

UC Merced

UC Merced Electronic Theses and Dissertations

Title

Explainable Artificial Intelligence Internet of Things (XAIoT) Enabled Smart Sensing of Soil Carbon Content for Smart Application of Biochar

Permalink

<https://escholarship.org/uc/item/0m61z4qs>

Author

An, Di

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

**EXPLAINABLE ARTIFICIAL INTELLIGENCE
INTERNET OF THINGS (XAIOT) ENABLED SMART
SENSING OF SOIL CARBON CONTENT FOR SMART
APPLICATION OF BIOCHAR**

by

Di An

A Dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Electrical Engineering & Computer Science

Committee in charge:
Dr. YangQuan Chen, Chair
Dr. Mukesh Singhal
Dr. Dong Li
Dr. Gerardo Diaz

©2024 Di An

©2024 Di An
All rights are reserved.

The dissertation of Di An is approved, and it is acceptable
in quality and form for publication on microfilm and electronically:

Dr. YangQuan Chen

Date

Dr. Mukesh Singhal

Date

Dr. Dong Li

Date

Dr. Gerardo Diaz

Date

University of California, Merced

©2024 Di An

Dedicated to my parents Xinxiang An and Wenjin Luo,
and my brother Zhenkun An.
Also to my deepest love Shuwen Zhong.

ACKNOWLEDGEMENTS

As I approach the conclusion of nearly thirty years of academic pursuit with the nearing end of my doctoral journey, I reflect on my growth from adolescence to maturity, seeing a path that was winding yet always moving forward. I am grateful for all the ordinary and extraordinary souls I have encountered along the way. As the time has come to pen this acknowledgment, a testament to the multitude of influences that have graced my doctoral passage, I am deeply grateful for every person and encounter that has shaped my academic and personal journey through these four and a half years.

First and foremost, I must extend my deepest gratitude to Professor YangQuan Chen, my doctoral supervisor, for the most precious gift of mentorship. It has been the greatest honor of my academic career to study under your guidance at UC Merced. Your excellence is consistent, evidenced by your continual reading, learning, innovating, and summarizing—a habit so ingrained that you often say excellence itself is a habit. No matter how busy, you have always made time for one-on-one guidance whenever a student is in need. Not only have I learned vast knowledge from you, but also life’s philosophies. Your profound insights have significantly influenced my choices and attitude towards life: “Our learning must be grounded in reality, possess the courage to withstand comparison, and strive for a sense of achievement recognized for its value. When you dedicate enough time and effort in one direction, you are not merely following someone’s research progress, but creating possibilities for that field.” You have shared not just your academic knowledge with me but also life experiences, leading me to a healthy and disciplined lifestyle. I fondly remember our walks around Castle to have drinks at Blaker, witnessing the seasonal landscapes of Merced; those late-night meetings that went into the early hours discussing academia; and how you read aloud from the collected works of Zhongshu Qian, “Poetry Can Be Resentful” or entitled “Our Sweetest Songs”, imparting traditional Chinese culture. I am especially grateful for your thoughtfulness towards me being away from home for five years, offering to visit my parents on my behalf to ensure their well-being. There are too many cherished memories to count. I hope we will always keep in touch and wish you and your family all the best for a smooth and peaceful life.

To my Ph.D. committee members, Professor Mukesh Singhal, Professor Dong Li, and Professor Gerardo Diaz, I extend my sincere thanks for their rigorous evaluations and constructive criticisms that greatly shaped my research. Thank you to

my MESA Lab colleagues for accompanying me through my doctoral studies. To Dr. Derek Hollenbeck and Rafal Krzysiak, with whom we've spent countless days and nights discussing research, conducting experiments, and sharing unlimited drinks. To Dr. Jairo Viola for his selfless help at any time, and to those lab mates whom I didn't get to work with as much due to COVID-19: Dr. Norelys Aguila-Camacho, Dr. Furkan Guc, Shiang Cao, Dr. Guoxiang Zhang, Dr. Haoyu Niu, Justus Nwoke, Osama Fuad; as well as the visiting scholars and students: Dr. Kai Cao, Dr. Lihong Guo, Dr. Yiheng Wei, Dr. Zhifu Li, Dr. Weiyuan Ma, Dr. Peng Wang. Thank you all for the beautiful times spent in pursuit of academics. I wish you all the best in your future endeavors.

I am grateful for the friends with whom I lived in Merced for seven years: Dr. Cheng Chen, Dr. Ruihao Li, Dr. Miaomiao Liu, Yue Zhang, Dr. Xueting Li, Dr. Guangfu Li. Over the years, we have traveled together, shared foods, engaged in deep conversations late into the night, and spent the most precious seven years together in the beautiful small town of Merced.

With the traditional principles of saving the most important for last, I turn to acknowledge my best friends, who have recognized each other throughout these years. A special thanks to Xue Yang, I will always cherish the days we spent watching the stars on the rooftop, remembering the many things you said and the enchanting music of your flute. To Ke Yang, who I've known for over a decade since my undergraduate days, is always so steady and supportive. To Dr. Yanting Zhao, for the brief two years of research and life in Merced, our shared interests laid the foundation for a quick recognition of each other, and I hope we can always maintain our youthful spirits. Thanks to Dr. Huifang Dou, whose warmth has always been as comforting as a family member's, providing me with the most sincere care beyond the rigors of doctoral time. I hope my best friends will never forget our beginnings, the journey we've shared, and the joys and sorrows we've experienced together. They are among the most treasured parts of my life, and I look forward to a future that remains inseparable.

Finally, my deepest reverence and heartfelt appreciation are reserved for my parents, whose unwavering sacrifices and dedication have profoundly influenced my very being and underpinned my academic accomplishments. My proudest achievement is to be your greatest pride. Moreover, my gratitude extends to my dear brother and to one who holds a special place in my heart, and she knows the depth of my appreciation. This acknowledgment is lovingly ascribed to the pillars of my life: my cherished parents and my deepest love.

Merced, CA, USA
March 20, 2024

Di An

CURRICULUM VITAE

Di An

Education

- **University of California, Merced.** Merced, CA, USA. (Jan. 2020 – Present).
Ph.D. in Electrical Engineering & Computer Science.
- **University of California, Merced.** Merced, CA, USA. (Aug. 2017 – May. 2019).
M.Sc. in Electrical Engineering & Computer Science.
- **Beijing Information Science & Technology University.** Beijing, China. (Sep. 2012 - Jun. 2016).
B.Sc. in Electronic Engineering.

Awards

- Won CITRIS Aviation Competition entitled “long endurance edge-AI platform for research opportunities and data gathering” (LEAP-FROG), Aug. 2022
- Won China-US Young Maker competition Third Prize, Aug. 2022
- Won finalist proposal UC Big Ideas competition, Top 10% - May. 2022
- Award Bobcat Fellowship of EECS from UC Merced - Summer 2022
- Won Student Success Internship (SSI) program for paying undergraduates spring research intern - Spring 2022
- Awarded Bobcat Fellowship of EECS from UC Merced - Summer 2021
- Won Student Success Internship (SSI) program for paying undergraduates summer research intern - Summer 2021
- Awarded Bobcat Fellowship of EECS from UC Merced - Summer 2020
- Awarded Scholarship for Outstanding Academic Performance - 2013-2015
- Awarded National Encouragement Fellowship - 2012-2014

TALKS

1. XAIoT Enabled Smart Sensing of Soil Carbon Content for Smart Application of Biochar
Poster, In Person, Sacramento, USA, Biochar Conference, Feb. 13th, 2024
2. XAIoT Enabled Smart Sensing of Soil Carbon Content for Smart Application of Biochar
Oral, In Person, UC Merced, USA, EECS Seminar, Dec. 15th, 2023
3. Explainable Artificial Intelligence (XAI) Empowered Digital Twin on Soil Carbon Emission Management Using Proximal Sensing
Oral, In Person, Orlando, USA, IEEE DTPI Conference, Nov. 7, 2023
4. A Proximal Point Sensing System for Mapping Soil Moisture Using A Miniaturized Spectrometer
Oral, In Person, Luxembourg, Luxembourg, IEEE ICCMA Conference, Oct. 10th, 2023
5. Long Endurance Site-Specific Management of Biochar Applications Using Unmanned Aircraft Vehicle and Unmanned Ground Vehicle.
Oral, In Person, Yokohama, Japan, IFAC World Congress, Jul. 3th, 2023
6. Battery-health-aware UAV mission planning using a cognitive battery management system.
Oral, In Person, Warsaw, Poland, ICUAS, Jun. 26th, 2023
7. A Miniature Millimeter-Wave Radar Based Contactless Lithium Polymer Battery Capacity Sensing with Edge Artificial Intelligence.
Oral, In Person, Taipei, China, IEEE/MESA Conference, Nov. 8th, 2022
8. A Soil Carbon Content Quantification Method Using A Miniature Millimeter Wave Radar Sensor and Machine Learning.
Oral, In Person, Taipei, China, IEEE/MESA Conference, Nov. 9th, 2022
9. A Digital Twin Enabled Internet of Living Things (IoLT) Framework for Soil Carbon Management.
Oral, In Person, Taipei, China, IEEE/MESA Conference, Nov. 15th, 2022

10. A Greenhouse Gas Proximity Sensing System Using Chemiresistive Strip and Miniaturized Radar Array.
Oral, In Person, Luxembourg, Luxembourg, IEEE ICCMA Conference, Nov. 23rd, 2022
11. Optimal Antenna Pairing of A Miniaturized Radar Array for Smart Sensing of Soil Carbon Content.
Poster, In Person, Shenyang, Liaoning, China, IEEE IAI Conference, Aug. 26th, 2022
12. A Field Study of Soil Biochar Treatment Response Using Small Unmanned Aerial Systems (sUAS).
Oral, Virtual Online, Dubrovnik, Croatia, ICUAS Conference, Jun. 14th, 2022
13. Smart Bi-eBikes (SBB): A low cost UGV solution for precision agriculture applications.
Oral, In Person, Orlando, USA, SPIE Conference, Apr. 5th, 2022
14. A Non-intrusive Quantification Method for Biochar Water Retention Capacity Using A Portable Microwave Sensor and Machine Learning.
Oral, Virtual Online, Luxembourg, ICCMA Conference, Nov. 27th, 2021
15. Digital Twin enabled methane emission abatement using networked mobile sensing and mobile actuation.
Oral, Virtual Online, Beijing, DTPI Conference, Aug. 15th, 2021
16. Application of Smart, Swarm and Small UAV's for Methane Emission. Reduction
Oral, Virtual Online, IDTEC/CIE ASME Conference, May, 2021
17. Regular group presentations on SGC biochar project, start with a title "Methane Mapping by Collaborative Stationary and Mobile Sensors - Unmanned Aerial Vehicles & Unmanned Ground Vehicles"
Oral, Hybrid, Feb. 2020 - 2023

PUBLICATIONS

Journal Publications:

1. **An, D., & Chen, Y. (2023).** Non-intrusive soil carbon content quantification methods using machine learning algorithms: A comparison of microwave and millimeter wave radar sensors. *Journal of Automation and Intelligence*, 2(3), 152-166.

2. **An, D.**, Hollenbeck, D., Cao, K., & Chen, Y. Q. (2022). Soil Methane Emission Suppression Control Using Unmanned Aircraft Vehicle Swarm Application of Biochar Mulch - A Simulation Study, *Journal of Information and Intelligence*, 1(1), 68-85.
3. Wang, P., **An, D.**, Chen, N., & Chen, Y. Q. (2021). Spatial path tracking controllers for autonomous ground vehicles: Conventional and nonconventional schemes. *Guidance, Navigation and Control*, 1(01), 2150003.
4. Gao, S., Harrison, B. P., Thao, T., Gonzales, M. L., **An, D.**, Ghezzehei, T. A., Diaz, G., & Ryals, R. A. (2022). Biochar co-compost improves nitrogen retention and reduces carbon emissions in a winter wheat cropping system. *GCB Bioenergy*, 15, 462– 477.
5. Cao, K., Chen, Y. Q., Gao, S., Yan, K., Zhang, J., & **An, D.** (2023). Omni-Directional Capture for Multi-Drone Based on 3D-Voronoi Tessellation. *Drones*, 7(7), 458.
6. Cao, K., Chen, Y., Gao, S., Dang, H., & **An, D.** (2023). Distributed Weighted Coverage for Multi-Robot Systems in Non-Convex Environment. *Applied Sciences*, 13(14), 8530.

Conference Publications:

1. **An, D.**, & Chen, Y. (2023, Nov.). Explainable Artificial Intelligence (XAI) Empowered Digital Twin on Soil Carbon Emission Management Using Proximal Sensing. In *2023 IEEE 3rd International Conference on Digital Twins and Parallel Intelligence (DTPI)* (pp. 1-5). IEEE.
2. Krzysiak, R., **An, D.**, & Chen, Y. (2023, Nov.). XCardio-Twin: An explainable framework to aid in monitoring and analysis of cardiovascular status. In *2023 IEEE 3rd International Conference on Digital Twins and Parallel Intelligence (DTPI)* (pp. 1-6). IEEE.
3. Hollenbeck, D., **An, D.**, Krzysiak, R., & Chen, Y. (2023, Nov.). Towards Cognitive Battery Monitoring on Hybrid VTOL Fixed-Wing sUAS with Maximized Safe Endurance. In *2023 11th International Conference on Control, Mechatronics and Automation (ICCMA)* (pp. 80-85). IEEE.
4. **An, D.**, Krzysiak, R., Cao, S., Zhao, T., & Chen, Y. (2023, Nov.). A Proximal Point Sensing System for Mapping Soil Moisture Using A Miniaturized Spectrometer. In *2023 11th International Conference on Control, Mechatronics and Automation (ICCMA)* (pp. 290-295). IEEE.

5. **An, D.**, & Chen, Y. (2023, Aug.). More Optimized Stochastic Configuration Networks (MOSCNS) Enabled Smart Sensing for Soil Carbon Content Quantification. In 2023 5th International Conference on Industrial Artificial Intelligence (IAI) (pp. 1-6). IEEE.
6. Cao, S., **An, D.**, Wang, J., & Chen, Y. (2023, Aug.). Portable, Non-contact, Proximal Sensing of Sands in Soil Using Microwave Radar Array, Fractional Fourier Transform, and Machine Learning. In 2023 5th International Conference on Industrial Artificial Intelligence (IAI) (pp. 1-5). IEEE.
7. **An, D.**, Krzysiak, R., Hollenbeck, D., & Chen, Y. (2023). Battery-health-aware UAV mission planning using a cognitive battery management system. In 2023 International Conference on Unmanned Aircraft Systems (ICUAS)(pp. 523-528). IEEE.
8. **An, D.**, Krzysiak, R., Hollenbeck, D., & Chen, Y. (2023). Long Endurance Site-Specific Management of Biochar Applications Using Unmanned Aircraft Vehicle and Unmanned Ground Vehicle. IFAC-PapersOnLine, 56(2), 8908-8913.
9. **An, D.**, & Chen, Y. (2022, Dec.). A Miniature Millimeter-Wave Radar Based Contactless Lithium Polymer Battery Capacity Sensing with Edge Artificial Intelligence. In 2022 18th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA) (pp. 1-6). IEEE.
10. **An, D.**, & Chen, Y. (2022, Dec.). A Soil Carbon Content Quantification Method Using A Miniature Millimeter Wave Radar Sensor and Machine Learning. In 2022 18th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA) (pp. 1-6). IEEE.
11. **An, D.**, & Chen, Y. (2022, Dec.). A Digital Twin Enabled Internet of Living Things (IoLT) Framework for Soil Carbon Management. In 2022 18th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA) (pp. 1-6). IEEE.
12. Ajcac, R., Viola, J., **An, D.**, & Chen, Y. (2022, Nov.). A Control System Benchmark for Biomass Burning Based Thermoelectric Generation: Modeling and Efficiency Maximization Algorithms. In 2022 10th International Conference on Control, Mechatronics and Automation (ICCMA) (pp. 129-134). IEEE.
13. **An, D.**, Liu, Z., Hollenbeck, D., & Chen, Y. (2022, Nov.). A Greenhouse Gas Proximity Sensing System Using Chemiristive Strip and Miniaturized Radar

- Array. In 2022 10th International Conference on Control, Mechatronics and Automation (ICCMA) (pp. 173-178). IEEE.
14. **An, D.**, Difrieri, M., & Chen, Y. (2022, Aug.). Optimal Antenna Pairing of A Miniaturized Radar Array for Smart Sensing of Soil Carbon Content. In 2022 4th International Conference on Industrial Artificial Intelligence (IAI) (pp. 1-6). IEEE.
 15. **An, D.**, Niu, H., Williams, L., & Chen, Y. (2022, Jun.). Smart Bi-eBikes (SBB): A low cost UGV solution for precision agriculture applications. In Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping VII (Vol. 12114, pp. 80-88). SPIE.
 16. **An, D.**, Hollenbeck, D., Gao, S., & Chen, Y. (2022, Jun.). A Field Study of Soil Biochar Treatment Response Using Small Unmanned Aerial Systems (sUAS). In 2022 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 1608-1613). IEEE.
 17. **An, D.**, Niu, H., & Chen, Y. (2021, Nov.). A Non-intrusive Quantification Method for Biochar Water Retention Capacity Using A Portable Microwave Sensor and Machine Learning. In 2021 9th International Conference on Control, Mechatronics and Automation (ICCMA) (pp. 152-157). IEEE.
 18. **An, D.**, & Chen, Y. (2021, Aug.). Application of Smart, Swarm and Small UAV's for Methane Emission Reduction. In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 85437, p. V007T07A050). American Society of Mechanical Engineers.
 19. **An, D.**, & Chen, Y. (2021, Jul.). Digital Twin enabled methane emission abatement using networked mobile sensing and mobile actuation. In 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI) (pp. 354-357). IEEE.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	v
CURRICULUM VITAE	vii
LIST OF FIGURES	xxii
LIST OF TABLES	xxxii
ABSTRACT	xxxiv
ABBREVIATIONS	xxxv

Chapter

1 INTRODUCTION	1
1.1 Motivations	1
1.1.1 What Are Soil Carbon and Its Impact?	2
1.1.2 What Is Smart Sensing in Estimation of the Soil Carbon?	4
1.1.3 What Is Smart Application of the Soil Carbon?	5
1.1.4 What Is XAIoT?	7
1.2 Dissertation Objectives and Methods	8
1.3 Dissertation Outline	10
1.4 Dissertation Contributions	11
1.5 Soil Organic Carbon Estimation: A Review	12
1.5.1 Introduction to Soil Carbon Sensing	12
1.5.2 Remote Sensing Technologies for Soil Carbon Estimation	13
1.5.2.1 Different Methodologies of Remote Sensing for Soil Carbon Estimation	13

1.5.2.2	Estimation Results	14
1.5.3	Proximal Sensing in Estimation of Soil Carbon	17
1.5.3.1	Different Types of The Proximal Sensors	17
1.6	Results Reproducibility	19
2	XAIOT EMPOWERED SMART DATA ACQUISITION PLATFORMS AND CASE STUDIES	20
2.1	The Microwave Radar Array Sensor	20
2.2	The Millimeter Wave Radar Array Sensor	22
2.3	The Unmanned Aircraft Vehicles	25
2.4	The Unmanned Ground Vehicles	26
2.5	Case Study I: A Non-intrusive Quantification Method for Biochar Water Retention Capacity	27
2.5.1	Microwave Reflectance Analysis	27
2.5.2	Microwave Reflective Signals Classification	28
2.5.3	Evaluation	29
2.5.3.1	Experiment Setup	29
2.5.3.2	Different Kinds of Biochar	30
2.5.3.3	Different Biochar Amount Classification	32
2.5.3.4	Features Extraction with Different Water Usage	33
2.5.3.5	Biochar Water Retention Capacity	35
2.6	Case Study II: Soil Carbon Content Quantification Method Using A Miniature Millimeter Wave Radar Sensor	36
2.6.1	Millimeter Wave Signal Classification Method	36
2.6.2	Experiments	37
2.6.2.1	Experiments Setup	37
2.6.3	Evaluation and Results	37
2.6.3.1	Different Soil Carbon Amount Recognition	37
2.6.3.2	Different Type of Soil Carbon Classification	38
2.6.3.3	Soil Moisture Study Using mm-waves	40

2.6.3.4	The Soil Moisture Detection Under Carbon Efforts	41
2.7	Case Study III: A Field Study of Soil Biochar Treatment Response Using Small Unmanned Aerial Systems (sUAS)	42
2.7.1	Introduction	42
2.7.2	Materials and Methods	44
2.7.2.1	Study Field Site and UASs Remote Sensing Configuration	44
2.7.2.2	Geospatial Interpolation with Kriging	46
2.7.2.3	Statistical Analysis	48
2.7.3	Experiments and Results	48
2.7.3.1	Correlation Analysis of The VARI and NDVI Indexes Versus Temperature	48
2.7.3.2	Correlation Analysis of The NDVI Index Versus Different Depths of VWC	50
2.7.4	Discussion and Challenges	51
2.8	Case Study IV: Smart Bi-eBikes (SBB) - A Low Cost UGV Solution for Precision Agriculture Applications	54
2.8.1	Introduction	54
2.8.2	The UGV Architecture	55
2.8.2.1	Fundamentals of Bikes	55
2.8.2.2	Chassis Design	57
2.8.2.3	SBB Smart Brain	59
2.8.3	Navigation Guidance from Image Segmentation	61
2.8.4	Pest Management Network	62
2.8.5	Discussion and Challenges	63
2.9	Chapter Summary	64
3	NON-INTRUSIVE SOIL CARBON CONTENT QUANTIFICATION METHODS USING MACHINE LEARNING ALGORITHMS: A COMPARISON OF	

MICROWAVE AND MILLIMETER WAVE RADAR SENSORS 66

3.1	Introduction	66
3.2	Related Work	68
3.2.1	Remote Sensing of Soil Carbon Content	68
3.2.2	Proximal Sensing of Soil Carbon Content	69
3.3	Materials and Methods	70
3.3.1	Sensor Hardware - Microwave Sensing Module	70
3.3.2	Sensor Hardware - Millimeter Wave Sensing System	71
3.3.3	Microwave Reflective Signal Classification	72
3.3.4	Millimeter Wave Reflectance Analysis	73
3.3.5	Conventional Classification Methods	74
3.3.5.1	Support Vector Machine	74
3.3.5.2	Decision Tree	76
3.3.5.3	Linear/Quadratic Discriminant Analysis	77
3.3.5.4	K Nearest Neighbors	77
3.3.5.5	Naive Bayes	78
3.3.5.6	Neural Network	78
3.3.6	More Optimized Stochastic Configuration Networks (MOSCNs)	80
3.3.6.1	Stochastic Configuration Networks	80
3.3.6.2	SC-I Classifier	82
3.3.6.3	SC-II Classifier	83
3.3.6.4	SC-III Classifier	83
3.3.6.5	DeepSCN Classifier	84
3.3.6.6	Random Variables - Uniform Distributions	85
3.3.6.7	Random Variables - Gaussian Distributions	86
3.3.6.8	Random Variables - Lévy Distributions	86
3.3.6.9	Random Variables - Cauchy Distributions	87
3.3.6.10	Random Variables - Weibull Distributions	87
3.3.6.11	Fractional Gaussian Distributions	87
3.3.6.12	Fractional Weibull Distributions	88
3.4	Experiments and Evaluation	89
3.4.1	Experimental Setup	89

3.4.2	Conventional Classification Results	91
3.4.2.1	Different Soil Carbon Amount Recognition	91
3.4.2.2	Different Types of Soil Carbon Classification	94
3.4.2.3	Soil Moisture Study and Comparison	97
3.4.2.4	Soil Moisture Influenced by Carbon Content Added .	100
3.4.2.5	Sensing Robustness for Soil Carbon Content Percentage	102
3.4.3	Optimal Distributions Parameters Tuning for SCNs Regression Model	108
3.4.4	Comparison of The Optimal Parameters of Fractional Distributions and Non-fractional Distributions for Classification Mission.	113
3.5	Discussion and Future Work	114
3.6	Chapter Summary	116
4	SITE SPECIFIC MANAGEMENT OF SOIL CARBON CONTENT SENSING ON LARGE-SCALE FIELD	117
4.1	What is Site Specific Management?	117
4.2	Digital Twin Enabled Methane Emission Abatement Using Networked Mobile Sensing and Mobile Actuation	118
4.2.1	Introduction	118
4.2.2	Digital Twin in Methane Abatement	120
4.2.2.1	The Physical Principle Behind Methane Emissions .	120
4.2.2.2	Uncertainty Conditions of Methane Emissions	120
4.2.3	Networked Mobile Sensing and Actuation	121
4.2.4	Preliminary Results	122
4.2.5	Discussion and Future Works	124
4.3	A Digital Twin Enabled Internet of Living Things (IoLT) Framework for Soil Carbon Management	124
4.3.1	Introduction	124
4.3.2	Digital Twin in Soil Carbon Management	126

4.3.3	Digital Twin Front End of Soil Carbon Accounting	127
4.3.3.1	Proximity Sensors for Measuring SOC	127
4.3.3.2	Swarm UAVs Mobile Platform	129
4.3.4	Digital Twin Back End for Soil Carbon Accounting	129
4.3.4.1	Soil Carbon Classification Using Machine Learning .	129
4.3.4.2	Smart Big Data for Soil Carbon Accounting	130
4.3.4.3	Digital Twin in Soil Carbon Accounting Assessment	130
4.3.5	Discussion and Future Work	131
4.4	A Case Study: Long Endurance Site-Specific Management of Biochar Applications Using Unmanned Aircraft Vehicle and Unmanned Ground Vehicle	131
4.4.1	Introduction	131
4.4.2	Framework and Architecture	133
4.4.2.1	Mission Planning	133
4.4.2.2	Site-specific Management Sensing System	134
4.4.2.3	Mobile Sensing Platform Payload	135
4.4.3	Approach Methods	136
4.4.3.1	Proximity Sensor Data Collection Configuration . . .	136
4.4.3.2	UAV Landing Process – A Simulation Approach . . .	137
4.4.3.3	Vision Process of Landing Target	138
4.4.4	Preliminary Results	139
4.4.4.1	Controller Performance Analysis	139
4.4.4.2	Tracking Moving Target Analysis	141
4.5	Conclusions	142
5	EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI) EMPOWERED DIGITAL TWIN ON SOIL CARBON	

EMISSION MANAGEMENT USING PROXIMAL SENSING . .	144
5.1 Introduction	144
5.2 XAIoT empowered Digital Twin in Soil Carbon Emissions Management	146
5.2.1 The XAIoT-DT Framework	146
5.2.2 XAI Explainer Strategy–Shapley Index	147
5.3 Preliminary Results	147
5.4 Discussion and Future Works	149
5.5 A Case Study: A Miniature Millimeter-Wave Radar Based Contactless Lithium Polymer Battery Capacity Sensing with XAI	149
5.5.1 Introduction	149
5.5.2 Materials and Methods	151
5.5.2.1 Sensing Hardware	151
5.5.2.2 Millimeter Wave Signal Classification Methods . . .	152
5.5.3 Experiments	153
5.5.3.1 Experiments Setup	153
5.5.3.2 Evaluation	154
5.5.4 Discussion and Future Work	156
5.6 Chapter Summary	158
6 SOIL METHANE EMISSION SUPPRESSION CONTROL USING AUTOMATIC DRONE SWARM APPLICATION OF BIOCHAR MULCH - A SIMULATION STUDY	160
6.1 Introduction	160
6.2 Problem Statement	163
6.3 Algorithm of Optimal Control Strategy & Simulation Platform	165
6.3.1 Optimal Actuation Algorithm	165
6.3.2 Wind Field Effects	166

6.3.3	Simulation Platform of FADE-MAS-2D	168
6.4	Simulation Experiments	169
6.4.1	Experiment Setup	169
6.4.2	Wind Effects	170
6.4.3	Time Fractional Diffusion Equation	171
6.4.4	Space Fractional Diffusion Equation	173
6.4.5	Time and Space Fractional Equations	174
6.4.6	An Optimal Control Strategy	175
6.4.7	Multiple Methane Emissions Sources	180
6.5	Conclusions	182
7	XAIOT ENABLED SMART APPLICATIONS	184
7.1	A Greenhouse Gas Proximal Sensing System Using Chemiresistive Strip and Miniaturized Radar Array	184
7.1.1	Introduction	184
7.1.2	Materials and Methods	186
7.1.2.1	Proximal Sensor Setup	187
7.1.2.2	Sensing Materials Preparation	188
7.1.3	Sensing Algorithms	190
7.1.3.1	Optimal Antenna Pair Selection	190
7.1.3.2	Savitzky-Golay (SG) Smoothing Filter	190
7.1.4	Experiments and Results	191
7.1.4.1	Classification of Different GHG Concentration	191
7.1.4.2	Maximum Gas Sensing Range	193
7.1.5	Discussion and Challenges	194
7.2	A Proximal Point Sensing System for Mapping Soil Moisture Using A Miniaturized Spectrometer	195
7.2.1	Introduction	195

7.2.2	Materials and Methods	197
7.2.2.1	Spectrometer Setup	197
7.2.2.2	Geospatial Interpolation with Kriging	199
7.2.2.3	Savitzky-Golay Smoothing Filter	199
7.2.3	Experiments and Results	200
7.2.3.1	Soil Moisture Uniformly Distributions	200
7.2.3.2	Soil Moisture Non-uniformly Distributions	203
7.2.3.3	Soil Moisture Regression Analysis	205
7.2.4	Discussion and Challenges	207
7.3	Chapter Summary	207
8	CONCLUSIONS AND FUTURE RESEARCH	209
8.1	Concluding Remarks	209
8.2	Future Research on XAIoT Empowered Smart Sensing of Soil Carbon Content	210
Appendix		
A	DATASET CONTRIBUTIONS	212
B	CODE CONTRIBUTIONS	213
	BIBLIOGRAPHY	214

LIST OF FIGURES

1.1	Trends in agriculture sector carbon emission sources [11].	3
1.2	Evolution of soil carbon sensing technologies.	12
2.1	Walabot sensing strategy ¹	20
2.2	Walabot of 18 antenna array distribution front and back.	21
2.3	Microwave radar array sensor reflectance signals	22
2.4	Millimeter wave radar sensor front and back.	23
2.5	Millimeter wave radar sensor front and back.	24
2.6	The drone image, left is a fixed-wing drone, and right is a Hover drone.	25
2.7	Five different kinds of biochars with soil mixture, the left column is the microwave in time domain signal, and the right column is the signal in the frequency domain.	28
2.8	Five different kinds of biochar with soil mixtures.	30
2.9	LDA results for different types of the biochar measurements.	31
2.10	LDA 3D results for different biochar types measurements.	32
2.11	Confusion matrix of different amounts of biochar recognition.	33
2.12	Five different water content levels with biochar mixture.	34
2.13	LDA 3D results for different water content levels classifications.	34

2.14	Microwave reflective signals in the frequency domain before and after 5 hours, respectively.	35
2.15	Original datasets scatter plot with all classes. Different amount of soil carbon.	38
2.16	Five different types of biochar.	39
2.17	Confusion matrix for all experiments.	39
2.18	Original datasets scatter plot with all classes. Different types of carbon.	40
2.19	Original datasets scatter plot with all classes. Different amounts of soil moisture.	41
2.20	Original dataset scatter plot with all classes. Different moisture levels with the same soil carbon.	42
2.21	(a) Replicated treatment plots (2 m x 12 m with a 2 m buffer) were laid out in a randomized block design in four blocks. biochar (10 t/ha), compost (20 t/ha), and co-compost (17.5 t/ha) were surface applied, and mixed. (b) Control, biochar, compost, and co-compost plots in one replicated block.	45
2.22	The TIR imagery is cropped in this diagram to enclose only the biochar plots. The geolocations of the ground truth measurements are identified and used to fit an experimental exponential semivariogram with a lag of 700 pixels. The resulting semivariogram is used to kriging the ground truth measurements to match the TIR imagery size. Lastly, the pixel-by-pixel correlation plot between the cropped TIR image and ground truth kriging is constructed with the best-fit line. This best-fit line's y -intercept is then corrected using the pixel intensity and a known temperature ice bath.	47
2.23	VARI vs. temperature. The x -axis indicates the VARI and NDVI values and the y -axis indicates the ground truth of soil temperature, The dark, red, blue, and green lines are represent control, biochar, manure, and mixture treatments, respectively.	49

2.24	NDVI vs. temperature. The x -axis indicates the VARI and NDVI values and the y -axis indicates the ground truth of soil temperature, The dark, red, blue, and green lines represent by control, biochar, manure, and mixture treatments, respectively.	50
2.25	VWC at 4cm. The x -axis the NDVI value, and the y -axis is the VWC (range from 0 to 2. VWC > 1 means water is greater than soil in the same unit, the same unit of y -axis for the rest of the VWC figures.	50
2.26	VWC at 7.5cm. The x -axis is the NDVI value, and the y -axis is the VWC.	51
2.27	VWC at 15cm. The x -axis is the NDVI value, and the y -axis is the VWC.	51
2.28	a) The R^2 and STD for the VARI and NDVI versus soil temperature. b) The R^2 and STD of NDVI versus three different depths of the VWC.	53
2.29	a) The general mountain bike with 26 inches tires,b) The electrical motor kit with battery, display, and controller.	56
2.30	(a) A CAD model of the finalized chassis prototype, (b) The assembly connecting a spanning rail to the top tube of the bicycle, (c) Adjustable screws controlling the tensions at each terminal of the T-joint, (d) A minimal superposition of chassis geometry with component numbers as in (a), (e) Chassis is equipped on our smart Bi-eBikes.	58
2.31	a) The STM32 board of motor controller and its extendable platform for related sensors, b) Switches and reserved holes.	60
2.32	Field road recognition, green color indicates true positive, red indicates false positive, and purple represents false negative.	61
2.33	Mission Planner windows and path planning implementation.	62
3.1	System hardware, left is the Raspberry Pi 4B with visualization user interface, and right is the Walabot microwave transmitter and receivers.	71

3.2	Millimeter wave radar sensor with Raspberry Pi Setup.	72
3.3	Millimeter wave reflectance signal analysis using PDP. (a) is the empty sensing signals (sensing background), (b)(c)(d)(e)(f) are the millimeter wave reflectance signals of different biochar materials. . .	74
3.4	Fractional order Gaussian distribution (in different H) vs. standard Gaussian distribution.	88
3.5	Fractional order Weibull distribution (in different ξ) vs. regular Weibull distribution.	89
3.6	Sampling experiments hardware setup.	90
3.7	The scatter plot of microwave signals for different percentage of carbon soil ratio.	91
3.8	The scatter plot of millimeter wave signals for different percentage of carbon soil ratio.	92
3.9	The scatter plot of microwave signals for different types of biochar (soil carbon).	94
3.10	Five different types of biochar (soil carbon).	95
3.11	The scatter plot of millimeter wave signals for different types of carbon.	97
3.12	The scatter plot of microwave signals for different levels of soil moisture.	98
3.13	The scatter plot of millimeter wave signals for different levels of soil moisture.	98
3.14	The scatter plot of microwave signals for different levels of soil moisture with soil carbon effects.	100
3.15	The scatter plot of millimeter wave signals for different levels of soil moisture with soil carbon effects.	101
3.16	Different containers with the same level of soil carbon content. . . .	103

3.17	The scatter plot of microwave signals for different container backgrounds.	103
3.18	The scatter plot of millimeter wave signals for different container backgrounds.	104
3.19	Confusion matrix of microwave signal classification using optimized DNN.	106
3.20	Confusion matrix of millimeter wave signal classification using optimized DNN.	107
3.21	Millimeter wave reflectance signal analysis of five different backgrounds comparison using PDP.	107
3.22	Optimal parameters tuning of a total of seven random distributions RMSE performance for SCN regression models.	109
3.23	Optimal parameters tuning of a total of seven random distributions RMSE performance for SCN regression models.	110
3.24	Confusion matrix of classification mission for training and testing generated by optimal fractional and non-fractional distributions. . .	111
3.25	Confusion matrix of classification mission for training and testing generated by optimal fractional and non-fractional distributions. . .	112
3.26	The performance of RMSE using tuned optimal parameters for all distributions. L represented the total number of hidden layers of SCN generated, and the y -axis represented the converted RMSE value by logarithmic.	112
3.27	The performance of classification accuracy and their RMSE using tuned optimal parameters for all distributions. L represented the total number of hidden layers of SCN used.	113
4.1	Digital twin in methane emissions quantification analysis.	119
4.2	Mobile robots perform various tasks in a dairy farm.	121
4.3	UAV waypoint flight path within UAV toolbox.	122

4.4	Swarming UAVs in Gazebo.	123
4.5	Methane diffusion process.	123
4.6	SCMDT-Digital twin in soil carbon accounting.	127
4.7	Proximity radar sensor, left is microwave radar sensor, right is millimeter wave radar sensor.	128
4.9	UAV and UGV for sensing carbon emissions diagram.	134
4.10	UAV and UGV sensing system.	135
4.11	The payload options of the mobile platforms. The biochar, microwave radar, millimeter wave radar, battery, and spray tank are shown individually.	136
4.12	Full trajectory of the UAV and UGV survey missions of size 10×20 m where the starting point is denoted by a magenta square. The black squares denote the UGV trajectory, while the blue line represents the UAV trajectory. The landing sequence of the UAV onto the UGV is denoted by red circles.	138
4.13	Landing sequence trajectory of a UAV and UGV survey mission. UGV trajectory is represented by black squares, while UAV point mass trajectory is represented by blue circles. The landing sequence of the UAV onto the UGV is denoted by red circles.	140
4.14	Controller performance for X , Y , and Z directions of the UAV towards landing on the UGV.	140
4.15	Testing a landing vision system with 5 aruco markers using the OpenCV Python library for detection. The coordinates denoted in the upper left corner represent the distance of the calculated centroid of all markers from the camera principle point. The red circle represents the calculated centroid of all aruco markers combined.	141
4.16	Testing the landing vision system with the camera close to the board observing a single aruco marker.	141

5.1	XAIoT empowered digital twin in soil carbon emissions quantification.	145
5.2	The Shapley value distributions of the microwave and millimeter wave signals that come from 20% of the soil and carbon mixture. . .	148
5.3	The Shapley value distributions of the microwave and millimeter wave signals that come from 20% soil moisture.	148
5.4	Drone battery 9500 mAh with high discharge rate.	151
5.5	Load circuit for battery capacity draining along with mm-wave radar sensor sampling hardware setup.	152
5.6	Original mm-wave data from varying battery capacity.	154
5.7	Confusion matrix of SOC classification of Li-Po battery	155
5.8	Comparison of mm-wave signal energy and battery voltage drops .	156
5.9	Failure case of using mm-wave sensing.	157
5.10	A failure case of using mm-wave radar sensing.	158
6.1	Spreading biochar mulch under wind field	167
6.2	Initialization of static mesh sensors, methane source and mobile actuators.	170
6.3	The diffusion behaviours under different α	172
6.4	Two numerical solutions	173
6.5	The diffusion behaviours under different β	174
6.6	The diffusion behaviours when $\alpha = 0.8$ and $\beta = 1.8$. Note that when $t = 2.25s$, biochar mulch spreading control is turned on.	175
6.7	Comparison different control process under different R , K , Q and no control baseline.	177

6.8	The total amount of biochar mulch sprayed based on different weights configuration	178
6.9	(a) The methane is diffusing, as shown in the green area under the wind field, and it is about to get the most concentration. (b) The sensors detect methane emissions and send them to UAVs. The UAVs start to spread biochar mulch as shown in the yellow area to curb the methane. The mulch is considered they have weights and can be affected by wind. (c) The control ends at the end of the simulation time.	179
6.10	(a)(b)(c) The multiple methane sources are diffusing as shown in the green area under the wind field, and it is about to get close to each other. (d) The sensors detect methane emissions and send them to UAVs. The UAVs start to spread biochar mulch as shown in the yellow area to curb the methane. (e) The control ends at the end of the simulation time.	180
6.11	(a)(b) The four methane sources are diffusing as shown in the green area under the wind field and are about to get close to each other. (c)(d) The UAVs start to spread biochar mulch as shown in the yellow area to curb the methane, and it will be eliminated at the end.	182
7.1	GHG sensing system framework.	187
7.2	GHG sensing system hardware setup. The left is our CO ₂ source, and the upper right is a proximal sensor with a Raspberry Pi and touchscreen. The lower right is a sensing chamber that contains CO ₂ airflow. The CO ₂ sensor is used to calibrate our system.	188
7.3	GHG sensing film. The W_a and L_a are silver antenna width (1 cm) and length (3 cm), respectively. The L_c and W_c are 1 cm and 1.6 cm, respectively.	189
7.4	Savitzky-Golay smoothing filtering.	191
7.5	a) The signal energy variations of different CO ₂ concentrations in time domain. b) The amplitude of different CO ₂ concentrations in frequency domain.	192

7.6	The resolution of amplitude after FFT per 100 PPM.	193
7.7	a) The signal energy variations of saturated CO ₂ concentrations in time domain with different exposure times. b) The amplitude of saturated CO ₂ concentrations in frequency domain.	194
7.8	Spectrometer hardware and captured spectrum.	198
7.9	Spectrum responses of surface soil moisture levels at 0%, 25%, 50%, 75% and 100%. The <i>y</i> -axis is light intensity, and <i>x</i> -axis is light wavelength.	200
7.10	Five different surface moisture levels and uniform distributions. (a) The moisture level increased one by one from the left, as the dry level, to the bottom, as the well-saturated level. (b) Sensing processing: SMA905 is capturing the reflectance signals from sunlight; the distance between the optical fiber and the sample is fixed at around 1 cm.	201
7.11	Kriging results of fitting captured spectrum absorption peak values.	202
7.12	Surface soil moisture in non-uniform distributions.	203
7.13	Kriging results of non-uniform distributions.	204
7.14	Regression analysis of spectrum absorption peaks intensity with surface soil moisture. (a) the first peak analysis for low, middle, and high moisture, and (b) the second peak analysis for low, middle, and high moisture, respectively.	206

LIST OF TABLES

1.1	Comparison of remote sensing platforms for SOC monitoring. . . .	14
1.2	Studies for SOC estimation with the use of spaceborne platforms. .	15
1.3	Studies for SOC estimation with the use of airborne platforms. . . .	16
1.4	Summary of proximal sensors for soil carbon estimation	18
1.5	Summary of benefits and drawbacks of remote and proximity sensing for soil organic carbon estimation	19
2.1	40 antenna pairs combinations, Tx and Rx stand for transmitter and receiver, respectively.	22
2.2	IMAGEVK-74 supports 3 predefined user-selectable recording profiles.	25
2.3	The comparison of water retention capacity for soil and soil with biochar mixture.	36
2.4	SBB system each component's price in detail.	64
3.1	Advantages and disadvantages of different classification algorithms.	79
3.2	Available hyperparameters options for all classifiers. N/A stands for not available.	80
3.3	Classification accuracy of different percentage of carbon soil ratio. .	92
3.4	The performance of optimized classifiers comparison for microwave signals of different percentage of carbon soil ratio.	93

3.5	The performance of optimized classifiers comparison for millimeter wave signals of different percentage of carbon soil ratio.	93
3.6	Classification accuracy of different types of biochar (soil carbon). .	95
3.7	The performance comparison of optimized classifiers for microwave signals of different types of biochar.	96
3.8	The performance comparison of optimized classifiers for millimeter wave signals of different types of biochar.	96
3.9	Classification accuracy of different moisture levels in the soil. . . .	99
3.10	The performance of optimized classifiers comparison for microwave signals of different moisture levels in the soil.	99
3.11	The performance of optimized classifiers comparison for millimeter wave signals of different moisture levels in the soil.	99
3.12	Classification accuracy of different moisture levels in soil under carbon effects.	101
3.13	The performance of optimized classifiers comparison for microwave signals of different levels of soil moisture with soil carbon effects. . .	102
3.14	The performance of optimized classifiers comparison for millimeter wave signals of different levels of soil moisture with soil carbon effects.	102
3.15	Classification accuracy of robustness investigation of sensing amount of soil carbon.	105
3.16	The performance of optimized classifiers comparison for microwave signals of sensing the amount of soil carbon (multi backgrounds). .	105
3.17	The performance of optimized classifiers comparison for millimeter wave signals of sensing the amount of soil carbon (multi backgrounds).	105
4.1	Selected nine machine learning algorithms with parameter setup. .	129

6.1	The weights of cost function for determining k_p , k_i , and k_s	176
6.2	The optimized k_p , k_i , and k_s under different weights configuration. .	177
7.1	Summary of GHG sensing system specification.	195
7.2	Technique details of Pleroma spectrometer.	199

ABSTRACT

Soil carbon content is vital to the way ecosystems function, and it plays an important role in soil properties in producing food, storing water, and mitigating climate change. Soil carbon is the key to unlocking the soil's various economic and environmental benefits, often known as the multi-functionality of soil carbon content. Human activities, particularly those that include the use of fire to burn biomass, deplete soil cover and result in immediate and enduring losses of soil organic carbon. We must minimize soil carbon loss due to erosion and greenhouse gas (GHG) emissions into the environment. As a way to capture and store carbon dioxide in the atmosphere, carbon sequestration can be used as a way to adapt to climate change because it has been shown to have positive effects on temperature, gross primary production, and the ability of soil to hold water. However, it is still a challenging problem to quantify the effect of how much carbon (biochar) is applied to the soil in terms of carbon sequestration accurately, quickly, cost-effectively, and on a large-scale field. Many studies use expensive element analyzers with intensive labor costs to measure soil carbon content. Several research efforts include utilizing artificial intelligence methods or applications to acquire carbon information from the soil and the atmosphere but lacking proper human interpretation (domain knowledge), the results might not be reliable. Therefore, Explainable Artificial Intelligence of the Internet of Things (XAIoT) is proposed to address the reliability issue and define a set of AI models that are interpretable by domain knowledge from environmental and engineering domain experts. Furthermore, the objectives that we would like to achieve are: 1) the XAIoT empowered microwave and millimeter-wave radar sensing as systematic data acquisition platforms for measuring soil carbon content; 2) in situ site-specific management (SSM) systems to manage soil carbon content and using machine learning algorithms; 3) the Digital Twin-enabled optimal soil carbon content management strategy in large-scale fields using mobile platforms such as UAVs and UGVs for the application of biochar and manure mixtures; 4) a cognitive battery management system for extending battery life for XAIoT smart sensing and smart actuation subsystems. The proposed applications and solutions have the positive potential to fight climate change, reduce carbon emissions, strengthen the quality of life on low-income and disadvantaged farms and adjacent communities, and identify means to gain acceptance among stakeholders in soil carbon management.

ABBREVIATIONS

AHS	Airborne Hyperspectral Sensor
AI	Artificial Intelligence
AIoT	Artificial Intelligence of Internet of Thing
AIoLT	Artificial Intelligence of Internet of Living Thing
ANOVA	Analysis of Variance
API	Application Programming Interface
BMS	Battery Management System
COTS	Commercially Off the Shelves
CPS	Cyber-Physical Systems
CSI	Camera Serial Interface
CW	Continuous Wave
CVT	Centroidal Voronoi Tessellations
DDS	Decision Support System
DI	Deionized
DT	Digital Twin
DNN	Deep Neural Network
DNA	Deoxyribonucleic Acid
ECM	Equivalent Circuit Model
EM waves	Electromagnetic Waves
EPA	Environmental Protection Agency

EV	Electric Vehicle
EVK	Evaluation Kit
FADE-MAS2D	Fractional Advection Diffusion Mobile Actuator and Sensor
FO-DIFF-MAS2D	Fractional Order Diffusion Mobile Actuator and Sensor
FDR	Frequency Domain Reflectometry
FFT	Fast Fourier Transform
FOV	Field of View
FOM	Fractional Order Models
GHG	Greenhouse Gas
GIS	Geographic Information Systems
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GPR	Ground Penetrating Radar
GPU	Graphics Processing Unit
IMU	Inertial Measurement Unit
INS	Inelastic Neutron Scattering
IoT	Internet of Things
IoLT	Internet of Living Things
IAI	Industrial Artificial Intelligence
KNN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
LIBS	Laser-induced Breakdown Spectroscopy
LR	Linear Regression
Li-Po	Lithium Polymer
MAS	Mobile Actuation and Sensor

MCU	Micro-controller Unit
MMT	Million Metric Tonnes
ML	Machine Learning
mm-wave	Millimeter wave
MOF	Metal Organic Framework
MPC	Model Predictive Control
NETs	Negative Emissions Technologies
NDVI	Normalized Difference Vegetation Index
NMR	Nuclear Magnetic Resonance
NIR	Near Infrared
NVM	Non Volatile Memory
PCA	Principal Component Analysis
PDE	Partial Differential Equation
PSS	Proximal Soil Sensing
PDF	Probability Density Function
PDP	Power Delay Profile
PI	Proportional Integral
PLSR	Partial Least Squares Regression
PPM	Parts Per Million
pXRF	Portable X-ray Fluorescence
QDA	Quadratic Discriminant Analysis
RAM	Random Access Memory
RBF	Radial Basis Function
RBW	Resolution Bandwidth
RF	Radio Frequency

RMS	Root Mean Square
RGB	Red Green Blue
ROS	Robotics Operational System
Rx	Receiver
RVFL	Random Vector Functional Link
S&AS	Smart and Autonomous Systems
SAW	Surface Acoustic Waves
SBB	Smart Bi-eBikes
SBC	Single Board Computer
SCN	Stochastic Configuration Network
SCMDT	Soil Carbon Management Digital Twin
SDK	Software Development Kit
SG	Savitzky-Golay
SNR	Signal to Noise Ratio
SLAM	Simultaneous Localization and Mapping
SOC	Soil Organic Carbon
SOC	State of Charge
SOH	State of Health
SoC	System-on-chip
SOM	Soil Organic Matter
SPM	Single Particle Model
SSM	Site-specific Management
STD	Standard Deviation
STM32	STMicroelectronics 32 Bits
SVM	Support Vector Machine

SVML	Support Vector Machine Linear
SVMR	Support Vector Machine RBF (Radial Basis Function)
SVMS	Support Vector Machine Sigmoid
TDR	Time Domain Reflectometry
TIR	Thermal Infrared
T_x	Transmitter
UAVs	Unmanned Aircraft Vehicles
UGVs	Unmanned Ground Vehicles
UV-B	Ultraviolet-B
UV-C	Ultraviolet-C
UV-VNIR-SWIR	Ultraviolet, visible, and infrared spectrum, short wave infrared
USB	Universal Serial Bus
UWB	Ultra-wide Band
VARI	Visible Atmospherically Resistant Index
VRT	Variable Rate Technology
VWC	Volume Water Content
WSAN	Wireless Sensor and Actuator Network
XAI	Explainable Artificial Intelligence
XAIoT	Explainable Artificial Intelligence of Internet of Things
XAIoT-DT	Explainable Artificial Intelligence of Internet of Things - Digital Twin
ZIF	Zeolitic Imidazolate Framework

Chapter 1

INTRODUCTION

1.1 Motivations

With the increase of the population, the corresponding demand for food is also increasing. This means that the soil is subjected to heavy farming activities, all of which will deprive the soil of its fertility, and worse, the loss of carbon from the soil can lead to failure to maintain soil health and climate change effects, for example, the water cycle in the soil, the maintenance of soil biodiversity, global warming, etc. Much research has proposed many management approaches from precision agriculture sector to meet these challenges [1–5]. However, few studies have concentrated on soil carbon loss and emissions, nor soil carbon content management. Environmental challenges such as climate change will be insurmountable for traditional agriculture, which dismisses the spatial and temporal complexity of the interacting factors that impact crop production and productivity.

The Internet of Things (IoT) technologies have been widely used in precision agriculture in recent years [6, 7]. IoT is described as the network of all items incorporated into devices, sensors, machines, software, and people that use the Internet to connect, share data, and interact to provide a holistic solution between the real and virtual worlds [8]. Compared with traditional agriculture, IoT-enabled agriculture applications such as irrigation and water quality monitoring, soil monitoring, pest control, crop and yield management, etc. [9] have determined the trends of agriculture 4.0 [10] which evolves towards smart and precision agriculture. As a result, when researching smart agricultural applications, two questions must be addressed in order to empower agricultural IoT devices:

- How smart is agricultural soil carbon sensing so that one can gain awareness not only of its current condition but also of its environmental development given the desired performance objectives?
- How does optimal sensing performance best serve the desired site-specific management requirements?

Addressing these questions is crucial to improving the sensing results of any IoT-enabled smart agricultural management applications. To do so, we introduced Artificial Intelligence plus the Internet of Things (AIoT) and Explainable Artificial

Intelligence plus the Internet of Things (XAIoT) to tackle the challenges. The AIoT comprises interconnected computing devices and equipment that can autonomously acquire, transport, and execute field or industrial operations while AI processes the collected data and assists in extracting the necessary actionable information. IoT boosts AI through machine learning and deep learning-based data analysis and learning capabilities, whereas IoT supplements AI with data acquisition, connectivity, and data exchange. Furthermore, to further convince human decision-maker, we have to introduce explanations of how machine learning (ML) algorithms work as part of having fairness, accountability, and transparency for the reasoning of their predictions, which can be explained in human-understandable ways through developing interpretable models, methods, and interfaces that are referred to as XAIoT.

Overall, the author proposes a systematic solution to tackle agricultural carbon emissions from smart sensing aspects by using XAIoT, which mainly focuses on sensors, sensing mechanisms, and sensing configurations. The author will justify the importance of agricultural carbon emissions, provide a definition of XAIoT and a review of the related research, and demonstrate smart sensing by combining XAIoT and its applications on agricultural carbon emissions.

1.1.1 What Are Soil Carbon and Its Impact?

Soil carbon is vital for fighting climate change, supplying water, and sustaining biodiversity. It is necessary for the important processes such as nutrient cycling and water filtration in the environment, which benefits not only plants but also humans. Human activities, particularly those that include the use of fire to burn biomass, deplete soil cover and result in immediate and enduring losses of soil organic carbon. Tillage and drainage both expose soil organic matter to oxygen and oxidation. We must minimize soil carbon loss due to erosion and greenhouse gas emissions into the environment, particularly carbon dioxide and other GHG.

The agriculture sector accounted for 593.4 MMT of CO₂ emissions in 2022 (EPA Report [11], Feb, 2024)¹. The primary contributor to nitrogen dioxide (N₂O) emissions in the United States was agricultural soil management, which involved fertilizer application and other practices designed to enhance nitrogen availability in the soil. This accounted for 75.2 percent of total N₂O emissions and 49.0 percent of emissions from the agriculture sector as a whole. Methane emissions originating from enteric fermentation and manure management constituted 27.4% and 9.2%, respectively, of the overall CH₄ emissions caused by human activities. These sources also accounted for 32.5% and 10.9% of the emissions from the agriculture sector (refer to Figure 1.1

¹In accordance with the present reporting obligations stipulated by the Paris Agreement and the United Nations Framework Convention on Climate Change (UNFCCC), this inventory report provides CO₂ equivalent values (or 9.4% of total U.S. greenhouse gas emissions) derived from the GWP values of the IPCC Fifth Assessment Report (AR5)

for further details). Among all categories of domesticated animals, beef and dairy cattle produced the most greenhouse gas emissions. Rice cultivation and agricultural residue field burning constituted negligible sources of CH_4 . Additionally, the management of manure and the field combustion of agricultural residues constituted negligible sources of N_2O emissions. Liming and urea fertilization contributed 0.1% and 0.06 percent, respectively, to total CO_2 emissions from anthropogenic activities. These high-level carbon emissions are supposed to be absorbed by soil as needed for carbon sequestration to increase soil health, which can improve agricultural productivity, increase climate resilience, and healthier soils make farms more resilient against both droughts and heavy rainfall.

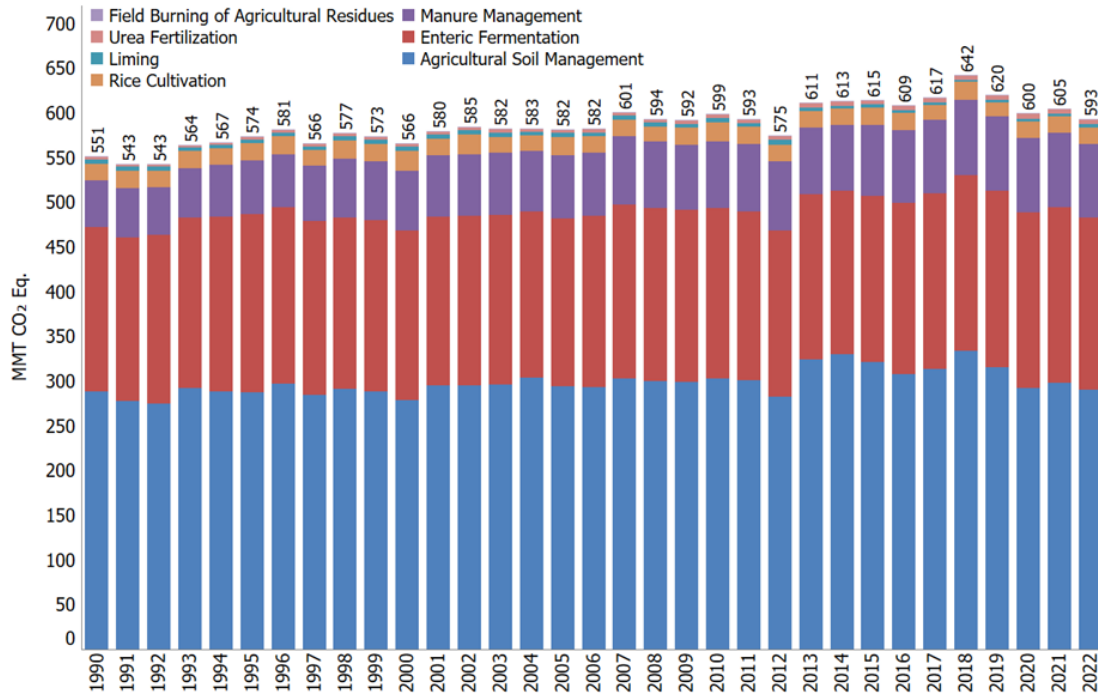


Figure 1.1: Trends in agriculture sector carbon emission sources [11].

Carbon sequestration is a process of capturing and storing atmospheric carbon dioxide, which means it can be used as a method for climate adaptation due to its established favorable impacts on soil erodibility, water retention capacity, temperature, and gross primary production [5]. Furthermore, sequestration of soil carbon on large-scale fields has the potential to significantly lower atmospheric GHG levels, improve the sustainability of working landscapes, and assure the availability of other ecosystem services such as water, food, and wildlife habitats [6]. Positive management of fields for increased carbon sequestration in soils, manure, and plants

is vital for California to improve climate resilience while lowering atmospheric GHG emissions as well.

One of the carbon sequestration techniques is applying biochar to soil. Biochar is a carbon-rich product of the thermochemical conversion of organic material that can be used as a soil amendment due to its resistance to decomposition, and its influence on nutrient dynamics. In other words, biochar can be applied to soil to improve soil functions and mitigate GHG from biomass that would otherwise decompose naturally into GHG. Through techniques like applying biochar production to the soil to minimize soil carbon loss, detecting which agricultural region is most in need has proven difficult. In other words, it is still a challenging problem to quantify the effect of biochar applied to the soil in terms of carbon sequestration accurately, quickly, cost-effectively, and on a large-scale field. This dissertation, for the first time, is to invent an extensive lab solution to demonstrate the fine-grained (detecting in local area) soil carbon sensing method (Microwave scanner) using a coarse-grained platform (Unmanned Aircraft Systems (UAS)).

1.1.2 What Is Smart Sensing in Estimation of the Soil Carbon?

The smart sensing in agricultural carbon emissions needs to consider the following characters, such as smart sensors, sensing mechanisms, and sensing configuration. Besides, according to the Smart and Autonomous Systems (S&AS) program of NSF (Natural Sciences Foundation) [12], the characteristics of a smart system are:

- **Cognizant:** should be self-aware of their own capabilities, understand their limitations, anticipate possible failures, undertake contingency planning, and recognize when they are operating incorrectly and react accordingly.
- **Taskable:** should be versatile, having the ability to accept high-level, possibly vague, instructions from humans and be capable of translating such goal-oriented directives into suitable plans for sensing, reasoning, communicating, and acting.
- **Adaptable:** should be capable of optimizing, reconfiguring, and repairing autonomously.
- **Ethical:** should adhere to an ethical system of societal and legal rules.
- **Knowledge-rich:** should be able to represent and reason with knowledge at multiple levels of abstraction.

Following the smart characteristics, we would empower the edge devices with AI to make them AIoT applications, and this would provide the comprehensive

smart sensors to solve the fundamental sensing problem in carbon emissions. Moreover, smart sensing not only relies on empowering the AI on sensors but also on the sensing mechanisms. In other words, for sensing and actuation, we also need to consider the cases of effective domains such as point-wise sensing and actuation, zone distributed sensing and actuation, and whole domain distributed sensing and actuation. Additionally, we can also consider sensors and actuators movable within the domain of interest. Therefore, the sensing problems need to ask the "WWWWWH" questions to get reasonable sensing results:

- Why Sense?
- What to Sense?
- Where to Sense?
- When to Sense?
- Who to Sense?
- How to Sense?

These questions lead to sensing mechanism exploration, where to put the smart sensors in the field, where are the optimal sensing locations, etc. Sensing configuration needs to consider how many sensors would be good enough to complete the sensing mission: one static sensor and one mobile sensor, or multiple mobile sensors and multiple static sensors to complete the sensing mission. The UAV and UGV, as our mobile platforms, play important roles in large scale field smart sensing. This will be explained in the following sections.

1.1.3 What Is Smart Application of the Soil Carbon?

Smart application in soil carbon (particularly biochar) management refers to the strategic use of cutting-edge digital technologies to optimize the application and effectiveness of biochar for carbon sequestration in soil. This innovative approach encompasses the deployment of AI, ML, the IoT, big data analytics, remote sensing, and geographic information systems (GIS) to not only apply biochar more effectively but also monitor and manage soil health and carbon levels with unprecedented precision. The goal is to harness these technologies to improve decision-making processes, enhance the sustainability of agricultural practices, and contribute significantly to climate change mitigation efforts.

The foundational characteristics that demarcate smart application within this domain are:

- **Precise Application:** Utilizing data analytics and AI to determine the optimal amounts, types, and timings for biochar application to different soil types, maximizing carbon sequestration and enhancing soil health without adverse effects.
- **Monitoring and Analysis:** Implementing IoT devices and remote sensing technology to continuously monitor soil conditions (e.g., moisture levels, carbon content, nutrient availability) post-biochar application, providing real-time data for timely adjustments.
- **Predictive Modeling:** Employing ML algorithms to predict the long-term effects of biochar on soil carbon sequestration, soil health, and crop yield, based on extensive datasets of soil characteristics, climate conditions, and biochar properties.
- **Decision Support Systems:** Offering interactive platforms that integrate GIS and spatial analysis with predictive insights, aiding stakeholders (farmers, land managers, policymakers) in making informed decisions about land use practices that enhance carbon sequestration and soil health.
- **Integrated Data Management:** Harmonizing diverse data sources, including satellite imagery, in-situ soil sensor data, and historical land use records, to create comprehensive models of soil carbon dynamics and biochar effects.

In essence, the smart application of soil carbon through biochar involves:

- **Optimizing Biochar Use:** Leveraging digital technologies to refine how biochar is applied, aiming for the most beneficial impact on carbon sequestration and soil vitality.
- **Data-Driven Insights:** Drawing on vast arrays of collected data to inform and enhance biochar application strategies, adapting to changing conditions and new research findings.
- **Sustainable Agricultural Practices:** Guiding stakeholders towards farming methods that bolster carbon sequestration, leveraging smart technologies to achieve a balance between agricultural productivity and environmental sustainability.

In practical terms, smart applications dedicated to soil carbon estimation could serve multiple purposes, such as: 1) Guiding farmers towards adopting agricultural methodologies that bolster soil health and augment carbon sequestration, including but not limited to cover cropping, minimal tillage practices, or rotational cropping,

2) Providing pivotal data on soil carbon reservoirs and dynamics to inform policy formulations related to land use, conservation strategies, and climate change abatement measures, and 3) Empowering researchers to investigate the repercussions of environmental alterations on soil carbon fluxes at various scales, from local to global.

Through the adoption of these cutting-edge technologies, the smart application of biochar in soil carbon management not only bolsters the role of biochar in carbon sequestration but also equips stakeholders with actionable insights for the pursuit of sustainable land management and the broader objective of climate change mitigation.

1.1.4 What Is XAIoT?

XAIoT stands for explainable artificial intelligence and the internet of things. The first time of the term XAI was used was in 2004 [13]. XAI can be defined as an effort to make AI systems more understandable to humans; however, there is no accepted technical definition of XAI at this time, and more clarity and consistency are needed in terms of the terminology in use [14–16]. Transparency, interpretability, and explainability are frequently used interchangeably, which is one of the concerns.

Interpretability relates to the extent to which a model may be comprehended [17], and it is also sometimes used in place of “explainability” [16]. Transparency alternatively pertains to a comprehensive attribute of providing decision makers with pertinent information about how the model operates; or it relates to the provision of such information. This comprises documentation of the training technique, analysis of training data distribution, code releases, and explanations at the feature level [14, 16, 18, 19]. Explainability provides insight into the rationale for the system’s decision-making, but is commonly confused with understandability, which was defined by consensus as “tools that empower a stakeholder to comprehend and, when necessary, debate the reasoning of model outputs.” In this work, we mainly focus on explainability [20].

AI systems, also often called black box systems, that make predictions without explanation are problematic for a number of reasons, including their lack of transparency and the fact that they conceal potential biases inside the system [21]. Numerous instances of bias in AI-based predictive systems have been identified. These systems have been demonstrated to reinforce social and historical human prejudices, with disproportionately harmful effects on traditionally underprivileged members of our society [22]. In [23], The authors concluded that present XAI algorithms cannot provide appropriate explanations for biased behavior in sensitive applications. Although, giving explanations can increase understanding of and trust in a system [15, 16, 24], Despite delivering explanations can boost users’ understanding and trust in a system, basic explanations might conceal the system’s negative characteristics and lead users to risky or unfounded conclusions that may eventually be unethical [17]. In other words, in agricultural carbon emission sector, providing

explainable results of carbon loss and where should put more carbon (biochar) using edge devices, as referred as XAIoT in agricultural need to be further explored.

The use of XAIoT in precision agricultural and related agricultural applications is rooted in demand for the added value that comes from agricultural and environmental professionals being able to understand how and why a machine-based decision has been made. Thus, there is growing demand for AI approaches that not only perform well, but are trustworthy, transparent, interpretable, and explainable for a human expert. This also has important implications for the public, policy, and governance as the explainability of AI tools will enhance the trust of agricultural and environmental professionals [25]. In many ways, agricultural professionals usually plays a guidance roles to the farmers and markets about the decision-making that is too complex for farmers to understand and act on. However, not many mature edge devices, and not many research studies [26] applied XAI on edge devices which towards to XAIoT.

Therefore, in this dissertation, we focus on XAI on edge devices (towards to XAIoT) as a technician solution that can help to expose systemic bias in agricultural carbon emissions. XAIoT can be used to help stakeholders evaluate if predictions are biased and defend algorithmic decisions as being fair and ethical [14–16, 27].

1.2 Dissertation Objectives and Methods

XAIoT is a new field with a lot of potential for agricultural carbon emission and biochar applications. The main goal of this dissertation is to create a methodological framework for XAIoT-enabled smart sensing of agricultural carbon emissions that is supported by: 1) XAIoT-enabled smart data acquisition platforms, such as microwave radar, millimeter wave radar, UAVs, and UGVs; 2) Site Specific Management and cognitive battery management systems to optimize XAIoT system extendable health and application on large-scale field sensing.

In order to accomplish the main objectives, the XAIoT enabled smart sensing for agricultural carbon emissions will be grouped into the following specific parts:

1. Employing XAI on microwave, millimeter wave radar sensors, UAVs and UGVs towards XAIoT empowered smart data acquisition platforms.
2. Establishing a set of assessment criteria to evaluate how smart and developmental the radar sensing is under site-specific management (SSM) framework.
3. Validating the XAIoT carbon emission sensing framework by simulation and experimental results on thermal systems.
4. Designing a cognitive battery management system for XAIoT carbon emissions sensing guarantees sensing accuracy, sustainability, and scalability for edge devices, UAVs, and UGVs.

5. Designing a site-specific management strategy for carbon emissions sensing, including quantification, in-situ sensing, real-time decision making, and deployment of swarm mobile sensors and actuators.
6. Developing advanced virtual representation models for process soil methane emission suppression control using UAVs.

To expand on the methodological framework for estimating soil carbon within the context of XAIoT-enabled smart sensing, the dissertation employs a multifaceted approach that integrates cutting-edge technologies and innovative strategies. This comprehensive method aims to enhance the accuracy, efficiency, and scalability of soil carbon estimation in agricultural settings, leveraging the capabilities of XAIoT to address the challenges associated with carbon emissions sensing and management. The methodology is structured around several key components, each designed to contribute to the overarching goal of advancing agricultural carbon emissions sensing.

XAIoT Empowered Smart Data Acquisition Platforms: A cornerstone of the methodological framework is the employment of XAI on various sensors and platforms, including microwave, millimeter wave radar sensors, UAVs, and UGVs. These XAIoT-empowered platforms facilitate smart data acquisition, enabling the collection of high-quality data on soil carbon content and environmental conditions. By integrating explainable AI, the research aims to improve the transparency and trustworthiness of the data acquisition process, allowing for more accurate and reliable soil carbon estimation.

Assessment Criteria: Establishing a set of XAIoT assessment criteria is essential for evaluating the smartness and development of the radar sensing technology under the SSM framework. These criteria assess the accuracy, reliability, and sustainability of the sensing technologies and strategies employed, providing a benchmark for continuous improvement.

Site-Specific Management Strategy: Designing a SSM strategy is crucial for targeted and effective carbon emissions sensing. This component involves the quantification of soil carbon content, in-situ sensing, real-time decision-making, and the deployment of a swarm of mobile sensors and actuators. The SSM strategy is tailored to the specific needs and characteristics of the agricultural site, ensuring that sensing activities are optimized for accuracy and efficiency.

Cognitive Battery Management System: A cognitive battery management system is proposed to ensure the sustainability and scalability of XAIoT carbon emissions sensing. This system manages the power supply of edge devices, UAVs, and UGVs, optimizing their operation time while minimizing the risk of data bias due to power shortages or failures. By intelligently managing battery life, the system supports continuous and reliable sensing activities, contributing to the overall effectiveness of the soil carbon estimation framework.

Advanced Virtual Representation Models: The development of advanced virtual representation models is pivotal to the methodology. These models are designed to simulate soil methane emission suppression control using UAVs, providing a virtual platform for testing and optimizing site-specific management strategies. By employing XAI to enhance the interpretability and effectiveness of these models, the research ensures that stakeholders can make informed decisions based on the insights generated.

Validation Through Simulation and Experimental Results: The final component of the methodology involves validating the XAIoT carbon emission sensing framework through simulation and experimental results on thermal systems. This validation process ensures that the proposed methods and technologies effectively estimate soil carbon content, contributing to the accuracy and reliability of agricultural carbon emissions sensing.

By integrating these components into a cohesive methodological framework, the dissertation aims to advance the field of XAIoT-enabled smart sensing for agricultural carbon emissions. The research not only addresses the technical aspects of soil carbon estimation but also considers the sustainability and scalability of the sensing solutions, ensuring their applicability and effectiveness in real-world agricultural settings.

1.3 Dissertation Outline

The dissertation delineates an innovative methodological framework for XAIoT that enabled smart sensing of agricultural carbon emissions. This work is situated at the confluence of XAIoT and agricultural science, aiming to advance the precision and effectiveness of carbon emission sensing in agricultural contexts. The overarching objective is to harness the potential of XAIoT to develop a comprehensive system that not only acquires data smartly but also manages and processes this data to optimize agricultural practices, thus contributing to the mitigation of climate change impacts. The dissertation is structured into several chapters, each addressing a critical component of the XAIoT-enabled smart sensing framework.

In Chapter 1, the concept of XAIoT is introduced and its application in smart sensing for agricultural carbon emissions is discussed. The significance of soil carbon in the broader context of climate change mitigation and soil health is underscored. A review of existing literature on soil carbon, coupled with an exploration of the opportunities and challenges presented by XAIoT in carbon emission sensing, sets the stage for subsequent discussions. Definitions of smart sensing and smart applications are articulated, alongside the delineation of future research directions in this domain.

Chapter 2 focuses on smart data acquisition platforms. It proposes the use of advanced edge devices, such as microwave and millimeter wave radar sensors,

alongside UAV and UGV platforms for remote and proximity sensing data collection. A case study provides practical insights into the application of these technologies.

In Chapter 3, the dissertation explores non-intrusive methods for quantifying soil carbon content using machine learning algorithms. The mechanics of common classifiers and lightweight model stochastic configuration networks are discussed in detail, supported by a comprehensive evaluation of proximal sensors' performance.

Chapter 5 delves into the role of XAI in empowering digital twins for managing soil carbon emissions via proximal sensing. A case study on cognitive battery management within the XAIoT framework highlights the necessity of sustainable power management for edge devices, UAVs, and UGVs to ensure data accuracy and mitigate additional carbon emissions.

Chapter 4 discusses the SSM strategy for soil carbon content sensing. The integration of MAS-net and the Internet of Living Things (IoLT) is examined for their impact on sensing accuracy. The potential of Digital Twins to facilitate real-time data exchange between physical and virtual entities is highlighted as a promising area for future work.

Chapter 6 presents a simulation platform for controlling soil carbon emissions using automatic drone swarms to apply biochar. The setup allows for the testing of SSM strategies and considers environmental factors such as wind field to simulate greenhouse gas emission variations accurately.

Chapter 7 explores XAIoT-enabled smart applications for measuring CO₂ concentrations and soil moisture and their integration into UAV systems to mitigate methane emissions.

Finally, Chapter 8 provides a comprehensive evaluation of the proposed XAIoT framework through real soil carbon data testing. Concluding remarks underscore the dissertation's contributions to the field and outline avenues for future research.

1.4 Dissertation Contributions

The main contribution of the dissertation is to have established foundation for using XAIoT for smart sensing in agricultural carbon emissions applications. The XAIoT sensing framework was created to enable the site-specific management of carbon emissions using smart data acquisition platforms and advanced data analytics. Specifically, the following contributions are listed below:

- A developing framework for XAIoT-empowered radar array sensors for sensing carbon emissions is applicable to any agricultural carbon sensing scenario.
- A set of data acquisition platforms such as UAVs, UGVs, and non-intrusive sensors such as microwave and millimeter wave radar sensors.
- A set of machine learning algorithms have been investigated and compared for improve assessment of identifying the microwave and millimeter wave signals.

- A simulation platform for assessing the sensing performance of the network of mobile sensors and actuators.
- The site-specific management strategy for agricultural carbon emissions utilizes the Digital Twin concept.
- Cognitive battery management system for guaranteed the XAIoT enabled smart sensing for agricultural carbon emissions with reliable sensing results.

1.5 Soil Organic Carbon Estimation: A Review

1.5.1 Introduction to Soil Carbon Sensing

The significance of soil carbon extends beyond its role as a crucial component of soil health; it is pivotal in agricultural productivity and climate change mitigation. Soil carbon serves as a key indicator of soil quality, influencing water retention, nutrient availability, and microbial activity, thus directly affecting plant growth and ecosystem sustainability [28, 29]. The measurement and monitoring of soil carbon content have evolved significantly with the advancement of sensing technologies marking a transition from traditional as shown in Figure 1.2

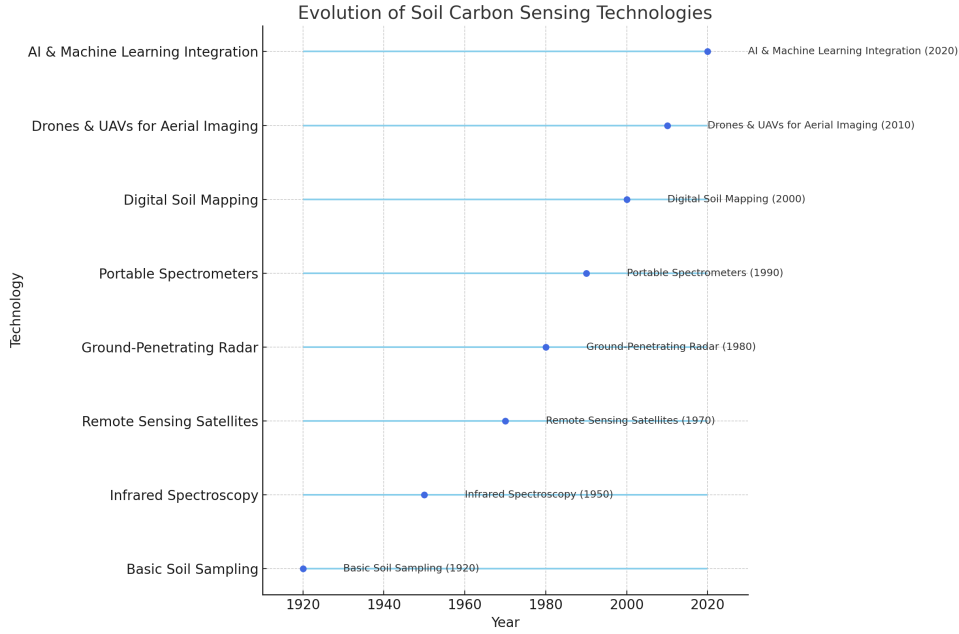


Figure 1.2: Evolution of soil carbon sensing technologies.

