ON THE USE OF INTERFEROMETRIC SYNTHETIC APERTURE RADAR FOR CHARACTERIZING THE RESPONSE OF RESERVOIRS TO FLUID EXTRACTION AND INJECTION AT WELLS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Karissa Suzanne Pepin
August 2022

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Howard Zebker) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(William Ellsworth)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Rosemary Knight)

Approved for the Stanford University Committee on Graduate Studies

Abstract

Interferometric synthetic aperture radar (InSAR) is a powerful tool used to measure displacements of the Earth's surface over expansive areas (hundreds of kilometers) and with high temporal (6-24 days) and spatial (5-20 m) resolution. These observations allow us to study the Earth's dynamic processes. One important application is the management of natural resources and hazards related to the injection and production of fluids. Ground deformation begins within the reservoir as a response to changes in fluid pressure, and propagates to the surface as either subsidence (ground sinking) or uplift. This deformation can damage infrastructure, produce ground fissures, induce earthquakes, and lead to a permanent loss in groundwater storage capacity or fluid productivity. InSAR can be used to identify these hazards and characterize the reservoir response to well activity, knowledge which can be used to develop tailored management practices. However, our analyses are only as good as the data we work with.

We first identify and characterize an overlooked error source in InSAR and its derived time series: aliasing during 2D phase unwrapping. Phase unwrapping refers to the process of transforming the InSAR measurement – the phase difference between two passes of the satellite at two points in time, wrapped to a 2π ambiguity – to the absolute phase relative to a reference point in the image. Aliasing occurs when we do not adequately sample the true displacement field, resulting in a reconstructed (unwrapped) phase field that is different than the true signal (the real surface displacements plus the noise in our measurements). Our work demonstrates the condition that lead to aliasing, namely, when true phase gradients greater than π rad form a closed loop, and shows that aliasing results in a biased loss in unwrapped magnitude that is proportional to the number of high-gradient ($> \pi$) loops encircling the pixel. Given a displacement field, the occurrence of a high-gradient loop in an interferogram is influenced by the radar system wavelength, spatial resolution, imaging geometry, and the local noise level. Furthermore, spatial filtering – a ubiquitous InSAR practice used to reduce noise – may induce aliasing because it decreases the spatial resolution and, consequently, the physical gradient tolerance.

We then extend our findings on aliasing in a single interferogram to the effects of including aliased images in small-baseline subset (SBAS) time series generation. We find that aliasing has an intimate relationship with the time between image acquisitions forming a given interferogram (the

temporal baseline), as the displacement field evolves in both time and space. Often, aliasing errors increase with increasing temporal baseline, so we observe a systematic decreases in SBAS solution magnitudes and a spatiotemporal distortion of displacement patterns as we increase the maximum temporal baseline included in our calculations. However, Sentinel-1 observations in three study areas (Kilauea Volcano, HI, the Delaware Basin, TX, and California's Central Valley) highlight that some of the best time series solutions with respect to ground-truth (GPS) include long-temporal baseline interferograms and others require their explicit exclusion. This indicates that there is a delicate balance between aliasing and noise reduction as we seek to identify the ideal phase set that best reproduces the true displacements. This phase set not only varies between study areas, but may even be unique to each pixel.

Next we present a case study of reservoir characterization with InSAR in the Delaware Basin, TX, an expansive oil field in the Permian Basin. The motivation for our analysis is a marked increase in earthquake frequency since a revitalization of oil pumping in 2010, using horizontal wells. Our work shows widespread deformation related to oil and gas activity, some of which can be directly related to volume changes from pumping and injection through the correlation of InSAR time series with monthly well volumes. Deformation primarily due to volume change near wells is prevalent in the northern portion basin, where there is a notable absence of earthquake activity. However, the southeastern portion of the basin displays short-wavelength, curvilinear displacement features that do not correlate temporally or spatially with well locations or activity in any obvious way. Rather, these features are spatially correlated with trends in seismicity and the local stress conditions, which indicates that they may be the surface expression of slip on normal faults. We use analytic models of edge dislocations to test this hypothesis in a small study area near the Reeves-Pecos county line. Our final three-fault, patched model reproduces the spatial patterns and magnitude of the main linear feature in both vertical and east-west horizontal displacement components. All three faults are high-angle (75°) and span the Delaware Mountain group (1000-2000 m depth). Two of the faults dip toward one another in a graben structure, matching the characteristics of a recently-identified larger graben network in the surrounding region. The areas of largest fault slip (up to 23 cm) are spatially correlated with nearby disposal wells, linking shallow wastewater injection to the reactivation of pre-existing faults. Though all fault planes contain the largest earthquakes in the study area, our analysis indicates that the fault slip is largely aseismic (stable), which will be an important mechanism to consider in modeling if we are to better understand induced earthquakes in the Delaware Basin.

Acknowledgments

My journey to and through my PhD has been long and indirect. Not only did I find myself on the scenic route, the path was not always well-maintained or clear. I would not have been able to make it to my destination had it not been for the multitude of people who provided light when the road got dark, steered me in the right direction when I took unwise turns, cheered me on from both near and far, and healed my wounds after a stumble. This work is dedicated to them, for it is as much their accomplishment as mine.

I would first like to thank my advisors, Howard Zebker, Bill Ellsworth, and Rosemary Knight. Not only did they provide scientific guidance and wisdom, I will never forget their essential emotional support as I struggled with setbacks and confidence. I would also like to thank them for sharing their unwavering passion for the Earth sciences – their enthusiasm helped bolster me when the going got tough. I am so glad they found some promise in me and took the risk of including me in their research programs. I would also like to thank the other members of my committee, Eric Dunham and Jane Willenbring. They exemplify what it means to be a professor who seeks to drive positive and inclusive change within the Earth Science community, a characteristic I will strive for in the future.

Faculty were not the only ones providing support in my journey. I would not have been able to make it to the end without the members of my primary research group: Yujie Zheng, Roger Michaelides, Stacey Huang, and Elizabeth Wig. Thank you so much for your patience with me and guidance for navigating the world of academia. To Rosemary's environmental geophysics group – Ryan Smith, Meredith Goebel, Ian Gottschalk, Alex Kendrick, Noah Dewar, Aakash Ahamed, and Matt Lees – thank you for including me as one of your own; it was the community I needed in my time at Stanford. Additionally, the COVID-19 pandemic was tough on everyone, but my 403 Tea group with Ying Qi Wong, Ben Mullet Strong, and Matt Lees got me through, and was certainly my silver lining. I cannot thank you enough and look forward to our many future adventures! Finally, I would like to thank my other friends in the School of Earth – particularly Jackson MacFarlane, Paul Summers, Jens Lund Snee, Lauren Abrahams, Rachael Engstrand, and Leighton Watson – for their friendship and support. All of the people mentioned in this section are some of the smartest and kindest people I have been privileged to know. Thank you for continuing to inspire me as a

scientist and human.

My acknowledgements would of course not be complete without a huge thank you to my family and friends who have supported me since the beginning of my journey. To my parents (Karen and Doug), siblings (Annaliese and Erik), in-laws (Grace, Bryan, Donna, Michael, and Madalyn), niece/nephews (June, Jackson, Beckett, and Fred), and far-away friend (Katie): thank you for always believing that I could achieve anything and for being there for me when I need you most (especially when I don't think I do). And finally, to the group I needed most of all. To my wonderful husband, Luke Holten – thank you for being a fantastic chef and neat-freak, forcing me to step away from my computer, making me laugh every day, and loving me always (even if I hadn't passed Quals). You allowed me to achieve this degree and to explore my passion for creating a better community in my department. And, to my cats (of course) – Emma and Margo – thank you for loving me (scientist or not), always being there when I get home, and knowing the perfect moment for a much-needed snuggle break.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Risks of Subsurface Fluid Extraction and Injection

Modern society relies on natural resources that are stored underneath the ground in subsurface reservoirs. Some of the most important of these resources come in fluid form: fresh groundwater, oil, and natural gas. Ingeniously, we have developed ways to extract these resources with minimal impact on the ground surface by using wells to pump fluids from their reservoirs. However, the change in pressure within the subsurface resulting from fluid withdrawal and its opposite – injection or recharge – perturbs the ambient stress field, leads to material deformation within the target reservoir and its surroundings, and alters fluid flow. These responses have consequences in the form of hazards at the ground surface and/or decreased access to the desired resources. Proper resource management requires the development of tailored operational practices that mitigate or avoid such consequences, but a major hurdle in this process is determining the unique subsurface response near each well, because we cannot directly observe processes beneath the ground surface. Therefore, it is important to advance our methods of indirect reservoir characterization because the risks of not doing so can be costly.

For example, one important consequence of fluid extraction and injection (i.e., well activity) is ground deformation, which begins within the reservoir as a response to changes in fluid pressure, and propagates to the surface as either subsidence (ground sinking) or uplift (Gambolati et al., 2005). If deformation within the reservoir is permanent, then there will be an unrecoverable loss of subsurface storage volume, which (in the case of groundwater) decreases our capacity to store freshwater for use in future droughts. Relatedly, the accompanying decrease in pore pressure can cause wells to go dry, severely limiting industry production or vulnerable communities' access to clean drinking water (depending on the resource). Deformation that reaches the surface can damage local infrastructure, including roads, buildings, aqueducts, and wells, which is expensive to repair. Furthermore, ground fissures may form even if infrastructure has been spared.

An additional hazard associated with well activity, particularly in the context of oil and gas (industry) operations, is the reactivation of faults (large subsurface cracks), which may induce earthquakes (Ellsworth, 2013). The pure motion of faults (ignoring the ground shaking) may cause ground deformation and damage to infrastructure, including the wells. But perhaps more importantly, the ground-shaking accompanying induced earthquakes has been associated with catastrophic structural damage and loss of life (see review in Foulger et al., 2018).

Sustainable use of natural resources requires that the associated hazards be well-understood and properly managed. If we could properly characterize the hydrogeologic properties of sediments, identify important structures, and see the movement of fluid and reservoir materials, then we could avoid or mitigate the associated risks by either i) identifying ideal locations to drill wells or ii) developing methods to counteract the negative consequences in non-ideal subsurface environments. Though we cannot do this directly with our eyes, geophysical techniques allow us to indirectly image the subsurface and infer its behavior. With this knowledge, we can develop the proper techniques to better manage our natural resources.

1.2 Approach

Often, the deformation we observe at the surface provides the clues we need to characterize subsurface properties and the reservoir response to well activity. One geophysical method that has been gaining popularity to quantify this deformation is interferometric synthetic aperture radar (InSAR), a microwave remote sensing tool that is capable of measuring cm-scale ground deformation over hundreds of kilometers at a resolution of meters to tens of meters. Traditional geodetic techniques – GPS or extensometers – provide very precise measurements of ground deformation, but they are sparsely-scattered point samples of the overall displacement field, often separated by tens of kilometers or more. InSAR is able to fill the gaps with its expansive coverage, though at decreased temporal resolution.

Many studies have demonstrated the benefits of knowing the overall displacement field. For example, Smith et al. (2017) used InSAR observations to estimate the proportion of unrecoverable reservoir compaction due to the over-pumping of groundwater resources in California's Central Valley. Additionally, a number of studies have derived the evolution of groundwater pressure conditions between wells directly from InSAR measurements (e.g., Chen et al., 2017). Beyond groundwater resources, InSAR has been useful for quantifying changes in pore pressure near induced earthquakes, illuminating the far-reaching effects of wastewater disposal (e.g., Shirzaei et al., 2019; Deng et al., 2020). All of these studies rely on some form of inverse modeling to relate the observed surface displacements to changes in the reservoir in response to well activity.

Despite the established success of InSAR to identify and characterize deforming regions, it is important to note that the inferences we make are limited by the quality of our data and its products (e.g. InSAR time series – the evolution of ground surface elevation over time). Furthermore, we risk misrepresenting the reservoir response to well activity if we restrict our modeling to the processes we *expect* to see, namely, subsidence and uplift driven by changes in pore fluid pressure. With these risks in mind, we can summarize the general workflow for using InSAR to manage an identified hazard or resource:

1. Generate high-quality InSAR data products
2. Predict the geomechanical mechanisms driving observed surface displacements
3. Test hypotheses with modeling
4. Validate inferences with ancillary data
5. Develop sustainable mitigation strategies

This thesis presents recent scientific advances related to each stage of this workflow, with a particular emphasis on steps 1-4.

1.3 Contributions

Our first research portion (Chapters 5 and 6) identifies and describes an underappreciated biased error source in InSAR data products: aliasing during 2D phase unwrapping. Though aliasing itself is a well-known phenomenon first characterized by Nyquist (1928) in his development of sampling theory, its occurrence and behavior in InSAR deformation-mapping has largely been overlooked for many reasons: i) it is challenging to identify the resulting errors; ii) ubiquitous InSAR practices have been passed down with little re-analysis; and iii) (as far as we know) nobody has taken the time to demonstrate the behavior or consequences of aliasing in time series generation. Our work reestablishes the risk of aliasing in InSAR products, characterizes its behavior, and presents possible identification techniques.

The next portion of this thesis (Chapters 7 and 8) presents an InSAR case study in the Delaware Basin, TX, a large oil and gas field in the Permian Basin. Our purpose is to characterize the subsurface response to industry activity in the context of a regional uptick in earthquake frequency since 2010. After developing quality InSAR data products, we characterize the deformation signatures and identify their potential driving mechanisms. Though much of the deformation in the northern portion of the basin is consistent with the behavior of subsurface volume changes, the region with the largest earthquake frequency shows distinct linear displacement features that do not correlate to wells. Using additional geophysical data, we determine that these features may be due to slipping faults. We test this hypothesis with analytic models and validate our findings with local earthquake and pore pressure data. Our results suggest that shallow wastewater injection is reactivating highly-pressurized faults and that a large component of stable (aseismic) slip needs to be included in future numerical models of this process.

The aforementioned studies encompass our discussion of reservoir characterization with InSAR. However, our workflow for hazard management would not be complete without an example of the development of mitigation strategies. We therefore present a method to aid groundwater management in California's Central Valley in Appendix E. InSAR has shown that this region is rapidly subsiding due the overpumping of groundwater resources. These processes can be offset (in part) through the reintroduction of fluids into the subsurface (recharge). Our study presents a new method to efficiently identify sites for recharge with electromagnetic resistivity imaging.

We summarize the contributions of this thesis in the following:

1. Identification of an over-looked, biased error source in InSAR: Aliasing during 2D phase unwrapping
2. Mathematical description of the conditions that lead to aliasing when using a common unwrapping algorithm, SNAPHU (Chen, 2001)
3. Demonstration of common InSAR practices that increase the likelihood of aliasing
4. Demonstration of how aliasing affects InSAR time series in time and space
5. Development of two-component InSAR deformation time series in the Delaware Basin, TX, in the presence of high-magnitude atmospheric noise
6. Identification of shallow aseismic slip in response to wastewater disposal in the Delaware Basin, TX
7. Development of a technique for the selection of sites for surface-spreading recharge with high-resolution electromagnetic imaging methods

1.4 Thesis Roadmap

The following chapters in this thesis are briefly summarized below.

Chapter 2 introduces the response of subsurface materials to fluid extraction and injection, or more specifically, the changes in effective stress as a result of anthropogenic well activity. We focus primarily on the sediment responses that lead to surface deformation, which we can measure with InSAR, and the hydrogeologic properties of subsurface materials that control fluid flow.

Chapter 3 describes the remote sensing technique used in this thesis: interferometric synthetic aperture radar. We heavily emphasize a few characteristics that are important to understanding aliasing: image resolution and time series techniques.

Chapter 4 presents an overview of the history and theory of two-dimensional (2D) phase unwrapping algorithms, culminating in the gold-standard algorithm known as SNAPHU (Chen, 2001). We focus on the objective and cost functions used in this unwrapping technique, leaving the exact solving method for the reader to pursue on one's own. SNAPHU is extensively used by the InSAR community, yet it has not been significantly updated since it was developed in 2002. We conclude with a description of its potential shortcomings in the context of modern InSAR applications.

Chapter 5 describes aliasing during 2D phase unwrapping. We begin with its occurrence in individual interferograms and show that, given a deformation field, the occurrence of aliasing will depend on the wavelength of the system, the resolution of the image, and the effects of noise.

Chapter 6 demonstrates how aliasing errors affect the development of InSAR time series and challenge the use of an important parameter in SBAS methods: the selection of a maximum temporal baseline. This work relies on synthetic interferograms and real data over three study areas: Kilauea Volcano, HI; the Delaware Basin, TX; and the Central Valley, CA.

Chapter 7 begins our demonstration of the use of InSAR to characterize the structure and behavior of the subsurface. In this chapter, we develop deformation time series over the Delaware Basin, TX to infer potential reservoir responses to fluid extraction and injection. We predict that some of the InSAR observations can be explained by induced slip on normal faults.

Chapter 8 tests our hypothesis of fault slip with analytic models. Our results suggest that earthquakes and stable slip along previously un-mapped faults are being induced by wastewater disposal in shallow formations. This is one of the first direct observations supporting the occurrence of stable fault slip in induced earthquakes.

Chapter 9 concludes our discussion of InSAR techniques for the characterization and management of subsurface reservoirs. We first summarize the key contributions, then look forward to future developments based on the results of our work.

Appendix E presents a new method to select sites for managed groundwater recharge with electromagnetic resistivity imaging.

Chapter 2

Scientific Background

This chapter provides a brief background of how reservoir materials respond to fluid extraction and injection. As a general rule, fluid extraction leads to compaction within the subsurface, which may manifest as surface subsidence (ground sinking). Similarly, fluid injection leads to expansion, which may propagate to the surface as uplift. In addition, well activity may cause far-reaching pressure changes that can directly and indirectly impact fault behavior and induce seismic and aseismic events. These processes are all responses to changes in effective stress, or the load experienced by the skeletal matrix. We begin with a discussion of effective stress and its relation to reservoir properties and deformation. We then provide an overview of how the reservoir response, controlled by its material properties, may manifest as displacement at the surface, first with a focus on volumetric changes from the extraction or injection of fluids, then with a discussion of the geomechanics of induced fault motion.

2.1 Effective Stress and Fluid Flow

2.1.1 Effective Stress

Compaction, uplift, and slip along faults are fundamentally controlled by changes in effective stress, which is the load carried by a porous matrix (Figure 2.1). Effective stress (σ') can be expressed in terms of the total stress (σ) and pore fluid pressure (p) (Terzaghi, 1923/1925, as summarized in Gambolati et al., 2005):

$$\sigma' = \sigma - p \tag{2.1}$$

Either a change in the total stress or a change in pore pressure can alter the effective stress, but because a Newtonian fluid is unable to sustain shear stress, only the effective normal stress (xx, yy, zz) can be directly influenced by a change in pore fluid pressure, as highlighted in the tensor version

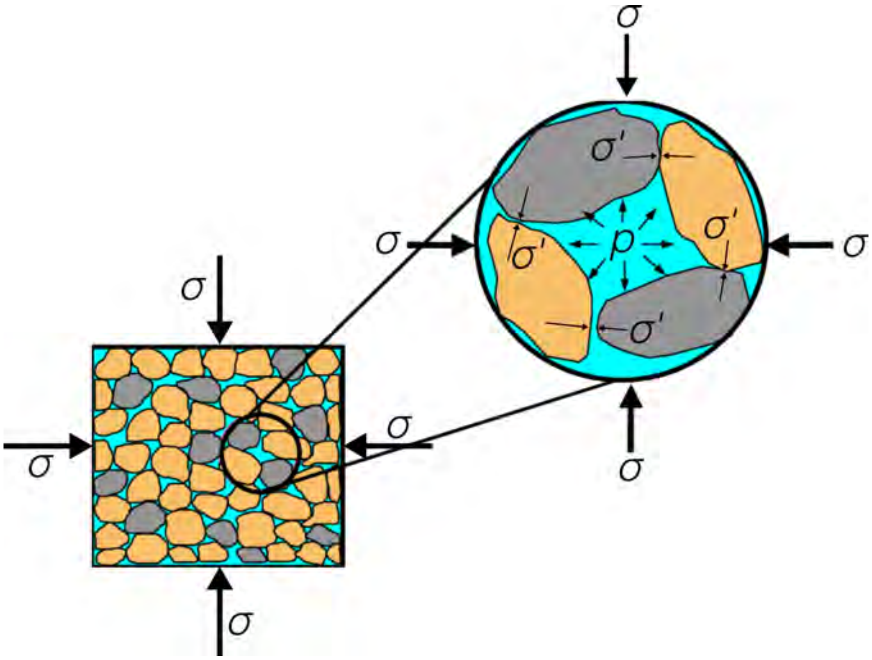


Figure 2.1: The effective stress (σ') experienced by a porous medium is a combination of total stress (σ) and pore fluid pressure (p), where $\sigma' = \sigma - p$. Effective stress is not strictly the stress of the grain-to-grain contacts, but instead is the distribution of load carried by the grains in the area considered.

of Eq. 2.1:

$$\begin{bmatrix} \sigma'_{xx} & \sigma'_{xy} & \sigma'_{xz} \\ \sigma'_{yx} & \sigma'_{yy} & \sigma'_{yz} \\ \sigma'_{zx} & \sigma'_{zy} & \sigma'_{zz} \end{bmatrix} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} - \begin{bmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{bmatrix} \tag{2.2}$$

It is easy to see that the off-diagonal elements in σ' can only be affected when the components of shear stress in σ change. However, pore pressure diffusion also affects the total stress tensor when poroelasticity is included in the model. In his theory of poroelasticity, Biot (1941) developed fully-coupled equations that describe how a change in pore pressure from fluid flow acts as a body force that alters the total stress, deforming the pore space. Deformation of the pore space, in turn, changes the pore pressure and the cycle continues until equilibrium is reached. Though poroelasticity is a more accurate representation of fluid flow dynamics, Gambolati et al. (1992) argue that at practical time scales (tens of years), a fully-coupled treatment of 3D poroelasticity can be ignored for the study of surface displacements arising from volumetric changes due to fluid withdrawal and injection. In these cases, we can characterize fluid flow and subsurface pressure distributions with an uncoupled diffusion equation, described next.

2.1.2 The Diffusion Equation

The standard diffusion equation governing 3D transient (time-varying) fluid flow in a confined, fully-saturated isotropic reservoir is presented in Eq. 2.3:

$$\nabla^2 h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 h}{\partial z^2} = \frac{S_s}{K}\frac{\partial h}{\partial t} \tag{2.3}$$

In this formulation, the pressure diffusion is expressed in terms of hydraulic head (h), which is defined as the potentiometric surface level of a fluid relative to a pre-defined datum. A change in hydraulic head (Δh) is directly proportional to a change in pore fluid pressure (Δp) by:

$$\Delta p = \rho g \Delta h, \tag{2.4}$$

where ρ is the fluid density and g is the acceleration due to gravity. The right side of Eq. 2.3 shows that the spatial distribution of pressure change is controlled by the specific storage (S_s), or the volume of fluid produced per unit volume of the aquifer system per unit change in head (Jacob, 1940), in addition to the hydraulic conductivity (K), the ease with which a fluid flows through a porous medium. Relating Eq. 2.4 to Eq. 2.1, a decrease in hydraulic head with time decreases the pore pressure and increases effective stress. Therefore, we can think of S_s as the response of the matrix and fluid due to an increase in effective stress (or drop in head). We first decompose S_s into its controlling variables:

$$S_s = \rho g(\alpha - \eta\beta), \tag{2.5}$$

where α is the matrix or skeletal compressibility, β is the compressibility of the fluid, and η is the porosity of the geologic unit. As head drops, the matrix compresses while the fluid expands; however, the compressibility of the matrix is generally much larger than that of the fluid ($\alpha >> \beta$), so for many practical purposes $S_s \approx \rho g \alpha = S_{sk}$. The skeletal specific storage (S_{sk}) describes the change in stored fluid that will result from deformation of the reservoir material in response to a change in head. Thus, a change in stored fluid, via pumping or injection, inherently implies deformation of the formation. The value of α determines the magnitude of deformation and is defined as the ratio of strain to a change in effective stress. In the vertical component, for example:

$$\alpha = \frac{\frac{\Delta b}{b_0}}{\Delta \sigma'_{zz}}. \tag{2.6}$$

Here, b_0 is the original thickness of the deforming unit and Δb is the change in thickness due to a change in vertical effective stress $\Delta \sigma'_{zz}$. If we are able to measure surface displacements (as we can with InSAR) and we have a formula to relate these measurements to Δb, then we have the means to begin characterizing parameters of the geologic model, such as S_s, and/or predict pressure changes due to pumping and injection.

2.2 Surface Displacement due to Volumetric Changes

We term a non-zero Δb, the change in thickness of reservoir materials due to changes in effective stress from pumping and injection, as a volumetric change. Since we would like to distinguish the resulting surface deformation due to volumetric changes from other processes like fault motion, we need a solid understanding of what controls its magnitude, spatial expression, and timing at the surface. The magnitude of reservoir deformation and its surface expression depends on the subsurface material's elastic properties, the boundary and initial stress conditions, and the volume of fluid produced (or injected). As we will see, surface displacement in space and time resulting from groundwater pumping and recharge will often behave differently to that of oil and gas activity, in large part due to the difference in reservoir depths and seasonal operations. We refer to Figures 2.2 and 2.3 to visually represent these concepts as we describe them in the following sections.

2.2.1 Material Compressibility

According to section 2.1, the magnitude of deformation within the compacting reservoir due to a change in pressure is a function of a material's elastic properties, with compressibility being one of the most important factors. Clay is up to two orders of magnitude more compressible than sand at shallow depths due to little prior consolidation of sediments (Chilingar & Knight, 1960). Therefore, subsidence due to groundwater pumping can largely be attributed to the compaction of, or rearrangement of grains within, clay units (aquitards), as depicted in Figure 2.2. The shallow depths of many aquifer systems (< 1 km) allows for a 1D approximation of 3D processes due to small overburden stresses (Gambolati et al., 2005). In this representation, the compaction of the aquitards directly propagates to the surface such that the integrated change in thickness of the underlying clays is approximately equal to the measured vertical subsidence. The same cannot be done for studies involving deformation due to oil and gas activity.

Oil and gas reservoirs are generally much deeper and more consolidated than groundwater systems. As the degree of consolidation increases, the compressibility of clay approaches that of sand and the difference between the two materials becomes negligible (Poland, 1984). In addition, stronger overburden effects means that settlement above oil and gas reservoirs is small in comparison to the compaction of sediments (Figure 2.3). Therefore, if subsidence is observed due to oil and gas production, considerable pore pressure declines or a large thickness of the deforming unit are expected (Gambolati et al., 2005). An additional contrast to groundwater systems is that the larger overburden effects associated with the depths of oil and gas activity lead to a wider footprint of subsidence at the surface in relation to the spatial extent of the deforming subsurface units, as highlighted in Figure 2.3. Therefore, we cannot expect to treat the deformation of oil and gas reservoirs as 1D systems.

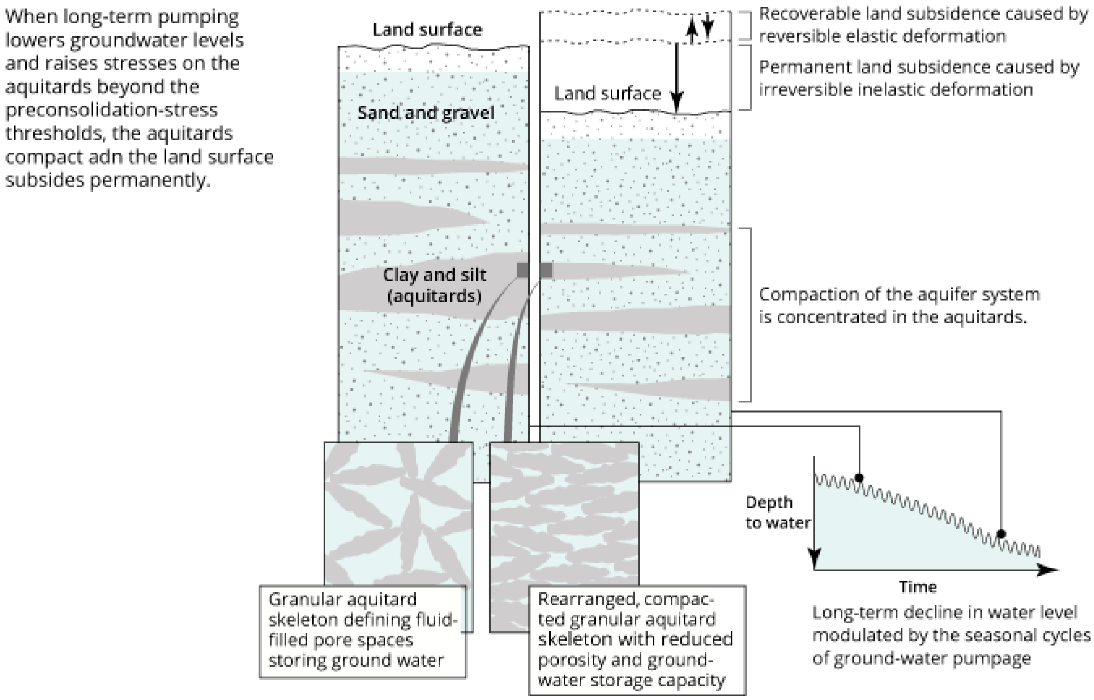


Figure 2.2: Land subsidence due to groundwater withdrawal is concentrated in the fine-grained layers (aquitards) within the confined aquifer system and it often treated as a 1D process. Deformation has both elastic and inelastic components if the pre-consolidation stress has been exceeded. Figure retrieved from the USGS Texas Water Science Center.

2.2.2 Confined vs. Unconfined Reservoirs

An unconfined reservoir is one in which the fluid is directly connected to atmospheric pressure, thus the hydraulic head is identical to the level of water in the sediments. In contrast, a confined reservoir has an impermeable overlying layer that separates the fluid from atmospheric conditions. Confined reservoirs usually have hydraulic head levels above the height of the pore-filling fluids. Because oil and gas reservoirs are often deep and their formation depends on confining layers, the term *unconfined* is usually associated with groundwater aquifers.

Deformation induced by fluid withdrawal is primarily associated with confined reservoirs because the head levels in these units change considerably more due to pumping than in unconfined systems. In addition, when fluid is produced from unconfined aquifers, the overburden load is reduced due to dropping water levels. Consequently, the net change in effective stress and subsequent deformation is not considerable. Therefore, if groundwater pumping is exclusively from unconfined units, little-to-no deformation is expected.

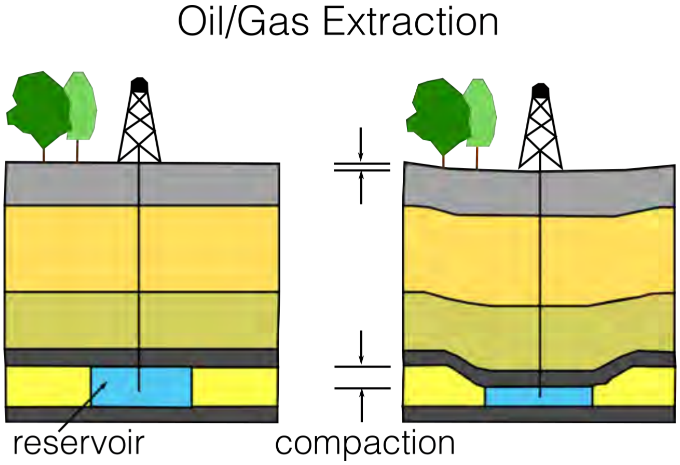


Figure 2.3: Oil and gas extraction is concentrated in deep, confined layers. Due to overburden effects, the displacement observed at the surface is typically not equal to the magnitude of compaction in the reservoir. In addition, the footprint of surface subsidence is wider than the compacting unit.

2.2.3 Inelasatic vs. Elastic Deformation

An additional control on the compressibility of subsurface materials, the rate at which they deform, and the recoverability of deformation, is the relation of the current effective stress to the reservoir's preconsolidation stress, the maximum effective stress that the system has experienced in its history since deposition (Poland et al., 1975; Holzer, 1981). If the current effective stress is less than its historic maximum, then the grains can accommodate the current load and no rearrangement of sediments occur. However, as the effective stress exceeds its historic maximum, the grains rearrange to more stable orientations, which serves as the primary source of compaction.

Compaction from the rearrangement of grains is unrecoverable, or inelastic. Alternatively, recoverable (elastic) deformation occurs when the materials swell or compress from the injection or extraction of fluids, respectively, with no rearrangement. When the pre-consolidation stress has been exceeded, both inelastic and elastic compaction will occur, with the largest magnitude attributed to the unrecoverable component, since inelastic compaction has compressibility factors one- to two orders of magnitude greater than that of elastic deformation (Sneed, 2001).

Whether the compaction is elastic or inelastic is an important consideration for groundwater managers, since unrecoverable deformation implies a permanent reduction in the volume of water the aquifer can store for future use. Trying to determine the elasticity of deformation using InSAR is an interesting topic in its own right and is gaining popularity as a tool for aquifer characterization (Smith et al., 2017; Smith and Li, 2021). However, for the purposes of this thesis, we wish to highlight that subsidence will include an inelastic component if the preconsolidation stress has been exceeded, in which case ground elevation recovery will be small in comparison to the magnitude of sinking. Conversely, uplift is largely reversible and may subside over time as the increase in pore pressure diffuses through the subsurface.

2.2.4 Time-Dependence

There are a number of factors in addition to the (in)elasticity of deformation that control the temporal behavior of surface displacements due to volumetric changes within the affected reservoir. As we noted in the previous section, some degree of subsidence is recoverable, so if pore pressures increase due to the reintroduction of fluids via direct injection or aquifer recharge we can expect to see a small amount of uplift or reduced rate in subsidence. Seasonal fluctuations in subsidence and uplift in regions of groundwater pumping often coincide with summer agricultural demands and winter recharge (Figure 2.2). These fluctuations would not be expected in displacement time series over industry operations, since production is more-or-less continuous. However, if wastewater disposal or CO_2 sequestration is relatively shallow in comparison to production (or within the producing formation), uplift may be observed.

An additional factor to consider when trying to relate observed surface displacements to specific well activities is that deformation may not be instantaneous. The delay time will be determined by some of the parameters we wish to constrain with our geodetic measurements: unit thickness, specific storage, and hydraulic conductivity. For example, the 1D time constant (t) describing the length of time for a drop in head to propagate through a doubly-draining aquitard is (Riley, 1969):

$$t = S_s \frac{(\frac{b_0}{2})^2}{K_z}, \tag{2.7}$$

where K_z indicates the vertical hydraulic conductivity, and S_s and b_0 are again the specific storage and original thickness of the deforming unit, respectively. If there were only a few wells performing one action (e.g. groundwater pumping), the delay time of subsidence may be quite informative for estimating material parameters (e.g. Chen et al., 2017). However, many interesting studies have multiple operations occurring simultaneously and/or in close proximity, thus this time delay can make it challenging to link surface displacement to subsurface volume changes.

2.2.5 Additional Controls of Volumetric Changes

Time delays are not the only complication when relating surface displacements to volumetric changes within the reservoir; faults and other structures may exhibit considerable control on the deformation observed at the surface (Hernandez-Marin & Burbey, 2012). Faults can serve as impermeable barriers to across-fault flow while acting as permeable conduits for flow along the fault plane. Thus, subsidence will likely be greater on the side of the fault from which fluids are being produced. These observable, non-symmetric surface displacements may occur without any movement of the fault, but fluid flow along the surface might also induce aseismic or seismic slip, which is covered in the next section.

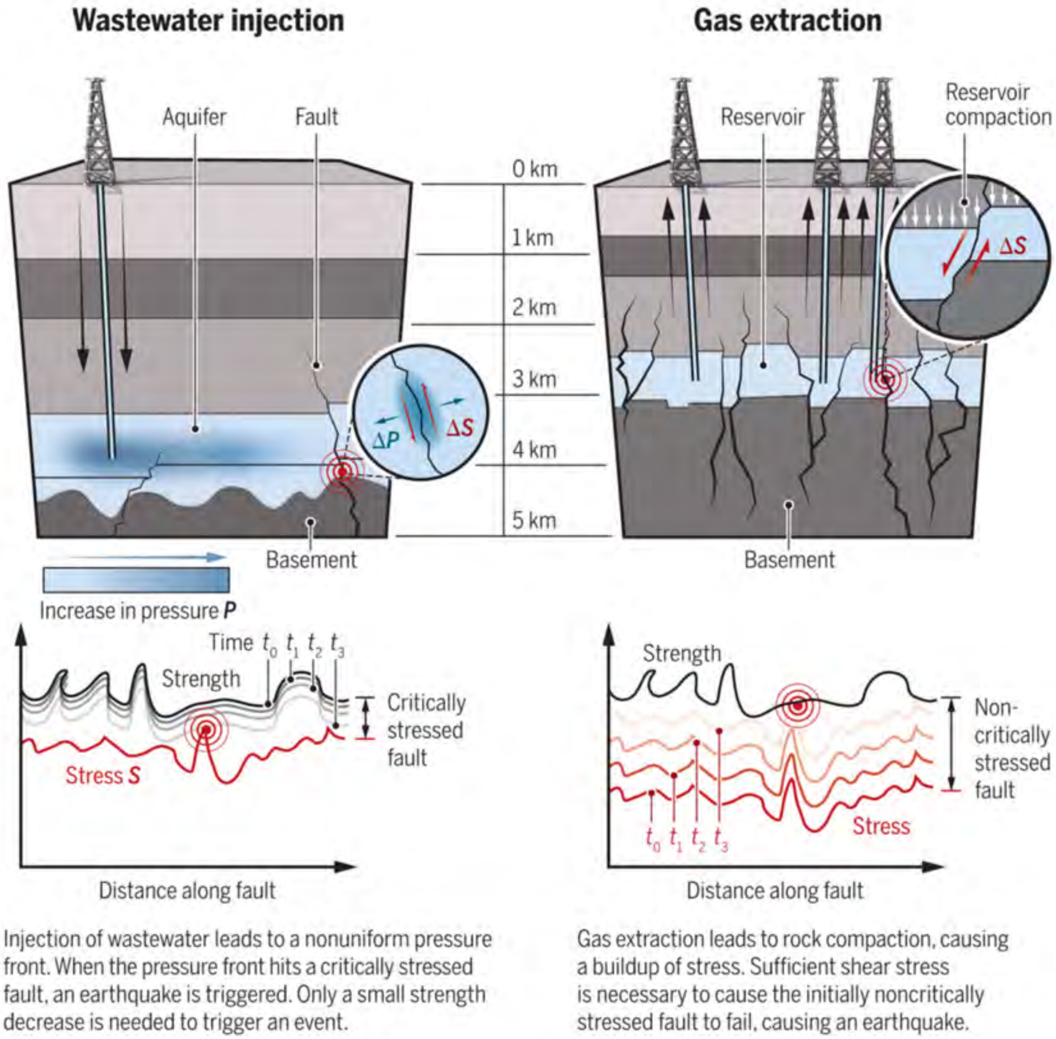
2.3 Induced Seismicity

2.3.1 Effective Stress and Fault Slip

Faults, or cracks in the Earth's crust, form and slip to relieve tectonic and overburden stresses, resulting in earthquakes (seismic events). While earthquakes often occur due to purely natural causes, fault slip may also be induced anthropogenically by the production of oil and gas or the injection of fluids for wastewater disposal, hydraulic fracturing (fracking), CO_2 sequestration, or enhanced oil recovery (Figure 2.4). Among these, the most-widely observed mechanism is wastewater injection (left), in which case increasing pore pressures diffuse directly into pre-existing faults, reducing the fault strength, bringing the fault closer to failure (Ellsworth, 2013). Theory and observations also show that indirect poroelastic stresses (right), either from production or injection, may propagate much further than the extent of pore pressure diffusion and induce slip on critically-stressed faults. Therefore, coupled poroelastic effects are often necessary to include in studies of induced seismicity (Chang and Segall, 2016a; Chang and Segall, 2016b).

The ambient stress conditions and orientation, frictional and hydraulic properties, and location of pre-existing faults relative to the well activity control whether or not seismicity will be induced. We can summarize the stress conditions, pore pressure, and frictional properties of a fault by its Mohr-Coulomb circle in the (τ, σ')-plane, where τ and σ are the shear and effective normal stresses, respectively (Figure 2.5). The total stress acting on a fault is represented by the circle, the diameter of which is controlled by the difference between the minimum and maximum normal stresses (σ_3 and σ_1, respectively). The pore pressure determines the shift of this circle along the σ'-axis. When the circle intersects the failure envelope defined by $\tau = c + \sigma' \tan\theta$ (c is the cohesion and θ is the angle of internal friction), properly-oriented faults will slip due to shear failure. If the circle falls below the intersection point of the failure envelope and σ'-axis, then new tensile fractures will form. The behavior of the Mohr-Coulomb circle mimics Eq. 2.1: when the total stress increases the circle expands and an increase in pore pressure moves the circle to the left, both of which bring the circle closer to the failure envelope.

The fault orientation, described by its azimuth and dip, relative to the ambient stress conditions contribute to its risk of failure under increasing effective stress. Figure 2.6 describes the ideal orientations for various faulting regimes, including normal, reverse or thrust, and strike-slip (Anderson, 1905). In each, the relative magnitudes of horizontal to vertical principal stresses determine the regime. Normal- and reverse faults form parallel and perpendicular to the maximum principal horizontal stress ($S_{H_{max}}$) orientation, respectively, with $\sim 60°$ and $30°$ - $45°$ dips, respectively. Strike-slip faults form at $\sim 30°$ angles to $S_{H_{max}}$ and are often close-to-vertical. The same stress conditions that formed the fault are the same as the stress conditions in which a fault is most favorable to slip again. Therefore, if we know the current total stress conditions in a region, we can predict the orientations of faults that are most susceptible to reactivation.



Injection of wastewater leads to a nonuniform pressure front. When the pressure front hits a critically stressed fault, an earthquake is triggered. Only a small strength decrease is needed to trigger an event.

Gas extraction leads to rock compaction, causing a buildup of stress. Sufficient shear stress is necessary to cause the initially noncritically stressed fault to fail, causing an earthquake.

Figure 2.4: Mechanisms of induced seismicity. Injection (left) or production (right) may induce an earthquake. Figure adapted from Candela et al., 2018.

2.3.2 Deformation Due to Fault Slip

New and slipping faults cause deformation. In geodetic studies, we primarily focus on the expression of this deformation at the surface. For small earthquakes, co- and post- seismic surface displacement is often negligible and will likely not be observed by InSAR measurements. However, if slow or aseismic slip continue to occur along the same fault, motion that builds up slowly over time might be measurable with InSAR observations. If we do measure surface displacements due to slipping faults, the sense of slip and orientation of the fault will determine the patterns of deformation we observe. Pure strike-slip faulting will result in horizontal surface displacements largely parallel to

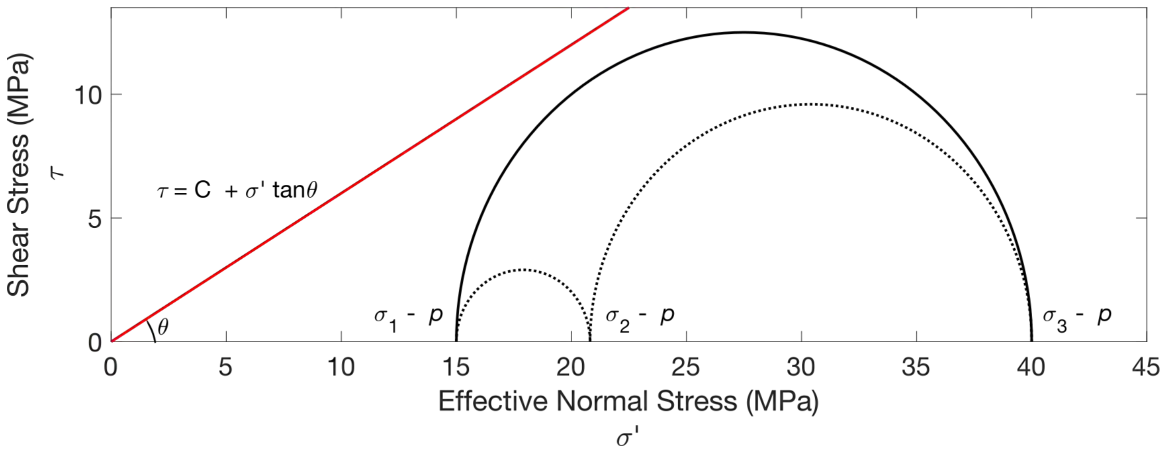


Figure 2.5: The Mohr-Coulomb circle. When pore pressure (p) increases due to fluid injection, the circle defined by the minimum, intermediate, and maximum principal normal stresses (σ_1, σ_2, and σ_3, respectively) moves left toward the failure envelope described by the line $\tau = c + \sigma \tan\theta$, where τ and σ' are the shear and normal stresses, respectively, c is cohesion (zero in this plot), and θ is the angle of internal friction. If the circle meets the failure envelope, properly-oriented faults will slip.

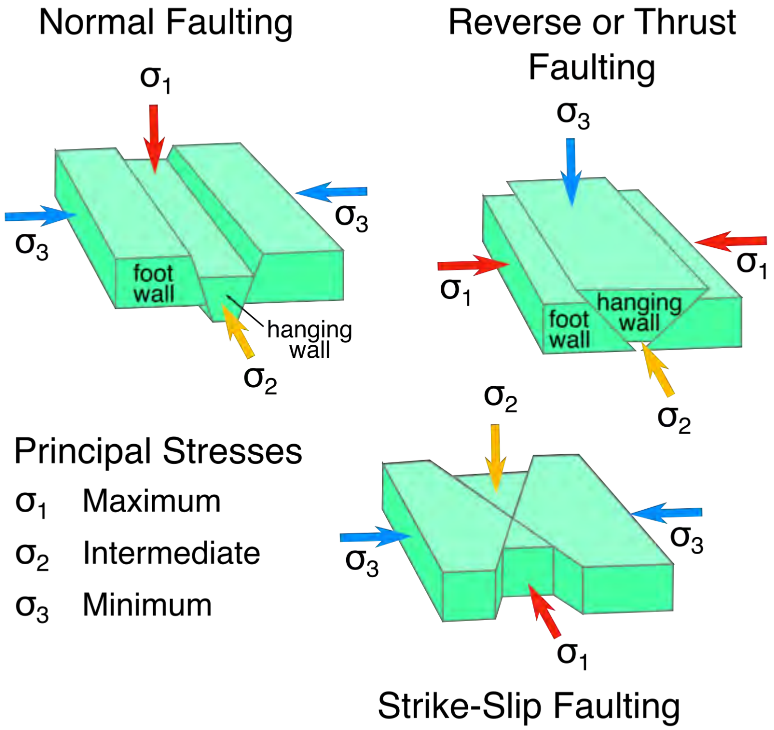


Figure 2.6: The relative magnitudes of the principal stresses (σ_1, σ_2, and σ_3) determine the azimuth and dip of fault formation or slip.

the fault trace. Horizontal displacement will also accompany normal- and reverse faulting, but its magnitude will be small in comparison to the vertical deformation. Normal faulting results in subsidence on the side of the hanging wall and uplift on the side of the foot wall, whereas the opposite will be observed for reverse faults. The wavelength and magnitude of deformation will depend on the fault dip and depth, area and magnitude of the slipping plane, and elastic properties of the subsurface. The goal of geodetic studies in the context of induced seismicity is to constrain these parameters and/or estimate the stress perturbations leading to rupture.

2.4 Groundwater Recharge

Geodetic techniques such as InSAR are often used to identify regions that are unsustainable in their use of groundwater resources. Sustainable use of groundwater requires that water pumped from the subsurface be replenished, or *recharged*, through natural or managed means. Natural recharge occurs when rain, floodwater, streams, lakes, or snow melt percolate through the subsurface to the aquifer. When the volume of these sources is not enough to offset the volume of groundwater pumped and lost via evapotranspiration, then techniques that fall under the umbrella of managed aquifer recharge (MAR) need to be used to maintain a balance between water in- and out of the aquifer system. As we saw in section 2.2, an imbalance may lead to a permanent reduction in groundwater storage capacity or ground deformation that may damage infrastructure. Therefore, in addition to achieving general sustainability, MAR is an important component of groundwater management in the prevention of these hazards.

An overview of MAR techniques can be found in Bouwer, 2002. Here, we highlight the main categories of MAR, which include: 1) injection recharge; 2) vadose zone recharge; and 3) surface spreading recharge. Injection recharge involves the re-injection of water directly into the aquifer via wells. This is often necessary for confined aquifer recharge, due to the impermeability of the confining clay layers that prevent water in the sediments above from reaching the aquifer. The second type of MAR directs the recharge water into the vadose zone, or the unsaturated sediments overlying an unconfined aquifer (Figure 2.7). The water is placed below the root zone in infiltration trenches, shafts, or wells, and it will then percolate through the vadose zone to the unconfined aquifer below. As shown in Figure 2.7, the recharge water will often travel through what are known as preferred pathways, which form along large cracks or other routes dominated by high hydraulic conductivity (Bianchi et al., 2011; Rizzo and de Barros, 2017). Finally, surface-spreading recharge also involves the percolation of water to the aquifer via the vadose zone, but the water is placed at the surface in a flood-plain, field, recharge pond, or ditch. Surface-spreading recharge is thus similar to vadose zone recharge, except the water must additionally flow through the soils at the surface. As these soils contain the root zones of plants, there is a risk of damaging the local vegetation if it is not suited for long-term saturation.

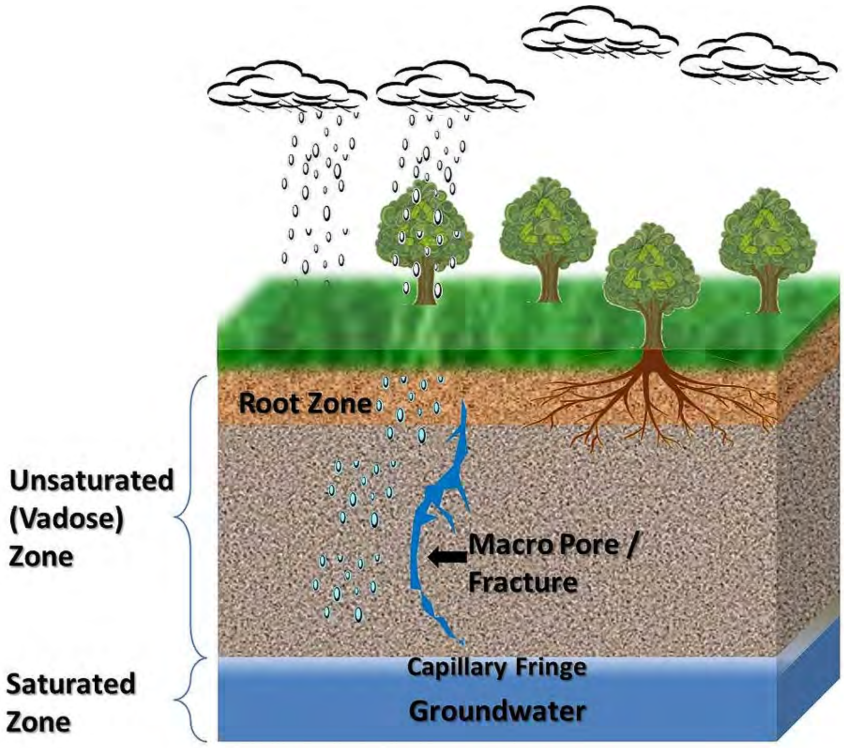


Figure 2.7: The basic anatomy of an unconfined aquifer.

The technique used in a MAR project will depend on the type of aquifer that needs to be recharged and the structure of the subsurface. For example, vadose zone recharge can be selected to augment replenishment of an unconfined aquifer if the overlying vegetation should not remain saturated for long periods of time. However, vadose zone recharge practices are often more expensive and/or destructive than surface-spreading techniques. Therefore, the selection of which method to use will be a calculated balance between cost and risks to the local ecosystem or vegetation. As we will see in Appendix E, geophysical imaging techniques can be used to characterize the structure of the subsurface and predict the behavior of recharge water within the vadose zone, which may be a powerful resource in the selection of appropriate MAR practices.

Chapter 3

Radar Imaging

A detailed history and theory of SAR and InSAR processing can be found in a multitude of texts (e.g. Cumming and Wong, 2004). In this chapter, we highlight only the concepts that will be crucial to the understanding of the work illustrated in this thesis. These topics include the resolution of the SAR image, the components of interferometric phase, the quality of phase measurements, and time series processing techniques.

3.1 Synthetic Aperture Radar

Synthetic aperture radar (SAR) is a form of radar imagery in which a known signal is transmitted from a moving platform (e.g. aircraft or spacecraft) to the ground, and a receiver measures the phase and amplitude of radar backscatter returned from the illuminated area. Figure 3.1 shows a typical imaging geometry of a monostatic system, in which the same antenna is used as the transmitter and receiver. This radar architecture requires a pulsed system, where the transmitted signal has finite duration, or pulse length (τ), and pulse repetition frequency, i.e. how often the pulse is transmitted as the platform travels along its flight path. These two parameters, along with the type of transmitted signal, influence the quality of the resulting image, which is often characterized by the signal-to-noise ratio (SNR) and the resolution in both the across- and along-track directions, or range and azimuth, respectively. While SNR can be related to the strength of the signal returned from scattering elements on the ground relative to background noise, the resolution can be thought of as the minimum distance between scattering elements required for them to be distinguishable in the SAR image.

We first consider the trade-off between SNR and range resolution, as this applies to any active radar system (one which transmits its own signal). A longer pulse length increases SNR by increasing the integrated signal power illuminating scatterers on the ground. Likewise, the range resolution is directly proportional to the pulse length, though in this case, a smaller resolution length is desired.

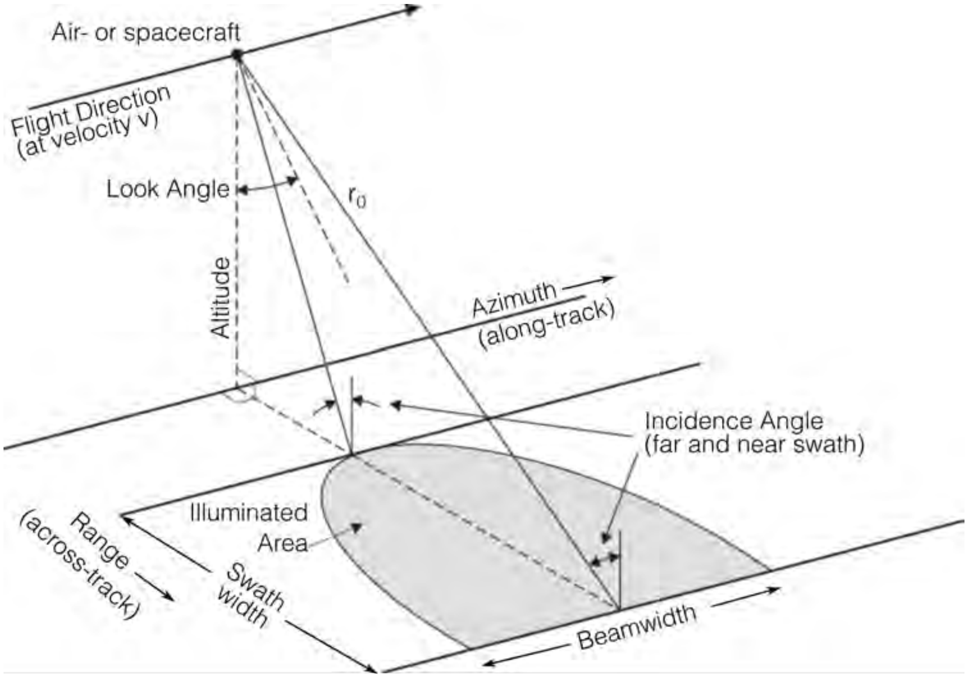


Figure 3.1: Simple schematic of synthetic aperture radar imaging geometry. As the air- or spacecraft travels along the flight path, the radar sends out a series of pulses. Though the illuminated footprint in the along-track (azimuth) direction of a single pulse is quite large (hundreds to thousands of meters), the combination of these overlapping swaths and Doppler compensation in the along-track direction allows us to generate fine resolution in azimuth (5 to tens of meters). Range compression with the known transmitted chirp signal generates fine resolution in the across-track (range) direction (also 5 to tens of meters).

This trade-off can be remedied by a processing technique known as matched filtering or range compression (RC). Here, we transmit a frequency-modulated (FM) chirp signal (a signal that sweeps through a range of frequencies with a linear chirp slope, S_r) and correlate it with the return echo, which is itself a convolution of the transmitted signal with distribution of the ground scatterers. This correlation effectively compresses the received signal into the system's expected impulse-response, which takes the form of a sinc function in the time domain. A simple example is demonstrated in Figure 3.2. This method results in a range resolution (δ_r) now defined by the speed of light (c) and bandwidth (BW) of the FM-signal:

$$\delta_r = \frac{c\tau}{2} \xrightarrow[RC]{} \frac{c}{2BW}. \tag{3.1}$$

Range compression allows us to achieve high SNR via a long pulse length without compromising resolution in range, which is now on the order of meters to tens of meters.

We desire a comparable resolution in the along-track, or azimuth, direction. In real-aperture

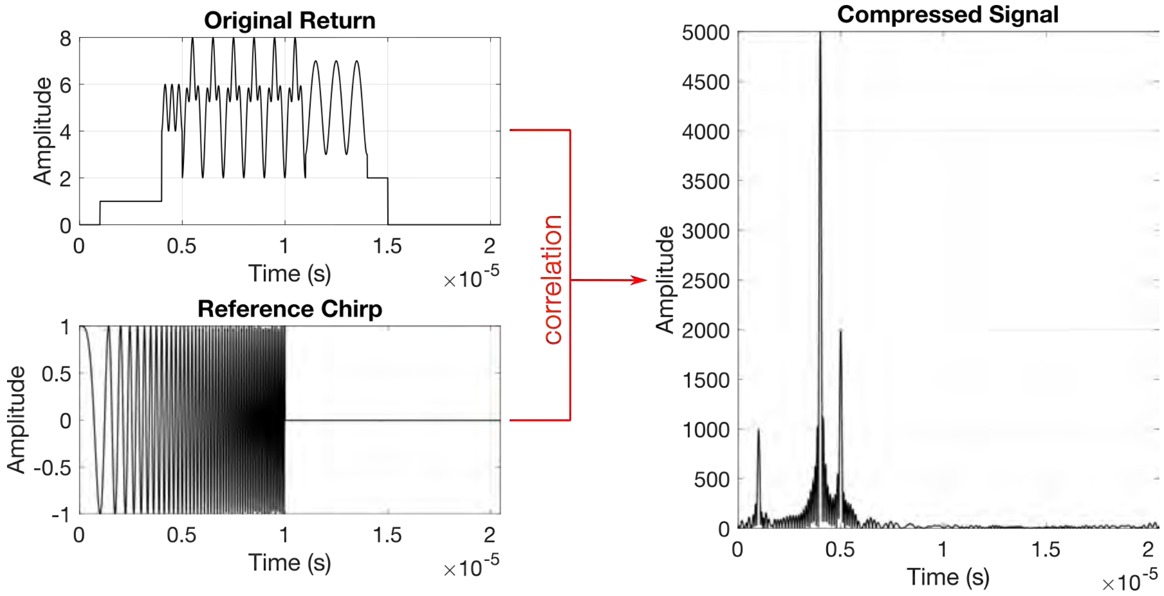


Figure 3.2: The effects of range compression in SAR processing. The returned echo (top left), which consists of no clear scatterers, is correlated with the known, reference chirp signal (lower left). The result is a signal consisting of three distinguishable scattering elements (right).

radar (RAR), the azimuth resolution is equivalent to the along-track width of the imaging swath, or beamwidth (Θ), and is thus dependent on the distance between the radar and the imaged surface, or range (r_0), in addition to the wavelength of the center-frequency of the transmitted chirp signal (λ) and the the length of the antenna (L):

$$\Theta = \frac{r_0 \lambda}{L}. \tag{3.2}$$

As an example, a C-band radar ($\lambda = 5.6$ cm) with a 10 m antenna on-board a satellite operating at 800 km altitude has a beamwidth (or azimuth resolution) of 8 km, which is three orders of magnitude greater than what we can achieve in range and far too coarse for any practical applications.

SAR, in contrast, exploits the movement of the platform and continuous pulsing of the transmitted signal to process the Doppler spectrum of the returned signals in much the same way as range compression. As the radar moves in the along-track direction at velocity v, it transmits its signal at a set pulse repetition frequency, which is engineered to illuminate each point on the ground multiple times, each at a different range. This creates the well-known Doppler shift phenomenon as the satellite passes by, resulting in an effective FM signal in the azimuth direction. Assuming that the radar is imaging directly perpendicular to the along-track direction (equivalent to saying zero squint angle), we can construct the phase history of a point on the ground as a function of time (t) with Eq. 3.3:

$$\phi(t) = -\frac{4\pi}{\lambda}(r_0 + \frac{v^2t^2}{2r_0^2}). \tag{3.3}$$

Here, r_0 represents the distance between the radar and imaged point at the closest approach, and λ is again the wavelength of the center-frequency of the transmitted chirp signal. Quadratic in time, this signal is equivalent to a chirp centered around the Doppler frequency (f_D), with rate (s),

$$S_{az} = -\frac{2v^2}{\lambda r_0}. \tag{3.4}$$

We can now exploit the same matched-filtering technique as we did in range by correlating the return echos in azimuth with this known FM signal. The resulting azimuth resolution (δ_{az}) can be derived from the effective azimuth bandwidth (BW_{az}):

$$BW_{az} = -\frac{2v}{L}. \tag{3.5}$$

Ignoring the negative sign, the azimuth resolution (δ_{az}) is then given by

$$\delta_{az} = \frac{v}{BW_{az}} = \frac{L}{2}. \tag{3.6}$$

Note that δ_{az} is now independent of wavelength and range, and depends solely on the length of the radar antenna, where a smaller antenna leads to finer resolution. This is due to our creation of a synthetically larger antenna by imaging each point many times with varying Doppler shifts (hence the term "synthetic aperture"). A smaller antenna leads to a larger beamwidth, which allows us to image a single point over a longer duration and higher bandwidth, resulting in a finer resolution. Given an antenna length of 10 m, as in our previous example, the azimuth resolution has been reduced to meters, which is comparable to, and often better than, the resolution we can achieve in range.

3.2 Interferometric Synthetic Aperture Radar

A single pass of the satellite results in a SAR image consisting of complex-valued scalars which record the amplitude and phase of the combined radar return reflected from all of the scatterers within each resolution element, or resel (the size of which is defined by δ_r and δ_{az}). We refer to this image as a single-look-complex, or SLC. The phase in each resel is itself a random value between $[-\pi, \pi]$; however, it is deterministic in the sense that if all of the scatters remained stationary between two points in time, two passes of the satellite from the same orbital path would record identical phase values. In this way, SAR is a coherent imaging technique.

We can use the coherent nature of the phases to measure the displacement of the Earth's surface between two points in time through the formation of interferograms. We combine two time-separated

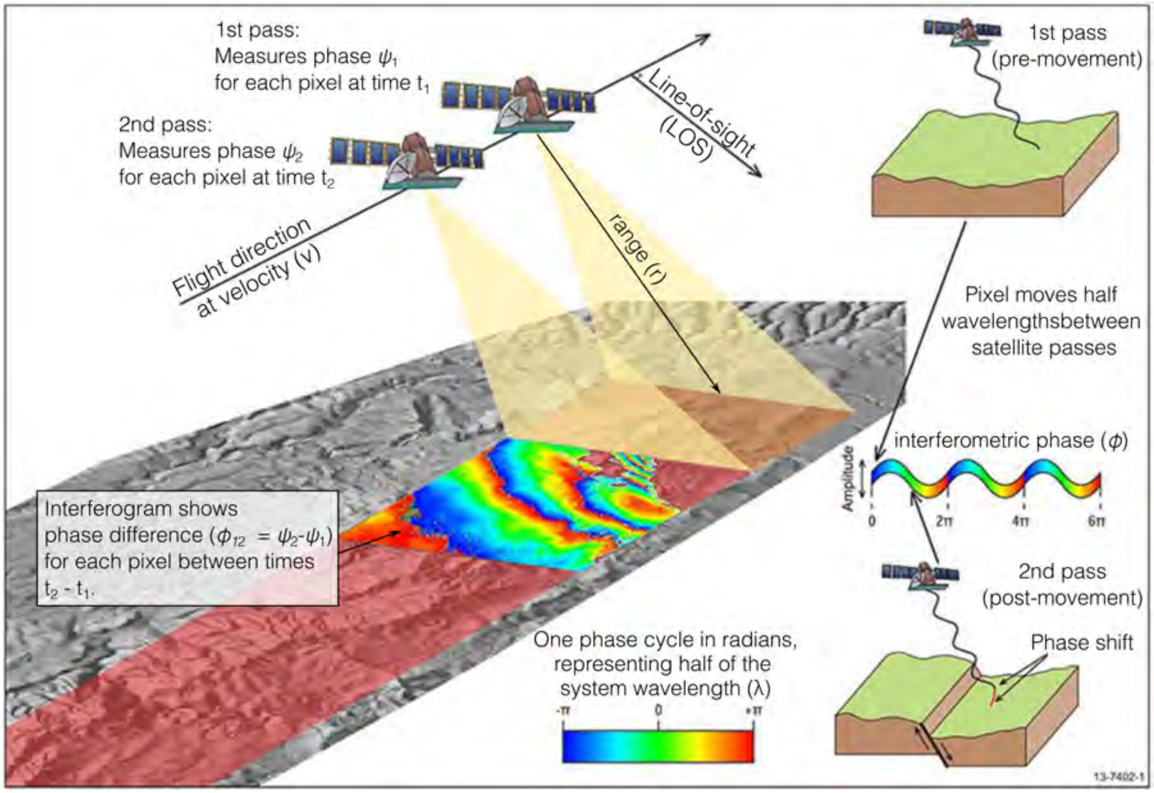


Figure 3.3: InSAR for measuring surface displacement requires two passes of a satellite over a deforming region. The measured phase difference between the two signals can be related to the change in range (in radar line-of-sight) between the two points in time. Each image pixel is approximately the size of a resolution element. Figure modified from Geoscience Australia (no date).

complex SAR signals (Figure 3.3), s_1 and s_2 with phases ψ_1 and ψ_2, respectively, to generate an interferometric complex signal z_{12} via the following relationship:

$$z_{12} = s_1 s_2^* = a_1 e^{-j\psi_1} a_2 e^{j\psi_2} = a_1 a_2 e^{j(\psi_2 - \psi_1)}, \tag{3.7}$$

where * indicates the complex conjugate and $j = \sqrt{-1}$. The phase (ϕ_{12}) of z_{12} is the difference in phase between the two SAR signals:

$$\phi_{12} = \psi_2 - \psi_1. \tag{3.8}$$

This simple calculation contains information about how the surface has changed between time 1 and time 2. In the ideal case, all of the scatterers within a resel remain stationary except for a constant shift of the surface, which corresponds to a non-zero interferometric phase for z_{12}. We can directly relate this phase to a change in range along the vector connecting the satellite and resel, i.e., the radar line-of-sight (LOS):

$$\Delta r = r_2 - r_1 = -\frac{\lambda(\phi_{12})}{4\pi}. \tag{3.9}$$

An image that displays the phase differences between two SAR scenes is called an interferogram, an example of which is shown in Figure 3.4a (amplitude) and b (phase scaled by the average amplitude). Of course, our interpretation of the interferogram is a little more complex than simply applying Eq. 3.9 everywhere and attributing the resulting value to the surface displacement. For example, there are other sources of phase-shifts in addition to pure ground deformation, and it is rarely the case that all of the scatterers within a resel remain stationary between two points in time, which may affect the interferometric signal quality. Furthermore, the cycled coloring, or fringes,

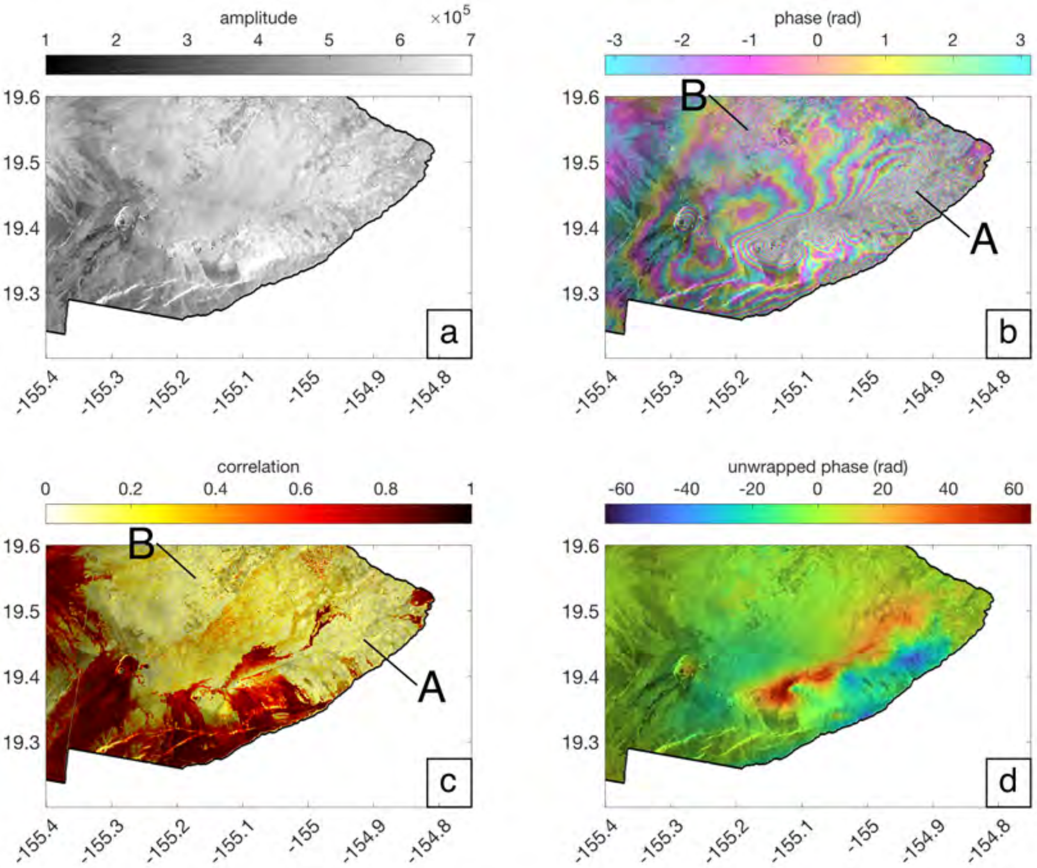


Figure 3.4: Example interferogram over Hawaii in 2018, where (a) shows the average amplitude of interferograms in this study area, (b) displays the interferometric phase values scaled by the average amplitude, (c) depicts the correlation from 24 looks in range and 6 looks in azimuth (approximately square pixels), and (d) is the unwrapped interferogram.

Phase Term	**Name**	**Behavior**
ϕ_d	displacement	correlated in space and time
ϕ_{top}	topography	systematic
ϕ_B	spatial baseline	systematic
ϕ_{atm}	atmospheric	correlated in space, random in time
ϕ_{dec}	decorrelation	random in space and time
ϕ_n	system noise	random in space and time
ϕ_{sys}	orbital/DEM errors	removable/easily identified

Table 3.1: Phase terms from Eqs. 3.10 and 3.11 and their typical behavior.

reflect the fact that our measurements are limited between $[-\pi\ \pi]$, thus there is an inherent 2π ambiguity in every phase. We discuss these limitations in the next two sections.

3.3 Phase Components

The native interferometric phase consists of many signals in addition to the signal-of-interest; in this thesis, the signal-of-interest is the surface displacement in time and space. In order to best isolate this displacement, we must understand how the other components may influence our measurements in each interferogram.

We begin with the interferometric phase measurement at the native posting of the radar, which we term the *measured* phase, ϕ_m. The measured phase consists of multiple components in addition to the desired signal (Eq. 3.10), which in the case of fluid-flow applications is the surface displacement, ϕ_d, in LOS.

$$\phi_m = \phi_d + \phi_{top} + \phi_B + \phi_{atm} + \phi_{dec} + \phi_n. \tag{3.10}$$

Each phase term is described in Table 3.1. The topography (ϕ_{top}) and spatial baseline (ϕ_B) terms are removed during the formation of the interferograms using a reference digital elevation model (DEM) and the measured orbital positions of the satellite. The removal of these terms often includes some small degree of error due to uncertainty in the DEM or orbital measurements. We lump these errors into a systematic error term, ϕ_{sys}, and update the components of our measured phase term:

$$\phi_m = \phi_d + \phi_{atm} + \phi_{dec} + \phi_n + \phi_{sys}. \tag{3.11}$$

We are now left with three additional terms that influence the relationship of the measured phase to the true displacement signal: the atmospheric (ϕ_{atm}), decorrelation (ϕ_{dec}), and thermal noise (ϕ_n) phases. Equation 3.11 describes the phases present in Figure 3.4b, which we use to demonstrate how each component may manifest in an interferogram.

