You are cordially invited to the public defence of my PhD thesis

Geodetic network design for InSAR: Application to ground deformation monitoring

on Friday 22 May 2015 from 10:00 - 11:00 in the Senaatzaal at the Aula of the TU Delft

I will give a brief introduction to the thesis at 9:30

A reception will follow immediately after the ceremony

Pooja S. Mahapatra
Geodetic network design for InSAR
Application to ground deformation monitoring

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K. C. A. M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 22 mei 2015 om 10:00 uur

door

Pooja Suman Mahapatra

M.Sc., Space Science and Technology
geboren te Bangalore, India.
Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. R. F. Hanssen

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof. dr. ir. Ramon Hanssen, Technische Universiteit Delft, promotor
Dr. Peter Fokker, TNO
Prof. dr. Andrea Monti Guarnieri, Politecnico di Milano
Prof. dr. Timothy Dixon, University of South Florida
Prof. dr. ir. Jan-Dirk Jansen, Technische Universiteit Delft
Prof. Dr. -Ing. Roland Klees, Technische Universiteit Delft
Prof. ir. Peter Hoogeboom, Technische Universiteit Delft

Keywords:
Satellite radar interferometry, ground deformation, geodesy, network design, persistent scatterers, transponders, corner reflectors, precision, InSAR-GNSS collocation, datum connection

Printed by: Ipskamp Drukkers

Front & back: Adapted from Sentinel-1 radar vision, photo credit ESA/ATG Medialab.

Copyright © 2015 by Pooja Suman Mahapatra.

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without prior permission of the author.

Typeset by the author with the LATEX documentation system.

ISBN/EAN: 978-94-6186-474-1

An electronic version of this dissertation is available at http://repository.tudelft.nl/.
To my parents
Summary

For the past two decades, interferometric synthetic aperture radar (InSAR) has been used to monitor ground deformation with subcentimetric precision from space. But the applicability of this technique is limited in regions with a low density of naturally-occurring phase-coherent radar targets, e.g. vegetated nonurbanized areas. Third-party end-users of InSAR survey results cannot, in a systematic way, determine *a priori* whether these coherent targets have adequate spatial distribution to estimate the parameters of their interest. Additionally, InSAR deformation estimates are referred to a local datum, meaning that the technique is sensitive only to the relative deformation occurring within the SAR images. This makes it difficult to compare these estimates with those from other techniques, e.g. historical levelling data or changes in the sea level. Here we propose the **design of a geodetic network for InSAR**, aimed at densifying the naturally-occurring measurement network and converting from a local datum to a global one.

**A practical solution for improving spatial sampling** is to deploy coherent target devices such as corner reflectors or transponders on ground, tailored to the specific monitoring application under consideration. The proposed method (1) provides a generic description of any deformation phenomenon; (2) determines whether the naturally-occurring InSAR measurements are adequate in terms of user-defined criteria; (3) finds the minimum number of additional devices to be deployed (if required); and (4) finds their optimal ground locations. It digests, as inputs, any prior knowledge of the deformation signal, the expected locations and quality of the existing coherent targets, and the quality of the deployed devices. The method is based on comparing different covariance matrices of the final parameters of interest with a criterion matrix (i.e., the ideal desired covariance matrix) using a predefined metric. The resulting measurement network is optimized with respect to precision, reliability and economic criteria; this is demonstrated via synthetic examples and a case of subsidence in the Netherlands.

As a basis for the choice and number of deployed devices, we evaluate the measurement precision of **compact active transponders** and demonstrate their **viability as alternatives to passive corner reflectors** through three field experiments, using different satellite data and geodetic validation techniques. Transponders are shown to be usable for subcentimetre-precision geodetic applications, while improving upon the drawbacks of corner reflectors in terms of size, shape, weight and conspicuousness.

For **transforming the spatially-relative InSAR deformation estimates** (local datum) to a **standard terrestrial reference frame** (global datum), we introduce a new concept involving the collocation of transponders with Global Navigation Satellite System (GNSS) measurements. The displacement of such a transponder is consequently determined in the standard reference frame used by GNSS, eliminating the need for any assumptions on reference-point stability in applications where the InSAR deformation estimates are compared with results from other techniques. The considerations, results and practical lessons learnt at several permanent GNSS stations in the Netherlands are described.
De laatste twee decennia is interferometrische synthetische apertuur radar (InSAR) gebruikt om bodembeweging met sub-centimeter precisie te meten vanuit de ruimte. De toepassing van deze techniek beperkt zich echter tot gebieden met een voldoende hoge dichtheid van fase-coherente objecten. Hierdoor kunnen eindgebruikers van InSAR producten niet op voorhand bepalen of de ruimtelijke bemonstering door coherente reflecties voldoende is om de deformatieparameters voor hun specifieke probleem te bepalen. Bovendien zijn deformatieschattingen door InSAR gegeven binnen een lokaal geodetisch datum, wat inhoudt dat de techniek slechts de relatieve beweging binnen een SAR-beeld waarneemt. Dit maakt het onmogelijk om deze schattingen direct te vergelijken met die van andere meettechnieken, zoals historische waterpasmetingen, of met veranderingen van het zeeniveau. In deze studie introduceren we het ontwerp van een geodetisch netwerk voor InSAR, gericht op de ruimtelijke verdichting van het meetnetwerk, en de transformatie van een lokaal naar een wereldwijd geodetisch datum.

Een praktische oplossing voor de meetpuntverdichting is het zodanig installeren van coherente objecten, zoals radarreflectoren en transponders, dat voldaan wordt aan het programma van eisen voor de specifieke toepassing. De voorgestelde methode (1) geeft een algemene omschrijving van de bodembeweging; (2) bepaalt of de reeds aanwezige coherente radarreflecties voldoende zijn om aan de criteria van de gebruiker te voldoen; (3) zoekt (indien nodig) het minimum aantal benodigde aanvullende coherente objecten; en (4) vindt hun optimale locatie. Als invoer wordt gebruikt: alle beschikbare kennis van het deformatiesignaal, de verwachte locatie en kwaliteit van reeds aanwezige coherente objecten en de kwaliteit van de (eventuele) plaatsen coherente objecten. De metode is gebaseerd op het vergelijken van verschillende covariantiematrices van de gewenste parameters met een kriteriummatrix (d.w.z. de ideale gewenste covariantiematrix) gebruik makend van een vooraf gedefinieerde metriek. Het resulterende meetnetwerk wordt geoptimiseerd naar nauwkeurigheid, betrouwbaarheid en kosten; dit wordt aangetoond aan de hand van simulaties en een voorbeeld van bodemdaling in Nederland.

De keuze en het aantal gebruikte coherente objecten wordt gebaseerd op de meetnauwkeurigheid van compacte active transponders, waarvan middels drie experimenten wordt aangetoond dat ze een geschikt alternatief zijn voor passieve radarreflectoren. Transponders blijken bruikbaar voor geodetische toepassingen waarbij een sub-centimeter precisie wordt vereist. Bovendien hebben zij, in vergelijking met radarreflectoren, een voordeel door hun beperkte afmetingen, vorm, laag gewicht en onopvallendheid.

Om de relatieve deformatieschattingen (in een lokaal datum) naar een standaard terrestrisch referentiesysteem (wereldwijd datum) te transformeren, wordt een geïntegreerd meetstation geïntroduceerd. Hierbij wordt een radar transponder fysiek gekoppeld aan een Global Navigation Satellite System (GNSS) antenne. De beweging van de transponder is hierdoor bekend in het GNSS referentiesysteem, waardoor aannames over de stabiliteit van het referentiepunt overbodig worden. De afwegingen die hieraan ten grondslag liggen en de resultaten van experimenten op verschillende permanente GNSS-stations in Nederland worden uitvoerig beschreven.
InSAR is a potent satellite application. The technique dates back to 1974, when it was first introduced for topographic mapping. Accurately inferring crustal deformation using spaceborne InSAR was an unexpected but welcome outcome of the ERS-1 satellite, which, in 1992, captured the co-seismic displacement due to an earthquake. In addition to providing data over a wide scale and fine resolution under all conditions of cloud cover and solar illumination, SAR images, when acquired worldwide at regular intervals, can be used to measure deformation without advance knowledge of its occurrence. Despite these strengths, InSAR has taken around two decades to become an operational tool for monitoring ground deformation, with articles still being published on InSAR validation by comparison with other geodetic techniques such as GNSS (which have their shortcomings too). In this dissertation, I have tried to bring out a few limitations of InSAR and provide practical solutions to them, in the hope that InSAR will be used someday as a benchmark to compare other geodetic techniques with.

A little more than five years ago, as a fresh graduate of a Master’s programme in space science and technology, I wanted to move on to the domain of satellite applications. A chain of circumstances led me to meet Ramon Hanssen at the TU Delft, to interview for a PhD position involving geodesy and InSAR — two topics about which I had hardly any knowledge then. Ramon’s role as my guide and mentor started at that meeting. Like he did several more times in the following years, he kindled my desire to delve into practical problems and come up with simple solutions (following Occam’s razor). Ramon, thank you very much for providing me with inspiration, insightful discussions, encouragement and exposure-building opportunities. I have learnt a lot from you, not limited to academic knowledge.

Several other people have been instrumental, directly or indirectly, in my research. I will remain indebted to Sami Samiei-Esfahany (TU Delft); sharing an office with him has been an enlightening experience. He is one of the best teachers and most thorough researchers I know. Sami, despite your own busy schedule, you were always ready to share your ideas and give your honest opinions of mine. Thank you for all the brainstorms, philosophical discussions and motivation (as well as the late-night pizza company).

I would like to express my sincere gratitude to Hans van der Marel (TU Delft) for sharing his GNSS and geodetic expertise, and for managing the practical aspects of transponder installation, maintenance and troubleshooting. I also owe a lot to Freek van Leijen (TU
Delft), who was of immense help with his advice, especially in the last phase my PhD work. Ch. 5 of this dissertation benefitted from software developed by both of them. Additionally, Freek translated to Dutch the propositions accompanying this dissertation. Hans and Freek, thank you for answering my questions whenever I knocked on your doors or wrote you emails.

