Title
Physical and Statistical Models in Deformation Geodesy

Permalink
https://escholarship.org/uc/item/4mv4f384

Author
Lipovsky, Brad

Publication Date
2011

Peer reviewed|Thesis/dissertation
Physical and Statistical Models in Deformation Geodesy

A Thesis submitted in partial satisfaction of the requirements for the degree of Master of Science in Geological Sciences by Bradley Paul Lipovsky

August 2011

Thesis Committee:
Dr. Gareth Funning, Chairperson
Dr. James Dieterich
Dr. Keith Richards-Dinger
The Thesis of Bradley Paul Lipovsky is approved:

__________________________________________

__________________________________________

__________________________________________

Committee Chairperson

University of California, Riverside
Acknowledgments

I am genuinely grateful from an want to acknowledge the many types of support that I have received from my friends, family, and colleagues. Hopefully I wont make any big omissions.

This thesis has been possible primarily because Gareth Funning has given me nearly complete freedom to pursue topics that have interested me, and for this I am genuinely appreciative. Interactions with Gareth have inspired many aspects of the work presented in this thesis.

The work presented in this thesis has benefitted tremendously because of guidance from and discussions with Keith Richards-Dinger and Jim Dieterich. Keith’s intuition in physics and geophysics and assistance with technical details have been invaluable. Jim has given me uncanny advice and perspective in many ways.

Michael Floyd has constantly provided a unique and critical point of view, and this work is better because of our discussions.

Thanks to the graduate students in the Department of Earth Science, especially the students in the earthquake group.

Elizabeth, Katherine, and David have been great company for many enlightening discussions and Roland Bürgmann and David Sandwell have been collaborators from whom I have learned a great deal.

I owe many thanks to those who inspired me to pursue geophysics, especially Rick Allmendinger for making it possible for a math major to do geology for a summer.

Many friendships, not excluding people already mentioned, have made life more enjoyable. In particular, Cassaundra has been a true friend.
My family has inspired me, given me unwavering support, and has made everything possible, and for this I am exceptionally grateful.
To my friends and family.
ABSTRACT OF THE THESIS

Physical and Statistical Models in Deformation Geodesy

by

Bradley Paul Lipovsky

Master of Science, Graduate Program in Geological Sciences
University of California, Riverside, August 2011
Dr. Gareth Funning, Chairperson

Geodetic techniques involving spaceborne measurement have revolutionized scientific understanding of Earth surface deformation, and this thesis presents several independent methodological developments concerning such data. The primary results are: a more general method for the optimal design of geodetic networks; an algorithm for detecting transient deformation events in large geodetic datasets; the identification of a previously unmapped fault in the San Francisco Bay Area, California; a new approach to quantifying the permeability structure of shallow, near-surface, strike slip faults; and a regional evaluation of correlated seasonal deformation in the Bay Area. These results are not strictly related, however, several themes are present throughout. Particular emphasis is placed on the development of statistical and numerical methodology. Although the results presented in this thesis are not strictly related, they represent a contribution to our scientific understanding of Earth surface deformation phenomena.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7</td>
<td>Conclusions</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>Strain Transient Detection</td>
<td>45</td>
</tr>
<tr>
<td>4.1</td>
<td>Eigenfunction analysis</td>
<td>46</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Interpretation of statistical eigenfunctions</td>
<td>48</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Separability in space and time</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>Noise modeling</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>Results</td>
<td>54</td>
</tr>
<tr>
<td>4.4</td>
<td>Conclusions</td>
<td>57</td>
</tr>
<tr>
<td>5</td>
<td>Correlated Hydrogeology</td>
<td>59</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>59</td>
</tr>
<tr>
<td>5.2</td>
<td>Ground surface velocities</td>
<td>61</td>
</tr>
<tr>
<td>5.3</td>
<td>Seasonal Deformation</td>
<td>64</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Empirical impulse response estimation</td>
<td>68</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Displacement equation estimation</td>
<td>71</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Model comparison</td>
<td>71</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>73</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Timing of hydrologic deformation</td>
<td>76</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Identification of fault traces</td>
<td>77</td>
</tr>
<tr>
<td>5.5</td>
<td>Conclusions</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>Optimal Design of Geodetic Networks</td>
<td>80</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction and Background</td>
<td>80</td>
</tr>
<tr>
<td>6.2</td>
<td>Model</td>
<td>82</td>
</tr>
<tr>
<td>6.3</td>
<td>Examples</td>
<td>85</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Network expansion at Parkfield, California</td>
<td>85</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Regional scale network design in southern California</td>
<td>89</td>
</tr>
<tr>
<td>6.4</td>
<td>Discussion</td>
<td>94</td>
</tr>
<tr>
<td>7</td>
<td>Synthesis</td>
<td>99</td>
</tr>
<tr>
<td>A</td>
<td>Solid Mechanics Background</td>
<td>102</td>
</tr>
<tr>
<td>A.1</td>
<td>The governing equations of linear elasticity</td>
<td>102</td>
</tr>
<tr>
<td>A.2</td>
<td>The governing equations of poroelasticity</td>
<td>104</td>
</tr>
<tr>
<td>A.2.1</td>
<td>State variables and material parameterization</td>
<td>104</td>
</tr>
<tr>
<td>A.2.2</td>
<td>Constitutive relations</td>
<td>105</td>
</tr>
<tr>
<td>A.3</td>
<td>Uncoupled diffusion</td>
<td>107</td>
</tr>
<tr>
<td>B</td>
<td>Alternative solution to the governing equations of diffusion in composite material</td>
<td>109</td>
</tr>
<tr>
<td>B.1</td>
<td>Solution formalism</td>
<td>111</td>
</tr>
<tr>
<td>B.2</td>
<td>Numerical Calculations</td>
<td>113</td>
</tr>
<tr>
<td>B.3</td>
<td>Solution and error estimation</td>
<td>114</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>C</td>
<td>One dimensional models of aquifer expansion</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>C.1 One dimensional unconfined aquifer</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>C.2 One dimensional confined aquifer</td>
<td>117</td>
</tr>
<tr>
<td>D</td>
<td>Appendices to Correlated Hydrogeology</td>
<td>118</td>
</tr>
<tr>
<td>E</td>
<td>Appendices concerning network design</td>
<td>124</td>
</tr>
<tr>
<td>E.1</td>
<td>Information content of simple models</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>E.1.1 One dimensional gravity anomaly</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>E.1.2 Interseismic deformation</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>E.1.3 Single observations with isotropic noise</td>
<td>128</td>
</tr>
<tr>
<td>E.2</td>
<td>Supplemental figures concerning the RSQSim dataset</td>
<td>132</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Comparison between three different model geometries of subsurface inflation: a plane-strain patch of inflation (solid lines), a spherical pressurized cavity (dots), and a penny shaped crack (dashed lines). The curves that start at the origin are the horizontal component of deformation and the curves that start at unity are the vertical component of deformation. The values of each model’s vertical and horizontal deformations are normalized to the maximum vertical deformation.

2.2 The model geometry of pore pressure diffusion through a fault zone. The model is not drawn to scale. Parameters are discussed in the text.

2.3 Pore pressure distributions at logarithmic time intervals (color scale) for four fault zone - aquifer permeability models. The left boundary condition is constant unit fluid mass.

2.4 Pore pressure distributions at logarithmic time intervals (color scale) for four fault zone - aquifer permeability models. The left boundary condition is constant unit fluid mass.

3.1 Map of the Santa Clara valley.

3.2 PS-InSAR data from the Santa Clara valley.

3.3 Well head levels from the Santa Clara Valley Central Aquifer.

3.4 PS-InSar data (dots) and modeled displacements (solid lines) using a an impermeable fault zone model. Note the variable vertical axis.

3.5 PS-InSar data and modeled displacements using a one layer diffusion model.

3.6 Sum of squared residual contours in parameter space, from the simulated annealing solution, for the one layer diffusion model.

3.7 PS-InSar data and modeled displacements using a two layer diffusion model.

3.8 Sum of squared residual contours in parameter space, from simulating annealing, for the two layer diffusion model.

4.1 Simple eigenfunction analysis example with synthetic GPS data.

4.2 Data showing a visually apparent transient event.

4.3 An example result of an eigenfunction-only transient detection method.
4.4 An example result of our transient detection method. The data used in this analysis is the SCEC Transient Detection Exercise Phase 3F dataset. From the upper right, the subplots show the spectra of the eigenvalues (or, Principal Components); the normalized spectra; the temporal eigenvector; and, the spatial eigenvector. ................................. 56

4.5 Results from our participation in the SCEC transient detection exercise. From the datasets provided we infer that the detection threshold for our method is approximately 2.5cm fault slip. ................................. 58

5.1 Overview map of the San Francisco Bay Area. ................................. 62
5.2 The first three eigenfunctions of the Bay Area PS-InSAR dataset. ................................. 63
5.3 Velocities from the Bay Area PS-InSAR dataset. ................................. 65
5.4 Comparison of the estimated impulse response functions and well level records. 70
5.5 Seasonal deformation indicates the presence of a previously unmapped fault. 78

6.1 Covariance functions discussed in this study. ................................. 86
6.2 The geometry of the geodetic network in Parkfield, California. ................................. 87
6.3 Information criterion versus number of campaign sites. ................................. 92
6.4 Contour plot of information criteria of proposed sites in southern California. 97

D.1 Sum of squared residuals between the detrended PS-InSAR time series and the models discussed in the text. ................................. 119
D.2 Model parameter histograms from the displacement equation estimation, $q = 2.120$
D.3 Model parameter histograms from the displacement equation estimation, $q = 3.121$

E.1 Fault model geometries discussed in the text. ................................. 126
E.2 Simplified example of displacements due to various fault models. ................................. 128
E.3 Information content for the magnitude of slip on a finite fault with various model geometries. ................................. 130
E.4 Information content for the locking depth of a finite fault showing various model geometries. ................................. 131
E.5 Eigenvalues of the data matrix. ................................. 133
E.6 The first four eigenvectors used in this study. ................................. 134
E.7 Example singular vector of the RSQSim simulation. ................................. 135
E.8 Information content in southern California. ................................. 136
## List of Tables

3.1 Table of material parameters used in this study. .......................... 34

6.1 Combinations of the campaign GPS sites that maximize the network’s information content. ................................................................. 90

6.2 Combinations of the campaign GPS sites that minimize the network’s information content. ................................................................. 91
Chapter 1

Introduction

The topics of this thesis are wide ranging yet related. The principal relationship between the topics is an interest in observations that are derived from spaceborne geodetic techniques, particularly from repeat-pass interferometric synthetic aperture radar (InSAR) and the Global Positioning System (GPS). InSAR and GPS have radically altered numerous scientific fields, including tectonics, seismology, Earth structure, meteorology, climatology, hydrology, glaciology, and atmospheric physics (Bürgmann et al., 2000; Blewitt, 2007; Galloway and Hoffmann, 2007; Segall, 2010). Although many Earth system phenomena are now routinely studied using these tools, this thesis attempts to shed new light on several outstanding problems in earth surface geodesy.

What justifies the existence of yet another thesis concerning surface deformation geodesy? To answer this question, we note a theme through this thesis is a focus on the development of new methodologies. Chapter 6 extends existing methods for the optimal design of experiments and applies these to GPS networks in California. Chapter 4 focuses
on the development of a transient detection algorithm, and Chapter 5 is, to the authors’ knowledge, the first use of this method in conjunction with a PS-InSAR data set. Chapter 2 develops a new model of the ground surface deformation induced by fluid flow near permeable fault zones. It is for these reasons that we maintain that this thesis represents a meaningful contribution to the scientific study of ground surface deformation.

1.1 The data

The observational constraints of InSAR and GPS data are unique yet complementary. Synthetic aperture radar (SAR) measures the change in distance between a SAR satellite and a radar reflector on the surface of the Earth, a quantity also called range change. Range change measurements made in the line-of-sight (LOS) of the satellite are inherently a one dimensional observation and have a measurement error on the order of millimeters. Data are acquired at best on a monthly time scale and this varies significantly. The InSAR data discussed in this thesis are processed using the Permanent Scatterer InSAR (PS-InSAR) technique (Ferretti et al., 2000; Colesanti et al., 2003), which tracks consistently radar-bright, phase-stable radar reflectors and can result in meter-scale positional uncertainty in the location of reflectors.

GPS measurements also characterize the distance between an earth orbiting satellite and a point on Earth’s surface, although here the use of a constellation of satellites permits three dimensional position estimates. GPS measurements are also associated with measurement uncertainty on the millimeter scale. Whereas InSAR measurements tend to be dense in space and sparse in time, GPS measurements are dense in time, with temporal
resolution as fine as 1 Hz, and sparse in space. The nominal accuracy of GPS measurements is 1mm horizontal, 3mm vertical, and 0.1mm velocity, averaged over decadal time series (Blewitt, 2007).

Both SAR and GPS observations are made on satellites in Earth orbit and both are therefore heavily reliant on the precise determination of satellite orbits. Scientific progress due to the use of both of these techniques has, in large part, been due improvements in the estimation of these orbits, as well as to an increased abundance of observations, refined global reference frames, and improved data processing tactics (Bürgmann et al., 2000; Altamimi et al., 2007; Blewitt, 2007).

Although this thesis does not extensively discuss the processing or analysis of the raw data from these sources, it should be constantly remembered that the inferred deformations are a sophisticated, derived product that are accompanied by a host of assumptions, constraints, and measurement uncertainties.

The scope of this thesis encompasses four topics in space geodesy: the design of GPS networks, transient detection in GPS data, regional-scale correlations in InSAR data, and poroelastic models of hydrologically-induced ground surface deformation. We now review the principal results of each topic of this document.

1.2 The optimal design of geodetic networks

Given a geodetic network that is intended to observe deformation similar to a preliminary source model, it is natural to ask about the sensitivity of the network to changes in the source model parameters. In Chapter 6, we address this concern by evaluating
the information content of two geodetic networks in California. In order to evaluate the information content in systems of mixed determinacy, we present a generalization of previous work.

We use several numerical experiments to offer several recommendations to practitioners seeking to deploy geodetic networks to observe such deformation. Our results suggest that the existing southern California GPS network would benefit from a greater density of observations in the near fault environment generally, and near the San Jacinto fault in particular. Although logistical and geologic concerns strongly affect station siting, parameter sensitivity offers a reasonable constraint on ideal geodetic observation locations.

The importance of careful design is shown by noting that in the case of the Parkfield, California local GPS network, a 90% reduction in model parameter uncertainty can be made by carefully choosing observation locations from among the numerous possible survey locations.

1.3 Transient detection

Increasing data wealth in the field of space geodesy has inspired the use of automated methods to detect patterns of deformation that may not be apparent through visual inspection. Chapter 4 discusses applies a basic pattern recognition technique to geodetic data in order to find anomalous events. The goal of this research is to identify deformation events that are near the noise level and may not be detectable through alternative geophysical means.

We develop a method for transient detection that relies on a filter-window-eigenfunction
approach. The problem of transient detection is recast as an optimization problem over the space of possible time- and space-windows in the data. In order to quantify the performance of this algorithm, we test our detector with dozens of synthetic datasets in a double-blind setting. We find that our method has a lower detection threshold for fault slip related events of about 2.5cm of fault slip.

1.4 InSAR observations of fluid diffusion through a permeable fault zone

Chapters 2 and 3 concern the development of a model of fluid diffusion through fault zones with a heterogeneous permeability structure. Chapter 2 develops a Laplace-domain semi-analytic method of solution for the relevant coupled poroelastic problem, and Chapter 3 describes the application of this method to InSAR data from the San Francisco Bay Area, California. The models that are presented are applicable to cases of groundwater and hydrocarbon flow and extraction.

The approach taken is to build models of increasing complexity as the data warrant. The first models considered explain why it is that appreciable ground surface deformations arise from confined aquifers but not unconfined aquifers. These models are one dimensional and solutions are presented in closed form.

We then discuss models of reservoirs that are bounded by faults. For the case of perfectly impermeable fault zones, closed form solutions are presented. The case of arbitrary, differing permeabilities in and around the fault zone is explored using a computationally
efficient semi-analytic method.

1.5 Correlated hydrogeological deformation in the San Francisco Bay Area, California

Oscillations of ground subsidence and rebound are ubiquitous features of Earth surface displacement time series. Such variations are due to a superposition of related processes including hydrologic, atmospheric, tidal, and correlated error effects. This complexity inspires the use of low dimensional characterizations to efficiently understand the basic properties of an otherwise intricate system. We present such a model that uses Principal Component Analysis (PCA), another name for the generalized eigenvalue problem. We find that in the San Francisco Bay Area, California, this method provides insights into the mechanisms of hydrogeologic seasonal deformation.

Our analysis suggests a three-component model of seasonal deformation in the Bay Area, whereby three distinct patterns of deformation occur with unique lag times with respect to forcing by precipitation. The first component of deformation occurs immediately after rainfall and consists of spatially extensive deformation. The second component is delayed by \( \sim 60 \) days, and well data support the conclusion that this deformation is due to artesian aquifer flow. The third phase occurs with a \( \sim 150 \) day lag time, and incorporates several smaller-scale processes including subsidence of anthropogenic fill soils and deep landslide creep. Comparison with well data shows that the majority of the seasonal deformation in the San Francisco Bay Area is caused by groundwater flow in the upper 20-30m of the
Earth.
Chapter 2

Analytical solutions for the Earth surface displacements due to fluid migration through a complex permeable barrier

2.1 Introduction and Background

Geodetic data often reveal the presence of groundwater aquifers that are bounded by faults (Schmidt and Bürgmann, 2003; Galloway and Hoffmann, 2007; Bell et al., 2008). Whereas unrestricted groundwater aquifers exhibit a radially symmetric pattern of uplift with diffuse boundaries, aquifers that are bounded by faults have one or more sharp, linear boundaries. Interferometric synthetic aperture (InSAR) data, due to their high spatial
density, are particularly well suited to observe this phenomenon. The goal of this work is to develop a model of the displacements at the surface of the Earth due to fluid diffusion through a permeable boundary such as a fault zone. The companion to this present work is Chapter 3, which presents an analysis of data from the Santa Clara valley using the method developed here.

We seek a simplified model of fault zones in order to model their poromechanical properties. The observations that we seek to describe are,

1. Fault zone architecture consists of a set of approximately fault parallel planar structures. Depending on the nature of the host rock and the history of the fault, these structures may include the damage, mixed-damage, and fault core zones. (Chester et al., 1993; Caine et al., 1996; Rawling et al., 2001; Evans and Bradbury, 2004; Caine and Minor, 2009).

2. The permeability of these layers may vary by more than seven orders of magnitude (e.g., between a clay rich gouge layer and a thickly-bedded gravel host rock).

3. One or more faults with a parallel trend and consisting of similar architectures may reside in close proximity (Jennings, 1994), and act as a system to control the flow of groundwater (Bawden et al., 2001; Lu and Danskin, 2001).

4. Surface displacements due to fluid flow in the near fault environment are often due to forces that act within the first few hundred meters of the surface (Schmidt and Bürgmann, 2003).

