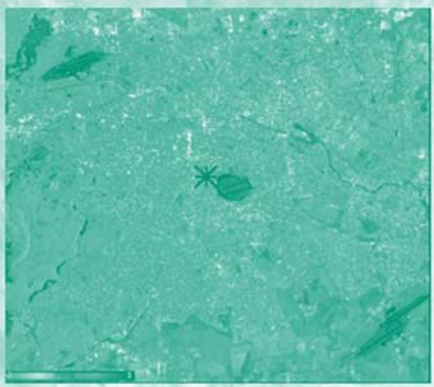
REMOTE SENSING AND DIGITAL IMAGE PROCESSING

Radar Interferometry

Persistent Scatterer Technique

by
Bert M. Kampes



Springer

CD-ROM INCLUDED

Radar Interferometry

Remote Sensing and Digital Image Processing

VOLUME 12

The titles published in this series are listed at the end of this volume

RADAR INTERFEROMETRY

Persistent Scatterer Technique

by

BERT M. KAMPES

German Aerospace Center (DLR), Germany

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4576-X (HB)
ISBN-13 978-1-4020-4576-9 (HB)
ISBN-10 1-4020-4723-1 (e-book)
ISBN-13 978-1-4020-4723-7 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Cover image: Estimated linear displacement rates for the Berlin test site for different thresholds on the a posteriori variance factor. Figure 6.6(c) (also see p. 96)

Printed on acid-free paper

To Jill and Madeline

Contents

Preface

Soon after the first attempts at Delft University of Technology to apply the radar interferometric technique for the monitoring of subsidence due to gas extraction in the province of Groningen, the Netherlands, it was recognized by Usai and Hanssen (1997) that man-made features remained coherent in radar interferograms over long time spans, while their surrounding was completely decorrelated. This particular area in the northern part of the Netherlands is well-known for its subsidence. Due to the slow subsidence rate—the maximum is approximately 1 cm/y—long temporal baselines needed to be used. Even though only interferograms with short perpendicular baselines were generated, temporal decorrelation severely limited the analysis, see (Usai, 1997, 2000; Usai and Klees, 1999). The Groningen data set was also used by Hanssen (1998), who analyzed artifacts of atmospheric origin in coherent interferograms with short temporal baselines. Aside from temporal and geometrical decorrelation, atmospheric signal is the main problem for the interpretation of interferometric signal of current day spaceborne sensors on board, e.g., ERS, ENVISAT and RADARSAT (Hanssen, 2001).

The Permanent Scatterers (PS) Technique was developed shortly after, see (Ferretti et al., 2000a, 2001). It aims to bypass the problem of geometrical and temporal decorrelation by considering time-coherent pixels. Furthermore, by using a large amount of data, atmospheric signal is estimated and corrected for. The PS technique offers a convenient processing framework that enables the use of all acquired images, irrespective of baseline, and a parameter estimation strategy for interferograms with low spatial coherence. The advantages of this method can be measured from the increasing attention it has received at major conferences For example, in the proceedings of the IGARSS conferences of 1999 to 2003 there are respectively 1, 5, 4, 17 and 26 direct references to the term Permanent Scatterer. The "Terrafirma" initiative further underlines the high potential of this technique. This project aims to provide a Pan-European ground motion hazard information service, to be distributed throughout Europe via the national geological surveys. All large towns in Europe are to be studied with the PS technique. In total, 189 towns

in total are identified, equalling 27% of the total population. In the longer term, areas will be included that suffer risks from ground motions caused by, for example, landslides or mining, see (Terrafirma, 2005).

Additionally, once the PS technique demonstrated that using a large number of images is a way to reduce atmospheric artifacts and to obtain highly precise estimates despite decorrelation, this sparked the development of a number of related techniques, e.g., Coherent Target Monitoring (Van der Kooij, 2003; Van der Kooij and Lambert, 2002), Interferometric Point Target Analysis (Wegmuller, 2003; Werner et al., 2003), Stable Point Network analysis (Arnaud et al., 2003), Small Baseline Subset Approach (Berardino et al., 2003, 2002; Lanari et al., 2003; Mora et al., 2002), and Corner Reflector Interferometry and Compact Active Transponders Interferometry (Nigel Press Associates, 2004). These techniques partly seek to improve the PS technique using a modified approach (some even assume distributed scattering of multi-looked pixels, although still use concepts similar to the PS technique), but also partly try to avoid disputes over the patent of the PS technique. The term *Persistent Scatterer Interferometry* (PSI) is now used to group techniques that analyze the phase time series of individual scatterers.

This book revisits the original PS technique and presents a new PSI algorithm, the STUN algorithm, which is developed to provide a robust and reliable estimation of displacement parameters and their precision.

Audience

This book is intended for scientists and students who want to understand and work with Persistent Scatterer Interferometry. Particularly of interest for this group of readers are the derivation of the functional and stochastic model, the description of the estimation using integer least-squares and variance components, and the alternative hypothesis testing procedure, see Chapter 2, 3, and 4, respectively. The software toolbox on the CDROM explain these key concepts using practical demonstrations, see also Appendix E. The modular programs can be easily adapted and be further developed by the interested reader for specific problems.

Secondly, this book is intended to provide insight in the problems and pitfalls of Persistent Scatterer Interferometry for users of PSI products and of commercially available PSI processing software, and to enhance their understanding of this technique. This group of readers includes geo-information professionals and high level decision makers who do not perform PSI processing themselves. The description of the reference PS technique and potential improvements upon it, see Chapter 2, and Chapter 6 on real data processing may prove to be most useful for this group.

The reader is assumed to be familiar with general radar concepts and conventional radar interferometric processing, as for example described in (Bamler and Hartl, 1998; Hanssen, 2001; Klees and Massonnet, 1999; Rosen

et al., 2000). The PS technique is regarded as an extension of the conventional differential interferometric technique. Background knowledge of the PS technique is not required to understand this work. A geodetic background is helpful, but not necessary. However, geodetic concepts are used, particularly concerning the integer least-squares estimator, variance component estimation, and alternative hypothesis testing. These issues are explained in detail in the chapters and appendices to make this work self-contained.

The technique described in this work can be applied to data obtained by current-day radar sensors and data of future systems. However, most emphasis is on ERS–1 and ERS–2, because data of these sensors are available for the test areas over extended time intervals.

Acknowledgments

This study was performed at the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, and at Delft University of Technology, Delft, The Netherlands. Many people helped me in various ways during this time, whom I would like to acknowledge here. I would like to thank Roland Klees, professor of Physical and Space Geodesy at the faculty of Aerospace Engineering at Delft University of Technology, for supervising my doctoral research. I am also most indebted to Ramon Hanssen, associate professor at Delft University, for his guidance through the early years of chaos and confusion, and for the many fun times we had discussing radar interferometry and the philosophy of life. For the same reasons I am thankful to Richard Bamler, director of the Remote Sensing Technology Institute at the DLR, and Michael Eineder, leader of the Image Science and SAR processing group.

The many people I met during these years have been a great motivation to me. I would like to mention Richard Lord, Andy Hooper, and Adele Fusco, visiting scientists at the DLR from the universities of Cape Town, Stanford, and Sannio, respectively. From Delft University of Technology I would like to thank Peter Teunissen and Peter Joosten for providing source code and useful discussions, mainly concerning the integer least-squares estimation. Also the people of the Delft radar group helped me a lot. Many thanks go to Gini Ketelaar, Petar Marinkovic, Yue Huanyin (Paul), and especially Freek van Leijen, who carefully checked parts of this work. This work would not have been possible without the inventors of the PS technique. I would like to thank Alessandro Ferretti, Fabio Rocca, and Marco Bianchi, for their encouragement and for providing me with reference processing results for the Las Vegas test site. I am also grateful to my colleagues at the DLR who are always prepared to listen to me. Particularly I learned a lot from Nico Adam, my office mate, whom I thank for patiently sharing his great understanding and skills in the fields of radar interferometry and software engineering. The people who used and evaluated the software during the development are also very much acknowledged for their feedback, specifically Michaela Kircher, Jirathana

Worawattanamateekul, Diana Walter, Franz Meyer, and Jörn Hoffmann. The latter I also thank for useful discussions on the Las Vegas test site and for providing me with the Lorenzi test site extensometer data. Finally, I thank Marie Lachaise for porting parts of the code from IDL to C.

This work benefited from the careful reading by Richard Bamler, Barry Parsons, Fabio Rocca, Peter Hoogeboom, Peter Teunissen, Ramon Hanssen, and Roland Klees, members of my defense committee. Thank you for your valuable comments.

Freek van der Meer, series editor, and Petra van Steenbergen at Springer are gratefully acknowledged for making publication of this work possible in the series on Remote Sensing and Digital Image Processing.

ESA is acknowledged for providing data of the ERS and ENVISAT sensors, and for the opportunity to work in the framework of the studies "Development of algorithms for the exploitation of ERS ENVISAT using the stable points network", "Retrieval of new Bio- and Geophysical variables", "Persistent Scatterer interferometry codes cross-comparison and certification for long term differential interferometry" (PSIC4), the "Terrafirma" project, and the category-one proposal "Atmospheric correction of interferometric data" (ACID).

But above all, I am grateful for the persisting support of my wife Jill.

Bert Kampes
Munich, March 2006

Summary

Persistent Scatterer Interferometry is the latest development in radar interferometric processing, which offers a practical way to reduce the main errors in conventional processing methods; temporal and geometrical decorrelation, and atmospheric artifacts. This is achieved by the analysis of the interferometric phase of individual long time-coherent scatterers in a stack of tens of differential interferograms with one master image. In this study the original PS technique is revisited and geodetic techniques are applied to improve the quality of estimated parameters that describe displacement. For this reason the STUN (Spatio-Temporal Unwrapping Network) algorithm is developed. The first step in this algorithm is to establish a reference network of coherent points. The points are initially selected based on their amplitude time series, which is expected to be related to the phase dispersion. A large number of estimations between points of the network are performed, followed by a least-squares adjustment to obtain the displacement and topography at the points. The estimations are performed between nearby points (distances less than ~2 km) in order to limit atmospheric signal, which could prevent successful estimation using the wrapped data. An alternative hypothesis testing strategy is carried out to identify incorrectly estimated parameters and incoherent points. The parameters are estimated with the integer least-squares estimator using the wrapped data. This estimator has the highest probability of finding the correct integer ambiguities for data with a multivariate normal distribution. A variance component model is developed to describe the dispersion of the double-differenced phase observations used in the estimation. This new model accounts for random noise and atmospheric signal at the acquisition times. The variance factors of the variance component model are estimated using the least-squares residuals of an initial estimation. The displacement is modeled using a linear combination of base functions. This generic approach allows for the estimation of non-linear displacements using wrapped data. Second—once the parameters at the points of the reference network are computed—more selected points are estimated with respect to the reference network. Based on the estimated a posteriori variance factor, a set of reliable points is selected

and a Minimal Cost Flow sparse grid phase unwrapping algorithm is used to obtain the unwrapped phase at these points. The final estimation is performed using the unwrapped data. The precision of the estimated parameters is described by the propagated variance-covariance matrix with respect to a chosen reference point.

The STUN algorithm is successfully applied to two urban test areas. Several tests are performed to assess the sensitivity of the algorithm to various parameters such as the number of available interferograms, the distance between points in the reference network, etc. The first test site, Berlin, was not expected to undergo significant displacements. It was selected to validate the developed algorithm and software. However, an uplift area is identified to the west of Berlin, with a maximum displacement rate of ~4 mm/y. Most likely, this uplift is related to underground gas storage at that location. Data of two adjacent tracks are used in a cross-comparison of the estimated displacement. Contrary, the second test site, Las Vegas, undergoes significant displacements. A combined linear and sinusoid displacement model is used to model the displacements. The maximum estimated subsidence rate is ~20 mm/y and the maximum amplitude of the seasonal component is ~20 mm. The results compare well with estimates by the reference PS technique. Finally, combined use of ERS and ENVISAT data is demonstrated.

1

Introduction

In the early 1990s, spaceborne radar interferometry (InSAR) was recognized as a powerful technique to measure the earth's topography and surface deformation. An era of continuous imaging started with the launch of the ERS–1 satellite in 1991, prolonged by ERS–2 (a replica of ERS–1), launched in 1995, and ENVISAT, launched in 2002. Radar interferometry has matured to an almost operational technique, among other reasons due to the large amount of accessible data over this time span. Aside from the aforementioned satellites that were launched by ESA, also the Canadian RADARSAT and the Japanese JERS satellites contributed to the success of radar interferometry. However, the data acquisition, archiving, and pricing policy of ESA, as well as precise orbit control and determination, are important advantages of the ESA satellites for the application of radar interferometry. InSAR is successfully applied to measure surface displacements caused by, for example, earthquakes (see, e.g., Hanssen et al., 2000; Massonnet et al., 1993; Zebker et al., 1994), volcanoes (e.g., Amelung et al., 2000; Massonnet et al., 1995), subsidence and uplift phenomena (e.g., Amelung et al., 1999; Hanssen et al., 1998; Shimoni et al., 2002) and glaciers (e.g., Joughin, 1995; Meyer, 2004). Without a doubt, radar interferometry is the only existing technique capable of observing these phenomena with a high resolution on a wide spatial scale. In addition, such analyses can be performed for practically any area using data starting in 1992, because these data have been acquired in a background mission and archived. The perfect example of radar interferometry being used to measure topography is given by the Shuttle Radar Topography Mission, where a near-global digital elevation model with unprecedented spatial resolution and accuracy was generated (Rabus et al., 2003; Rosen et al., 2000; Suchandt et al., 2001). Radar interferometry is also used to study the water vapor distribution in the atmosphere with high accuracy and unmatched spatial resolution, possibly leading to the application of interferometric radar meteorology (Hanssen, 2001; Hanssen et al., 1999).

The observed phase at a pixel in an interferogram is related to the difference in measured distances of a terrain element to the radar sensors

at the times of the acquisitions. This difference in turn is related to displacement, topography, and atmospheric delay. These phase components are superimposed in the interferogram. In most studies where radar interferometry is successfully applied, it is assumed that the component of interest is dominantly present, while the other components can either be estimated independently, or practically neglected. Moreover, in most studies the coherence of the interferogram (i.e., the precision of the phase observations) is relatively high, aiding interpretation. However, due to temporal and geometrical decorrelation (decrease of coherence due to incoherent movements of individual scattering elements and due to different viewing angles, see, e.g., Gatelli et al., 1994; Zebker and Villasenor, 1992) this is not the case in general, implying loss of correlation between the observed phases. The phase unwrapping process—necessary for unambiguous interpretation—becomes increasingly difficult when the coherence decreases.

In the late 1990s, the multi-image Permanent Scatterer (PS) technique was introduced (Ferretti et al., 1999c) that deals with these problems in an innovative way. This technique offers a systematic processing strategy, capable of utilizing all archived data of a certain area, by creating a stack of differential interferograms that have a common master image. Instead of analyzing the phase in the spatial domain, the phase of isolated coherent points is analyzed as a function of time and space.

1.1 Objectives

The PS technique, as described by Ferretti et al. (2000a, 2001), is the basis for this study. The principal estimation strategy is not questioned, i.e., a single master stack of complex differential interferograms is used, and the points are estimated using a preliminary and a final estimation step. Within this framework the central research question is formulated as:

"How can geodetic methodology aid displacement parameter estimation using Persistent Scatterer Interferometry?"

Improvements that are addressed in this study are related to the:

Functional model. In the reference PS technique a linear rate is used to model the displacement during the estimation using wrapped data. This has the disadvantage that unmodeled (non-linear) displacement may "leak" to the estimated atmospheric signal. Moreover, if the actual displacement deviates significantly from this model estimation may become impossible. In this study the functional model is extended to enable estimation of

kinematic displacement parameters (i.e., parameters describing displacements but not necessarily related to an underlying geophysical process)[1].

Stochastic model. The reference PS technique assumes the interferometric data have equal weight. This may be incorrect due to, e.g., differences in atmospheric conditions during the acquisitions. In this study a variance component stochastic model is formulated based on the originally observed phase. This model accounts for random noise and varying atmospheric conditions during the acquisitions. The variance factors of this model are estimated after an initial estimation using a default stochastic model.

Estimator. In the reference PS technique the parameters are estimated using ensemble coherence maximization. This estimator treats all data having the same weight. The algorithm used in this study utilizes the (weighted) integer least-squares estimator which has the highest probability of correct integer ambiguity resolution of all estimators for problems with a multivariate normal distribution.

Precision. The reference PS technique uses the coherence (a number between zero and one) to describe the precision of the estimated parameters. Although the coherence can be transformed to a phase dispersion it does not describe correlation between estimated parameters. In this study the full variance-covariance matrix of the estimated parameters is obtained by error propagation using the estimated variance factors of the stochastic model.

Reliability. In the reference PS technique estimates with a high coherence are assumed to be correct. During the preliminary estimation step it is assumed that all points are coherent and incorrect estimates have a negligible effect on the finally estimated parameters due to a weighted least-squares adjustment and low-pass filtering of the residual interferometric phase. In this study an alternative hypothesis testing procedure of a redundant network is performed to identify incoherent points and incorrect estimates.

1.2 Outline

This book is organized as follows. A review of the reference PS technique (Ferretti et al., 2000a, 2001) is given in Chapter 2. Potential improvements upon the reference technique are identified and the functional and stochastic model are derived. The next chapters focus on these improvements, making use of the derived mathematical model. Chapter 3 introduces the integer least-squares estimator, which is used in the developed algorithm to estimate unknown integer ambiguities and float parameters. The Spatio-Temporal Unwrapping Network (STUN) algorithm, which is developed for the estimation of

[1] In Ferretti et al. (2000a) the term non-linear deformation is used to indicate small deviations from the linear model that are obtained using filtering of the residual phase. In this study the displacement is completely parameterized.

displacement parameters, is described in Chapter 4. The developed algorithm and software are validated using realistic simulations, see Chapter 5. The STUN algorithm is applied to two urban areas with different characterization. Results for the Berlin and Las Vegas test sites are presented in Chapter 6. Finally, Chapter 7 reports the general results and conclusions, and gives recommendations for future research.

Proof of the variance component estimation equations is given in Appendix A, and Appendix B describes the Delft alternative hypothesis testing theory. A list of the radar data used during this study is provided in Appendix C. The software environment at the DLR and implementation specifics of the STUN algorithm are described in Appendix D. The software on the CDROM, demonstrating key concepts described in this book and providing the essential building blocks for Persistent Scatterer Interferometry, are described in Appendix E.

2

The Permanent Scatterer Technique

The Permanent Scatterer (PS) technique has been developed in the late 1990s by A. Ferretti, F. Rocca, and C. Prati of the Technical University of Milan (POLIMI) to overcome the major limitations of repeat pass SAR interferometry; temporal and geometrical decorrelation, and variations in atmospheric conditions. The main characteristics of this multi-image processing method are that it utilizes a *single master* stack of differential interferograms, and that only time–coherent pixels, i.e., "Permanent Scatterers," are considered. Furthermore, this technique distinguishes itself from other common interferometric processing methods by the fact that all acquired images can be used, including those with large baselines. This is the case since pixels with point-like scattering do not suffer from geometrical decorrelation as targets with a distributed scattering mechanism do, and such pixels thus remain coherent in all interferograms.

The PS technique, which is referred to as the reference technique in this book, is described in detail in section 2.1. The term "PS technique" is used to refer to this description. Potential improvements upon the PS technique are identified in section 2.2. These issues are addressed in the following chapters.

2.1 The reference PS technique

The key processing steps of the PS technique are the following (see, e.g., Ferretti et al., 1999b,c):

1. Computation of the interferograms.
2. Computation of the differential interferograms using a digital elevation model (DEM).
3. Preliminary estimation—at a coarse grid—of the presumably most coherent pixels. These pixels are referred to as Permanent Scatterer Candidates (PSCs).

4. Refinement of step 3. In the PS technique the long wavelength part of the atmospheric signal is estimated at the coarse grid of PSCs. After interpolation of these estimates, the differential interferograms are corrected, and additional PSs are computed.

The following description of the PS technique is based on (Colesanti et al., 2003a; Ferretti et al., 2000a, 2001). The PS technique is protected by a patent (Ferretti et al., 2000b), held by POLIMI, and the term "Permanent Scatterer technique" is trademarked. A commercial POLIMI spin-off company was founded that exploits the patent and performs ongoing research, see (Tele-Rilevamento Europa, 2004). Therefore, details of the implementation of the PS technique are not always clear. In the following sub-sections, occasionally, our approach is described, but only in cases where the details of the approach used in the PS technique are unclear, or a different approach only has a minor impact on the results. Non-trivial differences between the reference technique and our approach are described in the next chapters.

2.1.1 Interferogram formation

Given $K+1$ SAR images (all the available images on the same track), K interferograms are formed with respect to the same master image m. The SAR images are oversampled by a factor of two in range and azimuth direction before interferogram generation in order to avoid aliasing of the complex interferometric signal. Therefore, the amount of data that needs to be handled is considerable, even if the area of interest is limited to a city and its surroundings (i.e., typically less than five percent of the total area of a full-scene SAR image). For a typical project with $K=50$ interferograms, the required online storage space for convenient processing is approximately 100 GB. Although current computer systems have this storage available at low cost, the processing time to handle these amounts of data is a factor of importance when a processing environment is selected (regarding the amount of memory, the speed of the disk drives, and usage of multiple CPUs). Note that while the amount of data can be processed and stored by current computer systems without major difficulties, there are too many unknown parameters to perform their estimation in a single step. Aside from the amount of data, a second difference with conventional interferometric processing is that spectral range and azimuth filtering is not applied, since only targets with a point-like scattering mechanism are considered.

The use of a single master image implies that the temporal, geometrical, and/or Doppler baseline (difference in Doppler centroid frequency) will be large for a number of interferograms, leading to decorrelation of targets that have a distributed scattering mechanism. This may cause difficulties in the coregistration because standard algorithms require a certain level of coherence (see for example Hanssen, 2001). Therefore, in our implementation a newly developed geometric coregistration procedure is applied using a DEM of the

area and precise orbit data. The offset of the slave image with respect to the master image is computed on a grid of virtual tie-points using a zero-Doppler iteration scheme (see, e.g., Hanssen, 2001). Using this information, the higher order terms of the coregistration polynomial are determined. The zero-order terms of the polynomial are estimated using a point matching procedure, since timing errors in range and azimuth prevent an accurate geometric solution for these terms. This algorithm is described in detail in (Adam et al., 2003).

The master image is selected such that the dispersion of the perpendicular baselines is as low as possible, see (Colesanti et al., 2003a). In our implementation, the master image is selected maximizing the (expected) stack coherence of the interferometric stack, which facilitates visual interpretation of the interferograms and aids quality assessment. The stack coherence for a stack with master m is defined as

$$\gamma^m = \frac{1}{K}\sum_{k=0}^{K} g(B_\perp^{k,m}, 1200) \times g(T^{k,m}, 5) \times g(f_{\mathrm{dc}}^{k,m}, 1380), \tag{2.1}$$

where

$$g(x,c) = \begin{cases} 1-|x|/c & \text{if } |x|<c \\ 0 & \text{otherwise} \end{cases}, \tag{2.2}$$

and $B_\perp^{k,m}$ is the perpendicular baseline between images m and k at the center of the image, $T^{k,m}$ the temporal baseline (in years), and $f_{\mathrm{dc}}^{k,m}$ the Doppler baseline (the mean Doppler centroid frequency difference). The divisor c in Eq. (2.2) can be regarded as a critical baseline for which total decorrelation is expected for targets with a distributed scattering mechanism. The values given in Eq. (2.1) are typical for ERS, but they can be easily adapted to any other sensor with a different wavelength, look angle, and/or bandwidth. Fig. 2.1 shows an example of the stack coherence function for a real data stack of Berlin, where 70 SAR images are available. The images are sorted according to the acquisition time. Note that in general the stack coherence γ^m is larger when the master is selected more centrally in time, but that it decreases when it does not lie centrally regarding the perpendicular or Doppler baseline.

2.1.2 Differential interferogram formation

A reference digital elevation model (DEM) and precise orbit data are used to obtain K differential interferograms. The interferometric phase component that is induced by topography is largely eliminated using the differential technique, see, e.g., (Bamler and Hartl, 1998; Bürgmann et al., 2000; Eineder, 2003; Massonnet and Sigmundsson, 2000; Rosen et al., 2000). The differential interferometric phase is used in all further computations. In the following, the term (interferometric) phase refers to the differential interferometric phase, except when explicitly stated otherwise.

Instead of using an existing DEM, a height model can also first be generated from a subset of the available images, preferably with large perpendicular

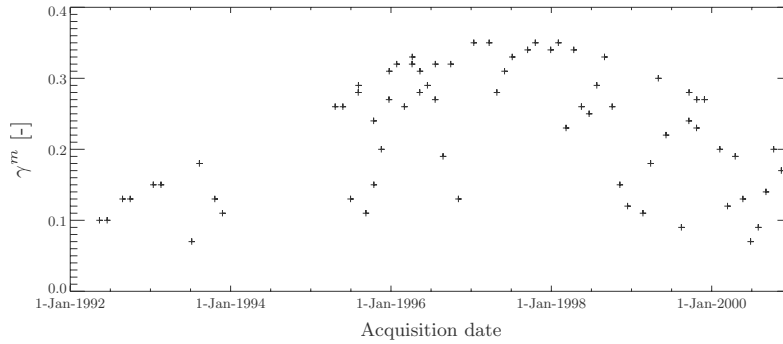


Fig. 2.1: Example of the stack coherence function, Eq. (2.1), for 70 available acquisitions of the Berlin area, track 165, frame 2547.

and small temporal baselines, see (Ferretti et al., 1999a). This is the standard approach in the PS technique (Colesanti et al., 2003a). However, after the Shuttle Radar Topography Mission (SRTM), a DEM of sufficient precision is readily available for practically any area of interest between -57° and 60° latitude (Suchandt et al., 2001). The DEMs have a vertical resolution of one meter (i.e., the elevation value is given in integer meters) and a horizontal spacing of 1 arc second (approximately 30 meters at the equator). The SRTM DEM accuracy specifications are 16 m absolute and 6 m relative for the vertical direction, and 20 m absolute and 16 m relative horizontally (90% confidence), see (Rabus et al., 2003). In our implementation, the SRTM X-band DEM is used for topographic correction since it is expected to be more precise than the C-band DEM, due to the shorter wavelength and the mode of operation used (Rabus et al., 2003). However, the X-band DEM does not have continuous coverage due to its smaller swath-width. If the area of interest is not fully covered by the X-band DEM the C-band DEM is (partially) used. Although the best available DEM is used, the results of the PS processing do not depend on the precision of the DEM, since for each pixel also the elevation with respect to the DEM is estimated. Keep in mind that even an extremely precise DEM does not allow to fully correct the interferometric phase, since the location of the scatterer is not known, e.g., the backscattered echo for a pixel at a certain range to the radar could come from the street, bounced via a wall, from a ledge in a window, or from a rooftop, or any combination thereof.

It was noted by Colesanti et al. (2003b) that the PS analysis can also be carried out without using a reference DEM, but only compensating the interferograms for a flat topography, since in the PS technique a topographic term is estimated anyway.

Functional model

The functional model that is used in the PS technique for the unwrapped differential interferometric phase Φ^k for a point in interferogram k is given in (Colesanti et al., 2003a) as

$$\Phi^k = \phi^k_{\text{topo}} + \phi^k_{\text{defo}} + \phi^k_{\text{atmo}} + \phi^k_{\text{noise}}, \tag{2.3}$$

where ϕ^k_{topo} is the phase due to inaccuracy of the reference DEM, ϕ^k_{defo} is the phase due to displacement of the point, ϕ^k_{atmo} is the phase due to atmospheric delays, and ϕ^k_{noise} is decorrelation noise. The topographic phase practically is a linear function of the perpendicular baseline, and can be written as

$$\phi^k_{x,\text{topo}} = \beta^k_x \cdot \Delta h_x, \tag{2.4}$$

where β^k_x is the height-to-phase conversion factor for point x, and Δh_x is the height of the point relative to the reference surface, referred to as DEM error (see Eq. (2.12) on page 17 for the definition of β). A time-linear model is used to model the displacement of each point x. Therefore,

$$\phi^k_{x,\text{defo}} = -\frac{4\pi}{\lambda} T^k \cdot \alpha(x), \tag{2.5}$$

where λ is the wavelength of the radar carrier signal, T^k is the temporal baseline with respect to the master acquisition, and $\alpha(x)$ is the average displacement rate at point x. The phase ϕ_{atmo} due to atmospheric signal is not modeled, but reduced considerably by considering phase differences between nearby points. The noise term contains all other phase contributions. If the displacement (difference between points) deviates from a time-linear behavior this signal is thus also contained in the noise term. In the PS technique a temporal high-pass filter is used to separate temporally correlated displacement signal from random noise. The next section describes the estimation of these signal components in detail.

2.1.3 Preliminary estimation

In the PS technique the preliminary estimation is limited to a coarse set of pixels, called Permanent Scatterer Candidates (PSCs). The goal of this step is to estimate the atmospheric phase at these pixel positions in all interferograms. This is accomplished by filtering of the residual phase after estimation of the modeled parameters, i.e., the DEM error and the displacement rate, taking advantage of the spatial correlation of the atmospheric signal. The estimations are performed between nearby points, because the phase contributions that are not modeled need to be smaller than π (since the observed data is not unwrapped), and the atmospheric signal is considerably reduced by taking this difference. After these estimations between points, the residual phase is

integrated, yielding the unwrapped residual phase at the PSC positions with respect to a reference point.

In the following, first the selection of the PSCs using the amplitude dispersion index is described, then the estimation of the parameters using the ensemble coherence, and finally the filtering that is performed to separate atmospheric signal from temporally correlated displacement and random noise.

Amplitude dispersion index thresholding

The amplitude dispersion index D_a, and its relation to the phase standard deviation σ_ϕ, is defined in (Ferretti et al., 2001) as

$$\hat{\sigma}_\phi = \frac{\sigma_a}{\bar{a}} = D_a, \tag{2.6}$$

where σ_a is the temporal standard deviation of the amplitude and $\bar{a}$ the temporal mean of the amplitude for a certain pixel. Thus, a pixel that consistently has a similar, relatively large, amplitude during all acquisitions is expected to have a small phase dispersion. This relation enables the identification of potentially coherent points without the need to analyze the phase. The latter would not be possible at this moment, since the phase still contains unknown signal contributions. Moreover, the amplitude dispersion index, Eq. (2.6), does not regard neighboring pixels. This enables the detection of isolated points, which is not possible if this detection is based on a spatially estimated coherence value, as, for example, done in (Usai, 1997; Usai and Hanssen, 1997).

Points are selected as PSC if the amplitude dispersion is below a threshold, typically between 0.25 and 0.4 (Colesanti et al., 2003a; Ferretti et al., 2001). Colesanti et al. (2003a) report that the PSC density must be at least ~3 PSC/km^2, since otherwise the atmospheric signal cannot reliably be interpolated. The estimation of the parameters is restricted to these selected pixels in the preliminary estimation step. Ferretti et al. (2001) have shown using a numerical simulation that the estimation of the phase stability based on the amplitude dispersion holds very well for $\sigma_\phi < 0.25$ rad (~15°) if $K = 33$. This experiment is repeated here, see Fig. 2.2. For larger values of the amplitude dispersion index there is no linear relation with the phase standard deviation. The amplitude dispersion index tends to 0.5 for low SNR, see also (Ferretti et al., 2001). Nonetheless, points with a smaller amplitude dispersion index are expected to have a smaller phase standard deviation. Therefore, thresholding on the dispersion index is a very practical way of selecting points that are expected to have the smallest phase dispersion.

The images need to be radiometrically calibrated in order to allow for the estimation of σ_a and $\bar{a}$ (Ferretti et al., 2001). In our implementation, the data are calibrated for antenna pattern, range spreading loss, and gain factor, (relevant to the sensor, acquisition time, and processing center), see

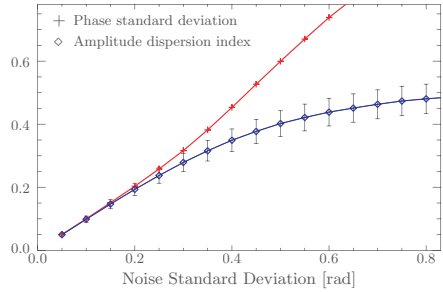


Fig. 2.2: Numerical simulation for the amplitude dispersion index following (Ferretti et al., 2001). A complex variable $z = s+n$ is simulated at 5000 points. The signal was fixed to $s = 1$, while the noise standard deviation on the real and imaginary parts of n was gradually incremented from 0.05 to 0.8. 34 data sets are supposed to be available ($K = 33$). The mean estimated dispersion index D_a (diamonds) and their standard deviations are plotted as function of the noise standard deviation, together with the phase standard deviation σ_ϕ (plus marks). Small values of the amplitude dispersion index are a good estimate for the phase standard deviation.

also (Adam et al., 2003; Laur et al., 1998). However, this blind calibration (i.e., based on annotated parameters without examining the data) may not work for all SLC images, likely due to an incorrectly annotated calibration constant in the leader file. To ensure that the calibrated images are comparable, in our implementation, the histograms of the calibrated intensity images are plotted on top of each other. If the modes of the histograms vary more than what could reasonably be expected, say 1 dB, then the histograms are all shifted to the mode of the first image. Since decibel is used as a unit, this is equivalent to multiplication of the intensity with a re-computed calibration constant. (To avoid large random variation of the backscatter due to changes in soil moisture and surface roughness, etc., the histograms are computed for a user-selected polynomial region, e.g., ~10 km^2 of inner city area.)

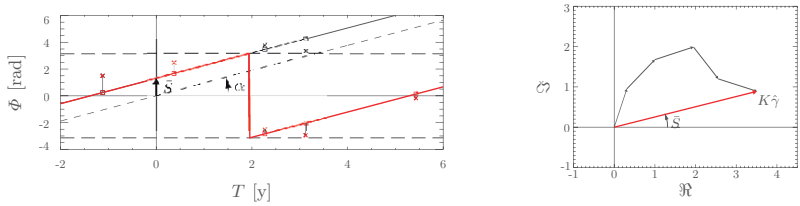
Ensemble coherence maximization

The estimation of the parameters from the observed wrapped phase data, see Eq. (2.3), is a non-linear inversion problem that cannot be solved by direct inversion (Teunissen et al., 1995c). A search through the solution space must always be performed. In the PS technique the absolute value of the complex ensemble coherence (i.e., in time)

$$\hat{\gamma}_{x,y} = \frac{1}{K} \sum_{k=1}^{K} \exp(j e^k_{x,y}), \tag{2.7}$$

is used as a norm, where j is the imaginary unit, and $e^k_{x,y}$ is the difference between the observed and modeled phase between points x and y in interferogram k. The "hat" in $\hat{\gamma}$ is used to stress that Eq. (2.7) is an estimate of the

unknown coherence. Fig. 2.3 gives a geometric interpretation of the complex coherence for a simplified case where the topographic phase is ignored and only 5 interferograms are considered. Normally a linear displacement model



(a) Interferometric residual phase

(b) Complex sum of residual phase

Fig. 2.3: Geometric interpretation of the complex coherence. The observed phase (bold ×) is wrapped between $-\pi$ and π. It is modeled (bold □) using a linear displacement rate α. The unwrapped phase and model are displayed using a normal font face. The residual phase is indicated by lines. The coherence is the complex sum (red arrow) of the residual phase between observations and model, depicted in 2.3(b). The angle of the complex coherence corresponds to the average residual $\bar{S}$.

is assumed in the PS technique, i.e.,

$$e^k_{x,y} = \phi^k_{x,y} \; - \; \big(\beta^k_x \cdot \Delta h_{x,y} - \frac{4\pi}{\lambda} T^k \cdot \alpha(x,y)\big), \tag{2.8}$$

see also Eqs. (2.4) and (2.5). The solution for the DEM error (difference) and displacement rate (difference), $\max_{\Delta h,\alpha} |\hat{\gamma}_{x,y}|$, is found in practice by sampling the two-dimensional solution space with a certain resolution and up to certain bounds, each time evaluating this norm.

Since the observed phase data are known only modulo 2π, estimation of the parameters can only be successfully performed if the residual phase (difference) terms are small enough, certainly $|e^k_{x,y}| < \pi$ in most interferograms, but a more reasonable number is said to be $\sigma_e \leq 0.6$ rad, see (Colesanti et al., 2003a) or (Colesanti et al., 2003b). The residual phase difference $e_{x,y}$ is assumed to be small, since all its components are small:

- Atmospheric difference signal:
  The atmospheric delay difference between two nearby points is small. For points less than 1 km apart the standard deviation of the difference commonly is below 0.3 rad (Ferretti et al., 2000a);
- Displacement that is not modeled:
  The linear model is assumed to be a good approximation for the actual displacement between two nearby points.

- Random noise:
  PSCs are selected by thresholding the amplitude dispersion index. If a 0.4 threshold is used for the dispersion index, then according to Eq. (2.6), the noise standard deviation is expected to be below $\sqrt{2} \cdot 0.40 = 0.56$ rad for all considered differences. (Note though that the amplitude dispersion underestimates the phase noise by approximately 50% for a value of 0.40, see Fig. 2.2.)

The absolute value of the coherence lies in the interval [0,1], where a coherence of 1 signifies complete correspondence of the modeled phase with the observed phase. The angle of the complex coherence is said to be an estimate for the master atmospheric signal in (Ferretti et al., 2001), but it would be more exact to refer to it as the average interferometric residual phase, see Fig. 2.3. The reason that it is called master atmosphere lies in the fact that the master image is present in all interferograms, and that, for example, a large atmospheric delay during the master acquisition would clearly be visible in this average. However, it is not true that the average residual phase is always caused by an atmospheric delay during the master acquisition.

Filtering to obtain the atmospheric phase

The atmospheric states during the radar acquisitions are assumed to be uncorrelated in time and correlated in space. Therefore, the atmospheric signal can be isolated from the other components of the residual phase (i.e., time non-linear displacement and random noise) by low-pass filtering in the spatial domain and high-pass filtering in the temporal domain. At least, when it is also assumed that unmodeled displacement is correlated in the temporal domain. But first the residual phase must be obtained at the PSCs points.

In the original PS algorithm, the parameters at the PSC points are estimated directly with respect to a reference pixel. The atmospheric phase, merged with possible phase contributions caused by orbit errors, was approximated (in first instance) for each differential interferogram as a linear phase term in range and azimuth direction. These phase components are estimated using an iterative approach, described in appendix A of (Ferretti et al., 2001). It was soon recognized by Ferretti et al. (2000a) that the planar approximation of the atmospheric phase restricted the application of the PS technique to small areas of approximately 5×5 km^2.

In order to avoid this problem, in later implementations of the PS technique, the DEM error and displacement rate differences are first estimated between all pairs of PSCs within a certain maximum distance, e.g., within 2 to 3 km, see (Colesanti et al., 2003a). As soon as these estimations are performed, the observed phase at the PSCs can be unwrapped in each interferogram with respect to a reference pixel, assuming $|e^k_{x,y}| < \pi$ (Colesanti et al., 2003a). If this condition is fulfilled, and all estimations between points are correct, then the unwrapped residual phase is obtained by integration along any path. However,

this is not guaranteed. Therefore, a weighted least-squares integration is used for this phase unwrapping in the PS technique (Ferretti et al., 2000a), yielding the unwrapped residual phase e_x^k at the PSC positions in all interferograms. Pairs with an ensemble coherence below a certain threshold, e.g., $|\hat{\gamma}| < 0.75$, see (Ferretti et al., 2001), are zero-weighted (discarded), while for the other pairs the coherence is used as weight in the least-squares adjustment (Ferretti et al., 2000a).

Now that the residual phase e_x^k is obtained at the PSC positions it is filtered to isolate the atmospheric component. First the temporal mean residual phase $\bar{e}_x$ is subtracted for each PSC, see (Ferretti et al., 2000a).

$$e_x^{k\prime} = e_x^k - \bar{e}_x. \tag{2.9}$$

This mean is an estimate for the atmospheric phase during the master acquisition, and it is removed because it will not pass the high-pass filter that is performed next. The temporal high-pass filtering is performed to remove possible temporally correlated displacement from the residual phase. Finally, a spatial low-pass filter is applied to the temporally filtered residuals in order to remove the random noise component. These filtering steps can be written symbolically, cf. (Ferretti et al., 2000a), as

$$\hat{\phi}_{x,\text{atmo}}^k = \left[\left[e_x^{k\prime}\right]_{\text{HP\_time}}\right]_{\text{LP\_space}} + \left[\bar{e}_x\right]_{\text{LP\_space}}, \tag{2.10}$$

where $\hat{\phi}_{x,\text{atmo}}^k$ is the estimated atmospheric phase at PSC position x in interferogram k. In (Ferretti et al., 2000a) a triangular window of length 300 days was used for the temporal filter and a 2×2 km^2 averaging window for the spatial filter. Note that the order of the filtering steps can be interchanged, and also that the temporal filter could be applied before the integration step.

Ferretti et al. (2000a) suggest that these filtering steps require the unwrapped residual phase, but this is not really necessary. By using a complex filter, i.e., filtering the real and imaginary parts of the complex residual signal separately, the (wrapped) low wavelengths can be easily obtained. In this case there is no need for phase unwrapping, since the complex filtered residuals can directly be subtracted from the original phase data, which is also wrapped. (Though if unwrapping is performed, the chance of occurrence of unwrapping errors is likely smaller after complex filtering, because the number of residuals are likely much lower, depending on the power of the signal in the higher frequencies.)

2.1.4 Final estimation

After the low wavelength part of the atmospheric delays is estimated at the PSC positions, it is interpolated at the original resolution of the differential interferograms. The interpolated atmospheric signal is referred to as "atmospheric phase screen" (APS). It is noted in (Colesanti et al., 2003a) that the

step of spatial low-pass filtering and interpolation of the residual phase can also be performed simultaneously using Kriging interpolation, instead of using the 2×2 km^2 moving averaging window.

The interpolated APSs are subtracted from the differential interferograms at full resolution, and additional PS points are searched for. This is done on a pixel-by-pixel basis (i.e., not between nearby pixels, although still with respect to a reference), since there is no need anymore to consider phase differences between nearby points. After all, the computations of the preliminary estimation step, described in the previous section, are performed between nearby points because otherwise the atmospheric signal would prevent a correct estimation, and this signal is now removed. The same functional model, Eq. (2.7), is used here as during the preliminary estimation step.

Points with an estimated ensemble coherence below a certain threshold are discarded, e.g., $|\hat{\gamma}| < 0.75$ (Ferretti et al., 2001). The number of points that finally can be used, is in the order of a few hundred points per square kilometer (in urban areas), according to Ferretti et al. (2001). The same strategy of low-pass temporal filtering, that is described at the end of the previous section, is used to estimate temporally correlated displacements that deviate from the linear displacement model.

2.2 Potential improvements

Despite its spreading application, the reference PS technique does not necessarily provide optimally estimated parameters under all circumstances, particularly in cases when the assumptions on the displacement model and properties of the signal components are not valid. Possible problem areas are identified here, related to the following assumptions made in the PS technique:

- The functional model contains all phase components, see Eq. (2.3).
  - As will be shown in section 2.2.1, the sub-pixel position of the PS point induces an additional phase that should be accounted for, particularly when there are significant differences between the radar frequencies and/or Doppler centroid frequencies of the acquired images. This additional phase term was also introduced in the PS technique, i.e., in the reference technique, when ERS–ENVISAT cross interferometry was discussed (Arrigoni et al., 2003; Colesanti et al., 2003d).
  - Moreover, in the model used in the PS technique, phase due to orbit errors is lumped with the atmospheric signal. These terms should be separated in the functional model.
- Displacement can be described using a constant rate, see Eq. (2.5).
  - The problem is over-parameterized in case a PS point does not undergo displacement. A significance test should be used to detect whether a displacement parameter can be significantly estimated, and if this would not be possible, the estimation should be repeated without such

a parameter. In principle, the strictest model, with the smallest amount of parameters, should be used as a null-hypothesis.
  - On the other hand, if a point undergoes a more complex motion, the danger is that this signal leaks to the estimated atmospheric signal (see also the next item). The displacement model should be adapted when the actual displacement significantly deviates from the model. Changing the displacement model is not a problem in the PS technique, but as noted by Colesanti et al. (2003b)—who estimated seasonal displacement using the PS technique—the ensemble coherence estimator becomes increasingly slow when more parameters are added to the displacement model.
- Atmospheric signal can be obtained by filtering the residual phase, see Eq. (2.10).
  - The residual phase at the PSCs in each interferogram is obtained by a least-squares unwrapping of the residual phase between selected pairs. This works well if all difference parameters are correctly estimated, and the residual phase between points is small (at least smaller than π). Possible incorrectly estimated parameters are not detected. Instead of relying on the assumption that a high coherence indicates correctly estimated parameters, an integrity test for the network could be introduced.
  - Unmodeled displacement is assumed to be filtered out by the application of a temporal high-pass filter. If the unmodeled displacement is not correlated in time it would thus be estimated as atmospheric signal. Contrary, atmospheric signal that is temporally correlated is not removed by this filter. Temporal correlation of the atmospheric signal could be due to, e.g., relief (in combination with wind or temperature), or the atmospheric water vapor content of an area near, e.g., a lake may be seasonally correlated. Moreover, in practice the temporal filter may not perform well for images at the start and end of the time series, and if there are temporal gaps. In the developed approach, to be described in Chapter 4, the atmospheric signal is not estimated, but dealt with in the stochastic model, as suggested by Hanssen (2001). The parameters of the final points are estimated with respect to a reference network, established in the preliminary estimation step.
- All observations have the same weight, see Eq. (2.7).
  - The ensemble coherence estimator treats the data having the same weight. If the observations do not have the same precision (for example due to differences in atmospheric circumstances, coregistration accuracy, or sensor type) a better estimate of the parameters can be obtained by using the actual precision of the observations.
- The quality of the estimated displacement is adequately described by the coherence, see Eq. (2.7).
  - Correlation between estimated parameters is not described by the coherence.

- Moreover, the coherence does not directly provide the precision of the estimated displacement at a certain epoch.

These issues can only be studied after rigorously deriving the functional and stochastic model. This is the subject of sections 2.2.1 and 2.2.2.

2.2.1 Functional model

The wrapped phase ϕ_x^k of a point in differential interferogram k can be decomposed to

$$\phi_x^k = W\{\phi_{x,\text{topo}}^k + \phi_{x,\text{defo}}^k + \phi_{x,\text{obj}}^k + \phi_{x,\text{atmo}}^k + \phi_{x,\text{orbit}}^k + \phi_{x,\text{noise}}^k\}, \tag{2.11}$$

where $W\{.\}$ is the wrapping operator[1], ϕ_{topo} is the phase caused by uncompensated topography, ϕ_{defo} is the phase caused by a displacement of the target in the time between the acquisitions, ϕ_{obj} is the object scattering phase related to the path length traveled in the resolution cell, ϕ_{atmo} is the atmospheric phase accounting for signal delays, ϕ_{orbit} is the phase caused by imprecise orbit data, and ϕ_{noise} is the additive noise term. The topographic phase is related to the elevation of the target with respect to the reference surface Δh_x, referred to as DEM error in this work, as (Rodriguez and Martin, 1992)

$$\begin{aligned}\phi_{x,\text{topo}}^k &= -\frac{4\pi}{\lambda}\frac{B_{\perp x}^k}{r_x^m \sin\theta_{x,\text{inc}}^m}\cdot \Delta h_x \\ &= \beta_x^k \cdot \Delta h_x,\end{aligned} \tag{2.12}$$

where λ is the wavelength of the carrier signal used by the radar system, $B_{\perp x}^k$ is the local perpendicular baseline, r_x^m is the range from master sensor to the pixel, and $\theta_{x,\text{inc}}^m$ is the local incidence angle, see also Fig. 2.4. The height to phase conversion factor β relates a change in height to a change in phase. This factor is computed for each pixel using a DEM of the area. It is equal to the phase difference between a synthetic interferogram computed from the DEM directly and from the DEM with a bias of one meter added to it. The displacement term equals

$$\phi_{x,\text{defo}}^k = -\frac{4\pi}{\lambda}\Delta r_x^k, \tag{2.13}$$

where Δr_x^k is the line-of-sight displacement toward the radar since the acquisition time of the master image. In order to limit the number of parameters that needs to be estimated, the displacement behavior needs to be modeled and parameterized. The displacement since the time of the master acquisition is modeled using a linear combination of base functions as

$$\Delta r_x^k = \sum_{d=1}^{D} \alpha_d(x)\cdot p_d(k). \tag{2.14}$$

[1] $W\{x\}=\angle\exp(jx)$

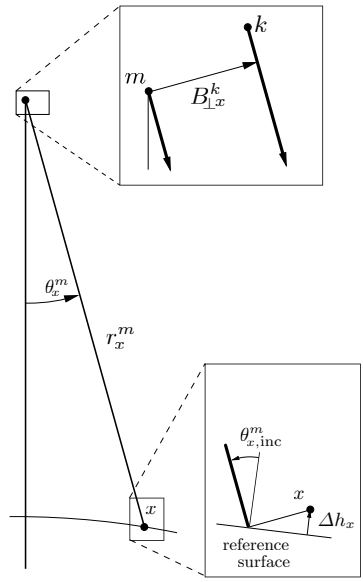


Fig. 2.4: Satellite configuration for across-track radar interferometry. The master sensor m and slave k go "into" the paper. A point x is observed in the pixel of the master image at a range r^m_x and under a look angle θ^m_x.

For example, algebraic polynomials of the temporal baseline T^k could be used

$$p_i(k) = (T^k)^i, \quad \text{for } i = 1, 2, \ldots \tag{2.15}$$

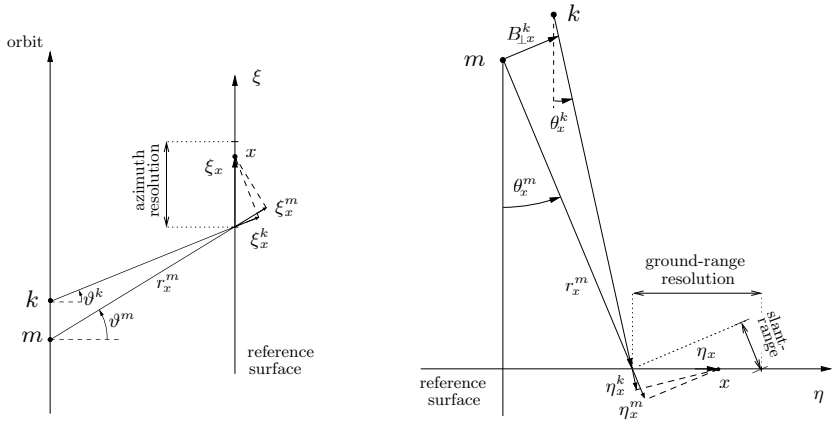
The condition $\Delta r^m_x = 0$, i.e., the observed displacement is relative to the acquisition time of the master image, is satisfied in this case because these base functions are all equal to zero at the time of the master acquisition. For more general base functions this must be forced by subtracting $p_d(T^m)$ from each base function d. In the PS technique a linear displacement rate $\alpha(x)$ is assumed by default, i.e.,

$$\Delta r^k_x = \alpha(x) \cdot T^k. \tag{2.16}$$

The object scattering term in Eq. (2.11) depends on the azimuth ξ_x and ground-range η_x sub-pixel position of the phase center of the point scatterer within the resolution cell as

$$\begin{aligned} \phi^k_{x,\mathrm{obj}} &= \phi^{k,m}_{\xi_x} + \phi^{k,m}_{\eta_x} \\ &= \big(\frac{4\pi}{\lambda^k} \sin \vartheta^k_x - \frac{4\pi}{\lambda^m} \sin \vartheta^m_x\big) \cdot \xi_x + \big(\frac{4\pi}{\lambda^k} \sin \theta^k_x - \frac{4\pi}{\lambda^m} \sin \theta^m_x\big) \cdot \eta_x, \end{aligned} \tag{2.17}$$

where λ^k is the wavelength of the carrier signal used by the radar, ϑ is the squint angle, and θ is the look angle. Fig. 2.5 shows the geometry for these terms.



(a) Azimuth **(b)** Range

Fig. 2.5: Geometry for a point scatterer located at a sub-pixel position in **(a)** azimuth and **(b)** ground-range. The phase in the interferogram is computed at the pixel position corresponding to the leading edge of the resolution cell, while the phase center of the scatterer x is actually located at the sub-pixel position ξ_x in azimuth and η_x in ground-range. The observed interferometric phase, corrected for the phase of the reference surface, still contains the contribution due to the path length difference that the signal traveled within the resolution cell, unless the phase is interpolated at the exact sub-pixel position of the point.

The azimuth term can also be expressed in terms of the Doppler centroid frequency. Using a simplified rectilinear imaging geometry, the Doppler centroid frequency can be written as (Bamler and Schättler, 1993; Fernandez et al., 1999)

$$f_{\mathrm{dc}}^k = \frac{-2v}{\lambda^k} \sin \vartheta^k, \tag{2.18}$$

where v is the instantaneous velocity of the satellite in an earth-fixed coordinate system. For a curved geometry a correction factor close to one needs to be applied, which accounts for the slightly smaller beam velocity on the ground, see also (Cumming and Wong, 2005; Raney, 1986). From Fig. 2.5(a) it is clear that the additional range from the start of the bin to the actual position is

$$\xi_x^k = \xi_x \sin \vartheta^k. \tag{2.19}$$

By substitution of Eq. (2.18) in Eq. (2.19) it follows that

$$\xi_x^k = \frac{\lambda}{-2v} f_{x,\mathrm{dc}}^k \cdot \xi_x, \tag{2.20}$$

which is equivalent to an additional phase of

$$\phi_{\xi_x}^k = \frac{2\pi}{v} f_{x,\mathrm{dc}}^k \cdot \xi_x, \tag{2.21}$$

using that $\phi/(2\pi) = -2r/\lambda$ for repeat pass interferometry. The interferometric phase, caused by the azimuth sub-pixel position of the point scatterer, can finally be expressed as

$$\phi_{\xi_x}^{k,m} = \frac{2\pi}{v}\left(f_{x,\mathrm{dc}}^{m} - f_{x,\mathrm{dc}}^{k}\right) \cdot \xi_x, \tag{2.22}$$

assuming equal sensor velocities v. Note that this phase depends on the wavelength used by the radar, even though this is not directly visible from Eq. (2.22).

In (Colesanti et al., 2003d), the interferometric phase caused by the range sub-pixel position was expressed as

$$\phi_{\eta_x} = \frac{4\pi}{c}\left(\eta_x^m \Delta f + \eta_x^m \frac{f^m B_\perp}{r^m \tan\theta^m}\right), \tag{2.23}$$

where c is the speed of light, f^m is the radar frequency of the master sensor, $\Delta f = f^k - f^m$ is the frequency offset of the slave sensor, and η_x^m is the slant-range sub-pixel position. Using $\lambda = c/f$, Eq. (2.23) can be written in terms of the wavelength as

$$\phi_{\eta_x} = \left(\frac{4\pi}{\lambda^k} - \frac{4\pi}{\lambda^m}\right) \cdot \eta_x^m + \frac{4\pi}{\lambda^m} \frac{B_\perp}{r^m \tan\theta^m} \cdot \eta_x^m. \tag{2.24}$$

The range sub-pixel term in Eq. (2.17) can be approximated using $\theta^m - \theta^k \approx B_\perp / r^m$ as

$$\begin{aligned} \phi_{\eta_x} &\approx \left(\frac{4\pi}{\lambda^k}\sin(\theta^m + \frac{B_\perp}{r^m}) - \frac{4\pi}{\lambda^m}\sin\theta^m\right) \cdot \eta_x \\ &\approx \left(\frac{4\pi}{\lambda^k}(\sin\theta^m + \frac{B_\perp}{r^m}\cos\theta^m) - \frac{4\pi}{\lambda^m}\sin\theta^m\right) \cdot \eta_x \end{aligned}. \tag{2.25}$$

The slant-range position of the scatterer in the master image η_x^m is related to the ground-range position as $\eta_x^m = \eta_x \sin\theta^m$, see Fig. 2.5(b). Substitution in Eq. (2.25) yields

$$\phi_{\eta_x} \approx \left(\frac{4\pi}{\lambda^k} - \frac{4\pi}{\lambda^m}\right) \cdot \eta_x^m + \frac{4\pi}{\lambda^k}\frac{B_\perp}{r^m}\cos\theta^m \cdot \eta_x \tag{2.26}$$

Moreover, it holds that $\eta_x \cos\theta^m = \eta_x^m / \tan\theta^m$, i.e., Eq. (2.26) can be written as

$$\phi_{\eta_x} \approx \left(\frac{4\pi}{\lambda^k} - \frac{4\pi}{\lambda^m}\right) \cdot \eta_x^m + \frac{4\pi}{\lambda^k} \frac{B_\perp}{r^m \tan\theta^m} \cdot \eta_x^m. \tag{2.27}$$

By comparison of Eq. (2.27) with Eq. (2.24), it follows that Eq. (2.17) and Eq. (2.23) are equivalent expressions for the interferometric phase caused by range sub-pixel position. The only difference is the usage of the master

wavelength for the baseline dependent term in (Colesanti et al., 2003d), although this has no practical consequences due to the small difference in wavelength.

The object scattering term $\phi^k_{x,\mathrm{obj}}$ vanishes if the phase is interpolated in the interferogram at the exact sub-pixel position of the PS point, which is the preferred approach. If such an analysis is not performed, e.g., due to computational constraints, the sub-pixel position can be estimated as an additional parameter. However, this may not be feasible:

- Estimation of the azimuth sub-pixel position may be cumbersome due to possible high correlation of this term with the displacement, since there is a systematic offset in Doppler centroid frequency between different sensors. For example, if ERS–1 and ERS–2 images are analyzed, the earlier ERS–1 images have a different Doppler centroid frequency. If all acquisitions have a larger variation of the Doppler centroid frequency, the correlation with time becomes less, making it more feasible to estimate the sub-pixel position from the observed phase data. This is the case for ERS–2 after February 2000, since around that time the gyroscopes of this sensor started to fail, causing problems with the yaw-steering. Note that such images cannot be used at all by processing techniques that rely on pixels with a distributed scattering mechanism.
- Moreover, the range sub-pixel position cannot be estimated if all acquisitions have the same wavelength, because there is a dependency on the perpendicular baseline, originating from the difference in look angle, see Eq. (2.25). The DEM error, Eq. (2.12), and the range sub-pixel position, Eq. (2.26), are both linear functions of the perpendicular baseline. Therefore, only the DEM error can be estimated if the wavelength is the same for all acquisitions. Specification of the wavelength used by the sensor, as done in Eq. (2.17), is important only when the sensors use a different wavelength, e.g., when cross interferometry is performed between ERS and ENVISAT, whose carrier frequencies differ approximately by one percent (the wavelength used by ERS is 5.66 cm and that by ENVISAT is 5.62 cm), see also (Adam et al., 2004; Arnaud et al., 2004). In that case, DEM error and range sub-pixel position should be estimated, because of the limited accuracy of the estimation of the peak position using a point target analysis. The cross-interferometric phase caused by the difference in wavelength of these sensors is approximately 1.3 rad/m in slant-range, which follows from the first term of the decomposition in Eq. (2.27).
  If sensors with different wavelengths are considered, the wavelength of the slave sensor should be substituted in the previous equations for ϕ_{topo} and ϕ_{defo}, although the effect is very small for these terms in the case of ERS–ENVISAT cross interferometry.

Fig. 2.6 shows the sensitivity of the interferometric phase to these signal components.

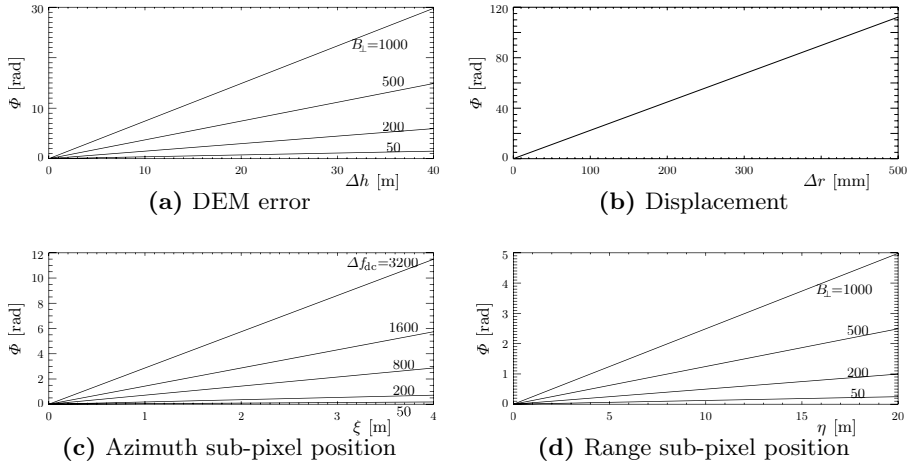


Fig. 2.6: Sensitivity of the differential interferometric phase to DEM error, displacement, and azimuth and range sub-pixel position for different values of the perpendicular and Doppler baseline, assuming equal wavelengths for master and slave. Typical ERS parameters are used for these computations (altitude 785 km, look angle $\theta = 21°$, wavelength $\lambda = 5.66$ cm, azimuth resolution 4 m, ground-range resolution 20 m).