Being part of the InSAR research group (‘radar group’) at the TU Delft and participating in the weekly meetings have enriched my knowledge of this field well beyond my own research topic. I would like to thank everyone who is or was a part of the radar group. Very special thanks to my officemates Anneleen Oyen and Joana Martins, who also turned into my close friends and were with me through thick and thin. I will always cherish the moments we shared, both professionally and personally. Anneleen also translated the summary of this dissertation to Dutch. Mahmut Arıkan provided me with help and advice, not limited to software and IT, as well as good food. My discussions with Andy Hooper (now at the University of Leeds) have been very motivating, especially in the early phase when I was defining the research questions. I have also had the pleasure of being associated with David Bekaert, Miguel Caro Cuenca, Shizhuo Liu, Manu Delgado Blasco, Ling Chang, Prabu Dheenathayalan, Karsten Spaans, Peter Buist, Lennard Huisman, Bram te Brake, Saygin Abdikan, Raluca Ianoschi, Alexandru Lepadatu, Roel van Bree, Gertjan van Zwieten, Piers van der Torren, Naresh Soni, Siavash Shakeri, Davide Imparato, Yanqing Hou and several other researchers, guests and students. All these people have helped me at some point in time. Many were also involved in the corner reflector and transponder experiments; their help during deployment and geodetic levelling is much appreciated. Thanks also to Lorenzo Iannini, Ramses Molijn and Ali Mousivand for the fun (and scary) times.

I would like to acknowledge other colleagues within and outside the TU Delft, with whom I have had professional contact over the past years. Chats with Petar Marinkovic (PPO Labs) have been informative on a wide variety of topics not limited to NEST and Sentinel-1, sometimes over drinks or Chinese food. Roderik Lindenbergh clarified my queries regarding spatial sampling and interpolation. Massimo Menenti gave me interesting international teaching opportunities. Roderik Koenders (responsible for discussions ranging from Martian terrain to bitcoins), Marcin Glegola and Edouard Martins were great officemates at different times. The technical support of Pascal Aubry and René Reudink for laboratory and field testing of the transponders is much appreciated. Thanks go out to all the Geoscience and Remote Sensing (GRS) PhDs and post-docs for their enjoyable lunchtime company. The administrative support provided by Lidwien de Jong, Rebeca Domingo, Marjolein de Niet-de Jager, File Koot-Stomp and Danko Roozemond is gratefully acknowledged.

This PhD project was funded primarily by CATO2, the national R&D programme on CO\(_2\) capture, transport and storage in the Netherlands. I would like to thank Jan Brouwer, Jan Hopman, Wilfred Visser, Sander van Eghmond, Mirjam van Deutekom, Marlies Verlinde and others at CATO2 for their encouragement and support. I enjoyed attending CATO2 events and excursions, where I could interact with experts and fellow researchers work-
ing on topics as varied as geophysics, geology, chemistry, infrastructure planning, policy making and human psychology.

The EU project I2GPS (Integrated Interferometry and GNSS for Precision Survey, FP7-GALILEO-2008-GSA-1) also played an important role. In addition to funding, it gave me access to the compact active transponders that are central to this dissertation. Rachel Holley from NPA (then Fugro, now CGG) gave valuable advice during course of this project, and meticulously planned the experiment in Slovenia, described in Ch. 3. Marko Komac, Blaz Milanic and their colleagues from the Geological Survey of Slovenia provided practical support at the experiment site in Slovenia. And thanks to Alan Fromberg and Chris Prior from SEA (now Thales Alenia Space), the transponders could be retained for experimentation beyond I2GPS.

The feedback provided by the members of my defence committee (Timothy Dixon, Andrea Monti-Guarnieri, Peter Fokker, Roland Klees, Peter Hoogeboom and Jan-Dirk Jansen) has improved the quality of this dissertation significantly; their efforts are highly appreciated. I used several InSAR software tools in the course of my research, such as Doris (with all the embedded humour), DePSI, and StaMPS. I thank everyone who has put the time and effort into developing and maintaining these (open-source) tools. I also gratefully acknowledge my colleagues at SkyGeo, where I have been working for the last five months, for providing a dynamic, enjoyable and flexible working environment.

Heartfelt thanks go out to my friends, even if I cannot mention them all here. Prem Prasad, I have greatly enjoyed my time with you, often over after-hour drinks and faculty coffees (as well as with Steven Engelen, Arash Noroozi and Raj Thilak). The support of, and fun with, Sandra Schumacher and Bernie Jantscher are warmly appreciated and cherished. And finally, SR Sudarsan has been my rock. SR (and through you, Saranya), your listening ears, sane advice, humour and (travel) company have been indispensable; thanks for everything.

I am deeply grateful to my parents, Purnima and Pravas Mahapatra, who have inspired me at every stage, and been with me in all my decisions despite the thousands of kilometres that separate us. My parents-in-law, Annemarie and Helmut Siegl, have provided me with a second home, moral support and many relaxing holidays in the Austrian Alps — vielen Dank. Many thanks also to my brother, Satya, and brother-in-law, Werner.

There is one person who has chugged alongside me on this PhD journey, through all the hills and dales. He is my source of pragmatic ideas, unwavering support, selfless encouragement, unerring navigation, hearty laughter and amazing experiences. Despite living in different countries, every weekend has been special, thanks to the Dutch and German railways. He has also proofread large parts of this dissertation, and climbed up a few permanent GPS stations for me. Martin — thank you, very very much.

Pooja S. Mahapatra
Delft, April 2015
1

Introduction

Die Geodäsie ist derjenige Teil der Geometrie, in dem die Approximationsmathematik ihre klarste und konsequenteste Durchbildung gefunden hat. Man untersucht bei ihr ununterbrochen einerseits die Genauigkeit der Beobachtungen und andererseits die Genauigkeit der Resultate, die aus den Beobachtungen folgen.¹

Felix Klein

1.1. Motivation

Spaceborne repeat-pass Interferometric Synthetic Aperture Radar (InSAR) is an operational geodetic technique for monitoring ground deformation. However, InSAR measurements are sometimes sparse (especially in vegetated areas) and inhomogeneous, making it unclear to an end-user whether the spatial density and distribution are adequate for the monitoring application at hand, or whether to place additional coherent targets on the ground to ensure measurements where desired. In the latter case, the questions of which kind of target to place, and where to most effectively place the minimum number of targets, also arise. Moreover, the final InSAR-derived estimates only reflect the relative deformation within the radar images, making it difficult to compare them with other datasets such as those from historical levelling campaigns or sea-level changes. In this thesis, we address these limitations using the concepts of geodetic network design.

¹Geodesy is that specific branch of geometry where the mathematics of approximation have found their clearest and most consistent development. One constantly examines, on the one hand, the accuracy of observations, and on the other, the accuracy of the results following from these observations.
1.2. Background

Humankind has always been interested in learning about the earth. Along with astronomy, *geodesy* — the study of Earth's size, shape, orientation and gravitational field, and the variations of these quantities over time ([1–3], extended from the classical definition of F. R. Helmert in 1880) — is one of the most ancient branches of science. For several centuries, information about the geometry of the earth was collected through astronomical observations of celestial bodies such as the sun, moon, planets and stars. Pioneering work by explorers and cartographers after the Middle Ages prompted rapid expansion in geographical knowledge, and terrestrial geodetic measurements (angles and distances) emerged as viable tools for relative positioning. The associated geodetic tasks such as triangulation, levelling, and the use of theodolites stimulated the best brains of the era [4]. Several contributions to the fields of mathematics and physics in the eighteenth and nineteenth centuries resulted from this, notably the least-squares method, the mechanics of physical bodies, the definition of the geoid and potential theory, differential geometry, analytical mechanics, as well as the application of electromagnetic waves to long distance measurements. In the mid-twentieth century, accurate and commercial electromagnetic distance measurement devices, based progressively on polarized light, radio waves, and finally lasers, had lasting effect on geodetic philosophy — horizontal angles which were measurable to much higher accuracy for centuries gave way to precise, easily- and remotely-operable distance measurements [4]. 'Radio detection and ranging' (radar) systems were also developed in this period, initially used primarily for military purposes, but quickly adopted for civil and scientific applications. The capabilities of radar were twofold: the two-way travel time of the electromagnetic pulse could be used to determine the range to an object, and the backscatter intensity used to infer physical characteristics of the object (e.g. surface roughness or size). These technological improvements were further aided by the contemporary invention of electronic computers, which facilitated complex numerical calculations that were impossible in the past.

The next quantum jump in geodesy was achieved in the latter half of the twentieth century, when artificial satellites were launched into space. The intervisibility barrier for accurate point positioning was surmounted, the earth's gravity field could be studied and mapped, and inertial navigation and positioning systems could be developed. The advent of space geodesy therefore opened up a large variety of applications, by providing observations at different spatial and temporal scales. Radar ranging from space took on a new dimension in the late 1980s, with the launch of Synthetic Aperture Radar (SAR) satellites and the development of InSAR [5], which has by now emerged as a powerful technique to measure the earth's topography and crustal deformation. Spaceborne InSAR in repeat-pass configuration, where the SAR antenna images the same ground area at specific time intervals, has enabled the observation of dynamic processes such as the surface displacement effects of earthquakes and tectonic motion, volcanism, anthropogenic ground subsidence or uplift (e.g. due to extraction or injection of oil, gas or water), groundwater flow, landslides and glaciers. InSAR possesses the unique capability of metre-level spatial resolution, temporal sampling of the order of days to weeks, and wide-area coverage distributed over the entire earth. In contrast with optical tech-
niques, InSAR can extract information from data acquired in the nighttime and through cloud cover. Large regions of the earth can also be studied retrospectively, with archived SAR data available since 1992. There are several SAR satellites orbiting the earth today, with the prospect of many more in the coming decades, making InSAR an operational technique for continuous deformation mapping of the earth.

