UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

InSAR Timeseries: Observations of Groundwater Recharge and Theoretical Framework of Uncertainty Quantification

Permalink

https://escholarship.org/uc/item/6gz404kj

ISBN

9798173723376

Author

Hurewitz, Zel

Publication Date

2026-11-03

Peer reviewed|Thesis/dissertation

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

UNIVERSITY OF CALIFORNIA SAN DIEGO

InSAR Timeseries: Observations of Groundwater Recharge and Theoretical Framework of Uncertainty Quantification

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

in

Earth Sciences

by

Zel Hurewitz

Committee in charge:

Professor Adrian A. Borsa, Chair
Professor Mattias Morzfeld
Professor David Sandwell
Professor Armin Schwartzman

2026

Copyright

Zel Hurewitz, 2026

All rights reserved.

The Dissertation of Zel Hurewitz is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

DEDICATION

To Leeloo, who turns 5 today.

EPIGRAPH

I'm afraid we need to use... MATH.

- Professor Hubert Farnsworth

TABLE OF CONTENTS

Dissertation Approval Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Figures ix

List of Tables xi

Acknowledgements xii

Vita xiv

Abstract of the Dissertation xv

Introduction 1

Chapter 1 Groundwater Extraction and Recharge at a Utah Semiconductor Plant Observed with Multi-Track InSAR 9
1.1 Introduction 9
1.1.1 Methodological Context 11
1.1.2 Study Area 13
1.2 Methods 15
1.2.1 Data Acquisition 15
1.2.2 InSAR Timeseries Processing 15
1.2.3 Inversion Using Mogi Source Distribution 17
1.2.4 Multi-Track Uncertainty Analysis 22
1.3 Results 22
1.4 Discussion 28
1.4.1 Sustainability Analysis 28
1.4.2 Permanent Compaction 30
1.4.3 Aquifer Properties 32
1.4.4 Applying the Mogi Model in an Aquifer Setting 33
1.5 Conclusions 34
1.6 Acknowledgements 35

Chapter 2 A Framework for GNSS-Informed InSAR Timeseries Uncertainty Quantification with a Case Study in the Mojave Desert 36
2.1 Introduction 36
2.2 Theory 40
2.2.1 Formalism and Notation 42

2.2.2 SBAS Inversion with Uncertainty Quantification 43
2.2.3 Atmospheric Error and the Likelihood Function 45
2.2.4 Constructing the Displacement Prior 47
2.2.5 Bayesian Displacement Timeseries Estimation 48
2.3 Methods 49
2.3.1 InSAR Pre-processing 49
2.3.2 GNSS Timeseries Preparation 52
2.3.3 Interferometric Noise Error Model 54
2.3.4 Atmospheric Delay Error Model 55
2.3.5 Prior for Displacement Parameters 57
2.3.6 Prior for Non-Parametric Component of Displacement 59
2.3.7 Timeseries Estimation Algorithm 60
2.4 Results 61
2.4.1 Uncertainty Quantification 61
2.4.2 Qualitative Comparison with GNSS 63
2.4.3 Quantitative Comparison with GNSS 64
2.5 Discussion 70
2.5.1 Extension for Multiple Subsets 70
2.5.2 Assumption of Normality and Linearity 72
2.5.3 Assumption of Spatial Uniformity 73
2.5.4 Relating the Prior to Regularization 75
2.5.5 Seasonal Displacements of the Mojave 76
2.5.6 Tropospheric Dynamics of the Mojave 78
2.6 Conclusion 82
2.7 Acknowledgements 82

Chapter 3 Theory of Uncertainty Quantification for InSAR Timeseries 84
3.1 Spatio-Temporal Covariance 84
3.1.1 Mathematical Formalism 84
3.1.2 Kronecker Product 86
3.1.3 Linear Transformations 87
3.1.4 Addition of Multiple Terms 90
3.1.5 Application to InSAR Processing 91
3.2 The SBAS Inversion 92
3.2.1 Uncertainty Quantification 94
3.2.2 The Implicit Zero 97
3.2.3 Re-referencing and the Day-Zero Atmosphere 98
3.2.4 Putting It All Together 100
3.3 Timeseries Smoothing and Interpolation 101
3.3.1 Basic Smoothing and Interpolation 102
3.3.2 Parameter-Augmented Kernel Smoothing 103
3.3.3 Regularization and Smoothing 106
3.4 Revisiting Common-Scene Stacking 106
3.4.1 Theory 107

3.4.2 CSS-Equivalent Smoothing 111
3.5 Spatial Referencing 114
3.5.1 Translation of Spatial Referencing Operations 116
3.5.2 Uncertainty Quantification for the Structure Function 121
3.5.3 Example 122
3.6 Covariance Construction 122
3.7 Summary 125
3.8 Appendix 126
3.8.1 Extracting Useful Quantities from Kronecker Covariances 126
3.8.2 Proof of Regularization as Matrix Multiplication 128

Bibliography 129

LIST OF FIGURES

Figure 1.1. a) Context of the Wasatch Front metropolitan area, with the faults (red lines) of the Wasatch Fault Zone highlighted. 11

Figure 1.2. Mogi Green's functions (MGF) derived from the depth profile (top left) inferred from the well log 19

Figure 1.3. Regularization trade off curve 21

Figure 1.4. Top) Total source volumes from areal integration of the MGF inversion, for each track individually 23

Figure 1.5. Surface deformation in Lehi, Utah 24

Figure 1.6. Mogi Green's Function (MGF) inversion results 26

Figure 1.7. a,b,c) Probability density function (PDF) of the three parameters of the model, displayed as normalized marginal distributions of each pair of parameters (central fields) and each parameter alone (bar charts) 29

Figure 1.8. Parameter exploration for an alternate model of total subsurface volume change with no permanent compaction 31

Figure 2.1. a) Sample unwrapped interferogram from June 28th, 2020 to December 25th, 2020 (thick line in b), demonstrating strong and spatially variable radar delay in the troposphere 37

Figure 2.2. Map of the Mojave Desert in Southern California 41

Figure 2.3. a) The P594 GNSS timeseries (grey dots) and the colocated InSAR SBAS timeseries (blue) 55

Figure 2.4. Method for estimation of the amplitude of the covariance kernel 58

Figure 2.5. In a), we show the posterior covariance matrix 60

Figure 2.6. A selection of colocated InSAR and LOS-projected GNSS timeseries, for four stations which demonstrate the range of agreement between the two data sources 63

Figure 2.7. a) Spatial plot of LOS secular velocities inferred from the GNSS (colored circles) and Bayesian InSAR 65

Figure 2.8. Spatial uniformity of the tropospheric delay noise depends on the spatial referencing scheme 74

Figure 2.9. Examples of the spatial distribution of tropospheric delay at six dates, inferred by subtracting the Bayesian displacement timeseries from the SBAS timeseries 78

Figure 2.10. a-b) The mean of the tropospheric delays in winter (December-February) and in summer (June-August) 81

Figure 3.1. Diagram of the InSAR processing chain and uncertainty quantification workflow 92

Figure 3.2. Schematic of the SBAS timeseries (black) and its implicit zero (blue), which may be made explicit. Deformation is estimated as a linear trend, either by including (green line) or excluding (red line) the zero. 97

Figure 3.3. Method of parameter-augmented kernel smoothing demonstrated on a synthetic example 104

Figure 3.4. Uncertainty quantification for parameter-augmented kernel smoothing . . . 106

Figure 3.5. Effect of successive iterations on the degree of smoothing of the common-scene stacking method, demonstrated on a synthetic SBAS timeseries. The red line is under-smoothed, while the yellow line is over-smoothed. 107

Figure 3.6. CSS is approximately equivalent to a forward mapping of the SBAS time-series through a smoothing matrix or through kernel convolution. The inverse mapping uses smooth timeseries calculated with CSS to invert for the CSS-equivalent smoothing matrix and kernel. 111

Figure 3.7. CSS-equivalent smoothing matrices for a range of window sizes and iterations. Only the bottom right corner of each matrix is shown for visual clarity. 113

Figure 3.8. CSS-equivalent smoothing kernels, as a function of window size and number of iterations. 114

Figure 3.9. Left) Classic CSS-equivalent smoothing kernels with a window size of 2. Right) Offline CSS analytically-derived kernels from the theory in the previous section. 115

Figure 3.10. a) Example synthetic atmospheric delay field with a power-law slope of 1.5. 115

Figure 3.11. a) Standard deviation, i.e., square root of the diagonal of the covariance matrix. b) Pearson correlation coefficient of a single pixel with the rest of the scene. c) Random sample drawn from the covariance matrix. 122

Figure 3.12. Transect method of power-law estimation 124

LIST OF TABLES

Table 2.1. Formal uncertainties of the parameters and Pearson correlation coefficients between pairs of parameters 62

Table 2.2. Summary statistics of the scatter plot, including the R^2 values, regression slopes, and regression intercepts. 67

Table 2.3. Summary statistics of the parameters: the 1:1 *p*-value statistic and pass/fail based on a *p*-value of 0.05. 68

ACKNOWLEDGEMENTS

I am lucky to have too many people to thank properly. I want to thank my advisor, Adrian Borsa, for endless support throughout this long process; my many excellent teachers over the years; and all my fellow grad students and post-doc friends who made it worth coming to work every day. Most of all, I want to thank my friends and family for the boundless love and support I will carry with me always. This little paragraph could never serve to convey the depth my feelings, so I will just have to show my love to you in person.

For Chapter 1, InSAR data was processed through the Alaska Satellite Facility's (ASF) Hybrid Pluggable Processing Pipeline 2025 [37] using hyp3_gamma plugin version 9.0.2 running GAMMA release 20240627. Contains modified Copernicus Sentinel-1 data 2014-2025, processed by the European Space Agency. Well log [2] and extraction volume timeseries [1] are publicly available through the Utah Division of Water Rights. This work was funded by NASA ROSES grant 80NSSC24K0670. The authors wish to thank Jim Reese at the Utah Division of Water Rights.

For Chapter 2, InSAR data was processed by the Advanced Rapid Imaging and Analysis (ARIA) Center at NASA Jet Propulsion Laboratory and California Institute of Technology. InSAR data was made available on ASF's Vertex website. Contains modified Copernicus Sentinel-1 data 2014-2025, processed by the European Space Agency. This work was funded by the National Aeronautics and Space Administration under NASA Grants NNX16AR07G (NISAR Science Team Extension) and 80NSSC24K0670 (Earth Surface and Interior). The authors would like to thank the European Space Agency, GPS networks with data processed by UNR's Nevada Geodetic Laboratory, and the U.S. Geological Survey for their open data policies regarding the data used in this work.

Chapter 1, in full, is currently being prepared for publication and is coauthored with Adrian A. Borsa and Tonie Van Dam. I was the primary investigator and author of this research material.

Chapter 2, in full, is currently being prepared for publication and is coauthored with Adrian A. Borsa. I was the primary investigator and author of this research material.

No AI tools were used in the making of this dissertation.

VITA

2026 Ph.D. in Earth Sciences, University of California, San Diego

2019 Bachelor of Science in Physics, University of Texas at Austin

ABSTRACT OF THE DISSERTATION

InSAR Timeseries: Observations of Groundwater Recharge and Theoretical Framework of Uncertainty Quantification

by

Zel Hurewitz

Doctor of Philosophy in Earth Sciences

University of California San Diego, 2026

Professor Adrian A. Borsa, Chair

Timeseries of ground surface displacement derived from Interferometric Synthetic Aperture Radar (InSAR) provide critical information about aquifer characteristics and groundwater sustainability. We observe subsidence and uplift in Lehi, Utah corresponding closely to industrial groundwater extraction at a semiconductor manufacturing plant (Chapter 1). We compare the surface deformation with public timeseries of extraction volumes to provide a sustainability analysis demonstrating that the aquifer has been pumped at a generally sustainable level but also that some permanent compaction has occurred. To investigate the effect of noises on InSAR timeseries, we compare InSAR with Global Navigation Satellite Systems (GNSS) timeseries in

the Mojave desert, using a formal approach which allowing us to place statistical error bounds on the InSAR-based temporal resolution of the post-seismic motion of the 2019 Ridgecrest earthquake (Chapter 2). We find spatial referencing schemes introducing non-uniform spatial patterns of uncertainty and seasonality in the bias and amplitude of turbulent and stratified tropospheric noise. Finally, we introduce a theoretical framework of uncertainty quantification for InSAR timeseries based on the Kronecker product of spatial and temporal covariance matrices (Chapter 3). We translate common InSAR processing techniques, such as the SBAS inversion, timeseries smoothing, common-scene stacking and spatial referencing into this framework. This uncertainty quantification is both powerful and practical for common use for InSAR timeseries processing.

Introduction

Groundwater and Surface Motion

Aquifer systems in water-stressed regions globally are stressed by over-extraction of groundwater. Groundwater is a critical resource for municipal, agricultural, and industrial use, and often serves as a buffer during times of drought. When pumped sustainably, natural recharge replenishes the aquifer and the water level or hydraulic head remains stable. Unsustainable extraction, by contrast, removes more water than is refilled, causing the water level to continuously fall and potentially damaging the health of the aquifer matrix by causing unrecoverable reductions in total aquifer capacity. Groundwater in sedimentary aquifers exists in pores between the sediment particles that comprise the aquifer matrix, which can range in size from fine clay particles to larger sands or gravels. When pore pressure falls beyond its previous lowest level, particles within lenses of clay rearrange and permanently compact, reducing the total amount of water that can be held by the aquifer.

The deformation of the ground surface above an aquifer contains critical information about the sustainability of extraction highly useful to groundwater managers and policy-makers. Permanent compaction within the aquifer causes subsidence, where the overlying ground surface sinks at rates up to several tens of centimeters per year, depending on the severity of extraction; subsidence is an indicator of poor aquifer health but also poses a direct hazard to regional infrastructure. In addition to permanent subsidence, there sometimes exists observable elastic deformation, where changes in the pressure exerted by the water on the aquifer matrix cause recoverable motion. Elastic uplift may be used to infer spatial patterns and timing of natural

recharge, information which may be leveraged to protect key refill pathways. The guiding research question of Chapter 1 of this dissertation will therefore be: How does one infer the sustainable level of extraction for a given aquifer using surface deformation?

Surface displacement cannot replace in-situ measurements of hydraulic head or hydrogeological properties; however, technical advancements in the remote measurement of surface displacement have elevated the utility of subsidence and uplift as proxies for groundwater dynamics. A highly effective tool for measuring surface displacement is satellite-based *Interferometric Synthetic Aperture Radar* or InSAR, which measures ground surface displacement. InSAR is an active remote sensing technique where an antenna mounted on a satellite in low earth orbit emits a pulse of radar waves towards the ground surface, then records the reflected signal. The reflected signal contains both amplitude and phase information; by differencing the phase between two separate passes of the satellite, the ground surface displacement in the direction of the satellite (line-of-sight of LOS) can be recovered. Satellite InSAR provides global coverage, high spatial resolution ($<$ 100 m), and centimeter-level precision, allowing for cheaper and denser measurements of deformation than can be obtained on the ground, including in difficult-to-access locations.

InSAR Processing

InSAR is one aspect of a more general technique called synthetic aperture radar (SAR), which attains high spatial resolution from the radar signal. The radar energy of a single burst spreads over a footprint well over a kilometer wide. Improvement of the resolution from kilometers to tens of meters is attained by firing many bursts as the satellite flies over. Deconvolution of the return signals allows the SAR satellite to operate as though it has a synthetic antenna the size of the orbit. A number of design features are implemented to extract spatial information from the deconvolution operation.

1. The radar beam is directed not straight down at the ground, but is slanted to either the left or right side of the satellite. As a result, the line-of-sight incidence angle changes from the

near field to the far field, with the exact sweep being dependent on the satellite in question.

2. Each pulse is emitted as a chirp, a sweep across multiple frequencies. The Doppler shift of the waveform is used to infer the location of the return signal along the azimuth, or direction of motion of the satellite.

3. The return time of the signal is then used to discern the location of the return signal along the range direction, or the direction perpendicular to the motion of the satellite.

The result of the SAR technique is a complex-valued image of the ground surface, known as a Single-Look Complex (SLC), where values for the return amplitude and phase are available at each pixel (or when not geo-coded, in Range-Doppler coordinates). The return or backscatter amplitude gives information about the roughness and electrical properties of the surface; backscatter amplitudes are used to study soil moisture, surface vegetation, landslides, and ice movement in glaciers and ice sheets. SAR backscatter is highly useful for providing first responders rapid information about natural hazards, such as flood maps. Additionally, the radar signal penetrates through clouds, permitting all-weather acquisition of SAR imagery, a highly useful feature for natural hazard response in cloudy conditions.

The SLC phase at a given pixel is the complex sum of the return phase from all the individual radar scatterers in the footprint of the pixel. These scatterers, which may be anything from pebbles to trees to buildings, are far smaller than the resolution of the SLC pixel and so the return phase is incoherent between pixels, causing the SLC phase to appear as pure noise. However, many of these scatterers are stable over time, meaning that if they move, they move in unison. The summed return phase between different passes of the satellite at different times will therefore remain the same if the ground surface has not moved, or will change coherently over time if motion has occurred. The phase difference is therefore often spatially *coherent*, under sufficiently stable, dry, and/or un-vegetated conditions, and gives information about the unified motion of the scatterers. Scatterers moving separately from each other causes *decoherence* or *decorrelation*, where the summed phase no longer reflects motion. Decoherence is often related

to vegetation growth or changes in the surface electrical properties such as from changes in soil or canopy water content.

The spatial field of differential phase, also called interferometric phase, between two passes of the satellite is known as an *interferogram*. The differential phase is wrapped, or modulo 2π, causing discontinuities at the $-\pi/+\pi$ boundary. Unwrapping algorithms take the wrapped interferogram and sequentially move pixel phases by 2π such that the unwrapped interferogram has a spatially continuous field of differential phase, in a process conceptually similar to estimating topography from a set of contours. Importantly, no information exists about the absolute number of phase cycles between the satellite and the ground surface, meaning that the differential phase has an unknown or arbitrary offset, essentially being free-floating. The unwrapping algorithm has to start somewhere, however, so it will lock in one reference point (typically the most coherent pixel) as having zero phase cycles. Every other pixel in the unwrapped interferogram will therefore contain information about the number of radians the phase moved towards or away from the satellite relative to the reference pixel. When many interferograms are present, it is therefore necessary to re-zero each one to a single common reference pixel, as the unwrapping algorithm is not guaranteed to choose the same pixel for each separate interferogram, thereby causing arbitrary offsets between images. This process is known as *spatial referencing* and has important implications for the spatial distribution of noise, as will be addressed in this work.

Noise in InSAR

Ideally for the purposes of ground motion studies, the differential phase of an interferogram should be a direct estimate of the ground surface displacement. There are other contributions to interferometric phase, however, which may be categorized as the following:

1. Known, static, removable signals, such as the contribution of topography, can be removed algorithmically.

2. Physical delay of the radar wave in the atmosphere introduces spatially continuous noise fields which are temporally variable. This noise source is often the limiting factor in InSAR studies and will largely be the focus of this work.

3. Interferometric noise is interferogram-specific noise not due to coherent radar wave delay, including decorrelation (which is spatially discontinuous, often resembling white noise) and unwrapping errors (which are internally spatially continuous but discontinuous with the rest of the image). When large amplitude and pervasive across a scene, interferometric noise renders interferograms virtually unusable; therefore, usable interferograms tend to have low levels of interferometric noise. Interferometric noise is the cause of non-zero closure phase in cycles of interferograms.

Atmospheric delay noise is often the main source of uncertainty in InSAR data. Radar waves are delayed in the ionosphere, the layer above 30 km altitude where electrons exist unbound to atoms, and in the troposphere, the layer below 15 km which hosts all weather phenomena, from large-scale circulation to individual clouds. Methods exist to directly calculate and remove ionospheric delay noise by leveraging the frequency-dependence of radio wave delay on variations in Total Electron Content (TEC). *Tropospheric noise* is caused by variations in pressure, temperature, and humidity in the lower atmosphere, has no such frequency dependence, and comprises two main components: 1) The stratified component of tropospheric noise is strongly correlated with ground elevation; layered changes in atmospheric pressure and water vapor content will affect low elevations more than high elevations, because the radar waves from the satellite travel through more atmosphere to reflect off the ground. 2) The second component is the turbulent component of tropospheric noise, which is the main focus of this work.

The turbulent component of tropospheric delay noise, hereafter referred to simply as tropospheric noise, is characterized by weather and turbulent flow in the lower atmosphere causing spatial variability of the radar wave speed. Crucially, tropospheric noise introduces noise at both long- and short-spatial scales, with the amplitude of noise generally growing with

wavelength. This exhibited self-similarity derives from the fractal nature of turbulence and is well-described by InSAR literature with a spatial power-law spectrum of the form $a\lambda^b$. The state of the atmosphere at a given pixel decorrelates temporally on a timescale of hours to days, leading the value of tropospheric delay to have effectively no correlation between successive passes of the satellite.

The self-similarity observed in the spatial patterns of tropospheric noise greatly affects the uncertainty of measurements of ground surface motion. The effect is most severe at longer wavelengths, where the amplitude of the noise gets particularly large. This issue particularly impacts large-scale studies of tectonic motion or isostatic rebound and leads to large uncertainties of crustal velocities, particularly when the geophysical system in question extends beyond the interferogram bounds. Sufficiently dense deployments of Global Navigation Satellite System (GNSS) stations, where available, can mostly compensate for this shortcoming. Small-scale features are overall better targets for InSAR, as they leverage the fine resolution of InSAR data relative to the sparsity of GNSS networks, they are completely contained within the interferogram bounds, and the noise level is lower at smaller wavelengths. However, tropospheric noise remains relevant at small scales where many geophysical systems have meaningful structure. The guiding research question of Chapter 2 of this dissertation will therefore be: How does noise affect InSAR estimates of ground surface displacement?

In part due to these considerations, InSAR work has historically focused on: 1) high-amplitude, unmistakable signals, such as co-seismic displacement fields or volcano deformation, or 2) signals which are predominantly static in time, such as tectonic motions, for which repeated, long-term observations can reduce uncertainties. The launch of the Sentinel-1 constellation by the European Space Agency (ESA) has transformed the field by acquiring data both frequently (as often as every 6 days) and over a long duration (surpassing 12 years as of writing). These dense observations have expanded the range of possibility for studying transient motions, including inter-annual changes in aquifer subsidence and slow-slip events on faults. As a result, the limits of InSAR data are being tested. The combination of short duration, small amplitude, and unknown

geometry in geophysical phenomena pushes against the bounds of tropospheric noise, providing an ever greater need for systematic uncertainty quantification in the InSAR field.

Uncertainty Quantification

Uncertainty quantification, in general, refers to formal treatment of the statistical properties of data to determine the error bounds on derived quantities. The basic question that uncertainty quantification helps answer is: Is this observation statistically significant? In our framework, data values are treated as random variables, mathematical objects described not by a static number but by a statistical distribution represented as a probability density function (PDF). Most often, the random variables considered here will be multivariate, meaning that their domain is multi-dimensional. We will mainly restrict our studies to the second-order moment of the probability distribution, known as the *covariance matrix*, which describes the variance of a random variable and correlations between random variables. A particular challenge in working with covariances in InSAR timeseries is the presence of spatially correlated noise, which becomes blended in time by temporal smoothing. Non-trivial covariance structure then exists between different pixels at different timesteps. The guiding research question of Chapter 3 of this dissertation will therefore be: How does one work with spatio-temporal uncertainties for InSAR timeseries?

In the opinion of the author, uncertainty quantification for InSAR exists in something of a Wild West state at the present moment. The field as a whole does not appear to have a consensus on a given approach. Various approaches to InSAR uncertainty quantification have been put forward over the years by isolated groups, but these appear to be either too highly specialized or too complex to be widely adopted. We believe that the opportunity remains for a theoretical framework addressing the treatment of InSAR uncertainties that is self-consistent, practical, and comprehensive. We intend this work to be a step towards this goal.

Overview

In Chapter 1, we use InSAR to observe hydrological surface subsidence and uplift at a semiconductor manufacturing plant in Lehi, Utah. This chapter is novel in its pairing of InSAR deformation data with extraction volume timeseries, as opposed to hydraulic head timeseries, to conduct a targeted analysis of groundwater sustainability and aquifer hydrogeological properties. This analysis offers a method for monitoring of aquifers when minimal other data are present.

In Chapter 2, we conduct an comparative analysis of InSAR data and GNSS with the goal of using GNSS to inform the uncertainty quantification of InSAR data. This chapter considers questions essential to the understanding of InSAR processing, such as the temporal and spatial patterns of tropospheric noise, the role of spatial referencing, and an analysis of interferometric noise.

In Chapter 3, we present a theoretical framework of uncertainty quantification for InSAR timeseries processing. This framework introduces a novel method for the practical storage and manipulation of the uncertainties of large interferogram and deformation timeseries by using the Kronecker product to separate the spatial and temporal components of the covariance matrix. We then translate into the language of this framework commonly used techniques in InSAR timeseries processing, such as the SBAS inversion, common-scene stacking, and spatial referencing.

Chapter 1

Groundwater Extraction and Recharge at a Utah Semiconductor Plant Observed with Multi-Track InSAR

1.1 Introduction

Industry is a significant consumer of global groundwater [74], totaling 5% or 4.4 trillion gallons per year of groundwater extraction in the U.S. [19]. Water use management and sustainability in the industrial sector are important to characterize in the context of the variable rainfall and long-term drought conditions experienced in the western United States.

The U.S. state of Utah, which is home to substantial industry, experiences considerable water stress with the potential for serious environmental hazard. Water supply in Utah, as in most of the southwestern US, is currently threatened by changing climate conditions, notably by increased likelihood of extreme droughts [77, 41]. Over-use of water, largely a result of extensive irrigated agriculture and industrial and municipal use, has resulted in the shrinkage of the Great Salt Lake [11], with consequences for the $60 million brine shrimp industry [49], bird migration patterns [8], human respiratory health [30, 6], and reduced local precipitation [31].

Remote sensing can help fill critical gaps in the global monitoring of aquifer health [74]. Remote observations of groundwater-related surface deformation provide insight into the natural and human characteristics of tapped aquifers; subsidence indicates potentially unsustainable

extraction, while uplift illuminates patterns of natural and managed recharge. The primary tool for remotely sensed measurements of surface motion is satellite-based Interferometric Synthetic Aperture Radar (InSAR). InSAR measures the ground surface displacement in the direction of the satellite (known as line-of-sight or LOS) from the phase difference between two radar images acquired during different passes of the satellite. Since 2014, the European Space Agency's (ESA) Sentinel-1 mission has provided global coverage, centimeter-scale precision, and 90 m spatial resolution on a 6-day repeat cycle, allowing for displacement timeseries capable of addressing geophysical phenomena on a decadal timescale.

Using InSAR to study surface deformation in the Salt Lake City metropolitan area, we observed a prominent, highly localized subsidence signal in the nearby city of Lehi, Utah. The existence of this subsidence signal has been noted by previous authors; we find mention and speculation on an anthropogenic origin in Hu et al., 2018 [38]. However, no detailed analysis of the source was presented. In this work, we demonstrate that the origin of this subsidence is groundwater extraction by the Texas Instruments semiconductor fabrication plant (LFAB) in Lehi, which produces 300 mm silicon wafers for the microchip industry. Silicon wafers and semiconductors need to be repeatedly cleaned with ultrapure water during their manufacture; this ultrapure water is obtained by water-intensive reverse-osmosis [85]. The LFAB, which had been previously owned and operated by other companies over the last two decades, operates two groundwater wells. Well logs and timeseries of water extraction volumes, provided to meet industrial groundwater well reporting requirements, are publicly available through the Utah Division of Water Rights (DWR).

The surface deformation accompanying the groundwater extraction at this plant demonstrates both sustainable and unsustainable characteristics. We find periods of substantial surface uplift, indicating natural aquifer recharge. However, the surface never recovers fully to its initial state, indicating that portions of the aquifer underwent permanent compaction, and the total groundwater capacity of the aquifer has been permanently reduced.

In addition to the implications for groundwater management, this research also highlights

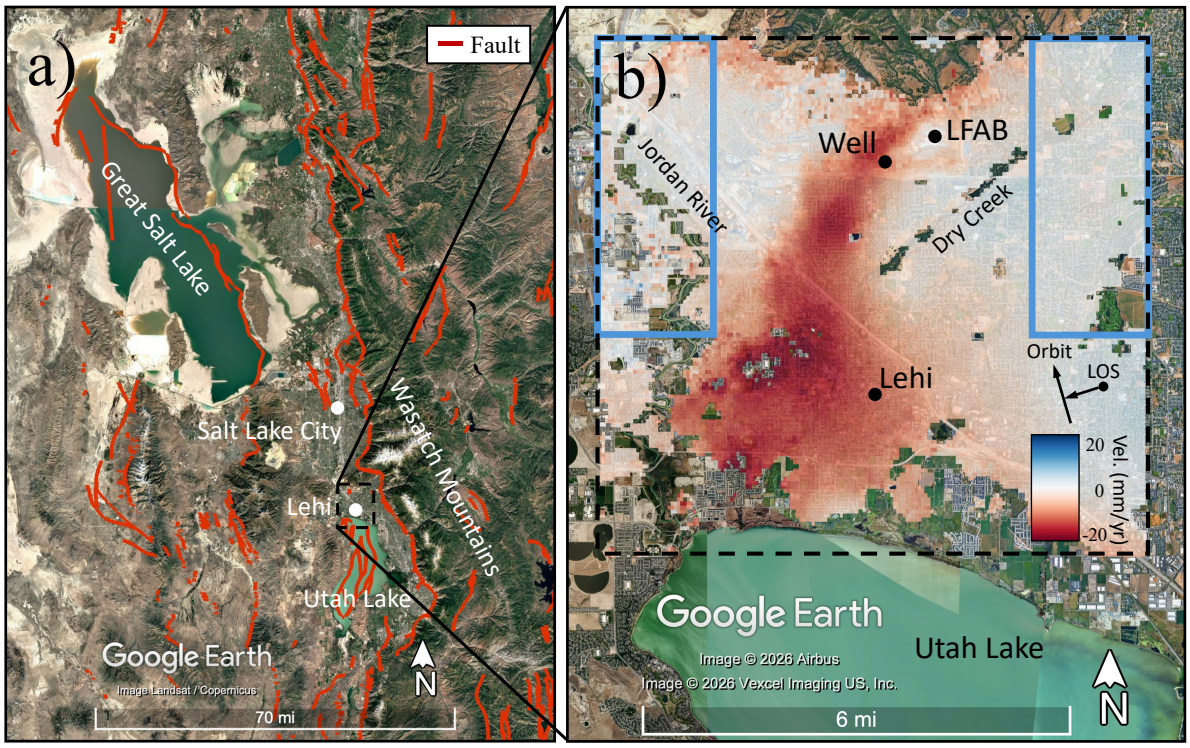


Figure 1.1: a) Context of the Wasatch Front metropolitan area, with the faults (red lines) of the Wasatch Fault Zone highlighted. Fault data are from USGS. b) Context of Lehi and the LFAB plant with line-of-sight (LOS) velocity from the ground towards the satellite in ascending track 122 for 2016. The black dashed line represents the study area, while the regions lined in solid blue denote the InSAR reference area.

the utility of this work as a case study for InSAR; horizontal motions, in the form of extension and contraction, have similar amplitudes to vertical motions, making this an ideal scenario to showcase the integration of multiple look geometries. In the same spirit, we also present a use of multiple tracks to perform practical uncertainty quantification for InSAR data interpretations.

1.1.1 Methodological Context

InSAR is a proven tool for monitoring groundwater-related surface deformation in the Western US. For decades, InSAR has observed subsidence in the California's Central Valley [22]; seasonal recharge in Las Vegas [3, 36]; aquifer compaction in the Mojave [26] and in Coachella Valley [68]. More recently, work investigating surface uplift has illuminated recharge dynamics in the Central Valley [55].

Groundwater extraction causes both recoverable and permanent subsidence of the ground surface. Water in aquifers exists in pore spaces between the solid material comprising the aquifer matrix and exerts an outwards pore pressure. When water is extracted from an aquifer, the deficit causes a reduction in pore pressure, which results in an increase in the effective stress ($\sigma_e = \sigma - pI$, where σ is the total stress tensor, p is the pore pressure, and I is the identity tensor) detected by the aquifer matrix [71]. This applied stress causes displacement of the aquifer structure towards the region of extraction and may be observed at the surface as subsidence. Similarly, when natural recharge or fluid injection occurs, the surface inflates and uplifts overlying material relative to its previous state. It is worth mentioning that in addition to the pressure effect, the weight of the water causes displacement of the ground due to elastic loading; however, we performed tests which demonstrated that the expected amplitude of this effect was orders of magnitude smaller than the pressure effect and so the loading signal will be neglected going forward.

In addition to recoverable poroelastic deformation, aquifers also experience permanent compaction during water level drawdown. Alluvial aquifers composed of sands and gravels may also contain lens-like inclusions of silts and clays [39]. These aquitard interbeds have lower hydraulic conductivity and react on a longer timescale to changes in hydraulic head [42]. When head drops, the aquitard lenses begin to leak water out into the surrounding matrix, causing their internal grain structure to settle and compact permanently [40]. This unrecoverable compaction represents a permanent loss to the total storage capacity of the aquifer. The draining process from the lenses can continue for years after head has recovered in the surrounding matrix and introduces a temporal lag between extraction and observed surface deformation [61, 66, 86]. Water extracted from these lenses can be considered to have been mined for one-time use, in contrast to the renewable water pumped from the elastic regions (sands and gravels) on the aquifer.

Fluid extraction in buried sediment will always result in some measure of 3-dimensional surface deformation (i.e., both vertical and horizontal motions); however, when the lateral

extent of extraction is much greater than the depth of extraction, the amplitude of the horizontal displacement is small compared to that of the vertical displacement and may in many situations be safely neglected. This approximation of 1D vertical deformation forms the basis for Terzaghi's original theory of soil consolidation [71], which remains in use today. Aquifers pumped for irrigation agriculture, for example the Central Valley of California, are typically good applications for Terzaghi's theory because many pumping wells are distributed over a wide area. The aspect ratio between extraction depth (often no more than a few hundred meters) and lateral extent (multiple kilometers) is therefore quite small and drawdown of the groundwater level is spatially homogeneous.

However, horizontal displacement cannot always be neglected. When the depth of fluid extraction is comparable to its breadth, horizontal displacement is on the same order of magnitude as vertical displacement. Biot's seminal work laid the theoretical foundation for the 3D treatment of poroelastic soil consolidation [12] by introducing the fundamental equations of poroelasticity. The Mogi model of displacement, originally developed to describe volcanic deformation associated with magma chambers, treats a spherical cavity with changing volume at depth in a uniform, isotropic, elastic half-space, in the limit where the depth is much greater than the radius of the sphere [52]. The Mogi model has been previously applied to groundwater extraction at the Geysers [54] and Cerro Prieto [14] geothermal fields. We follow the approach of Trugman et al., 2014 [72] by treating the observed deformation as resulting from a distribution of Mogi sources. We address criticisms of the applicability of the elastic Mogi model to an aquifer setting, where the effects of poroelasticity and permanent compaction are relevant, in the Discussion section.

1.1.2 Study Area

The city of Lehi (Fig. 1.1b) is located at the northern end of the Utah Valley, a sediment-filled basin that has been submerged in the past (18,000 - 16,000 BP) by the Greater Lake Bonneville. The valley is filled with several thousand meters of unconsolidated to semi-consolidated

sediments. The older Pre-Lake Bonneville Quaternary and Tertiary deposits form the bulk of these sediments by volume and include interbedded fluvial gravels, alluvial-fan deposits, lacustrine clays from earlier Pleistocene lake cycles, and Tertiary semi-consolidated sediments. They are highly heterogeneous laterally because they record repeated cycles of alluvial fan growth and lake flooding. The late Pleistocene Lake Bonneville group (Alpine and Bonneville formations) are deeper-water clays, silts, and marl deposited during the lake's rising and maximum lake level phases [57]. The shore embankments and deltas of the Lake Bonneville group are the most permeable units in the section. Post-Bonneville Holocene deposits contain modern alluvium, alluvial fans, marsh deposits, and aeolian silt.

The principal groundwater system of northern Utah Valley is found in unconsolidated basin-fill deposits. Near the mountain front, the groundwater can be located in an unconfined aquifer. As one moves from the mountain front to the center of the valley, the groundwater system transitions into a stacked sequence of confined aquifers separated by low-permeability layers [15]. Clark and Appel, 1985 [16] subdivided the principal reservoir in northern Utah Valley into three major confined aquifers separated by lacustrine fine-grained confining layers, plus an unconfined aquifer in pre-Bonneville coarse deposits at the mountain front. The confining beds separating the aquifers are predominantly fine-grained, although their lithology varies and includes clay, silt, and locally cemented or coarser horizons [16].

Lehi lies between a westward extension of the Wasatch Mountains to the north and Utah Lake to the south. The lake level and flow in the Jordan River, the outlet of Utah Lake, is controlled by a dam in southwest Lehi. Winter snow in the Wasatch Mountains melts in spring and feeds down to Utah Lake through a combination of mountain front recharge and surface runoff.

1.2 Methods

1.2.1 Data Acquisition

We use interferometric data from the European Space Agency's (ESA) Sentinel-1 mission, whose radar instrument has a wavelength of 5.6 cm (C-band). We generate Sentinel-1 interferogram products using the On-Demand service available from the Alaska Satellite Facility's (ASF) Hybrid Pluggable Processing Pipeline (HyP3). We generate 4,794 interferogram products from late 2014 to the present over the three Sentinel-1 tracks, which intersect the Lehi region: track 20 (ascending), track 100 (descending), and track 122 (ascending). Each interferogram product records the relative displacement in the line-of-sight (LOS) direction towards the satellite between a primary and a secondary date (a date pair) for each pixel within a given Sentinel-1 frame. The ascending tracks have data available from late 2014 to 2021, while the descending track has data available until the present.