The atmospheric phase $\phi^k_{x,\mathrm{atmo}}$ is caused by signal delay differences during the acquisitions, mainly due to water vapor in the troposphere. The amplitude of the atmospheric signal in the differential interferograms can be described by a power-law model of the form

$$P_{\Phi}(f) = P_0(f/f_0)^{-\beta}, \tag{2.28}$$

where f is some spatial frequency, P_0 and f_0 are normalizing constants, and $-\beta$ is the spectral index, see Hanssen (2001). The interferometric atmospheric signal can be considerable. For example, for the northern part of the Netherlands, typical values for the relative interferometric delay are 10 to 30 mm between points with a distance of approximately 10 km (Hanssen, 2001). This corresponds to ~2 to 6 rad in ERS interferograms, These phase values follow from Eq. (2.13), which relates displacement to interferometric phase, since a signal delay has the same effect on the phase as an actual increase of the path length due to displacement. The expectation of the atmospheric delay for a point x in acquisition k is $E\{S^k_x\} = 0$, and thus also $E\{S^{k,0}_{x,y}\} = 0$; the double-differenced atmospheric phase in interferogram k with master 0. This implies that it is not necessary to introduce an additional parameter to estimate the average atmospheric phase, which is done in the reference PS technique, see Fig. 2.3. However, if by chance the average atmospheric phase is relatively large—for example due to large spatial

variation of the atmospheric delays during the master acquisition—this phase is present in all interferograms, which can bias the estimation of the other parameters. Therefore, an additional parameter for the average interferometric atmospheric phase difference $\bar{S}_{x,y} = S^0_{x,y} - \frac{1}{K}\sum_{k=1}^{K} S^k_{x,y}$ can be introduced. The atmospheric phase in interferogram k is now written as

$$S^{k,0}_{x,y} = \bar{S}_{x,y} + (S^{k,0}_{x,y} - \bar{S}_{x,y}), \tag{2.29}$$

and $\bar{S}_{x,y}$ is estimated assuming $E\{S^{k,0}_{x,y} - \bar{S}_{x,y}\} = 0$.

It is well known that inaccuracies in the orbit data cause systematic phase errors in interferograms (Closa, 1998; Hanssen, 2001; Kohlhase et al., 2003). The reason is that the computed reference phase is incorrect (the so-called flat-earth phase, which is subtracted from the interferogram). Since the reference phase is mainly a function of range, orbit errors manifest mainly in range direction. For small areas (say smaller than 30 km), these errors can be well approximated by a linear phase ramp in range and azimuth direction

$$\phi_{\text{orbit}} = a + b \cdot \xi + c \cdot \eta, \tag{2.30}$$

see also (Hanssen, 2001). The bias a indicates that a reference point in the interferograms must be selected, with respect to which the other points are computed. In practice such a bias also absorbs differences in the absolute signal delay. Note that in Eq. (2.30) the symbols for azimuth ξ and range coordinate η are relative to the leading edges of the interferogram, while in Eq. (2.17) the symbols ξ_x and η_x for the sub-pixel positions are relative to the leading edges of the resolution cell. The orbit error phase is assumed to be small for most interferograms. Hanssen (2001) has shown that the maximum number of residual orbit fringes is less than one (95% confidence interval) in a 100×100 km^2 interferogram if 5 cm radial and 10 cm across-track rms is assumed for the orbit precision. Since we use precise orbits estimated by the GFZ (with comparable precision), in general the residual reference phase caused by orbit errors is smaller than a few radians over the area of interest. Note that a trend of the average displacement field cannot be distinguished from the average phase caused by orbit errors. This is the case for displacement estimation using a single interferogram as well as for the estimation using a data stack. However, the residual orbit trends are assumed to be uncorrelated between acquisitions, and their impact on the estimated displacement field is thus assumed to be small. To get an impression of the impact of this error on the estimated linear displacement rates, consider the case where ten interferograms are available, only containing phase ramps in range direction (caused by imprecise orbit data). Assume that the reference point is located at the left hand side, and that the standard deviation of the residual reference phase is one rad for points on the right hand side. If only a linear displacement rate α is estimated at a point x on the right, the following system of equations must be solved

$$\begin{bmatrix} \underline{\Phi}_x^1 \\ \underline{\Phi}_x^2 \\ \vdots \\ \underline{\Phi}_x^{10} \end{bmatrix} = \begin{bmatrix} -\frac{4\pi}{\lambda} T^1 \\ -\frac{4\pi}{\lambda} T^2 \\ \vdots \\ -\frac{4\pi}{\lambda} T^{10} \end{bmatrix} \alpha(x) + \underline{e}, \qquad D\{\underline{e}\} = I, \tag{2.31}$$

where T^k is the temporal baseline, see also Eq. (2.13), and $D\{\underline{e}\}$ denotes the dispersion of the unmodeled phase components, i.e., the residual reference phase caused by the orbit errors. (It is assumed that the unwrapped phase is available, and a least-squares estimation is performed.) The variance of the estimated linear displacement rate follows as

$$\hat{\underline{\sigma}}_\alpha^2 = \frac{1}{(-\frac{4\pi}{\lambda})^2 \sum_{k=1}^{10} (T^k)^2}. \tag{2.32}$$

Assuming $\lambda = 56.6$ mm, the wavelength used by ERS, and $T^k = k - 5.5$ years, for $k = 1, \ldots, 10$, it follows that $\hat{\underline{\sigma}}_\alpha^2 = 0.25$ mm^2/y^2. If it is assumed 50 interferograms are available, equally spaced in time over this nine year period, then $\hat{\underline{\sigma}}_\alpha^2 = 0.06$ mm^2/y^2. Depending on the application, this error cannot be neglected. It is easily derived that the error on the estimated displacement rates caused by orbit inaccuracies is a ramp just as the orbit error phase is. If it can be assumed that the displacement field does not contain a trend, or one is not interested in this component, the phase data can be detrended. This may be necessary for a few interferograms anyway, since orbit data is not always precise enough. For example, precise orbits may not yet be available for very recent acquisitions, or the quality of orbit data is degraded due to orbit maneuvers (causing problems for orbit propagation software). Moreover, the altimeter on board of the ERS–1 satellite was switched off June 3$^{\text{rd}}$, 1996, which severely degraded the quality of the estimated orbits after this date.

Finally, the noise term is caused by, among others, thermal noise, quantization of the signal in the D/A converter, approximations made during the processing, and coregistration errors. The phase noise at the considered pixels is assumed to have a zero-mean normal distribution.

The phase components induced by elevation with respect to the reference surface (DEM error), displacement, and the sub-pixel position are considered part of the functional model, whereas components due to inaccurate knowledge of the sensor position, atmospheric signal, and other effects are considered part of the stochastic model. If all acquisitions have the same radar frequency and only slightly different Doppler centroid frequencies, the functional model is written as

$$\begin{aligned} E\{\phi\} &= W\{\phi_{\text{topo}} + \phi_{\text{defo}}\}, \\ &= W\{\beta \cdot \Delta h - \frac{4\pi}{\lambda} \sum_{d=1}^{D} \alpha_d \cdot p_d\}. \end{aligned} \tag{2.33}$$

In case of a linear displacement model with displacement rate α, this reduces to the functional model that is typically used in the PS technique (Ferretti et al., 2000a)

$$E\{\phi\} = W\{\beta \cdot \Delta h - \frac{4\pi}{\lambda} T \cdot \alpha\}. \tag{2.34}$$

The stochastic model of interferometric observations is described in section 2.2.2. It is not used in the reference PS technique, which assumes equal weights for all observations, and no correlation between them.

2.2.2 Stochastic model

In this section the variance-covariance matrix (vc-matrix) of the original phase observations in the SLC images is propagated to that of the estimated parameters. The atmospheric phase is considered part of the stochastic model. It is assumed for this derivation that the unwrapped phase is available. The original phase observations (at H positions) in SLC image k are ordered in a vector as

$$\underline{\varphi}^k = \begin{bmatrix} \underline{\varphi}^k_1 \\ \underline{\varphi}^k_2 \\ \vdots \\ \underline{\varphi}^k_H \end{bmatrix}, \tag{2.35}$$

where $k = 0, \ldots, K$, and $k = 0$ denotes the master acquisition. The underlining indicates a stochastic variable. The dispersion of the phase observations is given by

$$\begin{aligned} D\{\underline{\varphi}^k\} &= Q_{\text{slc}^k} \\ &= Q_{\text{noise}^k} + Q_{\text{atmo}^k} \\ &= \begin{bmatrix} \sigma^2_{\text{noise}^k_1} & & \\ & \ddots & \\ & & \sigma^2_{\text{noise}^k_H} \end{bmatrix} + \begin{bmatrix} \sigma^2_{\text{atmo}^k}(0) & \sigma_{\text{atmo}^k}(l_{1,2}) & \sigma_{\text{atmo}^k}(l_{1,3}) & \cdots \\ . & \sigma^2_{\text{atmo}^k}(0) & \sigma_{\text{atmo}^k}(l_{2,3}) & \cdots \\ . & . & \ddots & \\ . & . & . & \sigma^2_{\text{atmo}^k}(0) \end{bmatrix}. \end{aligned} \tag{2.36}$$

Matrix Q_{noise^k} describes thermal noise, processing noise, etc.; the noise is assumed to be white. Q_{atmo^k} is the vc-matrix that describes the atmospheric state at acquisition k. The atmospheric signal S^k at the time of acquisition k is described by a probability density function having $E\{S^k_{a,b}\} = 0$, $D\{S^k_{a,b}\} = \sigma^2_{\text{atmo}^k}(l_{a,b})$, where $l_{a,b}$ is the distance between points a and b. Fig. 2.7 shows a covariance function that could be used to describe the residual phase in the SLC images. An empirical covariance function could be used to fill this matrix, for example one that is initialized using GPS measurements taken at the time of the radar acquisition, using the model derived in (Hanssen, 2001). This covariance function could also be parameterized by an analytical covariance function, of which its parameters are estimated using the residuals after estimation of DEM error and displacement. Numerical simulations

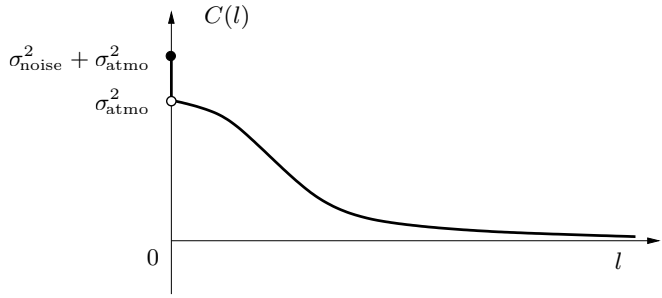


Fig. 2.7: Example covariance function for residual phase in an SLC image. The nugget at $l=0$ corresponds to the uncorrelated noise $\sigma^2_{\text{noise}^k}$.

using fractal surfaces with fractal dimension 2.67 showed that the empirical covariance function for the atmospheric signal can be modeled approximately by an exponential covariance function

$$C_{\text{atmo}}(l) = \sigma^2_{\text{atmo}} \exp(-l^2 w^2). \tag{2.37}$$

Argument l is the distance between two points in kilometers. The parameter w is related to the correlation length of the atmospheric signal as $l_c = \sqrt{\ln(2)}/w$ (which is a function of the fractal dimension), and σ^2_{atmo} is the variance of the atmospheric signal. It follows that the covariance function for the phase observations in an SLC image is written as

$$C(l) = \sigma^2_{\text{noise}} \delta_{l,0} + \sigma^2_{\text{atmo}} \exp(-l^2 w^2), \tag{2.38}$$

where $\delta_{l,0}$ is the Kronecker symbol[2]. This covariance function guarantees a positive-definite vc-matrix. If all phase observations at the points in the SLC images are collected in a $(K+1)H \times 1$ vector as

$$\underline{\varphi} = \begin{bmatrix} \underline{\varphi}^0 \\ \underline{\varphi}^1 \\ \vdots \\ \underline{\varphi}^K \end{bmatrix}, \tag{2.39}$$

then the phase $\underline{\phi}$ in the K interferograms is obtained by subtracting the phase at each point of the slaves $k = 1, \ldots, K$ from that of the master $k = 0$. This can be written in matrix notation as

$$\underline{\phi} = \Lambda \underline{\varphi}, \quad \text{where } \Lambda = [e_K, -I_K] \otimes I_H = \begin{bmatrix} I_H & -I_H & & \\ \vdots & & \ddots & \\ I_H & & & -I_H \end{bmatrix}. \tag{2.40}$$

[2] $\delta_{l,m} = 1$ for $l = m$, 0 otherwise.

The vector e_K is defined as $e_K = (1, 1, \ldots, 1)^*$, I_K is a $K \times K$ identity matrix, and $\otimes$ is the Kronecker tensor product. The estimation of the unknown parameters is performed with respect to an arbitrary reference point. Without loss of generality it can be assumed that the reference point is the first point. The vector of double-difference observations is then computed as

$$\underline{\psi} = \Omega \underline{\phi}, \quad \text{where } \Omega = I_K \otimes [e_{H-1}, -I_{H-1}] = \begin{bmatrix} e_{H-1} - I_{H-1} & & \\ & e_{H-1} - I_{H-1} & \\ & & \ddots \end{bmatrix}. \tag{2.41}$$

After this operation, the phase in vector $\underline{\psi}$ is still ordered per interferogram, i.e., for interferogram 1 all phase differences, for interferogram 2 all phase differences, etc. It is more convenient to order the phase per point instead of per interferogram. This can be achieved by multiplication with a square permutation matrix P with dimension $K(H-1)$, i.e.,

$$\underline{y} = P \underline{\psi}. \tag{2.42}$$

After this permutation, the order of the elements in vector $\underline{y}$ is that first all interferometric phases for the first arc are given, then for the second arc, etc. Since

$$\underline{y} = P \Omega \Lambda \, \underline{\varphi}, \tag{2.43}$$

it follows that the propagated vc-matrix for the interferometric phase differences with respect to the reference pixel is given by application of the law of propagation of variances[3] as

$$\begin{aligned} Q_{\text{ifg}} &= (P\Omega\Lambda)\, Q_{\text{slc}}\, (P\Omega\Lambda)^*, \\ &= (P\Omega\Lambda)\, Q_{\text{noise}}\, (P\Omega\Lambda)^* + (P\Omega\Lambda)\, Q_{\text{atmo}}\, (P\Omega\Lambda)^*, \end{aligned} \tag{2.44}$$

where Q_{slc}, Q_{noise}, and Q_{atmo} are the corresponding partioned matrices with dimension $(K+1)H \times (K+1)H$, e.g.,

$$Q_{\text{slc}} = \begin{bmatrix} Q_{\text{slc}^0} & & \\ & \ddots & \\ & & Q_{\text{slc}^K} \end{bmatrix}. \tag{2.45}$$

It is assumed that the noise is uncorrelated between the various SLC images. If a single large design matrix B with dimension $K(H-1) \times 2(H-1)$ is defined for the least-squares estimation of all parameters, for example in the case of DEM error and linear displacement

$$B = I_{H-1} \otimes B', \quad \text{where } B' = \begin{bmatrix} \beta^1 & -\frac{4\pi}{\lambda} T^1 \\ \vdots & \vdots \\ \beta^K & -\frac{4\pi}{\lambda} T^K \end{bmatrix}, \tag{2.46}$$

[3] $v = Uu \;\rightarrow\; Q_v = U Q_u U^*$

where β^k are the appropriate height-to-phase conversion factors, and T^k are the temporal baselines, see also Eqs. (2.12) and (2.13), then the full vc-matrix of all estimated parameters $\underline{\hat{b}}$ is given by

$$Q_{\hat{b}} = (B^* Q_{\text{ifg}}^{-1} B)^{-1}, \tag{2.47}$$

and the least-squares estimates for all parameters by

$$\underline{\hat{b}} = Q_{\hat{b}} B^* Q_{\text{ifg}}^{-1} \underline{y}. \tag{2.48}$$

The phase of a reference surface must be subtracted before Eq. (2.46) is valid. The subtraction of the reference phase does not affect the propagated vc-matrix, because it is considered to be a deterministic process, which is not shown here. Noise introduced by the processing, such as mis-registration of the slave images, also is not considered here. This can effectively be incorporated by increasing the noise level for the slave images. For example, if this noise is assumed to be equal to the inherent noise, this becomes $Q_{\text{noise}^k} = 2Q_{\text{noise}^0}$ for $k = 1, \ldots, K$.

Double-difference observations

Because all estimations are performed between two points, it is useful to consider this special case. According to Eq. (2.40) and Eq. (2.41), the matrix transforming the original observed phase to the interferometric double-difference is given by

$$P\Omega\Lambda = \begin{bmatrix} 1 & -1 & -1 & 1 & & & \\ 1 & -1 & 0 & 0 & -1 & 1 & \\ \vdots & & & & \ddots & \ddots & \\ 1 & -1 & \ldots & & & -1 & 1 \end{bmatrix}. \tag{2.49}$$

The permutation matrix P is the identity matrix in this case, because no re-ordering is required if only a single phase difference between two points is considered. Thus, the vc-matrix of the interferometric double-difference observations is given by

$$Q_{\text{ifg}} = 2\sigma^2_{\text{noise}^0} \begin{bmatrix} 1 & \ldots & 1 \\ 1 & \ldots & 1 \\ \vdots & & \\ 1 & \ldots & 1 \end{bmatrix} + 2\sigma^2_{\text{noise}^1} \begin{bmatrix} 1 & 0 & \ldots \\ 0 & 0 & \ldots \\ \vdots & & \\ 0 & 0 & \ldots \end{bmatrix} + 2\sigma^2_{\text{noise}^2} \begin{bmatrix} 0 & 0 & \ldots \\ 0 & 1 & \ldots \\ 0 & 0 & \ldots \\ \vdots & & \end{bmatrix} + \cdots$$
$$+ 2\big(\sigma^2_{\text{atmo}^0}(0) - \sigma_{\text{atmo}^0}(l)\big) \begin{bmatrix} 1 & \ldots & 1 \\ 1 & \ldots & 1 \\ \vdots & & \\ 1 & \ldots & 1 \end{bmatrix}$$

$$
\begin{aligned}
&+ 2\big(\sigma^2_{\text{atmo}^1}(0) - \sigma_{\text{atmo}^1}(l)\big) \begin{bmatrix} 1\,0\,\ldots \\ 0\,0\,\ldots \\ \vdots \\ 0\,0\,\ldots \end{bmatrix} + \ldots \\
&= 2\big(\sigma^2_{\text{noise}^0} + \sigma^2_{\text{atmo}^0}(0) - \sigma_{\text{atmo}^0}(l)\big)\, E_K \\
&\quad + \sum_{k=1}^{K} 2\big(\sigma^2_{\text{noise}^k} + \sigma^2_{\text{atmo}^k}(0) - \sigma_{\text{atmo}^k}(l)\big)\, i_k i_k{}^*, \qquad (2.50)
\end{aligned}
$$

where E_K is a $K \times K$ matrix filled with ones, and i_k is a $K \times 1$ vector with a single one at position k. It is assumed here that all points in an interferogram have the same inherent noise level $\sigma^2_{\text{noise}^k}$. From Eq. (2.50) it can be clearly seen that the double-difference observations are correlated. If it is assumed for small l that $\sigma^2_{\text{atmo}^k}(0) \approx \sigma_{\text{atmo}^k}(l)$, matrix Q_{ifg} reduces to

$$
Q_{\text{ifg}} = \sum_{k=0}^{K} \sigma^2_{\text{noise}^k} Q_k, \quad \text{where } Q_k = \begin{cases} 2E_K & \text{if } k = 0 \\ 2i_k i_k{}^* & \text{if } k = 1, \ldots, K \end{cases}. \qquad (2.51)
$$

Here, the vc-matrix of the double-difference observations is written as a variance component model, using $K+1$ cofactor matrices Q_k and $K+1$ variance components $\sigma^2_{\text{noise}^k}$. If these components are assumed to be unknown, a variance component estimation technique can be used to obtain estimates for them. Variance component estimation is summarized in Appendix A.

For example, if the inherent noise σ^2_{noise} would be equal for all acquisitions, and if mis-registration causes an equal additional amount of noise in the slave images, the vc-matrix Q_{ifg} for the double-difference observations becomes

$$
Q_{\text{ifg}} = \sigma^2_{\text{noise}} \begin{bmatrix} 6\,2 & 2 & 2 & \ldots 2 \\ 2\,6 & 2 & 2 & \ldots 2 \\ \vdots & \ddots & & \\ & & & \\ & & & \\ 2\,2 & 2 & \ldots & 2\;\; 6 \end{bmatrix}. \qquad (2.52)
$$

An analytical expression for the elements of the inverse of matrix Eq. (2.51) was not found. However, for the special case of Eq. (2.52) the inverse was found to be

$$
Q^{-1}_{\text{ifg}} = \frac{1}{(4K+8)\,\sigma^2_{\text{noise}}} \begin{bmatrix} K+1 & -1 & -1 & -1 \ldots & -1 \\ -1 & K+1 & -1 & -1 \ldots & -1 \\ \vdots & & \ddots & & \\ & & & & \\ & & & & \\ -1 & -1 & -1 & \ldots -1 & K+1 \end{bmatrix}. \qquad (2.53)
$$

This expression shows that the norm $e^* Q_{\text{ifg}}^{-1} e$ which is minimized using a least-squares approach is less sensitive to a bias in the double-differenced interferometric phase observations than a diagonal vc-matrix would be. Thus, if this stochastic model is used there is less need to include a parameter for the average atmospheric phase $\bar{S}$ in the functional model, as was suggested in section 2.2.1 (page 23).

3

The Integer Least-Squares Estimator

In the Persistent Scatterer Interferometry there is a need for an efficient and reliable non-linear parameter inversion algorithm. Present algorithms make use of a direct search of the solution space (ensemble coherence maximization), treating the observations as deterministic and equally weighted. Moreover, they do not provide a description of the quality of the estimated parameters.

In section 3.1 the integer least-squares estimator is introduced, which has the highest probability of correct integer estimation for problems with a multivariate normal distribution. For the application to Persistent Scatterer Interferometry the algorithm needs to be adapted, which is described in section 3.2. Numerical aspects of the implementation are described in section 3.3. The performance is demonstrated in section 3.4 using a simulation.

3.1 The LAMBDA method

The Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method was developed for fast GPS double-difference integer ambiguity estimation (Teunissen, 1995). In that problem, the carrier phase of the GPS signal is used to obtain a highly accurate distance measurement. The integer number of cycles is unknown, but the a priori knowledge of the integer nature of these ambiguities is used in the estimation to strengthen the solution. Aside from being a fast method, it also is the best, in the sense that it gives the highest probability of correct integer estimation (ambiguity success rate) for problems with a multivariate normal distribution, see (Teunissen et al., 1995c). For this problem, no direct inversion exists. The LAMBDA method makes use of a sequential conditional least-squares search, based on transformed ambiguities. It was first introduced in (Teunissen, 1994) and discussed in detail in (Teunissen, 1995). Source code is available at (Delft University of Technology, 2005). The following is an overview of the steps involved in integer ambiguity estimation to make this work self-contained. Point of departure is the linearized system of observation equations

$$\underline{y} = Aa + Bb + \underline{e}, \tag{3.1}$$

where:

- $\underline{y}$ is the vector of measurements (observed minus computed double-difference carrier-phase and code measurements in the case of GPS). The underlining indicates a vector of stochastic variables.
- a is the vector of integer-valued unknown ambiguities.
- b is the vector of real-valued unknowns for the parameters of interest. For GPS these are the three baseline components. Because the system of equations is linearized for GPS, this vector consists of increments with respect to a priori values or the previous iteration.
- A, B are the design matrices for the ambiguity terms and baseline components, respectively.
- $\underline{e}$ is the vector of measurement noise and unmodeled errors.

Since the estimation criterion is based on the principle of least-squares, the estimates for the unknown parameters of Eq. (3.1) follow from solving the minimization problem

$$\min_{a,b} \|\underline{y} - Aa - Bb\|^2_{Q_y} \quad \text{subject to } a \in \mathbb{Z},\ b \in \mathbb{R}, \tag{3.2}$$

where $\|.\|^2_{Q_y} = (.)^* Q_y^{-1} (.)$ and Q_y is the variance-covariance matrix of the observables (the asterisk denotes the transposition). This minimization problem is referred to as an integer least-squares problem (Teunissen, 1994). It is a constrained least-squares problem due to the integer constraint $a \in \mathbb{Z}$. The solution of the integer least-squares problem will be denoted as $\breve{\underline{a}}$ and $\breve{\underline{b}}$. The solution of the corresponding unconstrained least-squares problem will be denoted as $\hat{\underline{a}}$ and $\hat{\underline{b}}$. The estimates $\hat{\underline{a}}$ and $\hat{\underline{b}}$ are referred to as the "float solution", and the estimates $\breve{\underline{a}}$ and $\breve{\underline{b}}$ as the "fixed solution". The approach taken with the LAMBDA method, is to re-parameterize the integer least-squares problem such that an equivalent problem is obtained, but one that is much easier to solve. It consists of two steps. First, an ambiguity transformation Z^* is constructed that tries to decorrelate the ambiguities. This transformation increases the efficiency of the search for the (transformed) integer ambiguities that minimize Eq. (3.2). In the construction of Z^*, use is made of integer approximations to conditional least-squares transformations. The ambiguity transformation allows one to transform the original ambiguities, their least-squares estimates and their corresponding variance-covariance matrix as

$$z = Z^* a, \quad \hat{\underline{z}} = Z^* \hat{\underline{a}}, \quad Q_{\hat{z}} = Z^* Q_{\hat{a}} Z. \tag{3.3}$$

The computation of the integer minimizers $\breve{\underline{z}}$ is performed in the second step of the LAMBDA method. It follows from solving

$$\min_z \|\hat{\underline{z}} - z\|^2_{Q_{\hat{z}}} \quad \text{subject to } z \in \mathbb{Z}. \tag{3.4}$$

Since matrix Z consists of integers only and is volume preserving, the obtained solution also minimizes $\underline{\hat{a}}-a$ (Teunissen et al., 1995a). That is, the ambiguities that are of interest can be obtained by solving Eq. (3.4). The solution is obtained by means of a search using a set of bounds for the transformed ambiguities (Teunissen et al., 1995b). If the ambiguities would be totally decorrelated, the integer ambiguities would be given by means of a simple rounding of the float ambiguities, since that would minimize Eq. (3.4). However, this simple rounding scheme does not produce the required integer least-squares estimates when matrix $Q_{\hat{z}}$ is non-diagonal. It was shown by Teunissen (1994) that minimizing the objective function Eq. (3.4) is identical to minimizing

$$\min_{z_i \in \mathbb{Z}} \left(\frac{(\underline{\hat{z}}_1 - z_1)^2}{\sigma_1^2} + \frac{(\underline{\hat{z}}_{2|1} - z_2)^2}{\sigma_{2|1}^2} + \ldots + \frac{(\underline{\hat{z}}_{n|n-1} - z_n)^2}{\sigma_{n|n-1}^2} \right), \tag{3.5}$$

which makes use of a sequential conditional least-squares adjustment. The estimate $\underline{\hat{z}}_{i|I}$ is the least-squares estimate of z_i, conditioned on $z_j, j{=}1, \ldots, I$. In order to solve Eq. (3.5), a search is performed for the integer least-squares ambiguities, based on the set of bounds

$$(\underline{\hat{z}}_{i|I} - z_i)^2 \leq l_i \, \sigma_{i|I}^2 \, \chi^2, \quad \text{for } i = 1, \ldots, n \tag{3.6}$$

where

$$l_i = 1 - \frac{\chi_{i-1}^2}{\chi^2}, \quad \text{subject to } \chi_{i-1}^2 = \sum_{j=1}^{i-1} \frac{(\underline{\hat{z}}_{j|J} - z_j)^2}{\sigma_{j|J}^2}. \tag{3.7}$$

In order to perform the search, first a value for χ^2 needs to be determined, such that it is guaranteed that the search space contains at least one solution. Since the search takes place over the ambiguities, it will take longer when there are more ambiguities. Once the integer least-squares vector $\underline{\check{z}}$ is found, the corresponding integer least-squares vector of the original ambiguities can be found by invoking $\underline{\check{a}} = Z^{*-1}\underline{\check{z}}$.

3.2 Application of the LAMBDA method

In the case of radar interferometry, the observed wrapped phase (difference), $\underline{\phi}^k$, is unwrapped as

$$\underline{\Phi}^k = \underline{\phi}^k + 2\pi \cdot a^k, \quad \text{subject to } a^k \in \mathbb{Z}, \tag{3.8}$$

with integer ambiguity a^k for interferogram k. Therefore, application of the LAMBDA method seems straightforward, and the existence of this method was already pointed out by Hanssen and Ferretti (2002); Hanssen et al. (2001), and by Bianchi (2003), who performed a first simplified evaluation of a PS integer least-squares estimator using simulated data. However, the

problems of GPS and PS are not identical. The major difference is that for radar interferometry the problem is inherently under-determined, since each observation has an unknown ambiguity that needs to be estimated, aside from the parameters of interest, see Eq. (3.8). The solution to this problem can only be obtained by using the fact that the ambiguities are integers, while for GPS a (less precise) solution can also be obtained without using this information, using the code observations. A more practical constraint that must be kept in mind is that the number of estimations that need to be performed is much larger in the case of PS than it is for GPS, because the number of points is much larger. Moreover, the number of acquisitions, i.e., the number of ambiguities that need to be estimated, can be significantly larger than for GPS. The algorithm developed for this research is the first with convincing performance on real data (Kampes and Hanssen, 2004).

The model for the unwrapped phase in interferogram k is given in Eq. (2.11), repeated here for convenience, see also section 2.1.2

$$\underline{\Phi}_x^k = \phi_{x,\text{topo}}^k + \phi_{x,\text{defo}}^k + \phi_{x,\text{obj}}^k + \phi_{x,\text{atmo}}^k + \phi_{x,\text{orbit}}^k + \phi_{x,\text{noise}}^k. \tag{3.9}$$

The functional model for the phase difference $\Phi_{x,y}^k = \Phi_y^k - \Phi_x^k$ between two points x and y, is given by

$$E\{\underline{\Phi}_{x,y}^k\} = \beta_x^k \cdot \Delta h_{x,y} - \frac{4\pi}{\lambda} \sum_{d=1}^{D} \alpha_d(x,y) \cdot p_d(k) + \frac{2\pi}{v} f_{x,\text{dc}}^{k,m} \cdot \xi_{x,y}. \tag{3.10}$$

The atmospheric, orbit, and noise phase differences are lumped in a new random variable $\underline{e}$ with expectation $E\{\underline{e}\} = 0$. In matrix notation this system of observation equations is written as

$$E\{\begin{bmatrix} \underline{\phi}^1 \\ \underline{\phi}^2 \\ \vdots \\ \underline{\phi}^K \end{bmatrix}\} = \begin{bmatrix} -2\pi & & \\ & -2\pi & \\ & & \ddots \\ & & -2\pi \end{bmatrix} \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^K \end{bmatrix} + \begin{bmatrix} \beta_x^1 & p_1(1)..p_D(1) & \frac{2\pi}{v} f_{\text{dc}}^{1,m} \\ \beta_x^2 & p_1(2)..p_D(2) & \frac{2\pi}{v} f_{\text{dc}}^{2,m} \\ \vdots & & \\ \beta_x^K & p_1(K)..p_D(K) & \frac{2\pi}{v} f_{\text{dc}}^{K,m} \end{bmatrix} \begin{bmatrix} \Delta h \\ \alpha_1 \\ \vdots \\ \alpha_D \\ \xi \end{bmatrix} \tag{3.11}$$

The index $\{.\}_{x,y}$ is dropped, but note that this system of equations refers to the phase differences between two points. The basic task is to estimate the K integer ambiguities and the $2+D$ real-valued parameters from the K observed wrapped phase values. It is assumed here that there is significant variation in the Doppler centroid frequency. If this is not the case (or when the azimuth sub-pixel positions are already estimated using a point target analysis in the amplitude images), then the azimuth sub-pixel position needs not to be estimated here, leaving $1+D$ real-valued parameters. To solve this system of equations, additional constraints have to be introduced. As in (Bianchi, 2003; Hanssen et al., 2001), pseudo-observations $\underline{y}_2$ are used to achieve this

$$E\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}\} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} a + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} b, \qquad D\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}\} = \begin{bmatrix} Q_{y_1} & 0 \\ 0 & Q_{y_2} \end{bmatrix}. \tag{3.12}$$

Matrices $\underline{y}_1$, A_1, and B_1 are defined in Eq. (3.11). A_2 is a zero matrix $O_{D+3\times K}$, and B_2 is an identity matrix I_{D+3}. The value of the pseudo-observations is chosen as $\underline{y}_2 = 0$. Operator $D\{.\}$ denotes the dispersion of the observations. Unlike in GPS, the dispersion of the residual phase is not well known a priori, since it is not guaranteed that the points are coherent in time (because they are selected based on the amplitude dispersion index), and follow the displacement model that is used. Matrix Q_{y_1} follows from propagation of the variance-covariance matrices of the original phase observations in the SLC images, see section 2.2.2. The dispersion of the pseudo-observations follows from an appropriate, conservative, a priori standard deviation of the unknown parameters. Reasonable values are for example $\sigma = 20$ m for the DEM error difference, $\sigma = 20$ mm/y for the linear displacement rate difference, and $\sigma = 2$ m for the azimuth sub-pixel position difference (for sensors with azimuth resolution comparable to that of ERS). This augmented system of equations can be written symbolically similar to Eq. (3.1) as

$$E\{\underline{y}\} = Aa + Bb, \qquad D\{\underline{y}\} = Q_y. \tag{3.13}$$

Since all ambiguities are relative to each other, an arbitrary ambiguity could be fixed to zero, as done in (Kampes and Hanssen, 2004). However, in order to use all available a priori information, it is prefered here to introduce a pseudo-observation for each unknown parameter. The float solution for the ambiguities $\hat{\underline{a}}$ can be obtained using the solution of a partioned model as (Teunissen, 2000a)

$$Q_{\hat{a}} = \left(\bar{A}^* Q_y^{-1} \bar{A}\right)^{-1} \tag{3.14}$$

$$\hat{\underline{a}} = Q_{\hat{a}} \bar{A}^* Q_y^{-1}\, \underline{y}, \tag{3.15}$$

where $\bar{A} = P_B^{\perp} A$ and $P_B^{\perp} = I - B(B^* Q_y^{-1} B)^{-1} B^* Q_y^{-1}$. (A more efficient way is described in section 3.3.2.) This float solution is then transformed using Eq. (3.3), yielding $Q_{\hat{z}}$ and $\hat{\underline{z}}$. Then, a search is performed based on the bounds of Eq. (3.6), which yields the integer least-squares estimates for the ambiguities. These integer estimates are then used to compute the "fixed" solution $\check{\underline{b}}$ for the unknown real-valued parameters. This can be achieved by substitution of the estimated integer ambiguities $\check{\underline{a}}$ in Eq. (3.11) and moving this term to the left hand side, which leaves an ordinary unconstrained least-squares problem. That is, the observed phase values are unwrapped by

$$\check{\underline{y}}_1 = \underline{y}_1 - A_1\, \check{\underline{a}}, \tag{3.16}$$

leaving the systems of equations

$$E\{\check{\underline{y}}_1\} = B_1\, b, \qquad D\{\check{\underline{y}}_1\} = Q_{y_1}. \tag{3.17}$$

The least-squares estimator for the float parameters is thus given by

$$Q_{\hat{b}} = \left({B_1}^* Q_{y_1}^{-1} B_1\right)^{-1}, \tag{3.18}$$

$$\check{\underline{b}} = Q_{\hat{b}} {B_1}^* Q_{y_1}^{-1}\, \check{\underline{y}}_1. \tag{3.19}$$

Matrix $Q_{\hat{b}}$ is the full variance-covariance matrix that describes the precision of the estimated float parameters.

3.3 Computational aspects

In the case that each pixel in the interferograms needs to be estimated, the number of estimations that must be performed is approximately twenty million for an area of 20×20 km^2. If, for example, each estimation took one second, then the total computation time would be 231 days. In our implementation, a large number of points is discarded based on the fact that point scatterers are expected to have a relatively high amplitude, but a fast implementation is still extremely important. This is true even when a multi-CPU system would be used. Speed is considered more important than achieving the highest possible success rate. The points that are estimated lie in a "network of opportunity" anyway, i.e., their location is a result of how the radar signal happened to backscatter from the target to the satellite.

3.3.1 The bootstrap estimator

The integer bootstrapped estimator is used to determine a bound χ^2 on the search space, such that at least one solution is contained. This estimator takes some of the correlation between the ambiguities into account, but does not search the full hyper-ellipsoid up to all bounds. It results from a sequential conditioned least-squares adjustment and it is computed as follows (Teunissen et al., 1995b). If K ambiguities are available, the first ambiguity $\hat{\underline{z}}_1$ is rounded to its nearest integer. Having obtained the integer value of the first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected on the basis of their correlation with the first ambiguity. Subsequently, the second, now corrected, real-valued ambiguity is rounded to its nearest integer. Having obtained the value of the second ambiguity, the real-valued estimates of all remaining $K-2$ ambiguities are again corrected, but now on the basis of their correlation with the second ambiguity. This process of rounding and correcting is continued until all ambiguities are taken care of. The success rate for the bootstrap method can be computed as (Hanssen et al., 2001)

$$P(\hat{\underline{z}}{=}z) = \prod_{i=1}^{n} \left[2\Upsilon\Big(\frac{1}{2\sigma_{i|I}}\Big) - 1 \right], \tag{3.20}$$

where

$$\Upsilon(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2}} \exp\Big(-\frac{v^2}{2}\Big)\, dv. \tag{3.21}$$

The success rate can thus be computed in advance, using the baselines of the acquired images, and assuming a known data noise level. It is cumbersome to compute this probability for the LAMBDA method, but it can be shown

that it outperforms the bootstrapped method, which thus can be regarded as a lower bound for it (Teunissen, 1999).

After running the bootstrap estimator as described above, it is run slightly modified for $k = 1, \ldots, K$ extra times, in order to increase the success rate. During each extra run, the normal bootstrap estimate for the fixed ambiguity k is replaced by an integer either one higher or one lower, based on the difference between the original float value and the corrected float value. The remaining ambiguities are then computed based on their correlation with the previously fixed ones, as is done with the unmodified bootstrap method. This is similar to the search of the ambiguity solution space that is performed with the LAMBDA method, but more limited, and therefore faster. The final estimate from this extended bootstrap method is the one that has the smallest χ^2 according to Eq. (3.4). The success rate of this extended bootstrap method cannot be analytically quantified, but it is at least equal to that of Eq. (3.20), since the normal bootstrap estimate is included in the extended method.

Estimation strategy

After an estimate is obtained using the (extended) bootstrap method, the search of the LAMBDA method is performed with the bound χ^2 from the bootstrap method. Since this search can take an extremely long (CPU) time—particularly when the data contain more noise than described by the a priori variance-covariance matrix Q_{y_1}—it is discontinued after a maximum number of ambiguities are searched. Typically, K^3 is used for the maximum loop count. If the search is stopped and no candidate is found, the estimated ambiguities that resulted from the bootstrap method are used. If a solution is obtained using the LAMBDA search, the estimate is taken that has the smallest norm according to Eq. (3.4). Note that if a well-fitting solution cannot be found using the bootstrapped estimation, the search for the integer least-squares estimate also takes very long. Since the time required by the (extended) bootstrap estimator does not depend on the quality of the input data, alternatively, one could try to regularize the system with several sets of randomized pseudo-observations, each time using the bootstrap estimator. Then the search to obtain the integer least-squares estimate can optionally be performed for the solution with the smallest χ^2.

3.3.2 Reduction of the numerical complexity

Since the under-determined system of equations Eq. (3.11) is constrained with $D+2$ pseudo-observations (i.e., exactly the number of real-valued parameters), the constrained system Eq. (3.12) is exactly determined. This means that the float solution for the real-valued parameters must be equal to the added pseudo-observations. Since there is no redundancy in this case, the float solution for the ambiguities can be computed efficiently by

$$\hat{\underline{a}} = \frac{\underline{y}_1}{-2\pi}, \tag{3.22}$$

exploiting the diagonal structure of the matrix A_1 and the fact that the pseudo-observations are equal to zero. This expression avoids the matrix multiplication in Eq. (3.15). Since the observations in vector $\underline{y}_1 \in [-\pi, \pi)$, the float solution for the ambiguities $\hat{\underline{a}} \in [-0.5, 0.5)$. The float solution for all parameters is thus biased due to the introduction of the pseudo-observations. However, the fixed solution is not biased, as long as the correct integer ambiguities are found during the search. Due to the diagonal structure of A_1, the matrix multiplication in Eq. (3.16) can also be avoided. To obtain the unwrapped data (the "fixed" observations $\check{\underline{y}}_1$) expression Eq. (3.8) can be used instead

$$\check{\underline{y}}_1 = \underline{y}_1 + 2\pi\, \check{\underline{a}}. \tag{3.23}$$

Aside from this, a large amount of computation time can be saved by observing that the height conversion factor scales the same for all interferograms as function of range coordinate. This implies that $\Delta h' = \frac{\bar{\beta}}{\beta}\Delta h$ can be substituted in Eq. (3.11), with $\bar{\beta}$ the average height conversion factor for the whole interferogram. After the estimation of $\Delta h'$ the correct DEM error is obtained by rescaling it with this factor. By this substitution it is achieved that matrices A and B of Eq. (3.13) are the same for all estimations between points. This means that the transformation matrix Z needs to be computed only once, as well as the decorrelation of $Q_{\hat{a}}$, see Eq. (3.14), according to Eq. (3.3). Also, the $(D+2)\times K$ projection matrix $(B^* Q_y^{-1} B)^{-1} B^* Q_y^{-1}$ to obtain the final float solution for the parameters of interest can be computed in advance and re-used for each estimation, see Eq. (3.19).

3.3.3 Choice of the base functions

An appropriate set of base functions of the displacement model obviously depends on the signal that is expected. For example, if linear displacement in time is expected to be a good approximation of the displacement, a single base function using the temporal baseline is optimal, i.e.,

$$p_1(T) = T, \quad T \in [T_{\min}, T_{\max}]. \tag{3.24}$$

If it is known that an event happened at a certain epoch T_{i}, or when it is known that a linear displacement function does not likely represent the actual displacement over a long time, then a piecewise linear function may be a better choice. Such a displacement model can be created by defining a set of base functions as

$$p_1(T) = T, \quad T \in [T_{\min}, T_{\mathrm{i}}), \tag{3.25}$$
$$p_2(T) = 1, \quad T \in [T_{\min}, T_{\mathrm{i}}), \tag{3.26}$$
$$p_3(T) = T, \quad T \in [T_{\mathrm{i}}, T_{\max}], \tag{3.27}$$
$$p_4(T) = 1, \quad T \in [T_{\mathrm{i}}, T_{\max}]. \tag{3.28}$$

If not much information is available on the displacement signal, and a relatively large number of images are available, a good choice may be to use a piecewise linear function for, say, every two years of data. Higher order algebraic polynomials could be used instead of a piecewise linear function, if that is more appropriate. Moreover, trigonometric polynomials could be used to model periodical displacement. For example, the base functions to model seasonal displacement are given by

$$\begin{aligned} p_1(T) &= \sin(2\pi T), \\ p_2(T) &= \cos(2\pi T) - 1. \end{aligned} \tag{3.29}$$

If desired, a third base function for linear displacement could be added to this model. Finally, dedicated base functions could also be created, for example, by decomposing displacement that is observed with GPS or leveling into principal components.

However, not any choice of base functions is appropriate. It may be impossible to use a certain base function due to inadequate temporal sampling. For example, a piecewise linear function cannot be estimated in a domain where there is no data. Moreover, correlation between estimated parameters may prevent significant estimation. This depends on the distribution of the images in time, space, and Doppler frequency. For example, in the extreme case where the perpendicular baseline is a linear function of time, the DEM error is fully correlated with time-linear displacement, and both cannot be estimated. A measure for the correlation between estimated parameters is the cross-correlation coefficient that can be computed after selection of the base functions using Eq. (3.18). Furthermore, the statistical significance of an estimated parameter can be negligible. For example, for the DEM error this would be the case when all interferograms would have an almost zero baseline. Finally, sudden discontinues displacement larger than half the wavelength can never be detected due to the wrapped nature of the radar observations.

3.4 Validation

In order to validate the developed combined bootstrapped/integer least-squares estimator, and to quantify the processing time required for this algorithm, a large number of simulations are performed. For the simulated data, the number of available interferograms is set to $K = 10, 11, ..., 60$. The K acquisition times and perpendicular baselines are randomly selected from the configuration of a real ERS data set of the area of Berlin. The acquisition times are between May 1992 and November 2000, and the largest possible perpendicular baseline between the images is 2100 meters. For each K, the random selection of the baselines is repeated 10 times, each time applying the retrieval algorithm 100 times. This is done to reduce the dependency of the success rate on the actual baseline configuration, which is of particular

importance for small K. The second variable in the simulations is the amount of normally distributed noise that is added to the simulated input. The standard deviation of the noise $\underline{e}$ is set to $\sigma = 20, 30, 40, 50$ degrees. After the addition of the noise, the simulated phase is wrapped into the interval $[-\pi, \pi)$. In total, 204 different simulation scenarios are evaluated, for varying K and $\underline{e}$, where for each scenario 100 input sets are simulated. A linear displacement rate with a superimposed seasonal component is modeled using three base functions

$$\begin{aligned} p_1(T) &= T, \\ p_2(T) &= \sin(2\pi T), \\ p_3(T) &= \cos(2\pi T) - 1. \end{aligned} \tag{3.30}$$

The unwrapped model phase is then computed using the forward model $\Phi = Bb$, where the parameters are randomly simulated with standard deviations $\sigma_{\Delta h} = 20$ m, $\sigma_{\alpha_1} = 20$ mm/y, $\sigma_{\alpha_2} = 15$ mm, $\sigma_{\alpha_3} = 15$ mm. The standard deviation of the pseudo-observations, Q_{y_2}, used to retrieve the input is set to $\sigma_{\Delta h} = 40$ m, $\sigma_{\alpha_1} = 40$ mm/y, $\sigma_{\alpha_2} = 20$ mm, $\sigma_{\alpha_3} = 20$ mm, and the a priori standard deviation assumed for all interferometric phase differences Q_{y_1} is set to 50° in all scenarios, which is a conservative estimate.

Fig. 3.1 shows the individual CPU times required for the extended bootstrap method and for the integer least-squares search for all simulations. IDL[1] is used as programming language, running on a SUN workstation utilizing a single 750 MHz UltraSPARC-III CPU. Using C or FORTRAN codes would likely increase the speed by a factor of, maximally, ten. The reported CPU times originate from the IDL profiler. The time for the extended bootstrap method is $\mathrm{O}(K^2)$, since always K bootstraps are performed over the $K-1$ ambiguities. The time required for the integer least-squares search depends on both the quality and amount of data. For a low noise level, the correct ambiguities are found extremely quickly. This is caused also by the small search bound χ^2 that is returned from the bootstrap estimator. The computation time increases with an increasing noise level. The reason is that the search is performed for a solution that is in correspondence with the a priori precision, and in order to find such a solution, the bounds for the search of the hyper-ellipsoid get larger. If the maximum loop count would not be introduced, the computation time for the least-squares search would get extremely large for noisy data, and the method would become impractical.

Fig. 3.2 gives an overview of the success rate for all the simulations. The individual success rate for the bootstrap and integer least-squares method is not shown, since they have to be computed in all cases, and the combined success rate is always the highest. Only when the integer least-squares search is discontinued (using the maximum loop counter, which particularly occurs for higher noise levels), the success rate of the bootstrap method is sometimes larger than that of the integer least-squares estimator. An estimation is

[1] version 5.1

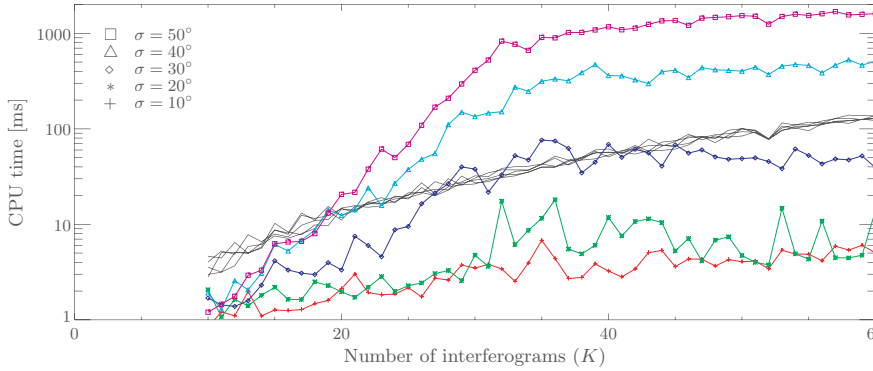


Fig. 3.1: CPU time required by the extended bootstrap and integer least-squares estimator as function of K and for different noise levels. The bootstrap method is represented by the bold solid lines for all noise levels; the computation time only depends on K. For the least-squares estimator the required computation time increases with increasing noise level.

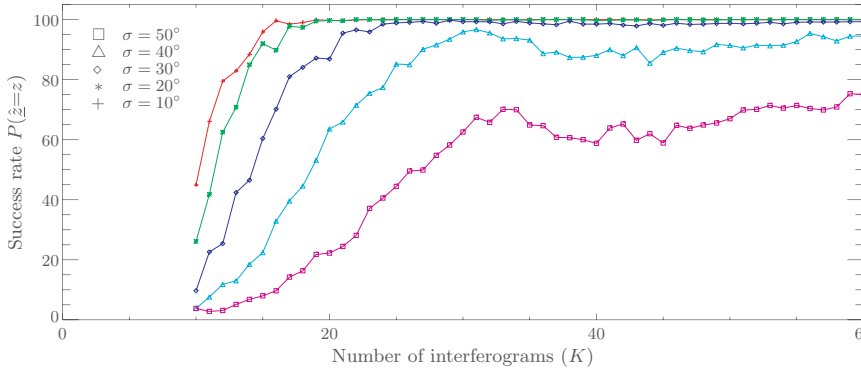


Fig. 3.2: Success rate of the combined bootstrap/integer least-squares method as function of K for different noise levels.

considered to be successful if all ambiguities are correctly estimated, and also if only a single ambiguity is estimated incorrectly by one cycle. The theoretic success rate is computed with Eq. (3.20), but it is not plotted in Fig 3.2. Since the a priori standard deviation of the noise on the observations is set to 50° in all simulations, while the actual standard deviation used to simulate the noise is lower in most cases, the theoretical success rate does not correspond very well with the one obtained in practice. Moreover, Eq. (3.20) is valid for the unmodified bootstrap method, while here a series of slightly altered bootstraps is used, which increases the probability of finding the correct ambiguities.

It can be observed that the success rate is very high for small noise levels, up to 30°, and more than 20 images. The success rate is low if there are only 10 images available, which can be explained by taking into account that 5 float parameters and 9 integer parameters (between, say, –15 and 15) are estimated in this case; there are simply too many possibilities that give a good fit in this case. Furthermore, it can be observed that the overall success rate increases with increasing number of images and decreasing noise level. The individual success rate of the extended bootstrap method is close to that of the integer least-squares search, and sometimes it is even higher, while theoretically the latter has at least the same success rate. This is caused by the maximum loop count, which is introduced in the least-squares search for speed considerations, causing the search to be discontinued at a certain point, which occurs in particular for higher noise levels. The same effect also explains why the success rate of the integer least-squares estimator decreases slightly with an increasing number of images for a constant noise level. The maximum loop count is kept constant, and thus for a smaller number of images K, the hyper-ellipsoid is searched through more completely before being discontinued. However, the least-squares search is more robust, since more possible solutions are searched for, and it is less affected by an individual noisy value.

4

The STUN Algorithm

The complex observations in a focused radar image (SLC) are given in a matrix of pixels. Since an interferogram is defined by a point-wise multiplication of the master image with the complex conjugated of the coregistered slave image, the phase in the interferogram is equal to the wrapped phase difference between the images. The unwrapped phase is not observed, since there is an unknown integer number of cycles that the signal traveled, and only the last fractional part can be measured. However, only the unwrapped phase can be related to the parameters of interest. Moreover, if a time series of unwrapped phase data is available, the estimation of these parameters is relatively straightforward. For example, Meyer (2004) uses a set of spatially unwrapped interferograms to estimate the topography and displacement of polar glaciers using a least-squares approach. Furthermore, the temporal and spatial filters used in the PS technique (see chapter 2) to estimate the atmospheric signal can as well be applied to the unwrapped phase time series.

In this chapter a newly developed algorithm is presented. This Spatial Temporal Unwrapping Network (STUN) algorithm performs phase unwrapping on a spatially sparse grid, utilizing the integer least-squares estimator and a temporal displacement model. A final parameter estimation is performed after the data are unwrapped, see also the flow chart in Fig. 4.1. After an introduction explaining the need for phase unwrapping in section 4.1, the pixel selection is described in section 4.2. Then, section 4.3 addresses the estimation of the variance components to obtain the stochastic model used by the integer least-squares estimator. The estimation of a reference network is described in section 4.4, after which section 4.5 explains the estimation of points with respect to this established network. Finally, section 4.6 describes the explicit phase unwrapping and estimation using the unwrapped data.

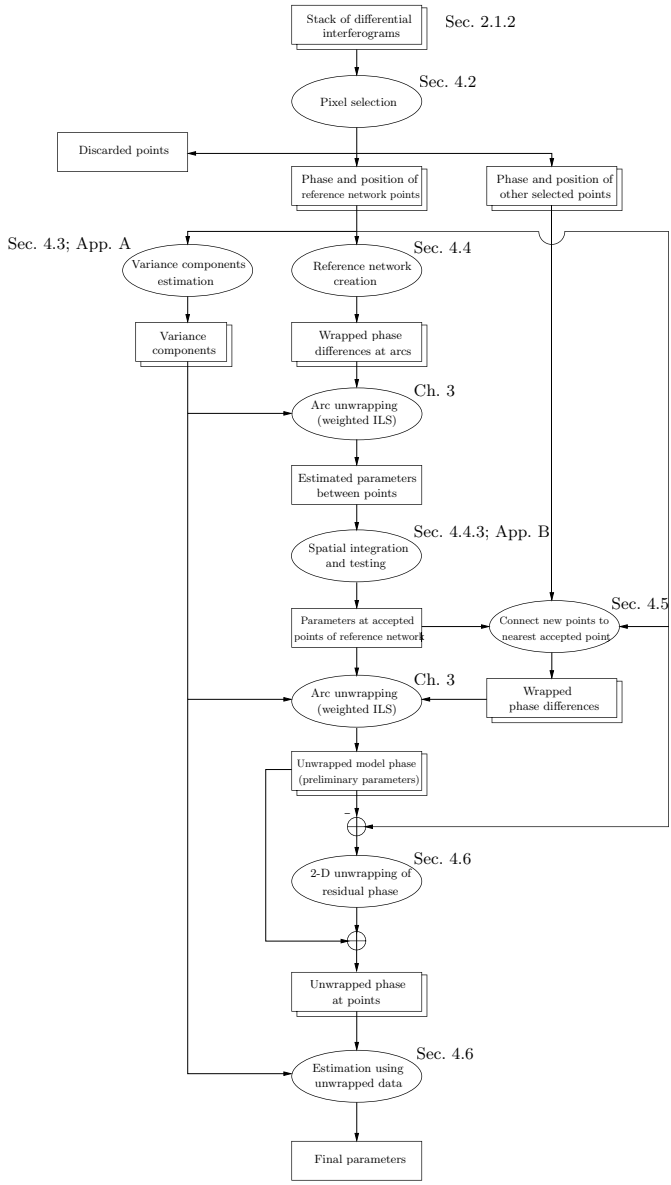
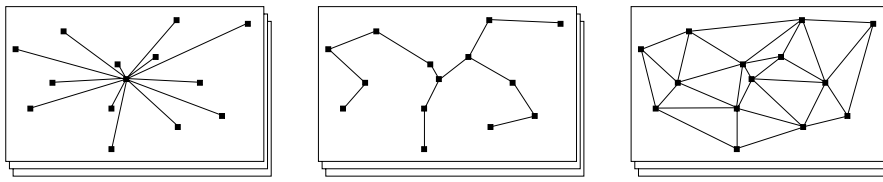


Fig. 4.1: STUN algorithm processing flow. Ovals represent processes, rectangles represent data. Temporal data are indicated by the double rectangles. First, the pixels are separated in three groups, and a variance component estimation is performed to obtain the vc-matrix of the observations. Second, the parameters are computed at the points of a reference network using the weighted integer least-squares estimator. Third, other selected points are estimated with respect to the established reference network. Finally, the phase is explicitly unwrapped, and a final parameter estimation is performed.

4.1 Three dimensional phase unwrapping

If unwrapped phase data are available the unknown parameters at the points can in principle be estimated directly by applying Eq. (2.48). Note that the phase in the interferograms must be relative to that of a reference pixel, because even if it is assumed that the unwrapped phase can be observed directly, these observations cannot be interpreted as absolute displacement observations due to inaccurate knowledge of the sensor position and signal delay. By taking spatial differences, these errors are reduced considerably, and the interferograms are related to each other by the reference point. This is depicted in Fig. 4.2(a). Moreover, Eq. (2.48) describes the simultaneous solution of a large number of points (possibly millions) using tens of images, which is not feasible due to numerical constraints. Instead, the (relative) parameters of each point can be estimated independently using the double-difference observations with respect to the reference point. In that approach the full vc-matrix of the estimated parameters at all points, cf. Eq. (2.47), is not obtained, i.e., the covariance between independently estimated points due to atmospheric signal.



(a) Star network **(b)** Minimal network **(c)** Redundant network

Fig. 4.2: Example networks of considered phase differences in interferograms. Shown are a small number of PSCs in (three) interferograms. **(a)** shows the case where the unwrapped phase relative to a reference pixel is available. **(b)** clarifies the situation for the wrapped case, when the phase difference between nearby points must be used to reduce atmospheric signal. Using the PS technique, the parameters are estimated using the phase time series. **(c)** introduces a network with redundant connections, which can be used to detect incoherent points and incorrect estimations between points.

Since the phase is observed in the interval $[-\pi, \pi)$, somehow the unwrapped phase must be obtained. Note that the slant-range (travel time) to the pixels cannot be used to obtain the unwrapped phase geometrically, since the wavelength is much smaller than the precision of this measurement, which currently is in the order of meters. It is also not possible, in general, to obtain the correctly unwrapped phase at all points using a conventional spatial unwrapping algorithm such as branch-and-cut (Zebker and Lu, 1998) or Minimum Cost Flow (MCF, see for example Chen, 2001; Chen and Zebker,

2002; Costantini, 1998) due to temporal and geometrical decorrelation. Since each point has its own unknown ambiguity, the unwrapped phase cannot be correctly estimated without using additional a priori information, see also, e.g., (Hanssen, 2001) or (Ghiglia and Pritt, 1998) for a general reference. To find the unwrapped solution for the interferogram, spatial unwrapping algorithms use the assumption that the unwrapped phase differences between adjacent pixels is generally limited to the interval $[-\pi, \pi)$, i.e., the phase is assumed to be induced by topography, displacement, or atmospheric signal, which is assumed to be similar for adjacent pixels. If this assumption is correct for all true unwrapped phase differences (between adjacent pixels) the wrapped phase differences are equal to the unwrapped phase differences, and the unwrapped solution is simply found by integration of the wrapped phase differences along an arbitrary path, starting from an arbitrary seed location. Unfortunately, undersampling of the signal and noise render this assumption incorrect, and currently best performing spatial unwrapping algorithms aim to find the optimal integration path, avoiding places where the unwrapped phase difference is not equal to the wrapped phase difference. Moreover, note that if there are isolated patches in the interferogram where the phase can be unwrapped correctly with respect to different reference points (i.e., seed locations to start the integration), it is not possible to correctly connect such regions without additional information.

Conventional spatial unwrapping algorithms cannot be used in Persistent Scatterer Interferometry, because interferograms are used that are severely affected by decorrelation. A sparse grid unwrapping algorithm, i.e., only applied to points that are expected to be coherent, proposed in (Costantini and Rosen, 1999; Eineder and Holzner, 1999; Yong et al., 2002), is also not likely to correctly unwrap the observed data in most circumstances, due to the increased distance between points, which makes it more likely that the (spatial, absolute) unwrapped phase differences are larger than π, particularly for large spatial and temporal baselines, due to larger topographic, displacement, or atmospheric signal differences. Therefore, first the largest part of these signal components needs to be removed before the phase can reliably be unwrapped spatially on a sparse grid.

In the PS technique, the phase differences between two points are considered in a time series of interferograms. The parameters of a model are estimated using the observed wrapped phase differences, see also section 2.1.3, in particular Fig. 2.3 on page 12. Using the assumption that the absolute residual phase with respect to the model phase, see Eq. (2.8), is smaller than π in all interferograms, this directly yields the unwrapped phase, which implies that the PS technique essentially is a model-based, temporal (one-dimensional) unwrapping strategy. However, this assumption is only valid if the displacement model is a good approximation for the actual displacement, and if the noise and atmospheric phase is small. Since the atmospheric signal—which is not modeled—is expected to increase for points further apart, the phase differences have to be considered between nearby points. This implies

that a network as sketched in Fig. 4.2(b) can be used. In this case, first the parameters can be estimated between nearby points, after which they can be integrated with respect to a reference point, which then yields the unwrapped phase, i.e., the same situation as sketched in Fig. 4.2(a). It is obvious that a possible incorrect estimation between two points propagates to the other points, while this cannot be detected. Moreover, if the residual phase in a single or a few interferograms is larger than π, this is not noticeable if only temporal unwrapping is used.

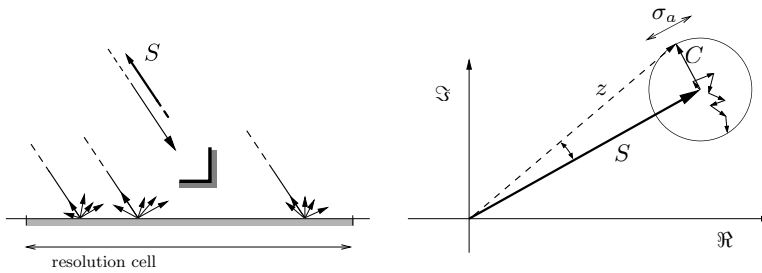
Therefore, a spatio-temporal unwrapping strategy is developed combining these two approaches of two-dimensional spatial sparse grid unwrapping and one-dimensional temporal unwrapping. This implies using a network as sketched in Fig. 4.2(c), where phase differences in space and time are used to obtain the unwrapped phase at the PS points. An algorithm that uses the three-dimensional network directly is likely to be the best possible approach to correctly estimate the unwrapped phase at all points, in all interferograms. Unfortunately, an efficient algorithm for this problem is not yet developed. The integer least-squares estimator can in principle be used to solve this problem, but the large amount of unknown ambiguities that have to be estimated prevents a direct application.