The first two points motivate a modeling strategy that can describe multiple scales
of heterogeneity, whether in the form of multiple fault zones, multiple architectural zones within a single fault zone, or both. To meet this requirement, we consider the case of diffusion that occurs in an arbitrarily complex one dimensional medium. The requirement that diffusion be one dimensional can be thought of as diffusion in a reservoir that is confined to a single stratigraphic unit.

The third point, concerning the depth of deformation, is directly relevant to the model formation. Many previous studies have focused on the ground deformation due to fluid mass movement within a reservoir at depth. To the authors’ knowledge, the results of all previous studies focus on systems where the distance between the loci of strain and the free surface is much greater than the length scale of the loci (Segall, 1985; Fialko et al., 2001; Segall, 2010). For the shallow, near-surface reservoir systems of interest here, this assumption is clearly not valid.

2.2 Survey of previous modeling work

Many analytical approaches to diffusion in composite media exist in the literature. Carslaw and Jaeger (1959) advocate a complex variable method that results in the application of the residue theorem. Although problems of interest are formulated within this paradigm, few solutions are available due to the complexity of the ensuing calculations. Mikhailov and Ozisik (1984) and Mulholland and Cobble (1972) develop general solutions by solving a Sturm-Liouville eigenvalue problem. In this approach, eigenvalues are found numerically using the method of Wittrick and Williams (1971) that is familiar in finding vibrational modes of structures. Johnston (1991) and Liu and Ball (1998) develop exact
closed-form solutions to more specialized problems in terms of series expansions. This approach, however, may require sufficiently many terms to be calculated in series expansion that purely numerical methods become computationally more efficient.

Finite difference methods may simulate groundwater barriers either through grid size reduction or by incorporating a notion of infinitely narrow, low permeability barriers with zero storage capacity (Hsieh and Freckleton, 1992; Harbaugh et al., 2000; Hanson et al., 2004). Such models may also be used in the context of parameter estimation (Hill et al., 2000).

We use two criteria to evaluate the usefulness of solution methods. First, solutions should provide direct information concerning the relative importance of the system parameters and their interactions. This is generally not the case for grid-based computation schemes such as the finite element method (Mikhailov and Ozisik, 1984). Second, the method should be as fast as possible. Fast calculations are especially important when a large number part of parameter space is to be explored, as is often the case when conducting statistical data inversions (Fialko et al., 2001).

Two developments are necessary to create a simple model of fault zone mass diffusion, the first being a model of poroelastic diffusion. Although the general governing equations of poroelasticity are bilaterally coupled, we consider a special case of unilateral coupling. Our strategy is to find the surface displacements due to a discretized pore fluid field of cross-sectional patches with prescribed, uniform poromechanical states. For this first problem we derive a simple analytical solution. The second problem is to determine the pore pressure field due to diffusion through layered media of heterogeneous poromechanical
properties. For this second problem we present a semi-analytical method with analytical solutions in the Laplace domain.

2.3 Reservoir discretization

A displacement formulation poroelasticity relates displacements, \( u_i \), and fluid mass change per reference volume, \( \Delta m \), by (Wang, 2000),

\[
G \nabla^2 u_i + \frac{G}{1 - 2\nu_u} \frac{\partial^2 u_k}{\partial x_i \partial x_k} = BK_u \frac{\partial \Delta m}{\partial x_i}.
\]  

(2.1)

In this expression and throughout we neglect body forces. Equation 2.1 introduces the material parameters \( G, \nu_u, B, \) and \( K_u \), which are the shear modulus, the undrained poisson ratio, Skempton’s coefficient, and the undrained bulk modulus, respectively. In the following several details will be omitted. For these, the reader is referred to Appendix A.

A useful interpretation of equation 2.1 is that gradients in fluid mass change act as body forces in the Navier-Cauchy formulation of classical elasticity (Rice and Cleary, 1976; Segall, 1985, 2010). The problem of determining the displacements due to a known fluid mass field is therefore equivalent to the problem of determining the displacements due to a known body force field, provided that this information is available. In general, one would not expect to have prior information about the fluid mass field, and solving for this quantity in full generality can itself be analytically challenging. The foci of this paper, however, are simple cases where one may assume that the fluid mass field is only unilaterally coupled to the stress field, and may therefore be computed independently. In this case, the solution to
Equation 2.1 may be given as a Greens function integral:

$$u_i(x, t) = BK_u \int_V g_i^k(x, \xi) \frac{\partial \Delta m(\xi, t)}{\partial \xi_k} dV_{\xi}, \quad (2.2)$$

Integration by parts and the requirement of zero fluid mass change at great distances results in,

$$u_i(x, t) = BK_u \int_V \Delta m(\xi, t) \frac{\partial g_i^k(x, \xi)}{\partial \xi_k} dV_{\xi}, \quad (2.3)$$

The term $\frac{\partial g_i^k(x, \xi)}{\partial \xi_k}$ represents a center of dilatation; it is the sum of three orthogonal opening mode force couples. The center of dilatation Greens function is given by, (Segall, 2010, correcting the omission of the elastic modulus),

$$\frac{\partial g_i^k(x, \xi)}{\partial \xi_k} = \frac{1}{8\pi G} \left( \frac{1}{1 - \nu} \right) \frac{x_i - \xi_i}{R^3}, \quad (2.4)$$

and $R$ is the euclidean distance from $x$ to $\xi$.

The Cauchy-Navier type equation, Equation 2.1 could have alternatively been written in terms of pore pressure gradients rather than fluid mass change gradients. This form was used, for example, by Jónsson et al. (2003) in their study of coseismic poroelasticity. The pore pressure formulation is given by,

$$G\nabla^2 u_i + \frac{G}{1 - 2\nu} \frac{\partial^2 u_k}{\partial x_i \partial x_k} = \alpha \frac{\partial p}{\partial x_i}, \quad (2.5)$$

and results in the Greens function representation,

$$u_i(x, t) = \alpha \int_V p(\xi, t) \frac{\partial g_i^k(x, \xi)}{\partial \xi_k} dV_{\xi}, \quad (2.6)$$

where we note that this kernel is identical to that used in the fluid mass change formulation.
2.3.1 Surface deformation due to a rectangular patch of uniform pore pressure or fluid mass

Groundwater aquifers have spatially variable pore fluid pressure fields. Such variability is due to aquifer heterogeneity and geometry with respect to the gravitational field, as well as from a non-equilibrium mass balance effects. In this section, we present a method to discretize a pore fluid pressure field and calculate the resulting surface deformations.

We take the basic unit of discretization to be a cross-sectional patch of uniform poromechanical properties and hydraulic state. We consider a uniform cross sectional patch, in plane-strain, with vertical thickness $H$, lateral extent $L$, and depth of burial, $D$. The surface $z = 0$ is assumed to be free of tractions and displacements are assumed to vanish at great depths. One end member case would be a shallowly buried stratigraphic that is spatially extensive compared to its depth of burial, e.g., $H < D \ll L$. We assume that the patch is either uniformly pressured to a pressure $P_0$ or has uniform fluid mass $M_0$. It should emphasized that the diffusion of fluid mass is not accounted for in this part of the model, and so this represents a type of undrained deformation.

We use Equation 2.6 to find the vertical deformation due to a pressurized horizontal layer. For the specified problem geometry, the solution is,

$$u_x(x) = \gamma \left\{ (L + 2x) \tan^{-1} \left( \frac{2D}{L + 2x} \right) + D \log \left[ D^2 + \frac{1}{4}(L + 2x)^2 \right] \right\}^*$$

$$u_z(x) = \gamma \left\{ 4D \tan^{-1} \left( \frac{L - 2x}{2D} \right) + (L + 2x) \log \left[ L^2 + 4Lx + 4(D^2 + x^2) \right] \right\}^*$$

(2.7)

Here, the notation $|^*$ is used to represent the substitution,

$$f(x, D)^* = f(x, D + H) + f(-x, D + H) - f(x, D) - f(-x, D)$$

(2.8)
The coefficient $\gamma$ is determined by the condition of either uniform fluid mass or uniform pore pressure,

$$\gamma = \begin{cases} 
P_0 \frac{3(\nu_0 - \nu)}{2\pi BG(1+\nu_0)} & \text{uniform pore pressure}, \\
M_0 \frac{2B(1+\nu_0)}{3\pi \rho_0} & \text{uniform fluid mass}.
\end{cases} \quad (2.9)$$

These expressions have been multiplied by a factor of $4(1 - \nu_r)$ to account for the free surface correction (Davies, 2003). The relevant Poisson’s ratio, $\nu_r$, is Poisson’s ratio for the pore pressure case and Poisson’s ratio under undrained conditions for the fluid mass case (Wang, 2000, p. 75). General poroelastic deformations are characterized by four material parameters, and volumetric sources by three (Wang, 2000). The first expression for $\gamma$ could therefore be written using fewer parameters, however, we choose this form for its relative simplicity.

Using this simple model, any arbitrarily complex pore fluid field may be discretized. In what follows, we couple this result with a fluid mass field determined by solving a mass diffusion problem. The result of Equation 2.7 could also be used to determine pore pressure distributions from surface displacements in the same way that the results of Okada (1985a) are used to statistically invert surface displacements for a distribution of fault slip. This application is, however, beyond the scope of the present work.

Figure 2.1 compares this solution to the deformations from a penny shaped crack and a pressurized spherical cavity, or Anderson-Mogi model (Fialko et al., 2001; Segall, 2010). The deformations due to the pressurized patch are geometrically distinct from the deformations due to the crack and cavity models. Only the patch model has a concave downward profile of vertical deformations, and both the horizontal and vertical deformations...
Figure 2.1: Comparison between three different model geometries of subsurface inflation: a plane-strain patch of inflation (solid lines), a spherical pressurized cavity (dots), and a penny shaped crack (dashed lines). The curves that start at the origin are the horizontal component of deformation and the curves that start at unity are the vertical component of deformation. The values of each model’s vertical and horizontal deformations are normalized to the maximum vertical deformation.
are more spatially extensive than the crack and cavity models. These models are not directly comparable, however, because the crack and cavity models are assumed to be buried to a depth much greater than their width (e.g., $D \gg L$). Additionally, the cavity model is asymptotically equivalent to a point source. Smaller aspect ratios for these models may result in a poor approximation to the boundary conditions along the surface of the crack or cavity (McTigue, 1987; Fialko et al., 2001).

### 2.4 Poroelastic diffusion of fluids through permeable barriers

We now focus on the particular case of time variable fluid mass change due to diffusion through a vertical barrier. In order to meet the requirements presented in the introduction, we formulate our solution in terms of an arbitrary number of vertical barriers. This will allow us to approximate a fault zone with any number of significant architectural components, as well as to consider the case of multiple parallel faults.

We consider a set of $n$ parallel, vertical layers, each with thickness $T_i$ and hydraulic diffusivity $\alpha_i$. We consider diffusion to occur as a one dimensional process, confined to a single stratigraphic layer in cross section. The assumption of one-dimensionality is describes, for example, diffusion between parallel impermeable layers that are much greater in extent than in their separation. In general we will assume that the fluid mass content or pore pressure is known as a function of time along one edge of the model, which we will take to be the origin. This model schematic is shown in Figure 2.2 for the case $n = 2$. 
The general governing equations for the $n$-layer problem is given by,

\[
\left( \alpha \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} \right) \Delta m = 0
\]

\[
\left( \alpha \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} \right) \Delta m_1 = 0
\]

\[
\left( \alpha_n \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} \right) \Delta m_n = 0
\]  

(2.10)

The general boundary conditions are given by,

\[
\Delta m_1(0,t) = \Delta m_0(t)
\]

\[
\Delta m_i(T_i, t) = \Delta m_{i+1}(T_i, t)
\]

\[
\frac{\partial}{\partial x} \Delta m_1(T_i, t) = \frac{\partial}{\partial x} \Delta m_2(T_{i+1}, t)
\]

\[
\Delta m_2(\infty, t) = 0
\]  

(2.11)

For the case $n = 1$, the analytical solution is given by the convolution integral

\[
\Delta m(x, t) = \int_0^t f(x, t - \tau) \Delta m_0(t) \, d\tau,
\]  

(2.12)

where the convolution kernel is given by,

\[
f(x, t - s) = \frac{x}{\sqrt{4\pi \alpha (t - \tau)^3}} \exp \left[ -\frac{x^2}{4\alpha (t - \tau)} \right]
\]  

(2.13)

For the case of a constant boundary condition, $\Delta m_0(t) = M_0$, the solution is given by,

\[
\Delta m(x, t) = M_0 \text{erfc} \left( \frac{x}{2\sqrt{\alpha t}} \right),
\]  

(2.14)

where erfc is the complimentary error function.

For more complex boundary conditions, Equation 2.12 may be solved analytically in some cases or by numerical quadrature otherwise. Since the latter may be computationally expensive, and since the former is not always available, we consider consider the
Figure 2.2: The model geometry of pore pressure diffusion through a fault zone. The model is not drawn to scale. Parameters are discussed in the text.
Laplace domain representation of the diffusion equation. The solution in the Laplace domain of Equation 2.12 is,

\[
\Delta \tilde{m}(x, t) = \Delta \tilde{m}_0(s) \exp\left(-\sqrt{\frac{s}{\alpha}}x\right)
\]  

(2.15)

In these expressions, the Laplace frequency variable is \(s\), \(\Delta \tilde{m}_0(s)\) is the Laplace transform of the boundary condition, and \(\Delta m(x, t)\) is the Laplace transform of the change in fluid mass.

We use the solution to verify our numerical implementation. Figure 2.3 shows benchmarking experiments for two values of numerical tolerance in the numerical Laplace transform inversion algorithm of de Hoog et al. (1982) implemented by Hollenbeck (1998). We find that the nominal error is globally maintained for tolerance of \(\epsilon = 10^{-5}\) (Figure 2.3a). Smaller values of \(\epsilon\) than this may contain local departures from the nominal tolerance (Figure 2.3b).

When using this solution, the pore pressure or fluid mass boundary condition may be derived from observational well data. When this is the case, values of the well data will be known at a discrete number of points. It will therefore be useful to have a representation for data in this form. We suppose that the data are available at times \(t_1, t_2, \ldots, t_n\), and that at these times the boundary condition takes the values \(a_1, a_2, \ldots, a_n\). Then a stepwise continuous representation of the discrete data is,

\[
\Delta m_0(t) = a_1 H(t - t_1) + \sum_{i=2}^{n} (a_{i+1} - a_i) H(t - t_i)
\]  

(2.16)

The Laplace transformed boundary condition function then becomes,

\[
\Delta \tilde{m}_0(s) = a_1 \frac{\exp(-t_1 s)}{s} + \sum_{i=1}^{n} (a_{i+1} - a_i) \frac{\exp(-t_i s)}{s}
\]  

(2.17)
2.4.1 Solution for the case \( n = 2 \)

For the particular case \( n = 2 \), the mass diffusion equations are given by,

\[
\left( \alpha_1 \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} \right) \Delta m_1 = 0 \\
\left( \alpha_2 \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} \right) \Delta m_2 = 0
\]

(2.18)

The boundary conditions are,

\[
\Delta m_1(0, t) = \Delta m_0(t) \\
\Delta m_1(L, t) = \Delta m_2(L, t) \\
\frac{\partial}{\partial x} \Delta m_1(L, t) = \frac{\partial}{\partial x} \Delta m_2(L, t) \\
\Delta m_2(\infty, t) = 0
\]

(2.19)

We take the temporal Laplace transform of these conditions, and then solve resulting system of ordinary differential equations for \( \tilde{\Delta} m_1 \) and \( \tilde{\Delta} m_2 \), the Laplace transformed fluid mass variables. The resulting expressions are,

\[
\Delta m_1 = \tilde{\Delta} m_0(0, s) \frac{\sqrt{\alpha_2} \cosh \left[ (L - x) \sqrt{\frac{s}{\alpha_1}} \right] + \sqrt{\alpha_1} \sinh \left[ (L - x) \sqrt{\frac{s}{\alpha_1}} \right]}{\sqrt{\alpha_2} \cosh \left[ L \sqrt{\frac{s}{\alpha_1}} \right] + \sqrt{\alpha_1} \sinh \left[ L \sqrt{\frac{s}{\alpha_1}} \right]}
\]

(2.20)

\[
\Delta m_2 = 2 \tilde{\Delta} m_0(0, s) \frac{\sqrt{\alpha_2} \exp \left[ \sqrt{\frac{s}{\alpha_1 \alpha_2}} \left( -x \sqrt{\alpha_1} + L(\sqrt{\alpha_1} + \sqrt{\alpha_2}) \right) \right]}{\sqrt{\alpha_1} \exp \left( 2L \sqrt{\frac{s}{\alpha_1}} \right) - 1 + \sqrt{\alpha_2} \exp \left( 2L \sqrt{\frac{s}{\alpha_1}} \right) + 1}
\]

(2.21)

Numerical inversion of Equations 2.20 and 2.21 are shown in Figure 2.4 for four parameter combinations. The boundary condition is the unit step function, zero before \( t = 0 \) and unity after. In this figure, time is represented by the colorscale and is taken
Figure 2.3: Pore pressure distributions at logarithmic time intervals (color scale) for four fault zone - aquifer permeability models. The left boundary condition is constant unit fluid mass.
in logarithmic steps. This figure shows that for plausible combinations of a permeable or impermeable fault and a permeable or impermeable aquifer, drastically different fluid mass distributions result after one year. For the case of a permeable fault and a permeable aquifer, a nearly uniform steady state, about 95% of the boundary condition, is reached at distances of up to 250m from the fault. In contrast, the case of an impermeable fault and an impermeable aquifer shows fluid mass achieving a value of less than 5% of the boundary condition within this time.

2.5 Conclusions

We obtain a closed form, analytical solution for the surface deformation due to a cross sectional, plain strain, patch with a uniform poromechanical state. We then couple this solution to models of mass diffusion through arbitrarily complex flow barriers. Whereas much previous work has focused on deformations due to deep magmatic processes, the solutions presented here are unique because they explore the near surface regime. The deformations from a confined aquifer are compared to models a uniformly pressured, spherically symmetric cavity, and a uniformly pressurized, penny shaped crack. The patterns of surface deformation due to the patch-model are shown to be highly distinguishable from these other models, especially for large width to depth ratios. This model is relevant to studies of shallow groundwater and hydrocarbon flow. Because of the strong effect of fault zone permeability on surface displacements, inversion of geodetic data to infer these parameters could replace costly ground based aquifer test methods.
Figure 2.4: Pore pressure distributions at logarithmic time intervals (color scale) for four fault zone - aquifer permeability models. The left boundary condition is constant unit fluid mass.
Chapter 3

InSAR observations of diffusion through a permeable fault zone in the Santa Clara Valley, CA

3.1 Introduction

Faults are fundamental controls of fluid flow in the crust, and numerous interferometric synthetic aperture radar (InSAR) studies have imaged faults acting as barriers to groundwater flow (Amelung et al., 1999; Lu and Danskin, 2001; Valentine et al., 2001; Bawden et al., 2001; Colesanti et al., 2003; Schmidt and Bürgmann, 2003; Bell et al., 2008). Despite the frequent observation of this phenomenon, it appears that the estimation of fault permeability from geodetic observation of ground surface deformation has not yet been attempted. The scope of the present work is to examine this inverse problem. This problem
is both intriguing in itself and also important for the study of groundwater extraction management, groundwater contaminant tracking and hydrocarbon exploitation. Because fault permeability structure may contain clues concerning the last activity of a fault and its slip history, this study may additionally inform seismic hazard analysis.