Conventional InSAR is based on interferograms, which are complex-valued images where the value of each pixel is related to the difference in range distances of a terrain element (resolution cell) to the SAR sensor at two different acquisition times. This difference is a superimposed effect of different phase components, predominantly from target displacement, topography and atmospheric delay. For deformation monitoring applications, errors are introduced by the spatio-temporal variability of the atmosphere, temporal decorrelation (i.e. changes in the scattering properties within a resolution cell over time) and geometric decorrelation (i.e. incomparable scattering characteristics arising from different satellite viewing angles) [6, 7]. Conventional InSAR using single interferograms therefore progressed to time-series InSAR, where a stack of interferograms of an area is utilized to minimize the effect of these errors.

One of the earliest time-series InSAR approaches, Persistent Scatterer Interferometry (PSI) [8–12], is based on radar targets, often man-made, that show stable phase behaviour over long time intervals, i.e. with minimal temporal decorrelation. These persistent scatterers (PS) dominate the reflection within a resolution cell, thereby also reducing the effect of geometric decorrelation. By using a stack of interferograms constructed with one SAR image in common (i.e. a single master image), the temporally-uncorrelated atmospheric phase contribution is estimated and removed. The term coherence is used as a measure of the accuracy of the interferometric phase; PSI therefore focuses on a subset of coherent targets to derive displacement time-series. This technique yields a high density of measurements in urbanized areas and in other regions containing radar scatterers that are coherent over long periods of time, such as rock outcrops. Applying PSI on rapidly decorrelating areas such as wetlands or regions with vegetation or snow cover, however, yields few PS.

An approach towards increasing the spatial density of measurements is to extract information not only from coherent point-like targets, but also from distributed scatterers (DS) [13–20]. These are resolution cells, such as desert areas or cultivated lands with short vegetation, that are subject to some amount of temporal and geometrical decorrelation, but from which coherent information can still be extracted. The Small Baseline Subset (SBAS) technique does this by making use of subsets of interferograms within the whole stack which have small baselines, i.e. shorter distances (both in satellite position and in acquisition time) between the two SAR images involved in each interferogram. Owing to the small baselines and the operation of multilooking (equivalent to replacing a group of complex interferogram pixels by their spatial average) that is often performed, SBAS reduces both temporal and geometric decorrelation. In SqueeSAR [19], clusters of pixels with similar scattering characteristics are selected, which then have lower overall noise level than single pixels. Estimation of deformation in the time domain is further optimized using a coherence matrix, which contains information on the correlation between radar images in the stack. To reach from wrapped phase observations in the \([-\pi, +\pi]\) interval to estimates of physical ground deformation, all InSAR methods per-
form phase unwrapping by making use of assumptions on the spatial and/or temporal smoothness of the expected deformation signal. Such *a priori* knowledge (APK) of the deformation can be expressed e.g. as a temporally linear or seasonal deformation rate, or as a limit on the maximum expected deformation between neighbouring pixels with a phase-equivalent less than \(\pi\). It is not intended here to provide a detailed description of (time-series) InSAR processing methodologies, for which dedicated books and articles are recommended, e.g. [6, 9–23].

InSAR algorithms are being aided by the newer generation of SAR satellite missions such as Sentinel-1 and TerraSAR-X, with shorter repeat cycles and narrower orbital tubes compared to the earlier generation of satellites such as ERS and Envisat, which further limit the temporal and geometrical decorrelation between acquisitions [24–26]. Even so, a limitation of the InSAR technique is that coherent PS and DS are not guaranteed to be present everywhere. There are still areas that heavily decorrelate between SAR acquisitions, such as forests, from where no coherent information can be extracted. This could result in a sparse and inhomogeneous spatial distribution of measurements, and it is not clear to an end-user if such areas can be monitored using these alone. If the user decides to densify the measurements by deploying passive or active coherent targets on ground, the most effective locations of the minimum number of targets for the monitoring application at hand still need to be determined.

A second property of InSAR is that the deformation estimates are inherently relative, i.e., the estimates are always with respect to a reference point or area within the SAR images under consideration. The technique cannot discriminate between the deformation at a certain location or at the reference point/area. For some applications, there is need for estimates which can be considered to be ‘absolute’, i.e., which can be expressed in a standard reference frame and related to the results of other techniques. This entails datum connection; the precise connection of InSAR to a global datum would enable tying overlapping or non-overlapping radar-derived deformation datasets together, yielding datasets comparable over large distances, even across oceans and continents. As a possible application, ‘absolute’ InSAR scatterer displacements would help to link ground deformation with changes in the sea-level at a global scale, thereby contributing valuable information towards flood risk assessment.

This thesis proposes practical methods based on the theory of geodetic network design and optimization, towards the planned densification of InSAR measurements in heavily decorrelating areas, as well as towards determining ‘absolute’ InSAR scatterer displacements.

1.3. Problem formulation and research objectives

The general problem statement addressed in this study is

*How can an optimal geodetic network be designed for InSAR, towards the goals of measurement densification in decorrelating areas, and integration of the deformation estimates into a standard terrestrial reference frame?*

From the perspective of third parties (end-users) who make decisions based on land survey results, the InSAR technique is of little use unless they are guaranteed to have
dense-enough measurements to estimate their final parameters of interest, e.g. ground displacements or deformation model parameters. A practical approach to densify InSAR measurements is to deploy in situ devices that behave as coherent targets, such as passive corner reflectors [27, 28] or active transponders [29]. But there is need for a systematic method to determine whether the standalone InSAR measurement density will be sufficient, and if not, how many additional devices to deploy and where to optimally locate them. Clearly, increasing the number of devices would lower the parameter estimation error, but would also increase the costs involved. There would also be significant value-addition for some applications if these additional coherent targets, whose locations are under our control, could be utilized towards connecting InSAR estimates to a well-defined terrestrial reference frame.

Therefore, more specific research questions are derived that give further direction to the treatment in this study:

1. Passive devices such as corner reflectors have been used in the past for InSAR measurement network densification. However, they suffer from drawbacks related to their large size and weight, conspicuousness, and loss of reliability due to geometric variations as well as material and maintenance-related degradation over several years of deployment. As an alternative, low-cost active radar transponders have been developed recently, which are smaller, lighter, and less conspicuous. Are such transponders a viable alternative to corner reflectors in terms of the measurement precision achieved?

2. Prior to device deployment, are the existing InSAR measurements adequately distributed for the monitoring application at hand? If not, what is the minimum number of additional devices (corner reflectors or transponders) required to be deployed? From a geodetic point of view, what are the optimal ground locations of these devices?

3. How can InSAR displacement estimates be precisely connected to a standard terrestrial reference frame (TRF)?

1.4. Thesis outline and research methodology

The current chapter provides a brief flashback into the fields of geodesy and InSAR, and motivates this thesis. The main research objectives are formulated as a general problem statement, divided into three specific questions. In order to answer these questions, the problem has been cast in the classical framework of geodetic network design [30–35], which has been applied in the past to other survey techniques such as levelling and Global Navigation Satellite Systems (GNSS). According to this framework, four classes of sub-problems are defined as first-order design (FOD, finding the optimum measurement locations), second-order design (SOD, deciding which observations to make and with what precision), third-order design (THOD, improving an existing network), and zero-order design (ZOD, defining a reference frame or datum).

Not all of these definitions are entirely applicable to InSAR, owing to an important difference of InSAR from, for example, levelling and GNSS: InSAR measurements are op-
portunistic in nature, in the sense that their inherent locations and precisions are not under our control. Only the measurements arising from deployed devices such as reflectors and transponders can be designed. The geodetic network design problem has therefore been modified for InSAR as SOD, combined FOD-THOD and ZOD. We attempt to provide practical solutions to these sub-problems in the following chapters. The flowchart in Fig. 1.1 outlines the methodology proposed in this thesis. Also indicated are the chapters in this thesis where the various blocks of this flowchart are covered.

**Chapter 2** is a brief review of the framework of geodetic design and optimization of deformation monitoring networks. The four relevant sub-problems (design orders) are explained, along with the general objectives of precision, reliability and cost.

**Chapter 3** deals with the SOD problem. In case of InSAR, the observations stem from naturally-occurring coherent targets, which are of predefined quality not under our control. However, determining the quality of InSAR observations, both naturally-occurring and deliberately-deployed, is an important step in the network design process. This chapter reviews the quality description of naturally-occurring coherent targets, and focuses on the measurement quality of the deployed devices, i.e. corner reflectors and transponders. Their respective advantages and limitations are discussed, and the geodetic applicability of low-cost transponders, which are relatively new technology, is verified. Since full-scale error propagation from the SAR sensor till the deformation estimates is not straightforward, we resort to empirical validation. Transponder precision estimates are derived through three field experiments performed under different conditions, by comparing transponder-InSAR deformation estimates from different SAR satellites with those from corner reflectors, levelling and GNSS.

**Chapter 4** covers a combination of FOD and THOD. FOD involves choosing the optimal measurement locations, which is redundant in InSAR owing to the inherent opportunistic nature. THOD is performed instead, where an existing network of InSAR measurements is improved or densified. At the core of the proposed algorithm is a comparison of different covariance matrices of the final parameters of interest with a criterion matrix (i.e. the ideal covariance matrix that is to be approximated by the design as best as possible), using a predefined metric. FOD is used to construct this criterion matrix. The combined method (1) provides a generic description of any deformation phenomenon, (2) determines whether the naturally-occurring InSAR measurements are adequate, (3) finds the minimum number of additional devices (if required) and (4) their optimal ground locations. The algorithm digests, as inputs, the expected locations and quality of existing coherent targets, the quality of the devices being deployed, and, if available, any prior knowledge of the deformation signal. The resulting network of devices is optimized with respect to precision, reliability and cost criteria. The methodology is demonstrated through several simulated case studies, as well as the real case of subsidence in Roswinkel, the Netherlands.

