Durham E-Theses

Geodetic Network Design for Low-Cost GNSS

WATSON, ANDREW,ROBERT

How to cite:

WATSON, ANDREW,ROBERT (2019) *Geodetic Network Design for Low-Cost GNSS*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/13064/

Use policy

This work is licensed under a Creative Commons Attribution Share Alike 3.0 (CC BY-SA)

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

Geodetic Network Design for Low-Cost GNSS

Andrew Watson

A Thesis presented for the degree of
Master of Science by Research

Department of Earth Sciences
Durham University
United Kingdom

April 2019

Geodetic Network Design for Low-Cost GNSS

Andrew Watson

Submitted for the degree of Master of Science by Research

April 2019

Abstract: Global Navigation Satellite System (GNSS) stations are a cornerstone of modern geodetic surveys, providing high temporal-frequency, sub-centimetric three-component measurements of surface displacement at fixed locations. However, the high cost of each instrument limits both spatial resolution and access for small-scale users.

Low-cost GNSS stations, in particular single-frequency instruments, provide a cheaper alternative to conventional systems, enabling the deployment of larger GNSS networks. Increased observation density around continental fault zones would improve our ability to recover distributed aseismic slip, in particular afterslip, on continental faults, which may be poorly constrained by other geodetic techniques such as InSAR.

To best recover aseismic slip using low-cost GNSS stations, a method for the estimation of optimal network layouts is required. For single-frequency GNSS stations, which present the greatest potential for increased GNSS network density, the reduced positional accuracy as a result of ionospheric delay must also be addressed.

In this work, a method for the automated design of single- and dual-frequency GNSS networks to recover distributed aseismic slip on continental faults is presented. Network layouts are generated using particle swarm optimisation and a criterion matrix technique to minimise the uncertainties on modelled slip values, relative to "best possible" values. These are estimated through non-uniform fault discretisation, in which a multi-objective genetic algorithm is utilised to explore the trade-off between the complexity of the discretisation and the associated model uncertainties. The reduced positional accuracy of single-frequency GNSS stations is mitigated through the network design, and an understanding of the spatial structure of the ionospheric delay.

Initial results demonstrate the potential of low-cost GNSS stations, in particular single-frequency GNSS stations, to recover distributed aseismic slip on continental faults. Future work should expand the methodology to included slip across multiple faults, and the generation of mixed GNSS networks.

Declaration

The work in this thesis is based on research carried out in the Department of Earth Sciences at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification.

Copyright © 2019 Andrew Watson.

“The copyright of this thesis rests with the author. No quotation from it should be published without the author’s prior written consent and information derived from it should be acknowledged.”

Acknowledgements

A great degree of gratitude goes out to my primary supervisor, Richard Walters, who provided an abundance of both time and guidance well above what was required of him. I also extended my gratitude to Ken McCaffrey and Max Wilkinson who served as secondary supervisors, the former providing a much need geological perspective on the project, and the latter proving a technical background and the opportunity to develop links with GRL.

Many thanks to my parents for their unending support throughout my entire time at Durham University, and to my brother for his suggestions and experience as a postgraduate researcher. Finally, I thank Danni, for all her support over the last four years, especially when writing this Thesis stole most of my free time.

Dedicated to

My parents, my brother,

and

Danni

Contents

Abstract **iii**

List of Figures **xv**

1 Introduction **1**

1.1 Overview 1

1.2 GNSS and Network Design 4

1.3 Thesis Aims 8

2 GNSS Network Design **9**

2.1 Discrete Linear Inverse Problems 9

2.1.1 Fault Model 10

2.1.2 Data and Model Uncertainties 12

2.1.3 Ill-posed and Ill-conditioned Inverse Problems 14

2.2 Network Design 16

2.2.1 Background 16

2.2.2 Reliability 17

2.2.3 Economy 19

2.2.4 Precision 19

2.2.5 Optimisation Techniques 20

3 Methodology **25**

3.1 Fault Discretisation 26

3.1.1 Idealised Network 28

3.1.2 Algorithm 28

3.2 Network Design 32

3.3 Single Frequency GNSS 35

3.3.1 Characterising the Spatial Structure of the Ionosphere . . . 39

3.3.2 Removal of Long-Wavelength Ionospheric Trends 44

4 Results **51**

4.1 Dual-Frequency GNSS 51

4.1.1 Idealised Network Structure 51

4.1.2 An Example Normal Fault 52

4.1.3 Network size 57

4.1.4 Fault geometry 60

4.2 Single-Frequency GNSS 65

4.2.1 Example networks 65

4.2.2 Alternative Fault Geometries 68

4.2.3 Dual-Frequency Comparison 73

5 Discussion **77**

5.1 Summary of Results 77

5.1.1 Potential of Single-Frequency GNSS 77

5.1.2 Optimisation Repeatability 78

5.1.3 Network Redundancy 82

5.2 Methodological Limitations 84
5.2.1 Single Rectangular Fault 84
5.2.2 Removal of Long-Wavelength Ionospheric Trends 85
5.3 Future Work 86
5.3.1 Mixed GNSS Networks 88
5.3.2 Multi-Fault Systems 88

6 Conclusion 91

Bibliography 93

List of Figures

3.1 Schematic diagram of a discretised fault plane. L and W are the length and width of the slip patch, σ is the dip angle, and u_1,u_2, and u_3 are the components of slip on each patch. 27

3.2 An example idealised network centred around a westward-dipping normal fault. Station spacings are 2 km for the high density area, and 4 km for the low density area. The bold red line indicates the fault trace, while the red area shows the vertically projected fault plane . 29

3.3 An example Pareto front for discretisations with four rows, where each point is a solution. A log axis is used for the model error, as this increases exponentially with decreasing patch size, for which the number of patches is a proxy. The front itself is not uniform, as this is only an estimate of the true Pareto front and is limited by the performance of the algorithm. 32

3.4 Combined Pareto front produced for a synthetic normal fault. Each circle is a discretisation, with the colour indicating the number of rows in the discretisation, with the red circles indicating the combined Pareto front. 33

3.5 Diagram of four single-frequency GNSS stations (red dots) arranged in a rectangle and paired (black lines). Blue lines indicate separation distances. 38

3.6 Experimental variograms taken from Huang et al. (2017). 40

3.7 Variograms fitted to squared positional uncertainties taken from Chen, Kuo et al. (2015), with non-negative quadratic fits, for the North (top), East (middle), and Up (bottom) components. Corrected$_1$ and Corrected$_2$ refer to the data corrected with the ionospheric model and with the correction terms, respectively. 42

3.8 Gaussian fit to the squared standard deviation error (variance) for uncorrected single-frequency stations provided by Chen, Kuo et al. (2015), for the North (top), East (middle) and Up (bottom) components of displacement. The corrected variance values, along with the dual-frequency variance values, are provided for reference. Each component is shown separately so that the quality of the fit to the data is more easily seen. 46

3.9 Gaussian fit to the squared standard deviation error (variance) for uncorrected single-frequency stations provided by Chen, Kuo et al. (2015), shown up to the full range of the functions. The lines for the North and Up components overlap. 47

3.10 An example random field generated by the Gaussian ionospheric model for the East componenet. 47

3.11 One hundred experimental variograms (black lines) generated from random fields with a planar trend removed, and the best fit spherical models (red lines). 48

3.12 Correlation values for a network of four single-frequency stations, using a non-negative quadratic variogram (left) and a spherical variogram with the long-wavelength ionospheric trends removed (right). 49

4.1 An idealised network of dual-frequency GNSS stations that is used to produce $\mathbf{Q}_{d,ref}$ in the following examples. The bold red line indicates the fault trace, while the red area shows the vertically projected fault plane. 52

4.2 The combined pareto front (top) created from three optimisations. The chosen discretisation (bottom) is highlighted in black, with 1σ model uncertainties for each patch given in metres. 54

4.3 The 1σ model uncertainties (m) for both the idealised (top) and designed networks (bottom), with the final optimal network of seven single-frequency GNSS stations. 55

4.4 Full correlation matrix for the idealised network (left) and designed network (right). 56

4.5 Model correlation between the selected fault patch (grey) and all other patches, for the idealised network (top) and the design network (bottom). 57

4.6 Three optimal dual-frequency GNSS networks consisting of five, ten, and twenty stations, with the model uncertainties derived from $\mathbf{Q}_{m,des}$ for each network shown on the discretisation. 58

4.7 Distance metric vs the number of dual-frequency GNSS stations in the optimally designed network. 61

4.8 Optimal networks of 20 dual-frequency GNSS stations generated for the normal fault geometry described in Table 4.1 and rotated to strikes of 180° (Top), 90° (Middle), and 135° (Bottom). Estimated model uncertainties (m) are shown on the discretisation. 62

4.9 Optimal networks of ten dual-frequency GNSS stations designed around a normal fault with a 15 km bottom depth (top), a 3 km bottom depth (second from the top), and a strike-slip fault with the a bottom depth of 15 km (second from the bottom), and a 3 km bottom depth (bottom). Estimated model uncertainties are given in metres. 64

4.10 Ten station optimal single-frequency network for a normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. The dotted lines indicate the pairs of stations between which relative measurements of ground displacement are made. 66

4.11 Ten station optimal single-frequency network for a normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. The dotted lines indicate the pairs of stations between which relative measurements of ground displacement are made. 67

4.12 The network at nine evenly spaced iterations throughout the inversion, with iterations where the distance metric did not change removed. The final plot (iteration 96) is the optimal network. Iteration number and distance metric value for each network layout is given. Dotted lines indicate station pairings. 69

4.13 Ten station optimal single-frequency network for a vertical strike-slip fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. . 70

4.14 Ten station optimal single-frequency network for a shallow (3km bottom depth) normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. Note that the x-axis of the network has been stretched so as to make the layout clearer. 71

4.15 Ten station optimal single-frequency network for an east-west striking normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. 72

4.16 Optimal networks and estimated model uncertainties for a 7 station dual-frequency network (top), a 7 station single-frequency network (middle), and a 12 station single-frequency network (bottom). . . 74

4.17 Model correlations from the single- and dual-frequency networks shown in Figure 4.16. 75

4.18 $\mathbf{Q}_{d,des}$ for the single- and dual-frequency networks shown in Figure 4.16. 75

5.1 Results from five runs of the MOGA for a two row discretisation of a normal fault with bottom depth of 3 km (e.g. Figure 4.14). Solutions are grouped by colour, with marker size varying so that overlapping solutions are visible. 79

5.2 Optimal single-frequency networks designed around a east-west striking normal fault. The maximum model uncertainty σ_{max} and distance metric values are provided. Networks are ordered from worst to best by distance metric, where one is the best and six the worst. Note that markers for two stations overlap in networks 3 and 4. 81

5.3 A comparison of the redundancy values for the dual- (left) and single-frequency (right) optimal networks shown in Figure 4.16. Values are for each station (circles) for dual-frequency, and each pair of stations (lines) for single-frequency. 83

5.4 A repeat of the network optimisation shown in Figure 4.15 for ten single-frequency GNSS stations around a east-west striking normal fault. The model uncertainties derived from $\mathbf{Q}_{m,ref}$ (Top) and $\mathbf{Q}_{m,des}$ (bottom) are displayed on the discretisation. 87

Chapter 1

Introduction

1.1 Overview

Global Navigation Satellite System (GNSS) stations are a key component of geodetic monitoring, allowing high temporal-frequency, subcentimetric three-component measurements of ground motion at fixed locations (Groves 2013). Deformation of the Earth's crust arising from a range of geophysical phenomena, such as earthquake and tectonic processes (Walker et al. 2004; Mazzotti et al. 2005; Reilinger et al. 2006; Hill et al. 2012; Karakhanyan et al. 2013; Kreemer et al. 2014; Bekaert et al. 2016; Metzger et al. 2017; Walters, England et al. 2017; Wilkinson et al. 2017; Scognamiglio et al. 2018), landsliding (Eyo et al. 2014; Cina et al. 2015; Bellone et al. 2016), and active volcanism (Janssen et al. 2002; Bartel et al. 2003) may be measured.

The last few decades have seen the development of large permanent GNSS networks such as the Plate Boundary Observatory (PBO) in the United States, the GPS Earth Observation Network (GEONET) in Japan, and the Central Andean Tectonic Observatory (CAnTO) Project in Chile and Peru. Temporal resolutions are high due to a measurement rate of up to 1 Hz, although spatial resolutions are commonly low relative to critical short-wavelength deformation processes such as localised fault slip (e.g. GEONET possesses an average station spacing of 20 km (Sagiya 2004)).

Interferometric Synthetic-Sperture Radar (InSAR) is a complementary geodetic technique that uses satellite-based radar to measure ground displacement (Elliott et al. 2016; Hamling et al. 2017; Hussain et al. 2018). Spatial resolutions are significantly higher than GNSS, with the newest Sentinel-1 satellites imaging at a 20 m resolution over 250 km wide swaths as standard (Torres et al. 2012). However, the temporal resolution is much lower, with a minimum repeat time of six days for Sentinel-1, and in the order of months for older instruments such as Envisat (Roustaei et al. 2010; Walters, Elliott et al. 2013). Measurements are also one-dimensional, in the satellite's line-of-sight, which makes InSAR only weakly sensitive to the north-south component of ground motion, as this is near-parallel to the satellite's orbital path.

High instrument cost is a primary cause of low spatial resolutions in GNSS networks. Low-cost GNSS stations, utilising “mass-market” receivers and antennas, allow for larger and denser networks at an equivalent cost (Günther et al. 2008; Biagi et al. 2016; Wilkinson et al. 2017). These devices suffer diminished accuracy compared to the specialist GNSS systems that are conventionally used in permanent and campaign networks, especially for single-frequency receivers (Takasu et al. 2008; Rademakers et al. 2016). Low-cost GNSS instruments enable the deployment of high-density networks to become practical, which provide higher spatial resolution while retaining high temporal resolution and three-component measurements. This is of particular interest for tectonic geodetic investigations, where measurements of surface deformation are used to estimate slip distributions on tectonically-active faults. Increased observation density around the fault allows estimated slip values to be better constrained. This is true for both seismic and aseismic slip, the latter of which cannot be observed using seismometers, and which may be poorly sampled by InSAR if the slip occurs at a higher temporal frequency than the repeat time of the satellite (Peng et al. 2010).

Single-frequency GNSS stations are the cheapest form of low-cost GNSS. The lack of a second frequency reduces software and hardware costs at the expense of positional accuracy and the ability to correct for ionospheric delay within independent stations. Instead, alternative methods such as pairing stations, ionospheric models (Bartel et al. 2003), or correction terms from reference dual-frequency stations (Chen, Kuo et al. 2015; Janssen et al. 2002) must be used, resulting in larger data uncertainties. The application of single-frequency GNSS stations to landslide (Tu et al. 2013; Eyo et al. 2014; Benoit et al. 2015; Cina et al. 2015; Bellone et al. 2016 and volcano monitoring (Janssen et al. 2002; Bartel et al. 2003; Lee et al. 2015) is well documented, however, their use in studying tectonic deformation is limited (Wilkinson et al. 2017). The further cost reduction over low-cost dual-frequency GNSS stations would allow for even denser networks than what can be practically achieved with conventional GNSS. This could lead to further improvements in our ability to study tectonic deformation, albeit with some limitations, if a method of mitigating positional accuracy can be found.

Coseismic slip has previously been studied using low-cost single-frequency GNSS stations. Wilkinson et al. (2017) used four stations, placed in pairs on either side of the fault break, to observe the 2016 Vettore earthquake in Italy. The short duration of coseismic slip means that many GNSS error sources have negligible impact on the observations. For example, tropospheric and ionospheric effects, the temporal variation of which contributes to positional uncertainties, may be considered static for the duration of the seismic event. In contrast, aseismic slip may occur over days or weeks. This requires a switch to average daily solutions for each station, so that the smaller tectonic signal (relative to coseismic slip) may be better isolated from the noise. Doing so introduces several error sources, of which ionospheric delay is the most important for single-frequency stations, which must now be corrected for. Daily solutions still provide a higher temporal resolution than InSAR, and so GNSS is preferential for studying aseismic slip.

Distributed aseismic slip on a fault may be estimated from geodetic observations of surface displacement using a discrete linear inversion (Aster et al. 2011). The inversion maps displacement measurements to estimated slip values by way of a design matrix. The design matrix contains Green's functions (e.g. Okada (1985)) that describe the relationship between slip on a fault plane and surface displacement. Faults are commonly modeled as multiple slipping patches so that distributed slip can be studied (e.g. Scognamiglio et al. (2018)). The data uncertainties, data locations, and the dimensions of the slipping patches all contribute to the uncertainties on the estimated slip values associated with these patches. The optimal design of both fault discretisations and GNSS networks is therefore important to studies of aseismic slip.

To investigate the potential of (single-frequency) low-cost GNSS stations to study aseismic slip on continental faults, we require a method of estimating optimal network layouts that best recover distributed slip using daily measurements of surface displacement. These are networks which produce the lowest model slip uncertainty across the greatest number of patches. This can be done through geodetic network design, a well-established technique within geophysics, originating from the work of Baarda (1973) and Grafarend and Schaffrin (1974). A range of methods exist depending upon the design context and the instruments involved, not being exclusive to GNSS stations (Amiri-Simkooei 2004; Berné et al. 2004; Doma 2013; Mahapatra et al. 2015; Sathiakumar et al. 2017; Klein et al. 2017). While network design for dual-frequency GNSS stations has been well studied, the same cannot be said for large-scale single-frequency GNSS networks, which are often deployed based on intuition and in small numbers (e.g. Chen, Kuo et al. (2015)).

1.2 GNSS and Network Design

GNSS refers to the collection of satellites which provide satellite navigation services, and their implementation. GNSS stations provided three-component spatial locations through trilateration with GNSS satellites. Vertical positional accuracy (altitude) is generally 1.5 to 3 times worse than horizontal positional accuracy (latitude-longitude), a result of satellite-receiever geometry (Groves 2013). GPS signals are transmitted over multiple wavebands, with the encoded information varying between them. Wavelengths were chosen based upon availability and the trade off between high (requires additional power, greater precipitation attenuation) and low (greater delay, localised attenuation from atmosphere) frequencies. Conventional survey GNSS stations are typically dual-frequency, meaning they contain the hardware and software necessary to process both the L1 (1575.42 MHz, 190 mm wavelength) and L2 (1227.60 Mhz, 244 mm wavelength) bands. This extends to low-cost dual-frequency GNSS, but not to low-cost single-frequency GNSS stations, which are only able to use the L1 band. The L2 band contains a greater amount of information, providing increased positional accuracy. However, processing this information requires proprietary algorithms, along with the hardware required to process a second frequency, increasing instrument costs. A second frequency also aids in mitigating certain errors, including ionospheric (described below), multipath, and cycle slip. Single-frequency stations are more susceptible to these errors, have decreased positional accuracy, and suffer more from carrier phase ambiguity (Brown et al. 2006; Takasu et al. 2008).

The task of generating optimal single- and dual-frequency low-cost GNSS networks can be broken down into a number of sub-problems. First, the ‘best-possible’ amount of retrievable slip information must be established, assuming perfect data, to provide a reference for the network design. This is achieved through discretisation of the fault plane using an idealised network of dual-frequency GNSS stations. Optimal GNSS networks can then be designed which attempt to mimic the results of the idealised network. In the case of single-frequency GNSS, the impact of ionospheric delay must also be mitigated and included within the network design.

Fault discretisation describes the splitting of the fault plane into the slip patches that make up the model parameters of our linear inversion. Commonly, a uniform grid is employed with interval size based upon the model resolution, observation noise, or expected magnitude (Wright et al. 2003; Funning et al. 2005; Hill et al. 2012; Sathiakumar et al. 2017; Scognamiglio et al. 2018). This method is sub-optimal as the displacement field decays proportional to the inverse square of the distance (Aki et al. 2002), which reduces the resolving power of the observations at depth. Poor sizing may also produce artifacts and sharp discontinuities, leading to unrealistic slip distributions (Page et al. 2009). A non-uniform discretisation, where patch size is proportional to depth, better represents the resolving power of the data. These may be generated manually (Pritchard et al. 2002; Custódio et al. 2009; Page et al. 2009), although automated methods are more efficient and likely to produce better results (Barnhart et al. 2010; Atzori et al. 2011; Hayes et al. 2014; Metzger et al. 2017). Regularisation, often in the form of smoothing, is commonly employed to reduce the model slip uncertainty on each patch. However, this comes at the cost of degraded model resolution (slip values become spread across multiple patches and are poorly constrained spatially) and the biasing of the inversion.

In this thesis, a new method for automated fault discretisation is presented, using a Multi-Objective Genetic Algorithm (MOGA) to minimise model error while maximising the number of slip patches in the discretisation. MOGAs function as an analogy of natural selection, and are capable of solving nonlinear optimisation problems with multiple objectives (Blum et al. (2003) and Konak et al. (2006)). Model error is negatively correlated with the number of patches, and so the algorithm produces a range of solutions based upon the relative weightings of these two objectives. By using optimisation to fit patch dimensions to the resolving power of the data, the use of regularisation and its associated problems may be avoided.

To estimate model uncertainty as part of the network design, knowledge of the structure of the data uncertainties is required. For dual-frequency GNSS stations, where the ionospheric delay may be removed in-station, common estimates of data uncertainty are used and each station is treated as an independent measure of surface displacement, with negligible covariance. For single-frequency stations the ionospheric delay cannot be removed, requiring that the stations are treated as paired relative observations of displacement, and that the structure of the ionospheric delay in known.

Ionospheric delay is caused by charged particles in the ionosphere interfering with the electromagnetic waves transmitted by GNSS satellites. This causes ionospheric divergence, where the signal information (ranging code, navigation data) is delayed and the carrier phase advanced. The number of charged particles, quantified as the total electron count, varies both spatially and temporally, with peaks in low- to mid-latitudes and during dawn and dusk. Variograms, which describe the spatial variation of ionospheric delay between two points, can be used to estimate the ionospheric delay and therefore the data uncertainty for single-frequency GNSS stations (Chen, Kuo et al. 2015; Huang et al. 2017).

Once a fault discretisation has been chosen and the best possible model uncertainties estimated from the idealised network, the layouts of optimal GNSS networks may be estimated. The quality of a given network is determined using a criterion matrix method (Mahapatra et al. 2015) paired with a distance metric (e.g. Förstner et al. (2003)). This expresses, as a single number, the difference in the model uncertainties and covariances between the designed network and the idealised network. A smaller distance metric value indicates that the network is approaching the "best possible" model uncertainty values than can be achieved.

Optimal networks are designed using particle swarm optimisation, a metaheuristic Evolutionary Algorithm (EA) inspired by swarm theory (Eberhart and Kennedy 1995). The locations of any fixed number of GNSS stations are optimised for simultaneously, both for single- and dual-frequency receivers, by minimising the distance metric. Previous works have utilised a grid search technique, where stations are tested in a discrete number of locations (Mahapatra et al. 2015; Klein et al. 2017; Sathiakumar et al. 2017). While simple to implement, this method is unable to test for optimal locations that lie between these points and can be time intensive. As a continuous function, PSO is not limited to gridded locations, which improves its ability to find true optimal solutions. It is also capable of handling the strong non-linearity and localisation of the design problem.

1.3 Thesis Aims

The aims of this thesis work are summarised as follows:

1. Implement a method for automated fault discretisation, utilising a MOGA to avoid the use of regularisation. Combined with a network of idealised "best possible" observations, this will be used to produce the lowest model uncertainties possible, that will act as the target values for optimally designed networks.
2. Estimate the spatial structure of the single-frequency GNSS data through the understanding of the spatial variability of the ionospheric delay.
3. Develop a method for estimating the optimal layout of low-cost single- and dual-frequency GNSS networks to recover distributed aseismic slip on continental faults.

Chapter 2

GNSS Network Design

Geodetic network design, for the purpose of recovering distributed aseismic slip on continental faults, is underpinned by the relationship between surface displacement and slip on a fault. This relationship takes the form of a *discrete linear inverse problem*, which is described below.

A range of methods for the optimal design of geodetic networks exist, depending upon the context of the design problem. Several of these methods, including both network objectives and optimisation algorithms, are summarised in the following Chapter.

2.1 Discrete Linear Inverse Problems

In general terms, the relationship between some data, d and a model, m can be described

$$G(m) = d \tag{2.1.1}$$

where G is a function based upon an understanding of the underlying physics relating d and m. As noise is inherent to physical observations, we expand Equation (2.1.1)

$$G(m_{true}) + \eta = d \tag{2.1.2}$$

where η is the data noise and m_{true} is the true model parameters.

Equation (2.1.1) may also be expressed as a discrete linear inverse problem

$$\mathbf{d} = \mathbf{G}\mathbf{m} \tag{2.1.3}$$

where **G** is the $m \times n$ *design matrix* that describe the relationship between the m-vector of observations, **d** and the n-vector of model parameters, **m** (Aster et al. 2011).

The expanded matrix notation for Equation (2.1.3), which highlights the discrete nature of the inverse problem, is given below.

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix} = \begin{bmatrix} f(d_1, m_1) & f(d_1, m_2) & \cdots & f(d_1, m_n) \\ f(d_2, m_1) & f(d_2, m_2) & \cdots & f(d_2, m_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(d_m, m_1) & f(d_m, m_2) & \cdots & f(d_m, m_n) \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix}$$

In the context of aseismic slip occurring on a fault plane, **d** is a vector of surface displacement values observed at discrete locations, **m** is a vector of slip values that describe the distributed slip on the fault plane, and **G** is a series of Green's functions that describe the linear relationship between each slip value, and the observed surface displacement. These functions are dependent upon the geometry of the fault model, and the locations of the data.

2.1.1 Fault Model

The fault is modelled as a finite rectangular half-space, as describe by Okada (1985). The Green's functions, which are used to build **G**, describe the surface deformation generated by slip of a fixed rake and magnitude on a rectangular fault plane. Multiple rectangular half-spaces may be mosaiced to define a single fault (e.g. Metzger et al. 2017; Sathiakumar et al. 2017), to model distributed slip. Going forward, each rectangular half-space in the model is referred to as a *slip patch*.

We desire a method for generating optimal GNSS stations layouts to model slip at any point on the fault plane. The location and magnitude of distributed slip is now known, and so **G** is generated by assuming a slip with a uniform rake and magnitude of 1 m. The tensile component of fault slip is assumed to be negligible, as is common with tectonically-driven earthquakes, and the ground surface to be a flat plane at $z = 0$.

The strike-slip components of the surface displacement are given as

$$u_x = -\frac{U_1}{2\pi}\left[\frac{\xi q}{R(R+\eta)} + tan^{-1}\frac{\xi\eta}{qR} + I_1 sin\,\delta\right] \tag{2.1.4}$$

$$u_y = -\frac{U_1}{2\pi}\left[\frac{\tilde{y}q}{R(R+\eta)} + \frac{q\ cos\,\delta}{R+\eta} + I_2 sin\,\delta\right] \tag{2.1.5}$$

$$u_z = -\frac{U_1}{2\pi}\left[\frac{\tilde{d}q}{R(R+\eta)} + \frac{q\ sin\,\delta}{R+\eta} + I_4 sin\,\delta\right] \tag{2.1.6}$$

$$\tag{2.1.7}$$

and the dip-slip components of the surface displacement as

$$u_x = -\frac{U_2}{2\pi}\left[\frac{q}{R} - I_3 sin\,\delta\ cos\,\delta\right] \tag{2.1.8}$$

$$u_y = -\frac{U_2}{2\pi}\left[\frac{\tilde{y}q}{R(R+\xi)} + cos\,\delta\ tan^{-1}\frac{\xi\eta}{qR} - I_1 sin\,\delta\ cos\,\delta\right] \tag{2.1.9}$$

$$u_z = -\frac{U_2}{2\pi}\left[\frac{\tilde{d}q}{R(R+\xi)} + sin\,\delta\ tan^{-1}\frac{\xi\eta}{qR} + I_5 sin\,\delta\ cos\,\delta\right] \tag{2.1.10}$$

$$\tag{2.1.11}$$

where

$$I_1 = \frac{\mu}{\lambda+\mu}\left[\frac{-1}{cos\,\delta}\frac{\xi}{R+\tilde{d}}\right] - \frac{sin\,\delta}{cos\,\delta}I_5 \tag{2.1.12}$$

$$I_2 = \frac{\mu}{\lambda+\mu}\left[-ln(R+\eta)\right] - I_3 \tag{2.1.13}$$

$$I_3 = \frac{\mu}{\lambda+\mu}\left[\frac{1}{cos\,\delta}\frac{\tilde{y}}{R+\tilde{d}} - ln(R+\eta)\right] - \frac{sin\,\delta}{cos\,\delta}I_4 \tag{2.1.14}$$

$$I_4 = \frac{\mu}{\lambda+\mu}\frac{1}{cos\,\delta}\left[ln(R+\tilde{d}) - sin\,\delta\ ln(R+\eta)\right] \tag{2.1.15}$$

$$I_5 = \frac{\mu}{\lambda+\mu}\frac{2}{cos\,\delta}tan^{-1}\frac{\eta(X+q\ cos\,\delta)+X(R+X)\ sin\,\delta}{\xi(R+X)\ cos\,\delta} \tag{2.1.16}$$

(2.1.17)

and

$$p = y\ cos\,\delta + d\ sin\,\delta \tag{2.1.18}$$

$$q = y\ sin\,\delta - d\ cos\,\delta \tag{2.1.19}$$

$$\tilde{y} = \eta\ cos\,\delta + q\ sin\,\delta \tag{2.1.20}$$

$$\tilde{d} = \eta\ sin\,\delta + q\ cos\,\delta \tag{2.1.21}$$

$$R^2 = \xi^2 + \eta^2 + q^2 = \xi^2 + \tilde{y}^2 + \tilde{d}^2 \tag{2.1.22}$$

$$X^2 = \xi^2 + q^2 \tag{2.1.23}$$

u_x, u_y, and u_3 are the components of displacement relative to the slipping patch, as shown in Figure 3.1, and caused by U_1, U_2, and U_3, which are the dislocations corresponding to the strike-slip, dip-slip, and tensile components of a dislocation at (ξ_1, ξ_2, ξ_3). L and W are the stike-parallel length and dip-parallel widths respectively, while δ is the dip angle of the fault, and λ and μ are the lameṕarameters of the elastic half-space, which are both assigned a value of 3.23×10^{10}.

