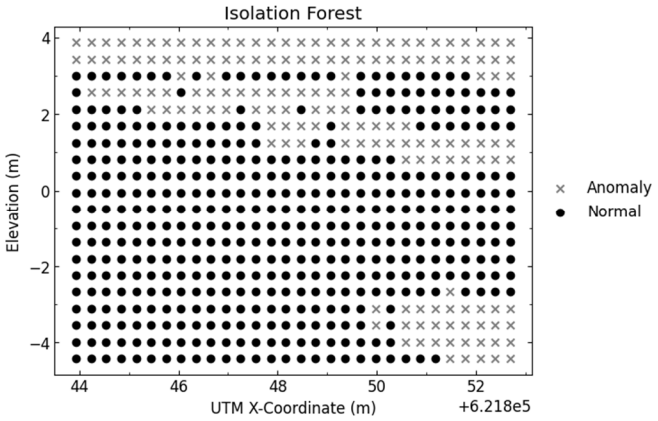


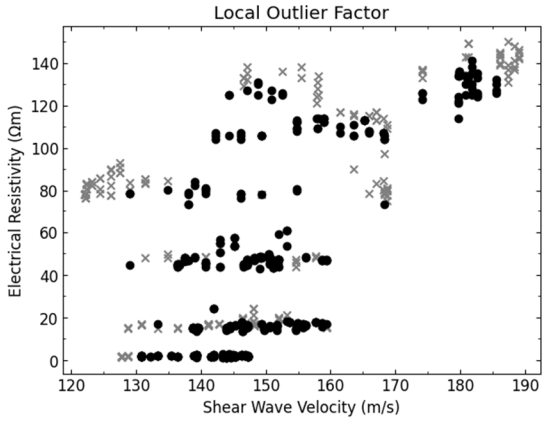


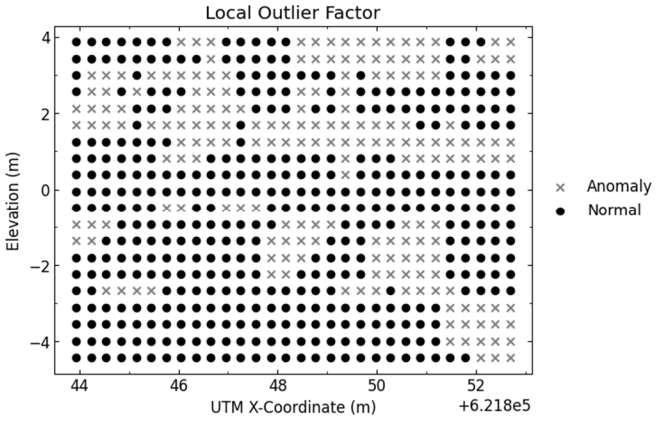


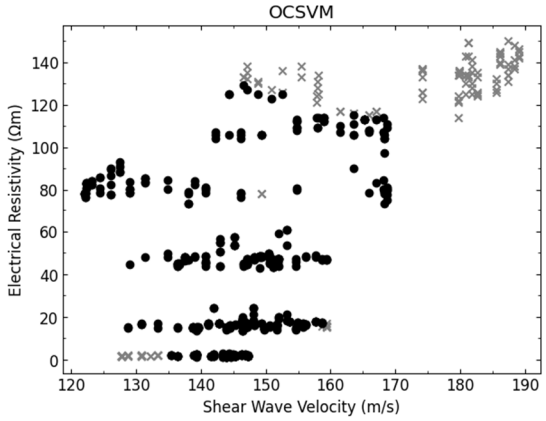
Profile: fdt_site2

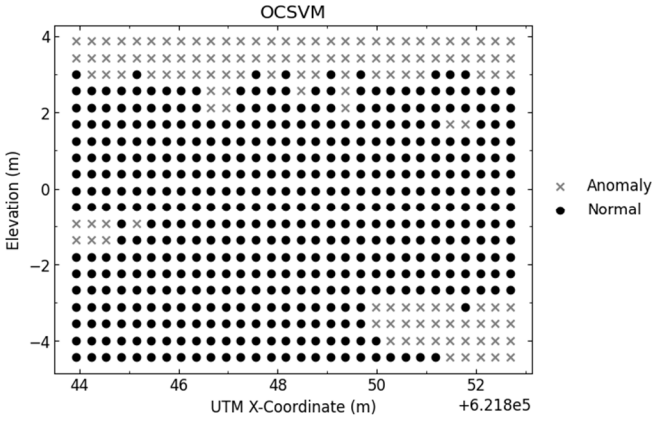




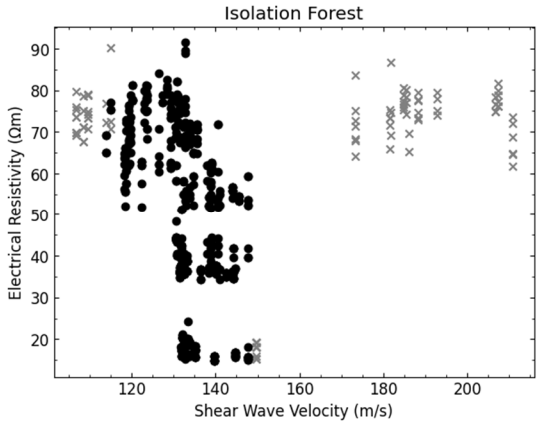


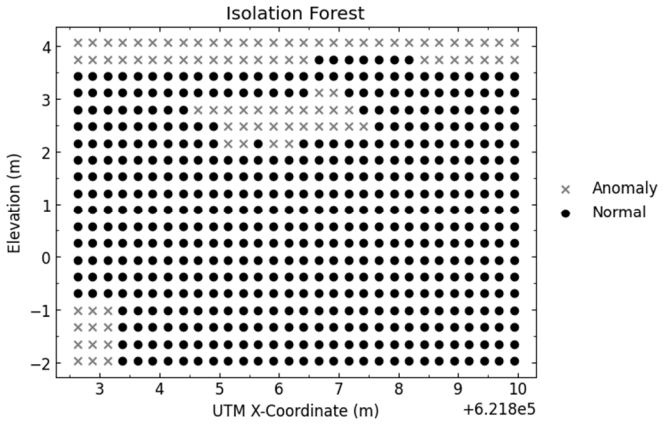


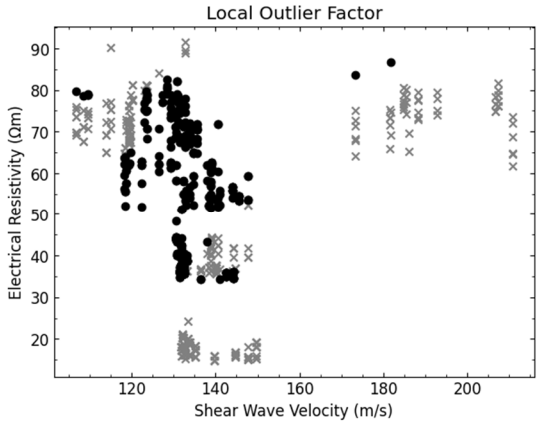


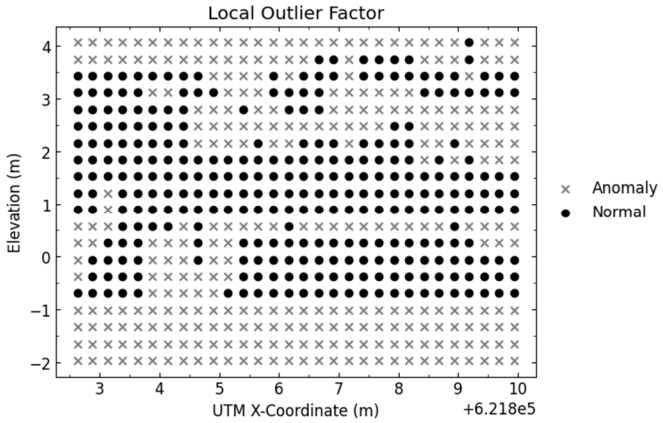


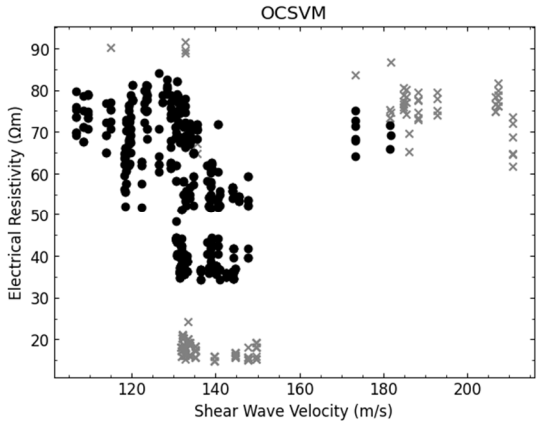
Profile: fwt_site2

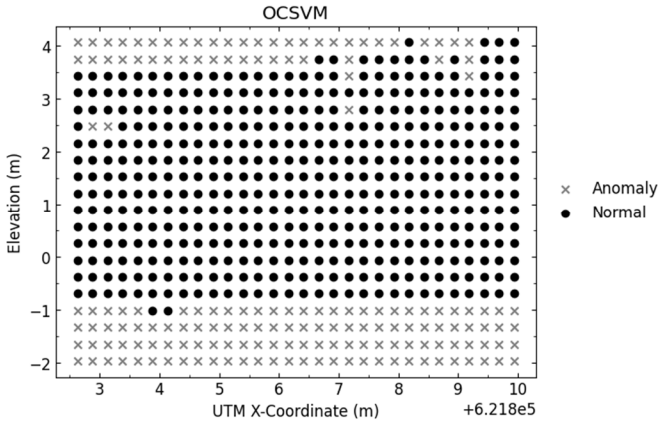




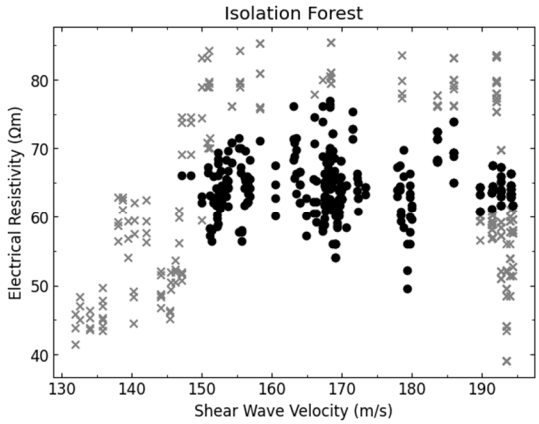


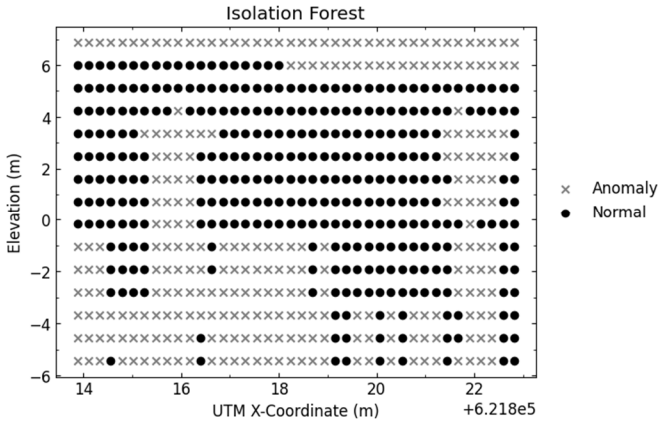


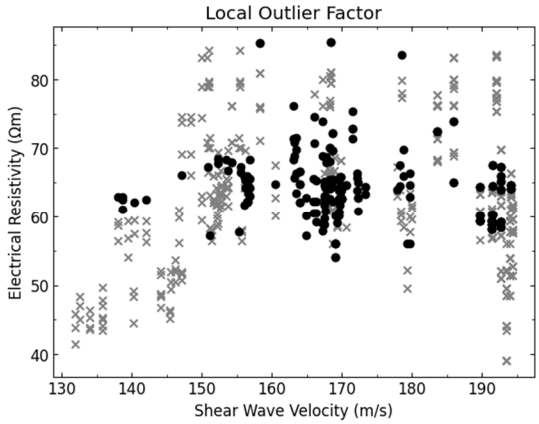


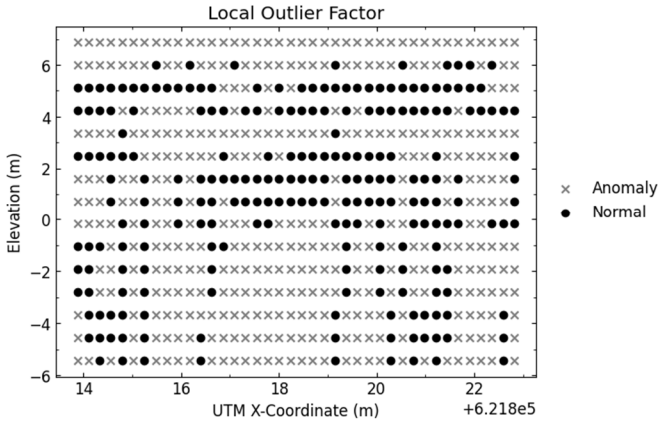


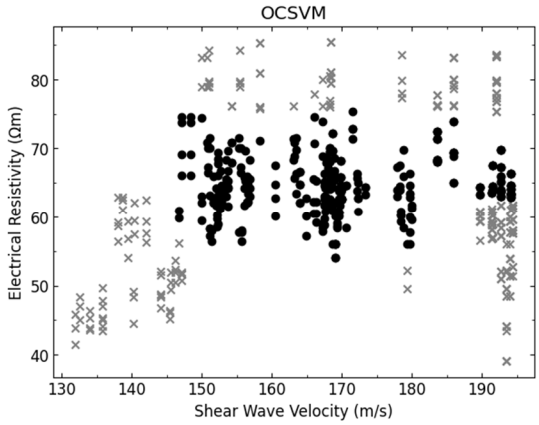
Profile: sc_site2

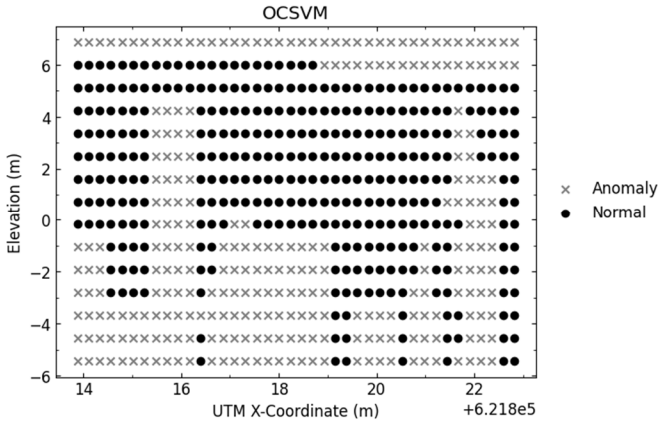




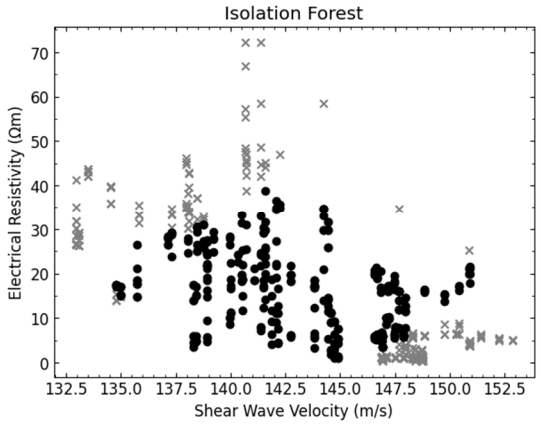


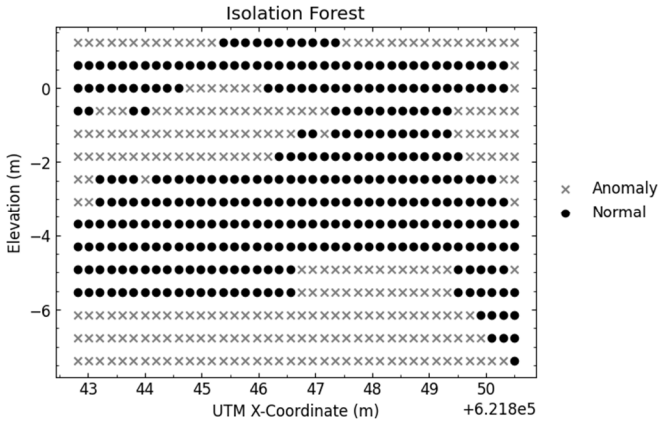


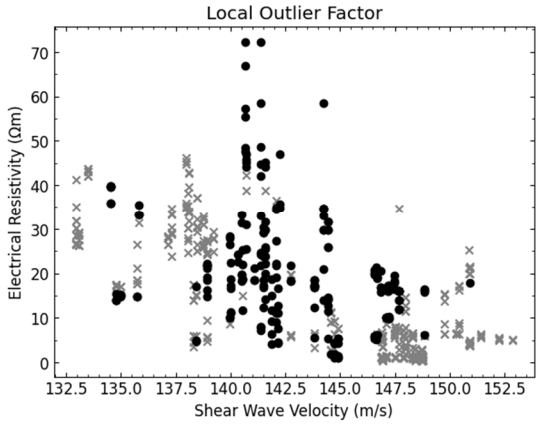


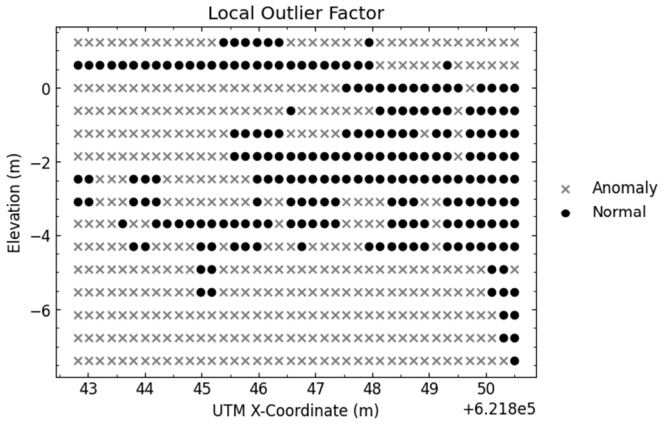


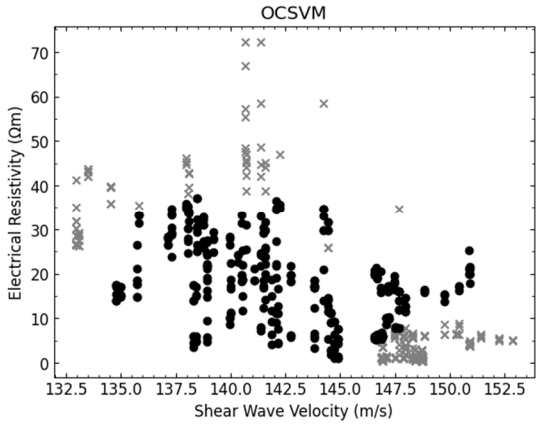
Profile: sdt_site2

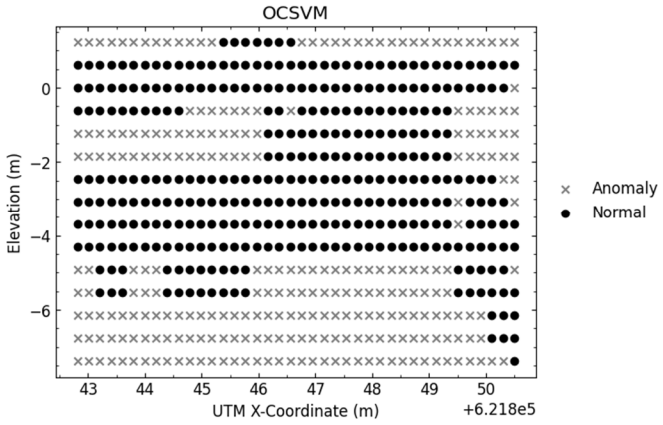




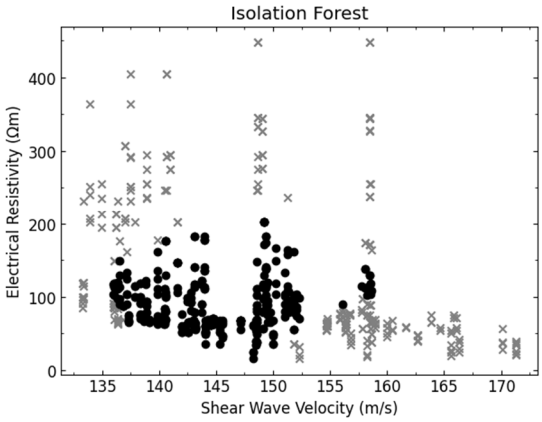


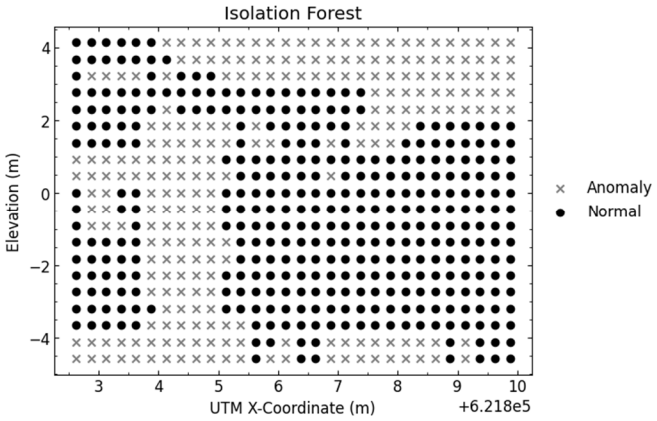


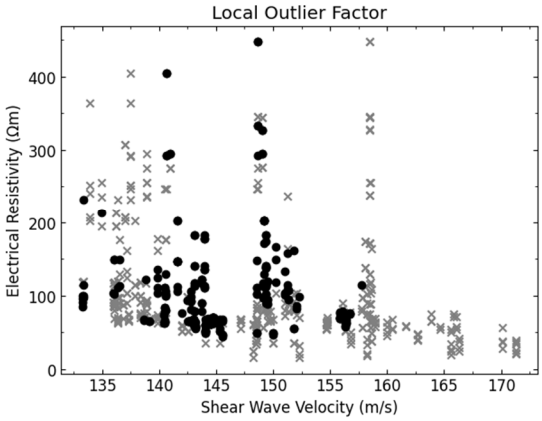


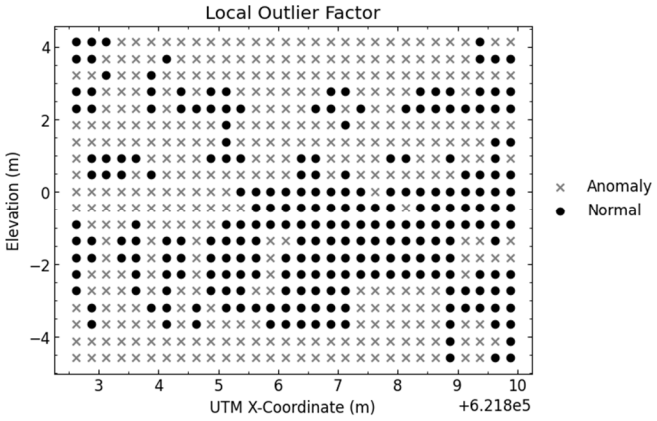


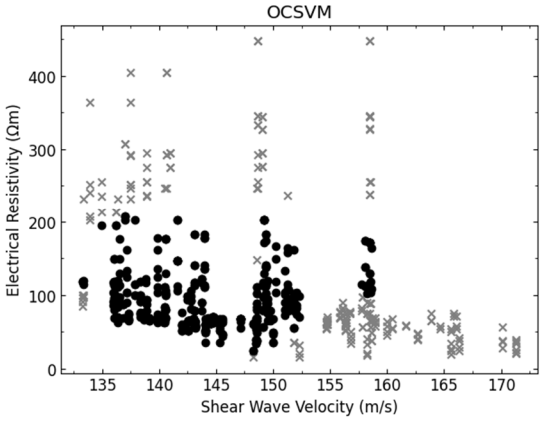
Profile: swt_site2

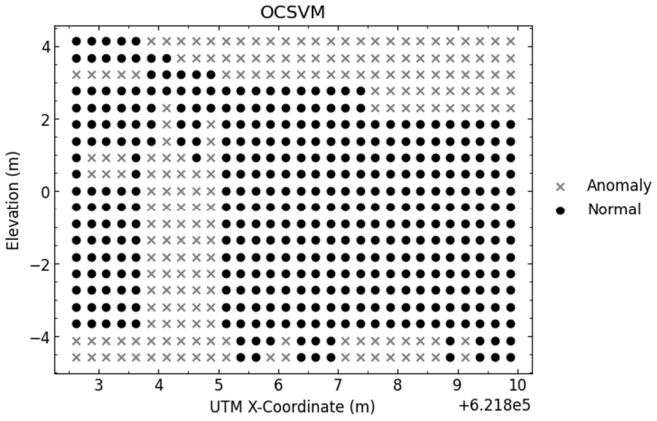






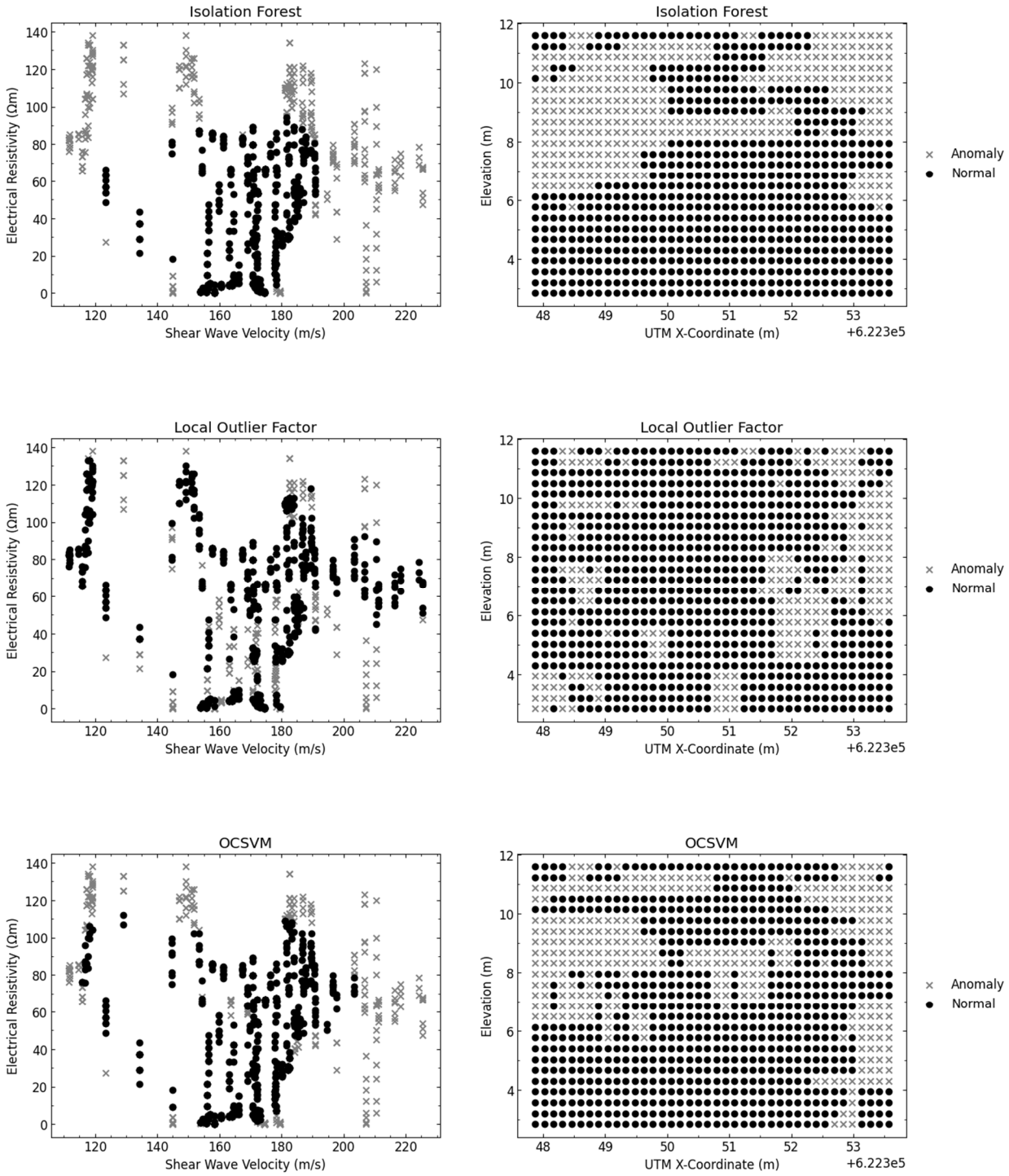




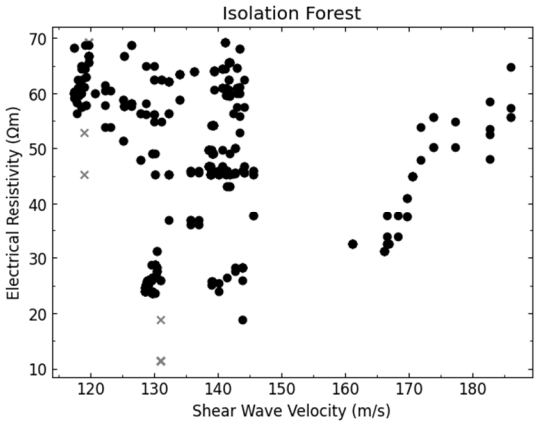


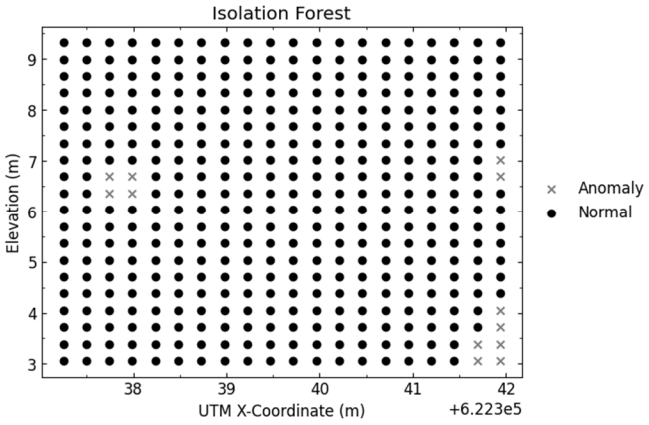
Fit: Site 3 Compilation Dataframe
Predicted: Site 3 Compilation Dataframe

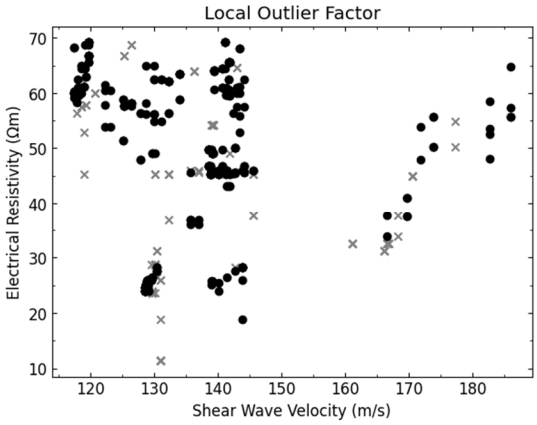
Profile: fc_site3

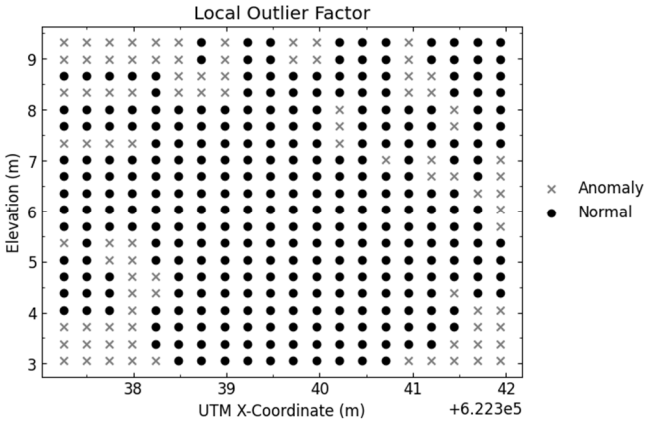


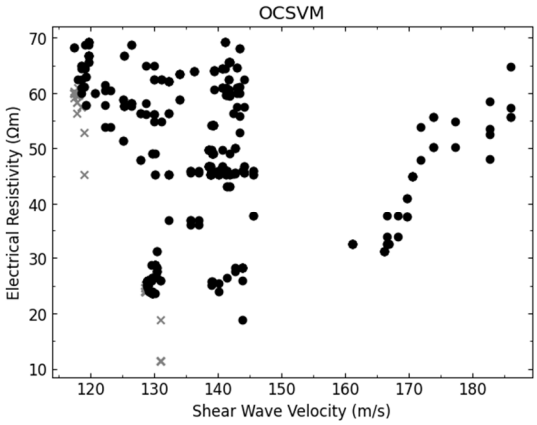
Profile: fwt_site3

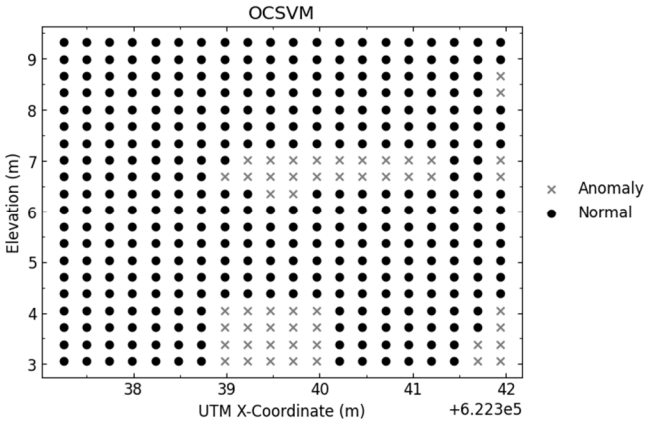




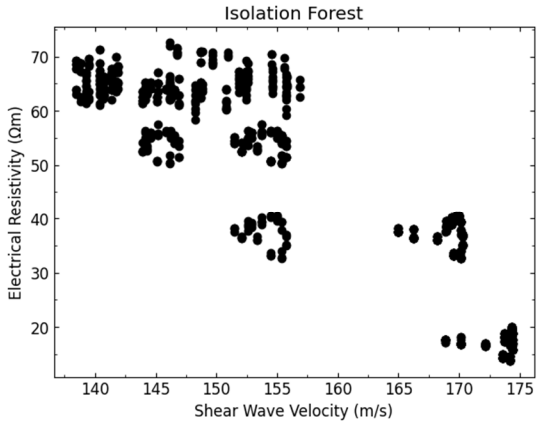


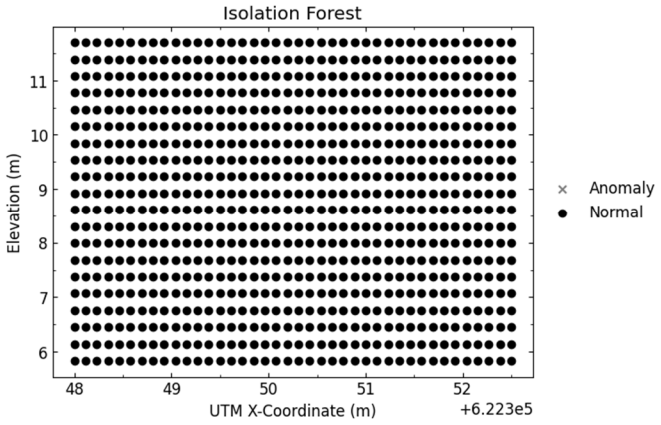


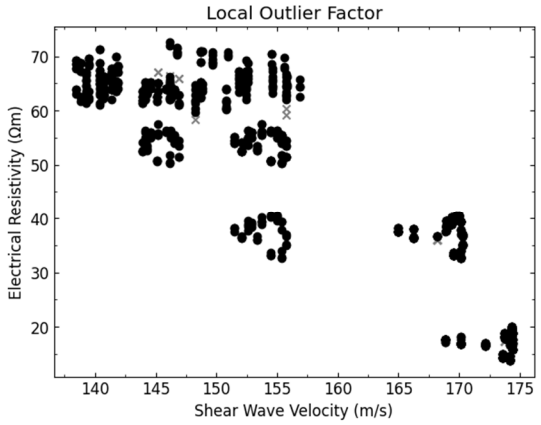


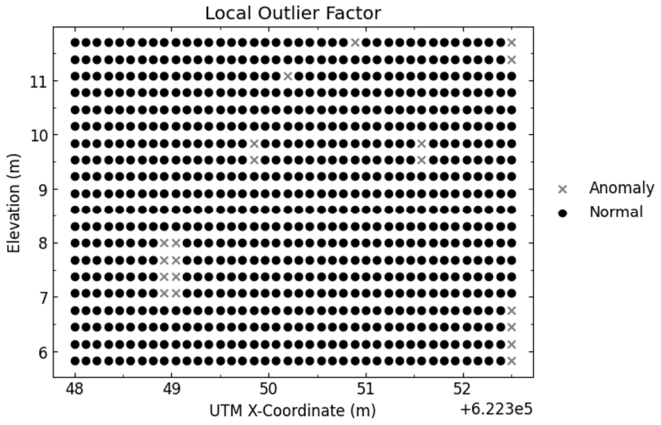


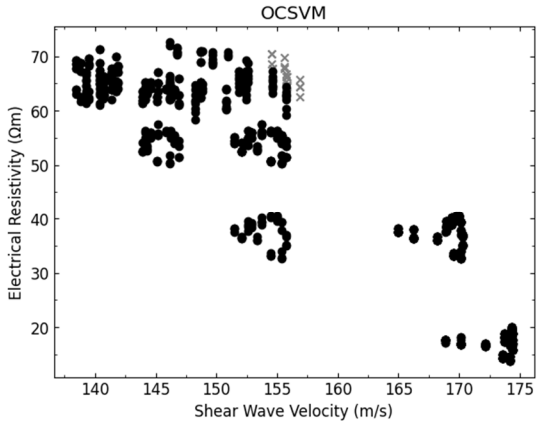
Profile: sc_site3

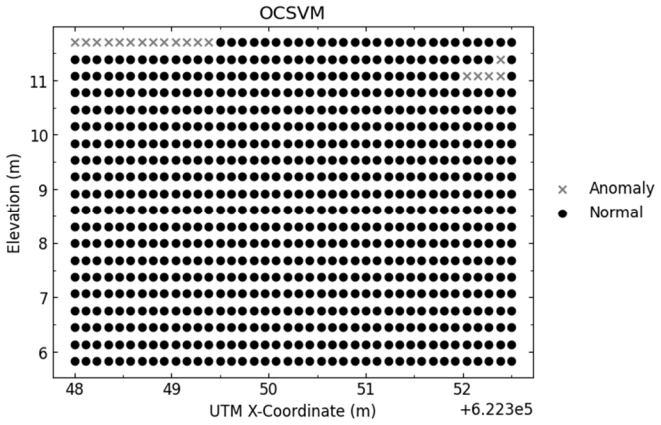




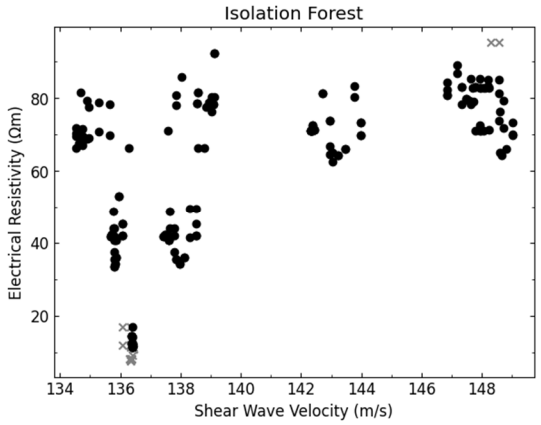


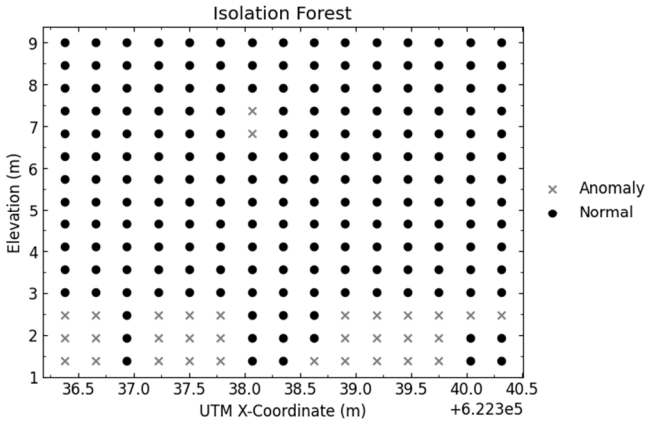


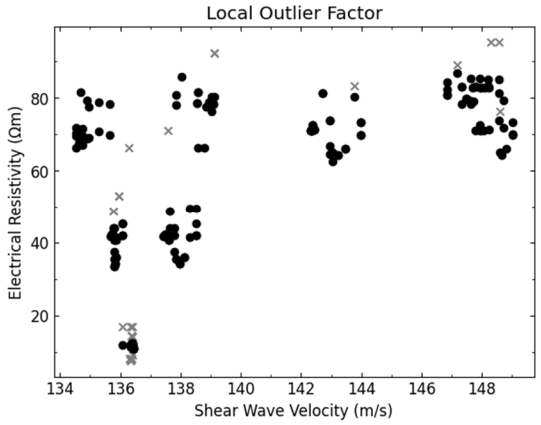


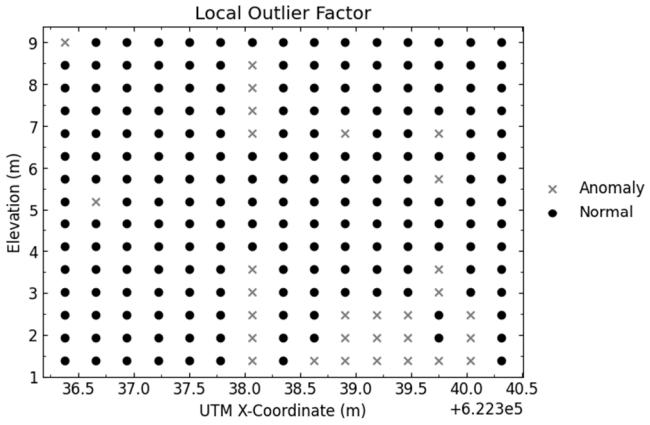


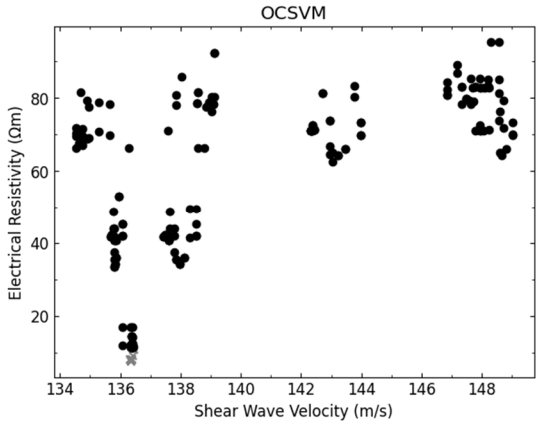
Profile: swt_site3

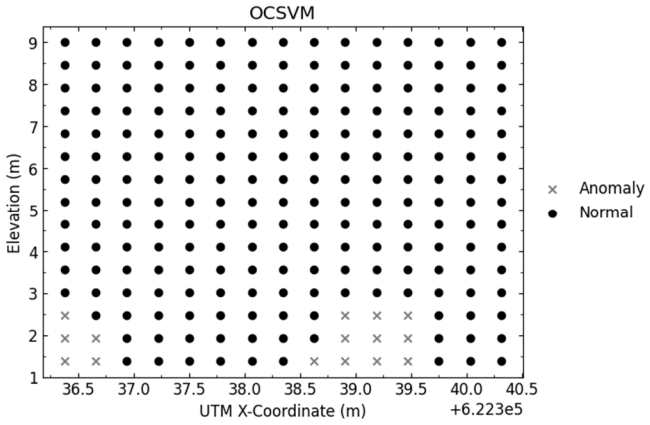






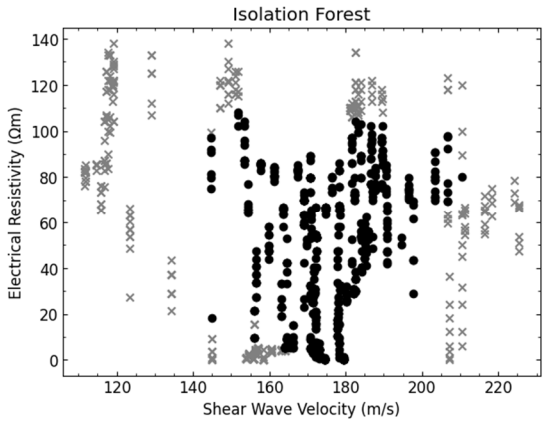


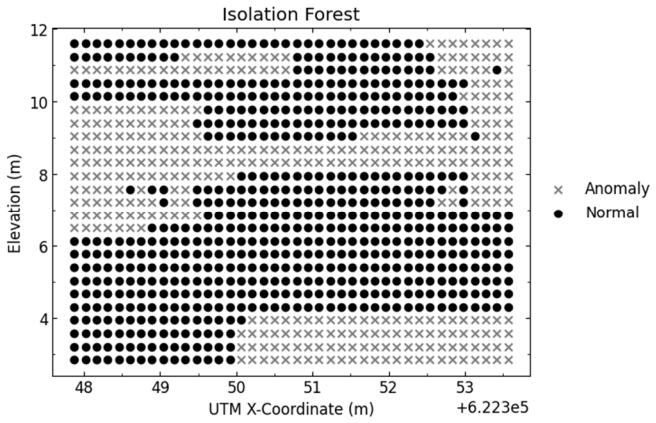


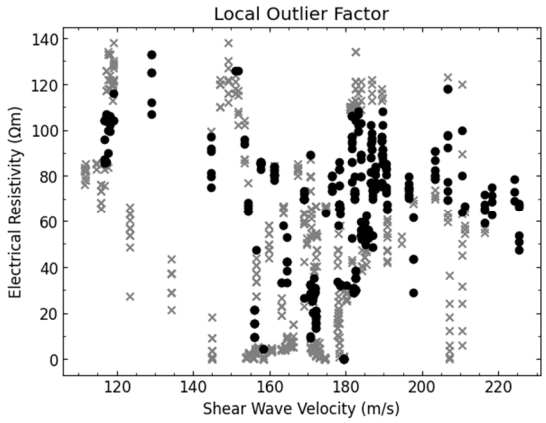


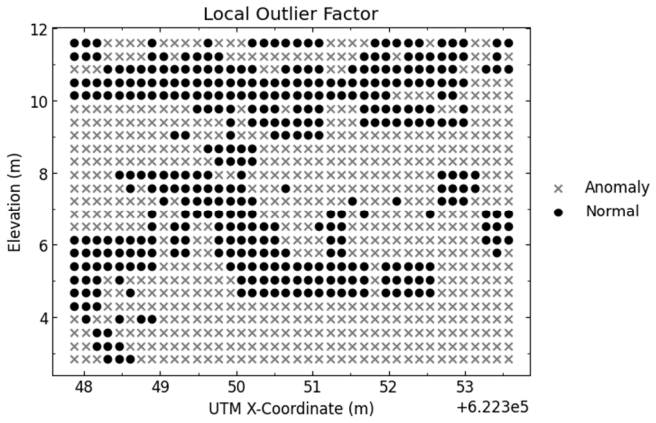
Fit: Site 3 Compilation Dataframe
Predicted: Individual Profiles

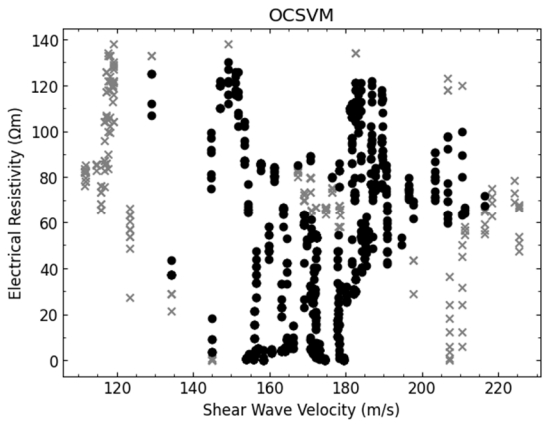
Profile: fc_site3

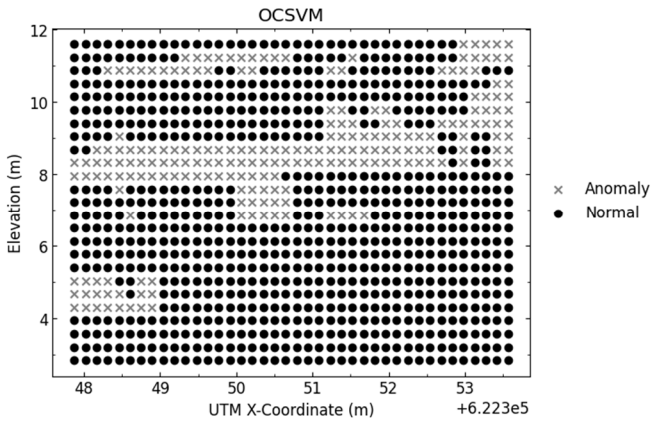




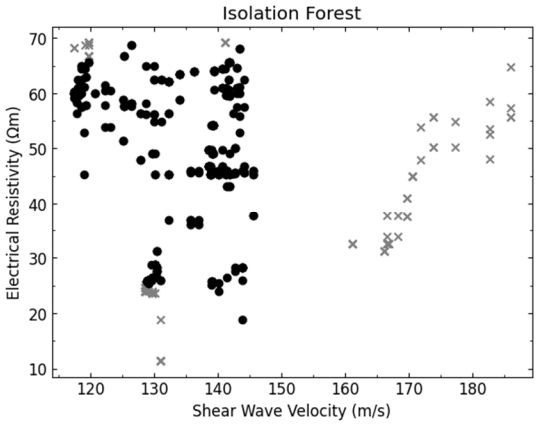


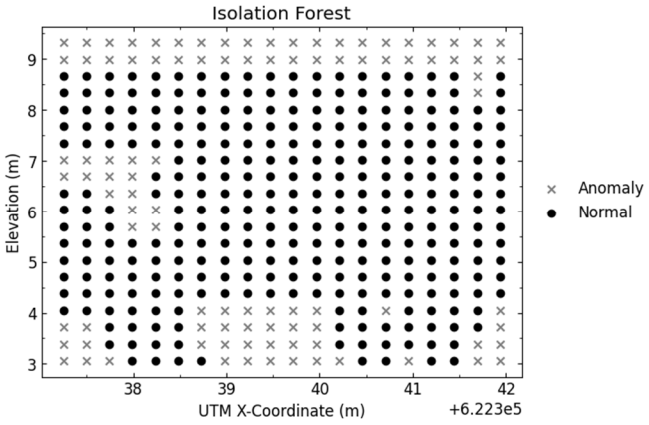


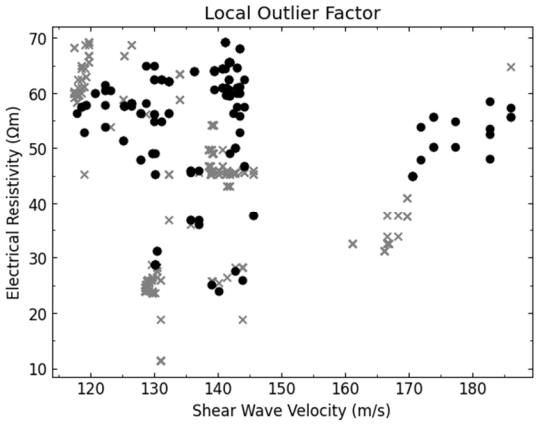


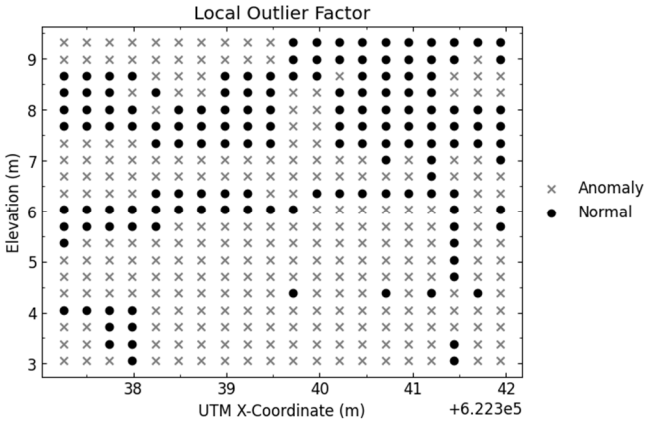


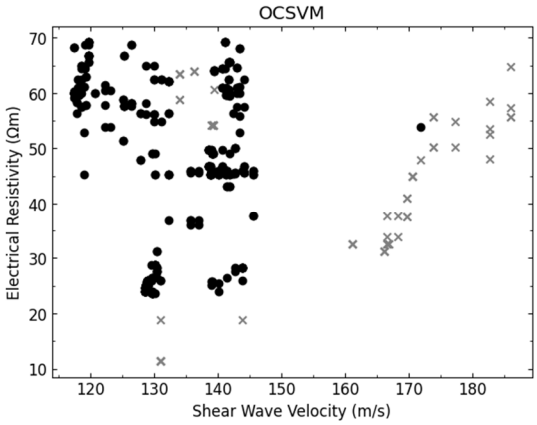
Profile: fwt_site3

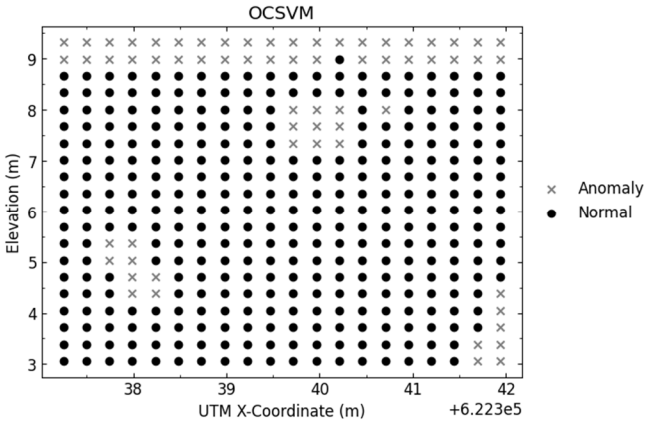




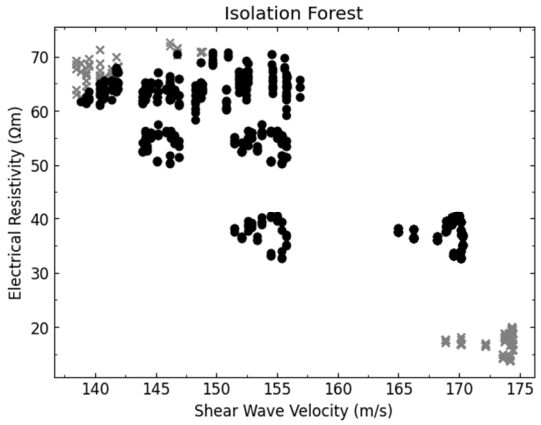


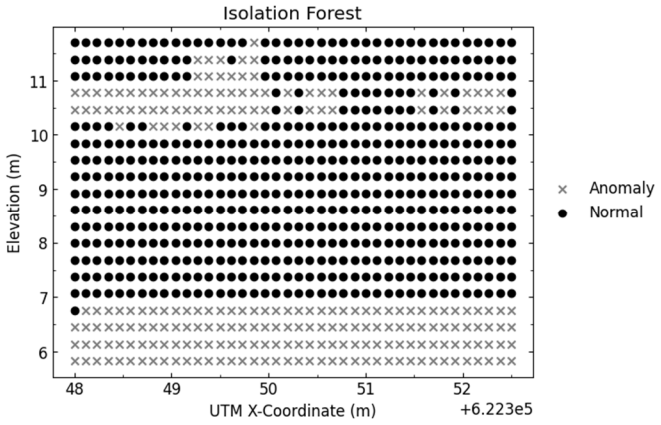


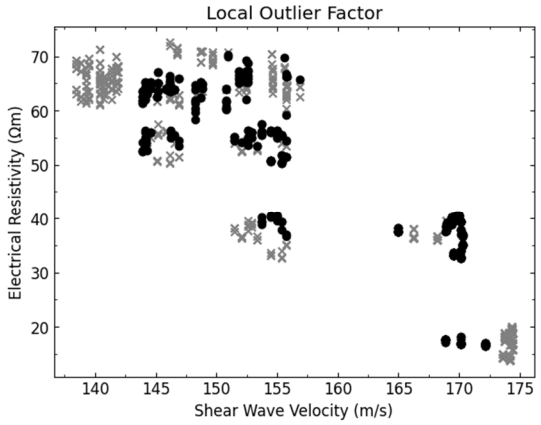


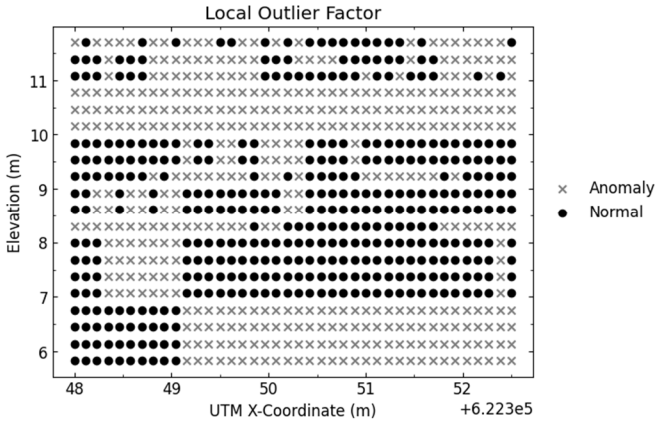


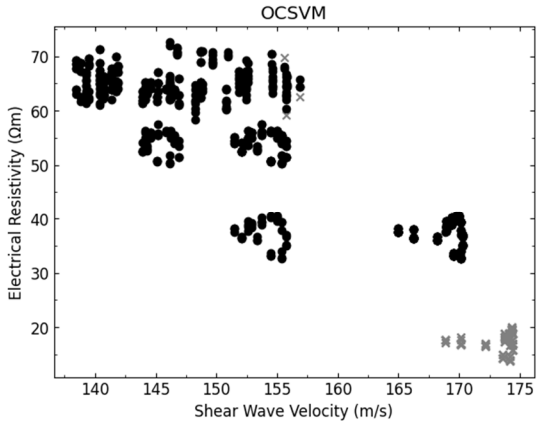
Profile: sc_site3

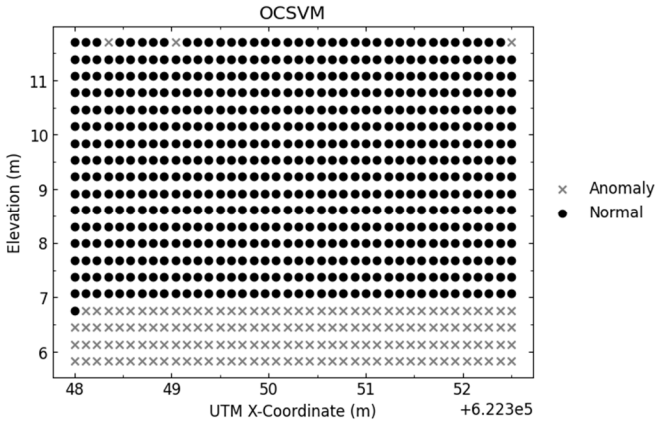




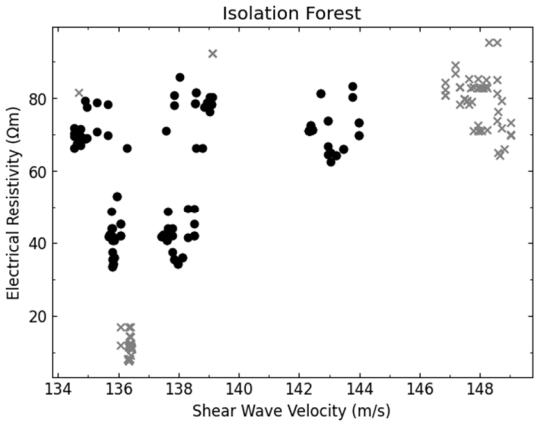


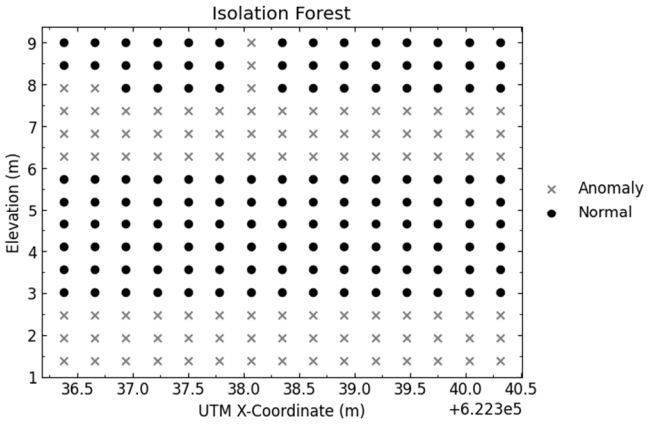


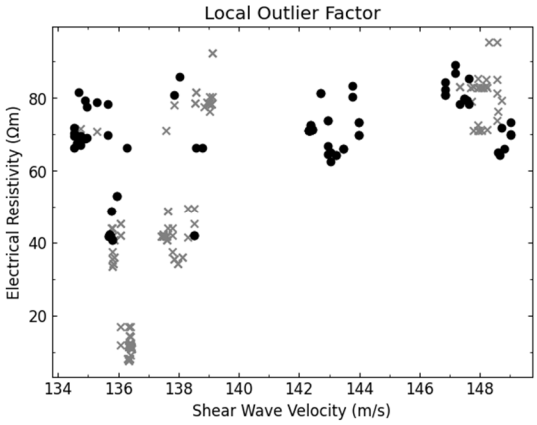


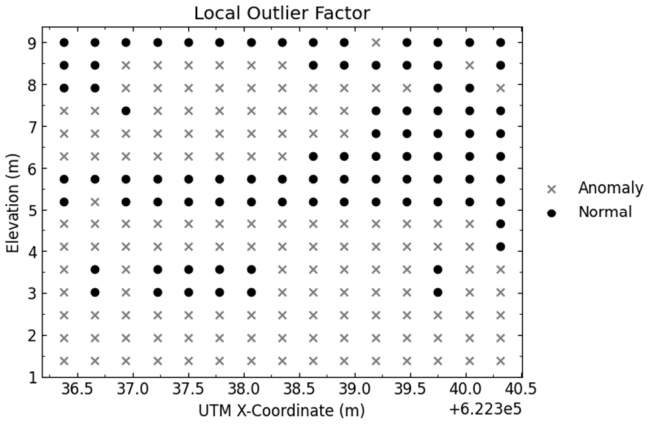


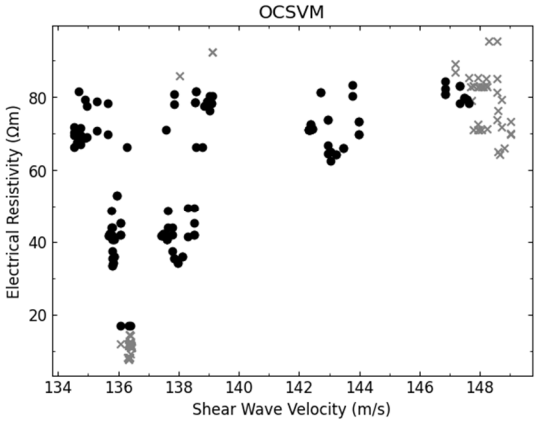
Profile: swt_site3

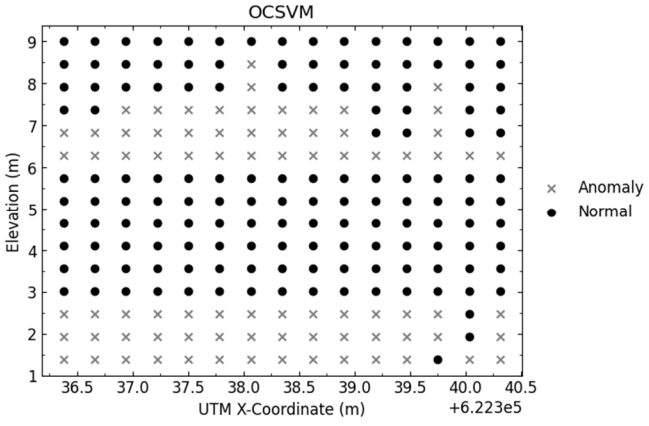




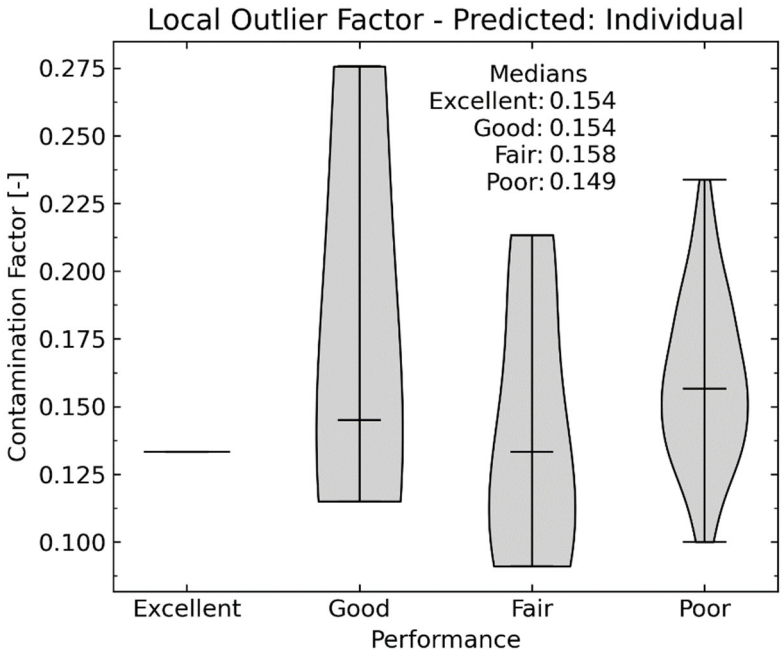


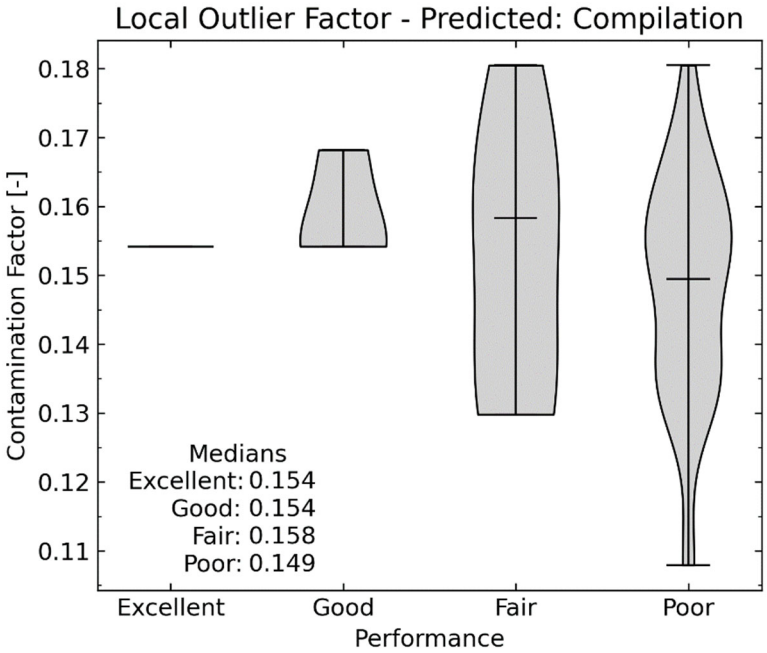




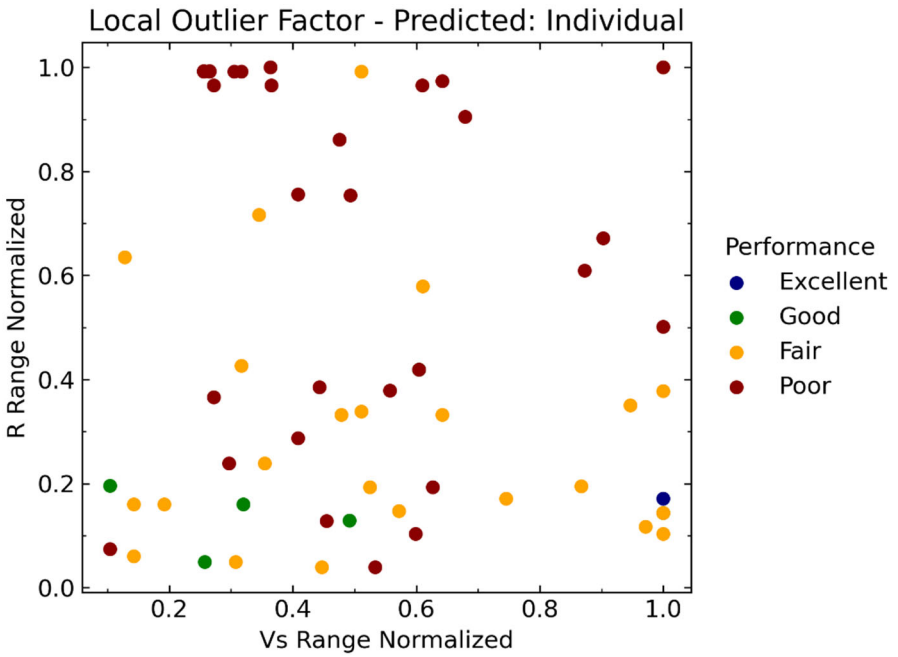


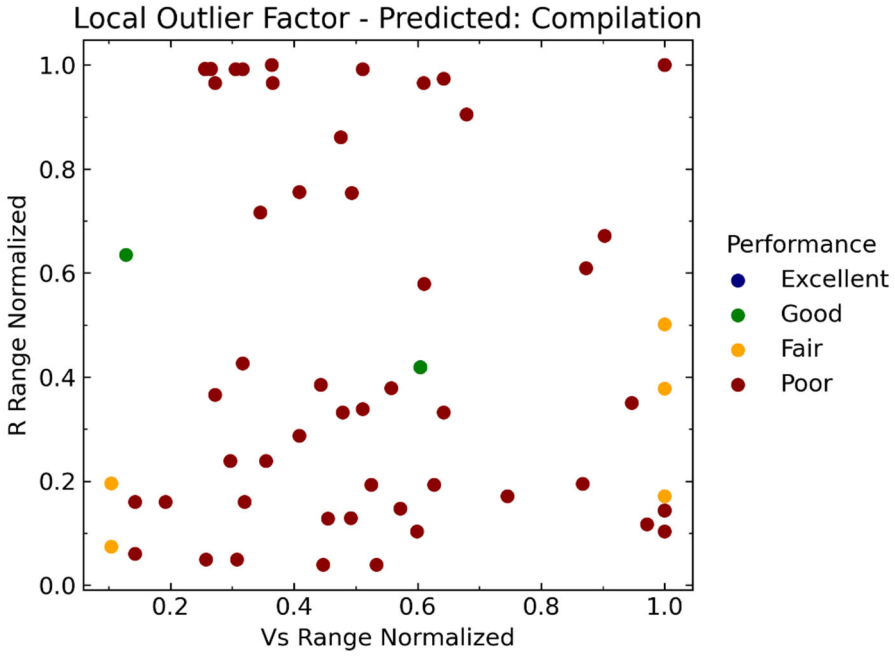
Effect of Contamination Factor on Performance





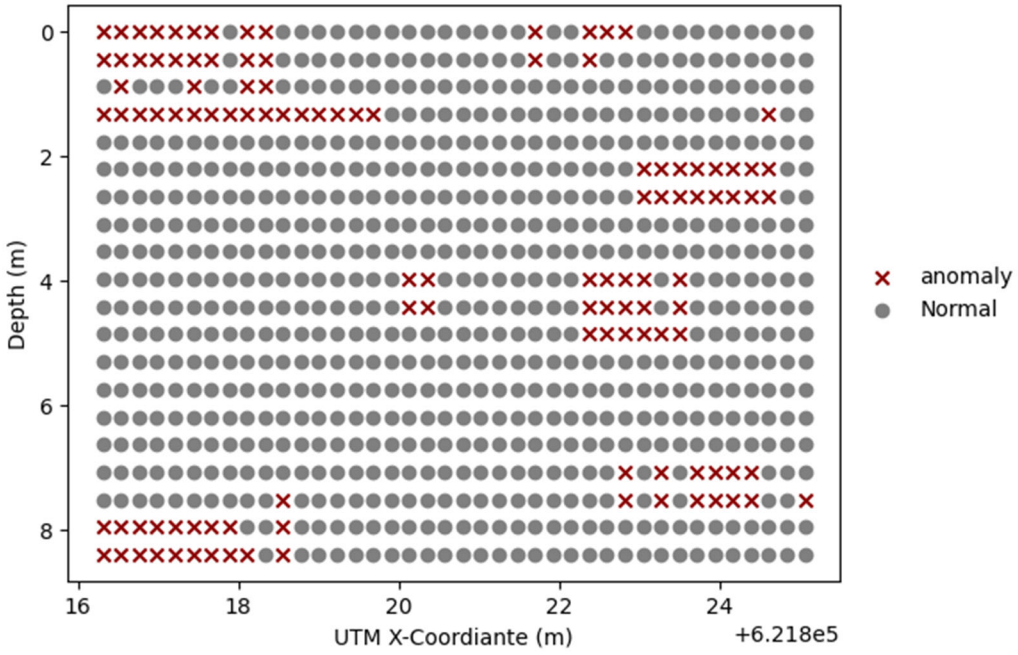
Effect of Shear Wave Velocity and Electrical Resistivity on Performance



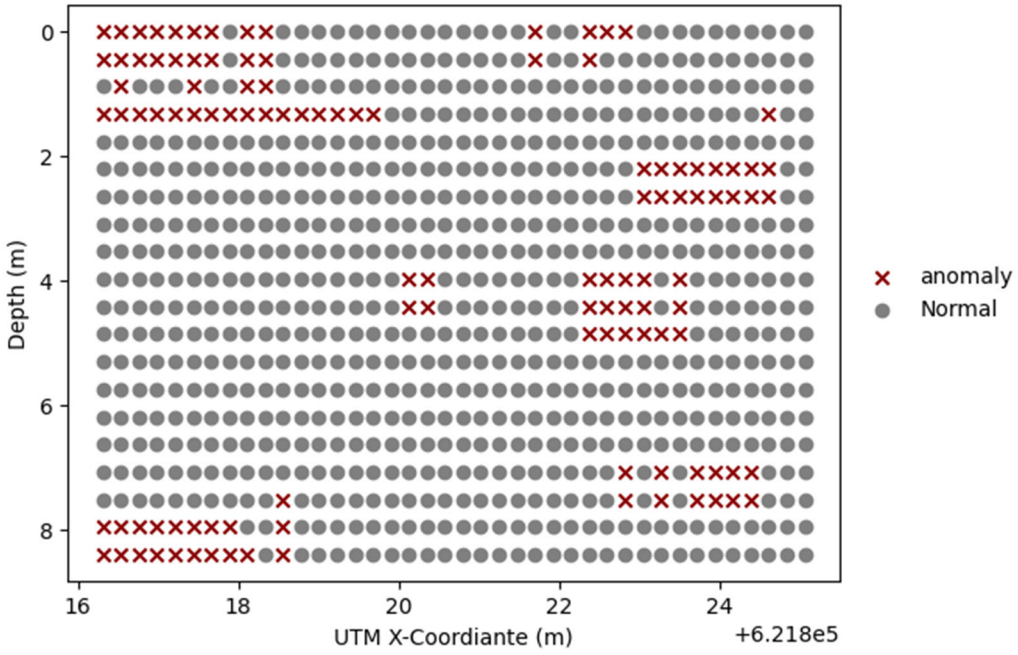


Appendix I – Clustering and Anomaly Detection Fusion

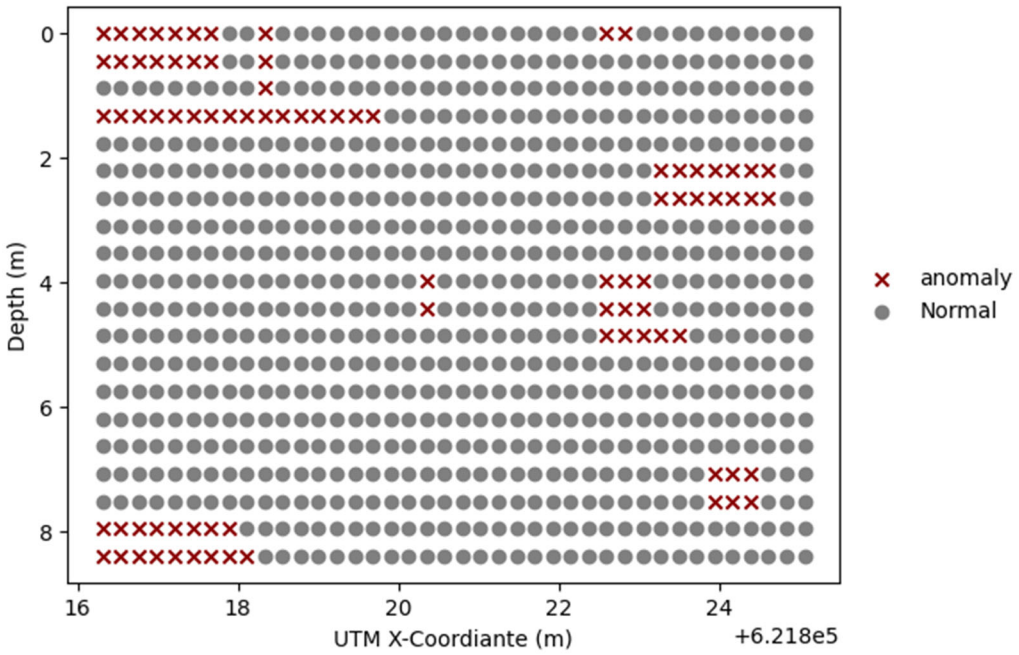
Anomaly Detection Noise Threshold Analysis



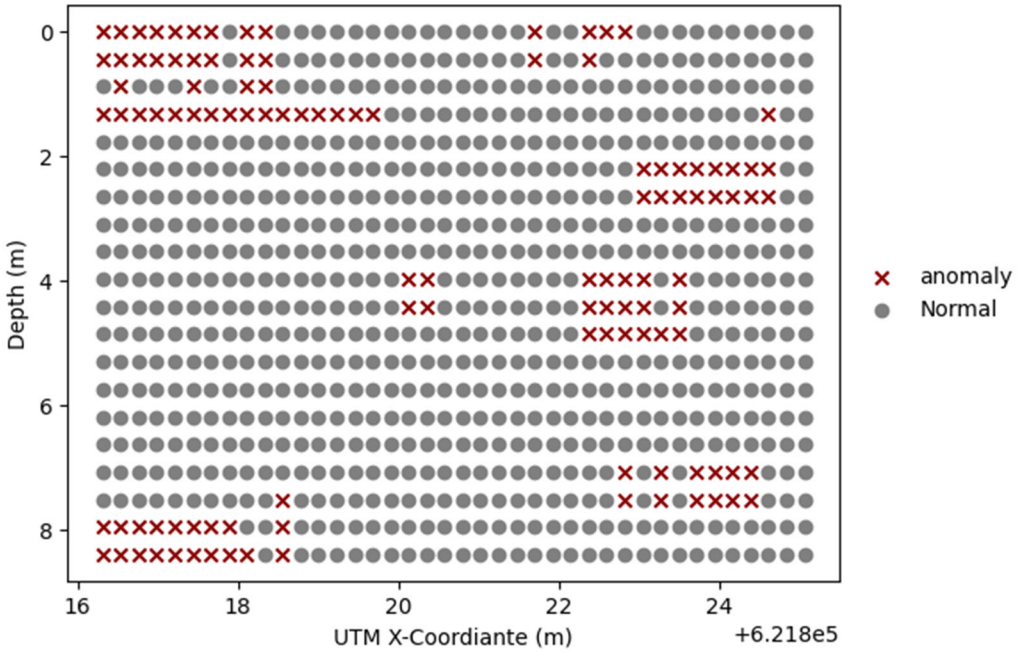
(a) Original distribution of anomalies in the subsurface