The plot above visualizes the evolution of soil carbon sensing technologies across different years, from the early 20th century to the present day. It outlines the progression from basic soil sampling methods to the latest advancements, including

basic Soil Sampling introduced in the 1920s [30,31]; Infrared Spectroscopy becoming prominent around the 1950s [32–34]; The launch of Remote Sensing Satellites in the 1970s [35–38]; The development of Ground-Penetrating Radar in the 1980s [39–43]; The introduction of Portable Spectrometers in the 1990s [44–46]; The emergence of Digital Soil Mapping in the 2000s [47–50]; The adoption of Drones & UAVs for Aerial Imaging in the 2010s [51–53] and the integration of AI & Machine Learning for data analysis in the 2020s [54–57].

1.5.2 Remote Sensing Technologies for Soil Carbon Estimation

Remote sensing has become an indispensable tool for soil carbon estimation, leveraging non-invasive techniques to evaluate and monitor soil organic carbon (SOC) levels efficiently over vast and varied landscapes [58,59]. By detecting soil’s reflectance and absorption characteristics, which vary with SOC content, remote sensing technologies such as multispectral and hyperspectral imaging enable the monitoring of soil carbon. These technologies exploit the influence of organic matter on soil’s optical properties to estimate SOC from aerial or satellite imagery, providing a means to assess soil carbon stocks and changes over time without extensive ground sampling [60].

The benefits of remote sensing include large-scale monitoring capabilities, cost-effectiveness by reducing the need for labor-intensive soil sampling, the ability to conduct temporal analyses by monitoring SOC changes over time, and the potential for integrated assessments when combined with other environmental data like land use and vegetation cover. However, challenges such as resolution limitations, calibration requirements with ground-truth data, and the complexity of data processing highlight the need for sophisticated algorithms and computational resources. Key studies by Angelopoulou et al. [61], Sharma et al. [62], and Singh et al. [63] have reviewed remote sensing techniques for SOC estimation, discussed the prediction ability of multispectral data with varying resolutions, and provided an overview of advances in remote sensing of the terrestrial carbon cycle, respectively, underscoring the evolution and potential of these technologies in carbon cycle studies.

1.5.2.1 Different Methodologies of Remote Sensing for Soil Carbon Estimation

Remote sensing techniques have revolutionized the way SOC is monitored and estimated, offering diverse platforms such as satellites, aircraft, and UAS to capture data over varying scales and resolutions [61]. These techniques enable the non-invasive assessment of SOC across different landscapes, leveraging the unique capabilities of each platform to gather critical environmental data. While satellites provide global coverage and regular monitoring capabilities, aircraft offer high-resolution imagery for detailed local-scale mapping, and UAS excels in ultra-high-resolution data collection for precise applications such as precision agriculture. The

choice of a specific remote sensing platform depends on the requirements of the SOC study, including the desired scale, resolution, budget, and the need to balance the benefits of broad coverage against the precision of high-resolution data. We summarized a Table 1.1 for comparing the specifics of these platforms in a table to highlight their altitude, resolution, coverage area, and some of the inherent benefits and drawbacks [64–71]:

Table 1.1: Comparison of remote sensing platforms for SOC monitoring.

Platform	Altitude	Spatial Resolution	Coverage Area	Benefits	Drawbacks
Satellite	400 km - 36,000 km	Coarse (> 1 km), Moderate (10-30 m), High (0.3-5 m)	Global	Global coverage, regular revisit times, diverse sensors	Limited resolution for small-scale variations, cloud cover issues, calibration needs
Aircraft	150 m - 10 km	< 1 m	Specific regions	High spatial resolution, flexibility, operates under cloud cover	Limited coverage, high costs, logistical challenges
UAS	< 120 m	< 30 cm (to a few cm)	Focused areas	Ultra-high-resolution, flexible deployment, lower costs	Limited by battery/weather, smaller coverage, regulatory restrictions

1.5.2.2 Estimation Results

Spaceborne sensors, encompassing multispectral, hyperspectral, and even simulated hyperspectral platforms, have progressively been utilized for SOC estimation, with research publications on this application increasing over time. These satellite sensors have been instrumental as auxiliary variables in mapping soil properties, offering more precision than traditional methods like ordinary kriging when paired with advanced geostatistical methods. For instance, Landsat 7 ETM+ data [72], especially the panchromatic band, were found to provide better SOC predictions compared to the normalized difference vegetation index (NDVI).

Hyperspectral data, such as from the now-deactivated Hyperion sensor, has shown promise in estimating and mapping SOC’s spatial variability. Gomez et al. [73] noted that while the Hyperion data offered useful insights similar to in-field VNIR–SWIR data, low Signal to Noise Ratio (SNR) and a 30 m spatial resolution were limiting factors. Castaldi et al.’s simulation studies [74] on forthcoming hyperspectral imagers such as EnMAP, PRISMA, and HypSIRI indicated that while resampled laboratory spectra could yield good results, actual satellite data would be necessary for more accurate estimations.

A key advancement was the use of data mining techniques to maximize bare soil area detection, leveraging the short revisit time of satellites. This approach allowed for improved reflectance retrieval from satellite images and, consequently, moderate accuracy in soil property predictions, including SOC, as demonstrated by Dematté [75] and Gallo’s work [76]. Gholizadeh et al. [77] further showcased the superiority of Sentinel-2 over airborne sensors in capturing SOC variability, particularly at higher levels, using a simple SVM model for the predictions.

In summary, while spaceborne sensors have shown significant potential for SOC estimation, the success of these methodologies hinges on several factors, including sensor resolution, atmospheric conditions, and the spectral characteristics of the soil. These methodologies and their respective benefits and limitations in terms of SOC estimation are detailed in the referenced studies and summarized in Table 1.2.

Table 1.2: Studies for SOC estimation with the use of spaceborne platforms.

Sensor	Spectral Range (nm)	Algorithms	R^2	RMSE (g·kg ⁻¹)	Reference
Hyperion	400–2500	PLSR	0.51	0.73	[60]
Landsat ETM+	450–2350	ANN5K	0.63	0.27	[78]
EnMAP	420–2500	PLSR	0.25–0.67	0.20–0.48	[79]
PRISMA	400–2500	PLSR	0.26–0.65	0.21–0.48	[79]
HyspIRI	380–2510	PLSR	0.23–0.60	0.22–0.48	[79]
EnMAP	420–2500	autoPLSR	0.67	2.8	[80]
Sentinel-2	440–2200	PLSR/RF	-	1.9–25.2/0.20–18.6	[74]
Sentinel-2	440–2200	PLSR	0.56	1.23	[73]
Sentinel-2	440–2200	SVM	-	0.08–0.24	[77]

In the pursuit of SOC estimation using airborne platforms, researchers have investigated various methodologies, each with distinct benefits and concerns. Stevens et al. [81] examined the Airborne Hyperspectral Sensor 160 (AHS-160) across different soil types, correlating spectra with soil samples to estimate SOC content. The study highlighted the influence of soil type heterogeneity on spectral reflectance and the need for local calibration due to soil type variations. Challenges included atmospheric correction, weather conditions, and sensor stability issues. Even with these problems, airborne hyperspectral imaging was seen as a promising way to estimate SOC, and changes to external parameters could make the model more accurate.

Other studies like those by Hbirkou et al. [82] and Franceschini et al. [83] focused on the effects of soil roughness, vegetation cover, and spectral mixture on SOC prediction models. These studies underscored the importance of controlled conditions, such as dry weather for reduced moisture content, and strategies for distinguishing between soil and vegetation or straw residues in spectral data. While laboratory conditions yielded more accurate organic matter content predictions, airborne hyperspectral sensors showed potential when external factors were accounted for.

The research consistently points to the necessity of atmospheric correction, as well as the alignment between remote and laboratory spectral measurements for reliable SOC estimation. Approaches that leverage existing large soil spectral libraries, such as the one by Castaldi et al. [84], suggest that airborne estimations can be time-effective without the need for additional laboratory measurements. However, the calibration of prediction models remains a crucial step that benefits from field data collection. These efforts illustrate the complexity of SOC estimation from airborne sensors and the various factors that must be meticulously controlled to improve accuracy and reliability. The details and specific results of these studies are encapsulated in referenced Table 1.3, which provides a more comprehensive comparison of methodologies and outcomes.

Table 1.3: Studies for SOC estimation with the use of airborne platforms.

Sensor	Spectral Range (nm)	Algorithms	R^2	RMSE ($\text{g}\cdot\text{kg}^{-1}$)	Reference
AHS-160	430–2540	PLSR, PSR, SVMR	0.53–0.89	3.13–6.22	[81]
AHS-160	430–2540	PLSR	-	1.7	[85]
HyMap	450–2500	PLSR	0.34–0.83	0.76–1.10	[82]
ProSpec TIR V-S	400–2500	PLSR	0.33	3.82	[83]
AHS-160	430–2540	PLSR	0.62	1.34	[86]
AISA-Eagle	400–1000	PLSR	0.44	4.05	[87]
AHS-160	430–2540	SLR, SMLR, PLSR	0.27–0.60	6.44–8.70	[88]
AISA Dual system	400–2450	SML	0.73	8.4	[89]
APEX	400–2500	PLSR	-	4.3	[84]
HyMap	450–2500	PLSR	0.73–0.85	0.19–0.25	[90]

In the realm of environmental and climate variable assessment, UAS has made significant strides, yet their integration into soil ecosystem monitoring has been limited. To date, only a solitary study [91] has been identified that explores SOC estimation using this technology. This research employed a multispectral Mini-MCA6 sensor by Tetracam Inc., with a spectral range of 450–1050 nm, mounted on a UAS to determine its effectiveness in predicting SOC levels. A meticulous operational plan was executed to ensure optimal conditions that mitigated the impact of soil moisture and surface roughness on the readings. The flight campaign was conducted under ideal circumstances—a cloudless sky, scant vegetation cover, and dry soil—to minimize the influence of soil moisture on the results.

The methodology adopted in the study showed considerable promise for SOC monitoring, utilizing an SVM algorithm. It achieved a mean coefficient of determination (R^2) of 0.95 and a Root Mean Square Error (RMSE) of 0.21% in cross-validation across 161 soil samples, demonstrating performance on par with standard dry combustion laboratory methods. However, it was noted that there was a tendency for overestimation at lower SOC concentrations and underestimation at higher SOC values.

1.5.3 Proximal Sensing in Estimation of Soil Carbon

Proximal soil sensing (PSS) is a technique involving the use of sensors in direct contact with or close to the soil for measuring a range of soil properties [92]. This method offers a quick, on-site analysis of soil attributes, reducing the reliance on labor-intensive lab tests. PSS is equipped with an array of sensors suited for different soil parameters, such as moisture, texture, compaction, and organic matter.

The advantages of PSS include its ability to provide high-resolution data that reflects the spatial heterogeneity of soil properties within a small area. It also facilitates real-time analysis, enabling immediate soil management decisions. This approach is often more cost-effective than traditional sampling, especially for regular monitoring, and it minimizes soil disturbance, thus maintaining its natural state [93].

Proximal sensors range from electromagnetic and optical sensors that infer soil properties through electrical conductivity and reflectance, to acoustic sensors and penetrometers that evaluate compaction [94]. These sensors are used in various formats, from handheld devices for spot checks to vehicle-mounted systems for wider area coverage, and even robotic systems for autonomous soil mapping.

1.5.3.1 Different Types of The Proximal Sensors

A summary and detailed explanation of various proximal sensor technologies used for estimating soil carbon content is presented in this section. The types of sensors discussed include digital cameras for soil color analysis, spectroscopy for chemical composition, laser-induced breakdown spectroscopy, inelastic neutron scattering, and ground-penetrating radar, each providing unique benefits for soil analysis.

1. Soil Color (Digital Camera [95,96]): Digital cameras capture the visible spectrum of the soil to infer carbon content. The color intensity and hue variations can indicate differences in SOC due to the darkening effect of organic matter. This method is non-destructive and provides immediate results but can be influenced by moisture, texture, and mineral content.
2. VIS-NIR Spectroscopy [71, 97]: This technique uses visible, near, and mid-infrared spectroscopy to detect SOC by measuring the reflectance of different wavelengths. It's a rapid method that provides information on various soil attributes, with mid-infrared offering the most direct measurement of SOC. While highly effective, its accuracy can be affected by soil moisture and surface roughness.
3. Laser-induced Breakdown Spectroscopy (LIBS [98]): LIBS uses a high-energy laser pulse to create a plasma on the soil's surface, then analyzes the light emitted to determine the concentration of elements, including carbon. It can analyze both the surface and subsurface layers and is effective in detecting

SOC. However, it requires careful calibration and can be sensitive to soil heterogeneity.

4. Inelastic Neutron Scattering (INS [99, 100]): INS involves bombarding the soil with high-energy neutrons and detecting scattered gamma rays, which provide a measure of carbon content within a large soil volume. It penetrates to depths of 20-30 cm and is suitable for non-invasive SOC assessment over large areas. The technique is less affected by soil moisture content but requires specialized equipment and safety precautions due to the use of high-energy particles.
5. Ground Penetrating Radar (GPR [101, 102]): GPR sends radio waves into the ground and analyzes the reflected signals to assess soil properties. Lower frequencies can penetrate up to 10 m, making them suitable for deep soil analysis, while higher frequencies offer higher resolution up to 1 m deep. GPR's ability to estimate SOC is indirect, often requiring correlation with physical properties that affect SOC, such as soil layering or water content.

These proximal sensing techniques offer diverse ways to estimate soil carbon, each with its benefits and limitations. While some provide high-resolution data and are cost-effective, others can analyze large volumes non-invasively. Their effectiveness can vary based on soil type, condition, and the specific sensor technology used [103]. The more details are in Table 1.4.

Table 1.4: Summary of proximal sensors for soil carbon estimation

Sensor Type	Characteristics
Soil Colour (Digital camera)	Surface analysis
Soil VIS-NIR Spectroscopy	vis: 400–700 nm, NIR: 700–2500 nm, mid-IR: 2500–25000 nm
Laser-induced Breakdown Spectroscopy (LIBS)	190 to 1000 nm
Inelastic Neutron Scattering (INS)	Volumetric sampling up to 0.3 m ³ , 20–30 cm depth, 150 cm footprint
Ground Penetrating Radar (GPR)	Lower freq.: 100 MHz to 200 MHz – 10 m, Higher freq.: 1 GHz to 2.6 GHz – 1 m

In conclusion, PSS marks a significant leap in soil science, providing a practical and scalable solution for soil analysis. With the ongoing advancement in technology, the integration of PSS data with digital soil mapping and modeling promises to further augment our understanding and stewardship of soil resources. In addition, the dissertation is focused on the PSS method and presents an entire system and comprehensive methodologies to evaluate the soil carbon in situ, and in real-time.

	Remote Sensing	Proximity Sensing
Coverage	Global to regional scale, allowing extensive area monitoring	Local scale, suitable for detailed, site-specific assessments.
Cost	Cost-effective for large-scale applications due to the free availability of satellite imagery (e.g., Copernicus program)	Potentially higher costs for equipment and on-the-ground operations, but effective for focused studies
Resolution	Limited by satellite or aerial sensor capabilities, affecting detail level	High-resolution data capturing fine-scale soil properties.
Data Acquisition	Non-destructive, rapid, and efficient over large areas.	Non-destructive and highly accurate but more time-consuming due to smaller coverage per unit time
Challenges	Affected by atmospheric conditions, soil moisture, vegetation cover, and the need for radiometric and geometric corrections	Limited by the need for physical access to the site, which may be hindered by terrain or land use
Applications	Suitable for mapping and monitoring SOC changes over extensive areas, identifying trends, and informing large-scale land management and policy	Ideal for in-depth soil health assessments, precision agriculture, and localized soil management practices

Table 1.5: Summary of benefits and drawbacks of remote and proximity sensing for soil organic carbon estimation

1.6 Results Reproducibility

The code and datasets of the results presented in this dissertation can be found in the website <https://andi00123456.github.com> where the users can download the codes for its use and evaluation.

Chapter 2

XAIOT EMPOWERED SMART DATA ACQUISITION PLATFORMS AND CASE STUDIES

2.1 The Microwave Radar Array Sensor

Using a radio frequency 3D sensor called Walabot Developer 2.2)(Vayyar Imaging Ltd, Yehud-Monosson, Israel¹), the reflectance of the surroundings being sensed can be obtained. The frequency range for the US/FCC model is 3.3 - 10 GHz while for the EU/CE model it is 6.3 - 8 GHz. The average transmitted energy of both models is less than -41 dB/MHz and poses no health risks to the user [104]. Multiple antennas allow Walabot to send, receive, and record microwave reflectance signals. (See Figure 2.1).

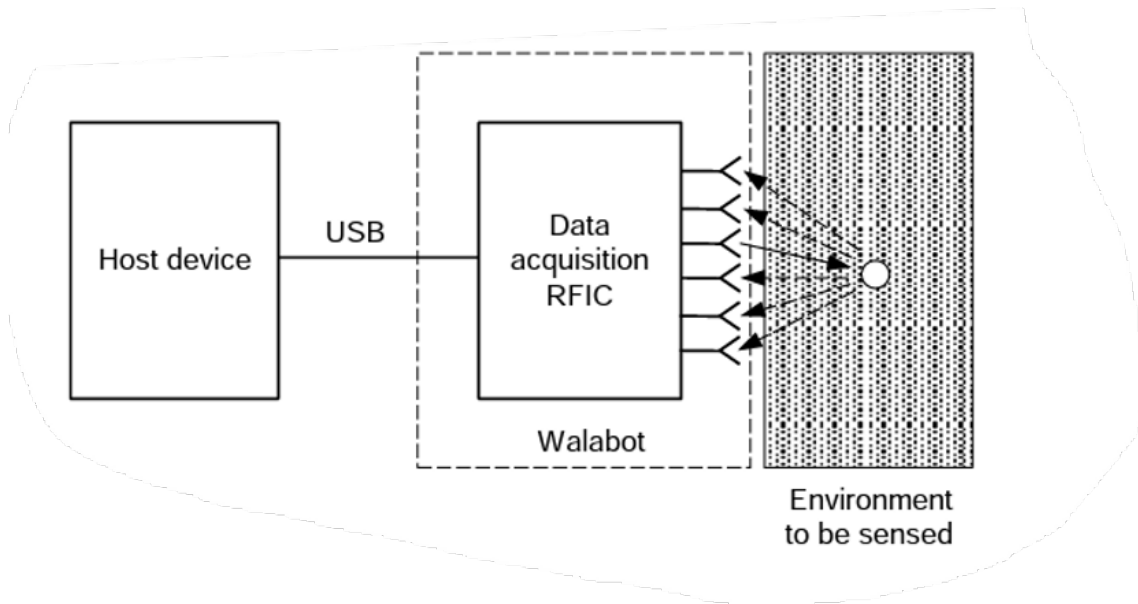


Figure 2.1: Walabot sensing strategy²

¹<https://walabot.com/>

²<https://site.walabot.com/docs/walabot-tech-brief-416>

Multiple transmit-receive antenna pairs' broadband recordings are processed to build a three-dimensional representation of the surroundings. Analysis of picture sequences permits the detection of environmental changes. Walabot is capable of short-range imaging in dielectric settings such as concrete and drywall. Then, researchers could detect environmental changes by examining image sequences. The Walabot employed an antenna array to transmit radio frequency to the region in front of it and then to receive the reflected signals (reflectance or response).

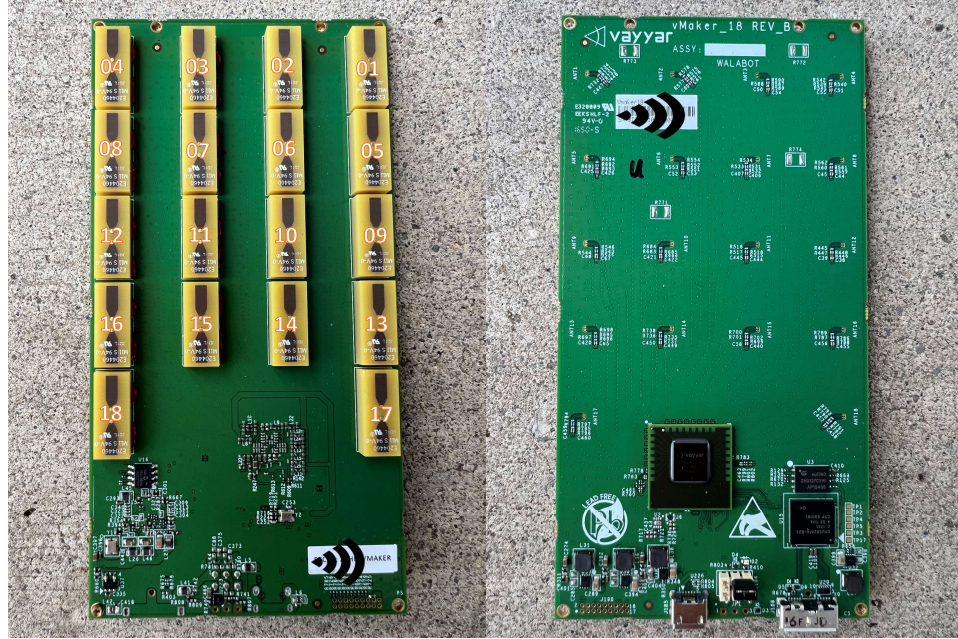
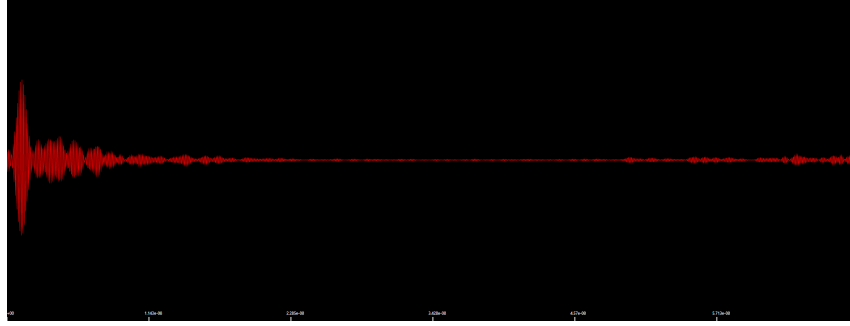


Figure 2.2: Walabot of 18 antenna array distribution front and back.

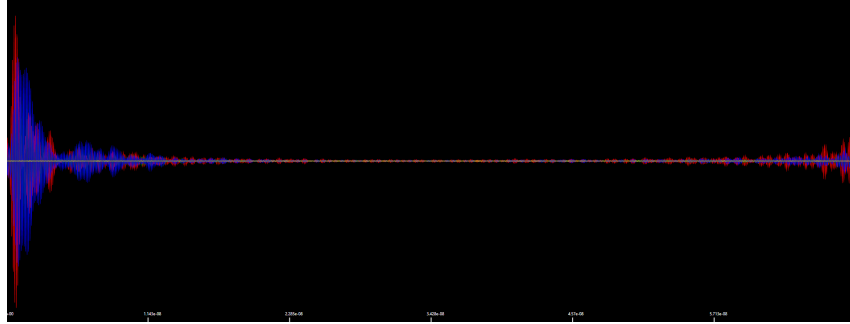
Additionally, Walabot is a programmable three-dimensional sensor that can gaze through objects utilizing previously inconceivable radio frequency technologies. Walabot lights the area in front of it using an antenna array (18 antennas, which are 4 transmitters and 14 receivers, respectively, the distributions are shown in Figure 2.2), detecting and recording reflected signals via the integrated VYR2401 A3 System-on-Chip unit. From the Walabot SDK, it offers 40 antenna pair combinations as shown in Table 2.1. The reflectance strength is measured by the receivers directly as shown in Figure 2.3.

Table 2.1: 40 antenna pairs combinations, Tx and Rx stand for transmitter and receiver, respectively.

Antenna Pairs	1	2	3	4	5	6	7	8	9	10
Tx, Rx	1,2	1,3	1,6	1,7	1,10	1,11	1,14	1,15	1,4	1,8
Antenna Pairs	11	12	13	14	15	16	17	18	19	20
Tx, Rx	1,12	1,16	1,18	17,2	17,3	17,6	17,7	17,10	17,11	17,14
Antenna Pairs	21	22	23	24	25	26	27	28	29	30
Tx, Rx	17,15	17,8	17,12	17,16	4,2	4,3	4,6	4,7	4,10	4,11
Antenna Pairs	31	32	33	34	35	36	37	38	39	40
Tx, Rx	4,14	4,15	18,2	18,3	18,6	18,7	18,10	18,11	18,14	18,15



(a) Microwave signal reflectance using signal antenna pair



(b) Microwave signal reflectance using multiple antenna pairs

Figure 2.3: Microwave radar array sensor reflectance signals

2.2 The Millimeter Wave Radar Array Sensor

The IMAGEVK-74 mm-waves sensor is from Vayyar³ designed and manufactured by Mini-circuits⁴. It has 20 transmitter (Tx) and 20 receiver (Rx) antennas

³<https://www.vayyar.com/>

⁴<https://www.minicircuits.com/>

(shown in Figure 2.4) that can be configured to transmit and receive signals in the 62-69 GHz frequency band. Three performance-optimized transmit profiles allow hardware engineers and researchers to change receiver resolution and imaging processing time with direct access to the Tx/Rx pair phasors for each swept frequency point. The high-resolution profile's 20 Tx and 20 Rx antennas and 400 raw signals make it perfect for high-resolution 3D imaging (20×20 of Rx and Tx pairs). The IMAGEVK-74 mm-waves sensor evaluation kit (EVK) has the whole mm-waves front-end and analog base band signal chain for up to 40 antenna paths (See the functional diagram in Figure 2.5). A high-performance synthesizer produces frequency-stepped continuous wave (CW) waveform and sweeps over transmitting frequencies and antennas. Moreover, the Rx resolution bandwidth, start and stop frequency points, and number of frequency points are totally customizable.

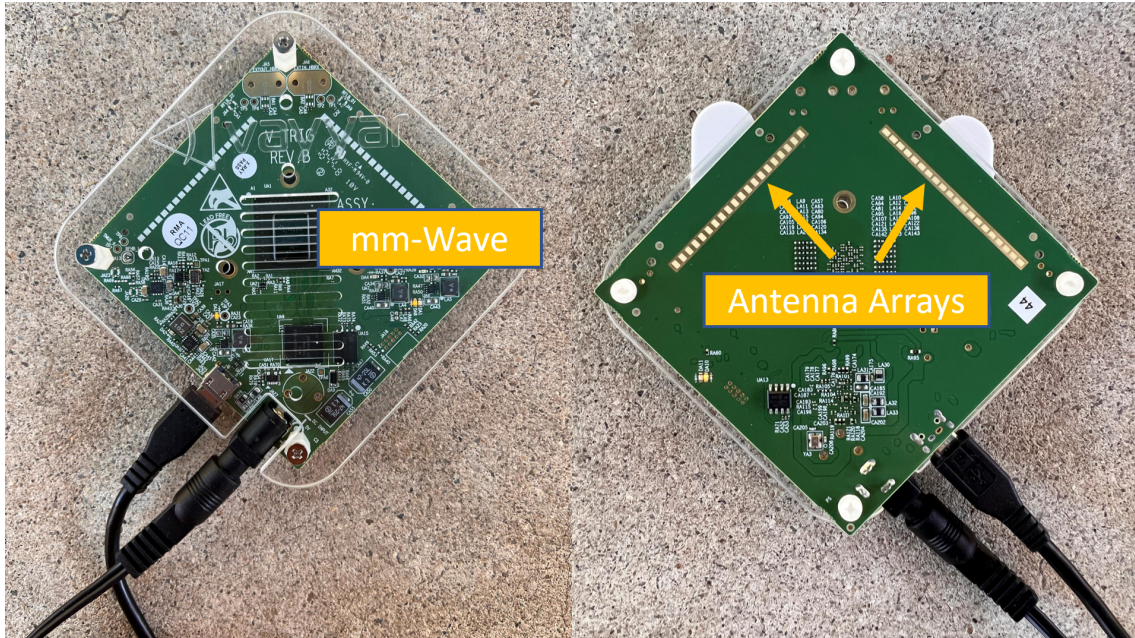


Figure 2.4: Millimeter wave radar sensor front and back.

Furthermore, the maximum power of the transmitter is -10 dBm, which meets the radar health standard [104]. In addition, this mm-waves radar sensor is a new product, and our literature review shows that there is currently only one work using these sensors for material classification [105].

In addition, when installed on a personal computer, the Vayyar Application Programming Interface (API) allows programmatic access to the Vayyar EVK engine, enabling the integration of Vayyar's non-contact RF sensors into user-developed systems. This simple API operates on Windows and is compatible with both Python

and Matlab to support a wide variety of academic and research environments. Overall, the total components are embedded in the sensor unit PCB are 1) Vayyar VYYR7201-A0 system-on-chip (SoC), 2) 40 embedded Tx/Rx linear polarized wide-band antennas, 3) digital interface that communicates with a host PC via a USB-2.0 connection. The interface is implemented using on-board Cypress micro-controller unit (MCU), 4) A Crystal used as a clocking source for the Cypress MCU, 5) Non Volatile Memory (NVM) used to store inventory and calibration information, and the recording profiles are shown in Table 2.2 for users to select due to requirements.

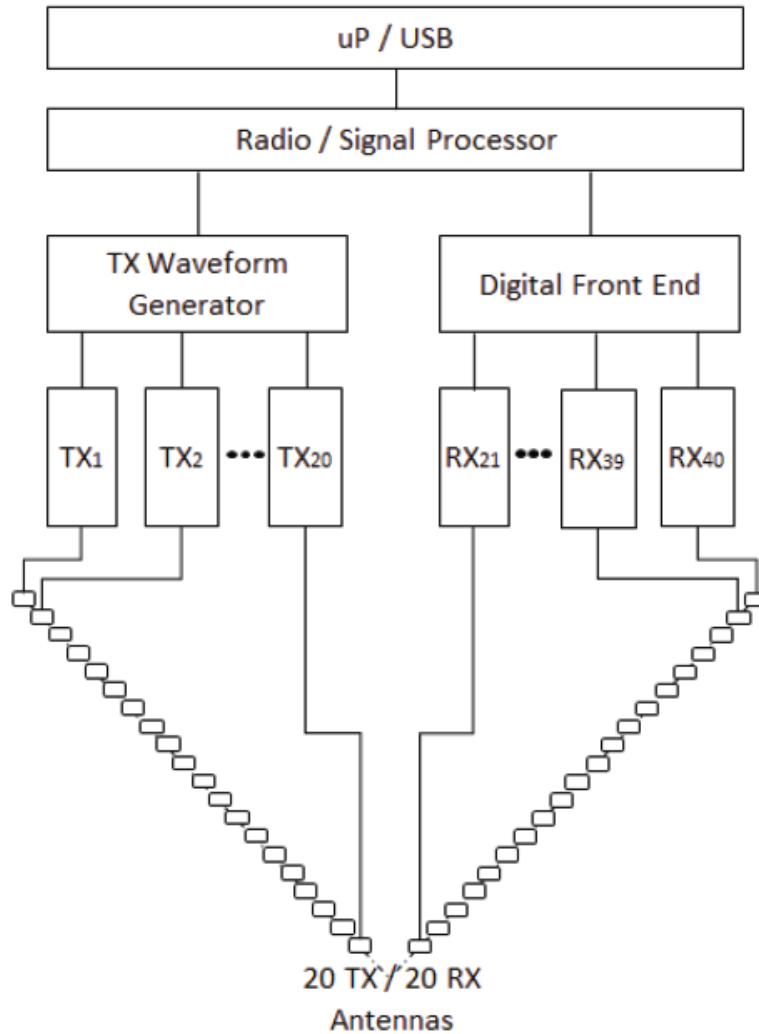


Figure 2.5: Millimeter wave radar sensor front and back.

Table 2.2: IMAGEVK-74 supports 3 predefined user-selectable recording profiles.

Profile Name	Number of Transmitting Antennas	Frame Rate	Transmitting Antenna Numbers	Examples User Application
LOW_RATE	20	Low	1-20	High resolution 3D imaging
MED_RATE	10	Med	1-10	Medium resolution 3D imaging
HIGH_RATE	4	High	1-4	2D imaging, People tracking

2.3 The Unmanned Aircraft Vehicles

Different types of UAVs are utilized for various research purposes, including ET evapotranspiration estimation. Several prominent UAV platforms are depicted in Figure 2.6. There are typically two types of unmanned aerial vehicle platforms: fixed-wing and multi-rotor. Generally, fixed-wing aircraft can fly longer with a greater payload. Typically, it can fly for around two hours, which is sufficient for a wide field. Multi-rotors can fly with a cargo for approximately 30 minutes, making them appropriate for short flight missions. Both have been utilized in agricultural studies, for example [106–108], which have tremendous promise for precision agriculture.



Figure 2.6: The drone image, left is a fixed-wing drone, and right is a Hover drone.

As depicted in Figure 2.6, the author primarily employed a quadcopter named “Hover” to capture aerial pictures. The ‘Hover’ has a Pixhawk flight controller, GPS, and telemetry antennae. It can fly in the mode waypoint mode over the field (designed by using Mission Planner software). The “Hover” lithium polymer battery has a capacity of 9500 mAh, allowing it to support a 30-minute flight mission with cameras attached. To provide a long flight time, the “Hover” is equipped with a highly efficient power system consisting of a T-Engine MN3508 KV380 motor, a 1552 folding propeller, and a Foxtech Multi-Pal 40A OPTP ESC.

2.4 The Unmanned Ground Vehicles

An UGV is an autonomous vehicle that is not only capable of driving without a human driver, but can also integrate artificial intelligence to undertake search, proximity sensing, mapping, and tracking missions autonomously [109]. Current UGVs are available in a variety of configurations and capacities, ranging from full-sized tractors capable of simulating human tasks to small systems for agricultural contexts [110]. Experts predicted a large increase in UGV deployment, particularly in the agriculture sector. Two important breakthroughs are anticipated to contribute to this exponential growth: the increased availability of improved sensor technologies [111] and cognitive applications enabling commercial use of UGVs [112]. Advanced sensor technology enables the automation of cameras, paving the door for UGVs to further cut agricultural labor expenses. Agricultural robots might become a \$45 billion business by 2038, according to a paper on robotics research [113].

Over the past decade, UGVs have demonstrated a great deal of potential for their future position as the principal data collection instrument for precision agriculture [114]. Commercial UGVs have the ability to meet the expanding global need for reduced labor costs. Cognitive swarming via UGVs with related system applications is fast growing in quantity and sophistication, regardless of whether they are utilized as a smart farmer’s companion or as a co-ecologist. Future uses are restricted only by our imagination, similar to the early days of personal computer.

The UGVs proposed in Case Study IV that can operate in the field for more than 5 hours (depends on battery capacity) and are equipped to collect remarkably high-resolution thermal, near-infrared, red-green-blue standard video and imaging, and LiDAR for 3D mapping. These data can then be used to make a number of timely decisions, such as determining not just the plant type but also their water and nutrient stress condition [115]. This type of applications is simply the beginning, as we believe there are other undiscovered agriculture and environmental research capabilities, including crop health checks (disease detection), pest control, irrigation management, crop harvesting, and mechanized weeding, to name a few.

Technological advancements are driving UGV applications in interesting new areas [116]. Numerous UGVs have been designed for a wide range of agricultural applications, including soil sampling, irrigation mapping, and crop harvesting [117]. Multiple UGVs can collaborate to establish a swarming [118] for large-scale fields in a short time where fleets of autonomous UGVs must work together to achieve a common goal.

Recent research on swarming has focused on features of communications and maintaining connectivity [119], such as swarm control and collision avoidance [120], task allocation and techniques for completing numerous tasks [121], and modeling swarm behavior by forecasting individual behaviors [122]. In these publications [123–126], control approaches for UGV swarms are discussed. In this chapter, research will be undertaken using the suggested UGV platform and a swarm mechanism with

numerous UGVs will be provided. In this mode, the followers can merely track the leader UGV based on its GPS position; however, detailed functionalities will be created based on the need for proximity sensing. For instance, multiple UGVs can work in tandem on a large-scale field, drastically reducing the duration of a mission. The acquired data can also be communicated in real-time between UGVs, hence enhancing the intelligence of the UGVs.

2.5 Case Study I: A Non-intrusive Quantification Method for Biochar Water Retention Capacity

2.5.1 Microwave Reflectance Analysis

Due to biochars made from different raw materials or feedstock that have different structures [127], especially their surface has many small porous structures. Therefore, when the microwave signal meets these structures, it will be penetrated, absorbed, and reflected out. In other words, we can get biochar information such as amount usage, different raw materials, water retention capacity, etc., by analyzing the received signal (reflected signals). To better understand what kind of features we are going to face, we draw the reflective signals in Figure 2.7. We collected reflective signals from different kinds of biochar and show them in the time domain, frequency domain, respectively.

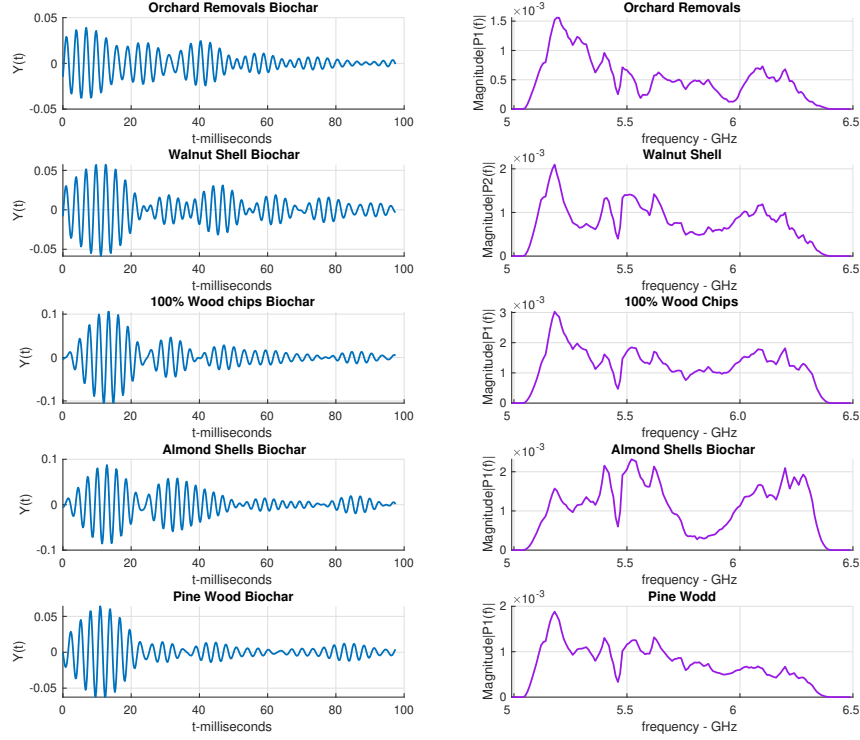


Figure 2.7: Five different kinds of biochars with soil mixture, the left column is the microwave in time domain signal, and the right column is the signal in the frequency domain.

In Figure 2.7, the reflective signals from different biochars show distinguishable variations in both time and frequency domains. Since the microwave attenuated very fast due to its shorter wavelength, here we only show 100-millisecond signal in the time domain (no other signals beyond this time interval), and all responded signals are distributed from 5.0 GHz to 6.5 GHz. As seen in Figure 2.7, the positions of peaks and troughs in the frequency domain vary across various biochars, which can be utilized to form features to classify them.

2.5.2 Microwave Reflective Signals Classification

To leverage the implicit physical knowledge of microwave reflectance relationship to biochar materials, our system takes a data-driven approach by selecting a commonly used supervised classifier, Support Vector Machine (SVM) [128], and

Linear Discriminant Analysis (LDA) [129] to classify and visualize the biochar features. SVM aims to find the maximum-margin hyperplane w and bias term b by minimizing the loss function,

$$\min_{w,b} \frac{\|w\|^2}{2} + C \sum_{i=1}^i \max(1 - y_i(w^T \phi(x_i) + b), 0), \quad (2.1)$$

where C denotes the penalty settings used to balance the regularization term and the training loss for labeled and unlabeled data, respectively. x_i is the input data and y_i is the label. ϕ is the active function. SVM typically discovers boundaries in regions with sparsely labeled and imbalanced data. With SVM, the system can determine which type and how much biochar is used.

2.5.3 Evaluation

2.5.3.1 Experiment Setup

Since we already have Walabot as a microwave transmitter and receiver, we need a processor unit to process data and store it. A portable sensing device that consists of a Raspberry Pi 4B 2GB (RAM) and Walabot developer module as a microwave signal transmitter and receiver are lightly placed on top of the sample, which is a soil and biochar mixture as shown in Figure 2.8. The soil was collected from almond tree fields and filtered for coarse-grained substances such as stone. The biochar is made from 100% of orchard removals, walnut shells, 100% wood, almond shell, and pinewood (labeled as 1, 2, 3, 4, 5, respectively). Following each experiment, the soil was uniformly distributed at 100g, and biochar was added from 20%, 40%, 60%, 80%, and 100% of the soil). Both of them are well mixed.



Figure 2.8: Five different kinds of biochar with soil mixtures.

The sampling frequency of the microwave sensor is 10.24 GHz. We sampled each sample cup five times. Our experiments utilize a classifier, LDA, to distinguish biochar’s unique features. To evaluate our system’s performance fairly, we introduced SVM to calculate average classification accuracy.

2.5.3.2 Different Kinds of Biochar

Based on the results from section 2.5.1, the different biochars have different features in the frequency domain. This fact motivated us to label the data for further research. As mentioned earlier, we have five kinds of biochar (five classes) and put 50g into every soil. Then we sampled 30 times for each class. Therefore, the total dataset contains 180 features, and each feature has 8192 sampling points [130].

Since our data has a high dimension that cannot be easily processed, the LDA algorithm can perform supervised dimension reduction by mapping the input data to a linear subspace consisting of the directions that maximize the separation of classes. The mathematical formulation of the multi-class LDA is shown in (2.2),

$$\sum_v = \frac{1}{M} \sum_{i=1}^M (\mu_i - \mu)(\mu_i - \mu)^T, \quad (2.2)$$

where M denotes classes that have a mean μ_i and the same covariance $\sum$. The specific algorithm of implementation LDA is explained in detail in algorithm 1.

Algorithm 1 LDA classification for multi-class

Require:

Given a dataset contains of matrix X which n (sampling points) by p (total features);

Matrix y (m by 1) that contains total labels.

Ensure:

- 1: Initialize X_{train} , X_{test} , y_{train} , y_{test} (default ratio is 4:1);
 - 2: Initialize predicting confidence scores of soil samples;
 - 3: Initialize parameters for the estimator;
 - 4: get parameters(X, y);
 - 5: fit (X, y) to transform model;
 - 6: **while** Convergence condition not satisfied **do**
 - 7: **for** each feature X_i **do**
 - 8: Transform(X);
 - 9: **end for**
 - 10: **end while**
 - 11: Predict test dataset X_j ;
 - 12: Calculate predict log probabilities $\log P$;
 - 13: plot classified classes in 2D or 3D.
-

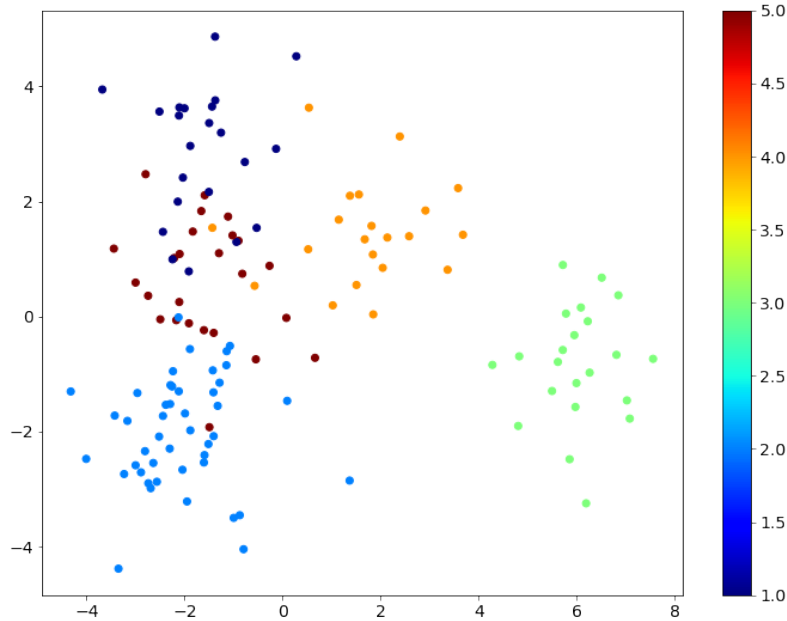


Figure 2.9: LDA results for different types of the biochar measurements.

The LDA results for different types of the biochar were shown in Figure 2.9. Label numbers represent different biochar, as mentioned earlier. There are five different kinds of biochars with different colors, which means different biochars can be classified by their features. However, from this 2D plot, we noticed that several sampling points had been overlapped. Thus we plotted out LDA in 3D (Figure 2.10) to see the distribution. As we can see Figure 2.10, biochars are classified and no overlapped with other sample points.

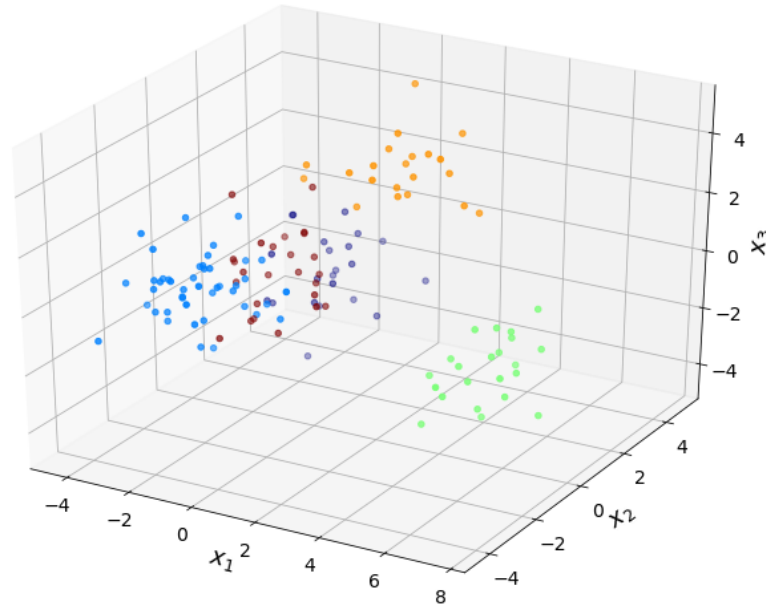


Figure 2.10: LDA 3D results for different biochar types measurements.

2.5.3.3 Different Biochar Amount Classification

In this section, we verify how different amounts of biochar corresponds to the microwave reflective signals in the frequency domain. We separated six sets which put 0g, 50g, 100g, 150g, 200g, and 250g biochar, respectively, into the same amount of soil (250g). The reflective signal has been recorded after mixed well. After the feature was collected, we divided the collected data into train and test set for SVM multi-model development and evaluation. The train and test sets ratio is 2:1. Figure 2.11 depicts the confusion matrix, and we found that the SVM classifier achieves 74.0% of biochar usage amount recognition accuracy. This result proves that it is feasible to identify biochar intake and other quantities such as moisture. This assessment establishes the practicality of deploying mobile microwave sensors to monitor soil conditions and demonstrates the potential for cost savings as the target area expands.

Accuracy: 74.0%

0g	100% 4	0.0% 0	0.0% 0	0.0% 0	0.0% 0	0.0% 0
50g	0.0% 0	100% 4	0.0% 0	0.0% 0	0.0% 0	0.0% 0
100g	0.0% 0	0.0% 0	100% 4	0.0% 0	0.0% 0	0.0% 0
150g	0.0% 0	0.0% 0	66.7% 2	33.3% 1	0.0% 0	0.0% 0
200g	0.0% 0	0.0% 0	20% 1	40% 2	40% 2	0.0% 0
250g	0.0% 0	0.0% 0	0.0% 0	0.0% 0	33.3% 1	66.7% 2
	0g	50g	100g	150g	200g	250g

Figure 2.11: Confusion matrix of different amounts of biochar recognition.

2.5.3.4 Features Extraction with Different Water Usage

Based on the idea in [131, 132], we tested our microwave signal to see whether different water capacities could be detected within the biochar mixture, as shown in Figure 2.12, and the LDA classifier results are shown in Figure 2.13. The magnitudes of reflective microwaves vary from different moisture levels and can be distinguished by using the LDA algorithm. Since the microwave attenuated fast during different mediums, the magnitudes can be treated as the energy of the signal that explicates the water amount (the medium amount for absorbing microwave is higher, the attenuation is higher [133]). Unfortunately, the result shows in not very clear and some prediction results overlapped in Figure 2.13. We think a Deep Neural Network could help to improve in the future work.

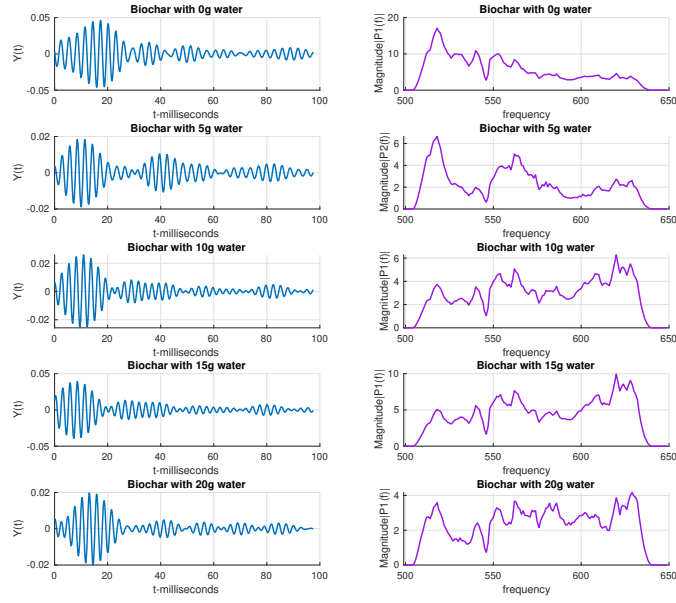


Figure 2.12: Five different water content levels with biochar mixture.

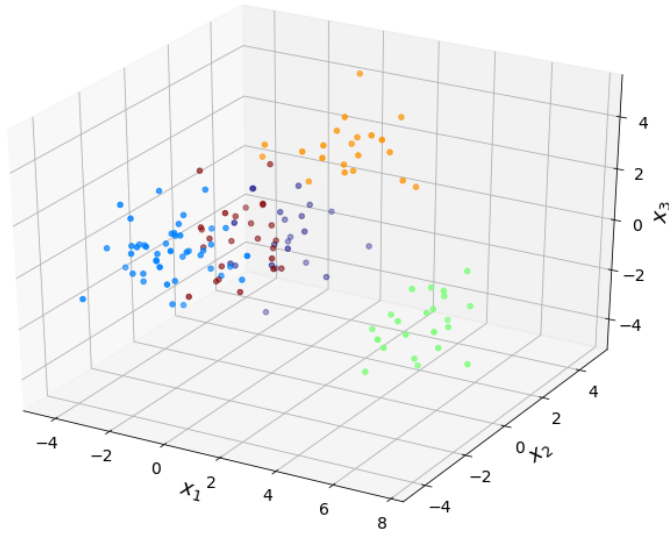


Figure 2.13: LDA 3D results for different water content levels classifications.

2.5.3.5 Biochar Water Retention Capacity

As mentioned earlier, combined with physical principles behind the microwave sensing and biochar properties, we validated the biochar with the soil of water retention capacity by comparing the same amount of soil (250g), and one of them added biochar (100g, mixed well). Our experiment method is, both of them adding water until it is well saturated, recorded current microwave reflective signals, then put both samples under the sun exposure for 5 hours. After that, we measure reflective signals again for the comparison as shown in Figure 2.14. We use a grid cover to cover the top of the sample during sun exposure to make sure no other substance polluted our samples.

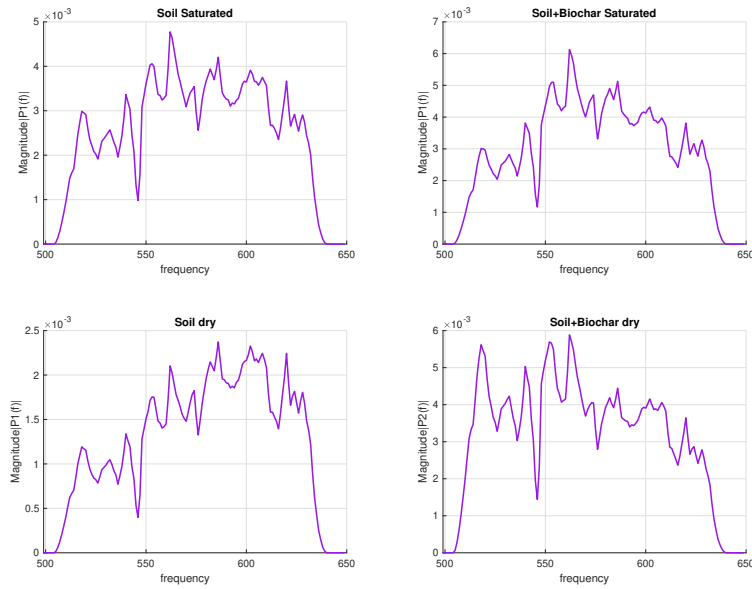


Figure 2.14: Microwave reflective signals in the frequency domain before and after 5 hours, respectively.

We utilized the average amplitude variation to inference the water level change in two samples. The results are shown in Table 2.3. As we can observe the variation, after 5 hours of sun exposure, the water reduced 49.34% whereas the biochar mixture sample only reduced 7.64% water content level. These results indicate that biochar has more substantial water retention capacity than plain soil, which benefits the soil for keeping moisture.

Table 2.3: The comparison of water retention capacity for soil and soil with biochar mixture.

	Soil	Soil+Biochar
Before	1.4691×10^{-4}	1.8484×10^{-4}
After 5 Hours	7.4419×10^{-5}	1.7071×10^{-4}
Percentage	50.66%	92.36%

2.6 Case Study II: Soil Carbon Content Quantification Method Using A Miniature Millimeter Wave Radar Sensor

2.6.1 Millimeter Wave Signal Classification Method

In this section, we discuss how to construct a classification method that, based on the time-domain characteristics of the received signal, can discriminate precisely between various signals. We have trained various classifiers, such as the logistic regression classifier, the one-versus-all SVM classifier, the K-Nearest Neighbors (KNN) classifier, and the Decision Tree classifier. Upon evaluation, we found that although each classifier offered unique advantages, their performance did not meet our criteria for precision and efficiency. we opted to utilize a finely tuned Decision Tree classifier for our classification needs.

The mathematical foundation of decision tree is that, given training vectors $x_i \in \mathbb{R}^n$, $i = 1, \dots, l$ and a label vector $y \in \mathbb{R}^l$, a decision tree recursively partitions the feature space such that the samples with the same labels or similar target values are clustered together.

Let the data at node m be represented by J_m with n_m samples. For each candidate split $\alpha = (j, t_m)$ consisting of the feature j and threshold t_m , partition the data into $J_m^{left}(\alpha)$ and $J_m^{right}(\alpha)$ subsets [134]:

$$\begin{aligned} J_m^{left}(\alpha) &= \{(x, y) | x_j \leq t_m\}, \\ J_m^{right}(\alpha) &= J_m \setminus J_m^{left}(\alpha). \end{aligned} \quad (2.3)$$

The quality of the candidate split of node m is then computed using an impurity function or loss function H , the choice of which depends on the task being solved (classification or regression)

$$G(J_m, \alpha) = \frac{n_m^{left}}{n_m} H(J_m^{left}(\alpha)) + \frac{n_m^{right}}{n_m} H(J_m^{right}(\alpha)). \quad (2.4)$$

Select the hyper parameters that minimize the cost

$$\alpha^* = \operatorname{argmin}_{\alpha} G(J_m, \alpha), \quad (2.5)$$

and act recursively for subsets $J_m^{left}(\alpha^*)$ and $J_m^{right}(\alpha^*)$ until the maximum allowable depth is reached, $n_m < \min_{samples}$ or $n_m = 1$.

The adoption of a decision tree classifier is motivated by its capacity to train a large number of largely uncorrelated models (trees) working as an ensemble of predictors more accurate than individual predictors, and it is a flexible, user-friendly supervised learning method that, even without hyper-parameter modification, typically yields excellent results. In addition, our final objective is to develop the machine learning model on Edge-AI devices to do on-site soil carbon content detection.

2.6.2 Experiments

2.6.2.1 Experiments Setup

The sensing module consists of three major parts: the mm-wave transmitter and receiver module (IMAGEVK-74), the Raspberry Pi 4B, and the visualization interface shown in Figure 2.4. The sample is placed underneath the mm-waves radar sensor with vertical gap less than 2 cm.

The soil was collected from almond tree fields and filtered to remove coarse-grained substances such as stone. The biochar (we mentioned earlier shown in Figure 2.16) made from 100% of orchard removals, walnut shells, 100% yellow wood, almond shells, pinewood. In the following experiments, the mixture ratio percentage of soil and biochar was varied from 0% to 100%.

The sampling frequency of the mm-wave sensor is from 62 GHz to 69 GHz. We sampled each sample 20 times for the purpose of reducing likelihood error. We set up 150 sampling points during the frequency band we mentioned, and the resolution bandwidth (RBW) is 100 KHz for reasonable sweep time and lower noise. Our experiments utilize principal component analysis (PCA) to reduce the dataset dimensions. To evaluate our system performance fairly, we introduced a confusion matrix to calculate average classification accuracy. The 5-fold cross-validation is also employed to adjust hyper-parameters and validate all models' performance on our soil carbon samples. The accuracy and confusion matrix are presented to evaluate the classified objects' performance.

2.6.3 Evaluation and Results

2.6.3.1 Different Soil Carbon Amount Recognition

The radar reflective signals can correlate the soil carbon content density, which has been proven in [135]. The wavelength of mm-waves is smaller than centimeter waves, which means it can detect more fine-grained information than microwave. As mentioned earlier, we select the carbon and soil ration from 0% to 100% including the air, that make up a total of 7 classes. Then we sampled 20 times for each class. Therefore, the total dataset contains 140 features, and each feature has 1024 sampling points (time domain features).

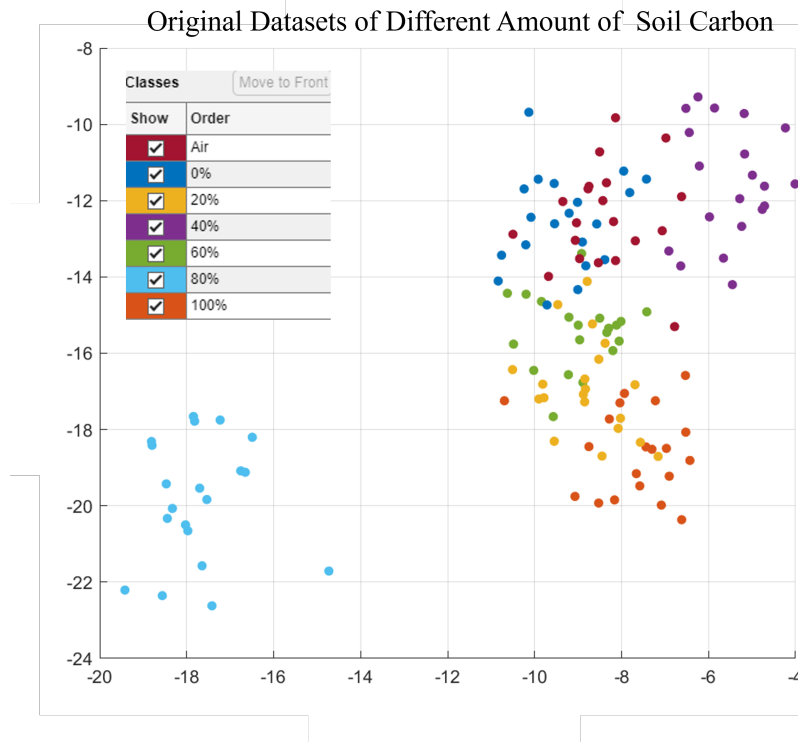


Figure 2.15: Original datasets scatter plot with all classes. Different amount of soil carbon.

The scatter plot of the raw data is shown in Figure 2.15. We observe with the mm-waves is less sensitive to the different amounts of biochar mixed with soil. The accuracy from the confusion matrix is 90.8%, as shown in Figure 2.17 a). Class 0%, air and 60% are sort of overlapped, which causes the relatively higher failure. However, compared with microwave sensing, the classification accuracy is pretty high [135].

2.6.3.2 Different Type of Soil Carbon Classification

In our study, we examined five distinct types of soil carbon (biochar), as depicted in Figure 2.16. The rationale behind this selection stems from the unique microporous structures inherent to each biochar variant, which are believed to influence water retention capabilities and nutrient storage among other properties. The scatter plot in Figure 2.18 illustrates the variability among the biochar types. When compared to Figure 2.15, it becomes evident that the data points for each biochar class cluster more tightly within their respective categories than when comparing different amounts of soil carbon. This clustering suggests that the distinct characteristics of each biochar type significantly enhance the classifier's ability to differentiate between them. Further validation of this observation is provided by

the confusion matrix in Figure 2.17 b), which reports an impressive classification accuracy of 95.8%. This high level of accuracy underscores the effectiveness of distinguishing between biochar types based on their unique properties, thereby facilitating more precise classification by the classifier.



Figure 2.16: Five different types of biochar.

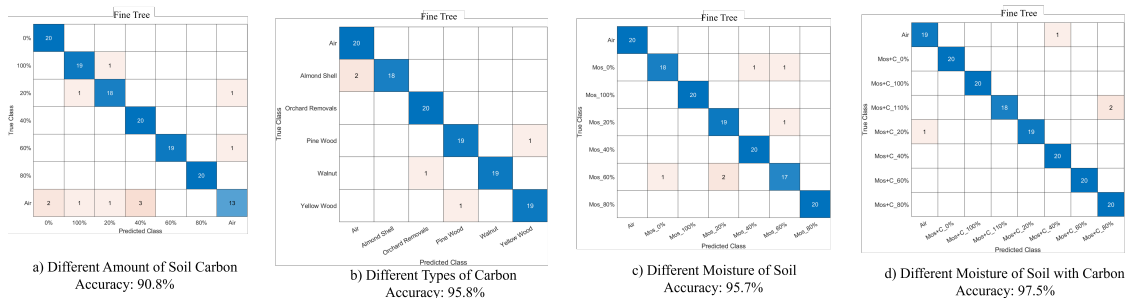


Figure 2.17: Confusion matrix for all experiments.

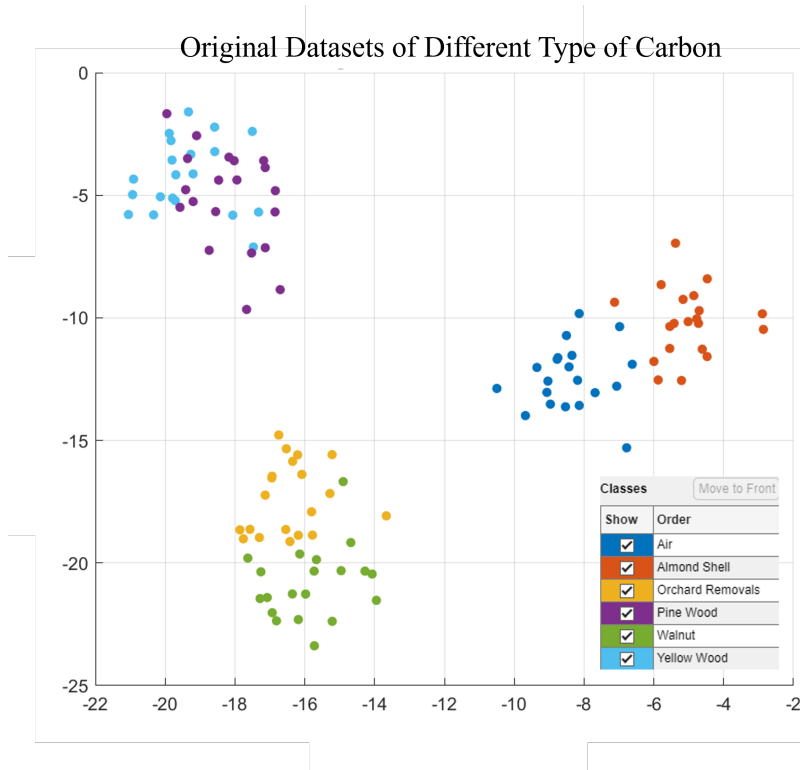


Figure 2.18: Original datasets scatter plot with all classes. Different types of carbon.

2.6.3.3 Soil Moisture Study Using mm-waves

Precision agriculture requires soil moisture monitoring, which has a considerable impact on crop evapotranspiration, water exchange, and energy fluxes. Numerous remote sensing and proximate sensing techniques, such as thermal, optical, and microwave measurements, can be used to measure soil moisture. Low spatial resolution, limited surface penetration, and vegetation are nonetheless limiting constraints for the use of these techniques. By using a contactless method to sense soil moisture, the work in [132] used a microwave sensor to sense the soil moisture. However, there is no work to use an mm-waves sensor to sense the soil moisture based on our best knowledge. We put different amounts of water on the soil until it is saturated (Mos_100%).

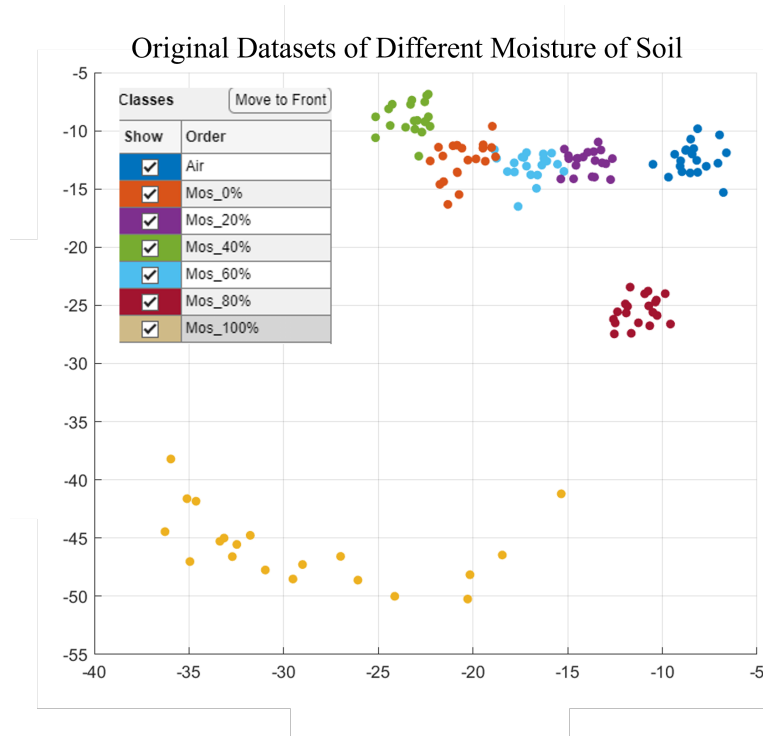


Figure 2.19: Original datasets scatter plot with all classes. Different amounts of soil moisture.

Figure 2.19 shows the sensing results in a scatter plot. As we all know, water can both absorb mm-waves and microwaves. On the other hand, soil and water together can coagulate into large pieces, which is good for sensing their unique features. The scatter plot also shows the sensed features are close to each other for the same class. The accuracy of soil moisture is 95.7% as shown in Figure 2.17 c).

2.6.3.4 The Soil Moisture Detection Under Carbon Efforts

As we mentioned earlier, combined with physical principles behind the mm-waves sensing and biochar properties, we validated that the biochar can improve the soil water retention capacity. The experiment shows that the same amount of soil with 50 g of biochar can absorb up to 110% water (well saturated). Due to biochar porous structures that can absorb soil and water, the reflectance signals are pretty consistent and close to each other in the same scatter as shown in Figure 2.20.

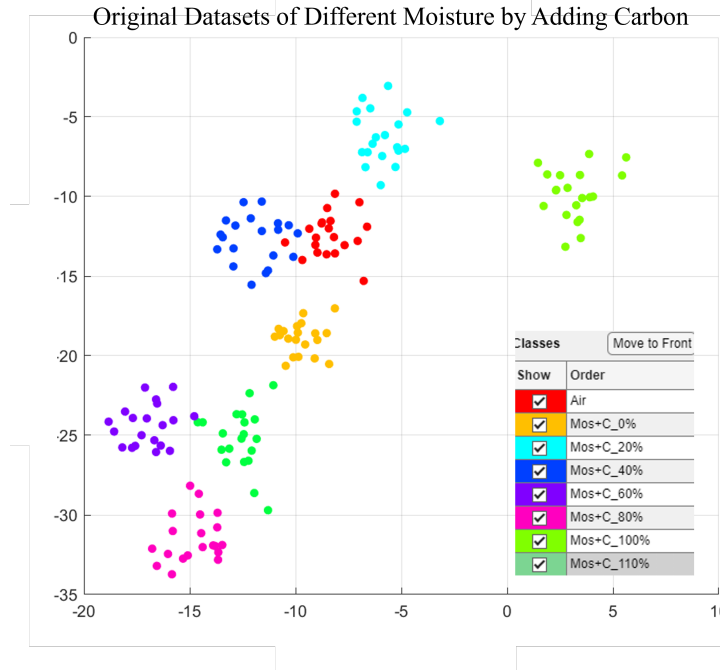


Figure 2.20: Original dataset scatter plot with all classes. Different moisture levels with the same soil carbon.

Furthermore, Figure 2.17 d) shows that the classification accuracy for classifying the different soil moisture with a fix amount of carbon is 97.5%. Overall, using mm-waves to do soil carbon, moisture, and other related missions is pretty reliable.

2.7 Case Study III: A Field Study of Soil Biochar Treatment Response Using Small Unmanned Aerial Systems (sUAS)

2.7.1 Introduction

Biochar is a solid material generated during the pyrolysis of biomass. Biochar can be applied to soil to enhance soil functions and offset greenhouse gas emissions caused by biomass that would otherwise decay naturally into GHG. The benefits of biochar to soil fertility include the extraordinarily high binding affinity of nutrients for biochar and other molecules (adsorption), and the biochar's remarkable retention in the ground (stability) [135]. Additionally, sustained biochar application can generate oil and gas byproducts that can be collected and used as fuel, resulting in clean, renewable energy and soil improvement. This is one of the negative emissions technologies (NETs) [136] that ensures that the final product becomes the real meaning of "carbon negative".

UAVs have been widely employed in agricultural applications as a novel remote sensing platform, including crop yield estimation [137], irrigation management [138,

139], and pest control [140, 141]. In comparison to satellites, UAVs may fly more freely and often in the field. The UAVs fly at a lower height and capture pictures of crops at a greater spatial resolution [142]. Additionally, UAVs reduce the cost of data collecting compared with using other remote sensing methods such as satellite sensing.

In addition, one of the objectives of our field study is to reveal the relationship of affection between soil temperature and soil moisture with soil amendment (biochar). This investigation requires ground truth data to map the temperature and moisture in the temporal and spatial domain, respectively. The UAV with the benefits we mentioned earlier is a great platform to implement our mission in the growing season.

Previous research on soil temperature and moisture has concentrated chiefly on obtaining a soil sample and then analyzing it using an expensive analyzer [143] to isolate the various components of soil [144]. Several researchers embedded portable devices on soil fields and inserted sensors in the soil to detect [145]. Additionally, a few studies employed remote sensing technology such as satellites and UAS [146, 147] to assess soil conditions and carbon emissions (such as GHG). These prior attempts are mainly relied on the expensive analyzer and limited to the environmental factor (which means cannot operate their analyzer or collect valid samples due to rainy day, human activities, etc [148]). Some works use UAV and NDVI to evaluate the growth, especially soil moisture, evapotranspiration, and their minor element contents [149–154]. However, these attempts didn’t explain which level of soil moisture can be accurately predicted by NDVI, and also whether these image indexes can correlate with soil temperature for further explaining carbon activities (such as GHG emissions). Therefore, we designed our field study with three benefits (innovations):

- The UAV equipped with thermal and near-infrared (NIR) cameras helps us monitor the soil amendment effects on large-scale soil in real-time.
- Compared with the different depths of VWC with NDVI to find the better correlation.
- Introduced the VARI to build a relationship with soil temperature even if there are disturbances (green plants) in between.

As a result, we target two challenges in this paper: 1) what’s the best sensing mechanism for using NDVI and VARI to reflect different depths of soil water content 2) what’s the optimized fitting model of these indexes with soil temperature and moisture. The contributions of this work are as follows:

1. To the best of our knowledge, this is the first study validating the optimized depth of VWC correlated by the NDVI.
2. After adding biochar amendment, we revealed a solid correlation between vegetation index and soil temperature.

3. We validated a field study in the winter growing season by conducting a series of intense sUAS flight and ground truth data collection.

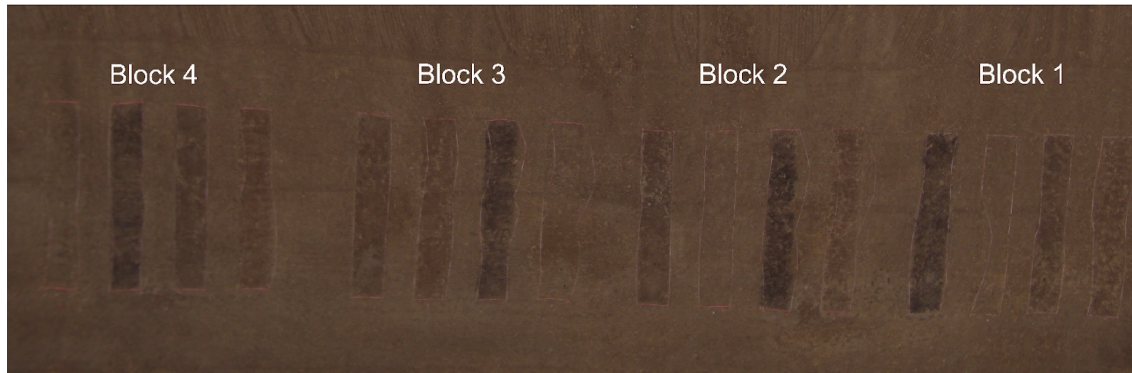
2.7.2 Materials and Methods

2.7.2.1 Study Field Site and UASs Remote Sensing Configuration

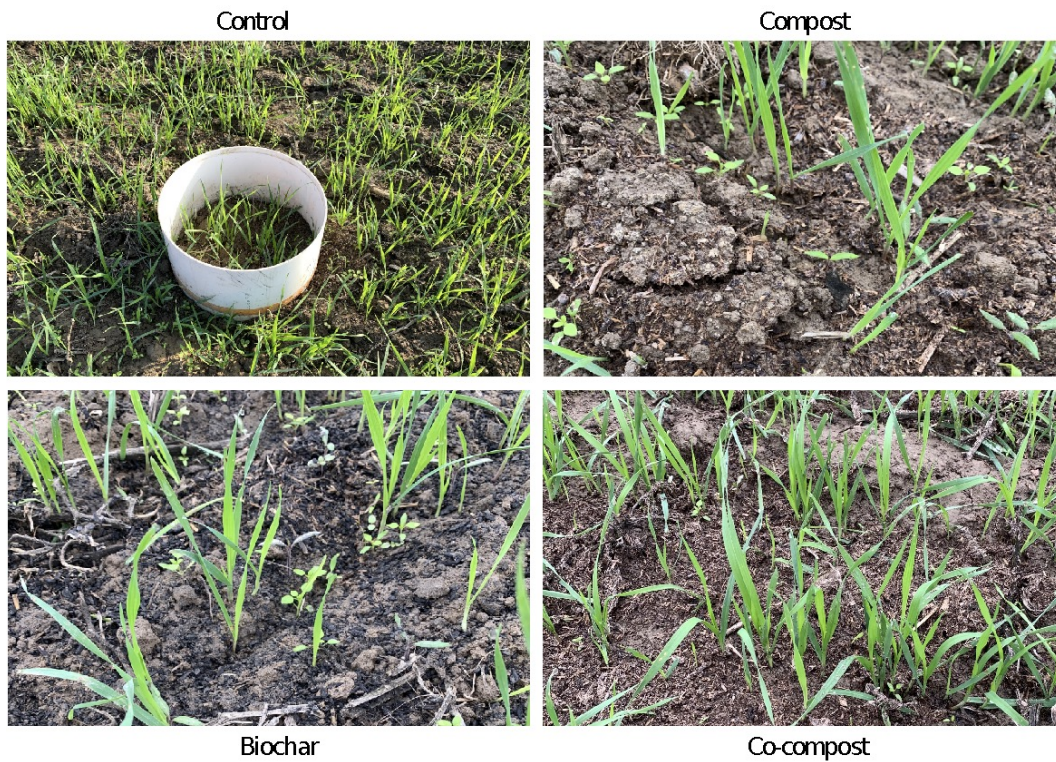
We established a field experiment at a wheat (winter) rotational cropping system near the Philip Verwey Farm in Madera, CA (36.94° N, 120.378° W) in October 2021. The study site exhibits a Mediterranean climate with a mean annual temperature of 18.2°C and mean annual precipitation of 311mm. Summer is hot and dry (e.g., July average temperature is 29.3°C with an average precipitation of 0 mm) while winter is relatively cool and damp (e.g., January average temperature is 8.5°C with a 56 mm monthly average precipitation).