2.1.2 Data and Model Uncertainties

The error and joint variability of both the observations, **d** and the model parameters, **m** can be quantified through two statistical properties: *variance*, and *covariance.*

The variance of a random variable, X describes the spread of the variable around it's expected value, *E[X]* (i.e. the mean)

$$Var(X) = E\left[(X - E\left[X\right])\right] \tag{2.1.24}$$

$$= \sigma_X^2 \tag{2.1.25}$$

where σ is the standard deviation of the variable.

The covariance of two random variables, X and Y is a measure of the joint variability

$$Cov(X,Y) = E\left[(X - E\left[X\right])(Y - E\left[Y\right])\right] \tag{2.1.26}$$

$$= \rho_{X,Y}\sqrt{Var(X)Var(Y)} \tag{2.1.27}$$

where $\rho_{X,Y}$ is the correlation coefficient, also known as the normalised covariance, which is a dimensionless measure of the linear relationship between X and Y. If both variables are independent, then $\rho_{X,Y} = 0$ and $Cov(X,Y) = 0$. The variance of a variable can also be expressed as the covariance with itself, where $\rho_{X,Y} = 1$ and $Var(X) = Var(Y)$.

Expanding X to a set of n independent random variables, the variances and covariances for each can be expressed as a *Variance-Covaraince Matrix* (VCM)

$$\mathbf{Q_X} = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{cccc|c} X_1 & X_2 & \dots & X_n & \\ \hline Var(X_1) & Cov(X_1,X_2) & \dots & Cov(X_1,X_n) & X_1 \\ Cov(X_2,X_1) & Var(X_2) & \dots & Cov(X_2,X_n) & X_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Cov(X_n,X_1) & Cov(X_n,X_2) & \dots & Var(X_n) & X_n \end{array} \tag{2.1.28}$$

where $\mathbf{Q_X}$ is a square symmetrical matrix, as $Cov(X,Y) = Cov(Y,X)$.

We establish two main VCMs of interest: the data VCM $\mathbf{Q}_d$, which describes the uncertainties on our observations, and the model VCM $\mathbf{Q_m}$, which describes the uncertainty on the slip values associated with each slip patch. $\mathbf{Q_d}$ is created from an *a priori* understanding of the GNSS positional uncertainties, which will differ between the single- and dual-frequency instruments.

$\mathbf{Q_m}$ may be derived from the design matrix and the data uncertainties

$$\mathbf{Q_m} = (\mathbf{G^T W_e G})^{-1} \tag{2.1.29}$$

where $\mathbf{W_e} = inv(\mathbf{Q}_d)$, which is termed the weight matrix. This is derived from the *weighted least squares* solution to the discrete linear inverse problem:

$$\mathbf{m}_{LS} = (\mathbf{G^T W_e G})^{-1}\mathbf{G^T W_e d} \tag{2.1.30}$$

where $\mathbf{m}_{LS}$ is the estimated model parameters. This is one method of solving a lienar inverse problem where no model perfectly fits the observations, and where the magnitude of the data uncertainties vary.

The data uncertainties are propagated through to the model uncertainties, based upon the geometry of the fault model and data locations. This is fundamental to the network design, as it means that model uncertainty can be minimised by finding the optimal locations for the GNSS stations. The fault discretisation will also affect the model uncertainties due to the propagation of the data errors through the model geometry.

2.1.3 Ill-posed and Ill-conditioned Inverse Problems

When calculating $\mathbf{Q}_m$, there are two problems that may be encountered with the structure of the discrete linear inverse problem itself: ill-posedness, and ill-conditioning.

Illl-posedness, also known as non-uniqueness, describies the situation where multiple solutions are able to exactly satisfy Equation (2.1.3). This occurs when the number of model parameters is greater than the number of independent observations ($n > m$). In this case, the data cannot uniquely constrain each model parameter, and so variations in one parameter may be offset by changes in other parameter. This results in an infinite number of solutions. However, ill-posedness can still occur when the number of observations is greater than the number of model parameters ($m > n$), if the equations within $\mathbf{G}$ are not linearly independent (i.e. $\mathbf{G}$ is rank-deficient). In this case, some observations may be linearly dependent, and so the true number of independent observations in less than m.

Ill-posed inverse problems may be uniquely solved by applying regularisation, often in the form of smoothing, and which is common in geodetic studies of aseismic slip (Barnhart et al. 2010; Atzori et al. 2011; Hayes et al. 2014; Metzger et al. 2017). Regularisation provides additional data and constrains the model parameters. This reduces the model uncertainties, but can also degrade the model resolutions, which is expressed

$$\mathbf{R} = \mathbf{G}^{-g}\mathbf{G} \tag{2.1.31}$$

where $\mathbf{R}$ is the model resolution matrix, and $\mathbf{G}^{-g}$ is the MoorePenrose pseudoinverse of the design matrix. $\mathbf{R}$ expresses how well the observations resolve the model parameters. An example resolution matrix for a model with three parameters is given below.

$$\mathbf{R} = \begin{array}{cccc} m_1 & m_2 & m_3 & \\ 1 & 0 & 0 & m_1 \\ 0 & 0.8 & 0.2 & m_2 \\ 0 & 0.2 & 0.8 & m_3 \end{array}$$

For m_1, the only non-zero value in the column and row that it is associated with is a 1 on the main diagonal. This expresses that the model parameter is fully resolved by the data. For m_2 and m_3, there exists non-zero values in the off-diagonal elements, showing that the values for both model parameters are a weighted average of each other. For slip patches on a fault plane, this is seen as slip values becoming smeared over multiple patches. If regularisation is used, and the model resolution degraded, then this can be corrected by scaling the size of the model parameters to match the resolution length of the data, which is the smallest dimension that the data can perfectly resolve. It is desired that $\mathbf{R} = \mathbf{I}$, where $\mathbf{I}$ is an $n \times n$ identify matrix, in which case all model parameters are perfectly resolved. $\mathbf{R}$ may also be poor if $\mathbf{G}$ is non-invertable, which can occur if there is a high degree of linear dependency between rows or columns.

The condition of a inverse problem describes how sensitive the model is to errors in the data. A problem that is ill-conditioned with be highly sensitive to any changes in the data, so that random noise may lead to significantly different results. This is expressed through high model uncertainties. As with ill-posed inverse problems, regularisation may be used to improve the condition of the problem.

In most cases, we can expect our inverse problem to be well-posed, as both the number of independent observations and the number of model parameters can be controlled. Therefore, the condition that $m > n$ can always be enforced, by either increasing the number of GNSS stations in the network, or by decreasing the number of slip patches in the discretisation.

It may be the case that our inverse problem is ill-conditioned with high model uncertainties. As previously stated, regularisation may then be used to improve the condition of the inverse problem and the model uncertainties, at the cost of degrading the model resolution and imposing an arbitary bias on the results. Instead of applying regularisation, we improve the condition of the inverse by adjusting the number and dimensions of the slip patches in the discretisation. This is done through optimisation, to minimise the model uncertainties by fitting the slip patches to the variable resolving power of the data.

2.2 Network Design

In this Section, a summary of network design theory is provided. Three potential network objectives are described; reliability, economy, and precision, alongside a range of potential optimisiation techniques.

2.2.1 Background

Geodetic network design originates from the work of Baarda (1973) and Grafarend and Schaffrin (1974), with a summary of the early development of the field provided by Schmitt (1982) and Grafarend (1985). The primary types of network design

Order	Fixed Parameters	Free Parameters
ZOD	G, Q_d	m, Q_m
FOD	Q_d, Q_m	G
SOD	G, Q_m	Q_d
THOD	Q_m	$G, Q_d(partlyfree)$

Table 2.1: Free and fixed parameters for each order of network design problem.

problems, along with three commonly employed optimality criteria, are summarised below.

Geodetic network design problems may be divided into four *Orders of Design* (Grafarend and Schaffrin 1974; Schmitt 1982)

1. **Zero-order design (ZOD):** selection of a datum for a free network of relative measurements (Teunissen 1985).
2. **First-order design (FOD):** finding of optimal measurement locations (Mahapatra et al. 2015).
3. **Second-order design (SOD):** selection of observations and their weights in an established network (Xu et al. 1995; Amiri-Simkooei 2004; Doma 2013).
4. **Third-order design (THOD):** improvement of an existing network through design of an optimal densification network (Chen, Rizos et al. 2001; Klein et al. 2017; Sathiakumar et al. 2017).

The free and fixed parameters for each order of design are summarised in Table 2.1.

In the following methodology, it is assumed that no existing GNSS stations are included as part of the network, making the network design problem one of THOD.

2.2.2 Reliability

Reliability is a measure of a network's ability to detect and resist model imperfections, such as outliers, and can be catagorised into *internal reliability* and *external* *reliability* (Amiri-Simkooei 2004; Amiri-Simkooei et al. 2012; Grafarend and Sansò 2012; Mahapatra et al. 2015). Internal reliability refers to the network's ability to detect gross errors in the observations, and can be measured by the Minimum Detectable Bias (MDB), which describes the size of the smallest model error that can be detected through appropriate statistical testing (Teunissen 1998)

$$|\nabla d_i| = \sqrt{\frac{\lambda_0}{(\mathbf{W_e Z})_{ii}}} \tag{2.2.1}$$

where λ_0 is the lower bound of the noncentrality parameter, and $\mathbf{Z}$ is the reliability matrix. $\mathbf{Z}$ is defined as

$$\mathbf{Z} = \mathbf{I} - \mathbf{G}(\mathbf{G^T W_e G})^{-1}\mathbf{G^T W_e} \tag{2.2.2}$$

where $\mathbf{I}$ is an identity matrix of equal dimensions to $\mathbf{W_e}$. External reliability refers to the maximum effect of undetected gross errors on the model parameters $\mathbf{m}$, and is expressed

$$\nabla m = \left(\mathbf{G^T W_e G}\right)^{-1} \mathbf{G^T W_e} |\nabla d_i| \tag{2.2.3}$$

It can be seen that both types of reliability are primarily based upon the model geometry (G) and the quality of the observations (W_e).

Both types of reliability may be expressed through $\mathbf{Z}$, the main diagonal of which contains the *redundancy numbers,* z_i. z_i expresses the level of detectable gross error in the ith observation, with $z_i = 1$ meaning that any gross error in the observation can be detected, and vice versa for $z_i = 0$. Redundancy values in the range $0.3 \leq z_i \leq 0.7$ are desired, representing reasonable controllability of the network observations (Mahapatra et al. 2015). Observations where $z_i \geq 0.7$ are over-controlled, and could be omitted from the network in the context of network reliability (Staudinger 2000). In more practical terms, an observation will have a high z value if multiple other observation measure the same model parameters, meaning any error can be detected through comparison with the other observations. Similarly, an observation which is alone in measuring a given parameter will posses a low z value as errors cannot be detected by comparison. Combined, these values express the sensitivity of a network to each individual observation.

Geodetic networks may be designed to maximise network reliability (e.g. Seemkooei (2001)). Reliability is not used as an objective in the following methodology, however, it is revisited in the discussion chapter to provide additional analysis of example networks.

2.2.3 Economy

Economy refers simply to the cost of operating a geodetic network. This includes purchasing of devices, deployment, maintenance, and potentially manual data collection. Generally, a network of fewer stations will have a lower cost, although there may be some minimum required expenditures such as traveling to the work area. Economy may be included simply as a maximum number of stations, or as a more complex cost function (Staudinger 2000; Amiri-Simkooei et al. 2012).

Economy will be a deciding factor when deploying GNSS networks in real life, however, it is not directly included in the following methodology. Instead, the performance returns from increasing the number of stations in a network are explored in Chapter 4.

2.2.4 Precision

Precision reflects the performance of a network in estimating the desired model parameters. Two different precision objectives are explored: model resolution, and model uncertainty.

As explain in Section 2.1.3, model resolution describes how well the observations resolve the model parameters. Poor model resolutions results in model slip values that are “smeared” across multiple slip patches, and so are poorly constrained spatially. GNSS networks may be designed for high model resolution by maximising the main diagonal of $\mathbf{R}$, or the number of patches above a given threshold (e.g. Sathiakumar et al. (2017)). Model uncertainty must then be controlled by other means, such as regularisation.

To use model uncertainty as a measure of network precision, the model VCM defined in Section 2.1.2 is returned to. An optimal GNSS network may be found by minimising the model variance and covariance values, found in $\mathbf{Q_m}$, to improve the precision of the estimated fault values.

Mahapatra et al. (2015) employs a criterion matrix method to minimise the model uncertainty in a network of coherent targets of InSAR studies. A criterion matrix is a model VCM that posses an idealised or desired structure ($\mathbf{Q}_{m,ref}$). This is created either manually, or through the use of a network of idealised “best possible” observations. In the latter case, $\mathbf{Q}_{m,ref}$ will then contain the best possibe model variance and covariance values than can be achieved.

To summarise the difference between the model VCM from the design network ($\mathbf{Q}_{m,des}$) and the criterion matrix ($\mathbf{Q}_{m,ref}$) as a scalar value, Mahapatra et al. (2015) employs a distance metric. This may then be minimised through optimisation to reduce the difference between $\mathbf{Q}_{m,des}$ and $\mathbf{Q}_{m,ref}$, which in turn maximises the precision of the estimated slip values on the model fault. The advantage of this method when designing GNSS networks to recover aseismic slip on a fault is that the variable resolving power of the observations, and therefore the minimum possible model uncertainty that may be achieved, is accounted for in the distance metric.

Network precision, expressed in terms of model uncertainty, is used in the network design described in this thesis. Further detail is provided in Section 3.2.

2.2.5 Optimisation Techniques

The optimisation of geodetic networks can be achieved through a range of different techniques and algorithms. I begin this section by providing a brief summary of the background theory, and then describing several previously used optimisation methods.

A simple optimisation problem consisting of a set of variables, $\mathbf{x}$, and an objective function f, can be defined as

$$\min f(\mathbf{x}) \qquad \text{subject to } \mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub} \text{ and } \mathbf{x} \in \mathbb{R}^n \tag{2.2.4}$$

where $\mathbf{lb}$ and $\mathbf{ub}$ are the upper and lower bounds on the variables $\mathbf{x}$. For a given solution $\mathbf{x}^\star$ that has an objective function value $f(\mathbf{x}^\star)$, it may be defined as a local solution if

$$f(\mathbf{x}^\star) \leq f(\mathbf{x}) \qquad \forall\, \mathbf{x} \in \mathbf{N} \tag{2.2.5}$$

where $\mathbf{N}$ is a neighbourhood of solutions around $f(\mathbf{x}^\star)$. In addition, the solution is the global optimum if

$$f(\mathbf{x}^\star) \leq f(\mathbf{x}) \qquad \forall\, \mathbf{x} \tag{2.2.6}$$

The problem may be defined as nonlinear if there is not a linear relationship between $\mathbf{x}$ and f, or if there exists some equality or inequality constraints on $f(\mathbf{x})$ that are nonlinear.

In the case where the optimisation is complex, possessing many local minima, nonlinearities, and an inability to calculate derivatives (non-smooth), *heuristic* algorithms may be employed to find solutions. These are a class of algorithms designed to solve problems where classical methods are either too slow, or fail to find exact solutions, often termed *hard* optimisation problems. The trade-off is that heuristic algorithms are not guaranteed to find the global optimum, meaning the solution is often an approximation, and establishing the true nature of the solution can be difficult. However, for many real-world applications a local minimum close to the global optimum is sufficient to solve the problem, and considered an acceptable compromise for the ability to efficiently solve complex problem where heuristcs may be the only viable option.

Metaheuristics are higher-level processes that may be used to guide or generate heuristic algorithms for solving a range of hard optimisation problems with minimum changes (Boussa´id et al. 2013). They are often inspired by some aspect of the natural world, such as physics or biology, and like heuristic algorithms do not use gradients of the objective function. The field saw rapid development throughout the 1980's and 90's, with the development of Simulated Annealing (Kirkpatrick et al. 1983; Berné et al. 2004), Ant Colony optimisation (Dorigo and Gambardella 1997; Dorigo and Birattari 2011), Genetic Algorithms (Goldberg et al. 1988; Holland 1992; Koza 1994; Mitchell 1998), and Particle Swarm optimisation (Eberhart and Kennedy 1995; Eberhart and Shi 1998; Bonyadi et al. 2017), the latter of which is used to perform the network optimisation in this thesis.

Grid search

Grid search optimisation employs a discrete, finite set of instrument locations which are then explored sequentially using a given objective function (Mahapatra et al. 2015; Klein et al. 2017). Uniformly-spaced grids of locations are common as they are simple to generate. Instruments may be placed one at a time, with previously add instruments retested to identify if they have been made redundant by later additions (Sathiakumar et al. 2017).

Grid search methods are sub-optimal for designing GNSS networks. The discrete set of station locations means that they are not able to test all possible network configurations and will likely miss the true optimal station locations. Placing stations individually, even with the ability to remove previously added stations when they become redundant, is also sub-optimal as it will not capture the full interplay between the networked stations. For these reasons, a grid search method is not used in this thesis.

Simulated annealing

Simulated annealing is an iterative metaheuristic optimisation algorithm that operates as an analogy of cooling metal, where freely moving particles will settle into the lowest energy configuration available (global minimum) if the cooling period is sufficiently long (Metropolis et al. 1953; Kirkpatrick et al. 1983). The algorithm will converge to the global optimum, if the optimisation is performed for a sufficient period of time (Geman et al. 1984). However, the required duration may be beyond practical limits, especially in the case of complex design problems with large numbers of local minima. Berné et al. (2004) demonstrates how simulated annealing can be used for simple geodetic network design.

Simulated annealing was tested as a potential optimisation algorithm for the network design problem described in this thesis. However, it proved less efficient than PSO, described below. Simulated annealing optimises a single solution at once, and so is less effective at exploring the solution space than PSO, which processes a population of solutions simultaneously.

Particle swarm

Particle Swarm Optimisation is an stochastic population-based optimisation algorithm designed to solve continuous nonlinear problems. The technique was developed by Eberhart and Kennedy (1995) as an analogy of swarms in nature, such as shoals of fish and flocks of birds, with foundations in swarm theory. PSO is classified as a evolutionary algorithm, much like MOGAs, due to its use of a population of solutions, although it does not employ analogies of natural selection. Bonyadi et al. 2017 provides an up-to-date summary of the algorithm and its uses. The Matlab Optimisation Toolbox contains the PSO function ‘particleswarm’, based upon the original algorithm by Eberhart and Kennedy (1995) with modifications from Pedersen 2010 and Mezura-Montes et al. 2011.

PSO is chosen to solve to network optimisation problem described in this thesis, with further details as to its use provided in Section 3.2.

Chapter 3

Methodology

To begin designing optimal low-cost GNSS networks to recover distributed aseismic slip on continental fault zones, it is first necessary to discretise the fault plane and estimate the "best possible" model uncertainties for each slip patch. The discretisation is performed using a multi-objective genetic algorithm, which is then be used to generate the design matrix $\mathbf{G}$ by treating each slip patch as a finite rectangular half-space (Okada 1985). An idealised network of dual-frequency GNSS stations, where each observation is treated as an independent absolute measurement of displacement, provides $\mathbf{Q}_{d,ref}$, from which the criterion matrix $\mathbf{Q}_{m,ref}$ is derived.

Once the discretisation and $\mathbf{Q}_{m,ref}$ have been fixed, the question of how to design optimal low-cost GNSS networks to recover slip on a modeled fault may be addressed. In the case of dual-frequency stations, the data VCM for the designed networks $\mathbf{Q}_{d,des}$ is generated based on independent absolute measurements of displacement, as with $\mathbf{Q}_{d,ref}$. For single-frequency GNSS stations, the effects of ionospheric delay must be included. $\mathbf{Q}_{m,des}$ may then be created using $\mathbf{Q}_{d,des}$, and GNSS networks designed so as to minimise the difference between $\mathbf{Q}_{m,des}$ and $\mathbf{Q}_{m,ref}$.

3.1 Fault Discretisation

Discretisation describes the splitting of the fault plane into slip patches, each of which is then modeled as a finite slipping fault plane. Slip patches form the model parameters of the discrete linear inverse problem described by Equation (2.1.3). Discretisation enables slip to be constrained to smaller spatial scales than the fault plane, and the varying slip uncertainty across the fault to be estimated. Rectangular patches are commonly adopted (e.g. Page et al. (2009), Atzori et al. (2011), Metzger et al. (2017) and Sathiakumar et al. (2017)) as they are simple to generate. An alternative choice is that of a triangular mesh, as described by Barnhart et al. (2010) and used in Hayes et al. (2014).

The fault plane is discretised into rectangular patches based upon horizontal boundaries, which define *rows* of patches, with an integer number of patches per row. The dimensions of each patch are uniform within each row, but differ between rows. Patch dimensions are allowed to vary with depth so that they better fit the resolving power of the observations, which generally decreases with increasing depth. Maintaining uniform patch size along each row means that slip sensitivity is consistent along strike. It is assumed that the fault geometry is perfectly constrained, but that the location of the slip on the fault plane is not known. Figure 3.1 shows a schematic diagram of a non-uniformly discretised fault plane. The fault itself is a fracture surface within some medium, with slip values that are relative displacements across the fault plane.

The quality of a discretisation can be assessed using two characteristics: the number of slip patches, and the uncertainties on the model slip associated with each patch. A high number of patches allows the spatial distribution of slip to be represented more smoothly, while lower model uncertainty increases the precision of modelled slip values for the same fixed data. A trade-off exists between the two characteristics, with greater numbers of patches (and thus smaller patch sizes) correlating with higher model uncertainty. These characteristics are used as fitness

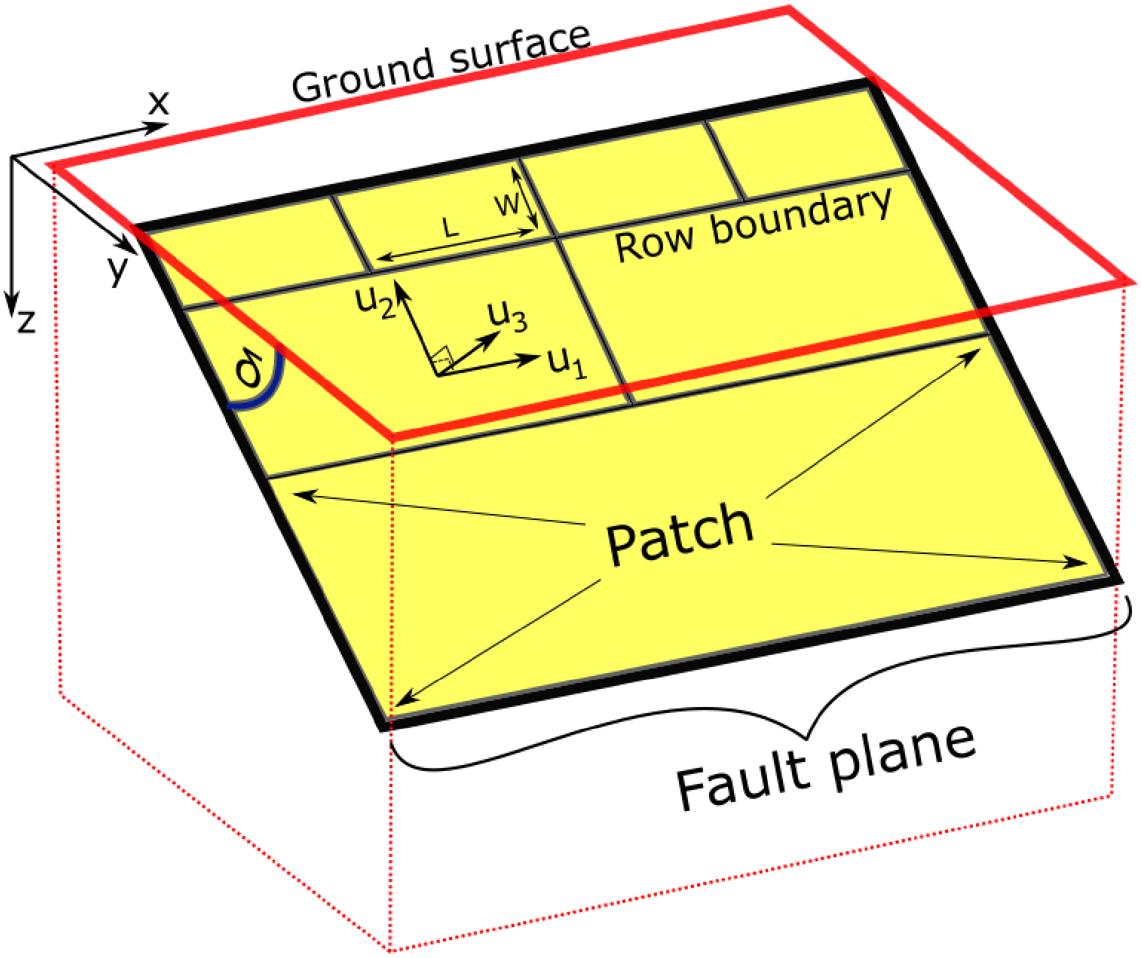

Figure 3.1: Schematic diagram of a discretised fault plane. L and W are the length and width of the slip patch, σ is the dip angle, and u_1,u_2, and u_3 are the components of slip on each patch.

functions in the optimisation.

Consider two extreme cases:

- For a discretisation with only a single patch, the associated variance value will be low. However, slip may only be constrained to the scale of the entire fault plane, and so no distinction can be made between slip events in different locations.

- For a discretisation where the number of patches tends to infinity, slip events may be discerned from each other and constrained to very small scales. However, in practice variance on the model parameters will also tend to infinity and effectively invalidate any modelled slip values.

Optimal fault discretisations will exist as a trade-off between these two characteristics, where neither characteristic can be improved without degrading the other. Generating these discretisations requires an optimisation algorithm capable of handling two objectives simultaneously, and efficiently exploring the possible solution space.

3.1.1 Idealised Network

Model uncertainty is used as a fitness function in the fault discretisation, and so it is necessary to define $\mathbf{Q}_{d,ref}$ before beginning the optimisation so that $\mathbf{Q}_{m,ref}$ may be estimated for each discretisation.

An idealised network of dual-frequency GNSS stations, which recovers the maximum practical amount of slip information about the fault, is used to generate $\mathbf{Q}_{d,ref}$. Two zones are defined within the idealised network - a high density area (small station separation) and a low density area (high station separation), with stations uniformly spaced within these zones (Fig. 3.2). This reduces processing requirements while ensuring that the short-wavelength deformation signal close to the fault is well sampled. The high density area may be rotated to remain parallel to the projected fault line, so that the observations are uniform along strike.

All observations within the network are assumed to be independent absolute measurements of surface displacement. Variance values equivalent to a 1σ positional error of 2 cm in the North and East components, and 4 cm in the Up component, are used to produce $\mathbf{Q}_{d,ref}$. These are reasonable average positional accuracies for dual-frequency GNSS stations (Groves 2013).

An example idealised network is shown in Figure 3.2.