The Silver Creek fault in the Santa Clara valley, California, is an excellent natural laboratory to study the diffusion of groundwater through a permeable fault zone. The largest ground surface displacements in the Bay Area occur in this region are due to the inflation of the Santa Clara aquifer (Bürgmann et al., March, 2006), and InSAR data plainly show that the Santa Clara aquifer is partitioned by the Silver Creek fault (Schmidt and Bürgmann, 2003).

To model the diffusion of groundwater in the Santa Clara Basin, we will use the formalism for the poroelastic diffusion of fluids through permeable boundaries in the shallow near surface that was developed in Chapter 2. This method analytically solves a one dimensional diffusion equation for a heterogeneous medium and then computes the resulting surface deformation.

3.2 Hydrogeological Context

The Santa Clara valley is situated between the Santa Cruz mountains to the southwest and the Diablo Range to the northeast, both margins being fault bounded. The northwestern end of the valley is the southern tip of San Francisco Bay, and we consider the southeastern end of the valley to be near the Coyote Narrows. The valley is underlain by the Santa Clara basin and aquifer which are composed of Pleistocene marine and fluvial
Figure 3.1: Overview map of the Santa Clara Valley showing major geographic and tectonic features. The red circles shows the location of the well used to infer fluid mass fluxes and the small green squares show the location of the USGS multi-well sites CCOC (west) and EVGR (east).
deposits that overlie Pleistocene and Pliocene fluvial deposits (Iwamura, 1995; Schmidt and Bürgmann, 2003; Newhouse et al., 2004).

In this study we consider the south central and south eastern portions of the Santa Clara aquifer, or simply the “Central Aquifer” and the “East Aquifer”. Newhouse et al. (2004) have determined from a generalized lithological characterization of multiple well logs that the Santa Clara aquifer consists of four to six coarse grained aquifer units in the top 260 feet of the near surface. The East Aquifer is best characterized as unconfined, although confined layers do exist at depth, and the Central Aquifer is dominated by confined aquifer units.

The total thickness of the aquifer units ranges from 50 to 200 feet in the Central Aquifer and from 10 to 25 feet in the Eastern Aquifer. Thermal gradient profiles show that the majority of lateral flow occurs shallower than 775 feet in the central aquifer and shallower than 510 feet in the eastern aquifer. Additionally, Newhouse et al. (2004) summarizes rock physical parameters of well core samples that we will describe later.

The aquifers of the Santa Clara Basin have experienced intense groundwater extraction since the beginning of the twentieth century. Groundwater extraction continues today, although much of the subsidence that accompanied earlier development has been mitigated through more careful management (Poland, 1984; Galloway et al.; Schmidt and Bürgmann, 2003).

The Silver Creek fault divides the central aquifer from the eastern aquifer, and these two zones have distinctly different hydrogeological properties. Fracture density increases near faults in lithified rock (Caine et al., 1996); unlithified host rock, in contrast,
does not experience a greater fracture density from faulting and instead suffers a loss of pore space (Caine and Minor, 2009). One might expect that this reduced porosity would result in lower permeability and therefore diminished flow, and indeed, this has been observed in field studies (Rawling et al., 2001; Caine and Minor, 2009). In situ studies have characterized the permeability structure of lithified rocks (Chester et al., 1993; Caine et al., 1996) and unconsolidated sediments (Caine and Minor, 2009), usually for normal faults. To date, no studies have attempted to use surface deformation measurements to estimate fault zone permeability. The present work is an investigation of this possibility.

3.3 InSAR Data

InSAR data are particularly germane to the problem of observing the spatial patterns of seasonal and long term aquifer deformation (Galloway and Hoffmann, 2007). We analyze InSAR data from the ERS-1 and ERS-2 satellites’ descending track 70 during the years from 1992 to 2001. Interferometric SAR time series sample Earth surface deformation with millimetric precision, monthly temporal sampling, and $\geq 10$ m spatial resolution (Mas-sonnet and Feigl, 1998; Bürgmann et al., 2000; Rosen et al., 2000). The data are processed using the PS-InSAR technique (Ferretti et al., 2000; Hanssen, 2001; Colesanti et al., 2003). In this approach persistently radar-bright, phase-stable natural and artificial reflectors (e.g., buildings, roads) are identified. The motion of these “scatterers” are then estimated simultaneously with atmospheric contribution, and according to a prescribed functional form. The final data set consists of 40-point time-series for the ground displacements of 178,275 reflectors. For a more detailed discussion of the data processing, the interested reader is
referred to (Colesanti et al., 2003).

The PS-InSAR data are shown in Figure 3.2. We consider only the data from the year 1999 because this time interval has a dense sampling of about one acquisition per month. As preliminary postprocessing we have removed the trend of each time series, which should have the effect of removing the interseismic and long term subsidence signals. We detrended the data over the nine year period, so it is likely that semi decadal subsidence signals remain in the data. Nonetheless, visual inspection of the data show that the dominant signal is the annual inflation of the Santa Clara aquifer. Uplift of the central aquifer reaches a maximum in March and a minimum in September.

For our analysis of the InSAR data, we consider a two dimensional cross sectional profile across the Silver Creek fault. The InSAR data are projected onto a plane that is perpendicular to the surface of the Earth and approximately perpendicular to the Silver Creek fault. The location of the profiles are shown in Figure 5.1. To improve the speed of our numerical calculations, the data are collected into 100m bins and averaged. The value of each bin is then smoothed with a lowpass filter that has a filter constant equal to the width of several bins.

In the following, all uplift profiles are converted to range change in the direction of the ERS-1 and ERS-2 satellites using a look angle of 20°.

3.4 Well Data

Well core data are available from several scientific multiple-well sites in the Santa Clara valley (Newhouse et al., 2004). We use this data to constrain the material parameters
Figure 3.2: PS-InSAR data from eleven epochs during 1999. The data show clearly show the seasonal inflation and deflation of the Santa Clara aquifer. The eastern boundary of the aquifer, the Silver Creek fault, is clearly seen as bounding the zone of inflation.
used in modeling InSAR surface displacements.

Fluid mass is a natural quantity to consider in the equations of poroelastic diffusion 
(Rice and Cleary, 1976; Segall, 1985), yet is not a quantity that is directly observable. In 
general, a change in hydraulic head measured at a well is not equal to the change in mass of 
the water displaced by the head level change. Rather, it has long been appreciated that the 
elasticity of the aquifer must be accounted for in order to accurately estimate the aquifer’s 
water mass flux (Meinzer and Hard, 1925). A first order approximation to this relationship 
is to consider the fluid mass flux per unit aquifer to be proportional to the differential head 
change. The constant of proportionality is given by the elastic specific storage coefficient, 

\[ S_s = \frac{1}{V_a} \frac{dV_w}{dh} \] 

(3.1)

where \( h \) is the hydraulic head. Values of specific storage will be used in our model to 
estimate the fluid mass content along the Silver Creek fault in the Santa Clara aquifer. 
The scientific multi-well site CCOC is located in the central aquifer and has a measured 
elastic specific storage as high as \( 4.9 \times 10^{-5} \text{m}^{-1} \) and inelastic specific storage as high as 
\( 8.9 \times 10^{-4} \text{m}^{-1} \). The site EVGR, located in the eastern aquifer, has measured elastic and 
inelastic specific storage values of \( 6.9 \times 10^{-5} \text{m}^{-1} \) and \( 3.6 \times 10^{-4} \text{m}^{-1} \). These values are at 
least an order of magnitude smaller than the value of 0.007 used by Hanson et al. (2004) in 
their large scale finite difference modeling of groundwater flow throughout the entire basin.

The bulk modulus under drained conditions may be estimated the well data. Although the bulk modulus isn’t directly used to calculate surface deformations, it can be 
used to assess the applicability of published values for other material parameters of interest. 
Assuming that Poisson’s ratio under drained conditions is in the range \( \nu = 0.2 \) and
$\nu = 0.3$, the mean core density and dilatational wave speed (Table 3.1) suggest that the
bulk modulus lies in the range 4.67 GPa and 3.5 GPa.

Skempton’s coefficient, $B$, is not usually directly observed. Skempton’s coefficient
is defined as the change in pore pressure per change in volumetric stress under undrained
conditions. For fluid saturated soils where the bulk modulus of the liquid phase is much
greater than that of the solid phase, one would expect $B = 1$. Wang (2000) reports values
of $B$ as large as 0.99 for an unnamed clay with bulk modulus $K = 0.062$ GPa. Also reported
is a value of $B = 0.83$ for an unnamed mudstone ($K = 2.13$ GPa) and $B = 0.5$ for the Boise
Sandstone ($K = 4.6$ GPa). Under the rationale that the strength of the shallow aquifer
rocks of interest is dominated by the fluid phase, we take $B = 0.9$.

The best information to constrain the depth of the deformation comes from the
multiple well site EVGR. Although the multiple well site CCOC (Figure 5.1) is closer to
the region of interest, since it is on the western side of the Silver Creek fault it is unlikely to
correspond to the juxtaposed lithology on the eastern side. The well core data from EVGR
show a gravelly-sand unit that is likely to represent the source of deformation. This unit
lies between the depths of 45m and 105m. Although temperature gradient measurements
in this well suggest multiple localized zones of flow, for simplicity we begin our modeling
with a simple homogeneous unit.

A comparison of available material parameter data, values used in previous studies,
and values used in this study are shown in Table 3.1.
Figure 3.3: Well head level data (red dots), interpolation between these data (blue dots) and their differential (black line, the latter magnified by a factor of ten) during the year 1999. Measurements are shown with red dots and the overlapping blue circles show interpolation between the measurements. The lower part of the figure marks the dates of SAR scene acquisitions.

<table>
<thead>
<tr>
<th>EVGR Mean</th>
<th>Basin Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>2.23</td>
</tr>
<tr>
<td>$\nu_p$</td>
<td>1.774</td>
</tr>
<tr>
<td>M</td>
<td>(7.0)</td>
</tr>
<tr>
<td>$S_{Se}$</td>
<td>$3.9 \times 10^{-5}$</td>
</tr>
<tr>
<td>$S_{Si}$</td>
<td>$4.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>Inflow</td>
<td>253,233,800</td>
</tr>
<tr>
<td>Recharge</td>
<td>193,779,997</td>
</tr>
</tbody>
</table>

Table 3.1: Table of material parameters used in this study.
3.5 Analysis of ground surface displacement data

We consider a family of models of increasing complexity, beginning with an analytical model of the surface deformations due to a uniformly pressurized aquifer bounded by a perfectly impermeable fault. The analytical expressions for this model are presented in Chapter 2.

The strategy for all subsequent models will be to compute the distribution of fluid mass in the Eastern Aquifer, given a known pore pressure distribution in the Central Aquifer. The fluid mass boundary condition is determined from the record of head level changes (Figure 3.3) and Equation 3.1. Our models all consider a two year record of well hydraulic head changes in order to avoid effects of being near the edge of the temporal domain.

Figure 3.4 shows the fit to the data of the impermeable boundary model. This fit is poor and the failings of this model are worth enumerating. Primarily, the model fails to describe deformations away from the fault. The model predicts a very rapid decay of surface displacement with distance across the fault whereas the observed displacements are gradual. Additionally, the model has an instantaneous response to changes in aquifer pressures, whereas the data show time-delay effects with respect to changes in head level (e.g., compare Figures 3.2 and 3.3).

This result also shows that although the Eastern Aquifer is typically considered to be unconfined, the observed deformation cannot be explained by a model of no pore-pressure induced surface displacements in that region. One explanation is that although the Eastern Aquifer is essentially unconfined for groundwater resource purposes, confined
layers exist at depth.

An alternative hypothesis is that both the time behavior and the geometry constraints may be explained by a model of a permeable fault boundary, and we now test this hypothesis.

We now consider models of groundwater diffusion through a permeable Silver Creek Fault. In this and the following models we introduce an offset parameter to affect an arbitrary shift in the data coordinate system. This shift ensures that the data and the model are compared in similar coordinate systems, and since the shift is constant it introduces no strains and therefore no hydromechanical changes. To find preferred models we use a simulating annealing method in order to conduct a global parameter search. Models are compared to the data using an $L^2$ norm.

We consider the simplest case of one dimensional diffusion in a homogeneous medium. This model would be a good approximation if, for example, the effect of fault related juxtaposition of materials with different poromechanical properties is the dominant regulator of cross-fault flow. The differing hydrogeological properties of the central and eastern aquifers (Newhouse et al., 2004) make such a situation plausible. This model has a single adjustable parameter, the permeability of the homogeneous medium, $\alpha$.

Figure 3.5 compares the data to a best fitting homogeneous model. The single best fitting model has a permeability of $\log(\alpha) = 0.5$, and the family of best fitting permeabilities all lie in the range $-0.5 < \log(\alpha) < 0.5$ (Figure 3.6). These values are in reasonable agreement with measured values from ground based studies (Newhouse et al., 2004). Although the permeable model does fit the data better than the impermeable fault model, several
Figure 3.4: PS-InSar data (dots) and modeled displacements (solid lines) using an impermeable fault zone model. Note the variable vertical axis.
prominent features of the data remain neglected by the model. Particularly, the model fares poorly in describing near-fault deformations.

We next consider a two-layer model with permeability $\alpha_1$ in a unit-thickness fault zone at $x = 0$ and permeability $\alpha_2$ in the remainder of the half-line. Figure 3.7 compares the preferred model to the data, and Figure 3.8 shows a subplane of the parameter space colored by the $L^2$ misfit function. The model provides a generally better fit to the data and introduces only one additional parameter compared to the one-layer model.

The family of preferred two-layer models has a fault zone permeability in the range $-3 < \log(\alpha_1) < -1$. These values, although quite variable, are in rough agreement with the value of $10^{-3}$ used in the study by Hanson et al. (2004) that determined fault permeability from well tests. Eastern Aquifer permeabilities lie in the range $-2 < \log(\alpha_2) < 0.25$, also in rough agreement with previous studies, and also in agreement with our results for a one-layer model.

The inversion results for the two-layer model generally show poor resolution in the fault zone permeability parameter. This is characterized by the horizontal, elongate regions of good fit in Figure 3.8. Although the model also contains a parameter for fault zone width, we fix this value at unity due to its significant trade-off against fault zone permeability.

3.6 Discussion

The uncertainty of our model parameter choices directly affects the uncertainty in our deformation calculations. We note that skempton’s coefficient, the specific storage coefficient, and $(1 - \nu_u)$ appear as scalar multiples in the Green’s function calculations. A
Figure 3.5: PS-InSar data (dots) and modeled displacements (solid lines) using a one layer diffusion model. Note the variation in the y-axis.
Figure 3.6: Sum of squared residual contours in parameter space, from simulating annealing, for the two layer diffusion model. Color scale shows sum of square residuals in (mm$^2$).
Figure 3.7: PS-InSar data (dots) and modeled displacements (solid lines) using a two layer diffusion model. Note the variation in the vertical axis scale.
Figure 3.8: Sum of squared residual contours in parameter space, from simulating annealing, for the two layer diffusion model. Color scale shows sum of square residuals in (mm²).
10% uncertainty in these quantities individually therefore corresponds directly to a 10% uncertainty in our final solution.

While evaluating the uncertainty in our solution due to uncertainties in model parameters is a straightforward task, evaluating the effect of our assumptions is a more formidable task. We assume that flow is two dimension when in reality groundwater flow occurs in three dimensions. We assume that the elastic properties within the reservoir rock are the same as those of the host rock, and we assume that these two units are perfectly welded together.

This study estimates fault zone permeability from a larger scale than previous studies. Evans et al. (1997) notes that the ∼cm scale of in situ measurements of fault zone permeability results in lower bound estimates. Rawling et al. (2001) suggests that, due to the lack of open fractures in unconsolidated material, the in situ estimate of permeability is likely to be in closer accord to the effective value over large length scales.

The data include a region with a small bend in the Silver Creek fault (Figure 5.1). The InSAR data do not show a marked increase in fluid diffusion at this bend, which suggests that the fault is continuous through this feature, rather than having a step over.

3.7 Conclusions

We have presented a method of estimating the permeability of faults that bisect aquifers. Estimates of fault zone permeability are in rough agreement with published figures, although poor parameter resolution limits careful comparison. Methodological developments, including more realistic simulation of groundwater sources and sinks should
enable more precise estimation of fault permeability parameters.
Chapter 4

Strain Transient Detection

Geodetic techniques involving spaceborne measurement have revolutionized scientific understanding of Earth surface deformation. Increasing data wealth in this field has inspired the use of automated methods to detect patterns of deformation that may not be immediately apparent through visual inspection. Here we present an algorithm to detect anomalous transient deformation events. Our challenge for creating this transient detector is to detect events that may be some combination of near the noise level of the data, undetectable through other geophysical means, or previously unnoticed.

We use a filter-window-eigenvalue method to detect transient deformation events. We first use a digital filter to attenuate certain regions of the noise spectrum and to interpolate over missing data. We then determine the eigenfunctions of the data matrix (Joliffe, 1986). The eigenfunction transformation allows us to find patterns that are highly correlated and are separable in space and time. We find that this latter criterion is often indicative of transient deformation. To pursue such results, we will derive a method to search for
events that satisfy the separability requirement. Whereas earlier efforts focused exclusively on eigenfunction analysis, we find that the combination of spectral filters, data windowing, and eigenfunction analysis greatly improves the success of our transient detection algorithm.

The present study has relied on testing transient detection algorithms on synthetic datasets provided by the Southern California Earthquake Center (SCEC) Transient Detection Exercise. The Transient Detection Exercise has provided teams with dozens of datasets with varying degrees of complexity that may or may not simulate a transient crustal deformation event. The datasets are provided in a double-blind setting, and because of this we have been able to test our detection algorithms in a manner more similar to the handling of real data. The Transient Detection Exercise datasets have been released in four phases, each of which has consisted of multiple datasets. Thus, when we refer to dataset “3F”, we are referring to one of the data sets released during Phase 3. With each successive phase the datasets became increasing complex and realistic, although the ordering of the data sets within a single phase was random.

4.1 Eigenfunction analysis

In this section we review several properties of the generalized eigenvalue transformation. The generalization of the eigen-decomposition to the case of nonsquare matrices is variously called Principal Component Analysis (PCA), the Principal Orthogonal Decomposition, Empirical Orthonormal Functions, the Singular Value Decomposition, the Proper Orthogonal Decomposition, the Hotelling Transform, the Karhunen-Loeve Transform, and probably other names, as well. We call it the generalized eigenvalue transformation because
this name seems to introduce the fewest new words.