**Chapter 5** comprises zero-order design (ZOD), i.e. defining the results of InSAR surveys and studies in a standard TRF, including variance-covariance information.
For this, we propose a methodology based on collocated InSAR-GNSS measurements, achieved by rigidly attaching a transponder to a GNSS antenna. The InSAR displacement estimates are referred to this transponder, whose velocity and initial position are estimated in a standard TRF using the GNSS measurements. Consequently all the other deformation estimates are also defined in the same TRF. The interpretation of InSAR-derived deformation is then independent of any postulation on reference scatterer stability, and the estimates can be compared with other datasets also defined in the same TRF (e.g. historical levelling data or sea-level changes). We demonstrate ZOD through a simulated example and a practical case study in IJmuiden, the Netherlands. A collocation feasibility map of all the permanent GNSS stations in the Netherlands is also presented, based on the characteristics of SAR amplitudes at and in the vicinity of the stations.

Chapter 6 reports the general conclusions of this thesis and provides recommendations for further research.
Specify area of interest (AOI) and duration of monitoring

Acquire corresponding SAR data stack

Are absolute deformation estimates (in a standard reference frame) desired?

- No
  - Use all available information to build a functional and/or stochastic model of expected deformation
  - Make reasonable assumptions on the smoothness / variability of expected deformation
  - Determine probable coherent target locations within AOI
  - Acquire corresponding SAR data stack

- Yes
  - Collocate a transponder with the GNSS station using a sturdy mechanical connection

Are additional coherent targets required?

- No
  - Add user-defined monitoring quality requirements

- Yes
  - Determine probable coherent target locations within AOI
  - Make reasonable assumptions on the smoothness / variability of expected deformation
  - Use all available information to build a functional and/or stochastic model of expected deformation

Are additional coherent targets required?

- No
  - Specify area of interest (AOI) and duration of monitoring

- Yes
  - Determine probable coherent target locations within AOI
  - Acquire corresponding SAR data stack

Are 'absolute' deformation estimates (in a standard reference frame) desired?

- No
  - Specify area of interest (AOI) and duration of monitoring

- Yes
  - Determine probable coherent target locations within AOI
  - Acquire corresponding SAR data stack

Figure 1.1: Simplified schematic summary of the geodetic network design methodology presented in this thesis. Light grey blocks denote the general steps of the algorithm, and dark grey blocks the input information. The yellow blocks are the final outputs. Also indicated in pink, green and blue are the chapters of this thesis, where the enclosed blocks are described. The abbreviation 'I2GPS' corresponds to a collocated transponder-GNSS unit, comprising a radar transponder and a GNSS antenna on a common baseplate [36, 37].
Geodetic network design may be defined as the design of a geodetic measurement configuration, before any actual measurements are performed, that will satisfy preset requirements on the estimation quality of the final parameters of interest, as well as on the subsequent (geophysical) interpretation of the results, with minimum or acceptable cost [31–33]. In the terminology of the design orders introduced in Ch. 1, the first stage involves determining the optimal configuration (location and type) and distribution (spatial extent, density, and relative positioning) of measurements within the network, and encompasses FOD, SOD and THOD. The second step involves deciding on an optimal datum using all available information and valid assumptions, and consists of ZOD [30, 31].

This chapter presents a brief review of the framework of geodetic design for deformation monitoring networks as relevant to InSAR. We start with the general optimality criteria of precision, reliability and economy, followed by the network design sub-problems (design orders). In the treatment of each topic, we begin with general geodetic concepts, and then proceed to the relevance or adaptation for the case of InSAR. This review is not intended to comprehensively cover all aspects of geodetic network design, but to provide the proper reference needed for applying these concepts in the context of InSAR. For further reading, [30–35, 39] are recommended.

Parts of this chapter have been published in the IEEE Transactions on Geoscience and Remote Sensing 53, 7 (2015) [38].
2.1. Optimality criteria

In classical geodesy, where directions and distances were the main observables, surveyors would design the measurement network prior to going out to the field; from the pioneering work in [40], it was recognized that for a given network configuration, parameter precision estimation was possible even before any observations were made. Network design would involve decisions on the required number and locations of measurement points, which measurement technique to use, and with what precision to make the measurements (since the effective measurement precision could be improved by increasing the number of observations).

Later on, networks were designed for deformation monitoring, where, in addition to the geometrical strength of the network, the requirements of the subsequent (geophysical) interpretation of the survey results had to be accounted for [33]. Solving for all these design aspects in a single analytical procedure was difficult, if not impossible, and therefore a priori planning was usually done by network simulation and comparing network versions differing in their configurations and measurement precisions. To find the ‘optimal’ network, optimality criteria had to be defined; the optimal network would then be the network version that satisfied these criteria with minimum or acceptable cost. In addition, the optimal network would help to identify and eliminate observational outliers (gross errors), as well as minimize the effect of undetectable gross errors [33, 41].

The optimality criteria for a deformation monitoring network are broadly characterized by precision, reliability and economy (or minimum/acceptable cost) [31–33].

2.1.1. Precision

Precision refers to the dispersion of a stochastic variable around its expectation value (i.e. the second-order central moment), and concerns the nominal system performance or quality. Precision is given and described by a variance-covariance matrix (VCM), a matrix whose \((i, j)\)th element is the covariance between the \(i\)th and \(j\)th elements of a random vector. The main diagonal elements therefore refer to the variance of the corresponding random variables. A VCM expresses the probability density function of a vector of (correlated) stochastic variables that are jointly normally-distributed.

In general, the functional and stochastic relationship between observations (measurements) and parameters of interest (e.g. geometrical/geophysical parameters or surface displacements) can be expressed by a Gauss-Markov model as [42, 43]:

\[
E\{y\} = F(x); D\{y\} = Q_y
\]

where \(y\) is the \(m\)-vector of observations, \(x\) the \(n\)-vector of parameters to be estimated, \(F(\cdot)\) the (non-linear) vector function from \(\mathbb{R}^n\) into \(\mathbb{R}^n\), \(E[\cdot]\) and \(D[\cdot]\) the expectation and dispersion operators respectively, and \(Q_y\) the \(m \times m\) VCM of observations which reflects their precision. An underlined vector indicates its stochastic character. The first part of this equation is known as the functional model, and the second part the stochastic model. \(Q_y\) can be propagated to the VCM of parameter estimates, \(Q_{\hat{x}}\), using a suitable scheme, e.g. weighted least-squares or Monte Carlo methods. Here we focus on the former (refer
to Appendix A), by considering a linear system, or a linearized version of (2.1), given by

$$E[y] = Ax; D[y] = Qy$$

(2.2)

where $A$ is the design matrix. $A$ (or $F(.)$ in Eq. (2.1)), the link between the observations and the parameters of interest, is derived from a priori information regarding the physics of the deformation process, which is a crucial (but often undervalued) requirement in the design of deformation monitoring networks. Paradoxically, geodetic network design needs some knowledge of the deformation itself; unexpected deformation processes will not be captured by the designed network [33, 44].

User-requirements on precision can be classified into (1) scalar functions of the elements of $Q_{\hat{x}}$, and (2) criterion matrices. Scalar precision measures derived from $Q_{\hat{x}}$, such as the norm, trace, maximum eigenvalue, spectral width or determinant, can serve as a global representation. However, they are rather coarse characteristics that lead to information loss, and do not control individual values of VCM elements, especially covariances. Their use in practice is therefore limited [32 –34]. Instead, a criterion matrix [31–35, 45] is used, which provides greater localized precision control. The criterion matrix, denoted here as $Q_{\hat{x}_{ref}}$, is an artificial VCM possessing an ideal structure, i.e., it represents the desired precision of the estimated parameters.

To choose the optimal network configuration from several possible alternatives, the VCM of estimates resulting from each configuration is compared to the criterion matrix; the VCM closest to the criterion matrix corresponds to the most optimal network configuration. Some measures of this ‘closeness’ or agreement are given in [31, 34, 35, 39, 45].

If $Q_{\hat{x}}$ and the criterion matrix $Q_{\hat{x}_{ref}}$ are two $n \times n$ VCMs that are symmetric, positive semi-definite and defined in the same datum, it was suggested in [34, 35] that the eigenvalues of $Q_{\hat{x}} Q_{\hat{x}_{ref}}^{-1}$ capture the difference in form of these two VCMs, $Q_{\hat{x}}$ and $Q_{\hat{x}_{ref}}$, completely. Accordingly, the two VCMs agree better if

$$\frac{\{\lambda_i(Q_{\hat{x}}, Q_{\hat{x}_{ref}})\}_{\max}}{\{\lambda_i(Q_{\hat{x}}, Q_{\hat{x}_{ref}})\}_{\min}} \rightarrow \text{smaller} \rightarrow 1$$

(2.3)

with the eigenvalues $\lambda_i(Q_{\hat{x}}, Q_{\hat{x}_{ref}})$ from $|AQ_{\hat{x}} - Q_{\hat{x}_{ref}}| = 0$. This was further analysed in [39], and extended into a distance metric. This metric gives the distance between the VCMs, defined as the sum of squared logarithms of the generalized eigenvalues [39], i.e.

$$d(Q_{\hat{x}}, Q_{\hat{x}_{ref}}) = \sqrt{n \sum_{i=1}^{n} \ln^2 \lambda_i(Q_{\hat{x}}, Q_{\hat{x}_{ref}})}.$$  (2.4)

The smaller the value of the distance metric, the more similar the VCMs are; e.g. if the two matrices are identical, the eigenvalues are all equal to 1, making the distance metric equal to 0. The reader is referred to [39] for the mathematical derivation as well as advantages of this metric.

**Criterion matrix for InSAR**

A diagonal VCM signifying uncorrelated parameter estimates can be used as a criterion matrix. For InSAR, however, this is unrealistic because of spatially correlated error sources, both in InSAR observations as well as unmodelled deformation, which are
propagated to the parameter estimates (a treatment on InSAR error sources is given in Sec. 3.1). Therefore, a full criterion matrix is constructed such that the desired variance thresholds are incorporated in the diagonal elements, with realistic covariances as off-diagonal elements. This will be further described in Sec. 4.2.