3.1.2 Algorithm

We chose a multi-objective genetic algorithm to perform the discretisation, specifically the ‘gamultiobj’ function from the Matlab Optimisation Toolbox. MOGAs are a

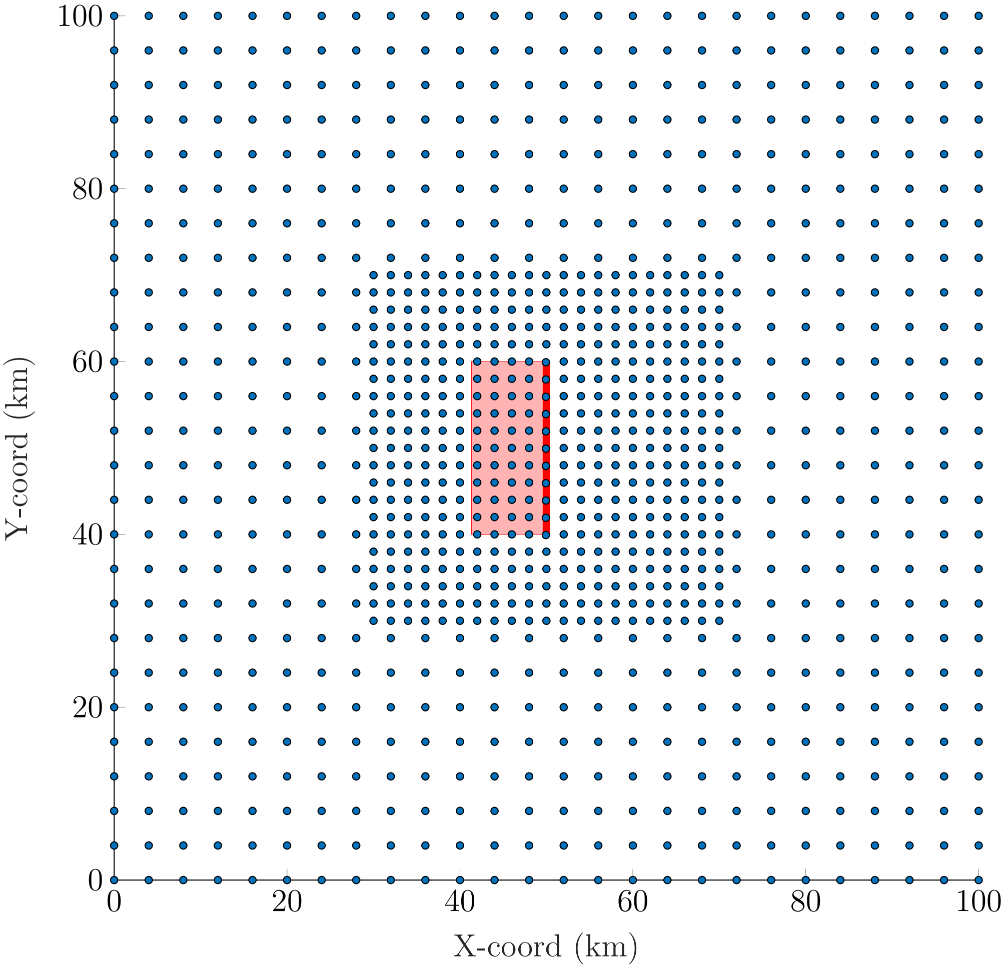

Figure 3.2: An example idealised network centred around a westward-dipping normal fault. Station spacings are 2 km for the high density area, and 4 km for the low density area. The bold red line indicates the fault trace, while the red area shows the vertically projected fault plane

metaheuristic optimisation technique, as described in Section 2.2.5, that are capable of solving complex optimisation problems with multiple objectives. The algorithm is provided model parameters and fitness functions that describe the structure and quality of each discretisation. Multiple discretisations, which define a population of solutions, are iteratively varied until a range of optimal solutions are found.

Model parameters are provided to the algorithm as a vector

$$\mathbf{m}_{GA} = \begin{bmatrix} z_1 & z_2 & \ldots z_{r-1} & c_1 & c_2 & \ldots & c_r \end{bmatrix} \tag{3.1.1}$$

where $\mathbf{m}_{GA}$ is the vector of parameters, z is vertical row boundary depths, c is the number of patches in each row, and r is the number of rows. As both the top and bottom depths of the fault plane are fixed, there is one less row boundary than the number of rows. The patch boundary depths are projected into dip-parallel depths using the dip angle of the fault plane. The algorithm is provided starting values for $\mathbf{m}_{GA}$ consisting of uniformly spaced row boundaries with a single patch on each row. This is done to ensure that the starting model does not contain an unrealistic number of slip patches, and to encourage the algorithm to explore discretisations with low model uncertainties.

The previously described measures of discretisation quality, those being model uncertainty and the number of slip patches, must be expressed as scalar values so that they may be used as fitness functions by the algorithm. For the model uncertainties, the algorithm is provided the maximum value of any model slip value in the discretisation. Using the maximum value observed encourages the generation of discretisations with uniform model uncertainties across all slip patches, an advantage when the location of slip on the fault plane is not know. The number of slip patches in maximised by minimising the inverse of the number of patches in the discretisation.

Discretisations are generated as a population of solutions. $\mathbf{G}$ is generated for each discretisation assuming each slip patch is a rectangular elastic half-space, as described in Section 2.1.1. $\mathbf{Q}_m$ is then estimated using $\mathbf{G}$ and $\mathbf{Q}_{d,ref}$. Each discretisation is tested for dominance and sorted. A solution (m_1) is said to dominate another solution (m_2) if:

- m_1 is at least equal in all objective functions compared to m_2.
- m_1 is better in at least one objective compared to m_2

The highest ranked discretisations are selected, randomly modified (*mutation*), and mixed together (*crossover*) to produce a new population that retains positive characteristics from the parent discretisations. The algorithm attempts to minimise both fitness functions, with lower values indicating a more optimal solution. Dominance between solutions is evaluated so as to define the estimated Pareto front, an example of which is provided in Figure 3.3.

The number of model parameters that the algorithm can handle is fixed during the optimisation, and so discretisations with different numbers of rows cannot be generated simultaneously .To overcome this limitation, the optimisation is performed multiple times, each time with a different number of rows in the discretisation. The solutions from each optimisation are than re-tested for dominance to define a new estimated Pareto front. An example of this combined Pareto front is shown in Figure 3.4.

It can be seen that for different ratios of the two fitness functions, different numbers of row boundaries dominate the Pareto front. A higher number of row boundaries leads to higher model uncertainty as patches are forced to smaller widths, and vice versa. This also indicates that patch aspect ratios are kept relatively consistent throughout the range of solutions.

A discretisation is chosen from the combined Pareto front based upon the magnitude of the target aseismic slip. The range of discretisations produced by the optimisation means that a wide range of slip magnitudes can be studied, as long as surface displacement above the minimum magnitude detectable of GNSS stations is produced. This selected discretisation is fixed, and the criterion matrix $\mathbf{Q}_{m,ref}$ is estimated using $\mathbf{Q}_{d,ref}$ and $\mathbf{G}$.

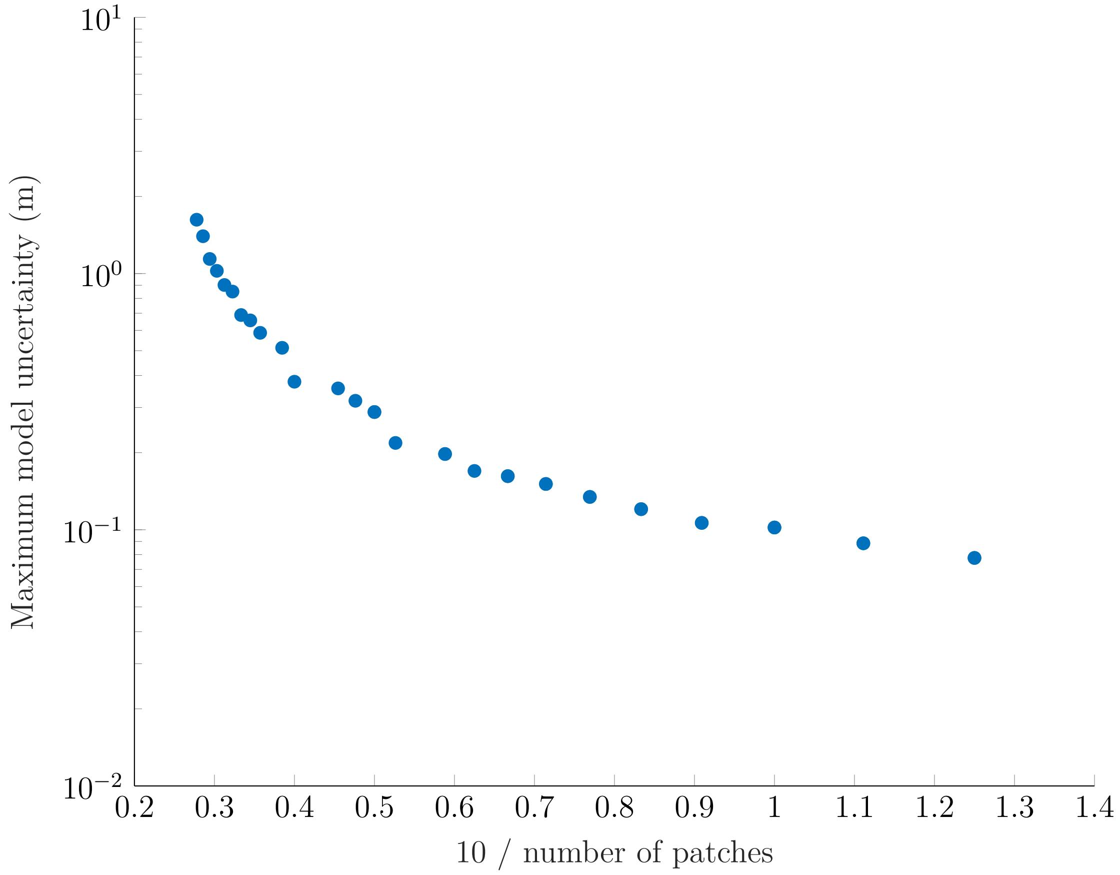

Figure 3.3: An example Pareto front for discretisations with four rows, where each point is a solution. A log axis is used for the model error, as this increases exponentially with decreasing patch size, for which the number of patches is a proxy. The front itself is not uniform, as this is only an estimate of the true Pareto front and is limited by the performance of the algorithm.

3.2 Network Design

Once a discretisation is chosen, and $\mathbf{Q}_{m,ref}$ has been estimated, it is possible to begin generating optimal GNSS networks. The network design is performed using the 'particleswarm' function from the Matlab Optimisation Toolbox, a PSO algorithm that is similar to the MOGA used in the fault discretisation.

The algorithm defines a population of particles, each of which is a GNSS network defined by a set of x and y coordinates that describe the location of each station. Starting values are provided randomly, within predefined bounds that define a rectangular region around the fault. Station locations are able to vary continuously in

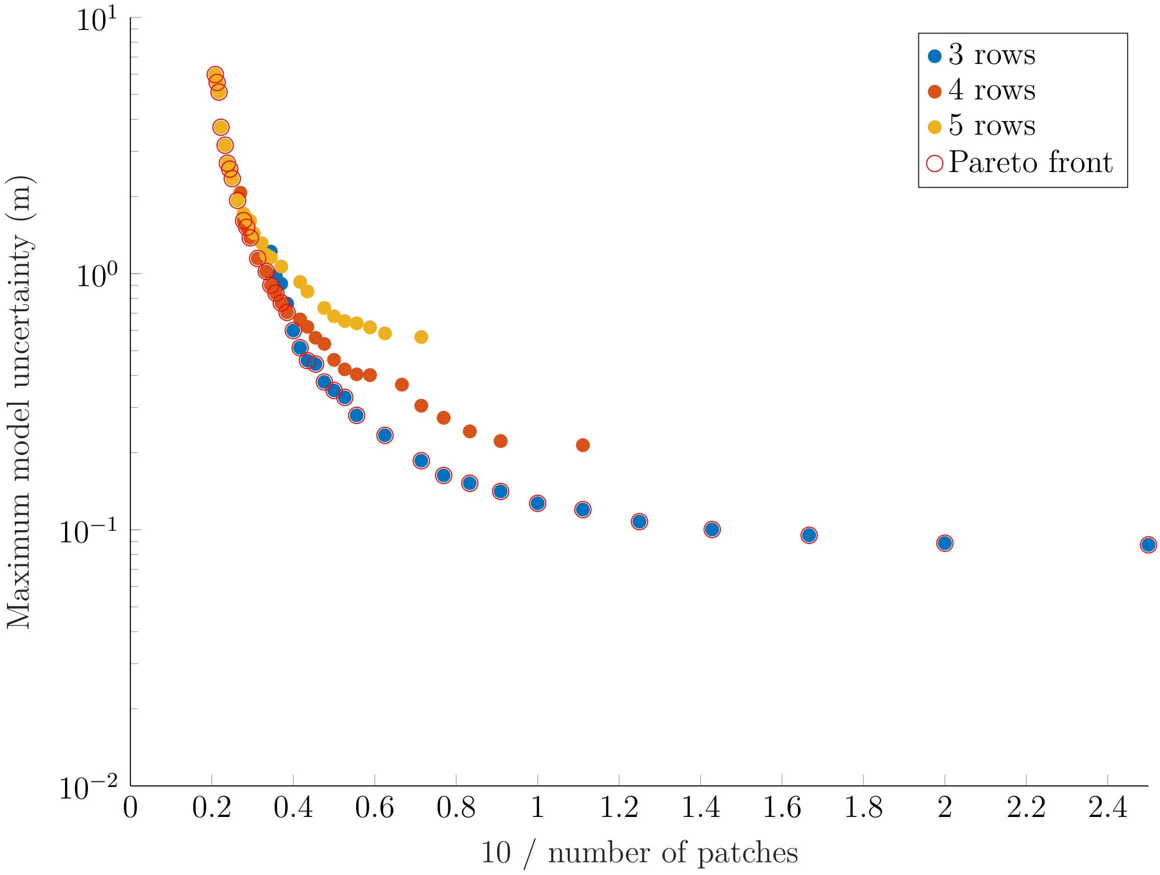

Figure 3.4: Combined Pareto front produced for a synthetic normal fault. Each circle is a discretisation, with the colour indicating the number of rows in the discretisation, with the red circles indicating the combined Pareto front.

this area, as opposed to at discrete intervals as is common in grid search methods.

A distance metric (e.g. Förstner et al. (2003)) is employed as a fitness function to express the quality of each network as a scalar value:

$$d\left(\mathbf{Q}_{m,des},\mathbf{Q}_{m,ref}\right)=\sqrt{\sum_{i=1}^{n} ln^2\lambda_i\left(\mathbf{Q}_{m,des},\mathbf{Q}_{m,ref}\right)} \tag{3.2.1}$$

where λ are the eigenvalues of $\mathbf{Q}_{m,des}$ and $\mathbf{Q}_{m,ref}$. The distance metric expresses in the difference in all variance and covariance values between $\mathbf{Q}_{m,des}$ and $\mathbf{Q}_{m,ref}$. A lower distance metric value expresses that the designed network is approaching the "best possible" model uncertainties. If both VCMs are identical, then $d\left(\mathbf{Q}_{m,des},\mathbf{Q}_{m,ref}\right)=0$.

Estimating $\mathbf{Q}_{m,des}$ requires $\mathbf{Q}_{d,des}$ and $\mathbf{G}$. $\mathbf{G}$ is again generated as described in Section 2.1.1. It is necessary to recalculate **G** for every network layout as it is dependent upon the observation locations. For dual-frequency GNSS networks, $\mathbf{Q}_{d,des}$ is generated in the same manner as $\mathbf{Q}_{d,ref}$. For single-frequency GNSS stations, the effects of ionospheric delay must be included. This is explored further in the next section.

The assumption that covariance values between dual-frequency GNSS stations are negligible can cause the optimisation algorithm to “stack” stations. This describes the behaviour where multiple GNSS stations are placed with little distance between adjacent stations (10’s of m) to minimise model uncertainty on a specific patch, with alternative locations proving less optimal . At very short separation distances, significant covariance would be present between adjacent stations due to localised error sources. To account for this, a bias is added to the $\mathbf{Q}_{d,des}$ for stations that are placed within 200 m of another station. This negatively impacts stations located within close proximity, forcing a different location to become optimal once the first has been filled. This is only performed for the dual-frequency GNSS stations, as the inclusion of ionospheric delay in the single-frequency stations also prevents stations from “stacking”.

Each particle is defined by four properties:

- Position - the position of the particle at a given iteration (current model parameter values).
- Velocity - the direction and length of movement of the particle at a given iteration (change in model parameter values).
- Personal best - the position of the best solution (minimum distance metric value) found so far by the particle.
- Global best - the position of the best solution observed by any particle in the swarm.

The movement of each particle, which controls how station locations change each iteration, is controlled by these four properties and an additional random component. Particles will on average move towards the global best while exploring the solution space. The global best at the end of the optimisation is considered the optimal solution. The optimisation finishes when either a maximum number of iterations is reached, or the change in the distance metric value of the global best solution over a given number of iterations decreases below a threshold.

3.3 Single Frequency GNSS

$\mathbf{Q}_{d,des}$ is required to estimate $\mathbf{Q}_{m,des}$, which itself describes the ability of a network to recover distributed aseismic slip on a fault. Accurate estimation of the variance and covariance values within $\mathbf{Q}_{d,des}$ is therefore critical to the network design. For dual-frequency GNSS stations, estimating $\mathbf{Q}_{d,des}$ is relatively simple because each observation can be assumed to be an independent absolute measurement of surface displacement. For single-frequency GNSS stations this is not the case because of ionospheric delay. Ionospheric delay is a major source of positional uncertainty for single-frequency GNSS stations. For dual-frequency stations, it is possible to mitigate this uncertainty by differencing estimates of the magnitude of the ionospheric delay from both frequencies. This is not possible for single-frequency, and so alternative methods must be used.

As the ionospheric delay has high spatial correlation over hundreds of kilometres, a simple way to mitigate this source of uncertainty is to consider relative differences in displacement between pairs of stations, rather than absolute measurements. The data uncertainty is then dependent upon the variation in ionospheric delay between pairs of stations, rather than the absolute magnitude.

If single-frequency GNSS stations are treated independently, then the effect of the ionospheric delay will be large. As a simple example, consider a vector of four observations

$$\mathbf{d} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 \end{bmatrix} \tag{3.3.1}$$

each of which is an absolute measurement of surface displacement in a given component (e.g. North, East, or Up), taken from four independent single-frequency GNSS stations. A VCM may be created in the same form as Equation (2.1.2):

$$\mathbf{Q}_d = \begin{array}{cc} \begin{array}{cccc} d_1 & d_2 & d_3 & d_4 \end{array} & \\ \begin{pmatrix} Var(d_1) & Cov(d_1, d_2) & Cov(d_1, d_3) & Cov(d_1 d_, 4) \\ Cov(d_2, d_1) & Var(d_2) & Cov(d_2, d_3) & Cov(d_2, d_4) \\ Covd_(3, d_1) & Cov(d_3, d_2) & Var(d_3) & Cov(d_3, d_4) \\ Cov(d_4, d_1) & Cov(d_4, d_2) & Cov(d_4, d_3) & Var(d_4) \end{pmatrix} & \begin{array}{c} d_1 \\ d_2 \\ d_3 \\ d_4 \end{array} \end{array} \tag{3.3.2}$$

If all other error sources are assumed to be negligible, then

$$Var(d) = \sigma^2_{d,max} \tag{3.3.3}$$

where $\sigma^2_{x,max}$ is the maximum variance of the ionospheric delay. This may be several squared metres in magnitude (Huang et al. 2017) and will propagate through to $\mathbf{Q}_{m,des}$, producing large model uncertainties and covariances.

In the case of relative measurements of displacement between pairs of stations, the variance may be expressed:

$$Var(d_1, d_2) = Var(d_1)_{max} + Var(d_2)_{max} - 2\,Cov(d_1, d_2) \tag{3.3.4}$$

Assuming that $Var(d_1)_{max} = Var(d_2)_{max}$, Equation (3.3) can be written as

$$Var(d_1, d_2) = \sigma^2_{d_1,max} + \sigma^2_{d_2,max} - 2\rho_{d_1,d_2}\sigma_{d_1,max}\sigma_{d_2,max} \tag{3.3.5}$$

where ρ is the correlation between any two stations. The left two terms represent the maximum variance of the ionospheric delay, with the right term acting to reduce this based upon the separation distance between the two stations. Two cases are considered, where l is the baseline distance between two stations, and x and y are the observations:

- if $l \rightarrow inf$, then $\rho = 0$ (observations are uncorrelated), and $Var(x, y) = 2Var_{max}$, which will dominate the signal.

- if $l = 0$, then $\rho = 1$ (observations are perfectly correlated), $Var(x, y) = 0$, and the observation will be unaffected by ionospheric delay.

In practice, separation distances between these extremes are desired. Shorter baselines result in lower data uncertainty, which is then propagated through to $\mathbf{Q}_{m,des}$, at the cost of higher covariance and poorer network coverage.

The covariance between two relative measurements (two pairs of single-frequency stations), can be expressed as

$$\begin{aligned} Cov(d_1, d_2 \,\&\, d_3, d_4) &= Cov(d_1, d_3) + Cov(d_2, d_4) - Cov(d_2, d_3) - Cov(d_1, d_4) \\ &= \rho_{d_1,d_3}\sigma_{d_1}\sigma_{d_3} + \rho_{d_2,d_4}\sigma_{d_2}\sigma_{d_4} - \rho_{d_2,d_3}\sigma_{d_2}\sigma_{d_3} - \rho_{d_1,d_4}\sigma_{d_1}\sigma_{d_4} \end{aligned} \tag{3.3.6}$$

As another example, two cases are considered for baseline distance in a setup of four stations as shown in Figure 3.5. For two pairs of stations arranged to form a rectangle, as the separation between the two pairs increases (L_1), the covariance decreases, and vice versa. The rate at which this occurs depends upon the pair separation (L_0) and the spatial autocorrelation function of the ionosphere, which describes the correlation between any two measurements of the ionospheric delay for a given separation distance. If the two pairs are separated to infinity, then the two relative measurements are independent. As the distance between the two pairs goes to zero, then the two relative measurements become perfectly correlated, and the resulting covariance value is the variance of either pair, as described by equation 3.3.

Using Equations (3.3) and (3.3.6), $\mathbf{Q}_{d,des}$ may be created for the relative measurements between the four observations in $\mathbf{d}$, assuming d_1 is paired with d_2, and d_3

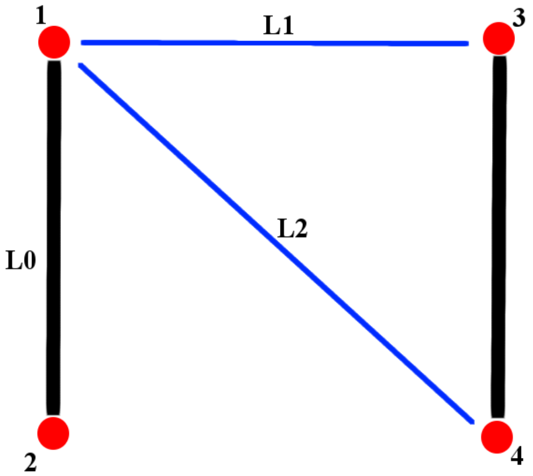

Figure 3.5: Diagram of four single-frequency GNSS stations (red dots) arranged in a rectangle and paired (black lines). Blue lines indicate separation distances.

with d_4 (Fig. 3.5)

$$\mathbf{Q}_{\mathbf{d}} = \begin{array}{cc} \begin{array}{cc} \qquad\qquad d_1, d_2 \qquad\qquad & \qquad\qquad d_3, d_4 \qquad\qquad \end{array} & \\ \left(\begin{array}{cc} Var(d_1) + Var(d_2) - 2Cov(d_1, d_2) & \begin{array}{c} Cov(d_1, d_3) + Cov(d_2, d_4) \\ -Cov(d_2, d_3) - Cov(d_1, d_4) \end{array} \\ & \\ \begin{array}{c} Cov(d_3, d_1) + Cov(d_4, d_2) \\ -Cov(d_3, d_2) - Cov(d_4, d_1) \end{array} & Var(d_3) + Var(d_4) - 2Cov(d_3, d_4) \end{array} \right) & \begin{array}{c} d_1, d_2 \\ \\ \\ \\ d_3, d_4 \end{array} \end{array} \tag{3.3.7}$$

This can be expanded for any network size, with each row and column corresponding to a different pair of observations from single-frequency stations.

A inherent problem with relative measurements is that columns of the $\mathbf{Q}_{d,des}$ which share a common station will be linearly dependnent. This results in a VCM of rank $n-1$, where n is the number of single-frequency stations, which causes the inversion to become ill-conditioned unless it is stabalised using regularisation. To avoid the use of regularisation, the rows and columns associated with certain pairs of stations are removed from the VCM until it once again full rank.

To select which pairs of stations to keep, and which to remove, a minimum spanning tree algorithm is employed (Moret et al. 1991; Neumann et al. 2007).

Station locations are converted to vertices on a graph, with weighted edges connecting each vertices to every other vertices. A minimum spanning tree is a subset of all possible edges that connects all vertices together without cycles (i.e. loops of stations). Multiple trees will exist, and so the tree with the lowest total weight is defined as the optimum. Edges are weighted using the separation distance between paired stations, which is roughly proportional to the data uncertainty resulting from the ionospheric delay. The optimal tree will therefore produce the $\mathbf{Q}_{d,des}$ with the lowest data uncertainties for any given station layout. Examples of minimum spanning trees are presented in Chapter 4.

3.3.1 Characterising the Spatial Structure of the Ionosphere

The difference in the ionospheric delay between paired single-frequency stations may be estimated from the average spatial structure of the ionosphere. It is assume that the ionospheric delay is isotropic and stationary for the duration of the observations, and so the difference in ionospheric delay between any two single-frequency GNSS stations is dependent only upon the separation distance between the two stations, and not the locations of the stations themselves.

The spatial structure of the ionospheric delay is best expressed as a variogram, which describes how separation distance affects the measurements of a spatial random field made in two locations. In terms of the ionosphere, the variogram provides the likely difference in value of two measurements of the ionospheric delay made by two GNSS stations. Measurements of the ionospheric delay from two stations at a large separation distance will differ more than those from two stations at a small separation distance. To avoid confusion between the terms *variogram* and *semivariogram*, which are sometimes used interchangeably, the work of Bachmaier et al. (2011) is followed, in which the function itself if referred to as the variogram, and the values taken from it as *gammavariance*, or just *gamma* for short. The term *semivariogram* is not used.

Huang et al. (2017) provides variograms for the ionospheric delay measured by the Crustal Movement Observation Network of China (CMONOC) (Fig. 3.6). The GNSS stations within this network were used to sample the ionosphere every two hours, with experimental variograms produced by differencing measurements across a range of separation distances . The shape of each variogram is defined by a Gaussian function.

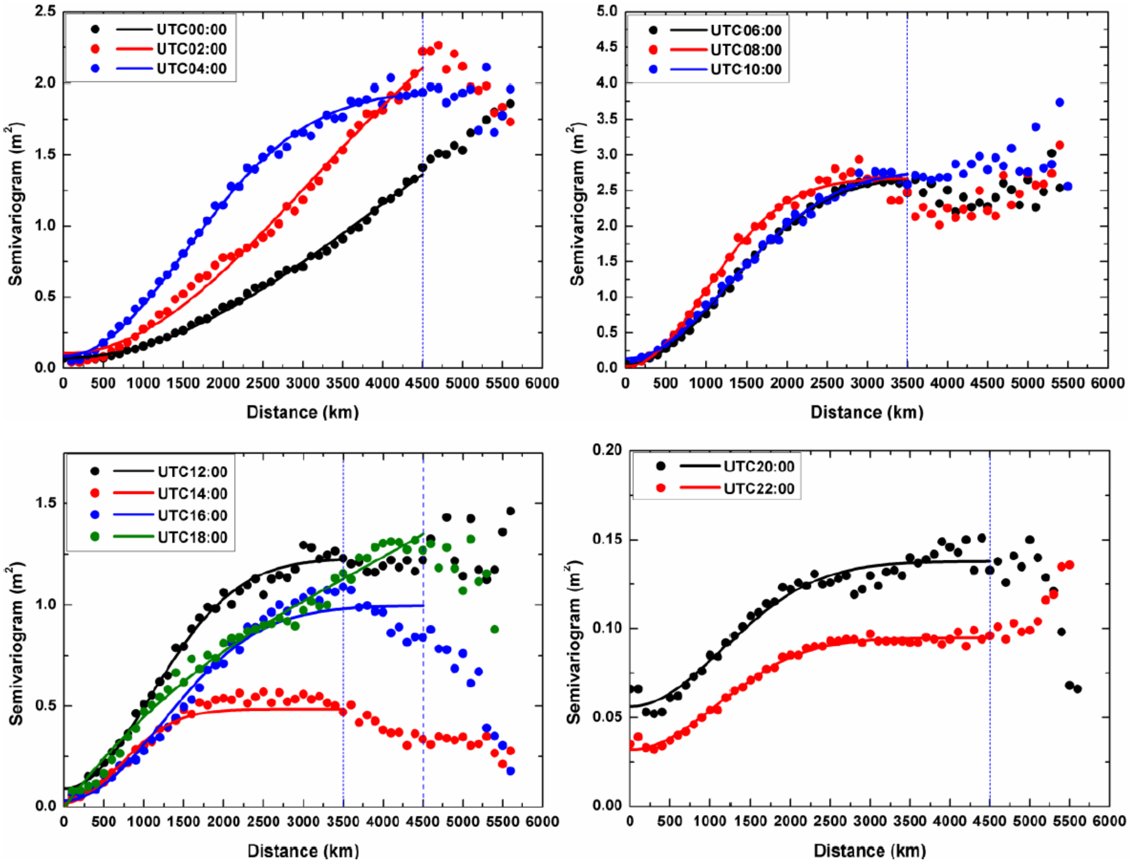

Figure 3.6: Experimental variograms taken from Huang et al. (2017).

The long-wavelength ionospheric delay varies on a scale of thousands of kilometers, and up to magnitudes of several metres. The short-wavelength trend is of greater interest, as the network design is primarily intended for continental faults that do not exceed one to two hundred kilometres in length. Due to the limited number of data points at these distances, additional information is required to constrain the short-wavelength ionospheric delay. It is worth noting that while these variograms provide a good example of how ionospheric delay varies throughout the day, they do not show longer timescale variations, which can be significant (Janssen et al. 2002).