The surface displacement phase, ϕ_d, is the signal-of-interest in this thesis, particularly as it relates to the subsurface response to fluid flow. These processes (expansion, compaction, and fault motion) most-often lead to smoothly varying signals in space, which we refer to as being spatially-correlated. While the wavelength of these displacement features may vary, there will rarely be abrupt changes in the displacement phases or phase gradients (the difference between adjacent resels in the x- and y-dimensions of the two-dimensional [2D] interferogram). An exception to this generalization would be motion along a fault which breaks the surface, in which case there would be a clear, curvilinear discontinuity amid spatially-correlated displacements. Since it is the signal-of-interest, we often refer to the physical surface displacements as the *true signal* and their gradients as the *true displacement gradients*.

An additional term that is correlated in space is the atmospheric phase, ϕ_{atm}. The atmospheric phase is due to a delay in signal propagation as a result of heterogeneous changes in water content and, to a lesser extent, pressure and temperature in the troposphere (lower atmosphere) (Zebker et al., 1997). Although spatially correlated, we often assume that ϕ_{atm} is temporally random, as long as the time between SAR acquisitions is greater than a day. Because of its spatial correlation, the atmospheric phase is difficult to distinguish from surface displacement, though a few techniques for atmospheric correction have been proposed. The simplest method includes stacking (averaging) sets of independent interferograms, effectively reducing the temporally random signals while highlighting the average displacement in the selected time period (Zebker et al., 1997; Emardson et al., 2003). This method is particularly useful for episodic events (e.g. coseismic displacement), but Chen et al. (2017) show that it can also be used in the development of displacement time series for temporally-dynamic systems, such as pumping-induced compaction. Other techniques for atmospheric corrections include the use of meteorological models (e.g. Delacourt et al., 1998), GPS zenith-delay measurements (e.g. Onn and Zebker, 2006), or global atmospheric models (Jolivet et al., 2011; Jolivet et al., 2014) to predict the expected phase delay from the troposphere. Empirically-derived corrections have also proven effective in certain settings, such as removing phase-elevation trends (Beauducel et al., 2000) or adjusting interferograms based on the residuals between InSAR and GPS measurements, assuming the GPS data are an adequate source of ground-truth (Neely et al., 2020). All of these methods are variably-effective and often depend on the validity of our assumptions in each study area, accuracy of our input models, nature of the deformation itself, or resolution and/or density of auxiliary data.

In contrast to the atmospheric phase term, the decorrelation and thermal noise terms (ϕ_{dec} and ϕ_n, respectively) are usually considered to be random in both time and space (Agram & Simons, 2015). Thermal noise is a property of the radar, well-understood, and modeled with Gaussian white-noise statistics. Decorrelation, on the other hand, is a function of the stability of scattering elements within each resel (Zebker and Villasenor, 1992; Just and Bamler, 1994). An example of decorrelation is demonstrated at points A and B in Fig. 3.4b. The fringe patterns at these points are less defined and more speckled than the spatially-correlated regions dominated by the deformation and atmospheric components. Some scatterer instability is purely stochastic, such as wind blowing branches into different orientations or animals and cars moving through the study area. Other processes are more systematic, like changes in water content altering the dielectric properties of the soils or crops that slowly grow with time (De Zan et al., 2014; De Zan et al., 2015; Michaelides et al., 2019). These latter examples challenge the assumption that decorrelation is a purely stochastic process. For the purpose of this thesis, we focus on the effects of decorrelation, as opposed to its phase statistics or mechanistic causes. The next section provides a few more details about the correlation coefficient, which is the main observable we use to quantify and visualize decorrelation.

3.4 Coherence

Decorrelation affects the quality of spatial agreement in phase measurements. We can often improve the quality by spatially averaging the intereferogram, which we refer to in the radar community as *multilooking*. In addition to improving the signal quality, or SNR, multilooking provides the additional benefit of reducing data storage requirements and increasing the efficiency of subsequent processing algorithms and analyses, though this is at the expense of spatial resolution (in other words, our resels increase in size). By multilooking, we are also able to infer the quality of the phase measurement within each resulting image pixel with its estimated coherence ($\hat{\rho}$):

$$\hat{\rho} = \left| \frac{\sum_{k=1}^{N_c} s_{1k} s_{2k}^*}{\sqrt{\sum_{k=1}^{N_c} |s_{1k}|^2 \sum_{k=1}^{N_c} |s_{2k}|^2}} \right|, \tag{3.12}$$

Here, s_1 and s_2 are the measured complex signals, $*$ denotes the complex conjugate, and N_c is the number of independent looks within the averaging window. An independent look is equivalent to the number of averaged resels, which may be different than the number of averaged image pixels. Coherence ranges between [0 1], where $\hat{\rho} = 1$ implies identical measurements and $\hat{\rho} = 0$ suggests pure noise. Thus, coherence describes the degree to which the interferometric phase measurements within the averaging window are similar. The coherence for the interferogram in Figure 3.4b is depicted in Figure 3.4c. Clearly, coherence varies considerably throughout an interferogram. Furthermore, low coherence values often coincide with high speckle characteristics (compare points A and B between Figure 3.4b and c). However, even in some of the low coherence areas, the fringe patterns are still observable, if not well-defined.

There are a few caveats: 1) Eq. 3.12 is an inherently biased estimates of the true coherence (ρ) at low values (Touzi et al., 1999) and 2) the deformation signal can also bias the results, where an increase in fringe rate (or spatial gradients) decreases our estimate of coherence when using Eq. 3.12. (Zebker & Chen, 2005). To overcome the former, Touzi et al. (1999) described the relationship between the estimated coherence and the expected true coherence as a function of N_c. This bias-correction is especially important for applications that directly use coherence as the main observable (e.g. damage detection after an earthquake [Yun et al., 2015]). The latter caveat can be remedied by removing the observed fringe rate from the averaging window prior to applying Eq. 3.12. However, there may be some applications that benefit from the direct use of Eq. 3.12 without the bias correction from Zebker and Chen (2005), which we will discuss in Chapter 5.

Decorrelation can have a significantly negative impact on our interpretation of interferograms. In the best case, it leads to a purely Gaussian noise component centered around the true phase value that is temporally random. This decorrelation noise could be reduced through the same stacking method described for the temporally random atmospheric component (see section 1.3). However, an issue may occur when the decorrelation phase is not purely random, and leads to a phase bias, as is observed with systematic decorrelation processes (De Zan et al., 2014; Michaelides et al., 2019).

We can assess whether a bias in the decorrelation phase may exist by considering the expected observations in correlated interferograms. As shown in Figure 3.5, three SAR scenes acquired at different times t form three interferograms. In theory, the three phases measured by this triplet of interferograms (in their wrapped state) should be consistent, where the sum of ϕ_{12} and ϕ_{23} is identical to ϕ_{13}. Indeed, this is the case for interferograms at their native posting (i.e., no multilooking). However, multilooking introduces a stochastic component to the resulting phases due to our assumption that each pixel in the averaging window is random sample from a phase distribution (Michaelides, 2020). Thus we often observe what we term *non-zero phase closure*, where $\phi_{12} + \phi_{23} \neq \phi_{13}$. Purely stochastic decorrelation will lead to non-zero closure phase randomly distributed around zero-mean. A systematic decorrelation term, say from changes in dielectric properties of the scattering surface, will lead to a non-zero mean distribution.

We introduce the concept of closure phase to demonstrate that inconsistencies between triplet pairs can occur in the multilooked, wrapped interferograms. We will explore how these inconsistencies may affect displacement time series in Chapter 6.

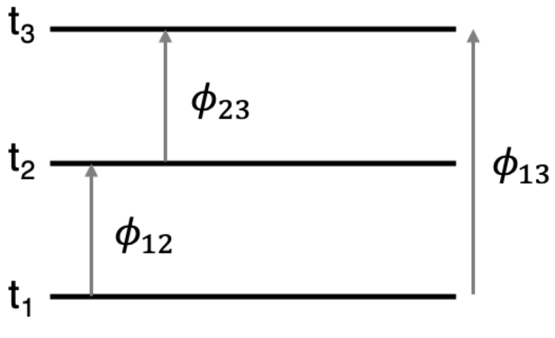


Figure 3.5: The closure phase between three interferograms at time t_1, t_2, and t_3 should be identically zero, or consistent, meaning that the sum of phases ϕ_{12} and ϕ_{23} should equal ϕ_{13}. However, random processes due to decorrelation can lead to non-zero closure phase for multilooked interferograms.

3.5 Phase Unwrapping

Thus far, we have considered only the *wrapped* measured phases, which are limited to values between $[-\pi, \pi]$, even if the true displacement at a point is well beyond those limits. In other words, there is a $2\pi n$-ambiguity to each of our measurements, where n is any integer. These ambiguities make it challenging for the geophysicist to interpret the interferograms, thus we require a method known as phase unwrapping to transform the wrapped phases (ϕ) to their true, unambiguous values, referred to as *unwrapped* phases (ψ). Figure 3.4d demonstrates the results of phase unwrapping for our example in Hawaii, which shows that we no longer have fringe patterns in the unwrapped interferogram. Phase unwrapping is a challenging process that has been extensively studied in recent decades, leading to many available algorithms (e.g. Goldstein et al., 1988; Buckland et al., 1995; Cusack et al., 1995; Ghiglia and Romero, 1994; Ghiglia and Romero, 1996; Pritt, 1996; Davidson and Bamler, 1999; Chen and Zebker, 2001; Yan et al., 2019), though only a handful are used in practice. In particular, the Statistical-Cost Network-flow Algorithm for PHase Unwrapping (SNAPHU; Chen and Zebker, 2001) has been adopted as the algorithm of choice in many InSAR processing software packages (ISCE - CalTech, 2022; SNAP - ESA, n.d.; StaMPS - Hooper, n.d.). Of course, phase unwrapping is not always successful and Chapters 4 - 5 are devoted to the causes and consequences of incorrect phase unwrapping. We use this brief section to note that the interpretation of interferograms and/or their derived time series often requires phase unwrapping at some point along the workflow. Thus, there are very few InSAR studies that are not susceptible to the risks resulting from these potential errors.

3.6 Time Series Techniques

A single interferogram is a snap-shot of the instantaneous LOS displacement of the surface between two points in time. These individual images can be illuminating for episodic events such as large earthquakes (e.g. Massonnet et al., 1993), where the majority of deformation (e.g. coseismic displacement) occurs virtually instantaneously. Studies of fluid flow, however, benefit from knowledge of the evolution of surface displacement as a function of both space and time. Thus, we require techniques that transform multiple interferograms into displacement time series. The most commonly-used techniques for this task are the Small BAseline Subset (SBAS) method (Berardino et al., 2002) and Permanent (or persistent) Scatterer (PS-) InSAR (Ferretti et al., 2001), each with various potential modifications. SBAS and PS-InSAR are ubiquitous in the InSAR world, thus it is important to understand how specific noise sources such as unwrapping errors may affect the quality of their results and subsequent interpretations. Both methods require unwrapped interferograms, albeit at different stages of the workflow, thus both are susceptible to this error source. We describe the fundamentals of each method in the following sections.

3.6.1 Small Baseline Subset Time Series

Small baseline subset techniques use many interferograms over a single area to develop InSAR time series (Berardino et al., 2002). The total number of interferograms (M) we can generate from N SAR scenes is:

$$M = \frac{N(N-1)}{2}. \tag{3.13}$$

This complete set of interferograms contains both good and bad measurements, most often related to the degree of decorrelation. The fundamental idea in SBAS is to reduce the total population of interferograms to smaller subsets with short spatial- and temporal baselines, which in theory will exclude extremely decorrelated measurements from the time series calculation. The spatial baseline refers to the distance between the satellite during the two SAR acquisitions forming the interferogram. Modern satellites such as Sentinel-1 have extremely tight control on the orbital paths such that the spatial baseline does not typically exceed 200 m (the orbit is controlled to a path with ~100 m radius, well below the critical baseline of ~5000 m, where the signals will be completely decorrelated). Therefore, selecting subsets with small spatial baselines is of little concern for modern satellite missions. The temporal baseline (TB) refers to the number of days between the two SAR scenes forming the interferogram. Since decorrelation tends to increase with increasing TB, the selection of subsets with exclusively short TB in theory reduces decorrelation noise. In practice, we select a maximum temporal baseline (TB_{max}) and use all interferograms with $TB \leq TB_{max}$ in our subset. The selection of an appropriate TB_{max} will be discussed in detail in Chapter 5.

Once we have selected our subset of interferograms, we solve for the linear velocity between SAR scenes at each pixel using the relationship between unwrapped phase measurements and time. For each pixel, this linear relationship can be described by Eq. 3.14:

$$\begin{bmatrix} \Phi_{12} \\ \Phi_{13} \\ \Phi_{23} \\ ... \\ \Phi_{(N-1)N} \end{bmatrix} = \begin{bmatrix} TB_{12} & 0 & 0 & ... \\ TB_{12} & TB_{23} & 0 & ... \\ 0 & TB_{23} & 0 & ... \\ ... & ... & ... & ... \\ 0 & 0 & ... & TB_{(N-1)N} \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{23} \\ v_{34} \\ ... \\ v_{(N-1)N} \end{bmatrix} \tag{3.14}$$

where Φ_{ij} is the unwrapped phase measurement between times t_i and t_j ($i = 1, 2, ..., N-1$; $j = 2, 3, ..., N$; $i < j$; and $(t_j - t_i) \leq TB_{max} \leq t_N - t_1$), $TB_{i(i+1)}$ represents the temporal baseline between adjacent SAR acquisitions ($t_{i+1} - t_i$), and $v_{i(i+1)}$ is the LOS displacement velocity between adjacent SAR acquisitions. Adopting the nomenclature from Aster et al. (2013), we reduce this matrix equation to a condensed linear relationship:

$$d = Gm, \tag{3.15}$$

where d is the vector of unwrapped phases, G is the relation-matrix of temporal baselines, and m is the vector of modeled velocities. Since G is often near-singular and/or not square, we solve for m using the generalized inverse solution of G ($G^{\dagger}$) derived from its singular-value-decomposition (SVD):

$$G = U\Sigma W^{T}. \tag{3.16}$$

Here, if rank$(G) = r$, then U consists of the r left singular vectors of G, W is a collection of the r right singular vectors of G, and Σ is a diagonal matrix containing the r singular values of G along its diagonal and zeros elsewhere. The Moore-Penrose pseudo-inverse of G is then:

$$G^{\dagger} = W\Sigma^{-1}U^{T}, \tag{3.17}$$

and we can estimate m through inversion:

$$\hat{m} = G^{\dagger}d, \tag{3.18}$$

where $\hat{m}$ is our modeled estimate of the true velocity vector. Now that we have an estimate of the linear displacement velocities between each SAR acquisition, we generate the time series through integration, assuming zero displacement at t_1.

3.6.2 Persistent Scatterer Techniques

Persistent scatterer InSAR (PS-InSAR) is an alternative method of conducting InSAR time-series analysis that optimizes for returns from strong scatterers as opposed to weaker distributed scatterers in SBAS (Ferretti et al., 2001; Osmanoğlu et al., 2016). Distributed scattering is characterized by the summation of many weak, random phase returns from scatterers within a single resolution element, such that the expected acquired phase in the resel can be represented statistically by a uniform phase distribution between [-π, π]. On the other hand, a strong scatterer is one which dominates the phase return such that the phase statistics more closely resemble a normal distribution. The standard deviation of this distribution will decrease with increasing strength of the dominant scatterer. Distributed and strong scatterers are represented visually in Figure 3.6.

In PS-InSAR, we consider a set of independent interferograms, which is usually the set of all interferograms formed using a single (primary) SAR acquisition (e.g. [ϕ_{12} ϕ_{13} ... ϕ_{1N}]). Using the non-multilooked, wrapped interferograms, we classify the pixels as either strong scatterers (i.e., the PS), which are the most phase-stable points across the set, or distributed scatterers. Assuming that the PS points have the highest-quality measurements, we mask out the distributed scatters and use only the PS pixels to unwrap the interferograms. Since most unwrapping algorithms require completely-filled interferograms, we often use interpolation methods to fill-in the masked portions of

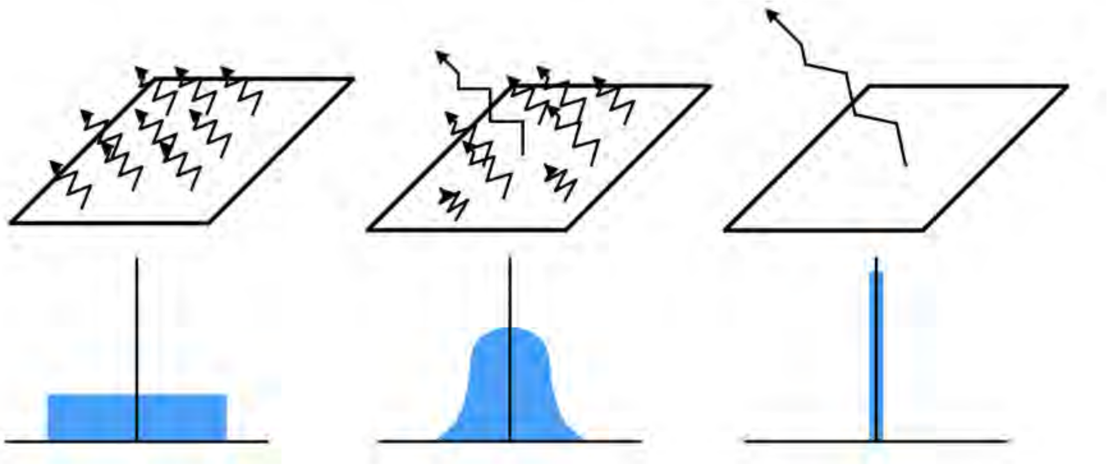


Figure 3.6: Scattering mechanism models (top) and their associated phase PDFs (bottom) as a function of scatterer dominance. Distributed scattering (left) consists of many weak scatterers resulting in a uniform-like phase distribution. As a scatterer becomes stronger and more dominant (middle), the distribution becomes Gaussian-like with narrower standard deviation, until the extreme case of a completely-dominant scatter (right), which has a phase distribution resembling a Dirac-delta. Figure from Michaelides (2020).

the image prior to unwrapping. Once we have our set of unwrapped images, we can easily generate the PS-InSAR time series.

While we use and discuss SBAS solutions in the majority of this thesis, we will briefly comment on why PS-InSAR is not immune to errors arising from phase unwrapping, despite its goals of identifying and using only the highest-quality measurements.

Chapter 4

2D Phase Unwrapping

An important step in virtually all InSAR studies is 2D phase unwrapping, which transforms our ambiguous, wrapped modulo 2π interferometric phase measurements into estimates of the absolute phases relative to each other. This is an inherently challenging task that is only tractable under strict assumptions; however, these assumptions can introduce inconsistencies in the unwrapped solution when compared to the true absolute phase field, which is the focus of Chapters 5 and 6. This chapter provides an overview of the general objectives and assumptions used in phase unwrapping, common algorithms, and an introduction to the potential shortcomings of these methods when applied to modern InSAR applications.

4.1 Adequate Sampling

Phase unwrapping is fundamentally the task of reconstructing a signal from samples. Within this framework, Nyquist theory tells us that we must adequately sample a target signal if we are to reconstruct it without ambiguity, where adequate sampling requires a sampling rate of at least twice the maximum frequency contained within the true signal (Nyquist, 1928). When the sampling rate falls below this threshold, the undersampled high-frequency components will be redistributed as added noise and the reconstructed signal will be different than the original. This phenomenon is formally known as *aliasing*.

A simple example of aliasing is demonstrated in Figure 4.1, which shows an original signal with a frequency of 1 cycle per 2π radians (rad). For this signal to be adequately sampled, we would need to sample at a rate of at least one per π rad (or one per $1/2$ cycle). Because the sampling rate here is lower than one per $1/2$ cycle, aliasing occurs and the reconstructed signal has reduced frequency compared to the original.

In InSAR, the signal we would like to reconstruct from samples is the true unwrapped phase field, at least in terms of the phases relative to each other. Here, we ignore the possibility of a

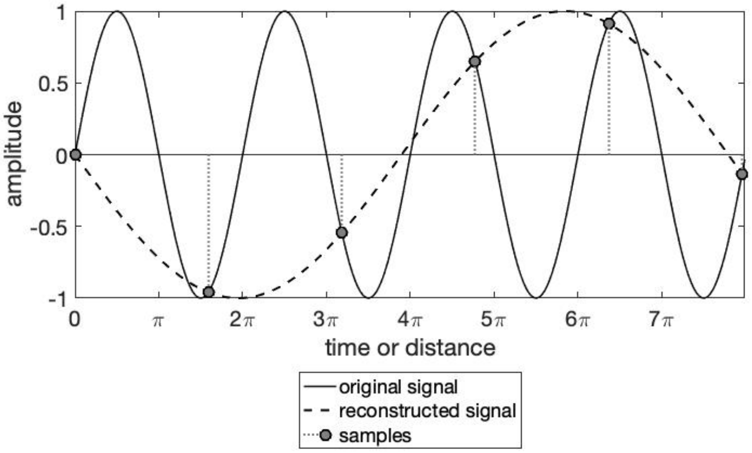


Figure 4.1: An example of aliasing. When we sample a signal at a rate below twice that of the maximum frequency of the true signal, the reconstructed signal will be incorrect.

constant, integer phase-cycle offset that applies to the entire phase field, as this constant is generally not considered part of the phase unwrapping problem. In other words, we care only about the absolute phases relative to each other. If we are to reconstruct this signal perfectly and without ambiguity, then Nyquist theory tells us that the phases must be adequately sampled everywhere in the interferogram. That is, we must have a sample at least every π rad of the true phase field, since the phase cycle in our wrapped measurements is 2π (Spagnolini, 1993). It is convenient to think of this criterion in terms of phase gradients ($\Delta\phi$), or the phase difference between adjacent samples (ϕ) along both dimensions x and y:

$$\Delta\phi^{(x)} = \phi_{i+1,j} - \phi_{i,j} \tag{4.1}$$

$$\Delta\phi^{(y)} = \phi_{i,j+1} - \phi_{i,j} \tag{4.2}$$

For simplicity at this stage, each phase sample $\phi_{i,j}$ is represented as a pixel whose size is equal to the resolution of the image (i.e., 1 pixel = 1 resel). The Nyquist criterion then requires that the magnitude of the true gradient between adjacent pixels not exceed π rad, which can be expressed mathematically as

$$|\Delta\psi| < \pi, \tag{4.3}$$

where ψ represents the true unwrapped phase in any arbitrary dimension. We note that if this holds true, then there is no ambiguity in the measured phase gradients (in contrast to the inherent ambiguity in the wrapped phases themselves), the unwrapped gradients are identical to their wrapped counterparts ($\Delta\psi = \Delta\phi$), and we can generate the unwrapped phase field with simple integration.

We demonstrate this workflow with the 1D example shown in Figure 4.2, in which the phases (top rows) and phase differences (bottom rows) are in terms of number of phase cycles. The wrapped phases in the top row of (a) each have an ambiguity of $+n$ cycles ($+2\pi n$ rad), where n is any integer. However, if we assume that the Nyquist sampling criterion has been satisfied everywhere, then all phase gradients are within ±0.5 cycles. We begin by selecting the third phase ($\phi = 0$) as the reference from which we will begin the integration. Though the second phase gradient appears to be +0.75 cycles, this would violate the Nyquist sampling criterion, so we assign to it its *complement gradient* ($\Delta\phi_c$), defined as

$$\Delta\phi_c = \begin{cases} \Delta\phi - 1 & \Delta\phi \geq 0 \\ \Delta\phi + 1 & \Delta\phi < 0 \end{cases} \tag{4.4}$$

when expressed in terms of phase cycles. A quick integration of the resulting gradients, relative to the selected reference pixel, provides a deterministic unwrapped phase field from ambiguous measurements, as shown in (b).

Though a quick and simple result, our method breaks down when at least one of the phase gradients does not lie within ±0.5 phase cycles, as is the case for the true absolute phase field in Figure 4.2c. Suppose the third phase gradient (depicted in red) is actually 1.25 cycles, but we

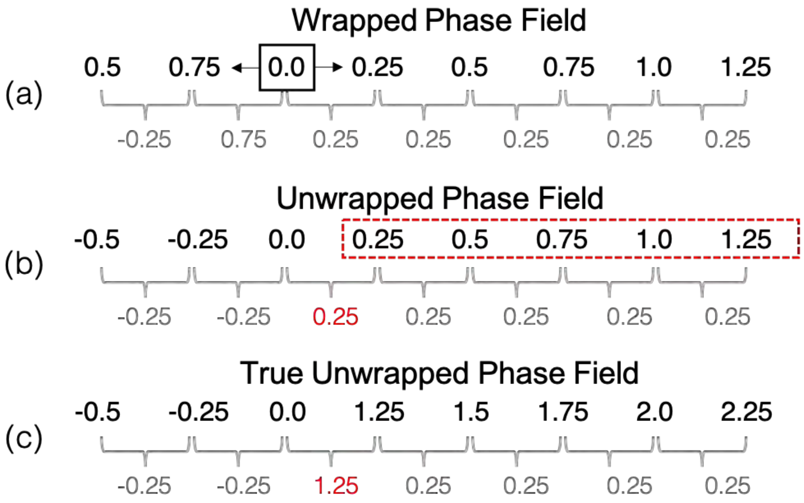


Figure 4.2: Example of phase unwrapping assumptions. (a) The wrapped phase field. (b) The unwrapped phase field solution under the assumption of adequate sampling. (c) Errors in the final solution when the assumption of adequate sampling is violated. If a discontinuity (phase gradient $> \pi$ rad) is not accommodated, the signal is aliased. These aliasing errors propagate beyond the point of the discontinuity (red box in b).

miss-assign it as 0.25 cycles when we assume adequate sampling. The integration leads to the same solution in (b) and all phases within the dashed red box are missing a phase cycle when compared to the true phase field. Thus, we have aliased the target signal in our reconstruction. We refer to a true phase gradient that exceeds 1/2 cycle (π rad) as a *discontinuity* and a dropped phase cycle as a *cycle slip*. As highlighted in (b), the cycle slip that results from an unidentified discontinuity is not contained to the point of its occurrence. Rather, the aliasing error propagates to all phases beyond it, such that the two segments separated by the discontinuity are in error relative to each other. We refer to this as a *global* error.

It is hopefully clear that assuming adequate sampling everywhere in the interferogram is not a good strategy, as there are many reasons that phase gradient magnitudes might exceed π rad. For example, certain artifacts resulting from the interaction of imaging geometry and steep elevation (*layover* to the InSAR specialist) require some gradients to exceed π rad during topographic mapping. Alternatively, added noise may make a phase gradient originally within $\pm\pi$ rad exceed the adequate sampling bounds. And of course, there is no geophysical reason why a true displacement field cannot have gradients that exceed π within the size of a resolution element of any arbitrary size. Thus, the general assumption in almost all 2D phase unwrapping algorithms is that the phases have been adequately sampled in most - *but not necessarily all* - parts of the interferogram. The challenge then becomes locating discontinuities (gradients that have not been adequately sampled) so that they can be accommodated in the unwrapped solution, lest aliasing-induced global errors occur. Indeed, these are the primary goals of phase unwrapping algorithms.

A multitude of unwrapping strategies based on the assumption of adequate sampling exist, many of which are reviewed in Yu et al. (2019). However, only a handful of these are used in practice. In this thesis, we focus on the Statistical-cost Network-flow Algorithm for PHase Unwrapping (SNAPHU), developed by Chen and Zebker (2001) at Stanford University. This formulation for phase unwrapping is considered by many to be the gold-standard and is the algorithm used in a number of practical InSAR-processing packages (ISCE - CalTech, 2022; SNAP - ESA, n.d.; StaMPS - Hooper, n.d.). To best understand SNAPHU's objectives, strategy, and (inevitable) shortcomings, it will help to dive into the concepts and algorithms that ultimately led to its development. We begin with describing a method that explicitly identifies discontinuities, then discuss the general idea behind *optimization-based* unwrapping techniques. SNAPHU draws on these basic concepts in its implementation.

4.2 Residues

The first step of unwrapping algorithms like SNAPHU is to identify the locations of potential discontinuities. For all 1D signals, including the example in Figure 4.2, this is an impossible task in the absence of additional information, but in 2D we have two directions (x- and y) from which to estimate an unwrapped phase, thus the gradients between its neighboring pixels are not strictly independent. For example, an unwrapped phase field that represents a physical surface is irrotational; that is, the curl calculated through any arbitrary loop is identically zero (Thomas & Finney, 1988). This property is not always satisfied for wrapped phase fields in the presence of a discontinuity, specifically under the assumption of adequate sampling. Therefore, we can use a non-zero curl to locate a phase set that contains a known discontinuity.

Almost all algorithms that exploit these inconsistencies calculate the curl via 2x2 loop integrals of wrapped phase gradients under the assumption of adequate sampling, as first proposed by Goldstein et al. (1988). The formula can be described by the four gradients included in the pixel set,

$$\oint \Delta\phi = \Delta\phi_1^{(x)} + \Delta\phi_1^{(y)} - \Delta\phi_2^{(x)} - \Delta\phi_2^{(y)}, \tag{4.5}$$

where superscripts (x) and (y) represent the dimension of the gradient (calculated using Eqs. 4.1 and 4.2) and the subscript corresponds to the order of the gradient in the set, as visually depicted in Figure 4.3a. An example of a calculated loop integral is shown in Figure 4.3b, where the phases are displayed in terms of number of cycles. When we use the native phase gradients, as shown in (b), the loop integral sums to zero, indicating we have a plausible surface. However, there is a true wrapped phase gradient outside the bounds of ±1/2 cycles (-0.8 cycles). The assumption of adequate sampling leads us to use its complement gradient of +0.2 cycles, as demonstrated in Figure 4.3c. The resulting non-zero loop integral is termed a *residue*, which takes on values of ±1

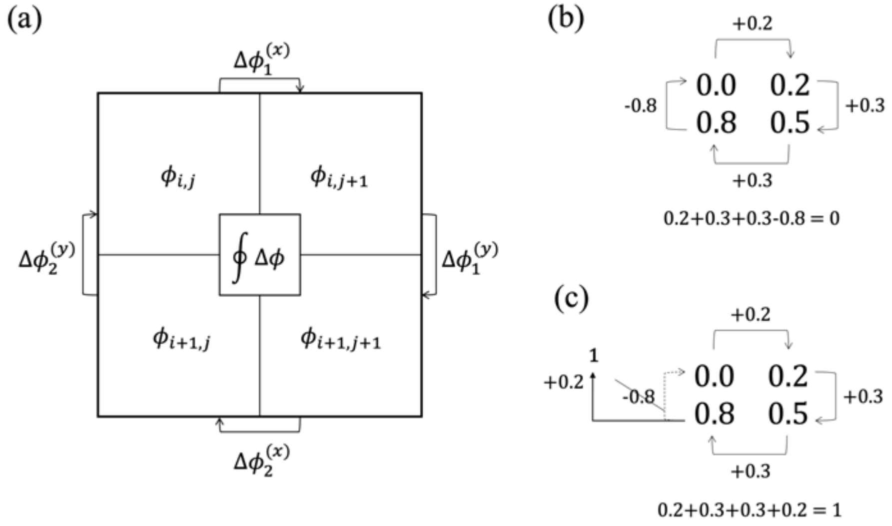


Figure 4.3: Calculation of loop integrals to identify residues. (a) Loop integral formulation in a 2x2 pixel set. (b) When we use the native phase gradients to calculate the loop integral, there will be no residue. However, there is a phase gradient greater than ±1/2 cycles, violating the assumption of adequate sampling. In (c), we use the complement gradient of +0.2 instead of -0.8 and the loop integral identifies a residue.

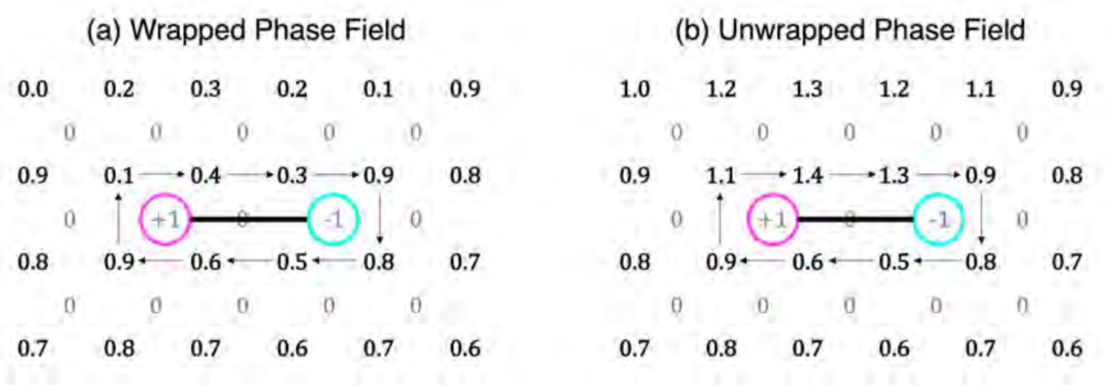


Figure 4.4: A simple example of the residue-cut algorithm. First, we identify residues through the calculation of 2x2 loop integrals. We then connect equal numbers of positive and negative residues with a cut line (barrier), shown in black. Across this line, the unwrapped gradients are greater than 1/2 cycle, whereas everywhere else, the wrapped and unwrapped gradients are identical.

and identifies the location of a discontinuity, at least to within the set of pixels used in the loop integral calculation.

An obvious shortcoming of this method is that, while a residue appropriately identifies a pixel set that contains a discontinuity, it cannot determine which of the four gradients is the discontinuity. Additionally, the absence of a residue does *not* necessarily indicate the lack of a discontinuity. Because of these shortcomings, it is typical to use residues as a guide for creating potential *networks* of discontinuities, along which the gradients are allowed to exceed ±1/2 cycles, examples of which we explore next.

4.3 Optimization-Based Methods

4.3.1 Residue-Cut

The first algorithm created to find and link residues with discontinuity networks was developed by Goldstein et al. (1988). They showed that a loop integral which encloses equal numbers of positive and negative residues will be self-consistent and integrate to zero. They exploited this property by developing a method, known as a residue-cut algorithm, that connects equal numbers of positive and negative residues with tree-like barriers (cuts), across which unwrapped phase gradients may fall outside $\pm\pi$ rad, as demonstrated in Figure 4.4. All other gradients which do not cross cuts are used in the integration step to generate the unwrapped phase field, and consequently, the wrapped and unwrapped gradients at these points are equal. Figure 4.4 uses arrows to highlight that the integration encloses the cut in a consistent, closed loop.

The challenge in residue-based methods is finding the best (in a general sense) network of barriers connecting balanced residue sets, which requires two basic steps. The first is to explicitly define the quantity to be optimized (i.e., what defines "best") in terms of an objective function. The second is to develop an algorithm that can reliably generate the optimal solution to this function (i.e., solve it). Thus, residue-based methods are optimization problems. In the residue-cut algorithm (Goldstein et al., 1988), "best" is minimizing the lengths of the barriers connecting residues. Viewed outside the context of barrier networks, its objective is to minimize the total number of unwrapped gradients that are unequal to their wrapped counterparts, such that corresponding wrapped and unwrapped phases differ by integer phase cycles. This latter condition, known as *congruence*, is explicitly forced in the residue-cut algorithm. However, congruence is not a condition of every phase unwrapping strategy (though it typically leads to more accurate solutions).

A number of methods, in addition to the original algorithm described by Goldstein et al. (1988), have been developed to (approximately) solve this optimization problem (Cusack et al., 1995; Buckland et al., 1995), though the exact techniques are outside the scope of this thesis. What is important is the general shared workflow, which is to 1) identify residues, and 2) connect sets of balanced residues such that the total length of cuts has been minimized. Though it is easy to determine the true minimized solution for simple phase fields as depicted in Figure 4.4, the exact minimum solution generally cannot be found for more complex problems (Chen, 2001). The aforementioned algorithms have nonetheless proven to efficiently generate close-to-minimum solutions of sufficient quality.

One major downside of the original residue-cut algorithm is that noisy, residue-dense regions may cause the networks of barriers to close in on themselves. Because there is no integration path through which the closed-off region can be connected to the rest of the phase field, residue-cut algorithms have been known to generate incomplete solutions. To overcome this problem, Goldstein et al. (1988) suggested the use of spatial low-pass filters prior to phase unwrapping as a means of reducing the number of residues. They showed that spatial filtering[1] leads to more complete solutions and, perhaps more importantly, increases the efficiency of unwrapping. Therefore, it has been adopted as a pre-processing technique in many InSAR workflows, regardless of whether or not incomplete solutions are a concern.

4.3.2 Least-Squares

While most residue-based methods are optimization problems, many optimization-based methods do not require the use of residues. Likewise, they do not always force congruence. One such group of algorithms uses a least-squares formulation as its objective function. Here, the objective (defined in Eq. 4.6) is to minimize the squared differences between the unwrapped and wrapped gradients, still under the assumption of adequate sampling:

---

[1] Spatial filtering can either take the form of a regular filter that smooths the image without reducing the number of pixels, or multilooking, which reduces the number of pixels by a factor of the number of looks.

$$\text{minimize}\left\{\sum_{i,j} w_{i,j}^{(x)}(\Delta\psi_{i,j}^{(x)} - \Delta\phi_{i,j}^{(x)})^2 + \sum_{i,j} w_{i,j}^{(y)}(\Delta\psi_{i,j}^{(y)} - \Delta\phi_{i,j}^{(y)})^2\right\}. \tag{4.6}$$

The unwrapped ($\Delta\psi$) and wrapped gradients ($\Delta\phi$) are defined based on their dimension (superscripts x and y) and the summations include all rows i and columns j of the gradient field. Optional weights (w) can be assigned to each individual gradient, which may be based on additional observables in the interferogram such as magnitude or coherence (e.g. Pritt, 1996; Ghiglia and Pritt, 1998; Zebker and Lu, 1998; Chen and Zebker, 2000).

The primary benefit of implementing least-squares objectives is that exact solutions can be found when all weights are uniform. Ghiglia and Romero (1994)[2] showed that the minimization can be directly calculated using Poisson's equation (Fried, 1977; Hudgin, 1977; Hunt, 1979) and the convolution theorem (Bracewell, 1986). However, the addition of weights requires the use of non-linear solver techniques, such as iterative preconditioned conjugate-gradient methods (Ghiglia & Romero, 1994) or multiresolution techniques (Pritt, 1996; Davidson and Bamler, 1999), among others.

Both weighted and unweighted least-squares algorithms do not typically use residues, explicitly define cut-networks, or force congruence on any unwrapped phase gradient. Therefore, another major benefit is that they always generate complete and spatially smooth solutions. Though this is aesthetically pleasing, least squares techniques have empirically under-performed compared to other objective functions. Specifically, they tend to underestimate the true phase gradients, producing biased results (Bamler et al., 1998; Zebker and Lu, 1998; Ghiglia and Pritt, 1998; Chen and Zebker, 2000).

4.3.3 L^p-Norm Framework

Though the strategies, benefits, and shortcomings of residue- and least-squares-based methods are quite different, Ghiglia and Romero (1996) showed that they are essentially the two practical end-members of what has come to be known as the minimum L^p-norm framework. This framework explicitly defines phase unwrapping as an optimization problem, in which the primary objective is to reduce the differences between wrapped and unwrapped gradients in some sense determine by the value of p, as described by Eq. 4.7:

$$\text{minimize}\left\{\sum_{i,j} w_{i,j}^{(x)}|\Delta\psi_{i,j}^{(x)} - \Delta\phi_{i,j}^{(x)}|^p + \sum_{i,j} w_{i,j}^{(y)}|\Delta\psi_{i,j}^{(y)} - \Delta\phi_{i,j}^{(y)}|^p\right\}. \tag{4.7}$$

As in Eq. 4.6, w are user-defined weights, $\Delta\psi$ and $\Delta\phi$ are the unwrapped and wrapped gradients,

---

[2] While the formulation of this method is outside the scope of this thesis, we encourage interested readers to refer to Ghiglia and Romero (1994) or its review in Chapter 2 of Chen (2001), as it demonstrates the elegance, and dare we say "magic"(?), of Fourier methods.

respectively, superscripts x and y indicate the dimension of the gradient, and summation includes all columns i and rows j of the interferogram. When all $w_{i,j}$ are uniform (i.e., $w_{i,j} = 1$) and $p = 0$, the objective is identical to that of the residue-cut algorithm, which can be considered as the L^0 problem. When $p = 2$ (the L^2-norm), the objective represents the weighted least-squares minimization.

It will be useful to re-frame the L^p objective as a cost function, which maps a solution value (e.g. the difference between a corresponding wrapped and unwrapped gradient) to a cost associated with that solution. Stated in this way, it is the cost of the entire solution set that we wish to minimize with our objective function. As an example, we begin with Φ and Ψ, which are the complete sets of wrapped and unwrapped gradients, respectively. The general objective is to minimize some arbitrary function ($G(\cdot)$) of these two sets:

$$\text{minimize}\left\{G_{\Psi,\Phi}(\Delta\Psi, \Delta\Phi)\right\}. \tag{4.8}$$

Each individual set includes all gradients in the x- and y-directions. Assuming that all of these gradients are separable (as is general practice), we can vectorize the matrices and display the minimization as a sum of the costs associated with each pair of corresponding unwrapped ($\Delta\psi$) and wrapped gradients ($\Delta\phi$):

$$\text{minimize}\left\{\sum_k g_k(\Delta\psi_k, \Delta\phi_k)\right\}, \tag{4.9}$$

where g_k is the arbitrarily-shaped cost function associated with each gradient pair k. In the L^p objectives, these cost functions take the familiar form of the summed term in Eq. 4.7:

$$g_k(\Delta\psi_k, \Delta\phi_k) = w_k|\Delta\psi_k - \Delta\phi_k|^p. \tag{4.10}$$

The cost functions of Eq. 4.10 for various p values are visually depicted in Figure 4.5.

Now that we have neat equations that summarize the L^p-norm framework in terms of its objective or cost functions, it is natural to ask which value of p leads to the most accurate solutions. This is a complicated question to answer because the success of any unwrapping scheme depends on multiple factors, including the objective (e.g. value of p), the inclusion of additional constraints (e.g. congruence), and the solver used to estimate its minimum.

Empirically, weighted L^0 solutions give the best results, but as mentioned in section 4.2, the truly optimal solution is generally impossible to find (Chen, 2001) (though approximate minimization is still valuable). The original residue-cut algorithm (Goldstein et al., 1988) (the unweighted L^0 objective) also leads to good approximate solutions, but only when residue density is low. On the other hand, there are a number of L^1 algorithms that can be efficiently used to find an exact optimal solution, but when the true phase field includes unwrapped gradients that differ by more that ±1 phase cycle, the L^1 solution does not perform well. Its under-performance can be attributed to the

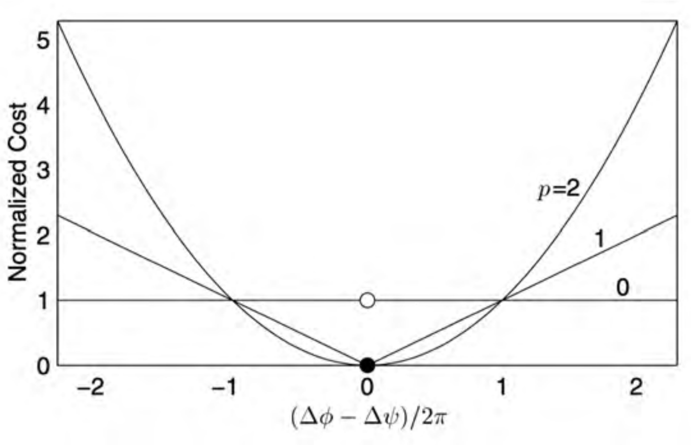


Figure 4.5: L^p objective functions expressed as costs for $p \in \{0, 1, 2\}$, all in their unweighted forms. The cost is a function of the difference between the unwrapped ($\Delta\psi$) and wrapped gradient ($\Delta\phi$), in terms of number of cycles. Figure from Chen (2001).

fact that it favors short, residue-pair discontinuities with ($\Delta\phi$ - $\Delta\psi$) $\in \{-1, 0, 1\}$, over connections of multiple residues that may contain larger cycle differences between corresponding gradients (Chen, 2001). Upon further increase to $p = 2$, the L^2 solution generally underestimates the magnitude of the unwrapped gradients. In addition, it loses optimality whenever congruence is included as an additional constraint. Therefore, despite the elegant and efficient algorithms we have to find an exact L^2 solution, they are only favorable to use in low-gradient scenarios.

This discussion brings to light an important distinction that can be made at this stage. An unwrapping *error* in the context of an optimization problem occurs when it is not possible to find a complete solution that minimizes (or approximately minimizes) the objective function. On the other hand, the *quality* of the unwrapped solution, which will vary between unwrapping methods even in the absence of errors, should be considered only relative to the true phase field. In general, errors may be more likely as p decreases, but solution quality (if it can be approximately computed) also tends to increase.

4.4 Network-Flow

The final concept needed to understand SNAPHU is a visualization of how the ideas and properties from the previous sections actually appear in the phase field. Costantini (1998) was the first to propose that the phase unwrapping problem be explicitly represented by a network flow model, which consists of nodes connected by arcs. As an example, suppose that we define the 2x2 loop integrals used to calculate residues (refer to Figure 4.3) as the nodes in our network. These nodes take on values of ±1 or 0, which can be generalized to representing the surplus (+), demand (-), or neutrality (0) of an arbitrary commodity. Any non-neutral node is considered to be *charged.* Arcs connect each neighboring node along the divide between neighboring phase values, as depicted in Figure 4.6, thus there is one arc for each phase gradient. Flow accumulates along arcs connecting charged nodes and this flow represents the difference between the wrapped and unwrapped phase gradients associated with that arc.

Phase unwrapping is then finding the optimal arrangement of flow, where flow paths can be thought of similarly to the barriers in the residue-cut algorithm. That is, integration of the phase gradients, the final step to generate the unwrapped solution, will not cross arcs included in the flow network. Therefore, there is a critical condition that the arrangement of flow needs to meet for there to be a valid unwrapped phase solution; namely, flow must be conserved. When conserved, flow into a node equals the flow out, plus or minus the surplus or demand associated with that node. Flow is considered feasible when this condition is met, where all feasible flow arrangements correspond to a unique, irrotational unwrapped phase field.

One benefit of the network-flow model is that existing minimum-cost-flow (MCF) solver routines can be readily applied to find a flow arrangement that approximately solves the desired objective in terms of cost functions. These solvers associate scalar costs (e.g. Eq. 4.10) to the arc flows and minimize the total cost associated with the solution (Eq. 4.9). Furthermore, the network-flow model is a completely general representation of the unwrapping problem, regardless of the commodity (e.g. residue), congruence condition, or cost functions used.

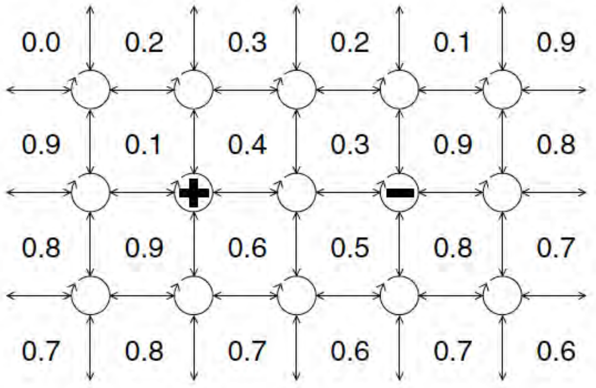


Figure 4.6: A network model representation of the phase unwrapping problem. The wrapped interferogram can be represented by its phase values (normalized here to be in terms of phase cycles). The network nodes are the loop integrals associated with each set of 2x2 pixels in the image, where the residues indicate supply (+) and demand (-) nodes, or charged nodes. Arcs connect neighboring nodes and represent potential flow paths of the tree branches containing the set of all charged nodes. Figure taken from Chen (2001).

4.5 SNAPHU

We now have all of the pieces needed to understand SNAPHU in terms of its objectives and the key conditions required of its solver. As we saw in the last few sections, both aspects of an unwrapping algorithm determine the quality of the solution with respect to the true absolute phase field. Here, we focus only on its main features and basic implementation, but refer the reader to the full thesis in Chen (2001) for more details.

4.5.1 SNAPHU Cost-Functions

The L^p optimization framework has a number of shortcomings, despite its success in many studies. First, the L^p-norm objectives and cost-function shapes (Figure 4.10) do not have any physical meaning. For example, despite the empirical success of L^0 solutions, there exists no geophysical requirement demanding that the solution with the fewest number of discontinuities be the most correct with respect to the true phase field. Second, though weights based on auxiliary information (e.g. amplitude or coherence) can be included in Eq. 4.7, the weighting function takes on a similar shape throughout the interferogram, a restriction which need not be the case for the generalized cost functions. SNAPHU's unwrapping objective is framed to address both of these shortcomings.

In an effort to develop objectives founded on physical principles related to additional observables in the interferogram, the optimization developed for SNAPHU is posed as a maximum *a posteriori* probability (MAP) estimation problem based on non-linear, generalized cost functions. That is, the cost functions represent conditional probability density functions (PDF) that take on different forms throughout the interferogram. Furthermore, the shape of the PDFs vary depending on the application.

The theoretical transformation of a generalized MAP objective to the cost functions used in SNAPHU can be found in Appendix A. Here, we simply state that the general cost function $g(\cdot)$ associated with each unwrapped gradient depends on the discrete conditional PDF of $\Delta\psi_k$ (evaluated only at values of $\Delta\psi_k = \Delta\phi_k + n2\pi$, n an integer), given the local intensity (I) and coherence (ρ):

$$g_k(\Delta\psi_k, \Delta\phi_k) = -\log(f(\Delta\psi_k|I,\rho)) \text{ for } \Delta\psi_k \equiv \Delta\phi_k + n2\pi. \tag{4.11}$$

Noting that maximization of a probability is equivalent to the minimization of the negative logarithm of the PDF $f(\cdot)$, the objective is then to minimize the total sum of costs associated with the complete set of unwrapped gradients. The nature of the discretization implies that congruence is an additional condition. Furthermore, for this objective to be feasibly solvable, it is necessary to assume that each unwrapped gradient is independent of all other unwrapped gradients and that the unwrapped solution is residue-free (i.e., irrotational).

The PDFs in Eq. 4.11 take different forms depending on the application (topography versus deformation mapping) and whether or not discontinuities are deemed likely given the values of intensity and coherence (which further depends on the number of looks). Since topographic mapping is outside the scope of this thesis, we will briefly review how the cost functions behave in deformation mode, but note that topography-mode PDFs intimately depend on both intensity and coherence.

In contrast to topography mode, the PDFs for deformation do not rely on intensity, since amplitude has no obvious relationship with surface displacement. The cost functions in this mode can thus be treated as conditional only on ρ:

$$f(\Delta\psi_{defo}|I,\rho) = f(\Delta\psi_{defo}|\rho), \tag{4.12}$$

where the subscript $defo$ highlights that this form of cost function applies only to deformation mapping. Chen (2001) claims that high displacement gradients are more likely to occur where there is low correlation, since deformation implies distortion of the surface. These distortions can be accompanied by changes in scattering properties (Zebker & Villasenor, 1992) and local misregistration of pixels (Just & Bamler, 1994).

When coherence is above some threshold ρ_{min}, the function behaves like a simple parabola centered at zero (i.e., the negative logarithm of a Gaussian noise PDF). In other words, the most likely unwrapped gradient is no different than the wrapped gradient. The width of the parabola depends on the modeled phase noise ($\sigma_{\Delta\psi}$), which is a function of measured coherence and the number of independent looks used in the coherence estimate (Li and Goldstein, 1990; Zebker and Villasenor, 1992; Rodriguez and Martin, 1992; Lee et al., 1994).

When coherence is below ρ_{min}, the cost function includes shelf-like regions centered around a zero-mean parabola. These shelves, which represent a discontinuity, have a cost of g_d and extend up to the maximum theoretical unwrapped gradient magnitude ($\Delta\psi_{max}$). Beyond $\Delta\psi_{max}$, the cost tails upward with a rate determined by C_{tail}. Both high- and low-coherence functions are displayed in Figure 4.7. The variables ρ_{min}, g_d, C_{tail}, and $\Delta\psi_{max}$ are empirically-derived constant based on many applications of SNAPHU.

Comparing Figure 4.7 to Figure 4.5, it is easy to see that that the statistical cost functions behave similarly to the L^0 objective when coherence is low (and a discontinuity is more likely to occur). Conversely, when coherence is high, a discontinuity is unlikely and the cost is similar to the L^2 objective (which works well when no physical discontinuities are expected). Therefore, the functions used in SNAPHU take advantage of the success of L^p-norm objectives in the scenarios in which they work the best. Moreover, they adapt to the local coherence conditions. Having improved the framework of the phase unwrapping objective, the only step left is to reliably find the solution that approximately minimizes it.

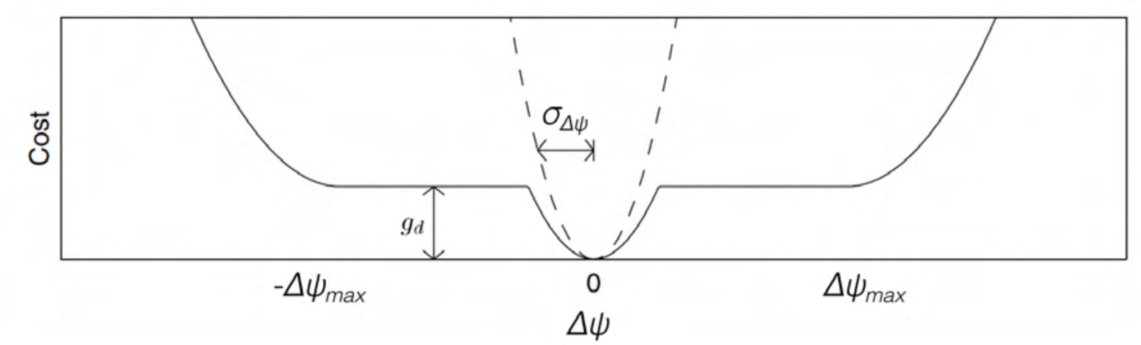


Figure 4.7: Example deformation-mode cost functions for SNAPHU. When coherence ρ is greater than threshold ρ_{min}, the cost function represents a parabola with a width determined by the phase noise ($\sigma_{\Delta\psi}$). When $\rho < \rho_{min}$, the cost function includes a shelf-like region which allow for potential discontinuities. Figure modified from Chen (2001).

4.5.2 Practical Implementation

Once again, the details regarding the SNAPHU optimization solver are beyond the scope of this thesis (for the full description of the SNAPHU solver techniques, see Chen (2001)). We therefore focus on the main steps, summarized in Figure 4.8, and solution criteria.

SNAPHU is a residue-based method; therefore, it begins by identifying residues and placing them at the nodes of the network-flow grid. The objective is to connect the entire set of charged nodes with a valid flow arrangement, such that the solution cost has been (approximately) minimized. This flow connection takes the form of a tree (or a subset of trees), which connect equal numbers of positive and negative nodes along arcs which do not form closed loops. A valid set of trees is depicted in Figure 4.9a, whereas the node connection in (b) is an invalid tree because there is no integration path to the pixels enclosed by the cuts.

```
algorithm SNAPHU
begin
    calculate residues
    find an initial feasible $L^1$ or $L^0$ flow (MCF or MST, respectively)
    while improvements can be made:
        reduce cost with pivot-and-grow algorithm
    end
    integrate the phase (no integration path crosses a tree branch)
end
```

Figure 4.8: Basic pseudo code for SNAPHU.

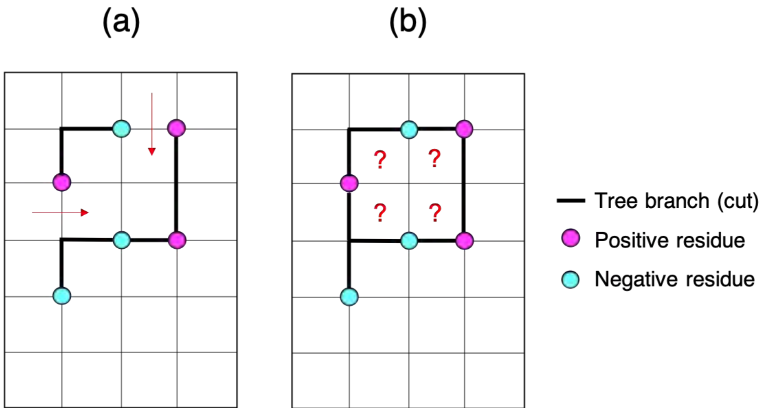


Figure 4.9: Branched-tree network. The solution in (a) is an allowable tree structure in SNAPHU, though it is not a minimum-spanning tree. The tree in (b), however, is not allowed because the cuts form a closed loop. A closed loop is an invalid tree formation because there is no longer an integration path to the pixels within the encircled segment, resulting in an incomplete solution.

The arrangement of flow begins with an initialization, either via a minimum cost flow (MCF) or minimum spanning tree (MST) algorithm. The former finds a valid L^1-norm solution. The latter develops a single tree connecting the entire set of charged nodes, which approximately minimizes the lengths (L^0) of branches emanating from the root node (simply the first residue arbitrarily selected by the algorithm). Both result in a valid flow network, which is neutral (positive and negative charges balance) and has a unique flow path between any arbitrary pair of nodes on a given (sub-)tree.

After initialization, the tree is incrementally improved by rearranging flow orientations such that they reduce the total cost of the solution. SNAPHU uses a pivot-and-grow technique, but other methods exist, such as dynamic-cost cycle canceling (both described in Chen, 2001). The improvements are made until no more can be found, in which case we have the final flow network defining the locations of potential discontinuities. Integration of the gradients around the tree cuts leads to the final unwrapped phase field. We note that while the integration is not allowed to cross the identified cuts, there may not be a discontinuity along every arc included in the tree, which is why we consider the final flow network to merely represent *potential* discontinuities.

4.6 Potential Shortcomings

As noted in the previous section, one critical requirement of the SNAPHU solver is that the tree network connecting residues cannot form closed loops. This condition is considered necessary for two reasons. First, it ensures that the path connecting charged nodes is unique and that flow is conserved along the path. In other words, it automatically generates a valid solution. Second, because the integration is designed to use only gradients that do not cross cuts, only a branched tree formulation can ensure a *complete* valid solution. If the branches of the tree did form closed-loops, there would be no integration path connecting the interior of the encircled segment to the rest of the phase field, as demonstrated in Figure 4.9.

Despite these benefits, it is important to highlight that neither justification references the implausibility of a closed-loop of high gradients appearing in an unwrapped solution. As it is, there is no geophysical principle precluding this from happening (just as there is no physical requirement that the L^0 solution should perform well). In fact, many of the models we use to represent underground volumetric changes produce radial patterns of displacement at the surface. So, if there are high gradients in pumping- or injection-induced deformation applications, excluding those induced by noise, it is actually likely that they will form closed loops.

The branched condition is thus a requirement of the *solver* as opposed to a requirement of the phase field. This distinction did not appear to be too important at the time of the major phase unwrapping developments discussed in this chapter. For example, the main InSAR application in the late 1990's was topographic mapping, in which the primary non-noise discontinuity is a curvilinear feature due to layover. At this time, the use of InSAR to measure the displacement of the Earth's surface was in its infancy and primarily used to study impressive deformation events such as large earthquakes. As earthquakes often rupture along curvilinear features (faults), the expected discontinuities were again assumed to follow this form. Therefore, many phase unwrapping algorithms leading up to and including SNAPHU were explicitly designed to accommodate only isolated, curvilinear discontinuities, specifically in the form of layover[3] or noise.

There is another potential shortcoming that is not related to a conditional feature of any specific unwrapping algorithm. Rather, it is the practice of applying spatial filters as a pre-unwrapping step to reduce the number of residues. This technique was first suggested in the original residue-cut algorithm (Goldstein et al., 1988) and it continues to be a recommended method today, even when using SNAPHU (which ensures a complete solution). Though filtering reduces noise and increases the efficiency of unwrapping, these benefits are at the expense of spatial resolution. While a discontinuity is always defined as a true gradient between adjacent resels that lies outside the bounds of $\pm\pi$ rad, the resolution of the image dictates the distance over which this gradient spans. Therefore, filtering will reduce the gradient magnitude that defines the threshold between adequate and under-sampling (e.g. $\frac{\pi}{20}$ rad m^{-1} to $\frac{\pi}{40}$ rad m^{-1}), which may increase the number of discontinuities, potentially in the form of closed loops, that need to be accommodated.

What will happen when discontinuities forming closed loops are not accommodated? Referring back to Figure 4.2, we can predict that the unaccommodated discontinuity will impose a cycle slip, resulting in biased global aliasing errors. Though it is indeed an error with respect to the true absolute phase field, there is a nuanced distinction between the well-known unwrapping *Error* (with a capital E) and what aliasing really is: an unwrapping *slip*. The former describes the root-failure

---

[3] This is highlighted by the fact that the original SNAPHU thesis devotes 22 pages to the derivation of cost functions for topographic mapping (layover) versus only 3.5 pages for the derivation of deformation cost functions.

of an algorithm to either 1) (approximately) solve its objective function, or 2) develop plausible geophysical surfaces because of noise. We define the latter as a failure of the algorithm to output a quality unwrapped solution due to a mismatch between the true phase field and the assumptions used in the phase unwrapping algorithm. More specifically, the solution solves the intended objective, but it includes biased regions that underestimate the true unwrapped magnitude – it is a PHase-Unwrapped, but Biased by Aliasing Result (a PHUBAR[4]).

Unwrapping Errors and slips are both significant sources of error, but slips should be of considerably more concern because they do not *look* like errors. Figure 4.10a shows an example displacement field. As shown in Figure 4.10b, the canonical unwrapping Errors can be thought of as random, sharp, and disjointed jumps that have no relationship to a real physical surface, much less the true displacements in (a). On the other hand, Figure 4.10c has up to 11 cm of unwrapping slips within the black circle, but shows a smooth surface, making the slip challenging to identify visually. In addition, slips are biased, whereas Errors are more-or-less random phase jumps. Unwrapping slips, and their effects on our ability to use InSAR to study subsurface fluid flow, are the focus of Chapter 5.

We do not wish to overlook that SNAPHU has empirically produced high-quality results in many InSAR studies since its development. Rather, we hope to highlight that there has been little re-analysis of its solution quality in the context of new InSAR applications, despite rapidly expanding in the last two decades. Furthermore, we anecdotally claim that the “problem” of phase unwrapping has been demoted to a simple “step” in the overall workflow for many InSAR users,

---

[4] It was too fun to come up with this acronym to not slip it into this thesis.

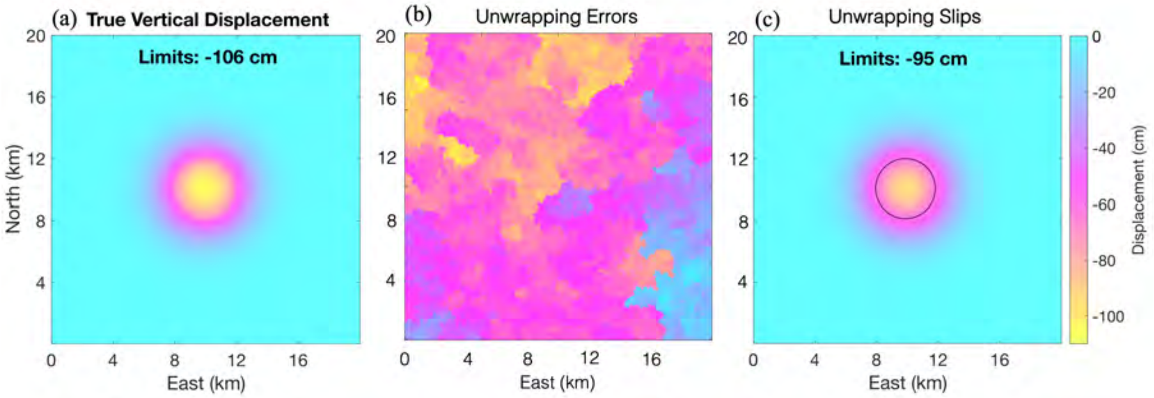


Figure 4.10: Unwrapping Errors versus slips. (a) The true vertical displacement. (b) Unwrapping Errors are distinguishable as unrealistic phase jumps with no physical meaning. (c) Unwrapping slips are not readily noticeable in the unwrapped interferogram, despite being incorrect relative to the true solution. The black circle delineates where the solution in (c) differs from the true displacement field in (a).

thus its importance is largely taken for granted. In other words, we happily press the standard "run" button to loop through our set of thousands of interferograms so that we can get to the fun part: deciphering the displacement field. The goal of this section and the following chapter is not to prove that the devaluation of phase unwrapping is a reality per se, but rather to describe the conditions in which aliasing still may occur and show why these errors can be overlooked when we adopt this attitude. First and foremost, undervaluing the importance of phase unwrapping contradicts a claim in the original thesis which overviews SNAPHU (Chen, 2001), namely that the objective to be solved in any phase unwrapping problem needs to be aligned with the target phenomenon; there is no universal set of cost functions or solver that will work well for every interferogram. In line with this sentiment, the next chapter re-evaluates the use of SNAPHU in modern InSAR applications and common workflows so that its users better understand the interferograms they are analyzing. While we do not test any other phase unwrapping algorithms, we believe this reanalysis may apply to other packages that require similar assumptions and/or solver techniques.

Chapter 5

Aliasing in InSAR

We may now begin our exploration of the management of hazards associated with fluid extraction and injection with the use of geophysical imaging techniques. We first use InSAR to identify and measure deforming areas in order to characterize the structure of the subsurface and its behavior in response to well activity. As with any geophysical study, our interpretations are heavily dependent on the quality of the data and data products we use, thus we begin our journey with a discussion of the quality of interferograms and their derived time series. In particular, we describe and demonstrate a biased error source in InSAR: aliasing during 2D phase unwrapping.

As we saw in Chapter 4, aliasing occurs when a signal reconstructed from samples is different than the original signal due to undersampling. Many 2D phase unwrapping algorithms (e.g. Goldstein et al., 1988; Ghiglia and Romero, 1994; Chen, 2001) attempt to avoid aliasing by identifying locations that violate the Nyquist sampling criterion (Nyquist, 1928): points at which the true phase gradient lies outside the bounds of $\pm\pi$ rad, which we refer to as a discontinuity. If properly identified, the algorithm may be able to accommodate these discontinuities in the unwrapped phase solution, but not always.

The Statistical-Cost Network-Flow Algorithm for PHase Unwrapping (SNAPHU; Chen, 2001), one of the most-commonly used phase unwrapping algorithms in the field (Yu et al., 2019), is equipped to find and accommodate isolated discontinuities associated with noise, curvilinear high-gradient displacements, and artifacts resulting from high topography (*layover* to the InSAR specialists). Our applications require consideration of only the former two, since we remove the topographic phase (including layover) during the formation of the deformation interferogram. In this chapter, we show that unwrapping errors due to aliasing can and do occur when high displacement gradients relative to the resolution of the image form *closed loops* (or, in some cases, approximately closed loops). These global aliasing errors (referred to as unwrapping slips) propagate well beyond the high-gradient boundary, as demonstrated in the simplified schematic in Figure 5.1. Furthermore, they are currently unaccounted for in InSAR uncertainty quantification, yet have the potential to

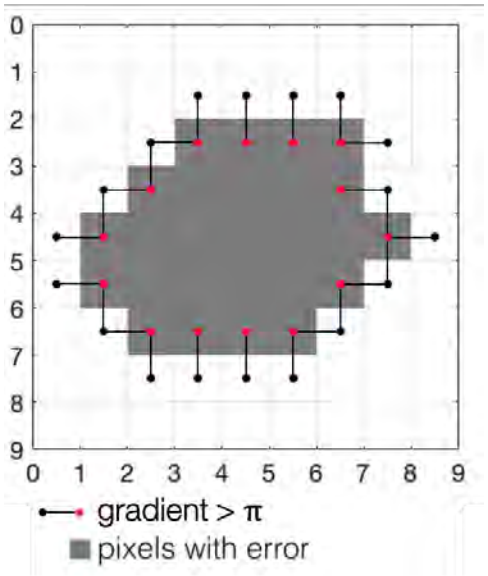


Figure 5.1: Schematic of aliasing errors during 2D phase unwrapping. When a high displacement gradient relative to the image resolution forms a closed loop in the wrapped interferogram, aliasing will lead to a cycle slip that propagates to all points within the enclosed loop (assuming that the reference phase lies outside of the closed loop)

be the largest source of error with respect to the true surface displacements. We first describe the conditions that lead to aliasing even in the absence of noise, highlight ubiquitous InSAR practices that increase the likelihood of aliasing, then demonstrate the interaction of aliasing with noise.

5.1 General Sources of Aliasing

Though interferograms in the real world include at least some degree of noise (e.g. thermal noise or decorrelation), it will be useful to overview the scenarios in which aliasing will occur in the absence of noise. We begin with examples that include only the signal-of-interest (e.g. the deformation field) and assume that a resolution element is the same size as a pixel in the image. In section 5.6 we discuss the scenario in which the pixel size is smaller than a resolution element.

Aliasing may occur in unwrapping algorithms such as SNAPHU because of (1) a direct mismatch between their objectives and the true displacement field, or (2) inadequate methods to identify potential discontinuities. We can demonstrate this with the two simple examples shown in Figure 5.2, in which the unwrapped (top row) and wrapped (bottom row) phase fields are depicted in terms of number of cycles. In this scenario, the intensity and coherence are uniform throughout the interferogram (note that the calculation of coherence requires multilooking, which we have not done; in this noise-free case we assume it is 1 everywhere). The uniform intensity and coherence indicate that the SNAPHU cost functions assigned to each gradient are identical within each mode (see Chapter 4.5). Therefore, the solution that minimizes the total cost will have the smallest possible number of discontinuities, mimicking the objective of the residue-cut algorithm (Goldstein et al., 1988).

The left column shows a phase field that SNAPHU can easily unwrap despite the presence of discontinuities. Here, the appropriate cuts (as shown by the dashed red lines in the top plot) align with the minimized connection of residues in the wrapped phase field; therefore, the placed cuts are identical to the true cuts and there are no observed cycle slips.

The right column demonstrates a slightly altered case in which SNAPHU fails for two reasons. We begin with the four sets of discontinuities on the corners of the yellow block in the true phase field. The corresponding wrapped phase field has the same number of residues as the left example and in identical locations (albeit with opposite sign). In an attempt to minimize the cut lengths, SNAPHU connects the residues along the sides, identically to the left example, rather than placing them in their correct positions along each corner. This results in the first cycle slip, which propagates to all of the colored pixels shown in the true unwrapped solution.

We now turn our attention to the inner high-gradient loop surrounding the orange region. The corresponding wrapped phase field shows no residues associated with these discontinuities. Consequently, there is no apparent violation of the Nyquist sampling criterion and no problem that SNAPHU must accommodate, leading to the second cycle slip. Note that the middle pixel has accumulated two cycle slips, one resulting from the misidentification of proper residue connections, the

0.50	0.80	0.90	0.80	0.50
0.80	1.28	1.41	1.28	0.80
0.90	1.41	1.50	1.41	0.90
0.80	1.28	1.41	1.28	0.80
0.50	0.80	0.90	0.80	0.50

0.50	0.80	0.90	0.80	0.50
0.80	1.31	1.28	1.31	0.80
0.90	1.28	2.50	1.28	0.90
0.80	1.31	1.28	1.31	0.80
0.50	0.80	0.90	0.80	0.50

0.50	0.80	0.90	0.80	0.50
0.80	0.28	0.41	0.28	0.80
0.90	0.41	0.50	0.41	0.90
0.80	0.28	0.41	0.28	0.80
0.50	0.80	0.90	0.80	0.50

0.50	0.80	0.90	0.80	0.50
0.80	0.31	0.28	0.31	0.80
0.90	0.28	0.50	0.28	0.90
0.80	0.31	0.28	0.31	0.80
0.50	0.80	0.90	0.80	0.50

Figure 5.2: Simple examples of high-gradient accommodation in SNAPHU. In the left example, minimization of the length of cuts connecting positive and negative residues leads to the correct unwrapped solution. Minimization of the cuts connecting residues in the right example leads to a misidentification of cut locations and the first cycle slip. The center, fully-closed high-gradient loop in the right example has no residues, thus the algorithm does not correctly unwrap the phase at that location either, and the cycle slips accumulate to two.

second from the absence of residues associated with the enclosed, high-gradient loop. These examples are the two general sources of unwrapping slips in the absence of noise and spatial filtering.

5.2 Residue Conditions

Both types of unwrapping slips are dependent on residues (or lack thereof); therefore, it is important to know what gradient conditions lead to the occurrence or absence of a residue. Recall that we only observe the ambiguous wrapped phase measurements, so the condition of a residue relies solely on the behavior of the measured wrapped phase gradients ($\Delta\phi$) within each 2x2 set of phases (ϕ), arranged as in Figure 5.3a. The formal equations for x- and y-gradients (Eq. 4.1-4.2), their complement gradients (Eq. 4.4), and loop integral (Eq. 4.5) can be found in Chapter 4. Here, we simplify our terminology and depiction to that represented in 5.3b.