2.1.2. Reliability

The reliability of a network is defined as its ability to detect and resist model imperfections via statistical hypothesis testing [31–33, 41, 43, 45–50]. A common kind of imperfection which can be generalized for different models and applications, and against which a network should be reliable, is the presence of gross errors or outliers in the observations [31].

Reliability is classified into internal and external reliability [31–33, 41, 43, 46–50]. Internal reliability refers to the ability of the network to allow for detecting errors. A measure of this, the minimum detectable bias, can be expressed for gross errors as

$$|\nabla y_i| = \sqrt{\frac{\lambda_0}{(Q_y^{-1}R)_{ii}}}$$

where $$(.)_{ii}$$ denotes the $$(i, i)$$th element, and $$\lambda_0$$ the lower bound of the non-centrality parameter, which is a function of the predefined probability of false and true detection of gross errors (see Appendix A). The idempotent matrix $$R$$ is given by:

$$R = P_A^\perp = I - A(A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1}$$

where $$I$$ is the identity matrix. $$R$$ is equivalent to $$P_A^\perp$$, which is the orthogonal projection of $$y$$ onto the vector of residuals $$\hat{e} = y - \hat{y}$$ in the metric of $$Q_y$$, and contains full information on the network geometry with respect to the influence of observational errors onto the residuals. External reliability refers to the maximum effect of undetected gross errors on the estimator $$\hat{x}$$, and can be expressed as

$$\nabla \hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} |\nabla y_i|.$$  \hspace{1cm} (2.7)

Thus, reliability defined with respect to gross errors depends on network geometry and the quality of observations.

From (2.5) and (2.7), the matrix $$R$$ is key to expressing both internal and external reliability. The redundancy numbers $$r_i$$, which are the diagonal elements of $$R$$, express the sensitivity of the network to individual observations. They have values between 0 and 1; when $$r_i = 0$$, no gross error in the $$i$$th observation can be detected irrespective of how large it is, and when $$r_i = 1$$, any gross error in the $$i$$th observation can be detected. The range $$0.3 \leq r_i \leq 0.7$$ signifies good controllability of the observations [32, 41, 50, 51].

**Gross errors in InSAR**

Gross errors or outliers can occur in InSAR observations; examples are temporal phase unwrapping errors [9, 10, 52]. An optimal network that is designed for InSAR accounting for the reliability criterion (examples shown in Sec. 4.2) would therefore be robust against such outliers. Errors from spatial phase unwrapping techniques [53–56] that can manifest as regional biases over larger areas are not considered in this context.
2.1.3. Economy

Economy is expressed as the costs incurred in implementing the designed network. In classical geodetic field measurements, this was closely related to the observation programme; costs could be minimized by reducing the surveying frequency, length, or man-hours involved.

Economic criteria for InSAR

For InSAR, the only additional costs by deploying coherent target devices are those of the device hardware, their installation and maintenance. Therefore, the fewer the deployed devices, the more economical the network.

2.2. Network design orders

As mentioned in Sec. 1.4, the geodetic network design problem has traditionally been divided into four sub-problems, ordered by the temporal processing of a network, as [30, 31]:

1. **First-order design (FOD):** finding the optimum measurement locations (configuration).
2. **Second-order design (SOD):** deciding which observations to make and with what precision (observation weights), i.e. finding the requisite quality of observations such that the quality of estimated parameters is as close as possible to the desired quality.
3. **Third-order design (THOD):** improving an existing network (additional observations), which covers two aspects: the optimal improvement of a network (e.g. reinforcing parts of the that turn out to be weak concerning the desired quality), and the optimal design of a densification network.
4. **Zero-order design (ZOD):** defining a reference frame or datum for a free network, which is a network in which the internal shape is given by measurements of a relative nature, e.g. height differences in one dimension [30].

This classification can be visualized by considering the free elements in the least-squares adjustment (Appendix A) of Eq. (2.2). Denoting $Q^{-1}$ as the weight matrix, the fixed and free parameters of the different design orders are summarized in Tab. 2.1. A fixed parameter implies that the object of the optimization problem is formulated in terms of that parameter. Free parameters are parameters that are free to assume different values for different network realizations in the corresponding design order.

In Sec. 1.4, we also saw that not all of these design orders are entirely applicable to InSAR, owing to an important difference of InSAR from traditional geodetic techniques: InSAR measurements are ‘opportunistic’ in nature, and their inherent locations and precisions are not under our control. This is a big advantage of using InSAR for deformation monitoring; ideally, no manpower is required at the monitoring site for installing benchmarks or performing in situ measurements. For example, a high density of naturally-occurring coherent targets or persistent scatterers (PS) of good quality is present in urban areas, owing to infrastructure such as buildings, bridges, roads and lamp-posts, which are stable radar targets.
Table 2.1: Fixed and free parameters in network design [30, 31].

<table>
<thead>
<tr>
<th>Design</th>
<th>Fixed parameters</th>
<th>Free parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-order design</td>
<td>$Q_y^{-1}, Q_{\text{xref}}$</td>
<td>$A$</td>
</tr>
<tr>
<td>Second-order design</td>
<td>$A, Q_{\text{xref}}$</td>
<td>$Q_y^{-1}$</td>
</tr>
<tr>
<td>Third-order design</td>
<td>$Q_{\text{xref}}$</td>
<td>$A, Q_y^{-1}$ (partly free)</td>
</tr>
<tr>
<td>Zero-order design</td>
<td>$A, Q_y^{-1}$</td>
<td>$\hat{x}, Q_{\hat{x}}$</td>
</tr>
</tbody>
</table>

Only the measurements arising from deployed devices such as reflectors and transponders, as well as the datum, can be designed for InSAR. Therefore, the primary design orders for InSAR are THOD (network densification) and ZOD (datum definition). FOD cannot be directly performed, but forms part of THOD, as will be shown in Ch. 4. SOD can, strictly speaking, not be performed; however, we term the accurate determination of $Q_y^{-1}$ as SOD for InSAR. $Q_y^{-1}$ forms an important input to the combined FOD-THOD step. The geodetic network design orders have therefore been modified for InSAR as SOD, combined FOD-THOD and ZOD, and are performed in that order.

**2.2.1. Second-order design for InSAR**

The InSAR measurements that stem from naturally-occurring coherent targets are of predefined quality that is not under our control. Using the notation of Tab. 2.1, the weight matrix $Q_y^{-1}$ is the free parameter for SOD, which is not possible in InSAR because the quality of the opportunistic PS observations cannot be designed. However, determining $Q_y^{-1}$ is an important step, and not a straightforward task because of the dependence of measurement precision on the physical properties of the radar scatterers, as well as various assumptions made during InSAR processing (see Appendix B). The quality of the deployed devices also affects $Q_y^{-1}$, and is estimated through empirical methods. Clearly, the less precise the device, the more the number of them required to be deployed. SOD for InSAR therefore involves the accurate quality description of the InSAR measurements, both naturally-occurring and deliberately deployed. Ch. 3 reviews InSAR error sources and the quality description of naturally-occurring coherent targets, as well as estimates the quality of deployed devices, particularly active radar transponders.

**2.2.2. Combined first- and third-order design for InSAR**

As mentioned before, the opportunistic nature of InSAR implies that the locations of naturally-occurring measurements are not under our control; some areas, especially if rural and vegetated, may yield a low measurement density. Several studies have assessed the factors affecting InSAR measurement density, such as [57–62]). Referring to Tab. 2.1, FOD for InSAR would make the design matrix $A$ a free parameter, which is not possible because the locations of the opportunistic PS observations cannot be designed. FOD is thus not directly applicable InSAR. However, FOD is used in THOD for constructing the criterion matrix $Q_{\text{xref}}$, which is the VCM of parameter estimates arising from an ‘ideal’ In-
SAR measurement network, regardless of where the naturally-occurring measurements are actually located (described in Ch. 4). We therefore term this design order for InSAR as combined FOD-THOD.

THOD for InSAR relates to the improvement or densification of an existing network of naturally-occurring coherent targets, by deploying additional devices if and where required. The improvement of an existing network could refer to reinforcing weak parts, e.g., if measurements are of poor quality. The design of a densification network pertains to areas that have insufficient InSAR measurements to estimate the parameters of interest with the desired quality. Referring to Tab. 2.1, both $A$ and $Q^{-1}$ are partly free parameters, meaning that only the location and quality of deployed devices (and not of the naturally-occurring measurements) can be altered.

The simulation method of solution

A complex design problem such as combined FOD-THOD can be solved in one of two ways: analytical or simulation methods [31, 33, 49, 63]. The former solves the design problem by a unique series of mathematical steps, and produces results optimal in a mathematical sense. The latter first postulates an initial network configuration solution to the problem, and the precision and reliability criteria are computed. If this solution does not satisfy the criteria requirements, the configuration is altered and the process is iterated until a satisfactory network is found. While the true absolute optimum in a mathematical sense may never be achieved, this does not usually have practical importance [32]. Also, simulation methods may find secondary optimal solutions to the problem, if the absolute optimum is restrictive or impractical in any way [31]. Another strong advantage of simulation methods is the flexibility to consider all objectives quasi-simultaneously, i.e. any arbitrary kind of precision or reliability criteria can be used and compared to reach to a solution. This contrasts with analytical methods that require these criteria to be in a (complicated) mathematical form, especially since geometry as a variable is difficult to deal with. Most analytical schemes using this variable have therefore turned out to be too restrictive for practical use [32]. Correct formulation of the mathematical model can be tedious, and the analytical solution can therefore have comparable computational cost as the simulation solution.