We begin by forming a data matrix, $X$, from column vector time series of GPS data. We assume that there are $n$ epochs of data and there are $m$ observations. The eigenvalue transformation will provide a more useful basis for the vector space represented by the columns of $X$. The new basis will compress the data such that the first coordinate axis will represent the greatest amount of data variance, the second coordinate axis the second greatest amount of the variance, and so on. The new basis will also be orthonormal, will be computationally efficient to compute, and will have useful relationships with least squares fitting.

To $X$ we apply the generalized eigen-decomposition,

$$X = EC^T.$$  \hspace{1cm} (4.1)

The existence of this decomposition is guaranteed by the spectral theorem (Halmos, 1963), the generalized eigen-decomposition. Uniqueness is ensured by requiring $C$ and $E$ to be orthonormal. The matrices $E$, $A$, and $C$ will have dimensions $m$-by-$m$, $m$-by-$s$, and $s$-by-$s$, and ranks of $m$, $n$, and $s$, respectively. $A$ has a $m$-by-$m$ diagonal matrix in the upper left block and has zero-valued entries elsewhere; the diagonal entries of upper left block of $A$ are called the singular values (SVs) of $X$. For the remainder of this work we assume that the rows of $E$, $A$, and $C$ are arranged such that the terms on the principal diagonal of $A$, $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ monotonically decrease down and to the right ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$, all $\lambda \geq 0$). The column-space of $E$ is an orthonormal basis of parameter space. Each column of $E$ is therefore itself a parameter set.

The following useful relationship with the sum of square residuals is notable. We
define the matrix $\mathbf{A}_q$ by truncating the matrix $\mathbf{A}$ to its first $q$ columns. Then the eigenvalue transformation has the useful property that the matrix, $\mathbf{X}_q$, defined by,

$$\mathbf{X}(q) = x_{ij}(q) = E_q A_q C_q, \quad (4.2)$$

has a concise relationship to the least-squares misfit to the original matrix $\mathbf{X}$. Particularly,

$$\sum \sum (x_{ij} - x_{ij}(q))^2 = \sum_{k=q+1}^{n-q} \lambda_k^2. \quad (4.3)$$

The sum of square residuals (SSR) reconstruction error is equal to the sum square value of the eigenvalues of the truncated eigenfunctions.

4.1.1 Interpretation of statistical eigenfunctions

The eigenfunctions can be considered a transformation of time-dependent data to a set of representative patterns of observations. Each representative observation has variable gain through time: at some observation times a pattern may not be present in the data at all, while at other times it may dominate. The data set at any epoch can be reconstructed by summing the product of each pattern and the corresponding gain for that pattern, at that time. In this way the entire data set may be reconstructed. It is often the case, however, that only the first few patterns represent the great majority of the salient features present in the data.

When analyzing eigenfunctions, it will be important to note that each eigenfunction does not necessarily represent a single physical process. Any particular eigenfunction should be interpreted as a building block of many physical processes because, in this representation, the $e_i$ form a basis for source model vector space. In this context, eigenfunction
analysis is an entirely statistical technique and so shared features of many different physical processes often become grouped together in a single pattern. This is both the benefit and limitation of the method: significant data compression is achieved by only considering a small subset of parameter space, yet the patterns represented in each eigenfunction may not be in one-to-one correspondence with any set of physical processes.

4.1.2 Separability in space and time

The eigenfunction decomposition is particularly well suited for functions that are linearly separable in space and time. The transformation of a data matrix, \( D \), containing geodetic time series in its rows, is given by, \( D = \sum \lambda_i X_i(x) T_i(t) \). Here, the \( X_i(x) \) and \( T_i(t) \) are functions of space and time, respectively, and the \( \lambda_i \) are the singular values. Space-time separability is therefore achieved when \( \lambda_1 \gg \lambda_i, i = 2, \ldots, n \). We will refer to this condition as the eigenvalue criterion. The search for transient deformation events may therefore be restated as the problem of finding a space-time window of data over which the eigenvalue criterion is extremal.

We consider a transient event to be defined by a transient centroid, defined to be a latitude (\( \phi \)) and longitude (\( \psi \)) range and a time period during which deformation has occurred, e.g., (\( \phi_1, \phi_2, \psi_1, \psi_2, \tau_1, \tau_2 \)). The search for transient deformation events is therefore equivalent to the optimization problem of maximizing the eigenvalue criterion over the space of the transient centroid parameters. Acceptable solutions to the optimization problem will have approximately space-time separability. The constraints of the optimization are that an acceptable data window must include a minimal number of sites, days of data, and/or
total number of data. The results of our optimization yield many possible data windows that may contain transient events. We find that the set of solutions often contain robust features that are shared between many individual solutions. The features that are the most persistent through various solutions are taken to be transient deformation events. Because this optimization looks for subsets, or windows, of the data that maximize the eigenvalue criterion, we call this process windowing.

Figure 4.1 shows a basic example detection using data from the SCEC transient detection exercise Phase 1a dataset. The transient event that is detected in this data set shows a clear spatial pattern associated with slip on the Imperial Fault. The corresponding temporal pattern shows a clear step pattern. This even was clearly detectable in the data by visual inspection, as is shown in Figure 4.2.

Although the data in this example are not as complex as the data sets considered later, they are nonetheless characterized by numerous superposed processes that could, in principal, thwart the transient detector. Many time series show errant signals that could mislead one’s visual inspection. The strong spatial correlation of the offset pattern, however, is clearly emphasized by the eigenfunction analysis, as is shown in Figure 4.1.

4.2 Noise modeling

Geodetic time series contain signals that remain elusive to model deterministically and have therefore been classified as belonging to one or a combination of stochastic processes. Such stochastic signals may degrade the signal to noise ratio of geophysically interesting events, and so it is therefore desirable to attenuate well characterized noise
Figure 4.1: Simple eigenfunction analysis performed on the SCEC transient detection dataset from Phase Ia.
Figure 4.2: One third of the stations from the SCEC Phase-1a dataset. The data show a wide variety of signals, including a clearly visible offset in 2002. A arbitrary time series, the North component of the station DAM3 demonstrates the clarity of the signal (top).
sources. Langbein (2008) summarizes work on this topic and evaluates the fitness of flicker, random-walk noise, power law, flicker plus random-walk, first-order Gauss-Markov plus random-walk, and bandpass plus power law processes for a regional GPS network.

First-order Gauss-Markov processes have received particular attention because of the advantage that they are easily facilitated in state estimation formalisms (Segall and Matthews, 1997). Ji and Herring (2011) exploit this relationship and present a state estimation-eigenfunction analysis detection method. Their method successfully detects a transient deformation episode at Akutan volcano in Alaska. Our approach differs from this effort in several ways, primarily in our modeling of noise.

The present work seeks a multi-filter approach to cope with the noise content of geodetic data. Neither this approach nor a Gauss-Markov model are able to attenuate all noise spectra that are known to be present in geodetic data (Langbein and Johnson, 1997; Langbein, 2004, 2008). However, by focusing on attenuation of particular noise spectral bands, we will test the relative importance of these sources for the application of transient detection.

In what follows, we apply two filters to the data. First, the data are bandstop filtered with a two-pole infinite response filter with stop frequencies at 0.5 and 2.0 cycles per year (Langbein, 2008). Low order filters are preferred because they minimize the output phase delay. We additionally apply a low pass filter with time constant chosen to lie between 50 and 125 days. Each of these filters is applied based on several assumptions. The bandstop filter is used to filter out residual seasonal noise, as this has been found to be present in many, although not all geodetic time series (Langbein, 2008). Additionally, it may that
poorly fitting sinusoid displacement equation models may result in seasonally correlated residuals (see, for example, the study in Chapter 5. The application of lowpass filter is used because it seems likely that fault slip should have long time spans. This might result in lower resolution when more fine-scale details could be resolved, however, for a detection algorithm, we believe that this is a reasonable tradeoff.

4.3 Results

We have applied both eigenfunction-only and filter-window-eigenfunction transient detectors to the synthetic deformation datasets of the SCEC Transient Detection Exercise. We find that while the eigenfunction-only approach is well suited for large amplitude events, it may often fail to detect more subtle signals. An example of our results, for the “3F” dataset, is shown in Figures 4.3 and 4.4. We note that the eigenfunction-only method (Figure 4.3) is unable to detect the transient deformation episode.

The event’s temporal evolution and spatial distribution is clearly shown by the filter-window-eigenfunction method (Figure 4.4). The temporal pattern shows a clear logarithmic evolution in time starting in the first days of July of 2002. The colored lines in the time series plot show the effect of the time and space windows on the inferred time series. Without windowing (e.g., the blue line), the deformation is nearly invisible. The map pattern is indicative of a dip slip faulting mechanism. Clockwise from the upper right, Figure 4.4 shows the eigenvalue distribution, the normalized eigenvalue distribution, the time series eigenfunction and the spatial eigenfunction. The ratio $\lambda_1/\lambda_2 \sim 0.5$ in this example. By successively windowing the transient deformation centroid we are able to make
Figure 4.3: An example result of an eigenfunction-only transient detection method. The results are not as readily interpretable as those from the filter-window-eigenfunction method shown in Figure 4.4. The data used in this analysis is the SCEC Transient Detection Exercise Phase 3F dataset (see text for discussion). Clockwise from the upper right, the subplots show the spectra of the eigenvalues; the temporal eigenvector; and, the spatial eigenvector.
Figure 4.4: An example result of our transient detection method. The data used in this analysis is the SCEC Transient Detection Exercise Phase 3F dataset. From the upper right, the subplots show the spectra of the eigenvalues (or, Principal Components); the normalized spectra; the temporal eigenvector; and, the spatial eigenvector.
\( \lambda_1 / \lambda_2 \sim 0.01. \)

The event’s temporal evolution shows a clear logarithmic evolution in time starting in early July of 2002. The colored lines in the time series plot show the effect of the time and space windows on the inferred time series. Without windowing (e.g., the blue line), the deformation is nearly invisible. The map pattern is indicative of a dip slip faulting mechanism.

Figure 4.3 shows the results of an eigenfunction-only transient detection algorithm for the same data set. The result shows no clear pattern in space or time. Previous attempts to interpret this result lead to a false-positive detection based on an apparent step in the temporal eigenvector at 2004/02/23. The ratio of the first to second eigenvalues is greater than 0.5, indicating that the result is not separable in space and time (upper-left subplot). This result highlights the utility of the multi-filter approach. As noted above, a eigenfunction-only approach was not able to detect this event.

A summary of our transient detection results for the exercise datasets is shown in Figure 4.5. This figure shows that the detection threshold of our method is approximately 2.5cm fault slip. Additional exercise sets in the very long period regime could help us better understand our method’s performance for such events.

### 4.4 Conclusions

We have presented a transient detection algorithm to identify anomalous deformation episodes that may be present in geodetic data. The method has detected such episodes that are as small as 2.5cm of fault slip. We anticipate applying this method to real data
Figure 4.5: Results from our participation in the SCEC transient detection exercise. From the datasets provided we infer that the detection threshold for our method is approximately 2.5cm fault slip.

and will pursue this topic in future research.
Chapter 5

Correlated Hydrogeology in the
San Francisco Bay Area, California

5.1 Introduction

Many forces act to cause seasonal ground surface deformation. Examples of such processes include groundwater migration, tides, landslide creep, and surface mass loading (Langbein et al., 1995; Heki, 2001; Blewitt et al., 2001; Dong et al., 2002; Schmidt and Bürgmann, 2003; Hilley et al., 2004). Such deformations could also be artificial and due to unmodeled geodetic noise sources (Langbein et al., 1995). Cyclic deformations often accompany steady displacements from interseismic tectonics and possibly deformations that are time dependent but not cyclic such as anthropogenic exploitation of geothermal or hydrocarbon reservoirs, postseismic relaxation, fault creep, surface-load rebound, and volcanic and earthquake motions (Bürgmann et al., 2000; Cavalié et al., 2007; Segall, 2010). Repeat-pass
interferometric synthetic aperture radar (InSAR) is routinely used to image ground surface deformation due to these and other processes. An outstanding challenge in understanding InSAR, and any, ground surface displacement time series is to quantify the correlations in space and time between the myriad of processes that induce ground surface deformation.

The San Francisco Bay Area of California (hereafter “Bay Area”) is a choice location to study an eclectic ensemble of processes acting in concert to deform the Earth’s surface. A regional overview map is shown in Figure 5.1. The area is crossed by the San Andreas Fault system, a plate boundary with $37.9 \pm 0.6 \text{ mm/yr} (\pm 2\sigma)$ relative motion distributed across several major structures (d’Alessio et al., 2005). The region is also subject to the hydraulic regime of a complex network of groundwater aquifers undergoing poroelastic and anelastic deformation (e.g., Poland, 1984; Schmidt and Bürgmann, 2003).

InSAR is well suited to capture the wide range of deformation processes known to exist in the Bay Area. InSAR samples Earth surface deformation with millimetric precision, monthly temporal sampling, and $\geq 10\text{m}$ spatial resolution (Massonnet and Feigl, 1998; Bürgmann et al., 2000; Rosen et al., 2000). We analyze SAR data from the European Space Agency ERS-1 and ERS-2 spacecraft between the years of 1992 and 2000. The data used are from descending track 70, frames 2835 and 2853. Although our processed data spans this entire interval, we analyze the shorter interval from 1995 to 2000. This interval avoids the period during 1993 and 1994 when the ERS-1 satellite was in a different acquisition mode, and was chosen to mitigate possible effects of inhomogeneous temporal sampling frequency. The spatial extent of the study area is shown in Figure 5.2. The data are processed using the PS-InSAR technique (Colesanti et al., 2003; Ferretti et al., 2000; Hanssen, 2001). In this
approach, atmospheric contribution to apparent displacement is simultaneously estimated from persistently radar-bright, phase-stable natural and artificial reflectors (e.g., buildings, roads). Our data consist of 40-point time-series for the ground displacements of 178,275 reflectors.

5.2 Ground surface velocities

We begin our analysis by analyzing the first order features of the dataset first order features that are described by the velocity (rate of range change) field. The velocity field is plotted in Figure 5.3. The most basic feature of the velocity field is a regional velocity gradient that is consist with the $37.9 \pm 0.6$ mm/yr ($\pm 2\sigma$) relative motion across the North American-Pacific tectonic plate boundary. The velocity field clearly shows that this deformation is partitioned across several major faults and dozens of smaller ones.

The greatest velocities are due to two anthropogenic sources: engineered soil and subsurface fluid extraction. The areas that show large subsidence signals due to engineered soil include Foster City, the San Francisco airport, the Oakland airport. The city of Stockton and an region north of the city of Woodland, both in the central valley, show large subsidence signals that are probably related to groundwater extraction.

Most of the San Joaquin River Delta contains few radar reflecting points. The points that are available in this region show subsidence which is locally in excess of 10 mm/yr. Subsidence in the San Joaquin River Delta is due to microbial oxidation of organic-rich peat material (Rojstaczer et al., 1991; Deverel and Rojstaczer, 1996; Mount and Twiss, 2005) as opposed to groundwater pumping. Similar subsidence has been seen in other delta
Figure 5.1: Overview map of the San Francisco Bay Area.
Figure 5.2: Relative surface displacements for the dominant correlated patterns of deformation (eigenfunctions 1-3 in A-C, respectively). Red colors show toward-satellite motion. The temporal variation of these patterns is shown in the normalized time series plots (D) in comparison with binned monthly precipitation in San Jose (D), and the singular value spectra (E). Note the varying color scales. The place labels are: SAF, San Andreas Fault; SCF, Silver Creek Fault; SCA, Santa Clara Aquifer; SR, San Ramon; P, Pleasanton; FC, Foster City; P, Pleasanton. The three magenta circles in (B) show the locations of well water level data, which are, in order from right to left, w53, w220, and w242.
systems, including the Nile (Becker and Sultan, 2009) and the Mississippi (Törnqvist et al., 2008). Spatial variations in subsidence rates appear to be correlated with the varying basin geomorphic features such as fan, channel, or basin (Strand and Koenig, 1971). Of particular interest are radar reflectors that lie atop levees, as the rate of subsidence of such points has been shown to be correlated with levee failure (Dixon et al., 2006; Dixon and Dokka, 2008).

Aquifers that recharge from the runoff of the Sierra Nevada mountains, e.g., those in the Central valley, tend to show subsidence, whereas groundwater that recharge from the Coast Range mountains tend to show uplift. This contrast between Central Valley and Coastal Range aquifers could be explained by either natural spatial variations in recharge or by differing management strategies, although it is not clear which of these causes is dominant. Uplift has been documented in the Santa Clara Aquifer by Schmidt and Bürgmann (2003), who studied the same track and time period of InSAR data that we discuss here, although with a different processing strategy. In the case of the Santa Clara aquifer, long term uplift is credited with higher groundwater levels due to changes in management practices in the last forty years (Schmidt and Bürgmann, 2003).

5.3 Seasonal Deformation

In this section we describe a method to isolate the dominant pattern of annual deformation from a set of interferograms or other time series records of ground surface displacement. Particularly, we seek a method to separate out the various cyclical deformation processes that may be present in the observed PS-InSAR measured range-change time series. In the following, we show that eigenfunction analysis is a method that may be used
Figure 5.3: Velocities from the Bay Area PS-InSAR dataset. The largest velocities are related to subsidence in Foster City and Sacramento. The largest, most coherent, natural signal is the uplift of the Santa Clara Aquifer.
to address this challenge. In what follows, we describe this method, apply it to PS-InSAR data from the Bay Area, and then compare the results to the technique of displacement equation fitting.

Surface displacements measured by InSAR are highly correlated in space and time (Williams et al., 1998; Hanssen, 2001; Lohman and Simons, 2005). These correlations imply that the displacements of many individual points on the surface of the Earth may be redundant and therefore may be approximated without significant loss of information. Many methods that achieve such an approximation could be chosen such as a quadtree based downsampling based on data-scatter (Jonsson et al., 2002). Another potential approach is to rely on the intrinsic statistical structure of the data rather than on a prior model formulation. By making fewer assumptions about the processes involved in an unknown system, such a strategy would have the benefit of being applicable to a wide class of data.

The eigenfunction analysis that we present, sometimes called principal component analysis, is a standard multivariate analysis technique that expresses highly correlated datasets as several representative linear functions. These functions are associated with a ranking by the amount of data set variance represented in each. An important advantage of is that there is no dependence on a prior functional form for the model. Such methods have been used to widely in geophysical geodesy, for example to decompose velocity fields in volcanic caldera (Savage, 1988), in areas undergoing postseismic deformation transients (Savage and Langbein, 2008; Savage and Svarc, 2009; Barbot et al., 2009; Savage and Svarc, 2010), and at a regional scale in southern California (Savage, 1995; Dong et al., 2006). A statistical treatment is given by Joliffe (1986). To the authors’ knowledge, this type of
analysis has not previously been conducted with permanent scatterer InSAR data.