4.2 Pixel selection

The phase stability of the pixels in the interferograms is not known beforehand. That is, a large number of pixels is likely to be decorrelated, particularly for interferograms with large temporal and perpendicular baselines. Reasons for this decorrelation are:

- The angle under which the resolution cell is observed during the two acquisitions is different (geometrical decorrelation).
- The elementary scatterers in the resolution cell move incoherently in the time between the acquisitions (temporal decorrelation).
- Processing induced decorrelation , e.g., due to mis-registration of the slave image. This cause is not further considered.

A distributed scatterer, i.e., a pixel for which the backscattered signal is the complex sum of many uncorrelated elementary scatterers of which none is dominant, cannot remain coherent in interferograms with either a large perpendicular baseline or squint angle (Doppler centroid frequency) difference, see also (Hanssen, 2001). Therefore, only pixels that have a dominant scatterer, i.e., point scatterers, are relevant. The signal model for such an observation is sketched in Fig. 4.3, see also (Adam et al., 2004). A dominant scatterer is spatially surrounded by incoherent clutter. Thus, the observed phase is composed of a dominant signal and a superposition of the clutter. The phase of the main scatterer is related to the distance to the sensor, while the resulting phase caused by the clutter is random. For a PS point, the phase center of the



(a) Spatial relation **(b)** Complex sum

Fig. 4.3: Signal model for a dominant point scatterer surrounded by incoherent background clutter. **(a)** depicts this in the spatial domain for a single resolution cell. **(b)** shows the final backscattered signal, i.e., the complex sum of the elementary scatterers in the resolution cell. The dashed vector z indicates a complex observation. Also indicated are the temporal variation σ_a of the amplitude $a = |z|$ and the phase error due to the clutter.

main scatterer is assumed to be at the same position under a wide range of angles, hence the pixel does not suffer from geometrical decorrelation. Moreover, the main scatterer is assumed to remain the dominant scatterer during the whole time span of all radar observations.

The majority (>~90%) of the pixels in an interferogram are expected to have a distributed scattering mechanism. Therefore, computation time is reduced significantly if the estimation is restricted to pixels with a point scattering mechanism. Moreover, if pixels with a distributed scattering mechanism are discarded the false alarm rate decreases (the probability that a pixel is regarded a PS while it is not), since it is possible that a number of pixels with a random phase have a good fit with the model.

Fig. 4.3(b) clearly demonstrates that the phase error decreases if the signal power increases. This suggests three possibilities for the detecting of point scatterers, i.e., methods based on thresholding of the amplitude, the signal to clutter ratio, or the amplitude dispersion index. These three methods are described in more detail in the following.

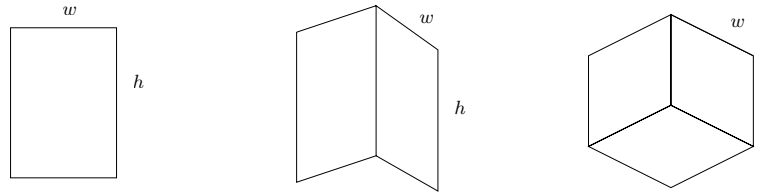
Thresholding on the amplitude

A pixel is selected as possible PS point if its normalized radar cross section σ^o (RCS), i.e., the calibrated intensity, is above a threshold N_2 in at least N_1 SLC images

$$\sum_{k=0}^{K} a_k \geq N_1, \qquad \text{with } a_k = \begin{cases} 1, & \text{if } \sigma_k^o > N_2 \\ 0, & \text{otherwise} \end{cases}, \tag{4.1}$$

see also (Kampes and Adam, 2004). The normalized RCS describes the mean reflectivity of an area of one square meter on the surface of the earth

(Ulaby et al., 1982). It is dimensionless and commonly expressed in decibels. Furthermore, it is independent of the image resolution or pixel size. Typical values for these thresholds are $N_1 = 0.65K$ and $N_2 = -2$ dB. Note that a value of $\sigma^o = 0$ dB for the normalized radar cross section indicates that the received backscatter is equal to the expected backscatter for a resolution cell with distributed scattering properties, see, e.g., (Elachi, 1987; Raney et al., 1994). Thus, it is unlikely that a pixel below 0 dB is a point scatterer, and by using a threshold $N_2 = -2$ dB it is expected that most point scatterers in the interferogram are detected. The main advantage of this method is that it is a simple method, although it requires the amplitude data to be calibrated.



(a) Planar mirror **(b)** Dihedral **(c)** Square trihedral

Fig. 4.4: Some fundamental reflector types.

In the following an interpretation of the threshold N_2 is given for a few fundamental reflector types, see Fig. 4.4. The radar cross section $\Delta\sigma$ of a square trihedral corner reflector with sides w is given as (Curlander and McDonough, 1991)

$$\Delta\sigma = \frac{12\pi w^4}{\lambda^2}. \tag{4.2}$$

The unit of the RCS is m^2. It is assumed that the point scatterer is much smaller than the resolution cell and that there is no background clutter. Then, the RCS can be approximated as the product of the normalized RCS with the ground-range resolution area A

$$\Delta\sigma = \sigma^o \cdot A. \tag{4.3}$$

Using the ERS ground-resolution of 5 m in azimuth and 25 m in ground-range direction, and a wavelength of 0.056 m, it follows that a pixel with a normalized RCS of –2 dB corresponds to a square trihedral corner reflector with sides of 0.28 m, since

$$w^4 = 10 \cdot 5 \cdot 10^{\frac{-2}{10}} \cdot \frac{\lambda^2}{12\pi} \;\leftrightarrow\; w = 0.28. \tag{4.4}$$

The normalized RCS value in decibel is converted back to a digital number using $\sigma^o = 10^{\sigma^o_{\mathrm{dB}}/10}$. The RCS of a dihedral, e.g., a double bounce of a street

and the wall of a house that are aligned with the flight path, is given as (Freeman, 1992)

$$\Delta\sigma = \frac{8\pi w^2 h^2}{\lambda^2}, \tag{4.5}$$

and that of a planar mirror, e.g., a metal rooftop with its normal direction in the line-of-sight to the sensor, is (Curlander and McDonough, 1991)

$$\Delta\sigma = \frac{4\pi w^2 h^2}{\lambda^2}, \tag{4.6}$$

where h and w are the height and width of the individual panels. Assuming these two dimensions are equal, a pixel with a normalized RCS of –2 dB thus corresponds to elements with sides of 31 and 37 cm for the dihedral and planar mirror, respectively.

Thresholding on the SCR

With this method pixels are selected if the average signal to clutter ratio (SCR) of a pixel is above a certain threshold. The relation of the SCR to the phase error is (Adam et al., 2004):

$$\hat{\sigma}_\phi = \hat{\sigma}_\Re = \frac{1}{\sqrt{2 \cdot \mathrm{SCR}}}. \tag{4.7}$$

Thus, a reasonable threshold SCR=2 selects points with a phase standard deviation $\sigma_\phi < 0.5$ rad ($\sim 30°$). The length of the vectors S and C is estimated using a point target analysis. In order to obtain estimates for the SCR a spatial estimation window is used. The assumption is that the power of the clutter around the pixel is equal to the power of the clutter inside the resolution cell. This technique to estimate the SCR was developed to check the phase stability of corner reflectors that were deployed for calibration purposes at specific locations with low clutter power, see also Fig. 4.5 and (Freeman, 1992). If this

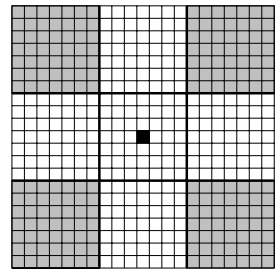


Fig. 4.5: Signal to clutter ratio estimation method for a corner reflector, see also, e.g. (Freeman, 1992). The shaded regions are used to estimate the power of the clutter, while the other pixels are used to estimate the power of the signal.

method is applied to a city area the clutter is likely to be overestimated using the spatial window, because there are more point scatterers in close proximity to each other. To compensate for this effect a relatively low threshold for the SCR should be used. However, note that a SCR equal to one indicates equal power for signal and clutter, which can be taken as a lower bound for this threshold. Moreover, the SCR method increasingly underestimates the phase error for decreasing SCR, see (Adam et al., 2004). The advantage of this method is that the amplitude data need not to be calibrated to estimate the SCR, and that it does not require assumptions on the temporal amplitude behavior of the considered pixel.

Thresholding on the amplitude dispersion

The amplitude dispersion index D_a is used in the reference PS technique to detect a relatively small number of points that have a high phase stability, see also section 2.1.3. Thus, a threshold on D_a can be used to detect point scatterers. The phase standard deviation is estimated using Eq. (2.6), repeated here for convenience, as

$$\hat{\sigma}_\phi = \frac{\sigma_a}{\bar{a}} = D_a,$$

which uses that $\sigma_a \approx \sigma_\Re$ and $\bar{a} \approx \sqrt{S}$ for high SCR, see (Ferretti et al., 2001). Adam et al. (2004) showed that the amplitude dispersion index method is based on the same signal model for point scatterers shown in Fig. 4.3. From this figure, the relation becomes clear by considering that for a certain point scatterer the temporal variation of the amplitude of the observable z is related to the clutter power C, while the amplitude of the observed variate is a function of both signal S and clutter power C. For high SCR, the dispersion index is a good approximation of the phase error. However, as shown in Fig. 2.2, the estimate of the phase error using the amplitude dispersion index only is a good approximation if this ratio is below ~0.35 rad (20°). Moreover, the dispersion index theoretically tends to the value 0.5 for low SCR (Ferretti et al., 2001), and a good value for a threshold to detect most point scatterers, but only a small amount of distributed scatterers may be difficult to find. Note that in the PS technique the dispersion index is only used to find the points with the highest SCR. The advantage of this method is that it can be directly applied to the calibrated amplitude data. It does not require assumptions on the amplitude behavior in the area around the considered pixel. However, at least ~20 temporal samples (images) are required for a reliable estimation of D_a, see (Adam et al., 2004).

Comparison of pixel selection methods

For high SCR the estimates for the phase variance using the amplitude dispersion method and the SCR method are equivalent. Both methods are biased, although the SCR method to a lesser extent, which is proven by the

following. Consider a complex variable $z = s + n$ where the signal $s = 1$, and n is complex circular Gaussian noise with $\sigma_{\Re} = \sigma_{\Im} \ll |s|$. Then, the signal power is $S = |s|^2 = 1$ and the clutter power is $C = |n|^2 = 2\sigma_{\Re}^2$. Thus, the SCR method estimates the phase standard deviation as $\hat{\sigma}_{\phi}^{\mathrm{SCR}} = \sigma_{\Re}$, which is a valid approximation for high SCR. The D_a method uses the estimated mean and standard deviation of the amplitude to estimate the phase standard deviation. The expectation for the estimated amplitude $E\{|z|\} > |s|$, which is due to the circular Gaussian noise n. Moreover, the expectation of the standard deviation of the amplitude $E\{\sigma_{|z|}\} \leq \sigma_{\Re}$. Therefore, the estimated phase error using the SCR method must be larger than that using the D_a method. Since the SCR method underestimates the phase error, it follows that the D_a method has a larger bias. This was confirmed by a numerical experiment performed by Adam et al. (2004).

The three described methods for pixel selection are applied to a 2×2 km^2 area of Berlin, Germany, where seventy images are available. The results are shown in Fig. 4.6. The thresholds that are used here and the resulting point density should not be compared with each other necessarily. However, clearly visible is that all methods do not tend to select scatterers that are not bright. The goal of the pixel selection is to select most point scatterers that were present during the whole time span of the acquisitions. This goal can be achieved with all methods by lowering the thresholds, in the extreme case until all pixels are selected. However, the percentage of distributed scatterers that are also selected may vary between the methods.

The threshold for the amplitude dispersion index method is set to 0.4, selecting 7357 of 200000 (oversampled) pixels, see also Fig. 4.6(f). The estimated minimum and maximum amplitude dispersion index is 0.12 and 2.19, respectively. The average is 0.54 and the standard deviation 0.10. Thus, for most pixels the estimated amplitude dispersion index is larger than the theoretical limit for D_a, which is 0.5. The reason for these larger values can be a calibration problem (which cannot be confirmed after a statistical analysis of all images), can be due to the uncertainty of the estimator for D_a, or can be due to temporal changes of the actual signal and clutter power. The latter reason is assumed to be the most likely cause. For example, the amplitude dispersion method does not select the extremely bright scatterer slightly off center in the average intensity image, see Fig. 4.6(a). This is easily explained by the fact that this scatterer was not present during earlier acquisitions, resulting in a large value for the amplitude standard deviation for this pixel. However, this method does not seem to be optimal for the selection of all point scatterers.

With the amplitude thresholding method 18191 of 200000 pixels were selected, see Fig. 4.6(d), likely including most point scatterers. However, the selected pixels do not have to be independent of each other, i.e., the high amplitude of a pixel may be due to side lobes of a nearby strong scatterer. Furthermore, it is likely that the same point is selected in at least four pixels with this method, since the data are oversampled by a factor two in both

directions. Moreover, layover effects may cause a large amplitude for a pixel that is not a point scatterer. After a point target analysis (sub-pixel local peak detection) at the selected pixels, 2308 unique points remain (~575 points per km^2). Fig. 4.6(e) shows the selected points using the SCR method for this area. The geometrical distribution of the points selected by this method seems to be better than that for the amplitude thresholding method, see Fig. 4.6(f), which can be seen, for example, in the dark triangular patch slightly off center, where the amplitude thresholding method does not select any pixel. After a point target analysis 1072 unique pixels were selected in this area, of originally 9489 selected points with an SCR > 2. The SCR method is used in the following for the pixel selection, followed by a point target analysis to obtain the phase at the estimated sub-pixel peak position.

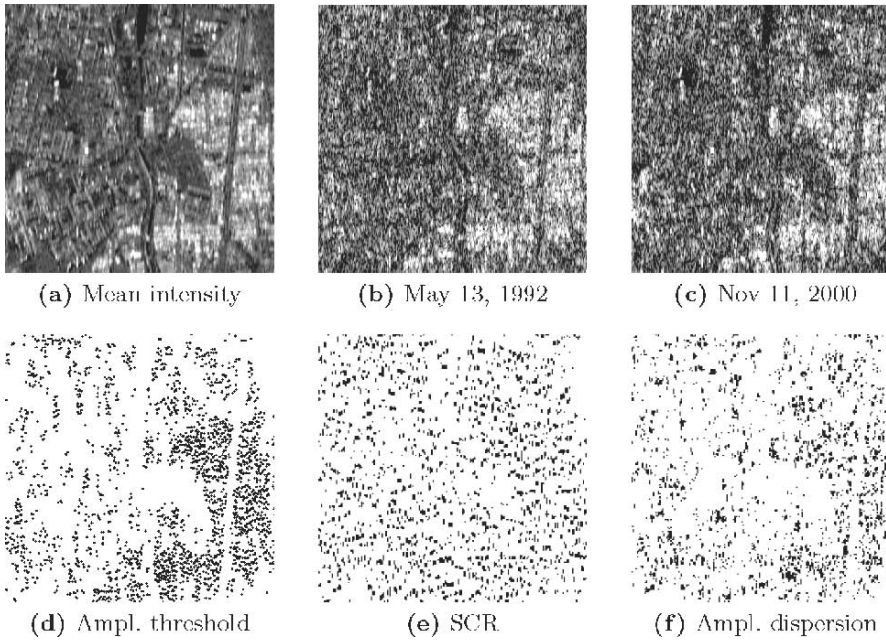


Fig. 4.6: Pixel selection with various methods for a 2×2 km^2 area of Berlin. The top row show the mean intensity of seventy available images in the interval $[-5, 15]$ dB. The intensity images for the first and last available acquisition are shown as well. Note that the radiometric resolution increases significantly by taking 70 temporal looks. The second row shows the selected pixels using the different methods. The threshold for the amplitude is $N_1 = 0.65K$ and $N_2 = -2$ dB, resulting in the selection of 18191 of 200000 (oversampled) pixels. Fig. 4.6(e) shows 9489 selected points using SCR >2. Finally, Fig. 4.6(f) shows 7357 pixels selected using $D_a > 0.4$.

4.3 Variance component estimation

Before the unknown parameters are estimated using the weighted ILS estimator (described in Chapter 3), first a variance component estimation is performed. The vc-matrix Q_{ifg} of the double-difference phase observations is parameterized by Eq. (2.51), i.e., the dependency on distance is ignored for the estimation of the relative parameters between nearby points. This is justified, since all arcs approximately have the same length. The estimated variance components of Eq. (2.51) in this case actually refer to $\sigma_k^2 = \sigma^2_{\text{noise}^k} + \sigma^2_{\text{atmo}^k}(0) - \sigma_{\text{atmo}^k}(l)$, cf. Eq. (2.50). The contribution of the atmospheric signal to the dispersion of the interferometric double-difference observations is supposed to be much smaller than that of the inherent noise. Moreover, if a relation of the variance component with distance is ignored, the least-squares projection matrix is identical for all estimations, which allows for a faster estimation.

An initial estimation of the unknown parameters is required before estimation of the variance components is possible. During this initial estimation, a stochastic model with a priori variance components[1] is used cf. Eq. (2.50). This a priori model is based on the assumptions that the interferometric phase standard deviation for point scatterers is expected to be below ~50°, and that slight mis-registration introduces a small amount of additional noise in the slave images. Note that the variance component estimation can be performed iteratively, and that the choice of the stochastic model during the initial estimation is not very important, e.g., a scaled identity matrix $\sigma^2_{\text{noise}} I_K$ could also be used. The vector of variance components of the SLC images $\sigma = [\sigma^2_{\text{noise}^0}, \ldots, \sigma^2_{\text{noise}^K}]^*$ is estimated by using the temporal least-squares interferometric phase residuals vector $\hat{\underline{e}}$ of the initial estimation between two points as (Verhoef, 1997)

$$\hat{\underline{\sigma}} = N^{-1}\underline{r}, \tag{4.8}$$

where N is a square $(K+1)\times(K+1)$ matrix. The elements of these matrices are given by (see also Appendix A)

$$\begin{aligned} \underline{r}(k+1) &= \hat{\underline{e}}^* Q_y^{-1} Q_k Q_y^{-1} \hat{\underline{e}}, \\ N(k+1, l+1) &= \text{trace}(Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} Q_l), \end{aligned} \tag{4.9}$$

for $k, l = 0, \ldots, K$ (Q_k is defined in Eq. (2.51)). The least-squares orthogonal projection matrix $P_B^{\perp}$ is given by $P_B^{\perp} = I_K - B(B^* Q_y^{-1} B)^{-1} B^* Q_y^{-1}$. The vc-matrix of the estimated variance components is given by (Verhoef, 1997)

$$D\{\hat{\underline{\sigma}}\} = 2N^{-1}. \tag{4.10}$$

This variance is reduced by using more estimations between different points. If each point in these estimations is used only once (as sketched in Fig. 4.7) the estimate for the variance components corresponding to the parameterization

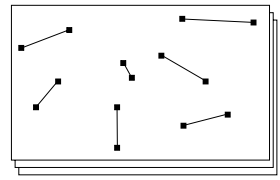


Fig. 4.7: "Network" used for the estimation of the variance components $\sigma^2_{\text{noise}^k}$ of the parameterization of Eq. (2.51). First the unknown parameters (displacement, DEM error) are estimated between the indicated points. Then, for each arc, the least-squares phase residuals are used to estimate the variance components cf. Eq. (4.8). Since each point is used only once, the mean of the estimated components is a more precise, unbiased, estimate for the unknown variance components.

(2.51) is given by the mean of these estimations. The variance of the estimated variance components is reduced by the number of estimations.

It is not guaranteed by Eq. (4.8) that the estimated variance components are larger than zero. If this happens it could indicate that initially $\sigma^2_{\text{noise}^0}$ is too large, that the number of estimates used to estimate the variance components is too small, or that the least-squares estimates are incorrect, i.e., the displacement model that is used. To avoid a possible non positive-definite vc-matrix, the estimated (mean) variance components $\sigma^2_{\text{noise}^k} < (10^\circ)^2$ are set to $(10^\circ)^2$.

4.4 Reference network computation

In the STUN algorithm, not all pixels in the interferograms are estimated, even not during a final estimation step as is done in the reference PS technique, see, e.g., (Ferretti et al., 2001). The pixels that are not selected in this step are discarded in the further processing, because it is not expected that they contain useful phase information, see section 4.2 for a more detailed explanation. Remember that the threshold for the pixel selection is low, such that a large amount of pixels is selected. The thresholds that are used for the pixel selection do not have to be extremely selective, because analysis of the phase data will reveal if the selected pixels are coherent or not. This selection reduces the number of points that needs to be estimated from hundreds of millions to a few hundred thousand. Memory requirements go down dramatically once this set is obtained. The points that are not discarded are further divided into two groups; reference points (PSCs) and other points, similar to the preliminary and final estimation step of the PS technique, see also Chapter 2. Most reference network points are expected to be coherent in time, based on their amplitude dispersion index. The parameters are first estimated between the points of the reference network, which is described

[1] $\sigma^2_{\text{noise}^0} = (20^\circ)^2$, and $\sigma^2_{\text{noise}^k} = (30^\circ)^2$ for $k = 1, \ldots, K$

in this section. The estimation of the other selected points is described in section 4.5.

4.4.1 Construction of the reference network

During the preliminary estimation step of the reference PS algorithm, all pixels are selected that have an amplitude dispersion index below a certain threshold, see section 2.1.3. The goal of this selection is to establish a reference network of coherent points, which preferably are distributed equally over the area of interest in order to interpolate the estimated atmospheric signal. In our approach the points of the reference network (PSCs) are selected differently. A grid is placed over the interferogram, and in each grid cell the point with the smallest amplitude dispersion index is selected, cf. Eq. (2.6). In order to avoid PSCs extremely close to each other, a grid shifted in both directions by half the cell width is placed over the initially selected points, and again the point with the smallest amplitude dispersion is selected. This guarantees a minimum distance between points equal to the cell width, which is normally set to ~500 m. This distance is a good trade-off between limiting atmospheric signal and obtaining an amount of reference network points that can be conviently processed. Finally, PSCs are discarded if their amplitude dispersion index is larger than a threshold (e.g., 0.67). This reduces the selection of incoherent points that otherwise may be selected, for example if the grid cell contains only water.

Once the PSC points are selected, a network needs to be constructed that indicates between which pairs of PSCs the estimation will be performed. In its simplest form, a network like in Fig. 4.2(b) can be used, but obviously it is better to be able to check whether an estimation between two points went wrong. In the STUN algorithm the network is constructed by connecting each point with the nearest N other points, but such that it is connected equally well in all directions. This is guaranteed by first connecting a point to its nearest neighbor in the first quadrant, then in the second, etc., until the required number of connections is reached. Arcs are not allowed to be longer than a certain maximum length, e.g., 2400 m, and possible double connections are removed in a final step. Moreover, this approach prevents that points at the borders of the estimated area are connected with less arcs than points in the middle, as happens with, e.g., Delaunay triangulation.

The reference network is thus computed using only a small portion of all selected pixels. It is possible to set up a system of equations with on the left hand side phase differences between the points of the reference network, and as unknown parameters the integer ambiguities and the float parameters at the points. Then, the integer least-squares estimation algorithm can be used to estimate these parameters directly. However, it is not straightforward to perform this estimation due to numerical constraints, because the number of points in the reference network is still a couple of thousand (assuming a distance between points of a few hundred meters), i.e., hundreds of thousands

ambiguities need to be estimated, aside from thousands of unknown float parameters. Therefore, the computation of the reference network is split in two parts. The DEM error and displacement parameters are first estimated between nearby points using the interferometric phase time series. After this temporal estimation, the parameters at the points are obtained by a spatial least-squares adjustment of the estimated difference parameters, similar to that conventionally used for leveling data. An alternative hypothesis testing procedure is followed to identify incorrect estimates in the first step. In both of these steps least-squares residuals are present. In the first step this is the residual phase in the interferograms, and in the second step it is the residual of the estimated parameters at the estimated arcs. Note that it may be confusing in the following, which least-squares residuals are addressed. In order to reduce the possible confusion, the residuals are referred to as temporal and spatial least-squares residuals, respectively. The design matrix relating the phase time series to the unknown parameters is B, and the design matrix for the spatial network is C. The symbols for the observables ($\underline{y}$), unknown float parameters (b), and the least-squares residuals ($\underline{e}$) will remain unchanged.

4.4.2 Spatial integration of estimated difference parameters

The DEM error and displacement parameters (differences) are first estimated at all arcs of the reference network, see, e.g., Fig. 4.2(c). The integer least-squares estimator, described in Chapter 3, is used for the estimation, instead of the maximization of the ensemble coherence, which is used in the reference PS technique. The dispersion of the double-difference phase observations is given by the vc-matrix of Eq. (2.51), where the estimated variance components are used to create the vc-matrix.

If all estimates at the arcs would be correct, the parameters at the points are obtained by integration along any given path. Then, there would be no need to perform these many estimations, instead, the estimations could be restricted to the arcs of network Fig. 4.2(b). It is noted here again that if the observed phase data at the points are not wrapped, then a network like in Fig. 4.2(a) can be used to obtain the parameters at each point (assuming a known reference point). In this case, the least-squares phase residuals can directly be used to identify incoherent points. Unfortunately, it would be impossible to identify incorrect estimations using a network like that of Fig. 4.2(c), since there would never be any inconsistency; the double-differences of the unwrapped phase at the arcs are linear combinations of the unwrapped phase observed at the points. However, if wrapped data inconsistencies occur, they can be used to identify incorrect estimations, as described in section 4.4.3.

When the network of Fig. 4.2(c) is considered, the parameters are estimated between the points indicated by the lines. It will be assumed in the following that only a DEM error (difference) and a displacement rate (difference) are estimated, although there is not such a restriction in the

algorithm. The DEM error at the points can be obtained with respect to the reference point (the first point; i.e., this unknown is removed from the vector of unknown parameter together with the corresponding column of the design matrix) by solving a system of observation equations like

$$\begin{bmatrix} \Delta h_{2,1} \\ \Delta h_{3,1} \\ \Delta h_{4,1} \\ \vdots \\ \Delta h_{3,2} \\ \Delta h_{4,2} \\ \vdots \\ \Delta h_{H-1,H} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & 0 \\ \vdots & & & & \\ 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots 1 & -1 \end{bmatrix} \begin{bmatrix} \Delta h_2 \\ \Delta h_3 \\ \Delta h_4 \\ \vdots \\ \Delta h_{H-1} \\ \Delta h_H \end{bmatrix}, \tag{4.11}$$

where the design matrix corresponds to the estimated arcs. The solution for the unknown parameters is given by Eq. (4.13). An identical design matrix is used to "integrate" the other estimated difference parameters independently, in this example the displacement rate differences at the arcs. This system of equations looks very similar to that of a leveling network. However, the difference is that here the least-squares residuals (the misclosures) at the arcs must be exactly zero, since there are no actual observations between points, as with leveling data. If all spatial residuals are indeed equal to zero the least-squares estimates at the points are the same as those that would be obtained after integration along any path.

4.4.3 Identification of incorrect estimates and incoherent points

In practice non-zero residuals are found at the arcs of the network, due to incorrect relative estimations at certain arcs. The problem is to identify the reason for this. Possible reasons are that a point is incoherent or that only an individual arc is estimated incorrectly. Consider the elementary case where three points (x, y, z) in one interferogram are available, see also Fig. 4.8. The phase is assumed to be induced by linear displacement only, i.e., the functional model is given as $E\{\phi_x\} = -\frac{4\pi}{\lambda} T \cdot \alpha_x$. Without loss of generality, the point x is selected as the reference point, and the double-difference observations $\in [-\pi, \pi)$ are given by $\phi_{x,y}$, $\phi_{y,z}$, and $\phi_{z,x}$. Signal aliasing cannot be inferred from these wrapped phase differences. For example, if the phase differences due to displacement (difference) are equal to $-\frac{4\pi}{\lambda} T \cdot \alpha_{x,y} = -\frac{4\pi}{\lambda} T \cdot \alpha_{y,z} = \pi - \epsilon$, where ϵ is a small positive number, then the wrapped phase difference $\phi_{z,x}$ is not equal to the unwrapped phase difference, which implies that the estimate $\underline{\hat{\alpha}}_{z,x}$ is incorrect, resulting in a misclosure of the estimated parameter at the arcs. This is equivalent to the way a residue is formed due to signal aliasing in the conventional two-dimensional phase unwrapping problem.

In this example with only one interferogram, the estimated difference parameters have an ambiguity that is one-to-one related to the phase differences.

Due to aliasing, a signal above the Nyquist frequency cannot be estimated, and it is assumed that the solution closest to zero is used, which is achieved by limiting the search space. In the general case, where there are K interferograms available, the estimation can also be incorrect, either due to signal aliasing or noise.

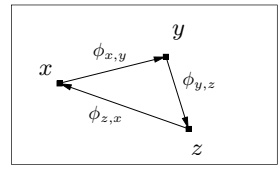


Fig. 4.8: Elementary network for which a misclosure can occur if parameters are estimated using the wrapped interferometric phase observed at points x, y, and z.

If the search space is not limited, there always is a solution for the ambiguities such that all least-squares phase residuals are zero. This is due to the fact that the problem is inherently under-determined, i.e., there are more unknown parameters than observations. In practice, the search space must be bounded, and the estimated parameters minimize some norm defined on the solution space. If it is assumed that only a single point is incoherent, and its parameters are estimated with respect to nearby points, then the location of the minimum (i.e., the estimated parameters) depends on the noise of the other points. This means that misclosures can occur. But even if all points are coherent, it may happen that a single estimation (an arc) is incorrect. Reasons for this could be that the true solution lies outside the search space for a specific arc, or that the wrong local minimum is found. The former reason becomes clear by considering Fig. 4.8 again. Suppose that the search space is limited to [–20, 20], while the true parameters are $\alpha_{x,y} = \alpha_{y,z} = 15$. Then, the estimated value $\underline{\hat{\alpha}}_{z,x}$ cannot be correct, and a misclosure occurs. The latter could happen if an algorithm is used which searches the float parameters with a too coarse resolution, for example the algorithm described in section 2.1.3. This problem does not exist for the integer least-squares estimator, since the integer ambiguities are searched instead of the float parameters.

Note that the least-squares phase residuals of the estimation between two points (i.e., the phase residuals in the interferograms of the integer least-squares estimation) are in the interval $[-\pi, \pi)$. These residuals do not necessarily have to be large if a point is incoherent. This is particularly true when only a small number of interferograms is available.

Alternative hypotheses testing

The integrity of the network should be checked for the reasons described above. Note again that the situation for the created network is not the same

as for, e.g., a leveling network. Here the misclosures must be perfectly zero, because there is no independent measurement (noise) on the arcs. A misclosure is solely due to incorrectly estimated parameters at an arc. Such an error should not be adjusted, but the cause must be found and rectified. Moreover, the network testing procedure does not have to be approached as strict as for, e.g., a leveling network. After all, if the misclosure is zero, removing a (correctly estimated) arc does not change the parameter solution or its precision. However, zero misclosures do not guarantee that the parameters are estimated correctly (they can all be consistently incorrect), but it does mean that all errors that could be found are dealt with.

In order to find outlier arcs and points, an alternative hypotheses testing strategy is used, known as DIA procedure (Teunissen, 2000b). First, in the *detection* step, the null-hypothesis is tested against general model mis-specifications. If this test is rejected alternative hypotheses are specified to *identify* the most likely cause. In the *adaption* step, either the stochastic model, or the functional model is changed to account for the identified cause. These steps are repeated until the null-hypothesis is accepted.

First, the case is considered where only DEM error differences are estimated, thus ignoring other estimated parameters. The system of observation equations, Eq. (4.11), is written as

$$H_0 : E\{\underline{y}\} = Cb, \qquad D\{\underline{y}\} = Q_y. \tag{4.12}$$

The design matrix C specifies the functional relation between the unknown parameters b at the points to the differences $\underline{y}$ between points, while matrix Q_y is the vc-matrix of the latter. The well-known least-squares formulas can be applied to this system of equations, see also (Teunissen, 2000b)

$$\begin{array}{ccc} Q_{\hat{b}} = (C^* Q_y^{-1} C)^{-1}, & Q_{\hat{y}} = C Q_{\hat{b}} C^*, & Q_{\hat{e}} = Q_y - Q_{\hat{y}}, \\ \underline{\hat{b}} = Q_{\hat{b}} C^* Q_y^{-1} \underline{y}, & \underline{\hat{y}} = C\underline{\hat{b}}, & \underline{\hat{e}} = \underline{y} - \underline{\hat{y}}, \end{array} \tag{4.13}$$

where $Q_{\hat{b}}$ is the estimated vc-matrix for the unknowns, $Q_{\hat{y}}$ that of the adjusted observations, $Q_{\hat{e}}$ that of the least-squares residuals, and $\underline{\hat{b}}$, $\underline{\hat{y}}$, and $\underline{\hat{e}}$ are the vectors of adjusted unknowns, observations and residuals, respectively. When there are more observations than unknown parameters, it is possible to test the null-hypothesis, Eq. (4.12), against alternative hypotheses. An alternative hypothesis is specified as a linear extension of the null-hypothesis

$$H_A : E\{\underline{y}\} = Cb + C_q \nabla, \qquad D\{\underline{y}\} = Q_y. \tag{4.14}$$

Matrix C_q defines the considered model mis-specification. One could think of an almost infinite number of alternative hypotheses to test, and in general a priori information is required to formulate them. The test statistic for a q-dimensional alternative hypothesis is given as (Teunissen, 2000b)

$$\underline{T}_q = \underline{\hat{e}}^* Q_y^{-1} C_q (C_q{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_q)^{-1} C_q{}^* Q_y^{-1} \underline{\hat{e}}. \tag{4.15}$$

This test statistic has a chi-squared distribution with q degrees of freedom

$$\begin{aligned} H_0 &: \underline{T}_q \sim \chi^2(q, 0), \\ H_A &: \underline{T}_q \sim \chi^2(q, \lambda_0), \end{aligned} \tag{4.16}$$

where λ_0 is the non-centrality parameter. The null-hypothesis is rejected if the test quotient, i.e., the test statistic normalized by its critical value, is larger than one

$$\underline{T}_q / \chi^2_{\alpha(q)} \begin{cases} >1 & \text{reject } H_0 \\ \leq 1 & \text{do not reject } H_0 \end{cases}. \tag{4.17}$$

Here, $\alpha(q)$ is the level of significance, i.e., the probability that the null-hypothesis is rejected while it is correct (type-I error), see also Appendix B. The B-method of testing (Baarda, 1968) is applied, which means that all tests have the same power. This implies that if the null-hypothesis is accepted, then all alternative hypotheses are rejected. The power of the test γ is the probability that the alternative hypothesis is accepted when indeed it is correct. The probability of accepting null-hypothesis while it is false (type-II error) is given by $\beta = 1-\gamma$. By fixing the level of significance for the one-dimensional test to a certain value, e.g., $\alpha = 0.001$, the level of significance of the other tests (and corresponding critical values) can be computed using the B-method of testing. It is essential that the power of the tests is set to equal $\gamma = 0.50$ in order to identify the most likely alternative hypothesis among tests of different dimensions, see also Appendix B.

Detection

The detection step of the DIA procedure uses the overall model test (OMT) to find model mis-specification in either or both the stochastic or functional model. The OMT is given by

$$\text{OMT} = \underline{\hat{e}}^* Q_y^{-1} \underline{\hat{e}}. \tag{4.18}$$

The dimension of the overall model test is equal to the redundancy r of the problem. It is an important safeguard to indicate the validity of the null-hypothesis. It is the most relaxed possible alternative hypothesis, imposing no constraints on the observables. Hence, no restrictions are imposed on the observables. Matrix C_q consists of r unit vectors for this alternative hypothesis, but note that it does not need to be explicitly specified to compute the OMT. If the OMT is rejected, alternative hypotheses are specified to identify the most likely cause.

Identification

The one-dimensional (conventional) alternative hypothesis that specifies an individual outlier observation is given by

$$C_q = \begin{bmatrix} 0 \\ . \\ 0 \\ 1 \\ 0 \\ 0 \\ . \\ 0 \end{bmatrix}. \tag{4.19}$$

The 1 at position i specifies an outlier of size ∇ for observation i. The normal procedure is to test all observations by increasing i from 1 to m (the number of observations), which is referred to as data-snooping. If the covariance matrix of the observations Q_y is diagonal, which is assumed here, the test statistic is given by

$$\underline{T}_{q=1} = \frac{\hat{e}_i^2}{\sigma_{e_i}^2}. \tag{4.20}$$

The critical value for this test is obtained from the chi-squared distribution with one degree of freedom. The alternative hypothesis testing for an incoherent point is specified as

$$C_q = \begin{bmatrix} 0\ 0\ \dots \\ \vdots\ \vdots\ \dots \\ 0\ 0\ \dots \\ 1\ 0\ \dots \\ 0\ 1\ \dots \\ 0\ 0\ \dots \\ \vdots\ \vdots\ \dots \\ 0\ 0\ \dots \end{bmatrix}, \tag{4.21}$$

where the ones correspond to observations that are connected to the point that is tested. This is a $(a-1)$-dimensional test, with a the number of connections to the point. One connection is required to fix the height (DEM error) of the point, see also, e.g., (Verhoef, 1997). If a point is connected by only two arcs this test reduces to the outlier test, Eq. (4.19). This alternative hypothesis is specified for each point. The test statistic $\underline{T}_{q\,=\,a-1}$ must be computed with Eq. (4.15).

If the overall model test is rejected alternative hypotheses are specified for each arc and for each point, and the corresponding test statistics are computed. In order to decide which alternative hypothesis is the most likely cause of the rejection these tests of different dimensions need to be compared. The alternative hypothesis with the largest test quotient $\underline{T}_q/\chi^2_{\alpha(q)}$ is selected as the most likely cause for the rejection (also see Appendix B)

$$\frac{\underline{T}^k_{q_i}}{\chi^2_{\alpha(q_i)}} > \frac{\underline{T}^l_{q_j}}{\chi^2_{\alpha(q_j)}} \qquad \forall\ k \neq l. \tag{4.22}$$

Adaption

After the most likely alternative hypothesis is identified as cause for rejection of the null-hypothesis, the functional model is adapted accordingly. If an outlier arc is identified it is removed from the reference network by deleting the appropriate row in Eq. (4.11). If a point is identified, it is removed from the vector of unknowns, together with all arcs connected to that point. The latter could lead to the formation of two isolated networks, which yields a singular problem that cannot be solved without introducing further constraints, e.g., selection of a second reference point. After this adaption, the estimation is performed again, and the DIA procedure starts the next iteration with the detection step.

Joint tests

For the purpose of finding incorrect estimates it is assumed that there is no correlation between the estimated parameters (in this example DEM errors and displacement rates). However, estimated parameters at arcs that have a common point are correlated, and also the estimated DEM error and displacement parameters at a single arc. This correlation depends on the distribution of the images in space and time. The justification for neglecting these correlations is that it is expected that this has no severe impact on the ability of finding the incorrectly estimated arcs. Recall that the misclosures must be perfectly zero for the problem at hand, as described at the beginning of this section. Moreover, the correlation between the estimated parameters is assumed to be small.

Since it is assumed that there is no correlation between the estimated parameters, the adjustment and testing for each parameter can be performed independently. This has the advantage that the factorization of the normal matrix can be re-used, and that the system of equations is a factor two smaller. This is particularly advantageous regarding the relatively large matrix $Q_{\hat{e}}$, which is required to compute the test statistic $\underline{T}_q$, see Eq. (4.15). However, the parameters could be estimated at the same time using a design matrix of the form

$$C' = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}, \tag{4.23}$$

where the sub-matrix C is defined in Eq. (4.11). It can be easily shown that the parameter solution for this model is equal to that of the adjustment in two separate computations. A large alternative hypothesis can be written similarly as

$$C_q{}' = \begin{bmatrix} C_q & 0 \\ 0 & C_q \end{bmatrix}, \tag{4.24}$$

where C_q is one of the previously defined alternative hypotheses, for example the conventional alternative hypothesis defined in Eq. (4.19). The dimension of this test statistic is two times that of the original one. The corresponding

critical value for this test is thus found under the chi-squared distribution with dimension two times that of the individual alternative hypotheses. It can be easily shown that the test statistic using the alternative hypothesis Eq. (4.24) is equal to the sum of the independently computed q-dimensional tests using Eq. (4.15)

$$\underline{T}_{2q} = \underline{T}_q + \underline{T}_q. \tag{4.25}$$

This is the case because it is assumed that there is no correlation between the estimated parameters. This joint test is used in order to identify outlier arcs or points. During each iteration of the DIA procedure the most likely alternative hypothesis is identified based on the test quotient of the joint test. In the adaption step, the concerned observations in the networks for DEM error and displacement rate are removed. Extension to more than two estimated parameters is straightforward.

4.5 Estimation of points relative to the reference network

Once the reference network is established and the DEM error and displacement parameters at the points in the reference network are computed and tested, the other points that are selected according to Eq. (4.1) can be estimated. This estimation is again performed with the integer least-squares estimator, identically to the estimation at the arcs of the reference network. Each new point is connected to the nearest point of the reference network and the wrapped phase difference time series is used to estimate the DEM error and displacement parameters differences between the reference point and the new point, see also Fig. 4.9. The parameters of the reference points are considered to be deterministic at this moment, which implies that the parameters of the new points are simply given by addition. The distance of the newly computed points to the points of the reference network is smaller than the typical distance between the points of the reference network. The atmospheric phase is thus also expected to be smaller. However, the random noise component is likely larger for the new points, because the reference network points were selected as the points with the smallest amplitude dispersion index. The same vc-matrix of Eq. (2.51) and estimated variance components are used during this estimation, see also section 4.3. The functional model is given in, e.g., Eq. (3.11), where integer ambiguities, DEM error, displacement rate, and azimuth sub-pixel position are estimated. To deselect incoherent points, an a posteriori variance factor is estimated for each arc (point) as

$$\underline{\hat{\sigma}}_x^2 = \frac{\underline{\hat{e}}^* Q_y^{-1} \underline{\hat{e}}}{r}, \tag{4.26}$$

where $\underline{\hat{e}}$ are the temporal least-squares residual phase (differences) vector, and r is the redundancy (i.e., the number of interferograms minus the number of

estimated parameters). An estimated variance factor of 1.0 signifies that the vc-matrix used during the least-squares adjustment correctly describes the dispersion of the observations, while a value of 2.0 would signify that the stochastic model is a factor two too optimistic (assuming that the functional model is correct). These a posteriori variance factors thus scale for each point the estimated components of the variance components stochastic model, given in Eq. (2.51).

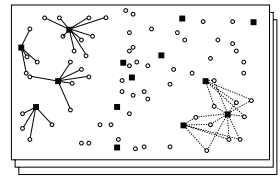


Fig. 4.9: Connection of new points (circles) to the reference network pixels (squares). Each selected point is connected to the nearest point in the reference network, and a relative estimation is performed using the integer least-squares estimator. These connections are indicated at the left hand side of this figure. Alternatively, each new point can be connected to a couple of points of the reference network, indicated at the right hand side.

If only a small number of interferograms is available, say less than fifteen, then the temporal least-squares residuals will always be relatively small, even when the phase is fully decorrelated. This is the case since some combination of ambiguities likely will give a relatively good fit with the model, also when these are not the correct ambiguities. The more the observations deviate from the model (due to noise), the higher the probability that several sets of ambiguities have an almost equally small norm. This depends on the size of the search space, in addition to the quality of the data and the correctness of the functional model. For example, if the search space would not have a bound at all, then there would always be a combination of integers ambiguities that would have zero residuals (because the problem is under-determined). In the integer least-squares estimation, the size of the search space is controlled by the variance of the pseudo-observations that are introduced, see Chapter 3. Therefore, the smaller the number of images that is available, the smaller this variance should be.

In order to obtain a more reliable estimate for the parameters of the new points, the new point could be connected to more than one reference point. In Fig. 4.9 this is indicated at the right hand side. The system of observation equations can be formed using

$$\begin{aligned}
\underline{\Phi}_x^k - \underline{\Phi}_R^k &= \beta_x^k \Delta h_x - \beta_R^k \Delta h_R \;\rightarrow W\{\underline{\Phi}_x^k - \underline{\Phi}_R^k + \beta_R^k \Delta h_R\} = 2\pi \cdot a_x^k + \beta_x^k \Delta h_x \\
\underline{\Phi}_x^k - \underline{\Phi}_P^k &= \beta_x^k \Delta h_x - \beta_P^k \Delta h_P \;\rightarrow W\{\underline{\Phi}_x^k - \underline{\Phi}_P^k + \beta_P^k \Delta h_P\} = 2\pi \cdot a_x^k + \beta_x^k \Delta h_x,
\end{aligned} \tag{4.27}$$

where only two reference network points R and P, and only the DEM error Δh, are considered for conciseness. The ambiguities are denoted by a, see Chapter 3 for a detailed description. The system of equations is more stable if more connections are used, because it is more likely that the parameters at the different arcs correspond with the a priori information. The fact that an incorrect set of ambiguities for a certain arc fits well with the model due to noise, is reduced by the second arc (or third, etc.). This system of equations can be solved using the integer least-squares estimator. However, note that in this case the reference points are demodulated for their known signal components, and that the unknown parameters of the new point are thus directly estimated with respect to a zero reference. This means that the value of the estimated parameters is likely to be larger than it would be when a relative estimation is performed, which should be taken into account in the values or the variances used for the pseudo-observations. Alternatively, the phase of the new point could be demodulated using the mean of the known parameters at the reference points, which then has to be added again when the ambiguities are estimated.

4.6 Final estimation

In the reference PS technique, the final estimation step is the inclusion of more points after subtraction of an estimated APS. Points with an ensemble coherence above a threshold are selected as reliable points, which is the final output. In the STUN algorithm an explicit unwrapping of the phase data is performed. This step follows after the inclusion of more points, already described in the previous section, and thus is an additional step compared to the reference PS technique.

Only points that have a small residual phase with respect to the (temporal) model should be used for the spatial unwrapping. The reason is that points that are incorrectly unwrapped in time can hardly be distinguished from correctly estimated points after addition of the unwrapped residual fields, because the temporal least-squares residuals are $\in [-\pi, \pi)$. by definition, while the spatially unwrapped residuals can be considerably larger due to, e.g., atmospheric artifacts. For example, consider a point that is fully coherent (and which does not undergo displacement), far away from the reference point, and a point that is fully incoherent, near to the reference point. Both points are temporally unwrapped by the ILS estimator with respect to a coherent reference pixel. The parameters that are estimated for the incoherent point are random, but they do minimize a norm with respect to the model. The wrapped residual phase differences of Eq. (4.26) are much smaller for the coherent point, as is the estimated variance factor. However, after the spatial unwrapping process, the unwrapped residuals at the incoherent point could be smaller due to its proximity to the reference point (and consequently the smaller expected atmospheric phase). Therefore, only pixels with an a posteriori

variance factor, cf. Eq. (4.26), below a certain threshold k are included in the spatial unwrapping,

$$\hat{\underline{\sigma}}_x^2 < k, \tag{4.28}$$

with, for example, $k = 2.0$. The value of this threshold depends on the vc-matrix used during the ILS estimation, because the estimated variance factor is a multiplication factor for this matrix.

Spatial phase unwrapping

After the parameters are estimated, the unwrapped model phase can be computed for each point in all interferograms, using the forward model, as

$$\hat{\underline{\Phi}}_x = B\hat{\underline{b}}_x. \tag{4.29}$$

Since the observed phase is wrapped, only the wrapped residual phase can be obtained, i.e.,

$$W\{\hat{\underline{e}}_x\} = W\{\underline{\phi}_x - \hat{\underline{\Phi}}_x\}. \tag{4.30}$$

The residual phase at the selected points in each interferogram are expected to contain a low-frequency component caused by interferometric atmospheric signal and possible unmodeled displacement, and a small amplitude, high-frequency component due to random noise. This property can be exploited by application of a spatial complex low-pass filter to the wrapped residuals. The residual phase per interferogram can then first be demodulated for the low-pass component, which can then be unwrapped separately from the high-pass component. The total unwrapped field is given by addition of the two unwrapped components. However, to unwrap the residual fields for each interferogram, a sparse grid MCF unwrapping algorithm (Eineder and Holzner, 1999) is directly applied in the STUN algorithm. The distance between the points is used to generate the cost function. Once the residual fields are unwrapped, the unwrapped phase at the selected point is obtained by addition of the unwrapped residual phase $\check{\underline{e}}_x$ to the model phase

$$\check{\underline{\Phi}}_x = \hat{\underline{\Phi}}_x + \check{\underline{e}}_x. \tag{4.31}$$

The unwrapped phase at a reference point must be set to zero in all interferograms. This reference point does not have to be identical to the previously selected one. However, it must be a point that has a relatively small noise component in all interferograms, i.e., the reference point must be present during all acquisitions.

Estimation using unwrapped data

There are many possibilities to estimate the parameters once the phase data at the selected points are unwrapped in all interferograms. For example,

a more complex displacement model could be used, e.g., an exponential displacement model which parameters can be geophysically interpreted, or additional parameters could be estimated once the data are unwrapped, e.g., parameters modeling orbit errors, or parameters describing the interferometric atmospheric signal. Phase components that are expected to have only a small contribution to the interferometric phase, such as the average interferometric atmospheric signal or phase caused by azimuthal sub-pixel position (assuming relative small Doppler centroid differences between acquisitions), could be regarded as a stochastic variable during the estimation using wrapped data, allowing for a faster estimation, while significance tests could be performed using the unwrapped data. Furthermore, additional a priori information could be introduced, for example that certain points do not undergo displacement.

On the contrary, less parameters could be used compared to the temporal model used to unwrap the data, for example a global space-time displacement model could be introduced at this point. Moreover, the temporal displacement model used to unwrap the data is the same for each point (actually for each difference), which may be an over-parameterization of the displacement. This over-parameterization is required because the observed phase data are wrapped, and it would be unfeasible to perform alternative hypotheses tests for all arcs using the wrapped data. If a model would be used that is too strict, e.g., neglecting seasonal displacement, incorrect parameters could be estimated that better fit the wrapped data. Now the phase data are unwrapped, alternative hypotheses tests for the significance of the computed parameters can be performed, for example using a null-hypothesis assuming no displacement. Using this hypothesis, a set of points could be identified that significantly is stable (that has the smallest test statistic under the null-hypothesis). This implies that the residual phase at these points is mainly composed of atmospheric signal (and phase due to orbit errors), which could be spatially low-pass filtered and interpolated, as is done in the reference PS technique. The advantage of such an approach is that the temporal high-pass filter to remove unmodeled displacement, see also section 2.1.3, would not have to be applied if points can be identified that are clearly stable. Alternatively, the same displacement model can be used to re-estimate the parameters (which should be almost identical to the previously computed parameters), and the non-displacement phase components, i.e., DEM error, could be removed from the unwrapped phase. This would yield a time series evolution of the displacement, where the deviations from the displacement model are caused by random noise, atmospheric phase, and un-modeled displacement. Similar to the filters used in the reference PS technique, a temporal low-pass filter, weighted using the estimated variance components, could then be applied to this time series to estimate the temporally correlated displacement component.

In this study a simple approach is proposed. After the spatio-temporal phase unwrapping, the parameters are again estimated using the least-squares principle. The same displacement model and the same vc-matrix are used as

for the estimation using the wrapped data. This system of equations is thus described by Eq. (3.17), see also Chapter 3.

Quality description

The vc-matrix of the estimated parameters is given by Eq. (3.18), assuming that the a priori vc-matrix of the double-difference observations is known. However, the vc-matrix that is used during the ILS estimation with wrapped data is only valid between nearby points, because the variance factors of this matrix were estimated using the wrapped phase differences between points of the reference network at a typical distance (see section 4.3). The precision of the unwrapped double-difference phase observations can be obtained by estimation of a final a posteriori variance factor $\underline{\hat{\sigma}}_x^2$ for each point x, cf. Eq. (4.26). This factor is computed using the unwrapped least-squares temporal phase residuals with respect to the reference point. The a posteriori variance factor scales the a priori vc-matrix given in Eq. (2.50) as

$$Q_{\mathrm{ifg}_x} = \underline{\hat{\sigma}}_x^2 \cdot Q_{\mathrm{ifg}}. \tag{4.32}$$

The use of a single factor implies that the precision of the double-difference observations scales in the same fashion for all interferograms. This is acceptable since the atmospheric signal is the main reason for the deteriorating precision at points with larger distances to the reference point, and the atmospheric signal is expected to have a power-law behavior, see Eq. (2.28). Such an approach was also suggested by Hanssen (2001), who constructed a generic stochastic model for the phase in the interferogram that can be initialized using a single scaling parameter (that could be initialized using, e.g., GPS observations or analysis of a small area in the interferogram). The propagated vc-matrix of the estimated parameters scales with the same a posteriori variance factor as the vc-matrix of the observations, i.e.,

$$Q_{\hat{b}_x} = \underline{\hat{\sigma}}_x^2 \cdot Q_{\hat{b}}. \tag{4.33}$$

The precision of the estimated parameters at a point x is thus described by this vc-matrix with respect to the reference point. Note that this expression does not describe covariances between estimated points.

5

Synthetic Data Experiments

The main purpose of performing an estimation using simulated data is to ensure that an algorithm functions as expected under controlled circumstances. Since the input signal is simulated, the estimated parameters can be compared with their true values. The true values of the parameters in case of real data are often not known, and, particularly for Persistent Scatterer Interferometry, not with the required precision and spatial resolution to allow for a good comparison. During software development, a standard test scenario is important to compare different versions of the developed software for correctness and performance. However, not all effects and noise sources that are present in real data can be simulated. In this chapter random noise, atmospheric signal, DEM errors and displacement signal is simulated. Each simulation builds upon the previous simulation with increasing complexity.

The simulation scenarios are described in section 5.1. Section section 5.2 describes the estimation when only random noise is simulated. Then, in section 5.3 atmospheric signal is added, and in section 5.4 topographic and displacement signal. Finally, conclusions are drawn in section 5.5.

5.1 Simulation scenarios

To make the simulation realistic, data are simulated for an existing configuration of perpendicular and temporal baselines for the ERS–1 and ERS–2 sensors. Frame 2547, track 165, corresponding to the Berlin area, Germany, is used to select 51 acquisition dates and sensor positions. Fig. 5.1 shows the distribution of the baselines. Using these baselines, input data is simulated at 41143 points, for an area of approximately 10×10 km^2, see also Fig. 5.2. The PS point positions at which the phase is simulated are obtained using a stack of real data, based on the thresholds, Eq. (4.1). The interferometric phase is simulated at the points as

$$\phi_x^k = W\{\phi_{x,\text{topo}}^k + \phi_{x,\text{defo}}^k + \phi_{x,\text{atmo}}^k + \phi_{x,\text{noise}}^k\}, \tag{5.1}$$

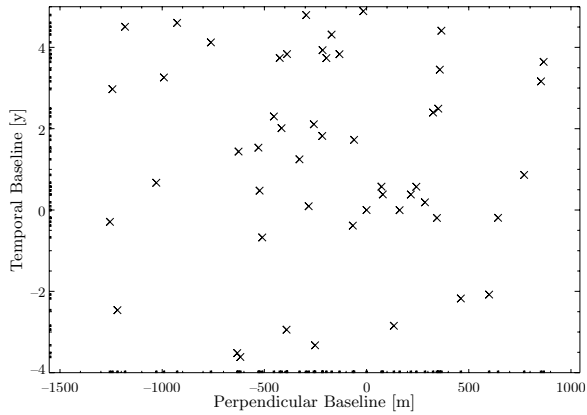


Fig. 5.1: Baseline configuration used in the simulation. The temporal baseline is given on the vertical axis, and the horizontal axis lists the perpendicular baselines, relative to the master acquisition at $(0, 0)$. The 51 selected acquisition times and perpendicular baselines correspond to actual acquired images for ERS frame 2547, track 165.

where the noise and atmospheric parts are simulated on the SLC images. The amplitudes of these pixels in the interferograms are set randomly. The pixels of the reference network are selected based on these amplitudes. A sparsification procedure with a grid of 250 by 250 m is used to select the pixels of the reference network, based on the amplitude dispersion index (i.e., randomly since this is based on the simulated amplitude). The reference network is constructed using the algorithm described in section 4.4 with four connections per point (one in each quadrant). The number of pixels in the reference network is 811, and the number of arcs is 2144. The reference pixel is arbitrarily selected near the Tempelhof airport, see Fig. 5.2.

5.2 Random noise

This section deals with the ϕ_{noise} component. The unwrapped input phase values are simulated as normally distributed noise at the 41143 positions in the 51 SLC images, see Fig. 5.1 and Fig. 5.2. The standard deviation of the noise in the master acquisition is set to $\sigma_{\text{noise}^0}=15°$, while the noise level in the slave acquisitions is set slightly higher to account for coregistration errors. The noise levels of the slave acquisitions are randomly simulated using a mean of twenty and a standard deviation of five degrees, i.e., $\sigma_{\text{noise}^k}=n$, $\bar{n}=20°$, $\sigma_n=5°$, for $k>0$. The DEM error and the linear displacement rate are estimated. The standard deviation used to regularize the system of equations with zero pseudo-observations for the integer least-squares estimator is 25 m and 25 mm/y, respectively. First, the variance components are estimated using

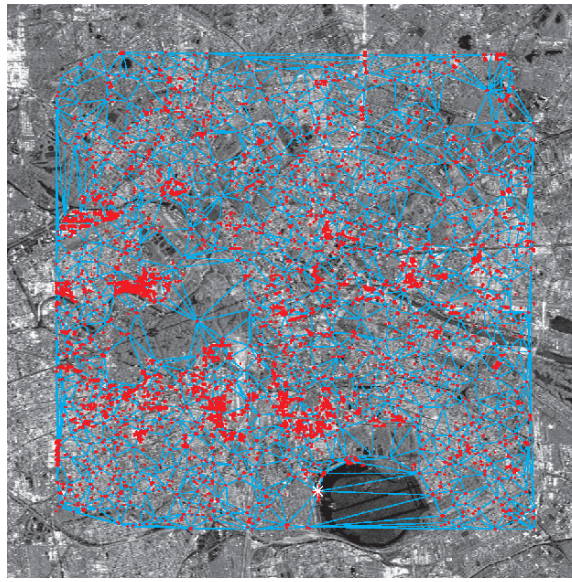


Fig. 5.2: Location of the points used in the simulation (red plus marks). 41143 points are selected in an area of ~10 × 10 km^2. These positions correspond to actually selected positions for acquired images for ERS frame 2547, track 165. The lines indicate the estimations that are performed between the 811 points of the reference network. The background shows the average intensity in dB for this area (Berlin, Germany). This image is mirrored in the horizontal and vertical direction to obtain an almost geo-referenced image. The dark circular structure at the bottom of this image is the Tempelhof airport located at the center of Berlin. The white asterisk indicates the location of the chosen reference point.

the least-squares residual phase (differences) of the estimations defined by the network shown in Fig. 5.3. The (difference) parameters at the arcs of this "network" are initially estimated using a priori variance components, see section 4.3. Then, the variance components are estimated using the least-squares residual phase, cf. Eq. (4.8). The results are shown in Fig. 5.4. The estimated variance components are used to create the vc-matrix for the double-difference phase observations. The a posteriori vc-matrix of the estimated parameters $Q_{\hat{b}}$ is then, see Eq. (3.18),

$$Q_{\hat{b}} = \begin{bmatrix} 0.0288 & -0.0010 \\ -0.0010 & 0.0140 \end{bmatrix}. \tag{5.2}$$

Here, the DEM error is the first estimated parameter in meters, and the displacement rate the second parameter in millimeters per year. The standard deviation of the estimated difference parameters in this case are thus 0.17 m and 0.12 mm/y. The correlation coefficient between the estimated parameters is as small as –0.03, which is caused by the almost uniform distribution of

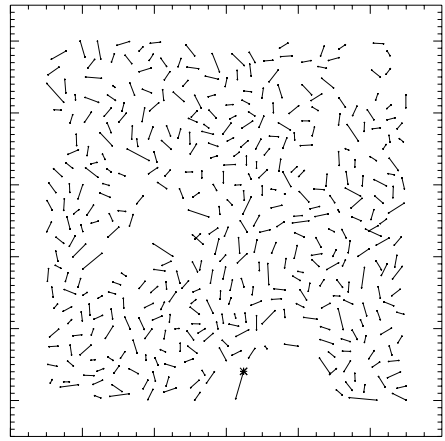


Fig. 5.3: Network for variance component estimation. The indicated arcs are used to perform an initial estimation of the parameters with a priori variance components. The temporal least-squares residuals are then used to estimate the variance components of the variance components model. The mean arc length is 290 m, the standard deviation is 107 m. The total number of arcs used for the estimation is 405. The asterisk indicates the location of the reference point.

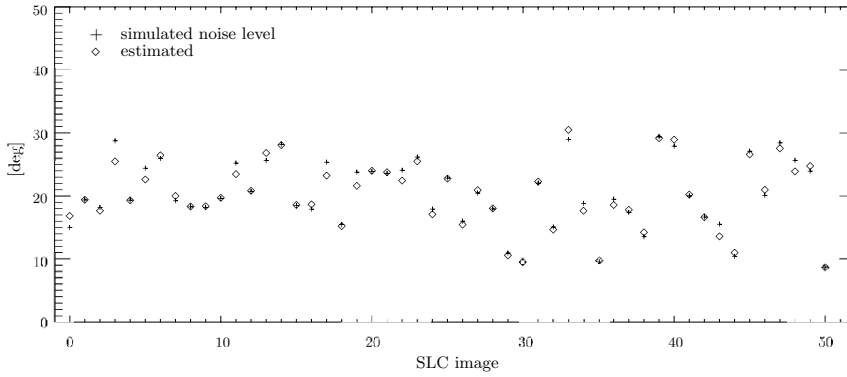


Fig. 5.4: Estimated variance components for random noise compared to simulated noise levels. The plus marks show the standard deviation of the Gaussian distributed noise that is added to the phase in the SLC images. The diamond represents the (square root) of the estimated variance components.