The simulation method of controlled variation is termed sequential optimization, and can be reduced to two approaches [30, 31]. Network construction starts with a minimal configuration of additional measurements, which is extended step-wise by measurements that contribute the most to the objective function, until the optimality criteria are met. Network reduction, on the contrary, starts with a maximal configuration of all possible measurements, and sequentially eliminates measurements that have the smallest influence on the objective function, until the lower bound of the user-defined criteria is attained. For combined FOD-THOD applied to InSAR, we choose the first approach. We first determine whether the naturally-occurring InSAR measurements are adequate. In case of inadequacy, the minimum number of additional devices to be deployed and their optimal ground locations are found by comparing different VCMs of the final parameters of interest (resulting from network construction) with the criterion matrix, using the distance metric of Eq. (2.4).
2.2.3. Zero-order design for InSAR

After SOD and combined FOD-THOD are performed, an InSAR measurement network of sufficient density and optimal spatial configuration has been designed. This network is foreseen to be able to estimate the parameters of interest with the desired quality. A stack of SAR data is then acquired and time-series InSAR processing is performed. The resulting InSAR deformation estimates are relative, because the first interpretable phase observation in InSAR is the double difference (in time between master and slave epochs, and in space between the measurement and the reference points) [64]. The deformation estimates are computed using free network adjustment to an arbitrary datum, e.g. by assuming a reference point in the image to be stable, or treating the average of all measurements as constant. This is termed as datum definition by inner constraints [30–33]. Changing to another reference point in the image does not affect the information contained in the system; it only refers the estimates to a different local datum.

The datum of a geodetic network is defined as the basic (minimum) parameters required to define the network in space, or to position the network with respect to a predefined coordinate system [33] (see Appendix C). There are 14 basic datum parameters for a deforming geodetic network, which are of translation, rotation and scale types, as well as their respective time derivatives. In case of one-dimensional deformation monitoring such as by InSAR, these 14 datum parameters reduce to 2: only the translation parameter (in z-direction) and its derivative are relevant. The estimation of these two parameters is done via datum connection, by which the local InSAR datum is connected to a global one. By doing so, all the InSAR-derived deformation estimates are represented in a standard TRF; this is the focus of ZOD for InSAR.

The process of geocoding transforms the radar coordinates (range, azimuth and height) of InSAR estimates to coordinates in a TRF. This uses information on the position of the satellite during the SAR image acquisition. Satellite state vectors provide this information, in a geocentric TRF. Geocoding uses additional information from the Doppler centroid (defining the squint angle), the local earth radius (from the ellipsoid) and the topographic height [6]. However, geocoding precision is limited by the precision of orbit determination, which is in the order of a few centimetres. Scatterer height estimation is approaching the precision of a few (tens of) centimetres, determined by experiments using corner reflectors with known phase-centres [65–68]; these errorbars are still too large for millimetre-level deformation monitoring.

Therefore, to define the InSAR free network in a TRF, external constraints by means of ‘absolute’ and precise geodetic measurements need to be introduced, preferably collocated with an InSAR measurement. We propose to use the concept of a worldwide surface network of GNSS stations. By (mechanically) attaching phase-stable radar transponders to GNSS stations, we obtain a network of radar beacons with observed elevation above the reference ellipsoid, and also observed motion with respect to the same reference ellipsoid. These reference GNSS stations are often located spatially dense enough to have one or more stations in a typical radar image. Alternatively, transponder-GNSS units (I2GPS [36, 37]) may be constructed and deployed specifically for this purpose. It is thus ensured that an ‘absolute’ InSAR reference point is present in each scene. The InSAR datum can then be changed to that of GNSS by applying a transformation that retains the network geometry (a similarity- or S-transformation [33, 35, 69, 70]). From
Tab. 2.1, the free parameters in ZOD are the estimated parameters \( \hat{x} \) and their quality \( Q_{\hat{x}} \); S-transformations can transform the parameters from one datum to another, and give the relation between VCMs arising from the choice of different datums. After ZOD, \( \hat{x} \) and \( Q_{\hat{x}} \) are defined in a standard TRF.

In the following chapter, we begin a detailed treatment of geodetic network design for InSAR, in temporal processing order, with SOD. We will focus on the quality description of InSAR measurements, both naturally-occurring and deliberately-introduced. This will form a basis for the design orders presented in the subsequent chapters.
Second-Order Design

Second-order network design, i.e. deciding which measurements to make and with what precision, is limited in InSAR to determining the quality of measurements, as explained in Ch. 2. InSAR differs from other classical geodetic techniques in that it is difficult to estimate measurement quality \textit{a priori}. This is because of the dependence of measurement precision on the physical properties of the radar scatterers. The exact origin of the reflection within a resolution cell is difficult to determine, and weaker reflections from the surrounding scatterers (clutter) can introduce further phase noise \cite{71}. Various assumptions (see Appendix B) made while solving the under-determined InSAR estimation problem also contribute to the precision of InSAR-derived deformation.

This chapter starts with a brief review of InSAR error sources and the quality description of the deformation estimated from coherent targets. Then follows a quality analysis of devices that may be deployed for network densification, beginning with corner reflectors, which have been well-studied in the past decades. More recently, research has been performed towards developing transponders for deformation monitoring \cite{72–76}, but experiments such as \cite{77, 78} have shown poor or partial agreement between transponder-InSAR and validation measurements using other geodetic techniques. In the last part of this chapter, we will demonstrate the applicability of a low-cost radar transponder for monitoring deformation, and empirically derive estimates of its precision using data from different satellites and geodetic validation measurements.

\begin{quote}
No measurement is exact, every measurement contains errors, the true value of a measurement is never known, and thus the exact sizes of the errors present are always unknown.
\end{quote}

Paul Wolf and Charles Ghilani

Parts of this chapter have been published in the IEEE Transactions on Geoscience and Remote Sensing 52, 1869 (2014) \cite{29}
3. Second-Order Design

3.1. InSAR error sources

SAR data contain information on the strength of the radar signal reflected from the Earth's surface (amplitude) as well as the travel time of the radar signal in the physical path to and from the Earth's surface (phase). Preprocessing (e.g., focusing and range migration) is performed on the raw SAR data to yield a single-look complex (SLC) image [79], where each resolution cell (or pixel) has a complex value formed by the summation of reflections from scatterers within the cell on the Earth's surface. The strength of these reflections depends on the physical properties (e.g., size, orientation, roughness, dielectric characteristics) of the individual scatterers. SAR system characteristics such as radar wavelength and incidence angle govern the interaction of the signal with the scatterer [6, 12, 71].

The scattering behaviour within a resolution cell may be distinguished into two extreme cases. **Point scattering** involves the presence of a dominant strong reflecting object within the resolution cell, called a point scatterer. The contribution of the surroundings is much weaker, and results in some level of noise or clutter. **Distributed scattering**, in contrast, is the effect of a large number of weakly-scattering objects that contribute to the total signal received from that resolution cell. Both these types of scattering can be temporally coherent or incoherent, depending on the change experienced by the scatterer in the period between the SAR acquisitions; **coherence** is a measure of the correspondence between two complex observations. InSAR extracts information from sufficiently-coherent pixels or groups of pixels. Fig. 3.1 schematically represents the various cases of coherent and incoherent point and distributed scattering mechanisms.

Each complex pixel in an SLC image consists of several contributions, which can be summarized as

$$
\phi = -2\pi a + \phi_{\text{range}} + \phi_{\text{atmo}} + \phi_{\text{scat}} + \phi_{\text{noise}}
$$

where \(a\) is the number of full phase cycles, \(\phi_{\text{range}}\) the range-dependent phase, \(\phi_{\text{atmo}}\) the signal delay caused by the atmosphere, \(\phi_{\text{scat}}\) the scattering phase related to the additional path length travelled due to the presence of multiple scattering objects within a resolution cell, and \(\phi_{\text{noise}}\) the thermal noise [6, 10, 12, 71].

Conventional InSAR involves interferograms created using pairs of SLC images; the phase value \(\phi_{i}^{01}\) of a pixel \(i\) in an interferogram is the phase difference between the corresponding pixels in the SLC images at master time \(t_0\) and slave time \(t_1\). However, the total atmospheric delay adds up to several metres, orbit accuracies are of (sub-)decimetre level, and the total number of integer phase cycles is unknown, making deformation measurement using a single interferogram pixel with millimetre-level precision impossible [64]. The first ‘interpretable’ phase observation is therefore the phase difference between two interferometric resolution cells, i.e. the double difference (in time between \(t_0\) and \(t_1\), and in space between pixels \(i\) and \(j\)). Therefore, InSAR observations require both a temporal and a spatial reference [64, 71], and we denote this double-difference interferometric phase as

$$
\phi_{ij}^{01} = -2\pi a + \phi_{\text{defo}} + \phi_{\text{topo}} + \phi_{\text{flat}} + \phi_{\text{atmo}} + \phi_{\text{orb}} + \phi_{\text{scat}} + \phi_{\text{noise}}
$$

where \(\phi_{\text{range}}\) of Eq. (3.1) is split into \(\phi_{\text{defo}}\), \(\phi_{\text{topo}}\), and \(\phi_{\text{flat}}\). \(\phi_{\text{defo}}\) is the phase component from deformation, \(\phi_{\text{topo}}\) from uncompensated topography, and \(\phi_{\text{orb}}\) from errors in
3.1. InSAR error sources

Figure 3.1: Examples depicting point and distributed scattering for coherent and incoherent cases. Top: Scattering objects within a resolution cell during two SAR acquisition epochs, indicated respectively by black and grey scatterers. Larger scatterers correspond to stronger reflections. Middle: Phasor diagrams for the two acquisitions, also respectively in black and grey. Bottom: Examples of scattering objects. Figure taken from [12].

the orbit parameters of the SAR acquisitions. $\phi_{\text{flat}}$ is due to a reference surface (ellipsoid), and is removed based on orbit and reference surface parameters. $\phi_{\text{scat}}$ together with $\phi_{\text{noise}}$ results in decorrelation, which has three dominant components [6, 7]: temporal decorrelation due to changes in the scattering properties within a resolution cell over time, geometric decorrelation due to incomparable scattering characteristics arising from different satellite viewing angles, and Doppler centroid decorrelation caused by different centre frequencies of the azimuth spectrum of the SAR acquisitions, dependent on the antenna squint angle. The phase contribution of our interest is $\phi_{\text{defo}}$; all other contributions are error sources for deformation studies. The ground deformation in the radar line of sight (LoS) for a radar signal of wavelength $\lambda$ between points $i$ and $j$ and times $t_0$ and $t_1$ is thus

$$d^{01}_{i,j} = -\frac{\lambda}{4\pi} \phi_{\text{defo}}. \tag{3.3}$$
3. Second-Order Design

3.2. Coherent targets

The major limiting factors in deformation estimation from single interferograms are often $\phi_{\text{atmo}}$ and $\phi_{\text{scat}}$. These error sources can be minimized by using a stack of multiple interferograms of the same area, i.e. time-series InSAR, and extracting information only from coherent radar targets. 