Biochar used in the field trial was the Rogue biochar produced by Oregon Biochar Solution (White City, OR, USA). The feedstock of biochar consisted of approximately 85% Douglas fir and ponderosa pine wood waste mixture, 14-15% almond and walnut tree pruning, and less than 1% of nutshells. Biochar particles had diameters ranging from 3 mm to 1 cm. Manure compost and biochar manure co-compost were prepared on-site at the Verwey farm during August - September 2021. Briefly, the manure-only pile consisted of approximately 15.34t of fresh solid manure and 1.32t of orchard clipping residues. The biochar co-compost pile consisted of 15.35t of fresh solid manure, 1.32t orchard clipping residues, and 1.0t of biochar.

Replicated treatment plots (2m x 12m with 2m buffer in between) were laid out in a randomized block design ($n = 4$) at our study site (Figure 2.21a). Treatments were applied to plots (a total of 16 plots) one week before seeding. All plots were seeded to winter wheat for the winter growing season on October 29th, 2021. The four treatments employed in this field trial were: 1) Control with no additional amendment; 2) Manure compost applied at 20 t/ha; 3) Biochar applied at 10 t/ha, and 4) Biochar manure co-compost applied at 17.5t/ha. All treatments were applied at equivalent C rates of 8 MgC/ha, see Figure 2.21b. Treatments were applied to the surface soil and incorporated to approximate 15cm depth with a rake and tines of a pitchfork. We also gently raked control plots to ensure consistency. A single time of flood irrigation was adopted following the landowners' common practices on November 11th, 2021. Our study site received no fertilizer input during the winter growing season, and the nutrient source for winter wheat was solely from the decomposition of crop residues left from the previous growing season.



(a) Biochar plots locates at Verwey Farm



(b) Four treatments of soil and their plants growth

Figure 2.21: (a) Replicated treatment plots (2 m x 12 m with a 2 m buffer) were laid out in a randomized block design in four blocks. biochar (10 t/ha), compost (20 t/ha), and co-compost (17.5 t/ha) were surface applied, and mixed. (b) Control, biochar, compost, and co-compost plots in one replicated block.

We employed a Hover quadcopter⁵ as our UAV platform. A Pixhawk flying controller, GPS, and telemetry antenna were installed on the Hover. Additionally, it is capable of flying over a biochar treatment field utilizing the waypoints setting (designed by using Mission Planner software⁶). The Hover’s lithium battery has a capacity of 9500 mAh, which is adequate for a 30-minute flight mission with the Hover’s cameras mounted (this battery also powers the camera).

In this field study, multi-spectral images were collected by Survey 2 (MAPIR, USA)⁷ cameras with 4 bands, blue, green, red (RGB), and near-infrared (NIR). The MAPIR camera has a resolution of 4608×3456 pixels and a spatial resolution of 1.01 cm/pixel. Besides, thermal images of the biochar treatment field were collected using the ICI 9640 P-Series thermal camera (ICI, USA). The thermal camera has a 640×480 pixel resolution. The precision was expected to be 1°C . The thermal camera will verify whether the soil temperature can be different through different treatments and plants. Furthermore, the flight height of the three cameras was settled at 30m, 30m, and 40m, respectively. The overlap of Field of View (FOV) was adjusted to 75% for all aircraft missions to ensure that the photos could be stitched together during image pre-processing. From late October 2021 to early February 2022, we did the flight mission biweekly on our biochar treatment field. All cameras’ reflectance was calibrated before the flight by using a color panel.

2.7.2.2 Geospatial Interpolation with Kriging

In geostatistics, one of the common problems encountered is to spatially interpret and estimate a set of unknown data locations, $z(s_0)$, from a set of sparse observations, $z(s_i)$, at locations s_0 and s_i , respectively. These N sparse sampling points are taken inside a domain $\Omega \in \mathbb{R}^2$ to find a set of weights, λ_i , to estimate the value at the unknown locations,

$$\hat{z}(s_0) = \sum_{i=1}^N \lambda_i z(s_i). \quad (2.6)$$

For ordinary Kriging [155], a semivariogram is used to represent the spatial variability. For a given spatial distance between observations, h , the point support semivariogram is written as,

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z(s_i) - z(s_i + h))^2, \quad (2.7)$$

⁵<https://www.foxtechfpv.com/foxtech-hover-1-quadcopter.html>

⁶<https://ardupilot.org/planner/>

⁷<https://www.mapir.camera/collections/survey2>

such that $N(h)$ is the number of observations with distance h . Using this experimental semivariogram data (2.7) a semivariogram model is typically fit using one of the several common functions (circular, spherical, exponential, Gaussian, or linear). Finally, given the semivariogram model, the kriging weights can be determined by solving the linear system,

$$\sum_{j=1}^N \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) + \mu(\mathbf{s}_0) = C(\mathbf{s}_i - \mathbf{s}_0), \text{ for } i = 1, 2, \dots, N, \quad (2.8)$$

where $C(\cdot)$ represents the point support covariance matrix. This covariance matrix has an important relationship to the semivariogram, such that, $\gamma(h) = C(0) - C(h)$ [155]. The estimation error variance is shown to be $\sigma_e^2 = \text{Var}(z(s_0) - \hat{z}(s_0))$, which for ordinary kriging, is minimized to make the estimated values $\hat{z}(s_0)$ optimal. In order for the estimator to be unbiased (e.g. $E[\hat{z}(s_0)] = E[z(s_0)]$), kriging requires $\sum_{i=1}^N \lambda_i = 1$ and the spatial mean to be stationary $E[z(s)] = \mu, \forall s \in \Omega$.

In this work we utilize kriging with the ground truth temperature measurements to compare with TIR imagery and ice water calibration. The workflow is illustrated in Fig. 2.22.

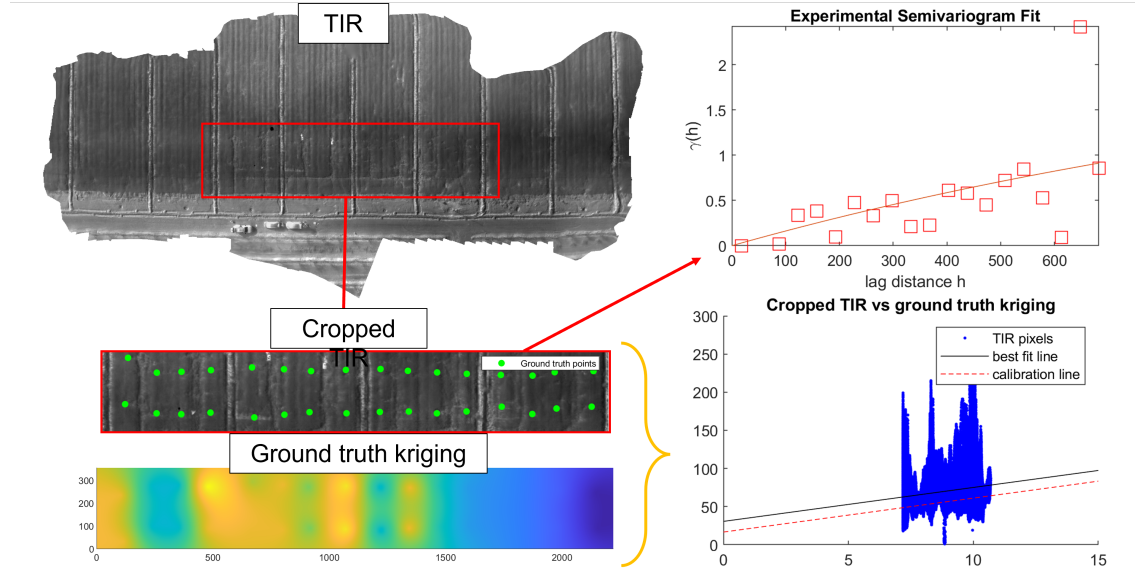


Figure 2.22: The TIR imagery is cropped in this diagram to enclose only the biochar plots. The geolocations of the ground truth measurements are identified and used to fit an experimental exponential semivariogram with a lag of 700 pixels. The resulting semivariogram is used to krige the ground truth measurements to match the TIR imagery size. Lastly, the pixel-by-pixel correlation plot between the cropped TIR image and ground truth kriging is constructed with the best-fit line. This best-fit line's y -intercept is then corrected using the pixel intensity and a known temperature ice bath.

2.7.2.3 Statistical Analysis

We decided to use the VARI and NDVI to verify whether these two indexes can indicate the soil temperature and its moisture at different treatments based on multiple-spectral images collected. The default equations to generate these indexes are given by (2.9) and (2.10), respectively.

$$\text{VARI} = \frac{\text{Green} - \text{Red}}{\text{Green} + \text{Red} - \text{Blue}}, \quad (2.9)$$

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}}. \quad (2.10)$$

The VARI method is a vegetation index that enables a quantitative estimate of vegetation proportions using only visible light. In other words, this index highlights the green vegetation on the ground to make the soil, rocks, and plants clearly distinguishable. The NDVI method is a standardized index that permits the construction of a visual representation of the degree of greenness in a given area (relative biomass). This index uses the contrast between chlorophyll pigment absorptions in the red band and the high reflectivity of plant components in the near-infrared region of a multispectral dataset. This index returns values in the range of [-1.0, 1.0] [156].

To illustrate the variety of different treatments, we plot all the data and use a linear regression model to seek the relationship between the index, soil temperature, and VWC at different depths of soil. The R^2 and standard deviation (STD) can be used to evaluate the strength of the relationship between soil temperature, VWC and image indexes.

2.7.3 Experiments and Results

We scheduled our flight missions bi-weekly during the winter growing season. Every time we captured RGB, NIR, and thermal information from the ground, the ground truth data such as soil temperature and VWC would also be collected by thermometer and soil suction lysimeter ⁸. The scatter would be first to plot to observe the potential trends relationship between these information. We extracted two points (Kriging, see details in Section 2.7.2.2) of VARI and NDVI values from our total plots individually. Please see details in the following sections.

2.7.3.1 Correlation Analysis of The VARI and NDVI Indexes Versus Temperature

In general, the soil temperature in November is higher than in January. The temperature in the central valley of California was still suffering from the dry season, and the farmers had just seeded the wheat in the early part of November, which

⁸<https://www.hannainst.com/suction-lysimeter-for-root-level-soil-monitoring.html>

meant the soil was exposed to sunlight, and there were no green plants on the ground. From Figure 2.28a, we can see that the average R^2 is the maximum ($R^2 = 0.5$), and its dispersion is the minimum ($STD = 0.19$), which indicates the VARI index has the potential ability to predict the soil temperature within biochar amendment.

The NDVI has a weaker indicating ability than VARI for the soil temperature. However, the plots of adding biochar amendment still have the most significant difference in R^2 and STD, which also verified that NDVI has the remarkable ability to predict the temperature better than the other treatment blocks.

Overall, the VARI and NDVI can both reveal the soil temperature. Biochar as an amendment would help quantify the soil temperature and VWC based on UAV sensing. The average enhanced R^2 of VARI from biochar treatment is 46.77%, compared with control, manure, and mixture treatments. Relative to the mixture treatment, the NDVI from biochar treatment enhanced the R^2 was 52.94%.

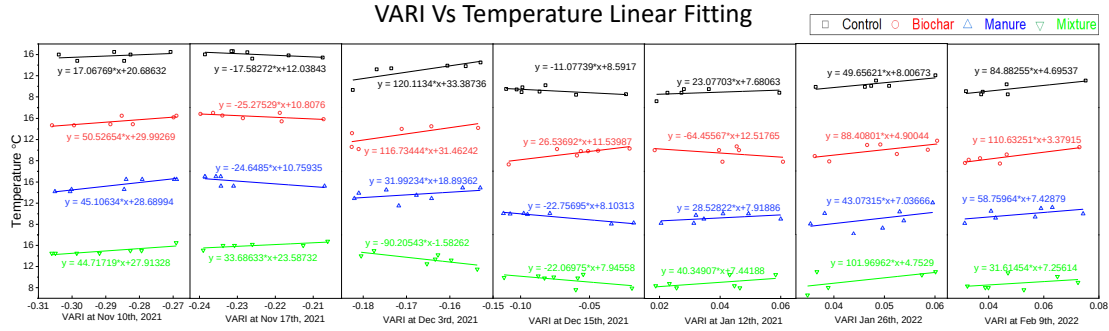


Figure 2.23: VARI vs. temperature. The x -axis indicates the VARI and NDVI values and the y -axis indicates the ground truth of soil temperature, The dark, red, blue, and green lines are represent control, biochar, manure, and mixture treatments, respectively.

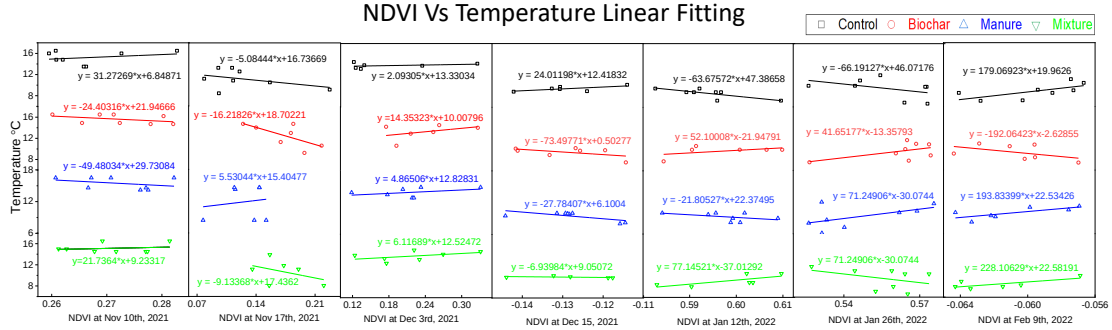


Figure 2.24: NDVI vs. temperature. The x -axis indicates the VARI and NDVI values and the y -axis indicates the ground truth of soil temperature, The dark, red, blue, and green lines represent by control, biochar, manure, and mixture treatments, respectively.

2.7.3.2 Correlation Analysis of The NDVI Index Versus Different Depths of VWC

As we know, biochar has strong water retention ability [135]. However, quantifying the biochar water retention ability on a large scale is still a challenge. Fortunately, the soil moisture can be indicated by the health and growth of the biomass (NDVI). Therefore, we compared different depths of VWC with NDVI values, see Figures 2.25, 2.26, and 2.27.

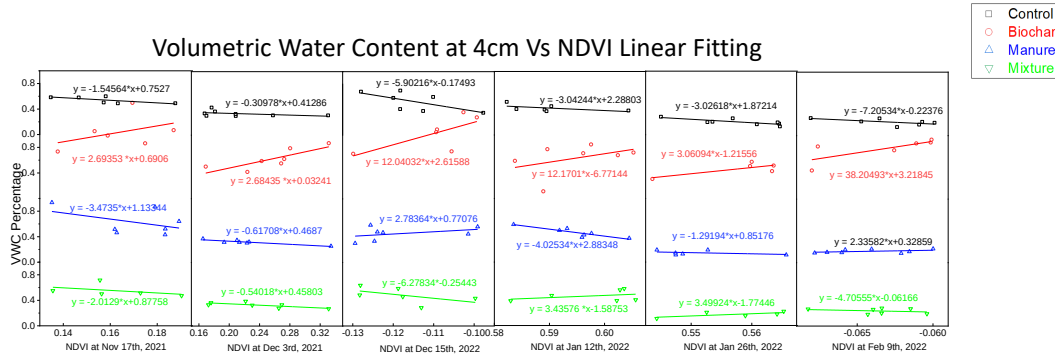


Figure 2.25: VWC at 4cm. The x -axis the NDVI value, and the y -axis is the VWC (range from 0 to 2. VWC > 1 means water is greater than soil in the same unit, the same unit of y -axis for the rest of the VWC figures.

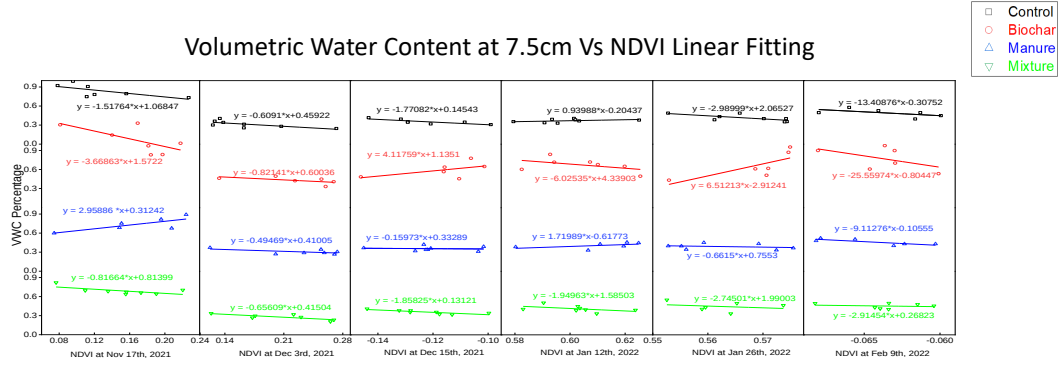


Figure 2.26: VWC at 7.5cm. The x -axis is the NDVI value, and the y -axis is the VWC.

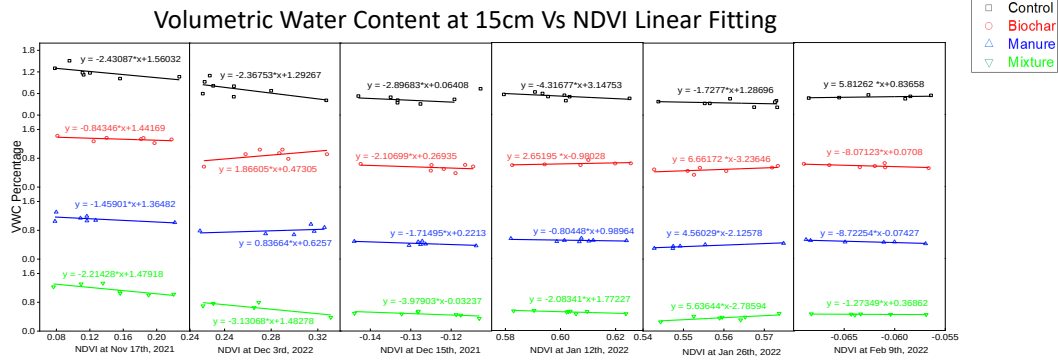


Figure 2.27: VWC at 15cm. The x -axis is the NDVI value, and the y -axis is the VWC.

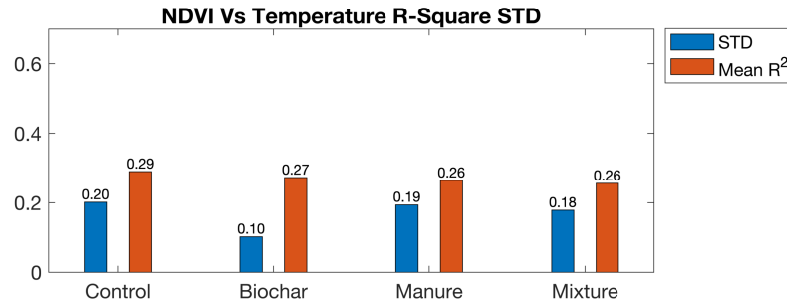
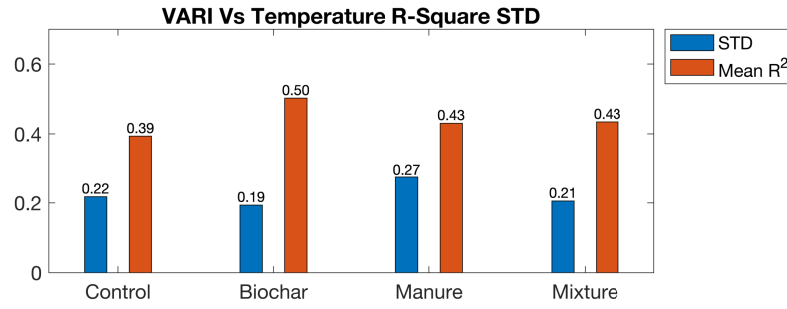
In Figure 2.28b, the correlation of NDVI and VWC at 7.5 cm is the most significant. Even if the control treatment has the maximum difference in STD and averages R^2 , the biochar treatment has the minimum STD and is not too low in R^2 . We believe the most reliable method is essential to improve the quantifying accuracy by using NDVI. In Figure 2.25 and Figure 2.26, the average VWC from biochar treatments is higher than the other treatment plots, which verifies that biochar could enhance water conservation and crop productivity. As seen in Figure 2.27, there is barely any relationship in these different treatments where VWC is at 15cm.

2.7.4 Discussion and Challenges

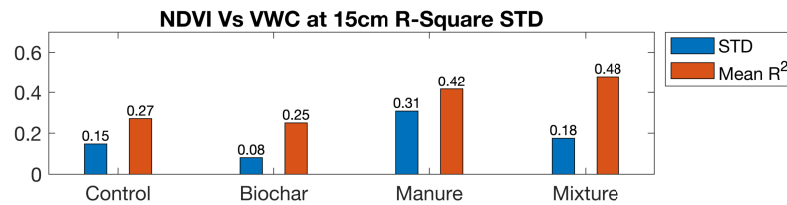
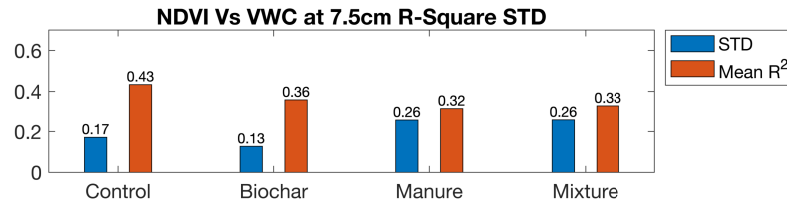
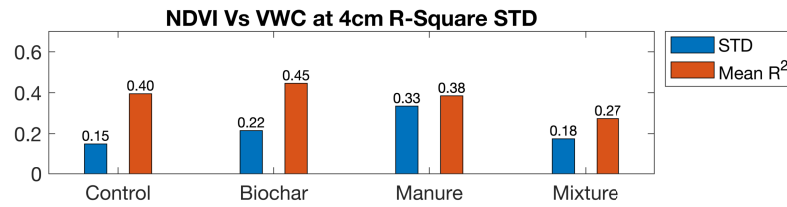
At Nov 17th, we can see the fitting lines of treatments are pretty short (Figure 2.24). Based on our literature review, biochar can reflect the near-infrared wave [157], this mechanism may affect the reflectance of biochar and natural green

plants. However, how it affects the reflectance of the biochar, mixture, and green plants needs more exploration in the future.

In November, the VWC is much higher than the rest of the month because farmers did flood irrigation once on the whole field. Although wheat growth is related to soil water content, correlation directions are different where VWC is at 4 cm and 7.5 cm. All biochar treatments have a positive correlation with VWC at 4 cm. However, the correlations of VWC at 7.5cm are various. This would also need to be further explored.



(a) The R^2 of VARI and NDVI versus temperature.



(b) The R^2 of VWC versus NDVI at different depth level.

Figure 2.28: a) The R^2 and STD for the VARI and NDVI versus soil temperature. b) The R^2 and STD of NDVI versus three different depths of the VWC.

In the future, we will provide further statistical analysis for our proposed regression model. We would plan to train the comprehensive model as well to fit these indexes with soil temperature and moisture so that we can use these indexes to explain soil carbon activities, such as methane and CO₂ emissions. We believe an effective method of quantifying carbon emissions still needs to be explored.

2.8 Case Study IV: Smart Bi-eBikes (SBB) - A Low Cost UGV Solution for Precision Agriculture Applications

2.8.1 Introduction

Crop pest management has vast economic consequences vital to agriculture, natural lands, and even public health. In this work, we contribute a new low-cost UGV solution to solve this precision agriculture problem. This UGV is made of two electric bikes (bi-eBike) that are commercially off the shelves (COTS) and driven by a ROS (robotics operational system) compatible UGV. Our bi-eBike system offers a mobile platform that can be used as a mobile sensing service with multiple sensors such as multispectral cameras, microwave scanners, etc., as well as a mobile actuation/application service with actuators such as UV-C light insecticide, beneficial bugs, growth stimulant spreaders or sprayers, etc. By mapping and imaging plants in the field, farmers can treat individual plants instead of troubleshooting the entire field, reducing both their costs and negative environmental impact. This smart bi-eBike system can be supplemented with solar panels (photovoltaic) and a UAV landing/charging pad. Thus, one can expect that the SBB can have a long operational duration (10 hours or more) and large coverage of acreages where UAVs are used for variability mapping for site-specific treatment. This section describes the system level concept, subsystem designs and integration, vehicle control electronics, autonomous navigation architecture, and some preliminary experimental results.

Previous pest management work has included biological control, artificial control, and organic control. Biological control involves using naturally occurring parasitic wasps and other spiders that prey upon *Lygus*' eggs and larvae. However, removing *Lygus* bugs in this manner is time-consuming and rarely achieves acceptable economic results. Artificial control is done by removing weeds from areas where strawberry fields, ditches, and field borders meet⁹. This stops *Lygus* from laying eggs in the weeds. Organic control utilizes non-chemical methods such as bug vacuums [158,159] to remove the *lygus* bugs. However, these methods have many disadvantages, and overall are inefficient at controlling insect populations. Bug vacuum machines are too expensive (approximately \$200,000), and are rarely economically viable for farmers. Thus, removing pests in a manner that maximizes strawberry yields with a low-cost solution is an important question. Although there are other similar products that use UGV for pest control management in the field [160], they

⁹<https://www2.ipm.ucanr.edu/agriculture/strawberry/Integrated-Weed-Management/>

are expensive or otherwise in the lab stage, which means they cannot scale up as necessary. Therefore, controlling pests that cause low strawberry yields with a low-cost solution is an essential challenge.

Ultraviolet (UV-C) irradiation has been shown in numerous studies to kill bugs and microorganisms. The UV-C light can deteriorate the pests' DNA with sufficient illumination, and the pests won't be able to repair their DNA if they are kept from direct sunlight for a period (dark period) greater than four hours [161]. Adjusting illumination time and frequency per week would significantly reduce the pests without impacting the overall health of the crop. Here, we present a low-cost solution using two mountain bikes converted to electrical bi-bikes. A fitted chassis turns them into a UGV mobile platform. By mounting UV-C light tubes and cameras, we enrich our SBB system's potential applications. Solar panels and UAV landing/charging pads would be future procedures to enhance the sensing and actuating coverage of the field.

The contributions of this work are as follows:

- We present a low-cost solution platformed on smart bi-eBikes for pest management (UV-C light) in large-scale fields.
- We perform a cost analysis to prove the cost efficacy of the SBB system in a commercial setting and demonstrate its value for the agricultural industry.
- We discuss the potential applications of mounting different sensors and actuators that make our SBB system scalable and robust.

2.8.2 The UGV Architecture

The UGV platform is made up of two electric bikes (bi-eBike) COTS driven by a ROS-compatible UGV. The bikes themselves are converted to electrical bikes from general bikes (COTS one). The total motor torque, power for max payload, and many other details are described in Section 2.8.2.1. Since we have two eBikes, a comprehensive chassis design is essential for eBikes to run smoothly and enhance the maximum pest management performance described in Section 2.8.2.2. Finally, the most important component is the eBike brain, which consists of a ROS-compatible platform on which the motor controller, eBike movement, and navigation system are implemented (Section 2.8.2.3).

2.8.2.1 Fundamentals of Bikes

To consider the off-road ability of our system, we decided to choose the mountain bike as our basis. It has a lightweight aluminum frame and a suspension fork to smooth out rough roads and trails. Alloy wheels with knobby tires for improved grip off-road. The original height is 40 inches, and the wheels are 26 inches in diameter, making them suitable for crop height (See Figure 2.29a).



(a) General mountain bike

(b) Electrical motor kit with battery

Figure 2.29: a) The general mountain bike with 26 inches tires, b) The electrical motor kit with battery, display, and controller.

An electrical motor kit for 26-inch bicycles is shown in Figure 2.29b. The thumb throttle controls acceleration by adjusting the duty cycle for the PWM wave. The nominal voltage and power are 48 V and 500 W, respectively, and the torque of the motor is 65 Nm. Therefore, based on these parameters, assuming ideal conditions, we can estimate the max payload of SBB as follows:

$$\tau = Fr, \quad (2.11)$$

$$f = \mu Mg = F, \quad (2.12)$$

where total torque should be 130 Nm (we have two motors as front wheel drive), radius r is 13 inches, g is 9.8 m/s^2 . The maximum payload means that, after the bikes achieve the desired speed, the motor force equals friction, given rolling friction coefficient $\mu = 0.08^{10}$, we have

$$M = \frac{\tau}{\mu rg} \approx 502.47 \text{ Kg} \approx 1107.76 \text{ lbs}. \quad (2.13)$$

Thus, our basic fundamental eBike framework has been established and the maximum payload is 1107.76 lbs, which would be optimal for accommodating a host of periphery such as crop-dusters [162], liquid sprayers, liquid tankers, generators, and so on. After that, we will aim to build a chassis to connect two bikes to form our UGV platform.

¹⁰https://www.school-for-champions.com/science/friction_rolling_coefficient.html

2.8.2.2 Chassis Design

A successful chassis design serves to provide a configurable, robust platform that optimizes for strength and cost. The proposed solution seeks to integrate the preexisting platform afforded by the use of the predefined elements (bicycles in parallel) and retain the functionality of a traditional chassis. A primary constraint guiding the design of the project is the emphasis placed on the use of prefabricated components. Rather than engineering an entirely novel approach, a straightforward solution comprised of readily available materials, open-source software, and an emphasis on affordability relative to existing solutions is the chief pursuit.

The final prototype utilizes a collection of prefabricated one-inch rails, fittings, and fasteners that together provide a rigid, relatively strong, and entirely modifiable frame that suitably joins the two bicycles together. Three points of contact are made at separate regions on the bike frame. The first six foot member, labeled 1 in Figure 2.30a, runs between the stems of each bicycle, replacing the fitted slot normally occupied by the handle bars. On each bicycle frame, a rubber-cushioned U-bolt with a top plate is fixed to the top tube. At both terminating ends of the U-bolt, rubber cushioned clamping mounts are fixed. The assembly is visible in Figure 2.30b.

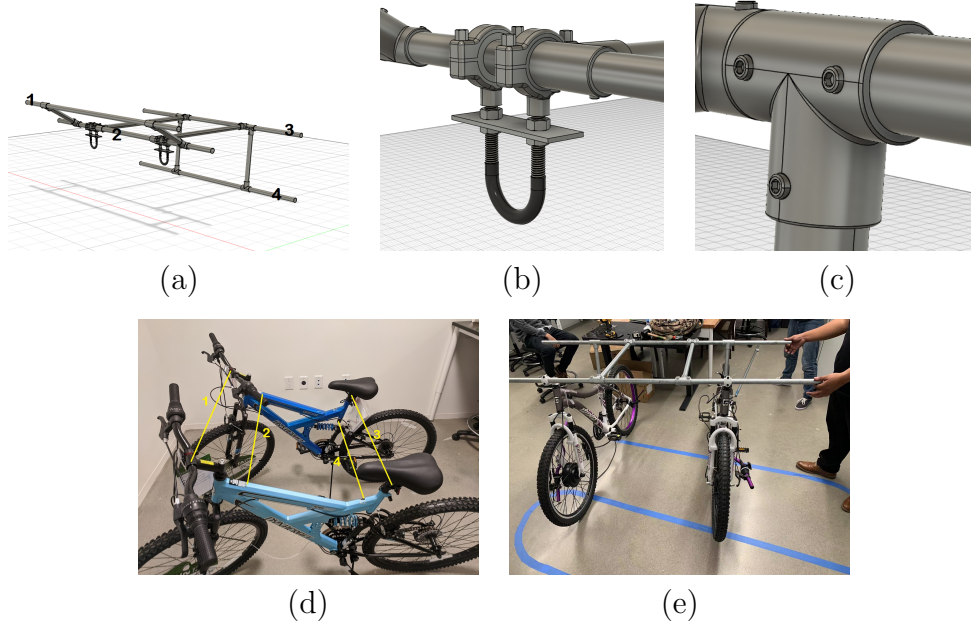


Figure 2.30: (a) A CAD model of the finalized chassis prototype, (b) The assembly connecting a spanning rail to the top tube of the bicycle, (c) Adjustable screws controlling the tensions at each terminal of the T-joint, (d) A minimal superposition of chassis geometry with component numbers as in (a), (e) Chassis is equipped on our smart Bi-eBikes.

A second six-foot rail travels the distance parallel to the first spanning member, inserted through the four clamp mounts of each U-bolt, labeled 2 in Figure 2.30a. The first and second rails are fixed together using a shorter member with through-hole T-joints on each end, the angled members running between points 1 and 2. The seat posts are replaced with a shorter rail with a T-joint on its terminating end, the rails being perpendicular to those as labeled 3 in Figure 2.30a. The final six-foot member spans the distance between the seat tubes of each frame, inserted into each T-joint provided by the replaced seat posts. In bicycles with through-hole seat tubes, as used in the project, a second T joint can be placed on the lower end of the member, replacing the seat post. An added six foot member can be placed through this last set of T-joints for increased stability. A rail traveling the length of the top tube connects the rail joining the seat posts and the rail attached top tube, the members spanning the transverse between points 2 and 3 in Figure 2.30a. Each T-joint used features three screws at each through-hole terminal, visible in Figure 2.30c. The correspondence of the chassis prototype to real hardware is shown in Figure 2.30d.

We also considered that using versatile fasteners allows for increased modularity

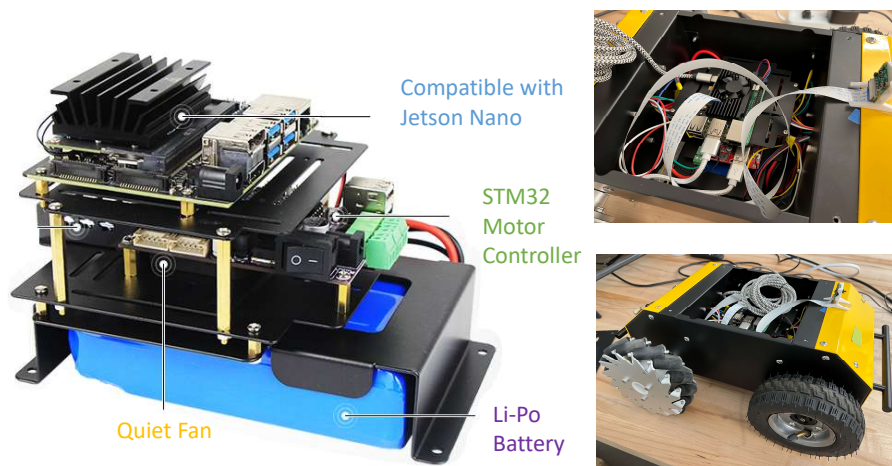
where each member can be swapped for longer or shorter components, and the span between bicycles can be modified as desired. These adjustable chassis dimension mean the vehicle can be manually adjusted to suit its environment. It is important to note that the stem and seat tube clearances are constraints for the design. Both dimensions were designed by the manufacturer for one inch diameter components. Finally, as shown in Figure 2.30e, we can see how the chassis is attached to the bi-eBikes.

2.8.2.3 SBB Smart Brain

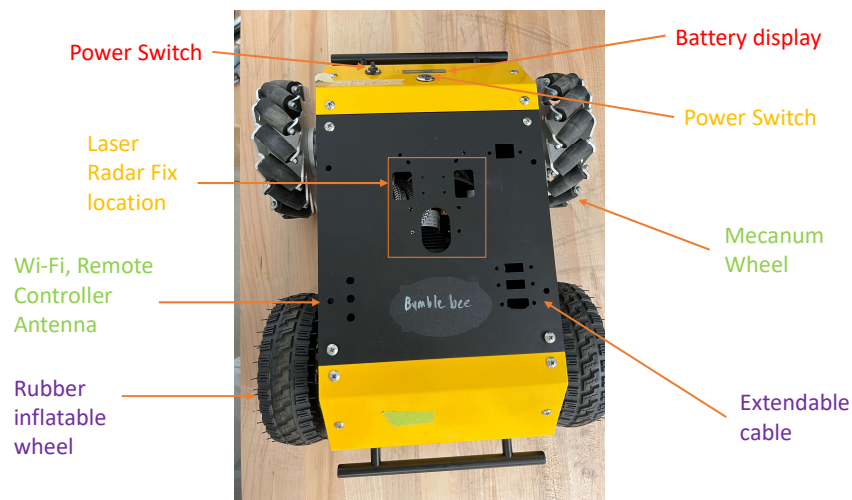
With a suitable chassis and means of locomotion, a control center is needed to allow for the UGV's primary capabilities. These functions are carried out through the use of computer vision and machine learning techniques on the edge, otherwise known as edge-AI device, by utilizing an on-board computer (e.g., Raspberry Pi¹¹) using ROS¹² to drive the UGV kinetic movement [163,164], visual signal processing, and pest management [165], such as controlling UV-C light presenting time. However, as building a control center as mentioned above from scratch is expensive, we decided to use an affordable ROS-ready robot to control our SBB ROS-ready, which is also referred to as a robot-driven robot. By hacking into this affordable robot and using the existing motor control logic to control the motor of our SBB, we can continue working on the navigation guidance on the robot's brain (Raspberry Pi). See Figure 2.31 for this affordable robot (less than \$250) we chose.

¹¹<https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>

¹²<https://www.ros.org/>



(a) STM32 controller and real hardware



(b) ROS-ready robot outside appearance

Figure 2.31: a) The STM32 board of motor controller and its extendable platform for related sensors, b) Switches and reserved holes.

The motor control board is equipped with the STM32F405RGT6 chips¹³. This STM32 can communicate with the Raspberry Pi through USB and provide 5V/5A for an external GPU such as the Jetson nano. This robot is open-source and can be programmed as needed to control the SBB. With this SBB “brain”, we can implement various behaviors, such as real-time obstacle avoidance, autonomous navigation, etc., with reduced effort.

2.8.3 Navigation Guidance from Image Segmentation

Our first mission is to train our SBB to recognize the road and strawberry fields. Then we can let our SBB move towards the correct path without colliding with any plants. Thus, image segmentation would be our first choice when it comes to visual processing. Image segmentation is a computer vision concept that labels certain sections of an image based on what is being presented. The purpose of semantic image segmentation is to assign a class to each pixel in a picture that represents what is being represented. This assignment is known as “dense prediction” because every pixel in the image is predicted.

Jetson TX2¹⁴ is used to implement an image segmentation algorithm to achieve the cognitive process in real time. Figure 2.32 depicts the findings of a field road detecting experiment. The Jetson TX2 is a very powerful and energy-efficient embedded AI computer device. It’s also based on an NVIDIA GPU with 8GB of RAM and 59.7 GB/s of memory bandwidth, allowing the SBB to react more quickly in real-time pest monitoring. If an object unexpectedly enters the SBB path, a cut-off command would be implemented immediately that can prevent danger to employees or crops.



Figure 2.32: Field road recognition, green color indicates true positive, red indicates false positive, and purple represents false negative.

¹³<https://www.digikey.com/en/products/detail/stmicroelectronics/STM32F405RGT6TR/4755958>

¹⁴<https://developer.nvidia.com/embedded/jetson-tx2>

Furthermore, researchers can use SBB, which contains an artificial intelligence platform, to undertake proximal sensing and mapping missions without any need for human involvement. This would allow the SBB versatility in many other scenarios. One of the most essential factors is real-time vision processing facilitated by cognition algorithms. The SBB movement behaviors are typically divided into three steps: stimulus perception, judgment, and decision making, and stimulus performance [141]. Applying cognitive algorithms creates a more comprehensive grasp of the scene based on objects and their spatial relationships. Using cognitive algorithms can even distinguish individual objects in the surroundings.

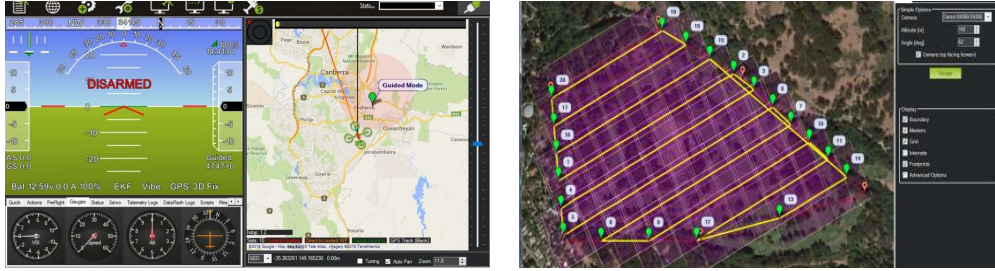


Figure 2.33: Mission Planner windows and path planning implementation.

Because the visual road sensing method based on image segmentation has been applied on the SBB, path planning or simultaneous localization and mapping (SLAM) on our SBB is possible. For our mission, we decided to use Mission Planner¹⁵ to set up survey grid waypoints that allow for autonomous driving to navigate the entire field. Because this method is relatively simple, it will enable our potential users to use it without extensive training (See Figure 2.33). The SLAM and other path planning can be applied to the SBB in the future for the related algorithm verification study.

2.8.4 Pest Management Network

The traditional methods of pest control management often involve applying chemicals to kill or mitigate the pests. However, these chemicals could have residual effects on crops and potentially impact human health. Therefore, using non-chemical techniques to mitigate the pests has been pursued. In several strawberry production techniques, UV-C irradiation has been employed to destroy microorganisms [166]. The UV-C light can kill the pests and prevent damage to crop [166]. In a greenhouse strawberry system, researchers [167] demonstrated the UV-C's potential in mitigating powdery mildew fungus and other fungal diseases. In [168],

¹⁵<https://ardupilot.org/planner/>

researchers demonstrated that UV-B radiation promoted resistance induction in strawberry plants against fungal infections. However, the damage inflicted upon plants at the scale needed to kill or reduce disease populations, and the use of UV-B technology in strawberry farming has been very limited.

When irradiated at night followed by a dark phase [169], the UV-C light has sufficient power to kill lygus [170], B. Cinerea, and P. Aphanis [161, 171] without harming the strawberry plant. Additionally, UV-C irradiation has also been demonstrated to promote plant growth and development [172], and low irradiation intensity of UV-C light will improve the productivity of the fruit [173]. Thus, we select UV-C light as our pest control management actuator since the literature shows UV-C light can mitigate the pests during the night with sufficient irradiation intensity and time, without damaging the plants.

Extensive research has addressed the optimal UV-C illumination intensity and exposure time on crops [174]. In our pest management section, we consider the UV-C light tubes as actuators, which dictates how many UV-C light tubes need to be placed and where they should be placed so that the irradiation effects can be maximized. Research shows an optimal configuration of 55-watt UV-C lights, located above the plants (30 cm away from the UV-C light). In this case, the UV-C light irradiation intensity is 0.206 W/m^2 , and in 60 s illumination corresponds to 12 J/m^2 [175]. To achieve the cover of the large-scale field during the night (after applying UV-C light, the crops must be subject to a 4-hour dark period [161] during which the bugs cannot sufficiently recover [175] as soon as possible). Two UV-C lights would be mounted at 55 watts each under chassis 2 and 3 members, respectively (Figure 2.30). The exterior of the SBB will utilize a reflective panel to increase the overall irradiation and cover the undersides of leaves: maximizing performance (no more than 50 J/m^2 and twice a week [175]). The exposure will have a negligible effect on the nutritional value of the crops [176]. The UV-C light performance evaluation will be implemented in the future.

2.8.5 Discussion and Challenges

The objective here is to explore the economic viability of using commercialized products to provide a low-cost solution for pest management. See Table 2.4 of each part price in detail. Compared with the existing solutions, such as bug vacuum machines (\$200,000¹⁶), and other personal customized solutions, we are competitive in cost-effectiveness and save lots of low-level motor control work, especially bus control, relative to other products. Potential customers, such as related agriculture research institutions and farmers, may be eager to pursue a solution amounting to only \$2,000 will mitigate the bugs and crop diseases without the purchase of chemicals.

¹⁶<https://organicgrower.info/article/california-betters-its-bug-vacuums-for-lygus/>

Item	Price	Unit
Mountain Bike	\$155	2
Motors Combo with Battery	\$379.9	2
Chassis Members Combo	\$649.49	1
lingao UGV robot	\$235	1
Raspberry Pi Camera	\$40	1
Raspberry Pi 4B	\$45	1
Total	\$2039.29	

Table 2.4: SBB system each component’s price in detail.

We currently use differential steering [177] to control the SBB direction. However, most scenarios for our SBB prove to be rough and uneven, which may cause rollover and failure to steer at sharp corners [178]. To make our SBB turn more efficiently, we will explore angular steering instead of differential steering. Angular steering needs to use an extra electric motor to control the steering, which reduces the force applied to chassis connectors. Although this procedure would increase the cost (extra motor and connectors), it would also increase the system’s robustness, especially when we are trying to transport a heavy load.

Because we are using machine learning algorithms to process the navigation strategy in real-time, an important issue is that we need to build a specific image dataset of the strawberry field and perform the benchmark that labels the road. This will be a tedious work to do in the future, and we need to seek cooperation with the industry to do the autonomous collection of images for preparing datasets.

Further areas of interest would be integrating other functional groups into the control system: solar panels and a UAV landing/charging pad [179]. The can greatly increase the operational duration and have coverage of large areas where drones could be employed for variability mapping for site-specific management.

2.9 Chapter Summary

In this chapter, we addressed the XAIoT data acquisition platform, such as microwave and millimeter radar array sensors, UAV and UGV, etc. We also demonstrated the possibility of identifying different biochars with their raw materials, usage amount, and water retention capacity. Our classifier based on our feature extraction achieves 74.0% identification accuracy when identifying the biochar amount of applying it into the soil. The water retention capacity also is validated that biochar can almost double the water retention ability than plain soil. Our final goal is to improve the simplicity and accuracy of biochar application, and soil condition monitoring in real time of such sensing systems in the large-scale field.

In addition, we demonstrated the ability to identify different amounts of soil carbon with their origin, usage, soil moisture, and soil moisture under the effects of

carbon application using mm Waves. The machine learning classifier we used is a lightweight decision tree that achieves an accuracy of up to 97.5% for soil moisture with carbon application. The millimeter wave has been verified to distinguish soil carbon, soil moisture, water retention ability, and soil carbon content sensing. In the future, we plan to use the mm-waves sensor in UAVs and UGVs to make such sensing systems easier to use and more accurate for biochar application and soil condition monitoring in large-scale field.

In case study III, we validated a field study in which UAVs were used to check on the soil for a whole growing season after biochar was added as a soil amendment. The relationship revealed between soil temperature and volume of water content can be described by linear regression models. After applying different treatments, we illustrated that the soil temperature and VWC can still be correlated by the VARI and NDVI indexes. Biochar as a soil amendment would enhance the sensing accuracy of those two indexes mentioned, and the correlation using VARI is about 46.77% higher than the non-treatment plots, and the NDVI index is 52.94% higher than the other treatments. The NDVI shows good prediction results where VWC will be measured at 7.5 cm depth. All in all, our results show that image index analysis has the potential to become a fast way to measure the condition of soil when adding biochar.

For case study IV, we present a low-cost UGV solution for precision agriculture applications, specifically for pest management. We first retrofitted two mountain bikes into bi-eBikes, equipped with a controller as the center of the edge-AI ROS-ready device. The overall bi-eBikes system is capable of carrying a heavy payload of up to 1107.76 lbs. These properties would enable the SBB system with rich possibilities in agriculture applications, and allow for a host of industrial tasks in precision agriculture. Economic analysis shows that using COTS would reduce hardware and software solution costs significantly. Future work will extend the system's functionalities and applications, enhancing its contribution to precision agriculture.

Chapter 3

NON-INTRUSIVE SOIL CARBON CONTENT QUANTIFICATION METHODS USING MACHINE LEARNING ALGORITHMS: A COMPARISON OF MICROWAVE AND MILLIMETER WAVE RADAR SENSORS

3.1 Introduction

Carbon in the soil plays a crucial role in ecosystems by retaining and delivering plant nutrients, improving soil structure, reducing erosion, and enhancing water retention capacity. Generally speaking, the soil is regarded as a carbon-rich sink. Unfortunately, poor management, vegetation shifts, and climatic changes have led to a decline in soil carbon reserves worldwide. The vast worldwide area of depleted soil carbon reserves has focused on the management of these ecosystems by carbon sequestration as a method of mitigating climate change. Many agricultural practices, including conservation tillage and cropping, have been studied to assess their effectiveness in promoting soil carbon sequestration. However, assessing the effect of soil carbon, particularly the widely used biochar application [127], on carbon sequestration in a large-scale field remains a difficult, time-consuming, labor-intensive, and expensive task.

Biochar is a solid material that results from the pyrolysis of biomass. When applied to soil, it can improve soil functions and help to reduce GHG emissions that would otherwise occur from the natural decomposition of biomass. The goal of producing biochar is to create fine-grained and highly porous charcoal that aids in the retention of nutrients and water in the soil. Biochar is a stable, solid, carbon-rich substance that can persist in soil for thousands of years. By using biochar as soil amendment, it is possible to create a “carbon-negative” system. The soil fertility benefits of biochar include its exceptional ability to adsorb nutrients and other molecules, as well as its extraordinary stability in soil. In addition, sustainable methods of producing biochar can result in the production of oil and gas byproducts that can be used as clean, renewable energy. However, to achieve a carbon-negative system, it is essential to carefully measure and quantify the amount and type of biochar used, along with other monitoring and measurement methods related to soil

carbon. Therefore, a cost-effective quantification method is necessary for fighting these challenges.

Previous studies on measuring soil carbon and related components have mainly focused on collecting soil samples and using expensive analyzers [143], costing millions of dollars, to determine the various components of the soil [144]. Some researchers have used portable devices that can be placed directly in soil fields or embedded sensors in the soil for detection [145]. Additionally, a few studies have utilized remote sensing technologies, such as satellites [180], UAS [146, 147], and UGVs [181], to map soil conditions and soil carbon emissions [182], including GHG emissions. However, some prior attempts have used contactless methods [182], such as microwaves, to detect polymer composites and compare electrical characteristics based on biochar. These methods are not related to agriculture and are not relevant to our mission. These previous attempts have primarily relied on expensive analyzers and are limited by environmental factors, such as soil moisture level, soil temperature, wind, etc., making it difficult to operate or collect valid samples during rainy days, human activities, etc. [148]). Additionally, collecting soil samples requires persistent sampling, which can be labor-intensive and challenging to scale up on a large-scale field. Therefore, we have presented a microwave and a millimeter wave-based soil carbon content sensing system and the following benefits:

- Compared with the traditional method of sensing soil carbon or biochar, the non-intrusive method is less labor-intensive.
- Compared with the traditional method and microwave of sensing soil carbon content, the millimeter wave radar sensing method is less labor-intensive and more accurate.
- A portable mobile device can be easily equipped on an autonomous drone or vehicle to detect soil on a large scale in a cost-effective manner, without negatively impacting the soil.

As a result, two challenges are addressed: 1) how to perform the feature extraction of soil carbon content amount and different types of soil carbon from microwave and millimeter wave reflection signals, and 2) how to leverage these unique microwave and millimeter wave signal features to infer soil conditions. Our system employs microwave/millimeter-wave transmitters and receivers, integrated with a microprocessor, to sense and classify characteristics of biochar including material usage quantities, and water retention capacity. The contributions of this work are delineated as follows:

1. To the best of our knowledge, this is the first time to demonstrate that microwave/millimeter wave signals being reflected by soil biochar mixture can be used to identify soil carbon content and its footprints.

2. We revealed a potential relationship between features from microwave/millimeter wave reflection and soil carbon content usage based on physical principles.
3. We validated our system by conducting real-world experiments and using machine learning algorithms.

3.2 Related Work

Almost all measurement methods for SOC and GHG emissions involve collecting samples from the soil at selected sampling positions. After air-drying, grinding, and weighing samples [183], the researchers usually put these samples into element analyzers. The process is usually considered costly and labor-intensive. Moreover, some researchers also use portable devices to detect carbon-based gases such as methane, CO₂, and other GHGs that are leaking from soil [184]. This manuscript is mainly focused on soil organic carbon contactless sensing. The related work of soil carbon content sensing can be categorized into two areas: 1) remote sensing of soil carbon content, 2) proximal sensing of soil carbon content. In the following review, we mainly explore the existing approaches and the remaining research gap for soil carbon content sensing mechanisms and sensing configurations.

3.2.1 Remote Sensing of Soil Carbon Content

Remote sensing techniques for SOC content mainly use spectral sensors to estimate SOC, and methods difference based on sensing temporal, spatial, and spectral resolution. Based on the spectral sensor mounting platforms, we can simplify and categorize them into spaceborne, airborne, and UAVs [61].

The spaceborne usually uses satellites equipped with multispectral or hyperspectral [185] to map soil properties. Several studies use regression models to analyze spectral data acquired from satellites. For example, Gomez et al. [60] used data from the Hyperion sensor on board the EO-1 satellite (400–2500 nm) and compared the predictions of SOC to in-situ VNIR-SWIR measurements. The algorithm usually contains partial least squares regression (PLSR) [60], ANNSK [186], auto-PLSR [79], PLSR/RF [74], PLSR [73], SVM [77] for estimation and classification of the SOC. However, there are still limitations to the spaceborne measurement of SOC due to temperature, air pressure, spectral calibration, and the height of the crop. Moreover, the satellite is expensive [187].

The airborne sensors for sensing SOC are more flexible than satellites, and sensing results aren't highly affected by clouds, air pressure disturbances, and so on for the spectral sensors [188]. Using airborne sensors to sense SOC has first been verified by Stevens et al. [81] that equipped spectral range of 430–2540 nm over large soil areas (around 420 km²) of various soil types. The spectral data acquisition from airborne still needs sensing calibration due to large pixels and sensing stability and sensitivity [188]. Furthermore, it is easier and better to sense

SOC on bare soil, but it is difficult to find large-scale bare soil. Thus, several studies estimated the SOC by using several vegetation indexes such as NDVI [108] and other plant-based model [89]. Moreover, since this is still spectral data, the algorithm for spaceborne used to estimate the SOC can also be applied to airborne. The advantages of using airborne sensing methods are mainly on rapid time windows, the ability to select optimal sensing conditions such as temperature and moisture, and the comprehensive payload options for multiple spectral sensors and other sensors for sensor fusion that are benefits for sensing SOC further.

The UAVs have been becoming more and more popular in precision agriculture in recent years [108, 189–191]. The UAVs are small compared with aircraft, they have less revisit sensing time and high spectral resolution due to their adjustable flight altitude, which is more flexible than satellites and aircraft. The most recent studies [192–194] used UAVs to estimate SOC under different scenarios. Beltrán-Marcos et al. [192] highlights the benefits of combining satellite and UAV images to produce spatially and spectrally enhanced images, which are relevant for estimating the main impacts on SOC in burned forest areas. Biney et al. [193] demonstrated that SOC can be estimated with high accuracy in completely bare soil UAV imagery with other auxiliary data, and Näsi et al. [194] evaluated the relationship between drone image data, field, and soil quality indicators.

Overall, remote sensing for estimating SOC has several advantages, including no negative impact on the soil. Furthermore, it can scan large-scale fields in an efficient way and gather soil properties in those difficult-to-reach areas. Nevertheless, remote sensing has a low signal-to-noise ratio [195], low spectral resolution [60], and is affected by atmospheric and topography disturbance [196]. Moreover, remote sensing is also highly affected by soil moisture, crops, soil structure, etc. [197] which needs further improvements in estimating SOC.

3.2.2 Proximal Sensing of Soil Carbon Content

Proximal soil sensing is defined as the use of field-based sensors that are in close proximity to the ground, or plants, such as on a harvester, or ground-based sensor. Proximal soil sensing may be described by their measurement types, which is invasive and non-invasive, in situ and ex situ, active sensing and passive sensing (depending on their transmitter source, from sunlight or artificial sources (microwave or millimeter wave)), stationary sensing, and mobile sensing, and the obtained data for inferring soil properties in a direct or indirect way [198, 199]. Additionally, the high sampling density of proximal sensors makes them better at detecting field variability. This solves the problem of picking the right soil sampling strategy that ensures representative samples.

Based on the wavelength range that is being used by proximal sensors, the sensing types can be categorized by their sensing type and application scenarios,

such as radio spectrum, ultraviolet, visible, and infrared spectrum, short wave infrared (UV-VNIR-SWIR), and X-rays [198]. For the radio spectrum, several studies have been conducted, such as time domain reflectometry (TDR) [200], frequency domain reflectometry (FDR) [201, 202], nuclear magnetic resonance (NMR) [203], and ground penetrating radar (GPR) [101, 204, 205]. However, these radio waves are usually used for sensing soil water content and not yet for soil carbon content.

Using UV-VNIR spectroscopy to sense soil carbon content is widespread since the sensing mechanism overlaps with remote sensing (hyper-spectral and multiple spectral sensing), whereas a VNIR-SWIR spectrometer should be placed as close as possible to the sample. UV and visible light are usually used for classifying inorganic species [206], and the visible and near-infrared (VIS-NIR) applied for soil and crop analysis due to most plants reflecting in NIR band [207–211]. The mid-IR spectra are applied for analyzing soil properties, minerals, and other organic compositions. This is because the molecular size of the soil components approaches the mid-IR wavelength.

The portable X-ray fluorescence (pXRF) is used to estimate the clay, sand, silt, heavy metals in soil, and other inorganic species [212, 213]. pXRF is dependent on the fluorescence at a specific energy of atoms that are stimulated by X-rays. The detection of specific fluorescent photons permits the qualitative and quantitative study of a sample’s constituent parts.

To sum up, each of the proximal sensing technologies described above has its own strengths and weaknesses. Compared to remote sensing, proximal sensing is less limited by environmental restrictions and can collect fine-grained information about the soil. However, proximal sensing is not efficient on large-scale fields, in which it needs to know where to sense, why to sense, who to sense, etc., and the currently market-available proximal sensors are heavy and expensive compared with hyperspectral spectrometers in general.

3.3 Materials and Methods

In this section, we will illustrate the sensor hardware setup and physical principle of using microwave and millimeter waves for sensing SOC.

3.3.1 Sensor Hardware - Microwave Sensing Module

The sensing module consists of three major parts: the microwave transmitter and receiver (Walabot¹), the Raspberry Pi 4B, and the visualization interface shown in Figure 3.1. Walabot is a programmable three-dimensional sensor that can penetrate into objects using radio frequency technology. Walabot illuminates the space in front of it with an antenna array (18 antennas), detecting reflected signals and recording them with the VYYR2401 A3 System-on-Chip integrated unit.

¹<https://walabot.com/>

In addition, Walabot operates over an ultra-wideband (UWB) range of frequencies corresponding to the regulatory domain of the model, with a 3.3–10 GHz range. The average transmission power is below -41 dB/MHz, which will not impact human health [104]. Note that since the Walabot is working in the range of 3.3–10 GHz, we better operate it without Wi-Fi (in the 5 GHz band).



Figure 3.1: System hardware, left is the Raspberry Pi 4B with visualization user interface, and right is the Walabot microwave transmitter and receivers.

3.3.2 Sensor Hardware - Millimeter Wave Sensing System

The sensing module consists of three major parts: the mm-wave transmitter and receiver (IMAGEVK-74), the Raspberry Pi 4B, and the visualization interface shown in Figure 3.2. The sample was placed under the direct mm-waves radar sensor, less than 2 cm.



Figure 3.2: Millimeter wave radar sensor with Raspberry Pi Setup.

The soil was collected from almond tree fields and filtered to remove coarse-grained substances such as stone. The biochar (we mentioned earlier, and it also shows in Figure 2.16) made from 100% of orchard removals, walnut shells, 100% yellow wood, almond shells, pinewood. In the following experiments, the mixture rate of soil and biochar was varied from 0% to 100%.

The sampling frequency of the mm-wave sensor is from 62 GHz to 69 GHz. We sampled each sample 20 times for the purpose of reducing likelihood error. We set up 150 sampling points during the frequency band we mentioned, and the RBW is 100 kHz for reasonable sweep time and lower noise. Our experiments utilize PCA to reduce the dataset dimensions so that we can reduce the later classifier computation resources. To evaluate our system’s performance fairly, we introduced a confusion matrix to calculate average classification accuracy. The 5-fold cross-validation is also employed to adjust hyper-parameters and validate all models’ performance on our soil carbon samples. The accuracy and confusion matrix are presented to evaluate the classified objects’ performance.

3.3.3 Microwave Reflective Signal Classification

Biochar is made from different raw materials, which means it has different structures [127], especially its surface has many small porous structures. Therefore, when the microwave signal goes through these structures, it will be penetrated,

absorbed, and reflected in our microwave signals. In other words, we can get biochar information such as amount usage, different raw materials, water retention capacity, etc., by analyzing the receiving signal (reflective signals). To better understand what kind of features we are going to face, we draw the reflective signals shown in Figure 2.7. We collected reflective signals from different kinds of biochar and showed them in the time domain and frequency domain, respectively.

In Figure 2.7, the reflective signals from different biochar show distinguishable variations in both time and frequency domains. Since the microwave attenuated very fast due to its shorter wavelength, here we only show a 100-millisecond signal in the time domain (no other signals beyond this time interval), and all magnitudes are distributed from 5.0 GHz to 6.5 GHz. As seen in Figure 2.7, the positions of peaks and troughs in the frequency domain vary across various biochar types, which can be utilized to classify them.

3.3.4 Millimeter Wave Reflectance Analysis

As mentioned in Section 3.3.3, different biochar structures and soil carbon content can be detected by microwave signals. The wavelength of the millimeter wave is shorter than the microwave, which means it should gather more fine-grained spectrum information on biochar/SOC, as shown in Figure 3.3. In addition, as mentioned earlier, this millimeter-wave sensor has a total of 20 transmitters and 20 receivers, which means 400 signals are received from multiple paths. Therefore, we used a power delay profile (PDP) that generates the magnitude of the signal received through a multi-path channel as a function of time delay, where the time delay is referred to as the travel time between multi-path arrivals. In Figure 3.3, the y -axis is the normalized amplitude, which helps read all 20 paths reflections. The x -axis is the distance that is generated by the time vector and then multiplied by light speed (3×10^8 m/s). The overall transmitting frequency is from 62 to 69 GHz and uniformly sampled 100 points of the signals.

From Figure 3.3, we obtained that the reflectance of the millimeter wave can also be used to identify the different biochar/SOC. The physical principle behind this sensing mechanism is the same as microwave sensing. We will utilize these features to justify whether millimeter-wave signals are better than microwave signals for sensing SOC.

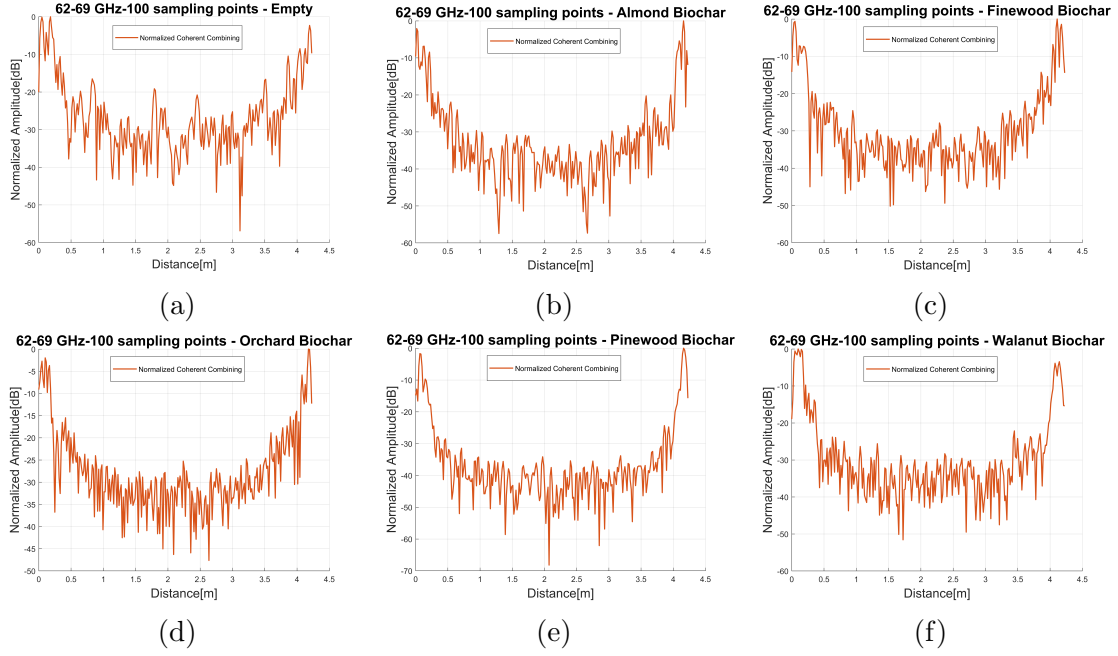


Figure 3.3: Millimeter wave reflectance signal analysis using PDP. (a) is the empty sensing signals (sensing background), (b)(c)(d)(e)(f) are the millimeter wave reflectance signals of different biochar materials.

3.3.5 Conventional Classification Methods

In this section, we discuss how to construct classification models that, based on the time-domain and frequency-domain characteristics of the received signal, can discriminate precisely between various objects. Furthermore, since this is the first time utilizing microwave and millimeter wave reflectance to estimate the SOC, we don't know which classifiers are best suited to classify these features. Thus, we will discuss the most common classifiers regarding advantages, disadvantages, and options for hyperparameters, and they are shown in Table 3.1 and Table 3.2.

To further emphasize the mathematical foundation of processing microwave and millimeter wave reflectance signals and leverage the implicit physical knowledge of microwave and millimeter wave relations, our system takes a data-driven approach by selecting a commonly used supervised classifier, such as SVM [128], the decision tree, KNN, LDA/QDA, naive Bayes classifiers, and a DNN for identifying microwave and millimeter wave reflectance.

3.3.5.1 Support Vector Machine

SVMs are a type of supervised learning models designed for classification tasks. The fundamental principle of SVM is identifying the hyperplane that most effectively

divides the data points into distinct classes.

To define a linear SVM, consider a dataset $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ with each $x_i \in \mathbb{R}^p$ representing an input feature vector of p dimensions, and $y_i \in \{-1, 1\}$ denoting the class label of x_i . The SVM aims to discover the hyperplane that separates the classes, which can be represented as:

$$w^T x + b = 0. \quad (3.1)$$

Here, w is the weight vector, x is the input vector, and b is the intercept. The classification function becomes:

$$f(x) = \text{sign}(w^T x + b). \quad (3.2)$$

To identify the optimal hyperplane, the SVM tackles the following minimization problem:

$$\min_{w, b} \frac{1}{2} \|w\|^2, \quad (3.3)$$

under the constraint:

$$y_i(w^T x_i + b) \geq 1, \text{ for all } i. \quad (3.4)$$

For data that cannot be separated linearly, the soft margin approach introduces flexibility by allowing certain misclassifications. This is done through slack variables $\xi_i \geq 0$, leading to an optimization problem:

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i, \quad (3.5)$$

with the constraints:

$$y_i(w^T x_i + b) \geq 1 - \xi_i, \text{ for all } i, \quad (3.6)$$

$$\text{and, } \xi_i \geq 0, \text{ for all } i. \quad (3.7)$$

In this case, C serves as a regularization parameter to balance the margin width and the classification error.

To handle non-linearly separable data, SVM utilizes the kernel technique, which involves mapping the input features into a higher dimensional space where linear separation is possible. The optimization framework remains similar, but the linear dot product $w^T x$ is substituted with a kernel function $K(x_i, x_j)$ that computes the dot product in the higher-dimensional space. Examples of kernel functions include:

- Linear: $K(x_i, x_j) = x_i^T x_j$,

- Polynomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d$,
- RBF: $K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$,

where γ , r , and d are kernel-specific parameters that can be fine-tuned.

SVM is typically used to discover boundaries in regions with sparsely labeled and imbalanced data. With SVM, the system can determine which brand and how much biochar are being used. The motivation for using SVM to classify the microwave/millimeter wave is that the SVM works well with high-dimension features captured (each sample test has 8049 sampling points).

3.3.5.2 Decision Tree

Given a dataset D comprising instances x_i and corresponding labels y_i , for $i = 1, \dots, N$, where N is the number of instances, the objective is to segment D into subsets through decision rules that enhance the purity of y_i within these subsets.

The objective function $J(D, f, \tau)$ assesses the quality of partitioning dataset D by a feature f and a split point τ , where the choice of J depends on the task (classification or regression) and may use metrics such as Gini impurity, entropy, or mean squared deviation.

The decision tree is built from the root with the entire dataset D , proceeding recursively as follows:

1. **Optimal Partition Selection:** For each feature f and each candidate split point τ , divide D into D_{left} and D_{right} where:

$$D_{left} = \{(x, y) \in D \mid x_f \leq \tau\}, \quad D_{right} = D \setminus D_{left} \quad (3.8)$$

Select the feature and split point that maximize the objective function:

$$(f^*, \tau^*) = \underset{f, \tau}{\operatorname{argmax}} J(D, f, \tau) \quad (3.9)$$

2. **Perform the Split:** Use the best feature and split point to divide D into two subsets.
3. **Recursive Division:** Apply the above steps to both D_{left} and D_{right} until a stopping criterion is reached, such as maximum depth, minimum node size, or minimal increase in J .

The recursion stops upon reaching a predefined stopping condition, creating a leaf node. In classification, the leaf's prediction is often the majority class among its instances. For regression, the prediction is typically the average of the target values

in the leaf. Predictions for new instances are made by following the path from the root to a leaf, applying the decision rules which are as follows:

Classification:

$$\hat{y} = \text{majority}\{y_i | (x_i, y_i) \in D_{\text{terminal}}\}$$

Regression:

$$\hat{y} = \text{average}\{y_i | (x_i, y_i) \in D_{\text{terminal}}\}$$

This framework properly represents the fundamental principles of decision tree algorithms, emphasizing the iterative division of the dataset according to an evaluative objective function, resulting in a structure that enables predictions to be made.

3.3.5.3 Linear/Quadratic Discriminant Analysis

Both LDA and QDA can be derived from straightforward probabilistic models that represent the class-conditional distribution of the data $P(X|y = c)$ for each class c . Predictions can then be obtained by using Bayes' rule, for each training sample $x \in \mathbb{R}^d$:

$$\begin{aligned} P(y = c|x) &= \frac{P(x|y = c)P(y = c)}{P(x)} \\ &= \frac{P(x|y = c)P(y = c)}{\sum_l P(x|y = l)P(y = l)}, \end{aligned} \quad (3.10)$$

we select the class whose posterior probability is maximized. Specifically, $P(x|y)$ is characterized as a multivariate Gaussian distribution with density for linear and quadratic discriminant analysis.

$$P(y = c|x) = \frac{1}{(2\pi)^{d/2} |\sum_k|^{1/2}} e^{-\frac{1}{2}(x - \mu_c)^t \sum_c^{-1} (x - \mu_c)}, \quad (3.11)$$

where d is the number of features, μ_c is the mean of the sample, and $\sum_c$ is the covariance matrix.

3.3.5.4 K Nearest Neighbors

The principle underlying nearest neighbor methods is to locate a predetermined number of training samples that are geographically closest to the new site and predict its label. The number of samples may be a user-defined constant (k-nearest neighbor learning), the distance may be any metric measure, and the most common option is the standard Euclidean distance. Regarding the classification rule, we denote p_i as the probability that sample i will be accurately classified in accordance with stochastic nearest neighbor rules in the learned space, and the objective is to

maximize the sum over all samples of the probability p_i that sample i is accurately classified:

$$\begin{aligned} & \underset{L}{\operatorname{argmax}} \sum_{i=0}^{N-1} p_i, \\ & \text{where, } p_i = \sum_{j \in C_i} p_{ij}, \end{aligned} \quad (3.12)$$

where N is total number of samples, C_i is the set of points in the same class as sample i , and p_{ij} is the softmax over Euclidean distance in the space. In particular, each point i selects another point j as its neighbor with probability p_{ij} [214].

3.3.5.5 Naive Bayes

Naive Bayes method is a set of supervised learning algorithms based Bayes' theorem with the "naive" assumption of conditional independence between every pair of features given the class variable value. Given class variable y and dependent feature vectors x_1 to x_n , Bayes' theorem asserts the following relationship [134]:

$$P(y|x_1, \dots, x_n) = \frac{P(y)P(x_1, x_2, \dots, x_n|y)}{P(x_1, \dots, x_n)}, \quad (3.13)$$

using the naive assumption of conditional independence, this classification rule can be simplified for all i :

$$\begin{aligned} P(y|x_1, \dots, x_n) &= \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_1, \dots, x_n)}, \\ \Rightarrow \tilde{y} &= \underset{y}{\operatorname{argmax}} P(y) \prod_{i=1}^n P(x_i|y). \end{aligned} \quad (3.14)$$

The primary difference between the various naive Bayes classifiers is the assumptions they make regarding the distribution of $P(x_i|y)$. In this study, we employed Gaussian Naive Bayes, where $P(x_i|y)$ is defined as follows:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right), \quad (3.15)$$

where parameters σ_y and μ_y are estimated using maximum likelihood.

3.3.5.6 Neural Network

We used a feedforward, fully connected neural network for classification. The first fully connected layer of the neural network has a connection from the network input (predictor data), and each subsequent layer has a connection from the previous

layer. Each fully connected layer multiplies the input by a weight matrix and then adds a bias vector ($y = wx + b$). An activation function follows each fully connected layer. The network's classification scores (posterior probabilities) and predicted labels are generated by the final fully connected layer and following the softmax activation function.

Classifiers	Advantages	Disadvantages
Fine Decision Tree Max number splits=100	Good predictive accuracy in low dimensions.	Prone to overfitting and instability, sensitive to small changes in data.
Coarse Decision Tree Max number splits=4	Easy to interpret and visualize, works well with non-linearly separable data, good at handling missing values	Less accurate than fine decision tree, may overfitting.
Discriminant Analysis (DA)	Fast, accurate and easy to interpret. Good for most high dimension data.	Assumes linear separation between classes, may not work well for non-linearly separable data.
Gaussian Naive Bayes	Easy to interpret and useful for multiclass classification, good for text classification and spam filtering	Assumes feature independence may not be true in reality, may not work well for highly correlated features.
SVM	Effective in high dimensional data, effective number of dimensions is greater than number samples.	Less effective for non-linear separable data, sensitive to outliers and misclassified points.
Fine K-Nearest Neighbors (KNN) Numbers of Neighbors =1	Accurate for complex data and variable relationships.	Computationally expensive for large datasets, sensitive to outliers
Coarse K-Nearest Neighbors (KNN) Numbers of Neighbors =100	Good predictive accuracy in low dimensions, less computation source required.	Less accurate than fine KNN.
DNN	Capability to Handle Complex Data Designed to handle large-scale datasets.	Overfitting, Computational Complexity, lack of interpretability

Table 3.1: Advantages and disadvantages of different classification algorithms.

Table 3.2 shows the hyperparameter options for all classifiers. These classifiers come from the Matlab 2023b classifier learner APP. Optimal classifiers are the processes that APP automatically searches among those hyperparameters and their optimal values. Selecting these classifiers is motivated by their capacity to train a large number of largely uncorrelated models working as an ensemble of more accurate predictors than individual predictors, and it is a flexible, user-friendly supervised learning method that, even without hyper-parameter modification, typically yields excellent results. In addition, our final objective is to try to develop the machine learning model on Edge-AI devices to do in situ soil carbon content quantification. Low cost is also what we are concerned about for millimeter waves since the sampling features and dimensions are not high (400×1024) compared to microwave data sampling.