We form a data matrix, $D$ in which each column contains the PS-InSAR derived line-of-sight displacements for a given epoch and each row contains the displacement time series of a particular point on the surface. Time series that have rates greater than 15mm/yr, a level about equal to $3\sigma$ in RMS scatter, are removed from the analysis. This corresponds to 325 of 178,600, or 0.18%, of the data points. This step is justified because outliers significantly bias the eigenfunction decomposition. The resulting data matrix has dimensions of 178275x40. Each time series is detrended to remove the best fitting linear velocity ramp; it will be important to recall this when interpreting results. This step is taken in order to increase the signal to noise ratio of cyclical phenomena.

We find the singular value decomposition (SVD), $D = EC^T$. The resulting matrices $E$ and $C$ have as their rows the left and right singular vectors of $D$. These singular vectors represent sets of spatial patterns and the linear variation of these patterns in time, respectively. The matrix $A$ is diagonal and its entries are the singular values.

The data matrix $D$ may be reconstructed using approximations to $E$, $A$, and $C$. An important property of this reconstruction is that it satisfies the following optimal criterion. If $A^\alpha$ is the matrix $A$ with all diagonal entries beyond the $\alpha^{th}$ entry changed to zero, then it is easily verified that the square of each singular value is equal to the $L^2$ misfit between $D$ and $EA^\alpha C^T$. The singular values are therefore a measure of “importance” in the sense that they give the improvement in $L^2$ error achieved by reconstruction including that element.

Figure 5.2 shows the first three eigenfunctions of the Bay Area data set: the
first three spatial patterns are shown in figure 5.2a, 5.2b, and 5.2c, and the corresponding
time series are shown in figure 5.2d. We observe two striking features of the first three
eigenfunction time series (right singular vectors). First, each of these time series show a
strongly annual periodic frequency. Second, each of these periodic signals appears to be
shifted in phase with respect to the others. We choose to analyze only the first three
eigenfunctions because their associated singular values (SVs, the diagonal entries of Λ),
are distinctly larger than the less significant eigenfunctions (Figure 5.2e), and therefore
represent the majority of the variance of the data set.

5.3.1 Empirical impulse response estimation

We hypothesize that these features are linearly related to hydraulic forcing by pre-
cipitation. To examine this model we compare deformation to a record of binned monthly
precipitation at the San Jose Airport (shown with a white dot in Figure 5.2c). This pre-
cipitation record is shown in gray in the lower-left subplot of Figure 5.2. We note that the
first eigenfunction time series appears to describe peak uplift at the epoch of maximal pre-
cipitation, the second eigenfunction time series reaches peak uplift about one month after
the peak in precipitation, and the third eigenfunction time series reaches peak uplift at the
onset of the dry season.

We estimate precipitation-deformation impulse response functions (IRFs) for each
of the first three eigenfunction. The response functions are assumed to describe a linear
and time-invariant model, as visual inspection of the data appears to confirm. The impulse
response is modeled as a simple linear function of the form,

\[ C_{kt} = \phi_0 + \sum_{i=1}^{p} \phi_{ki} R_{t-i} + \epsilon_t, \]

(5.1)

where \( R_{\tau} \) is the rainfall at epoch \( \tau \), \( \epsilon_t \) is white noise, and \( p \) is the number of epochs over which the impulse response is estimated. The parameters \( \phi_{ki} \) relate rainfall at time \( t - i \) to deformation in the \( k^{th} \) eigenfunction at time \( t \). In order to compare the monthly precipitation data with the somewhat irregular InSAR acquisition dates, the eigenfunction time series are linearly interpolated onto a monthly time grid. Although this procedure could affect extreme values of the time series, in practice we do not find that this is a problem.

Figure 5.4 shows the estimated IRFs for each of the first three eigenfunction time series. We compute the IRFs for \( p = 11 \) epochs, or about 260 days. This cutoff was chosen because a longer interval would begin to include the harmonic effects of wrapping into the following year’s rainy season.

The results of the IRF analysis quantify what is visually apparent in Figure 5.2d: cyclic deformation occurs with several distinct annual phases. Particularly, the first component of deformation shows no resolvable delay with respect to the onset of precipitation. This deformation tends to have a residence time of about 60 days after precipitation before it begins to diminish. The second component of deformation has a more broad peak than the first and reaches its epoch of maximal uplift after a delay of about 60 days. The third component of deformation has the sharpest peak of the first three components, and this peak occurs about 150 days after precipitation.
Figure 5.4: (A) Precipitation-deformation impulse response functions for the first, second and third eigenfunctions shown in blue, green, and red, respectively. (B) Well levels at three sites in the Santa Clara valley. The locations of these sites are shown in Figure 5.2b. See text for discussion.
5.3.2 Displacement equation estimation

Quantifying the oscillations of geodetic time series is perhaps most naturally accomplished by estimating the amplitudes and phases of a series of sinusoidal functions. While any smooth function may be represented as a sum of scaled sinusoidal functions, because of their cyclic phenomenology, geodetic time series with seasonal contributions are often well represented by about four superposed sinusoidal functions (Nikolaidis, 2002). Neglecting offsets and linear trend, a best fit may be determined with respect to the displacement function,

\[ y(t_i) = \sum_{j=1}^{q} \left[ a_j \sin(\omega_j t_i) + b_j \cos(\omega_j t_i) \right]. \] (5.2)

In this expression the \( a_j \) and \( b_j \) determine the amplitude of oscillation and the \( \omega_j \) determine the frequency of oscillation. For values of \( t \) given in years, it is common to have \( q = 2 \) with \( \omega_1 = 2\pi \) and \( \omega_2 = 4\pi \). We find that allowing the \( \omega_j \) to be free parameters in the estimation greatly improves the goodness of fit. We additionally set \( q = 3 \). We therefore determine the best fitting six parameters for each time series in the PS-InSAR data set. This is done using the least squares criterion and the Nelder-Meade low-dimensional simplex method of Lagarias et al. (1998).

5.3.3 Model comparison

The sum of square residuals (SSR) is much lower between the eigenfunction model and the data than between the displacement equation model and the data. It is not immediately clear, however, that this comparison is equal. While the eigenfunction model was derived simultaneously using the entire data set at once, the displacement equation model
was fit to each time series individually. We therefore seek a statistic that will offer a fair comparison of these two models despite their differing approaches.

Every statistic has an associated number of degrees of freedom. The degrees of freedom indicates how many independent pieces of information were need to compile the statistic.

The number of degrees of freedom for the sinusoid model is simply the total number of data minus the total number of parameters fit. Faber (2008) has clarified the relevant definition of the degrees of freedom for the residuals of a eigenfunction analysis. For a eigenfunction analysis that uses the leading \( k \) eigenfunctions of a data set with \( m \) objects of dimensions \( n \) and has centered (zero-mean) rows, the number of degrees of freedom is given by

\[
df = (m - k) \times (n - k - 1).
\] (5.3)

The SSR/df for the eigenfunction model is 7.148 mm\(^2\). The displacement equation model with \( q = 2 \) has an SSR/df of 9.867 mm\(^2\), and increasing to \( q = 3 \) gives an SSR/df of 9.717 mm\(^2\). Tables of these statistics are given in Appendix D. The SSR statistic for each individual radar scatterer time series is plotted in map view in Figure D.1. Figure D.1a shows the SSR for the eigenfunction model, Figure D.1b-d show the displacement equation model for \( p = 1, 2, 3 \), respectively. This figure shows that although the sum of sinusoid model fits time series quite well in many areas, areas of poor fit (high SSR) are highly spatially correlated. Furthermore, these correlated areas of poor fit appear in precisely the locations of highly coherent signals in the first few eigenfunctions (e.g., compare Figure 5.2a to Figure D.1), a topic of later discussion.
5.4 Discussion

The deformation of the Santa Clara Valley is the most prominent feature of our PS-InSAR data set. This deformation is apparent in the first and second eigenfunctions. In the first eigenfunction, deformation occurs over the entire breadth of the valley and with uplift (range change) coinciding with increased rain fall. (Figure 5.4a). In the second eigenfunction, deformation occurs in a wide swath near the middle of the basin, and temporally occurs with a $\sim 60$-day delay. Although we will focus the discussion on the Santa Clara Aquifer, other features that are mapped in the same eigenfunction tend to reoccur simultaneously in the data set. Inclusion in the first eigenfunction, for example, indicates that aquifers in San Ramon, San Leandro, and Pleasanton tend to move in unison with the Santa Clara Aquifer.

In order to investigate the timing of the first two eigenfunction time lags, we examine well levels at three frequently sampled wells in the Santa Clara Aquifer. The data are plotted in Figure 5.4b and the location of the wells are noted in Figure 5.2b. The well w53 is located in the extensive area of peak deformation of the second eigenfunction, and the data show that this well is artesian (zero depth to water) during much of the year. The broad peak in the second eigenfunction’s IRF (Figure 5.4a) may therefore be interpreted as the time during which groundwater is very near or at the surface.

In contrast, regions that are further away from the valley axis experience peak deformation at roughly the same time that peak well levels are observed at wells such as w220 and w242. The rapid seasonal retreat of water levels at these wells corresponds to the decline of the peak amplitude of the first eigenfunction’s IRF. Because of the seasonal
longevity of the artesian units, we speculate that the decline in peripheral levels, and associate deformation, is due to lateral groundwater flow towards the valley axis. More dense temporal sampling of InSAR data, such as will be possible with the future SENTINEL-1 and DESDynI SAR satellite missions, may permit the evaluation of this hypothesis entirely with remote sensing data.

The third eigenfunction contains numerous correlated signals that are generally of shorter spatial wave length than those contained in the first two eigenfunctions. The time series of this eigenfunction contains a characteristic $\sim 150$ day time lag. In this pattern, upward displacements in San Jose, Sunnyvale, and Palo Alto occur simultaneously with downward displacements in Foster City, the San Francisco International Airport, and other regions built upon reclaimed land. Additionally present are cyclical subsidence patterns on Treasure Island and near the Oakland Airport, as well as modulated slip on deep seated landslides.

Fault boundaries are readily visible in each of the first three eigenfunction, most prominently the Silver Creek and the San Jose Faults in the first eigenfunction and the San Andreas and Hayward Faults in the third eigenfunction. The relative displacements that occur across these structures are unlikely to represent fault slip because the pattern of displacement is cyclical in time. From the frequent occurrence of structurally bounded aquifers we conclude that such features represent a first order control on regional groundwater flow. This conclusion is similar to the results of previous studies in this area (Schmidt and Bürgmann, 2003) as well as comparable studies in other localities (Bell et al., 2008).

We find that the eigenfunction method has lower SSR/df than the displacement
equation method when the leading three eigenfunctions are used. This result might be expected because eigenfunction models satisfy numerous relevant optimal algebraic properties (Joliffe, 1986). Residuals in Figure D.1 shows that areas of poor fit are more spatially correlated for the displacement equation approach, and furthermore, these areas of poor fit tend to coincide with the significant features of the first few eigenfunctions that have been discussed. Although more terms could be added to the displacement equation (increased values of \( q \)), the better fit to the data would be outweighed by the loss of analytical simplicity. This result suggests that seasonal deformation is significantly more complex than a model of constant-amplitude annual and semiannual oscillations. We speculate that as more data become available, decadal oscillations (e.g., related to El Niño-Southern Oscillation) will become more visible in the data. It is possible that the widespread use of simple displacement equation estimation in the literature has been successful in explaining geodetic data largely because these data have not analyzed dense, multi-decade datasets.

We assume that precipitation induced groundwater redistribution is the dominant agent responsible for the observed ground surface displacements, however, the timing of precipitation is strongly correlated with other possible causes of seasonal surface displacements. The model of ground surface displacements in each eigenfunction is a combination of many processes that tend to occur simultaneously. Other seasonally variable phenomena that might be present in the data include atmospheric loading, seasonally correlated errors (e.g., atmospheric mis-modeling), and tides (Blewitt et al., 2001; Dong et al., 2002). Despite the presence of numerous sources of deformation in each eigenfunction, some features of the eigenfunctions can unmistakably be attributed to certain mechanisms. It is highly
unlikely, for example, that any mechanism besides groundwater redistribution could be responsible for the surface deformation that is bounded by the Silver Creek Fault in the first two eigenfunctions.

The calculation of all eigenfunctions is $10^2$ to $10^3$ times computationally faster than estimating the unknown parameters of the displacement equation (Equation 5.2). Computing only the first few eigenfunctions is even faster, usually requiring only a few tens of seconds on a single 2009 MacBook Pro with a 2.26GHz Intel Core 2 Duo processor and 8GB 1067 MHz DDR3 memory. Fitting the parameters of the displacement equation required at least several to a dozen hours on the same computer, depending on the number of terms included in the formulation.

### 5.4.1 Timing of hydrologic deformation

Using in situ measurements of the Santa Clara aquifer’s mechanical properties we may estimate the expected timing of hydrologic deformation.

Horizontal hydraulic conductivities in the aquifer were measured to be as great as $2.06 \times 10^{-3}$ m/s. The hydraulic diffusivity, $\alpha$ is given by $\alpha = k/S_s$, where $k$ is the hydraulic conductivity and $S_s$ is the specific storage. A typical value for the elastic specific storage of a sandy aquifer unit is $9 \times 10^{-6}$ m$^{-1}$. These values yield a hydraulic diffusivity of 233 m$^2$/s.

The RMS diffusive distance is given by $\sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{2n\alpha t}$. Assuming 2D (confined) diffusion and groundwater migration distances of 15km to 20km, this implies diffusion times of two to five days. This is plausible because we used an extreme value of the hydraulic conductivity. Using a more moderate value of $k = 5 \times 10^{-4}$ m/s gives diffusion
times of 12 to 21 days.

5.4.2 Identification of fault traces

InSAR observations often show that faults in the shallow, near-surface environment act as barriers to groundwater flow (Bawden et al., 2001; Schmidt and Bürigmann, 2003; Bell et al., 2008). We use our transient detection algorithm to create a less noisy seasonal deformation signal. In applying this method to real data, we have identified a previously unmapped fault in the San Francisco Bay Area. This fault is shown in Figure 5.5.

The seasonal deformation signal shown in Figure 5.5 was created by filtering the processed data from ERS Track 70 and finding the first eigenfunction of a windowed selection of the data. The plot shows the difference between the maximum and minimum values of the seasonal deformation pattern, i.e., the difference between the rainy season and the dry season.

5.5 Conclusions

Measurements of surface displacements in the Santa Clara Valley show strong seasonal and annual modulation. We have shown that eigenfunction analysis greatly facilitates the interpretation of ground surface displacements due to such cyclic deformation. In our dataset, the first and second eigenfunctions correspond strongly with aquifer units that reach peak water heights during different parts of the year. This interpretation is justified by examination of well water level data. A statistical juxtaposition shows that our eigenfunction method is statistically competitive with the frequently used displacement equation.
Figure 5.5: Seasonal deformation from a subset of the ERS T70 InSAR dataset. The deformation pattern shows a confined aquifer that is bounded by two previously unknown faults. Known faults are shown with thin, black lines.
method. We therefore conclude that eigenfunction-based approaches might be used in conjunction with displacement equation methods when examining seasonal deformations. Our results indicate that a three-phase model of deformations elegantly describes the majority of seasonal correlated deformation in the San Francisco Bay Area, California.
Chapter 6

Optimal Design of Geodetic Networks

6.1 Introduction and Background

The increasing quality of space-based geodetic networks has radically altered numerous scientific fields, including tectonics, seismology, Earth structure, meteorology, climatology, hydrology, glaciology, and atmospheric physics (Blewitt, 2007). This scientific impact has been the result of increased observation density (e.g., the Global Positioning System [GPS] Earth Observation Network, the Plate Boundary Observatory, the Southern California GPS network, the Continually Operating Reference Service, the International GNSS Service) and, fundamentally, the increasing quality of global reference frames (Al-tamimi et al., 2007).

The optimal design of geodetic networks is a classical problem in geodesy (Gru-
farend and Sanso, 1985; Vanicek and Krakiwsky, 1986). Indeed, early work in the field of optimal experimental design was inspired by problems of optimal geodetic investigation (Nordstrom, 1999). Geodetic networks using GPS need not be designed to enhance positional accuracy but may be designed to better observe particular geophysical events (Blewitt, 2000). Basic questions concerning the quality of such networks include: how might one most efficiently alleviate subnetworks which are insufficiently monitored? Alternatively, when conducting statistical data inversions with the products of these networks, could similar results be achieved using fewer stations? Is it possible to design networks whose structure is more resistant to atmospheric and other correlated noise sources? These questions concern optimal geodetic network design and are the focus of the present work.

One natural way to evaluate the performance of noisy geodetic networks is to conduct simulations with noisy synthetic data (Johnson and Wyatt, 1994; Segall and Schmidt, 2003; Schmidt et al., 2003). This approach has been used to study the effects of white and random walk noises (Segall and Schmidt, 2003; Schmidt et al., 2003) as well as spatially correlated noise (Johnson and Wyatt, 1994). In the case when the data kernel and data covariance matrices are known, or are at least assumed, then studying the forms of the resolution matrices and the covariance or information matrix may be advantageous. Osten-sibly, the primary benefit in this regard is to avoid the brute-force approach of iteratively calculating realizations of a noise model (Menke, 1986; Blewitt, 2000; Gerasimenko et al., 2000). However, the global optimization problem of maximizing a statistic of one of the matrices is also a computationally intensive endeavor (Johnson and Wyatt, 1994; Berne and Baselga, 2004).
We choose to analyze the information content of geodetic networks because it is a single statistic that, as we shall show, describes a desirable property of geodetic networks. This paper is organized as follows. Section 6.2 develops a model of the information content of geodetic networks. We select an information criterion that describes the amount of information about geophysical parameters contained in a geodetic network (Pukelsheim, 1993). Using this metric we then evaluate geodetic network performance at two different scales. In Section 6.3.1 we evaluate the importance of each local campaign GPS observation of the Parkfield, California earthquake. We find that a small subset of observations contribute the majority of uncertainty reduction during geophysical parameter estimation. In Section 6.3.2 we evaluate a dense regional network spanned by 1256 continuously operating GPS stations in the western United States, and a regional fault model consisting of $\sim 12,500$ fault elements.

The main conclusion of this study is that, although geodetic GPS observations are commonly perceived as low cost and easy to deploy, their optimal placement nonetheless strongly affects the quality of geophysical model parameter estimates. To cope with any finite pool of engineering resources, consideration of the information content of geodetic network design provides an efficient means to maximize the quality of inferred geophysical parameters.

6.2 Model

We consider the well known discrete, linear inverse problem with Gaussian statistics: to estimate the vector of model parameters, $\mathbf{x}$, given Gaussian noise $\epsilon$, data vector $\mathbf{y}$,
and the linear stochastic model (Jackson, 1972; Menke, 1986),

\[ A(r, x)x = y + \epsilon. \]  \hspace{1cm} (6.1)

The vector \( r \) contains the geographical coordinates of the geodetic network.