Chen, Kuo et al. (2015) describes methods for reducing ionospheric errors in mixed networks of low-cost single- and dual-frequency GNSS stations. The positional uncertainties caused by the ionospheric delay for both corrected and uncorrected single-frequency GNSS stations are provided, up to a maximum baseline distance of 30 km. Two methods of correcting single-frequency GNSS data are described. The first is a local one-layer ionospheric model, generated from reference dual-frequency GNSS stations. For the second, station specific corrections terms are produced for each single-frequency GNSS stations, using a combination of information from both station types.

Returning to Huang et al. (2017), at short baseline distances it can be seen that the Gaussian curve is convex upwards, and so it is possible to approximate the curve using a second order polynomial. Chen, Kuo et al. (2015) shows that the short-wavelength ionospheric delay is linear when expressed as the standard deviation, measured in millimeters, which becomes quadratic when converted to m^2. Therefore, non-negative quadratics are fit to the squared data uncertainties provided by Chen, Kuo et al. (2015) to produce variograms for the short-wavelength ionospheric delay. The quadratics are constrained to only positive values because negative variance does not make physical sense. Figure 3.7 shows the squared positional uncertainties taken from Chen, Kuo et al. (2015), and the quadratic fits to these data. The corrected variograms share the same quadratic shape as the uncorrected varigorams, an observation that will be returned to further on in this chapter.

Variance values are measured between a single-frequency GNSS station, and a dual-frequency reference GNSS station. The variance of the dual-frequency station is assumed to be zero, and so the positional uncertainties provided are solely dependent upon the single-frequency station. It is therefore necessary to double any values of gamma taken from Figure 3.7 when using two single-frequency GNSS stations, so that the variance of both stations is accounted for. Only the uncorrected single-frequency data is used to produce $\mathbf{Q}_{d,ref}$.

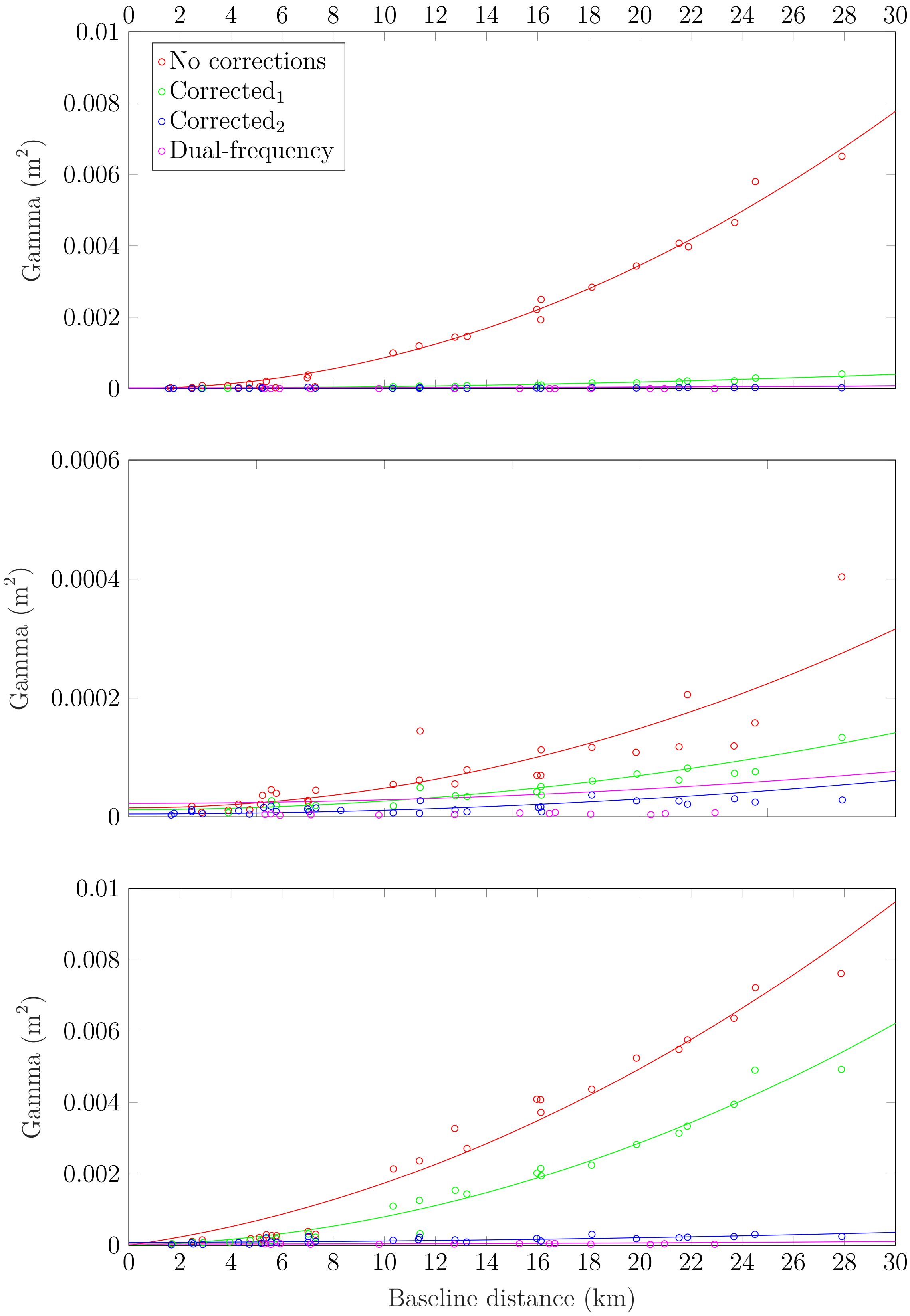

Figure 3.7: Variograms fitted to squared positional uncertainties taken from Chen, Kuo et al. (2015), with non-negative quadratic fits, for the North (top), East (middle), and Up (bottom) components. Corrected$_1$ and Corrected$_2$ refer to the data corrected with the ionospheric model and with the correction terms, respectively.

From these variograms, the required variance and covariance values needed to produce $\mathbf{Q}_{d,ref}$ can be derived. The covariance at distance h can be expressed

$$\begin{aligned} C(h) &= \sigma^2 - \gamma(h) && (3.3.8) \\ &= \gamma(\infty) - \gamma(h) && (3.3.9) \end{aligned}$$

where $\gamma(\infty)$ is the gamma value at infinite distance, which is the sill of the variogram. The correlation at h can be expressed

$$\rho(h) = \frac{C(h)}{\sigma^2} = \frac{\gamma(\infty) - \gamma(h)}{\gamma(\infty)} \qquad (3.3.10)$$

rearranging for variance

$$\begin{aligned} Var &= 2\gamma(\infty)(1 - \rho(h)) \\ &= 2\gamma(\infty)(\frac{\gamma(\infty)}{\gamma(\infty)} - \frac{\gamma(\infty) - \gamma(h)}{\gamma(\infty)} \\ &= 2\gamma(\infty)(\frac{\gamma(h)}{\gamma(\infty)}) \\ &= 2\gamma(h) && (3.3.11) \end{aligned}$$

It can be seen that the variance is independent of the sill ($\sigma^2 = \gamma(\infty)$) assuming that h is less than the range (the distance to the sill).

Covariance may also be expressed in terms of γ

$$\begin{aligned} Cov(d_1, d_2 \,\&\, d_3, d_4) &= \frac{\gamma(\infty)}{\gamma(\infty)}[(\gamma(\infty) - \gamma(h_{d_1,d_4})) + (\gamma(\infty) - \gamma(h_{d_2,d_3})) \\ &\quad - (\gamma(\infty) - \gamma(h_{d_1,d_3})) - (\gamma(\infty) - \gamma(h_{d_2,d_4}))] \\ &= \gamma(h_{d_1,d_4}) + \gamma(h_{d_2,d_3}) - \gamma(h_{d_1,d_3}) - \gamma(h_{d_2,d_4}) && (3.3.12) \end{aligned}$$

Again, it can be seen that the covariance is independent of σ^2, assuming that all station separation distances are less than h_{max}.

To produce $\mathbf{Q}_{d,des}$, the data VCM for the ionospheric delay is produced using Equations (3.3.11) and (3.3.12). Additional variance values, equivalent to twice the estimated dual-frequency station variance, are then added. This represents additional error sources associated with the observations from each single-frequency GNSS stations, and which aren't otherwise included in the ionospheric delay uncertainties. As with the dual-frequency stations, it is assumed that these errors are independent to each station (zero correlation), and so the covariance values within the VCM are unaffected.

3.3.2 Removal of Long-Wavelength Ionospheric Trends

The convex upward shape of the variograms shown in Figure 3.7 indicates that the ionospheric delay contains a long-wavelength planar trend (i.e. is not stationary). This results in high covariance between pairs of stations that are aligned parallel to each other, and low correlation between pairs aligned perpendicular. High data covariance is propagated through to high model covariance, degrading the ability of the network to resolve distributed slip. It was previously noted that both corrected and uncorrected single-frequency data sets share this convex upward shape, indicating that the long-wavelength trend is not removed by local ionospheric delay corrections.

The Chen, Kuo et al. (2015) and Huang et al. (2017) data sets are combined by fitting Gaussian functions to the short-wavelength data from Chen, Kuo et al. (2015), and constraining the range, sill, and nugget of these functions using the long-wavelength trends from Huang et al. (2017). Minimum and maximum values for each are provided in Table 3.1. These curves are shown in Figures 3.8 and 3.9. It can be seen that the fit to the Chen, Kuo et al. (2015) is worse than when using an otherwise unconstrained non-negative quadratic (Fig. 3.7), which is a trade-off to have the long-wavelength trend constrained by the values provided by Huang et al.

	Range (km)	Sill (m^2)	Nugget (m^2)
Min	0	0	0
Max	5000	10	0.5

Table 3.1: Constraints for the Gaussian functions shown in Figures 3.8 and 3.9.

(2017). If the curves are unconstrained, a better fit is achieved, but with unrealistic values for the sill, indicating that it is poorly constrained by the short-wavelength data.

These variograms are then used to generate one hundred random spatial fields, over a 100 km by 100 km area. An example field is shown in Figure 3.10. Each field is a synthetic ionospheric delay map, where the spatial structure of the delay can be characterised by the provided variogram. A planar trend is then removed from each of these fields, and a experimental variogram is then fit to the residual delay. Finally, a spherical model is fit to experimental variograms for each component, to produce updated variograms with the long-wavelength ionospheric delay removed (Fig. 3.11).

Figure 3.12 shows the correlation values calculated from the data VCMs of a four station single-frequency network. The first is from a VCM generated using the original non-negative quadratic variograms, and the second using the spherical variograms with the long-wavelength trends removed. The values in the former are significantly higher than the latter, which posses near-zero correlation values for all but the main diagonal. High correlation values will propagate to high model correlation, and so removal of the planar trend before the network design is a necessity.

The new variograms may now be used to produce $\mathbf{Q}_{d,des}$, and so $\mathbf{Q}_{m,des}$ may be estimated for the network design. Removal of the planar trend from the variograms significantly reduces the correlation between pairs of stations. The magnitude of the data uncertainties are also reduced, which propagates through to lower model uncertainties.

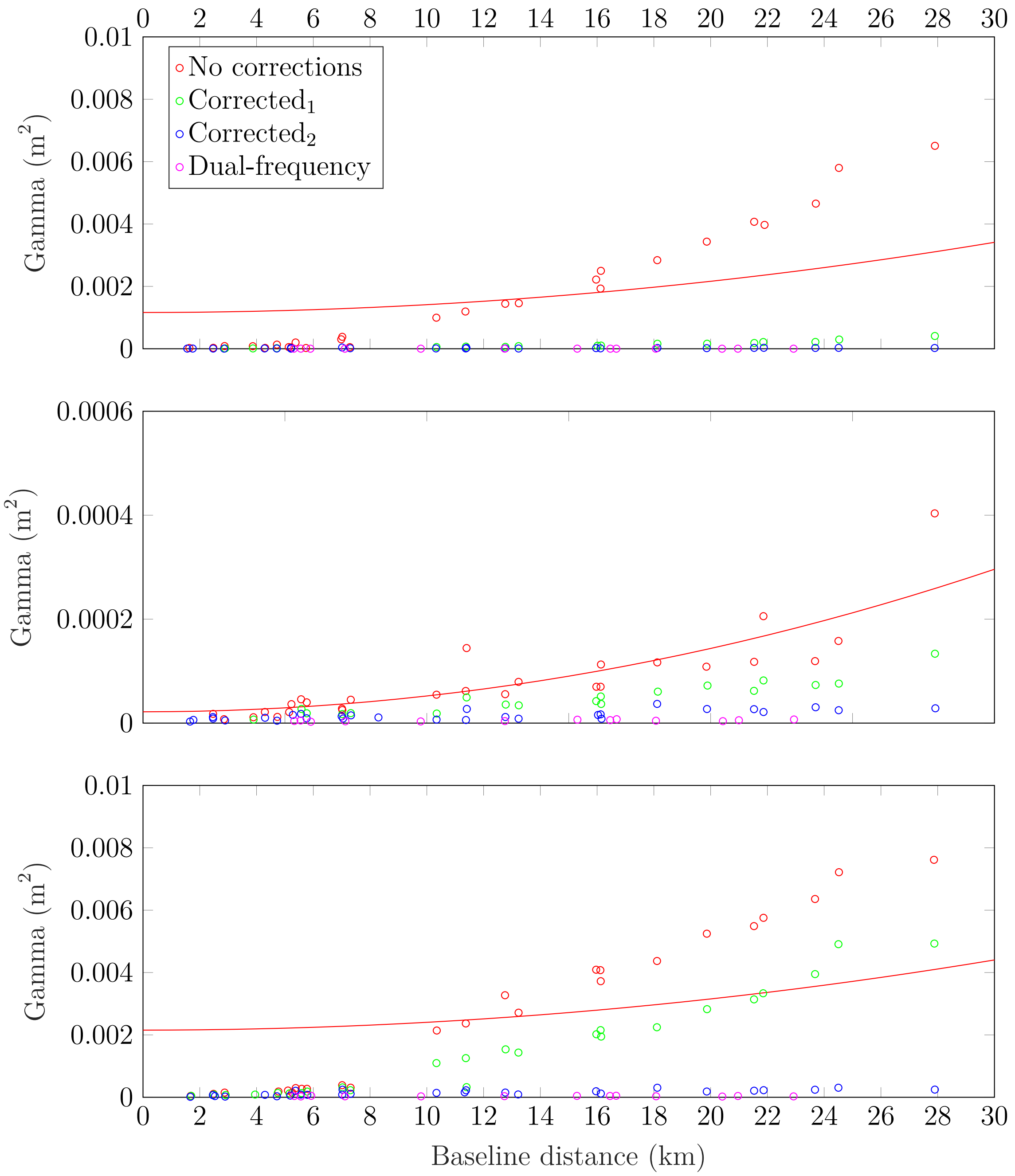

Figure 3.8: Gaussian fit to the squared standard deviation error (variance) for uncorrected single-frequency stations provided by Chen, Kuo et al. (2015), for the North (top), East (middle) and Up (bottom) components of displacement. The corrected variance values, along with the dual-frequency variance values, are provided for reference. Each component is shown separately so that the quality of the fit to the data is more easily seen.

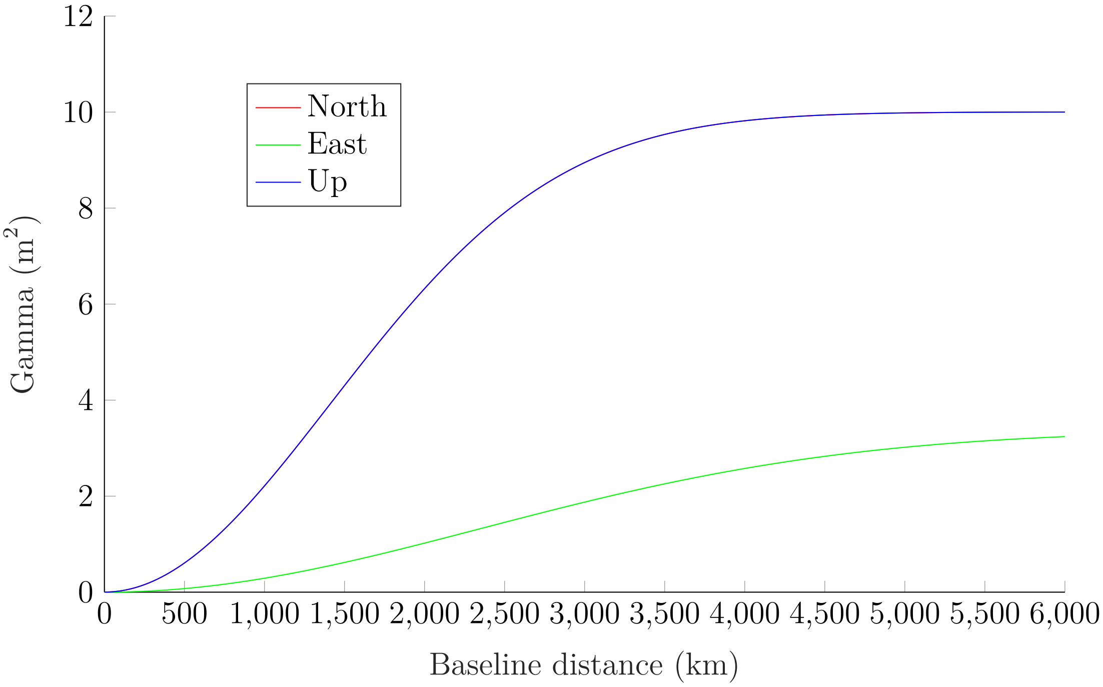

Figure 3.9: Gaussian fit to the squared standard deviation error (variance) for uncorrected single-frequency stations provided by Chen, Kuo et al. (2015), shown up to the full range of the functions. The lines for the North and Up components overlap.

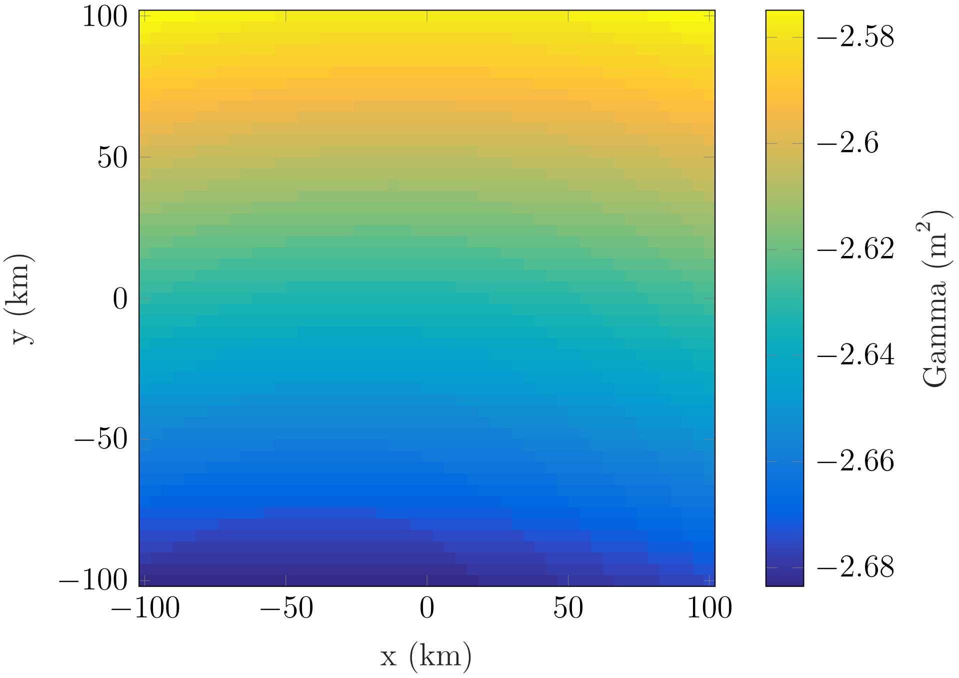

Figure 3.10: An example random field generated by the Gaussian ionospheric model for the East componenet.

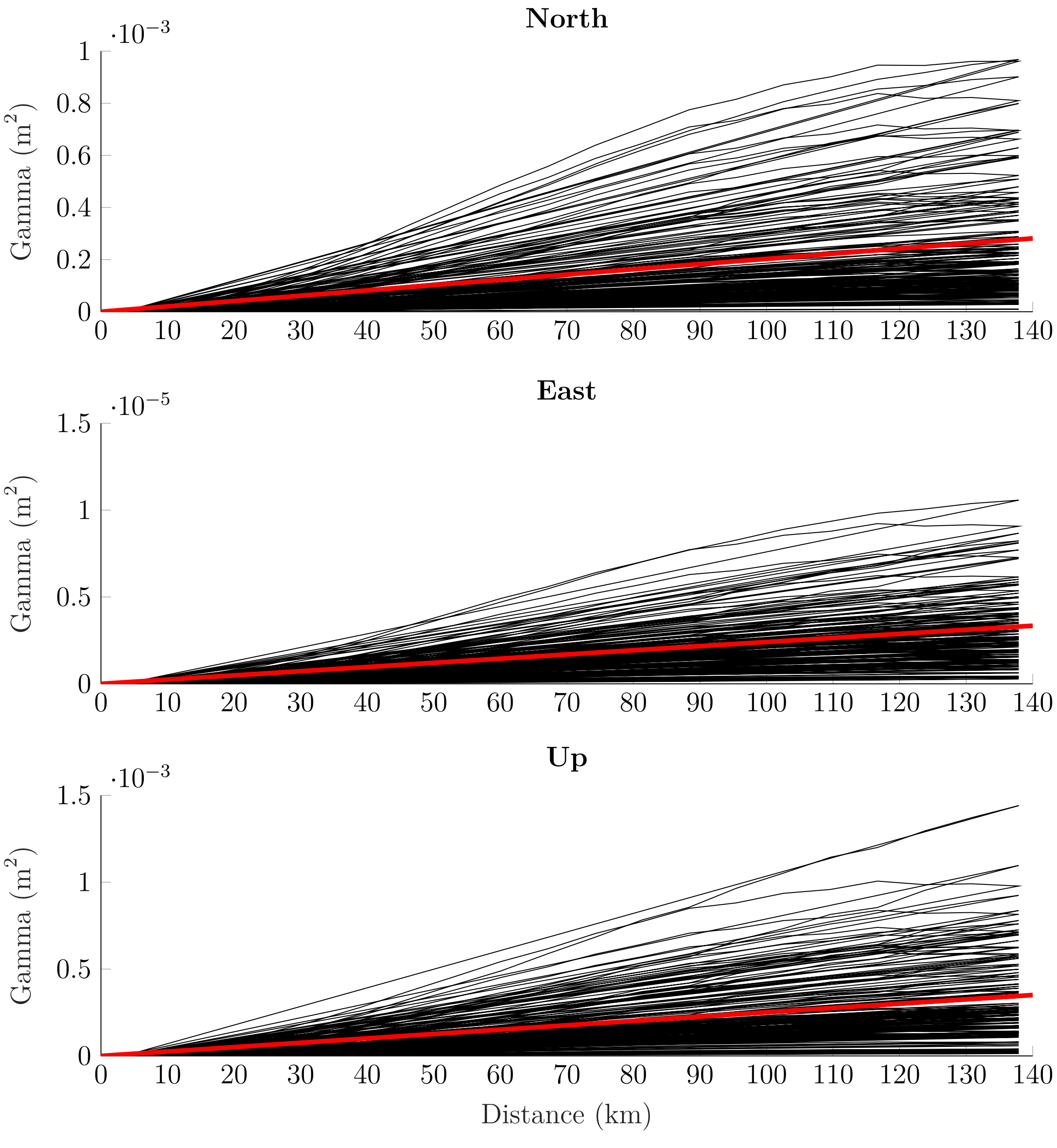

Figure 3.11: One hundred experimental variograms (black lines) generated from random fields with a planar trend removed, and the best fit spherical models (red lines).

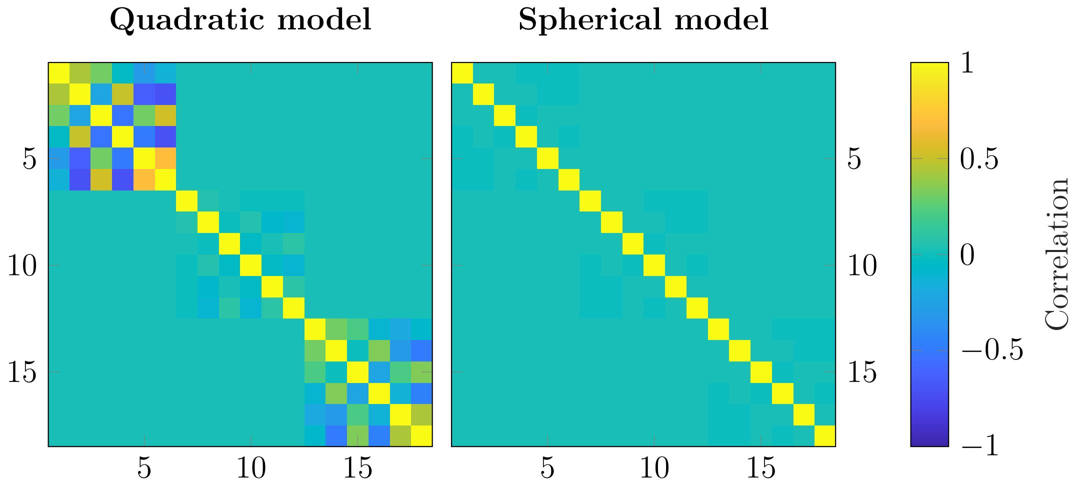

Figure 3.12: Correlation values for a network of four single-frequency stations, using a non-negative quadratic variogram (left) and a spherical variogram with the long-wavelength ionospheric trends removed (right).

Chapter 4

Results

The methodology described in Chapter 3 is now applied to a range of synthetic model faults. Optimal networks of low-cost dual-frequency GNSS stations are generated first, to act as a reference for the more complicated single-frequency GNSS networks.

4.1 Dual-Frequency GNSS

Dual-frequency examples are presented first, starting with a run through of the full network design method for a simple normal fault. The affects of varying network size and fault geometry are then explored.

4.1.1 Idealised Network Structure

The idealised network is used to derive $\mathbf{Q}_{d,ref}$, which is used to estimate $\mathbf{Q}_{m,ref}$ as part of the fault discretisation. The idealised network consists of two regions of uniformly-spaced GNSS stations; a high station density area centred on the fault trace, and a low station density area that extends to the limits model. To ensure consistency across the results, we define two fixed station separations for these regions; 2 km for the high-density area, and 4 km for the low density area. An idealised network possessing these station separation values in shown in Figure 4.1.

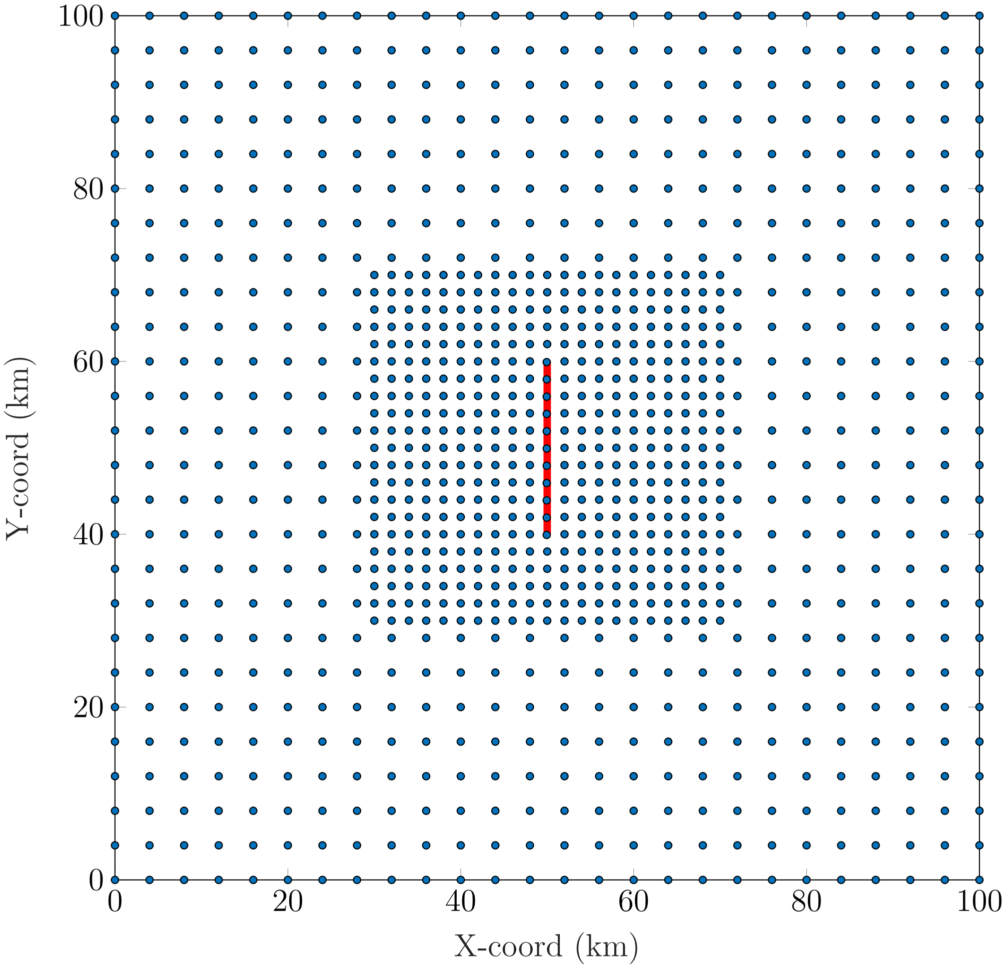

Figure 4.1: An idealised network of dual-frequency GNSS stations that is used to produce $\mathbf{Q}_{d,ref}$ in the following examples. The bold red line indicates the fault trace, while the red area shows the vertically projected fault plane.