Classifiers	Hyperparameters options	Value
Fine tree	Maximum number of splits (Use Gini's diversity index)	100
Coarse tree		4
LDA/QDA	Covariance structure	Full/Diagonal
Naive Bayes	N/A	N/A
SVM	Linear Gaussian Quadratic Cubic	N/A
Fine KNN	Number of neighbors	1
	Distance Metric	Euclidean
	Distance Weight	Equal
Coarse KNN	Number of neighbors	100
	Distance Metric	Euclidean
	Distance Weight	Equal
DNN	Number of fully connected layers	3
	First layer size	10
	Second layer size	10
	Third layer size	10
	Activation function	ReLU
	Iteration limit	1000

Table 3.2: Available hyperparameters options for all classifiers. N/A stands for not available.

3.3.6 More Optimized Stochastic Configuration Networks (MOSCNs)

Beyond the conventional classification methods mentioned in previous subsections, this subsection mainly explores a new framework known as SCN (stochastic configuration network) and an interesting investigation of using different randomness, including heavy-tailed randomness, that is shown to achieve more optimal SCN performance than the existing SCN.

3.3.6.1 Stochastic Configuration Networks

A SCN is a relatively new neural network compared to the existing randomized learning algorithms for single-layer feed-forward neural networks (since 2017 [215]). A supervisory mechanism (an inequality constraint) enables the SCN to randomly assign the input weights ω and biases b of the hidden nodes, thereby enabling universal approximation and learning capabilities for classification and regression. To determine proper (optimal) input weights and biases for further enhancing the generalization of SCN itself and proposing a suitable architecture when processing various scenarios datasets still remains a significant challenge.

SCN can deliver high generalization performance at a faster rate than other well-known randomized learning algorithms, such as the Radial Basis Function (RBF) Networks [216] and the Random Vector Functional-link (RVFL) [217]. There are three distinct varieties of SCN algorithms: SC-I, SC-II, and SC-III. The SC-I algorithm employs a constructive scheme to evaluate output weights only for newly added hidden nodes [218]. All previously obtained output weights are maintained unchanged. The SC-II algorithm recalculates a portion of the output weights by analyzing a local least squares problem with a shifting window size specified by the user. By resolving a global least-squares problem, the SC-III algorithm discovers all output weights simultaneously.

SCN is an incremental fast modeling method. Its network structure consists of a hidden layer node, using the data-driven supervision mechanism to randomly configure the input weight and bias of the hidden layer node, and gradually increase the number of hidden layer nodes. SCN uses the least squares method to calculate the output weight to complete the network construction. The network structure of SCN is similar to the traditional single hidden layer feed-forward neural network, and its mathematical expression [215, 219] is shown below.

Given a set of training data $X = \{x^1, x^2, x^3, \dots, x^N\}$, where each $x^i = \{x_1^i, x_2^i, x_3^i, \dots, x_d^i\} \in \mathbb{R}^d$, and the corresponding labels $Y = \{y^1, y^2, y^3, \dots, y^N\}$ where $y^i = \{y_1^i, y_2^i, y_3^i, \dots, y_m^i\} \in \mathbb{R}^m$ for $i = 1, 2, 3, \dots, N$. Given a target mapping $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$, assume the $L - 1$ th hidden layer node of the Stochastic Configuration Network (SCN) has been established. The function of this layer, $f_{L-1}(X)$, is then defined by:

$$f_{L-1}(X) = \sum_{k=1}^{L-1} \beta_k g_k(\omega_k^\top X + b_k), \quad \text{where } L = 1, 2, \dots, \text{ and } f_0 = 0, \quad (3.16)$$

with β_k as the weight of the k th hidden layer node, g denoting an activation function, and ω_k and b_k representing the input weight and bias of the k th hidden layer node respectively, for $k = 1, 2, 3, \dots, L_{\max}$. To calculate the network's current residual error vector, we use:

$$e_{L-1}(X) = f - f_{L-1}(X) = [e_{L-1,1}(X), \dots, e_{L-1,m}(X)]^\top \in \mathbb{R}^{N \times m}. \quad (3.17)$$

If the norm $\|e_{L-1}\|^2$ has not yet reached the predefined error ξ or the maximum node count $L_{\max}$, the algorithm incorporates an additional L th hidden layer node. This is based on a supervised strategy to determine the input weight ω and bias b , expressed as:

$$h_L = [g_L(\omega_L^\top x^1 + b_L), g_L(\omega_L^\top x^2 + b_L), \dots, g_L(\omega_L^\top x^N + b_L)]^\top \in \mathbb{R}^{N \times m}, \quad (3.18)$$

and

$$\xi_{L,p} = \frac{(e_{L-1,p}^\top \cdot h_L)^2}{h_L^\top h_L} - (1 - r - \mu_L) \|e_{L-1,p}^\top\|^2, \quad (3.19)$$

where $p = 1, 2, 3, \dots, m$, h_L signifies the L th layer's output, ω_L and b_L are the candidate parameters for node L , $r \in (0, 1)$, and μ_L is a non-negative number satisfying $\mu_L \leq 1 - r$, with $\lim_{L \rightarrow \infty} \mu_L = 0$. The parameters for the L th node are then defined by maximizing $\xi_L = \sum_{p=1}^m \xi_{L,p} \geq 0$. Subsequently, the hidden layer's output weight β is determined by:

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 = H^+Y, \quad (3.20)$$

where H^+ denotes the Moore-Penrose generalized inverse of $H = [h_1, h_2, \dots, h_L]$. Thus, the network's output f is computed

3.3.6.2 SC-I Classifier

The main variants of the SCN focus on the strategy of generating the output weights β . In [215], they total introduced three different strategies of calculating the β . In a particular way, SC-I utilizes a constructive scheme to assess the output weights exclusively for the recently added hidden node, while maintaining the previously acquired output weights unchanged. SC-II performs a local least squares method with a shifting window size specified by the user to recalculate a portion of the current output weights. Lastly, SC-III determines the output weights in their entirety by solving a global least squares problem utilizing the current model. In addition, Wang et al., 2019 also proposed DeepSCN [218] for enriching learning features by adding multiple hidden layers and hidden nodes.

Suppose that $\text{span}(\mathcal{T})$ is dense in L_2 space and $\forall g \in \mathcal{T}, 0 < \|g\| < b_g$ for some $b_g \in \mathbb{R}^+$. Given $0 < r < 1$ and a non-negative real number sequence $\{\mu_L\}$ with $\lim_{L \rightarrow \infty} \mu_L = 0$ and $\mu_L < (1 - r)$. For $L = 1, 2, \dots$, stands for [215]

$$\delta_L = \sum_{q=1}^m \delta_{L,q}, \delta_{L,q} = (1 - r - \mu_L) \|e_{L-1,q}\|^2, \quad q = 1, 2, \dots, m. \quad (3.21)$$

If the random basis function g_L is generated to satisfy the following inequalities:

$$\langle e_{L-1,q}, g_L \rangle^2 \geq b_{g_0}^2 \delta_{L,q}, \quad q = 1, 2, \dots, m, \quad (3.22)$$

therefore, the output weights are constructively evaluated by

$$\beta_{L,q} = \frac{\langle e_{L-1,q}, g_L \rangle}{\|g_L\|^2}, \quad q = 1, 2, \dots, m. \quad (3.23)$$

Then, we have $\lim_{L \rightarrow \infty} \|f - f_L\| = 0$, where $f_L = \sum_{j=1}^L \beta_j g_j$, $\beta_j = [\beta_{j,1}, \dots, \beta_{j,m}]^T$.

3.3.6.3 SC-II Classifier

Suppose that $\text{span}(\mathcal{T})$ is dense in L_2 space and $\forall g \in \mathcal{T}, 0 < \|g\| < b_g$ for some $b_g \in \mathbb{R}^+$. Given $0 < r < 1$ and a nonnegative real number sequence $\{\mu_L\}$ with $\lim_{L \rightarrow \infty} \mu_L = 0$ and $\mu_L \leq (1 - r)$. For a given window size K and $L = 1, 2, \dots$, denoted by [215]

$$\delta_L^* = \sum_{q=1}^m \delta_{L,q}^* \delta_{L,q}^* = (1 - r - \mu_L) \|e_{L-1,q}^*\|^2, \quad q = 1, 2, \dots, m. \quad (3.24)$$

If the random basis function g_L is generated to satisfy the following inequalities:

$$\langle e_{L-1,q}^*, g_L \rangle^2 \geq b_{g_0}^2 \delta_{L,q}^*, \quad q = 1, 2, \dots, m, \quad (3.25)$$

and, as $L \leq K$, the output weights are evaluated by

$$[\beta_1^*, \beta_2^*, \dots, \beta_L^*] = \arg \min_{\beta} \left\| f - \sum_{j=1}^L \beta_j g_j \right\|. \quad (3.26)$$

Otherwise (i.e., $L > K$), the output weights will be selectively evaluated (i.e., keep $\beta_1^*, \dots, \beta_{L-K}^*$ unchanged and renew the left $\beta_{L-K+1}, \dots, \beta_L$) by

$$[\beta_{L-K+1}^*, \beta_{L-K+2}^*, \dots, \beta_L^*] = \arg \min_{\beta_{L-K+1}, \dots, \beta_L} \left\| f - \sum_{j=1}^{L-K} \beta_j^* g_j - \sum_{j=L-K+1}^L \beta_j g_j \right\|. \quad (3.27)$$

Then, we have $\lim_{L \rightarrow \infty} \|f - f_L^*\| = 0$, where $f_L^* = \sum_{j=1}^L \beta_j^* g_j$, $\beta_j^* = [\beta_{j,1}^*, \dots, \beta_{j,m}^*]^T$.

3.3.6.4 SC-III Classifier

Suppose that $\text{span}(\mathcal{T})$ is dense in L_2 space and $\forall g \in \mathcal{T}, 0 < \|g\| < b_g$ for some $b_g \in \mathbb{R}^+$. Given $0 < r < 1$ and a nonnegative real number sequence $\{\mu_L\}$ with $\lim_{L \rightarrow \infty} \mu_L = 0$ and $\mu_L \leq (1 - r)$. For $L = 1, 2, \dots$, stands for [215]

$$\delta_L^* = \sum_{q=1}^m \delta_{L,q}^* \delta_{L,q}^* = (1 - r - \mu_L) \|e_{L-1,q}\|^2, \quad q = 1, 2, \dots, m. \quad (3.28)$$

If the random basis function g_L is generated to satisfy the following inequalities:

$$\langle e_{L-1,q}, g_L \rangle^2 \geq b_{g_0}^2 \delta_{L,q}^*, \quad q = 1, 2, \dots, m, \quad (3.29)$$

Therefore, the output weights are constructively evaluated by

$$\beta_L^* = [\beta_{L,1}^*, \beta_{L,2}^*, \dots, \beta_{L,m}^*] = \arg \min_{\beta} \left\| f - \sum_{j=1}^L \beta_j g_j \right\|. \quad (3.30)$$

Then, we have $\lim_{L \rightarrow \infty} \|f - f_L^*\| = 0$, where $f_L^* = \sum_{j=1}^L \beta_j^* g_j$, $\beta_j^* = [\beta_{j,1}^*, \dots, \beta_{j,m}^*]^T$.

3.3.6.5 DeepSCN Classifier

Given a training dataset with inputs $X = \{x_1, x_2, \dots, x_N\}$; $x_i = [x_{i1}, \dots, x_{id}]^T \in \mathbb{R}^d$ and outputs $T = \{t_1, t_2, \dots, t_N\}$, where $t_i = [t_{i1}, \dots, t_{im}]^T \in \mathbb{R}^m$, $i = 1, \dots, N$. Denote $e_{L_{n-1}}^{(n)}(X) = [e_{L_{n-1},1}^{(n)}(X), \dots, e_{L_{n-1},q}^{(n)}(X)]^T$ as the corresponding residual error vector before the L_n -th new hidden node of the n -th hidden layer is added, where $e_{L_{n-1},q}^{(n)}(X) = [e_{L_{n-1},q}^{(n)}(x_1), \dots, e_{L_{n-1},q}^{(n)}(x_N)] \in \mathbb{R}^N$, $q = 1, \dots, m$. After adding the L_n -th hidden node in the n -th hidden layer [218], we have

$$h_{L_n}^{(n)} = h_{L_n}^{(n)}(X) = [\phi_{n,L_n}(x_1^{(n-1)}), \dots, \phi_{n,L_n}(x_N^{(n-1)})]^T, \quad (3.31)$$

where $\phi_{n,L_n}(x_i^{(n-1)})$ is used to simplify $\phi_{n,L_n}(x_{i1}^{(n-1)}, \dots, x_{id}^{(n-1)}, b_n^{(n-1)})$, and $x_{ij}^{(0)} = x_{ij} = x_i$; $i = \{1, \dots, N\}$, $j = \{1, \dots, d\}$, and $\Phi(\cdot)$ is $\Phi(x^{(n-2)}) = W^{(n-1)}\Phi(x^{(n-2)}) + B^{(n-1)}$ for $n \geq 2$.

Let $H_{L_n}^{(n)} = [h_1^{(n)}, h_2^{(n)}, \dots, h_{L_n}^{(n)}]$ represent the hidden layer output matrix and denote a temporary variable $\delta_{L_n,q}^{(n)} (q = 1, 2, \dots, m)$ where

$$\delta_{L_n,q}^{(n)} = \left(\frac{e_{L_{n-1},q}^{(n)} h_{L_n}^{(n)}}{\|h_{L_n}^{(n)}\|^2} \right)^2 = (1-r) \|e_{L_{n-1},q}^{(n)}\|^2, \quad (3.32)$$

Therefore, suppose that $\text{span}(\mathcal{T})$ is dense in L_2 space and $\forall \phi \in \mathcal{T}$, $0 < \|\phi\| < c$ for some $c \in \mathbb{R}^+$. Given $0 < r < 1$ and a non-negative decreasing sequence $\{\mu\}$ with $\lim_{\mu \rightarrow +\infty} \mu = 0$ and $\mu < (1-r)$. For $n = 1, 2, \dots$, and $j = 1, 2, \dots, L_n$, denoted by

$$\delta_{j,q}^{(n)} = (1-r-\mu_j) \|e_{j-1,q}^{(n)}\|^2, \quad q = 1, 2, \dots, m. \quad (3.33)$$

Stochastically configuring the j -th hidden node $\phi_{n,j}$ within the n -th hidden layer ($j = 1, 2, \dots, L_n$) to satisfy the following inequalities:

$$\langle e_{j,q}^{(n)}, \phi_{n,j} \rangle^2 \geq c^2 \delta_{j,q}^{(n)}, \quad q = 1, 2, \dots, m. \quad (3.34)$$

Fix the random basis functions $\phi_{n,1}, \phi_{n,2}, \dots, \phi_{n,L_n}$, and start to add the first hidden node $\phi_{n+1,1}$ in the $(n+1)$ -th hidden layer according to the following inequalities

$$\langle e_{L_n,q}^{(n)}, \phi_{n+1,1} \rangle^2 \geq c^2 \delta_{L_n,q}^{(n)}, \quad q = 1, 2, \dots, m. \quad (3.35)$$

Keep adding new hidden nodes within the $(n+1)$ -th hidden layer based on 3.34, followed by generating the first hidden node in a new hidden layer via 3.35. After adding one hidden node (either in the present hidden layer or starting a new hidden layer), the read-out weights are evaluated by the least squares method,

$$\beta^* = \arg \min_{\beta} \left\| F - \sum_{k=1}^n \sum_{j=1}^{L_k} \beta_j^{(k)} \phi_{k,j} \right\|^2. \quad (3.36)$$

Hence, we have $\lim_{L \rightarrow \infty} \|F - F_{S_n}^{(n)}\| = 0$, where

$$F_{S_n}^{(n)}(x) = \sum_{k=1}^n \sum_{j=1}^{L_k} \beta_j^{(k)} \phi_{k,j} \left(x^{(k-1)}; w_j^{(k-1)}, b_j^{(k-1)} \right). \quad (3.37)$$

The SCNs mentioned above offer the following benefits: 1) The infinite approximation of the conventional randomized neural network is contingent on the range of random parameters and the number of hidden layer nodes. Insufficiently configured parameters hinder the model's capacity to effectively converge toward the target function. The SCNs utilize an incremental modeling approach. Using the supervisory mechanism to set the input weights and biases of the hidden layer nodes and sticking to the set parameter range ensures a good overall approximation. 2) In contrast to the deep neural network model, it computes the output weight of hidden layer nodes using pseudo-inverse operations instead of iterative updates through back-propagation. This approach enhances the learning rate and effectively prevents issues such as local optima, gradient vanishing, or exploding. It is well-suited for application situations that have demanding real-time needs and can easily accommodate simultaneous and integrated activities. 3) The setting of SCN settings involves minimal human participation, and its random parameters are dependent on data supplied by the supervisory mechanism. If the model fails to achieve the predetermined error, it can dynamically augment the number of hidden layer nodes or grow into DeepSCN.

3.3.6.6 Random Variables - Uniform Distributions

Uniform distribution probability density function (PDF) is given by

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & \text{for } a \leq x \leq b, \\ 0, & \text{otherwise,} \end{cases} \quad (3.38)$$

where a and b represent the distribution's minimum and maximum limits, respectively. Continuous uniform distribution is a probability distribution in which all values within a given interval have the same probability of occurring. It has a constant PDF over the support interval $[a, b]$ and is frequently employed in simulations and statistical inference when there is no prior knowledge of the data distribution.

The uniform distribution is simple and easy to use, which is one of its advantages. Its PDF is straightforward to comprehend and compute, and it possesses a number of advantageous properties, such as being invariant under location and scale transformations. It is also frequently used as a reference point for comparing other distributions and testing statistical hypotheses.

3.3.6.7 Random Variables - Gaussian Distributions

Gaussian distribution PDF is given by

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad (3.39)$$

where μ is the mean (location parameter) and σ is the standard deviation (scale parameter).

In statistics and probability theory, the Gaussian distribution, also known as the normal distribution, is one of the most frequently employed probability distributions. It is frequently used to approximate other distributions and to model random variables that are the sum of many tiny, unrelated, and identically distributed components, such as measurement errors. In addition to being symmetric and unimodal, having a bell-shaped curve, and being entirely determined by its mean and standard deviation, the Gaussian distribution possesses several beneficial characteristics. Moreover, the wide applications of the Gaussian distribution and its use in numerous statistical methods, such as the Central Limit Theorem and Maximum Likelihood Estimation, is one of its advantages. It is also computationally feasible, facilitating its use in a variety of applications.

3.3.6.8 Random Variables - Lévy Distributions

Lévy distributions PDF is given by

$$f(x; c, \alpha) = \sqrt{\frac{c}{2\pi}} \frac{1}{(x-c)^{3/2}} e^{-\frac{\alpha}{2(x-c)}}, \quad (3.40)$$

where x is the random variable, c is the location parameter, and α is the scale parameter.

The Lévy distribution is an example of a heavy-tailed distribution, which has a greater likelihood of producing extreme values than other distributions with finite variance. A significant property of the Lévy distribution is that it is used to model systems with power-law behavior, such as earthquakes and financial markets. It also serves as the foundation for the Lévy flight, a mathematical model used to characterize the movement patterns of animals, particles, and other non-Gaussian entities.

Among the intriguing mathematical properties of the Lévy distribution are its closedness under convolution and its infinite variance. It is also related to the Cauchy distribution, which is a special case of the Lévy distribution when $\alpha = 1$ and the location parameter is c .

3.3.6.9 Random Variables - Cauchy Distributions

Cauchy distribution PDF is given by

$$f(x; x_0, \gamma) = \frac{1}{\pi} \frac{\gamma}{(x - x_0)^2 + \gamma^2}, \quad (3.41)$$

where x is the random variable, x_0 is the location parameter, and γ is the scale parameter.

The Cauchy distribution is an example of a heavy-tailed distribution, which has a greater likelihood of producing extreme values than other distributions with finite variance. The Cauchy distribution is invariant under translation and dilation, which means that if $X \sim \text{Cauchy}(x_0, \gamma)$, then the random variable $Y = aX + b$ follows a Cauchy distribution with location parameter $x_0 + b$ and scale parameter $|a|\gamma$.

3.3.6.10 Random Variables - Weibull Distributions

The PDF of the Weibull distributions is given by:

$$f(x) = \begin{cases} \frac{\alpha}{\lambda} \left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-(x/\lambda)^\alpha}, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad (3.42)$$

where x is the random variable, α is the shape parameter, and λ is the scale parameter. The Weibull distribution is often used to model failure rates in reliability engineering, which is also a heavy-tailed.

3.3.6.11 Fractional Gaussian Distributions

The fractional Gaussian distribution is a probability distribution associated with the fractional Brownian motion, a continuous-time stochastic process with long-range dependence. The fractional Gaussian distribution is a one-dimensional distribution that characterizes the statistical properties of fractional Brownian motion increments. The PDF formula for the fractional Gaussian distribution is as follows:

$$f_H(t) = \frac{1}{\Gamma(H + 1/2)} \int_t^0 (t - s)^{H-0.5} df(s), \quad (3.43)$$

where Γ is the Gamma function, and $df(s)$ is the general definition of white noise. H is the Hurst exponent and it's a real number (value internal is (0,1)). It describes the raggedness of the resultant motion, with a higher value leading to a smoother motion [220]. Several properties of the H value need to be discussed here, if $H = 0.5$, then it is a standard Brownian motion or Wiener process; if $0.5 < H < 1$, then increments of the process are positively correlated; and if $0 < H < 0.5$, then increments of the process are negatively correlated [221]. The H index variations and comparison with the Gaussian distribution are shown in Figure 3.4.

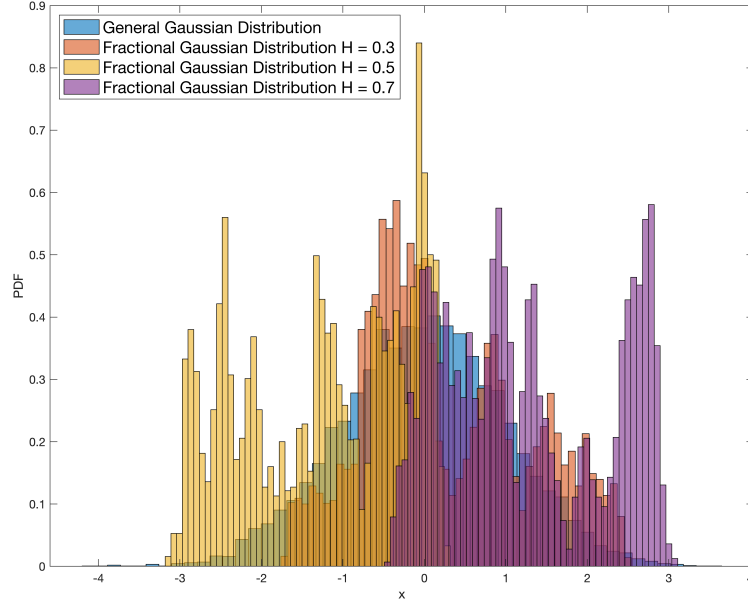


Figure 3.4: Fractional order Gaussian distribution (in different H) vs. standard Gaussian distribution.

3.3.6.12 Fractional Weibull Distributions

The PDF of the fractional Weibull distribution can be expressed mathematically as [222]:

$$f(x; \alpha, \beta, \rho, \xi) = \frac{\rho}{\beta} (1 - \xi) \left(\frac{x - \alpha}{\beta} \right)^{\rho-1} e^{-\left(\frac{x - \alpha}{\beta} \right)^\rho}, x \geq \alpha, \quad (3.44)$$

where the shape parameter is ρ and scale parameter is β , ξ is the fractal parameter, and α is the location parameter. The scaling factor $0 \leq (1 - \xi) \leq 1$ represents the reserved proportion of the data. If δ goes to 0, then fractional Weibull distributions degenerate into Weibull distributions. Changing ξ and keeping the other parameters the same with Weibull distribution would move the peak vertically (See Figure 3.5 with different ξ). The performance of different variants of the SCNs is shown in Figure 3.5.

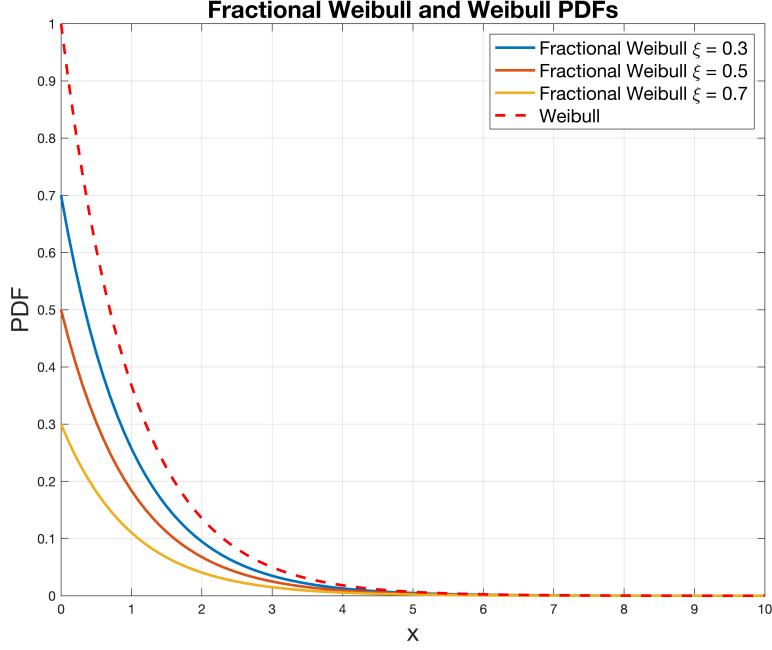


Figure 3.5: Fractional order Weibull distribution (in different ξ) vs. regular Weibull distribution.

3.4 Experiments and Evaluation

3.4.1 Experimental Setup

Since we already have Walabot and IMAGEVK-74 as microwave and millimeter wave transmitters and receivers, we need a processor unit to process and store the data. The portable sensing device, which consists of a Raspberry Pi 4B with 2 GB of RAM, a Walabot developer module, and millimeter wave signal transmitters and receivers, is lightly placed on top of the sample, which is a soil and biochar mixture, as shown in Figure 3.6. Overall, we proposed five classification experiments, which are: 1) different soil carbon amounts, 2) different types of soil carbon, 3) different soil moisture levels, 4) soil moisture under soil carbon effects, and 5) different containers (background) for soil carbon classification.

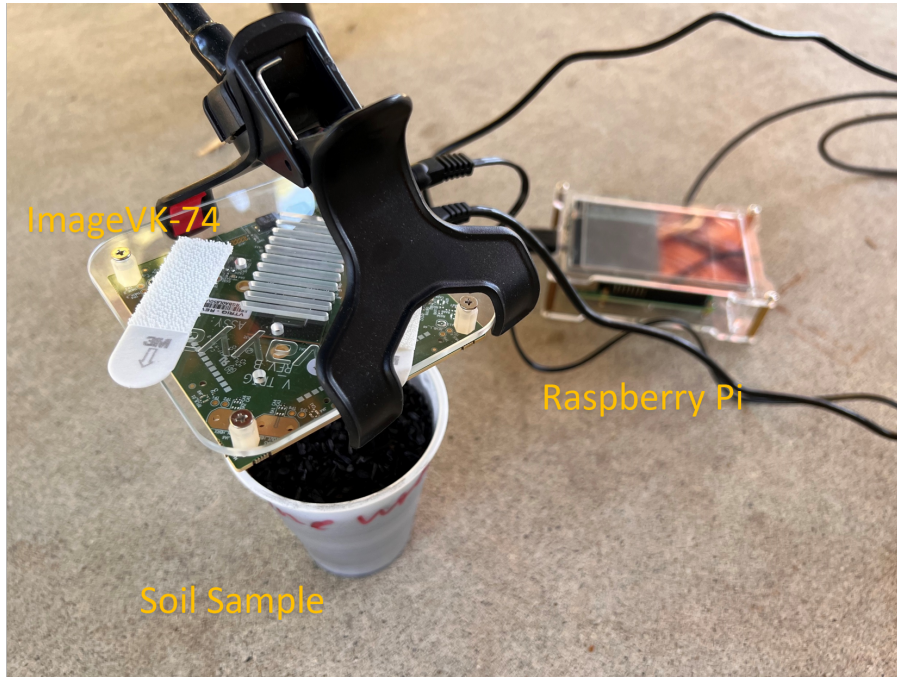


Figure 3.6: Sampling experiments hardware setup.

The sampling frequency of the microwave sensor is 10.24 GHz. We sampled each sample five times to reduce random errors. The sampling frequency of the millimeter wave sensor is from 62 GHz to 69 GHz. We used the same five sampling times for the millimeter wave sensor to reduce the random error. We set up 150 sampling points in the frequency band we mentioned, and the RBW is 100 kHz for reasonable sweep time and lower noise. Moreover, the sample was placed under the direct millimeter wave radar sensor by less than 2 cm.

The soil was collected from almond tree fields and filtered for coarse-grained substances such as stone. The biochar is made from 100% of orchard removals, walnut shells, 100% wood, almond shells, and pinewood (labeled as 1, 2, 3, 4, and 5, respectively). In the following experiments, the soil was uniformly distributed at 50g (g stands for gram), and biochar was placed at 2g (20% of the soil). Both of them are well mixed. As for the millimeter wave experiment setup, we had the new biochar and soil mixture rate in the following experiments, which varied from 0% to 100%.

Our experiments utilize PCA to reduce the dimensions of the dataset so that we can reduce the subsequent classifier computation resources. To evaluate the performance fairly, we introduced a confusion matrix to calculate average classification accuracy. The 5-fold cross-validation is also employed to validate all models' performance on our soil carbon samples. The accuracy and confusion matrix are presented to evaluate the classification's performance.

3.4.2 Conventional Classification Results

We conducted real-world experiments using microwave and millimeter waves to quantify the soil carbon content and related experiments.

3.4.2.1 Different Soil Carbon Amount Recognition

We selected the carbon and soil ratio from 0% to 100% including the air, which makes up a total of 7 classes for millimeter wave sensing. With microwave sensing, we have a carbon and soil ratio of 20% to 100%. Then we sampled 20 times for each class. The millimeter wave datasets have been split into training datasets and test datasets in a 5:1 ratio. Therefore, the total dataset contains 140 samples. Each microwave signal was sampled 10 times for each class, and the datasets have the same split training and testing ratio of 5:1 for classification. Each microwave signal of each feature has 8192 sampling points.

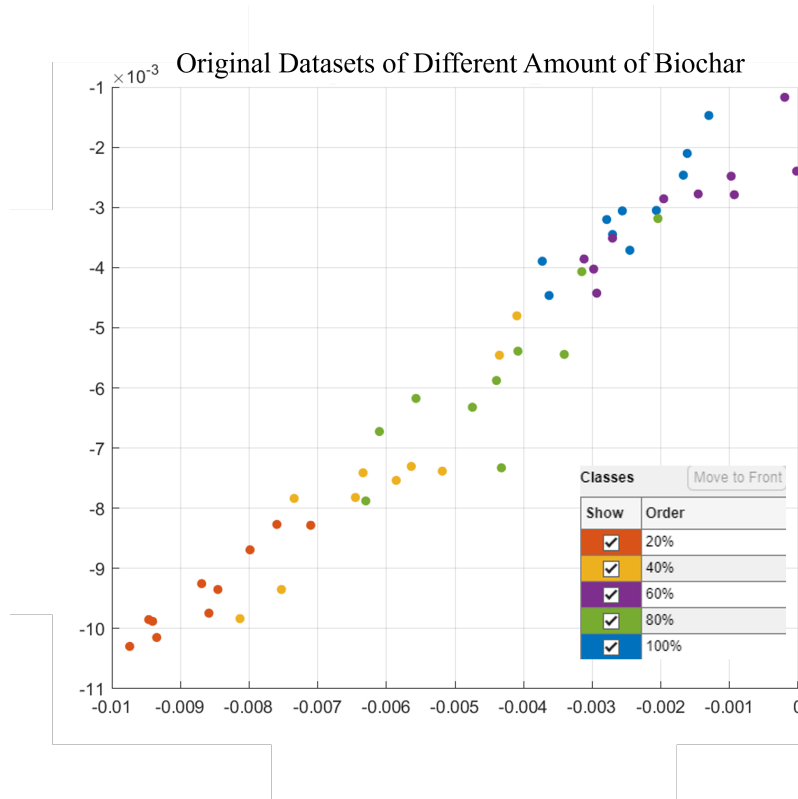


Figure 3.7: The scatter plot of microwave signals for different percentage of carbon soil ratio.

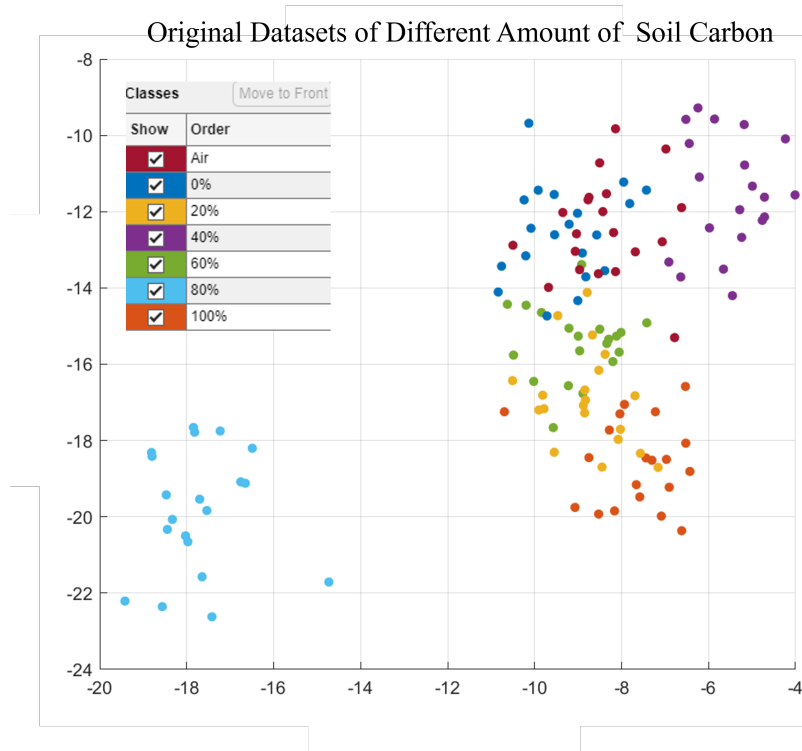


Figure 3.8: The scatter plot of millimeter wave signals for different percentage of carbon soil ratio.

Classifiers	Microwave Accuracy	Millimeter wave Accuracy
Fine Tree	74.0%	92.1%
Coarse Tree	74.0%	69.3%
LDA	92.0%	100.0%
Naive Bayes	98.0%	100.0%
Linear SVM	98.0%	100.0%
Quadratic SVM	94.0%	100.0%
Cubic SVM	94.0%	100.0%
Gaussian SVM	96.0%	100.0%
Fine KNN	90.0%	100.0%
Coarse KNN	20.0%	21.4%
DNN	88.0%	93.6%

Table 3.3: Classification accuracy of different percentage of carbon soil ratio.

Optimal Classifiers	Microwave	Optimization Configuration
Optimized DNN	94.0%	Number of fully connected layers: 1 Activation: Tanh First layer size: 186
Optimized KNN	96.0%	Number of neighbors: 5 Distance metric: Euclidean Distance weight: Inverse
Optimized Tree	88.0%	Maximum number of splits: 13
Optimized Naive Bayes	98.0%	Distribution: Gaussian
Optimized Discriminant	98.0%	Discriminant type: Diagonal Quadratic
Optimized SVM	98.0%	Kernel function: Linear

Table 3.4: The performance of optimized classifiers comparison for microwave signals of different percentage of carbon soil ratio.

Optimal Classifiers	Millimeter wave	Optimization Configuration
Optimized DNN	100.0%	Number of fully connected layers: 1 Activation: Sigmoid First layer size: 66
Optimized KNN	100.0%	Number of neighbors: 11 Distance metric: Euclidean Distance weight: Squared Inverse
Optimized Tree	96.4%	Maximum number of splits: 20
Optimized Discriminant	100.0%	Discriminant type: Diagonal Linear
Optimized SVM	100.0%	Kernel function: Gaussian

Table 3.5: The performance of optimized classifiers comparison for millimeter wave signals of different percentage of carbon soil ratio.

The scatter plot of the raw data of microwave and millimeter waves are shown in Figures 3.7 and 3.8, respectively. Our sensors captured raw data which are the electromagnetic wave reflectance of the sample. In classification learner app, the app applied a dimensionless reduction technique such as PCA to project the data onto a 2D plane. The axes would represent the first two principal components, which capture the most variance in the data. We found that millimeter waves are much more sensitive to the different amounts of biochar mixed with soil. After we applied the eight most common classifiers, the accuracy shown in Tables 3.3, 3.4 and 3.5 is for the optimized classifiers comparison. The highest accuracy of sensing the amount of soil carbon using millimeter waves is 100% whereas the highest sensing accuracy using the microwave is 98% by using linear SVM and naive Bayes. Furthermore, from the scatter plots, the classes 60%, and 100% overlap, which causes a relatively higher

failure rate. However, compared with microwave sensing, using millimeter waves to categorize the amount of soil carbon is much high using optimized classifiers.

3.4.2.2 Different Types of Soil Carbon Classification

We used total five different types of soil carbon (biochar), as shown in Figure 3.10. This is because different biochars have different microporous structures that could cause different water retention abilities, nutrient storage, etc. In Figures 3.9 and 3.11, the scatter plot demonstrates the distribution of different types of biochar using microwave and millimeter wave sensing, respectively. Compared to the previous different amounts of soil carbon classification, we can see that the variance of each class's data is small than the different percentage of carbon soil ratio, which means this would make the classifier easier to classify these biochars. Moreover, from Table 3.6, the highest classification accuracy is 100% for both microwave and millimeter wave. The fine decision tree's accuracy may change or fluctuate, but there is still around 90% sensing accuracy for the classification of different types of biochar. The optimized classification results from Tables 3.7 and 3.8 also revealed that millimeter wave results are better than the microwave.

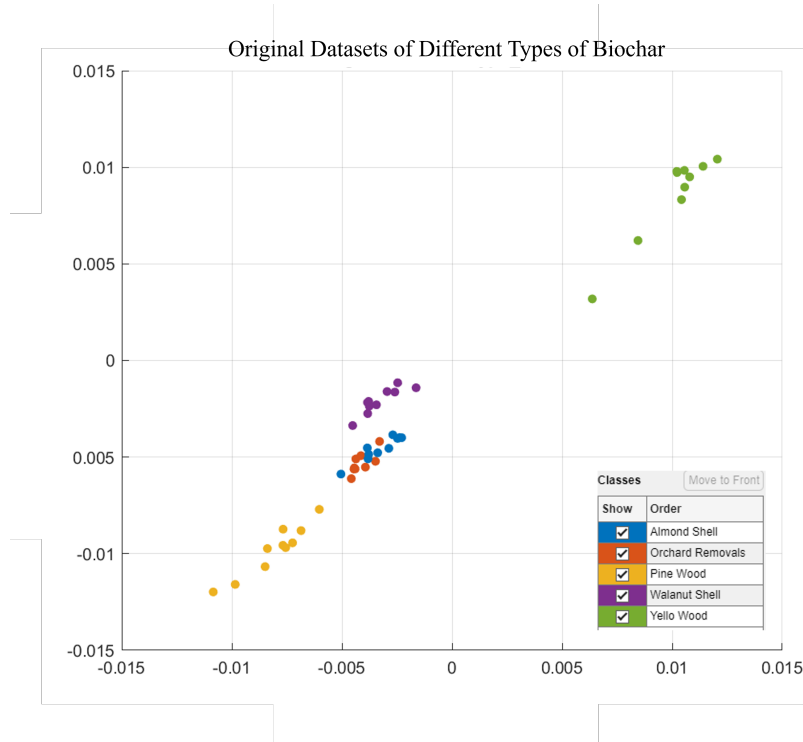


Figure 3.9: The scatter plot of microwave signals for different types of biochar (soil carbon).



Figure 3.10: Five different types of biochar (soil carbon).

Classifiers	Microwave Accuracy	Millimeter wave Accuracy
Fine Tree	94.0%	95.8%
Coarse Tree	94.0%	80.0%
LDA	100.0%	100.0%
Naive Bayes	100.0%	100.0%
Linear SVM	100.0%	100.0%
Quadratic SVM	100.0%	100.0%
Cubic SVM	100.0%	100.0%
Gaussian SVM	100.0%	100.0%
Fine KNN	100.0%	100.0%
Coarse KNN	20.0%	16.7%
DNN	100.0%	100.0%

Table 3.6: Classification accuracy of different types of biochar (soil carbon).

Optimal Classifiers	Microwave	Optimization Configuration
Optimized DNN	98.0%	Number of fully connected layers: 3 Activation: Tanh First layer size: 2 Second layer size: 6 Third layer size: 9
Optimized KNN	100.0%	Number of neighbors: 25 Distance metric: Cosine Distance weight: Squared Inverse
Optimized Tree	94.0%	Maximum number of splits: 6
Optimized Discriminant	100.0%	Discriminant type: Diagonal Quadratic
Optimized SVM	100.0%	Kernel function: Linear

Table 3.7: The performance comparison of optimized classifiers for microwave signals of different types of biochar.

Optimal Classifiers	Millimeter wave	Optimization Configuration
Optimized DNN	100.0%	Number of fully connected layers: 1 Activation: ReLU First layer size: 77
Optimized KNN	100.0%	Number of neighbors: 15 Distance metric: Correlation Distance weight: Inverse
Optimized Tree	95.8%	Maximum number of splits: 11
Optimized Discriminant	100.0%	Discriminant type: Diagonal Quadratic
Optimized SVM	100.0%	Kernel function: Linear

Table 3.8: The performance comparison of optimized classifiers for millimeter wave signals of different types of biochar.

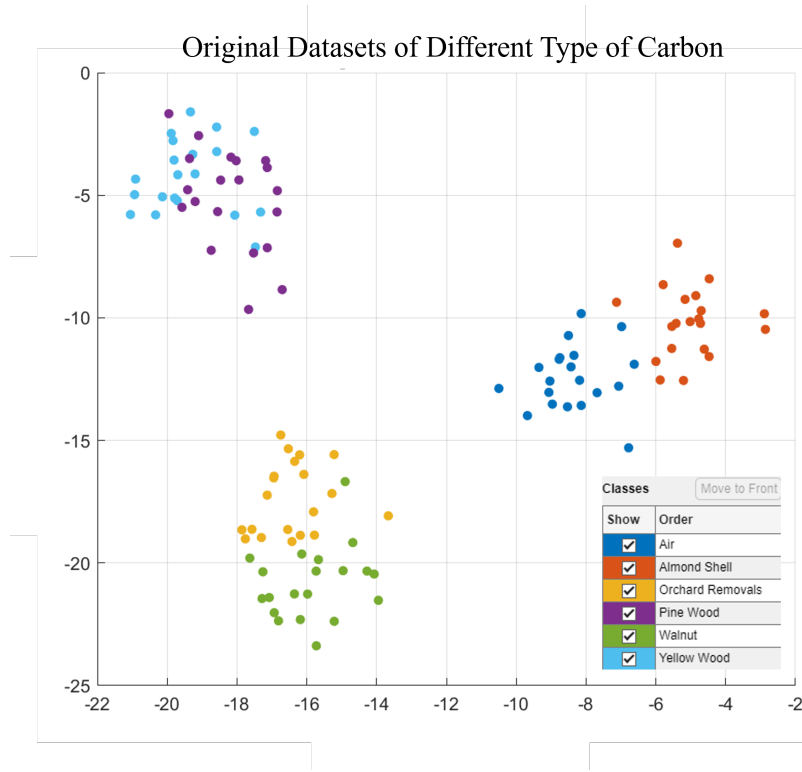


Figure 3.11: The scatter plot of millimeter wave signals for different types of carbon.

3.4.2.3 Soil Moisture Study and Comparison

Precision agriculture requires reliable soil moisture monitoring, which has a considerable impact on crop evapotranspiration, water exchange, and energy fluxes. Numerous remote sensing or proximal sensing techniques, such as thermal, optical, and microwave measurements, can be used to measure or infer soil moisture levels. Low spatial resolution, limited surface penetration, and vegetation are nonetheless limiting factors for the actual use of these techniques. By using a contactless method to sense soil moisture, so far, [132] used a microwave sensor to sense the soil moisture has been studied. However, there is no work to use a millimeter wave sensor to sense the soil moisture to our best knowledge. Furthermore, [132] did not compare the performance of using different classifiers. We put different amounts of water on the soil until it is saturated (labeled as Mos_100%).

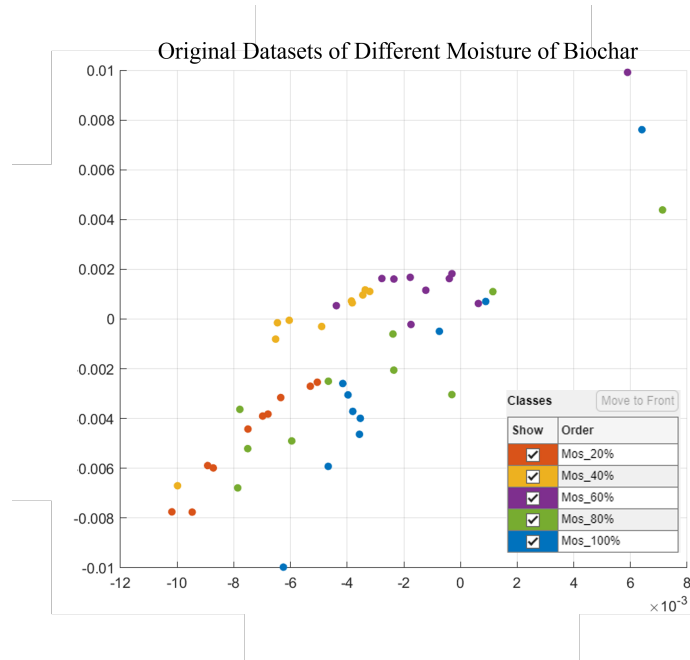


Figure 3.12: The scatter plot of microwave signals for different levels of soil moisture.

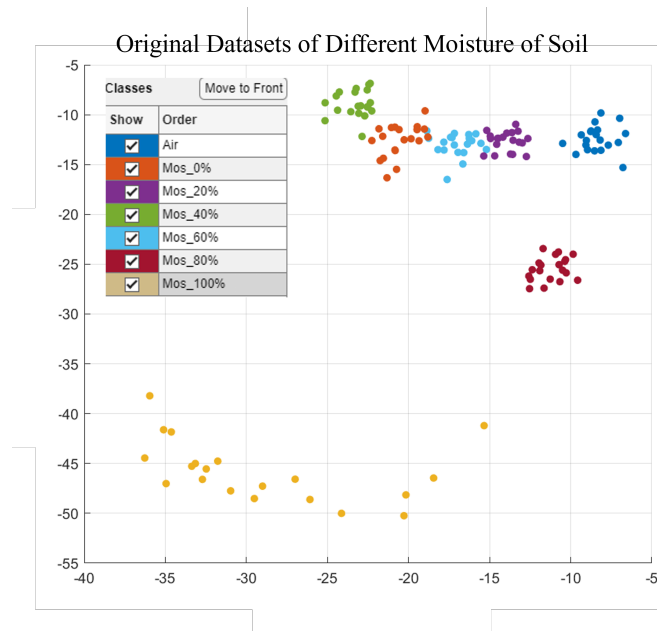


Figure 3.13: The scatter plot of millimeter wave signals for different levels of soil moisture.

Classifiers	Microwave Accuracy	Millimeter wave Accuracy
Fine Tree	56.0%	95.7%
Coarse Tree	56.0%	70.0%
LDA	90.0%	100.0%
Naive Bayes	72.0%	100.0%
Linear SVM	78.0%	98.6%
Quadratic SVM	76.0%	99.3%
Cubic SVM	72.0%	99.3%
Gaussian SVM	70.0%	100.0%
Fine KNN	78.0%	97.1%
DNN	68.0%	96.4%

Table 3.9: Classification accuracy of different moisture levels in the soil.

Optimal Classifiers	Microwave	Optimization Configuration
Optimized DNN	70.0%	Number of fully connected layers: 3 Activation: Tanh First layer size: 210 Second layer size: 68 Third layer size: 8
Optimized KNN	76.0%	Number of neighbors: 1 Distance metric: Correlation Distance weight: Squared Inverse
Optimized Tree	62.0%	Maximum number of splits: 5
Optimized Discriminant	82.0%	Discriminant type: Linear
Optimized SVM	84.0%	Kernel function: Gaussian

Table 3.10: The performance of optimized classifiers comparison for microwave signals of different moisture levels in the soil.

Optimal Classifiers	Millimeter wave	Optimization Configuration
Optimized DNN	100.0%	Number of fully connected layers: 1 Activation: Tanh First layer size: 95
Optimized KNN	100.0%	Number of neighbors: 57 Distance metric: City Block Distance weight: Squared Inverse
Optimized Tree	98.6%	Maximum number of splits: 19
Optimized Discriminant	100.0%	Discriminant type: Linear
Optimized SVM	100.0%	Kernel function: Linear

Table 3.11: The performance of optimized classifiers comparison for millimeter wave signals of different moisture levels in the soil.

Figures 3.12 and 3.13 both show the classification sensing results of the soil moisture in scatter plots. As we know, water can absorb millimeter waves and microwaves. On the other hand, soil and water together can coagulate into large pieces, which is good for forming their unique features. The scatter plot of millimeter wave shows that features in each class are close to each other, whereas Figure 3.12 (microwave) demonstrates that features of all moisture levels overlap with each other. Tables 3.9, 3.10, and 3.11 also revealed that microwave sensing accuracy is lower than millimeter wave sensing accuracy in general.

3.4.2.4 Soil Moisture Influenced by Carbon Content Added

As we mentioned earlier, based on the physical principles behind the millimeter wave sensing and biochar properties, we validated the biochar with the soil water retention capacity by comparing it with the same amount of soil without biochar. The experiment shows that the same amount of soil with 50 g of biochar can absorb up to 110% water (well saturated). Due to biochar's porous structures that can absorb soil and water, the reflectance signals are pretty stable and close to each other, as shown in Figure 3.14 (microwave) and 3.15 (millimeter waves).

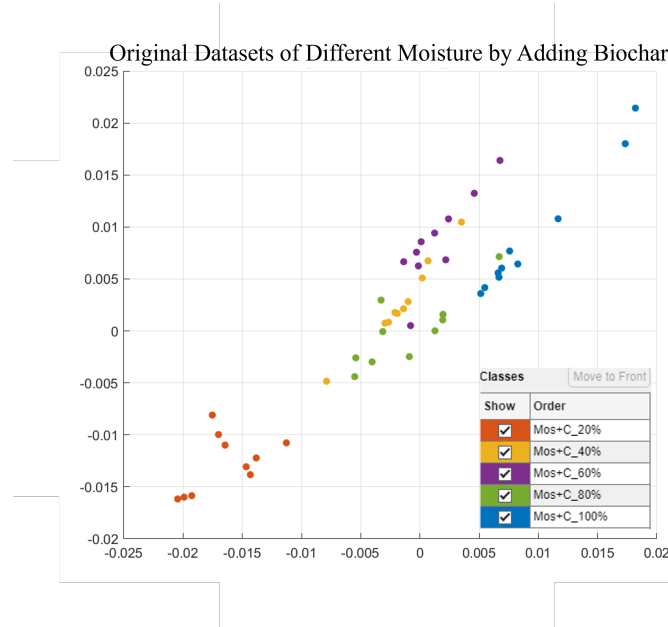


Figure 3.14: The scatter plot of microwave signals for different levels of soil moisture with soil carbon effects.

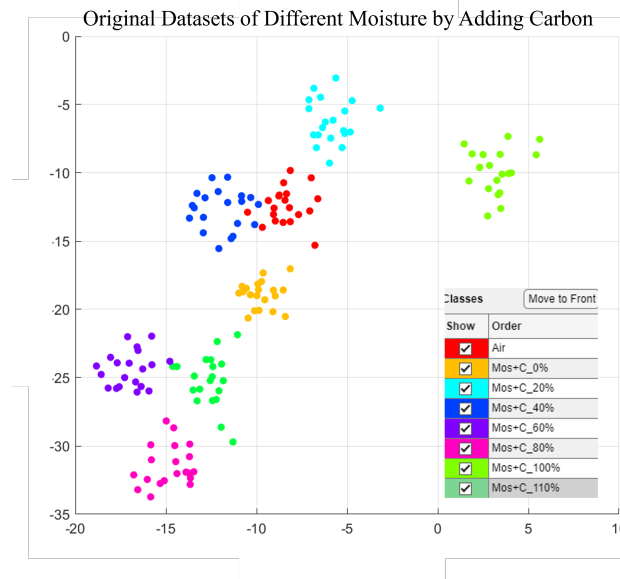


Figure 3.15: The scatter plot of millimeter wave signals for different levels of soil moisture with soil carbon effects.

Classifiers	Microwave Accuracy	Millimeter wave Accuracy
Fine Tree	74.0%	97.5%
Coarse Tree	74.0%	61.3%
LDA	88.0%	100.0%
Naive Bayes	86.0%	100.0%
Linear SVM	86.0%	100.0%
Quadratic SVM	76.0%	100.0%
Cubic SVM	72.0%	100.0%
Gaussian SVM	84.0%	100.0%
Fine KNN	86.0%	100.0%
DNN	74.0%	95.0%

Table 3.12: Classification accuracy of different moisture levels in soil under carbon effects.

Optimal Classifiers	Microwave	Optimization Configuration
Optimized DNN	78.0%	Number of fully connected layers: 2 Activation: ReLU First layer size: 101 Second layer size: 142
Optimized KNN	86.0%	Number of neighbors: 3 Distance metric: Cubic Distance weight: Inverse
Optimized Tree	76.0%	Maximum number of splits: 8
Optimized Discriminant	94.0%	Discriminant type: Linear
Optimized SVM	88.0%	Kernel function: Linear

Table 3.13: The performance of optimized classifiers comparison for microwave signals of different levels of soil moisture with soil carbon effects.

Optimal Classifiers	Millimeter wave	Optimization Configuration
Optimized DNN	100.0%	Number of fully connected layers: 1 Activation: Tanh First layer size: 23
Optimized KNN	100.0%	Number of neighbors: 80 Distance metric: Cosine Distance weight: Squared Inverse
Optimized Tree	97.5%	Maximum number of splits: 38
Optimized Discriminant	100.0%	Discriminant type: Diagonal Quadratic
Optimized SVM	100.0%	Kernel function: Linear

Table 3.14: The performance of optimized classifiers comparison for millimeter wave signals of different levels of soil moisture with soil carbon effects.

Tables 3.12, 3.13, and 3.14 illustrated the classification accuracy of the soil moisture under the microwave and millimeter wave sensing, respectively. Compared to Table 3.12, adding biochar would benefit the sensor in capturing the soil moisture level. Moreover, millimeter wave sensing accuracy is still higher than microwave sensing accuracy in sensing soil moisture under carbon effects and other related biochar applications.

3.4.2.5 Sensing Robustness for Soil Carbon Content Percentage

In this section, we investigated whether different sample containers will affect the soil carbon content sensing accuracy and which one has better performance with the microwave and millimeter wave. As we can see in Figure 3.16, we used different containers and put the same level of soil carbon content into these containers. We labeled these five containers as having the same soil carbon content level. After finishing the sampling procedure, we empty these containers, put the next level of soil carbon content into them, and do the sampling again. We used the same

settings for microwave and millimeter wave sensing for this investigation, and the total features for classification of the microwave and millimeter wave are 8192 and 1024, respectively. The microwave and millimeter wave scatter plots are shown in Figures 3.17 and 3.18, respectively.



Figure 3.16: Different containers with the same level of soil carbon content.

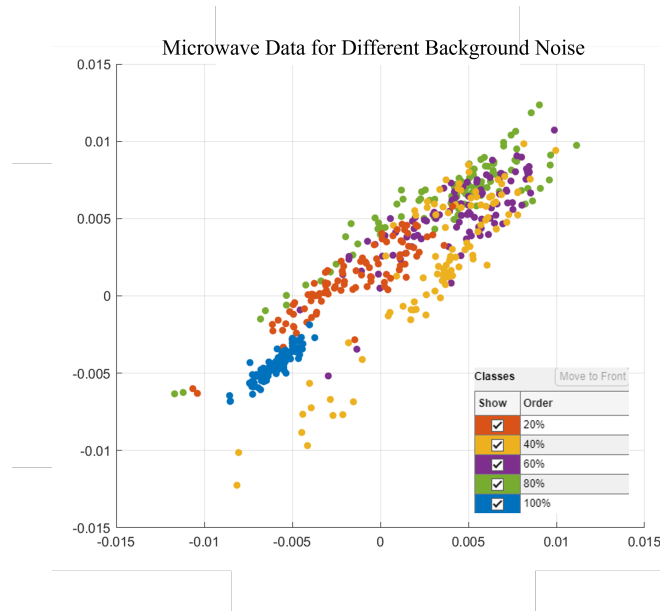


Figure 3.17: The scatter plot of microwave signals for different container backgrounds.

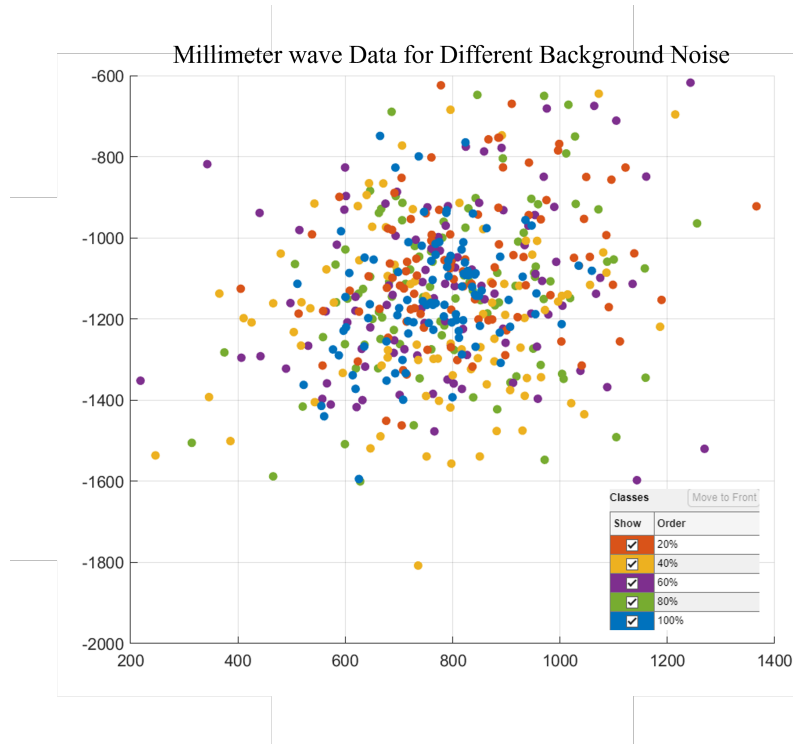


Figure 3.18: The scatter plot of millimeter wave signals for different container backgrounds.

From Figure 3.17 (microwave), we obtained the different soil carbon content levels distinguished from each other, and Table 3.15 shows the microwave sensing accuracy achieves at least 62%. However, Figure 3.18 (millimeter wave) shows that data points (sample points) of different soil carbon levels are sort of overlap with each other, which is difficult to classify. Table 3.15 also reveals that millimeter wave sensing accuracy is lower than microwave in general. This is supposed to be due to the low anti-disturbance ability of the millimeter wave.

Classifiers	Microwave Accuracy	Millimeter wave Accuracy
Fine Tree	66.4%	39.0%
Coarse Tree	62.0%	34.3%
LDA	83.6%	51.0%
Naive Bayes	67.2%	46.8%
Linear SVM	82.4%	67.0%
Quadratic SVM	86.4%	76.2%
Cubic SVM	85.4%	73.8%
Gaussian SVM	82.4%	67.6%
Fine KNN	80.8%	56.6%
DNN	77.0%	65.0%

Table 3.15: Classification accuracy of robustness investigation of sensing amount of soil carbon.

Optimal Classifiers	Microwave	Optimization Configuration
Optimized DNN	87.6%	Number of fully connected layers: 3 Activation: Tanh First layer size: 2 Second layer size: 6 Third layer size: 9
Optimized KNN	81.0%	Number of neighbors: 6 Distance metric: Euclidean Distance weight: Squared Inverse
Optimized Tree	71.4%	Maximum number of splits: 78
Optimized Discriminant	85.4%	Discriminant type: Linear
Optimized SVM	87.6%	Kernel function: Linear

Table 3.16: The performance of optimized classifiers comparison for microwave signals of sensing the amount of soil carbon (multi backgrounds).

Optimal Classifiers	Millimeter wave	Optimization Configuration
Optimized DNN	71.0%	Number of fully connected layers: 2 Activation: Tanh First layer size: 203 Second layer size: 267
Optimized KNN	61.2%	Number of neighbors: 162 Distance metric: Correlation Distance weight: Squared Inverse
Optimized Tree	40.2%	Maximum number of splits: 92
Optimized Discriminant	51.0%	Discriminant type: Linear
Optimized SVM	71.1%	Kernel function: Gaussian

Table 3.17: The performance of optimized classifiers comparison for millimeter wave signals of sensing the amount of soil carbon (multi backgrounds).

We validated our different background experiments by applying the most common classifiers we mentioned earlier. However, both microwave and millimeter wave sensing accuracy is lower than usual. We also applied optimal classifiers that let them optimize their hyperparameters during validation.

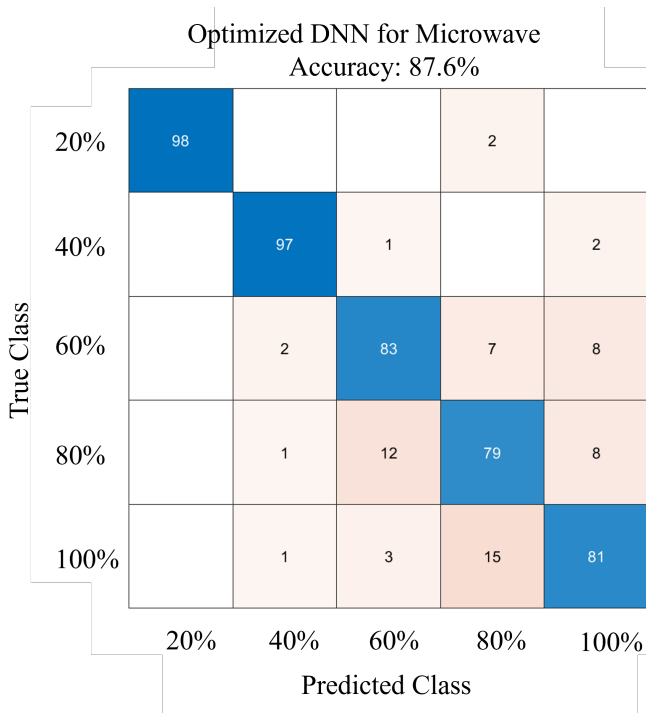


Figure 3.19: Confusion matrix of microwave signal classification using optimized DNN.

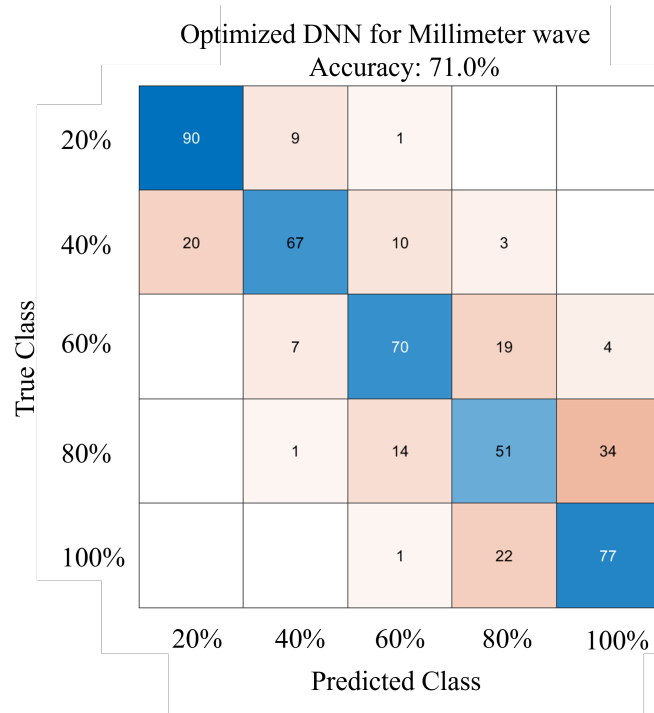


Figure 3.20: Confusion matrix of millimeter wave signal classification using optimized DNN.

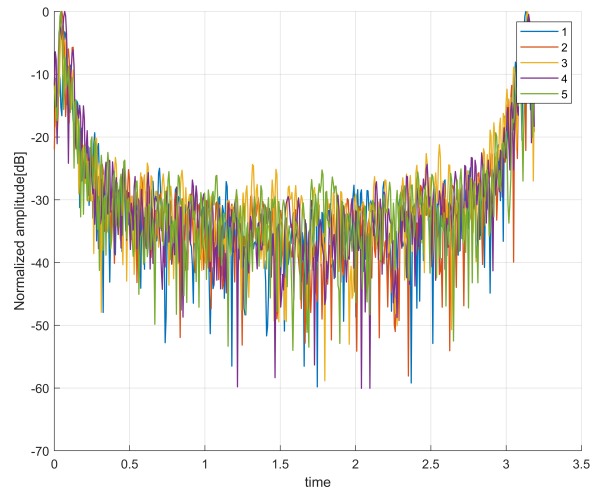


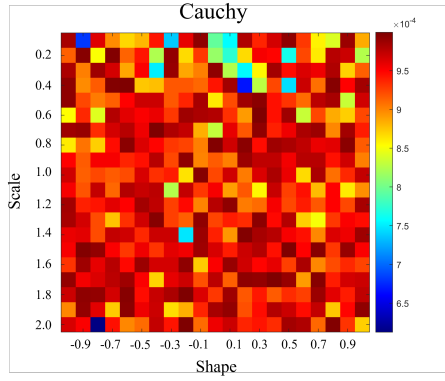
Figure 3.21: Millimeter wave reflectance signal analysis of five different backgrounds comparison using PDP.

Figure 3.19 shows the confusion matrix of the microwave sensing accuracy using optimized DNN is 87.6%, so does the optimized SVM, whereas compared to Figure 3.20, the millimeter wave sensing accuracy of optimized DNN is only 71.0%, and optimized SVM is 71.1% which are shown in Tables 3.16 and 3.17. Moreover, the sensing accuracy of the millimeter wave is worse than that of the microwave under different background noises. The reason for this phenomenon is that we believe that the microwave signal has a stronger anti-disturbance than the millimeter wave, which means the reflectance of the millimeter wave signal contains the background noise that is contained in the different soil carbon content signals. Figure 3.21 also shows the PDP is hard to distinguish from each other (different backgrounds and different soil carbon levels).

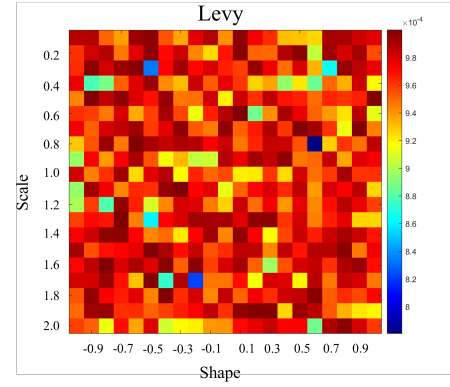
3.4.3 Optimal Distributions Parameters Tuning for SCNs Regression Model

As we mentioned in Section 3.3.6, we introduced 7 different random distributions for selecting weights and bias. Since different parameters of the distributions would have different behaviors of the generated random variables. Therefore, we permute all parameters to train SCNs by comparing RMSE value. RMSE is an important indicator of the regression model, it measures the average deviation between values predicted by a model and actual values.

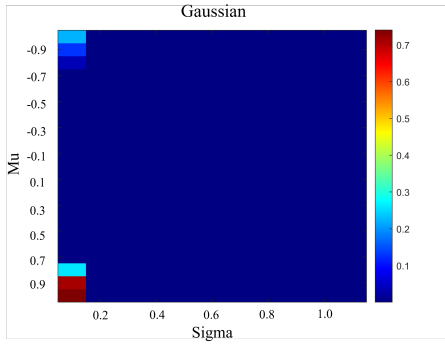
Based on our experiment results as shown in Figures 3.22 and 3.23: The optimal parameter of SCN-Uniform (also default setting) for soil carbon content data is 8 (for lower boundary would be -8, the upper boundary is 8). The optimal $\mu = 1$ and $\sigma = 0.2$ for SCN-Gaussian is shown in Figure 3.22c. For the heavy tail distributions, the optimal parameters of SCN-Lévy that scale parameter is 0.8 and shape (location) parameter is 0.6. The optimal parameters of SCN-Cauchy are 2 (scale parameter) and -0.8 (for location parameter), SCN-Weibull occupied optimal scale parameter is 1, and the optimal shape parameter is 1.5. As of the fractional Gaussian and fractional Weibull, we used the optimal parameters from general Gaussian and Weibull distributions as the baseline, and obtained the optimal Hurst value of fractional Gaussian (fGn) is 0.6, and fractal of the fractional Weibull (fWeibull) is 0.3. Please see details in sub-figures of Figures 3.22 and 3.23. As for the generated hidden layers, they are all around 450 ± 50 . Moreover, in this paper, we mainly considered the better accuracy of the SCNs with different distributions, the complexity of the network is not the main concern.



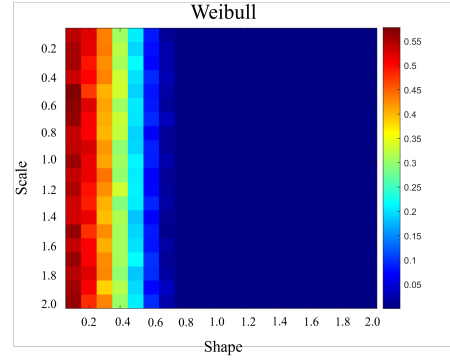
(a) Cauchy distributions of RMSE performance are generated by each scale and shape permutation.



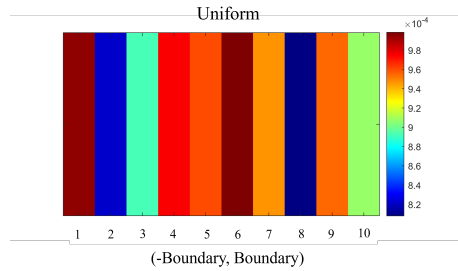
(b) Lévy distributions of RMSE performance generated by each scale and shape permutation.



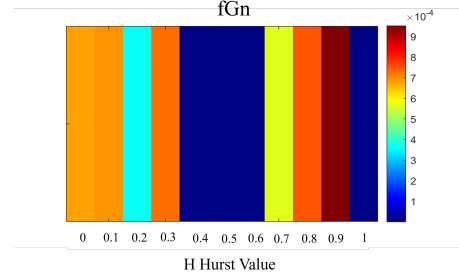
(c) Gaussian distributions of RMSE performance generated by each μ and σ permutation.



(d) Gaussian distributions of RMSE performance generated by each scale and shape permutation.

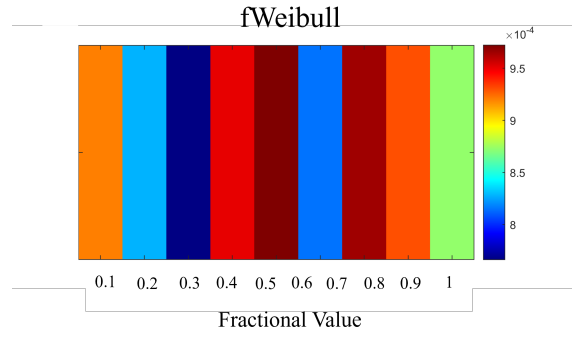


(e) Uniform distributions of RMSE performance generated by (-boundary, boundary)



(f) Fractional Gaussian distributions of RMSE performance generated by H fractal.

Figure 3.22: Optimal parameters tuning of a total of seven random distributions RMSE performance for SCN regression models.



(a) Fractional Weibull distributions of RMSE performance generated by different ξ values.

Figure 3.23: Optimal parameters tuning of a total of seven random distributions RMSE performance for SCN regression models.

Training Confusion Matrix							
Output Class	1	71 17.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	76 19.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	76 19.0%	0 0.0%	0 0.0%	100% 0.0%
	4	6 1.5%	3 0.8%	2 0.5%	77 19.2%	2 0.5%	85.6% 14.4%
	5	3 0.8%	1 0.2%	2 0.5%	3 0.8%	78 19.5%	89.7% 10.3%
		88.8% 11.3%	95.0% 5.0%	95.0% 5.0%	96.2% 3.7%	97.5% 2.5%	94.5% 5.5%
		Target Class					
		1	2	3	4	5	

(a) Gaussian training.

Training Confusion Matrix							
Output Class	1	76 19.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	79 19.8%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	80 20.0%	0 0.0%	0 0.0%	100% 0.0%
	4	2 0.5%	1 0.2%	0 0.0%	80 20.0%	0 0.0%	96.4% 3.6%
	5	2 0.5%	0 0.0%	0 0.0%	0 0.0%	80 20.0%	97.6% 2.4%
		95.0% 5.0%	98.8% 1.2%	100% 0.0%	100% 0.0%	100% 0.0%	98.8% 1.2%
		Target Class					
		1	2	3	4	5	

(b) Fractional Gaussian training.

Training Confusion Matrix							
Output Class	1	78 19.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	80 20.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	80 20.0%	0 0.0%	0 0.0%	100% 0.0%
	4	1 0.2%	0 0.0%	0 0.0%	80 20.0%	0 0.0%	98.8% 1.2%
	5	1 0.2%	0 0.0%	0 0.0%	0 0.0%	80 20.0%	98.8% 1.2%
		97.5% 2.5%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	99.5% 0.5%
		Target Class					
		1	2	3	4	5	

(c) Weibull training.

Training Confusion Matrix							
Output Class	1	79 19.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	80 20.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	80 20.0%	0 0.0%	0 0.0%	100% 0.0%
	4	0 0.0%	0 0.0%	0 0.0%	80 20.0%	0 0.0%	100% 0.0%
	5	1 0.2%	0 0.0%	0 0.0%	0 0.0%	80 20.0%	98.8% 1.2%
	98.8% 1.2%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	99.8% 0.2%	
		Target Class					
		1	2	3	4	5	

(d) Fractional Weibull training.

Test Confusion Matrix							
Output Class	1	14 14.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	10 10.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	2 2.0%	1 1.0%	5 5.0%	1 1.0%	11.1% 88.9%
	4	1 1.0%	3 3.0%	10 10.0%	10 10.0%	3 3.0%	37.0% 63.0%
	5	5 5.0%	5 5.0%	9 9.0%	5 5.0%	16 16.0%	40.0% 60.0%
		70.0% 30.0%	50.0% 50.0%	5.0% 95.0%	50.0% 50.0%	80.0% 20.0%	51.0% 49.0%
		Target Class					
		1	2	3	4	5	

(e) Gaussian test.

Test Confusion Matrix							
Output Class	1	15 15.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	1 1.0%	6 6.0%	0 0.0%	0 0.0%	0 0.0%	85.7% 14.3%
	3	0 0.0%	1 1.0%	0 0.0%	2 2.0%	0 0.0%	0.0% 100%
	4	4 4.0%	3 3.0%	10 10.0%	15 15.0%	3 3.0%	42.9% 57.1%
	5	0 0.0%	10 10.0%	10 10.0%	3 3.0%	17 17.0%	42.5% 57.5%
		75.0% 25.0%	30.0% 70.0%	0.0% 100%	75.0% 25.0%	85.0% 15.0%	53.0% 47.0%
		Target Class					
		1	2	3	4	5	

(f) Fractional Gaussian test.

Figure 3.24: Confusion matrix of classification mission for training and testing generated by optimal fractional and non-fractional distributions.

Test Confusion Matrix						
Output Class	1	2	3	4	5	
	5 5.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	12 12.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	2 2.0%	1 1.0%	1 1.0%	2 2.0%	16.7% 83.3%
	5 5.0%	3 3.0%	9 9.0%	10 10.0%	10 10.0%	27.0% 73.0%
	10 10.0%	3 3.0%	10 10.0%	9 9.0%	8 8.0%	20.0% 80.0%
						25.0% 75.0%
						60.0% 40.0%
						5.0% 95.0%
						50.0% 50.0%
						40.0% 60.0%
						36.0% 64.0%
						Target Class

(a) Weibull test.

Test Confusion Matrix						
Output Class	1	2	3	4	5	
	11 11.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	11 11.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	2 2.0%	1 1.0%	4 4.0%	0 0.0%	14.3% 85.7%
	1 1.0%	3 3.0%	10 10.0%	9 9.0%	9 9.0%	28.1% 71.9%
	8 8.0%	4 4.0%	9 9.0%	7 7.0%	11 11.0%	28.2% 71.8%
						55.0% 45.0%
						55.0% 45.0%
						5.0% 95.0%
						45.0% 55.0%
						55.0% 45.0%
						43.0% 57.0%
						Target Class

(b) Fractional Weibull test.