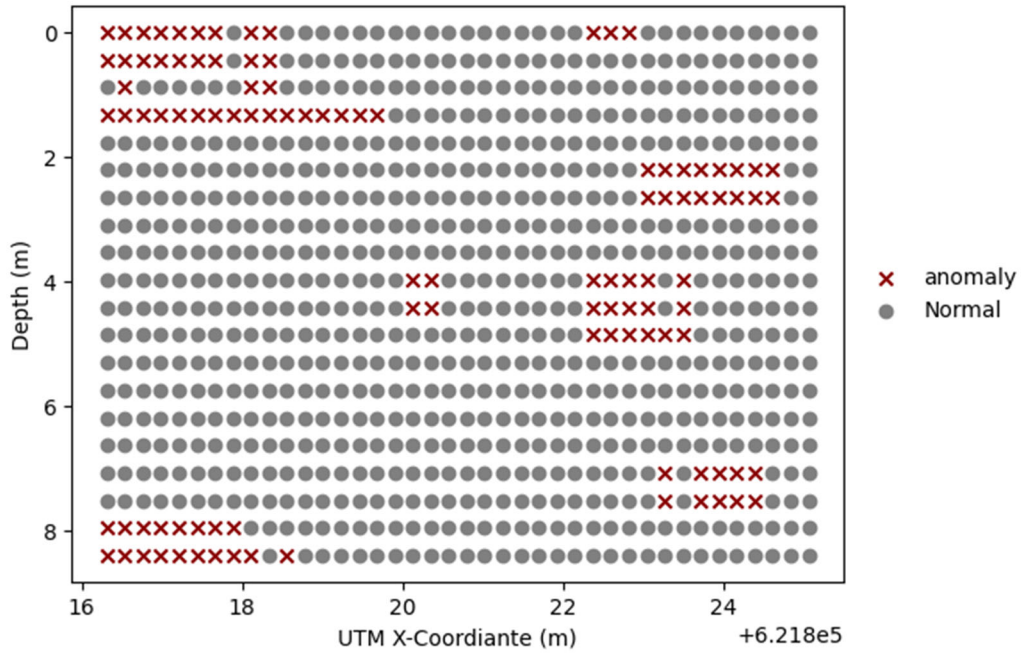
(b) N = 2, M = 1


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

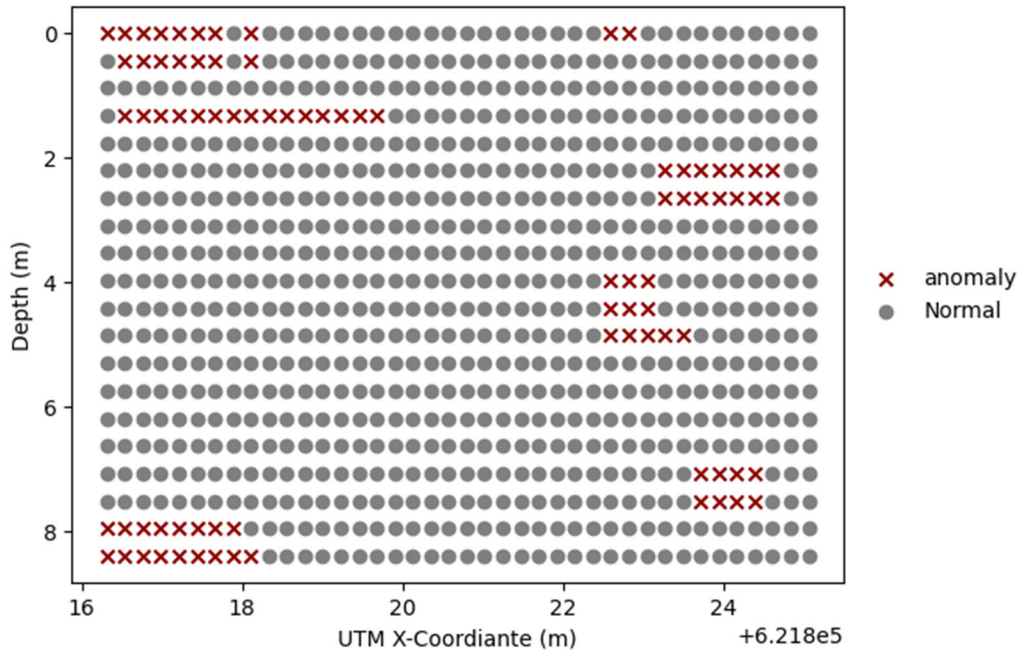
(c) N = 2, M =2


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

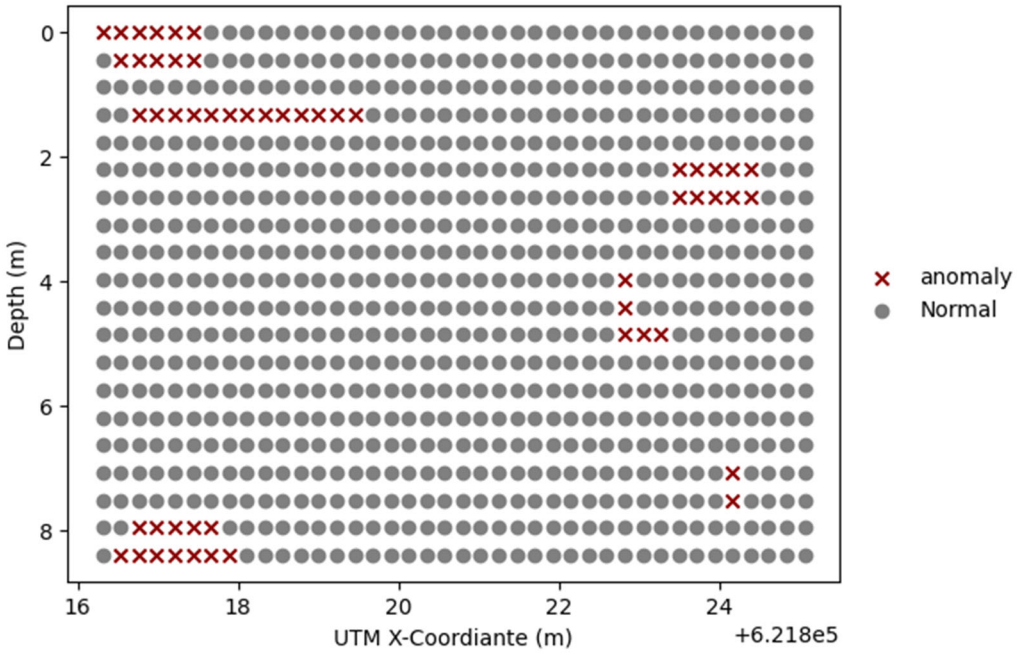
(d) N = 4, M = 1


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

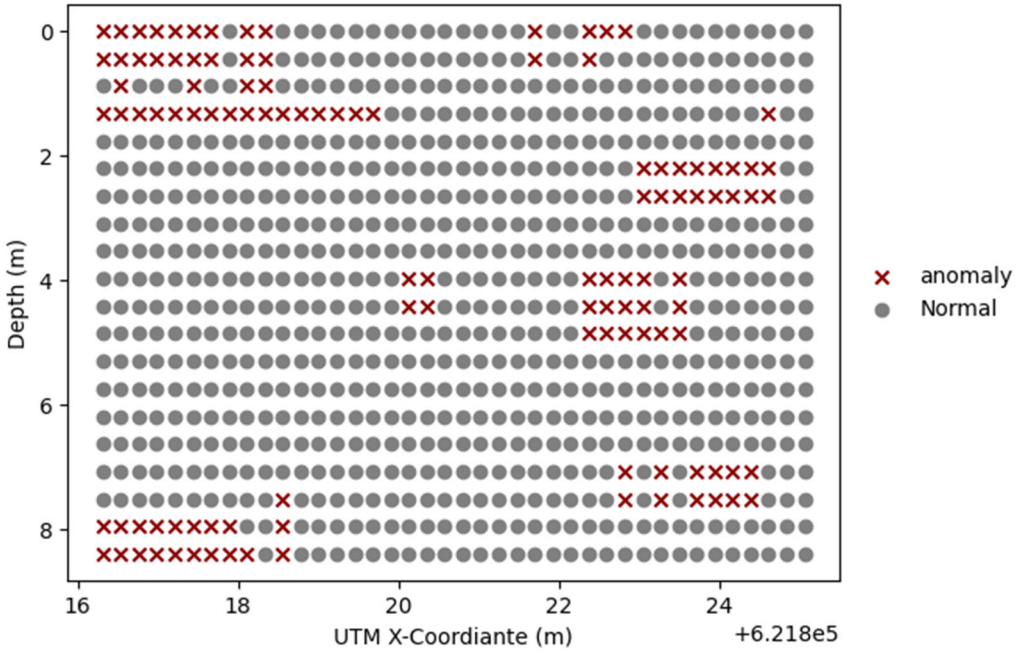
(e) N = 4, M = 2


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

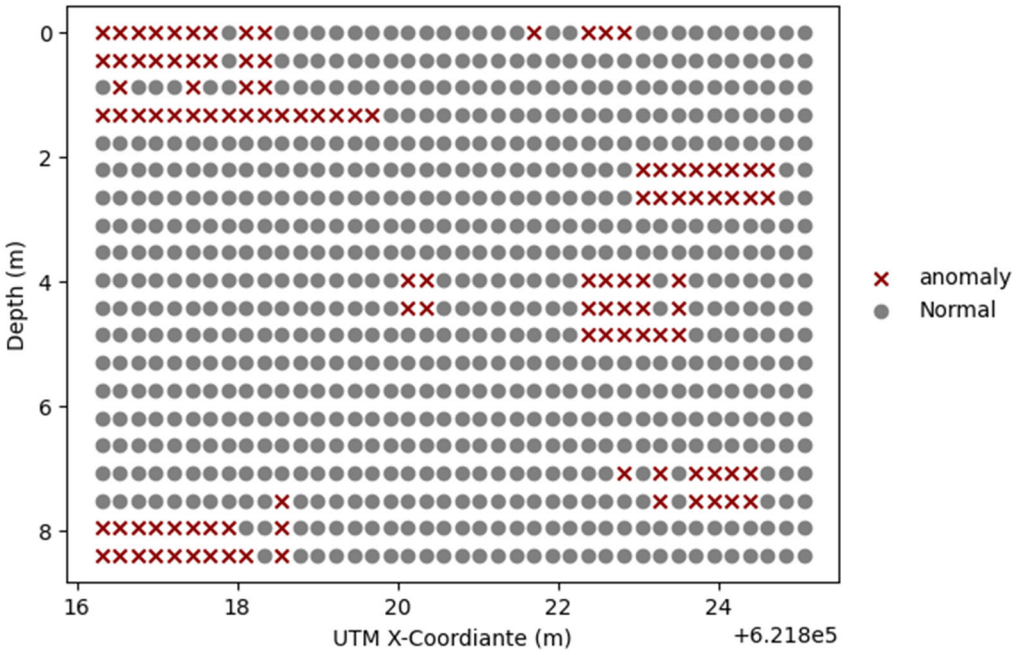
(f) N = 4, M = 3


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

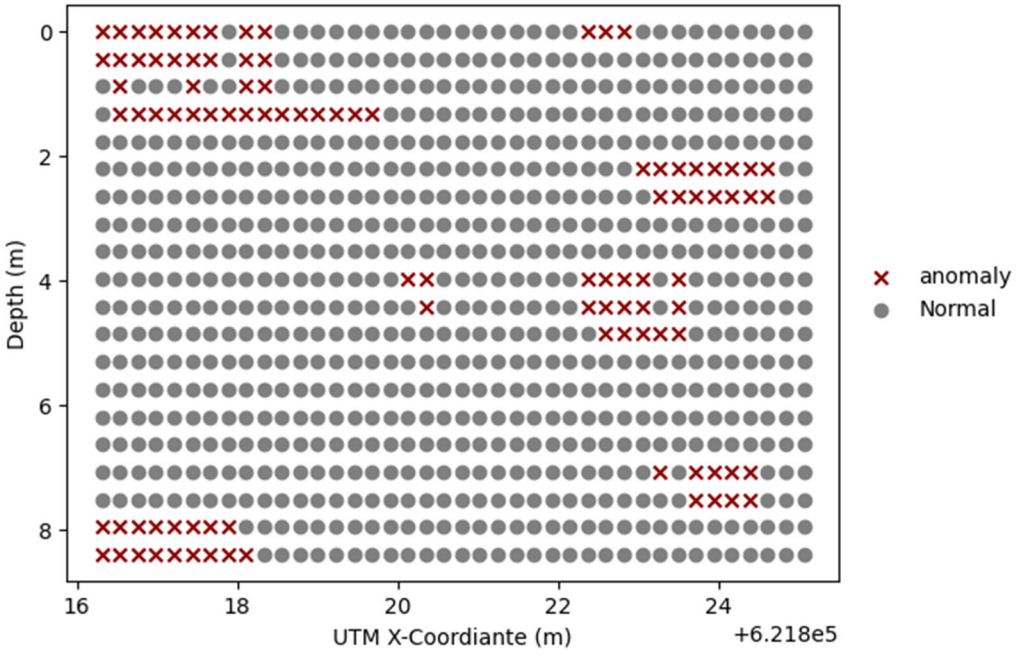
(g) N = 4, M = 4


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

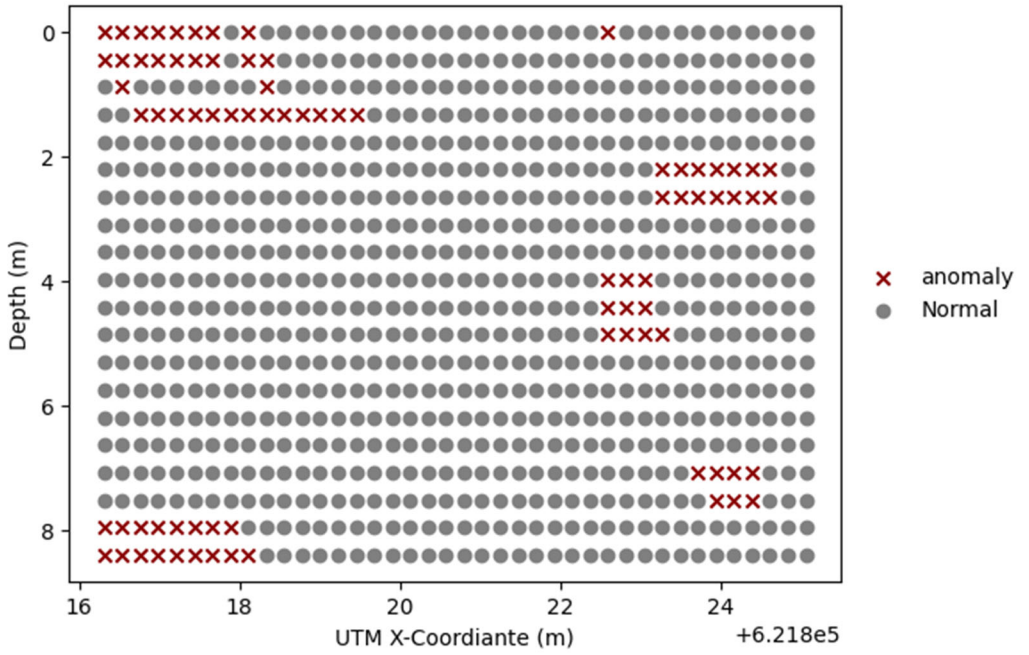
(h) N = 6, M = 1


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

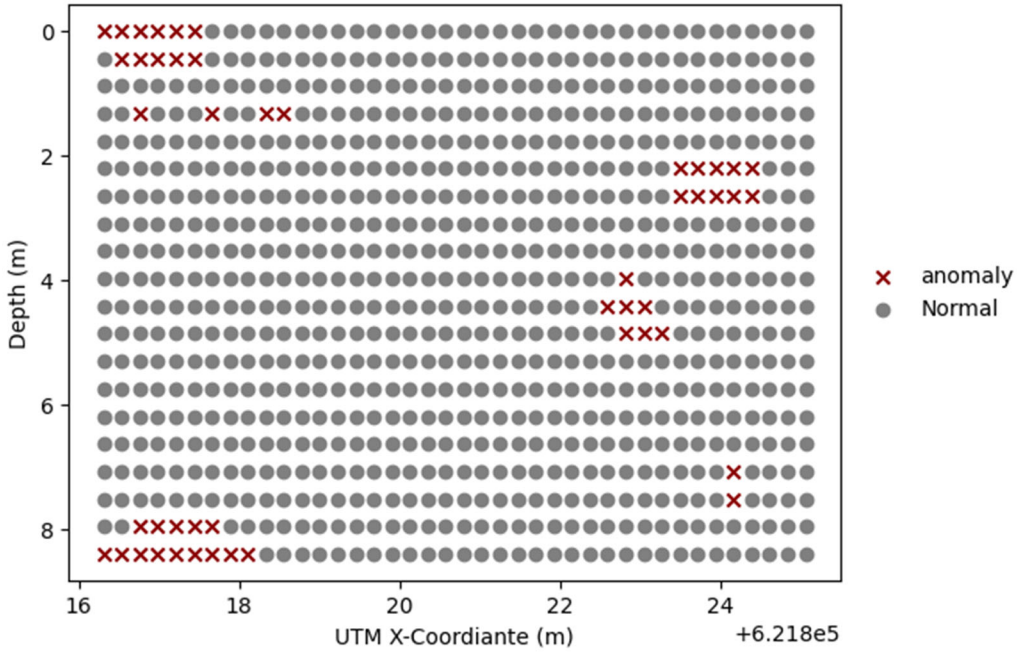
(i) N = 6, M = 2


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

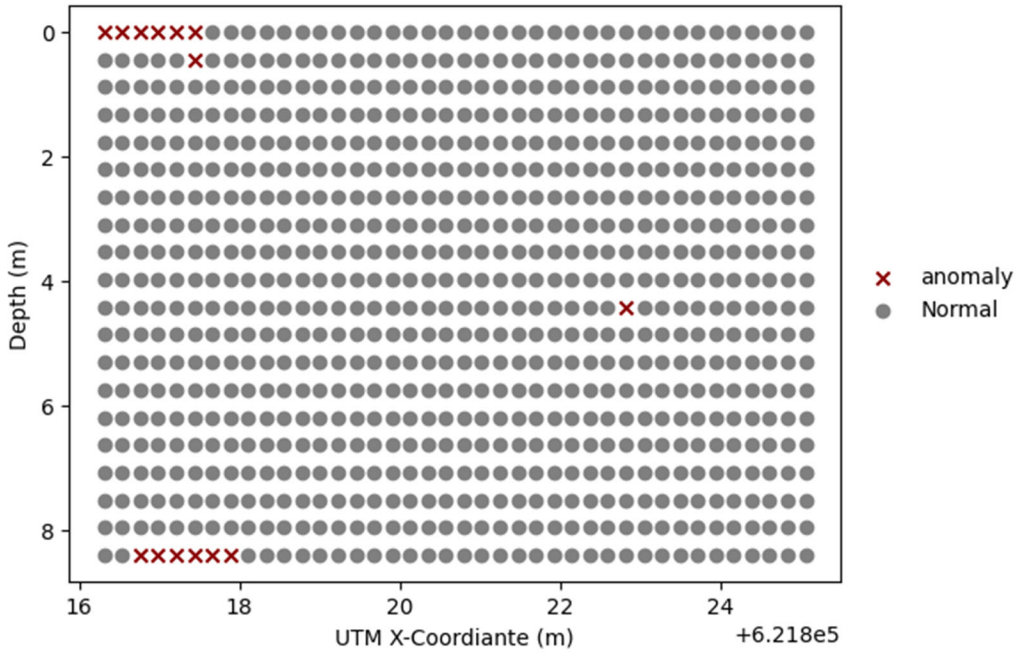
(j) N = 6, M = 3


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

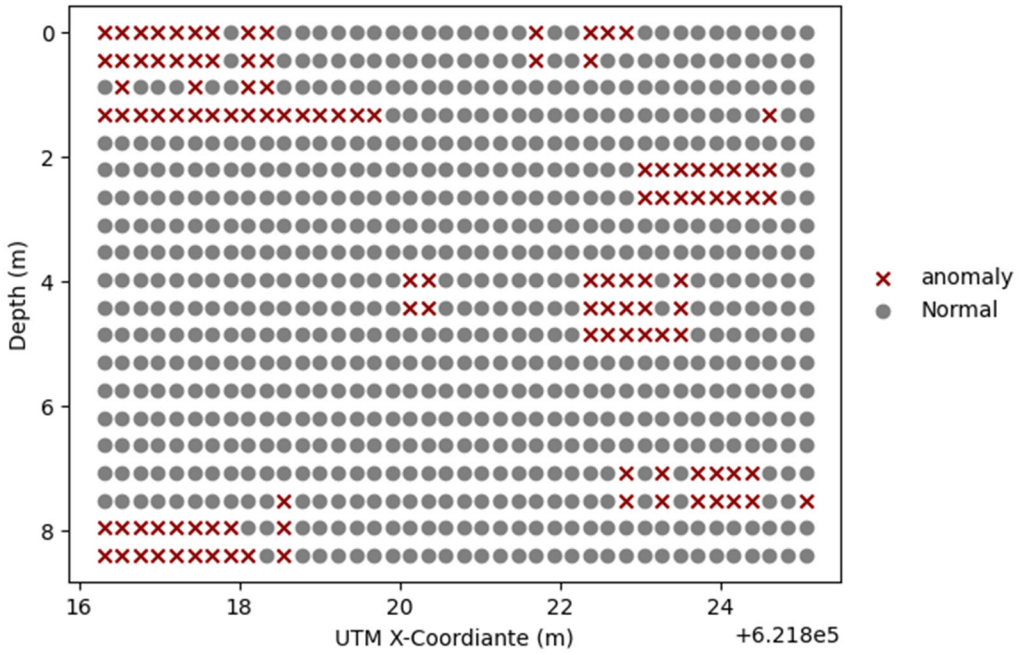
(k) N = 6, M = 4


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

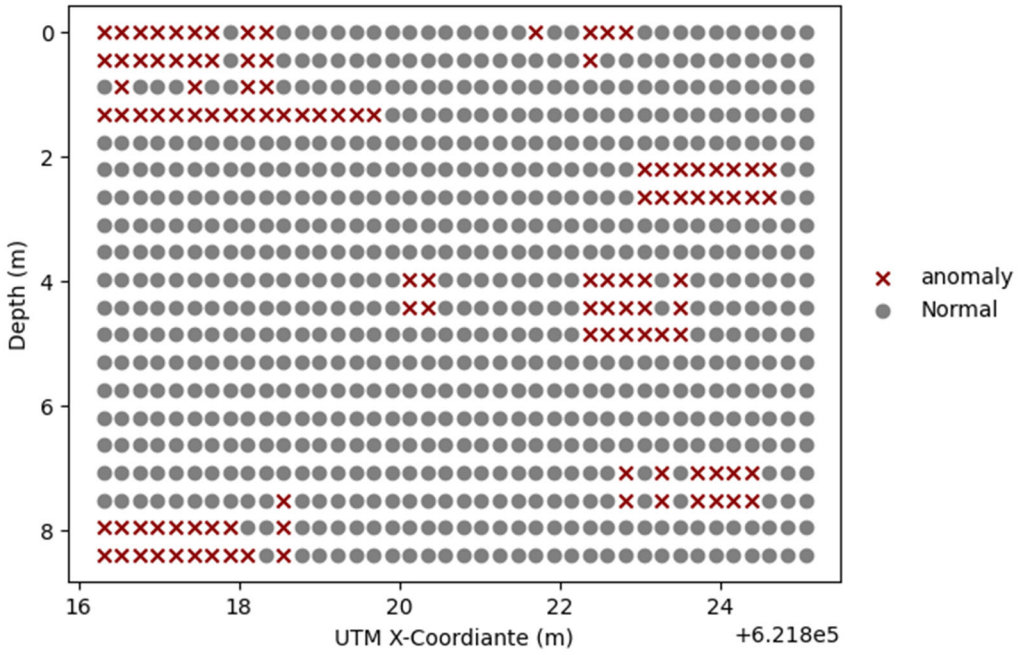
(l) N = 6, M = 5


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

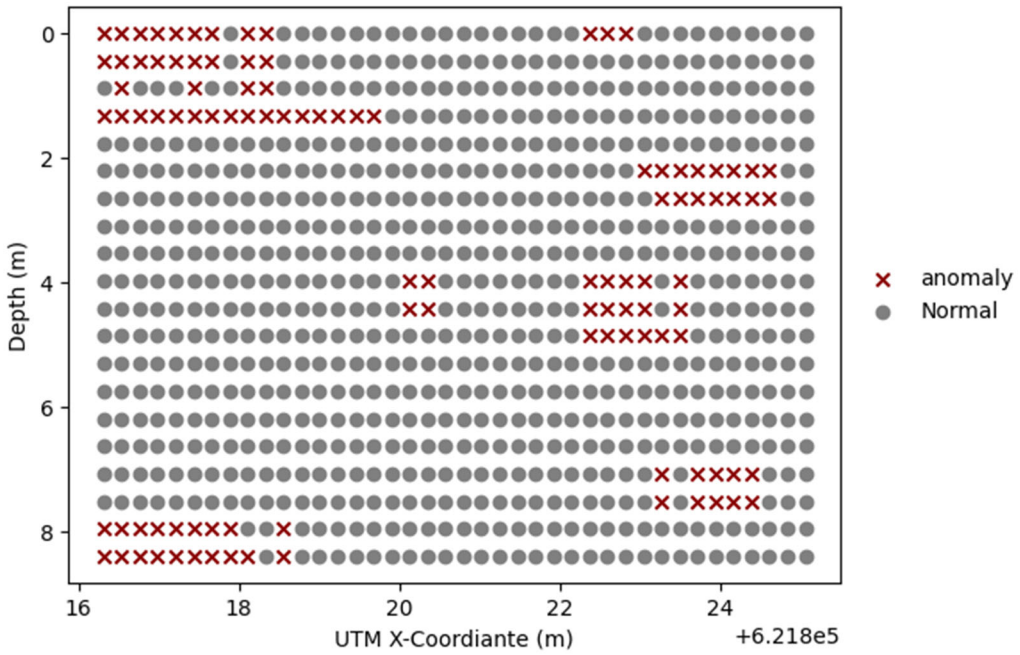
(m) N = 6, M = 6


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

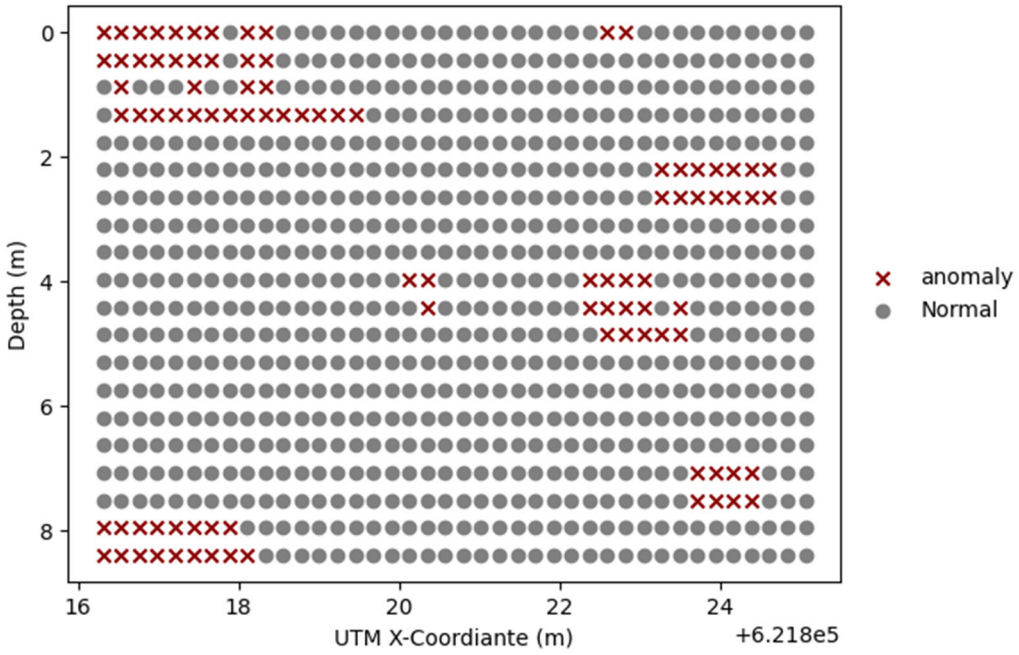
(n) N = 8, M = 1


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

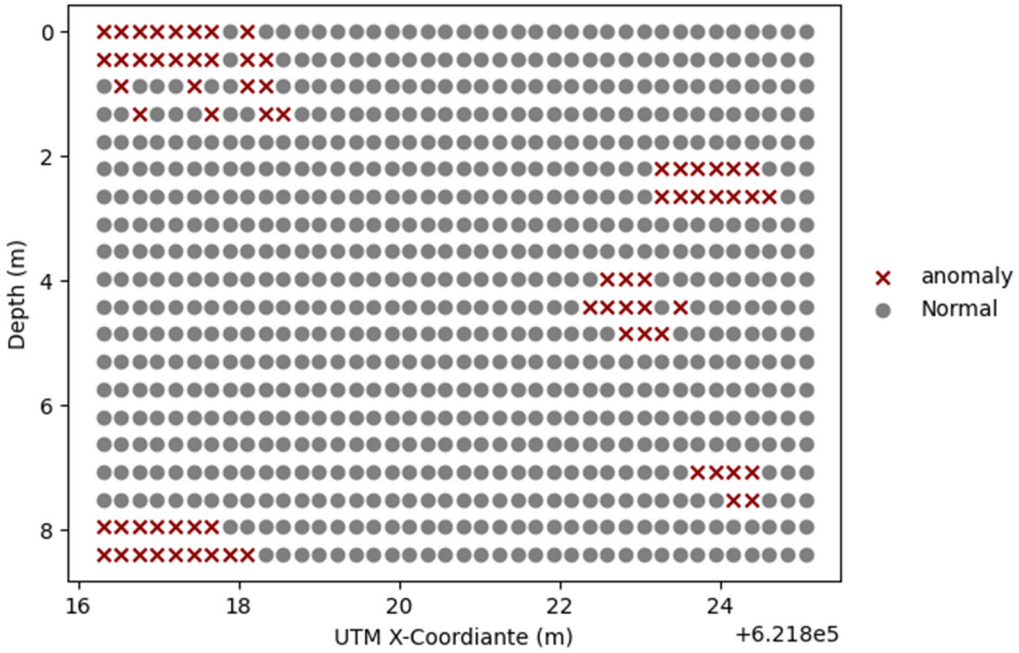
(o) N = 8, M = 2


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

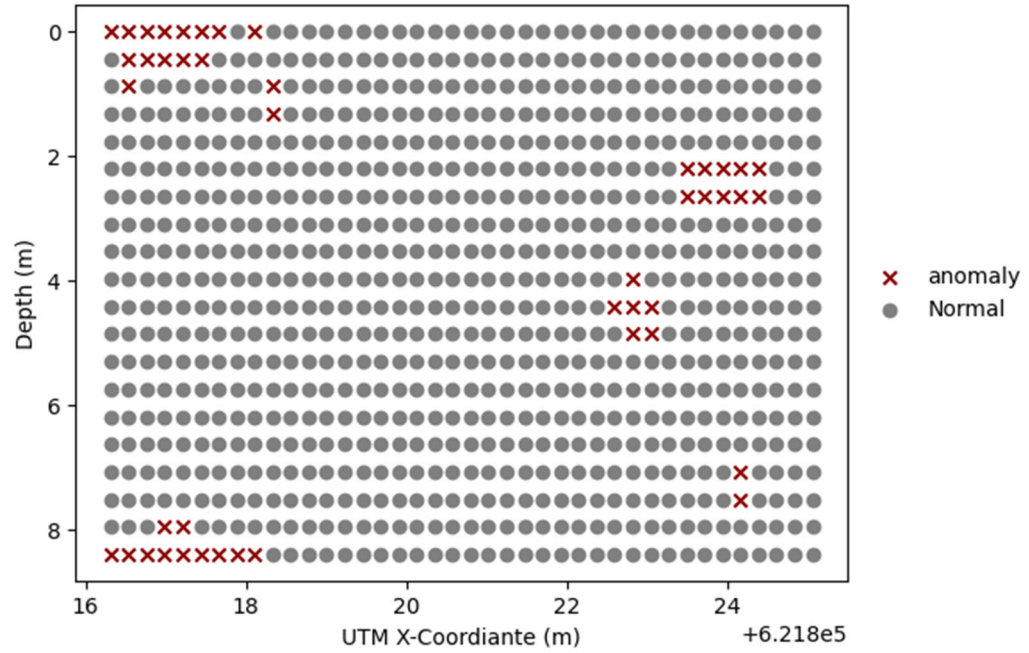
(p) N = 8, M = 3


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

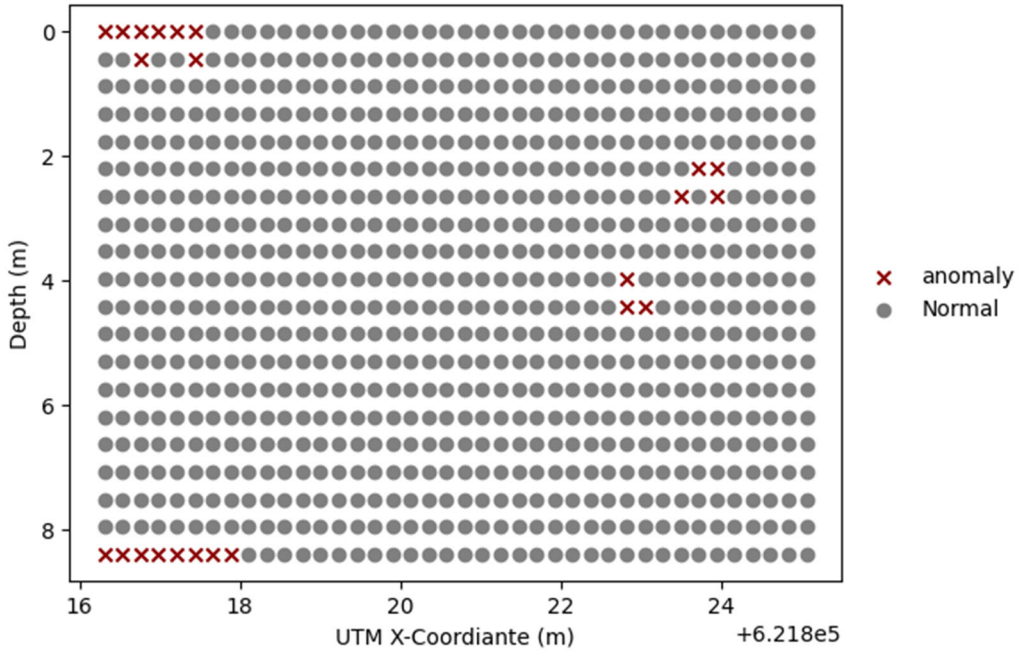
(q) N = 8, M = 4


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

(r) N = 8, M = 5


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

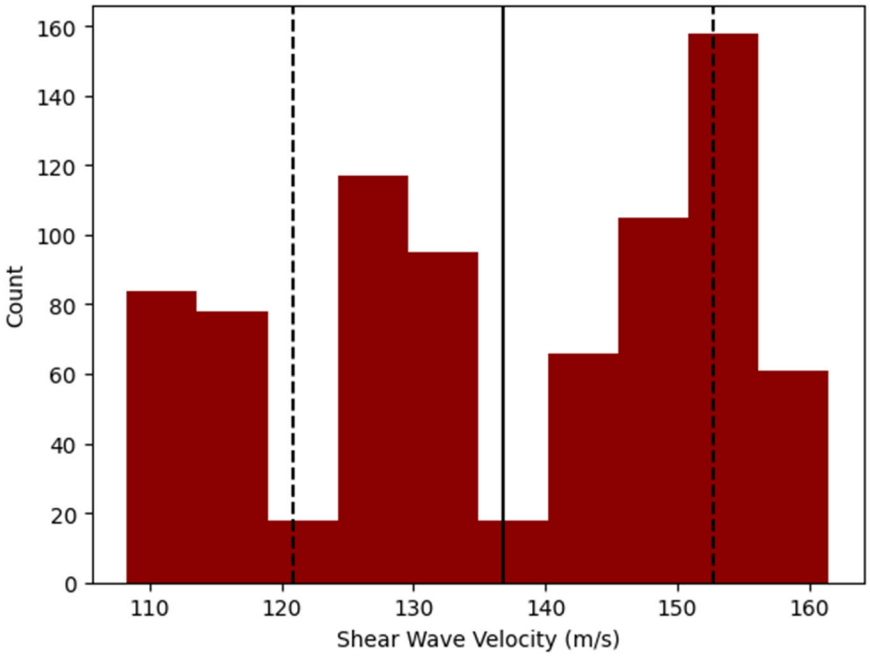
(s) N = 8, M = 6


Depth (m)
0
2
4
6
8
16
18
20
22
24
UTM X-Coordiante (m)
+6.218e5
anomaly
Normal

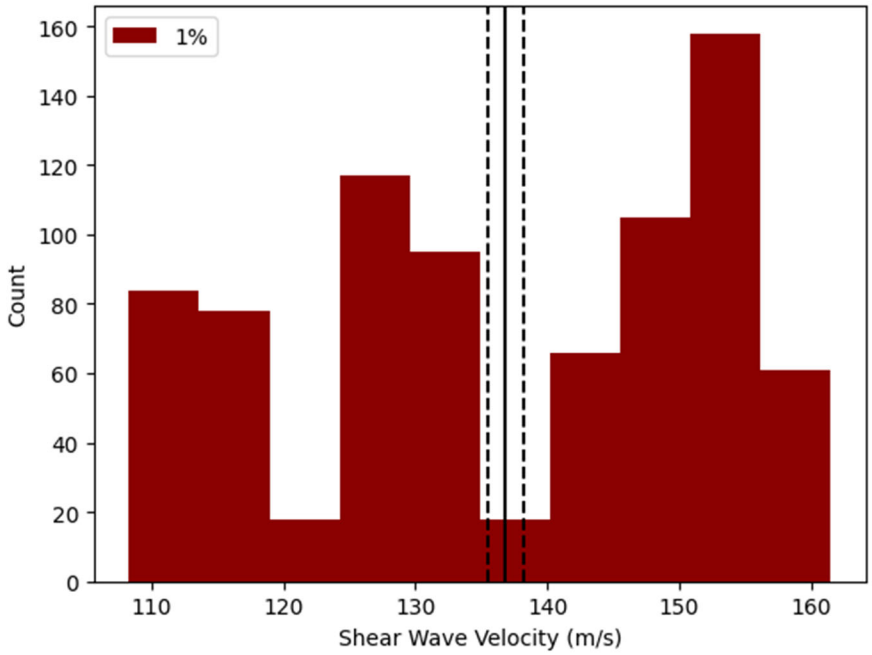
(t) N = 8, M = 7

Subsurface Change Threshold Histograms and Subsurface Figures

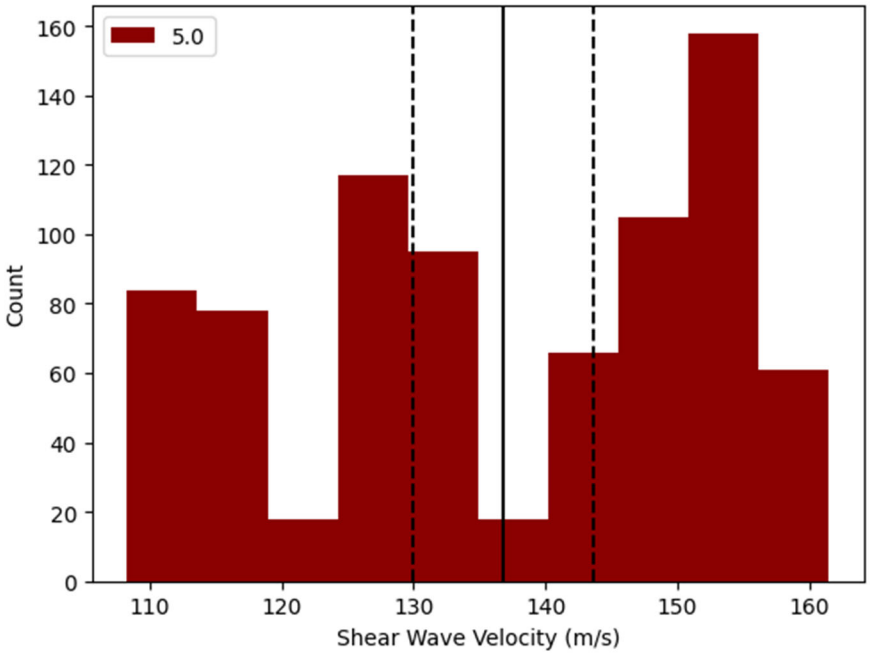
Case Study 1: fc_site2



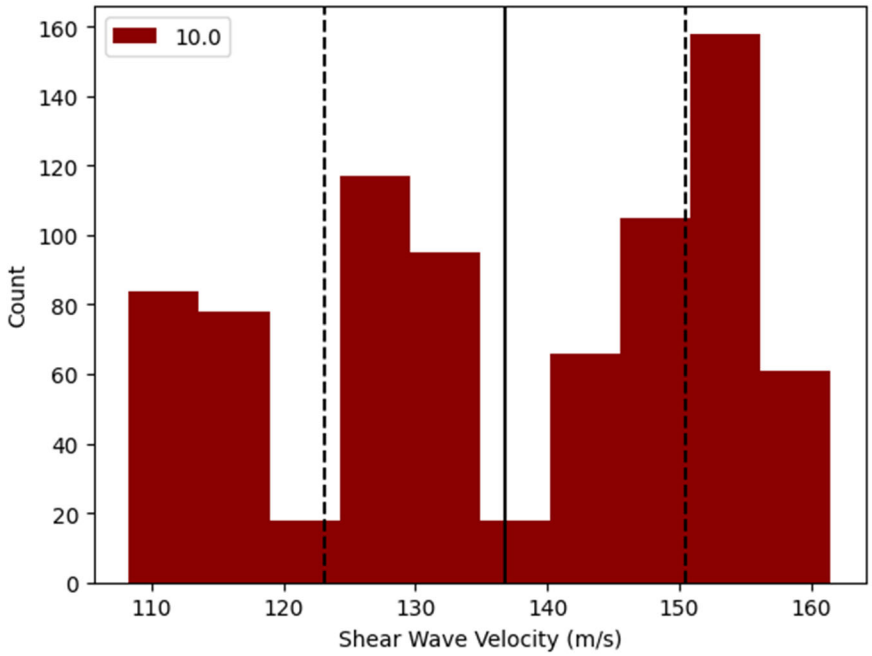
+/- 1 Standard deviation of change from the mean for shear wave velocity



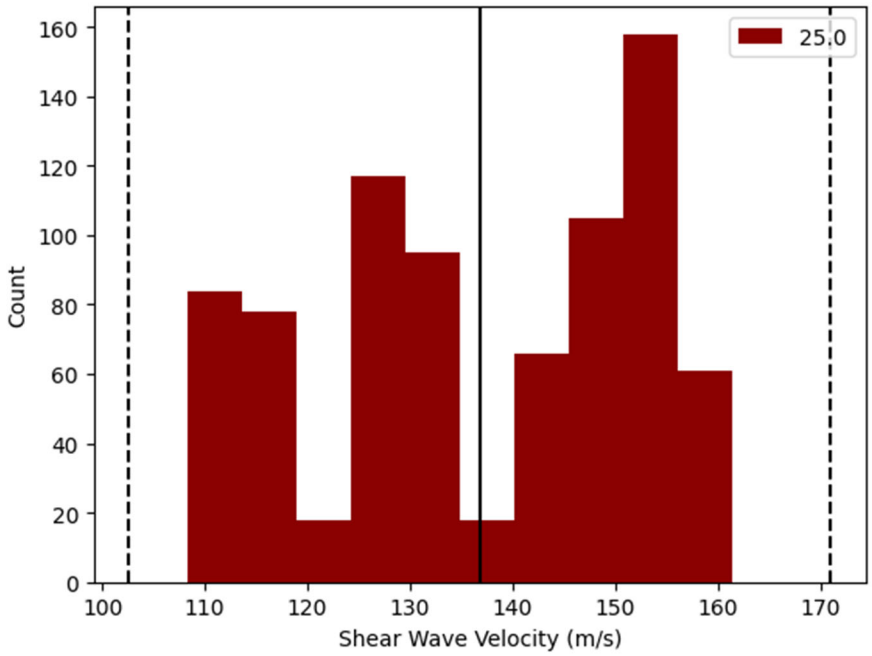
+/- 1% change from the mean for shear wave velocity



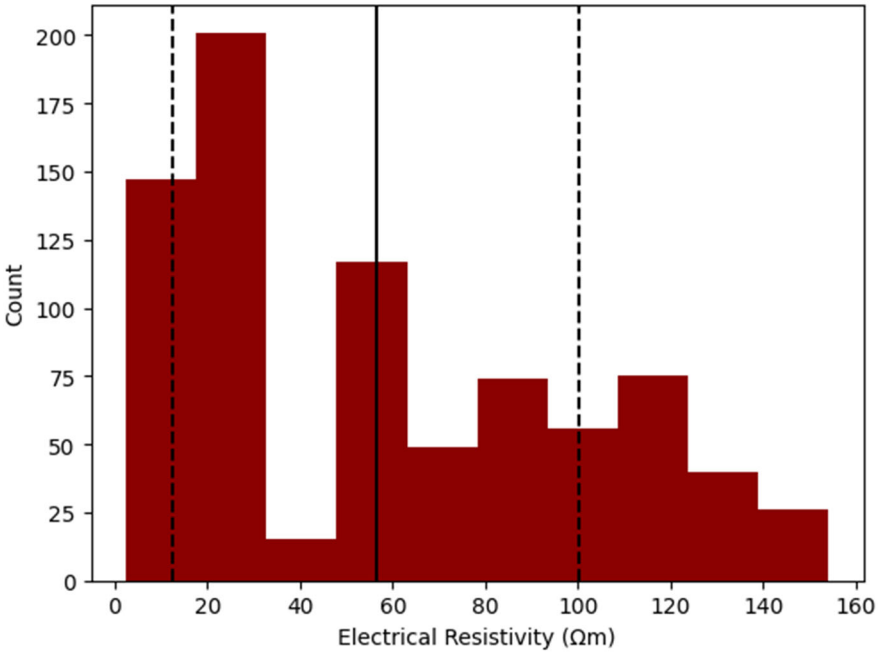
+/- 5% change from the mean for shear wave velocity



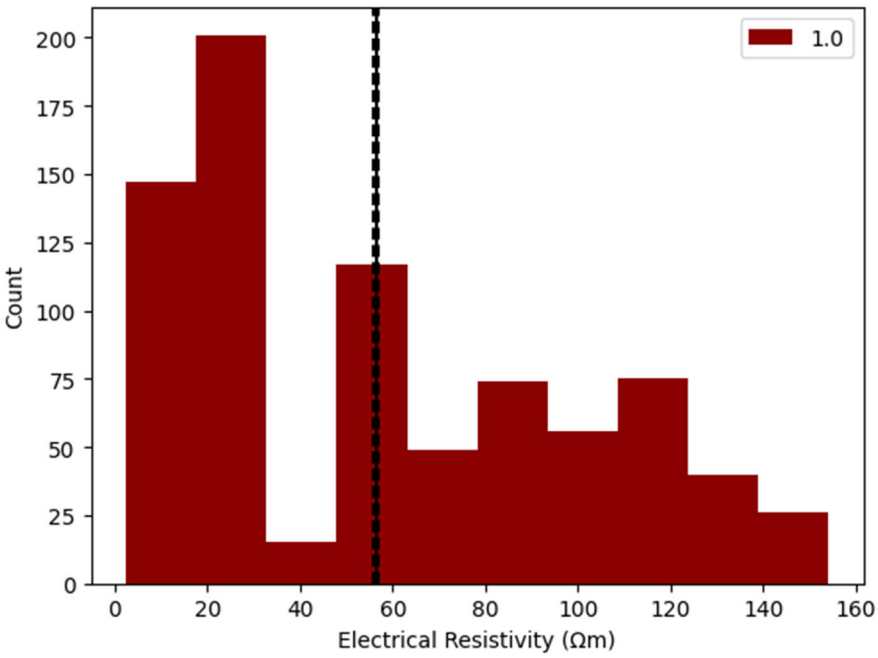
+/- 10% change from the mean for shear wave velocity



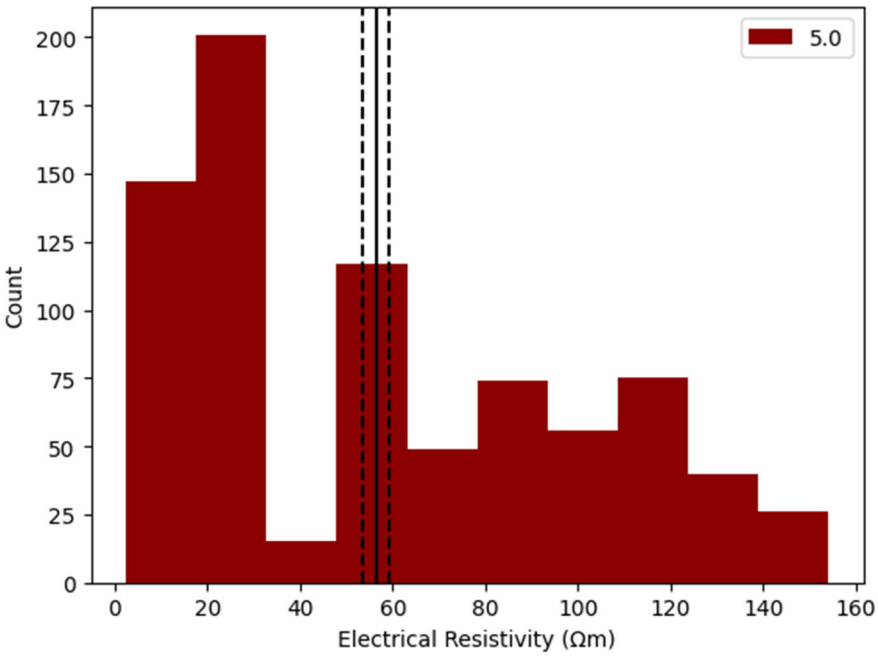
+/- 25% change from the mean for shear wave velocity



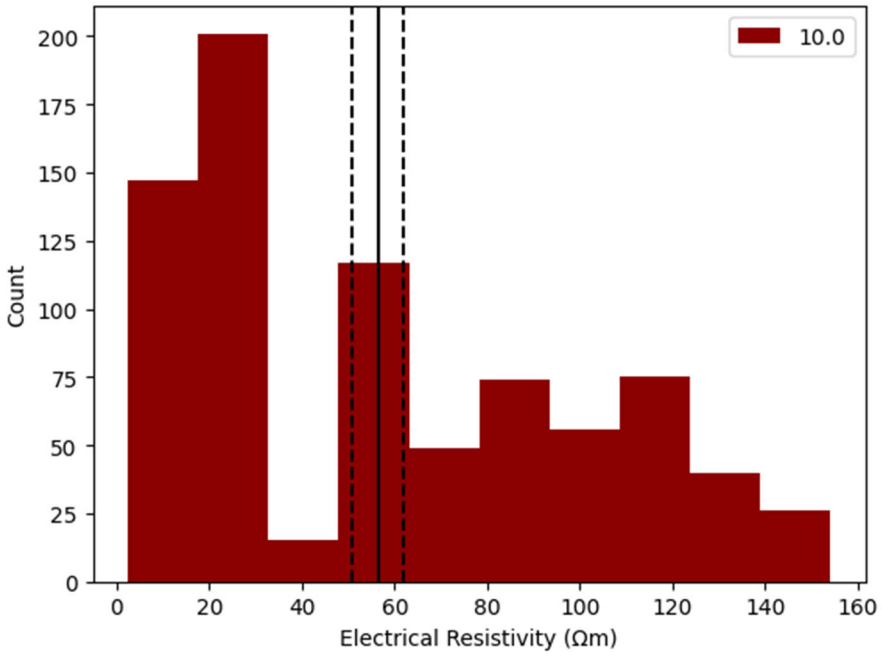
+/- 1 Standard deviation of change from the mean for electrical resistivity



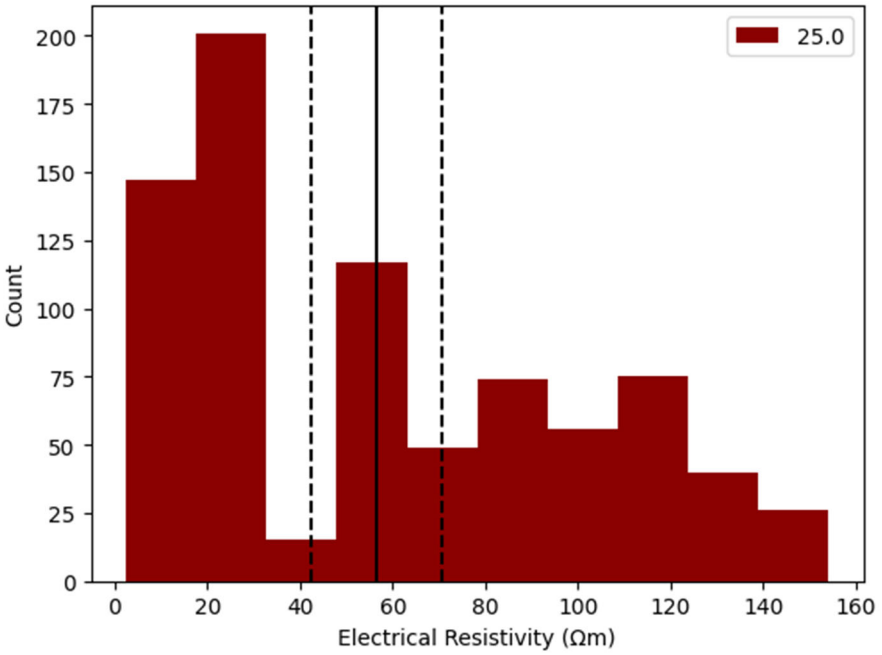
+/- 1% change from the mean for electrical resistivity



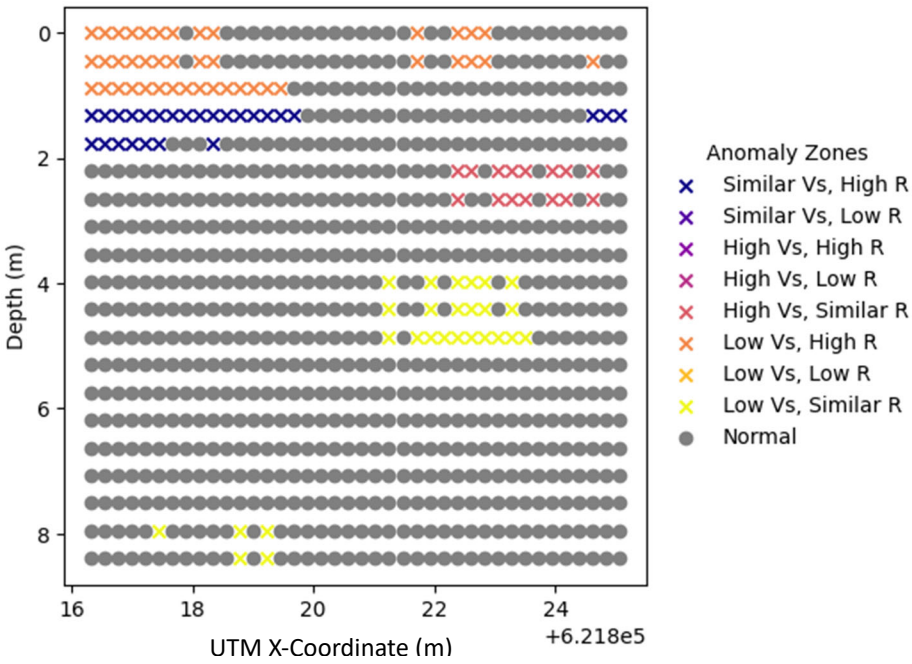
+/- 5% change from the mean for electrical resistivity



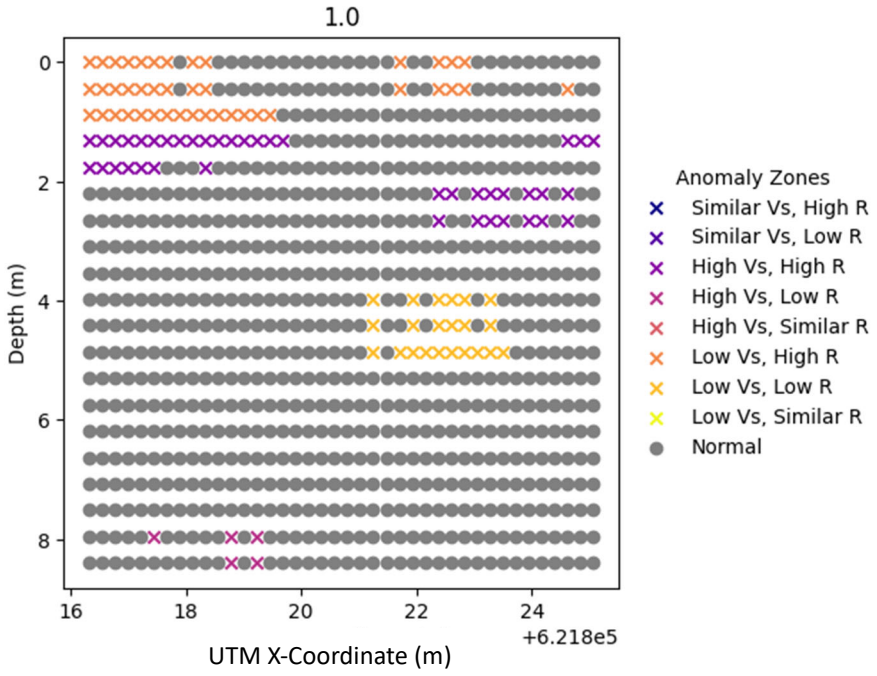
+/- 10% change from the mean for electrical resistivity



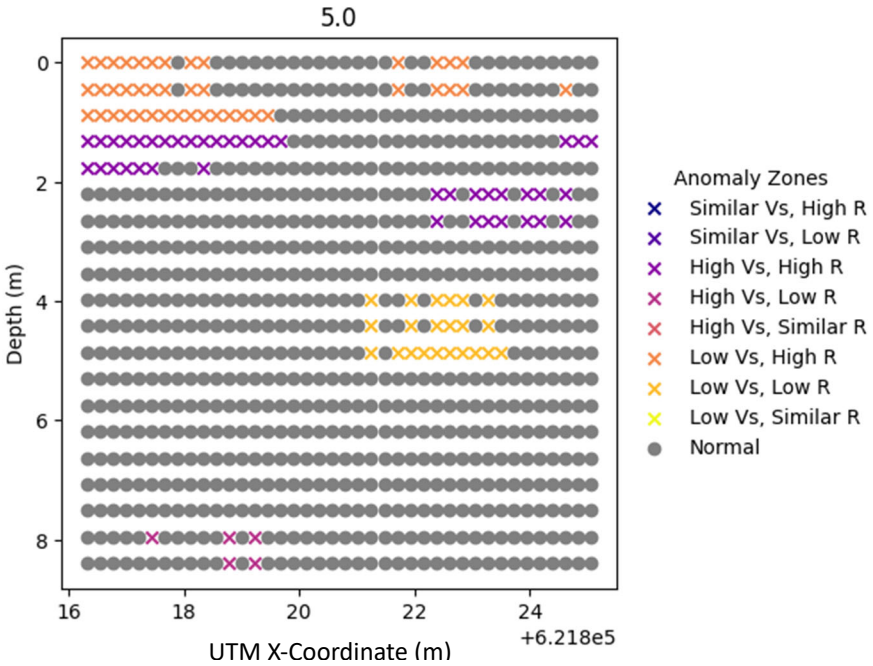
+/- 25% change from the mean for electrical resistivity



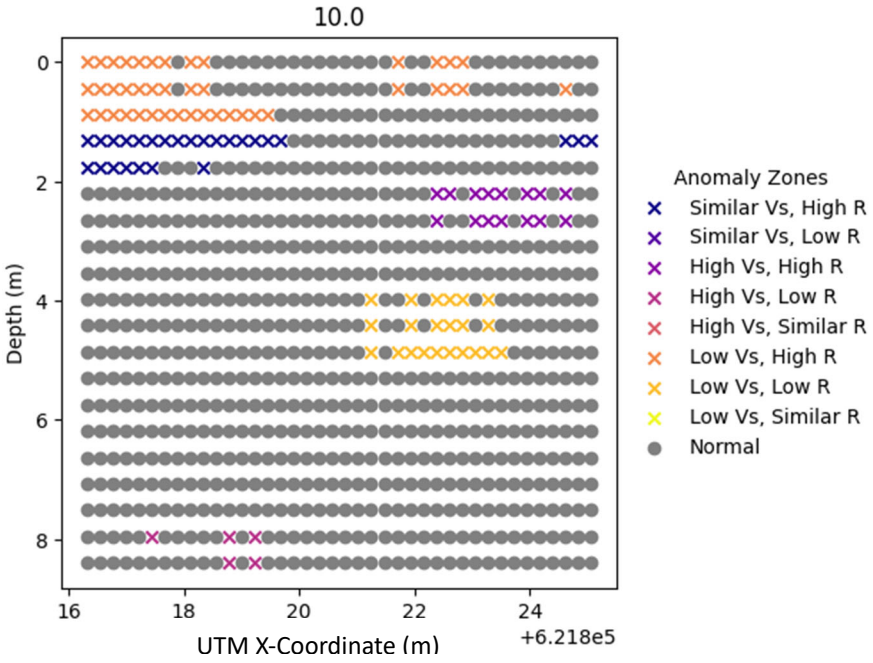
Subsurface profile with 1 standard deviation change threshold



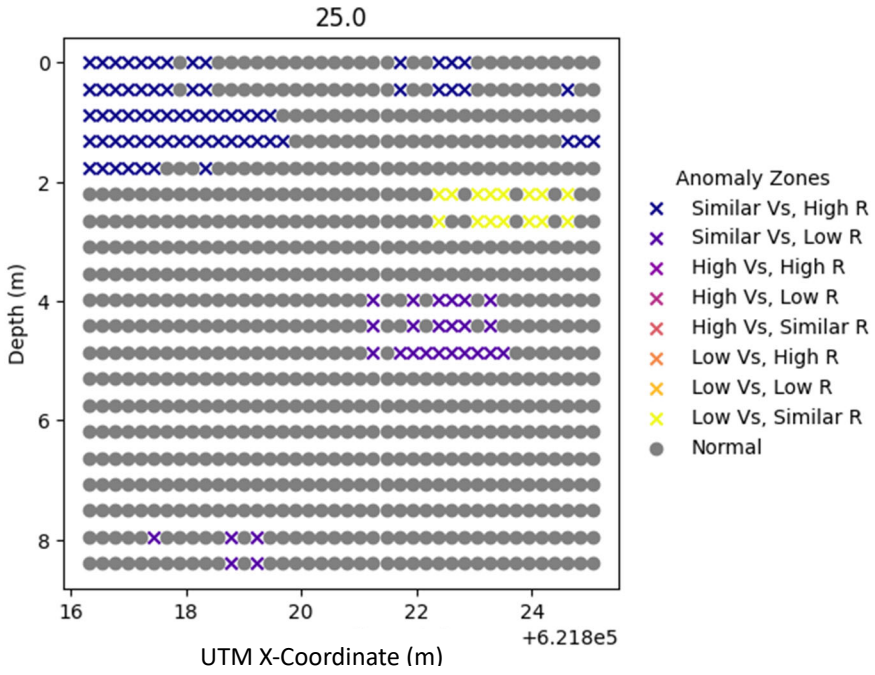
Subsurface profile with 1% change threshold



Subsurface profile with 5% change threshold

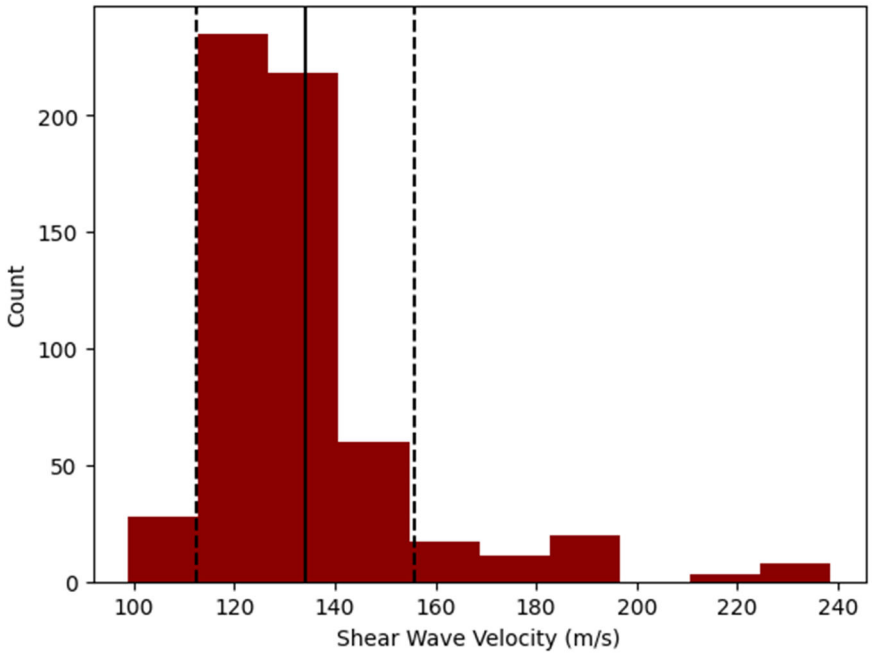


Subsurface profile with 10% change threshold

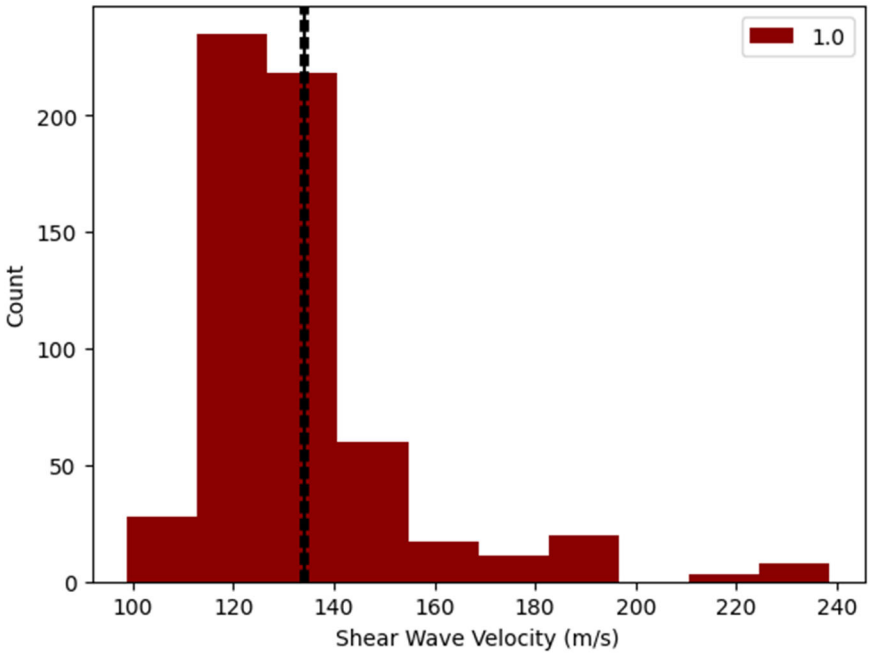


Subsurface profile with 25% change threshold

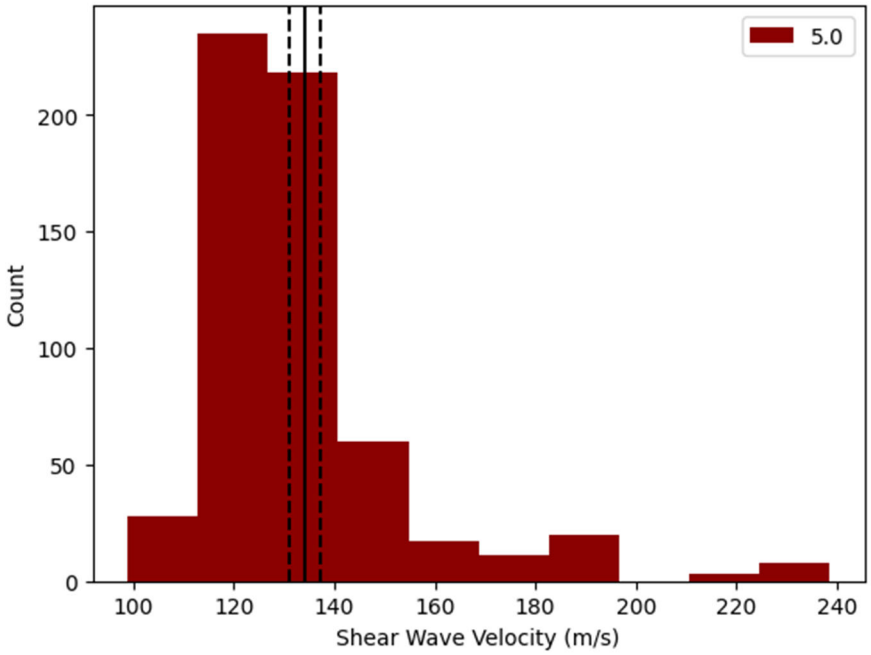
Case Study 2: swt_site1



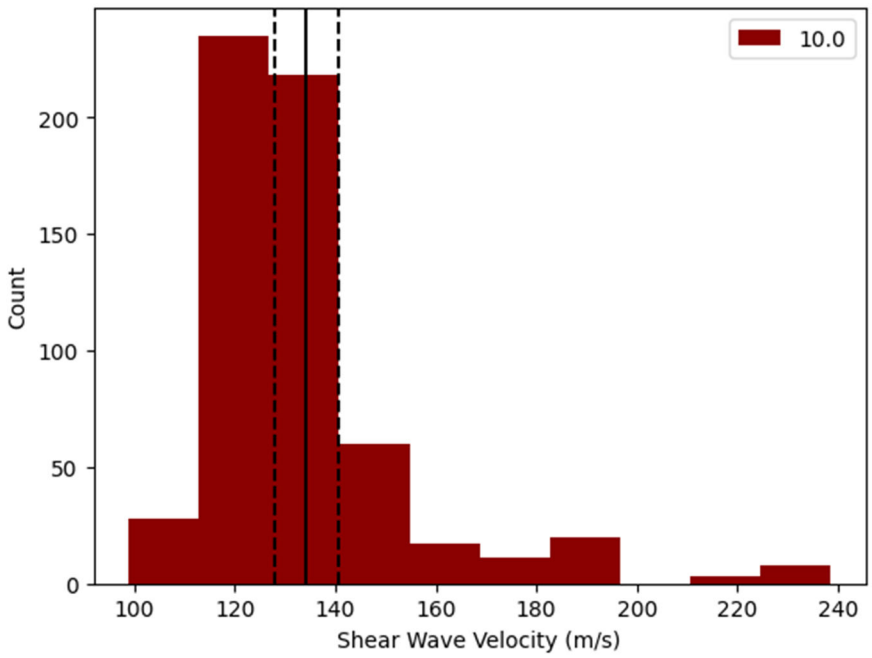
+/- 1 Standard deviation of change from the mean for shear wave velocity



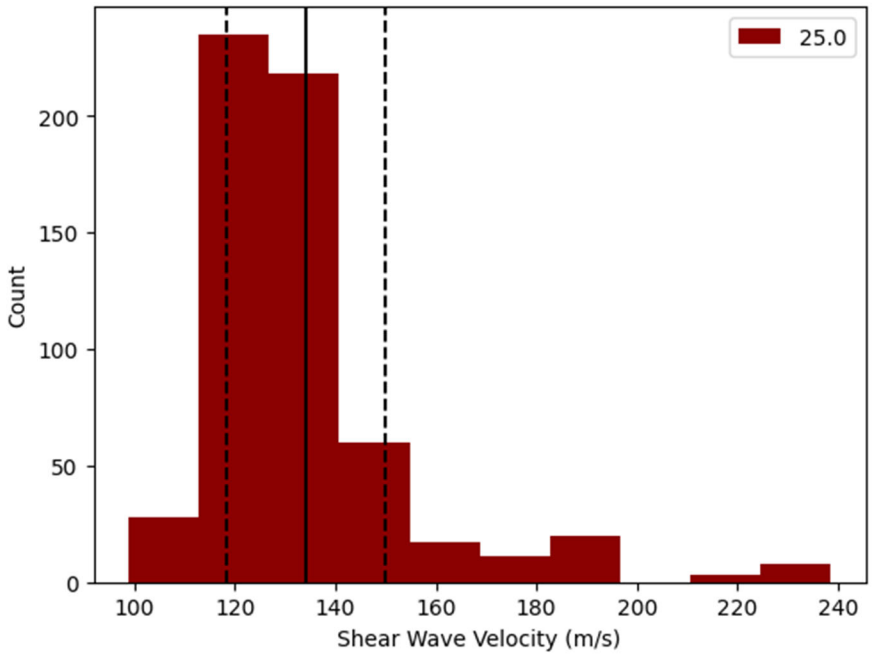
+/- 1% change from the mean for shear wave velocity



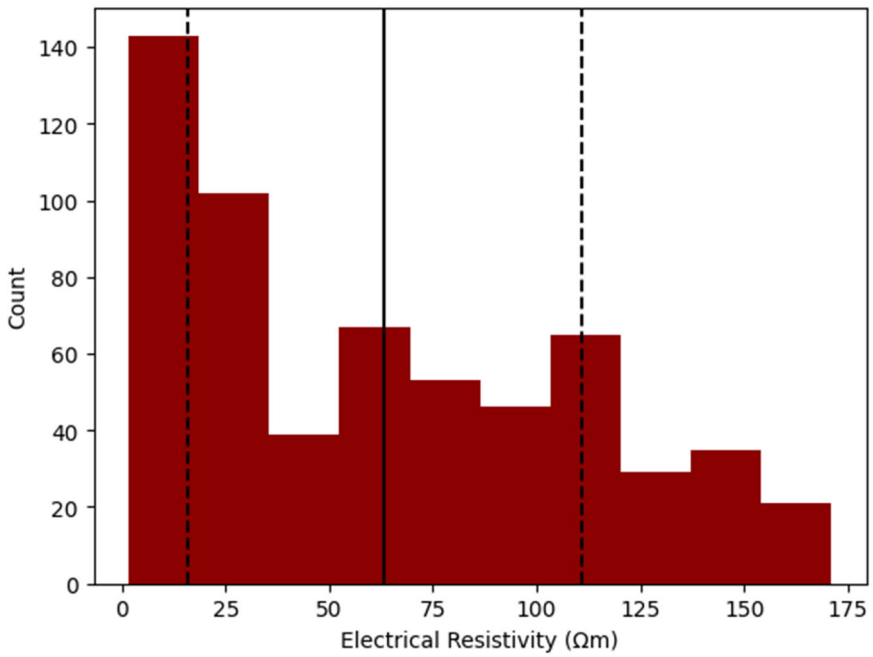
+/- 5% change from the mean for shear wave velocity



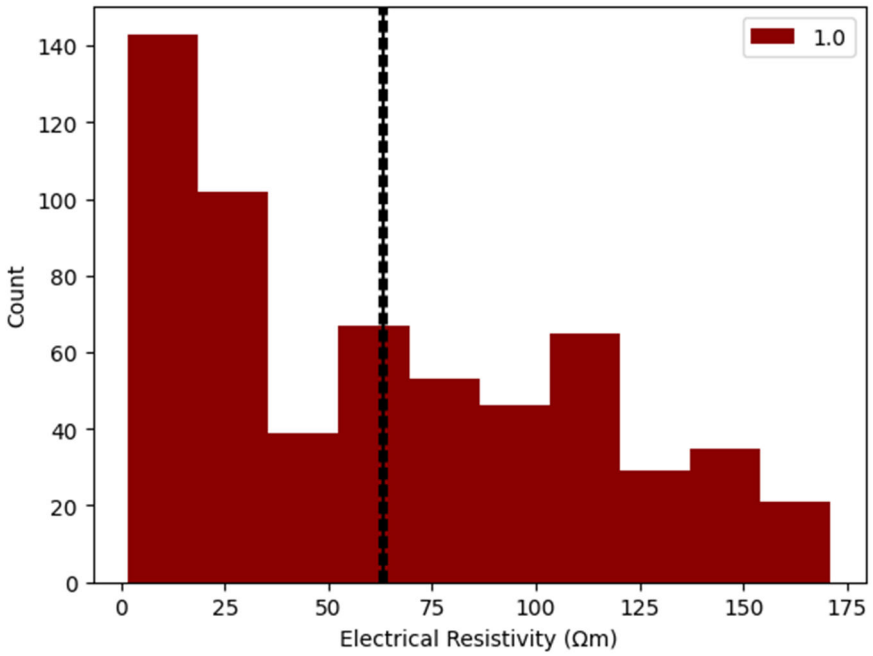
+/- 10% change from the mean for shear wave velocity



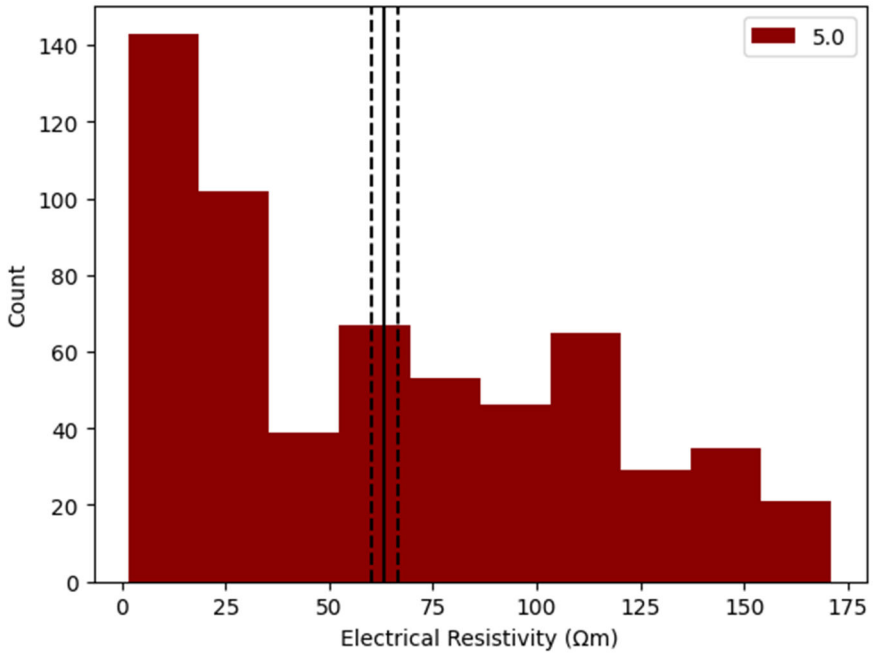
+/- 25% change from the mean for shear wave velocity



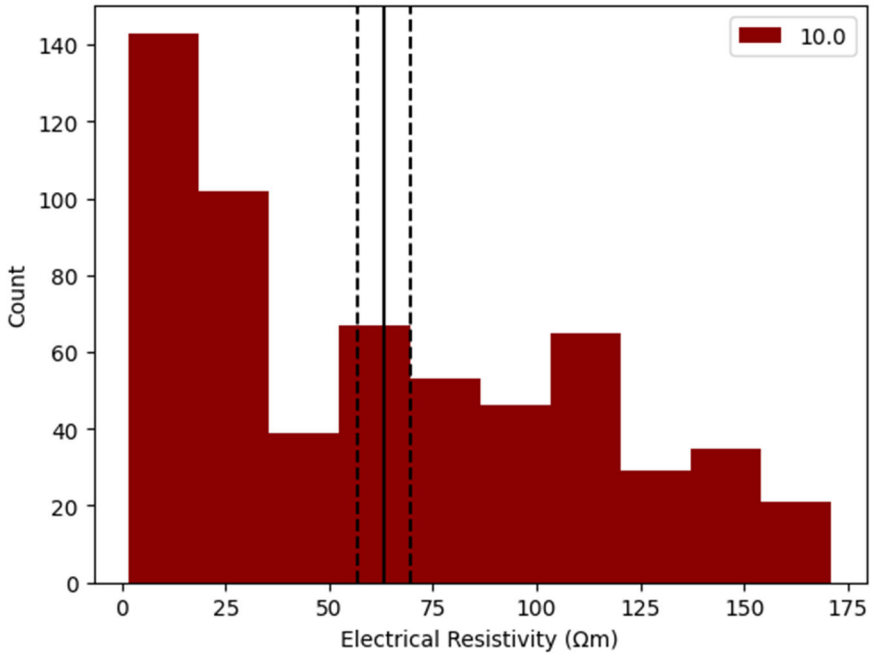
+/- 1 Standard deviation of change from the mean for electrical resistivity



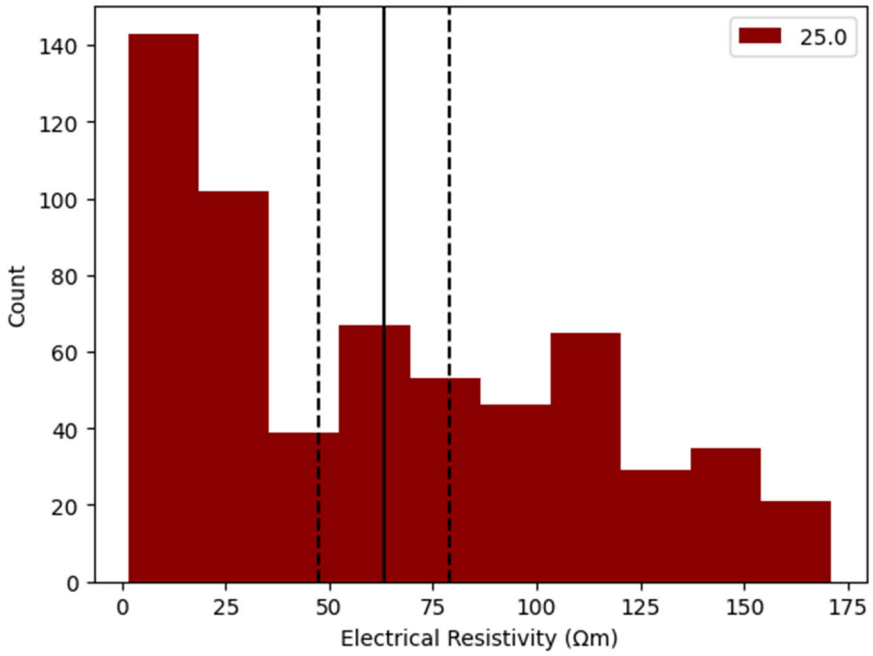
+/- 1% change from the mean for electrical resistivity



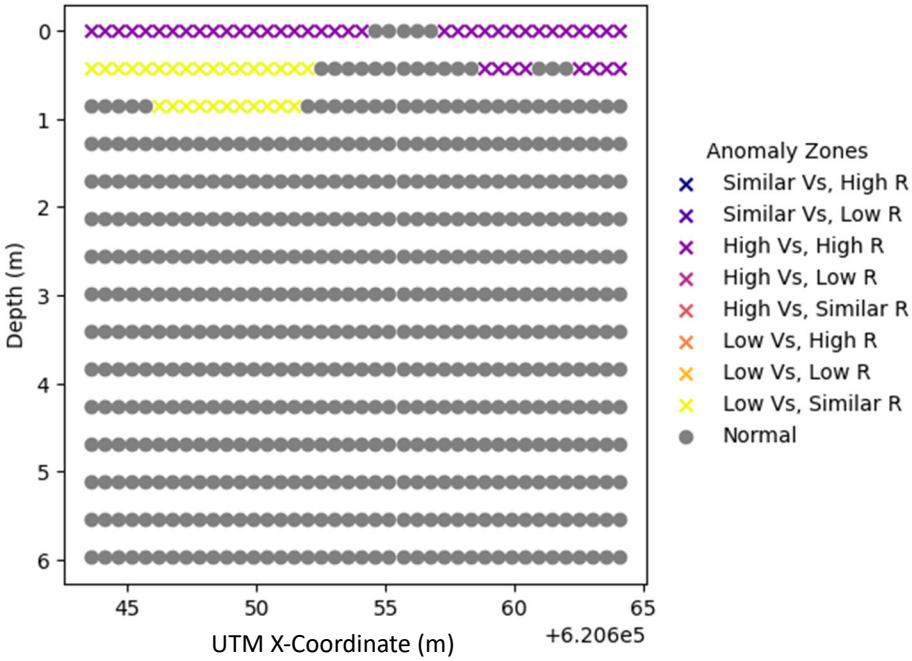
+/- 5% change from the mean for electrical resistivity



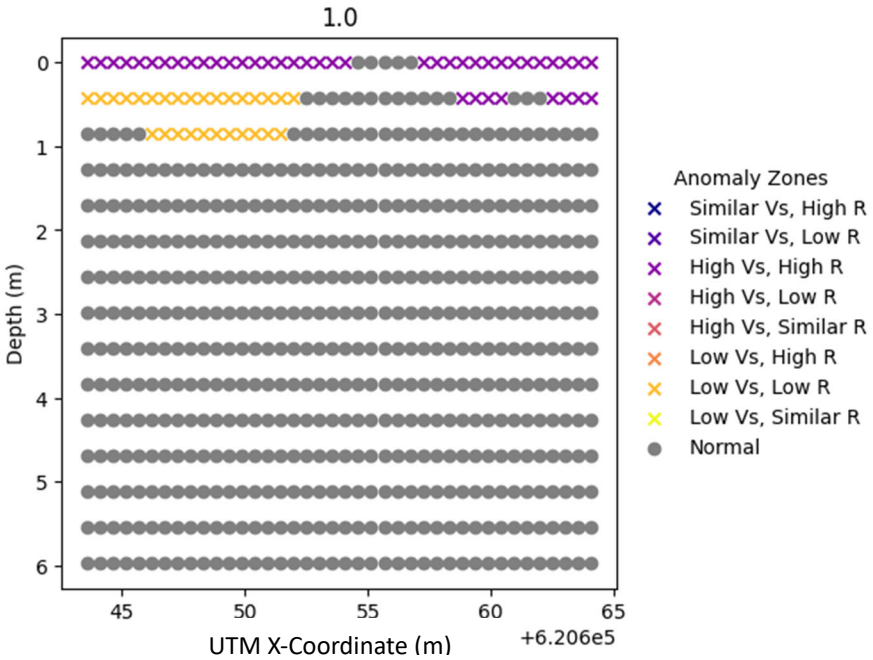
+/- 10% change from the mean for electrical resistivity



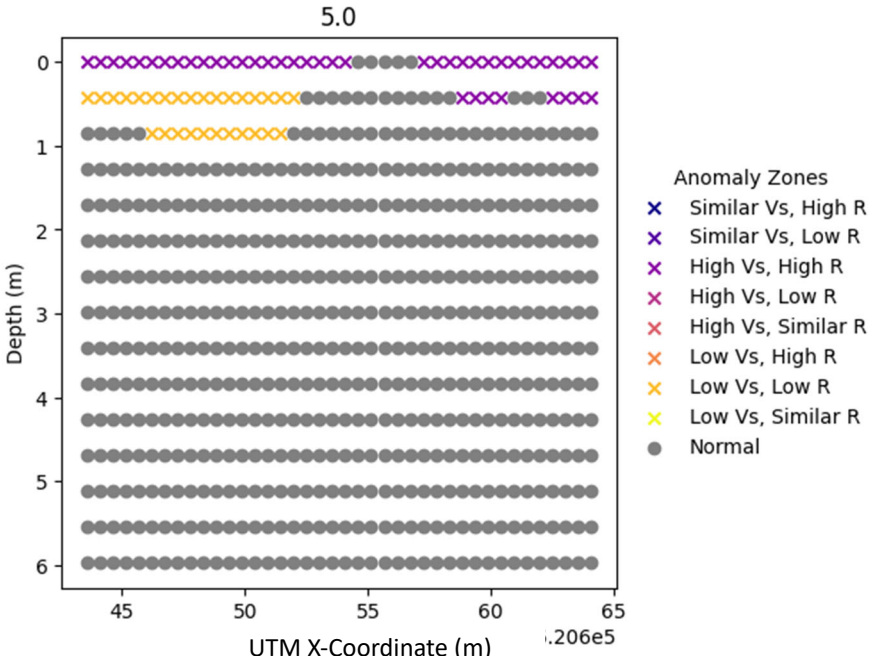
+/- 25% change from the mean for electrical resistivity



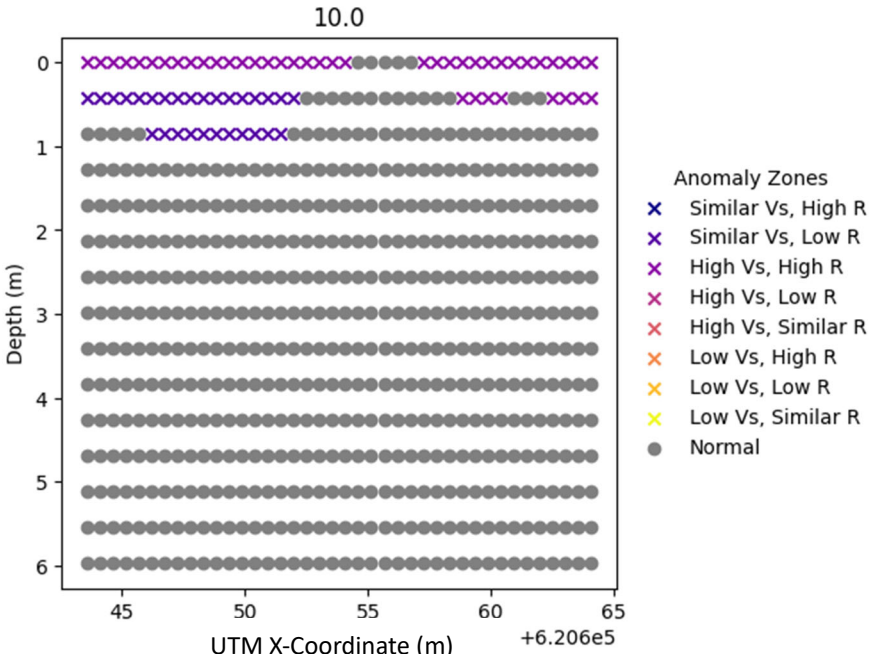
Subsurface profile with 1 standard deviation change threshold



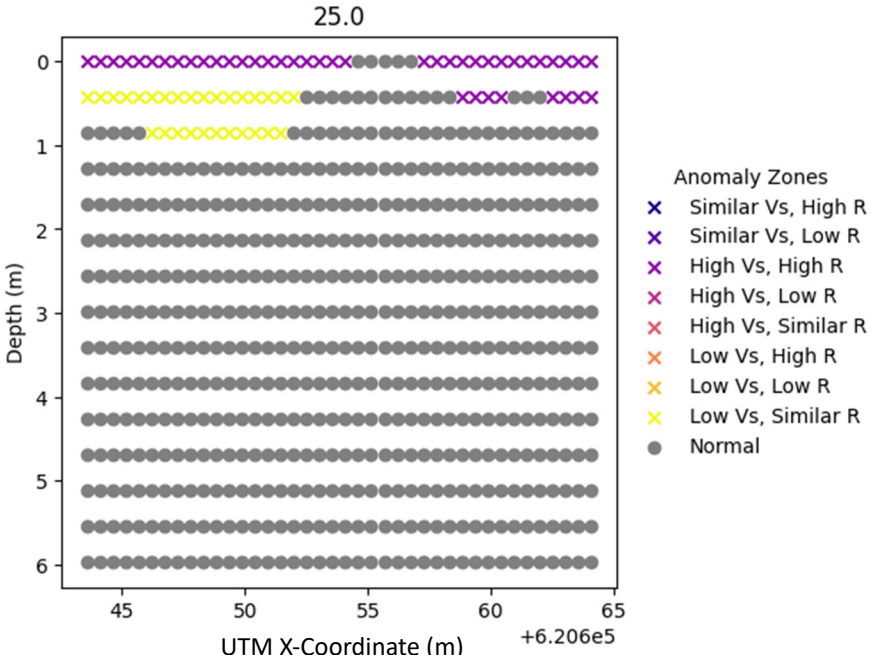
Subsurface profile with 1% change threshold



Subsurface profile with 5% change threshold

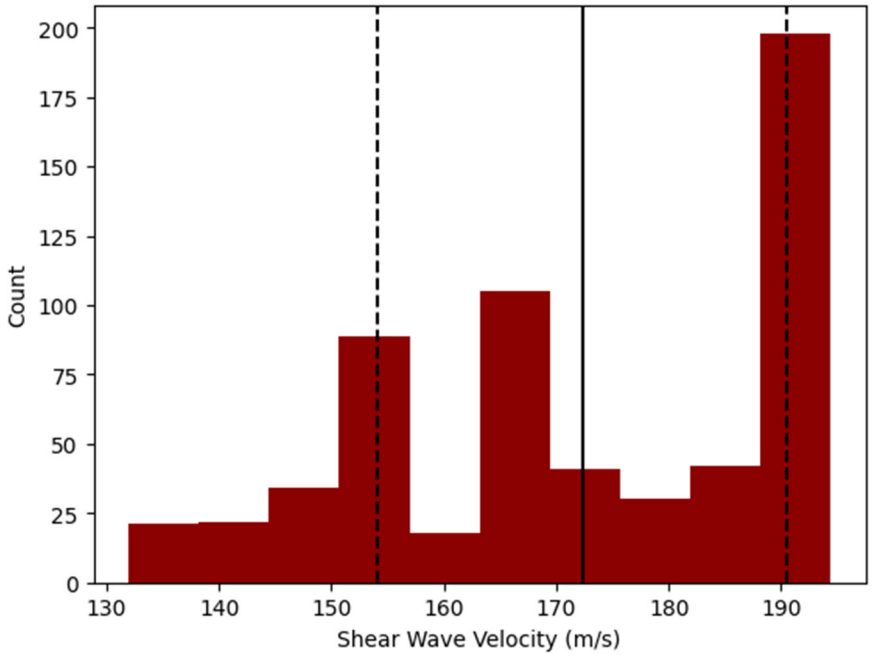


Subsurface profile with 10% change threshold

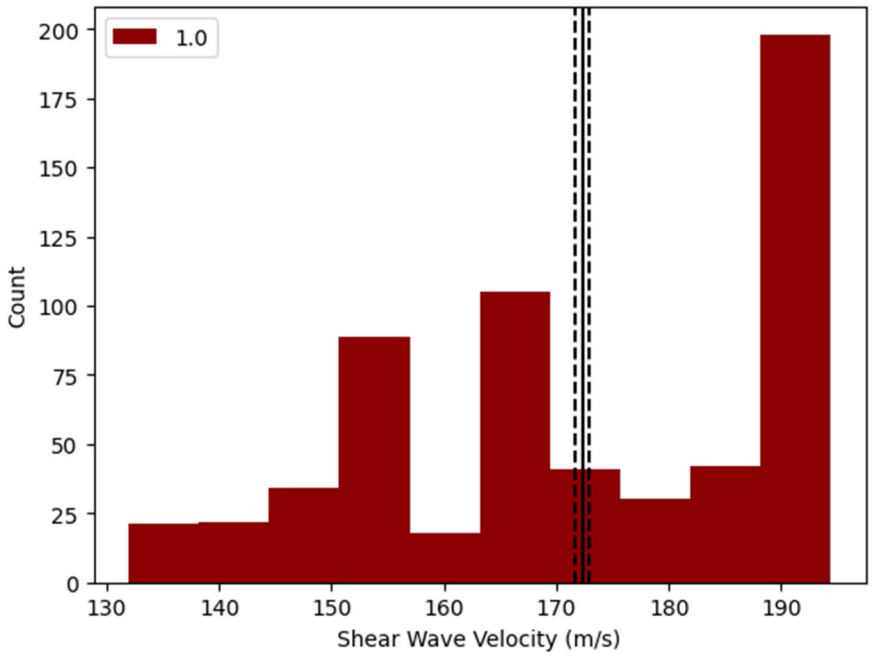


Subsurface profile with 25% change threshold

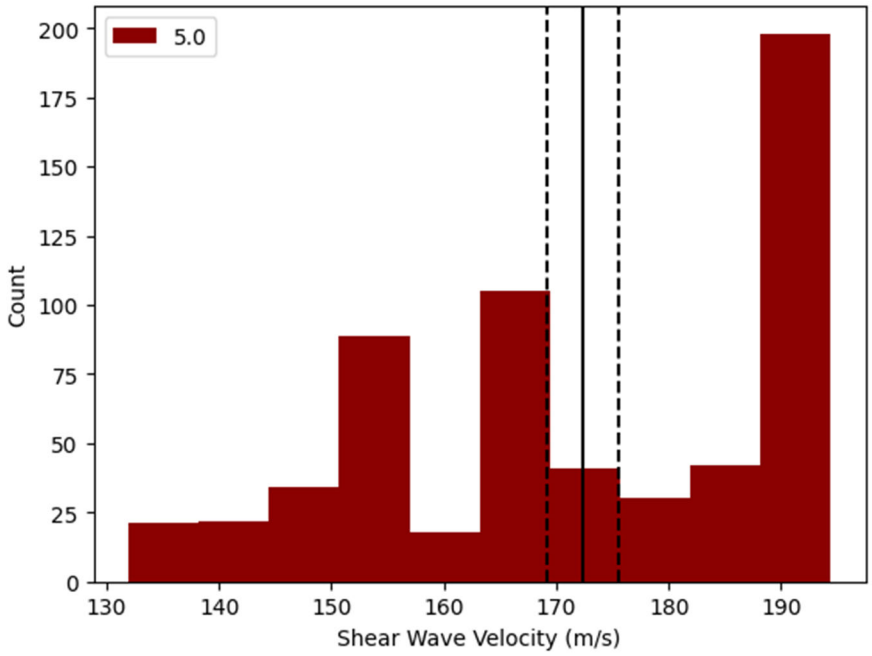
Case Study 3: sc_site2



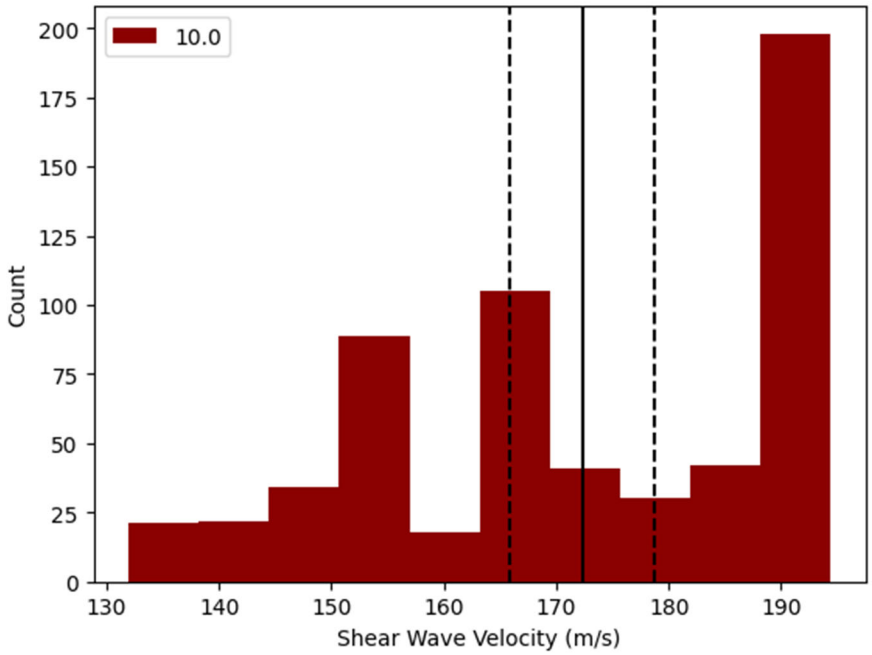
+/- 1 Standard deviation of change from the mean for shear wave velocity



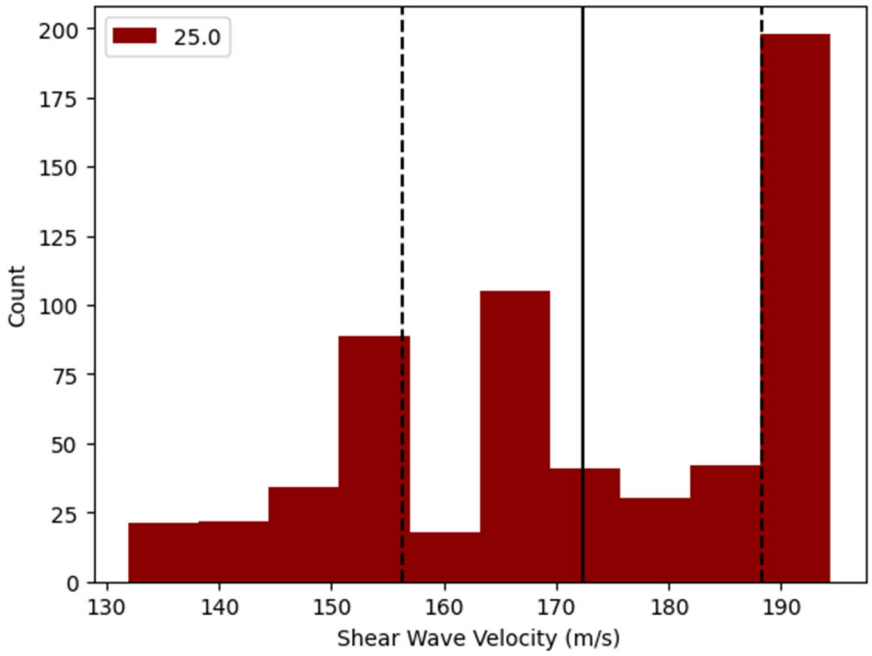
+/- 1 % change from the mean for shear wave velocity



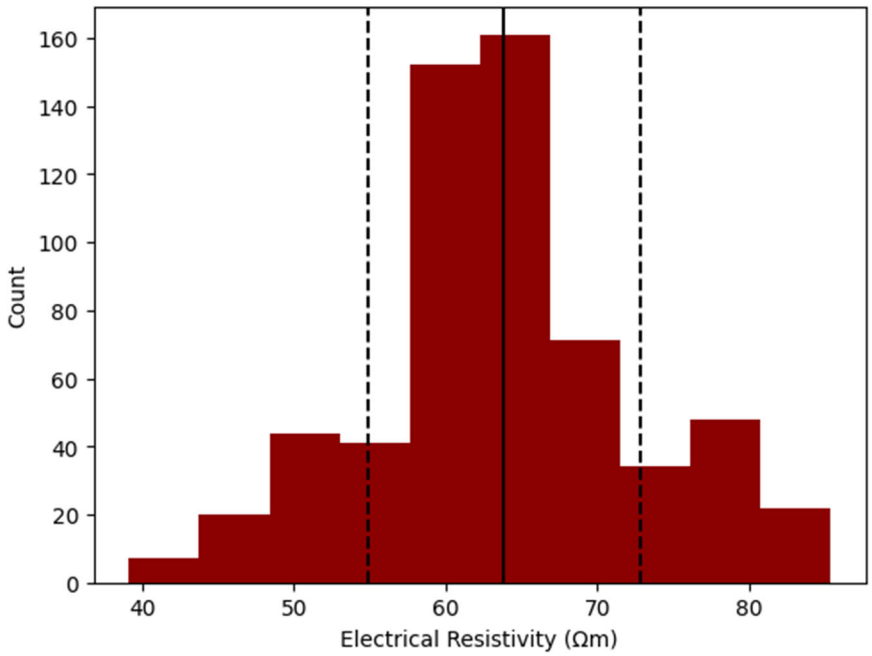
+/- 5% change from the mean for shear wave velocity



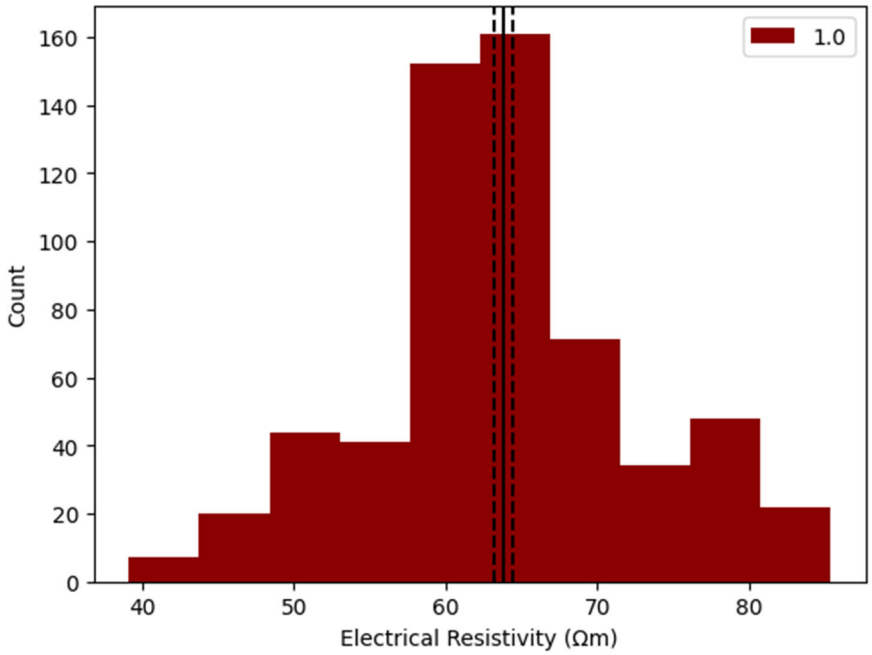
+/- 10% change from the mean for shear wave velocity



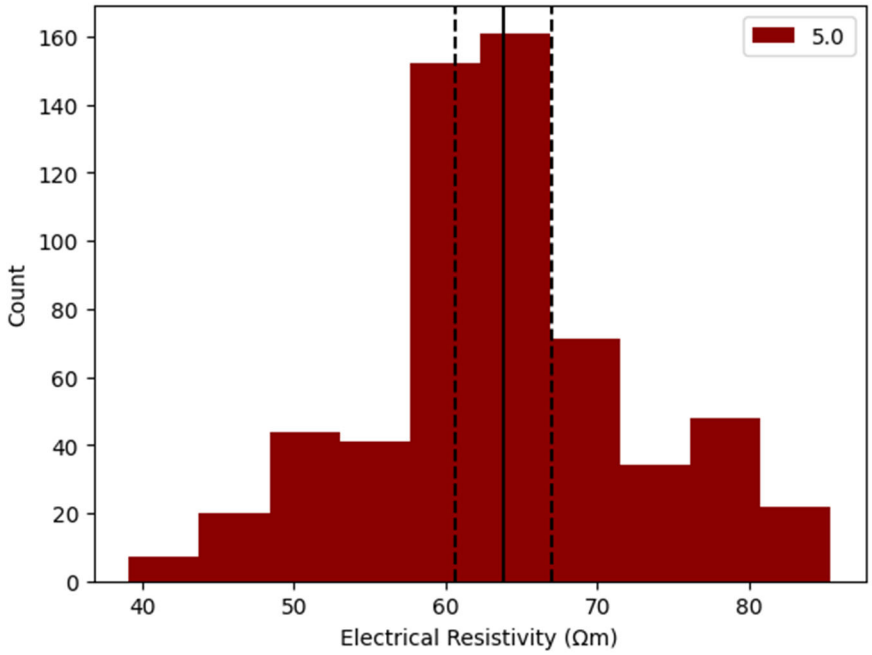
+/- 25% change from the mean for shear wave velocity



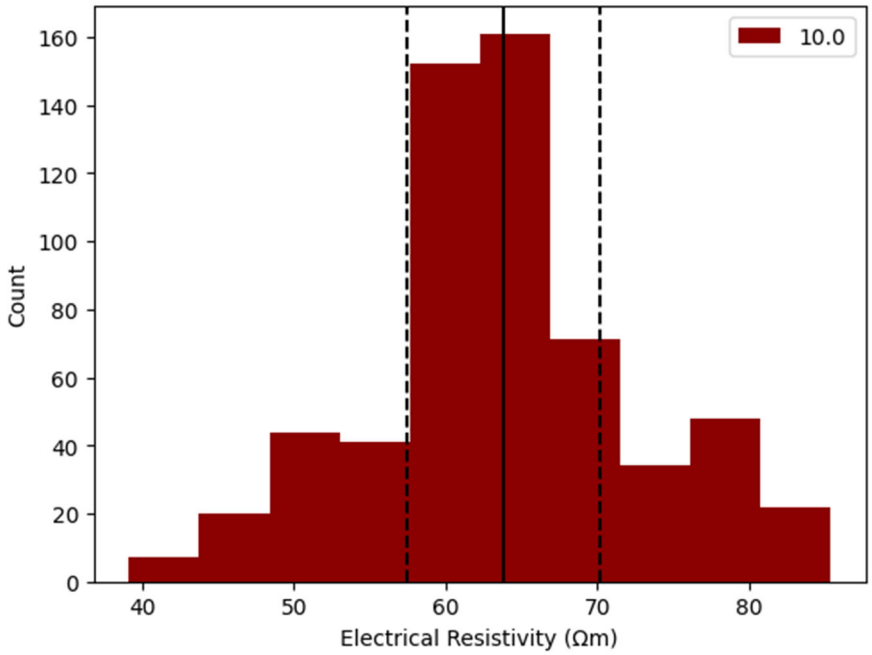
+/- 1 Standard deviation of change from the mean for electrical resistivity



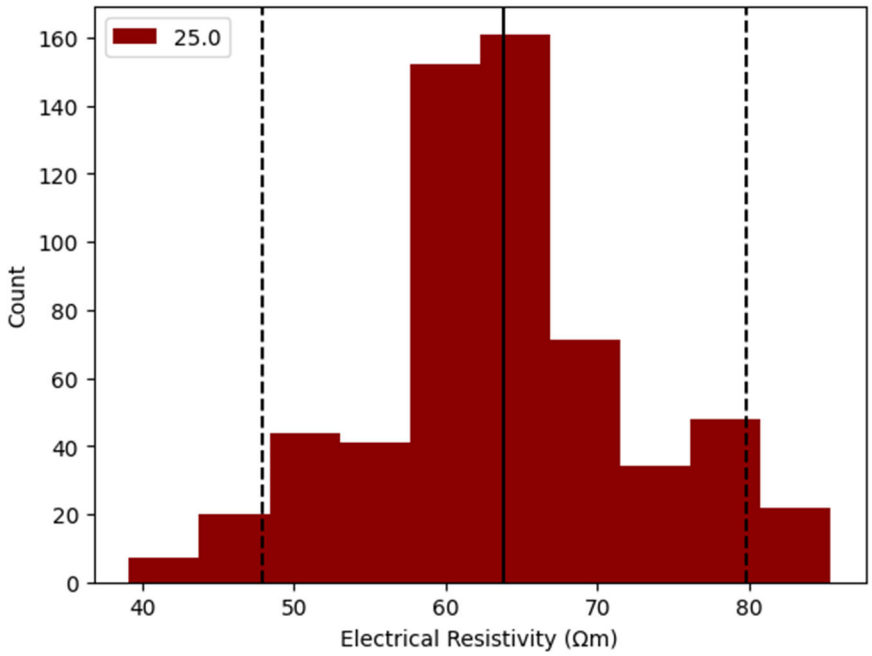
+/- 1% change from the mean for electrical resistivity



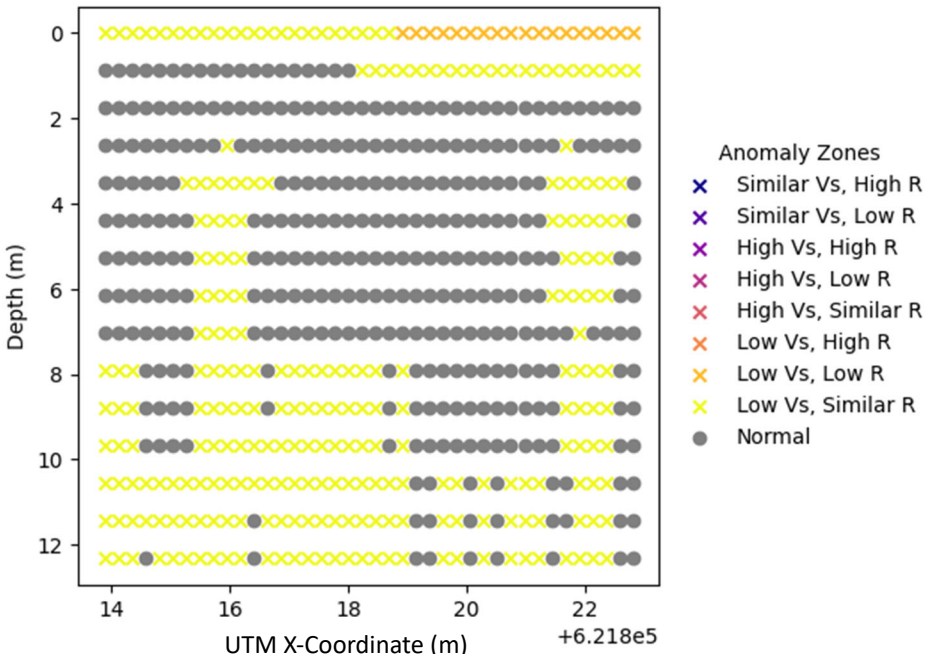
+/- 5% change from the mean for electrical resistivity



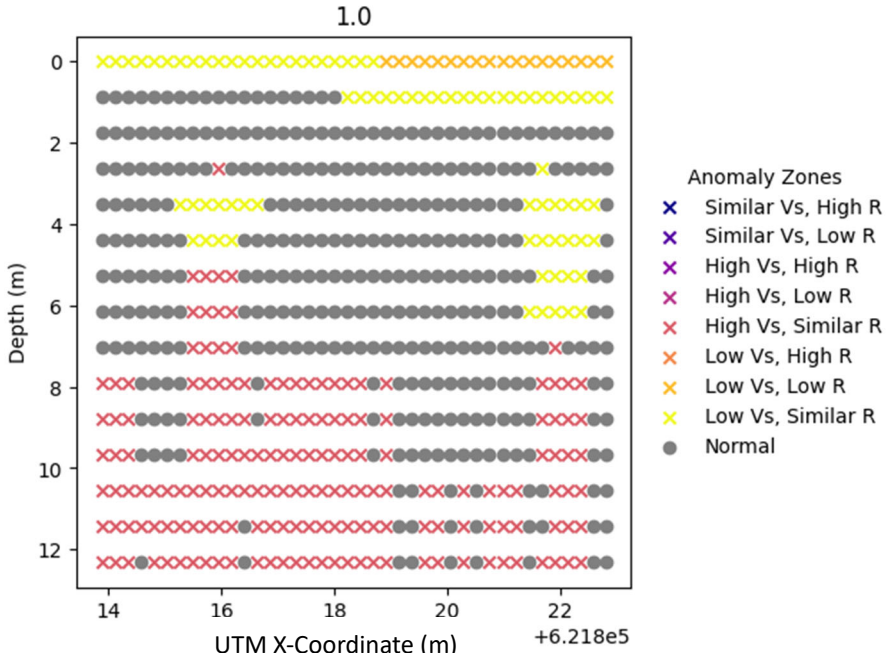
+/- 10% change from the mean for electrical resistivity



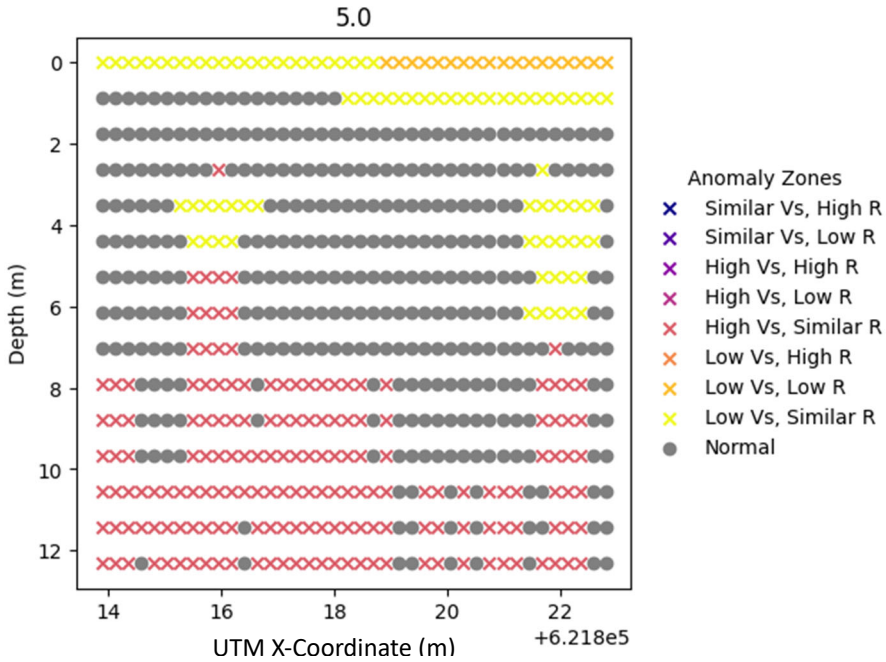
+/- 25% change from the mean for electrical resistivity



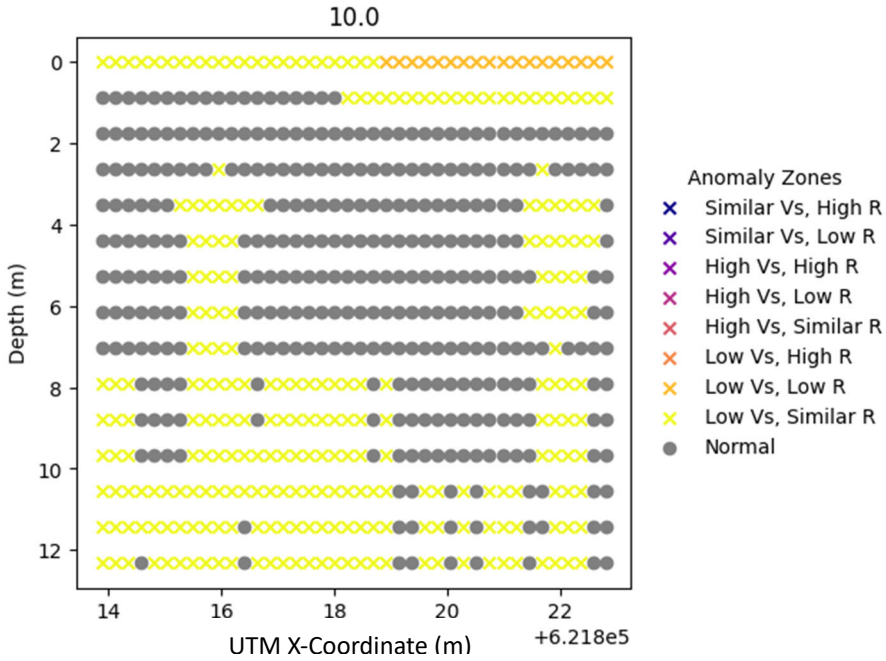
Subsurface profile with 1 standard deviation change threshold



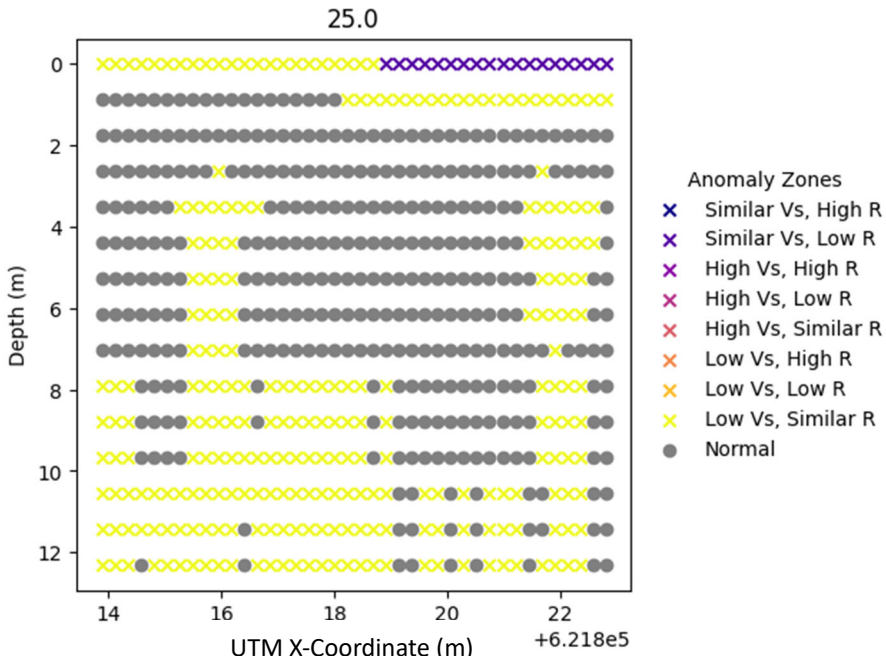
Subsurface profile with 1% change threshold