We consider the general case of a mixed-determined inverse problem, e.g., some linear combinations of the model parameters are overdetermined and some are underdetermined. The so called “natural solution” is given by,

\[ x = A^{-1}y \]  \hspace{1cm} (6.2)

where,

\[ A^{-1} = V_p \Lambda_p^{-1} U^T \]  \hspace{1cm} (6.3)

is the Moore-Penrose, or “natural” inverse. The matrices \( U, S, \) and \( V \) form the generalized (non-square) eigendecomposition of \( A \) (Joliffe, 1986). In general, given a matrix \( X \), the notation \( X_p \), will refer to the matrix that is formed from the first \( p \) columns of \( X \). A superscript \( T \) indicates matrix transpose.

The corresponding model parameter covariance matrix is given by,

\[ C_m = A^{-1} C_d A^{-T} \]  \hspace{1cm} (6.4)

\[ = (V_p \Lambda_p^{-1} U^T) C_d (V_p \Lambda_p^{-1} U^T)^T \]

Control over the parameter \( p \) is useful because we may use it to adjust the tradeoff between resolution and precision (Menke, 1986).

The inverse \( C_m^{-1} = P \) is the information matrix of the model parameter vector \( x \). The introduction of the information matrix is useful because it represents the amount
of information about $\mathbf{x}$ contained in the geodetic network. We seek a summary statistic to characterize the “size” of the information matrix $\mathbf{P}$. We choose the determinant $|\mathbf{P}|$ as our statistic, or information criterion, because it has several useful properties. The quantity $\sqrt{|\mathbf{P}|}$ is inversely proportional to the confidence region of the estimated model parameters; a large value for $|\mathbf{P}|$ therefore secures a small volume of the confidence ellipsoid. The determinant is invariant under an internally consistent change of units (Grafarend and Sanso, 1985; Pukelsheim, 1993; Blewitt, 2000) and change of parameter basis (Meyer, 2000).

The structure of the data covariance strongly affects the information content of a geodetic network, and noise that models simple spatial covariation may be understood to represent a combination of atmospheric processes that corrupt GPS position estimates (Johnson and Wyatt, 1994). We note that the simplest case of statistically independent observations imply independence of optimal observation locations, clearly an oversimplification. In contrast, a basic spatially correlated noise model dictates that the covariance between two observations is inversely proportional to the distance between the observation sites. We additionally suppose that covariation occurs over a characteristic length scale, $L$. We choose a covariance function with finite support so as to provide greater numerical stability (Sansò and Schuh, 1987). One such class of covariance functions are of the form (Watkins, 1992):

$$W_{ij}(\delta) = \begin{cases} \sigma^2 \left(1 - \frac{h_{ij}}{L}\right)^m & 0 \leq h_{ij} < L \\ 0 & h_{ij} \geq L \end{cases}$$  \hspace{1cm} (6.5)

A comparison of related covariance functions is shown in Figure 6.1. In the remainder of the present work, we choose to simulate covariation with the simple linear ramp covariance...
function specified by Equation 6.5 with \( m = 1 \). A more general discussion of the information content of observations of an elastic dislocation is given in Appendix E.1.

6.3 Examples

6.3.1 Network expansion at Parkfield, California

Johanson et al. (2006) conduct a fault slip inversion of interferometric synthetic aperture (InSAR) and GPS data for the 2004 September 28 \( M_w 6.2 \) Parkfield, California earthquake. There was a high density of GPS data coverage for this earthquake, with seventeen continuous and five campaign GPS sites used. Using the fault model and network configuration described by Johanson et al. (2006), we evaluate the importance of the campaign GPS sites in this data inversion. This investigation could be useful if, for example, one were to choose the most useful campaign site to be converted to a continuous site.

The geometry of the GPS network in Parkfield, California, is shown in Figure 6.2. We use a fault model that is very similar to that of Johanson et al. (2006) and Barbot et al. (2009). The model consists of a vertical right-lateral strike slip fault with a strike of 315° that is discretized into \( \sim 5 \)km square fault patches.

We first compute the information criterion of the continuous GPS network alone. We then compare this value to the information criterion of the continuous GPS network plus each one-, two-, three-, and four-site subnetworks of the campaign GPS network. In this way we determine which campaign sites add the most information to the continuous GPS network. Results will always be normalized to the information criterion value of
Figure 6.1: Comparison of three covariance functions considered in this study. The linear ramp function is favored for its simplicity and compact support on the real line.
Figure 6.2: The geometry of the geodetic network in Parkfield, California. Continuous GPS sites from the SCIGN and PBO networks are shown with squares and diamonds, and campaign GPS sites are shown with triangles. The gray line denotes the surface projection of the fault that ruptured in the 2006 event, and black lines denote other mapped faults. Also pictured are the 2006 epicenter (gray star), the SAFOD fault drilling experiment (gray circle) and highways (dashed lines). Black dots indicate a relocated seismic catalogue (Thurber et al., 2006). Figure cropped from Johanson et al. (2006).
the continuous-only GPS network. The relationship between information and confidence regions implies that if the continuous network, augmented with a subnetwork of campaign GPS sites, takes an information criterion value of unity, then this subnetwork does not alter the size of the confidence region. In this case we would conclude that the augmented subnetwork contributes minimal information to the network.

The results of this comparison are shown in Figure 6.3 and in Tables 6.1 and 6.2. Because the spatially varying noise parameter $L$ is not known a priori, we compare four values of $L$: 1km, 5km, 10km, and 50km. For each number of sites, $n$, we find the best and worst combination of $n$ sites. These are shown in Figure 6.3 as solid lines and dashed lines, respectively. For each noise model considered, the single site that contributes the least takes a value near unity.

Our analysis shows that the site 05QJ is the most informative site of the campaign GPS subnetwork for all cases except the long wavelength noise model ($L = 50km$). In the long wavelength model, the site 05SH is preferred. We note that in the presence of long wavelength noise, the station that is further from the center of the network (05SK) is preferred to the station that is more proximal (05QJ). The campaign sites that contribute the least information are 510 and 05SH, both of which are somewhat far from the rupture.

The gap between the best and worst performance cases is typically on the order of ten normalized information units. This is a striking performance difference: it implies that a ten-fold reduction in the uncertainty of a slip model could be achieved by choosing well placed geodetic observation sites. The difference is the greatest for the case of a large spatially covariant noise parameter.
We conclude from these calculations that 1. The longest scales of spatially correlated noise favor more compact GPS arrays, and 2. that a typical improvement from an arbitrarily-placed site to a well-chosen site could achieve a ten-fold reduction in model parameter uncertainty.

6.3.2 Regional scale network design in southern California

Southern California is home to one of the world’s most spatially dense networks of continuously operating GPS stations, the SCIGN/PBO network. This fact notwithstanding, it may still be inquired whether additional sites would improve the crustal deformation observing capacity of this network.

One approach to improve the existing SCIGN/PBO network would be to evaluate its performance against several recent earthquakes. However, given that most faults in the region have not yet experienced a complete earthquake cycle within the lifetime of the network, such an effort would be fundamentally limited. Instead, we choose as the basis for this experiment a simulated record of fault slip on 12,502 discretized fault elements (each approximately 3 km x 3 km) in a system of faults meant to approximate the actual California fault system. The geometry and long-term average slip rates for these faults are based on those in the UCERF2 report (Working Group on California Earthquake Probabilities (WGCEP), 2007) and are used in a Southern California Earthquake Center project comparing several different earthquake simulators (Tullis et al., 2010). This simulated record is created with RSQSim, a fast, physics-based multi-cycle earthquake simulator described by Dieterich and Richards-Dinger (2010). From a simulated RSQSim catalog containing
<table>
<thead>
<tr>
<th>L</th>
<th>Sites</th>
<th>$\sqrt{P}$</th>
<th>L</th>
<th>Sites</th>
<th>$\sqrt{P}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1km</td>
<td>05QJ</td>
<td>3.0730</td>
<td>10km</td>
<td>05QJ</td>
<td>3.0328</td>
</tr>
<tr>
<td></td>
<td>05SK</td>
<td>7.7021</td>
<td></td>
<td>05QJ</td>
<td>6.9211</td>
</tr>
<tr>
<td></td>
<td>05SK</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>508</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
<tr>
<td>1km</td>
<td>05QJ</td>
<td>14.7592</td>
<td>10km</td>
<td>05SK</td>
<td>14.5262</td>
</tr>
<tr>
<td></td>
<td>508</td>
<td></td>
<td></td>
<td>510</td>
<td></td>
</tr>
<tr>
<td></td>
<td>05QJ</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
<tr>
<td>1km</td>
<td>05SK</td>
<td>20.0322</td>
<td>10km</td>
<td>05SK</td>
<td>18.9287</td>
</tr>
<tr>
<td></td>
<td>508</td>
<td></td>
<td></td>
<td>510</td>
<td></td>
</tr>
<tr>
<td></td>
<td>05SH</td>
<td></td>
<td></td>
<td>05SH</td>
<td></td>
</tr>
<tr>
<td>5km</td>
<td>05QJ</td>
<td>2.9726</td>
<td>50km</td>
<td>05SK</td>
<td>3.3839</td>
</tr>
<tr>
<td></td>
<td>05QJ</td>
<td>7.0389</td>
<td></td>
<td>05SK</td>
<td>10.4765</td>
</tr>
<tr>
<td></td>
<td>05QJ</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
<tr>
<td>5km</td>
<td>05SK</td>
<td>13.4291</td>
<td>50km</td>
<td>05QJ</td>
<td>22.9471</td>
</tr>
<tr>
<td></td>
<td>508</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>508</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
<tr>
<td>5km</td>
<td>05SK</td>
<td>17.8536</td>
<td>50km</td>
<td>05QJ</td>
<td>24.3473</td>
</tr>
<tr>
<td></td>
<td>05SH</td>
<td></td>
<td></td>
<td>05QJ</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: Combinations of the campaign GPS sites that maximize the network’s information content. Although all site combinations were tested, only the extremal site combinations are shown.
Table 6.2: Combinations of the campaign GPS sites that minimize the network’s information content. Although all site combinations were tested, only the extremal site combinations are shown.
Figure 6.3: Plot of information criterion (IC, normalized to reference value) versus the number of campaign sites augmented to the continuous network. The figure shows the information criterion for each of several noise models (color scale), for the worst- (dashed line) and best- (solid lines) performing networks.
1.5 million events over 55,000 years, we extract descriptions of fault slip in each of 12,092 events with magnitudes greater than 6.5. We analyze the region of the southern California fault system bounded by 121°W, 115.5°W, 32.5°N, and 35.5°N. Our description of the current GPS network in southern California was obtained from UNAVCO\(^1\). In order to avoid edge effects, we include in our analysis 1256 GPS sites that lie throughout western North America.

For computational efficiency, we consider a subset of the full parameter family of interest. We consider a linear combination of the model parameter vector \(x\) (Pukelsheim, 1993),

\[
\theta = Kx. \tag{6.6}
\]

A reasonable choice for the new parameter basis could be obtained through an eigenvalue transformation of a real or synthetic dataset (Joliffe, 1986). Such a dataset would consist of a large collection of slip distributions, and the matrix \(K\) could then be interpreted as a collection of representative slip patterns. We proceed with this method, applying the eigen-decomposition to a data matrix of slip distributions produced by the RSQSim earthquake simulator. We note that this is the second use of the generalized eigendecomposition in the present work. The first was used in a typical manner: in Equation 6.3 to define a generalized inverse. The second, in the present context, is somewhat less traditional. Several examples of the representative patterns derived from this method are shown in Figures E.5, E.6, and E.7. A similar methodology is used in Chapter 5, and references concerning this topic are given there.

\(^1\)http://facility.unavco.org/data/dai2/app/dai2.html
Our optimization method is to perform a grid search, on a 1.5km grid, to find the location of the best additional site. The results of this experiment are shown in Figure 6.4, which plots the information content of site choices in southern California. Warmer colors in this plot show the locations of potential GPS sites that would increase the information content of the network. Every fault pictured in Figure 6.4 was used in the RSQSim simulation, however not every fault patch is included in the parameterization of Equation 6.6. We would therefore not necessarily expect all regions with sparse station coverage to be regions of high information content.

The near fault environment surrounding the San Jacinto fault is that with the greatest potential to improve the information content of the existing GPS network in southern California to future earthquakes. Figure E.8 shows the San Jacinto fault region of Figure 6.4 in greater detail. It is interesting to note that the precise location of maximal information contribution to likely coseismic events occurs in the Anza Valley. This area has a high density of observation on the east side of the San Jacinto Fault, nearer the San Andreas fault, and lower density on the west side of the fault.

6.4 Discussion

The current study has modestly generalized previous results to analyze inverse problems of mixed determinacy (Equation 6.3) and orthogonal reparameterization (Equation 6.6). In contrast with earlier work, the present study has focused mostly on the problem of network design for the observation of coseismic deformation. Most previous studies of geodetic network design for fault mechanics observation have focused on the problem of
interseismic deformation (Johnson and Wyatt, 1994; Gerasimenko et al., 2000; Blewitt, 2000).

We note that the maximal increase in information varies dramatically between the Parkfield and southern California examples. In the Parkfield case we were able to achieve a normalized information criterion value of $\sim 25$, whereas the maximal value throughout southern California was $\sim 1.1$. This observation is expected given the length scales and station distributions involved in these calculations. In a topological sense (e.g., with normalized length scales) the southern California regional network is both more dense and more uniform than the Parkfield local network. For observing large, regional earthquakes, the southern California regional network is rather well designed, as evidenced by the marginal improvements offered by augmentation. For observing a local earthquake, the local Parkfield network does not fare as well.

We therefore offer the recommendation that GPS networks be designed at multiple scales. Although this recommendation may seem somewhat obvious, it is currently only adopted in a small handful of localities in the California study areas. Our analysis suggests that the southern California regional network is well designed for regional deformation studies but could be improved by densification in the near fault environment.

Selection of sites for network augmentation is strongly affected by logistical and geologic concerns, such as site accessibility, land rights, and substrate stability. Not all sites that enhance a network’s information content will be practical options for expansion. Nonetheless, our network design experiment offers a significant constraint on the preferred siting of extensions to the southern California continuous GPS network.
Our southern California experiment demonstrates that the near-fault environment offers the greatest potential for improvement of the existing network. Figure 6.4 shows that regions of high information content are found exclusively in areas around major faults. This agrees with our intuition that coseismic rupture is best observed near faults and that observations of this phenomena are not well constrained in the very far field. We conclude that the single most important complement to the existing network is densification in the near-fault environment. It has previously been noted that similar GPS installations are similarly deficient (e.g., d’Alessio et al., 2005, §3.3).

The method presented in this study could be extended in a straightforward manner to other causes of ground surface deformation including volcanic, hydrological, interseismic, or postseismic forces.

The information criterion field in southern California shows variation along the plate boundary strike. In particular, the near-fault environment surrounding the San Jacinto Fault is the location that offers the greatest potential to improve the sensitivity of the southern California GPS network. This region of special interest is shown in greater detail in Figure E.8. The San Jacinto fault is one of the most active in southern California (Petersen and Wesnousky, 1994), and yet inspection of the current GPS network reveals that the area surrounding the San Jacinto fault has a relatively low density of GPS observation.

The information content of geodetic networks informs their optimal design. We have used this premise to offer recommendations concerning the continued maintenance of geodetic networks in the southern California GPS network. The southern California regional GPS network is an Earth observation tool of unprecedented calibre, and we show
Figure 6.4: Contour plot of information criteria of proposed sites in southern California. Warm colors suggest deficiencies in the existing network, e.g., locations where additional observations would most improve the information content of the existing network. The IC field is computed using the RSQSim synthetic earthquake catalogue. Not all faults are supported by the ten eigenfunctions used in this analysis. See text for discussion.
that it can be significantly improved by making more dense observations in the near-fault environment.
Chapter 7

Synthesis

Spaceborne measurement has revolutionized our understanding of Earth surface deformation. Previous generations of scientists doubted, for example, that strike-slip earthquakes could cause large postseismic deformations (Nur and Mavko, 1974) or that faults could act as barriers to groundwater flow (Iwamura, 1995) because they lacked modern datasets that clearly indicate such phenomena. These scientific misunderstands occurred not because previous workers were inferior scientists, but rather because they lacked the striking data sets that are becoming commonplace today.

Several themes that cross cut the chapters are now noted. Each chapter deals to varying degree with data, and each chapter must therefore account for measurement uncertainty, precision, and accuracy through statistical analysis. Rather than taking a passive role with statistical methods, these tools have been actively developed throughout. The statistics that are used are highly varied, ranging from the heavily model-based analysis of optimal geodetic network design to the “hypothesis-free” eigenfunction analysis used to
study correlated hydrogeology.

Both the study of network design and the study of strain transient detection (Chapters 6 and 4) involve global, constrained optimization problems. Both optimization problems have similar results in the following general sense. Although two local minima may be ordered (one has a smaller absolute objective function value), both may be reasonable solutions if they are convey similar information. Rather than noting only the most extreme value of the information criterion field in southern California, we are instead interested in the consistent pattern of extreme values. Similarly, rather than noting only the most separable part of a deformation data set as a transient event, we consider the family of events that share extreme values. In both of these cases, the “solution” to the optimization problem isn’t a single pattern, but rather it is the robust features that are persistent through the family of plausible solutions.

The data that are analyzed from the San Francisco Bay Area contain information concerning geological phenomena that pose hazards to society, including the presence of a previously unmapped fault details surrounding the slow collapse of levies in a sediment-starved delta. Dixon et al. (2006) notes that while scientific understanding favors statements emphasizing uncertainty, when speaking of hazards findings that convey clear risk to society it should be stated plainly. In summary of these findings, we note that:

- Levies in the San Joaquin Delta show highly variable subsidence rates. Levy subsidence is correlated with failure. Based on the analysis in Chapter 5, the safety of quickly subsiding levies should be verified with field inspection.

- The unmapped fault underneath Vacaville could pose significant risk to this city.
Fieldwork should be undertaken to confirm the age of most recent activity of this structure.

The computation approach favored throughout this thesis is to derive analytical expressions until these become unfeasible, and then to switch to numerical methods. This scheme has the advantage of maintaining the benefits of both numerical and analytical methods: analytical insights such as the relationships between parameters are retained, and yet the solutions are easily attainable and fast due to tactical use of numerical methods.