Figure 3.25: Confusion matrix of classification mission for training and testing generated by optimal fractional and non-fractional distributions.

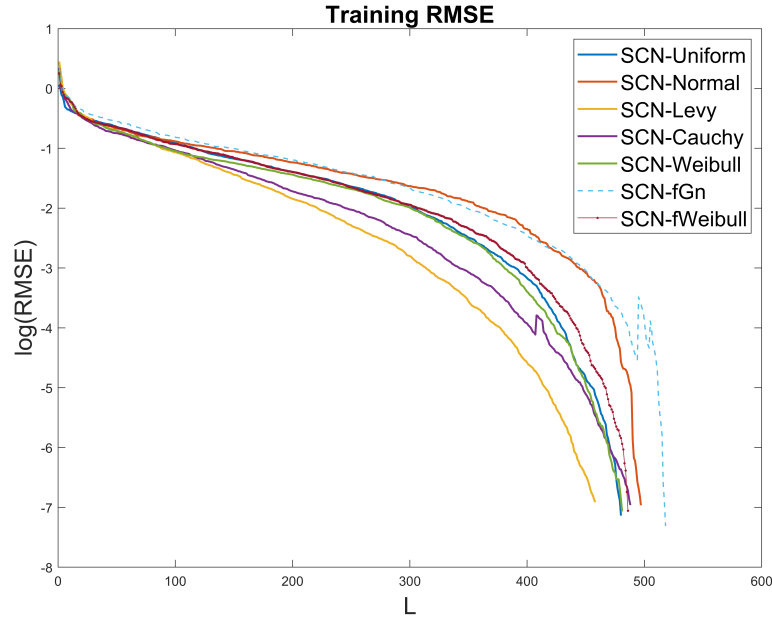


Figure 3.26: The performance of RMSE using tuned optimal parameters for all distributions. L represented the total number of hidden layers of SCN generated, and the y -axis represented the converted RMSE value by logarithmic.

We also compared RMSE behaviors of 7 distributions and applied a logarithmic scale to the y -axis for clear observations. The x -axis represents the total number of hidden layers generated. As we can see SCN-fGn (with $H = 0.5$) has the lowest value of the RMSE but also has the maximum hidden layers which makes the model more complex than the other. The SCN-Lévy achieves the minimum hidden layers and RMSE is the maximum compared with other distributions. The fWeibull is around medium performance among these distributions.

3.4.4 Comparison of The Optimal Parameters of Fractional Distributions and Non-fractional Distributions for Classification Mission.

Our objective in using different neural networks is to classify our soil carbon content data precisely. Thus, investigating better classification is necessary. We randomly split our dataset into 80% as a training dataset and 20% as a testing dataset. We found that fGn and fWeibull achieved great performance in the regression model, thus, we used the optimal parameters we found in Section 3.4.3 for achieving great classification performance. The confusion matrix comparison is shown in Figures 3.24 and 3.25.

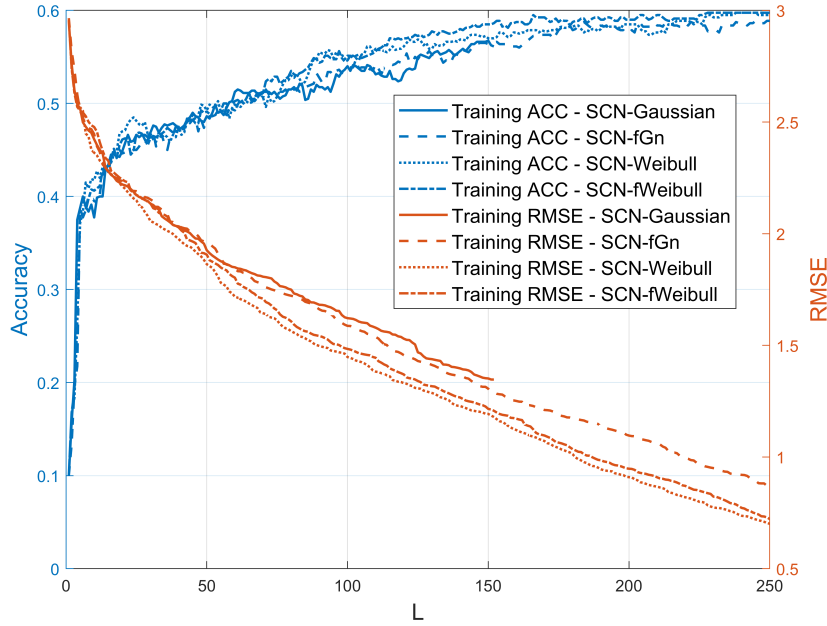


Figure 3.27: The performance of classification accuracy and their RMSE using tuned optimal parameters for all distributions. L represented the total number of hidden layers of SCN used.

From these confusion matrices, we observe that fractional distributions of Gaussian and Weibull achieve better results than Gaussian and Weibull by around 4% and 0.2% improvements, respectively (all above 99% accuracy). They have around 20% improvements by using SVM and LDA. The test accuracy is a little bit lower than they were around 50%. But all fractional distributions still achieve better results.

We also compared all training accuracy with RMSE of all distributions in Figure 3.27. We observed that fWeibull and Weibull have lower RMSE and higher training accuracy. These results have also been revealed in the confusion matrix.

3.5 Discussion and Future Work

In this work, we performed several experiments for validating our microwave and millimeter wave sensing mechanisms and sensing configurations. The comparison of the two sensors for soil carbon content has also been given. Here, we will present some discussions and future work.

Datasets/Sampling preparation: Since we used machine learning methods to classify the different soil carbon content levels and related biochar applications, proper sampling would play a crucial role in the training model by using microwave and millimeter sensors. As we know, transmitting signals from different angles/locations would generate different reflectances. To achieve reliable results, we must collect samples from various angles and repeat the process numerous times to reduce the likelihood of errors. In our work, we always fixed the relative positions and held them both stable for the sensor and samples. In addition, our previous work [223] revealed that different sensing backgrounds, such as glass and metal tables, would absorb or reflect differently, yielding uncertain information. Therefore, when users operate the microwave and millimeter wave sensing, pay attention to the background (sample containers, see Section 3.4.2.5). Moreover, the bandwidth of the microwave sensor is from 3.3 GHz to 10 GHz, which means the Wi-Fi signal may affect the sensing results.

Classification accuracy analysis: From Table 3 to Table 14, almost all classifiers meet expectations that millimeter wave classification accuracy is better than the microwave. However, all coarse classifiers, such as the coarse tree and coarse KNN, performed worse in millimeter wave classification than in microwave classification. Based on our setting, the difference between the coarse tree and the fine tree is the maximum number of splits (the coarse tree is 4, the fine tree is 100), which controls the tree depths. In other words, less proper splits of the decision tree would cause underfitting of the model. Moreover, from the scatter plots, the data distributions of the millimeter wave are close to each other, like a tight cluster and microwave data followed by a line. Based on Scikit Learn to provide the classification comparison on synthetic datasets, the decision tree usually has bad performance on

the cluster format input data². This comparison explains possible reasons but needs further proof. The coarse KNN (the number of neighbors is set to 100 for coarse KNN) and fine KNN (the number of neighbors is 1) are having the same issue, which is the under-fitting for coarse KNN with worse performance.

Machine learning method selection: After correctly setting the sensing configuration, the machine learning algorithms would play an important role in classification. We illustrate the benefits and disadvantages of the eight classifiers and their suitable datasets. LDA and SVM have high accuracy in classifying the multi-class mission in a cost-effective way (in training time). Coarse KNN and Coarse Tree perform badly in all classification missions (coarse KNN only has around 20% accuracy), and from the tables of the sensing accuracy, we observed that no matter what sort of classifiers used, the millimeter wave results are always better than the microwaves except for different container tests. Therefore, for soil carbon content sensing, we recommended selecting the SVM and LDA for microwave and millimeter wave signal rapid classification. In addition, we plan to run the model on the Raspberry Pi for in-situ sensing. A lightweight model is exactly the right choice for these objectives. In the future, we will validate several lightweight models to check their performance.

Impact of sample containers: Due to the fact that microwaves and millimeter waves can be absorbed, penetrated, and have reflective properties, the carrier or container would be much more important for influencing the signals captured and feature extraction. As discussed in Section 3.4.2.5, the different containers do affect the sensing accuracy, and both microwave and millimeter sensing accuracy are lower than using the same container (discussed in Section 3.4.2.1). Moreover, we always expect that millimeter waves can achieve better-sensing results compared with microwaves due to the millimeter wavelength being less than soil carbon (biochar) size. However, millimeter wave has the worse anti-disturbance ability and is more sensitive to different containers.

Moreover, we experienced that microwaves and millimeter waves can penetrate glass, wood, and plastic with bare absorption, the reflection would be absorbed through the water, and the signals would be highly absorbed and scattered through the mental medium [223, 224]. In our sensing practice, we should pay attention to avoid this interference.

The IMAGEVK-74 (millimeter wave sensor) has total 40 antennas for transmitting and receiving signals, where each receiver would receive all the transmitters generating signals, also referred to as a sensor array. The advantage of the sensor array is that it provides new dimensions to the observation, allowing more parameters to be estimated and improving estimation performance. An array of radio antenna

²https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

elements used for beam-forming, for instance, can raise antenna gain in the direction of the signal while decreasing antenna gain in other directions, thereby boosting the signal-to-noise ratio (SNR) by amplifying the signal coherently. Because our target is soil carbon, which is particles in the soil, using a sensor array to sense the soil carbon content and distribution would give more guidance for precision agriculture applications.

In the future, we plan to explore more proximal sensors, such as spectrometers, for soil carbon content quantification. As we know, the spectrum of most soil, carbon, water, etc., is focused on infrared bandwidth. We believe that combining microwave, millimeter wave, and infrared response spectrum reflectance would further improve soil carbon management. In addition, for proximal, in situ sensing exploration, we would explore the XAI (eXplainable artificial intelligence) research opportunities for predicting soil carbon content spatial distribution levels much more precisely.

3.6 Chapter Summary

In this chapter, we demonstrated the comparison of using microwave and millimeter waves to identify different amounts of soil carbon with their feedback origin, usage, soil moisture, soil moisture under the effects of carbon application, and background characteristics of soil carbon content sensing. For the machine learning classifier, we selected commonly used classifiers for classifying millimeter-wave and microwave reflectance signals. In general, millimeter wave sensing is better than microwave sensing for soil carbon content, which is around 10–15% improvement. We illustrated the sensing mechanism and configuration from the proximal sensing perspective of using microwave and millimeter waves for sensing soil carbon and related biochar applications. The optimal algorithm parameters and the model regression abilities are investigated. Our objective is to improve the simplicity and accuracy of biochar application and the soil condition monitoring capabilities using such sensing systems in real-time. In the future, we plan to attach the mentioned two sensors to UAVs and UGVs to make such sensing systems easier to use and more accurate for biochar application and soil condition monitoring in large-scale fields.

Note that all reported results are reproducible and the collected datasets are available by visiting the DRYAD³ [130].

³<https://doi.org/10.6071/M3M092>

Chapter 4

SITE SPECIFIC MANAGEMENT OF SOIL CARBON CONTENT SENSING ON LARGE-SCALE FIELD

4.1 What is Site Specific Management?

Agricultural site-specific management (SSM) entails the changeable treatment of soils and crops according to field-specific conditions. SSM, also known as “Grid Farming”, “Farming by Soils”, and “Variable Rate Technology (VRT),” is a fast developing combination of technologies that enables farmers to manage their soils and crops as equipment moves across a field. SSM implies doing the right thing at the right time, in the right place, and in the right quantity. Consequently, SSM is intuitively attractive since it represents a method for enhancing the economic and environmental performance of cropping systems. Despite its enormous potential, SSM is an emerging technology yet in its infancy. SSM has its strengths (Pros) and its drawbacks (Cons), which together describe the current status of SSM for agriculture (Realities). The following is a brief summary of some fundamental features of SSM and the degree to which these elements have been developed or not [225].

Five core components comprise a site-specific management system: (1) spatial reference; (2) measurement and monitoring of crop, soil, and environmental attributes; (3) attribute mapping; (4) decision support system (DDS); and (5) differential action. In the mid-1990s, notably with the maturity of global navigation satellite systems (GNSS), in situ and on-the-go sensor technologies, geographic information systems (GIS), and variable-rate technology. These enabling technologies make SSM widely adapted [225].

SSM helps to specify inputs to the demands of each management zone within a field once the variation within a field has been identified and documented. This involves three essential steps: (1) the identification of management zone within a field, (2) an acceptable recommendation for each managed input for each management zone inside a field, and (3) the capacity to precisely apply the managed inputs such as water irrigation, fertilizer, pesticide etc. (precision agriculture) [226].

Overall, the SSM can be summarized in three actions: to know current location, to gather needed information, and to make decision and action at the location. In our scenario, the soil carbon content, soil carbon erosion, and carbon emissions may be spatially varying. Therefore, utilizing the SSM strategy is essential to tackle

the various environmental challenges. XAIoT empowered edge devices, UAVs, and UGVs can help to perform the site-specific management. Note that UAVs and UGVs not only considered as sensing platform, but also considered as actuators, which means if any zone has lower soil carbon than desired level, the UAVs and UGVs would spread or apply biochar to increase the carbon content. Here, we introduce Digital Twin to as a uniform platform enhance the efficiency in mitigating carbon emission. We presented several preliminary results of using Digital Twin to implement SSM in the following subsections.

4.2 Digital Twin Enabled Methane Emission Abatement Using Networked Mobile Sensing and Mobile Actuation

4.2.1 Introduction

Global Methane Challenge¹ as an international campaign of reducing methane has attracted much attention for responding to potential global warming. Methane is the second most significant known source of greenhouse gases, accounting for most contribution to global warming. It is well understood that if the global climate warms enough, permafrost will melt, releasing methane and CO₂ plus unknown ancient viruses/bacteria [227], which will have an impact on human health. Methane is abundant in the natural world, derived from bacteria in the soil, water, and ruminant animals. Besides, human activities such as fossil fuel burning, dairy farms, landfills, oil wells, and other sources of methane are primarily found in ecosystems industry and agriculture. As a result, the main objective of methane abatement is to reduce methane emissions into the atmosphere, especially from human activities. However, there are still many impediments to abating methane emissions. The primary research challenges, in particular, are a lack of an appropriate method for measuring time-varying methane concentrations (in a spatial and temporal sense).

Prior research into methane detection has resulted in numerous advances in sensing and actuation methods, including remote sensing by satellite [180], aircraft, or sensors on the tower to monitor methane spreading. However, these remote sensing techniques are constrained by space weather at fixed-location [148]. Due to bad weather, methane concentrations may be jammed or shifted, making monitoring more complicated and efforts hard in abating methane emissions. Another common technique is regional surveillance, which entails putting several such as clouds sensors on the ground to collect methane data simultaneously. However, the installation, repair, and calibration criteria for these approaches [228] will substantially increase their deployment expense and complexity.

¹<https://globalmethane.org/challenge/index.html>

Digital Twin in Methane Emissions/GHG within UAVs

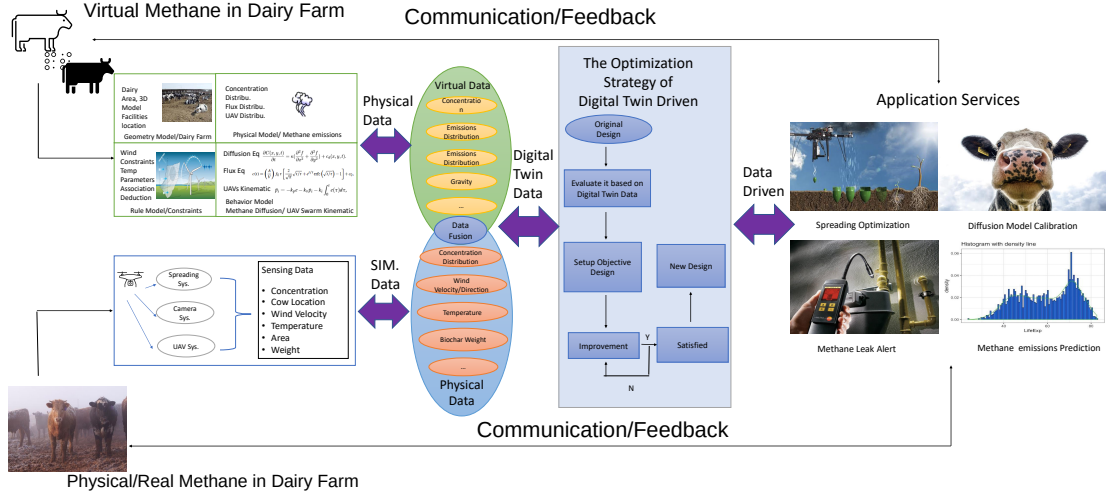


Figure 4.1: Digital twin in methane emissions quantification analysis.

To overcome these shortcomings, we introduced a Digital Twin (DT) framework for reducing methane emissions with swarming UAVs. The DT is a virtual representation of the physical asset or process that provides real-time data services. Real-time data from the UAV moving sensor is extremely beneficial for real-time methane distribution modeling using DT. This DT approach, when behavior matching is well done, can be advantageous for mapping and forecasting methane emissions because it requires fewer field trials but has a high-level prediction precision. To do so, we utilize Matlab UAV toolbox [229], ROS², Gazebo³ and Web App [230] to deploy our DT platform, and we use the UAVs as mobile sensors and actuators to track methane diffusion and will be able to spray methane neutralizers (e.g., biochar [28]) to abate methane emissions. The following are the challenges of using UAVs as mobile sensors and deploying DT: 1) Connectivity instability and security concerns about real-time data transfer from UAV to DT platform, 2) The uncertainty associated with methane distribution concentrations that are time-varying, adding ambiguity to the Digital Twin. To tackle these challenges, we designed our DT system to: 1) transmit data opportunistically while UAVs are not performing tasks (e.g., path planning, hovering, etc.), 2) apply the physical principle to develop a core model of methane emissions, and 3) utilize machine learning algorithms to iterate the DT or, behavior matching using real-time data efficiently. In summary, the contributions of this work are as follows:

²<https://www.ros.org/>

³<http://gazebosim.org/>

- We presented a DT that contains Matlab UAV toolbox, ROS, and Gazebo for methane emission abatement;
- We tested our DT prototype of methane emissions to demonstrate the feasibility, challenges, and opportunities of using the DT.

4.2.2 Digital Twin in Methane Abatement

Due to the ineffectiveness and inaccuracy of the conventional approach of monitoring methane emissions (More in Section 4.3.1), we need to develop a method that accurately represents methane emissions processes and is also cost-effective. As a result, DT as an Industry 4.0 component has come to our attention, and to our best knowledge, there is no relevant work using DT to abate methane emissions. DT utilizes various subsystems, including sensors and actuators, to gather data from the natural environment. DT will analyze these data in order to make a decision and give actuators instructions. Figure 4.1 illustrates the role of each subsystem of DT in methane emissions. Although DT has been shown to be effective and valuable, there are numerous unresolved questions, which will be addressed in the following subsections.

4.2.2.1 The Physical Principle Behind Methane Emissions

In this section, let us assume wind speed is at zero, then the methane as gas diffusion satisfies fluid dynamics, which that the dynamically spreads in the time and space domain, respectively. Partial differential equation (PDE) (4.1) is used to describe methane emission diffusion dynamics:

$$\frac{\partial f(x, y, t)}{\partial t} = \kappa \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \right) + c_d(x, y, t), \quad (4.1)$$

where $f(x, y, t)$ is methane concentration, x and y are directions in space, t is the time, κ is the diffusivity constant and c_d is methane source.

We can easily construct virtual methane emission prototypes using the equation above and assume that the actual wind speed is none-zero. The following step is to consider the impact of the uncertainty disturbances.

4.2.2.2 Uncertainty Conditions of Methane Emissions

In general, the uncertainty conditions include but are not limited to wind, local temperature, different types of methane emission sources (such as point sources, uniform distribution area sources, non-uniform distribution area sources, etc.), and other unknown issues. Different uncertainty conditions need to add specific terms to equation (4.1) to approach accurate methane emissions.

Here is the idea of how to apply the wind field to the PDE. Once we get the current wind speed and direction, we use the fluid dynamics equation as wind force

split into x and y directions, respectively. This term needs to be added to methane disturbance and actuators' sides. Hence, based on different wind speeds and their directions, it would approach natural methane emissions better.

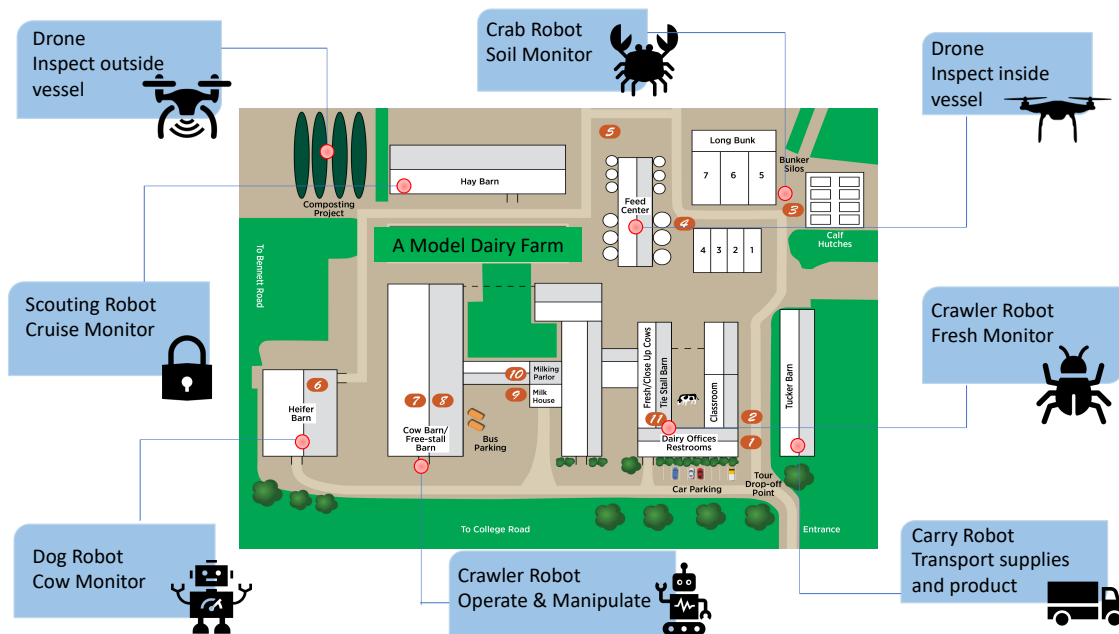


Figure 4.2: Mobile robots perform various tasks in a dairy farm.

Variations in local temperatures either accelerate or decelerate the kinetic activity of methane molecules, thereby exerting a significant influence on methane emissions. Notably, these temperature fluctuations can be attributed to anthropogenic sources such as dairy farms, landfills, and oil wells. Incorporating temperature variables into differential temperature models is essential for effectively monitoring methane emissions and assessing methane leakage.

To sum up, we discussed the essential elements of DT, which include the ability to process real-time data, behavior matching, a modular structure, trackability, and re-programmable. These components are feasible and are crucial for addressing uncertainty in measuring environmental methane emissions. To address any unforeseen or unfamiliar disruptions, we can employ an algorithm to train real-time data and integrate it into our decision making, such as through machine learning techniques.

4.2.3 Networked Mobile Sensing and Actuation

Mobile sensors have advantages in acquiring fine-grained information and an extensive detection spatial range for dynamic system characterization. Our UAVs fly back and forth to map desired land fields, dairy farms, and so on. According to the DT philosophy, multiple robots, each performing different mission tasks,

equip themselves with sensors to perform sensing fusion and provide data to the DT system. Figure 4.2 demonstrates the possibility that we can have multiple mobile sensors to finish sensing tasks. This figure suggests that we should equip more diverse mission sensors to achieve more accurate data for DT.

Besides, the UAVs can also be used as mobile actuators to help DT find the optimal location to spread biochar to abate methane emissions as much as the UAVs can. Furthermore, to make UAVs fly to the desired locations as fast as possible, we need to explore an optimal coverage control problem to find a better actuator distribution strategy to effectively sample methane data and exchange information between DT and physical entities. The related coverage control methods include lead follower [231], potential field [232], virtual structure [233], centroidal Voronoi tessellations (CVT) [234], etc.

After we combined mobile sensors and closed-loop control actuators, we have a robust wireless sensor and actuator network (WSAN) for information exchange. In addition, the battery management system decides whether to stay or act based on remaining energy, which needs to be further noticed.

4.2.4 Preliminary Results

To validate our DT, we 1) conduct UAV toolbox, Gazebo, and ROS experiments for UAV flight paths and waypoints following, and 2) write a script for applying PDE (4.1) as the platform of DT to build methane emissions dynamic processes.

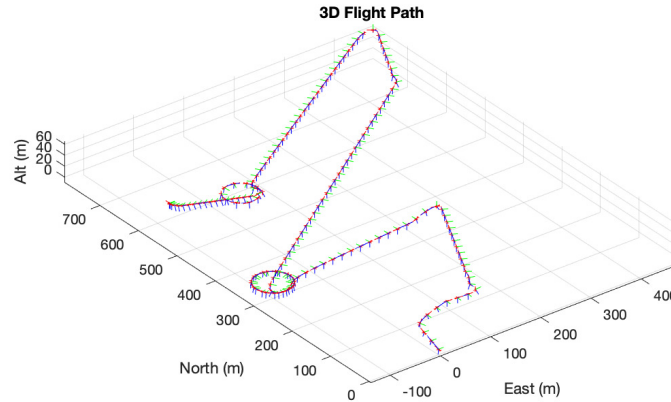


Figure 4.3: UAV waypoint flight path within UAV toolbox.

We utilized the ROS toolbox in Matlab to make the connection between the UAV toolbox and Gazebo. The coverage area and flight paths of UAVs were tested. Gazebo drove the physical engine to demonstrate flight status and how UAVs fly under uncertain conditions. Figure 4.3 shows examples of UAV fixed-point flight under different altitudes. The swarming UAVs were implemented in Gazebo at Figure 4.4 to check swarming UAVs of coverage control feasibility. Figure 4.5 demonstrates the prototype of methane emissions under no wind effects, and we observed a methane diffusion process that approaches the maximum diffusion concentration (in the spatial domain) almost at the middle of simulation time (total simulation time is 4s), which demonstrates the methane diffusion characteristics.

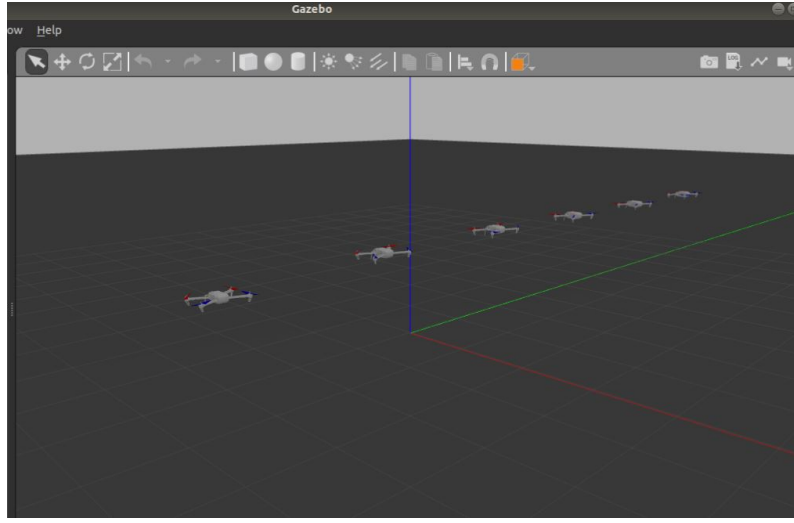


Figure 4.4: Swarming UAVs in Gazebo.

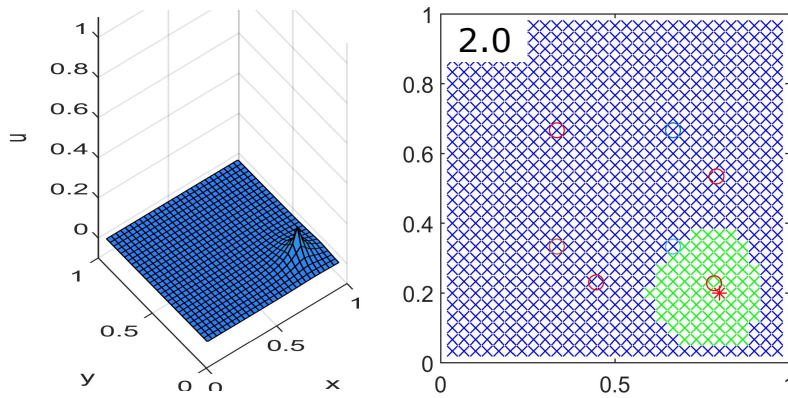


Figure 4.5: Methane diffusion process.

Since there is no standard evaluation metric for DT, we only check task completion on each subsystem. The 3D flight path of testing fixed-point flight that performs well, the coverage swarming such as the CVT algorithm in Gazebo perform efficiently. The PDE experiments was successful in simulating the methane diffusion process. The only thing we need to notice about the order of PDE is that it affects diffusion temporal and spatial information.

4.2.5 Discussion and Future Works

Networked sensors and actuators are mainly composed of UAVs, and their battery life is limited to the total mission lifetime. Besides, the payload of UAVs also needs to be considered, as we need to use UAVs to spread biochar to abate methane emissions. Thus, selecting a coverage strategy to cover the entire field as fast as the UAVs can is the critical component of implementing real-time data updating.

When we simulate the methane diffusion process, the order of the PDE does matter for methane emissions in the time and spatial domains, respectively. There are super diffusion and subdiffusion that correspond to different orders. Besides, adjusting the order of PDE to meet those unexpected issues (wind field has many types, turbulence, eddy, meandering [235], etc.) as referred to fractional order is one way to approach DT in the practical situation of methane emissions. Once we have configured the DT of methane emissions, we can easily predict and abate methane by controlling our UAVs.

The opportunities and challenges of DT introduced by a combination of UAVs as mobile sensors and actuators, real-world data input, etc., as virtual representation can be further explored. We plan to explore the DT in the following directions: 1) optimizing the UAV sampling operation schedule to increase the opportunity to capture methane emissions and minimize the total mapping and prediction cost; 2) deploying the Web App as user interaction to combine these subsystems (methane emissions, UAVs, sensing, actuators, etc.) into the DT; and 3) collecting actual methane emission data to enhance the DT by selectively hovering at desired locations that allow high-fidelity data acquisition and transmission.

4.3 A Digital Twin Enabled Internet of Living Things (IoLT) Framework for Soil Carbon Management

4.3.1 Introduction

Numerous environmental challenges that lead to climate or agricultural crises on a global scale prompt us to consider innovative technologies that facilitate a quicker recovery of our natural ecosystems. The first question is always how to observe our ecosystem efficiently. We have recently become interested in a new concept known as the Internet of Living Things (IoLT). The key idea is we envision Internet-enabled cyber interfaces (or cyber insects) paired with living insects that

would revolutionize the way we manage natural ecosystems by managing their behavior and communication with the surrounding environment [236]. However, it's difficult to employ sensors on insects to monitor soil conditions, and that would also break ecosystem balance. Therefore, we employ AI on the Internet of Living Things (IoLT), which has the potential to revolutionize soil carbon management. AI-driven IoT generates intelligent technologies that mimic intelligent behavior and facilitate decision-making with minimal or no human involvement. We declare this combination Artificial Intelligent Internet of Living Things (AIoLT) [237].

Digital Twin (DT) [238] is a popular concept in Industry 4.0. It is the digital replicas of real-world objects designed to simulate, evaluate, and explain their physical equivalents' behavior. In soil carbon accounting management, DT duplicates past objects and simulates future objects to provide farmers with actionable scenarios. It enables farmers and stakeholders to respond to unexpected deviations, for instance, by diagnosing problems on site and in advance, scheduling predictive maintenance such as irrigation and fertilizer at the optimal time, optimal place, and optimal amount, and by offering quick-reaction solutions to problems that are accumulating. For carbon-neutral and carbon-negative techniques, DT speeds up the production, marketing, and sales processes.

Previous techniques for assessing changes in soil carbon have used direct measurement [183], indirect measurement (such as remote sensing) [108], model estimates (using process-based or empirical models) [239], and hybrid methods that combine direct sampling and modeling [240]. In general, soil carbon accounting includes sampling, which involves randomly selecting a few bulks at a depth of approximately 30 cm (with some recommending sampling down to 1 m) [241], air drying samples, grinding samples, and then putting them into expensive element analyzers, which require extensive lab resources and hardware costs. Besides, the main items of the soil carbon accounting baseline take into account the time domain and spatial domain, respectively, to account for fluctuations resulting from past management decisions and current distribution, including temperature variation, human activities, moisture variation, etc. These methodological inconsistencies can result in vastly different estimates of carbon net sequestration rates over the life cycle of soil carbon footprints.

In addition, only a few studies have discussed soil carbon accounting using DT. In [239, 242], the authors discussed digital twin-enabled methane emission by using networked UAV mobile sensing and model simulator-based estimation of methane emission. A study of soil microbiomes for climate mitigation [243] has been illustrated by deploying a DT platform to monitor the GHG emissions situation through microbial metabolism. These studies all target the single scenario of carbon accounting, which cannot meet the needs of soil carbon accounting.

To overcome these drawbacks, we utilized the DT as an essential analytical platform to manage soil carbon accounting. The physical world data collection

would use AIoLT-enabled edge devices (soil carbon sensors, which are microwave and millimeter wave)⁴⁵, mobile platforms such as swarming UAVs and UGVs to obtain long-range and fine-grained information that would be helpful for soil carbon management in large-scale fields.

To deploy the DT platform for virtual soil carbon management, ML algorithms would be employed to analyze reflective signals from microwave and millimeter waves that contain the soil carbon content. In the meantime, instructions would be sent to UAVs for spraying the proper substances, such as water, biochar, and fertilizer, that sequester carbon in the soil. The main challenges of using AIoLT-enabled sensors and UAVs for deploying DT are: 1) how to assign artificial intelligence to soil living things such as insects, plants, and microbiomes in order to reveal soil carbon footprints; and 2) the AIoLT evaluation metric must meet DT properties [244]. To tackle these challenges, we proposed a DT system to: 1) deploy proximity sensors (AIoT) to observe the insects, plants, and microbiome distribution situations that compose it into AIoLT; 2) develop a kinematic model that performs the implication of the affection to soil carbon footprint by the living things; and 3) deploy ML algorithms to power on DT using real-time data for interaction with the real world. In summary, the contributions of this work are as follows:

- To our best knowledge, this is the first time to utilize the DT as the essential framework to engage the edge-AI proximity radar sensors to do soil carbon management referred to as AIoLT.
- We present an SCMDT evaluation metric to validate AIoLT with standardized soil carbon accounting management.

4.3.2 Digital Twin in Soil Carbon Management

Most soil carbon exists in the form of soil organic matter (SOM) [245]. Soil organic matter consists of microorganisms such as bacteria and fungi, decomposing matter from once-living species such as plant and animal tissues, feces, and the products of their decomposition. SOM is a diverse collection of elements ranging from fresh plant remains to a highly decomposed substance, humus. SOM is composed of carbon-heavy organic molecules. Soil organic carbon (SOC) levels [246] are directly proportional to the amount of organic matter in soil, and SOC is frequently used to assess organic matter in soils.

Several ecosystem processes, including photosynthesis, respiration, and decomposition, interact to determine SOC levels. Photosynthesis is the process through which atmospheric CO₂ is converted into plant biomass. A plant's root biomass

⁴<https://shop.walabot.com/products/walabot-developer-pack-new>

⁵<https://blog.vayyar.com/ignite-intelligence-vayyar-element>

primarily determines SOC input rates, but plant shoots also deposit litter. Climate conditions, particularly soil temperature and moisture, influence the rates of photosynthesis, decomposition, and respiration. Therefore, our soil carbon management digital twin (SCMDT) will mainly focus on SOC, living things, soil temperature, soil moisture measurement, and our proposed SCMDT solution, as illustrated in Figure 4.6.

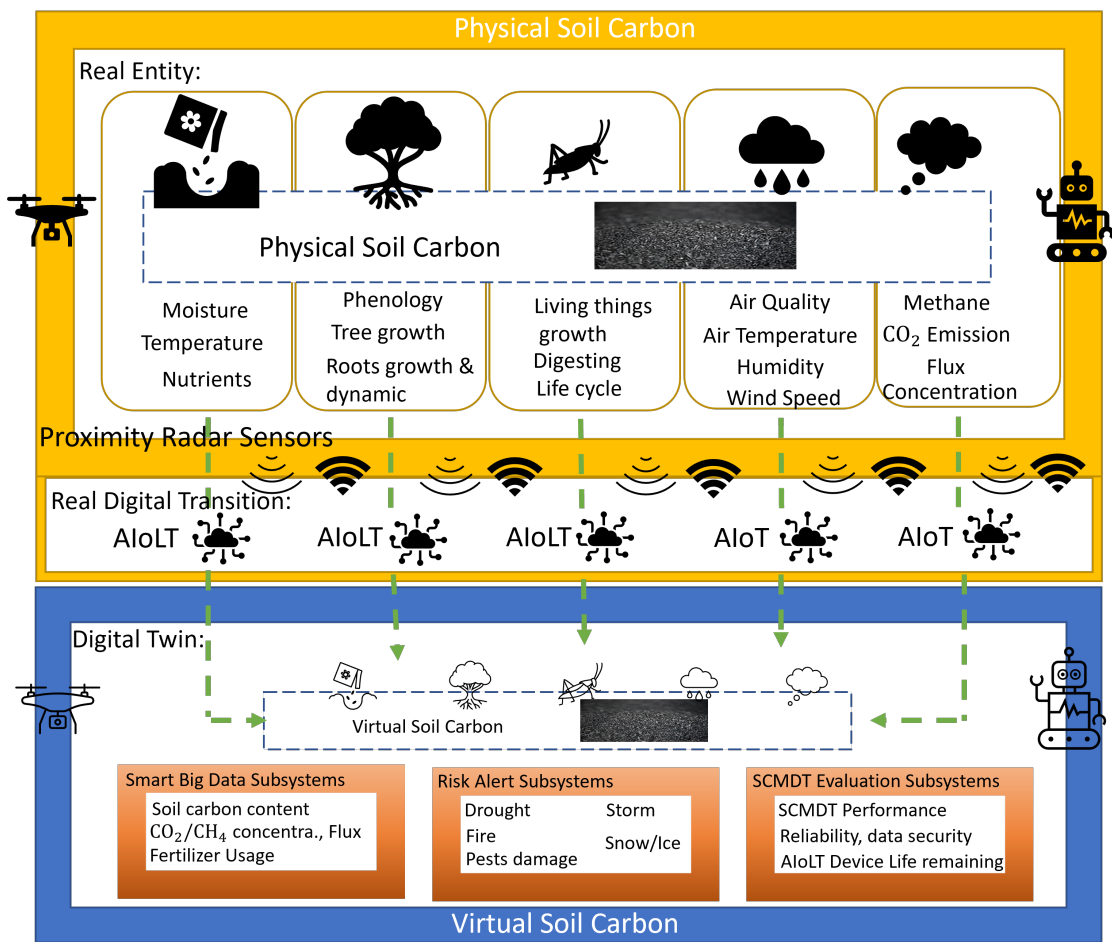


Figure 4.6: SCMDT-Digital twin in soil carbon accounting.

4.3.3 Digital Twin Front End of Soil Carbon Accounting

4.3.3.1 Proximity Sensors for Measuring SOC

As we previously mentioned, almost all approaches to measuring SOC and greenhouse GHG emissions include collecting soil samples at designated sampling

locations. Scientists typically place these samples into element analyzers after air-drying, grinding, and weighing them. Some researchers also employ portable devices⁶ to detect carbon-based gases such as methane, CO₂, etc. However, these methods face limitations such as low spatial resolution, high labor costs, and additional hardware costs. As a result, we would use proximity sensors to tackle the challenges we mentioned and measure SOC both in practice and in real-time.

As shown in Figure 4.7, the microwave radar sensor is shown on the left, and the millimeter wave radar sensor is shown on the right. In [135], the researchers already verified possibility of using proximity radar sensors to measure soil carbon and soil moisture content. Using these proximity radar sensors to measure soil carbon has many benefits. For instance, quantifying soil carbon is less labor-intensive, and the proximity radar sensor is a small device (hand-sized) that is easily equipped on autonomous drones or vehicles to detect large-scale soil cost-effectively. Lastly, the implementation of the proximity sensor will not disrupt standard agricultural activities, including irrigation and fertilization. This sensor captures signal reflections in real-time, making it an ideal data collection device for DT analysis.

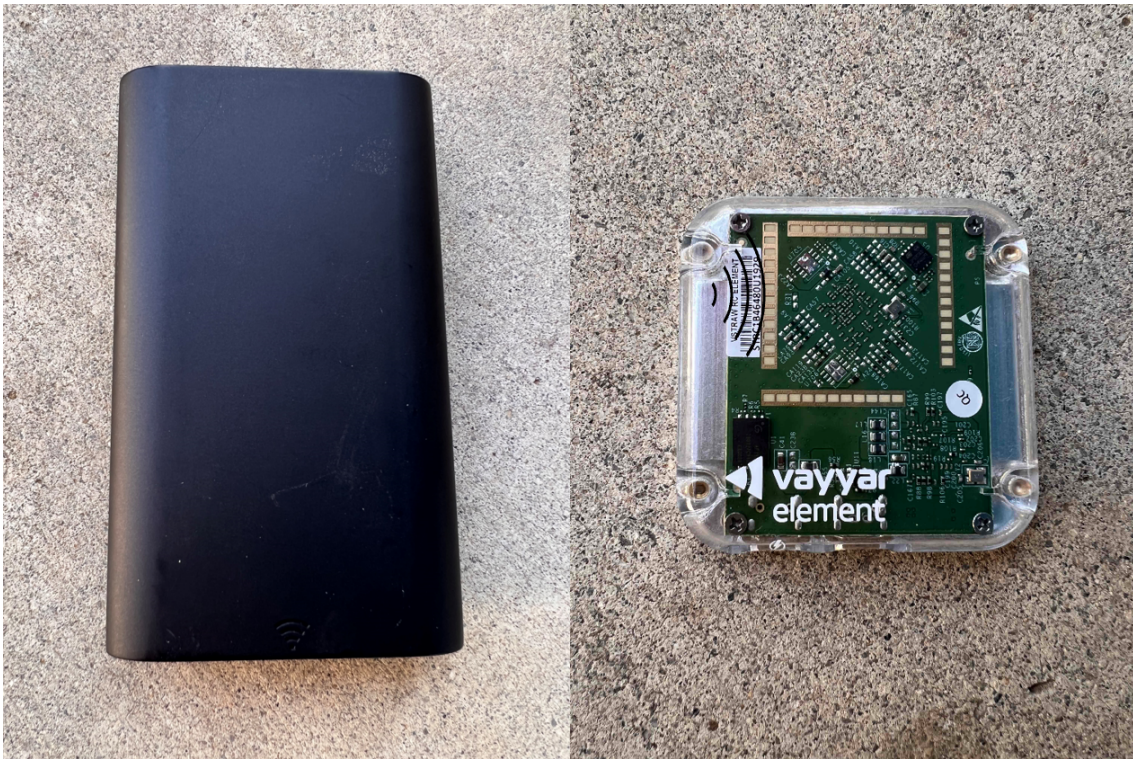


Figure 4.7: Proximity radar sensor, left is microwave radar sensor, right is millimeter wave radar sensor.

⁶<https://www.gazomat.com/en/products/gazoscan/>

4.3.3.2 Swarm UAVs Mobile Platform

For the characterization of dynamic systems, mobile sensors offer the benefit of gathering local area data (fine-grained) and a broad detection spatial range. Our UAVs fly back and forth in order to map desired sensing areas in large-scale fields. The DT philosophy dictates that we equip multiple robots with sensors to perform sensing fusion and transmit data to SCMDT. Previous research [242] has shown that several mobile sensors can complete long sensing range tasks.

In addition, for UAVs to reach their destinations as quickly as possible, we should investigate an optimal coverage management issue in order to identify a better actuator distribution method for efficiently sampling soil carbon content data and exchanging information between digital and physical entities.

4.3.4 Digital Twin Back End for Soil Carbon Accounting

4.3.4.1 Soil Carbon Classification Using Machine Learning

As indicated previously, soil carbon exists in a variety of formats, signifying that it has a variety of structures [127]. The surface of living organisms, in particular, is comprised of numerous small, porous structures. Therefore, the microwave signal permeates, absorbs, and reflects these structures when it encounters them. In other words, by analyzing the reflective signals, we may obtain soil carbon data like the amount remaining, different formats of soil carbon (solid, liquid, gas), water retention capacity, etc.

The next stage is to analyze and classify the radar reflective signals. To reduce the possible effect of classification algorithms on sensing accuracy, we looked at nine popular classification algorithms in Table 4.1 in detail in [135, 223].

Table 4.1: Selected nine machine learning algorithms with parameter setup.

Name
K-nearest Neighbors (KNN)
Decision Tree
Gaussian Bayes Naive (GNB)
Random Forest (RF)
Linear Regression (LR)
Support Vector Machine -Linear(SVML)
Support Vector Machine - RBF(SVMR)
Support Vector Machine - Sigmoid (SVMS)
Linear Discriminant Analysis (LDA)

These classifiers all have great performance in soil carbon accounting. We aim to create tools for data labeling, data preprocessing, and data analysis to support DT and facilitate the implementation of further UAV commands.

4.3.4.2 Smart Big Data for Soil Carbon Accounting

IoT and edge devices (sensors and actuators) help researchers solve complicated agricultural challenges [247]. Farmers or related researchers may track critical data for farm management, such as soil moisture, wind speed, air temperature, humidity, etc., The volume of data might be enormous and time-consuming to process. Therefore, it becomes necessary to "smarten up" this vast amount of data, a question that will also arise in the context of soil carbon accounting. Soil temperature, soil moisture, soil living things distribution, edge devices power life remaining, soil carbon content, and so on need to be processed in real time and given useful information. Therefore, we propose the concept of smart big data analysis. Smart big data in soil carbon is an interdisciplinary research topic concerned with the extraction of meaningful information, utilizing techniques from multiple fields, such as UAV image processing, deep learning, pattern recognition, high-performance computing, and statistics. We can filter and transform big data into smart big data before examining it for insights, leading to more effective decision-making. In Figure 4.6, we deployed the three main subsystems, which are SOC measurement, risk alert, and SCMDT evaluation, to compose the smart big data system and towards SCMDT for running up the soil carbon accounting mission in the meantime.

4.3.4.3 Digital Twin in Soil Carbon Accounting Assessment

The verification and assessment of SCMDT need thorough exploration. We believe that evaluating the performance of SCMDT involves several critical areas, including system fault tolerance, soil data validation, data storage, data security, data pre-processing, smart data processing, and the optimized antenna combination of radar sensors. Additionally, effective communication strategies for edge AI sensors and UAVs must also be considered to ensure comprehensive functionality and reliability. Therefore, to enable a standardized and quantifiable evaluation of SCMDT, we created a model that takes into account the components we mentioned earlier and accuracy of prediction for soil carbon content to determine the SCMDT performance:

$$SCMDT = q - \sum_{a=1}^b \omega_a f_{sig}(d_a), \quad (4.2)$$

where $f_{sig}(d_a)$ represent the sigmoid function applied to the b^{th} components value we mentioned above, and f_{sig} is the sigmoid function to restrict the value to $[-1, 1]$. The ω_a is the weight of each component a , which could be adjusted dynamically due to different missions. The total SCMDT value range is $[-1, 1]$, and q is the soil

carbon classification accuracy. Overall, the SCMDT represents data reliability and DT data processing performance.

4.3.5 Discussion and Future Work

Edge sensors, AIoLT, proximity radar sensors, and UAVs can all serve as networked sensors and/or actuators, but their battery life limits the duration of the mission. When we evaluate the SCMDT's remaining life, battery consumption becomes an essential concern. Thus, establishing an optimal strategy and self-health awareness battery management system to cover the entire field as quickly as possible using these AIoLT sensors is an interesting future research effort.

The opportunities and challenges of SCMDT introduced by a combination of edge AI sensor devices, AIoLT, proximity radar sensors, and virtual soil carbon content can be investigated further. We intend to further investigate the SCMDT in the following areas: 1) increasing the opportunity to capture soil carbon accounting costs by optimizing the AIoLT proximity radar sensor distribution; 2) deploying a uniformed user interaction to combine the subsystems (soil carbon content, proximity radar sensor, SCMDT evaluation, sensor life remaining, etc.); and 3) Enhancing the strategy for monitoring soil carbon on a site-specific basis, where high-quality data can be systematically observed and recorded.

4.4 A Case Study: Long Endurance Site-Specific Management of Biochar Applications Using Unmanned Aircraft Vehicle and Unmanned Ground Vehicle

4.4.1 Introduction

In recent years, the world has paid attention to climate change, and one of the main reasons is that carbon emissions have lost control. According to the Environmental Protection Agency (EPA), in 2020, methane (CH_4) accounted for about 11% of all U.S. GHG emissions from human activities⁷, and most of them were the methane generated from dairy farms. Methane is the second-greatest contributor to global warming. To mitigate climate change, we must sequester carbon emissions in the soil. Biochar is a carbon-rich product that can improve soil health by enriching its soil structures [248]. However, determining which places require carbon and which generate lots of GHG remains a huge challenge.

Carbon sequestration is the process of trapping and storing atmospheric carbon dioxide. Consequently, it can be employed as a form of climate adaptation due to its proven positive effects on soil erodibility, water retention capacity, temperature, and gross primary production [249]. Moreover, SSM in agriculture has been discussed widely and considered a characteristic of agricultural 4.0 [250]. Using SSM

⁷<https://www.epa.gov/ghgemissions/overview-greenhouse-gases>

to capture, measure, quantify, and mitigate carbon emissions is suitable for agricultural applications, especially on large-scale fields. The essential elements of SSM are site spatial information and obtaining needed information, then making the correct decision to manage. Several studies work on site-specific management using UAVs and UGVs [181, 251–253]. In [253] and [252], these works confirmed the feasibility of using UAV and UGV to manage nitrogen using remote sensing image processing. In [251] used UAVs for imaging remote sensing and UGVs as actuators to spray chemicals, and [181] created a smart bi-bike as a UGV for strawberry pest control management. However, none of these studies combined UAV and UGV to perform agricultural and biochar applications from an SSM standpoint.

Many studies have been conducted on UAV landing on moving targets, and we chose several pieces of literature to summarize their contributions [254–263]. The UAV’s landing on a moving target can be categorized mainly in three ways: vision-based control, optimal landing control strategy, and model predictive control (MPC). Vision-based control includes traditional image processing, deep learning image processing, etc. [254]. The optimal landing control strategy is based on multiple sensor fusion data sources, such as IMU, GPS, and ultrasonic sensors, to get the landing coordinates [262]. The MPC method is based on UAV and UGV kinematic models to estimate and predict the vehicles’ status for landing [259]. These studies are able to solve the landing on moving targets problem, but they are costly; they require lots of computation resources, labor intensive work, and well-trained people to manipulate the UAVs.

To tackle the challenges, we proposed a low-cost and long-endurance agricultural carbon emissions site-specific management strategy using the combination of UAV and UGV. In our system, long endurance has referred to UAVs and UGVs having long flight times for the practice of biochar application on large-scale fields. The significance of the system lies in the UGV’s role as a moving platform, enabling UAVs to gather fine-grained information without power limitations. The sensor payload consists of millimeter-wave (mm-wave) radar and microwave radar sensors that have been proven to capture soil carbon content and GHG without causing harm to the soil [135]. Thus, we designed our sensing system with two benefits:

- Our system requires less labor to operate on a large-scale field than other SSMs that use a UAV and a UGV for agricultural purposes.
- To cut costs, the edge devices would be a co-pilot system that would measure the amount of carbon in the soil and process the moving landing mark.

As a result, we target two challenges in this paper: 1) how to provide a reliable landing strategy by using a simple controller (a proportional integral (PI) controller); 2) how to process the sensing data on time and the actuation (spray biochar or other necessary allowable chemicals). Our system provides an in-situ SSM strategy for soil carbon content and biochar applications. The contributions are as follows:

- To the best of our knowledge, this is the first time that long-endurance sensing and actuation for biochar applications have been implemented by a UAV landing on a moving UGV as a backup station.
- To do proximity sensing, we equipped the UAV and UGV with mm-wave radar and microwave radar sensors as payloads.
- We conducted experiments to evaluate our mission-guided biochar application system.

The rest of the paper is organized as follows: Section 4.4.2 depicts the system’s framework architecture from a mission perspective. Section 4.4.3 presents the evaluation setup, and the preliminary results are shown in Section 4.4.4.

4.4.2 Framework and Architecture

In this section, we first introduce the complete scenario of mission planning using a UAV and UGV combination to detect field carbon emissions. Then we explain our sensing system, including the UAV, UGV, and co-pilot system that help complete the sensing mission. The payload, such as carbon emissions sensors, batteries, and a spray tank, would also be introduced later.

4.4.2.1 Mission Planning

We mainly focus on UAV autonomous landing on UGV while sensing carbon emissions to improve SSM endurance. In Figure 4.9, the UGV would follow the field path, slowly moving on the trail, and spray necessary chemicals and biochar if needed. To collect carbon emission information, the UAV would fly toward the survey area. The specific sensing process involves the UAV flying by the top of the survey area, hovering or landing on the crop or field, sampling, and then flying to the next survey area. In the meantime, we consider two cases: 1) In the event of the UAV’s battery depletion, it has the capability to autonomously return to the UGV landing location to avoid a collision and undergo a battery replacement; and 2) the UAV can talk to the UGV to do sensing and actuation procedures by sending GPS coordinates and pre-processed survey information (by AI devices on the UAV), such as less or more carbon emissions. In the end, after finishing the survey mission, we should be able to get a carbon emissions map of the field.

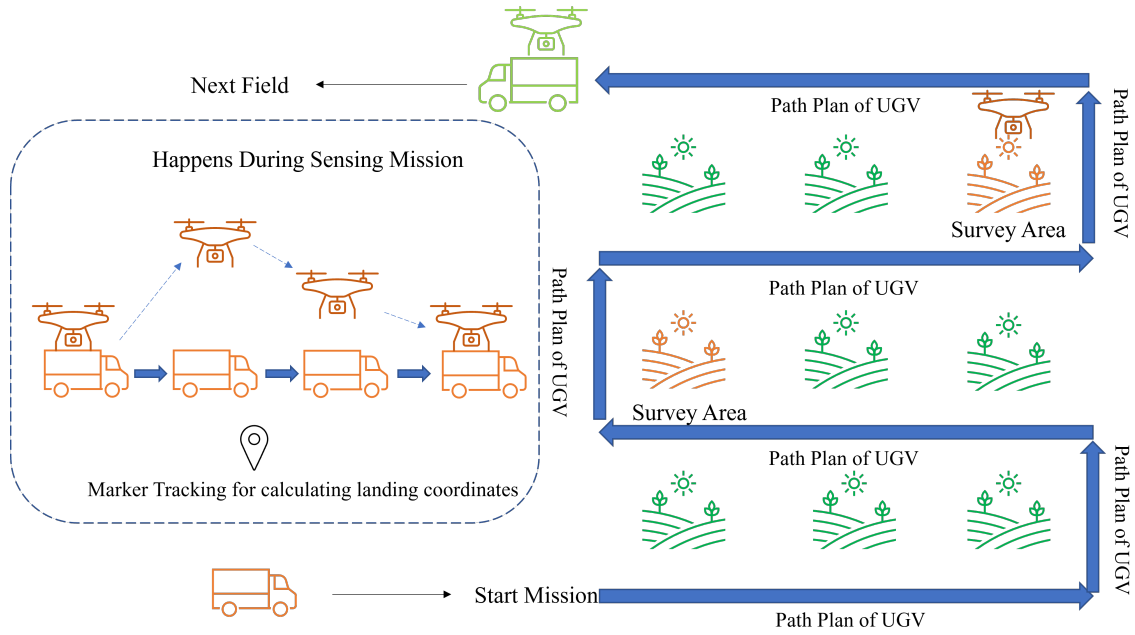


Figure 4.9: UAV and UGV for sensing carbon emissions diagram.

4.4.2.2 Site-specific Management Sensing System

Our sensing system mainly consists of UAV and UGV, as shown in Figure 4.10. We build a UAV from scratch by using a lightweight and low-cost TAROT 650 Carbon Fiber frame⁸. The 4 motors we used are T-Motor MN3508 KV380, which would give a maximum takeoff payload of 1.4 kg. Using a 9500 mAh (6S 1P) battery would support at least 40 minutes (no extra payload). In addition, we customized the co-pilot system installation structure by using the 3D printer (Figure 4.10 red words and points direction). The Raspberry Pi is the essential processor for running landing pad marker coordinate tracking, receiving GPS signals from UGV, and also being able to send UAV GPS location through Pixhawk's telemetry to UGV for both communications. The camera we used was the Raspberry Pi camera, facing downward to the ground. The UGV was retrofitted from a toy to a tool [164]. The UGV would be programmed by mission planner⁹ for survey path planning and waypoint setup that allows for autonomous driving to the needed survey area. The details are shown in Figure 4.9 and 4.10.

⁸<https://www.rc-wing.com/tarot-tl65b01.html>

⁹<https://ardupilot.org/planner/>

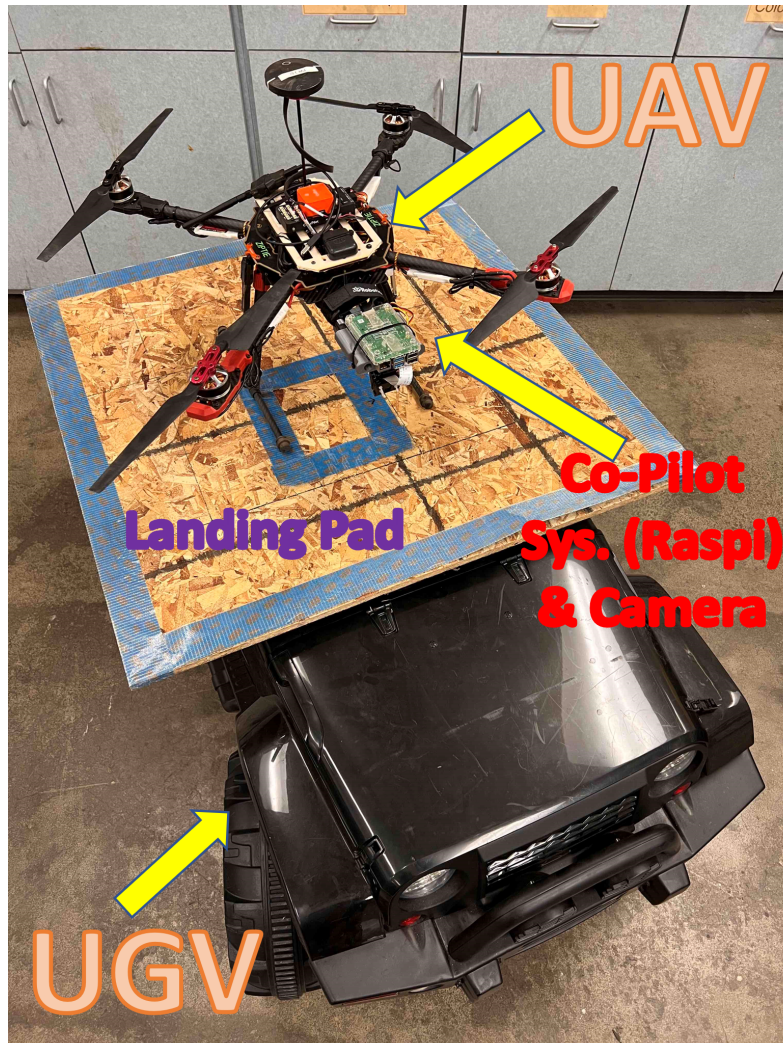


Figure 4.10: UAV and UGV sensing system.

4.4.2.3 Mobile Sensing Platform Payload

The payload of the carbon sensing mission mainly consists of sensors, biochar, a spray tank, and backup batteries, as shown in Figure 4.11. The sensors and microwave and millimeter-wave sensors would be equipped on the UAV and UGV. For more details, please see our previous work [135,264]. We would primarily equip the biochar with a UAV, enabling in-situ spraying after analyzing the carbon results. The spray tank and backup batteries are on the UGV. The spray tank can carry anything, such as water, chemicals, and biochar, to spray them at a specific location.



Figure 4.11: The payload options of the mobile platforms. The biochar, microwave radar, millimeter wave radar, battery, and spray tank are shown individually.

4.4.3 Approach Methods

The experiment setup and methods would be discussed in this section, including proximity sensor data collection, controller performance, mission simulation, and vision-based marker tracking.

We designed two fields for demonstrating entire experiments on the long-endurance SSM functionality of our system. In the first field, the UAV and UGV demonstrate the touch-and-go survey method for SSM biochar applications. In the second field, we zoomed in on the touch-and-go method, which involves the UAV landing on the moving UGV with a visual tracking algorithm target and a properly tuned PI controller. We ran the simulation on Matlab 2022b and conducted the data analytics on the same platform.

4.4.3.1 Proximity Sensor Data Collection Configuration

As shown in Figure 4.11, we deployed the microwave and millimeter wave radar sensors, respectively. The microwave sampling rate was set at 102.9 GHz, and the millimeter sampling rate was set at 100 kHz. We record the microwave and millimeter-wave radar data when the UAV approaches the desired survey area. The sampling repeats each sensor 20 times at the same location (technically, the UAV would have a small vibration due to wind or GPS error). The Raspberry Pi receives the data for processing, enabling the UAV to determine whether to spray biochar and

forward the data to the UGV. Please see [135] for information on carbon emissions from radar signal processing.

4.4.3.2 UAV Landing Process – A Simulation Approach

At the beginning of the simulated mission survey developed, the objective was for the UGV to navigate between two fields of size 10×10 m, separated by a distance of 1 m (see Figure 4.12). We achieved this by positioning waypoints at the extreme points of each field and directing the UGV to minimize the distance between itself and these waypoints. In the first half of the survey, we set up waypoints for the UAV to survey Field 1, which involved landing on the field, taking off, landing again within the same field, and then taking off again. The waypoints would be picked at either a point of interest or with equal spacing throughout the field. This method was to display a close encounter with the soil sample by conducting a touch-and-go survey, which creates a grid-style survey. After completing the survey in Field 1, the UAV proceeds to Field 2, utilizing a hover-sweep survey methodology instead of the touch-and-go approach. We once again introduced waypoints, but selected the inner corners of the field to facilitate a sweeping method for data collection. This hover-sweep represents collecting data samples from a more overgrown field where landing cannot be done. The waypoint navigation for both the UAV and UGV of the distance-driven velocity controller is handled using (4.3):

$$\mathbf{d}^{i,\theta} = \theta^{i,j} - \mathbf{X}_k^i, \quad (4.3)$$

where $\mathbf{d}^{i,\theta}$ denotes the distance between the i^{th} vehicle state and waypoint state $\theta^{i,j}$, where $j \in \mathcal{J} \in \mathbb{R}$. $\mathbf{X}_k^i$ denotes the i^{th} vehicle state which contains (x, y, z) position at timestep k and $\theta^{i,j}$ denotes the waypoint position j in 3D space also containing (x, y, z) position for the i^{th} vehicle. Once the distance between the waypoint and the vehicle is calculated, the velocity of i^{th} vehicle is calculated using (4.4):

$$\mathbf{v}^i = \frac{\mathbf{d}^{i,\theta}}{\|\mathbf{d}^{i,\theta}\|} V_{max}^i, \quad (4.4)$$

where $\mathbf{v}^i$ denotes the velocity of the i^{th} vehicle and V_{max}^i represents the max velocity of the i^{th} vehicle. In the case of the simulations, the maximum velocity for the UGV and UAV is 0.5 m/s and 1.5 m/s, respectively.

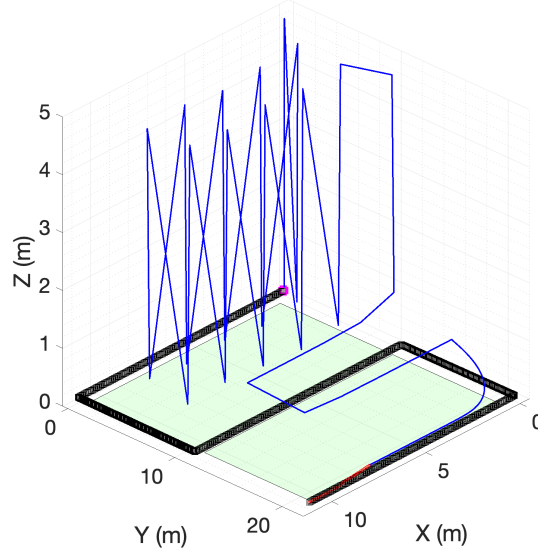


Figure 4.12: Full trajectory of the UAV and UGV survey missions of size 10×20 m where the starting point is denoted by a magenta square. The black squares denote the UGV trajectory, while the blue line represents the UAV trajectory. The landing sequence of the UAV onto the UGV is denoted by red circles.

Once the UAV completed the survey mission, the new objective was to land on the UGV while the UGV was moving. This process was done in two parts, first lining up with the UGV from 0.1 m above by matching UGV velocity and position, then landing on the UGV by reducing the UAV's vertical velocity (see Figure 4.13). The distance-driven velocity controllers mentioned previously in (4.3) and (4.4) are coupled to be used with a PI controller.

$$\mathbf{d}^{UAV,UGV} = k_p e + k_i \int_{t_1}^{t_2} e dt, \quad (4.5)$$

The error e is calculated using (4.3) and then passed into the PI controller to output $\mathbf{d}^{UAV,UGV}$, which denotes the distance between the UAV and UGV at a height of 0.1 m above the UGV. We then used this to calculate the UAV's velocity in the (x, y, z) directions, aiming to initially catch up with the UGV before landing.

4.4.3.3 Vision Process of Landing Target

Once the UAV matches its horizontal speed with the UGV from above while beginning in the frame, the landing sequence commences. There would be 5 arcuo markers on a flat surface of the UGV with the same orientation but positioned in

an X pattern. Four large markers are at the corners of the flat surface while a single smaller aruco marker is in the center. This was done to allow the Raspberry Pi 4 camera to see the large markers from a far distance during the landing descent stage. As the UAV reduces its altitude closer to the UGV, the larger markers would become omitted from the field of view (FOV) of the camera, leaving the smaller aruco marker left as the driving factor of localizing the landing point.

During the initial descent stage, the centroid of the marker pattern is calculated in (4.6):

$$\mathbf{C} = \frac{\sum_{q=0}^{N_m} \mathbf{x}_k^q}{N_m}, \quad (4.6)$$

where $\mathbf{C}$ is the measured (x, y, z) location of the centroid of the aruco markers with respect to the camera on the UAV. $\mathbf{x}_k^q$ represents the measured (x, y, z) location of the q^{th} aruco marker at time step k , where $q \in \mathcal{Q} \in \mathbb{R}$ and N_m denotes the number of markers used.

4.4.4 Preliminary Results

We conducted a series of experiments to evaluate our long-endurance site-specific management system. A simulation approach includes two controllers, one for controlling the landing process and another for tracking UGV movement.

4.4.4.1 Controller Performance Analysis

We evaluated our PI controller's performance based on the landing sequence. Figure 4.13 shows the 3D of the UAV landing on the moving UGV process. When the UAV approaches its last survey waypoint, it starts the landing sequence. We divided the PI control for the landing sequence into two stages: the transition from the survey to the UGV's above-ground landing pad, and the final landing on the UGV. During the first stage, shown in Figure 4.14, the y and z directions converge to 0, while the UAV's x position is catching up with the x position of the UGV. Once the UAV position is caught up with the UGV, the next stage is engaged (see Figure 4.14) of the landing sequence. The UAV then matched the horizontal velocity of the UGV while reducing vertical velocity to land on the UGV, as shown in Figure 4.13 and 4.14. The timestep size dt we set is 0.1 s, the UGV moves at 0.5 m/s, the maximum UAV speed is 1.5 m/s for survey missions, and the UAV's velocity is constrained to 0.6 m/s during the landing sequence. The overall controller performs well and is cost-effective, making it suitable for use on the UAV's edge devices.

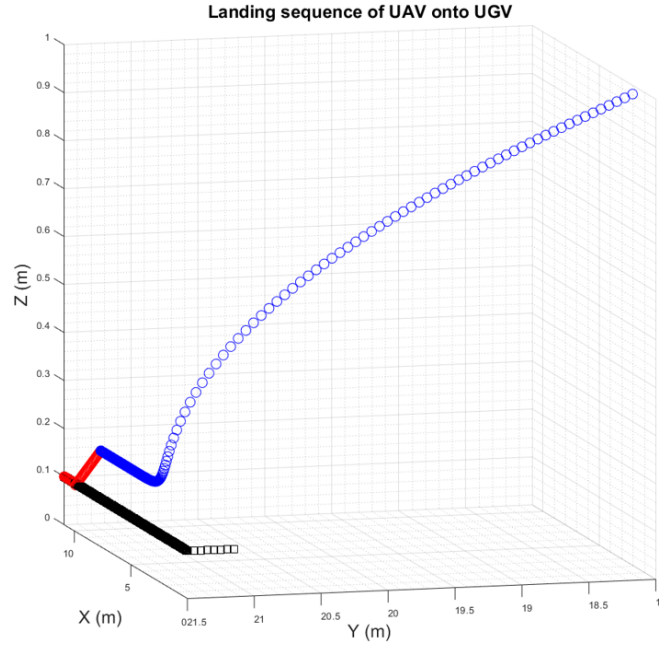


Figure 4.13: Landing sequence trajectory of a UAV and UGV survey mission. UGV trajectory is represented by black squares, while UAV point mass trajectory is represented by blue circles. The landing sequence of the UAV onto the UGV is denoted by red circles.

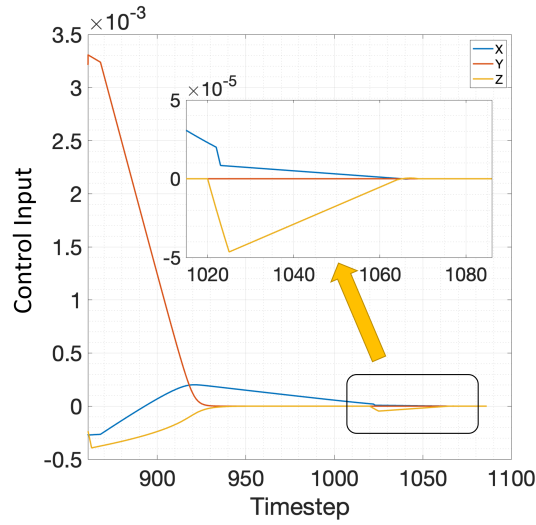


Figure 4.14: Controller performance for X , Y , and Z directions of the UAV towards landing on the UGV.

4.4.4.2 Tracking Moving Target Analysis

Testing the vision system for the landing sequence consisted of using OpenCV on a Raspberry Pi 4B with a Picamera. A board outfitted with 5 aruco markers was used to test the accuracy of the vision system's capability in determining the distance and position of the markers with respect to the camera.

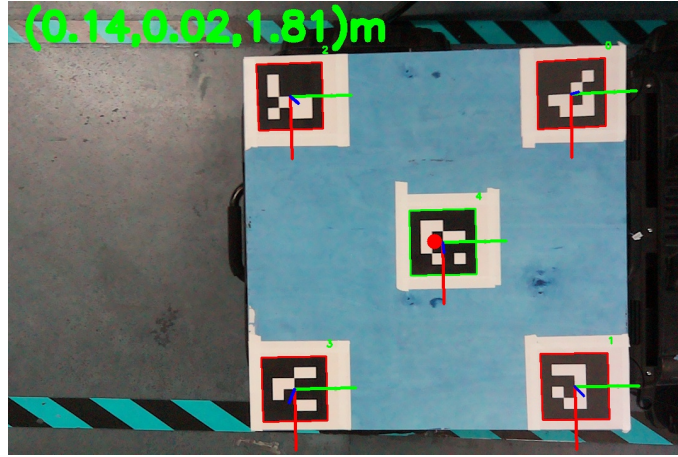


Figure 4.15: Testing a landing vision system with 5 aruco markers using the OpenCV Python library for detection. The coordinates denoted in the upper left corner represent the distance of the calculated centroid of all markers from the camera principle point. The red circle represents the calculated centroid of all aruco markers combined.

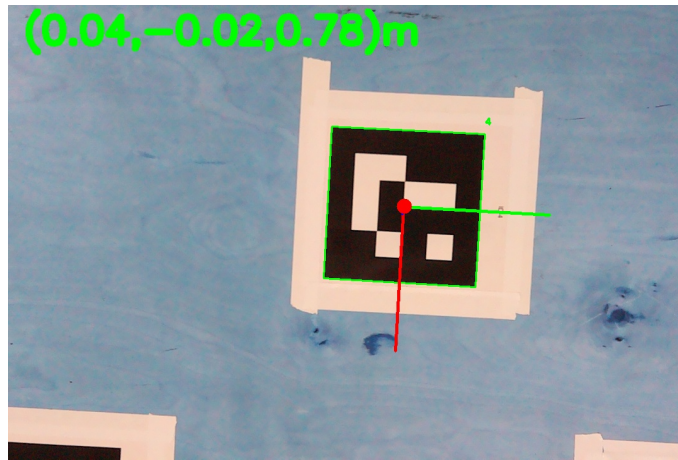


Figure 4.16: Testing the landing vision system with the camera close to the board observing a single aruco marker.

The frames captured while determining the accuracy of marker measurement are shown in Figures 4.15 and 4.16. For the far distance case (Figure 4.15), the measured distance the vision system calculated the markers to be was 1.81 m, while the true distance between the markers and camera was 1.78 m. The centroid of the group of markers on the board showed a slight shift from the center of the middle marker. While for the close distance case (Figure 4.16), the measured distance the vision system calculated the markers to be was 0.78 m, while the true distance between the markers and camera was 0.77 m.

4.5 Conclusions

This chapter begins by introducing the concept of SSM, a sophisticated system designed for biochar applications. The system leverages an array of technologies including UAVs, UGVs, and proximity sensors, which are integrated into various payload options to optimize the practice of SSM. Our design is aimed at enhancing the mission life and operational feasibility of carbon sequestration initiatives, centering on the sustained endurance of the mobile sensing platform. This is achieved through an innovative approach enabling UAV autonomous landing on a moving UGV, as verified through simulation of the landing sequence and vision-based marker tracking.

Building on the foundation of SSM, we delve into the DT concept, applying it to dairy farms for networked sensors and actuators within the realm of the IoT. The DT-enabled methane emission abatement framework that we present employs a UAV toolbox, ROS, and Gazebo, in conjunction with mobile sensors and actuators. We provide a detailed exposition of the DT’s core physical principles, demonstrating how mobile sensing and biochar spreading can significantly mitigate methane emissions. Preliminary results showcase the efficiency of UAVs in flight path management, swarming actuators, and in tracking methane diffusion processes, indicating DT’s substantial potential to enhance methane abatement.

Subsequently, we expand the scope of DT into the IoLT proposing a soil carbon accounting management system (SCMDT). This system integrates proximity radar sensors, AIoLT, and UAVs as mobile sensors and actuators. We outline the fundamental physical principles of SCMDT, alongside an evaluation metric for verifying soil carbon content accounting. The utilization of common machine learning algorithms for determining soil carbon content is discussed, positing SCMDT as a unifying framework poised to revolutionize soil carbon management.

These developments are exemplified in a case study on long-endurance site-specific management sensing and actuation systems for biochar applications. Our system offers multiple benefits over traditional carbon emission management strategies, such as non-intrusive sensing and suppression of carbon emissions, extended operational endurance through UAV-UGV interfacing, and cost-effectiveness in edge

computing. The system’s controller exhibits rapid response times, with vision marker tracking errors maintained below ± 0.02 m on average.

In the future, our efforts will pivot towards real-world system validation, with an emphasis on devising comprehensive evaluation metrics for biochar applications. To bolster the realism of our simulations, further exploration into ROS2 and Gazebo is needed. The development of AI-enabled edge devices for advanced radar signal processing is also on the horizon. We plan to expand our machine learning algorithms to establish an exhaustive site-specific management model, incorporating UAV and UGV sensing data to elevate sensing precision. Moreover, the incorporation of Explainable AI (XAI) will enhance decision-making processes, empowering agricultural and environmental experts to tackle insights from UAV and UGV data streams and other sensor inputs.