Subsurface profile with 5% change threshold



Subsurface profile with 10% change threshold

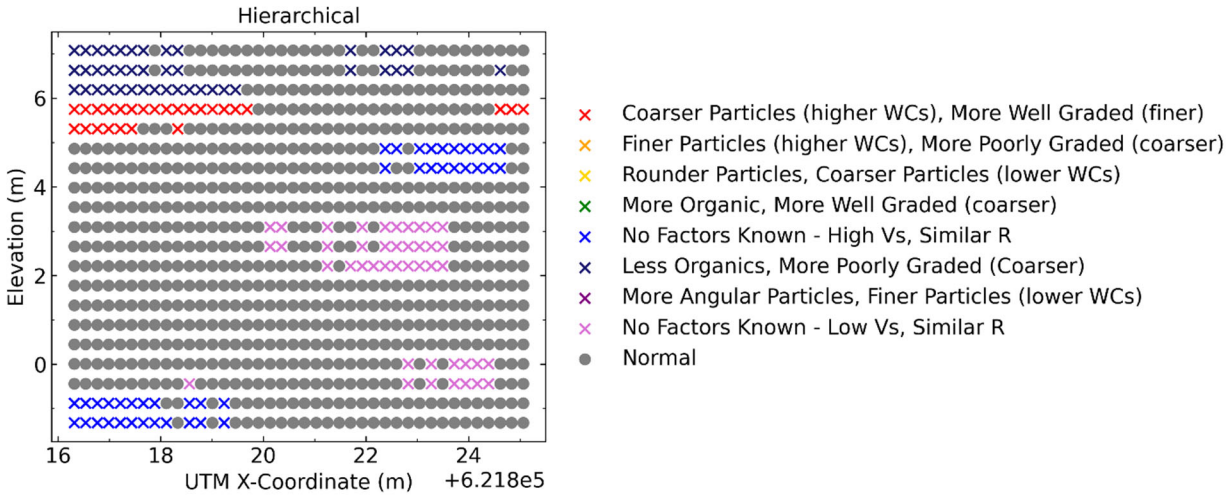


Subsurface profile with 25% change threshold

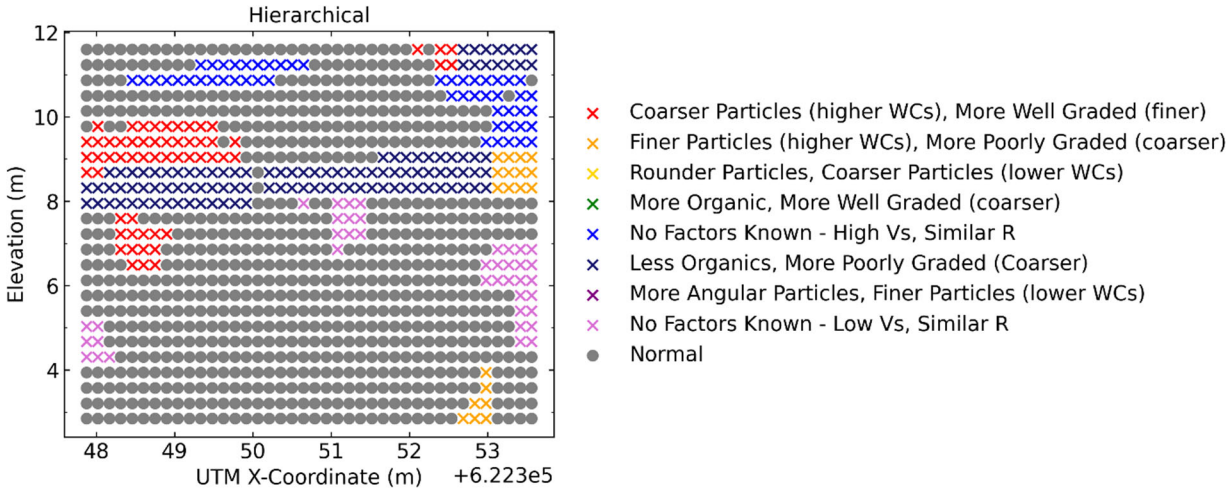
Clustering and Anomaly Detection Profiles