We find residues by calculating the loop integrals ($\oint dx$) for each 2x2 set of phases (x_i) in the interferogram (expressed here in number of cycles), under the condition of adequate sampling. That is, when the measured gradient magnitude ($|dx_i|$) is larger than 1/2 cycle, we re-assign it to be its complement gradient ($\hat{dx_i}$):

$$\hat{dx_i} = \begin{cases} dx_i - 1 & dx_i \geq 0 \\ dx_i + 1 & dx_i < 0. \end{cases} \tag{5.1}$$

The loop integral under the condition of adequate sampling ($\oint dx_{as}$) is the sum of the four restricted gradients:

$$\oint dx_{as} = \sum_i^n dx_i + \sum_j^m \hat{dx}_j, \tag{5.2}$$

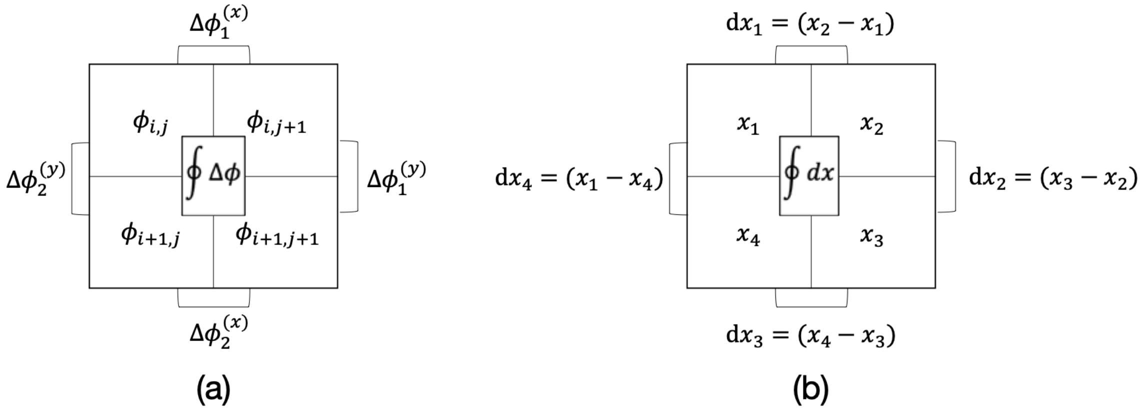


Figure 5.3: Loop integrals expressed as our original definition (a) and a simplified version (b).

where n is the number of natural gradients, m is the number of complement gradients, $i \neq j$, $n + m = 4$, $|dx_i| < 0.5$, and $|\hat{dx}_i| < 0.5$. The conditions we have imposed on our gradients lead to a simple criterion that will dictate whether or not the loop integral will result in a residue. Namely, we will observe a residue if and only if we use an odd number of complement gradients, and there will be no residue when we use an even number of complement gradients. The derivation of this criterion can be found in Appendix B, but we summarize the findings here with an update to Eq. 5.2:

$$\oint dx_{as} = \sum_{i}^{n} dx_i + \sum_{j}^{m} \hat{dx}_j = \begin{cases} 0 & m \text{ even} \\ \pm 1 & m \text{ odd.} \end{cases} \tag{5.3}$$

5.3 Residue Patterns

Now that we know that the occurrence of a residue depends only on the number of complement gradients enforced in the loop integral calculation, it is useful to see the patterns of residues on noise-free surfaces with contoured displacements. We begin with the 2D exponential in Figure 5.4. Figure 5.4a and b show the true, unwrapped gradients in x- and y-components, respectively. The gradients in x are $\Delta\psi^{(x)} = \psi_{(i+1,j)} - \psi_{(i,j)}$ for all columns i and rows j, and similarly, the y-component gradients are $\Delta\psi^{(y)} = \psi_{(i,j+1)} - \psi_{(i,j)}$. Figure 5.4c and d show the corresponding wrapped gradients in x- ($\Delta\phi^{(x)} = \phi_{(i+1,j)} - \phi_{(i,j)}$) and y-components ($\Delta\phi^{(y)} = \phi_{(i,j+1)} - \phi_{(i,j)}$), under the assumption of adequate sampling. When the magnitude of a calculated gradient exceeds π rad, its complement gradient is assigned instead (see Eq. 4.4).

Included in Figure 5.4 is the location of residues which we find using Eq. 4.5. Though there are no residues in the unwrapped phase field (all valid unwrapped surfaces are irrotational), we include them in these plots to highlight the conditions that lead to their occurrence in the wrapped phase field. Figure 5.4a and b show that residues exclusively appear along the boundaries of the lobes in which the true x- or y-gradients exceed $\pm\pi$ rad, and the sides of each lobe have opposite residue polarities (+ or -).

The corresponding lobes in the wrapped phase field have the opposite sign compared to the true unwrapped gradients. That is, the gradients that should be less than $-\pi$ are now greater than π, and vice versa. Some of these sign reversals are simply a result of wrapping the true phase field, the others are forced to have opposite sign when we use the condition of adequate sampling. For example, consider we have a set of phase cycles [0.8 1.3 1.9 2.7] with positive gradients increasing to the right ([0.5 0.6 0.8]). Once wrapped into the phase cycle set [0.8 0.3 0.9 0.7], the gradients appear to be [-0.5 0.6 -0.2] cycles. Two of the wrapped gradients (-0.5 and -0.2) are within the adequate sampling bounds, but have opposite sign to their true unwrapped counterparts. On the other hand, the gradient with the appropriate sign (0.6) exceeds 1/2 cycle, so we force it to be its complement gradient (-0.4) when we calculate residues. We refer to any gradient that appears to be

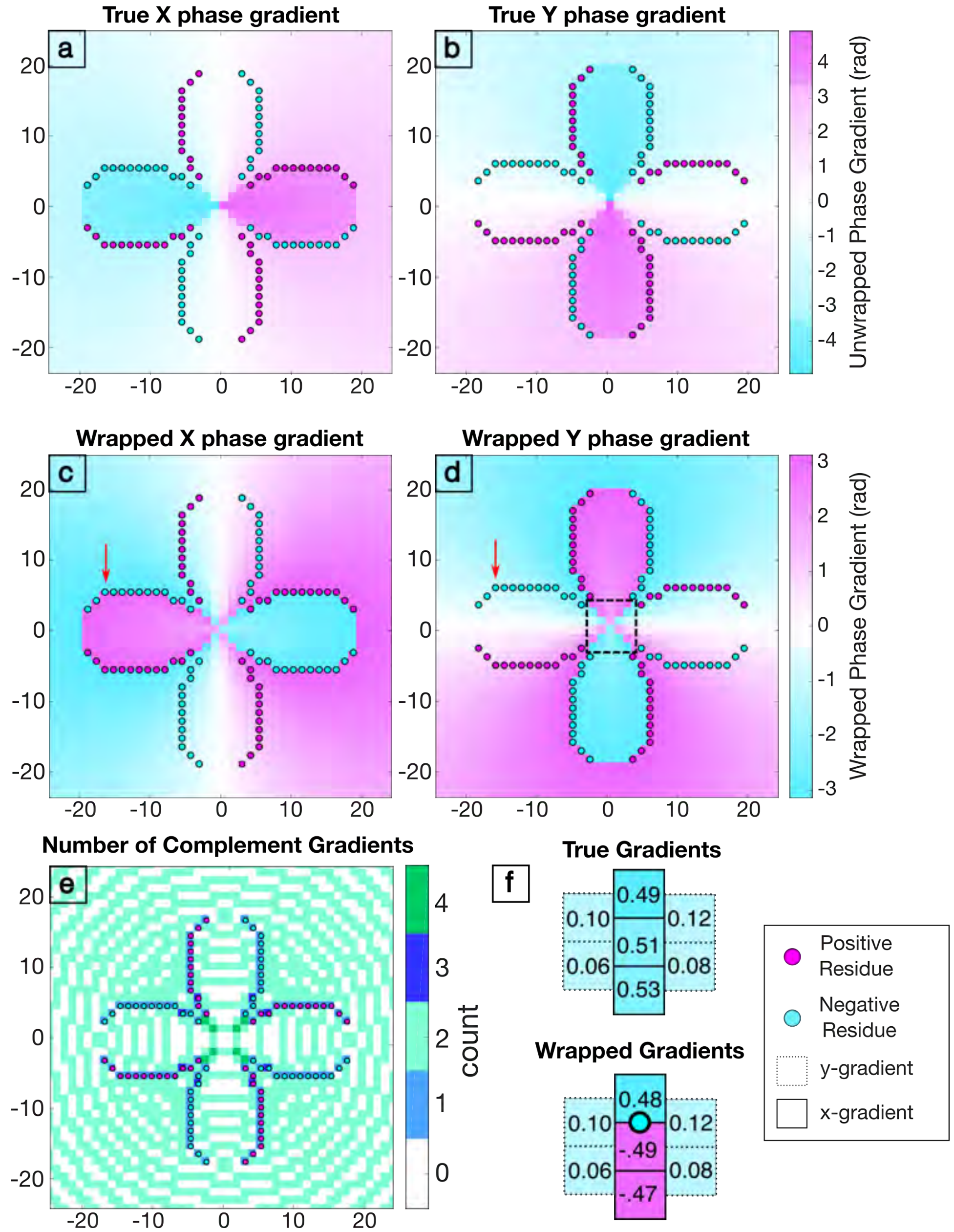


Figure 5.4: Residues on a noise-free surface. (a) and (b) show that residues appear along the boundaries of the lobes containing true x- and y-gradients, respectively, greater than $\pm\pi$ rad. The wrapped gradients (c and d) show a sign flip within these lobes. (e) Only an odd number of complement gradients in the residue calculation leads to a residue. (f) A residue occurs when there are an odd number of apparent complement gradients compared to the true phases, whether they were wrapped into or forced to be their complement. Within the black box in (d), there are high gradients but no residues because of the overlap in apparent complement gradients.

a complement gradient relative to the true unwrapped gradients as *apparent* complement gradients. We also note that additional abrupt sign-reversals would be observed if the true gradients continued to increase through $\pi + 2n\pi$, where n is an integer.

Why do residues only occur at the boundaries of these sign-reversals? As we saw in the last section, the formal criterion is that residues only appear when the number of complement gradients used in the calculation is odd. This is demonstrated in Figure 5.4e, where we show the number of complement gradients used to calculate each loop integral. Residues only appear at locations with odd numbers (blue), as expected.

Though the occurrence of a residue is well-defined, it may also be useful to think of this criterion as the number of apparent complement gradients in the pixel set. For example, Figure 5.4f zooms in on the first few gradient sets (expressed as number of cycles) below the red arrow in c and d. Moving vertically downward, a residue occurs when one gradient in the set of four appears to have been flipped to its complement. However, one step further down shows a set containing two apparent complement gradients and no residue. This version of the criterion makes it easier to visualize why residues only appear at the boundaries of each of the high-gradient lobes. Within the lobes, the gradients are self-consistent.

It is important to highlight that there is not a residue at every point along these boundaries. For example, the left edge of the high x-gradient lobe in Figure 5.4c has no residues because each column is independent of the others. Therefore, residues will only occur when there is an abrupt sign flip in the orthogonal direction to the given gradient (in this case, y). Similarly, we see no residues within the region surrounded by the black box in Figure 5.4d, despite having abrupt sign changes along the orthogonal direction (in this case, x). There are no residues here because the apparent complement gradients in both dimensions begin to overlap (and referring back to Figure 5.4e, the overlap coincides with loop integrals that require the use of all four complement gradients). In the next section, we will see how these residue placements influence the unwrapped solution.

5.4 The Origin of Unwrapping Slips

If we can precisely determine the conditions that lead to a residue, can we predict where exactly we will observe unwrapping slips given a displacement field? We claimed at the beginning of this chapter that aliasing will occur when gradients exceed ±π rad in closed loops. We return to our 2D exponential to develop a more-precise definition of a closed, high-gradient loop, and to determine the conditions that spark aliasing.

An obvious place to start is to assume that aliasing will originate when the combined gradient of the true displacement field exceeds π rad in a closed loop, where the combined gradient ($\Delta\psi^{xy}$) is defined as:

$$\Delta\psi^{xy} = \sqrt{(\Delta\psi^{(x)})^2 + (\Delta\psi^{(y)})^2}. \tag{5.4}$$

We show this potential aliasing boundary as the dashed black circle in Figures 5.5a and b, which display the true unwrapped x- and y-gradients, respectively, for the same 2D exponential phase field used in Figure 5.4. This boundary encloses the lobes with high gradients in both dimensions, but some of the individual gradients within this contour are adequately sampled regardless of dimension. Furthermore, the lobes that contain high x-gradients often have adequately-sampled y-gradients, and vice versa. Indeed, this is the advantage of considering the unwrapping problem in two dimensions. We can avoid discontinuities by leveraging the low gradients parallel to them. As we learned in the last chapter, this is accomplished by connecting balanced sets of residues.

Figure 5.5c and d show the wrapped x- and y-gradients with the correct residue connections given the true unwrapped phase gradients in a and b. The proper cuts connect the residue pairs within each of the high-gradient lobes. These cuts allow us to take advantage of the low gradients along the directions indicated by the arrows. Note, however, that the three inner-most cuts are not associated with any residues and they form closed loops. These three loops will definitely be aliased for two reasons. First, there is no indication that there are discontinuities that need to be accommodated and, second, enclosed cuts are an invalid tree formation in SNAPHU. We can thus conclude that aliasing will occur when the phase field requires a looped discontinuity cut.

We have placed the largest looped discontinuity in Figure 5.5a and b as the inner-most red box. It is reasonable to assume that this will be the aliasing origin, and all pixels within this loop will have cycle slips. However, we have also included the true aliasing boundary as the solid red rectangle, which clearly encompasses a larger area. The reason for this extension is the other general source of aliasing: the misplacement of cuts. Figure 5.5e and f show the residue-pair connections that SNAPHU selects based on the objective to minimize the cut lengths. Only some of the appropriate residue pair connections (c and d) are the shortest; at the eighth residue inward from the combined gradient contour, the cut lengths between oppositely-charged residues on adjacent lobes become shorter than the correct cuts within each lobe. Consequently, the cuts are misplaced and aliasing begins. The maximum aliasing limit would be when all of the cuts are misplaced, corresponding to the outer red box in Figure 5.5a and b. This box aligns with the first discontinuities from each direction.

We can now provide a precise definition of a high-gradient loop and formulate bounds on where aliasing will originate. A closed high-gradient loop specifically refers to a looped arrangement of true gradients that exceeds ±π rad ($\Delta\psi^{(x)} > \pi$ and $\Delta\psi^{(y)} > \pi$), similar to what is depicted in the simple schematic from Figure 5.1. We hereafter refer to this as a *looped discontinuity*, requiring an invalid cut formation to be accommodated. When there is a series of multiple looped discontinuities, the outermost contour will be the minimum aliasing boundary. The minimum number of cycle slips observed at a point within this enclosed region will be the number of looped discontinuities

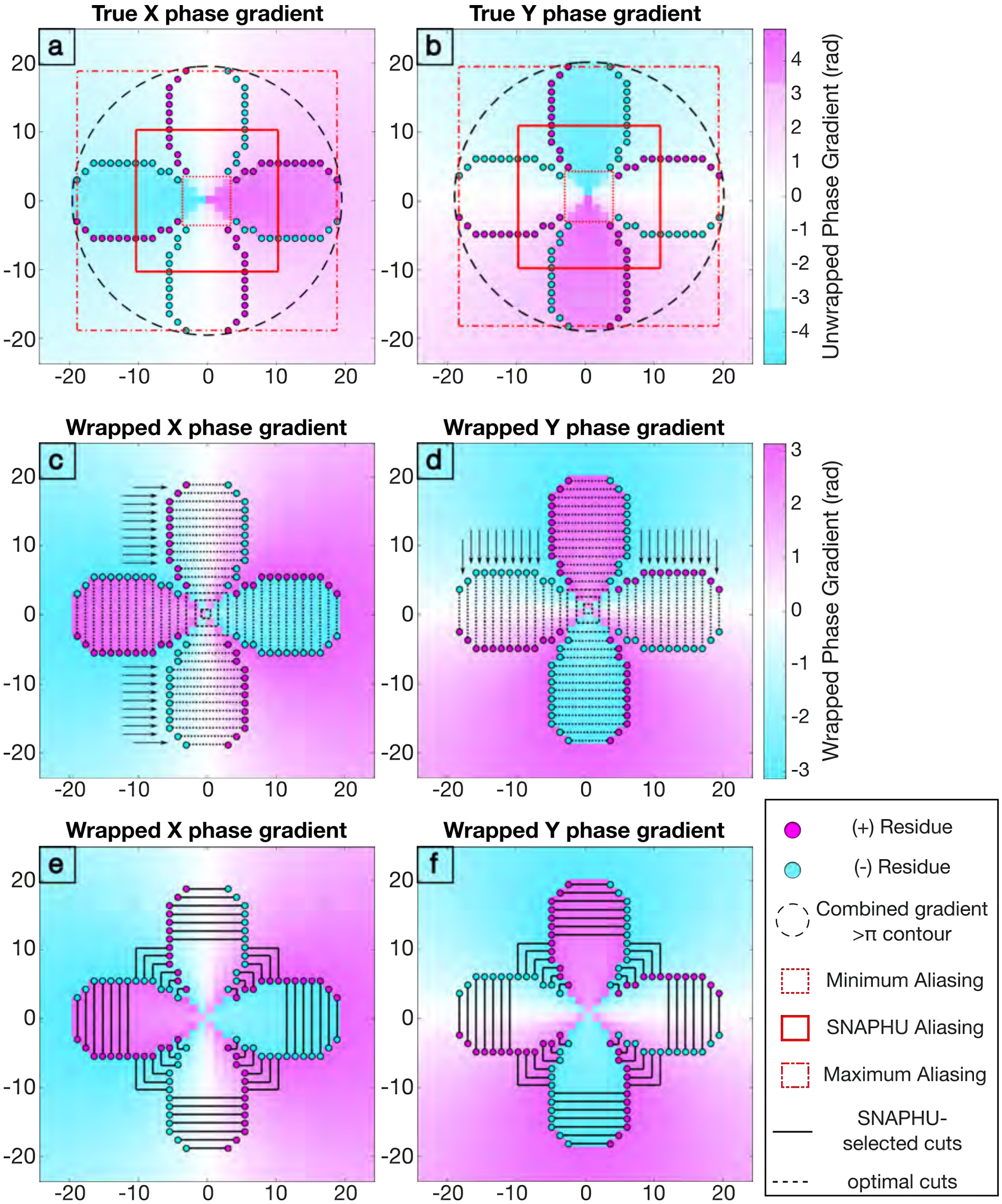


Figure 5.5: Controls on aliasing boundaries. (a) and (b) show the true x- and y- unwrapped phase gradient fields, respectively. (c) and (d) display the wrapped phase fields with the proper cut locations, whereas (e) and (f) show where SNAPHU truly places the cuts. The first cycle slip (solid red box in a and b) coincides with the first cut misplacements. When gradients are high in both directions there are no residues and the proper cuts form closed loops. This boundary (innermost-dotted red box in a and b) is the minimum aliasing contour. The maximum aliasing contour is the box corresponding to the misplacement of all cuts (outermost-dotted red box in a and b).

surrounding it. Often, the true aliasing origin will arise prior to the minimum boundary at the point where proper cut connections are no longer optimal in terms of the objective function. Theoretically, the maximum limit of the aliasing boundary extends up to the misplacement of all of the cuts, or the rectangle containing the true combined gradient $> \pi$ contour ($\Delta\psi^{xy} > \pi$-contour).

5.5 Wavelength Dependence

Now that we have a solid understanding of where and why aliasing may occur in the presence of contoured displacement patterns, we can turn our attention to factors that influence the gradients we observe in an interferogram. For example, the observed phase gradient for a given displacement field will depend on the system wavelength (λ). We saw in Chapter 3 that the surface displacement (r) in radar line of sight (LOS) is related to the interferometric phase (ϕ) by:

$$r = \frac{\phi\lambda}{4\pi}. \tag{5.5}$$

Likewise, the displacement and phase gradients (Δr and $\Delta\phi$, respectively) can be expressed as:

$$\Delta r = \frac{\Delta\phi\lambda}{4\pi}. \tag{5.6}$$

Substituting $\Delta\phi = \pi$ for the phase gradient limit that defines a discontinuity ($\Delta\psi_{lim} = \pi$ rad pixel^{-1}), we can see that the corresponding displacement gradient limit (Δr_{lim}) is $\lambda/4$ m pixel^{-1}. Given the direct proportionality between Δr_{lim} and λ, the maximum displacement gradient that unwrapping algorithms can unambiguously unwrap decreases with decreasing system wavelength.

The phase gradient limit leads to relatively stringent displacement gradient limits for common radar systems. An L-band system, with $\lambda = 23$ cm, has $\Delta r_{lim} = 5.75$ cm pixel^{-1}. However, at X-band, with $\lambda = 3$ cm, the Δr_{lim} reduces to only 0.75 cm pixel^{-1}. We demonstrate the wavelength-dependence in Figure 5.6 using a synthetic Mogi source (Mogi, 1958; commonly used to model point-source subsurface volume changes), projected into LOS. The maximum LOS displacement is 0.91 m and we use a native posting (pixel size) of 20 m. The wrapped interferograms for L-, C- ($\lambda = 5.6$ cm), and X-band systems are displayed in the top panel and the corresponding SNAPHU-unwrapped phases are in the bottom panel. It is clear from the wrapped phase fields that the phase gradients increase significantly with decreasing wavelength. Consequently, unwrapping leads to variable success. The L-band signal leads to perfect unwrapping, but at C-band, the maximum displacement is only 0.88 m (one cycle slip), and at X-band, the maximum displacement has dropped to a mere 0.37 m (19 cycle slips). The observed wavelength dependence is a direct result of the increased phase gradient with decreasing wavelength, for the same true displacement field.

Figure 5.6 also highlights how the imaging geometry affects gradients in the interferogram. A Mogi source produces radially-symmetric displacements at the surface, but with InSAR we are

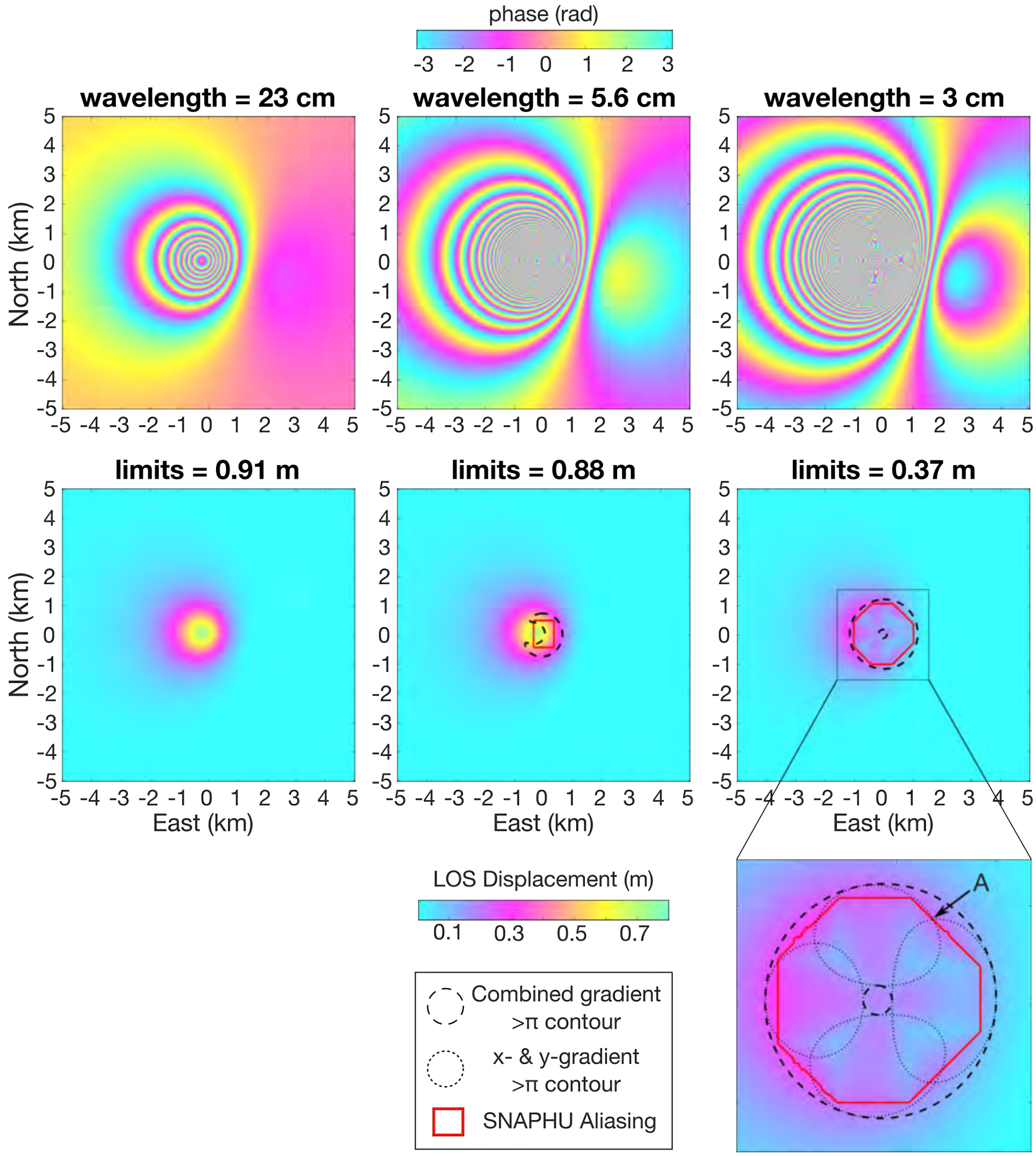


Figure 5.6: Aliasing dependence on system wavelength. A given displacement field will be imaged differently in terms of the number of fringe cycles as a function of the system wavelength (top row). The risk of aliasing increases with decreasing wavelength, as depicted in the middle row. L-band (23 cm) perfectly unwraps the displacement field, whereas C-band (5.6 cm) and X-band (3 cm) show ∼3 and ∼54 cm dropped, or 1 and 19 cycle slips, respectively. The bottom row zooms in on the gradient patterns for the X-band interferogram. The aliasing origin begins at the first instance where x- and y-gradients both exceed π (point A).

limited to seeing only the displacement component parallel to the vector connecting the point on the ground to the satellite (the LOS vector). This limitation can increase the gradients on one side of the Mogi source while decreasing the gradients on the other. Despite the asymmetry, the aliasing limits still apply. For example, the bottom panel of Figure 5.6 zooms in on the unwrapped displacements and adds the individual contours for x- or y-gradients exceeding π rad. The SNAPHU aliasing boundary aligns with the outermost-instance of overlapping high x- and y-gradients, as highlighted by point A.

The cycle slip observed at C-band is challenging to visually identify with the color scale limits we have chosen. On the other hand, the unwrapping slips at X-band are far more noticeable as clearly-defined geometric features. A seasoned InSAR user might recognize that this unwrapped solution has aliasing artifacts, but note that they still look quite different to the unwrapping Error demonstrated in Figure 4.10a. Rather than sharp, random jumps, the solution appears smooth, albeit with unlikely geophysical features. It is also important to highlight that we unwrapped these interferograms at their native posting (i.e., we did not apply any spatial filters or use any multilooking). Typical InSAR practice incorporates some degree of spatial filtering prior to unwrapping per the suggestion from Goldstein et al. (1988). Spatial filters reduce the number of residues in the interferogram by decreasing noise, and unwrapping efficiency is greatly improved. Furthermore, spatial filtering in the form of multilooking reduces data storage requirements, which is particularly beneficial when we have a large set of interferograms (as is common with modern satellites) and/or extensive study areas. As we will see next, obvious aliasing artifacts are less likely to appear when we apply spatial filters prior to unwrapping, in which case aliasing is more likely to be overlooked.

5.6 Resolution and Spatial Filtering

The discussion thus far has defined the phase gradients and the adequate-sampling limit in terms of arbitrarily-sized pixels. Given that a displacement field has fixed gradients in terms of some physical unit distance, the phase gradients we observe (or adequately sample) in the interferogram will directly depend on the size of a pixel. Suppose we form an interferogram at a native posting equal to the maximum resolution of the radar (δ_{p_0}; here we assume that the individual range and azimuth resolutions are identical). The maximum gradient that can be adequately sampled ($\Delta\psi_{max}$) will depend on δ_{p_0} by:

$$\Delta\psi_{max} = \frac{\pi}{\delta_{p_0}}. \tag{5.7}$$

If we elected to use a larger pixel size, then the maximum true gradient we can tolerate without ambiguity will clearly decrease. As such, common InSAR practice is to form the interferograms with a pixel size approximately equal to δ_{p_0}; however, interferograms at their native posting are typically noisy and/or extremely large. We can improve both of these problems with low-pass spatial filters.

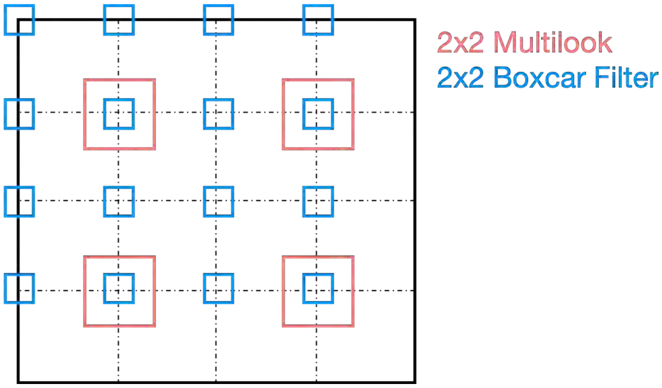


Figure 5.7: Multilooking versus a boxcar filter with 2×2 looks. Multilooking samples the average within every independent 2×2 set of pixels, reducing the number of pixels by the number of looks (2) in each dimension (pink boxes). The equivalent boxcar filter samples the average in every 2×2 moving window. A boxcar filter does not reduce the number of pixels (blue boxes). The colored boxes touch each original pixel included in the average.

In this section we consider two forms of filtering. The first is *multilooking*, which spatially averages each independent set of $n \times n$ pixels in the interferogram, where n is referred to as the number of looks. Figure 5.7 displays an example schematic of 2×2 multilooking (pink boxes). Each independent set of 2×2 pixels becomes one single pixel of size $2\delta_{p_0}$. Alternatively, we can use a moving window that averages every $n \times n$ set of pixels in the interferogram. This is commonly known as a moving average or a *boxcar filter*, and a 2×2 example is demonstrated by the blue boxes in Figure 5.7. The boxcar filter smooths the interferogram (reduces noise) without reducing the number of pixels, whereas multilooking reduces both noise and data storage requirements. Despite these benefits, both filters decrease the resolution of the interferogram and, consequently, the maximum physical gradient that can be adequately sampled; however, the effect that each has on aliasing differs slightly. We discuss these behaviors next, leaving the effects of other types of filters to future work.

5.6.1 Multilooking

SNAPHU defines unwrapping parameters (e.g. discontinuities, residues, gradients) by the phase difference between pixels, assuming that a pixel is equal to a resel (Chen, 2001). If we begin with an interferogram at its native posting, the maximum adequately-sampled gradient is $\Delta\psi_{max}$ (Eq. 5.7). When we multilook the interferogram by n pixels in both range and azimuth, the new pixel size becomes $\delta_{p_n} = n\delta_{p_0}$. Because each multilooked pixel represents an independent set of phase samples, we can treat each as a new resolution element; thus, the definitions and conditions we applied to discontinuities and residues before continue to apply here. The only change is that the new maximum adequately-sampled gradient reduces to:

$$\Delta\psi^{(n)}_{max} = \frac{\pi}{n\delta_{p_0}}, \tag{5.8}$$

where superscript (n) indicates the number of looks along each dimension. Thought of in a slightly different way, we will create a discontinuity if the averaging window is larger than half the width of a phase cycle at some point along its contour.

Figure 5.8 demonstrates the effects of multilooking on a noise-free LOS interferogram generated with a Mogi source (maximum displacement of 0.45 m). The top panel shows the wrapped phases with an increasing number of looks (1, 2, and 3, respectively), the middle row displays the number of complement gradients used to calculate residues (blue=odd=residue), and the bottom panel displays the unwrapped interferograms with displacement in meters. At the native posting (20 m), there are no residues (blue pixels) and we perfectly unwrap the interferogram. However, note that the green contours, which represent the contours of fringe cycles, are tightly-spaced near the middle of the image (i.e., gradients are high). When we take looks, the averaging window will exceed half the width of some of the tightly-spaced cycles, causing them to interfere with each other. This introduces residues (as shown in the middle panel), which indicates that we have discontinuities at the new resolution. However, SNAPHU cannot generate the required looped-discontinuities and we observe unwrapping slips. The aliasing origin (red) aligns with the contour delineating when the combined gradient exceeds $\pi/n\delta_{p_0}$ rad m^{-1}, as expected. As n increases, the largest magnitude in the unwrapped interferogram decreases. Furthermore, the aliasing origins are no longer polygonal (as was the case in Figures 5.5 and 5.6) and there are no longer obvious geometrically-shaped aliasing artifacts. When we introduce aliasing with filters, it is challenging to visually identify that we have underestimated the displacement magnitude.

5.6.2 Boxcar Filter

A boxcar filter reduces the resolution of an interferogram without changing the pixel size. Although the pixels are no longer independent phase samples (resels), the resolution is slightly better than the number of looks (n) (the resolution of the multilooked interferogram). Consequently, SNAPHU is tolerant of higher gradients than in the equivalent multilooked case. For example, when there are more than 1.5 phase cycles within an averaging window of size n, we will alias the fringe cycle. The corresponding gradient limit $(\Delta\psi^{(n)}_{upper-bx})$ is:

$$\Delta\psi^{(n)}_{upper-bx} = \frac{3\pi}{n\delta_{p_0}}. \tag{5.9}$$

We refer to this as an upper gradient limit for two reasons. The first is that the gradients will have already exceeded the adequate sampling bound for a filter size up to $n = 3$. That is, if the gradients exceed $\Delta\psi^{(2)}_{upper-bx}$ or $\Delta\psi^{(3)}_{upper-bx}$, they will have already exceeded 1/2 cycle (π rad) at the native posting. The second is that, empirically, we find that the the aliasing origin may begin at slightly

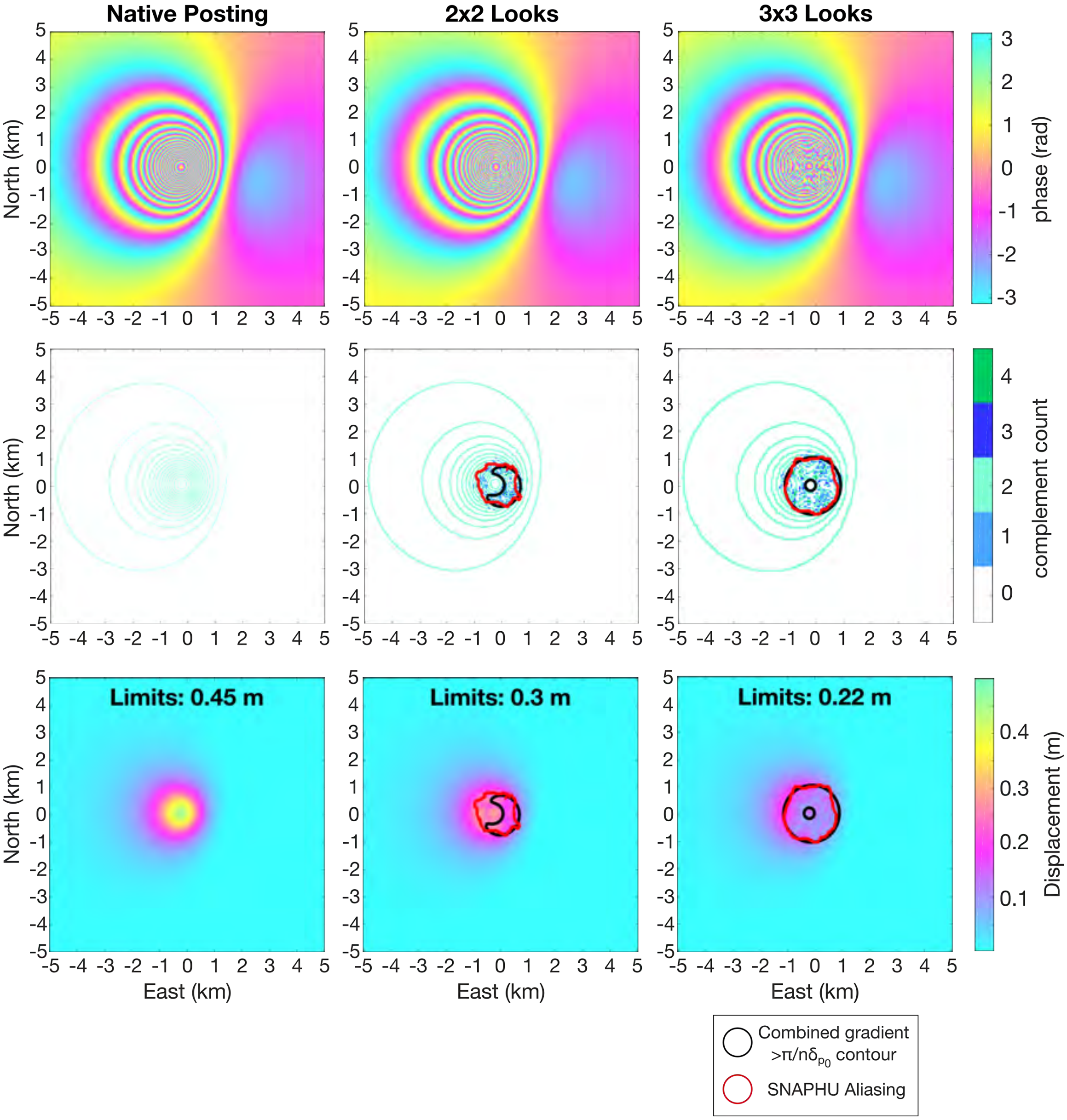


Figure 5.8: Aliasing dependence on multilooking. Multilooking decreases the resolution of an interferogram (top row), which reduces the physical phase gradient that defines a discontinuity. Consequently, the degree of aliasing increases with a higher degree of multilooking, as depicted in the bottom row. Though the original displacement field was adequately sampled, averaging over tight fringe cycles (middle row) introduces residues (blue pixels) and aliasing. The interferograms with 2x2 and 3x3 looks show ∼15 and ∼23 cm dropped, or ∼5 and ∼8 cycle slips, respectively. The aliasing boundary approximately aligns with the contour enclosing $\Delta\psi^{(n)} > \pi/n\delta_{p_0}$

lower gradients. We begin with the maximum adequately-sampled gradient for the multilooked case ($\Delta\psi^{(n)}_{max} = \frac{\pi}{n\delta_{p_0}}$) and include a scaling factor of $(2n\text{-}1)/n$. The lower gradient limit for a boxcar filter ($\Delta\psi^{(n)}_{lower-bx}$) is then:

$$\Delta\psi^{(n)}_{lower-bx} = \frac{(2n-1)\pi}{n^2\delta_{p_0}}. \tag{5.10}$$

The observed aliasing origin will begin when gradients first exceed a value between the lower and upper bounds. We summarize the gradient limits for multilooked and filtered interferograms in Figure 5.9. As the number of looks increases, the gradient tolerance (in terms of cycles/δ_{p_0}) decreases. Furthermore, it is clear that the boxcar filter is tolerant of slightly higher gradients in comparison to its multilooked counterpart, though this difference is only significant up to about fifteen looks or so.

We demonstrate the behavior of the boxcar filter for a Mogi source with a LOS maximum displacement of 0.4 m in Figure 5.10. We begin with a filter size of 2x2 and increase to the right to 4x4 and 10x10 (top row). The bottom row depicts the corresponding unwrapped displacements in meters. Our 2x2 filter does not induce aliasing; however, the combined gradients do exceed the multilooked gradient tolerance (solid black line) – had we mutlilooked this same interferogram, we would have induced aliasing. The 4x4 boxcar filter does induce aliasing along the contour depicting when the combined gradients begin to exceed the lower gradient tolerance (dotted black line), which

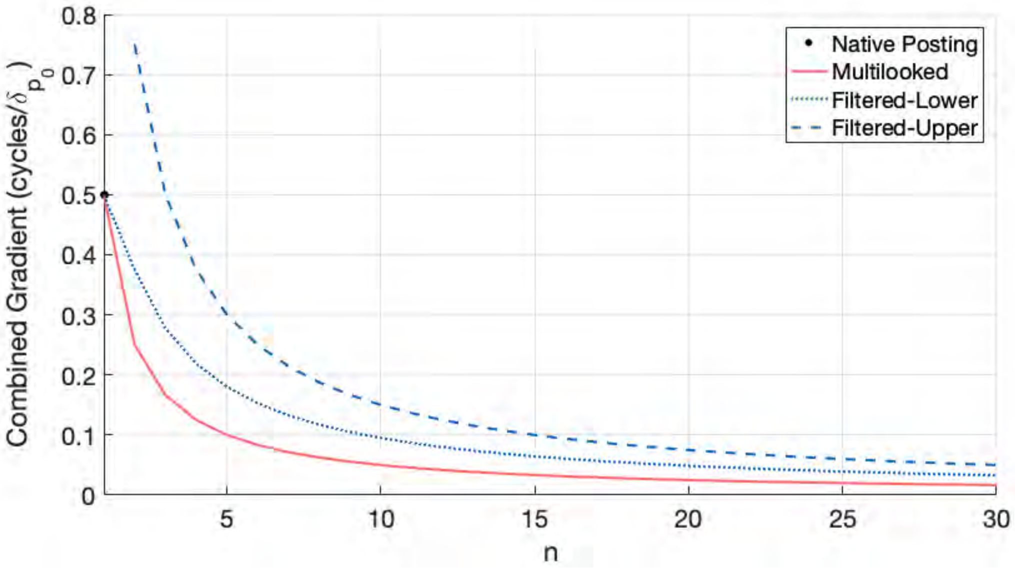


Figure 5.9: Gradient limits as a function of the number of looks (n). Filtering is tolerant of slightly higher displacement gradients than the equivalent multilooked case. Any contoured fringes with larger gradients than these thresholds will also be aliased at the corresponding number of looks.

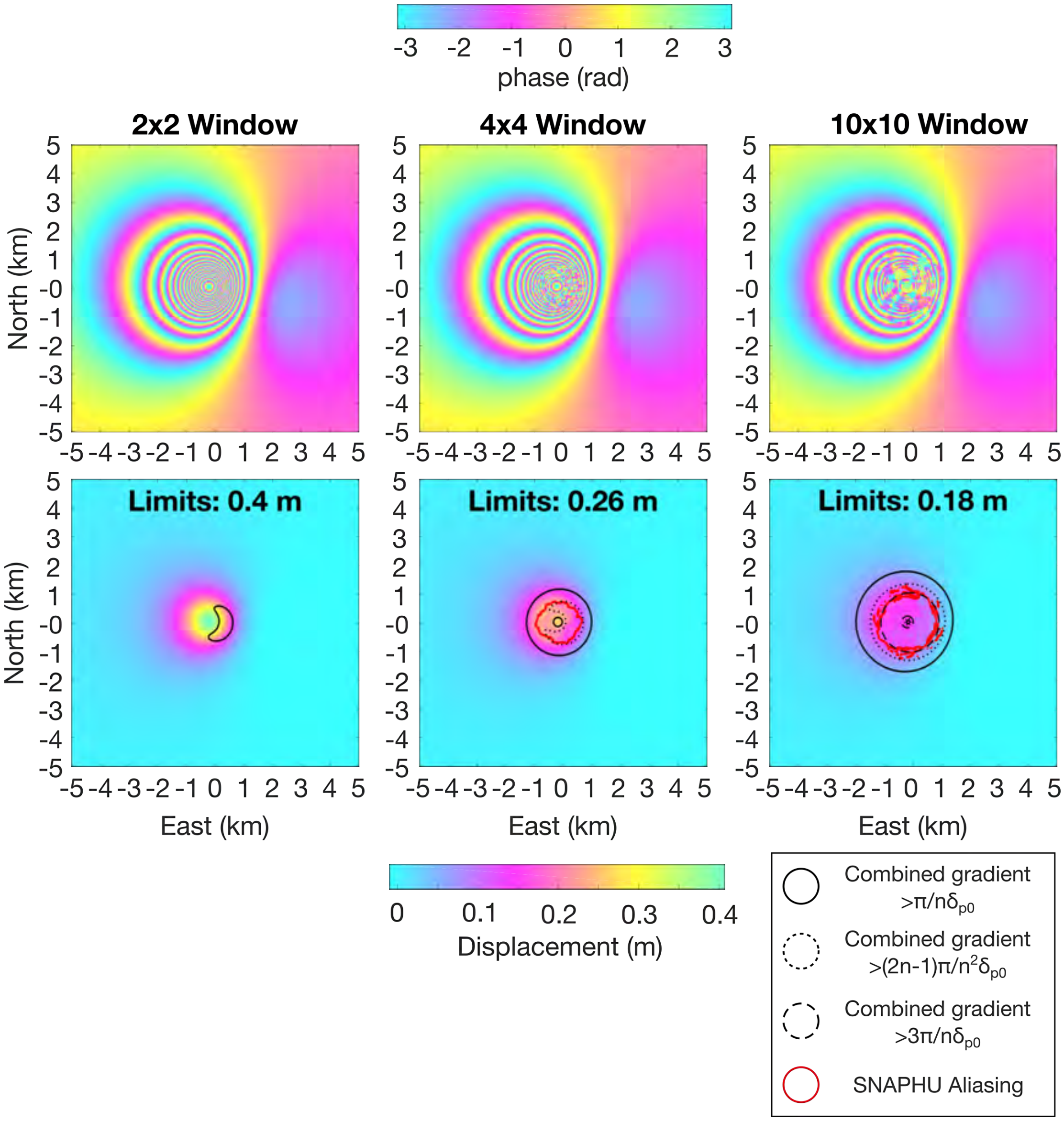


Figure 5.10: Aliasing dependence on a boxcar filter, which reduces the resolution of the interferogram while maintaining the original pixel size. The aliasing origin is associated with higher gradients (between the dotted and dashed contours) than in the equivalent multilooked case (solid black contour). As the filter size increases from left to right (top row), the number of cycle slips increases (bottom row).

is well within the multilooking bounds. When we increase our filter to 10x10 looks, the extents and magnitude of aliasing increases and the aliasing origin is contained within the boundaries of the lower and upper (dashed black line) contours. At a high number of looks, it appears that a little bit of error is introduced at the lower gradient limit and everything within the upper gradient contour will necessarily be aliased.

These gradient limits are not restricted to Mogi source deformation patterns. Figure 5.11 demonstrates a similar exercise for an Okada edge dislocation (Okada, 1985), which is used to model surface displacements caused by a slipping fault (see Chapter 7 for more details). Here, we observe both subsidence (up to -0.3 m) and uplift (up to 0.15 m), and use a diverging color scale for the unwrapped displacement fields (bottom row). Beginning with the interferogram at its native posting (20 m), we increase the filter size to 4x4 and 7x7 looks (top row). The aliasing origins demonstrated in the unwrapped interferograms follow the same gradient rules as the Mogi source example from Figure 5.10, so we will not belabor the point here. Instead, we note two important observations. Here, it is the magnitude of subsidence that is underestimated when aliasing begins, highlighting that a cycle slip refers to a loss in displacement *magnitude* (as opposed to a negative displacement becoming more negative after aliasing). Second, we do not alias the LOS uplift even though the highest uplift magnitude appears directly along the edge of the aliasing boundary. SNAPHU has identified an appropriate curvilinear discontinuity that allows perfect unwrapping from the direction of the arrows shown in the 7x7-filtered interferogram, at least up to the depicted line. This underscores that only the displacements within high-gradient contours are aliased, not large displacements themselves (though high gradients are usually accompanied by large displacement magnitudes).

5.6.3 Apparent Decorrelation

Figures 5.8, 5.10, and 5.11 show that the regions impacted by aliasing appear to be decorrelated in the corresponding wrapped, filtered images. That is, the fringe patterns become less distinct and the phases look more random as the number of looks increases. In a sense, we have decorrelated the images by filtering, as demonstrated by the introduction of residues in an adequately sampled interferogram at native posting (middle row in Figure 5.8). However, in contrast to true decorrelation which represents changes in scattering processes at the surface, this is "decorrelation" of our own making – it is *filtering-induced aliasing.*

If we were to multilook or filter an interferogram without observing the original (as is common practice), we may misinterpret aliasing as surface decorrelation and not recognize that it could have been avoided. This matters because it is often the regions with the highest displacement gradients that are the most interesting to study, thus the most important to not misrepresent with biased measurements. Next, we will explore added noise (the reason we apply filters prior to unwrapping at the expense of resolution in the first place) and how it may affect aliasing behavior.

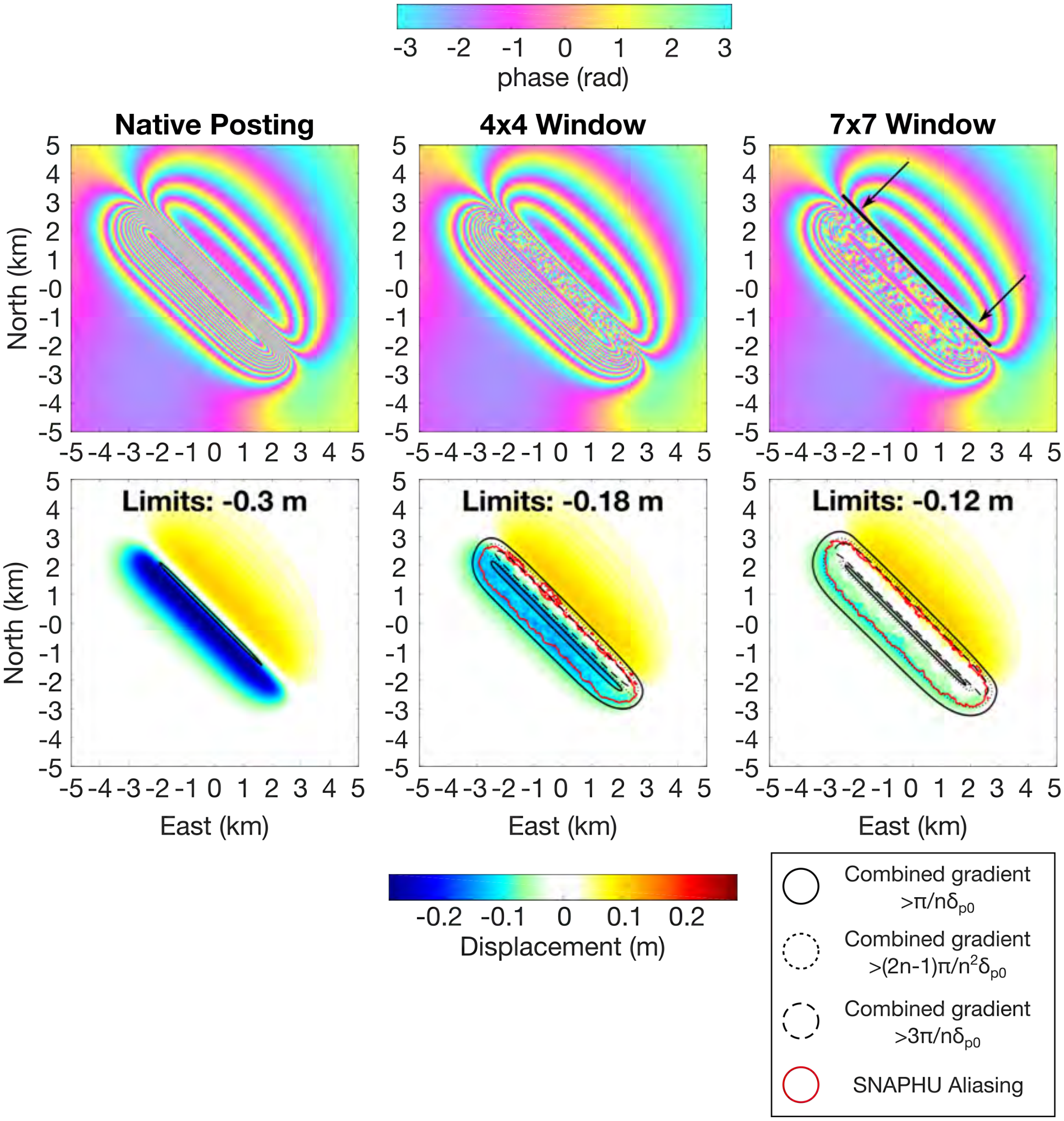


Figure 5.11: Filter-induced aliasing on an Okada edge dislocation. As the filter size increases from left to right (top row), the number of cycle slips increases (bottom row), where this time, cycle slips are associated with underestimated subsidence magnitude. The arrows in the upper right plot show that SNAPHU appropriately unwraps up to fault trace from one side; however, the contoured high gradients on the other side of this discontinuity are aliased.

5.7 Added Noise

The previous section highlights that aliasing can occur when we have contoured deformation patterns with high gradients in an interferogram, even in the absence of noise. Furthermore, common InSAR practices (multilooking or general spatial filtering) may induce aliasing because they reduce the image resolution and, consequently, the maximum gradient we can adequately sample. We apply these filters to reduce system and decorrelation noise, which are inevitable in real interferograms (see Chapter 3.3-3.4). These additional phase terms will impact the unwrapping process because they affect our ability to "resolve" the fringe patterns and true deformation gradients, even if they do not decrease the resolution of the image. That is, a resolution element still represents a measurement of the combined return phase from all scatterers within that resel, but adjacent measurements may no longer appear spatially-correlated. The major questions are how does noise impact aliasing and to what degree do spatial filters improve the unwrapping results? We explore these questions with system and decorrelation noise, which are mostly temporally- and spatially-random.

5.7.1 System Noise

We begin our discussion with system noise because the random phase terms at each pixel are drawn from the same distribution (i.e., the same "level" of noise affects the whole interferogram). We typically model this distribution as white Gaussian noise, the standard deviation of which is determined by the signal-to-noise ratio (SNR) of the system. Figure 5.12 demonstrates the effects of added noise on an interferogram of a Mogi source with a maximum LOS displacement of 0.65 m (top row). As SNR decreases to the right (40, 20, and 5 dB, respectively), the spatial correlation between adjacent pixels also decreases, which we interpret as noise. The gradients of the true displacement field are all less than $\Delta\psi_{max}$ (π/δ_{p_0}) and thus there are no true discontinuities that would need to be accommodated in the noise-free interferogram. However, the added noise creates local discontinuities, particularly in regions with gradients that are already close to the adequate sampling bounds. Near these high gradients, there is a higher density of residues, as shown in the middle row of Figure 5.12. Some of these discontinuities can be accommodated by SNAPHU; indeed, when SNR is relatively high (40 dB), we can perfectly unwrap the noisy displacements (corresponding bottom row). However, as SNR decreases, the noise in the measurements increases and SNAPHU can no longer determine the correct network of cuts to properly unwrap the interferogram within the regions enclosed by the red contours, which happen to overlap the highest gradients.

An important question to consider at this stage is whether or not the observed unwrapping errors are slips or Errors (with a capital E; see Chapter 4.6). Technically, we have adequately-sampled the gradients in terms of the necessary resolution to avoid what we have defined as aliasing – unaccommodated looped discontinuities or misplacement of cuts. On the other hand, the observed errors follow similar patterns to the *outcome* of aliasing. That is, there are integer cycle slips

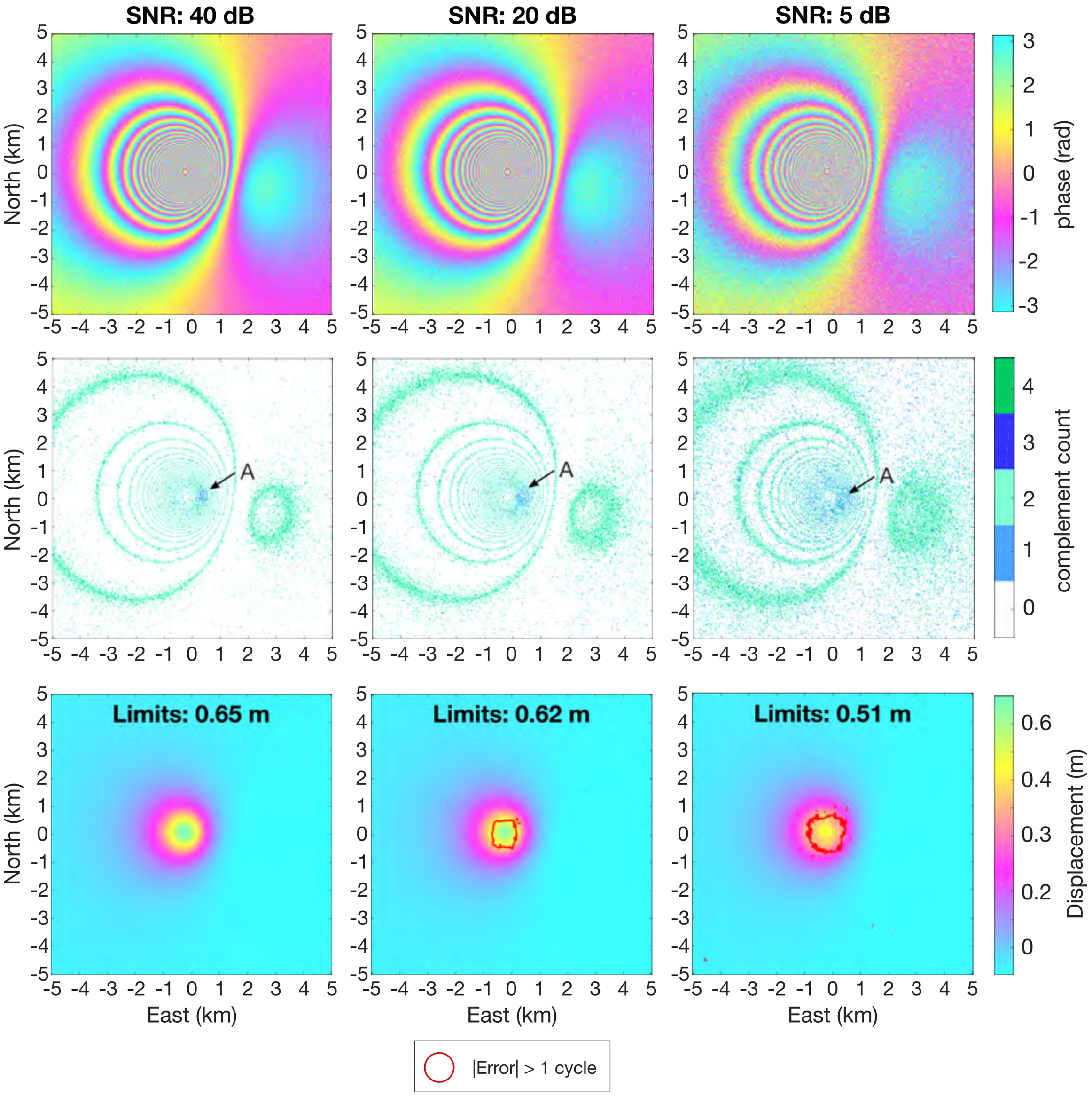


Figure 5.12: Mogi source with increasing system noise. A Mogi source interferogram with increasing added noise as SNR decreases to the right (top row). As a higher level of noise is added to the adequately sampled interferogram (with no noise), more residues are introduced (middle row), as shown by the odd complement gradients used in the residue calculation (blue pixels). Residue densities are highest in the regions indicated by point A in each plot. As SNR increases, so do unwrapping errors (bottom row), but only in regions with high gradients.

contained within an enclosed boundary, these errors propagate to all pixels within that boundary, and these global errors systematically underestimate the true displacements (i.e., they are biased).

These details are summarized in Figure 5.13, which in the top row shows the difference between the true and unwrapped displacement fields from Figure 5.12 as a function of the true gradients, all in terms of number of phase cycles. The bottom row displays similar plots for the noise-free interferogram filtered with a boxcar filter of $n = 2$ (left) and an aliased, noise-free interferogram from a Mogi source with gradients that exceed the adequate sampling bounds at the native posting (right). We have already established that the errors in the bottom row represent unwrapping slips. The filtering-induced aliasing displays a component of apparent random noise, a result from averaging pixels with non-uniform interferometric phase (note that the apparent noise increases with increasing gradient – or increasing difference between the phases in the averaging window). The errors in the top row mimic the behavior of the established unwrapping slips of the bottom right, combined with the added noise. We conclude that the noisy interferograms with SNR of 20 and 5 have *noise-induced aliasing*, not random displacement jumps that signify unwrapping Errors.

Will filtering improve the unwrapping of our noisy measurements? It depends. Filters will indeed reduce noise and improve the quality of unwrapping when the filter is appropriately-sized for the local gradients. However, a filter that would induce aliasing on the true, noise-free displacement field will induce the exact same aliasing behavior when there is noise, regardless of the local SNR. We demonstrate this in Figure 5.14, which shows the maximum unwrapped displacement for interferograms of a Mogi source with 9 cm (a) and 72 cm (b) maximum displacement, as a function of the number of looks in multilooking (pinks) and boxcar (blues) filters. Increasing the number of looks for the low-gradient field (a) benefits the quality of the unwrapping, up to about eight looks or so. A further increase in n induces aliasing, but the unwrapping limits show no dependence on the SNR of the interferogram (open circles); rather, they follow the same trend as the noise-free cases (filled circles). This is highlighted again in the high-gradient interferogram (b). The noise-free interferogram at native posting is not aliased, though aliasing is induced to a larger degree as SNR decreases (the multilooked examples overlap the boxcar filter examples, as we have not applied the filters here). However, two or more looks with either filter induces aliasing, and the noisy examples mimic the corresponding noise-free behavior.

5.7.2 Decorrelation Noise and Coherence

Decorrelation noise is typically associated with changes in scattering behavior on the ground surface (Zebker & Villasenor, 1992). Though it has been shown to have some spatial and temporal correlation (e.g., Michaelides, 2020; De Zan et al., 2015), it is still an added phase component with largely random behavior and has the ability to affect unwrapping in exactly the same manner as discussed in section 5.7.1. Therefore, we use this section to comment on the effects of spatially-variable noise levels due to decorrelation and common practices used to adapt our analyses to this varying signal

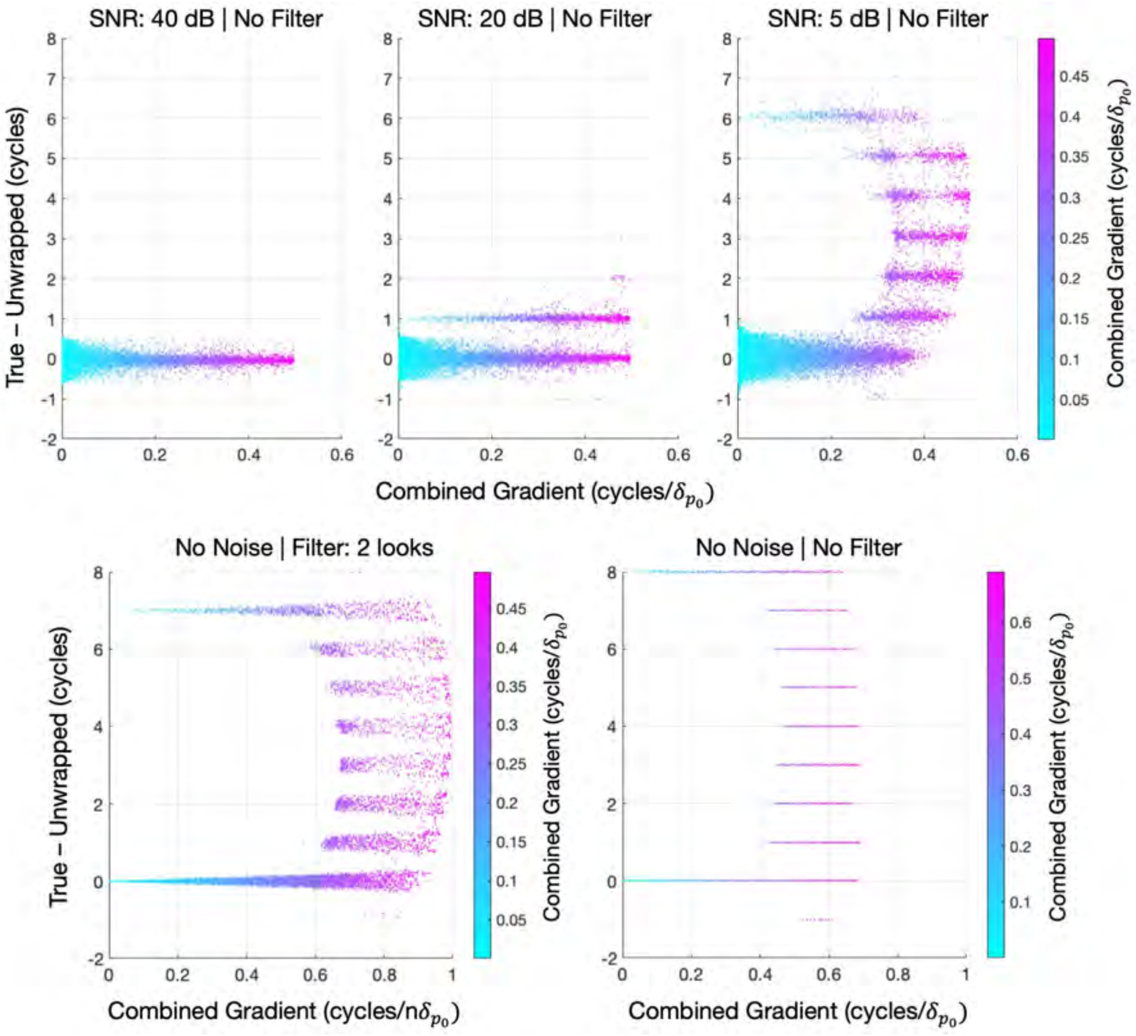


Figure 5.13: Unwrapping errors versus combined gradients for noisy, filtered, and noise-free interferograms. The top row displays the unwrapping errors from the interferograms in Figure 5.12. The bottom left displays the unwrapping errors for the noise-free version of the same Mogi source, filtered with a boxcar filter of size $n = 2$. The bottom right demonstrates aliasing errors from a noise-free, unfiltered interferogram.

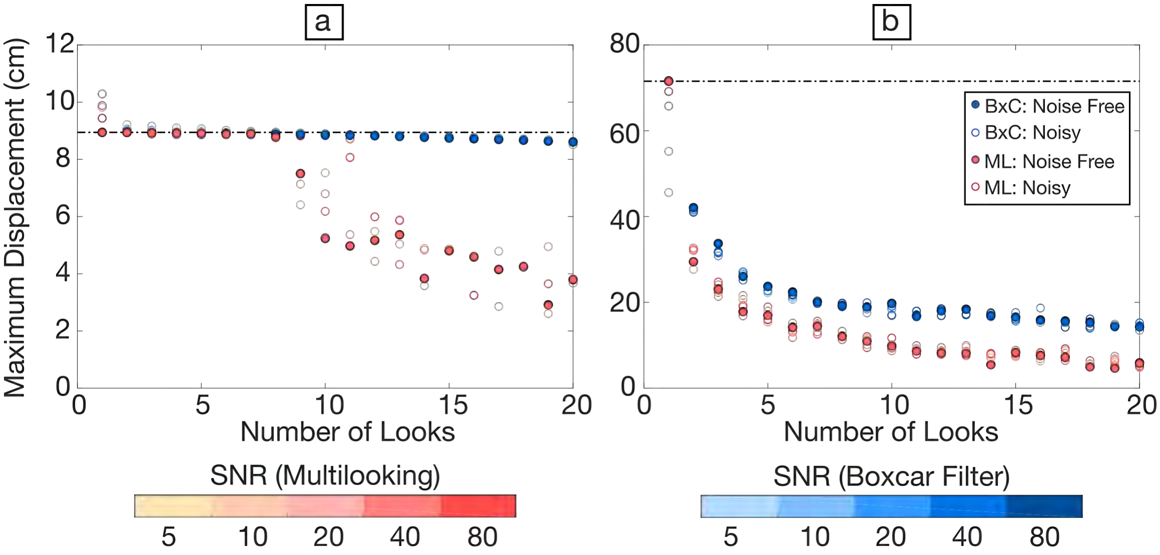


Figure 5.14: Filtering and multilooking-induced unwrapping with and without noise. A small displacement signal in (a) benefits from some filtering when there are noisy measurements, up to about 7 looks or so. More looks in the multilooking examples induces aliasing. For a signal close to its gradient limits at native posting (b), noise introduces aliasing even though we have adequately sampled the gradients at native resolution. No amount of filtering (boxcar or multilooking) can improve the unwrapping result, nor does noise impact the aliasing induced by a filter.

quality.

It is well known that surface characteristics influence the level of decorrelation throughout an interferogram. For example, highly-vegetated regions typically decorrelate to a greater degree than bare land, roads, or stable infrastructure (Zebker and Villasenor, 1992; Zheng and Zebker, 2017; De Zan et al., 2015). As such, some regions of an image consistently have high coherence, whereas others are systematically decorrelated. Because we don't want to misinterpret areas with noisy measurements, common InSAR practice is to mask out areas with low coherence in order to focus the analysis on the high-quality measurements. Alternatively, we can weight the importance or reliability of a measurement by its coherence, either in inverse modeling or time series generation (we discuss the latter in the next chapter).

There are a few potential problems with this practice in the context of phase unwrapping and aliasing. First, coherence measurements *require* some degree of spatial averaging, which we know has the potential to induce aliasing. Second, aliasing causes *global* errors in the unwrapped interferogram, which means that the error propagates to all pixels within the aliasing boundary, even if they have perfect coherence. Furthermore, spatial filters work remarkably well to decrease noise in regions with low spatial gradients, so low coherence does not necessarily mean the unwrapped phases have low quality.

Figure 5.15 demonstrates these potential problems using another Mogi source as an example.

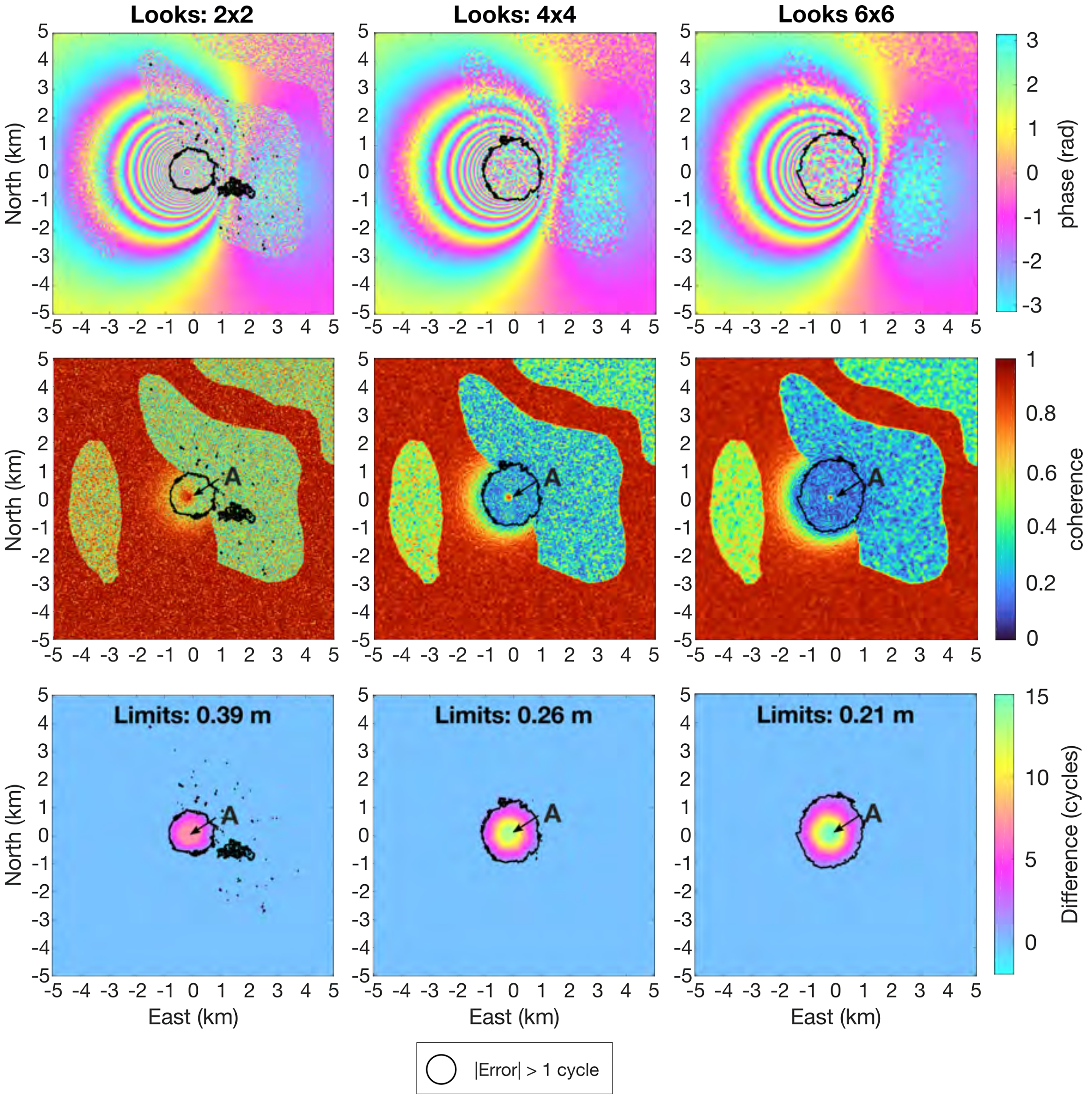


Figure 5.15: Spatially-variable coherence and aliasing. Top row: interferograms of a Mogi source with maximum LOS displacement of 0.58 m, with increasing multilooking window size to the right. Middle row: Coherence of the multilooked interferograms. Bottom row: Error between with unwrapped interferogram and the true displacement field in number of phase cycles. Black contours show regions with error greater than one phase cycle. Label A points to areas with relatively high coherence but large errors because they are within the aliasing boundary.