These values are chosen as a trade-off between run time, as a higher density of stations contains more observations to process, and the model error values that the idealised network is able to achieve. The structure and variance values of $\mathbf{Q}_{d,ref}$ are given in Section 3.1.1.

4.1.2 An Example Normal Fault

The fault discretisation and network design methods are first tested on a synthetic normal fault. The model fault geometry is described in Table 4.1.

$\mathbf{Q}_{d,ref}$ is produced using the idealised network shown in Figure 4.1. The fault

Parameter	Value
Strike	180°
Dip	60°
Rake	−90°
Fault length	20 km
Top depth	1 m
Bottom depth	15 km

Table 4.1: Parameters for a normal fault. The top depth of the fault is set to 1 m rather than 0 m, as a value of zero causes the half-space equations to fail.

discretisation is ran three times with varying numbers of rows. The resulting solutions are re-tested for dominance to produce the combined estimated Pareto front, and an optimal fault discretisation is selected (Fig. 4.2). The chosen discretisation was picked based upon the desire for a maximum model error value of 5 cm. As expected, patch size increases with depth, and model uncertainty is close to uniform across the fault plane, ranging from roughly 3.55 - 4 cm. These values do vary along rows, a result of the station separation which means that the proximity of stations to each patch varies by a small amount within each row.

With a discretisation selected, the network optimisation is performed with seven dual-frequency GNSS stations, taking a total of 206 iterations to achieve acceptable convergence. The designed network, along with the model uncertainties for both the idealised network and the designed network, are shown in Figure 4.3.

The lowest model uncertainty achieved by the designed network is 5 cm for the top left patch, with the majority possessing values that are between two to five times larger than was observed by the idealised network. The largest values are observed on the deepest patches, as the network does not contain enough stations in the far-field of the region to provide resolving power at depth. A maximum model uncertainty of 18 cm is still be considered a success, however, when considering that this is achieved using only seven stations compared to the 917 stations present in the idealised network.

The correlation between uncertainties on either observations or model parameters

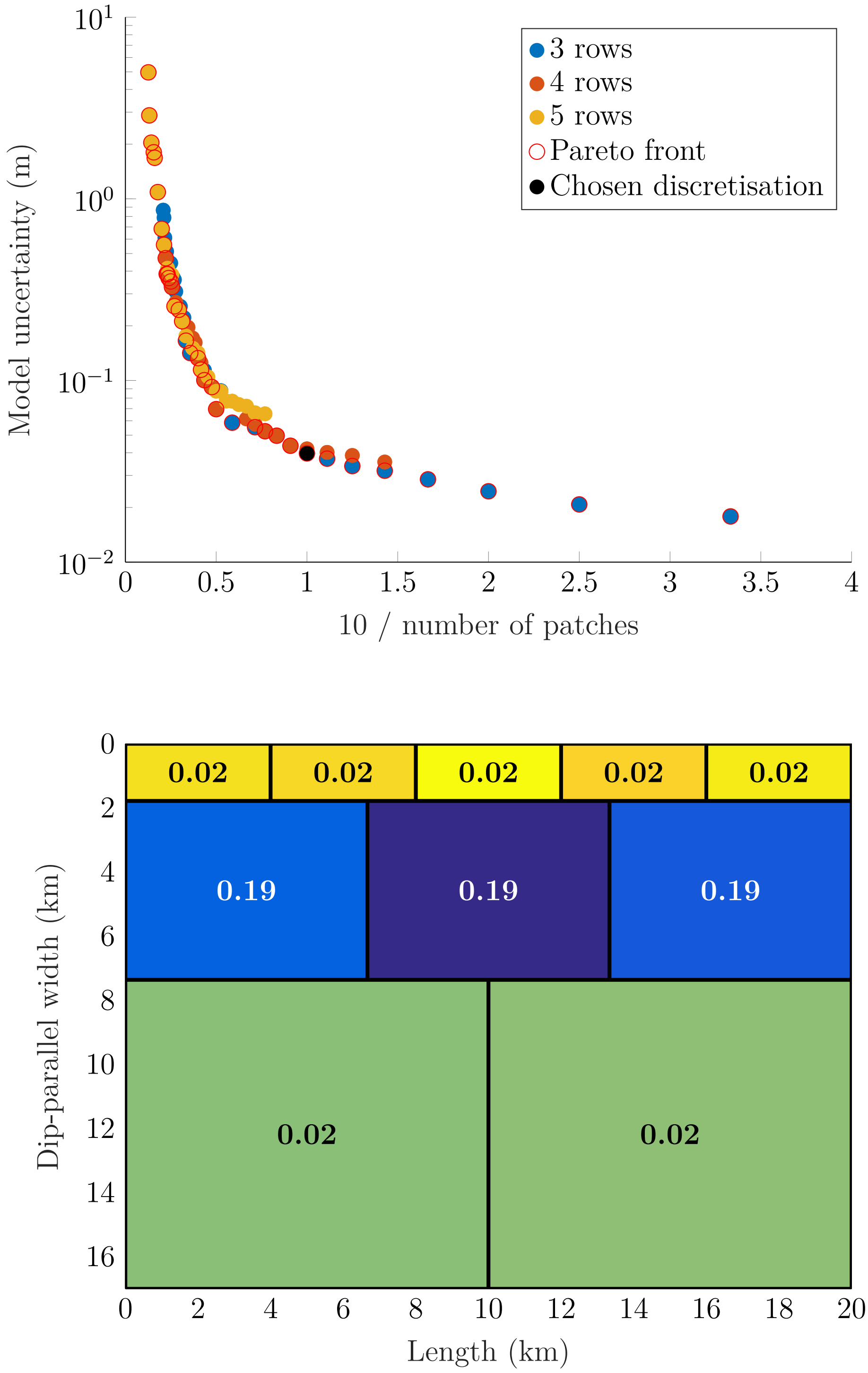

Figure 4.2: The combined pareto front (top) created from three optimisations. The chosen discretisation (bottom) is highlighted in black, with 1σ model uncertainties for each patch given in metres.

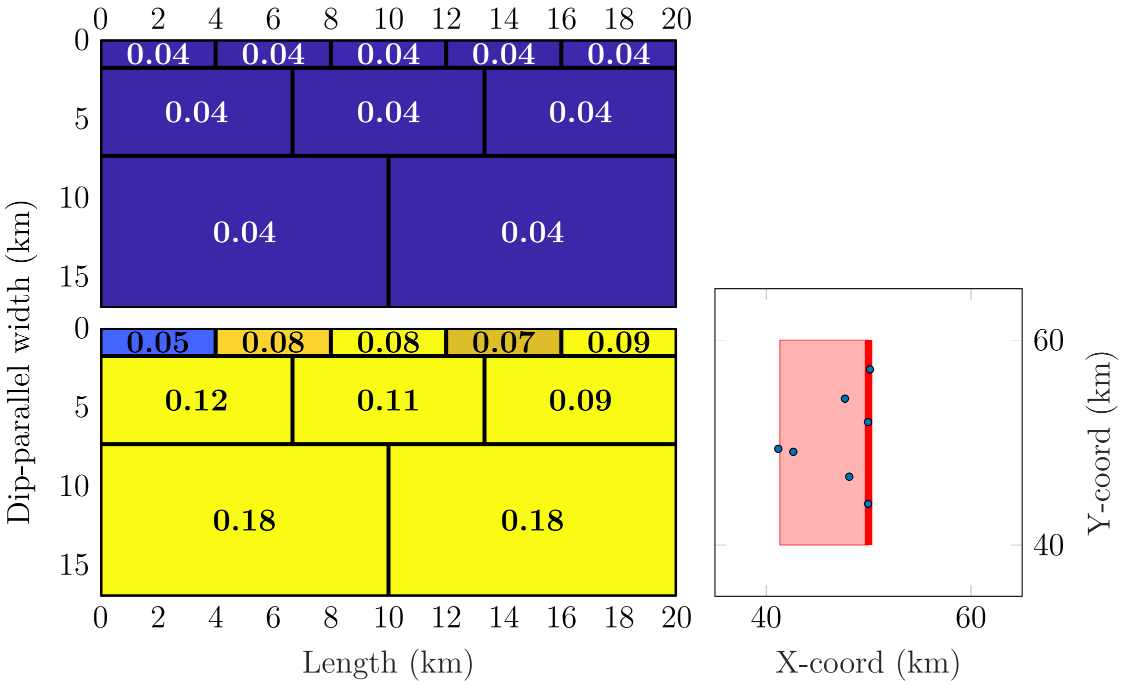

Figure 4.3: The 1σ model uncertainties (m) for both the idealised (top) and designed networks (bottom), with the final optimal network of seven single-frequency GNSS stations.

can be calculated using

$$\rho_{a,b} = \frac{Cov(a,b)}{Var(a)Var(b)} \tag{4.1.1}$$

This can be repeated for a full VCM to produce a correlation matrix. Figure 4.4 shows the correlation matrices produced from $\mathbf{Q}_{m,ref}$ and $\mathbf{Q}_{m,des}$. Both matrices show similar patterns, however, the designed network has multiple adjacent patches with high positive correlation, shown by the sets of four adjacent squares along the main diagonal. While correlation values are lower for the idealised network, there are still non-zero values.

This demonstrates the ill-conditioned nature of the inverse problem. Given perfect data, it would be possible to recover the slip distribution perfectly with the current network of stations, as we have 21 observations (three components for each station) and 10 model parameters. However, real-world data will always contain some amount of measurement uncertainty, which is mapped through to model parameters by the linear inversion.

Figure 4.5 shows the correlation values betweeen the middle patch (coloured

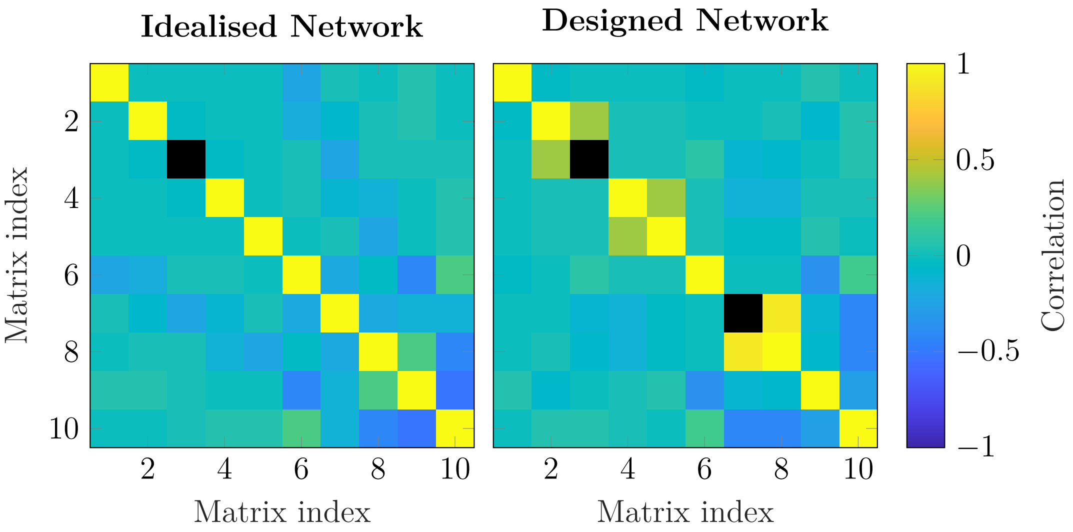

Figure 4.4: Full correlation matrix for the idealised network (left) and designed network (right).

grey) of the discretisation and every other slip patch, again for both the idealised and designed networks.

Both networks result in entirely non-zero correlation values, as expected. These values are overall higher for the designed network, especially in the case of the right adjacent patch which is near to fully correlated with the centre patch. This is likely caused by a lack of stations measuring displacement from these patches, with those that do lying fully in a the region deformed by both patches, and so it is difficult to discern between the displacement caused by either patch. Also of interest is that some patches, specifically some of those on the top row, have lower correlation values for the designed network than the idealised network. This is a potential benefit of having few stations, as fewer observations in regions of maximum displacement overlap will lead to lower correlation. The downside is that the model error increases, as seen in Figure 4.2.

Given the magnitude of the correlations observed in Figure 4.4, it is important that the model VCM is fully examined for any optimal generated networks. While covariance values will be minimised during the optimisation, as they are included in the distance metric, they may still be problematically high, despite the network producing acceptable variance values.

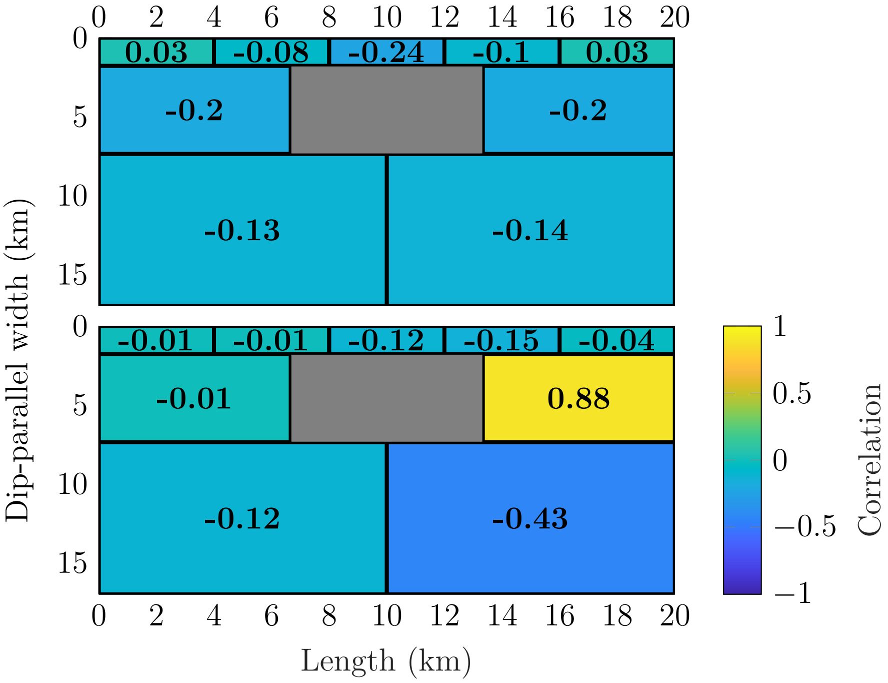

Figure 4.5: Model correlation between the selected fault patch (grey) and all other patches, for the idealised network (top) and the design network (bottom).

4.1.3 Network size

Following on from the previous example, the effects of network size on both the layout of optimal GNSS networks, and the magntiude of the model uncertainties estimated by these networks, are explored. Increasing the number of GNSS stations in a network is expected to reduce the model uncertainties estimated on the fault. The normal fault geometry described in Table 4.1 is repeated, although with a new discretisation (4.6). The new discretisation consists of 17 slip patches with a maximum model uncertainty of 20 cm, an increase from the 10 patches and 5 cm maximum uncertainty observed in the previous discretisation. The new discretisation is selected so that the affects of increasing network size are clearer.

Figure 4.6 shows three optimal networks consisting of five, ten, and twenty dual-frequency stations. The maximum model error ranges from 1.2 m for the five station network, down to 30 cm for the twenty station network. Increasing the number of stations in the network leads to a reduction in the average model uncertainty across the discretisation. Note that the five station network is an ill-posed problem, as the

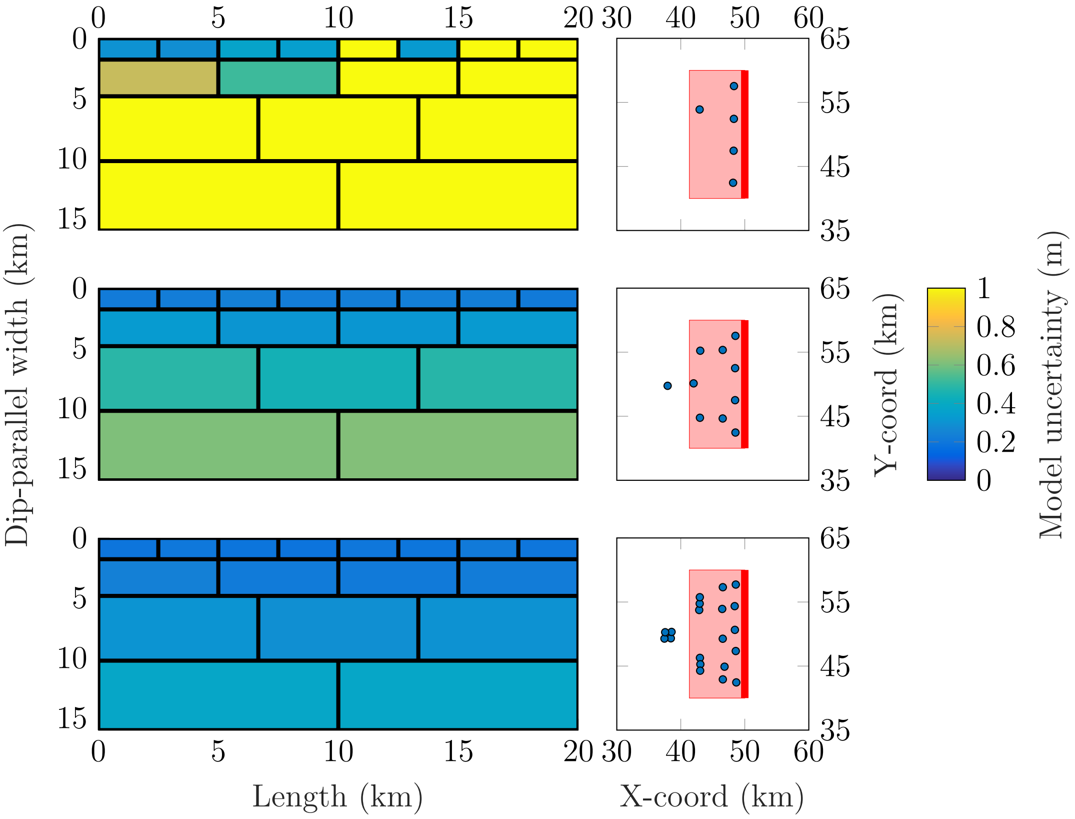

Figure 4.6: Three optimal dual-frequency GNSS networks consisting of five, ten, and twenty stations, with the model uncertainties derived from $\mathbf{Q}_{m,des}$ for each network shown on the discretisation.

number of model parameters (17) is greater than the number of observations (15), which will contribute to the high model uncertainties. Model resolution may also be degraded.

Similarities in design exist between all three networks. Each is close to symmetrical about a line perpendicular to the fault top, which is to be expected given that the discretisation is itself symmetrical about the vertical centre line of the fault plane. Four to five stations are placed parallel to the projected fault top, at on offset of several kilometers. These stations lie above the first two rows of fault patches, lowering the model error across them. Both the ten and twenty stations networks posses a second row of stations beyond these, and over the second and third rows of the discretisation. When not enough stations are available (10 station network), this second row is placed over the midpoints of both halves of the discretisation. With more stations available, a more uniform spacing is achieved similarly to the first parallel row.

For deeper patches, stations are located increasingly centrally and further out. In the twenty station network, we see that the algorithm begins to place stations very close together as opposed to uniformly spaced lines. This reflects the decreasing number of patches at depth, and the need to reduce the model error by attempting to minimise the data error at a given point, by placing multiple observations in the same location. This demonstrates the effect of the artificial variance curve for close together stations. Without this there stations would become very densely packed, with separations of tens to hundreds of meters, which in the real-world would be begin to cause issues with short-wavelength correlated errors such as multipath. However, it does still suggest that for sensing deep slip on fault planes, there is some benefit for locating a number of stations within relatively close proximity, as opposed to spacing them further out.

The end result of this is networks that are triangular in shape, pointing towards the deepest end of the fault plane, and mirroring the decreasing number of patches with depth in the discretisation. Within each network, stations are also placed in triangular arrangements due to the decreasing number of slip patches with decreasing row depth.

Klein et al. (2017) generates optimal GNSS locations to support an existing network of dual-frequency GNSS stations for the study of aseismic slip on the Marmara fault. The fault is discretised manually, with increasing slip patch size with depth. The network design is performed using a grid search and swapping method, as described in Section 2.2.5, to maximise model resolution across the discretisation. The optimal GNSS networks presented by Klein et al. (2017) show similar features to the dual-frequency GNSS networks shown in Figure 4.6. The majority of the GNSS stations are placed close to the fault trace, with a small number located at a greater distance from the fault trace. This creates a triangular structure of stations, as observed in Figure 4.6.

As described in Section 2.2.3, network economy refers to the costs associated with deploying a geodetic network. Economy is not directly included in the network design, however, the trade-off between network size and distance metric value is of interested for assessing the returns given by deploying more GNSS stations. Figure 4.7 shows the trade-off between the distance metric and the number of dual-frequency stations in the designed network. Increasing the number of stations in the network decreases the distance metric value achieved, bringing $\mathbf{Q}_{m,des}$ closer to $\mathbf{Q}_{m,ref}$. The change in the distance metric value decreases with increasing network size. The reduced return in how well distributed aseismic slip can be resolved by the network will therefore impact the size of deployed networks.

4.1.4 Fault geometry

Both $\mathbf{Q}_{m,des}$ and $\mathbf{Q}_{m,ref}$ are dependent upon $\mathbf{G}$, which is a function of the fault geometry and observation locations. It is therefore expected that changing the fault geometry will alter the layout of the optimally generate GNSS networks. Four fault parameters are selected to vary fault geometry: the stike angle, the dip angle, the

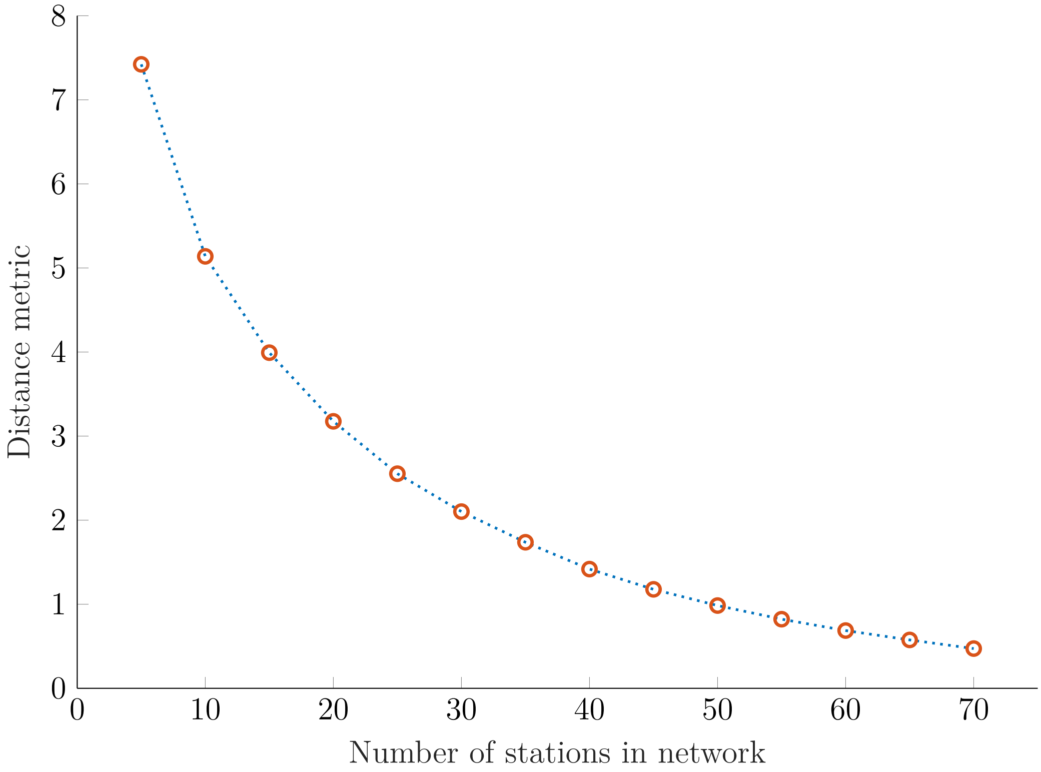

Figure 4.7: Distance metric vs the number of dual-frequency GNSS stations in the optimally designed network.

rake angle, and the bottom depth of the fault plane. All other fault parameters are fixed, so that the impact of varying any one parameter may be better isolated.

Orientation

Figure 4.8 shows three networks, each containing 20 dual-frequency stations, used to model the same normal fault at three different strikes. The model fault geometry is otherwise as described in Table 4.1. $\mathbf{Q}_{d,ref}$ is generated using the three idealised networks shown in Figure 4.1, with the alignment of the high-station-density area matched to the respective fault strike.

All three networks show very little change in model error based on fault orientation, with the exception of a high-valued slip patch in the shallowest row that occurs in different locations for two networks, although this has a lower error in the third. This patch may be lacking in the east-west striking fault due to a small shift in the locations of two stations. As the data VCM is not affected by fault orienta-

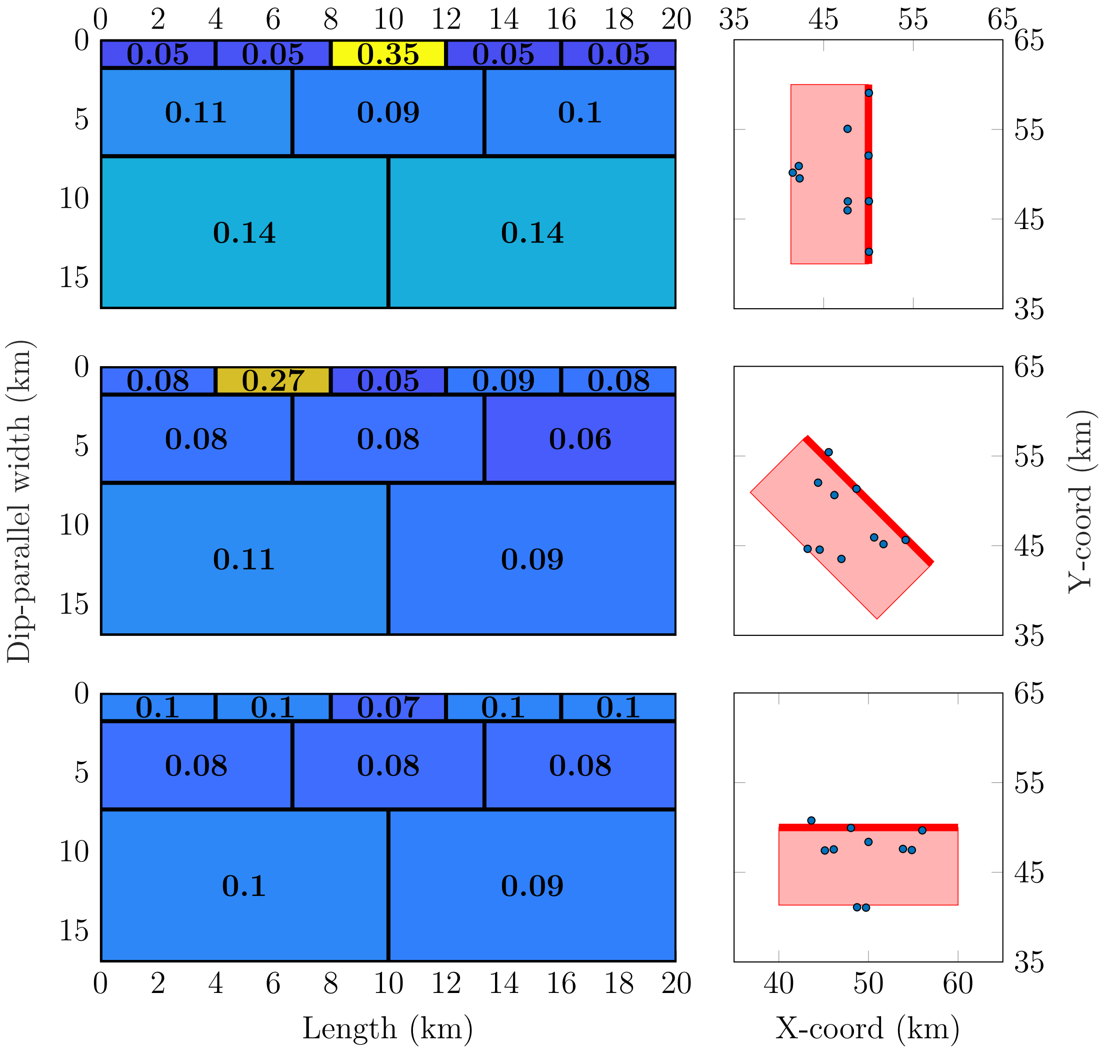

Figure 4.8: Optimal networks of 20 dual-frequency GNSS stations generated for the normal fault geometry described in Table 4.1 and rotated to strikes of 180° (Top), 90° (Middle), and 135° (Bottom). Estimated model uncertainties (m) are shown on the discretisation.

tion, and so no effect is expected on the model VCM, it is likely that this network achieved a slightly more optimal state than the other two, likely due to random chance within the optimisation. These results are as expected for the dual-frequency GNSS stations.