Fit: Individual Compilation
Algorithm: Hierarchical

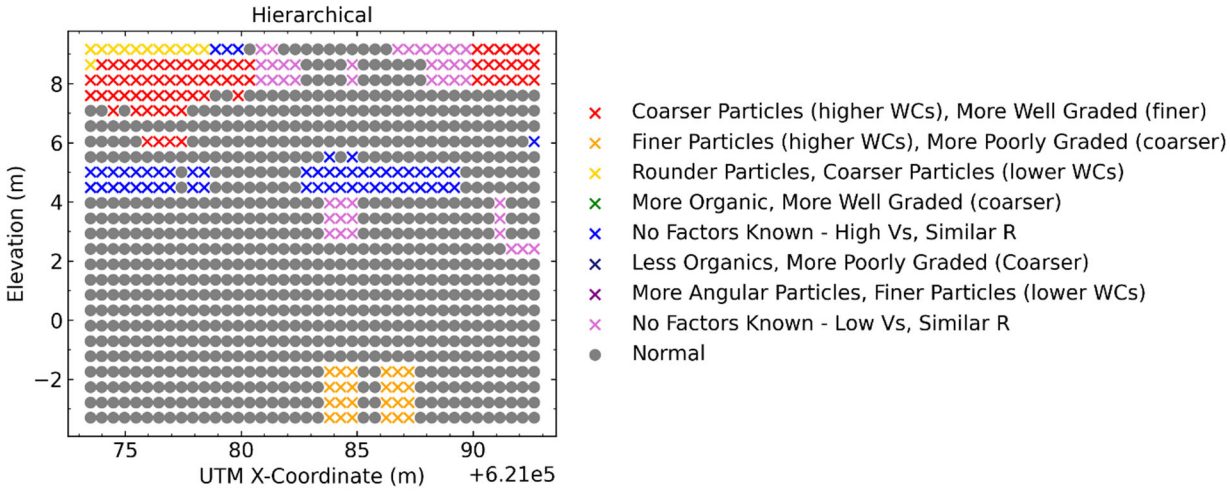
Profile: fc_site2



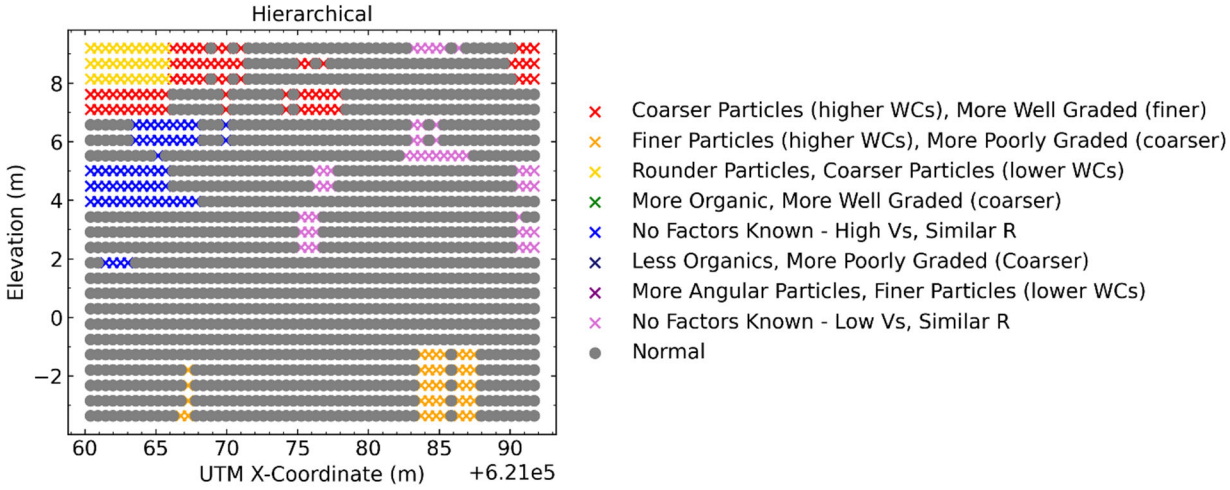
Profile: fc_site3



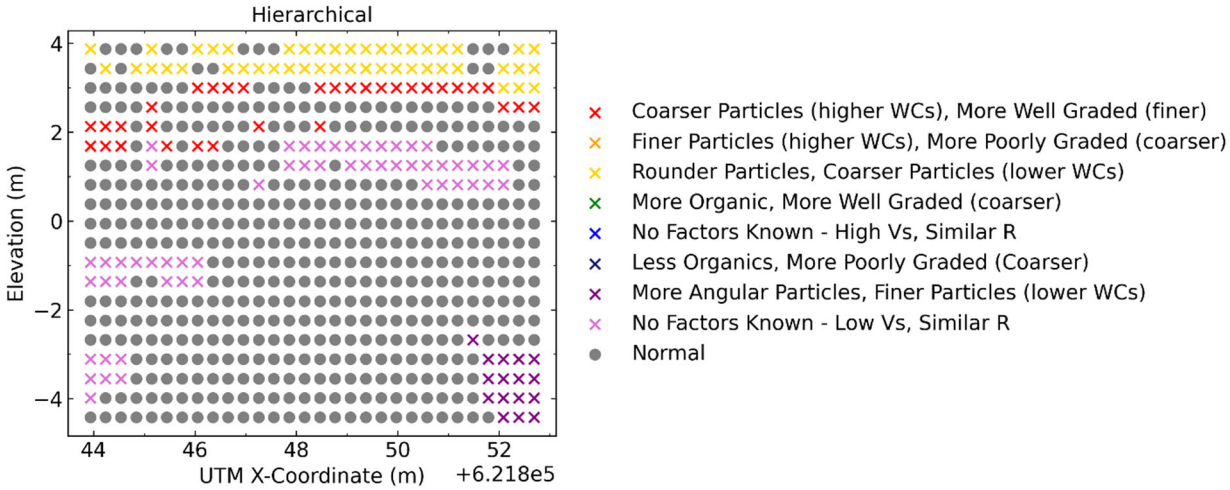
Profile: fc3_site1



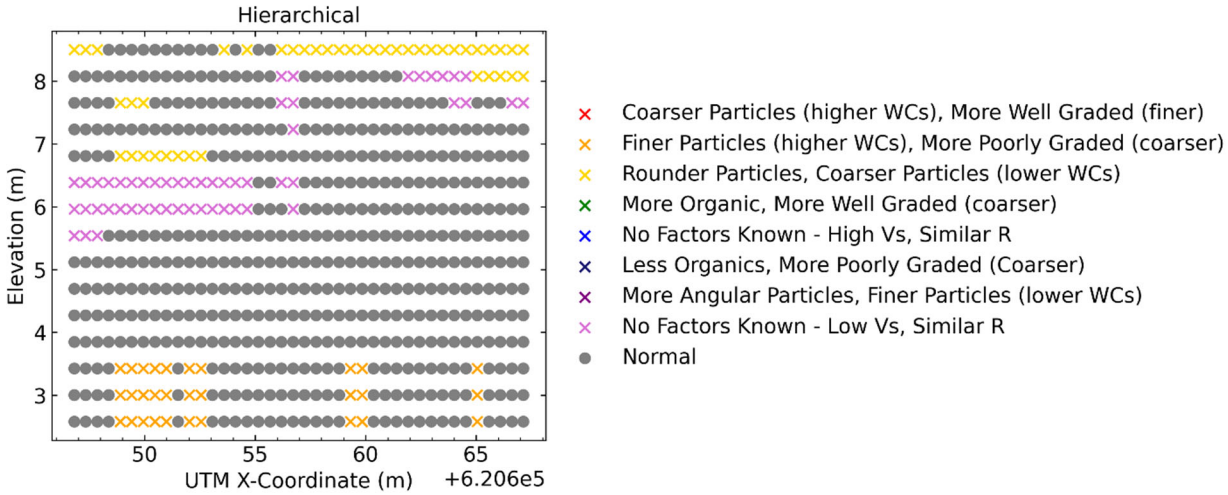
Profile: fc5_site1



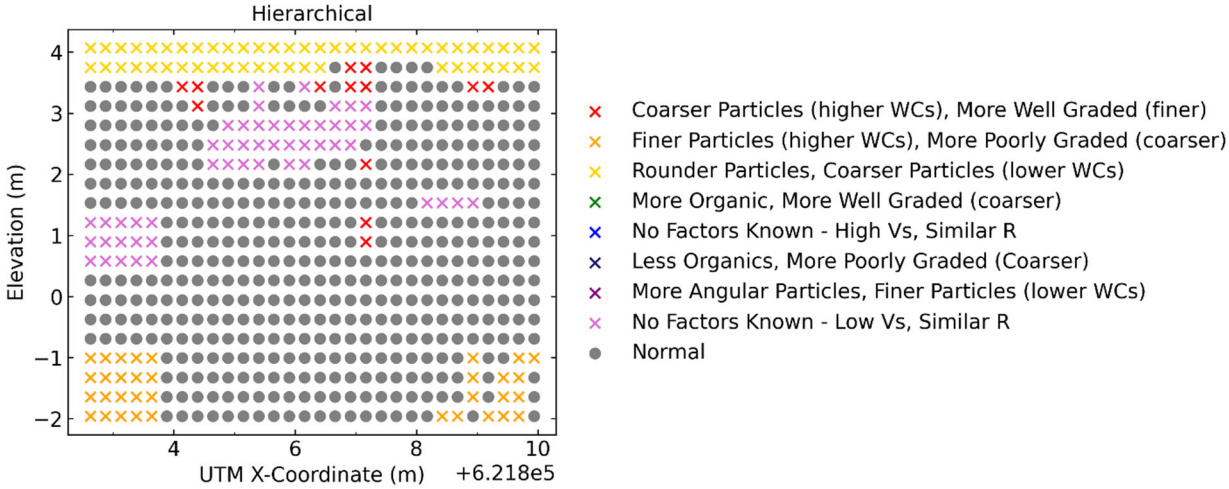
Profile: fdt_site2



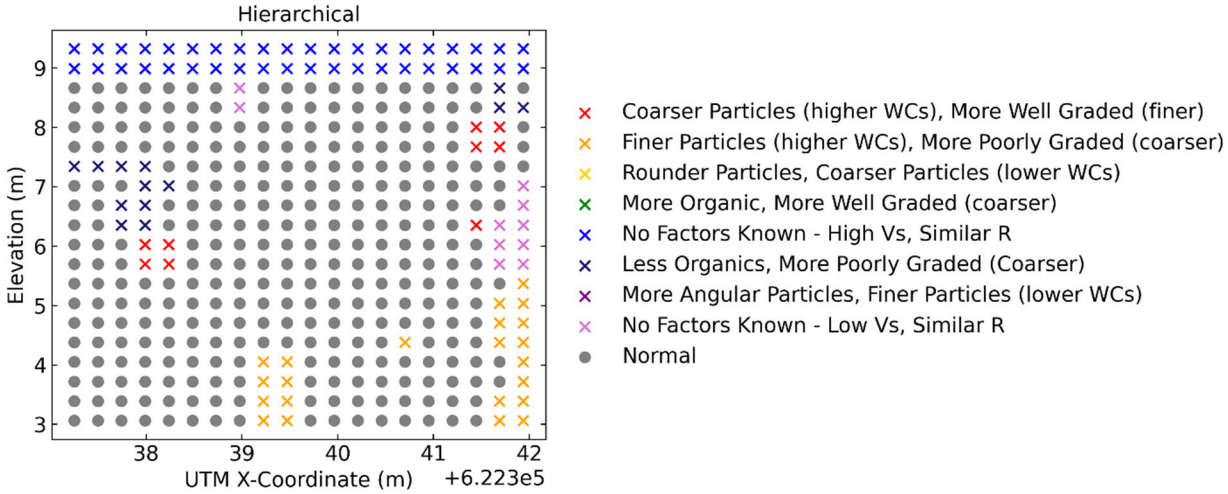
Profile: fwt_site1



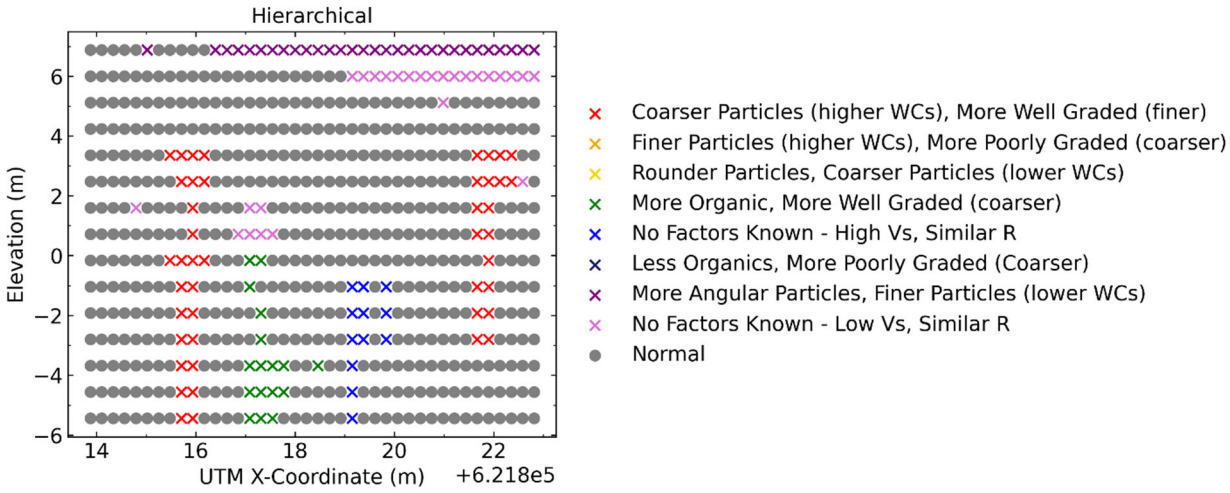
Profile: fwt_site2



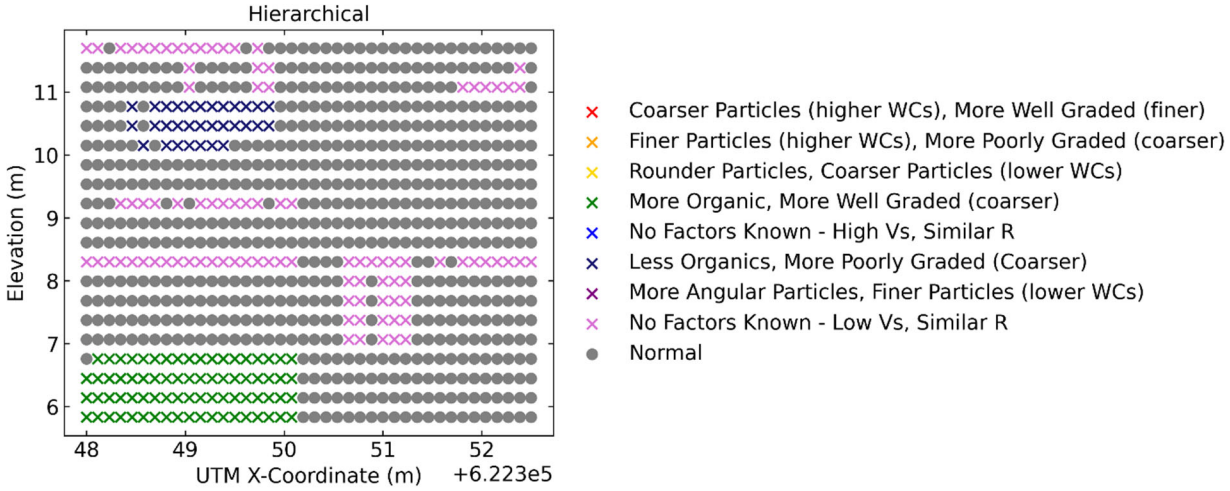
Profile: fwt_site3



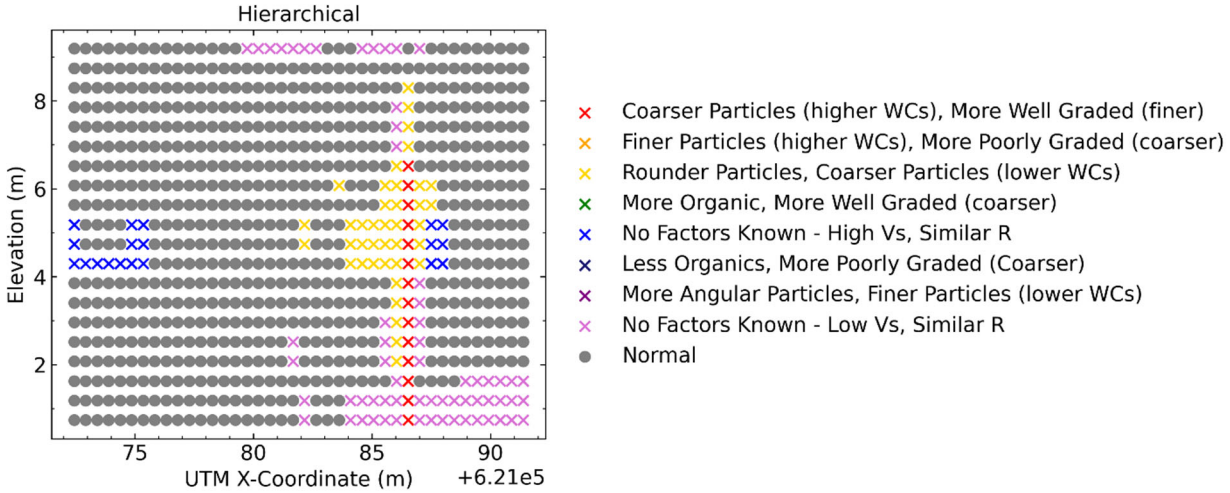
Profile: sc_site2



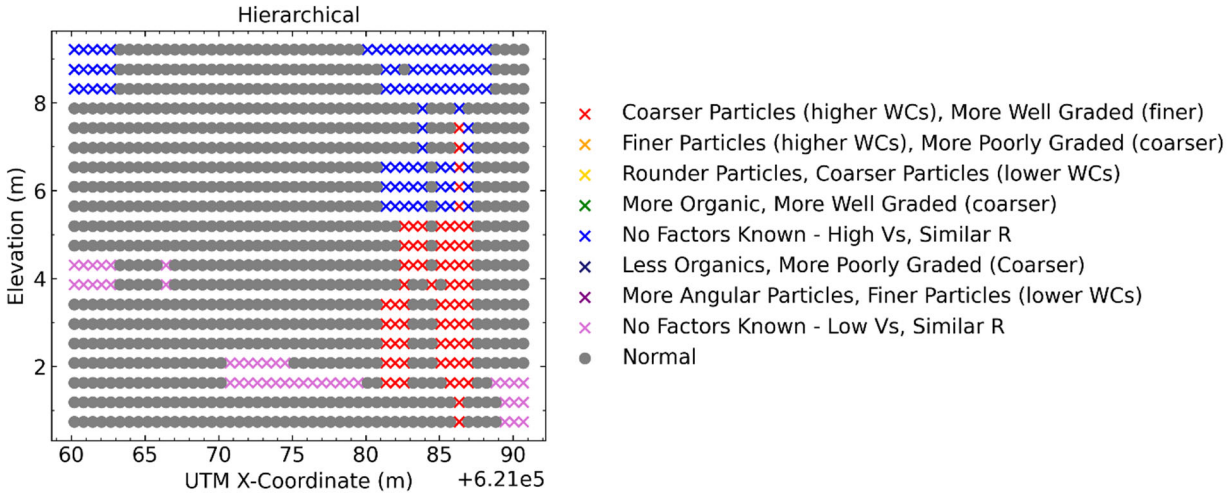
Profile: sc_site3



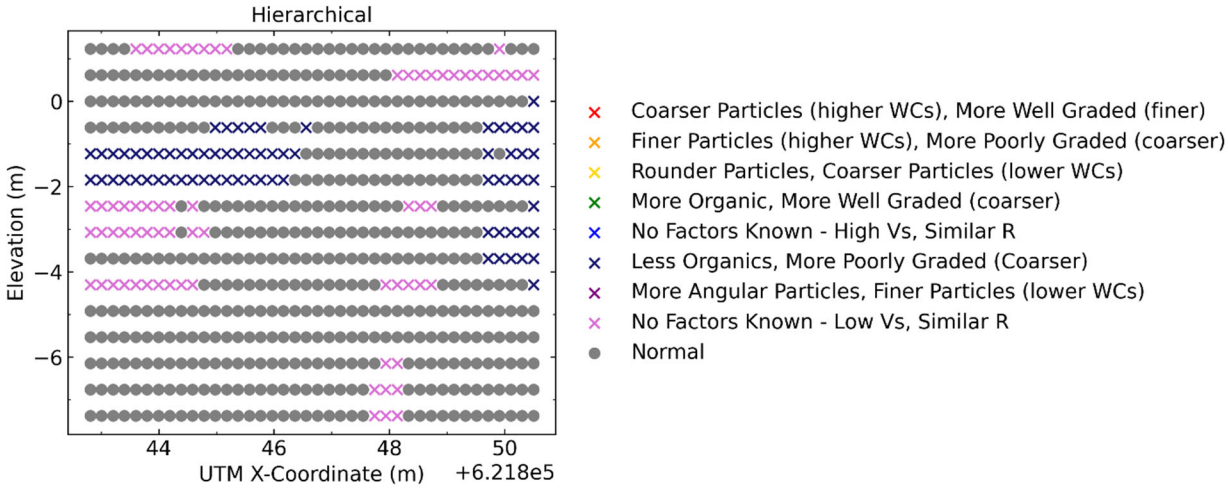
Profile: sc3_site1



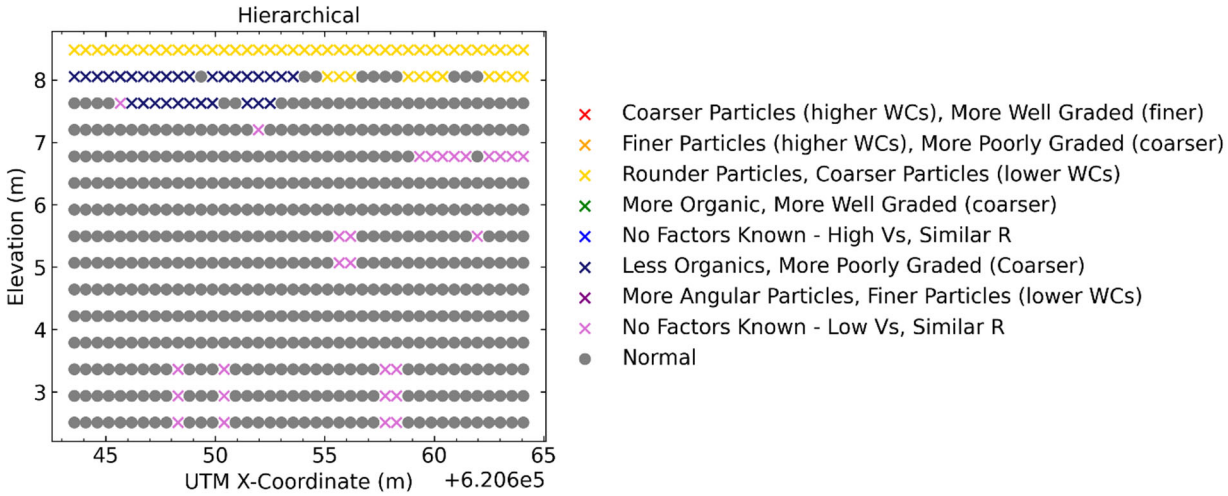
Profile: sc5_site1



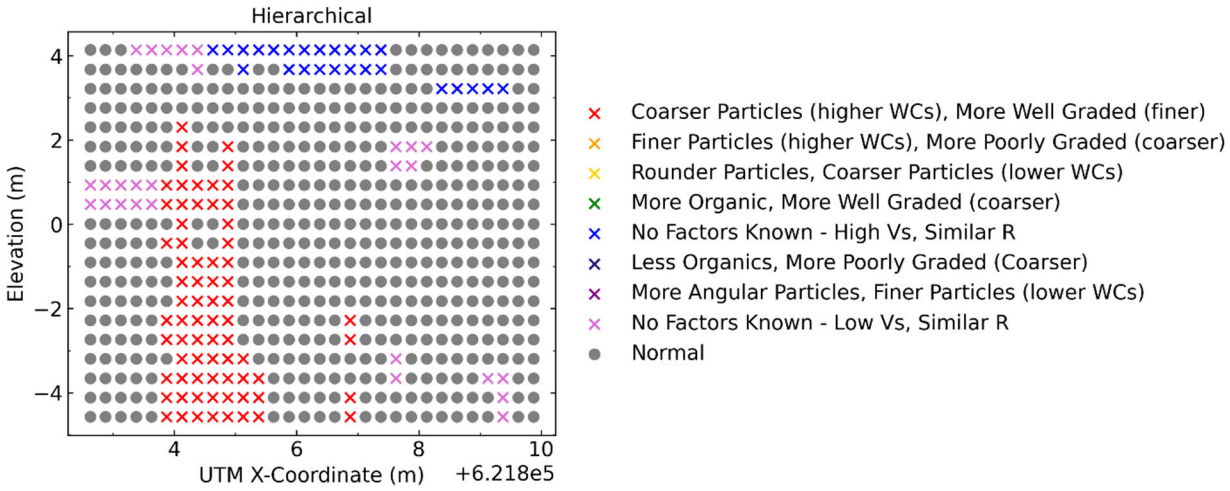
Profile: sdt_site2



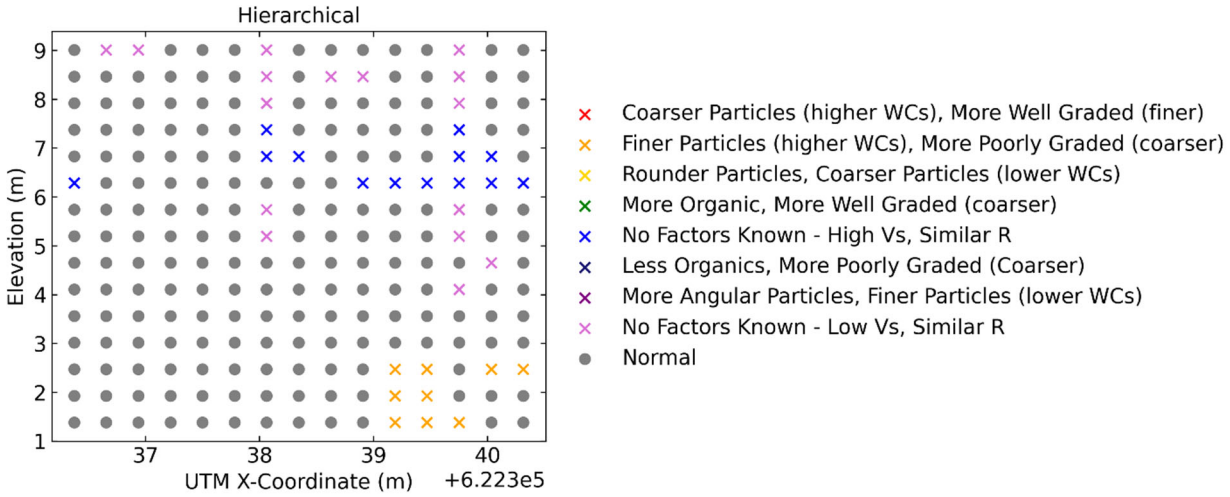
Profile: swt_site1



Profile: swt_site2

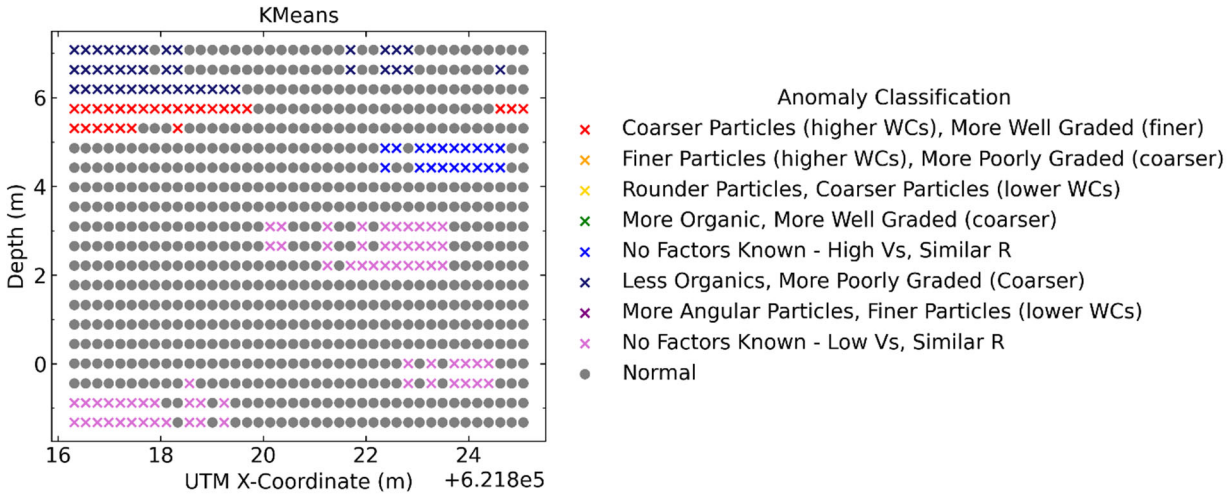


Profile: swt_site3

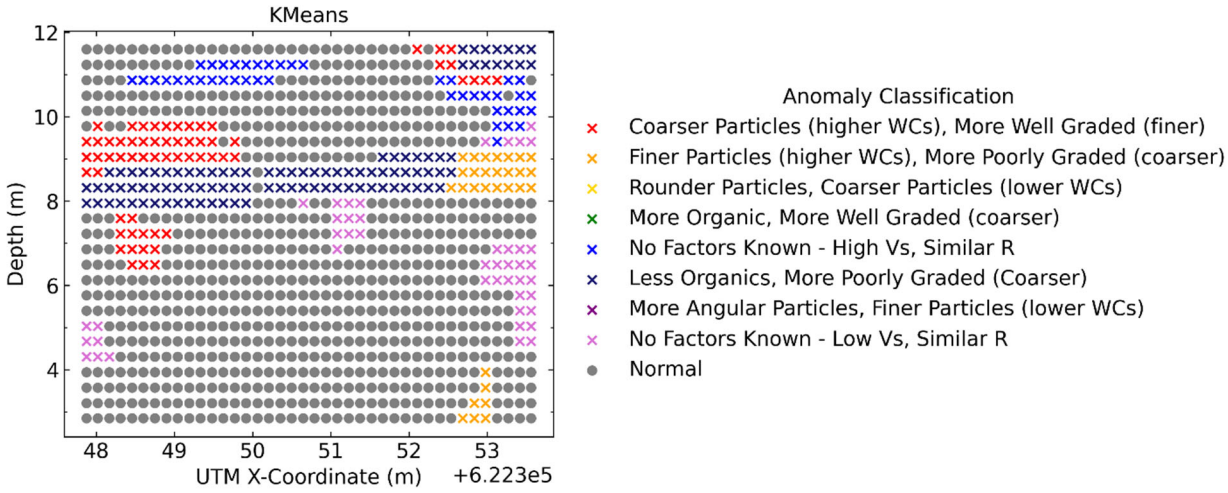


Fit: Individual Compilation
Algorithm: K-Means

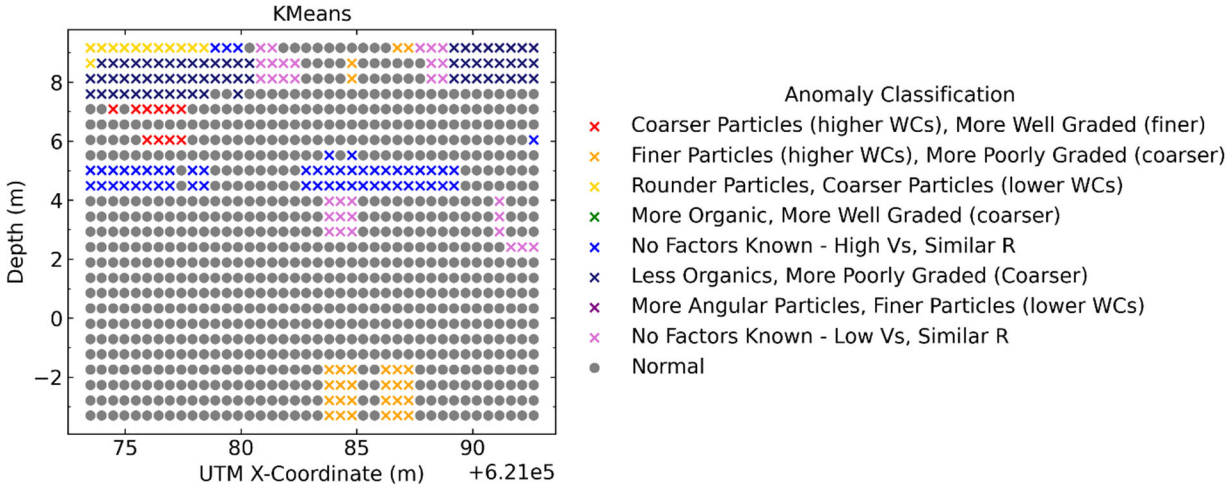
Profile: fc_site2



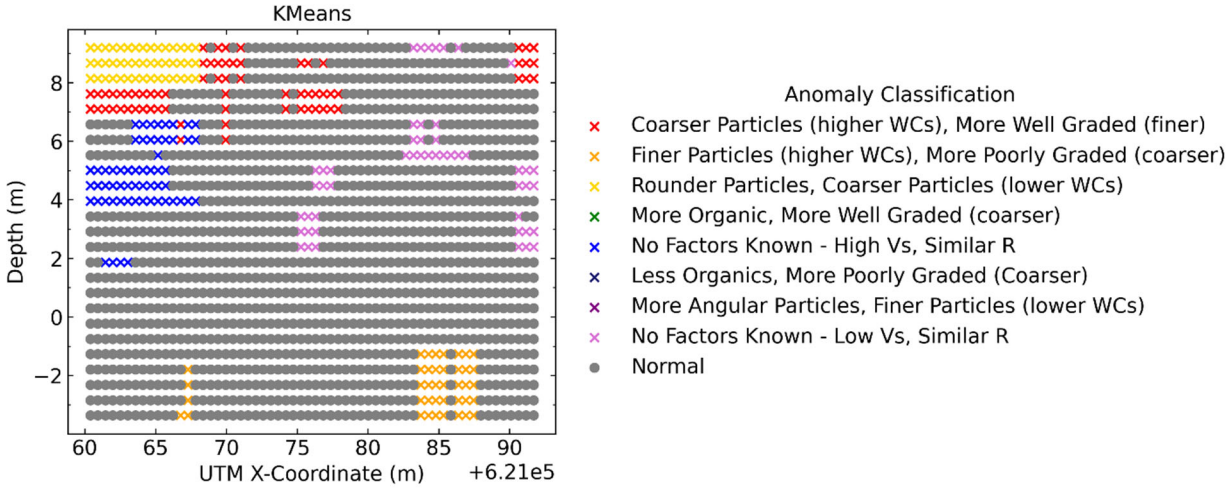
Profile: fc_site3



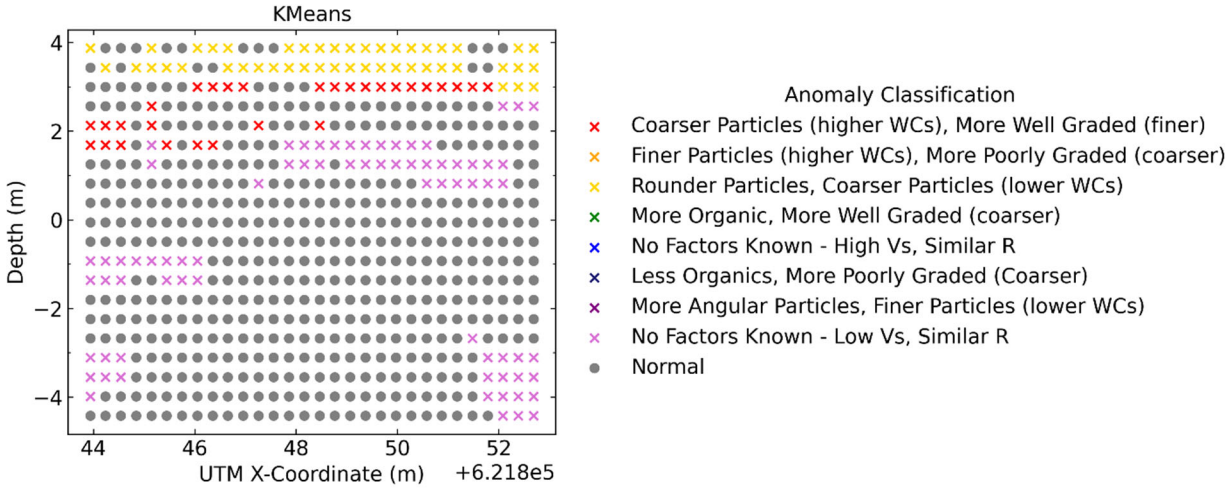
Profile: fc3_site1



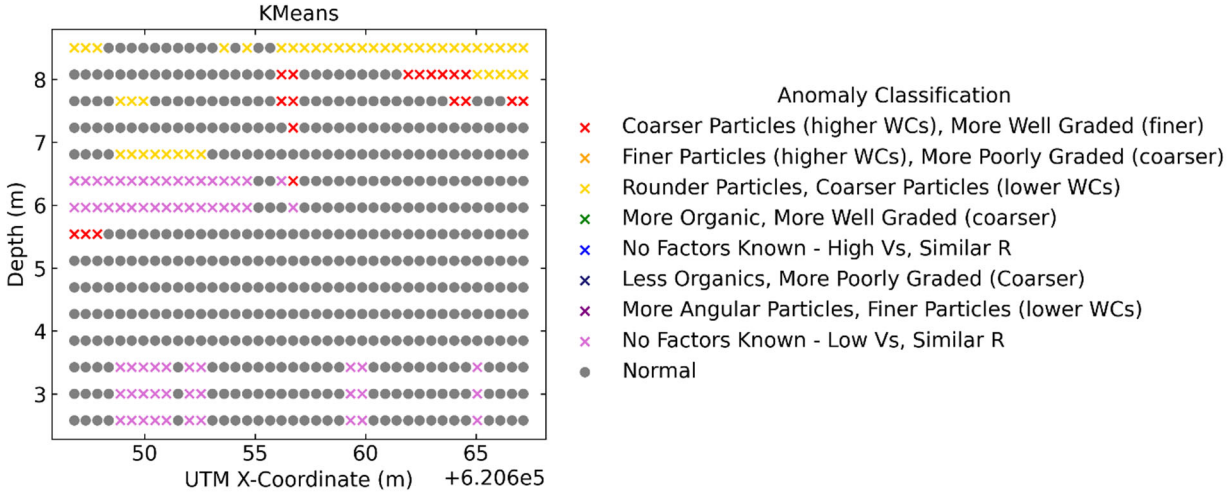
Profile: fc5_site1



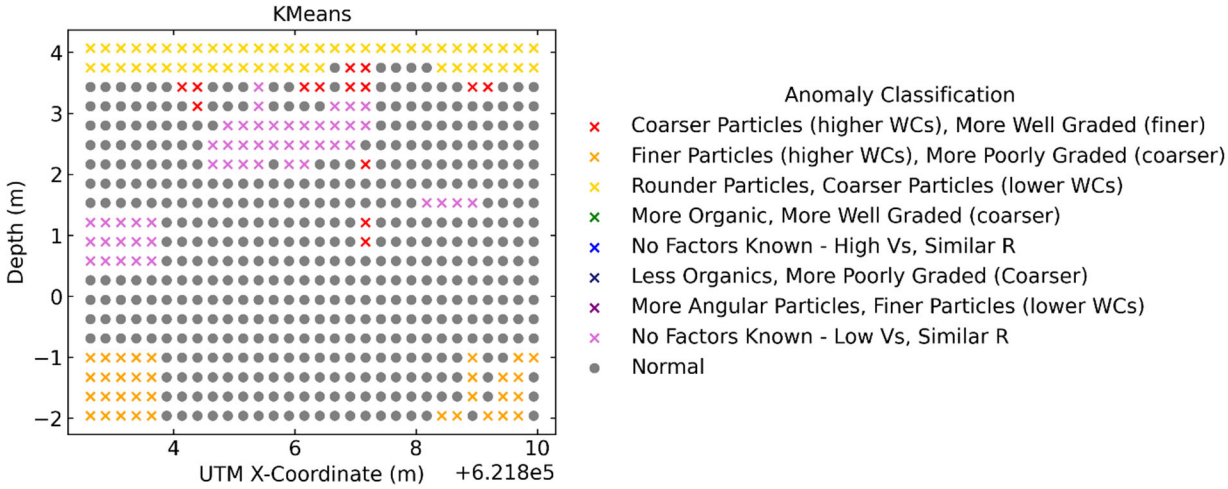
Profile: fdt_site2



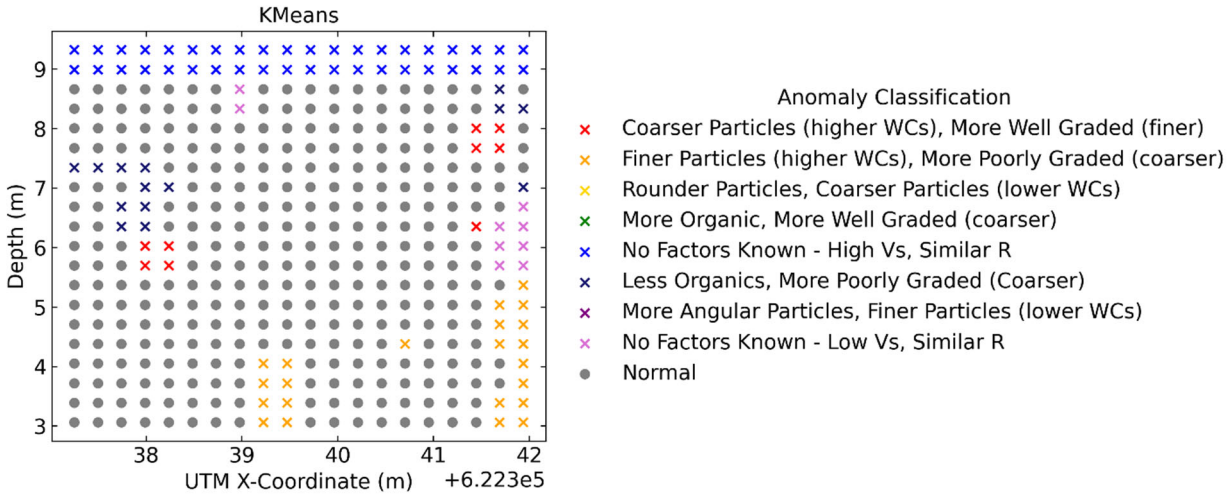
Profile: fwt_site1



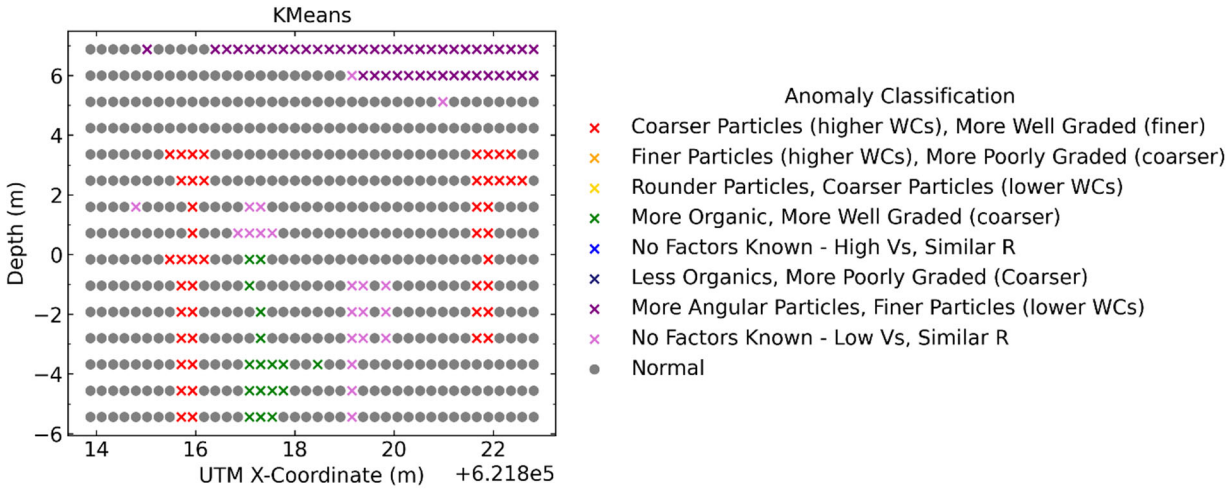
Profile: fwt_site2



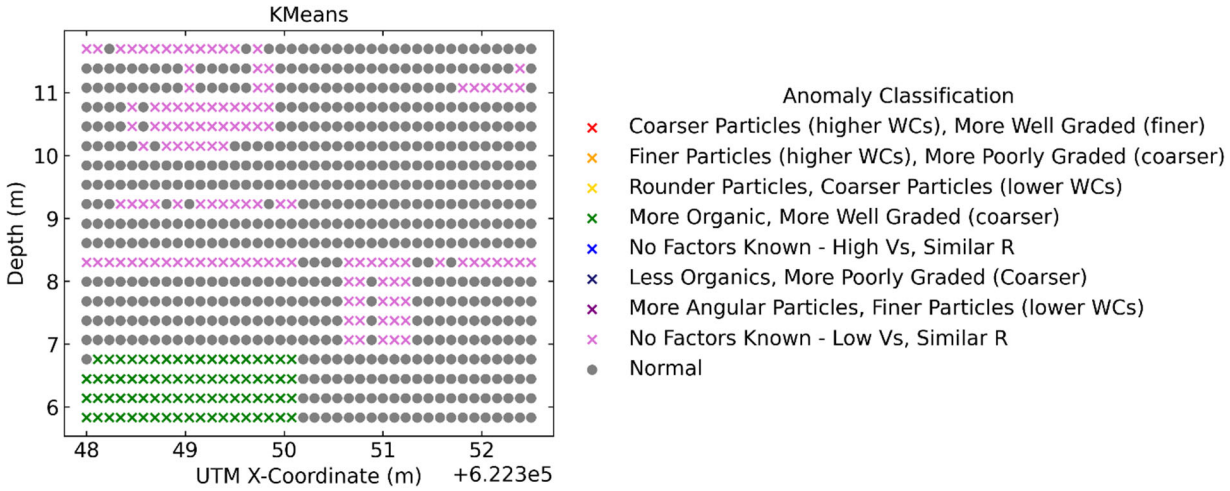
Profile: fwt_site3



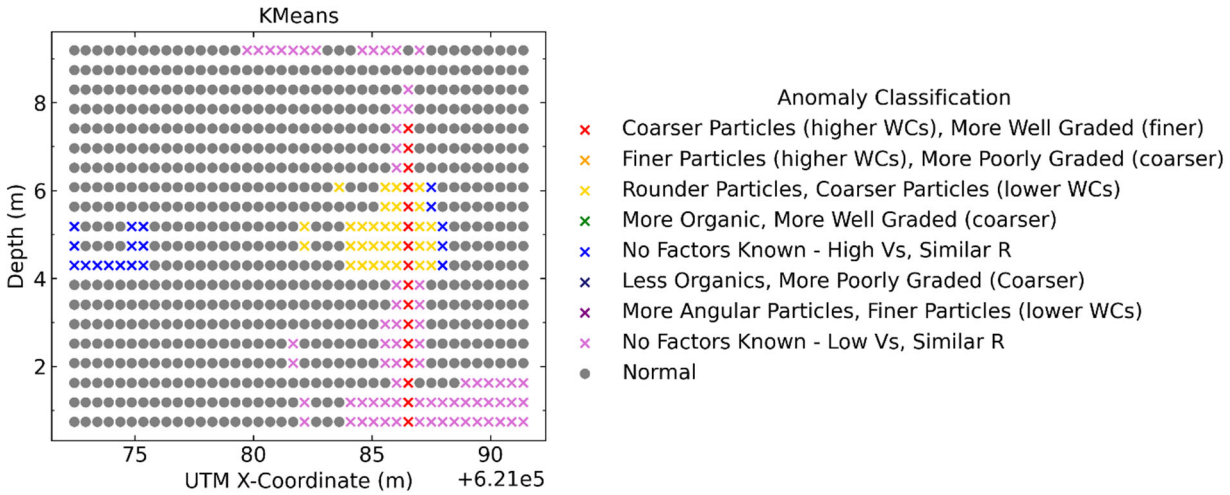
Profile: sc_site2



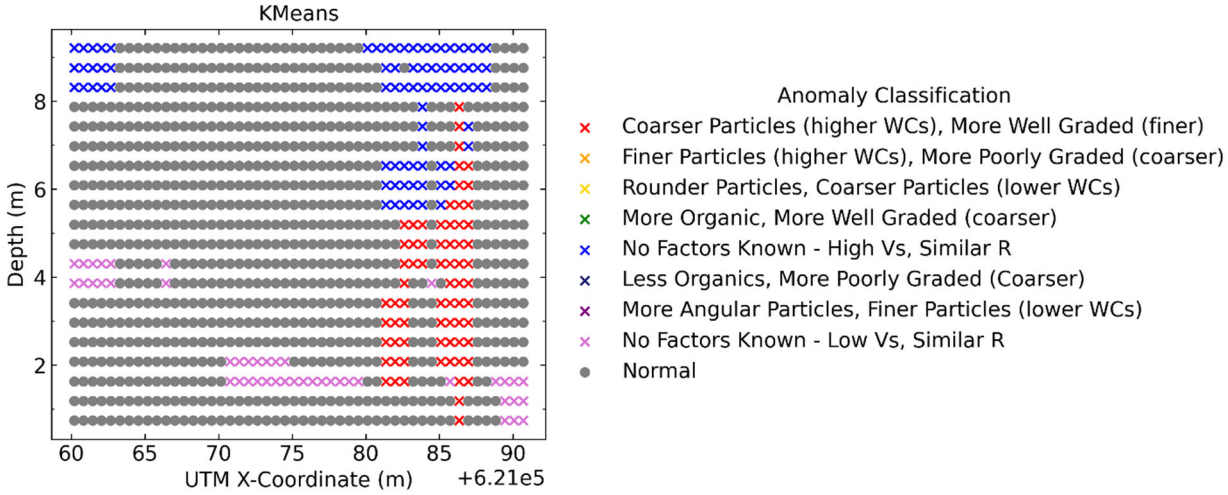
Profile: sc_site3



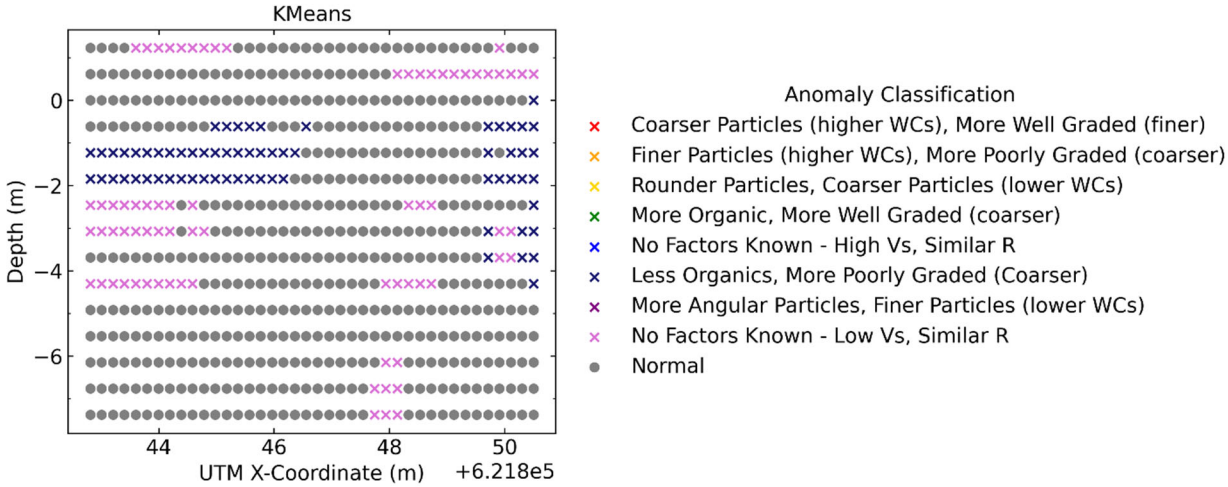
Profile: sc3_site1



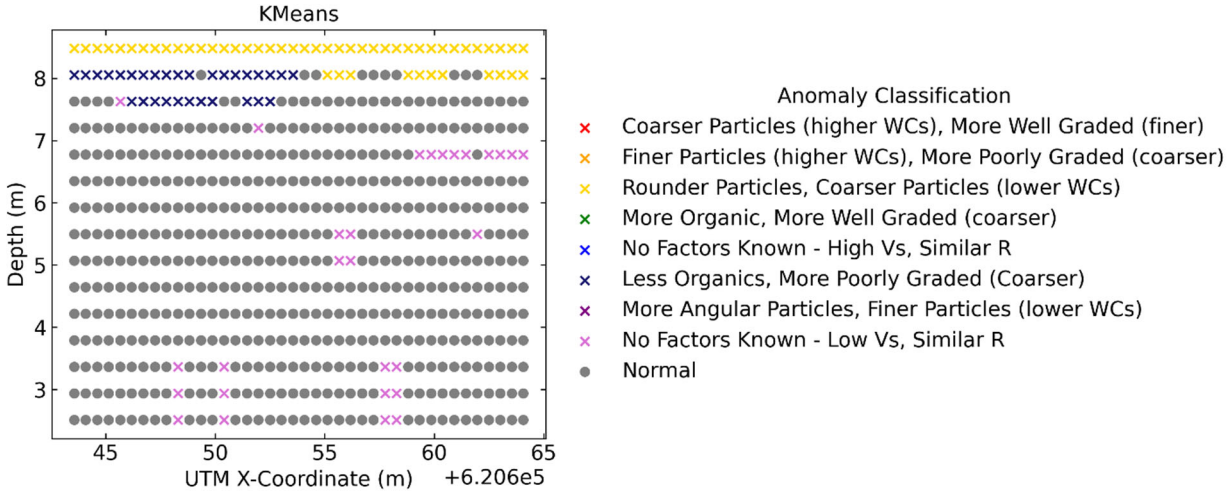
Profile: sc5_site1



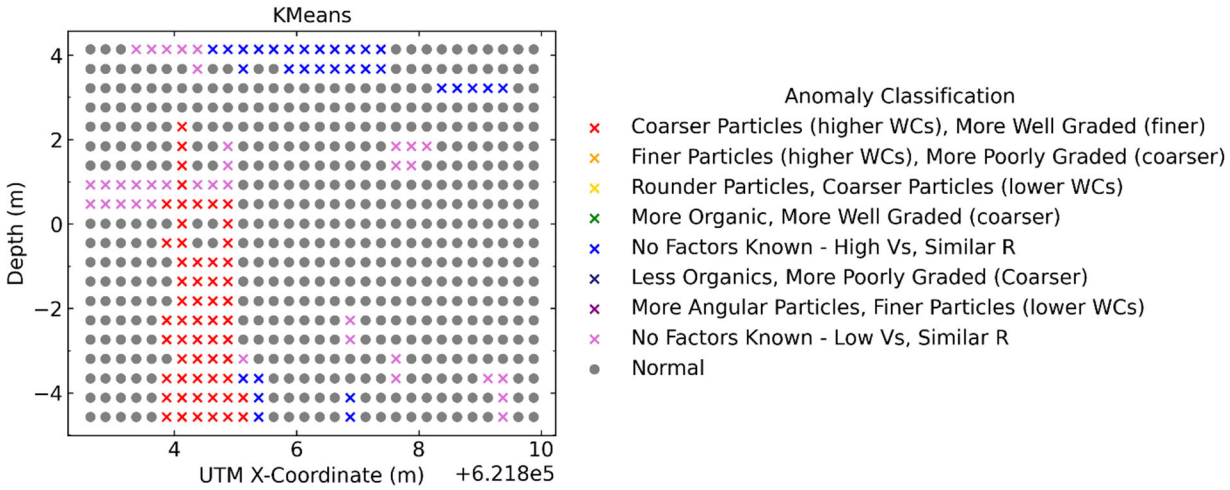
Profile: sdt_site2



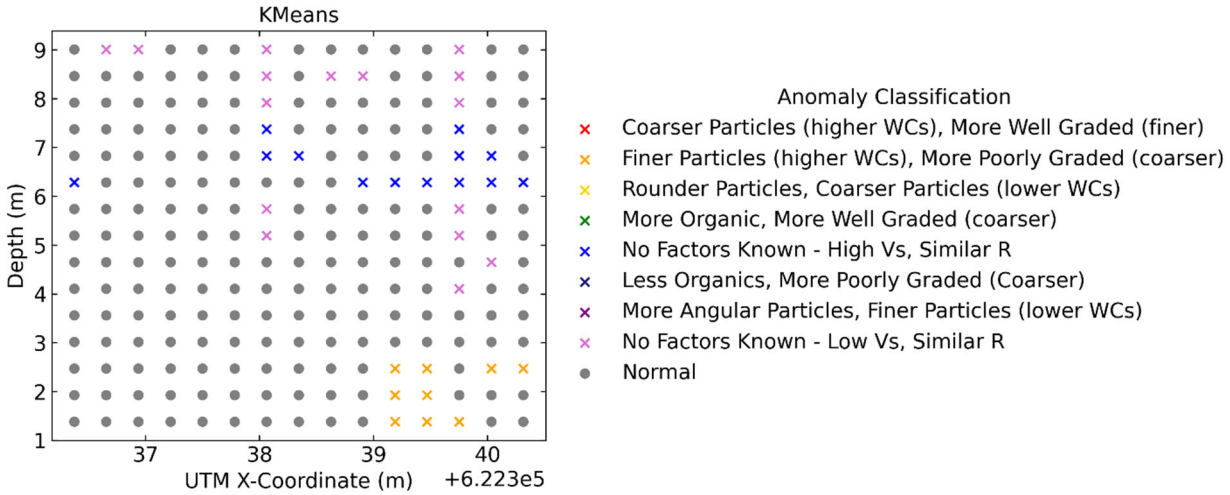
Profile: swt_site1



Profile: swt_site2

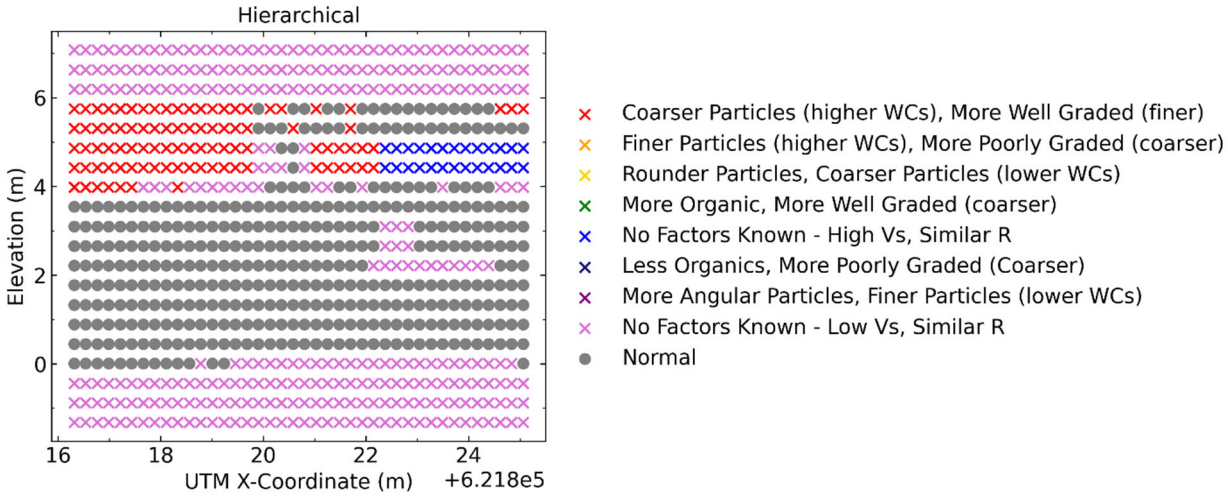


Profile: swt_site3

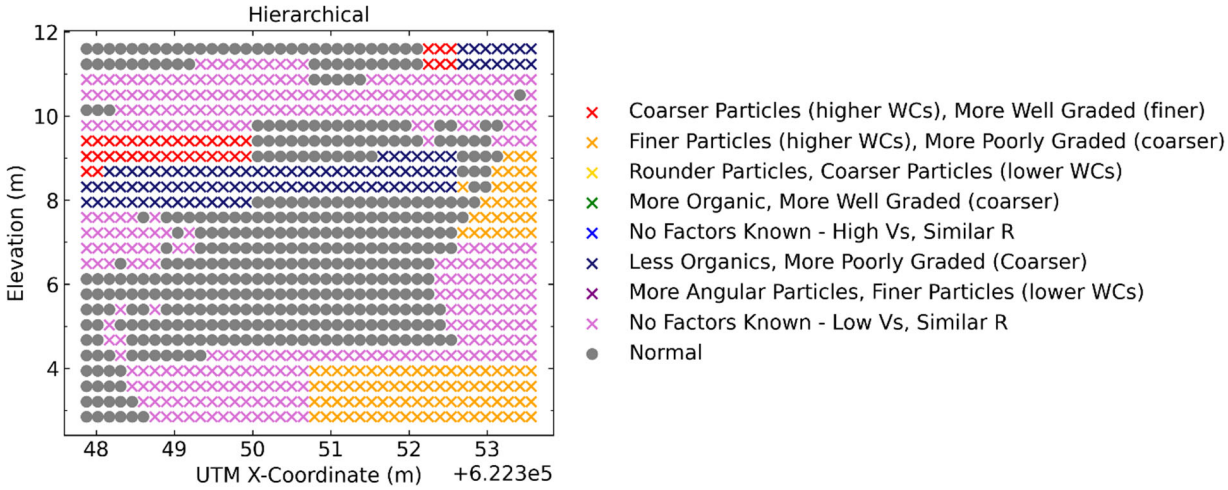


Fit: Crest Compilation
Algorithm: Hierarchical

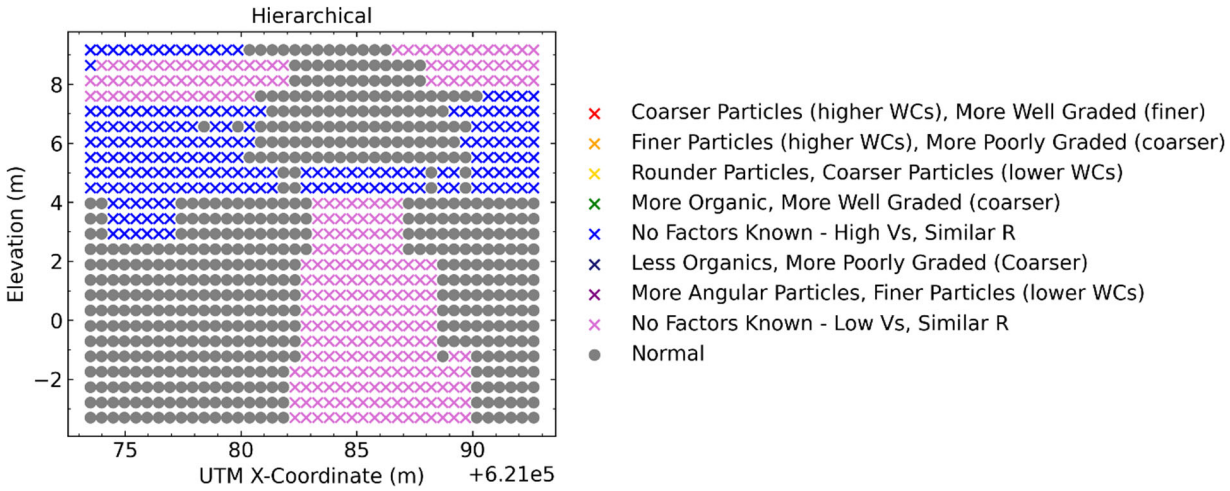
Profile: fc_site2



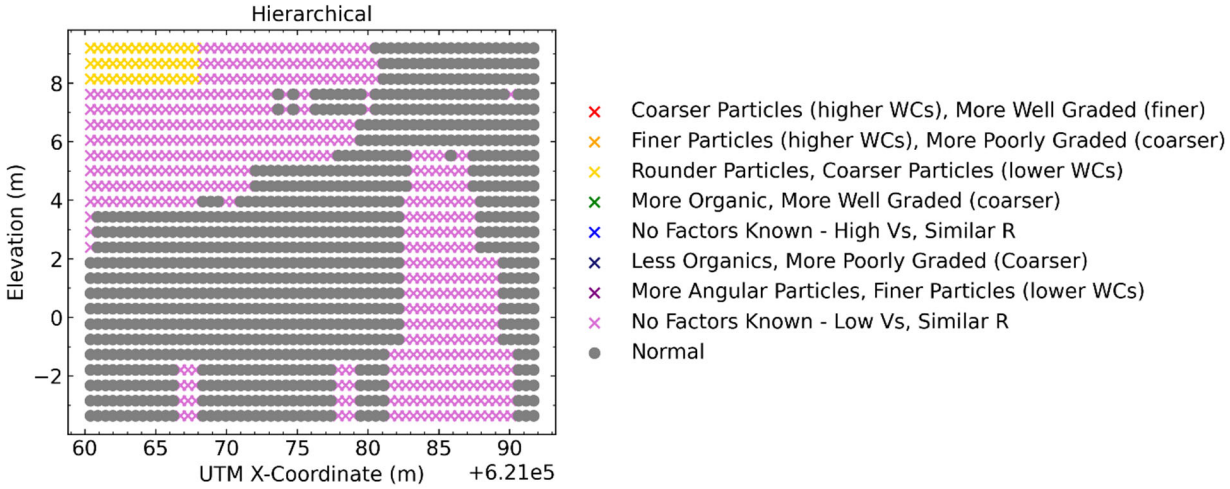
Profile: fc_site3



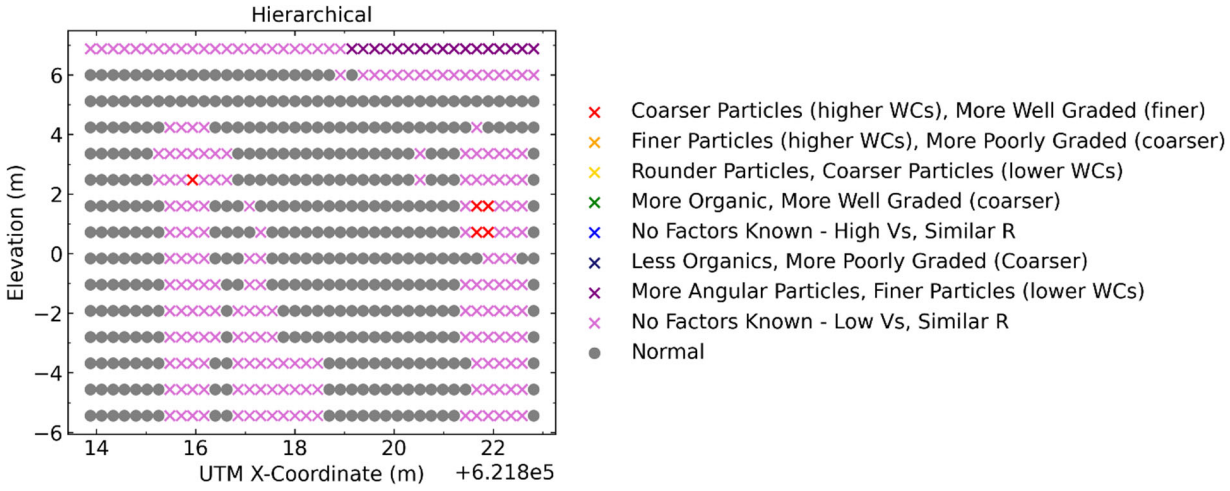
Profile: fc3_site1



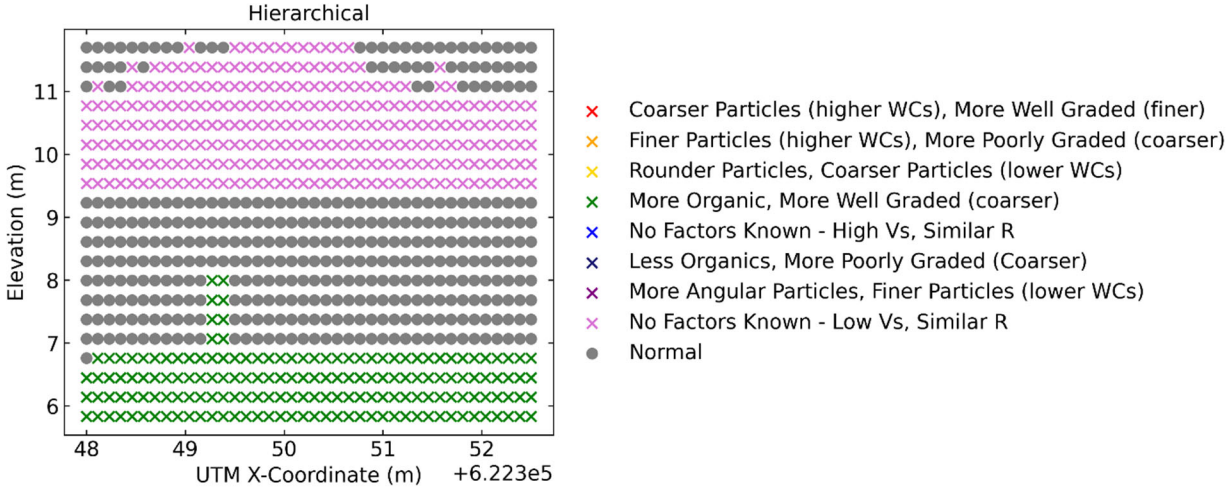
Profile: fc5_site1



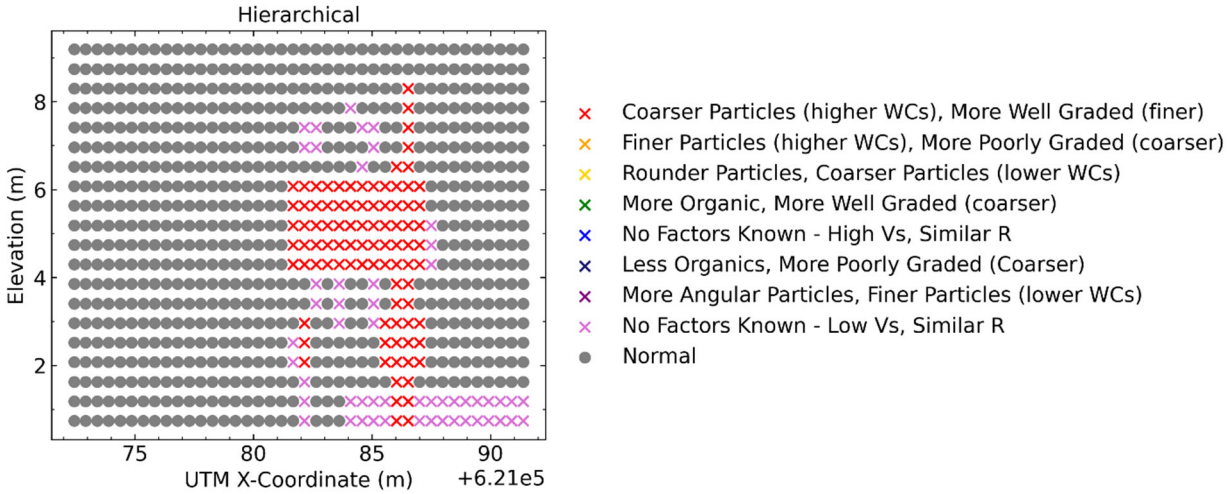
Profile: sc_site2



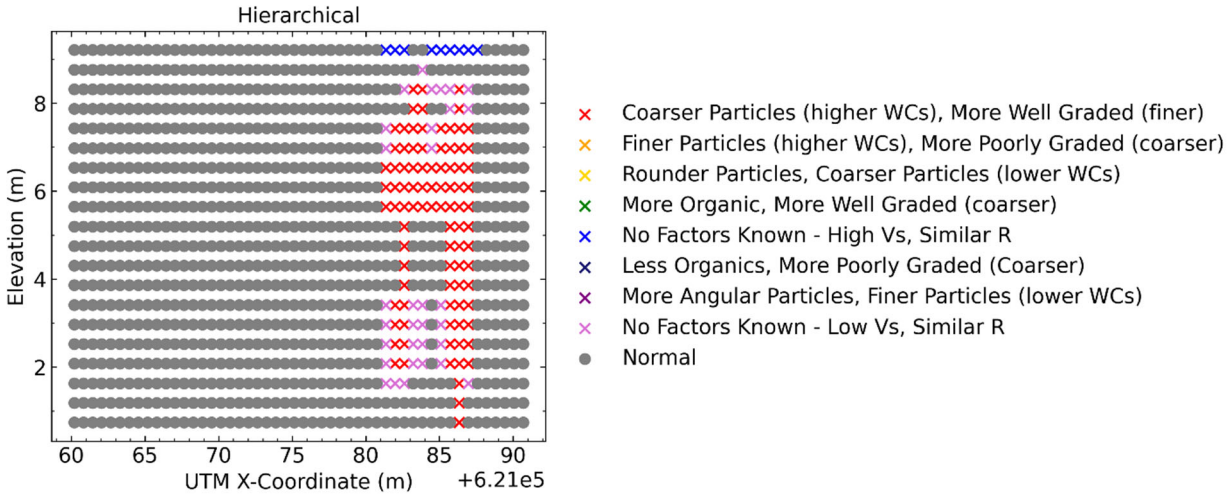
Profile: sc_site3



Profile: sc3_site1

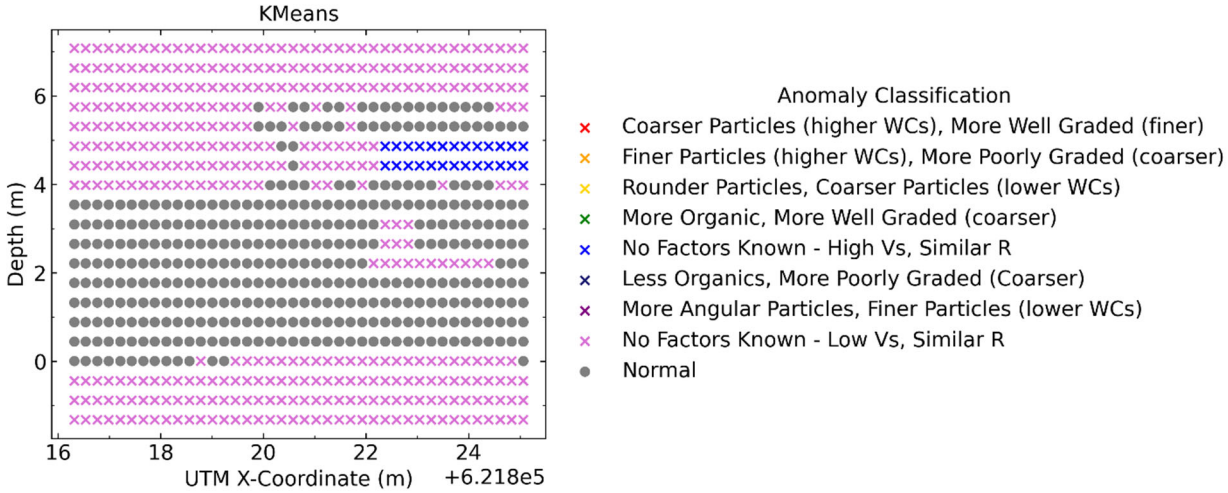


Profile: sc5_site1

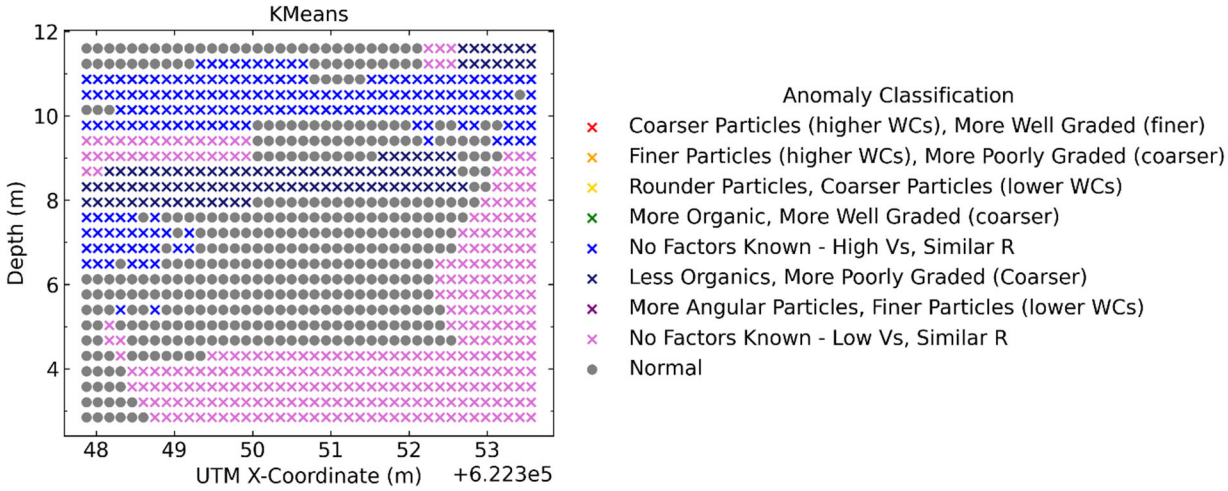


Fit: Crest Compilation
Algorithm: K-Means

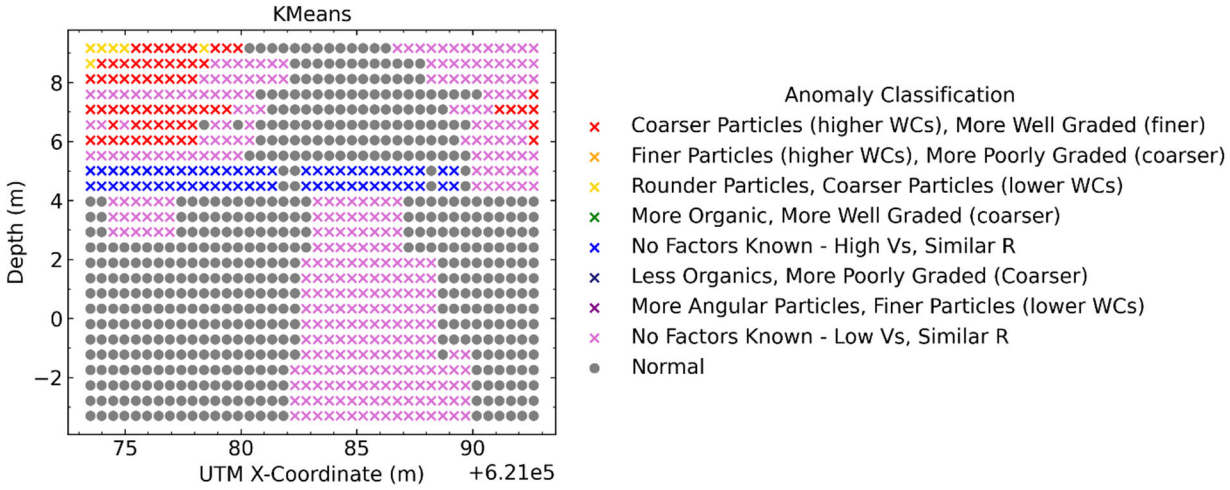
Profile: fc_site2



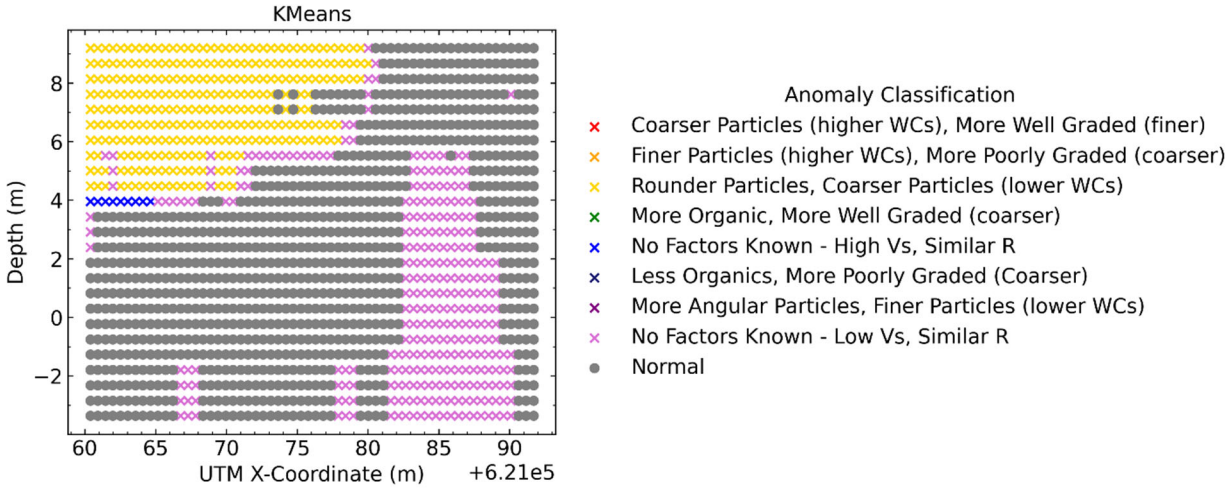
Profile: fc_site3



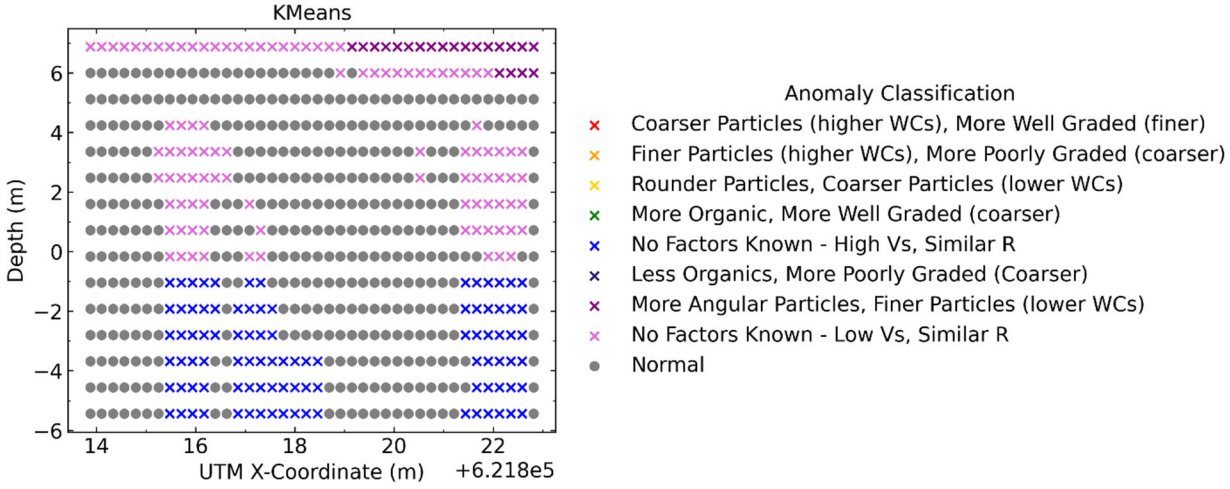
Profile: fc3_site1



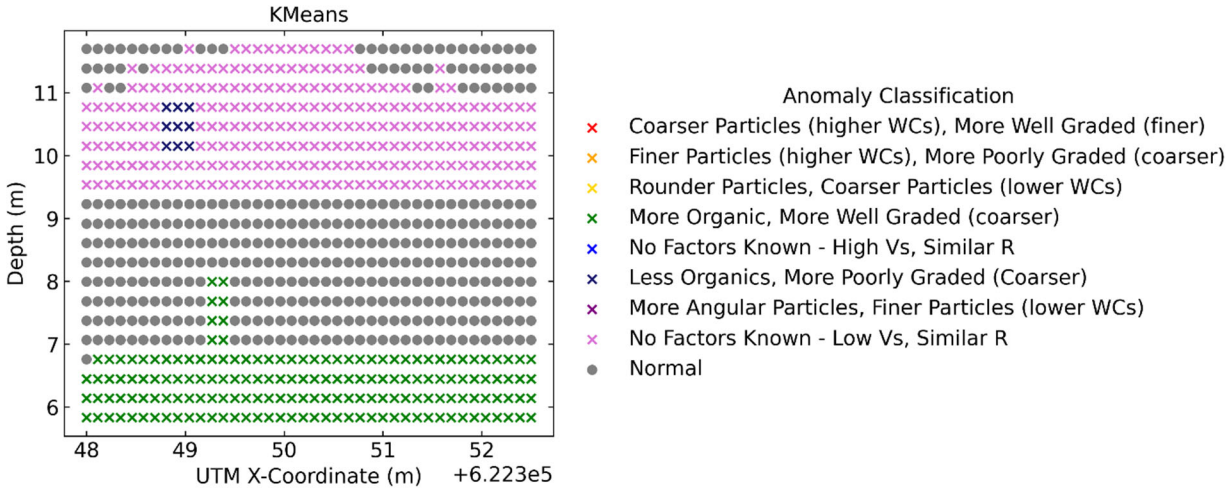
Profile: fc5_site1



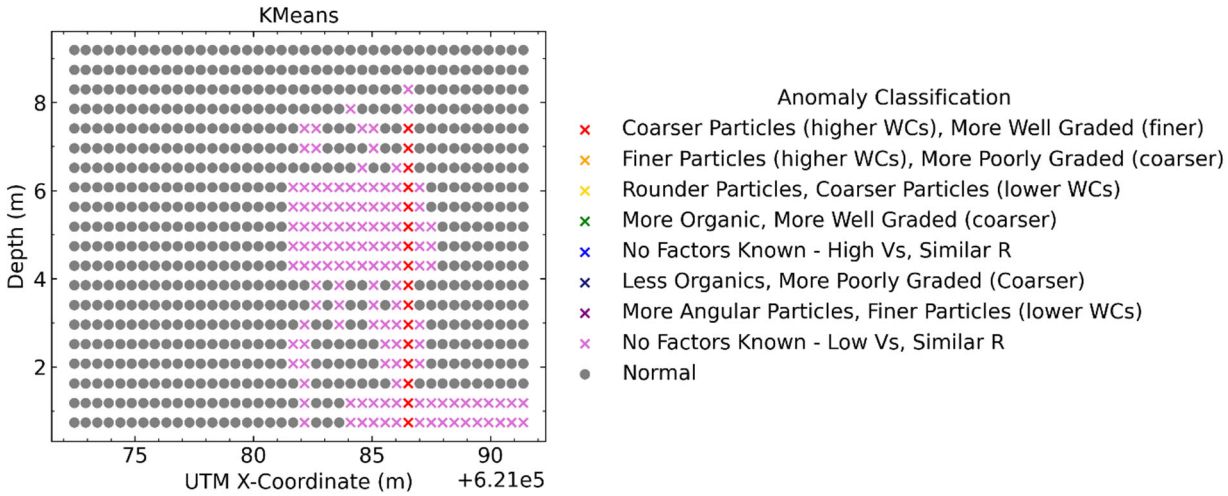
Profile: sc_site2



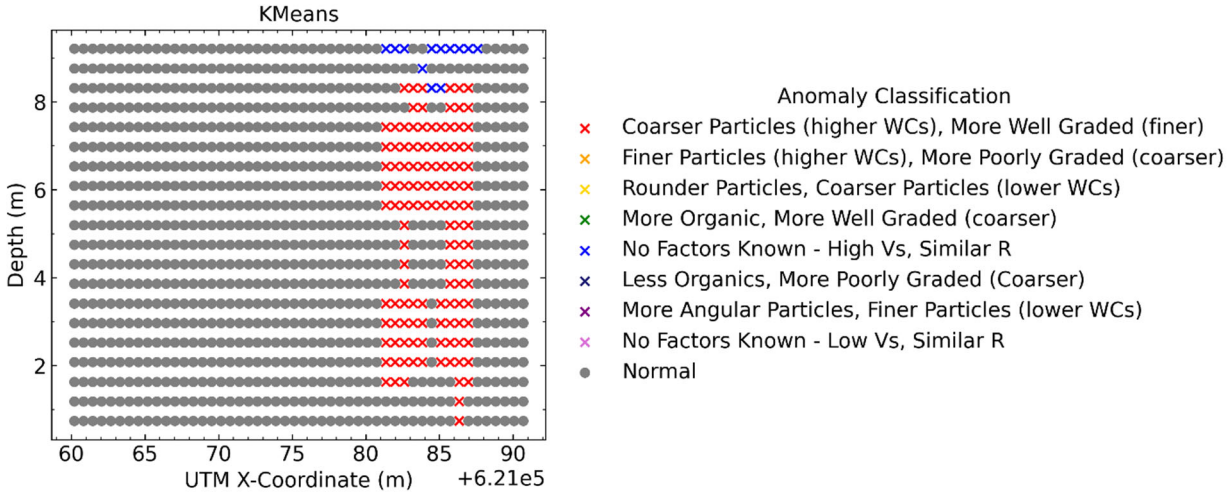
Profile: sc_site3



Profile: sc3_site1

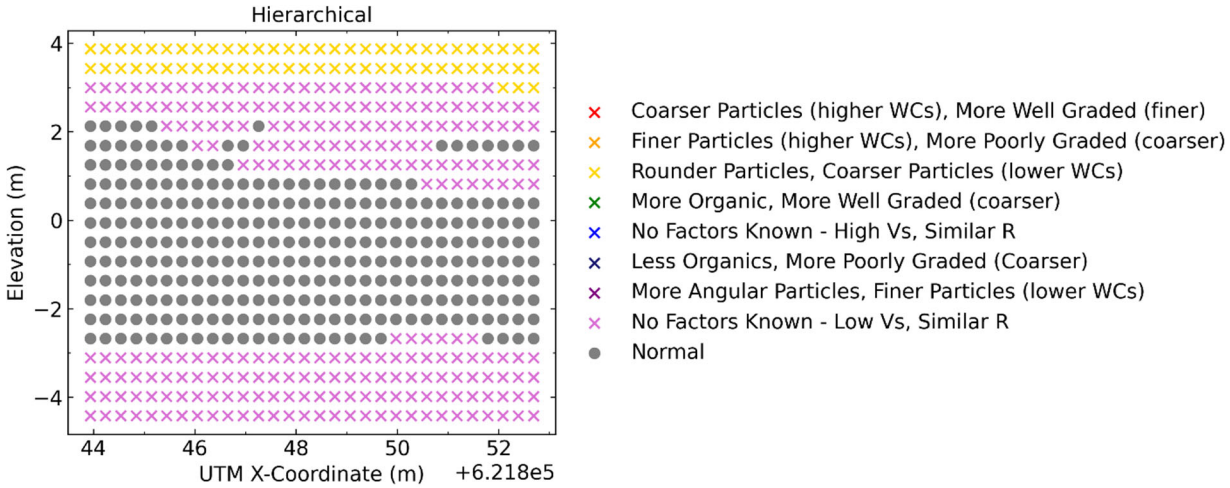


Profile: sc5_site1

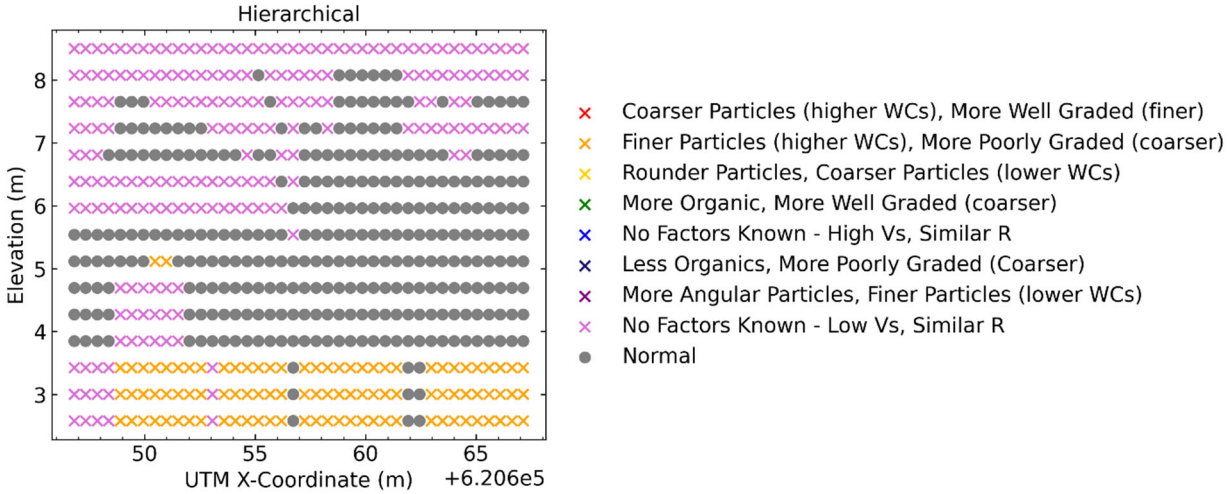


Fit: Toe Compilation
Algorithm: Hierarchical

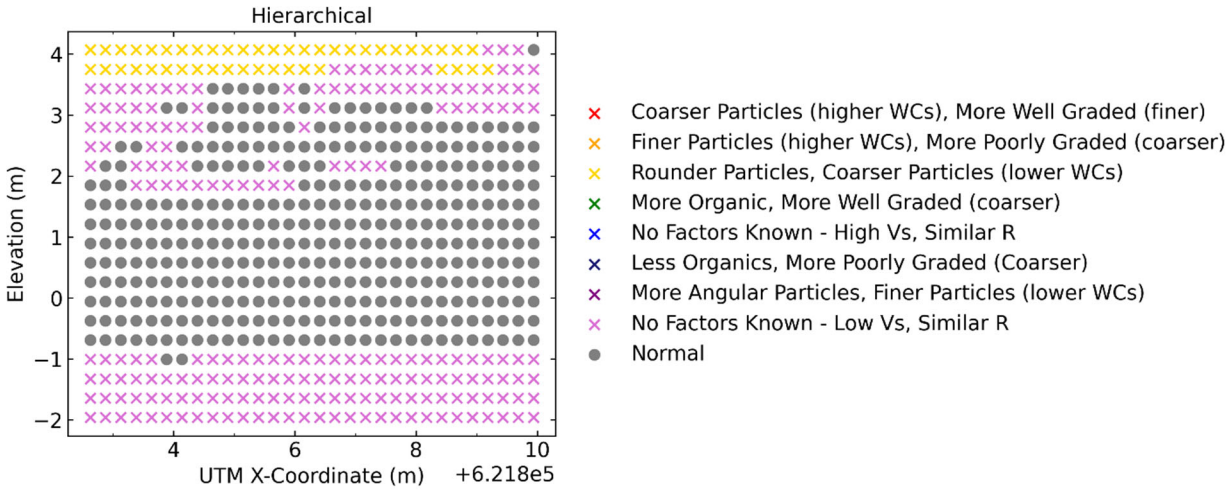
Profile: fdt_site2



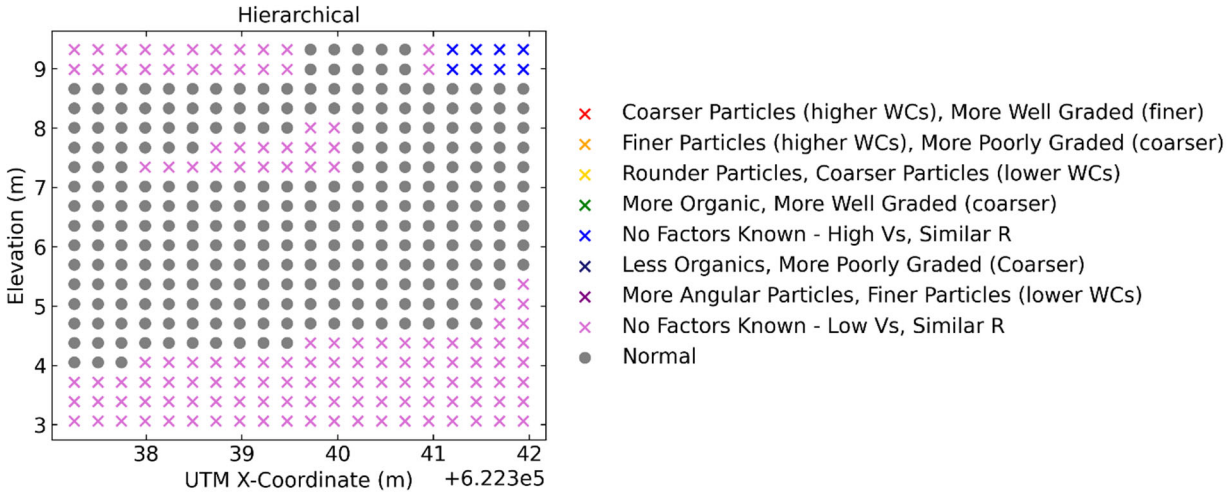
Profile: fwt_site1



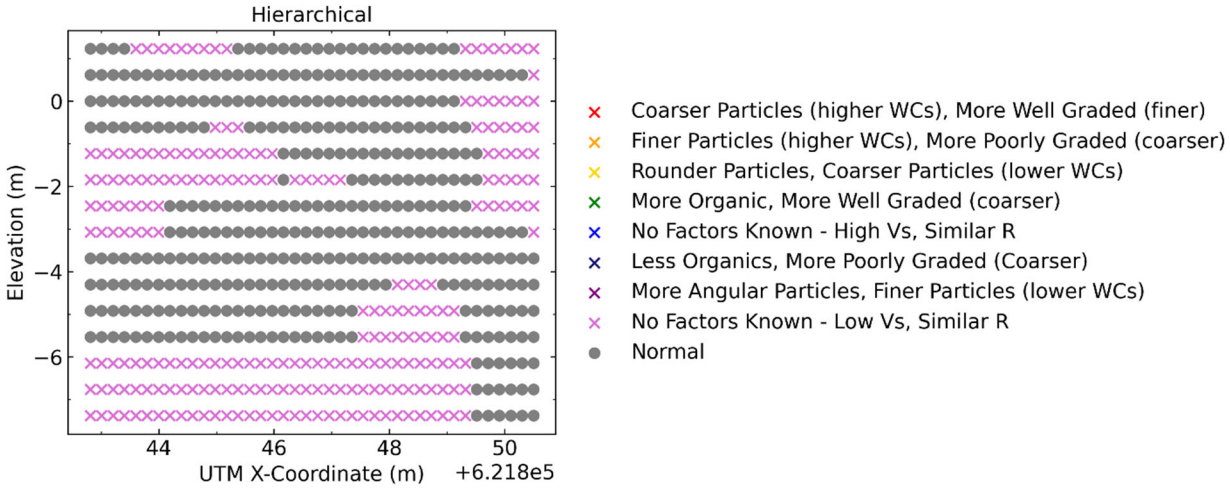
Profile: fwt_site2



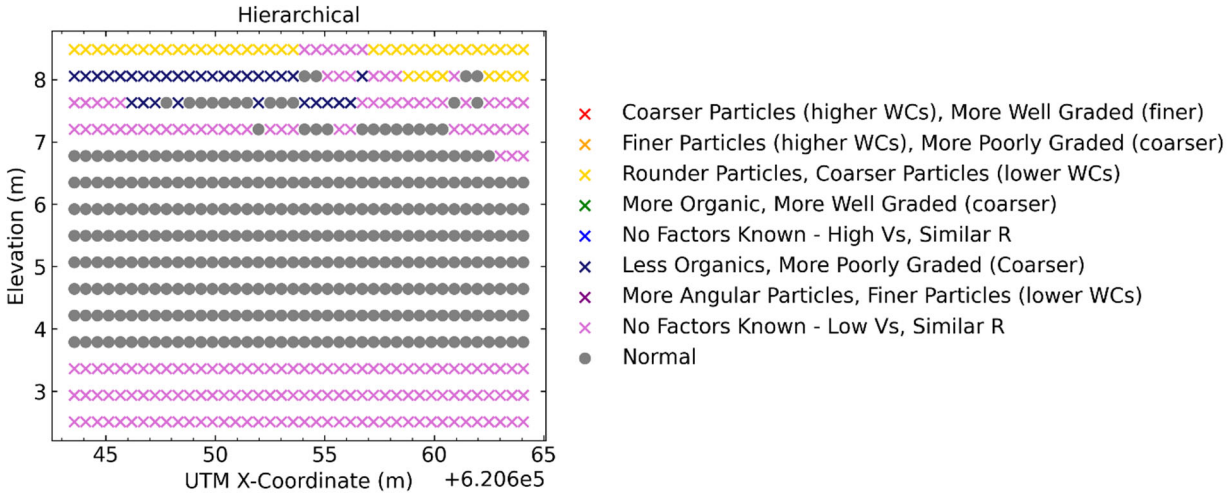
Profile: fwt_site3



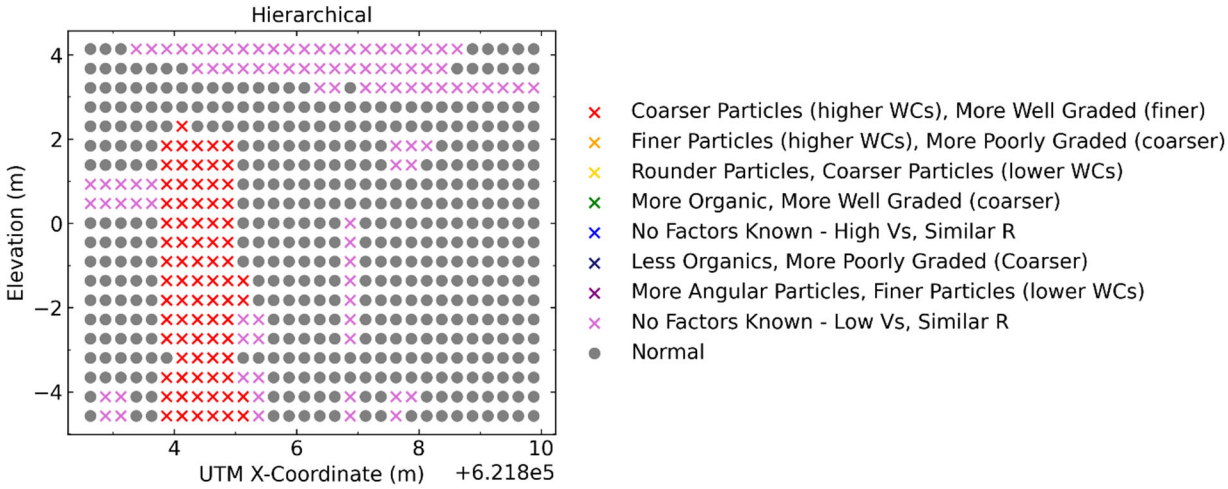
Profile: sdt_site2



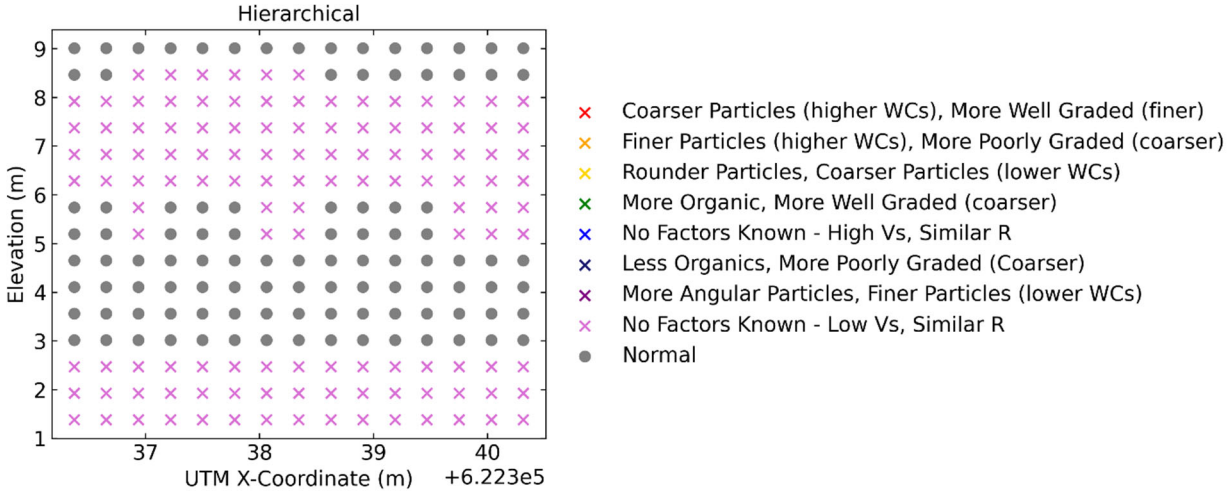
Profile: swt_site1



Profile: swt_site2

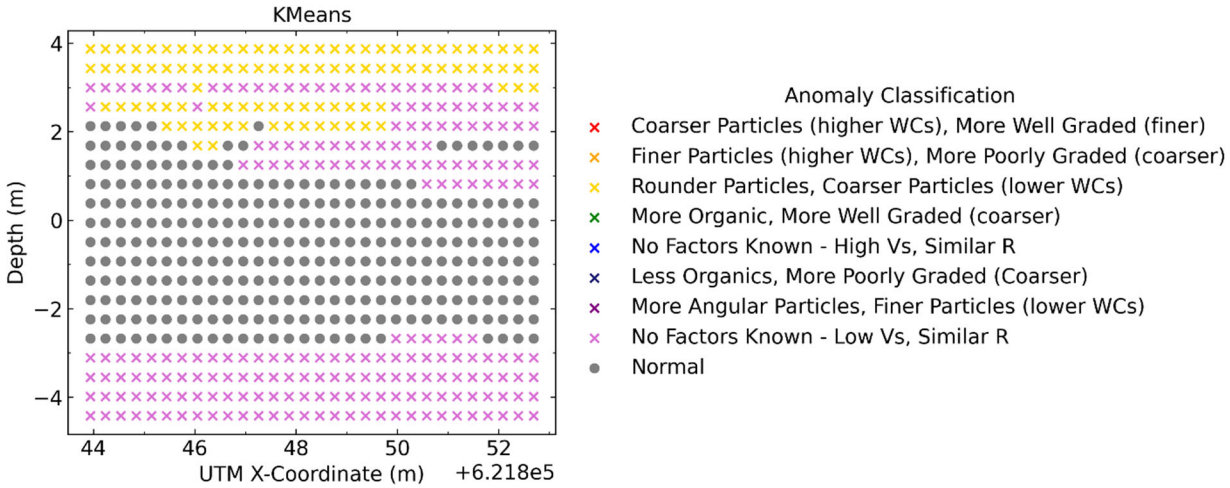


Profile: swt_site3

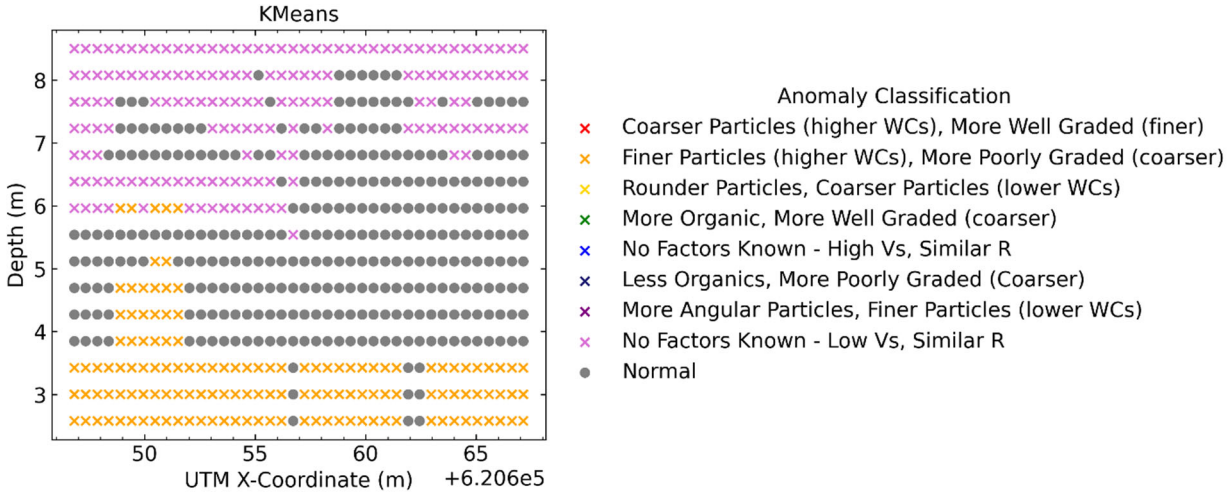


Fit: Toe Compilation
Algorithm: K-Means

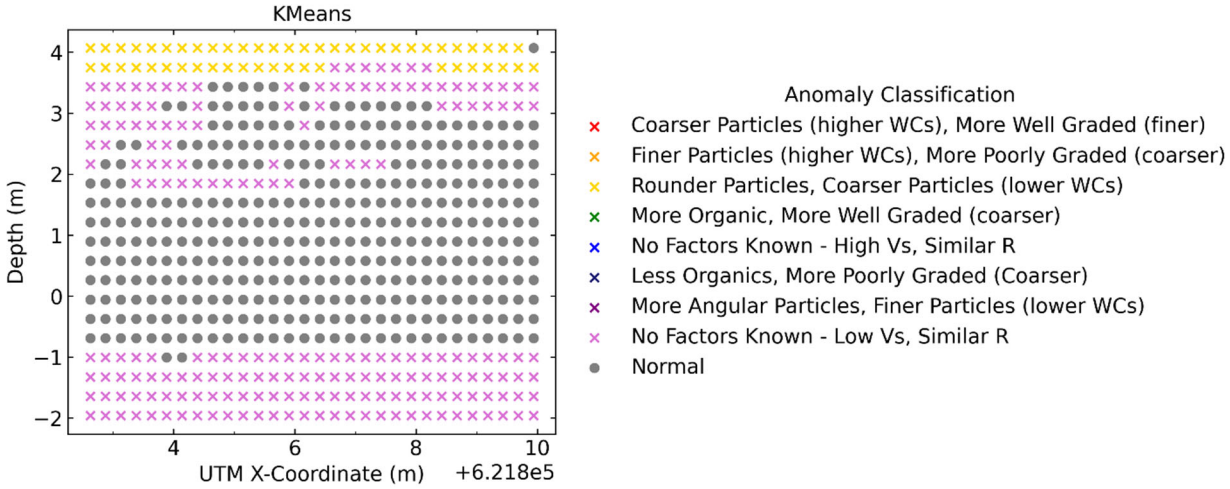
Profile: fdt_site2



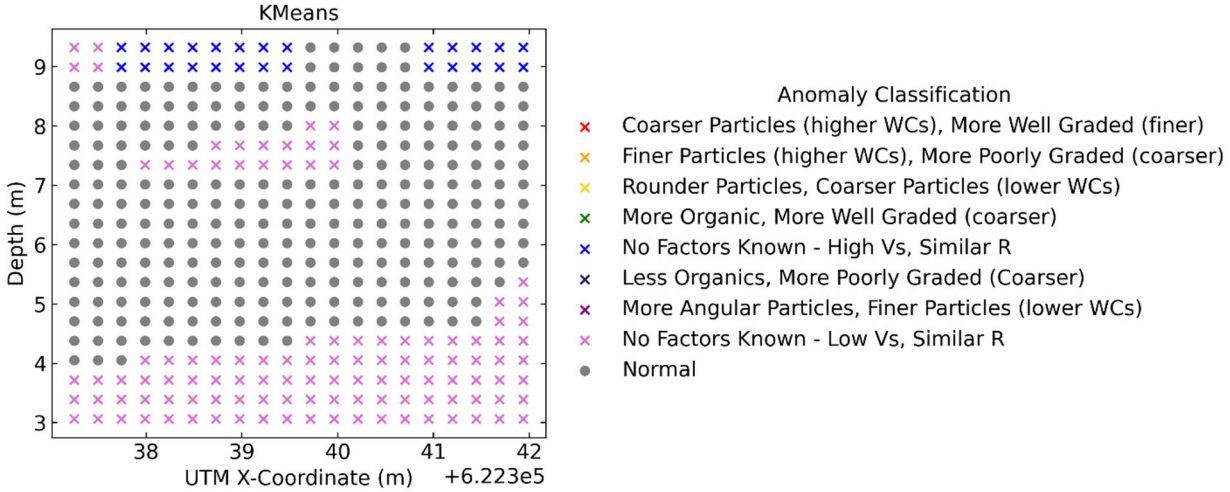
Profile: fwt_site1



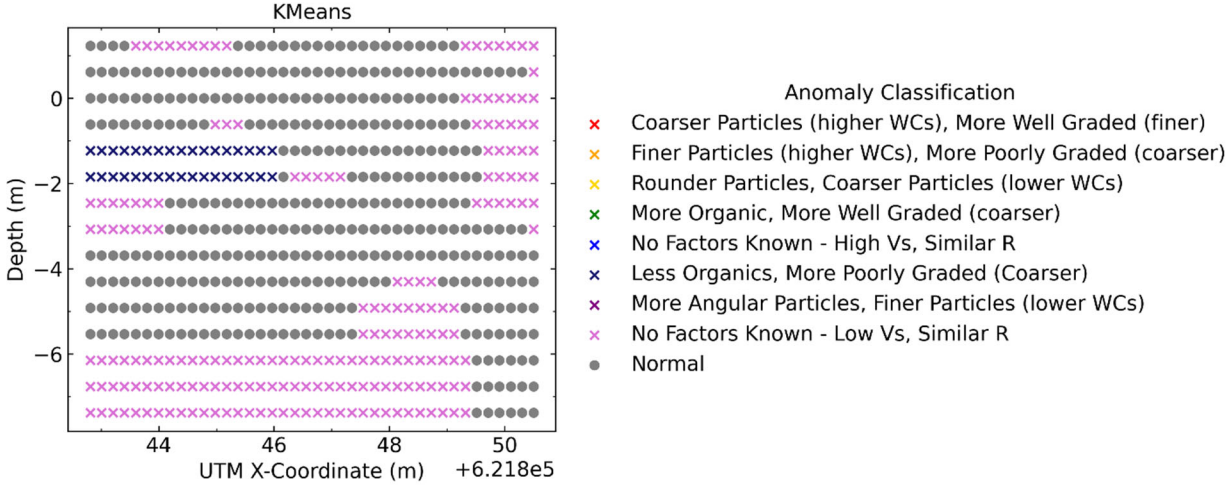
Profile: fwt_site2



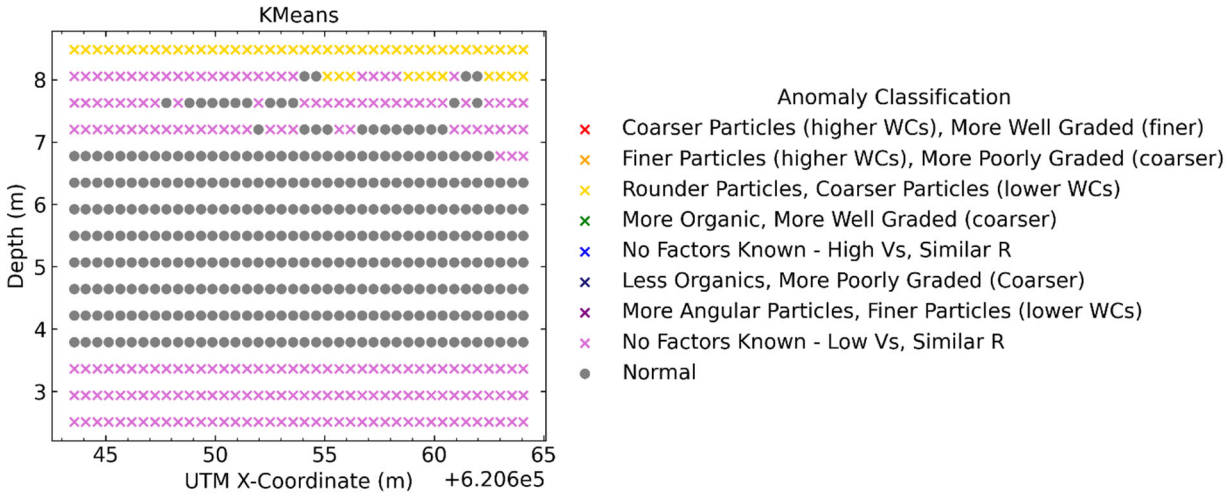
Profile: fwt_site3



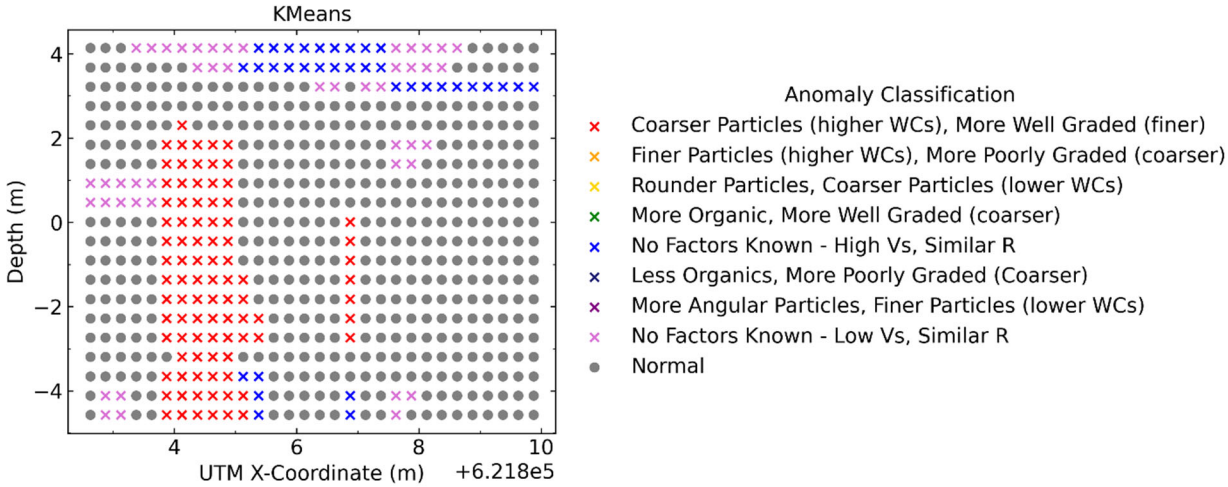
Profile: sdt_site2



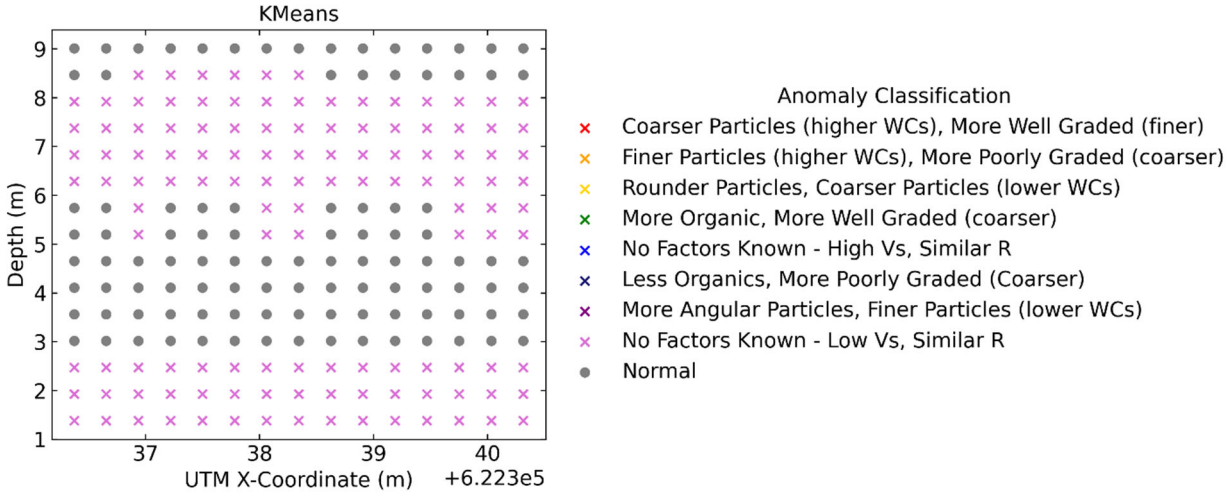
Profile: swt_site1



Profile: swt_site2

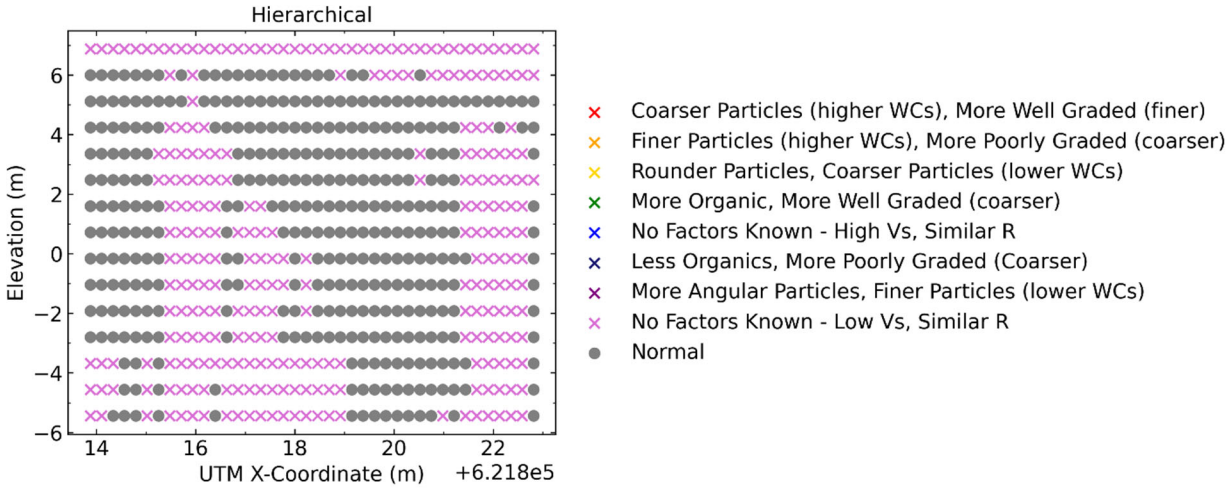


Profile: swt_site3

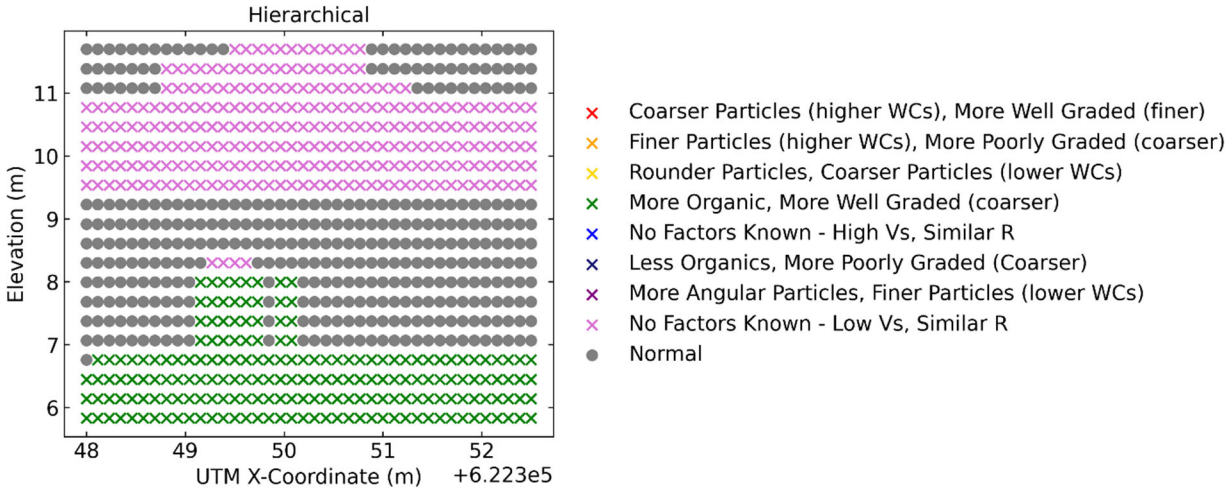


Fit: Dry Compilation
Algorithm: Hierarchical

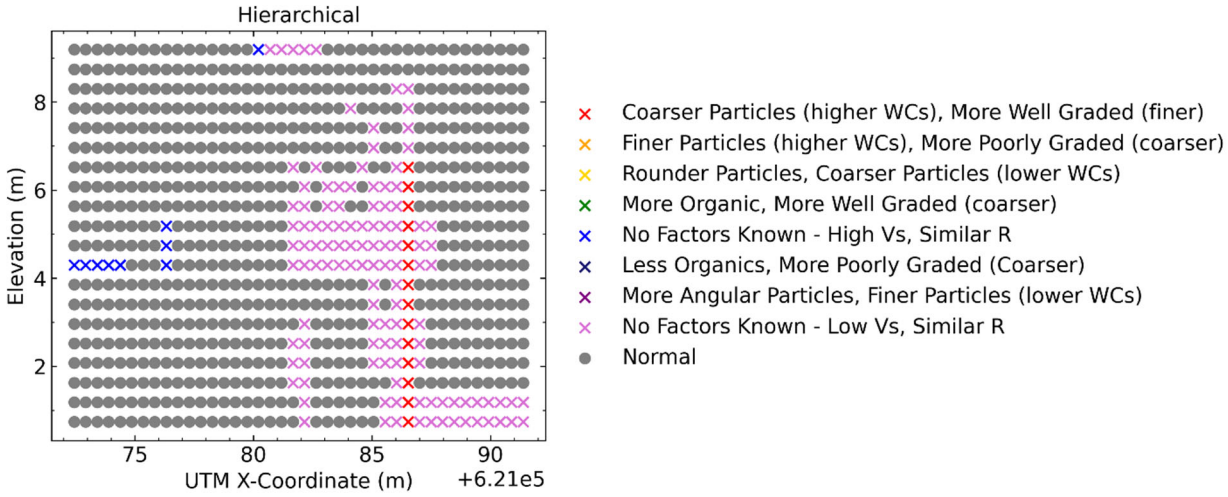
Profile: sc_site2



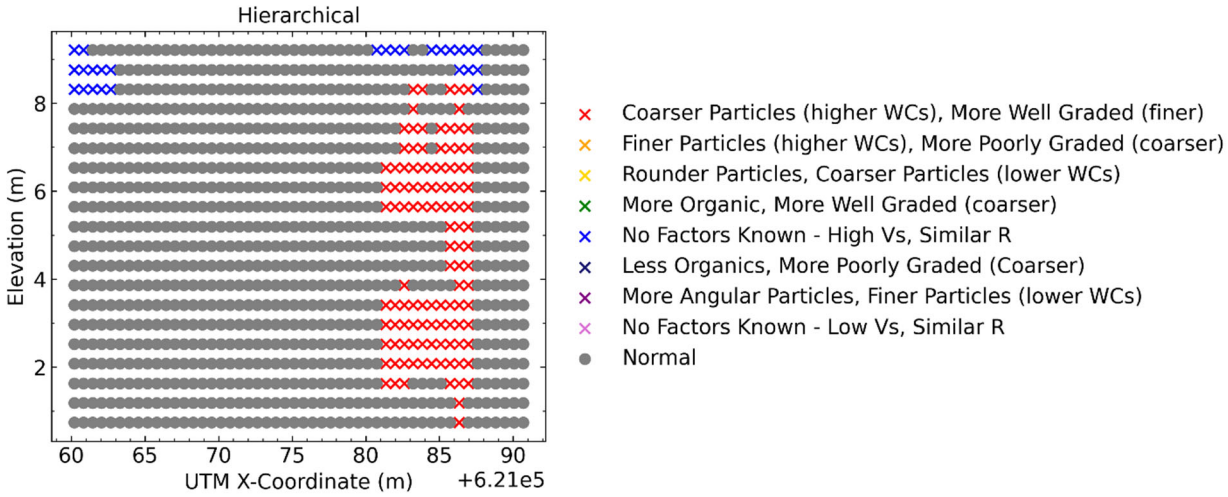
Profile: sc_site3



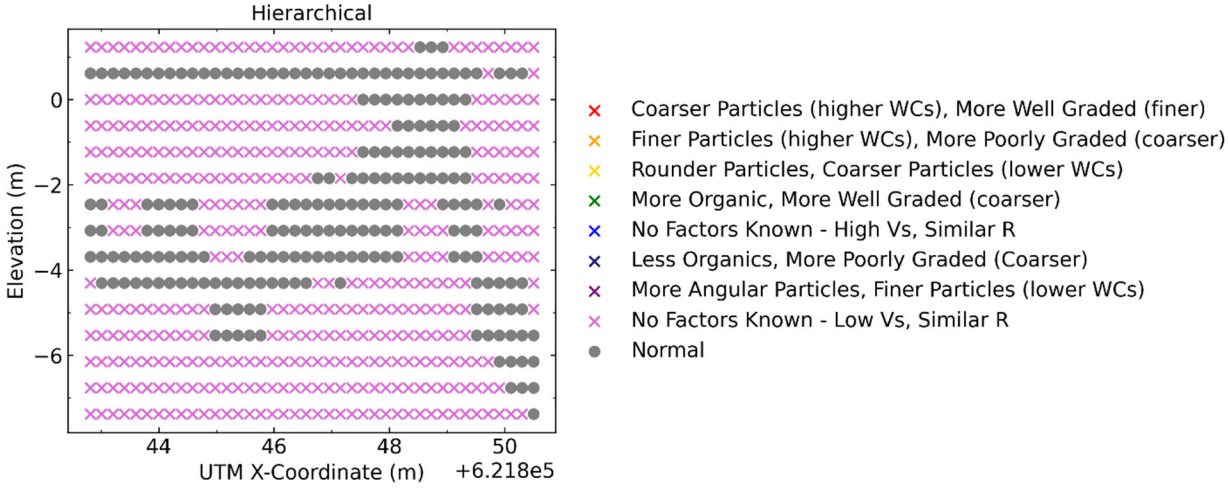
Profile: sc3_site1



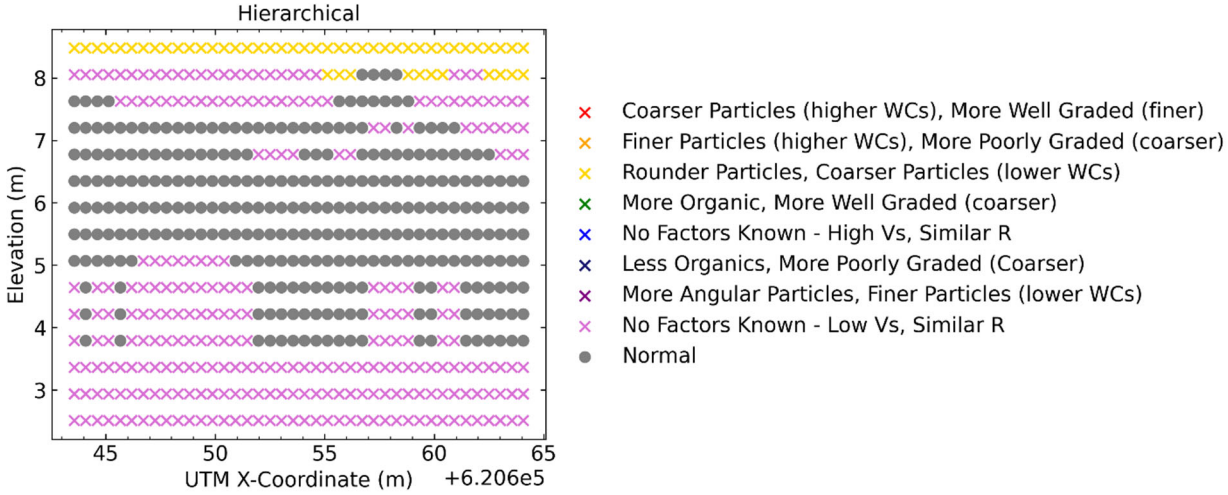
Profile: sc5_site1



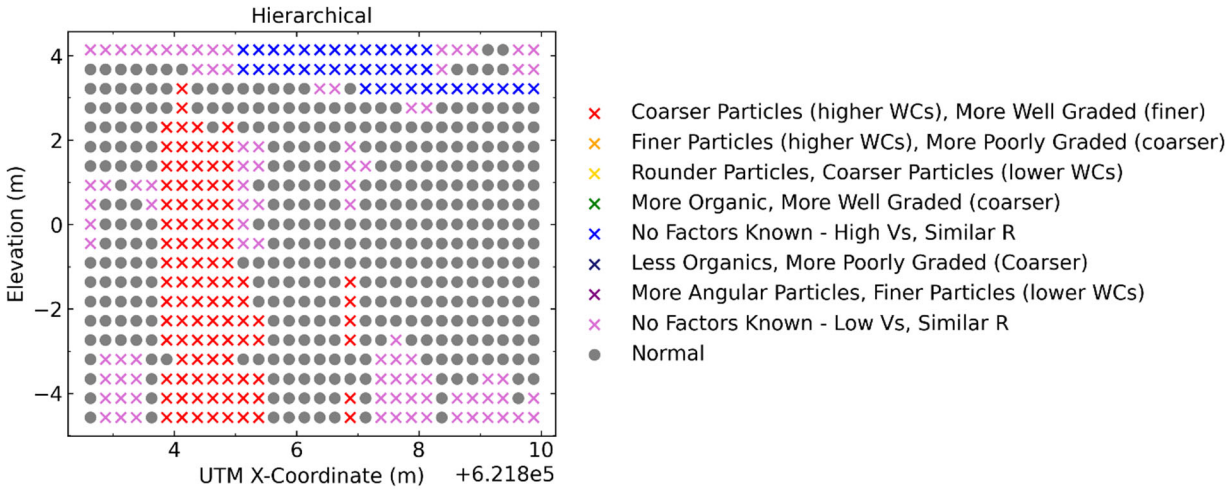
Profile: sdt_site2



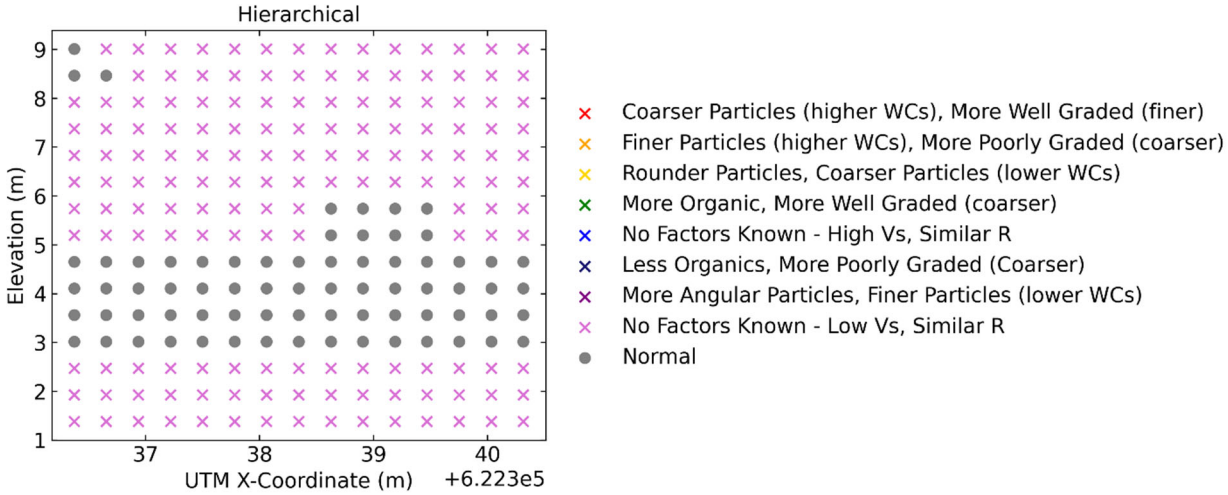
Profile: swt_site1



Profile: swt_site2

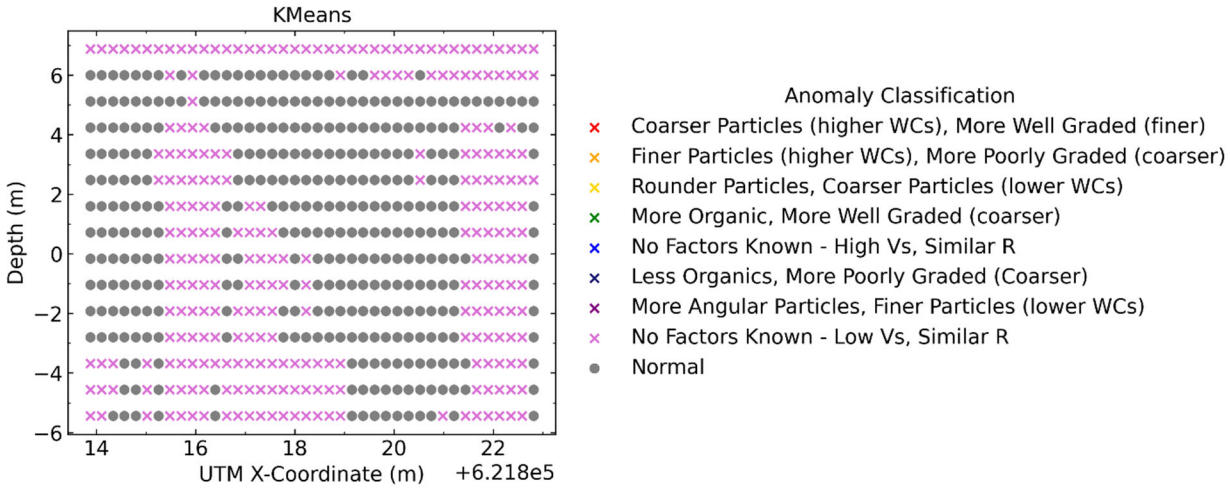


Profile: swt_site3

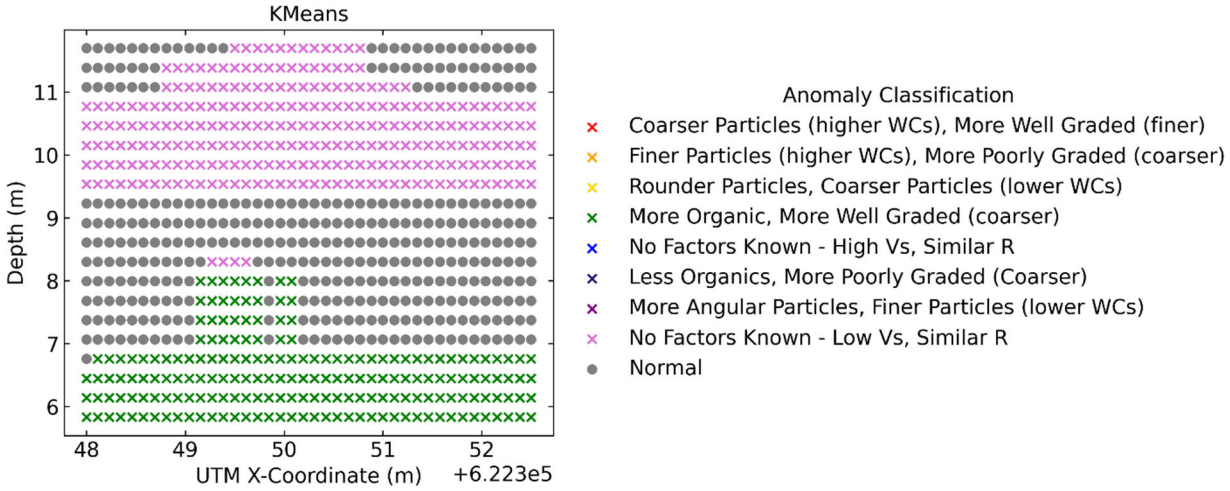


Fit: Dry Compilation
Algorithm: K-Means

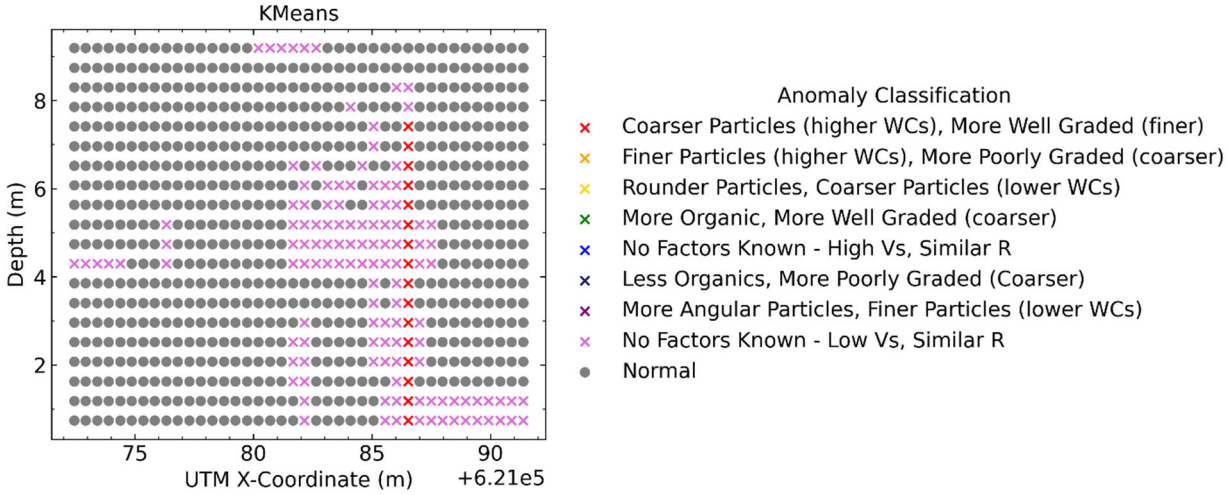
Profile: sc_site2



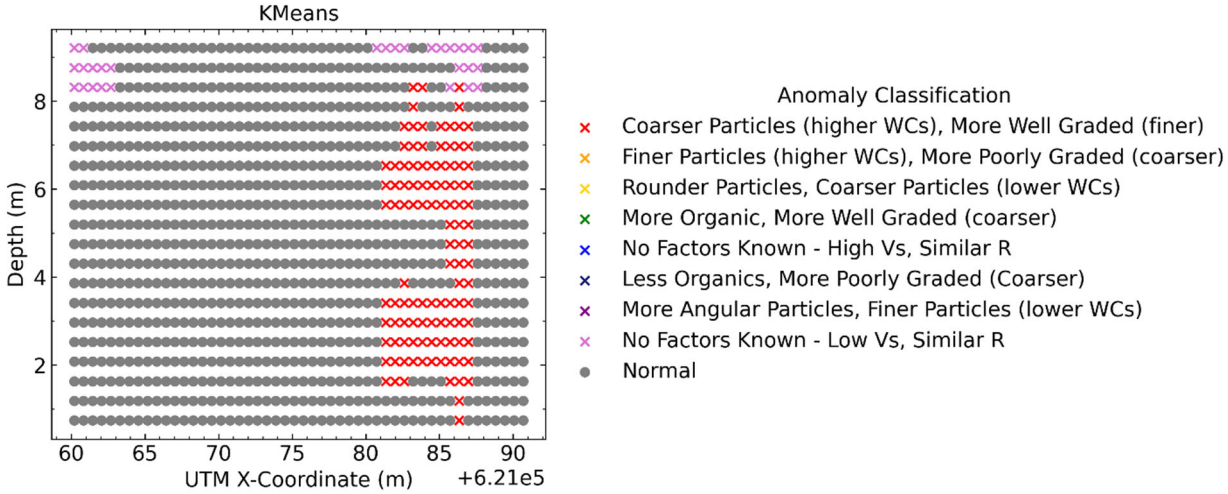
Profile: sc_site3



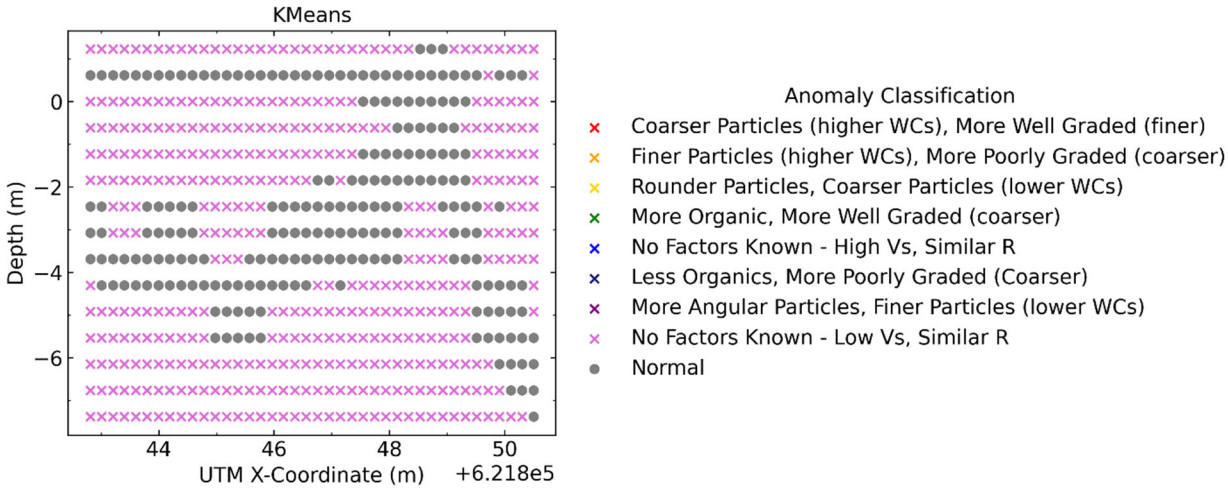
Profile: sc3_site1



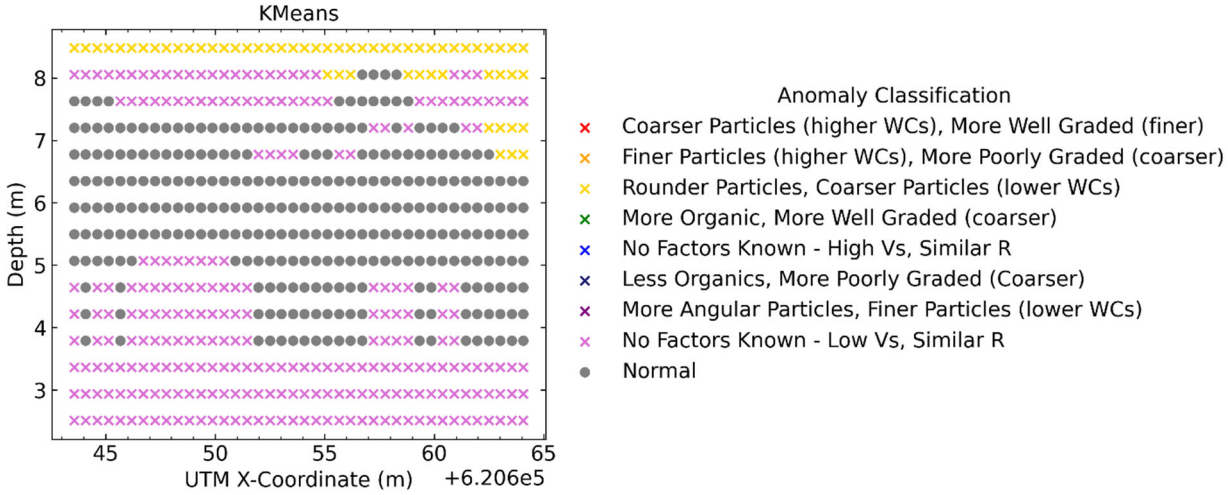
Profile: sc5_site1



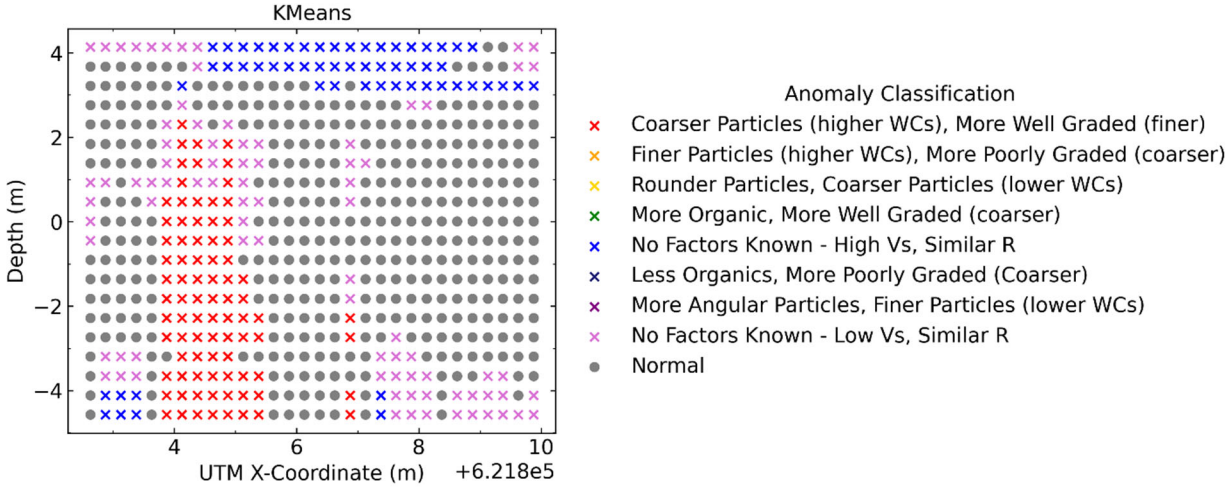
Profile: sdt_site2



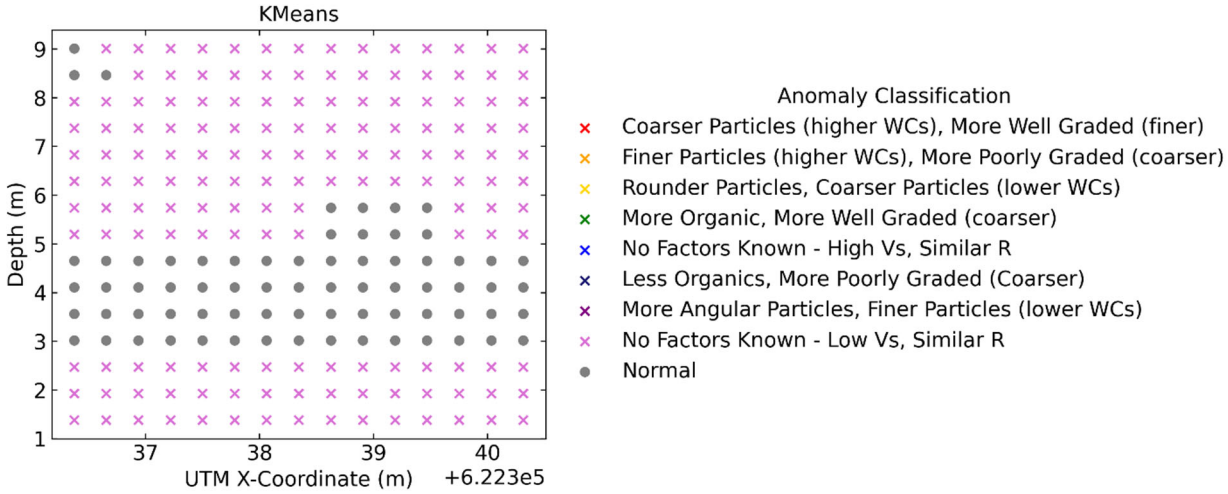
Profile: swt_site1



Profile: swt_site2

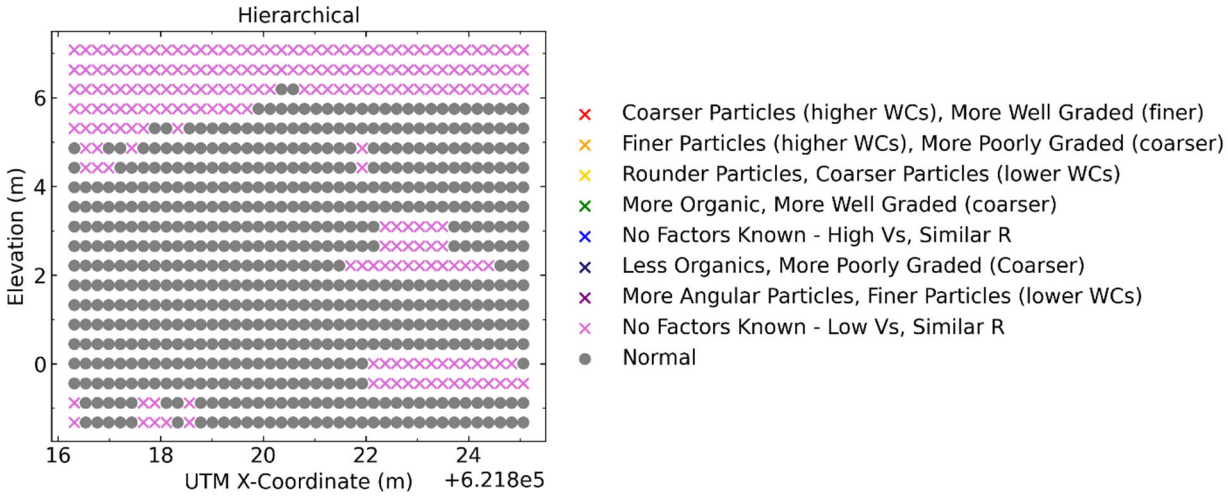


Profile: swt_site3

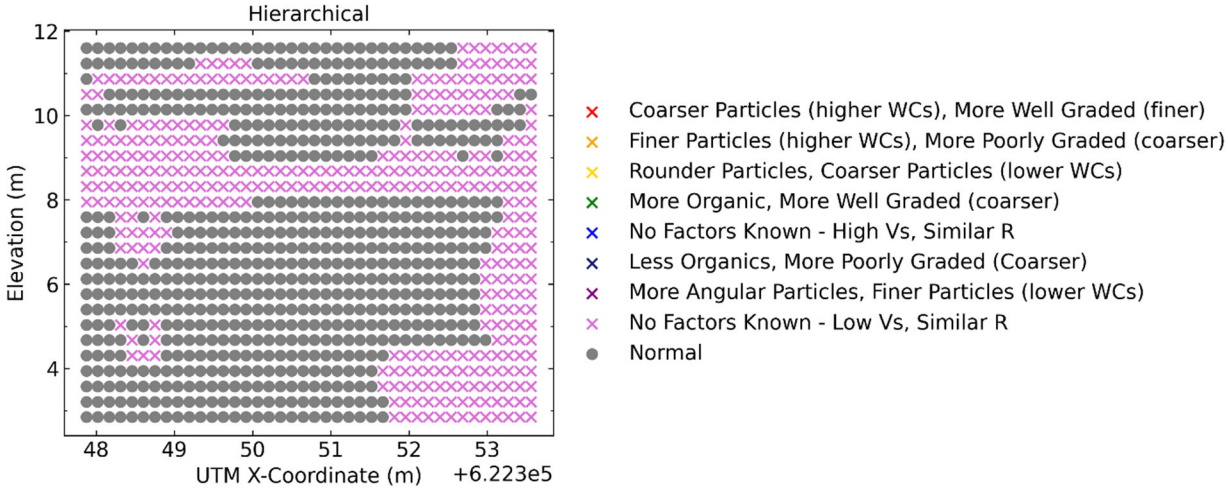


Fit: Wet Compilation
Algorithm: Hierarchical

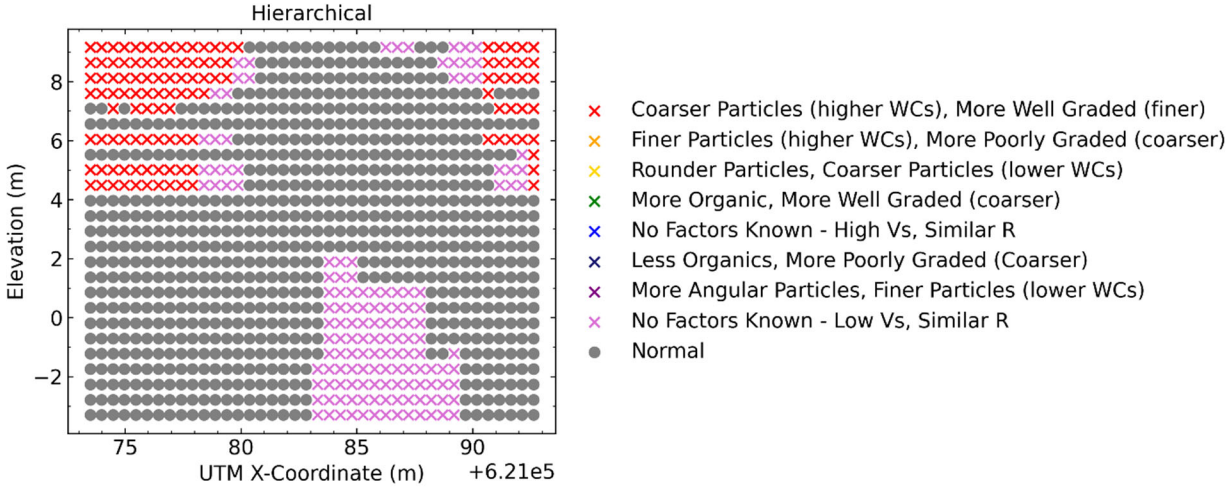
Profile: fc_site2



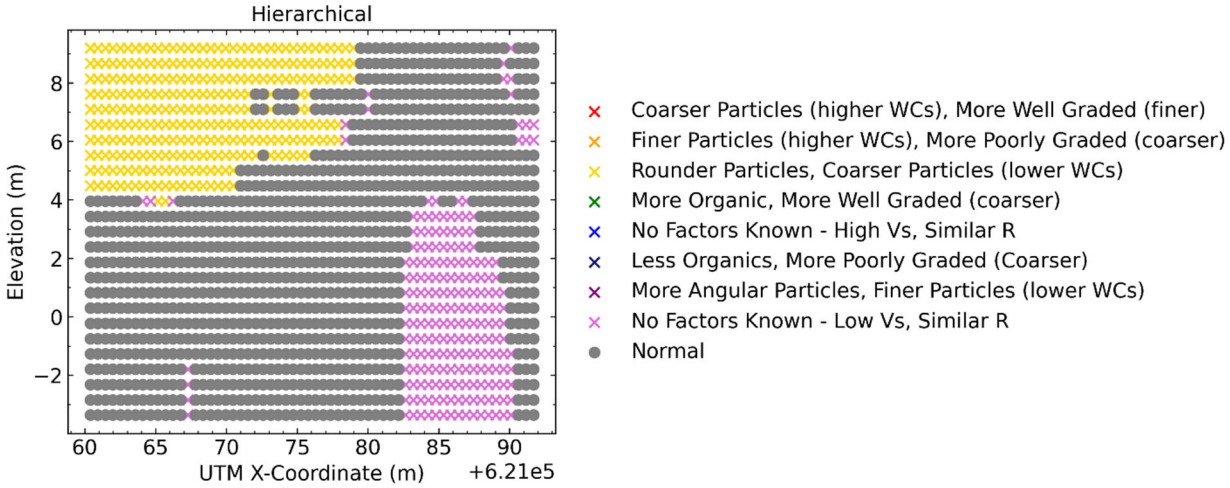
Profile: fc_site3



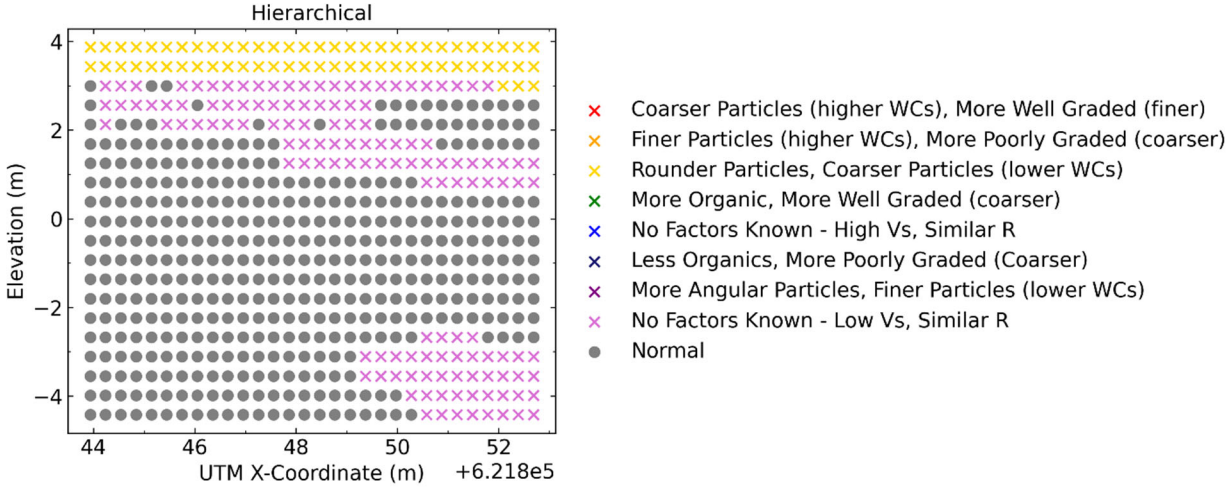
Profile: fc3_site1



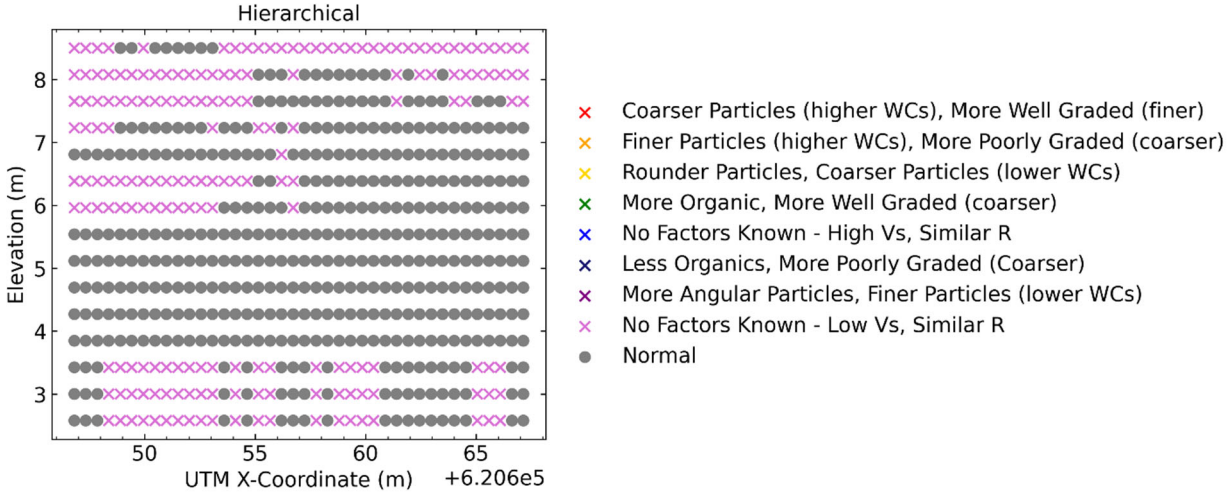
Profile: fc5_site1



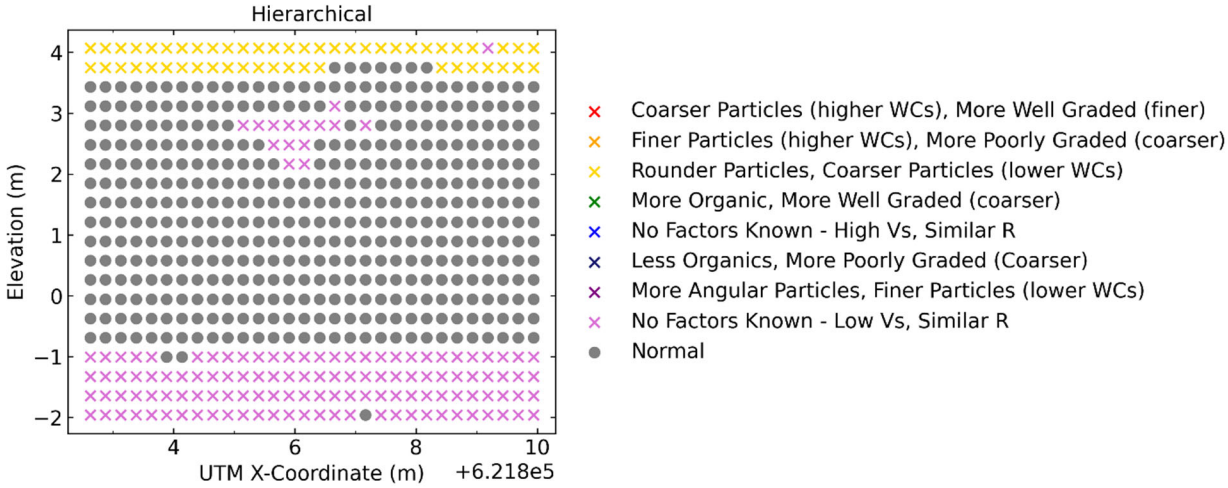