On its website, the Utah Division of Water Rights (DWR) makes publicly accessible well logs for the LFAB, which detail the layered geological structure observed during the well drilling operation, in addition to the depths of the perforated sections of the well casing. Also made public on the website are monthly reportings of water extraction volumes over the lifetime of the well.

1.2.2 InSAR Timeseries Processing

Due to the limitations of HyP3 processing at the time of interferogram generation, multiple interferogram products are generated which have the same track, primary date, and secondary date, but are spatially limited to a given frame and must be stitched together. We begin our processing by stitching the interferogram products together into interferograms that fully cover the study area. We mask out each pixel for which 30% or more of interferograms have a coherence less than 0.7. This masking procedure removes the most persistently incoherent pixels, most notably in the foothills of the Wasatch mountains where snowpack over much of the year

causes decoherence in the InSAR data. We then remove entire interferograms from processing using an automated procedure designed 1) to remove certain individual interferograms, as well as all interferograms associated with certain dates, which have poor coherence, and 2) to preserve the integrity of the interferogram network structure. This procedure is designed to minimize decoherence noise. Details of the algorithm may be found in the Supplemental Materials.

We spatially reference each interferogram by subtracting the mean value of the interferometric displacement over the set of pixels which have coherence greater than 0.5 for every interferogram. This many-pixel referencing scheme is chosen to minimize the contribution of the interferometric noise. We then perform a Small-Baseline Subset (SBAS) inversion [10] to transform the stack of interferograms into an SBAS timeseries, which represents a timeseries of displacement perturbed by tropospheric noise.

After the SBAS inversion, we perform a second spatial referencing whose purpose is to minimize the impact of tropospheric noise on the study area. Tropospheric noise is spatially self-correlated over all length scales, but with correlations that decay with distance according to the structure function, whose functional form is a power-law [78, 21, 47]. To optimally reduce the impact of tropospheric noise, we fit and subtracted a planar trend to the SBAS displacement value at each date over the reference area, which is a region closely surrounding, but not intersecting with, the observed localized deformation within the study area (Fig. 1.1b, solid blue lines). This planar referencing scheme reduces the uncertainty due to tropospheric noise over the region of interest.

We smooth the LOS SBAS timeseries in time (on a pixel-by-pixel basis) with a moving weighted average with a Gaussian kernel. We choose the kernel timescale to be 0.1 yr to maximize smoothing without degrading the displacement signal. We then interpolate onto a uniform 6-day posting to reduce the effect of tropospheric noise.

For interpretation only (i.e., not for the inversion described in the next section), we project the line-of-sight displacements onto an east/up coordinate system. The geometry of the available look vectors ensures that the InSAR data are almost insensitive to motion in the

north-south direction. The three look vectors are effectively co-planar, meaning that we do not have independent estimates of motion exactly in all three dimensions; moreover, the plane in which they lie is nearly perpendicular to the north-south axis, with a normal vector only 6° off. North-south motion of the same order of magnitude of east-west or up-down motions would therefore appear about ten times smaller in the LOS. For these reasons, we make the approximation that no north-south motion is observable in the timeseries and project into the east/up plane. The line-of-sight displacements d_i are then the dot product of the displacements in the x (east) and z (up) directions (d_x and d_z respectively) with the look vector components (e.g., $v_{x,100}$), shown collected in the matrix equation below. With as many equations as unknowns, the displacement in xz-coordinates may be inverted using the normal equations.

$$\begin{pmatrix} d_{100} \\ d_{122} \end{pmatrix} = \begin{pmatrix} v_{x,100} & v_{z,100} \\ v_{x,122} & v_{z,122} \end{pmatrix} \begin{pmatrix} d_x \\ d_z \end{pmatrix} \tag{1.1}$$

1.2.3 Inversion Using Mogi Source Distribution

We invert for the spatial pattern of subsurface volume deficit using a distribution of sources derived from the Mogi model. This inversion allows us to 1) localize the deformation to its subterranean source region and 2) partition the deformation between the deep extraction of interest and unrelated broad-scale subsidence in the near-surface. The Mogi model as formulated by [65] is described by the following formula, where z is the depth of the source, r is the radial distance (in cylindrical coordinates) from the source, u_r is displacement in the radial direction, u_z is the displacement in the vertical direction, and ΔV_{sphere} is the change in volume of the spherical source, and ν is Poisson's ratio.

$$\begin{pmatrix} u_r \\ u_z \end{pmatrix} = \Delta V_{sphere} \frac{1-\nu}{\pi (r^2+z^2)^{\frac{3}{2}}} \begin{pmatrix} r \\ z \end{pmatrix} \tag{1.2}$$

In the above formulation, there is a direct trade-off between the value of Poisson's ratio and the source volume change ΔV_{sphere}. Poisson's ratio is commonly taken to be 0.25 (e.g., [51]), but there is actually substantial variation among aquifer materials [63, 86]. In-situ values for Poisson's ratio in aquifers inferred from the relative velocities of P- and S-waves range from 0.3 or below [43, 82] to near the theoretical maximum of 0.5 [58, 87].

Here, we present a formulation which eschews use of Poisson's ratio entirely by putting the source volume in terms of the total change in volume of surface, i.e., the total amount of subsidence or uplift. By using this formulation, we apportion observed surface deformation to its origin at a given spatial location at depth, without directly considering the volume change of the structure of the subsurface. Integrating the vertical displacement results in the following relation: $\Delta V_{surf} = 2(1-\nu)\Delta V_{sphere}$ [65], where ΔV_{surf} is the volume change of the surface. We end with the following set of equations for the Mogi model which is free of any unknown physical constants.

$$\begin{pmatrix} u_r \\ u_z \end{pmatrix} = \Delta V_{surf} \frac{1}{2\pi(r^2+z^2)^{\frac{3}{2}}} \begin{pmatrix} r \\ z \end{pmatrix} \tag{1.3}$$

We account for known vertical layering of the aquifer system in our estimation of the subsurface source distribution. The well log for the LFAB well records that the well casing is perforated in two sections; from 329 to 457 meters, and again from 534 to 665 meters. As a first-order approximation, we assume that this represents the layering of the hydraulically conductive portions of the aquifer, i.e., that any pumped water will be extracted equally from the full vertical range of perforations and at no other depths (Fig. 1.2, top left). We construct a Mogi Green's function (MGF) by vertically integrating the Mogi model across the perforation depths (Fig. 1.2, top right). The resulting MGF is functionally indistinguishable from a point Mogi source at 460 m depth, indicating a minimal sensitivity of the MGF to the exact layering structure.

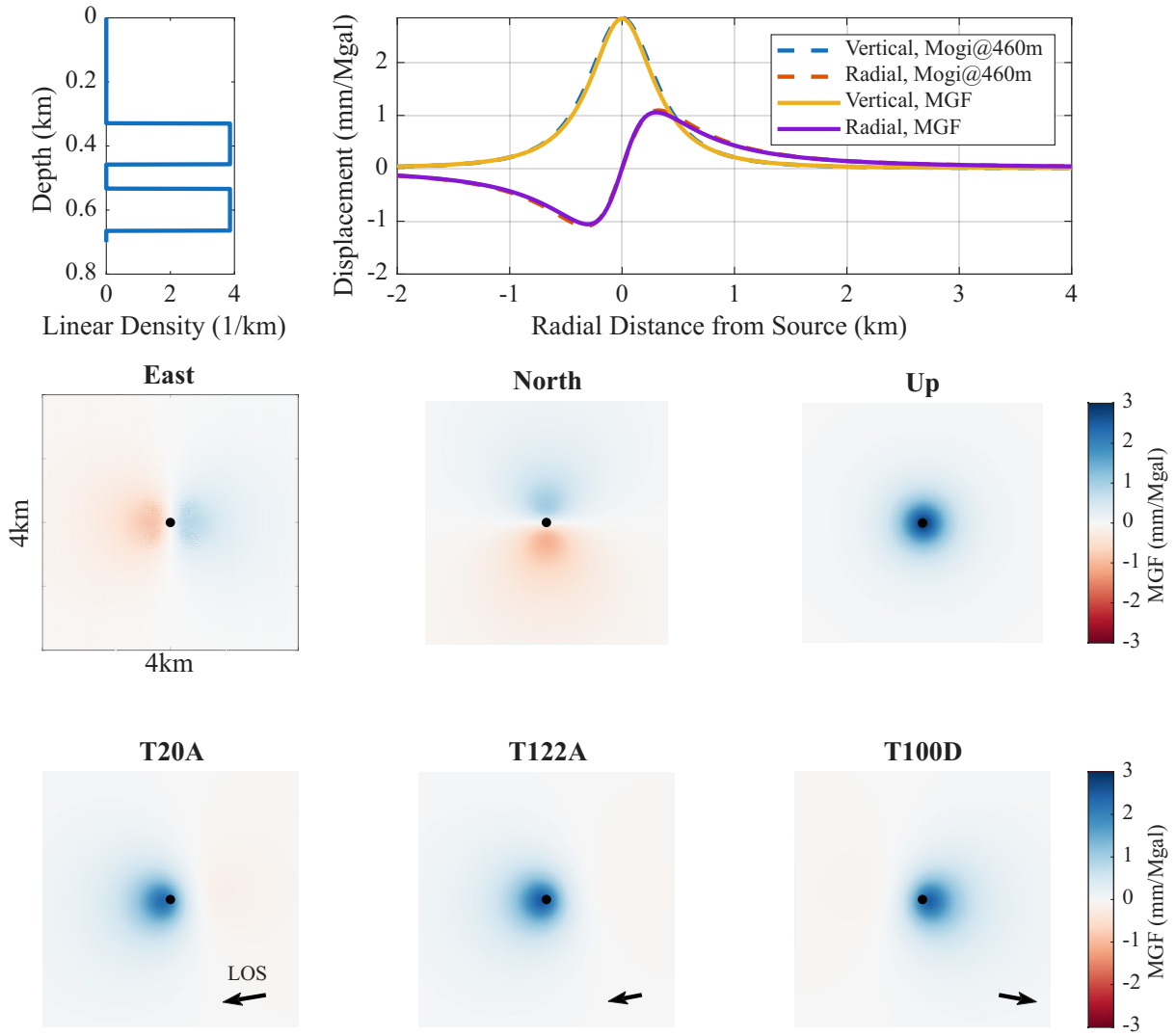


Figure 1.2: Mogi Green's functions (MGF) derived from the depth profile (top left) inferred from the well log. The MGF is nearly identical to a point source Mogi function at 460m depth, shown for comparison in the dashed lines. Middle row: map view of the MGF in the east, north, and up directions. The black dot indicates the position of the source. Bottom row: MGF projected into the LOS direction for each track. As the look vector varies slightly across the scene, the above examples are sampled in the scene's center. However, this variation is not neglected when constructing the Green's function matrices for the inversion.

For the spatial inversion, we set up a horizontal distribution MGF sources with 200 m spacing (circles in Fig. 1.6b). We construct the Green's function matrices between each source and each pixel, projected into the line-of-sight directions, for tracks 100 (descending) and 122 (ascending) ($\mathbf{G}_{100}$ and $\mathbf{G}_{122}$ respectively). For the portion of the timeseries when we have multiple look directions, we combine the data for the two tracks into the joint displacement vector $\mathbf{u}$ and the Green's function matrices into the joint Green's function matrix $\mathbf{G}$ (eq. 1.5). We invert for the optimal distribution of volume changes $\mathbf{V}^*$ (eq. 1.6), including a model-size regularization (i.e., Tikhanov regularization/ridge regression). To determine the regularization parameter λ, we constructed a model-size/RMSD trade-off curve using the spatial field of displacement on December 12$^{\text{th}}$, 2016 as a representative datum (Fig. 1.3). We use the regularization value at which the model size begins to grow ($\lambda = 400$ mm^2/Mgal^2). We perform this regularized inversion for each timestep, with the regularization parameter held constant.

$$\text{Single Track } k : \mathbf{u} = \mathbf{u}_k, \mathbf{G} = \mathbf{G}_k \tag{1.4}$$

$$\text{Joint} : \mathbf{u} = \begin{pmatrix} \mathbf{u}_{100} \\ \mathbf{u}_{122} \end{pmatrix}, \mathbf{G} = \begin{pmatrix} \mathbf{G}_{100} \\ \mathbf{G}_{122} \end{pmatrix} \tag{1.5}$$

$$\mathbf{V}^* = \text{argmin}_{\mathbf{V}}(||\mathbf{u} - \mathbf{G}\mathbf{V}||^2 + \lambda||\mathbf{V}||^2) \tag{1.6}$$

We integrate the source volumes over the integration region (box in Fig. 1.6b) to get a timeseries of total volume change (Fig. 1.6a). As a result of our formulation of Mogi source in terms of the surface volumes, this timeseries represents the total volume of subsidence or uplift. A nearly equivalent timeseries can be obtained by integrating the vertical component of displacement directly from the InSAR (Fig. S3). However, the vertical displacement-derived volume will likely include some portion of the shallow subsidence, which is present around the shore of Utah Lake (see Fig. 1.5a). The MGF source inversion is not sensitive to this shallow signal, and so the volume timeseries derived in this manner better represents the portion of

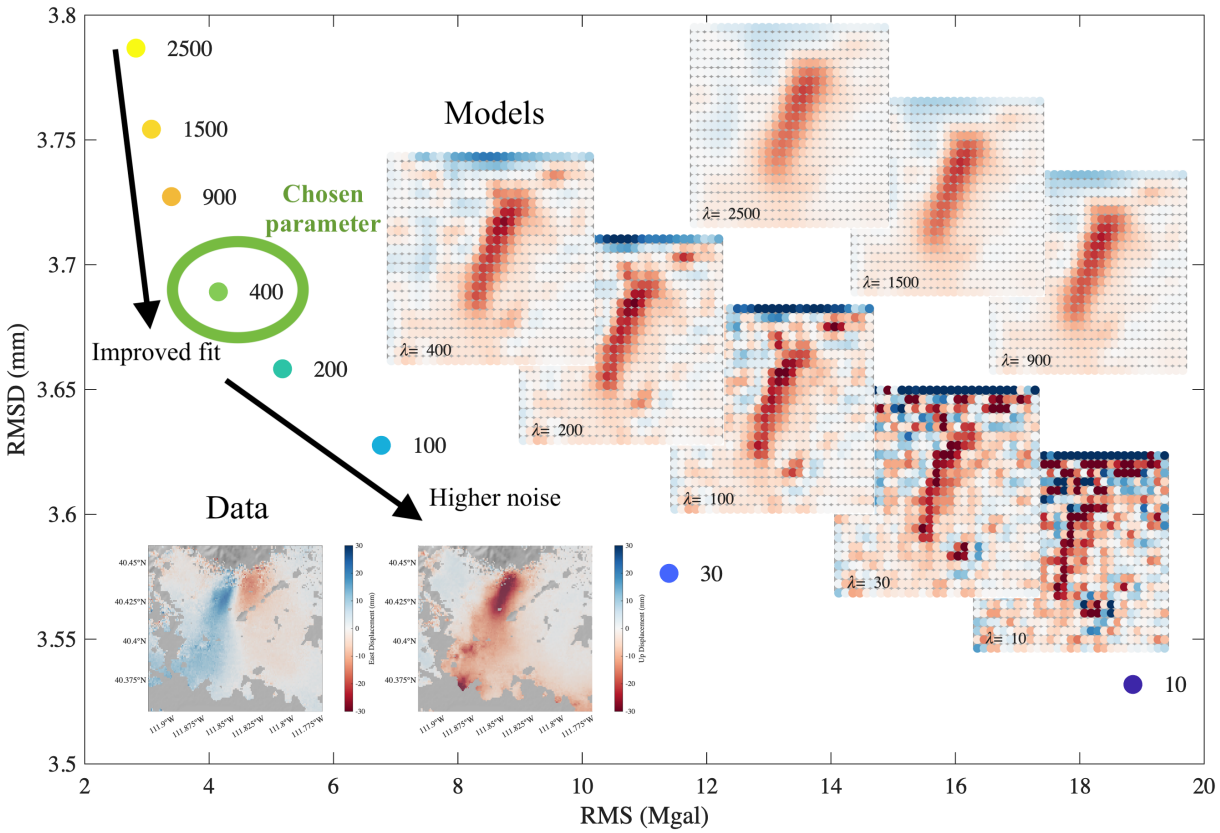


Figure 1.3: Regularization trade off curve. The larger the regularization parameter (color/text), the smaller the total model size (RMS, x-axis) and the larger the data misfit (RMSD, y-axis). The data are shown in the lower right inset. Highlights of the inverted models are shown in the upper right panels for each regularization parameter. The chosen regularization parameter value of 400 was chosen to be the point on the trade-off curve where the improvement in fit is balanced by the increasing noise.

subsidence attributed to the deep extraction.

1.2.4 Multi-Track Uncertainty Analysis

We perform the same MGF inversion and areal integration for the two ascending tracks separately, giving us two independent timeseries of volume change. The two timeseries differ as a result of noise, mostly tropospheric noise, propagating through the various stages of processing. We assess that there are two distinct periods to consider in the uncertainty analysis; the period before 2017 experiences large deviations due to relatively infrequent SAR acquisitions, while the period after shows smaller deviation. For pre- and post- January 1st, 2017, we take the standard deviation of the difference between the volume timeseries of the two tracks and divide by $\sqrt{2}$. We assign these values as the uncertainty for a timeseries derived from a single track, for the pre- and post-2017 periods respectively, giving a timeseries $\sigma(t)$ (Fig. 1.4). For the jointly-inverted timeseries, we divide $\sigma(t)$ by $\sqrt{2}$ to account for the two independent data sources, which inform the inversion equally.

1.3 Results

In the up-projected component of displacement, we observe a prominent localized signal of subsidence (Fig 1.5a). The amplitude of the localized signal peaks in the north, then spreads and dissipates along an elongated oval oriented towards the southwest. In the east-west displacement field, we observe a contraction whose axis coincides with the strike of the subsidence signal (Fig 1.5c). Taken together, the two components of displacement indicate an extraction of fluid from a localized region at depth, with a pattern spatially similar to a single Mogi source/sink [52]. Like the up field, each LOS displacement field (Fig 1.5b) shows localized negative signals which may be interpreted as subsidence; however, the signals in the ascending and descending tracks are offset from each other (Fig. S5), and it is only by combining the ascending and descending tracks that we obtain the interpretation of a source at a significant depth, relative to its lateral extent.

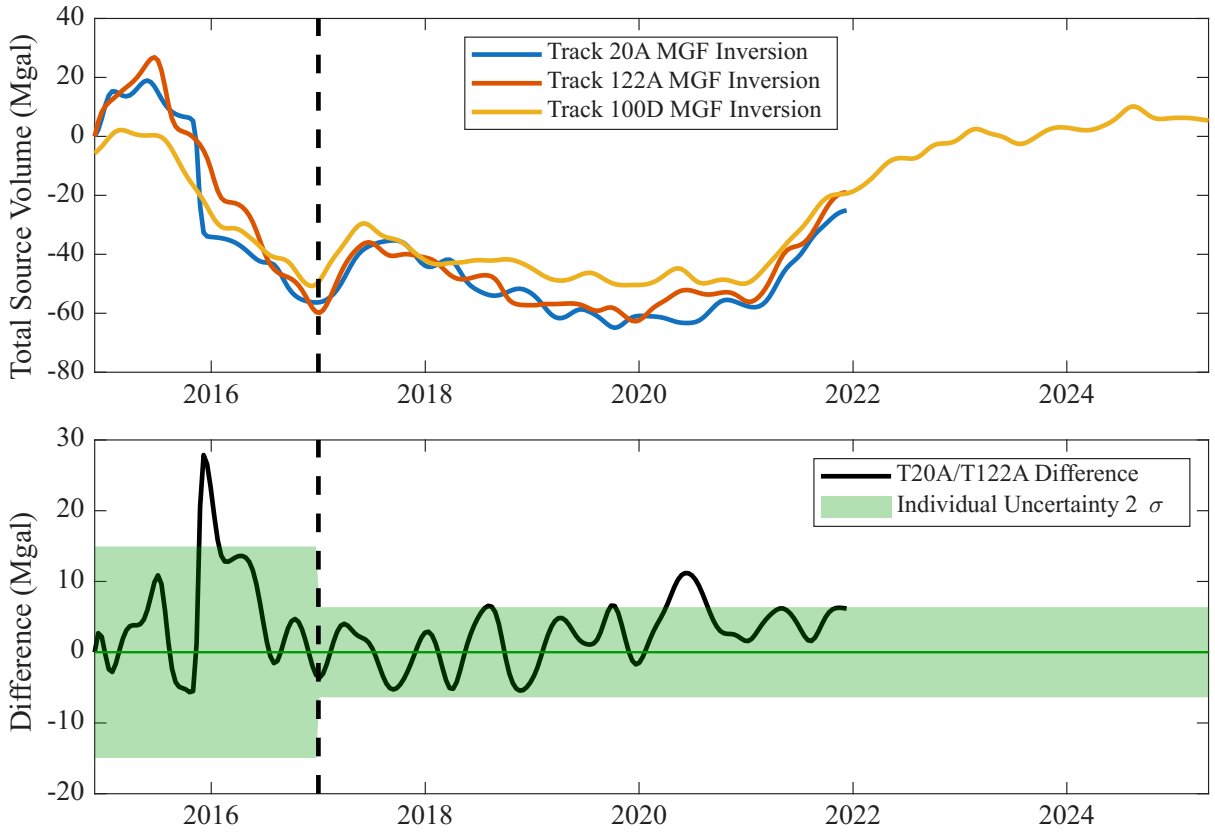


Figure 1.4: Top) Total source volumes from areal integration of the MGF inversion, for each track individually. Bottom) The difference between ascending tracks 20 and 122 (solid black line). The dashed line indicates the chosen divide between the earlier and later periods, differentiated by reduced SAR acquisitions and higher uncertainties in the earlier period. The green shaded region indicates the estimated uncertainty for an individual track at the 2σ level. For the joint inversion, the use of two independent tracks to estimate the same quantity means that the resulting uncertainty is reduced by a factor of $\sqrt{2}$.

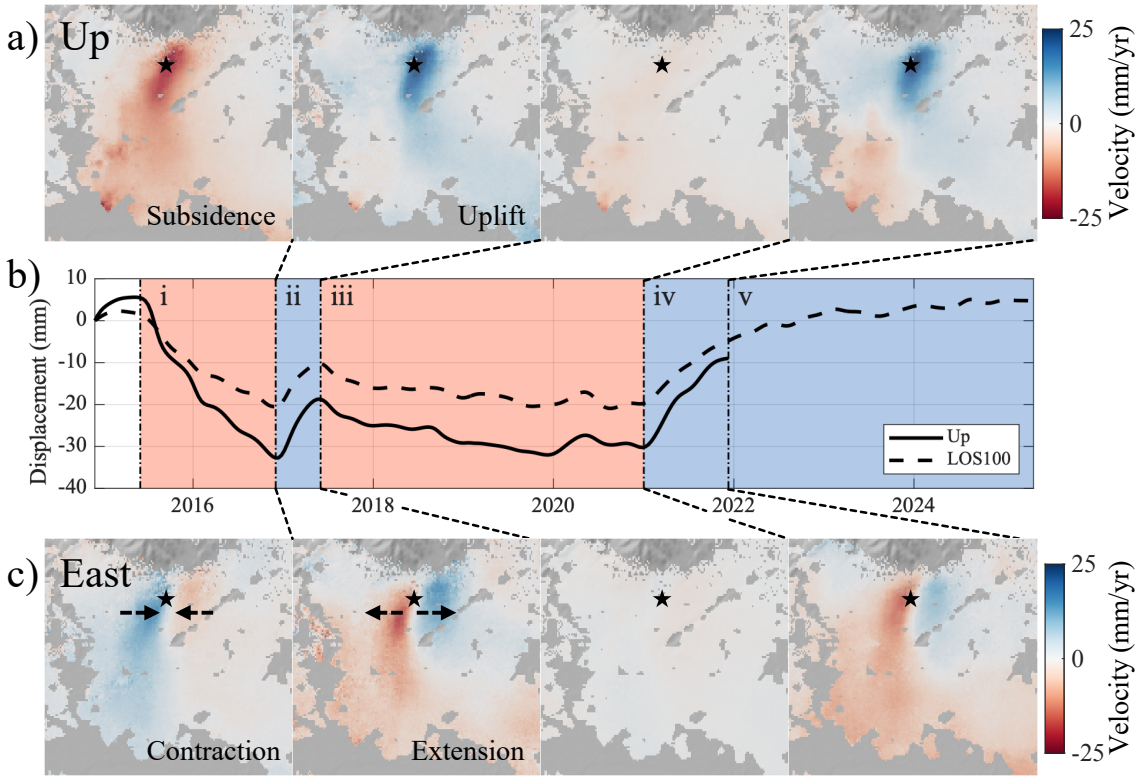


Figure 1.5: Surface deformation in Lehi, Utah. In a), we show map plots of the vertical component of surface velocity averaged over the time periods marked. In b), we show the timeseries of displacement at the location of LFAB well (starred in the map plots), in the up direction (solid) and the descending line-of-sight direction (dashed). Periods of subsidence are shaded red and periods of uplift are shaded blue. Period (v) is inferred to be a period of uplift from the descending track data alone, as there is no accompanying ascending track data during this period to estimate an east/up decomposition. In c) we show map plots of the horizontal velocity in the east direction for the corresponding time periods. Rates and patterns of contraction and extension closely match subsidence and uplift respectively.

We separate the timeseries into five distinct periods of subsidence and uplift (Fig. 1.5b). After an initial uplift, the surface subsides rapidly from June 2015 to December 2016 (i). This subsidence is followed by a half-year period of steep uplift (ii). Starting in June 2017, a steady decline in the ground surface lasts for three and a half years until January 2021 (iii). The surface then undergoes a second period of rapid uplift (iv). The final part of the timeseries, after June 2021 (v), is a gentle uplift to a steady state and is inferred from the descending data alone, as there is no accompanying ascending data during this period with which to perform the east/up decomposition. The spatial pattern of the localized deformation is consistent throughout. In each case, there is horizontal contraction paired with subsidence and extension paired with uplift in the east-west direction.

The MGF inversion allows us to explore the spatial distribution of subsurface water change (Fig. 1.6c). The inversion demonstrates a clear linear feature arching gently from the well towards the southwest. We interpret this feature as a portion of the aquifer with high hydraulic conductivity into which the well happened to penetrate.

There is a clear correspondence between the surface volume timeseries and the monthly well extraction rate, which is published on the Utah Division of Water Rights (DWR) website. The heaviest extraction, in period (i), corresponds to the fastest subsidence. By contrast, when the well stops pumping in periods (ii), (iv), and (v), we observe reinflation of the ground surface.

We model the relationship between the published extraction timeseries and the total MGF source volume timeseries. The model, presented in terms of the total volume of water in the aquifer within the integration region $w(t)$, has three terms. 1) The extraction term removes fluid from the aquifer at the specified time-dependent rate $dw/dt = -x(t)$, which is the published extraction timeseries from DWR. 2) The $dw/dt = -w/\tau$ term is an exponential decay to zero from the initial condition, representing the diffusion of water into the aquifer over recovery timescale τ. 3) However, we modify the exponential decay term to $dw/dt = (f(t) - w)/\tau$ so that instead of decaying to zero, w will decay to a time-dependent function $f(t)$. Our choice for this function, $f(t) = -\dot{V}t$, is a single-parameter model which is linear in time, where the capacity

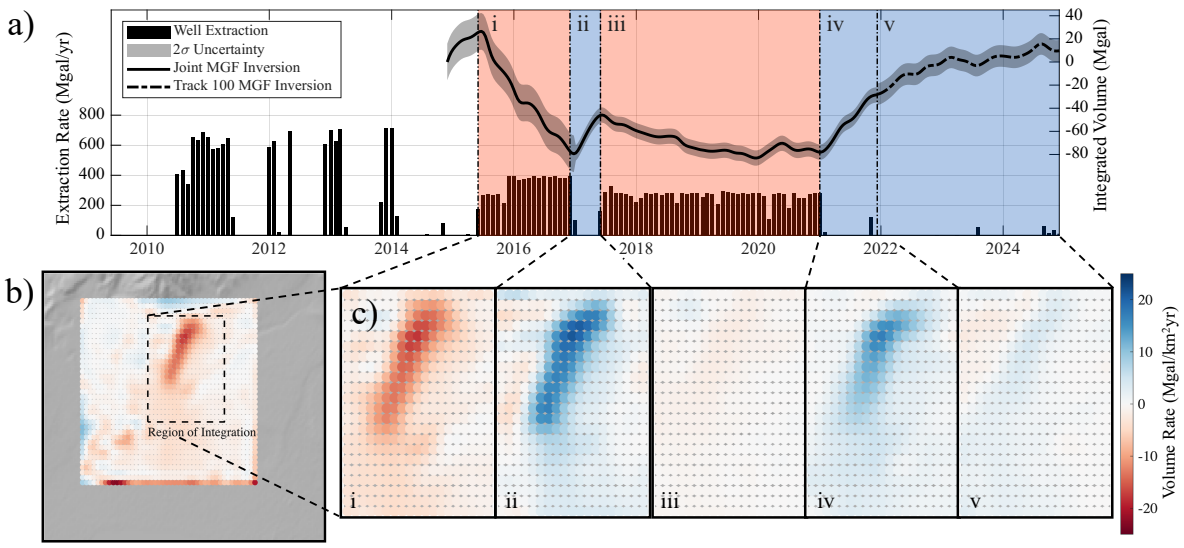


Figure 1.6: Mogi Green's Function (MGF) inversion results. a) The extraction rate at the well (bar chart) is publicly available through Utah DWR and extends earlier than the beginning of the InSAR timeseries. The source volumes are summed over the integration region delineated in (b) to produce the timeseries of total MGF source volumes using the joint inversion (solid) and the Track 100-only extension (dashed). The inferred 2σ error (transparent black region) is included. b) The full spatial extent of the inversion is placed in context. Missing data in the outer regions causes overfitting of noise, leading to unphysical oscillatory overfitting on the boundaries and certain portions of the interior. c) Spatial fields of the average rates of source volume change over the periods (i) through (v).

loss rate $\dot{V}$ is the average rate of change in the amount of fluid volume that the aquifer can store (see discussion of our use of a linear model in Section 1.4.2). All together, the differential equation governing the total subsurface volume of groundwater is below.

$$\frac{dw}{dt} = -x(t) + \frac{-\dot{V}t - w}{\tau} \tag{1.7}$$

To match the observed surface displacement, we require a third parameter, a scaling parameter; the coefficient of compaction c is the ratio between an incremental unit of volume of fluid extracted and the incremental observed subsidence volume change. The displacement is then $y(t) = c * w(t)$.

For any given triplet of parameter values $q = \{c, \dot{V}, \tau\}$, we integrate the differential equation to get a timeseries $y(t) = M(t|q)$, where M refers to the model function, which may be compared against the timeseries of surface volume change from the MGF inversion $v(t)$. As $v(t)$ is derived from a timeseries of displacement, it has an arbitrary offset; therefore our initial condition for $y(t)$ may also have an arbitrary offset. In practice, we integrate from an initial condition of $y(0) = 0$, then correct the entire timeseries by the mean difference between $y(t)$ and $v(t)$.

We estimate the value of the probability density function (PDF) with respect to the parameters directly with $p(q) \propto \exp(-||(v(t) - M(t|q))/\sigma(t)||^2)$; in other words, the probability is Gaussian with respect to the root-mean-square difference between the model and the data timeseries, weighted by the error timeseries. Proportionality is sufficient because, being a probability density function, $p(q)$ must integrate to unity. The parameter PDF is displayed in Figures 1.7 a), b), and c) as a set of normalized marginal distributions, PDFs integrated with respect to one or more parameters then normalized to the maximum value. The 2D marginal distributions, represented as the central fields, may be thought of as 2D histograms and are functions of two parameters. The bar charts are the 1D marginal distributions for each parameter. As a result of the non-linearity of the model with respect to the parameters, the optimal set

of parameters (i.e., the PDF mode) does not lie exactly at the PDF mean. In Figure 1.7d, the timeseries of the optimal model (solid purple line) is plotted against the total source volumes derived from the MGF inversion (black line). The shaded purple region represents the 2σ variation around the PDF mean.

We find a coefficient of compaction of 0.33 ± 0.07, indicating that there is an approximately 3:1 ratio between volume of water extracted and volume of surface subsidence. We find a recovery timescale τ of 1.5 ± 0.4 years. We find that the average rate of capacity loss is 11 ± 4 Mgal/yr, corresponding to a total aquifer capacity loss of 170 ± 60 Mgal since extraction began in 2010.

1.4 Discussion

1.4.1 Sustainability Analysis

Our analysis gives insight into the overall sustainability of the groundwater extraction taking place. In total, over 2600 Mgal of water were pumped at the LFAB well since 2010. Our model suggests that 160 ± 60 Mgal of capacity was lost permanently, implying that roughly 94% of water extracted was renewed by recharge. Poland et al., 1975 [59] discusses pumpage derived from the "water of compaction," or the one-time mining of water from aquitard interbeds, as the ratio between subsidence volume and total pumpage volume. Presenting our result in these terms, the observed fraction of 6% corresponds to regions described by Poland in California's Central Valley where natural recharge rates are adequate to largely prevent permanent head decline and surface subsidence.

The model suggests that, at its minimal point, the aquifer had been drawn down by about 500 ± 90 Mgal relative to the start of the timeseries (Fig. 1.7d). However, by the end of the timeseries, only 300 ± 70 Mgal, or 60%, had been recovered, with an additional 40 ± 30 Mgal of current water debt on track to be refilled in the future (raising the recovery fraction to 67%). If we assume that the minimum volume experienced by the aquifer over the whole timeseries

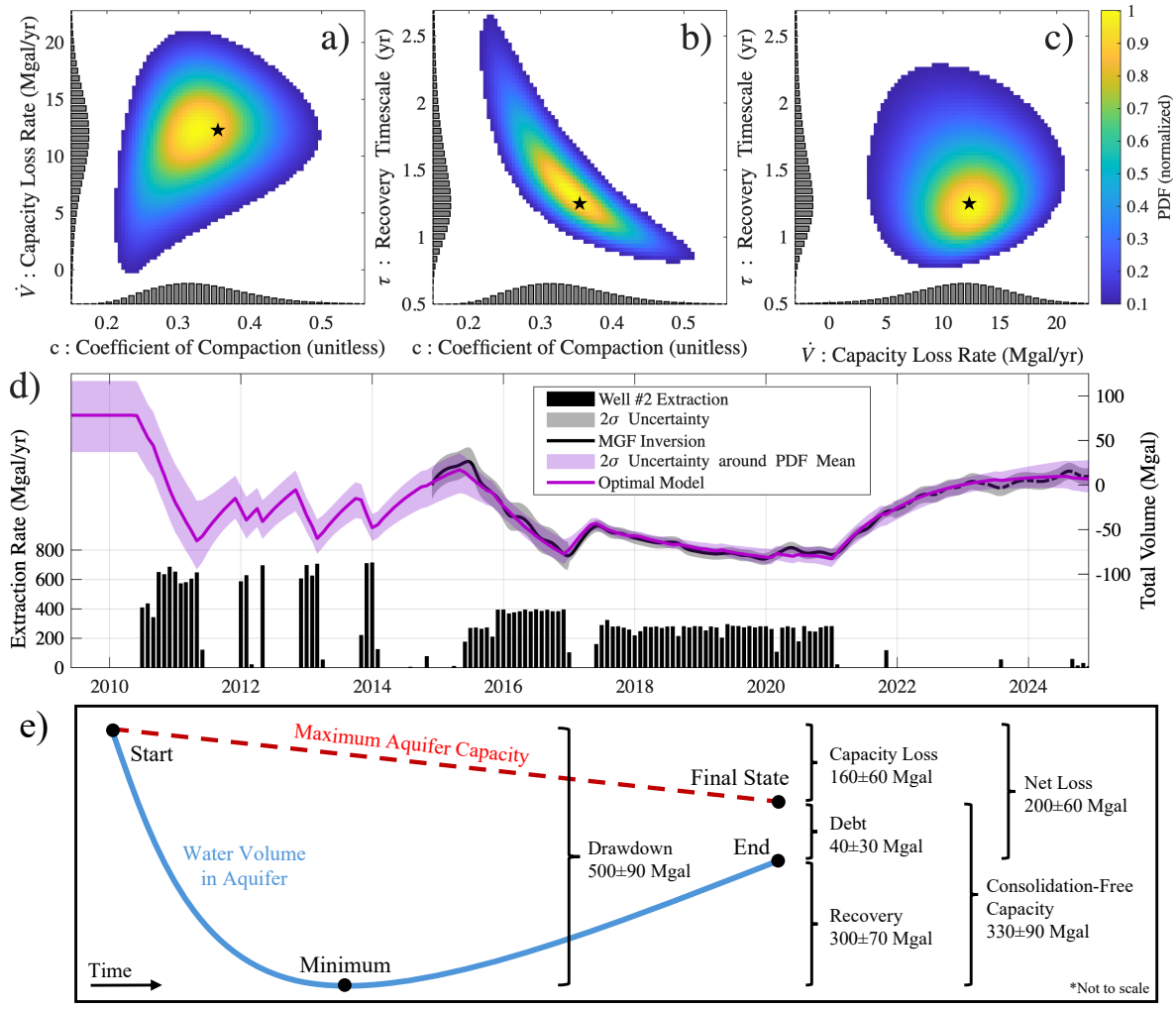


Figure 1.7: a,b,c) Probability density function (PDF) of the three parameters of the model, displayed as normalized marginal distributions of each pair of parameters (central fields) and each parameter alone (bar charts). The star indicates the parameter values of the optimal model (i.e., PDF mode). d) Extraction rate from Utah DWR (black bar chart) is used to produce the model of the timeseries of MGF source volumes. e) Schematic representation of the model, showing the water volume present in the aquifer (blue line) and the reduction in aquifer capacity (red dashed line) over time.

is now the present-day preconsolidation point, the total volume by which the aquifer could be drained without causing additional permanent consolidation, relative to the final state, is the recovery plus the current water debt, or 330 ± 90 Mgal. In steady state at that level of drawdown and assuming no further capacity loss (adapting from equation 1.7, $x = -w/\tau$), an extraction rate of 230 ± 20 Mgal/yr would balance recharge sustainably. This is slightly less than the average extraction over the 2017-2021 period, during which mild surface subsidence was observed. On the other hand, if extraction so far exceeded recharge as to drop the head below the present-day preconsolidation point, the recovery fraction of 67% might again come into play; each additional unit of drawdown below that point would result in unrecoverable compaction of 0.33 units, assuming the characteristics of past consolidation can be applied to a future analysis.