Depth and fault type

We next explore the effects of fault depth and fault plane area, and the type of fault slip. When creating the fault model, it is not required that the dimensions match

Parameter	Normal	Strike-slip
Dip	60°	90°
Rake	$-90°$	0°

Table 4.2: Table showing the change in rake and slip values required to alternate between a normal and stike-slip fault geometry.

those of the real-world fault it is attempting to replicate. Instead, the fault model can be created to target slip on a given part of the fault plane, for example to target shallow afterslip following an earthquake vs afterslip anywhere on the fault plane. Changing the depth and surface area of the fault plane may affect the layout of optimal networks, as this will change the displacement field around the fault. The same is true for the type of faulting, whether that be normal faulting, strike-slip, or thrust faulting.

Previously we have used a normal fault geometry, as described in Table 4.1, which we now modify to create a strike-slip fault. This is done by varying the dip and rake angles. The values used are given in Table 4.2. The depth of the fault is varied by changing the bottom depth to 3 km, while fixing the top depth.

Figure 4.6 shows four networks, generated for four different fault geometries. The top two rows are from a normal fault and a strike-slip fault, both with a bottom depth of 15 km. The bottom two rows are the same, but with a bottom depth of 3 km.

As previously shown in Figure 4.6, the networks for normal faults show a degree of symmetry about a line that intersects the fault trace at it's centre, and with a perpendicular orientation. For strike slip faults, this symmetry is also present about a line that runs along and parallel to the fault trace, as can be seen in Figure 4.9. This is caused by the symmetrical nature of deformation around strike-slip faults.

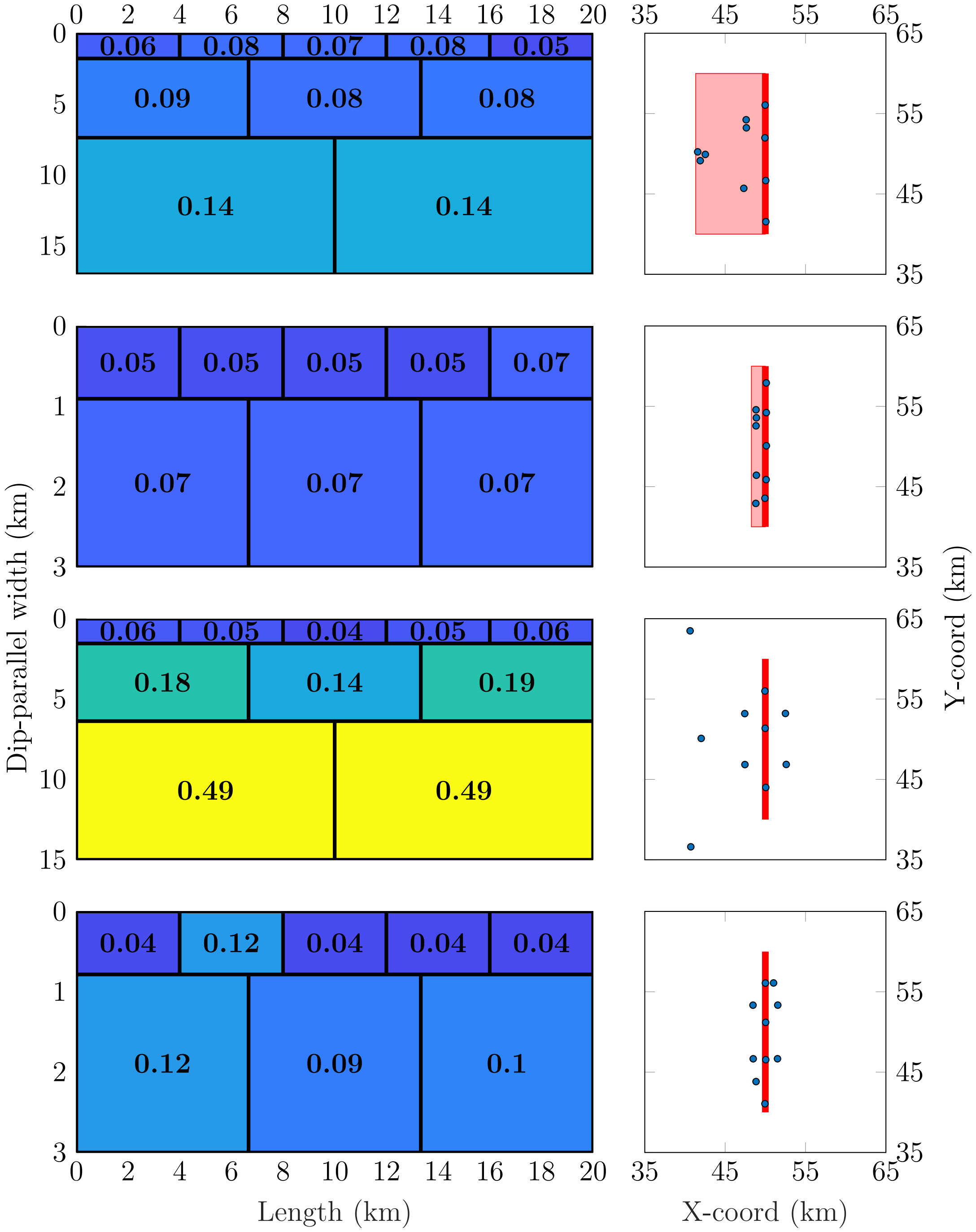

Figure 4.9: Optimal networks of ten dual-frequency GNSS stations designed around a normal fault with a 15 km bottom depth (top), a 3 km bottom depth (second from the top), and a strike-slip fault with the a bottom depth of 15 km (second from the bottom), and a 3 km bottom depth (bottom). Estimated model uncertainties are given in metres.

4.2 Single-Frequency GNSS

4.2.1 Example networks

The starting fault geometry is that of the normal fault described in Table 4.1, with a manually generated discretisation of three patches over two rows. This is to provide a very simple discretisation to begin with, while still possesing enough complexity to require proper design of the network. $\mathbf{Q}_{d,ref}$ is generated by the idealised network shown in Figure 4.1 and used to estimate $\mathbf{Q}_{m,ref}$.

Figure 4.10 shows the the estimated layout of an optimal four station single-frequency GNSS network. $\mathbf{Q}_{m,ref}$ contains very low model uncertainties of 1 - 2 cm, which is expected given the sizes of the slip patches. The designed network produces model uncertainties that are several times larger, but with a maximum value of 11 cm. A discernible structure is present within the designed network, with the centre-points of the baselines between pairs of stations positioned directly above the boundary between the two rows of fault patches. While the slip patches in the discretisation are very large, the model uncertainties achieved are promising.

The number of slip patches in the discretisation is now increased, from three to seven slip patches, and the number of stations in the designed network are increased from four to ten. $\mathbf{Q}_{d,ref}$ is unchanged, however, a new $\mathbf{Q}_{m,ref}$ is generated because the change in discretisation affects $\mathbf{G}$. Figure 4.11 shows the estimated layout of the optimal single-frequency GNSS network.

$\mathbf{Q}_{m,ref}$ again contains very low model uncertainties at a uniform 3 cm across all slip patches in the discretisation. The designed network achieves roughly double that at 5 - 7 cm on the top row of patches, with model uncertainty then increasing with depth. A return to the triangular stations layouts previously identified in the dual-frequency GNSS networks (Fig. 4.6) is observed, where two stations are placed close to the fault trace with a third station placed over the deeper fault patches. Four stations are not located directly over the fault plane, three of which are at a significant

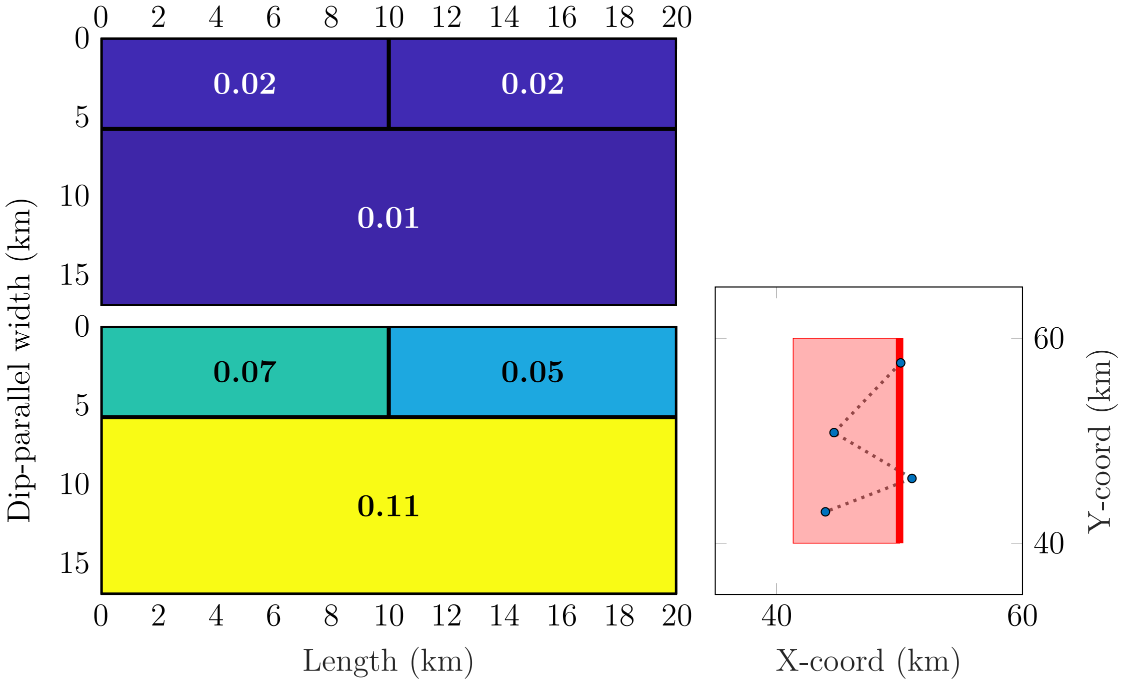

Figure 4.10: Ten station optimal single-frequency network for a normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. The dotted lines indicate the pairs of stations between which relative measurements of ground displacement are made.

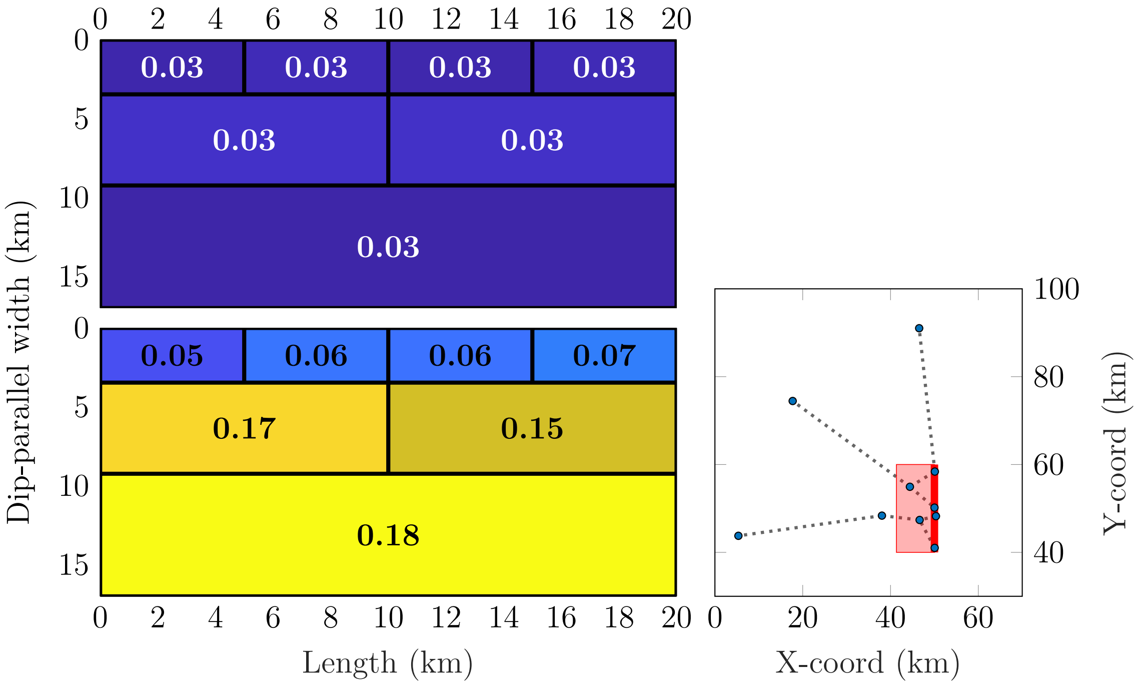

Figure 4.11: Ten station optimal single-frequency network for a normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. The dotted lines indicate the pairs of stations between which relative measurements of ground displacement are made.

distance (30 km) away from the fault. These stations sample the displacement in the far-field of the region, where the greatest difference between the displacement contributions from the deep and shallow patches is observed. Displacement from the shallow patches is still present, but at a significantly lower magnitude than the displacement from the deep patches.

Figure 4.12 shows from nine evenly-spaced iterations throughout the network design optimisation. The optimisation allows iterations where the network neither improves or declines as the solution space is explored, and so these have been removed to smooth out the optimisation process and to avoid exactly repeated network layouts.

The distance metric improves rapidly at the start of the optimisation, decreasing by 39.6% in the first 13 iterations and by 51.6% by the end of the inversion (96 iterations). The final structure of the network is recognisable by the 37th iteration, with only the locations of the far-field stations varying significantly for the remainder of the optimisation. The final third of the optimisation consists of small station movements which are not visible at this scale, and which produce changes to the distance metric on the order of 10^{-3} .

4.2.2 Alternative Fault Geometries

As demonstrated by the dual-frequency GNSS results, the geometry of the fault will impact the layout of the optimal GNSS networks. This is expected to hold true for the single-frequency GNSS stations. The normal fault from Table 4.1 is again modified to give a strike-slip fault with rake and dip values given in Table 4.2. $\mathbf{Q}_{d,ref}$ and the fault discretisation are retained from the previous example. $\mathbf{Q}_{m,ref}$ is re-estimated because the changes to the fault geometry affect $\mathbf{G}$.

Figure 4.13 shows the the estimated layout of an optimal ten station single-frequency GNSS network. $\mathbf{Q}_{m,ref}$ contains values equivalent to model uncertainties of 2 - 7 cm and which increase with depth. The increase in model uncertainty compared to Figure 4.11 demonstrates the increased difficulty in resolving slip on strike-slip faults, both due to the high dip angle and the surface-parallel slip direction. Similarly, the model uncertainties estimated by the designed network are higher, especially at depth, with a maximum value of 59 cm on the deepest slip patch.

The network again contains triangular layouts of stations which now wrap around the fault plane at either end. Far-field stations are also present, and the network shows a high degree of symmetry about a line that extends perpendicular from the midpoint of the fault trace. In contrast to the network shown in Figure 4.11 which was designed for a normal fault geometry, GNSS stations are located on both sides of the fault trace.

The fault model is now returned to that of the normal fault geometry described in Table 4.1, with the bottom depth adjusted to 3 km as previously shown in Figure 4.9. The fault is discretised using the automated discretisation method. $\mathbf{Q}_{d,ref}$ is retained

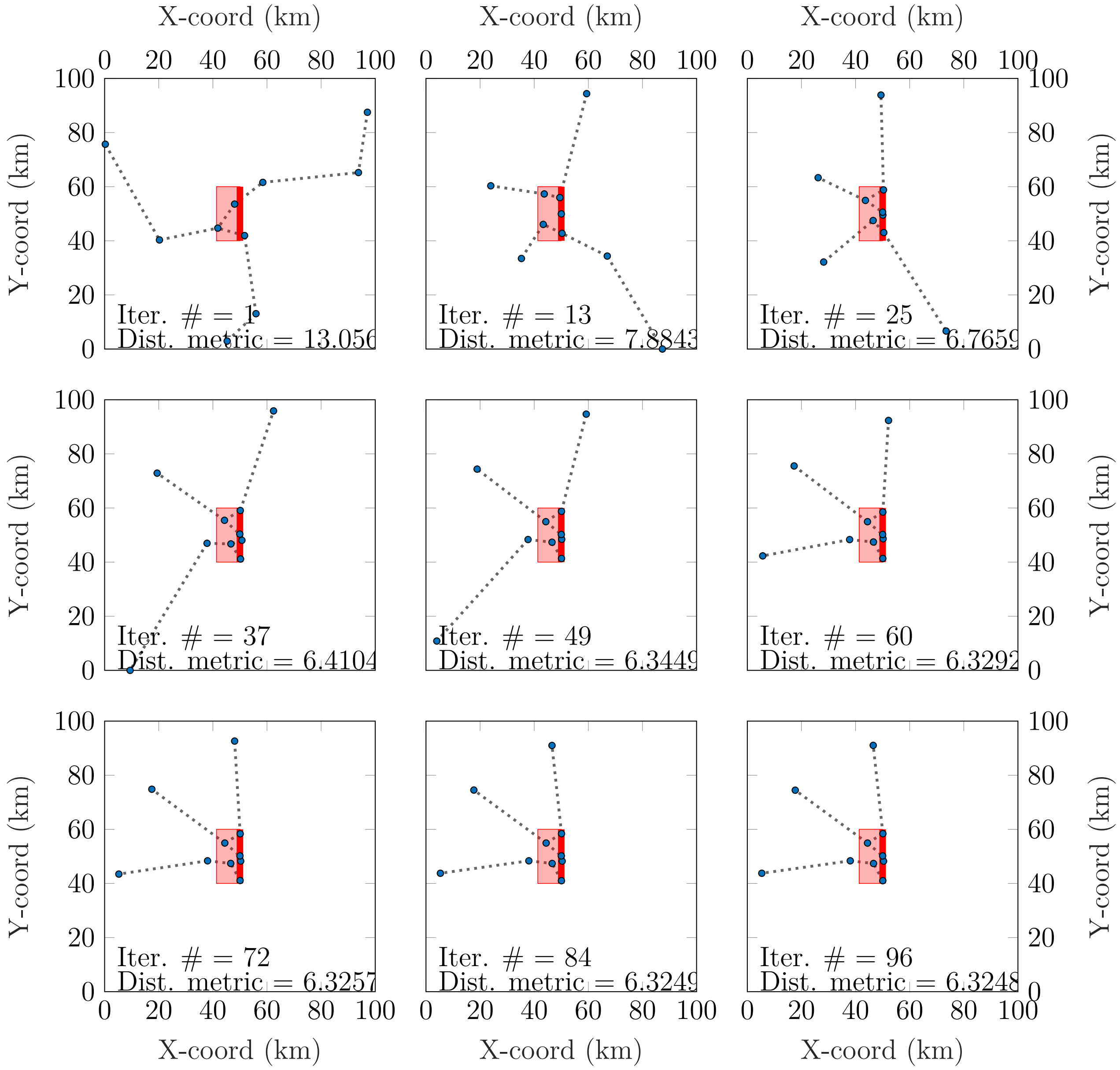

Figure 4.12: The network at nine evenly spaced iterations throughout the inversion, with iterations where the distance metric did not change removed. The final plot (iteration 96) is the optimal network. Iteration number and distance metric value for each network layout is given. Dotted lines indicate station pairings.

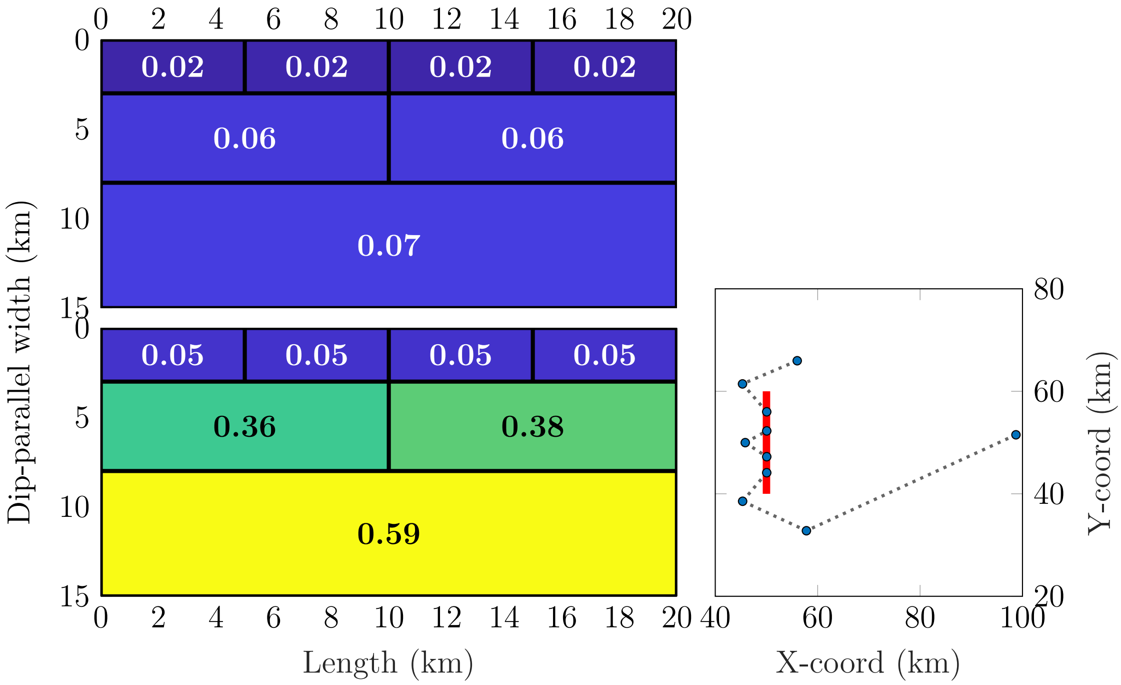

Figure 4.13: Ten station optimal single-frequency network for a vertical strike-slip fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation.

and $\mathbf{Q}_{m,ref}$ remade to account for the modified fault geometry and discretisation.

Figure 4.14 shows the the estimated layout of an optimal ten station network. $\mathbf{Q}_{m,ref}$ contains values equivalent to a uniform model uncertainty of 5 cm across all slip patches. Model uncertainties estimated by the design single-frequency network range from 6 - 9 cm, generally increasing with depth. Values are also overall higher on the northern half of the fault plane compared to the southern half, which is caused by a greater number of stations being located directly above the southern half the fault plane. Estimated model uncertainties are similar in magnitude to those observed in the top row of the discretisation in Figure 4.11, but achieved over a greater number of slip patches. This reflects the greater number of observations directly above the fault plane, as no far-field stations are required to measure displacement from deeper slip patches. This highlights the importance of designing networks only for the section of the fault that is of interested (e.g. only modeling the top of the fault plane if only shallow slip is desired).

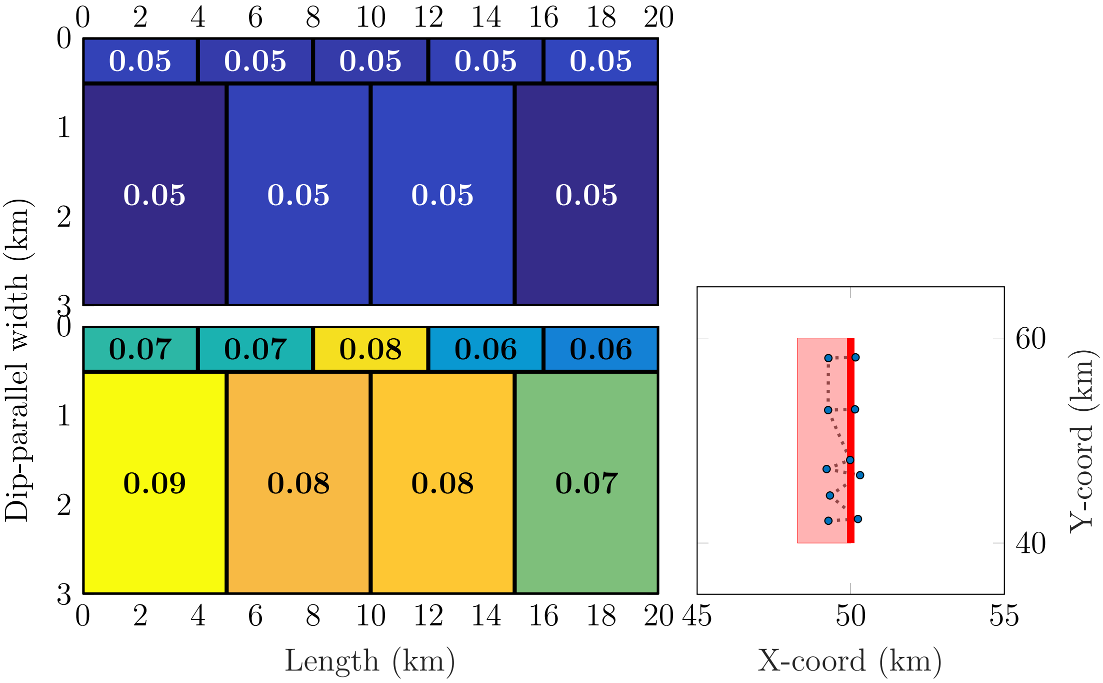

Figure 4.14: Ten station optimal single-frequency network for a shallow (3km bottom depth) normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation. Note that the x-axis of the network has been stretched so as to make the layout clearer.

The structure of the design network varies from those shown in previous examples. Stations located over the northern half of the fault plane are placed at right angle to each other. Stations over the southern half of the fault plane are roughly placed in two strike-parallel lines either side of the fault trace. As previously stated, no far-field stations are present due to the lack of deeper slip patches. The change from triangular layouts of stations may be caused by the similarity in the number of patches between both rows, and also due to a greater clustering of stations over the fault plane.

For the final fault geometry variation, the normal fault geometry described in Table 4.1 is reused with the strike changed to 90°. The discretisation shown in Figure 4.11 is retained, along with $\mathbf{Q}_{d,ref}$. $\mathbf{Q}_{m,ref}$ is estimated again due to the change in fault geometry, although the resulting model uncertainties are almost identical to those shown in Figure 4.11.

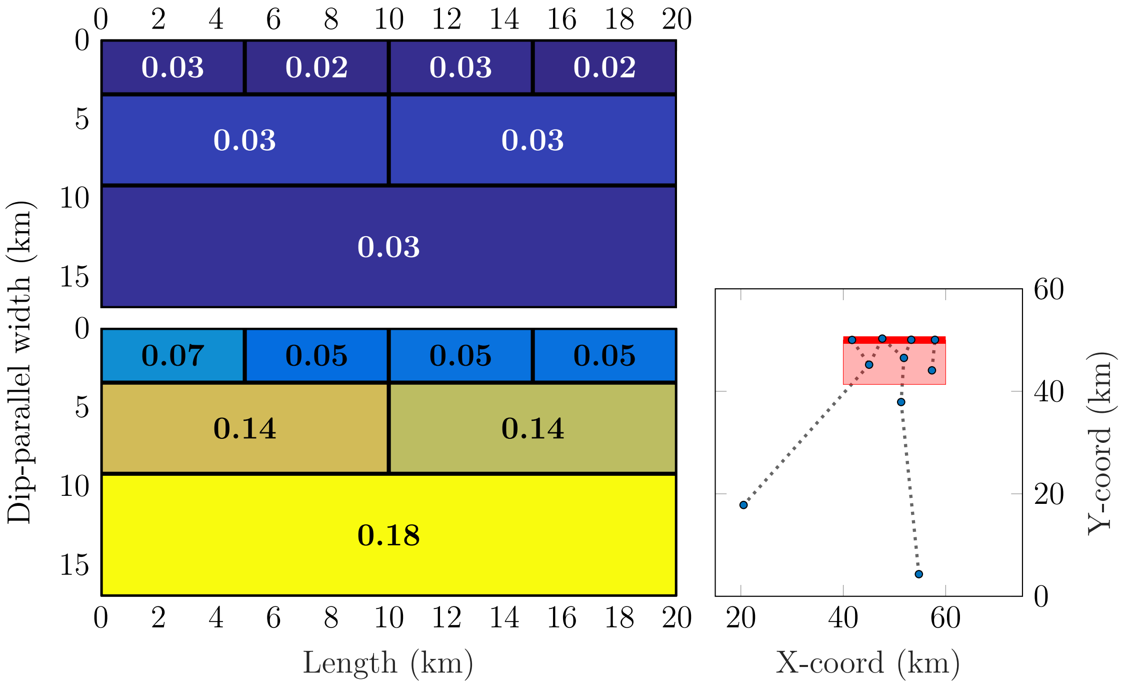

Figure 4.15: Ten station optimal single-frequency network for an east-west striking normal fault, with the model uncertainties (m) estimated by the idealised (top) and designed (bottom) networks shown on the discretisation.