Profile: fdt_site2



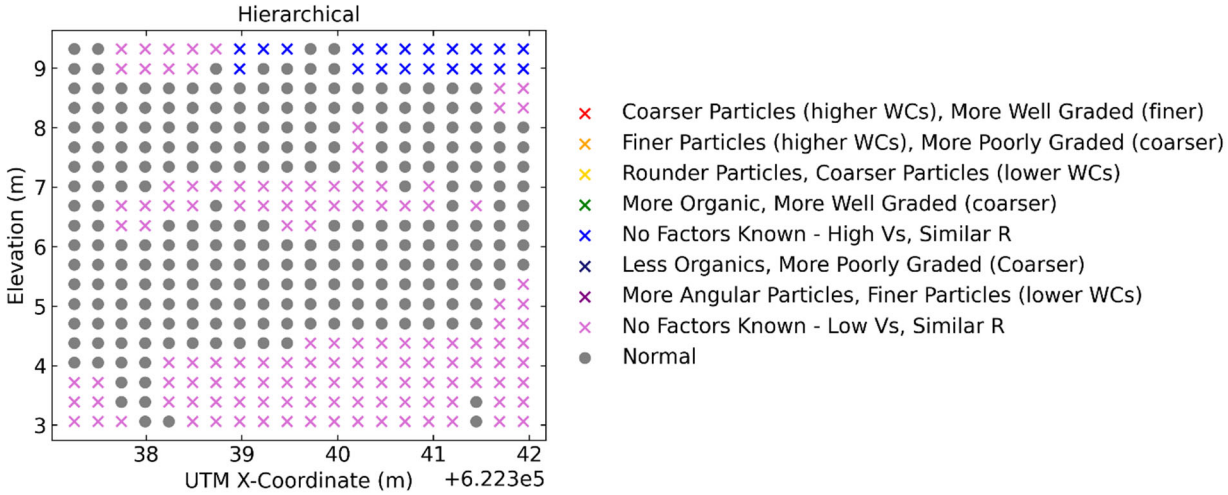
Profile: fwt_site1



Profile: fwt_site2

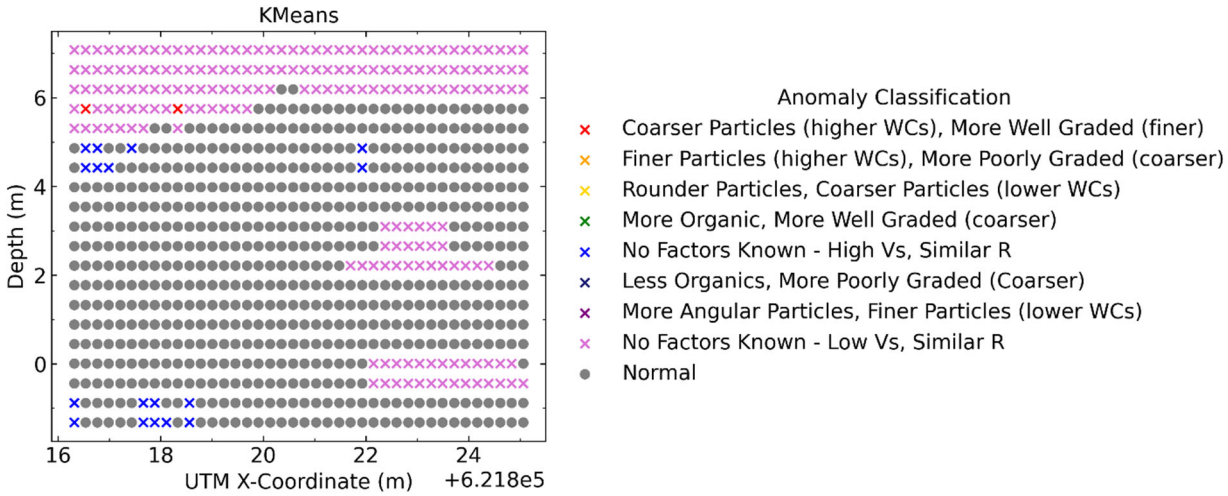


Profile: fwt_site3

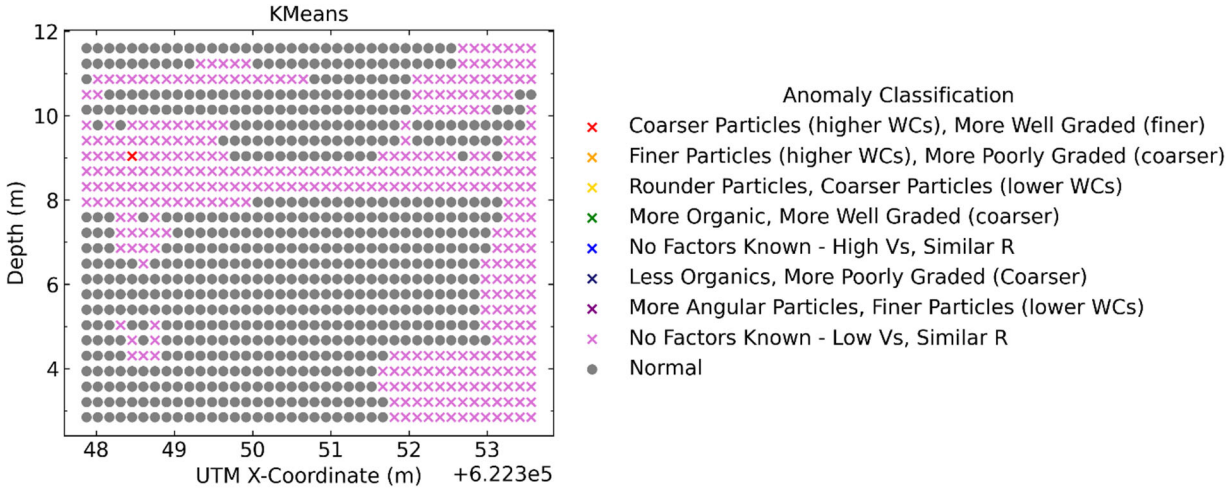


Fit: Wet Compilation
Algorithm: K-Means

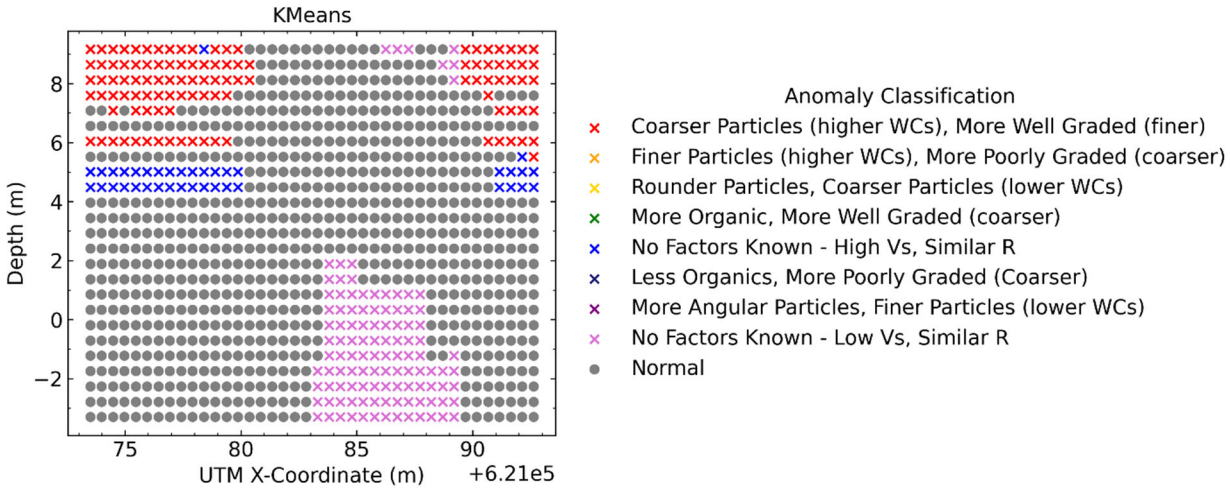
Profile: fc_site2



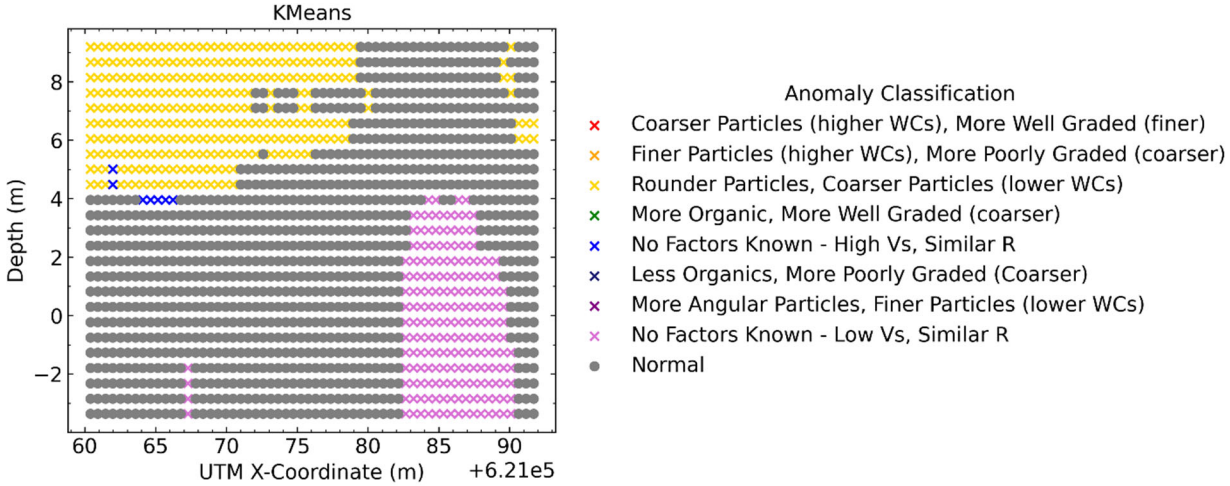
Profile: fc_site3



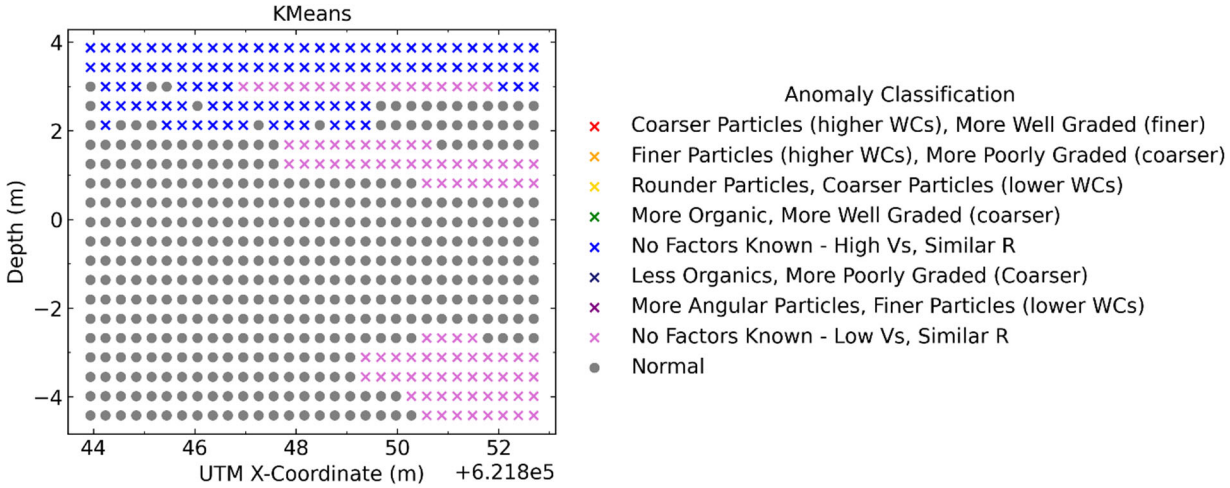
Profile: fc3_site1



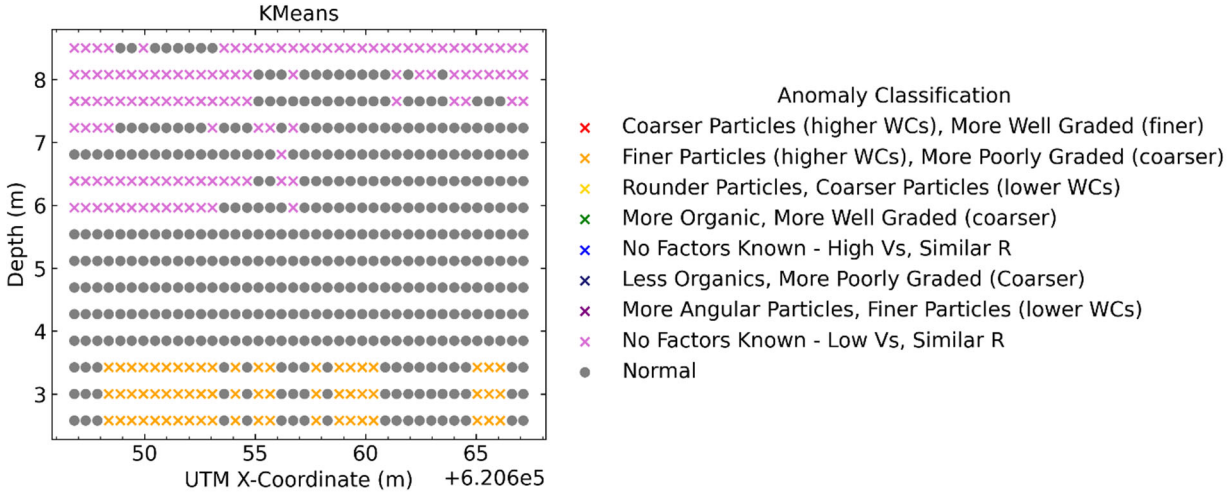
Profile: fc5_site1



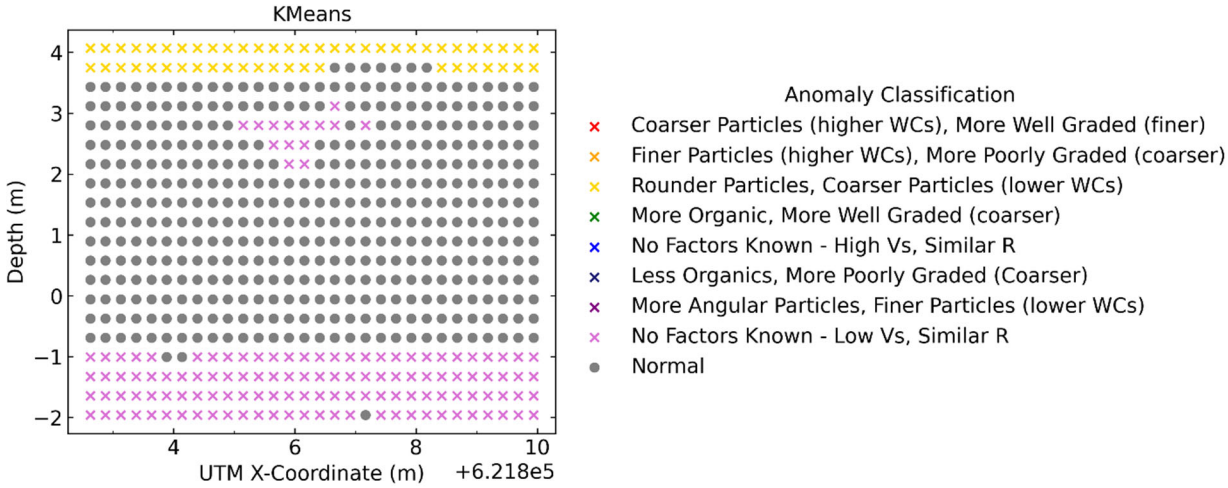
Profile: fdt_site2



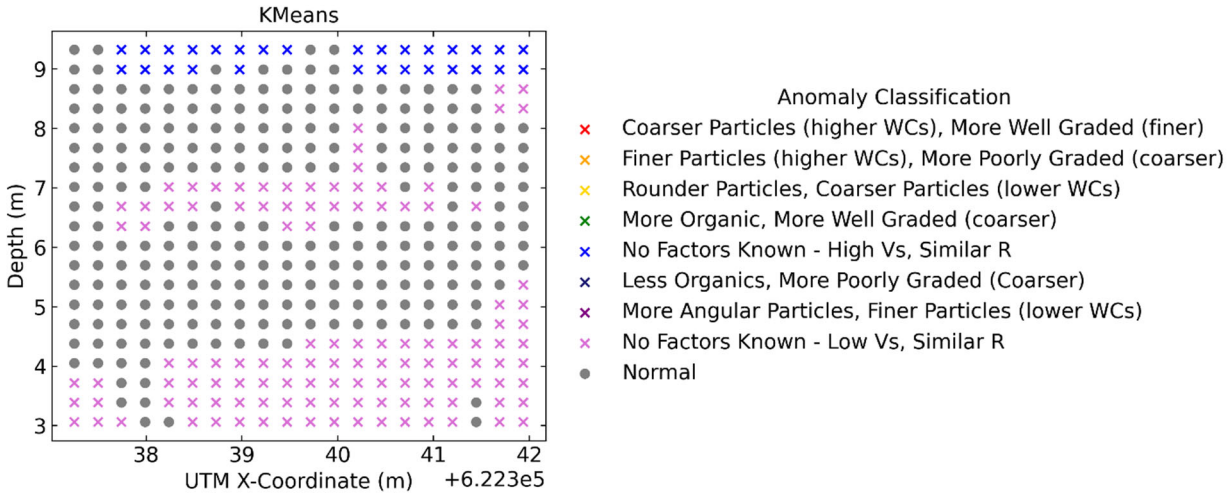
Profile: fwt_site1



Profile: fwt_site2

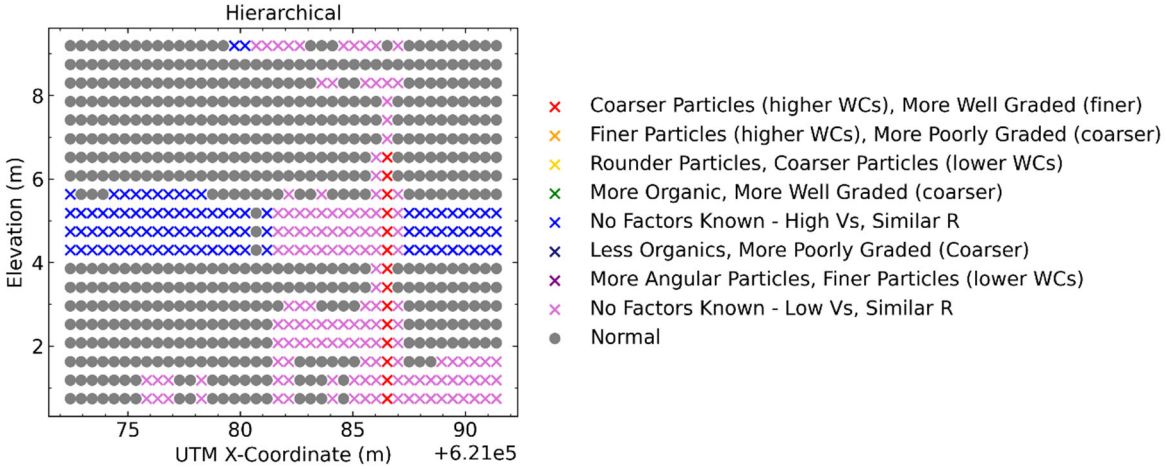


Profile: fwt_site3

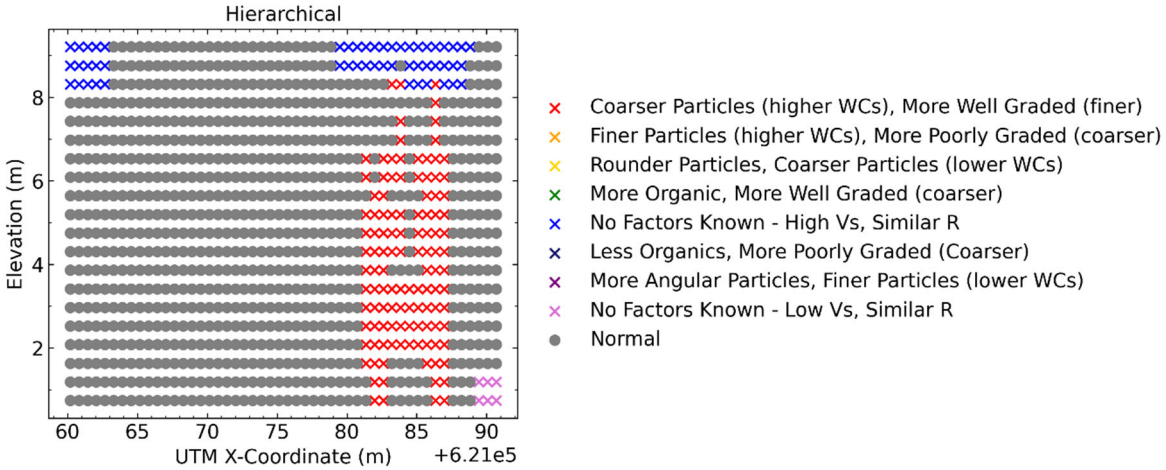


Fit: Dry Site 1 Compilation
Algorithm: Hierarchical

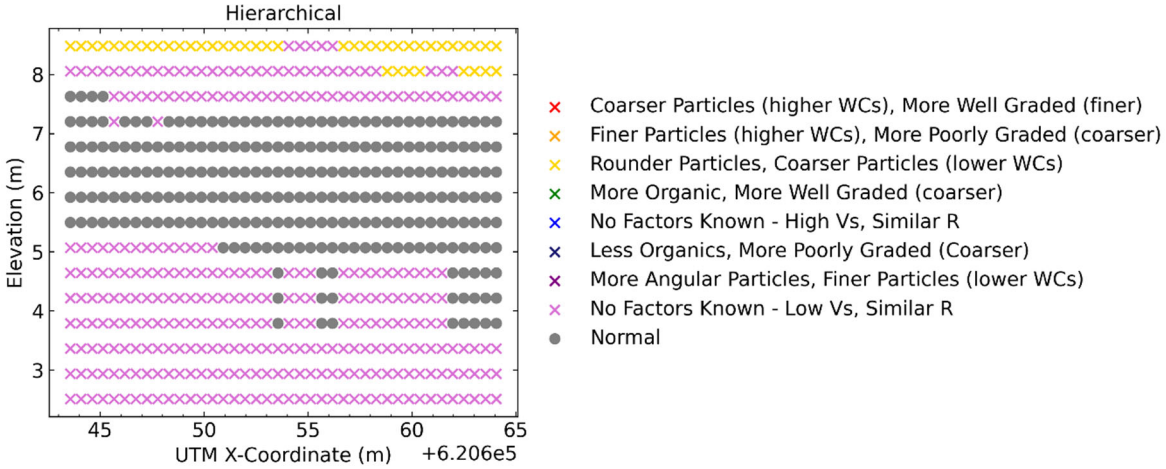
Profile: sc3_site1



Profile: sc5_site1

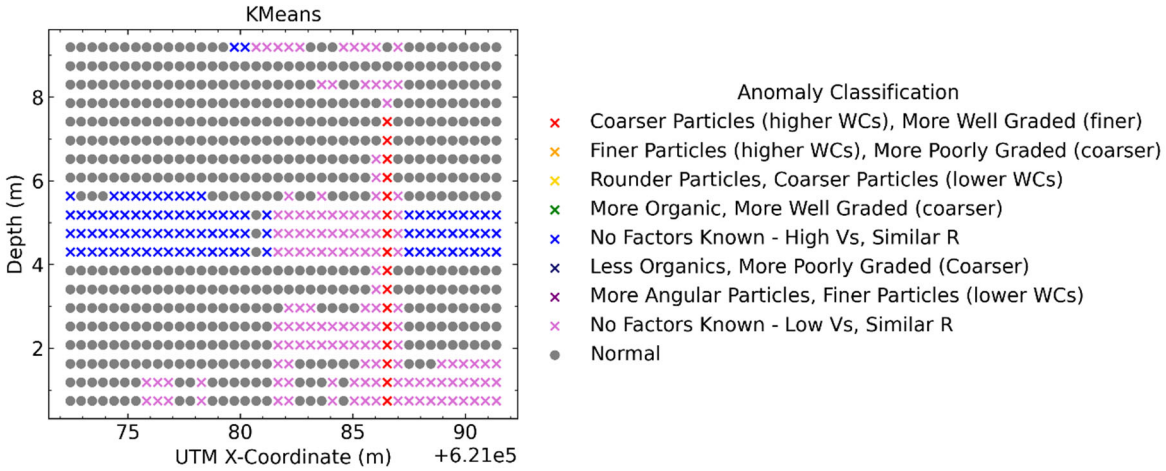


Profile: swt_site1

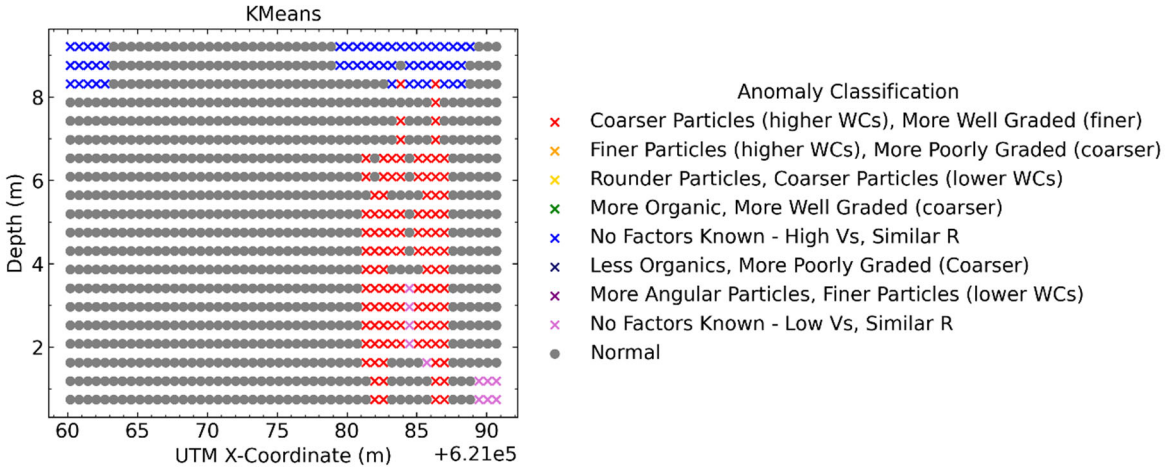


Fit: Dry Site 1 Compilation
Algorithm: K-Means

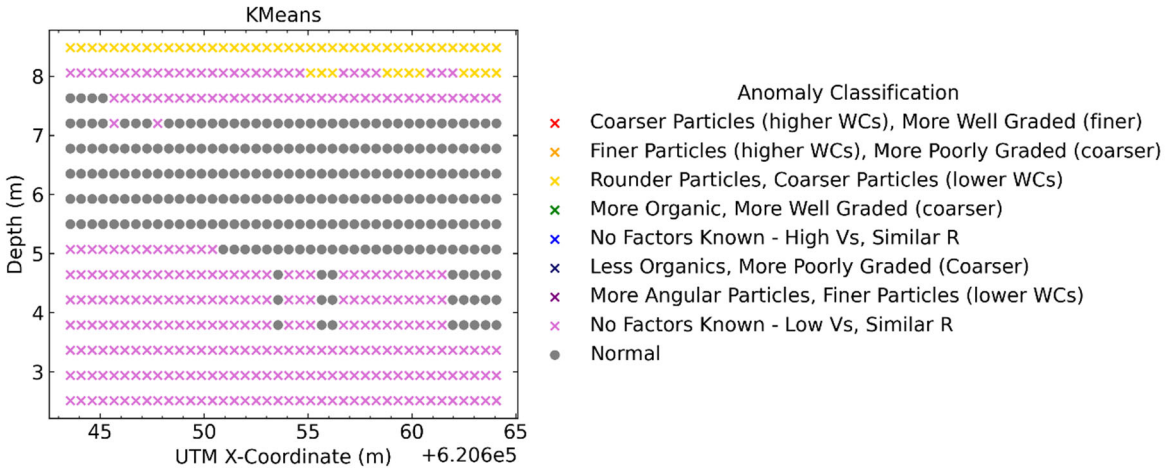
Profile: sc3_site1



Profile: sc5_site1

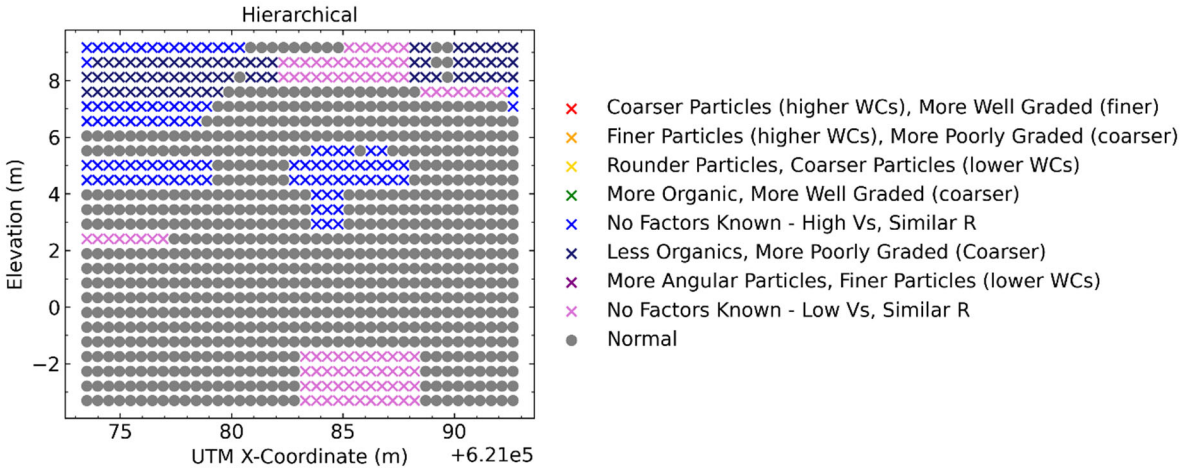


Profile: swt_site1

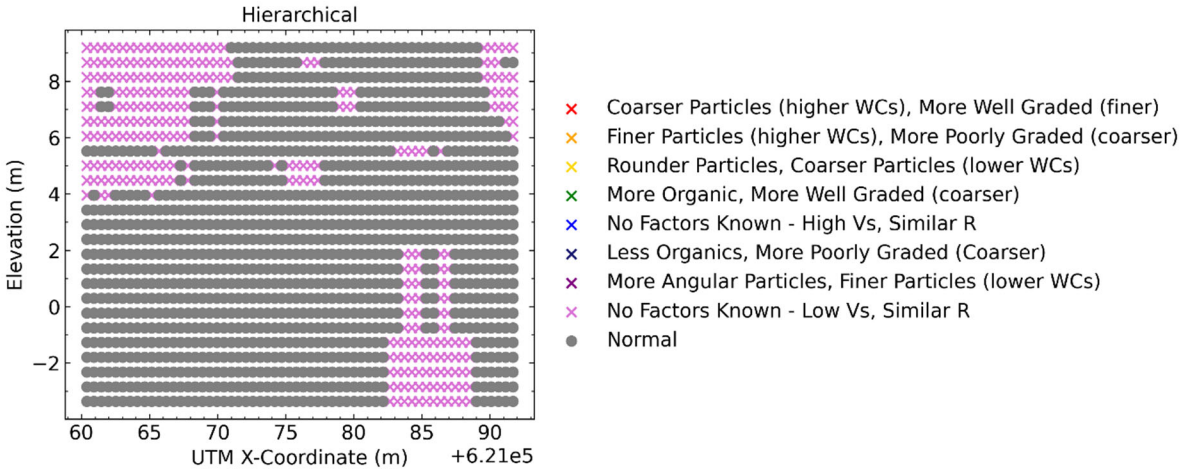


Fit: Wet Site 1 Compilation
Algorithm: Hierarchical

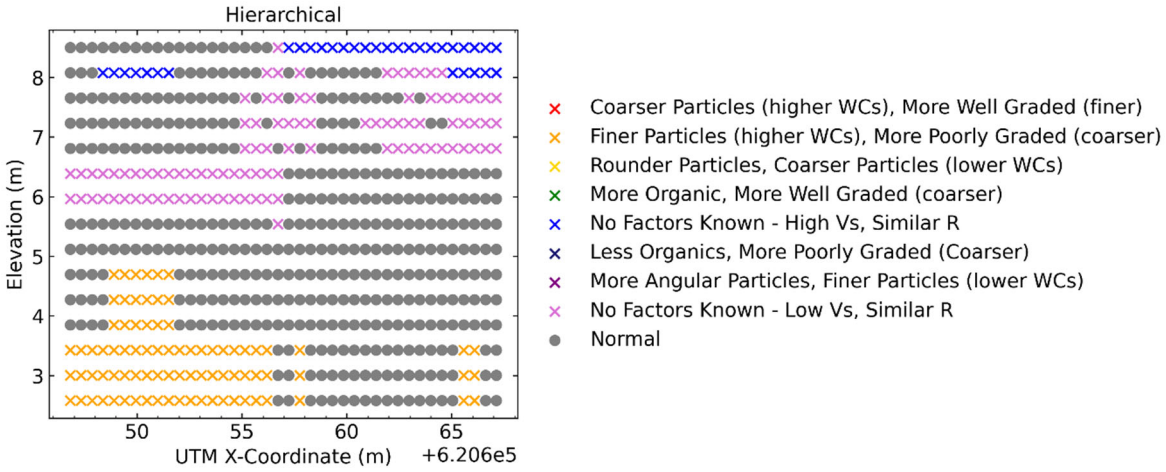
Profile: fc3_site1



Profile: fc5_site1

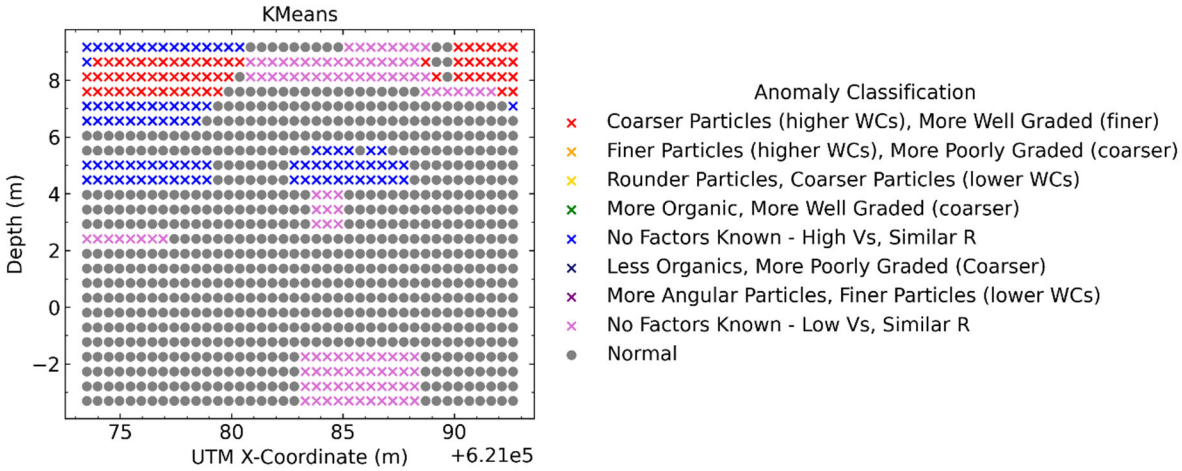


Profile: fwt_site1

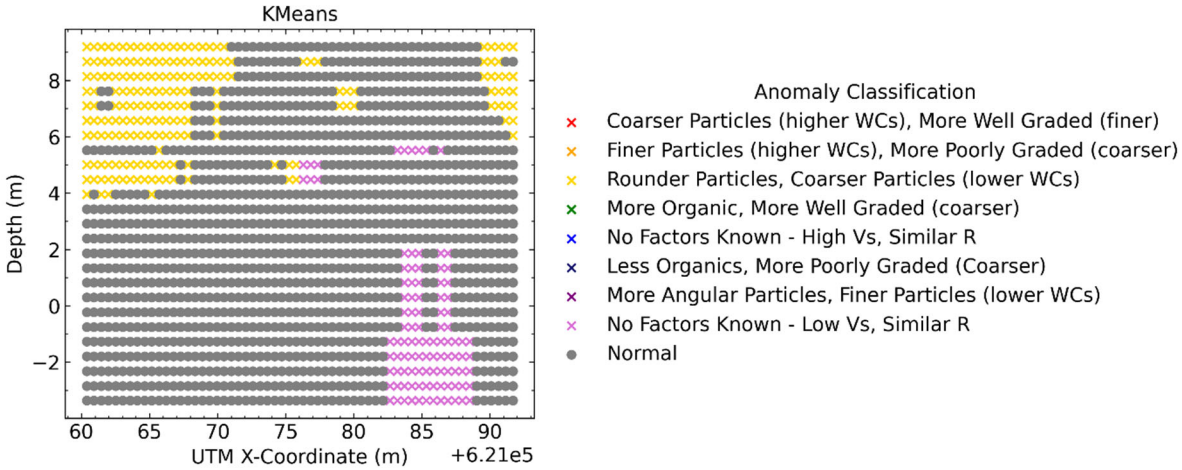


Fit: Wet Site 1 Compilation
Algorithm: K-Means

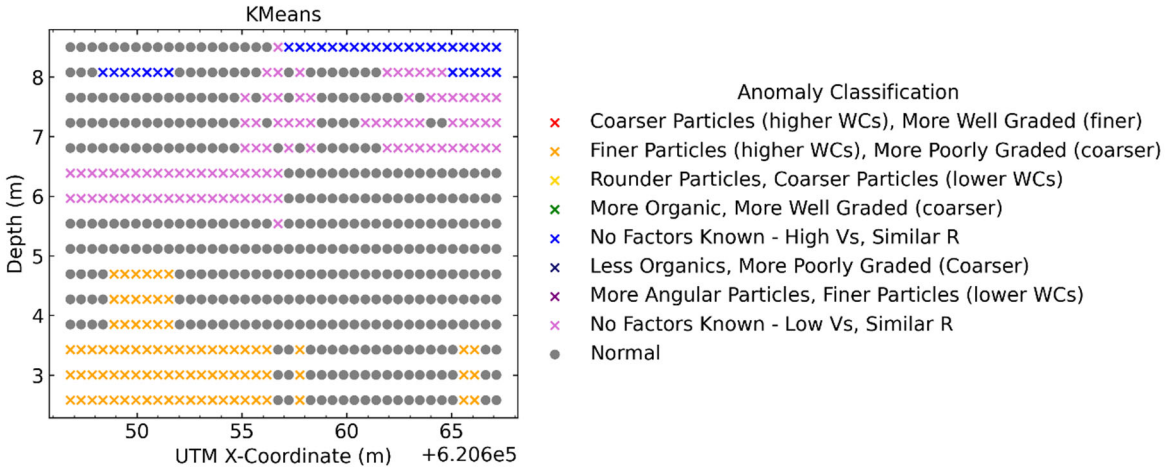
Profile: fc3_site1



Profile: fc5_site1

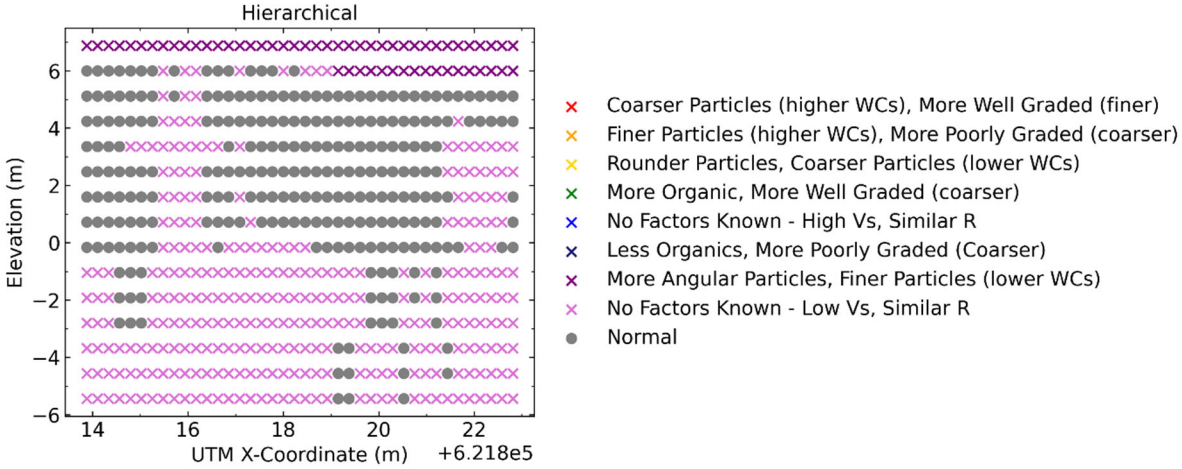


Profile: fwt_site1

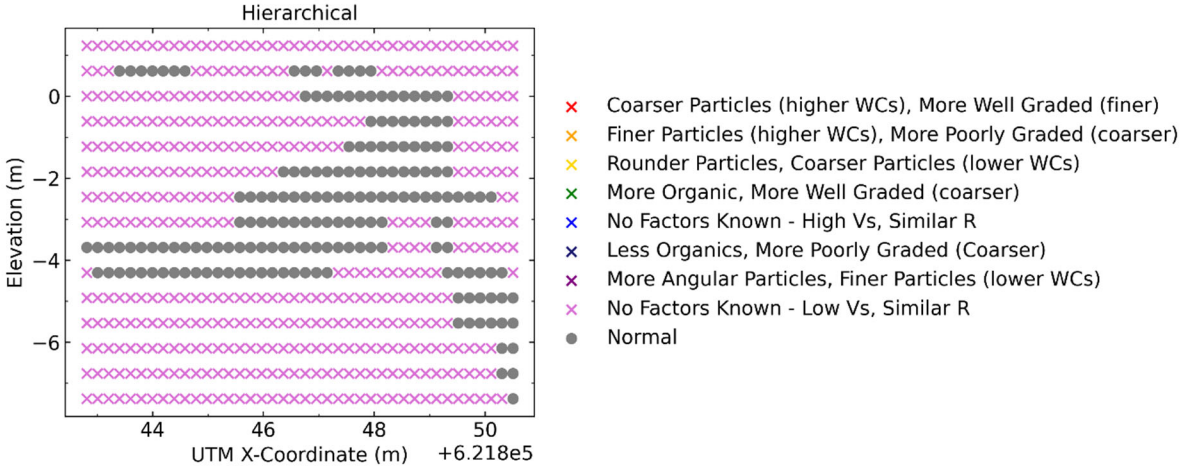


Fit: Dry Site 2 Compilation
Algorithm: Hierarchical

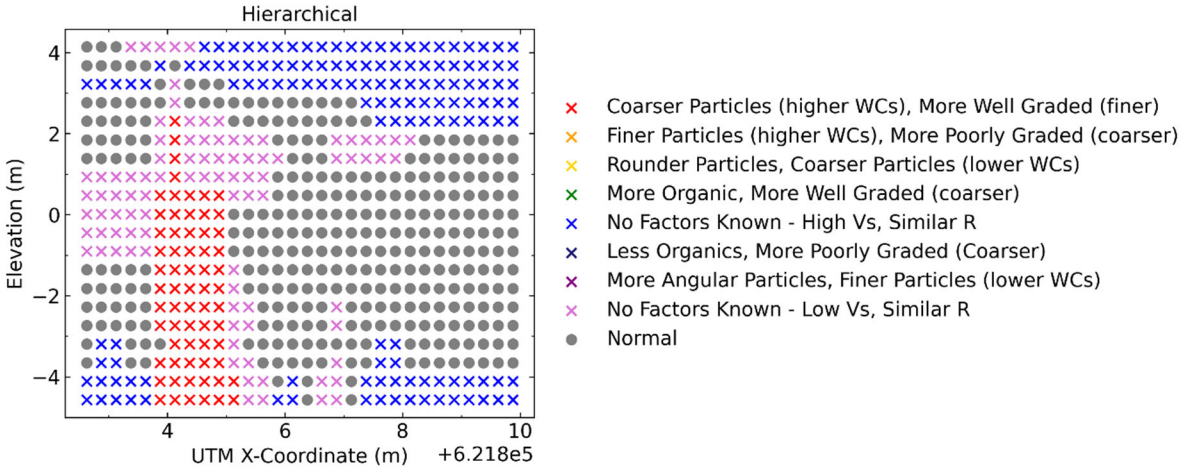
Profile: sc_site2



Profile: sdt_site2

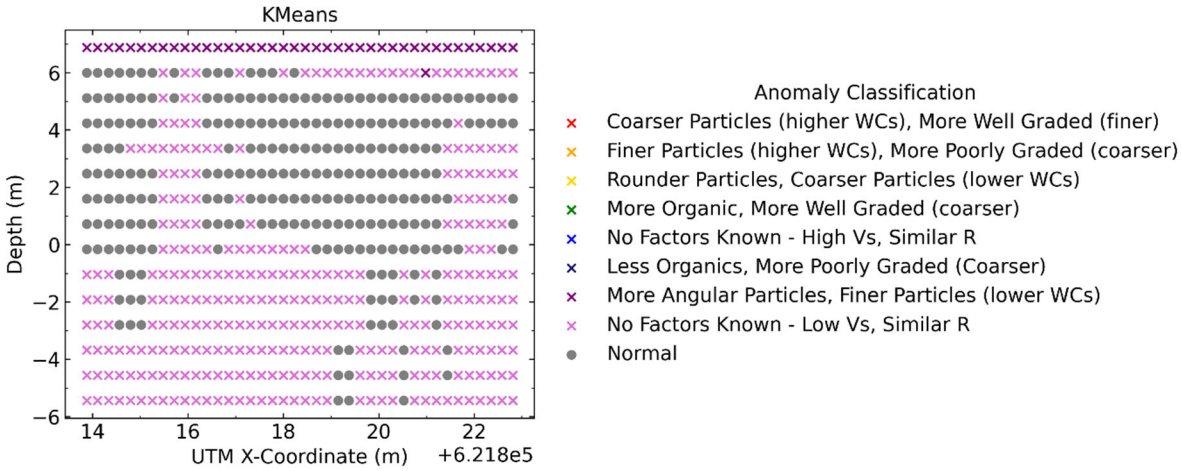


Profile: swt_site2

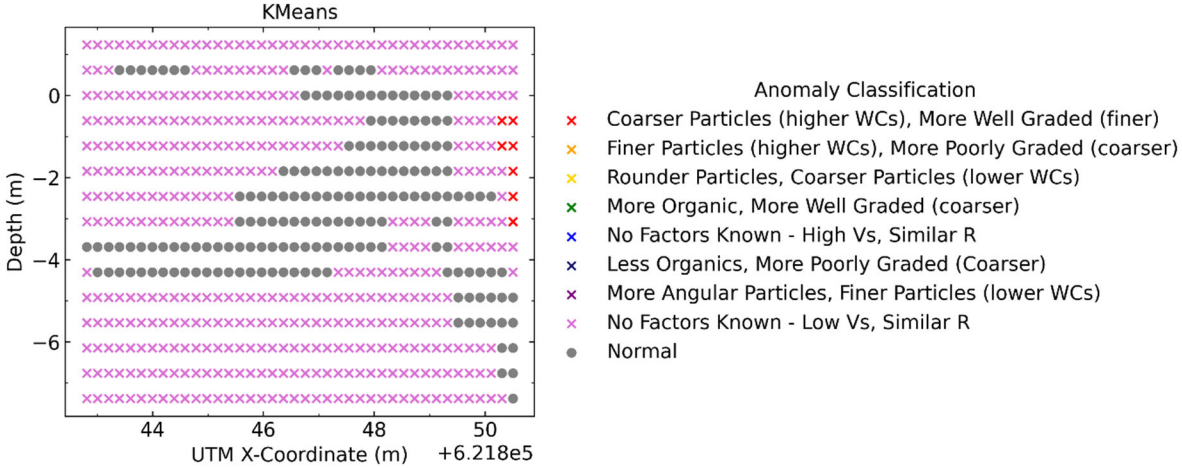


Fit: Dry Site 2 Compilation
Algorithm: K-Means

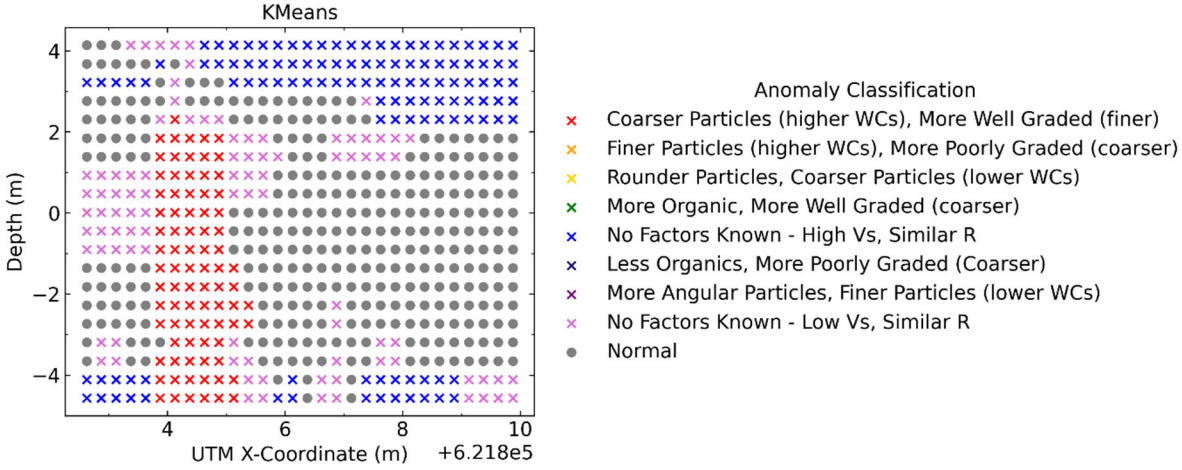
Profile: sc_site2



Profile: sdt_site2

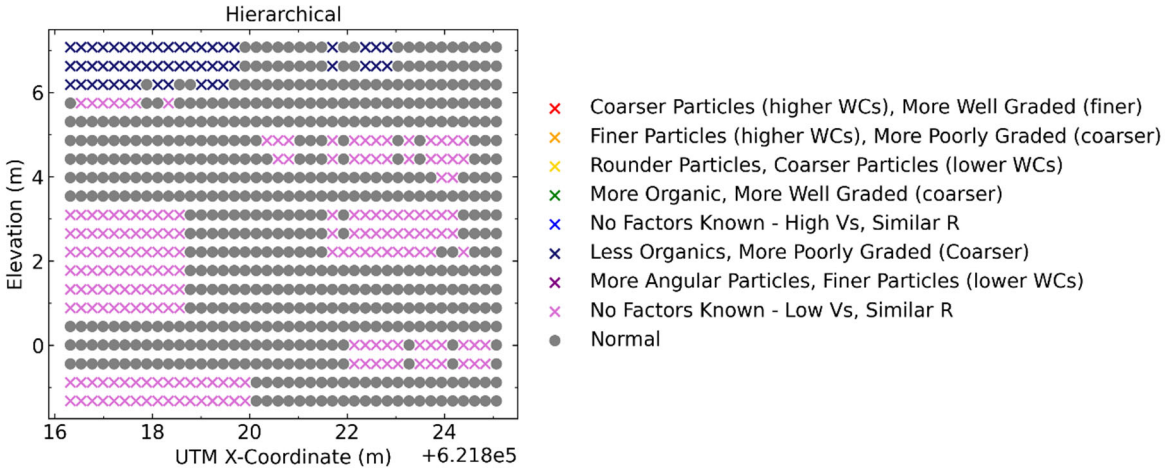


Profile: swt_site2

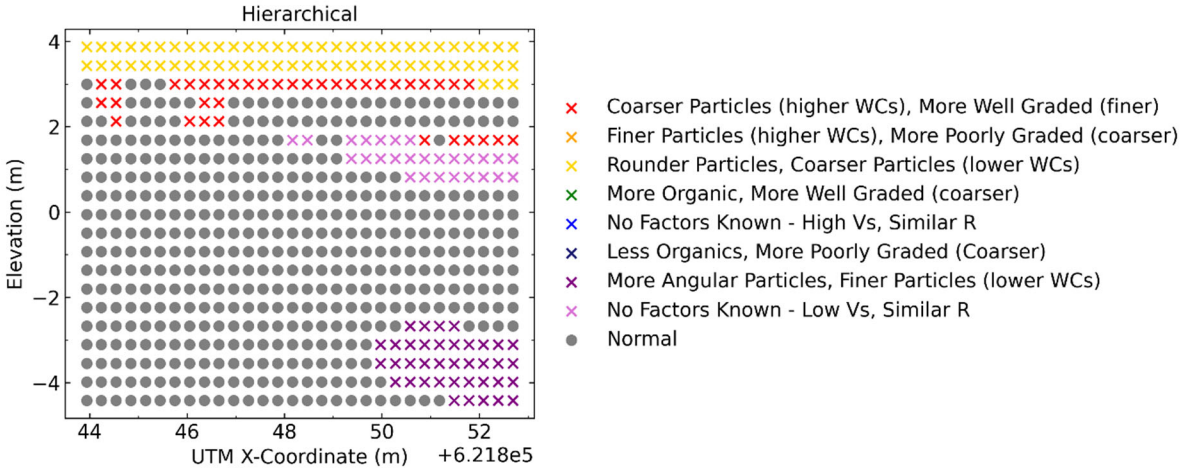


Fit: Wet Site 2 Compilation
Algorithm: Hierarchical

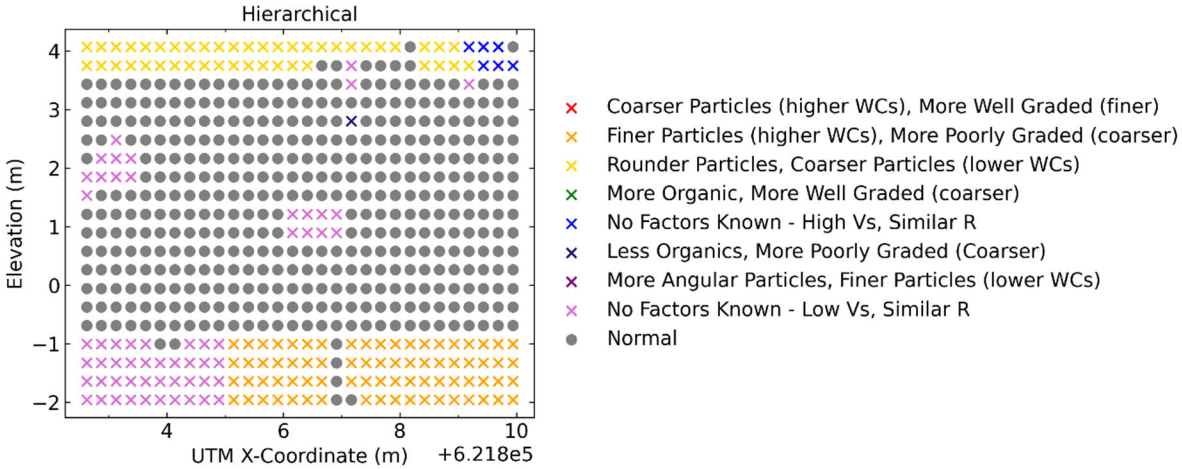
Profile: fc_site2



Profile: fdt_site2

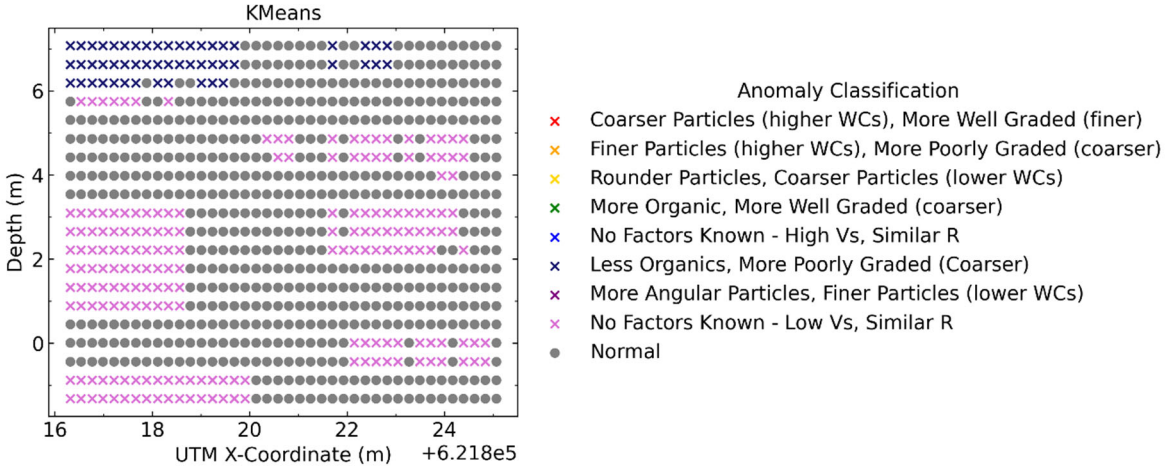


Profile: fwt_site2

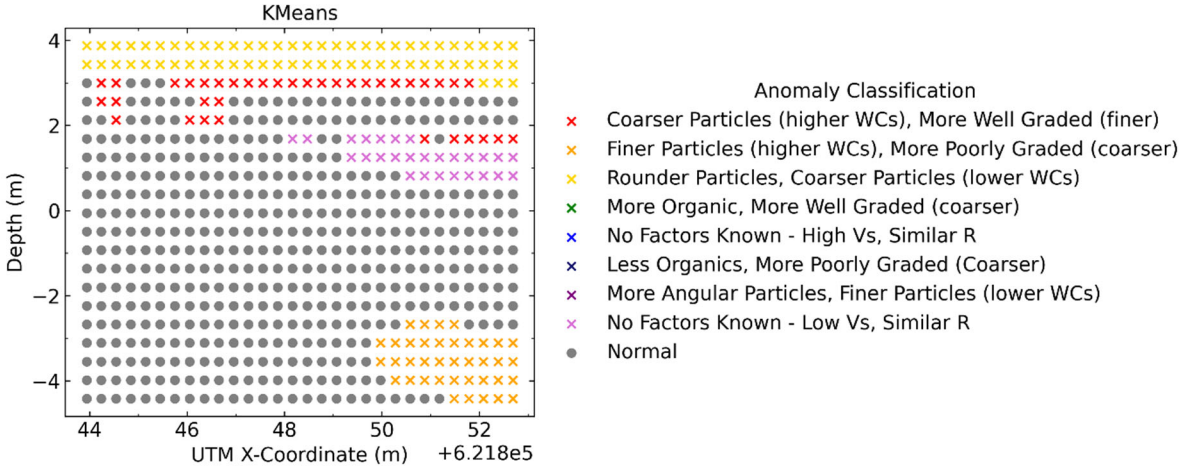


Fit: Wet Site 2 Compilation
Algorithm: K-Means

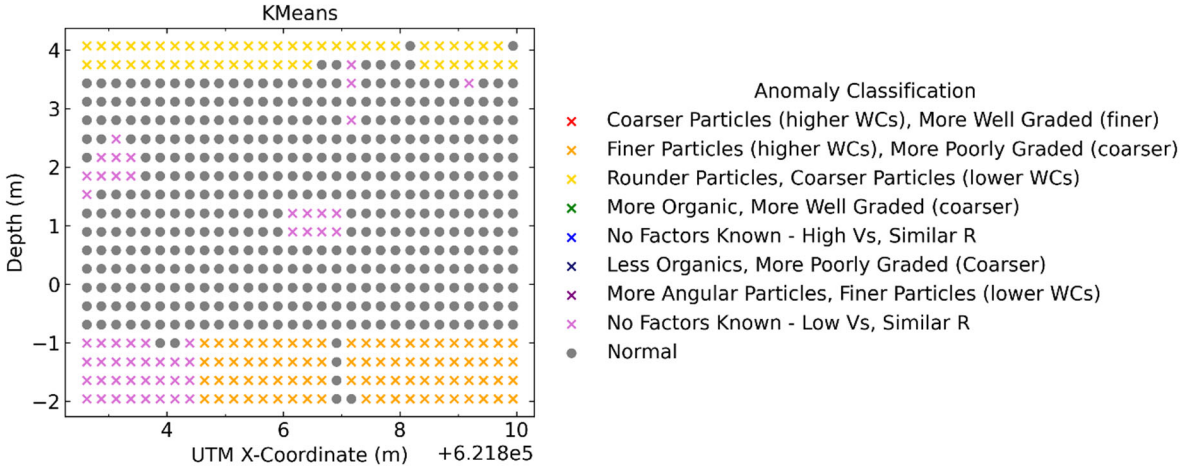
Profile: fc_site2



Profile: fdt_site2

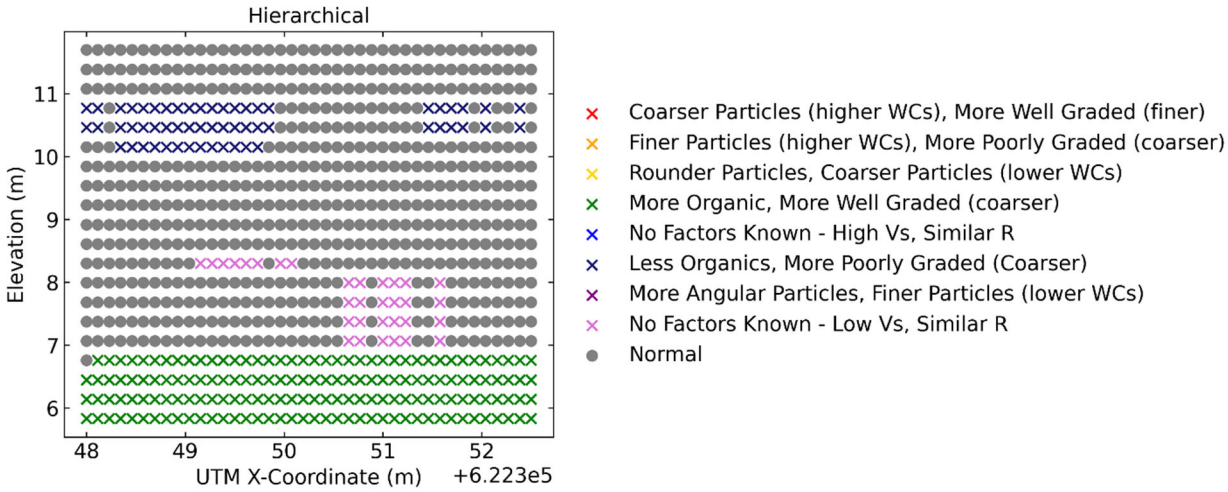


Profile: fwt_site2

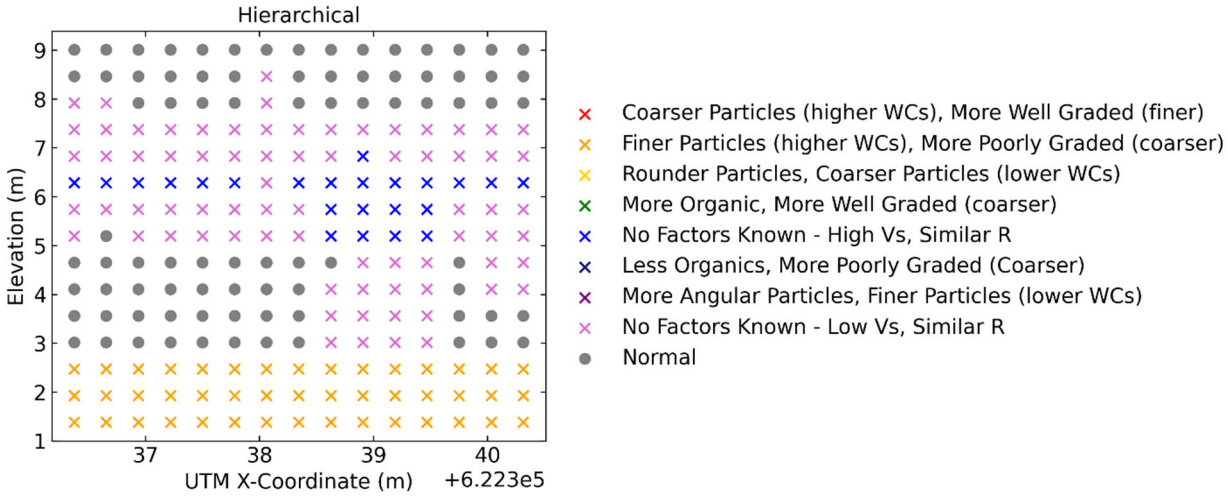


Fit: Dry Site 3 Compilation
Algorithm: Hierarchical

Profile: sc_site3

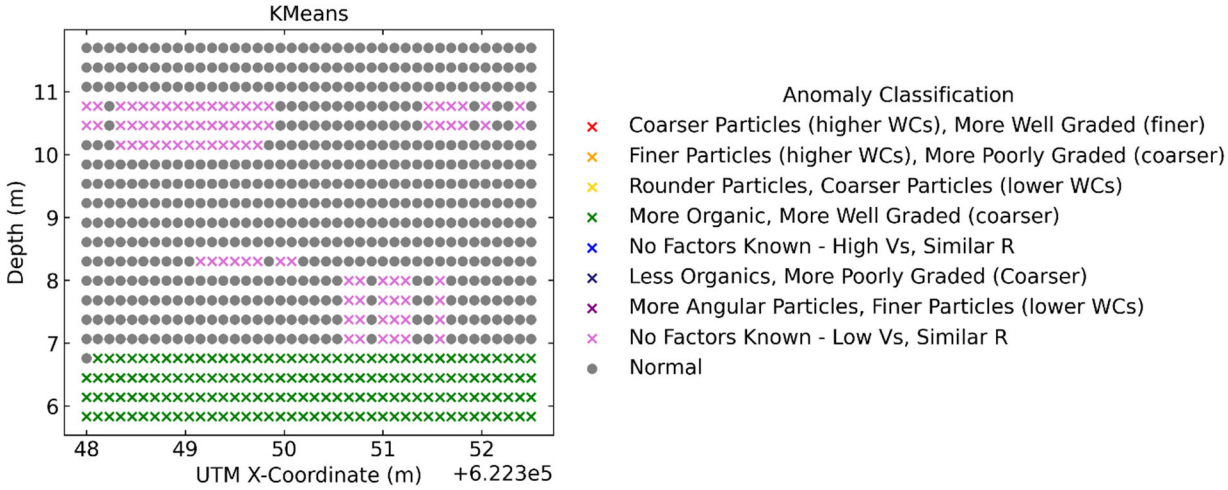


Profile: swt_site3

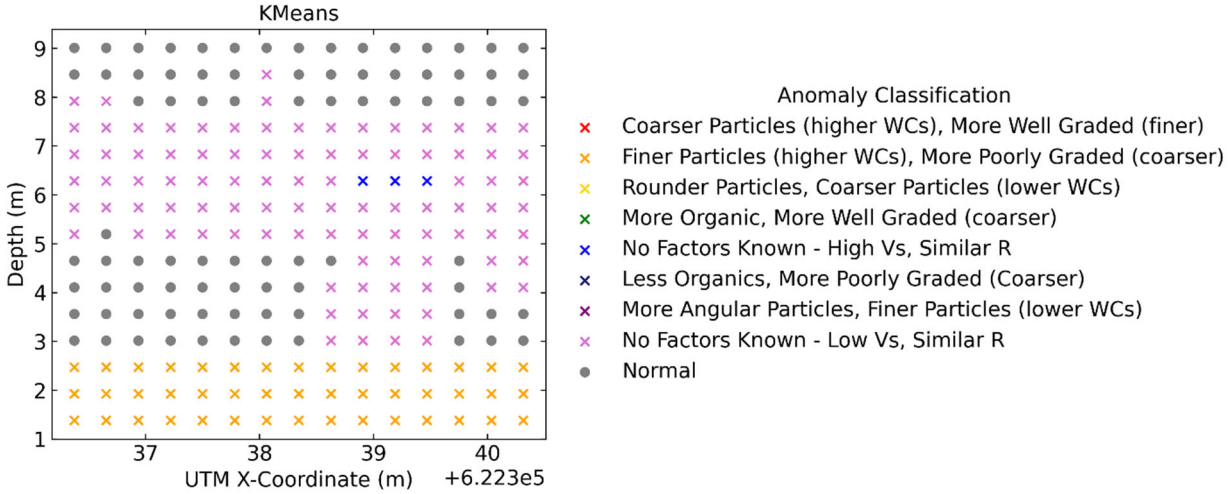


Fit: Dry Site 3 Compilation
Algorithm: K-Means

Profile: sc_site3

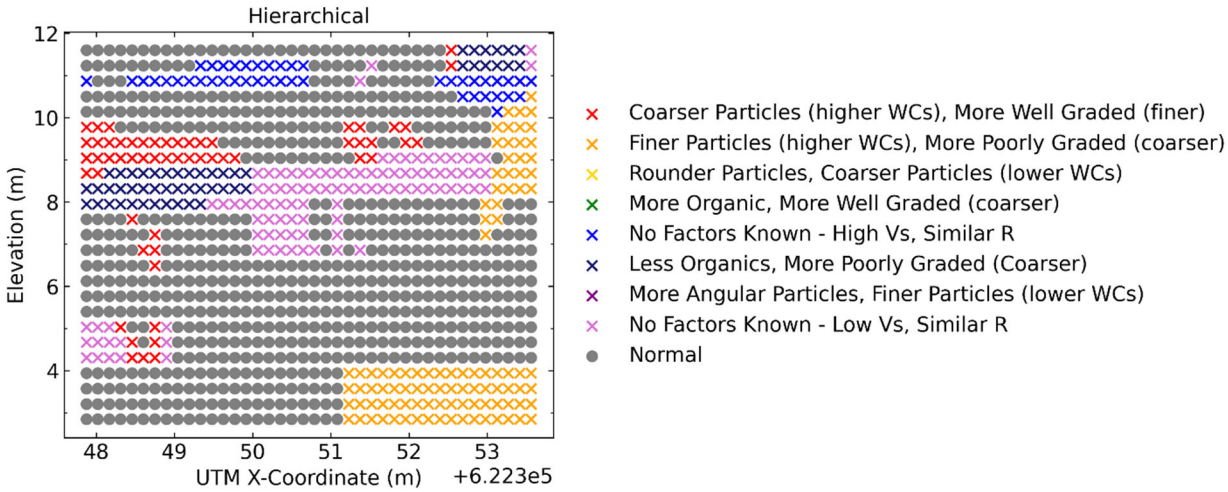


Profile: swt_site3

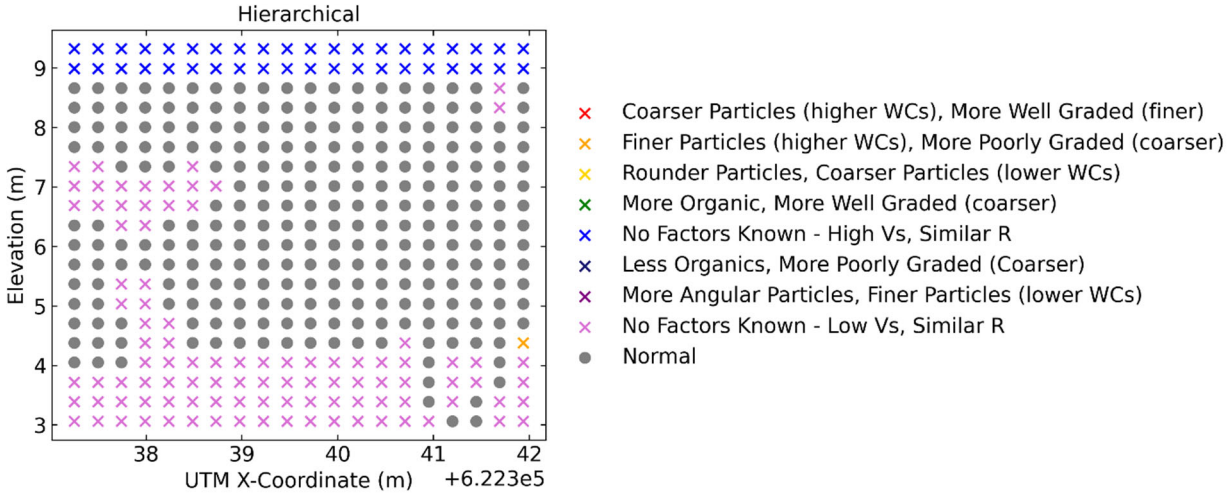


Fit: Wet Site 3 Compilation
Algorithm: Hierarchical

Profile: fc_site3

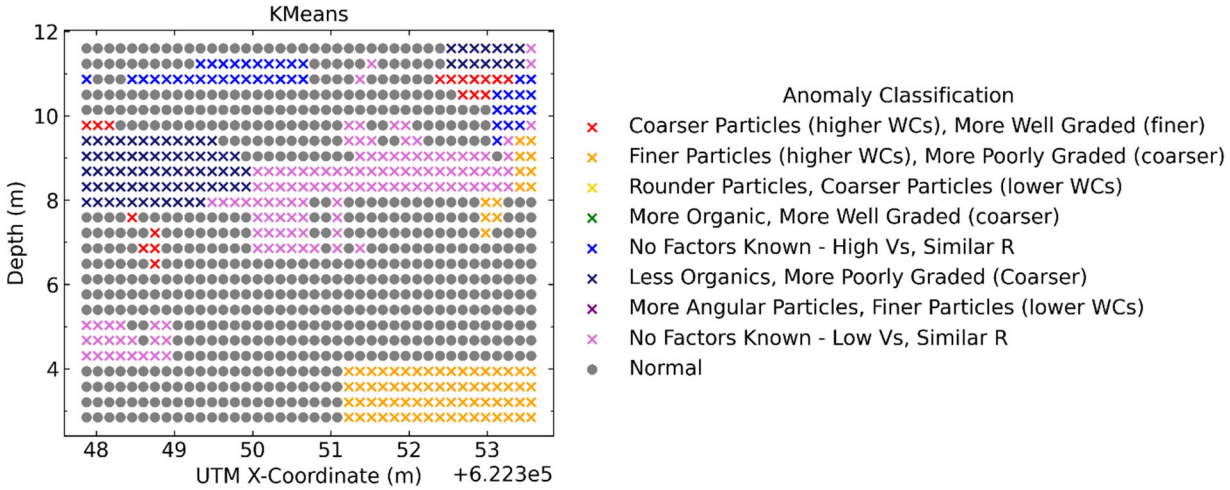


Profile: fwt_site3

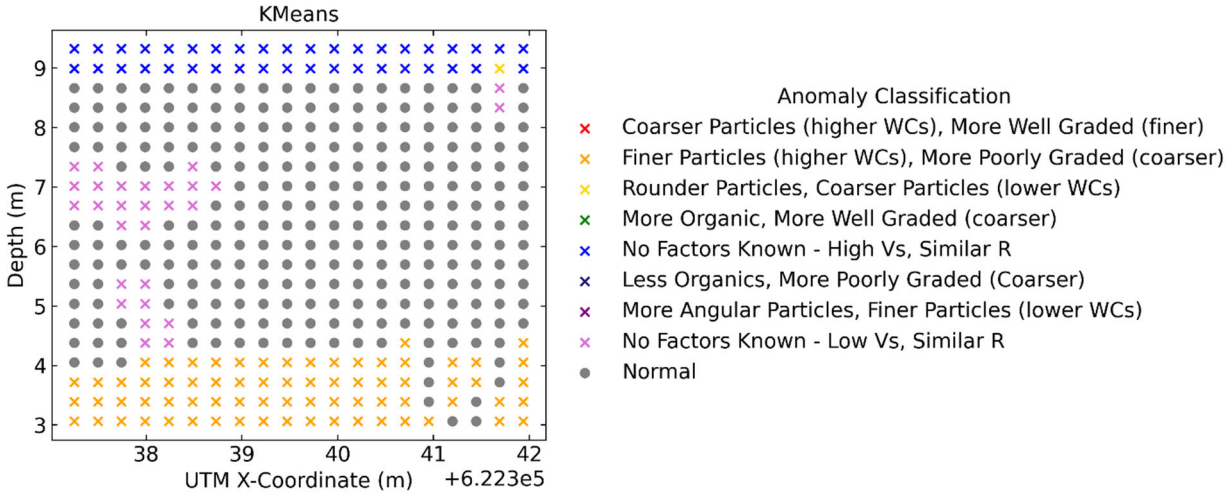


Fit: Wet Site 3 Compilation
Algorithm: K-Means

Profile: fc_site3

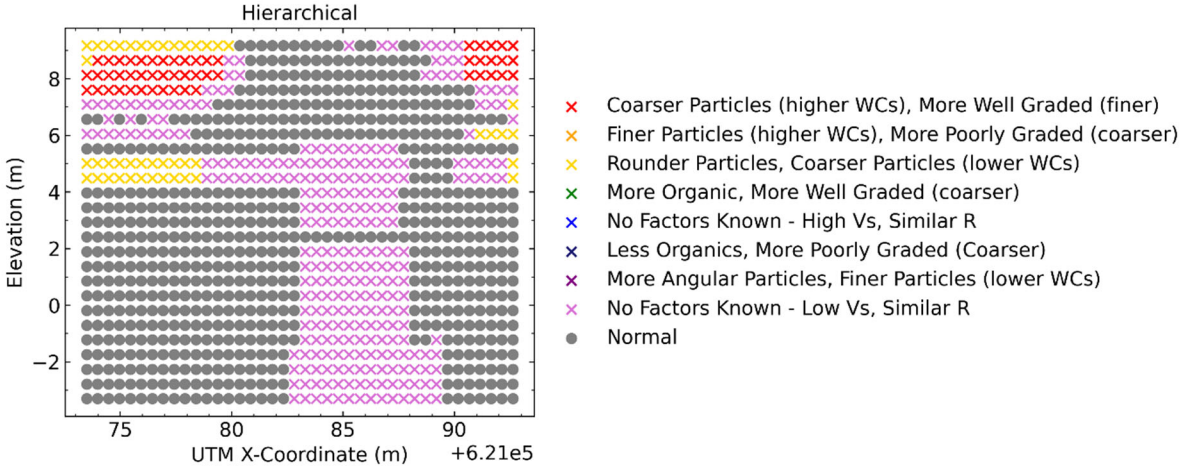


Profile: fwt_site3

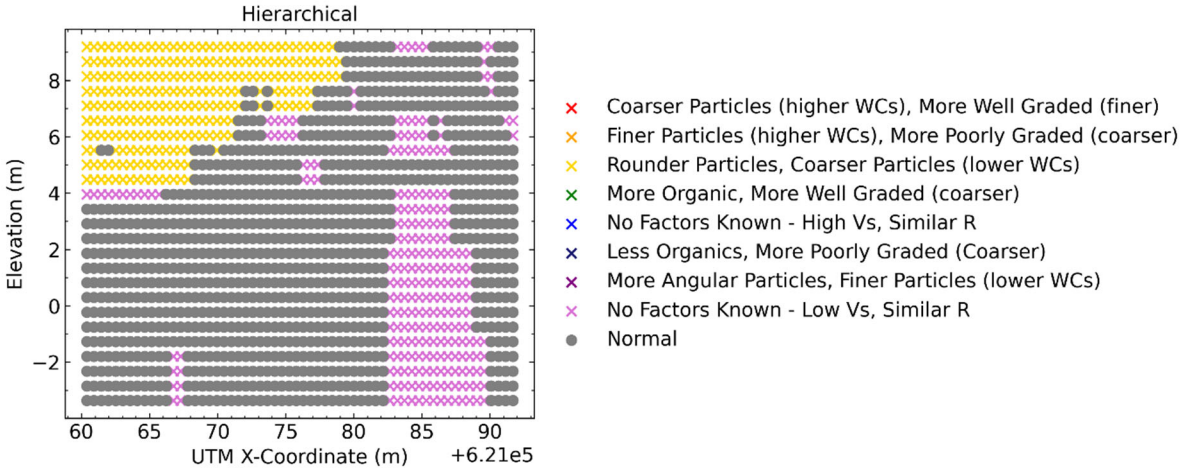


Fit: Site 1 Compilation
Algorithm: Hierarchical

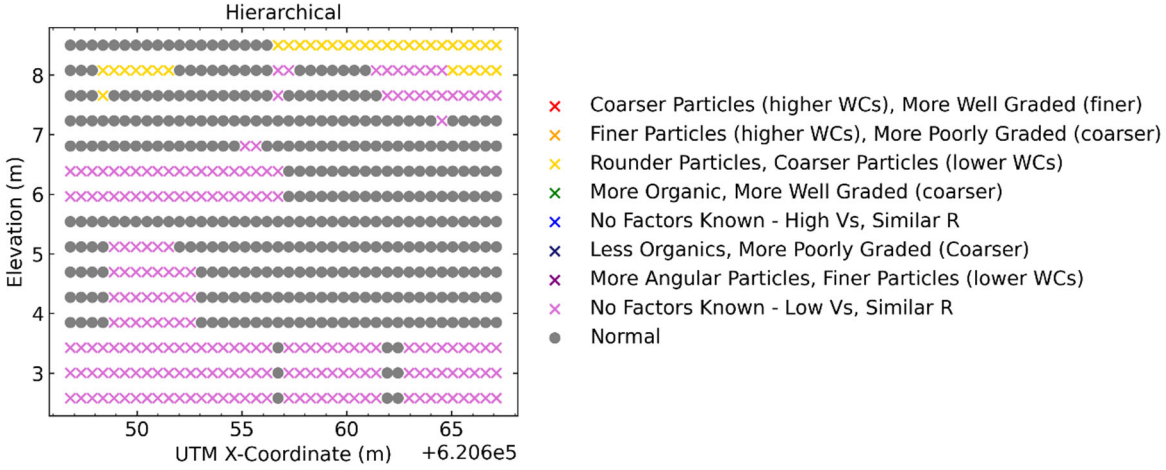
Profile: fc3_site1



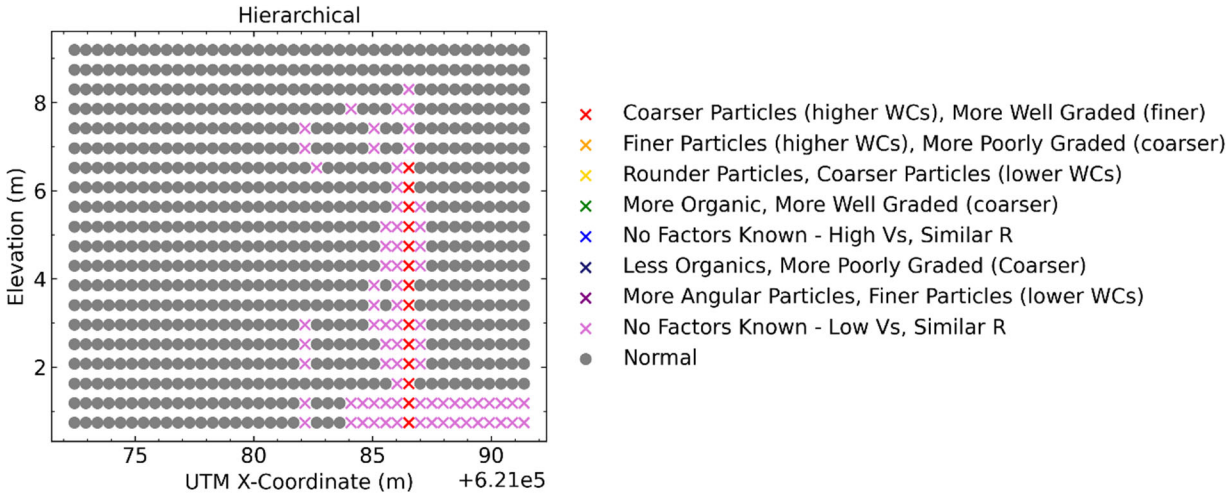
Profile: fc5_site1



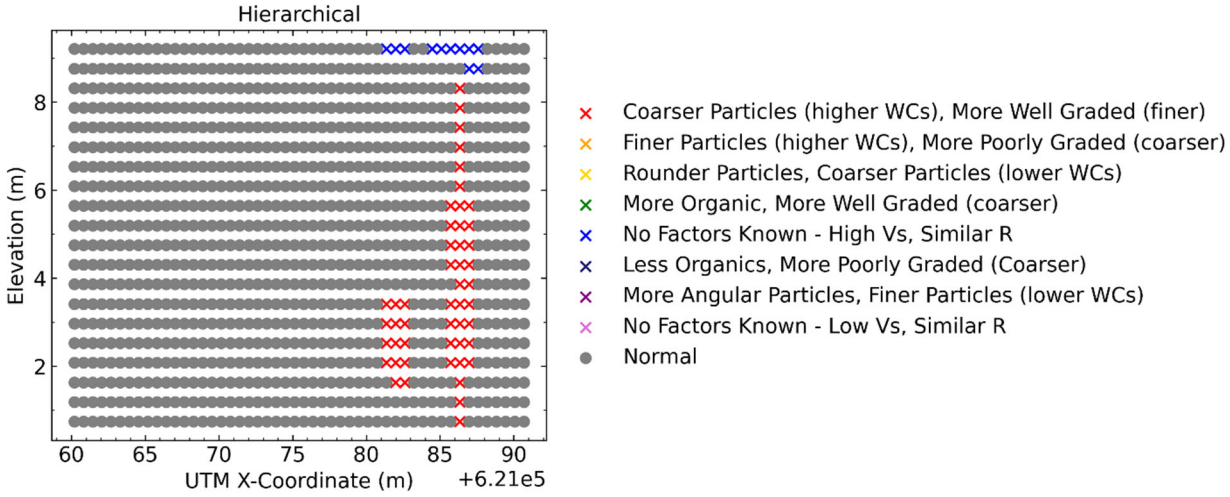
Profile: fwt_site1



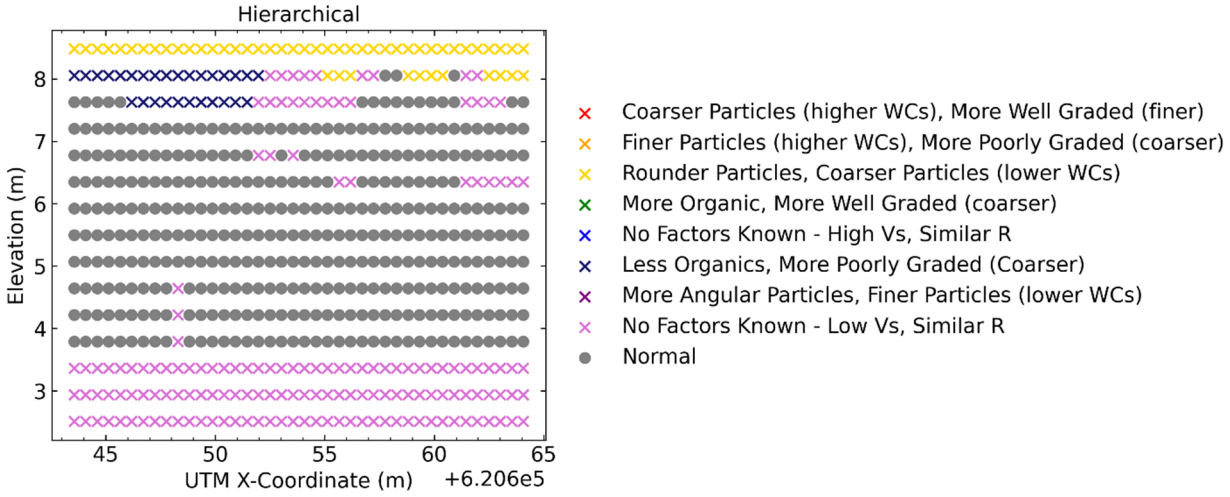
Profile: sc3_site1



Profile: sc5_site1

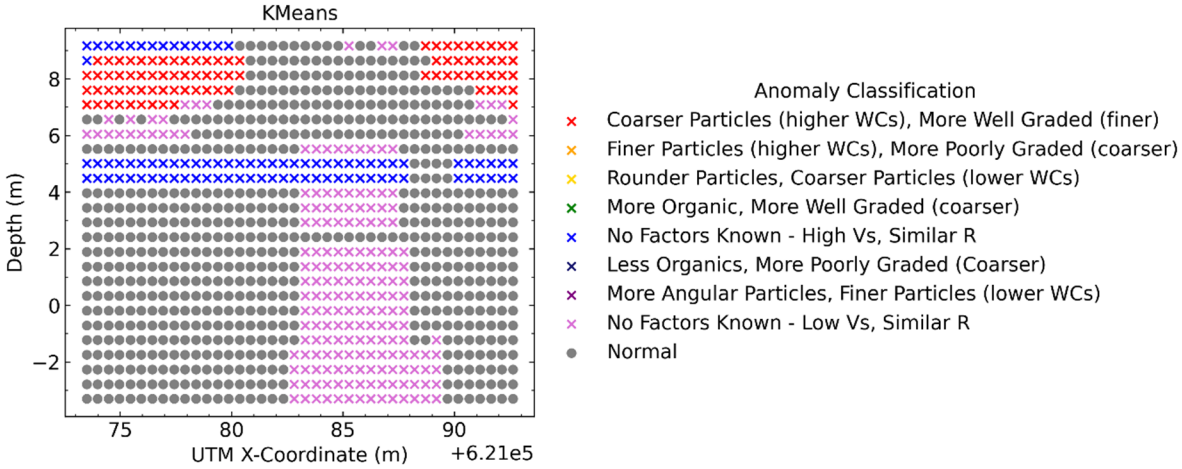


Profile: swt_site1

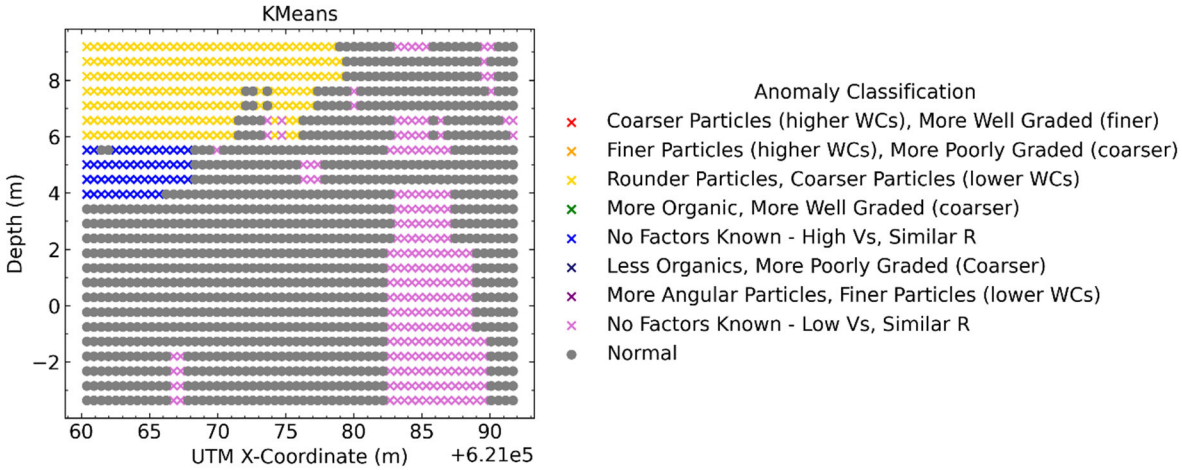


Fit: Site 1 Compilation
Algorithm: K-Means

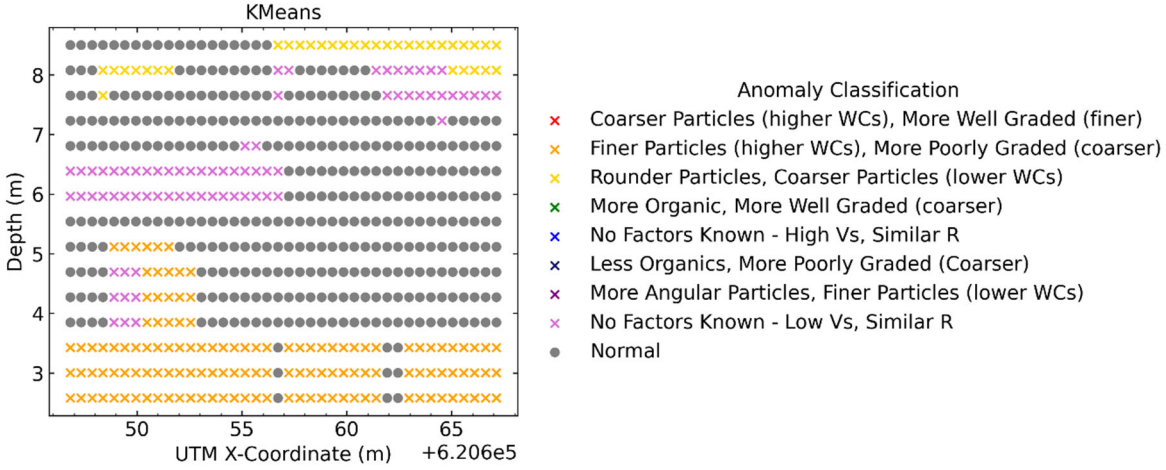
Profile: fc3_site1



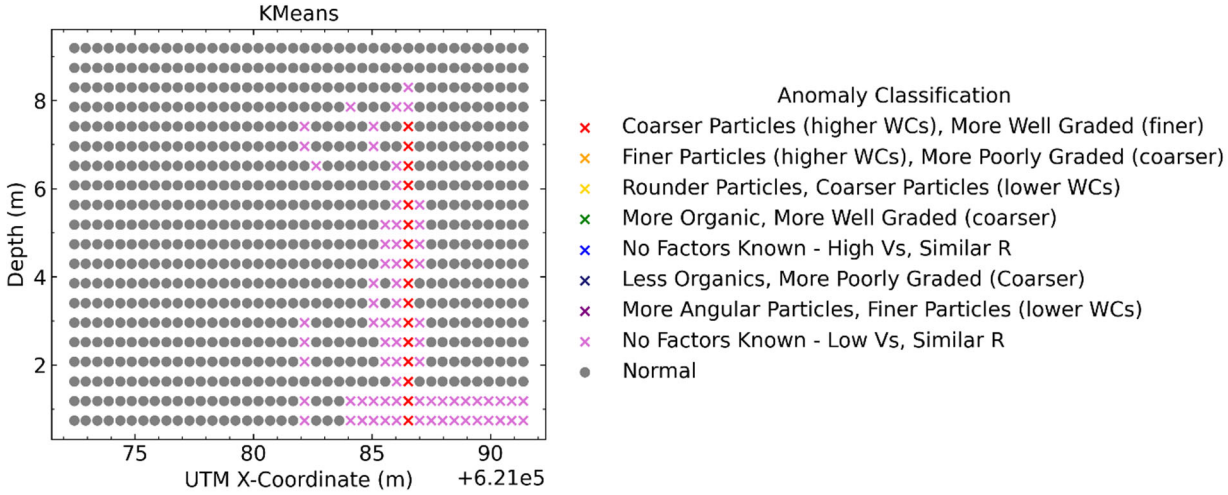
Profile: fc5_site1



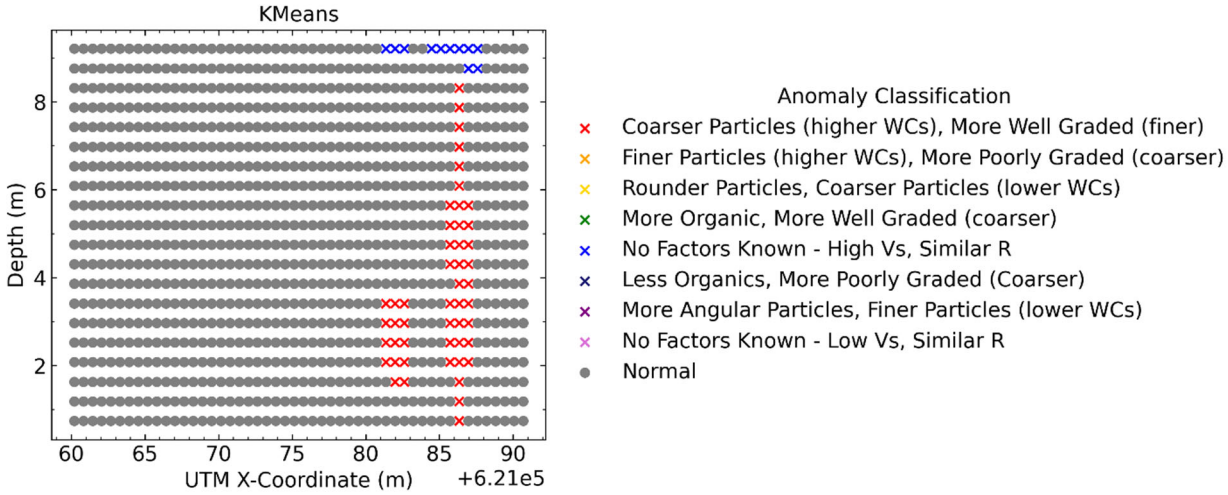
Profile: fwt_site1



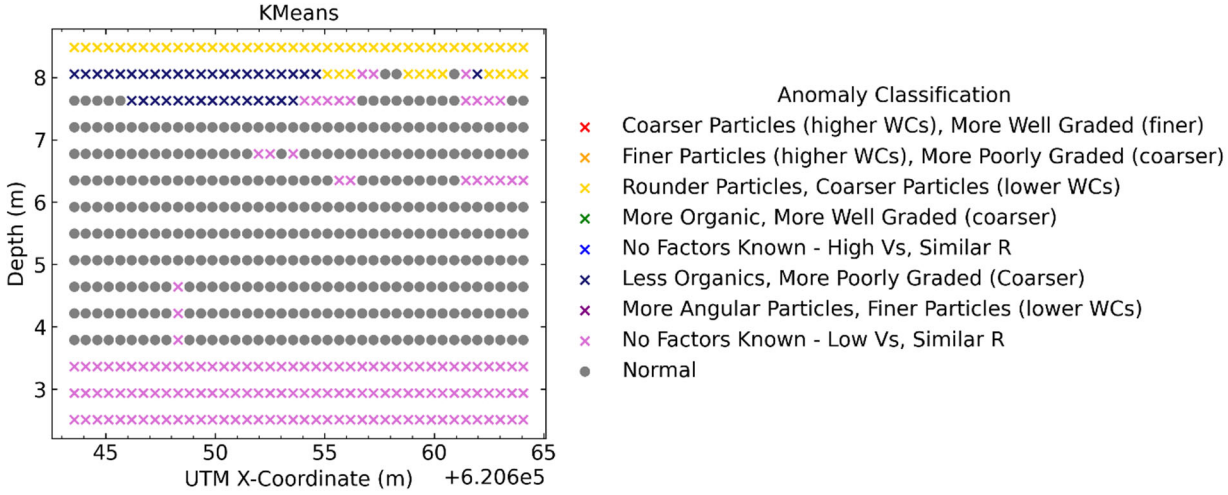
Profile: sc3_site1



Profile: sc5_site1

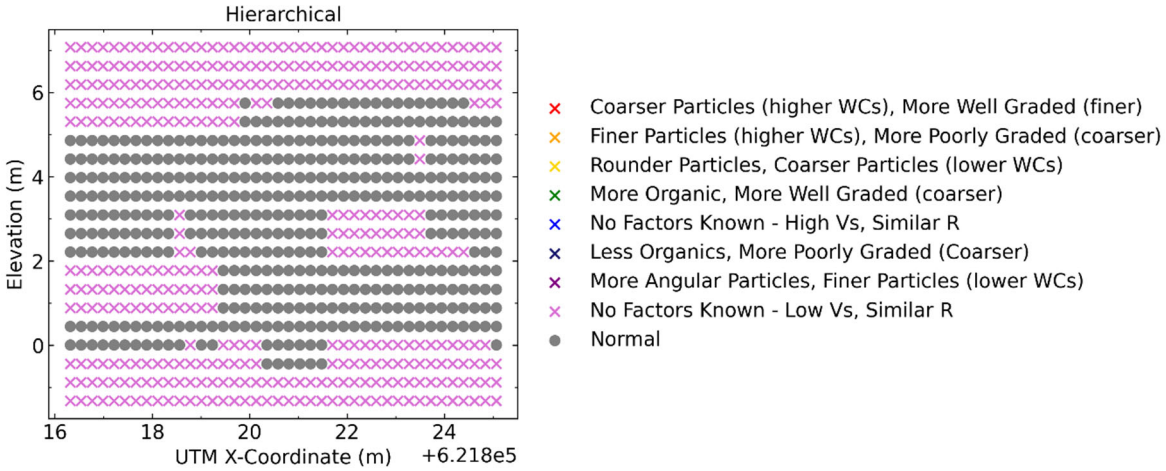


Profile: swt_site1

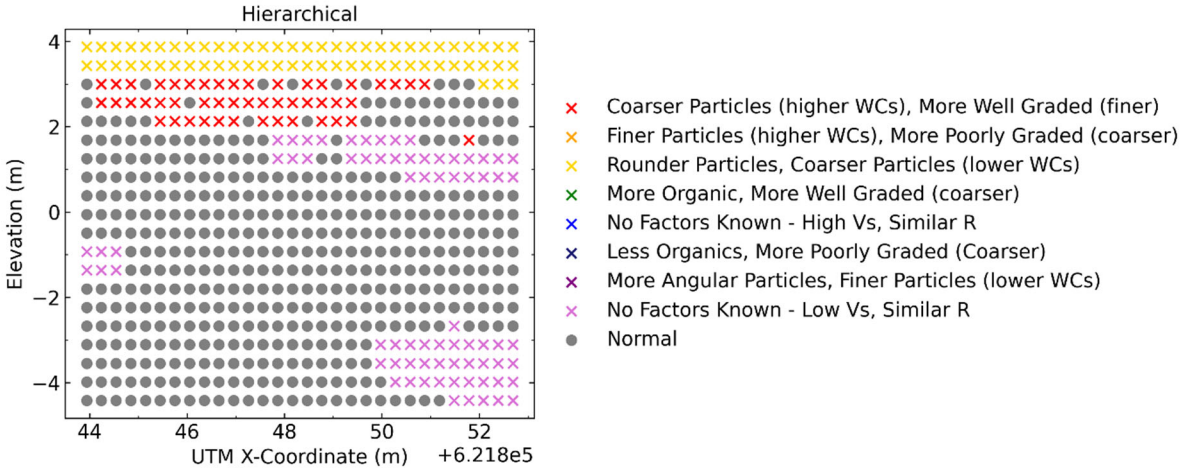


Fit: Site 2 Compilation
Algorithm: Hierarchical

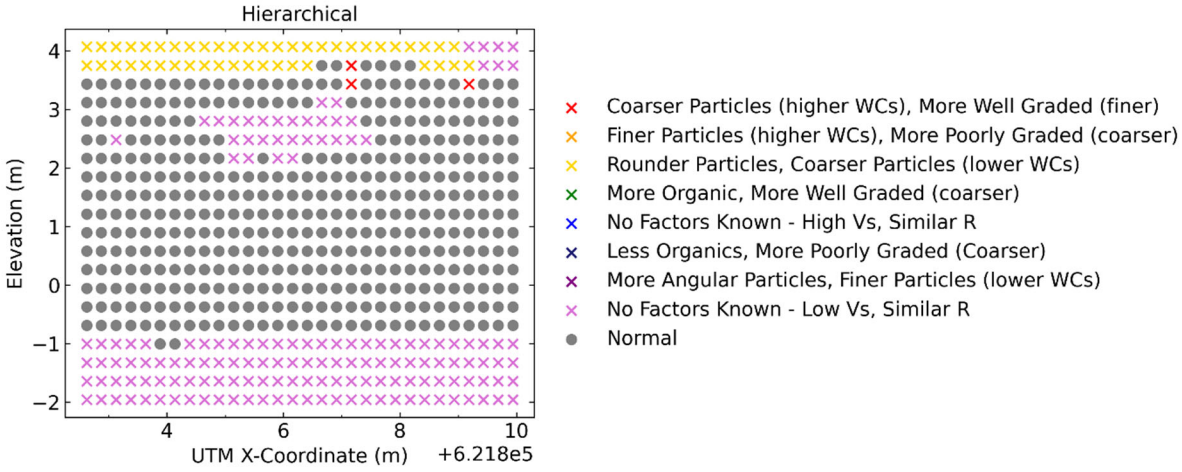
Profile: fc_site2



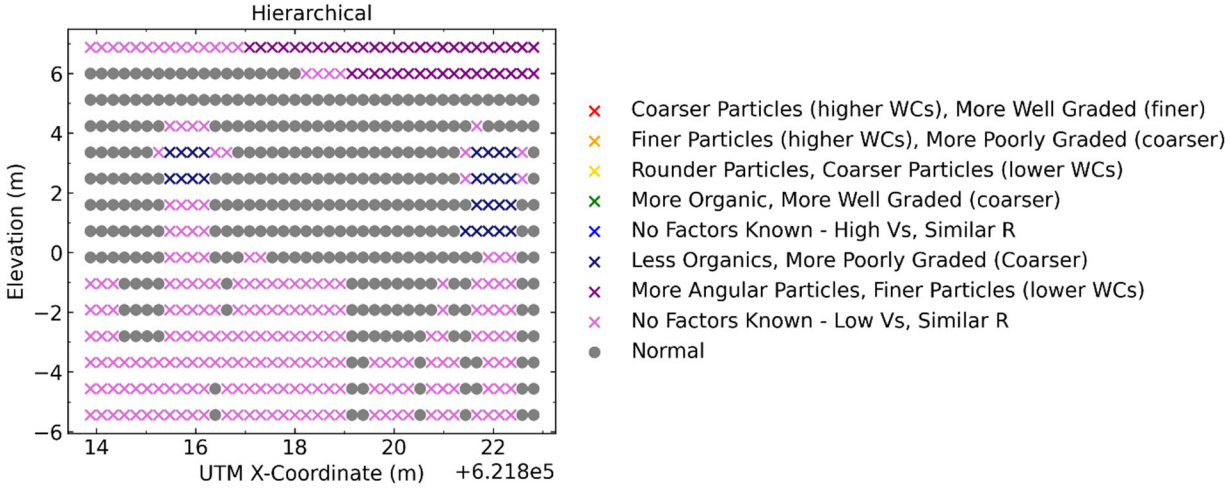
Profile: fdt_site2



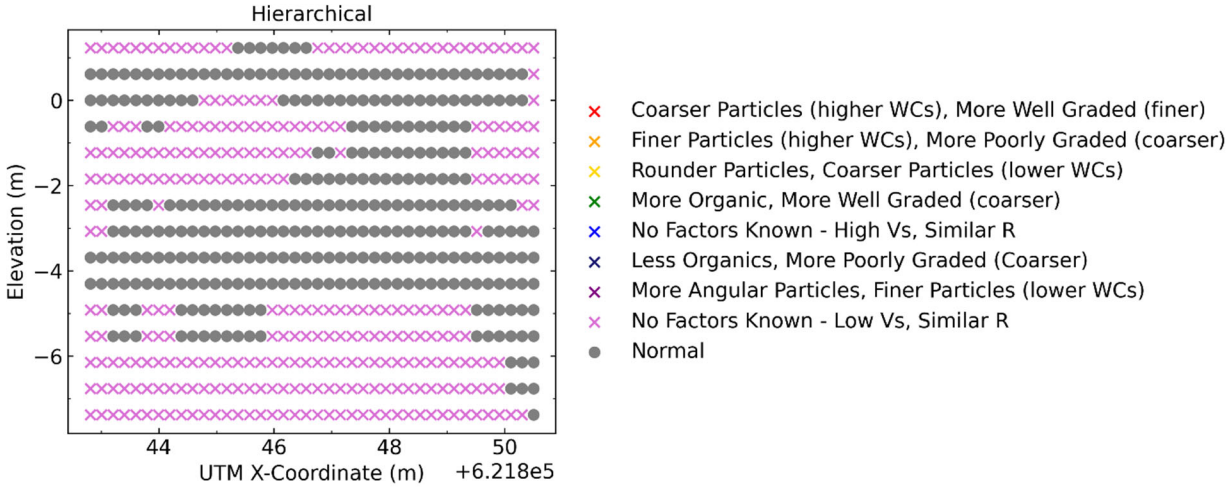
Profile: fwt_site2



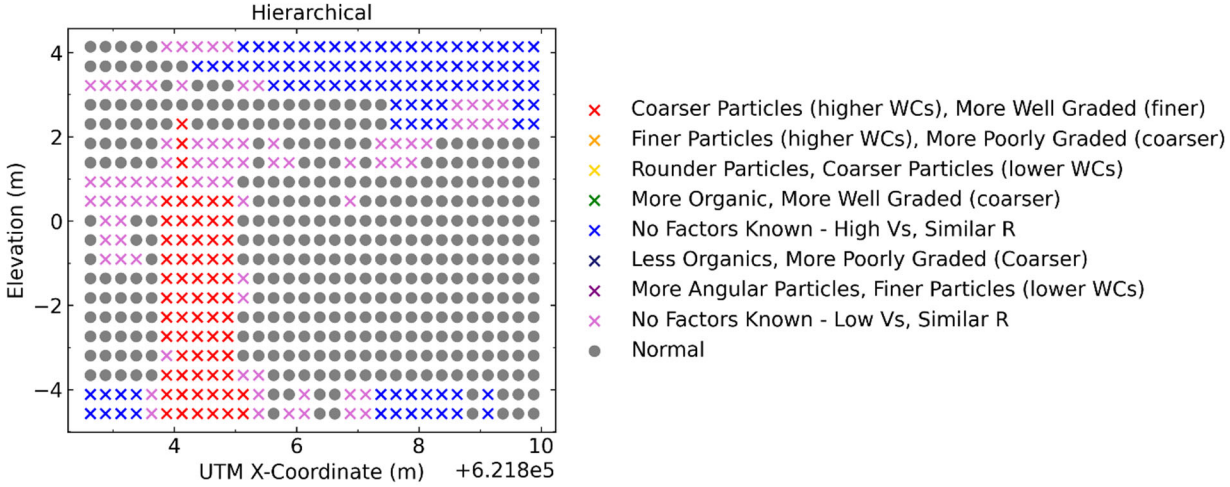
Profile: sc_site2



Profile: sdt_site2

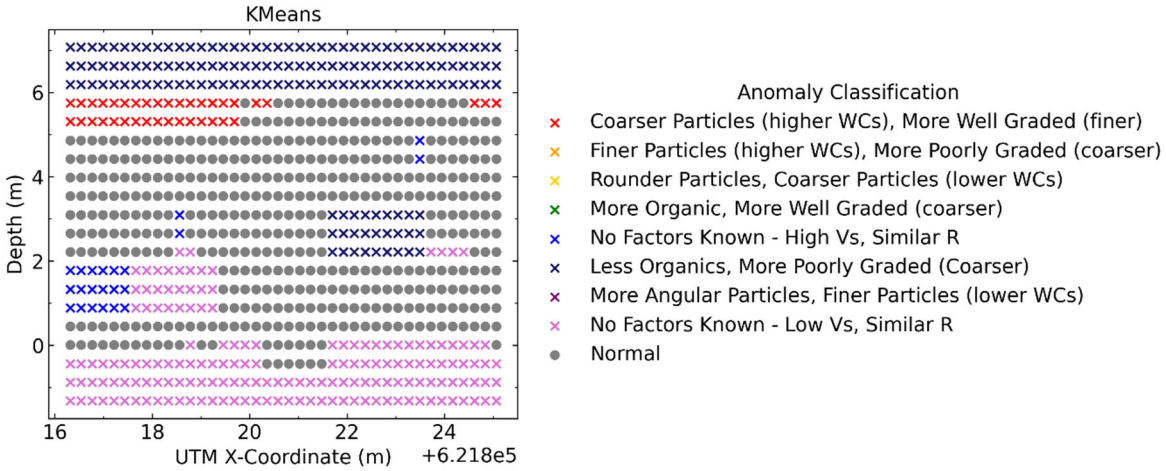


Profile: swt_site2

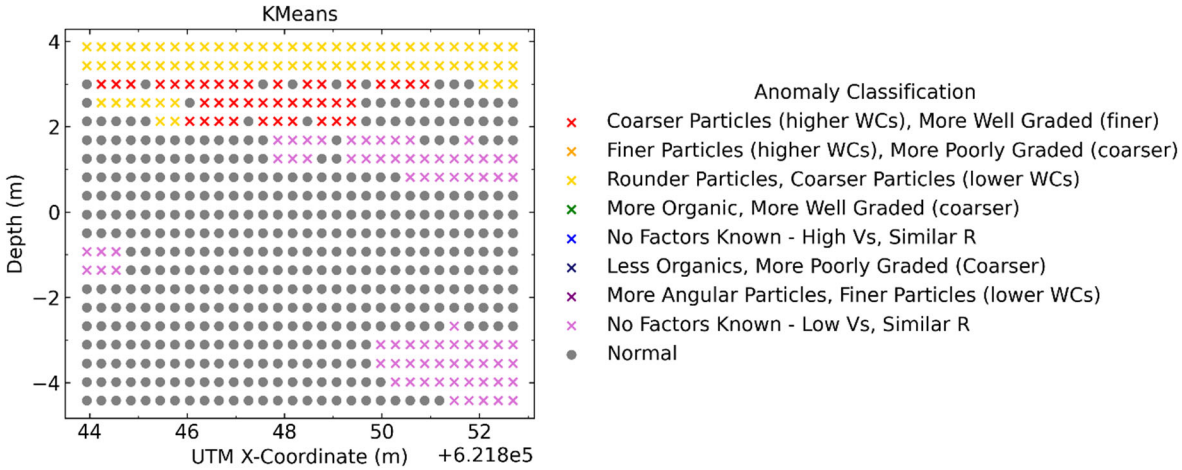


Fit: Site 2 Compilation
Algorithm: K-Means

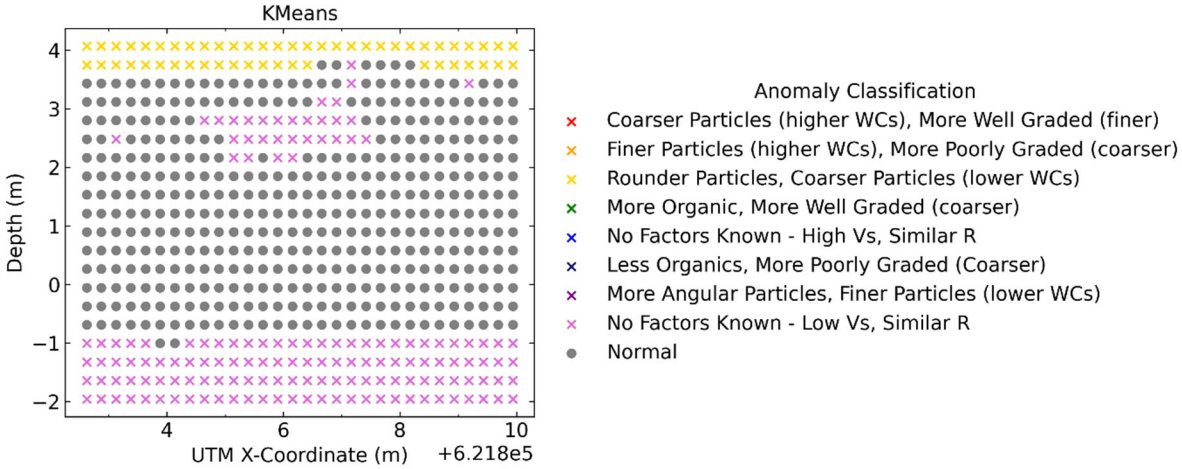
Profile: fc_site2



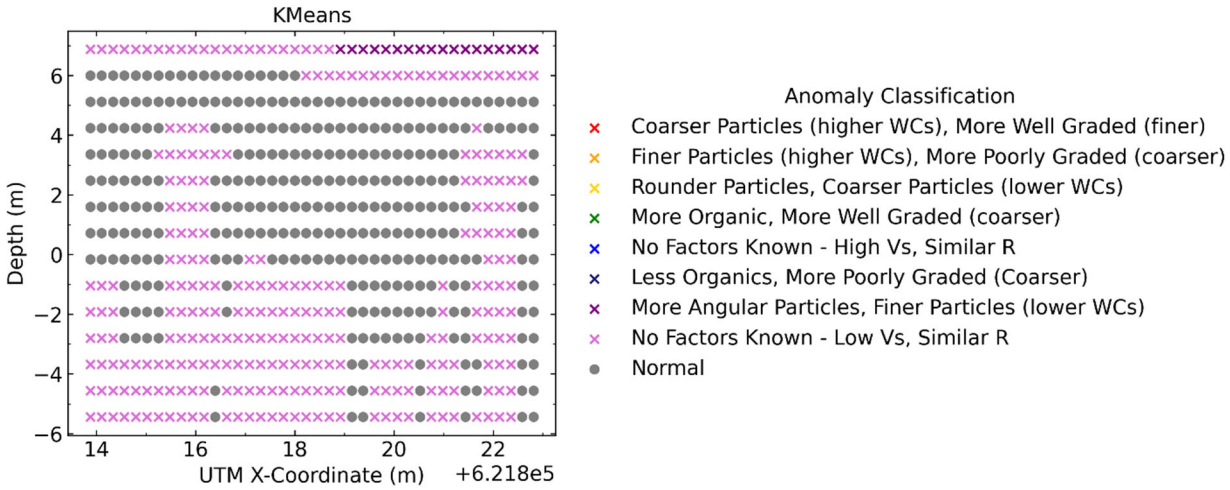
Profile: fdt_site2



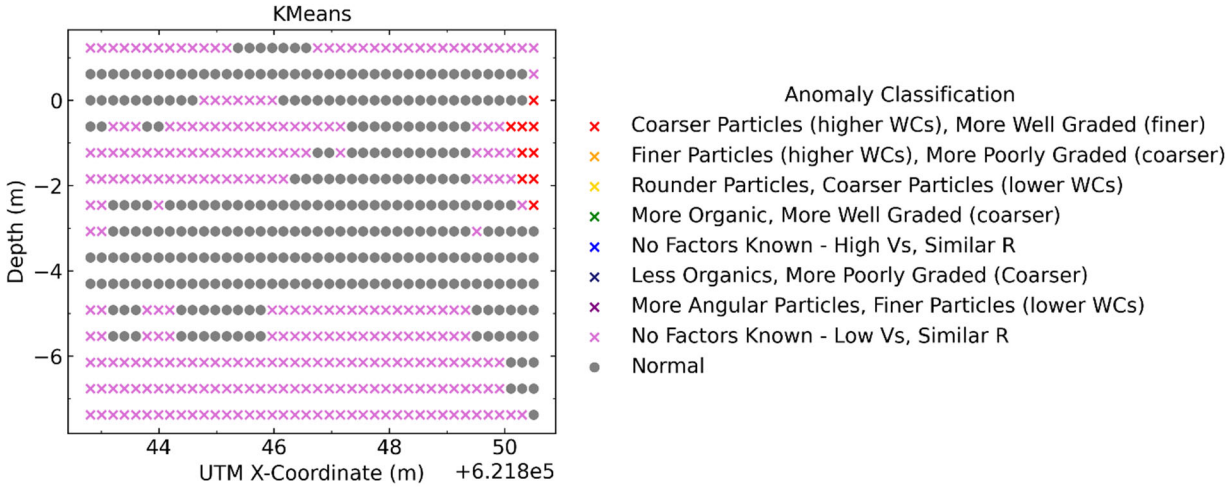
Profile: fwt_site2



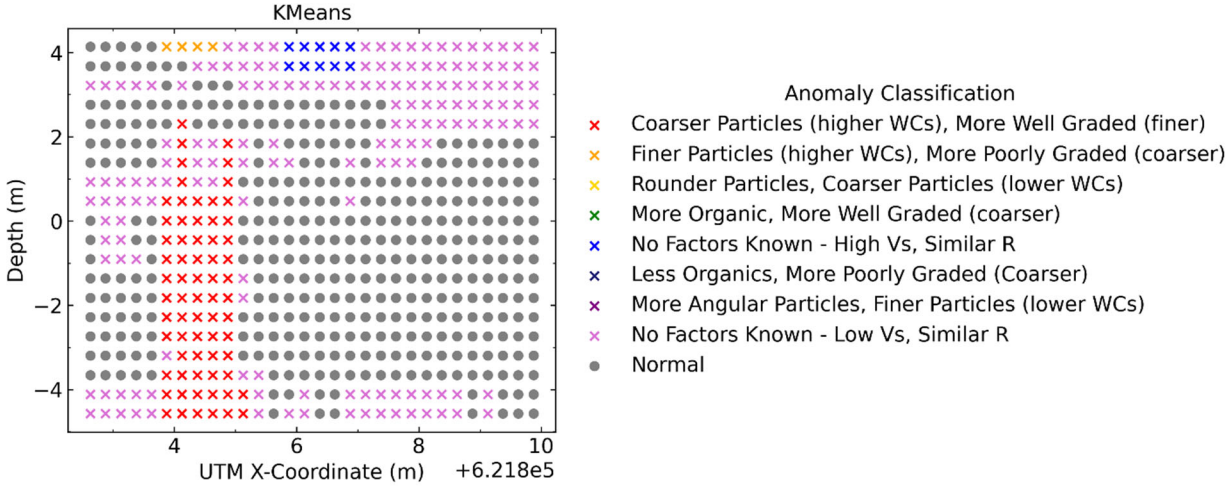
Profile: sc_site2



Profile: sdt_site2

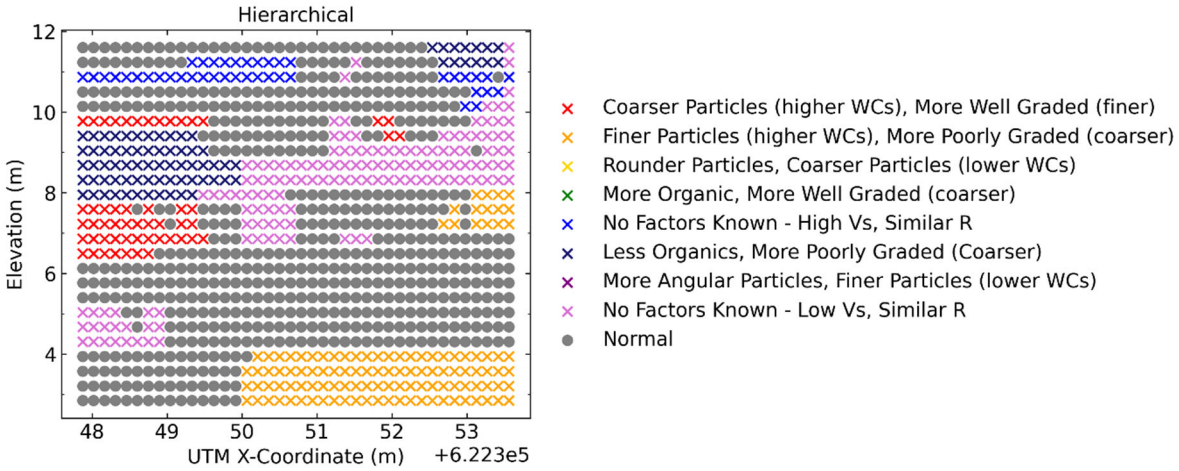


Profile: swt_site2

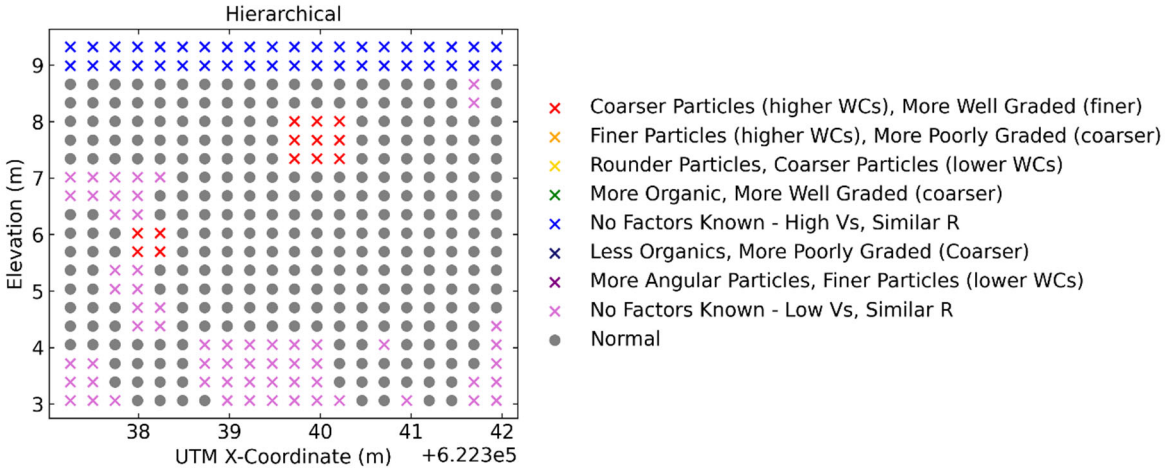


Fit: Site 3 Compilation
Algorithm: Hierarchical

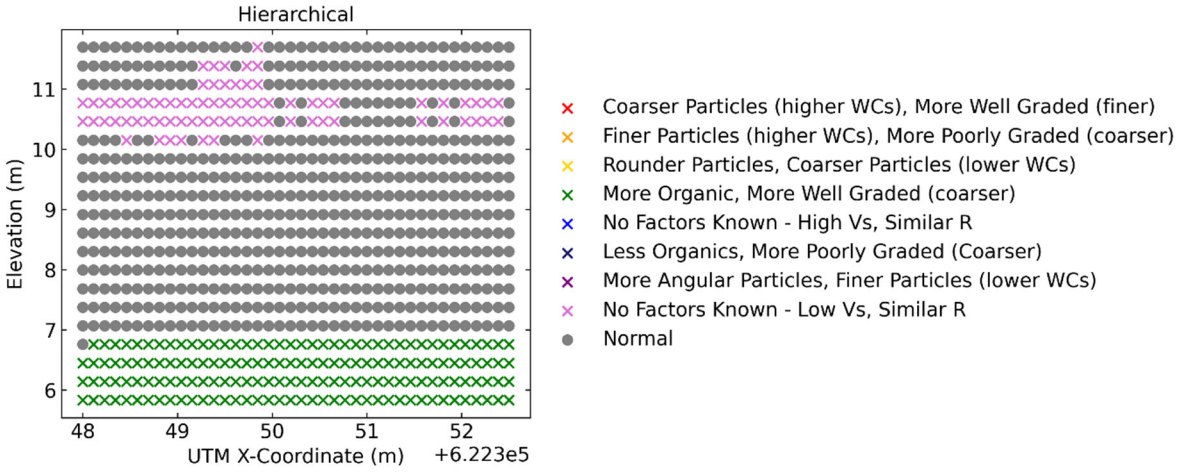
Profile: fc_site3



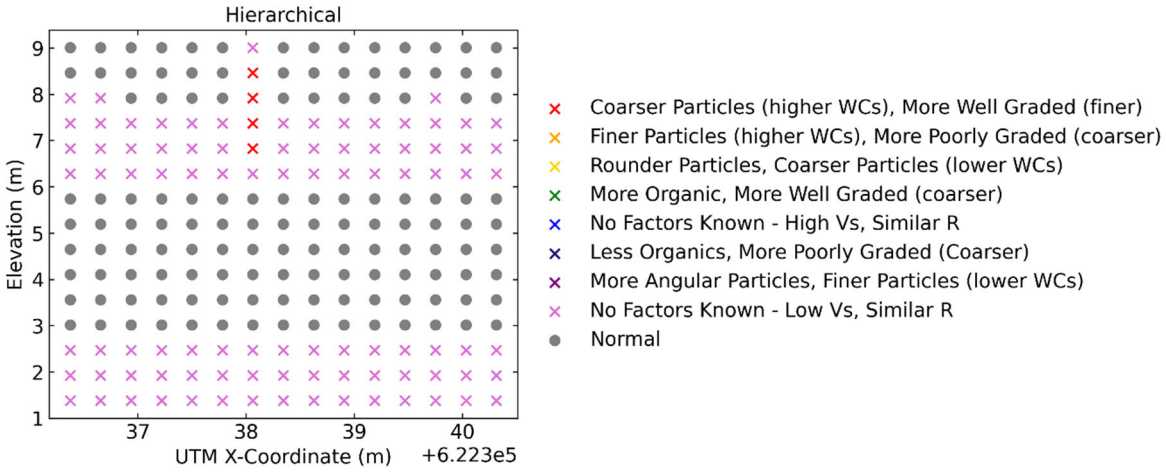
Profile: fwt_site3



Profile: sc_site3

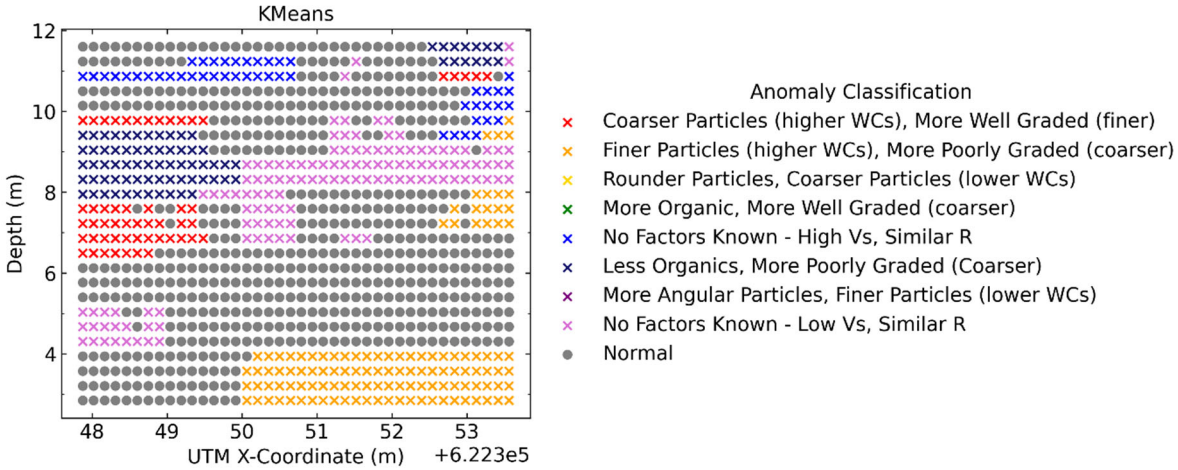


Profile: swt_site3

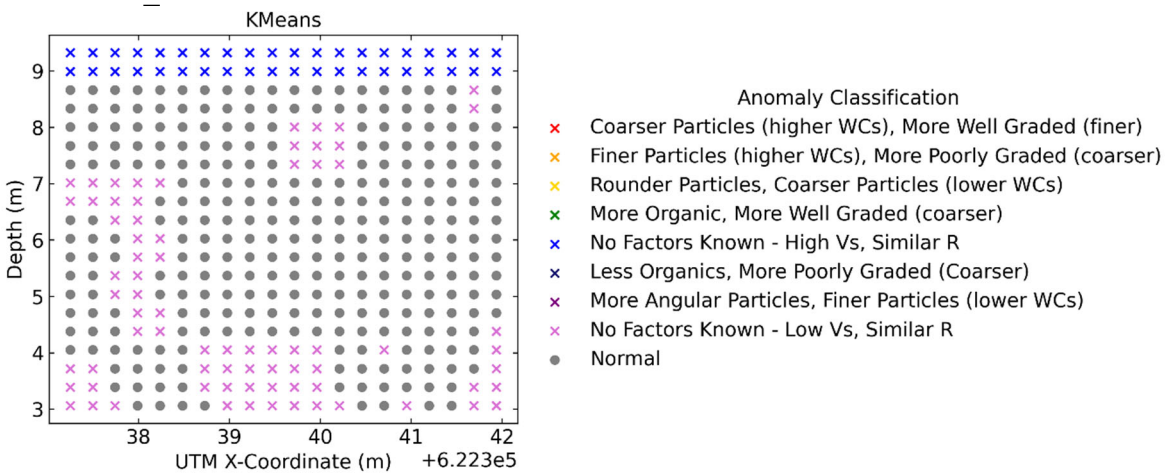


Fit: Site 3 Compilation
Algorithm: K-Means

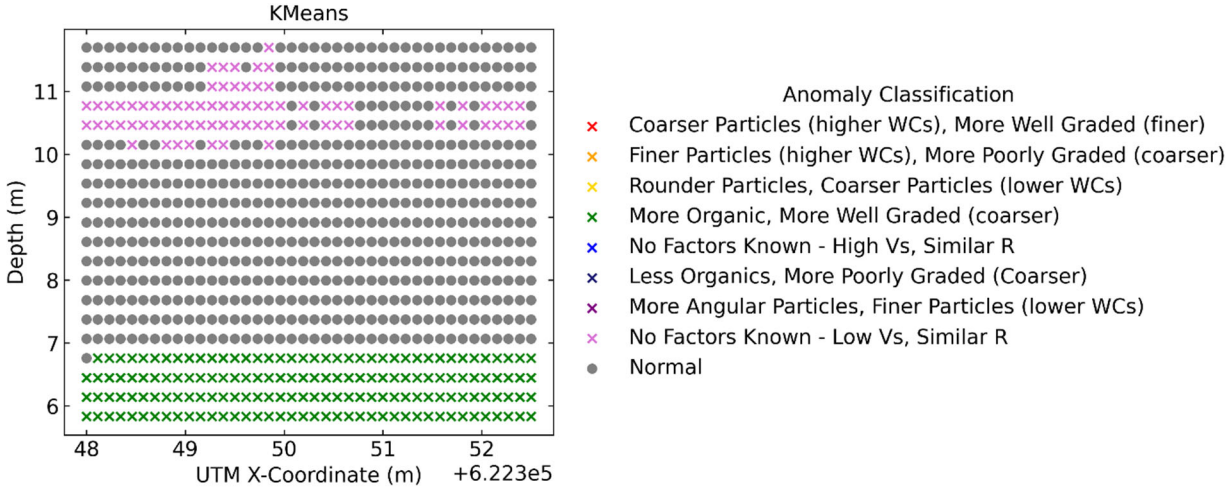
Profile: fc_site3



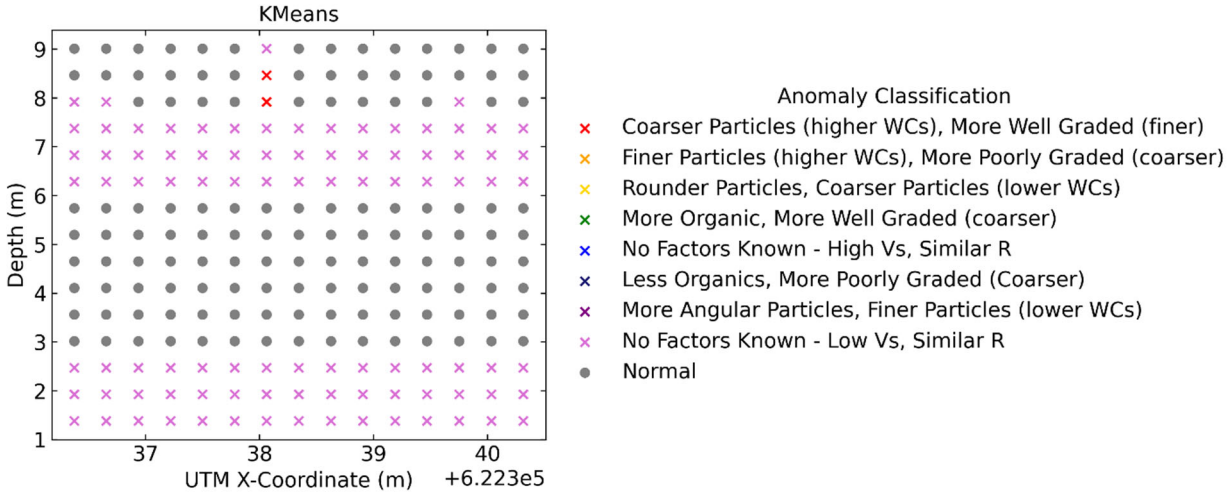
Profile: fwt_site3



Profile: sc_site3



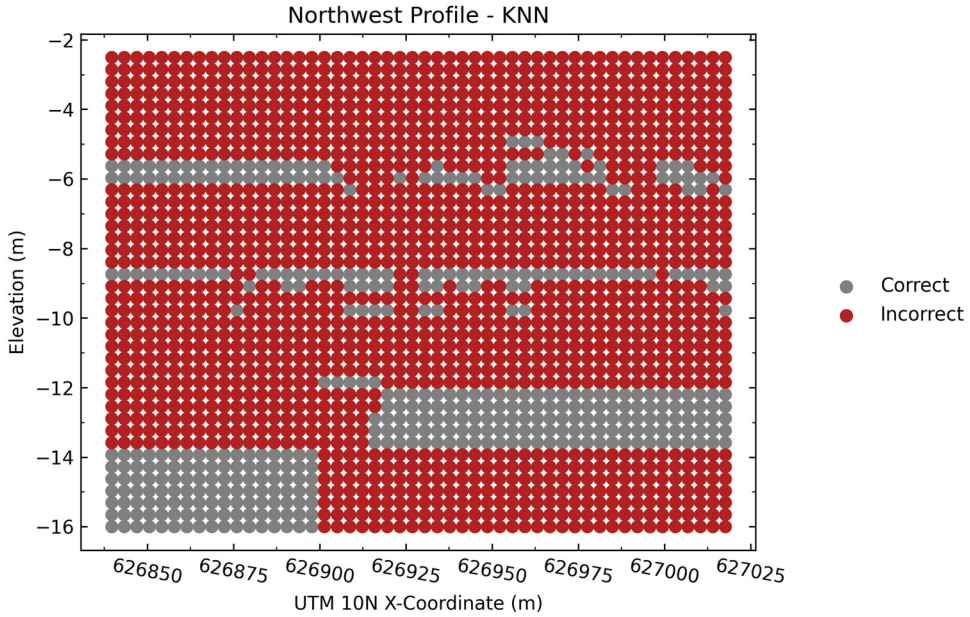
Profile: swt_site3



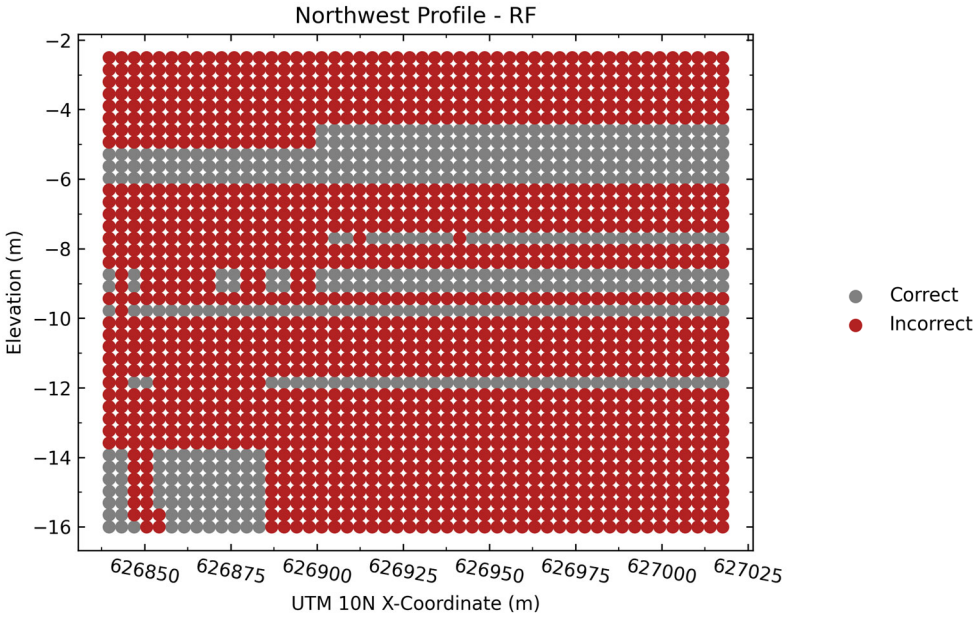
Appendix J – Supervised Machine Learning