Table 5.1: Statistics for estimated parameters at the arcs of the reference network and for all points connected to the reference network after integration of the parameters. Statistics are for estimates with an estimated a posteriori variance factor $\underline{\hat{\sigma}}^2 < 2.0$ (i.e., 2143 of 2144 arcs for the reference network, 41137 of 41143 estimated points). The standard deviation that follows from the propagated vc-matrix is given in parentheses.

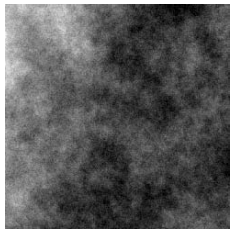
	min	max	mean	std
arcs DEM	–0.59	0.52	–0.0032	**0.166 (0.17)**
arcs defo	–0.35	0.39	0.0005	**0.119 (0.12)**
points DEM	–0.65	0.65	–0.0132	**0.166 (0.17)**
points defo	–0.52	0.52	–0.0046	**0.120 (0.12)**

the images in time and space, see Fig. 5.1. The parameters at the arcs of the reference network are estimated using the integer least-squares algorithm. These difference parameters at the arcs are integrated using the algorithm described in section 4.4.2, and more points are estimated with respect to this reference network. No arcs or points are rejected by the alternative hypotheses tests in this case. For this simple scenario the residuals are not spatially integrated, since it can be assumed that the data are correctly unwrapped temporally. Reported in Table 5.1 is the mean and standard deviation of the estimated parameters. Estimates with an a posteriori variance factor $\underline{\hat{\sigma}}^2 \geq 2$, cf. Eq. (4.26), are not considered. The estimated precision of the accepted estimated parameters corresponds very well with the formally propagated vc-matrix. This is not remarkable, since the simulated signal only consists of normally distributed noise. The theoretical success rate of the simple bootstrap estimator, according to Eq. (3.20), is $P(\underline{\hat{z}} = z) = 0.999$. The theoretical success rate for the simple bootstrap estimator using a priori variance components is $P(\underline{\hat{z}} = z) = 0.843$. The computations are performed using six CPUs operating at 750 MHz. The total CPU time required for the estimation of the 41143 points is 348 seconds, or approximately one minute for each processor. An implementation in C of the bootstrap and integer least-squares estimator is used.

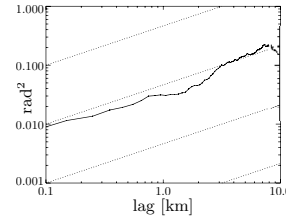
5.3 Atmospheric phase

For the second scenario, the same simulated noise is used as in section 5.2. On top of this, atmospheric signal is simulated at the acquisition times. Again, the DEM error and the linear displacement rate are estimated while no signal is simulated. The goal of this scenario is to demonstrate the quality estimation of the network algorithm. Since there is no signal, most ambiguities are equal to zero and this scenario thus cannot be used to demonstrate the correctness of the ILS estimator. The atmospheric phase is simulated using fractal surfaces

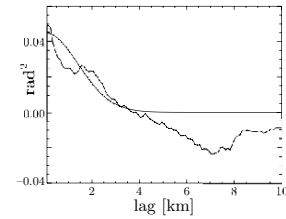
with a dimension 2.67, typically for atmospheric signals, see for example (Hanssen, 2001). The maximum variation of the simulated atmospheric signal during the acquisitions is set randomly with a standard deviation of two rad. The variation of the atmospheric signal in an interferogram is typically about one fringe, since the difference between the atmospheric states during master and slave acquisition is observed. Fig. 5.5 shows an example of the simulated atmospheric signal. The variance components are estimated using



(a) Fractal atmosph.



(b) Structure function



(c) Covariance function

Fig. 5.5: Example of simulated atmospheric signal using a fractal with fractal dimension 2.67. Shown is the simulated master atmosphere. This APS has a maximum variation of 1.72 rad (1.50 rad at the considered points). Fig. 5.5(b) shows the structure function using logarithmic axes. Indicated by the dotted lines is the theoretical slope of 2/3, which follows from the fractal dimension 8/3, see also (Hanssen, 2001). Fig. 5.5(c) shows the empirical covariance function for this APS. This covariance function is modeled using an exponential model. The covariance typically gets negative at a certain distance, since the simulated atmosphere has most power in the long wavelengths. For the estimation of the covariance and structure function 5000 randomly selected points are used.

the same network as before, see Fig. 5.3. The estimated variance components, noise level, and variation of the simulated APS are shown in Fig. 5.6. Clearly visible is that the estimated variance components are larger if the atmospheric variation is larger. This is expected, since the estimated variance component accounts for both components. The variance components previously estimated in the scenario without atmospheric signal are also plotted in this figure for reference. In general, they are smaller than the variance components estimated here, as expected. The vc-matrix of the estimated parameters is in this case given by

$$Q_{\hat{b}} = \begin{bmatrix} 0.0411 & -0.0018 \\ -0.0018 & 0.0180 \end{bmatrix}. \tag{5.3}$$

This vc-matrix is approximately a factor 1.4 larger, compared to the vc-matrix Eq. (5.2) of the scenario without atmospheric signal, which is due to the atmospheric phase that is added. During the alternative hypothesis testing step of the spatial integration of the estimated parameters, four arcs are removed that clearly are estimated incorrectly. Table 5.2 reports the

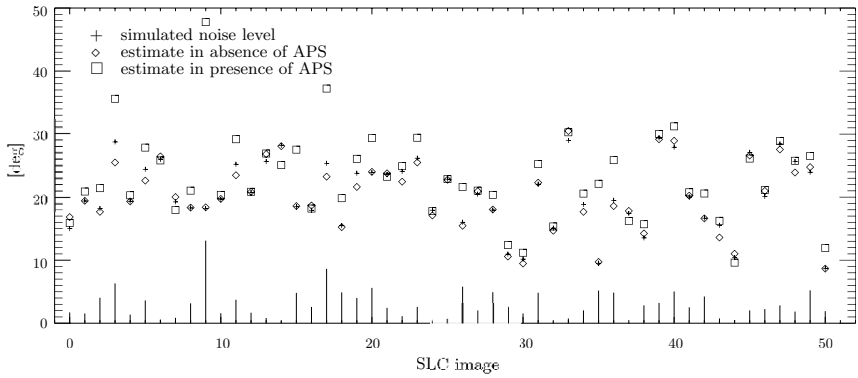


Fig. 5.6: Estimated variance components for random noise and atmospheric signal. The plus marks show the standard deviation of the Gaussian distributed noise that is added to the phase in the SLC images. The squares show the estimated variance components using the short arcs. On the bottom, the maximum variation of the atmospheric signal is plotted for each SLC image. For comparison reasons, the diamond again shows the previously estimated variance components for the simulation without the atmospheric signal, see section 5.2.

statistics of the estimated parameters. Estimates with an estimated variance factor $\underline{\hat{\sigma}}^2 \geq 2$ are excluded from the statistics. Such parameters can have large estimated values, in this case up to 102 m estimated DEM error. This suggests that the estimates themselves could also be used to identify unreliable estimates. The standard deviation of the estimated parameters

Table 5.2: Statistics for estimated parameters for the simulation scenario including atmospheric signal. Given are the estimates at the arcs of the reference network and for all points connected to the reference network after integration of the parameters. Statistics are for estimates with an estimated a posteriori variance factor $\underline{\hat{\sigma}}^2 < 2.0$ (i.e., 2138 of 2144 arcs for the reference network, 41139 of 41143 estimated points). The standard deviation that follows from the propagated vc-matrix is given in parentheses.

	min	max	mean	std
arcs DEM	–0.66	0.80	0.0094	**0.211 (0.20)**
arcs defo	–0.48	0.58	–0.0014	**0.143 (0.13)**
points DEM	–0.78	0.72	–0.0326	**0.187 (0.20)**
points defo	–0.50	0.42	–0.0296	**0.111 (0.13)**

again corresponds very well with the formally propagated vc-matrix. This is somewhat unexpected, since the power of the atmospheric signal is expected to increase the further away the points are from the reference point, while the

propagated vc-matrix is valid only for relatively nearby points. However, the estimated parameters are spatially correlated (not shown). Fig. 5.7 shows the covariance function of the estimated displacement rate. The theoretical success

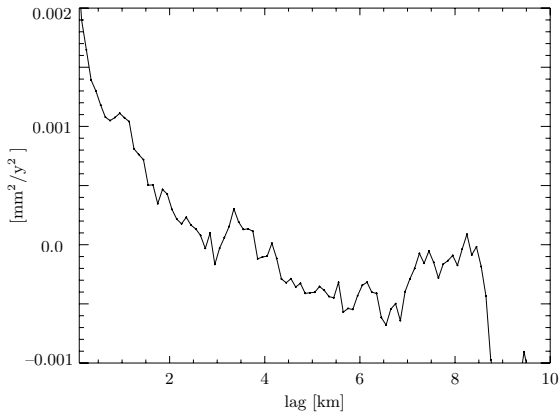


Fig. 5.7: Covariance function of the estimated displacement rates in the presence of atmospheric signal. The estimated parameters are spatially correlated, because the atmospheric signal is spatially correlated.

rate of the simple bootstrap estimator for this case of satellite distribution and estimated precision, according to Eq. (3.20), is $P(\underline{\hat{z}} = z) = 0.990$. The total CPU time required for the estimation of the 41143 points is 340 seconds, practically equal to the previous simulation scenario.

Final estimation using unwrapped data

The residual phase at the points is unwrapped using a sparse grid MCF algorithm, see (Eineder and Holzner, 1999). Since, in this case, the displacement model fully describes the actual displacement (indeed over-parameterizes it), it would be feasible to estimate the low frequency atmospheric component by applying a spatial low-pass filter to the residuals. For this study, the parameters are simply re-estimated, but now using the unwrapped data. Also a variance factor is estimated for each point, cf. Eq. (4.26). This a posteriori factor describes the precision of the estimated parameters, taking into account the atmospheric and random noise component, but assuming a correct displacement model. The estimated variance factors are shown in Fig. 5.8. The further away from the reference point, the worse the precision. Fig. 5.7 shows the covariance function of the estimated displacement rate. In order to verify the description of the precision, Table 5.3 reports the percentage of points for which the actual error e is outside the given confidence interval. Practically all estimates are within the two-sigma level.

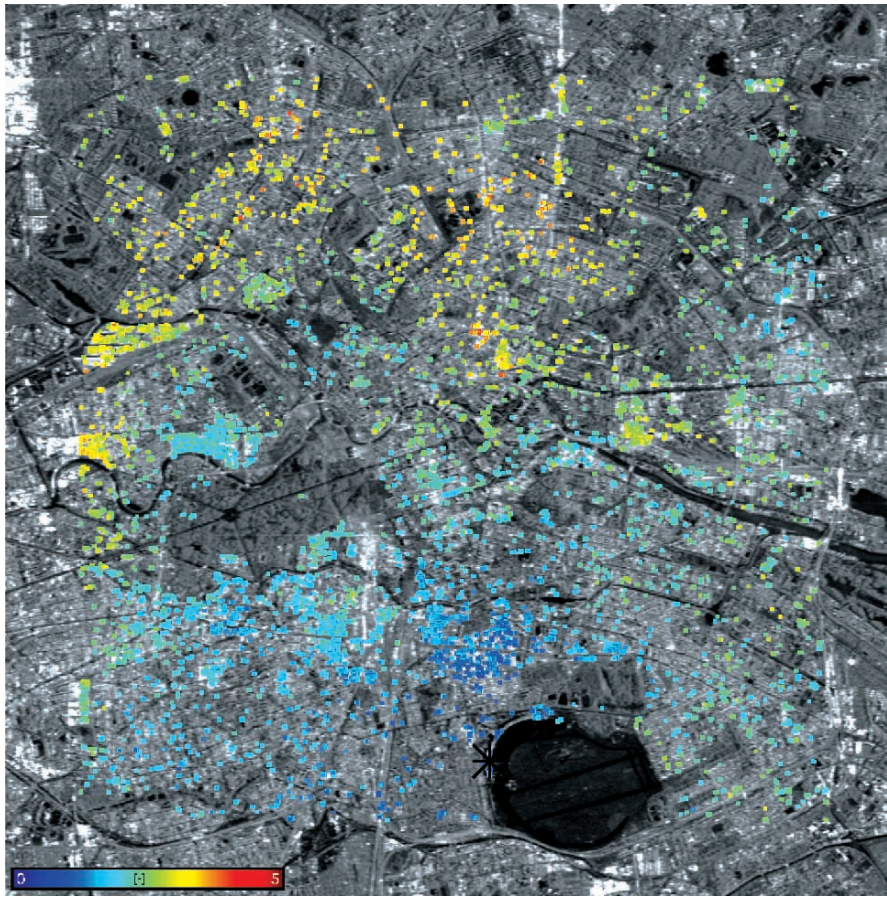


Fig. 5.8: A posteriori variance factors in presence of atmospheric signal. The precision of the estimated parameters (with respect to the reference point, indicated by the asterisk near Tempelhof airport) decreases the further away from the reference point. Plotting range between 0 and 5.

Table 5.3: Quality description of estimated parameters for simulation with atmospheric signal. Reported is the percentage (of 41143 points) for which the actual error on the estimated parameters is below the given threshold.

	$\lvert e_x\rvert < 0.5\hat{\underline{\sigma}}_x$	$\lvert e_x\rvert < \hat{\underline{\sigma}}_x$	$\lvert e_x\rvert < 2\hat{\underline{\sigma}}_x$	$\lvert e_x\rvert < 3\hat{\underline{\sigma}}_x$
DEM	54.7	87.0	99.8	100.0
defo	60.8	90.0	99.7	100.0
both	33.3	78.3	99.5	100.0

5.4 DEM error and displacement signal

In this section the most realistic simulation is described. All components of Eq. (5.1) are simulated. The simulated noise of section 5.2 and the atmospheric signal of section 5.3 are used. The uncompensated topographic phase ϕ_{topo} is computed according to Eq. (2.12), using a simulated DEM error that has a uniform distribution between –30 and 30 m. The simulated DEM error is assumed to be reasonably realistic for urbanized areas, since it accounts for the elevation with respect to the reference height of the actual position of the PS point, which can be at rooftops, window ledges, etc. In practice this obviously also depends on whether the reference surface that is used during the differential interferometric processing refers to the topography of the landscape (DEM, DTM), or whether it includes the height of man-made objects and canopy (digital surface model, DSM, sometimes also referred to as DEM). Using a highly precise DEM thus does not necessarily decrease the DEM error that is estimated in PS InSAR, since the location of the scatterers is unknown. However, in practice one may expect to see a two-peaked histogram for the estimated DEM error, since most PS points are likely located either on the ground, or on top of buildings, which are likely to have a typical height. The height to phase conversion factor β^k is set to a constant value for each interferogram. The interferometric displacement phase is modeled using Eq. (2.13). The surface displacement in the line-of-sight since the master acquisition is described with a function

$$d(T, x, y) = 100 \cdot f(T) \cdot g(x, y), \tag{5.4}$$

with

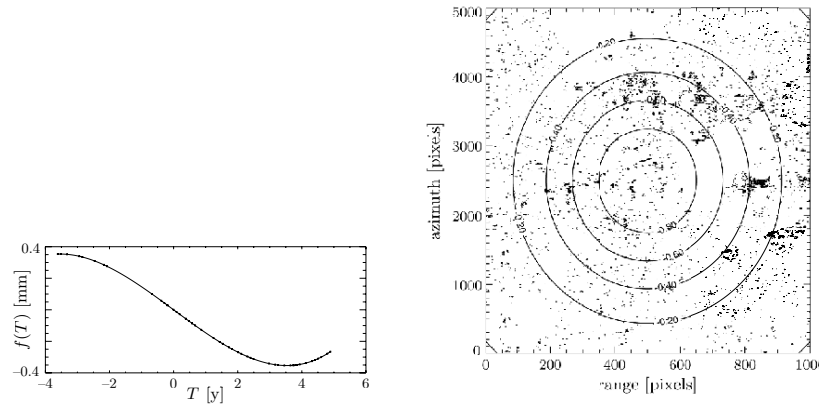
$$f(T) = \text{-}0.15 \cdot T + 0.004 \cdot T^3, \tag{5.5}$$

and

$$g(x, y) = -\exp\Big(\frac{x^2 + y^2}{-(N_\eta/3)^2}\Big), \tag{5.6}$$

where x, y are spatial coordinates, specifically $x = \eta - N_\eta/2$, $y = (\xi - N_\xi/2)/5$, with $N_\eta = 1000$ and $N_\xi = 5000$ the number of pixels in range and azimuth, respectively. Fig. 5.9 shows this spatio-temporal displacement function. As can be seen, the spatial displacement pattern is very smooth, with the maximum located at the center of the area. For the center point, the subsidence rate first increases to approximately 7 cm in six years, followed by ~1 cm uplift in two years. Additionally, a fully decorrelated phase is simulated for 2.5% of the points, i.e., at 1029 points. During the alternative hypothesis testing step of the network algorithm these incoherent points must be detected and removed from the reference network. Finally, also a random bias is added to each SLC image to account for inaccurate knowledge of the absolute signal delay and sensor position.

In the following, first base functions are used that are capable of correctly modeling the simulated displacement at the points and between points,



(a) Temporal displacement function **(b)** Spatial displacement function

Fig. 5.9: Simulated signal. Fig. 5.9(a) shows $f(T)$, the temporal model of the displacement function, and Fig. 5.9(b) the spatial model $g(x, y)$. Indicated are the acquisition times and position of the PS points.

cf. section 5.4.1. Second, section 5.4.2 describes the processing using only a single base function for linear displacement, as is routinely used in the PS technique. Since during the estimation differences between nearby points are considered, and the spatial displacement pattern is very smooth, it can be expected that the simulated displacement can be adequately modeled by linear displacement. That is, the phase data at the PS points could potentially be correctly unwrapped using a linear displacement model and the spatio-temporal unwrapping steps of the network algorithm. However, the estimated parameters for the linear displacement will not correspond to Eq. (5.5). The proper way to assess the success of the method is thus a comparison between the simulated unwrapped phase values and the estimated unwrapped phase, and not by using the standard deviation of the estimated parameters. After all, once the data are correctly unwrapped, the displacement model that is used during the unwrapping is no longer relevant. However, the estimated DEM error is used to validate the estimation. Since the DEM error and displacement parameters are to a great extent uncorrelated (depending on the space-time distribution of the acquisitions, i.e., it is assumed that the displacement is a smooth function in time, and not smooth as function of the perpendicular baseline), a correctly estimated DEM error does not necessarily imply that the displacement is correctly estimated.

5.4.1 Estimation using the correct base functions

The first analysis of the simulated data is performed using base functions that are capable of exactly reconstructing the simulated displacement. These base

functions are

$$\begin{aligned} p_1(k) &= -\frac{4\pi}{\lambda} T^k \\ p_2(k) &= -\frac{4\pi}{\lambda} (T^k)^2 \\ p_3(k) &= -\frac{4\pi}{\lambda} (T^k)^3. \end{aligned} \tag{5.7}$$

The vc-matrix of the estimated parameters (DEM error, displacement parameters) is

$$Q_{\hat{b}} = \begin{bmatrix} 0.084 & 0.004 & 0.006 & -0.001 \\ 0.004 & 0.252 & 0.025 & -0.018 \\ 0.006 & 0.025 & 0.013 & -0.003 \\ -0.001 & -0.018 & -0.003 & 0.002 \end{bmatrix}, \tag{5.8}$$

where a priori variance components are used, see also Fig. 5.3. The corresponding correlation matrix ρ is given by

$$\rho = \begin{bmatrix} 1.000 & 0.026 & 0.187 & -0.072 \\ 0.026 & 1.000 & \mathbf{0.447} & \mathbf{-0.885} \\ 0.187 & \mathbf{0.447} & 1.000 & \mathbf{-0.713} \\ -0.072 & \mathbf{-0.885} & \mathbf{-0.713} & 1.000 \end{bmatrix}. \tag{5.9}$$

Clearly, the second and third estimated parameter are very correlated with each other and with the first base function. Still, the correlation coefficients between the estimated DEM error and the displacement parameters are small. However, for this estimation the base functions Eq. (5.7) are used. The standard deviation for the pseudo-observations used by the integer least-squares estimation is 25 mm/y for α_1, 5 mm/y^2 for α_2, and 1 mm/y^3 for α_3. These values result in the same magnitude of the displacement for $T = 5$ year for all three base functions.

Reference network computation

The estimated variance components per SLC image are practically identical to the variance components estimated in the previous simulation scenario (which are shown in Fig. 5.6). These variance components are used to create the vc-matrix for the integer least-squares estimation to obtain the parameters between the points of the reference network. After the ILS estimation, eighteen points are removed from the reference network. These points are clearly incoherent, since the mean of the estimated variance factors of the connecting arcs to these points is larger than three, cf. Eq. (4.26). This pre-processing step is performed in order to save time during the alternative hypothesis tests. During the spatial integration step of these parameters, additionally two points and sixteen arcs are removed based on the DIA testing procedure. In total twenty of 811 points (2.5%) are detected and removed from the reference network, i.e., all incoherent points are correctly identified.

Final estimation using unwrapped data

The parameters at the remaining points are estimated using a single connection to the reference network, cf. section 4.5. The total CPU time (sum of six CPUs) for the estimation of the 41143 points is 411 seconds. The extra time in comparison with the previous simulation scenarios is caused by the larger amount of noise on the approximately thousand points, which causes the integer least-squares search to take longer. The estimated variance factors of each point are used to select reliable points. Table 5.4 shows the number of points that would be selected if a certain threshold is for the a posteriori variance factor. Here, a threshold $\underline{\hat{\sigma}}^2 < 2.0$ is used, i.e., 39913

Table 5.4: Number of selected points for different thresholds for the a posteriori variance factor. The percentage is relative to the number of coherent points (40114) used in the simulation.

	$\underline{\hat{\sigma}}^2 < 1.0$	$\underline{\hat{\sigma}}^2 < 2.0$	$\underline{\hat{\sigma}}^2 < 3.0$	$\underline{\hat{\sigma}}^2 < 4.0$
#	27761	39913	39945	40217
%	69.21	99.50	99.58	100.26

points of originally 41143 points are selected as reliable points. All 1029 incoherent points are identified, but also 201 coherent points are removed. The wrapped residual phase is computed in each interferogram on the accepted points, which is unwrapped using the sparse MCF algorithm. The DEM error and displacement parameters are finally estimated using the unwrapped data.

Comparison with simulated input

First, the estimated DEM error is compared with the simulated topographic signal. The parameters are estimated with respect to the reference point R of the reference network, for which $\Delta h_R = -0.10$, $\alpha_1(R) = 3.03$, $\alpha_2(R) = 0.0$, and $\alpha_3(R) = -0.081$. The statistics of the difference between the simulated input with the estimated DEM error and displacement parameters are given in Table 5.5. It is interesting to see that all the estimated DEM errors are very small, while some (actually eighteen) estimated displacement parameters are incorrect. This is clearly visible when the minimum and maximum error of the quadratic term are considered, since the simulated input has only linear and cubic terms. The error on the estimated parameters is again spatially correlated due to the simulated atmospheric phase, which cannot be inferred from this table. Though, in general the estimated standard deviation of the error agrees well with the propagated vc-matrix, see Eq. (5.8).

A better measure for the performance of the network algorithm is the error in the unwrapped phase, since it allows for a direct comparison between several estimations using difference base functions. Of 39916 unwrapped phase values in 50 interferograms (nearly two million values), 122 are not identical to the

Table 5.5: Statistics of the error on the estimated parameters at accepted points for the simulation including displacement. The standard deviation that follows from the propagated vc-matrix Eq. (5.8) is given in parentheses.

	min	max	mean	std	
DEM error [m]	–0.78	0.76	0.00	0.194	(0.3)
linear term [mm/y]	–6.73	5.41	–0.23	0.374	(0.5)
quadratic term [mm/y^2]	**–3.43**	**3.41**	–0.029	0.108	(0.1)
cubic term [mm/y^3]	–0.49	0.52	–0.021	0.032	(0.04)

simulated input. These incorrectly unwrapped values are all due to incorrectly estimated parameters at eighteen points. The integer least-squares estimator, i.e., the temporal unwrapping step, found a better fit with the simulated data at these points when the quadratic term is approximately 3 mm/y^2, instead of zero, in combination with also incorrect linear and cubic coefficients. The incorrectly estimated points are spatially separated, and it is by pure chance that this occurs. Note that it would be relatively simple to identify these incorrectly estimated points by using a threshold for the estimated coefficients or for the estimated total amount of displacement. In conclusion it can be said that the network algorithm is successful at 39895 of 39913 points (99.95%).

5.4.2 Estimation using a linear displacement model

The probability of correct temporal phase unwrapping using the ILS estimator is larger when the difference between the displacement model and the actual displacement is smaller. The absolute value of this model error should at least be smaller than π in most interferograms, since otherwise it will be wrapped, preventing the retrieval of the displacement from the observed wrapped phase values. The error is expressed in radians here. For ERS and ENVISAT, π corresponds to ~1.4 cm of displacement. Furthermore, other error sources, such as atmospheric phase and random noise, are ignored. For this simulation, the temporal displacement function, Eq. (5.5), can be approximated reasonably well using a linear displacement rate. Additionally, a linear displacement rate model may yield correctly unwrapped data, since, during the estimation, the displacement between nearby points is considered, and here the spatial displacement pattern is very smooth, see Eq. (5.6) and Fig. 5.9.

One could envision a processing strategy which uses the linear model, or piecewise linear model, during the temporal phase unwrapping step, and thereafter another displacement model during the final estimation using the unwrapped data, for example using alternative hypothesis test to test for significance. For the displacement simulated here, the uplift at the end of the time interval would not be detected if the model would not be changed once the data are unwrapped, since the displacement dominantly consists of subsidence. Note that in the standard PS technique, where a linear

displacement model is routinely used, the danger is that the deviations from the model are interpreted as atmospheric phase. This would certainly happen when a temporal high-pass filter would not be used before applying the spatial low-pass filter to isolate the APS. Moreover, note that this data set contains a temporal gap, and that the uplift occurs at the end of the time interval. Both effects makes the application of a temporal filter cumbersome.

The following describes the estimation using a linear displacement model. The vc-matrix of the estimated parameters (DEM error, linear displacement) is given by

$$Q_{\hat{b}} = \begin{bmatrix} 0.081 & 0.002 \\ 0.002 & 0.037 \end{bmatrix}, \tag{5.10}$$

where a priori variance components are used. The estimated variance components are practically equal for this model as for the higher order polynomial model used in section 5.4.1, indicating that indeed the "actual" displacement between nearby points is well approximated by a linear model. In a pre-processing step, again, eighteen points are removed. The overall model test of the spatial integration of the parameters (DEM error and linear displacement rate) is accepted after removing four more arcs. All incoherent points are removed from the reference network by the alternative hypothesis testing. The threshold for the a posteriori variance factor is again set to $\hat{\underline{\sigma}}^2 < 2$, which for this estimation meant that 40067 points are accepted (all 1029 incoherent points are rejected, as well as 47 coherent points). Using the same threshold for the a posteriori variance factor in this estimation thus yields 154 more points compared to when two more displacement base functions are used. This is explained by Eq. (4.26), since the weighted squared sum of least-squares residuals is divided by the redundancy, which is larger in this case. The unwrapped phase at these points is again obtained by the MCF algorithm. For this estimation, 100 phase values are incorrectly unwrapped, at 9 points (99.98%). Likely, the temporal unwrapping (estimation) is incorrect at these points. The estimation using a linear displacement model thus is able of unwrapping more points correctly than the estimation using the correct base functions. The reason is that the simulated displacement between nearby points can be well approximated by linear displacement. More points are estimated incorrectly when the higher order polynomial base functions are used, since the degree of freedom is larger in that case, i.e., the data can be unwrapped using quadratic and cubic coefficients, leading to a better fit.

5.5 Conclusions

The STUN algorithm performs well using simulated data sets. Data are simulated on 41143 points in 51 acquisitions. The simulated positions of the points and the perpendicular and temporal baselines correspond to an existing data set. It is shown by a simulation using random noise only that the variance components of the stochastic model are correctly estimated, see Fig. 5.4. As

expected, the estimated variance factors increase when atmospheric signal (fractal surfaces with fractal dimension 2.67) is added to the simulated phase—visible in Fig. 5.6—because this signal is considered stochastic. Table 5.3 shows that the precision of the estimated parameters is described well by the a posteriori vc-matrix and estimated variance factors for each point. For the most complex case, the data is nearly perfectly unwrapped (see section 5.4; simulation of displacement, DEM errors, atmospheric signal, and random noise). Only 122 of nearly two million phase values are incorrectly unwrapped when the displacement is modeled using base functions capable of fully reconstructing the simulated displacement signal. This affected the estimated parameters at eighteen points. All ~1000 incoherent points are detected, while only 201 of 41143 points are adversely rejected. The maximum error of the estimated DEM error on the accepted points is ~80 cm. The standard deviation for the error of this parameter is 0.19 cm, while the standard deviation for the error of the estimated displacement rate is 0.37 mm/y. The estimation using only one base function modeling linear displacement performs slightly better. This is explained by the fact that the simulated displacement signal is approximately linear, particularly the difference in displacement between nearby points. The CPU time for the computation of these 41143 points is less than six minutes using a 750 MHz processor for each simulation scenario.

Autonomous movement (i.e., points with locally deviating displacement behavior, for example, points that do not undergo displacement) is not simulated. There is no reason to assume that such points would be estimated incorrectly, at least when their behavior can be described by the chosen base functions and the amplitude of the deviation is in the search space. Furthermore, the construction of the network is not varied. A denser network would not improve the results, and the effect of an increase of the distance between the reference points is described in Chapter 6, using real data. Variation of the number of available images is also described using real data.

6

Real Data Processing

The PS technique is very well suited to the analysis of displacements in urban areas because many man-made constructions are likely to act as stable reflectors. However, the PS technique is also applied to more rural areas, and the patent (Ferretti et al., 2000b) claims application to landsliding areas. Landsliding areas are a severe problem in, e.g., Italy, and are studied in, for example, (Colesanti et al., 2003c; Farina, 2003). Tectonic motion is also successfully studied using the PS technique, see, for example, (Colesanti et al., 2000, 2003c; Hilley et al., 2004; Musson et al., 2004). The study of volcano deformations is attempted, see, e.g., (Salvi et al., 2004). Temporal decorrelation is a major problem to apply conventional interferometric techniques to volcanic areas. The flanks of volcanoes are generally forested, causing temporal decorrelation, except where there are lava flows. The PS technique could possibly offer an alternative by using scatterers such as tree trunks or bare rocks that might act as PS points. However, this does not seem to be the case. Moreover, the displacement model of volcanoes is not necessarily linear. A new processing algorithm based on the concepts of the PS technique that could be applied in such situations was recently introduced in (Hooper et al., 2004).

In this chapter the focus is on the application of the developed STUN algorithm to urban areas. In section 6.1 the Berlin test site is presented. This area is considered to be a typical area with no, or moderate, displacement, well suited to the application of the PS technique. Several experiments are performed to assess the sensitivity of the STUN algorithm to different parameter settings and circumstances. Furthermore, a cross-comparison using data of an adjacent track is made to assess the precision of the estimates. Section 6.2 reports about a study of the city of Las Vegas. This data set is used to study the effect of increasing the number of estimated parameters. The Las Vegas area undergoes severe displacements, partly of seasonal nature, which are modeled using trigonometric base functions. A comparison with the reference PS technique is made and the use of ENVISAT data is demonstrated. Finally, section 6.3 summarizes the main results of the performed experiments.

Application of the STUN algorithm to rural areas is described in, e.g., (Kircher et al., 2003a,b; Walter et al., 2004).

6.1 Berlin

Berlin is chosen as the first test site at the start of the DLR Permanent Scatterer software project, mainly as a validation site for the developed software. Berlin is the largest city and the capital of Germany. It is located at 52°30′ northern latitude, 13°20′ eastern longitude, and lies ~200 km inland at 100 m height (WGS84). The topographic variation is approximately 50 m in the processed area, see also the DEM of the area provided in Fig. 6.1. The urbanized area is approximately 20×20 km^2 and Berlin has a population of ~3.5 million people. Thus, Berlin is a typical urban area that can be well processed using the PS technique. Moreover, a large number of ERS acquisitions are available for Berlin. No significant displacement signal is expected for the Berlin test site. Atmospheric circumstances are typical for a land climate (moderate). The average annual precipitation is 5.7 cm and the average temperature is between 1°C in January and 20°C in July (Berlin Tempelhof station, data 1991–2004, see Deutscher Wetterdienst, 2004). These factors make Berlin a good candidate for an initial test site.

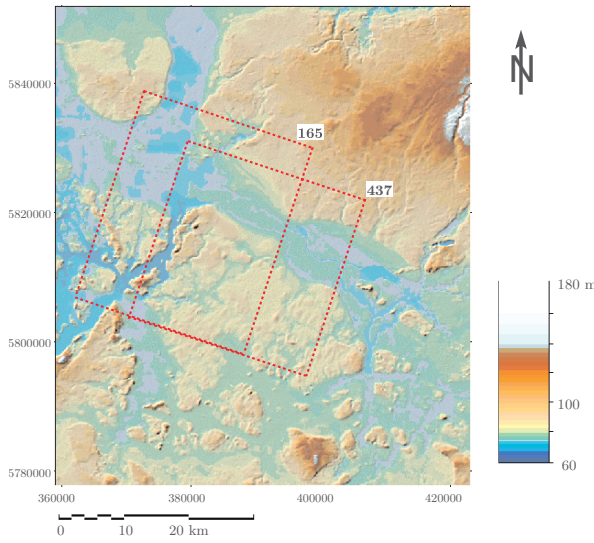


Fig. 6.1: DEM used for the Berlin test site (color shaded). Projection: UTM, zone 33, WGS84 ellipsoid. The rectangles indicate the processed areas for data of track 165 and 437, respectively.

First, the data availability for the Berlin test site is described in section 6.1.1. Next, section 6.1.2 describes a reference processing for the Berlin area where the estimated parameters are DEM error and linear displacement rate. In section 6.1.3 the sensitivity of the STUN algorithm to several algorithm settings is investigated. Finally, section 6.1.4 describes a cross-comparison of the estimated linear displacement rates using a second stack of data from an adjacent track. The experiments that are performed for this test site are summarized in Table 6.1.

Table 6.1: Berlin test site experiments. Listed are the approximate number of interferograms used K, the estimated parameters $\hat{\underline{b}}$, and the purpose of the experiment. The estimated parameters are coded as H for estimated DEM error and V for linear displacement rate. The size of the area for all tests is ~20×20 km^2.

#	K	$\hat{\underline{b}}$	Purpose
I	50	H, V	Reference processing. Distance between points in the reference network is ~1000 m and each point has ~6 connections. (The other scenarios are compared to this processing).
IIa	50	H, V	Sensitivity to the stochastic model.
IIb	50	H, V	Sensitivity to the number and length of the arcs between points in the reference network.
IIc	10	H, V	Sensitivity to the number of available acquisitions.
IId	50	H, V	Sensitivity to the choice of the testing parameters for the arc and point test.
III	30	H, V	Cross-comparison of linear displacement rates estimated using data of an adjacent track.

6.1.1 Data availability

For the Berlin test site data of two descending ERS tracks are available. The first data set, frame 2547 and track 437, is used in the experiments described in the following sections. This data set contains more acquisitions, and has the advantage that the city of Berlin is completely covered. The second data set, frame 2547 and track 165, is used to perform a cross-comparison between the estimated linear displacement rates in the overlapping area. These data and the comparison are described in section 6.1.4. All available acquisitions are listed in Appendix C.

The baseline distribution for the first track is shown in Fig. 6.2. In total, 70 ERS–1 and ERS–2 acquisitions are available. All 69 differential interferograms are generated with orbit 10039 as the master image, acquired at March 22$^{\text{nd}}$, 1997. As an example, fourteen differential interferograms for track 437 are shown in Fig. 6.3. This stack is centered on the city of Berlin, hence the higher coherence in the middle and lower coherence due to temporal decorrelation

at the edges. The interferograms are sorted according to the perpendicular baseline from left to right. Geometric decorrelation can clearly be observed as function of the perpendicular baseline. Note that application of a conventional interferometric processing method on these data is expected to be cumbersome for interferograms with larger perpendicular or temporal baselines because of severe decorrelation.

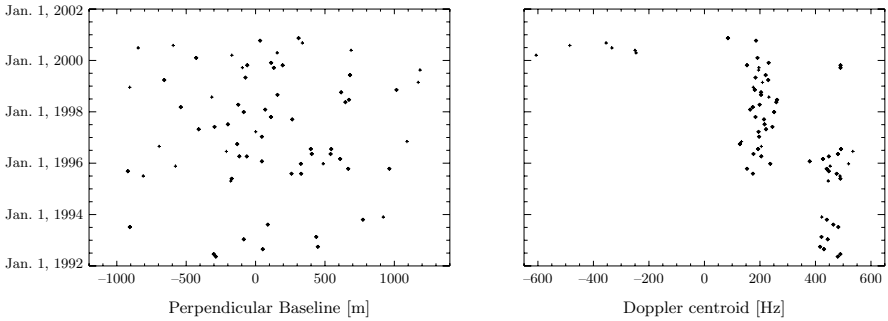


Fig. 6.2: Baseline distribution for available ERS data (70 acquisitions) for the Berlin area (track 165, frame 2547). Doppler centroid frequency for the Berlin data set. The (earlier acquired) ERS–1 images have a consistently larger Doppler centroid frequency than the ERS–2 images. The Doppler centroid frequencies of ERS–2 data acquired after February 2000 is not stable due to gyroscope failures. Selected 50 data for the experiments are shown with a diamond.

6.1.2 Reference processing

In this section a standard processing using the STUN algorithm is described. The estimated parameters are the DEM error and the linear displacement rate. The number of interferograms is limited to 50 during this processing, see also Fig. 6.2. This amount of ERS images is archived by ESA for practically all European and American cities, spanning the time period 1992–2002. The selected points and the reference network for the Berlin test site are shown in Fig. 6.4. There are 78779 points selected which are further analyzed. These points have an SCR larger than two and final selection is done after a point target analysis, see also section 4.4 For the reference network, the grid cell width is set to ∼500 m, resulting in the selection of 1066 points. The network between these points is constructed as described in section 4.4.1 using a maximum of ten connections per point and a maximum distance of 2400 m between points. This procedure results in 6650 arcs with an mean arc length of 970 m and a standard deviation of 370 m. The smallest arc is 123 m and the largest arc is 2396 m. Thus, the final arcs per point ratio of the reference network is 6.2 using these parameters. The variance components of the stochastic model Eq. (2.51) are estimated using an initial estimation

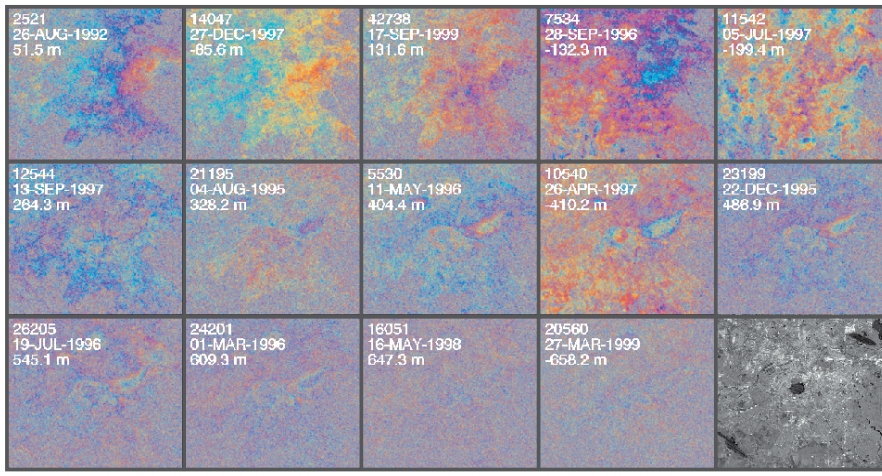
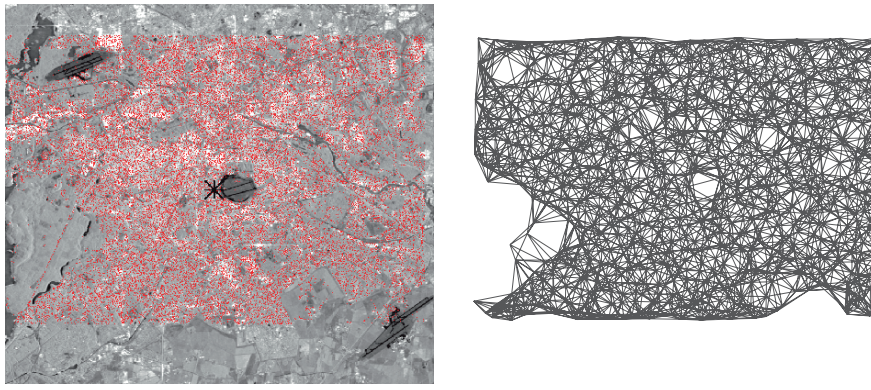


Fig. 6.3: Some differential interferograms for the Berlin area, track 437. The interferograms are sorted from left to right according to (the absolute value of the) perpendicular baseline, $|B_\perp| \in [52, 658]$ m. The master image was acquired at March 22$^{\text{nd}}$, 1997. The last panel shows the average amplitude of the processed area. The images are in the radar coordinate system. In this case, they are roughly geo-referenced after mirroring in the vertical axis.



(a) Selected points **(b)** Reference network

Fig. 6.4: Selected points and reference network for the Berlin test site. The total interferometrically processed area is ~26 km wide by 24 km high. Based on their SCR, 78779 points are selected in an area of ~26 km wide by 18 km high (170 PS/km^2). In the reference network there are 1066 points (2.3 per km^2), and 6650 arcs.

using a priori values, see also section 4.3. They are estimated, cf. Eq. (4.8), at 526 independent arcs between points of the reference network. The mean arc length during this estimation is 528 m with a standard deviation of 153 m. The vc-matrix of the estimated parameters using the a priori stochastic model is given as

$$Q_{\hat{b}} = \begin{bmatrix} 0.123 & -0.004 \\ -0.004 & 0.042 \end{bmatrix}. \tag{6.1}$$

The DEM error is the first parameter and the displacement rate the second, i.e., the estimated standard deviation using this a priori model is ∼0.35 m for the estimated DEM error and ∼0.20 mm/y for the displacement rate between points. The correlation between the estimated parameters can be neglected in this scenario.

Fig. 6.5 shows the estimated variance components of the stochastic variance component model plotted as function of perpendicular, temporal, and Doppler baseline, respectively. The ERS–2 images seem to have a slightly better precision than the ERS–1 images, which could be due to improved hardware and sensor settings. However, errors in the functional model (e.g., the linear displacement model may be too simple, or some points may not be coherent during the earlier acquisitions) also lead to larger estimated variance components for the earlier ERS–1 images. The latter cause is more likely, since the estimated variance components are only larger for the earlier acquired ERS–1 data. Note that the estimation of the variance components could become self-fulfilling, i.e., data that do not fit the functional model are down-weighted, which in turn leads to a better fit with the model. The vc-matrix using the estimated stochastic model is given as

$$Q_{\hat{b}} = \begin{bmatrix} 0.080 & -0.005 \\ -0.005 & 0.030 \end{bmatrix}, \tag{6.2}$$

The standard deviations of the estimated difference in DEM error and linear displacement rate between nearby points are thus estimated to be ∼0.28 m and ∼0.17 mm/y, respectively.

Next, The DEM error and the linear displacement rate (differences) are computed at the arcs of the reference network. The ILS estimator and estimated stochastic model are used for this estimation using wrapped data. The standard deviation for the pseudo-observation to regularize the ILS estimator is set to 25 m for the DEM error and to 10 mm/y for the displacement rate. The theoretical success rate for the bootstrap estimator, cf. Eq. (3.20), in this case is $P(\hat{\underline{z}} = z) = 0.992$. The mean of the estimated DEM error is –0.11 m and –0.06 mm/y for the displacement rate. The mean a posteriori variance factor at the arcs is 1.26, which suggests that the estimated variance components realistically describe the actual precision. The estimated DEM error and displacement rate at the arcs are plotted in Fig. 6.12 (top row).

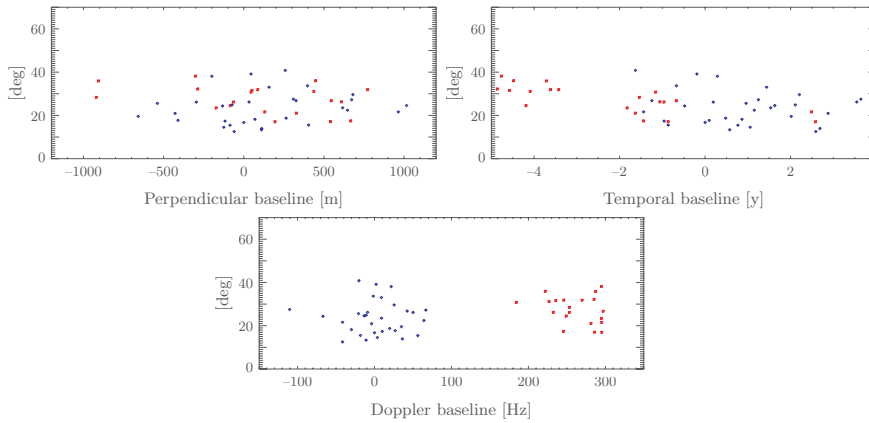
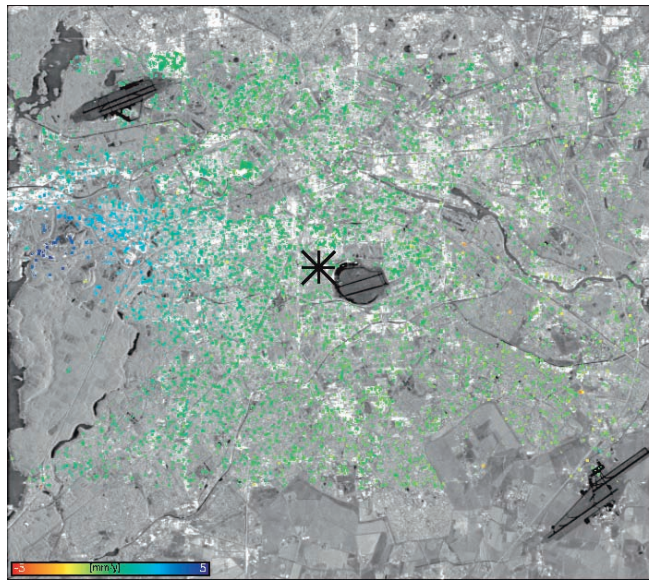


Fig. 6.5: Square roots of the estimated variance components for the Berlin test site as function of perpendicular, temporal, and Doppler baseline. A red asterisk corresponds to an ERS–1, and a blue diamond to an ERS–2 acquisition.

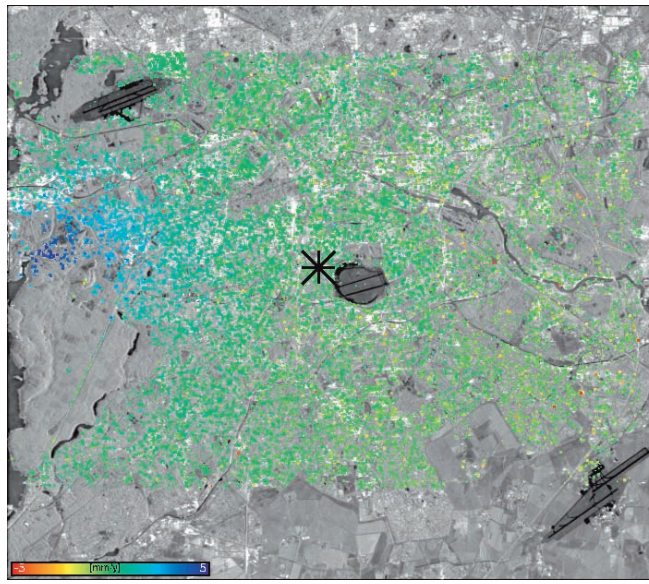
Network integration

After the estimation of the DEM error and the displacement rate at the arcs of the reference network, they are integrated using the least-squares adjustment and testing procedure described in section 4.4.2. A pre-processing step is performed first to save time during the alternative hypothesis step, and to guarantee that clearly incoherent points are removed from the reference network. As visible in Fig. 6.12 (top row), the estimated a posteriori variance factor is less than two for most of the estimated arcs. Therefore, all points are removed for which the mean of the a posteriori variance factors of the connecting arcs is larger than three. Seventeen points are removed of the originally 1066 reference network points. Thereafter, additionally 46 arcs are removed (of 6439 remaining arcs) with an a posteriori variance factor $\hat{\sigma}_x^2 > 3.0$, but only if both connecting points have at least three other connections. For the redundant network used here, each point is still connected with at least eight arcs after this procedure. One additional point and sixteen arcs are removed during the alternative hypothesis testing. The testing parameters used are $\gamma = 90\%$ and $\alpha_1 = 0.05$, see also Appendix B. The mean least-squares residual at the arcs is 0.001 m and –0.001 mm/y for the DEM error and linear displacement rate, respectively. The standard deviation is 0.09 m and 0.11 mm/y, and the maximum absolute error after the hypothesis testing is 1.35 m and 1.78 mm/y. After integration of the parameters, the estimated DEM error at the (remaining 1048) points is between –39.14 and 44.55 m and between –3.48 and 4.34 mm/y for the displacement rate. The estimated parameters are relative with respect to the selected reference point. In this case, the reference point is selected at the center of the image. The mean intensity of the selected point is 9.2 dB and the amplitude dispersion index

is $D_a = 0.12$, i.e., the random noise component of the reference point seems to be small. Note that the least-squares residuals at the arcs are not all exactly equal to zero, which is expected from the theory, and confirmed by the simulation experiments. This could be caused by a small number of points that are partially incoherent, e.g., points that are not visible during all acquisitions. The residual phase at such points during these acquisitions could be around π during the ILS estimation of the parameters at the arcs. This could have lead to small misclosures, although it is unclear why. Another cause could be rounding errors or numerical instability during the adjustment of this relatively large system of equations (designmatrix is $\sim 6000 \times 1000$, computations are performed in IDL, using eight and four byte floating point arithmetic). However, it is considered unlikely that this can cause residuals in the order of one meter, particularly because no instability is reported by the software. By continuation of the alternative hypothesis tests it could be forced that the least-squares residuals become zero at all arcs of the reference network. In the extreme case, the iterations continue until a network results as sketched in Fig. 4.2(b) remains. However, no obvious outlier arcs could be detected anymore, and it is not expected that doing so would significantly affect the final parameter solution using the unwrapped phase. The parameters at the 77731 selected points that are not part of the reference network are now estimated with respect to the nearest point of the reference network using the ILS estimator. The standard deviation for the pseudo-observation to regularize the ILS estimator is set to 25 m for the DEM error and 10 mm/y for the displacement rate, i.e., the same values as those used during the estimation at the arcs of the reference network. The next step of the STUN algorithm is the phase unwrapping at selected (reliably estimated) points. These points are selected based on the estimated variance factor, cf. Eq. (4.26). In this case, a threshold on the estimated variance factor of 1.0 selects points with a variance below 0.080 m^2 and 0.030 mm^2/y^2 for the relatively estimated DEM error and the linear displacement rate, respectively, see also Eq. (6.2). If a threshold of 2.0 is used, these variances would be multiplied by a factor two. Fig. 6.6 shows the estimated displacement rates at the selected points using different thresholds for the a posteriori variance factor. The parameters are estimated using the wrapped data. An uplift area to the west of Berlin can be clearly identified (close to the Olympic Stadium). This uplift was not anticipated since the Berlin test site was not expected to undergo significant displacements. Most likely this uplift is related to underground gas storage. This area was in the news at April 26^{th}, 2004, after a gas explosion occurred (Berliner Zeitung, 2004). The reservoir is located under a densely populated area (particularly Berlin-Charlottenburg and Berlin-Spandau). The reservoir is in use since 1992 and can provide the city of Berlin with gas for one year. Therefore, it is expected that the linear displacement model does not fully describe the actual displacements in that area. Some of the more localized subsidence points visible in Fig. 6.6 are likely to be incorrectly estimated.

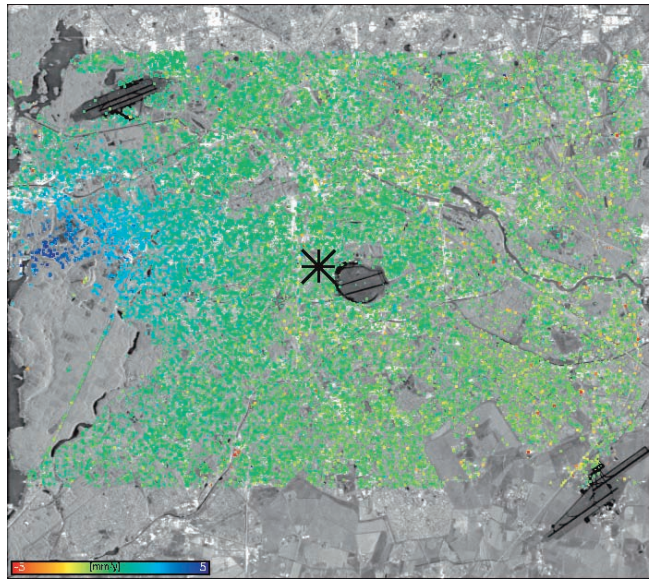


(a) Threshold $\hat{\underline{\sigma}}^2 < 1.0$

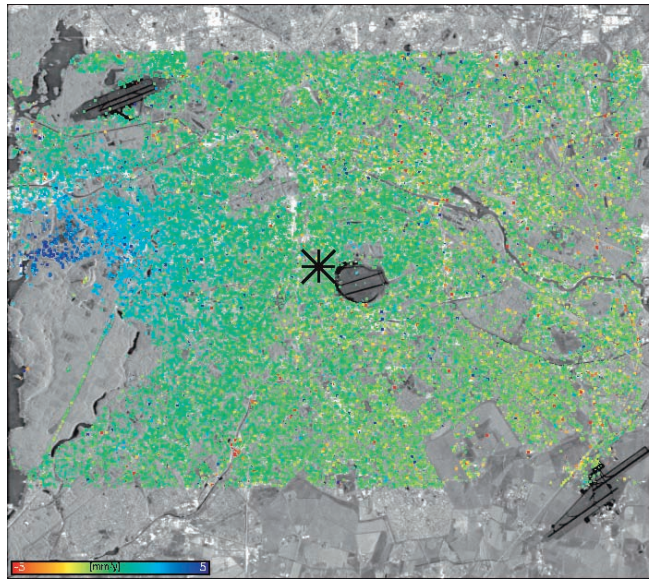


(b) Threshold $\hat{\underline{\sigma}}^2 < 2.0$

Fig. 6.6: Estimated linear displacement rates for the Berlin test site for different thresholds on the a posteriori variance factor. In total 78779 points are computed. Red corresponds to 5 mm/y subsidence and blue to 5 mm/y uplift. **(a)** shows 11078 points with an a posteriori variance factor below 1.0 (estimates between –4.2 and 4.9 mm/y). **(b)** shows 28724 points below 2.0 (estimates between –18.2 and 6.6),



(c) Threshold $\hat{\underline{\sigma}}^2 < 3.0$



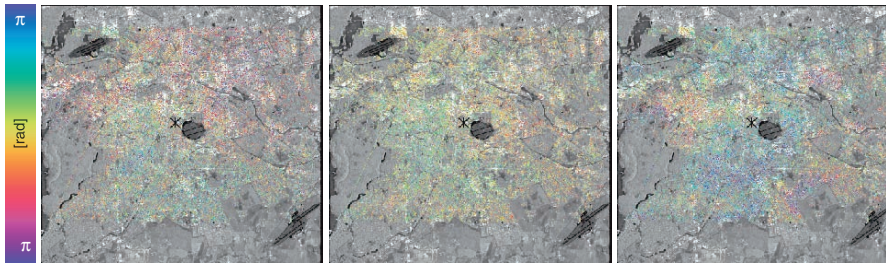
(d) Threshold $\hat{\underline{\sigma}}^2 < 4.0$

Fig. 6.6: (cont.) Fig. **(c)** 42802 points below 3.0 (estimates between –18.2 and 8.9), and **(d)** 54384 points below 4.0 (estimates between –95.6 and 93.1). The location of the reference point is marked by the black asterisk. The (blue) uplift area west of Berlin is the most striking displacement feature. Note that for a larger threshold the amount of incorrectly estimated points increases.

However, the estimates at most of the points plotted in Fig. 6.6(a) and Fig. 6.6(b) are likely to be correct.

Final estimation using unwrapped data

The final estimation is performed using unwrapped data at selected points. A threshold $\hat{\underline{\sigma}}_x^2 < 3.0$ on the estimated a posteriori variance factor is used to select 42802 points (~90 points/km^2). At these points the (wrapped) residual phase is computed in each interferogram by subtracting the model phase (preliminary estimated parameters). The residual phase at the points is then unwrapped using the sparse grid MCF algorithm. The average number of residues (i.e., detected inconsistencies between wrapped and unwrapped phase differences at the arcs of a triangle) in the wrapped residual phase data is ~1%. This indicates that the residual phase at the 42750 selected points in the interferograms is rather smooth. This is expected, since the residual phase is assumed to be composed of (spatially correlated) atmospheric phase and a relatively small amount of random noise. The (wrapped) residual phase of a few interferograms are shown in Fig. 6.7. The residual phase is spatially correlated, i.e., it likely consists of atmospheric signal. However, note that the residual phase seems to be correlated with the uplift area too, particularly visible in Fig. 6.7(a) and 6.7(c). These residuals could possibly be caused by unmodeled displacement, too. If the atmospheric signal is estimated using the plotted residuals by the application of a low-pass filter, this displacement cannot be recovered. A possibility for a more in depth study of this area is to select a nearby reference point, thus largely eliminating the atmospheric signal.



(a) 10330 **(b)** 10540 **(c)** 10831

Fig. 6.7: Residual phase for first three interferograms (sorted numerically according to orbit number). The spatial correlation suggests that atmospheric signal is present in the residual phase. The location of the reference point is marked by the black asterisk. A cyclic colorbar is used (one color cycle corresponds to a 2π phase difference). **(a)** corresponds to interferogram with slave orbit 10330, acquired at July 7$^{\text{th}}$, 1993, $B_\perp = 890$ m, **(b)** to orbit 10540, acquired at April 26$^{\text{th}}$, 1997, $B_\perp = 450$ m, and **(c)** to orbit 10831, acquired at August 11$^{\text{st}}$, 1993, $B_\perp = 85$ m.

The unwrapped phase at the selected points is obtained using the MCF sparse grid unwrapping algorithm. This phase is used to estimate the DEM errors and the linear displacement rates. The parameter solution is visually identical to the solution using wrapped data, see Fig. 6.6. The a posteriori variance factors, estimated using the unwrapped data, are shown in Fig. 6.8. These are multiplication factors for the vc-matrix given in Eq. (6.2), i.e., a factor of five indicates a variance of about 0.40 m^2 and 0.15 mm^2/y^2 for the estimated DEM error and linear displacement rate, respectively. The precision is described with respect to the reference point. In general, the further away from the reference point the worse the estimated precision of the double-difference observations and the estimated displacement rate differ ence. The a posteriori variance factors are relatively small, i.e., the atmospheric variation is likely to be small for the Berlin test site. However, the estimated precision suddenly increases in the uplift area. This supports the theory that the functional model does not fully describe the displacement in this area.

6.1.3 Sensitivity to algorithm settings

There are a large number of settings that can be changed when applying the STUN algorithm. The following list gives an overview of the variables that can influence the final parameter estimation. Thereafter, several experiments are performed using the Berlin data set to investigate the sensitivity of the STUN algorithm to different values for these variables. Note that aside from these settings there are also factors that generally cannot be influenced. These include the amount of acquisitions and their distribution in space and time, the Doppler centroid frequency, the amount and position of point scatterers, the magnitude of the displacement signal, the correctness of the displacement model, and the data noise. There often is a link between these variables and the optimal settings of the STUN algorithm. For example, there is a trade-off between the amount of acquisitions and the bound chosen for the search space. Such dependencies are also described in the following.

- Distance between points of the reference network.
  The points of the reference network are selected using the sparsification strategy described in section 4.4.1. The larger the cell width that is used, the less reference network points are selected, and the larger the distance between points. Advantages of a larger cell width are that the computations are faster, particularly the alternative hypothesis testing step, and that it is less likely that the reference network contains incoherent points. The disadvantage is that the atmospheric (difference) signal increases with distance, which could lead to incorrect estimated parameters using wrapped data, and that if points are removed from the reference network the distance between remaining points may become too large. It is not always possible to influence the number of points in the reference network due to lack of selected points, particularly in rural areas. Moreover, if a small number of acquisitions is available, the initial selection of reference

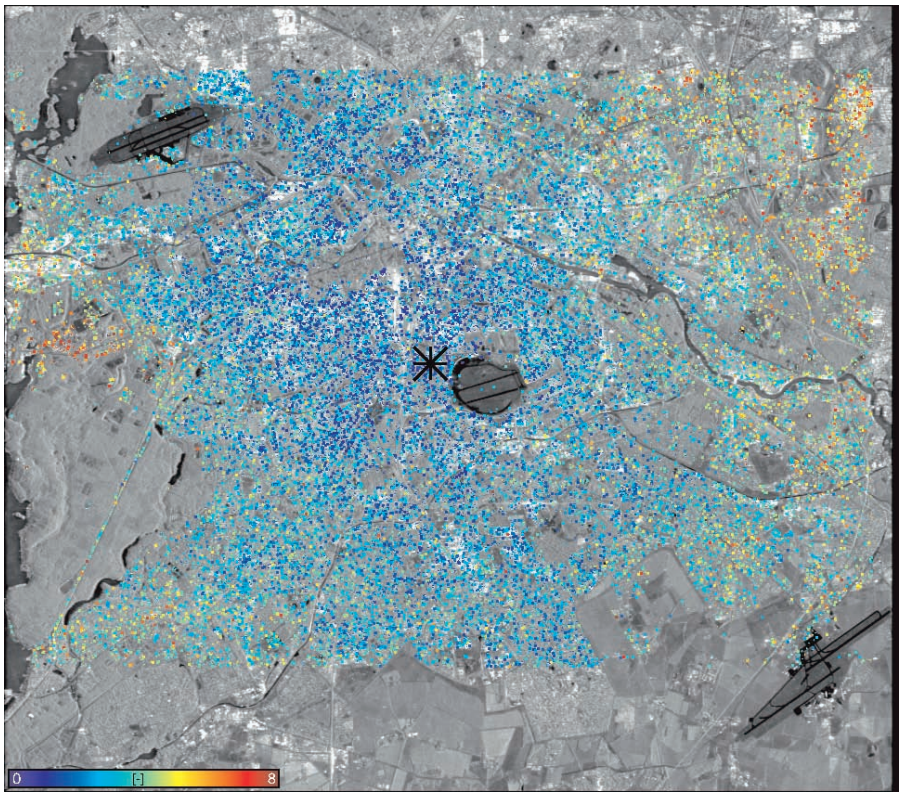


Fig. 6.8: A posteriori variance factors for the Berlin test site estimated using unwrapped data. (blue: 0, red: $\geq$8).

network points becomes cumbersome because this selection is based on the temporally estimated amplitude dispersion index.