Figure 4.15 shows the optimal network layout with ten stations. The structure of the network is markedly similar to the network shown in Figure 4.11, both possessing triangular layouts of stations directly above the fault plane and additional stations in the far-field. The primary difference between the two networks is the relocation of a previously far-field station to be directly over the fault plane. The estimated model uncertainties have improved slightly, with a maximum change of 3 cm on the middle left patch (14 cm vs 17 cm). This improvement is likely the result of chance in the optimisation finding a more optimal solution on this run. The difference may also be the result of the different variograms for the north and east components of the ionospheric delay.

4.2.3 Dual-Frequency Comparison

The normal fault geometry described in Table 4.1 is returned to again to provide a comparison between the single- and dual-frequency network designs. The fault discretisation and optimal dual-frequency network shown in Figure 4.2 are reused. A new single-frequency network of seven stations is generated for comparison.

Figure 4.16 shows the two optimal networks and their respective estimated model uncertainties. Model uncertainties estimated by the single-frequency network are higher on all slip patches except for two in the shallowest row, which are 1 - 2 cm lower. The difference in model uncertainties between the two networks becomes larger with increasing depth.

The same triangular layout of stations that has been observed in previous examples is again present in both networks. The single-frequency network contains fewer stations placed directly over the lowest patches which contributes to the larger model uncertainties on the deeper slip patches. This is a result of having fewer observations in the single-frequency network, and indicates that the shallower slip patches have a greater impact on the distance metric than the deeper slip patches.

The model uncertainties estimated by the dual-frequency network can be achieved by an optimal single-frequency network of 12 stations (Fig. 4.16). The triangular layout of stations above the fault plane is again repeated, with the addition of two stations in the far-field that constrain slip on the deepest patches.

Figure 4.17 shows the model correlation values derived from $\mathbf{Q}_{m,des}$ for both the single- and dual-frequency networks. While the overall correlation between fault patches appears to be higher for the single-frequency network, the correlation values for some adjacent patches are actually lower than those observed from the dual-frequency network.

Figure 4.18 is a graphic comparison of $\mathbf{Q}_{d,des}$ for both networks. Variance values are higher for the single-frequency network, with very low covariance values observed for both networks.

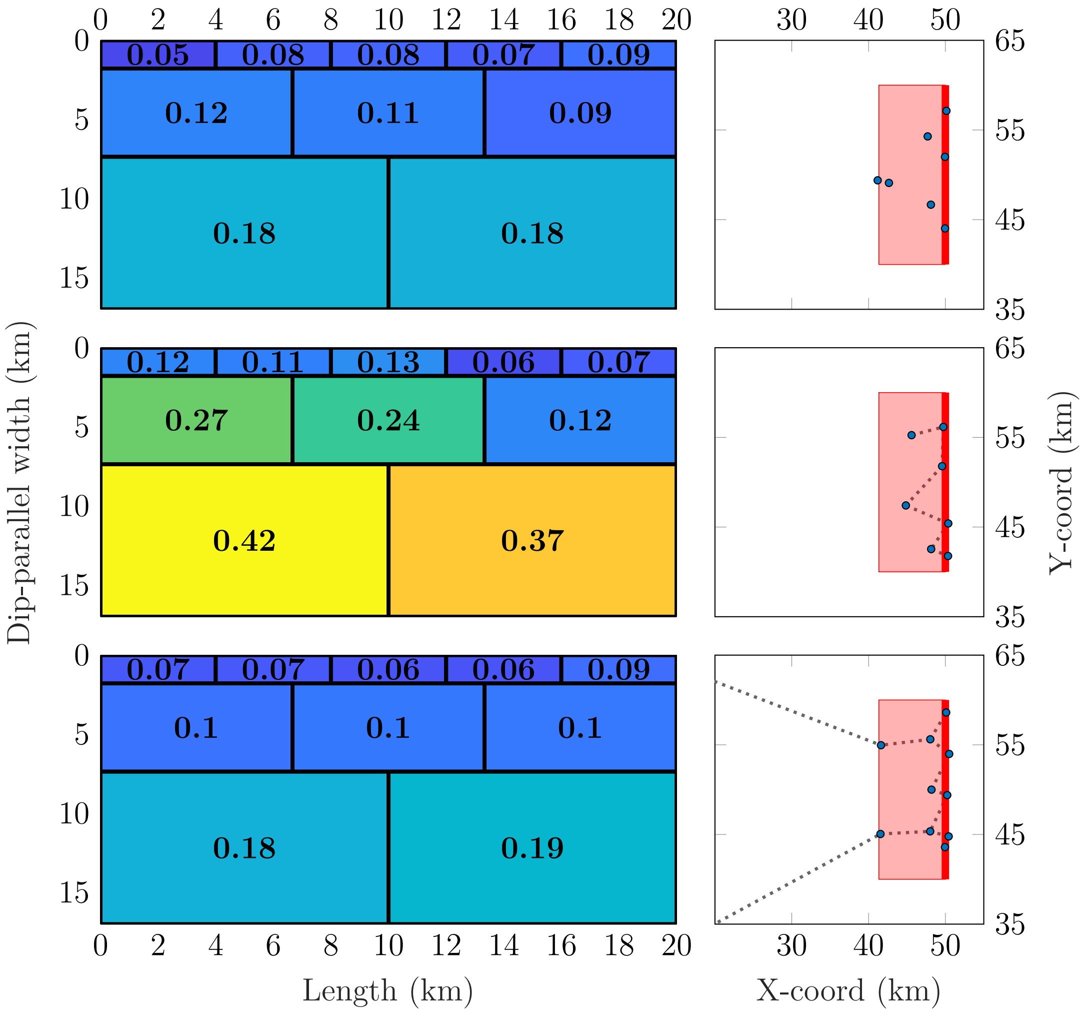

Figure 4.16: Optimal networks and estimated model uncertainties for a 7 station dual-frequency network (top), a 7 station single-frequency network (middle), and a 12 station single-frequency network (bottom).

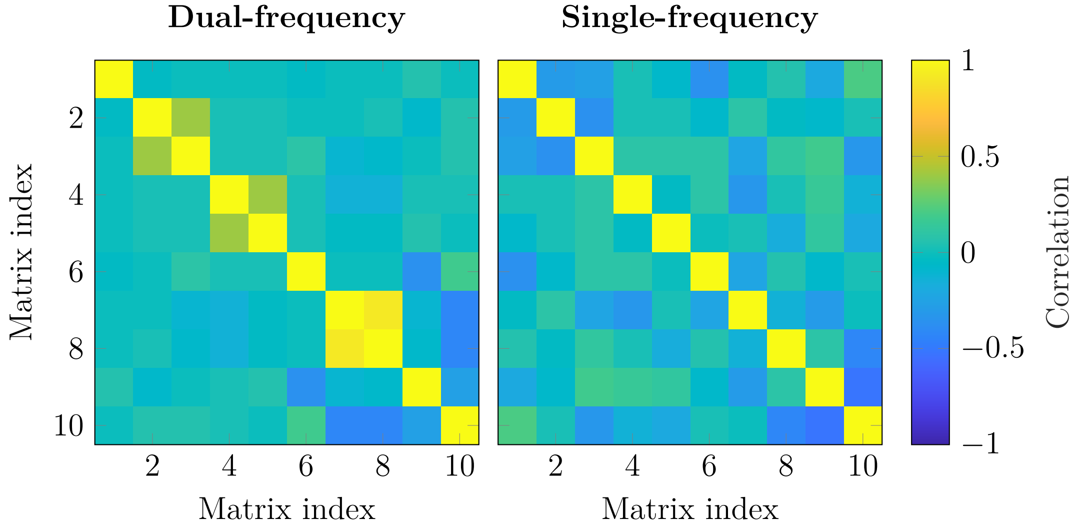

Figure 4.17: Model correlations from the single- and dual-frequency networks shown in Figure 4.16.

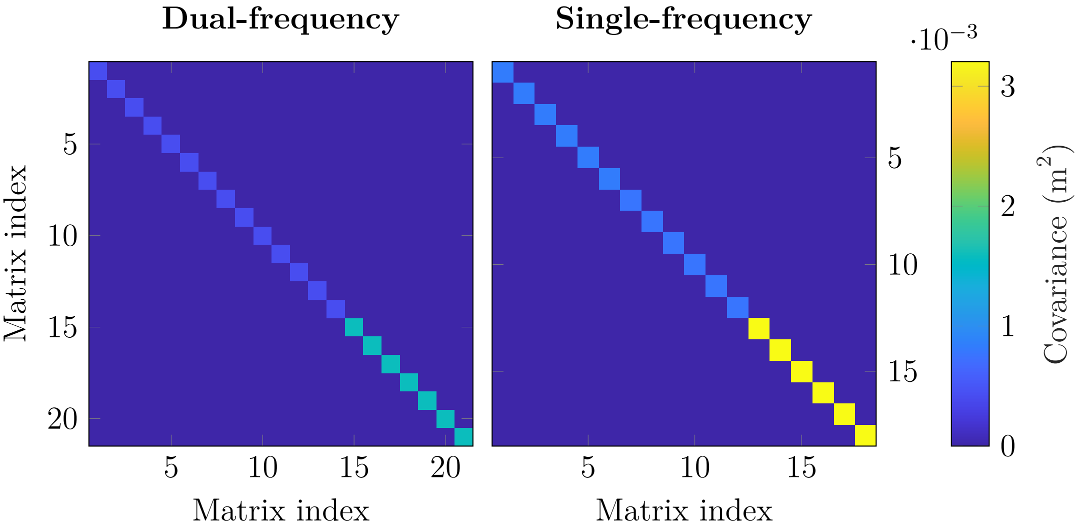

Figure 4.18: $\mathbf{Q}_{d,des}$ for the single- and dual-frequency networks shown in Figure 4.16.

Chapter 5

Discussion

In this Chapter, the previous results are summarised and further examined. Several limitations of the methodology are explored, and potential solutions provided. Finally, future expansions of the methodology are described.

5.1 Summary of Results

5.1.1 Potential of Single-Frequency GNSS

Through Chapters 3 and 4, a method for the estimating the optimal layouts of low-cost GNSS networks to recover distributed aseismic slip on continental faults has been provided and demonstrated. Of particular interest is the use of single-frequency GNSS stations, for which a method for correcting the ionospheric delay has been presented and included within the network design. Single-frequency GNSS stations present the largest potential for improvement in network size and density.

The examples presented in Chapter 4 demonstrate that average model uncertainties of 5 - 25 cm may be achieved by relatively small networks of single-frequency GNSS stations, depending upon the fault geometry and the discretisation. Increasing network sizes further would allow for lower model uncertainties on estimated slip values, and a greater number of slip patches in the discretisation. If the number of stations is fixed, then the number of slip patches in the discretisation can be varied until the model uncertainties are reasonable for the expected magnitude of aseismic slip. This flexability means that single-frequency stations may be used to study any magnitude of aseismic slip thta produces surface deformation above the observation uncertaintiy, although potentially with a very coarse discretisation. If the number of observations is kept above the number of slip patches, so that the inverse problem is not ill-posed, then all slip values will be perfectly resolved by the observations.

5.1.2 Optimisation Repeatability

A key characteristic of the metaheuristic optimisation algorithms used to perform the fault discretisation and the network design, is that it cannot be guaranteed that the final solution is the true global optimum. It is also not possible to definitively prove that a solution is the global optimum, only that it is more optimal than all other observed local minima. The final solution may also vary for over multiple runs of the optimisation, due to the randomised components inherent to both algorithms (e.g. starting model parameter values). It is therefore desirable that, while the model parameters that define the final solution may change, the quality of the solution is repeatable.

The MOGA used to discretise the fault plane estimates a range of solutions that define the pareto front. Given that the pareto front defines a continuous spectrum between the two fitness functions, the generated discretisations are able to vary along this curve while remaining optimal. Figure 5.1 shows the solutions from five runs of the MOGA. Solutions are repeated for much of the pareto front, with the greatest variation where discretisations have the highest number of slip patches. Where solutions are not perfectly repeated, variation is commonly along the pareto front, representing equally optimal solutions with a different weighting of the fitness functions. Some solutions are fully dominated by solutions found in different runs.

In the context of the network design, it is desired that the distance metric value,

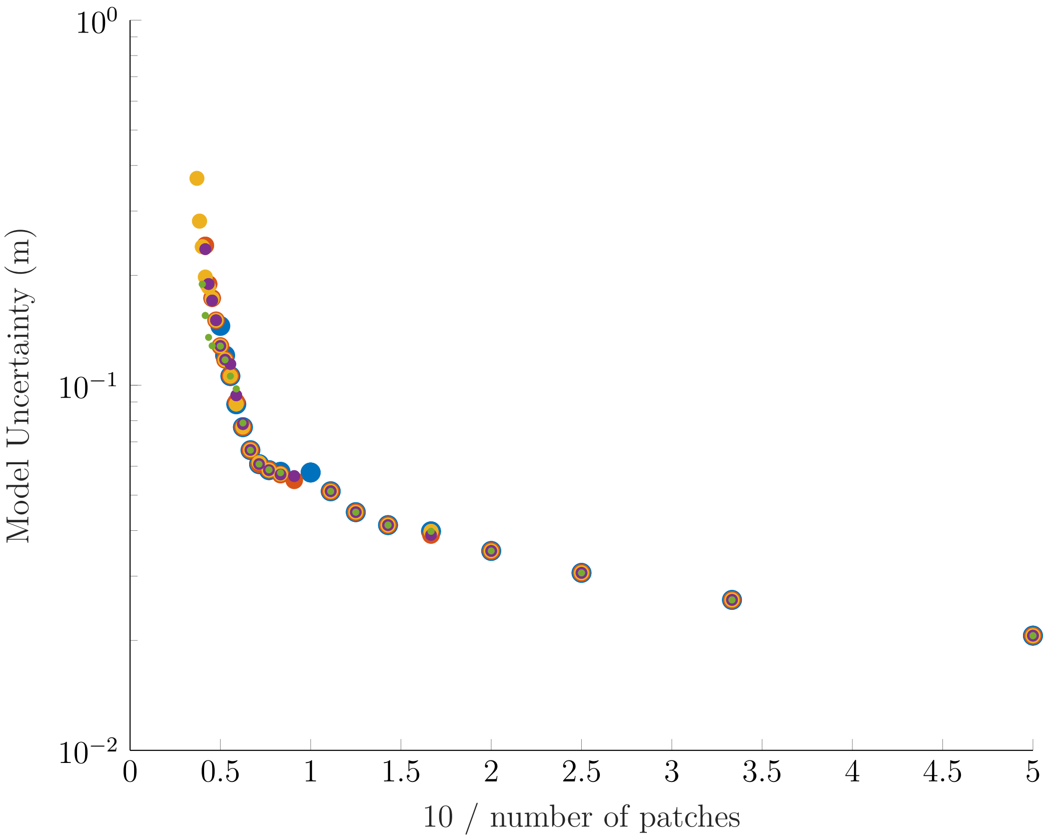

Figure 5.1: Results from five runs of the MOGA for a two row discretisation of a normal fault with bottom depth of 3 km (e.g. Figure 4.14). Solutions are grouped by colour, with marker size varying so that overlapping solutions are visible.

and therefore the model uncertainties, is repeatable even if the exact network layout is not. Repeated structures within optimal networks are also desired, both as proof that the network design is working, and for the deployment of stations, which will be discussed further in this section.

In Chapter 4, it was noted that similar structures appeared in many of the example networks. These primarily consisted of triangular layouts of stations positioned directly over the fault plane, and far-field stations, which were observed for both types of GNSS station. This strongly implies some logic to the network design, and that the optimisation is working as intended.

Figure 5.2 shows six single-frequency GNSS networks designed around an east-west striking normal fault. The highest ranked network is also shown in Figure 4.15. Both the distance metric and the maximum estimated model uncertainty are provided, with a reasonable degree of positive correlation present between both values as expected.

The smallest variation in location is observed for stations placed directly over the fault plane. These as again arranged in triangular shapes in all six networks, although the exact positions do vary. All six networks also contain stations in the far-field of the region, which show the greatest variation in locations, and the three highest-ranked networks contain at least one station placed within 20 km of the fault trace but not directly over the fault plane itself. The presence of stations directly above the fault, arranged in similar layouts, in all six networks indicates the importance of these stations to reducing the distance metric. The variability in far-field locations shows that these stations have a smaller impact on the distance metric, and that their effect is less dependent upon the exact location. However, these far-field stations are still present in all six networks, reflecting their importance for constraining slip on the deeper slip patches.

The ability to vary exact station locations while maintaining network quality is beneficial for the deployment of real-world GNSS networks. The optimisation estimates optimal station locations assuming that all positions around the fault are

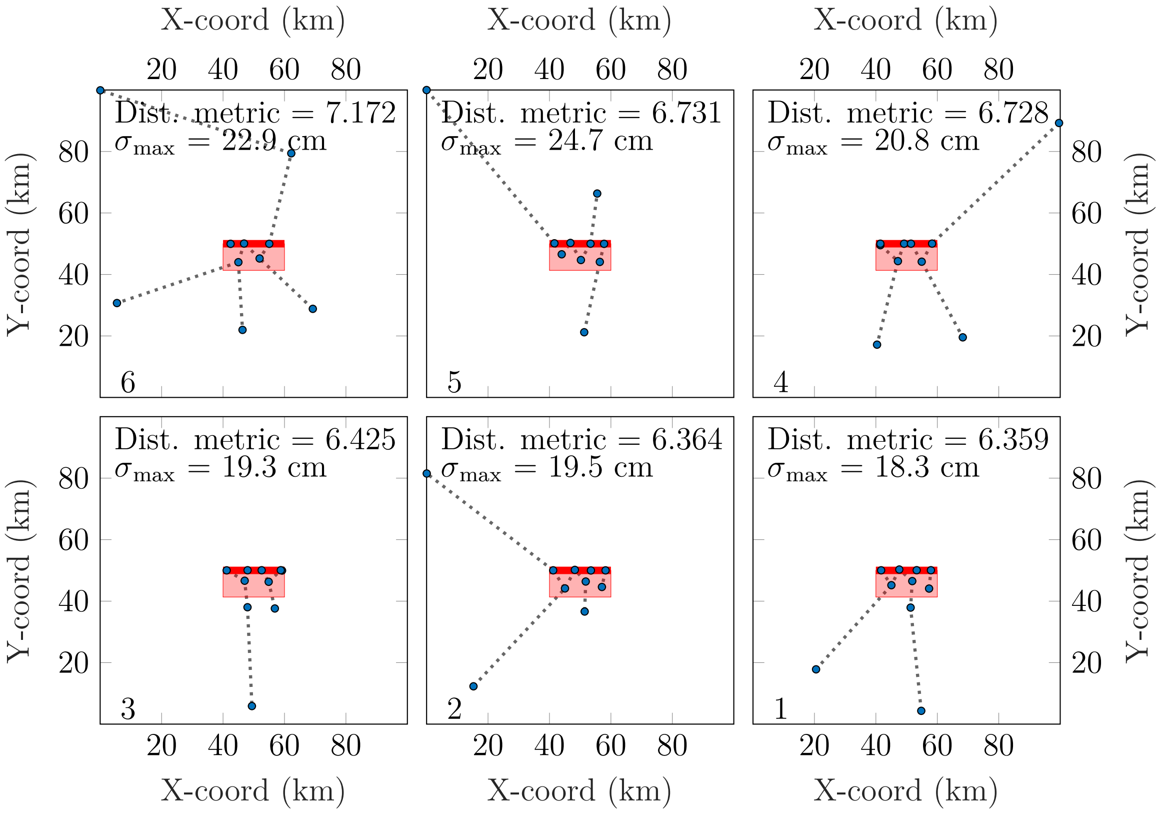

Figure 5.2: Optimal single-frequency networks designed around a east-west striking normal fault. The maximum model uncertainty σ_{max} and distance metric values are provided. Networks are ordered from worst to best by distance metric, where one is the best and six the worst. Note that markers for two stations overlap in networks 3 and 4.

equally valid. This is not true in a real world setting where the ability to deploy GNSS stations may be limited, for example by private property, lack of suitable instrument foot (e.g. steep terrain, bodies of water), or poor access. Therefore, it may be necessary to manually modify optimal GNSS networks, or to purposely deploy sub-optimal networks, so that the maximum number of stations may be deployed.

5.1.3 Network Redundancy

As described in Section 2.2.2, the redundancy values of a network can be calculated using Equation 2.2.2. Each value expresses the level of detectable gross error in the associated observation, ranging from all gross errors being fully detectable ($z_i = 1$) through to no gross errors being detectable ($z_i = 0$).

Returning to Figure 4.16, the redundancy values for both designed networks are calculated using Equation 2.2.2, which are shown in Figure 5.3. These are the mean values for the redundancy of the north, east, and up components for each station. In the case of single-frequency stations, the mean redundancy values are for each pair of stations.

For the dual-frequency GNSS network, average redundancy decrease with proximity to the fault trace, and with increasing separation distance between stations. Stations in closer proximity will measure similar parts of the displacement field, and so errors in one observation may then be identified by comparison to another. As separation distance increases the difference between the observed displacement also increases, and so changes in the observation due to gross error become harder to distinguish. The decreasing redundancy trend towards the fault trace, although the most northern station is an anomaly to this trend, is likely the result of increasing slip patch density towards the top of the fault plane. Higher slip patches density results in more tightly overlapping displacement fields and increasing the difficulty of isolating an error. This exasperates the problem of increasing station separation de-

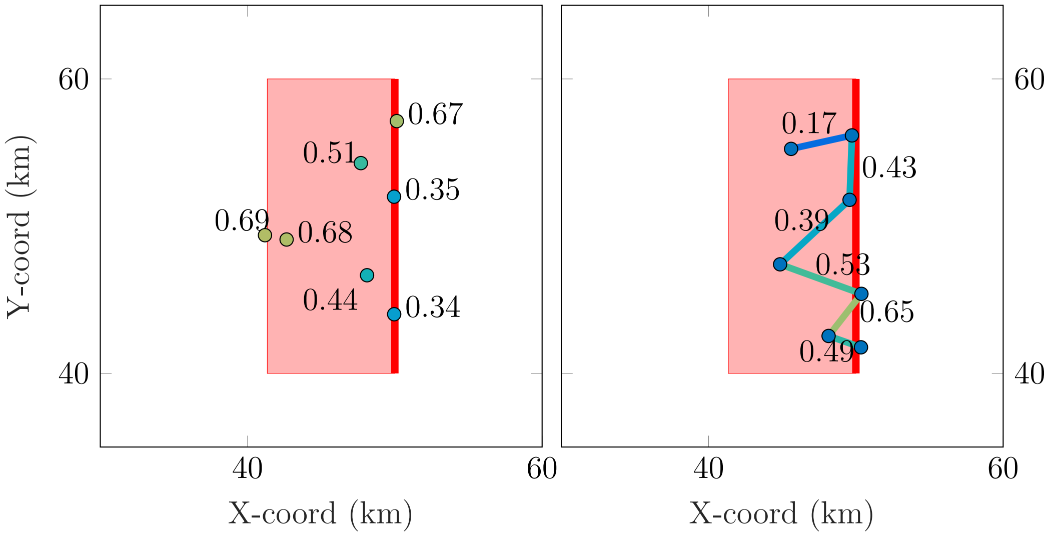

Figure 5.3: A comparison of the redundancy values for the dual- (left) and single-frequency (right) optimal networks shown in Figure 4.16. Values are for each station (circles) for dual-frequency, and each pair of stations (lines) for single-frequency.

scribed above. All values are greater than 0.3, representing reasonable controllability of the network observations (Staudinger 2000; Mahapatra et al. 2015).

The trend of increasing redundancy away from the fault trace is reversed in the single-frequency network. Redundancy values also increase towards the southern end of the fault as a result of higher station density. Values are overall lower than those observed in the dual-frequency network, although all but one pair remain above 0.3. This may be the result of having one less set of observations (north, east, up), and the layout of the network itself. The northern most pair of stations, where $z = 0.17$, are a weakness in the network that would need to be addressed if the network were to be deployed, so that significantly amount of undetectable error are not propagated to the model parameters. This would most easily be done by adding an additional station in the same region.

5.2 Methodological Limitations

The presented methodology is is reliant upon a number of simplifications and assumptions, that limit the real-world applications. These are explored, and potential solutions suggested.

5.2.1 Single Rectangular Fault

Three assumptions are currently made about the model fault: that the geometry is perfectly constrained, that only a single fault is present, and that the relationship between fault slip and surface displacement may be accurately represented by a rectangular elastic half-space (Okada 1985).

The focus of this work has been the design of GNSS networks for the recovery of distributed aseismic slip on a fault, in particular postseismic slip. When deploying a GNSS network to capture postseismic slip following an earthquake, it is a reasonable assumption that other geodetic data sets (e.g. InSAR, seismic) have already been analysed. Therefore, the geometry of the fault may already be constrained to a reasonable degree of accuracy.

As part of the fixed fault geometry, it is assumed that the rake angle is uniform across all slip patches. In reality, the rake may be poorly constrained and vary across the fault plane, which would alter the surface displacement and therefore the model uncertainties. The rake for each slip patch could be included as a model parameter, which would result in an uncertainty on the rake angle for each slip patch. Care must be taken that that the problem does not become ill-posed (more model parameters than observations), by either decreasing the number of slip patches in the discretisation, or by increasing the number of stations in the designed network.

All example presented so far have focused on a single fault plane, where all surface displacement is the result of slip on this one fault. Aseismic slip may occur over multiple adjacent faults, especially in the case of postseismic afterslip where multiple faults have ruptured (e.g. Walters, Gregory et al. (2018)). Overlapping surface displacement fields from multiple slipping faults would be expected to impact the optimal layout of GNSS networks, requiring an increase in the number of stations needed to produce comparable model uncertainties across multiple faults. Attempting to represent a system of slipping faults as a single fault plane would result in sub-optimal networks and model uncertainties that may deviate significantly from their true values. Expansion of the methodology to explore multi-fault setups is considered in Section 5.3.2.

The fault model used to produce $\mathbf{G}$ consists of a number of rectangular elastic half-spaces with uniform stress properties and no tensile component (Okada 1985). This model under-represents the complexity of the Earth and fault slip, resulting in surface deformation that does not exactly reflect reality. This will impact the network design, which is reliant upon the fault model and surface displacement. However, Sathiakumar et al. (2017) demonstrates a method for the optimal design of dual-frequency GNSS stations, using model resolution as a design criteria. While an elastic half-space is used as the primary fault model, a more complex layered earth model is also explored, using velocity and density values taken from the Preliminary Reference Earth Model (Dziewonski et al. 1981). The impact of the more complex fault model on the network design is minimal, with similar results achieved with networks of equal size. The network design is therefore only weakly sensitive to the fault model.

5.2.2 Removal of Long-Wavelength Ionospheric Trends

In Section 3.3, it was shown that a planar trend must be removed from the single-frequency observations before they can utilised. This is to mitigate long-wavelength ionospheric trends that result in high correlations between single-frequency GNSS station pairs aligned with similar orientations. High data correlation will propagate through to high model correlation as a result of the linear inverse.

	North	East	Up
X	0.0977	0.0942	0.3824
Y	0.0902	0.0868	0.3524

Table 5.1: Gradient variances for each component of the linear plane resolved by the network in Figure 5.4. Units are m^2km^{-1}.

The single-frequency network design is performed assuming that the observations have been corrected for the ionospheric delay *a priori*. This is a reasonable assumption when studying aseismic slip on major continental faults, as regional GNSS data will be available. However, if the observations cannot be corrected *a priori*, then the removal of the planar trend can instead be included in the linear inversion. Including the planar trend adds an additional six model parameters, those being the x and y gradients of the planar trend for each components (north, east, and up).

The single-frequency network optimisation shown in Figure 4.15 is repeated, with the planar trend parameters included in both $Q_{m,ref}$ and $Q_{m,des}$. The model uncertainty on each gradient is minimised as part of the network design. The updated optimal network is shown in Figure 5.4.

The updated optimal network contains five stations placed at the extremes of the region, as the network design attempts to constrain the gradients. Model uncertainties have increased across all slip patches, with the biggest changes observed on the deepest patches, due to the relocation of the far-field stations. The variance values for the planar trend are given in Table 5.1, which are also poorly constrained.

The increase in the model uncertainties as a result of the inclusion of the planar trend in the model parameters highlights the importance of correcting the single-frequency GNSS data *a priori*.

5.3 Future Work

A number of expansions to the current methodology, to increase its practicality and accuracy in design optimal GNSS networks, are presented below.