Supervised Machine Learning Spatial Performance

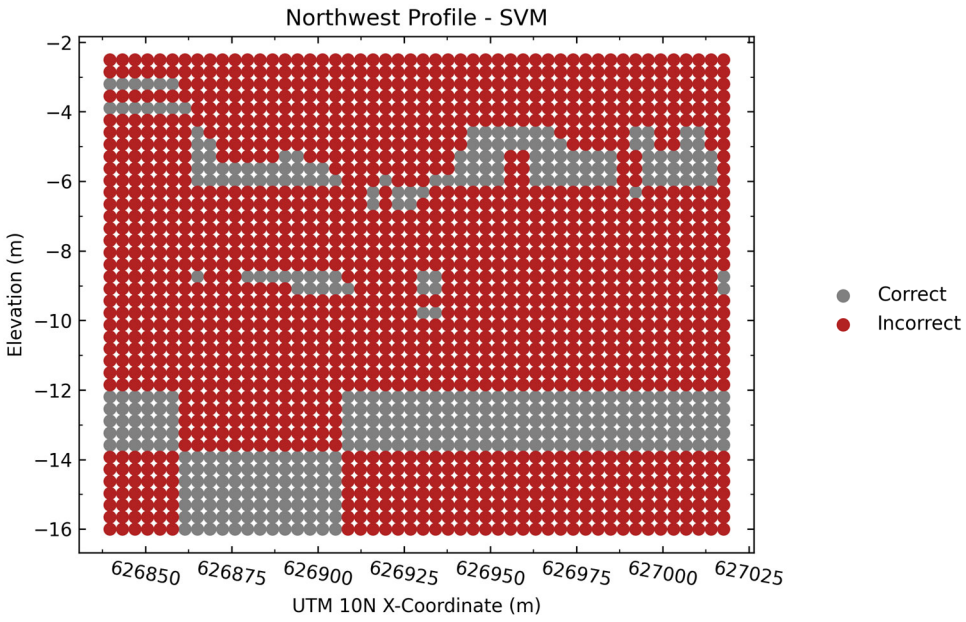
Bacon Island
Profile: Northwest
Algorithm: K-Nearest Neighbors



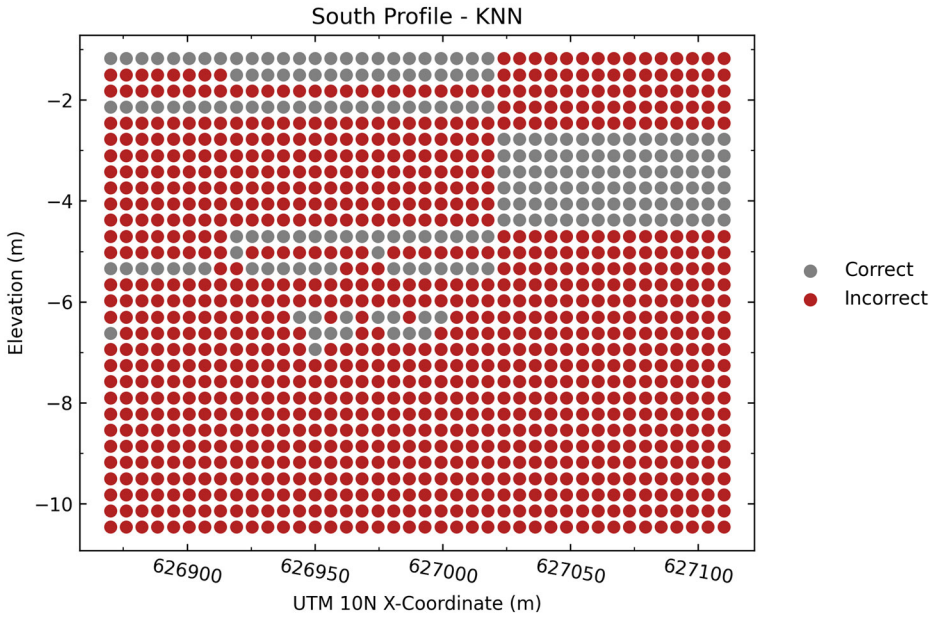
Bacon Island
Profile: Northwest
Algorithm: Random Forest



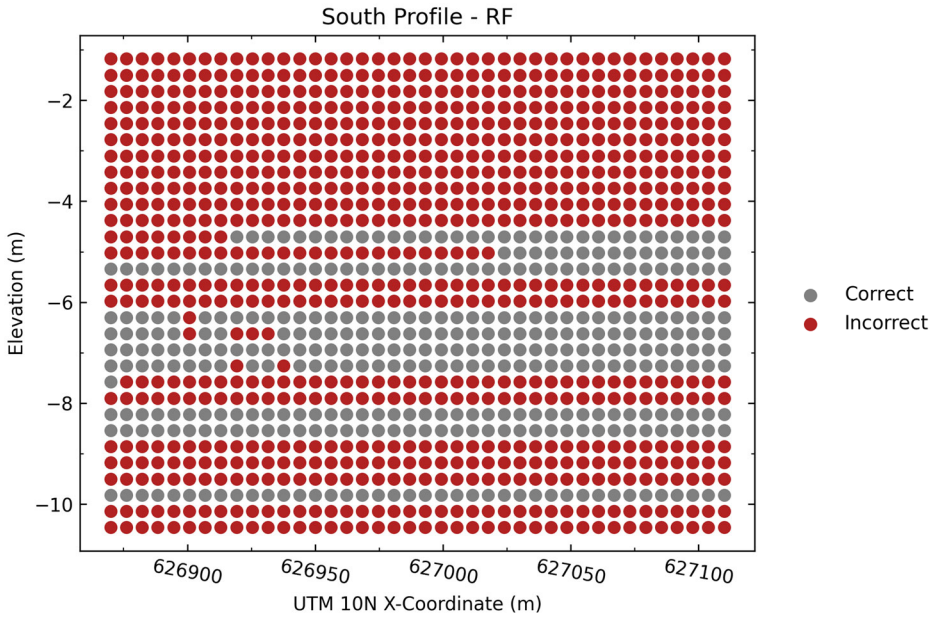
Bacon Island
Profile: Northwest
Algorithm: Support Vector Machines



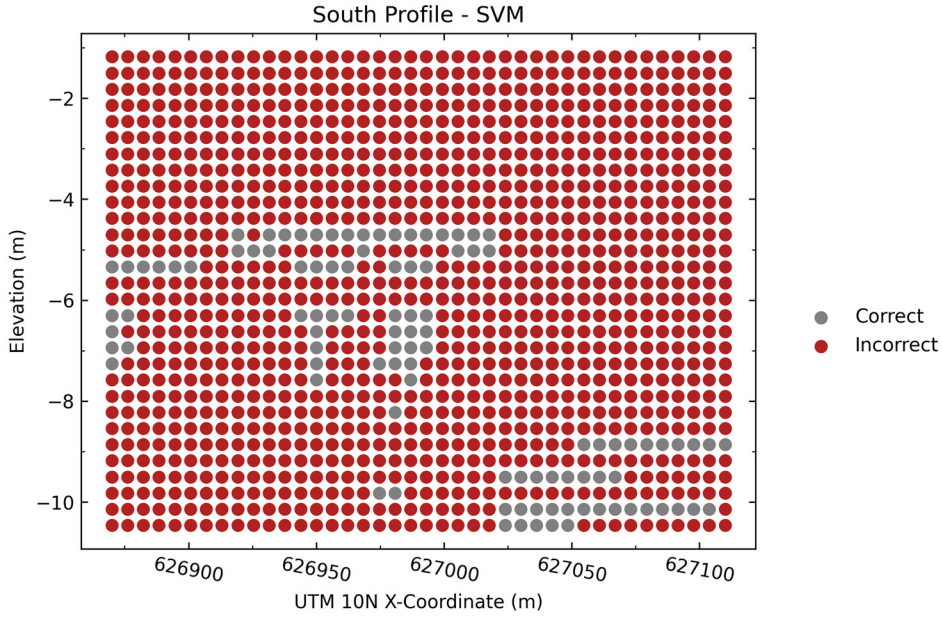
Bacon Island
Profile: South
Algorithm: K-Nearest Neighbors



Bacon Island
Profile: South
Algorithm: Random Forest

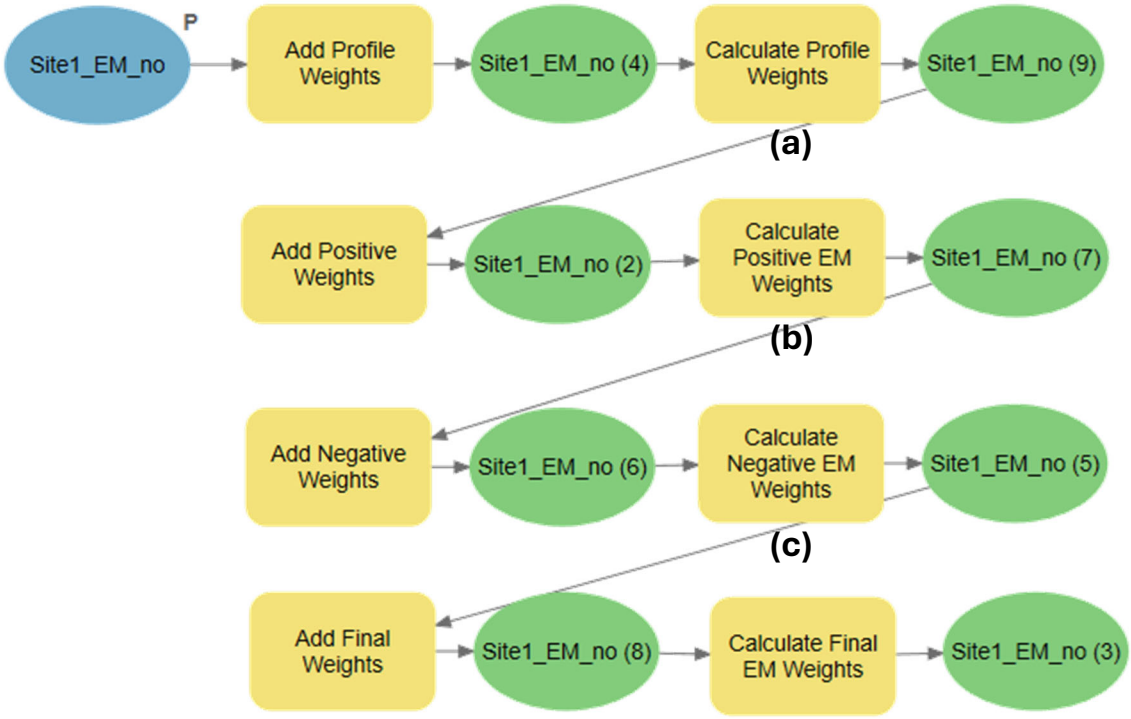


Bacon Island
Profile: South
Algorithm: Support Vector Machines



Appendix K – ModelBuilder Models

Assigning Weights to EMI Data



(a) – adding the profile weights to each point

```
def Reclass(arg):
    if arg is None:
        return 0
    elif (arg >= 0 and arg < 20):
        return -1
    elif (arg >= 20 and arg <
40):
        return -0.75
    elif (arg >= 40 and arg <
60):
        return -0.50
    elif (arg >= 60 and arg <
80):
        return -0.25
    elif (arg >= 80):
        return -0.05
    return arg
```

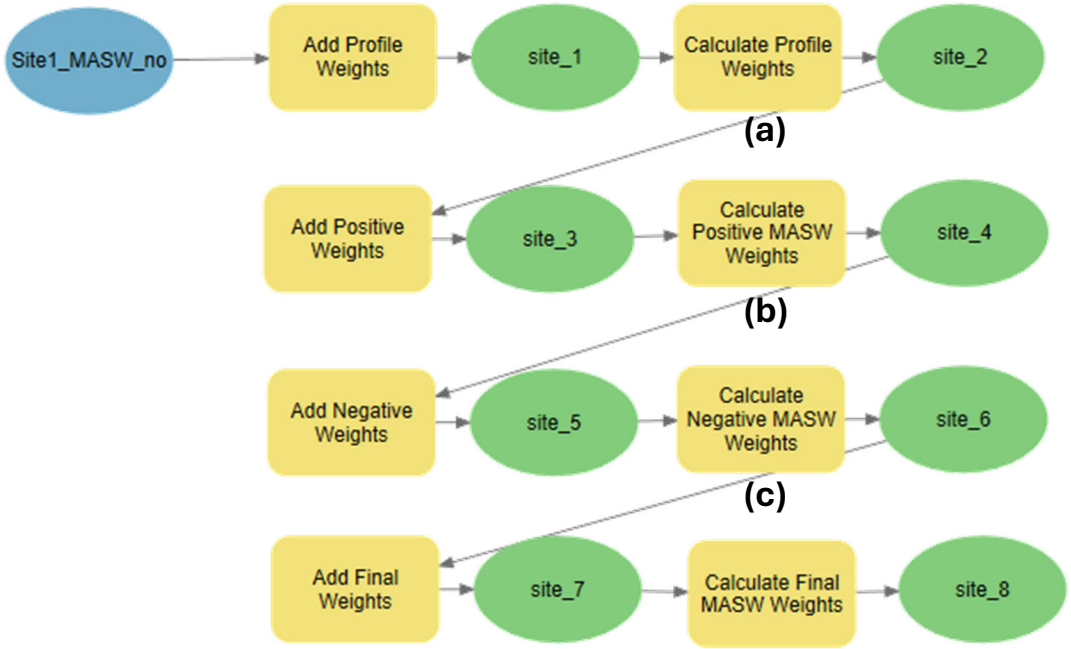
(b) – calculating the positive EMI weights for each point

```
def Reclass(arg):
    if arg is None:
        return 0
    elif (arg >= 10 and arg < 50):
        return 0.05
    elif (arg >= 50 and arg < 100):
        return 0.25
    elif (arg >= 100 and arg < 500):
        return 0.50
    elif (arg >= 500 and arg < 1000):
        return 0.75
    elif (arg >= 1000):
        return 1
    return arg
```

(c) – calculating the negative EMI weights for each point

```
def Reclass(arg):
    if arg is None:
        return 0
    elif (arg <= -10 and arg > -50):
        return 0.05
    elif (arg <= -50 and arg > -100):
        return 0.25
    elif (arg <= -100 and arg > -500):
        return 0.50
    elif (arg <= -500 and arg > -1000):
        return 0.75
    elif (arg <= -1000):
        return 1
    return arg
```

Assigning Weights to MASW Data



(a) – adding the profile weights to each point

```
def Reclass(arg):
    if arg is None:
        return 0
    elif (arg >= 0 and arg < 20):
        return -1
    elif (arg >= 20 and arg < 40):
        return -0.75
    elif (arg >= 40 and arg < 60):
        return -0.50
    elif (arg >= 60 and arg < 80):
        return -0.25
    elif (arg >= 80):
        return -0.05
    return arg
```

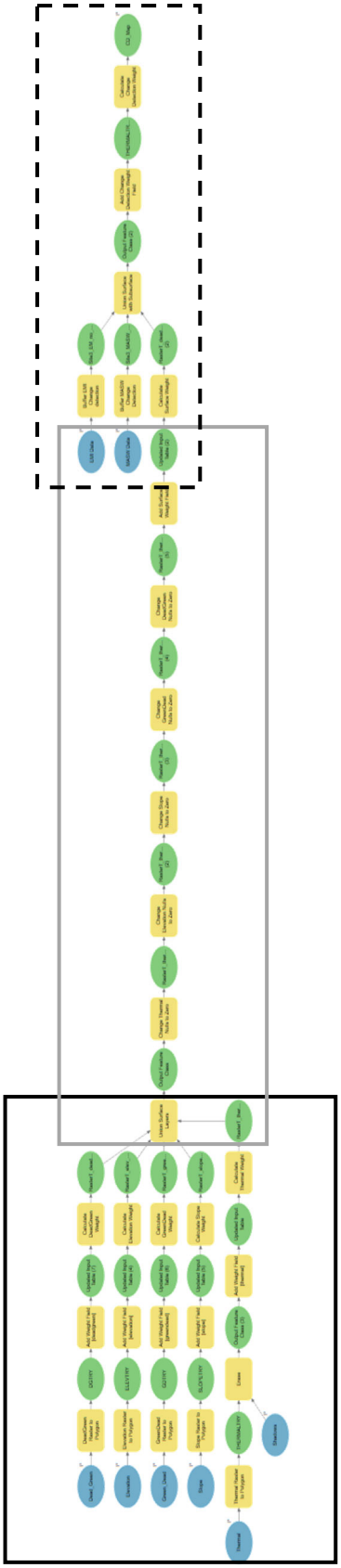
(b) – calculating the positive MASW weight for each point

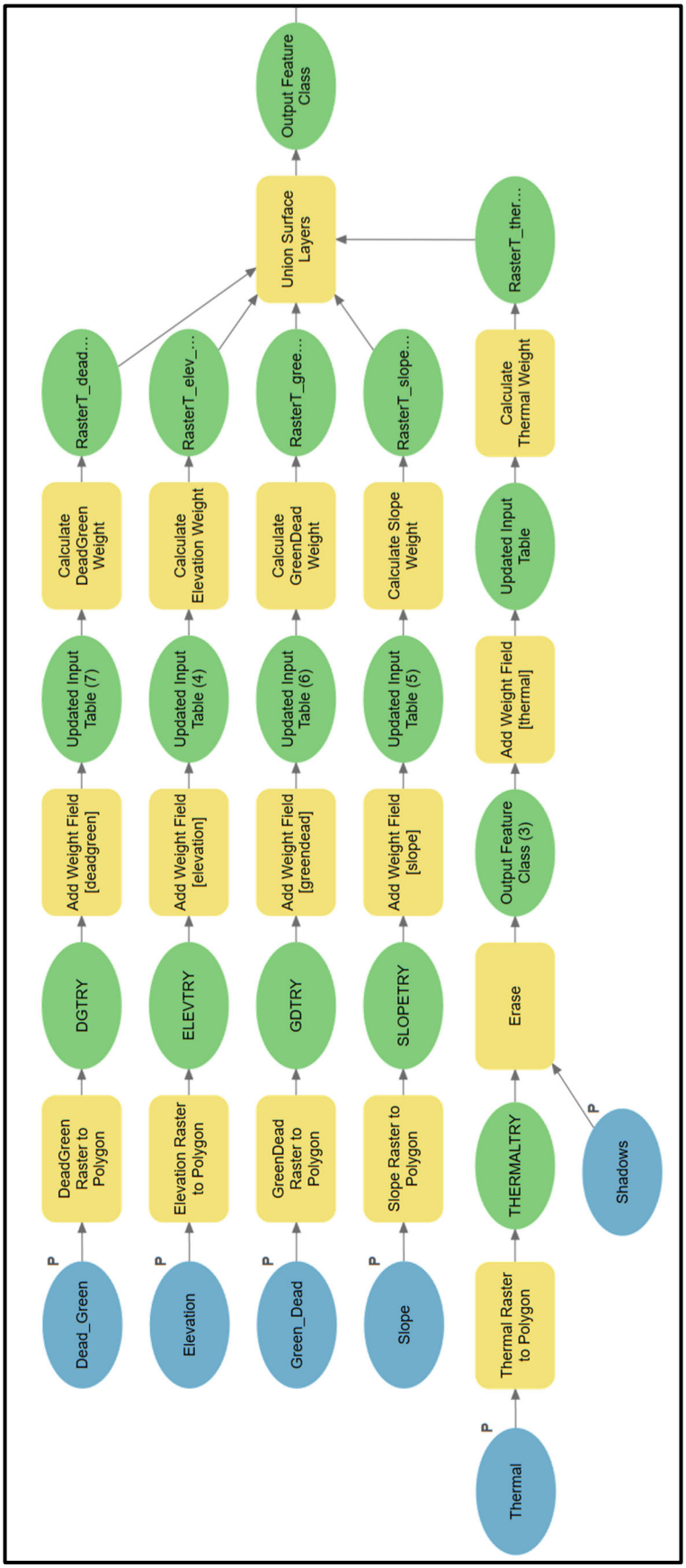
```
def Reclass(arg):
    if arg is None:
        return 0
    elif (arg >= 10 and arg < 30):
        return 0.05
    elif (arg >= 30 and arg < 50):
        return 0.25
    elif (arg >= 50 and arg < 70):
        return 0.50
    elif (arg >= 70 and arg < 90):
        return 0.75
    elif (arg >= 90):
        return 1
    return arg
```

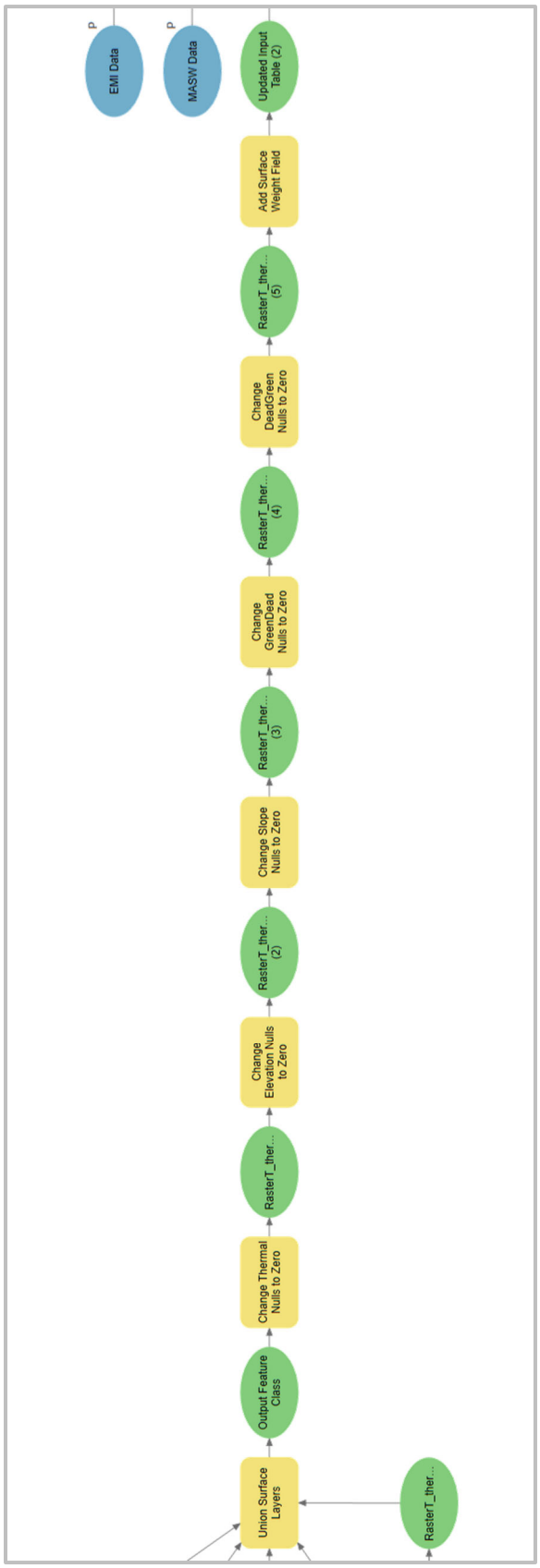
(c) – calculating the negative MASW weight for each point

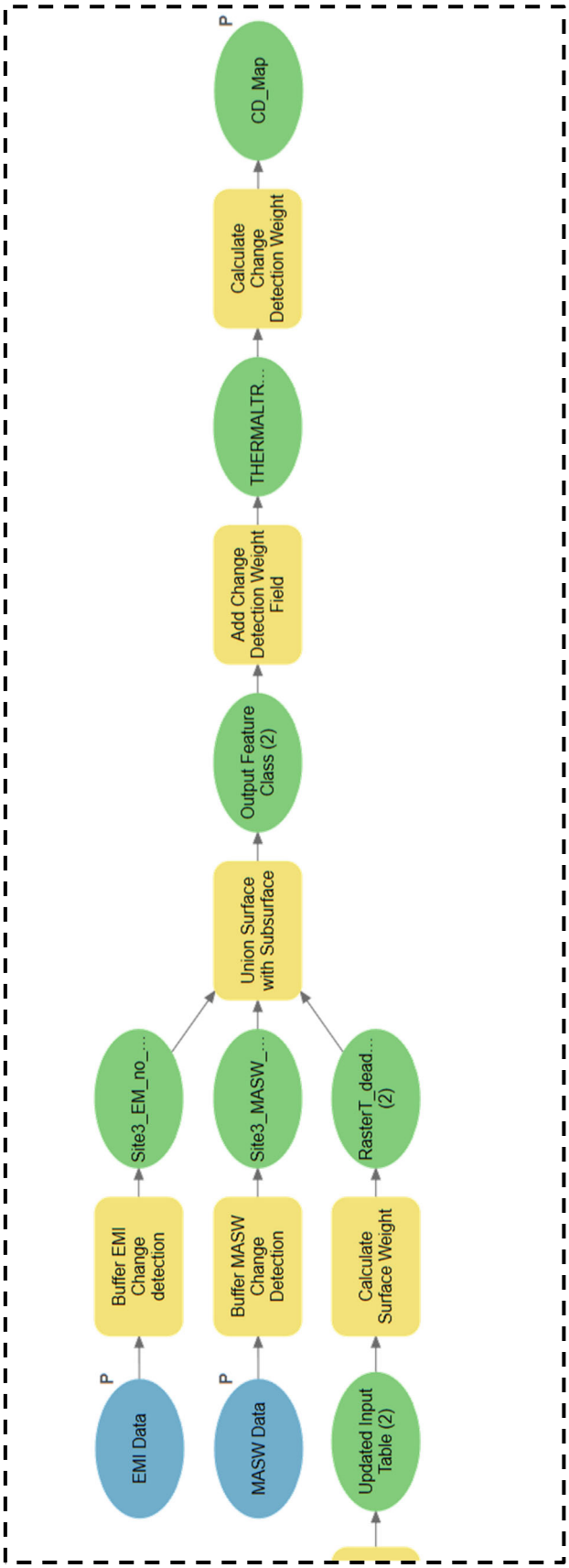
```
def Reclass(arg):
    if arg is None:
        return 0
    elif (arg <= -10 and arg > -30):
        return 0.05
    elif (arg <= -30 and arg > -50):
        return 0.25
    elif (arg <= -50 and arg > -70):
        return 0.50
    elif (arg <= -70 and arg > -90):
        return 0.75
    elif (arg <= -90):
        return 1
    return arg
```

Change Detection Model – Surface and Subsurface

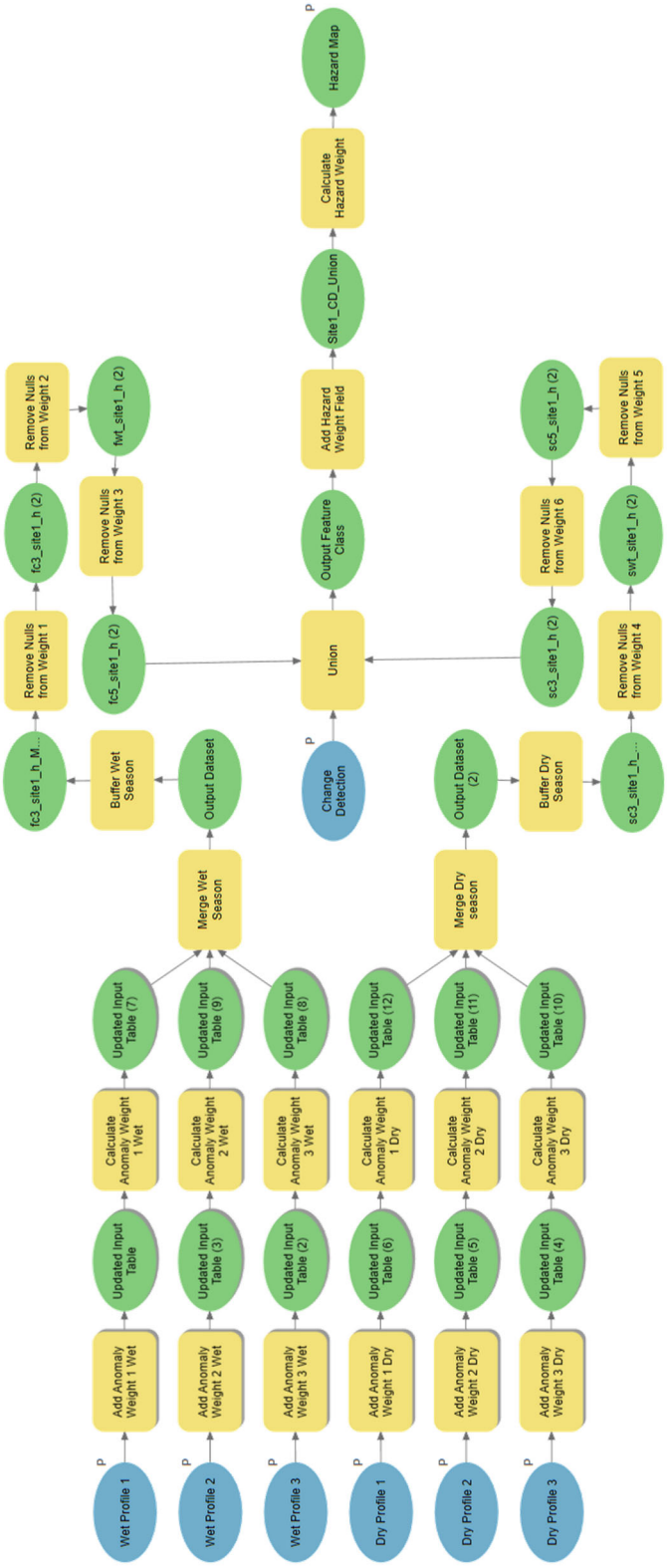



Dead_Green
DeadGreen Raster to Polygon
DGTRY
Add Weight Field [deadgreen]
Updated Input Table (7)
Calculate DeadGreen Weight
RasterT_dead...
Elevation
Elevation Raster to Polygon
ELEVTRY
Add Weight Field [elevation]
Updated Input Table (4)
Calculate Elevation Weight
RasterT_elev_...
Green_Dead
GreenDead Raster to Polygon
GDTRY
Add Weight Field [greendead]
Updated Input Table (6)
Calculate GreenDead Weight
RasterT_gree...
Slope
Slope Raster to Polygon
SLOPETRY
Add Weight Field [slope]
Updated Input Table (5)
Calculate Slope Weight
RasterT_slope...
Union Surface Layers
Output Feature Class
Thermal
Thermal Raster to Polygon
THERMALTRY
Erase
Output Feature Class (3)
Add Weight Field [thermal]
Updated Input Table
Calculate Thermal Weight
RasterT_ther...
Shadows
P

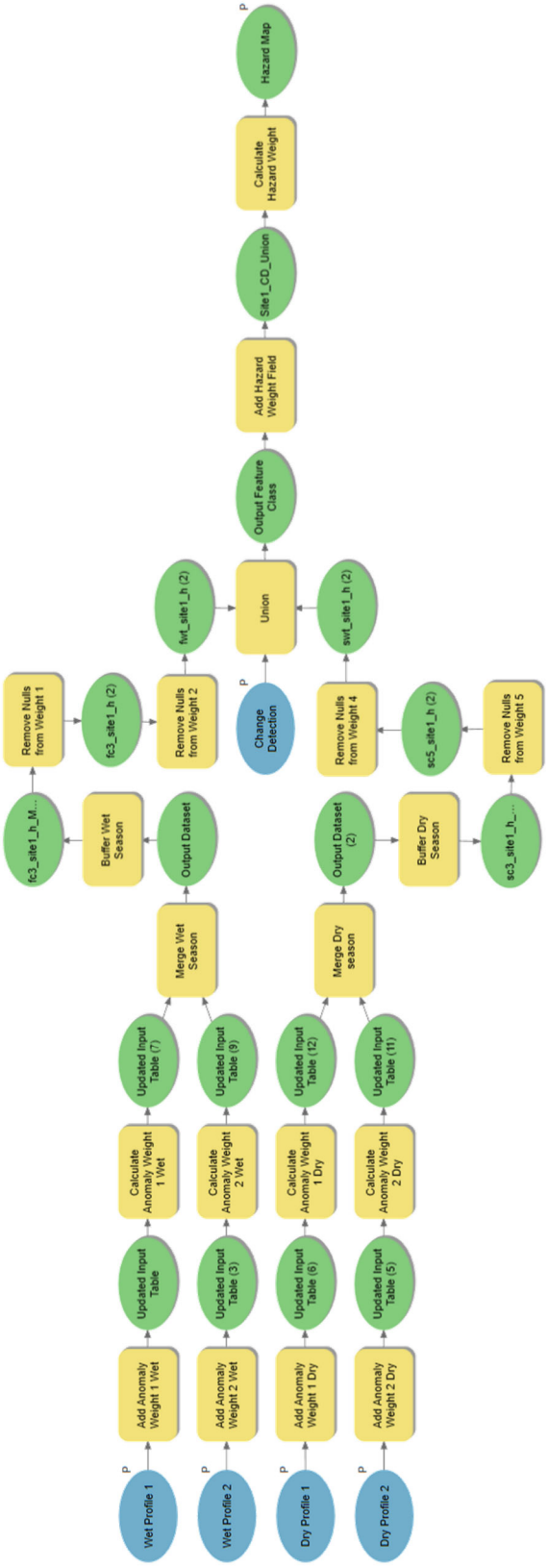

P
EMI Data
P
MASW Data
Updated Input Table (2)
Add Surface Weight Field
RasterT_ther... (5)
Change DeadGreen Nulls to Zero
RasterT_ther... (4)
Change GreenDead Nulls to Zero
RasterT_ther... (3)
Change Slope Nulls to Zero
RasterT_ther... (2)
Change Elevation Nulls to Zero
RasterT_ther...
Change Thermal Nulls to Zero
Output Feature Class
Union Surface Layers
RasterT_ther...


P
EMI Data
Buffer EMI Change detection
Site3_EM_no_...
P
MASW Data
Buffer MASW Change Detection
Site3_MASW_...
Updated Input Table (2)
Calculate Surface Weight
RasterT_dead... (2)
Union Surface with Subsurface
Output Feature Class (2)
Add Change Detection Weight Field
THERMALTR...
Calculate Change Detection Weight
P
CD_Map

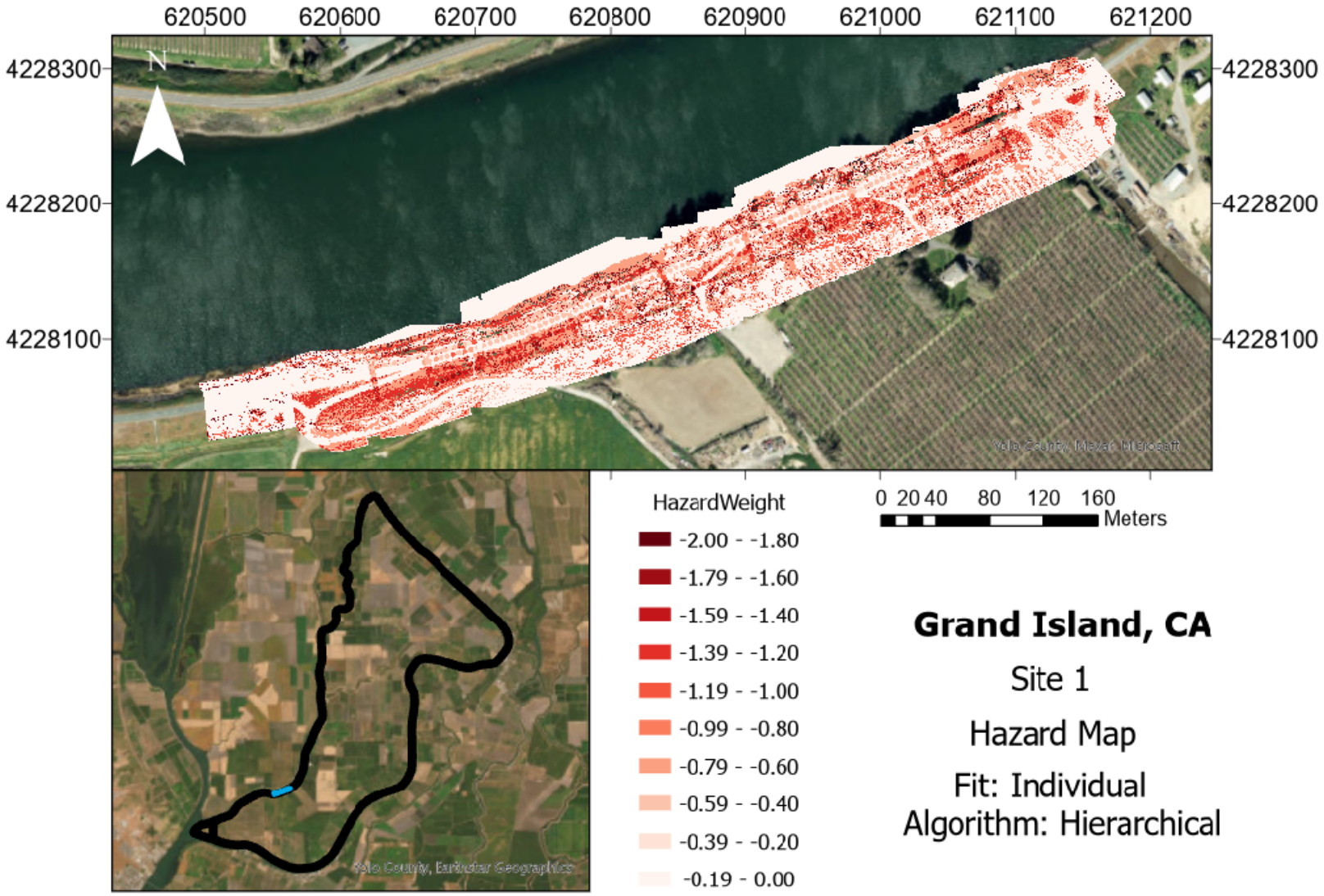
Final Hazard Model for 3 MASW Profiles at a Site

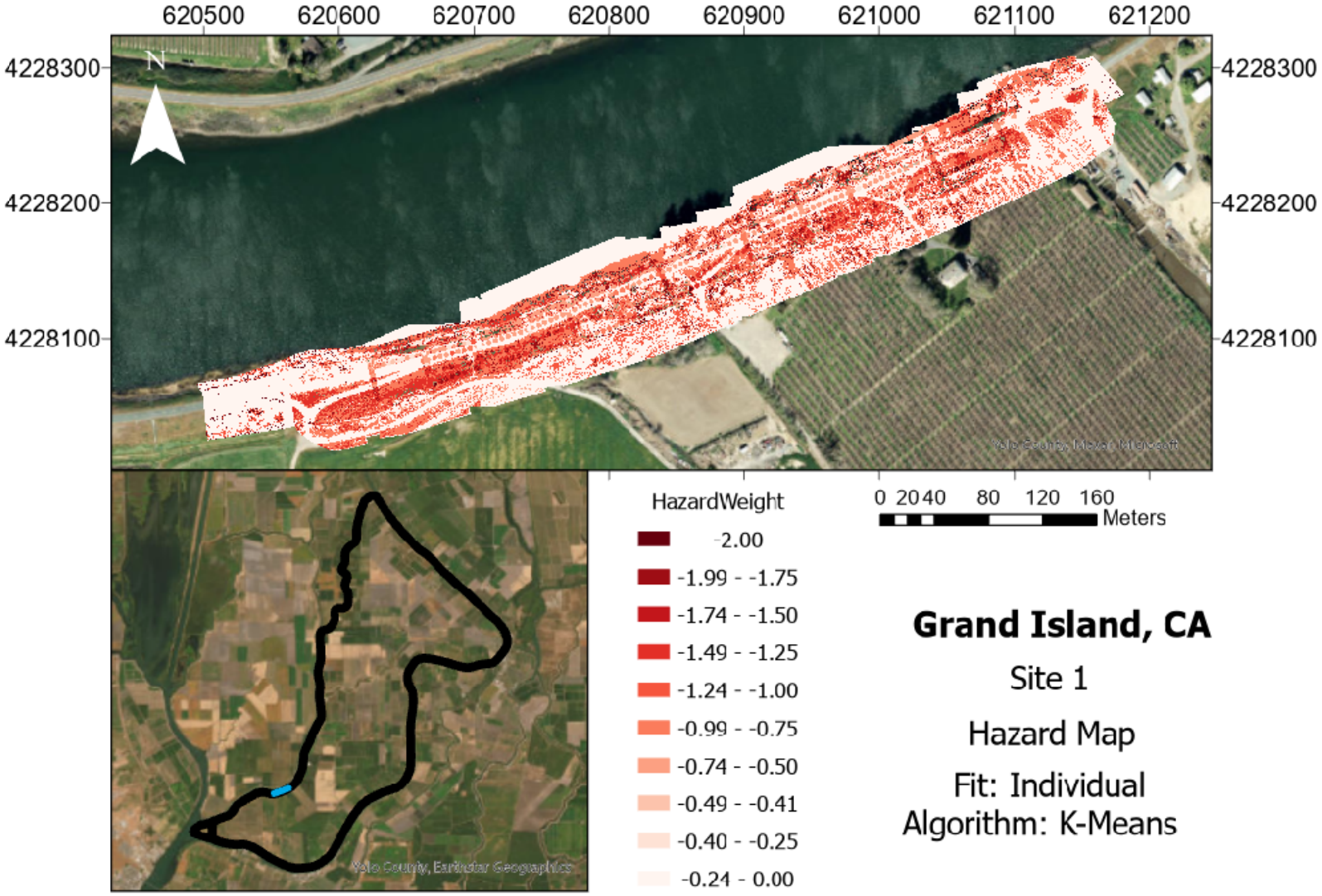


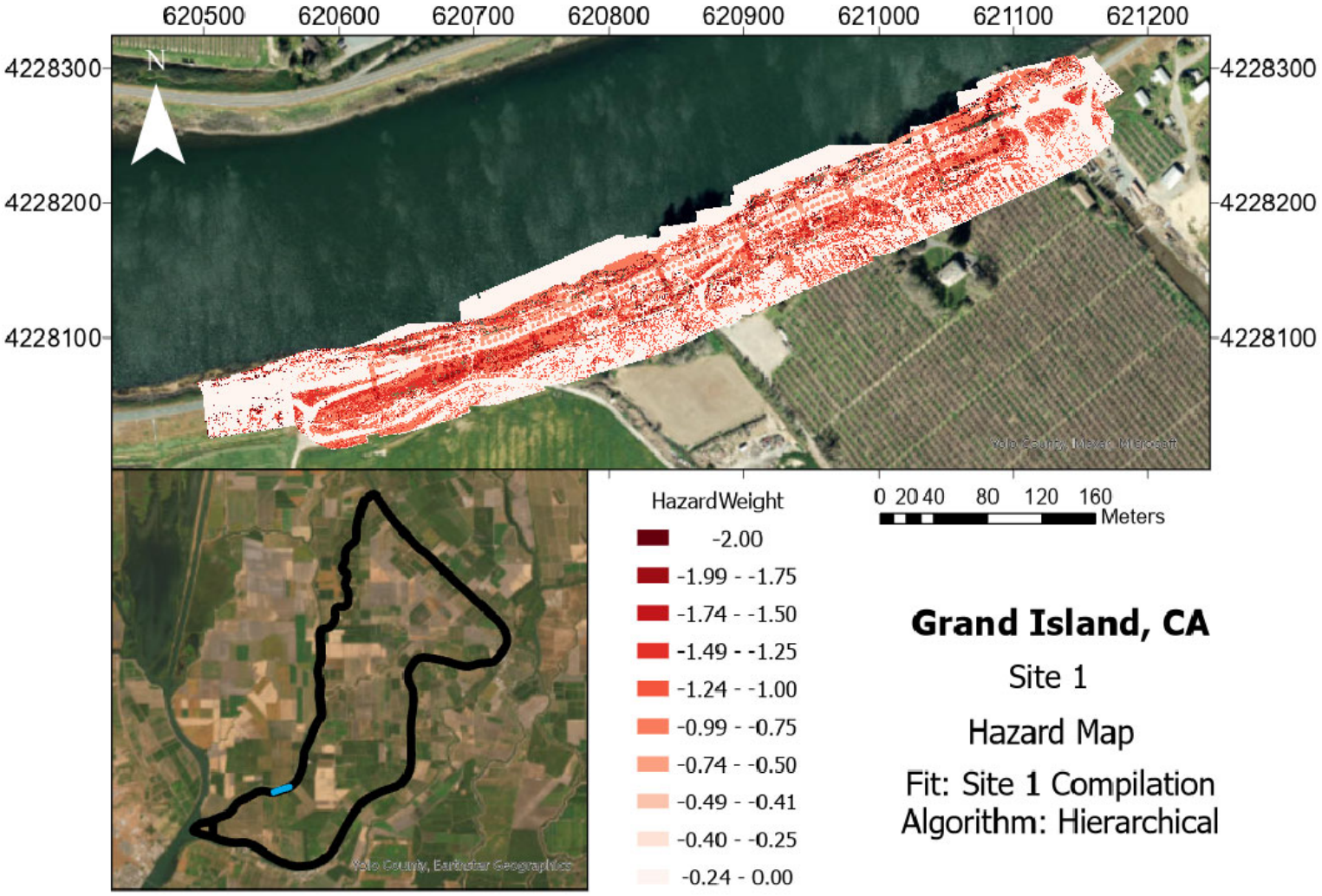
Final Hazard Model for 2 MASW Profiles at a Site

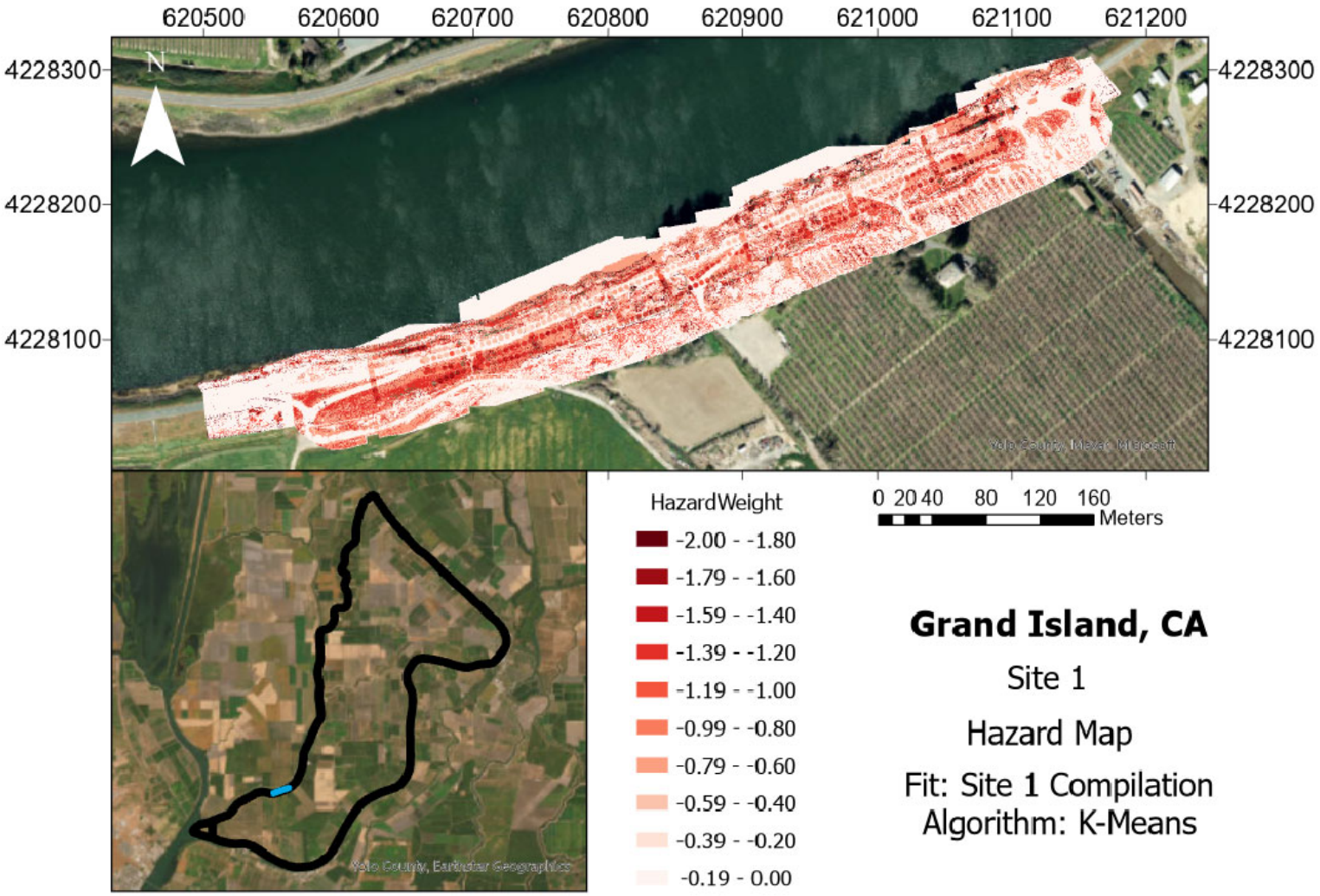


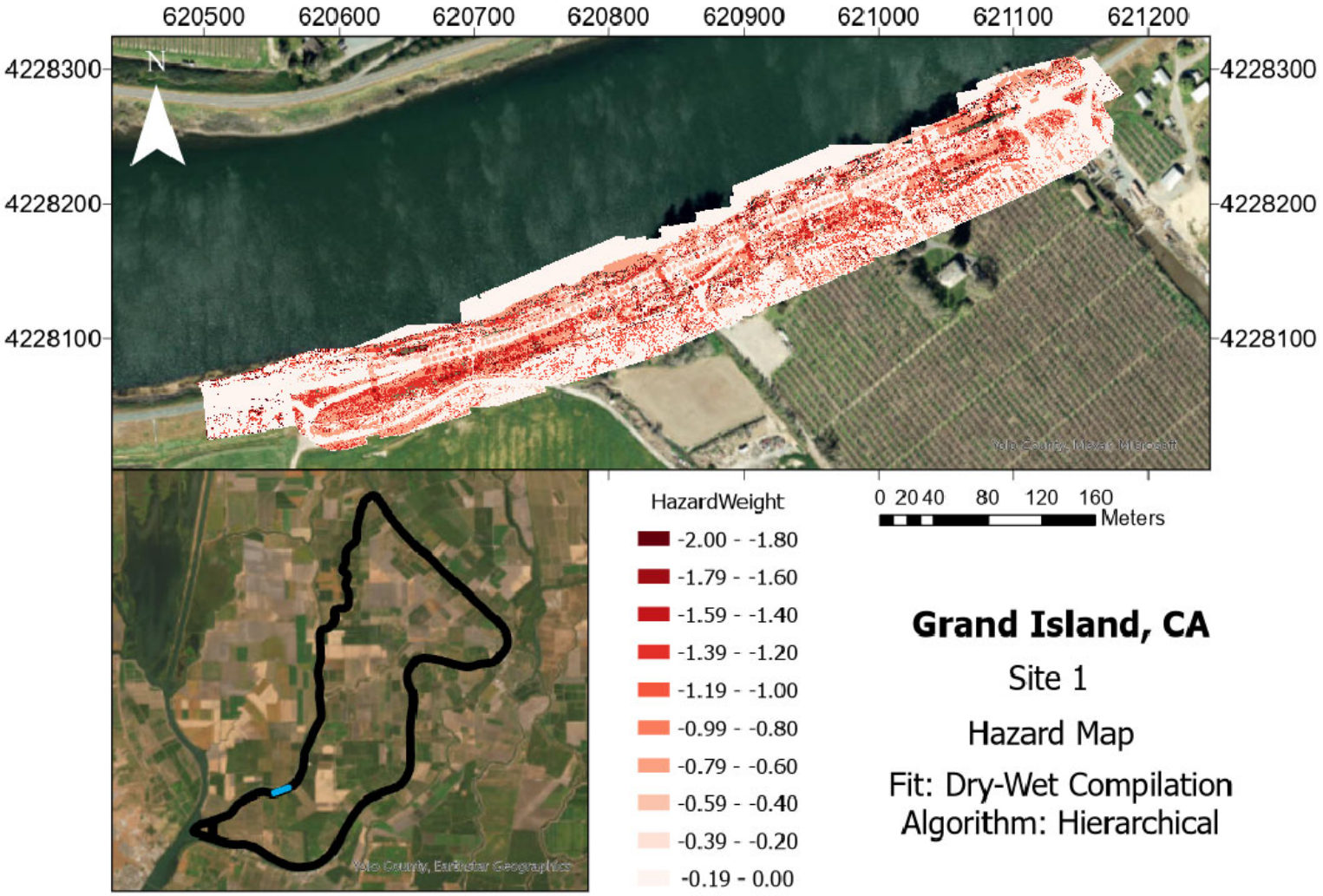
Appendix L – Hazard Maps and Comparisons

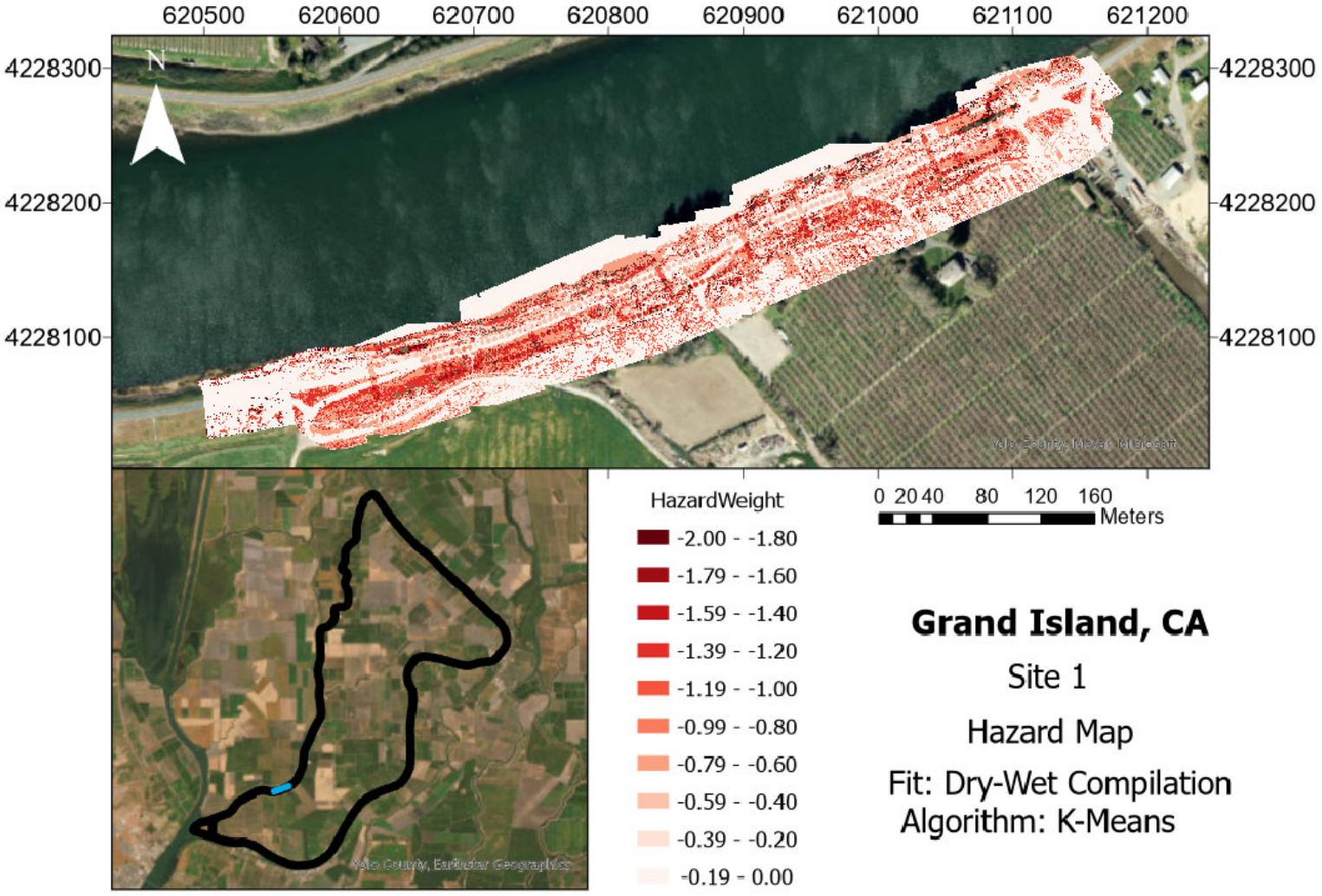

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Individual
Algorithm: Hierarchical

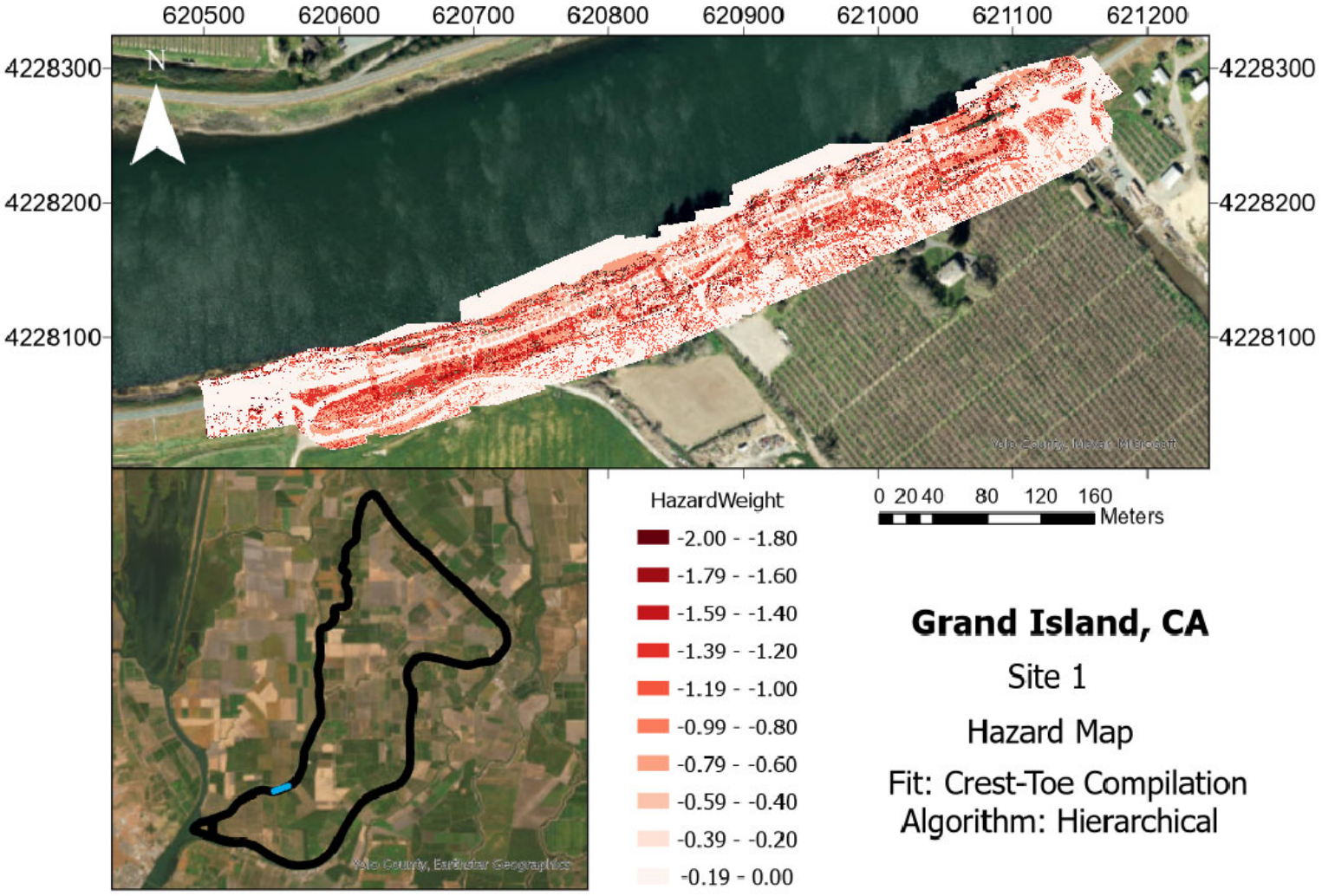

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00
-1.99 - -1.75
-1.74 - -1.50
-1.49 - -1.25
-1.24 - -1.00
-0.99 - -0.75
-0.74 - -0.50
-0.49 - -0.41
-0.40 - -0.25
-0.24 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Individual
Algorithm: K-Means

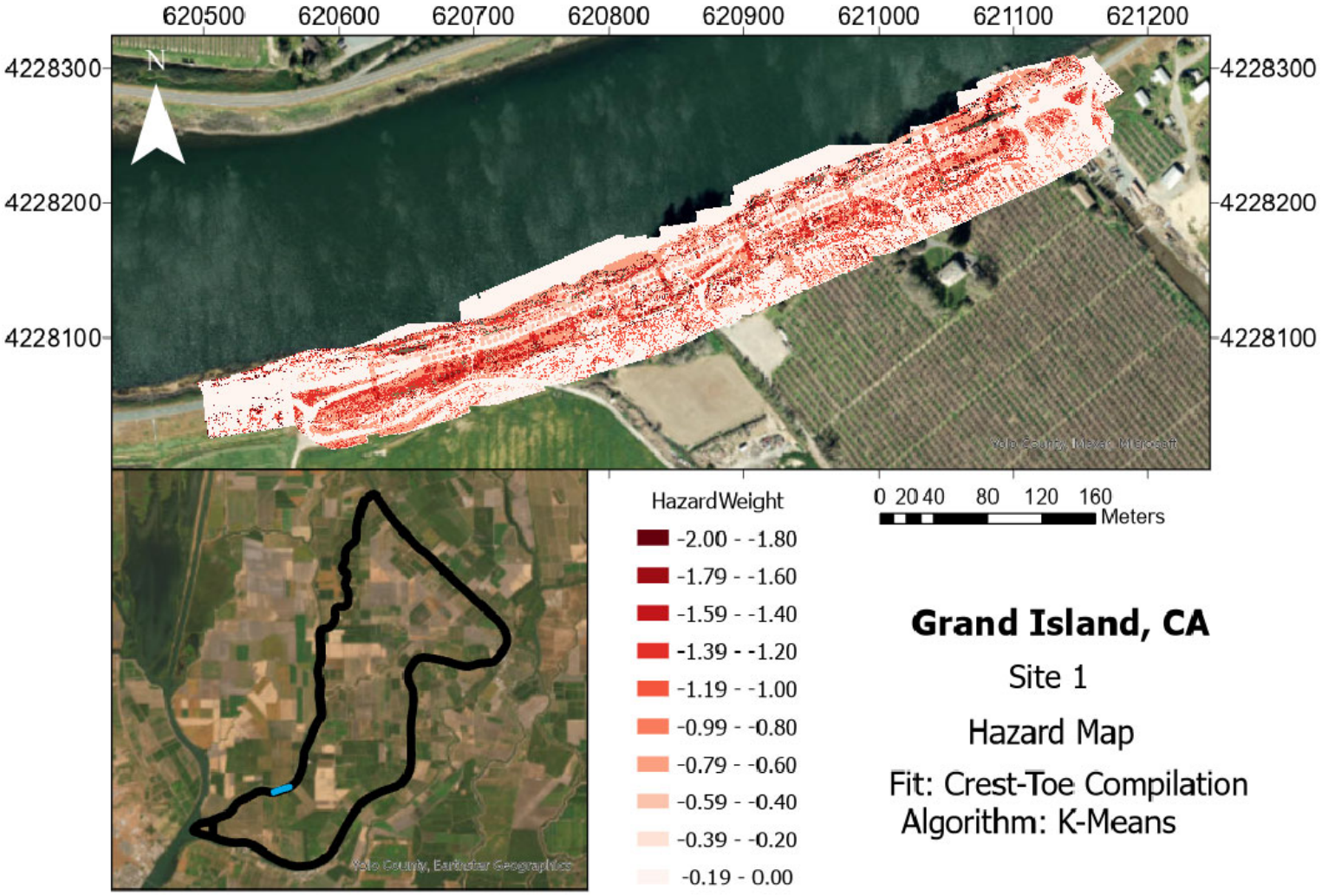

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00
-1.99 - -1.75
-1.74 - -1.50
-1.49 - -1.25
-1.24 - -1.00
-0.99 - -0.75
-0.74 - -0.50
-0.49 - -0.41
-0.40 - -0.25
-0.24 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Site 1 Compilation
Algorithm: Hierarchical

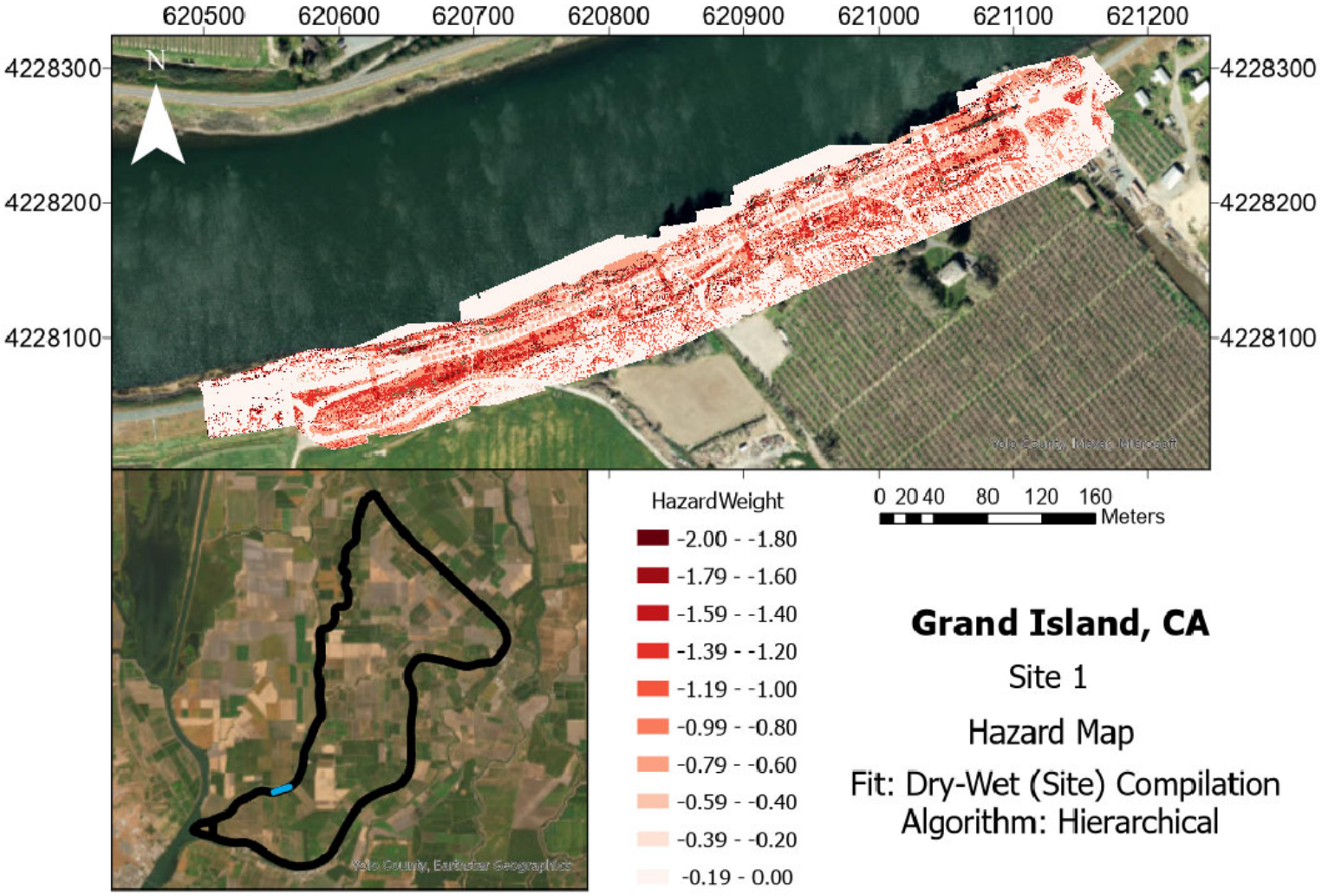

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Site 1 Compilation
Algorithm: K-Means

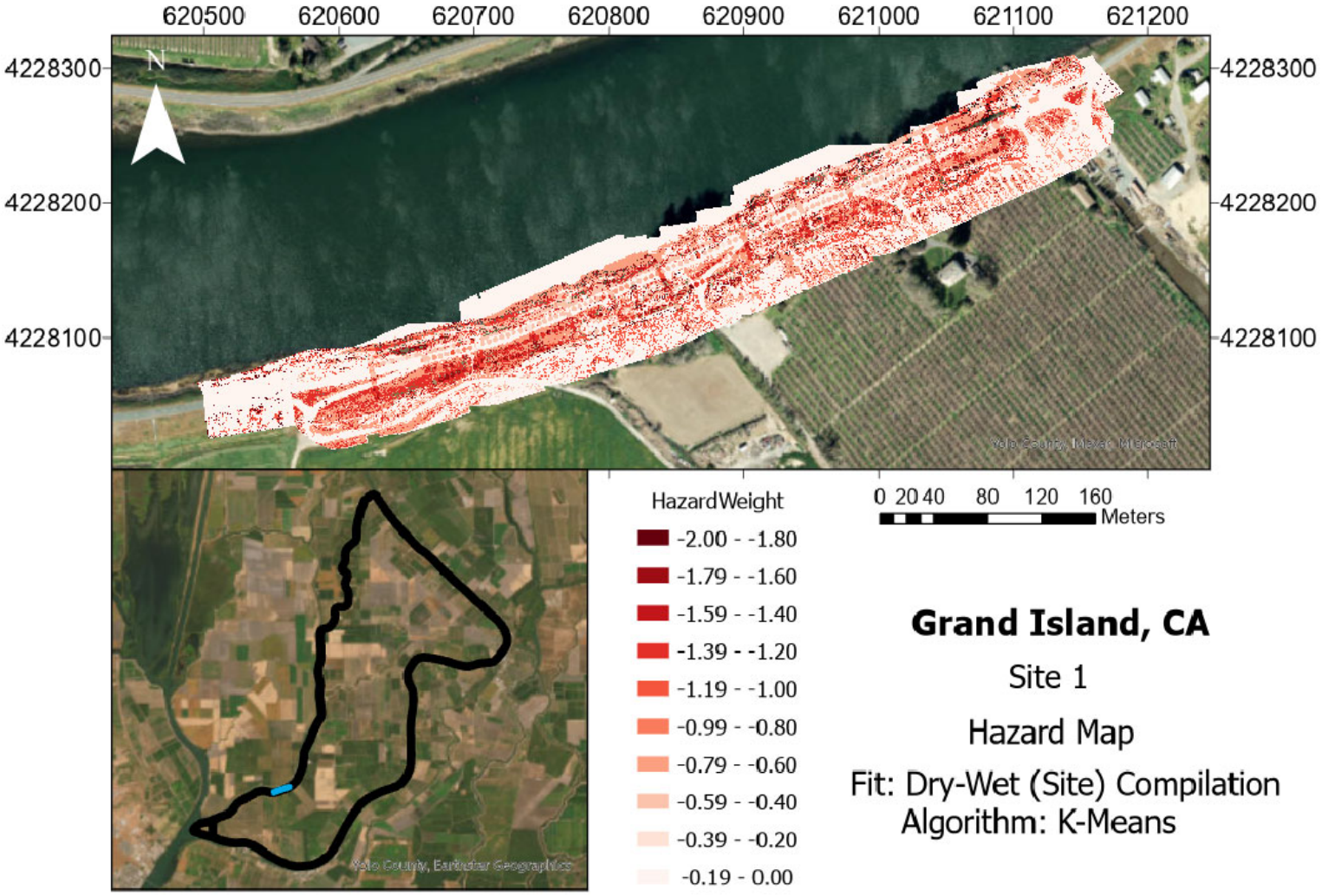

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Dry-Wet Compilation
Algorithm: Hierarchical

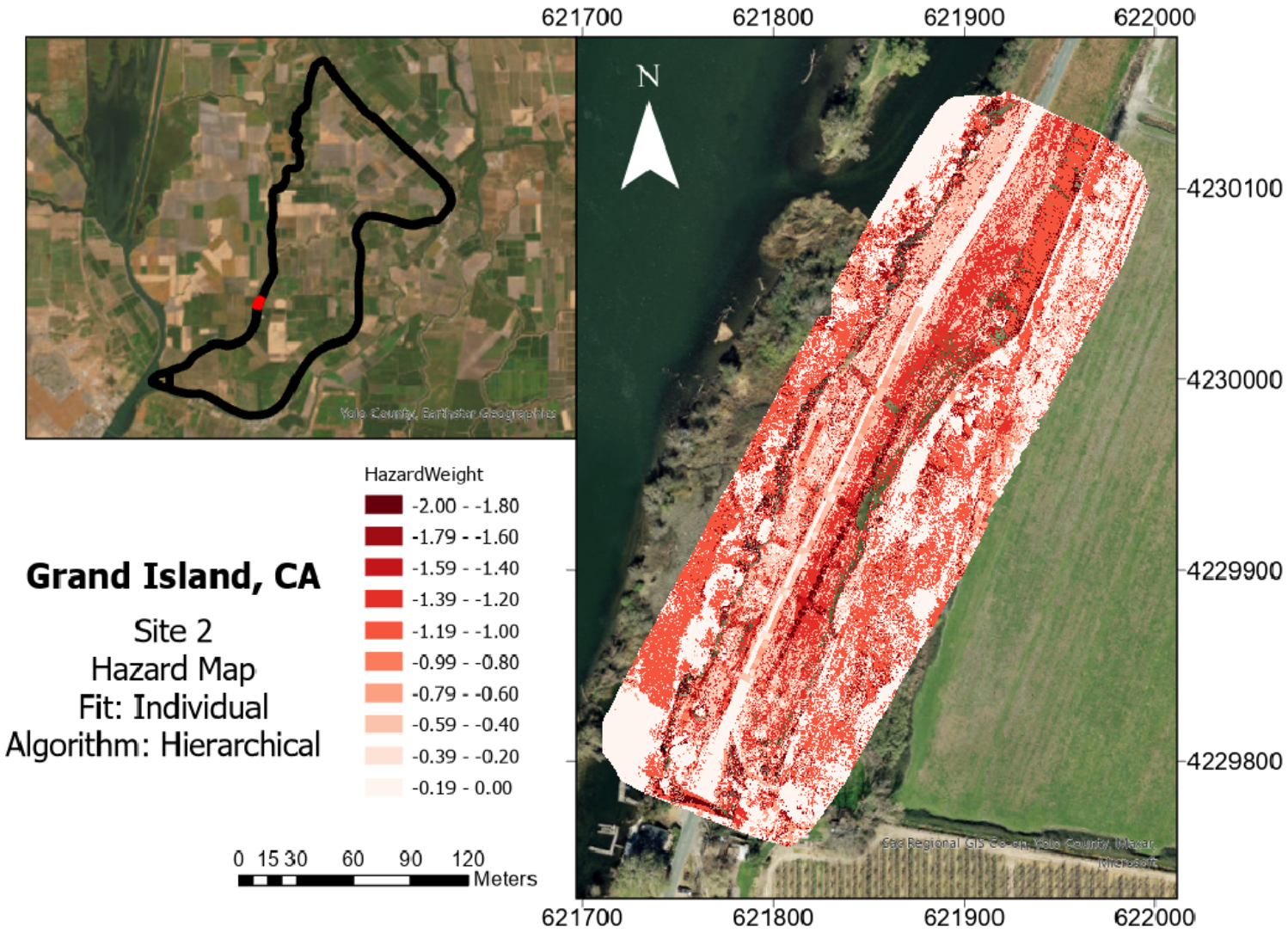

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Dry-Wet Compilation
Algorithm: K-Means

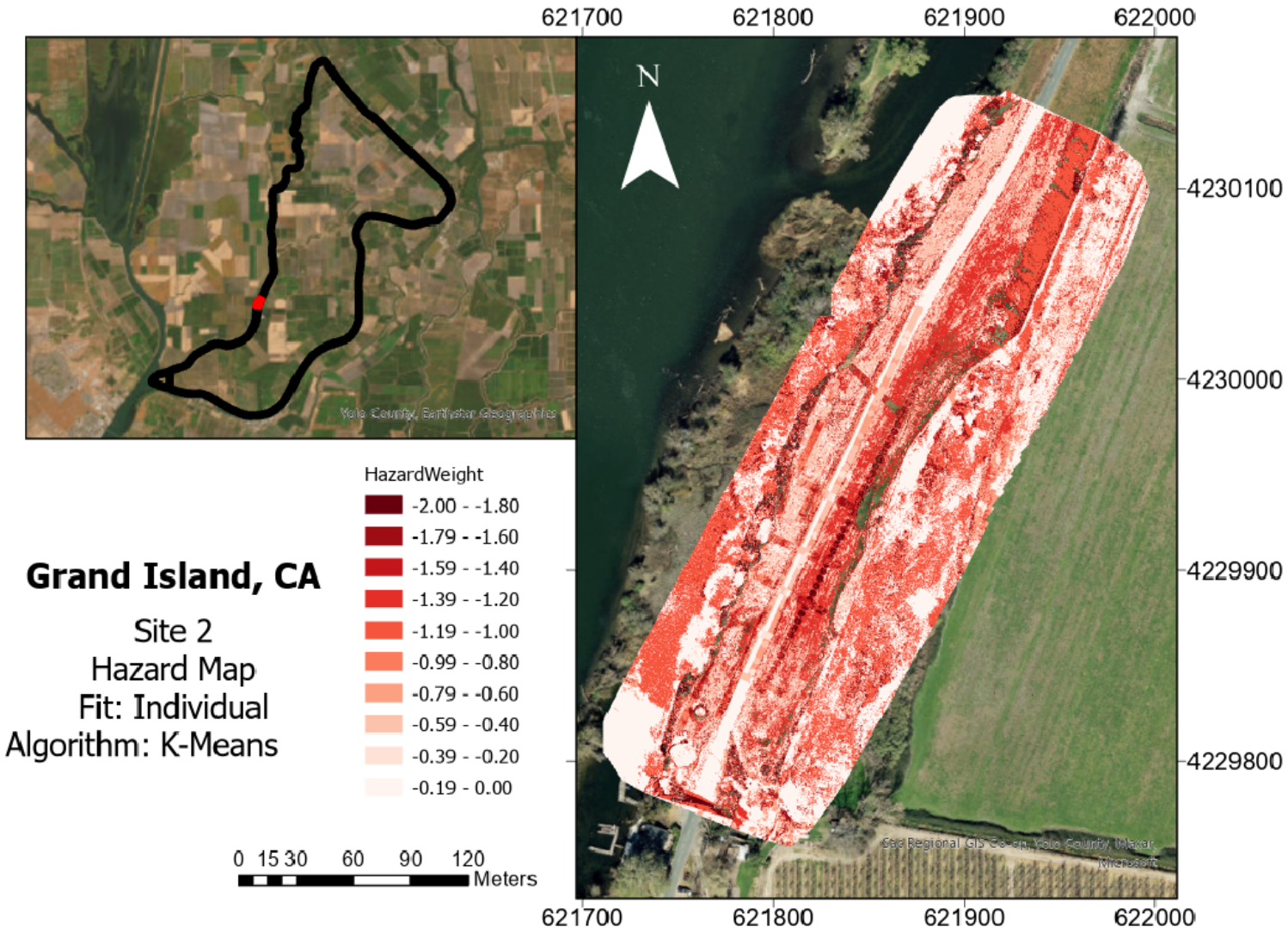

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Crest-Toe Compilation
Algorithm: Hierarchical

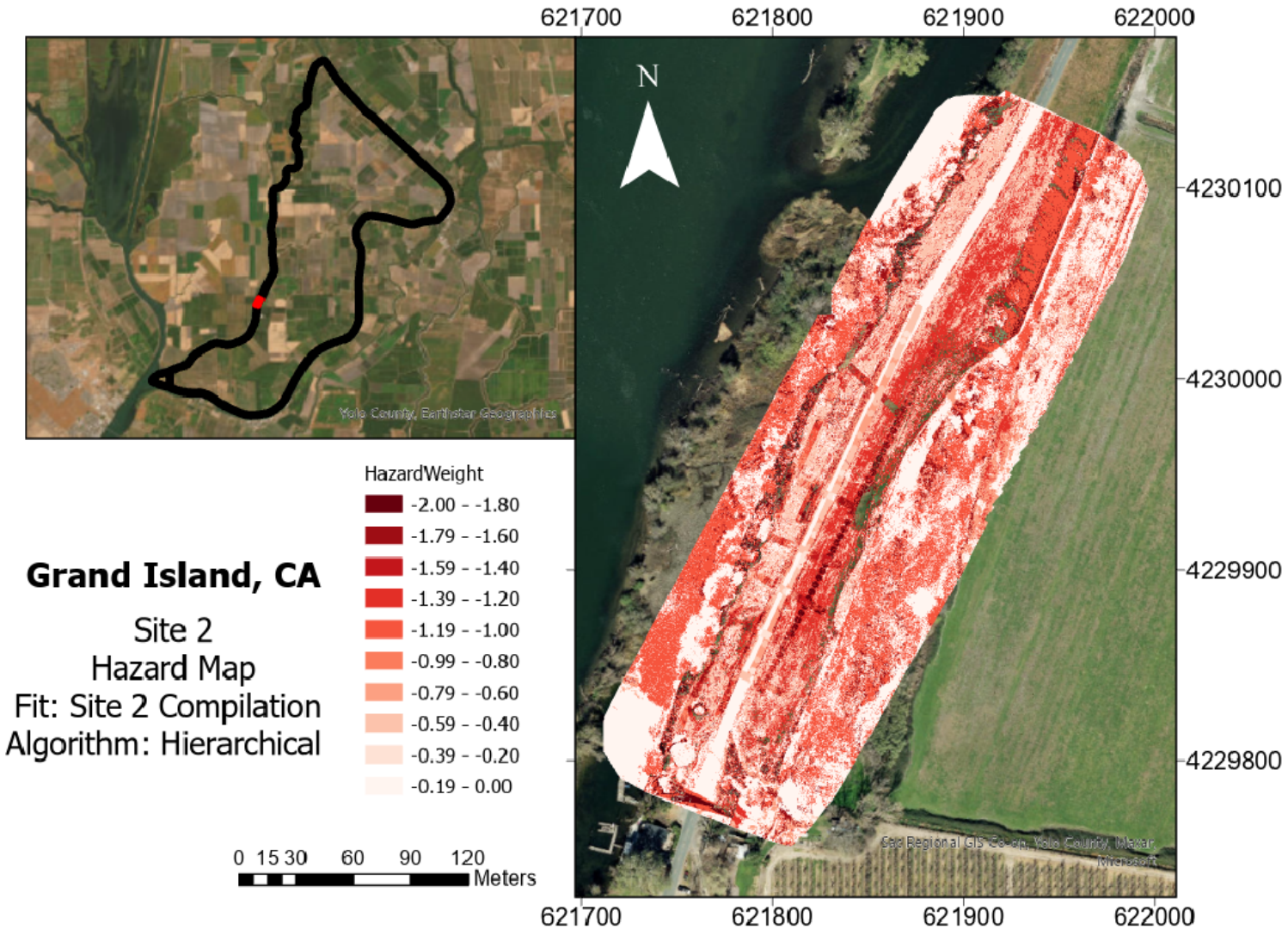

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Crest-Toe Compilation
Algorithm: K-Means

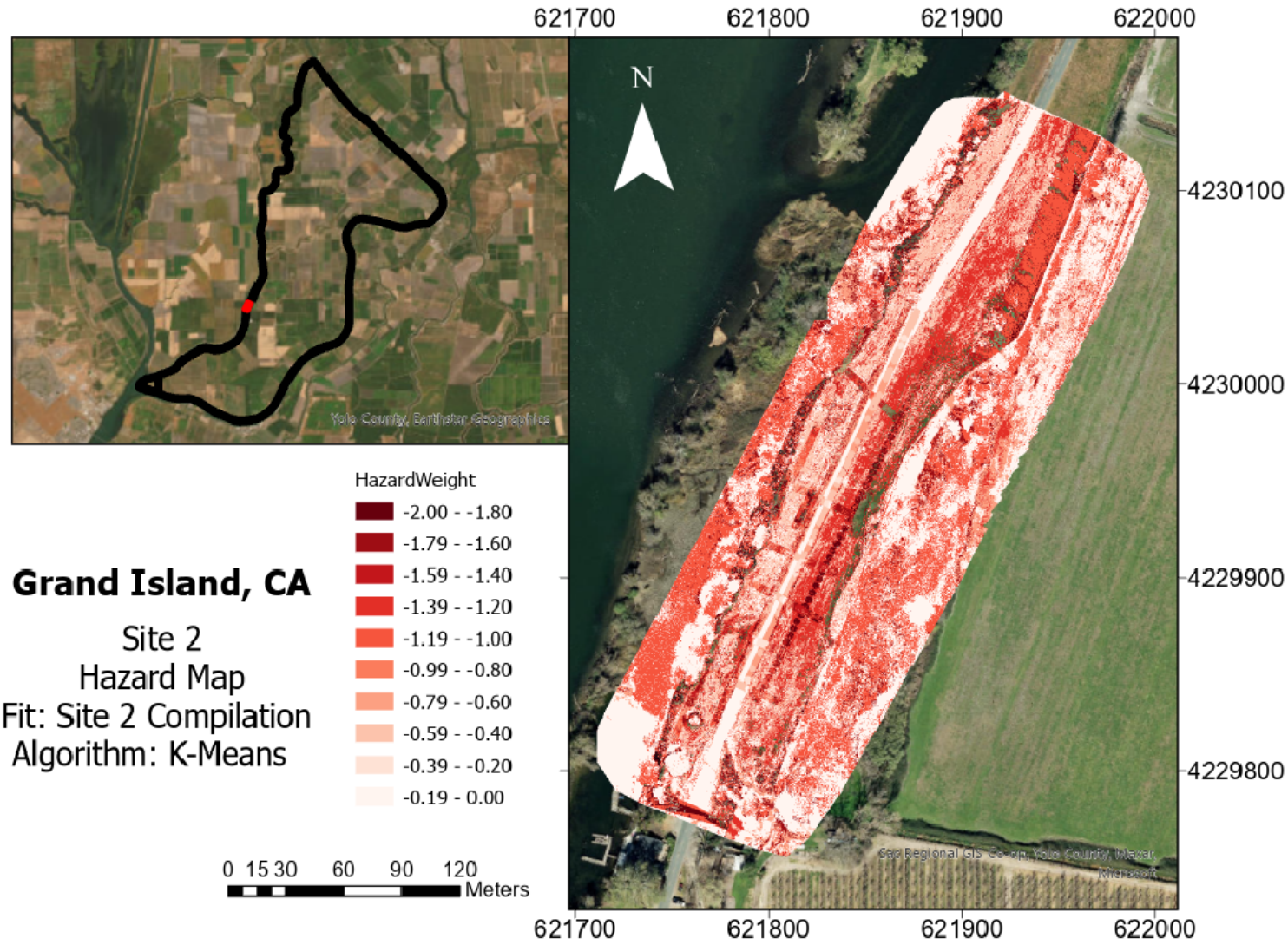

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Dry-Wet (Site) Compilation
Algorithm: Hierarchical

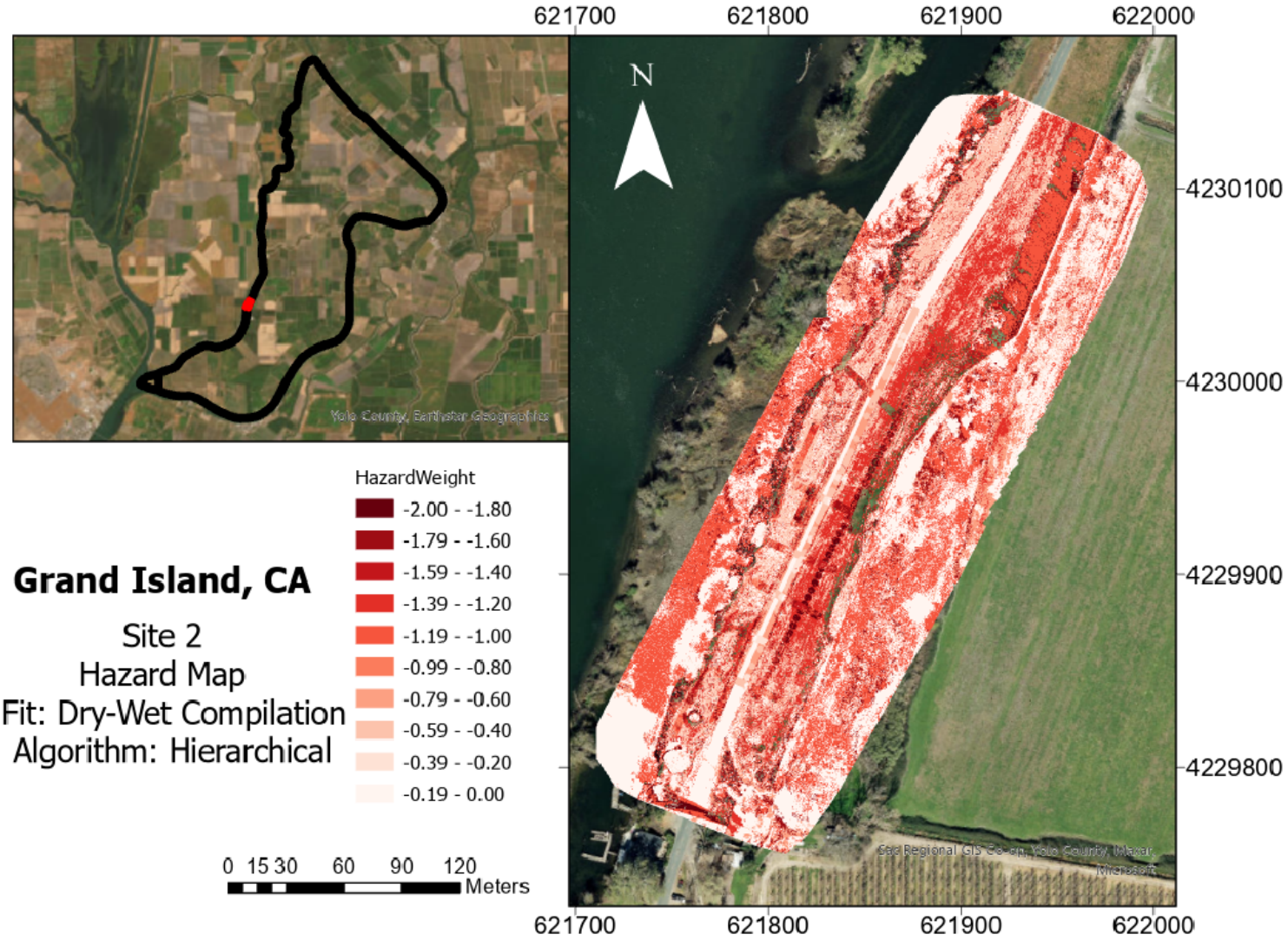

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Hazard Map
Fit: Dry-Wet (Site) Compilation
Algorithm: K-Means

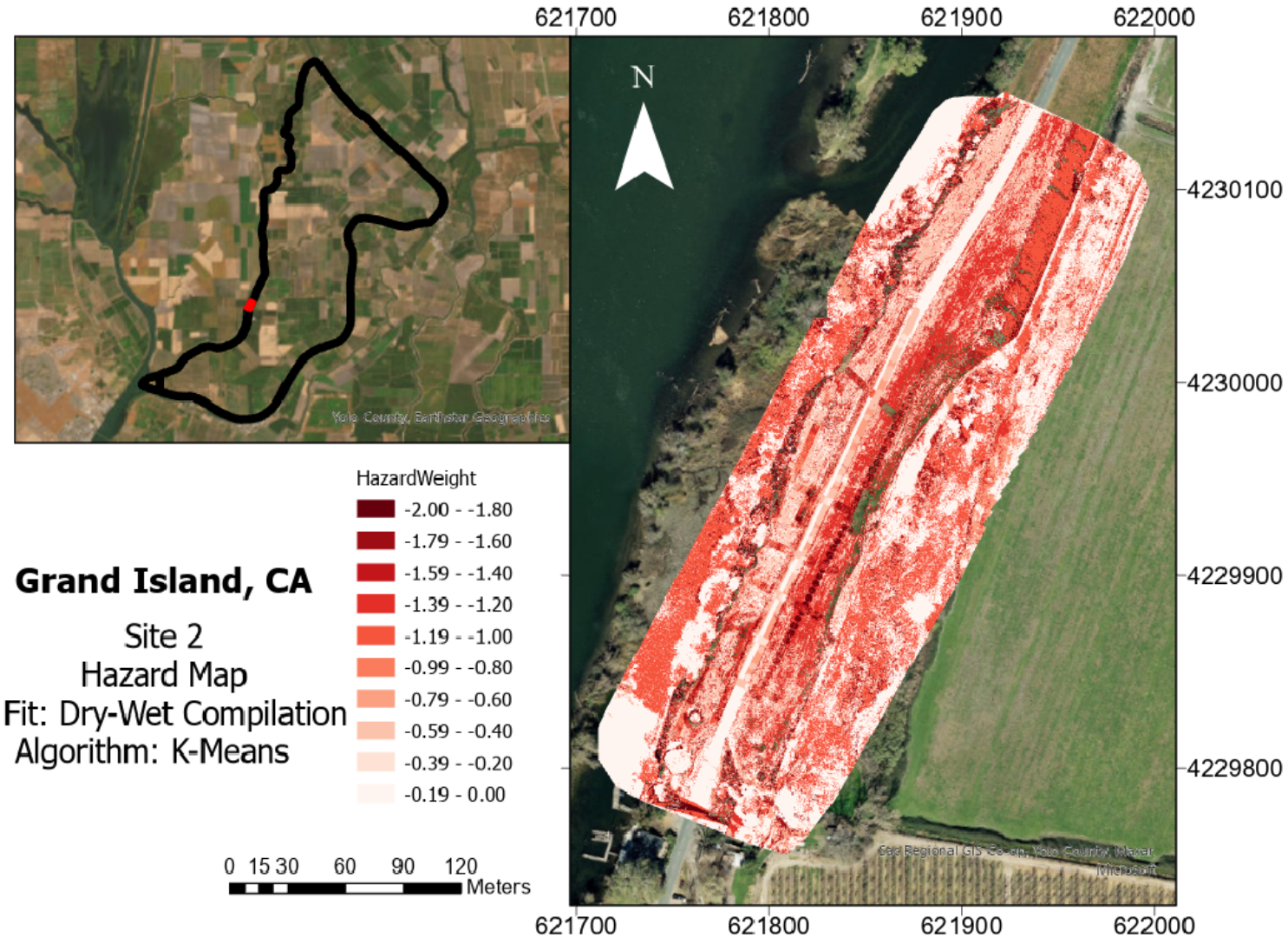

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Individual
Algorithm: Hierarchical
0 15 30 60 90 120
Meters

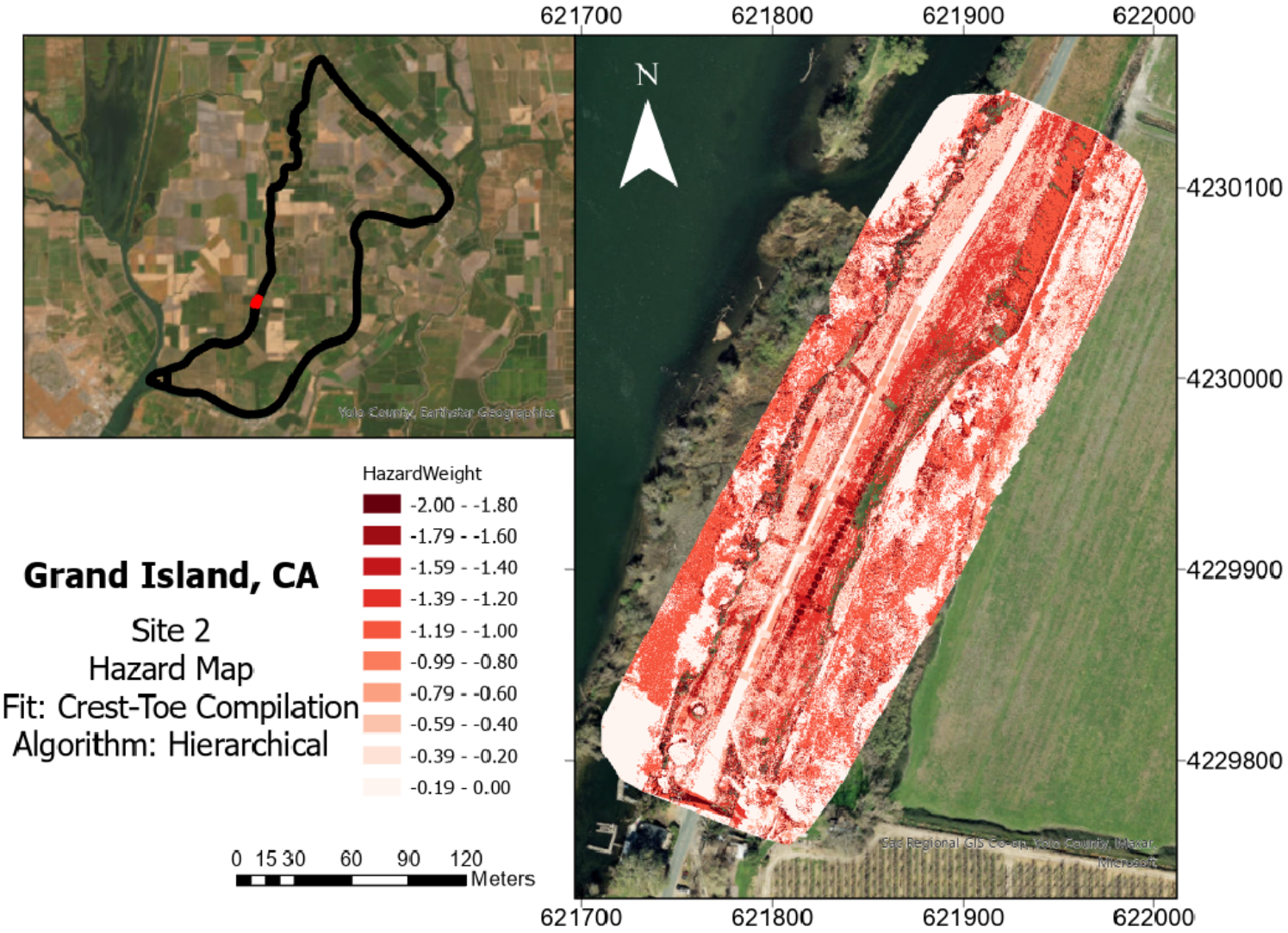

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Individual
Algorithm: K-Means
0 15 30 60 90 120
Meters