This thesis has presented a wide range of material that is underlain by an interest in Earth surface deformation geodesy. Although not strictly related, these topics nonetheless explore several outstanding challenges in deformation geodesy, and represent a contribution to this field.
Appendix A

Solid Mechanics Background

A.1 The governing equations of linear elasticity

The statement of conservation of momentum for an extended body is given by,

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} + \rho f_i.$$  \hspace{1cm} (A.1)

This equation introduces the notation that will be used throughout: $u_i$ is the displacement field, $\sigma_{ij}$ is the stress tensor field, $x_j$ are the spatial coordinates, $\rho$ is the density, $f_i$ is the body force field, and $t$ denotes time. The left hand side of Equation A.1 corresponds to inertial resistance to deformation. This term is responsible for the radiation of mechanical waves through a solid. For the present work, it suffices to neglect inertial terms. Conservation of momentum therefore implies the equilibrium equation,

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho f_i = 0$$  \hspace{1cm} (A.2)

For an isotropic elastic solid, the relationship between stress and strain is given
\[ \sigma_{ij} = \lambda \epsilon_{kk} \delta_{ij} + 2 \mu \epsilon_{ij} \]  

(A.3)

Here, \( \epsilon_{ij} \) denotes the strain tensor field, \( \delta_{ij} \) is the Kronecker delta function, and \( \lambda \) and \( \mu \) are the Lame elastic parameters. Since this is a linear relationship, it may be inverted to describe strains in terms of stresses. This is more easily represented with the elastic constants \( E \), Young’s modulus, and \( \nu \) Poisson’s ratio,

\[ \epsilon_{ij} = \frac{1}{E} \left[ (1 + \nu) \sigma_{ij} - \nu \delta_{ij} \sigma_{kk} \right] \]  

(A.4)

We assume that strains are small, and therefore the strain-displacement relationship becomes,

\[ \epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \]  

(A.5)

where this form arises by neglecting displacement gradients of order greater than unity.

Finally, we may combine Equations A.2, A.3, and A.5 to arrive at the elastostatic equation of motion,

\[ (\lambda + \mu) u_{j,ij} + \mu u_{i,jj} + \rho f_i = 0 \]  

(A.6)

In this last equation, the notation has been introduced that,

\[ \frac{\partial \cdot}{\partial x_j} = \cdot j \]  

(A.7)

Two simplified cases of the general three dimensional case are anti-plane strain and plane strain. Anti-plane strain is the condition that, \( u_1 = u_2 = 0, u_3 = u_3(x_1, x_2) \).

From Equation A.5, the only nonzero strains are therefore,

\[ \epsilon_{13} = \frac{1}{2} \frac{\partial u_3}{\partial x_1}, \epsilon_{23} = \frac{1}{2} \frac{\partial u_3}{\partial x_2} \]  

(A.8)
The equation of motion (Equation A.6) reduces to $\nabla^2 u_3 = 0$, and, introducing the bulk modulus $G$, Hooke’s Law yields $\sigma_{i3} = G\epsilon_{i3}$ ($i = 1, 2$). The function $u_3(x_1, x_2)$ may therefore be represented by an analytic function $\omega(z)$, $z = x_1 + ix_2$,

$$u_3 = G^{-1}\text{Im}[\omega(z)],$$

(A.9)

where Im denotes the imaginary part of a complex value. Stresses may therefore be represented as,

$$\omega'(z) = \sigma_{23} + i\sigma_{13}$$

(A.10)

A.2 The governing equations of poroelasticity

This section is concerned with a basic exposition of the governing equations of a fluid infiltrated elastic solid that undergoes quasi-static deformation. These relationships may be developed through multiple lines of reasoning (Biot, 1941; Rice, 1968; Wang, 2000; Segall, 2010). It is interesting to note, as pointed out by Rice and Cleary (1976), that images of an “array of grains” or a “solid skeleton” are not necessary, and, in fact, are somewhat contrary to a continuum viewpoint.

A.2.1 State variables and material parameterization

A poroelastic system may be described by two pairs of conjugate variables: stress and strain, and pore pressure and increment of fluid content. One variable from each pair may be chosen as an independent variable. Strain is defined as for a linear elastic solid.
The increment of fluid content, $\zeta$, is defined by,

$$
\zeta = \frac{\delta V_p - \delta V_f}{V} = \frac{\delta m_f}{\rho_f} = \frac{m_f - m_{f_0}}{\rho_{f_0}}
$$

(A.11)

The increment of fluid content describes the change in fluid mass content from a reference state, per reference volume. Because $\zeta$ is a normalized kinematic quantity, its introduction is somewhat analogous to the use of strain in classical elasticity. This quantity has the advantage that it eliminates density from the governing equations.

A poroelastic system is described by four material parameters. Following Rice and Cleary (1976), we use on a parameterization that focuses on the asymptotic drained and undrained states. The undrained state is the condition immediately after a sudden, quasi-static disturbance when no fluid flow has yet occurred. Two material properties that arise from this scenario are the undrained Poisson’s ratio,

$$
\nu_u = \frac{\partial \epsilon_i}{\partial \epsilon_j} |_{\zeta=0},
$$

(A.12)

and Skempton’s coefficient,

$$
B = \frac{\epsilon_{kk}}{\zeta} |_{\sigma=0} = \frac{p}{\sigma_{kk}} |_{\zeta=0}.
$$

(A.13)

### A.2.2 Constitutive relations

Before considering the complexity of the three dimensional constitutive equations, we first note a more simple case. For an isotropic applied stress field, $\sigma$, the constitutive relationships between the conjugate variables are given by,

$$
\begin{align*}
\epsilon &= a_{11}\sigma + a_{12}p \\
\zeta &= a_{21}\sigma + a_{13}p
\end{align*}
$$

(A.14)
The quantity $\epsilon$ is the volume strain, and generalized material parameters have been used to emphasize the simple form of the governing equations. This example is shown because it emphasizes that the independent and dependent variables are linearly related.

The three-dimensional isothermal pore pressure-stress-strain constitutive relationship, derived by Biot (1941) by symmetry and thermodynamic considerations, is given by,

$$2 \, G \epsilon_{ij} = \sigma_{ij} - \frac{\nu}{1+\nu} \sigma_{kk} \delta_{ij} + \frac{3(\nu_u - \nu)}{B(1+\nu)(1+\nu_u)} p \delta_{ij}$$

(A.15)

and

$$\zeta = \frac{3(\nu_u - \nu)}{2GB(1+\nu)(1+\nu_u)} \left[ \sigma_{kk} + \frac{3}{B} p \right]$$

(A.16)

The linear combination on the right hand side of this relationship generalizes the notion of effective stress to the three dimensional case.

The equations of strain compatibility ensure that the strain field induces a physically admissible displacement field. Rice and Cleary (1976) show that the strain compatibility equations in $\sigma_{ij}$ and $p$, six mutually independent equations, emit the relationship,

$$\nabla^2 \left[ \sigma_{kk} + \frac{6(\nu_u - \nu)}{B(1-\nu)(1+\nu_u)} p \right] = 0$$

(A.17)

The linear relationship between pore pressure gradients and fluid flux is known as Darcy’s law. This relationship is analogous to Fourier’s law (linearity between heat flow and temperature gradient) and Fick’s law (linearity between mass flux and concentration gradient).

The relationship is given by,

$$q_i = -\rho f \kappa \frac{\partial p}{\partial x_i}$$

(A.18)

and the conservation of fluid mass, stated as,

$$\frac{\partial q_i}{\partial x_i} + \frac{\partial m}{\partial t} = 0$$

(A.19)
we arrive at the diffusion equation,

\[ c \nabla^2 \left( \sigma_{kk} + \frac{3}{B} \rho \right) = \frac{\partial}{\partial t} \left( \sigma_{kk} + \frac{3}{B} \rho \right), \quad (A.20) \]

where \( c \) is Terzaghi’s coefficient of consolidation, given by,

\[ c = \frac{\kappa}{\eta} \left\{ \frac{2G(1 - \nu)}{1 - 2\nu} \left[ B^2(1 + \nu)^2(1 - 2\nu) \right] \frac{1}{9(1 - \nu_u)(\nu_u - \nu)} \right\}. \quad (A.21) \]

Since the quantity in brackets in Equation A.20 is proportional to the change in fluid mass content, we also have the simplified diffusion equation,

\[ \left( c \nabla^2 - \frac{\partial}{\partial t} \right) \zeta = 0. \quad (A.22) \]

**A.3 Uncoupled diffusion**

The general governing equations exhibit two-way coupling in the sense that gradients in pore pressure induce stresses in the solid phase and, also, stresses in the solid phase create gradients in pore fluid pressure. Under several special cases, however, the latter coupling is eliminated. This situation, in which coupling only occurs by the creation of solid elastic stresses pore pressure due to pore pressure gradients, is referred to as the uncoupled state. This state may occur if any of three general conditions is met (Wang, 2000):

1. The fluid phase is highly compressible in comparison to the solid phase.

2. The system is under uniaxial strain with constant vertical stress

3. The flow field is irrotational and exists in an infinite or semi-infinite medium without body forces
For the systems of interest, the first condition is not met. The second condition will be used to examine the cases of two one dimensional systems, and the third condition will be used to examine the situation of diffusion through a permeable fault zone.
Appendix B

Alternative solution to the governing equations of diffusion in composite material

Incremental fluid content diffusion in this geometry is described by the set of \( n \) coupled partial differential equations,

\[
\frac{\partial \zeta_k}{\partial t} = \alpha_k \frac{\partial^2 \zeta_k}{\partial x^2}
\]  \hspace{1cm} (B.1)

in \( x_{k-1} < x < x_k \), for \( t > 0 \) and \( k = 1, 2, \ldots, n \). The value of \( \alpha_k \) is the groundwater diffusivity constant.

Three types of boundary conditions are necessary. The first two are at the boundaries of the model: at the origin, a prescribed value and/or flux (e.g., Dirichlet-type, von Neumann-type, or both types) condition will be imposed; at infinity, incremental fluid mass
goes to zero. The former will allow more flexibility later, for example, by allowing either groundwater recharge or groundwater recharge rate boundary conditions. These conditions are given by,

\[ \alpha_0 \zeta(0, t) - \beta_0 k_1 \frac{\partial \zeta(0, t)}{\partial x} = \phi_0 \]
\[ \zeta(\infty, t) = 0 \]  \hspace{1cm} (B.2)

Boundary conditions are required at the interfaces between layers, at the points \( x = x_k \). The conditions at these points are essentially prescribed continuity or prescribed jump discontinuity.

\[ -\kappa_k \frac{\partial \zeta(x_k, t)}{\partial x} = h_k \left[ \zeta_k(x_k, t) - \zeta_{k+1}(x_k, t) \right] \]
\[ \kappa_k \frac{\partial \zeta(x_k, t)}{\partial x} = \kappa_{k+1} \frac{\partial \zeta(x_k, t)}{\partial x} \]  \hspace{1cm} (B.3)

The concept of a jump discontinuity in permeability at a boundary between dissimilar poroelastic materials has not been considered in the poroelasticity literature. The analogous quantity, \( h \), used in Equation B.3 is called the film coefficient, or the Biot number in the study of heat conduction, so named after the 19th century French physicist. The value \( h = \infty \) corresponds to the case of perfect hydraulic conductivity across the boundary.

Lastly, the initial condition is given by

\[ T_k(x, 0) = f_k(x) \]  \hspace{1cm} (B.4)

for \( t = 0 \), in the region \( x_{k-1} \leq x \leq x_k \), and \( k = 1, 2, \ldots, n \).
B.1 Solution formalism

A method for the solution of Equation B.1 is described by Mikhailov and Ozisik (1984). The general outline of the method of solution is familiar: the system is reduced to a homogeneous eigenvalue problem and then a solution is found that satisfies the boundary and initial conditions. In the study of heat conduction the analogous problem is often referred to as conduction in composite materials. Our goal is to present the most general solution that is relevant to the problems of geophysical interest.

We begin by taking considering the homogeneous eigenvalue problem given by,

$$ \frac{d\psi_k(\lambda, x)}{dx} = \frac{\lambda^2}{\alpha_k} \psi_k(\lambda, x) $$

(B.5)

in the region $x_{k-1} \leq x \leq x_k$, and $k = 1, 2, \ldots, n$. The boundary conditions are given by,

$$\begin{align*}
\alpha_0 \psi_1(\lambda, x_0) - \beta_0 \psi_1(\lambda, x_0) &= 0 \\
\kappa_k \psi_k(\lambda, x_k) &= \kappa_{k+1} \psi_{k+1}(\lambda, x_k) = h_k \psi_{k+1}(\lambda, x_0) - \psi_k(\lambda, x_k) \\
\kappa_n \psi_n(\lambda, x_k) &= \kappa_n \psi_n(\lambda, x_n) = 0
\end{align*}$$

(B.6)

The eigenvalues and eigenfunctions are $\lambda_k$ and $\psi_k$. The eigenfunctions can be written in terms of the arbitrary basis functions $u_k(\lambda, x)$ and $v_k(\lambda, x)$ as,

$$\psi_k(\lambda, x) = C_k u_k(\lambda, x) + D_k v_k(\lambda, x),$$

(B.7)

It is convenient to have the coefficients $C_k$ and $D_k$ take the values $C_k = \psi_k^{\ast} \equiv \psi_k(\lambda, x_{k-1})$ and $D_k = \psi_k^{\ast} \equiv \psi_k(\lambda, x_k)$. It can be shown that the correct basis functions for this
requirement are given by,

\[
U_k = \frac{\sin[\omega_k(x_k - x)]}{\sin[\omega_k(x_k - x_{k-1})]}
\]
\[
V_k = \frac{\sin[\omega_k(x - x_{k-1})]}{\sin[\omega_k(x_k - x_{k-1})]}
\]

(B.8)

where \( \omega_k = \lambda / \sqrt{\alpha_k} \).

The problem is now to find the appropriate eigenvalues and eigenfunctions. The eigenvalues are the characteristic roots of the equation \( \text{det}[K(\lambda)] = 0 \), where \( K \) is the banded matrix,

\[
\begin{pmatrix}
a_0 & b_1 & 0 & \ldots & 0 & 0 \\
b_1 & a_1 & b_2 & \ddots & \vdots \\
0 & b_2 & a_2 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
& b_{n-1} & a_{n-1} & b_n \\
0 & 0 & \ldots & 0 & b_n & a_n
\end{pmatrix}
\]

(B.9)

where,

\[
a_0 = \frac{\alpha_0}{\beta_0} - P_1(\lambda, x_0)
\]

\[
b_k = P_k(\lambda, x_k), \quad k = 1, 2, \ldots, n
\]

(B.10)

\[
a_k = Q_k(\lambda, x_k) - P_{k+1}(\lambda, x_k), \quad k = 1, 2, \ldots, (n - 1)
\]

\[
a_n = Q_n(\lambda, x_n) + \frac{\alpha_n}{\beta_n}
\]
The functions $P$ and $Q$ are given by,

\[ P_k(\lambda, x) = \begin{cases} 
-h_k & \text{if } x = x_k, \\
-\omega_k\kappa_k \cos[\omega_k(x_k - x)] \sin[\omega_k(x_k - x)] & \text{otherwise.} 
\end{cases} \]

\[ Q_k(\lambda, x) = \begin{cases} 
h_k & \text{if } x = x_k, \\
\omega_k\kappa_k \cos[\omega_k(x - x_k)] \sin[\omega_k(x_k - x_k - 1)] & \text{otherwise.} 
\end{cases} \quad (B.11) \]

It should be emphasized that the eigenvalue problem at hand is not approachable through conventional means, such as the Householder method (Wittrick and Williams, 1971; Mikhailov and Ozisik, 1984). We employ the Wittrick-Williams method, sometimes called the “sign-count” method, to determine the values of the $\lambda_i$. This numerical scheme is presented in Section B.2.

Once the eigenvalues are available, the eigenfunctions may be computed by the recurrence relationship,

\[ \psi_{k+1}^* = \frac{[P_{k+1}(\lambda, x_k) - Q_k(\lambda, x_k)]\psi_k^* - P_k(\lambda, x_k)\psi_{k-1}^*}{P_{k+1}(\lambda, x_k+1)} \quad (B.12) \]

The values of $\psi_0^*$ and $\psi_1^*$ will be determined depending on the boundary conditions.

### B.2 Numerical Calculations

The appropriate eigenvalues are determined using the method of Wittrick and Williams (1971). The Wittrick-Williams method uses the fact that the number of eigenvalues in the range from zero to $\tilde{\lambda}$ is given by,

\[ N(\tilde{\lambda}) = N_0(\tilde{\lambda}) + s[K(\tilde{\lambda})]. \quad (B.13) \]
B.3 Solution and error estimation

The solution is given by,

\[
ζ_k(x, t) = \left\{ \beta_n + \alpha_n \left( \frac{x_k - x}{\kappa_k} + \sum_{i=k+1}^{n} \frac{x_i - x_{i-1}}{k_i} + \sum_{i=k}^{n-1} \frac{1}{h_i} \right) \right\} φ_0 \\
+ \left\{ \beta_0 + \alpha_0 \left( \frac{x_k - x_{k-1}}{\kappa_k} + \sum_{i=1}^{k-1} \frac{k - x_{k-1}}{k_i} + \sum_{i=k}^{k-1} \frac{1}{h_i} \right) \right\} φ_n \right\} \times \left( \alpha_0 β_n + \alpha_n β_0 + \alpha_0 \alpha_n \sum_{k=1}^{n} \frac{x_k - x_{k-1}}{\kappa_k} + \alpha_0 \alpha_n \sum_{k=1}^{n-1} \frac{1}{h_k} \right)^{-1} \\
+ \sum_{i=1}^{∞} \frac{ψ_k(λ_i, x)}{N_i} e^{-λ_i^2 t} \left\{ \tilde{f}_i - \frac{1}{λ_i^2} \left( \phi_i Ω_i(x_0) + φ_i nΩ_i(x_n) \right) \right\}
\]

(B.14)

This solution introduces several helper functions:

\[
N_i = \sum_{k=1}^{n} \rho_k C_k \int_{x_{k-1}}^{x_k} ψ_k^2(λ_i, x) dx \\
\tilde{f}_i = \sum_{k=1}^{n} \rho_k C_k \int_{x_{k-1}}^{x_k} ψ_k(λ_i, x) f_k(x) dx \\
Ω_i(x_0) = \frac{ψ_1(λ_i, x_0) + k_1 ψ_1(λ_i, x_0)}{α_0 + β_0} \\
Ω_i(x_n) = \frac{ψ_n(λ_i, x_n) - k_n ψ_n(λ_i, x_n)}{α_n + β_n}
\]

(B.15)

Lastly, the global error involved in the solution is given by the expression,

\[
P_n(λ, x_0) ψ_{n-1}^* + \left[ Q_n(λ, x_n) + \frac{α_n}{β_n} \right],
\]

(B.16)

where, in the case of a perfect approximation, this value will exactly cancel.
Appendix C

One dimensional models of aquifer expansion

C.1 One dimensional unconfined aquifer

The most simple model of ground surface displacements due to groundwater flow at depth is described by one dimensional systems. We now consider the asymptotic states of a one dimensional unconfined aquifer. The resulting model will show the difference between a two layer-cake near surface models: one with a small water-saturated layer, and one with a large water-saturated layer. The difference between these models models the variation in ground surface displacements between the wet and dry seasons.

If we require that an elastic solid only be displaced in the $x_3$ direction and that variation in displacement only occur along that direction, then $u_3 = u_3(z) = u(z)$, and a one dimensional system results. By Equation A.5, the only non-zero strain is $\epsilon_{zz} = \partial u / \partial z = \ldots$
\( \varepsilon(z) \).

We assume a simple heterogeneous geometry to model the height of groundwater. The deformation of each layer is modeled individually as a layer that deforms under its overburden. The sum of the deformations of each layer equal the displacement at the surface. It is therefore only required to derive the deformations in a single layer.