- Number of connections in the reference network.
  The larger the number of connections, the larger the number of estimations that must be performed, and consequently the larger are processing time and memory requirements. However, it is expected that a denser network allows for a better discrimination of alternative hypotheses, i.e., that the probability increases that the correct alternative hypothesis is identified. For example, if each point would be connected to only one other point, the arc and point test are identical and they cannot be discriminated. Moreover, if the number of redundant connections is large, the impact of removing an incoherent point only has a small impact on the ability to detect other incoherent points. Note that if the success rate of the estimator is low (say, 30%, e.g., due to a small amount of available acquisitions), increasing the number of connections does not improve the ability of

finding incoherent points, or of correctly identifying incorrectly estimated parameters.

- The stochastic model used during estimation.
  Using the correct vc-matrix during the estimation is expected to provide the best estimates and a good description of the precision. The variance components model used by the STUN algorithm must first be estimated using an initial estimation, see section 4.3. Although this estimation takes some time, it only has to be performed once. Due to the numerical approach described in section 3.3.2, the use of the full vc-matrix does not take additional time during the parameter estimation. Moreover, the estimation of the variance components enables an automated detection of incorrectly processed images that can then be re-processed or removed from the data set. Alternatives to this approach are to use the same variance components or a priori variance components for the SLC images, or using the ensemble coherence estimator, which uses the same weights for all interferometric phase observations. Note in particular that the (wrapped) phase data do not have a normal distribution, although this distribution is a good approximation for the points used in this study.
- Choice of testing parameters.
  The choice of the testing parameters is described in detail in Appendix B. It is expected that a larger power of the tests will favor the point test, i.e., more often a point will be removed from the network. However, the sensitivity of the STUN algorithm to these parameters is expected to be small, i.e., it is expected that most outliers can be identified regardless of the choice of the testing parameters. The reason is that the residuals of the network adjustments of the DEM error and the displacement parameters are expected to be large if the arcs are estimated incorrectly, and to be exactly zero otherwise, as described in section 4.4.2.
- Threshold to select reliable points.
  After estimation using the wrapped data, points are selected based on the a posteriori variance factor. These points are unwrapped using the MCF algorithm. The larger the threshold, the more points are selected, and the more incoherent points are selected, see also Fig. 6.6. The selection of incoherent points can lead to unwrapping errors. If these errors have a global effect this may influence the estimation of other points. In addition, these points will be incorrectly estimated, which cannot be detected during the final estimation using the unwrapped data.
- Bound on the search space.
  This bound is controlled in the STUN algorithm by the variance chosen for the pseudo-observations. A smaller variance implies a larger constraint, and thus a smaller probability that an incorrect maximum is found (assuming the correct maximum is in the search space). Aside from the amount of signal, this setting is also closely related to the amount of available acquisitions. The larger the amount of available data, the larger the probability that the parameters are estimated correctly, and the less

sensitive the STUN algorithm is for this setting. This effect is confirmed by the experiments described in section 3.4 and is explained in more detail below.

- Number of estimated (displacement) parameters.
  The larger the number of estimated parameters, the smaller the degree of freedom, and the better the fit with the observed data. However, if the problem is over-parameterized, the probability increases that parameters are fitted that describe the wrapped data better, but have no relation with the actual displacement. The more acquisitions are available the more complex the model for the displacement can be. The vc-matrix and corresponding correlation matrix can be used to assess whether it is feasible to estimate individual parameters. Note that this analysis can be performed without using actual data. Sensitivity to the estimated parameters is studied in detail using the Las Vegas data set, see section 6.2.

Sensitivity to the number of points and arcs of the reference network

The larger the distance between the points in the reference network, the more likely it is that the parameters are incorrectly estimated using the wrapped data, due to increased atmospheric difference signal. Moreover, if the redundancy of the network decreases, it is more difficult to identify outlier arcs and points. Fig. 6.9 shows the much sparser reference network used in this section. The cell width used during the sparsification to select reference points is set to 1500 m, resulting in the selection of 154 points. The network is constructed using Delaunay triangulation. The mean distance between the points is 2107 m and the standard deviation is 726 m. The number of arcs is 435, i.e., there are ~2.8 connections per point. Since the distance between points in the reference network is increased on purpose, the stochastic model previously estimated using shorter distances could be used here. However, the stochastic model is estimated using the 154 selected points to simulate a situation in which the point density cannot be increased, for example in rural environments. The variance components are estimated using 77 independent estimations, cf. Eq. (4.8). The average distance between points is 1623 m with a standard deviation of 933 m, i.e., the arcs are approximately three times longer than during the reference processing. The estimated variance components are shown in Fig. 6.10. Compared to the variance components estimated previously, see Fig. 6.10, these components seem to be slightly larger. This is particularly noticeable for some of the ERS–2 acquisitions. It is likely that the atmospheric conditions during these acquisitions were more severe than during the others, which remained unchanged. However, the parameters seem to be estimated correctly using the wrapped phase differences between points on these distances for the Berlin area. The vc-matrix of the estimated parameters is propagated as

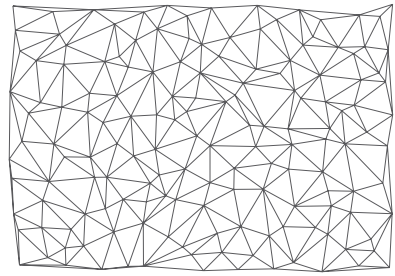


Fig. 6.9: Reference network using large distances and small redundancy for the Berlin test site (network created using Delaunay triangulation). The redundancy of this network is much smaller than that used for the reference processing, shown in Fig. 6.4(b).

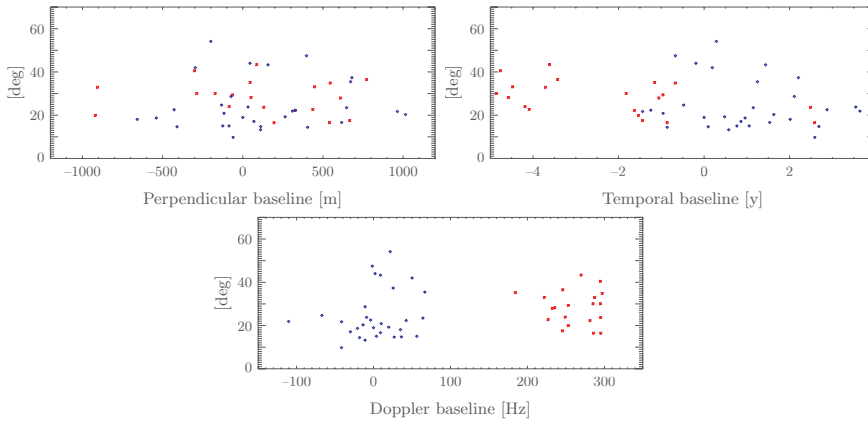


Fig. 6.10: Square roots of the estimated variance components for the Berlin test site using a reference network with large distances between points, as function of perpendicular, temporal, and Doppler baseline. A red asterisk corresponds to an ERS–1, and a blue diamond to an ERS–2 acquisition.

$$Q_{\hat{b}} = \begin{bmatrix} 0.067 & -0.003 \\ -0.003 & 0.027 \end{bmatrix}, \tag{6.3}$$

Compared to the vc-matrix using the shorter arcs, see Eq. (6.2), the propagated precision using these variance components is slightly better, although the atmospheric (double-difference) signal is expected to be larger. The reason may be that the atmospheric signal is limited for the Berlin test site or that the (fewer) selected points contain less noise.

During the same pre-processing procedure as used during the reference processing scenario, 24 arcs and three points are removed from the reference network. During the alternative hypothesis tests one more point is removed. The standard deviation of the least-squares residuals at the arcs (misclosures) is 0.04 m and 0.04 mm/y for the DEM error and displacement rate, respectively. For the selection of the points, an a posteriori variance factor threshold of $\underline{\hat{\sigma}}_x^2 < 3.33$ is used. This threshold is chosen such that selected points have the same threshold on the variance of the displacement rate, see the vc-matrices in Eq. (6.2) and Eq. (6.3). Using this threshold, 43169 points are selected, i.e., approximately the same number as during the reference processing. Fig. 6.11 shows the estimated linear displacement rates using this reference network. No difference with the estimation using a denser reference network can be observed. The estimated precision is practically identical to that estimated during the reference processing described in section 6.1.2, particularly Fig. 6.8. The reason is that the unwrapped phase is identical in both cases. Apparently, the atmospheric signal for the Berlin area is small.

Sensitivity to the stochastic model

For comparison, the estimation at the arcs of the reference network is also performed using the a priori stochastic model, and the ensemble coherence estimator used in the reference PS technique, see Eq. (2.7). The theoretical success rate for the bootstrap estimator, cf. Eq. (3.20), is $P(\underline{\hat{z}} = z)$=0.907 using the a priori stochastic model, i.e., somewhat smaller than the success rate obtained using the estimated variance components. However, the meaning of the computed success rate is limited in this case since it depends on the precision of the observations, which is described using the a priori model. The mean of the estimated DEM error is –0.10 m and is 0.03 mm/y for the displacement rate using this estimator. The average a posteriori variance factor at the arcs is 0.96, which suggests that the precision of the observations is described well by the a priori stochastic model. Finally, the DEM error, displacement rate, and average interferometric residual phase (i.e., termed master atmosphere in Ferretti et al., 2001) are estimated using the ensemble coherence estimator used in the reference PS technique. The search space is bound to [–50,50] m for the DEM error and [–50,50] mm/y for the linear displacement rate. Using the ensemble coherence estimator, the mean of the estimated parameters are –0.29 m, –0.08 mm/y, and 0.96° for the DEM error,

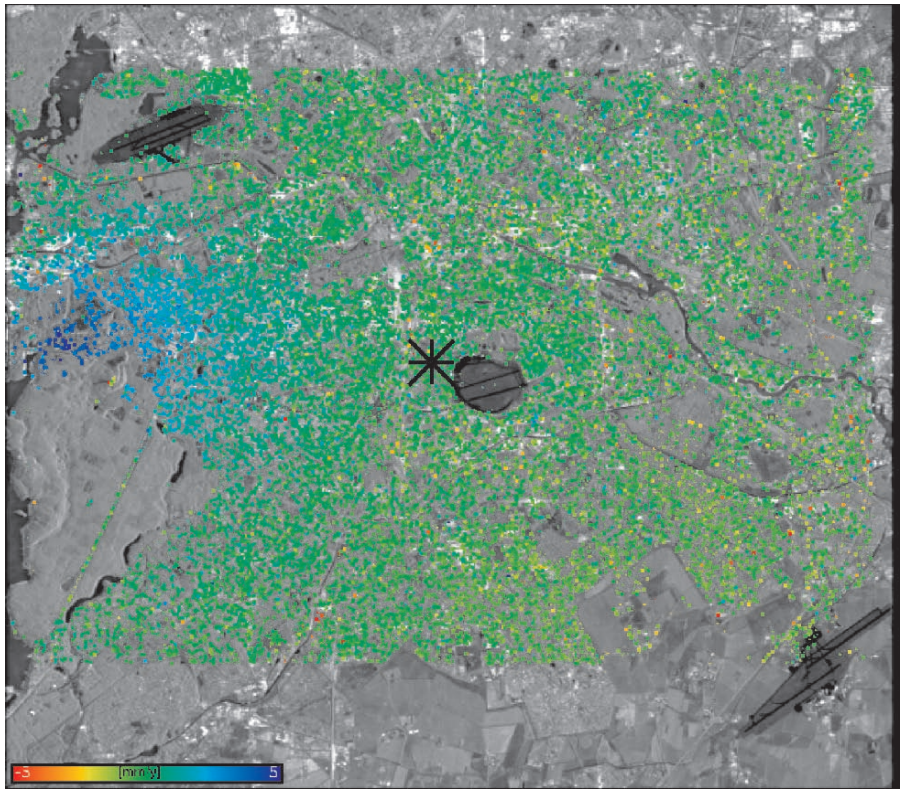
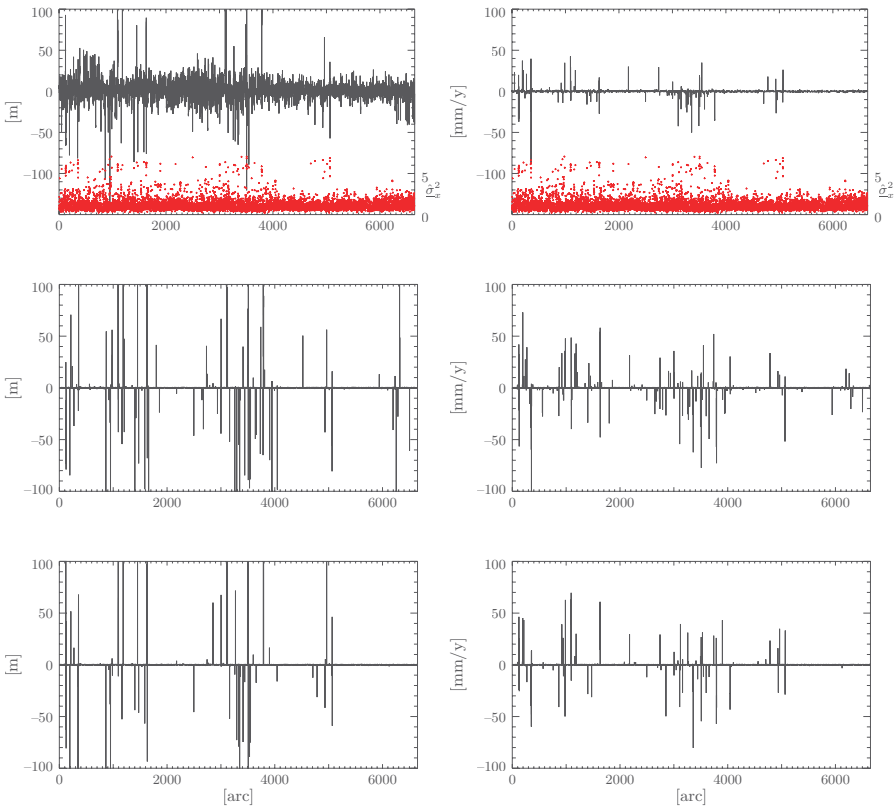


Fig. 6.11: Estimated displacement rates for the Berlin test site using a sparse reference network with ~3 arcs per point and an average distance of ~1600 m between points (red: 5 mm/y subsidence, blue: 5 mm/y uplift).

the displacement rate, and the bias, respectively. The average coherence is 0.80 with a standard deviation of 0.09. The average of the coherence corresponding to the second best fitting set of parameters is 0.39, i.e., it is likely that at most arcs the correct parameters are estimated.

The estimated DEM errors and displacement rates at the arcs of the reference network using the different methods are plotted in Fig. 6.12. The histograms of the estimates using the different methods are visually identical and are not shown. However, there are clear differences between the estimated parameters using the different methods. In general these differences are larger if the data do not agree with the mathematical model. The estimated parameters often get unrealistically large (e.g., 100 m DEM error difference between nearby points) when the precision of the observations is low. Although the estimated parameters do not differ very much at most arcs using this configuration of acquisitions and points of the reference network, an additional

benefit of the estimation of the variance components is that incorrectly processed interferograms are detected.



(a) DEM error **(b)** Displacement rate

Fig. 6.12: Sensitivity of the estimated parameters to the stochastic model. The top row shows the estimated DEM errors (left) and linear displacement rates (right) at the arcs of the reference network using the ILS estimator and the a posteriori stochastic model. The a posteriori variance factors $\underline{\hat{\sigma}}_x^2$, cf. Eq. (4.26), are plotted on the bottom for each arc. The second row shows the differences between these estimated parameters and estimates obtained using the a priori model (see footnote at page 55). The third row shows the differences with parameters estimated using the unweighted ensemble coherence estimator.

Sensitivity to the choice of the testing parameters

During the alternative hypothesis testing, points and arcs are removed based on the computed test statistics using an iterative procedure. This testing procedure is described in detail in section 4.4.3 and Appendix B. The choice of the level of significance α and the power of the (all) tests γ_0 influences when the null-hypothesis is rejected, and, if so, which alternative hypothesis is selected as most likely cause for the rejection. A higher level of significance implies that the false alarm rate increases, i.e., that more often points and arcs are removed that do not need to be removed. Here, the level of significance for the one-dimensional test is fixed to $\alpha_1 = 5\%$. The effect of changing the power is studied using the network and data of the reference processing. The power of the test is gradually increased from 20% to 90%. The results are reported in Table 6.2. Initially, the reference network consists of 1066 points

Table 6.2: Experiments with testing parameters. For different values of the power γ the level of significance for the OMT test, the non-centrality parameter λ_0, the number of iterations before acceptance of the OMT test, and the total number of removed points and arcs are reported. The level of significance for the one-dimensional test $\alpha_1 = 5\%$ is fixed during all experiments.

γ_0	α_b	λ_0	Iterations	Removed points	Removed arcs
0.20	0.073	1.24	90	5	137
0.30	0.079	2.06	90	5	137
0.40	0.084	2.91	87	5	137
0.50	0.087	3.84	80	7	142
0.60	0.090	4.90	75	9	164
0.70	0.092	6.17	74	9	164
0.80	0.095	7.85	74	9	164
0.90	0.099	10.51	74	9	164

and 6650 arcs. As expected, more points are removed if the power increases, because the higher dimensional alternative hypothesis is more likely to be accepted. The number of iterations also decreases for increasing power. This is likely caused by the fact that if a point is removed, also all incorrectly estimated arcs are removed. However, the differences are marginal, which may be caused by the fact that most points in the reference network are coherent, and most estimations at the arcs are correct. A large number of arcs are clear outliers, which are identified using all settings for the test parameters. Moreover, the errors at the arcs are not expected to have a normal distribution. Nonetheless, the procedure that is followed offers a way to automatically remove the inconsistencies in the network.

During the reference processing, a power $\gamma_0 = 5\%$ and a level of significance $\alpha_1 = 5\%$ is used. During a pre-processing step, 17 points and 257 arcs are removed. After this, another point and 16 arcs (in total) are removed by

the alternative hypothesis testing procedure, see also section 6.1.2. This pre-processing step is not performed here. Apparently, more points are removed if this pre-processing step is used. However, it is advised to use some kind of pre-processing step in order to guarantee the removal of all clearly incoherent points, and to make this step faster. The alternative hypothesis testing procedure takes approximately one hour for 80 iterations. The computation of the vc-matrix of the least-squares residuals and the point tests takes the longest.

Sensitivity to the number of acquisitions

In order to assess the sensitivity of the algorithm to the number of available acquisitions, first a hypothetical phase signal $f(t) = \alpha t$ is considered. It is of interest to derive the amount of equidistant wrapped phase samples that are required to reconstruct this signal (or given a certain number of samples what the maximum slope α is of the signal that can be recovered). The equidistant samples are given as

$$f(t_i) = W\{\alpha t_i\}, \quad t_i = \frac{2i}{N}, i = -N/2 + 1, \ldots, N/2, \tag{6.4}$$

where N (even) is the number of samples. It is easily understood that a single sample at $t_i = 1$ suffices to fully reconstruct the unwrapped continuous signal provided that $|\alpha| < \pi$, and that aliasing occurs otherwise, see also Fig. 6.13. Generalization implies that the correct slope α can only be estimated from N equidistant samples if $|\alpha| < N\pi$, which follows from the property of the wrapping operator

$$W\{\alpha\pi t_i\} = W\{(\alpha + 2kN)\pi t_i\}, \quad t_i = \frac{i}{N}, N \in \mathbb{Z} > 0, k, i \in \mathbb{Z}. \tag{6.5}$$

This is equivalent to the Nyquist sampling rate or Shannon sampling frequency known from signal processing (Shannon, 1948) which states that the sampling frequency must be at least twice that of the highest frequency contained in the signal for exact reconstruction.

In this case the difference in ordinates of the sampled signal must be smaller than π between adjacent samples to avoid signal aliasing. For the linear displacement model used in PS interferometry, this result can be interpreted as follows: if K interferograms are available, acquired over ΔT years, the average sampling distance is given as $\Delta T/K$. Since, in repeat pass interferometry, a displacement with a magnitude of half a wavelength induces a full phase cycle, the maximum displacement rate that can be estimated using equidistant sampled wrapped phase data is

$$\alpha^{\max} = \frac{\lambda}{4} \cdot \frac{K}{\Delta T} \quad \leftrightarrow \quad K_{\alpha}^{\min} = \frac{4}{\lambda} \cdot \Delta T \cdot \alpha^{\max}. \tag{6.6}$$

The wavelength used by ERS is λ=56.6 mm, i.e., if one sample is available per year the maximum displacement rate that can be estimated unambiguously is

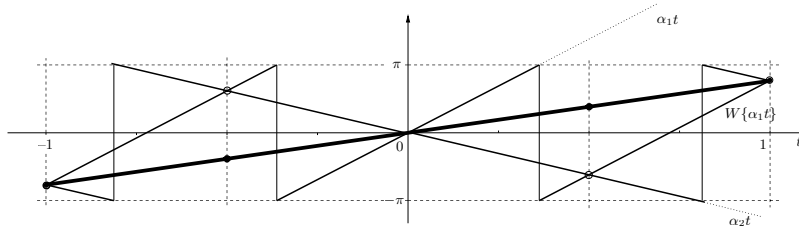


Fig. 6.13: Signal aliasing. Shown are three phase signals $\phi = \alpha_i t$ that have (some) coinciding wrapped phase values at the sample points. This figure demonstrates that the (unwrapped) signal can only be recovered if at least $|\alpha_i|/\pi$ regularly spaced samples are available.

~14 mm/y. Moreover, the nominal orbit repeat cycle of the ERS satellites is 35 days, which implies that aliasing occurs for a displacement rate (difference) larger than

$$\alpha_{\text{ERS}}^{\max} = \frac{\lambda}{4} \cdot \frac{1}{35} \approx 150 \text{ mm/y}. \tag{6.7}$$

The phase induced by a DEM error is a linear function of the baseline. A phase cycle is induced by a DEM error equal to the height of ambiguity (Hanssen, 2001) which follows from Eq. (2.12) by substitution of $\phi_{\text{topo}} = 2\pi$ as

$$H_{\text{amb}} = \frac{\lambda}{-2} \frac{r \sin(\theta)}{B_\perp}. \tag{6.8}$$

The (average) sampling distance of the perpendicular baseline is analogously defined as $\Delta B_\perp / K$, where $\Delta B_\perp$ is the baseline span. Thus, the maximum DEM error that can be estimated unambiguously using equidistant samples is given as

$$\Delta h^{\max} = \frac{\lambda}{4} \cdot \frac{r \sin(\theta)}{\Delta B_\perp} \cdot K \quad \leftrightarrow \quad K_{\Delta h}^{\min} = \frac{4}{\lambda} \cdot \frac{\Delta B_\perp}{r \sin(\theta)} \cdot \Delta h^{\max}. \tag{6.9}$$

For example, for K=10 and typical ERS parameters[1], the maximum DEM error between points that can be estimated is $\Delta h^{\max} = 21.6$ m. Since both the linear displacement rate and the DEM error are estimated using the wrapped phase data, the minimum required number of interferograms is the sum $K_\alpha^{\min} + K_{\Delta h}^{\min}$ of the minimum in each dimension. For example, this is the case if $K_\alpha^{\min}$ temporally equally spaced samples with a zero perpendicular baseline and $K_{\Delta h}^{\min}$ spatially equally spaced samples are available with a zero temporal baseline. In this special case, the two frequencies can be estimated independently. Due to the irregular sampling, the sampling distance for certain interferograms is smaller than the average. This, and the fact that only a single frequency is estimated in each dimension, allows for the estimation of DEM error and linear displacement rates above the Nyquist frequency

[1] $\lambda = 56.6$ mm, $r = 850$ km, $\theta = 21°$, $\Delta B_\perp = 2000$ m

(i.e., the frequency corresponding to the Nyquist sampling rate). The effect of irregular sampling has another advantage, namely that aliasing does not occur at a single frequency, but the power is spread out over the estimated spectrum.

Noise is ignored in this derivation. However, from the numerical simulation described in section 3.4, it is clear that it can have a large impact on finding the correct ambiguities. If the solution space is searched above the Nyquist frequency, aliasing occurs in case of regular sampling, and the correct DEM error or displacement rate cannot be distinguished from the aliased solution. In case of irregular sampling, aliasing occurs to a lesser extent, but certain solutions aside from the true solution are still likely to have a higher amplitude in the spectrum. Due to observation noise, the incorrect solution may actually have a higher amplitude than the true solution, which implies that the search bound on the solution space should be chosen appropriately low if the number of interferograms is small.

To demonstrate this, an estimation is performed using $K = 10$ interferograms; the solution space is bounded using a standard deviation $\sigma_{\Delta h} = 20$ m for the DEM error and $\sigma_\alpha = 10$ mm/y for the linear displacement rate. These "soft" bounds imply that most parameters are expected to be between 40 m and 20 mm/y (two-sigma level). For regularly sampled data the number of required interferograms would be (using ERS parameters and $\Delta T = 5$ y, $\Delta B_\perp = 1000$ m)

$$K^{\min} = K_\alpha^{\min} + K_{\Delta h}^{\min} = 7.1 + 9.3 = 16.4, \tag{6.10}$$

i.e., signal aliasing is expected to occur. The ten interferograms used during this estimation are randomly selected from all the images used during the reference processing (see section 6.1.2) for which $|B_\perp| < 500$ m, and $|T| < 2.5$ y. The selected data set has $\Delta B_\perp = 795$ m and $\Delta T = 3.1$ mm/y. The same reference network is used as during the reference processing. The points in the reference network are thus not selected using the amplitude dispersion index estimated using the reduced data set. The variance components of the stochastic model are estimated as described during the reference processing. The variance components have similar values. The vc-matrix of the estimated parameters using the estimated stochastic model is given as

$$Q_{\hat{b}} = \begin{bmatrix} 1.016 & 0.471 \\ 0.471 & 1.035 \end{bmatrix}, \tag{6.11}$$

and the corresponding correlation matrix as

$$\rho = \begin{bmatrix} 1.000 & 0.460 \\ 0.460 & 1.000 \end{bmatrix}. \tag{6.12}$$

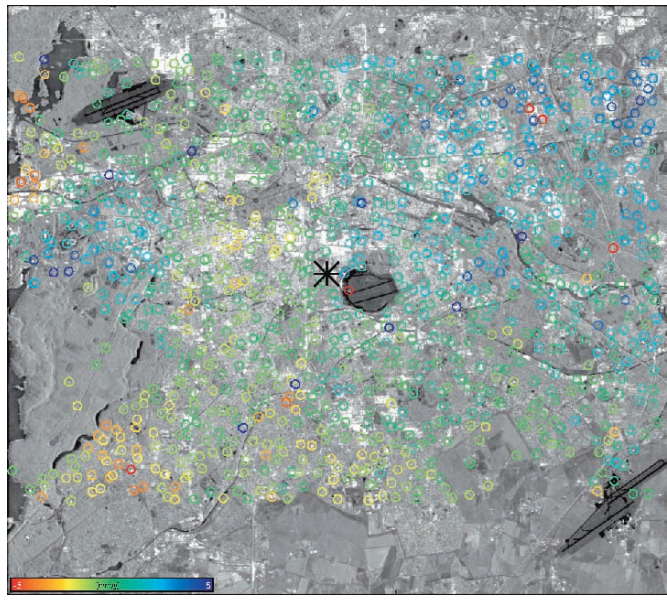
The correlation between the estimated DEM error and linear displacement rate increased compared to previous estimations using more data. However, the correlation is still reasonably small. The theoretical success rate using

the bootstrap estimator, cf. Eq. (3.20), is $P(\hat{\underline{z}} = z) = 0.958$ (and 0.84 using the a priori stochastic model). However, aliasing effects are not taken into account by this estimate for the success rate.

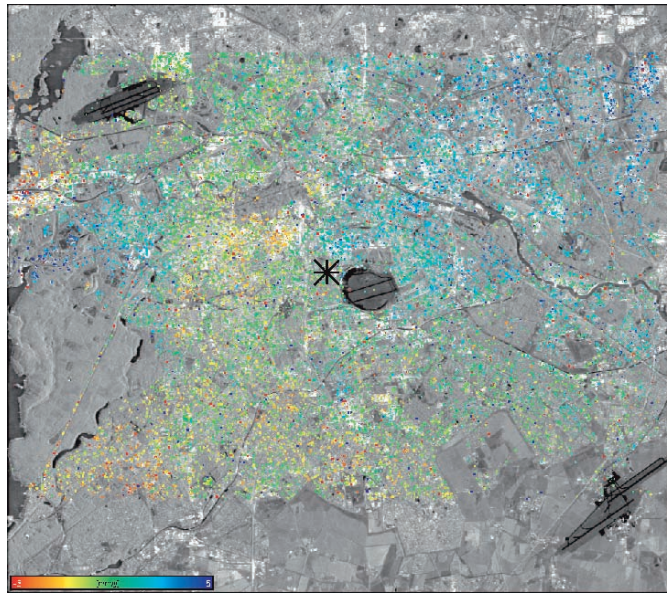
The DEM error and displacement rate are estimated at the 6650 arcs of the reference network, using the integer least-squares estimator and the estimated stochastic model. After this estimation, points and arcs that clearly are incorrect are removed in a pre-processing step. This procedure is described in section 6.1.2 (network integration). In total, one point and 178 arcs are removed from the reference network. During the alternative hypothesis testing step, the same testing parameters are used as during the reference processing. Additionally, eighteen points and 395 arcs are removed. For this scenario, a total of nineteen points and 573 arcs are removed, while during the reference processing eighteen points and 62 arcs are removed. Thus, the network using $K = 10$ is less consistent then the network using $K = 50$ interferograms, which is to be expected. However, an internally consistent reference network could be established by removing the identified arcs and points. The standard deviation of the misclosures at the arcs of the reference network are 0.28 m and 0.40 mm/y for the DEM error and displacement rate, respectively. The estimated displacement rates at the points of the reference network are plotted in Fig. 6.14(a). Some points of the reference network seem to be estimated incorrectly (large values), but apparently consistently. This demonstrates that a small closing error does not necessarily mean a correct estimated parameter, which is due to the fact that the arcs are not independently observed. Fig. 6.14(b) shows the estimated displacement rates at 28462 points with an a posteriori variance factor below one. The estimated variance factors are significantly smaller for the estimation using only ten interferograms. The reason for this is the reduced redundancy, i.e., the least-squares residuals are expected to be smaller, and thus the a posteriori variance factor. The variation of the estimated displacement rates is much larger in this case, compare, e.g., with Fig. 6.6. This larger variation is likely caused by unmodelled atmospheric signal. Note that the uplift area cannot significantly be detected from this result.

6.1.4 Cross-comparison between adjacent tracks

For the Berlin test site, data of the ERS–1 and ERS–2 satellites are available for two descending tracks (i.e., adjacent tracks, with approximately 40 km overlap at this latitude). The difference in viewing angle between these two ERS tracks is approximately 3°. Due to this difference in viewing angle, also the ground-range pixel spacing is different. For the considered area this is approximately 9.64 m vs. 10.69 m for the first stack towards the East (points at larger slant range) and West, respectively. A joint processing of all data with respect to a single master image is not attempted, because the height ambiguity would be extremely small for interferograms with such large baselines. Aside from this, such an approach would severely limit the amount



(a) Reference network



(b) All points

Fig. 6.14: Estimated line-of-sight displacement rates for the Berlin test site using ten interferograms. **(a)** shows the estimates at the 1047 points of the reference network in the range from –5 to 5 mm/y (red to blue). **(b)** shows 28462 selected points with an estimated a posteriori variance factor smaller than 1.0.

of points that can be estimated (only targets with a extremely wide opening angle are expected to be coherent in both stacks). Colesanti et al. (2002) estimated that less than 30% of PS points that are visible in data of one track are also observed in data of the other track. Therefore, each stack is processed on its own master image. However, it is assumed that there is a common set of PS points located at the same objects in both stacks. The overall displacement pattern is assumed to be spatially smooth, thus allowing a cross-comparison of the estimated displacements at different PS points in both stacks. The estimated DEM error cannot be compared, because the PS points are not expected to be the same in both stacks. The processed area for both tracks is shown in Fig. 6.15. It is not identically cropped in both stacks. The main

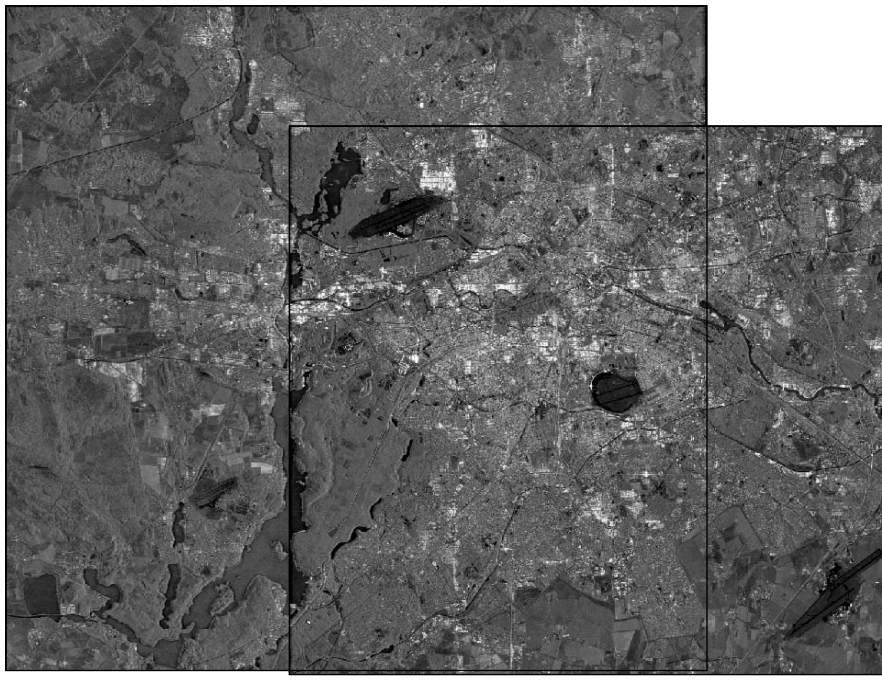


Fig. 6.15: Processed area for the Berlin test site for the two adjacent tracks. Shown is the mean amplitude of all available data for each track. The data is coarsely geo-referenced by mirroring in azimuth and range direction. The stack in the West (left), which only partially covers the city of Berlin, corresponds to track 165 (~28 km wide by 29 km high). The stack in the East (right) corresponds to track 437 (~26 km wide by 24 km high).

reason for this is that the city of Berlin is not fully covered by the second stack (the scene ends at the right side of the crop shown in Fig. 6.15). Moreover, the first stack was processed before the second stack became available, and

the master image and processed area were selected without considering the data of the second stack.

Scene selection

Fig. 6.16 shows all available data for both tracks. The first stack, track 437, contains 70 scenes and is processed on a master acquired at March 22$^{\text{nd}}$, 1997 (ERS–2). The second stack, track 165, contains 43 scenes, and is processed relative to a master acquired at October 22$^{\text{nd}}$, 1998 (ERS–2). Data are acquired at approximately 10:03 and 10:06 UTC for the first and second track respectively, i.e., around 8:00 am local time. For the cross-comparison of the estimated displacement in the two data stacks, data are selected that had the largest possible temporal overlap, see also Table 6.3. Data before December 23$^{\text{rd}}$, 1995, are thus not used during this not cross-validation, because they are not available for track 437. Data after December, 1999, are not considered for this cross-validation due to the instability of the ERS–2 platform after this date, and resulting large variation of the Doppler centroid frequency. The Doppler centroid frequency of the selected acquisitions of track 437 is shown in Fig. 6.17. For ERS 1 acquisitions, the Doppler centroid frequency

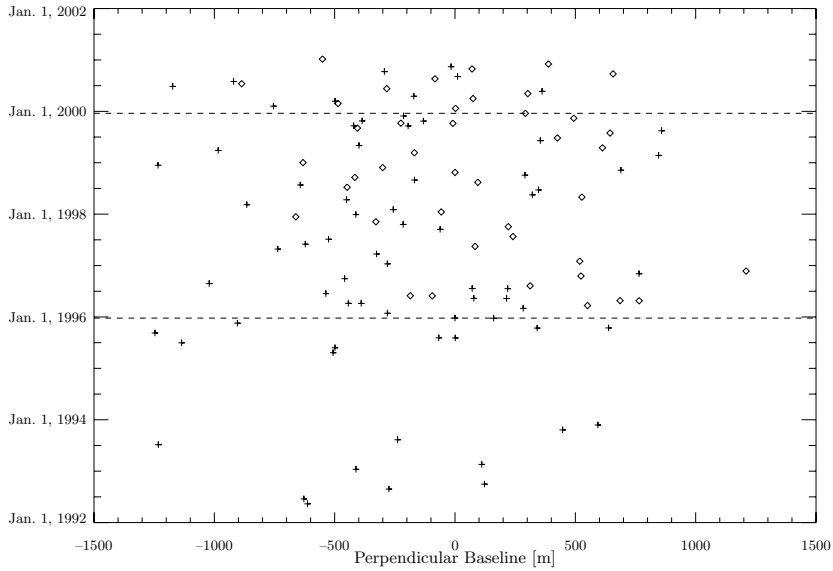


Fig. 6.16: Data selection for the Berlin test site. Data for two adjacent tracks are available; track 437 (plus marks) and track 165 (diamonds). Only data acquired between December 23$^{\text{rd}}$, 1995, and after February 5$^{\text{th}}$, 1999, are selected to make the two processed stacks more comparable (dashed lines).

is approximately 400 Hz, while for ERS–2 it is approximately 200 Hz. Due to this difference, a difference of one meter of the azimuth sub-pixel position between two PS points induces an interferometric phase of approximately 10°, see Fig. 2.6 at page 22. The azimuth resolution of ERS is ∼4 m, which implies that this effect should be included in the functional model. However, the precision of an estimated azimuth sub-pixel position is small, because the variation in Doppler centroid frequency is small. Assuming a priori variance components (see section 4.3), the vc-matrix of the DEM error, the linear displacement rate, and the azimuth sub-pixel position can be computed in advance. For these 41 images it is given, cf. Eq. (3.18), as

$$Q_{\hat{b}} = \begin{bmatrix} 0.127 & -0.007 & 0.033 \\ -0.007 & 0.167 & -0.110 \\ 0.033 & -0.110 & 1.648 \end{bmatrix}, \tag{6.13}$$

and the corresponding correlation matrix is

$$\rho = \begin{bmatrix} 1.000 & -0.047 & 0.072 \\ -0.047 & 1.000 & -0.209 \\ 0.072 & -0.209 & 1.000 \end{bmatrix}. \tag{6.14}$$

The units are m, mm/y, and m for the DEM error, linear displacement rate, and azimuth position, respectively. The standard deviation of the estimated azimuth sub-pixel position thus would be ∼1.3 m (assuming the a priori variance components correctly describe the precision of the data), which is rather large compared to the azimuth resolution of ERS. In addition, as can be inferred from Fig. 6.17, the correlation between possible unmodeled displacement in 1996 and estimated azimuth sub-pixel position is large if these data would be used, since almost all ERS–1 images are acquired in that year. Therefore, only ERS–2 acquisitions are used during the comparison of the two stacks, and only a DEM error and linear displacement rate are estimated.

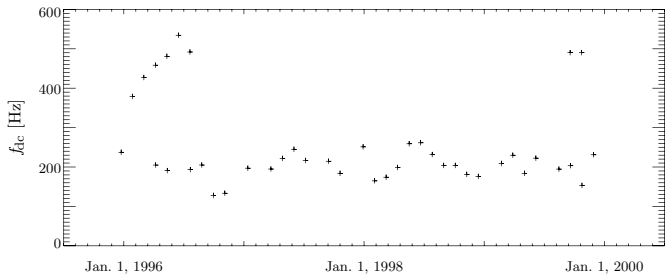


Fig. 6.17: Doppler centroid frequency for selected data of the first stack of the Berlin test site. The ERS–1 images have a larger Doppler centroid frequency than the ERS–2 images.

Table 6.3 gives an overview of the finally selected data for the comparison based on the temporal and the Doppler centroid frequency constraint.

Table 6.3: Berlin test site data used during cross-comparison of adjacent tracks. Listed are the acquisition dates of the master, the first and the last scene. In Appendix C a full listing of all available acquisitions is given.

track	#scenes	first	last	master
437	33	23–Dec–1995	27–Nov–1999	22–Mar–1997
165	29	21–Mar–1996	20–Jan–2000	22–Oct–1998

Estimation of track 437

The city of Berlin is in the center of the processed area, which is approximately 26 km wide by 24 km high, see also Fig. 6.16 and Fig. 6.15. Using 33 ERS–2 acquisitions, the DEM error and the linear displacement rate are estimated at the selected points. The estimated vc-matrix is

$$Q_{\hat{b}} = \begin{bmatrix} 0.101 & -0.002 \\ -0.002 & 0.152 \end{bmatrix}. \tag{6.15}$$

The parameters are estimated in a similar manner as is done for the reference processing, described in section 6.1.2. The main difference is that during this estimation less data from a smaller time span are used. Finally, 51269 of 78779 estimated points are selected as reliable points using a threshold $\underline{\hat{\sigma}}_x^2 < 3.0$ on the a posteriori variance factor (estimated using the wrapped data). The estimated linear displacement rates at these points are shown in Fig. 6.18(b).

Estimation of track 165

Track 165 is the most West track, only partially covering the city of Berlin, see also Fig. 6.15. The processed area is approximately 20 km wide by 20 km high. For this track, 29 ERS–2 acquisitions are selected, see also Table 6.3. The processing is performed using similar parameters as for the first stack. The number of points in the reference network is 968. The ratio of the arcs per point is 6.2. The average distance between points in the reference network is 1016 m, with a standard deviation of 382 m. These numbers are very similar for both stacks, even though this crop contains more rural area west of the city. After estimation of the variance components, orbit 8307 is removed from the data set because it clearly is less precise than the other images. The perpendicular baseline for the interferogram with this slave image is ~1200 m. Likely a coregistration problem occurred, or not all points are ideal point scatterers, causing geometrical decorrelation for this interferogram. The

propagated vc-matrix of the estimated parameters is (using the estimated variance components of the variance component stochastic model)

$$Q_{\hat{b}} = \begin{bmatrix} 0.118 & 0.024 \\ 0.024 & 0.174 \end{bmatrix}. \tag{6.16}$$

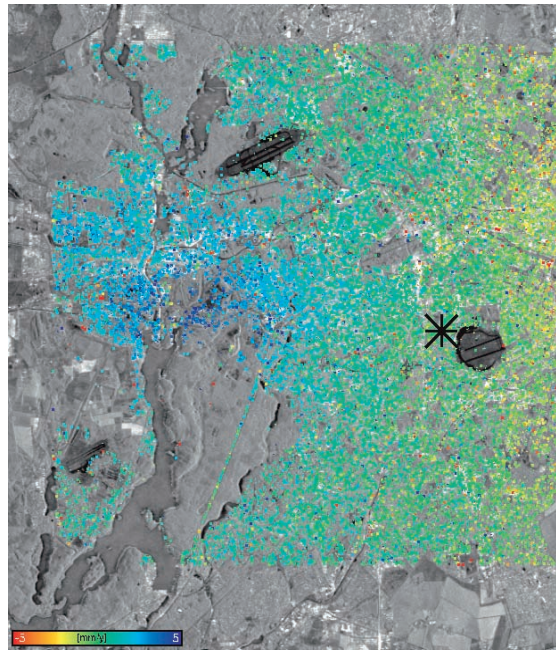
The estimated precision is very similar for this stack as for the first stack, see Eq. (6.15). The larger correlation between the estimated parameters is caused by the different distribution of the acquisitions in time and space. Particularly, the perpendicular baselines of the interferograms of the first stack is a bit larger. Anyway, the correlation coefficient is 0.17, which is small. The points are selected using the same threshold on the a posteriori variance factor. In total, 41620 of 65137 estimated points are selected. The finally estimated displacement rates using the unwrapped data are shown for both tracks in Fig. 6.18(a). The reference points are chosen near to the Tempelhof airport for both tracks.

Cross-comparison of the results

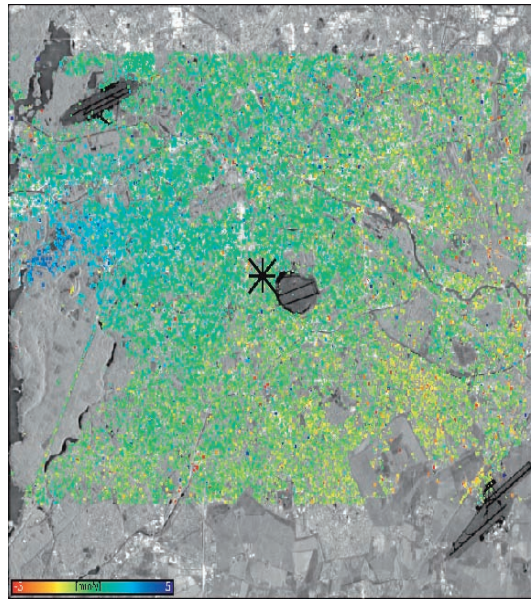
The linear displacement rates at the PS points are estimated from fully independent data, and can be cross-validated using the assumption that the displacement is spatially correlated. The only variable that is not independent is the DEM that is used during the differential interferometric processing. However, it is not expected that the DEM affects the estimated displacement rates, because a DEM error is also estimated. Moreover, the correlation coefficient between these two estimated parameters is small. The estimated points are geo-referenced to enable a comparison in the same reference frame. Furthermore, vertical displacement is assumed, i.e., the estimated line-of-sight displacements are mapped to the local vertical direction using the incidence angle as

$$\hat{\underline{\alpha}}_{x,\mathrm{VERT}} = \frac{\hat{\underline{\alpha}}_x}{\cos\theta_{x,\mathrm{inc}}}. \tag{6.17}$$

The impact of this mapping is negligible, because the difference in look angle is only a few degrees, but it is performed nonetheless. For the comparison ~2000 selected points around the uplift area are used for which the a posteriori variance factor is $\hat{\underline{\sigma}}_x^2 < 3.0$, see Fig. 6.19. Each estimated point of track 437 is compared to the closest point estimated in track 165. The averages of the estimated displacement fields of both tracks are set to zero in the overlapping are where no displacement is expected, to avoid the influence of different displacement rates of the reference points used during the independent estimations. A histogram of the difference between the estimates of both tracks is plotted in Fig. 6.20(a). The variance of the difference is ~1 $\mathrm{mm}^2/\mathrm{y}^2$. Assuming equal variance for all points in both tracks, and neglecting individual points that may have a deviating displacement behavior, the variance of the estimated displacement rates is ~0.5 $\mathrm{mm}^2/\mathrm{y}^2$. This is in



(a) Track 165



(b) Track 437

Fig. 6.18: Line-of-sight displacement rates for Berlin, estimated independently using data from adjacent tracks.

reasonable agreement with the expected variance for this area, see Eq. (6.15) and Eq. (6.16), the vc-matrix of the first and second track, respectively, which have to be scaled with the estimated a posteriori variance factors for each point (which is between two and five for most points).

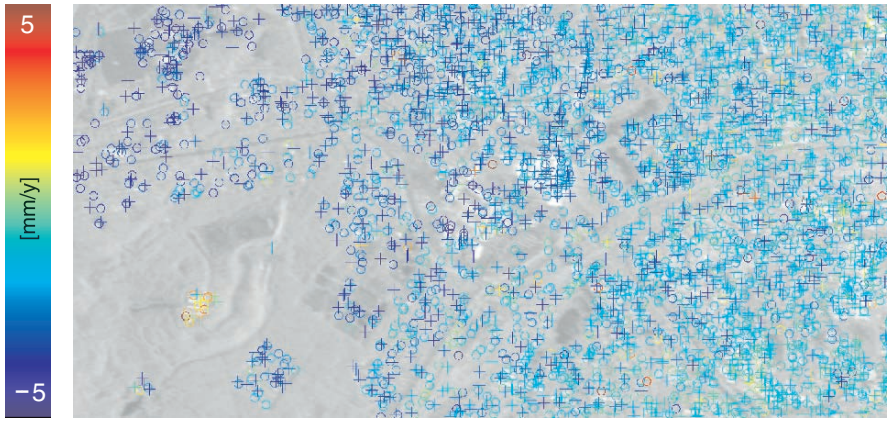


Fig. 6.19: Estimated displacement rates for the city of Berlin using data from two adjacent tracks. Estimates are converted to vertical displacement and plotted in the range between –5 and 5 mm/y in UTM projection. The plotted area is $\sim 6 \times 3$ km^2. A plus mark is used to plot 2187 selected estimates of track 437 and an $\times$-mark to plot 2071 estimates of track 165.

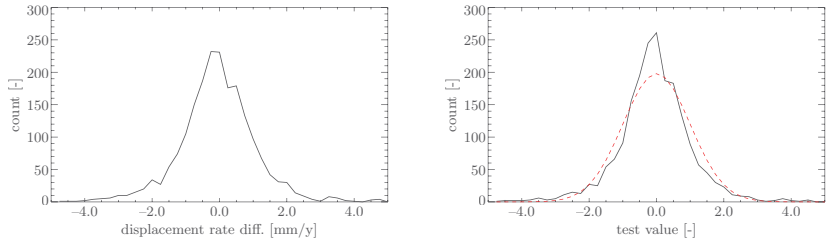


Fig. 6.20: Histograms of estimated displacement rates using data from two adjacent tracks. **(a)** shows the histogram of the difference between the estimated vertical displacement rates. **(b)** shows the difference confronted with the estimated precision, cf. Eq. (6.18). The dashed line shows the standard normal distribution.

To confront the estimated values with the estimated precision (i.e., to assess the quality of the estimated precision), the following statistic is computed for each difference

$$w_x = \frac{\hat{\underline{\alpha}}^I_{x,\mathrm{VERT}} - \hat{\underline{\alpha}}^{II}_{x,\mathrm{VERT}}}{\sqrt{\hat{\underline{\sigma}}^2_{x^I} + \hat{\underline{\sigma}}^2_{x^{II}}}}, \tag{6.18}$$

where $\hat{\underline{\alpha}}^I_{x,\mathrm{VERT}}$ is the estimated vertical displacement rate using data of track 437, and $\hat{\underline{\alpha}}^{II}_{x,\mathrm{VERT}}$ that of the closest point in the second stack. The variance $\hat{\underline{\sigma}}^2_{x^i}$ of the estimated vertical displacement rate is given as

$$\hat{\underline{\sigma}}^2_{x^i} = \frac{1}{\cos\theta^i_{x,\mathrm{inc}}} \cdot \hat{\underline{\sigma}}^2_x \cdot \hat{\underline{\sigma}}^2_\alpha, \tag{6.19}$$

where $\hat{\underline{\sigma}}^2_x$ is computed cf. 4.26, using the a posteriori variance factor $\hat{\underline{\sigma}}^2_x$ which is estimated using the unwrapped data, and $\hat{\underline{\sigma}}^2_\alpha$ is the propagated precision of the estimated line-of-sight displacement, i.e., 0.152 and 0.174 $\mathrm{mm}^2/\mathrm{y}^2$ for track 437 and 165, respectively, see Eq. (6.15) and Eq. (6.16). The factor $1/\cos(\theta_{\mathrm{inc}}) \approx 1.07$ maps the estimated precision of the line-of-sight displacement to that of vertical displacement. The test Eq. (6.18) is expected to have a standard normal distribution. Values of w smaller than one indicate that the actual precision is better than described by $\sigma^2_{z_i}$ and values larger than one that the description is too optimistic. The histogram of the computed values for the uplift area is show in Fig. 6.20(b). The precision is described correctly for this area, because the test values closely match the standard normal distribution.

6.2 Las Vegas

The second test site is Las Vegas city, located at 36°10′ northern latitude, 115°10′ western longitude. Las Vegas is one of the fastest growing metropolitan areas in the United States of America. Between 1990 and 2000 the population almost doubled. In the metropolitan area live ~1.4 million people (Las Vegas Metropolitan Statistical Area, see Evans et al., 2000). Currently, the urbanized area is approximately 20×20 km^2. Las Vegas lies in a broad desert valley in southern Nevada. Mountains surrounding the valley extend to ~3500 m above the valley floor, see also the DEM of the area shown in Fig. 6.21. The average daily temperature is between 5°C in January and 30°C in June. The average annual precipitation varies significantly from year to year but typically is between 5 and 20 cm (National Weather Service, 2004). The Las Vegas area undergoes large displacements dominantly linear and locally seasonal of nature, see, e.g., (Amelung et al., 1999; Bell et al., 2002; Hoffmann et al., 2001; Pavelko, 2003). The local subsidence is primarily related to groundwater withdrawal. Between 1948 and 1963 the center of the valley had subsided ~1.0 m, and by 1980 ~1.5 m, and it still continues to do so (Bell et al., 2002). First, section 6.2.2 describes a standard processing with the STUN algorithm of the Las Vegas data set, estimating DEM errors and linear displacement rates. In the next section the number of estimated

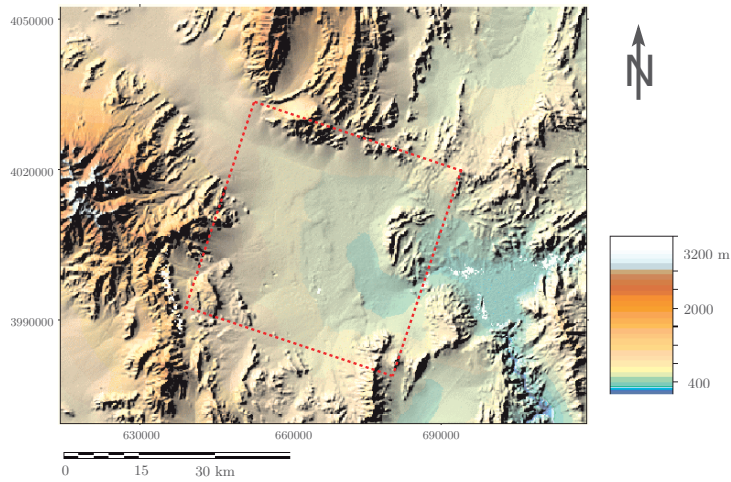


Fig. 6.21: DEM used for the Las Vegas test site (color shaded). Projection: UTM, zone 11, WGS84 ellipsoid. The area covered by the interferograms is indicated by the rectangle.

parameters is increased. Section 6.2.4 reports about the estimation of the seasonal displacement component using trigonometric base functions. Finally, in section 6.2.5 ENVISAT acquisitions are included in the estimation, which requires the additional estimation of the range sub-pixel position. Table 6.4 lists the experiments that are performed for the Las Vegas test site.

6.2.1 Data availability

The distribution of the 45 usable ERS acquisitions is shown in Fig. 6.22. All available data for this track, including the ENVISAT data used in experiment IV, are listed in Appendix C. The first image of this data set was acquired at April 21st, 1992, and the last image at February 18th, 2000. The selected master image was acquired at June 13th, 1997. Data after this date are not used for the estimation because of the large Doppler centroid frequencies of these ERS–2 acquisitions (1000–4000 Hz, due to failures of the gyroscopes and resulting yaw-steering problems). Note that the earlier images are acquired by ERS–1, and that there is a temporal gap in the data set approximately from 1994 to 1995. During this time the "Geodetic" and "Shifted Geodetic" phase were executed, which required a 168 days orbit repeat cycle for ERS–1. The perpendicular baselines vary between –1100 and 1100 m, i.e., distributed scatterers are fully decorrelated in the large baseline interferograms. The interferogram with slave orbit 8226 is not included in the estimation because the coregistration did not succeed for this interferogram, which has largest perpendicular baseline, $B_\perp \approx$1350 m. The nominal Doppler

Table 6.4: Las Vegas test site experiments. Listed are the number of interferograms used (K), the estimated parameters ($\hat{\underline{b}}$), and the goal of the experiment. The estimated parameters are coded as H for estimated DEM error, V for linear displacement rate, A for average atmosphere, D for Doppler dependent azimuth sub-pixel position, S and C for sine and cosine terms of seasonal displacement, and R for range sub-pixel position. The size of the area for all tests is ~23 × 20 km^2. The time range is from 1992 to 2000 for experiments I–III, and up to 2004 for experiment IV.

#	K	$\hat{\underline{b}}$	Purpose
I	45	H, V	Reference processing. This estimate is compared with the reference PS technique.
IIa	45	H, V, A	Additionally estimate average atmosphere (demonstration that estimated H and V are not very sensitive to this parameter, but that the residual phase is reduced considerably).
IIb	45	H, V, D	Additionally estimate azimuth sub-pixel position (demonstration that this parameter should not be estimated for this data set).
III	45	H, V, S, C, A	Additionally estimate seasonal displacement (demonstration of using trigonometric base functions).
IV	55	H, V, D, R	Include ERS–ENVISAT cross interferograms (demonstration of continuation of the ERS phase time series with ENVISAT data).

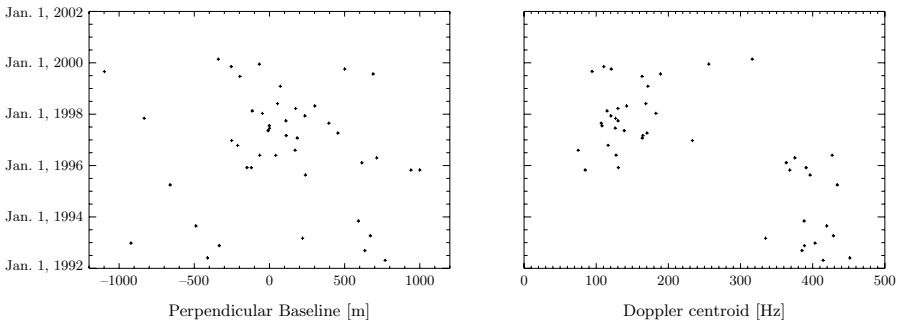


Fig. 6.22: Baseline distribution for available ERS data (45 acquisitions) for the Las Vegas area (track 356, frame 2871). Note that the (earlier acquired) ERS–1 images have a consistently larger Doppler centroid frequency than the ERS–2 images.

centroid frequencies for this area are centered around ~125 Hz for the ERS–2, and around ~400 Hz for the ERS–1 acquisitions. This implies that the interferometric phase of point scatterers contains a term induced by azimuth sub-pixel position. One meter position (difference) causes $\phi_\xi \approx 15^\circ$ phase in interferograms with an ERS–1 slave image, see Eq. (2.17). Note that there is only a small temporal overlap between data of ERS–1 and ERS–2, which implies that there is a large correlation between estimated azimuth sub-pixel position and displacement that occurred around January 1st, 1996. The amplitude of the processed area and fourteen differential interferograms are shown in Fig. 6.23. The NS–EW street pattern, typical for American cities, can be clearly seen in the (average) amplitude image, as well as highway 95 (upper left to lower right), highway 15 (center to upper right), and the mountains surrounding Las Vegas. The Las Vegas area appears very coherent, even for interferograms with temporal baselines of more than five years. Furthermore, significant atmospheric signal is visible in the interferograms. Fig. 6.24 shows the selected pixels and the constructed reference network.