Here, the maximum displacement is 0.58 m and the gradients are adequately-sampled in the noise-free displacement field. The system SNR is 10 dB, and we create three highly decorrelated subregions with additional added random noise (SNR$<$ 2). The top row shows the interferogram with an increasing number of looks using a multilooking filter, the middle displays the coherence of the corresponding interferogram, and the bottom row depicts the error of the unwrapped interferogram relative to the true displacement field (in number of cycles). As we saw before, increasing the number of looks increases the spatial extent of the aliasing boundary and the magnitude of error for pixels within that boundary. Alternatively, the error decreases in the low-gradient regions outside of this boundary, highlighting that spatial filtering can improve unwrapping quality when appropriately sized for the local gradients. However, the coherence patterns shown in the middle plots do not adequately reflect the amount of error in the unwrapped interferogram. The pixels highlighted by label A, for example, maintain high coherence despite having the highest error. If we were to create a mask that only displays pixels with a coherence value above, say, 0.5, we would be left with displacement values with over six cycles of biased error. Alternatively, the regions outside of the aliasing boundary with low coherence become (almost) perfectly unwrapped. Should these be disregarded as bad measurements just because they have low coherence? We argue no.

The examples in Figure 5.15 highlight that there is an important distinction between measurement quality and unwrapping quality that needs to be made. Coherence estimators reflect an estimate of the similarity in phase between pixels within the local averaging window (Zebker and Villasenor, 1992; Zebker and Chen, 2005), which we can relate to an estimate of the phase standard deviation based on the number of looks used in the estimate (Lee et al., 1994). However, coherence estimators do not describe the similarity in averaged phase values between adjacent pixels, which is ultimately what drives the quality of unwrapping. Noisy measurements in low-gradient areas may be perfectly unwrapped (especially if filtered) relative to a reference pixel, whereas high quality measurements that are separated from the reference by any aliasing boundary may have the highest errors. Coherence reflects the quality of wrapped phase measurements, *not* the quality of the unwrapped phase field.

5.8 Chapter Summary

This chapter has demonstrated the conditions that lead to aliasing during 2D phase unwrapping with SNAPHU. Aliasing is driven by arrangements of high gradients that require looped discontinuities to accommodate (an invalid cut formation in SNAPHU) or the misplacement of cuts connecting balanced sets of residues. Unwrapping errors resulting from aliasing are really unwrapping slips, as they are biased toward underestimating the true displacement field by integer phase cycles. Unfortunately, these slips can occur even in the absence of noise and they propagate to all pixels within the enclosed aliasing boundary, regardless of the local measurement and/or unwrapping quality. We also described the factors that influence the gradients we measure in an interferogram: the imaging wavelength, the application of filters that decrease resolution, and added noise.

Given the large errors we have shown in our examples, it may appear as though we have a snaphu on our hands – that everything we have been doing with InSAR analysis up to this point has been riddled with biased errors. This is likely not true; however, these examples and observations highlight that the methods we use to enhance and analyze interferograms need to be appropriate for the local study area, including the wavelength of the radar and the local noise relative to the true displacement field. It is also important that we develop methods to detect aliasing. This is extremely challenging when we do not know what the true displacement field is supposed to be, which is often the case. Indeed, for episodic events such as large earthquakes, a comparison between InSAR and ancillary geodetic data such as GPS may be the only way to detect whether phase unwrapping is a success. However, we may be able to avoid aliasing and/or leverage information directly from the interferograms when we generate InSAR time series over dynamically-deforming regions. The next chapter discusses the consequences of including aliased images in InSAR time series and how we may be able to directly detect and avoid aliasing in these circumstances.

Chapter 6

Aliasing in InSAR Time Series

Chapter 5 describes the conditions of the displacement field that lead to aliasing during 2D phase unwrapping with SNAPHU (Chen, 2001). Given that aliasing produces global errors that systematically underestimate displacement magnitudes, it will come as no surprise that the inclusion of aliased interferograms in a time series analysis will also produce biased results. This chapter explores the consequences of aliasing in InSAR time series, with a strong focus on Small BAseline Subset techniques (Berardino et al., 2002).

Our work demonstrates that aliased interferograms systematically bias the SBAS time series solution, most often leading to an underestimation of the true displacement magnitudes. Furthermore, the set of aliased interferograms show distinct correlation with the temporal, seasonal, and spatial trends of the displacements, in addition to surface properties. Therefore, we can often avoid aliasing with the selection of a proper subset of interferograms to use in the time series analysis, though this is a challenging task because this subset is unique to each study area and may even vary pixel-by-pixel.

We begin this chapter with a discussion of SBAS and what conditions motivate the use of a large set of interferograms in its calculation, despite having an increase in computational load. We then demonstrate the effect that aliasing errors have on SBAS time series and discuss the correlation of these errors with time. Throughout this chapter, we use synthetic data to demonstrate aliasing concepts in the context of SBAS time series, and to support our interpretation of real Sentinel-1 data over three study areas: Hawaii's Kilauea volcano, the Delaware Basin in West Texas, and the San Joaquin Valley in California.

6.1 SBAS

SBAS techniques use many interferograms over a single area to develop InSAR time series (Berardino et al., 2002). The total number of interferograms (M_{max}) we can generate from N SAR scenes is:

$$M_{max} = \frac{N(N-1)}{2}. \tag{6.1}$$

Even in the absence of aliasing, this complete set of interferograms likely contains both good and bad measurements, most often dependent on the degree of decorrelation. The fundamental idea in SBAS is to reduce the total population of interferograms to smaller subsets with short spatial- and temporal baselines (distance and time between the satellite acquisitions forming the interferogram, respectively), with the hope that we have enough measurements to reduce random noise, while excluding extremely decorrelated values that no longer reflect the nature of the displacement field. Though historic satellites require us to consider the spatial baseline between the satellites on their repeat passes, the orbits on modern systems (such as Sentinel-1) are so tightly controlled that we need only consider the temporal baseline, referred to here as TB. General practice is to select a maximum TB (TB_{max}) and use all interferograms with $TB \leq TB_{max}$ in the subset.

Once we have selected the subset of interferograms, we calculate the linear velocity between SAR scenes at each pixel individually, using the relationship between *unwrapped* phase measurements and time. This linear relationship can be described by:

$$\begin{bmatrix} \psi_{12} \\ \psi_{13} \\ \psi_{23} \\ ... \\ \psi_{(N-1)N} \end{bmatrix} = \begin{bmatrix} TB_{12} & 0 & 0 & ... \\ TB_{12} & TB_{23} & 0 & ... \\ 0 & TB_{23} & 0 & ... \\ ... & ... & ... & ... \\ 0 & 0 & ... & TB_{(N-1)N} \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{23} \\ v_{34} \\ ... \\ v_{(N-1)N} \end{bmatrix}, \tag{6.2}$$

where ψ_{ij} is the unwrapped phase measurement between times t_i and t_j ($i = 1, 2, ..., N-1$; $j = 2, 3, ..., N$; $i < j$), $(t_j - t_i) \leq TB_{max} \leq (t_N - t_1)$, $TB_{(i-1)i}$ represents the temporal baseline between adjacent SAR acquisitions ($t_i - t_{i-1}$), and $v_{(i-1)i}$ is the line-of-sight (LOS) displacement velocity between temporally-adjacent SAR acquisitions. As we saw in Chapter 3.6, we can represent this linear relationship in its vector notation as:

$$\psi = Gv. \tag{6.3}$$

We compute a model of the linear velocities between each step in the time series using the generalized pseudo-inverse of G ($G^{\dagger}$). We then generate the displacements by integrating the resulting velocity vector, often assuming zero displacement at t_1.

6.2 SBAS with Independent and Redundant Interferograms

Figure 6.1 depicts an example SBAS time series at a single pixel in an arbitrary displacement field (blue) compared to its true displacement (black). Random, zero-mean noise should be evenly-distributed (approximately) about the true displacements; however, calibrating the beginning of the time series to zero at t_1 may cause an offset (red). We ignore this constant shift for the time being.

Let us first assume that we generated the blue time series in Figure 6.1 using only data from temporally-adjacent interferograms, i.e., subset $[\psi_{12},\ \psi_{23},...,\psi_{(N-1)N}]$. This is an *independent* set of interferograms, meaning that no time frame is spanned by more than one image. Consequently, the unwrapped phase measurement (ψ_{mes}) between each single-look-complex (SLC) acquisition date includes the true ground motion ψ_{def} between t_{i-1} to t_i with all of the potential sources of noise, assuming that the displacement field has been perfectly unwrapped:

$$\psi_{mes_{(i-1)i}} = \psi_{def_{(i-1)i}} + \psi_{atm_{(i-1)i}} + \psi_{decor_{(i-1)i}} + \psi_{n_{(i-1)i}}, \tag{6.4}$$

where subscripts *atm*, *decor*, and *n* indicate the atmospheric, decorrelation, and system noise, respectively (for simplicity, we ignore potential DEM and orbital errors, as these are typically small and systematic). It is common to assume that all of the added noise terms are random in time, whereas only ψ_{decor} and ψ_n are are approximately spatially random as well. We focus first on the temporal component of random noise because we calculate the SBAS solution at each pixel,

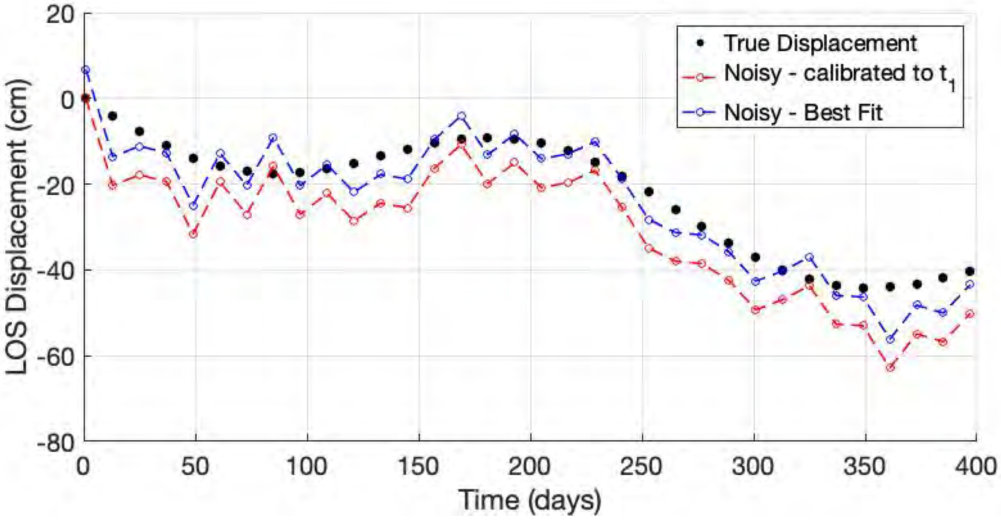


Figure 6.1: A noisy displacement time series at one pixel. We would like to estimate the true displacement (black), assuming zero displacement at time t_1. The true noisy displacements (blue) should be evenly distributed about the true time series. Note, however, that calibrating the noisy time series to zero displacement at t_1 typically causes a slight constant offset (red).

independent of all others in the image.

Because we only have one measurement over each subsequent time span, the displacement time series (d) solution from SBAS using only independent data ($d^{(ind)}$) is equivalent to the cumulative sum of the unwrapped phases between adjacent interferograms at each time step T:

$$d_T^{(ind)} = \sum_{i=2}^{T} v_{(i-1)i}^{(ind)}(t_i - t_{i-1}) = \sum_{i=2}^{T} d_i - d_{(i-1)} = \sum_{i=2}^{T} \psi_{mes_{(i-1)i}}, \tag{6.5}$$

where $T \in 1, 2, ..., N$, and N is the number of SLC acquisitions.

Can we do better than $d^{(ind)}$? Common SBAS thinking and practice suggests that the answer is "yes". It is believed that including more interferograms beyond the independent set reduces temporally-random noise and, consequently, the blue time series in Figure 6.1 will approach the black samples. For this theory to strictly hold, the added noise in each interferogram (not just the SLCs) must be: (i) statistically independent and (ii) drawn from a zero-mean distribution. The key to understanding the benefits (or consequences) of including more interferograms is to characterize the noise with respect to these two criteria, while also considering the increase in computational load as the number of measurements rises.

6.2.1 Unwrapped Closure Phase

Each additional interferogram we add to Eq. 6.2 is not an independent measurement of the displacement; rather, each represents some degree of *redundancy*. We define a redundant interferogram as one formed between non-adjacent SLCs. That is, $\psi_{(i-p)i}$ is redundant if $p > 1$ and independent if $p = 1$. Phase $\psi_{(i-p)i}$ is considered redundant because it can be reproduced by summing the independent interferograms between time t_{i-p} to t_i, if the measurements are perfectly correlated. We can determine the degree of correlation by considering the unwrapped phase closure between phase triplets sharing three common SLC acquisitions (ψ_{ij}, ψ_{jk}, and ψ_{ik}), regardless of their strict designation of redundancy or independence:

$$\psi_{clos_{(ijk)}} = \psi_{ij} + \psi_{jk} - \psi_{ik}. \tag{6.6}$$

Equation 6.6 will evaluate to zero when there is perfect correlation between the unwrapped phase triplet set. Note that this construction is exactly parallel to the idea of *closure phase* using wrapped interferogram phase triplets (see Chapter 3.4) in the study of decorrelation and soil moisture (e.g. Agram and Simons, 2015; De Zan et al., 2015; Michaelides et al., 2019; Zheng and Zebker, 2017). These studies have demonstrated that the wrapped closure phase between non-multilooked interferograms is identically zero everywhere (Samiei-Esfahany et al., 2016), suggesting that the set of independent measurements completely characterizes the observations. However, spatial-filtering introduces a random component in each interferogram that often leads to non-zero wrapped closure phase. This non-correlated phase is likely reflective of the decorrelation noise (Michaelides et al., 2019) and, if unwrapping is perfect, should also appear in the unwrapped interferograms (with a possible 2π ambiguity, like any phase term). A non-zero unwrapped phase closure suggests that the redundant measurements contain some information that cannot be fully characterized by the independent set.

The implications of correlated measurements in SBAS are rather straightforward: If the additional interferograms added to Eq. 6.2 are perfectly correlated to their triplet counterparts (i.e., $\psi_{clos_{(ijk)}} = 0$), then the inclusion of ψ_{ik} provides no additional benefits to using only ψ_{ij} and ψ_{jk}, and the independent SBAS time series will be equal to the redundant one ($d^{(ind)} = d^{(red)}$). If, however, there is some non-correlated component of ψ_{ik} with respect to the sum of the other two, then the independent and redundant time series will be different. The degree to which they differ will depend on the level of non-correlated noise, which is not equivalent to the level of noise from each non-displacement term in Eq. 6.4. This concept is best understood by decomposing each phase measurement not into its signal and noise terms (as in Eq. 6.4), but rather into its correlated and uncorrelated components, referred to here as *static* and *incidental*, respectively.

6.2.2 Static and Incidental Phase Terms

The static term refers to the phase components that are consistent between redundant triplet sets. That is, the sum of ψ_{ij} and ψ_{jk} will be equal to ψ_{ik}, thus Eq. 6.6 evaluates to zero and our estimate of the phase difference between times i and k is unchanging. Static terms include the displacement, in addition to sources of noise such as the atmospheric phase screen and non-random decorrelation that affects all scatterers within a resolution element equally (Emardson et al., 2003; Zwieback et al., 2016; Agram and Simons, 2015; De Zan et al., 2015; Michaelides et al., 2019). In addition to static phase terms, spatial filtering may introduce an inconsistent – or random – component that leads to a non-zero unwrapped closure phase. So as to differentiate this random phase term from the random noise in an SLC (e.g. atmosphere, decorrelation, system, etc.), which may be static, we refer to this component as incidental (ψ_{inc}). Assuming perfect unwrapping, we can summarize the measured unwrapped phase as a sum of the static (*stat*) and incidental (*inc*) terms:

$$\psi_{mes} = \psi_{stat} + \psi_{inc}. \tag{6.7}$$

The SBAS velocity solution can also be separated accordingly,

$$v = G^{\dagger}(\psi_{stat} + \psi_{inc}), \tag{6.8}$$

and v can thus likewise be decomposed into its static and incidental terms:

$$v = v_{stat} + v_{inc}, \tag{6.9}$$

where

$$\begin{aligned} v_{stat} &= G^{\dagger}\psi_{stat} \\ v_{inc} &= G^{\dagger}\psi_{inc}. \end{aligned} \tag{6.10}$$

Noting that v_{stat} will be identical whether we use the $N-1$ independent interferograms or include additional redundant terms (for a total of M measurements), the redundant SBAS time series ($d^{(red_M)}$) is:

$$\begin{aligned} d_T^{(red_M)} &= \sum_{i=2}^{T} v^{(ind)}_{stat_{(i-1)i}}(t_i - t_{i-1}) + \sum_{i=2}^{T} v^{(ind)}_{inc_{(i-1)i}}(t_i - t_{i-1}) + \sum_{i=2}^{T} v^{(red_{M-(N-1)})}_{inc_{(i-1)i}}(t_i - t_{i-1}) \\ &= \sum_{i=2}^{T} v^{(ind)}_{stat_{(i-1)i}}(t_i - t_{i-1}) + \sum_{i=2}^{T} v^{(M)}_{inc_{(i-1)i}}(t_i - t_{i-1}) \end{aligned} \tag{6.11}$$

See Appendix C for full derivation. The superscript *ind* instead of M in v_{stat} highlights that this term is not affected by the inclusion of redundant interferograms.

If the incidental noise terms are drawn from a zero-mean distribution, which is often the case, then the incidental velocities will approach zero as M increases, and Eq. 6.11 shows that the “best” SBAS solution we can achieve is the integrated static velocity terms. Therefore, the inclusion of redundant measurements is necessarily driving the solution to approach the true displacement time series; rather, it is driving the solution to approach the sum of the displacement and static noise terms, which may or may not reduce the error with respect to the true displacement.

The fact that redundant interferograms only change the SBAS solution when there is incidental noise has important implications for what we should expect to gain by trading computational efficiency for incidental noise reduction. For example, we cannot expect to reduce atmospheric noise, which may be on the order of 10 centimeters, due to its characterization as a static phase term (Emardson et al., 2003; Zheng, 2019). Rather, we likely need to rely on other methods, such as atmospheric corrections or inversion regularizations, to enhance the displacement signal. Note that this contradicts the behavior of atmospheric noise in weighted averages (or stacks), which has been shown to reduce to ±2 cm (Zebker et al., 1997). Atmospheric noise is random in time, thus the weighted average of these noise terms should approach zero as N (the number of SLCs) increases (Zheng, 2019).

On the other hand, we can expect to decrease the influence of decorrelation noise, as this is the main contributor to incidental noise (if the interferograms have been spatially filtered). If decorrelation is extremely low, then incidental noise may be negligible enough to use only the independent phase set in SBAS, increasing computational efficiency. However, if decorrelation is high, then incidental noise may be significant enough to warrant the use of redundant measurements. This is parallel to the findings in Zheng (2019), which demonstrates a similar cost-benefit behavior for interferometric stacks: decorrelation noise in the stack reduces as M increases, but only for study areas with low and/or rapidly decreasing decorrelation as a function of temporal baseline.

6.2.3 Aliasing as an Additional SBAS Term

It is clear that the decision to include redundant interferograms in SBAS – and how many (i.e. selecting an appropriate TB_{max}) – is a complicated balancing act. This task is made even more challenging when we also consider the potential for imperfect phase unwrapping. Thus far, we have assumed that the static and incidental phase terms come from the components on the right-hand side of Eq. 6.4, namely the deformation (displacement), atmosphere, decorrelation, and other temporally-random noise. However, there is potentially an additional phase term driven by unwrapping errors, including the biased, aliased measurements we overview in Chapter 5. We avoid making assumptions at this time about whether these terms are static or incidental, and include unwrapping errors (aliased or not) as a separate term in our final representation of the SBAS time series:

$$d_T^{(M)} = \sum_{i=2}^{T} v_{stat_{(i-1)i}}^{(ind)}(t_i - t_{i-1}) + \sum_{i=2}^{T} v_{inc_{(i-1)i}}^{(M)}(t_i - t_{i-1}) + \sum_{i=2}^{T} v_{unw_{(i-1)i}}^{(M)}(t_i - t_{i-1}) \tag{6.12}$$

where subscript unw is the unwrapping error.

If either the random noise or unwrapping errors are biased, which we know to be the case in aliasing, then the SBAS time series will be as well. However, this bias is relative to the static component of the solution. Therefore, a bias in SBAS is only "bad" if it affects the quality of the solution relative to the signal-of-interest (the displacements). For example, if the static terms lead to a biased time series relative to the true displacement, then an opposite bias from either the incidental or unwrapping error phase terms may actually improve the quality of the solution. We will discuss this trade-off in further detail when we show Sentinel-1 observations in Section 6.4.

We summarize this section by stating the (potentially) obvious: the benefits gained from increasing the number of phase measurements in SBAS depends first on the quality of the static time series with respect to the true displacements, then on the magnitude and statistical distribution of the incidental noise and unwrapping error terms. Before we explore the inevitable trade-offs such a set-up entails, we will first demonstrate how aliasing errors will affect SBAS solutions in the absence of noise and discuss the correlation of aliasing with time. This correlation will be important to consider when selecting TB_{max}, or more broadly, the set of phases to use at each pixel.

6.3 Inclusion of Aliased Measurements in SBAS

We learned in Chapter 5 that, given a known system wavelength, aliasing during 2D phase unwrapping depends first on the true displacement gradients. Each SLC pair views a unique snapshot of these gradients as the displacement field evolves over time. Therefore, there is reason to suspect that the risk of aliasing has an important relationship with the temporal baseline (TB) between the SLC pairs.

For example, a Mogi source that loses volume (contracts) at a constant rate over time will produce a surface displacement field with higher gradients as time increases. Interferograms formed from SLCs with short TB will view smaller gradients than those with longer TB, and as gradients increase, so too does the risk of aliasing. Here, if aliasing begins at $TB = TB_{thresh}$, then all interferograms with $TB \geq TB_{thresh}$ will also be aliased, likely with larger errors as TB increases. Because of this temporal correlation, we find that increasing the maximum temporal baseline (TB_{max}) used in SBAS systematically decreases the cumulative displacement magnitude in the presence of aliasing. We demonstrate this effect in Figure 6.2 with a cumulative displacement field (top left) from a noise-free Mogi source (viewed in radar line-of-sight (LOS)) contracting at a linear rate. Labels B, C, and D highlight pixels for which we generate multiple SBAS time series by systematically increasing the TB_{max} (see color scale).

The time series at Pixel B (top right) can be perfectly reproduced with any TB_{max}; there are no unwrapping errors at this point in the study area. The time series for Pixels C and D (bottom panel), however, show no such consistency. Pixel C demonstrates a systematic decrease in cumulative displacement magnitude as TB_{max} increases beyond ∼400 days, culminating in an error of 50% when we use the highest TB_{max} (936 days). This bias is a direct result of including aliased measurements. For example, the top plot of Figure 6.3a demonstrates the error between the unwrapped and true displacements at Pixel C, for each measurement as a function of TB. Recalling that these are noise-free interferograms, the errors beginning at TB=400 reflect biased, unwrapping slips due to aliasing, which increase in magnitude as TB increases. The increase in error is consistent with our expectation that the gradients will be higher over larger time spans. We also note that if these were real interferograms, we may expect aliasing to begin earlier than a TB of 400 days and/or larger errors due to higher decorrelation noise as TB increases (see noise-induced aliasing in Chapter 5.7).

The inclusion of aliased measurements not only influences the time series displacement magnitudes, but the temporal behavior as well. Figure 6.2c shows that the true linear rate of subsidence becomes increasingly non-linear as we include more aliased measurements by increasing TB_{max}. Consequently, aliasing may lead us to erroneously interpret and explain temporal patterns that are not truly occurring.

Pixel D demonstrates a similar behavior to Pixel C, though to a lesser degree. We include this example to highlight that the time series displays signs of aliasing despite having less cumulative

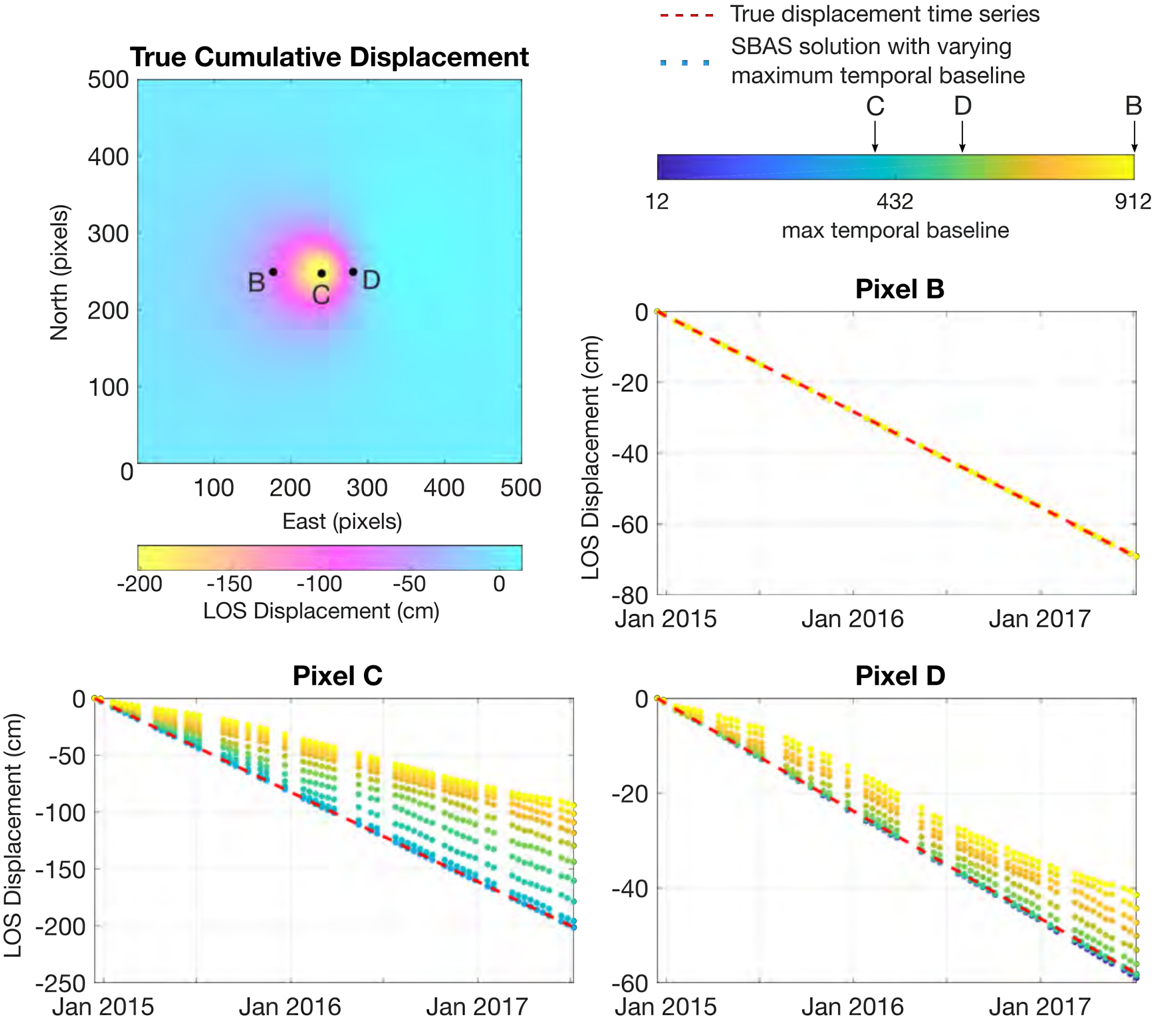


Figure 6.2: Time series for a linearly-deforming Mogi source. The cumulative displacement field is shown in the upper left plot. The other plots display multiple time series for pixels B (top right), C (bottom left), and D (bottom right). The time series are colored by the maximum temporal baseline used in the SBAS calculation. The dashed red line in each plot shows the true displacement at each point.

magnitude than Pixel B, which is perfectly reproduced with any TB_{max}. This disparity emphasizes that gradients (as viewed by the imaging geometry) control aliasing, not the magnitude of displacement.

The example from Figure 6.2 suggests that there is a simple solution to the problem of aliasing; namely, select a TB_{max} that is sufficiently low to exclude aliased measurements. If we categorize the unwrapped phases as either *consistent* (agrees with the true phase to within one cycle) or *inconsistent* (has more than one cycle of error), then there is indeed a strict TB cutoff separating consistent and

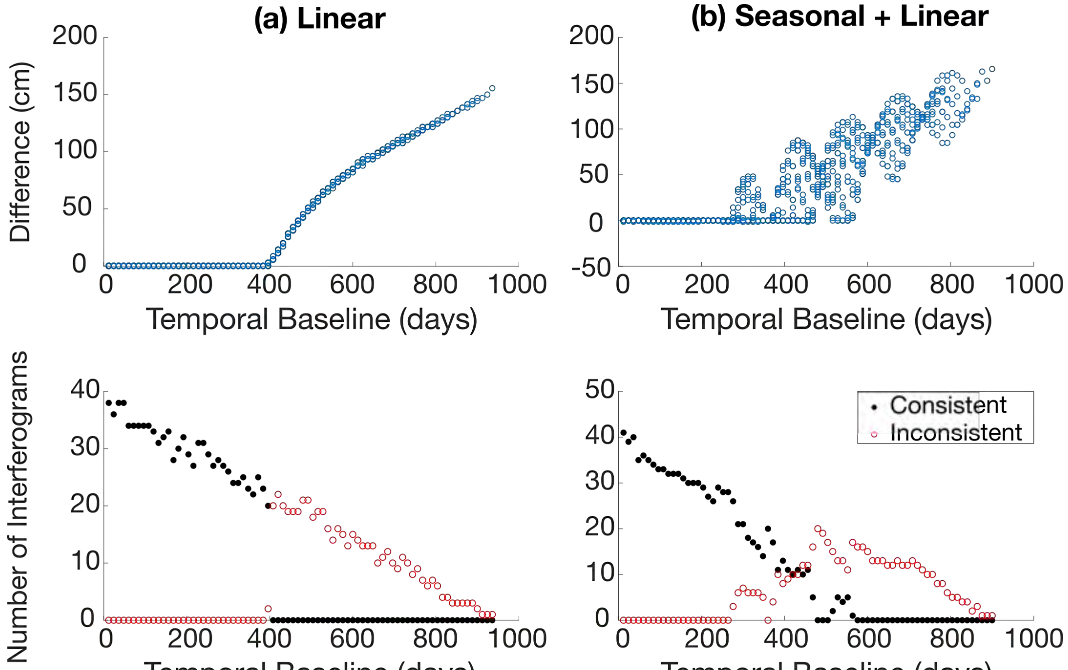


Figure 6.3: Consistent versus inconsistent interferograms in SBAS time series for a Mogi source with linear (left) and seasonal (right) behavior. (a) the difference between the unwrapped and true displacements at pixel C in Figure 6.2 (top) and the number of consistent versus inconsistent interferograms for the same pixel (bottom). (b) shows the same for pixel C from Figure 6.3. A consistent interferogram (black) is within one cycle of the true displacement, whereas inconsistent (red) is a measurement outside ± one cycle.

inconsistent phase measurements for the linearly-deforming Mogi source. This is demonstrated in the bottom plot of Figure 6.3a. If we select a $TB_{max} < 400$ days, there will be no bias in the SBAS time series. However, it is rare that we have perfect linear rates of displacement, which make our selection of the proper TB_{max} a little more complicated.

Figure 6.4, for example, demonstrates the time series for a Mogi source that has a seasonal component in addition to a linear subsidence trend. Pixels B, C, and D demonstrate the familiar systematic decrease in cumulative magnitude, but in addition, the amplitude of the seasonal signal is also dampened by the inclusion of strongly aliased measurements. Another difference between the linear and seasonal examples is that the seasonal signal does not have a strict cutoff between TB below and above which there are only consistent or inconsistent measurements, respectively. Rather, as demonstrated in Figure 6.3b, there is a transition zone between a TB of 250 to 600 days in which there is a mix of consistent and inconsistent measurements.

It may be obvious to select a TB_{max} of 250 days in the seasonal example, since all of the noise free measurements below this threshold are perfectly correlated (indeed, recall that we can actually

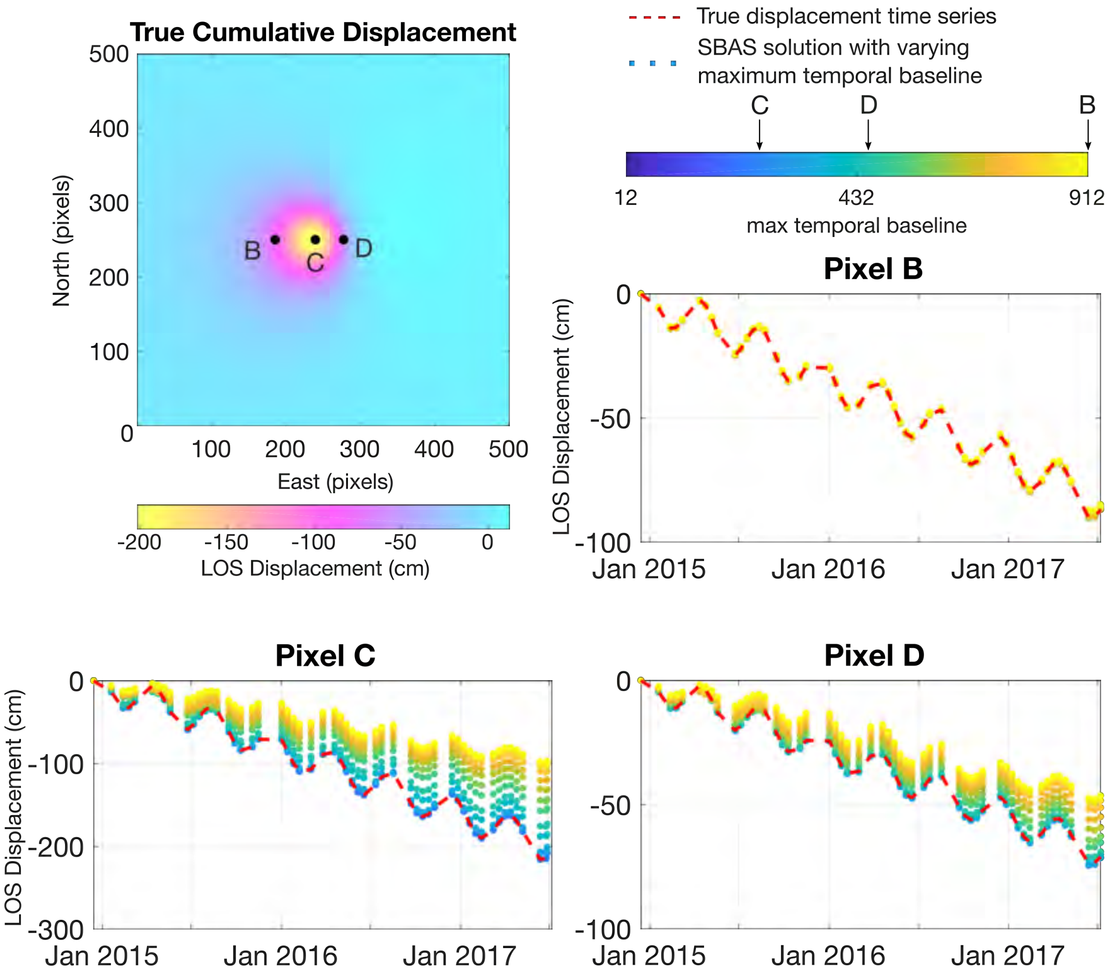


Figure 6.4: Time series for a deforming Mogi source with a seasonal component and linear trend. The cumulative displacement field is shown in the upper left plot. The other plots display multiple time series for pixels B (top right), C (bottom left), and D (bottom right). The time series are colored by the maximum temporal baseline used in the SBAS calculation. The dashed red line in each plot shows the true displacement at each point.

reproduce the time series with only the independent interferograms when they are noise-free and perfectly unwrapped). However, the motivation to include redundant measurements is to reduce the random incidental noise present in each interferogram. If there is incidental noise and the time series has more complex behavior than a simple linear trend with a seasonal component, it may not be clear which TB_{max} provides the best balance between noise reduction and aliasing avoidance. To further complicate matters, some bias from aliasing may actually correct biased noise contained in the independent or static solutions. We demonstrate the complexity of this balancing act in the next section with real Sentinel-1 data.

6.4 Sentinel-1 Observations

Our observations of aliasing and its systematic effect on SBAS time series is not limited to synthetic cases. Indeed, we were driven to explore this topic based on observations in three study areas using Sentinel-1 data. We begin with the Kilauea volcano deformation event in 2018, then demonstrate aliasing in the Delaware Basin, a large oil and gas field in west Texas, and finish our examples in California's Central Valley.

6.4.1 Methods

We follow the same basic steps to generate the SBAS time series in each study area. We first form all available interferograms at approximately 3.75 m x 15 m posting (pixel size) within a selected site-specific maximum temporal baseline (Kilauea - 360 days; Delaware Basin - 900 days; Central Valley - 1000 days). After multilooking and unwrapping with SNAPHU (Chen and Zebker, 2001), we calculate and remove the linear trend between the unwrapped phases and surface elevation from each interferogram individually, in an effort to reduce the dry atmospheric noise. In Kilauea and the Central Valley, we select a reference pixel corresponding to the location of a GPS station in the study area, whereas in the Delaware Basin, we do not calibrate the phases to any reference pixel, and instead use the the de-trended set of interferograms, as is. We then calculate multiple SBAS solutions in each study area by systematically increasing the TB_{max} beyond the study area's minimum TB_{max}, which is the largest temporal baseline in the independent (temporally-adjacent) set of interferograms.

Though there are modest differences between SBAS results using different degrees of multilooking, we restrict the discussion here to a single pixel size at each study area, allowing us to focus on the effects of TB. In Kilauea, we take 8 looks in range and 2 looks in azimuth (~ 30 m^2 pixels); in the Delaware Basin we use 60 looks in range and 15 looks in azimuth (~ 225 m^2 pixels); and in the Central Valley, we use 52 range looks and 13 azimuth looks (~ 195 m^2 pixels). The decision for each set of looks is largely driven by the size of the study area, where a larger area requires more multilooking to make data storage manageable. We leave a detailed discussion about the effects of multilooking (or general spatial filtering) to future work.

In Kilauea and the Central Valley, there are many GPS stations to which we can compare the SBAS time series. The Delaware Basin has a few GPS stations, but as described in Chapter 7.3, they are located in non-deforming regions and we leave our analysis of InSAR versus GPS to that section. We obtain the GPS data from the Nevada Geodetic Laboratory's repository and generate the GPS time series by first projecting the east, north, and vertical components into radar line-of-sight (LOS) at each time step, then smoothing the data with a 10-day moving window. After selecting a reference station and subtracting its time series from each of the other stations, we calibrate each time series to zero displacement at the date of the first SLC acquisition in the respective study area. In Hawaii and the Central Valley, we use the pixel containing the reference station as the reference pixel in the InSAR analysis.

6.4.2 Kilauea Volcano, HI

Our first example of aliasing comes from the 2018 volcanic event at Kilauea volcano in Hawaii. Though the surface displacements here are not due to anthropogenic well activity (the focus of this thesis), the region has an abundance of GPS measurements to which we can compare our InSAR time series. Furthermore, as an extreme deformation event, we can use it as an example of rapid and widespread subsurface volume changes – a pseudo end-member case – to which we can compare our other study areas.

Figure 6.5 demonstrates the effect of increasing the TB_{max} by showing the cumulative displacement in 2018 as estimated from SBAS using 12-, 60-, and 360-day TB_{max}. The volcanic event included a caldera collapse at the western-most LOS subsidence feature (motion away from the satellite). As TB_{max} increases, this feature spatially shrinks, as does the cumulative magnitude. In contrast, there appears to be noisy, but relatively small, displacements to the north in the 12-day TB_{max} result, whereas the other two solutions demonstrate widespread, smoothly-varying subsidence. Furthermore, the subsidence observed between the two strongest features actually increases slightly in magnitude and becomes more widespread. Without ancillary information, it would be challenging to know which image reflects the true displacement field best. Given the reduction in apparent decorrelation noise, we may favor the solutions with longer TB (e.g. 60 day).

The 12-day solution from Figure 6.5 shows the location of GPS stations. We compare the InSAR time series to the labeled red stations in Figure 6.6. We have calibrated the GPS stations to zero displacement at t_1 and calculated the shift in SBAS time series that leads to the best fit to the GPS data (in a least-squares sense). First, we consider the time series results at stations HOLE and UWEV, which do not demonstrate signs of aliasing. Each shows almost identical time series regardless of the TB_{max}, including the sharp, random jumps between each time step, which likely reflect atmospheric noise. As a static term, the atmospheric noise will not reduce as we increase the number of measurements, as shown here. The slight variations between the different results reflects the influence of incidental noise; however, these influences are so small in comparison to the static noise terms that an increase in TB_{max} does not produce a significantly better time series with respect to the GPS.

The other time series in Figure 6.6 are a little more interesting, as they demonstrate the clear influence of aliasing. Stations OUTL, CRIM, and BRYL show a distinct systematic decrease in cumulative displacement magnitude as the TB_{max} increases, exactly as we saw in the synthetic example from Figure 6.2. Though station OUTL experiences 1 m of displacement in the span of five months, we are able to generate a quality time series with the shortest TB_{max}. All others begin to introduce aliased measurements and lead to increasingly biased results. Note, however, that each

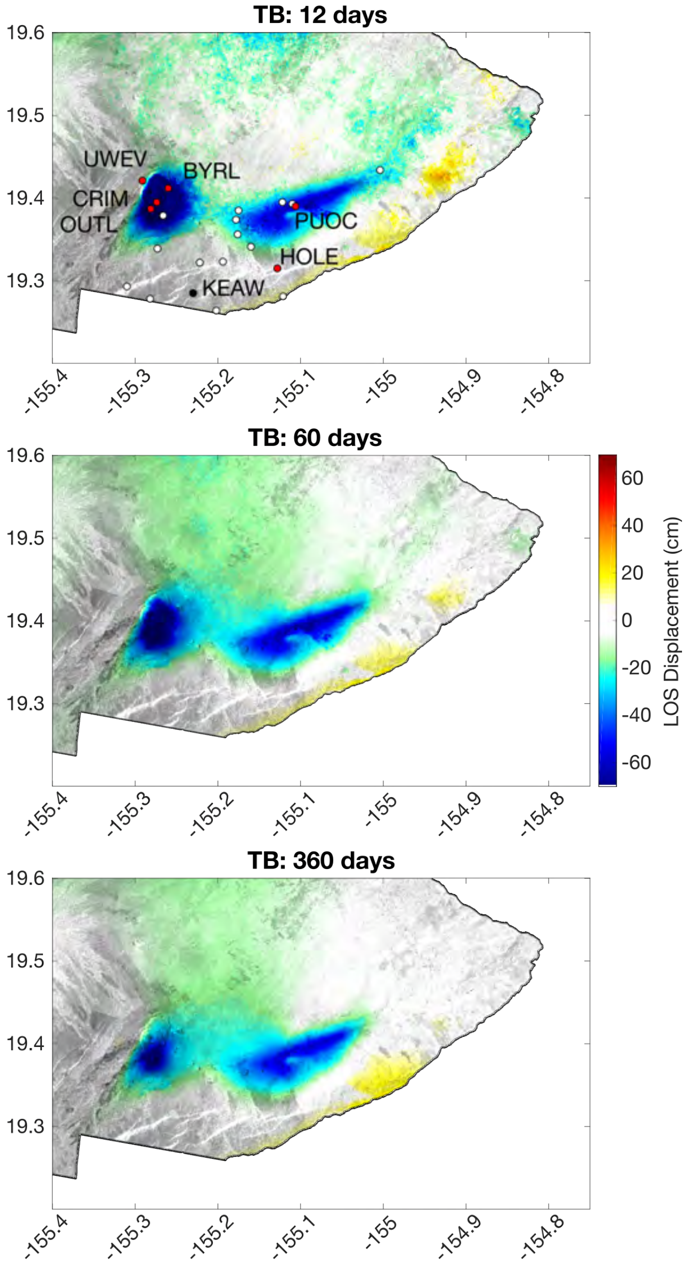


Figure 6.5: InSAR at Kilauea Volcano - Descending LOS. The magnitude and extent of motion away from the satellite changes with increasing maximum temporal baseline (TB). Sample GPS stations for time series (red) are labeled, whereas additional GPS stations are shown in white. All stations and InSAR are relative to zero displacement at KEAW (black), and the GPS station data is projected into LOS.

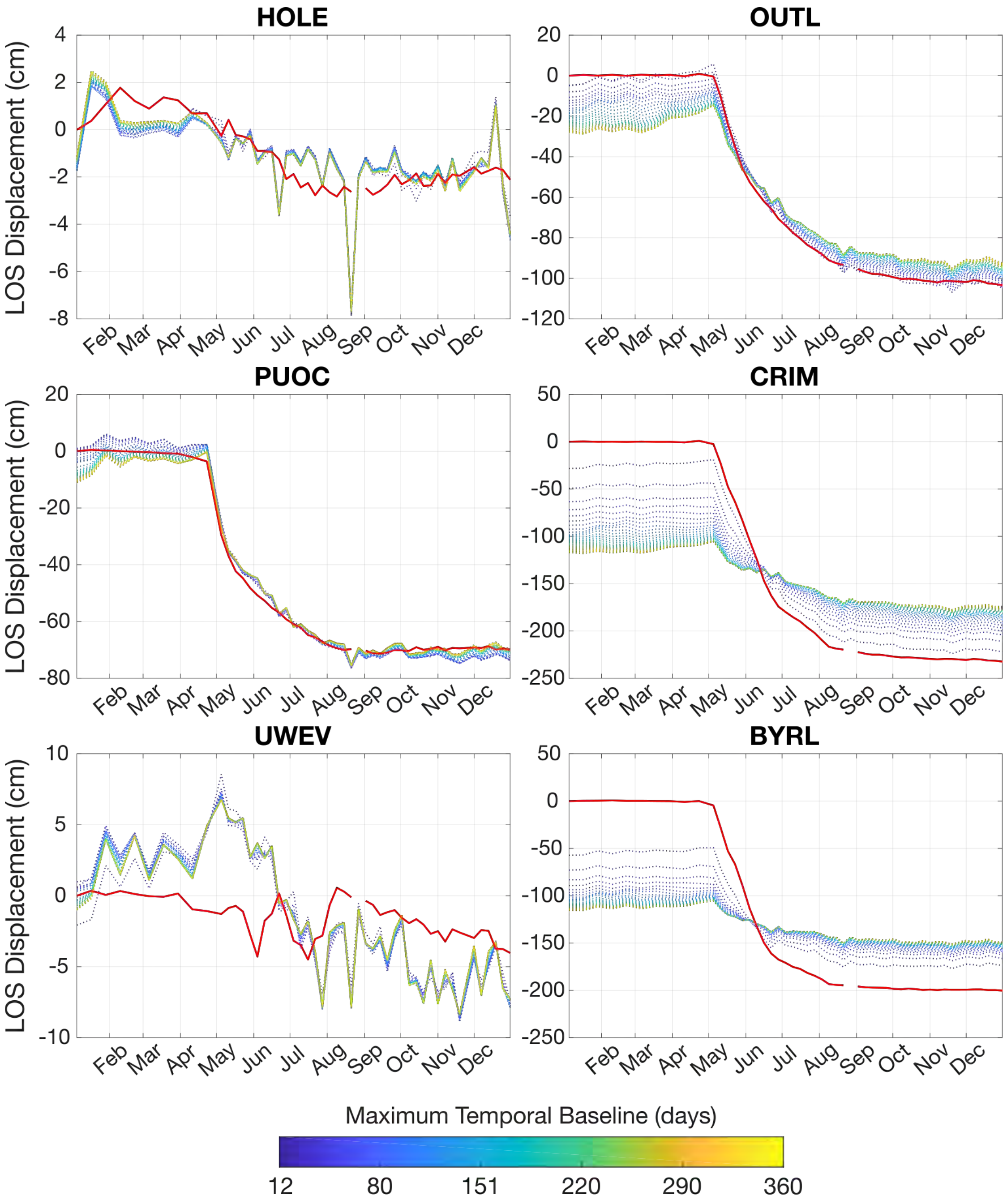


Figure 6.6: Aliasing in Hawaii: Kilauea Volcano. Each time series come from the corresponding labeled points in Figure 6.5. The maximum temporal baseline increases from blue to yellow (12-360 days), and the GPS data, projected into LOS, is depicted in red.

solution demonstrates similar sharp jumps between adjacent time steps, again reflecting that static noise sources cannot be reduced directly by the inclusion of more measurements. These sources of noise must be addressed by other means.

In contrast to station OUTL, we cannot fully reproduce the true displacement magnitudes at stations CRIM and BYRL, even when using the shortest TB_{max} (12 days). Although the spatial distance between station CRIM and OUTL is rather small, CRIM has almost an additional meter of displacement in the same time span. The resulting high displacement gradients between these two stations leads to aliasing at the CRIM station in even the shortest TB interferograms spanning the deformation event (May to August). Consequently, none of the time series can perfectly reproduce the GPS data. Station BYRL is similarly aliased, but it is interesting to see that its results have greater displacement loss with respect to GPS, despite having ~ 50 cm less cumulative displacement magnitude than CRIM. This disparity can be explained by recalling once again that gradients control aliasing, not magnitude, and noting that station BYRL is positioned at a highly decorrelated point in the study area. It is likely that the influence of larger decorrelation noise on the high gradients induced a greater degree of aliasing at this location.

The results thus far suggest that the shortest temporal baseline is best. Stations HOLE and UWEV show marginal gains by increasing the TB_{max}, suggesting that we need not expend additional computational resources by including redundant measurements. Alternatively, the shortest TB_{max} produces the best time series with respect to GPS at stations OUTL, CRIM, and BYRL due to the influence of aliasing. However, we cannot always assume that the SBAS solution using the shortest TB_{max} will be best. Station PUOC, for example, demonstrates a slight systematic shift with increasing TB_{max} that suggests the occurrence of aliasing. However, here it is the longest TB_{max} that produces the best solution compared to GPS, enough to warrant the increase in computation load. We hypothesize that the static and/or incidental noise leads to a slight bias that causes us to overestimate the displacement magnitude when using short TB_{max}. The inclusion of aliased measurements "corrects" this bias by reducing the total displacement observed within the deformation event. Alternatively, it is possible that there is no aliasing but rather an elevated incidental noise level, in which case increasing the number of redundant measurements improves the solution. Regardless of the reason, it is clear from these examples that pixels may require slightly different sets of phases to generate the highest-quality time series, with respect to the true displacements.

6.4.3 Delaware Basin, TX

We now shift our attention to a study area where deformation is driven largely by fluid extraction and injection, and has much smaller magnitudes than observed at Kilauea volcano. Indeed, it may be surprising to know that our first observation of aliasing was from our next study area: the Delaware Basin, TX (a large oil field in the Permian Basin). Chapters 7 and 8 provide more details about our motivation to analyze InSAR observations in this region, but as a summary, we expected to see

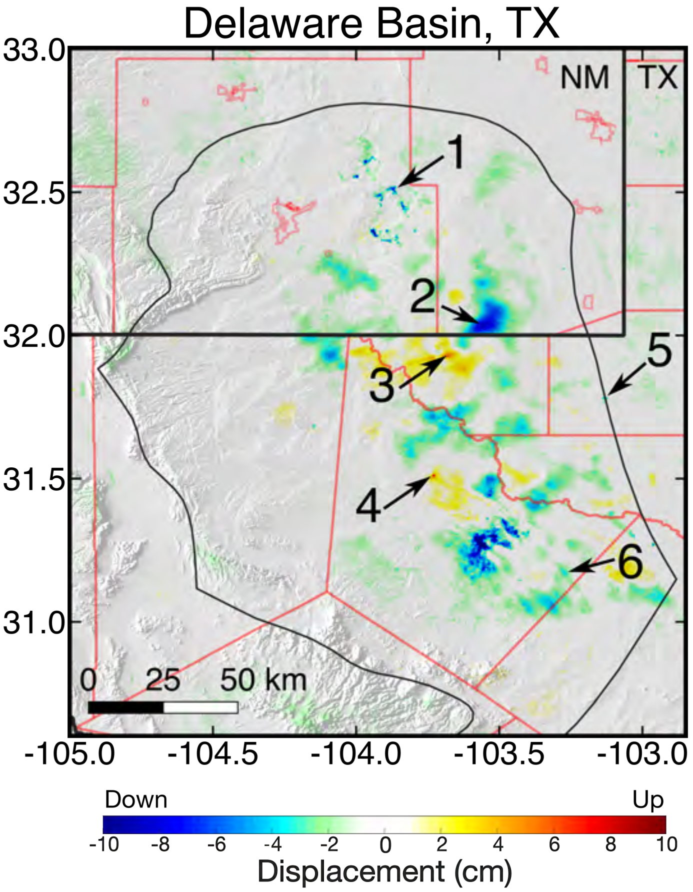


Figure 6.7: InSAR in the Delaware Basin - Descending LOS. We observe both subsidence (motion away from the satellite) and uplift (motion toward the satellite) in LOS. Sample points for time series are labeled and indicated with arrows.

modest displacements from subsurface volume changes due to oil and gas pumping and wastewater disposal, which we would then use to model the geomechanical response to the well activity. Our results, depicted in Figure 6.7, show up to only ∼12 cm of LOS displacement in 4.5 years and most features have rather small geographic footprints. Compared to the deformation event at Kilauea with 2 m of displacement in 5 months, it might be reasonable to assume that no aliasing would be observed in the Delaware Basin. However, we must recall that aliasing is not driven by magnitude, but rather our sampling and filtering of the displacement gradients.

The Delaware Basin is an expansive area, leading us to use a great deal of multilooking prior to unwrapping, so as to reduce data storage requirements and computation time (60 looks in range by 15 looks in azimuth). This degree of spatial averaging significantly reduces our aliasing gradient tolerance from 1.4/20 cm m^{-1} to 1.4/225 cm m^{-1} (see Chapter 5.6). Consequently, we do observe signs of aliasing in this region, despite the modest displacement magnitudes and rates.

The time series for the labeled points in Figure 6.7 are shown in Figure 6.8. The signals that display the familiar systematic shift (Points 1, 4, 6, and 7) are the features with the smallest spatial wavelengths. Point 1, for example, coincides with shallow potash mining in New Mexico. Being so close to the surface, the subsidence is largely one-dimensional (vertical); therefore, there is an abrupt spatial change (i.e. high gradients) between stationary ground and the subsidence over the potash quarry. Similarly, Point 5 corresponds to a sinkhole forming due to relatively shallow karst dissolution (Kim et al., 2019). Such features have very small spatial footprints, but relatively large displacement. Though we observe 10 cm of LOS subsidence with the shortest TB_{max}, we predict that the actual magnitude is much larger and filtering-induced aliasing is limiting the displacement we are

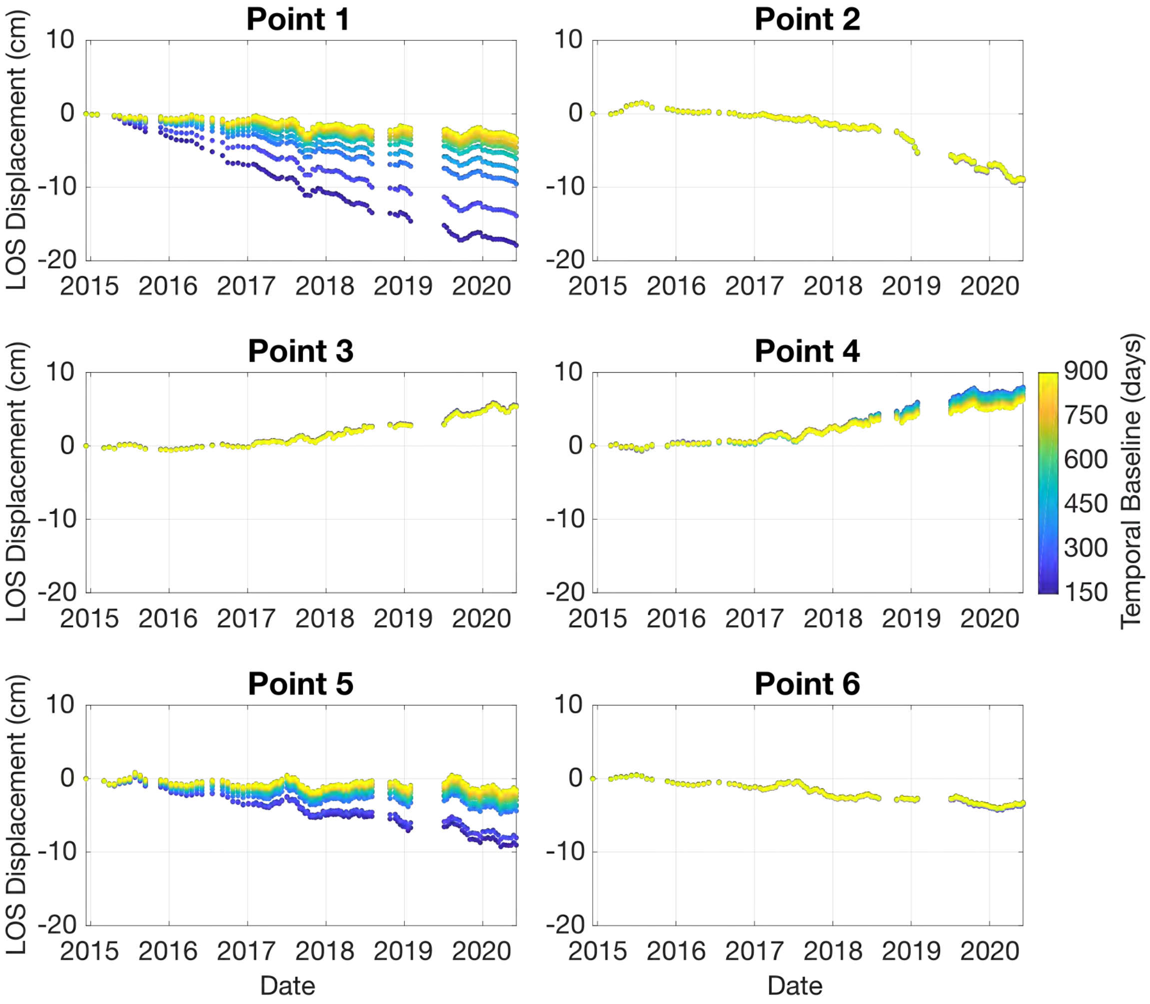


Figure 6.8: Aliasing in the Delaware Basin. Each time series come from the corresponding labeled points in Figure 6.7.

able to observe. The uplift feature at point 4 shows only a slight amount of potential aliasing, even though this feature has a similarly short spatial wavelength. Compare these occurrences of aliasing to Points 2 and 3, which demonstrate no aliasing but have cumulative displacement magnitudes similar to Points 4 and 5 (10 cm). Points 2 and 3 correspond to displacement features with much larger spatial wavelengths, highlighting again that aliasing depends on the gradients, not magnitude.

Points 2, 3, and 6 not only show no signs of aliasing, they show no signs of any change when we vary the TB_{max} from 115 to 900 days. This lack of variation indicates that the additional redundant interferograms are largely – if not completely – correlated to their independent counterparts at these locations. This suggests that the incidental noise in the Delaware Basin is mostly negligible, allowing us to use the shortest TB_{max} almost everywhere.

6.4.4 Central Valley, CA

Aliasing

Widespread subsidence in California's southern Central Valley (the San Joaquin Valley) can be attributed to drastic over-pumping of shallow groundwater resources for over a century. The Sentinel-1 InSAR cumulative time series between January 2015 through June 2019 demonstrates this ongoing subsidence, of which we show the 114-, 300-, and 1000-day TB_{max} SBAS solutions in Figure 6.9. Similarly to the Kilauea time series, the LOS displacement patterns become smoother and lessen in magnitude as TB_{max} increases. The 114-day TB_{max} solution highlights a few GPS stations to which we compare various SBAS time series in Figure 6.10. Clearly, each example demonstrates the systematic shift we expect from aliasing, where the shortest TB_{max} has the largest displacement magnitude and the longest TB_{max} has the smallest.

The characteristics that make the Central Valley time series interesting with respect to the previous study areas is that the best TB_{max} varies significantly throughout the valley; it is not exclusively the shortest temporal baseline that performs the best compared to the GPS data (red). Stations CHOW, P056, CRCN, and LEMA demonstrate the behavior we expect: the shortest TB_{max} produces the best result with respect to GPS, but even these do not always reproduce the true ground motion. Stations P056, CRCN, and LEMA underestimate the displacements at all TB_{max} we consider. Notably, station P056 demonstrates some of the highest error (LEMA is comparable), but has only half the cumulative displacement of CRCN.

The best SBAS solutions at the remaining stations in Figure 6.10 are not generated by the shortest TB_{max}. A moderate TB_{max} of ~400-600 days leads to the most accurate time series at stations P307 and MULN, whereas the longest longest TB_{max} (>~800 days) performs best at stations ALTH and P303. At this time it is unclear why there should be such drastic differences in the best TB_{max} throughout the Central Valley, especially considering that all of these points are situated within the region of widespread subsidence, but we can speculate. For example, the best solutions at stations P307 and MULN still correspond to the largest cumulative displacement magnitudes, suggesting a

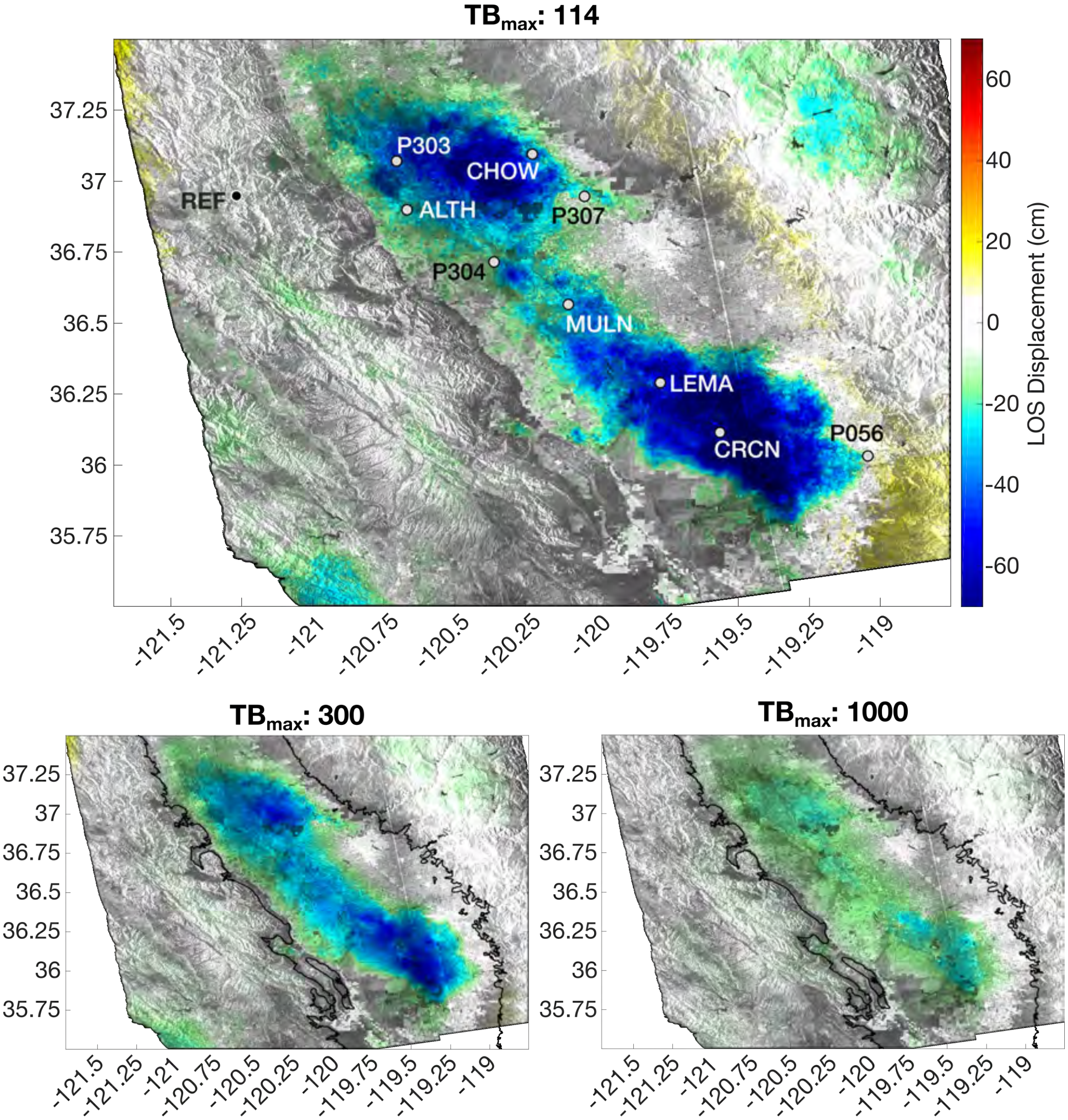


Figure 6.9: InSAR in the Central Valley - Ascending LOS SBAS solution using a TB_{max} of 114-, 300, and 1000 days. We primarily observe subsidence in the Central Valley related to over-pumping of groundwater resources. Labeled points in the top plot show GPS stations to which we compare the InSAR time series in Figure 6.10. The reference pixel ("REF") is shown in black on the northwestern side of the Valley.

balance between competing factors: high incidental noise decreases with the inclusion of more phase measurements, the benefits of which we see only up to a TB_{max} of 400-600 days. After this point,

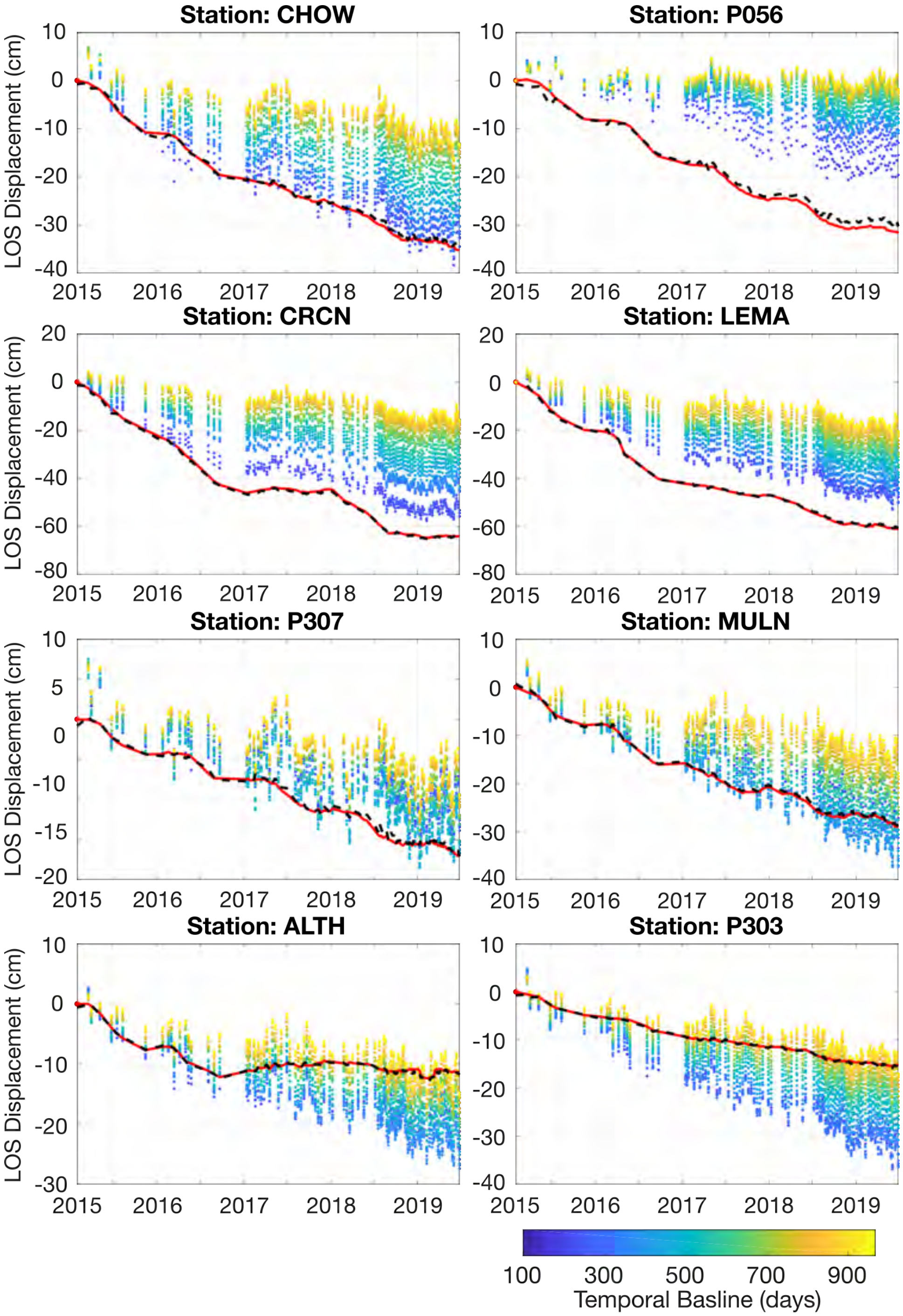


Figure 6.10: Aliasing in the Central Valley. Each InSAR (color scale) and GPS (red) time series come from the corresponding labeled points in Figure 6.10. The dashed black time series is generated using only consistent phase measurements (phases with one cycle of the corresponding GPS).

aliasing errors are introduced and the time series quality degrades. In contrast, the best time series solutions at stations ALTH and P303 require the longest TB_{max} but have the smallest displacement magnitudes. One explanation for this apparent contradiction could be that these stations are within a region of the study area that has biased measurements with respect to the reference pixel, perhaps due to decorrelation, persistent atmospheric trends, or possibly introduced by our removal of the phase-versus-elevation trend in each interferogram. In these cases, the aliasing errors offset the bias, leading to higher-quality solutions with longer TB_{max}. Another potential explanation is that the observed improvement with large TB_{max} is from a reduction in incidental noise, which may be particularly large in the Central Valley due to rapid decorrelation from changes in vegetation associated with agricultural practices. If so, it will be important to investigate why an increase in TB_{max} exhibits the same systematic decrease in cumulative displacement magnitude as aliasing.

Even though we can often reproduce the general temporal displacement patterns with at least one SBAS solution, each example from Figure 6.10 still contains a large amount of static noise (the consistent displacement jumps between adjacent time steps). This raises an important question: aside from utilizing potential atmospheric corrections or regularization techniques (e.g. Tikhonov regularization while inverting for velocity), is the static velocity solution the best we can hope to achieve with SBAS? Our analysis in the Central Valley suggests that if there is a sufficient level of non-static noise, we may be able to achieve much better results if we abandon the practice of selecting the set of unwrapped phases based on a maximum temporal baseline. Instead, we should seek to determine the best phase set with respect to the displacements.

We demonstrate this idea by generating an SBAS solution at each station in Figure 6.10 using all InSAR phase measurements that agree to their corresponding GPS displacement to within one phase cycle (1.4 cm), which we refer to as “consistent”. The resulting time series are displayed as the dashed black line at each station, all of which are almost indistinguishable from the GPS data (red line), indicating that we have successfully removed most of the noise (both static and incidental) from the SBAS time series at these points. Of course, identification of the proper phase set is simple when we have co-located ground-truth measurements such as GPS, as is the case in these examples. However, some components of static noise are spatially correlated (e.g. atmospheric phase screen), as are unwrapping slips. Therefore, it may be sufficient to identify the proper phase set at a few select points, and leverage this knowledge to generate more-accurate time series at the pixels in between.