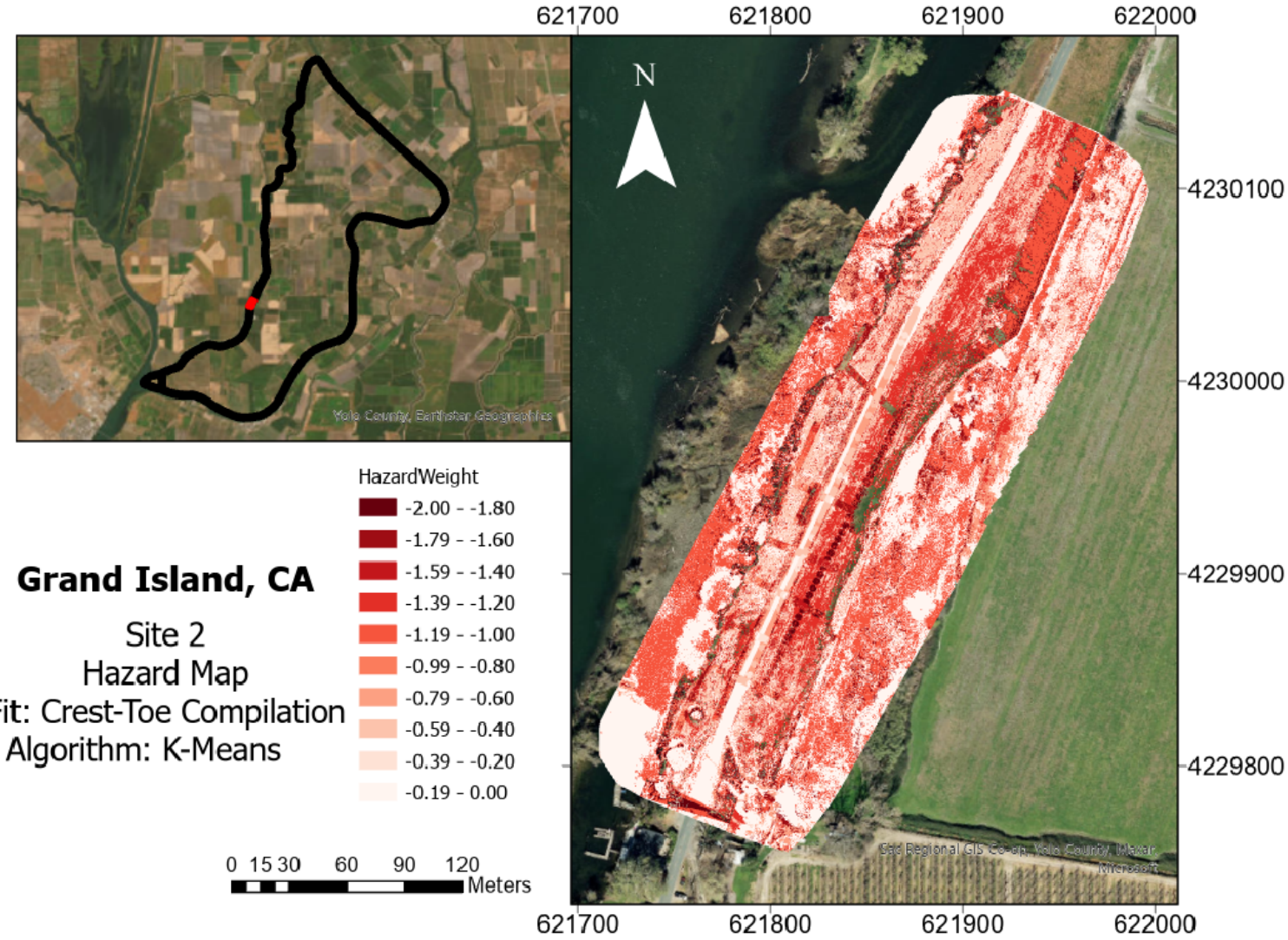

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Site 2 Compilation
Algorithm: Hierarchical
0 15 30 60 90 120
Meters

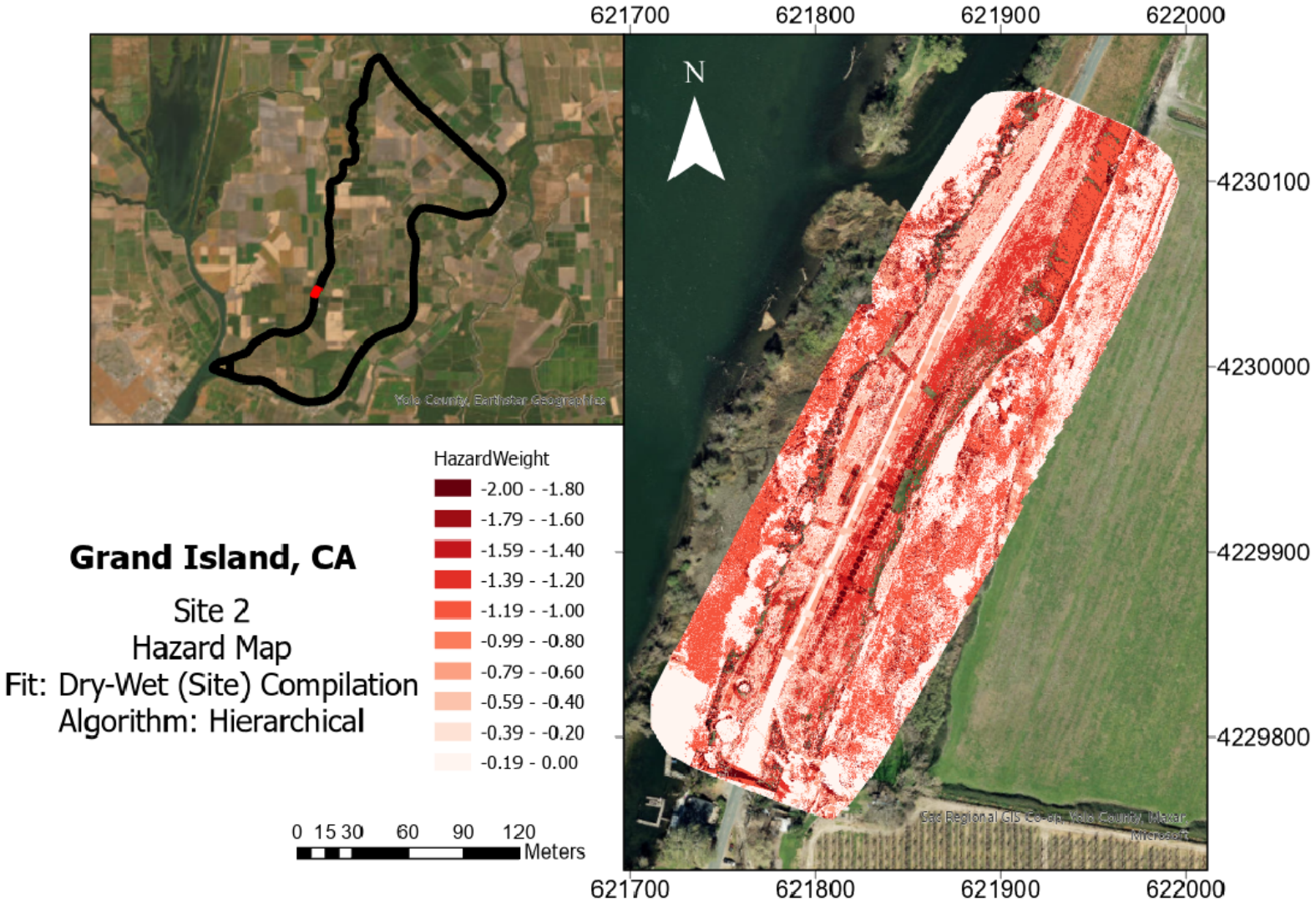

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Site 2 Compilation
Algorithm: K-Means
0 15 30 60 90 120
Meters

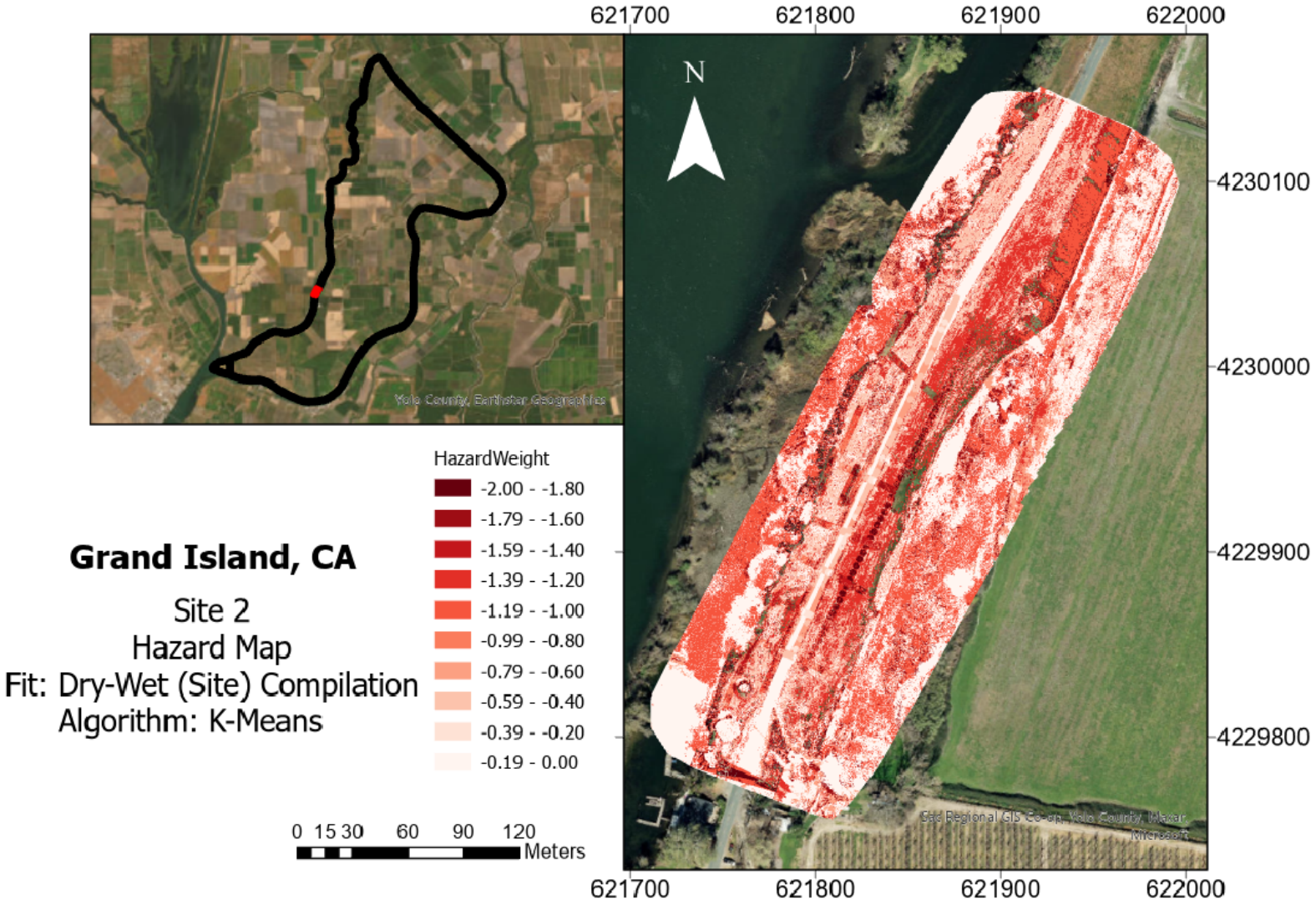

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Dry-Wet Compilation
Algorithm: Hierarchical
0 15 30 60 90 120
Meters

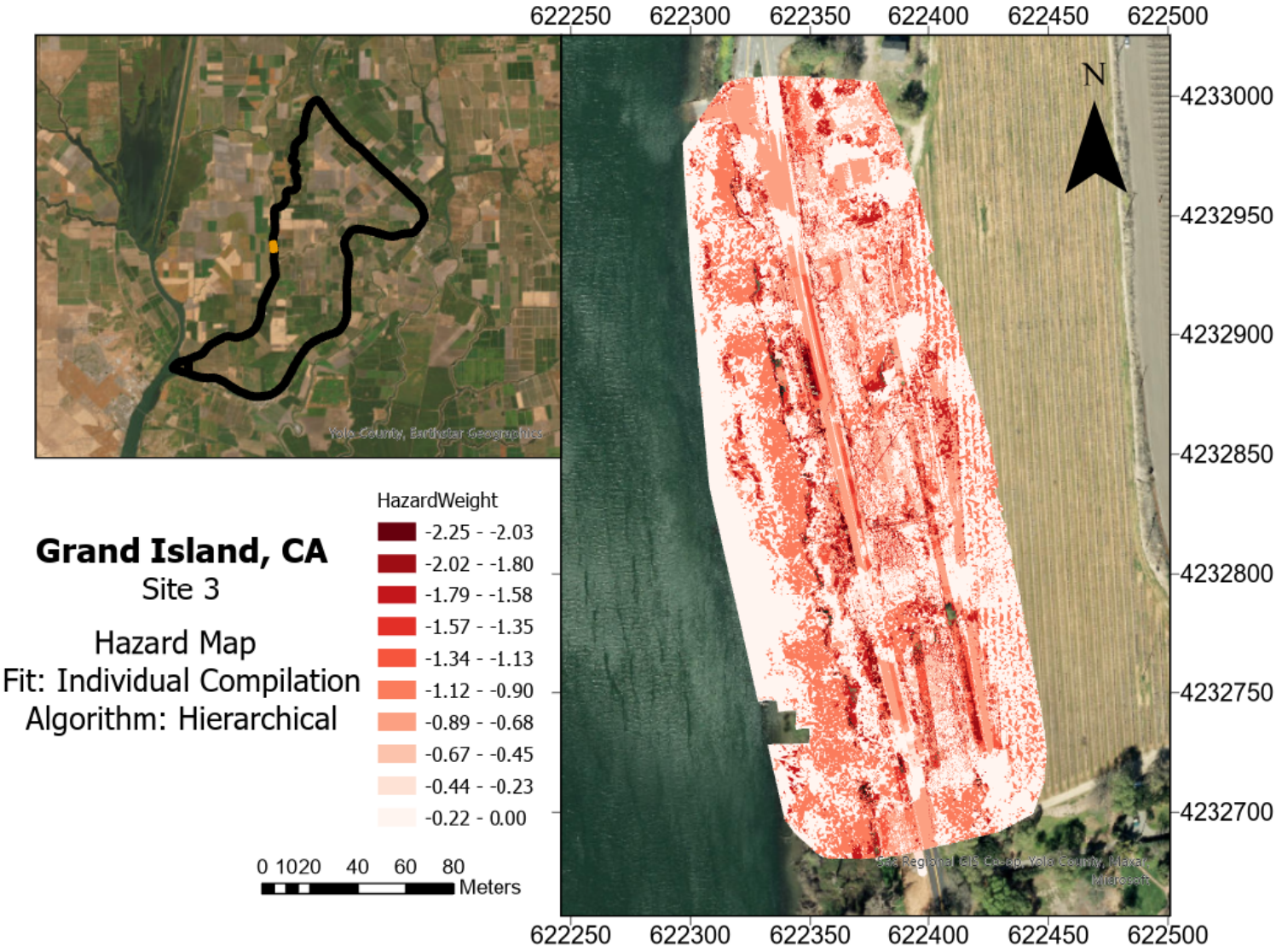

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Dry-Wet Compilation
Algorithm: K-Means
0 15 30 60 90 120
Meters

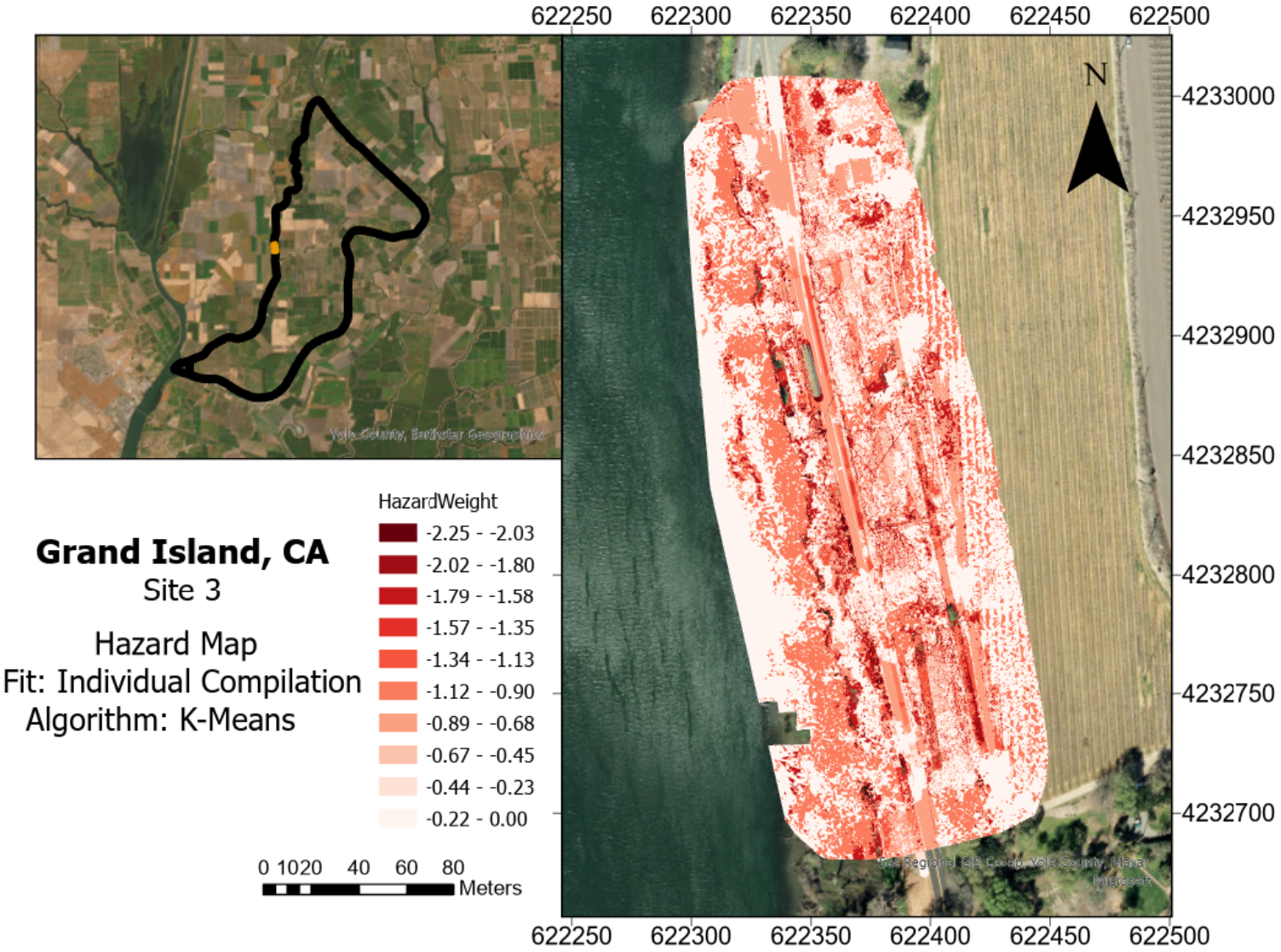

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Crest-Toe Compilation
Algorithm: Hierarchical
0 15 30 60 90 120
Meters

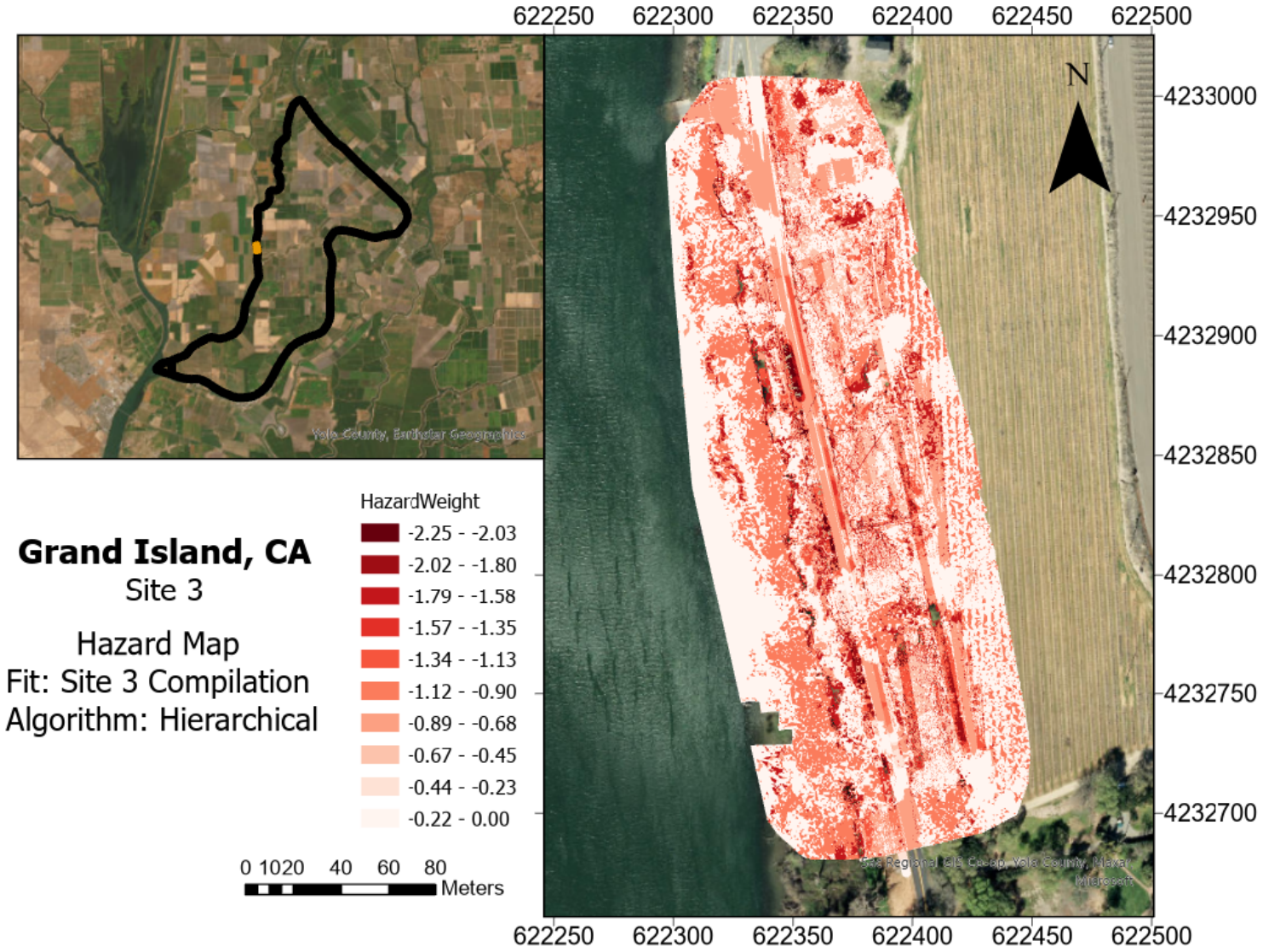

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Crest-Toe Compilation
Algorithm: K-Means
0 15 30 60 90 120
Meters

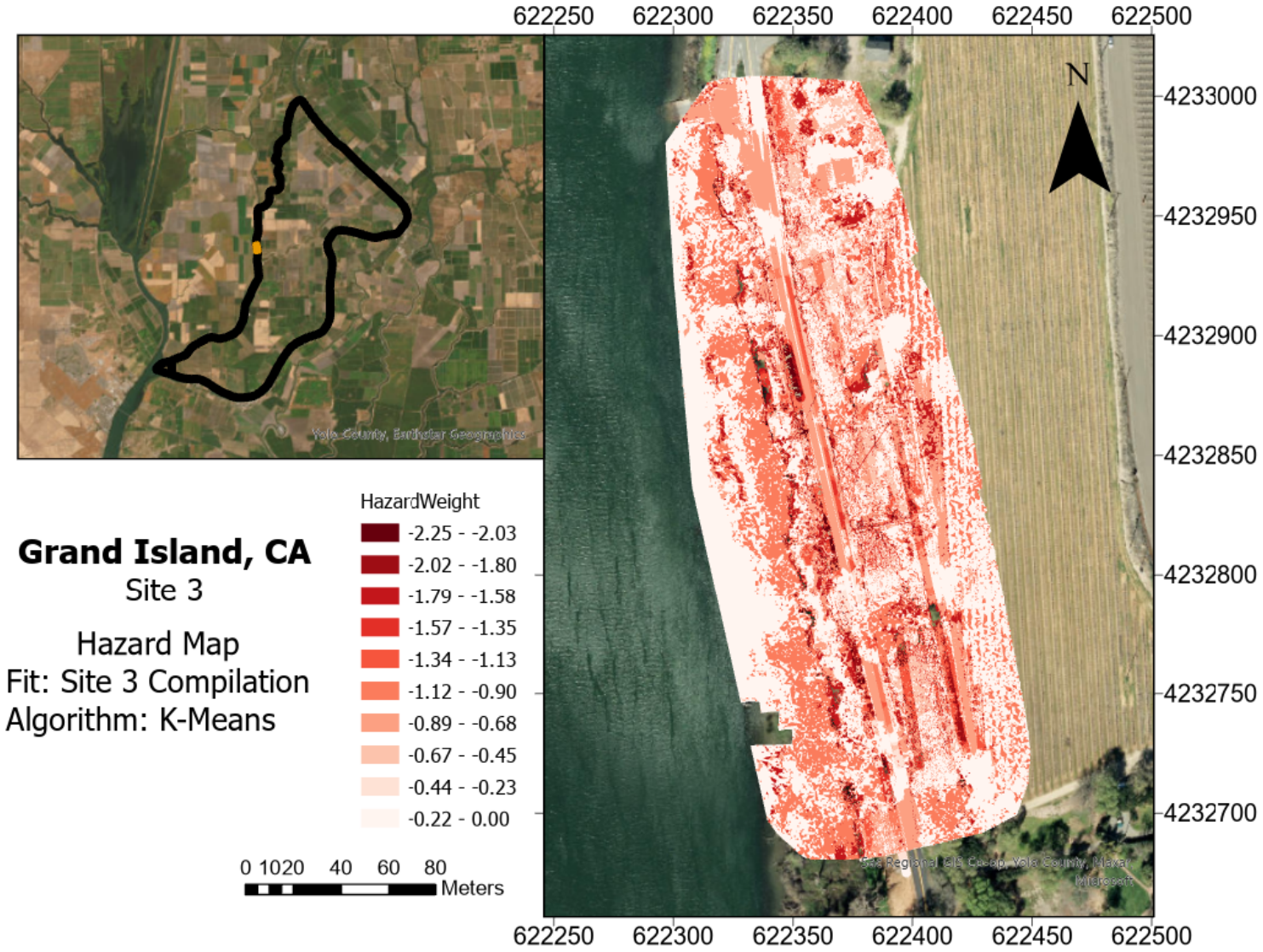

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Dry-Wet (Site) Compilation
Algorithm: Hierarchical
0 15 30 60 90 120
Meters

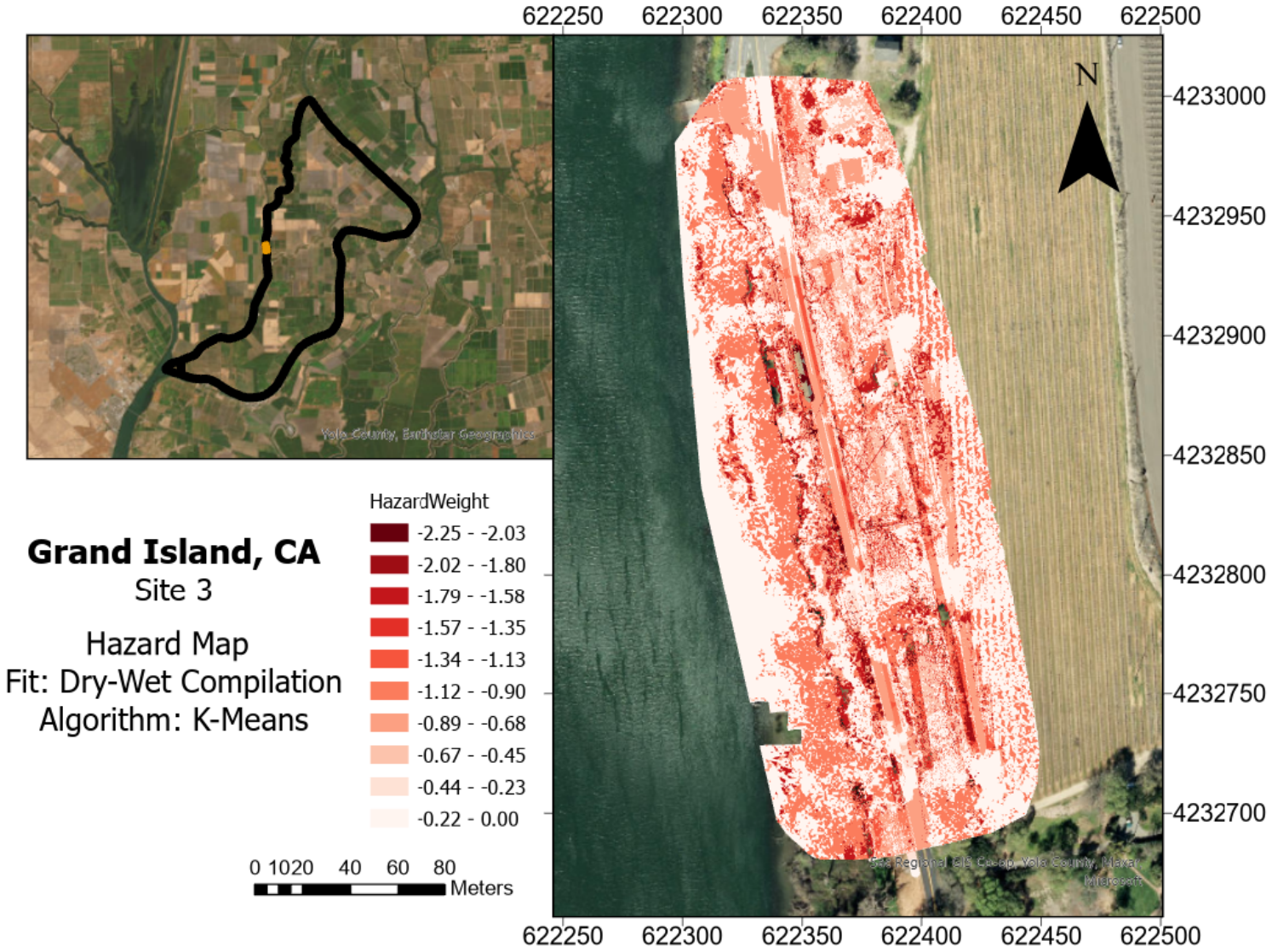

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
HazardWeight
-2.00 - -1.80
-1.79 - -1.60
-1.59 - -1.40
-1.39 - -1.20
-1.19 - -1.00
-0.99 - -0.80
-0.79 - -0.60
-0.59 - -0.40
-0.39 - -0.20
-0.19 - 0.00
Grand Island, CA
Site 2
Hazard Map
Fit: Dry-Wet (Site) Compilation
Algorithm: K-Means
0 15 30 60 90 120
Meters

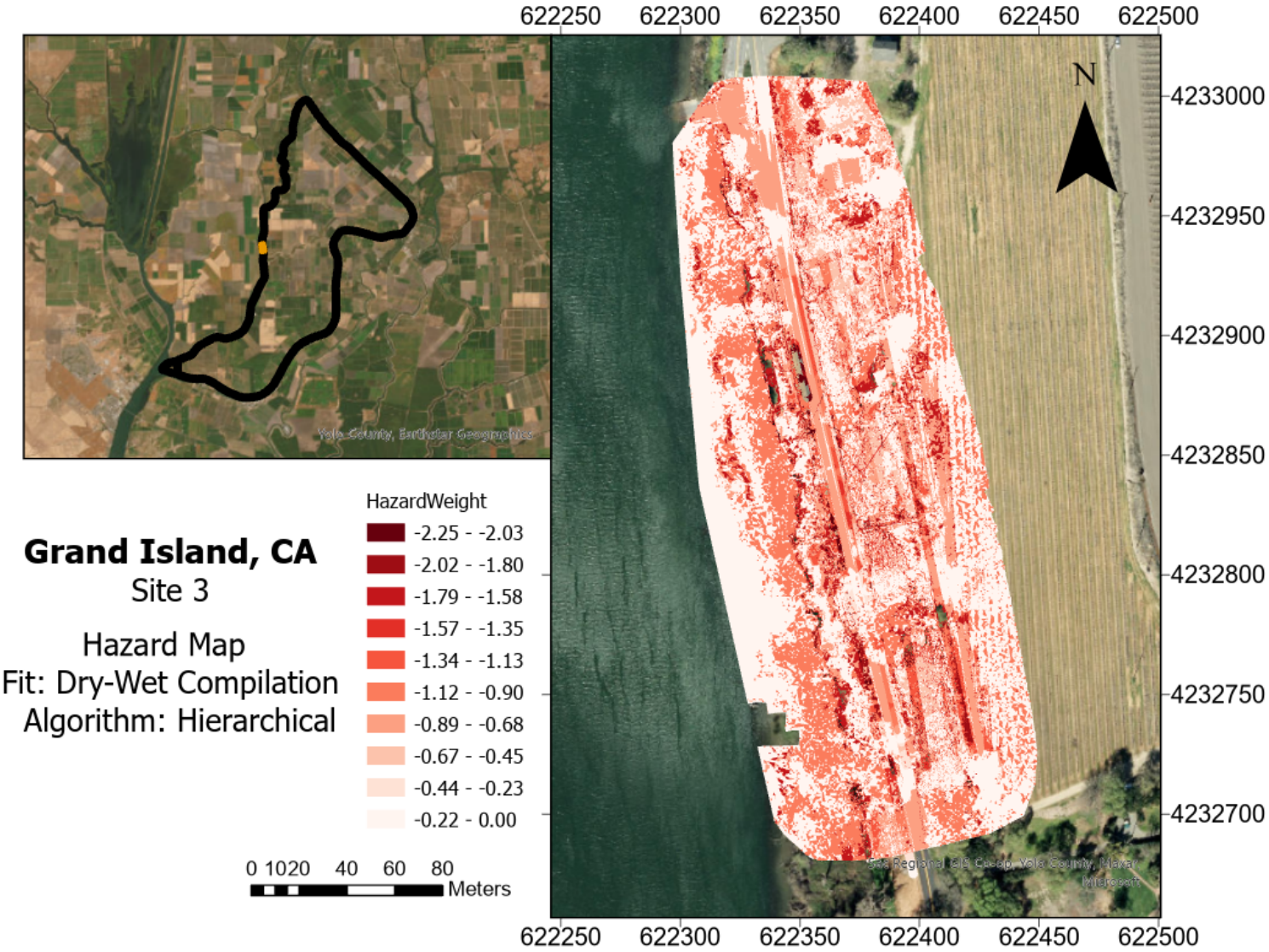

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
Grand Island, CA
Site 3
Hazard Map
Fit: Individual Compilation
Algorithm: Hierarchical
0 1020 40 60 80
Meters

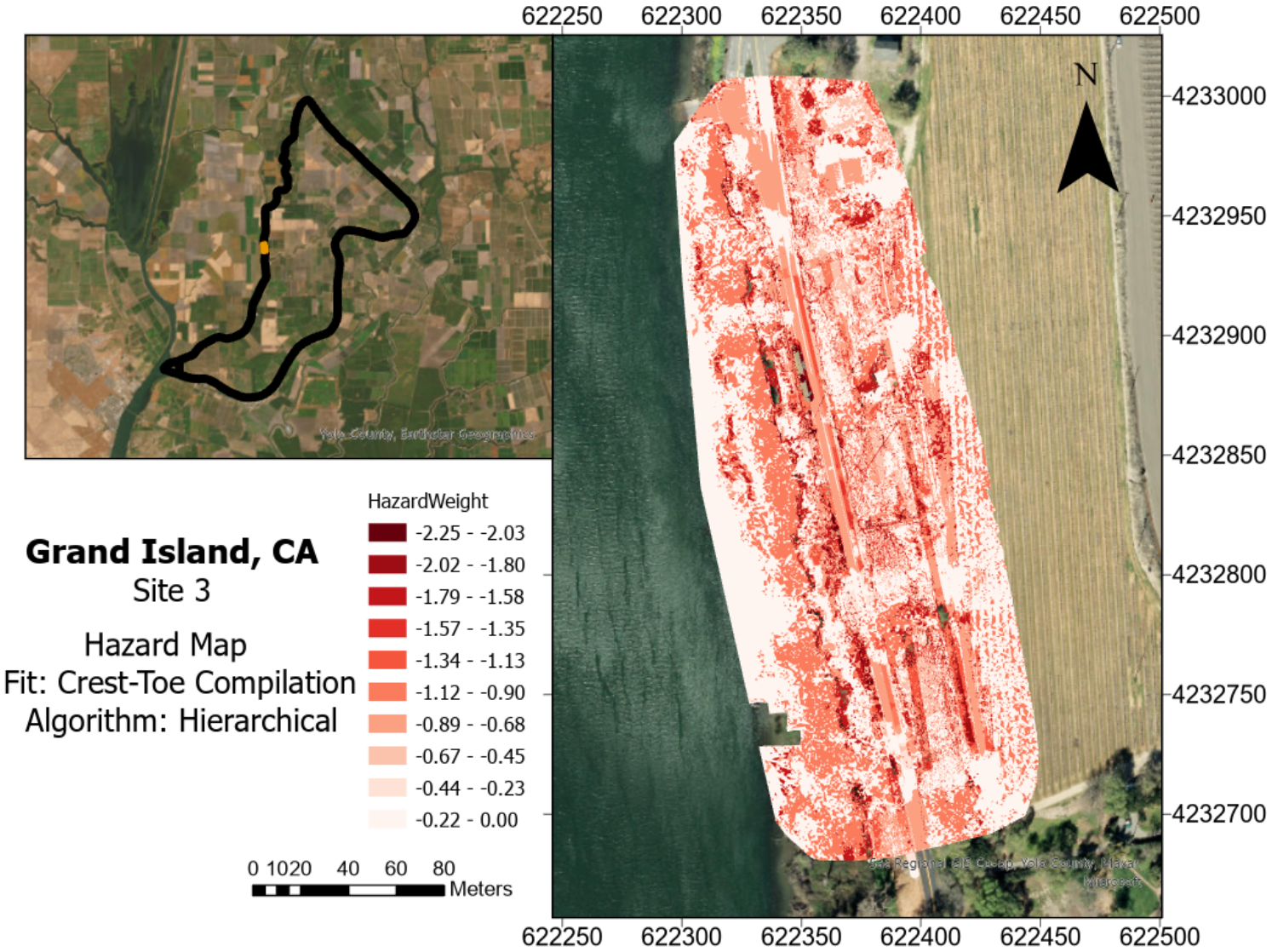

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Individual Compilation
Algorithm: K-Means
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters

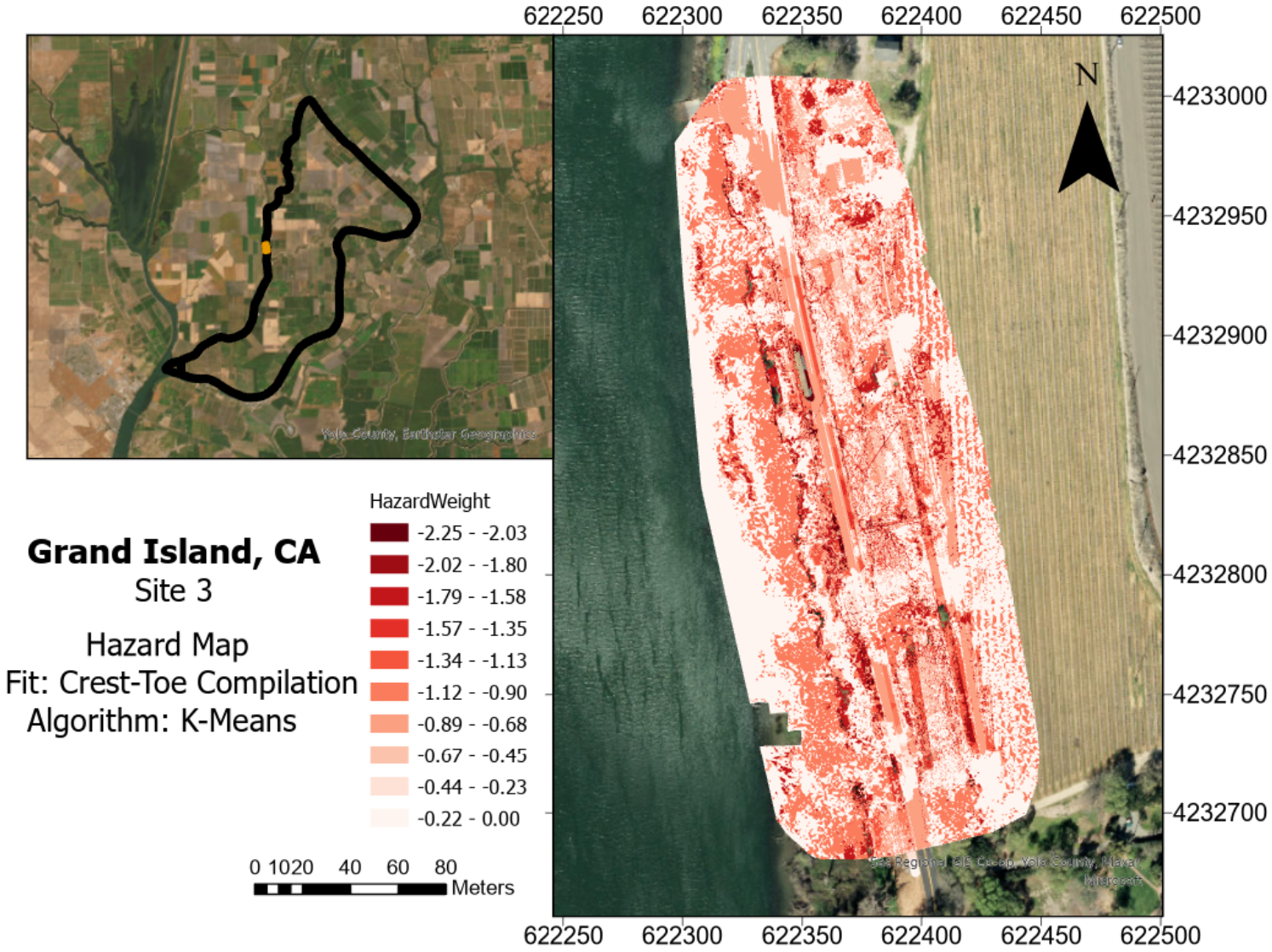

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Site 3 Compilation
Algorithm: Hierarchical
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Site 3 Compilation
Algorithm: K-Means
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Dry-Wet Compilation
Algorithm: K-Means
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
Grand Island, CA
Site 3
Hazard Map
Fit: Dry-Wet Compilation
Algorithm: Hierarchical
0 1020 40 60 80
Meters

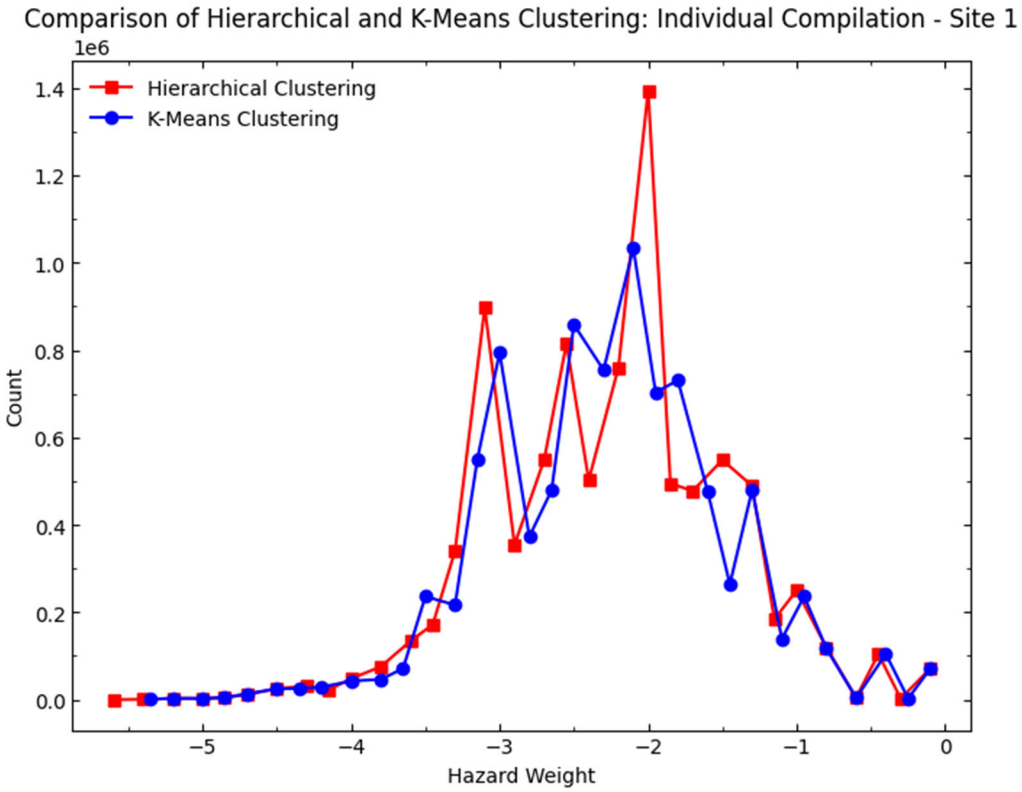

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Crest-Toe Compilation
Algorithm: Hierarchical
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters

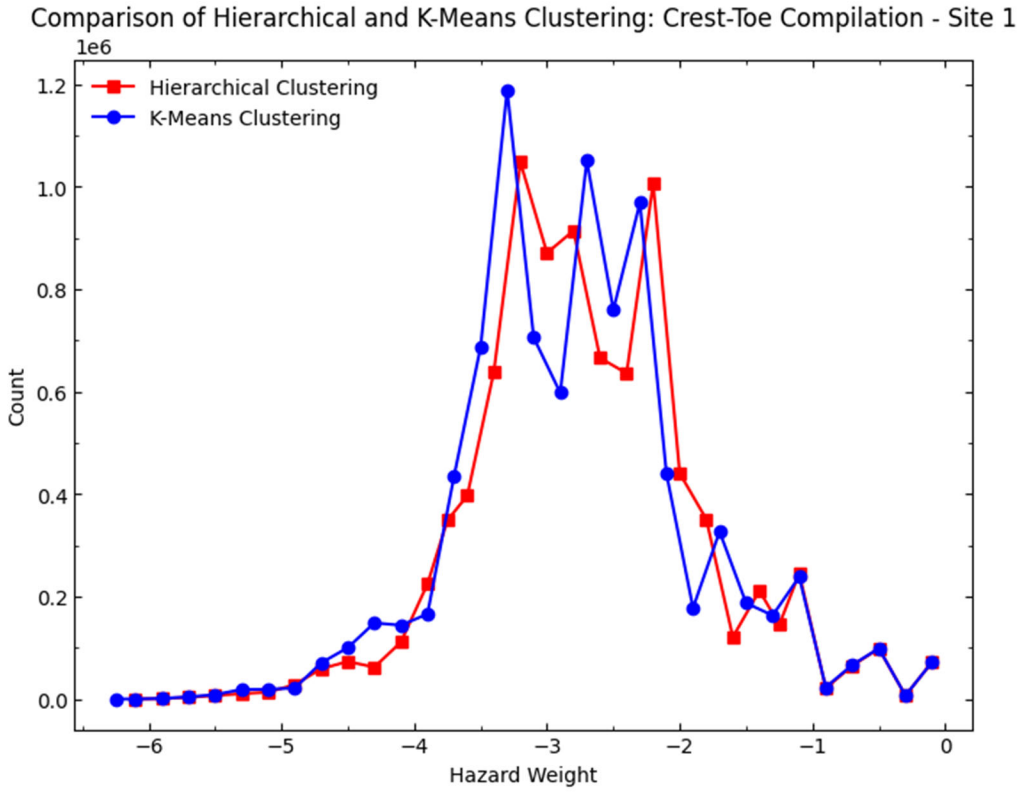

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Crest-Toe Compilation
Algorithm: K-Means
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 1020 40 60 80
Meters

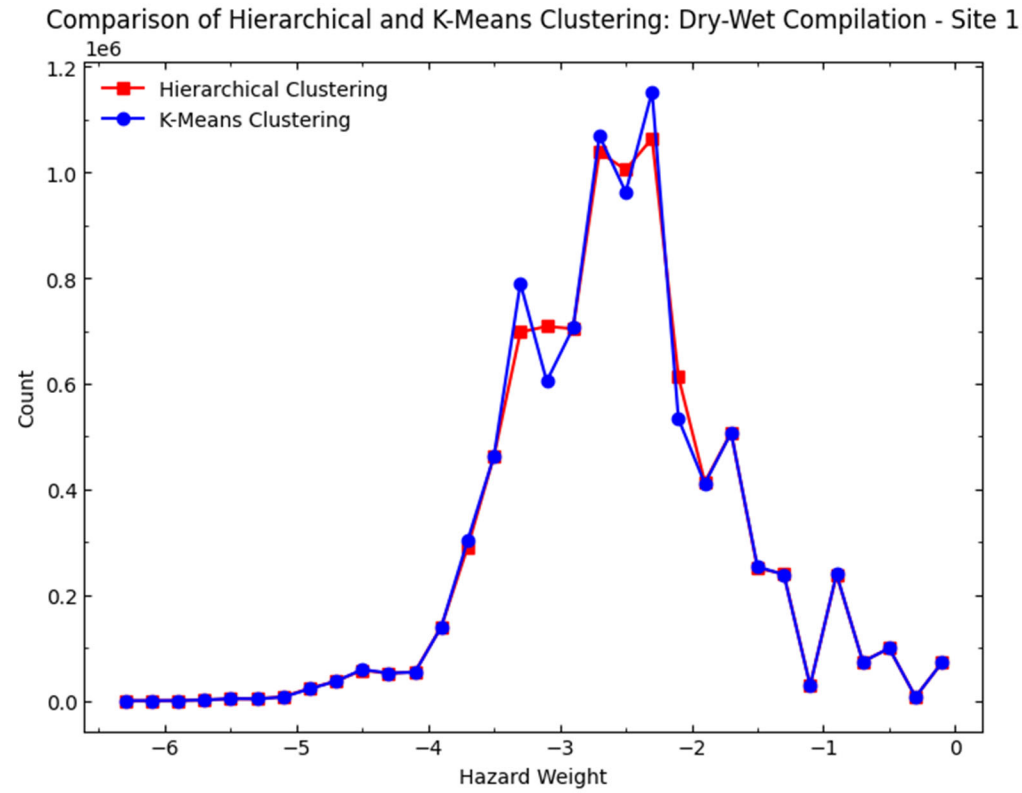

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Dry-Wet (Site) Compilation
Algorithm: Hierarchical
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters

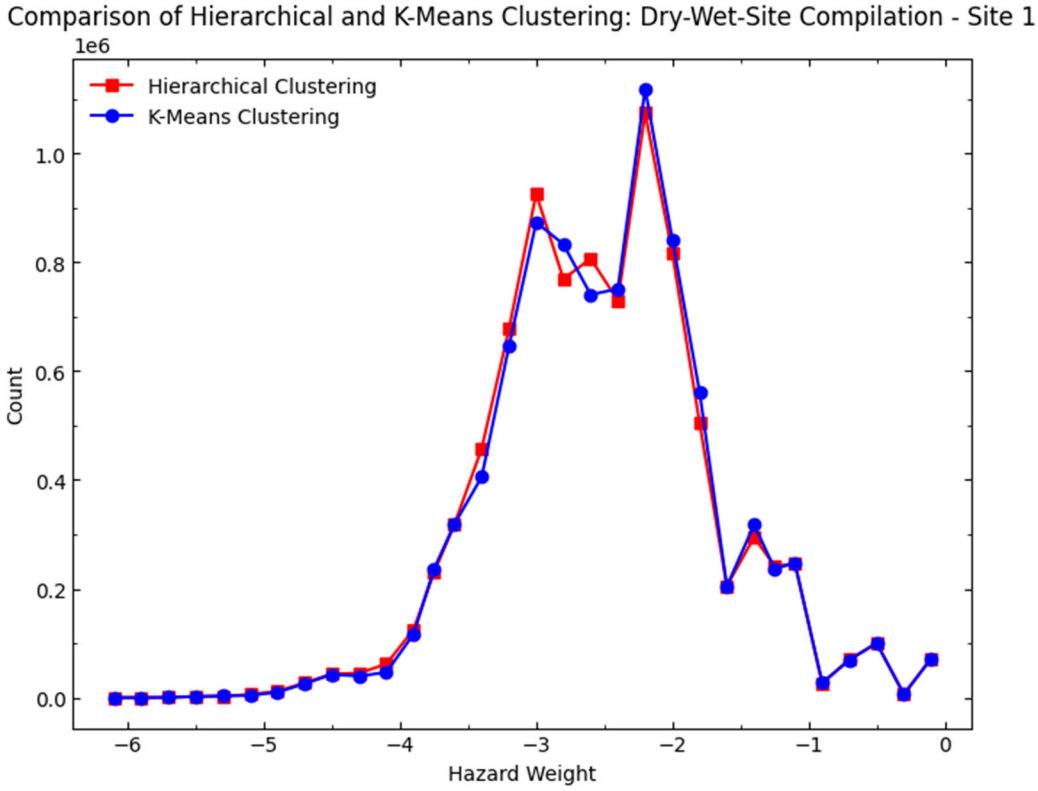

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hazard Map
Fit: Dry-Wet (Site) Compilation
Algorithm: K-Means
HazardWeight
-2.25 - -2.03
-2.02 - -1.80
-1.79 - -1.58
-1.57 - -1.35
-1.34 - -1.13
-1.12 - -0.90
-0.89 - -0.68
-0.67 - -0.45
-0.44 - -0.23
-0.22 - 0.00
0 10 20 40 60 80
Meters

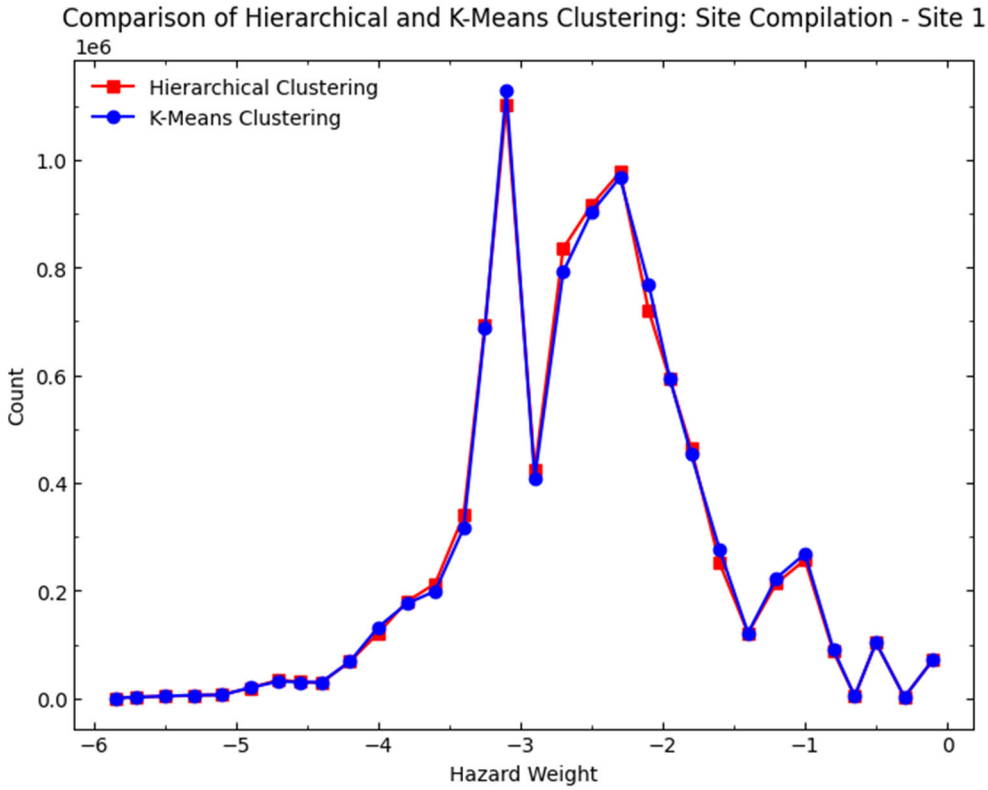
Comparison Between Hierarchical Clustering and K-Means Clustering

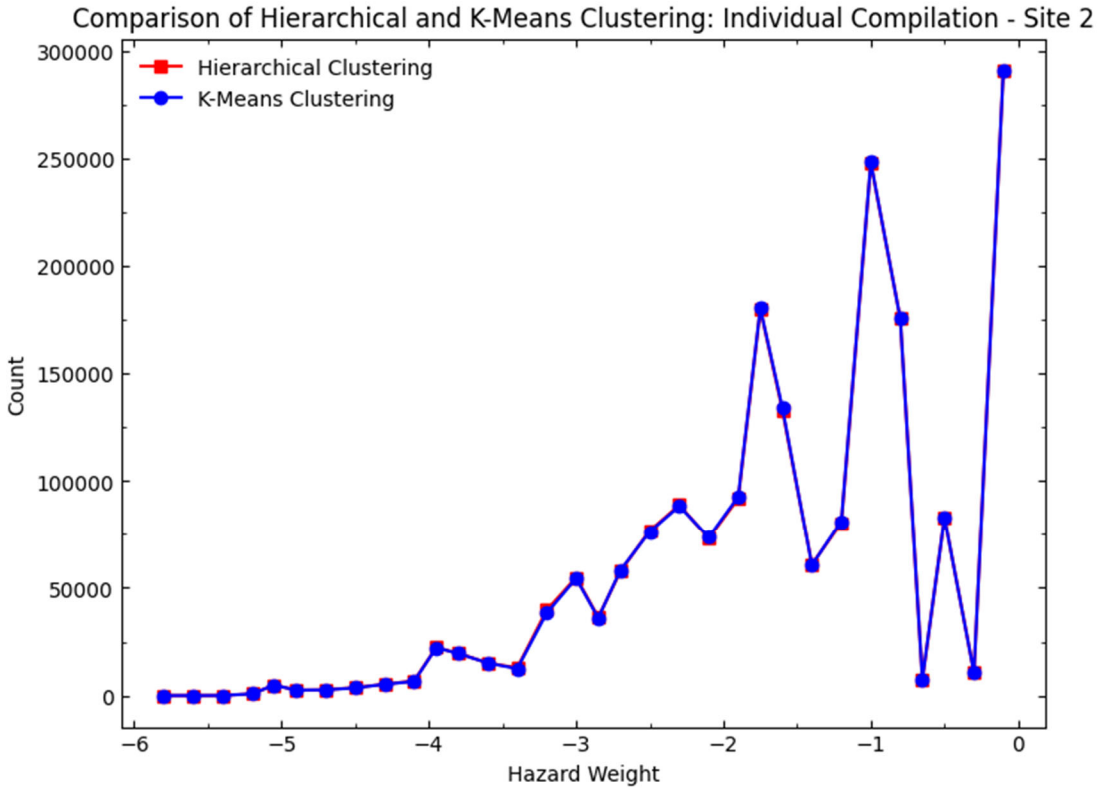


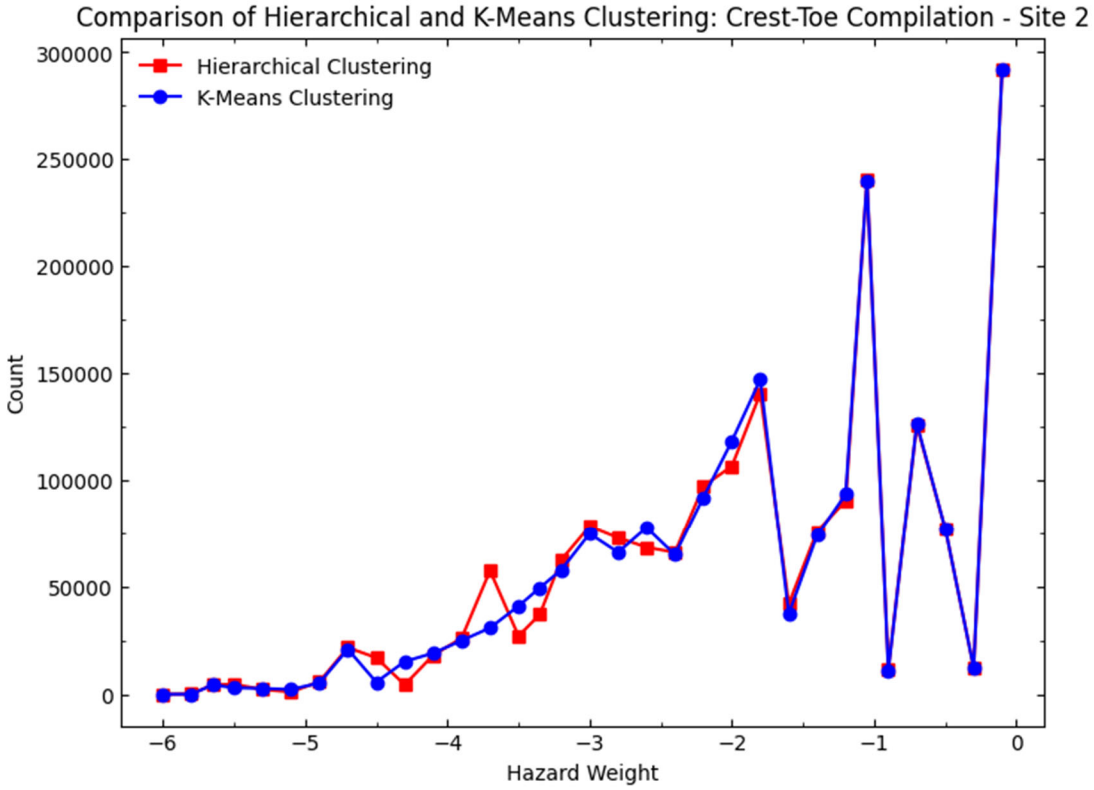


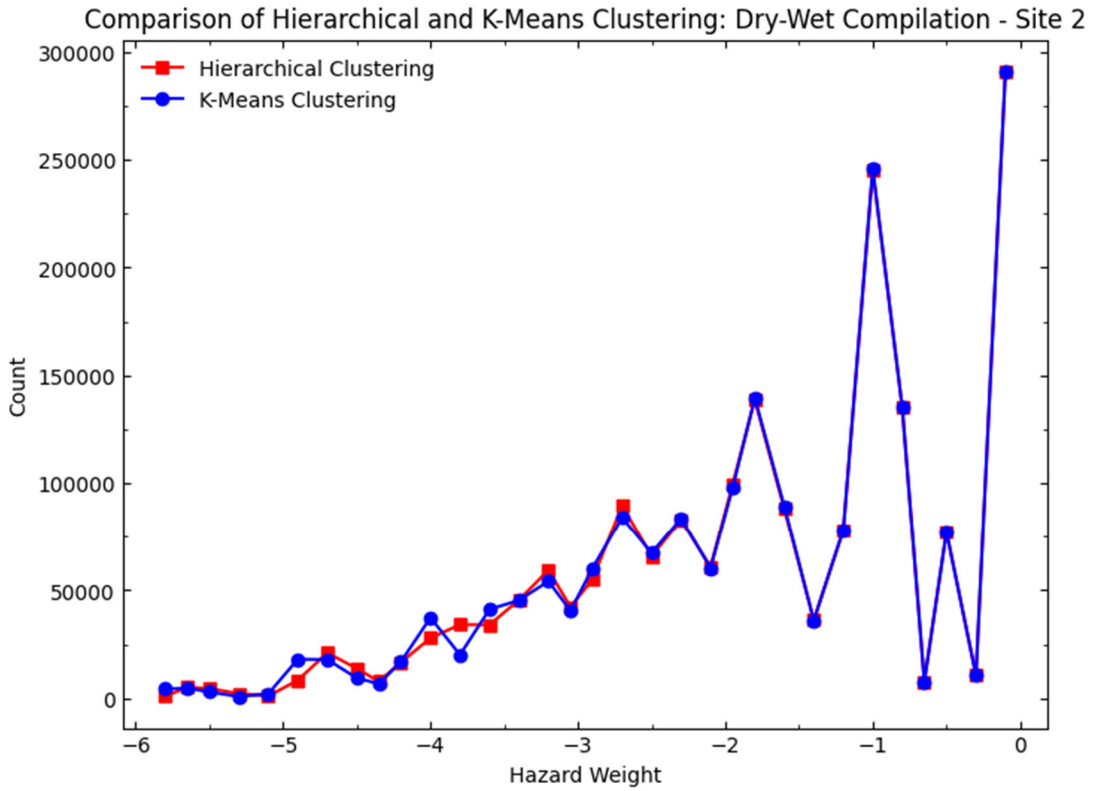

Comparison of Hierarchical and K-Means Clustering: Dry-Wet Compilation - Site 1
1e6
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight

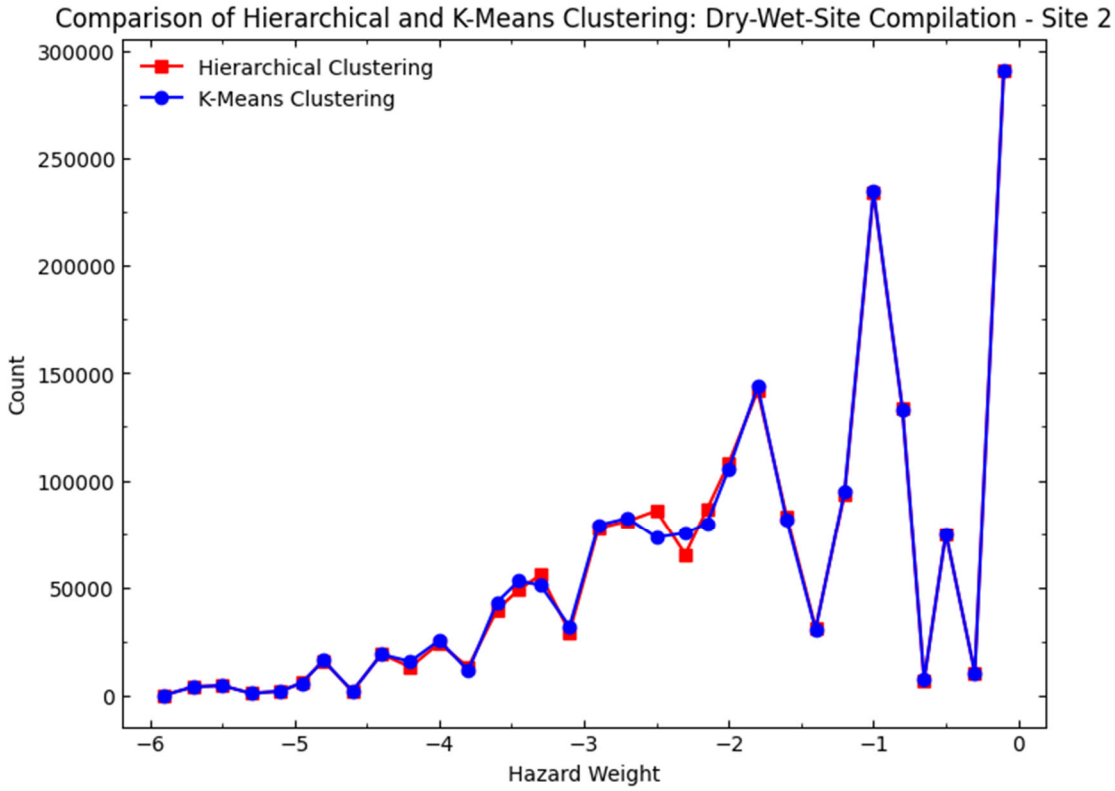

Comparison of Hierarchical and K-Means Clustering: Dry-Wet-Site Compilation - Site 1
1e6
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight

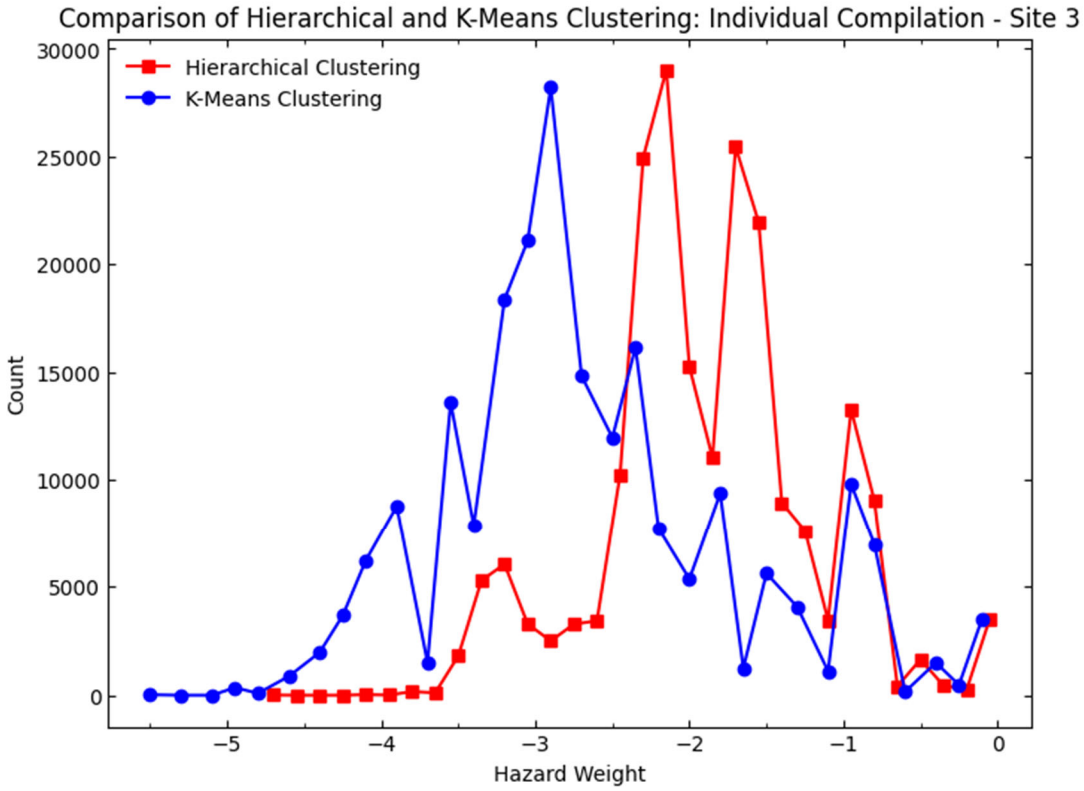

Comparison of Hierarchical and K-Means Clustering: Site Compilation - Site 1
1e6
Hierarchical Clustering
K-Means Clustering
1.0
0.8
0.6
0.4
0.2
0.0
Count
−6
−5
−4
−3
−2
−1
0
Hazard Weight

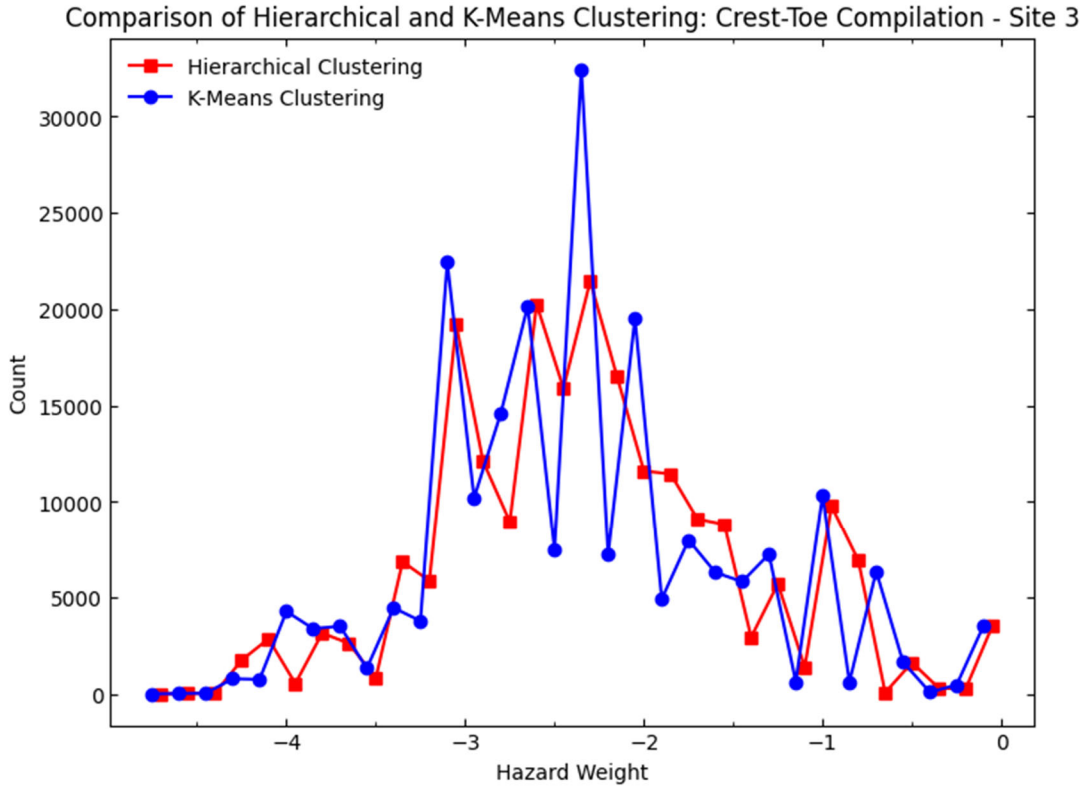

Comparison of Hierarchical and K-Means Clustering: Individual Compilation - Site 2
Hierarchical Clustering
K-Means Clustering
300000
250000
200000
150000
100000
50000
0
Count
−6
−5
−4
−3
−2
−1
0
Hazard Weight

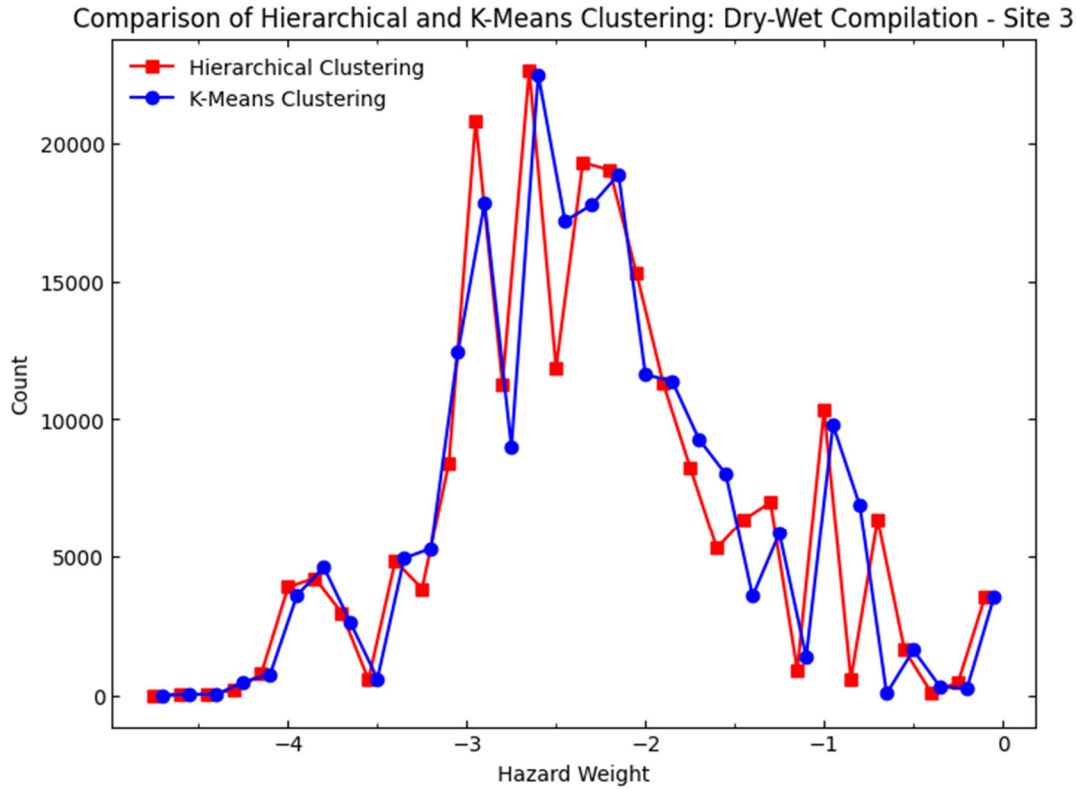

Comparison of Hierarchical and K-Means Clustering: Crest-Toe Compilation - Site 2
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
300000
250000
200000
150000
100000
50000
0
−6
−5
−4
−3
−2
−1
0

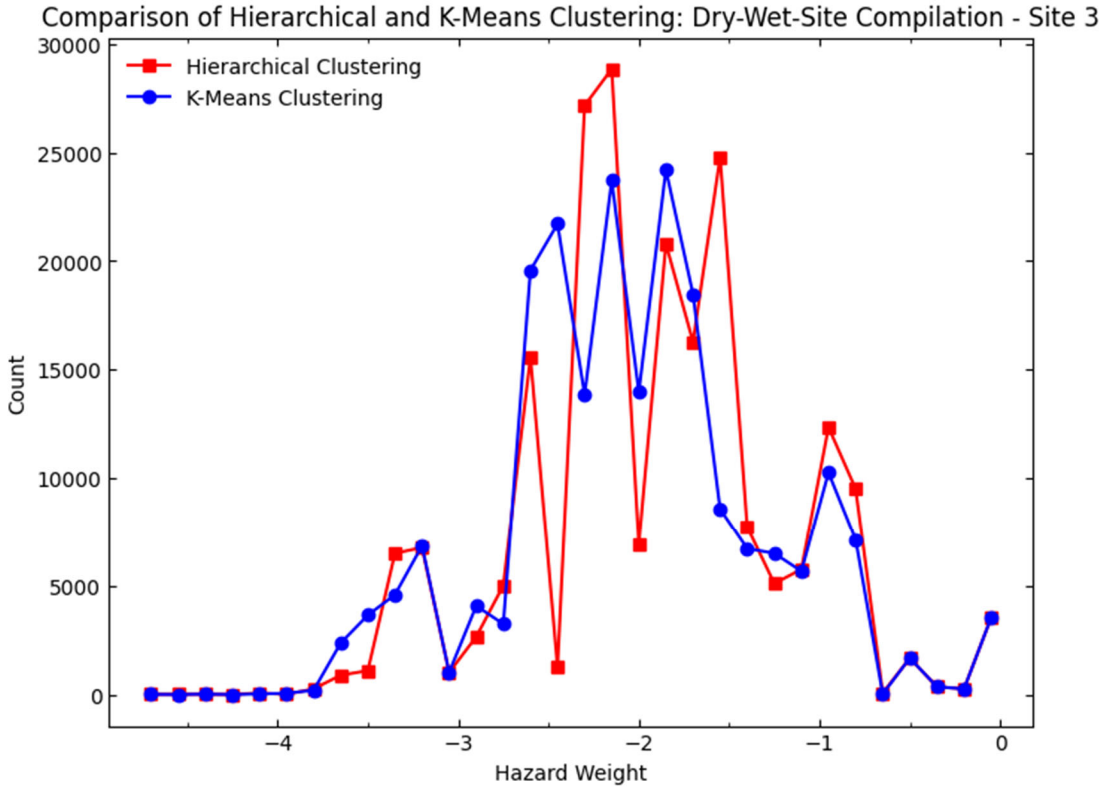

Comparison of Hierarchical and K-Means Clustering: Dry-Wet Compilation - Site 2
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
300000
250000
200000
150000
100000
50000
0
−6
−5
−4
−3
−2
−1
0

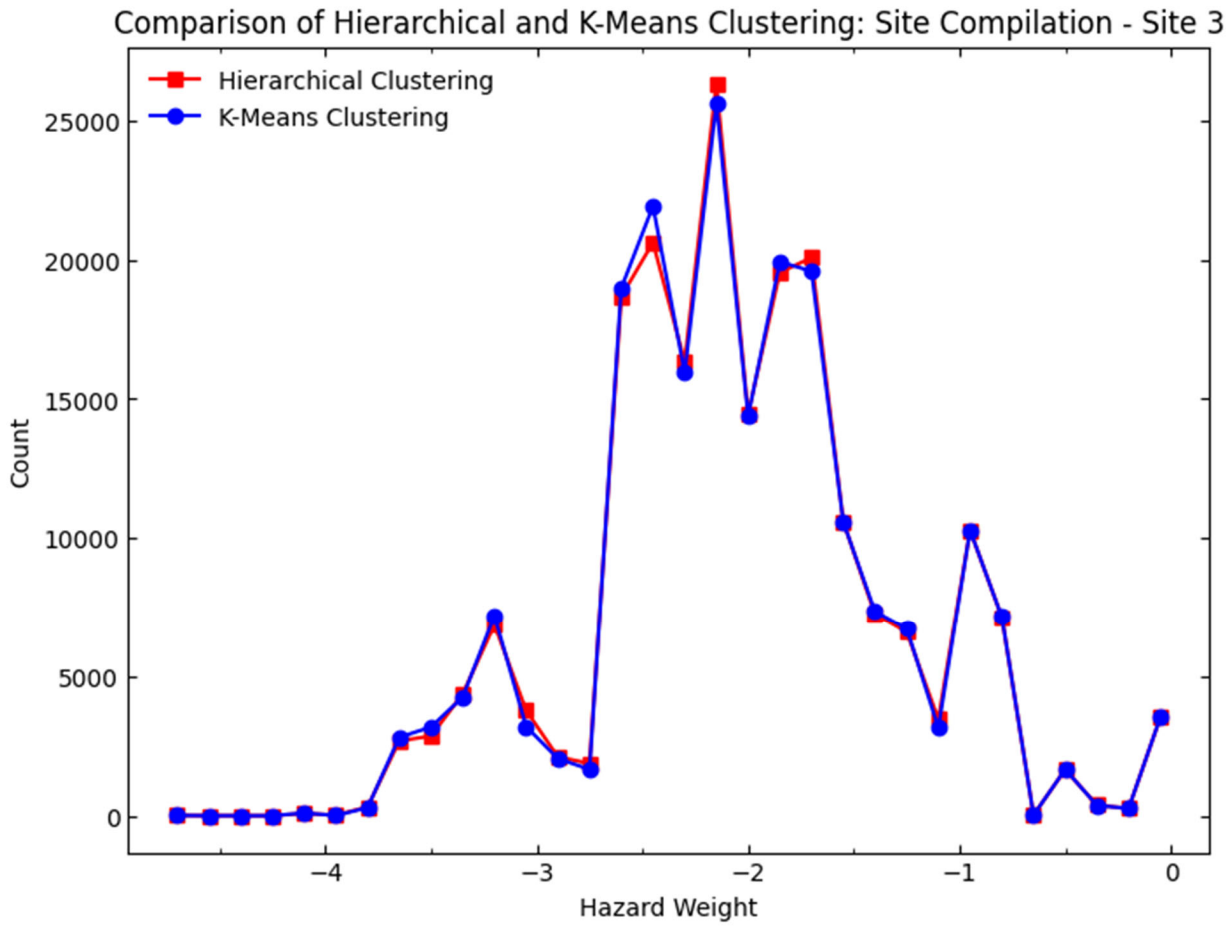

Comparison of Hierarchical and K-Means Clustering: Dry-Wet-Site Compilation - Site 2
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
300000
250000
200000
150000
100000
50000
0
−6
−5
−4
−3
−2
−1
0

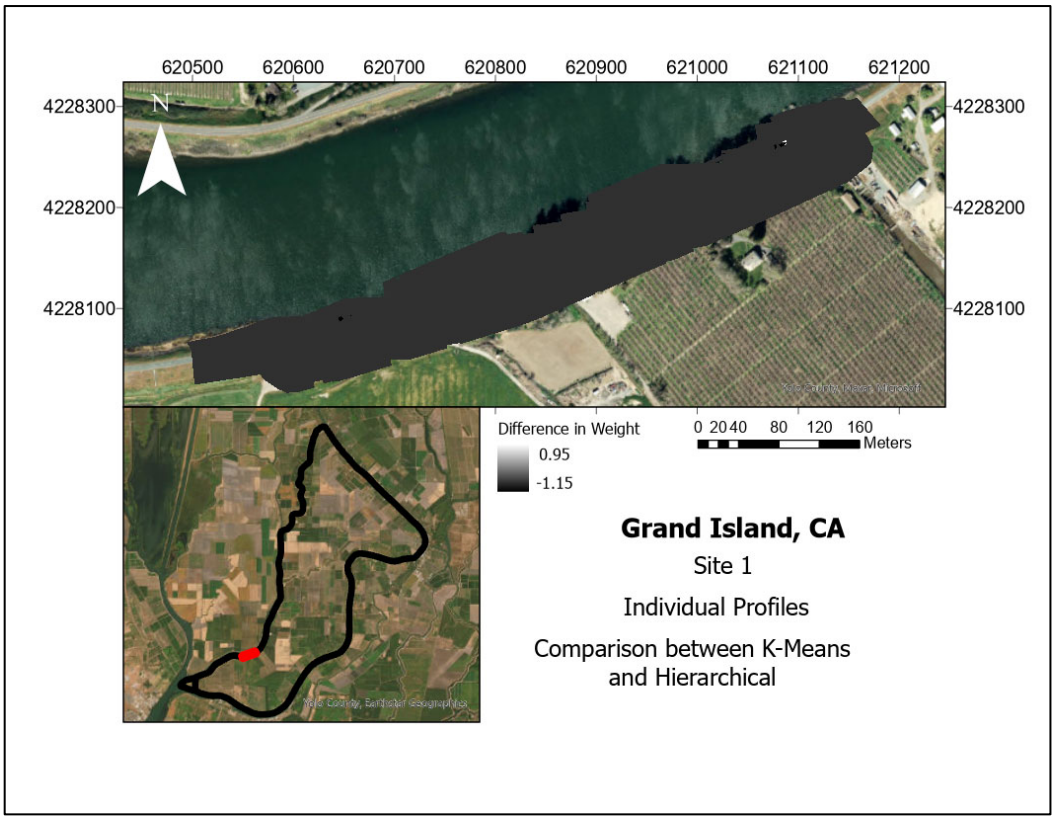

Comparison of Hierarchical and K-Means Clustering: Individual Compilation - Site 3
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
30000
25000
20000
15000
10000
5000
0
−5
−4
−3
−2
−1
0

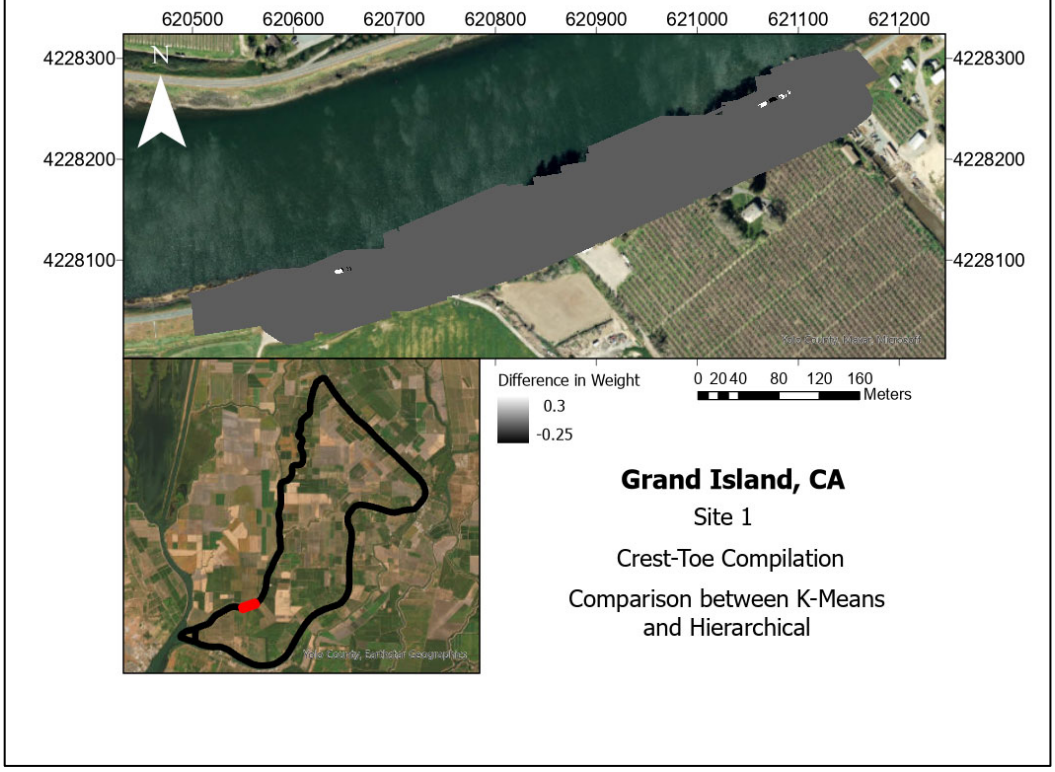

Comparison of Hierarchical and K-Means Clustering: Crest-Toe Compilation - Site 3
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
30000
25000
20000
15000
10000
5000
0
−4
−3
−2
−1
0


Comparison of Hierarchical and K-Means Clustering: Dry-Wet Compilation - Site 3
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
0
5000
10000
15000
20000
−4
−3
−2
−1
0

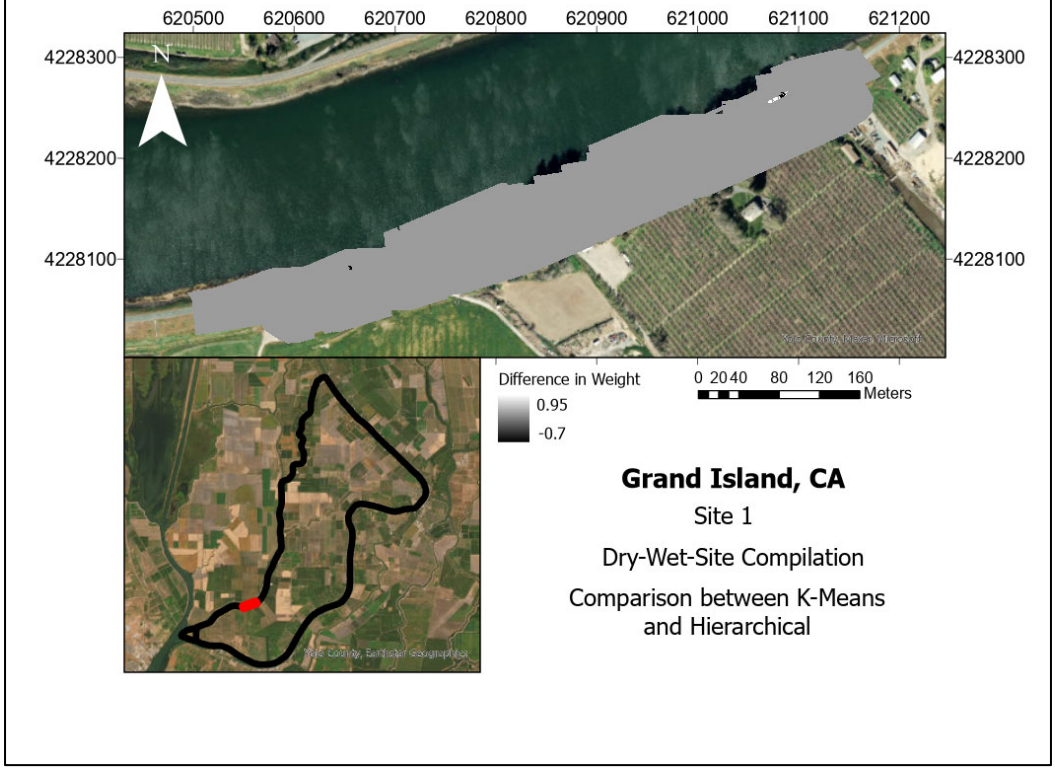

Comparison of Hierarchical and K-Means Clustering: Dry-Wet-Site Compilation - Site 3
Hierarchical Clustering
K-Means Clustering
Count
Hazard Weight
0
5000
10000
15000
20000
25000
30000
−4
−3
−2
−1
0


Comparison of Hierarchical and K-Means Clustering: Site Compilation - Site 3
Hierarchical Clustering
K-Means Clustering
Count
25000
20000
15000
10000
5000
0
−4
−3
−2
−1
0
Hazard Weight

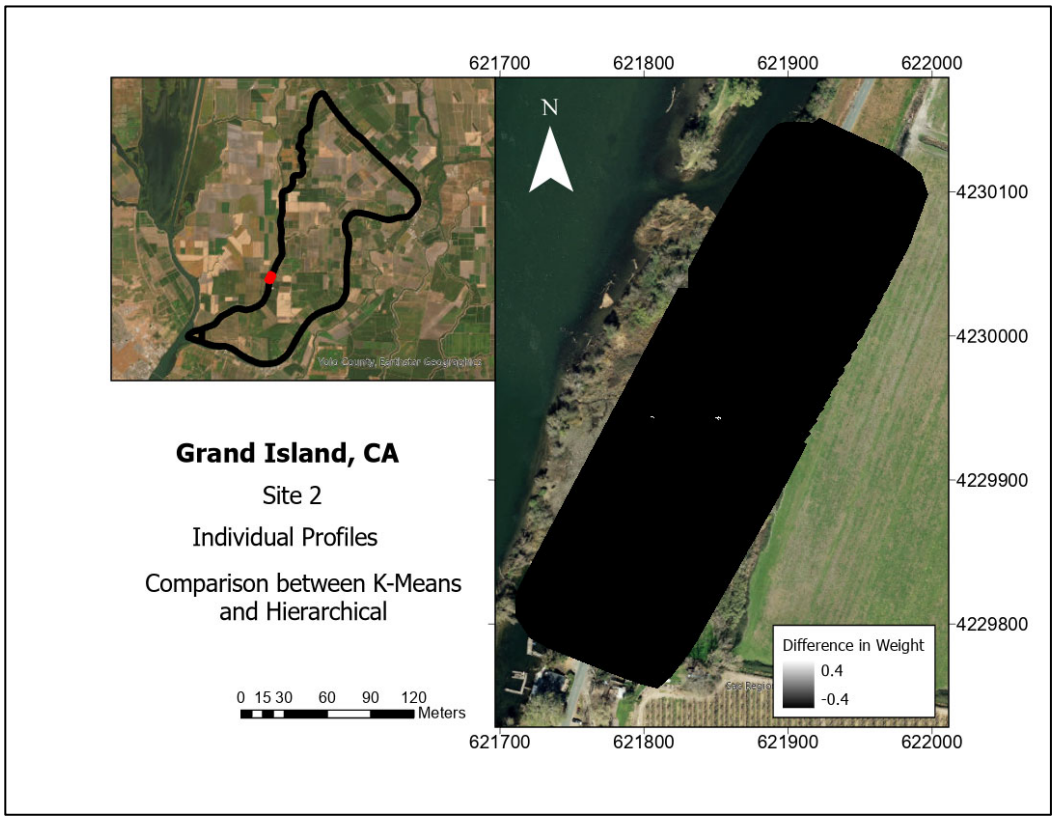

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
Difference in Weight
0.95
-1.15
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Individual Profiles
Comparison between K-Means
and Hierarchical

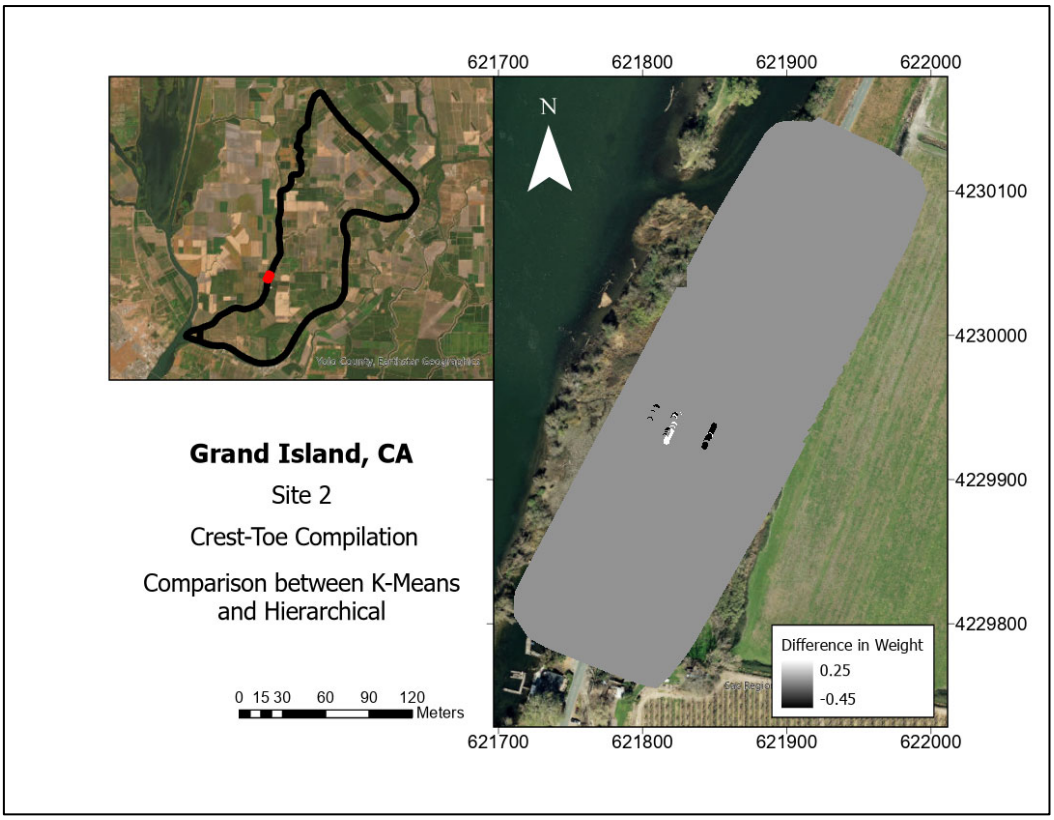

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
Difference in Weight
0.3
-0.25
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Crest-Toe Compilation
Comparison between K-Means
and Hierarchical

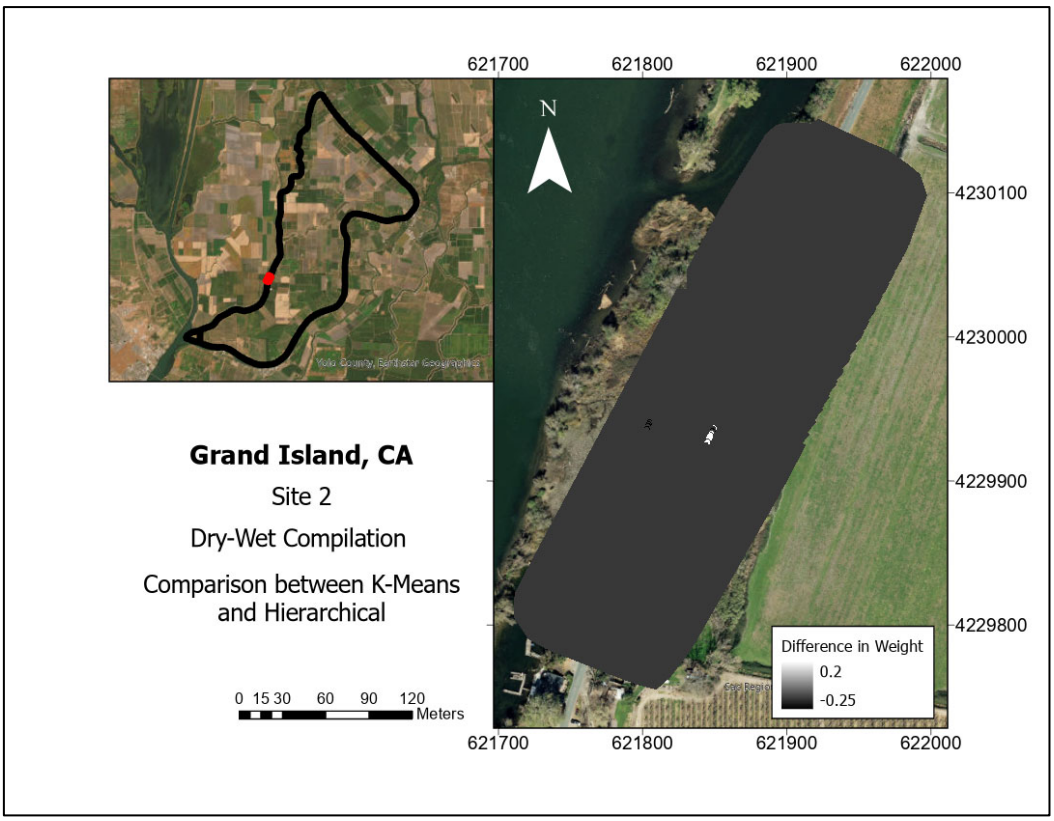

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
Difference in Weight
0.95
-0.7
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Dry-Wet Compilation
Comparison between K-Means
and Hierarchical

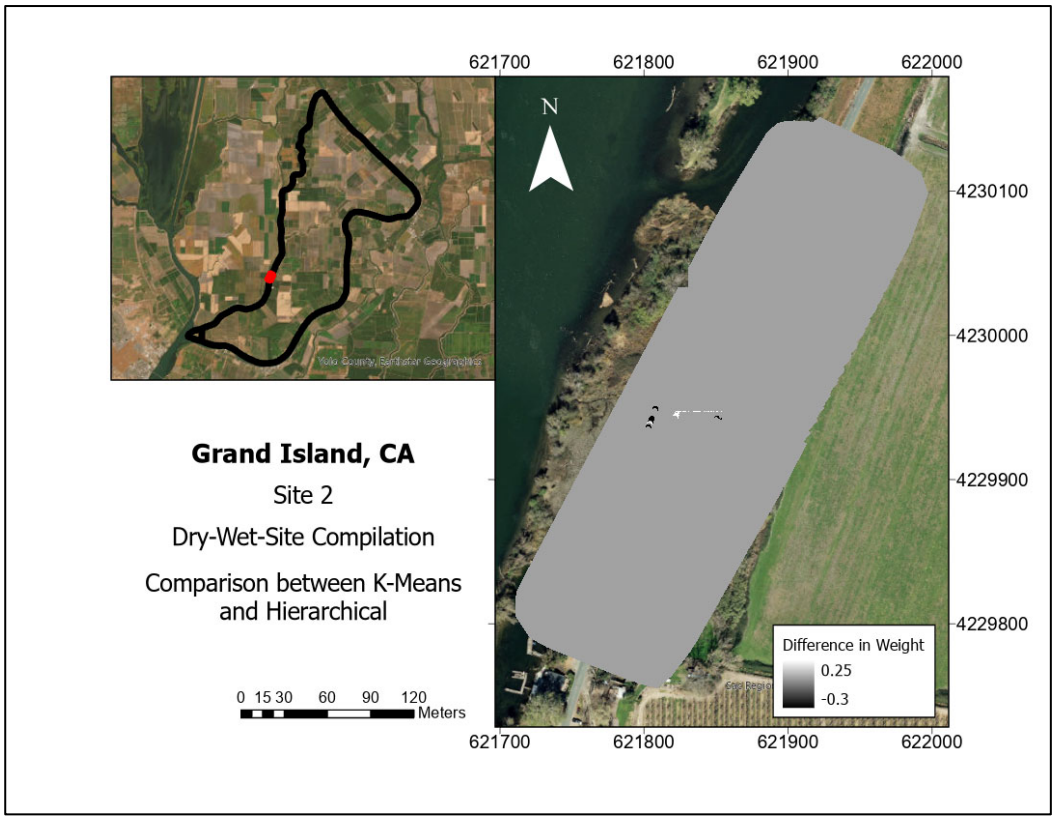

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
Difference in Weight
0.95
-0.7
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Dry-Wet-Site Compilation
Comparison between K-Means
and Hierarchical