6.2.2 Estimation of linear displacement

During this first estimation, four parameters are considered; DEM error, linear displacement rate, azimuth sub-pixel position, and average atmosphere. The estimation strategy is to only model the DEM error and the displacement rate for the estimation using the wrapped data, and to additionally estimate the other parameters after phase unwrapping. Ignoring the azimuth sub-pixel position is not expected to have a severe impact on the estimation using wrapped data, because the phase induced by this parameter is relatively small. Aside from this, it is correlated with linear displacement rate, which can be observed from Fig. 6.22; the Doppler frequency is not random in time. Since the azimuth sub-pixel position is not included in the functional model during the initial estimation, the initially estimated linear displacement rates can be slightly biased. However, this bias is expected to be small and spatially uncorrelated, because the azimuth sub-pixel position of the PS points is assumed to have a uniform distribution. Moreover, the azimuth sub-pixel position cannot be estimated with a high precision due to the small variation of the Doppler frequencies. Using a priori variance components (see section 4.3), the vc-matrix of the parameters DEM error (meter), displacement rate (mm/year), and azimuth sub-pixel position (meter) is

$$Q_{\hat{b}} = \begin{bmatrix} 0.114 & 0.016 & -0.050 \\ 0.016 & 0.100 & -0.318 \\ -0.050 & -0.318 & 2.049 \end{bmatrix}, \tag{6.20}$$

which clearly shows that the azimuth sub-pixel position cannot be estimated with the required precision (the standard deviation for this parameter is ~1.4 m while the azimuth resolution for ERS is ~4 m). The corresponding correlation matrix is

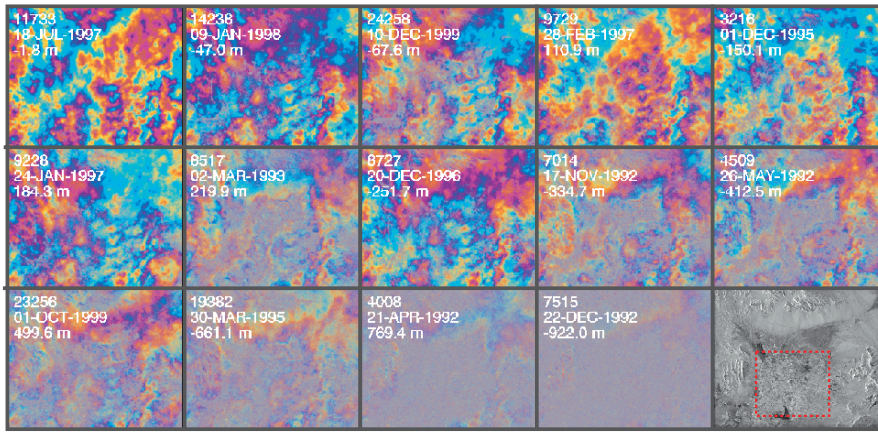
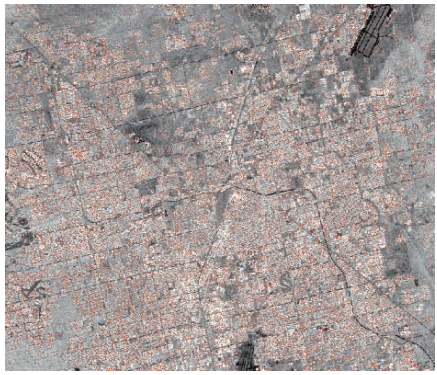
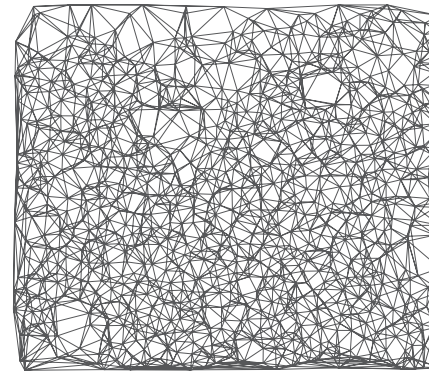


Fig. 6.23: Some differential interferograms for the Las Vegas area. The interferograms are sorted from left to right according to (the absolute value of the) perpendicular baseline, $|B_\perp| \in [2, 1098]$ m. The master image was acquired at June 13$^{\text{th}}$, 1997. The bottom right image shows the mean intensity of the 45 images, scaled to the interval [–20,0] dB. The city area where the estimation is performed is indicate by the rectangle. The images are in the radar coordinate system, i.e., in this case, the images are roughly geo-referenced when they are mirrored in the vertical axis.



(a) Selected points



(b) Reference network

Fig. 6.24: Selected points (red) and constructed reference network for the Las Vegas test site. The images are coarsely geo-referenced by mirroring in vertical axis. The estimation is restricted to 100592 pixels with SCR > 2, see also section 4.2. The points are selected in the city area of approximately 23×20 km^2 (~220 points per km^2). The reference network contains 1084 points, selected using a sparsification with a grid cell width of ~500 m, see section 4.4.1. The number of arcs per point is set to six, which resulted in 4475 arcs with an average length of 880 m.

$$\rho = \begin{bmatrix} 1.000 & 0.150 & -0.104 \\ 0.150 & 1.000 & \mathbf{-0.703} \\ -0.104 & \mathbf{-0.703} & 1.000 \end{bmatrix}. \tag{6.21}$$

This clearly expresses the large correlation between estimates of the linear displacement rate and the azimuth sub-pixel position. Therefore, during the estimation using wrapped data only a DEM error and linear displacement rate are estimated.

First, the variance components for each SLC image are estimated. The 1084 points of the reference network are used to perform 534 independent estimations of the DEM error and the linear displacement rate differences, using the a priori stochastic model. The mean distance between points is 521 m, with a standard deviation of 166 m. The variance components are estimated at each arc separately, cf. Eq. (4.8). The mean of the estimated variance components at these 534 arcs are used to construct the stochastic model, see Eq. (2.51). They are plotted as function of perpendicular, temporal, and Doppler baseline, see Fig. 6.25. The earliest ERS–1 images seem to have a slightly worse precision, as do images with a large perpendicular or Doppler baseline. This can be due to pixels that are not ideal point scatterers, i.e., pixels that (slightly) decorrelate with these baselines, even though the reference network points are selected based on their amplitude dispersion index.

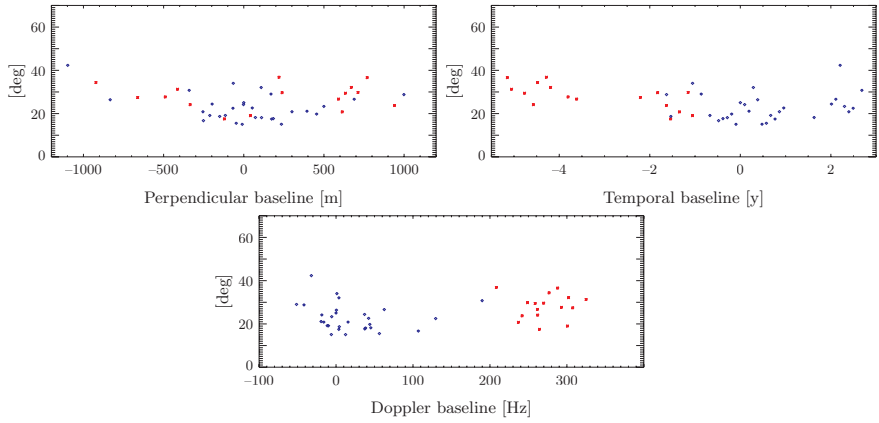
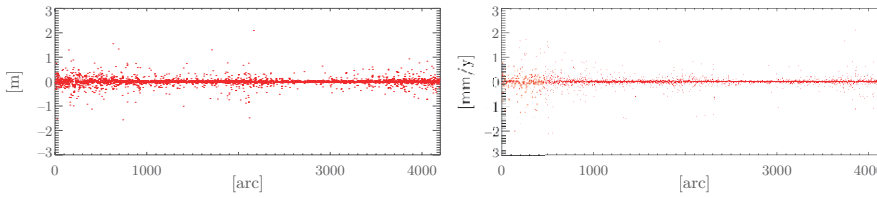


Fig. 6.25: Square roots of the estimated variance components for the Las Vegas test site as function of perpendicular, temporal, and Doppler baseline. A red asterisk corresponds to an ERS–1, and a blue diamond to an ERS–2 acquisition.

Next, the DEM error and the linear displacement rate are estimated with the ILS estimator at the arcs of the reference network. The vc-matrix of the double-difference phase observations is constructed using the estimated variance components. The standard deviation of the pseudo-observations used

for the regularization of the ILS estimator is set to 25 m for the DEM error and 25 mm/y for the linear displacement rate differences. The theoretical success rate for the bootstrap estimator, cf. Eq. (3.20) is $P(\underline{\hat{z}} = z) = 0.990$. After these estimations, a pre-processing step is performed to speed up the alternative hypothesis testing procedure during the network integration. Firstly, points for which the connecting arcs have a mean estimated a posteriori variance factor (cf. Eq. (4.26)) larger than three are iteratively removed. Secondly, arcs with an a posteriori variance factor $\underline{\hat{\sigma}}_x^2 > 3.0$ are removed, as long as each point is still connected with at least three arcs. In this case, in total seventeen points and 244 arcs are removed by this procedure, leaving 1067 points and 4231 arcs in the reference network. After this pre-processing step, the reference point is selected in an area that is known to be stable. Then, the parameters are least-squares adjusted (integrated), as described in section 4.4.2. The testing parameters that are used are $\gamma_0 = 0.90$ and $\alpha_0 = 0.05$, see also Appendix B. During these tests four more points and 43 arcs are removed (~1%). The residuals for the parameters at the arcs of the reference network are shown in Fig. 6.26. Note that not all residuals are exactly equal to zero, as they are expected to be. The maximum absolute residual is 2.09 m and 2.14 mm/y for the DEM error and the linear displacement rate, respectively. The standard deviation of the residual is 0.15 m for the DEM error and 0.20 mm/y for the linear displacement rate. More arcs could be removed until all residuals are zero, but this would hardly have any effect on the estimated parameters at the points of the reference network, and no obvious outlier could be detected anymore. A reason for these non-zero misclosures could be the use of the a posteriori variance factor to create the vc-matrix for the estimated parameters at the arcs, see Eq. (4.12). This is done to down-weight arcs that are estimated with a large a posteriori variance factor since this information would otherwise be lost. Next, the ~100000 points, initially selected based on their SCR, which are not part of the reference network, are estimated relatively to the established reference network. The



(a) DEM error **(b)** Displacement rate

Fig. 6.26: Least-squares residuals at arcs of reference network for the Las Vegas test site. Plotted are 4188 residuals for the DEM error and the linear displacement rate.

CPU time required for these estimations is ~0.03 s per point. The same standard deviation for the pseudo-observations used to regularize the ILS estimator and the same stochastic model are used for these estimations. 58888 points with an a posteriori variance factor $\hat{\underline{\sigma}}_x^2 < 3.0$ are selected as reliable points. Furthermore, eighteen points are removed because the estimated DEM error is outside the interval $[-80, 80]$ m, or the estimated displacement rate is outside the interval $[-30, 30]$ mm/y. The residual phase at these points is unwrapped in the interferograms using the sparse grid MCF algorithm, after which the unwrapped interferometric phase is obtained by addition of the unwrapped residual phase to the model phase, see also section 4.6. Using the unwrapped data, the same or additional parameters can be estimated, see also section 6.2.4, where seasonal displacement is estimated using these unwrapped data. However, the estimated parameters using wrapped data can be inspected already. Fig. 6.27 shows a plot of the estimated linear displacement rates at the accepted points. Wrapped data is used to obtain these estimates. The point density is approximately 130 points per km^2. Clearly visible is the overall

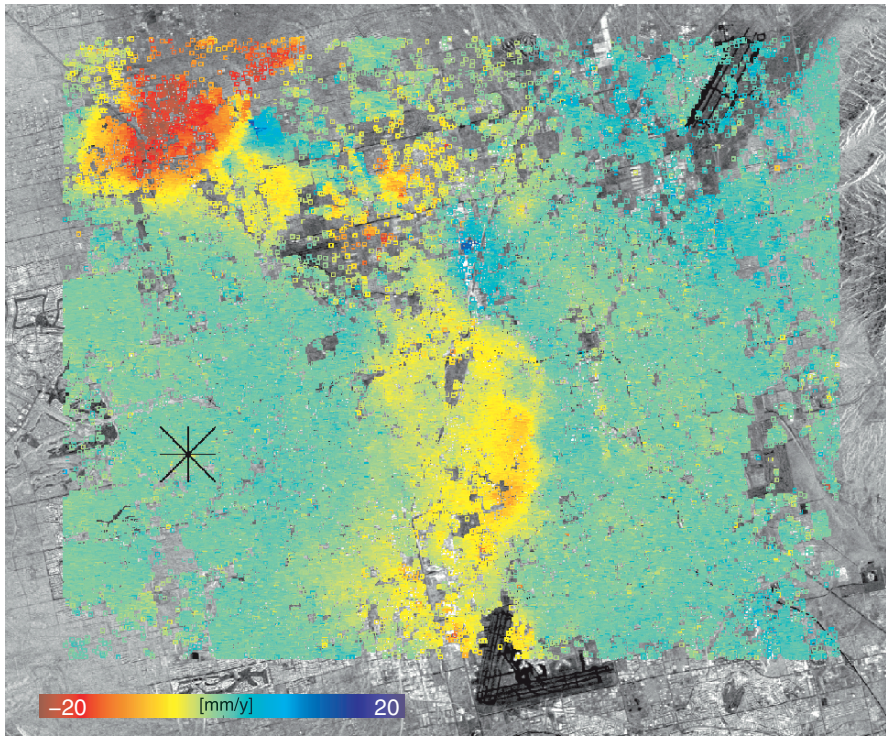
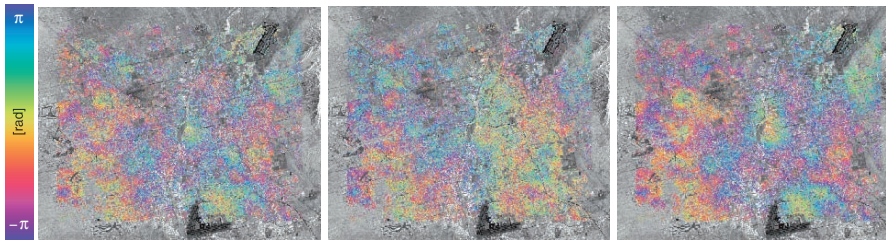


Fig. 6.27: Estimated linear displacement rates using wrapped data at ~60000 points for the Las Vegas test site. Area is approximately 23×20 km^2. 45 ERS acquisition are used, acquired between April 21st, 1992, and February 18th, 2000.

subsidence, occurring from the bottom of this image, to the center, and then to the upper left corner. This displacement pattern is mainly caused by three localized subsidence bowls (Northern, Central, and Southern bowl), which were recognized before 1980 (Bell et al., 2002). Amelung et al. (1999) measured a maximum subsidence of 190 mm between April 1992 and December 1997 (33 mm/y average) using four differential interferograms for the Northern bowl, and 110 mm (19 mm/y average) for the Central bowl, which reasonably agrees with these estimates. Furthermore, there seems to be a subsidence bowl located roughly between the Northern and Central bowls, and some localized uplift areas (lower right, slightly above the center, on the right of the major subsidence bowl in the Northwest, and in the upper right corner). The two uplift areas on the right were also identified by Bell et al. (2002).

The phase residuals at the points after estimation of the DEM error and the linear displacement rate are shown for the first three interferograms in Fig. 6.28. The spatial correlation of the phase residuals suggests that interferometric atmospheric signal is contained in the residuals. A spatial low-pass filter could be used to estimate this signal. However, spatially correlated unmodeled displacement could also be present in the residuals.



(a) 12024 **(b)** 22254 **(c)** 3216

Fig. 6.28: Residual interferometric phase after estimation of ithe DEM error and the displacement rate for the first three interferograms of the Las Vegas data set, orbits 12024, 22254 and 3216. The residuals are spatially correlated, which suggests that atmospheric signal is contained in these residuals.

Estimation of average atmospheric phase

The average atmospheric interferometric phase can be estimated using the wrapped data by defining an additional base function as

$$p_2(T^k) = 1, \quad k = 1, \ldots, K. \tag{6.22}$$

With this base function a bias in the double-difference interferometric phase observations can be estimated, i.e., the average interferometric atmospheric

phase, see also Eq. (2.29). Note that the base functions are normally used to model displacement, cf. Eq. (2.14), and this parameter is not related to displacement. However, the generic concept of defining base functions to estimate parameters can be used for this parameter as well. This has the advantage that the software does not need to be adapted for this choice, but only the input.

There is a small correlation between the estimated parameter for the average atmosphere and the other parameters in this case, which is evident from the vc-matrix

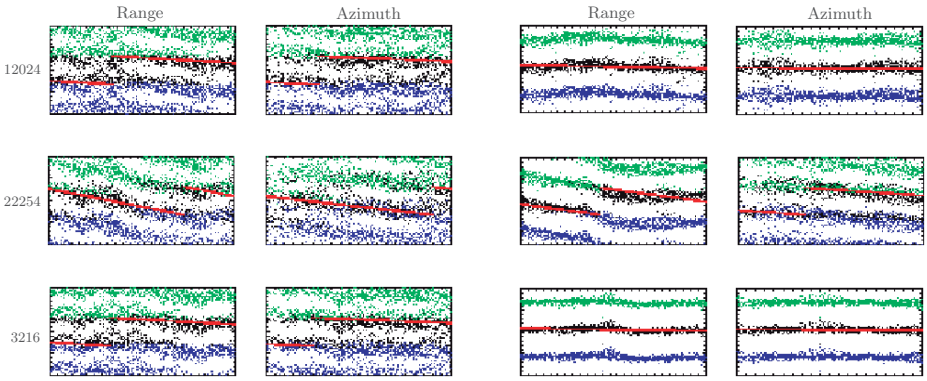
$$Q_{\hat{b}} = \begin{bmatrix} 0.113 & 0.008 & -0.006 \\ 0.008 & 0.041 & -0.008 \\ -0.006 & -0.008 & 0.258 \end{bmatrix}, \tag{6.23}$$

and corresponding correlation matrix

$$\rho = \begin{bmatrix} 1.000 & 0.106 & -0.037 \\ 0.106 & 1.000 & -0.073 \\ -0.037 & -0.073 & 1.000 \end{bmatrix}. \tag{6.24}$$

A priori variance components are used to compute these matrices. The parameters are DEM error in meters, linear displacement rate in mm/y, and average atmosphere in radians. It is verified that the estimated DEM error and the displacement rate are practically identical with and without using this base function. Of the 4475 arcs of the reference network, 4240 have an estimated a posteriori variance factor $\underline{\hat{\sigma}}_x^2 < 3.0$. For these arcs, the mean difference between the estimates is 0.000 m and 0.000 mm/y for the DEM error and the linear displacement rate, respectively, and the standard deviation is 0.021 m and 0.012 mm/y. The maximum absolute difference is 0.58 m and 0.43 mm/y for these arcs.

Although the estimated DEM errors and linear displacement rates are similar, the least-squares interferometric phase residuals are not equal for both cases. Adding a parameter for the average atmospheric phase to the functional model significantly reduces the residual phase for the Las Vegas test site. This is shown in Fig. 6.29, where the residual phase at the points of the reference network are plotted as function of range and azimuth for a few selected interferograms. Clearly, the residual phase is smaller when the average atmosphere is estimated. Fig. 6.30 shows the estimated average atmosphere spatially for all points. The spatial correlation indicates that these estimates can be interpreted as atmospheric signal. This parameter can be estimated using the wrapped data or the unwrapped data. Since the atmospheric signal is spatially correlated, the correctly unwrapped phase can be obtained also if this parameter is not estimated using the wrapped data. Therefore, the base function for this parameter is introduced after the phase unwrapping.



(a) Avg. atmosphere not estimated **(b)** Avg. atmosphere estimated

Fig. 6.29: Residual phase after estimation of DEM errors and displacement rates (left), compared to additional estimation of average atmosphere (right). For the first three interferograms of the Las Vegas stack, orbits 12024, 22254 and 3216, the residual phase at the points of the reference network is shown as function of range and azimuth coordinates. Two ambiguity levels are plotted for easier interpretation (green and blue). The red line corresponds to the estimated trend. Clearly, the residuals are much smaller if an additional parameter is estimated accounting for the average interferometric atmosphere.

Estimation of azimuth sub-pixel position

The DEM errors, linear displacement rates, and azimuth sub-pixel positions are simultaneously estimated using the unwrapped phase data at the ~60000 unwrapped points. The vc-matrix for this choice of parameters is given in Eq. (6.20). From this matrix, and the correlation matrix, Eq. (6.21), it is clear that the azimuth sub-pixel position should not be estimated, due to the relatively small variation and temporal correlation of the Doppler centroid frequencies of the acquisitons, see also Fig. 6.22. Nonetheless, this estimation is performed. The estimated DEM errors and linear displacement rates are visually identical to the solutions obtained using a functional model that does not include the azimuth sub-pixel position, see also Fig. 6.27. The mean of the estimated DEM error at these points is –1.68 m, with a standard deviation of 3.90 m. The mean and standard deviation of the estimated displacement rates are –0.91 mm/y and 2.70 mm/y, respectively.

The estimated azimuth sub-pixel positions for these points is shown in Fig. 6.31 in the range between –4 and 4 m, the azimuth resolution of ERS. Clearly, unmodeled displacement is leaked to the estimated azimuth sub-pixel position. The sub-pixel position is expected to be spatially uncorrelated, but the estimates are not. This is caused by the distribution of the Doppler centroid frequencies of this dataset. Consider Fig. 6.22 again. The estimated azimuth sub-pixel position is a linear function of the Doppler centroid

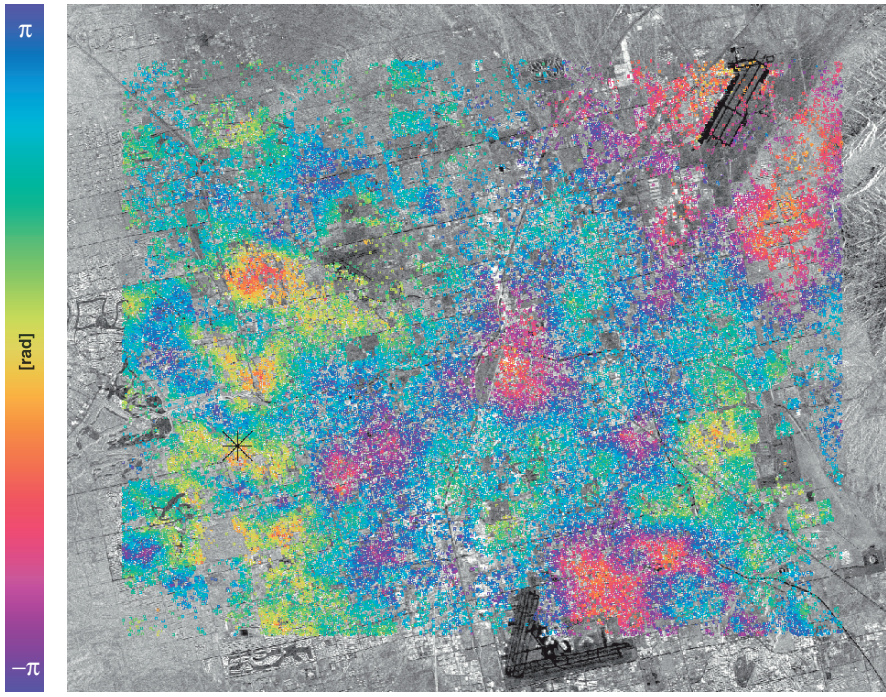


Fig. 6.30: Average atmospheric delay estimates at ~60000 points for the Las Vegas test site. Note that some features of the average atmospheric phase are visible in the interferometric residual phase images shown in Fig. 6.28, particularly at the center, lower right, and upper right, for orbits 12024 and 3216.

frequency difference between master and slave image, see Eq. (2.22). Thus, displacement that occurs according to the pattern of the Doppler centroid frequencies, i.e., around January–June, 1996, is estimated as azimuth position. Therefore, the azimuth sub-pixel position should not be estimated if data are used with a Doppler centroid frequencies that is strongly correlated with time. Instead, the sub-pixel positions should be estimated using a point target analysis, and the phase interpolated at these positions in the interferograms.

6.2.3 Comparison with the reference PS technique

A Standard PS Analysis (SPSA) of the Las Vegas area was carried out by Tele-Rilevamento Europa[2] (TRE), independently from the estimation with the STUN algorithm. In their terminology, an SPSA analysis is intended for large scale application, mainly to identify stable areas and to highlight possible risk areas. The minimum size of the area is approximately 100 km^2. The estimated parameters are DEM error, displacement rate, and the APS for each

[2] a POLIMI spin-off company

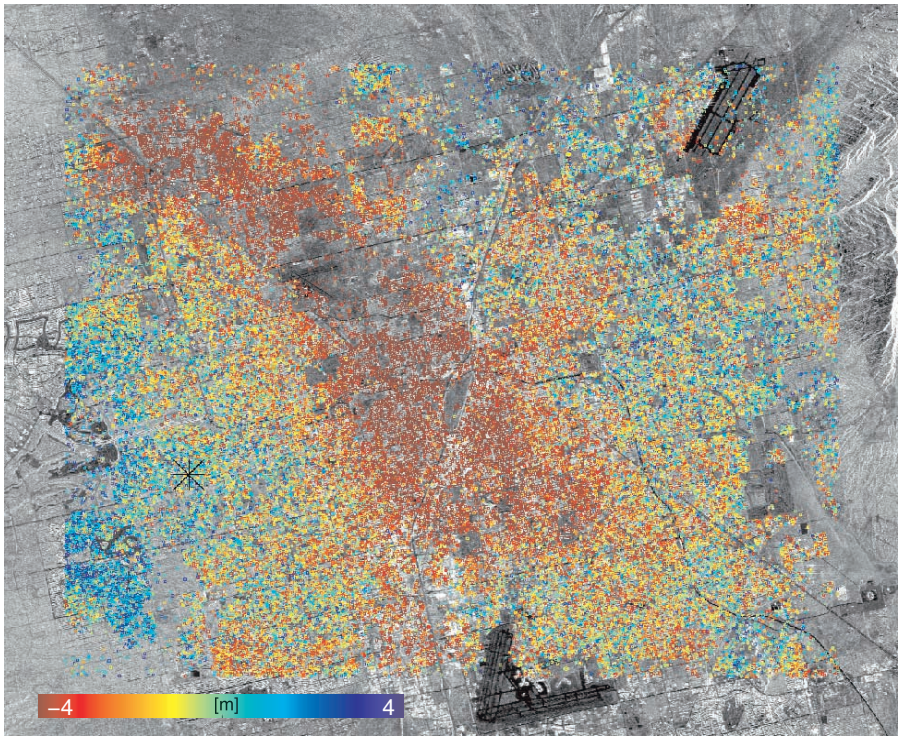


Fig. 6.31: Estimated azimuth sub-pixel positions for the Las Vegas test site. Clearly, unmodeled displacement is leaked to the estimated positions, which are expected to be random. This is due to the small variation and temporal correlation of the Doppler centroid frequencies of the acquisitions. For this dataset, azimuth sub-pixel position should not be estimated. The azimuth direction corresponds to the vertical axis.

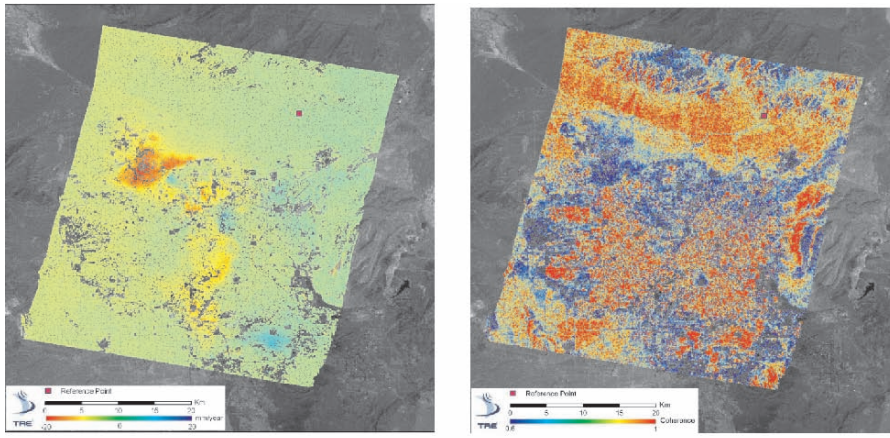
acquisition, see also (Tele-Rilevamento Europa, 2004). After the SPSA, an Advanced PS Analysis (APSA) can be carried out. This advanced processing is a high-resolution analysis of small areas, of ∼1 km^2, to assess the stability of individual buildings or structures, such as dams, industrial plants, etc.

Nearly the same data set is used for the estimation with the reference technique and with the STUN algorithm. TRE has three additional SLC images available, acquired at September 21st, 1995, January 4th, 1996 (both ERS–1), April 28th, 2000 (ERS–2), and included the acquisition of August 11th, 2000, while the last image used for the STUN estimation was acquired at February 28th, 2000, see also section 6.2.1 and Appendix C. The master used for the SPSA was acquired at February 28th, 1997, i.e., 105 days before the master used in the STUN algorithm. However, the displacement rates that are estimated from the acquisitions between 1992 and 2000 are not expected to be affected by the choice of the master image. The reference point that is used by TRE is located on a geologically stable layer, outside the displacement

area, whereas the reference point used by the STUN algorithm is located in an area possibly undergoing slight subsidence. This different choice can introduce a constant bias between the displacement rates that are estimated by both algorithms. Moreover, it can cause differences in the (displacement) phase time series, since these are relative to the selected reference point. The parameters that are estimated using the SPSA are the DEM error, the line-of-sight displacement rate, and the APS for each acquisition. No temporal filtering is performed to compute the APS. The reference PS technique uses the ensemble coherence maximization for the estimation of these parameters. The estimated precision is better than 1 mm/y for the displacement rates and in the order of a few mm for the individual APS corrected measurements, for PS within ~5 km from the reference point (Ferretti et al., 2000a, 2001).

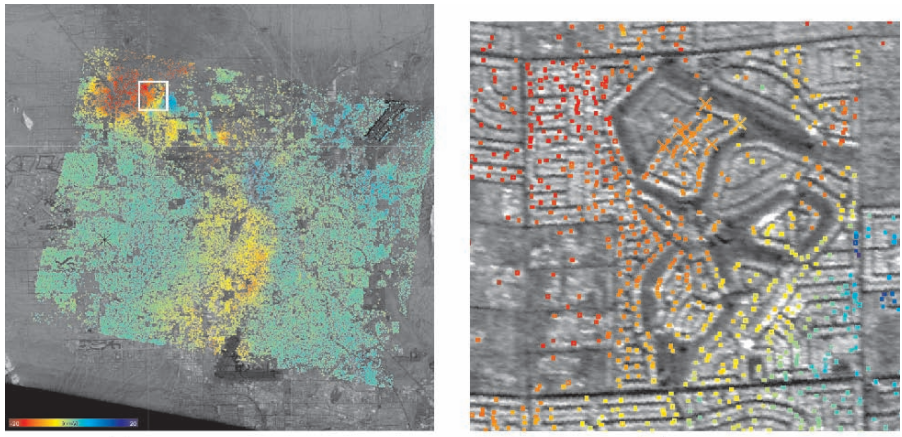
Fig. 6.32 shows the estimated linear displacement rates and coherence using the reference technique. Data are geo-referenced and superimposed on a LandSat image. For comparison, the estimates using the STUN algorithm are geo-referenced as well, see Fig. 6.33. The area processed by TRE is clearly somewhat larger than the area processed using the STUN algorithm. Plotted are 157018 points with a minimum estimated coherence $|\hat{\gamma}| = 0.63$. For visualization purposes, only the point with the highest coherence in each 100×100 m^2 cell is plotted by TRE. However, the density of the estimated points using the SPSA seems larger, also outside the urbanized area. For example, a large amount of points are estimated with high coherence north of the city. This suggests that pixels with a distributed scattering mechanism are estimated as well, apparently with good results. Recall that the estimation using the STUN algorithm is limited to an initially selected set of pixels, see also section 4.2. However, the estimated displacement rates compare well to each other. The same spatial displacement features and similar magnitudes of the displacement rates are estimated by both processing techniques. Note that the colorbar used by TRE contains more green in the middle and slightly less intense red and blue at the edges.

Finally, the displacement time series of a few points are compared. The position of the points is near the Northern subsidence bowl, see Fig. 6.33. The estimated displacement rates and quality of these points are listed in Table 6.5 for the SPSA and STUN algorithm. In general there is a good agreement between the estimated displacement rates at the selected points. A constant bias between the estimated displacement rate is acceptable, because it can be due to relative movement between the reference points used during both estimations. The mean difference between the estimated displacement rates at these eight points is –1.36 mm/y (supporting the theory that the reference point used by the STUN algorithm is subsiding slightly) and the standard deviation of the difference is 0.77 mm/y. The difference between the estimated displacement rates is the largest for the sixth point. If this point is not considered, the mean and standard deviation are 1.13 and 0.47 mm/y, respectively, which is within the estimated precision for the displacement rates using the STUN algorithm. Furthermore, the estimated coherence (SPSA) is



(a) Displacement rate (b) Coherence

Fig. 6.32: Linear displacement rates and ensemble coherence estimated using the Standard PS Analysis (reference technique). Data processed by ©Tele-Rilevamento Europa.



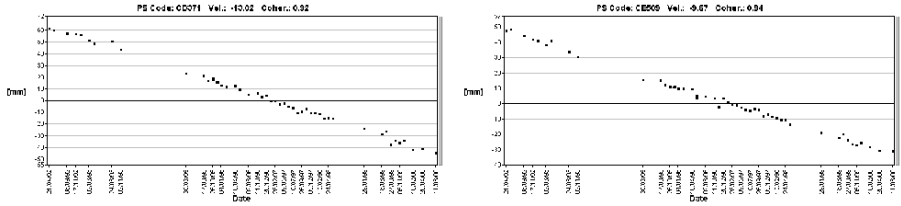
(a) Displacement rate (b) Zoom

Fig. 6.33: Estimated linear displacement rates using the STUN algorithm. Data are geo-referenced (UTM projection, zone 11) and plotted on top of the temporally averaged radar intensity map. The white rectangle indicates the zoom area shown in **(b)**. The area is $\sim 2 \times 2$ km^2. Displacement time series estimated using the reference PS technique are available for the eight points marked with an $\times$.

Table 6.5: Line-of-sight displacement rates estimated with the STUN and reference technique (SPSA estimates and geographical coordinates provided by TRE). Position of the PS points given in Fig. 6.33. Estimates using STUN after unwrapping; estimation included a parameter for the average atmosphere. Estimated precision $\hat{\underline{\sigma}}_\alpha$ using STUN algorithm relative to reference point at distance ~10 km. A bias of the difference between both estimates can be due to relative motion between the different reference points used.

	SPSA		STUN		$\hat{\underline{\alpha}}^{\mathrm{SPSA}} - \hat{\underline{\alpha}}^{\mathrm{STUN}}$
#	$\hat{\underline{\alpha}}$ [mm/y]	$\lvert\hat{\gamma}\rvert$ [-]	$\hat{\underline{\alpha}}$ [mm/y]	$\hat{\underline{\sigma}}_\alpha$ [mm/y]	[mm/y]
1	–12.21	0.92	–10.92	0.47	–1.29
2	–12.05	0.92	–10.57	0.45	–1.48
3	–12.12	0.91	–11.17	0.49	–0.95
4	–10.41	0.93	–10.15	0.50	–0.26
5	–11.33	0.91	–10.30	0.48	–1.03
6	**–13.02**	**0.92**	**–10.11**	**0.47**	**–2.91**
7	–11.36	0.92	–10.19	0.46	–1.17
8	**–9.67**	**0.94**	**–7.91**	**0.56**	**–1.76**

the highest for the eighth point while the estimated precision using the STUN algorithm is the worst for this point. It is possible that point identification errors are made causing these differences.



(a) Point 6 **(b)** Point 8

Fig. 6.34: Displacement time series using the Standard PS Analysis (reference technique). Data processed by ©Tele-Rilevamento Europa.

The displacement times series for these two interesting points are shown in Fig. 6.34 (as provided by TRE). The results of the STUN algorithm at these two points are shown in Fig. 6.35. The estimated mean displacement rate is also plotted for the STUN and for the SPSA algorithm in this figure. The time series of the STUN algorithm appears more noisy. This is expected since the APS is estimated and removed during the SPSA using a spatial low-pass filter, while this is not done for the estimation using the STUN algorithm. Note that a possible phase unwrapping error would manifest as a jump of ~28 mm, i.e., half the wavelength used by ERS, in these plots. Such errors are not apparent, but could explain the difference between the estimated

displacement rates at these points. Particularly, a larger subsidence rate would likely be estimated by the STUN algorithm if 28 mm (i.e., phase of 2π) would be added to some of the earlier displacement observations. There is no reason to do so, i.e., to assume that a phase unwrapping error is made by the STUN algorithm. However, it could be that the PS point is not present in some or all of the earlier acquisitions, which would lead to random phase values, which cannot be expected to be unwrapped correctly. Moreover, the time series shown for the STUN algorithm are relative to the selected reference point. Instability of this point (i.e., unmodeled displacement or random phase during a few acquisitions) would be visible as a random phase in the displacement time series. To eliminate this effect and to reduce the atmospheric phase contribution in the estimation using the STUN algorithm, Fig. 6.36 shows the difference between relatively nearby points. Furthermore, the intensity of the observed data is plotted. Indeed, the small intensity values of approximately –15 dB for some of the earlier acquisitions for point six suggest that the corresponding phase may not be reliable. However, a clear relation between intensity and phase precision is not apparent. The error bars indicate the one-sigma level. They are computed using the a posteriori variance factors. Moreover, the deviation of the linear model (noise) is considerable, even for the difference between the extremely nearby points six and eight. It is interesting that the noise (i.e., atmospheric signal) does not seem to increase if the distance between the points is increased to ~2 km.

6.2.4 Estimation of linear and seasonal displacement

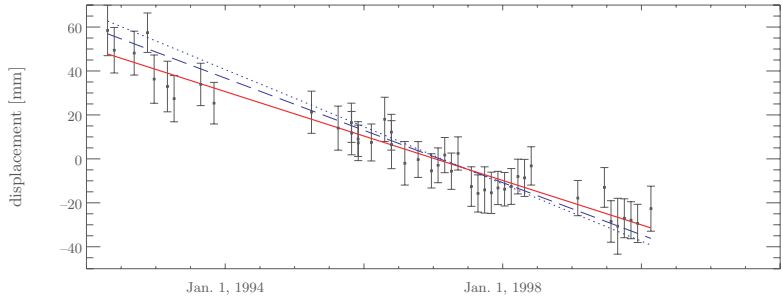
The displacement of the Las Vegas test site is known to have a seasonal component, see, for example, (Hoffmann et al., 2001). Therefore, one linear and two trigonometric base functions are used to model the displacement

$$\begin{aligned} p_1(T^k) &= -\frac{4\pi}{\lambda} T^k, \\ p_2(T^k) &= -\frac{4\pi}{\lambda} \sin(2\pi T^k), \\ p_3(T^k) &= -\frac{4\pi}{\lambda} \left(\cos(2\pi T^k) - 1\right). \end{aligned} \tag{6.25}$$

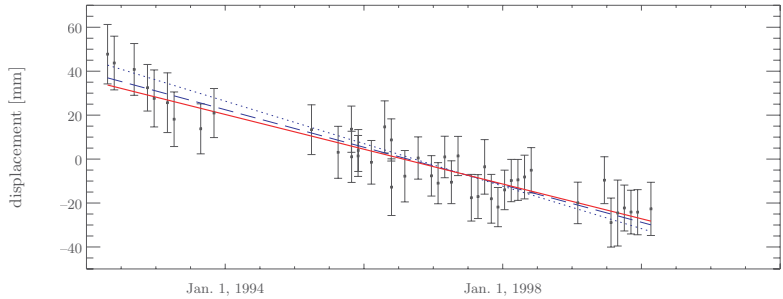
For easier interpretation, the coefficients of the trigonometric base functions are transformed to a seasonal displacement function with certain amplitude A (in mm) and certain temporal offset t_0 (in years) as

$$\begin{aligned} \Delta r_x^k &= \alpha_2(x) \cdot p_2(k) + \alpha_3(x) \cdot p_3(k) \\ &= -\frac{4\pi}{\lambda} (A \cdot (\sin(2\pi(T^k - t_0))) + \sin(2\pi t_0)), \end{aligned} \tag{6.26}$$

where $t_0 = \arctan(-\alpha_3/\alpha_2)/(2\pi)$ and $A = \alpha_2/\cos(2\pi t_0)$. The term $\sin(2\pi t_0)$ in Eq. (6.26) ensures that the displacement function equals zero at the time of the

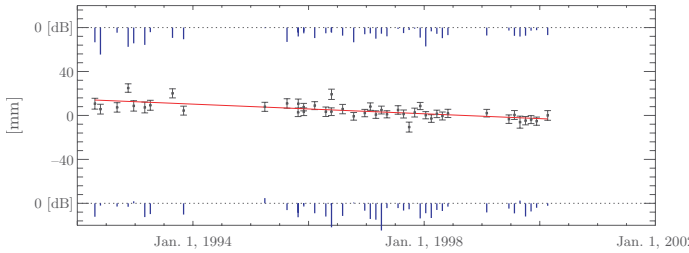


(a) Point 6

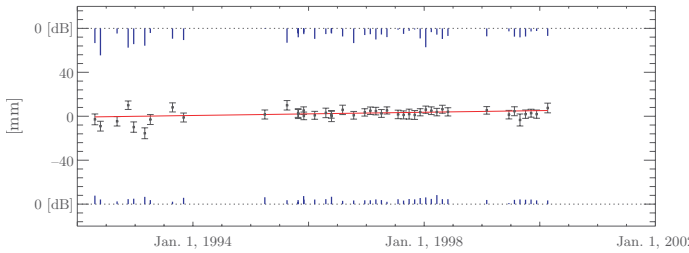


(b) Point 8

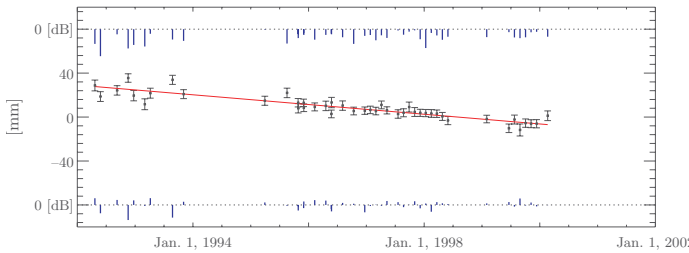
Fig. 6.35: Displacement time series with respect to the reference point using the STUN algorithm. The unwrapped phase data are plotted, corrected for the estimated DEM error and the average atmospheric phase, at the two points that deviate the most from estimates obtained using the reference technique. The error bars show the a posteriori error on the interferometric double-difference phase (one-sigma level, converted to mm). The estimated displacement rate is plotted as a red line. The SPSA estimate obtained by TRE is plotted in blue (dotted line: original estimate; dashed line: corrected by 1.13 mm/y to account for the mean difference between STUN and SPSA).



(a) Distance ∼200 m



(b) Distance ∼500 m



(c) Distance ∼1150 m

Fig. 6.36: Relative displacement time series between nearby points using the STUN algorithm. **(a)** shows the difference between the nearby points 6 and 8, see also Fig. 6.35. The error bars are computed for the double-difference phase observations using the estimated variance components (valid for nearby points). Note the large deviations from the displacement model although atmospheric phase is not expected to be present. The intensity (in dB) of the points in the slave images are given above and below the displacement values for point 6 and point 8, respectively. The intensity for these points varies, but it is clearly above the average for this area, which is approximately –11 dB. **(b)** shows the difference between point 6 and a bright point ∼500 m to the South, and **(c)** the difference with a point ∼1150 m to the South.

master acquisition. Note that these parameters cannot be directly estimated using wrapped data.

The cause for the (localized) seasonal displacement is water extraction from underground reservoirs in summer, and natural and artificial recharge of the groundwater system in winter. The sinusoidal model does not perfectly model the displacement, because the amount of water extraction and refill varies from year to year, and the duration of the extraction is longer than that of the refill period. Using this information, a single dedicated base function could be used locally for each well. For example, Fig. 6.37 shows borehole extensometer data of the Lorenzi site that could be used as a proxy for the displacement, particularly near that site. The location of this site is approximately at the Northern subsidence bowl, see also Fig. 6.27. An extensometer measures the distance (change) between an anchor point at a certain depth and the surface. The typical precision of these measurements is better than 1 mm. However, the amplitude of the measurements made by the extensometer does not necessarily have to agree with the InSAR data, depending on the stability of the anchor point. The extensometer is not sensitive to movement below the anchor point and the difference between extensometer and InSAR data depends on how such movements manifest as surface displacement (which is observed by InSAR). However, the two-parameter sinusoidal displacement model is used as an approximate model. The vc-matrix is given as

$$Q_{\hat{b}} = \begin{bmatrix} 0.102 & 0.005 & 0.002 & -0.018 \\ 0.005 & 0.036 & 0.011 & 0.004 \\ 0.002 & 0.011 & 0.265 & 0.026 \\ -0.018 & 0.004 & 0.026 & 0.241 \end{bmatrix}, \tag{6.27}$$

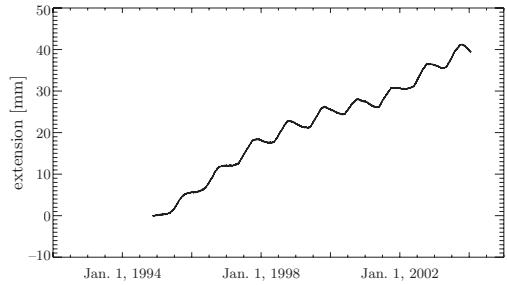


Fig. 6.37: Borehole extensometer data of the Lorenzi test site. Data courtesy USGS.

where the order of the parameters is DEM error (m), linear displacement rate (mm/y), amplitude of sinusoid term (mm), and cosine term (mm). The maximum correlation coefficient is $\rho_{2,3} = 0.23$, i.e., there is no significant

correlation between estimated parameters. The estimated a posteriori variance components are used to compute this matrix. These variance components are almost identical to the ones that are estimated during the reference processing, see Fig. 6.25. The theoretical success rate for the bootstrap estimator is $P(\hat{\underline{z}} = z) = 0.973$, cf. Eq. (3.20), using these variance components.

The base functions given in Eq. (6.25) can be used for the estimation using wrapped and unwrapped data. Alternatively, the unwrapped data could be obtained using a linear displacement model, followed by estimation of seasonal displacement using the unwrapped phase. In the following these two processing appoaches are described. First, these base functions are used from the start, i.e., the reference network is computed using these three displacement parameters as well as the other points. The data is then unwrapped using the sparse grid unwrapping algorithm, and the final estimation is performed using the unwrapped data. As an alternative, the second estimation of the seasonal component is performed using the unwrapped phase that is obtained after estimation using the linear displacement model, see section 6.2.2.

Estimation of seasonal displacement using wrapped data

The displacement is modeled with the three base functions given in Eq. (6.25). The parameters of these base functions are estimated simultaneously with the DEM error at the arcs of the reference network. The a posteriori variance components are used to construct the vc-matrix used by the ILS estimator. The standard deviation for the pseudo-observations for the DEM error is set to 25 m, to 10 mm/y for the linear displacement, and to 10 mm for the sine and cosine terms. Note that a smaller value for the linear displacement parameter is used here than in the previous scenarios. This is done to prevent estimation of unrealistically large coefficients for the linear and seasonal terms that, when combined, could better fit the wrapped data. After the ILS estimation at the arcs, the estimated differences are spatially integrated and tested, as described in section 6.2.2. During a pre-processing step, 21 points and 330 arcs are removed from the reference network. One more point and eleven arcs are removed during the following alternative hypothesis testing. The least-squares residuals for the estimated parameters at the arcs after the alternative hypothesis testing step are shown in Fig. 6.38. The mean residual for all terms is 0.000, while the standard deviations are 0.16 m, 0.28 mm/y, 0.18 mm, and 0.15 mm, respectively. After the integration of the estimated difference parameters, the same four parameters are estimated at all other points with respect to the network. Finally, 52242 points are selected with an estimated a posteriori variance factor $\hat{\underline{\sigma}}_x^2 < 3.0$. Previously, when less parameters were estimated, ~60000 points were detected using the same threshold for the a posteriori variance factor. The reason that more points are removed is the smaller redundancy due to the two additional base functions. Points that do not undergo seasonal displacement have identical least-squares phase residuals, but a larger estimated variance factor. Consequently, less points are

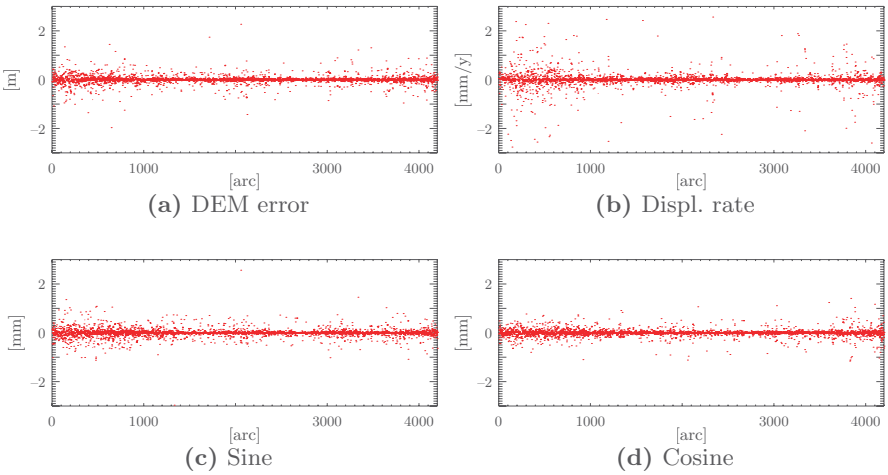
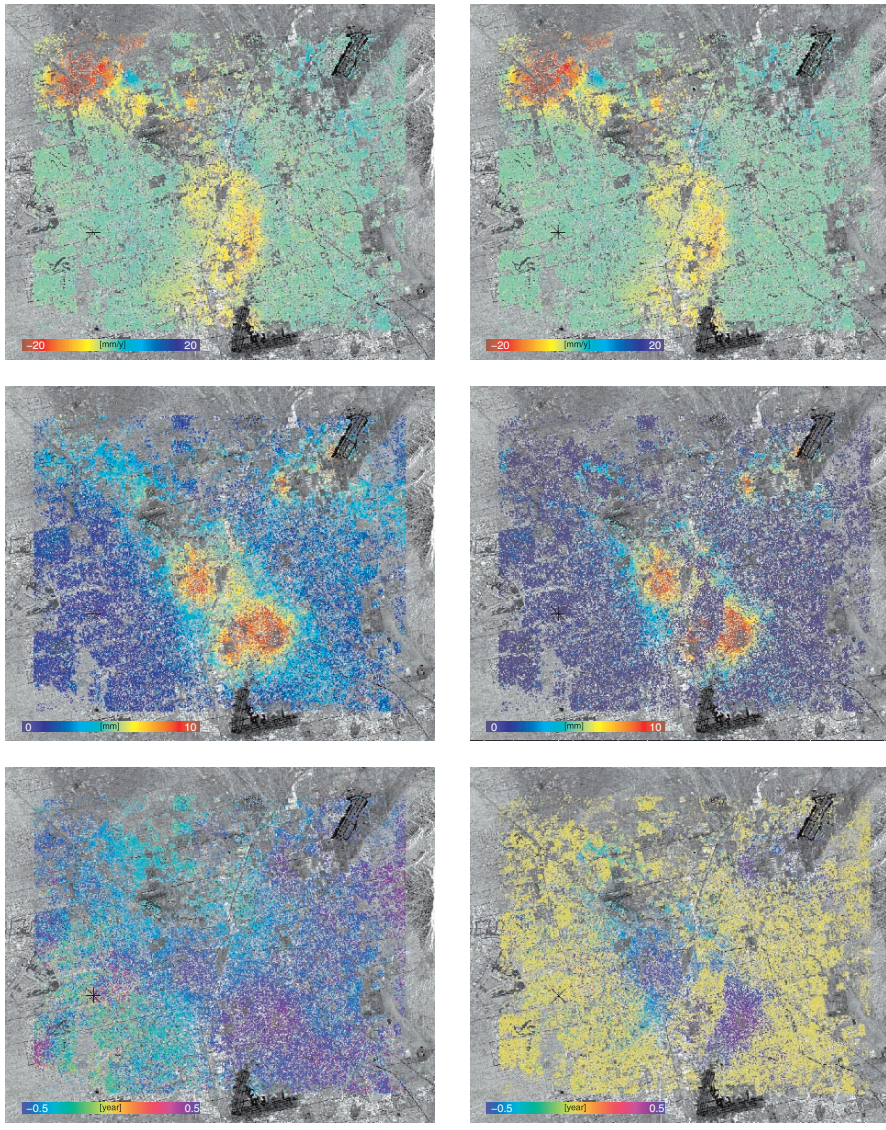


Fig. 6.38: Least-squares residuals at arcs of reference network for the Las Vegas test site. Plotted are 4216 residuals for the DEM error, linear displacement rate, sine and cosine term of the seasonal displacement.

selected. Moreover, the estimated variance factors are now marginally smaller, which also implies that more points will be removed if the same threshold for the a posteriori variance factor is used. The phase at the selected points is unwrapped using the MCF sparse grid unwrapping algorithm. The final estimation is performed after unwrapping of the data at the selected points. The same three base functions are used to model the displacement as for the estimation using wrapped data. Additionally, the average interferometric atmosphere is estimated using the unwrapped data, see also Eq. (6.22). The vc-matrix for this choice of estimated parameters (DEM error, linear displacement, sinus and cosine term, average interferometric phase.) is given as

$$Q_{\hat{b}} = \begin{bmatrix} 0.088 & 0.005 & 0.002 & -0.019 & -0.010 \\ 0.005 & 0.036 & 0.012 & 0.004 & 0.004 \\ 0.002 & 0.012 & 0.266 & 0.029 & 0.016 \\ -0.019 & 0.004 & 0.029 & 0.255 & 0.070 \\ -0.010 & 0.004 & 0.016 & 0.070 & 0.371 \end{bmatrix}. \tag{6.28}$$

The estimated displacement parameters are shown in Fig. 6.39(a). Clearly, the estimated amplitude of the seasonal displacement is significant for the central subsidence bowl, i.e., ∼10 mm or ∼2.25 rad for ERS. Furthermore, the estimated offset is spatially very consistent, while it is estimated independently for each point. The average estimated offset is ∼0.5 year at the positions with the largest amplitude. Since the master image is acquired June 13^{th}, the maximum (relative uplift) of the seasonal term thus occurs around March and the minimum (additional subsidence) around September.



(a) No significance tests **(b)** With significance tests

Fig. 6.39: Las Vegas linear and seasonal displacement. **(a)** shows the estimated linear displacement rate, the amplitude and the offset of the seasonal displacement, cf. Eq. (6.26). **(b)** shows the same parameters after significance tests are performed.

The precision of the estimated parameters using unwrapped data is given by the a posteriori variance factor, according to Eq. (4.33), see Fig. 6.40. This variance factor is a scaling factor for the vc-matrix given in Eq. (6.28). For example, if the variance factor is equal to ten, then, in this case, the variance of the estimated DEM error is 0.88 m^2 and 0.36 $(mm/y)^2$ for the linear displacement rate. The precision of the estimated parameters decreases with distance to the reference points, which is due to the atmospheric signal. The variance factors at the location of the Northern subsidence bowl are larger than the average of the factors surrounding it. This indicates that unmodeled displacement is still present in the residuals. However, if it is assumed that the residuals only contain atmospheric signal and uncorrelated noise, a spatial low-pass filter could be applied to separate these components. The variance factors of the estimation using the corrected data are expected to be much smaller, approximately one. It must be accepted in this case that spatially correlated unmodeled displacement cannot be estimated.

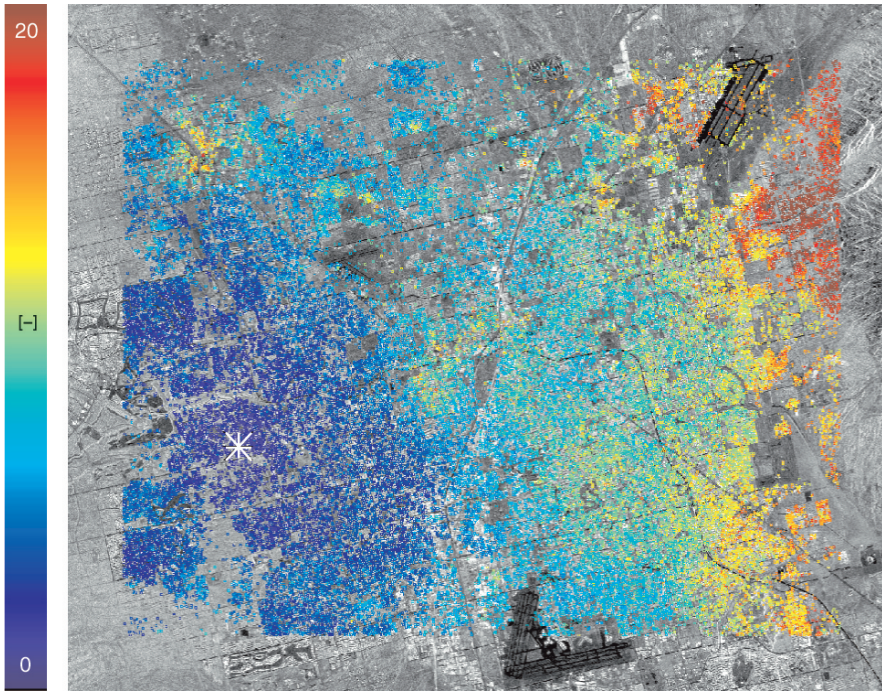


Fig. 6.40: A posteriori variance factors for the Las Vegas test site. The asterisk indicates the reference point. In general, the precision decreases with distance from the reference point.

Significance tests

However, the parameters cannot be estimated significantly for all points. For example, the estimated offsets of the seasonal term are probably meaningless at the sides of the processed area where the amplitudes are practically zero, see also Fig. 6.39(a). Therefore, a hypothesis testing procedure is followed, to test the significance of the estimated displacement parameters. The least relaxed model is used as the null-hypothesis, i.e., no displacement is assumed, but only parameters for a DEM error and average interferometric atmosphere are modeled. The null-hypothesis is thus given by

$$H_0 : E\{\underline{y}\} = \begin{bmatrix} \beta^1 & 1 \\ \beta^2 & 1 \\ \vdots & \vdots \\ \beta^K & 1 \end{bmatrix} \begin{bmatrix} \Delta h \\ \bar{S} \end{bmatrix}. \tag{6.29}$$

The first alternative hypothesis extends the null-hypothesis using a linear displacement model, i.e.,

$$H_A^1 : E\{\underline{y}\} = \begin{bmatrix} \beta^1 & 1 \\ \beta^2 & 1 \\ \vdots & \vdots \\ \beta^K & 1 \end{bmatrix} \begin{bmatrix} \Delta h \\ \bar{S} \end{bmatrix} + \begin{bmatrix} p_1(1) \\ p_1(2) \\ \vdots \\ p_1(K) \end{bmatrix} \begin{bmatrix} \alpha_1 \end{bmatrix}. \tag{6.30}$$

The second alternative hypothesis extends the first alternative hypothesis further to account for seasonal displacement

$$H_A^2 : E\{\underline{y}\} = \begin{bmatrix} \beta^1 & 1 \\ \beta^2 & 1 \\ \vdots & \vdots \\ \beta^K & 1 \end{bmatrix} \begin{bmatrix} \Delta h \\ \bar{S} \end{bmatrix} + \begin{bmatrix} p_1(1) & p_2(1) & p_3(1) \\ p_1(2) & p_2(2) & p_3(2) \\ \vdots & \vdots & \vdots \\ p_1(K) & p_2(K) & p_3(K) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}. \tag{6.31}$$

The procedure which is usually followed to perform these tests is to first set the power for all tests and the level of significance for the one-dimensional test. If the null-hypothesis is rejected, the test quotient is computed for the specified alternative hypotheses. The alternative hypothesis with the largest test quotient is selected as the most likely one, see Appendix B. The vc-matrix of the observations is generally assumed to be known. (Because otherwise, for example, the null-hypothesis would almost never be rejected if the precision of the observations would be described very pessimistically.) However, the vc-matrix is not known in this case, since the atmospheric signal (at lower frequencies) is not accounted for in the estimated variance components. Therefore, the following procedure is followed during the significance tests:

1. Perform the least-squares adjustment under the most relaxed model and estimate a variance factor for the vc-matrix used during the estimation with wrapped data.

2. Perform the least-squares adjustment under the null-hypothesis and compute the overall model test using the updated stochastic model.
3. If the overall model test is rejected, compute the test statistics for the alternative hypotheses and select the one with the largest test quotient as the most likely hypothesis.
4. Compute the parameters for the estimated hypothesis and an a posteriori variance factor.

The power of all tests is set to $\gamma_0 = 50\%$ and the level of significance for the one-dimensional test to $\alpha_0 = 1\%$. The tests are coupled using the B-method of testing.

The null-hypothesis (no displacement) is not rejected at 28378 points. At the 23864 points where the null-hypothesis is rejected, at 11964 the first alternative hypothesis is selected (linear displacement), and at 11900 points the second alternative hypothesis (linear and seasonal displacement). Fig. 6.39(b) shows the estimated displacement parameters. Compared to the estimated parameters without significance tests, see Fig. 6.39(a), the most striking difference is that the seasonal displacement could not be estimated at the center of the center bowl. An explanation for this could be that the (seasonal) displacement does not follow the sinusoid model in that area.

Estimation using the unwrapped phase of the linear model

The unwrapped phase that is obtained using a simpler displacement model could also be used to estimate the seasonal component. This is only possible if the phase residuals are spatially correlated, i.e., if the unmodeled displacement is spatially correlated. Furthermore, the phase residuals must be small enough to estimate the correct coefficients of the simpler model using wrapped data. In this case, the unwrapped phase obtained after the estimation using a linear displacement model (see section 6.2.2), is used to estimate the coefficients of the base functions for seasonal displacement. Thus, the estimation procedure is the same as before, only the unwrapped phase is obtained in a different manner. The estimated parameters are not shown here, because they are practically identical to the displacements obtained using the seasonal base functions already with the wrapped data. The reason that the results are practically identical is that the phase could be unwrapped correctly after demodulation for the estimated DEM error and linear component.

6.2.5 ERS–ENVISAT cross interferometry

For the Las Vegas test site, nine ENVISAT acquisitions are available, see Appendix C. ENVISAT swath IS2 data with a comparable looking angle as ERS are used. In this section it is demonstrated that this ENVISAT data can be used on the PS points by generating cross interferograms with the same ERS master that is used before. The ENVISAT sensor does not exactly

use the same carrier frequency and sampling rates as the ERS satellites, see Table 6.6. These slight differences cause three problems when data of these

Table 6.6: Sensor parameters for ERS and ENVISAT. Derived quantities are given in parenthesis. The ground-range is computed using a look angle $\theta = 21°$ (i.e., swath IS2 of ENVISAT).