1.4.2 Permanent Compaction

Physically, permanent compaction happens after the hydraulic head falls below the preconsolidation head, i.e., the lowest head previously experienced in that region of the aquifer. At that point, aquitard lenses, with excess pore pressure relative to their environment, begin slowly leaching water out into the surrounding aquifer. Due to their low hydraulic conductivity, there is a large time lag introduced by the slow propagation of the pressure front into the lenses. In the meantime, the reduction of the support from the pore pressure causes the clay grains to rearrange, collapsing permanently under the weight of the material above, which is observed as subsidence. The net effect is that after a period of low groundwater head, the surface can continue to decline for years or decades afterwards, even if head has recovered.

The central piece of evidence in support of the presence of permanent compaction in this region is that the surface in 2025, after years with little to no pumping, remains slightly subsided compared to the beginning of the InSAR timeseries. Permanent compaction is a necessary feature in our model to explain the observed deformation timeseries; the total probability that the capacity loss rate is less than 2 Mgal/yr (obtained by integrating the PDF) is only 2.5%. We tested a modified timeseries model which excluded the capacity loss rate term ($\dot{V}$) (Fig. 1.8).

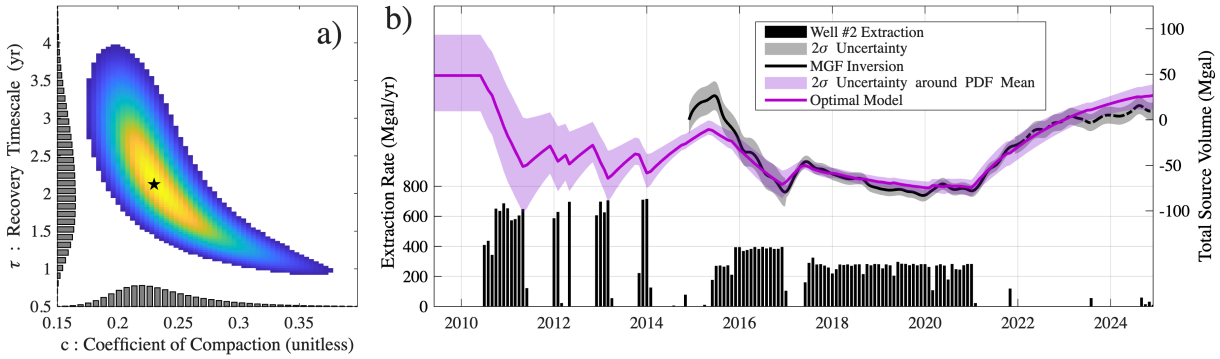


Figure 1.8: Parameter exploration for an alternate model of total subsurface volume change with no permanent compaction. The model (purple) fits the period from 2016 to 2023 reasonably well, although the rates are reduced compared to the data (black). However, the fit is poor in the period before 2016 and the period after 2023. In the case of the pre-2016 period, the observed subsidence lies considerably outside the two-sigma range. This plot is intended to make the case that permanent compaction is a necessary addition to the parameterized model.

This modified model is unable to match the relative surface deformation from the beginning to the end of the timeseries.

Although there is enough supporting information to assert the necessity of including the process of permanent compaction, there remains little possibility of discriminating between various models which may resemble a physical model. Moreover, without access to head data, any information about the previous and present state of consolidation would be highly uncertain. In our model, we eschew these issues by reducing the modeled compaction to a single variable describing the average rate of aquifer capacity loss. The simplicity of our approach comes at the loss of any predictive power; it is not within the scope of the model to answer questions about whether the aquifer will continue to lose capacity in the future, for example. The model as presented does include some implicit assumptions, however, namely, that the timescale of compaction is long enough to produce a roughly linear observed compaction. The large water drawdown in 2010-2011 almost certainly caused the preconsolidation head to drop to a value near to which it remained for the rest of the timeseries, so if the timescale were significantly shorter than 5 years or thereabouts, all resulting compaction would have been complete before the start of the InSAR timeseries and consequently would not be observed.

1.4.3 Aquifer Properties

Our analysis also provides insight into the nature of the subsurface. From the pattern of surface deformation, we infer a relatively bounded linear feature in the aquifer, which is most likely a paleo-bed of high-conductivity sands and gravels. It is also notable that the bed appears to be relatively laterally isolated, i.e., surrounded by low-conductivity sediments in the horizontal direction in addition to being confined above and below. If this were not the case, for example if the entire aquifer layer had high horizontal conductivity, we would expect to see a more radially symmetric pattern of diffusion. Numerical tests with a 2D diffusion model show that a large contrast in conductivity (by a factor of at least 10-100) is necessary to sustain this elongated pattern.

Our model identified a recovery timescale τ of 1.5 ± 0.4 years. Experiments with a 2D diffusion model indicate that this value is largely derived from the horizontal diffusivity of the low-conductivity portion of the aquifer surrounding the high-conductivity material; in other words, the region exterior to the paleo-bed determines how fast water can flow into the paleo-bed from outside. The value of the horizontal diffusivity, which results in the observed timescale, is 1.5 ± 0.5 km^2/yr (1.2 ± 0.4 m^2/s).

In a general sense, the hydrogeology of the aquifer plays a controlling role in the surface expression of groundwater change. The relative hydraulic isolation of the tapped bed amplifies the amplitude and spatial compactness of the surface displacement by causing greater drawdown in a small area, as opposed to less drawdown over a wider area. Moreover, as the tapped bed's connection to the broader aquifer system is low conductivity, it recharges more slowly than would otherwise be the case, causing greater water drawdown. Our detection of the surface motion is indebted to this amplifying effect.

The ratio between the volume of water extracted and the volume of surface subsidence is calculated to be 3:1 (i.e., the coefficient of compaction of 0.33). This ratio cannot be fully explained by the effects of Poisson's ratio with a value between 0.25 and 0.5, which would cause

a ratio of at most 2:1 (see Section 1.2.3), i.e., a coefficient of compaction of 0.5. From the model PDF (Fig. 1.7), the total probability of a coefficient of compaction greater than 0.45 is only 5.2%. The compressibility of water, no more than a few percent at these pressures, cannot also account for the discrepancy. One possibility is that some water may be derived not from the confined portion of the aquifer but instead from the upper unconfined layer; if the upper layer is composed of large-grained sediments such as sand, shallow withdrawal may not cause enough associated deformation to appear in the InSAR deformation pattern, which we attributed entirely to deep withdrawal.

1.4.4 Applying the Mogi Model in an Aquifer Setting

The Mogi model describes the deformation observed at the surface from a deep spherical body (radius $\ll$ depth) changing volume in a uniform, isotropic, elastic half-space. A single Mogi source does not account for 1) a spatially distributed source field (Figure 1.5), 2) temporal variability, or 3) poroelasticity. We address these concerns in this section.

The elastostatic equations describe how a continuous medium displaces in response to forces, in this case the force exerted by pore pressure on the aquifer skeleton ($\nabla \cdot \sigma = \mathbf{F}$). Pore pressure evolves in time according to the groundwater diffusion equation ($\partial_t h = \nabla \cdot (\alpha \nabla h) + s$). In a full poroelastic formulation, the diffusion equation includes a term to account for the volumetric dilatation of the aquifer matrix, thus coupling the pressure evolution to the elastic stress field. Gambolati et al., 2000 find that the coupling introduces a slight overpressure of only a few percent in shallow ($\approx$ 500 m) conditions, which dissipates quickly with time. They concluded that "on a scale of practical interest... the uncoupled and coupled predictions of land subsidence are approximately the same" [27]. As a further practical consideration, the poroelastic equations only describe reversible motions, and therefore themselves are incompletely describe aquifers undergoing permanent consolidation, often called inelastic deformation for that reason.

Using the uncoupled form of the elastic and diffusion equations, we may treat the pressure term as part of the forcing. The equations of both elasticity and poroelasticity are linear, meaning

that any linear combination of solutions is also a solution. The equations can be therefore solved through the method of Green's functions, wherein a fundamental analytic solution for an infinitesimal source is integrated over the source distribution. Analytic solutions are exact and are particularly useful for inversions due to their accuracy and low computational cost [24]. The Mogi solution for surface displacement operates as a Green's function solution to the elastic equations for the case of a small/infinitesimal volume source. In our analysis, we construct an aquifer-specific Mogi Green's function to account for the vertical layering, then invert for the horizontal distribution of these sources instantaneously for each moment, therefore accounting for the source field being both spatially distributed and temporally dynamic. The elastic equations are satisfied by the use of the Mogi solution; and the diffusion equation is used not in the forward sense, but instead the time evolution of the solution is inferred by the continuous inversion. For these reasons, we assert that shallow aquifers constitute an appropriate use of the Mogi model when inverting for the spatially and temporally variable volume field.

1.5 Conclusions

From a decade of InSAR data from Sentinel-1, we generate a surface displacement timeseries and find a prominent localized signal of subsidence and uplift in the city of Lehi, Utah. The use of multiple InSAR tracks allows us to decompose the signal into the east/up coordinate system, revealing that the displacement has horizontal motions of comparable magnitude to the vertical motions. The revealed pattern is akin to that of a Mogi source, which describes subsurface volume change at depths much greater than the horizontal extent of the source. In the time domain, we see alternating periods of subsidence and uplift, with the final period of uplift exponentially plateauing to a value near or above the starting displacement of the timeseries. We find an industrial well operated by a semiconductor manufacturing plant located at the center of the subsidence/uplift signal, whose publicly available groundwater extraction timeseries correlates strongly with the surface displacement timeseries.

To isolate the subsurface signal from near-surface subsidence, we invert for the spatial pattern of subsurface sources with Green's functions derived from integrating Mogi sources vertically along the layers of perforation as described in the public well log. We find that the subsidence originates from a narrow, elongated feature within the aquifer, which we interpret as an isolated bed of high-hydraulic conductivity material.

We construct a three-parameter box model to represent the total water content of the aquifer, which takes into account the known extraction, exponential recharge, permanent compaction of the aquifer, and a scaling factor between volume extracted and volume change observed on the surface. We find that both significant recharge and aquifer capacity loss has occurred. Using this model, we perform a sustainability analysis which infers that 6% of water extracted was mined (i.e., cannot be recovered) and that an extraction rate of 230 ± 20 Mgal/yr could be maintained without additional damage to the aquifer. Although many studies using InSAR to investigate aquifer related subsidence rely on in-situ hydraulic head data, we wish to convey that crucial sustainability analyses can be performed when only volume extraction data is available.

1.6 Acknowledgements

For Chapter 1, InSAR data was processed through the Alaska Satellite Facility's (ASF) Hybrid Pluggable Processing Pipeline 2025 [37] using hyp3_gamma plugin version 9.0.2 running GAMMA release 20240627. Contains modified Copernicus Sentinel-1 data 2014-2025, processed by the European Space Agency. Well log [2] and extraction volume timeseries [1] are publicly available through the Utah Division of Water Rights. This work was funded by NASA ROSES grant 80NSSC24K0670. The authors wish to thank Jim Reese at the Utah Division of Water Rights. Chapter 1, in full, is currently being prepared for publication and is coauthored with Adrian A. Borsa and Tonie Van Dam. I was the primary investigator and author of this research material.

Chapter 2

A Framework for GNSS-Informed InSAR Timeseries Uncertainty Quantification with a Case Study in the Mojave Desert

2.1 Introduction

Measurements of Earth surface motions using Interferometric Synthetic Aperture Radar (InSAR) remain limited by the lack of a consistent framework of uncertainty quantification. The high spatial resolution of InSAR has greatly improved understanding of volcanic and seismic hazard [60, 79]; landslides [45]; and aquifer extraction [3] and recharge [55], among other topics. The increasing duration of modern Synthetic Aperture Radar (SAR) missions, particularly ESA's Sentinel-1 constellation, have resulted in a shift towards multi-year timeseries analysis. However, the presence of noise, particularly from radar delay in the turbulent troposphere, hinders estimation of surface deformation timeseries. Although many methods have been proposed to correct or mitigate tropospheric noise, little attention has been given to the specific problem of quantifying the uncertainty arising from tropospheric noise. In this work, we present a framework for the systematic inclusion of error models and uncertainty propagation for the problem of displacement timeseries estimation from InSAR.

Noise in InSAR arises from many sources but may be separated into two categories: interferometric noise, which is unique to the construction of each interferogram; and atmospheric

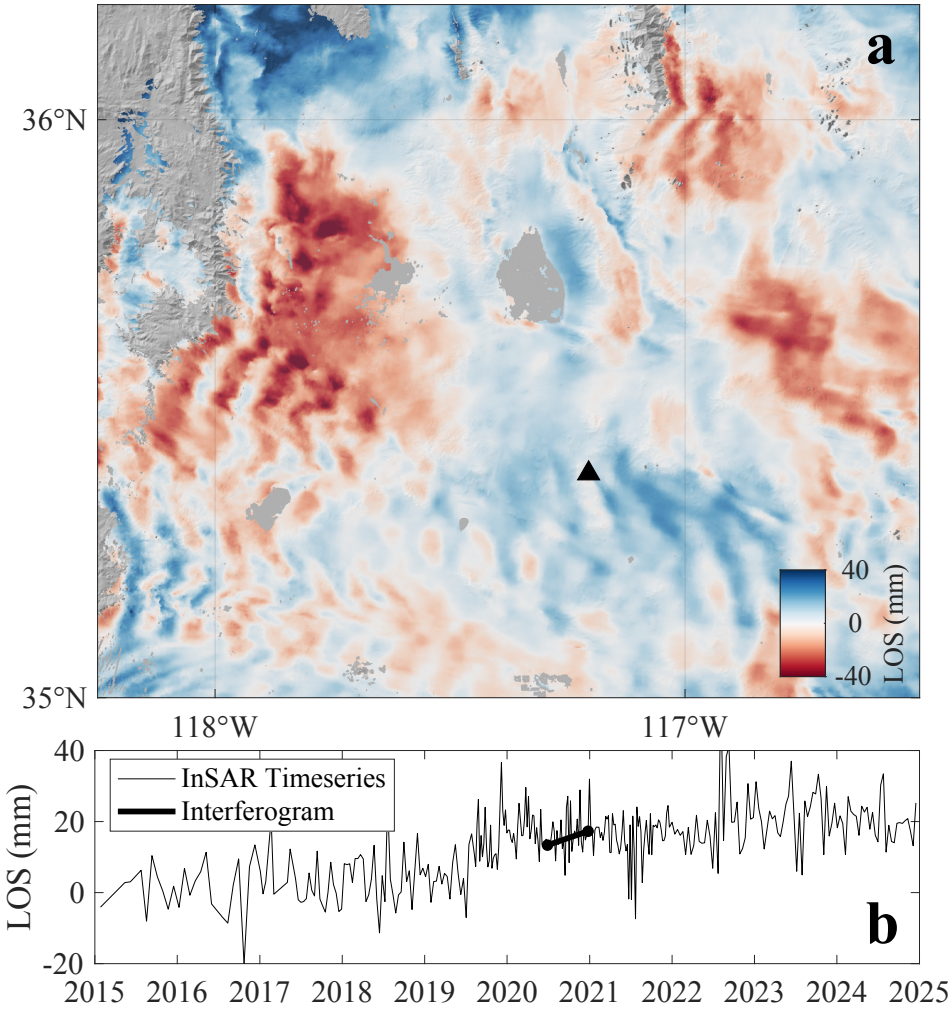


Figure 2.1: a) Sample unwrapped interferogram from June 28th, 2020 to December 25th, 2020 (thick line in b), demonstrating strong and spatially variable radar delay in the troposphere. b) The SBAS timeseries of apparent displacement at the location of the black triangle in the top figure. Tropospheric noise is both high amplitude relative to deformation and is highly variable both spatially and temporally.

noise, which is unique to each SAR acquisition but, crucially, which has the same value for all interferograms that share a given SAR acquisition. Interferometric noise results from strong phase gradients, surface electrical changes, and spatial averaging of the complex phase (i.e. multi-looking). Since interferometric noise is unique to each interferogram, generation of more interferogram pairs from the same set of Single Look Complices (SLCs) can help reduce its impact. Atmospheric delays, by contrast, are physical signals present within the SLC data and cannot be averaged out by the addition of more interferograms. Radar signals propagating through the atmosphere are delayed by variations in total electron content in the ionosphere and by humidity, temperature, and pressure variations in the troposphere [83]. For C-band SAR missions like Sentinel-1, the troposphere is generally considered to be the source of the largest errors in InSAR timeseries analysis [46]. Stratification of the troposphere causes a topography-correlated component of noise [9] which is often corrected by removing a linear trend in elevation. Atmospheric turbulence causes fractal noise with a Kolmogorov power-law slope of -8/3 [28], although anisotropy is often present [34]. A common way to mitigate tropospheric noise is to estimate the radar phase delay from zenith delays [78] or from weather models of atmospheric water vapor[81]. Such approaches often suffer from a lack of spatial resolution [20].

The other major avenue for tropospheric noise mitigation involves timeseries smoothing, for which all methods assume that the tropospheric noise is a random perturbation to a underlying smooth displacement timeseries. One of the earliest steps towards a smoothing approach was taken by Berardino et al. (2002) [10], whose Small-Baseline Subset (SBAS) algorithm remains foundational to all InSAR timeseries generation techniques. The authors impose a minimum-norm velocity condition, thus accomplishing a degree of regularization, but state that this step is only applied when the system of equations is under-determined. Tymofyeyeva and Fialko (2015) [73] apply a common-scene stacking approach to tropospheric noise mitigation. To be effective, the user must terminate the iteration to achieve smoothness (as in [32]) or remove a spline first (which may fail to adequately represent the timeseries [84]); otherwise, taking the stacking approach to convergence results in a pure linear trend. Schmidt and Bürgmann (2003) [64]

regularize their timeseries using a finite-difference smoothing constraint. Neely et al. (2020)[56], in addition to removing the long-wavelength component of tropospheric delay by referencing interferograms to a low-order GNSS polynomial fit, also use Tikhanov regularization to smooth the timeseries. In an improvement to [64], however, they use an L-curve analysis to select the regularization parameter in a data-driven manner. Guns et al. (2022)[32], recognizing that discontinuous displacement signals from earthquakes interfere with tropospheric noise mitigation algorithms, use a parametric inversion to estimate and remove the earthquake step.

We synthesize the above approaches into a consistent and powerful framework based on Bayesian statistics. In the Bayesian interpretation of probability, we condition the data on a prior model which expresses the degree of belief about the behavior of the system. The Bayesian approach is conceptually different from the regularization-based approaches above because the prior is endowed with a direct physical interpretation relating to the expected variance of the timeseries. In turn, the physical interpretation allows us to use an external data source to estimate the prior.

Global Navigation Satellite Systems (GNSS) stations offer an independent measurement of ground surface displacements with which to inform our prior and error models. GNSS timeseries provide significantly better displacement accuracy at the cost of far lower spatial coverage. Although GNSS carrier wave signals are delayed by the atmosphere like InSAR, the GNSS signals arrive from multiple directions (spatially averaging tropospheric delay), feature a dual-frequency ionospheric correction, and are acquired at high rates (temporally averaging tropospheric delay). Since GNSS and InSAR both observe the same deformation signal, discrepancies between the two observations directly measure the amount of noise present, with the majority attributable to the InSAR delay.

We demonstrate the utility of our method of uncertainty quantification with a case study in the Mojave Desert, CA. On July 4$^{\text{th}}$, 2019, a fault in the Mojave near Ridgecrest ruptured in a M6.4 earthquake, followed the next day by a larger M7.1 earthquake on the orthogonal Little Lake fault [23] (Fig. 2.2). The resulting deformation pattern has been extensively studied

with InSAR [48, 44, 80, 76, 32]. Beyond the large amplitude, well-characterized coseismic displacement, the Mojave is an excellent study site for InSAR because its dry, sparsely vegetated surface results in coherent interferograms and because it is well-instrumented with GNSS stations. Furthermore, the NASA-ISRO Synthetic Aperture Radar (NISAR) mission has selected the Mojave as a calibration/validation site, highlighting the importance of characterizing its surface dynamics. In the Mojave, we show how the derived timeseries uncertainties support statistical equality between the InSAR and GNSS displacements. Furthermore, we demonstrate how our approach is necessary to mitigate large-amplitude seasonal tropospheric delays that affect all InSAR-based approaches discussed above but which does not appear in GNSS.

In Part 2.2, we present our theoretical framework for Bayesian uncertainty propagation for InSAR displacement timeseries estimation. In Part 2.3, we detail our data processing steps, with particular focus on how to use GNSS timeseries to estimate models for interferometric error, atmospheric error, and displacement in a straightforward and widely applicable manner. In Part 2.4, we present the results of our uncertainty analysis and how it can be used to understand surface deformation of the Mojave. We end with Discussion (Part 2.5) and Conclusion (Part 2.6) sections.

2.2 Theory

In this section, we present a theoretical framework for the treatment of noise and uncertainty in InSAR timeseries. In Section 2.2.1, we present the Bayesian formalism and notation. In Section 2.2.2, we discuss the uncertainty quantification for the SBAS inversion that converts interferogram stacks to timeseries. In Section 2.2.3, we characterize the formal relationship between SBAS timeseries, atmospheric error and surface displacement timeseries. In Section 2.2.4, we present a generalized prior for the surface displacement. Finally, in Section 2.2.5, we combine the prior with the likelihood and present the analytic expression of the displacement timeseries.

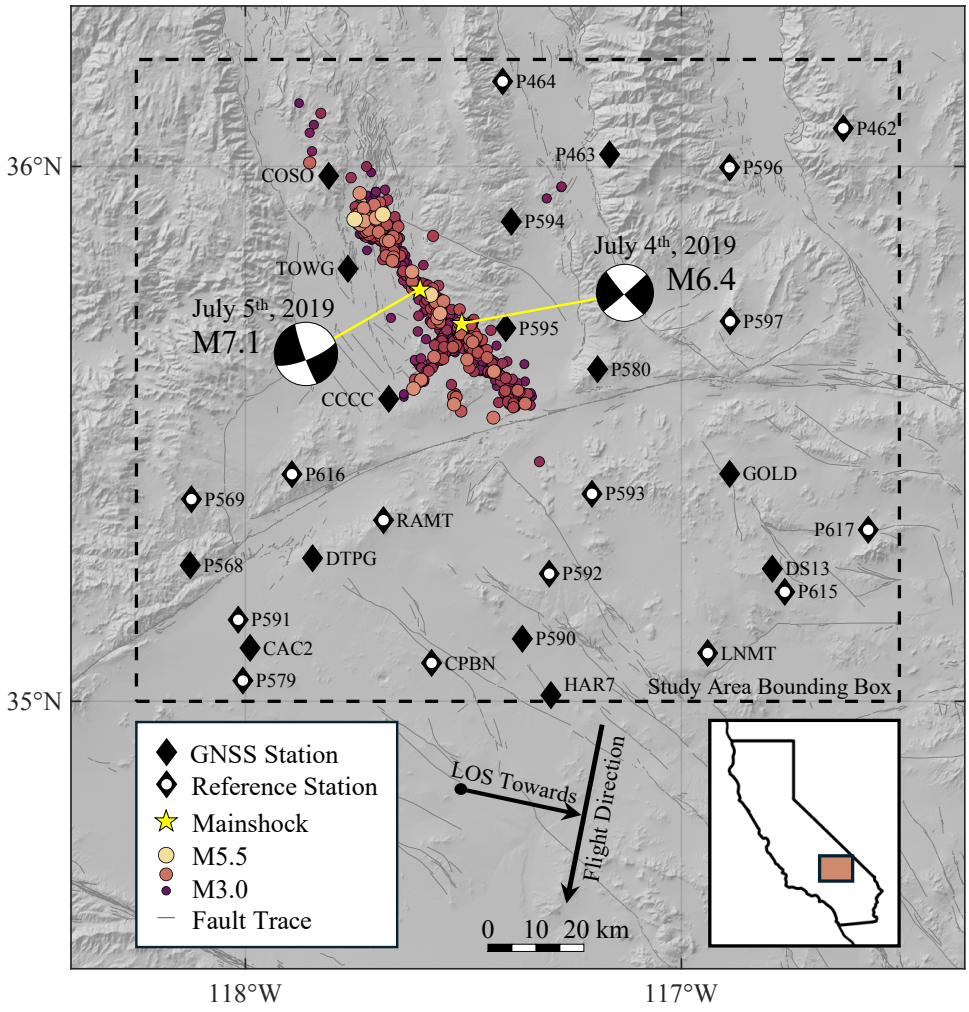


Figure 2.2: Map of the Mojave Desert in Southern California. The study area (dashed black line) is entirely contained within the Sentinel-1 descending frame (not shown). The flight direction and LOS vector are shown. GNSS stations are the black diamonds. GNSS stations used as references are marked with white dots. The Ridgecrest earthquake sequence (from USGS, [69]) is shown, with the aftershocks (>M3.0) colored and sized by their magnitude. The position of the two mainshocks are displayed as yellow stars, with the focal mechanisms attached. Fault traces (grey lines) are from USGS Quaternary Fault Database [70].

2.2.1 Formalism and Notation

In this work, we formalize the problem of displacement timeseries estimation in a statistical sense. We model the timeseries, noises and relevant parameters as random variables (r.v.), which are mathematical objects that take not one value, but encompass a range of values with a statistical likelihood described by a probability density function (PDF). The random variables in this discussion are multi-variate, meaning that the domain of their PDF has multiple dimensions. Multi-variate random variables may be though of as vectors composed of scalar random variables. In the Bayesian framework, a prior is a PDF chosen before an estimation problem to represent the best knowledge-based understanding of the values that the random variable representing the model can take. The likelihood is a function representing the probability of observing the observed data given a particular model. The posterior is the PDF of the model given the likelihood function conditioned on the prior.

In this discussion, scalar values are denoted with normal text (e.g. x). Matrices are denoted with bold capital letters, such as $\mathbf{M}$. Multivariate random variables are denoted with bold lowercase letters such as $\mathbf{x}$ or with bold Greek letters (e.g. $\boldsymbol{\eta}$), the latter of which are largely reserved for noise terms. The prior mean value of a random variable is denoted with the symbol μ_x (with appropriate subscript), while the posterior mean is denoted by a star (e.g. $\mathbf{x}^*$). The means are vectors of real numbers, not random variables. The symbol $\mathbf{R}$ is reserved for error covariance matrices, $\mathbf{B}$ for prior covariance matrices, and $\mathbf{p}$ for posterior covariance matrices, with descriptive subscripts as needed. The symbol $\mathbf{I}$ refers to the identity matrix, where the dimension is inferred from the context.

As a first-order approximation, we describe all random variables in this discussion as having multi-variate normal (i.e. Gaussian) probability distributions. A normal distribution has the PDF of equation 2.1, stating that the probability density of the random variable $\mathbf{v}$ is proportional to a Gaussian function with mean $\mathbf{v}^*$ (a vector) and covariance matrix $\mathbf{V}$. Here $||\mathbf{v}||^2 = \mathbf{v}^T\mathbf{v}$ is the square of the L_2-norm and $\mathbf{V}^{-\frac{1}{2}}$ is the matrix square root of the inverse of

$\mathbf{V}$. Proportionality is sufficient because of the requirement that $p(\mathbf{v})$ integrate to unity over the entire domain in order to be a probability distribution.

$$p(\mathbf{v}) \propto \exp\left(-\frac{1}{2}\left|\left|\mathbf{V}^{-\frac{1}{2}}\left(\mathbf{v}-\mathbf{v}^{*}\right)\right|\right|^{2}\right) \tag{2.1}$$

We commonly use the equivalent but more concise notation below;

$$\mathbf{v} \sim \mathcal{N}\left(\mathbf{v}^{*}, \mathbf{V}\right) \tag{2.2}$$

where $\mathbf{N}$ indicates the Gaussian distribution.

A sample from a random variable is a realization of that variable that is consistent with the corresponding PDF. A sample drawn from a normal distribution is realized as below, where $\mathbf{v}_k$ is the sample, $\mathbf{V}^{\frac{1}{2}}$ is the matrix square root (calculated by eigenvalue decomposition) of the covariance, and φ_k is a sample from the unit normal distribution.

$$\mathbf{v}_k = \bar{\mathbf{v}} + \mathbf{V}^{\frac{1}{2}}\varphi_k \tag{2.3}$$

$$\varphi \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right) \tag{2.4}$$

Each step in our analysis is performed on a pixel-by-pixel basis unless otherwise specified. For clarity, spatial indices are omitted.

2.2.2 SBAS Inversion with Uncertainty Quantification

Displacement is estimated from SAR data by taking the difference in the phase of the radar return signal between two passes of the satellite. Every interferogram is therefore a measure of the apparent displacement between a given primary and secondary date. Apparent displacement refers to physical ground surface displacement perturbed by atmospheric phase delay noise, or in other words, the sum of all signals which change the arrival time of the radar wave to the satellite. Distinct from atmospheric noise, interferometric noise is a numerical signal

which arises in the interferogram generation and multi-looking processes, and which perturbs each interferogram separately.

To construct a timeseries of apparent displacement from a stack of interferograms, we perform a Small-Baseline Subset (SBAS) [10] inversion. The SBAS system of equations with the inclusion of the interferometric error is presented in matrix form in equation 2.5, where y_{ij} is the interferometric LOS displacement value from the i-th to the j-th SAR acquisition, z_k is the SBAS timeseries value at the k-th SAR acquisition date, and η_{ij} is the interferometric noise r.v. associated with the ij-th interferogram. Since displacement is a relative measure, we can set z_0 to any value; by construction, we set it to zero and interpret the other z_k values as the displacement relative to the date of zeroth SAR acquisition. This system of equations can also be collected into vector form (equation 2.6), where $\mathbf{y}$ is the vector of interferogram values, $\mathbf{M}$ is the SBAS matrix, $\mathbf{z}$ is the SBAS timeseries vector at the SAR postings, and η is the interferometric noise random variable.

$$\begin{pmatrix} y_{01} \\ y_{02} \\ y_{12} \\ y_{13} \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ -1 & 1 & 0 & \dots \\ -1 & 0 & 1 & \\ & \vdots & & \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix} + \begin{pmatrix} \eta_{01} \\ \eta_{02} \\ \eta_{12} \\ \eta_{13} \\ \vdots \end{pmatrix} \tag{2.5}$$

$$\mathbf{y} = \mathbf{M} \mathbf{z} + \eta \tag{2.6}$$

We assume the interferometric error is normally distributed and zero-mean with covariance matrix $\mathbf{R}_\eta$.

$$\eta \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_\eta) \tag{2.7}$$

Precision orbit control of modern SAR constellations allows for connected interferogram

networks where the entire stack of interferograms is part of the same small baseline subset. Under these conditions, the SBAS system of equations above is over-determined and no regularization is needed. (Additional desired details may be found in the derivation of equation 9 of Berardino et al. 2002 [10].) In Bayesian terminology, lack of regularization is equivalent to choosing a uniform prior ($p(\mathbf{z}) = const.$). Applying Bayes' rule, the posterior distribution will be equal to the likelihood $p(\mathbf{z}|\mathbf{y}) = p(\mathbf{y}|\mathbf{z})$. As such, equation 2.6 is solved exactly as a weighted least-squares problem in a frequentist manner, where the weight matrix is equal to $\mathbf{R}_\eta^{-1}$. The solution is the normal probability distribution of $\mathbf{z}$ described below in equation 2.8. Below, $\mathbf{z}^*$ is the mean of the normal distribution (eq. 2.9) and $\mathbf{p}_z$ is the covariance matrix (eq. 2.10).

$$\mathbf{z} \sim \mathcal{N}(\mathbf{z}^*, \mathbf{p}_z) \tag{2.8}$$

$$\mathbf{z}^* = \mathbf{p}_z \mathbf{M}^T \mathbf{R}_\eta^{-1} \mathbf{y} \tag{2.9}$$

$$\mathbf{p}_z = \left(\mathbf{M}^T \mathbf{R}_\eta^{-1} \mathbf{M}\right)^{-1} \tag{2.10}$$

We reincorporate the zero-th SAR acquisition $z_0 = 0$, with the condition of zero uncertainty arising from having defined the value z_0 to be zero.

$$\mathbf{z}^* \rightarrow \begin{pmatrix} 0 \\ \mathbf{z}^* \end{pmatrix} \tag{2.11}$$

$$\mathbf{p}_z \rightarrow \begin{pmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{p}_z \end{pmatrix} \tag{2.12}$$

2.2.3 Atmospheric Error and the Likelihood Function

Atmospheric noise is the contribution of all radar delays in the ionosphere and troposphere to the apparent displacement of the SBAS timeseries. In the case that we sample the displacement timeseries $\mathbf{x}$ on the same temporal posting as the SAR acquisitions, the expression relating the SBAS timeseries $\mathbf{z}$ to the displacement timeseries $\mathbf{x}$ would merely be $\mathbf{z} = \mathbf{x} + \xi$, the displacement

plus the atmospheric noise ξ. However, since SAR acquisitions are often irregular, we wish to interpolate the timeseries on a regular posting. We accomplish this by introducing a data availability matrix $\mathbf{G}$, defined in equation 2.13, which maps the more densely sampled times at which $\mathbf{x}$ is queried to the more sparse SAR acquisition times. Below, T_i is the i-th SAR posting and t_j is the j-th query epoch.

$$[\mathbf{G}]_{ij} = \begin{cases} 1 & T_i = t_j \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

The matrix equation relating the SBAS timeseries $\mathbf{z}$ to the physical displacement $\mathbf{x}$ and tropospheric noise ξ then becomes the following equation.

$$\mathbf{z} = \mathbf{G}\mathbf{x} + \xi \tag{2.14}$$

The atmospheric error ξ is assumed to be zero-mean and normally distributed. The exact form of the covariance matrix $\mathbf{R}_\xi$ may be tailored to each application but will generally be a diagonal matrix. See Section 2.3.3 for an example.

$$\xi \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_\xi\right) \tag{2.15}$$

The associated likelihood function for equation 2.14, denoted $p(\mathbf{z}|\mathbf{x})$ (eq. 2.16), describes the probability of finding a value of the SBAS timeseries mean $\mathbf{z}^*$ for a given displacement timeseries $\mathbf{x}$. Here the error covariance $\mathbf{R}_x$ accounts not only for the atmospheric error but also the uncertainty of the SBAS timeseries (eq. 2.10).

$$p(\mathbf{z}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}\left|\left|\mathbf{R}_x^{-\frac{1}{2}}\left(\mathbf{z}^* - \mathbf{G}\mathbf{x}\right)\right|\right|^2\right) \tag{2.16}$$

$$\mathbf{R}_x := \mathbf{R}_\xi + \mathbf{p}_z \tag{2.17}$$

2.2.4 Constructing the Displacement Prior

The displacement prior is the probability distribution representing our understanding of the ground surface displacement timeseries. The practical effect of a prior will be to regularize, or smooth, the timeseries. Unlike with the SBAS inversion, the extraction of the ground surface displacement timeseries from the noisy SBAS timeseries cannot be performed without a prior model because the problem is under-determined (i.e. there are more unknowns than knowns). In principle, the apparent displacement signal in an SBAS timeseries may arise equally from true ground motion or atmospheric noise. In practice, however, atmospheric noise in the time domain has a pattern of randomness well approximated by white noise, permitting estimation of the true displacement by using a smoothing prior. In this section, we describe the general theory behind the construction of such a prior from multiple components.