Consider, for instance, stations ALTH and P303. Given their spatial proximity, we hypothesize that the set of consistent phases for one station will generate a reasonable time series at the other. Figure 6.11a displays various time series generated for station P303. The GPS (red) and the SBAS solution using the longest TB_{max} (dotted blue) are identical to what is shown in Figure 6.10. We generate the third time series using the P303 phases from the consistent set of interferograms for station ALTH, which are highlighted in Figure 6.11b. The consistent time series is, indeed,

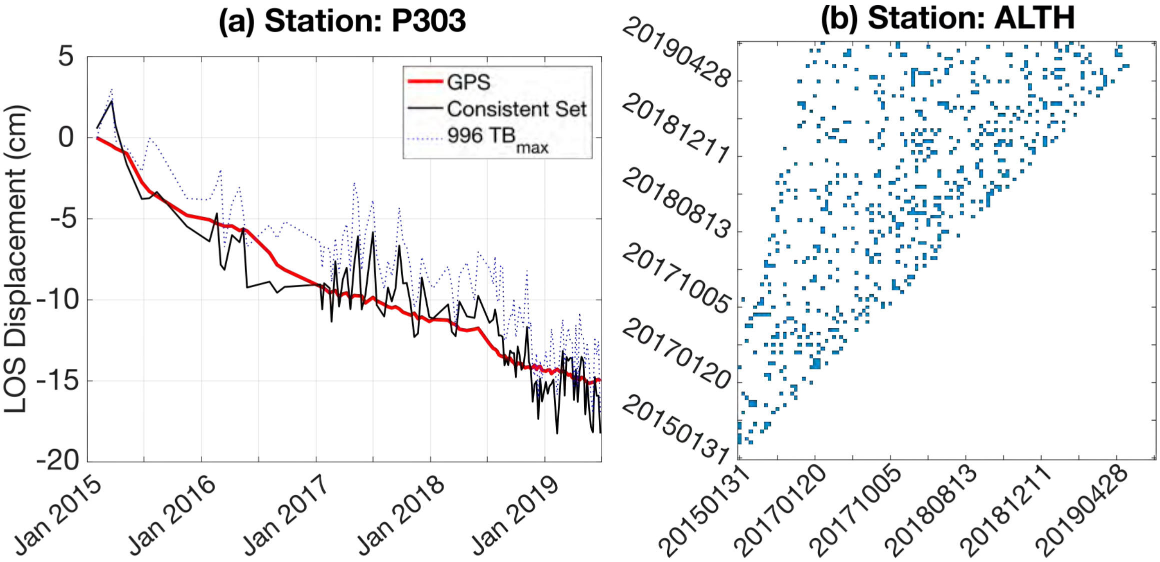


Figure 6.11: Generation of P303 time series using the set of consistent phases found at station ALTH. (a) Comparison of station P303 GPS (red) to two SBAS solutions: one generated using phase measurements from the set of consistent intereferograms found at station ALTH (as shown in b), and the other representing the time series with the longest maximum temporal baseline (996 days).

marginally better than the original best SBAS time series. This first suggests that we should be able to predict quality phase sets at pixels between GPS stations, which reduces the aliasing-driven uncertainty. Second, this indicates that these noisy interferograms often have the data we require to reduce atmospheric and decorrelation noise throughout the study area.

Though we can identify pixels that demonstrate signs of aliasing (a systematic shift as TB_{max} increases), we currently require knowledge of the true displacements at some distributed points within the study area if we are to isolate quality phase sets that can reduce both incidental and static noise. Luckily, California's dense GPS network allows InSAR users to quickly check or calibrate their time series. However, this restriction is not useful if we are interested in using InSAR to measure the displacement field in GPS-deficient regions. While we encourage a widespread expansion of global GPS networks, the ideal development from an InSAR perspective is to be able to identify quality phase sets at each pixel with knowledge derived exclusively from the interferograms. Such variables may be coherence, comparison of the wrapped and unwrapped closure phase values between corresponding phase triplets, or nearby displacement gradients. The Central Valley, with its hundreds of GPS stations, would make an excellent study area to begin developing such an automatic detection method.

6.5 Chapter Summary

This chapter has shown the effects that aliased measurements have on InSAR time series generated with SBAS techniques. The biased measurements decrease the cumulative displacement magnitudes as more aliased phases are included in the calculation. Because many of the phenomena we study demonstrate increased displacement over time, aliasing may also be temporally correlated. Therefore, we often observe a systematic shift in SBAS time series as the TB_{max} increases. Though the patterns are clear, it is not always the smallest TB_{max} that leads to the most accurate measurements when we have the luxury of comparing InSAR to ground-truth (GPS). Sometimes, as we showed in the Kilauea and Central Valley study areas, longer temporal baselines lead to the best results in some portions of the region.

Our results highlight that there is often not one magic TB_{max} that will work for every study area nor even every pixel. Though we have demonstrated how to detect aliasing, we must develop techniques that allow us to generate the appropriate sets of unwrapped phases for each pixel in our image. Without such methods, our analyses (and potential mitigation practices) may be misguided, especially in regions of the world where ground-truth measurements are not readily available to validate or guide our SBAS decisions.

Chapter 7

The Subsurface Response to Industry Activity in the Delaware Basin Inferred from InSAR

This thesis has thus far focused on the first step in the management of hazards using InSAR: ensuring high data quality by identifying and mitigating error sources in InSAR data products (aliasing). We now shift our attention to the application of InSAR to study the subsurface response to fluid extraction and injection, focusing on the Delaware Basin, TX, an extensive oil and gas field in the Permian Basin. We first identify the hazard-of-interest (induced seismicity) and motivate our use of InSAR to characterize the subsurface response to the basin's industry activity. We then describe our methods to generate basin-scale cumulative displacements, in both vertical and east-west components, and a vertical time series. Our results suggest that the northern and southern portions of the basin respond differently to fluid extraction and injection, which we use to support the potential relationship between shallow wastewater disposal and induced seismicity in Chapter 8.

7.1 Induced Seismicity in the Delaware Basin

The Delaware Basin is a giant oil and gas field in the Permian Basin, covering an extensive portion (22,000 km^2) of West Texas and southeastern New Mexico (Figure 7.1a inset). After being heavily exploited in the 20th century via conventional vertical production, 2009 brought a resurgence in oil and gas activity due to the development of organic rich shale beds using horizontal drilling and hydraulic fracturing (a.k.a. "unconventional") techniques. Similar to what has been observed in oil fields around the world, the Delaware Basin experienced an uptick in seismic activity coincident with unconventional development, leading many seismologists to infer those earthquakes were being induced by the development itself (e.g., Frohlich et al., 2016; Skoumal et al., 2020). Consequently, the state of Texas funded deployment of a regional seismic network, the TexNet array (Savvaidis et al., 2019), to better detect the regional seismicity and determine the underlying causes. The network has recorded thousands of small-to-moderate earthquakes in the Delaware Basin since its deployment in January 2017, including a M_W 4.8 event on 26 March 2020 (Figure 7.1a). These events are mainly concentrated in the southeastern portion of the Delaware Basin in Reeves county, despite widespread oil and gas activity throughout the basin (Figure 7.1b).

The low density of earthquakes to the north of the Grisham fault (Figure 7.1a) is likely due to low pore pressure conditions caused by decades of conventional oil and gas activity prior to the 21st century; however, the geomechanical mechanisms leading to the onset of seismic activity to the south of the Grisham fault since 2009 remain under debate (Dvory and Zoback, 2021; Hennings et al., 2021). Within the seismically active portion of the basin, the dense concentration of old vertical, new horizontal, and disposal wells (Figure 7.1b) makes it challenging to determine the most probable industrial drivers, since hydraulic fracturing ("fracking"), fluid production, and wastewater injection can all lead to induced seismicity (see Ellsworth, 2013; Schultz et al., 2020; Suckale, 2010, respectively, for reviews on these topics). For instance, hydraulic fracking is considered a major cause of induced events in western Canada (e.g., Farahbod et al., 2015), whereas earthquakes near the Wilmington Field in California have been linked to extreme subsidence and stress changes from oil and gas production (Kovach, 1974). The most-commonly cited mechanism, however, is wastewater injection, where fluid and increased pore pressure propagate to pre-existing faults, reducing normal stress and allowing seismic rupture (Ellsworth, 2013). Indeed, in nearby Oklahoma, where there have been a number of large-magnitude induced earthquakes ($M_W > 5$), studies strongly suggest that deep wastewater disposal near basement faults is the driving mechanism (Grandin et al., 2017; Keranen et al., 2013; Yeck et al., 2017).

In the Delaware Basin, the majority of wastewater disposal occurs in the Delaware Mountain Group (DMG), which lies above the producing shales (Wolfcamp) and 3–4 km above the basement in much of the producing portion of the basin (Figure 7.1c). In addition, there are few publicly-mapped faults in Reeves county and none of them extend from the DMG into the basement. Therefore, it is unlikely that basement faults are being induced to failure by wastewater disposal, as observed in nearby Oklahoma or elsewhere in Texas (Frohlich et al., 2014; Frohlich et al., 2016; Hornbach et al., 2015) unless poroelastic effects are the dominant mechanism (Zhai et al., 2021). An additional hurdle is the difficulty of linking specific events to any group of wells, due to the large depth uncertainty in earthquake hypocenters. Earthquakes in the southern Delaware Basin in the TexNet catalog range in depth between 0 and 19 km relative to ground surface (Figure 7.1c), with an average depth of 6 km and mean uncertainty of 1.9 km. Lomax and Savvaidis (2019) studied absolute depth errors in the basin and found a narrower depth range when a near station provided some depth control, but also large uncertainties of approximately 4–5 km. Because the average depth separation between

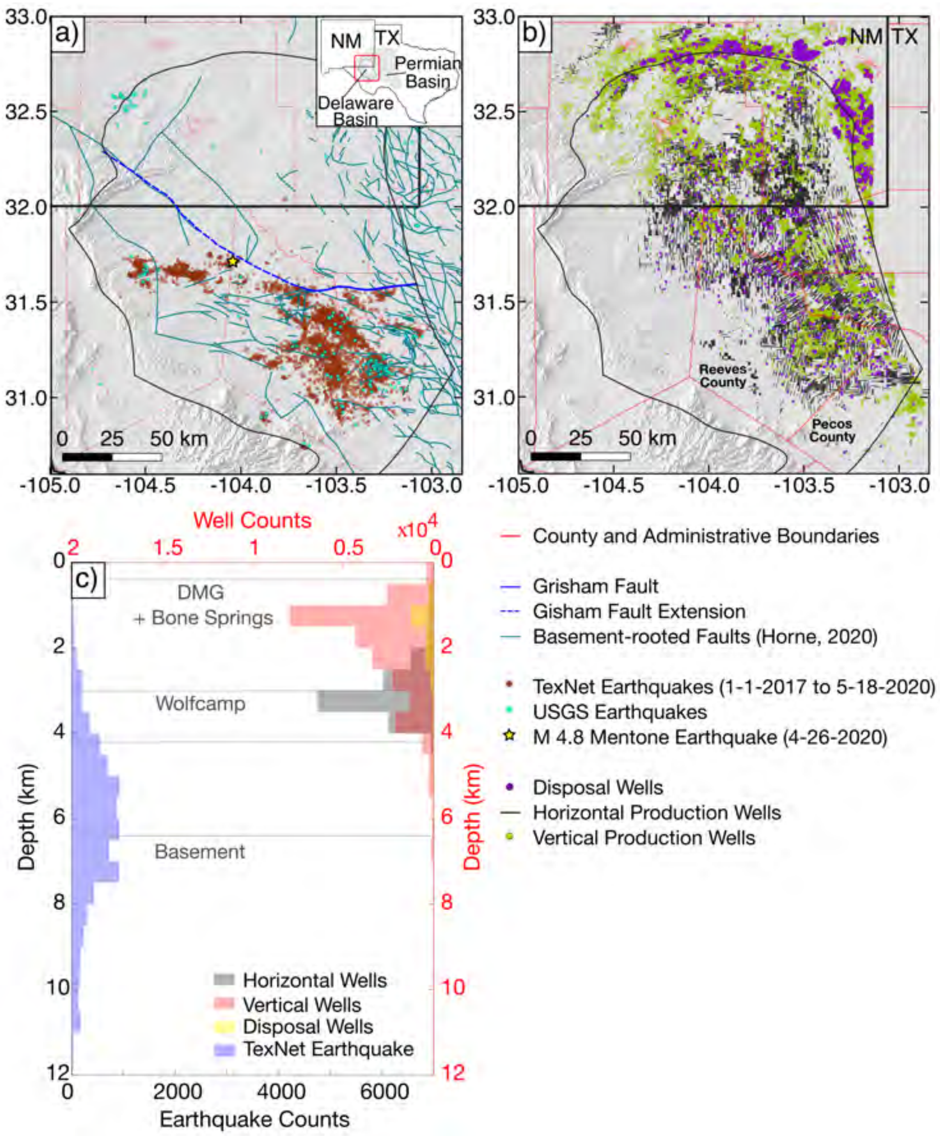


Figure 7.1: Seismicity, fault-mapping, and oil and gas activity in the Delaware Basin. (a) The seismic activity recorded by the USGS and TexNet arrays is concentrated in the southeastern Delaware Basin, below the Grisham Fault and its extension. Besides this distinction, few other faults show spatial correlations with seismic trends. (b) shows the disposal and productions wells that were active at some point between Dec 2014–June 2020 and assigned to the Delaware Basin (downloaded from Enverus, 1999. In contrast to the seismicity, oil and gas activity is widespread throughout the basin. (c) Earthquake depths from the TexNet catalog span a wide range, including into the basement, but these depths are highly uncertain. Most of the injection is concentrated above 3 km depth and above the producing shales. The formation depths depicted in (c) are approximate averages that we developed from formation surfaces provided by Enverus (1999); the true depth ranges vary throughout the basin. Fault traces provided by Horne, 2020.

disposal wells in the DMG and the production wells in the Wolfcamp is only 1.2 km, the formal uncertainty can move an earthquake from an injection formation to a producing one, or from a producing formation to the basement, and vice versa.

Despite these challenges, recent works attribute seismicity in the Delaware Basin to both hydraulic fracking and wastewater injection. Savvaidis et al. (2020) link clusters of events to fracking operations via temporal and spatial correlations and also highlight a causal link between wastewater disposal and seismicity in a few isolated cases where hydraulic fracking could be ruled out. However, in regions where fracking and disposal overlap, it remains challenging to distinguish between the two sources. On the other hand, Skoumal et al. (2020) attributed the majority of the seismicity to wastewater disposal, with just $\sim 5\%$ of the earthquakes induced by hydraulic fracturing operations. Another recent study uses poroelastic modeling to show that wastewater disposal at selected wells leads to pore pressure changes sufficient to induce earthquakes (Deng et al., 2020); however, they do not yet explain the absence of earthquakes near the majority of disposal wells in the basin, other than to posit that there are no favorably oriented pre-existing fault planes nearby. Zhai et al. (2021) also use poroelastic modeling to hypothesize that basement seismicity could be explained by poroelastic effects due to shallow injection within the DMG, though the evidence supporting earthquakes in the basement remains highly uncertain.

To better understand the geomechanical connections between industry operations and induced seismicity, it is essential to constrain earthquake depths, determine how the subsurface is responding to well activity, and locate faults hosting seismicity. In some instances, addressing one or both of the first two objectives may illuminate the geometry and behavior of unmapped faults. For example, precision earthquake locations and focal mechanisms, and measured surface deformation from co- and inter-seismic displacements can be combined to define faults and determine the sense and magnitude of slip (e.g., Massonnet and Feigl, 1995; Weston et al., 2012). These techniques are especially feasible when investigating shallow, large-magnitude earthquakes, of which there are currently none in the Delaware Basin. Most of the observed events only have $M_W < 4$, making co-seismic deformation analysis challenging, though not impossible (for instance, Staniewicz et al. (2021) showed $\sim$0.7 cm of co-seismic deformation related to the M_W 4.8 Mentone earthquake (Figure 7.1a). Nevertheless, using InSAR methods we will show that deformation in the Delaware Basin defines fault geometries and sheds light on the difference between the northern (non-seismic) and southeastern (seismic) zones of the region.

The use of InSAR to study the Delaware Basin has been growing in recent years. Kim and Lu (2018) used Sentinel-1 InSAR to map spatially isolated deformation signals and attributed them to subsurface causes, including karst dissolution at the Wink Sink and oilfield activity (see also Kim et al., 2019). In particular, the authors identified local instances of production-induced subsidence and injection-induced uplift. Both Deng et al. (2020) and Zhai et al. (2021) measured one component of deformation (a single Sentinel-1 radar line-of-sight look direction) to analyze the poroelastic pressure changes due to pumping and disposal, and included groundwater extraction as a possible source of subsidence. They also provided a wider look at the general deformation features in Reeves County. Staniewicz et al. (2020) further extended InSAR to the greater Permian Basin, using two overlapping Sentinel-1 passes (two look directions) over the Central Basin Platform and the eastern half of the Delaware Basin. They predicted a significant component of east-west horizontal motion in Reeves and Pecos counties, where the highest density of induced earthquakes occurs. These studies highlight the existence of non-tectonic deformation in the basin and demonstrate that geodesy may be an invaluable tool for understanding the subsurface response to oil and gas operations in this region.

In this chapter, we use Sentinel-1 InSAR to obtain estimates of the cumulative vertical and east-west horizontal displacements, and the vertical time series, in the Delaware Basin. The measurements reveal widespread deformation spatially correlated with production and injection wells; however, the northern and southern portions of the basin appear to respond differently to well activity, at least in terms of the observed surface deformation patterns. After analyzing the quality of our InSAR measurements, we begin to characterize the potential subsurface processes at play. Our results suggest that the displacements in the northern portion of the basin are driven by subsurface volume changes, whereas the observations in the southeastern region require an alternative geomechanical explanation. These observations motivate our hypothesis that the displacements can be explained by slip on normal faults, which we test with analytic models in the next chapter.

7.2 InSAR Methods

The InSAR processing method we use to study the Delaware Basin consists of four main parts:

1. Generate geocoded SLCs from the Sentinel-1 data and create the interferograms.
2. Reduce atmospheric noise.
3. Generate time series for each Sentinel-1 orbit using SBAS techniques (Berardino et al., 2002).
4. Decompose time series into vertical and east-west horizontal displacements.

We describe each briefly in turn.

7.2.1 Sentinel-1 Data Coverage

Our first step is to create geocoded single-look-complex (SLC) images at fine resolution (approximately 3.75 x 15 m) in three orbit sets (ascending paths 151 and 78, and descending path 85), using software developed by the Stanford Radar Group (Zebker, 2017; Zheng and Zebker, 2017). Figure 7.2 shows that these Sentinel-1 orbits have excellent spatial coverage. Descending Path 85

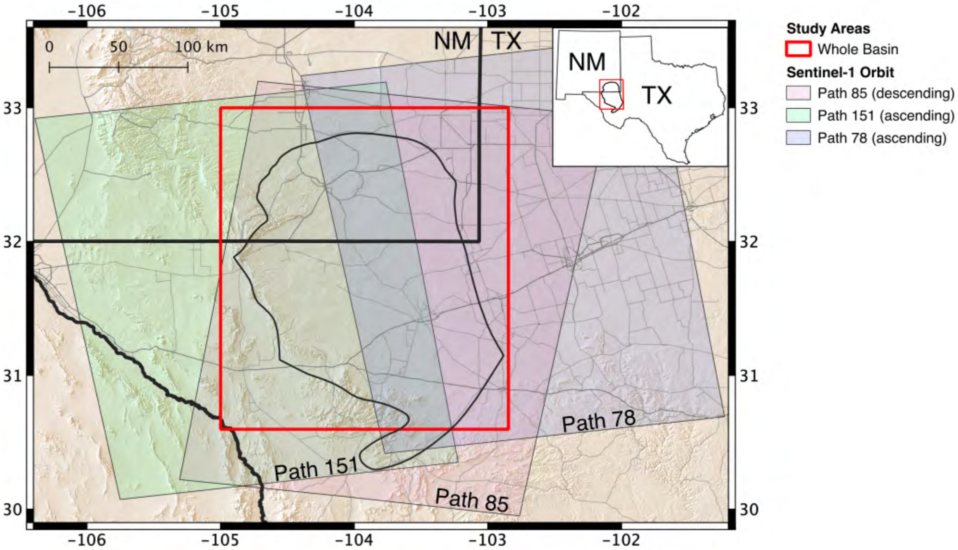


Figure 7.2: Sentinel-1 spatial coverage of the Delaware Basin. Descending Path 85 covers the entire basin, whereas the two ascending Path 151 (west) and 78 (east) split the basin, requiring both orbits for full coverage.

fully covers the entire basin, missing only a sliver of the northwest corner of the selected study area (solid red line). However, two ascending orbits (Path 151 (west) and Path 78 (east)) are needed to image the full study area. We remove SLCs with extremely high atmospheric noise based on visual, qualitative observations, resulting in 100 (Path 151), 108 (Path 78), and 109 (Path 85) SLCs between December 2014–June 2020. The dates of these SLCs are shown in Figure 7.3. In general, the repeat frequency in 2014-2016 is 24+ days, decreasing to 12 days by 2017. The large data gaps in 2018-2019 are either due to removal of SLCs with high atmospheric noise (Path 151) or lulls in acquisition (Path 85).

Next, we calculate all interferograms formed from SLCs spaced 600 or fewer days apart (resulting in 1791 (Path 151), 2448 (Path 78), and 2197 (Path 85) interferograms). We then spatially-average the interferograms to ∼225 m pixel spacing (60 range x 15 azimuth looks). Finally, we unwrap the interferograms using SNAPHU (Chen & Zebker, 2001).

7.2.2 Reduction of Atmospheric Noise with High-Pass Filters

Broad atmospheric signals up to ∼10 cm affect the unwrapped phases in many of our interferograms. To reduce the contribution of these signals, we first estimate and remove the dry atmospheric phase by fitting a linear trend between unwrapped phase and surface elevation and removing that trend from the interferogram. Next we must address the turbulent atmospheric component. Chapter 3 discusses a number of methods for atmospheric correction, each of which work in certain contexts.

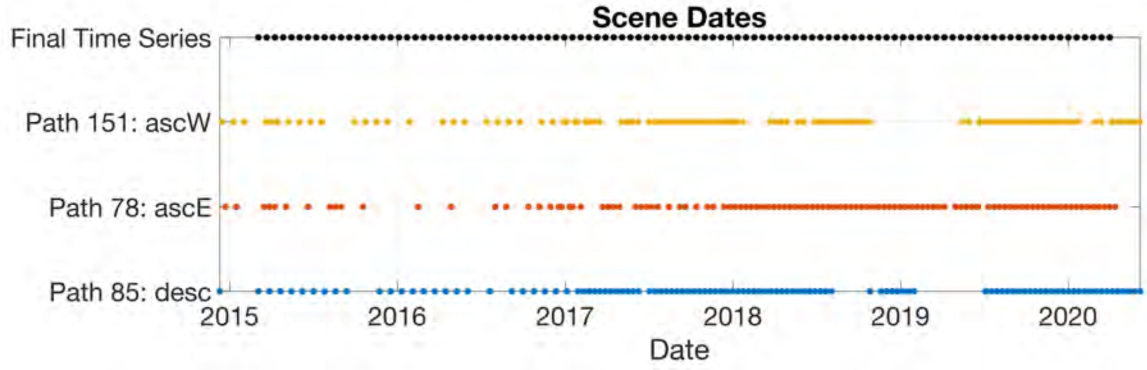


Figure 7.3: Dates for each single-look-complex (SLC) acquisition in the three line-of-sight (LOS) subsets. The Final Time Series line shows the dates chosen for the common interpolated time series in the vertical and east-west horizontal decomposition.

Here, we use a simple GPS-informed high-pass filtering method to remove long wavelength errors such as atmospheric noise, developed originally by Wei et al. (2010). Adopting this method, we develop the parameters for the Gaussian high-pass filter local to our study area with data acquired from nearby GPS stations and a preliminary interferometric stack (weighted average of our set of unwrapped interferograms). Our Gaussian high-pass filter has a standard deviation of 120 pixels and we apply this to each interferogram.

7.2.3 SBAS Time Series

The next step is to generate a unique time series for each of the Sentinel-1 orbits in our study area. We use an SBAS strategy (Berardino et al., 2002), which relates the unwrapped InSAR phases to the linear velocity between each SLC acquisition via a transform matrix consisting of interferogram time spans (Eq. 3.14). We select seven maximum temporal baselines (120, 156, 204, 300, 396, 504, and 600 days) in order to test for aliasing during the 2D phase unwrapping (Chapters 4-6). After solving for the linear velocity at each pixel, we integrate the velocities over the appropriate number of days for the line-of-sight (LOS) displacement time series. We generate one set of time series using the same maximum temporal baseline at each pixel (400 days). In this set, we also apply a second-order Tikhonov regularization in the SBAS calculation (α=150) to further reduce temporally-random signals (such as atmosphere and decorrelation), resulting in a smoother time series. We use the results of this time series for the cumulative displacements (and modeling in Chapter 8). However, we saw in Chapter 6 that the Delaware Basin is not immune to aliasing. Therefore, we generate a second set of LOS time series, using a different maximum temporal baseline at each pixel. The temporal baseline we select is the one with the greatest cumulative displacement magnitude. We use this set to generate the time series demonstrated in section 7.4.1[1].

7.2.4 Vertical and Horizontal Displacements

In the last stage, we combine and decompose each set of three LOS time series into time series of vertical and east-west horizontal displacements. For each set, we first resample the LOS time series to a uniform set of dates between 4 March 2015–31 March 2020 with 18 days spacing (Figure 7.3) and reference the time series to zero displacement on 4 March 2015. We then combine the two ascending data sets into a "composite" ascending time series by scaling Path 151 such that it approximately corresponds to the expected displacement as viewed by Path 78. Since the majority of the displacement appears to be vertical, we approximate the scaling factor at each pixel based on the relative projection of the LOS vectors onto the vertical component; that is, the ratio of the cosines of the incidence angles (θ_i) (in this case, $\frac{\cos\theta_{i-78}}{\cos\theta_{i-151}} \approx 0.86$). We then calculate the arithmetic mean at pixels where the two orbits overlap. We adopt the LOS unit vectors for Path 78 as the composite ascending unit vectors in further analyses. Finally, we decompose the descending and composite ascending LOS time series into vertical (V) and east-west horizontal (H_{ew}) displacements via the following relationship, which assumes zero north-south motion:

$$\begin{bmatrix} d \\ a \end{bmatrix} = \begin{bmatrix} los_{d-v} & los_{d-ew} \\ los_{a-v} & los_{a-ew} \end{bmatrix} \begin{bmatrix} V \\ H_{ew} \end{bmatrix}, \tag{7.1}$$

where d and a are the descending and ascending LOS measurements, respectively, at a single pixel and time step. Descending (los_d) and ascending (los_a) LOS unit vectors include only their vertical (v) and east-west horizontal (ew) components. We apply Eq. 8.1 to estimate V and H_{ew} at each pixel and time step.

7.3 Analysis of InSAR Methods

The cumulative displacements for each orbit (using a maximum temporal baseline of 400 days) are shown in Figure 7.4, where (a) is ascending west (Path 151), (b) is ascending east projected into the ascending west LOS (Path 78), (c) is the composite ascending displacements (we averaged the region of overlap), and (d) is descending (Path 85). In the region of overlap between the two ascending orbits, (a) and (b) agree to within 2 cm, which is the expected atmospheric noise level. Moreover, the displacement patterns are markedly similar. We demonstrate the agreement of the two time series in Figure 7.5, which shows samples from the points indicated in Figure 7.4a.

---

[1] The original paper published in Pepin et al. (2022a) uses the 400-day maximum temporal baseline results. Though there is some aliasing scattered throughout the basin, our modeling study area was not affected. Therefore, we consider our analysis and modeling with the 400-day time series to be accurate, not requiring an update. We do, however, include some time series analysis in this chapter. Here, we use the updated results that account for aliasing.

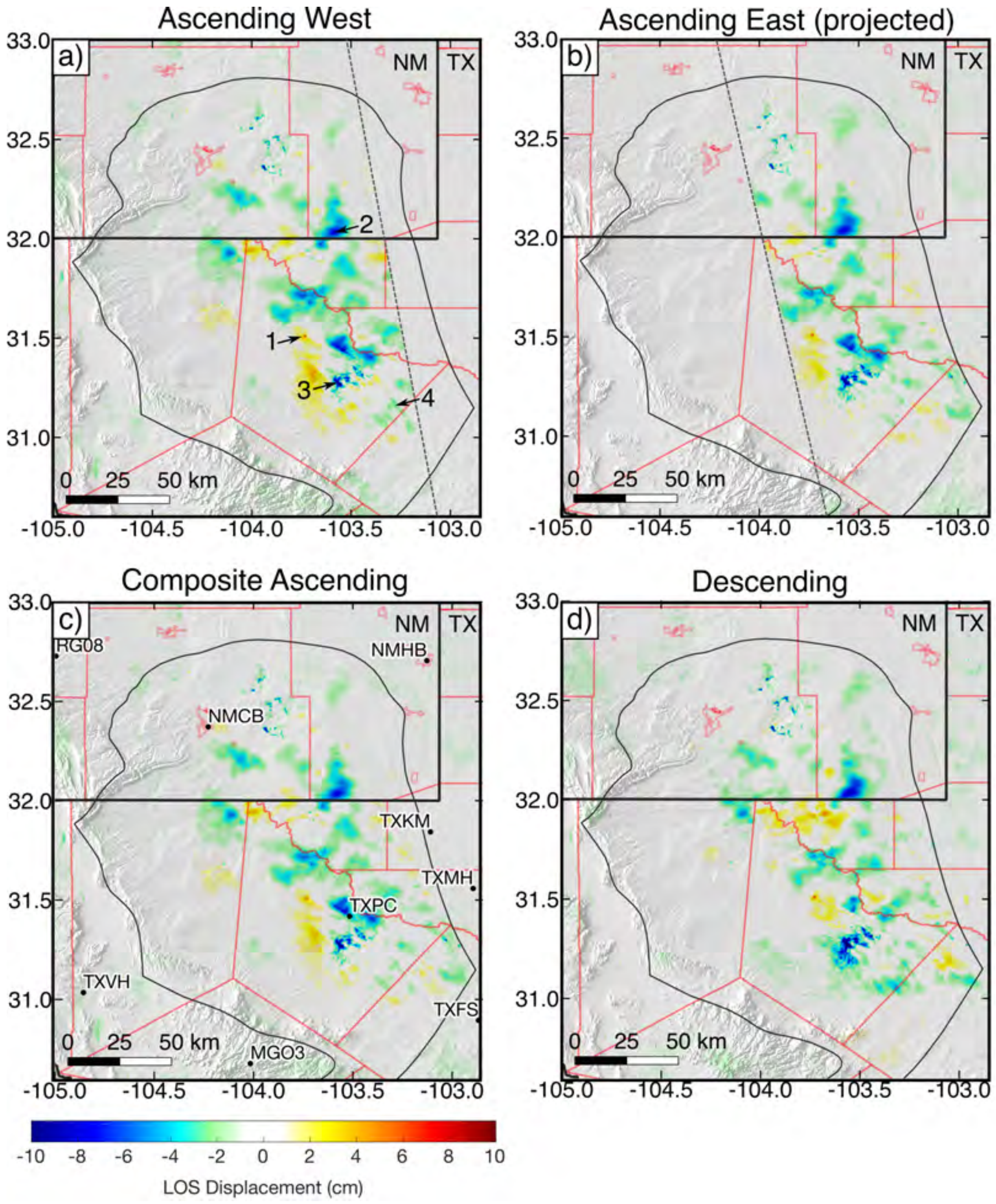


Figure 7.4: Cumulative displacements for orbits (a) ascending west, (b) projected ascending east, (c) composite ascending, and (d) descending, between 4 March 2015 to 31 March 2020. The dashed black lines in (a) and (b) indicate the bounds of the SLCs. The numbers and arrows in (a) show the locations where we compare the ascending time series in Figure 7.5. The local GPS stations are displayed in (c); we compare the GPS and InSAR time series in Figures S5 and S6. Red lines in all plots delineate the county and administrative boundaries.

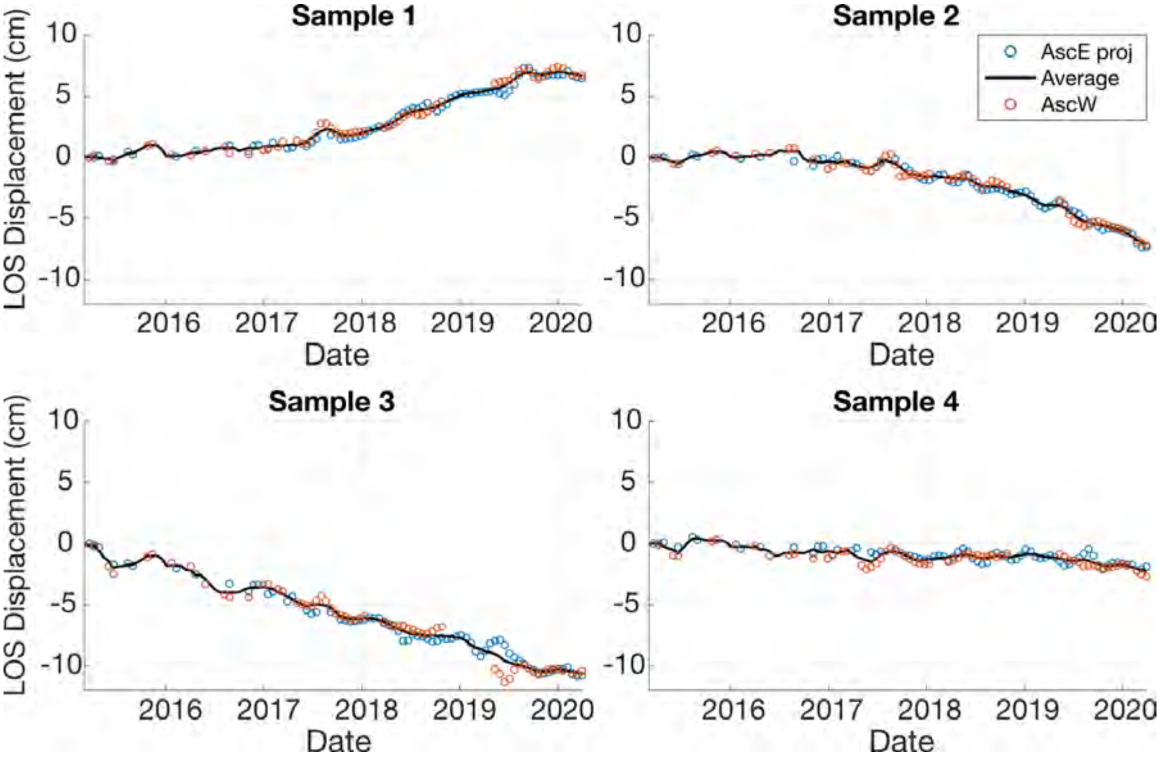


Figure 7.5: Comparison of time series for samples (for locations numbered in Figure 7.4a) from the two ascending orbits (ascending west and the projected ascending east). We calculated the average displacement at time steps which had data from both orbits. We then applied a moving average (window size = 3 samples) to develop the final average time series, as shown by the solid black lines. The time series from the two orbits agree within ±2 cm at each time step, which is the expected error due to atmospheric signals.

In our comparison, we show only the resampled time steps for which there was at least one ascending measurement within six days. The averaged time series is generated using either the single measurement (when only one is available) or the average between the two acquisitions (one from each orbit). We then apply a moving average (window = 3 samples) to develop the final time series shown in black. Each example shows high agreement between the two orbits at all time steps. Only Sample 3 shows a slightly higher difference between the two orbits (3 cm) for a short time in 2019. The cumulative displacements for all of these examples agree within 1 cm, well below the expected atmospheric noise. The observed consistency suggests that our processing decisions result in robust estimates of surface displacements, and that the majority of displacement is vertical.

Many InSAR studies support the accuracy of derived time series with the use of independent displacement estimates from GPS stations (as we did in Chapter 6). We show the locations of local GPS stations in Figure 7.4c and compare their time series (projected into InSAR LOS and smoothed with a moving average of 12 days) to the composite ascending (Figure 7.6) and descending (Figure

7.7) InSAR time series at these locations, noting that only a few GPS stations have coverage over our entire study period. Unfortunately, eight of the nine GPS stations are located in non-deforming areas, though we note that the InSAR data typically follow the seasonal trends within expected noise levels. Thus, we are unable to support the higher-magnitude InSAR displacements with GPS. The only station for which we may glean some insight is TXPC, for which we have GPS data through 2016. At this location, the ascending InSAR data show much stronger LOS displacement away from the satellite than the descending LOS time series (Figure 7.4c and 7.4d). The projected GPS time series through 2016 support this disparity, in which the projected ascending time series trends downwards, whereas the corresponding descending projection remains relatively flat. At this location we may thus expect a difference in LOS magnitude between the two look orbits, which is observed in the cumulative InSAR images.

7.4 Basin-Scale Deformation

We depict the cumulative vertical and east-west horizontal displacements between 4 March 2015–31 March 2020 in Figures 7.8a and 7.8b, respectively. In general, the vertical component is larger than the horizontal counterpart, consistent with previously proposed mechanisms of surface displacement in this region (e.g., poroelastic fluid flow [Deng et al., 2020; Staniewicz et al., 2020] and normal faulting [Staniewicz et al., 2020]). We find that the land surface both rises and falls in the portions of the Delaware Basin where there is a high density of horizontal and disposal wells, and is relatively static elsewhere (see Figure 7.8b inset and horizontal well region outline in Figure 7.8a). We note that the deforming areas include both seismically active and aseismic areas (see Figure 7.1a). This spatial correlation implies that the deformation can be linked to oil and gas operations, but variations in displacement patterns suggest that, depending on the region, different mechanisms may be dominating, which may explain the occurrence or absence of seismicity.

In Figure 7.8c, we modify the scale for cumulative vertical displacement to highlight narrow, short-wavelength linear deformation features in the southern portion of the basin, below the Grisham Fault. These features strike northwest-southeast with a gradual clockwise rotation to the south. In contrast, displacements north of the Grisham Fault have longer spatial wavelengths and no apparent preferred orientation. The horizontal deformation shows a similar regional distinction. To the north of the Grisham Fault, horizontal displacement magnitudes are only up to $\sim 1/2$ of the associated vertical magnitudes, but usually $< 1/4$, and form pairs of east-west displacement centered around subsidence and uplift features (e.g., westward motion on the right and eastward motion on the left of a subsidence bowl). Below the Grisham Fault, horizontal displacements are typically 1/2 to 3/4 of the associated vertical displacements (in some instances the horizontal even exceeds the nearby vertical), the preferred orientation of features is northwest-southeast, and there are fewer pairs of horizontal displacements around strong subsidence features. Thus, surface deformation in the zones

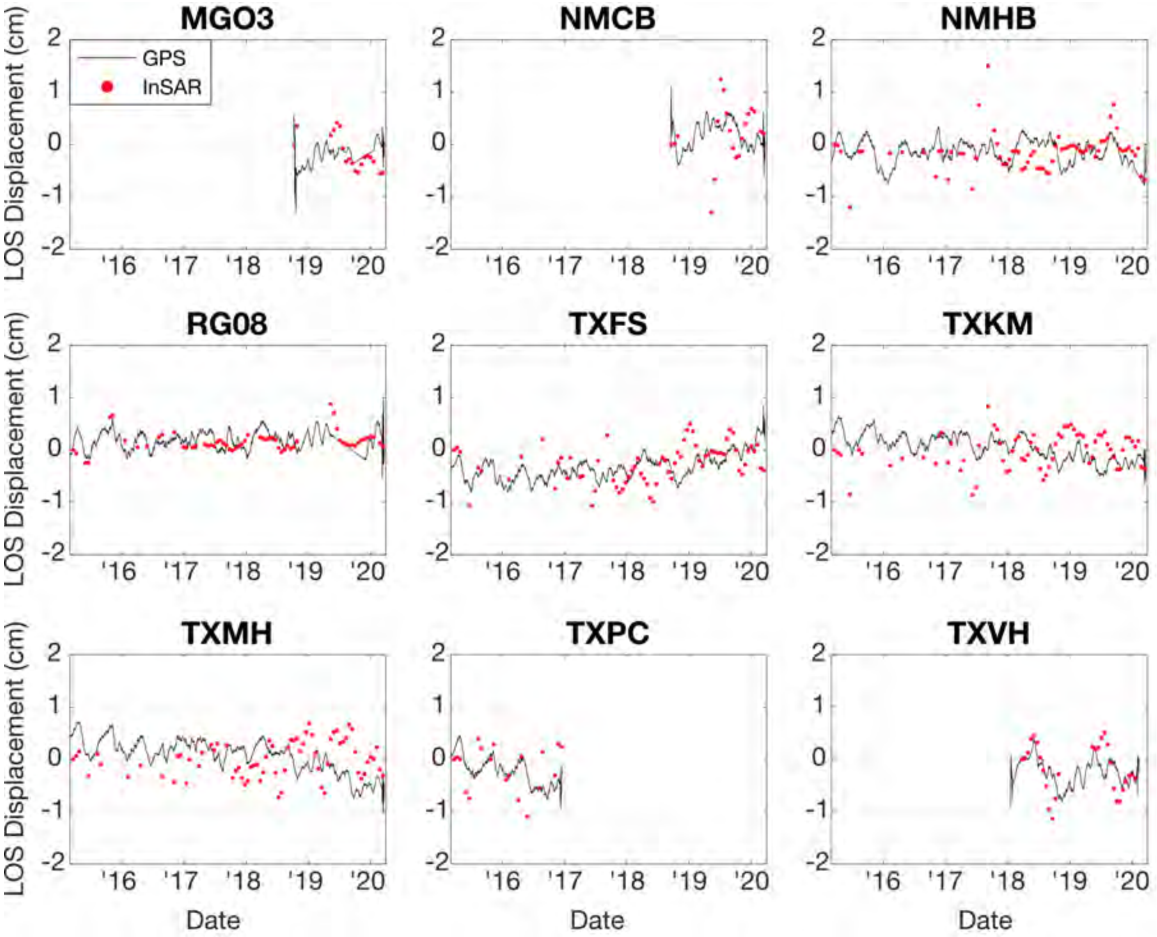


Figure 7.6: Comparison of the ascending composite time series to the GPS station data (locations depicted in Figure 7.4c), both in line-of-sight of the ascending west orbit. We shifted the GPS data to best-fit the InSAR time series in a least-squares sense. For stations MGO3, NMCB, and TXVH, we referenced the InSAR time series to the first date of the GPS data and show only the dates with GPS.

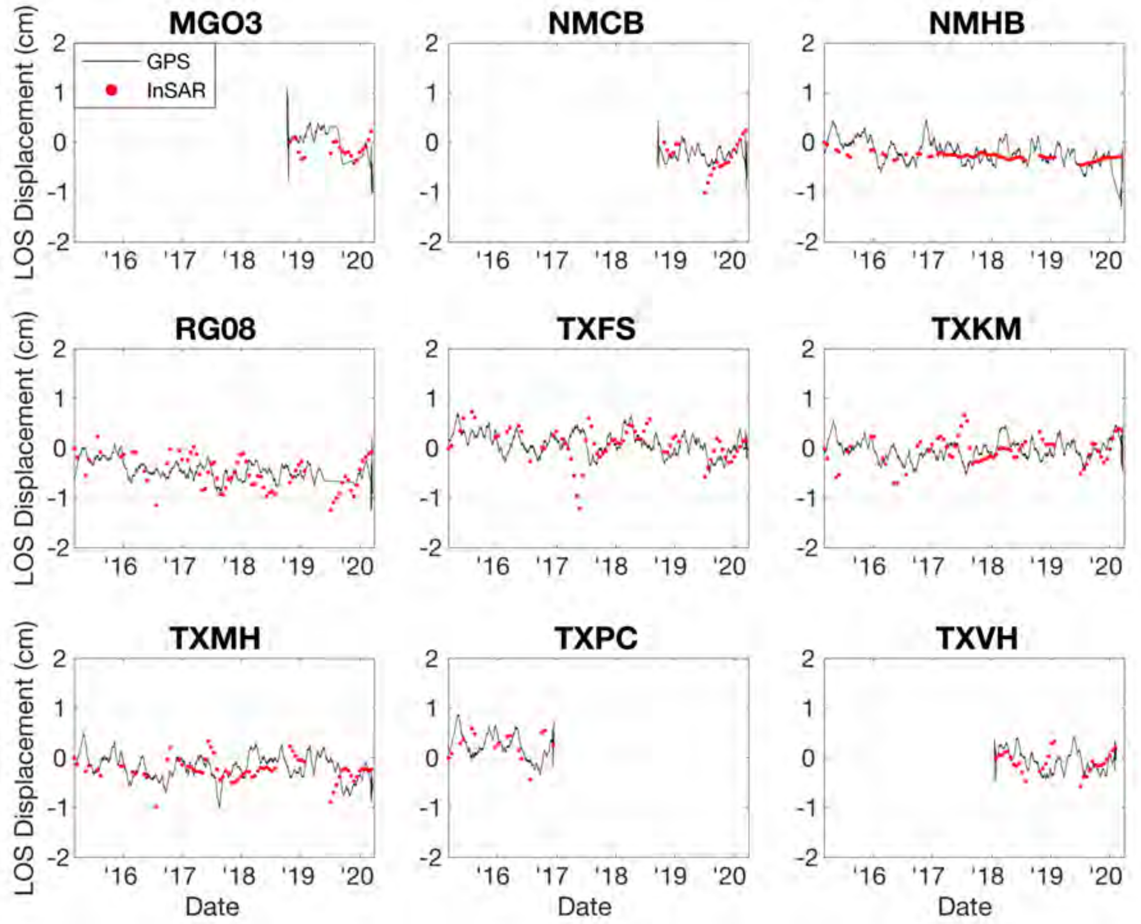


Figure 7.7: Comparison of the descending composite time series to the GPS station data (locations depicted in Figure 7.4c), both in line-of-sight of the descending orbit. We shifted the GPS data to best-fit the InSAR time series in a least-squares sense. For stations MGO3, NMCB, and TXVH, we referenced the InSAR time series to the first date of the overlapping GPS data and show only the dates with GPS.

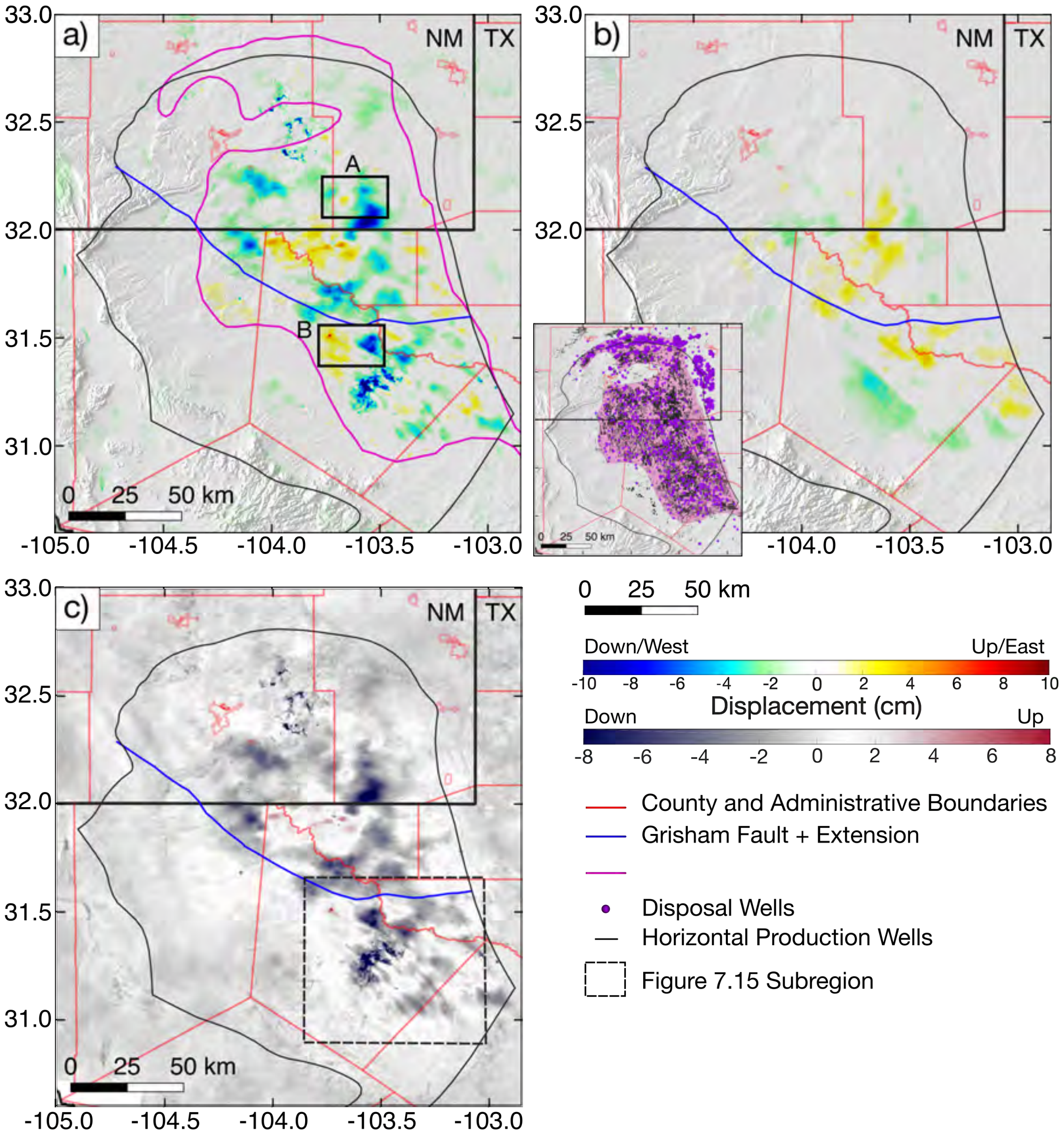


Figure 7.8: InSAR results in the Delaware Basin. (a) vertical and (b) east-west horizontal cumulative displacement between 4 March 2015–31 March 2020. In (c), We modified the color scale of the vertical displacement to highlight the linear features in the southeastern portion of the basin. Vertical displacements north of the Grisham fault and its extension have longer wavelengths and no preferred orientation. In (a) and (c), warm colors are uplift and cool colors are subsidence, whereas in (b), warm colors indicate eastward motion and cool represent westward. The inset in (b) shows the locations of disposal and horizontal production assigned to the Delaware Basin. The outline containing the region with a high density of horizontal wells is in pink, which is also included in (a) for reference; there is little measured deformation outside this region.

to the north and south of the Grisham fault apparently respond differently to industrial operations.

7.4.1 Volumetric Response to Fluid Extraction and Injection

Displacements to the north of the Grisham fault have a strong spatial and temporal correlation with local well activity. The InSAR data from subregion A in Figure 7.8a is depicted in Figure 7.9, where (a) and (b) show the vertical displacement and (b) additionally shows the locations of local wells, and (c) displays the horizontal motion. with an adjusted color scale. It is clear from (b) that the disposal wells (red diamonds) are typically associated with uplift or reduced subsidence, whereas horizontal production wells (black lines) correlate with the strongest subsidence features. These displacements are likely due to the combined effects of shallow wastewater disposal in the Delaware Mountain Group and the shale production in the Wolfcamp. That is, they reflect the volumetric changes from fluid extraction and injection at wells.

This hypothesis is supported by the nature of the horizontal displacements (Figure 7.9c). The stretched color scale highlights that eastward and westward motion come in pairs that flank the vertical subsidence and uplift features. For example, the uplift corresponding to point 1 in (a) is centered at approximately 12 km east and 9 km north. The same area in horizontal shows westward and eastward motion on the left and right of the uplift, respectively, exactly as we would expect from an expanding subsurface volume.

The temporal correlations between the InSAR time series and monthly well volumes further supports our hypothesis. We first focus on the region highlighted by the black box from Figure 7.9a. Figure 7.10 depicts the combined monthly volumes for all wells within the black box. Production is split into the three various produced fluids: oil (top left), gas (top right), and water (bottom left). The bottom right plot shows the combined wastewater disposal volumes in the same region. The colors highlight that each well produces various volumes of each fluid. We also include the vertical time series from point 1 in Figure 7.9a as the black line. The uplift signal is largely contained within the time frame of local wastewater injection, up to the beginning of 2018. At this time, injection abruptly stops while production modestly increases. Consequently, it appears as though the increased pore pressure from the injection diffuse away from the disposal wells and the ground surface gradually subsides.

There is also a strong temporal correlation between subsidence and production in the small region outlined by the white box in Figure 7.9a. Figure 7.11 mimics Figure 7.10, this time for the region within the white box and with point 2 as the InSAR time series. The horizontal dashed line in each plot highlights the zero-datum for the surface displacement. Production appears to be relatively constant up to the beginning of 2018 and we see a small rate of subsidence through this time. When production begins to ramp up, the rate of subsidence increases. Note also that the increased rate in subsidence corresponds to when the nearby disposal well stops injection in the middle of 2018.

These examples all support the hypothesis that the the local surface displacements are driven by subsurface volume changes from the well activity. Displacements are most strongly correlated with disposal volumes, which are shallower than production volumes in this region. This is consistent with our knowledge of how increased loading effects with depth dampen the propagation of deformation to the surface (Gambolati et al., 2005).

Our InSAR observations to the south of the Grisham fault do not correlate as well to local wells as the previous examples to the north. For illustration, we turn our attention to the region outlined by box B in Figure 7.8a. This area depicts a sharp uplift feature at the end of a curvilinear trend oriented northwest-southeast, as shown in Figure 7.12a. There is also a wider, strong subsidence feature that aligns with this same trend. The horizontal displacements (b) are a little harder to interpret. Though we see the expected westward and eastward displacement pair flanking the sharp uplift at point 1, the other horizontal motion is largely eastward near the strong subsidence feature,

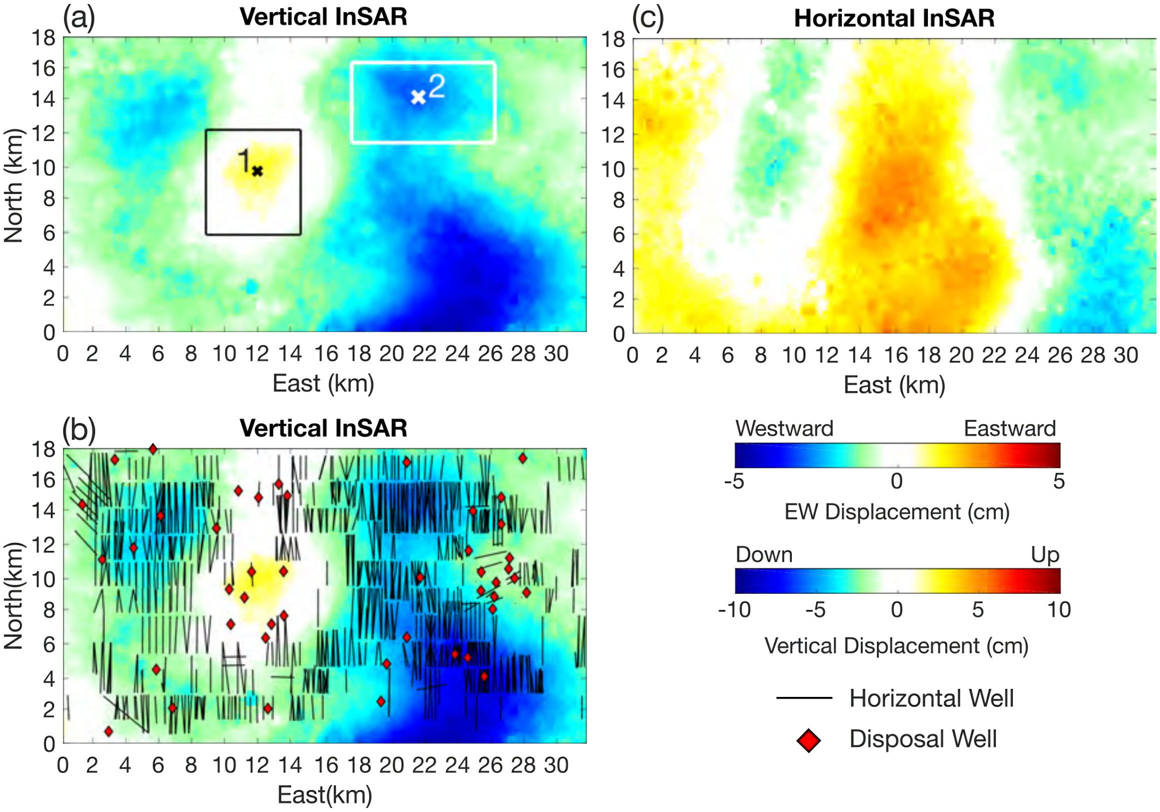


Figure 7.9: Comparison of InSAR and wells in a small New Mexico subregion. The vertical InSAR is depicted in a and b, where b includes the location of horizontal production and vertical disposal wells. The horizontal displacements are depicted in c. Horizontal motion behaves has expected relative to the locations of vertical subsidence and uplift. The boxes in a delineate the regions from which we selected wells for comparison to the vertical time series at points 1 (Figure 7.10) and 2 (Figure 7.11).

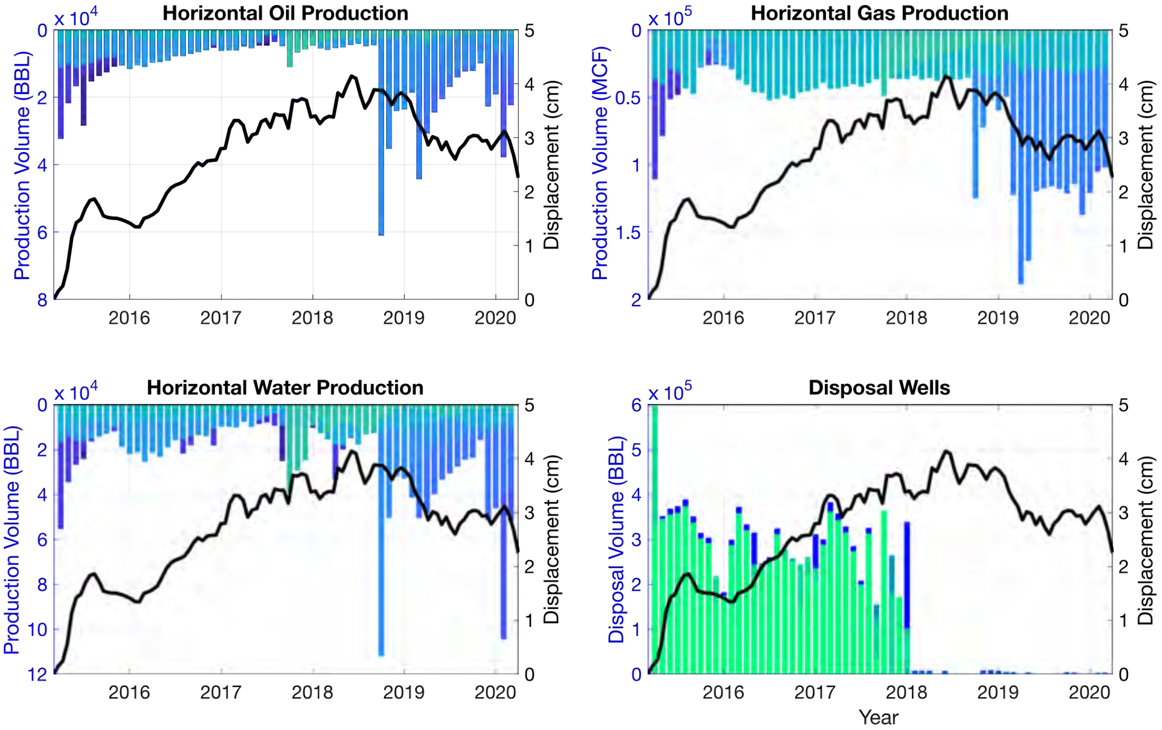


Figure 7.10: Comparison of InSAR time series and monthly well data at point 1 in the New Mexico subregion (Figure 7.9). Each plot shows the vertical time series with scale on the right axis. The left axes display the monthly well volumes for various produced fluids, with oil on the top left, gas on the top right, and water on the bottom left. The bottom right depicts the monthly wastewater disposal volumes. The search region for the wells is depicted as the black outline in Figure 7.9a

.

without an appropriate westward pair, and extend farther than the subsidence to the northwest end. Furthermore, the magnitude of eastward motion is approximately half the magnitude of the corresponding subsidence, which is higher than we would expect from volumetric contraction.

We explore the temporal relationship of the features at points 1 and 2 in Figures 7.13 and 7.14, respectively, using the same format as Figures 7.10-7.11. The uplift feature in Figure 7.13 is not surrounded by many wells (solid black box in Figure 7.12a), though there is a disposal well directly to the south. Surprisingly, the nearby wells did not begin injection until 2019, but the rapid uplift begins near the end of 2015 (lower right). However, there are many disposal wells along the curvilinear feature oriented to the southeast of the uplift, within the dashed black box from Figure 7.12a. These well volumes are compared to the uplift time series in the lower left plot of Figure 7.13; they show a much stronger temporal correlation to the uplift signal. Note also that there is little uplift near these wells. It is possible that the fluid is being injected near a hydraulically-conductive fault oriented along the northwest-southeast trend. The observed uplift may be a junction point of

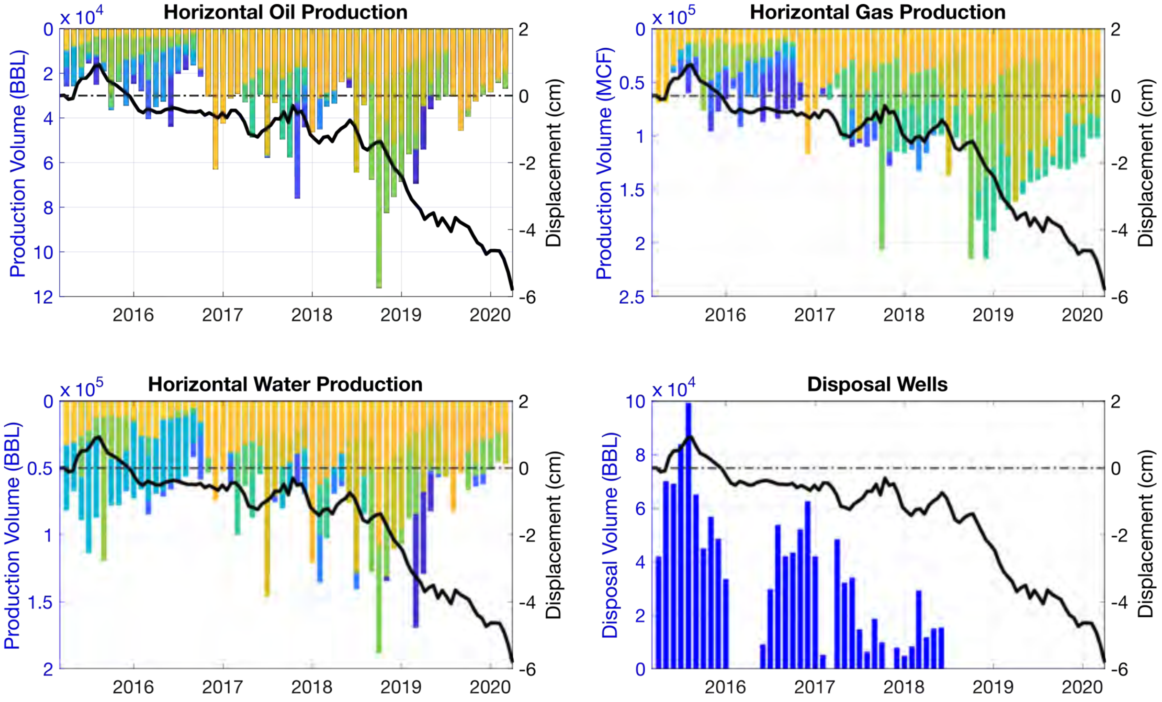


Figure 7.11: Comparison of InSAR time series and monthly well data at point 2 in the New Mexico subregion (Figure 7.9). Each plot shows the vertical time series with scale on the right axis. The left axes display the monthly well volumes for various produced fluids, with oil on the top left, gas on the top right, and water on the bottom left. The bottom right depicts the monthly wastewater disposal volumes. The search region for the wells is depicted as the white outline in Figure 7.9a

.

pore pressure build-up. Though we do not test this hypothesis here, it is clear that the deformation does not show the clear spatial and temporal correlations that we see to the north of the Grisham fault.

We see a similar disparity between the spatially-correlated production wells near the subsidence feature at point 2 (Figure 7.12a) and the observed time series, as shown in Figure 7.14. There is little to no displacement at this point until the beginning of 2016, at which point subsidence begins. However, monthly production is strong throughout the time series, with no appreciable change corresponding to the onset of subsidence. On the other hand, wastewater disposal increases markedly as subsidence continues. These InSAR observations do not clearly correspond to the behavior we expect from volumetric changes in the subsurface due to production and injection. Therefore, we may require additional geomechanical mechanisms to explain the deformation to the south of the Grisham fault.

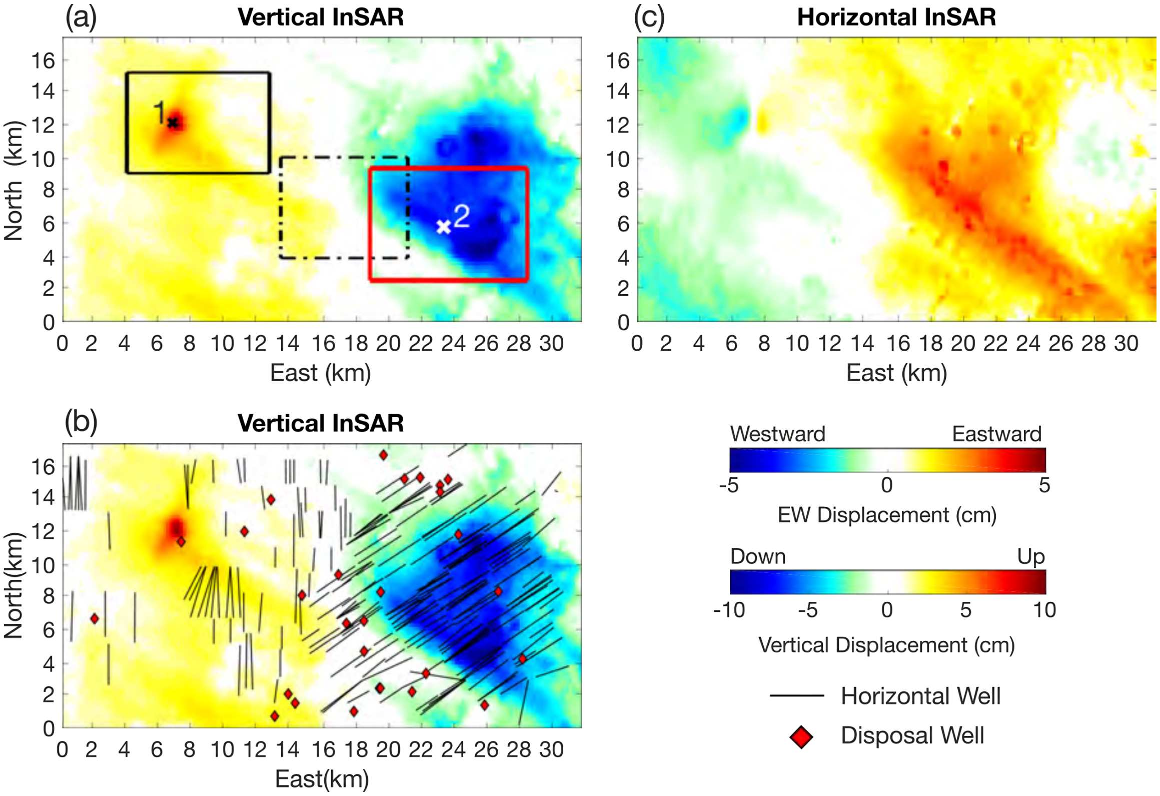


Figure 7.12: Comparison of InSAR and wells in a small region of Reeves county, south of the Grisham fault. The vertical InSAR is depicted in a and b, where b includes the location of horizontal production and vertical disposal wells. The horizontal displacements are depicted in c. Horizontal motion behaves has expected near the sharp uplift at point 1, but not elsewhere. The boxes in a delineate the regions from which we selected wells for comparison to the vertical time series at points 1 (Figure 7.13) and 2 (Figure 7.14).

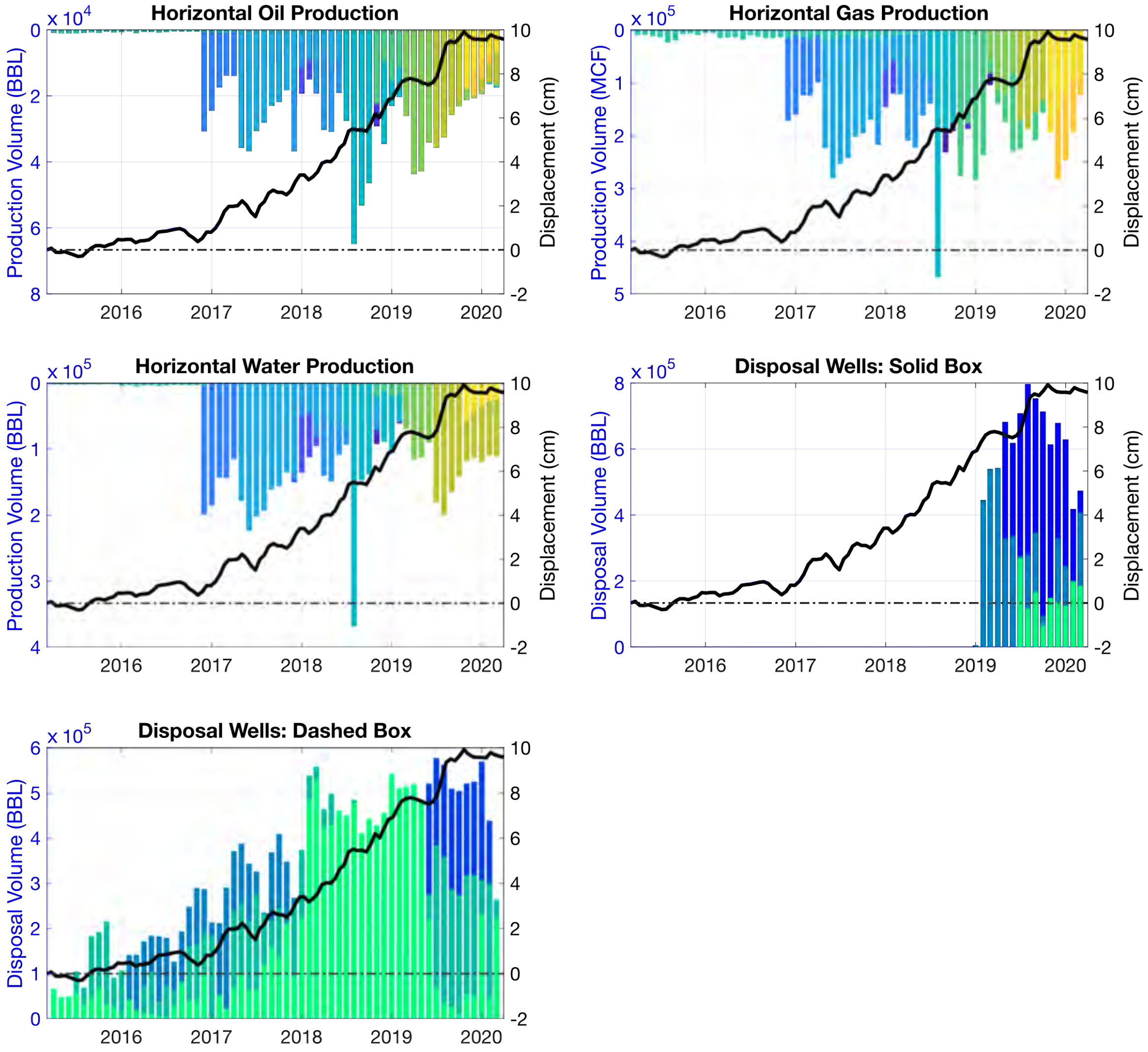


Figure 7.13: Comparison of InSAR time series and monthly well data at point 1 in the Reeves County subregion (Figure 7.12). Each plot shows the vertical time series with scale on the right axis. The left axes display the monthly well volumes for various produced fluids, with oil on the top left, gas on the top right, and water in the middle left. The middle right depicts the monthly wastewater disposal volumes for the disposal wells in the solid black box from Figure 7.12a, whereas the bottom left shows the monthly wastewater injection volume for the disposal wells in the dashed black box.