Sensor	Radar frequency	(Wavelength)	Pixel spacing range	(ground-range)	azimuth
ERS	5.300 GHz	(5.656 cm)	7.90 m	(22.04 m)	4.00 m
ENVISAT	5.331 GHz	(5.624 cm)	7.80 m	(21.77 m)	4.05 m

sensors are combined using cross interferometry:

- The coregistration is more difficult because the pixel spacing is not the same for both images. This is solved by interpolating the ENVISAT data before coregistration to a grid with the same pixel spacing as that of the oversampled ERS image.
- Distributed targets appear decorrelated in zero baseline ERS–ENVISAT cross interferograms due to the spectral shift property (assuming a zero terrain slope), see (Prati and Rocca, 1994). This spectral shift is compensated by a perpendicular baseline of ~2000 m, although note that such a baseline cannot be maintained during a full orbit revolution. The first cross interferogram showing coherence at distributed scatterers—thus experimentally proving this property—was computed at the DLR using data of the Las Vegas area, see also (Arnaud et al., 2003).
- For point scatterers, a new interferometric phase term as function of range sub-pixel position is induced, caused by the difference in radar frequency, see Eq. (2.17).

A DEM error, linear displacement rate, azimuth and range sub-pixel position are estimated using the wrapped data. Four ERS acquisitions with a Doppler centroid frequency difference larger than 1600 Hz are excluded, leaving 58 acquisitions. Fig. 6.41 shows the estimated variance components using these data. Clearly, the ENVISAT images can be used together with the ERS data to create a long time series of displacement information. The estimated precision of one ERS and two ENVISAT acquisitions is above a threshold (orbits 8226, 10931 and 5921). The variance components for these data are not shown. Most likely, a coregistration problem occurred for these images. However, the variance component for the seven remaining ENVISAT images are similar in magnitude as those of the ERS images, which implies that the phase data of ENVISAT is of good quality.

Fig. 6.42 shows a few typical cross interferometric phase time series at arcs of the reference network. Atmospheric signal can be neglected for these estimates, since the distance between the points in the reference network is ~900 m (except the average interferometric atmospheric phase which appears

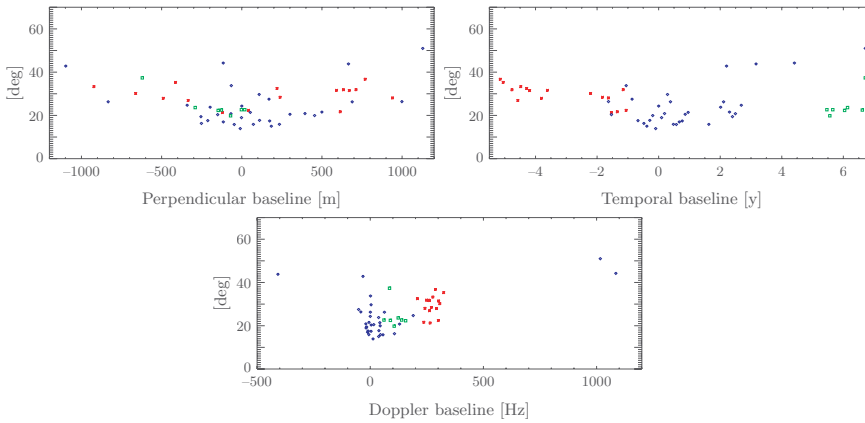


Fig. 6.41: Square roots of the estimated variance components for the Las Vegas test site as function of perpendicular, temporal, and Doppler baseline. A red asterisk corresponds to an ERS–1, a blue diamond corresponds to an ERS–2, and a green square to an ENVISAT acquisition.

as a bias of the residual phase). Each row shows a phase difference time series between two points. From the left to the right, the phase differences are corrected for an additional component. The left column shows the phase differences corrected for the estimated DEM error. The second column corrects the phase also for the estimated linear displacement. The difference between the second and third column is the correction for the estimated azimuth sub-pixel position. Note that this has the largest effect in the interferogram with the last ERS–2 acquisition as slave image, which has a Doppler centroid frequency difference of ~1000 Hz. The effect in (all) the interferograms with ERS–1 slaves is approximately 25% of that, due to the smaller Doppler centroid frequency differences, see also Fig. 6.22. Finally the range sub-pixel position is corrected for, i.e., the last column shows the residual phase. The range sub-pixel position only affects the ENVISAT acquisitions due to the difference in wavelength of ERS and ENVISAT.

6.3 Conclusions

The STUN algorithm is successfully applied at two test sites using real data of ERS–1 and ERS–2. For the Berlin test site a linear model is used to estimate displacement over a time period from 1992 to 2000. A bowl-shaped uplift area with a diameter of approximately 4 km and a maximum displacement rate of ~4 mm/y is identified to the west of the city of Berlin, as well as some individual points with apparent displacement. Using 50 images, the standard deviation of the displacement rate is estimated to be about 0.3 mm/y between points less than one kilometer apart, and to be below 0.9 mm/y for all

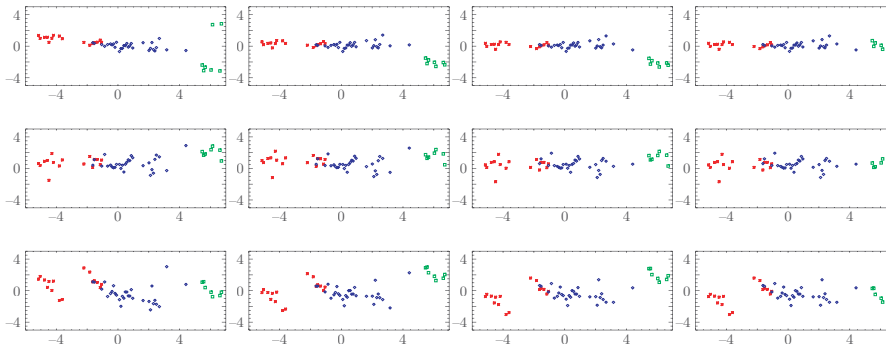


Fig. 6.42: Time series for estimation including ENVISAT for the Las Vegas test site. Plotted are interferometric phase difference data between three pairs (rows) of nearby points. The most left panel shows the data corrected for the estimated DEM error. The next panel additionally corrects for linear displacement rate. The third panel corrects for the azimuth sub-pixel position and the last panel for range sub-pixel position. The estimate shown in the first row has an a posteriori variance factor $\underline{\hat{\sigma}}^2 = 0.3$; second row $\underline{\hat{\sigma}}^2 = 1.1$; third row $\underline{\hat{\sigma}}^2 = 2.1$. A red asterisk corresponds to an ERS–1, a blue diamond to an ERS–2, and a green square to an ENVISAT acquisition.

points (with respect to the reference point). Different settings for the STUN algorithm are described and experimented with. The estimated parameters are influenced mainly by the amount of available acquisitions, although this dependency is related to the actual signal and the bounds on the search space. Results are relatively insensitive to the number of points and arcs in the reference network. This is likely related to limited atmospheric signal for the Berlin area and small displacement rates. The choice of the testing parameters also has only a small effect on the finally estimated parameters. A cross-comparison using data from two adjacent tracks confirms the presence of the uplift area and the validity of the estimated precision.

For the Las Vegas test site, approximately 50 images are available, acquired between 1992 and 2000. Using these data, experiments are performed regarding the choice of estimated parameters. First, the displacement is modeled using a linear rate. Three (known) subsidence bowls and various uplift areas are identified. The maximum estimated displacement is ~20 mm/y (Northern bowl). It is demonstrated for this test site that estimation of an average atmospheric interferometric phase significantly reduced the phase residuals, but not the estimated displacement rates. Azimuth sub-pixel position could not be estimated reliably, due to the small variation of the Doppler centroid frequencies of the Las Vegas data that are used. Seasonal displacement is modeled using trigonometric base functions. The seasonal displacement mainly occurred in the area of the main subsidence bowls. The maximum

amplitude of this seasonal term is ~10 mm. The precision of the estimated displacement is estimated to be ~0.2 mm/y for the linear component, and ~1 mm for the amplitude of the seasonal component (standard deviation of the difference between point less than ~1 km apart). The use of significance tests is demonstrated using the unwrapped data. The estimated parameters using the STUN algorithm compared well with results using the reference PS technique that is performed by TRE. Finally it is shown that ENVISAT data can be used to extend the data stack by performing cross interferometry using the same ERS master image.

7

Conclusions and Recommendations

The central research question is defined in Chapter 1 as:

"How can geodetic methodology aid displacement parameter estimation using Persistent Scatterer Interferometry?"

In section 7.1 the conclusions of the research described in this work are given. Section 7.2 presents recommendations for further research that are outside the scope of this study.

7.1 Conclusions

The conclusion is provided in this section using the items that already are identified in section 1.1:

Functional model. The functional relationship between the observed, wrapped, double-difference phase values and unknown parameters is derived and written in matrix notation using the model of observation equations, as commonly used in geodesy. The model contains parameters for a DEM error, displacement, average atmospheric phase, azimuth and range sub-pixel position. The displacement is modeled as a function of temporal baseline in a generic fashion using base functions. This allows for a broad range of applications. Using synthetic data, an algebraic polynomial is used to model the displacement. For the Berlin test site the displacement is modeled using a linear rate, and for the Las Vegas test site a combination of a linear rate and a seasonal model is successfully applied. The unknown DEM error is practically a linear function of the perpendicular baseline, while the azimuth sub-pixel position is a linear function of Doppler centroid frequency difference. The range sub-pixel position is mainly a linear function of the difference in radar frequency, i.e., this parameter is only of importance when data of sensors with different frequencies are used.

Stochastic model. The stochastic model is formulated as a variance component model. This model describes the precision of the double-difference phase observations and accounts for random noise and varying atmospheric conditions during the radar acquisitions. The components are estimated using a variance component estimation. Experiments with synthetic data demonstrate that this technique yields good estimates of the actual precision of the data (within a few degrees of the simulated noise level). Aside from a realistic data weighting during parameter estimation, an additional advantage of the variance component estimation is that incorrectly processed images are automatically detected. Using this stochastic model during parameter estimation hardly costs additional time, because it remains unchanged during all computations (except for a scaling factor).

Estimator. The integer least-squares estimator is adapted to the problem of PS interferometry and is used for the estimation with wrapped data. Pseudo-observations must be used to limit the search space, because the problem of finding the correct ambiguities is inherently under-determined. Each estimated parameter is constrained in this way using a "soft" bound for its expected value. A bootstrap estimator is initially used to set the bounds for the search space used by the integer least-squares estimator. This algorithm performs well using simulated data. The success rate depends on the number of images, the amount of displacement, the correctness of the displacement model, the amount of noise on the data, and the search bounds. Since the integer least-squares estimator searches the space of the unknown integer ambiguities, the processing time is mainly a function of the number of acquisitions, and not of the number of estimated parameters.

The same functional model is used for all points during the estimation with the wrapped data. A Minimal Cost Flow spatially sparse grid phase unwrapping is applied after the parameters are estimated using the wrapped data. The unwrapped phase, with respect to a single reference point in all interferograms, is used during a final estimation step. Significance tests can be performed using the unwrapped data. Moreover, additional parameters can be introduced once the data are unwrapped, for example accounting for orbit errors as function of spatial coordinates. Experiments with real and synthetic data show that using a simple (linear) displacement model during the estimation using wrapped data is often better as it avoids unrealistic parameter estimates, while the relative displacement between nearby points can often be well approximated using a linear model.

Precision. The precision of the estimated parameters follows by propagation of the estimated variance component stochastic model. A final a posteriori variance factor is estimated for each point using the unwrapped data. The variance-covariance matrix that is obtained in this way describes the precision of, and the correlation between, the estimated parameters. Since all unwrapped phase values are relative to a reference point, the precision

is also described with respect to this point. The precision deteriorates the further the distance of a point from the reference point. This is due to atmospheric signal, which is accounted for in the stochastic model. The precision of the observed phase in the acquired images is estimated to be between ∼15°–40°. The formal relative precision of estimated DEM error and linear displacement rate between nearby points is typically ∼0.3 m and ∼0.2 mm/y (standard deviation) for the test sites Berlin and Las Vegas, using approximately 50 ERS–1 and ERS–2 acquisitions over a nine year time period. For points approximately 25 km apart, these standard deviations are a factor three and five worse for Berlin and Las Vegas, respectively. The correlation between these two estimated parameters is small for these data stacks. Azimuth sub-pixel position could not be estimated because of two reasons. First, the precision would be too small (compared to the azimuth resolution), due to small variation of the Doppler centroid frequency. Second, the correlation with the displacement is largely caused by a lack of overlap between the available ERS–1 and ERS–2 data, which have a systematically different Doppler centroid frequency.

Reliability. The Spatio-Temporal Unwrapping Network (STUN) algorithm is a robust method for three-dimensional phase unwrapping. Due to numerical constraints, estimations are first performed between selected points of a reference network. The parameters are then obtained at the points by a least-squares adjustment and testing procedure. During experiments with synthetic data, all simulated incoherent points are detected and removed. Also during experiments with real data all significant misclosures could be handled by removing points and arcs. The main parameter that affects the outcome of the STUN algorithm is the number of available interferograms, although displacement for the Berlin test site could be estimated using ten interferograms, albeit with a smaller precision. The sensitivity to the number of points and arcs in the reference network is small for the cases presented in this study. Furthermore, the STUN algorithm is shown to be insensitive to the choice of the testing parameters during the alternative hypothesis testing.

7.2 Recommendations

The approach presented in this study concentrates on the use of individual Persistent Scatterer points in radar interferometric data stacks. In this section a few topics for further study are outlined to improve and extend the developed algorithm.

- The signal model that is used in this study describes the radar backscatter as caused by a dominant scatterer in the resolution cell surrounded by incoherent background clutter. Moreover, it is assumed that each estimated

point is visible during the entire time span of the observations. Both assumptions can be relaxed. There could be more than one dominant point in the resolution cell, and it could be that points, for example, on top of new buildings, become "persistent" after a certain time. Relaxation of the first assumption leads to the application of tomography, i.e., the estimation of the position of multiple scatterers based on observations with a small variation of the viewing angle. Concerning the second assumption, recently the concepts of semi-PS and temporary-PS are introduced in the reference PS technique, see (Basilico et al., 2004). These are PS points that are only visible in a subset of the interferometric stack. In the reference PS technique, such points are now identified based on sudden changes of the amplitude, using the (reasonable) assumption that the phase stability is directly related to the amplitude stability.

It should be studied how these issues are best dealt with in the STUN algorithm. For example, a different weight for each point in each acquisition could be introduced. However, this increases the numerical complexity and processing time considerably. The recent concept of Integer Aperture Estimation (see, e.g., Teunissen, 2003a,b, 2004) can possibly be used to identify points that are not coherent during certain acquisitions. This class of estimators provides an overall approach of integer estimation and validation. Each estimated ambiguity can be integer or non-integer, though it is known that this parameter is integer valued. This choice can be made based on the distance of the float solution to the closest integer, i.e., if the integer solution fits badly with the model, the float solution can be used instead, or the observation can be ignored. Moreover, the fail rate can be controlled using this estimator, i.e., the user can set a limit to the amount of incorrectly fixed ambiguities.

- In the near future, a new class of 1–3 m high-resolution spaceborne radar sensors will be launched, such as RADARSAT–2, TerraSAR–X, and COSMO–SkyMed. To demonstrate the level of detail that will be visible in such imagery, Fig. 7.1 shows a high-resolution radar acquisition. Some interesting aspects can be derived from this image. First, it seems that there must be many scatterers in the resolution cell of current day sensors. Such high-resolution images could be used to study the physical properties of PS points and the interaction of the radar signal with the object, something not yet fully understood. Second, using high-resolution data, each object may contain many PS points. This could allow for the observation of stress increase in buildings, which could be used for civil protection. However, suitable algorithms need to be developed to deal with the large amount of data. The point selection used by the STUN algorithm may prove to be a good way to achieve a considerable data reduction, while not losing information. Finally, a recursive estimation scheme needs to be developed to enable updating a current solution with newly acquired data.
- The phase caused by atmospheric heterogeneities is treated as a stochastic signal, and the estimation of displacement parameters is focused upon.

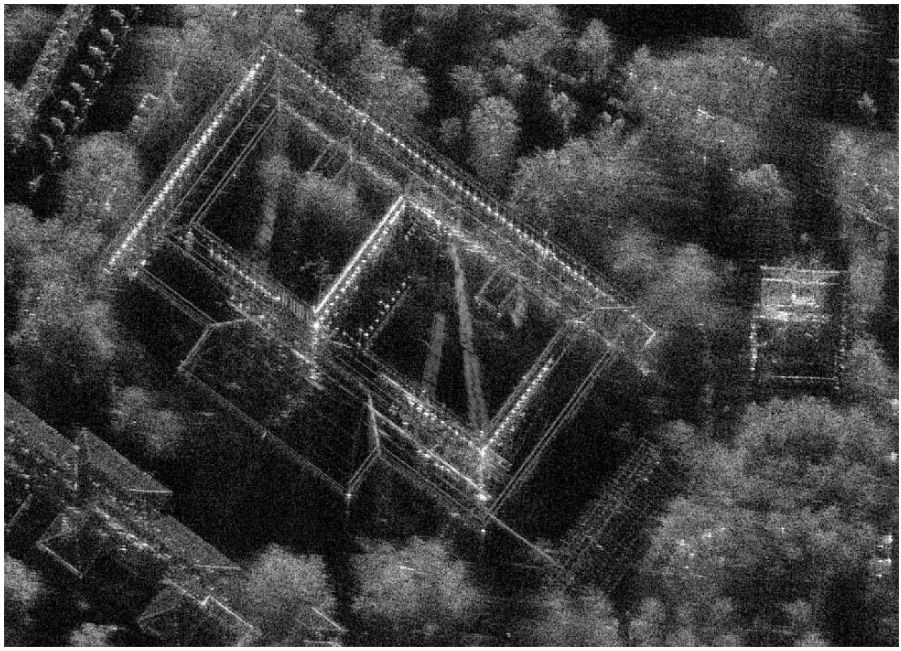


Fig. 7.1: (Image courtesy of A. Brenner.) High-resolution radar image of the campus of Karlsruhe University, Germany. The SAR data were acquired by the X-band, airborne, PAMIR sensor, in August, 2002, and processed to 20 cm resolution, see also (Brenner and Ender, 2004; Soergel et al., 2004).

However, atmospheric signal is the most limiting error source, preventing interpretation of the individual interferometric phase observations at the PS points at the millimeter level. Therefore, it is important to further investigate whether this signal can be estimated and corrected for, which would yield highly precise time series of displacement at individual points. One possible approach is to apply temporal and spatial filters to the unwrapped phase, similar to those used by the reference PS technique. As described, this approach does not work optimally if the residuals contain displacement signal. Since all measurements are relative to each other, a solution to analyze a small area is to choose a nearby stable reference point, ignoring atmospheric signal. An alternative approach could be developed using the unwrapped phase in the data stack, for example at the points of the reference network. Or, for a particular application, only points that seemingly do not undergo displacement could be used to estimate the atmospheric phase. Particularly with large data sets in urbanized areas, it is likely that a set of points can be identified that remain stable. At these points there is no need for a temporal filter, which is of particular importance for near real-time monitoring applications when new acquisitions are continuously added to the interferometric stack.

Addenda

A

Variance Component Estimation

This appendix gives the proof for Eq. (4.26) on page 64, following (Verhoef, 1997). The expectation and covariance of the quadratic form of normally distributed observables are required for this proof, which are derived in section A.1. The proof, given in section A.2, is based on the concept of $\underline{y}^R$-variates. The definition of $\underline{y}^R$-variates is that they are either functionally or stochastically related to another set of observables $\underline{y}$ (Teunissen, 2000a).

A.1 The quadratic form of normally distributed observables

The mathematical model for a set of m linear observation equations can be written as

$$\underline{y} = Bb + \underline{e}; \quad E\{\underline{e}\} = 0; \quad D\{\underline{y}\} = Q_y. \tag{A.1}$$

where $\underline{y}$ is an $m\times1$ vector of observations, B is the design matrix with dimension $m\times n$, b are the unknown parameters in a vector of size $n\times1$, $\underline{e}$ is the $m\times1$ vector of stochastic errors, and Q_y is the $m\times m$ vc-matrix of the observations. The expectation and dispersion of the quadratic form $\underline{y}^*M\underline{y}$, and the covariance between two quadratic forms

$$E\{\underline{y}^*M\underline{y}\}; \quad D\{\underline{y}^*M\underline{y}\}; \quad C\{\underline{y}^*M\underline{y}, \underline{y}^*N\underline{y}\}; \tag{A.2}$$

are now derived (M and N are symmetric matrices with dimension $m\times m$). The trace operator[1] is used for this derivation, as well as the following of its properties

[1] $\mathrm{trace}(A) = \sum_i A_{ii}$

$$\begin{aligned}
\operatorname{trace}(A) &= \operatorname{trace}(A^*) \\
\operatorname{trace}(A+B) &= \operatorname{trace}(A) + \operatorname{trace}(B) \\
\operatorname{trace}(CD) &= \operatorname{trace}(DC) \\
\operatorname{trace}(Abb^*) &= b^*Ab \\
E\{\operatorname{trace}(A)\} &= \operatorname{trace}(E\{A\}),
\end{aligned} \tag{A.3}$$

where A and B are square matrices with dimension $m \times m$, C is $m \times n$, D is $n \times m$, and b is $m \times 1$. Using these properties, together with $E\{\underline{e}\} = 0$ and $D\{\underline{e}\} = E\{\underline{e}^*\underline{e}\} = Q_y$, the expectation of a quadratic form is

$$\begin{aligned}
E\{\underline{y}^*M\underline{y}\} &= E\{(Bb+\underline{e})^*M(Bb+\underline{e})\} \\
&= b^*B^*MBb + 2b^*B^*ME\{\underline{e}\} + E\{\underline{e}^*M\underline{e}\} \\
&= b^*B^*MBb + E\{\operatorname{trace}(\underline{e}^*M\underline{e})\} \\
&= b^*B^*MBb + E\{\operatorname{trace}(M\underline{e}^*\underline{e})\} \\
&= b^*B^*MBb + \operatorname{trace}(ME\{\underline{e}^*\underline{e}\}) \\
&= b^*B^*MBb + \operatorname{trace}(MQ_y).
\end{aligned} \tag{A.4}$$

The first to fourth moment of the probability density function of the observables is used to derive the dispersion of the quadratic form. A normal distribution is assumed here. The covariance between two quadratic forms of normally distributed observables is given as

$$C\{\underline{y}^*M\underline{y}, \underline{y}^*N\underline{y}\} = E\{(\underline{y}^*M\underline{y} - E\{\underline{y}^*M\underline{y}\})\,(\underline{y}^*N\underline{y} - E\{\underline{y}^*N\underline{y}\})^*\}. \tag{A.5}$$

With Eq. (A.4) this becomes

$$C\{\underline{y}^*M\underline{y}, \underline{y}^*N\underline{y}\} = E\{(\underline{y}^*M\underline{y} - E\{\underline{y}^*M\underline{y}\})\,(\underline{y}^*N\underline{y} - E\{\underline{y}^*N\underline{y}\})^*\}, \tag{A.6}$$

which can be further expanded to

$$\begin{aligned}
C\{\underline{y}^*M\underline{y}, \underline{y}^*N\underline{y}\} = {} & 4b^*B^*ME\{\underline{e}\underline{e}^*\}NBb \\
& + 2b^*B^*ME\{\underline{e}\underline{e}^*N\underline{e}\} \\
& - 2b^*B^*ME\{\underline{e}\}\operatorname{trace}(NQ_y) \\
& + 2E\{\underline{e}M\underline{e}\underline{e}^*\}NBb \\
& + E\{\underline{e}^*M\underline{e}\underline{e}N\underline{e}^*\} \\
& - E\{\underline{e}^*M\underline{e}\}\operatorname{trace}(NQ_y) \\
& - 2\operatorname{trace}(MQ_y)E\{\underline{e}\}NBb \\
& - \operatorname{trace}(MQ_y)E\{\underline{e}N\underline{e}^*\} \\
& + \operatorname{trace}(MQ_y)\operatorname{trace}(NQ_y).
\end{aligned} \tag{A.7}$$

From the assumption of a normal distribution it follows that the third central moment is zero. The fourth central moment is known to be (Koch, 1988)

$$E\{\underline{e}^*M\underline{e}\,\underline{e}^*N\underline{e}\} = \operatorname{trace}(MQ_y)\operatorname{trace}(NQ_y) + 2\operatorname{trace}(MQ_yNQ_y). \tag{A.8}$$

Using this information, Eq. (A.7) reduces to

$$\begin{aligned} C\{\underline{y}^*M\underline{y}, \underline{y}^*N\underline{y}\} = {} & 4b^*B^*ME\{\underline{e}\underline{e}^*\}NBb \\ & + \operatorname{trace}(MQ_y)\operatorname{trace}(NQ_y) \\ & + 2\operatorname{trace}(MQ_y\,NQ_y) \\ & - E\{\underline{e}^*M\underline{e}\}\operatorname{trace}(NQ_y) \\ & - \operatorname{trace}(MQ_y)E\{\underline{e}N\underline{e}^*\} \\ & + \operatorname{trace}(MQ_y)\operatorname{trace}(NQ_y). \end{aligned} \tag{A.9}$$

Finally, using $E\{\underline{e}^*\underline{e}\}=Q_y$ and $E\{\underline{e}^*M\underline{e}\}=\operatorname{trace}(MQ_y)$, the covariance between two quadratic forms of normally distributed observables is derived as

$$C\{\underline{y}^*M\underline{y}, \underline{y}^*N\underline{y}\} = 4b^*B^*MQ_yNBb + 2\operatorname{trace}(MQ_y\,NQ_y). \tag{A.10}$$

The dispersion (variance) of a quadratic form of normally distributed observables is analogously derived as

$$D\{\underline{y}^*M\underline{y}\} = 4b^*B^*MQ_yMBb + 2\operatorname{trace}(MQ_y\,MQ_y). \tag{A.11}$$

A.2 Proof of the variance component estimation formula

The $m\times1$ vector of stochastic errors of the model defined in Eq. (A.1) is now written as a linear combination of p groups of elementary errors as

$$\underline{e} = \sum_{k=1}^{p} U_k\,\underline{\epsilon}_k, \tag{A.12}$$

where

$\underline{e}$ is a $m\times1$ vector of stochastic errors.
U_k is a $m\times c_k$ transformation matrix describing the influence of the k$^{\text{th}}$ group of errors on the observations.
$\underline{\epsilon}_k$ is a $c_k\times1$ vector of stochastic errors of group k.

It is assumed that the groups of errors are not correlated and that the errors within a group have equal variance and are uncorrelated, i.e.,

$$\begin{aligned} E\{\underline{\epsilon}_k\} &= 0 \quad \forall\, k = 1, \cdots, p; \\ C\{\underline{\epsilon}_k, \underline{\epsilon}_l\} &= 0 \quad \forall\, k = 1, \cdots, p,\ l \neq k; \\ D\{\underline{\epsilon}_k\} &= Q_{\epsilon_k} = \sigma_k^2\, I_{c_k} \quad \forall\, k = 1, \cdots, p. \end{aligned} \tag{A.13}$$

Using the propagation law of variances and covariances it follows for the vc-matrix of the observations and the covariance matrix of the k$^{\text{th}}$ group of errors with the observations that

$$Q_y = \sum_{k=1}^{p} \sigma_k^2 \, U_k U_k{}^* = \sum_{k=1}^{p} \sigma_k^2 \, Q_k, \tag{A.14}$$

$$Q_{\epsilon_k,y} = \sigma_k^2 \, U_k{}^*. \tag{A.15}$$

The p unknown vectors of errors $\underline{\epsilon}_k$ can be regarded as constituent $\underline{y}^R$-variates of the resulting vector of stochastic errors $\underline{e}$. The general expression for the $\underline{y}^R$-correction is given as (Teunissen, 2000a)

$$\underline{\hat{e}}^R = Q_{y^R,y} Q_y^{-1} \, \underline{\hat{e}}, \tag{A.16}$$

where $\underline{\hat{e}}^R$ and $\underline{\hat{e}}$ are the vectors of least-squares residuals and $\underline{y}^R$-variates, respectively, and $Q_{y^R,y}$ is the matrix of covariance between the $\underline{y}$-variates and the $\underline{y}^R$-variates. According to the theory of $\underline{y}^R$-variates, the least-squares estimator for the k$^{\text{th}}$ group of errors $\underline{\epsilon}_k$ can thus be computed from the least-squares vector of corrections $\underline{\hat{e}} = P_B^{\perp} \underline{y}$, with $P_B^{\perp} = I - B(B^* Q_y^{-1} B)^{-1} B^* Q_y^{-1}$, as

$$\begin{aligned} \underline{\hat{\epsilon}}_k &= Q_{\epsilon_k,y} Q_y^{-1} \, \underline{\hat{e}} \\ &= \sigma_k^2 U_k{}^* Q_y^{-1} P_B^{\perp} \, \underline{y}. \end{aligned} \tag{A.17}$$

The shifting variate of the k$^{\text{th}}$ group of errors is defined as

$$\begin{aligned} \underline{\hat{\epsilon}}_k{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_k &= \underline{y}^* P_B^{\perp *} Q_y^{-1} U_k \sigma_k^2 (\sigma_k^2 I_{c_k})^{-1} \sigma_k^2 \, U_k{}^* Q_y^{-1} P_B^{\perp} \, \underline{y} \\ &= \underline{y}^* Q_y^{-1} P_B^{\perp} \sigma_k^2 Q_k Q_y^{-1} P_B^{\perp} \, \underline{y}. \end{aligned} \tag{A.18}$$

With Eq. (A.4) the expectation of this quadratic form can be shown to be

$$E\{\underline{\hat{\epsilon}}_k{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_k\} = \operatorname{trace}(Q_y^{-1} P_B^{\perp} \sigma_k^2 Q_k Q_y^{-1} P_B^{\perp} Q_y). \tag{A.19}$$

The covariance between the shifting variates of the k$^{\text{th}}$ and l$^{\text{th}}$ group of errors follows from Eq. (A.10) as

$$\begin{aligned} C\{\underline{\hat{\epsilon}}_k{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_k, \underline{\hat{\epsilon}}_l{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_l\} &= 2 \operatorname{trace}(Q_y^{-1} P_B^{\perp} \sigma_k^2 Q_k Q_y^{-1} P_B^{\perp} Q_y Q_y^{-1} P_B^{\perp} \sigma_l^2 Q_l Q_y^{-1} P_B^{\perp} Q_y), \\ &= 2 \operatorname{trace}(Q_y^{-1} P_B^{\perp} \sigma_k^2 Q_k Q_y^{-1} P_B^{\perp} \sigma_l^2 Q_l). \end{aligned} \tag{A.20}$$

In this case, the following relation is valid

$$\begin{aligned} \sum_{l=1}^{p} \frac{1}{2} C\{\underline{\hat{\epsilon}}_k{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_k, \underline{\hat{\epsilon}}_l{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_l\} &= \operatorname{trace}(Q_y^{-1} P_B^{\perp} \sigma_k^2 Q_k Q_y^{-1} P_B^{\perp} \sum_{l=1}^{p} \sigma_l^2 Q_l, \\ &= \operatorname{trace}(Q_y^{-1} P_B^{\perp} \sigma_k^2 Q_k Q_y^{-1} P_B^{\perp} Q_y), \\ &= E\{\underline{\hat{\epsilon}}_k{}^* Q_{\epsilon_k}^{-1} \underline{\hat{\epsilon}}_k\}. \end{aligned} \tag{A.21}$$

Therefore, by interchanging the equations on the left and right side and canceling σ_k^2, a linear system of observation equations with p unknown variance components arises as

$$E\{\underline{y}^* Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} \underline{y}\} = \sum_{l=1}^{p} \operatorname{trace}(Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} Q_l)\, \sigma_l^2 \quad \forall\, k = 1, \cdots, p. \tag{A.22}$$

Symbolically this system of equations can be written as

$$E\{\underline{r}\} = N\sigma, \tag{A.23}$$

where

$$\underline{r} = \begin{bmatrix} \underline{r}_1 \\ \vdots \\ \underline{r}_l \\ \vdots \\ \underline{r}_k \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & \dots & N_{1l} & \dots & N_{1p} \\ \vdots & & \vdots & & \vdots \\ N_{k1} & \dots & N_{kl} & \dots & N_{kp} \\ \vdots & & \vdots & & \vdots \\ N_{p1} & \dots & N_{pl} & \dots & N_{pp} \end{bmatrix}, \quad \sigma = \begin{bmatrix} \sigma_1^2 \\ \vdots \\ \sigma_l^2 \\ \vdots \\ \sigma_k^2 \end{bmatrix}, \tag{A.24}$$

and

$$\begin{aligned} \underline{r}(k{+}1) &= \hat{\underline{e}}^* Q_y^{-1} Q_k Q_y^{-1} \hat{\underline{e}}, \\ N(k{+}1, l{+}1) &= \operatorname{trace}(Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} Q_l). \end{aligned} \tag{A.25}$$

If the $p \times p$ matrix N is regular, then an unbiased estimator for σ is given by

$$\hat{\underline{\sigma}} = N^{-1}\, \underline{r}, \tag{A.26}$$

which is identical to Eq. (4.8).

Variance of the estimator for the components

Using Eq. (A.10), it follows that

$$\begin{aligned} C\{\underline{r}_k, \underline{r}_l\} &= C\{\underline{y}^* Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} \underline{y}, \underline{y}^* Q_y^{-1} P_B^{\perp} Q_l Q_y^{-1} P_B^{\perp} \underline{y}\} \\ &= 2 \operatorname{trace}(Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} Q_y^{-1} P_B^{\perp} Q_l Q_y^{-1} P_B^{\perp} Q_y), \\ &= 2 \operatorname{trace}(Q_y^{-1} P_B^{\perp} Q_k Q_y^{-1} P_B^{\perp} Q_l), \\ &= 2N. \end{aligned} \tag{A.27}$$

Thus,

$$D\{\underline{r}\} = 2N. \tag{A.28}$$

The vc-matrix of the estimated components follows by application of the propagation law of variances as

$$D\{\hat{\underline{\sigma}}\} = N^{-1} D\{\underline{r}\} N^{-1} = 2N^{-1}. \tag{A.29}$$

B

Alternative Hypothesis Testing

During hypothesis testing, the probability of correctness of the null-hypothesis is tested against one or more alternative hypotheses. This appendix deals with the question when to reject the null-hypothesis, and if so, in favor of which alternative hypothesis. First, in section B.1 a brief overview is given of the testing parameters and the types of errors that can be made during alternative hypothesis testing. Then, section B.2 describes the consequences of the choice of a specific value for the power of the tests. This appendix is based on (De Heus et al., 1994; Teunissen, 2000b; Verhoef, 1997).

B.1 The Delft method of testing

Using a linear model of observation equations, the null-hypothesis is given as

$$H_0 : E\{\underline{y}\} = Bb; \qquad D\{\underline{y}\} = Q_y, \tag{B.1}$$

where $\underline{y}$ is a $m\times1$ vector of observations, B is the design matrix with dimension $m\times n$, $\underline{b}$ is the $n\times1$ vector of unknown parameters, $\underline{e}$ is the $m\times1$ vector of measurement noise, and Q_y is the $m\times m$ vc-matrix of the observations. An alternative hypothesis is specified as a linear extension of the null-hypothesis, i.e.,

$$H_A : E\{\underline{y}\} = Bb + C_q\nabla; \qquad D\{\underline{y}\} = Q_y. \tag{B.2}$$

Matrix C_q defines the type of the considered model mis-specification, and ∇ is the vector of additional parameters. The dimension of the alternative hypothesis is q, i.e., C_q is $m\times q$ and ∇ is $q\times1$. Here, the stochastic model, i.e., matrix Q_y, is assumed to be the same under both hypotheses. Furthermore, it is assumed that the precision of the observations is correctly described by Q_y. Alternative hypothesis could be specified to test (De Heus et al., 1994):

- An individual observation. $C_{q=1}$ is a zero vector with a single one corresponding to the observation that is tested, see Eq. (4.19). In chapter 4 this was referred to as the arc test.

- A group of related observations. For example the point test, see Eq. (4.21).
- All observations. This is the overall model test, see Eq. (4.18).
- A specific model deviation. For example, significance of increasing the degree of a polynomial displacement model.

The test statistic $\underline{T}_q$ can be used to decide whether the alternative hypothesis is a significant extension of the null-hypothesis. It is given as (Teunissen, 2000b)

$$\underline{T}_q = \underline{\hat{e}}^* Q_y^{-1} C_q (C_q{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_q)^{-1} C_q{}^* Q_y^{-1} \underline{\hat{e}}, \tag{B.3}$$

where $\underline{\hat{e}}$ is the vector of least-squares residuals under the null-hypothesis. This test statistic has a chi-squared distribution with q degrees of freedom

$$H_0 : \underline{T}_q \sim \chi^2(q, 0), \tag{B.4}$$

$$H_A : \underline{T}_q \sim \chi^2(q, \lambda), \tag{B.5}$$

where λ, the non-centrality parameter of the chi-squared distribution, is given by

$$\lambda = \nabla^* C_q{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_q \nabla. \tag{B.6}$$

The null-hypothesis is rejected when this test statistic is larger than some critical value $\chi^2_{\alpha(q)}$, which depends on the dimension of the test q and the chosen level of significance α, see also Fig. B.1. From this figure it is clear that two types of errors can be made when performing hypothesis testing, namely:

- Type-I error. The probability of incorrect rejection of the null-hypothesis in favor of the alternative hypothesis is given by α. This probability is referred to as the size or level of significance of the test, and is also known as the false alarm rate.
- Type-II error. It is also possible that the null-hypothesis is not rejected while it should be, which can be considered the neglected alarm rate. The probability of this occurrence is β. In geodesy, instead of β, normally the power $\gamma = 1-\beta$ is used, i.e., the probability that the alternative hypothesis is accepted when indeed it was correct.

Aside from these two errors, also a third kind of error can be made, particularly when more than one alternative hypothesis is considered. This error is the selection of the wrong alternative hypothesis if the null-hypothesis is rejected, which is sometimes referred to as a type-III error (Verhoef, 1997). The chance on this error cannot be expressed in terms of α and γ, but it can be easily understood that this probability increases when the alternative hypotheses are more alike. Section B.2 deals with this error in more detail, particularly regarding the comparison of tests of different dimensions.

Ideally, α and β should be as small as possible. However, if α is chosen smaller, then β increases, and vice versa, see also Fig. B.1. In order to define a test that best deals with this paradox, the Neyman-Pearson principle is applied (Neyman and Pearson, 1933). This principle states that from all

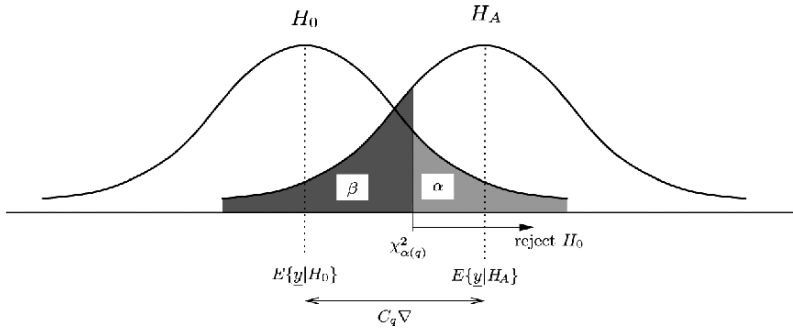


Fig. B.1: Relation between testing parameters. Sketched is the null-hypothesis H_0 and an alternative hypothesis H_A. The null-hypothesis is rejected if the test statistic is larger than the critical value $\chi^2_{\alpha(q)}$. The probability of rejecting the null-hypothesis while it was correct (false alarm or type-I error) is given by α. The probability that the null-hypothesis is not rejected while the alternative hypothesis was correct is given by the area β (neglected alarm or type-II error). From the distance $C_q\nabla$ between the two hypotheses it is clear that these probabilities decrease if this model error becomes larger.

possible tests possessing the same size type-I error, the test for which the type-II error is as small as possible should be used. The test statistic $\underline{T}_q$, cf. Eq. (B.3), is a consequence of this principle (Teunissen, 2000b).

The B-method of testing

As stated above, the power γ is computed from a non-central chi-squared distribution with q degrees of freedom, and depends on the chosen level of significance α. This functional relation can be symbolized as $\gamma = \gamma(\alpha, q, \lambda)$. Thus, if ∇ is known, the non-centrality parameter λ can be computed by Eq. (B.6), and thereafter the power of a test with level of significance α and dimension q can be computed. However, instead of obtaining the power as a function of a certain model error which is unknown, it is of more interest to know what size of model error can be found with a fixed probability. By fixing $\gamma = \gamma_0$, the non-centrality parameter can be computed from the relation

$$\lambda_0 = \lambda(\alpha, q, \gamma = \gamma_0). \tag{B.7}$$

Then, the model error that can just be detected follows from solving for ∇ in Eq. (B.6). For a one-dimensional test, $C_{q=1}$ reduces to a vector and ∇ to a scalar, and the minimal detectable bias (MDB) is given by

$$|\nabla| = \sqrt{\frac{\lambda_0}{C_{q=1}{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_{q=1}}}. \tag{B.8}$$

Note that only the size of the MDB can be computed, not its sign. To compare the model errors that can be found by different alternative hypotheses, it is

required that all tests have the same power. Otherwise it would be difficult to assess which alternative hypothesis to select. For example, consider the case that a specific hypothesis is able to detect a model error of, say, 2 cm with a probability $\gamma_1 = 60\%$ while another hypothesis can detect another model error of, say, 3 cm with $\gamma_2 = 90\%$. It cannot be decided which of these two alternative hypothesis must be selected. Choosing the same power for all tests is the essence of how tests of different dimensions are related to each other in the B-method of testing, see (Baarda, 1968; Teunissen, 2000b). The non-centrality parameter λ is the connection between the tests, enabling the computation of appropriate testing parameters and the corresponding critical values. First, the power γ_0 for all tests and a value α_1 for the level of significance for the one-dimensional test is fixed, and the corresponding non-centrality parameter λ_0 is computed. Then, the level of significance for a test of dimension q_i is computed using the relation

$$\lambda_0 = \lambda(\alpha = \alpha_1, q = 1, \gamma = \gamma_0) = \lambda(\alpha = \alpha_{q_i}, q = q_i, \gamma = \gamma_0). \qquad \text{(B.9)}$$

Then, the corresponding critical value can be computed from the chi-squared distribution using this level of significance. The use of equal values for the non-centrality parameter $\lambda = \lambda_0$ and power $\gamma = \gamma_0$ in all tests implies that a certain model error can be found with the same probability by all tests.

B.2 Selecting the testing parameters

It is shown by De Heus et al. (1994) that if several alternative hypotheses of different dimensions are specified, then the test with the largest test quotient (the test statistic divided by its critical value) is the most likely alternative hypothesis, at least if a power $\gamma_0 \leq 50\%$ is chosen. In this section the proof for this statement is repeated, and consequences of a choice $\gamma > 50\%$ are derived.

In order to choose the testing parameters for a specific application (alternative hypothesis), the consequences of an erroneous decision, i.e, occurrence of a type-I or type-II error, should be taken into account. For surveying applications, it is common practice to select $\gamma_0 = 80\%$ and $\alpha_1 = 0.001$. If, like in surveying, it can reasonably be expected that maximally 10% of the observations are incorrect, then a choice α_1=0.001 implies that 1% of identified blunder observations actually were correct observations. Thus, this choice prevents unnecessary re-measurements in the field, which in general are expensive. The choice of $\gamma_0 = 80\%$ implies that it can be expected that in 20 percent of the cases the null-hypothesis is incorrectly accepted.

The problem at hand is the detection of incoherent points and incorrect estimated parameters between points, as described in Chapter 4. The alternative hypotheses that are specified are referred to as the point and arc test, respectively. Note that the standard choice of the testing parameters do not have to apply to this particular problem, because re-measurements cannot be

performed. Moreover, compared to conventional leveling networks, the point density is very high, although note that they do not lie at a specified location, but in a "network of opportunity". In order to choose appropriate values for α_1 and γ_0, the question is, which of the following alternatives has a more severe impact for this application:

1. Rejecting H_0 while this was not correct, i.e., removing arcs and points unnecessary.
2. Not rejecting H_0 while it should be done, i.e., leaving arcs and points in the network that should be taken out (even if that means also taking out correct observations).

In our case, item 2 is more important than item 1, thus α_1 can be chosen relatively large, and $\beta_0 = 1-\gamma_0$ relatively small. However, note that a selection of $\gamma_0 > 50\%$ favors acceptance of the more-dimensional alternative hypothesis, which is proven in the following.

Tests of equal dimensions

If the null-hypothesis is only tested against a single alternative hypothesis, rejection of the null-hypothesis leads in general to acceptance of the alternative hypothesis (the stochastic model could also be adapted). However, in practice more than one alternative hypothesis is specified, and the statistically most significant one must be selected. Suppose two alternative hypotheses H_A^i and H_A^j are specified

$$H_A^i : E\{\underline{y}\} = Bb + C_{q_i}\nabla_i, \tag{B.10}$$

$$H_A^j : E\{\underline{y}\} = Bb + C_{q_j}\nabla_j, \tag{B.11}$$

where the dimension of the alternative hypothesis are equal to q_i and q_j, respectively. If these tests are of the same dimension, $q=q_i=q_j$, then the alternative hypothesis H_A^k with the largest test statistic $\underline{T}_q^k$ is the most likely alternative hypothesis. If the level of significance α is equal for both alternative hypotheses, then also the hypothesis with the largest test quotient is the most likely one, i.e., select H_A^k for which

$$\frac{\underline{T}_q^k}{\chi^2_{\alpha(q)}} > \frac{\underline{T}_q^l}{\chi^2_{\alpha(q)}} \quad \forall\, k \neq l. \tag{B.12}$$

This is the case when the B-method of testing is applied.

Tests of unequal dimensions

If the dimensions of the alternative hypotheses are not equal, then the most likely alternative hypothesis does not necessarily have the largest test statistic $\underline{T}_{q_i}^k$, since the probability density function of the test statistics, cf. Eq. (B.5),

are not the same. Therefore, to select the most likely alternative hypothesis, the test quotient is a better criterion, confronting each test statistic with its critical value. Test quotients with a value smaller than one do not have to be considered, since the null-hypothesis is more likely for these alternative hypotheses. For the remaining alternative hypotheses, it is assumed that the test quotient of the most likely alternative hypothesis is most rejected, i.e., that it has the largest test quotient. This assumption is only true when a power $\gamma_0 \leq 50\%$ is used, as proven in (De Heus et al., 1994). The proof is given below.

In the following, it is assumed that H_A^i is the arc test, and H_A^j is the point test. The dimension of the alternative hypothesis H_A^j is thus larger than that of H_A^i. Moreover, the point test is an extension of the arc test, i.e., $C_{q_i} \in C_{q_j}$. Two cases can be distinguished:

1. H_A^i is the correct hypothesis (H_A^j is too relaxed).
   If the testing parameters are chosen according to the B-method of testing, i.e., $\lambda(\alpha_{q_i}, q = q_i, \gamma = \gamma_0) = \lambda(\alpha_{q_j}, q = q_j, \gamma = \gamma_0)$, then both test statistics $\underline{T}_{q_i}^k$ and $\underline{T}_{q_j}^l$ are rejected with the same probability. Furthermore, if the choice is made for $\gamma_0 = 50\%$, then it is expected that when an error occurs of the size of the minimal detectable bias that

$$\frac{\underline{T}_{q_i}^k}{\chi^2_{\alpha(q_i)}} = \frac{\underline{T}_{q_j}^l}{\chi^2_{\alpha(q_j)}} = 1. \tag{B.13}$$

   If a larger error would occur, it would be rejected with a larger probability γ' by H_A^i. Namely, if also H_A^j would reject this error with the same probability, then it should hold that

$$\lambda(\alpha_{q_i}, q = q_i, \gamma = \gamma') = \lambda(\alpha_{q_j}, q = q_j, \gamma = \gamma'). \tag{B.14}$$

   Since $\gamma' > \gamma_0$, this can only be the case if $\alpha_{q_j'} > \alpha_{q_j}$, where $\alpha_{q_j'}$ is the new value for α. Therefore,

$$\chi^2_{\alpha(q_j')} < \chi^2_{\alpha(q_j)}. \tag{B.15}$$

   Since the test quotient is computed with the larger, original, value $\chi^2_{\alpha(q_j)}$, the test quotient for the lower dimensional alternative hypothesis is expected to be the larger than that of the higher dimensional test quotient

$$\frac{\underline{T}_{q_i}^k}{\chi^2_{\alpha(q_i)}} > \frac{\underline{T}_{q_j}^l}{\chi^2_{\alpha(q_j)}} > 1. \tag{B.16}$$

   The choice $\gamma_0 \leq 50\%$ is essential, since the above reasoning could be reversed for occurrence of an error smaller than the minimal detectable bias. This implies that then the test quotient for the higher dimensional test is larger, and an erroneous conclusion would be drawn (the higher dimensional alternative hypothesis would incorrectly be selected as most probable). By restricting $\gamma_0 \leq 50\%$, the test quotient for an error smaller

than the minimal detectable bias is always smaller than 1, which implies that this alternative hypothesis should be rejected in favor of the null-hypothesis.

2. H_A^j is the correct hypothesis (H_A^i is too narrow).
   Since $C_{q_i} \in C_{y_j}$, a model error according to H_A^j implies that

$$\nabla_i{}^* C_{q_i}{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_{q_i} \nabla_i < \nabla_j{}^* C_{q_j}{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_{q_j} \nabla_j. \tag{B.17}$$

   For a just detectable model error ∇_j for which

$$\nabla_j{}^* C_{q_j}{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_{q_j} \nabla_j = \lambda(\alpha_{q_j}, q = q_j, \gamma = 50\%), \tag{B.18}$$

   the test quotients are expected to be

$$\frac{\underline{T}_q^j}{\chi^2_{\alpha(q_j)}} = 1 > \frac{\underline{T}_q^i}{\chi^2_{\alpha(q_i)}}. \tag{B.19}$$

   For an increasing model error, at a certain point the situation is reached that the lower dimensional test is just rejected

$$\begin{aligned} \nabla_i{}^* C_{q_i}{}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C_{q_i} \nabla_i &= \lambda(\alpha_{q_i}, q = q_i, \gamma = 50\%) \\ &= \lambda(\alpha_{q_j}, q = q_j, \gamma = 50\%). \end{aligned} \tag{B.20}$$

   Then, the test quotients are expected to be

$$\frac{\underline{T}_q^j}{\chi^2_{\alpha(q_j)}} > \frac{\underline{T}_q^i}{\chi^2_{\alpha(q_i)}} = 1. \tag{B.21}$$

   If the model error ∇_j increases more, the test quotient of the lower dimensional test increases more than that of the hypothesis H_A^j. It could even happen that the test quotient of H_A^i becomes larger than that of H_A^j, depending on the values of the elements in the vector ∇_j. This means that it could happen that the wrong alternative hypothesis (the lower dimensional) is selected, also when $\gamma = 50\%$, particularly if some elements of the model error are large. The chance that this occurs is reduced if the dimension of H_A^j is only one larger than that of H_A^i. If the significance of extension of a displacement model is tested, e.g., by increasing the degree of an algebraic polynomial, the alternative hypotheses testing procedure thus must be performed in small steps. The first alternative hypothesis should specify an increase of the degree of the displacement model by one, and if the null-hypothesis is rejected, this becomes the new null-hypothesis, and a new alternative hypothesis is specified, again increasing the degree of the displacement model by one.

Thus, for selection of the most likely alternative hypothesis among hypotheses of unequal dimensions, it is mandatory that $\gamma_0 \leq 50\%$. Since a larger power enables detection of smaller errors, in (De Heus et al., 1994) the power is

chosen as $\gamma_0 = 50\%$. This choice implies that the minimum detectable bias of an alternative hypothesis of dimension q can be interpreted as the error that just is rejected/not rejected (i.e., the expected test quotient is one).

However, if a larger value for γ_0 is chosen, more often a higher dimensional alternative hypothesis is selected, even when the lower dimensional alternative hypothesis correctly specifies the model error. For the application of point tests and arc tests this implies that points are removed from the reference network, even when this would not be necessary. This is not a severe drawback, since it is considered more important that the points that remain in the network are correctly computed. In the developed software, the user can select whether to do only arc tests, only point tests, or to do both. Only in the latter case the above is of importance.

C

Used SAR Data

This appendix lists the ERS and ENVISAT data that are relevant to the experiments performed during this study, see Chapter 6.

Table C.1: ERS data for the Berlin area (track 165, frame 2547). Parameters are relative to the master acquisition, orbit 10039, acquired at 22–MAR–1997 10:03 am (UTC). Data are sorted on acquisition time, except for the master image, which is listed first.

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**22–MAR–1997**	**10039**	**ERS–2**	**0**	**194.84**
2	13–MAY–1992	4318	ERS–1	–285	480.10
3	17–JUN–1992	4819	ERS–1	–300	489.83
4	26–AUG–1992	5821	ERS–1	50	430.52
5	30–SEP–1992	6322	ERS–1	445	416.68
6	13–JAN–1993	7825	ERS–1	–85	443.91
7	17–FEB–1993	8326	ERS–1	435	421.66
8	07–JUL–1993	10330	ERS–1	–905	482.27
9	11–AUG–1993	10831	ERS–1	85	464.58
10	20–OCT–1993	11833	ERS–1	770	440.89
11	24–NOV–1993	12334	ERS–1	915	422.19
12	21–APR–1995	19692	ERS–1	–180	445.82
13	26–MAY–1995	20193	ERS–1	–170	489.88
14	30–JUN–1995	20694	ERS–1	–805	487.97
15	04–AUG–1995	21195	ERS–1	325	476.14
16	05–AUG–1995	1522	ERS–2	255	174.54
17	08–SEP–1995	21696	ERS–1	–920	447.90
18	13–OCT–1995	22197	ERS–1	665	440.28
19	14–OCT–1995	2524	ERS–2	960	153.34
20	17–NOV–1995	22698	ERS–1	–575	452.72
21	22–DEC–1995	23199	ERS–1	485	518.89
22	23–DEC–1995	3526	ERS–2	325	237.29

Table C.1: (continued)

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**22–MAR–1997**	**10039**	**ERS–2**	**0**	**194.84**
23	26–JAN–1996	23700	ERS–1	45	379.07
24	01–MAR–1996	24201	ERS–1	605	427.07
25	05–APR–1996	24702	ERS–1	–60	447.99
26	06–APR–1996	5029	ERS–2	–115	204.89
27	10–MAY–1996	25203	ERS–1	540	480.89
28	11–MAY–1996	5530	ERS–2	400	176.62
29	14–JUN–1996	25704	ERS–1	–210	534.36
30	19–JUL–1996	26205	ERS–1	545	491.91
31	20–JUL–1996	6532	ERS–2	395	193.27
32	24–AUG–1996	7033	ERS–2	–690	204.97
33	28–SEP–1996	7534	ERS–2	–130	127.81
34	02–NOV–1996	8035	ERS–2	1090	133.36
35	11–JAN–1997	9037	ERS–2	45	196.91
36	26–APR–1997	10540	ERS–2	–410	221.76
37	31–MAY–1997	11041	ERS–2	–295	244.95
38	05–JUL–1997	11542	ERS–2	–195	216.54
39	13–SEP–1997	12544	ERS–2	260	214.54
40	18–OCT–1997	13045	ERS–2	110	183.88
41	27–DEC–1997	14047	ERS–2	–85	251.12
42	31–JAN–1998	14548	ERS–2	65	164.80
43	07–MAR–1998	15049	ERS–2	–535	174.02
44	11–APR–1998	15550	ERS–2	–120	198.63
45	16–MAY–1998	16051	ERS–2	645	259.04
46	20–JUN–1998	16552	ERS–2	670	261.54
47	25–JUL–1998	17053	ERS–2	–315	231.61
48	29–AUG–1998	17554	ERS–2	155	203.86
49	03–OCT–1998	18055	ERS–2	615	203.90
50	07–NOV–1998	18556	ERS–2	1015	181.29
51	12–DEC–1998	19057	ERS–2	–905	176.02
52	20–FEB–1999	20059	ERS–2	1170	209.00
53	27–MAR–1999	20560	ERS–2	–655	229.84
54	01–MAY–1999	21061	ERS–2	–70	183.91
55	05–JUN–1999	21562	ERS–2	680	220.38
56	14–AUG–1999	22564	ERS–2	1180	194.61
57	17–SEP–1999	42738	ERS–1	130	490.15
58	18–SEP–1999	23065	ERS–2	–95	195.70
59	22–OCT–1999	43239	ERS–1	190	490.26
60	23–OCT–1999	23566	ERS–2	–55	153.28
61	27–NOV–1999	24067	ERS–2	110	231.10
62	05–FEB–2000	25069	ERS–2	–425	191.09

Table C.1: (continued)

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**22–MAR–1997**	**10039**	**ERS–2**	**0**	**194.84**
63	11–MAR–2000	25570	ERS–2	–170	–606.49
64	15–APR–2000	26071	ERS–2	155	–246.87
65	20–MAY–2000	26572	ERS–2	685	–250.47
66	24–JUN–2000	27073	ERS–2	–845	–334.01
67	29–JUL–2000	27574	ERS–2	–590	–485.53
68	02–SEP–2000	28075	ERS–2	335	–354.54
69	07–OCT–2000	28576	ERS–2	30	185.85
70	11–NOV–2000	29077	ERS–2	305	84.58

Table C.2: ERS data for the Berlin area (track 437, frame 2547). Parameters are relative to the master acquisition, orbit 18327, acquired at 22–OCT–1998 10:06 am (UTC). Data are sorted on acquisition time, except for the master image, which is listed first.

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**22–OCT–1998**	**18327**	**ERS–2**	**0**	**150.31**
2	21–MAR–1996	4800	ERS–2	550	124.47
3	24–APR–1996	24974	ERS–1	765	424.55
4	25–APR–1996	5301	ERS–2	685	133.42
5	29–MAY–1996	25475	ERS–1	–90	478.43
6	30–MAY–1996	5802	ERS–2	–185	116.46
7	08–AUG–1996	6804	ERS–2	310	96.37
8	17–OCT–1996	7806	ERS–2	520	96.06
9	21–NOV–1996	8307	ERS–2	1205	141.26
10	30–JAN–1997	9309	ERS–2	515	113.50
11	15–MAY–1997	10812	ERS–2	80	152.56
12	24–JUL–1997	11814	ERS–2	240	123.88
13	02–OCT–1997	12816	ERS–2	220	155.99
14	06–NOV–1997	13317	ERS–2	–325	104.89
15	11–DEC–1997	13818	ERS–2	–660	155.36
16	15–JAN–1998	14319	ERS–2	–55	153.45
17	19–FEB–1998	14820	ERS–2	150	208.70
18	30–APR–1998	15822	ERS–2	525	149.10
19	09–JUL–1998	16824	ERS–2	–445	166.87
20	13–AUG–1998	17325	ERS–2	90	153.90
21	17–SEP–1998	17826	ERS–2	–415	157.36
22	26–NOV–1998	18828	ERS–2	–300	68.26
23	31–DEC–1998	19329	ERS–2	–630	138.28
24	11–MAR–1999	20331	ERS–2	–165	136.32
25	15–APR–1999	20832	ERS–2	610	135.58

Table C.2: (continued)

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**22–OCT–1998**	**18327**	**ERS–2**	**0**	**150.31**
26	24–JUN–1999	21834	ERS–2	425	217.04
27	29–JUL–1999	22335	ERS–2	640	165.75
28	02–SEP–1999	22836	ERS–2	–405	123.96
29	06–OCT–1999	43010	ERS–1	–5	452.46
30	07–OCT–1999	23337	ERS–2	–220	87.71
31	11–NOV–1999	23838	ERS–2	490	88.44
32	16–DEC–1999	24339	ERS–2	290	147.68
33	20–JAN–2000	24840	ERS–2	0	104.75
34	24–FEB–2000	25341	ERS–2	–485	297.59
35	30–MAR–2000	25842	ERS–2	75	–360.72
36	04–MAY–2000	26343	ERS–2	300	–105.86
37	08–JUN–2000	26844	ERS–2	–280	–377.23
38	13–JUL–2000	27345	ERS–2	–885	–451.72
39	17–AUG–2000	27846	ERS–2	–80	–505.75
40	21–SEP–2000	28347	ERS–2	655	–229.43
41	26–OCT–2000	28848	ERS–2	70	288.49
42	30–NOV–2000	29349	ERS–2	385	408.88
43	04–JAN–2001	29850	ERS–2	–550	326.15

Table C.3: Data for the Las Vegas area (track 356, frame 2871). Parameters are relative to the master acquisition, orbit 11232, acquired at 13–JUN–1997 18:22 pm (UTC). Data are sorted on acquisition time, except for the master image, which is listed first.