Persistent Scatterer Interferometry (PSI) [8–12] is based on coherent point or point-like targets, often man-made, that show stable phase behaviour over long time intervals, i.e. with minimal temporal decorrelation. These persistent scatterers (PS) dominate the reflection within a resolution cell, thereby also reducing the effect of geometric decorrelation. By using a stack of interferograms, the temporally-uncorrelated atmospheric phase contribution is estimated and removed. This stack may be constructed for PSI with one SAR image in common (i.e. a single master image), or in any arbitrary way in general (e.g. by minimizing the perpendicular baselines between the combinations). For PSI, where the slave acquisitions are $1, 2, ..., (K - 1)$ and $K$ is the number of SAR acquisitions,

$$\phi_{ij}^1 = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \phi_i^0 \\ \phi_i^1 \\ \phi_j^0 \\ \phi_j^1 \end{bmatrix}$$

(3.4)

with the superscript 1 now denoting the value at slave time $t_1$ referred to master time $t_0$ by default. Including its wrapped nature, this double difference phase $\phi_{ij}^1$ can be quantified as [10, 12, 71]

$$\phi_{ij}^1 = -2\pi a_{ij}^1 - \frac{4\pi}{\lambda} d_{ij}^1 - \frac{4\pi}{\lambda} B_i^1 \frac{1}{R_i^0 \sin \theta_{\text{inc}}^0} H_{ij} + \phi_{ij}^1_{\text{atmo}} + f_{\phi_{\text{orbit}}}^1 \left( \xi_0^0, \eta_0^0 \right) + f_{\phi_{\text{DC}}}^0, i - f_{\phi_{\text{DC}}}^1, i + \frac{2\pi}{\nu} \left( f_{\phi_{\text{DC}}}^0 - f_{\phi_{\text{DC}}}^1 \right) \xi_{ij}^0 + \eta_{ij}^0$$

(3.5)

where $a_{ij}^1$ is the integer ambiguity between points $i$ and $j$, $d_{ij}$ the radar LoS deformation between points $i$ and $j$ and times $t_0$ and $t_1$, $B_i^1$, $R_i^0$, $\theta_{\text{inc}}^0$ the perpendicular baseline, range and incidence angle respectively for point $i$, $H_{ij}$ the (residual) topographic height between points $i$ and $j$, $\phi_{ij}^1_{\text{atmo}}$ the (residual) atmospheric signal, $f_{\phi_{\text{orbit}}}^1$ the (residual) orbital trend, $f_{\phi_{\text{DC}}}^0, i - f_{\phi_{\text{DC}}}^1, i$ the Doppler centroid frequencies of master and slave acquisitions respectively, $\nu$ the satellite velocity, $\xi_0^0, \eta_0^0$ the subpixel positions in azimuth and range respectively, and $n_{ij}^1$ includes measurement noise and processing-induced errors.

An explicit distinction between these different phase components is important to isolate the component of our interest, i.e. the deformation $d_{ij}^1$. This can be done using external information, e.g. from ground truth or meteorological data, or by exploiting the different correlation properties of the phase components, summarized in Fig. 3.2. Topographic, atmospheric and orbital phase contributions are hence estimated and removed. To solve the under-determined problem of phase unwrapping, some knowledge of the
expected deformation model is required. There is a trade-off between the preservation of possible unmodelled deformation (i.e. any deviation of the actual deformation from the expected model) and the removal of the atmospheric signal, since the unmodelled deformation is not known \textit{a priori}. Appendix B gives an impression of the sensitivity of InSAR results to various assumptions made during PSI processing. Further treatment on InSAR error sources, the methodology of PS deformation estimation and error propagation is beyond the scope of this thesis; the interested reader is referred to [6, 10, 12, 71, 80, 81] for these aspects in the Delft implementation of the PSI algorithm (DePSI).

For the purpose of network design, it suffices to know that the final PS deformation estimated from the interferometric phase is disturbed primarily by decorrelation, any residual atmospheric signal and unmodelled deformation, as well as eventual unwrapping errors. Decorrelation is the effect of changes in scattering within a resolution cell as well as noise, lumped together and termed henceforth simply as ‘noise’. The PS deformation estimates are the InSAR measurements referred to in Ch. 2, i.e. the vector of observations \( y \) in Eqs. (2.1) and (2.2). Assuming unwrapping errors to be deterministic, the quality of PS measurements can therefore be written as

\[
Q_y = Q_w + Q_s + Q_n
\]

(3.6) using the notation of Eqs. (2.1) and (2.2). \( Q_w, Q_s \) and \( Q_n \) are respectively the stochastic descriptors (VCMs) of residual atmospheric error, unmodelled deformation and noise, propagated through the PSI processing chain. \( Q_w \) can be estimated as described in [6, 10, 12, 81]. \( Q_s \) is unknown \textit{a priori}, but intuitive estimates may be provided from experience based on geophysical knowledge of the deforming area. \( Q_n \) depends on the physical properties of the radar scatterer and interference due to surrounding clutter, which, along with uncertainties in identifying the phase centres, makes it difficult to be estimated \textit{a priori}. It can be approximated using metrics such as the \textit{signal-to-clutter ratio}.
ratio (SCR). SCR is defined as \[82, 83\]

\[
SCR = \frac{s^2}{c^2}
\]

where \(s\) is the amplitude of the pixel, in this context referred to as the signal, and \(c\) the amplitude of the clutter within the pixel. A spatial estimation window is used to estimate the SCR per pixel, assuming stationary stochastic behaviour of scatterer surroundings, i.e. that the surrounding pixels contain clutter equal to the clutter within the resolution cell. The amplitude of the central pixel represents the signal. A relation can be derived between the phase variance and the SCR as \[83\]

\[
\sigma^2_\phi = \frac{1}{2 \times SCR}.
\]  

A disadvantage of this method is that the surrounding pixels (representing the clutter) are needed for the assessment. The clutter can be overestimated in urban areas, because pixels around the PS may also contain point scatterers \[83\]. \(Q_n\) can be more precisely determined \textit{a posteriori}, by making use of deformation estimates from strong point scatterers with known phase-centres, and comparing them with the corresponding estimates from alternative geodetic techniques. Such an \textit{a posteriori} approach is described in the following sections, for corner reflectors and transponders.

A more recent family of time-series InSAR approaches deals with extracting coherence information from distributed scatterers, by processing only interferograms with small perpendicular and temporal baselines \[13–20\]. These distributed scatterers (DS) often occur with a higher density compared to PS, and the information content can be further enhanced by the filtering operations, at the cost of spatial resolution. Error propagation from interferometric phase to the final DS deformation estimates is usually based on a coherence matrix describing the correlation between the radar images, and is much more complicated and computationally intensive than for PSI. Some attempts have been made in \[18,20, 84, 85\]. In this thesis, we focus on improving PS networks. We make use of PS quality description, although in principle the methodology can also be applied for improving a DS network.

As a rule of thumb, including all the InSAR error sources (especially atmospheric error and decorrelation), naturally-occurring PS can be assumed to have a precision of < 5 mm \[86, 87\]. The precision of PS displacement rates also depend on the distance between the PS and the number of interferograms in the stack. As an example, it is shown in \[71\] that the precision of PS displacement rates derived from ERS/Envisat stacks of 24 to 74 interferograms are 0.04–0.1 mm per year, per \(\sqrt{\text{km}}\). This precision can be worsened by unmodelled error sources and unwrapping errors. Neglecting the azimuth sub-pixel position can lead to an additional error of \(\sim 0.5\) mm/year, and orbit errors of 5 and 8 cm respectively in radial and across-track directions can result in errors up to \(\sim 1\) mm/year between near and far range. The precision of deployed devices such as corner reflectors and transponders, as estimated from empirical studies, is discussed in the following sections.
3.3. Corner reflectors

The precision of corner reflectors has been validated via many experiments conducted in the past [27, 28, 88–90]. A corner reflector experiment performed in Delft over a period of five years (2003–2007) compared InSAR deformation estimates with that of precise levelling. The impact of atmospheric error was minimized by placing the reflectors a few tens of metres apart. This distance is below the correlation length of tropospheric and ionospheric error components, i.e. any possible phase disturbance was cancelled out by the double-differencing operation. Topographic errors were also negligible owing to the flat characteristics of the site chosen in Delft. The experiment concluded that the 1σ precision of corner reflector InSAR double-difference phase measurements (vertical) is 2.8 mm for ERS-2 (in zero-gyro mode [91]), and 1.6 mm for Envisat. The correlation between levelling and Envisat double-difference displacements was found to be 94% [28, 71]. Another experiment performed in Milan as described in [27] used data from multiple viewing geometries of Radarsat-1 and Envisat. The effective precision of measuring movement was estimated to be 0.75 mm in the vertical direction, and 0.58 mm in the East-West direction. It should be noted that the smaller numbers reported here are derived using SAR data from two satellites and four tracks, unlike in [28, 71] where a single track was used to derive precision estimates per satellite.