We define the top and bottom of the \( i \)th layer to be the points \( z_{i-1} \) and \( z_i, z_0 = 0 \). The overburden acting on this layer is \( M_i = g \sum_{j<i} \rho_j (z_j - z_{j-1}) \) The appropriate boundary conditions state that there is zero displacement at the bottom of the layer, \( u(z_i) = 0 \), and that the stress at the top of the layer is equal to the overburden, \( \sigma_z(z_{i-1}) = M_i \). Elementary manipulations of the equation of motion (Equation A.7) yield,

\[
  u_i^z(z) = \frac{1}{9K + 12\mu} \left[ \frac{\rho_i g}{2} (z^2 - x_i^2) + M_i (z - x_i) + \rho_i g x_{i-1} (x_i - z) \right] \quad (C.1)
\]

The displacement at the surface is then given by

\[
  u_z(0) = \sum_{i=1}^{N} u_i^z(z_{i-1}), \quad (C.2)
\]

where \( N \) is the number of layers.

Using this model, we determine that unconfined aquifers are unlikely to cause deformations measurable using spaceborne geodetic techniques. The maximum displacements due to the poroelastic inflation of an unconfined aquifer under hydrostatic pressure is determined to be on the order of 0.1mm. Calculations were conducted using material parameters of Hart and Wang (1995) for the Berea sandstone: drained bulk modulus 6.6 GPa; undrained bulk modulus, 15.8 GPa; drained Poisson’s ratio, 0.17; and undrained Poisson’s ratio, 0.34.
C.2 One dimensional confined aquifer

We now use the method of solution for the case of a known pore pressure distribution.

We consider a layer of pore fluid pressure $P$ that has infinitely lateral extent. By symmetry, it suffices to only consider vertical deformations. We are interested in surface observations and so $x = (0, 0, z)$. The integral of equation C.3 then becomes,

$$
u_z(z) = \frac{3(\nu_u - \nu)}{G B(1 + \nu_u)(1 - 2\nu)} \int_V P(z) \frac{\partial g_z(z, \zeta)}{\partial \zeta} dV$$

$$= \gamma P \int_{D_1}^{D_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{z - \zeta_z}{(\zeta_x^2 + \zeta_y^2 + (z - \zeta_z)^2)^{3/2}} d\zeta_x d\zeta_y d\zeta_z$$

$$= 2\gamma P \int_{D_1}^{D_2} \int_{-\infty}^{\infty} \frac{(z - \zeta_z)^2}{(\zeta_y^2 + (z - \zeta_z)^2)} d\zeta_y d\zeta_z$$

$$= 2\pi \gamma P \int_{D_1}^{D_2} (z - \zeta_z) d\zeta_z$$

$$= 2\pi \gamma P \left[ z(D_2 - D_1) - \frac{(D_1^2 + D_2^2)}{2} \right]$$

Here we have introduced a constant, $\gamma$, for notational simplicity. This constant is given by,

$$\gamma = \frac{3(\nu_u - \nu)}{8\pi G^2 B(1 + \nu_u)(1 - \nu)}$$

As a rough estimate of the deformations due to a confined aquifer, we consider the material properties considered in the previous section. For an aquifer of thickness $D_2 - D_1 = 100m$, with an excess pore pressure sufficient to generate a 10m column of water, buried at 20m depth therefore generates mm of surface deformation.
Appendix D

Appendices to Correlated Hydrogeology

In this appendix we list several tables concerning the statistical comparison of the displacement equation and eigenfunction models discussed in Chapter 5.
Figure D.1: Sum of squared residuals between the detrended PS-InSAR time series and (A) the representation of the time series by the first three eigenfunctions, (B) the best fit time series by the displacement equation (Equation 5.2), $q = 2$, (C) Equation 5.2, $q = 3$, (D) Equation 5.2, $q = 4$. Areas of high misfit are more strongly spatially correlated in the displacement equation model and tend to occur in the areas of interest identified by the eigenfunction model (e.g., aquifers, and landfill regions).
Figure D.2: Histograms of model parameter estimates for the displacement equation, Equation 5.2, with \( q = 2 \). The amplitudes are generally distributed around zero (first two rows). The second-order shape of the amplitude distributions show two peaks, one on either side of zero. Two dominant frequencies are estimated (third row). The frequency of \( 2\pi \) (period of one year) is almost always present with a strongly peaked distribution. Other dominant frequencies include \( 4\pi \) and \( 4.3\pi \).
Figure D.3: Histograms of model parameter estimates for the displacement equation, Equation 5.2, with $q = 3$. The amplitudes are generally distributed around zero (first two rows). Three dominant frequencies are estimated (third row). The frequency of $2\pi$ (period of one year) is almost always present with a strongly peaked distribution. Other dominant frequencies include $4\pi$, $4.3\pi$, $\pi$, $\pi/2$, and $0$. 
<table>
<thead>
<tr>
<th>Model</th>
<th>q</th>
<th>$\chi^2/n$</th>
<th>f</th>
<th>$\chi^2/f$</th>
<th>$\omega_1$</th>
<th>$\omega_2$</th>
<th>$\omega_3$</th>
<th>$\omega_4$</th>
<th>$\omega_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>2</td>
<td>22.11</td>
<td>36n</td>
<td>0.61</td>
<td>2$\pi$</td>
<td>4$\pi$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>3</td>
<td>22.11</td>
<td>34n</td>
<td>0.65</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td>4$\pi$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>2</td>
<td>23.27</td>
<td>36n</td>
<td>0.65</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>4</td>
<td>22.11</td>
<td>32n</td>
<td>0.32</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td>4$\pi$</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>3</td>
<td>19.83</td>
<td>34n</td>
<td>0.58</td>
<td>$\pi$</td>
<td>2$\pi$</td>
<td>4$\pi$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>2</td>
<td>20.89</td>
<td>36n</td>
<td>0.61</td>
<td>$\pi$</td>
<td>2$\pi$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>4</td>
<td>19.83</td>
<td>32n</td>
<td>0.62</td>
<td>$\pi$</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td>4$\pi$</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>5</td>
<td>19.83</td>
<td>30n</td>
<td>0.66</td>
<td>$\pi$</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td>2$\pi$</td>
<td>4$\pi$</td>
</tr>
</tbody>
</table>

**F-Statistic**

<table>
<thead>
<tr>
<th>Model</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.08e-05</td>
<td>-</td>
<td>-</td>
<td>0.921</td>
</tr>
<tr>
<td>b</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.56e-04</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>c</td>
<td>-</td>
<td>0.892</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.420</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>d</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>e</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>f</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.907</td>
<td>-</td>
<td>0.428</td>
<td>-</td>
</tr>
<tr>
<td>g</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>h</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Confidence of null hypothesis**

<table>
<thead>
<tr>
<th>Model</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>b</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>c</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
</tr>
<tr>
<td>e</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>f</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>h</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Model</td>
<td>$\chi^2/n$</td>
<td>$f$</td>
<td>$\chi^2/f$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(1)</td>
<td>27.25</td>
<td>39(n - 2)</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(2)</td>
<td>20.32</td>
<td>38(n - 3)</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(3)</td>
<td>16.46</td>
<td>37(n - 4)</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(4)</td>
<td>14.93</td>
<td>36(n - 5)</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(5)</td>
<td>13.76</td>
<td>35(n - 6)</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(6)</td>
<td>12.89</td>
<td>34(n - 7)</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(7)</td>
<td>12.45</td>
<td>33(n - 8)</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**F-Statistic**

<table>
<thead>
<tr>
<th>Model</th>
<th>PC(2)</th>
<th>PC(3)</th>
<th>PC(4)</th>
<th>PC(5)</th>
<th>PC(6)</th>
<th>PC(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC(1)</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(2)</td>
<td></td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(3)</td>
<td></td>
<td></td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(4)</td>
<td></td>
<td></td>
<td></td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>PC(6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.47</td>
</tr>
</tbody>
</table>

**Confidence of null hypothesis**

<table>
<thead>
<tr>
<th>Model</th>
<th>PC(2)</th>
<th>PC(3)</th>
<th>PC(4)</th>
<th>PC(5)</th>
<th>PC(6)</th>
<th>PC(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC(1)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(2)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(3)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(4)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(5)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(6)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
</tbody>
</table>
Appendix E

Appendices concerning network design

E.1 Information content of simple models

Geodetic networks may be designed with the goal of making observations with high information content (IC). In this section we will present a series of applications with increasing complexity, beginning with a simple model of a one-dimensional gravity anomaly, and then proceeding to interseismic and coseismic deformation.

E.1.1 One dimensional gravity anomaly

When only a single observation is made and only a single model parameter is to be estimated, the information criterion reduces to the magnitude of the derivative, $|\frac{\partial y}{\partial x}|$. Consider the well-known gravity anomaly $y$ due to a spherical body of radius $R$, depth $b$ and
density contrast $\Delta p$ (Turcotte and Schubert, 2002). The anomaly observed at a horizontal distance $x$ from the body is given by,

$$y = \frac{4\pi}{3} G \Delta p \frac{R^3 b}{(x^2 + b^2)^{\frac{3}{2}}}$$

We will find the optimal observation point $x_o$ to observe the radius of the anomaly. We first take the parameter derivative,

$$\frac{\partial y}{\partial R} = 4\pi G \Delta p \frac{R^2 b}{(x^2 + b^2)^{\frac{3}{2}}}$$

and then find the maximum of this expression,

$$\frac{\partial^2 y}{\partial R \partial x} = 0 = -12\pi G \Delta p \frac{R^2 b x}{(x^2 + b^2)^{\frac{5}{2}}}$$

$$\Rightarrow x_o = 0.$$

This result indicates that maximum information about the anomaly’s radius occurs directly above the anomaly.

### E.1.2 Interseismic deformation

The well-known interseismic deformation model describes antiplane surface displacements at a perpendicular distance $Y$ from an infinite-length vertical fault with deep strike-slip dislocation $U$ and locking depth $D$ as (Savage and Burford, 1973),

$$y = M(Y, U_1, D) = \frac{U_1}{\pi} \tan^{-1} \frac{Y}{D} + \frac{U}{2}$$

The geometry of this model is shown in Figure E.1a.

Blewitt (2000) presents analytical expressions for the derivatives of Equation E.4 and finds the maxima of these functions. Specifically, it is shown that the maximal IC
for a one-observation network to changes in the locking depth parameter, $D$, occur at a distance away from the fault equal to the value of $Y = D$. It is also found that peak IC about slip occurs in the far field ($Y \rightarrow \infty$). The basic methods of calculus suffice to determine the extrema of sensitivity in this one-dimensional case. Applications involving more complicated and realistic models will require numerical methods of solution.

We begin our use of numerical methods with the present example so that we may compare the result to the known analytical solution. Our numerical method is the simple parameter-space grid search: the IC field is sampled at a finite number of locations and then interpolated. The present application seeks only the location of the highest IC without an existing network. Therefore configurations with only a single observation will be evaluated.

This method recovers the analytical solution, and graphical inspection is often an efficient and informative means for determining maxima. This agreement is shown in
Figure E.2, which shows the information about the locking depth parameter occurring at a distance $D$ away from the fault. This figure shows the interseismic surface deformation due to several models, each only different in locking depth. In each model the surface displacement directly on top of the fault is constant. A site placed in this location would therefore experience the same deformation due to each model, and so would be a poor choice of site location for the estimation of locking depth.

This notion of poor site choice is reflected by the zero valued IC about the locking depth directly atop the fault. Comparison between the black line (IC) shows that maximum IC occurs precisely where the difference between similar deformation models is the greatest. Blewitt (2000) has show by analytical methods that this distance is equal to the locking depth. This is a recurring theme in parameter sensitivity for fault mechanics studies: locations of peak sensitivity are often located a proportional distance away from the fault to the depth of the deformation source process.

The information field due to a single dislocation on a homogenous fault informs one’s intuition concerning more complicated scenarios. We therefore consider the deformation due to an idealized single homogenous fault. Ground surface displacements due to a single vertical strike-slip rectangular dislocation in a homogeneous elastic half-space are described by six model parameters (Okada, 1985b): the geographical coordinates of the observation point in the along-strike, $X$, and fault-perpendicular, $Y$, directions; the length of the fault trace, $L$; the upper and lower depths of the fault, $D_1$ and $D_2$; and slip magnitude, $U$. The geometry of this model is shown in Figure E.1b. Fault dip is fixed to be vertical, rake is right-lateral, strike oriented along the $x$-axis, and all three components of
Figure E.2: Displacements due to various fault models (orange and blue lines) and the derivative of the deformation model (black line). This figure shows that the region of maximal IC occurs where the derivative of the model with respect to the parameter of interest (in this case, locking depth) is greatest. The provisional model has the locking depth at 10 km. See text for accompanying discussion.

deformation are measured and analyzed.

E.1.3 Single observations with isotropic noise

Examination of single-site network configurations is useful to understand how IC fields change with changes in the provisional model parameters. The limiting case of a single observation and a single parameter gives a collapsed information criterion of the form,

\[ J = \left| \frac{\partial m(r)}{\partial x} \right|. \]  \hspace{1cm} (E.5)

In the two following examples, we will consider two parameters, slip and locking depth, that have dimension of length. The form of criterion of Equation E.5 is therefore a particular type of strain. In this particular case maximal IC coincides with maxima of this strain. We therefore expect that many of the properties of the parameter IC field should be shared in common with the elastic dislocation field.
Figures E.3 and E.4 show IC fields for fault slip and locking depth, respectively. Figure E.3 shows IC fields for vertical homogeneous strike-slip faults that extend 5 km in down dip extent, for the case in which slip magnitude is the estimated quantity. Various experiments are conducted that vary the upper depth of the fault. These two figures are useful because, by varying the estimated parameters, aspects of the general behavior of IC fields may be illustrated.

The resulting IC field shows symmetric “butterfly” patterns that are expected because of the simple relationship to the surface displacement field. Figure E.3 shows optimal slip observation occurs at a distance from the fault approximately proportional to the depth of the top of the fault, $D_1$, and in a direction that is nearly perpendicular to the fault. Particular cases of this generalization are shown in Figure E.3.

The IC fields in Figures E.3 and E.4 share several notable features. There is an observed geometric relationship that the depth of a particular fault feature is correlated with the optimal distance away from a fault to observe that feature. Peak IC to locking depth is an example. This IC maximum occurs at a perpendicular distance away from the fault that is slightly less than the locking depth. This IC field is shown in Figure E.4. The field is somewhat, but not exactly, similar to the analytical result achieved for the infinite-interseismic model considered in Section E.1.2. Whereas interseismic locking depth and coseismic slip are both best observed near one locking depth away from the fault, interseismic slip is best observed in the far field.

Information content decreases as the depth to a feature is increased. This corresponds with our intuition that a diminishing response is observed to phenomena at greater
Figure E.3: Information content for the magnitude of slip on a finite fault (black line) for faults with $D_1 = 100\text{m } (A)$, $D_1 = 1\text{ km } (B)$, $D_1 = 2\text{ km } (C)$, $D_1 = 3\text{ km } (D)$, $D_1 = 4\text{ km } (E)$, $D_1 = 5\text{ km } (F)$, $D_1 = 10\text{ km } (G)$, $D_1 = 15\text{ km } (H)$, $D_1 = 30\text{ km } (I)$, $D_1 = 45\text{ km } (J)$. For reference, the vertical and horizontal displacement fields are shown in subplots (K) and (L) that correspond to the deformations of (B). The units of the scale for (K) are mm, and the maximum displacement in (L) is 21cm. The total vertical extent of the fault is held constant as $D_2 - D_1 = 5\text{ km}$. A geodetic observation at any location in these plots will be endowed with the IC indicated by the color bar. All values are normalized to the maximum value of the field in (A).
Figure E.4: Information content for the locking depth of a finite fault (black line) with the same model geometries as in Figure E.3. The total vertical extent of the fault is held constant as $D_2 - D_1 = 5$ km. $U = 1.0$ m. A geodetic observation at any location in this plot will be endowed with the IC indicated by the color bar. The text presents extended discussion.
depth. Information content about changes in slip, for example, tend to be at least an order of magnitude greater than IC to changes in locking depth. This may be explained by the logic that, by definition, the locking depth parameter effects changes at the deepest extent of the fault. In comparison, slip acts uniformly from $D_2$ to $D_1$ on each fault patch, and therefore is associated with greater IC. Similarly, peaks of IC tend to become more broad as depth increases. This occurs as a result of an “elastic filter” that acts to spread out small changes at depth over a larger area on the surface.

IC fields become more symmetric and therefore less specific with increasing depth to source process. Several transitions occur as depth increases. All transitions are characterized by a decreasing influence of the free surface. As the fault is moved immediately away from the free surface, the pattern of the IC field changes from an elliptical region about the entire fault to a field with peaks near the fault’s center (Figures E.3a and b). The second transition occurs when the fault is buried to a depth that is roughly equivalent to the characteristic length scale of the fault. As this occurs, the maxima in the IC field become more focused away from the fault endpoints. Finally, as the fault is buried beyond the depth of several characteristic fault lengths, $90^\circ$ rotation symmetry is achieved. This final transition might be expected by consideration of St. Vernant’s principle, the rule-of-thumb that the influence of an elastostatic feature becomes indiscernible beyond several times its characteristic length.

E.2 Supplemental figures concerning the RSQSim dataset
Figure E.5: Eigenvalues of the data matrix from an RSQSim synthetic earthquake catalog.

Figure shows all eigenvalues in linear scale (top) and semi-log scale (middle), as well as the first 100 SVs (bottom, linear scale).
Figure E.6: The first four eigenvectors from the RSQSIm data set. Linear combinations of these slip patterns represent the simulated earthquakes. “Negative slip” is show in blue, which recalls the fact that each vector is only a single component of an actual simulated earthquake, not an earthquake in itself. The IC study presented in the text uses the first fifty singular vectors, not all of which are shown.
Figure E.7: The fourth eigenvector from the RSQSim data set. “Negative slip” is show in blue, which recalls the fact that each singular vector is only a single component of a synthetic slip distribution but not a synthetic slip distribution in itself.
Figure E.8: Information content in southern California. Warm colors suggest deficiencies in the existing network, e.g., locations where additional observations would most improve the information content of the existing network. The IC field is computed using the RSQSim synthetic earthquake catalogue. Not all faults are supported by the ten PCs used in this analysis. See text for discussion.
Bibliography


Dixon, T. H., F. Amelung, A. Ferretti, F. Novali, F. Rocca, R. Dokka, G. Sella, S.-W. Kim,


Ferretti, A., C. Prati, and F. Rocca (2000), Nonlinear subsidence rate estimation using


Hollenbeck, K. (1998), Invlap. m: A matlab function for numerical inversion of laplace transforms by the de hoog algorithm, unpublished work.


Langbein, J. (2008), Noise in GPS displacement measurements from Southern California


McTigue, D. F. (1987), Elastic stress and deformation near a finite spherical magma body:


Segall, P., and D. Schmidt (2003), Final report to UNAVCO and the National Science Foundation, Plate Boundary Observatory design studies, manuscript available at unavco.org.