We model surface displacement timeseries as the sum of two types of signal: 1) a function with well-defined temporal behavior $\mathbf{a}(\theta)$, whose value is known up to a number of parameters θ; and 2) a stochastic term denoted by the r.v. β, which accounts for non-parametric behavior.

$$\mathbf{x} = \mathbf{a}(\theta) + \beta \tag{2.18}$$

In general, the parameterized timeseries $\mathbf{a}(\theta)$ will be a non-linear function of the parameter vector θ. We restrict our discussion here to the special case $\mathbf{a}(\theta) = \mathbf{A}\theta$, where $\mathbf{A}$ is a constant design matrix whose columns are basis functions. For example, the columns of $\mathbf{A}$ would be unit cosine and sine functions with a period of one year if the signal had a predominantly annual behavior. Under this condition, inversion linearity is preserved. The displacement timeseries $\mathbf{x}$ then takes the following form:

$$\mathbf{x} = \mathbf{A}\theta + \beta \tag{2.19}$$

We assume normal prior distributions for θ and β, the exact form of which may be

specific to each use case (see Section 2.3.5 for details). We only note here that by selecting a sufficiently broad (i.e. high-variance) prior for θ, the user may essentially allow for a constraint-free inversion with respect to the parameter values.

$$\theta \sim \mathcal{N}(\mu_\theta, \mathbf{B}_\theta) \tag{2.20}$$

$$\beta \sim \mathcal{N}(\mu_\beta, \mathbf{B}_\beta) \tag{2.21}$$

We combine the two priors into a unified prior $p(\mathbf{x})$ for the displacement timeseries using the additive and multiplicative properties of normal random variables.

$$p(\mathbf{x}) \propto \exp\left(-\frac{1}{2}\left|\left|\mathbf{B}_x^{-\frac{1}{2}}(\mathbf{x}-\mu_x)\right|\right|^2\right) \tag{2.22}$$

$$\text{where } \mu_x = \mathbf{A}\mu_\theta + \mu_\beta \tag{2.23}$$

$$\text{and } \mathbf{B}_x = \mathbf{A}\mathbf{B}_\theta\mathbf{A}^T + \mathbf{B}_\beta \tag{2.24}$$

2.2.5 Bayesian Displacement Timeseries Estimation

With the likelihood $p(\mathbf{z}|\mathbf{x})$ (eq. 2.16) and prior $p(\mathbf{x})$ (eq. 2.22) defined, we estimate the probability distribution of the posterior $p(\mathbf{x}|\mathbf{z})$ as the likelihood conditioned on the prior, following Bayes' rule:

$$\begin{aligned} p(\mathbf{x}|\mathbf{z}) &\propto p(\mathbf{z}|\mathbf{x})p(\mathbf{x}) \\ &\propto \exp\left(-\frac{1}{2}\left(\left|\left|\mathbf{R}_x^{-\frac{1}{2}}(\mathbf{z}^* - \mathbf{G}\mathbf{x})\right|\right|^2 + \left|\left|\mathbf{B}_x^{-\frac{1}{2}}(\mathbf{x}-\mu_x)\right|\right|^2\right)\right) \end{aligned}$$

The derivation of the analytical form of solution below will not be presented here, but may be found in many standard texts on Bayesian statistics (e.g. [29], [67]). The solution for the posterior PDF is normal with posterior mean $\mathbf{x}^*$ and posterior covariance matrix $\mathbf{p}_x$ (eq. 2.25). To simplify the expressions for the mean and covariance, we calculate the Kalman gain matrix $\mathbf{K}$

(eq. 2.28), which may be thought of as the degree to which a data point updates the prior.

$$\mathbf{x}|\mathbf{z} \sim \mathcal{N}(\mathbf{x}^*, \mathbf{p}_x) \tag{2.25}$$

$$\mathbf{x}^* = \mu_x + \mathbf{K}(\mathbf{z}^* - \mathbf{G}\mu_x) \tag{2.26}$$

$$\mathbf{p}_x = (\mathbf{I} - \mathbf{K}\mathbf{G})\mathbf{B}_x \tag{2.27}$$

$$\text{where } \mathbf{K} := \mathbf{B}_x\mathbf{G}^T(\mathbf{G}\mathbf{B}_x\mathbf{G}^T + \mathbf{R}_x)^{-1} \tag{2.28}$$

The mean of the posterior r.v. $\mathbf{x}^*$ represents the regularized timeseries of displacement derived from the InSAR data. The posterior covariance $\mathbf{p}_x$ captures the information about the uncertainty of that timeseries.

2.3 Methods

In this section, we describe the methods used to apply the theoretical framework above to real InSAR data in the Mojave Desert. In Section 2.3.1, we describe the interferogram acquisition and pre-processing, including the referencing scheme. In Section 2.3.2, we discuss the processing of the GNSS timeseries used in this project. In Sections 2.3.3 and 2.3.4, we construct error models for the interferometric noise η and atmospheric delay noise ξ respectively. In Section 2.3.5, we construct a prior model for the parameterized component of displacement θ to include secular, coseismic, and postseismic terms. In Section 2.3.6, we use GNSS to create a prior for the non-parametric component of displacement β. Finally, in Section 2.3.7, we present the practical details of the algorithm to estimate displacement timeseries.

2.3.1 InSAR Pre-processing

Sourcing InSAR Data Products

We source Sentinel-1 interferograms from the Alaska Satellite Facility's (ASF) Vertex portal. NASA JPL's Advanced Rapid Imaging and Analysis project (ARIA) has pre-processed and made available on Vertex unwrapped interferograms in NISAR mission format in various

regions, including those selected for NISAR calibration/validation. We download interferograms intersecting a study area in the Mojave Desert in southern California, spanning 35°N to 36.2°N and 118.25°W to 116.5°W. We use 1936 interferograms from descending track 71 over 332 unique SAR acquisitions (postings) ranging from January 26th, 2015 to December 16th, 2024.

Interferogram Masking and Referencing

Decorrelation noise is a spatial noise present in interferograms where the gradient of the interferometric phase is so large that it is not possible to correctly unwrap the phase, resulting in non-deformation signals and introducing numerical artifacts. Regions within interferograms can decorrelate in the presence of strongly variable displacements (such as result from large earthquakes) or changes to the near-surface electrical environment from factors like snowfall, vegetation growth, or moisture changes. We generate a pixel mask based on the stack of interferogram coherences. We mask out all pixels for which fewer than 90% of interferograms have a coherence value above 0.5. The mask removes most pixels at high elevations, particularly in the Sierra Nevada in the northwest corner of the study area, as well as over Searles Lake and other periodically flooded depressions. This pixel mask is applied identically to all interferograms.

The remaining pixels are highly coherent in most interferograms. However, some interferograms show much poorer coherence overall. To ensure that we are working with a stack of high-coherence interferograms, we keep only interferograms with a temporal baseline shorter than 200 days and for which 90% of pixels have a coherence above 0.9. This reduces the number of interferograms used to 1326 over 328 unique SAR acquisitions.

Due to the ambiguity of phase unwrapping, each interferogram is unique only up to a constant. To ensure consistency with the GNSS, we reference each interferogram to a set of colocated GNSS stations. We exclude from the reference set the five stations closest to the earthquake epicenter (CCCC, P580, P594, P595, and TOWG) as well as any station with more than 2% of data missing during the study period. In total, we use 15 of the 29 available GNSS stations as our reference GNSS stations, leaving the remaining 14 for analysis only.

To reference each interferogram, we calculate the mean value of all pixels within 1 km of all of the reference stations, and subtract this mean from the entire interferogram. Since the 15 reference stations are widely and evenly spaced across the study area, the mean value of the reference pixels for the InSAR are more representative of the mean value of the tropospheric delay across the whole region than a single reference pixel would be. See [84] for a discussion about the errors resulting from choosing a single point reference.

We do not correct interferograms for stratified tropospheric delay. See the Discussion section for details about characterizing the turbulent and stratified components of tropospheric delay.

Closure Phase and Masking

As an additional quality measure, we examine the internal consistency of the interferogram stack, essentially asking how well the interferograms agree with each other. The circular sum of interferometric phase over a closed cycle in the interferogram network (e.g. $y_{ij}+y_{jk}-y_{ik}$), known as closure phase, often has a non-zero value in stacks of multi-looked interferograms. Non-zero closure phase has been attributed to both statistical [53] and systemic physical effects [5], most notably soil moisture changes [18]. For our purposes, large non-zero closure phase for a given pixel indicates that the apparent displacement at that location is being obscured by decoherence noise.

We generalize the concept of closure phase by defining the SBAS residual $\mathbf{r}_\eta$ below. The SBAS residual is the amount by which the interferometric values $\mathbf{y}$ differ from the SBAS timeseries (eq. 2.30), which we calculate using the unweighted ordinary linear least-squares solution $\bar{\mathbf{z}}$ to the SBAS matrix equation (eq. 2.29). Whereas closure phase which is defined per cycle, the SBAS residual is defined per interferogram. When the network of interferograms is

internally inconsistent, both non-zero closure phase and non-zero SBAS residual will result.

$$\bar{\mathbf{z}} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{y} \tag{2.29}$$

$$\mathbf{r}_\eta := \mathbf{y} - \mathbf{M}\bar{\mathbf{z}} \tag{2.30}$$

We calculate the root-mean-square deviation (RMSD) of the SBAS system of equations as $\text{RMSD}_{\text{SBAS}} = \text{RMS}(\mathbf{r}_\eta)$ to be a metric for the overall internal inconsistency of the interferogram stack at a given pixel. We mask out all pixels with a $\text{RMSD}_{\text{SBAS}}$ greater than 1.3 mm, which is the 95$^{\text{th}}$ percentile of the $\text{RMSD}_{\text{SBAS}}$ distribution and the point where the corresponding histogram drops off. This RMSD mask eliminates additional pixels in the Sierra Nevada, high elevations within the Mojave (e.g. the Panamint Mountains), decorrelating bodies of water such as Searles Lake, and the immediate fault zone of the Ridgecrest earthquake.

2.3.2 GNSS Timeseries Preparation

We source GNSS timeseries from University of Nevada Reno's Nevada Geodetic Laboratory (NGL). We use daily East-North-Up (ENU) position solutions in the IGRS-14 reference frame from January 1$^{\text{st}}$, 2015 to September 28$^{\text{th}}$, 2024. There are 40 total GNSS stations in the study area which have any data available during the time window above. We reject stations BEPK, KENN, and WORG because they are located in the Sierra Nevada where the InSAR data is entirely decorrelated. We reject stations P811, P812, GOL2, GOL3, and GOL4 because they are nearly colocated with (i.e., within 100 m of) another GNSS station. Finally, we reject stations BEER, COSJ, and PEAR because they have lack sufficient temporal coverage for timeseries analysis, leaving 29 remaining GNSS stations. To remove outliers, we take the 60-day moving window median and median absolute deviation (MAD), then remove data points whose absolute difference from the moving median is larger than $10\times$MAD. We additionally remove the Ridgecrest earthquake dates, July 5$^{\text{th}}$ and 6$^{\text{th}}$, 2019 (UTC), from the GNSS timeseries.

Offsets in GNSS timeseries are step signals which may arise from antenna or other

equipment changes, nearby earthquakes, or unknown causes. We follow the steps below to both remove the offsets and temporally reference the GNSS timeseries to allow consistent comparison with InSAR.

1. From the list of offsets provided by UNR, select offsets which are the result of antenna changes. We exclude the offsets at the Ridgecrest earthquake dates because we wish to preserve, not correct, the coseismic step-like displacement.

2. For each station and along each coordinate axis, fit the GNSS timeseries with the functional fit in equation 2.31, which includes constant, secular velocity, co-seismic offset, post-seismic, and flagged offset terms.

$$f(t) = a_0 + a_1 t + a_2 H(t - t_{\mathrm{EQ}}) + a_3 f_{\mathrm{PS}}(t) + \sum_{i=1}^{N_{\mathrm{Offset}}} b_i H(t - t_i) \tag{2.31}$$

3. Subtract from the GNSS timeseries the constant a_0 and flagged offset terms $b_i H(t - t_i)$.

In the above, t is the time in years relative to the reference date of January 1$^{\mathrm{st}}$, 2015, t_{EQ} is the time of the Ridgecrest earthquake, t_i is the time of the i-th flagged offset, and $H(x)$ is the Heaviside step function. Following Guns et al (2022).[32], the post-seismic displacement function $f_{\mathrm{PS}}(t)$ is exponential with characteristic timescale τ_{PS}:

$$f_{\mathrm{PS}}(t) = H(t - t_{\mathrm{EQ}}) \left(1 - \exp\left(-\frac{t - t_{\mathrm{EQ}}}{\tau_{\mathrm{PS}}}\right)\right) \tag{2.32}$$

Guns et al. (2022) [32] use a post-seismic timescale $\tau_{\mathrm{PS}} = 0.5$ years. However, we performed a 1D grid search and found that $\tau_{\mathrm{PS}} = 1.8$ years minimized the root-mean-square deviation, which we used in this analysis.

After the temporal referencing, we project the GNSS displacements into the InSAR line-of-sight (LOS) direction, then spatially reference the GNSS. Our method of GNSS spatial

referencing is identical to the procedure used for the InSAR spatial referencing, enforcing consistency and ensuring that we are always comparing like-to-like:

1. For each of the 15 reference stations, we linearly interpolate the LOS displacement timeseries over all data gaps less than 60 days long to avoid abrupt changes in the mean when a station has a missing value. Note that the interpolated values are discarded after the referencing procedure.

2. We calculate the mean value of the LOS displacement across the 15 reference GNSS stations for each date separately.

3. We then subtract this value from all 29 stations for each date separately.

By performing this spatial referencing, we recognize that we will only be able to observe relative displacements and that the motions observed should not be considered to be in an absolute reference frame.

2.3.3 Interferometric Noise Error Model

In our formalism, noise comes in two flavors: interferometric noise, which has a different value for each interferogram; and atmospheric noise, which has a different value for each SAR acquisition but the same value for all interferograms which share a given SAR acquisition. In this and the next subsection, we will present straightforward methods for investigating both sources of noise and will develop error models to include in the timeseries inversion.

Interferometric noise is a signal which takes a different value for each interferogram and causes internal inconsistencies in stacks of interferograms. We assume a uniform and uncorrelated PDF with a standard deviation of r_η. The interferometric error covariance matrix $\mathbf{R}_\eta$ is then proportional to the identity matrix.

$$\mathbf{R}_\eta = r_\eta^2 \mathbf{I} \tag{2.33}$$

In Section 2.3.1, we defined the SBAS residual $\mathbf{r}_\eta = \mathbf{y} - \mathbf{M}\bar{\mathbf{z}}$ (eq. 2.30) as a measure of the internal inconsistency of a stack of interferograms. Equally, the SBAS residual may be thought of as sampling the interferometric noise random variable η. To select a value for r_η, the amplitude of the interferometric noise, we take the root-mean-square of the SBAS residual $\mathbf{r}_\eta$ over all interferograms and pixels $r_\eta = \text{RMS}(\mathbf{r}_\eta)$. The amplitude r_η has a value of approximately 0.56 mm.

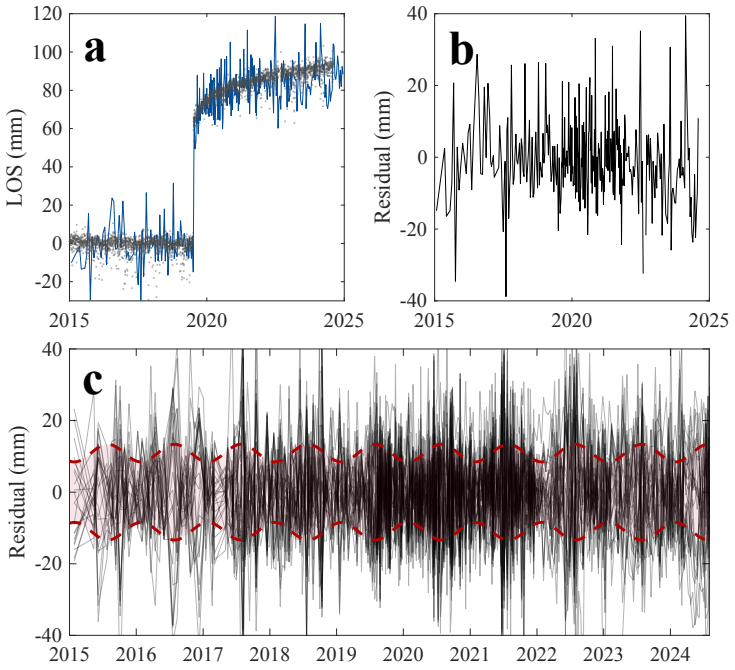


Figure 2.3: a) The P594 GNSS timeseries (grey dots) and the colocated InSAR SBAS timeseries (blue). b) The residual timeseries formed by subtracting the GNSS timeseries from the SBAS timeseries. c) We aggregate residual timeseries (black) from all 29 GNSS stations, demonstrating their distinct seasonal character. The 1σ uncertainty bounds of our seasonal tropospheric error model is shown in red.

2.3.4 Atmospheric Delay Error Model

Tropospheric delay, the largest component of noise present in our SBAS timeseries, is the primary cause of mismatch between GNSS and InSAR timeseries. The difference between the timeseries can be leveraged to give direct insight into the temporal behavior of the tropospheric delay noise at the GNSS station locations.

For each GNSS station, we subtract the GNSS displacement values on the SAR acquisition dates from the InSAR SBAS timeseries $\bar{\mathbf{z}}$ at the station-colocated pixels (Fig. 2.3a), giving residual timeseries corresponding to tropospheric delay (Fig. 2.3b). We observe that the residual tropospheric delay is dominated by random high-frequency noise, consistent with weather patterns which change on a daily basis. However, the tropospheric delay timeseries also have a strong seasonal character (Fig. 2.3c). The amplitude of the noise is much larger during the summer months than the winter (see Discussion).

To construct our atmospheric delay error model, we aggregate the tropospheric delay timeseries values across all 29 GNSS stations for each SAR acquisition. We take the root-mean-square (RMS) of these values, generating a timeseries of RMS. We then fit a constant and a seasonal sine and cosine (eq. 2.34) to the RMS timeseries (dashed red lines in Fig. 2.3c).

$$\sigma(t) = \sigma_0 + \sigma_1 \cos(2\pi t) + \sigma_2 \sin(2\pi t) \tag{2.34}$$

We construct our error model PDF (eq. 2.35) for the r.v. of tropospheric delay ξ to account for the observed seasonal behavior. We assume the error is zero-mean and uncorrelated (meaning a diagonal covariance matrix) but that the standard deviation changes seasonally according to $\sigma(t)$ (eq. 2.36) evaluated at the vector of SAR posting times $\mathbf{T}$.

$$\xi \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_\xi\right) \tag{2.35}$$

$$\text{where } \mathbf{R}_\xi = \text{diag}(\sigma(\mathbf{T})^2) \tag{2.36}$$

We then test that the additional complexity of our seasonal error model, relative to a constant error model, is justified. We perform two Komlogorov-Smirnov (KS) tests, which test the normality of a distribution (i.e. how close a PDF is to the unit Gaussian PDF). For the first test, we normalize the residual timeseries by the overall standard deviation (i.e. $\sigma(t) = const.$), returning a p-value of 10^{-20}. This result implies that uniform white noise is a poor error model

for the tropospheric delay. For the second test, we normalize the residual timeseries by the seasonal fit $\sigma(t)$ to the standard deviation (eq. 2.34), returning a p-value of 10^{-5}. Although this result shows that the seasonal error model is also not a perfect statistical description of the tropospheric delay, we assess that the $\mathcal{O}(10^{15})$ improvement in p-value exhibits its utility as an error model.

2.3.5 Prior for Displacement Parameters

We parameterize our timeseries using a function with the following terms: a constant, a secular velocity, a co-seismic step function centered on July 5$^{\text{th}}$, 2019, and a post-seismic exponential function (eq. 2.32).

$$a(t) = \theta_0 + \theta_1 t + \theta_2 H(t - t_{\text{EQ}}) + \theta_3 f_{\text{PS}}(t) \tag{2.37}$$

The corresponding design matrix $\mathbf{A}$ is:

$$\mathbf{A} = \begin{pmatrix} 1 & t_1 & H(t_1 - t_{\text{EQ}}) & f_{\text{PS}}(t_1) \\ 1 & t_2 & H(t_2 - t_{\text{EQ}}) & f_{\text{PS}}(t_2) \\ 1 & t_3 & H(t_3 - t_{\text{EQ}}) & f_{\text{PS}}(t_3) \\ & & \vdots & \end{pmatrix} \tag{2.38}$$

For each of the parameters, we have the option of selecting a "strongly informative" prior, based on the actual value expected from GNSS or another data source, or a "weakly informative" prior, which allows the InSAR data to entirely drive the inversion with respect to that parameter. We choose to use weakly informative priors for each parameter (see Discussion). For the prior model for $\theta = [\theta_0\ \theta_1\ \theta_2\ \theta_3]^{\text{T}}$, we choose a zero-mean normal distribution with standard deviation (σ_θ) of 20000 (millimeters for the constant, co-seismic, and post-seismic parameters, millimeters per year for the velocity parameter). The value of 20000 mm is chosen to be ten times larger than the largest co-seismic displacements ($\approx \pm$2m, see [7]). The prior covariance matrix $\mathbf{B}_\theta$ is

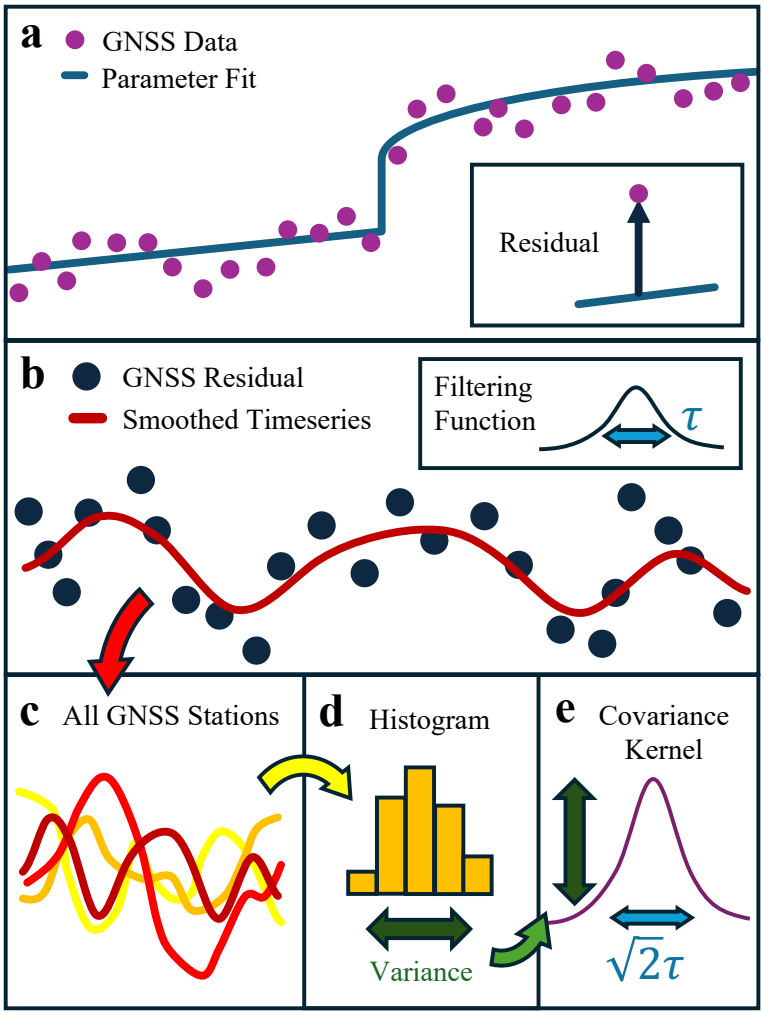


Figure 2.4: Method for estimation of the amplitude of the covariance kernel. We take the residual to the parameter fit (a), smooth the residual timeseries (b), and take the variance across all times and stations (c-d). The timescale of the Gaussian covariance kernel (e) is a factor of $\sqrt{2}$ greater than the filtering timescale.

proportional to the identity matrix with four dimensions $\mathbf{I}_4$.

$$\theta \sim \mathcal{N}(\mathbf{0}, \mathbf{B}_\theta) \tag{2.39}$$

$$\text{where } \mathbf{B}_\theta = \sigma_\theta^2 \mathbf{I}_4 \tag{2.40}$$

2.3.6 Prior for Non-Parametric Component of Displacement

The stochastic non-parametric signal $\beta \sim \mathcal{N}\left(\mathbf{0}, \mathbf{B}_\beta\right)$ is the residual component of the surface displacement timeseries, or in other words, everything which is not explicitly included in the parametric equation 2.37. We assume this residual component is zero-mean. We construct the covariance matrix $\mathbf{B}_\beta$ with a Gaussian kernel to enforce smoothness of the displacement timeseries. The covariance matrix

$$[\mathbf{B}_\beta]_{ij} = b^2 \exp\left(-\frac{\left(t_i - t_j\right)^2}{2(\sqrt{2}\tau)^2}\right) \tag{2.41}$$

is constructed with two hyper-parameters, the timescale τ and the covariance scale b. The factor of $\sqrt{2}$ is the difference between the filter timescale and the equivalent covariance timescale (see Supplemental Materials).

We select the timescale τ to be 30 days, which is significantly longer than the satellite repeat time of 6 to 12 days. This choice prevents aliasing of tropospheric noise while preserving sub-seasonal dynamics. We use the GNSS displacement timeseries to estimate the covariance scale hyper-parameter b:

1. For each GNSS station, we fit the GNSS timeseries using equation 2.37 (i.e. the same parameter fit as the InSAR) and take the residual (Fig. 2.4a).

2. We filter these residual timeseries with a Gaussian filter of width τ (Fig. 2.4b).

3. The covariance amplitude b is the root-mean-square of every epoch of the smoothed timeseries across all stations (Fig. 2.4c-e).

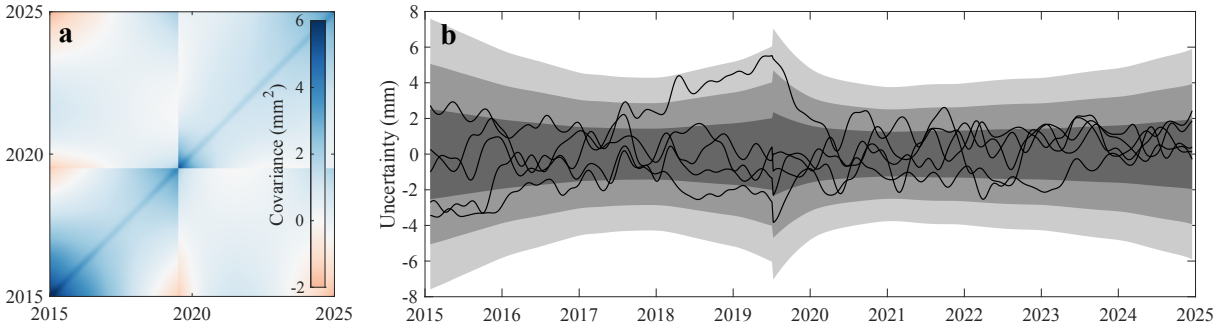


Figure 2.5: In a), we show the posterior covariance matrix. The timeseries of uncertainty at the 1σ (68.3%), 2σ (95.4%), and 3σ (99.7%) levels are shown in b) as the dark, medium, and light solid bands respectively. Five sample timeseries (black lines) are drawn from the posterior distribution and displayed with the uncertainty.

We find a covariance amplitude of 0.74 mm.

2.3.7 Timeseries Estimation Algorithm

In the previous sections, we constructed the necessary error models and priors to apply to the Bayesian framework. We now briefly describe the practical steps taken to estimate the displacement timeseries from them.

1. Using the interferometric error covariance $\mathbf{R}_\eta$, calculate the SBAS posterior covariance $\mathbf{p}_z$ (eq. 2.10), then the SBAS posterior mean $\mathbf{z}^*$ (eq. 2.9).

2. Temporally reference the SBAS posterior mean to the reference date of January 1$^{\text{st}}$, 2015 using an approach equivalent to the GNSS temporal referencing (Section 2.3.2).

3. Sum the atmospheric error covariance and the SBAS posterior covariance to get the error covariance $\mathbf{R}_x$ (eq. 2.17).

4. Calculate the prior mean (eq. 2.23) and prior covariance (eq. 2.24) for $\mathbf{x}$.

5. Calculate the Kalman gain matrix (eq. 2.28). For numerical reasons, we perform a matrix solve with MATLAB's slash operator (i.e. `mrdivide`), which uses a Cholesky solver, instead of an explicit matrix inverse.

6. Calculate the posterior mean (eq. 2.26) and posterior covariance (eq. 2.27).

2.4 Results

In this section, we demonstrate the results of our uncertainty quantification and comparison with GNSS. In Section 2.4.1, we describe the uncertainties determined by our method. In Section 2.4.2, we qualitatively compare the InSAR timeseries to GNSS. In Section 2.4.3, we generate maps of the velocity, coseismic, and postseismic components of InSAR-derived displacement and compare the InSAR to the GNSS in a statistically robust manner.

2.4.1 Uncertainty Quantification

Perhaps counter-intuitively, the timeseries uncertainties are completely uniform in space. This result directly follows from: 1) our choice to make the prior and error models spatially uniform (see Discussion) and; 2) the lack of dependence of the equation for the posterior covariance matrix (eq. 2.27) on the actual values of the data $\mathbf{y}$. This choice brings the immediate benefit of reducing the storage requirements for the posterior covariance from $N_t \times N_t \times N_{pix}$ to $N_t \times N_t$ and also greatly speeds up computation by permitting vectorization of the system of equations.

We show three different but equally valid ways of understanding the uncertainty quantification in Figure 2.5. The posterior covariance (Fig. 2.5a) shows mostly non-zero covariances and a clear discontinuity at the earthquake date resulting from the Heaviside function in the parameter fit. The timeseries of uncertainty (Fig. 2.5b, solid bands) for a normal probability distribution is the square root of the diagonal of the covariance matrix. The average uncertainty at any given date is around 3-4 mm (at the 2σ level), but rises to 5 mm at the earthquake date and the beginning and end of the timeseries. At the beginning of the timeseries, the uncertainty is greatest, which we attribute to the relative lack of observations compared to the later portion of the timeseries. We attribute the larger uncertainty around the date of the earthquake to the additional degrees of freedom due to the coseismic and postseismic parameter fits. Finally, in

Table 2.1: Formal uncertainties of the parameters and Pearson correlation coefficients between pairs of parameters

Parameter	1σ Uncertainty	Parameter Pair	Pearson C.C.
Velocity	0.79 mm/yr	Velocity & Coseismic	-0.19
Coseismic Displacement	2.4 mm	Velocity & Post-seismic	-0.83
Post-seismic Displacement	5.2 mm	Co- & Post-seismic	-0.27

Fig. 2.5b, we draw 5 samples from the posterior distribution to demonstrate the characteristic smoothness resulting from our choice of covariance timescale τ of 30 days.

We can leverage the timeseries uncertainties to understand their impact on the uncertainty of derived parameters. We draw 10,000 samples from the posterior, find the linear-least squares best fit to our parameterized equation 2.37 for each, and calculate the empirical parameter covariance matrix. From this covariance matrix, we calculate the parameter uncertainty as well as the Pearson correlation coefficient (ρ) between each pair of parameters (Table 2.1). For brevity, we exclude the constant parameter (which is an arbitrary displacement) from this analysis.

Each pair of parameters are negatively correlated with each other, but the trade-off is greatest between the velocity and post-seismic parameters, which are strongly anti-correlated with a ρ of -0.83. We believe that this effect may be most easily understood by considering the frequency content of the various signals. The velocity and post-seismic functions are both most sensitive to long period changes and trade off with each other which they are fit to the data. The co-seismic displacement, by contrast, is the most impulsive signal, preferentially sampling higher frequencies. This results in lower correlation between the coseismic parameter on the one hand and the velocity and post-seismic parameters on the other.

For completeness sake, we calculate the residual RMSD from the least-squares fit to the 10,000 samples above. The RMSD has a value of 0.67 mm, which is consistent with the non-parametric component of the signal in the prior which, has an amplitude of 0.74 mm. This posterior non-parametric residual component may be found as the positive (blue) band along the diagonal of the covariance matrix (Fig. 2.5a) or in the short-period variability of the samples in Figure 2.5b.

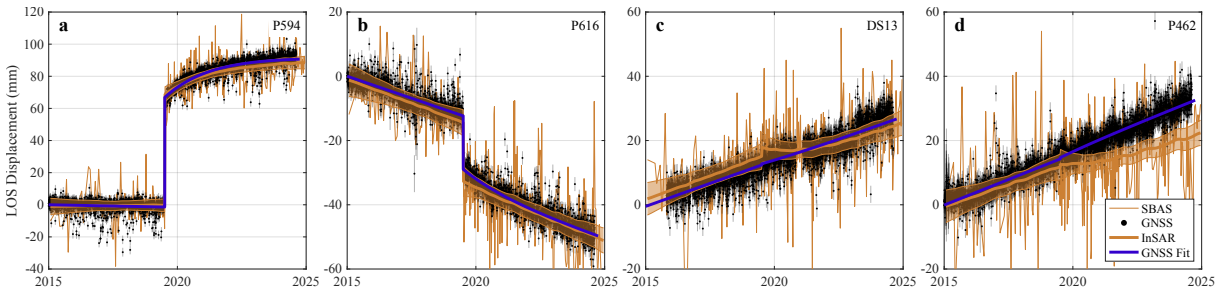


Figure 2.6: A selection of colocated InSAR and LOS-projected GNSS timeseries, for four stations which demonstrate the range of agreement between the two data sources. In each plot, a timeseries of GNSS (black, with 2σ formal uncertainties) and accompanying best-fit parameterized timeseries (blue) is compared against the posterior distribution of the InSAR displacement timeseries (thick orange line, with 2σ uncertainties). The SBAS timeseries are plotted as thin orange lines. Error bounds on the GNSS fit are approximately 0.1 mm and are omitted because they are too small to resolve. Note that the vertical scale of Figure a) is twice that of the other figures.

2.4.2 Qualitative Comparison with GNSS

To compare the posterior InSAR displacement timeseries against GNSS, we extract the InSAR displacements co-located with the GNSS stations. At each station location, we filter the InSAR displacements at each timestep with a Gaussian spatial kernel with a standard deviation of 500 m and a window size of 61 by 61 pixels, or approximately 2.4 km (E-W) by 2.7 km (N-S). We renormalize the kernel to account for any masked pixels within the kernel window.

Stations near the earthquake epicenter have large co-seismic displacements and well-defined post-seismic trends. For example, Station P594 (Fig. 2.6a) is located 24 km from the M7.1 earthquake epicenter (see Figure 1) and shows a co-seismic step of over 60 mm, with excellent agreement between InSAR (orange) and GNSS (blue). Station P616 (Fig. 2.6b), located 47 km away, has a smaller co-seismic step and little discernible post-seismic signal. Note that the sign of the coseismic step depends on the relative position of the GNSS station with respect to the fault and that the sign and amplitude of the velocity trend varies across the scene.

Station DS13 (Fig. 2.6c), further from the epicenter at 93 km, shows no discernible coseismic step in the GNSS timeseries. In the InSAR, however, there is both a step and a post-seismic signal in the opposite direction of the coseismic displacement, an unphysical behavior

for a strike-slip fault. This disagreement is a consequence of the inversion attempting to fit, in the presence of significant noise, a signal (e.g., a coseismic offset) which is not actually present. The trade-off between the co- and post-seismic behavior is localized to a small period of time after the earthquake, during which consistently positive tropospheric delay is observed, affecting the parameter estimation. Despite this disagreement, however, the GNSS parameter fit lies well within the 2σ uncertainty bounds of the InSAR. We suggest that this case highlights the necessity for uncertainty quantification; by including error bars, we can see that the InSAR and GNSS agree to within uncertainty. This case also indicates that poor choice of model can lead to physically unrealistic results.

Station P462 (Fig. 2.6d) in the far corner of the study region (93 km from the earthquake) shows generally poor agreement between the InSAR and GNSS, with the InSAR underestimating the velocity and overestimating the amplitude of the post-seismic displacement. Unlike DS13, it is unclear that a better parameterized model would have improved the fit. For example, selecting a velocity-only model would result in an InSAR-derived velocity smaller than the GNSS velocity by approximately 1 mm/yr. On the other hand, this discrepancy may be explained by an underestimation of the uncertainty at this location at the periphery of the study area. See discussion for details.

2.4.3 Quantitative Comparison with GNSS

We use the Bayesian posterior timeseries to perform a parameter estimation for the velocity, co-seismic displacement, and post-seismic displacement at each pixel in the InSAR and at each GNSS station. For each InSAR pixel, we find the linear least-squares best fit of the posterior mean to the parameterized equation 2.37 to get the parameter mean values. Since the InSAR uncertainties are uniform in space, we use the results of the uncertainty quantification above (Table 2.1) to assign the parameter standard deviation values. For each GNSS station, we use the formal GNSS uncertainties from NGL to draw 10,000 samples, find the linear least-squares parameter fit as above, and calculate the mean and standard deviation of the parameters

Figure 2.7: a) Spatial plot of LOS secular velocities inferred from the GNSS (colored circles) and Bayesian InSAR. The color scale is the same between the two data sources. b) Scatter plot comparison of the two velocities. The 1:1 line is shown with grey dashes. Uncertainties in both the GNSS (horizontal error bars) and InSAR (vertical error bars) are shown. Simple linear regression (red line) and uncertainties (transparent red region) are calculated with a 10,000 sample Monte Carlo approach. c-d) Map and scatter plot comparison of coseismic LOS displacement inferred from GNSS and InSAR. The inset plot in d) highlights the stations with small coseismic displacements. e) Spatial plot of post-seismic LOS displacement. The solid black line is the 2σ contour, indicating where the post-seismic signal can be reliably estimated. The dashed black circle marks 60 km away from the M7.1 mainshock epicenter. f) Scatter comparison of post-seismic displacement between GNSS and InSAR, where the simple linear regression analysis using all available stations is in red. We show an additional linear regression analysis in blue where only stations within the 60 km circle (circled dots) are included. Station TOWG is an outlier due to limited GNSS data availability after the earthquake and is omitted from the post-seismic analyses.