Chapter 5

EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI) EMPOWERED DIGITAL TWIN ON SOIL CARBON EMISSION MANAGEMENT USING PROXIMAL SENSING

5.1 Introduction

Digital Twin (DT) becomes a more and more important management tool in Industry 4.0 and has enabled many applications. DT is a digital replica of a physical entity, system, or process that enables real-time monitoring, analysis, and optimization. By creating a high-fidelity virtual model that accurately mirrors the real world, it is updated from real-time data and uses simulation, machine learning, and domain knowledge to help decision-making accurately [265]. DT enables many applications of precision agriculture in that it represents certain land areas or ecosystems for managing carbon emissions. DT would take into account various parameters such as the type of soil, vegetation, climate, human interaction, and many more. With the help of sensors and satellite Imagery, real-time data can be collected and fed into the digital model. However, the models used in DT are usually trained by machine learning or deep learning algorithms, which are more or less black boxes, which means there is less transparency and interoperability to help DT make reasonable decisions. Thus, we decided to involve explainable artificial intelligence (XAI) in DT to better manage soil carbon emissions.

Explainable AI is a subfield of AI that focuses on creating models that provide clear, understandable explanations for their decisions and actions [266]. XAI can address some of the shortcomings of DT for carbon emission management. With XAI, the decision-making process of the DT becomes more transparent and understandable. This is particularly useful when the DT is used for complex scenarios where it's crucial to understand why a particular prediction or decision has been made.

Prior research on DT applications for carbon emissions management has focused on the DT framework proposal. In [242] was proposed DT for methane emissions using mobile sensing and mobile actuation on dairy farms. The Internet of Living Things (IoLT) has been investigated and validated for soil carbon emission, and DT is suitable for managing soil carbon emission as a core platform [267]. Papers

above systematically illustrated how to apply DT to carbon emissions management, including gas and solid carbon emissions, which is the holistic idea of how to utilize DT and its benefits. Recent works using DT for precision agriculture and soil condition diagnosis are explained in [268,269]. These works discussed applications and challenges in agriculture by combining data science. Several works, such as [270] applied XAI on DT for managing the versatile corn production ecosystem, while [271,272] discussed potential applications of using DT on agriculture activities. However, these works did not explain how to enhance the DT's performance on prediction or decision-making in carbon emission management and did not even introduce potential functions if XAI is involved.

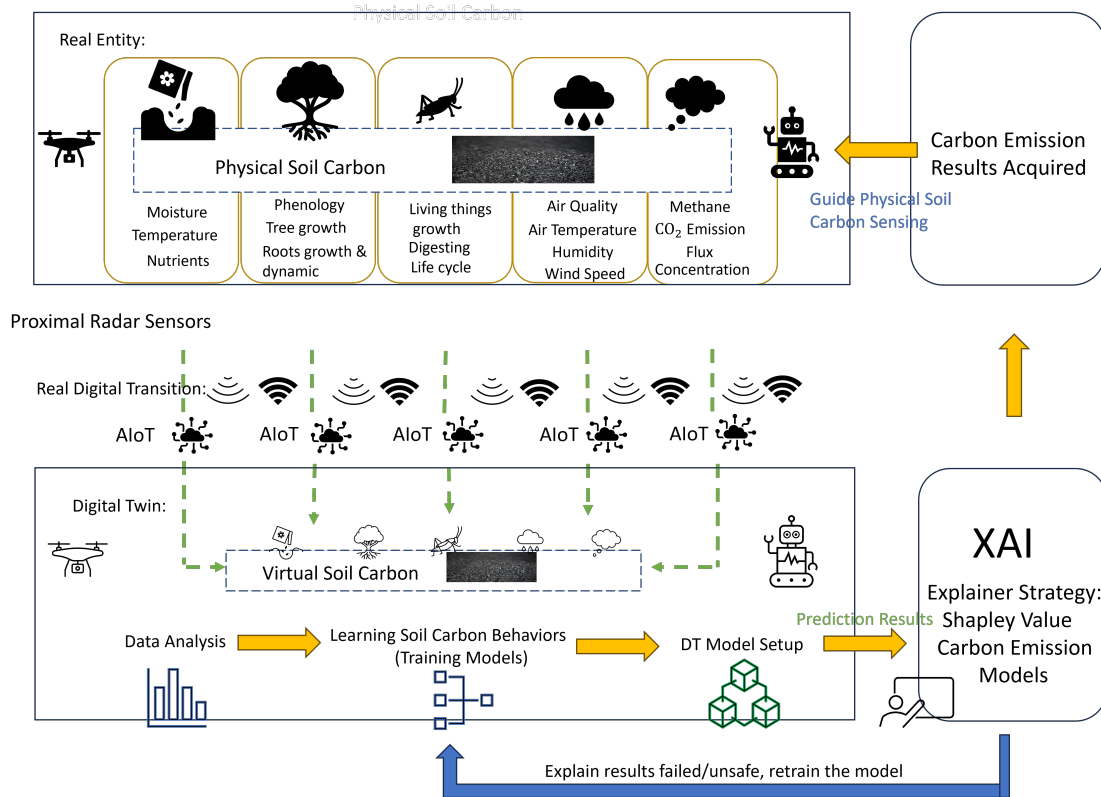


Figure 5.1: XAIoT empowered digital twin in soil carbon emissions quantification.

To address these shortcomings, we introduced an XAIoT-empowered DT framework for managing soil carbon emissions with advanced proximal sensing (AI can be applied to the collection data section at local devices (AloT) that justify whether data is reliable or not). In addition, we also explored mobile sensing and mobile actuation in [242,267], and the proximal sensors can be attached to these mobile platforms for acquiring fine-grained information. In summary, the contributions of this chapter are as follows:

- We present an XAIoT-DT that contains explainable artificial intelligence components to explain, and interpret the DT prediction results and improve them.
- We validated our XAIoT-DT framework based on real-world soil carbon content datasets to demonstrate the feasibility, challenges, and opportunities of our XAIoT-DT.

5.2 XAIoT empowered Digital Twin in Soil Carbon Emissions Management

Industry 5.0 seeks to reintroduce human, environmental, and social elements [273]. Therefore, we inherit DT from Industry 4.0 and add XAIoT to DT to better solve precision agriculture and environmental challenges. We will discuss the details of the XAIoT-DT framework for soil carbon emissions management in this section. The entire framework is shown in Figure 5.1, and the definition of XAIoT-DT is shown in Section 5.2.1. We will illustrate XAI components in the next section.

5.2.1 The XAIoT-DT Framework

As shown in Figure 5.1, we conducted physical twin data collection for soil carbon emissions activities such as moisture, temperature, nutrients, tree growth, living things activities, air quality, GHGs, etc. Each of these variables is monitored, including but not limited to the proximal sensors. In our previous work, we applied AI to edge devices to process, label, and verify the collected data in advance. Next, the preprocessed data would feed into DT. To clearly address DT, we simply revealed the all-data relationship in the following:

$$f(S_{XAIoT-DT}) = \{S_T, S_M, S_P, S_{AD}, S_F, \dots\}, \quad (5.1)$$

where $f(S_{XAIoT-DT})$ is the virtual soil carbon emissions activity generated by DT. $S_T, S_M, S_P, S_{AD}, S_F$ are temperature, moisture, phenology, air quality, flux, and so on, respectively.

In the DT process, DT will do data analysis and learn soil carbon behaviors, which means training the model by feeding it real-world data. After that, DT would generate the prediction results for the next level of carbon emissions, and we would add the XAI component to judge whether the results are reasonable and reliable. If it gives unreliable results, we will go back to the training stage to retrain the model by adding XAI feedback. After that, we acquire the carbon emissions results, and these results can guide the physical soil carbon sensing, quantification, mitigation, and so on in agriculture activities.

5.2.2 XAI Explainer Strategy–Shapley Index

Shapley value was first proposed by Lloyd Shapley in Game Theory [274], and it is a fair allocation of the total gain generated by a cooperative game to individual participants. The Shapley value of a feature for a specified query point explains the contribution of the feature to a prediction (the response for regression or the score for each class for classification) at the specified query point. The Shapley value represents the deviation between the prediction for the query point and the average prediction caused by the feature. For each query point, the total deviation of the prediction from the mean is equal to the sum of the Shapley values for all features [275, 276].

The Shapley value of the i^{th} feature for the query point x is defined by the value function f :

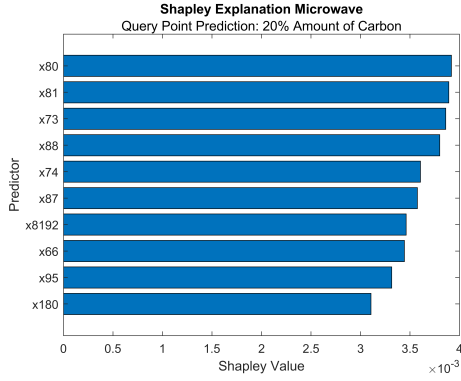
$$\phi_i(x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f_{S \cup \{i\}}(x) - f_S(x)] \quad (5.2)$$

where $\phi_i(x)$ is the Shapley value of the i^{th} feature for the query point x , S is a subset of features not containing the i^{th} feature, N is the set of all features, and the sum is over all subsets S of N that do not contain the i^{th} feature. $|S|$ is the number of features in subset S , f_S is the prediction model trained only on the features in subset S , $f_{S \cup \{i\}}$ is the prediction model trained on the features in subset S and the i^{th} feature.

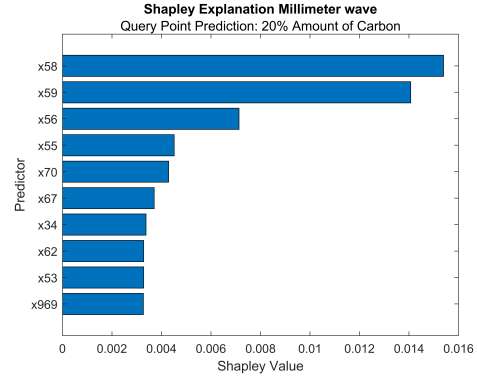
In this paper, to implement the Shapley algorithm, we used intervention algorithms [275] that compute the expected contribution by using artificial samples created from the specified data x . The results are presented in Section 5.3.

5.3 Preliminary Results

To validate our XAIoT-DT and especially XAI components, we mainly conducted an explanation effort of signal features extracted from proximal sensor signals, which are microwave and millimeter wave signals. The datasets we used are already published [130], and they were collected from the soil carbon content by using microwave [135] and millimeter wave [264]. The microwave signal contains 8192 features, and the millimeter wave signal contains 1024 features. Here we selected the amount of soil carbon content that contains 5 classes and the amount of soil moisture that contains 6 classes for the classification mission. At this time, we will mainly focus on the XAIoT-DT performance, while the details regarding data processing workflow can be found on our previous work [135, 264, 267].

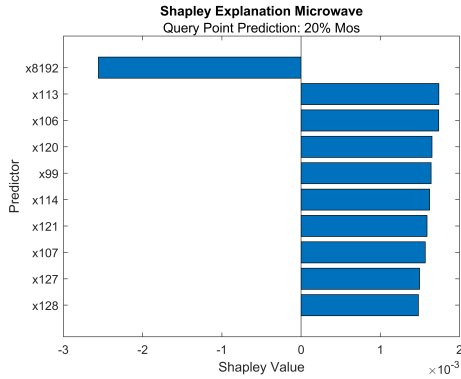


(a) The Shapley value explanation of the microwave signals for sensing the amount of soil carbon.

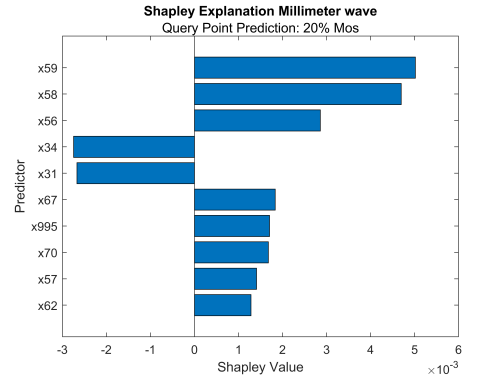


(b) The Shapley value explanation of the millimeter wave signals for sensing the amount of soil carbon.

Figure 5.2: The Shapley value distributions of the microwave and millimeter wave signals that come from 20% of the soil and carbon mixture.



(a) The Shapley value explanation of the microwave signals for sensing soil moisture.



(b) The Shapley value explanation of the millimeter wave signals for sensing soil moisture.

Figure 5.3: The Shapley value distributions of the microwave and millimeter wave signals that come from 20% soil moisture.

As shown in Figures 5.2 and 5.3, we demonstrated the first 10 highly correlated features extracted from the signals that affect the prediction results. Positive values mean positive correlations, and negative values mean negative correlation or bad contributions to the results. We checked the prediction results for the first class (20% of the soil and carbon mixture and 20% soil moisture). The correlated signal distribution is from the first few milliseconds to the last few milliseconds. This is

an important observation: we can train our models based on these highly correlated features to reduce costs. In addition, we trained the model using the SVM classifier by only feeding these 10 features into the model; the classification accuracy is 90.6% in microwave and 95.2% in millimeter wave, respectively, which are close to previous results.

5.4 Discussion and Future Works

In this work, we propose an XAIoT-DT framework for better monitoring and managing soil carbon emissions so that we can mitigate climate change. The priority should be data collection and sensor calibration. Users should ensure that the sensors used for data collection are regularly calibrated and maintained to provide accurate and reliable data and also consider expanding the range of variables monitored in (5.1) if new research indicates other factors significantly impact soil carbon emissions. Moreover, the iterative process of model training and XAI feedback is crucial for improving the accuracy of predictions. We would like to consider implementing advanced machine learning techniques rather than classifiers, such as reinforcement learning, which could potentially improve the efficiency of this iterative process.

In the future, we will keep investigating how the XAIoT-DT framework could be integrated with other systems, such as weather forecasting systems or satellite imagery, to provide additional data that could improve the accuracy of soil carbon emission predictions. User-friendly interfaces and data visualization tools would also be developed to help end-users, such as farmers or land managers. Finally, long-term studies to assess the impact of the XAIoT-DT framework on soil carbon emissions and overall environmental health also need to be conducted and validated. This could provide valuable data for refining the framework and demonstrating its effectiveness.

5.5 A Case Study: A Miniature Millimeter-Wave Radar Based Contactless Lithium Polymer Battery Capacity Sensing with XAI

5.5.1 Introduction

In recent years, the use of drones and UGVs for precision agriculture applications has become a significant area of research, while the use of rechargeable batteries in electrical systems, such as portable devices, etc. [277], has become more widespread. As a significant electrochemical energy storage solution for drones and electric vehicles (EV), lithium-ion batteries are among the most commonly used rechargeable batteries. For numerous applications, Li-Po (Lithium Polymer) batteries have been chosen due to their exceptional advantages, including critical high single-cell voltage, critical high discharging rate (often several times greater than the charging current rate), critical long service life, and critical wide typically operating temperature range. However, it is still a challenging problem to

accurately sense SOC and estimate how much capacity left for safety critical mission in use of the battery.

To tackle the challenges, several researchers and industry companies proposed battery management systems (BMS) for EVs and drones [278]. It examines battery performance indicators such as SOC, SOH, power state, and battery discharging rate condition [279–281]. Modeling with parameter identification, and additional studies on optimal charging design [278] are the conventional techniques for Li-Po. Two most commonly used models for Li-Po are the equivalent circuit model (ECM) [282] and the electrochemical model [283]. The ECM is an impedance-based model that focuses primarily on electrical properties. In recent years, ECM has been expanded by including temperature equations or other chemical reaction equations [283], and the single particle model (SPM) and its extended forms have gotten a lot of attention in recent research [284]. Some research has proposed fractional order models (FOM) for Li-ion batteries to improve the performance and accuracy of BMS, such as fractional circuit impedance models [285] and fractional electrochemical models [286]. During charging and discharging, these fractional models can better explain the nonlinear, stochastic electrochemical behavior of Li-Po.

Most BMS systems are built according to the simulation model to estimate the battery life. We all know that the life time of a battery will reduce over charging cycles. In other words, estimating the capacity is not accurate in battery use, which is not reliable to provide safety instructions. There is only one work that tests drone batteries using a contactless method (Microwave sensing [287]). However, the sensing accuracy is not consistent, and the models cannot run on edge devices (such as the Raspberry Pi, which can easily be mounted on drones or UGVs). Other research [105, 288–291] and [135] used mm-waves and microwave for materials classification and soil carbon sensing, but not for battery monitoring. So, here we apply millimeter-wave radar sensor for battery capacity monitoring system, which has two innovations:

- Compared with the traditional method and microwave of sensing battery capacity, the mm-wave radar sensing method is less labor-intensive and more accurate compared with existing contactless solutions.
- A portable mobile device is easily equipped on an autonomous drone or vehicle to detect SOC and SOH in a cost-effective way without causing a negative impact on the battery.

As a result, we target two challenges: 1) How to convert a mm-waves signal indicating battery capacity, 2) how to use these unique mm-wave signal characteristics to predict battery life. The contributions of this work are as follows:

1. To the best of our knowledge, this is the first time work to show that mm-wave radar can detect Li-Po battery capacity on site.

2. We revealed a potential relationship between features from mm-wave reflectance and battery capacity and voltage based on physical knowledge.
3. We validated our system by conducting real-world experiments and machine learning algorithms.

5.5.2 Materials and Methods

5.5.2.1 Sensing Hardware

The ImageVK-74 mm-wave sensor is from Vayyar¹ and was designed and manufactured by Mini-circuits². Please see more details in Chapter 3

The drone battery we used is 6S 1P (6 cells in series connection, no parallel connection) with fully charged voltage 25.1 V, and the capacity is 9500 mAh (as shown in Figure 5.4 left). A battery contains an anode, a cathode, a separator, an electrolyte, and two current collectors (positive and negative). Both the anode and the cathode contain lithium. Through the separator, the electrolyte transports positively charged lithium ions from the anode to the cathode and vice versa. The mobility of lithium ions generates free electrons in the anode, which in turn produces a charge in the positive current collector. The current then flows from the positive current collector through the load device (cell phone, computer, etc.) to the negative current collector. The separator prevents electrons from flowing through the battery.

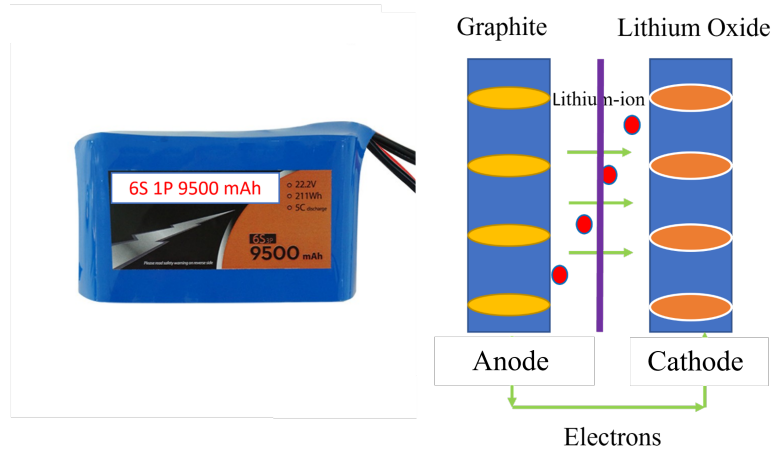


Figure 5.4: Drone battery 9500 mAh with high discharge rate.

While the battery is draining and supplying an electric current, the anode discharges lithium ions to the cathode, resulting in an electron flow from one side

¹<https://www.vayyar.com/>

²<https://www.minicircuits.com/>

to the other. When charging the battery, the converse occurs: the cathode releases lithium ions and the anode receives them (See Figure 5.4 right).

Since we understand the charging and discharging process of the Li-Po battery: the variation of terminal density during charging and discharging would be captured by mm-waves radar. This physical principle inspired us to believe that battery capacity monitoring through mm-waves is feasible.

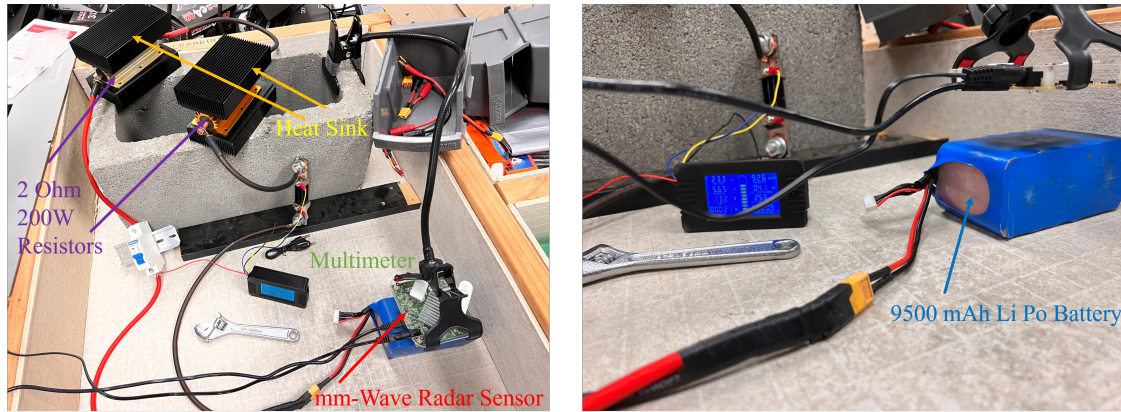


Figure 5.5: Load circuit for battery capacity draining along with mm-wave radar sensor sampling hardware setup.

5.5.2.2 Millimeter Wave Signal Classification Methods

In this section, we proposed using machine learning algorithms to process the sample of the mm-wave signals that categorize the battery capacity and voltage. We select PCA to reduce high dimensions and use linear SVM to do classification. For the details of linear SVM, see Chapters 2 and 3.

To solve the multi-class classification problem, N binary SVM classifiers are created in the case of N -class issues ($N > 2$). The i^{th} SVM is trained so that samples in the i^{th} class are labeled as positive and all other samples are labeled as negative. In the classification process, a test sample is obtained from all N SVMs and is labeled based on the classifier with the highest output [292].

In addition, we used the Classification Learner App from Matlab³ to implement the classification mission. The Classification Learner is used to train classifier models using many popular models such as decision trees, discriminant analysis, support vector machines, logistic regression, closest neighbors, naive Bayes, kernel approximation, ensembles, and neural networks. Moreover, 10-fold cross-validation is used to tweak hyper-parameters and validate the performance of all models on

³<https://www.mathworks.com/help/stats/classificationlearner-app.html>

our soil carbon samples. The accuracy and confusion matrices are used to sense how well the objects have been categorized.

5.5.3 Experiments

5.5.3.1 Experiments Setup

To further capture the variation of the battery capacity during charging and discharging, we need to use an equivalent discharging circuit to simulate the battery used for drone power consumption. The objectives are to test the batteries, including: 1) real battery capacities, 2) obtaining the real cutoff voltage. The load circuit we deployed is shown in Figure 5.5. We connect two 2Ω resistors in series (total external resistance is 4.1Ω). The fully charged battery voltage is 25.1 V (Open Circuit Voltage). Therefore, this load circuit would generate around 5 to 6 A of current for fast discharging. Moreover, the mm-waves sensor is placed above the battery at less than 2 cm height to keep sampling at every 1000 mAh of consumption by monitoring the loading circuit, such as the power cost.

The sampling frequency of the mm-wave sensor is from 62 GHz to 69 GHz. We sampled each sample 20 times for the purpose of reducing likelihood error. We set up 150 sampling points during the frequency band we mentioned, and the RBW is 100 KHz for reasonable sweep time and lower noise. To evaluate our system's performance fairly, we introduced a confusion matrix to calculate average classification accuracy. The 10-fold cross-validation is also employed to adjust hyper-parameters and validate all models' performance on our battery capacity samples. The accuracy and confusion matrix are presented to evaluate the classification objects' performance.

5.5.3.2 Evaluation

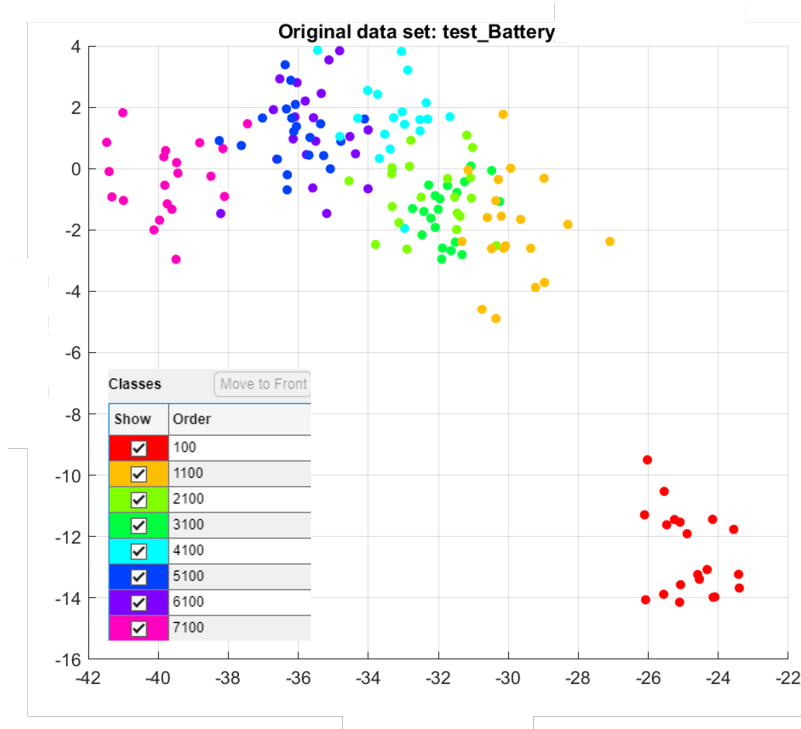


Figure 5.6: Original mm-wave data from varying battery capacity.

As mentioned earlier, we used PCA and linear SVM to reduce the data dimension and perform multi-class classification, respectively. The original data scatter plot is shown in Figure 5.6. Totally, it has 8 classes from 100 mAh to 7100 mAh. The scatters of each class is pretty independent of each other.

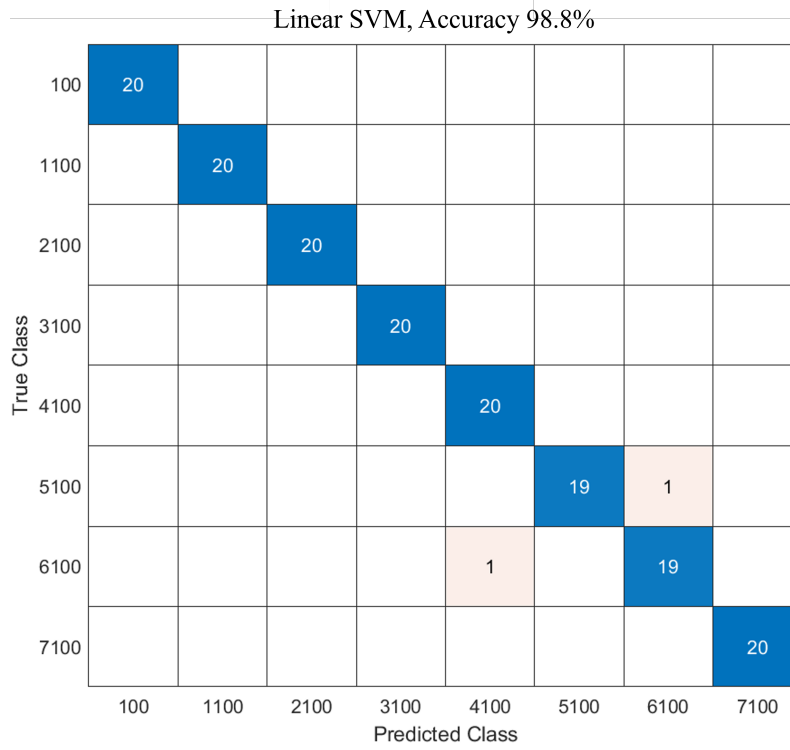


Figure 5.7: Confusion matrix of SOC classification of Li-Po battery

The accuracy of categorized capacity is 98.8% from the confusion matrix after 10-fold cross-validation (See Figure 5.8). In [287], their classification accuracy of voltage is 93.06% by using the microwave radar, which is lower than using mm-waves radar.

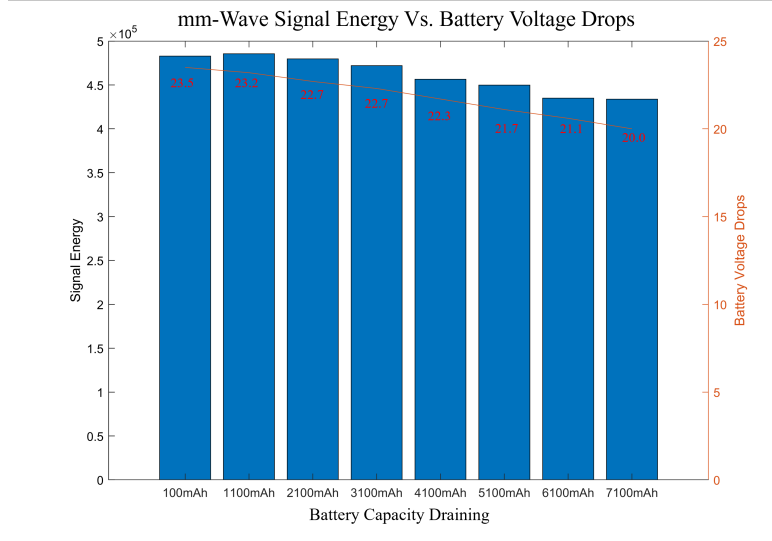


Figure 5.8: Comparison of mm-wave signal energy and battery voltage drops

In addition, we also investigated the relationship between mm-wave signal energy and battery voltage drops in Figure 5.8. As shown in the figure, the mm-wave signal energy decreases slightly as capacity decreases and the voltage drops from 23.5 V to 20.0 V. As we know, the voltage drops of Li-Po batteries are non-linear, and we need to further explore the relationship of mm-wave signal energy with the capacity to predict the battery health, which may help the earlier battery alarm to whether to trigger emergency landing of the drone or not.

5.5.4 Discussion and Future Work

In this study, we proposed a contactless method of measuring battery capacity by identifying the reflectance signals of mm-wave radar. The accuracy of measuring capacity is higher than microwaves radar. However, there are still some issues that need to be discussed.

The sensing distance affects the sensing accuracy. We tried many times to place the mm-wave radar sensor on top of the battery, and we discovered that if we place the sensor more than 3 or 4 cm above from the battery, all of the scatter points are the same (See Figure 5.9). To achieve the most accurate results, we will design more experiments to validate the sensing accuracy affected by the placement distance.

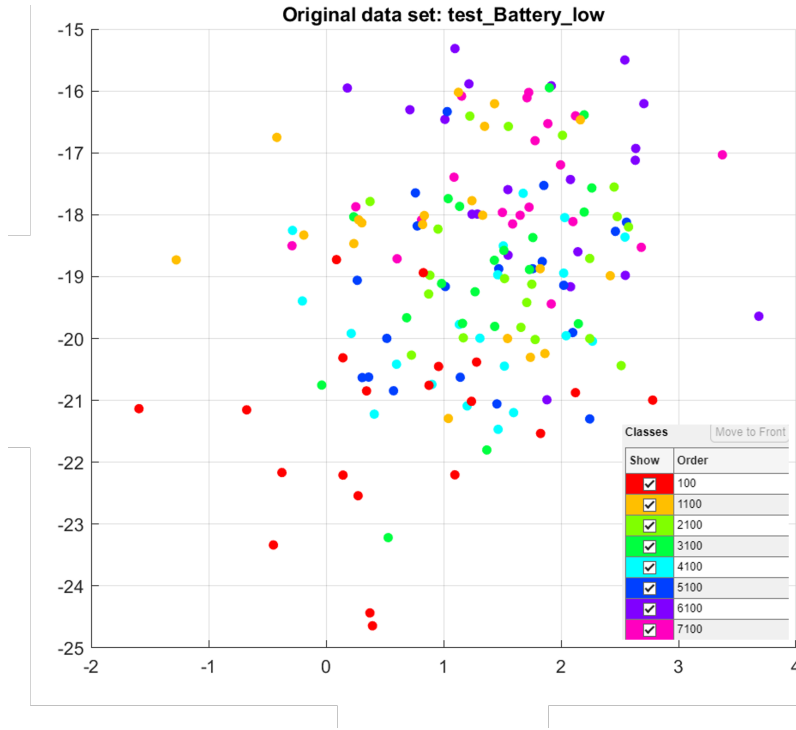


Figure 5.9: Failure case of using mm-wave sensing.

In addition, when testing battery capacity, the battery should not be moved for at least 10 minutes for stable electrolyte density. In Figure 5.10, we tested a fully charged battery six times every 10 min. As we can see, the signals of first time sensing are far away from the rest of the tests. These unique features should be paid attention to when measuring battery activities such as SOC, SOH, and voltage and capacity left. Moreover, this method has also been validated by different batteries (same capacity, same connection, same materials), and they also achieved the same results.

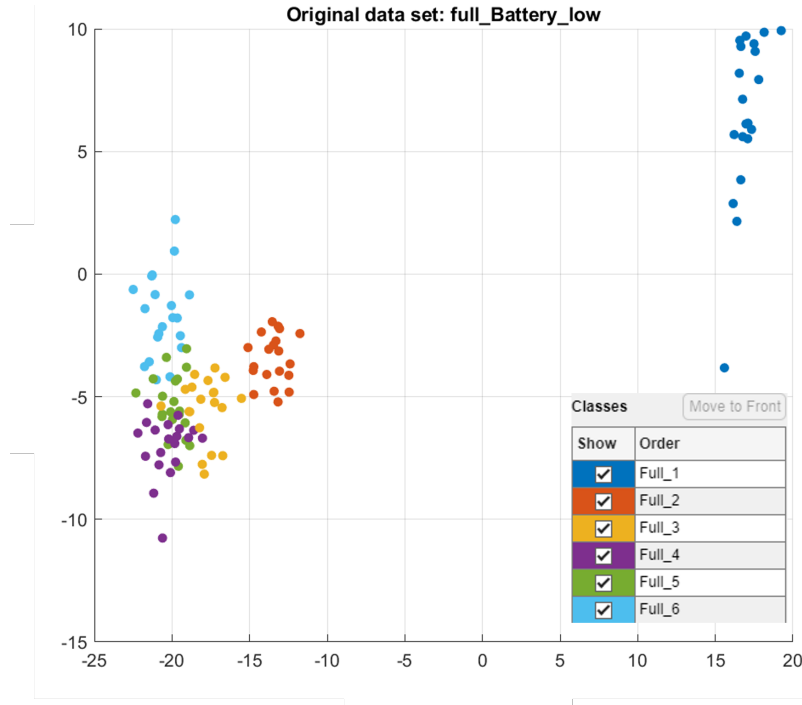


Figure 5.10: A failure case of using mm-wave radar sensing.

Another consideration is that the ImageVK-74 has a total of 40 antennas for transmitting and receiving signals, where each receiver would receive all the transmitters generating signals, also referred to as a sensor array. The advantage of the sensor array is that it provides new dimensions to the observation, allowing more parameters to be estimated with improved estimation performance. An array of radio antenna elements used for beamforming, for instance, can raise antenna gain in the direction of the signal while decreasing antenna gain in other directions, thereby boosting the SNR by amplifying the signal coherently⁴. Because our objective here is materials with varying Li-ion density, using a sensor array to sense the electrolyte density and Li-ion distribution would give reliable battery health awareness for all electrical systems, including EV safety control.

5.6 Chapter Summary

In this chapter, we presented an XAIoT-DT framework for managing soil carbon emissions. DT can help understand the intricate dynamics of these emissions and manage them more effectively. Combining DT with XAI by introducing Shapley value as the preliminary explainable method for addressing soil carbon content

⁴https://en.wikipedia.org/wiki/Sensor_array

sensing results can greatly enhance users' ability to understand, monitor, and manage soil carbon emissions, helping us address the pressing issue of climate change more effectively. In addition, we also extended the XAI concept to XAIoT, which enables many potential smart applications that not only work for the precision agriculture sector but also work in many other environmental, urban, and infrastructure construction scenarios.

In addition, the cognitive battery management system and a contactless method of measuring Li-Po battery capacity and SOC are presented. We employed a mm-wave radar sensor, designed a load emulation circuit, and machine learning algorithms (XAI) for building groundtruth. After conducting a series of experiments, we found that our system has a higher sensing accuracy (98.8%) than using a microwave radar sensor.

In the future, this mm-wave radar sensor can be placed on the drone to get the on-site battery capacity monitoring to form a battery health awareness system to prevent crash damage. Moreover, the proposed cognitive BMS framework needs to be explored further to estimate the reliable capacity by introducing working temperature, future mission requirements, etc.

Chapter 6

SOIL METHANE EMISSION SUPPRESSION CONTROL USING AUTOMATIC DRONE SWARM APPLICATION OF BIOCHAR MULCH - A SIMULATION STUDY

6.1 Introduction

Methane (CH_4) is a GHG that is getting more and more attention as people try to curb global warming. Natural methane can be produced extensively by bacteria [293] that live in the soil, water, or ruminant animals (the stomach of cattle has four compartments that ferment food as a significant part of the digestion process.). As a GHG, methane traps infrared radiation in the atmosphere and contributes to global temperature warming and thus climate change. Industrial and agricultural methane can also be regarded as the second most abundant GHG made by humans, which ultimately comes from fossil fuel combustion [294], dairy farms, landfills, etc. Once it is emitted, methane will exist in the atmosphere for at least 12 years [295]. As a result, the global climate is getting warm enough to potentially release some unknown ancient bacteria/virus [227], which existed in permafrost. Keeping away from these unknown viruses by curbing global warming is becoming a more common sense. Besides, methane is extremely flammable and could result in an explosion, causing damage to people or industrial infrastructures, raising unexpected financial costs [296]. It seems that reducing methane emissions into the atmosphere is the best way to inhibit global warming potential. However, the main research challenge is how to curb methane emissions in the soil by where, when, and how much to apply biochar mulch from a team of UAVs. The research questions include: what's the best way of measuring time-varying methane concentration (in a spatial sense) and how to control methane emissions under wind field disturbances effectively. These two main research challenges are the keys to curbing methane emissions in this application.

Prior research into methane detection (or remote sensing in general) has led to many developments in sensing and actuation approaches, such as methane emission maps provided by satellite [180], [297], aircraft, or sensors on the tower to visualize the methane distribution. However, these remote sensing methods are susceptible to space weather constraints [148]. Bad weather can delay or change methane information, which makes it harder to take steps to control the emission source (like

methane emissions). Another popular method is regional monitoring, such as placing multiple fixed-location sensors on the ground to simultaneously acquire methane information. However, the installation, maintenance, and calibration requirements of these methods (e.g., [228]) would significantly increase cost and complexity. More recently, companies such as SenseAir or Figaro have developed small, cheap, and low-power GHG sensors that can be used in regional monitoring. Although these sensors are calibrated to measure much higher quantities than what may be found in natural emissions, several works [298–300] have made good progress in exploring the ability of these sensors to measure at the natural emission levels. For this work, we assume that sensors like these can be used for methane sensing simulation purpose.

In this chapter, we present a soil methane emissions suppression approach using swarms of UAVs by spreading biochar mulch on top of the detected methane emissions area/source. Microorganisms in the soil can produce methane and release it into the atmosphere, causing climate change, such as global warming. However, people lack methods to manage soil methane emissions, especially the quantification of methane emissions from the soil. Current methods for measurement and suppression of methane (e.g., remote sensing using satellites, regional sensing with multiple embedded sensors, etc.) are often limited due to the maintenance, installation, and calibration requirements. Besides, none of these methods put forward how to suppress methane emissions. To overcome these drawbacks, we present a new method called FADE-MAS2D (Fractional Advection Diffusion Mobile Actuator and Sensor) in which swarming UAVs are applied as an optimal coverage control to various methane release scenarios (from single to multi-source disturbances) utilizing an anomalous diffusion model with different time and space fractional orders subject to wind fields. This strategy is based on the premise that methane diffusion can be modeled as an anomalous diffusion equation to which swarming UAVs can be applied to tackle the optimal coverage control issue. To simulate methane diffusion under the wind, we utilize fractional calculus to solve the anomalous diffusion equation and define wind force with the drag equation. However, this approach involves two challenges. First, biochar mulch would also be affected by the wind when spreading it. Second, due to the continuous influence of wind, swarming UAVs have lots of fluctuations when flying to the desired location. These issues may lead to a large-scale system failure. To address the wind challenge, we propose a PI controller to reduce the suppressing methane emissions’ compensation effects. To determine the optimal PI controller coefficients, we must incorporate emissions control, UAV control efforts, and UAV position error into the cost function. We evaluated our approach using simulation experiments with methane diffusion and multiple methane emission sources in the time and space domains, respectively. The results show that when $\alpha = 0.8$ and $\beta = 1.8$, the emission dynamics of methane look realistic. Furthermore, our approach resulted in great control performance with multiple methane emission sources and different wind velocities and directions.

As a result, we designed our FADE-MAS2D system with easily accessible materials and comprehensive algorithms that provide two advantages (innovations):

- The wind field was added to the fractional advection diffusion model for the first time to simulate the effect of area source diffusion under the wind disturbance. Drone flights and the spray materials (biochar) are also assumed to be affected by wind fields.
- Consider the total emissions area, UAVs control efforts, and actuators position error as guidance for finding the optimal controller performance.

More specifically, the main research challenges are: first, the traditional anomalous diffusion model [301] can only describe a stationary diffusion process, which means when the methane source is moving (subject to wind), it fails to demonstrate the methane diffusion, and UAVs would not know where to apply the biochar.

Second, the optimal UAV coverage control strategy can deal with disturbances and delays. The UAVs need to move towards the methane diffusion source given the UAV's optimal distribution location based on measured methane data from the regional sensors, and then apply the biochar mulch to suppress emissions. Third, in the real world, the most common situation is that methane diffusion would be affected by wind and turbulence. Therefore, since we already have a mathematical model for anomalous diffusion and want to see how multiple UAVs correspond to the methane diffusion subject to wind, we believe that adding a wind field to our anomalous diffusion model can be a substantial computational challenge. To overcome this, we only look at the advection-diffusion equation on a larger scale.

To address the modeling challenges, we proposed a method that applies fractional-order calculus to describe the methane diffusion in the time domain and space domain, respectively, which corresponds to our first challenge. Fractional order calculus has non-local and hereditary properties, making fractional differential equations powerful to model the methane diffusion process. We also applied CVT proposed by Qiang Du et al. [234], which serves as a multiple UAVs distribution strategy to command our UAVs to spread biochar to accommodate the methane diffusion area. We then added the advection term to the anomalous diffusion model [301] to accommodate the wind field conditions. The modeling is validated by using Simulink in the time and space domain with UAVs. The main contributions are summarized as follows:

1. Proposed a practical biochar spreading application algorithm based on CVT using UAVs swarms as networked moving actuators or spreaders.
2. In FO-DIFF-MAS2D [302] simulation platform, multiple methane emission sources can be considered to better characterize the real methane emission dynamics in the region of interest.

3. We added wind in FO-DIFF-MAS2D to create a fractional advection diffusion mobile actuation and sensing (FADE-MAS2D) model which is more realistic.

6.2 Problem Statement

Suppose that anomalous diffusion processes occur in a convex polytope $\Omega : \Omega \subset \mathbb{R}^2$. $\mathbb{R}$ is the set of all real numbers. Denote $\mathbb{R}_+$ the set of all positive real numbers. The time-domain $t \geq 0$ or $t \in \mathbb{R}_+ \cup \{0\}$. Let $\rho(x, y, t) : \Omega \times \mathbb{R}_+ \cup \{0\} \rightarrow \mathbb{R}_+ \cup \{0\}$ represent the methane concentration. The corresponding spatial-temporal dynamic process can be modeled by using the following advection-diffusion equation is:

$$\frac{\partial \rho}{\partial t} + U \frac{\partial \rho}{\partial x} + V \frac{\partial \rho}{\partial y} = D \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} \right), \quad x, y \in \mathbb{R}, t > 0, \quad (6.1)$$

where $\rho(x, y, t)$ is the methane concentration, U and V are the advection velocities along x and y directions, respectively, and D ($D > 0$) is the diffusion coefficient. If $U = 0$, $V = 0$, (6.1) is the standard diffusion equation. If $|U| \gg 0, |V| \gg 0$, then the equation (6.1) is called the advection-dominated diffusion [303]. The fractional anomalous diffusion model is described as follows:

$${}_0^C D_t^\alpha \rho(x, y, t) = k_\alpha \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} \right) + f_d(\rho, x, y, t) + f_c(\tilde{\rho}, x, y, t), \quad (6.2)$$

and space fractional diffusion equation is:

$$\frac{\partial \rho(x, y, t)}{\partial t} = k_\beta \left(\frac{\partial^\beta \rho}{\partial |x|^\beta} + \frac{\partial^\beta \rho}{\partial |y|^\beta} \right) + f_d(\rho, x, y, t) + f_c(\tilde{\rho}, x, y, t), \quad (6.3)$$

where $\rho(x, y, t)$ is the methane concentration, k_α and k_β are positive constants representing the diffusion rates, $f_d(\rho, x, y, t)$ is the source, $\tilde{\rho}$ is the measured data of ρ from the sensors, $f_c(\tilde{\rho}, x, y, t)$ is the control input by mobile actuators to spread the biochar, and its dynamic equation depends on the closed-loop control that designed by the user based on certain control performance requirements, and we will introduce it in Section 6.3. ${}_0^C D_t^\alpha \rho$ is Caputo fractional derivative of order α ($0 < \alpha \leq 1$) which is defined in [304]. Its mathematical presentation is shown below:

$${}_0^C D_t^\alpha \rho = \begin{cases} \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\tau)^{-\alpha} \rho_\tau(x, y, \tau) d\tau, & 0 < \alpha < 1, \\ \frac{\partial \rho}{\partial t}, & \alpha = 1. \end{cases} \quad (6.4)$$

The operators $\frac{\partial^\beta \rho}{\partial |x|^\beta}$ and $\frac{\partial^\beta \rho}{\partial |y|^\beta}$ are Riesz fractional derivatives. The representation is defined as follows [304]:

$$\frac{\partial^\beta \rho}{\partial |x|^\beta} = \begin{cases} -c_\beta ({}_{RL}D_{a,x}^\beta \rho + {}_{RL}D_{x,b}^\beta \rho), & 1 < \beta < 2, \\ \frac{\partial^2 \rho}{\partial x^2}, & \beta = 2, \end{cases} \quad (6.5)$$

$$\frac{\partial^\beta \rho}{\partial |y|^\beta} = \begin{cases} -c_\beta ({}_{RL}D_{c,y}^\beta \rho + {}_{RL}D_{y,d}^\beta \rho), & 1 < \beta < 2, \\ \frac{\partial^2 \rho}{\partial y^2}, & \beta = 2, \end{cases} \quad (6.6)$$

where ${}_{RL}D_{a,x}^\beta$, ${}_{RL}D_{x,b}^\beta$, ${}_{RL}D_{c,y}^\beta$, ${}_{RL}D_{y,d}^\beta$ are the left/right Riemann-Liouville derivatives for variable x and y defined in [304]. The parameter c_β in (6.5) and (6.6) is $c_\beta = \frac{1}{2 \cos(\beta\pi/2)}$.

Let $P = (p_1, \dots, p_n)$ be the location of n actuators and $|\cdot|$ be the Euclidean distance. Each actuator at position p_i will receive information from sensors and then move towards the methane diffusion area and release or spread the biochar by a control law. The objectives are:

- Control the methane diffusion to a confined area.
- Spread biochar optimally while avoiding overspread the area.
- Suppress the methane emissions.

n actuators will divide Ω into a set of n polytopes $V = v_1, \dots, v_n$, $p_i \in v_i$, $v_i \cap v_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^n \bar{v}_i = \bar{\Omega}$ ($\bar{v}_i = v_i \cup \partial v_i$ and $\bar{\Omega} = \Omega \cup \partial \Omega$) [301]. The actuators should be close to areas with high methane concentrations in order to control the diffusion process and reduce the methane emitted area. This will allow the methane to quickly absorb the biochar and prevent further diffusion. But putting all the actuators very close to the methane source is not a good strategy. Hence, to decide the positions of the actuators, we consider minimizing the following cost function [301]

$$\mathcal{K}(P) = \sum_{i=1}^n \int_{V_i} \rho(q) |q - p_i|^2 dq, \text{ for } q \in \Omega. \quad (6.7)$$

Obviously, we needed to minimize $\mathcal{K}$, the distance between actuators and desired location generated by CVT, $|q - p_i|$ should be small when the methane density $\rho(q)$ is big. The density function $\rho(q)$ determines the optimal positions of the actuators. A necessary condition for $\mathcal{K}$ to be minimized is that $\{p_i, V_i\}_{i=1}^k$ is a CVT of Ω [234]. Our algorithm is based on a discrete version of (6.7), and the concentration of methane comes from the measurements of mesh sensors.

6.3 Algorithm of Optimal Control Strategy & Simulation Platform

6.3.1 Optimal Actuation Algorithm

In this section, we give the algorithm to compute the locations of actuators by using CVT [305], [306], [307], as mentioned before, we use Lloyd's method to determine CVT. The algorithm is described as follows [234]: Give a region Ω , a density function $\rho(x)$ defined for all $x \in \bar{\Omega}$, and a positive integer k . Therefore, the algorithm 2 has the details of how to generate a CVT map to command UAVs moving towards the mass centroids points.

Algorithm 2 CVT generation based on Lloyd's method.

Require:

Initial set of K points $\{Z_i\}_{i=1}^k$ as the generators

Ensure:

Voronoi generators are in region Ω

- 1: Initialize UAVs positions $p_i(0) \in \Omega$;
 - 2: Initialize sensors positions $q_j \in \Omega$;
 - 3: **while** Convergence condition not satisfied **do**
 - 4: **for** each sensor q_j **do**
 - 5: Find the nearest UAVs $p_i(k)$;
 - 6: Check $q_j \in \Omega$ and send it to UAV p_i ;
 - 7: **end for**
 - 8: Each UAV calculate the mass centroids $p'_i(k)$;
 - 9: Each UAV move towards to $p'_i(k)$;
 - 10: Update $k = k + 1$;
 - 11: Update $p_i(k) = p'_i(k)$
 - 12: **end while**
-

Although each UAV can calculate its own desired mass centroid, it still requires computational resources, which could slow down convergence or, even worse, cause divergence. Thus, to overcome these limitations, we assume that an actuator can communicate with the sensors and other actuators within a changeable radius r_i . Here we introduce an improved algorithm from [305] with mild modification. After each Voronoi cell and its centroid have been generated, each UAV will do the following:

1. Assign its detection range r_i with a small initial value, detect all its neighboring sensors with radius r_i .
2. Generate its own Voronoi cell within the radius r_i .
3. For every sensor q_j , compute $L_i = \max_j |q_j - p_i|$.

4. If $r_i > 2 \times L_i$, stop. Otherwise, set $r_i = 2 \times r_i$, go to the next iteration.

Considering UAVs indeed have a detection range, this procedure can lower the computational cost due to the real application such as spreading biochar. r_i obtained at the first step will be used as the initial values for the following steps. More specifically, these procedures explained the different strategies in the For loop of Algorithm 2. This improvement on the algorithm from [305], [308] helps to reduce the computational requirements.

The mobile actuators are regarded as virtual particles and obey the second-order dynamical equation:

$$\begin{aligned}\ddot{p}_i &= k_p e - k_v \dot{p}_i + k_i \int_0^t e(\tau) d\tau, \\ e &= (\bar{p}_i - \Delta p'_i) - p_i,\end{aligned}\tag{6.8}$$

where $\Delta p'$ is the desired offset due to the wind, the first term of (6.8) on the right-hand side is a proportional control law used as the force input to control the motion, and $\bar{p}_i$ is the desired location from computing the mass centroid of the current Voronoi cell. The second term is the viscous friction artificially introduced [309]. k_v is the friction coefficient, and $\dot{p}_i$ denotes the velocity of the actuator i . The last term is the integral error to reduce the fluctuation of UAVs when approaching the methane source. As for how to determine k_p and k_v , we will explain in the experiment subsection.

6.3.2 Wind Field Effects

We consider how methane diffusion is affected under the wind field, as we mentioned earlier. In general, the wind field not only affects methane diffusion, but it also affects the actuators to spread biochar mulch. Since applying wind to diffusion equation, it has been claimed to be advection-diffusion. Here, we introduce how the wind affects spreading biochar mulch.

As we know, the trajectories of the objects, which are the biochar mulch falling from the height, would be parabolas, as shown in Figure 6.1.

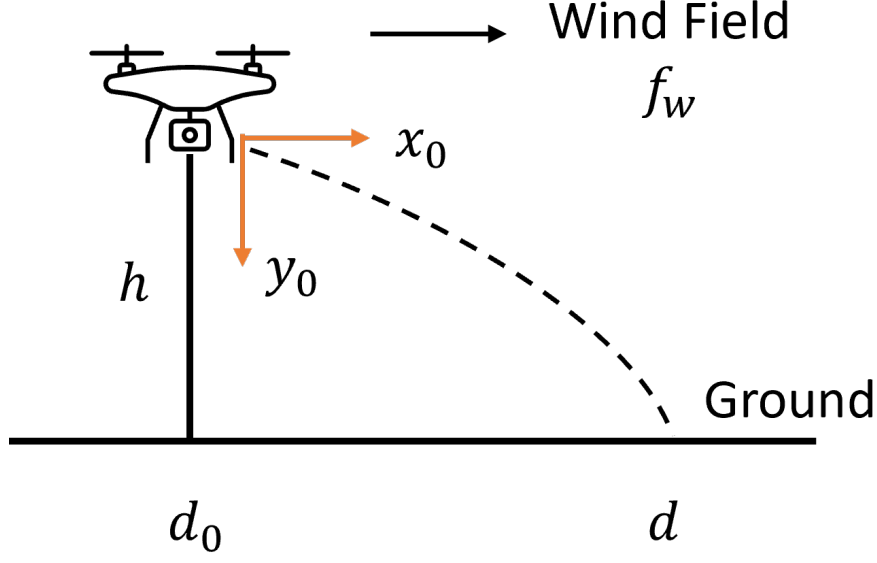


Figure 6.1: Spreading biochar mulch under wind field

The wind force can be split within x and y directions. Based on the drag equation from fluid dynamics, the wind force is represented as below:

$$f_w = \frac{1}{2} C_d \rho_{air} S V^2, \quad (6.9)$$

where C_d indicates drag coefficient, ρ_{air} denotes air density, S represents the cross-sectional area, and V is wind velocity relative to biochar mulch. Since we have wind force, we need to add the term on control law f_c , at the following location:

$$\begin{cases} x^* = \frac{1}{2} \ddot{x}_0 (t^*)^2 + \dot{x}_0 t^* + x_0, \\ y^* = \frac{1}{2} \ddot{y}_0 (t^*)^2 + \dot{y}_0 t^* + y_0, \end{cases} \quad (6.10)$$

the x_0 and y_0 represents the current mulch location, and $t^* = \sqrt{\frac{2h}{g}}$ is the time that is needed for biochar mulch to drop to the ground. Therefore, with wind, the methane concentration in our measurement should be modified to:

$$\rho(x, y, t) \rightarrow \rho(x_m, y_m, t), \quad (6.11)$$

where $x_m = x + x^*$, $y_m = y + y^*$. We will introduce the methane diffusion, biochar mulch spreading effects under wind in Section 6.3.3 and Section 6.4.

6.3.3 Simulation Platform of FADE-MAS-2D

The simulation platform FADE-MAS2D is mainly based on FO-DIFF-MAS2D [302] for fractional-order modeling and control with mobile sensors and actuators. FADE-MAS2D is deployed fractional module to solve the PDE (6.4)

Based on (Section 6.11), $f(\tilde{\rho}, x, y, t)$ consists of control and methane source.

$$f(\tilde{\rho}, x, y, t) = f_c(\tilde{\rho}(x_m, y_m, t), x, y, t) + f_d(x, y, t), \quad (6.12)$$

where $\tilde{\rho} = \rho(x_m, y_m, t)$ is the measured data of $\rho(x, y, t)$ from mobile sensors, $f_c(\tilde{\rho}(x_m, y_m, t), x, y, t)$ is the control law that applied to mobile sUAVs, and its closed-loop control law is provided by the user. $f_d(x, y, t)$ is the methane source.

The following two types of boundary conditions can be used in the problem formulation.

- Dirichlet Boundary Condition

$$\rho = C,$$

where C is a real constant.

- Neumann Boundary Condition

$$\frac{\partial \rho}{\partial n} = C_1 + C_2 \rho$$

where C_1 and C_2 are two real constants. n stands for the outward normal direction.

FADE-MAS2D uses the finite difference method to discretize the spatial derivative in (Section 6.2), then leaves the time domain integration to Matlab/Simulink. Specifically, FADE-MAS2D is used to solve a two dimensional fractional diffusion equation. As an extension of FO-Diff-MAS2D [302], the main features of FADE-MAS2D are listed as follows:

- Sensors and actuators can be collocated or non collocated.
- Disturbances can be movable, and they can change over time.
- The mobility platform dynamics of sensors and actuators can be modeled as either the first or second orders.
- Movement of sensors and actuators can be open-loop or closed-loop.
- Control algorithms can be applied in $f_c(\tilde{\rho}(x_m, y_m, t), x, y, t)$.

6.4 Simulation Experiments

In this section, we conducted a set of experiments with methane emissions and spreading mulch to check system performance. We first introduce the simulation configuration and initialization of sensors and actuators in (Section 6.4.1). Next, we present and analyze the time and space fractional diffusion processes that were compared with and without wind fields using our approach (in Section 6.4.3, 6.4.2 6.4.4, 6.4.5). We then explore the effect of the different control law parameters we proposed to seek the optimal control strategy (Section 6.4.6). Finally, we validate our FADE-MAS2D platform approach to simulate the multiple diffusion sources (Section 6.4.7).

6.4.1 Experiment Setup

To evaluate our methane emissions under different situations, we proposed FADE-MAS2D for this simulation. The boundary of simulation is given by $\Omega = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1\}$. The methane source is modeled as a point disturbance f_d to the fractional diffusion system Section 6.14 with its position at (0.8, 0.2) and the representation of f_d is shown below:

$$f_d(t) = 20e^{-t}|_{x=0.8, y=0.2}. \quad (6.13)$$

The methane source begins to move at $t = 0$ in Ω , four actuators that can release or spread the biochar mulch are deployed with initial positions at (0.33,0.33), (0.33,0.66), (0.66,0.33), (0.66,0.66), respectively. There are 30×30 sensors evenly distributed in Ω , and they form a grid over the area. The wind coefficients U and V are initialized with 0.2, the unit of wind velocity is normalized based on simulation length. In our simulation, we suppose that once the sensors are deployed, they remain static. In Section 6.2, we can see the initial positions of actuators, the positions of the sensors, and the position of the methane source.

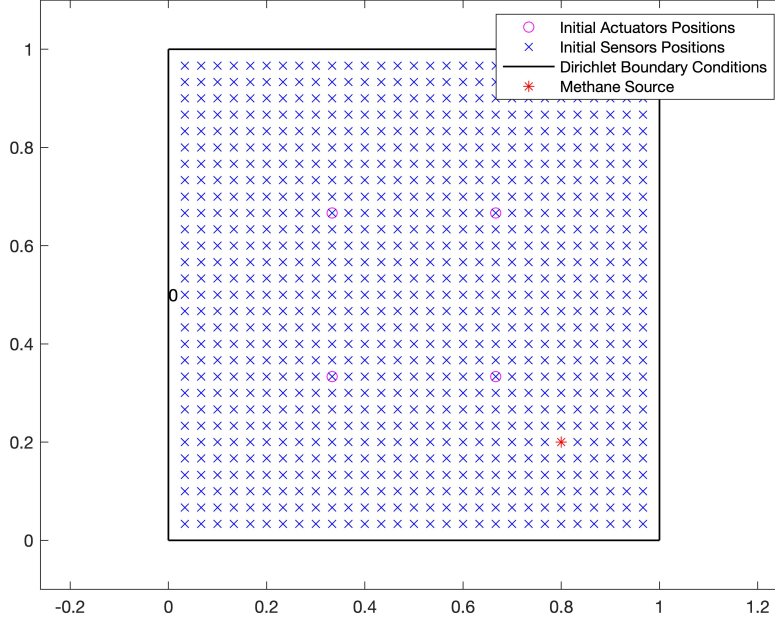


Figure 6.2: Initialization of static mesh sensors, methane source and mobile actuators.

Several preparation steps are needed before simulation studies. The entire simulation time varies from $t = 4s$ for a single source, to $t = 5s$ for multiple sources. The time step is set to $\Delta t = 0.005s$ for the system's numerical algorithm. The UAVs recompute their desired position every 0.1s. To demonstrate how the UAVs inhibit methane diffusion, we turn on the actuators and at $t = 2s$. The evolution of the methane diffusion and biochar mulch released by UAVs are demonstrated in the following sections.

6.4.2 Wind Effects

Let us recall the advection-diffusion equation. When $|U| \gg 0$ and $|V| \gg 0$, the standard diffusion equation is transformed into an advection-diffusion equation, where U and V denote the velocity in the U and V directions, respectively. Based on this, and given that the coefficient of diffusion D is 0.01, we choose the velocity along the x and y axes of the U and V areas to be 0.1 with direction, which is ten times greater than D . The simulated wind effects are shown in Figure 6.4a and no wind effects are demonstrated in Figure 6.4b.

6.4.3 Time Fractional Diffusion Equation

The dynamic transport system with control input is modeled as the following time-fractional diffusion equation:

$${}_0^C D_t^\alpha \rho(x, y, t) = 0.01 \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} \right) - 0.1 \frac{\partial \rho}{\partial x} + 0.1 \frac{\partial \rho}{\partial y} + f_d(\tilde{\rho}, x, y, t) + f_c(\tilde{\rho}, x, y, t), \quad (6.14)$$

with homogeneous Neumann boundary, given by

$$\frac{\partial \rho}{\partial n} = 0.$$

Therefore, the initial value and boundary condition are:

$$\begin{aligned} \rho(x, y, 0) &= 0, \quad 0 < x < 1, \quad 0 < y < 1, \\ \rho(0, y, t) &= \rho(1, y, t) = 0, \\ \rho(x, 0, t) &= \rho(x, 1, t) = 0. \end{aligned} \quad (6.15)$$

The fractional order α dictates the time-fractional diffusion behaviours. To illustrate the effect of α , we tested different α values to show the behaviours of the diffusion process. We plot total methane emissions to show the different behaviors with different α values. The different diffusion behaviors with different α are shown in Figure 6.3.

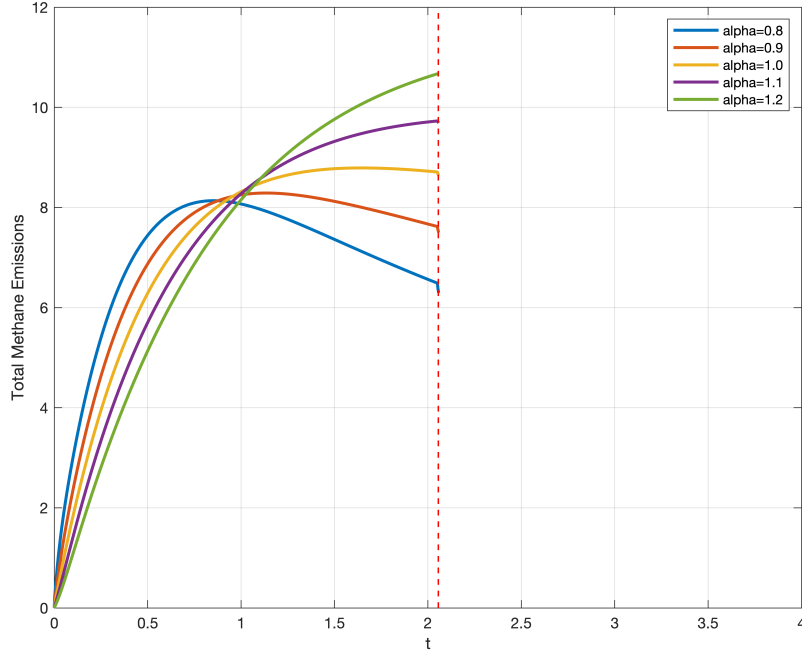
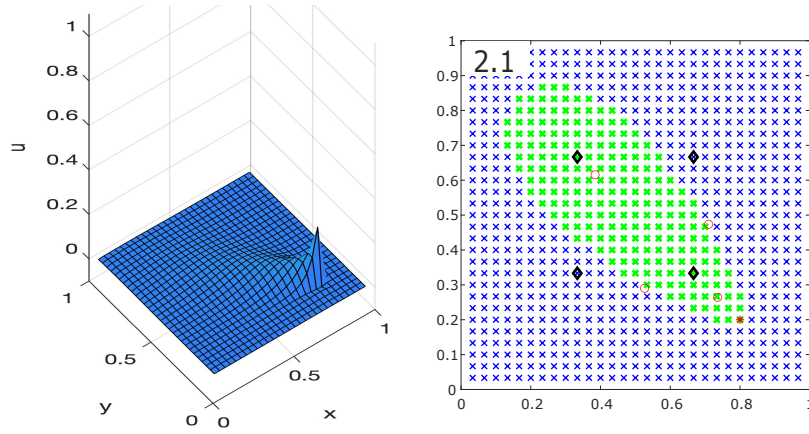
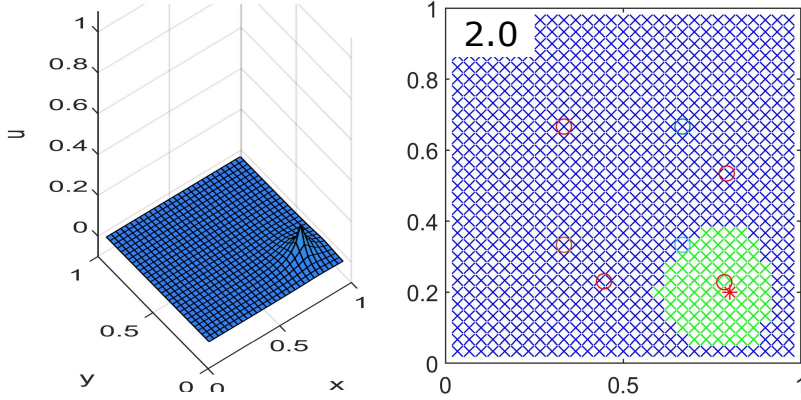


Figure 6.3: The diffusion behaviours under different α .

As we can see in Figure 6.3, we choose α from 0.8 to 1.2 since we want to show the different behaviors with integer order and fractional orders. When $\alpha < 1$, the methane emissions grow fast as sub-diffusion but easily reach the critical value when $t < 2.25s$. As we mentioned earlier, the controller's delay in lowering the methane after 2 seconds is because the UAVs need to move towards the pollution area (CVT desired locations are marked as red dots in Figure 6.4). The super-diffusion happens when $\alpha > 1$, which is growing slowly, but they still need time to reach the maximum concentration.



(a) Methane diffusion with wind field.



(b) Methane diffusion without wind.

Figure 6.4: (a) The solution of time-fractional partial differential equations for simulating the methane diffusion system with the upper left wind, its velocity is 0.1, and the diffusion coefficient is 0.01. The alpha is fixed $\alpha = 0.8$. (b) Same as (a) but no wind effect.

6.4.4 Space Fractional Diffusion Equation

The system is modeled by the following space fractional diffusion equation, that is presented as follows:

$$\frac{\partial \rho(x, y, t)}{\partial t} = 0.01 \left(\frac{\partial^\beta \rho}{\partial |x|^\beta} + \frac{\partial^\beta \rho}{\partial |y|^\beta} \right) - 0.1 \frac{\partial \rho}{\partial x} + 0.1 \frac{\partial \rho}{\partial y} + f_a(\rho, x, y, t) + f_c(\tilde{\rho}, x, y, t), \quad (6.16)$$

where the boundary condition with homogeneous Dirichlet boundary condition is given by:

$$\rho(x, y, t) = 0,$$

where $x, y \in \partial\Omega$. The initial condition is:

$$\rho(x, y, 0) = 0, \quad 0 < x < 1, 0 < y < 1, \quad (6.17)$$

Here, we choose several values of β and plot the methane diffusion over time. The graph is presented in Figure 6.5.

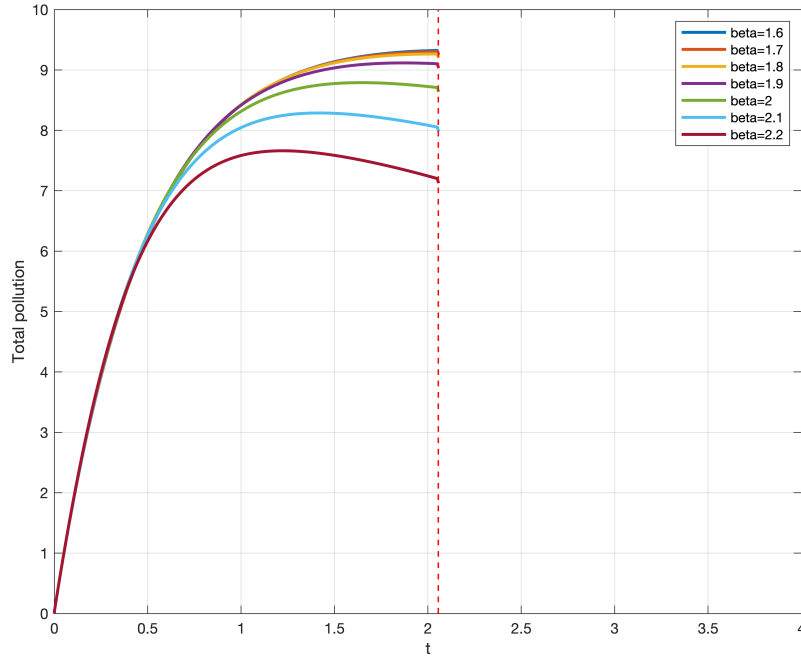


Figure 6.5: The diffusion behaviours under different β .

As we can see from 6.5, the higher β corresponds to a smaller diffusion process.

6.4.5 Time and Space Fractional Equations

The diffusion process can be fractional in both the time and space domains, respectively. The transport dynamic is defined as below:

$${}_0^C D_t^\alpha \rho(x, y, t) = 0.01 \left(\frac{\partial^\beta \rho}{\partial |x|^\beta} + \frac{\partial^\beta \rho}{\partial |y|^\beta} \right) - 0.2 \frac{\partial \rho}{\partial x} + 0.2 \frac{\partial \rho}{\partial y} + f_d(\rho, x, y, t) + f_c(\tilde{\rho}, x, y, t), \quad (6.18)$$

where its boundary condition is:

$$\rho|_{\partial\Omega} = 0.$$

From the previous Section 6.4.3, and Section 6.4.4, we choose that $\alpha = 0.8$ in Figure 6.3 and $\beta = 1.8$ in Figure 6.5 for control study. Based on the literature review describing diffusion of methane that satisfied sub-diffusion [310], we select here $\beta = 1.8$. The total methane emissions are described in Figure 6.6.

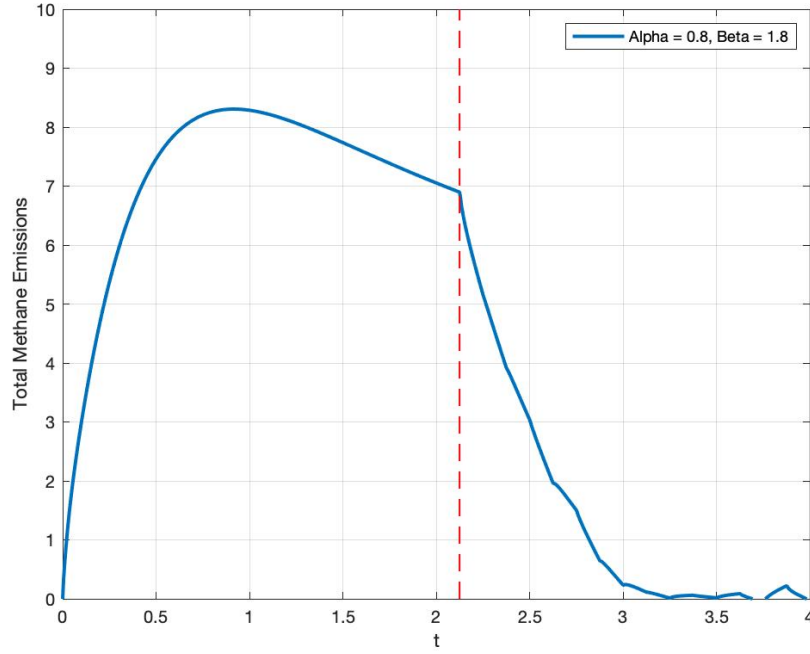


Figure 6.6: The diffusion behaviours when $\alpha = 0.8$ and $\beta = 1.8$. Note that when $t = 2.25s$, biochar mulch spreading control is turned on.

6.4.6 An Optimal Control Strategy

In this section, we will validate our controller with different k_p and k_i from (6.8). At $t = 2.25s$ in the simulation, the UAVs begin to respond by spreading biochar

mulch k_s (spray rate) to absorb methane diffusion. There are several PI controller tuning methods, such as heuristic tuning, rule-based tuning, and model-based tuning. Due to different scenarios, we would generate a cost function for tuning k_p and k_i that considers three main components, which are total methane emissions area, control efforts, and UAVs position error. The cost function is defined as follows:

$$\min_{k_p, k_i, k_s} J(k_p, k_i, k_s) = \min_{k_p, k_i, k_s} \{R \sum (\frac{X_c}{X_{no-c}}) + K \sum (\frac{u}{u_{max}})^2 \frac{1}{NA} + Q \sum (e_k^2)\},$$

$$R + K + Q = 1, \quad (6.19)$$

where R, K, Q are the weights of each components. The first term is the total methane emissions, the second term is the total control input signals (control efforts) per number of actuators; and the third term is the total actuators (UAVs) position error. The goal of this cost function is that we want to minimize these three items to get the optimized k_p and k_i parameters for the actuator controller. The spraying rate of the amount of biochar k_s in f_c must also be optimized.

To find the optimal control performance, we need to set different weights for each item in (6.19), The table containing the different weights is shown in Table 6.1. We proposed four scenarios that consider all items and assign the same weight to each item, as well as extreme scenarios involved. The friction parameter (the noise of UAVs, such as drifting) has been set at $k_v = 1$.

Table 6.1: The weights of cost function for determining k_p , k_i , and k_s .

	Methane Emissions	Control Efforts	Position Error	Average
R	1	0	0	0.33
K	0	1	0	0.33
Q	0	0	1	0.33

Next, we create the matrix for k_p , k_i , and k_s , respectively, and use their permutation of these parameters to find minimization of the cost function. By selected ten number from each parameters uniformly by given range are $0.1 \leq k_p \leq 80$, $0 \leq k_i \leq 80$, and $0 \leq k_s \leq 80$, we would find the minimum J among 1000 grids. The minimum J of each cases is shown in Table 6.2. As we can see, when the friction coefficient $k_v = 1$ and the weights of the cost function have been averaged, the controller coefficients $k_p = 0.1$, $k_i = 0$, and sprayer rate $k_s = 8.8899$ in Figure 6.7 are consider as the optimized coefficients for f_c . The controller effects with optimized k_p , k_i , and k_s under different R, K, Q is shown in Figure 6.7 as well.. The methane diffusion has been controlled, as demonstrated by simulations in Figure 6.9.

Table 6.2: The optimized k_p , k_i , and k_s under different weights configuration.

	K_p	K_i	K_s
Ave R, K, Q	0.1	0	8.8899
Only R	0.1	0	80
Only K	0.1	0	0
Only Q	80	0	0

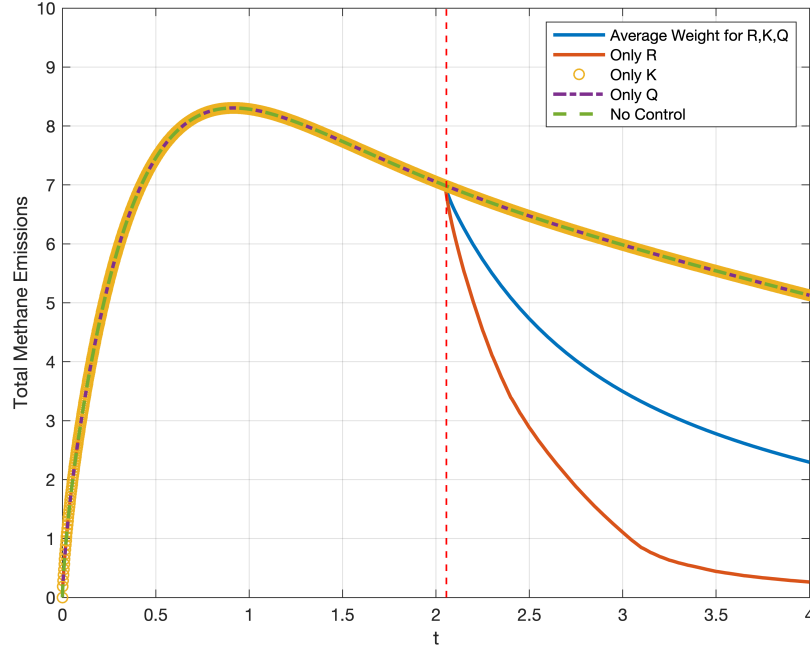


Figure 6.7: Comparison different control process under different R , K , Q and no control baseline.