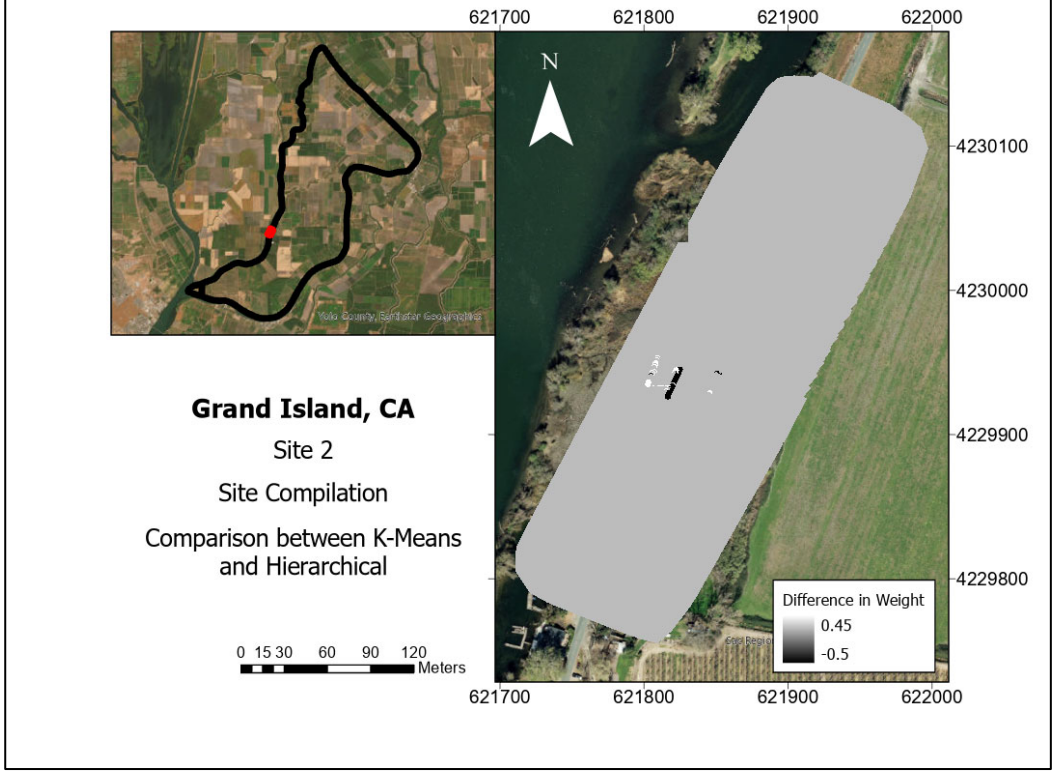

620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
N
Difference in Weight
0.95
-0.75
0 20 40 80 120 160
Meters
Grand Island, CA
Site 1
Site Compilation
Comparison between K-Means and Hierarchical

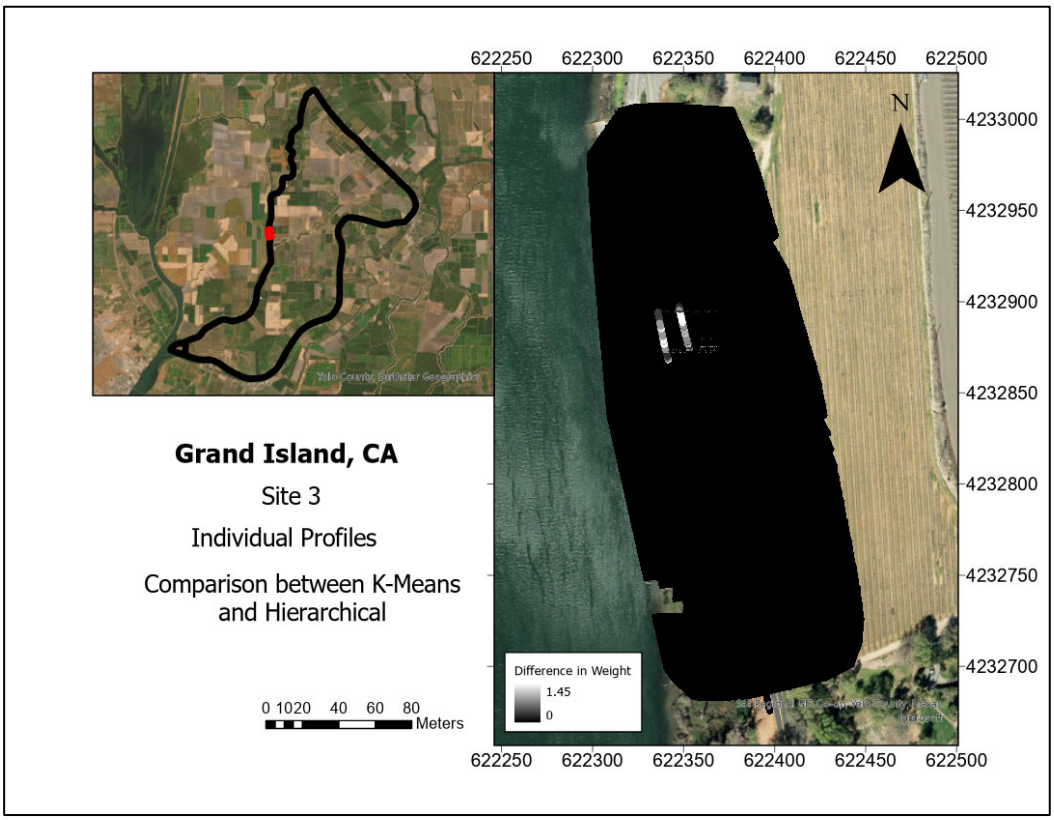

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
Grand Island, CA
Site 2
Individual Profiles
Comparison between K-Means and Hierarchical
0 15 30 60 90 120
Meters
Difference in Weight
0.4
-0.4

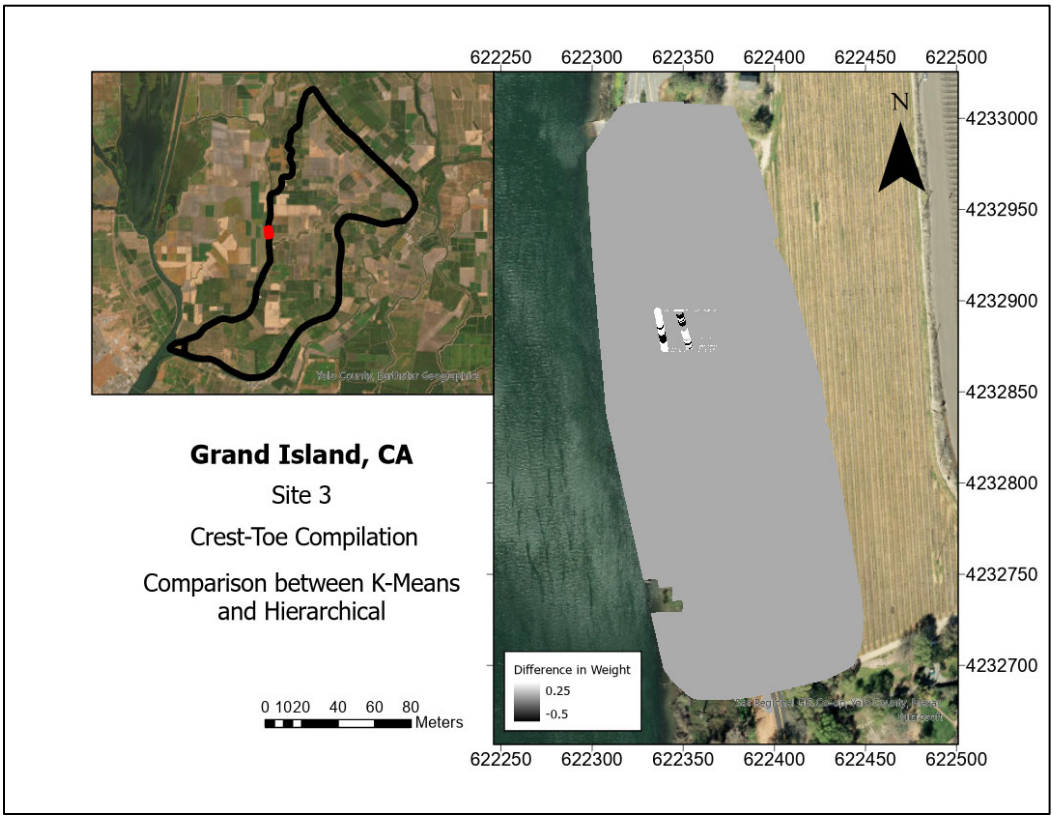

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
Grand Island, CA
Site 2
Crest-Toe Compilation
Comparison between K-Means
and Hierarchical
0 15 30 60 90 120
Meters
Difference in Weight
0.25
-0.45

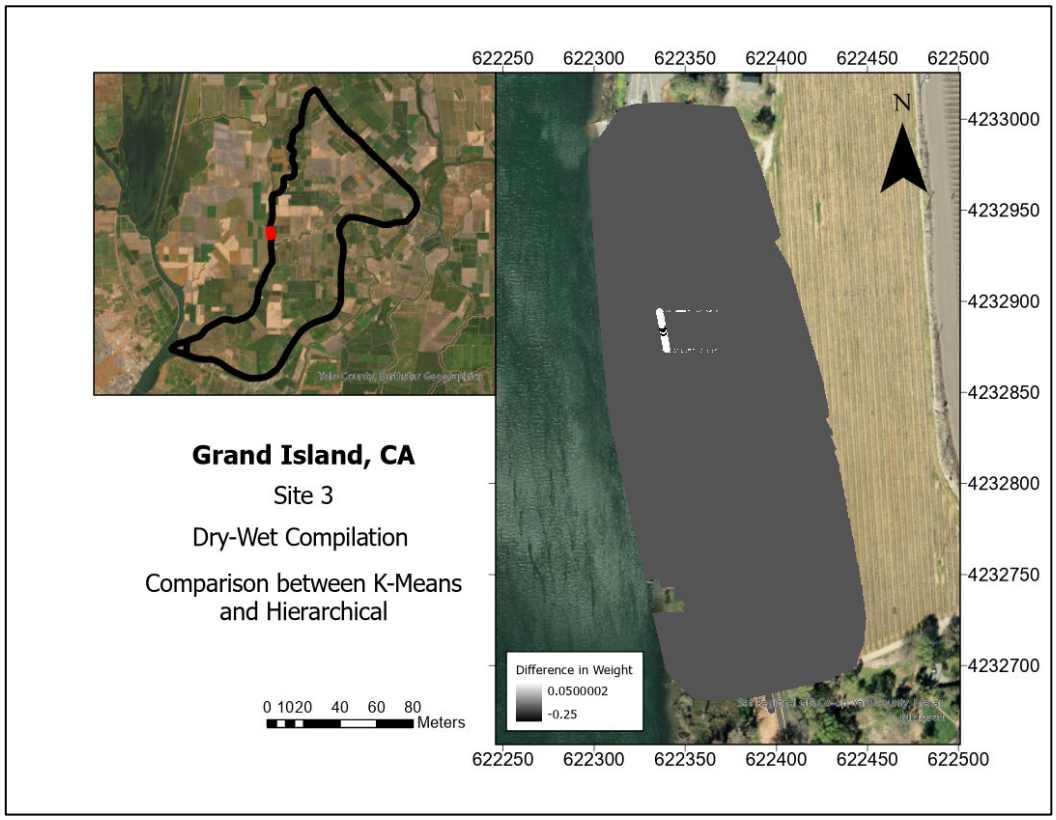

621700
621800
621900
622000
4230100
4230000
4229900
4229800
N
Grand Island, CA
Site 2
Dry-Wet Compilation
Comparison between K-Means
and Hierarchical
0 15 30 60 90 120
Meters
Difference in Weight
0.2
-0.25

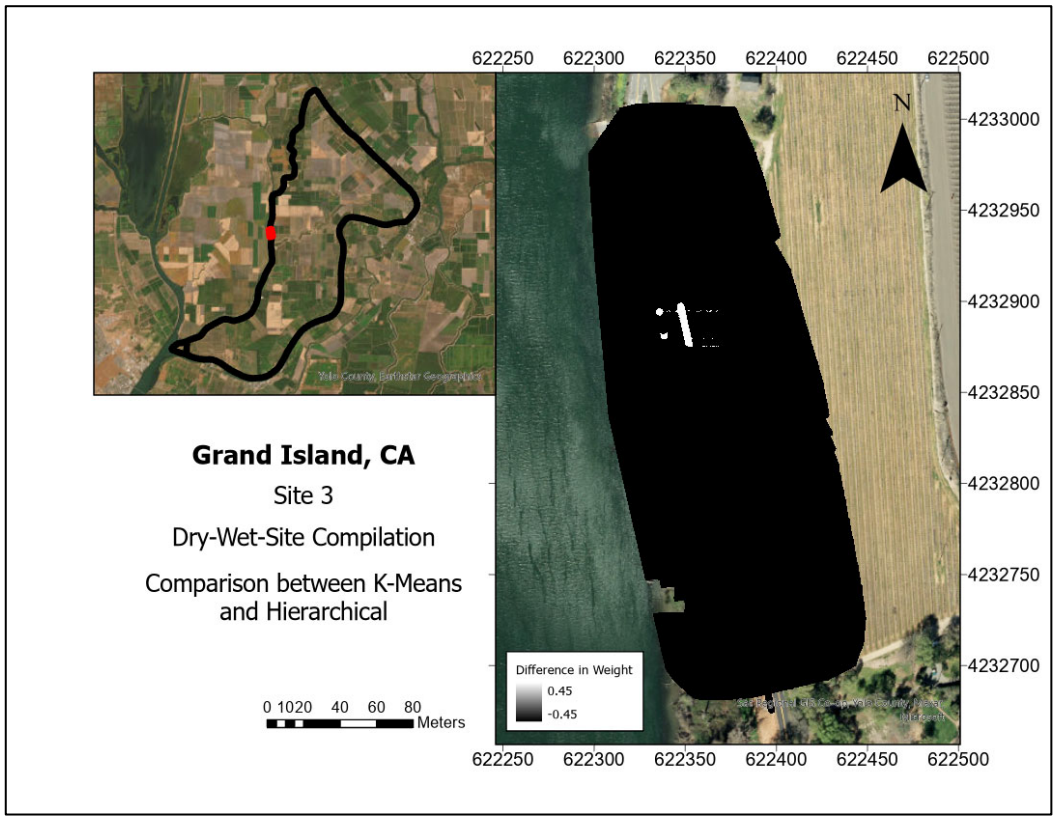

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
Dry-Wet-Site Compilation
Comparison between K-Means
and Hierarchical
0 15 30 60 90 120
Meters
Difference in Weight
0.25
-0.3
621700
621800
621900
622000

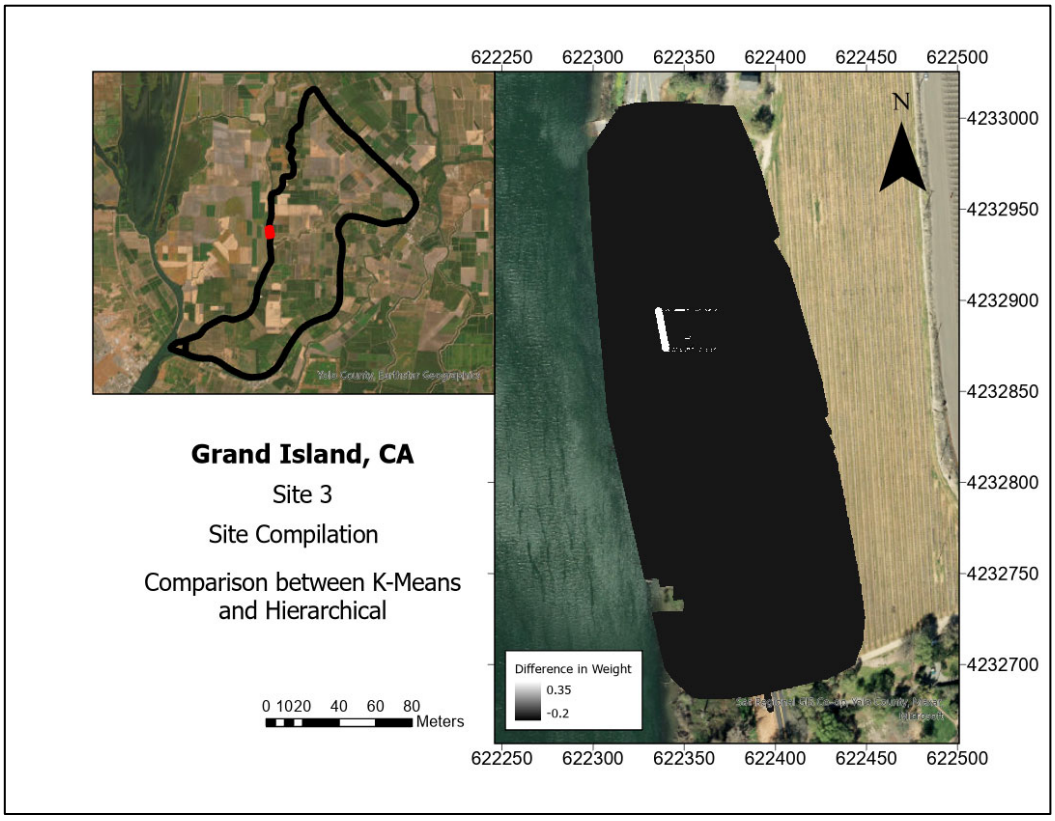

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
Site Compilation
Comparison between K-Means
and Hierarchical
0 15 30 60 90 120
Meters
Difference in Weight
0.45
-0.5
621700
621800
621900
622000


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Individual Profiles
Comparison between K-Means and Hierarchical
0 1020 40 60 80
Meters
Difference in Weight
1.45
0

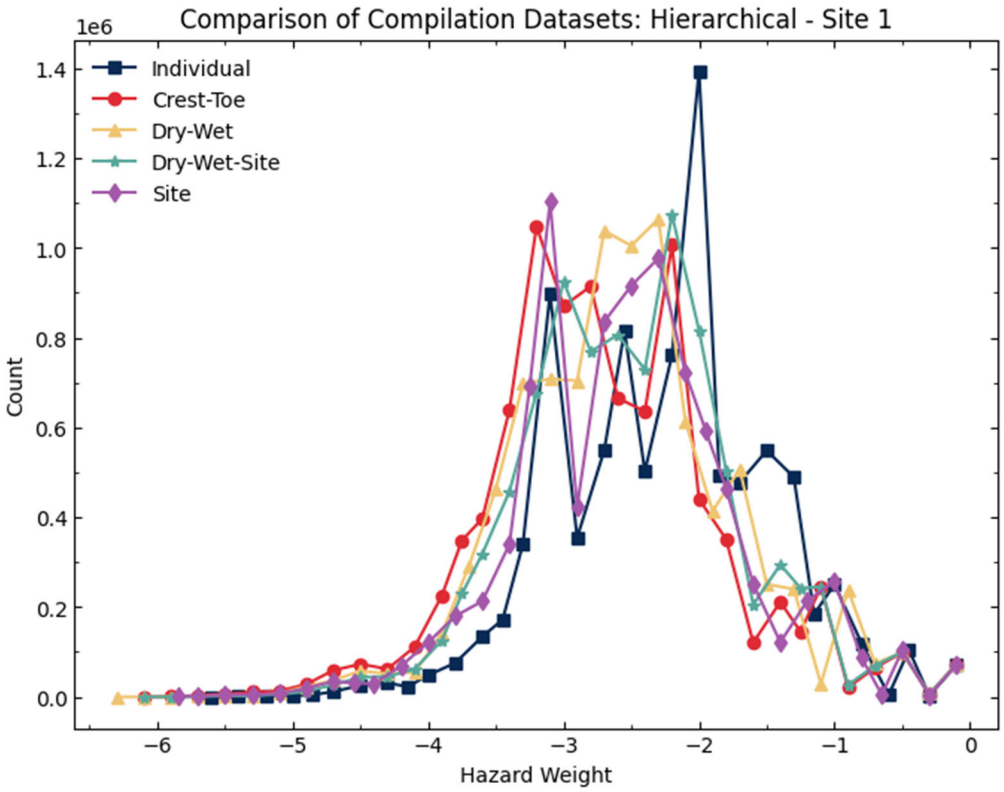

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Crest-Toe Compilation
Comparison between K-Means and Hierarchical
0 1020 40 60 80
Meters
Difference in Weight
0.25
-0.5

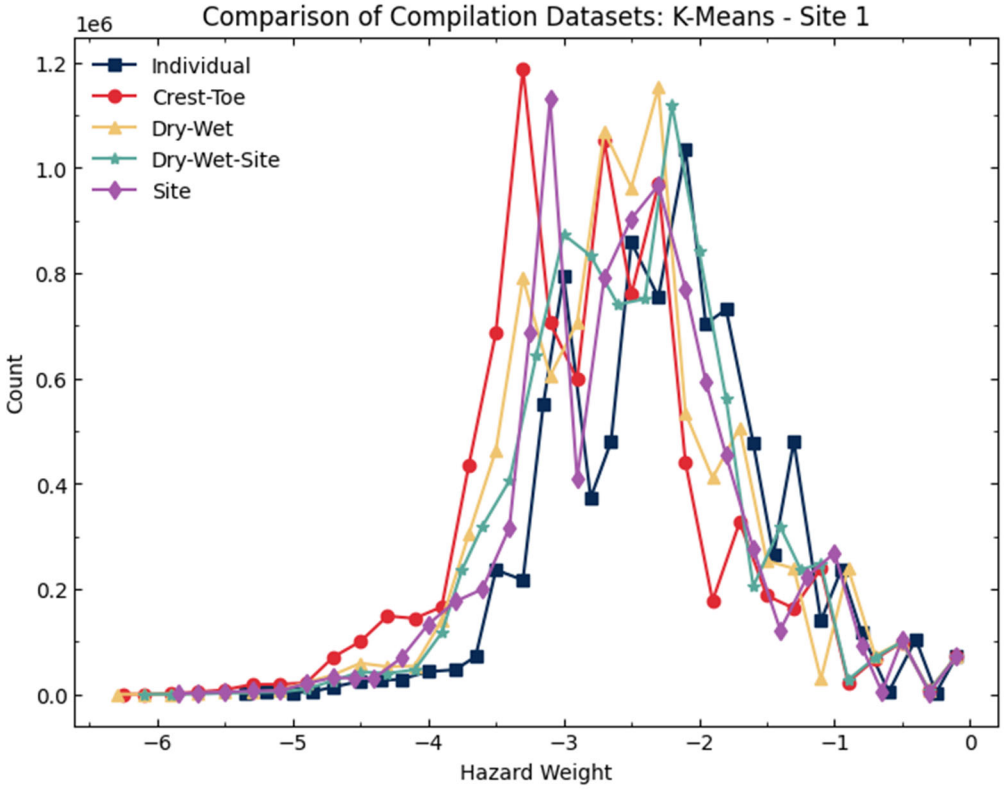

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Dry-Wet Compilation
Comparison between K-Means and Hierarchical
0 1020 40 60 80
Meters
Difference in Weight
0.0500002
-0.25

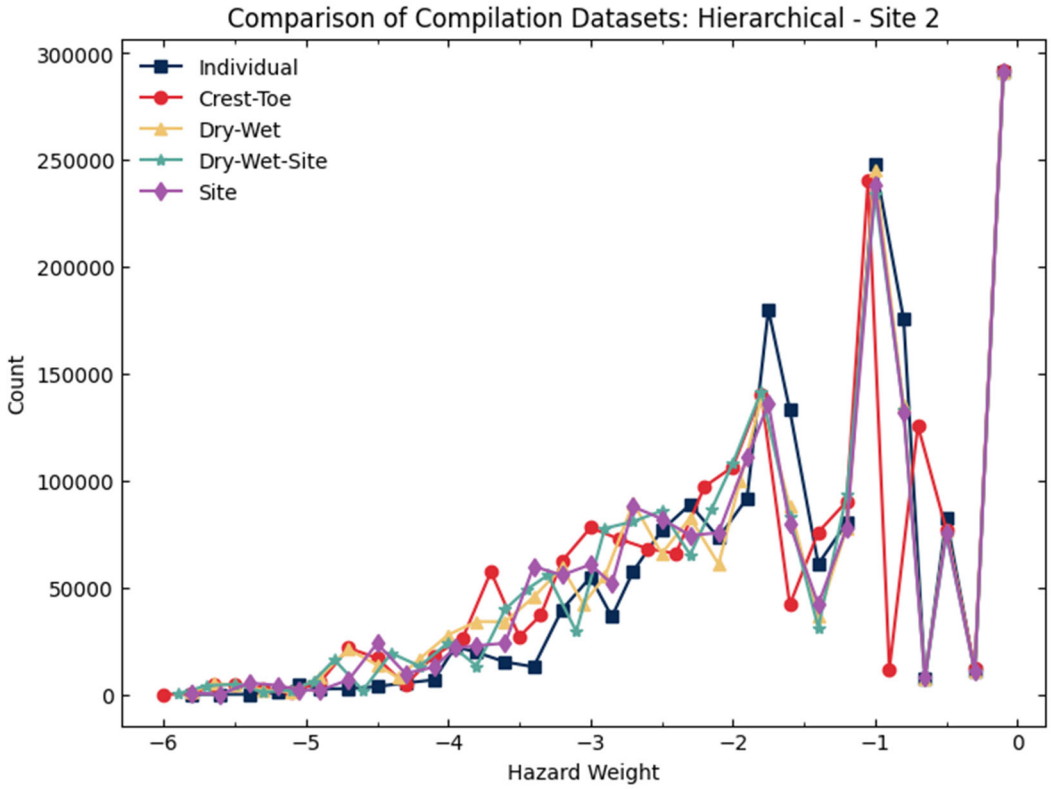

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Dry-Wet-Site Compilation
Comparison between K-Means and Hierarchical
0 1020 40 60 80
Meters
Difference in Weight
0.45
-0.45

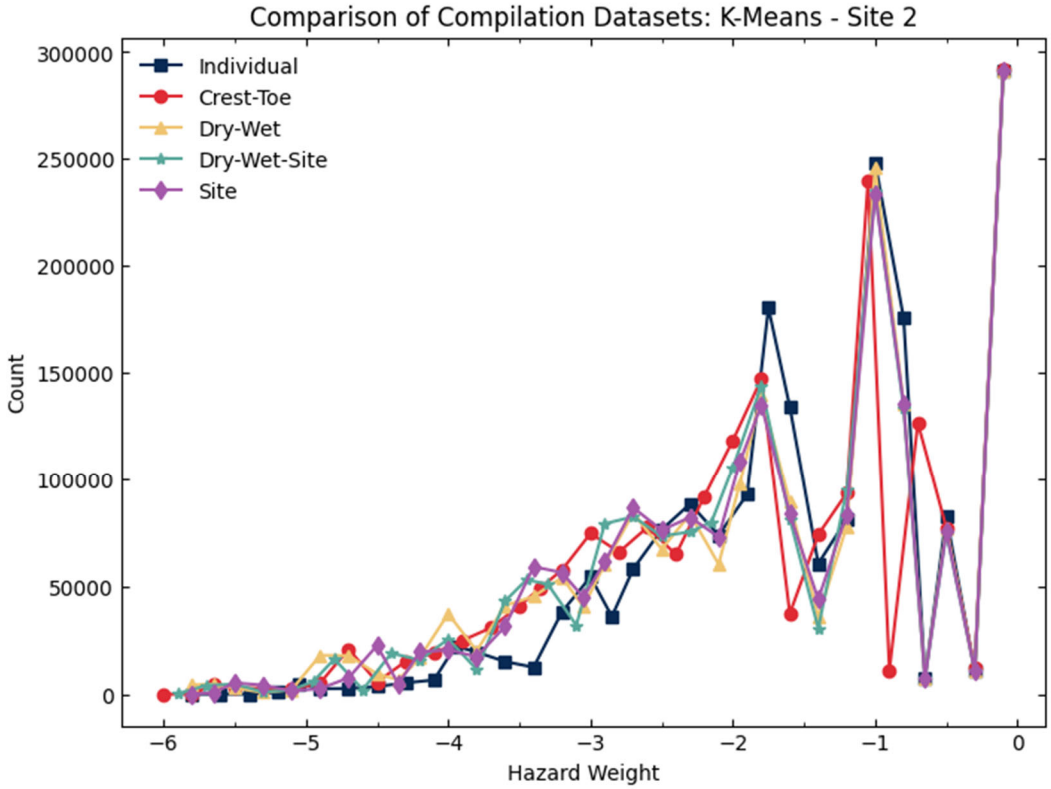

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Site Compilation
Comparison between K-Means and Hierarchical
0 1020 40 60 80
Meters
Difference in Weight
0.35
-0.2

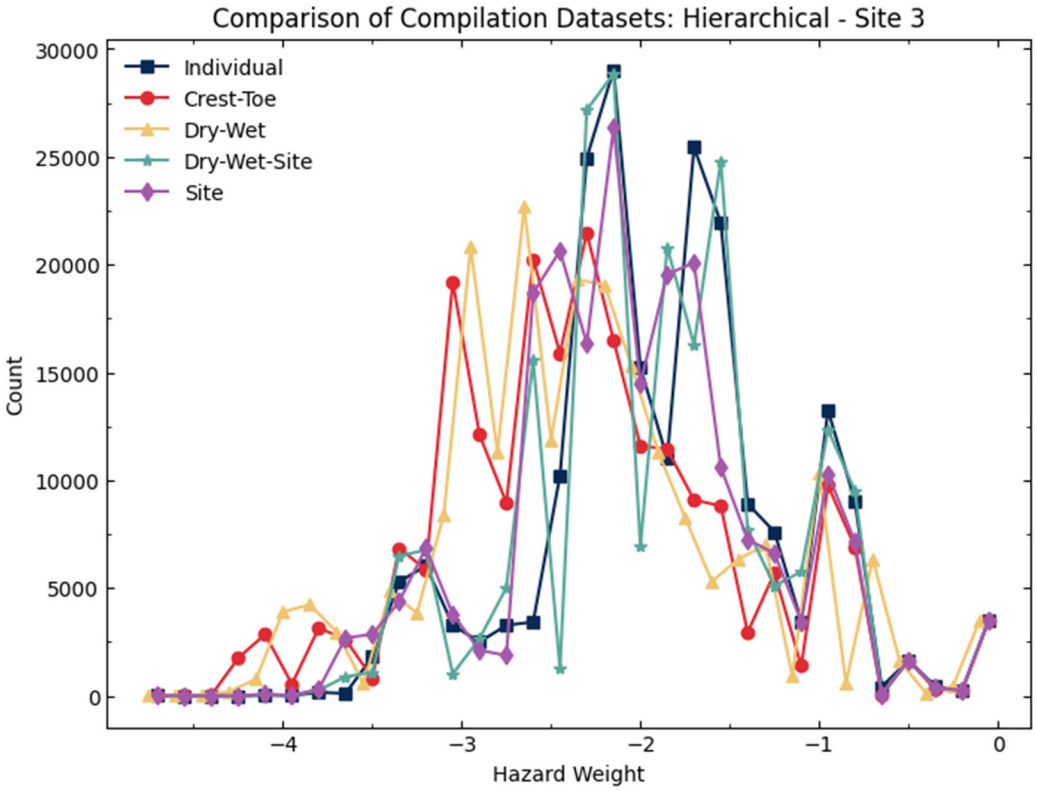
Comparison Between Compilation Datasets

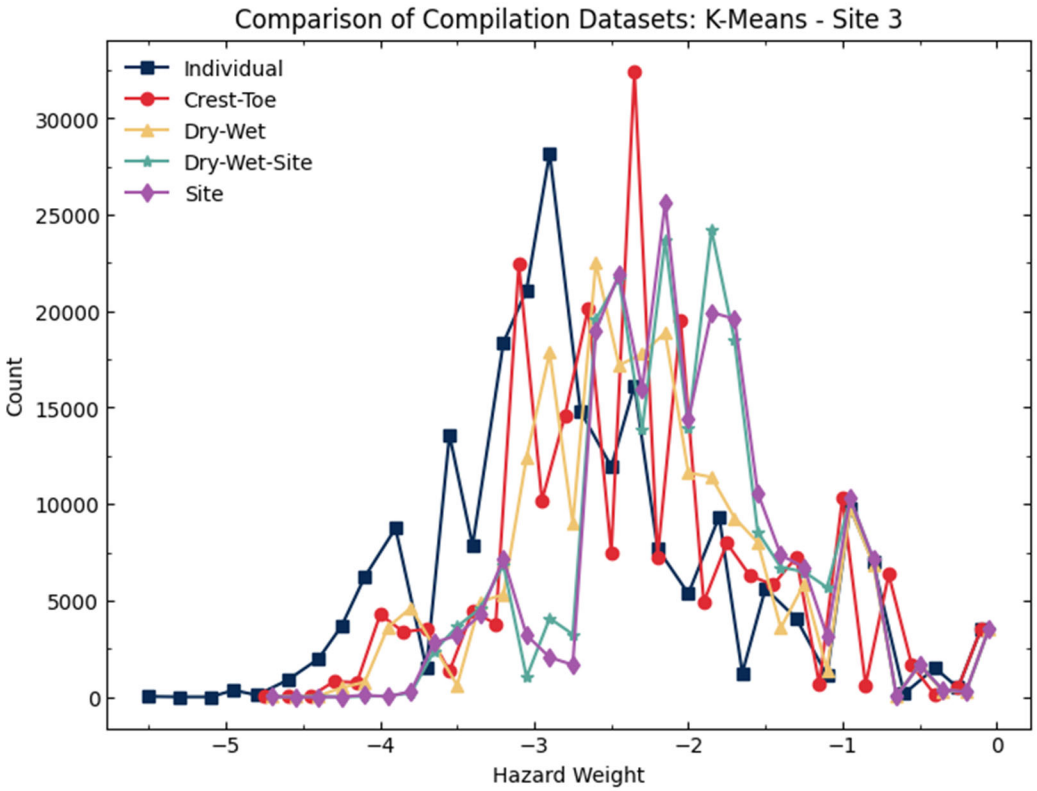


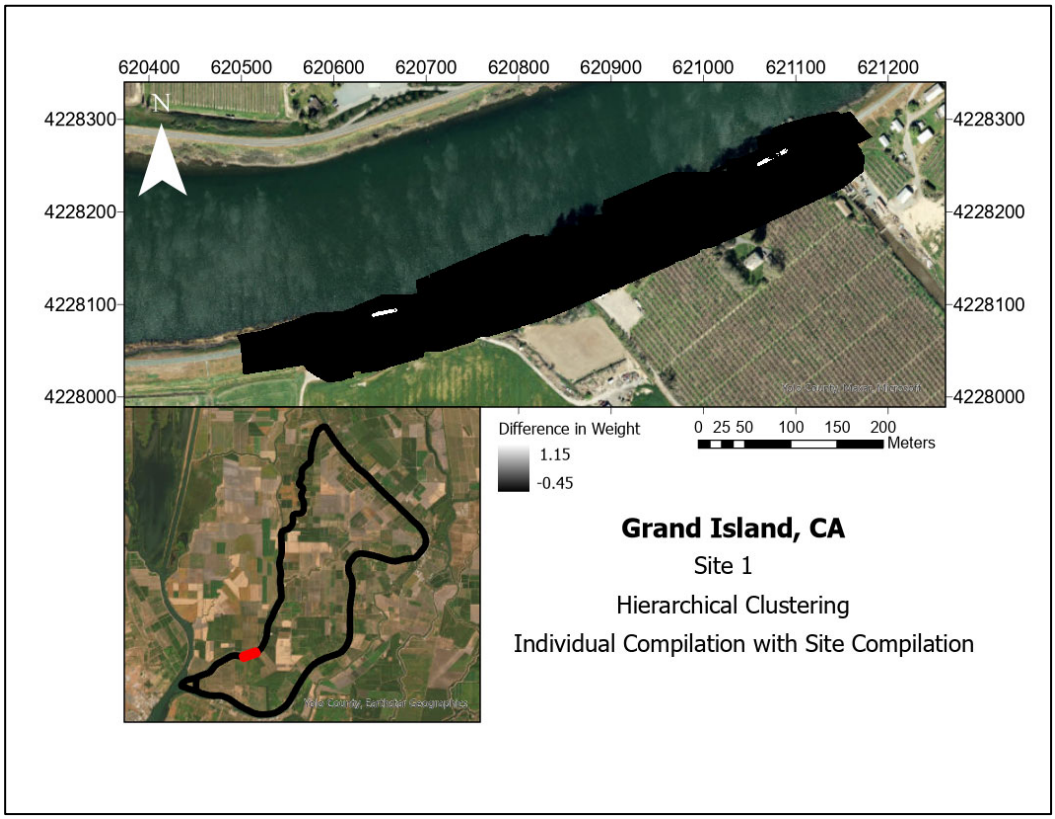


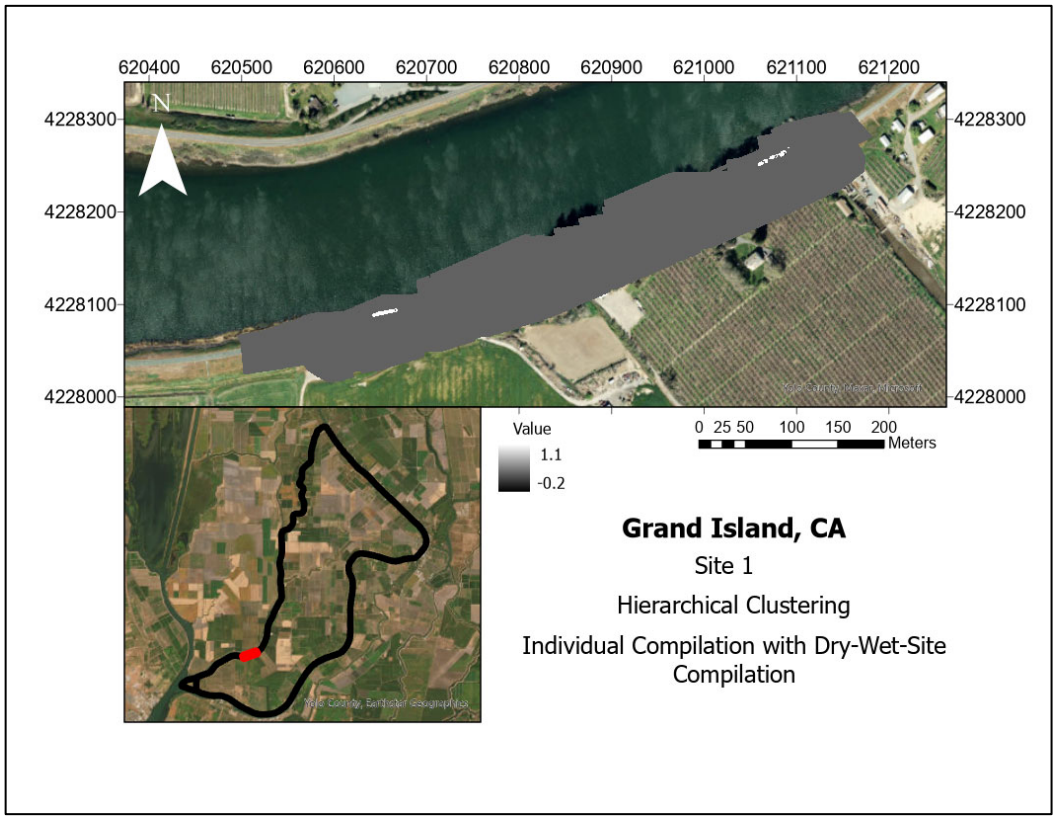

Comparison of Compilation Datasets: Hierarchical - Site 2
Individual
Crest-Toe
Dry-Wet
Dry-Wet-Site
Site
Count
Hazard Weight
300000
250000
200000
150000
100000
50000
0
−6
−5
−4
−3
−2
−1
0

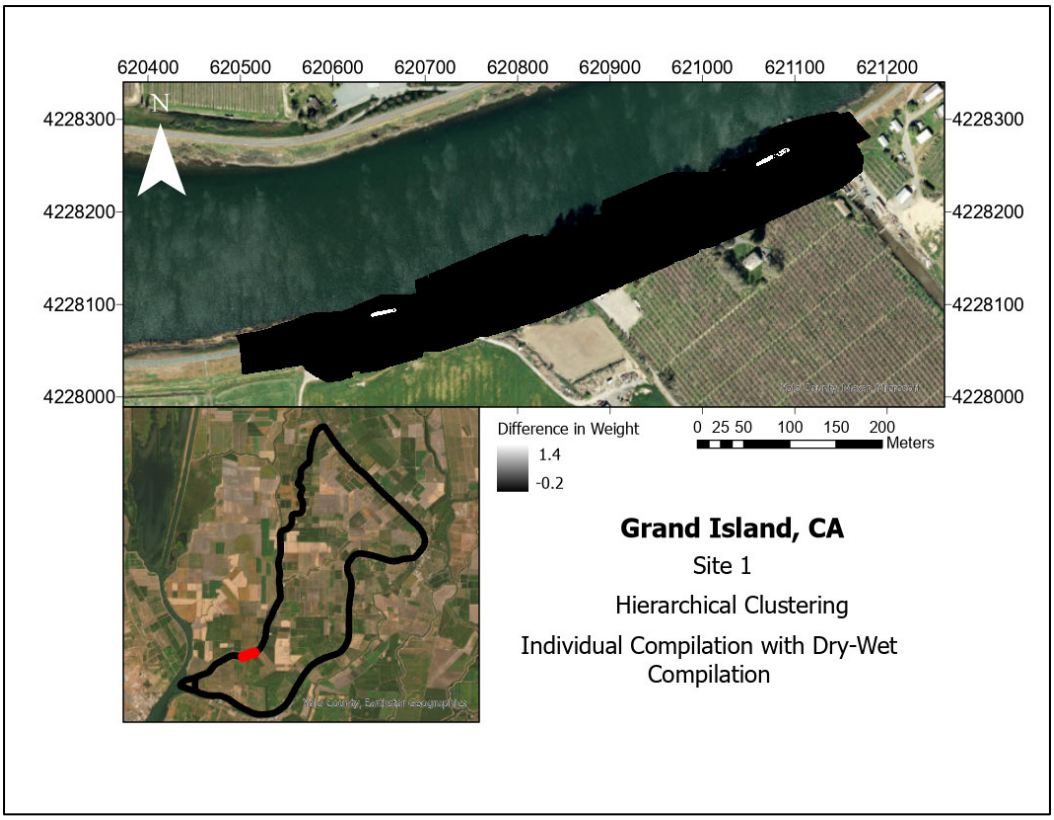

Comparison of Compilation Datasets: K-Means - Site 2
Individual
Crest-Toe
Dry-Wet
Dry-Wet-Site
Site
Count
Hazard Weight
300000
250000
200000
150000
100000
50000
0
−6
−5
−4
−3
−2
−1
0

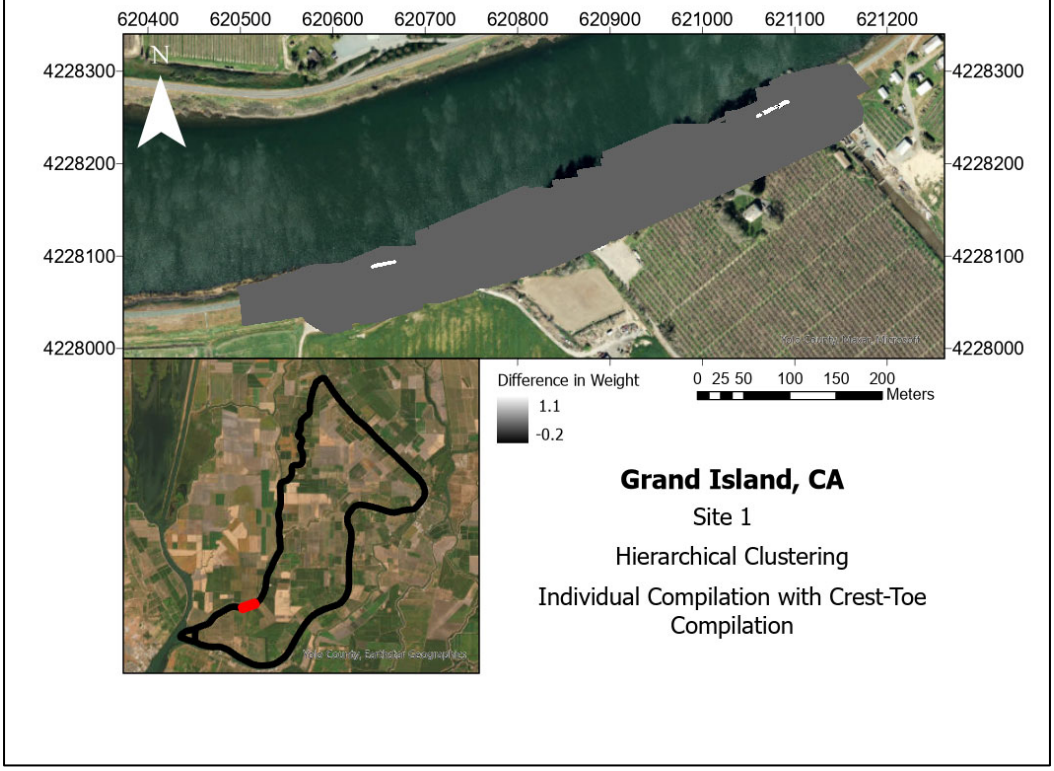

Comparison of Compilation Datasets: Hierarchical - Site 3
Individual
Crest-Toe
Dry-Wet
Dry-Wet-Site
Site
Count
Hazard Weight


Comparison of Compilation Datasets: K-Means - Site 3
Individual
Crest-Toe
Dry-Wet
Dry-Wet-Site
Site
Count
Hazard Weight

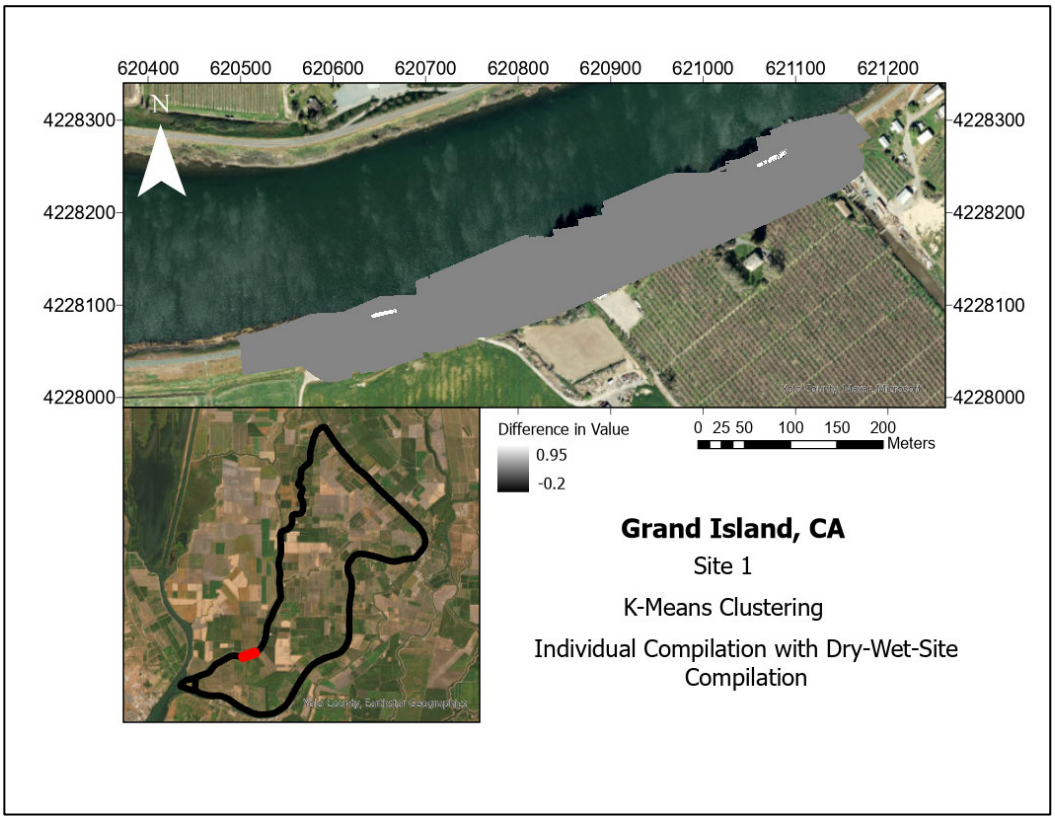

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Weight
1.15
-0.45
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
Hierarchical Clustering
Individual Compilation with Site Compilation

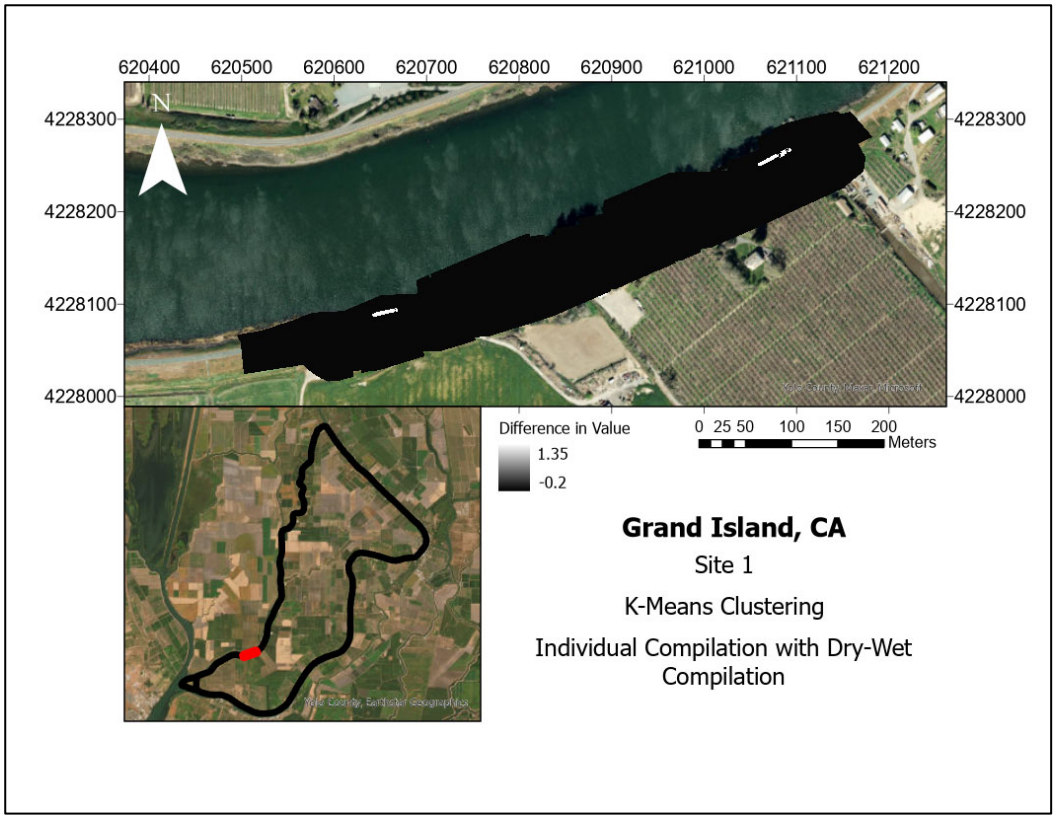

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Value
1.1
-0.2
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
Hierarchical Clustering
Individual Compilation with Dry-Wet-Site Compilation

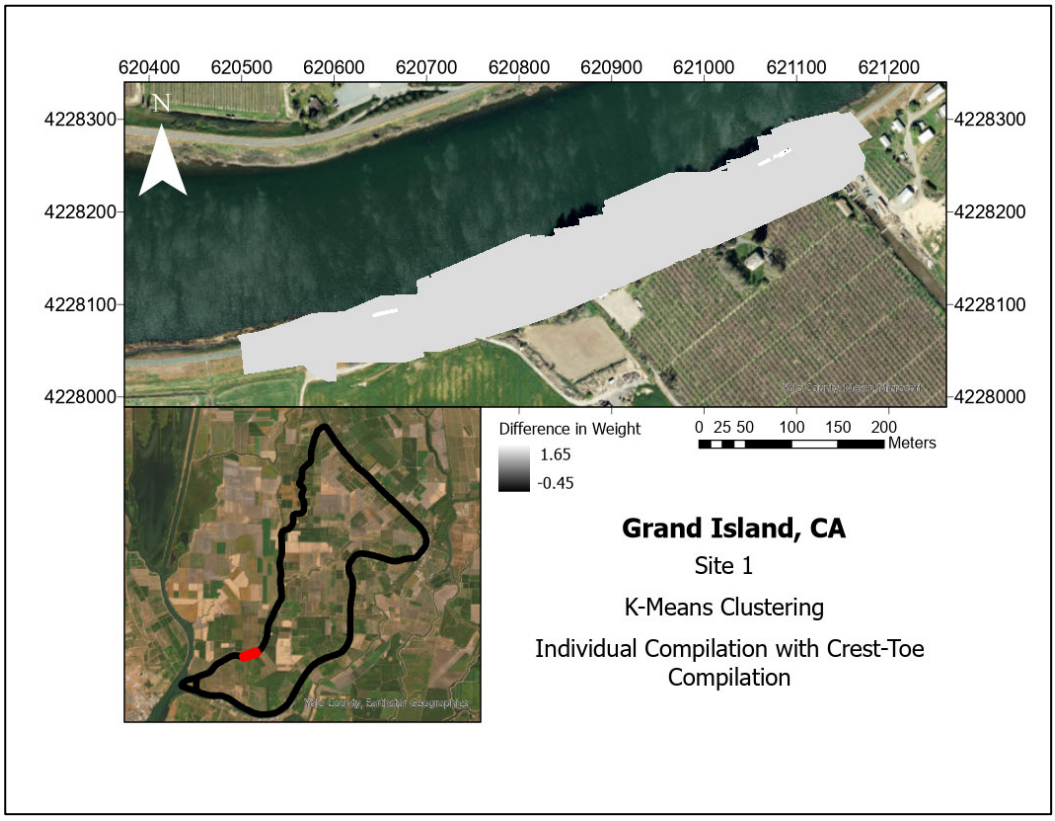

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Weight
1.4
-0.2
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
Hierarchical Clustering
Individual Compilation with Dry-Wet Compilation

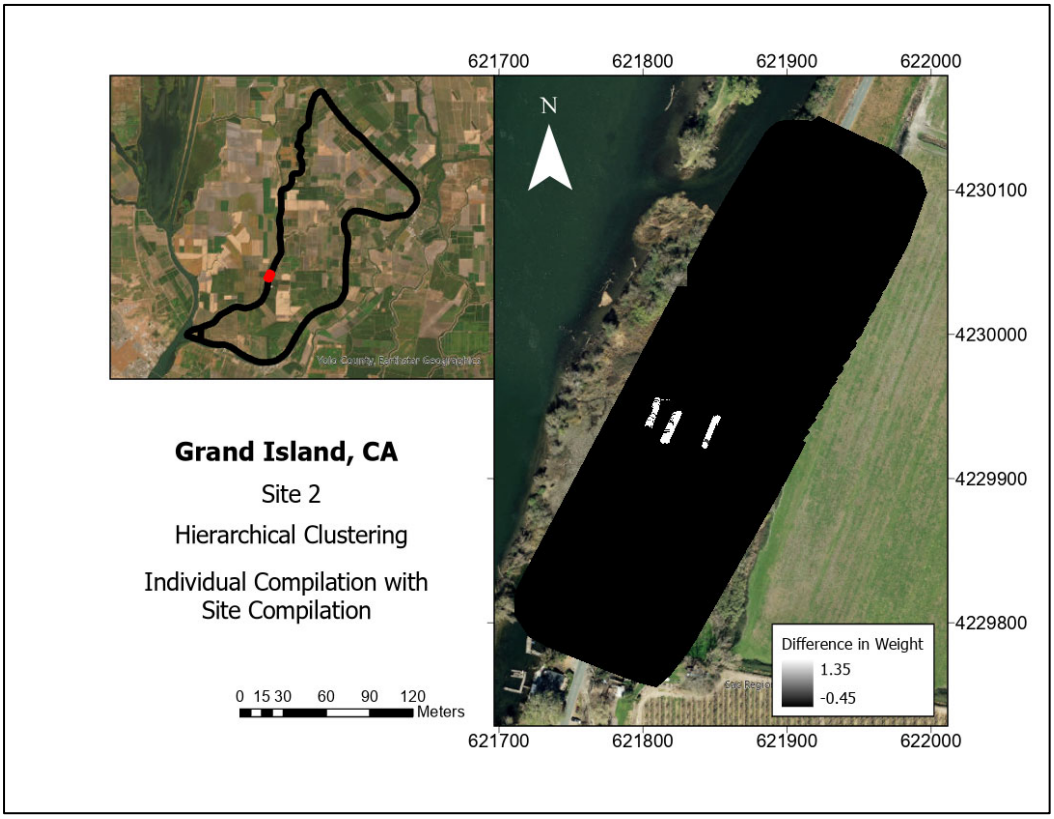

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Weight
1.1
-0.2
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
Hierarchical Clustering
Individual Compilation with Crest-Toe Compilation

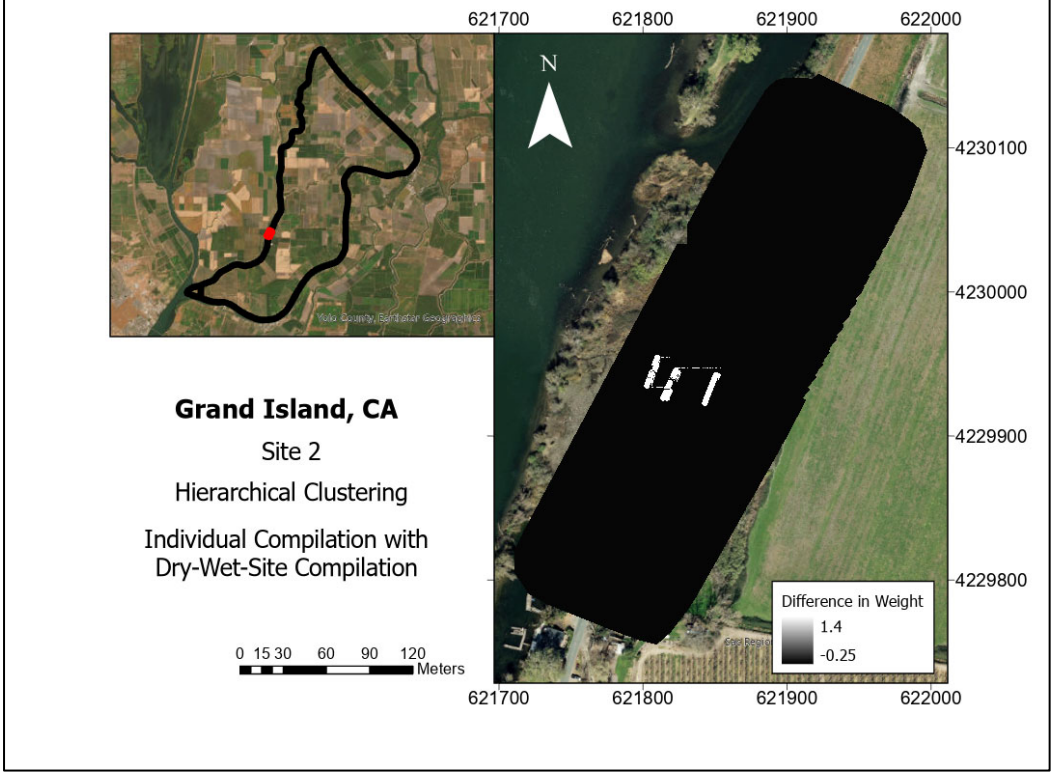

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Weight
0.9
-0.45
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
K-Means Clustering
Individual Compilation with Site Compilation

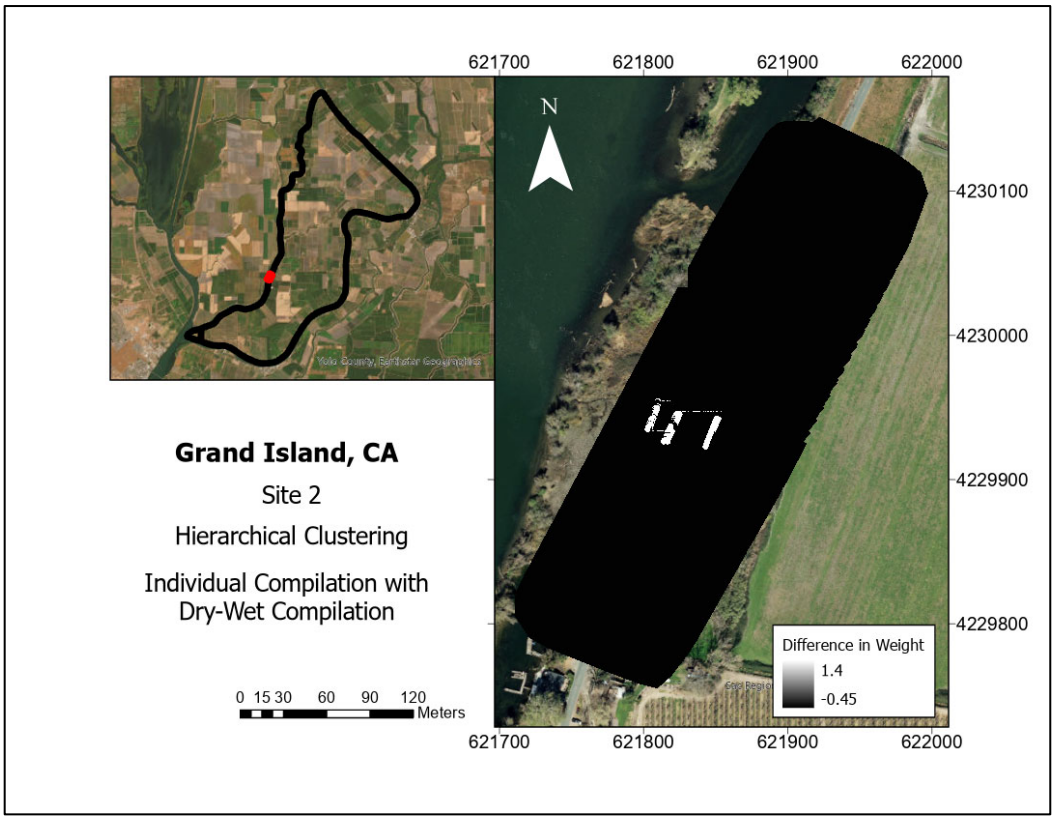

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Value
0.95
-0.2
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
K-Means Clustering
Individual Compilation with Dry-Wet-Site Compilation

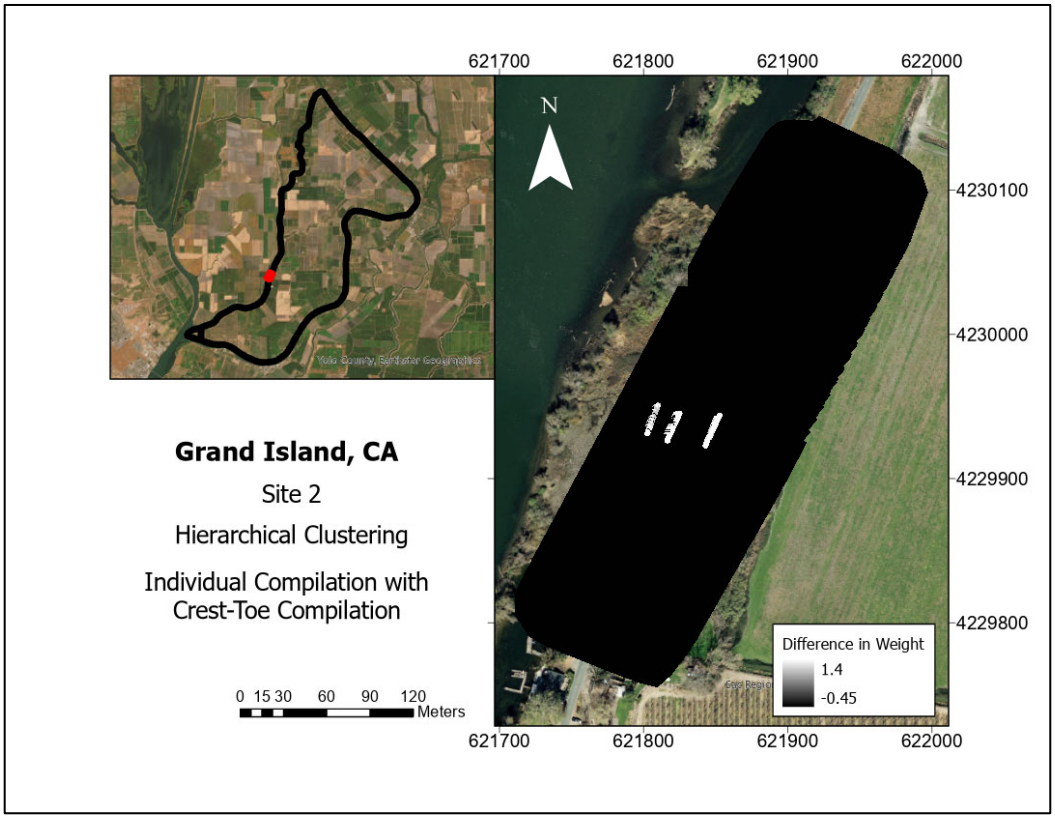

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Value
1.35
-0.2
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
K-Means Clustering
Individual Compilation with Dry-Wet Compilation

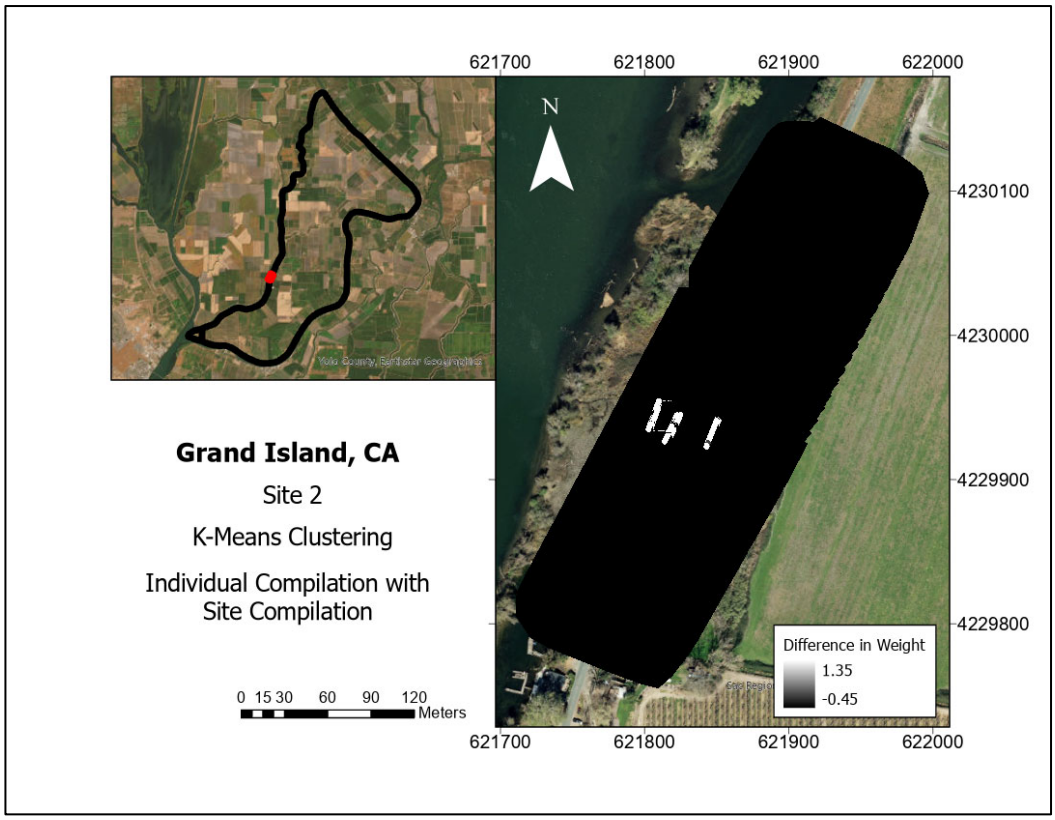

620400
620500
620600
620700
620800
620900
621000
621100
621200
4228300
4228200
4228100
4228000
N
Difference in Weight
1.65
-0.45
0 25 50 100 150 200
Meters
Grand Island, CA
Site 1
K-Means Clustering
Individual Compilation with Crest-Toe Compilation

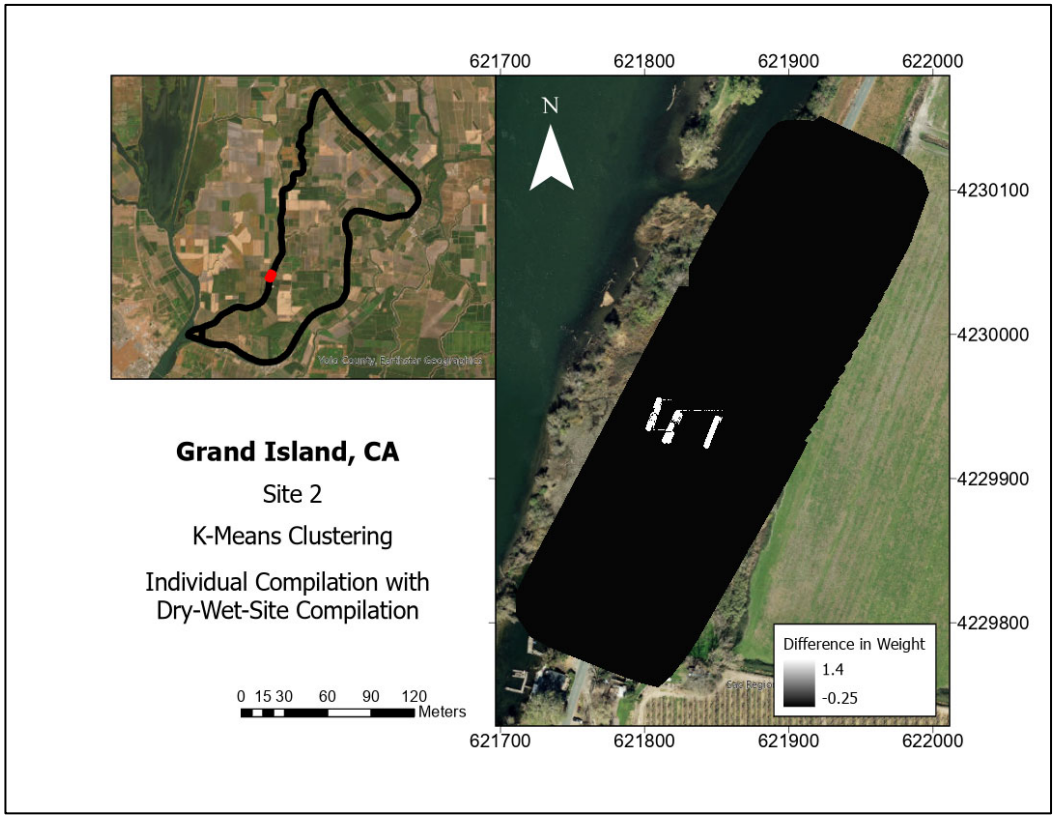

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
Hierarchical Clustering
Individual Compilation with
Site Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.35
-0.45

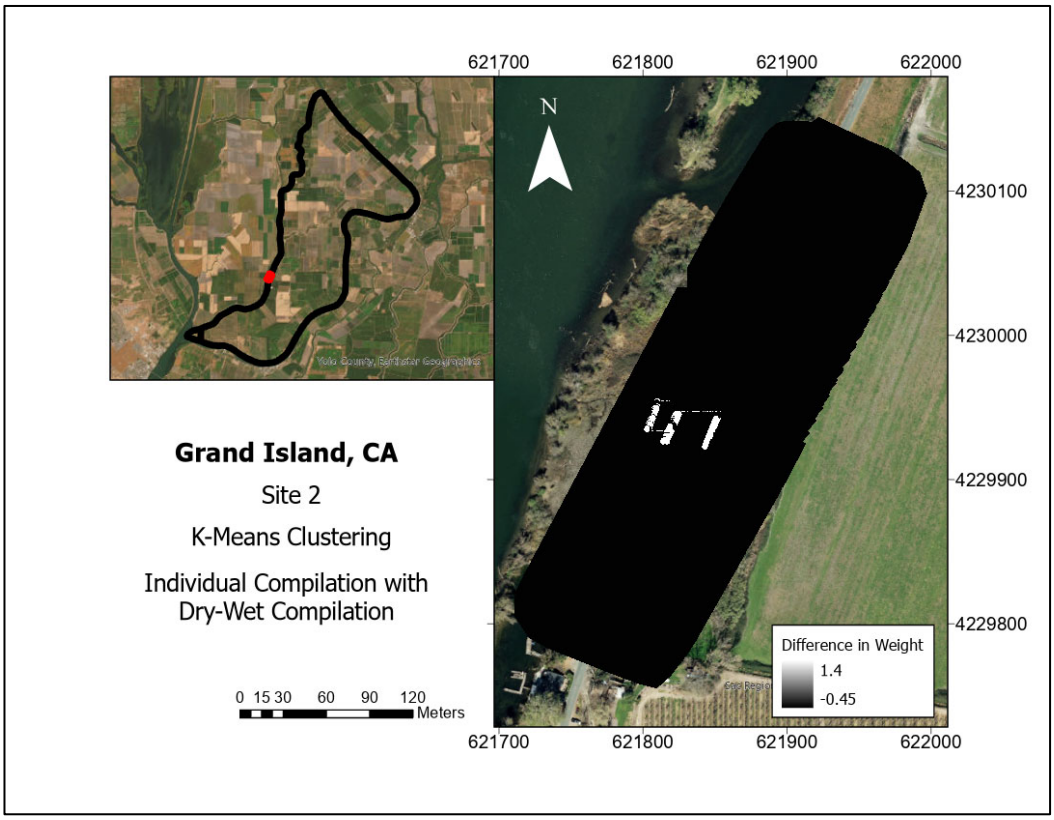

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
Hierarchical Clustering
Individual Compilation with
Dry-Wet-Site Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.4
-0.25

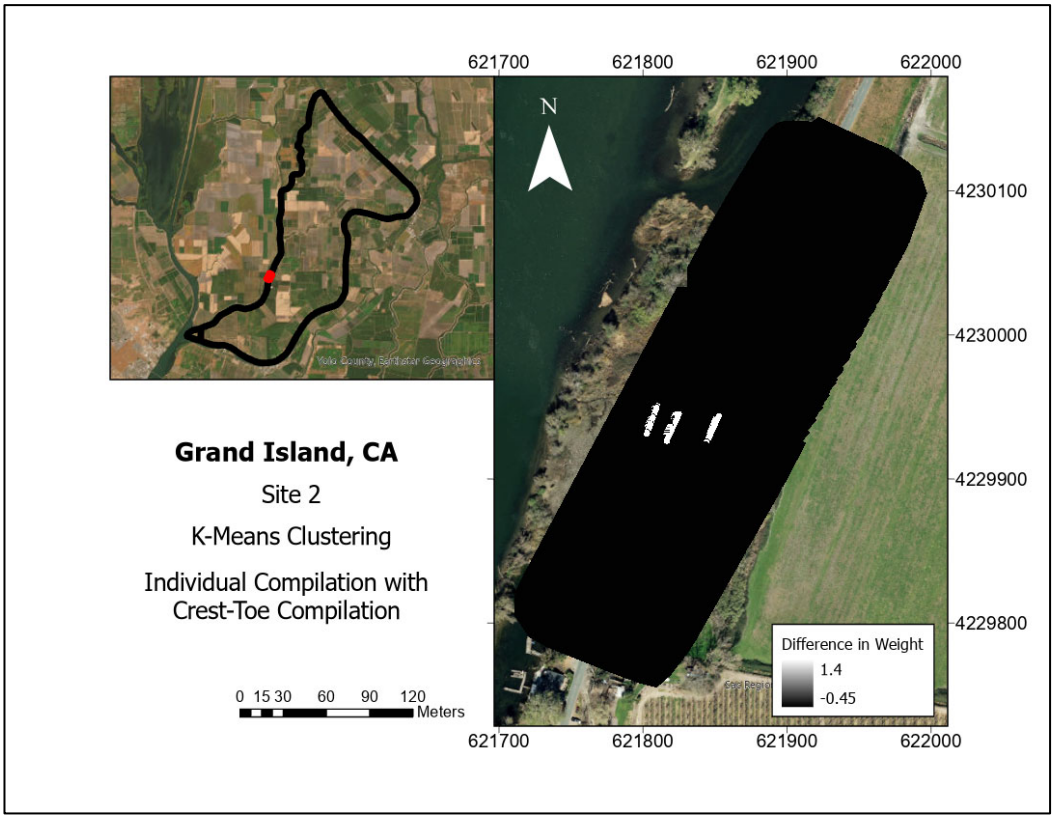

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
Hierarchical Clustering
Individual Compilation with
Dry-Wet Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.4
-0.45
621700
621800
621900
622000

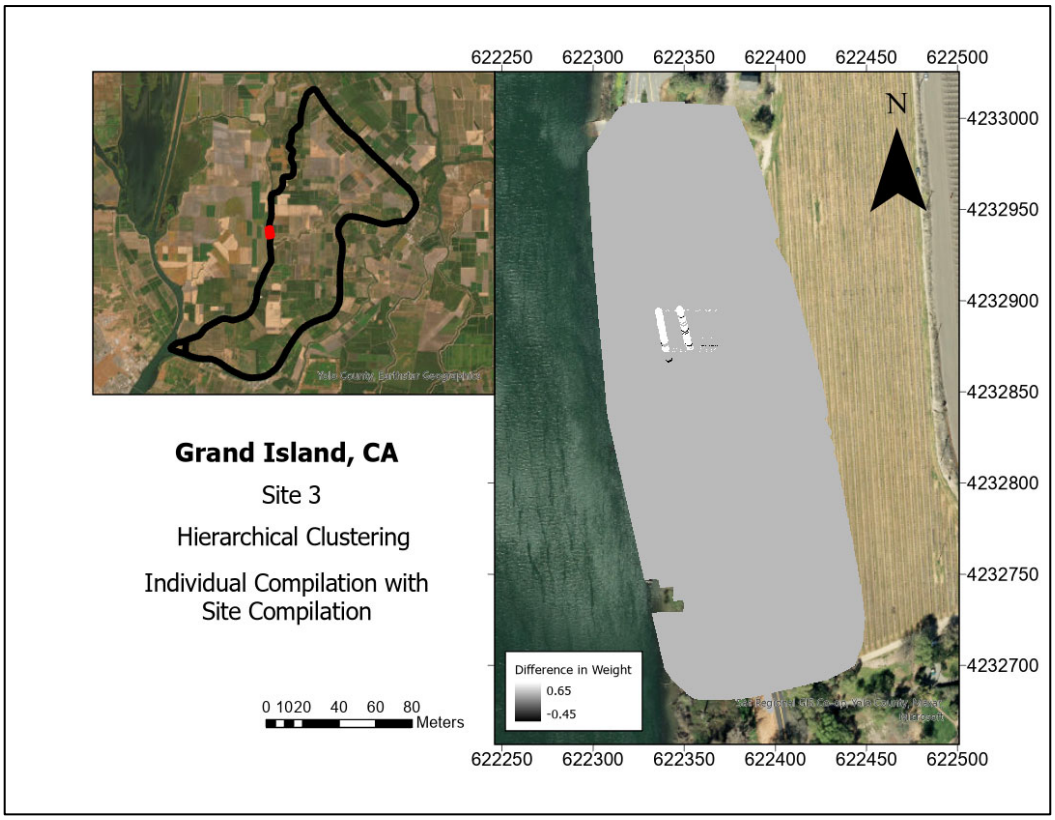

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
Hierarchical Clustering
Individual Compilation with
Crest-Toe Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.4
-0.45
621700
621800
621900
622000

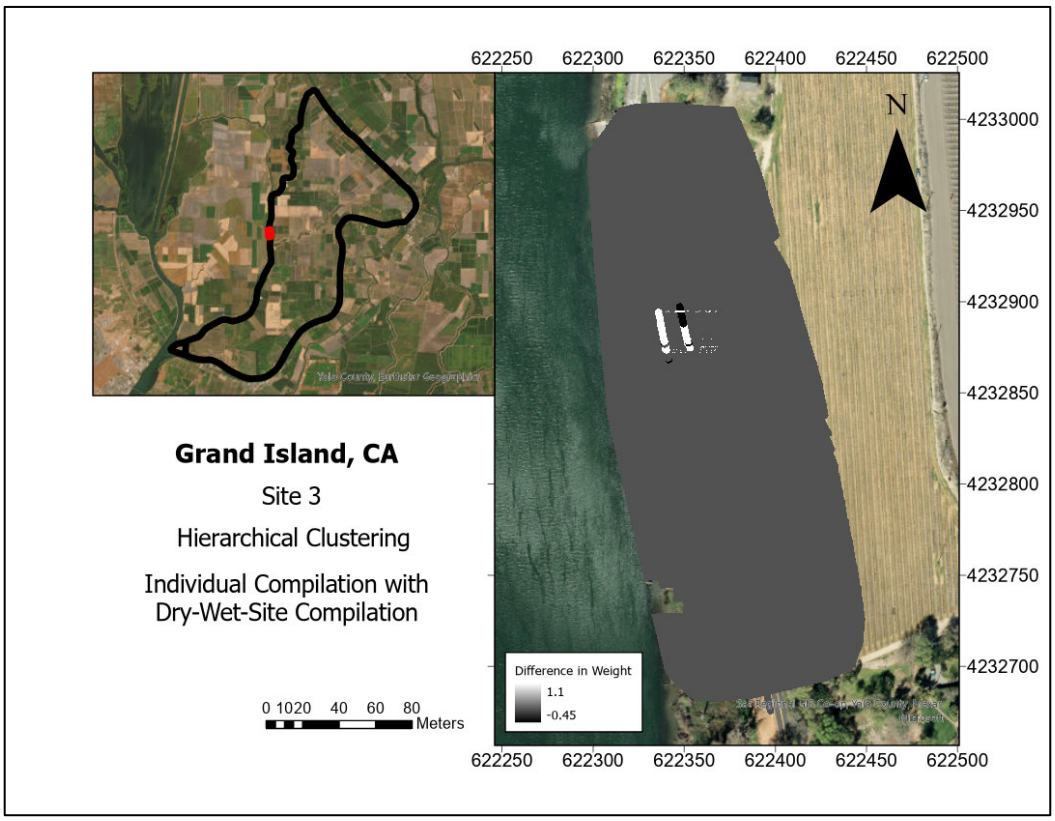

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
K-Means Clustering
Individual Compilation with
Site Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.35
-0.45
621700
621800
621900
622000

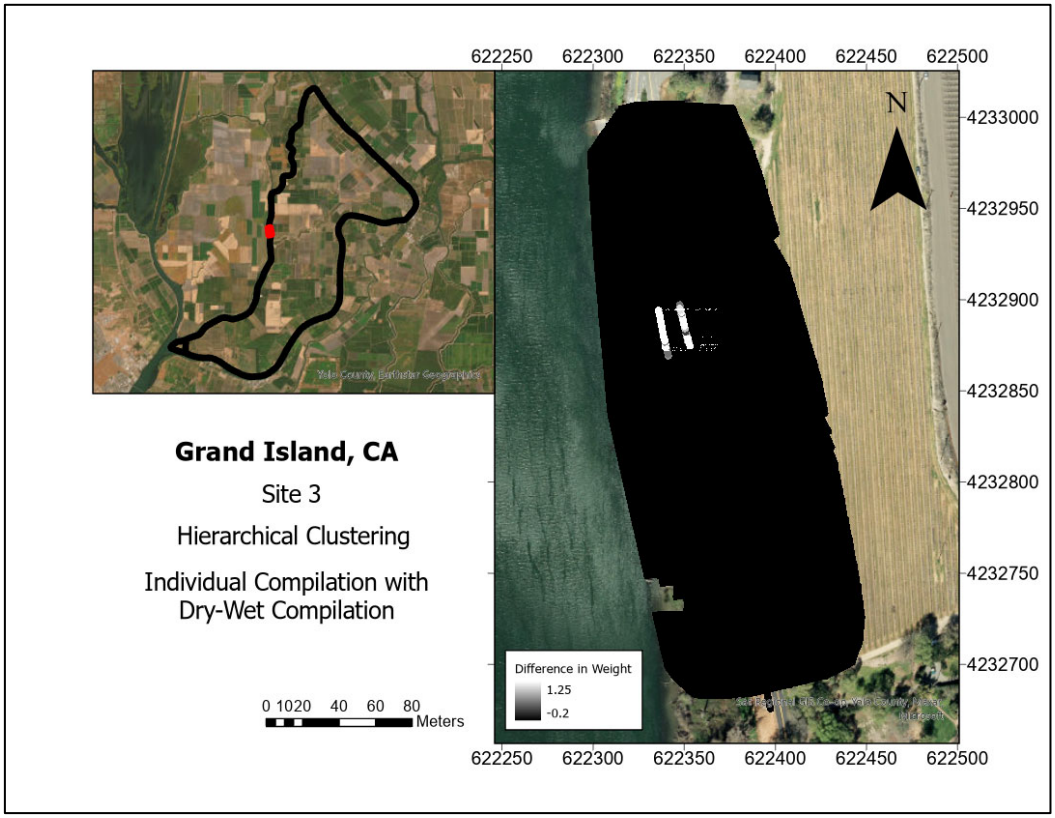

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
K-Means Clustering
Individual Compilation with
Dry-Wet-Site Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.4
-0.25
621700
621800
621900
622000

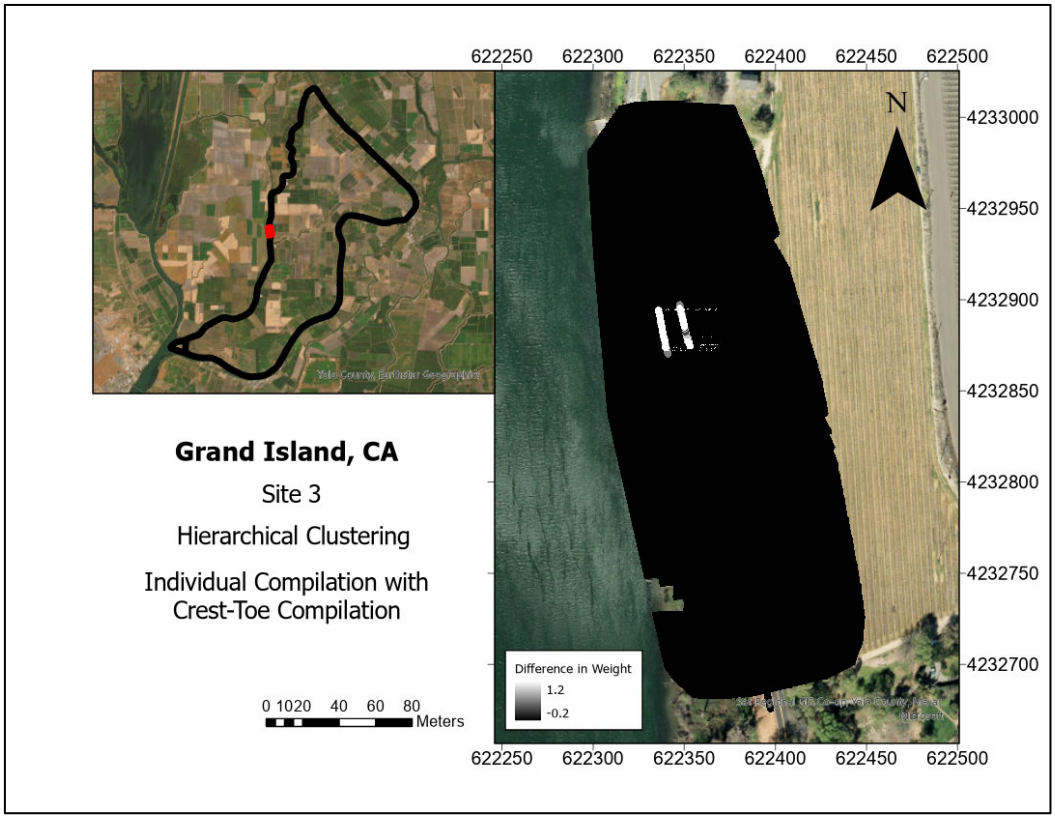

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
K-Means Clustering
Individual Compilation with
Dry-Wet Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.4
-0.45

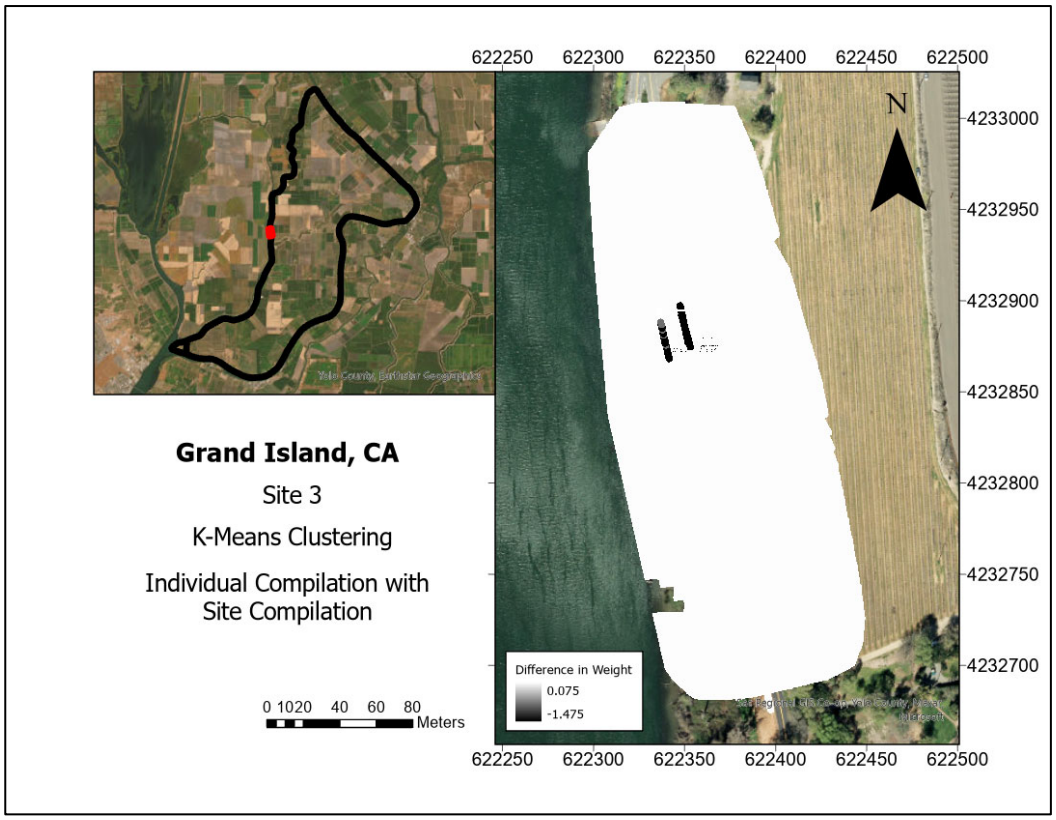

621700
621800
621900
622000
N
4230100
4230000
4229900
4229800
Grand Island, CA
Site 2
K-Means Clustering
Individual Compilation with
Crest-Toe Compilation
0 15 30 60 90 120
Meters
Difference in Weight
1.4
-0.45

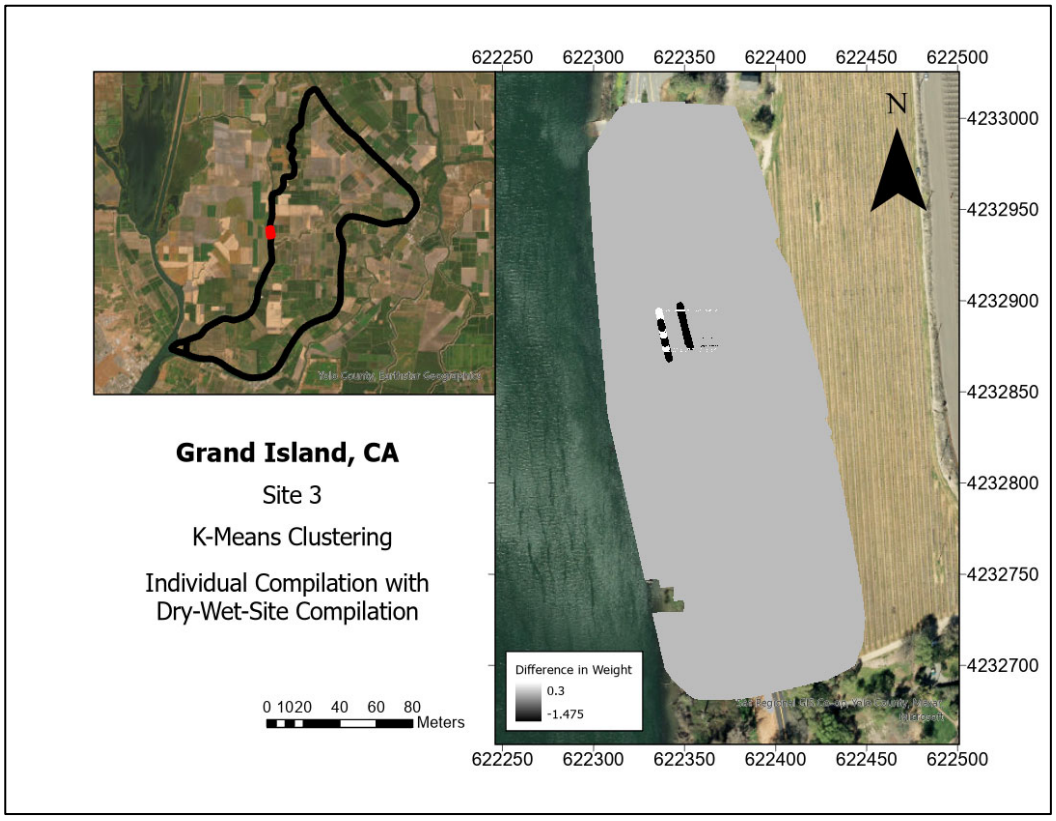

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hierarchical Clustering
Individual Compilation with
Site Compilation
0 1020 40 60 80
Meters
Difference in Weight
0.65
-0.45

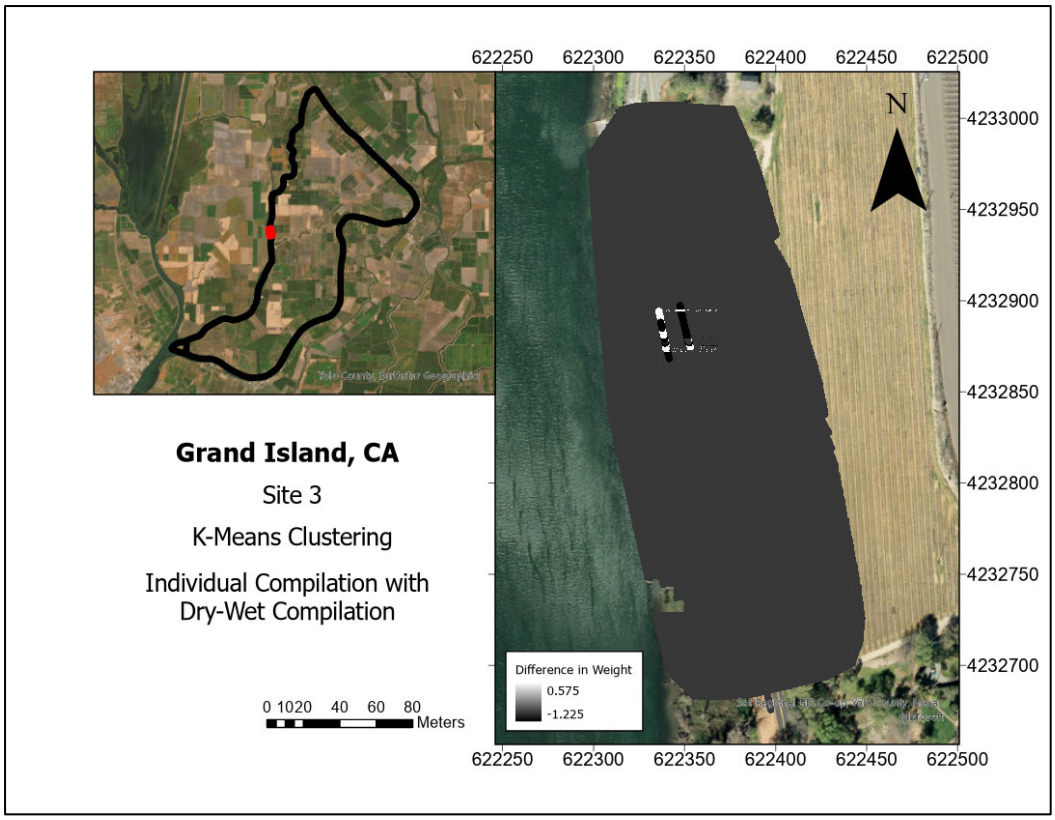

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hierarchical Clustering
Individual Compilation with
Dry-Wet-Site Compilation
0 1020 40 60 80
Meters
Difference in Weight
1.1
-0.45

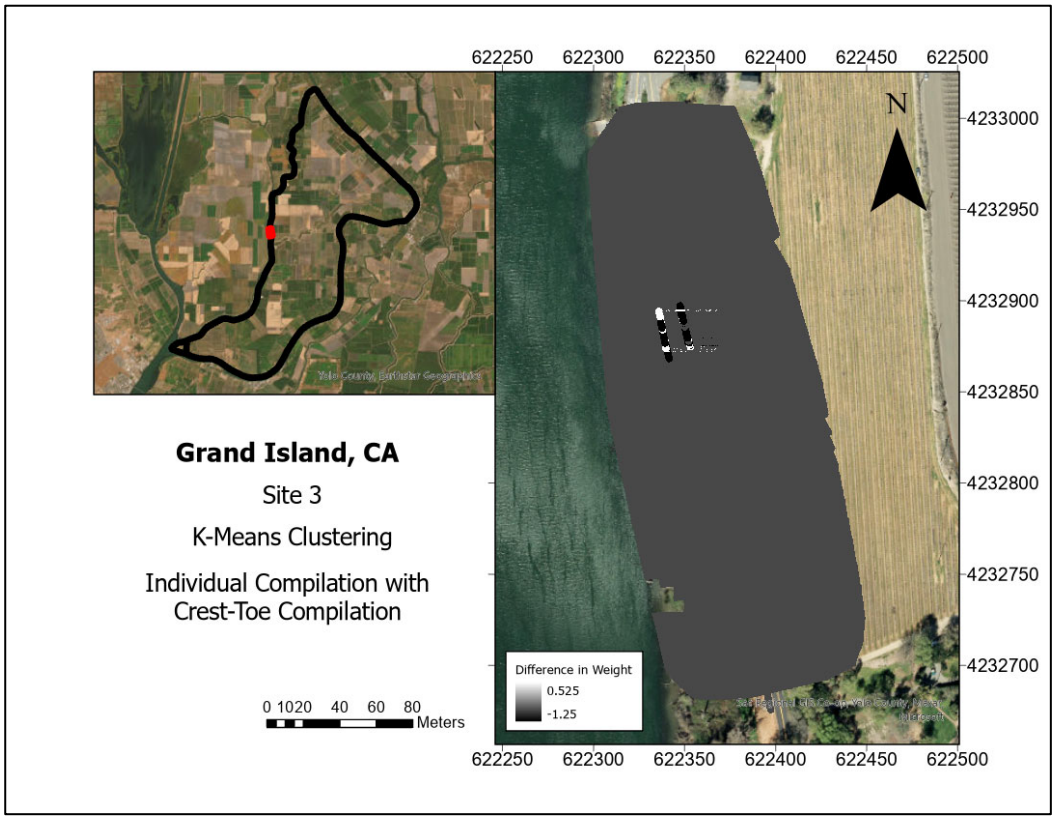

622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hierarchical Clustering
Individual Compilation with
Dry-Wet Compilation
0 1020 40 60 80
Meters
Difference in Weight
1.25
-0.2


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
Hierarchical Clustering
Individual Compilation with
Crest-Toe Compilation
0 1020 40 60 80
Meters
Difference in Weight
1.2
-0.2


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
K-Means Clustering
Individual Compilation with
Site Compilation
0 1020 40 60 80
Meters
Difference in Weight
0.075
-1.475


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
K-Means Clustering
Individual Compilation with
Dry-Wet-Site Compilation
0 1020 40 60 80
Meters
Difference in Weight
0.3
-1.475


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
K-Means Clustering
Individual Compilation with
Dry-Wet Compilation
0 1020 40 60 80
Meters
Difference in Weight
0.575
-1.225


622250
622300
622350
622400
622450
622500
4233000
4232950
4232900
4232850
4232800
4232750
4232700
N
Grand Island, CA
Site 3
K-Means Clustering
Individual Compilation with
Crest-Toe Compilation
0 1020 40 60 80
Meters
Difference in Weight
0.525
-1.25