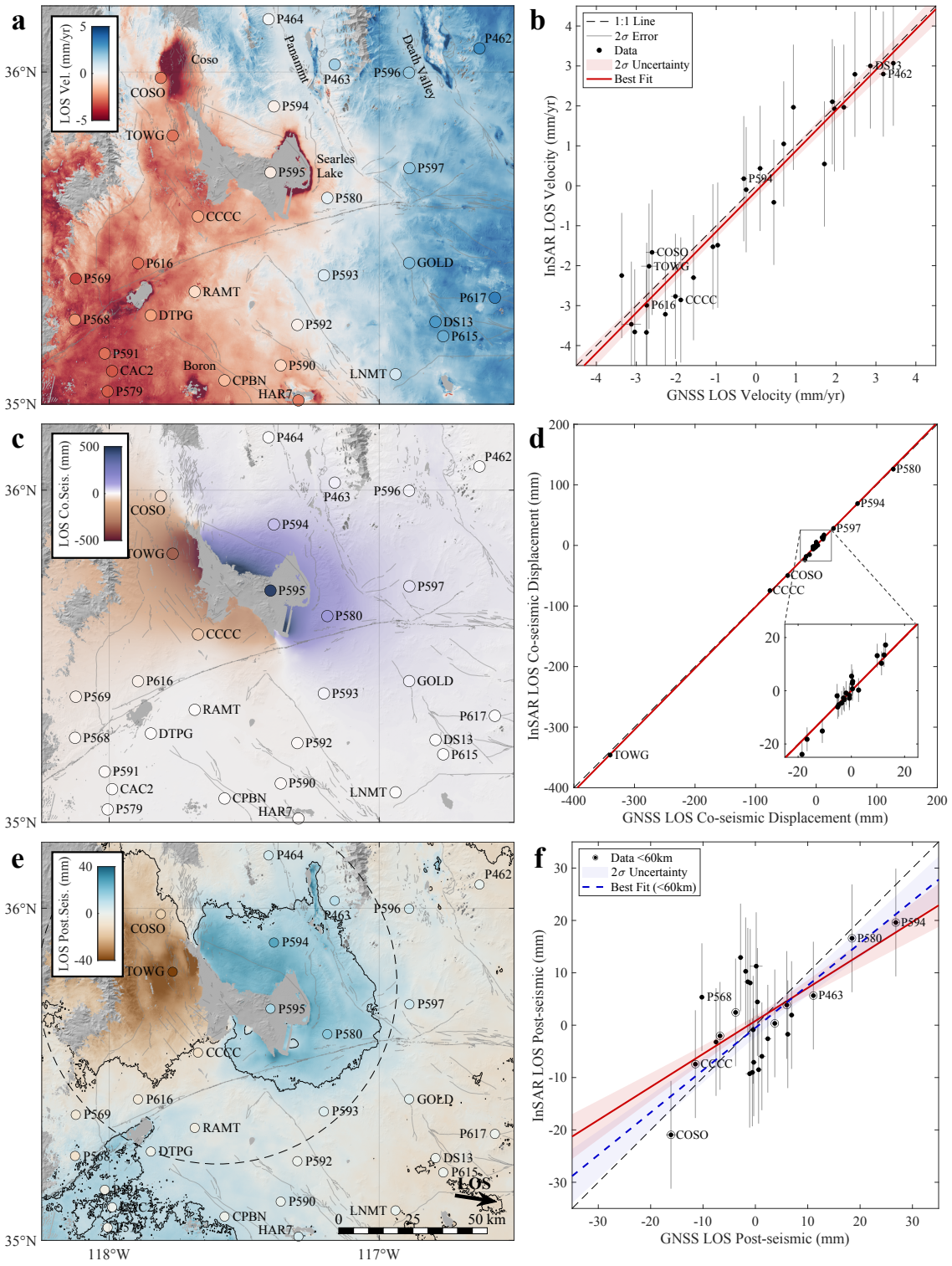

a
LOS Vel. (mm/yr)
5
0
-5
Coso
Panamint
Death Valley
Searles Lake
P464
P462
P463
P596
P594
COSO
TOWG
P595
P597
P580
CCCC
P616
P569
GOLD
P593
RAMT
P617
P568
DTPG
P592
DS13
P615
P591
CAC2
Boron
CPBN
P590
LNMT
P579
HAR7
36°N
35°N
b
1:1 Line
2σ Error
Data
2σ Uncertainty
Best Fit
InSAR LOS Velocity (mm/yr)
GNSS LOS Velocity (mm/yr)
c
LOS Co.Seis. (mm)
500
0
-500
d
InSAR LOS Co-seismic Displacement (mm)
GNSS LOS Co-seismic Displacement (mm)
e
LOS Post.Seis. (mm)
40
0
-40
LOS
0
25
50 km
118°W
117°W
f
Data <60km
2σ Uncertainty
Best Fit (<60km)
InSAR LOS Post-seismic (mm)
GNSS LOS Post-seismic (mm)

Table 2.2: Summary statistics of the scatter plot, including the R^2 values, regression slopes, and regression intercepts.

Parameter	R^2	Regression Slope	Regression y-Intercept
Velocity	0.92	1.01 ± 0.07	-0.14 ± 0.15
Co-seismic Displacement	0.999	1.01 ± 0.01	-0.01 ± 0.44
Post-seismic Displacement	0.37	0.63 ± 0.12	0.33 ± 0.43
Post-seismic Displacement (<60 km)	0.89	0.81 ± 0.13	-0.21 ± 0.75

empirically. We aggregate the comparison of parameters into the map and scatter plots in Figure 2.7, where the secular LOS velocity is shown in the two sub-figures at the top, the coseismic LOS displacement is shown in the middle, and the post-seismic displacement is shown at the bottom. Summary statistics for the scatter plots are displayed in Table 2.2.

In Figure 2.7a, we observe a large-scale velocity gradient of approximately 8 mm/yr across the scene ($\sim$150 km) in the LOS direction which arises from the predominately horizontal tectonics of the Mojave. By contrast, small-scale velocity features in this scene arise from vertical motions. Large negative velocities are observed at the Coso volcanic field and geothermal plant, where groundwater extraction is leading to rapid subsidence of the ground surface [25]; Searles Lake, where brine has been extracted industrially for decades [50]; and the borax mine in Boron, CA. Notable regional postive signals observed in the transient lakes of Panamint Valley and Death Valley may correspond to groundwater recharge causing surface uplift over the 2015-2025 decade or may have other causes related to the nature of the brine and salt deposit environment.

In the corresponding scatter plot (Fig. 2.7b), the InSAR and GNSS velocities show excellent agreement. Using the velocity parameter uncertainties (grey lines), we calculate a simple linear regression (red line) with 2σ uncertainties (transparent red region) with a 10,000 sample Monte-Carlo approach like was done for the parameter estimation above. We calculate the R^2 value of the regression, which shows high correlation ($R^2 = 0.92$) between GNSS and InSAR. The 1:1 slope and near-zero intercept of the regression line also support the agreement between the two data sources. These metrics are found in Tables 2.2 and 2.3.

We calculate one additional metric, which we call the 1:1 p-value, which is a direct

Table 2.3: Summary statistics of the parameters: the 1:1 p-value statistic and pass/fail based on a p-value of 0.05.

Parameter	1:1 p-value	Pass?
Velocity	0.37	Yes
Co-seismic Displacement	0.54	Yes
Post-seismic Displacement	0.07	Yes
Post-seismic Displacement (<60 km)	0.30	Yes

measure of the probability that the GNSS and InSAR velocity values are statistically equal. For each station, we calculate the difference between the x (GNSS) and y (InSAR) velocity values and normalize by the velocity uncertainties $(y-x)/\sqrt{\sigma_x^2+\sigma_y^2}$. The p-value comes from performing a Kolmogorov-Smirnov test on the samples, testing if they are consistent with the unit normal distribution. A p-value of less than 0.05 indicates that the x and y values are statistically different, i.e. that they do not lie along the 1:1 line (grey dashed line). By contrast, a p-value greater than or equal to 0.05 indicates that the null hypothesis (i.e. that the GNSS and InSAR are equal) cannot be rejected, or in other words that they are statistically equal at the 0.05 level. For the velocity parameter, we find a 1:1 p-value of 0.37, well above the 0.05 threshold and so we conclude that the velocities calculated from InSAR and GNSS are statistically identical. We emphasize again the importance of the uncertainty quantification in permitting this analysis. We include the 1:1 p-values as well as the 0.05 pass/fail result in the table of scatter plot summary statistics (Table 2.2).

The co-seismic displacement signal (Fig. 2.7c) is the largest signal present in the data. Near the earthquake epicenter (in regions which remain unmasked), the LOS ground surface displacement during the earthquake sequence approaches 500 mm, while displacements exceed 100 mm in a nearly 60 km diameter region. Unfortunately, the high fringe rate in the highest displacement regions directly adjacent to the fault caused significant unwrapping errors, causing the SBAS residual mask to remove those pixels, including those near GNSS station P595. To confirm that the SBAS residual mask is needed to prevent large errors, we performed a test without it, which showed an InSAR/GNSS co-seismic discrepancy at station P595 of 78 mm

(approx. 3 fringes), with the InSAR underestimating. We caution, however, that the SBAS residual mask method is not guaranteed to flag all unwrapping errors, notably those which are consistent between interferograms. With the exclusion of station P595 ensured by the SBAS residual mask, the InSAR/GNSS comparison scatter plot (Fig. 2.7d) shows remarkable agreement, with an R^2 value of 0.999 and a regression slope of 1.01 ± 0.01. The 1:1 p-value is 0.53, indicating that the two measurements are equal to well within uncertainty.

We show the map view and scatter plot of the post-seismic displacement parameter in Figures 2.7e-f. The post-seismic displacement at a given time t can be estimated by multiplying the parameter value by the value of the exponential post-seismic function (eq. 2.32) at that time, or equivalently the parameter value may be thought of as the limit of the post-seismic displacement as $t \to \infty$. The post-seismic displacement observed by the InSAR shows a similar spatial pattern to the co-seismic displacements. However, there is significant noise present, as the signal is small (maximum of $\pm$40 mm) relative to the InSAR parameter uncertainty (5.2 mm, see Table 2.1). We show the contour (solid black line) within which the absolute value of the post-seismic displacement exceeds 2σ. The contour extends approximately 60 km from M7.1 earthquake epicenter (dashed circle).

We exclude station TOWG from the post-seismic analysis because our temporal model for post-seismic displacement is not applicable for this station, causing it to appear as an outlier. The GNSS data at station TOWG terminates on December 4$^{\text{th}}$, 2019, 153 days after the earthquake. As a consequence, the rapid displacement in the hours and days after the earthquake, caused by afterslip on the fault surfaces [23], is weighted more heavily than for the other GNSS stations. The other stations have nearly 5 years of post-earthquake data, which weights their post-seismic fit towards the longer term viscous relaxation and poroelastic processes.

We show two scatter analyses for the post-seismic displacement (Fig. 2.7f). The first analysis (red) uses all stations within the study region and shows much poorer agreement between the InSAR and GNSS than was the case for the velocity and co-seismic cases. Despite the poor agreement, we see in Table 2.2 that the 1:1 p-value is 0.07, meaning that we cannot reject the

null hypothesis of statistical equality between the InSAR and GNSS for the post-seismic case.

Despite the statistical equality, we are still committing errors of comission. Beyond the 60 km circle, a parameterization which includes a post-seismic function is simply not a good model for the displacement timeseries. As we see in Figure 2.6c, overfitting of tropospheric noise causes a poor estimate of the surface displacement dynamics. To test this hypothesis, we perform a second scatter analysis (blue) using only the 9 stations (excluding TOWG) within the 60 km circle (stations are circled in the scatter plot). This analysis demonstrates much improved statistics, including a much larger p-value of 0.30, indicating that our approach including the post-seismic function is sound in the regions where the noise level is clearly exceeded.

2.5 Discussion

In this section, we discuss the context and interpretation of our findings. In Section 2.5.1, we discuss an extension of our theoretical framework. In Sections 2.5.2 and 2.5.3, we detail and present justifications for assumptions made as part of the theoretical framework and as part of the specific methods presented here. In Section 2.5.4, we discuss how our Bayesian framework may be interpreted in the context of regularized inversion. In Section 2.5.5, we defend our lack of observations of seasonality in the displacement timeseries of the Mojave. In Section 2.5.6, we conclude by using our method to investigate the spatial and temporal characteristics of the troposphere in the Mojave.

2.5.1 Extension for Multiple Subsets

Our method as presented above is only valid for the case where are SAR acquisitions are part of a single subset – that is, all SAR acquisitions have a relative spatial baseline of less than approximately 400 m. This restriction is because, for conceptual clarity, we chose to perform a two-step inversion where we solve for the SBAS timeseries $\mathbf{z}$ first, then leverage that to solve for the displacement timeseries $\mathbf{x}$. To do so, we assumed that equation 2.6 is a solvable system of equations without regularization, i.e. that the SBAS matrix $\mathbf{M}$ is full-rank. However, this

intermediate step is not required to solve for the displacement directly, as we will show now.

We begin by combining the two equations below:

$$\mathbf{y} = \mathbf{M}\mathbf{z} + \eta \tag{2.42}$$

$$\text{and } \mathbf{z} = \mathbf{G}\mathbf{x} + \xi \tag{2.43}$$

$$\rightarrow \mathbf{y} = \mathbf{M}(\mathbf{G}\mathbf{x} + \xi) + \eta \tag{2.44}$$

$$= \mathbf{M}\mathbf{G}\mathbf{x} + \mathbf{M}\xi + \eta \tag{2.45}$$

We collect the errors into a combined error term γ, then express the likelihood function. Below, $\mathbf{R}_\xi$ and $\mathbf{R}_\eta$ are the error covariance matrices discussed in Section II.

$$\gamma := \mathbf{M}\xi + \eta \tag{2.46}$$

$$\gamma \sim \mathcal{N}\left(\mathbf{0}, \mathbf{M}\mathbf{R}_\xi\mathbf{M}^T + \mathbf{R}_\eta\right) \tag{2.47}$$

$$\sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_\gamma\right) \tag{2.48}$$

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}\left|\left|\mathbf{R}_\gamma^{-\frac{1}{2}}\left(\mathbf{y} - \mathbf{M}\mathbf{G}\mathbf{x}\right)\right|\right|^2\right) \tag{2.49}$$

We combine the likelihood with the prior on $\mathbf{x}$, which remains unchanged. The posterior solution (i.e. the exact expression for $p(\mathbf{x}|\mathbf{y})$) is not presented here but may be derived analogously to the solution in the text (Section II-E).

$$\begin{aligned} p(\mathbf{x}|\mathbf{y}) &\propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \\ &\propto \exp\left(-\frac{1}{2}\left(\left|\left|\mathbf{R}_\gamma^{-\frac{1}{2}}\left(\mathbf{y} - \mathbf{M}\mathbf{G}\mathbf{x}\right)\right|\right|^2 + \left|\left|\mathbf{B}_x^{-\frac{1}{2}}(\mathbf{x} - \mu_x)\right|\right|^2\right)\right) \end{aligned}$$

Although we consider this result to be arrived at in a less intuitive manner than the approach presented in the text, it is equally valid. Importantly, it is no longer a requirement that $\mathbf{M}$ be full-rank, because the system of equations is now conditioned by the prior. Therefore, this

framework may be equally used under multiple small-baseline subset conditions, such as with older SAR constellations.

2.5.2 Assumption of Normality and Linearity

In this framework, we have assumed that every probability distribution can be described with a multi-variate normal PDF. This choice is greatly enabling when working with the sheer volumes of data associated with InSAR work. Assuming normal probability distributions gives the following advantages:

1. We may exclusively use linear matrix solvers. Non-Gaussian probability distributions would require non-linear solvers or other computationally expensive techniques, resulting in orders of magnitude greater computational resources per pixel.

2. All probability distributions, including posterior solutions, have analytic expressions.

3. We preserve all information up to the second order moment of the probability distributions.

4. Our posterior distribution will be the Best Linear Unbiased Estimator [35] of the true probability distribution, given our priors.

As part of the normality assumption, we require our parameterization equation to be a linear function of the parameters. This is often an acceptable restriction, as for the most common use cases (for example, velocity, seasonal, and co- and post-seismic functions) the temporal behavior is rigidly defined. This would not have been the case if we had included the post-seismic timescale estimation in the inversion, if for example we thought that it varied spatially. Instead, we estimated the timescale using GNSS data and assumed it was spatially constant.

Furthermore, the stochastic component β (Section II-D) permits flexibility in the recovery of the true output timeseries, the extent of which is determined by the chosen prior residual model. In our case, the sub-millimeter amplitude of β means that very little non-parametric displacement will be inferred by the inversion.

Aside from enforcing linearity, the major limitation of normality is that we cannot fully represent certain distributions. The KS-test in Section 2.3.4 demonstrated that even with seasonal correction, the Gaussian error model for tropospheric delay was slightly under-estimating the prevalence of days with high-amplitude tropospheric delay. A more accurate PDF would capture the heavier tail. In this case, however, we judge that the non-Gaussianity is too small to significantly affect the results. Similarly, while a more accurate interferometric error model might better reflect the expected error statistics of a multi-looked pixel, we expect that in this case the contribution of the non-Gaussian component is small relative to the other sources of error.

2.5.3 Assumption of Spatial Uniformity

Additionally, by assuming that our error and prior models are spatially uniform, we attain a large computational speed-up by vectorizing the matrix inversion. There are four separate instances in which we assume uniformity: the interferometric error model, the atmospheric error model, the prior model for the parameterized component of displacement, and the prior model for the stochastic component of displacement. In the following paragraphs, we briefly justify the assumption for each case.

Interferometric error, as measured by the SBAS residual, is spatially concentrated in regions of large phase gradients and surface electrical changes. However, the SBAS residual mask omits the regions where these effects are strongest. The remaining interferometric error is small and more evenly distributed. Residual spatial patterns in individual interferograms indicate that much of the interferometric error at this scale aligns with radar burst boundaries.

The spatial uniformity of atmospheric error, in particular the turbulent component of tropospheric delay, is largely controlled by the spatial referencing scheme. Since the turbulent troposphere follows a spatial power law [20], nearby pixels experience smaller relative tropospheric variance than more distant pixels. Referencing interferograms to a single GNSS station, as is commonly done, results in highly non-uniform tropospheric noise pattern with a

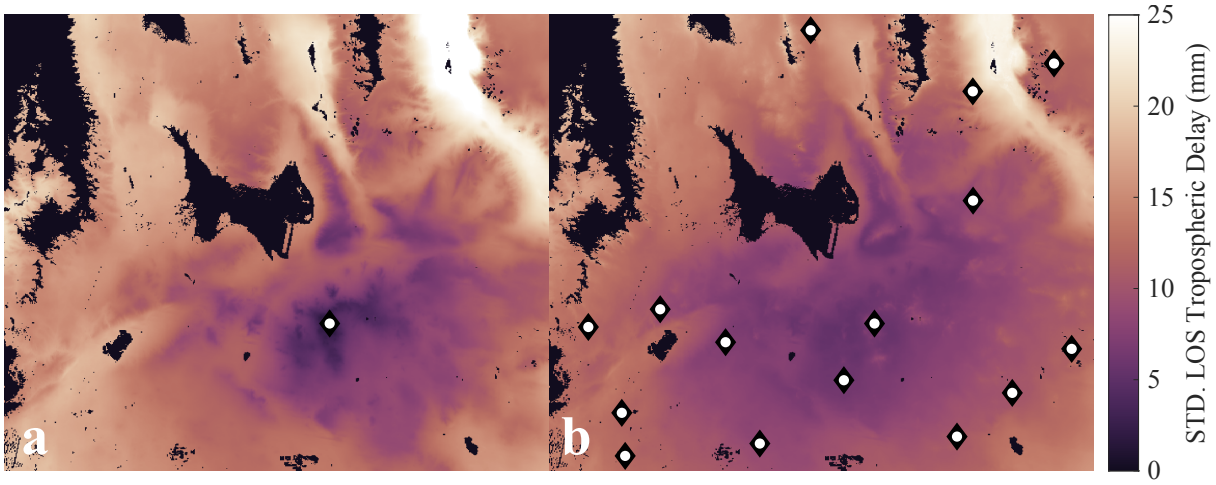


Figure 2.8: Spatial uniformity of the tropospheric delay noise depends on the spatial referencing scheme. Figure a), showing a test case with a reference point at GNSS station P593 (diamond with white center), has high variance and shows strong correlations with distance and topography, but could be useful for a localized study. Figure b), the referencing scheme used in this paper, with 15 reference stations, is smoother and has lower amplitude and variance, making it more suitable for a large scale study.

predominantly radial structure centered on the reference station (Fig. 2.8a). To minimize this effect, we reference to the mean of multiple GNSS station locations spanning the study area. The resulting tropospheric noise is more spatially homogeneous in the central portion of the scene (Fig. 2.8b). The LOS tropospheric delay σ has a standard deviation of 4.3 mm in the single station case (Fig. 2.8a) and 3.4 mm in the multi-station case (Fig. 2.8b). Similarly, using a single station reference amplifies the noise at larger distances; the single station case has a maximum of 29 mm LOS σ while the multi-station case has a maximum of 25 mm.

We chose the prior for the parametric fit (Section III-E) to be weakly informative, meaning that we can expect the data (i.e., the SBAS timeseries) to entirely inform the posterior result. The reasons for this choice are twofold: the temporal density of InSAR is sufficient to estimate the parameterized signals without prior information; and the spatial density of the GNSS is insufficient, relative to the spatial variability of the earthquake signal, to correctly interpolate a strongly informative prior onto the InSAR grid without large deviations near the fault zone. For the velocity parameter, although GNSS are spatially dense enough to estimate a strongly

informative prior, we still use a weakly informative prior to allow for InSAR/GNSS comparison. By choosing this weakly informative prior, we ensure spatial uniformity at no extra cost; the prior is uninformative, in other words, so it might as well be equally uninformative everywhere.

The stochastic component of the displacement prior (Section III-F) is inferred from the residual to the parameter fit of the GNSS timeseries. The estimated scale of the stochastic component is small compared to the parameterized component and has no obvious spatial structure. In most of the surrounding desert environment, without other information, we assume that this component is spatially constant. Certain locations undoubtedly buck the trend: Searles Lake and the other Mojave lakebeds experience poroelastic signals on hydrological timescales. However, since such signals are better studied with a more localized approach, we ignore their contribution in this study and focus on large scale deformation.

2.5.4 Relating the Prior to Regularization

In Section 2.2.4, we introduced a prior for the displacement timeseries which had both a parameterized and a stochastic component. It is worth discussing the meaning of the stochastic term β and its relationship to regularization in a standard inverse problem. Although the stochastic component of displacement is small in the Mojave, this is not the case for all geophysical settings. For example, in a region with significant hydrological surface deformation, there may be large deformation on inter-annual (e.g., from multi-year drought cycles) or sub-seasonal timescales (e.g., from year-to-year irregularities in seasonal precipitation) not captured by a seasonal fit. Even if this residual deformation does not have a specific functional form, it is still observed by InSAR. The stochastic term β allows the user to include this deformation in the output timeseries.

Conversely, the β term may be thought of as a regularization term, allowing us to regularize the noisy SBAS timeseries, while the parameterized term $\mathbf{A}\theta$ prevents the sacrifice of known features in the timeseries. As a concrete example, performing a smooth regularization over a known impulse deformation signal from an earthquake will result in obvious and unnecessary

smearing of the earthquake signal in time. The joint approach presented here is akin to removing the earthquake step, performing a smoothing regularization on the residual, then replacing the step function, thus preserving the sharpness of the earthquake deformation.

The relative amplitudes of β and the atmospheric error model ξ will weight the resulting timeseries to be more or less sensitive to impulsive changes in the SBAS timeseries. The difficulty, as in traditional regularized inversions, lies in selecting an appropriate weighting of regularization. Our method is conceptually and practically different from techniques such as Occam's inversion [17] which use the data to estimate the regularization parameter. In our method, the prior term β as well as the error model ξ are physically informed by the GNSS timeseries.

2.5.5 Seasonal Displacements of the Mojave

In our parameterization of the displacement timeseries for both GNSS and InSAR, we notably omit a seasonal term. Although seasonal signals are present in GNSS timeseries, with amplitudes ranging from 0-2 mm, the seasonal phase shows little consistency and varies strongly between stations, leading us to believe that the observed seasonality is not due to real wide-scale ground surface displacement. We hypothesize that other sources of seasonality in GNSS timeseries, such as multi-path or monument effects [75], or draconitic signals [4], are contaminating the GNSS displacements.

In our study area, GNSS stations P591 and P812 near California City, CA in the Mojave are part of the Plate Boundary Observatory Multi-Monument Experiment [13]. The stations are located approximately ten meters from each other, making them ideal for studying GNSS noise in the absence of relative deformation. We take the difference between the GNSS timeseries projected into LOS over the time period 2019 to 2024, during which there are no logged antenna changes that could cause offsets. We fit a linear trend and a seasonal sinusoid to the differenced timeseries, resulting in a seasonal amplitude of 0.45 mm. For comparison, the median seasonal amplitude of all GNSS timeseries in the study area is 0.43 mm; in other words, half of GNSS

stations have seasonal motions smaller than what might be considered to be the noise floor. Although P591 and P812 are of different monument types (deep-drilled braced and short-drilled braced respectively), these results are consistent with the possibility that essentially none of the GNSS timeseries seasonality in this region reflects actual ground surface motion. At best, such motion would be restricted to a localized region. Consequentially, we assume in our inversion that there is no large-scale seasonal motion.

The InSAR SBAS timeseries shows large seasonal amplitudes relative to the GNSS. As a test, we added a seasonal sinusoid term to the parameterization function (eq. 2.37) and fit the SBAS timeseries at the pixels co-located with GNSS stations. The amplitude of the seasonal component ranged from 0.59 mm to 5.0 mm (median 2.1 mm), far higher than the GNSS values. Unlike the GNSS, however, the histogram of the seasonal phase is sharply peaked around late May and late November. Furthermore, there is no correlation between the seasonal phase of the InSAR and GNSS ($R^2 = 0.009$) as might be expected if they both reflected surface deformation. From these results, we infer that there is a large, seasonal noise source in the InSAR, likely arising in the troposphere. A similar seasonality of tropospheric noise was observed in Oman by Zebker et al. (2023) [84]. We note that since each interferogram is referenced, seasonality implies a seasonally changing spatial gradient across the scene.

Unlike high-frequency noise due to short-term weather, a seasonal noise source is difficult to removal using standard methods of InSAR tropospheric noise mitigation. Common-scene stacking [73] and regularization-based approaches [64, 56] are typically applied in order to smooth InSAR timeseries to sub-annual timescales, under the implicit assumption that any longer-period signal is deformation. Likewise, a parameterized fit that includes a seasonal component will see a large signal not reflective of surface motion. Such methods will incorrectly label seasonal tropospheric noise as surface motion [84]. This difficulty highlights one of the strengths of our proposed method, in particular the use of GNSS to inform the prior model for the displacement (Section 2.3.5). By doing so, we recover much smaller seasonal signals. Drawing samples from our posterior, we estimate small seasonal amplitudes (median 0.20 mm) with large

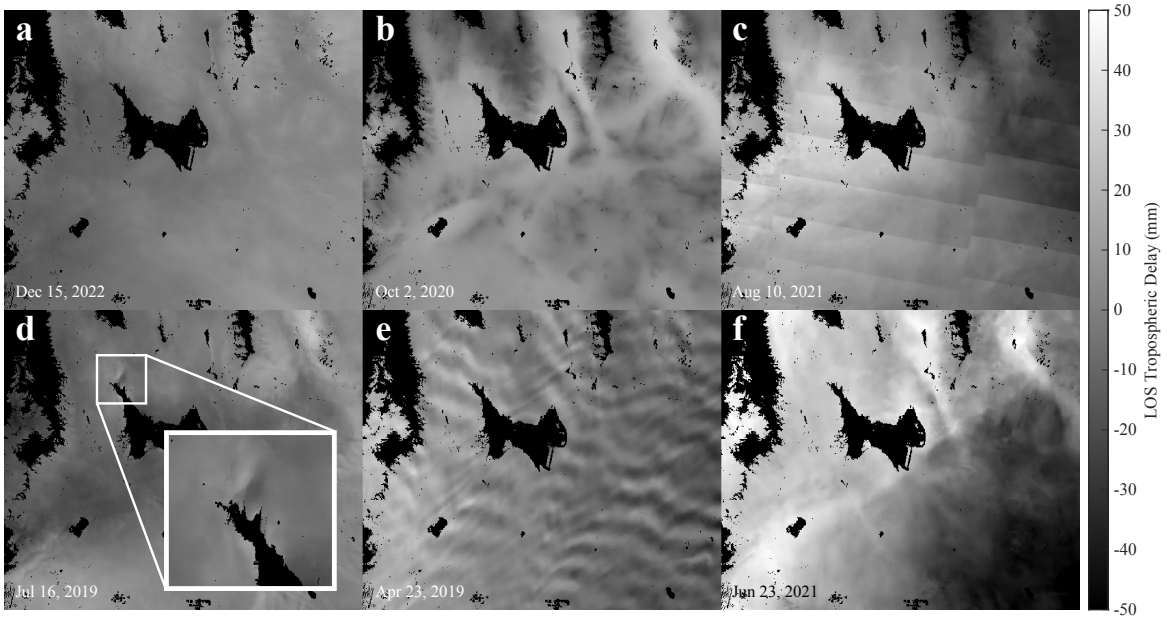


Figure 2.9: Examples of the spatial distribution of tropospheric delay at six dates, inferred by subtracting the Bayesian displacement timeseries from the SBAS timeseries. a) A winter date with mild troposphere. b) A date with strongly stratified troposphere showing high correlation with elevation. c) A date with large burst errors. d) The date immediately following the earthquake, where a small-amplitude localized signal from the rapid post-seismic deformation is present. e) A large-amplitude troposphere with clear cloud-like banding, showing strong short-wavelength, anisotropic behavior. f) A high-amplitude turbulent summer troposphere, where the strong northwest/southeast gradient may indicate the boundary between two air masses.

uncertainties (IQR 0.13 mm-0.28 mm).

2.5.6 Tropospheric Dynamics of the Mojave

By producing a reduced-noise timeseries of surface displacement, we can use the noisy SBAS timeseries to recover the spatio-temporal characteristics of tropospheric noise in the Mojave. Similar methods have been proposed before. Hanssen et al. (1999) [33] used individual interferograms to infer water vapor content at high resolution. Zebker et al. (2023) [84] use common-scene stacking over multiple subsets to estimate displacement signal linearity and corresponding atmospheric noise. Our method of troposphere recovery is very simple: we subtract the smoothed displacement timeseries from the SBAS timeseries. This procedure may

be applied broadly to any method where a smoothed timeseries of deformation is calculated, such as common-scene stacking; however, we caution users to ensure (via a GNSS comparison) that any long-period noises, such as seasonal troposphere, are not accidentally absorbed into the deformation signal, or else a component of the tropospheric signal will be omitted.

The recovered troposphere shows a diversity of clear spatial patterns. In Figure 2.9a, we see a day with smooth, low amplitude troposphere. The amplitude of delay is often smaller during the winter, but not as a rule: the atmosphere can exhibit almost any state at any time of year. Figure 2.9b shows a day with no obvious tropospheric turbulence, but with a strong stratified component highly correlated with elevation. Figure 2.9c demonstrates large burst stitching errors. Although not uncommon in individual interferograms, this phenomena appears to have been present for each interferogram associated with this date and is therefore most likely to be present in the corresponding SLC. There was only one occurrence of this magnitude during the study period.

Figure 2.9d is the troposphere snapshot immediately following the earthquake. We highlight a region near the north edge of the fault zone which shows evidence of rapid post-seismic deformation due to fault slip [23]. Since our parameterization accounts for both the co-seismic deformation and the long-period, viscoelastic post-seismic deformation, this signal is a residual which happened too quickly to be captured by our method. The amplitude of noise outside of this localized region (which is clearly atmospheric in nature) demonstrates why such rapid signals should not be included in InSAR parameter estimations, including in our method. The tropospheric noise in Figure 2.9d. has a large enough amplitude that, given the timeseries of two pixels without knowledge of their spatial location, it would be impossible to tell which had true rapid post-seismic motion. In other words, there is a choice between capturing the rapid post-seismic signal and falsely including a large amount of tropospheric noise in the timeseries (co-mission), or rejecting both the signal and noise (omission). Such a signal is best studied on a localized scale where relative tropospheric noise is sufficiently small.

Finally, there are days in which the turbulent component of tropospheric noise has very

large amplitudes. Figure 2.9e shows a anisotropic troposphere with a characteristic cloud banding pattern. The wavelength of the bands are far below the resolution of tropospheric correction products such as GACOS, even if the products perfectly captured the state of the atmosphere. Figure 2.9f shows an extremely high amplitude day, where the tropospheric delay has a range of nearly 100 mm. The strong difference between the northwest and southeast corners of the scene possibly indicate that there was a moving air front present in the scene during the satellite fly-over. Such extreme states of the troposphere are common during the summer and have been observed before in other regions [84].

The summer/winter dichotomy is shown in Figure 2.10, in which we take the mean and standard deviation of the derived tropospheric delay maps from June to August and from December to February. The seasonal difference in the mean value of the tropospheric delay shows that the mean shift shows a predominantly east/west gradient with some elevation dependence. There is also a large difference between the summer and winter standard deviation, with the summer showing a much larger variability by about 5 mm. Death Valley in the upper right corner of the study area shows extreme differences in the seasons relative to the southern Mojave, showing that the high relief of the Panamint and Amargosa Ranges isolates the weather in the valley [62].

Finally, we briefly investigate the effect of the vertical stratification of the atmosphere. We regress the tropospheric delay against elevation as well as latitude and longitude, to account for the effect of large turbulent gradients across the scene. To avoid the extreme tropospheric dynamics of Death Valley and the Sierras, we restrict our regression to pixels south of the Panamint Mountains, east of the Sierras, and with a relatively high elevation slope (100 m per km or greater). The delay trend in elevation has large variability around zero (-0.3 ± 13.8 mm/km), meaning that the contribution of the stratified atmosphere flips sign (see Figure S1). The difference between summer and winter is small, with summer trends (-2.4 ± 12.8 mm/km) showing slightly higher variance and more negative slope than winter trends (1.7 ± 10.7 mm/km). Conventional InSAR processing often removes the stratified component of tropospheric delay.

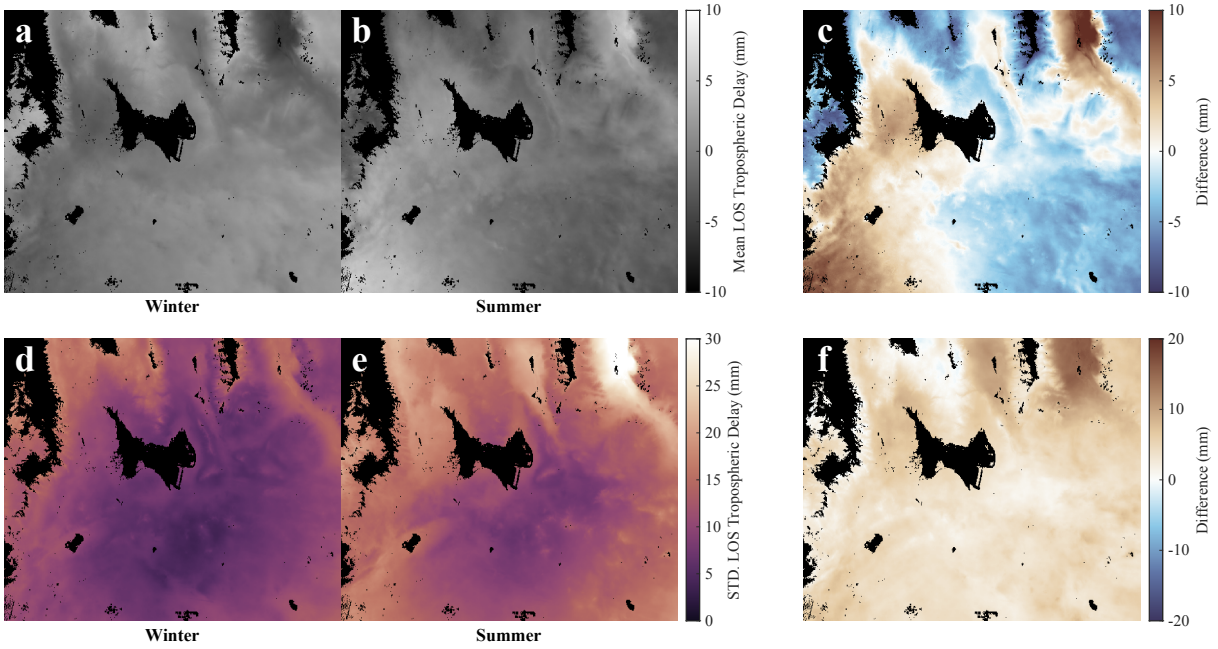


Figure 2.10: a-b) The mean of the tropospheric delays in winter (December-February) and in summer (June-August). c) The summer minus winter difference between b) and a). The difference in the mean value shows an east/west gradient across the scene, corresponding to the seasonal shift in the mean of the tropospheric delay. d-e) The standard deviation of the tropospheric delay. f) The amplitude of tropospheric noise is almost uniformly higher in the summer, by around 5 mm. Death Valley in the upper right of the study area stands out as having extreme summer/winter differences in both the mean and standard deviation.

By preserving it in our processing, we are better able to isolate and study it in the presence of the turbulent noise.

2.6 Conclusion

InSAR timeseries estimation in geophysical applications would greatly benefit from the systematic use of uncertainty quantification. In this paper, we have proposed a theoretical framework for Bayesian uncertainty analysis of ground surface deformation timeseries for InSAR. Our theory incorporates interferometric and atmospheric noise, functional fitting of the components of displacement timeseries with simple parameterizations, and analysis of the non-parametric component of displacement. We used the residual from the SBAS inversion to estimate the interferometric noise, the difference between GNSS and SBAS timeseries to estimate a seasonal error model for the troposphere, and the residual from the GNSS parameter fit to find the amplitude of stochastic displacements. We propagated uncertainties from the SBAS inversion through to the estimation of the displacement timeseries.

We compared the estimated InSAR timeseries with the GNSS timeseries. Decomposing the signal into velocity, co-seismic, and post-seismic components, we found that the two deformation signals are statistically indistinguishable. In our discussion, we showed how, given the high levels of tropospheric noise we observe in the Mojave, real seasonal displacements as inferred from GNSS are unobservable by InSAR. We also demonstrated how to infer the tropospheric dynamics of the region and concluded that there is a large seasonal change in both the mean and variance of tropospheric delay in the Mojave. The troposphere product presented has the potential to be useful in the future for high-resolution weather characterization.