In Figure 6.7, we can see the control effects of the extreme situation and it would not be suitable for the real methane emission suppression process. The R term is unconcerned about the total amount of biochar carried by UAVs, which means it could spray indefinitely. The drone's battery life and speed are limited (limited by K term), which means the drone cannot just fly towards the desired location within a second. The Q term, like the K term, only considers the UAVs' position error, which has no concern on UAV missions (suppression of methane emissions). As a result, this cost function can be used as an optimal control strategy for better control effects when determining which parts are most important or costly during the process of reducing methane emissions.

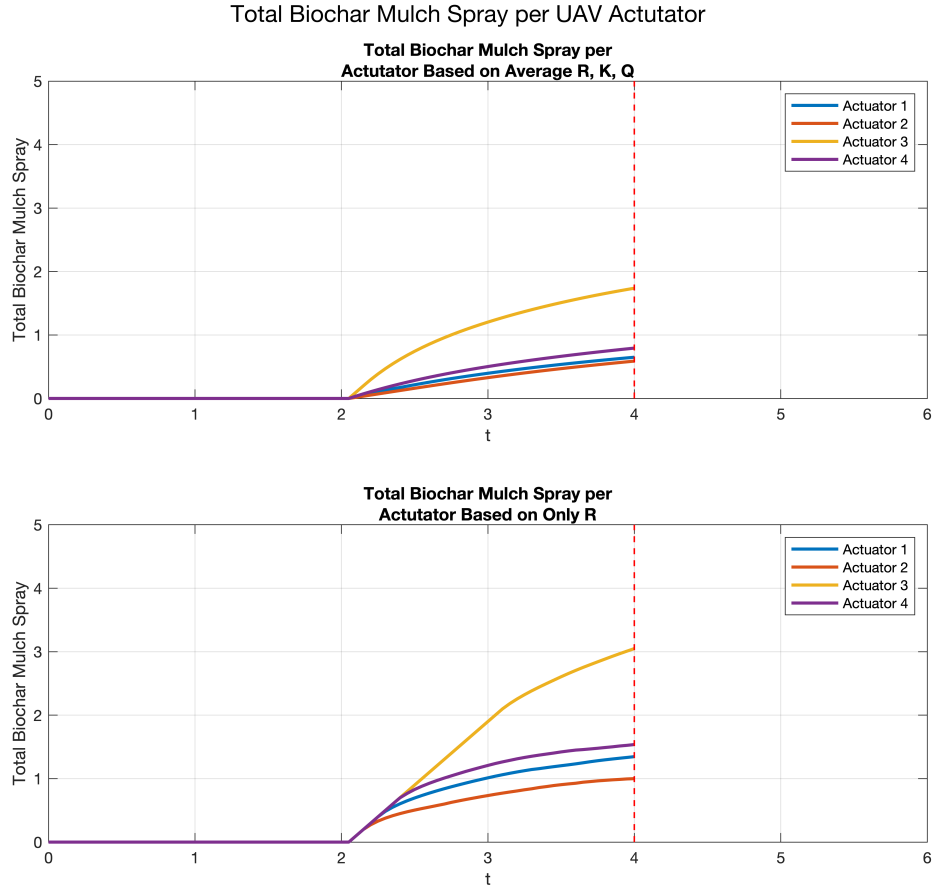
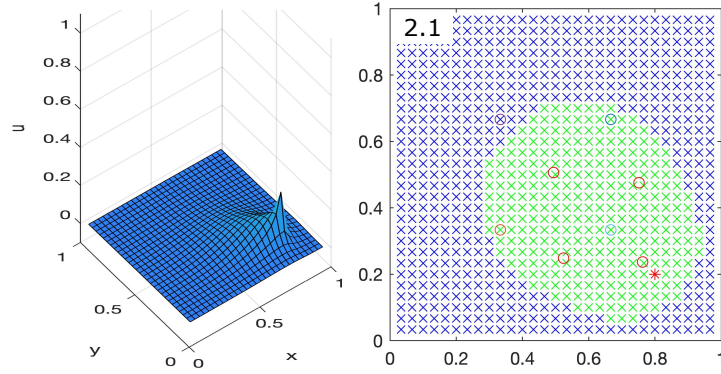
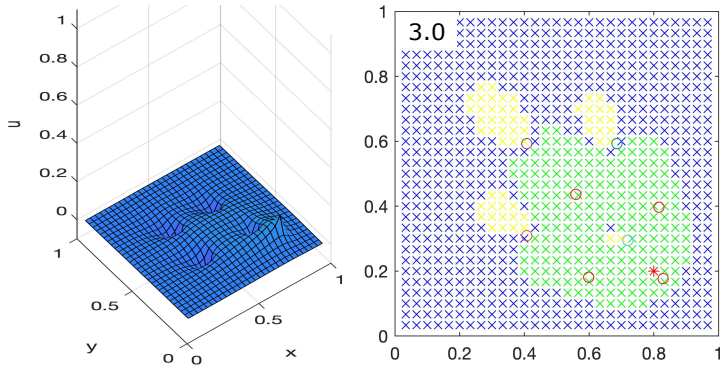


Figure 6.8: The total amount of biochar mulch sprayed based on different weights configuration

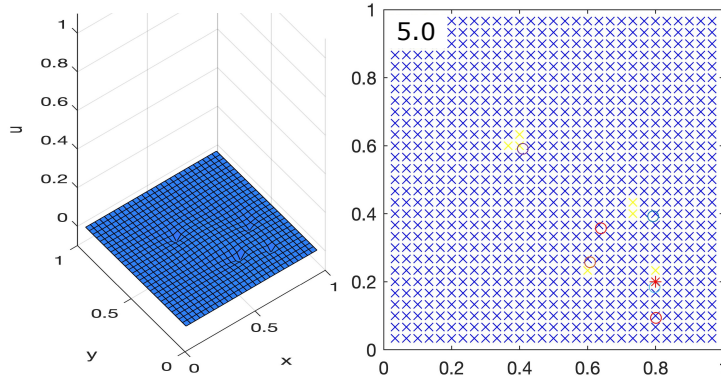
The total spray per UAV based on Table 6.2 has been demonstrated in Figure 6.8. As we can see, the different actuators would spray different amounts of biochar mulch due to their different initial positions, whether close to the emission source, etc. The Only K , the only Q , and the no control cases have not taken k_s which means no spray at all during the suppression mission. This would be useful when we do real emission control, as the UAVs' payload needs to be taken into account in different spray situations based on the simulation results.



(a) Methane diffusion with wind field effect.



(b) The UAVs start to spread biochar mulch.



(c) The UAVs fly to their desired CVT locations and finish the control.

Figure 6.9: (a) The methane is diffusing, as shown in the green area under the wind field, and it is about to get the most concentration. (b) The sensors detect methane emissions and send them to UAVs. The UAVs start to spread biochar mulch as shown in the yellow area to curb the methane. The mulch is considered they have weights and can be affected by wind. (c) The control ends at the end of the simulation time.

6.4.7 Multiple Methane Emissions Sources

We now demonstrate the diffusion of multiple methane emission sources and their suppression process. Here we demonstrated the diffusion under wind field with two and four sources to explain how our controller and fractional anomalous diffusion equations work together.

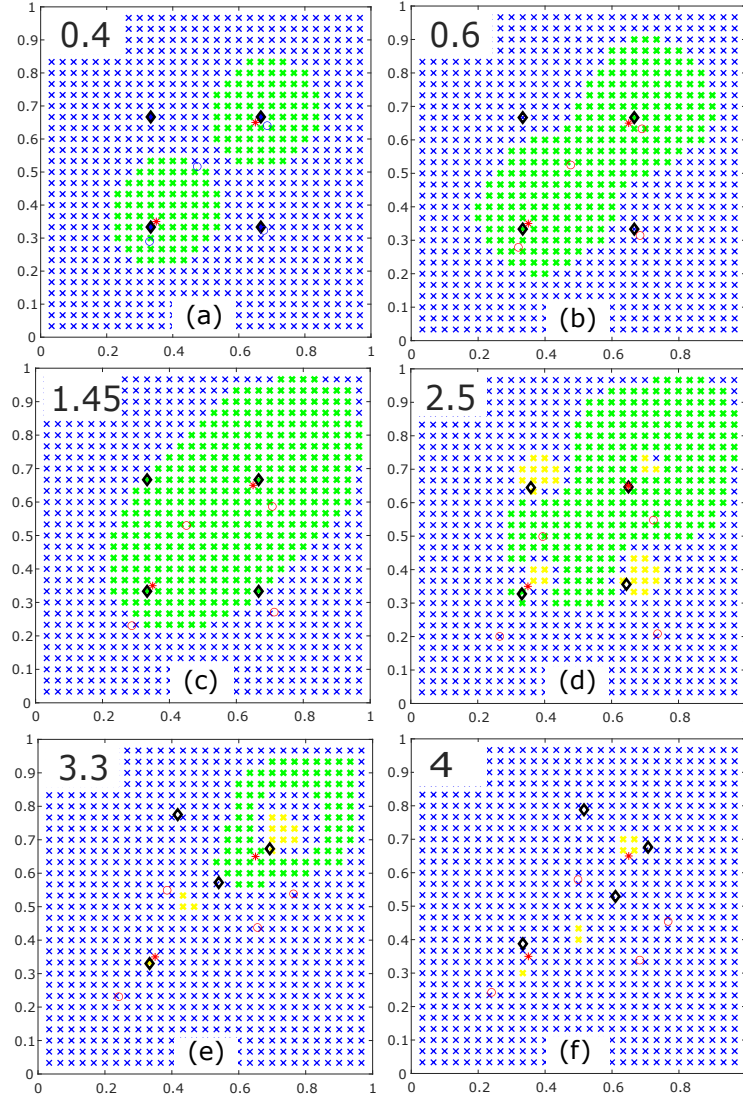


Figure 6.10: (a)(b)(c) The multiple methane sources are diffusing as shown in the green area under the wind field, and it is about to get close to each other. (d) The sensors detect methane emissions and send them to UAVs. The UAVs start to spread biochar mulch as shown in the yellow area to curb the methane. (e) The control ends at the end of the simulation time.

We set up methane sources at $(0.35, 0.35)$ and $(0.65, 0.65)$ in Figure 6.10, and we also changed the wind direction with $U = -0.2$ and $V = -0.2$ to see how the two emissions combine. As we can see in Figure 6.10, the diffusion of multiple sources is fused as expected with each other, and our actuators with the desired control law perform better than previous work [301]. We can obtain the UAVs tracing the desired CVT location and emission sources to spread the biochar mulch dynamically. Similar experiments with four sources are demonstrated in Figure 6.11. To see the multi-source emissions diffusion process, we set up the sources in $(0.3, 0.3)$, $(0.7, 0.3)$, $(0.3, 0.7)$, and $(0.7, 0.7)$, with wind coefficients of $U = -0.2$ and $V = 0.2$. The actuators take control efforts for a longer time than single methane emission sources since the diffusion area is larger. Overall, these experiments illustrate that our FADE-MAS2D platform is capable of multiple scenario simulations of methane emission suppression and can provide clear instructions for real-time methane emissions management.

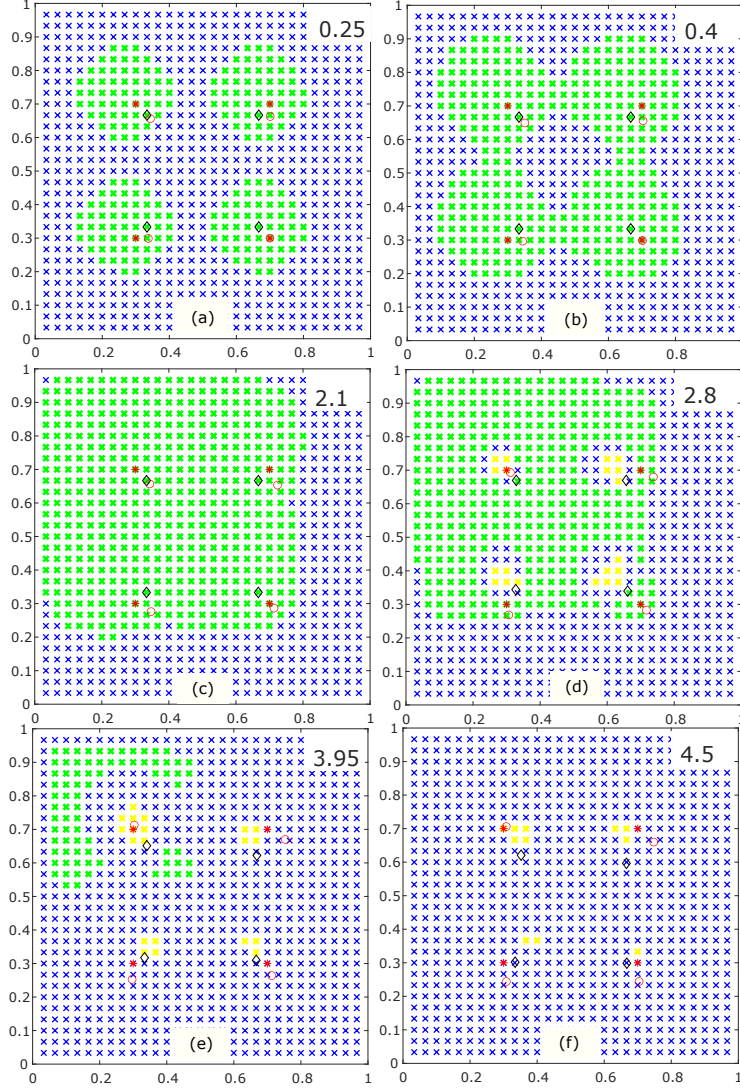


Figure 6.11: (a)(b) The four methane sources are diffusing as shown in the green area under the wind field and are about to get close to each other. (c)(d) The UAVs start to spread biochar mulch as shown in the yellow area to curb the methane, and it will be eliminated at the end.

6.5 Conclusions

In this Chapter, we introduce an FADE-MAS2D platform for simulating soil methane emission suppression control using swarming UAVs that spread biochar mulch based on fractional anomalous diffusion with wind field effect. Our approach uses fractional calculus to solve the diffusion equation by adding a wind field. The main challenges are that methane concentration is time-varying due to gas plume properties, and the optimal control strategies for methane emissions are difficult to

determine due to various disturbances and delays. We address the control strategy by combining CVT and PI controllers to minimize a cost function.

Our results show that when $0 < \alpha \leq 1$ and $1 < \beta \leq 2$, sub-diffusion occurs, and when $\alpha > 1$ and $\beta > 2$, super-diffusion occurs. α and β are determined based on the methane diffusion behaviours. Second, our optimal control strategy is based on a cost function that takes into account three situations: total methane emissions, control efforts, and UAV position errors ϵ . This approach to suppressing methane outperforms the proportional-only controller approach with CVT which significantly reduces the system oscillation. Third, our approach performs well across multiple emissions sources (with different sources' locations and different wind directions). Overall, our system provides an accurate and robust model for estimating methane emissions and suppressing emissions processes. In the future, our platform could be used to efficiently manage soil methane emissions in practice.

Chapter 7

XAIOT ENABLED SMART APPLICATIONS

7.1 A Greenhouse Gas Proximal Sensing System Using Chemiresistive Strip and Miniaturized Radar Array

7.1.1 Introduction

Soil carbon sequestration techniques play a vital role in addressing climate change, providing water, and promoting biodiversity. It is necessary for crucial environmental processes that have positive impacts on both plants and humans. Human activities, particularly those that involve the combustion of biomass using fire, lead to a decrease in soil coverage and cause both immediate and long-term depletion of soil organic carbon. Both soil cultivation and drainage facilitate the exposure of organic molecules to oxygen, leading to their oxidation. We must prevent the loss of carbon from the soil caused by erosion and the production of greenhouse gases.

According to a report by the California Air Resources Board in 2019, the agricultural sector in California was responsible for 7.6 percent of the state's GHG emissions. The majority of these emissions came from sources of methane (CH_4) and nitrous oxide (N_2O) [311]. Approximately 71 percent of the greenhouse gas emissions in the agriculture industry are generated by animals. In 2019, livestock emissions have increased by 18% compared to the levels seen in 2000. Enteric fermentation and manure management are the primary sources of methane emissions, with dairy activities being responsible for the majority of emissions from cattle. From 2000 to 2007, GHG emissions resulting from the management of dairy manure and enteric fermentation increased. Since 2007, the annual fluctuations in emissions have been minimal. The soil is expected to absorb these significant carbon emissions in order to facilitate carbon sequestration, which can improve soil health, enhance agricultural productivity, bolster climate resilience, and strengthen the resistance of farms to droughts and severe rainfall. If we are unable to measure GHG emissions, we cannot effectively manage or pursue more carbon sequestration efforts. Quantifying GHG emissions at a parts per million (PPM) level remains a difficult task, since it requires precise, rapid, cost-effective, and large-scale sensing capabilities.

Researchers have explored a variety of methods and sensors for measuring GHG. The portable carbon analyzer is typically utilized to identify gas leaks in the field or manure compost. These sensors may be utilized and provide results in a short

amount of time. The primary sensing technique relies on gases that exhibit observable infrared bands. Based on these features include the G2508 Picarro gas flux analyzer¹, the Irwin portable pump-based² infrared sensor, and the GAZOSCAN³, etc.

Typically, at least one month is required for an appropriate monitoring period to track GHG performance. In other words, researchers must transport a large number of heavy instruments to the field for measurement and data recording, which requires a significant amount of resources, such as intensive labor, additional power supply for the Picarro analyzer, etc. In addition, it takes considerable effort to prepare the detection procedure (equipment warm-up, calibration, etc.) for the analyzer due to varied measurement objectives. The sluggish measurement results may delay the entire project, potentially leading to delays in typical agricultural operations such as irrigation planning and fertilizer management. Besides, the equipment are expensive, making it less affordable.

To tackle the above challenges we mentioned, we are looking for a low-cost solution that doesn't require complex preparation yet is effective and accurate enough to measure GHG concentrations in real-time. Our intuition behind our solution was inspired by air quality monitoring using UWB radar [312], where UWB radar is used to determine the concentration of ammonia by using dedicated materials to absorb ammonia gas. Therefore, we considered another material, metal-organic frameworks (MOFs), which can absorb CO₂ significantly. MOFs are 3D-ordered crystal structures that can be built by combining metal ions as nodes and functionalized organic ligands as struts [313]. Gas storage and separation have widely employed MOFs due to their organized high-surface-area geometry, structural flexibility, and enormous pores throughout the structure. The zeolitic imidazolate framework (ZIF) is a class of MOFs where metal cations are bridged by imidazole-based ligands. In particular, the ZIF-67, which is made of Co²⁺ ions and 2-methylimidazole, has been extensively researched for CO₂ capture applications recently as a result of its well-defined and tunable morphology, a uniform pore size distribution, high thermal stability, and high capacity for CO₂ adsorption [314,315].

Previous research efforts on MOFs absorbing methane and carbon dioxide can be found in [316–321]. These studies verified the possibility of using MOFs to sense methane and other GHGs. However, they used lab analyzers or surface acoustic waves (SAW) to measure dielectric variations, which can be costly and difficult to manipulate. In our sensing system, the mechanism for using proximal radar sensor to measure CO₂ is that ZIF-67 has electrical conductivity that is very sensitive to the presence of small quantities of gas. Upon exposure to a small amount of gas,

¹https://www.picarro.com/g2508_gas_concentration_analyzer

²<https://www.inficon.com/en/products/mobile-methane-leak-detector-irwin>

³<https://www.gazomat.com/en/products/gazoscan/>

such as CO₂, the conductivity of the ZIF-67 will increase significantly. Our proximal sensor⁴ would capture this small change by analyzing microwave reflectance. A few studies using microwaves to identify blood type [322], determine electrolytic content in the distilled water [323], and determine how serious diabetes is by sensing exhaled acetone and ethanol [324] all utilized the properties of dielectric changes to sense gas concentration fluctuations. These studies proved our solution is feasible and supported by its underlying physical principle. As a result, we designed our low-cost gas sensing system with easily accessible materials and machine learning algorithms that can provide two advantages (innovations) as follows:

- MOFs material can be applied on a large-scale field to sense GHG emissions concentration without compromising sensing accuracy.
- We use a filtering algorithm for noise reduction to filter out disturbances and enhance sensitivity to various gas concentrations.

Here, we target two challenges: 1) How long can we capture the variation of concentrations within our sensing system’s sensing range in PPM? 2) How do we accurately label the gas concentration for the purpose of calibrating our sensing system? The contributions to this work are as follows:

1. We introduce a gas sensing system that uses a proximal radar sensor with ZIF-67 material to detect GHG at the PPM level.
2. We present an algorithm to quantify the gas concentration at the PPM level from the variation of radar signal amplitude acquired.
3. We conduct real-world experiments to evaluate our sensing system and the sensor processing algorithm.

7.1.2 Materials and Methods

We present a GHG sensing system based on the proximal radar sensor illustrated in Figure 7.1. The sensing system is a small, wireless GHG gas sensor for accurate, real-time gas detection at the PPM level, and we combined the proximal radar sensor with a piece of gas sensing material, which consists of a photo paper, ZIF-67, and printed silver paste as an antenna.

⁴<https://shop.walabot.com/products/walabot-developer-pack-new>

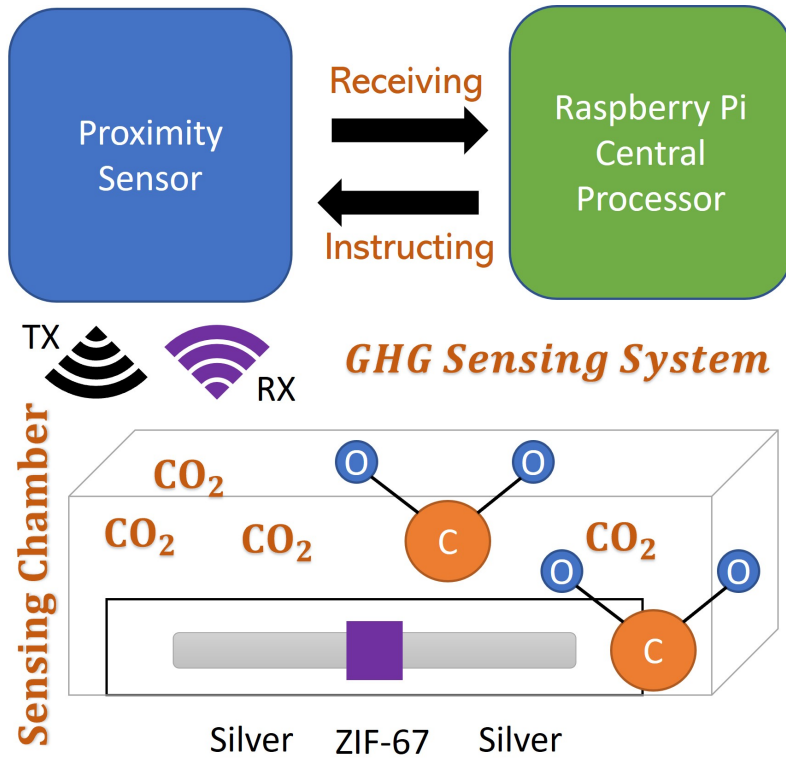


Figure 7.1: GHG sensing system framework.

7.1.2.1 Proximal Sensor Setup

The sensing module consists of four essential components: a Walabot microwave transmitter and receiver, a Raspberry Pi 4B, the display touchscreen, and a CO₂ Sensor (KEYESTUDIO CCS811 Carbon Dioxide⁵) for our system calibration in Figure 7.2. Walabot is a reprogrammable device capable of capturing three-dimensional information that utilizes novel radio frequency technology to see through objects [135]. As shown in Figure 7.2, the Walabot API offers 40 possible antenna pair combinations, which we use to detect variations in CO₂ concentrations and obtain accurate results.

⁵<https://wiki.keyestudio.com/KS0457/keyestudio/CCS811/Carbon/Dioxide/Air/Quality/Sensor>

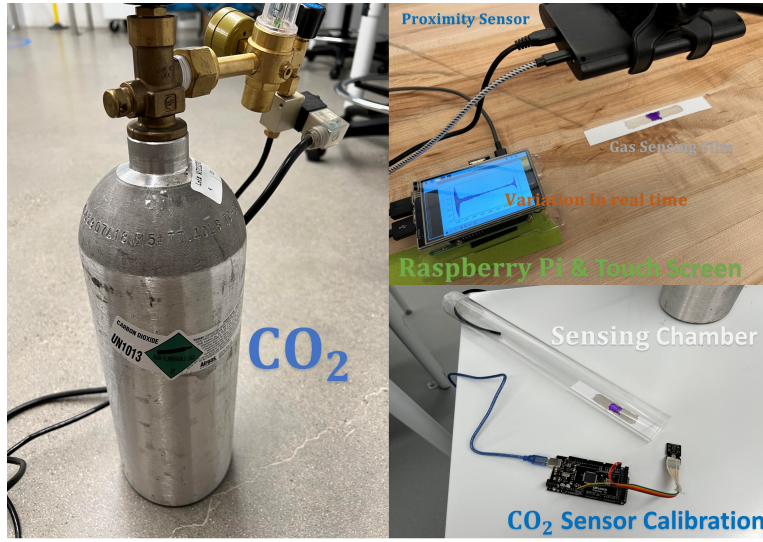


Figure 7.2: GHG sensing system hardware setup. The left is our CO₂ source, and the upper right is a proximal sensor with a Raspberry Pi and touchscreen. The lower right is a sensing chamber that contains CO₂ airflow. The CO₂ sensor is used to calibrate our system.

As for the sensing background, obviously, our target is to sense the variations in magnitude when we measure various gas concentrations. Thus, the minimum sensing background noise needs to be determined. In our previous work [223], we validated the influence for different backgrounds and different sensing distances. In this work, we selected the wood table as our background, and the Walabot is located on top of sensing materials at 3 inches to reduce sensing background noise as much as possible.

7.1.2.2 Sensing Materials Preparation

It has been reported that the adsorption of CO₂ will increase the conductivity of ZIF-67 [325]. The main reasons are: 1) The process of adsorption of CO₂ leads to a considerable temperature increase at the corresponding adsorption location, which means the conductivity has increased; 2) CO₂ molecular is an electron donor, which provides negative charge to the carrier material [326] so that the conductivity increases. As shown in Figure 7.3, we made a gas-sensitive film to test CO₂ concentration.

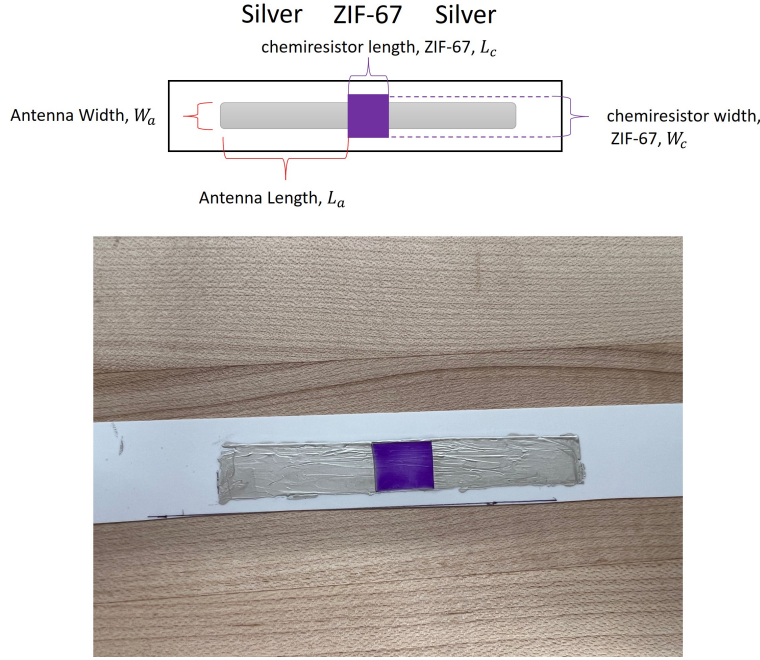


Figure 7.3: GHG sensing film. The W_a and L_a are silver antenna width (1 cm) and length (3 cm), respectively. The L_c and W_c are 1 cm and 1.6 cm, respectively.

As previously stated, the conductivity of ZIF-67 changes after CO_2 adsorption, so we created a set of silver pastes that act as conductors when brought into contact with the ZIF-67, increasing the induced current by the proximal sensor. The underlying physical principle is that electromagnetic waves (EM waves) have electric and magnetic field components. Under the action of an electric field, metal electrons may leap to higher energy levels, and this process absorbs EM waves. Meanwhile, as the electrons jump to higher energy levels, they will also spontaneously decline to lower levels. This process emits energy in the form of EM waves, which our proximal sensor will capture. In Figure 7.3, we made our ZIF-67 attached to silver paste, which is composed of parallel resistors, in order to enhance the current induced by silver paste and capture resistance variation.

The essential material we introduced in Section 7.1.1 is ZIF-67. The synthesis process for ZIF-67 sample is identical to our previous research [327]. Typically, 7 g of $\text{Co}(\text{NO}_3)_2 \cdot \text{H}_2\text{O}$ and 8 g of 2-methylimidazole ($\text{C}_4\text{H}_6\text{N}_2$) were dissolved in a mixed solution of 80 mL ethanol and 80 mL methanol separately. They were further mixed rapidly and stirred for 10 seconds, and then the final solution was kept at room temperature for 24 hours. The resulting purple precipitates collected by centrifugation were washed with ethanol and DI water (deionized water) three times and finally dried at 80 °C for 12 hours to form ZIF-67.

The as-prepared ZIF-67 (10 mg) was fully dissolved in 1000 μm of ethanol and then ultrasounded for 20 minutes to obtain a homogeneous solution. For the preparation of each gas sensing material, 5 drops of ZIF-67 solution were uniformly deposited on the surface of self-made silver printing photo paper using a 1 mL disposable syringe, forming an area of $1 \times 1 \text{ cm}^2$ in the center. The sample was dried under a vacuum for 10 minutes for further CO_2 capture measurement.

7.1.3 Sensing Algorithms

7.1.3.1 Optimal Antenna Pair Selection

The proximal sensor has 40 antenna Tx/Rx pair combinations in total, and as our previous study of optimal antenna pair [223] showed, the different antenna pair will affect the measurement results. However, in our case, we don't know which antenna pair performs better since the induced currents may vary by the generated EM waves from different antenna pairs. Additionally, the signal energy of EM waves can represent their reflectance. We use the trapezoidal numerical integration method to calculate signal energy. The approximation for an integration with $N+1$ evenly spaced points is as follows:

$$\begin{aligned} \int_a^b f(x)dx &\approx \frac{b-a}{2N} \sum_{n=1}^N (f(x_n) + f(x_{n+1})) \\ &= \frac{b-a}{2N} [f(x_1) + 2f(x_2) + \dots + 2f(x_N) + f(x_{N+1})], \end{aligned} \quad (7.1)$$

where the spacing between each point is $\frac{b-a}{N}$.

7.1.3.2 Savitzky-Golay (SG) Smoothing Filter

We established a relationship between captured signal energy and gas concentration in section 7.1.4.1, and the amplitude of the signal is our most concerning data. In the signal processing domain, the FFT is a powerful tool to transform the time domain signal to the frequency domain in order to observe the distribution of amplitude. However, reducing unknown noise and extracting useful signals from raw signals is still a challenge. As a result, we decided to tackle this challenge by using the RMS (Root Mean Square) method and the SG smoothing filter [328].

The SG filter is a digital filter that can be applied to a set of data to smooth the data and improve the accuracy of the data without changing the trend of the signal. Convolution implements this filter by fitting a continuous subset of adjacent data points to a low-order polynomial using linear least squares. Arbitrarily selected different window widths can be at any position, which is the advantage of the SG filter.

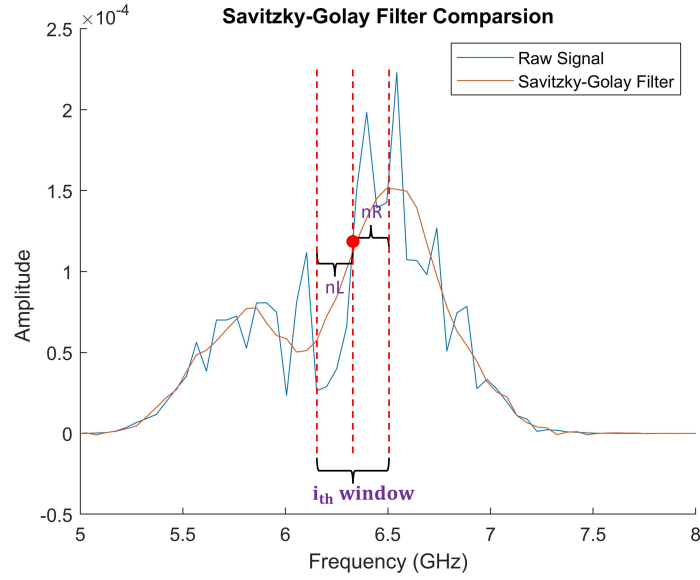


Figure 7.4: Savitzky-Golay smoothing filtering.

As shown in Figure 7.4, the nL and nR are selected sampling point sets from raw data to fit a polynomial model, where in our case, the parameters are 3 (order) and 21 (frame length), respectively.

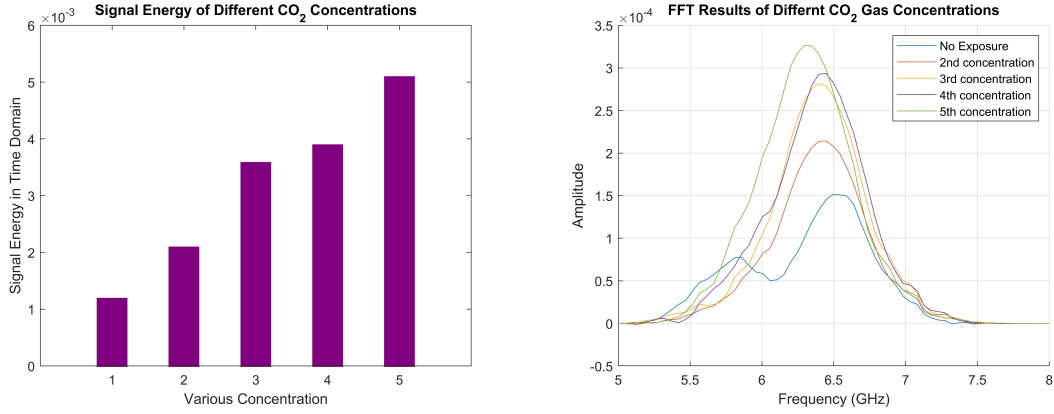
7.1.4 Experiments and Results

Our proposed sensing system has been tested through deployments of real-world experiments: 1) we identified the different CO_2 concentrations in time and frequency domains, respectively; 2) we tested the maximum sensing range of our sensing system.

7.1.4.1 Classification of Different GHG Concentration

The difference in CO_2 concentrations at the PPM level can be seen in Figure 7.5a. We can also see how they change by sending and receiving radar signals from our proximal sensor. The first bar is always no exposure (exposure in an ambient air environment where CO_2 is approximately 410 ppm⁶) and as for the rest of the bar plot from 2 to 5, we exposed our gas film to different CO_2 concentrations, which are {570, 765, 793, 900} PPM, and each exposure lasted 50 s. We obtained the amplitude that can represent the difference in CO_2 concentrations shown in Figure 7.5b.

⁶<https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide>



(a) Variation of CO₂ concentrations in time domain. (b) Variation of CO₂ concentrations in frequency domain.

Figure 7.5: a) The signal energy variations of different CO₂ concentrations in time domain. b) The amplitude of different CO₂ concentrations in frequency domain.

In addition, the frequency amplitude of the raw radar signal has a linear positive correlation with different concentrations. The higher the gas concentration is, the larger amplitude is. We also verified our observation by deploying 5 times independent environments within 50 s of exposure time. The adsorption of CO₂ is linear, which means that our sensing range could reach up to an average around 1500 PPM, in which case the sensing system needs to do calibration first (setup no exposure (ambient environment) amplitude as the 410 PPM). The resolution of amplitude per 100 PPM is shown in Figure 7.6. It is difficult to control the same concentrations at every experiment. However, we found that all independent experiments achieved similar resolution from expected sensing results (the blue bar is the ambient air CO₂ concentration) when compared with our groundtruth data.

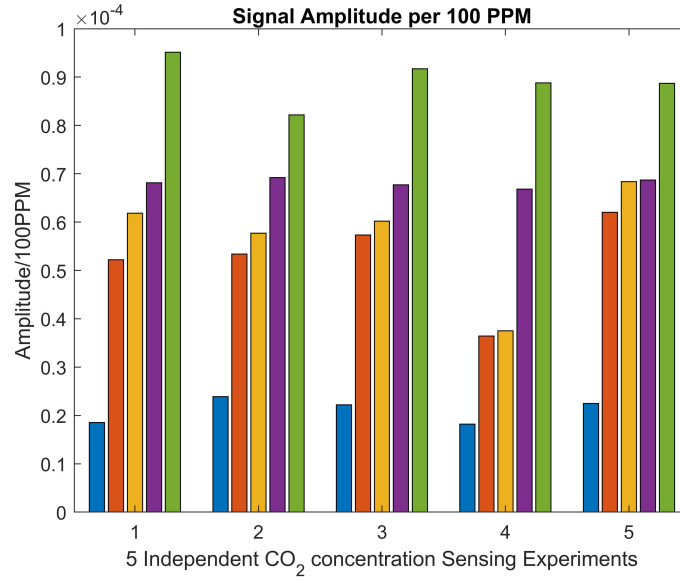
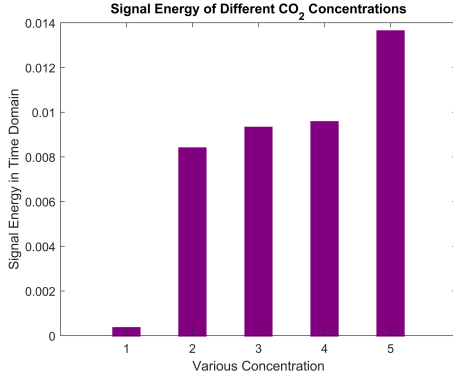


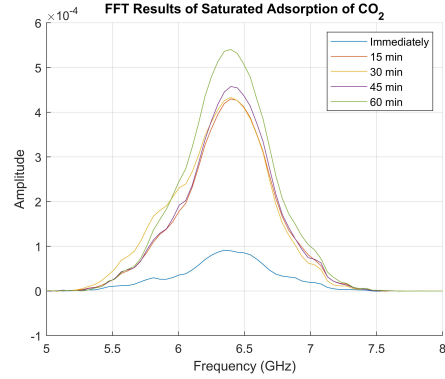
Figure 7.6: The resolution of amplitude after FFT per 100 PPM.

7.1.4.2 Maximum Gas Sensing Range

We also explored the saturated adsorption of our sensing system. As shown in Figure 7.7, we exposed each gas sensing film to the same CO₂ concentrations (4000 PPM) for {0, 15, 30, 45, 60} minutes. After 15 to 30 min, the adsorption of CO₂ of the certainty usage of ZIF-67 from our gas sensing film was temporarily saturated, which is approximately from 2100 PPM to 2300 PPM. Therefore, our sensing system should be able to sense CO₂ at 2200 PPM within 15 min.



(a) Variation of CO₂ concentrations in time domain.



(b) Variation of CO₂ concentrations in frequency domain.

Figure 7.7: a) The signal energy variations of saturated CO₂ concentrations in time domain with different exposure times. b) The amplitude of saturated CO₂ concentrations in frequency domain.

After 45 min, we can see the amplitude is a little higher than the first stage saturation, and the total amount of adsorption for CO₂ which is up to 132.5 mg. In the meantime, the concentration is meaningless since the current amplitude cannot reflect the real concentration. The reason for causing additional adsorption will be explained in Section 7.1.5.

7.1.5 Discussion and Challenges

CO₂ adsorption is primarily a result of physical adsorption and chemical adsorption. As can be seen from Figure 7.7, there are mainly two stages for the adsorption of CO₂. The first stage, where the adsorption starts from 0 to 45 min, is mainly attributed to the chemical adsorption of CO₂ into the pores inside the ZIF-67 structure. This is because the porous ZIF-67 structure made it easier for CO₂ to adsorb into specific binding sites through chemical adsorption, whereas the adsorption mainly occurred inside the structure. In the second stage, the adsorption rate suddenly increased to a very high value after 45 min. Because the adsorption process is limited to the surface of ZIF-67, Van der Waals forces must be in operation for this to occur [329]. The physical adsorption is not selective since the inter-molecular force would happen between all molecules even if we sealed other gases than CO₂ inside the container. So we will mainly focus on the first 45 minutes for the selective adsorption of CO₂.

As a result, we validated that our sensing system would give reliable sensing results within 50 s of exposure time. After conducting 5 independent experiments, our sensing range is from 410 PPM and up to around 1500 PPM, and the resolution is 15 PPM compared with our groundtruth CO₂ sensor. The reasons that could cause

significant variations in resolution are: 1) our groundtruth results are inaccurate because our CO₂ sensor has an offset and is inaccurate; 2) our sensing system needs to do calibration for every sensing activity to get reliable results. Additionally, the gas sensing film should not be exposed to the air for a long time since it absorbs CO₂ from the air. Our GHG sensing system is summarized in Table 7.1.

Table 7.1: Summary of GHG sensing system specification.

Cost	Estimate \$500	proximal Sensor + Gas Sensing Film
Sensitivity	PPM Level	
Resolution	400-1500 PPM in 50 s	2600 PPM in 15 min (Saturated under fixed gas sensing film)
Portable	Yes, hand-sized. easy equipped on UAV as a payload option	

To increase our sensing range and sensing sensitivity, the material coverage and silver paste area need to be enlarged as much as possible. Based on the previous analysis, the adsorption of CO₂ would be affected by the usage of ZIF-67 and its total surface area, and the silver paste area would affect the induced current as well. However, simply increasing these two components would raise relatively high cost (currently, the cost of one gas sensing film is worth no greater than \$0.2), which would be a trade-off of our sensing system. In the future, we will investigate potential solutions for reducing costs and increasing the sensing range.

7.2 A Proximal Point Sensing System for Mapping Soil Moisture Using A Miniaturized Spectrometer

7.2.1 Introduction

Surface soil moisture is referred to as the volume of water content (VWC) in the uppermost layer of soil, typically up to a few centimeters deep. This layer of soil is the first to interact with precipitation, irrigation, and evaporation, making it an essential part of the water cycle. Surface soil moisture plays a significant role in determining the rate of evapotranspiration, which is the combination of transpiration from plants and evaporation from the land surface. For many plants, seeds germinate and young roots develop in the soil’s uppermost stratum. Consequently, the moisture content of this stratum can have a significant effect on plant growth and agricultural output. Additionally, moist surface soil is more susceptible to wind and water erosion. Understanding surface soil moisture is necessary for soil erosion prediction and management. However, it is still a challenge because it lacks a high-precision sensing method for measuring surface soil moisture in situ and in real time.

A spectrometer is a scientific instrument used to measure the spectral response of a physical phenomenon. The term “spectrometer” is a broad term often used to describe instruments that measure a continuous variable of a phenomenon where the spectral components are somehow mixed. Spectrum usually refers to a plot of light intensity or power as a function of frequency or wavelength, also known as a

spectral density plot. In other words, almost every material has its own spectrum response that can be used to identify its signature, and surface soil moisture can also have its own spectrum response [330] which can be used as identification to classify different levels of VWC.

Many studies utilize spectrum response to measure soil moisture, such as hyperspectral imagers on satellites [185], aircraft [188], and UAVs [108, 189–191], and these methods are also referred to as remote sensing. They used regression models to find the correlation between indices such as NDVI or NMDI and soil moisture [131, 331]. Moreover, some researchers used a thermal camera to sense soil moisture through soil temperature [332]. Remote sensing for estimating soil moisture has several advantages, including no negative impact on the soil. Furthermore, it can efficiently scan large-scale fields and gather soil properties in those difficult-to-reach areas. Nevertheless, remote sensing has a low signal-to-noise ratio [195], low spectral resolution [60], and is affected by atmospheric and geometric disturbance [196]. As for the proximal sensing of soil moisture, [131, 135] used microwave and [264] used millimeter wave sensors for soil moisture quantification. They verified the feasibility of using microwave and millimeter waves to measure soil moisture but found a lack of accuracy. Furthermore, it is difficult to scale up their proximal sensing into large-scale fields.

To tackle these challenges, we proposed a sensing system that consists of a proximal spectrometer with a spectral range of 300 - 1000 nm, a Raspberry Pi for spectrum analysis in real-time, and lastly, a SMA905 optical fiber for extending sensing range and location (see Figure 7.8a). This system enables a point sensing mechanism for soil moisture that combined proximal sensing and remote sensing benefits such as in situ, real-time, high visit frequency, etc., for large-scale field surface soil moisture mapping.

As a result, we target two challenges: 1) how to interpolate reasonable measurement data from the spectrum response to those unknown sensing locations, and 2) how to build the relationship between the spectrum response and surface soil moisture. This work has the following contributions:

1. We proposed a point-sensing system for soil moisture using a portable spectrometer by analyzing the spectrum response in real-time.
2. We investigated the performance of sensing surface soil moisture under different spatial distributions.
3. We validated the spectrum absorption response to the surface soil moisture level by comparing it with groundtruth data.

7.2.2 Materials and Methods

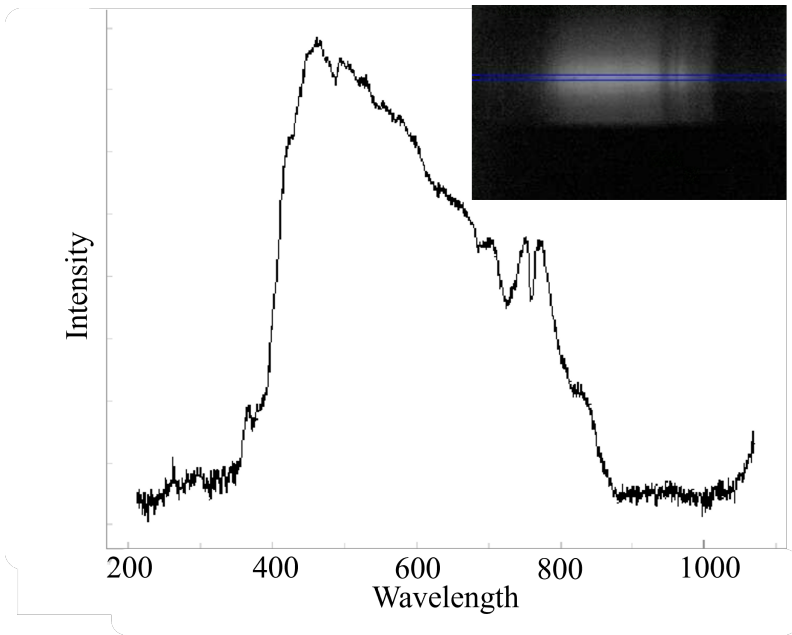
7.2.2.1 Spectrometer Setup

As mentioned earlier, our point sensing system consists of three components: a miniaturized spectrometer, a Raspberry Pi, and an optical fiber for extending its sensing range and location shown in Figure 7.8a. The spectrometer we used is called Pleroma⁷. Spectrochip Inc. developed Pleroma, a spectrometer that integrates an optical MEMS chip and an image sensor. Several studies use this spectrometer to detect COVID-19 virus concentrations [333, 334]. Pleroma guides the light to measure into its entrance port (e.g., through an optical fiber), and the built-in image sensor outputs the spectrum for data acquisition from the CSI port. Pleroma allows for accurate measurement with low noise. Pleroma comes with evaluation software (Python) that allows setting measurement conditions, acquiring and saving data, and displaying graphs. Pleroma takes full advantage of an open-source feature of the Raspberry Pi Single Board Computer (SBC), the most widely distributed SBC, to quickly customize applications. The technical details are shown in Table 7.2.

⁷<http://www.spectrochips.com/>



(a) Spectrometer hardware.



(b) Spectrometer captured.

Figure 7.8: Spectrometer hardware and captured spectrum.

Figure 7.8b shows the spectrum response in real-time at a specific band and their intensity. The image sensor is shown in the upper-right image, revealing how much light it captures. In other words, we could use captured images to feed into

machine learning models for classification, extract information from spectrum responses, and build relationships between absorption peaks and desired sensing targets. The light source we used here is sunlight, enabling a full-band spectrum.

Table 7.2: Technique details of Pleroma spectrometer.

Item	Specification	Unit
Spectral range	300 - 1000	nm
Spectral resolution	5-7	nm
Wavelength reproducible	± 0.375	nm
Spectral Stray light	19.3 (1.17%)	dB
A/D conversion	8	bits
Number of pixels	1280	pixel
Optical solid angle	24	°
Power consumption	158	mW

7.2.2.2 Geospatial Interpolation with Kriging

In geostatistics, it is common to interpret and estimate a set of unknown data locations, $p(s_0)$, from a set of sparse observations, $p(s_i)$, at locations l_0 and l_i . These N sparse sampling points are taken inside a domain $\Omega \in \mathbb{R}^2$ and determine a set of weights, λ_i , to estimate the value at the unknown locations. For the details of Kriging methods, please refer Chapter 2.

$$\hat{p}(s_0) = \sum_{i=1}^N \lambda_i p(s_i). \quad (7.2)$$

In this work, we utilize ordinary Kriging by feeding into absorption peak intensity for interpolating surface soil moisture levels.

7.2.2.3 Savitzky-Golay Smoothing Filter

We found a relationship between absorption magnitude and surface soil moisture, with the amplitude of the signal being our most important piece of information. In the field of signal processing, the FFT is a potent instrument for transforming signals from the time domain to the frequency domain in order to observe amplitude variation. Nevertheless, reducing unknown noise and extracting accurate signal's from unprocessed signals remains a challenge. Thus, we chose to use the RMS and SG smoothing filter [328] to tackle this challenge.

In this study, the SG algorithm has been applied during the sensing phase to choose sample point sets from raw data for fitting a polynomial model. Specifically,

the polynomial model has been chosen to be cubic and the frame length is chosen to be 18.

7.2.3 Experiments and Results

We validated our point sensing system by collecting real-world data. We investigated 1) surface soil moisture at different distributions and 2) the relationship between surface soil moisture and spectrum absorption peaks.

7.2.3.1 Soil Moisture Uniformly Distributions

In this experiment, our main target is to verify the feasibility of interpreting surface soil moisture by spectral response. We proposed five different moisture levels, as shown in Figure 7.10a. We define well-saturated soil as adding water to dry soil after mixing well and seeing liquids on top (our record shows 60g of water reached well saturation). Then we selected the moisture ratio based on well-saturation: 0%, 25%, 50%, 75%, 100%, respectively.

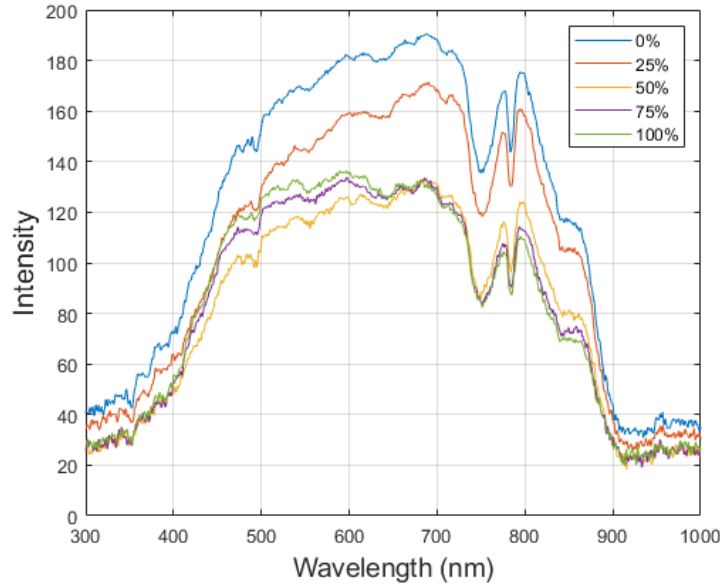


Figure 7.9: Spectrum responses of surface soil moisture levels at 0%, 25%, 50%, 75% and 100%. The y -axis is light intensity, and x -axis is light wavelength.

Figure 7.9 shows the spectral responses of the five different surface soil moisture levels under sunlight exposure. It has two obvious absorption peaks at 750 nm (the first peak) and 785 nm (the second peak). According to the spectral response, the absorption peaks correspond to surface soil moisture levels.

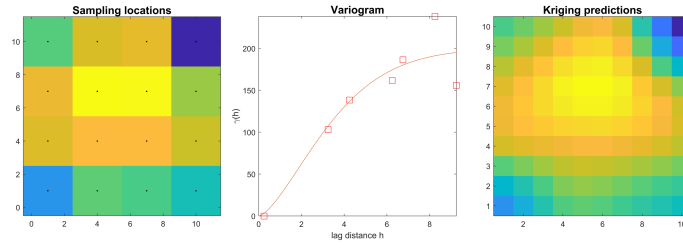


(a) Spectrometer hardware.

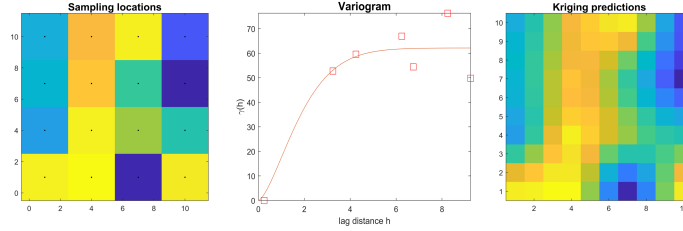


(b) sensing process

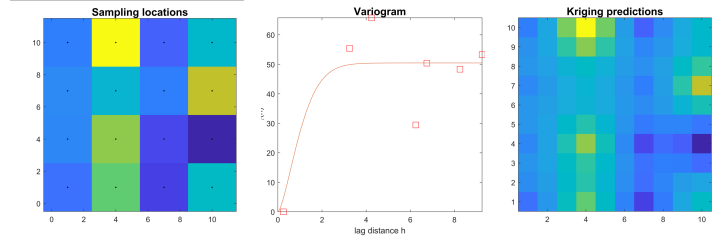
Figure 7.10: Five different surface moisture levels and uniform distributions. (a) The moisture level increased one by one from the left, as the dry level, to the bottom, as the well-saturated level. (b) Sensing processing: SMA905 is capturing the reflectance signals from sunlight; the distance between the optical fiber and the sample is fixed at around 1 cm.



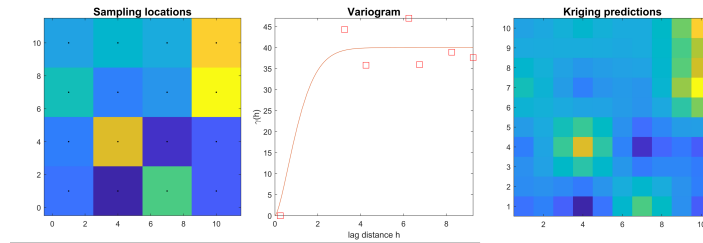
(a) Dry soil surface Kriging results.



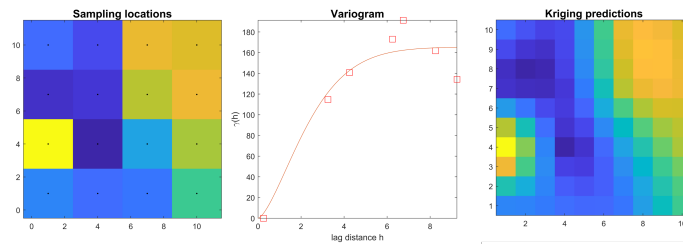
(b) 25% soil moisture surface Kriging results.



(c) 50% soil moisture surface Kriging results.



(d) 75% soil moisture surface Kriging results.



(e) 100% soil moisture surface Kriging results.

Figure 7.11: Kriging results of fitting captured spectrum absorption peak values.

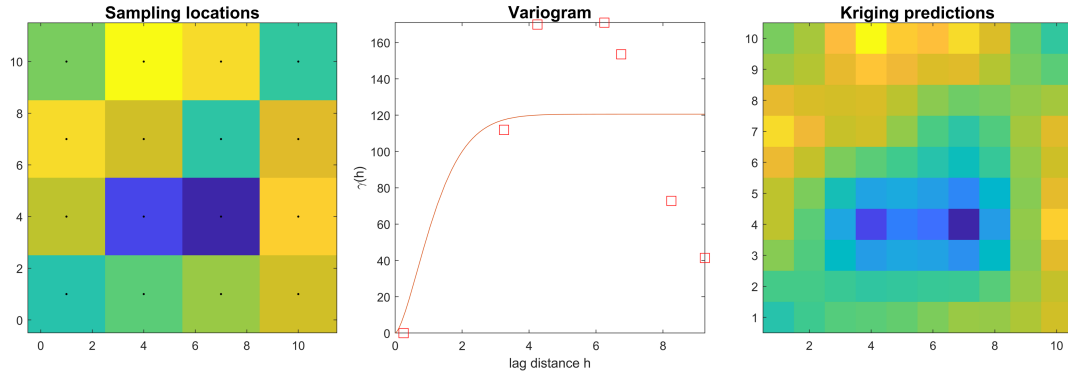
Figure 7.11 reveals the Kriging results of using absorption peak values from dry soil to well-saturated soil. We sampled 16 points (4 by 4) so that we could fit this data into the Kriging method. We present several conclusions: 1) the point sensing results extracted from the spectrum response can indeed reveal the surface soil moisture levels, and 2) the Kriging method demonstrates the surface moisture distributions. In addition, we believe that the Kriging results, when compared to the results of these five experiments, reveal the soil moisture levels in terms of average values. We also observed that different angles and surfaces result in varying reflectance intensities, which in turn reveals that the surface soil moisture distributions may not appear uniformly in the sensing results.

7.2.3.2 Soil Moisture Non-uniformly Distributions

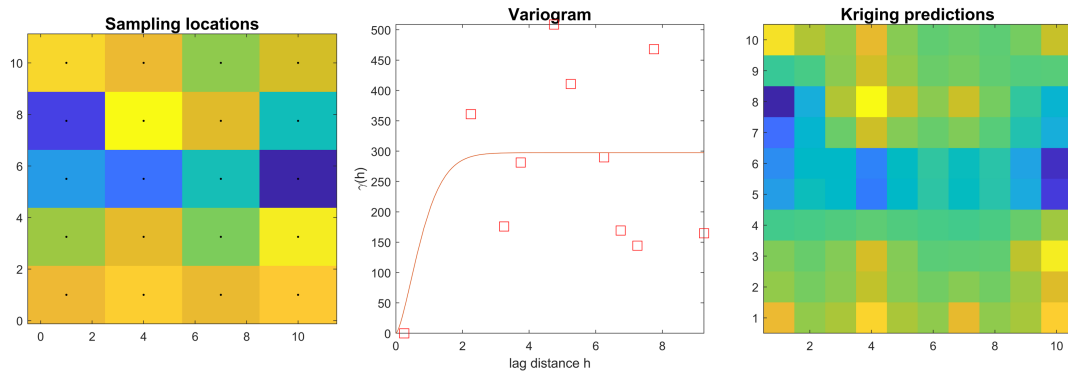
In this section, we explored the performance of the spectrometer to capture non-uniform surface soil moisture distributions. Figure 7.12 demonstrates a total of three different moisture distributions. Figure 7.13 shows the sensing and Kriging results of these three distributions.



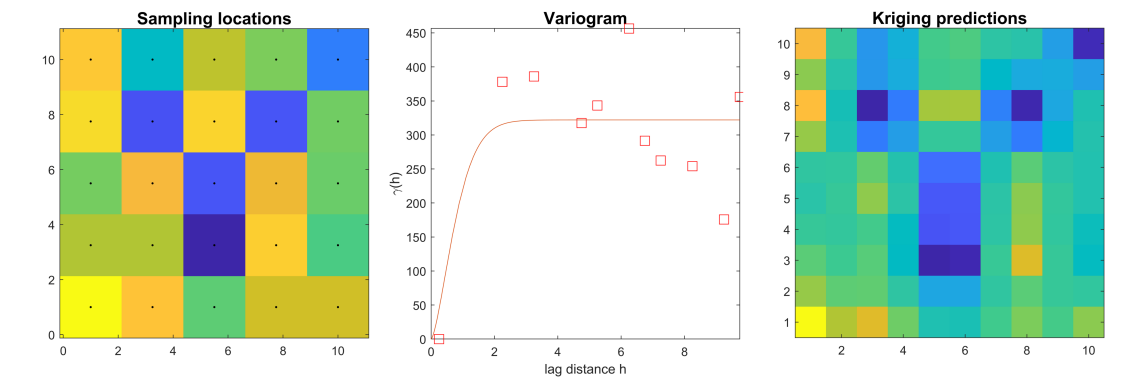
Figure 7.12: Surface soil moisture in non-uniform distributions.



(a) Single locations of soil moisture.



(b) Double locations of soil moisture.



(c) Triple locations of soil moisture.

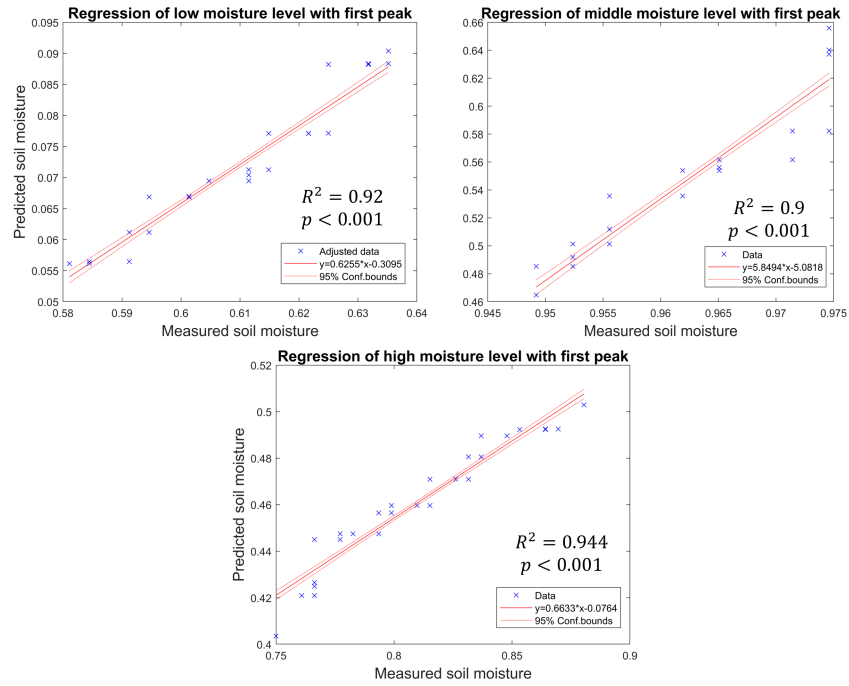
Figure 7.13: Kriging results of non-uniform distributions.

Figure 7.13a took samples from 16 places (4 by 4), and Kriging displays the outcomes on the lower middle position, which is the same as Figure 7.12 for a single location's moisture. Figure 7.13b sampled a total of 20 locations (5 by 4) for more details on moisture location. We obtained a lower Kriging value, which indicates that the moisture level is higher than in other places in the boundary area. We

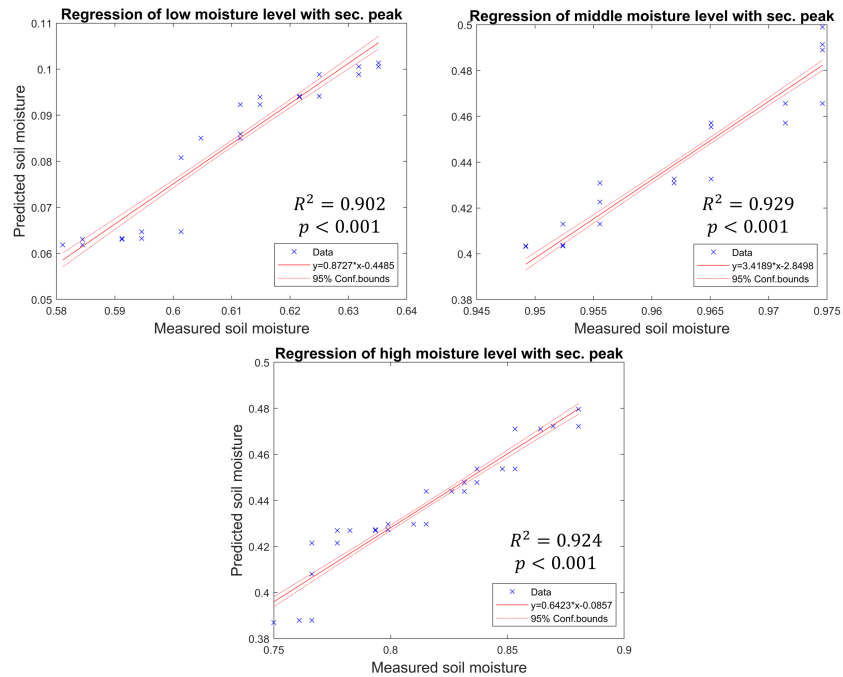
believe this phenomenon stems from the fact that water, when dipped into the soil, absorbs not only the vertical soil but also the neighboring soil. Lastly, Figure 7.13c reveals three moisture locations through sampling 25 locations (5 by 5). Overall, we maintain that our point-sensing system is capable of demonstrating surface soil moisture levels and spatial distributions.

7.2.3.3 Soil Moisture Regression Analysis

We validated further the regression analysis of the spectrum response by simultaneously measuring the surface soil moisture and the spectrum response at the sample site. We sampled 500 points and 100 points of each soil moisture level (in total, we have 5 levels. See Section 7.2.3.1). As we mentioned earlier, we observed that there are two significant absorption peaks shown in Figure 7.9. Thus, we investigated the relationship between surface soil moisture levels and two peaks, respectively. An Arduino board powered a resistive moisture sensor to sample the groundtruth data.



(a) Regression analysis of the first peak (750nm) with soil moisture level.



(b) Regression analysis of the second peak (785nm) with soil moisture level.

Figure 7.14: Regression analysis of spectrum absorption peaks intensity with surface soil moisture. (a) the first peak analysis for low, middle, and high moisture, and (b) the second peak analysis for low, middle, and high moisture, respectively.

The relationship between surface soil moisture and two absorption peaks' sizes is shown in Figure 7.14. After fitting into the linear regression model, the average $R^2 = 0.9213$ for the first peak and the average $R^2 = 0.9183$ for the second peak. The first peak can indicate that the surface soil moisture level is slightly better than the first peak, and the low moisture level is less correlated with the spectrum absorption peak. We also obtained the same conclusion from observing Figure 7.9 that the second peak at 785 nm shows a more clear ascending order of moisture levels. To evaluate whether our obtained results have statistical significance since we performed regression model analysis, we used analysis of variance (ANOVA) to illustrate that our data has significance, which means it is reproducible and meaningful in statistical analysis. The Pearson value (p -value) is less than 0.001, which indicates these data are significant and totally different from each other.

7.2.4 Discussion and Challenges

We provided various analyses and experiments in this work to discuss the feasibility of spectrometer applications. Therefore, we would like to address a few issues here.

Dataset collection and spectrometer usage instructions: The light source, sensing angles, and light strength are the essential factors affecting spectrum response. In this work, we conducted all experiments during noon, when the sun shines directly, resulting in a relatively large amount of radiation. In addition, we found several works that use general RGB as a light source. This could work on several materials; however, we recommend that users better investigate the spectral response of the target before using an LED light source.

Moreover, to capture signals with the lowest noise, we increased the shutter speed to 1000000 μs (maximum) so that more lights would come into the SMA905 connector, which would show the spectrum response more clearly than other settings.

Feature extraction: Multi-modal employment: Numerous pervasive computing systems and Internet of Things applications utilize multi-modal data. We can use various data formats to store a diverse range of information from various sources. Our information extract method is just the Kriging method for implementing point sensing of surface soil moisture distribution. In the future, we will utilize more than one spectrometer sensor to build a spectral sensing matrix to measure soil moisture, soil carbon content, etc. We would also consider using multi-modal and XAI to autonomously tackle soil carbon sensing tasks (including but not limited to soil carbon content, soil moisture, nutrients, etc.).

7.3 Chapter Summary

In this Chapter, we present several smart applications using XAIoT. A low-cost GHG sensing system that employs a proximal sensor and a low-cost gas sensing film. The system detects the CO₂ concentrations in a contact less approach through

variation of radar signal amplitude after gas sensing film adsorption of CO₂, and we employ a filter and antenna selection algorithm to enhance the sensing accuracy. After conducting a series of experiments, we found that our system sensing range is from 410 PPM up to around 1500 PPM based on our gas sensing film setup and proper calibration. The environment we tested is easy to expand to a large field so that GHG emissions from soil or compost can be measured. In the future, we will collect the soil CO₂ groundtruth data to calibrate our sensing system, which will provide much more reliable sensing results for targeting soil carbon emissions. Furthermore, we want to set up verification metrics for sensing concentrations and fluxes of GHG emissions in large-scale fields.

Additionally, we present a point sensing system using a pocket-sized spectrometer to sense surface soil moisture distributions for proof concept of how to use a point sensing system. We illustrated the sensing mechanism of using a spectrometer to sense surface soil moisture distributions and also discussed other potential spectrometer-enabled agricultural applications. The point sensing system was validated by conducting real-world data and revealed that surface soil moisture can be identified by spectral absorption response significantly (R^2 is around 0.92 and p value is less than 0.001). The Kriging method has been applied and verified on geospatial interpolations, and it achieves good results.

Chapter 8

CONCLUSIONS AND FUTURE RESEARCH

8.1 Concluding Remarks

This dissertation has articulated the advancement of XAIoT in the research domain of soil carbon analysis and precision agriculture. Across the diverse methodologies and studies presented, a consistent theme has been the utilization of smart sensing and applications to understand and improve soil health and carbon management.

Chapter 1 presented the foundational concepts underscoring the significance of soil carbon and its environmental impacts, introducing the concept of smart sensing and applications within the XAIoT framework. Chapter 2 discussed the XAIoT data acquisition platforms, revealing a classifier based on feature extraction with 74.0% accuracy in identifying biochar amounts, underscoring the substantial water retention capacity of biochar.

The case studies elucidated the potential of XAIoT in practical field scenarios. The use of UAVs in monitoring soil conditions post biochar application, as discussed in case study III, showed that image index analysis could become a fast method for assessing soil conditions. Case study IV introduced a low-cost UGV solution for precision agriculture, specifically pest management, with significant scalability and cost-saving potential.

In Chapter 3, the comparative efficacy of microwave and millimeter waves in sensing soil carbon was presented, indicating a 10–15% classification accuracy improvement of millimeter wave sensing over microwave sensing for soil carbon content. Chapter 4 detailed a DT-enabled methane emission abatement framework using UAVs, illustrating the potential of DT for more effective methane management. The chapter also proposed a SCMDT-enabled soil carbon accounting management system framework, indicating the promise of SCMDT as the future standard for soil carbon management.

Chapter 5 introduced an XAIoT-DT framework for managing soil carbon emissions, leveraging the Shapley value method to enhance user comprehension of soil carbon sensing results.

The FADE-MAS2D platform approach presented in Chapter 6 focused on simulating soil methane emission suppression using swarming UAVs, demonstrating an

optimal control strategy with a PI controller that showed considerable promise for practical methane management.

Finally, Chapter 7 showcased various smart applications utilizing XAIoT, including a low-cost GHG sensing system, and discussed the potential of point sensing systems using spectrometers for soil moisture and SOC estimation.

8.2 Future Research on XAIoT Empowered Smart Sensing of Soil Carbon Content

The future of XAIoT in soil carbon estimation and management is bright with huge potential. Building on the successes of the current research, future work will delve deeper into enhancing the simplicity and accuracy of biochar application and soil condition monitoring in large-scale fields. The integration of mm-wave sensors into UAVs and UGVs can be further pursued to advance the operational ease and precision of sensing systems. Real-world experiments are vital for validating the proposed systems, and a comprehensive evaluation metric for biochar applications is imperative. To our current thinking, the future research efforts plans are suggested in details as follows:

- **Enhanced XAIoT Frameworks:** Development of XAIoT frameworks tailored for dynamic soil analysis, integrating data from multispectral imaging, millimeter wave sensing, and in-situ measurements. Advanced machine learning models will be devised to interpret complex data patterns and provide actionable insights into soil health projects.
- **Enhanced XAI in Soil Carbon Sensing:** Bridging the gap between AI predictions and human interpretability by advancing XAI techniques in soil carbon sensing. This includes refining algorithms for better clarity on model decisions, thereby enhancing trust and reliability in automated soil health assessments.
- **Biochar Characterization:** Delving deeper into the material properties of biochar using XAIoT for a comprehensive understanding of its impact on soil health enhancement, water retention, and carbon sequestration. Explore biochar's various effects based on its origin and treatment processes.
- **Sensing Technology Advancements:** Innovating sensor technologies to accurately quantify soil carbon levels and biochar effectiveness. This involves advancing both proximal and remote sensing methods and improving calibration techniques for more precise soil property measurements.
- **Soil Carbon Measurement Standards:** Establishing standard methodologies for soil carbon measurement using XAIoT, bridging the gap between

research findings and practical applications. Efforts may include the development of universal metrics for soil carbon assessment that can be applied across different landscapes and conditions.

- **Autonomous Sensing Platforms:** Utilizing autonomous UAVs and UGVs equipped with advanced sensing payloads for continuous monitoring of soil conditions. XAI will play a crucial role in processing the collected data, offering explanations for patterns related to soil carbon fluctuations and biochar impact.
- **Integrated Soil Health Management:** Creating an integrated framework for soil health management that combines soil carbon data with other vital parameters such as moisture levels, nutrient content, and microbial activity, leveraging XAIoT for a holistic approach to precision agriculture.

In addition, the cognitive BMS should be further studied to estimate reliable capacity, accounting for factors like working temperature and mission requirements. Furthermore, the AI-enabled edge devices for radar signal processing could be developed to process more complex datasets with advanced machine learning algorithms.

The application of XAI will be expanded to empower UAV and UGV systems with advanced agricultural and environmental insights. This expansion includes refining the control strategies for methane emission suppression and broadening the soil carbon accounting management system to encompass large-scale field implementations.

Lastly, it is necessary to establish verification metrics for sensing concentrations and fluxes of GHG emissions in large-scale fields. This will involve collecting extensive ground truth data to calibrate sensing systems and developing robust models for real-time monitoring and management of methane and CO₂ emissions, thereby contributing significantly to the mitigation of climate change impacts. The potential research gaps in soil carbon and biochar, particularly from the sensing perspective, include the need for real-time, high-resolution data acquisition, and the development of sensors capable of distinguishing between different types of biochar and their effects on soil properties. Addressing these gaps will require interdisciplinary research, combining sensor technology, artificial intelligence, agronomy, and environmental science to achieve a more sustainable agricultural practice.

Appendix A

DATASET CONTRIBUTIONS

- **Dataset:** An, Di; Chen, YangQuan (2023). Microwave and millimeter wave signals reflectance of soil carbon content [Dataset]. Dryad. <https://doi.org/10.6071/M3M092>
- **Dataset:** An, Di; Zhang, Patrick; Chen, YangQuan (2024). Exploring the impact of varied background on quantification of soil carbon content using microwave and millimeter wave signal reflectance [Dataset]. Dryad. <https://doi.org/10.5061/dryad.ghx3ffbx3>

Appendix B

CODE CONTRIBUTIONS

- **Chapters 2 and 3:** The MathWorks Inc. (2024). Deep Learning Toolbox version: R2023b, Natick, Massachusetts: The MathWorks Inc. <https://www.mathworks.com>
- **Chapter 3:** https://www.deeppscn.com/SCN_release_v1.zip
- **Chapters 5 and 4:** <https://andi00123456.github.com>
- **Chapters 6 and 7:** <https://andi00123456.github.com>

BIBLIOGRAPHY

- [1] F. J. Pierce, P. Nowak, Aspects of precision agriculture, *Advances in Agronomy* 67 (1999) 1–85.
- [2] N. Zhang, M. Wang, N. Wang, Precision agriculture—a worldwide overview, *Computers and Electronics in Agriculture* 36 (2-3) (2002) 113–132.
- [3] R. Gebbers, V. I. Adamchuk, Precision agriculture and food security, *Science* 327 (5967) (2010) 828–831.
- [4] J. V. Stafford, Implementing precision agriculture in the 21st century, *Journal of Agricultural Engineering Research* 76 (3) (2000) 267–275.
- [5] U. Shafi, R. Mumtaz, J. García-Nieto, S. A. Hassan, S. A. R. Zaidi, N. Iqbal, Precision agriculture techniques and practices: From considerations to applications, *Sensors* 19 (17) (2019) 3796.
- [6] L. García, L. Parra, J. M. Jimenez, J. Lloret, P. Lorenz, IoT-based smart irrigation systems: An overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture, *Sensors* 20 (4) (2020) 1042.
- [7] R. K. Singh, R. Berkvens, M. Weyn, Agrifusion: An architecture for IoT and emerging technologies based on a precision agriculture survey, *IEEE Access* 9 (2021) 136253–136283.
- [8] V. K. Quy, N. V. Hau, D. V. Anh, N. M. Quy, N. T. Ban, S. Lanza, G. Randazzo, A. Muzirafuti, IoT-enabled smart agriculture: Architecture, applications, and challenges, *Applied Sciences* 12 (7).
- [9] P. Nayak, K. Kavitha, M. Rao, et al., IoT-enabled agricultural system applications, challenges and security issues, in: *Proceedings of IoT and Analytics for Agriculture*, Springer, 2020, pp. 139–163.
- [10] M. De Clercq, A. Vats, A. Biel, Agriculture 4.0: The future of farming technology, *Proceedings of the World Government Summit, Dubai, UAE* (2018) 11–13.