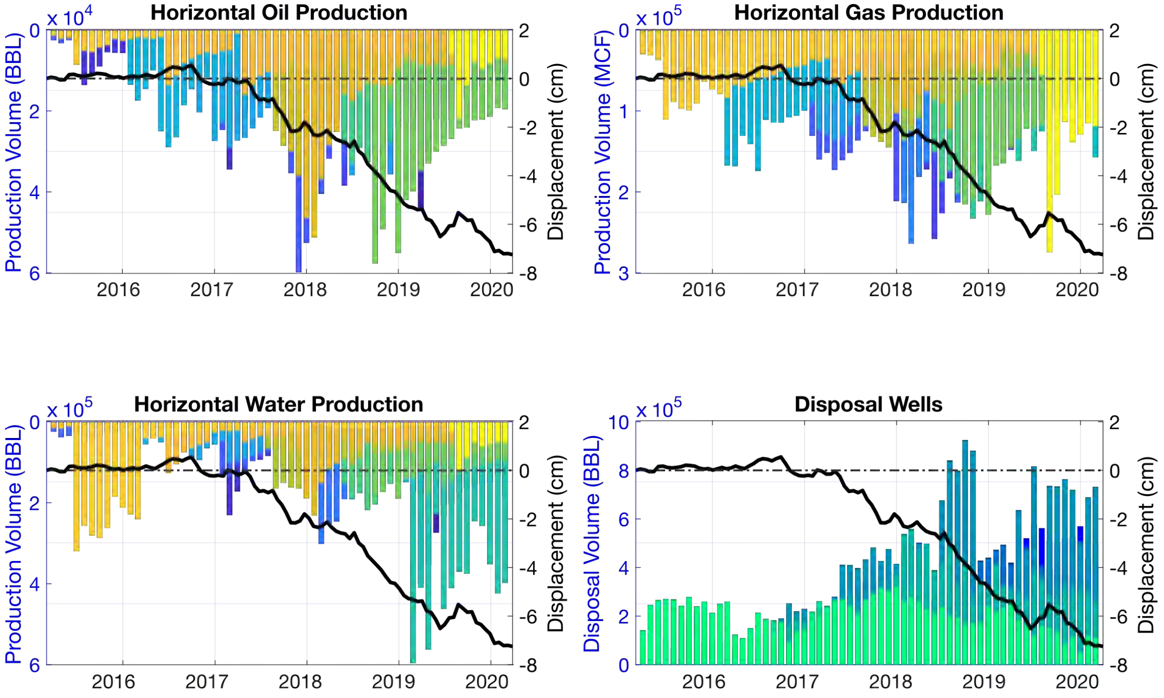


Figure 7.14: Comparison of InSAR time series and monthly well data at point 2 in the Reeves County subregion (Figure 7.12). Each plot shows the vertical time series with scale on the right axis. The left axes display the monthly well volumes for various produced fluids, with oil on the top left, gas on the top right, and water in the lower left. The lower right depicts the monthly wastewater disposal volumes for the disposal wells in the solid red box from Figure 7.12a.

7.4.2 Curvilinear Displacement Patterns

The outlined subregion in Figure 7.8c corresponds to the highest density of seismic activity in the southeastern quadrant of the basin, suggesting that the linear InSAR displacement features could be related to the earthquakes. In Figure 7.15, we display this subregion to compare these linear features with the tectonic stress field and seismicity from the TexNet catalog. Lund Snee and Zoback (2018) compiled measurements of maximum principal horizontal stress ($S_{H_{max}}$) orientations, depicted as red lines in Figure 7.15a, and ranked their quality based on the number, depth range, and agreement of measured stress indicators (the authors consider only orientations with A-C ranking sufficiently robust for plotting and analysis). The highest-quality $S_{H_{max}}$ orientations ("A" and "B" lines) are parallel to the linear deformation features. As shown in Figure 7.15b, seismicity also tends to align with the InSAR deformation patterns. All three data sets independently display the same rotation in strike from ∼300° in the northwest corner of the subregion to ∼330° in the southeast. Lund Snee and Zoback (2018) classify the stress state of the Delaware Basin as a predominantly normal-faulting regime. Under these stress conditions, normal faults striking parallel to $S_{H_{max}}$ are the most-susceptible to fail. Thus, the spatial relationship of these three data sets suggests that slip on

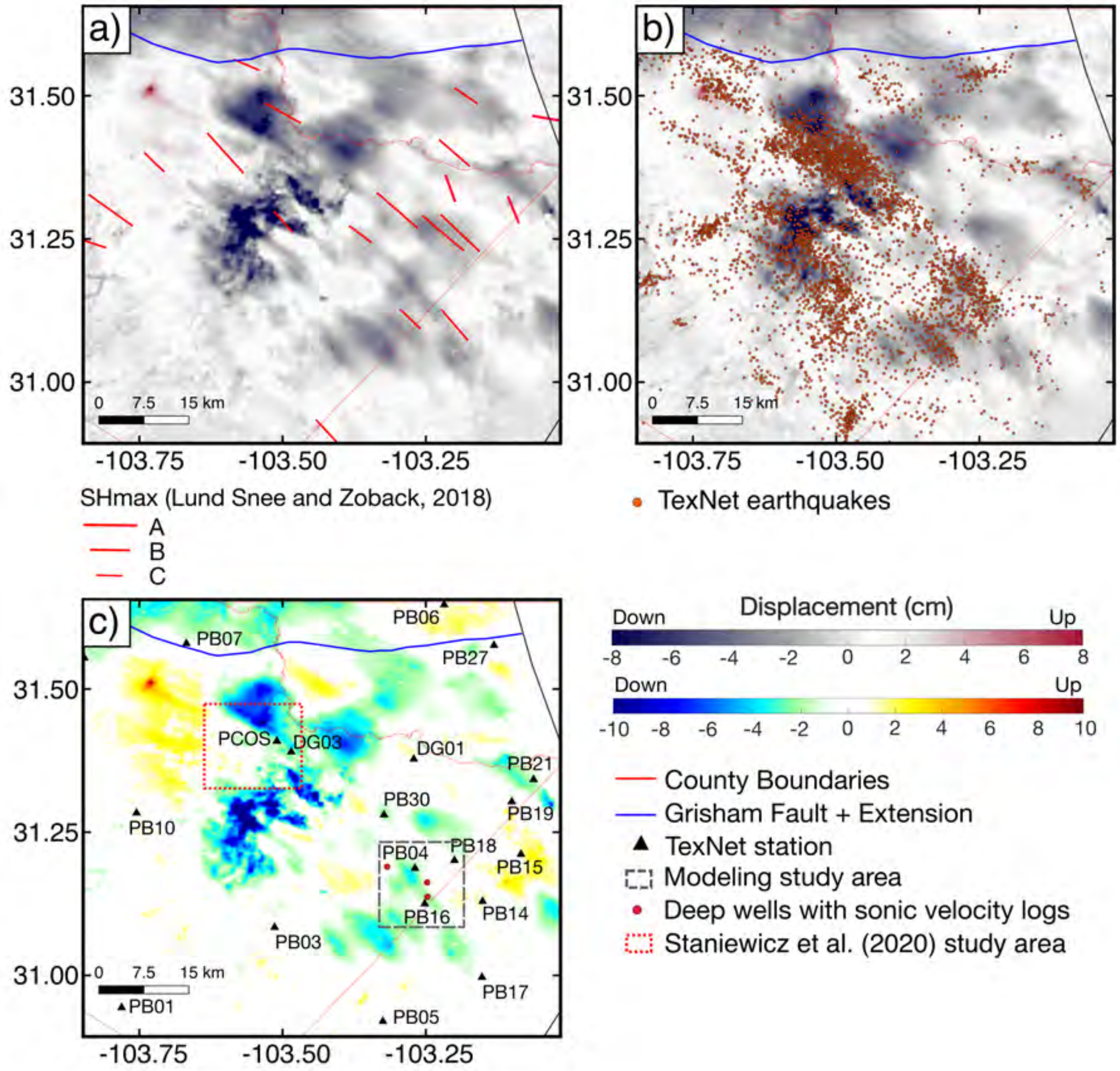


Figure 7.15: Subregion with saturated color scale comparing linear deformation features to (a) $S_{H_{max}}$ orientations (Lund Snee & Zoback, 2018) and (b) TexNet events. In (a), the quality of $S_{H_{max}}$ measurement is indicated by the length of the vector, where “A” is the highest quality, “B” is good, and “C” is moderate. We exclude lower-quality measurements from our analysis. (c) The subregion with normal color scale. Our study area is the gray, dashed box, with four TexNet stations (black triangles) near moderately-sized earthquakes. Deep wells with sonic logs used to create the 1D velocity model are the red dots.

pre-existing normal faults is a potential mechanism for the observed deformation in the southeastern zone of the Delaware Basin. The next chapter is devoted to testing this hypothesis with analytic models.

7.5 Chapter Summary

This chapter has introduced a case study in which we use InSAR to characterize the subsurface response to fluid extraction and injection in the Delaware Basin, TX, a hot-spot for industry-related induced seismicity. We demonstrated our generation of a basin-scale InSAR time series using SBAS techniques. The agreement between the three LOS look directions we used strongly suggest that our techniques produced robust results, despite having little opportunity to validate our time series with ancillary GPS measurements. Our InSAR results show widespread deformation that is spatially correlated with disposal and horizontal production wells. However, the displacement patterns differ between the portions of the basin to the north and south of the Grisham fault. Despite clear signs that the northern displacements can be explained by fluid volume changes near wells, the southeastern basin does not demonstrate such strong correlation to monthly well volumes. However, short-wavelength linear features are oriented parallel to horizontal stress conditions and trends in seismicity. We hypothesize that the deformation in the south includes a component of fault slip, which is the topic of the next chapter.

Chapter 8

Modeling Fault Slip in the Delaware Basin

The InSAR results in the Delaware Basin (Chapter 7) strongly suggest that subsurface volume changes are not the only response to fluid extraction and injection near regions with high seismicity. Curvilinear displacement features in the southeastern portion of the basin align with principal horizontal stress orientations and seismic trends, suggesting that they may be explained by slip on normal faults. This chapter is devoted to testing whether fault motion can explain these displacement features and, if so, characterizing the geometry and behavior of fault motion, within the context of local pore pressure and seismicity.

We determine the geometry and slip of potential faults using analytic modeling of both vertical and east-west horizontal displacements, focusing on a small study area along the border of Reeves and Pecos counties (see Figure 7.1b). We compare our results to an improved seismic analysis in the same study area, which is presented in a related paper by Sheng et al. (2022). In that work, they used a moment tensor analysis to determine focal mechanisms and depths for nine moderate events ($M_W > 2.7$), and phase arrival times to determine the depths of smaller earthquakes. When considered together, our study and Sheng et al. (2022) suggest high-angle normal faults in the Delaware Mountain Group are activated by wastewater injection. We conclude with a discussion of the implications for the nature of induced seismicity in the greater Delaware Basin.

8.1 Choice of Modeling and Study Area

We use an Okada edge dislocation analytic model (Okada, 1985) to test the hypothesis that normal fault slip is the source of linear deformation features in the southeastern zone of the Delaware Basin. In this model the surface displacements are caused by a slipping plane contained within a homogeneous, elastic half-space. Comparing such a fault model with the InSAR displacement field will indicate whether fault slip is a plausible mechanism for the expected i) geometry and location of the planes, and ii) range of slip magnitudes. These model results, however, need to make sense in the larger geophysical context, including the earthquake depths, focal mechanisms, and the spatial relationship of these earthquakes to the deformation. Therefore, to define a suitable study area, we identified a region satisfying the following criteria:

1. A simple, yet distinct, deformation feature with a clear preferred orientation in vertical and east-west horizontal InSAR components
2. Sufficient seismic station coverage to provide accurate focal depths
3. Earthquakes large enough to determine focal mechanisms ($M_W > \sim 3$)
4. Deep wells with sonic logs to define the local geologic and velocity structure

The first criterion defines the characteristics of the deformation feature we seek to reproduce using Okada edge dislocations. The latter three criteria address the required accuracy for the earthquake data, if we are to compare the deformation modeling results to seismicity.

The study area we selected is outlined by the dashed gray box in Figure 7.15 from Chapter 7. Although there are larger deformations elsewhere nearby (Staniewicz et al. (2020) modeled the area outlined in red), the area we have selected contains a relatively isolated, clear linear feature that exhibits both vertical and east-west horizontal components (Figures 8.1a and 8.1b, respectively) in the InSAR measurements, and aligns well with both seismicity from the TexNet catalog and the $S_{H_{max}}$ direction. However, the local wells show poor spatial correlation with the expected deformation from fluid volume and pore pressure changes. For example, as described in Appendix D and depicted in Figure D.1, there are few production wells (oil or groundwater) collocated with the observed subsidence along the linear feature of interest, and there is little-to-no uplift near active disposal wells. Therefore, explaining this deformation feature needs geomechanical mechanisms other than (or in addition to) radial changes in fluid volume. Also of note is that our selected study area coincides with the region identified by Teng and Baker (2020) as having the highest seismic hazard in the Delaware Basin. Thus, it is a region of particular importance for operation managers to understand.

We present the related seismic analysis in a companion paper by Sheng et al. (2022). In our study area, we determined moment tensors for nine events (Table 8.1) along with the relocation of numerous smaller earthquakes. This analysis used sonic logs from three deep wells in our study area (red circles in Figure 7.15c) to develop the local velocity model that tightly controls earthquake focal depth and moment tensor solutions. Earthquake focal depths concentrate between 1.5 and 3.0 km below ground level, with approximately 80% of the events located in the DMG; fewer than 2% are as deep as the Wolfcamp formation and none locate in the basement. All of the moment tensor solutions

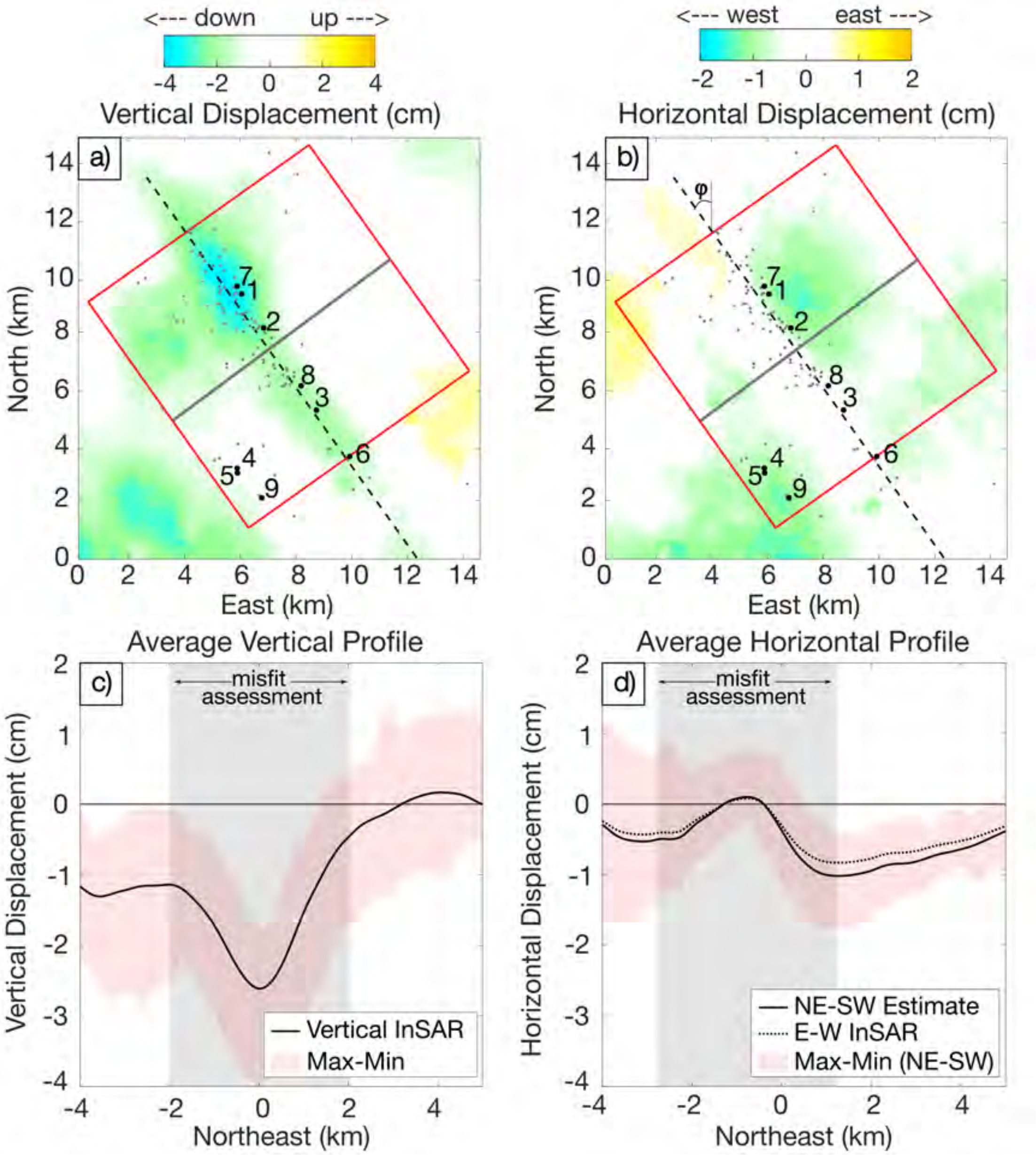


Figure 8.1: InSAR results in selected study area. (a) Vertical and (b) east-west horizontal cumulative InSAR deformation, with relocated moment tensors (black, numbered dots) and earthquakes (gray dots). Within the red boxes, we calculated the average vertical and horizontal profiles along the gray line, perpendicular to the midline (dashed black line), which we assume to be the azimuth (φ) of the predicted faults. The bottom panel shows (c), the average vertical profile, and (d), the average east-west horizontal and estimated northeast-southwest horizontal profiles. During modeling, we calculate the misfit within the shaded gray regions in (c) and (d).

are consistent with normal faulting on high-angle planes striking northwest-southeast, with the dip direction split almost evenly between northeast and southwest dips (Table 8.1). Sheng et al. (2022) also found no spatiotemporal correlation between fracked wells and the earthquakes, suggesting that they were not induced by hydraulic fracking; rather, they need to be explained by another driving mechanism, such as wastewater disposal, oil and gas production, or perhaps a combination of the two.

Hypocenters determined by Sheng et al. (2022) align with the linear deformation feature in our study area, as shown in Figures 8.1a and 8.1b. The dashed black line delineates the midline of the displacement feature of interest for initial analysis. Epicenters of the nine events with moment tensors are the numbered black dots, whereas smaller earthquakes determined through conventional location analysis are the gray dots. Earthquakes numbered 1–3 and 6–8 lie along the midline, thus we define them as Group 1, and the relocated smaller earthquakes are densely packed around the same feature. Events 4–5 and 9 (Group 2) form a smaller linear trend to the southwest of the midline, but striking in the same azimuthal direction. In addition, the strikes of the moment tensor solutions are sub-parallel to the azimuth of the midline and earthquake location trends, with predominantly normal slip. We now need to determine whether fault slip can also explain the deformation, if it is consistent with the seismicity, and how it might be related to oilfield activity. The next section is devoted to answering these questions.

8.2 Okada Edge Dislocation Modeling

We model surface deformation due to slip on normal faults using Okada edge dislocations (Okada, 1985), using the dmodels Matlab package (modified for ease of use with our data formats) from Battaglia et al. (2013). As shown in Figure 8.2, the basic 2D model is a plane of infinite length (extending into the page), parameterized by the dip direction and angle (θ), and depths to the top

ID#	Focal depth (km)	Strike	Dip	Rake	M$_W$
1	2.4 ± 0.1	152	82	-77	2.95
2	1.8 ± 0.2	146	68	-80	2.90
3	2.0 ± 0.2	150	70	-82	2.70
4*	1.4 ± 0.1	326	75	-83	2.84
5*	1.4 ± 0.1	327	74	-82	3.18
6	1.6 ± 0.2	326	70	-81	2.89
7	1.6 ± 0.1	336	63	-76	3.18
8	2.0 ± 0.1	166	81	-65	2.81
9*	1.6 ± 0.1	338	68	-78	2.76

Table 8.1: Moment tensor solutions (adapted from Sheng et al., 2022). Stars indicate the earthquakes that belong to Group 2; the others belong to Group 1. All solutions strike sub-parallel to one another and have predominantly dip-slip motion.

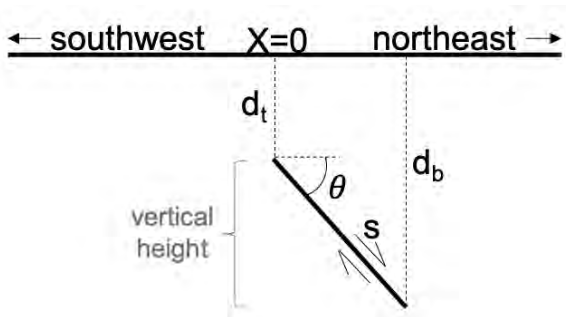


Figure 8.2: Schematic diagram of the fault geometry for a 2D edge dislocation in a homogeneous elastic half-space. In 3D, the predicted fault strikes northwest, thus the fault-perpendicular profile is in the northeast-southwest direction. Fault parameters are described in Table 8.2; we define the vertical height as the distance between the top and bottom edges of the fault in depth.

and bottom edges (d_t and d_b, respectively), contained within an elastic halfspace. In our approach, X is the lateral distance between the midline at x = 0 and the top edge of the fault, and s is the magnitude of slip in the down-dip direction. This 2D analytical model of surface deformation consists of only two components: vertical and fault-perpendicular horizontal. When extended to the 3D analytic model, the edge dislocation is a plane of finite length (L) and the surface deformation includes vertical, eastward, and northward components of motion. Due to the limitations of polar orbital paths, InSAR is less sensitive to northward motion, and we exclude this component from our modeling.

8.2.1 2D Modeling

We use the 2D model to constrain the approximate depth intervals of slip by comparing forward models of Okada edge dislocations to the measured InSAR data using a parametric sweep. Our initial assumption is that the linear feature of interest can be explained by a single infinitely long fault plane oriented parallel to the midline in Figures 8.1a and 8.1b. However, the study area undoubtedly consists of multiple deformation sources in addition to a single slipping fault which dominates the signal. In order to reduce the sensitivity of our analysis to these other sources, we created an average InSAR profile parallel to the solid gray line in Figures 8.1a and 8.1b, using data from within the red boxes. The resulting profiles are shown in Figures 8.1c and 8.1d. In (c), the vertical profile is the black line; however, in (d), the average east-west displacement depicted by the dashed black line is not strictly fault-perpendicular, as required in the data for the 2D modeling. It is not possible to determine the true northeast-southwest deformation from only two InSAR components; however, if we assume that the measured displacements along the linear feature are due to pure dip-slip motion on a fault parallel to the midline, then there is a unique solution to the required northeast-southwest displacements (H_{ne-sw}) via the trigonometric relationship in Eq. 8.1:

Parameter	Values	Notes
Dip direction	northeast or southwest	Strike parallel to midline (dashed line in Figure 8.1a)
Dip magnitude (θ)	5-90°	θ is an integer
Depth to top edge (d_t)	100, 200, ..., 6,300 (m)	—
Depth to bottom edge (d_b)	200, 300, ..., 6,400 (m)	100 m $\leq$ (d_b - d_t) $\leq$ 6,300 m
Location of top edge (X)	—	Determined from vertical model and InSAR
Poisson ratio (ν)	0.25	Kept constant

Table 8.2: Parameter space for 2D edge dislocation models.

$$H_{ne-sw} = \frac{H_{e-w}}{\cos\varphi}, \tag{8.1}$$

where H signifies horizontal motion and subscript e-w indicates east-west motion. Variable φ is the angle between North and the strike of the midline (35.3°), as shown in Figure 8.1b. The resulting fault-perpendicular displacement profile is the solid black line in 8.1d. In our model, we use the vertical and estimated northeast-southwest horizontal profiles as the reference data for misfit assessment within the gray regions in 8.1c and 8.1d. The chosen regions in each profile have the same number of measurements (n), but are offset from each other, such that the area in vertical is centered around the valley at 30 m and in horizontal is centered around the peak at -777 m. Beyond these regions, the InSAR profiles deviate from the expected deformation due to a single edge dislocation and are more likely to be influenced by other sources.

In the parametric sweep, we assess the fit of all forward Okada edge dislocation models characterized by the parameter sets developed from the values listed in Table 8.2. We selected a common value for the Poisson ratio (0.25) and kept it constant during modeling to simplify the parameter space. We determine the X-location for the top edge of the fault relative to the midline (x = 0) directly from the model: for a given parameter set i consisting of d_t, d_b, θ, and dip direction, we compute the vertical forward model of the dislocation with the top edge at x = 0 and 10 cm of normal slip, and then adopt the lateral offset between the minima in the vertical forward model and InSAR profile as the appropriate X-location.

With the full geometry for parameter set i defined, we determine the magnitude of slip (s) best-fitting the InSAR profiles by minimizing a modified RMS error (E), which we refer to as *misfit*, as defined in Eq. 8.2:

$$E_i = \sqrt{\frac{\sum_{i=1}^{n}((\hat{v} - (v_i + DS_{v_i}))/2)^2 + (\hat{h_i} - (h_i + DS_{h_i}))^2}{2n}}. \tag{8.2}$$

Here, $\hat{v}$ and $\hat{h}$ are the vertical and horizontal displacements, respectively, from the forward model, the un-hatted v and h are from the InSAR profiles, and n is the number of samples in the InSAR profile, within the misfit assessment bounds. Since our main goal in the 2D modeling is to fit the wavelength and relative amplitudes of the vertical and horizontal data, we allow datum shifts in each (DS_v and DS_h, respectively) during measurement of the misfit, such that the minima in vertical and maxima in horizontal between the forward model and data are equal (see Figure 8.7). We also weight the vertical differences by 1/2 in order to account for the higher amplitude in vertical motion compared to horizontal and better allow the latter to influence the solution. We prefer this weighted misfit assessment because a dip-slip edge dislocation results in vertical displacements that are approximately twice the amplitude of the horizontal, within our chosen misfit bounds, which is also the proportion observed in the InSAR profiles. Weighting the vertical differences between data and model by 1/2 results in a solution in which the proportion of differences to amplitude in each displacement component are comparable.

8.2.2 3D Modeling

While the 2D modeling is useful for constraining appropriate edge dislocation parameters, we require the 3D model to analyze the relationship of proposed faults to the local seismicity and well locations, and better understand the deformation due to slipping faults in the context of the InSAR displacements in the full study area. Using the dmodels package (Battaglia et al., 2013), we are able to extend any of the 2D, one-fault forward models to the full 3D space by adopting the X-location and uniform slip magnitude resulting from 2D modeling, and assigning finite length L (equal to the length of the midline) and strike direction (parallel to the midline). Observations from comparing these 3D, one-fault models to the full InSAR data inform our development of increasingly complex multi-fault models.

In the first stage of multi-fault modeling, we assume uniform slip on numerous edge dislocations of varying length. After selecting the number of faults (N) to include in the modeling, we manually select the endpoints of the top edge of each, thus defining their locations in the 3D space. For simplicity, we then select and assign identical d_t, d_b, and θ to each fault plane, but permit the strikes (as determined by the endpoints) and dip direction to vary on each, noting that we do not allow significant deviations ($\sim 10^\circ$) from the strike of the midline or linear trends created by the moment tensor solutions from Sheng et al. (2022). We then solve for the magnitude of uniform slip on each fault plane using the relationship in Eq. 8.3:

$$Wd = WGs', \tag{8.3}$$

where d is a vector of vertical and east-west horizontal InSAR data, s' is the unknown $N \times 1$ vector of slip magnitude on each fault plane, and G is the Green's function matrix relating slip magnitude to vertical and east-west horizontal surface deformation at each pixel, via the Okada (1985) equations. Matrix W is a diagonal weighting matrix that prioritizes data pixels near the fault segments. Along its diagonal is $1/R_i^2$, where R_i is the distance between data pixel i and the top edge of the nearest fault segment. We use dmodels (Battaglia et al., 2013) to generate the appropriate G matrix and apply Eq. 8.3 to find the vector s' of uniform dip-slip magnitudes that best fits the selected InSAR data in a least-squares sense.

After developing a uniform-slip, multi-fault model, we introduce additional complexity by discretizing each plane into finite patches approximately 1,000 m in length along strike and 200 m in down-dip width. The slip vector s' is now equal in length to the number of discretized patches. Equation 8.3 is significantly underdetermined, leading to an unrealistically rough solution of vector s'. Therefore, we include a smoothing operator that minimizes the 2D second-derivative of fault slip in the patch model, resulting in the regularized inversion relation shown in Eq. 8.4:

$$\begin{bmatrix} Wd \\ 0 \end{bmatrix} = \begin{bmatrix} WG \\ \alpha^2 D \end{bmatrix} s', \tag{8.4}$$

where α is the Lagrange operator that determines the weight put on the smoothing, and D is the second-order finite difference operator such that $\nabla^2 s' = Ds'$.

8.3 Modeling Results

8.3.1 2D Modeling

The purpose of the 2D, one-fault modeling is to constrain the approximate depth intervals (d_t to d_b) of slip. As indicated in Table 8.2, we explore vertical slip heights (d_b–d_t) ranging from 100 to 6,300 m, contained between depths of 100–6,400 m. The chosen range of d_t and d_b approximately corresponds to the complete geologic section above the basement (Figure 7.1c).

Our simple 2D model demonstrates significant parameter trade-offs. However, some parameter sets are geologically more realistic than others, despite having similar misfit to the InSAR data. To explore these trade-offs, we condense our parameter space to include fault width ($w = (d_b - d_t)/\sin\theta$), the approximate 2D stress drop ($\Delta\sigma_{2D} = 0.85\mu s/w$) (Kanamori and Anderson, 1975; Starr, 1928), and the midpoint depth of the dislocation. The 2D stress drop calculation requires knowledge of the shear modulus (μ); we used the P velocity ($V_p = 4.3$ km/s) and V_p/V_s ratio of 1.89 for depths between 2 and 3 km, from Sheng et al. (2022), and a density of 2700 kg m^{-3} to estimate a shear modulus of 14 GPa.

Figure 8.3 shows the trade-offs between stress drop and fault width for four subsets of southwest-dipping faults (vertical heights ranging between 100 and 1,500 m), colored by the midpoint depth range in gray-scale. All subsets display similar trends. For example, greater fault widths and shallower depths require lower stress drops to fit the data. Likewise, stress drops decrease across the subsets as vertical height increases.

We constrain the vertical height upper bound to 1500 m based on observations of the misfit

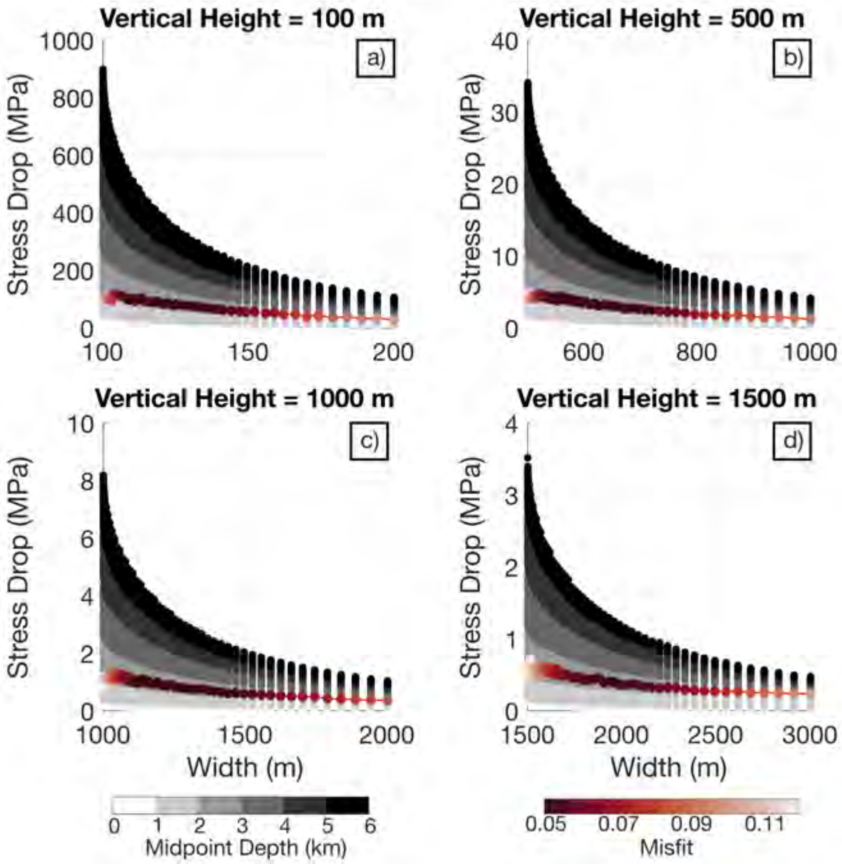


Figure 8.3: Results for the southwest-dipping faults with vertical heights of (a) 100 m, (b) 500 m, (c) 1000 m, and (d) 1500 m. As the vertical height increases, stress drop decreases due to widening of fault widths. For each subset, faults with deeper midpoints result in larger stress drops. In each subset, we also plot the best-fitting model in each width bin, colored by its misfit value (Eq. 8.2). For all widths, the best model has a midpoint around a depth of ∼2200 m, with a slight increase in depth with increasing vertical height.

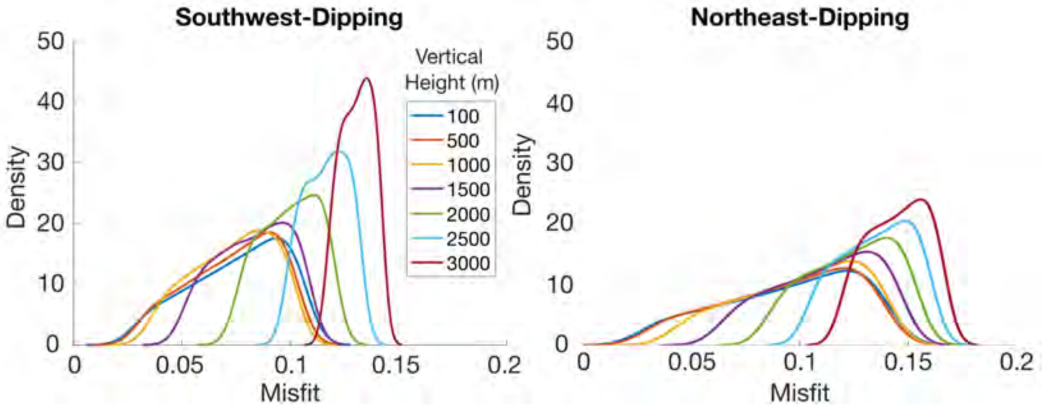


Figure 8.4: Densities of misfit values for the top 20% of models (with dips between 30°-90°) in specified vertical height (colors) and dip direction. The misfit values become increasingly higher with vertical heights greater than 1500 m.

densities in each vertical height subset, depicted in Figure 8.4. We split the full parameter sets into subsets by either southwest- or northeast-dipping (left and right columns, respectively), and by vertical height (colors). The densities correspond to the misfit values of the top 20% of models overall within the specified subset. In all cases, smaller vertical heights result in smaller misfits, and densities for vertical heights 100 m – 1000 m are similar. There begins to be significant deviation in misfit densities at 1500 m, with increasing deviation for all larger vertical heights. We thus set the upper bound of vertical heights at 1500 m for both dip directions. Next we require a lower bound.

Despite fitting the InSAR data, Figure 8.3 shows that most stress drops for models with vertical heights of 100 m exceed 100 MPa, which is unrealistically high. The stress drops reduce to < 35 MPa (much more reasonable values) for vertical heights of 500 m, and all explored models with vertical heights of 1,500 m have stress drops < 4 MPa. We thus constrain our parameter space to include only those models with vertical heights ranging between 500 and 1,500 m, to maintain a balance between plausible stress drops and model fit.

In addition to highlighting important trade-offs, Figure 8.3 shows that the best-fitting models have midpoint depths around 2200 m in all subsets of vertical height, regardless of fault width. While the best-fitting midpoint depth appears to be invariant to vertical heights and width, the midpoint depth also has low sensitivity to dip angle. In Figure 8.5, we show model subsets with vertical heights of 500 m (left) and 1,500 m (right), and with either southwest- or northeast-dip (top and bottom panels, respectively). The model misfit is shown as a function of midpoint depth and dip angle. Depending on the vertical height, the best 20% of models in each dip bin have mean midpoint depths between 1,900–2,500 m. These depths coincide with the Delaware Mountain Group, the formation in which wastewater disposal is concentrated and the majority of the earthquakes occur, suggesting a connection between fluid injection and fault movement.

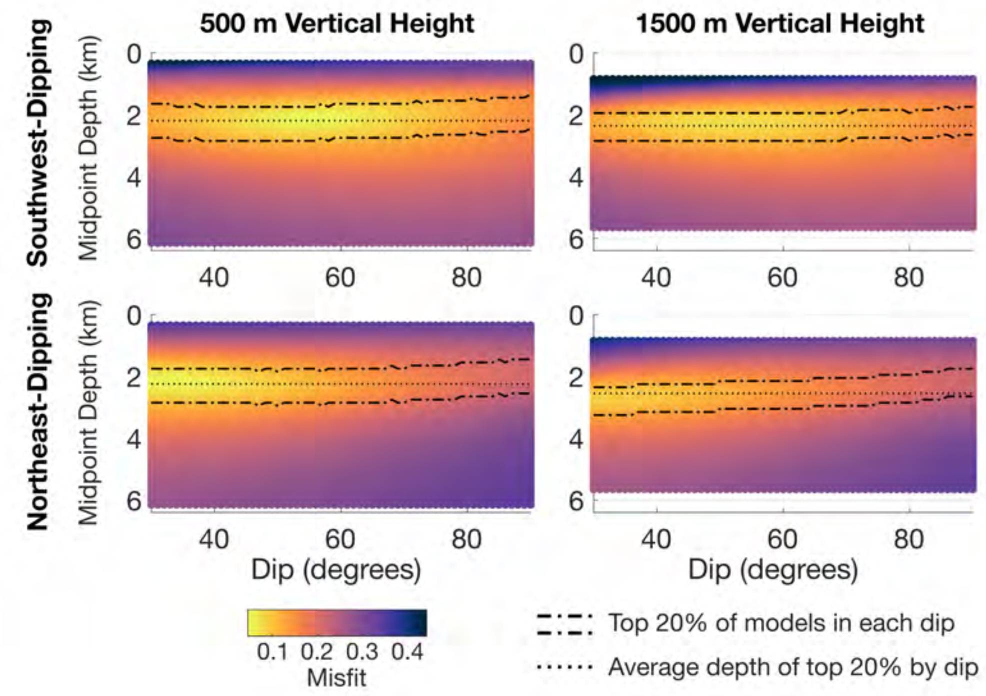


Figure 8.5: Misfit values for models as a function of midpoint depth, dip magnitude and direction, and vertical height. Regardless of parameter set, the top 20% of models in each integer dip bin have midpoints between ∼1.5 and ∼3 km. We exclude dips below 30° based on the local stress conditions (see Figure 8.6).

We are able to constrain the depth intervals and vertical heights from misfit assessment and geomechanical arguments about stress drop. We can do a similar exercise to constrain the expected dip magnitudes. Table 8.1 lists the high-angle moment tensor solutions from Sheng et al. (2022), which have a median dip of 70°. For each earthquake, there exists an auxiliary low-angle plane; these low-angle planes have a median dip of 22°. Although the moment tensor analysis alone cannot distinguish between the two dips, we can eliminate the low-angle dips based on the local stress conditions. In a predominantly normal-faulting stress regime, as is the case in the Delaware Basin (Lund Snee & Zoback, 2018), low-angle faults are the furthest from failure. Figure 8.6 shows (a), the Mohr circle derived from measurements of the principal stress components in the southern Delaware Basin from Dvory and Zoback (2021), and (b), the minimum increase in pore pressure (dP) required for fault failure as a function of dip. Not only are low-angle faults the least likely to slip, faults with dips $< 30°$ are precluded from slipping by the local stress conditions, since the change in pore pressure required would exceed the fracking threshold (dashed line in Figure 8.6b) and create microfractures in lieu of fault-reactivation. Therefore, we expect to see active high-angle faults with dips $> 60°$, consistent with the high-angle fault planes from the moment tensor solutions in Table

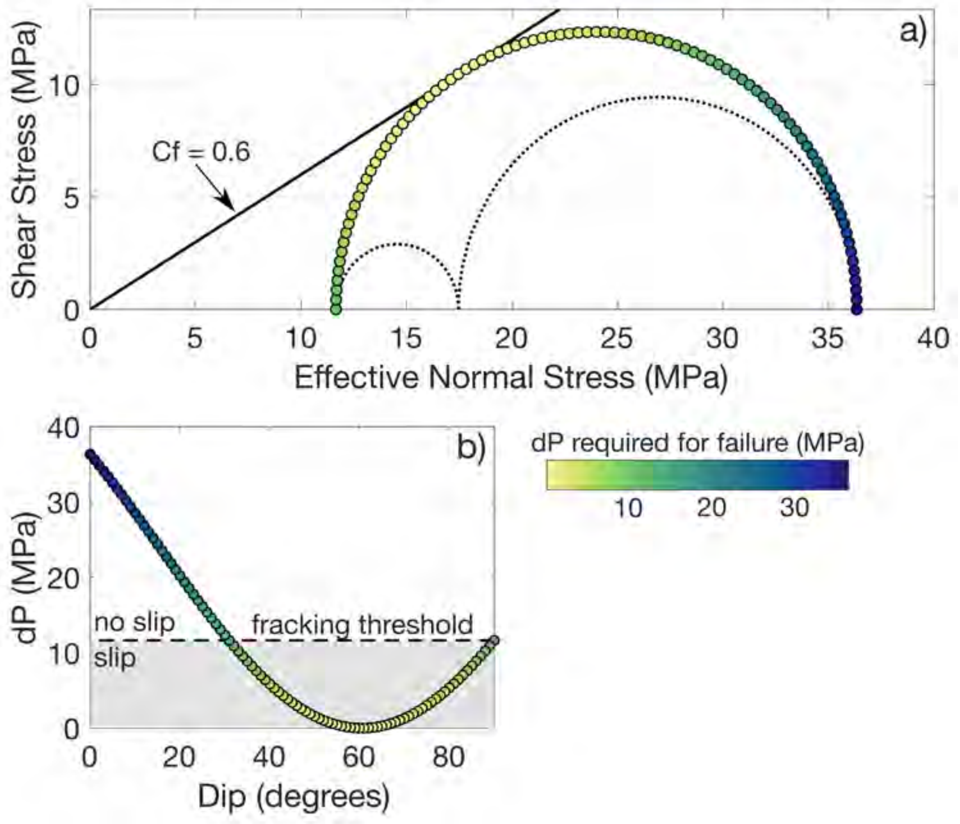


Figure 8.6: Minimum change in pore pressure (dP) required to reactive faults of specified dip. (a) Mohr circle and relative stresses for the Delaware Mountain group (Dvory & Zoback, 2021), assuming a coefficient of friction (Cf) of 0.6. (b) Minimum dP for slip as a function of dip. Any dP exceeding the fracking threshold (dashed line in b) will produce new microfractures, significantly reducing the local effective stress. Thus, faults with orientations of dP greater than the fracking threshold are not expected to slip.

8.1.

Figure 8.5 shows that the best-fitting one-fault models in our constrained parameter space have dips between 30° and 40° (northwest-dipping) or 50°–60° (southwest-dipping), suggesting that the southwest-dipping faults fit the InSAR data better under the constraint of high-angle dips. However, it is important to highlight that we allow a datum shift of the InSAR data during the misfit assessment, as demonstrated in Figure 8.7, which compares the 2D forward models of the best-fitting southwest- and northeast-dipping edge dislocations with dips of 75° and vertical heights of 1,000 m (see Table 8.3 for additional parameters). The dashed lines in each subplot are the InSAR measurements without datum adjustment. The southwest-dipping fault does indeed fit the datum-shifted InSAR profiles better than the northeast-dipping example. In contrast, we note that the horizontal InSAR profile as measured (i.e., no datum shift) is better represented by the northeast-dipping fault,

Parameter	Southwest-Dipping (Figure 8.7a/b)	Northeast-Dipping (Figure 8.7c/d)
Dip Magnitude (θ)	75°	75°
Vertical Slip Height (d_b-d_t)	1000 m	1000 m
Depth to Top Edge (d_t)	1400 m	1800 m
Depth to Bottom edge (d_b)	2400 m	2800 m
Slip Magnitude (s)	12.9 cm	13.2 cm

Table 8.3: Model parameters for example 2D slip results in Figure 8.7 and the 3D models results in Figure 8.8. We selected to find the best-fitting model given dip magnitude and vertical slip height. The other parameters listed are associated with the model that resulted in the smallest misfit.

though there is a sacrifice in vertical fit. These results suggest an ambiguity in the dip direction for a single fault that truly best fits the InSAR data.

8.3.2 Development of a 3D Model

The next step is to consider slip models of finite length and uniform slip in the full 3D space. Figure 8.8 depicts the 3D finite-fault model for each 2D model from Figure 8.7. Both models reasonably reproduce the vertical subsidence along the linear deformation feature of interest, albeit with lower magnitude than the InSAR observations. In the east-west horizontal component, however, the northwest end of the midline in the InSAR data appears to be dominated by a southwest-dipping fault, whereas the southeast end may be dominated by slip on a northeast dipping fault. Therefore, we explore the possibility of a two-fault model consisting of a combination of the oppositely-dipping single-fault models from Figure 8.8.

Using these observations and the expected slip interval depths constrained from the 2D model, we develop a model with two high-angle finite edge dislocations dipping toward each other in a graben structure, each with uniform slip, determined using Eq. 8.3. This intermediate, two-fault model is depicted in Figure 8.9. The southwest-dipping fault is rather short, but its extension along strike would contain the Group 1 earthquakes from Table 8.1, suggesting that the fault plane may be much longer, despite slip being concentrated in an isolated section. We thus extend each fault plane along its strike, increase the vertical width to 1,400–2800 m (to account for the range of the best 1-fault models from Table 8.4), and discretize each into multiple patches. We also include a small northeast-dipping fault parallel to the Group 2 earthquakes from Table 8.1. Using the regularized solution described in Eq. 8.4, we solve for the dip slip magnitude on each patch of the three defined faults, using $\alpha = 70$ due to its position on the bend of the L-curve of the solution semi-norm versus residual norm log-scale plot, depicted in Figure 8.10.

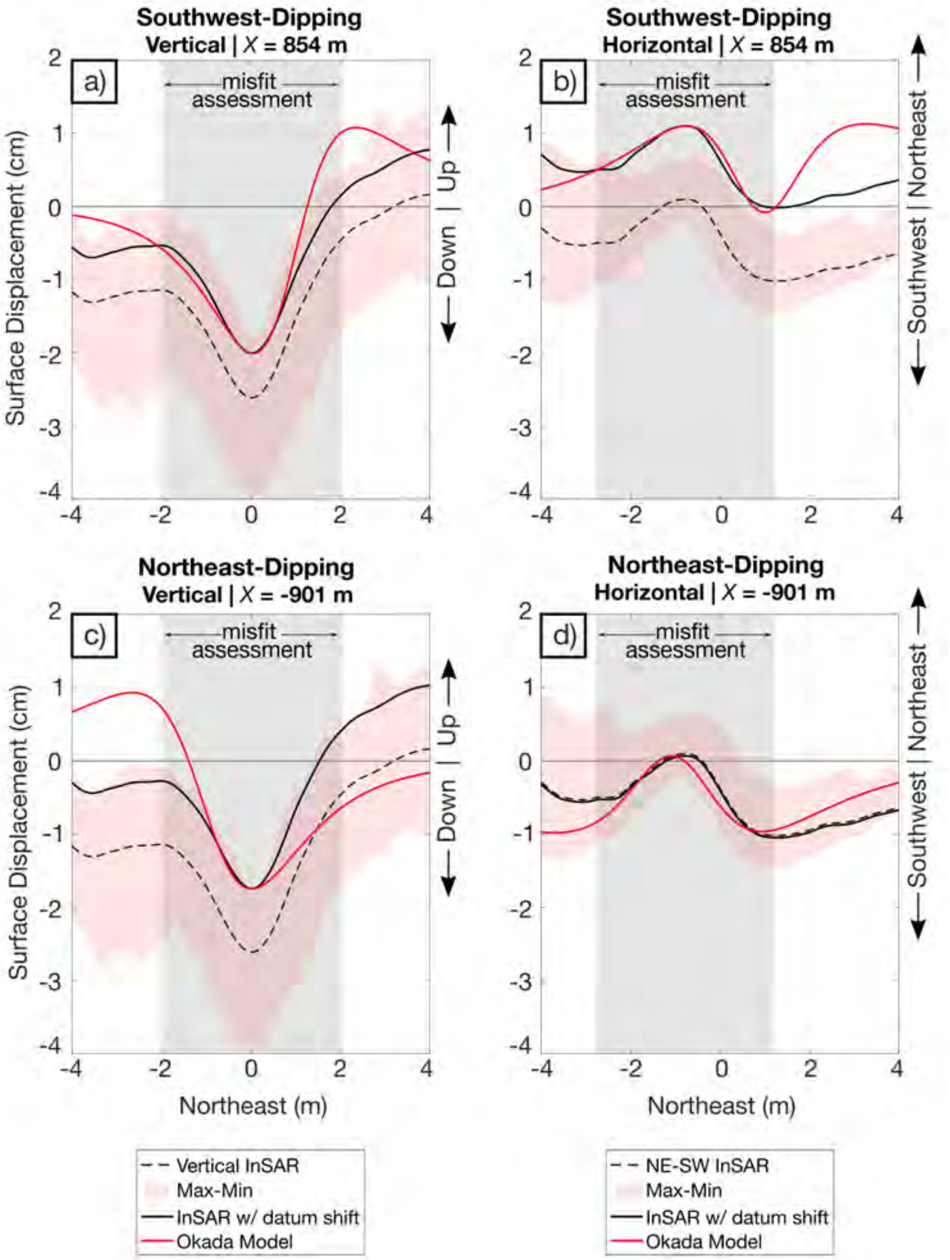


Figure 8.7: Forward models of two selected high-angle edge dislocations (parameter values described in Table 8.3). The vertical and northeast-southwest horizontal results are in (a) and (b), respectively, for the southwest-dipping fault and in subplots (c) and (d) for the northeast-dipping fault. The forward Okada models are depicted as red lines, whereas the InSAR profiles with and without datum shifts are shown as solid- and dashed-black lines, respectively. The misfit assessment bounds are the shaded gray regions.

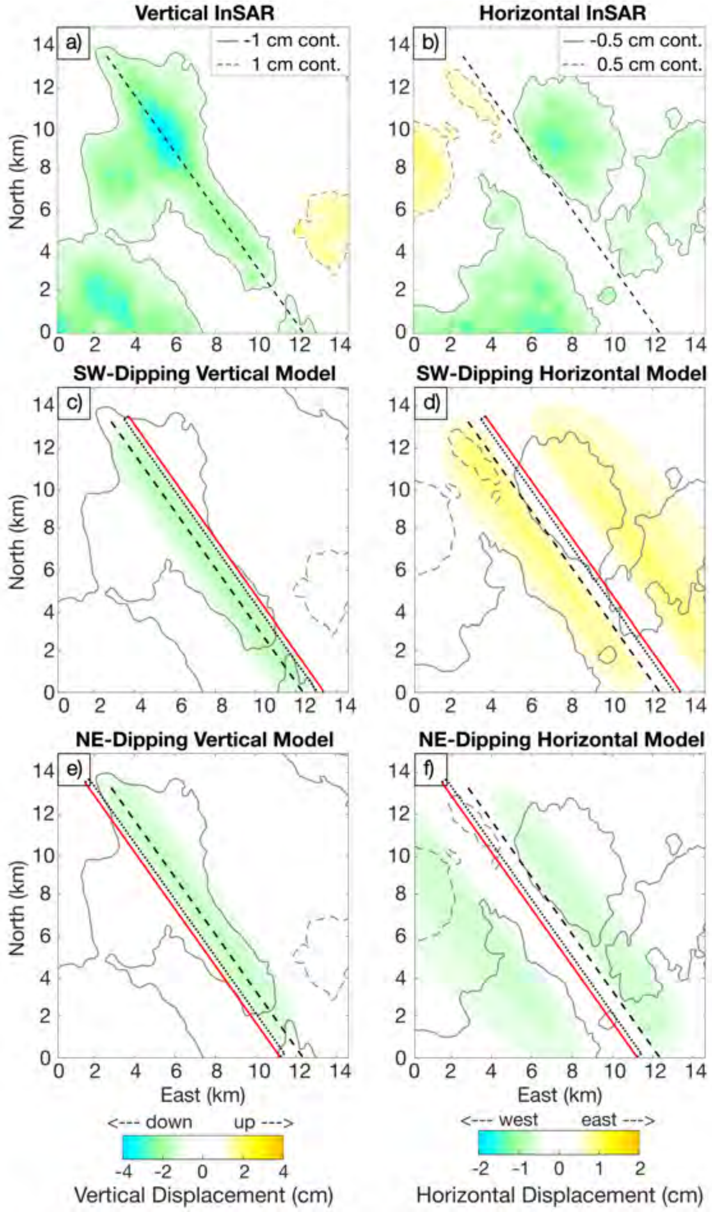


Figure 8.8: 3D edge dislocation modeling results. The top panel, (a) and (b), show the original InSAR vertical and east-west horizontal displacements, respectively. The middle panel, (c) and (d) are the forward model results for the southwest-dipping fault from Figure 8.7a and 8.7b and the bottom panel, (e) and (f), are the forward model results for the northeast-dipping fault from Figure 8.7c and 8.7d. The linear feature of interest is highlighted by the midline (dashed black line) and the gray lines outline the main deformation shapes as observed in the InSAR data. The extents of the finite edge dislocations are shown by the red lines (top edge) and the dotted black lines (map-view extent).

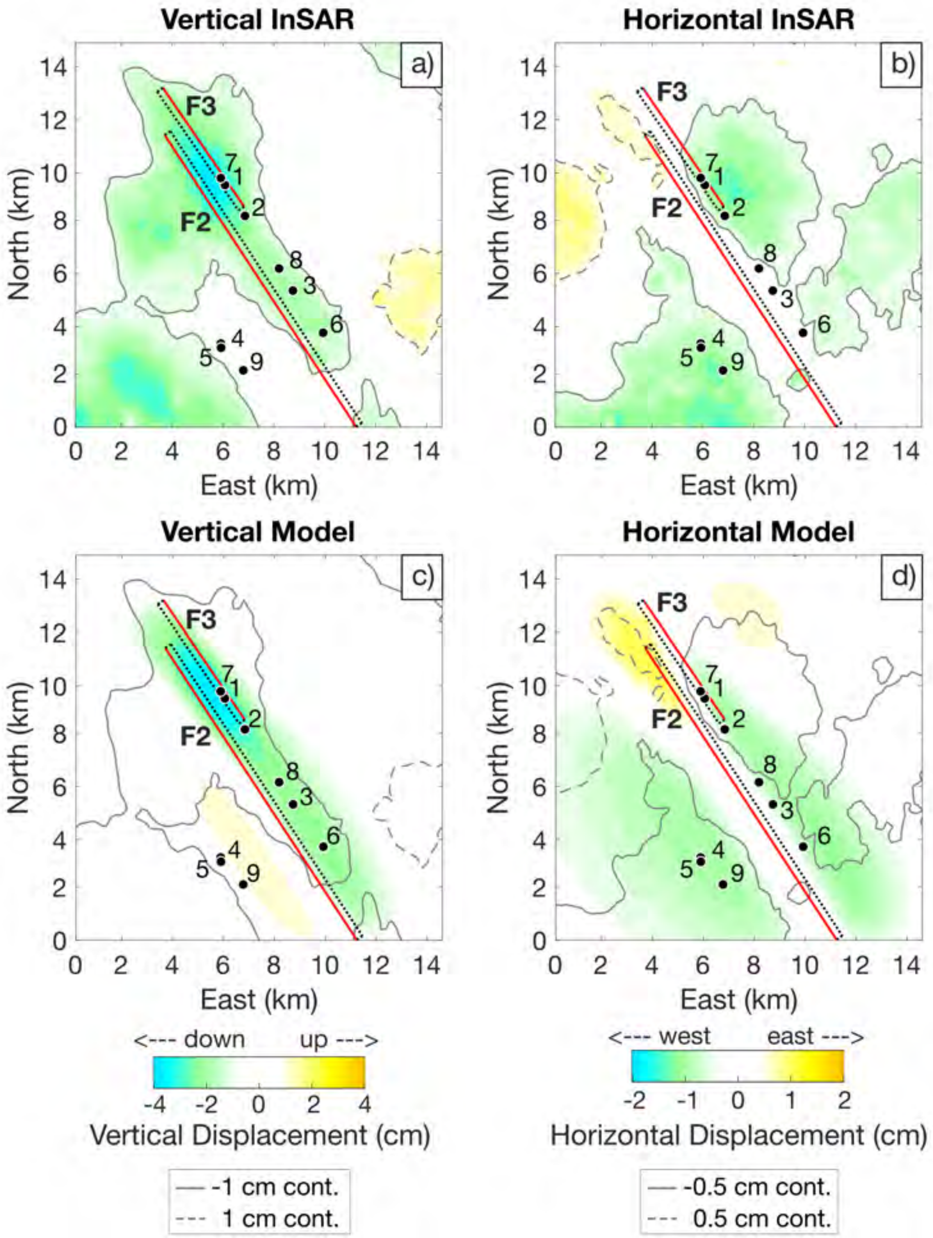


Figure 8.9: Two-fault forward Okada model. The original vertical and East-West horizontal InSAR measurements are depicted in plots (a) and (b), respectively. Plots (c) and (d) show the vertical and east-west horizontal forward models, respectively, from two finite edge dislocations. The parameters for each fault are listed in Table 8.4, where fault F3 is southwest-dipping and fault F3 is northeast-dipping. The red line is the top edge of the fault, and the dotted lines depict the map-view extent. The numbered moment tensor points from Sheng et al. (2022) are also pictured for reference.

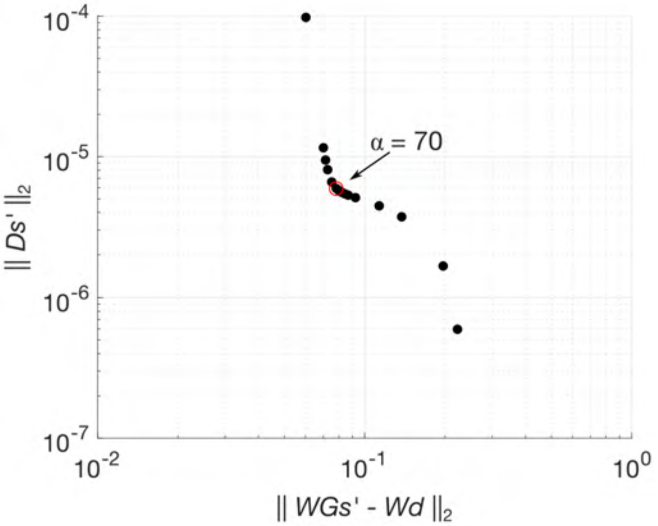


Figure 8.10: L-curve for determining an appropriate value for regularization parameter α, used in Equation 8.4. For the three-fault model, we select $\alpha = 70$, due to its position on the bend of the curve between the norm of the residuals (x-axis) and the solution semi-norm (y-axis), both plotted on log-scale. See Equations 8.3 and 8.4 for explanation of variables.

8.3.3 Three-Fault, 3D Model

We compare the forward model of vertical and east-west horizontal surface displacements from the patched, three-fault model to the measured InSAR data in Figure 8.11. The top edge of each fault (F1–F3) is marked by a solid red line and its downward-looking extent is outlined by the dotted black line. The slip distributions along each fault are shown in Figure 8.12, where (a) depicts the bird's eye view of the average slip along each fault's down-dip direction, and (b)–(d) display the side-view of each fault from the perspective of the arrow in Figure 8.12a. In Figures 8.12c–8.12d, we also include the along-strike profile of modeled (black line) and InSAR (red line) surface deformation directly

Parameter	Southwest-Dipping (F3)	Northeast-Dipping (F2)
Dip Magnitude (θ)	75°	75°
Vertical Slip Height (d_b-d_t)	1000 m	1000 m
Depth to Top Edge (d_t)	1600 m	1600 m
Depth to Bottom edge (d_b)	2600 m	2600 m
Slip Magnitude (s)	16.8 cm	18.0 cm
Length (L)	5622 m	13628 m

Table 8.4: Model parameters for example 3D, two-fault model in Figure 8.9. We selected the geometry of the two faults: dip magnitude, vertical slip height, depths to top and bottom edges, and length. The shown slip magnitudes produce the best fit to the InSAR data given Eq. 8.3.

above the top edge of traces F2 and F3, which flank the linear deformation feature of interest. In Figures 8.11 and 8.12a, we include the locations of earthquakes from Table 8.1, which highlight that Group 1 falls along the trace of F3, and Group 2 aligns with F1. Therefore, in Figures 8.12b–8.12d, we show only the earthquake locations in the side-view plots for the faults with which they are associated (Group 1 in Figure 8.12d and Group 2 in 8.12b). The final detail in Figure 8.12 is the location of active disposal wells (gray dots), which are labeled by the volume of injected fluid in millions of barrels (MMbbl) during the time span of our study. In Figures 8.12b–8.12d, we include only the disposal wells within 2 km of the fault planes. The perforation interval of each well is indicated by the solid black lines.

The main linear deformation feature in vertical (Figure 8.11a) is reproduced well in the forward model (Figure 8.11c); likewise, the horizontal deformation from the forward model (Figure 8.11d) agrees with the westward sense of motion flanking the linear feature in the InSAR data (Figure 8.11b), without requiring a datum shift. This model, however, cannot explain the two subsidence features to the southwest of F2, nor the uplift to the southeast of F3. Consequently, there are unmodeled displacement features in the horizontal component which coincide with the same geographical areas. In Figures 8.12c–8.12d, the comparison of the model and InSAR profiles also highlight some residual deformation that has not been captured by the model. These residuals are a direct result of our decision to favor smoothly varying slip models to prevent over-fitting the data with unrealistic slip distributions. The maximum amount of slip along F1 is mostly to the northwest of the Group 2 earthquakes, all of which have a northeast dip, agreeing with the dip of F1 (Figures 8.12a and 8.12b). Although there are no large earthquakes that spatially locate along the trace of F2, this fault has the greatest slip magnitude (25 cm) and extent of slip, as shown in Figures 8.12a and 8.12c. The majority of slip along F3 is confined between earthquakes #2, #7, and #1, and there is a small amount of slip ($\sim$7 cm) near earthquakes #8 and #3. We note, however, that the dip for earthquakes #6 and #7 are northeast, suggesting that they may belong to F2 or an additional unmodeled fault within the graben structure. In the former case, both northeast-dipping earthquakes would locate above the two local slip maxima on F2, whereas the latter case requires further modeling to draw any conclusions regarding the relationship to slip. The largest earthquakes do not collocate directly with the patches hosting the greatest predicted slip magnitudes, suggesting that the faults are principally slipping aseismically. Additional evidence stems from the timing of earthquakes #6–9, which all occurred after the end of the InSAR study period (post-March 2020). Thus, the observed slip only has the potential to be attributed to earthquakes #1–5, which have a peripheral relation to the greatest slip magnitudes.

While the majority of proposed slip cannot be attributed to the earthquakes, the regions of large slip along each fault trace do coincide with the location of disposal wells. In Figures 8.12b–8.12d, local areas of maximum slip lie between adjacent disposal wells. For instance, the patches of maximum slip on F2 lie between wells with disposal volumes of 6.4 and 17.5 MMbbl, with the absolute maximum

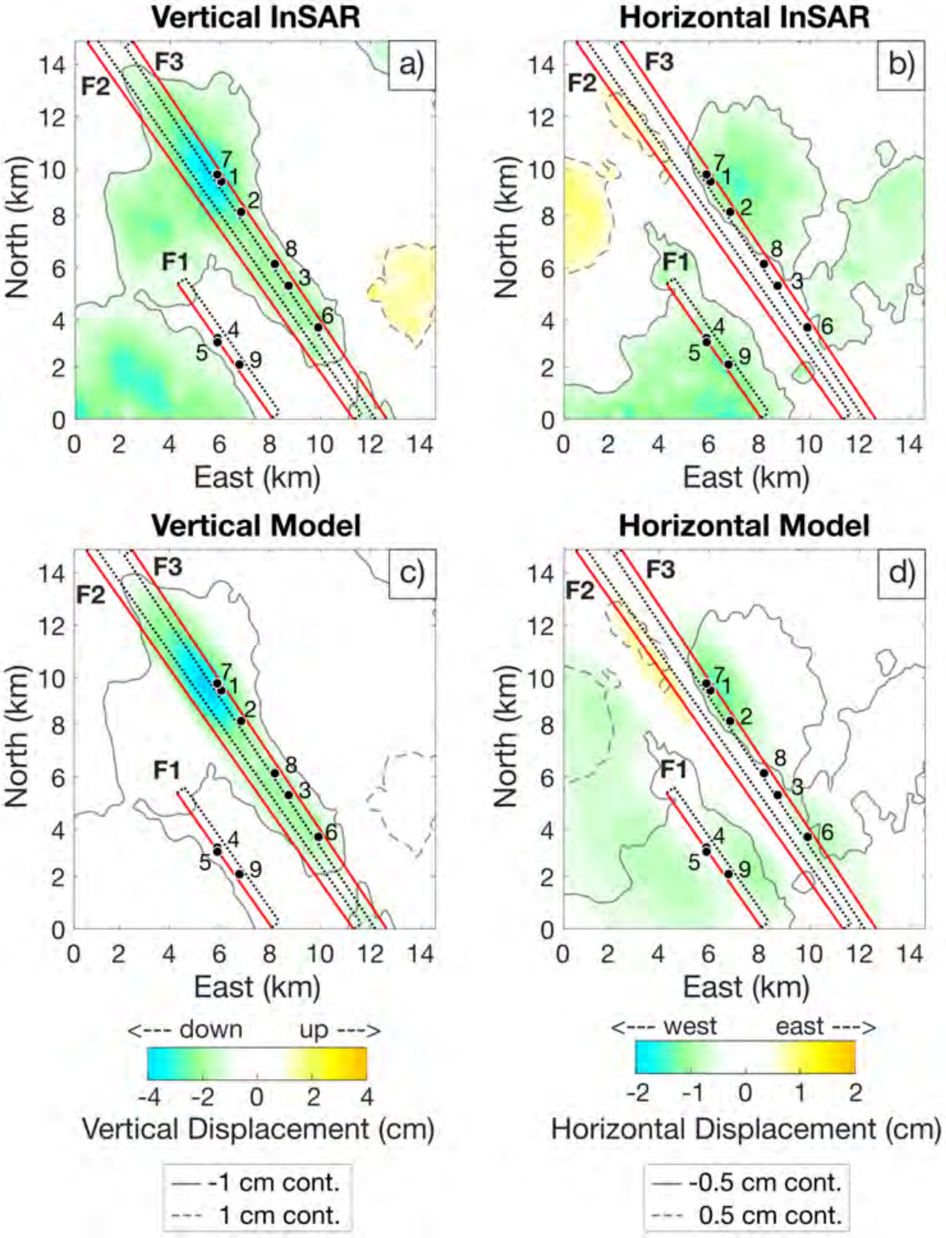


Figure 8.11: Three-fault model in a 3D space. The top panel is the original InSAR data, where (a) is vertical and (b) is east-west horizontal, with gray contours highlighting major features in order to better compare with the forward model in the lower panel, where (c) is the vertical forward model and (d) is the horizontal forward model. The two edge dislocations are represented by the red lines (top edge of fault) with map-view extent depicted by the dotted lines. The gray solid and dashed shapes highlight the displacement contours.

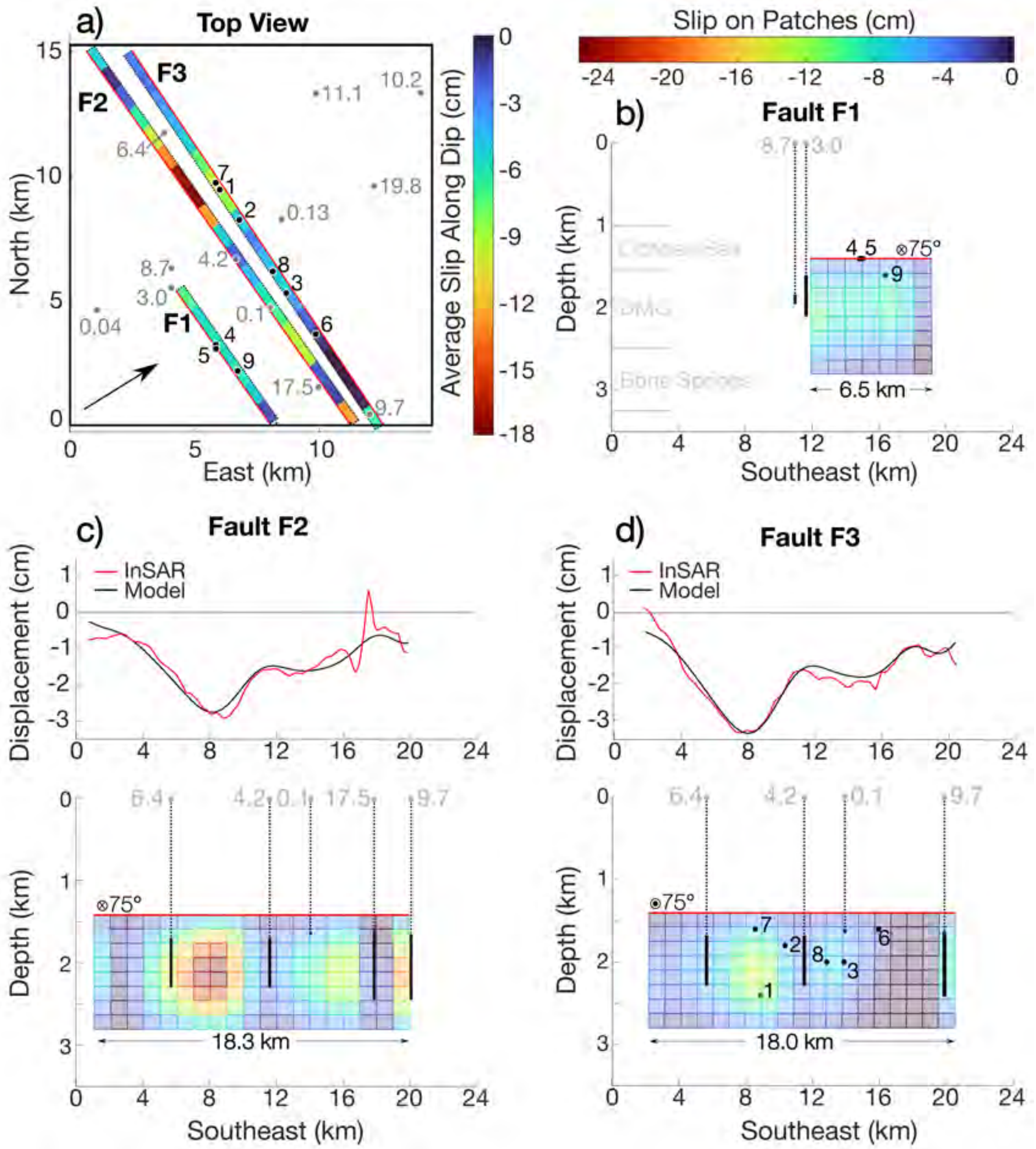


Figure 8.12: Slip distribution on the three-fault edge dislocation model. The top view in (a) shows the spatial relationship of the faults with the earthquake moment tensors (black dots) and disposal wells (gray dots). The numbers that accompany disposal wells are the values of cumulative injection volume between March 2015–March 2020, in millions of barrels (MMbbl). Plots (b)–(d) show the side view of each fault from the perspective of the black arrow in (a). Faults F2 and F3 (c and d, respectively) also display the InSAR and model surface deformation directly above the top edge of each respective fault. Earthquake moment tensors and disposal wells within 2 km of each fault are included in plots (b)–(d). Formation intervals are also indicated on the cross-sectional profiles in (b).

falling directly between wells with 6.4 and 4.2 MMbbl. Even on F3, where the maximum slip also lies adjacent to the well with 6.4 MMbbl, there is an observable increase in slip at the right edge of the fault that coincides with the well with 9.7 MMbbl of injection volume. Consequently, there is evidence for a link between fault slip and fluid injection in our study area.

Although the spatial relationship between fault slip and disposal is clear, there does not appear to be a direct correlation between the amount of slip and disposal volumes. However, there are many other variables to consider, including disposal rate, distance from the fault, and hydraulic and frictional properties of both the fault and surrounding subsurface medium. We note that the vertical InSAR profile along F2 (Figure 8.12c) shows signs of uplift directly above the disposal wells with the largest injection volumes, suggesting that the measured deformation may be due to the combination of many effects. In this case, it appears that injection-related uplift is superimposed on the subsidence signal from fault slip. The combined effects pose a challenge for isolating the true magnitude of slip on each fault patch. For example, on fault F2 near the disposal well with 17.5 MMbbl injection volume, there is a distinct column of little fault motion interrupting an otherwise smooth slip distribution on either side. It is possible that uplift related to the injection wells is causing an underestimation of the slip magnitudes, at this location and near other disposal wells along the fault traces.