2.7 Acknowledgements

For Chapter 2, InSAR data was processed by the Advanced Rapid Imaging and Analysis (ARIA) Center at NASA Jet Propulsion Laboratory and California Institute of Technology.

InSAR data was made available on ASF's Vertex website. Contains modified Copernicus Sentinel-1 data 2014-2025, processed by the European Space Agency. This work was funded by the National Aeronautics and Space Administration under NASA Grants NNX16AR07G (NISAR Science Team Extension) and 80NSSC24K0670 (Earth Surface and Interior). The authors would like to thank the European Space Agency, GPS networks with data processed by UNR's Nevada Geodetic Laboratory, and the U.S. Geological Survey for their open data policies regarding the data used in this work.

Chapter 2, in full, is currently being prepared for publication and is coauthored with Adrian A. Borsa. I was the primary investigator and author of this research material.

Chapter 3

Theory of Uncertainty Quantification for InSAR Timeseries

3.1 Spatio-Temporal Covariance

In InSAR timeseries, the spatial correlations of tropospheric noise become blended in time by temporal processing algorithms such as timeseries smoothing. This spatio-temporal blending presents a particular challenge for InSAR uncertainty quantification, given the quantity of data involved. In this chapter, we will present a theoretical framework for uncertainty quantification of InSAR timeseries. We will then translate commonly-used InSAR processing algorithms into this framework. This framework, based on the Kronecker product, is both predictive and practical for common use.

3.1.1 Mathematical Formalism

We begin with a data matrix $\mathbf{X}$ of size $N_{time} \times N_{space}$, which may represent an interferogram stack or InSAR timeseries. An element X_{ji} of this matrix is a random variable at timestep j and position i. Each row represents the spatial field, or image, at a given time and each column represents the timeseries at a given spatial position.

For a single vector $\mathbf{x}$ in isolation from any other vector, the covariance matrix has a

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \\ x_{21} & x_{22} & x_{23} & \dots \\ x_{31} & x_{32} & x_{33} & \end{pmatrix}$$

Image at time 1

Timeseries at pixel 1

simple structure.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix} \qquad \text{cov}(\mathbf{x}) = \begin{pmatrix} \text{var}(x_1) & \text{cov}(x_1,x_2) & \text{cov}(x_1,x_3) & \\ \text{var}(x_1,x_2) & \text{var}(x_2) & \text{cov}(x_2,x_3) & \dots \\ \text{cov}(x_1,x_3) & \text{cov}(x_2,x_3) & \text{var}(x_3) & \\ & \vdots & & \ddots \end{pmatrix}$$

$$= E[(\mathbf{x} - \text{mean}(\mathbf{x}))(\mathbf{x} - \text{mean}(\mathbf{x}))^{\text{T}}]$$

The covariance of a data matrix $\mathbf{X}$ is more involved. We need a way of representing the covariance between each pair of data points in different rows and columns. We accomplish this using a *vectorization operation* $\text{vec}(\mathbf{X})$ on the matrix $\mathbf{X}$ to flatten it into a vector. The covariance of $\text{vec}(\mathbf{X})$ then follows from the covariance of a vector as above, where we now have a covariance between each pair of vector elements of $\text{vec}(\mathbf{X})$. Expressed in terms of the matrix elements of $\mathbf{X}$, the full structure of the *data matrix covariance* is shown below.

$$\text{vec}(\mathbf{X}) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{12} \\ x_{22} \\ \vdots \end{pmatrix}$$

Image at time 1

Timeseries at pixel 1

$$\text{cov}(\text{vec}(\mathbf{X})) = \begin{pmatrix} \text{var}(x_{11}) & \text{cov}(x_{11},x_{21}) & \dots & \text{cov}(x_{11},x_{12}) & \text{cov}(x_{11},x_{22}) & \dots \\ \text{cov}(x_{11},x_{21}) & \text{var}(x_{21}) & \dots & \text{cov}(x_{21},x_{12}) & \text{cov}(x_{21},x_{22}) & \dots \\ \vdots & \vdots & & \vdots & \vdots & \\ \text{cov}(x_{11},x_{12}) & \text{cov}(x_{21},x_{12}) & \dots & \text{var}(x_{12}) & \text{cov}(x_{12},x_{22}) & \dots \\ \text{cov}(x_{11},x_{22}) & \text{cov}(x_{21},x_{22}) & \dots & \text{cov}(x_{12},x_{22}) & \text{var}(x_{22}) & \dots \\ \vdots & \vdots & & \vdots & \vdots & \end{pmatrix}$$

Timeseries at pixel 1

Image at time1

For a vector with N elements, the covariance matrix has N^2 elements. The full covariance

matrix of vec($\mathbf{X}$) will therefore have $N_{data}^2 = (N_{space}N_{time})^2$ elements. For even small datasets, the numbers involved here are prohibitive; a 100-interferogram stack with a mere 100 pixels per horizontal dimension would have 10^{12} elements in its covariance matrix, requiring 4 terabytes of storage in single-precision.

It is important to be able to work with the full covariance matrix, or a suitable approximate of it. Covariances in either the spatial domain or the temporal domain are not particularly challenging to work with by themselves. Even the storage of a different spatial domain covariance matrix for each timestep, while challenging, is not necessarily prohibitive. The difficulty is when space and time are coupled. InSAR work starts with spatially non-trivial noise structure; turbulence in the troposphere imparts spatial covariances between all pairs of pixels in the spatial domain. Meanwhile, commonly used timeseries smoothing techniques to reduce this noise smear these covariances across the temporal domain, causing spatio-temporal coupling.

3.1.2 Kronecker Product

The Kronecker product creates a large block matrix from any two input matrices and is a generalization of the tensor product for vectors to matrices.

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{pmatrix}$$

To address the challenges of spatio-temporal coupling and computational storage, we make the assumption that cov(vec($\mathbf{X}$)) can be expressed as the Kronecker product between matrix $\mathbf{C}_s$ (size $N_{space} \times N_{space}$) on the left and matrix $\mathbf{C}_t$ (size $N_{time} \times N_{time}$) on the right. When this assumption holds, the data matrix covariance will be referred to as a *Kronecker covariance*.

$$\text{cov}(\text{vec}(\mathbf{X})) = \mathbf{C}_s \otimes \mathbf{C}_t$$

This assignment has the following consequences: 1) The Kronecker covariance does not need to be stored in full, but may be represented by a pair of two small matrices $\mathbf{C}_s$ and $\mathbf{C}_t$. 2) Every spatial field will have the same spatial covariance matrix ($\mathbf{C}_s$) as every other spatial field, up to a constant factor. 3) Every timeseries will have the same temporal covariance matrix ($\mathbf{C}_t$) as every other timeseries, again up to a constant factor. 4) This representation of the covariance matrix is approximate. The spatial pattern of covariance of tropospheric noise (specifically, the exponent of the power-law in the structure function) varies on a day-to-day basis in a manner which cannot be represented in this manner.

3.1.3 Linear Transformations

A single representation of the covariance in terms of a Kronecker product may be useful; however, if some transformation of the data spreads the covariances in ways which cannot be represented with a Kronecker product structure, we lose the advantages that come with it. As we show in this section, the category of transformations which preserve Kronecker structure are *linear transformations*.

A linear transformation of a matrix $\mathbf{X}$ is multiplication by a matrix from the left or the right. Although they are among the simplest examples of transformations of matrix data, linear transformations comprise an important and highly useful class. In our experience, the majority of transformations of InSAR data either 1) are explicitly stated as linear transformations; 2) are not explicitly stated as linear transformations but can be restated as such; or 3) are well enough approximated by linear transformations that using one would make no significant difference.

Starting with our data matrix $\mathbf{X}$, applying a matrix multiplication from the left $T(\mathbf{X},\mathbf{A}) = \mathbf{AX}$ performs a linear temporal transformation equally to each pixel's timeseries. Applying a matrix multiplication from the right $S(\mathbf{X},\mathbf{B}) = \mathbf{XB}^{\mathrm{T}}$ performs a linear spatial transformation equally to each timestep's spatial field. We generalize both transformations into one below.

$$\mathbf{X} \rightarrow \mathbf{AXB}^{\mathrm{T}}$$

We perform a vectorization operation on the matrix $\mathbf{AXB}^{\mathrm{T}}$ to flatten it into a single column vector. The *mixed Kronecker matrix-vector product* identity allows it to be expressed in terms of the vectorization of $\mathbf{X}$.

$$\mathrm{vec}(\mathbf{AXB}^{\mathrm{T}}) = (\mathbf{B} \otimes \mathbf{A})\mathrm{vec}(\mathbf{X})$$

The general rule for the transformation of covariance matrices under multiplication by the matrix $\mathbf{M}$ on the variable $\mathbf{x}$ is as follows:

$$\mathrm{cov}(\mathbf{Mx}) = \mathbf{M}\mathrm{cov}(\mathbf{x})\mathbf{M}^{\mathrm{T}}$$

We apply this transformation rule to the vectorized equation above and distribute the transpose[1] over the Kronecker product.

$$\begin{aligned}\mathrm{cov}(\mathrm{vec}(\mathbf{AXB}^{\mathrm{T}})) &= (\mathbf{B} \otimes \mathbf{A})\mathrm{cov}(\mathrm{vec}(\mathbf{X}))(\mathbf{B} \otimes \mathbf{A})^{\mathrm{T}} \\ &= (\mathbf{B} \otimes \mathbf{A})\mathrm{cov}(\mathrm{vec}(\mathbf{X}))(\mathbf{B}^{\mathrm{T}} \otimes \mathbf{A}^{\mathrm{T}})\end{aligned}$$

We make use of the assumption that $\mathrm{cov}(\mathrm{vec}(\mathbf{X}))$ can be expressed as the Kronecker product between left matrix $\mathbf{C}_s$ and right matrix $\mathbf{C}_t$. We see now that this assumption is necessary to twice apply the *mixed-product property* of the Kronecker product.[2]

$$\begin{aligned}\mathrm{cov}(\mathrm{vec}(\mathbf{X})) &= \mathbf{C}_s \otimes \mathbf{C}_t \\ \mathrm{cov}(\mathrm{vec}(\mathbf{AXB}^{\mathrm{T}})) &= (\mathbf{B} \otimes \mathbf{A})(\mathbf{C}_s \otimes \mathbf{C}_t)(\mathbf{B}^{\mathrm{T}} \otimes \mathbf{A}^{\mathrm{T}}) \\ &= (\mathbf{BC}_s \otimes \mathbf{AC}_t)(\mathbf{B}^{\mathrm{T}} \otimes \mathbf{A}^{\mathrm{T}}) \\ &= \mathbf{BC}_s\mathbf{B}^{\mathrm{T}} \otimes \mathbf{AC}_t\mathbf{A}^{\mathrm{T}}\end{aligned}$$

[1] The transpose operation is distributive over the Kronecker product: $(\mathbf{A} \otimes \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} \otimes \mathbf{B}^{\mathrm{T}}$.

[2] The mixed-product property of the Kronecker product is: $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$.

Therefore the covariance of the transformed variable $\mathrm{cov}(\mathrm{vec}(\mathbf{AXB}^{\mathrm{T}}))$ preserves the same structure as $\mathrm{cov}(\mathrm{vec}(\mathbf{X}))$: a Kronecker product between matrices. Transformations of this form consequently maintain all the associated advantages regarding covariance storage and spatio-temporal coupling.

We restate this result in TS notation to demonstrate that the temporal and spatial transformation operations have the dual effect of data matrix transformation and covariance transformation. The left matrix $\mathbf{C}_s$, which represents the spatial covariance, transforms under the spatial transformation matrix $\mathbf{B}$, while the temporal covariance matrix $\mathbf{C}_t$ transforms under the temporal transformation matrix $\mathbf{A}$. In this sense, the temporal and spatial aspects of the data may be treated as separate.

$$T(\mathbf{X},\mathbf{A}) = \mathbf{AX} \qquad \mathrm{cov}(\mathrm{vec}(T(\mathbf{X},\mathbf{A}))) = \mathbf{C}_s \otimes \mathbf{AC}_t\mathbf{A}^{\mathrm{T}}$$

$$S(\mathbf{X},\mathbf{B}) = \mathbf{XB}^{\mathrm{T}} \qquad \mathrm{cov}(\mathrm{vec}(S(\mathbf{X},\mathbf{B}))) = \mathbf{BC}_s\mathbf{B}^{\mathrm{T}} \otimes \mathbf{C}_t$$

We wish to emphasize that the uncertainty analysis, i.e., the transformation of the covariances, is inherently implied by the stating of the transformation in TS notation. Therefore one of the main goals of the remainder of this chapter is to translate the relevant InSAR processing transformations into the correct matrix multiplication form for TS notation. With the accomplishment of this task, the uncertainty analysis follows directly from the rules above.

Associativity

Here we show that temporal (T) and spatial (S) operations are associative with each other. The practical effect is that spatial and temporal processing chains may be analyzed separately.

$$T(S(\mathbf{X},\mathbf{B}),\mathbf{A}) = \mathbf{A}(\mathbf{XB}^{\mathrm{T}}) = (\mathbf{AX})\mathbf{B}^{\mathrm{T}} = S(T(\mathbf{X},\mathbf{A}),\mathbf{B})$$

However, transformation operations are not associative with respect to two temporal

operations or two spatial operations. Therefore the analysis of the spatial and temporal processing chains must remain internally ordered.

$$T(T(\mathbf{X},\mathbf{A}_1),\mathbf{A}_2) = \mathbf{A}_2(\mathbf{A}_1\mathbf{X}) \neq \mathbf{A}_1(\mathbf{A}_2\mathbf{X}) = T(T(\mathbf{X},\mathbf{A}_2),\mathbf{A}_1)$$
$$S(S(\mathbf{X},\mathbf{B}_1),\mathbf{B}_2) = (\mathbf{X}\mathbf{B}_1^{\mathrm{T}})\mathbf{B}_2^{\mathrm{T}} \neq (\mathbf{X}\mathbf{B}_2^{\mathrm{T}})\mathbf{B}_1^{\mathrm{T}} = S(S(\mathbf{X},\mathbf{B}_2),\mathbf{B}_1)$$

3.1.4 Addition of Multiple Terms

Cases arise when it is appropriate to add multiple data matrices; for example, when considering interferometric noise in addition to tropospheric noise, or when referencing to a GNSS field which includes its own uncertainty. The data matrix covariances add, but their component spatial or temporal covariance matrices do not generally do so.

$$\begin{aligned}\mathbf{X}_3 &= \mathbf{X}_1 + \mathbf{X}_2 \\ \mathrm{cov}(\mathrm{vec}(\mathbf{X}_3)) &= \mathrm{cov}(\mathrm{vec}(\mathbf{X}_1)) + \mathrm{cov}(\mathrm{vec}(\mathbf{X}_2)) \\ &= \mathbf{C}_{s,1}\otimes\mathbf{C}_{t,1} + \mathbf{C}_{s,2}\otimes\mathbf{C}_{t,2} \\ &\neq \mathbf{C}_{s,3}\otimes\mathbf{C}_{t,3}\end{aligned}$$

The exception is when either the spatial or temporal covariance matrices are identical up to a common factor. Below, the linear combination of two terms with a common spatial covariance matrix is shown as an example.

$$\begin{aligned}(a\mathbf{C}_s)\otimes\mathbf{C}_{t,1} + (b\mathbf{C}_s)\otimes\mathbf{C}_{t,2} &= \mathbf{C}_s\otimes(a\mathbf{C}_{t,1}) + \mathbf{C}_s\otimes(b\mathbf{C}_{t,2}) \\ &= \mathbf{C}_s\otimes(a\mathbf{C}_{t,1} + b\mathbf{C}_{t,2})\end{aligned}$$

Addition in general does not preserve the single Kronecker covariance structure. However, the above results regarding linear transformations and associativity generalize trivially to Kronecker-sum matrices. As a practical matter, therefore, when encountering multiple terms,

simply store multiple spatial and temporal covariance matrices, one each for each noise term, and replicate the processing steps for each term individually.

3.1.5 Application to InSAR Processing

The interferogram-to-deformation timeseries processing chain generally involves a user-customized pipeline with a mix of spatial and temporal operations, including interferogram correction, SBAS inversion, spatial referencing, and temporal smoothing. For example, Xu et al., 2021 [79] performs the following steps after interferogram construction, unwrapping, and masking. Below, we have indicated whether each step is spatial or temporal in nature based on how it affects the covariance of the data matrix.

1. Correct interferograms for stratified tropospheric delay (spatial).
2. Reference interferograms to an interpolated, filtered field of GNSS LOS velocity (spatial).
3. Perform common-scene stacking correction (temporal).
4. Perform SBAS inversion (temporal).
5. Subtract horizontal GNSS-derived velocity field (spatial).

The uncertainty analysis follows from the InSAR processing chain, whatever its particular customization. The Kronecker covariance of the tropospheric noise can be constructed after the completion of the InSAR processing chain using information about the residual atmosphere, or it can be supplied another way. Then the individual operations must be translated into the form of a matrix multiplication and applied to the tropospheric Kronecker covariance to get the displacement Kronecker covariance. By the associativity of spatial and temporal operations, uncertainty analysis of such processing chains may be split into a spatial chain and a temporal chain. See Figure 3.1 for the workflow diagram and locations of specific translations within this work.

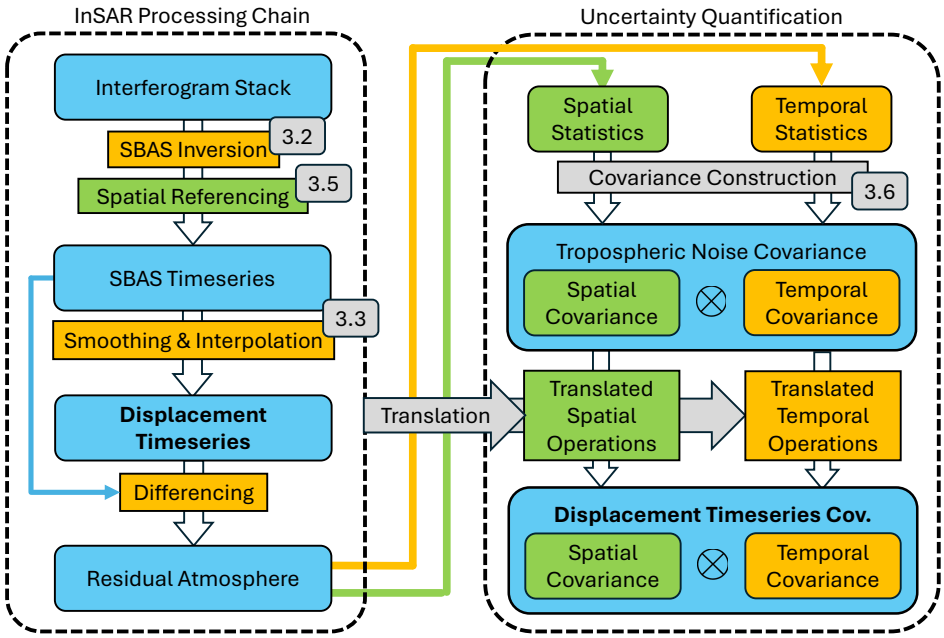


Figure 3.1: Diagram of the InSAR processing chain and uncertainty quantification workflow. Blue indicates a data matrix or associated Kronecker covariance. Orange indicates a temporal operation or object, while green indicates a spatial operation or object. The gray boxes show the section where the translation of a specific operation may be found in this work.

After completing the InSAR processing, derived quantities such as a velocity map or timeseries at a single position are often of interest. The uncertainties associated with such quantities may be found by performing a spatial or temporal transformation with the correct matrix. Examples of some such matrices for common quantities are found in the Appendix (Section 3.8.1).

3.2 The SBAS Inversion

The Small-Baseline Subset (SBAS) inversion [10] transforms a set of interferometric displacements into a timeseries of surface deformation and is the core of InSAR timeseries work. It is a temporal domain transformation which operates on a pixel-by-pixel basis; for clarity, we will consider only a single pixel timeseries. We leave the (trivial) generalization to full data matrices and TS notation until the end.

The differential phase $\delta\phi_{i,j}$ between two SAR acquisitions i and j is the following, where d_i represents the ground surface displacement at time i, a_i represents the atmospheric noise at time i, $\eta_{i,j}$ is the interferometric noise associated with this interferogram, and λ is the central wavelength of the radar signal.

$$\delta\phi_{i,j} = \frac{4\pi}{\lambda}(d_j + a_j - d_i - a_i + \eta_{i,j})$$

Crucially, the atmospheric noise term is differential, in precisely the same manner as the displacement. For now, we will unite them into one object, the *noisy displacement*; in later sections we will describe how these terms may be separated and atmospheric noise removed. In terms of the noisy displacement $x_i = d_i + a_i$, the differential noisy displacement $y_{i,j}$ is:

$$y_{i,j} = x_j - x_i + \eta_{ij}$$

Given a set of interferograms, we may construct a matrix equation, where $\mathbf{y}$ is the vector of the noisy interferogram stack, $\mathbf{M}$ is the SBAS matrix, and $\mathbf{x}$ is the vector of the *SBAS timeseries*. Following the usual definitions, the displacements x_i are relative to the position at the start time t_1. The exact form of the SBAS matrix depends on the network structure of the interferogram stack; a representative example of one such is shown below.

$$\underbrace{\begin{pmatrix} y_{1,2} \\ y_{1,3} \\ y_{2,3} \\ y_{2,4} \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}}_{\mathbf{M}} \underbrace{\begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}}_{\mathbf{x}}$$

Assuming that $\mathbf{M}$ is full-rank (i.e., all interferograms belong to the same subset), this system of equations is invertible using least-squares, i.e., by multiplication with the Moore-Penrose

pseudo-inverse $\mathbf{M}^+$, thus completing the SBAS inversion.

$$\mathbf{x} = \mathbf{M}^+\mathbf{y} = (\mathbf{M}^{\mathrm{T}}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{y}$$

We emphasize that regularization, timeseries smoothing, and other methods of incorporating atmospheric noise mitigation *are not essential* to the SBAS inversion (at least, where only one subset is concerned). Of course, what we have inverted for here is the SBAS timeseries (which is nearly identical to the timeseries of noisy displacement, in a way to be made precise in the next subsection), not the timeseries of deformation.

3.2.1 Uncertainty Quantification

What is the distinction between the timeseries of noisy displacement and the SBAS timeseries, and why is it important? The difference comes down to three factors. 1) The SBAS timeseries is referenced to the first time t_1 and thus is relative, whereas the timeseries of noisy displacement includes the position at that time and therefore is absolute. 2) As a consequence, the timeseries of noisy displacement is technically unobservable; we use it here as a theoretical tool. 3) The SBAS timeseries, by inverting the noisy interferograms, includes perturbations from interferometric noise. As we will show in this subsection, being precise about this distinction will improve our understanding of the uncertainty quantification.

We begin with the uncertainty quantification of the timeseries of noisy displacement, which is entirely due to atmospheric noise. Although highly correlated in space (which will be treated in a later section), over time atmospheric noise at a given spatial position rapidly decorrelates on a minute-to-hour timescale as the weather evolves. The off-diagonal terms in our noise covariance matrix will therefore be zero. The diagonal terms (i.e., the variance of the atmospheric noise) will not be zero and need not be identical. For an example with four timesteps, we have the following covariance matrix, where the superscript on $\mathbf{x}^a$ denotes the

absolute nature of the timeseries of noisy displacement.

$$\mathrm{cov}(\mathbf{x}^a) = \begin{pmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{pmatrix}$$

To reference the timeseries to the start time, we can introduce a referencing matrix $\mathbf{R}$.

$$\underbrace{\begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}}_{\mathbf{R}} \begin{pmatrix} x_1^a \\ x_2^a \\ x_3^a \\ x_4^a \end{pmatrix} = \begin{pmatrix} x_2^a - x_1^a \\ x_3^a - x_1^a \\ x_4^a - x_1^a \end{pmatrix} \approx \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

The covariance of the referenced noisy displacement timeseries is shown below. Note that the atmospheric noise at the first timestep percolates through the entire timeseries.

$$\begin{aligned} \mathrm{cov}(\mathbf{R}\mathbf{x}^a) &= \mathbf{R}\mathrm{cov}(\mathbf{x}^a)\mathbf{R}^{\mathrm{T}} \\ &= \begin{pmatrix} c_1 + c_2 & c_1 & c_1 \\ c_1 & c_1 + c_3 & c_1 \\ c_1 & c_1 & c_1 + c_4 \end{pmatrix} \end{aligned}$$

The covariance of the interferogram stack is related to the above covariance matrix via the SBAS matrix, with the addition of the covariance matrix of the interferometric noise η. Lacking knowledge of the structure of the interferometric noise covariance matrix, we allow it to

be completely general, with non-zero off-diagonal terms.

$$\begin{aligned}\mathbf{y} &= \mathbf{M}\mathbf{x} + \eta \\ &= \mathbf{M}\mathbf{R}\mathbf{x}^a + \eta \\ \mathrm{cov}(\mathbf{y}) &= \mathbf{M}\mathrm{cov}(\mathbf{R}\mathbf{x}^a)\mathbf{M}^{\mathrm{T}} + \mathrm{cov}(\eta)\end{aligned}$$

The SBAS timeseries is estimated by multiplying the interferogram stack with the Moore-Penrose pseudo-inverse of the SBAS matrix $\mathbf{M}^+$. Assuming $\mathbf{M}$ is full-rank, the two matrices cancel when $\mathbf{M}^+$ is left-multiplied. The structure of the SBAS network consequently affects the interferometric noise term, but has no effect on the atmospheric noise term.

$$\begin{aligned}\mathbf{x} &= \mathbf{M}^+\mathbf{y} \\ &= \mathbf{M}^+\mathbf{M}\mathbf{R}\mathbf{x}^a + \mathbf{M}^+\eta \\ &= (\mathbf{M}^{\mathrm{T}}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{M}\mathbf{R}\mathbf{x}^a + \mathbf{M}^+\eta \\ &= \mathbf{R}\mathbf{x}^a + \mathbf{M}^+\eta\end{aligned}$$

The covariance of the SBAS timeseries is therefore as below. When the interferometric noise is small, the covariance is approximately the same as that of the referenced noisy displacement timeseries; going forward, we will neglect interferometric noise entirely.

$$\begin{aligned}\mathrm{cov}(\mathbf{x}) &= \mathrm{cov}(\mathbf{R}\mathbf{x}^a) + \mathbf{M}^+\mathrm{cov}(\eta)(\mathbf{M}^+)^{\mathrm{T}} \\ &\approx \begin{pmatrix} c_1 + c_2 & c_1 & c_1 \\ c_1 & c_1 + c_3 & c_1 \\ c_1 & c_1 & c_1 + c_4 \end{pmatrix}\end{aligned}$$

3.2.2 The Implicit Zero

The temporal referencing of the SBAS timeseries to the start time has non-trivial (albeit, admittedly minor) implications for the deformation timeseries estimation and the uncertainty quantification. To see this effect, we show a synthetic example in Figure 3.2. For this short SBAS timeseries (or it could be considered to be the start of a longer timeseries), we use a linear parameter fit to estimate the deformation timeseries. The red line is estimated by excluding the implicit zero, while the green line is estimated by including the implicit zero. They have noticeably different slopes.

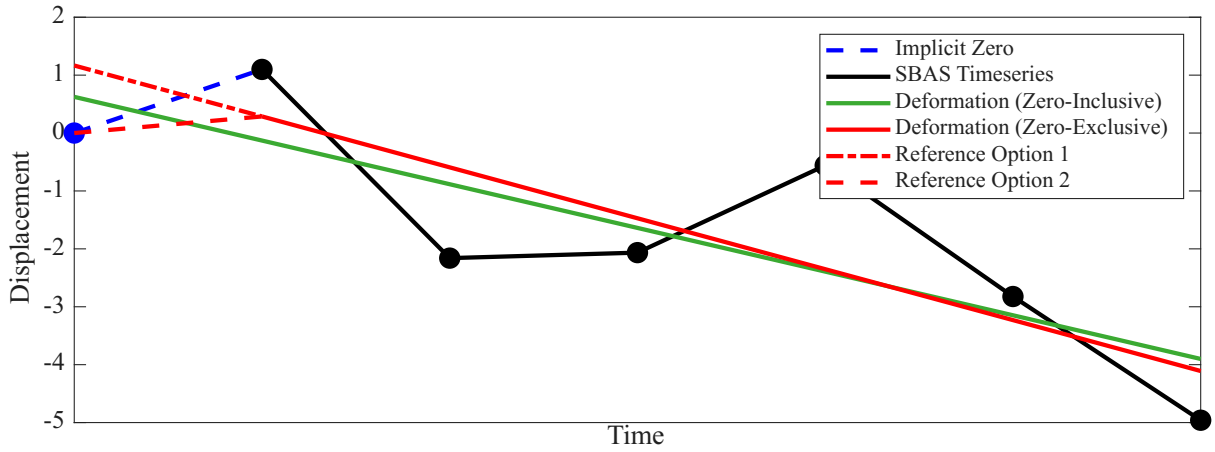


Figure 3.2: Schematic of the SBAS timeseries (black) and its implicit zero (blue), which may be made explicit. Deformation is estimated as a linear trend, either by including (green line) or excluding (red line) the zero.

That the zero is implicit in the SBAS timeseries $\mathbf{x}$ as formulated above does not imply that it should be excluded. It has physical meaning; the deformation plus noise at the start time is defined to be zero, not missing. Exclusion will remove a data point from estimate of the deformation timeseries, leading to a different velocity from the inclusive case.

The exclusionary case also leads to an undefined start-time deformation. The different start-time deformation possibilities are shown as the red dashed lines in Figure 3.2. Should the deformation be projected backwards to time zero (Reference Option 1)? Or is the start deformation zero by definition as well (Reference Option 2)? Note that neither agrees with the

inclusive case. The estimate of the tropospheric noise at time zero, as the difference between the SBAS timeseries and the estimated deformation timeseries, will be significantly different as well.

To avoid this ambiguity, we formalize the reincorporation of the start timestep with the matrix operation below. The timeseries $\mathbf{x}'$ is similar enough to $\mathbf{x}$ also be referred to as the SBAS timeseries.

$$\mathbf{x}' = \begin{pmatrix} 0 \\ \mathbf{x} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{A}_0} \mathbf{x}$$

The covariance structure is as below:

$$\mathrm{cov}(\mathbf{x}') = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & c_1 + c_2 & c_1 & c_1 \\ 0 & c_1 & c_1 + c_3 & c_1 \\ 0 & c_1 & c_1 & c_1 + c_4 \end{pmatrix}$$

3.2.3 Re-referencing and the Day-Zero Atmosphere

The SBAS timeseries with the zero reference makes sense up until this point in the processing. However, a more natural reference point going forward would be the *deformation* value at the start time. Using this value, the SBAS timeseries would no longer start at zero, but the deformation timeseries would. The deformation timeseries would here be intended to represent the displacement from the start time, and therefore it is natural that the displacement at the start time is zero, since no time has elapsed. The value of the SBAS timeseries at time zero would then be the noise value at time zero, or a close approximation of it.

Of course, we do not know *a priori* the deformation value at time zero. At this stage,

choices must be made by the user about how best to estimate the deformation from the SBAS timeseries. Common options include parameter estimation or timeseries smoothing (explored in Section 3.3). For the purposes of this subsection, it is enough to state that the deformation estimate can most often be reduced to a single matrix multiplication by a matrix which we denote **S** to evoke a smoothing matrix. This matrix is of size $n \times N$, where n is the number of times in the output deformation timeseries, while N is the number of SAR acquisition times.

$$\mathbf{d}' = \mathbf{S}\mathbf{x}'$$

The deformation at the start time may be selected with a data selection matrix.

$$\mathbf{G}_1 = \begin{pmatrix} 1 & 0 & 0 & \dots \end{pmatrix}_{1\times n}$$
$$d'_1 = \mathbf{G}_1\mathbf{d}'$$

Subtracting the reference deformation from the deformation timeseries is as below, where $\mathbf{1}_{n\times 1}$ is the column vector of ones of height n and $\mathbf{I}_{n\times n}$ is the square identity matrix of the same dimension; this subtraction is referred to as re-referencing.

$$\begin{aligned}\mathbf{d}'' &= \mathbf{d}' - \mathbf{1}_{n\times 1} d'_1 \\ &= (\mathbf{I}_{n\times n} - \mathbf{1}_{n\times 1}\mathbf{G}_1)\mathbf{d}'\end{aligned}$$

The reference deformation should likewise be subtracted from the SBAS timeseries.

$$\begin{aligned}\mathbf{x}'' &= \mathbf{x}' - \mathbf{1}_{N\times 1} d'_1 \\ &= \mathbf{x}' - \mathbf{1}_{N\times 1}\mathbf{G}_1\mathbf{d}' \\ &= (\mathbf{I}_{N\times N} - \mathbf{1}_{N\times 1}\mathbf{G}_1\mathbf{S})\mathbf{x}'\end{aligned}$$

The SBAS timeseries is sampled at the set of irregular SAR acquisition times $\{T_i\}$. However, we can choose the displacement timeseries $\mathbf{d}$ to be sampled on a set of regularly spaced intervals $\{t_j\}$, which would involve a matrix $\mathbf{S}$ which is taller than it is wide. The *data selection matrix* $\mathbf{G}$ selects the displacements on the SAR acquisition times T_i and is of size $N \times n$.

$$[\mathbf{G}]_{ij} = \begin{cases} 1 & T_i = t_j \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{d}''(T) = \mathbf{G}\mathbf{d}''$$

The noise timeseries is then estimated to be the difference between the noisy SBAS timeseries and the smooth displacement timeseries (at the SAR acquisition times). The reference values cancel because they are subtracted from both the SBAS timeseries and deformation timeseries.

$$\begin{aligned} \mathbf{a} &= \mathbf{x}'' - \mathbf{d}''(T) \\ &= \mathbf{x}'' - \mathbf{G}\mathbf{d}'' \\ &= \mathbf{x}' - \mathbf{G}\mathbf{d}' \\ &= \mathbf{x}' - \mathbf{G}\mathbf{S}\mathbf{x}' \\ &= (\mathbf{I}_{N\times N} - \mathbf{G}\mathbf{S})\mathbf{x}' \end{aligned}$$

3.2.4 Putting It All Together

Going from the interferograms to the temporally referenced displacement timeseries, we have the following expression. From right to left, we have 1) the SBAS inversion ($\mathbf{M}^+$), 2) the inclusion of the start time zero ($\mathbf{A}_0$), 3) smoothing or parameter estimation ($\mathbf{S}$, user-defined), and 4) the re-referencing ($\mathbf{I}_{n\times n} - \mathbf{1}_{n\times 1}\mathbf{G}_1$).

$$\mathbf{d}'' = (\mathbf{I}_{n\times n} - \mathbf{1}_{n\times 1}\mathbf{G}_1)\mathbf{S}\mathbf{A}_0\mathbf{M}^+\mathbf{y}$$

The similarly referenced SBAS timeseries has the following expression, except that the last step is the re-referencing by the displacement value.

$$\mathbf{x}'' = (\mathbf{I}_{N\times N} - \mathbf{1}_{N\times 1}\mathbf{G}_1\mathbf{S})\mathbf{A}_0\mathbf{M}^+\mathbf{y}$$

The tropospheric noise timeseries, as the difference between the above two expressions, has the form given below.

$$\mathbf{a} = (\mathbf{I}_{N\times N} - \mathbf{GS})\mathbf{A}_0\mathbf{M}^+\mathbf{y}$$

The generalization to TS notation is trivial and is shown below, where $\mathbf{Y}$ is the interferogram stack and $\mathbf{D}''$ is the displacement timeseries data matrix. Similar expressions for the SBAS timeseries data matrix and tropospheric noise timeseries data matrix (not shown) likewise follow from the above expressions.

$$\mathbf{D}'' = T(\mathbf{Y}, (\mathbf{I}_{n\times n} - \mathbf{1}_{n\times 1}\mathbf{G}_1)\mathbf{S}\mathbf{A}_0\mathbf{M}^+)$$

3.3 Timeseries Smoothing and Interpolation

The need for timeseries smoothing in InSAR work follows directly from the spatial noise structure of the troposphere. Since tropospheric noise is spatially correlated, spatial averaging does not meaningfully reduce the noise level. However, since tropospheric noise is temporally uncorrelated, temporal averaging may be applied to do so. Of course, any applied temporal averaging equally affects estimates of the actual displacement timeseries. However, it is generally the case that the timescales of surface displacement and atmospheric noise are different enough (the former long, the latter short) that temporal smoothing is an effective way of separating the two timeseries. The one major exception is co-seismic deformation, which occurs instantly relative to InSAR repeat timescales; therefore generic temporal smoothing cannot be applied in

such a case and special consideration must be taken.

In this section, we will present the matrix multiplication form of generic timeseries smoothing. We will then develop parameter-augmented kernel smoothing or PAKS, which melds co-seismic and smooth displacement estimation. Finally, we will address regularization and argue that it too falls under the matrix multiplication framework of timeseries smoothing.

In InSAR work, atmospheric noise mitigation is often accomplished with the common-scene stacking (CSS) method [73], which produces similar results to timeseries smoothing. The full treatment of CSS requires its own section, immediately following this one. There, we will demonstrate that CSS may be understood as convolution, and therefore it as well is encompassed within the matrix multiplication framework used here.