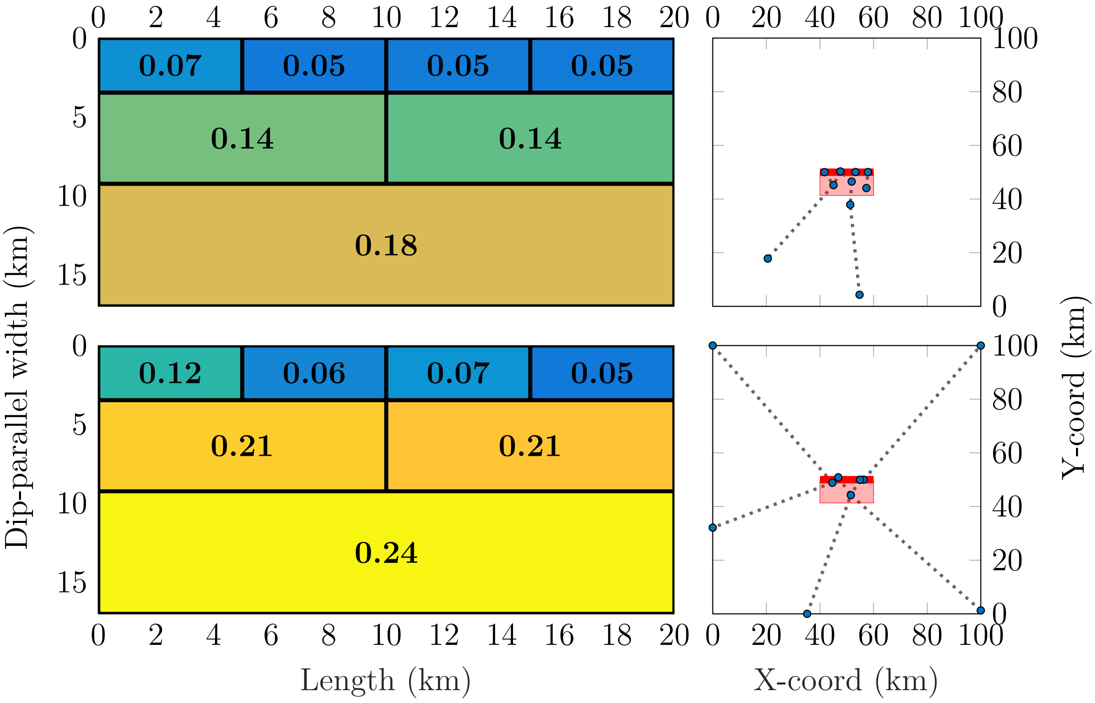

Figure 5.4: A repeat of the network optimisation shown in Figure 4.15 for ten single-frequency GNSS stations around a east-west striking normal fault. The model uncertainties derived from $\mathbf{Q}_{m,ref}$ (Top) and $\mathbf{Q}_{m,des}$ (bottom) are displayed on the discretisation.

5.3.1 Mixed GNSS Networks

The current methodology assumes that all stations within the designed network are of uniform type (i.e. all single- or all dual-frequency stations). This was done so as to simplify the design problem so that a method could be devised within the time constraints of the project.

When studying aseismic slip on a major continental fault, in particular postseismic deformation, it is reasonable to assume that other conventional GNSS stations will be present. The designed network may then be generated to support these existing GNSS stations, which is a third-order design problem. Mixed-frequency networks may also be generated with no existing GNSS stations present. In both cases, the structure of $\mathbf{Q}_{d,des}$ will change to includes both levels of data uncertainty, which will propagate to $\mathbf{Q}_{m,des}$.

Mixed-frequency networks are of special interest for single-frequency GNSS stations. Mitigating the increased data uncertainties associated with single-frequency stations is a key problem in their use for studying aseismic slip. Dual-frequency GNSS stations could be used to support single-frequency stations, providing reference measurements of surface displacement that possess low data uncertainties, and ionospheric delay corrections. Mixed-frequency networks would combine the advantages of both station types, with dual-frequency GNSS stations providing more accurate reference measurements of surface displacement, and single-frequency stations allowing for more observations relative to a dual-frequency only network due to their lower instrument cost. An example is provided by Chen, Kuo et al. (2015), who uses of a mixed network of 17 single- and 10 dual-frequency stations to monitor active faults in Taiwan.

5.3.2 Multi-Fault Systems

The current assumption that aseismic slip occurs only on a single fault plane limits the application of the network design to many real-world fault zones that may consist of multiple slipping faults (e.g. Walters, Gregory et al. (2018)). Expansion of the methodology to include multiple fault planes would be relatively minor, requiring the modification of **G** to include the geometry of each fault plane, so that the surface deformation can be modelled for all faults. The more complex surface dispalcement field would complicate the network design, however, this may result in more interesting and less intuitive network layouts that may better constrain the aseismic slip.

Chapter 6

Conclusion

Low-cost GNSS stations, in particular single-frequency instruments, offer the potential for increased ability to resolve distributed aseismic slip on continental fautls zones, through the deployment of larger and denser GNSS networks. To fully utilise the increased number of stations, and to mitigate teh reduced positional accuracy of single-frequency GNSS stations, a method for the estimation of optimal GNSS network layouts is required.

In this thesis, a method for the automated design of optimal GNSS networks has been presented. Networks are generated using a PSO algorithm and a criterion matrix method, where model uncertainties are minimised relative to their “best possible” values. PSO enables the efficient optimisation of complex non-linear problems, and allows stations to be placed in a continuous region, rather than at discrete points.

To generate the “best possible” model uncertainties, a method for automated fault discretisation is also presented. The use of a MOGA allows a range of optimal solutions to be generated, based upon the trade-off between the number of slip patches and the model uncertainty. The use of optimisation to minimise model uncertainty means that no regularisation is required, which would otherwise bias the results.

For the design of single-frequency GNSS stations, the structure of the ionospheric delay is modeled using variograms, and long-wavelength trends removed. This mitigates the high observations correlations associated with single-frequency GNSS stations and reduces the data uncertainty, allowing the layouts of optimal single-frequency GNSS networks to be estimated.

In Chapter 4, it is shown that optimal low-cost GNSS networks can be estimated that provide low model uncertainties across varying fault geometries and discretisations. The flexability of the fault discretisation means that all magnitudes of aseismic slip can be resolved, given that the magnitude is above the minimum detection level of the observations. Estimated networks show repeated structures and resilience to varying fault geometry. Single-frequency GNSS networks may be generated that provide similar results to dual-frequency GNSS stations, with a reasonable increase in the number of stations deployed.

Further expansions of the methodology should include the estimation of mixed GNSS networks that benefit from the advantages of both station types, and the inclusion of multiple slipping fault plane so that real fault zone may be better represented.

Bibliography

Aki, Keiiti and Paul G Richards (2002). *Quantitative seismology.*

Amiri-Simkooei, AliReza (2004). 'A new method for second order design of geodetic networks: Aiming at high reliability'. In: *Survey review* 37.293, pp. 552–560.

Amiri-Simkooei, AR et al. (2012). 'Basic concepts of optimization and design of geodetic networks'. In: *Journal of Surveying Engineering* 138.4, pp. 172–183.

Angrisano, A et al. (2011). 'Ionospheric models comparison for single-frequency GNSS positioning'. In: *Proceedings of the European navigation conference*, pp. 92–106.

Ansberque, Claire et al. (2016). 'The Longriqu fault zone, eastern Tibetan Plateau: Segmentation and Holocene behavior'. In: *Tectonics* 35.3, pp. 565–585.

Aster, Richard C, Brian Borchers and Clifford H Thurber (2011). *Parameter estimation and inverse problems.* Vol. 90. Academic Press.

Atzori, Simone and Andrea Antonioli (2011). 'Optimal fault resolution in geodetic inversion of coseismic data'. In: *Geophysical Journal International* 185.1, pp. 529–538.

Audet, Charles and John E Dennis Jr (2002). 'Analysis of generalized pattern searches'. In: *SIAM Journal on optimization* 13.3, pp. 889–903.

Baarda, W (1973). *S-transformation and criterion matrices.* Delft, Netherlands: Netherland Geodetic Commission.

Bachmaier, Martin and Matthias Backes (2011). 'Variogram or semivariogram? Variance or semivariance? Allan variance or introducing a new term?' In: *Mathematical Geosciences* 43.6, pp. 735–740.

Bagnardi, Marco, Pablo J González and Andrew Hooper (2016). 'High-resolution digital elevation model from tri-stereo Pleiades-1 satellite imagery for lava flow volume estimates at Fogo Volcano'. In: *Geophysical Research Letters* 43.12, pp. 6267–6275.

Barnhart, WD and RB Lohman (2010). 'Automated fault model discretization for inversions for coseismic slip distributions'. In: *Journal of Geophysical Research: Solid Earth* 115.B10.

Bartel, Beth A et al. (2003). 'Dynamics of active magmatic and hydrothermal systems at Taal Volcano, Philippines, from continuous GPS measurements'. In: *Journal of Geophysical Research: Solid Earth* 108.B10.

Baselga, Sergio (2011). 'Second order design of geodetic networks by the simulated annealing method'. In: *Journal of Surveying Engineering* 137.4, pp. 167–173.

Bekaert, DPS et al. (2016). 'A Network Inversion Filter combining GNSS and InSAR for tectonic slip modeling'. In: *Journal of Geophysical Research: Solid Earth* 121.3, pp. 2069–2086.

Bellone, Tamara et al. (2016). 'Real-time monitoring for fast deformations using GNSS low-cost receivers'. In: *Geomatics, Natural Hazards and Risk* 7.2, pp. 458–470.

Benoit, L et al. (2015). 'Monitoring landslide displacements with the Geocube wireless network of low-cost GPS'. In: *Engineering Geology* 195, pp. 111–121.

Berné, JL and S Baselga (2004). 'First-order design of geodetic networks using the simulated annealing method'. In: *Journal of Geodesy* 78.1-2, pp. 47–54.

Biagi, Ludovico, Florin Ctlin Grec and Marco Negretti (2016). 'Low-cost GNSS receivers for local monitoring: experimental simulation, and analysis of displacements'. In: *Sensors* 16.12, p. 2140.

Blanch, Juan (2002). 'An ionosphere estimation algorithm for WAAS based on kriging'. In: *proceedings of ION GPS*. Vol. 2002, pp. 24–27.

Blum, Christian and Andrea Roli (2003). ‘Metaheuristics in combinatorial optimization: Overview and conceptual comparison’. In: *ACM computing surveys (CSUR)* 35.3, pp. 268–308.

Bonyadi, Mohammad Reza and Zbigniew Michalewicz (2017). *Particle swarm optimization for single objective continuous space problems: a review.*

Boussa´id, Ilhem, Julien Lepagnot and Patrick Siarry (2013). ‘A survey on optimization metaheuristics’. In: *Information Sciences* 237, pp. 82–117.

Broughton, Robert et al. (2010). ‘Determinant and exchange algorithms for observation subset selection’. In: *IEEE Transactions on Image Processing* 19.9, pp. 2437–2443.

Brown, N et al. (2006). ‘Advances in RTK and Post Processed Monitoring with SingleFrequency GPS’. In: *Positioning* 1.10, pp. 1–8.

Cheloni, Daniele et al. (2017). ‘Geodetic model of the 2016 Central Italy earthquake sequence inferred from InSAR and GPS data’. In: *Geophysical Research Letters* 44.13, pp. 6778–6787.

Chen, HY, LC Kuo et al. (2015). ‘Reducing distance dependent bias in low-cost single frequency GPS network to complement dual frequency GPS stations in order to derive detailed surface deformation field’. In: *Survey Review* 47.340, pp. 7–17.

Chen, HY, Chris Rizos and Shaowei Han (2001). ‘From simulation to implementation: low-cost densification of permanent GPS networks in support of geodetic applications’. In: *Journal of Geodesy* 75.9-10, pp. 515–526.

Cina, Alberto and Marco Piras (2015). ‘Performance of low-cost GNSS receiver for landslides monitoring: test and results’. In: *Geomatics, Natural Hazards and Risk* 6.5-7, pp. 497–514.

Custódio, Susana, Morgan T Page and Ralph J Archuleta (2009). ‘Constraining earthquake source inversions with GPS data: 2. A two-step approach to combine seismic and geodetic data sets’. In: *Journal of Geophysical Research: Solid Earth* 114.B1.

D’Agostino, N (2014). ‘Complete seismic release of tectonic strain and earthquake recurrence in the Apennines (Italy)’. In: *Geophysical Research Letters* 41.4, pp. 1155–1162.

D’Agostino, Nicola et al. (2011). ‘Evidence for localized active extension in the central Apennines (Italy) from global positioning system observations’. In: *Geology* 39.4, pp. 291–294.

Deb, Kalyanmoy et al. (2002). ‘A fast and elitist multiobjective genetic algorithm: NSGA-II’. In: *IEEE transactions on evolutionary computation* 6.2, pp. 182–197.

DoD, US (2001). ‘Global positioning system standard positioning service performance standard’. In: *Assistant secretary of defense for command, control, communications, and intelligence.*

Doma, MI (2013). ‘Particle swarm optimization in comparison with classical optimization for GPS network design’. In: *Journal of Geodetic Science* 3.4, pp. 250–257.

Dorigo, Marco and Mauro Birattari (2011). ‘Ant colony optimization’. In: *Encyclopedia of machine learning.* Springer, pp. 36–39.

Dorigo, Marco and Luca Maria Gambardella (1997). ‘Ant colony system: a cooperative learning approach to the traveling salesman problem’. In: *IEEE Transactions on evolutionary computation* 1.1, pp. 53–66.

Duputel, Zacharie et al. (2014). ‘Accounting for prediction uncertainty when inferring subsurface fault slip’. In: *Geophysical Journal International* 197.1, pp. 464–482.

Dziewonski, Adam M and Don L Anderson (1981). ‘Preliminary reference Earth model’. In: *Physics of the earth and planetary interiors* 25.4, pp. 297–356.

Eberhart, Russell C and Yuhui Shi (1998). ‘Comparison between genetic algorithms and particle swarm optimization’. In: *International conference on evolutionary programming.* Springer, pp. 611–616.

Eberhart, Russell and James Kennedy (1995). ‘A new optimizer using particle swarm theory’. In: *Micro Machine and Human Science, 1995. MHS’95., Proceedings of the Sixth International Symposium on*. IEEE, pp. 39–43.

Elliott, JR, RJ Walters and TJ Wright (2016). ‘The role of space-based observation in understanding and responding to active tectonics and earthquakes’. In: *Nature communications* 7, p. 13844.

Eyo, EE et al. (2014). ‘Application of Low-Cost GPS Tools and Techniques for Landslide Monitoring: A Review’. In: *Jurnal Teknologi (Sciences & Engineering)* 71.4, pp. 71–78.

Förstner, Wolfgang and Boudewijn Moonen (2003). *A metric for covariance matrices*. Springer, pp. 299–309.

Funning, Gareth J et al. (2005). ‘Surface displacements and source parameters of the 2003 Bam (Iran) earthquake from Envisat advanced synthetic aperture radar imagery’. In: *Journal of Geophysical Research: Solid Earth* 110.B9.

Geman, Stuart and Donald Geman (1984). ‘Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images’. In: *IEEE Transactions on pattern analysis and machine intelligence* 6, pp. 721–741.

Gembicki, F and Y Haimes (1975). ‘Approach to performance and sensitivity multiobjective optimization: The goal attainment method’. In: *IEEE Transactions on Automatic control* 20.6, pp. 769–771.

Goldberg, David E and John H Holland (1988). ‘Genetic algorithms and machine learning’. In: *Machine learning* 3.2, pp. 95–99.

Grafarend, E and B Schaffrin (1974). ‘Unbiased free net adjustment’. In: *Survey Review* 22.171, pp. 200–218.

Grafarend, Erik W and Fernando Sansò (2012). *Optimization and design of geodetic networks*. Springer Science & Business Media.

Grafarend, EW (1985). ‘Criterion matrices for deforming networks’. In: *Optimization and design of geodetic networks*. Springer, pp. 363–428.

Groves, Paul D (2013). *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Artech house.

Günther, Jessica et al. (2008). 'Developments towards a low-cost GNSS based sensor network for the monitoring of landslides'. In: *Proceedings of the 13th FIG International Symposium on Deformation Measurements and Analysis, Lisbon, Portugal*, pp. 12–15.

Hamling, Ian J et al. (2017). 'Complex multifault rupture during the 2016 M w 7.8 Kaikura earthquake, New Zealand'. In: *Science* 356.6334, eaam7194.

Hardt, Matthias and Frank Scherbaum (1994). 'The design of optimum networks for aftershock recordings'. In: *Geophysical Journal International* 117.3, pp. 716–726.

Hayes, Gavin P et al. (2014). 'Continuing megathrust earthquake potential in Chile after the 2014 Iquique earthquake'. In: *Nature* 512.7514, p. 295.

Hill, Emma M et al. (2012). 'The 2010 Mw 7.8 Mentawai earthquake: Very shallow source of a rare tsunami earthquake determined from tsunami field survey and near-field GPS data'. In: *Journal of Geophysical Research: Solid Earth* 117.B6.

Holland, John H (1992). 'Genetic algorithms'. In: *Scientific american* 267.1, pp. 66–73.

Hoque, MM and N Jakowski (2015). 'An alternative ionospheric correction model for global navigation satellite systems'. In: *Journal of Geodesy* 89.4, pp. 391–406.

Huang, Ling et al. (2017). 'Kriging with unknown variance components for regional ionospheric reconstruction'. In: *Sensors* 17.3, p. 468.

Hussain, Ekbal et al. (2018). 'Constant strain accumulation rate between major earthquakes on the North Anatolian Fault'. In: *Nature communications* 9.1, p. 1392.

Janssen, Volker et al. (2002). 'Low-cost GPS-based volcano deformation monitoring at Mt. Papandayan, Indonesia'. In: *Journal of Volcanology and Geothermal Research* 115.1-2, pp. 139–151.

Kaplan, Elliott D and Christopher Hegarty (2017). *Understanding GPS/GNSS: Principles and Applications*. Artech House.

Karakhanyan, A et al. (2013). ‘GPS constraints on continental deformation in the Armenian region and Lesser Caucasus’. In: *Tectonophysics* 592, pp. 39–45.

Kirkpatrick, Scott, C Daniel Gelatt and Mario P Vecchi (1983). ‘Optimization by simulated annealing’. In: *science* 220.4598, pp. 671–680.

Klein, Emilie et al. (2017). ‘Aseismic slip and seismogenic coupling in the Marmara Sea: What can we learn from onland geodesy?’ In: *Geophysical Research Letters* 44.7, pp. 3100–3108.

Kolda, Tamara G, Robert Michael Lewis and Virginia Torczon (2003). ‘Optimization by direct search: New perspectives on some classical and modern methods’. In: *SIAM review* 45.3, pp. 385–482.

Komac, Marko et al. (2015). ‘Coupling of GPS/GNSS and radar interferometric data for a 3D surface displacement monitoring of landslides’. In: *Landslides* 12.2, pp. 241–257.

Konak, Abdullah, David W Coit and Alice E Smith (2006). ‘Multi-objective optimization using genetic algorithms: A tutorial’. In: *Reliability Engineering & System Safety* 91.9, pp. 992–1007.

Koza, John R (1994). ‘Genetic programming as a means for programming computers by natural selection’. In: *Statistics and computing* 4.2, pp. 87–112.

Kreemer, Corné, Geoffrey Blewitt and Elliot C Klein (2014). ‘A geodetic plate motion and Global Strain Rate Model’. In: *Geochemistry, Geophysics, Geosystems* 15.10, pp. 3849–3889.

Kruskal, Joseph B (1956). ‘On the shortest spanning subtree of a graph and the traveling salesman problem’. In: *Proceedings of the American Mathematical society* 7.1, pp. 48–50.

Lee, Seung-Woo et al. (2015). ‘Real-time volcano monitoring using GNSS single-frequency receivers’. In: *Journal of Geophysical Research: Solid Earth* 120.12, pp. 8551–8569.

Li, Zhenhong, John R Elliott et al. (2011). 'The 2010 MW 6.8 Yushu (Qinghai, China) earthquake: Constraints provided by InSAR and body wave seismology'. In: *Journal of Geophysical Research: Solid Earth* 116.B10.

Li, Zhenhong, Wanpeng Feng et al. (2008). 'The 1998 Mw 5.7 Zhangbei-Shangyi (China) earthquake revisited: A buried thrust fault revealed with interferometric synthetic aperture radar'. In: *Geochemistry, Geophysics, Geosystems* 9.4.

Lohman, Rowena B and Mark Simons (2005). 'Some thoughts on the use of InSAR data to constrain models of surface deformation: Noise structure and data downsampling'. In: *Geochemistry, Geophysics, Geosystems* 6.1.

Mahapatra, Pooja S, Sami Samiei-Esfahany and Ramon F Hanssen (2015). 'Geodetic network design for InSAR'. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.7, pp. 3669–3680.

Marler, R Timothy and Jasbir S Arora (2004). 'Survey of multi-objective optimization methods for engineering'. In: *Structural and multidisciplinary optimization* 26.6, pp. 369–395.

Mazzotti, Stéphane et al. (2005). 'GPS crustal strain, postglacial rebound, and seismic hazard in eastern North America: The Saint Lawrence valley example'. In: *Journal of Geophysical Research: Solid Earth* 110.B11.

Metropolis, Nicholas et al. (1953). 'Equation of state calculations by fast computing machines'. In: *The journal of chemical physics* 21.6, pp. 1087–1092.

Metzger, Sabrina et al. (2017). 'The 2015 Mw7. 2 Sarez Strike-Slip Earthquake in the Pamir Interior: Response to the Underthrusting of India's Western Promontory'. In: *Tectonics* 36.11, pp. 2407–2421.

Mezura-Montes, Efrén and Carlos A Coello Coello (2011). 'Constraint-handling in nature-inspired numerical optimization: past, present and future'. In: *Swarm and Evolutionary Computation* 1.4, pp. 173–194.

Mitchell, Melanie (1998). *An introduction to genetic algorithms.* MIT press.

Moraglio, Alberto, Julian Togelius and Sara Silva (2013). ‘Geometric differential evolution for combinatorial and programs spaces’. In: *Evolutionary computation* 21.4, pp. 591–624.

Moret, Bernard ME and Henry D Shapiro (1991). ‘An empirical analysis of algorithms for constructing a minimum spanning tree’. In: *Workshop on Algorithms and Data Structures*. Springer, pp. 400–411.

Neumann, Frank and Ingo Wegener (2007). ‘Randomized local search, evolutionary algorithms, and the minimum spanning tree problem’. In: *Theoretical Computer Science* 378.1, pp. 32–40.

Okada, Yoshimitsu (1985). ‘Surface deformation due to shear and tensile faults in a half-space’. In: *Bulletin of the seismological society of America* 75.4, pp. 1135–1154.

Page, Morgan T et al. (2009). ‘Constraining earthquake source inversions with GPS data: 1. Resolution-based removal of artifacts’. In: *Journal of Geophysical Research: Solid Earth* 114.B1.

Pedersen, Magnus Erik Hvass (2010). *Good parameters for particle swarm optimization*. Tech. rep.

Peng, Zhigang and Joan Gomberg (2010). ‘An integrated perspective of the continuum between earthquakes and slow-slip phenomena’. In: *Nature Geoscience* 3.9, p. 599.

Pizzi, A et al. (2017). ‘Fault segmentation as constraint to the occurrence of the main shocks of the 2016 Central Italy seismic sequence’. In: *Tectonics* 36.11, pp. 2370–2387.

Pritchard, ME et al. (2002). ‘Co-seismic slip from the 1995 July 30 M w= 8.1 Antofagasta, Chile, earthquake as constrained by InSAR and GPS observations’. In: *Geophysical Journal International* 150.2, pp. 362–376.

Rademakers, Erwin et al. (2016). ‘Obtaining real-time sub-meter accuracy using a low cost GNSS device’. In: *Navigation Conference (ENC), 2016 European*. IEEE, pp. 1–8.

Reeves, Stanley J and Zhao Zhe (1999). ‘Sequential algorithms for observation selection’. In: *IEEE Transactions on Signal Processing* 47.1, pp. 123–132.

Reilinger, Robert et al. (2006). ‘GPS constraints on continental deformation in the Africa-Arabia-Eurasia continental collision zone and implications for the dynamics of plate interactions’. In: *Journal of Geophysical Research: Solid Earth* 111.B5.

Roustaei, M et al. (2010). ‘The 2006 March 25 Fin earthquakes (Iran)insights into the vertical extents of faulting in the Zagros Simply Folded Belt’. In: *Geophysical Journal International* 181.3, pp. 1275–1291.

Sagiya, Takeshi (2004). ‘A decade of GEONET: 1994-2003’. In: *Earth, planets and space* 56.8, pp. xxix–xli.

Sathiakumar, Sharadha, Sylvain Denis Barbot and Piyush Agram (2017). ‘Extending Resolution of Fault Slip With Geodetic Networks Through Optimal Network Design’. In: *Journal of Geophysical Research: Solid Earth* 122.12.

Schmitt, Günter (1982). ‘Optimization of geodetic networks’. In: *Reviews of Geophysics* 20.4, pp. 877–884.

Scognamiglio, Laura et al. (2018). ‘Complex Fault Geometry and Rupture Dynamics of the MW 6.5, 30 October 2016, Central Italy Earthquake’. In: *Journal of Geophysical Research: Solid Earth* 123.4, pp. 2943–2964.

Seemkooei, Alireza Amiri (2001). ‘Strategy for designing geodetic network with high reliability and geometrical strength’. In: *Journal of Surveying Engineering* 127.3, pp. 104–117.

Segall, Paul (2010). *Earthquake and volcano deformation*. Princeton University Press.

Sentinel-1 Team (2013). *ESA Sentinel 1 handbook*. European Space Agency, pp. 1–80. DOI: `10.1017/CBO9781107415324.004`.

Serafini, Paolo (1994). ‘Simulated annealing for multi objective optimization problems’. In: *Multiple criteria decision making*. Springer, pp. 283–292.

Simonetto, Elisabeth et al. (2014). ‘Combination of INSAR and GNSS measurements for ground displacement monitoring’. In: *Procedia Technology* 16, pp. 192–198.

Siwale, Isaac (2013). *Practical multi-objective programming.* Tech. rep. Tech. Rep. RD-14-2013.

Srensen, Kenneth (2015). ‘Metaheuristicsthe metaphor exposed’. In: *International Transactions in Operational Research* 22.1, pp. 3–18.

Staudinger, Martin (2000). *A cost oriented approach to geodetic network optimisation.* Institute for Geoinformation, Vienna University of Technology.

Swan, Jerry et al. (2015). ‘A research agenda for metaheuristic standardization’. In: *Proceedings of the XI Metaheuristics International Conference.*

Takasu, Tomoji and Akio Yasuda (2008). ‘Evaluation of RTK-GPS performance with low-cost single-frequency GPS receivers’. In: *Proceedings of international symposium on GPS/GNSS*, pp. 852–861.

Teunissen, Peter (1985). ‘Zero order design: generalized inverses, adjustment, the datum problem and S-transformations’. In: *Optimization and design of geodetic networks.* Springer, pp. 11–55.

Teunissen, PJG (1998). ‘Minimal detectable biases of GPS data’. In: *Journal of Geodesy* 72.4, pp. 236–244.

Torres, Ramon et al. (2012). ‘GMES Sentinel-1 mission’. In: *Remote Sensing of Environment* 120, pp. 9–24.

Tu, R et al. (2013). ‘Cost-effective monitoring of ground motion related to earthquakes, landslides, or volcanic activity by joint use of a single-frequency GPS and a MEMS accelerometer’. In: *Geophysical Research Letters* 40.15, pp. 3825–3829.

Walker, Richard and James Jackson (2004). ‘Active tectonics and late Cenozoic strain distribution in central and eastern Iran’. In: *Tectonics* 23.5.

Walters, RJ, JR Elliott et al. (2013). ‘Rapid strain accumulation on the Ashkabad fault (Turkmenistan) from atmosphere-corrected InSAR’. In: *Journal of Geophysical Research: Solid Earth* 118.7, pp. 3674–3690.

Walters, RJ, PC England and GA Houseman (2017). ‘Constraints from GPS measurements on the dynamics of the zone of convergence between Arabia and Eurasia’. In: *Journal of Geophysical Research: Solid Earth* 122.2, pp. 1470–1495.

Walters, RJ, Laura C Gregory et al. (2018). ‘Dual control of fault intersections on stop-start rupture in the 2016 Central Italy seismic sequence’. In: *Earth and Planetary Science Letters* 500, pp. 1–14.

Wang, Ningbo et al. (2016). ‘Improvement of Klobuchar model for GNSS single-frequency ionospheric delay corrections’. In: *Advances in Space Research* 57.7, pp. 1555–1569.

Wilkinson, Maxwell W et al. (2017). ‘Near-field fault slip of the 2016 Vettore M w 6.6 earthquake (Central Italy) measured using low-cost GNSS’. In: *Scientific Reports* 7.1, p. 4612.

Wright, Tim J, Zhong Lu and Chuck Wicks (2003). ‘Source model for the Mw 6.7, 23 October 2002, Nenana Mountain Earthquake (Alaska) from InSAR’. In: *Geophysical Research Letters* 30.18.

Xu, Peiliang and Erik Grafarend (1995). ‘A multi-objective second-order optimal design for deforming networks’. In: *Geophysical Journal International* 120.3, pp. 577–589.

Zheng, Yanli, Rui Zhang and Shengfeng Gu (2014). ‘A new PPP algorithm for deformation monitoring with single-frequency receiver’. In: *Journal of earth system science* 123.8, pp. 1919–1926.

Zhou, Yu, John R Elliott et al. (2015). ‘The 2013 Balochistan earthquake: An extraordinary or completely ordinary event?’ In: *Geophysical Research Letters* 42.15, pp. 6236–6243.

Zhou, Yu, Barry Parsons et al. (2015). ‘Assessing the ability of Pleiades stereo imagery to determine height changes in earthquakes: A case study for the El

Mayor-Cucapah epicentral area'. In: *Journal of Geophysical Research: Solid Earth* 120.12, pp. 8793–8808.