8.4 Implications of Seismicity and Fault Motion

Our 2D and 3D edge dislocation model results show that the observed InSAR surface deformation can in part be explained by slip on high-angle normal faults within the DMG, with possible extension into the overlying Ochoan salts and underlying Bone Springs. In our small study area, our model consists of a long, shallow graben structure, and at least one other fault plane approximately 3–4 km to the southwest of the graben. Although there have been no detailed structural analyses in our study area, recent studies using 3D seismic arrays have mapped similar graben structures throughout Reeve's county (Charzynski et al., 2019; Hennings et al., 2021). All occurrences show graben structures mainly spanning the DMG, with slight extension into the Ochoan and Bone Springs. The grabens are all high-angle, $\sim$0.25–1 km wide (as measured by their top edges), and spaced 2–4 km apart. The three-fault model we developed has identical characteristics, suggesting it is a part of this larger graben network. The graben features align with the modern state of stress (Lund Snee and Zoback, 2018; Hennings et al., 2021) and are consistent with the stress state of the regional basin and range extensional episode in the middle to latest Miocene (Charzynski et al., 2019). However, there have been multiple extensional and contractional events since the formation of the basin during the Mississippian through the Permian (Hennings et al., 2021, and references therein).

The occurrence of deformation and the improved focal depth analysis from Sheng et al. (2022) highlight that these shallow grabens are not only present, but also active. In Figure 8.13, we

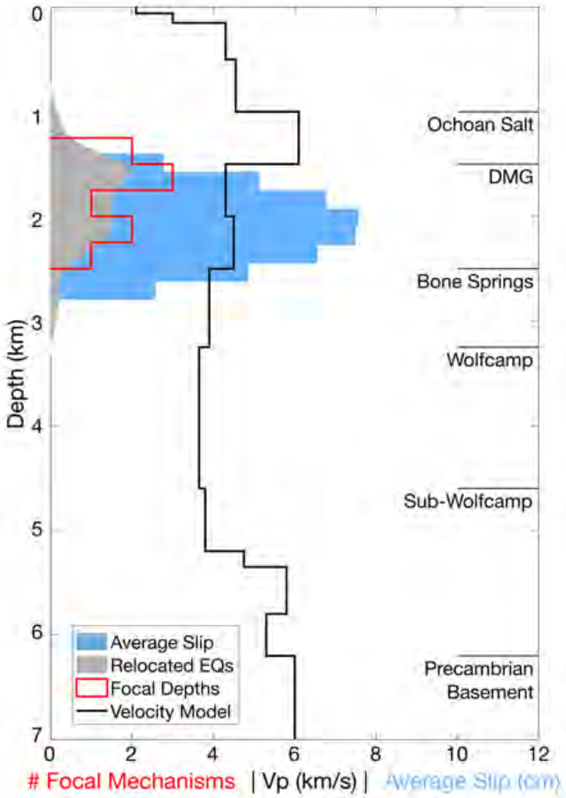


Figure 8.13: Summary of slip intervals from Okada modeling compared to the relocated earthquake depths (Relocated EQs), moment tensor centroid depths (Focal Depths), and velocity model from Sheng et al. (2022). All fault motion (seismic and aseismic) extends through the Delaware Mountain group (DMG), the main formation used for wastewater injection. We developed the local formation intervals using the average depth of each formation surface (Enverus, 1999) within our study area.

summarize the depth distribution of average slip (blue histogram), moment tensor centroids (red histogram), and relocated earthquake hypocenters (gray histogram), along with the 1D geological model Sheng et al. (2022) developed from the P-wave velocity profile (black line). All data peak at a depth of ∼2000 m in the middle of the DMG, which hosts all the local wastewater disposal. Not only do these data fall within the same formation, they have strong spatial relationships to one another. We were able to develop a discretized fault model that aligns with the larger earthquakes in our study area and agrees with the moment tensor solutions in terms of high-angle dip, as suggested by the local stress conditions, and sense of predominantly dip-slip motion. Furthermore, though we did not constrain our model with the available well data, wastewater disposal wells are located near patches of greatest slip on each fault. Therefore, it seems likely that the nearby fluid injection is activating these normal faults; however, the displacement is clearly not all seismic.

We calculate the cumulative geodetic moment along the patched surfaces of all three faults

F1–F3, using equations for seismic moment:

$$M_0 = \mu AS, \tag{8.5}$$

where μ is the shear modulus (we use a density of 2000 and 2700 kg m^{-3} to calculate a reasonable range of shear moduli between 10 and 14 GPa), A is the rupture area, and S is the average slip. To convert seismic moment (M_0) to moment magnitude (M_W), we use the definition from Hanks and Kanamori (1979) with M_0 in Newton-meters:

$$M_W = \frac{2}{3}(\log_{10} M_0 - 9.1). \tag{8.6}$$

The combined equivalent magnitude released during slip on all patches ranges between $M_W = 4.9$ to $M_W = 5.05$, whereas the combined equivalent magnitude of all earthquakes recorded by the TexNet array (between 01 January 2017 to 31 March 2020) in our study area is $M_W = 3.9$. Hence seismicity accounts for only 1.8%–2.5% of the predicted fault slip. If normal slip is contributing to the InSAR observations, as suggested by our model, it is predominantly aseismic.

To date, the role of aseismic slip in induced seismicity has been largely limited to indirect inference and associated with hydraulic fracturing (Cornet et al., 1998; Eyre et al., 2019; Eyre et al., 2020; Guglielmi et al., 2015; Zhu et al., 2020), so the implications of its occurrence in the Delaware Basin are challenging to know. Though Sheng et al. (2022) and our work suggest that wastewater disposal is likely inducing seismic and aseismic slip on normal faults in the DMG, it is unclear whether both are a direct consequence of the fluid injection, or whether aseismic slip triggers seismic events and/or vice versa. Based solely on our static 3D model, it is clear that the largest earthquakes along F1 and F3 do not coincide with the patches hosting the largest cumulative displacements (up to 25 cm), but rather are located around the periphery in patches with slip <12 cm. This suggests that hydraulic and frictional conditions vary along the faults.

Although our focus here has been on a small area in the Delaware Basin, we can extend our findings to the rest of the basin, which has contrasting deformation and seismicity patterns between the southern and northern sections. As demonstrated in the full-basin InSAR results (Figure 7.8), the linear deformation features only occur where there is seismic activity, suggesting that aseismic and seismic slip are intimately linked. Thus, the lack of seismicity and linear deformation features to the north of the Grisham fault could indicate that favorably oriented normal faults in the DMG are absent. However, this explanation lacks supporting evidence and is rather ad hoc. Dvory and Zoback (2021) analyzed the stress state and frictional stability of faults in the basin. They found that the fluid pressure in the DMG in the northern portion of the basin was diminished by conventional oil and gas production in that formation in the decades before unconventional exploitation began. Under this explanation, pressures are currently too low to induce fault slip, even under conditions of wastewater injection in the presence of favorably oriented faults. In contrast, Dvory and Zoback

(2021) found that the stress state is near-critical south of the Grisham fault, where very little production has occurred in the DMG. Thus, a modest pressure rise of a few MPa due to wastewater disposal in the DMG would bring favorably-oriented normal faults to failure, both seismically and aseismically. However, we note that not all of the observed deformation in the southeastern portion of the basin may be attributed to fault motion. There is evidence that fault slip is prevalent along the main subsidence feature in the study area investigated by Staniewicz et al. (2020) (Figure 7.15c), despite having a shape that is only sharply linear on the southwestern edge (though some of the subsidence associated with this feature may be due to horizontal production). On the other hand, the strong subsidence signals to the south of their study area (approximately longitudes 103.6 to 103.4 and latitudes 31.1–31.3) have wavelengths and magnitude that may be more consistent with shallow production, such as groundwater withdrawal used for hydraulic fracturing. It is also possible that some of the linear features in the southeastern basin are due to faults acting as hydraulic barriers to cross-flow, which may cause asymmetric deformation in the presence of nearby pumping or disposal. While the occurrence of seismicity along these features suggests at least some degree of fault motion, it may be a combination of these deformation mechanisms that are highlighting fault traces in this region.

Consequently, one limitation of our model is the assumption that the observed surface deformation is due exclusively to fault slip. More likely it results from the combined effects of fault slip (both seismic and aseismic), oil and gas production, wastewater and CO2 injection, and groundwater pumping for municipal, agricultural, and industrial purposes. Further evidence for multiple causes is clearer in the northern portion of the basin where there is observable deformation but no obvious patterns suggestive of fault movement. The contribution of other deformation sources within our own study area are clear in Figures 8.12c and 8.12d, where the smoothed slip model has up to 2 cm of misfit to the InSAR data, suggesting additional mechanisms contribute to the surface displacement. In particular, there is less subsidence in the InSAR data than predicted near some disposal wells, suggesting uplift from fluid injection. If the latter contributes to surface deformation, then we cannot rule out production-related subsidence as well, especially from shallow groundwater wells. Staniewicz et al. (2020) addressed the possibility of multiple deformation sources by removing the predicted vertical deformation from normal fault motion and computing residual vertical displacements resulting from subsurface volume changes. While forming a useful approach for modeling volumetric changes from fluid extraction and injection, including these in our model would not change our primary conclusion that high-angle normal faults in the DMG are moving.

It is essential to highlight the importance of including both InSAR components in the development of our model. The observations we made about the east-west horizontal deformation patterns produced from the single faults in Figure S7 directly guided us to the two-fault graben structure in Figure S8, both in Supporting Information S1. In addition, faults F1 and F3, which we in part defined to align with the focal depths and sense of slip of the nine larger earthquakes, cannot reproduce the observed InSAR deformation without the inclusion of fault F2. Had we used only the vertical deformation in the development of our model, we would have lacked the information needed to determine the geometry of all three faults, which altogether create a consistent story with the additional geophysical data available and recent works showing shallow graben structures in the DMG (Charzynski et al., 2019; Hennings et al., 2021).

8.5 Chapter Summary

Our InSAR analysis shows a stark contrast in deformation patterns between the northern and southeastern portions of the Delaware Basin. The three-fault model we developed from both components of these InSAR data suggests that fault motion is responsible for the linear deformation features in the southeastern portion of the Delaware Basin. Based on the spatial relationship between wastewater disposal wells, critically stressed faults, and relocated earthquakes, we have shown that wastewater injection in the DMG has likely been inducing both aseismic and seismic fault movement in this area. However, it remains unclear whether the aseismic slip and seismic events are both a direct result of pore pressure increase, or if induced aseismic slip triggers the seismicity or vice versa. Theoretical numerical modeling of injection-induced aseismic slip will be paramount to understanding the complex subsurface response to wastewater disposal, and our work provides observation-based slip models that can be used to constrain and contextualize these efforts. As we continue to explore the evidence for aseismic slip in the rest of the southern Delaware Basin and determine the likely geomechanical mechanisms contributing to deformation in the northern portion of the basin, it may be possible to constrain the conditions that lead to aseismic and seismic slip, so operators can better plan the location and operating standards for future wells.

Chapter 9

Conclusion

This thesis has demonstrated multiple aspects of the workflow guiding the use of radar imaging in hazard management, with an emphasis on characterizing the reservoir response to anthropogenic fluid extraction and injection. Once characterized, we can begin developing tailored management practices that mitigate the hazards associated with the response (e.g. ground deformation, induced earthquakes, or depletion of groundwater resources). The significant research contributions of this thesis are summarized as follows:

1. Identification of an over-looked, biased error source in InSAR: Aliasing during 2D phase unwrapping
2. Mathematical description of the conditions that lead to aliasing when using a common unwrapping algorithm, SNAPHU (Chen, 2001)
3. Demonstration of common InSAR practices that increase the likelihood of aliasing
4. Demonstration of how aliasing affects InSAR time series in time and space
5. Development of two-component InSAR deformation time series in the Delaware Basin, TX
6. Identification of shallow aseismic slip in response to wastewater disposal in the Delaware Basin, TX
7. Development of a technique for the selection of sites for surface-spreading recharge with high-resolution electromagnetic imaging methods

9.1 Summary of Contributions and Perspectives

The first step in InSAR applications is to ensure we are using high-quality deformation measurements. Our work identifies a source of biased error – aliasing during 2D phase unwrapping – that can greatly impact our analyses, even in the absence of noise. We first characterize the occurrence of aliasing in InSAR and demonstrate that ubiquitous InSAR practices, such as spatial filtering, increase the occurrence of aliasing due to a decrease in spatial resolution. Furthermore, we show that coherence, a common InSAR observable used to quantify the quality of our data, does not adequately describe the quality of unwrapped phase measurements because aliasing is a global error that can impact pixels with perfect signal-to-noise ratio.

If left unidentified, aliased measurements will lead to a systematic underestimation of displacement magnitudes, which will likewise bias a time series generated with SBAS techniques. We identify a correlation between temporal baseline and aliasing that is driven by the evolution of the displacement field, with examples from synthetic data and real Sentinel-1 observations in Hawaii, West Texas, and the Central Valley in California. Our results show that the selection of a maximum temporal baseline (or a quality set of unwrapped phases) in SBAS must balance the risk of aliasing with the potential benefits of reducing random noise. This phase set will be unique to each study area and will possible vary pixel-to-pixel. This work not only characterizes a significant error source, but calls into question our treatment of InSAR data. While the idea of "InSAR everywhere, all the time"[1] is a noble goal trumpeted by leading InSAR specialists, there is clearly not a single workflow that applies to every study area nor every application. Aliasing itself, for instance, is controlled by the gradient field in time and space, in addition to the local surface conditions relative to the displacements. We must adapt our methods to each study area and rigorously assess the results for signs of biased inaccuracies. Often, this requires generating multiple potential solutions and seeking ancillary data for validation. The rewards are worth the effort.

Generating high quality InSAR data is only the first step in our InSAR workflow; next we must interpret the observations in the context of the hazard we wish to characterize. We demonstrate an example of this process with a case study of induced seismicity in the Delaware Basin, TX. Using two InSAR look directions, we develop time series of both vertical and east-west horizontal displacements, which is the first of its kind in this study area. Having multiple displacement components and additional geophysical data (e.g. earthquake mechanisms) allowed us to infer that slip on normal faults may be contributing to the InSAR observations. We support our hypothesis with analytic models of ground deformation due to moving faults (edge dislocations), finding slip magnitudes of up to 23 cm. Our final three-fault model links wastewater disposal to fault-reactivation in shallow formations (the Delaware Mountain Group) and shows that the majority of slip is aseismic (i.e. does not produce an earthquake). This work is one of the first direct observations of asesimic slip as a response to wastewater disposal, which has already proven useful in constraining numerical models of induced aseismic slip. We are now one step closer to understanding the seismic risk in the Delaware Basin and how to mitigate it.

Interested readers may also have perused our new method for identifying quality sites for surface

---

[1] A nod to a presentation given by my advisor, Howard Zebker, at an AGU Fall Meeting.

spreading recharge with an electromagnetic imaging technique, tTEM. This project demonstrates the last step in our workflow: development of mitigation strategies to offset the effects of fluid pumping. We first generate simplified models of sediment properties that control fluid flow (flow or no-flow), then predict the paths that water will follow from the surface to the aquifer below. Characterizing these pathways allows us to compare sites in terms of their suitability for managed aquifer recharge (in the context of the goals and risks involved). This project highlights an important compromise between scientific rigor and practical methods. Though a complete hydrogeologic analysis and numerical model would better represent fluid flow at any given site, these require time and money – luxuries that are not often readily available to groundwater managers. With these limitations in mind, our method leverages generalized characteristics of sediment type and fluid flow to efficiently identify promising sites while quickly eliminating others. These will save both time and money in groundwater management projects.

9.2 Future Directions

Though we have made important advances of the use of InSAR in this thesis, it is clear that our work is far from complete. Indeed, many of the results raise more questions than they answer. While we have identified and characterized aliasing in InSAR and InSAR time series, we require better identification techniques. One promising method is the comparison of wrapped and unwrapped closure phase to illuminate triplet sets with aliased measurements. Next, we wish to develop a method that uses this comparison to automatically generate a set of interferograms that is free of aliased data – a unique set for each pixel in the study area. Alternatively, our work highlights the need to update the cost functions in SNAPHU so that we may avoid aliasing in the first place. We imagine that these cost functions, unique to contoured displacement fields, will account for the type and degree of spatial filtering used on the interferograms. Likewise, it would be interesting to explore whether the knowledge of nearby gradients and/or the number of complement gradients required to satisfy the condition of adequate sampling can be used to inform the residue connections in the branch-cut network. Furthermore, we need to develop a new parameter that quantifies the quality of an unwrapped phase, because coherence, as we learned, can only be applied to the wrapped measurements. We plan to explore the derivation of such a variable using the agreement of gradients within a search window.

Our work in the Delaware Basin also raises important questions that need to be addressed. For example, we need to use more modeling to determine the connection between seismic and aseismic slip – does one induce the other, is there a feedback loop, or are they simply two potential responses to the nearby fluid injection? We should also validate the assumptions used in our modeling approach. Specifically, we assumed that all of the displacement could be explained by slip on normal faults. Our slip model would likely be improved through a joint inversion of well data and InSAR to simultaneously solve for volume changes (or changes in pore pressure) and fault slip. Finally, we hope to extend our analysis to the other linear features in the southeastern portion of the basin, which may allow us to infer the presence and geometry of additional unmapped faults. Given the widespread deformation in this basin, it will likely require a team of scientists to decipher the combined effects of fluid extraction and injection. We are not lacking interesting signals to study.

Appendix A

SNAPHU Cost Functions

The general probability density function (PDF) used to frame SNAPHU's (Chen, 2001) statistically-based cost functions represents the likelihood that the solution of the unwrapped gradient field ($\Delta\Psi$) is correct given the wrapped gradient field ($\Delta\Phi$):

$$f(\Delta\Psi|\Delta\Phi). \tag{A.1}$$

There are two general assumptions that make it possible to develop this PDF and find the optimal solution. First, the individual unwrapped gradients ($\Delta\psi$) are separable, or statistically independent given their wrapped counterparts ($\Delta\phi$). In addition, the resulting unwrapped phase field is irrotational (i.e., free of residues). Given these assumptions (the former of which is not strictly correct) the objective can be expressed as a product of the individual PDFs associated with each unwrapped and wrapped gradient pair k:

$$f(\Delta\Psi|\Delta\Phi) = \prod_k f(\Delta\psi_k|\Delta\phi_k). \tag{A.2}$$

Expressed in this way, the objective is to maximize the product shown in Eq. A.2; however, we can transform it to a more-familiar representation which minimizes a sum by using the negative logarithm of the PDFs:

$$\text{minimize} \left\{ -\sum_k \log(f(\Delta\psi_k|\Delta\phi_k)) \right\}. \tag{A.3}$$

If we assume congruence, then we need only consider the non-conditional PDFs of $\Delta\psi_k$ and the cost function used in SNAPHU becomes:

$$g_k(\Delta\psi_k, \Delta\phi_k) = -\log(f(\Delta\psi_k)) \;\text{ for } \Delta\psi_k \equiv \Delta\phi_k + n2\pi, \tag{A.4}$$

where n is any integer.

While we have separated out the conditionality on the wrapped gradient, there are other observables in the interferogram that may be useful as additional conditions in $f(\Delta\psi)$, depending on the application. These observables are the nearby amplitude (i.e. intensity, I) and coherence (ρ); we can update Eq. A.4 accordingly to its general conditional PDF:

$$g_k(\Delta\psi_k, \Delta\phi_k) = -\log(f(\Delta\psi_k|I, \rho)) \text{ for } \Delta\psi_k \equiv \Delta\phi_k + n2\pi. \tag{A.5}$$

Appendix B

Residue Conditions

When we calculate residues, we require that all gradients in the loop integral of Figure 5.3b are less than 1/2 cycle (π rad). The conditions we have imposed on our gradients lead to a simple criterion that will dictate whether or not the loop integral will result in a residue. Namely, we will observe a residue if and only if we use an odd number of complement gradients, and there will be no residue when we use an even number of complement gradients.

We demonstrate this criterion by first showing that any set of $x_i \in [0,\ 1)$ cycles will have no residue when we calculate the general loop integral with zero substitutions of complement gradients ($\oint dx^{(0)}$):

$$\begin{aligned} \oint dx^{(0)} &= dx_1 + dx_2 + dx_3 + dx_4 \\ &= (x_2 - x_1) + (x_3 - x_2) + (x_4 - x_3) + (x_1 - x_4) \\ &= (x_1 - x_1) + (x_2 - x_2) + (x_3 - x_3) + (x_4 - x_4) \\ &= 0. \end{aligned} \tag{B.1}$$

If we then substitute any one of these gradients dx_i with its arbitrary complement ($\hat{dx_i} = dx_i + r_i$, $r \in \{-1, 1\}$), then the loop integral with one substitution ($\oint dx^{(1)}$) will necessarily have a residue:

$$\begin{aligned}\oint dx^{(1)} &= dx_1 + \hat{dx}_2 + dx_3 + dx_4 \\ &= dx_1 + (dx_2 + r_2) + dx_3 + dx_4 \\ &= (dx_1 + dx_2 + dx_3 + dx_4) + r_2 \\ &= r_2 \\ &= \pm 1.\end{aligned} \tag{B.2}$$

Two substitutions will force our loop integral ($\oint dx^{(2)}$) to be $\in \{-2, 0, 2\}$, depending on whether the two r_i values cancel or add:

$$\begin{aligned}\oint dx^{(2)} &= dx_1 + \hat{dx}_2 + \hat{dx}_3 + dx_4 \\ &= dx_1 + (dx_2 + r_2) + (dx_3 + r_3) + dx_4 \\ &= (dx_1 + dx_2 + dx_3 + dx_4) + (r_2 + r_3) \\ &= (r_2 + r_3) \\ &= \begin{cases} \pm 2 & r_2 = r_3 = \pm 1 \\ 0 & r_2 = -r_3 \end{cases}\end{aligned} \tag{B.3}$$

Continuing in this manner will lead us to find the remaining two cases for three and four uses of complement gradients, $\oint dx^{(3)}$ and $\oint dx^{(4)}$, respectively:

$$\oint dx^{(3)} \in \{-3, -1, 1, 3\} \tag{B.4}$$

$$\oint dx^{(4)} \in \{-4, -2, 0, 2, 4\}. \tag{B.5}$$

Given Eqs. B.3-B.5, it appears that we might see loop integrals outside the set of $\{-1, 0, 1\}$; however, we claimed in Chapter 4.2 that residues only take on values of ±1. This restriction is a result of our use of complement gradients only when the original gradient exceeds 1/2 cycle, rather than arbitrarily-selected gradients. If there are two or four gradients greater than 1/2 cycle, the included r_i will cancel. Consider, for example, when $|x_1| \leq |x_2| < 0.5$, and $0.5 \leq |x_3| \leq |x_4| < 1$. Then,

$$\begin{aligned}
\oint dx^{(0)} &= dx_1 + dx_2 + dx_3 + dx_4 = 0 \\
(dx_1 + dx_2) &= -(dx_3 + dx_4) \\
1 > |dx_2 + dx_2| &= |-(dx_3 + dx_4)| \\
1 > |dx_3 + dx_4| &= |(\hat{dx}_3 + r_3) + (\hat{dx}_4 + r_4)| \quad \text{(substitute)} \\
1 &> |(\hat{dx}_3 + \hat{dx}_4) + (r_3 + r_4)| \\
\text{(since } |\hat{dx}_3 + \hat{dx}_4| &< 1, \text{ and } (r_3 + r_2) \in \{-2, 0, 2\}) \\
r_3 + r_4 &= 0 \\
\Rightarrow r_3 &= -r_4
\end{aligned} \tag{B.6}$$

We can perform a similar exercise to show that three required uses of complement gradients will result in the cancellation of two of the r_i, and four required uses will have all four r_i balance to zero.

The occurrence of a residue thus depends solely on the number of complement gradients required under the condition of adequate sampling and we can update Eq. 5.2 accordingly:

$$\oint dx_{as} = \sum_i^n dx_i + \sum_j^m \hat{dx}_j = \begin{cases} 0 & m \text{ even} \\ \pm 1 & m \text{ odd.} \end{cases} \tag{B.7}$$

Appendix C

Effects of Random Noise in SBAS

To determine the degree of correlation between independent and redundant interferograms, we begin with the wrapped measurements and the concept of wrapped closure phase. Many studies have shown that the wrapped phases in non-filtered interferograms are completely correlated between phase triplets (Agram and Simons, 2015; De Zan et al., 2015; Michaelides et al., 2019). That is, the wrapped set of phase triplets $[\phi_{ij}, \phi_{jk}, \phi_{ik}]$ will have a closure phase identically equal to zero, where we calculate the wrapped closure phase (ϕ_{clos}) in exactly the same way as Eq. 6.6:

$$\phi_{clos_{(ijk)}} = \phi_{ij} + \phi_{jk} - \phi_{ik}. \tag{C.1}$$

However, when we apply spatial averaging (either in the form of multilooking or other types of filters, $\phi_{clos}^{(n)}$ is typically *not* identically zero (where superscript (n) denotes the number of looks). This suggests that the filter has introduced some degree of random noise that is not perfectly correlated between phase triplet sets (Agram and Simons, 2015; De Zan et al., 2015; Michaelides et al., 2019). Regardless of the source of this random noise (e.g. soil or vegetation moisture), these random, independent phase terms should propagate to the unwrapped interferograms if we assume perfect unwrapping. If so, then the unwrapped closure phase ($\psi_{clos}^{(n)}$) will equal the wrapped closure phase (with a potential $2m\pi$ ambiguity, m an integer).

We avoid making any assumptions about which phase components should be completely correlated versus statistically independent in redundant sets, and instead generalize these components similarly to Michaelides et al., 2019, by:

$$\psi_{ij}^{(n)} = \psi_{stat_{ij}} + \psi_{inc_{ij}}, \tag{C.2}$$

where all of the correlated, or static terms (e.g. deformation, atmosphere, topographic errors), are indicated by subscript *stat*, and the random noise unique to each interferogram (i.e. incidental) is denoted by subscript *inc*. With this formulation, we can decompose the unwrapped closure phase into its constituents:

$$\psi^{(n)}_{clos_{(ijk)}} = (\psi_{stat_{ij}} + \psi_{inc_{ij}}) + (\psi_{stat_{jk}} + \psi_{inc_{jk}}) - (\psi_{stat_{ik}} + \psi_{inc_{ik}}). \tag{C.3}$$

Because this linear combination of correlated components should be identically zero, this reduces to only including the random, incidental terms:

$$\psi^{(n)}_{clos_{(ijk)}} = \psi_{inc_{ij}} + \psi_{inc_{jk}} - \psi_{inc_{ik}}, \tag{C.4}$$

which will typically not be zero.

Given the separable properties of linear inversions, we can separate the terms in Eq. 6.3:

$$\psi^{(n)}_{stat} + \psi^{(n)}_{inc} = Gv, \tag{C.5}$$

and the inversion to model the linear velocity between each time step is:

$$v = G^{\dagger}(\psi^{(n)}_{stat} + \psi^{(n)}_{inc}) \tag{C.6}$$

and

$$\begin{aligned} v^{(n)}_{stat} &= G^{\dagger}\psi^{(n)}_{stat} \\ v^{(n)}_{inc} &= G^{\dagger}\psi^{(n)}_{inc} \end{aligned} \tag{C.7}$$

Finally, because we know that the velocity vector of static terms $v^{(n)}_{stat}$ will be identical whether we use only the independent set of interferograms or include redundant interferograms, we can summarize the redundant SBAS time series with the following:

$$d^{(red)}_T = \sum_{i=2}^{T} v_{stat_{(i-1)i}}(t_i - t_{i-1}) + \sum_{i=2}^{T} v_{inc_{(i-1)i}}(t_i - t_{i-1}) = d^{(ind)}_T + \sum_{i=2}^{T} v^{(red)}_{inc_{(i-1)i}}(t_i - t_{i-1}), \tag{C.8}$$

where $v^{(red)}_{inc_{(i-1)i}}$ indicates the additional incidental noise from including the redundant interferograms in SBAS.

If the independent noise in every interferogram $\psi_{inc_{ij}}$ is drawn from a zero-mean distribution, then the $d^{(red)}_T$ will approach

$$d^{(red)}_T = d^{(ind)}_T + \sum_{i=2}^{T} v^{(red)}_{inc_{(i-1)i}}(t_i - t_{i-1}) \Rightarrow \sum_{i=2}^{T} v_{stat_{(i-1)i}}(t_i - t_{i-1}), \tag{C.9}$$

as the number of redundant interferograms increases. If not, then the time series will be biased toward the mean of the incidental-noise distribution.

Appendix D

InSAR and Wells in Delaware Basin Study Area

There is little-to-no obvious spatial correlation between the InSAR surface displacements in our study area and the wells that were active during the time frame of our study. Supplemental Figure D.1 breaks down the well data into type: groundwater wells (acquired from the BRACS database and Texas Water Development Board websites), disposal wells, vertical production wells, and horizontal production wells. We acquired the data for the latter four categories (and their cumulative volumes) from the Enverus DrillingInfo database (Enverus, 1999). The upper left subplot displays all of the wells (blue = groundwater; magenta = production; red = disposal).

Deng et al. (2020) suggested that the subsidence in our study area is due to groundwater withdrawal. While that may be the case, in part, for the subsidence signal near A in the groundwater subplot, there are few wells that align with the main linear deformation feature of interest in our study. However, we note that the groundwater well database in Texas may not be complete. The disposal well plot shows little evidence of injection-related uplift, except for a small correlation near point B. The large uplift near signal C has the same preferred orientation as the main linear deformation feature, though any associated wells may be off the bounds of our study area. We do not explicitly address this feature in our study, except to note that our final model is unable to reproduce the uplift at the magnitude observed by the InSAR. The vertical production wells show little correlation with subsidence signals, except for, perhaps, the few wells near D. However, these wells have true vertical depths that exceed 6000 m; any observable subsidence from these depths is unlikely. Finally, the horizontal production wells have north-south and east-west orientations, in contrast to the preferred orientation of the deformation features, which strike northeast. It is possible that production contributes to the high magnitude of subsidence near E, though there are few other locations where subsidence and horizontal wells seem to be correlated. While a full analysis

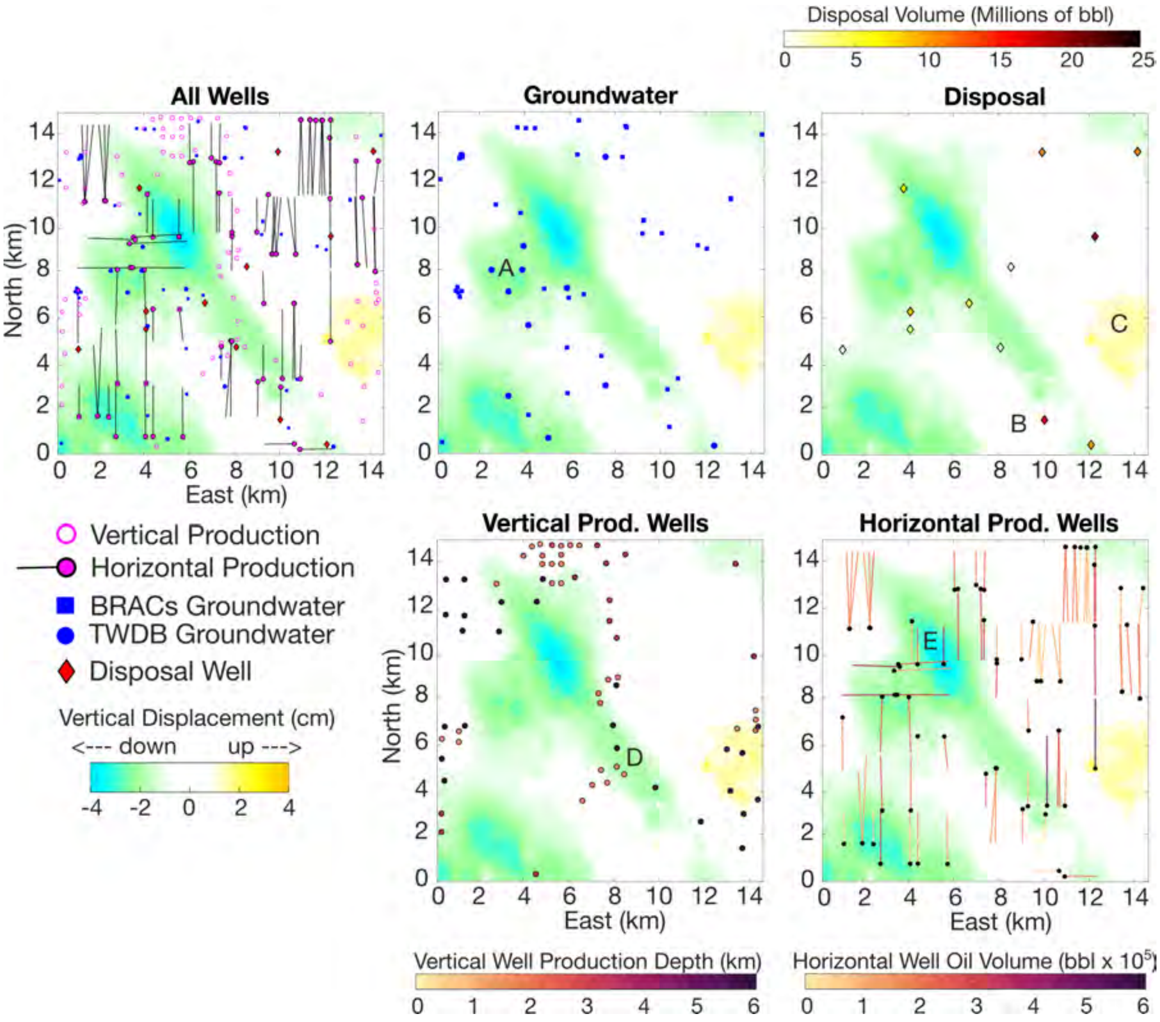


Figure D.1: Well data in relationship to the vertical InSAR data in our study area. We show groundwater, disposal, vertical production, and horizontal production wells. Although there are a few potential spatial correlations between wells and displacement features, none of them fully explain the linear deformation feature of interest.

of volume-change-related uplift and subsidence is required, our study area lacks spatial correlation with wells that would suggest the main linear deformation feature is directly related to poroelastic fluid flow. We believe the deformation in this region requires other geomechanical mechanisms as explanation, such as slip on normal faults.

Appendix E

Identification of Sites for Surface Spreading Recharge with Electromagnetic Resistivity Imaging

Characterizing the reservoir response to fluid extraction and injection is not the final step in hazard management. Rather, it marks the beginning of the next stage: the development of mitigation techniques. In many cases, mitigation falls on the shoulders of government agencies, which are significantly limited by financial restrictions and low man-power. Furthermore, time is of the essence. Therefore, the methods we develop must be efficient, cost-effective, and relatively easy to implement. Often, this requires a compromise between rigorous scientific analysis and practical approximations. This chapter demonstrates the utility of this compromise in groundwater management.

As we saw in Chapter 6, InSAR has demonstrated that California's Central Valley (CV) is sinking due to the over-pumping of groundwater resources, with recent rates reaching up to 20 cm yr^{-1}. Figure E.1 demonstrates the widespread extent of subsidence between 2015-2019, which has damaged infrastructure and led to a permanent loss in groundwater storage capacity (Smith et al., 2017). Having characterized the driving cause (compaction of clays in confined aquifers), the next stage is to manage groundwater levels to prevent further damage and loss of freshwater resources. Managed aquifer recharge (MAR) is the targeted restoration of groundwater levels to maintain a balance between pumping and injection. Various methods of MAR exist, as discussed in Chapter 2, including the direct injection of fluids into the aquifers, or through the surface spreading of reclaimed floodwater. Our focus in this chapter is to demonstrate how electromagnetic imaging can illuminate

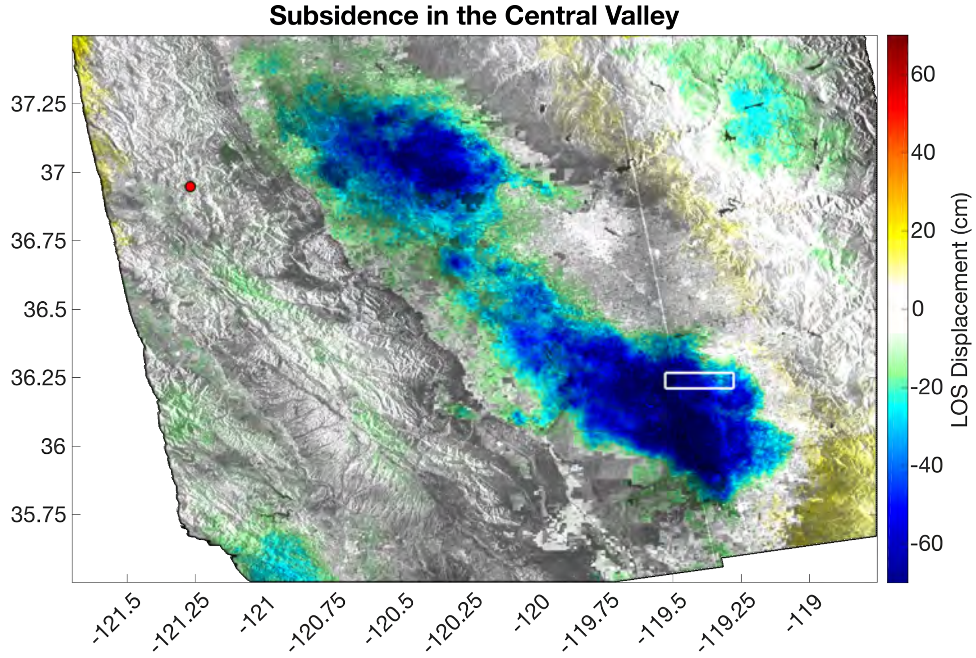


Figure E.1: Subsidence in the Central Valley as measured by Sentinel-1 InSAR between 01-31-2015 and 06-27-2019. The reference station used to calibrate the data to GPS is shown by the red dot and the white box highlights our study area, where we seek quality sites for surface-spreading recharge.

properties of the subsurface that can be related to sediment type. We can use this knowledge to estimate material parameters that control the preferred pathways through which fluids will flow: hydraulic conductivity. Our simple characterization of fluid flow through the shallow subsurface will allow groundwater managers to quickly identify potential sites for recharge while efficiently eliminating others. These finding have been published in Pepin et al. (2022b).

E.1 Groundwater Management in the Central Valley, CA

The sustainability of agricultural practices in California's Central Valley (CV), an industry valued at US\$17 billion yr^{-1} (USGS, 2016), depends on maintaining a balance between the amount of groundwater extracted for irrigation and the amount of water returned to the groundwater system through recharge. In the San Joaquin Valley, the southern part of the CV, extraction over decades has greatly exceeded recharge, resulting in dropping groundwater levels and subsidence, which is demonstrated in Figure E.1. With the passage of the Sustainable Groundwater Management Act, there is now considerable interest in finding ways to supplement natural recharge processes through managed aquifer recharge.

A detailed overview of various recharge methods can be found in Bouwer (2002). One of these methods, which we refer to as surface spreading recharge, involves the use of water to flood an available area on the ground surface, such as an agricultural field during non-growing seasons, open land that has been converted to a recharge basin, or a natural or restored floodplain. The objective is to have the applied water percolate through the subsurface to the water table, which marks the top of aquifer and intended storage location, thus replenishing the groundwater system and storing groundwater for future use in dry years. However, this method involves certain risks that are related to the conditions and hydrogeologic properties of the selected site:

1. Slow infiltration of the water could lead to ponding at the surface or lingering of water in the root zone, both of which may weaken the roots and bases of trees, leaving them susceptible to damage in strong winds (Dahlke et al., 2018; O'Geen et al., 2015; Ganot and Dahlke, 2021).
2. Slow infiltration and subsequent ponding can lead to increased evaporative loss.
3. If the spatial distribution of the sediment types favors lateral over vertical movement, the water may be unable to reach the intended storage location. If the goal is to store groundwater for future use or reduce subsidence, it is important that the water reaches the intended location. Furthermore, barriers inhibiting downward flow can result in perched water tables or water mounds that will slow infiltration rates (Bouwer, 2002).
4. When contaminants are a concern, either as result of agricultural practices or the chemistry of the recharge water, recharge can have a negative effect on water quality (Bouwer, 2002).

The first risk applies to agricultural fields, whereas the latter three apply to any site under consideration for surface spreading recharge. Given the above, one of the key questions that needs to be addressed before starting the design and operation of a surface spreading project is where to recharge to reduce these risks. In this study, our focus was on the identification and assessment of recharge flow paths, with the objective being to maximize the efficiency with which water can move from the surface to the water table. Related to this, Bouwer (2002) described two major criteria for the hydrogeologic properties of the vadose zone that should be satisfied: (a) fast infiltration in the shallow subsurface and (b) downward, uninhibited flow connections to the intended storage location. In this study, we developed a methodology to aid in the selection of a recharge area — an entire site, or part of a site — that satisfies these criteria. The effect of surface spreading on water quality will also influence the selection of a recharge area. While our methodology can provide information relevant for assessing potential changes in water quality, far more information about subsurface conditions and properties would be required. Adequately addressing water quality issues is thus beyond the scope of the developed methodology.

O'Geen et al. (2015) developed a map of the recharge suitability of agricultural regions of the CV using the Soil Agricultural Groundwater Banking Index (SAGBI), which is a scoring system used to rank regional areas for surface spreading recharge by assessing the hydrogeologic properties of the soils along with other surface features, such as topography and the chemicals used in farming. Two of the five factors used in the scoring, *root-zone residence time* and *deep percolation*, were designed to predict where water would quickly infiltrate the surface soils and move past the root zone to deeper sediments, respectively. These two factors are determined using estimates of the hydraulic conductivity (K) of the soil horizons specific to the region, where high K is better for recharge. These soil horizons and assigned K estimates are obtained from the USDA–NRCS digital soil survey (USDA–NRCS, 2014) and, in general, are only valid estimates up to the first 2 m below the ground surface.

While SAGBI has great utility for assessing the first 1 or 2 m, root zones may extend deeper (up to 4 m) and water tables in the CV are often tens of meters below the surface, so knowledge of the sediment types at greater depths within the vadose zone is important for predicting where there are uninhibited flow connections in the shallow subsurface and to the aquifer. An additional limitation in the use of SAGBI to assess specific sites is the fact that it was developed from soil surveys consisting of map units ranging in size from 0.02 to 2 km^2 (5–500 acres), and while many units contain multiple soil profiles, SAGBI used only the most abundant type in each map unit. Therefore, the K estimates may not best represent the local conditions within any individual site, which are often on the order of <0.5 km^2 (tens of acres). In fact, the USDA–NRCS explicitly recommends onsite investigations when planning the use of a small, local area (USDA–NRCS, no date).

One approach to assessing the suitability of a site for surface spreading recharge would be to conduct the modeling required to simulate flow in the vadose zone. Such modeling can be complex and time-consuming and requires the accurate estimate of parameters — infiltration rates and hydraulic conductivity — that are challenging to obtain, especially in the absence of nearby wells. However, after the initial saturation, flow patterns become similar to those observed in the saturated zone, where water travels in preferential flow paths that are dominated by networks of sediments with high K, while occasionally “jumping” through smaller regions of low K (e.g., Bianchi et al., 2011; Rizzo and de Barros, 2017). These findings suggest that, upon initial wetting of the vadose zone, regions dominated by coarse-grained sediments (sand and gravel) that have high K are the conduits for flow, whereas areas dominated by fine-grained sediments (silt and clay) that have low K inhibit flow. If it is possible to resolve the spatial distribution of sediment type in terms of the amount of coarse-grained material, then it may be feasible to predict where there will be predominantly vertical preferential flow paths both in the first few meters for fast infiltration and connecting the ground surface to the intended storage location, using sediment type as a proxy for K. What this requires is a model of sediment type with sufficiently high spatial coverage and resolution to evaluate the suitability of sites for surface spreading recharge.

The CV hydrologic model (CVHM) (Faunt, 2009) provides information about the unconsolidated sediment type in the CV. The CVHM includes an estimated fraction of coarse-grained material, which we refer to as coarse fraction, throughout the CV, extending to a maximum depth of ∼800 m with cells of lateral dimension 1,600 by 1,600 m and vertical dimension of 15 m. Because of its widespread spatial coverage, Jankowski et al. (2018) used the CVHM to augment SAGBI with a factor for scoring deep sediment type, where deep refers to the zone between the soil (described with SAGBI) and the water table. At each CVHM cell, they calculated the harmonic mean of coarse fraction in this zone, with higher fractions receiving better scores in the assessment. With a recharge site typically smaller than a single cell in the CVHM, all that would be obtained from CVHM is an average value of coarse fraction in 15-m-depth intervals. In order to identify recharge pathways from the surface to the water table, we need to resolve a finer-scale spatial distribution of coarse-grained material, that is, on the scale of a few meters, rather than kilometers, per cell.

Over the past 10 yr, there has been growing interest in the use of geophysical methods as a cost-efficient way of characterizing sediment type to support managed aquifer recharge operations, including surface spreading recharge. Studies have used a variety of geophysical methods including surface and borehole electrical resistivity tomography, time domain electromagnetics, and nuclear magnetic resonance (Cook et al., 1992; Gottschalk et al., 2017; Maliva et al., 2015; Mawer et al., 2013; Parsekian et al., 2014; Sendrós et al., 2020; Lawrie et al., 2013). Two of the most significant impediments with the use of many of these methods are the time it takes to acquire data and the challenge of efficiently and accurately deriving the information needed about the subsurface to assess the suitability of a site for recharge.

In our research, we are using a new geophysical system, towed time-domain electromagnetics (tTEM), which is a time-domain electromagnetic system (Auken et al., 2019) towed behind an all-terrain vehicle that has been found to be a highly effective approach for subsurface characterization at the scale of an agricultural field. The tTEM system contains a pair of rectangular transmitter and receiver loops of size 2 by 4 m and 1 by 1 m, respectively, spaced with a 9-m offset. Time-varying currents are injected into the transmitter loop; this induces eddy currents in the subsurface with the resulting signals measured at the receiver loop. Through inversion, a model of the electrical resistivity of the subsurface is recovered from the measured signals.

The tTEM system was used at seven sites (four recharge basins, two nut tree orchards, and one open field) in the Tulare Irrigation District (TID) in the San Joaquin Valley of California (Behroozmand et al., 2019). This area is outlined by the white box in Figure E.1. The relatively small size of the tTEM system meant that data were easily acquired by driving the all-terrain vehicle between rows of trees in almond [*Prunus dulcis* (Mill.) D. A. Webb] and pistachio (*Pistacia vera L.*) orchards. Recovered through inversion of the acquired data were 1D models of the electrical resistivity of the subsurface. Horizontal spacing of the 1D resistivity models was ∼10 m along each acquisition line, with lines separated by 10–20 m depending on the site. Although the 1D vertical resistivity models are continuous in depth, the noncontinuous horizontal spacing between tTEM measurements results in sparsely filled, 3D resistivity models when all are combined to represent the full site. Vertical discretization of the resistivity models begins at ∼1 m at the surface and gradually increases to 6–12 m by the depth of investigation, which was between 50 and 70 m; the depth of investigation is considered to be the depth to which the estimates of electrical resistivity are reliable and depends on the subsurface electrical resistivity. Using measurements of sediment behavior, acquired using cone-penetrometer testing, the resistivity data in the vadose zone from one of the sites were transformed to the fraction of coarse-grain-dominated material, referred to as coarse-dominated material (i.e., material consisting mainly of sands and gravels), by empirically determining a resistivity-to-sediment-type transform (Goebel & Knight, 2021).

In the study by Goebel and Knight (2021), the sparse 3D model displaying the spatial variation in the fraction of coarse-dominated material was qualitatively interpreted to determine where there appeared to be preferential flow paths from the surface to the water table at a site. An example is shown in Figure E.2 (modified from Goebel and Knight, 2021), where the authors identified a potential connected pathway of coarse-dominated material through the subsurface. This qualitative interpretation required manually scrolling through the data to identify these pathways and relied on visual interpolation of sediment type in vertical sections lacking tTEM data. While valuable for illustrative purposes, this approach did not provide a way to easily locate or quantify the length of these pathways nor did it consider the uncertainty associated with identifying continuous pathways in a sparsely filled model.

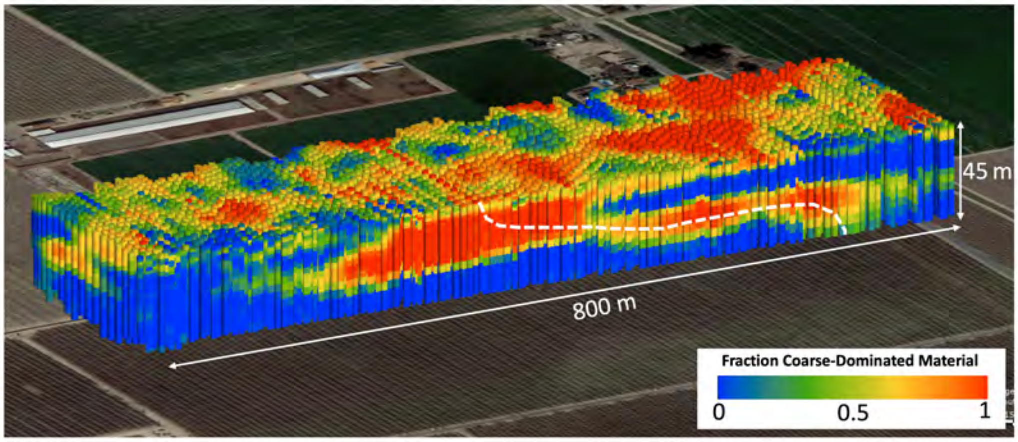


Figure E.2: Model of the fraction of coarse-dominated material at a site in the Tulare Irrigation District. The preferential flow path depicted by the dashed white line was manually selected using the spatial distribution of units with a high fraction of coarse-dominated material. Figure modified from Goebel and Knight (2021).

In this study, we advanced the use of tTEM data for assessing potential recharge sites by developing and demonstrating a more quantitative methodology, most of which can be automated and efficiently used, for identifying preferential flow paths that connect the surface and the water table and quantifying their length. We assessed the quality of the flow paths by considering first, their ability to rapidly transmit water beyond the root zone and, second, their distance to reach the water table. In this context, shorter preferential flow paths with no shallow low-K barriers were deemed to be of higher quality, as these will transmit water most efficiently to the intended storage location. This method allowed us to identify areas within a site on which recharge efforts should be concentrated and evaluate the relative suitability of potential sites for surface spreading recharge. While a binary system (high-K and low-K) is a simplified representation of the subsurface, it is a useful model for rapidly identifying preferred sites based on our metrics. These sites can then be selected for further tests (e.g., infiltration rates) rather than performing these time-consuming and often expensive assessments at all available sites.

Our methodology of finding high-quality preferential flow paths consisted of two parts. We began with resistivity models derived from the inversion of tTEM data and used a geostatistical simulation to transform them into many 3D models of sediment type, classified as either high-K or low-K, which could represent the subsurface (described in Methods I section). By developing numerous models with varied input parameters for each simulation run we were able to capture the uncertainty resulting from the assumptions we needed to make in their development.

We then assessed the quality of preferential flow paths in the produced models by locating those that had no shallow barriers and provided the shortest connections of high-K units between the surface and the water table (described in Methods II section). This allowed us to evaluate the preferential flow paths at each site and compare the different sites to each other.

E.2 Data and Study Area

The tTEM data used in this study were acquired in October 2017 by Behroozmand et al. (2019) at seven sites in the TID (Figure E.3). Four of the sites — Swall-North, Swall-South, Martin-West, and Martin-East — have been used as active recharge basins. Another cluster of two sites consists of an almond grove, Almond-West, and a smaller field across the street, Almond-East, which is under consideration for surface spreading recharge. The final site is a large pistachio orchard given the designation of Nichols. The land surfaces at all of the sites are close to level in terms of elevation.

We defined the depth to the water table at each site using the depth-to-groundwater map created by the TID, 2017. At sites Almond-West and Almond-East, the depth obtained from this map agreed well with a detailed analysis conducted by Goebel and Knight (2021) to identify the water table depth using the tTEM data.

A regional assessment of our study area indicates that the top 30 m of the subsurface consists predominantly of young, continental alluvium distributed in interbedded deposits of clay, silt, silty sand, and gravelly sand (MKGA, 2019). Below these deposits is a transition zone to less-permeable, clay-rich sediments that would greatly inhibit the flow of recharge water. Thus, there is significant spatial heterogeneity in sediments that needs to be assessed in order to characterize the recharge potential of sites in this area.

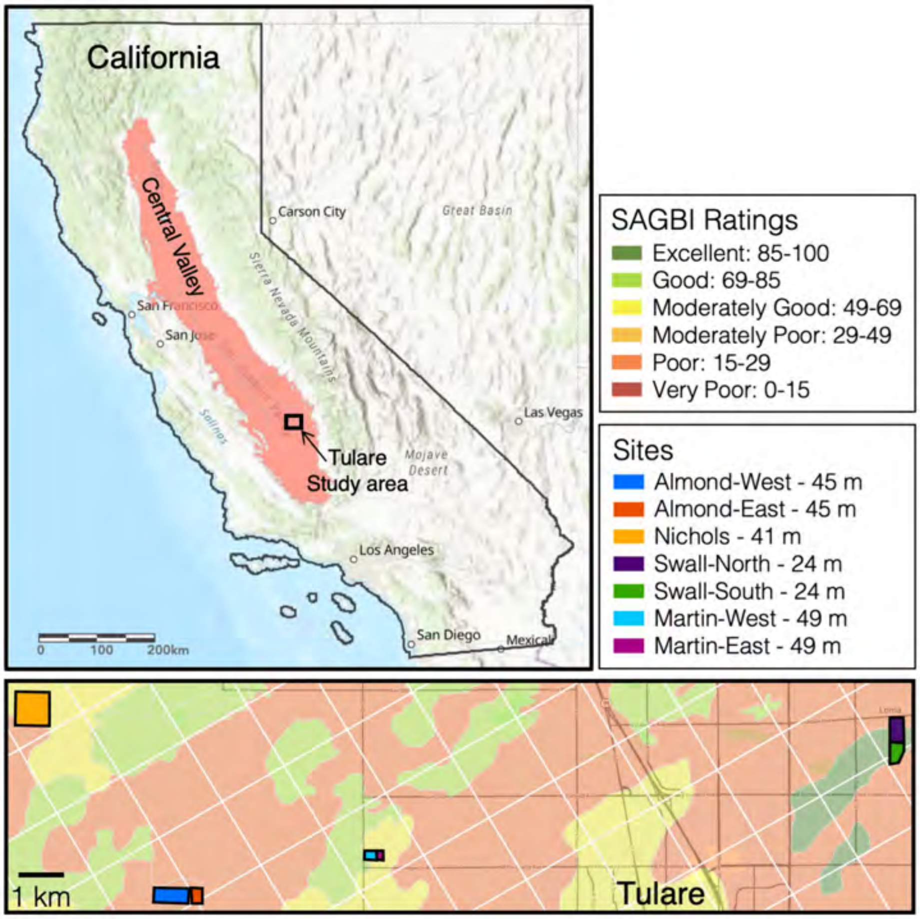


Figure E.3: California's Central Valley and study sites in the Tulare Irrigation District (TID). The Central Valley (red in top panel) contains our study area in Tulare. The bottom panel shows the locations of sites and the nearby Soil Agricultural Groundwater Banking Index (SAGBI) categorical designation (O'Geen et al., 2015). Values in the sites legend indicate the local depth to the water table, obtained from the TID depth-to-groundwater map (TID, 2017). The white grid lines show the map-view footprint of the Central Valley Hydrogeologic Model's (CVHM) texture model (Faunt, 2009). The sites are consistently smaller than the CVHM cell size.

E.3 Methods

The methods we developed can be split into two parts: (a) Methods I, the development of multiple models of subsurface sediment type and (b) Methods II, the assessment of subsurface models for surface spreading. The full Methods I workflow is visually depicted in Figure E.4, which can be referenced throughout section E.3.1.

E.3.1 Methods I: Development of Subsurface Models

Recovering Uniformly-Gridded 3D Resistivity Models from tTEM Data

We began this study with the resistivity models from the study by Behroozmand et al. (2019). The resistivity model at each site, recovered through inversion of the tTEM data, is non-unique, meaning that many resistivity models could fit the data equally well. We accounted for the uncertainty in the inversion by generating 100 resistivity models at each site using a posterior sampling (PS) technique (Kang et al., 2021; Rue, 2001), which used the recovered resistivity model as an input. This technique uses a linear approximation and assumes a Gaussian posterior distribution. We refer to these 100 models as the PS resistivity models.

Next, we resampled these PS resistivity models onto uniform 5- × 5- × 1-m 3D grids, leaving a 25-m buffer zone extending beyond the bounds of the resistivity models, which were contained within the edges of the site. This resulted in 100 sparse uniform grids of resistivity — one for each of the 100 PS resistivity models. We assigned resistivity values only to those voxels for which there was a resistivity measurement within 5 m during the resampling, leaving all other voxels, including the buffer zone, empty.

Transforming Resistivity to Probabilities of Sediment Type

The next step for each site was to transform the uniformly gridded resistivity models to subsurface models in which each voxel containing a resistivity value was assigned a probability of being high-K material ($\mathrm{P}_{\text{high-}K}$). We began with the transform developed by Goebel and Knight (2021), which defines distributions of the fraction of coarse-dominated material ($\mathrm{FRAC}_{\mathrm{CD}}$) that correspond to integer resistivity values as shown in Figure E.5a. While this relationship was developed using sediment samples at the Almond-West site, we assumed that it was applicable to the other six sites where the sediments are similar in terms of composition and depositional environment. If a highly accurate subsurface model had been required, the procedures followed by Goebel and Knight (2021), using cone-penetrometer testing, could have been repeated at a site so as to develop a site-specific transform. In order to define $\mathrm{P}_{\text{high-}K}$, we needed to select a cutoff between values of $\mathrm{FRAC}_{\mathrm{CD}}$ that are low-K vs. high-K. Setting an appropriate cutoff should take advantage of any information about the properties of the site and will be dependent on the objective of the investigation and the risk being managed. The number of voxels identified as high-K in the workflow would increase as the

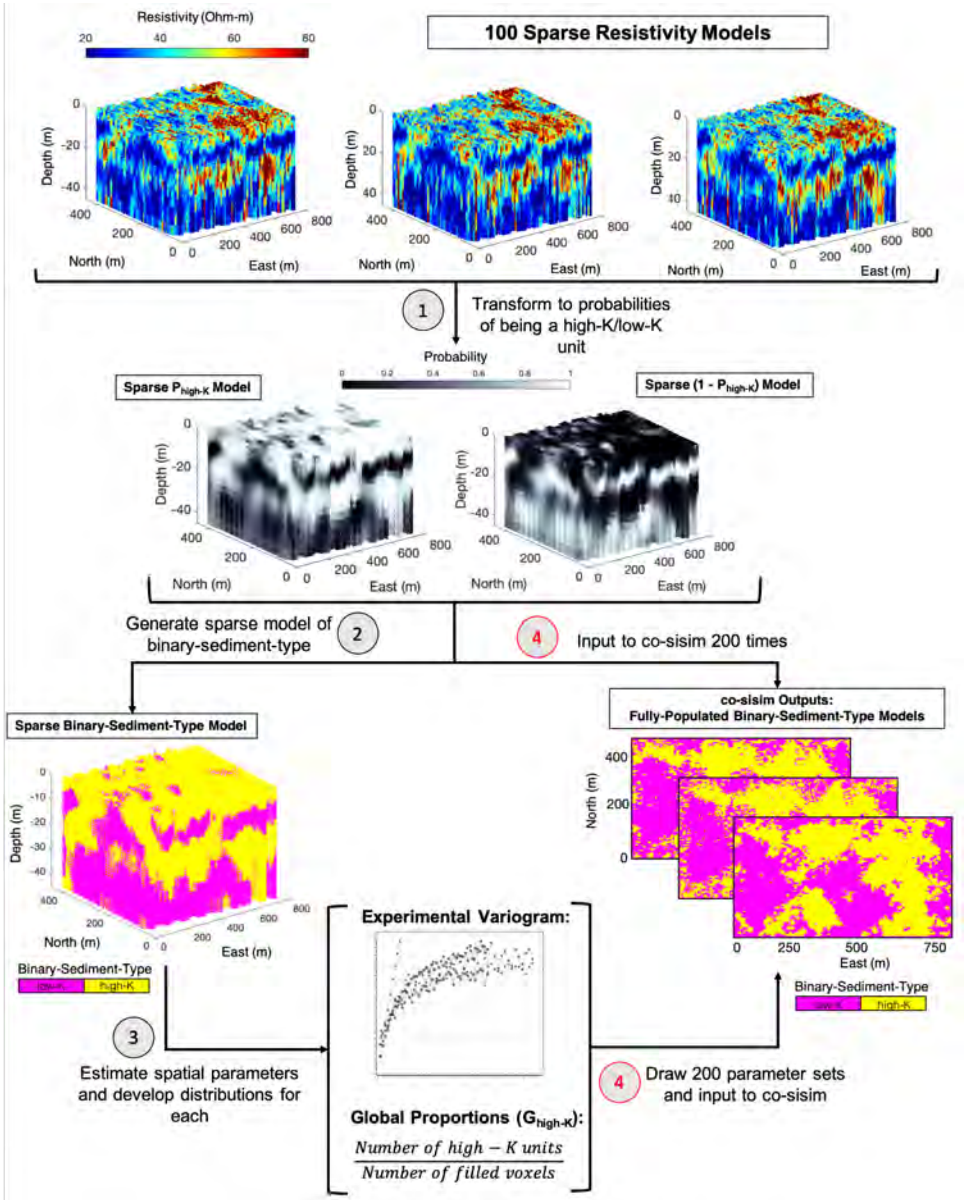


Figure E.4: Visual workflow for Methods I. Steps 1-2 create sparsely-filled models representing K, including (1) mean $\mathrm{P}_{\text{high-}K}$ and (1-$\mathrm{P}_{\text{high-}K}$) and (2) a model of sparse binary-sediment-type. Step 3 uses the sparse binary-sediment-type model to estimate the spatial parameters and their distributions for use in co-sisim, including the global proportion of high-K units ($\mathrm{G}_{\text{high-}K}$) and the variogram model parameters (maxHR, minHR, VR, and angle max HR). The final step (4), denoted by the two red circles, runs co-sisim 200 times, using the $\mathrm{P}_{\text{high-}K}$ and (1-$\mathrm{P}_{\text{high-}K}$) models and randomly-selected spatial parameter sets (drawn from their respective distributions), to create 200 fully-populated realizations of binary-sediment-type, each with varying spatial correlation.

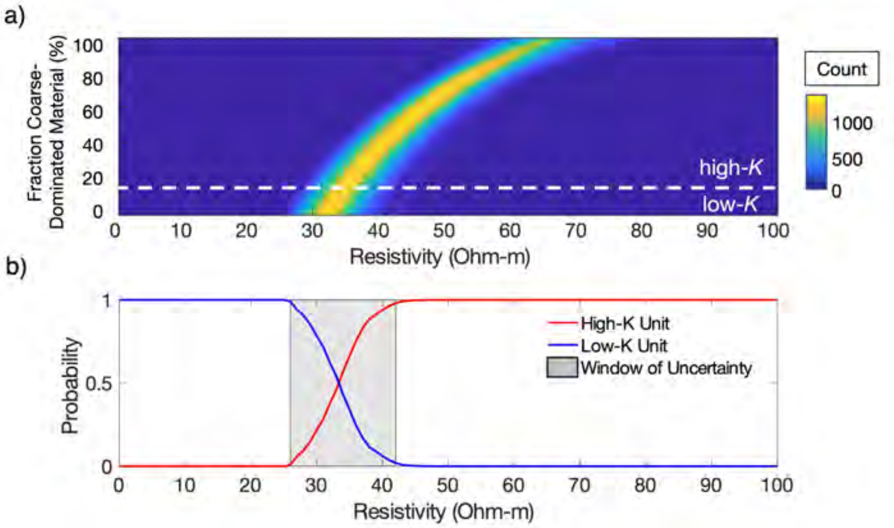


Figure E.5: Transform from resistivity to categorical probabilities. We started with (a) the transform from resistivity to fraction coarse-grain-dominated material ($\mathrm{FRAC_{CD}}$) as obtained in Goebel and Knight (2021). From these distributions, we derived (b) the transform to categorical probabilities by comparing the number of counts at or above the high-K threshold (white dashed line in (a)) with the total counts at each resistivity bin. Resistivity values within the window of uncertainty in (b) may represent either a high-K or a low-K unit. We applied the transform in (b) to each resistivity value.

$\mathrm{FRAC_{CD}}$ value selected for the cutoff is decreased. If it is important not to overestimate the quality of preferential flow paths, so as to reduce the risk of surface ponding, then higher values of $\mathrm{FRAC_{CD}}$ should be used. Alternatively, to reduce the risk of overlooking a suitable site by underestimating the quality of preferential flow paths, lower values of $\mathrm{FRAC_{CD}}$ should be used.

Lacking information about the way in which K varied with $\mathrm{FRAC_{CD}}$ at any of the sites, we referred to the work of Fogg et al. (2000), which showed that a volume of sediments in the CV requires a minimum of 18% coarse-grained material to have a high-K pathway for flow. We thus set $\mathrm{FRAC_{CD}} = 0.18$ as the cutoff in our study area noting that this value may be readily adjusted to fit the needs of the study.

After selecting the cutoff between high-K and low-K, we transformed all of the resistivity values to $\mathrm{P_{high\text{-}\mathit{K}}}$ using the probability vs. resistivity relationships shown in Figure E.5b; these were extracted from the $\mathrm{FRAC_{CD}}$ distributions in Figure E.5a by relating the number of counts above the $\mathrm{FRAC_{CD}}$ cutoff to the total number of counts in each resistivity bin. This step resulted in 100 sparse $\mathrm{P_{high\text{-}\mathit{K}}}$ models, which we reduced to a single mean $\mathrm{P_{high\text{-}\mathit{K}}}$ model by calculating the arithmetic mean value at each populated cell.

Generating Fully-Populated Models of Binary Sediment Type

The final stage in our workflow was to generate fully populated models of binary-sediment-type at each site, which could be used to find the preferential flow paths. We sought to generate many binary-sediment-type models to capture the uncertainty resulting from varying the assumptions about spatial correlation of sediment type. To accomplish this, we used a geostatistical module known as the co-sequential indicator simulation (co-sisim) in Stanford's Geostatistical Modeling Software (SGeMS) (Remy et al., 2009), which outputs fully populated categorical models (e.g., of binary-sediment-type) using sparse input data and various spatial parameters that describe the expected spatial correlation and proportion of the output categories. In our workflow, we used the sparse mean $\mathrm{P}_{\text{high-}K}$ model as the sparse input data; for the spatial parameters, we defined (a) the expected total (global) proportion of high-K units in the fully populated model ($\mathrm{G}_{\text{high-}K}$) and (b) a variogram model, which designates the spatial correlation (range) of sediment type. We characterized the variogram model by its mutually orthogonal maximum horizontal, minimum horizontal, and vertical ranges, and the angle between the axis of the maximum horizontal range and north. The parameters we used to define the model grid for each site ($\mathrm{G}_{\text{high-}K}$, the experimental variogram model, and depth to water table) are listed in Table E.1 in the chapter appendix (section E.8).

To generate multiple models, we ran co-sisim 200 separate times. During each run, we used the same sparse input data but varied the values of the spatial parameters, as we were able to estimate them only roughly from the available data. After first estimating the five spatial parameters, we developed uniform distributions for each. These distributions allowed us to randomly select values for these parameters during each simulation run, resulting in 200 different models of binary-sediment-type. These steps, founded in Monte Carlo principles, follow the same basic structure as other geostatistical modeling studies that attempt to account for the uncertainty in simulation parameters (e.g., Soutter and Musy, 1998; Christensen, 2004; Caers and Hoffman, 2006; Wang et al., 2020). Appendix E.8 at the end of this chapter describes in detail the steps we used to obtain the initial estimates of $\mathrm{G}_{\text{high-}K}$ and the variogram model, develop the distributions for each spatial parameter, and randomly generate a parameter set during each run of co-sisim.

E.4 Methods II: Assessment of Subsurface Models for Surface Spreading Recharge

The binary-sediment-type models at each of our seven sites provided the data needed to evaluate preferential flow paths and identify optimal areas for surface spreading recharge within each site. At each site, we evaluated each of the 200 models using two metrics, described below, and selected the optimal rectangular recharge areas based on the arithmetic mean of the outputs. We chose rectangular areas for this study for practical reasons; it is easier to select and analyze rectangular areas on gridded data, and if only a portion of a site were selected for recharge, barriers could be positioned to follow the row crops, which are gridded as well.

Defining Metrics and Parameters for the Selection of Recharge Areas

In order to avoid ponding and root damage, water must infiltrate past the root zone as quickly as possible. While we cannot accurately quantify infiltration rates with our models, we can identify regions of the site that have high-K units beyond the root zone, thus serving as indicative of fast infiltration rates because the presence of low-K units slows infiltration. Consequently, we developed Metric 1 to quantify the depth to the shallowest low-K unit beneath each surface pixel, here referred to as $\mathrm{DEPTH}_{\text{low-}K}$ and demonstrated in Figure E.6a. Values for $\mathrm{DEPTH}_{\text{low-}K}$ will range from 0 m to the depth of the water table, where a larger value is better for recharge.

In addition to Metric 1, we needed a metric that identified flow paths with short connections to the water table. We developed Metric 2 to quantify the lengths of preferential flow paths, which we defined as the routes through high-K units that minimized the distance between points at the surface and the water table. A 2D schematic showing examples of preferential flow paths (which will extend into 3D) is given in Figure E.6b. An ideal recharge area will have short preferential flow

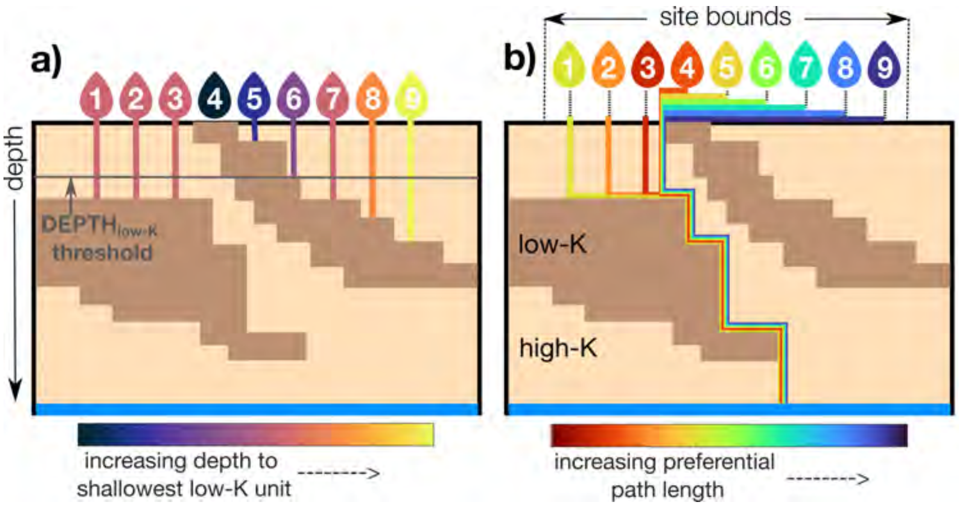


Figure E.6: Conceptual illustration of metrics used to assess the models of fully populated binary-sediment-type. (a) Metric 1, the depth to the shallowest low-K unit beneath each pixel ($\mathrm{DEPTH}_{\text{low-}K}$). We consider only pixels with $\mathrm{DEPTH}_{\text{low-}K}$ exceeding the selected threshold (gray horizontal line) as suitable for recharge. (b) The normalized path lengths (NPL) from each pixel to the water table. For both metrics, we only analyzed surface pixels within the bounds of the site noting that preferential paths may extend beyond these bounds once they infiltrate past the surface. Though pixels eight and nine have the deepest $\mathrm{DEPTH}_{\text{low-}K}$, there are no connections to the water table, thus they have the longest NPL. In our model, pixels are 5 by 5 by 1 m (illustration not to scale).

paths where the minimum length would be equal to the depth to the water table. It is important to note that these paths are not necessarily the routes through which water would flow, rather, we considered the path length to be a proxy for the time to reach the water table (i.e., shorter is faster).

Quantifying both metrics allowed us to identify optimal areas for recharge within an individual site. We first used Metric 1 to locate areas with deep $\text{DEPTH}_{\text{low-}K}$, then used Metric 2 to additionally prioritize regions with short preferential flow paths. We chose thresholds that defined suitably "deep" and "short," and then selected a rectangular optimal area that satisfied both thresholds. The $\text{DEPTH}_{\text{low-}K}$ threshold should be chosen by considering the depth of the root zone for the crops at the sites and should err on the side of being greater than deemed necessary if there is a high level of concern about ponding and root damage. We elected to use $\text{DEPTH}_{\text{low-}K}$ threshold as 8 m, which is twice the maximum depth of the root zone for the almond and pistachio trees in our study area. In practice, the choice of the $\text{DEPTH}_{\text{low-}K}$ threshold will also depend on the volume of recharge water, the surface area over which the water is applied, and the infiltration rates.