3.3.1 Basic Smoothing and Interpolation

In general timeseries smoothing, this operation is performed with convolution, which is mathematically equivalent to a matrix multiplication. An element d_j of $\mathbf{d}$, a smoothed and interpolated estimate of the true displacement, is a weighted mean of the elements of the noisy SBAS timeseries x_i. The weights w_{ji} are normalized to unity. For clarity, we omit the primes from the previous subsection.

$$d_j = \sum_i w_{ji} x_i \qquad\qquad \sum_i w_{ji} = 1$$

The exact construction of the weights must be determined by the user. Any filter appropriate for the task will suffice. As an example, a Gaussian filter with timescale τ is shown below.

$$w_{ji} = \frac{e^{-(t_j - t_i)^2/(2\tau^2)}}{\sum_k e^{-(t_j - t_k)^2/(2\tau^2)}}$$

The summation can be interpreted as a matrix multiplication $\mathbf{d} = \mathbf{S}\mathbf{x}$, where the smoothing

matrix $\mathbf{S}$ has elements w_{ji}. This smoothing matrix has been used in a generic sense in the previous section. We may also generalize to TS notation, which encompasses both the transformation of the full data matrix and that of the Kronecker covariance.

$$\mathbf{D} = T(\mathbf{X}, \mathbf{S})$$

3.3.2 Parameter-Augmented Kernel Smoothing

Step-like displacements from earthquakes present a particular challenge to noise mitigation by smoothing. Smoothing over an earthquake signal results in an obviously erroneous displacement timeseries. To apply smoothness to the remainder of the timeseries, therefore, the earthquake signal must be first removed [32], or as in our approach, jointly estimated. The conceptual approach we take is to first estimate a parameterization of the timeseries, remove it, apply smoothing, then reconstruct the full timeseries. Similar methods have been explored before by Guns et al. [32]; however, we present a general and powerful approach which encompasses these previous methods and fits into the mathematical framework of this chapter.

The parameters θ, the parameter matrix $\mathbf{P}$, and the SBAS timeseries $\mathbf{x}$ form the least-squares equation $\mathbf{x} = \mathbf{GP}\theta$, where the data availability matrix $\mathbf{G}$ converts between differences in temporal coverage. Accounting for a weighting matrix $\mathbf{W}$, we perform a weighted parameter estimation (Fig. 3.3a). The weights allow us to localize the parameter estimation, for example by ensuring the step function is calculated in the times shortly before and after the earthquake.

$$\begin{aligned}\theta &= \left((\mathbf{GP})^{\mathrm{T}}\mathbf{W}(\mathbf{GP})\right)^{-1}(\mathbf{GP})^{\mathrm{T}}\mathbf{Wx} \\ &= \left(\mathbf{P}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{WGP}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{Wx}\end{aligned}$$

Below is the residual $\mathbf{r}$ to the parameter fit, to which we will apply the smoothing matrix (Fig.

3.3b).

$$\mathbf{r} = \mathbf{x} - \mathbf{G}\mathbf{P}\theta$$

The reconstructed estimate of the displacement timeseries is the sum of the parameterized estimate and the residual times the smoothing matrix (Fig. 3.3c).

$$\mathbf{d} = \mathbf{P}\theta + \mathbf{S}\mathbf{r}$$

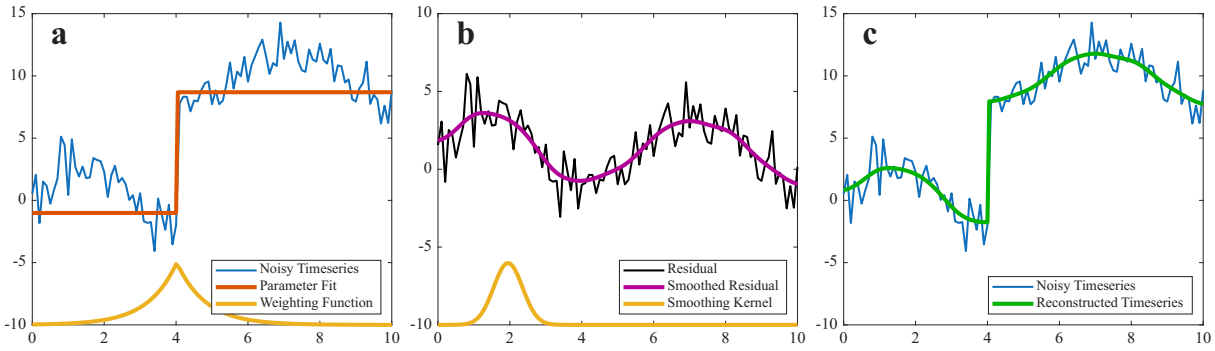


Figure 3.3: Method of parameter-augmented kernel smoothing demonstrated on a synthetic example. a) Parameter fit of a step function (red) to the noisy timeseries (blue) given an exponential weighting function (yellow). b) The parameter fit residual (black) and smoothing function (yellow) used to produce the smoothed residual (purple). c) Reconstructed timeseries (green) from addition of parameter fit and smoothed residual.

We expand and put the above expression entirely in terms of $\mathbf{x}$. The augmented smoothing matrix Λ combines both the parameterization and smoothing operations in one object. The matrix Λ

may take the place of the matrix **S** from the previous subsection where appropriate.

$$
\begin{aligned}
\mathbf{d} &= \mathbf{P}\theta + \mathbf{S}\mathbf{r} \\
&= \mathbf{P}\theta + \mathbf{S}(\mathbf{x} - \mathbf{G}\mathbf{P}\theta) \\
&= \mathbf{P}\theta + \mathbf{S}\mathbf{x} - \mathbf{S}\mathbf{G}\mathbf{P}\theta \\
&= (\mathbf{I} - \mathbf{S}\mathbf{G})\mathbf{P}\theta + \mathbf{S}\mathbf{x} \\
&= (\mathbf{I} - \mathbf{S}\mathbf{G})\mathbf{P}\left(\mathbf{P}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{G}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{x} + \mathbf{S}\mathbf{x} \\
&= \underbrace{\left[(\mathbf{I} - \mathbf{S}\mathbf{G})\mathbf{P}\left(\mathbf{P}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{G}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{W} + \mathbf{S}\right]}_{\Lambda}\mathbf{x} \\
\mathbf{d} &= \Lambda\mathbf{x}
\end{aligned}
\tag{3.1}
$$

Example

In Figure 3.3, we use a synthetic noisy timeseries where unit normal random noise (with a standard deviation of 1 unit) is imposed on a smooth timeseries and a step function with a value of 10 units. Above, we use an exponential weighting function with a timescale of 0.8 centered around the earthquake timestep. For smoothing, we use a Gaussian function with timescale 0.4.

Both the smoothing function and the weighting function have an impact on the uncertainty of the reconstructed timeseries. We see the uncertainty in Figure 3.4a for the reconstructed timeseries with the same hyper-parameters as above. In Figure 3.4b, the smoothing timescale has been shrunk to 0.1. The reconstructed timeseries fits the wiggles of noisy timeseries much more closely, but the uncertainty is noticeably larger throughout the timeseries. In Figure 3.4b, the weighting timescale has been shrunk to 0.02. The uncertainty is again larger, but only in the immediate region of earthquake step function.

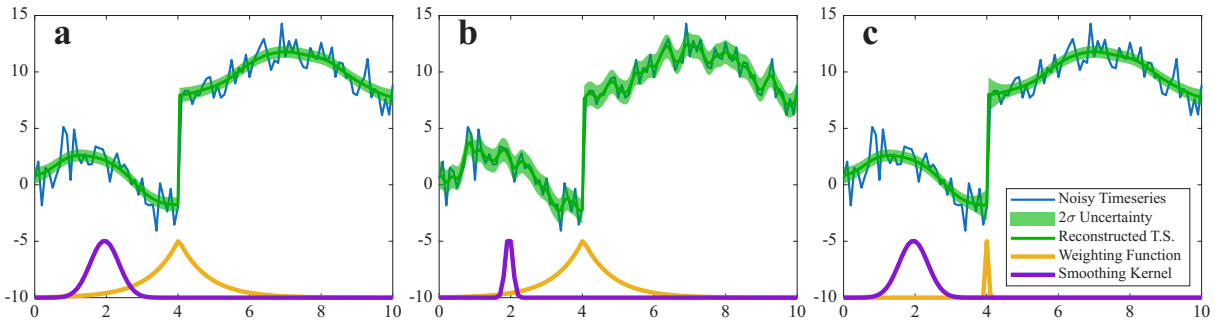


Figure 3.4: Uncertainty quantification for parameter-augmented kernel smoothing. a) Baseline case with weighting function (yellow) and smoothing kernel (purple) giving uncertainty bounds (green shaded region). b) Case with narrower smoothing kernel, demonstrating larger uncertainty. c) Case with narrower weighting function, demonstrating larger uncertainty in the immediate vicinity of the co-seismic step signal.

3.3.3 Regularization and Smoothing

Tikhanov regularization, also called ridge regression, may be reframed as smoothing by matrix multiplication and so fits within the framework presented here. We seek to minimize the misfit to the noisy data $\mathbf{x}$ subject to a regularization matrix $\mathbf{L}$ and regularization parameter μ. Then we wish to optimize the function below over all possible models $\mathbf{d}'$ to get the optimal smooth model $\mathbf{d}$. The result is equivalent to matrix multiplication by the matrix $\mathbf{S}$ below. The proof may be found in Appendix Section 3.8.2.

$$\mathbf{d} = \text{argmin}_{\mathbf{d}'}(||\mathbf{x} - \mathbf{M}\mathbf{d}'||^2 + \mu||\mathbf{L}\mathbf{d}'||^2)$$

$$\mathbf{d} = \underbrace{(\mathbf{M}^{\mathrm{T}}\mathbf{M} + \mu\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}\mathbf{M}^{\mathrm{T}}}_{\mathbf{S}}\mathbf{x}$$

3.4 Revisiting Common-Scene Stacking

Common-scene stacking (CSS) [73] is a frequently used method of atmospheric noise estimation and mitigation, which has the effect of smoothing out InSAR timeseries (Fig. 3.5). Here, we show that it can be described as a convolution operation and consequently, that one can implement uncertainty quantification using our framework. The CSS method requires two inputs,

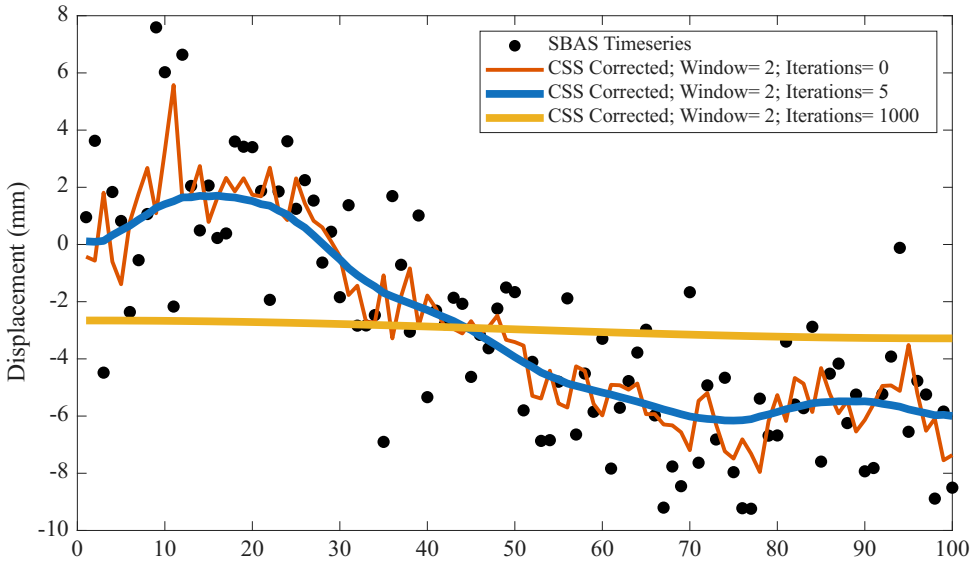


Figure 3.5: Effect of successive iterations on the degree of smoothing of the common-scene stacking method, demonstrated on a synthetic SBAS timeseries. The red line is under-smoothed, while the yellow line is over-smoothed.

the stencil size and the number of iterations, which both impact the degree of smoothing.

3.4.1 Theory

Common-Scene Stacking Algorithm

The atmospheric noise or atmospheric phase screen (APS) estimate for a given date is determined from the average of interferograms which are constructed with data from that date. Below, a_i is the APS for the i-th date, $y_{i,k}$ is the interferogram constructed from primary date i to secondary date k, and N is the one-sided window size (i.e., stencil width).

$$a_i = \frac{1}{2N} \sum_{j=1}^{N} y_{i-j,i} - y_{i,i+j} \tag{3.2}$$

Common-scene stacking involves an iterative version of the above operation, which runs as follows:

1. Initial APS is estimated for each date and removed from the interferograms.

2. The overall magnitude of the APS for each date is calculated and the dates are sorted from

largest magnitude to smallest.

3. In order, the residual APS for a given date is estimated using equation 3.2, where the interferograms $y_{i,j}$ have been corrected. It is then added to the previous APS estimate, and the interferograms are re-corrected (to be referred to as the *online correction*). The estimation and correction is applied to each date sequentially.

4. Steps 2 and 3 are repeated until a given convergence criterion or iteration limit is reached.

5. SBAS inversion is performed using the corrected interferograms to generate a timeseries of displacement.

SBAS Timeseries Formulation

Common-scene stacking may also be reformulated using the SBAS timeseries. A stack of interferograms and the corresponding SBAS timeseries may be converted between each other using the equations $\mathbf{y} = \mathbf{M}\mathbf{x}$ and $\mathbf{x} = \mathbf{M}^{+}\mathbf{y}$; neglecting interferometric noise, the conversion is exact. The interferogram $y_{i,k} \approx x_k - x_i$ is the difference between the SBAS timeseries value at time i and that at time k.

$$\begin{aligned} a_i &= \frac{1}{2N}\sum_{j=1}^{N}(x_i - x_{i-j}) - (x_{i+j} - x_i) \\ &= x_i + -\sum_{j=1}^{N}\frac{1}{2N}(x_{i-j} + x_{i+j}) \\ &= \begin{pmatrix} \cdots & -\frac{1}{2N} & -\frac{1}{2N} & 1 & -\frac{1}{2N} & -\frac{1}{2N} & \cdots \end{pmatrix} \begin{pmatrix} \vdots \\ x_{i-1} \\ x_i \\ x_{i+1} \\ \vdots \end{pmatrix} \end{aligned} \tag{3.3}$$

Offline CSS

We begin by investigating a modified version of CSS where the interferogram/SBAS timeseries correction step happens *offline*, i.e., after the entire set of residual APS have been estimated. This offline CSS procedure has cleaner theoretical qualities, which we will use to illuminate some aspects of classic CSS; in the next section we will compare results between offline and classic CSS.

Offline CSS allows us to consider the vector form of equation 3.3 below in equation 3.4, which replaces the online correction with an offline matrix multiplication. Here $\mathbf{I}$ is the identity matrix and $\mathbf{B}$ is a modified boxcar filter where the central element is zero. An example matrix $\mathbf{B}$ with a one-sided window size of 2 is shown. Note that the top two rows of the matrix, which represent the initial two dates in the timeseries, compensate for the lack of interferograms previous to the start date with re-balanced values.

$$\mathbf{a} = (\mathbf{I} - \mathbf{B})\mathbf{x} \tag{3.4}$$

$$\mathbf{B} = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & \dots \\ 0.33 & 0 & 0.33 & 0.33 & 0 & 0 & 0 & \dots \\ 0.25 & 0.25 & 0 & 0.25 & 0.25 & 0 & 0 & \dots \\ 0 & 0.25 & 0.25 & 0 & 0.25 & 0.25 & 0 & \dots \\ 0 & 0 & 0.25 & 0.25 & 0 & 0.25 & 0.25 & \dots \\ & & & \vdots & & \ddots & & \end{pmatrix}$$

Offline CSS follows the iterative procedure below. In each iteration, the atmospheric noise $\mathbf{a}$ is estimated, then removed from the SBAS timeseries $\mathbf{x}$ giving the residual timeseries $\mathbf{r}$, which is an approximation of the displacement timeseries. A new residual atmospheric noise

map $\Delta\mathbf{a}$ is calculated from the corrected SBAS timeseries, then added to the previous estimate.

$$\begin{aligned}
\mathbf{a}^{(0)} &= \mathbf{0} && \text{Initial APS Estimate} \\
\mathbf{r}^{(k)} &= \mathbf{x} - \mathbf{a}^{(k)} && \text{Residual} \\
\Delta\mathbf{a}^{(k)} &= (\mathbf{I} - \mathbf{B})\mathbf{r}^{(k)} && \text{Incremental APS Estimate} \\
\mathbf{a}^{(k+1)} &= \mathbf{a}^{(k)} + \Delta\mathbf{a}^{(k)} && \text{Next APS Estimate}
\end{aligned}$$

Some simplification highlights that the residual follows an especially simple iterative rule.

$$\begin{aligned}
\mathbf{r}^{(k+1)} &= \mathbf{x} - \mathbf{a}^{(k+1)} \\
&= \mathbf{x} - \mathbf{a}^{(k)} - \Delta\mathbf{a}^{(k)} \\
&= \mathbf{r}^{(k)} - (\mathbf{I} - \mathbf{B})\mathbf{r}^{(k)} \\
\mathbf{r}^{(k+1)} &= \mathbf{B}\mathbf{r}^{(k)}
\end{aligned}$$

Since $\mathbf{r}^{(0)} = \mathbf{x}$, the closed form solution is $\mathbf{r}^{(k)} = \mathbf{B}^k\mathbf{x}$. With each iteration, the residual timeseries $\mathbf{r}$ becomes successively smoothed by the modified boxcar matrix. The corresponding APS estimate is $\mathbf{a}^{(k)} = (\mathbf{I} - \mathbf{B}^k)\mathbf{x}$. For a CSS procedure with a given stencil and number of iterations, the degree of smoothing can be predicted by constructing the $\mathbf{B}$ matrix and raising it to the power of the number of iterations.

In the limit of convergence, the matrix $\mathbf{B}^\infty = \lim_{k\to\infty}\mathbf{B}^k$ becomes so smooth as to be completely flat; see the effect of over-smoothing on the yellow line in Figure 3.5. Each element of $\mathbf{B}^\infty$ more than N dates away from the beginning and end of the timeseries is $1/N_t$, where N_t is the number of dates in the SBAS timeseries. The corresponding value for the APS estimate in the convergence limit is $\mathbf{a}^{(\infty)} = \mathbf{x} - \text{mean}(\mathbf{x})$, which is to say the estimate of atmospheric noise is simply the SBAS timeseries, temporally referenced by its mean value. To preserve the smoothing feature of common-scene stacking without over-smoothing, the iteration should be terminated

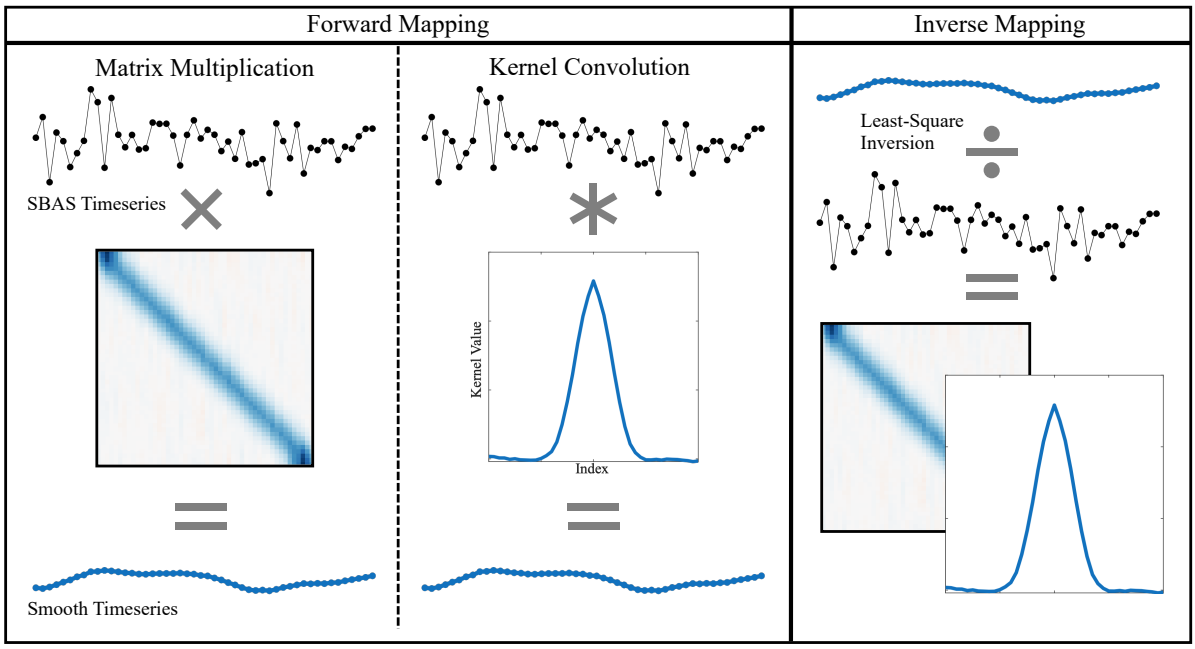


Figure 3.6: CSS is approximately equivalent to a forward mapping of the SBAS timeseries through a smoothing matrix or through kernel convolution. The inverse mapping uses smooth timeseries calculated with CSS to invert for the CSS-equivalent smoothing matrix and kernel.

before convergence.

3.4.2 CSS-Equivalent Smoothing

If classic CSS can be viewed as a smoothing method, what would be the functionally equivalent smoothing kernel? We address this question with the following procedure: 1) We construct synthetic timeseries of atmospheric noise which have power-law spatial structure and realistic temporal statistics. 2) We perform the CSS method on the synthetic timeseries. 3) We invert for the map between the input timeseries and the smooth result, to be referred to as the CSS-equivalent smoothing matrix. 4) We average the cross-sections of this matrix to get a CSS-equivalent smoothing kernel. The relationship of CSS to the mappings of the smoothing matrix and kernel is diagrammed in Figure 3.6.

We construct synthetic timeseries of atmospheric noise which is generated with a power-law slope of 1.5. The variance of each scene is scaled by $|\xi|^{0.7}$, where ξ is a sample from a

normal random distribution. This ensures that many dates have relatively small variance, while some have large variance, similar to observations of the real atmosphere. For simplicity, the synthetic SAR acquisitions happen regularly and every interferogram is exactly the difference between the SBAS values with no interferometric noise.

The inverse mapping is accomplished by constructing matrices of the input timeseries (eq. 3.5) and CSS-smoothed timeseries (eq. 3.6). The CSS-equivalent smoothing matrix $\mathbf{S}_{CSS}$ is then given by equation 3.7.

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \dots \end{bmatrix} \tag{3.5}$$

$$\mathbf{Y} = \begin{bmatrix} CSS(\mathbf{x}_1) & CSS(\mathbf{x}_2) & CSS(\mathbf{x}_3) & \dots \end{bmatrix} \tag{3.6}$$

$$\mathbf{Y} = \mathbf{S}_{CSS}\mathbf{X} \implies \mathbf{S}_{CSS} = (\mathbf{Y}\mathbf{X}^{\mathrm{T}})(\mathbf{X}\mathbf{X}^{\mathrm{T}})^{-1} \tag{3.7}$$

The CSS-equivalent smoothing matrices are shown in Figure 3.7. For visual clarity, only the bottom right corner of the matrices are displayed, allowing observation of the effects at the beginning of the timeseries due to one-sided windowing. Both larger window size and a greater number of iterations contribute to the smoothing. Successive iterations act to smooth out the residual timeseries completely.

When constructing a smoothing matrix from a known kernel, the kernel is placed in each row of the smoothing matrix such that it is centered on the diagonal. Conversely, to estimate the kernel from a CSS-equivalent smoothing matrix, we averaged together the rows of the matrix. The resulting CSS-equivalent smoothing kernels are shown in Figure 3.8. The zeroth iteration shows the modified boxcar filter. The first iteration approximates a triangular kernel, appropriate for the convolution of a boxcar filter with itself. Successive iterations continually broaden and soften the kernel. In the limit of convergence, the kernel is flat. In Figure 3.9, we also compare the classic CSS (left) to the analytical solutions of the offline CSS method of the previous section (right). We see that the online correction introduces additional smoothing and serves to soften

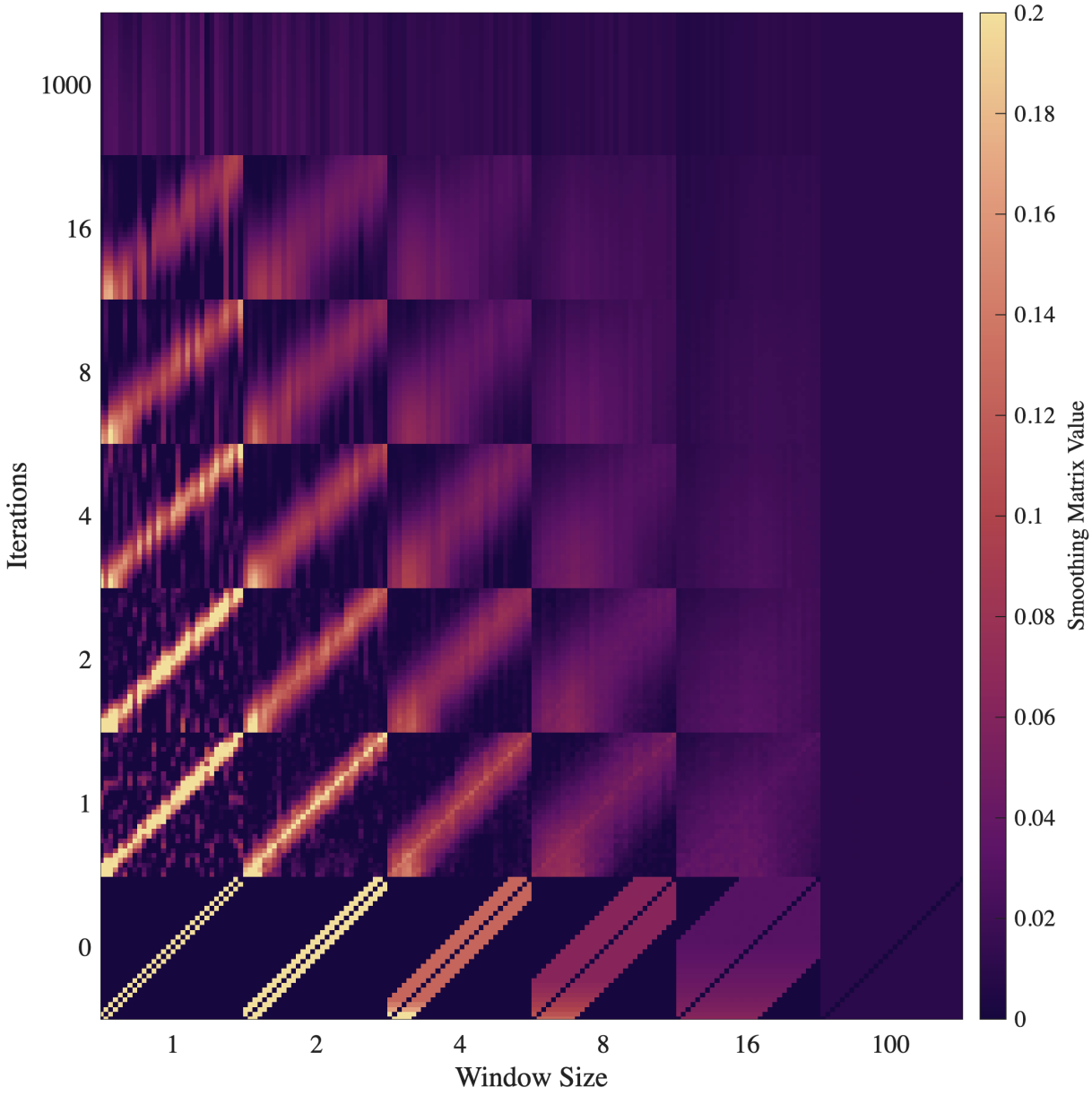


Figure 3.7: CSS-equivalent smoothing matrices for a range of window sizes and iterations. Only the bottom right corner of each matrix is shown for visual clarity.

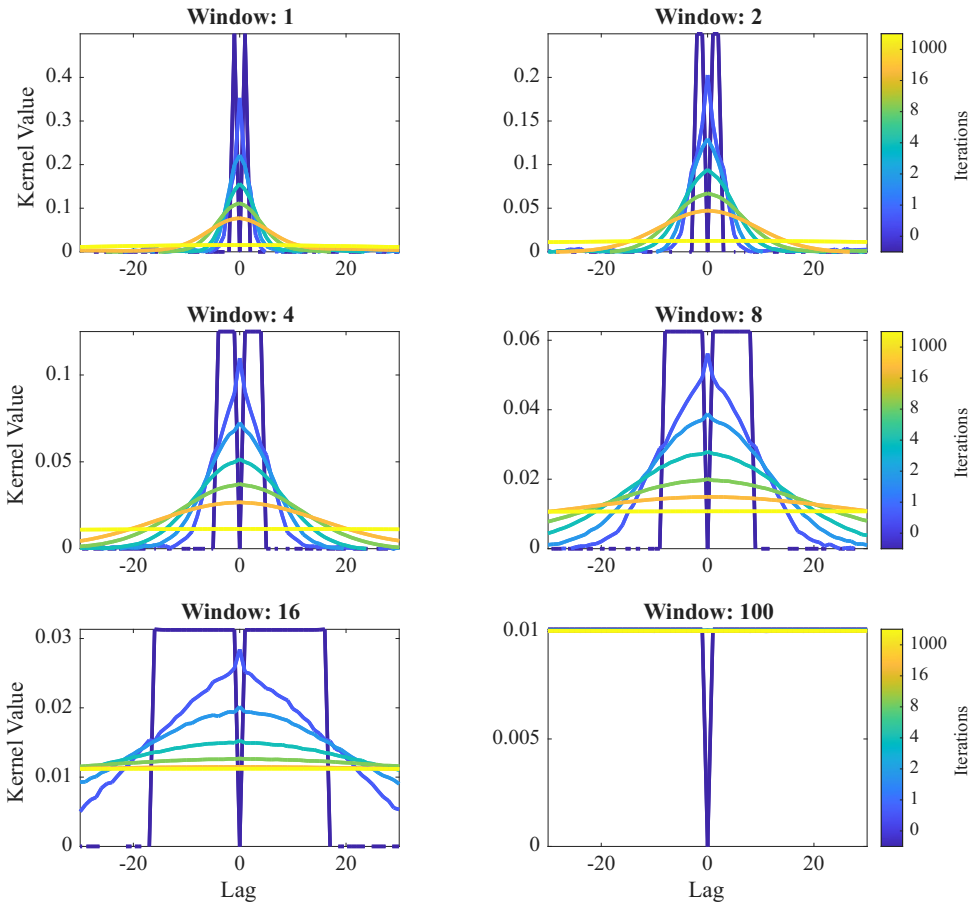


Figure 3.8: CSS-equivalent smoothing kernels, as a function of window size and number of iterations.

the sharp peaks observed in the offline case for low number of iterations. Classic CSS is better approximated by the offline CSS theory at larger iteration values.

3.5 Spatial Referencing

Spatial referencing describes the procedure of subtracting from every interferogram a reference, ensuring that the deformation is relative to that reference. As described in the Introduction, spatial referencing is necessary because of the fundamental phase ambiguity of wrapped interferograms. A reference can be a single pixel, the mean value of the whole interferogram, the planar trend across the interferogram, a long-wavelength function derived from GNSS, or any similar user-defined choice. Referencing can also include correction for

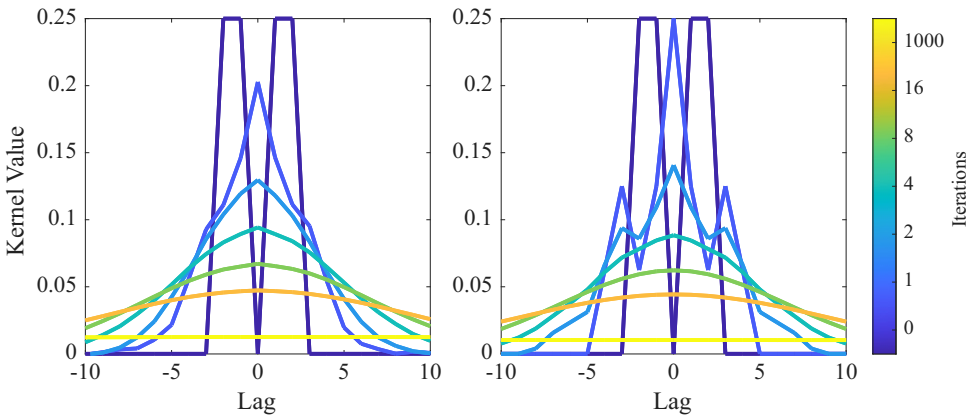


Figure 3.9: Left) Classic CSS-equivalent smoothing kernels with a window size of 2. Right) Offline CSS analytically-derived kernels from the theory in the previous section.

stratified tropospheric noise.

Spatial referencing interacts with the power-law nature of tropospheric noise to have a strong effect on the spatial distribution of uncertainty in InSAR data. In Figure 3.10, we demonstrate this effect with a stack of synthetic tropospheric noise with a power-law slope of 1.5, an example field of which is shown (Fig. 3.10a). Referencing is performed relative to the mean value of the scene (Fig. 3.10b), a single point (Fig. 3.10c), and the mean value of four points in the corners of the scene (Fig. 3.10d).

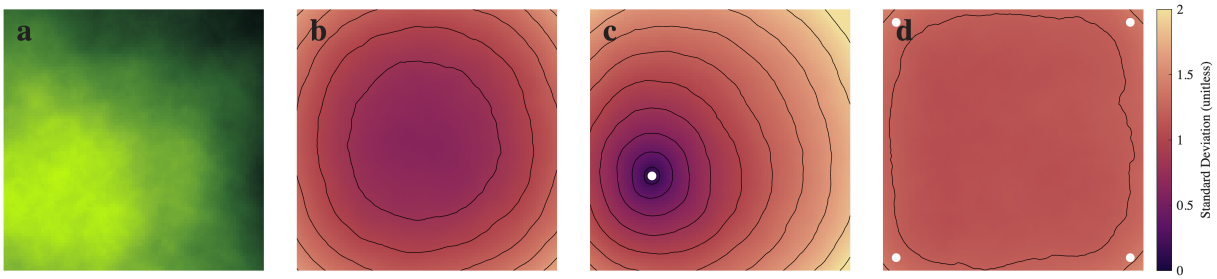


Figure 3.10: a) Example synthetic atmospheric delay field with a power-law slope of 1.5. b) The pixel-wise standard deviation of a stack (in the temporal domain) of such delay fields is not uniform, but highly dependent on the geometry of the scene. Here each delay field has been spatially demeaned and normalized by its standard deviation, so the pixel-wise standard deviation of the stack is unitless. c) The scenes are referenced to a single point, causing stronger gradients and a lower minimum and higher maximum standard deviation. d) The scenes are referenced by demeaning relative to four pixels in the corners of the scene, causing a more uniform spatial standard deviation pattern.

There are trade-offs for each case. The first case demonstrates that: 1) using the maximum number of points does not meaningfully minimize the total variance, which is due to the spatial correlation of the noise; and 2) that the uncertainty pattern is not uniform but is roughly circular and is centered on the bounds of the scene. The second case demonstrates that minimal uncertainties may be obtained in the immediate vicinity of a single reference point, at the cost of higher uncertainties further away. The third case demonstrates that uniform uncertainties may be obtained by referencing to the surrounding region; however, the minimal level of that uncertainty is higher than that of the other cases.

3.5.1 Translation of Spatial Referencing Operations

Spatial fields are two-dimensional, but since spatial covariance matrices contain information about the correlations between each pair of pixels in the scene, the spatial fields must be vectorized so as to be one-dimensional. Unfortunately, although the process is nominally similar to the data matrix case, the spatial covariance matrix $\mathbf{C}_s$ cannot itself be regarded as a Kronecker product between two smaller covariance matrices for the two horizontal axes. Spatial covariance matrices therefore must be stored in full, meaning that they have N_{pixels}^2 values in total. A large, high-resolution InSAR scene may easily have over a million pixels, a size which may cause practical difficulties in matrix manipulation for some users. In such a case, it is often sufficient for the purposes of the spatial uncertainty quantification to down-sample to a more manageable number of pixels.

Parameter Fitting, Elevation Detrending, Demeaning, and Single Point Referencing

One approach is to reference to a low-order polynomial, i.e., polynomial detrending. For example, Neely et al, 2019 [56] fits a low-order polynomial field to a set of InSAR/GNSS residuals.