- [11] U. E. P. Agency, EPA draft inventory of U.S. greenhouse gas emissions and sinks: 1990-2022, Tech. rep., United States Environmental Protection Agency (2024).
- [12] NSF, Smart and Autonomous Systems (S&AS) (2018).
URL <https://new.nsf.gov/funding/opportunities/smart-autonomous-systems-sas/505325/nsf18-557/solicitation>
- [13] M. Van Lent, W. Fisher, M. Mancuso, An explainable artificial intelligence system for small-unit tactical behavior, in: Proceedings of the National Conference on Artificial Intelligence, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2004, pp. 900–907.
- [14] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbadó, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, *Information Fusion* 58 (2020) 82–115.
- [15] A. Adadi, M. Berrada, Peeking inside the black-box: A survey on explainable artificial intelligence (XAI), *IEEE Access* 6 (2018) 52138–52160.
- [16] Z. C. Lipton, The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery., *Queue* 16 (3) (2018) 31–57.
- [17] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, L. Kagal, Explaining explanations: An overview of interpretability of machine learning, in: Proceedings of 2018 IEEE 5th International Conference on data science and advanced analytics (DSAA), IEEE, 2018, pp. 80–89.
- [18] U. Bhatt, A. Xiang, S. Sharma, A. Weller, A. Taly, Y. Jia, J. Ghosh, R. Puri, J. M. Moura, P. Eckersley, Explainable machine learning in deployment, in: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 2020, pp. 648–657.
- [19] A. Richard, B. Mayag, F. Talbot, A. Tsoukias, Y. Meinard, Transparency of classification systems for clinical decision support, in: Proceedings of International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, 2020, pp. 99–113.
- [20] U. Bhatt, M. Andrus, A. Weller, A. Xiang, Machine learning explainability for external stakeholders, *arXiv preprint arXiv:2007.05408* (2020) n.

- [21] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, *ACM Computing Surveys (CSUR)* 51 (5) (2018) 1–42.
- [22] A. Birhane, Algorithmic injustice: A relational ethics approach, *Patterns* 2 (2) (2021) 100205.
- [23] D. Slack, S. Hilgard, E. Jia, S. Singh, H. Lakkaraju, Fooling lime and shap: Adversarial attacks on post hoc explanation methods, in: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2020, pp. 180–186.
- [24] T. Miller, P. Howe, L. Sonenberg, Explainable AI: Beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences, *arXiv preprint arXiv:1712.00547* (2017) n. pag.
- [25] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, H. Müller, Causability and explainability of artificial intelligence in medicine, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9 (4) (2019) e1312.
- [26] X. Dai, M. T. Keane, L. Shalloo, E. Ruelle, R. M. Byrne, Counterfactual explanations for prediction and diagnosis in XAI, in: *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 2022, pp. 215–226.
- [27] F. Doshi-Velez, B. Kim, Towards a rigorous science of interpretable machine learning, *arXiv preprint arXiv:1702.08608* (2017) n. pag.
- [28] K. Draper, H.-P. Schmidt, the Biochar Journal, biochar-journal.org, 2016.
- [29] A. Demirbas, Pyrolysis mechanisms of biomass materials, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 31 (13) (2009) 1186–1193.
- [30] G. G. Pohlman, Methods of counting the number of legume bacteria in the soil, *Agronomy Journal* 23 (1931) 70–77.
- [31] R. H. Coats, Enumeration and sampling in the field of the census, *Journal of the American Statistical Association* 26 (1931) 270–284.
- [32] K. Kumada, K. Aizawa, The infrared absorption spectra of soil components, *Soil Science and Plant Nutrition* 4 (1959) 181–188.
- [33] R. L. Mitchell, R. O. Scott, The application of spectrochemical methods to agricultural problems, *Applied Spectroscopy* 11 (1957) 12 – 6.

- [34] R. C. Blinn, F. A. Gunther, M. S. Mulla, Infrared determination of aldrin and dieldrin in aldrin-treated soil, *Journal of Economic Entomology* 53 (1960) 1129–1131.
- [35] H. L. Mathews, R. L. Cunningham, J. E. Cipra, T. R. West, Application of multispectral remote sensing to soil survey research in southeastern Pennsylvania, *Soil Science Society of America Journal* 37 (1973) 88–93.
- [36] T. J. Schmugge, J. M. Meneely, A. Rango, R. A. Neff, Satellite microwave observations of soil moisture variations, *Journal of The American Water Resources Association* 13 (1977) 265–282.
- [37] D. T. Lewis, P. M. Seevers, J. V. Drew, Use of satellite imagery to delineate soil associations in the sand hills region of Nebraska, *Soil Science Society of America Journal* 39 (1975) 330–335.
- [38] P. Anuta, S. Kristof, D. Levandowski, T. Phillips, R. MacDonald, Crop, soil, and geological mapping from digitized multispectral satellite photography., in: *Proceedings of International Symposium on Remote Sensing of Environment*, 7th, University of Michigan, 1971, p. n. pag.
- [39] J. L. Davis, A. P. Annan, Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy, *Geophysical Prospecting* 37 (1989) 531–551.
- [40] C. C. Truman, H. F. Perkins, L. E. Asmussen, H. D. Allison, Using ground-penetrating radar to investigate variability in selected soil properties, *Journal of Soil and Water Conservation* 43 (1988) 341–345.
- [41] R. W. M. Johnson, R. Glasscum, R. J. Wojtasinski, Application of ground penetrating radar to soil survey, *Soil Horizons* 23 (1982) 17–25.
- [42] J. Doolittle, Using ground-penetrating radar to increase the quality and efficiency of soil surveys, *Soil Survey Techniques* 20 (1987) 11–32.
- [43] G. Vellidis, M. C. Smith, D. L. Thomas, L. E. Asmussen, Detecting wetting front movement in a sandy soil with ground-penetrating radar, *Transactions of the ASABE* 33 (1990) 1867–1874.
- [44] L. Løvborg, L. Bøtter-Jensen, P. Kirkegaard, Experiences with concrete calibration sources for radiometric field instruments, *Geophysics* 43 (1978) 543–549.
- [45] S. L. Ustin, B. N. Rock, R. A. Woodward, Analysis of substrate and plant spectral features of semi-arid shrub communities in the Owens Valley, California (1986).

- [46] I. Wendt, The hentiesbaai range for Airborne Gamma spectrometer calibration, in: *Proceedings of 5th SAGA Biennial Conference and Exhibition*, European Association of Geoscientists & Engineers, 1997, pp. cp–223.
- [47] T. Behrens, K. Schmidt, A.-X. Zhu, T. Scholten, The ConMap approach for terrain-based digital soil mapping, *European Journal of Soil Science* 61 (1) (2010) 133–143.
- [48] J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, S. Kienast-Brown, *Digital soil mapping: Bridging research, environmental application, and operation*, Springer Science & Business Media, 2010.
- [49] C. Brungard, J. Boettinger, Conditioned latin hypercube sampling: Optimal sample size for digital soil mapping of arid rangelands in Utah, USA, *Digital Soil Mapping: Bridging Research, Environmental Application, and Operation* (2010) 67–75.
- [50] S. Kienast-Brown, Z. Libohova, J. Boettinger, *Digital soil mapping*, *Soil Survey Manual* (2017) 295–354.
- [51] A. C. Teodoro, J. Fernandes, P. Santos, L. Duarte, J. A. Gonçalves, D. Flores, Monitoring of soil movement in a self-burning coal waste pile with UAV imagery, in: *Proceedings of Earth Resources and Environmental Remote Sensing/GIS Applications XI*, Vol. 11534, SPIE, 2020, pp. 130–136.
- [52] L.-M. Misthos, L. Papada, G. Panagiotopoulos, N. Gakis, D. Kaliampakos, Estimating climate change-based soil loss using erosion models and uav imagery in the metsovo mountain region, in: *Proceedings of 4th Joint International Symposium on Deformation Monitoring (JISDM)*, 2019, p. 1.
- [53] E. G. M. Pluer, D. T. Robinson, B. U. Meinen, M. L. Macrae, Pairing soil sampling with very-high resolution UAV imagery: An examination of drivers of soil and nutrient movement and agricultural productivity in southern Ontario, *Geoderma* 379 (2020) 114630.
- [54] A. Baghbani, M. D. Nguyen, A. Alnedawi, N. Milne, T. Baumgartl, H. Abuel-Naga, Improving soil stability with alum sludge: An AI-enabled approach for accurate prediction of California bearing ratio, *Applied Sciences* 13 (8) (2023) 4934.
- [55] P. Sultana, A. K. Dey, D. Kumar, Empirical approach for prediction of bearing pressure of spread footings on clayey soil using artificial intelligence (ai) techniques, *Results in Engineering* 15 (2022) 100489.

- [56] Q. Chen, L. Li, C. Chong, X. Wang, AI-enhanced soil management and smart farming, *Soil Use and Management* 38 (2021) 13 – 7.
- [57] Y. Li, S. K. Vanapalli, Prediction of soil-water characteristic curves using two artificial intelligence (AI) models and AI aid design method for sands, *Canadian Geotechnical Journal* 59 (1) (2022) 129–143.
- [58] H. Zayani, Y. Fouad, D. Michot, Z. Kassouk, N. N. Baghdadi, E. Vaudour, Z. Lili-Chabaane, C. Walter, Using machine-learning algorithms to predict soil organic carbon content from combined remote sensing imagery and laboratory Vis-NIR spectral datasets, *Remote Sensing* 15 (2023) 4264.
- [59] B. van Wesemael, S. Chabrillat, A. S. Dias, M. Berger, Z. Szantoi, Remote sensing for soil organic carbon mapping and monitoring, *Remote Sensing* 15 (2023) 3464.
- [60] C. Gomez, R. A. V. Rossel, A. B. McBratney, Soil organic carbon prediction by hyperspectral remote sensing and field Vis-NIR spectroscopy: An Australian case study, *Geoderma* 146 (3-4) (2008) 403–411.
- [61] T. Angelopoulou, N. Tziolas, A. Balafoutis, G. Zalidis, D. Bochtis, Remote sensing techniques for soil organic carbon estimation: A review, *Remote Sensing* 11 (6) (2019) 676.
- [62] R. Sharma, D. R. Mishra, M. R. Levi, L. A. Sutter, Remote sensing of surface and subsurface soil organic carbon in Tidal wetlands: A review and ideas for future research, *Remote Sensing* 14 (2022) 2940.
- [63] S. Singh, Remote sensing applications in soil survey and mapping: A review, *International Journal of Geomatics and Geosciences* 7 (2) (2016) 192–203.
- [64] K. C. L. Wong, L. J. Klein, A. F. da Silva, H. Wang, J. Singh, T. F. Syeda-Mahmood, Image-based soil organic carbon remote sensing from satellite images with fourier neural operator and structural similarity, *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium* (2023) 1305–1308.
- [65] T. G. Morais, M. Jongen, C. Tufik, N. R. Rodrigues, I. Gama, J. Serrano, M. C. Gonçalves, R. Mano, T. Domingos, R. F. M. Teixeira, Satellite-based estimation of soil organic carbon in Portuguese grasslands, *Frontiers in Environmental Science* (2023) n. pag.
- [66] C. Sothe, A. Gonsamo, J. Arabian, J. Snider, Large scale mapping of soil organic carbon concentration with 3D machine learning and satellite observations, *Geoderma* 405 (2022) 115402.

- [67] T. Chan, C. A. Gomez, A. Kothikar, P. Baiz, Joint study of above ground biomass and soil organic carbon for total carbon estimation using satellite imagery in scotland, *ArXiv abs/2205.04870* (2022) n.pag.
- [68] L. Guo, P. Fu, T. Shi, Y. Chen, H. Zhang, R. Meng, S. Wang, Mapping field-scale soil organic carbon with unmanned aircraft system-acquired time series multispectral images, *Soil & Tillage Research* 196 (2020) 104477.
- [69] D. Sullivan, J. Fulmer, T. Strickland, M. Masters, H. Yao, Field scale evaluation of crop residue cover distribution using Airborne and satellite remote sensing, in: *Proceedings of the Georgia Water Resources Conference*, held March, 2007, pp. 27–29.
- [70] D. Sullivan, J. Shaw, P. Mask, D. Rickman, J. Luvall, J. Wersinger, E. Guertal, Rapid assessment of in situ wheat straw residue via remote sensing platforms, *NASA Technial Reports* (2003) n. pag.
- [71] J. K. M. Biney, M. Saberioon, L. Borůvka, J. Houška, R. Vašát, P. Chapman Agyeman, J. A. Coblinski, A. Klement, Exploring the suitability of UAS-based multispectral images for estimating soil organic carbon: Comparison with proximal soil sensing and Spaceborne imagery, *Remote Sensing* 13 (2) (2021) 308.
- [72] A. Mondal, D. Khare, S. Kundu, S. Mondal, S. Mukherjee, A. Mukhopadhyay, Spatial soil organic carbon (SOC) prediction by regression kriging using remote sensing data, *The Egyptian Journal of Remote Sensing and Space Science* 20 (1) (2017) 61–70.
- [73] E. Vaudour, C. Gomez, Y. Fouad, P. Lagacherie, Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems, *Remote Sensing of Environment* 223 (2019) 21–33.
- [74] F. Castaldi, A. Hueni, S. Chabrillat, K. Ward, G. Buttafuoco, B. Bomans, K. Vreys, M. Brell, B. van Wesemael, Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands, *ISPRS Journal of Photogrammetry and Remote Sensing* 147 (2019) 267–282.
- [75] J. A. M. Demattê, A. C. Dotto, L. G. Bedin, V. M. Sayão, A. B. e Souza, Soil analytical quality control by traditional and spectroscopy techniques: Constructing the future of a hybrid laboratory for low environmental impact, *Geoderma* 337 (2019) 111–121.
- [76] B. C. Gallo, J. A. M. Demattê, R. Rizzo, J. L. Safanelli, W. de Sousa Mendes, I. F. Lepsch, M. V. Sato, D. J. Romero, M. P. C. Lacerda, Multi-temporal

satellite images on topsoil attribute quantification and the relationship with soil classes and geology, *Remote Sensing* 10 (2018) 1571.

- [77] A. Gholizadeh, D. Žižala, M. Saberioon, L. Borůvka, Soil organic carbon and texture retrieving and mapping using proximal, Airborne and Sentinel-2 spectral imaging, *Remote Sensing of Environment* 218 (2018) 89–103.
- [78] S. Mirzaee, S. Ghorbani-Dashtaki, J. Mohammadi, H. Asadi, F. Asadzadeh, Spatial variability of soil organic matter using remote sensing data, *Catena* 145 (2016) 118–127.
- [79] F. Castaldi, A. Palombo, F. Santini, S. Pascucci, S. Pignatti, R. Casa, Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon, *Remote Sensing of Environment* 179 (2016) 54–65.
- [80] A. Steinberg, S. Chabrillat, A. Stevens, K. Segl, S. Foerster, Prediction of common surface soil properties based on Vis-NIR Airborne and simulated en-MAP imaging spectroscopy data: Prediction accuracy and influence of spatial resolution, *Remote Sensing* 8 (2016) 613.
- [81] A. Stevens, T. Udelhoven, A. Denis, B. Tychon, R. Liroy, L. Hoffmann, B. Van Wesemael, Measuring soil organic carbon in croplands at regional scale using Airborne imaging spectroscopy, *Geoderma* 158 (1-2) (2010) 32–45.
- [82] C. Hbirkou, S. Pätzold, A.-K. Mahlein, G. Welp, Airborne hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale, *Geoderma* 175 (2012) 21–28.
- [83] M. H. D. Franceschini, M. H. D. Franceschini, J. A. M. Demattê, D. F. S. Terra, L. E. Vicente, H. M. Bartholomeus, D. C. S. Filho, Prediction of soil properties using imaging spectroscopy: Considering fractional vegetation cover to improve accuracy, *International Journal of Applied Earth Observation and Geoinformation* 38 (2015) 358–370.
- [84] F. Castaldi, S. Chabrillat, A. Jones, K. Vreys, B. Bomans, B. van Wesemael, Soil organic carbon estimation in croplands by hyperspectral remote APEX data using the LUCAS topsoil database, *Remote Sensing* 10 (2018) 153.
- [85] A. Stevens, B. van Wesemael, H. M. Bartholomeus, D. Rosillon, B. Tychon, E. Ben-Dor, Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils, *Geoderma* 144 (2007) 395–404.

- [86] H. M. Bartholomeus, L. Kooistra, A. Stevens, M. van Leeuwen, B. van Wesemael, E. Ben-Dor, B. Tychon, Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy, *International Journal of Applied Earth Observation and Geoinformation* 13 (2011) 81–88.
- [87] E. Vaudour, J. M. Gilliot, L. Bel, J. Lefevre, K. Chehdi, Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra, *International Journal of Applied Earth Observation and Geoinformation* 49 (2016) 24–38.
- [88] J. Peón, C. Recondo, S. Fernández, J. F. Calleja, E. D. Miguel, L. Carretero, Prediction of topsoil organic carbon using Airborne and satellite hyperspectral imagery, *Remote Sensing* 9 (2017) 1211.
- [89] L. Homolová, M. E. Schaepman, P. Lamarque, J. G. Clevers, F. de Bello, W. Thuiller, S. Lavorel, Comparison of remote sensing and plant trait-based modeling to predict ecosystem services in subalpine grasslands, *Ecosphere* 5 (8) (2014) 1–29.
- [90] M. Vohland, M. Ludwig, S. Thiele-Bruhn, B. Ludwig, Quantification of soil properties with hyperspectral data: Selecting spectral variables with different methods to improve accuracies and analyze prediction mechanisms, *Remote Sensing* 9 (2017) 1103.
- [91] E. Aldana-Jague, G. Heckrath, A. Macdonald, B. van Wesemael, K. V. Oost, UAS-based soil carbon mapping using Vis-NIR (480–1000nm) multi-spectral imaging: Potential and limitations, *Geoderma* 275 (2016) 55–66.
- [92] J. R. England, R. A. Viscarra Rossel, Proximal sensing for soil carbon accounting, *Soil* 4 (2) (2018) 101–122.
- [93] E. Erazo-Mesa, A. Echeverri-Sánchez, J. G. Ramírez-Gil, Advances in hass avocado irrigation scheduling under digital agriculture approach, *Revista Colombiana de Ciencias Hortícolas* 16 (1) (2022) n. pag.
- [94] T. Angelopoulou, A. T. Balafoutis, G. C. Zalidis, D. D. Bochtis, From laboratory to proximal sensing spectroscopy for soil organic carbon estimation—A review, *Sustainability* 12 (2020) 443.
- [95] C. I. D. Rodrigues, L. M. Brito, L. J. Nunes, Soil carbon sequestration in the context of climate change mitigation: A review, *Soil Systems* 7 (3) (2023) 64.
- [96] R. A. V. Rossel, Y. Fouad, C. Walter, Using a digital camera to measure soil organic carbon and iron contents, *Biosystems Engineering* 100 (2008) 149–159.

- [97] T. Li, A. Xia, T. I. McLaren, R. Pandey, Z. Xu, H. Liu, S. Manning, O. Madgett, S. Duncan, P. Rasmussen, F. Ruhnke, O. Yüzügüllü, N. Fajraoui, D. Beniwal, S. Chapman, G. Tsiminis, C. Smith, R. C. Dalal, Y. P. Dang, Preliminary results in innovative solutions for soil carbon estimation: Integrating remote sensing, machine learning, and proximal sensing spectroscopy, *Remote. Sens.* 15 (2023) 5571.
- [98] S. Mohajan, Y. lei Huang, N. Beier, M. Dyck, F. A. Hegmann, A. Bais, A. E. Hussein, Effect of laser wavelength on soil carbon measurements using laser-induced breakdown spectroscopy, *Optics Express* 31 20 (2023) 32335–32349.
- [99] G. Yakubova, A. Kavetskiy, S. A. Prior, H. A. Torbert, Benchmarking the inelastic neutron scattering soil carbon method, *Vadose Zone Journal* 15 (2) (2016) n. pag.
- [100] G. Yakubova, A. Kavetskiy, S. A. Prior, H. A. Torbert, Applying monte-carlo simulations to optimize an inelastic neutron scattering system for soil carbon analysis, *Applied Radiation and Isotopes* 128 (2017) 237–248.
- [101] R. Knight, Ground penetrating radar for environmental applications, *Annual Review of Earth and Planetary Sciences* 29 (1) (2001) 229–255.
- [102] M. Wolfe, D. Huo, H. Ruiz-Guzman, B. Teare, T. Adams, I. Dobрева, M. Everett, M. Bishop, R. Jessup, D. Hays, Ground penetrating radar for below-ground phenotyping of high-biomass grasses for soil carbon sequestration, *Research Square* (Preprint) (2021) n. pag.
- [103] N. E. Silvero, J. A. Demattê, B. Minasny, N. A. Rosin, J. G. Nascimento, H. S. R. Albarracín, H. Bellinaso, A. M. Gómez, Sensing technologies for characterizing and monitoring soil functions: A review, *Advances in Agronomy* 177 (2023) 125–168.
- [104] J. J. McGill, A. Agarwal, The impact of cell phone, laptop computer, and microwave oven usage on male fertility, in: *Male Infertility*, Springer, 2014, pp. 161–177.
- [105] R. N. Khushaba, A. J. Hill, Radar-based materials classification using deep wavelet scattering transform: A comparison of centimeter vs. millimeter wave units, *IEEE Robotics and Automation Letters* 7 (2) (2022) 2016–2022.
- [106] T. Zhao, H. Niu, E. de la Rosa, D. Doll, D. Wang, Y. Chen, Tree canopy differentiation using instance-aware semantic segmentation, in: *Proceedings of 2018 ASABE Annual International Meeting*, American Society of Agricultural and Biological Engineers, 2018, p. 1.

- [107] T. Zhao, Y. Yang, H. Niu, D. Wang, Y. Chen, Comparing U-Net convolutional network with mask R-CNN in the performances of pomegranate tree canopy segmentation, in: *Proceedings of Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII*, Vol. 10780, SPIE, 2018, pp. 210–218.
- [108] D. An, D. Hollenbeck, S. Gao, Y. Chen, A field study of soil biochar treatment response using small unmanned aerial systems (sUAS), in: *Proceedings of 2022 International Conference on Unmanned Aircraft Systems (ICUAS)*, IEEE, 2022, pp. 1608–1613.
- [109] R. Rault, D. Trentesaux, Artificial intelligence, autonomous systems and robotics: legal innovations, in: *Service Orientation in Holonic and Multi-Agent Manufacturing*, Springer, 2018, pp. 1–9.
- [110] R. Oberti, M. Marchi, P. Tirelli, A. Calcante, M. Iriti, E. Tona, M. Hočevár, J. Baur, J. Pfaff, C. Schütz, et al., Selective spraying of grapevines for disease control using a modular agricultural robot, *Biosystems Engineering* 146 (2016) 203–215.
- [111] D. J. Romero, E. Ben-Dor, J. A. Demattê, A. B. e Souza, L. E. Vicente, T. R. Tavares, M. Martello, T. F. Strabeli, P. P. da Silva Barros, P. R. Fiorio, et al., Internal soil standard method for the brazilian soil spectral library: Performance and proximate analysis, *Geoderma* 312 (2018) 95–103.
- [112] P. Kashyap, Machine learning algorithms and their relationship with modern technologies, in: *Machine Learning for Decision Makers*, Springer, 2017, pp. 91–136.
- [113] K. Ghaffarzadeh, H. Zervos, Conductive ink markets 2013-2018: Forecasts, technologies, players, IDTechEx Market Report.
- [114] P. Tokekar, J. Vander Hook, D. Mulla, V. Isler, Sensor planning for a symbiotic UAV and UGV system for precision agriculture, *IEEE Transactions on Robotics* 32 (6) (2016) 1498–1511.
- [115] S. Sankaran, A. Mishra, R. Ehsani, C. Davis, A review of advanced techniques for detecting plant diseases, *Computers and Electronics in Agriculture* 72 (1) (2010) 1–13.
- [116] J. D. Foster, Swarming unmanned aerial vehicles (UAVs): extending marine aviation ground task force communications using UAVs, Tech. rep., Naval Postgraduate School, Monterey–CA–USA (2015).

- [117] S. Bonadies, A. Lefcourt, S. A. Gadsden, A survey of unmanned ground vehicles with applications to agricultural and environmental sensing, in: *Proceedings of Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping*, Vol. 9866, SPIE, 2016, pp. 142–155.
- [118] G. Danoy, M. R. Brust, P. Bouvry, Connectivity stability in autonomous multi-level UAV swarms for wide area monitoring, in: *Proceedings of the 5th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications*, 2015, pp. 1–8.
- [119] M. Saska, T. Baca, J. Thomas, J. Chudoba, L. Preucil, T. Krajník, J. Faigl, G. Loianno, V. Kumar, System for deployment of groups of unmanned micro aerial vehicles in GPS-denied environments using onboard visual relative localization, *Autonomous Robots* 41 (4) (2017) 919–944.
- [120] H. Yu, R. Beard, A vision-based collision avoidance technique for micro air vehicles using local-level frame mapping and path planning, *Autonomous Robots* 34 (1) (2013) 93–109.
- [121] P. Fazli, A. Davoodi, A. K. Mackworth, Multi-robot repeated area coverage, *Autonomous Robots* 34 (4) (2013) 251–276.
- [122] A. F. Winfield, W. Liu, J. Nembrini, A. Martinoli, Modelling a wireless connected swarm of mobile robots, *Swarm Intelligence* 2 (2) (2008) 241–266.
- [123] A. Marjovi, L. Marques, Optimal spatial formation of swarm robotic gas sensors in odor plume finding, *Autonomous Robots* 35 (2) (2013) 93–109.
- [124] M. Kloetzer, C. Belta, Temporal logic planning and control of robotic swarms by hierarchical abstractions, *IEEE Transactions on Robotics* 23 (2) (2007) 320–330.
- [125] N. Cai, J.-X. Xi, Y.-S. Zhong, Swarm stability of high-order linear time-invariant swarm systems, *IET Control Theory & Applications* 5 (2) (2011) 402–408.
- [126] C. C. Cheah, S. P. Hou, J. J. E. Slotine, Region-based shape control for a swarm of robots, *Automatica* 45 (10) (2009) 2406–2411.
- [127] M. Keiluweit, P. S. Nico, M. G. Johnson, M. Kleber, Dynamic molecular structure of plant biomass-derived black carbon (Biochar), *Environmental Science & Technology* 44 (4) (2010) 1247–1253.
- [128] C.-C. Chang, C.-J. Lin, LIBSVM: A library for support vector machines, *ACM Transactions on Intelligent Systems and Technology (TIST)* 2 (3) (2011) 1–27.

- [129] S. Balakrishnama, A. Ganapathiraju, Linear discriminant analysis-a brief tutorial, Institute for Signal and Information Processing 18 (1998) (1998) 1–8.
- [130] D. An, Y. Chen, Microwave and millimeter wave signals reflectance of soil carbon content, Dryad, Dataset, doi:<https://doi.org/10.6071/M3M092>.
- [131] H. Niu, T. Zhao, C. Zuber, J. Vasquez-Mendoza, D. Doll, K. Arnold, Y. Chen, A low-cost stem water potential monitoring method using proximate sensor and Scikit-learn classification algorithms, in: Proceedings of 2020 ASABE Annual International Virtual Meeting, American Society of Agricultural and Biological Engineers, 2020, p. 1.
- [132] H. Niu, Y. Wang, T. Zhao, Y. Chen, A low-cost soil moisture monitoring method by using Walabot and machine learning algorithms, IFAC-PapersOnLine 53 (2) (2020) 15784–15789.
- [133] K. L. S. Gunn, T. W. R. East, The microwave properties of precipitation particles, Quarterly Journal of the Royal Meteorological Society 80 (346) (1954) 522–545.
- [134] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, G. Varoquaux, API design for machine learning software: experiences from the Scikit-learn project, in: Proceedings of ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 2013, pp. 108–122.
- [135] D. An, H. Niu, Y. Chen, A Non-intrusive quantification method for biochar water retention capacity using a portable microwave sensor and machine learning, in: Proceedings of 2021 9th International Conference on Control, Mechatronics and Automation (ICCMA), IEEE, 2021, pp. 152–157.
- [136] K. Anderson, G. Peters, The trouble with negative emissions, Science 354 (6309) (2016) 182–183.
- [137] T. Zhao, Z. Wang, Q. Yang, Y. Chen, Melon yield prediction using small unmanned aerial vehicles, in: Proceedings of Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping II, Vol. 10218, SPIE, 2017, pp. 53–58.
- [138] T. Zhao, A. Koumis, H. Niu, D. Wang, Y. Chen, Onion irrigation treatment inference using a low-cost hyperspectral scanner, in: Proceedings of Multi-spectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII, Vol. 10780, International Society for Optics and Photonics, 2018, p. 107800D.

- [139] H. Niu, T. Zhao, D. Wang, Y. Chen, A UAV resolution and waveband aware path planning for onion irrigation treatments inference, in: Proceedings of 2019 International Conference on Unmanned Aircraft Systems (ICUAS), IEEE, 2019, pp. 808–812.
- [140] R. Näsi, E. Honkavaara, P. Lyytikäinen-Saarenmaa, M. Blomqvist, P. Litkey, T. Hakala, N. Viljanen, T. Kantola, T. Tanhuanpää, M. Holopainen, Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level, *Remote Sensing* 7 (11) (2015) 15467–15493.
- [141] H. Niu, T. Zhao, Y. Chen, Intelligent bugs mapping and wiping (iBMW): An affordable robot-driven robot for farmers, in: Proceedings of 2019 IEEE International Conference on Mechatronics and Automation (ICMA), IEEE, 2019, pp. 397–402.
- [142] T. Zhao, H. Niu, A. Anderson, Y. Chen, J. Viers, A detailed study on accuracy of uncooled thermal cameras by exploring the data collection workflow, in: Proceedings of Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping III, Vol. 10664, SPIE, 2018, pp. 105–112.
- [143] J. Anibal, C. Veiga-Pires, E. Esteves, Effects of spoilage on nitrogen and carbon stable isotopes signatures of the clam, in: INCREASE: Proceedings of the 1st International Congress on Engineering and Sustainability in the XXI Century-INCREASE 2017, Springer, 2018, p. 241.
- [144] P. Rochette, L. Flanagan, E. Gregorich, Separating soil respiration into plant and soil components using analyses of the natural abundance of carbon-13, *Soil Science Society of America Journal* 63 (5) (1999) 1207–1213.
- [145] M. Hanada, H. Koda, K. Onaga, K. Tanaka, T. Okabayashi, T. Itoh, H. Miyazaki, Portable oral malodor analyzer using highly sensitive In_2O_3 gas sensor combined with a simple gas chromatography system, *Analytica Chimica Acta* 475 (1-2) (2003) 27–35.
- [146] D. An, Y. Chen, Application of smart, swarm and small UAV’s for methane emission reduction, in: Proceedings of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol. 85437, American Society of Mechanical Engineers, 2021, p. V007T07A050.
- [147] D. An, Y. Chen, Digital Twin enabled methane emission abatement using networked mobile sensing and mobile actuation, in: Proceedings of 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), 2021, pp. 354–357.

- [148] J. Green, J. Likar, Y. Shprits, Impact of space weather on the satellite industry, *Space Weather* 15 (6) (2017) 804–818.
- [149] R. Heidarian Dehkordi, H. Pelgrum, J. Meersmans, High spatio-temporal monitoring of century-old biochar effects on evapotranspiration through the ET-Look model: a case study with UAV and satellite image fusion based on additive wavelet transform (AWT), *GIScience & Remote Sensing* 59 (1) (2022) 111–141.
- [150] I. Zafeiriou, D. Gasparatos, D. Ioannou, D. Kalderis, I. Massas, Selenium biofortification of lettuce plants (*lactuca sativa* l.) as affected by Se Species, Se Rate, and a Biochar co-application in a Calcareous soil, *Agronomy* 12 (1) (2022) 131.
- [151] Á. Horel, E. Tóth, Changes in the soil–plant–water system due to biochar amendment, *Water* 13 (9) (2021) 1216.
- [152] H. Jin, C. J. Köppl, B. M. Fischer, J. Rojas-Conejo, M. S. Johnson, L. Morillas, S. W. Lyon, A. M. Durán-Quesada, A. Suárez-Serrano, S. Manzoni, et al., Drone-based hyperspectral and thermal imagery for quantifying upland rice productivity and water use efficiency after biochar application, *Remote Sensing* 13 (10) (2021) 1866.
- [153] R. Heidarian Dehkordi, V. Burgeon, J. Fouche, E. Placencia Gomez, J.-T. Cornelis, F. Nguyen, A. Denis, J. Meersmans, Using UAV collected RGB and multispectral images to evaluate winter wheat performance across a site characterized by century-old biochar patches in Belgium, *Remote Sensing* 12 (15) (2020) 2504.
- [154] C. Nzediegwu, S. Prasher, E. Elsayed, J. Dhiman, A. Mawof, R. Patel, Effect of biochar on the yield of potatoes cultivated under wastewater irrigation for two seasons, *Journal of Soil Science and Plant Nutrition* 19 (4) (2019) 865–877.
- [155] H. Wackernagel, Ordinary kriging, in: *Proceedings of Multivariate Geostatistics*, Springer, 2003, pp. 79–88.
- [156] A. A. Gitelson, Y. J. Kaufman, R. Stark, D. Rundquist, Novel algorithms for remote estimation of vegetation fraction, *Remote Sensing of Environment* 80 (1) (2002) 76–87.
- [157] J. B. Reeves, P. Falcao, N. Comerford, Spectral effects of biochar on NIR and mid-IR spectra of soil/char mixtures, in: *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for A Changing World*, 2010, p. 1.

- [158] C. Pickel, F. G. Zalom, D. B. Walsh, N. C. Welch, Efficacy of vacuum machines for *Lygus hesperus* (Hemiptera: Miridae) control in coastal California strawberries, *Journal of Economic Entomology* 87 (6) (1994) 1636–1640.
- [159] C. Pickel, F. Zalom, D. Walsh, N. Welch, et al., Vacuums provide limited *Lygus* control in strawberries, *California Agriculture* 49 (2) (1995) 19–22.
- [160] F. Takeda, W. Janisiewicz, B. Short, T. Leskey, A. Stager, Ultraviolet-C (UV-C) for disease and pest management and automating UV-C delivery technology for strawberry, in: *Proceedings of IX International Strawberry Symposium* 1309, 2021, pp. 533–542.
- [161] W. J. Janisiewicz, F. Takeda, D. M. Glenn, M. J. Camp, W. M. Jurick, Dark period following UV-C treatment enhances killing of *Botrytis cinerea* conidia and controls gray mold of strawberries, *Phytopathology* 106 (4) (2016) 386–394.
- [162] R. McConnell, A. Pacheco Anton, R. Magnotti, Crop duster aviation mechanics: high risk for pesticide poisoning., *American Journal of Public Health* 80 (10) (1990) 1236–1239.
- [163] Z. B. Rivera, M. C. De Simone, D. Guida, Unmanned ground vehicle modelling in Gazebo/ROS-based environments, *Machines* 7 (2) (2019) 42.
- [164] P. Wang, J. Tian, H. Niu, Y. Chen, Smart agricultural in-field service robot: From toy to tool, in: *Proceedings of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 59292, American Society of Mechanical Engineers, 2019, p. V009T12A050.
- [165] J. Tian, H. Niu, P. Wang, Y. Chen, Smart and autonomous farm field scouting service robot as an edge device under \$1000: Challenges and opportunities, in: *Proceedings of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 59292, American Society of Mechanical Engineers, 2019, p. V009T12A055.
- [166] R. Bhat, R. Stamminger, Preserving strawberry quality by employing novel food preservation and processing techniques - Recent updates and future scope - An overview, *Journal of Food Process Engineering* 38 (6) (2015) 536–554.
- [167] T. Van Delm, P. Melis, K. Stoffels, W. Baets, Control of powdery mildew by UV-C treatments in commercial strawberry production, in: *Proceedings of VII International Strawberry Symposium* 1049, 2012, pp. 679–684.

- [168] T. Kanto, K. Matsuura, M. Yamada, T. Usami, Y. Amemiya, UV-B radiation for control of strawberry powdery mildew, in: *Proceedings of VI International Strawberry Symposium* 842, 2008, pp. 359–362.
- [169] W. J. Janisiewicz, F. Takeda, B. Nichols, D. M. Glenn, W. M. Jurick II, M. J. Camp, Use of low-dose UV-C irradiation to control powdery mildew caused by *podosphaera aphanis* on strawberry plants, *Canadian Journal of Plant Pathology* 38 (4) (2016) 430–439.
- [170] N. Johansen, I. Vänninen, D. M. Pinto, A. Nissinen, L. Shipp, In the light of new greenhouse technologies: 2. direct effects of artificial lighting on arthropods and integrated pest management in greenhouse crops, *Annals of Applied Biology* 159 (1) (2011) 1–27.
- [171] M. Zhu, M. Riederer, U. Hildebrandt, UV-C irradiation compromises conidial germination, formation of appressoria, and induces transcription of three putative photolyase genes in the barley powdery mildew fungus, *Blumeria graminis* F. sp. *hordei*, *Fungal Biology* 123 (3) (2019) 218–230.
- [172] M. Forges, M. Bardin, L. Urban, J. Aarouf, F. Charles, Impact of UV-C radiation applied during plant growth on pre-and postharvest disease sensitivity and fruit quality of strawberry, *Plant Disease* 104 (12) (2020) 3239–3247.
- [173] A. I. Darras, G. Tsikaloudakis, I. Lycoskoufis, C. Dimitriadis, D. Karamountas, Low doses of UV-C irradiation affects growth, fruit yield and photosynthetic activity of tomato plants, *Scientia Horticulturae* 267 (2020) 109357.
- [174] L. Urban, F. Charles, M. R. A. de Miranda, J. Aarouf, Understanding the physiological effects of UV-C light and exploiting its agronomic potential before and after harvest, *Plant Physiology and Biochemistry* 105 (2016) 1–11.
- [175] F. Takeda, W. Janisiewicz, B. Smith, B. Nichols, et al., A new approach for strawberry disease control, *European Journal of Horticultural Science* 84 (1) (2019) 3–13.
- [176] J. Sun, W. J. Janisiewicz, F. Takeda, B. Evans, W. M. Jurick II, M. Zhang, L. Yu, P. Chen, Effect of nighttime UV-C irradiation of strawberry plants on phenolic content of fruit: Targeted and non-targeted metabolomic analysis, *Journal of Berry Research* 10 (3) (2020) 365–380.
- [177] C. Hu, L. Zhao, L. Cao, P. Tjan, N. Wang, Steering control based on model predictive control for obstacle avoidance of unmanned ground vehicle, *Measurement and Control* 53 (3-4) (2020) 501–518.

- [178] P. Wang, D. An, N. Chen, Y. Q. Chen, Spatial path tracking controllers for autonomous ground vehicles: Conventional and nonconventional schemes, *Guidance, Navigation and Control* 1 (01) (2021) 2150003.
- [179] K. E. Wenzel, A. Masselli, A. Zell, Automatic take off, tracking and landing of a miniature UAV on a moving carrier vehicle, *Journal of Intelligent & Robotic Systems* 61 (1) (2011) 221–238.
- [180] A. Turner, D. J. Jacob, K. J. Wecht, J. D. Maasakkers, E. Lundgren, A. E. Andrews, S. C. Biraud, H. Boesch, K. W. Bowman, N. M. Deutscher, et al., Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmospheric Chemistry and Physics* 15 (12) (2015) 7049–7069.
- [181] D. An, H. Niu, L. Williams, Y. Chen, Smart Bi-eBikes (SBB): a low cost UGV solution for precision agriculture applications, in: *Proceedings of Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping VII*, Vol. 12114, SPIE, 2022, pp. 80–88.
- [182] M. Giorcelli, A. Khan, A. Tagliaferro, P. Savi, F. Berruti, Microwave characterization of polymer composite based on biochar: A comparison of composite behaviour for biochar and MWCNTs, in: *Proceedings of 2016 IEEE International Nanoelectronics Conference (INEC)*, IEEE, 2016, pp. 1–2.
- [183] S. Gao, T. H. DeLuca, C. C. Cleveland, Biochar additions alter phosphorus and nitrogen availability in agricultural ecosystems: A meta-analysis, *Science of The Total Environment* 654 (2019) 463–472.
- [184] J. Kamieniak, E. P. Randviir, C. E. Banks, The latest developments in the analytical sensing of methane, *TrAC Trends in Analytical Chemistry* 73 (2015) 146–157.
- [185] J. Zhao, W. Huang, D. Zhang, J. Luo, J. Zhang, L. Huang, S. Chen, Characterization and identification of leaf-scale wheat powdery mildew using a ground-based hyperspectral imaging system, *Disaster Adv* 5 (2012) 1657–1662.
- [186] S. Mirzaee, S. Ghorbani-Dashtaki, J. Mohammadi, H. Asadi, F. Asadzadeh, Spatial variability of soil organic matter using remote sensing data, *Catena* 145 (2016) 118–127.
- [187] H. Croft, N. Kuhn, K. Anderson, On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems, *Catena* 94 (2012) 64–74.

- [188] A. Brook, E. B. Dor, Supervised vicarious calibration (SVC) of hyperspectral remote-sensing data, *Remote Sensing of Environment* 115 (6) (2011) 1543–1555.
- [189] A. Capolupo, L. Kooistra, C. Berendonk, L. Boccia, J. Suomalainen, Estimating plant traits of grasslands from UAV-acquired hyperspectral images: A comparison of statistical approaches, *ISPRS International Journal of Geo-Information* 4 (4) (2015) 2792–2820.
- [190] W. Li, Z. Niu, H. Chen, D. Li, M. Wu, W. Zhao, Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system, *Ecological Indicators* 67 (2016) 637–648.
- [191] K. C. McGwire, M. A. Weltz, J. A. Finzel, C. E. Morris, L. F. Fenstermaker, D. S. McGraw, Multiscale assessment of green leaf cover in a semi-arid rangeland with a small unmanned aerial vehicle, *International Journal of Remote Sensing* 34 (5) (2013) 1615–1632.
- [192] D. Beltrán-Marcos, S. Suárez-Seoane, J. M. Fernández-Guisuraga, V. Fernández-García, E. Marcos, L. Calvo, Relevance of UAV and Sentinel-2 data fusion for estimating topsoil organic carbon after forest fire, *Geoderma* 430 (2023) 116290.
- [193] J. K. M. Biney, J. Houška, J. Volánek, D. K. Abebrese, J. Cervenka, Examining the influence of bare soil UAV imagery combined with auxiliary datasets to estimate and map soil organic carbon distribution in an erosion-prone agricultural field, *Science of The Total Environment* 870 (2023) 161973.
- [194] R. Näsi, H. Mikkola, E. Honkavaara, N. Koivumäki, R. A. Oliveira, P. Peltonen-Sainio, N.-S. Keijälä, M. Änäkälä, L. Arkkola, L. Alakukku, Can basic soil quality indicators and topography explain the spatial variability in agricultural fields observed from drone orthomosaics?, *Agronomy* 13 (3) (2023) 669.
- [195] S. Minu, A. Shetty, B. Gopal, Review of preprocessing techniques used in soil property prediction from hyperspectral data, *Cogent Geoscience* 2 (1) (2016) 1145878.
- [196] S. Jakob, R. Zimmermann, R. Gloaguen, The need for accurate geometric and radiometric corrections of drone-borne hyperspectral data for mineral exploration: Mephysto—A toolbox for pre-processing drone-borne hyperspectral data, *Remote Sensing* 9 (1) (2017) 88.

- [197] E. Ben-Dor, K. Patkin, A. Banin, A. Karnieli, Mapping of several soil properties using DAIS-7915 hyperspectral scanner data-a case study over clayey soils in Israel, *International Journal of Remote Sensing* 23 (6) (2002) 1043–1062.
- [198] R. V. Rossel, V. Adamchuk, K. Sudduth, N. McKenzie, C. Lobsey, Proximal soil sensing: An effective approach for soil measurements in space and time, *Advances in Agronomy* 113 (2011) 243–291.
- [199] R. V. Rossel, A. McBratney, Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content, *Geoderma* 85 (1) (1998) 19–39.
- [200] K. Noborio, Measurement of soil water content and electrical conductivity by time domain reflectometry: A review, *Computers and Electronics in Agriculture* 31 (3) (2001) 213–237.
- [201] V. I. Adamchuk, C. R. Hempleman, D. G. Jahraus, On-the-go capacitance sensing of soil water content, in: *Proceedings of Mid-Central Conference, American Society of Agricultural and Biological Engineers*, 2005, p. 1.
- [202] T. J. Dean, J. P. Bell, A. Baty, Soil moisture measurement by an improved capacitance technique, part I. sensor design and performance, *Journal of Hydrology* 93 (1-2) (1987) 67–78.
- [203] R. Paetzold, G. Matzkanin, A. De Los Santos, Surface soil water content measurement using pulsed nuclear magnetic resonance techniques, *Soil Science Society of America Journal* 49 (3) (1985) 537–540.
- [204] D. Daniels, D. Gunton, H. Scott, Introduction to subsurface radar, in: *IEE Proceedings F (Communications, Radar and Signal Processing)*, Vol. 135, IET, 1988, pp. 278–320.
- [205] S. Lambot, J. Rhebergen, I. Van den Bosch, E. Slob, M. Vanclooster, Measuring the soil water content profile of a sandy soil with an off-ground monostatic ground penetrating radar, *Vadose Zone Journal* 3 (4) (2004) 1063–1071.
- [206] J. B. Dixon, S. B. Weed, R. Parpitt, Minerals in soil environments, *Soil Science* 150 (2) (1990) 562.
- [207] E. Ben-Dor, D. Heller, A. Chudnovsky, A novel method of classifying soil profiles in the field using optical means, *Soil Science Society of America Journal* 72 (4) (2008) 1113–1123.

- [208] C. D. Christy, Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy, *Computers and Electronics in Agriculture* 61 (1) (2008) 10–19.
- [209] J. Reeves, G. McCarty, W. Hively, Mid-versus near-infrared spectroscopy for on-site analysis of soil, *Proximal Soil Sensing* (2010) 133–142.
- [210] R. V. Rossel, R. Webster, Predicting soil properties from the Australian soil visible–near infrared spectroscopic database, *European Journal of Soil Science* 63 (6) (2012) 848–860.
- [211] B. Stenberg, R. A. V. Rossel, A. M. Mouazen, J. Wetterlind, Visible and near infrared spectroscopy in soil science, *Advances in Agronomy* 107 (2010) 163–215.
- [212] A. Gianoncelli, J. Castaing, L. Ortega, E. Dooryhee, J. Salomon, P. Walter, J.-L. Hodeau, P. Bordet, A portable instrument for in situ determination of the chemical and phase compositions of cultural heritage objects, *X-Ray Spectrometry: An International Journal* 37 (4) (2008) 418–423.
- [213] P. Sarrazin, D. Blake, S. Feldman, S. Chipera, D. Vaniman, D. Bish, Field deployment of a portable X-ray diffraction/X-ray fluorescence instrument on mars analog terrain, *Powder Diffraction* 20 (2) (2005) 128–133.
- [214] J. Goldberger, G. E. Hinton, S. Roweis, R. R. Salakhutdinov, Neighbourhood components analysis, in: L. Saul, Y. Weiss, L. Bottou (Eds.), in *Proceedings of Advances in Neural Information Processing Systems*, Vol. 17, MIT Press, 2004, p. 1.
- [215] D. Wang, M. Li, Stochastic configuration networks: Fundamentals and algorithms, *IEEE Transactions on Cybernetics* 47 (10) (2017) 3466–3479.
- [216] D. Broomhead, D. Lowe, A. Casadevall, Multivariable functional interpolation and adaptive networks, crisis in infectious diseases-time for a new paradigm, *Complex System* 2 (1996) 312–355.
- [217] Y.-H. Pao, Y. Takefuji, Functional-link net computing: theory, system architecture, and functionalities, *Computer* 25 (5) (1992) 76–79.
- [218] D. Wang, M. Li, Deep stochastic configuration networks with universal approximation property, in: *Proceedings of 2018 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2018, pp. 1–8.
- [219] C. Zhang, S. Ding, L. Guo, J. Zhang, Research progress on stochastic configuration network, *Journal of Software* (2022) 1.

- [220] B. B. Mandelbrot, J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, *SIAM Review* 10 (4) (1968) 422–437.
- [221] J. Viola, D. Hollenbeck, C. Rodriguez, Y. Chen, Fractional-order stochastic extremum seeking control with dithering noise for plasma impedance matching, in: *Proceedings of 2021 IEEE Conference on Control Technology and Applications (CCTA)*, IEEE, 2021, pp. 247–252.
- [222] W. Deng, W. Song, C. Cattani, J. Chen, X. Chen, On the fractional Weibull process, *Frontiers in Physics* (2022) 852.
- [223] D. An, M. Difrieri, Y. Chen, Optimal antenna pairing of a miniaturized radar array for smart sensing of soil carbon content, in: *Proceedings of 2022 4th International Conference on Industrial Artificial Intelligence (IAI)*, IEEE, 2022, pp. 1–6.
- [224] J. Sun, W. Wang, Q. Yue, C. Ma, J. Zhang, X. Zhao, Z. Song, Review on microwave–metal discharges and their applications in energy and industrial processes, *Applied Energy* 175 (2016) 141–157.
- [225] D. L. Corwin, Site-specific management and delineating management zones, in: *Precision Agriculture for Sustainability and Environmental protection*, Routledge, 2013, pp. 153–175.
- [226] F. Pierce, P. Robert, G. Mangold, Site specific management: The pros, the cons, and the realities, in: *Proceedings of the Integrated Crop Management Conference*, Vol. 3, 1994, pp. 17–21.
- [227] S. Houwenhuyse, E. Macke, L. Reyserhove, L. Bulteel, E. Decaestecker, Back to the future in a petri dish: Origin and impact of resurrected microbes in natural populations, *Evolutionary applications* 11 (1) (2018) 29–41.
- [228] T. Lauvaux, N. L. Miles, A. Deng, S. J. Richardson, M. O. Cambaliza, K. J. Davis, B. Gaudet, K. R. Gurney, J. Huang, D. O’Keefe, et al., High-resolution atmospheric inversion of urban CO₂ emissions during the dormant season of the indianapolis flux experiment (INFLUX), *Journal of Geophysical Research: Atmospheres* 121 (10) (2016) 5213–5236.
- [229] Matlab, UAV Toolbox, The MathWorks Inc., 2020.
- [230] Matlab, Web Apps, The MathWorks Inc., 2020.
- [231] Q. Lu, Z. Miao, D. Zhang, L. Yu, W. Ye, S. X. Yang, C.-Y. Su, Distributed leader-follower formation control of nonholonomic mobile robots, *IFAC-PapersOnLine* 52 (15) (2019) 67–72.

- [232] M. Zhang, Y. Shen, Q. Wang, Y. Wang, Dynamic artificial potential field based multi-robot formation control, in: *Proceedings of the 2010 IEEE Instrumentation & Measurement Technology Conference*, IEEE, 2010, pp. 1530–1534.
- [233] Q. Chen, Y. Meng, J. Xing, Shape control of spacecraft formation using a virtual spring-damper mesh, *Chinese Journal of Aeronautics* 29 (6) (2016) 1730–1739.
- [234] L. Ju, Q. Du, M. Gunzburger, Probabilistic methods for centroidal Voronoi tessellations and their parallel implementations, *Parallel Computing* 28 (10) (2002) 1477–1500.
- [235] D. Hollenbeck, Y. Chen, Characterization of ground-to-air emissions with sUAS using a digital twin framework, in: *Proceedings of 2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2020, pp. 1162–1166.
- [236] T. Pflanzner, M. Hovari, I. Vass, A. Kertesz, Designing an IoT-Cloud gateway for the internet of living things, in: *Proceedings of International Conference on Cloud Computing and Services Science*, Springer, 2019, pp. 23–41.
- [237] A. Ghosh, D. Chakraborty, A. Law, Artificial intelligence in internet of things, *CAAI Transactions on Intelligence Technology* 3 (4) (2018) 208–218.
- [238] F. Tao, H. Zhang, A. Liu, A. Y. Nee, Digital twin in industry: State-of-the-art, *IEEE Transactions on Industrial Informatics* 15 (4) (2018) 2405–2415.
- [239] D. Hollenbeck, D. Zulevic, Y. Chen, MOABS/DT: Methane odor abatement simulator with digital twins, in: *Proceedings of 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, IEEE, 2021, pp. 378–381.
- [240] T. W. Nauman, M. C. Duniway, A hybrid approach for predictive soil property mapping using conventional soil survey data, *Soil Science Society of America Journal* 84 (4) (2020) 1170–1194.
- [241] C. Kramer, G. Gleixner, Soil organic matter in soil depth profiles: Distinct carbon preferences of microbial groups during carbon transformation, *Soil Biology and Biochemistry* 40 (2) (2008) 425–433.
- [242] D. An, Y. Chen, Digital twin enabled methane emission abatement using networked mobile sensing and mobile actuation, in: *Proceedings of 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, IEEE, 2021, pp. 354–357.

- [243] H. Mukhtar, R. F. Wunderlich, Y.-P. Lin, Digital twins of the soil microbiome for climate mitigation, *Environments* 9 (3) (2022) 34.
- [244] J. Viola, Y. Chen, Digital twin enabled smart control engineering as an industrial AI: A new framework and case study, in: *Proceedings of 2020 2nd International Conference on Industrial Artificial Intelligence (IAI)*, IEEE, 2020, pp. 1–6.
- [245] J. Baldock, P. Nelson, *Soil organic matter*, CRC Press, 2000.
- [246] U. Stockmann, M. A. Adams, J. W. Crawford, D. J. Field, N. Henakaarchchi, M. Jenkins, B. Minasny, A. B. McBratney, V. d. R. De Courcelles, K. Singh, et al., The knowns, known unknowns and unknowns of sequestration of soil organic carbon, *Agriculture, Ecosystems & Environment* 164 (2013) 80–99.
- [247] A. Sourav, A. Emanuel, Recent trends of big data in precision agriculture: A review, in: *Proceedings of IOP Conference Series: Materials Science and Engineering*, Vol. 1096, IOP Publishing, 2021, p. 012081.
- [248] J. Lehmann, S. Joseph, *Biochar for Environmental Management: Science, Technology and Implementation*, Routledge, 2015.
- [249] R. Ryals, M. Kaiser, M. S. Torn, A. A. Berhe, W. L. Silver, Impacts of organic matter amendments on carbon and nitrogen dynamics in grassland soils, *Soil Biology and Biochemistry* 68 (2014) 52–61.
- [250] Z. Zhai, J. F. Martínez, V. Beltran, N. L. Martínez, Decision support systems for agriculture 4.0: Survey and challenges, *Computers and Electronics in Agriculture* 170 (2020) 105256.
- [251] A. Ribeiro, C. Fernandez-Quintanilla, J. Dorado, F. López-Granados, J. Peña, G. Rabatel, M. Pérez-Ruiz, J. Conesa-Muñoz, P. Santos, A fleet of aerial and ground robots: a scalable approach for autonomous site-specific herbicide application, *Precis Agric* 15 (2015) 167–173.
- [252] F. Liebisch, J. Pfeifer, R. Khanna, P. Lottes, C. Stachniss, T. Falck, S. Sander, R. Siegwart, A. Walter, E. Galceran, Flourish-A robotic approach for automation in crop management, in: *Workshop Computer-Bildanalyse und Unbemannte Autonom Fliegende Systeme in Der Landwirtschaft*, Vol. 21, Wernigerode Harz University Wernigerode, Germany, 2016, p. 2016.
- [253] D. L. Antille, C. R. Lobsey, C. L. McCarthy, J. A. Thomasson, C. P. Baillie, A review of the state of the art in agricultural automation. part IV: Sensor-based nitrogen management technologies, in: *Proceedings of 2018 ASABE Annual*

International Meeting, American Society of Agricultural and Biological Engineers, 2018, p. 1.

- [254] X. Chen, S. K. Phang, M. Shan, B. M. Chen, System integration of a vision-guided UAV for autonomous landing on moving platform, in: Proceedings of 2016 12th IEEE International Conference on Control and Automation (ICCA), IEEE, 2016, pp. 761–766.
- [255] A. Rodriguez-Ramos, C. Sampedro, H. Bavle, P. De La Puente, P. Campoy, A deep reinforcement learning strategy for UAV autonomous landing on a moving platform, *Journal of Intelligent & Robotic Systems* 93 (1) (2019) 351–366.
- [256] A. Rodriguez-Ramos, C. Sampedro, H. Bavle, Z. Milosevic, A. Garcia-Vaquero, P. Campoy, Towards fully autonomous landing on moving platforms for rotary unmanned aerial vehicles, in: Proceedings of 2017 International Conference on Unmanned Aircraft Systems (ICUAS), IEEE, 2017, pp. 170–178.
- [257] T. Baca, P. Stepan, V. Spurny, D. Hert, R. Penicka, M. Saska, J. Thomas, G. Loianno, V. Kumar, Autonomous landing on a moving vehicle with an unmanned aerial vehicle, *Journal of Field Robotics* 36 (5) (2019) 874–891.
- [258] O. Araar, N. Aouf, I. Vitanov, Vision based autonomous landing of multirotor UAV on moving platform, *Journal of Intelligent & Robotic Systems* 85 (2) (2017) 369–384.
- [259] Y. Feng, C. Zhang, S. Baek, S. Rawashdeh, A. Mohammadi, Autonomous landing of a UAV on a moving platform using model predictive control, *Drones* 2 (4) (2018) 34.
- [260] D. Lee, T. Ryan, H. J. Kim, Autonomous landing of a VTOL UAV on a moving platform using image-based visual servoing, in: Proceedings of 2012 IEEE International Conference on Robotics and Automation, IEEE, 2012, pp. 971–976.
- [261] T. Baca, P. Stepan, M. Saska, Autonomous landing on a moving car with unmanned aerial vehicle, in: Proceedings of 2017 European Conference on Mobile Robots (ECMR), IEEE, 2017, pp. 1–6.
- [262] F. Alarcon, D. Santamaria, A. Viguria, UAV helicopter relative state estimation for autonomous landing on moving platforms in a GPS-denied scenario, *IFAC-PapersOnLine* 48 (9) (2015) 37–42.

- [263] T. Yang, Q. Ren, F. Zhang, B. Xie, H. Ren, J. Li, Y. Zhang, Hybrid camera array-based UAV auto-landing on moving UGV in GPS-denied environment, *Remote Sensing* 10 (11) (2018) 1829.
- [264] D. An, Y. Chen, A soil carbon content quantification method using a miniature millimeter wave radar sensor and machine learning, in: *Proceedings of 2022 18th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, IEEE, 2022, pp. 1–6.
- [265] M. Grieves, *Virtually intelligent product systems: Digital and physical twins*, *Complex Systems Engineering: Theory and Practice* (2019) 175–200.
- [266] F. K. Došilović, M. Brčić, N. Hlupić, Explainable artificial intelligence: A survey, in: *Proceedings of 2018 41st International Convention on Information and Communication Technology, electronics and Microelectronics (MIPRO)*, IEEE, 2018, pp. 0210–0215.
- [267] D. An, Y. Chen, A digital twin enabled internet of living things (IoLT) framework for soil carbon management, in: *Proceedings of 2022 18th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, IEEE, 2022, pp. 1–6.
- [268] C. Krupitzer, T. Noack, C. Borsum, Digital food twins combining data science and food science: System model, applications, and challenges, *Processes* 10 (9) (2022) 1781.
- [269] W. Purcell, T. Neubauer, K. Mallinger, Digital twins in agriculture: Challenges and opportunities for environmental sustainability, *Current Opinion in Environmental Sustainability* 61 (2023) 101252.
- [270] P. Bhattacharya, M. S. Obaidat, S. Sanghavi, V. Sakariya, S. Tanwar, K.-F. Hsiao, Internet-of-explainable-digital-twins: A case study of versatile corn production ecosystem, in: *Proceedings of 2022 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, IEEE, 2022, pp. 1–5.
- [271] R. Ferdousi, F. Laamarti, M. A. Hossain, C. Yang, A. El Saddik, Digital twins for well-being: An overview, *Digital Twin* 1 (2022) 7.
- [272] S. P. Ramu, P. Boopalan, Q.-V. Pham, P. K. R. Maddikunta, T. Huynh-The, M. Alazab, T. T. Nguyen, T. R. Gadekallu, Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions, *Sustainable Cities and Society* 79 (2022) 103663.

- [273] X. Xu, Y. Lu, B. Vogel-Heuser, L. Wang, Industry 4.0 and Industry 5.0—inception, conception and perception, *Journal of Manufacturing Systems* 61 (2021) 530–535.
- [274] A. E. Roth, *The Shapley value: essays in honor of Lloyd S. Shapley*, Cambridge University Press, 1988.
- [275] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: *in Proceedings of Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc., 2017, p. 1.
- [276] I. E. Kumar, S. Venkatasubramanian, C. Scheidegger, S. Friedler, Problems with Shapley-value-based explanations as feature importance measures, in: *Proceedings of International Conference on Machine Learning*, PMLR, 2020, pp. 5491–5500.
- [277] P. Liu, C. Lu, C. Wang, X. Wang, W. Xu, Y. Zhou, H. Li, Research status and development of battery management system, in: *Proceedings of International Conference in Communications, Signal Processing, and Systems*, Springer, 2020, pp. 1422–1429.
- [278] K. Liu, K. Li, Q. Peng, C. Zhang, A brief review on key technologies in the battery management system of electric vehicles, *Frontiers of Mechanical Engineering* 14 (1) (2019) 47–64.
- [279] J. Tian, R. Xiong, Q. Yu, Fractional-order model-based incremental capacity analysis for degradation state recognition of lithium-ion batteries, *IEEE Transactions on Industrial Electronics* 66 (2) (2018) 1576–1584.
- [280] Q. Zhong, F. Zhong, J. Cheng, H. Li, S. Zhong, State of charge estimation of lithium-ion batteries using fractional order sliding mode observer, *ISA Transactions* 66 (2017) 448–459.
- [281] P. Shen, M. Ouyang, L. Lu, J. Li, X. Feng, The co-estimation of state of charge, state of health, and state of function for lithium-ion batteries in electric vehicles, *IEEE Transactions on Vehicular Technology* 67 (1) (2017) 92–103.
- [282] H. Mu, R. Xiong, H. Zheng, Y. Chang, Z. Chen, A novel fractional order model based state-of-charge estimation method for lithium-ion battery, *Applied Energy* 207 (2017) 384–393.
- [283] A. N. Eddine, B. Huard, J.-D. Gabano, T. Poinot, A. Thomas, S. Martemianov, Frequential identification of an electrochemical cell impedance using fractional modeling, *IFAC-PapersOnLine* 51 (15) (2018) 802–807.

- [284] X. Li, G. Fan, K. Pan, G. Wei, C. Zhu, G. Rizzoni, M. Canova, A physics-based fractional order model and state of energy estimation for lithium-ion batteries. part I: Model development and observability analysis, *Journal of Power Sources* 367 (2017) 187–201.
- [285] C. Zou, X. Hu, S. Dey, L. Zhang, X. Tang, Nonlinear fractional-order estimator with guaranteed robustness and stability for lithium-ion batteries, *IEEE Transactions on Industrial Electronics* 65 (7) (2017) 5951–5961.
- [286] H. Perez, S. Dey, X. Hu, S. Moura, Optimal charging of Li-ion batteries via a single particle model with electrolyte and thermal dynamics, *Journal of The Electrochemical Society* 164 (7) (2017) A1679.
- [287] Y. Wang, H. Niu, T. Zhao, X. Liao, L. Dong, Y. Chen, Contactless Li-ion battery voltage detection by using Walabot and machine learning, in: *Proceedings of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 59292, American Society of Mechanical Engineers, 2019, p. V009T12A002.
- [288] J. Bhatia, A. Dayal, A. Jha, S. K. Vishvakarma, S. Joshi, M. Srinivas, P. K. Yalavarthy, A. Kumar, V. Lalitha, S. Koorapati, et al., Classification of targets using statistical features from range FFT of mmwave FMCW radars, *Electronics* 10 (16) (2021) 1965.
- [289] G. Agresti, S. Milani, Material identification using RF sensors and convolutional neural networks, in: *Proceedings of ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 3662–3666.
- [290] J. Weiß, A. Santra, One-shot learning for robust material classification using millimeter-wave radar system, *IEEE Sensors Letters* 2 (4) (2018) 1–4.
- [291] J. Weiß, A. Santra, Material classification using 60-GHz radar and deep convolutional neural network, in: *Proceedings of 2019 International Radar Conference (RADAR)*, IEEE, 2019, pp. 1–6.
- [292] F. F. Chamasemani, Y. P. Singh, Multi-class support vector machine (SVM) classifiers – an application in hypothyroid detection and classification, in: *Proceedings of 2011 Sixth International Conference on Bio-Inspired Computing: Theories and Applications*, 2011, pp. 351–356. doi:10.1109/BIC-TA.2011.51.
- [293] R. Conrad, C. Erkel, W. Liesack, Rice cluster I methanogens, an important group of Archaea producing greenhouse gas in soil, *Current Opinion in Biotechnology* 17 (3) (2006) 262–267.

- [294] IPCC Working Group III, Climate Change 2014-Mitigation of Climate Change: Summary for Policymakers, Intergovernmental Panel on Climate Change, 2014.
- [295] T. Stocker, Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2014.
- [296] H. Allcott, M. Greenstone, Is there an energy efficiency gap?, *Journal of Economic Perspectives* 26 (1) (2012) 3–28.
- [297] A. R. Brandt, G. A. Heath, E. A. Kort, F. O’Sullivan, G. Pétron, S. M. Jordaan, P. Tans, J. Wilcox, A. M. Gopstein, D. Arent, S. Wofsy, N. J. Brown, R. Bradley, G. D. Stucky, D. Eardley, R. Harriss, Methane leaks from North American natural gas systems, *Science* 343 (6172) (2014) 733–735.
- [298] K. Haase, G. Casile, R. Kauble, B. Andraski, D. Stonestrom, N. Simon, N. Rybicki, W. Sanford, Data from laboratory and field tests of an economical, environmentally sealed, self-contained pCO₂ sensor for environmental studies, U.S. Geological Survey Data Release (2018) n. pag.
- [299] D. Bastviken, I. Sundgren, S. Natchimuthu, H. Reyier, M. Gålfalk, Cost-efficient approaches to measure carbon dioxide (CO₂) fluxes and concentrations in terrestrial and aquatic environments using mini loggers, *Biogeosciences* 12 (12) (2015) 3849–3859.
- [300] D. Bastviken, J. Nygren, J. Schenk, R. Parellada Massana, N. T. Duc, Facilitating the use of low-cost methane (CH₄) sensors in flux chambers—calibration, data processing, and an open-source make-it-yourself logger, *Biogeosciences* 17 (13) (2020) 3659–3667.
- [301] Jianxiong Cao, Y. Chen, Changpin Li, Multi-UAV-based optimal crop-dusting of anomalously diffusing infestation of crops, in: *Proceedings of 2015 American Control Conference (ACC)*, 2015, pp. 1278–1283.
- [302] Jianxiong Cao, A numerical simulation platform for the control of anomalous diffusion process, in: *MATLAB Central File Exchange*, 2020, p. n. pag.
- [303] C. Li, Q. Yi, J. Kurths, Fractional convection, *Journal of Computational and Nonlinear Dynamics* 13 (2017) n. pag.
- [304] I. Podlubny, An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications, *Math. Sci. Eng* 198 (1999) 340.

- [305] Y. Chen, Z. Wang, J. Liang, Optimal dynamic actuator location in distributed feedback control of a diffusion process, *International Journal of Sensor Networks* 2 (3-4) (2007) 169–178.
- [306] YangQuan Chen, Zhongmin Wang, K. L. Moore, Optimal spraying control of a diffusion process using mobile actuator networks with fractional potential field based dynamic obstacle avoidance, in: *Proceedings of 2006 IEEE International Conference on Networking, Sensing and Control*, 2006, pp. 107–112.
- [307] YangQuan Chen, Zhongmin Wang, Jinsong Liang, Actuation scheduling in mobile actuator networks for spatial-temporal feedback control of a diffusion process with dynamic obstacle avoidance, in: *Proceedings of IEEE International Conference Mechatronics and Automation*, 2005, Vol. 2, 2005, pp. 752–757 Vol. 2. doi:10.1109/ICMA.2005.1626644.
- [308] J. Cortes, S. Martinez, T. Karatas, F. Bullo, Coverage control for mobile sensing networks, *IEEE Transactions on Robotics and Automation* 20 (2) (2004) 243–255. doi:10.1109/TRA.2004.824698.
- [309] A. Howard, M. J. Matarić, G. S. Sukhatme, Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem, in: *Proceedings of Distributed Autonomous Robotic Systems 5*, Springer, 2002, pp. 299–308.
- [310] A. A. Zhokh, P. E. Strizhak, Investigation of the time-dependent transitions between the time-fractional and standard diffusion in a hierarchical porous material, *Transport in Porous Media* 133 (3) (2020) 497–508.
- [311] EPA, California greenhouse gas 2000-2019 emissions trends and indicators report, Tech. rep., California Air Resources Board (07 2021).
- [312] F. Haider, S. Liu, G. Shaker, Air quality monitoring using UWB radar, in: *Proceedings of 2018 18th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*, IEEE, 2018, pp. 1–4.
- [313] A. M. Andrade, Z. Liu, S. Grewal, A. J. Nelson, Z. Nasef, G. Diaz, M. H. Lee, MOF-derived Co/Cu-embedded N-doped carbon for trifunctional ORR/OER/HER catalysis in alkaline media, *Dalton Transactions* 50 (16) (2021) 5473–5482.
- [314] Y. Yang, X. Yan, X. Hu, R. Feng, M. Zhou, W. Cui, Development of zeolitic imidazolate framework-67 functionalized Co-Al LDH for CO₂ adsorption, *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 552 (2018) 16–23.

- [315] J. Ethiraj, S. Palla, H. Reinsch, Insights into high pressure gas adsorption properties of ZIF-67: Experimental and theoretical studies, *Microporous and Mesoporous Materials* 294 (2020) 109867.
- [316] J. Devkota, D. W. Greve, T. Hong, K.-J. Kim, P. R. Ohodnicki, An 860 MHz wireless surface acoustic wave sensor with a metal-organic framework sensing layer for CO₂ and CH₄, *IEEE Sensors Journal* 20 (17) (2020) 9740–9747.
- [317] J. Devkota, P. R. Ohodnicki, J. A. Gustafson, C. E. Wilmer, D. W. Greve, Designing a SAW sensor array with MOF sensing layers for carbon dioxide and methane, in: *Proceedings of 2019 Joint Conference of the IEEE International Frequency Control Symposium and European Frequency and Time Forum (EFTF/IFC)*, IEEE, 2019, pp. 1–4.
- [318] R. Cao, H. Ding, Z. Peng, K.-J. Kim, P. Ohodnicki, A. Yan, K. P. Chen, Fiber optical sensor for methane detection based on metal-organic framework/silicone polymer coating, in: *Proceedings of CLEO: Applications and Technology*, Optical Society of America, 2018, pp. JW2A–168.
- [319] H. Zhou, D. Li, X. Hui, D. Hu, X. Chen, X. He, X. Mu, Metamaterial gas sensing platform based on surface-enhanced infrared absorption, in: *Proceedings of 2020 IEEE 33rd International Conference on Micro Electro Mechanical Systems (MEMS)*, IEEE, 2020, pp. 717–720.
- [320] S. Hanf, R. Keiner, D. Yan, J. Popp, T. Frosch, Fiber-enhanced raman multi-gas spectroscopy: a versatile tool for environmental gas sensing and breath analysis, *Analytical Chemistry* 86 (11) (2014) 5278–5285.
- [321] Y. Zhao, B. Chen, Q. Liu, X. Li, D. Zheng, F. Feng, Micro gas chromatographic columns with metal-organic frameworks as stationary phase, in: *Proceedings of 2021 IEEE 34th International Conference on Micro Electro Mechanical Systems (MEMS)*, IEEE, 2021, pp. 1034–1035.
- [322] A. E. Omer, G. Shaker, R. Hughson, S. Safavi-Naeini, Microwave-based non-destructive sensing approach for blood type identification, in: *Proceedings of 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, IEEE, 2020, pp. 504–508.
- [323] A. Ebrahimi, J. Scott, K. Ghorbani, Highly sensitive microwave-based biosensor for electrolytic level measurement in water, in: *Proceedings of 2019 IEEE Asia-Pacific Microwave Conference (APMC)*, IEEE, 2019, pp. 759–761.

- [324] A. Rydosz, E. Maciak, K. Wincza, S. Gruszczynski, Microwave-based sensors for exhaled acetone and ethanol detection, in: Proceedings of 2018 International Conference on Electromagnetics in Advanced Applications (ICEAA), IEEE, 2018, pp. 32–34.
- [325] H. Babaei, J.-H. Lee, M. N. Dods, C. E. Wilmer, J. R. Long, Enhanced thermal conductivity in a diamine-appended metal–organic framework as a result of cooperative CO₂ adsorption, *ACS Applied Materials & Interfaces* 12 (40) (2020) 44617–44621.
- [326] S. Tang, E. W. Rothe, G. P. Reck, Negative ion formation from energetic collisions of cesium with CO₂, CS₂, and COS, *The Journal of Chemical Physics* 61 (7) (1974) 2592–2595.
- [327] Z. Liu, A. M. Andrade, S. Grewal, A. J. Nelson, K. Thongrivorong, H.-S. Kang, H. Li, Z. Nasef, G. Diaz, M. H. Lee, Trace amount of ceria incorporation by atomic layer deposition in Co/CoOx-embedded N-doped carbon for efficient bifunctional oxygen electrocatalysis: Demonstration and quasi-operando observations, *International Journal of Hydrogen Energy* 46 (77) (2021) 38258–38269.
- [328] A. Savitzky, M. J. Golay, Smoothing and differentiation of data by simplified least squares procedures., *Analytical Chemistry* 36 (8) (1964) 1627–1639.
- [329] R. Kusumastuti, M. Pancoko, S. Butar-Butar, G. E. Putra, H. Tjahjono, et al., Study on the mechanism of CO₂ adsorption process on zeolite 5A as a molecular sieve in RDE system: An infrared investigation, in: *Journal of Physics: Conference Series*, Vol. 1198, IOP Publishing, 2019, p. 032009.
- [330] E. Muller, H. Decamps, Modeling soil moisture–reflectance, *Remote Sensing of Environment* 76 (2) (2001) 173–180.
- [331] T. Zhao, B. Stark, Y. Chen, A. L. Ray, D. Doll, A detailed field study of direct correlations between ground truth crop water stress and normalized difference vegetation index (NDVI) from small unmanned aerial system (sUAS), in: Proceedings of 2015 International Conference on Unmanned Aircraft Systems (ICUAS), IEEE, 2015, pp. 520–525.
- [332] M. Friedl, F. Davis, Sources of variation in radiometric surface temperature over a tallgrass prairie, *Remote Sensing of Environment* 48 (1) (1994) 1–17.
- [333] W.-H. Chiu, W.-Y. Kong, Y.-H. Chueh, J.-W. Wen, C.-M. Tsai, C. Hong, P.-Y. Chen, C.-H. Ko, Using an ultra-compact optical system to improve lateral flow immunoassay results quantitatively, *Heliyon* 8 (12) (2022) e12116.

- [334] W.-H. Chiu, W.-Y. Kong, Y.-J. Su, J.-W. Wen, C.-M. Tsai, C. Hong, P.-Y. Chen, C.-H. Ko, A faster, novel technique to detect COVID-19 neutralizing antibodies, *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* 28 (2022) e935812–1.