To select a threshold for Metric 2, we first defined a normalized path length (NPL) as a preferential flow path length divided by the local depth to water table. The NPL is thus a measure of the tortuosity of the path where a value of 1 is the minimum and deemed "best". In this study, we selected a NPL threshold of 2 as the cutoff for short paths. A higher NPL corresponds to predominantly lateral flow, thus increased residence time, which may result in perched water tables that ultimately slow infiltration rates (Bouwer, 2002). We note that the tolerance for lateral flow will depend on each individual case.

Calculation of Metric Maps

For each model of binary-sediment-type at a site, we calculated both metrics at every 5- × 5-m surface pixel within the lateral bound of the site (25 m inward from the edges of the 3D grid), displaying the result in plan-view maps. This resulted in 200 maps per metric. We obtained a single representative map for each metric by calculating the arithmetic mean value at each pixel across the 200 maps.

Metric 1 involved a simple calculation of the depth to the shallowest low-K unit in each vertical column. For Metric 2, we calculated the length of the preferential path lengths using the Python package scikit-fmm (Furtney, 2015), referred to here as the "Eikonal Solver," which uses the fast-marching method (Sethian, 1996) to solve the eikonal equation, an approximation of the wave equation. This equation is a nonlinear partial differential that can be used to find the shortest distance (or travel time) between points and surfaces. When calculating distance, as done in our workflow, the simplifying assumption made was that there were only flow- and no-flow units. In our models of binary-sediment-type, high-K units were defined as flow units and low-K units were defined as no-flow units.

For each binary-sediment-type model, we configured the Eikonal Solver to find the shortest distance between each pixel at the surface to the water table. Once a path went below the surface, it could extend beyond the bounds of the site and into the 25-m buffer zone. Additional constraints included no-flow boundary conditions on the vertical sides of the 3D grid and an added layer of flow units at the surface to mimic the ability of water to move laterally across the surface of the site. We normalized the resulting maps of preferential flow path lengths to display the NPL at each pixel.

We used the representative metric maps to identify pixels with mean $\text{DEPTH}_{\text{low-}K} \geq 8$ m and mean NPL ≤ 2. We refer to the contours that contain these pixels as the 8-m-$\text{DEPTH}_{\text{low-}K}$ contour and 2-NPL contour, respectively, and overlaid both on the site's representative maps. At each site, we used these maps to manually select an optimal rectangular area for recharge. These optimal areas could range in size from a small rectangular area to the entire site. We sought to select optimal areas within which all metric values would have mean $\text{DEPTH}_{\text{low-}K} \geq 8$ m and mean NPL ≤ 2. In practice, some optimal areas included values outside these thresholds because of the challenge of selecting a rectangular area with irregularly distributed metrics.

E.5 Results and Discussion

E.5.1 Comparison of Areas Within a Site

Identifying an optimal area for recharge can involve assessment of a single site to select an area within the site. This might be done, for example, if there was low tolerance for the risk of ponding (e.g., an agricultural field). A selected area could be separated with flood barricades from the rest of the site.

We demonstrate the use of the two metrics for selection of the optimal area for surface spreading within the Almond-West site. Figure E.7 shows the (a) mean $\text{DEPTH}_{\text{low-}K}$ and (b) mean NPL maps calculated from the 200 model results and our manually selected optimal recharge areas. It is clear that Almond-West displays high variability in $\text{DEPTH}_{\text{low-}K}$ and NPL, suggesting significant spatial heterogeneity and underscoring the importance of resolving sediment type at high resolution. The CVHM texture model, for example, can only show that the entire almond grove has a harmonic mean value of percentage coarse-grained material of 58% between the surface to a depth of 45 m. The representation we get from tTEM is a much more informative view of sediment variation and how that may affect recharge efforts.

In selecting the optimal area for surface spreading at Almond-West, we first consider Metric 1. The mean $\text{DEPTH}_{\text{low-}K}$ map indicates that only ∼50% of the site may have sufficiently deep low-K units for fast infiltration during recharge, with areas outside of the 8-m-$\text{DEPTH}_{\text{low-}K}$ contour considered to be at risk of slow infiltration and subsequent ponding. When we then consider Metric 2, the southern half of the field has no NPL ≤ 2, despite having sufficiently deep $\text{DEPTH}_{\text{low-}K}$. Only the northern half of the site has any areas with NPL ≤ 2. An additional consideration is the small size of the two eastern-most areas. This suggests that a high number of the preferential flow paths

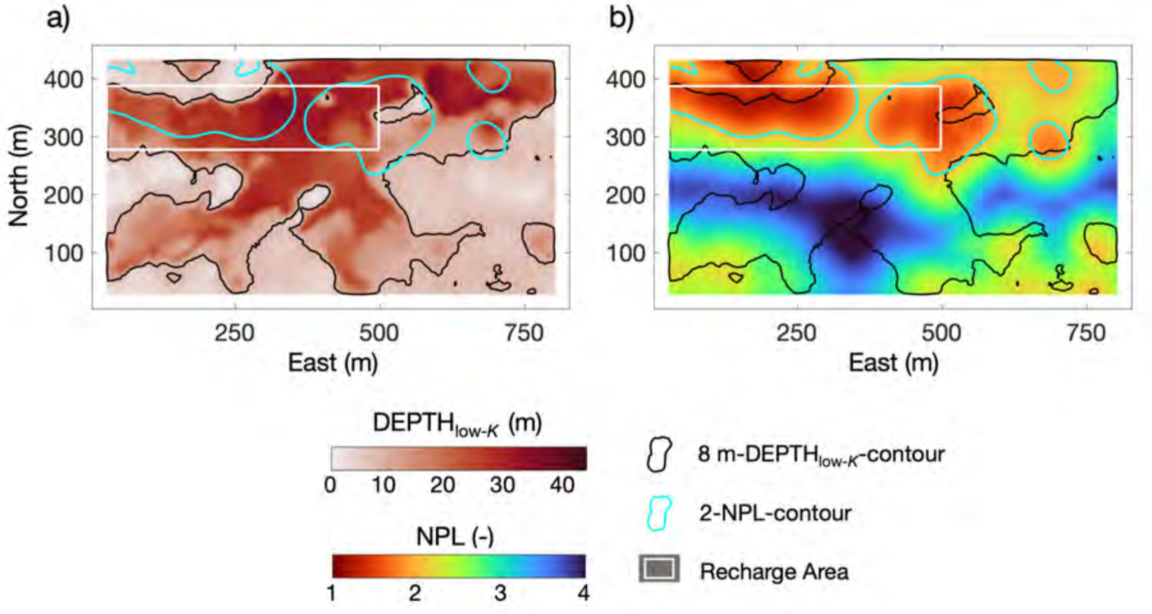


Figure E.7: Metric maps for Almond-West site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models.

merge along these routes, which might affect infiltration rates. We used the overlap between the two metrics to delineate the optimal area shown in Figure E.7, which largely satisfies both recharge criteria.

Similar figures for each of the other sites can be found in Figures E.8–E.13. When there were no, or extremely small, overlapping contours, as is the case for Martin-West and Martin East (Figures E.12 and E.13), we prioritized DEPTH$_{\text{low-}K}$ while attempting to include the local minimum of NPL, noting that these sites would not be ideal for recharge under our criteria.

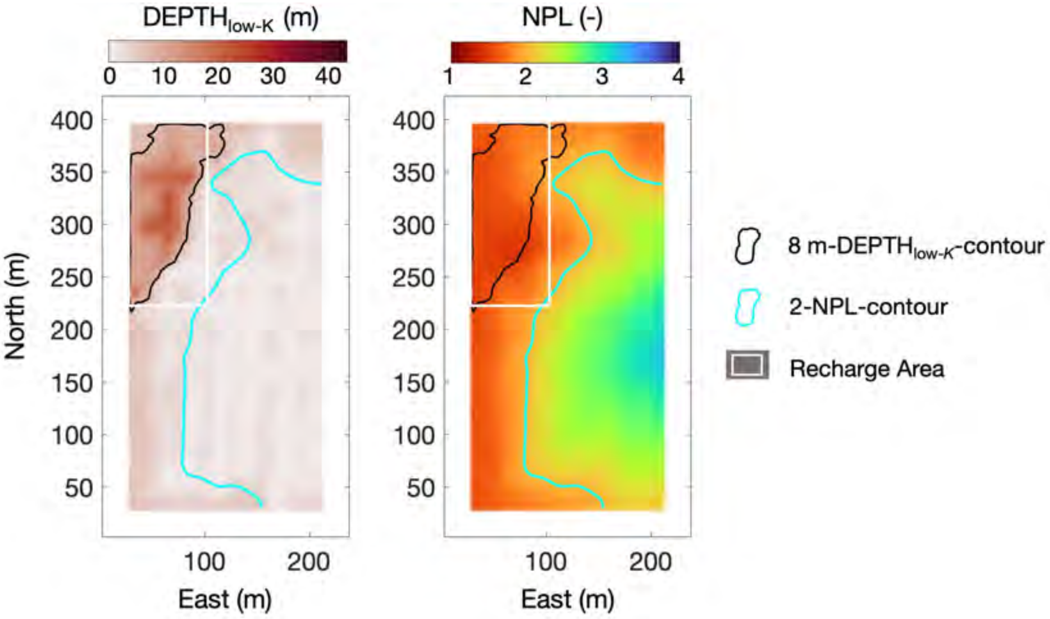


Figure E.8: Metric maps for Almond-East site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models.

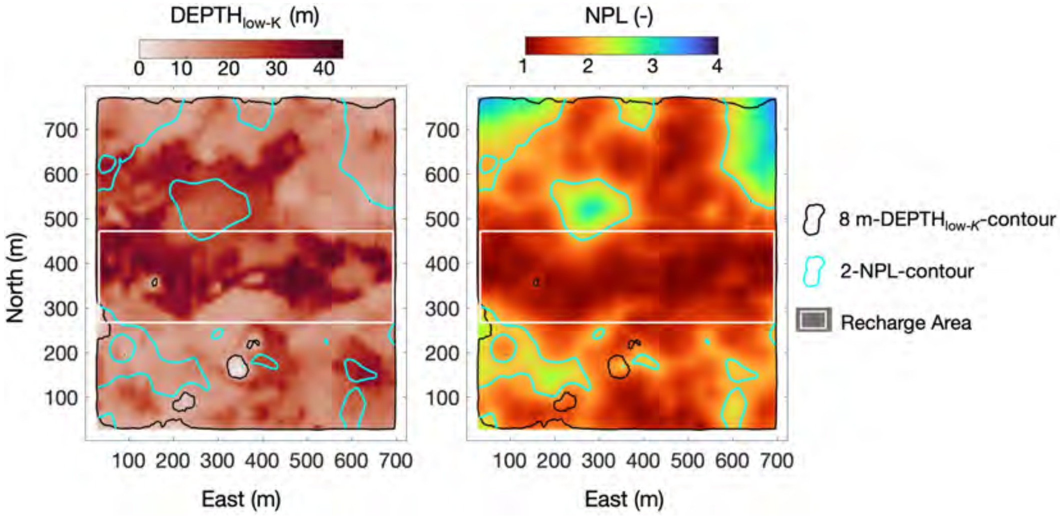


Figure E.9: Metric maps for Nichols site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models.

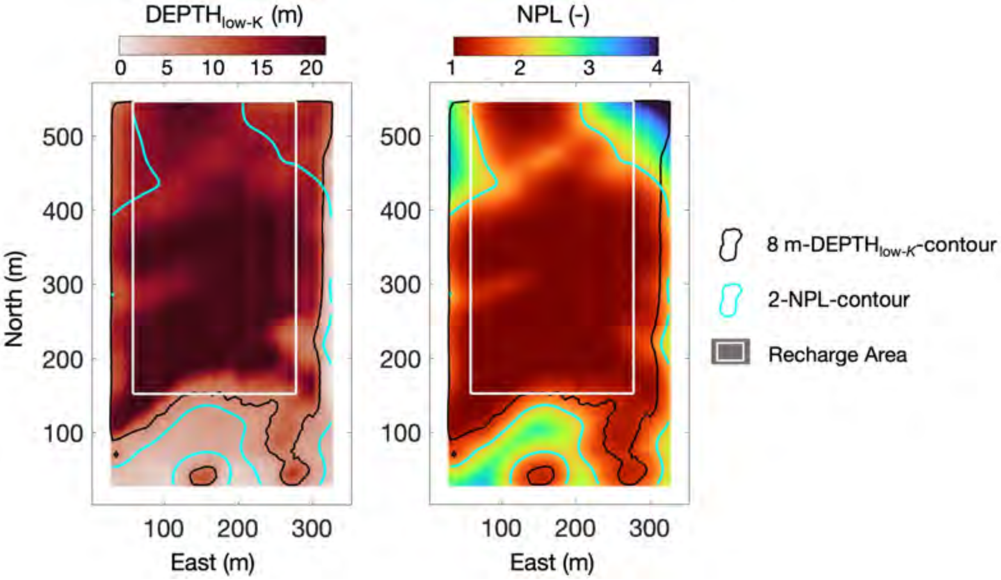


Figure E.10: Metric maps for Swall-North site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models.

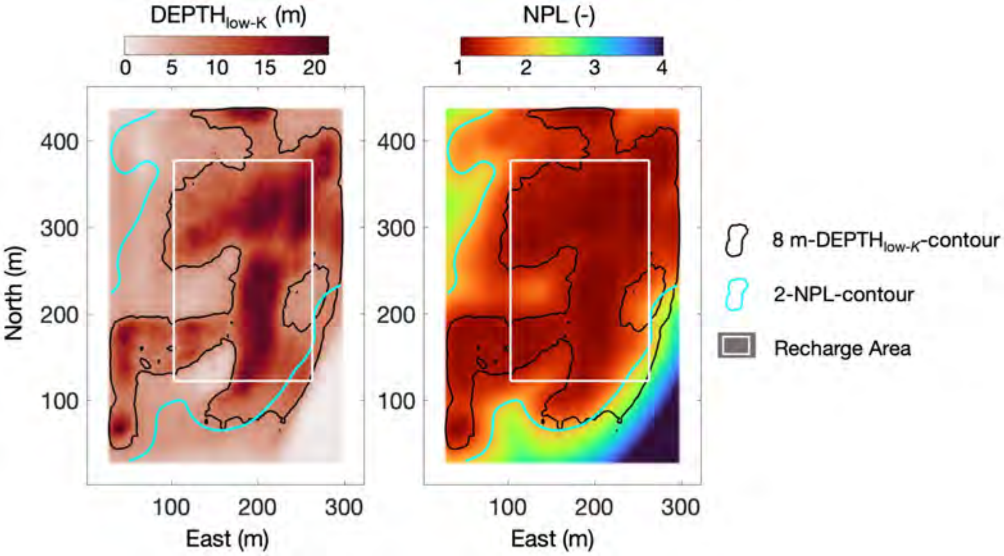


Figure E.11: Metric maps for Swall-South site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models.

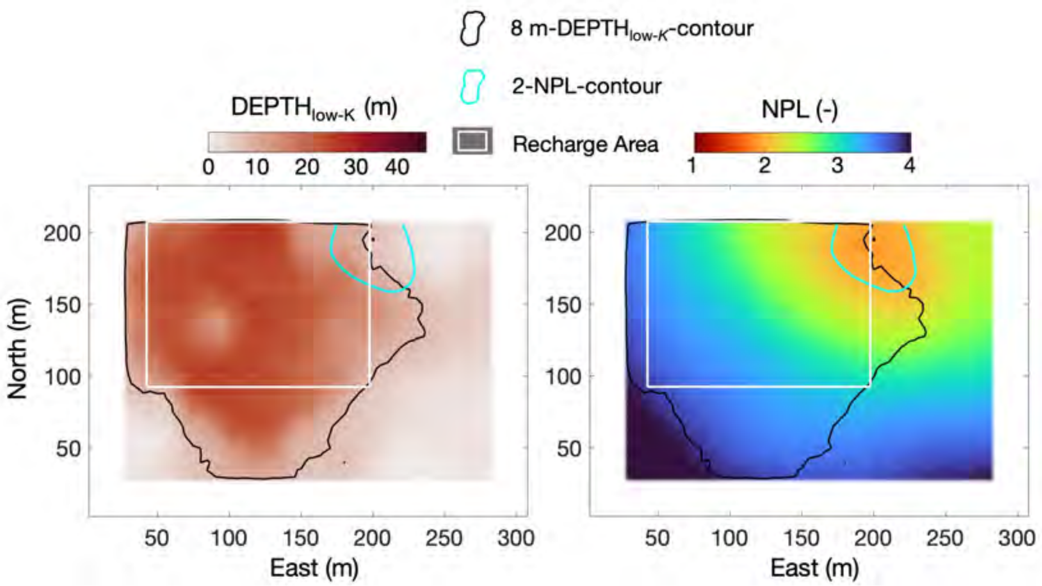


Figure E.12: Metric maps for Martin-West site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models.

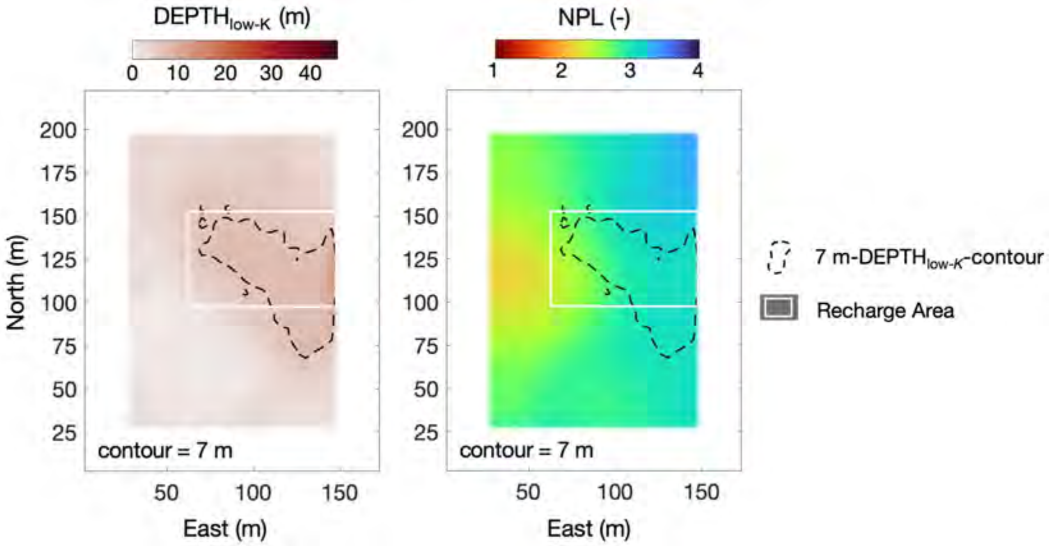


Figure E.13: Metric maps for Martin-East site. (a) Mean depth to the first low-K unit (DEPTH$_{\text{low-}K}$) and (b) mean normalized path lengths (NPL) are the average metric results calculated from the 200 generated models. There are no 8-m-DEPTH$_{\text{low-}K}$ or 2-NPL contours for this site; we do, however, show the 7-m-DEPTH$_{\text{low-}K}$ contour as the dashed black lines.

E.5.2 Comparison of Multiple Sites

We now consider the case where a number of sites are under consideration for recharge. Our approach can also be used to obtain a relative ranking of the sites, but it is difficult to draw any conclusions based on the metric maps alone because of variations in the size of sites and water table depths, except for scenarios in which we can eliminate sites from consideration because of little-to-no overlap in the metric contours (Martin-West and Martin-East; Figures E.12 and E.13, respectively). We developed a way of quantitatively comparing the sites by calculating the number and percentage of vertical high-K paths as a function of depth from the ground surface to the water table. The total number of possible vertical paths is the number of pixels in the $\mathrm{DEPTH}_{\text{low-}K}$ map, where the value at each pixel denotes the expected depth at which the first low-K unit will be encountered, indicating that the path vertically connecting the surface to the pixel's $\mathrm{DEPTH}_{\text{low-}K}$ consists only of high-K units. Moving from the surface to the local water table at 1-m depth slices, we defined the number of vertical paths at each depth as the number of pixels that had not yet exceeded their $\mathrm{DEPTH}_{\text{low-}K}$ value. The same exercise can be done for any area within a site, where the total number of possible vertical paths is the number of pixels within the selected area.

A decay in the number of vertical paths as a function of depth indicated that vertical flow paths initiating from multiple points at the surface were merging. Thus, a low rate of decay indicated a better site. Using both the number and percentage of vertical paths, where the percentage is the number of vertical paths at some depth normalized by the number at the surface, allowed us to identify, respectively, sites with (a) the largest areas well-suited for recharge and (b) the areas that retain the highest proportion of vertical paths with depth, independent of the site size. Figure E.14 presents this quantitative comparison of the seven sites. For each site, we assessed both the whole site (Figure E.14a,b) and the optimal area within the site (Figure E.14c,d), which was chosen using the approach described in the preceding section. Here, to demonstrate the utility of this approach to site comparison, we will present a few important observations that can be made.

As seen in Figure E.14a and c, some fields have many more vertical paths because of their larger starting surface areas. In each plot, the number of vertical paths at 0-m depth indicates the ground surface area of the site or selected optimal areas. The Nichols site has the largest surface area (1.5–24× the size of the other sites) and consistently has more vertical paths with depth than all other sites despite having significant decay beginning at $\sim$10 m. The same is true when comparing the selected optimal areas for the sites. Nichols has a large subsurface volume through which water can quickly move past the root zone, and it has a larger optimal area than any other site. Were there a desire to select a single site based on the criteria we are using and under the assumption that the recharge operations were not limited by water availability, Nichols would be ranked as the best of the sites.

Another key observation, seen in Figure E.14b and d, is that some sites retain a much higher percentage of vertical paths as depth increases regardless of the site size. A scenario in which this

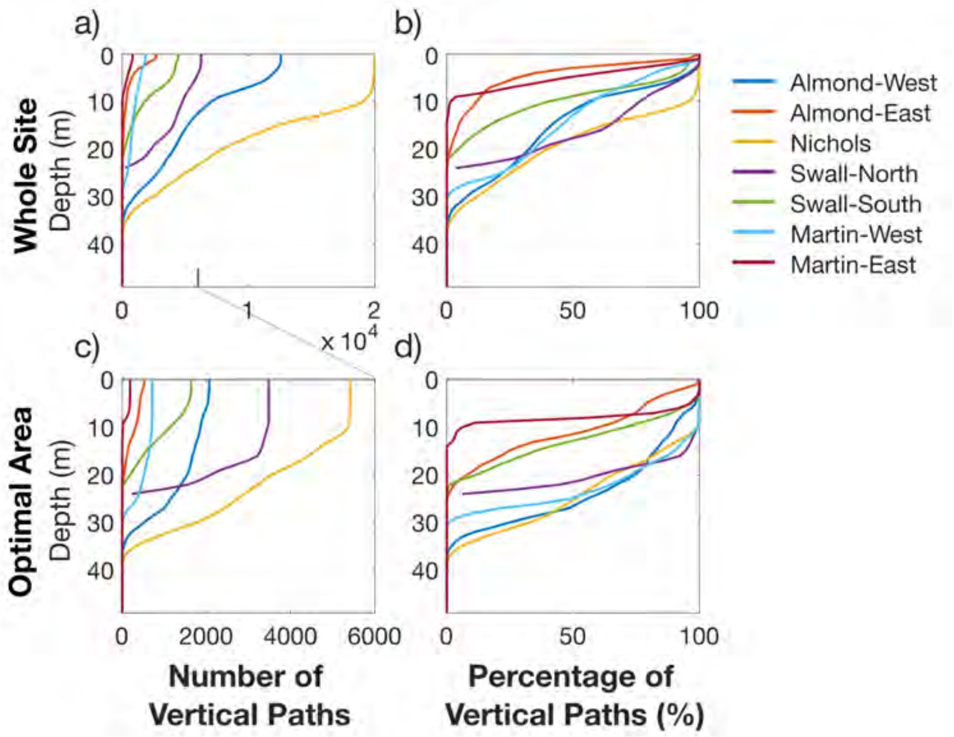


Figure E.14: Quantitative comparison of vertical paths vs. depth between all seven sites. We compare the number and percentage of vertical paths in two areas: Row 1, the whole site; Row 2, the optimal area; where percentage is the number of vertical paths normalized by the number at zero depth. The selection of an optimal area typically reduces the decay rate of vertical paths especially in the shallow (<8 m) portion of the subsurface

might be important would be one where the objective was to spread water over an entire site, or its optimal recharge area, and to maximize the value of the site (or the optimal area) per unit area covered with water, where value is defined as the retention of vertical paths. If we considered only the number of vertical paths, as noted previously, Nichols would seem to be the best choice. Looking at the percentage of vertical paths, however, Swall-North would emerge as comparable; for example, despite its significantly smaller size, the whole site retains ∼5% vertical paths to the depth of its water table at 24 m and ∼10% within its optimal recharge area. Nichols, on the other hand, loses 100% of its vertical paths by the depth of its water table at 41 m, even in the optimal area.

Assuming that all of these sites are farmland where ponding is of high concern, the number and percentage of vertical paths available from the surface to the water table in both the entire site and the optimal area would lead us to rank Nichols, Swall-North, and Almond-West as the best sites for recharge. Although Martin-West was comparable in terms of the results shown in Figure E.14d, this site showed little overlap between contours in its metric maps (Figure E.12) and thus should not be selected for recharge using our criteria. We note that the top three sites are also the largest. When searching for quality recharge areas, investigating larger fields increases the likelihood of finding an optimal area of sufficient size for recharge.

E.6 The Value of Generating Multiple Sediment-Type Models

The results from the previous section were derived from averaging across 200 generated sediment-type models for each site; therefore, the metric maps (Figure E.7) and vertical-path decay curves (Figure E.14) represent expected values. However, we could also use the metric maps and vertical-path decay curves from each individual model to capture the uncertainty in these curves as a result of differences in the models. In Figure E.15 we show (as gray lines) the decay in the percentage of vertical paths with depth for each model within each site's optimal area. The variation in the vertical-path decay curves displays the differences that could be present at a site and can be used to tailor the assessment of the site so as to avoid specific undesired outcomes. For example, if avoiding shallow ponding was a high priority, a choice could be made to reject sites where any model shows a significant decay in the number of vertical flow paths near the surface. Alternatively, if shallow ponding were of little concern, such as for reclamation of a floodplain, sites with many models that retain a high percentage of vertical paths to the water table (e.g., Swall-North and Swall-South) could be selected for further analysis regardless of the level of uncertainty in percentage of vertical paths in the shallow subsurface.

In Figure E.15 we show for each optimal area the mean of the individual vertical path decay curves (blue line) and the result from Figure E.14d (red line). If the 200 binary-sediment-type models for the optimal area at a site are similar, the red and blue curves will be similar. Alternatively, the red and blue curves will be significantly different when there is high variability in the individual models. For example, the individual vertical path decay curves from Almond-West are tightly clustered, and its blue and red curves show good agreement, particularly beyond a depth of 20 m. Swall-South, on the other hand, has a large spread in its individual curves. Consequently, its blue and red curves differ in magnitude and shape throughout the subsurface, suggesting we have little certainty about the subsurface and its effect on recharge at this site. The difference between the red and blue curves represents the extent to which the spatial structure varies from the average. In contributing models of the subsurface to be used by others in planning and implementing recharge operations, it is important to highlight the level of uncertainty in terms of the spatial distribution of sediments in the subsurface or how water will flow during recharge. Plots such as those in Figure E.15 can be used to assess and demonstrate this uncertainty and guide informed decisions for locating recharge sites.

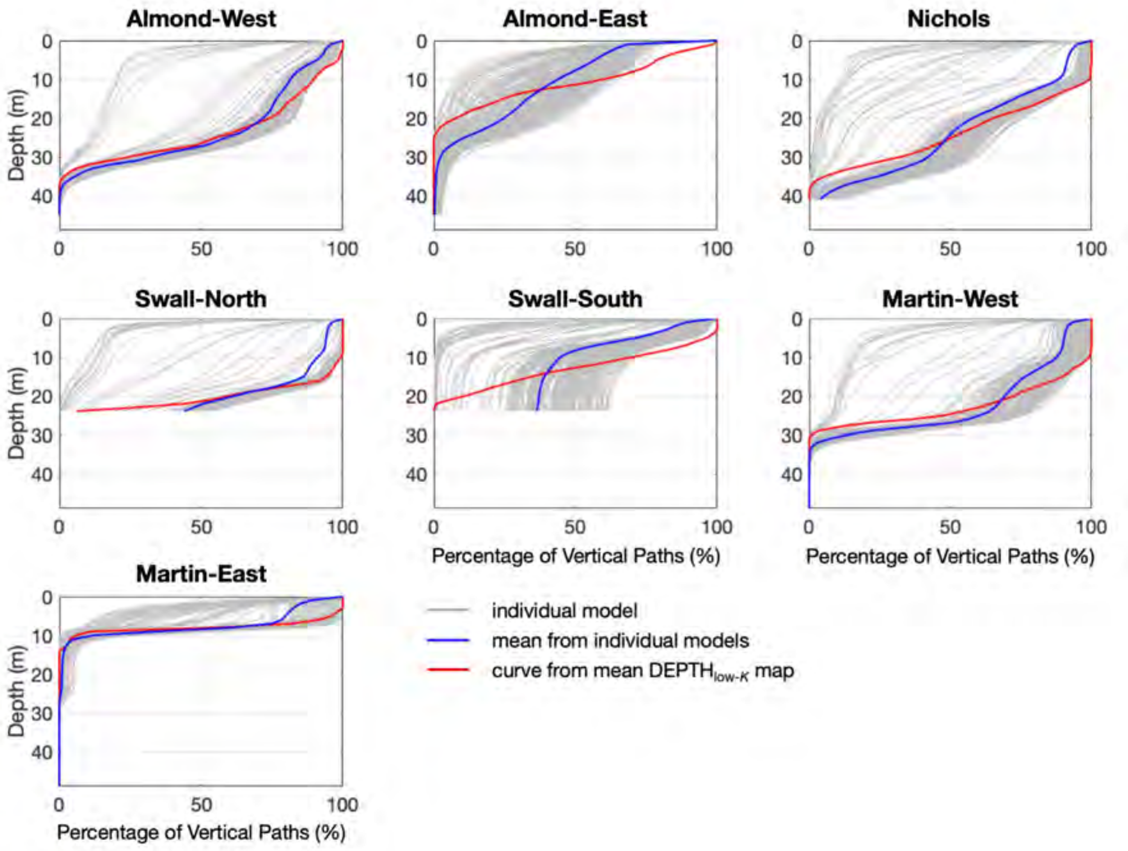


Figure E.15: Vertical path decay curves in the optimal areas. The individual model results highlight the variability around the various average representations of the sites.

E.7 Conclusions

In this chapter, we have demonstrated the use of tTEM data to select the optimal area within a site for surface spreading recharge based on the vertical distance between the surface and the shallowest low-K unit underlying each point at the surface (Metric 1) and the length of preferential flow paths between each point at the surface and the water table (Metric 2). Even with our simplified representation of the subsurface using a binary system, our results show high spatial variability in the distribution of binary-sediment-type at each site and, consequently, high variability in our metric assessments, underscoring the need for high-resolution geophysical imaging. The comparison of sites using these metrics becomes challenging because of differences in the size of the sites and the depth to the water table. We therefore developed an additional way to assess the relative suitability of each site using the decay in the number of vertical flow paths as a function of depth.

A limitation of our method is that we reduced the sediment types in the subsurface to a binary system (high-K and low-K). Although this is a simplified representation of the subsurface, it allowed us to demonstrate the use of the geostatistical methods and the Eikonal Solver. Our methods could be extended to create fully populated models of the fraction of coarse-grain-dominated material. With available field data that makes it possible to determine a relationship between sediment type and hydraulic properties (e.g., infiltration rate and K), we could then reconfigure the Eikonal Solver to determine actual travel times within the vadose zone.

We have developed methods that can be used with electromagnetic imaging data to rapidly identify promising sites for surface spreading recharge with respect to their ability to efficiently transmit water from the surface to the water table. Further studies could then be conducted at these sites including, for example, infiltration tests, modeling and sampling to better quantify flow rates and flow paths, and an assessment of potential effects on water quality. We wish to emphasize that the factors deemed most important in the selection of a site or an area within a site will vary depending on the goals and risk tolerance of the recharge operation. With the concerns of groundwater depletion in the CV, new approaches that can reduce the time and expense involved are needed to support the growing commitment to surface spreading for recharge. Electromagnetic imaging techniques with the complementary methods for data analysis have an important role to play.

E.8 Chapter Appendix: The Geostatistical Simulation

To obtain the spatial parameters required for the geostatistical simulation, co-sisim, we used the mean $\mathrm{P}_{\text{high-}K}$ model to develop a 'most-likely', sparse categorical model containing high-K and low-K units by classifying $\mathrm{P}_{\text{high-}K} \geq 0.5$ as a high-K unit and $\mathrm{P}_{\text{high-}K} < 0.5$ as a low-K unit. We refer to this categorical model as the sparse binary-sediment-type model.

Co-sisim requires a variogram model to define the spatial correlation of output categories. We constructed an experimental variogram in SGeMS using the sparse binary-sediment-type model and fit it with a model, allowing the model type to be either an exponential or spherical model (whichever fit best). We characterized the variogram model by its maximum horizontal range (maxHR), minimum horizontal range (minHR), and vertical range (VR), all defined in meters, and the angle of the maximum horizontal range (angle of maxHR), defined in degrees from north. The experimental variogram represented only an estimate of the spatial correlation using a sparse, most-likely model, and therefore introduced a major source of uncertainty. A way to account for this uncertainty is to create distributions for the input parameters from which to randomly draw during many simulation runs. This process results in many potential realizations of the subsurface, allowing us to identify where units are consistently high-K or low-K. In our workflow, we developed uniform (U) distributions for maxHR, minHR, and VR, each denoted with subscript j, as defined in Equations E.1-E.3:

$$maxHR_j \sim U[5, maxHR] \tag{E.1}$$

$$minHR_j \sim U[\frac{maxHR_j}{(maxHR/minHR)}, maxHR_j] \tag{E.2}$$

$$VR_j \sim U[1, min(VR, minHR_j)] \tag{E.3}$$

All of these distributions have units of meters. We defined our distributions assuming the calculated values for maxHR and VR were the maximum range values for the site, since EM methods typically result in smooth representations of the subsurface resistivity. We expected this to be the case with our tTEM data given the smoothing constraints used in the inversion of these data.

The other spatial parameters required as input to co-sisim were the global proportions of the two sediment types for each site, with that of the high-K category referred to as $\mathrm{G}_{\text{high-}K}$. This was developed by comparing the number of high-K units in the sparse binary-sediment-type model to its total number of filled voxels. Again, to account for the uncertainty in our estimate of $\mathrm{G}_{\text{high-}K}$, we generated a uniform distribution of values for $\mathrm{G}_{\text{high-}K}$, denoted with subscript j, as defined in Equation E.4:

$$G_{(high-K)_j} \sim U[G_{high-K} - 0.1, G_{high-K} + 0.1] \tag{E.4}$$

The final step for each site was to generate 200 fully-populated realizations of binary-sediment-type with 200 runs of co-sisim, each time using the sparse $\mathrm{P}_{\text{high-}K}$ and (1-$\mathrm{P}_{\text{high-}K}$) models and a randomly-selected set j of variogram parameters (maxHR_j, minHR_j, and VR_j) and $\mathrm{G}_{(high-K)_j}$, drawn from their respective distributions. Parameters $\mathrm{G}_{(high-K)_j}$ and maxHR_j do not depend on any of the other variables; however, to ensure $\mathrm{maxHR}_j \geq \mathrm{minHR}_j \geq \mathrm{VR}_j$, the distribution of minHR_j depended on the selected maxHR_j value; likewise, the distribution for VR_j depended on the selected minHR_j value. Thus, Equations E.1-E.3 were used sequentially in the selection of parameter set j. For simplicity, the angle of maxHR was held constant across the simulation runs.

Site	$G_{\text{high-}K}$	Variogram Model	maxHR	minHR	VR	Angle of maxHR	Depth to Water Table
	—		m	m	m	°	m
Almond-West	0.58	Exponential	187	125	57	90	45
Almond-East	0.40	Spherical	90	70	65	45	45
Nichols	0.66	Exponential	177	120	60	90	41
Swall-North	0.80	Spherical	244	116	39	45	24
Swall-South	0.76	Spherical	131	81	32	45	24
Martin-West	0.49	Exponential	199	169	70	45	49
Martin-East	0.38	Exponential	93	42	17	45	49

Table E.1: Calculated global proportion of high-K units ($G_{\text{high-}K}$) and variogram model parameters for each site. Depths to water table are based on the Tulare Irrigation District depth-to-groundwater map (TID, 2017).

Bibliography

Agram, P. S., & Simons, M. (2015). A noise model for InSAR time series. *Journal of Geophysical Research: Solid Earth*, *120*(4), 2752–2771. https://doi.org/10.1002/2014JB011271

Anderson, E. M. (1905). The dynamics of faulting. *Transactions of the Edinburgh Geological Society*, *8*(3), 387–402. https://doi.org/10.1144/transed.8.3.387

Aster, R. C., Borchers, B., & Thurber, C. H. (2013). Linear regression. *Parameter estimation and inverse problems* (pp. 25–53). Elsevier. https://doi.org/10.1016/B978-0-12-385048-5.00002-1

Auken, E., Foged, N., Larsen, J. J., Lassen, K. V. T., Maurya, P. K., Dath, S. M., & Eiskjær, T. T. (2019). tTEM — a towed transient electromagnetic system for detailed 3d imaging of the top 70 m of the subsurface. *GEOPHYSICS*, *84*(1), E13–E22. https://doi.org/10.1190/geo2018-0355.1

Bamler, R., Adam, N., Davidson, G., & Just, D. (1998). Noise-induced slope distortion in 2-d phase unwrapping by linear estimators with application to SAR interferometry [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(3), 913–921. https://doi.org/10.1109/36.673682

Battaglia, M., Cervelli, P. F., & Murray, J. R. (2013). Modeling crustal deformation near active faults and volcanic centers - a catalog of deformation models: US geological survey techniques and methods, book 13, ch. b1, 96 p., https://pubs.usgs.gov/tm/13/b1/pdf/tm13-b1.pdf

Beauducel, F., Briole, P., & Froger, J.-L. (2000). Volcano-wide fringes in ERS synthetic aperture radar interferograms of etna (1992-1998): Deformation or tropospheric effect? *Journal of Geophysical Research: Solid Earth*, *105*, 16391–16402. https://doi.org/10.1029/2000JB900095

Behroozmand, A. A., Auken, E., & Knight, R. (2019). Assessment of managed aquifer recharge sites using a new geophysical imaging method. *Vadose Zone Journal*, *18*(1), 1–13. https://doi.org/10.2136/vzj2018.10.0184

Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). 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. https://doi.org/10.1109/TGRS.2002.803792

Bianchi, M., Zheng, C., Wilson, C., Tick, G. R., Liu, G., & Gorelick, S. M. (2011). Spatial connectivity in a highly heterogeneous aquifer: From cores to preferential flow paths. *Water Resources Research*, *47*(5). https://doi.org/10.1029/2009WR008966

Biot, M. A. (1941). General theory of three-dimensional consolidation. *Journal of Applied Physics*, *12*(2), 155–164. https://doi.org/10.1063/1.1712886

Bouwer, H. (2002). Artificial recharge of groundwater: Hydrogeology and engineering. *Hydrogeology Journal*, *10*(1), 121–142. https://doi.org/10.1007/s10040-001-0182-4

Bracewell, R. (1986). *The fourier transform and its applications*. McGraw-Hill.

Buckland, J. R., Huntley, J. M., & Turner, S. R. E. (1995). Unwrapping noisy phase maps by use of a minimum-cost-matching algorithm [Publisher: Optica Publishing Group]. *Applied Optics*, *34*(23), 5100–5108. https://doi.org/10.1364/AO.34.005100

Caers, J., & Hoffman, T. (2006). The probability perturbation method: A new look at bayesian inverse modeling. *Mathematical Geology*, *38*(1), 81–100. https://doi.org/10.1007/s11004-005-9005-9

CalTech. (2022, March 27). *ISCE2* [original-date: 2019-02-09T23:14:54Z]. isce-framework. Retrieved March 28, 2022, from https://github.com/isce-framework/isce2

Candela, T., Wassing, B., ter Heege, J., & Buijze, L. (2018). How earthquakes are induced. *Science*, *360*(6389), 598–600. https://doi.org/10.1126/science.aat2776

Chang, K. W., & Segall, P. (2016a). Injection-induced seismicity on basement faults including poroelastic stressing. *Journal of Geophysical Research: Solid Earth*, *121*(4), 2708–2726. https://doi.org/10.1002/2015JB012561

Chang, K. W., & Segall, P. (2016b). Seismicity on basement faults induced by simultaneous fluid injection–extraction. *Pure and Applied Geophysics*, *173*(8), 2621–2636. https://doi.org/10.1007/s00024-016-1319-7

Charzynski, K., Faith, K., Fenton, Z., Shedeed, A., McKee, M., Bjorlie, S., & Richardson, M. (2019). Delaware basin horizontal wolfcamp case study: Mitigating h2s and excessive water production through isolating densely fractured intervals correlative to seismically mapped shallow graben features in the delaware mountain group. *Proceedings of the 7th Unconventional Resources Technology Conference*. https://doi.org/10.15530/urtec-2019-1037

Chen, C. W. (2001). *Statistical-cost network-flow approaches to two-dimensional phase unwrapping for radar interferometry* (Doctoral dissertation).

Chen, C. W., & Zebker, H. (2000). Network approaches to two-dimensional phase unwrapping: Intractability and two new algorithms [Publisher: Optica Publishing Group]. *JOSA A*, *17*(3), 401–414. https://doi.org/10.1364/JOSAA.17.000401

Chen, C. W., & Zebker, H. (2001). Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization [Publisher: Optica Publishing Group]. *JOSA A*, *18*(2), 338–351. https://doi.org/10.1364/JOSAA.18.000338

Chen, J., Knight, R., & Zebker, H. (2017). The temporal and spatial variability of the confined aquifer head and storage properties in the san luis valley, colorado inferred from multiple InSAR missions. *Water Resources Research*, *53*(11), 9708–9720. https://doi.org/10.1002/2017WR020881

Chilingar, G., & Knight, L. (1960). Relationship between pressure and moisture content of kaolinite, illite, and montmorillonite clays. *AAPG Bulletin*, *44*. https://doi.org/10.1306/0BDA5F81-16BD-11D7-8645000102C1865D

Christensen, O. F. (2004). Monte carlo maximum likelihood in model-based geostatistics. *Journal of Computational and Graphical Statistics*, *13*(3), 702–718. https://doi.org/10.1198/106186004X2525

Cook, P., Walker, G., Buselli, G., Potts, I., & Dodds, A. (1992). The application of electromagnetic techniques to groundwater recharge investigations. *Journal of Hydrology*, *130*(1), 201–229. https://doi.org/10.1016/0022-1694(92)90111-8

Cornet, F. H., Helm, J., Poitrenaud, H., & Etchecopar, A. (1998). Seismic and aseismic slips induced by large-scale fluid injections. In S. Talebi (Ed.), *Seismicity associated with mines, reservoirs and fluid injections* (pp. 563–583). Birkhäuser. https://doi.org/10.1007/978-3-0348-8814-1_12

Costantini, M. (1998). A novel phase unwrapping method based on network programming [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(3), 813–821. https://doi.org/10.1109/36.673674

Cumming, I. G., & Wong, F. H. (2004). *ARTECH HOUSE USA : Digital processing of synthetic aperture radar data*. Artech House. Retrieved March 28, 2022, from https://us-artechhouse-com.stanford.idm.oclc.org/Digital-Processing-of-Synthetic-Aperture-Radar-Data-P1549.aspx

Cusack, R., Huntley, J. M., & Goldrein, H. T. (1995). Improved noise-immune phase-unwrapping algorithm [Publisher: Optica Publishing Group]. *Applied Optics*, *34*(5), 781–789. https://doi.org/10.1364/AO.34.000781

Dahlke, H. E., Brown, A. G., Orloff, S., Putnam, D., & O'Geen, T. (2018). Managed winter flooding of alfalfa recharges groundwater with minimal crop damage. *California Agriculture*, *72*(1), 1–11. https://doi.org/10.3733/ca.2018a0001

Davidson, G., & Bamler, R. (1999). Multiresolution phase unwrapping for SAR interferometry [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(1), 163–174. https://doi.org/10.1109/36.739150

De Zan, F., Parizzi, A., Prats-Iraola, P., & Lopez-Dekker, P. (2014). A SAR interferometric model for soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 418–425. https://doi.org/10.1109/TGRS.2013.2241069

De Zan, F., Zonno, M., & Lopez-Dekker, P. (2015). Phase inconsistencies and multiple scattering in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(12), 6608–6616. https://doi.org/10.1109/TGRS.2015.2444431

Delacourt, C., Briole, P., & Achache, J. A. (1998). Tropospheric corrections of SAR interferograms with strong topography. application to etna. *Geophysical Research Letters*, *25*(15), 2849–2852. https://doi.org/10.1029/98GL02112

Deng, F., Dixon, T. H., & Xie, S. (2020). Surface deformation and induced seismicity due to fluid injection and oil and gas extraction in western texas. *Journal of Geophysical Research: Solid Earth*, *125*(5). https://doi.org/10.1029/2019JB018962

Dvory, N. Z., & Zoback, M. D. (2021). Prior oil and gas production can limit the occurrence of injection-induced seismicity: A case study in the delaware basin of western texas and southeastern new mexico, USA. *Geology*, *49*(10), 1198–1203. https://doi.org/10.1130/G49015.1

Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science*, *341*(6142), 1225942. https://doi.org/10.1126/science.1225942

Emardson, T. R., Simons, M., & Webb, F. H. (2003). Neutral atmospheric delay in interferometric synthetic aperture radar applications: Statistical description and mitigation. *Journal of Geophysical Research: Solid Earth*, *108*. https://doi.org/10.1029/2002JB001781

Enverus, D. (1999). *Enverus gallery* [Enverus DrillingInfo gallery]. Retrieved March 10, 2022, from https://app.drillinginfo.com/gallery/

ESA. (n.d.). *STEP – science toolbox exploitation platform*. Retrieved March 28, 2022, from http://step.esa.int/main/

Eyre, T. S., Eaton, D. W., Garagash, D. I., Zecevic, M., Venieri, M., Weir, R., & Lawton, D. C. (2019). The role of aseismic slip in hydraulic fracturing–induced seismicity. *Science Advances*, *5*(8), eaav7172. https://doi.org/10.1126/sciadv.aav7172

Eyre, T. S., Zecevic, M., Salvage, R. O., & Eaton, D. W. (2020). A long-lived swarm of hydraulic fracturing-induced seismicity provides evidence for aseismic slip. *Bulletin of the Seismological Society of America*, *110*(5), 2205–2215. https://doi.org/10.1785/0120200107

Farahbod, A. M., Kao, H., Cassidy, J. F., & Walker, D. (2015). How did hydraulic-fracturing operations in the horn river basin change seismicity patterns in northeastern british columbia, canada? *The Leading Edge*, *34*(6), 658–663. https://doi.org/10.1190/tle34060658.1

Faunt, C. (2009). *Professional paper* (Professional Paper) [Series: Professional Paper].

Ferretti, A., Prati, C., & Rocca, F. (2001). Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(1), 8–20. https://doi.org/10.1109/36.898661

Fogg, G. E., Carle, S. F., & Green, C. (2000). Connected-network paradigm for the alluvial aquifer system. *Theory, modeling, and field investigation in hydrogeology: A special volume in honor of shlomo p. neumans 60th birthday*. Geological Society of America. https://doi.org/10.1130/0-8137-2348-5.25

Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R., & Davies, R. J. (2018). Global review of human-induced earthquakes. *Earth-Science Reviews*, *178*, 438–514. https://doi.org/10.1016/j.earscirev.2017.07.008

Fried, D. L. (1977). Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements [Publisher: Optica Publishing Group]. *JOSA*, *67*(3), 370–375. https://doi.org/10.1364/JOSA.67.000370

Frohlich, C., DeShon, H., Stump, B., Hayward, C., Hornbach, M., & Walter, J. I. (2016). A historical review of induced earthquakes in texas. *Seismological Research Letters*, *87*(4), 1022–1038. https://doi.org/10.1785/0220160016

Frohlich, C., Ellsworth, W., Brown, W. A., Brunt, M., Luetgert, J., MacDonald, T., & Walter, S. (2014). The 17 may 2012 *M* 4.8 earthquake near timpson, east texas: An event possibly triggered by fluid injection: 2012 may 17 east texas earthquake. *Journal of Geophysical Research: Solid Earth*, *119*(1), 581–593. https://doi.org/10.1002/2013JB010755

Furtney, J. (2015). *Scikit-fmm* [GitHub]. Retrieved March 25, 2022, from https://github.com/scikit-fmm

Gambolati, G., Lewis, R. W., Schrefler, B. A., & Simoni, L. (1992). Comment on 'coupling versus uncoupling in soil consolidation'. *International Journal for Numerical and Analytical Methods in Geomechanics*, *16*(11), 833–837. https://doi.org/10.1002/nag.1610161105

Gambolati, G., Teatini, P., & Ferronato, M. (2005, October 21). Anthropogenic land subsidence. In M. G. Anderson & J. J. McDonnell (Eds.), *Encyclopedia of hydrological sciences* (hsa164b). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470848944.hsa164b

Ganot, Y., & Dahlke, H. E. (2021). Natural and forced soil aeration during agricultural managed aquifer recharge. *Vadose Zone Journal*, *20*(3). https://doi.org/10.1002/vzj2.20128

Ghiglia, D., & Pritt, M. (1998). *Two-dimensional phase unwrapping: Theory, algorithms, and software*. A Wiley Interscience Publication. Retrieved April 12, 2022, from https://analyticalscience.wiley.com/do/10.1002/sepspec.9780471249351

Ghiglia, D. C., & Romero, L. A. (1994). Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *Journal of the Optical Society of America A*, *11*(1), 107. https://doi.org/10.1364/JOSAA.11.000107

Ghiglia, D. C., & Romero, L. A. (1996). Minimum lp-norm two-dimensional phase unwrapping. *Journal of the Optical Society of America A*, *13*(10), 1999. https://doi.org/10.1364/JOSAA.13.001999

Goebel, M., & Knight, R. (2021). Recharge site assessment through the integration of surface geophysics and cone penetrometer testing. *Vadose Zone Journal*, *20*(4). https://doi.org/10.1002/vzj2.20131

Goldstein, R. M., Zebker, H., & Werner, C. L. (1988). Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Science*, *23*(4), 713–720. https://doi.org/10.1029/RS023i004p00713

Gottschalk, I. P., Hermans, T., Knight, R., Caers, J., Cameron, D. A., Regnery, J., & McCray, J. E. (2017). Integrating non-colocated well and geophysical data to capture subsurface heterogeneity at an aquifer recharge and recovery site. *Journal of Hydrology*, *555*, 407–419. https://doi.org/10.1016/j.jhydrol.2017.10.028

Grandin, R., Vallée, M., & Lacassin, R. (2017). Rupture process of the M_{w} 5.8 pawnee, oklahoma, earthquake from sentinel-1 InSAR and seismological data. *Seismological Research Letters*, *88*(4), 994–1004. https://doi.org/10.1785/0220160226

Guglielmi, Y., Cappa, F., Avouac, J.-P., Henry, P., & Elsworth, D. (2015). Seismicity triggered by fluid injection–induced aseismic slip. *Science*, *348*(6240), 1224–1226. https://doi.org/10.1126/science.aab0476

Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research*, *84*, 2348. https://doi.org/10.1029/JB084iB05p02348

Hennings, P., Dvory, N., Horne, E., Li, P., Savvaidis, A., & Zoback, M. (2021). Stability of the fault systems that host-induced earthquakes in the delaware basin of west texas and southeast new mexico. *The Seismic Record*, *1*(2), 96–106. https://doi.org/10.1785/0320210020

Hernandez-Marin, M., & Burbey, T. J. (2012). Fault-controlled deformation and stress from pumping-induced groundwater flow. *Journal of Hydrology*, *428-429*, 80–93. https://doi.org/10.1016/j.jhydrol.2012.01.025

Holzer, T. L. (1981). Preconsolidation stress of aquifer systems in areas of induced land subsidence. *Water Resources Research*, *17*(3), 693–703. https://doi.org/10.1029/WR017i003p00693

Hooper, A. (n.d.). *StaMPS*. Retrieved March 28, 2022, from http://homepages.see.leeds.ac.uk/~earahoo/stamps/

Hornbach, M. J., DeShon, H. R., Ellsworth, W. L., Stump, B. W., Hayward, C., Frohlich, C., Oldham, H. R., Olson, J. E., Magnani, M. B., Brokaw, C., & Luetgert, J. H. (2015). Causal factors for seismicity near azle, texas. *Nature Communications*, *6*(1), 6728. https://doi.org/10.1038/ncomms7728

Horne, E. (2020). Data for: Horne et al. (2021) basement-rooted faults of the delaware basin and central basin platform, permian basin, west texas and southeastern new mexico [Type: dataset]. https://doi.org/10.18738/T8/UHOUX8

Hudgin, R. H. (1977). Wave-front reconstruction for compensated imaging [Publisher: Optica Publishing Group]. *JOSA*, *67*(3), 375–378. https://doi.org/10.1364/JOSA.67.000375

Hunt, B. R. (1979). Matrix formulation of the reconstruction of phase values from phase differences [Publisher: Optica Publishing Group]. *JOSA*, *69*(3), 393–399. https://doi.org/10.1364/JOSA.69.000393

Jacob, C. E. (1940). On the flow of water in an elastic artesian aquifer. *Transactions, American Geophysical Union*, *21*(2), 574. https://doi.org/10.1029/TR021i002p00574

Jankowski, S., Faunt, C. C., & Phillips, S. P. (2018). Optimal managed aquifer recharge to address land subsidence within california's central valley. Retrieved March 25, 2022, from https://agu.confex.com/agu/fm18/meetingapp.cgi/Paper/468557

Jolivet, R., Grandin, R., Lasserre, C., Doin, M.-P., & Peltzer, G. (2011). Systematic InSAR tropospheric phase delay corrections from global meteorological reanalysis data. *Geophysical Research Letters*, *38*(17), n/a–n/a. https://doi.org/10.1029/2011GL048757

Jolivet, R., Agram, P. S., Lin, N. Y., Simons, M., Doin, M.-P., Peltzer, G., & Li, Z. (2014). Improving InSAR geodesy using global atmospheric models. *Journal of Geophysical Research: Solid Earth*, *119*(3), 2324–2341. https://doi.org/10.1002/2013JB010588

Just, D., & Bamler, R. (1994). Phase statistics of interferograms with applications to synthetic aperture radar. *Applied Optics*, *33*(20), 4361. https://doi.org/10.1364/AO.33.004361

Kanamori, H., & Anderson, D. L. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America*, *65*(5), 1073–1095.

Kang, S., Knight, R., Greene, T., Buck, C., & Fogg, G. (2021). Exploring the model space of airborne electromagnetic data to delineate large-scale structure and heterogeneity within an aquifer system. *Water Resources Research*, *57*(10). https://doi.org/10.1029/2021WR029699

Keranen, K. M., Savage, H. M., Abers, G. A., & Cochran, E. S. (2013). Potentially induced earthquakes in oklahoma, USA: Links between wastewater injection and the 2011 mw 5.7 earthquake sequence. *Geology*, *41*(6), 699–702. https://doi.org/10.1130/G34045.1

Kim, J.-W., & Lu, Z. (2018). Association between localized geohazards in west texas and human activities, recognized by sentinel-1a/b satellite radar imagery. *Scientific Reports*, *8*(1), 4727. https://doi.org/10.1038/s41598-018-23143-6

Kim, J.-W., Lu, Z., & Kaufmann, J. (2019). Evolution of sinkholes over wink, texas, observed by high-resolution optical and SAR imagery. *Remote Sensing of Environment*, *222*, 119–132. https://doi.org/10.1016/j.rse.2018.12.028

Kovach, R. L. (1974). Source mechanisms for wilmington oil field, california, subsidence earthquakes. *Bulletin of the Seismological Society of America*, *64*(3), 699–711. https://doi.org/10.1785/BSSA0643-10699

Lawrie, K. C., Brodie, R. S., Dillon, P., Tan, K. P., Gibson, D., Magee, J., Clarke, J. D., Somerville, P., Gow, L., Halas, L., Apps, H. E., Page, D., Vanderzalm, J., Abraham, J., Hostetler, S., Christensen, N. B., Miotlinski, K., Brodie, R. C., Smith, M., & Schoning, G. (2013).

Assessment of conjuctive water supply options to enhance the drought security of broken hill, regional communities and industries: Summary report. Geoscience Australia.

Lee, J.-S., Hoppel, K., Mango, S., & Miller, A. (1994). Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *32*(5), 1017–1028. https://doi.org/10.1109/36.312890

Li, F., & Goldstein, R. (1990). Studies of multibaseline spaceborne interferometric synthetic aperture radars [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(1), 88–97. https://doi.org/10.1109/36.45749

Lomax, A., & Savvaidis, A. (2019). Improving absolute earthquake location in west texas using probabilistic, proxy ground-truth station corrections. *Journal of Geophysical Research: Solid Earth*, *124*(11), 11447–11465. https://doi.org/10.1029/2019JB017727

Lund Snee, J.-E., & Zoback, M. D. (2018). State of stress in the permian basin, texas and new mexico: Implications for induced seismicity. *The Leading Edge*, *37*(2), 127–134. https://doi.org/10.1190/tle37020127.1

Maliva, R. G., Herrmann, R., Coulibaly, K., & Guo, W. (2015). Advanced aquifer characterization for optimization of managed aquifer recharge. *Environmental Earth Sciences*, *73*(12), 7759–7767. https://doi.org/10.1007/s12665-014-3167-z

Massonnet, D., & Feigl, K. L. (1995). Satellite radar interferometric map of the coseismic deformation field of the m = 6.1 eureka valley, california earthquake of may 17, 1993. *Geophysical Research Letters*, *22*(12), 1541–1544. https://doi.org/10.1029/95GL01088

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., & Rabaute, T. (1993). The displacement field of the landers earthquake mapped by radar interferometry. *Nature*, *364*(6433), 138–142. https://doi.org/10.1038/364138a0

Mawer, C., Kitanidis, P., Pidlisecky, A., & Knight, R. (2013). Electrical resistivity for characterization and infiltration monitoring beneath a managed aquifer recharge pond [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.2136/vzj2011.0203]. *Vadose Zone Journal*, *12*(1), vzj2011.0203. https://doi.org/10.2136/vzj2011.0203

Michaelides, R. J., Zebker, H., & Zheng, Y. (2019). An algorithm for estimating and correcting decorrelation phase from InSAR data using closure phase triplets. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(12), 10390–10397. https://doi.org/10.1109/TGRS.2019.2934362

Michaelides, R. J. (2020). *Quantifying permafrost processes and soil moisture with interferometric phase and closure phase* (Doctoral dissertation). Stanford University. Stanford, California.

MKGA. (2019). Groundwater sustainability plan - mid-kaweah groundwater sustainability agency. https://www.midkaweah.org/%20documents

Neely, W. R., Borsa, A. A., & Silverii, F. (2020). GInSAR: A cGPS correction for enhanced InSAR time series. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(1), 136–146. https://doi.org/10.1109/TGRS.2019.2934118

Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, *47*(2), 617–644. https://doi.org/10.1109/T-AIEE.1928.5055024

O'Geen, A., Saal, M., Dahlke, H., Doll, D., Elkins, R., Fulton, A., Fogg, G., Harter, T., Hopmans, J. W., Ingels, C., Niederholzer, F., Solis, S. S., Verdegaal, P., & Walkinshaw, M. (2015). Soil suitability index identifies potential areas for groundwater banking on agricultural lands. *California Agriculture*, *69*(2), 75–84. https://doi.org/10.3733/ca.v069n02p75

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space, 20.

Onn, F., & Zebker, H. (2006). Correction for interferometric synthetic aperture radar atmospheric phase artifacts using time series of zenith wet delay observations from a GPS network. *Journal of Geophysical Research*, *111*, B09102. https://doi.org/10.1029/2005JB004012

Osmanoğlu, B., Sunar, F., Wdowinski, S., & Cabral-Cano, E. (2016). Time series analysis of InSAR data: Methods and trends. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 90–102. https://doi.org/10.1016/j.isprsjprs.2015.10.003

Parsekian, A. D., Regnery, J., Wing, A. D., Knight, R., & Drewes, J. E. (2014). Geophysical and hydrochemical identification of flow paths with implications for water quality at an ARR site. *Groundwater Monitoring & Remediation*, *34*(3), 105–116. https://doi.org/10.1111/gwmr.12071

Pepin, K. S., Ellsworth, W. L., Sheng, Y., & Zebker, H. (2022a). Shallow aseismic slip in the delaware basin determined by sentinel-1 InSAR. *Journal of Geophysical Research: Solid Earth*, *127*(2). https://doi.org/10.1029/2021JB023157

Pepin, K., Knight, R., Goebel-Szenher, M., & Kang, S. (2022b). Managed aquifer recharge site assessment with electromagnetic imaging: Identification of recharge flow paths. *Vadose Zone Journal*. https://doi.org/10.1002/vzj2.20192

Poland, J. F. (1984). Mechanics of land subsidence due to fluid withdrawal. *Guidebook to studies of land subsidence due to ground-water withdrawal: United Nations Educational, Scientific, and Cultural Organization (UNESCO), Studies and Reports in Hydrology*, *40*, 37–54.

Poland, J. F., Lofgren, B. E., Ireland, R. L., & Pugh, R. G. (1975). *Land subsidence, in the san joaquin valley, california, as of 1972: A history of land subsidence caused by water-level decline in the san joaquin valley, from the 1920's to 1972* [Google-Books-ID: ACUUJsHZ_c4C]. U.S. Government Printing Office.

Pritt, M. (1996). Phase unwrapping by means of multigrid techniques for interferometric SAR [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *34*(3), 728–738. https://doi.org/10.1109/36.499752

Remy, N., Boucher, A., & Wu, J. (2009). *About SGeMS — SGeMS*. Retrieved March 25, 2022, from http://sgems.sourceforge.net/?q=node/20

Riley, F. S. (1969). Analysis of borehole extensometer data from central california, 9.

Rizzo, C. B., & de Barros, F. P. J. (2017). Minimum hydraulic resistance and least resistance path in heterogeneous porous media. *Water Resources Research*, *53*(10), 8596–8613. https://doi.org/10.1002/2017WR020418

Rodriguez, E., & Martin, J. (1992). Theory and design of interferometric synthetic aperture radars. *IEE Proceedings F Radar and Signal Processing*, *139*(2), 147. https://doi.org/10.1049/ip-f-2.1992.0018

Rue, H. (2001). Fast sampling of gaussian markov random fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 325–338. https://doi.org/10.1111/1467-9868.00288

Samiei-Esfahany, S., Martins, J. E., van Leijen, F., & Hanssen, R. F. (2016). Phase estimation for distributed scatterers in InSAR stacks using integer least squares estimation [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(10), 5671–5687. https://doi.org/10.1109/TGRS.2016.2566604

Savvaidis, A., Lomax, A., & Breton, C. (2020). Induced seismicity in the delaware basin, west texas, is caused by hydraulic fracturing and wastewater disposal. *Bulletin of the Seismological Society of America*, *110*(5), 2225–2241. https://doi.org/10.1785/0120200087

Savvaidis, A., Young, B., Huang, G.-c. D., & Lomax, A. (2019). TexNet: A statewide seismological network in texas. *Seismological Research Letters*, 1225942. https://doi.org/10.1785/0220180350

Schultz, R., Skoumal, R. J., Brudzinski, M. R., Eaton, D., Baptie, B., & Ellsworth, W. (2020). Hydraulic fracturing-induced seismicity. *Reviews of Geophysics*, *58*(3). https://doi.org/10.1029/2019RG000695

Sendrós, A., Himi, M., Lovera, R., Rivero, L., Garcia-Artigas, R., Urruela, A., & Casas, A. (2020). Geophysical characterization of hydraulic properties around a managed aquifer recharge system over the llobregat river alluvial aquifer (barcelona metropolitan area). *Water*, *12*(12), 3455. https://doi.org/10.3390/w12123455

Sethian, J. A. (1996). A fast marching level set method for monotonically advancing fronts. [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *93*(4), 1591–1595. https://doi.org/10.1073/pnas.93.4.1591

Sheng, Y., Pepin, K. S., & Ellsworth, W. L. (2022). On the depth of earthquakes in the delaware basin: A case study along the reeves–pecos county line. *The Seismic Record*, *2*(1), 29–37. https://doi.org/10.1785/0320210048

Shirzaei, M., Manga, M., & Zhai, G. (2019). Hydraulic properties of injection formations constrained by surface deformation. *Earth and Planetary Science Letters*, *515*, 125–134. https://doi.org/10.1016/j.epsl.2019.03.025

Skoumal, R. J., Barbour, A. J., Brudzinski, M. R., Langenkamp, T., & Kaven, J. O. (2020). Induced seismicity in the delaware basin, texas. *Journal of Geophysical Research: Solid Earth*, *125*(1). https://doi.org/10.1029/2019JB018558

Smith, R., Knight, R., Chen, J., Reeves, J. A., Zebker, H., Farr, T., & Liu, Z. (2017). Estimating the permanent loss of groundwater storage in the southern san joaquin valley, california. *Water Resources Research*, *53*(3), 2133–2148. https://doi.org/10.1002/2016WR019861

Smith, R., & Li, J. (2021). Modeling elastic and inelastic pumping-induced deformation with incomplete water level records in parowan valley, utah. *Journal of Hydrology*, *601*, 126654. https://doi.org/10.1016/j.jhydrol.2021.126654

Sneed, M. (2001). *Hydraulic and mechanical properties affecting ground-water flow and aquifer-system compaction, san joaquin valley, california*. U.S. Department of the Interior, U.S. Geological Survey.

Soutter, M., & Musy, A. (1998). Coupling 1d monte-carlo simulations and geostatistics to assess groundwater vulnerability to pesticide contamination on a regional scale. *Journal of Contaminant Hydrology*, *32*(1), 25–39. https://doi.org/10.1016/S0169-7722(97)00075-2

Spagnolini, U. (1993). 2-d phase unwrapping and phase aliasing [Publisher: Society of Exploration Geophysicists]. *GEOPHYSICS*, *58*(9), 1324–1334. https://doi.org/10.1190/1.1443515

Staniewicz, S., Chen, J., Lee, H., Olson, J., Savvaidis, A., & Hennings, P. (2021, January 27). *Cumulative and transient surface deformation signals in the permian basin* (other). Geodesy. https://doi.org/10.1002/essoar.10505979.1

Staniewicz, S., Chen, J., Lee, H., Olson, J., Savvaidis, A., Reedy, R., Breton, C., Rathje, E., & Hennings, P. (2020). InSAR reveals complex surface deformation patterns over an 80,000 km ^{2} oil-producing region in the permian basin. *Geophysical Research Letters*, *47*(21). https://doi.org/10.1029/2020GL090151

Starr, A. T. (1928). Slip in a crystal and rupture in a solid due to shear. *Mathematical Proceedings of the Cambridge Philosophical Society*, *24*(4), 489–500. https://doi.org/10.1017/S0305004100014626

Suckale, J. (2010). Moderate-to-large seismicity induced by hydrocarbon production. *The Leading Edge*, *29*(3), 310–319. https://doi.org/10.1190/1.3353728

Teng, G., & Baker, J. W. (2020). Short-term probabilistic hazard assessment in regions of induced seismicity. *Bulletin of the Seismological Society of America*, *110*(5), 2441–2453. https://doi.org/10.1785/0120200081

Terzaghi, K. (1923). Die berechnung der durchlassigkeitsziffer des tones aus dem verlauf der hydrodynamischen spannungserscheinungen, akad wissensch wien sitzungsber mathnaturwissensch klasse IIa. *142*, 125–138.

Terzaghi, K. (1925). Structure and volume of voids in soils, translated from erdbaumechanik auf bodenphysikalischer grundlage. from theory to practice in soil mechanics. new york, NY: John wiley.

Thomas, G., & Finney, R. (1988). *Calculus and analytic geometry*. Addison-Wesley Publishers Limited.

TID. (2017). *Tulare irrigation district depth to groundwater map*. Retrieved March 25, 2022, from https://tulareid.org/fall-2017-gw-1pdf

Touzi, R., Lopes, A., Bruniquel, J., & Vachon, P. (1999). Coherence estimation for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(1), 135–149. https://doi.org/10.1109/36.739146

USGS. (2016). *California's central valley — USGS california water science center*. Retrieved March 25, 2022, from https://ca.water.usgs.gov/projects/central-valley/about-central-valley.html

Wang, Z., Yin, Z., Caers, J., & Zuo, R. (2020). A monte carlo-based framework for risk-return analysis in mineral prospectivity mapping. *Geoscience Frontiers*, *11*(6), 2297–2308. https://doi.org/10.1016/j.gsf.2020.02.010

Wei, M., Sandwell, D., & Smith-Konter, B. (2010). Optimal combination of InSAR and GPS for measuring interseismic crustal deformation. *Advances in Space Research*, *46*(2), 236–249. https://doi.org/10.1016/j.asr.2010.03.013

Weston, J., Ferreira, A. M., & Funning, G. J. (2012). Systematic comparisons of earthquake source models determined using InSAR and seismic data. *Tectonophysics*, *532-535*, 61–81. https://doi.org/10.1016/j.tecto.2012.02.001

Yan, L., Zhang, H., Zhang, R., Xie, X., & Chen, B. (2019). A robust phase unwrapping algorithm based on reliability mask and weighted minimum least-squares method. *Optics and Lasers in Engineering*, *112*, 39–45. https://doi.org/10.1016/j.optlaseng.2018.08.024

Yeck, W. L., Hayes, G. P., McNamara, D. E., Rubinstein, J. L., Barnhart, W. D., Earle, P. S., & Benz, H. M. (2017). Oklahoma experiences largest earthquake during ongoing regional wastewater injection hazard mitigation efforts. *Geophysical Research Letters*, *44*(2), 711–717. https://doi.org/10.1002/2016GL071685

Yu, H., Lan, Y., Yuan, Z., Xu, J., & Lee, H. (2019). Phase unwrapping in InSAR : A review [Conference Name: IEEE Geoscience and Remote Sensing Magazine]. *IEEE Geoscience and Remote Sensing Magazine*, *7*(1), 40–58. https://doi.org/10.1109/MGRS.2018.2873644

Yun, S.-H., Hudnut, K., Owen, S., Webb, F., Simons, M., Sacco, P., Gurrola, E., Manipon, G., Liang, C., Fielding, E., Milillo, P., Hua, H., & Coletta, A. (2015). Rapid damage mapping for the 2015 M_{w} 7.8 gorkha earthquake using synthetic aperture radar data from COSMO–SkyMed and ALOS-2 satellites. *Seismological Research Letters*, *86*(6), 1549–1556. https://doi.org/10.1785/0220150152

Zebker, H. (2017). User-friendly InSAR data products: Fast and simple timeseries processing. *IEEE Geoscience and Remote Sensing Letters*, *14*(11), 2122–2126. https://doi.org/10.1109/LGRS.2017.2753580

Zebker, H., & Chen, K. (2005). Accurate estimation of correlation in InSAR observations. *IEEE Geoscience and Remote Sensing Letters*, *2*(2), 124–127. https://doi.org/10.1109/LGRS.2004.842375

Zebker, H., & Lu, Y. (1998). Phase unwrapping algorithms for radar interferometry: Residue-cut, least-squares, and synthesis algorithms [Publisher: Optica Publishing Group]. *JOSA A*, *15*(3), 586–598. https://doi.org/10.1364/JOSAA.15.000586

Zebker, H., Rosen, P. A., & Hensley, S. (1997). Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps. *Journal of Geophysical Research: Solid Earth*, *102*, 7547–7563. https://doi.org/10.1029/96JB03804

Zebker, H., & Villasenor, J. (1992). Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(5), 950–959. https://doi.org/10.1109/36.175330

Zhai, G., Shirzaei, M., & Manga, M. (2021). Widespread deep seismicity in the delaware basin, texas, is mainly driven by shallow wastewater injection. *Proceedings of the National Academy of Sciences*, *118*(20), e2102338118. https://doi.org/10.1073/pnas.2102338118

Zheng, Y. (2019). *Imaging cascadia slow slip events with modern interferometric synthetic aperture radar datasets* (Doctoral dissertation). Stanford University.

Zheng, Y., & Zebker, H. (2017). Phase correction of single-look complex radar images for user-friendly efficient interferogram formation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(6), 2694–2701. https://doi.org/10.1109/JSTARS.2017.2697861

Zhu, W., Allison, K. L., Dunham, E. M., & Yang, Y. (2020). Fault valving and pore pressure evolution in simulations of earthquake sequences and aseismic slip. *Nature Communications*, *11*(1), 4833. https://doi.org/10.1038/s41467-020-18598-z

Zwieback, S., Liu, X., Antonova, S., Heim, B., Bartsch, A., Boike, J., & Hajnsek, I. (2016). A statistical test of phase closure to detect influences on DInSAR deformation estimates besides displacements and decorrelation noise: Two case studies in high-latitude regions [Conference Name: IEEE Transactions on Geoscience and Remote Sensing]. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(9), 5588–5601. https://doi.org/10.1109/TGRS.2016.2569435