We begin with the recognition that the subtraction of the residual is the same as the subtraction of the referencing of the InSAR to the GNSS locations, then the re-addition of the

GNSS field. Therefore we may investigate the effects of the InSAR-only referencing step and the GNSS addition on the uncertainty quantification separately. The logic is demonstrated below.

$$\begin{aligned}
\text{InSAR}_{\text{GNSS}} &= \text{InSAR} - \text{paramFit}(\text{Residual}) \\
&= \text{InSAR} - \text{paramFit}(\text{InSAR} - \text{GNSS}) \\
&= \text{InSAR} - \text{paramFit}(\text{InSAR}) + \text{paramFit}(\text{GNSS}) \\
&= \text{InSAR}_{\text{Referenced}} + \text{paramFit}(\text{GNSS})
\end{aligned}$$

We begin with the least-squares equation $\mathbf{x} = \mathbf{P}\theta$, where the parameter vector is θ and the associated parameter matrix is $\mathbf{P}$. In a change from the previous section, the vector $\mathbf{x}$ now represents not a timeseries, but a spatial field. Shown below is a planar trend parameter matrix (left) and a second-order polynomial parameter matrix (right), for the horizontal coordinates x and y.

$$\mathbf{P} = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ & \vdots & \end{pmatrix} \qquad \mathbf{P} = \begin{pmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2 y_2 & y_2^2 \\ 1 & x_3 & y_3 & x_3^2 & x_3 y_3 & y_3^2 \\ & & & \vdots & & \end{pmatrix}$$

Stratified tropospheric estimation and correction can also be included by appending a function of elevation in the parameter matrix. Shown below is a planar trend parameter matrix with linear elevation dependence (left), an elevation-only parameter matrix with exponential elevation dependence (center), and an elevation-only parameter matrix with linear elevation dependence but where the elevation trend varies linearly across the scene horizontally. The

elevation at the location (x_i, y_i) is E_i and Γ is the lapse rate.

$$\mathbf{P} = \begin{pmatrix} 1 & x_1 & y_1 & E_1 \\ 1 & x_2 & y_2 & E_2 \\ 1 & x_3 & y_3 & E_3 \\ & & \vdots & \end{pmatrix} \qquad \mathbf{P} = \begin{pmatrix} 1 & e^{-E_1/\Gamma} \\ 1 & e^{-E_2/\Gamma} \\ 1 & e^{-E_3/\Gamma} \\ & \vdots \end{pmatrix} \qquad \mathbf{P} = \begin{pmatrix} 1 & E_1 & x_1E_1 & y_1E_1 \\ 1 & E_2 & x_2E_2 & y_2E_2 \\ 1 & E_3 & x_3E_3 & y_3E_3 \\ & & \vdots & \end{pmatrix}$$

Demeaning (subtracting the mean value) and referencing to a single point are special cases of the parameter referencing approach where the parameter matrix has only one column. Demeaning across the whole scene is accomplished by choosing the parameter matrix to be the vector of ones $\mathbf{P} = \mathbf{1}_{N_{pixels} \times 1}$. If there are multiple pixels to consider, the parameter matrix is a vector where ones mark included pixels and zeros mark excluded pixels. Below, I_{ref} is the set of reference indices. Single point referencing is a further subset of demeaning for the case where the set I_{ref} contains only a single index.

$$[\mathbf{P}]_i = \begin{cases} 1 & i \in I_{\text{ref}} \\ 0 & \text{otherwise} \end{cases}$$

When solving the least-squares equation, we include a weighting matrix $\mathbf{W}$, which is useful to include or exclude certain points when a diagonal matrix; points to be excluded from the fit may be given a diagonal value of zero, while included points receive a value of unity.

$$\theta = (\mathbf{P}^{\mathrm{T}}\mathbf{W}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{W}\mathbf{x}$$

We then subtract the parameter fit to get the matrix expression of polynomial detrending

below, where $\mathbf{B}$ is the spatial referencing matrix.

$$\begin{aligned}\mathbf{x}_R &= \mathbf{x} - \mathbf{P}\theta \\ &= \mathbf{x} - \mathbf{P}(\mathbf{P}^{\mathrm{T}}\mathbf{W}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{W}\mathbf{x} \\ &= \underbrace{(\mathbf{I} - \mathbf{P}(\mathbf{P}^{\mathrm{T}}\mathbf{W}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{W})}_{\mathbf{B}}\mathbf{x}\end{aligned}$$

Generalizing to TS notation for data matrix manipulation, we have the following:

$$\mathbf{X}_R = S(\mathbf{X}, \mathbf{B}) \tag{3.8}$$

The last step is the re-addition of the GNSS-derived field. If provided with an uncertainty quantification which includes a covariance matrix, the two covariance matrices may simply be added (see Section 3.1.4).

Interpolation

Another similar method in Xu et al, 2021 [79] is to subtract an interpolated, filtered version of the residual between InSAR and GNSS at points co-located with the GNSS stations, thus tying the InSAR to the GNSS frame. We provide a translation of this procedure to our matrix multiplication framework here. Following the same logic as above, we focus on the InSAR-only referencing component.

We start by creating an interpolation matrix $\mathbf{L}$ (size $N_{pixels} \times N_{stations}$), which is straightforward to do simply by matching inputs and outputs. We can construct the i-th column of the interpolation matrix by setting the value of each GNSS station to zero except for the i-th station, which is set to unity, then running those values through the interpolation algorithm. We then

repeat this procedure for each station.

$$\mathbf{L} = \begin{bmatrix} \text{interp} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} & \text{interp} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} & \dots \end{bmatrix}$$

We also create a station transfer matrix $\mathbf{G}$ with size $N_{stations} \times N_{pixels}$, which accounts for the averaging scheme used to go from the InSAR grid to the GNSS station. As an example, the averaging scheme could be the mean value of all pixels within a certain distance R (eq. 3.9) or a Gaussian weighted average (eq. 3.10). The constant k_j normalizes the rows of the matrix to unity. Note that $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ can take the place of the weight matrix $\mathbf{W}$ of the previous section when a similar form of station averaging is used.

$$[\mathbf{G}]_{ji} = k_j * (\text{dist}(X_j, Y_j, x_i, y_i) \leq R) \tag{3.9}$$

$$[\mathbf{G}]_{ji} = k_j * \exp\left(-\frac{\text{dist}(X_j, Y_j, x_i, y_i)^2}{2R^2}\right) \tag{3.10}$$

$$\sum_i [\mathbf{G}]_{ji} = 1$$

Following the procedure in Xu et al., 2021 [79], we add a smoothing matrix $\mathbf{S}$ to smooth the interpolating functions. The matrix multiplication form of referencing is below, where $\mathbf{B}$ is the spatial referencing matrix. The TS formulation is identical to equation 3.8.

$$\begin{aligned} \mathbf{x}_R &= \mathbf{x} - \mathbf{SLGx} \\ &= \underbrace{(\mathbf{I} - \mathbf{SLG})}_{\mathbf{B}}\mathbf{x} \end{aligned}$$

3.5.2 Uncertainty Quantification for the Structure Function

Following Lohman et al., 2005 [47], the spatial covariance matrix of $\mathbf{C}$ is the variance of the scene σ^2 minus half the structure function matrix Σ. The dimension of the column vector of ones $\mathbf{1}$ is $N \times 1$.

$$[\mathbf{C}]_{ij} = \sigma^2 - \frac{1}{2}[\Sigma]_{ij}$$
$$\mathbf{C} = \sigma^2 \mathbf{1}\mathbf{1}^{\mathrm{T}} - \frac{1}{2}\Sigma$$

Here we show that the overall variance σ^2 cancels under any referencing procedure. We express the referenced covariance $\mathbf{C}_R$ in terms of the spatial referencing matrix $\mathbf{B}$ and the original covariance, then in terms of the structure function matrix.

$$\begin{aligned}\mathbf{C}_R &= \mathbf{B}\mathbf{C}\mathbf{B}^{\mathrm{T}} \\ &= \mathbf{B}(\sigma^2 \mathbf{1}\mathbf{1}^{\mathrm{T}} - \frac{1}{2}\Sigma)\mathbf{B}^{\mathrm{T}} \\ &= \sigma^2(\mathbf{B}\mathbf{1})(\mathbf{B}\mathbf{1})^{\mathrm{T}} - \frac{1}{2}\mathbf{B}\Sigma\mathbf{B}^{\mathrm{T}}\end{aligned}$$

The term $\mathbf{B1}$ is the spatial field that would result from referencing/detrending a uniform field with the value unity. Since any referencing procedure should be able to perfectly nullify such a simple field, we conclude that $\mathbf{B1} = \mathbf{0}$. As a result, the referenced covariance matrix $\mathbf{C}_R$ is expressed entirely in terms of the structure function matrix and has no dependence on the overall scene variance σ^2. We observe in all cases (but do not prove) that $\mathbf{C}_R$ is positive-definite.

$$\begin{aligned}\mathbf{C}_R &= \sigma^2 \mathbf{0}\mathbf{0}^{\mathrm{T}} - \frac{1}{2}\mathbf{B}\Sigma\mathbf{B}^{\mathrm{T}} \\ &= -\frac{1}{2}\mathbf{B}\Sigma\mathbf{B}^{\mathrm{T}}\end{aligned}$$

3.5.3 Example

We demonstrate the utility of the spatial referencing uncertainty quantification in Figure 3.11. For this figure, we use the interpolation scheme of Section 3.5.1 over the GNSS stations of Southern California. We assume a power-law slope of 1.1. We use a linear scattered interpolation and filter using a 2D Gaussian filter with standard deviation of 80 km. As we are only working with the spatial covariance in this example, the results are normalized to the power-law amplitude. Figure 3.11a shows the uncertainty of the referenced covariance matrix, which is relatively uniform across the interior of scene, although there are edge effects. However, there are substantial correlations on short-wavelengths, ranging up to 100 km (Fig. 3.11b). A sample drawn from the covariance matrix (3.11c) demonstrates the expected noise structure present for a single date.

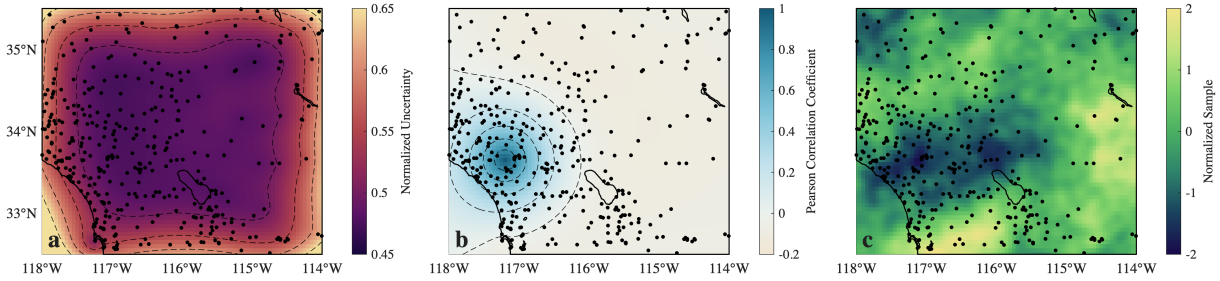


Figure 3.11: a) Standard deviation, i.e., square root of the diagonal of the covariance matrix. b) Pearson correlation coefficient of a single pixel with the rest of the scene. c) Random sample drawn from the covariance matrix.

3.6 Covariance Construction

Until this point, we assumed that the tropospheric noise covariance structure in the temporal and spatial domains is known. In practice, uncertainties are not know but must be estimated, a topic which is the focus of this section. Where the previous sections focused on the translation of InSAR processing operations could be mathematically precise, this section of necessity involves a number of user choices and simplifications. Therefore we do not claim our approach to be the only, or only valid, procedure; rather, we seek only to present an example

of an internally-consistent, data-driven covariance construction framework which addresses the major InSAR processing operations.

The general approach taken here is to use the data itself to calculate statistics used to construct the covariance matrices. The most relevant quantity is the residual atmosphere, which is the SBAS timeseries minus the estimated displacement timeseries. The full processing chain is listed below:

1. Reference all interferograms to a single common point.
2. Perform SBAS inversion to get the SBAS timeseries.
3. Estimate displacement timeseries by performing timeseries smoothing with an appropriate timescale. If there is a co-seismic displacement signal, use the PAKS method.
4. Subtract displacement from SBAS timeseries to get an estimate of the atmospheric noise timeseries.
5. For each date, estimate the stratified component of noise by regressing the atmospheric noise against elevation, then subtract it.
6. For each date, estimate the atmospheric structure function with the transect method.
7. For each date, estimate the power-law amplitude $a_t(t)$ and power-law slope $\gamma(t)$ of the structure function (Fig. 3.12).
8. Using the median value of the power-law slope, generate a spatial covariance matrix as the matrix of pixel distances ρ to the element-wise power γ (eq. 3.11).
9. Using the timeseries of the power-law amplitude, generate a diagonal temporal covariance matrix (eq. 3.12). The combined Kronecker matrix is shown in equation 3.13.
10. Using the desired spatial referencing scheme, re-reference the SBAS timeseries and repeat step 3.

11. Apply the translated transformations to the covariance matrix to perform the uncertainty quantification.

$$\mathbf{C}_s = -\frac{1}{2}\rho^{\circ\gamma} \tag{3.11}$$

$$\mathbf{C}_t = \mathrm{diag}(a(t)^2) \tag{3.12}$$

$$\mathbf{C} = \mathbf{C}_s \otimes \mathbf{C}_t \tag{3.13}$$

Other valid approaches could build the uncertainty quantification by incorporating outside information, such as weather reanalysis data. As was demonstrated in Chapter 2, GNSS data can provide information about seasonal biases which otherwise might be missed in a InSAR-only uncertainty framework. However, a full treatment of externally driven approaches is beyond the scope of this work.

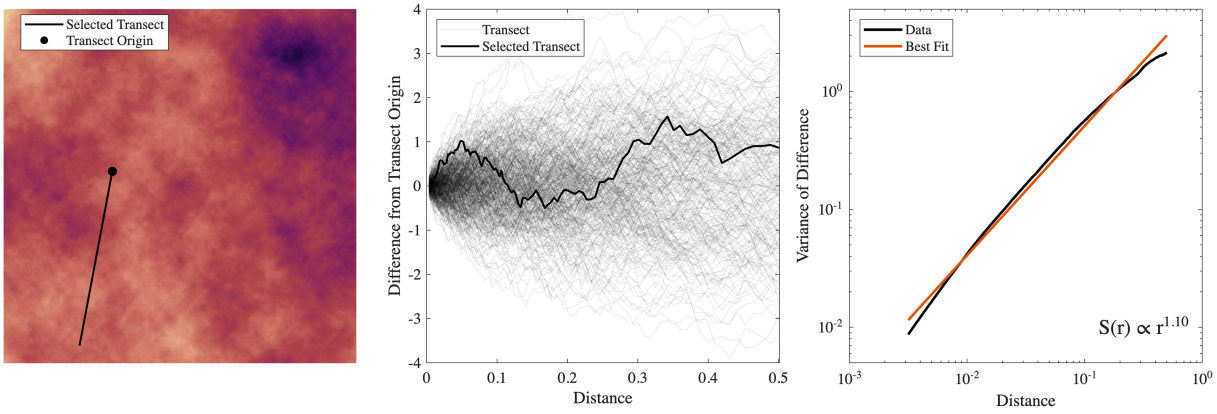


Figure 3.12: Transect method of power-law estimation. Left) A synthetic field of tropospheric delay with a power-law slope of 1.1, along with a selected transect. Center) The difference in the value of the selected transect from the transect origin (bold line) and that of other random transects (thin lines). Right) Variance of the values in the center plot at each distance (black line) and linear best fit in log-log space (red).

3.7 Summary

We present a theoretical framework of uncertainty quantification for InSAR interferogram stacks and displacement timeseries based on the Kronecker product of spatial and temporal covariance matrices. Our framework is mathematically cohesive, useful for theoretical studies of InSAR uncertainties, and practical for common use in InSAR processing. We also translate commonly used InSAR processing methods into the matrix multiplication framework of this chapter. We revisit the SBAS inversion, common-scene stacking, a variety of spatial referencing schemes. We also introduce parameter-augmented kernel smoothing to simultaneously solve for a co-seismic offset and a smoothly varying timeseries. We believe that this work has the potential to serve as a foundation of uncertainty quantification for the InSAR community.

3.8 Appendix

3.8.1 Extracting Useful Quantities from Kronecker Covariances

It is one matter to simply have the full Kronecker covariance matrix at hand, but another to visualize uncertainties for derived quantities. In this section we will describe how to extract the covariances and uncertainties for 1) single timeseries or images, 2) instantaneous differences (such as co-seismic displacement), 3) mean values, and 4) least-squares parameter fits (such as velocity estimation).

To extract the timeseries at the third pixel, apply the data selection matrix $\mathbf{G}$ below in a spatial transformation. The result is the temporal covariance matrix multiplied by the spatial variance of that pixel.

$$\begin{aligned} \mathrm{cov}(\mathrm{vec}(\mathbf{X})) &= \mathbf{C}_s \otimes \mathbf{C}_t \\ \mathbf{G} &= \begin{pmatrix} 0 & 0 & 1 & 0 & \dots \end{pmatrix} \\ \mathrm{cov}(\mathbf{x}_3(t)) &= \sigma_3^2 \mathbf{C}_t \end{aligned}$$

To estimate the spatial distribution of uncertainty due for a co-seismic displacement signal, take the temporal difference between the two timesteps bracketing the earthquake. The result is proportional to the spatial covariance matrix.

$$\begin{aligned} \mathbf{G} &= \begin{pmatrix} 0 & -1 & 1 & 0 & \dots \end{pmatrix} \\ \mathrm{cov}(\mathbf{x}(x,y)) &= (\sigma_3^2 + \sigma_2^2 - 2C_{23})\mathbf{C}_x \end{aligned}$$

A spatial or temporal average is accomplished with the appropriate transformation using

the averaging matrix **G** below.

$$\mathbf{G} = \frac{1}{N}\begin{pmatrix} 1 & 1 & 1 & 1 & \dots \end{pmatrix}$$

To estimate the velocity using a parameter fit which accounts for seasonal motion, generate the parameter matrix **P** below. Then the multiplying matrix **G** estimates the parameters and selects for the velocity.

$$\mathbf{P} = \begin{pmatrix} 1 & t_1 & \cos(2\pi t_1) & \sin(2\pi t_1) \\ 1 & t_2 & \cos(2\pi t_2) & \sin(2\pi t_2) \\ 1 & t_3 & \cos(2\pi t_3) & \sin(2\pi t_3) \\ & & \vdots & \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}$$

3.8.2 Proof of Regularization as Matrix Multiplication

In this section we prove that Tikhanov regularization or ridge regression can be reformulated as matrix multiplication.

$$
\begin{aligned}
\mathbf{d} &= \text{argmin}_{\mathbf{d}'}(||\mathbf{x}-\mathbf{M}\mathbf{d}'||^2+\mu||\mathbf{L}\mathbf{d}'||^2) \\
&= \text{argmin}_{\mathbf{d}'}\left(\left|\left|\begin{pmatrix}\mathbf{M}\mathbf{d}'-\mathbf{x}\\ \sqrt{\mu}\mathbf{L}\mathbf{d}'\end{pmatrix}\right|\right|^2\right) \\
&= \text{argmin}_{\mathbf{d}'}\left(\left|\left|\underbrace{\begin{pmatrix}\mathbf{M}\\ \sqrt{\mu}\mathbf{L}\end{pmatrix}}_{\tilde{\mathbf{M}}}\mathbf{d}'-\underbrace{\begin{pmatrix}\mathbf{x}\\ \mathbf{0}\end{pmatrix}}_{\tilde{\mathbf{x}}}\right|\right|^2\right) \\
&= \text{argmin}_{\mathbf{d}'}(||\tilde{\mathbf{M}}\mathbf{d}'-\tilde{\mathbf{x}}||^2) \\
&= (\tilde{\mathbf{M}}^{\mathrm{T}}\tilde{\mathbf{M}})^{-1}\tilde{\mathbf{M}}^{\mathrm{T}}\tilde{\mathbf{x}} \\
&= (\mathbf{M}^{\mathrm{T}}\mathbf{M}+\mu\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{x}
\end{aligned}
$$

Bibliography

[1] Industrial water supplier information, 2026. [Dataset].

[2] Well driller's report, 2026. [Dataset].

[3] Falk Amelung, Devin L Galloway, John W Bell, Howard A Zebker, and Randell J Laczniak. Sensing the ups and downs of las vegas: Insar reveals structural control of land subsidence and aquifer-system deformation. *Geology*, 27(6):483–486, 1999.

[4] AR Amiri-Simkooei. On the nature of gps draconitic year periodic pattern in multivariate position time series. *Journal of Geophysical Research: Solid Earth*, 118(5):2500–2511, 2013.

[5] Homa Ansari, Francesco De Zan, and Alessandro Parizzi. Study of systematic bias in measuring surface deformation with sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2):1285–1301, 2020.

[6] Reuben Attah, Kamaljeet Kaur, Kevin D Perry, Diego P Fernandez, and Kerry E Kelly. Assessing the oxidative potential of dust from great salt lake. *Atmospheric Environment*, 336:120728, 2024.

[7] William D Barnhart, Ryan D Gold, and James Hollingsworth. Localized fault-zone dilatancy and surface inelasticity of the 2019 ridgecrest earthquakes. *Nature Geoscience*, 13(10):699–704, 2020.

[8] Bonnie K Baxter and Jaimi K Butler. Climate change and great salt lake. In *Great Salt Lake biology: a terminal lake in a time of change*, pages 23–52. Springer, 2020.

[9] François Beauducel, Pierre Briole, and Jean-Luc Froger. Volcano-wide fringes in ers synthetic aperture radar interferograms of etna (1992–1998): Deformation or tropospheric effect? *Journal of Geophysical Research: Solid Earth*, 105(B7):16391–16402, 2000.

[10] Paolo Berardino, Gianfranco Fornaro, Riccardo Lanari, and Eugenio Sansosti. A new algorithm for surface deformation monitoring based on small baseline differential sar interferograms. *IEEE Transactions on geoscience and remote sensing*, 40(11):2375–2383, 2002.

[11] Siiri Bigalke, Paul Loikith, and Nicholas Siler. Explaining the 2022 record low great salt lake volume. *Geophysical Research Letters*, 52(2):e2024GL112154, 2025.

[12] Maurice A Biot. General theory of three-dimensional consolidation. *Journal of applied physics*, 12(2):155–164, 1941.

[13] Frederick Blume, Thomas Herring, Glen S Mattioli, Karl Feaux, Christian P Walls, Kenneth Emil Austin, and Stephen T Dittmann. Stability of gnss monumentation: Analysis of co-located monuments in the unavco plate boundary observatory. In *AGU Fall Meeting Abstracts*, volume 2017, pages G14A–02, 2017.

[14] Claudie Carnec and Hubert Fabriol. Monitoring and modeling land subsidence at the cerro prieto geothermal field, baja california, mexico, using sar interferometry. *Geophysical Research Letters*, 26(9):1211–1214, 1999.

[15] Jay R Cederberg, Philip M Gardner, and Susan A Thiros. Hydrology of northern utah valley, utah county, utah, 1975-2005. Technical report, US Geological Survey, 2009.

[16] David Wayne Clark and Cynthia L Appel. Ground-water resources of northern utah valley, utah. *USGS Res. on Mineral Resources, 1985*, page 82, 1985.

[17] Steven C Constable, Robert L Parker, and Catherine G Constable. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52(3):289–300, 1987.

[18] Francesco De Zan, Mariantonietta Zonno, and Paco Lopez-Dekker. Phase inconsistencies and multiple scattering in sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 53(12):6608–6616, 2015.

[19] Cheryl A Dieter, Molly A Maupin, Rodney R Caldwell, Melissa A Harris, Tamara I Ivahnenko, John K Lovelace, Nancy L Barber, and Kristin S Linsey. *Estimated use of water in the United States in 2015*. Number 1441. US Geological Survey, 2018.

[20] Xiao-li Ding, Zhi-wei Li, Jian-jun Zhu, Guang-cai Feng, and Jiang-ping Long. Atmospheric effects on insar measurements and their mitigation. *Sensors*, 8(9):5426–5448, 2008.

[21] TR Emardson, M Simons, and FH Webb. Neutral atmospheric delay in interferometric synthetic aperture radar applications: Statistical description and mitigation. *Journal of Geophysical Research: Solid Earth*, 108(B5), 2003.

[22] Tom G Farr and Zhen Liu. Monitoring subsidence associated with groundwater dynamics in the central valley of california using interferometric radar. *Remote sensing of the terrestrial water cycle*, pages 397–406, 2014.

[23] Wanpeng Feng, Sergey Samsonov, Qiang Qiu, Yuqing Wang, Peizhen Zhang, Tao Li, and Wenjun Zheng. Orthogonal fault rupture and rapid postseismic deformation following 2019 ridgecrest, california, earthquake sequence revealed from geodetic observations. *Geophysical Research Letters*, 47(5):e2019GL086888, 2020.

[24] Yuri Fialko, Yakov Khazan, and Mark Simons. Deformation due to a pressurized horizontal circular crack in an elastic half-space, with applications to volcano geodesy. *Geophysical Journal International*, 146(1):181–190, 2001.

[25] Yuri Fialko and Mark Simons. Deformation and seismicity in the coso geothermal area, inyo county, california: Observations and modeling using satellite radar interferometry. *Journal of Geophysical Research: Solid Earth*, 105(B9):21781–21793, 2000.

[26] Devin L Galloway, Kenneth W Hudnut, SE Ingebritsen, Steven P Phillips, Gilles Peltzer, Francois Rogez, and PA Rosen. Detection of aquifer system compaction and land subsidence using interferometric synthetic aperture radar, antelope valley, mojave desert, california. *Water resources research*, 34(10):2573–2585, 1998.

[27] Giuseppe Gambolati, Pietro Teatini, Domenico Baú, and Massimiliano Ferronato. Importance of poroelastic coupling in dynamically active aquifers of the po river basin, italy. *Water resources research*, 36(9):2443–2459, 2000.

[28] Richard Goldstein. Atmospheric limitations to repeat-track radar interferometry. *Geophysical research letters*, 22(18):2517–2520, 1995.

[29] Mohinder S Grewal and Angus P Andrews. *Kalman filtering: Theory and Practice with MATLAB*. John Wiley & Sons, 2014.

[30] Sara E Grineski, Derek V Mallia, Timothy W Collins, Malcolm Araos, John C Lin, William RL Anderegg, and Kevin Perry. Harmful dust from drying lakes: Preserving great salt lake (usa) water levels decreases ambient dust and racial disparities in population exposure. *One Earth*, 7(6):1056–1067, 2024.

[31] Hongping Gu, Wei Zhang, and Robert Gillies. The shrinking great salt lake may exacerbate droughts by reducing local precipitation: A case study. *Journal of Hydrometeorology*, 25(7):1099–1109, 2024.

[32] Katherine Guns, Xiaohua Xu, Yehuda Bock, and David Sandwell. Gnss-corrected insar displacement time-series spanning the 2019 ridgecrest, ca earthquakes. *Geophysical Journal International*, 230(2):1358–1373, 2022.

[33] Ramon F Hanssen, Tammy M Weckwerth, Howard A Zebker, and Roland Klees. High-resolution water vapor mapping from interferometric radar measurements. *Science*, 283(5406):1297–1299, 1999.

[34] Raymond Franciscus Hanssen. *Atmospheric heterogeneities in ERS tandem SAR interferometry*, volume 3. Delft University Press Delft, 1998.

[35] Charles R Henderson. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, pages 423–447, 1975.

[36] Jörn Hoffmann, Howard A Zebker, Devin L Galloway, and Falk Amelung. Seasonal subsidence and rebound in las vegas valley, nevada, observed by synthetic aperture radar interferometry. *Water Resources Research*, 37(6):1551–1566, 2001.

[37] Kirk Hogenson, Heidi Kristenson, Joseph Kennedy, Andrew Johnston, James Rine, Thomas Logan, Jiang Zhu, Forrest Williams, Jake Herrmann, Jacquelyn Smale, and Franz Meyer. Hybrid pluggable processing pipeline (hyp3): A cloud-native infrastructure for generic processing of sar data, 2026. [Software].

[38] Xie Hu, Zhong Lu, and Teng Wang. Characterization of hydrogeological properties in salt lake valley, utah, using insar. *Journal of Geophysical Research: Earth Surface*, 123(6):1257–1271, 2018.

[39] CE Jacob. On the flow of water in an elastic artesian aquifer. *Eos, Transactions American Geophysical Union*, 21(2):574–586, 1940.

[40] Donald G Jorgensen. *Relationships between basic soils-engineering equations and basic ground-water flow equations*, volume 2064. US Government Printing Office, 1980.

[41] Krishna B Khatri and Courtenay Strong. Climate change, water resources, and potential adaptation strategies in utah. *Salt Lake City, UT: Division of Water Resources, Utah Department of Natural*, 2020.

[42] SA Leake. Interbed storage changes and compaction in models of regional groundwater flow. *Water Resources Research*, 26(9):1939–1950, 1990.

[43] Jonathan M Lees and Huatao Wu. Poisson's ratio and porosity at coso geothermal area, california. *Journal of volcanology and geothermal research*, 95(1-4):157–173, 2000.

[44] Shuiping Li, Gang Chen, Tingye Tao, Ping He, Kaihua Ding, Rong Zou, Jingwei Li, and Qi Wang. The 2019 m w 6.4 and m w 7.1 ridgecrest earthquake sequence in eastern california: Rupture on a conjugate fault structure revealed by gps and insar measurements. *Geophysical Journal International*, 221(3):1651–1666, 2020.

[45] Xiang Li, Alexander L Handwerger, Gilles Peltzer, and Eric Fielding. Exploring the behaviors of initiated progressive failure and slow-moving landslides in los angeles using satellite insar and pixel offset tracking. *Geophysical Research Letters*, 51(13):e2024GL108267, 2024.

[46] Zhiwei Li, Meng Duan, Yunmeng Cao, Minzheng Mu, Xin He, and Jianchao Wei. Mitigation of time-series insar turbulent atmospheric phase noise: A review. *Geodesy and Geodynamics*, 13(2):93–103, 2022.

[47] Rowena B Lohman and Mark Simons. Some thoughts on the use of insar data to constrain models of surface deformation: Noise structure and data downsampling. *Geochemistry, Geophysics, Geosystems*, 6(1), 2005.

[48] Yohai Magen, Alon Ziv, Asaf Inbal, Gidon Baer, and James Hollingsworth. Fault rerupture during the july 2019 ridgecrest earthquake pair from joint slip inversion of insar, optical imagery, and gps. *Bulletin of the Seismological Society of America*, 110(4):1627–1643, 2020.

[49] Brad Marden, Phil Brown, and Thomas Bosteels. Great salt lake artemia: ecosystem functions and services with a global reach. In *Great Salt Lake Biology: A terminal lake in a time of change*, pages 175–237. Springer, 2020.

[50] Michael A McKibben. *The Diversity of Mineral and Energy Resources of Southern California*. Society of Economic Geologists, 1991.

[51] Camilla Medici, Matteo Del Soldato, Gabriele Fibbi, Lorenzo Bini, Pierluigi Confuorto, Gaddo Mannori, Alessandra Mucci, Vania Pellegrineschi, Silvia Bianchini, Federico Raspini, et al. Insar data for detection and modelling of overexploitation-induced subsidence: application in the industrial area of prato (italy). *Scientific Reports*, 14(1):17950, 2024.

[52] Kiyoo Mogi. Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Earthq Res Inst*, 36(99):e134, 1958.

[53] Yusuf Eshqi Molan, Zhong Lu, and Jin-Woo Kim. Influence of the statistical properties of phase and intensity on closure phase. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10):7346–7354, 2020.

[54] Antony Mossop and Paul Segall. Subsidence at the geysers geothermal field, n. california from a comparison of gps and leveling surveys. *Geophysical Research Letters*, 24(14):1839–1842, 1997.

[55] Wesley R Neely, Adrian A Borsa, Jennifer A Burney, Morgan C Levy, Francesca Silverii, and Michelle Sneed. Characterization of groundwater recharge and flow in california's san joaquin valley from insar-observed surface deformation. *Water resources research*, 57(4):e2020WR028451, 2021.

[56] Wesley R Neely, Adrian A Borsa, and Francesca Silverii. Ginsar: A cgps correction for enhanced insar time series. *IEEE Transactions on Geoscience and Remote Sensing*, 58(1):136–146, 2019.

[57] Charles G Oviatt. Chronology of lake bonneville, 30,000 to 10,000 yr bp. *Quaternary Science Reviews*, 110:166–171, 2015.

[58] Sylvain Pasquet, Ludovic Bodet, Amine Dhemaied, Amer Mouhri, Quentin Vitale, Fayçal Rejiba, Nicolas Flipo, and Roger Guérin. Detecting different water table levels in a shallow aquifer with combined p-, surface and sh-wave surveys: Insights from vp/vs or poisson's ratios. *Journal of Applied Geophysics*, 113:38–50, 2015.

[59] Joseph Fairfield Poland, BE Lofgren, RL Ireland, and RG Pugh. Land subsidence in the san joaquin valley, california, as of 1972. Technical report, 1975.

[60] Michael P Poland and Howard A Zebker. Volcano geodesy using insar in 2020: the past and next decades. *Bulletin of Volcanology*, 84(3):27, 2022.

[61] Francis S Riley. Analysis of borehole extensometer data from central california. In *Land subsidence: Proceedings of the Tokyo Symposium*. IASH-UNESCO, 1969.

[62] Steven Roof and Charlie Callagan. The climate of death valley, california. *Bulletin of the American Meteorological Society*, 84(12):1725–1740, 2003.

[63] R Kerry Rowe and Prebaharan Nadarajah. Evaluation of the hydraulic conductivity of aquitards. *Canadian Geotechnical Journal*, 30(5):781–800, 1993.

[64] David A Schmidt and Roland Bürgmann. Time-dependent land uplift and subsidence in the santa clara valley, california, from a large interferometric synthetic aperture radar data set. *Journal of Geophysical Research: Solid Earth*, 108(B9), 2003.

[65] Paul Segall. *Earthquake and volcano deformation*. princeton university press, 2010.

[66] TR Shearer. A numerical model to calculate land subsidence, applied at hangu in china. *Engineering Geology*, 49(2):85–93, 1998.

[67] Dan Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[68] Michelle Sneed and Justin T Brandt. Detection and measurement of land subsidence using global positioning system surveying and interferometric synthetic aperture radar, coachella valley, california, 1996-2005. Technical report, Geological Survey (US), 2007.

[69] US. Geological Survey. Search earthquake catalog. Accessed on Jan 30th 2025.

[70] US. Geological Survey and California Geological Survey. Quaternary fault and fold database for the united states. Accessed on Aug 4th 2025.

[71] Karl Terzaghi. *Principles of Soil Mechanics*. Engineering News Record, New York, 1925.

[72] Daniel T Trugman, Adrian A Borsa, and David T Sandwell. Did stresses from the cerro prieto geothermal field influence the el mayor-cucapah rupture sequence? *Geophysical Research Letters*, 41(24):8767–8774, 2014.

[73] Ekaterina Tymofyeyeva and Yuri Fialko. Mitigation of atmospheric phase delays in insar data, with application to the eastern california shear zone. *Journal of Geophysical Research: Solid Earth*, 120(8):5952–5963, 2015.

[74] UNESCO. *The United Nations World Water Development Report 2022: Groundwater: Making the invisible visible*. United Nations Educational, Scientific and Cultural Organization, Paris, 2022.

[75] Kaihua Wang, Weiping Jiang, Hua Chen, Xiangdong An, Xiaohui Zhou, Peng Yuan, and Qusen Chen. Analysis of seasonal signal in gps short-baseline time series. *Pure and Applied Geophysics*, 175:3485–3509, 2018.

[76] Kang Wang and Roland Bürgmann. Co-and early postseismic deformation due to the 2019 ridgecrest earthquake sequence constrained by sentinel-1 and cosmo-skymed sar data. *Seismological Research Letters*, 91(4):1998–2009, 2020.

[77] A Park Williams, Benjamin I Cook, and Jason E Smerdon. Rapid intensification of the emerging southwestern north american megadrought in 2020–2021. *Nature Climate Change*, 12(3):232–234, 2022.

[78] Simon Williams, Yehuda Bock, and Peng Fang. Integrated satellite interferometry: Tropospheric noise, gps estimates and implications for interferometric synthetic aperture radar products. *Journal of Geophysical Research: Solid Earth*, 103(B11):27051–27067, 1998.

[79] Xiaohua Xu, David T Sandwell, Emilie Klein, and Yehuda Bock. Integrated sentinel-1 insar and gnss time-series along the san andreas fault system. *Journal of Geophysical Research: Solid Earth*, 126(11):e2021JB022579, 2021.

[80] Xiaohua Xu, David T Sandwell, and Bridget Smith-Konter. Coseismic displacements and surface fractures from sentinel-1 insar: 2019 ridgecrest earthquakes. *Seismological Research Letters*, 91(4):1979–1985, 2020.

[81] Chen Yu, Zhenhong Li, and Nigel T Penna. Interferometric synthetic aperture radar atmospheric correction using a gps-based iterative tropospheric decomposition model. *Remote Sensing of Environment*, 204:109–121, 2018.

[82] Gatot Yuliyanto, Tony Yulianto, Udi Harmoko, M Irham Nurwidyanto, and Sugeng Widada. Identification of the presence of groundwater aquifers using poisson's ratio analysis of inversion of microtremor data, a case study of nglekok, asemrudung, geyer, grobogan, central java, indonesia. In *11TH INTERNATIONAL SEMINAR ON NEW PARADIGM AND INNOVATION ON NATURAL SCIENCES AND ITS APPLICATION (ISNPINSA): Developing Science and Technology Toward the New Normal Era*, volume 2738, page 030024. AIP Publishing LLC, 2023.

[83] Howard A Zebker, Paul A Rosen, and Scott Hensley. Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps. *Journal of geophysical research: solid earth*, 102(B4):7547–7563, 1997.

[84] Molly S Zebker, Jingyi Chen, and Marc A Hesse. Robust surface deformation and tropospheric noise characterization from common-reference interferogram subsets. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–14, 2023.

[85] Xinbo Zhang, Yuanying Yang, Huu Hao Ngo, Wenshan Guo, Haitao Wen, Xiao Wang, Jianqing Zhang, and Tianwei Long. A critical review on challenges and trend of ultrapure water production process. *Science of The Total Environment*, 785:147254, 2021.

[86] Xuejun Zhou and Thomas J Burbey. Deformation characteristics of a clayey interbed during fluid injection. *Engineering geology*, 183:185–192, 2014.

[87] Achmad Syaifuddin Zuhri, Udi Harmoko, and Gatot Yuliyanto. Modeling of 3d aquifer layers using poisson's ratio analysis hvsr method in tembelang village, candimulyo district, magelang regency. *Journal of Physics and Its Applications*, 3(1):124–128, 2020.