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**13–JUN–1997**	**11232**	**ERS–2**	**0**	**126.31**
2	21–APR–1992	4008	ERS–1	765	414.40
3	26–MAY–1992	4509	ERS–1	–410	451.23
4	08–SEP–1992	6012	ERS–1	630	385.02
5	17–NOV–1992	7014	ERS–1	–330	388.46
6	22–DEC–1992	7515	ERS–1	–920	403.13
7	02–MAR–1993	8517	ERS–1	215	334.59
8	06–APR–1993	9018	ERS–1	670	428.58
9	24–AUG–1993	11022	ERS–1	–490	419.27
10	02–NOV–1993	12024	ERS–1	590	387.99
11	30–MAR–1995	19382	ERS–1	–660	433.98
12	17–AUG–1995	21386	ERS–1	235	396.23
13	26–OCT–1995	22388	ERS–1	940	368.07
14	27–OCT–1995	2715	ERS–2	995	84.66
15	30–NOV–1995	22889	ERS–1	–120	390.69

Table C.3: (continued)

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**13–JUN–1997**	**11232**	**ERS–2**	**0**	**126.31**
16	01–DEC–1995	3216	ERS–2	–150	130.54
17	08–FEB–1996	23891	ERS–1	610	363.23
18	18–APR–1996	24893	ERS–1	710	375.13
19	23–MAY–1996	25394	ERS–1	40	426.95
20	24–MAY–1996	5721	ERS–2	–65	127.41
21	02–AUG–1996	6723	ERS–2	165	74.96
22	11–OCT–1996	7725	ERS–2	–210	116.45
23	15–NOV–1996	8226	ERS–2	1355	112.92
24	20–DEC–1996	8727	ERS–2	–250	233.27
25	24–JAN–1997	9228	ERS–2	180	163.68
26	28–FEB–1997	9729	ERS–2	110	164.59
27	04–APR–1997	10230	ERS–2	450	170.20
28	09–MAY–1997	10731	ERS–2	–5	138.68
29	18–JUL–1997	11733	ERS–2	0	107.82
30	22–AUG–1997	12234	ERS–2	390	106.88
31	26–SEP–1997	12735	ERS–2	105	130.04
32	31–OCT–1997	13236	ERS–2	–830	126.80
33	05–DEC–1997	13737	ERS–2	230	120.17
34	09–JAN–1998	14238	ERS–2	–45	182.63
35	13–FEB–1998	14739	ERS–2	–115	114.85
36	20–MAR–1998	15240	ERS–2	170	129.95
37	24–APR–1998	15741	ERS–2	300	141.95
38	29–MAY–1998	16242	ERS–2	50	168.65
39	29–JAN–1999	19749	ERS–2	70	171.49
40	18–JUN–1999	21753	ERS–2	–195	163.39
41	23–JUL–1999	22254	ERS–2	685	188.95
42	27–AUG–1999	22755	ERS–2	–1095	94.21
43	01–OCT–1999	23256	ERS–2	495	120.72
44	05–NOV–1999	23757	ERS–2	–250	110.40
45	10–DEC–1999	24258	ERS–2	–65	255.80
46	18–FEB–2000	25260	ERS–2	–340	316.16
47	11–AUG–2000	27765	ERS–2	665	–281.54
48	13–APR–2001	31272	ERS–2	–595	3835.75
49	09–NOV–2001	34278	ERS–2	–115	1211.36
50	14–DEC–2001	34779	ERS–2	–555	2598.71
51	18–JAN–2002	35280	ERS–2	–760	3712.29
52	22–FEB–2002	35781	ERS–2	475	3040.06
53	29–NOV–2002	3917	ENVISAT	15	186.99
54	03–JAN–2003	4418	ENVISAT	–65	232.05
55	07–FEB–2003	4919	ENVISAT	0	265.17

Table C.3: (continued)

#	Acq. date	Orbit	Sensor	$B_\perp$ [m]	f_{dc} [Hz]
1	**13–JUN–1997**	**11232**	**ERS–2**	**0**	**126.31**
56	18–APR–2003	5921	ENVISAT	1165	332.68
57	27–JUN–2003	6923	ENVISAT	–150	287.29
58	01–AUG–2003	7424	ENVISAT	–290	249.71
59	23–JAN–2004	9929	ENVISAT	–125	215.23
60	27–FEB–2004	10430	ENVISAT	–620	211.41
61	27–FEB–2004	46302	ERS–2	1125	1142.66
62	02–APR–2004	10931	ENVISAT	870	219.02

D

Developed Software

The algorithms described in this research are implemented at the German Aerospace Center (DLR-IMF) and integrated with existing software. First, section D.1 describes the software environment at the DLR and the programming philosophy that is applied. Implementation aspects are given in section D.2. A basic description of the developed modules is given in section D.3.

D.1 Computer environment at the DLR

The interferometric processing system GENESIS is developed since 1995 at DLR to perform basic SAR interferometric tasks, and, since then, is extended for specific projects, such as differential interferometry and the SRTM mission (Adam et al., 2003). The GENESIS system is designed for operational mass data processing (Eineder and Adam, 1997), and thus is a good starting point for a PS processing system, considering the large amounts of data that need to be handled. It runs on high-performance multiple CPU and standard SUN workstations as well as on Linux PCs. Many software modules take advantage of multiple CPU computers by supporting multi-threading. The SUN system available for the development of the PS processing software has eight (750 MHz UltraSPARC) CPUs, 32 GB of main memory and approximately 5 TB of online data storage available.

The GENESIS software is highly flexible due to its modular structure. The individual sub-systems use data of previous processing steps, but work independently from each other. Advantages of a modular approach are that algorithms can be conveniently compared (regarding accuracy, performance and robustness) and interchanged, both during development and at run time. The data dependencies and processing status are checked automatically during processing. For example, if an error occurs in the coregistration of an individual interferogram in the stack, only this single interferogram will be reprocessed. This is accomplished by using a Makefile or csh-scripts that check zero size status files that are written by each program on exit.

D.1.1 Passing information to programs

Passing information between programs can be done via parameter pools or via command line arguments. A parameter pool is a simple human-readable ASCII file with key/format/value entries. Each program can read from and write to the parameter pool. The format of the parameters can, for example, be string, 4 byte real, 8 byte real, 4 byte integer, etc., and be a scalar, vector or matrix. Block and line comments can also be written to the parameter pool.

A parameter pool is not suited to pass large amounts of data between programs. Instead, data are written to binary files, and the file names are passed using a key/value argument to the program. Input to programs can be given in several ways. In order of increasing priority, an argument is read from the environment, from a program control file (PCF), from the command line, or interactively from the prompt. The latter is only invoked when an exception occurs, e.g., a mandatory key is not present, or the name of a non-existing file is given. Normally PCF files are used to pass input, for reasons of standardization and repeatability of the processing. The PCF file can contain comments using common conventions. Due to the increasing order of priority, the value of an argument set in a PCF file is overwritten if it is also specified on the command line. Table D.1 lists common keys to all programs.

Table D.1: Common keys to all programs. "prg" stands for the name of the program.

key	default	meaning
-lfl	info	Logging file level
-lfn	log.prg	Logging file name
-lsl	info	Logging screen level
-pcf	pcf.prg	Name of program control file
-pid	prg	Program identificator
-status	prg	Base name of status file created on exit

D.1.2 Logging

Logging is extremely important for software developers to obtain intermediate values for all sorts of variables, and for users to obtain more general information on the processing. Normally, users are not interested in the value of a variable, except if it is an important variable, or when an unexpected event has occurred. To conveniently handle these different requirements, the logging is handled by increasing levels that can be set at run-time. These levels are trace, debug, info, warning, and alarm. For example, if the level is set to info,

then trace and debug information are suppressed. Typically, trace is used to output the value of a variable, debug for more important variables and for locating where something is executed in a program, info for user information on the process and important parameters, warning for manageable unexpected events, and alarm for fatal events. Furthermore, these levels are defined for logging to the screen and to a file. The level for the screen is normally set equal to or higher than that of the file in order to have more detailed information available in the log file without disrupting the overview of the processing.

D.1.3 Generic file format

All data stored in binary files use an adapted form of the SUN raster file (SRF) format. The SRF format is lossless and mainly used to store rectangular image data. The data is written after a header as a raw binary stream in row major order using big endian byte order. The header consists of eight 4 byte integers, optionally followed by a color map. These integers specify the magic number, width, height, depth, length, type, color map type, and color map length. The width must be even. The standard is defined for a data depth of 1, 8, 24, and 32 bit. The type specifies the way the raster is stored.

The GENESIS format is an extension of the SRF format that enables storing data of any type, including complex data, in big and small endian byte order. The header contains additional information on the data type and byte order using bits that are not used in the SRF format. This format is backward compatible with the SRF format for the old data types, except that the data are allowed to be of odd width, which is correctly handled by most image viewing software anyway.

D.1.4 Archiving

Archiving source code is important for continuous development, particularly with an increasing number of programs and users. At the DLR a CVS system is used for this purpose. CVS is a version control system that can record the history of source files. It can keep a log of who, when, and why changes occurred, etc. Since older versions can be easily retrieved, this allows, for example, to identify what modification of the software caused a bug. The advantage of CVS is that it allows to archive all kind of files (also binary files such as images for documentation purposes), and that several developers can work on the same code simultaneously. A web server is used to get an overview of all code within a project. Differences with previous version can be highlighted using CVSview, an online browser tool. It can also visualize the relations between functions in source code files and between classes. The program doxygen is used to create documentation directly from the source code.

When a collection of programs, intended to perform a specific task, such as PS processing or differential interferometry, has reached a stable level, a

frozen version is created. This is a snapshot of the code, which is copied to a new place. Together with the programs, PCF files are written for standard processing, as well as a setup script and a Makefile to run it. Since comments are allowed in the PCF files, they can be used to explain the intended usage. Users can rely on the characteristics of this version since it is not developed further. New algorithms are tested in the development versions of the programs. A new frozen version is created when significant improvements or additional functionality compared to the previous one is reached.

D.1.5 Implementation

Library functions handle argument parsing and i/o with the parameter pool and binary files. Using these functions prevents programming mistakes and guarantees conformity of the modules. Moreover, it puts a buffer between implementation and realization, i.e., a change in the library function updates all modules. For example, if a variable of type string was written to a parameter pool without surrounding quotes (by the library function), it may be awkward to work with strings containing blanks. By updating the definition of a string in the library such that it has quotes this can conveniently be remedied, without the need for changing the source code of every individual module. The computing environment IDL is used at the DLR for algorithm development. IDL is an interactive array-oriented language with numerous routines for mathematical analysis and visualization. The possibility of interactive examination of the content of variables and graphical representation allows for rapid prototyping. When a program is fully developed, the IDL code is ported to C++ for performance reasons. Similar library functions for argument parsing, parameter pool handling, and binary i/o are available in IDL and C++.

Although IDL is an interactive language, it can be run in batch mode from the UNIX prompt. This concept is used extensively to make the software (quasi-) operational. Each program written in IDL has a (csh) wrapper script that starts IDL and parses the command line input to it. The syntax for executing C++ and IDL code is therefore identical, and the user does not notice a difference between the two, except in execution time.

D.2 Specifics of the software

The general concepts described in the previous section are used for the development of the PS software. Further requirements imposed on each program are that its name, version, and location is logged, and that an online help function is present. For this reason a new common key, "-help", is introduced, see also Table D.1. The help displays all allowed keys, their default value and purpose. Furthermore, input keys given by the user are checked and a warning is issued if a key is not recognized. The external

documentation contains at least a description of the purpose of the program, a list of all allowed input keys and their meaning, and an algorithm flow description. Internally the programs are documented using comments and logging statements.

D.2.1 File formats

After creation of the differential interferometric stack, a limited number of points are selected, based on their amplitude time series, for further analysis. The range and azimuth sub-pixel position, temporal baseline, Doppler centroid frequency and orbit number are stored on disk in vectors, and the differential phase, amplitude, local height to phase conversion factor, and local look angle are stored in matrices. Additionally, flag arrays (vectors) are defined, which signal the status of points and interferograms. Figure D.1 shows the definition of these files. The required disk space is reduced considerably after this selection, since only a few percent of all points in the interferograms are selected. The definition of the bits in the flag arrays is given in Table D.2.

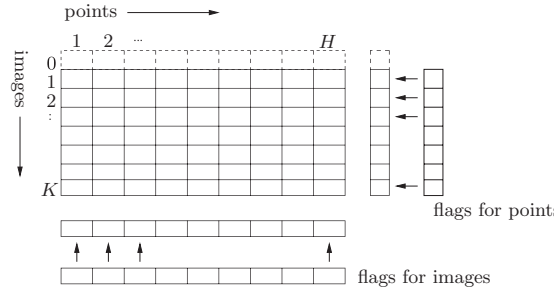


Fig. D.1: Definition of PS files. Information related to points is stored in the horizontal direction, while the vertical is related to images (SLC images or interferograms. The first row refers to the master acquisition, where appropriate.) Additionally, flag arrays are defined, which signal the status of points and interferograms.

The spare bits could be used in future version to flag, e.g., points that do not undergo significant displacement, interferograms that are corrected for phase trends, etc.

D.2.2 Parallelization

The flag array for the points can be conveniently used to divide the workload between several available processors. The first bit in the flag array is used to signal that a point needs to be computed. Since the computations can be performed independently of each other, each processor can be assigned

Table D.2: Meaning of bits used in flag arrays. A flag array is a byte vector that contains information on the usage of points or interferograms. A dash indicates that a bit is undefined.

	1	2	3	4	5	6	7	8
point	use	ref.	network	ref.	point	accepted	unwrapped	–	–	–
image	use	–		–		–			–	–	–

an equal number of points to compute. This is accomplished by splitting the original flag array in two new arrays and starting two simultaneous processings, see also Figure D.2. When the computations are finished, the results are combined as if they were computed by a single processor. This can be easily implemented as a shell around the original programs, which do not have to be changed themselves. The increase in speed is practically equal to the number of processors used.

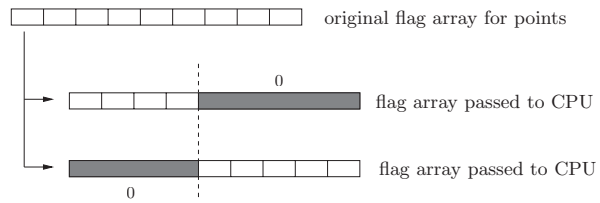


Fig. D.2: High-level parallelization using the flag array for points. Example for two CPUs. The points in the right half of the original flag array are flagged with a zero (gray) in the flag array passed to the first CPU, and vice versa for the second CPU.

D.3 Synopsis of the modules

Table D.3 provides an overview of the software developed at the DLR for PS processing. Most stack generation software are developed by Nico Adam and the estimation software by the author, both at the DLR.

Table D.3: Software modules, ordered alphabetically.

Module name	Purpose
Stack generation	
auto_doc	Extract information from parameter pools.
calibrate_ampl	Absolute calibration for range spreading loss, antenna pattern, calibration constant.
coregister	Estimate polynomial to align slave on master image using cross-correlation of intensity images.

Table D.3: (continued)

Module name	Purpose
coreg_tool	Manual clicking tool to identify tie points in two scenes.
cr_size	Estimate equivalent trihedral corner reflector dimension based on intensity.
crop_scene	Crop image to rectangular area of interest.
dem2insar	Compute synthetic phase image from input DEM (UTM) to radar coordinate system.
dfringe	Topographic correction using synthetic phase of reference DEM.
fringe	Compute phase of interferogram by complex multiplication.
geo_coregister	Estimate polynomial to align slave on master image using orbit geometry and reference DEM.
geo_est	Compute geometry of interferogram based on precise orbital information (baseline, coarse offset, etc.)
get_area	Interactive tool to select an area of interest to create the interferometric data stack.
get_dem	Obtain DEM (UTM) from data base for any area world wide.
osmpl_slc	Harmonic interpolation of image with factor two.
ps_get_pscat	Extract points to estimate from stack.
ps_import_slc	Initial import of SLC data in SUN raster format; create initial parameter pool.
ps_process_all	Perform all processing steps to create the differential interferometric stack.
ps_views_all	Create view graphs.
resample	Interpolation of slave image on master grid.
Estimation software	
aps_analyze	Compute and plot covariance and structure functions.
aps_blockfilt	Spatial low-pass filter using an averaging kernel.
aps_complexfilt	Spatial low-pass complex filter.
aps_low_pass_time	Temporal low-pass filter.
baselineplot	Plot baseline distribution and selects master based on the total coherence function.
calibrate_analyze	Plot histograms of amplitude of user selected region.
construct_network	Construction of the reference network.
correct_dem	Remove holes and invalid values from a DEM.
create_base_f	General purpose utility to generate base functions.
estimate_all_points	Estimate parameters at points with respect to the reference network.

Table D.3: (continued)

Module name	Purpose
estimate_arcs_network	Integer least-squares estimation of arcs of the reference network.
estimate_unwrapped	Estimate of parameters using the unwrapped data.
geocode_bg	Geo-referencing of a background image.
ls_integrate_network	Least-squares integration of the estimated parameters at the arcs of the reference network, including alternative hypothesis testing.
ppool_update	Change parameter pool(s) based on a master pool.
ps_analyze	Interactive tool to analyze the phase time series.
ps_los2vert	Convert line-of-sight displacement to vertical displacement using the local incidence angle.
ps_movie	Create (interpolated) movies of data.
ps_plot	Generic plotting program for x,y,z data.
ps_x2ps_wgs	Compute point positions in WGS84.
radarcode_dem	Transform DEM to radar geometry.
residual_phase	Correct phase for estimated parameters or trends.
select_ifgs	Initial selection of interferograms to use in estimation.
select_pscs	Initial selection of points to estimate.
select_ref_point	Select the reference point of the reference network.
select_reliable_points	Select reliable points based on thresholds.
setup_ps_estimation	Start estimator, copy Makefile, template configuration.
sparsify_data	Select points for reference network.
speedup_ls_integration	Remove points and arcs in pre-processing step.
srf_math	General operations on SUN raster files.
trend_analyze	Analyze trends in phase at selected points.
unwrap_sparse	Sparse MCF unwrapper for a stack.
var_factors_aps	Estimate variance components as function of distance.
var_factors_noise	Estimate variance components.

E

Software on the CDROM

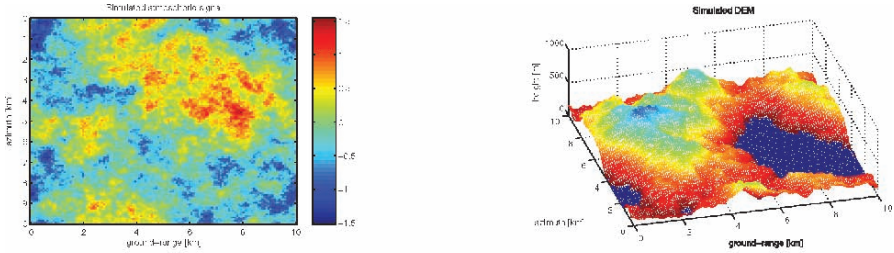
This appendix provides a quick-start to the programs on the CDROM that accompanies this book. These programs demonstrate the capabilities of integer least-squares estimation and of variance component estimation. It is hoped that these programs increase the reader's understanding of the theoretical concepts that are described in the Chapters 2, 3, and 4, that they make these techniques accessible for a wide audience, and that they make them practically applicable, particularly for displacement parameter estimation using Persistent Scatterer Interferometry.

After a brief introduction and an overview of the key features given in section E.1, installation of the software is described in section E.2. Finally, section E.3 gives specific hints how to start working with the programs.

E.1 Introduction

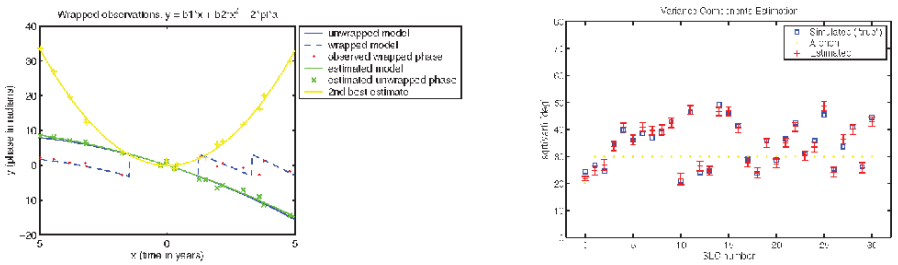
The software on the CDROM consists of the STUN Matlab Toolbox. The fundamental capabilities of this toolbox are variance component estimation and integer least-squares estimation. Furthermore, test data and demonstration scripts are provided that explain the key concepts and usage of these functions. These functions are intended to demonstrate the practical application of the theory described in this book. They serve as essential building blocks in a persistent scatterer processing system. The ILS routines are adapted from original code distributed in the lambda toolbox, see (Delft University of Technology, 2005). The changes make the code better suited for PSI processing. Because the dimension of the problem is bigger for PSI (there are more ambiguities to fix compared to GPS), the code is optimized to handle frequent calls using the same mathematical model. Furthermore, the input parsing is changed to be more consistent and to use pre-computed matrices which previously were computed inside loops. Optimization for speed is also achieved by vectorization of loops and by assuming that in general only the best candidate needs to be found during the ILS search of the

hyper-ellipsoid. Moreover, additional help is included which is available via the online Matlab help interface, as comments in the source code, and by using the demonstration scripts. The source code also contains references to appropriate chapters, sections and equations.



(a) (b)

Fig. E.1: Fractal atmospheric delay and topography generated by the phase simulation program "simphi". While in Matlab, type help simphi for more information and ready-to-use examples.



(a) ilsdemo1db.m (b) vcedemo.m

Fig. E.2: Screenshot of the demonstration program "ilsdemo1db". To run this program, start Matlab, and type ilsdemo1db .

Key Features

- Simulation scripts to generate data for a single master stack of differential interferograms, according to the functional model given in section 2.2.1.
- Variance Component Estimation, described in section 4.3 and Appendix A, using the stochastic model derived in section 2.2.2.
- Integer Least-Squares ambiguity resolution using the LAMBDA method described in Chapter 3.

- Demonstration scripts that can be used to learn the basic call sequences and can be customized for a specific implementation.
- Online help and back-references to this book to find more information.

Examples of the data that can be simulated with the program "simphi" are shown in Fig. E.1. See Fig. E.2(a) for a screenshot of the demonstration script "ilsdemo1db". With this script the basics of ILS are demonstrated by simulation of a second degree polynomial with random parameters and noise. The coefficients of the polynomial are then estimated using wrapped data, together with a second best fitting set of parameters. Fig. E.2(b) shows a screenshot of the demonstration program to estimate the variance components using a stack of simulated data. Help on these demonstration scripts can be obtained from within Matlab.

E.2 Installation

The program Matlab is required to run the routines and demonstrations. The M-file scripts have been tested with Matlab versions 5.3 upward. To learn more about Matlab, see (MathWorks, 2005).

To install this toolbox, only the directory `stun` on the CDROM needs to be copied to your hard drive. After you have copied the toolbox, start Matlab. Make sure that the toolbox can be found by adding the directory where you copied it in the Matlab search path. For example, you can do this by using the script `addpath`. Type:

```
>> help addpath
```

for more information. After this, typing `help stun` should give:

```
>> help stun

  STUN algorithm for Persistent Scatterers Interferometry

  Elementary Functions.
    bs_success  - Success rate of the bootstrap.
    enscoh      - Ensemble coherence.
    plotarc     - Plot arcs of network in color.
    plotps      - Plot PS points in color.
    sparsify    - Bin points in grid cells.
    wrap        - Wrap phase data.

  Functional Model.
    simacq      - Simulate acquisition baselines.
    simphi      - Simulate ERS-like phase observations.
    simpos      - Simulate 2D positions.

  Stochastic Model.
    psivce      - Variance Component Estimation for PSI.
```

```
  psivcmtx    - VC-matrix for double-difference observations.

Ambiguity Resolution.
  ebs         - Extended bootstrap fixed solution.
  ils         - Integer least-squares fixed solution.
  ltdl        - LTDL decomposition Q=L.'*D*L.
  zt          - Z-transformation (decorrelation).

Demonstrations.
  ilsdemo1d   - ILS estimation of the slope of wrapped line.
  ilsdemo1db  - ILS estimation of 2nd degree polynomial.
  stundemo    - Main demonstration script.
  vcedemo     - Variance Component Estimation.

Data Sets.
  poly_good.mat    - Example data set for ilsdemo1db.
  poly_wrong.mat   - Example data set for ilsdemo1db.
  vcedemodat.mat   - Reference results for vcedemo.
  stundemodat1.mat - Fractal displacement example data set.
  stundemodat2.mat - Example data with atmosphere.
```

E.3 How to begin

After installation of the STUN toolbox, it is advised to first read the general help as described in section E.2. To obtain the help of a specific function in the STUN toolbox, type `help` followed by the M-file name, for example

```
>> help ils
```

gives a synopsis of the program "ils" and explains the input and output variables. It also provides some short examples of how to run it.

- People interested in variance component estimation are encouraged to first run the program "vcedemo" and to inspect the code of this M-file.
- For those people mainly interested in integer least-squares estimation, a good starting point is to run the demonstration program "ilsdemo1d" and follow the instructions on the screen.

The demonstration program "stundemo" combines these concepts, i.e., estimation of variance components and estimation of DEM error and linear displacement rate differences. Furthermore, this program creates a reference network, and the estimated parameters at the arcs are integrated to obtain them with respect to the reference point. The source code of this demonstration program can be extended for your own applications.

References

Adam, N., Kampes, B. M. and Eineder, M.: 2004. The development of a scientific persistent scatterer system: Modifications for mixed ERS/ENVISAT time series. *ENVISAT & ERS Symposium,* Salzburg, Austria, 6–10 September, 2004. pp. 1–9 (cdrom).

Adam, N., Kampes, B. M., Eineder, M., Worawattanamateekul, J. and Kircher, M.: 2003. The development of a scientific permanent scatterer system. *ISPRS Workshop High Resolution Mapping from Space, Hannover, Germany, 2003.* pp. 1–6 (cdrom).

Amelung, F., Galloway, D. L., Bell, J. W., Zebker, H. A. and Laczniak, R. J. 1999. Sensing the ups and downs of Las Vegas: InSAR reveals structural control of land subsidence and aquifer-system deformation. *Geology* 27(6), 483–486.

Amelung, F., Jónssen, S., Zebker, H. and Segall, P. 2000. Widespread uplift and trap door faulting on Galápagos volcanoes observed with radar interferometry. *Nature* 407(6807), 993–996.

Arnaud, A., Adam, N., Hanssen, R., Inglada, J., Duro, J., Closa, J. and Eineder, M.: 2003. ASAR ERS interferometric phase continuity. *International Geoscience and Remote Sensing Symposium,* Toulouse, France, 21–25 July, 2003. pp. 1–3 (cdrom).

Arnaud, A., Closa, J., Hanssen, R., Adam, N., Eineder, M., Inglada, J., Fitoussi, G. and Kampes, B.: 2004. Development of algorithms for the exploitation of ERS-Envisat using the stable points network. *Technical report.* Altamira Information. Barcelona, Spain. European Space Agency Study report ESA Contract Nr. 16702/02/I-LG.

Arrigoni, M., Colesanti, C., Ferretti, A., Perissin, D., Prati, C. and Rocca, F.: 2003. Identification of the location phase screen of ERS-ENVISAT permanent scatterers. *Third International Workshop on ERS SAR Interferometry 'FRINGE03',* Frascati, Italy, 1–5 December, 2003. pp. 1–3 (cdrom).

Baarda, W.: 1968. *A testing procedure for use in geodetic networks.* Vol. 5 of *Publications on Geodesy.* 2 edn. Netherlands Geodetic Commission. Delft.

Bamler, R. and Hartl, P. 1998. Synthetic aperture radar interferometry. *Inverse Problems* 14, R1–R54.

Bamler, R. and Schättler, B.: 1993. SAR data acquisition and image formation. *in* G. Schreier (ed.), *SAR Geocoding: data and systems.* Wichmann Verlag. Karlsruhe. pp. 53–102.

Basilico, M., Ferretti, A., Novali, F., Prati, C. and Rocca, F.: 2004. Advances in permanent scatterer analysis: semi and temporary PS. *European Conference on Synthetic Aperture Radar,* Ulm, Germany, 25–27 May, 2004. pp. 349–349.

Bell, J. W., Amelung, F., Ramelli, A. R. and Blewitt, G. 2002. Land subsidence in Las Vegas, Nevada, 1935–2000: New geodetic data show evolution, revised spatial patterns, and reduced rates. *Geoscience and Engineering Geoscience* VIII(3), 155–174.

Berardino, P., Casu, F., Fornaro, G., Lanari, R., Manunta, M., Pepe, A. and Sansosti, E.: 2003. Small baseline DIFSAR techniques for earth surface deformation analysis. *Third International Workshop on ERS SAR Interferometry 'FRINGE03',* Frascati, Italy, 1–5 December, 2003. p. 6.

Berardino, P., Fornaro, G., Lanari, R. and Sansosti, E. 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing* 40(11), 2375–2383.

Berliner Zeitung: 2004. Berliner Zeitung internet news archive. http://www.berlinonline.de/berliner-zeitung/archiv/.bin/dump.fcgi/2004/0426/lokales/0027/ (Accessed March, 2004).

Bianchi, M.: 2003. *Phase ambiguity estimation on permanent scatterers in SAR interferometry: the integer least-squares approach.* Master's thesis. Politecnico di Milano/Delft University of Technology.

Brenner, A. R. and Ender, J. H. G.: 2004. Airborne SAR imaging with subdecimeter resolution. *European Conference on Synthetic Aperture Radar,* Ulm, Germany, 25–27 May, 2004. pp. 267–270.

Bürgmann, R., Schmidt, D., Nadeau, R. M., d'Alessio, M., Fielding, E. J., Manaker, D., McEvilly, T. V. and Murray, M. H. 2000. Earthquake potential along the Northern Hayward Fault, California. *Science* 289, 1178–1182.

Chen, C. W.: 2001. *Statistical-cost network-flow approaches to two-dimensional phase unwrapping for radar interferometry.* PhD thesis. Stanford University.

Chen, C. W. and Zebker, H. A. 2002. Phase unwrapping for large SAR interferograms: Statistical segmentation and generalized network models. *IEEE Transactions on Geoscience and Remote Sensing* 40, 1709–1719.

Closa, J.: 1998. The influence of orbit precision in the quality of ERS SAR interferometric data. *Technical Report ES-TN-APP-APM-JC01.* ESA.

Colesanti, C., Ferretti, A., Ferrucci, F., Prati, C. and Rocca, F.: 2000. Monitoring known seismic faults using the Permanent Scatterers (PS) technique. *International Geoscience and Remote Sensing Symposium,* Honolulu, Hawaii, 24–28 July, 2000. pp. 2221–2223.

Colesanti, C., Ferretti, A., Locatelli, R. and Savio, G.: 2003a. Multi-platform permanent scatterers analysis: first results. *Second GRSS/ISPRS Joint Workshop on "Data Fusion and Remote Sensing over Urban Areas",* Berlin, Germany, 22–23 May, 2003. pp. 52–56.

Colesanti, C., Ferretti, A., Novali, F., Prati, C. and Rocca, F. 2003b. SAR monitoring of progressive and seasonal ground deformation using the Permanent Scatterers Technique. *IEEE Transactions on Geoscience and Remote Sensing* 41(7), 1685–1701.

Colesanti, C., Ferretti, A., Prati, C. and Rocca, F.: 2002. Full exploitation of the ERS archive: Multi data set permanent scatterers analysis. *International Geoscience and Remote Sensing Symposium,* Toronto, Canada, 24–28 June, 2002. pp. 1–3 (cdrom).

Colesanti, C., Ferretti, A., Prati, C. and Rocca, F. 2003c. Monitoring landslides and tectonic motions with the Permanent Scatterers Technique. *Engineering Geology* 68, 3–14.

Colesanti, C., Ferretti, A., Prati, C., Perissin, D. and Rocca, F.: 2003d. ERS-ENVISAT Permanent Scatterers Interferometry. *International Geoscience and Remote Sensing Symposium,* Toulouse, France, 21–25 July, 2003. Vol. 2. pp. 1130–1132.

Costantini, M. 1998. A novel phase unwrapping method based on network programming. *IEEE Transactions on Geoscience and Remote Sensing* 36(3), 813–821.

Costantini, M. and Rosen, P.: 1999. A generalized phase unwrapping approach for sparse data. *International Geoscience and Remote Sensing Symposium,* Hamburg, Germany, 28 June–2 July, 1999. pp. 1–3 (cdrom).

Cumming, I. and Wong, F.: 2005. *Digital Processing Of Synthetic Aperture Radar Data: Algorithms And Implementation.* Artech House Publishers. New York. ISBN 1580530583.

Curlander, J. C. and McDonough, R. N.: 1991. *Synthetic aperture radar: systems and signal processing.* John Wiley & Sons, Inc. New York.

De Heus, H. M., Joosten, P., Martens, M. H. F. and Verhoef, H. M. E.: 1994. Geodetische deformatie analyse: 1d- deformatieanalyse uit waterpas-netwerken. *Technical Report 5.* Delft University of Technology, LGR Series. Delft.

Delft University of Technology: 2005. Mathematical Geodesy and Positioning web page. http://enterprise.lr.tudelft.nl/mgp/ (Accessed April, 2005).

Deutscher Wetterdienst: 2004. Deutscher Wetterdienst. http://www.dwd.de/de/FundE/Klima/KLIS/daten/online/nat/ausgabe_monatswerte.htm (Accessed March, 2004).

Eineder, M. 2003. Efficient simulation of SAR interferograms of large areas and of rugged terrain. *IEEE Transactions on Geoscience and Remote Sensing* 41(6), 1415–1427.

Eineder, M. and Adam, N.: 1997. A flexible system for the generation of interferometric sar products. *International Geoscience and Remote Sensing Symposium,* Singapore, 3–8 August, 1997.

Eineder, M. and Holzner, J.: 1999. Phase unwrapping of low coherence differential interferograms. *International Geoscience and Remote Sensing Symposium,* Hamburg, Germany, 28 June–2 July, 1999. pp. 1–4 (cdrom).

Elachi, C.: 1987. *Introduction To The Physics and Techniques of Remote Sensing*. 2 edn. John Wiley & Sons. New York.

Evans, D. L., Price, J. L. and Barron, W. G.: 2000. Profiles of general demographic characteristics; 2000 census of population and housing; Nevada. *Technical report* U.S. Census Bureau. http://www.census.gov/ (Accessed March, 2004).

Farina, P.: 2003. Integration of permanent scatterers analysis and high resolution optical images within landslide risk analysis. *Third International Workshop on ERS SAR Interferometry 'FRINGE03',* Frascati, Italy, 1–5 December, 2003. pp. 1–3 (cdrom).

Fernandez, D. E., Meadows, P. J., Schaettler, B. and Mancini, P.: 1999. ERS attitude errors and its impact on the processing of SAR data. *CEOS SAR Workshop,* ESA-CNES, Toulouse, France, 26–29 October, 1999. pp. 1–9 (cdrom).

Ferretti, A., Prati, C. and Rocca, F. 1999a. Multibaseline InSAR DEM reconstruction: The wavelet approach. *IEEE Transactions on Geoscience and Remote Sensing* 37(2), 705–715.

Ferretti, A., Prati, C. and Rocca, F.: 1999b. Non-uniform motion monitoring using the permanent scatterers technique. *Second International Workshop on ERS SAR Interferometry 'FRINGE99',* Liège, Belgium, 10–12 November, 1999. ESA. pp. 1–6.

Ferretti, A., Prati, C. and Rocca, F.: 1999c. Permanent scatterers in SAR interferometry. *International Geoscience and Remote Sensing Symposium,* Hamburg, Germany, 28 June–2 July, 1999. pp. 1–3.

Ferretti, A., Prati, C. and Rocca, F. 2000a. Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing* 38(5), 2202–2212.

Ferretti, A., Prati, C. and Rocca, F.: 2000b. Process for radar measurements of the movement of city areas and landsliding zones. International Application Published under the Patent Cooperation Treaty (PCT).

Ferretti, A., Prati, C. and Rocca, F. 2001. Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing* 39(1), 8–20.

Freeman, A. 1992. SAR calibration: An overview. *IEEE Transactions on Geoscience and Remote Sensing* 30(6), 1107–1121.

Gatelli, F., Monti Guarnieri, A., Parizzi, F., Pasquali, P., Prati, C. and Rocca, F. 1994. The wavenumber shift in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing* 32(4), 855–865.

Ghiglia, D. C. and Pritt, M. D.: 1998. *Two-dimensional phase unwrapping: theory, algorithms, and software*. John Wiley & Sons, Inc. New York.

Hanssen, R.: 1998. *Atmospheric heterogeneities in ERS tandem SAR interferometry*. Delft University Press. Delft, the Netherlands.

Hanssen, R., Amelung, F. and Zebker, H. 1998. Geodetic interpretation of land subsidence measurements at the Cerro Prieto geothermal field monitored by radar interferometry. *EOS Transactions, AGU* 79(45), F37.

Hanssen, R. F.: 2001. *Radar Interferometry: Data Interpretation and Error Analysis*. Kluwer Academic Publishers. Dordrecht.

Hanssen, R. F. and Ferretti, A. 2002. Parameter estimation in PS-InSAR deformation studies using integer least-squares techniques. *EOS Trans. AGU, Fall Meet. Suppl., Abstract G62A-06* 83(47), F37.

Hanssen, R. F., Teunissen, P. J. G. and Joosten, P.: 2001. Phase ambiguity resolution for stacked radar interferometric data. *International Symposium on Kinematic Systems in Geodesy,* Geomatics and Navigation, Banff, Canada, 5–8 June, 2001. pp. 317–320.

Hanssen, R. F., Weckwerth, T. M., Zebker, H. A. and Klees, R. 1999. High-resolution water vapor mapping from interferometric radar measurements. *Science* 283, 1295–1297.

Hanssen, R., Vermeersen, B., Scharroo, R., Kampes, B., Usai, S., Gens, R. and Klees, R. 2000. Deformatiepatroon van de aardbeving van 17 augustus 1999 in Turkije gemeten met satelliet radar interferometrie. *Remote Sensing Nieuwsbrief* 90, 42–44. In Dutch.

Hilley, G. E., Burgmann, R., Ferretti, A., Novali, F. and Rocca, F. 2004. Dynamics of slow-moving landslides from Permanent Scatterer analysis. *Science* 304, 1952–1955.

Hoffmann, J., Zebker, H. A., Galloway, D. L. and Amelung, F. 2001. Seasonal subsidence and rebound in Las Vegas Valley, Nevada, observed by synthetic aperture radar interferometry. *Water Resources Research* 37(6), 1551–1566.

Hooper, A., Zebker, H., Segall, P. and Kampes, B. 2004. A new method for measuring deformation on volcanoes and other non-urban areas using InSAR persistent scatterers. doi: 10.1029/2004GL021737 *Geophysical Research Letters*.

Joughin, I. R.: 1995. *Estimation of Ice-Sheet Topography and Motion Using Interferometric Synthetic Aperture Radar*. PhD thesis. University of Washington.

Kampes, B. M. and Adam, N.: 2004. Deformation parameter inversion using permanent scatterers in interferogram time series. *European Conference on Synthetic Aperture Radar,* Ulm, Germany, 25–27 May, 2004. pp. 341–344.

Kampes, B. M. and Hanssen, R. F. 2004. Ambiguity resolution for permanent scatterer interferometry. *IEEE Transactions on Geoscience and Remote Sensing* 42(11), 2446–2453.

Kircher, M., Adam, N., Hoffmann, J., Kampes, B. M., Neugebauer, H. J. and Roth, A.: 2003a. Application of permanent scatterers on mining induced

subsidence. *Third International Workshop on ERS SAR Interferometry 'FRINGE03',* Frascati, Italy, 1–5 December, 2003. pp. 1–6 (cdrom).

Kircher, M., Roth, A., Adam, N., Kampes, B. M. and Neugebauer, H. J.: 2003b. Remote sensing observation of mining induced subsidence by means of differential SAR-interferometry. *International Geoscience and Remote Sensing Symposium,* Toulouse, France, 21–25 July, 2003. pp. 1–3 (cdrom).

Klees, R. and Massonnet, D. 1999. Deformation measurements using SAR interferometry: potential and limitations. *Geologie en Mijnbouw* 77, 161–176.

Koch, K.-R.: 1988. *Parameter estimation and hypothesis testing in linear models.* Springer-Verlag. New York.

Kohlhase, A. O., Feigl, K. L. and Massonnet, D. 2003. Applying differential InSAR to orbital dynamics: a new approach for estimating ERS trajectories. doi: 10.1007/s00190-003-0336-3 *Journal of Geodesy* 77(9), 493–502.

Lanari, R., Mora, O., Manunta, M., Mallorqui, J. J., Berardino, P., Guarino, S., Zeni, G., Pepe, A. and Sansosti, E.: 2003. A differential SAR interferometry approach for monitoring urban deformation phenomena. *Second GRSS/ISPRS Joint Workshop on "Data Fusion and Remote Sensing over Urban Areas",* Berlin, Germany, 22–23 May, 2003. pp. 57–61.

Laur, H., Bally, P., Meadows, P., Sanchez, J., Schättler, B., Lopinto, E. and Esteban, D.: 1998. Derivation of the backscattering coefficient σ^0 in ESA ERS SAR PRI products. *Technical Report ES-TN-RS-PM-HL09.* ESA. Issue 2, Rev. 5b.

Massonnet, D. and Sigmundsson, F.: 2000. Remote sensing of volcano deformation by radar interferometry from various satellites. *in* P. J. Mouginis-Mark, J. A. Crisp and J. H. Fink (eds), *Remote sensing of active volcanism.* Geophysical Monographs 116. American Geophysical Union. Washington, DC. pp. 207–221.

Massonnet, D., Briole, P. and Arnaud, A. 1995. Deflation of Mount Etna monitored by spaceborne radar interferometry. *Nature* 375, 567–570.

Massonnet, D., Rossi, M., Carmona, C., Adagna, F., Peltzer, G., Feigl, K. and Rabaute, T. 1993. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature* 364(8), 138–142.

MathWorks, T.: 2005. Company web page. http://www.mathworks.com/ (Accessed October, 2005).

Meyer, F.: 2004. *Simultane Schaetzung von Topographie und Dynamik polarer Gletscher aus multi-temporalen SAR Interferogrammen.* PhD thesis. Munich University of Technology. Munich.

Mora, O., Mallorqui, J. J. and Duro, J.: 2002. Generation of deformation maps at low resolution using differential interferometric SAR data. *International Geoscience and Remote Sensing Symposium,* Toronto, Canada, 24–28 June, 2002.

Musson, R. M. W., Haynes, M. and Ferretti, A. 2004. Space-based tectonic modeling in subduction areas using PSInSAR. *Seismological Research Letters* 75(5), 598–606.

National Weather Service: 2004. National Weather Service. http://www.nws.noaa.gov/ (Accessed March, 2004).

Neyman, J. and Pearson, E. S. 1933. On the problem of the most efficient tests of statistical hypotheses. *Philosophic Transactions of the Royal Society London, Series A* 231, 289–337.

Nigel Press Associates: 2004. Company web page. http://www.npagroup.com/ (Accessed March, 2004).

Pavelko, M. T.: 2003. Estimates of hydraulic properties from a one-dimensional numerical model of vertical aquifer-system deformation, lorenzi site, las vegas, nevada. *Technical Report Water-Resources Investigations Report 03–4083*. U.S. Geological Survey.

Prati, C. and Rocca, F.: 1994. Use of the spectral shift in SAR interferometry. *Second ERS-1 Symposium—Space at the Service of our Environment,* Hamburg, Germany, 11–14 October, 1993. ESA SP-361. pp. 691–696.

Rabus, B., Eineder, M., Roth, A. and Bamler, R. 2003. The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing* 57, 241–261.

Raney, R. K. 1986. Doppler properties of radars in circular orbits. *International Journal of Remote Sensing* 7(9), 1153–1162.

Raney, R. K., Freeman, T., Hawkins, R. W. and Bamler, R.: 1994. A plea for radar brightness. *International Geoscience and Remote Sensing Symposium,* Pasadena, CA, USA, 8–12 August, 1994. pp. 1090–1092.

Rodriguez, E. and Martin, J. M. 1992. Theory and design of interferometric synthetic aperture radars. *IEE Proceedings-F* 139(2), 147–159.

Rosen, P., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S., Rodríguez, E. and Goldstein, R. 2000. Synthetic aperture radar interferometry. *Proceedings of the IEEE* 88(3), 333–382.

Salvi, S., Atzori, S., Tolomei, C., Allievi, J., Ferretti, A., Rocca, F., Prati, C., Stramondo, S. and Feuillet, N. 2004. Inflation rate of the colli albani volcanic complex retrieved by the Permanent Scatterer SAR interferometry technique. *Geophysical Research Letters* 31, 1–4 (cdrom).

Shannon, C. E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 623–656.

Shimoni, M., Hanssen, R., Van der Meer, F., Kampes, B. and Ben-Dor, E.: 2002. Salt diapir movements using SAR interferometry in the Lisan Peninsula, Dead Sea rift. *Proc. SPIE Vol. 4543, SAR Image Analysis, Modeling, and Techniques IV*. International Society for Optical Engineering. pp. 151–160.

Soergel, U., Thoennessen, U., Stilla, U. and Brenner, A. R.: 2004. New oppertunities for analysis of urban areas in high resolution SAR data. *European Conference on Synthetic Aperture Radar,* Ulm, Germany, 25–27 May, 2004. pp. 415–418.

Suchandt, S., Breit, H., Adam, N., Eineder, M., Schättler, B., Runge, H., Roth, A. and Mikusch, E.: 2001. The shuttle radar topography mission. *ISPRS*

Workshop, High Resolution Mapping from Space, Hannover, Germany, 19–21 September, 2002. ISPRS. pp. 235–242.

Tele-Rilevamento Europa: 2004. Company web page. http://www.treuropa.com/ (Accessed March, 2004).

Terrafirma: 2005. Project web page. http://www.terrafirma.eu.com/ (Accessed December, 2005).

Teunissen, P. J. G.: 1994. Least-squares estimation of the integer GPS ambiguities. *Publications and annual report 1993.* number 6 in *LGR-Series.* Delft geodetic computing centre. pp. 59–74.

Teunissen, P. J. G. 1995. The least-squares ambiguity decorellation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy* 70(1–2), 65–82.

Teunissen, P. J. G. 1999. A theorem on maximizing the probability of correct integer estimation. *Artificial Satellites* 34(1), 3–9.

Teunissen, P. J. G.: 2000a. *Adjustment theory; an introduction.* 1 edn. Delft University Press. Delft.

Teunissen, P. J. G.: 2000b. *Testing theory; an introduction.* 1 edn. Delft University Press. Delft.

Teunissen, P. J. G. 2003a. A carrier phase ambiguity estimator with easy to evaluate fail-rate. *Artificial Satellites* 38(3), 89–96.

Teunissen, P. J. G. 2003b. Integer aperture GNSS ambiguity resolution. *Artificial Satellites* 38(3), 79–88.

Teunissen, P. J. G. 2004. Penalized GNSS ambiguity resolution. *Journal of Geodesy* 78, 235–244.

Teunissen, P. J. G., de Jonge, P. J. and Tiberius, C. C. J. M. 1995a. The invertible GPS ambiguity transformation. *Manuscripta Geodaetica* 20, 489–497.

Teunissen, P. J. G., de Jonge, P. J. and Tiberius, C. C. J. M.: 1995b. The LAMBDA-method for fast GPS processing. *GPS Technology Applications, Bucharest, Romania, Sep 26–29.*

Teunissen, P. J. G., de Jonge, P. J. and Tiberius, C. C. J. M. 1995c. A new way to fix carrier-phase ambiguities. *GPS World* pp. 58–61.

Ulaby, F. T., Moore, R. K. and Fung, A. K.: 1982. *Microwave remote sensing: active and passive. Vol. 2. Radar remote sensing and surface scattering and emission theory.* Addison-Wesley. Reading.

Usai, S.: 1997. The use of man-made features for long time scale insar. *International Geoscience and Remote Sensing Symposium,* Singapore, 3–8 August, 1997. pp. 1542–1544.

Usai, S. 2000. An analysis of the interferometric characteristics of anthropogenic features. *IEEE Transactions on Geoscience and Remote Sensing* 38(3), 1491–1497.

Usai, S. and Hanssen, R.: 1997. Long time scale INSAR by means of high coherence features. *Third ERS Symposium—Space at the Service of our Environment,* Florence, Italy, 17–21 March, 1997. pp. 225–228.

Usai, S. and Klees, R. 1999. SAR interferometry on very long time scale: A study of the interferometric characteristics of man-made features. *IEEE Transactions on Geoscience and Remote Sensing* 37(4), 2118–2123.

Van der Kooij, M. W. A.: 2003. Coherent target analysis. *Third International Workshop on ERS SAR Interferometry 'FRINGE03',* Frascati, Italy, 1–5 December, 2003.

Van der Kooij, M. W. A. and Lambert, A.: 2002. Results of processing and analysis of large volumes of repeat-pass InSAR data of Vancouver and Mount Meager (B.C.). *International Geoscience and Remote Sensing Symposium,* Toronto, Canada, 24–28 June, 2002.

Verhoef, H. M. E.: 1997. Geodetische deformatie analyse. Lecture notes, Delft University of Technology, Faculty of Geodetic Engineering, in Dutch.

Walter, D., Hoffmann, J., Kampes, B. and Sroka, A.: 2004. Radar interferometric analysis of mining induced surface subsidence using permanent scatterer. *ENVISAT & ERS Symposium,* Salzburg, Austria, 6–10 September, 2004. pp. 1–8 (cdrom).

Wegmuller, U.: 2003. Potential of interferometry point target analysis using small data stacks. *Third International Workshop on ERS SAR Interferometry 'FRINGE03',* Frascati, Italy, 1–5 December, 2003. pp. 1–3 (cdrom).

Werner, C., Wegmuller, U., Strozzi, T. and Wiesmann, A.: 2003. Interferometric point target analysis for deformation mapping. *International Geoscience and Remote Sensing Symposium,* Toulouse, France, 21–25 July, 2003. pp. 1–3 (cdrom).

Yong, Y., Chao, W., Hong, Z., Zhi, L. and Xin, G.: 2002. A phase unwrapping method based on minimum cost flows method in irregular network. *International Geoscience and Remote Sensing Symposium,* Toronto, Canada, 24–28 June, 2002.

Zebker, H. A. and Lu, Y. 1998. Phase unwrapping algorithms for radar interferometry: residue-cut least-squares, and synthesis algorithms. *Journal of the Optical Society of America A.* 15(3), 586–598.

Zebker, H. A. and Villasenor, J. 1992. Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing* 30(5), 950–959.

Zebker, H. A., Rosen, P. A., Goldstein, R. M., Gabriel, A. and Werner, C. L. 1994. On the derivation of coseismic displacement fields using differential radar interferometry: The Landers earthquake. *Journal of Geophysical Research* 99(B10), 19617–19634.

About the Author

Bert Kampes (1974, The Netherlands) was awarded the Master of Science title from the department of Geodetic Engineering at Delft University of Technology in 1998, and the PhD title from the department of Aerospace Engineering in 2005. His MS graduation work focused on physical geodesy, particularly on different methods for the estimation of spherical harmonic coefficients that describe the gravity field of the earth. He worked as researcher at Delft University of Technology from 1998 to 2001. In this time he developed the "Delft Object-oriented Radar Interferometric Software" (*Doris*). This software is in the public domain and used worldwide for the generation of Digital Elevation Models and deformation maps. In 2001 he started his PhD research on the application of the Permanent Scatterer technique, with the aim to incorporate geodetic methodology in it, in order to increase its reliability. In the same year he started working at the Remote Sensing Technology Institute of the German Aerospace Center (DLR) as a project scientist to develop algorithms for this technique, and to integrate them in the existing operational radar interferometric processing chain. At present, he is with Vexcel Canada Inc., located in Ottawa, Canada. His research interests are the application of Persistent Scatterer techniques of radar interferometry, the integration with optical data and Geographical Information Systems (GIS), and structured software development.

Nomenclature

List of acronyms

am	Ante meridian
APS	Atmospheric Phase Screen
APSA	Advanced Permanent Scatterer Analysis
ASCII	American Standard Code for Information Interchange
C	Computer language
C++	Computer language
C-band	Frequency band with wavelength ~6 cm
COSMO–SkyMed	Italian next generation radar satellite constellation
CPU	Central Processor Unit
csh	C-shell (a UNIX shell)
CVS	Concurrent Versions System
cycle	Normalized phase difference ($\phi \,/\, (2\pi)$)
D/A	Digital/Analog
DEM	Digital elevation model
DIA	Detection Identification Adaption alternative hypotheses testing procedure
DLR	Deutsches Zentrum für Luft und Raumfahrt e.V. (German Aerospace Center)
DLR-IMF	DLR-Institut für Methodiek der Fernerkundung (Remote Sensing Technology Institute)
DSM	Digital Surface Model
DTM	Digital Terrain Model
ENVISAT	Environmental Satellite (ESA)
ERS–1	First European Remote Sensing satellite (ESA)
ERS–2	Second European Remote Sensing satellite (ESA)
ESA	European Space Agency
EW	East-West
FORTRAN	FORmula TRANslation (computer language)
fringe	Phase difference of 2π

GENESIS	Generic SAR Interferometric Software
GFZ	GeoForschungsZentrum
GPS	Global Positioning Satellite
IDL	Interactive Data Language (prototyping software)
IGARSS	International Geoscience and Remote Sensing Symposium
ILS	Integer Least-Squares
InSAR	Interferometric Synthetic Aperture Radar
i/o	Input/output
ITT	Invitation To Tender
JERS	Japanese Earth Resource Satellite
LAMBDA	Least-squares AMBiguity Decorrelation Adjustment
LandSat	USA satellite with thematic mapper (visible and infra red)
Linux	Free operating system similar to UNIX (not an acronym)
MCF	Minimal Cost Flow
MDB	Minimal Detectable Bias
NS	North-South
OMT	Overall Model Test statistic
PAMIR	Phased Array Multifunctional Imaging Radar
PC	Personal Computer
PCF	Program Control File
POLIMI	Politecnico di Milano (Technical University of Milan, Italy)
PS	Permanent Scatterer
PSC	Permanent Scatterer Candidate
PSI	Persistent Scatterer Interferometry
PSIC4	Persistent Scatterer Interferometry Codes Cross-Comparison and Certification for long term differential interferometry
RADAR	Radio detecting and ranging
RADARSAT	Canadian radar satellite
RCS	Radar Cross Section
rms	Root mean square
SAR	Synthetic Aperture Radar
SCR	Signal to Clutter Ratio
SLC	Single Look Complex
SNR	Signal to Noise Ratio
SPSA	Standard Permanent Scatterer Analysis
SRF	SUN Raster File
SRTM	Shuttle Radar Topography Mission
STUN	Spatio-Temporal Unwrapping Network
SUN	Stanford University Network (SUN Microsystems)
TerraSAR–X	German radar satellite

TRE	Tele-Rilevamento Europ (a POLIMI spin-off company)
UNIX	A standard computer operating system (not an acronym)
USGS	United States Geological Survey
UTC	Universal Time Coordinated (Greenwich Mean Time)
UTM	Universal Transverse Mercator (map projection)
vc-matrix	Variance-covariance matrix
WGS84	World Geodetic System 1984 (ellipsoid)
X-band	Frequency band with wavelength ~3 cm
y	Year
°C	Degree Celcius
cm	Centimeter
GB	Giga byte
Hz	Hertz
km	Kilometer
m	Meter
mm	Millimeter
rad	Radian
TB	Tera Byte

List of symbols

$'$	minute
$^\circ$	degree
$\in$	Element of
$[a, b)$	Half-open interval $\{x \vert a \leq x < b\}$
$\otimes$	Kronecker tensor product
$\delta_{l,0}$	Kronecker symbol ($\delta_{l,m}=1$ for $l = m$, 0 otherwise)
$\Im$	Imaginary part
$\Re$	Real part
j	Imaginary unit ($j^2 = -1$)
$\angle c$	Phase of complex number c
$\vert . \vert$	Absolute value; Amplitude of complex number
$\Vert . \Vert$	Norm
$\{.\}^{-1}$	Inversion
$\{.\}^*$	Transposition
$C\{.\}$	Covariance
$D\{.\}$	Dispersion
$E\{.\}$	Expectation
$\exp(p)$	Irrational number e (2.71828...) to power p
$\ln(p)$	Natural logarithm $\log_e p$

trace(.)	Trace
$W\{.\}$	Wrap
ξ	Azimuth coordinate
η	Range coordinate
λ	Wavelength of carrier signal; Non-centrality parameter
γ	Complex coherence; Power of test
σ^o	Normalized radar cross section
$\Delta\sigma$	Radar cross section
θ_x^k	Local look angle (viewing angle, off-nadir angle)
$\theta_{x,\mathrm{inc}}^k$	Local incidence angle
ϑ_x^k	Local squint angle
$B_{\perp x}^k$	Local perpendicular baseline
D_a	Amplitude dispersion index
$f_{x,\mathrm{dc}}^k$	Local Doppler centroid frequency
H_{amb}	Height ambiguity
k	Sensor in orbit, corresponding position, and/or time
m	Master sensor; Number of observations
r_x^k	Geometric slant range from sensor k to point x
T^k	Temporal baseline
v	Sensor velocity
x	Observed point
α_d	Coefficient of displacement base function d
β_x^k	Local height-to-phase conversion factor
Δh_x	DEM error (height above reference surface)
$\Delta r_x^{1,0}$	Line-of-sight deformation at x in the time interval $t_0 - t_1$
$d(T, x, y)$	Spatio-temporal displacement function
$f(T)$	Temporal displacement function
$g(x, y)$	Spatial displacement function
H	Number of PSC points
K	Number of interferograms
$p_d(t)$	Displacement base function
S_x^k	Atmospheric delay
φ	Phase in SLC image
ϕ	Wrapped interferometric phase
Φ	Unwrapped interferometric phase
ϕ_{atmo}	Phase induced by atmospheric delay
ϕ_{topo}	Interferometric phase caused by (uncompensated) topography
ϕ_{defo}	Interferometric phase caused by displacement

ϕ_{noise}	Random noise phase
ϕ_{obj}	Phase caused by sub-pixel position of the scatterer
ϕ_{orbit}	Phase caused by orbit errors
ρ	Correlation matrix
$\underline{\hat{\sigma}}^2$	A posteriori variance factor
$\sigma^2_{\text{atmo}}(0)$	Variance of atmospheric phase
σ^2_{noise}	Variance of random noise phase
$C(l)$	Covariance function
l	Distance
l_c	Correlation length
Q_{atmo}	Vc-matrix of atmospheric phase
Q_{ifg}	Vc-matrix of interferometric phase
Q_{noise}	Vc-matrix of noise
Λ	Transformation matrix for temporal differences
Ω	Transformation matrix for spatial differences
A	Design matrix for integer parameters of the partioned model of observation equations
$\bar{A}$	Reduced design matrix ($\bar{A} = P_B^{\perp} A$)
B	Design matrix for the model of phase observation equations
C	Design matrix describing a spatial network
e_K	Vector $(1, 1, \ldots, 1)^*$ of length K
E_K	Matrix of dimension $[K \times K]$ filled with ones
i_K	Vector $(0, \ldots, 1, 0, \ldots, 0)^*$ of length K with single 1 at position i.
I	Identity matrix
O	Zero matrix
P	Permutation matrix
$P_B^{\perp}$	Least-squares orthogonal projection matrix
Q_k	Co-factor matrix of variance component model
Z	Z-transformation matrix for decorrelation
χ^2	Bound on search space for ILS; Chi-squared distribution
a	Vector of unknown integer ambiguities
$\underline{\hat{a}}$	Float solution for ambiguities
$\underline{\check{a}}$	Integer solution for ambiguities
b	Vector of unknowns
$\underline{\hat{b}}$	Vector of adjusted unknowns
$\underline{\check{b}}$	Vector of adjusted unknowns using $\underline{\check{a}}$
e	Vector of measurement noise
$\underline{\hat{e}}$	Vector of adjusted residuals

$\underline{\check{e}}$	Vector of unwrapped residuals
m	Number of observations; Master sensor
n	Number of estimated parameters
$Q_{\hat{e}}$	A posteriori vc-matrix of least-squares residuals
$Q_{\hat{b}}$	A posteriori vc-matrix of estimated parameters
Q_y	A priori vc-matrix of observations
$Q_{\hat{y}}$	A posteriori vc-matrix of adjusted observations
r	Redundancy $r = m - n$
$\underline{y}$	Vector of observations
$\underline{\hat{y}}$	Vector of adjusted observations
$\underline{\check{y}}$	Vector of adjusted unknowns using $\underline{\check{a}}$
$\underline{y}^R$	Observable that is functionally or stochastically related to another set of observables
$\underline{\check{z}}$	Integer solution for Z-transformed unknown parameters
$\underline{\hat{z}}$	Float solution for Z-transformed unknown parameters
α	Level of significance of a test (probability that the null-hypothesis is rejected when it should not have been)
β	Probability that the null-hypothesis is not rejected when it should have been
γ	Power of a test ($\gamma = 1-\beta$); Complex coherence
λ	Non-centrality parameter of chi-squared distribution; Wavelength
χ^2	Chi-squared distribution; Bound on search space
$\chi^2_{\alpha(q)}$	Critical value for test of dimension q with level of significance α
∇	Vector of additional parameters of length q
C_q	Matrix specifying alternative hypothesis of dimension q
H_0	Null-hypothesis
H_A	Alternative hypothesis
$P(\underline{\hat{z}} = z)$	Success rate
$\underline{T}_q$	Test statistic of dimension q

Index

Remote Sensing and Digital Image Processing

1. A. Stein, F. van der Meer and B. Gorte (eds.): *Spatial Statistics for Remote Sensing.* 1999 ISBN: 0-7923-5978-X
2. R.F. Hanssen: *Radar Interferometry.* Data Interpretation and Error Analysis. 2001 ISBN: 0-7923-6945-9
3. A.I. Kozlov, L.P. Ligthart and A.I. Logvin: *Mathematical and Physical Modelling of Microwave Scattering and Polarimetric Remote Sensing.* Monitoring the Earth's Environment Using Polarimetric Radar: Formulation and Potential Applications. 2001 ISBN: 1-4020-0122-3
4. F. van der Meer and S.M. de Jong (eds.): *Imaging Spectrometry.* Basic Principles and Prospective Applications. 2001 ISBN: 1-4020-0194-0
5. S.M. de Jong and F.D. van der Meer (eds.): *Remote Sensing Image Analysis.* Including the Spatial Domain. 2004 ISBN: 1-4020-2559-9
6. G. Gutman, A.C. Janetos, C.O. Justice, E.F. Moran, J.F. Mustard, R.R. Rindfuss, D. Skole, B.L. Turner II, M.A. Cochrane (eds.): *Land Change Science.* Observing, Monitoring and Understanding Trajectories of Change on the Earth's Surface. 2004 ISBN: 1-4020-2561-0
7. R.L. Miller, C.E. Del Castillo and B.A. McKee (eds.): *Remote Sensing of Coastal Aquatic Environments*. Technologies, Techniques and Applications. 2005 ISBN: 1-4020-3099-1
8. J. Behari: *Microwave Dielectric Behaviour of Wet Soils*. 2005 ISBN 1-4020-3271-4
9. L.L. Richardson and E.F. LeDrew (eds.): *Remote Sensing of Aquatic Coastal Ecosystem Processes*. Science and Management Applications. 2006 ISBN 1-4020-3967-0
10. To be published
11. To be published
12. B.M. Kampes: *Radar Interferometry*. Persistent Scatterer Technique. 2006 ISBN 1-4020-4576-X