Some advantages of using corner reflectors as coherent targets are that they are conceptually simple and can be constructed quite economically. They do not require a source of power, and do not possess electronic components whose performance could decay/drift over time. However, they also suffer from a few drawbacks. They are large and heavy (particularly for longer wavelengths such as C-band), making them cumbersome to deploy especially in poorly-accessible areas. Their autonomous motion (settling effect due to their weight) may contribute to the deformation signal being measured. For long-term deployment spanning several years, these large structures can be disturbed or geometrically altered by weather or thermal conditions, fauna or even by vandalism or theft. They require protection from the accumulation of snow, rain and general debris, and/or need periodic maintenance. Additionally, only ascending or descending satellite passes can be utilized with a fixed reflector setup.

Corner reflectors need to be relatively large in order to return sufficient power back to the satellite. Active devices such as radar transponders, however, can be more compact (Fig. 3.3) to achieve the same radar cross-section (RCS); a transponder can be of the size of a shoe box for the same RCS as a trihedral corner reflector with a side of about 1 m [92]. Low-cost radar transponders developed for deformation monitoring applications are described further in the following section.

3.4. Active radar transponders

Radar transponders have been used extensively in the past for SAR external calibration, both radiometric [93–97] and geometric [98, 99]. But calibration transponders are expensive (of the order of several hundred thousand Euros [100]) because of the power stability needed for calibration. For example, Sentinel-1 specifications are power
accurate up to 1 dB, and therefore the calibration should be accurate to about 0.1 dB (equal to 3σ), with an RCS of about 70 dBm² [100]. However, if there is a less stringent requirement on amplitude precision or stability, transponder cost can be lowered dramatically, down to a thousand Euros, or even less if produced in large quantities [101].

A transponder designed for deformation monitoring operates as follows. It wakes up from sleep mode according to a preprogrammed schedule. When it is illuminated by the radar satellite, it receives the signal on a receiver antenna, and instantaneously retransmits a band-pass filtered and amplified version of the received signal on the transmit antenna. After the satellite overpass, the transponder enters sleep mode again.

Being active devices, transponders face the undesirable possibility of their electronic component performance drifting/decaying over time. Therefore, the use of transponders for geodetic applications heavily depends on their short and long term phase stability (reliability) and precision in the relevant operating and environmental conditions. In other words, it is of paramount importance that the transponder shows phase stability, i.e. it transmits a signal with the same frequency and a constant phase relationship with the received signal from the satellite, for all overpasses.

Transponders have several advantages over corner reflectors. They can be compact, lightweight (< 4 kg) and inconspicuous. They may be sealed, function autonomously and over a wide temperature range with internal power for more than a year, and are less susceptible to environmental impact such as strong winds, precipitation and debris accumulation. Hence they require little or no maintenance; visits would only be required to change/charge the battery, check for clock drift, or upload a new SAR acquisition schedule, if needed. Moreover, since a transponder is transmitter-specific and only turned on at the time of the satellite overpass, it offers little interference to other radar or radio targets. Transponders can be used for both ascending and descending satellite modes in a single fixed setup, providing two components of motion vector as well as doubling the frequency of measurements. Since the signal frequency can be preprogrammed, they can be used for various C-band SAR sensors in a single setup.

A limitation of current transponders is battery maintenance: the battery needs to be recharged once in several months, and may eventually need replacement if deployed over several years. This may be circumvented by the use of external power sources such as small solar panels. Also, the operational use of transponders may require transmission
3.4. Active radar transponders

licenses, which vary across national boundaries. Despite these limitations, transponders can be a more practical alternative to corner reflectors, as long as phase stability is ensured.

3.4.1. Hardware

The prototype transponder used in this study, a compact active transponder (CAT) consists of receiving and transmitting antennas, a phase-stable amplification chain designed for providing around 40 dB of gain (equivalent to an RCS of around 29 dBm$^2$) and the means to power and control the unit. It has four C-band microstrip transmit and receive antenna arrays, and radio frequency (RF) switches control the received and transmitted signals from the satellite during overpass. The switching is programmed for signal throughput in vertical or horizontal polarization for the west- and east-pointing antennas. Similar switches also control the selection of switched bandpass filters designed to be compatible with Envisat/ERS or Sentinel-1/Radarsat-2. DC power is provided via a rechargeable deep discharge lead acid battery, and alternative options with solar panels may also be implemented. A schematic of the hardware design is depicted in Fig. 3.4.

Transponder design addresses two key technical issues [101]: sustenance of phase stability in all circumstances, and isolation between receive and transmit channels. Phase stability implies that the phase relationship between received and transmitted radar signals remains constant over successive satellite acquisitions, irrespective of operating temperature or other changing ambient conditions. This is implemented by two means: selection of phase-stable components in the amplification chain, and calibration of residual variations (for example with temperature). In order to avoid cross-coupling between the receive and transmit channels, over 70 dB of isolation is required between the receive and transmit antennas. Again, designing to avoid RF propagation and specific mitigation steps prevent feedback which otherwise disturbs phase stability. Transponders may need to be left unattended for extended periods, which requires a protective...
housing: it is currently designed for operation between -10 and +40°C. Low power consumption is achieved by turning the units on for short windows around satellite overpasses, requiring a local clock to be synchronized versus satellite overpasses. The transponder microcontroller reads both user-programmed calendar and real-time clock information in order to synchronize with satellite overpasses, and maintains synchronization by means of a commercial quartz oscillator. The 3 dB beamwidth of the transponder in azimuth and elevation directions is about 20° and 40° respectively. For comparison—a trihedral corner reflector has a 3 dB beamwidth of about 40° in both directions [102], and is therefore less sensitive to azimuth pointing errors than a transponder.

As mentioned before, it is of paramount importance in geodetic applications to ensure the phase stability of transponders. Three field experiments were performed towards this objective. First, experiments in the Netherlands (Delft) and Slovenia were set up in 2011, making use of ERS-2 and Envisat data respectively; at that time, no other C-band satellite data was available to us. A methodology was therefore developed to account for the inaccuracies introduced in InSAR (owing to the zero-gyro mode of ERS-2 and the drifting perpendicular baselines of Envisat prior to satellite decommissioning) as well as in the geodetic techniques used for validation (levelling and GNSS). Later on, in 2013, the lessons learnt from these two experiments and access to Radarsat-2 data prompted us to perform another iteration of transponder validation in Wassenaar, the Netherlands. In the remainder of this chapter, we present the methodology and results from these three experiments, in chronological order.

3.4.2. Validation experiments

Transponder setup

Delft—controlled validation experiment: A field experiment comprising three transponders (T1, T2 and T3) and two corner reflectors (C1 and C2) was set up in a pasture used for dairy farming in Delft, as shown in Fig. 3.5. The area was chosen because of the flatness of the ground (to exclude topographic phase contribution), low background clutter, and some amount of surface dynamism (swelling and shrinkage of peat), in addition to logistical convenience. The transponders were deployed by attaching them to 50 cm long iron stakes and boring these stakes into the soil. The experiment involved capturing the relative motion between transponder-corner reflector pairs with both InSAR and levelling, to check if InSAR measures the same deformation signal as levelling does, and if so, to what precision.
Device pairs were identified as T1-C2, T2-C1 and T3-C1. Levelling between these pairs (Fig. 3.5) was performed shortly before or after the ERS-2 satellite overpasses, usually within 24 hours. One of the transponders (T3) was subjected to an intentional displacement, to validate the ability of InSAR to capture this motion.

Slovenia—validation of operational performance: To validate transponder operation in a practical case of deformation monitoring, a landslide site was chosen in a heavily vegetated area in the Slovenian Alps (Figs. 3.6 and 3.7). The landslide is known to have produced historical debris flows, with the potential for future slides presenting a risk to the village in the valley [103].

One reference and two landslide points were selected for this experiment, with a further location across the Sava fault chosen for tectonic application (Fig. 3.6). At each of these locations, a transponder and a Global Positioning System (GPS) receiver mounted on a common baseplate were installed, as shown in the inset of Fig. 3.6. This ensured that the InSAR and GPS observations were collocated, i.e. that the two independent techniques were measuring the same deformation signal.
InSAR measurements

Delft experiment: The experiment spanned the duration of the ERS-2 Ice-Phase Mission (before satellite decommissioning, Apr–Jul 2011), where SAR images were acquired with a three-day repeat cycle. Single master interferograms are generated for 19 SAR images for which the corresponding levelling was performed. The master (13 May 2011) is chosen based on maximal stack coherence, which is a function of perpendicular and temporal baselines and Doppler centroid frequency [10]. InSAR double differences are computed for transponder-corner reflector pairs using Eq. (3.4), with 13 May 2011 as the time reference.

In Eq. (3.5), $a_{ij}^1$ is solved using a testing procedure, described in Sec. 3.4.2; $H_{ij}$ is estimated from levelling; $\phi_{ij,\text{atmo}}^1$ is regarded to be less than 1–2 mm, due to the short device distances (< 100 m) [6]; $f_{\phi_{\text{orbit}}}^1$ is rendered negligible for the same reason; $\xi_{ij}^0, \eta_{ij}^0$ are determined by oversampling by a factor of 32; and $n_{ij}^1$ is assumed zero-mean. The only unknown, deformation $d_{ij}^1$, can therefore be estimated.

Slovenia experiment: Envisat SAR images were acquired in two tracks (108 and 381) in the periods Feb 2011–Mar 2012 (14 images) and Mar 2011–Dec 2011 (9 images), respectively. This dataset was acquired after the orbital manoeuvre of Oct 2010, which implied drifting perpendicular baselines at the Slovenian latitude [104], as visible in the baseline plots of Fig. 3.8.

Small baseline (SBAS) interferograms are therefore created for both Envisat tracks, and time-series InSAR processing is performed using StaMPS [13, 17]. The topographic phase is removed using an SRTM 3-arcsecond digital elevation model (DEM). InSAR
3.4. Active radar transponders

Figure 3.8: Baseline plots of the available Envisat SAR data and small baseline combinations for track 108 (left) and track 381 (right) over Slovenia. The SAR acquisitions are marked with red circles, and the green lines denote the small baseline combinations.

double differences are computed for transponder pairs using Eq. (3.4), with time reference to 01 August 2011 (track 108) and 21 July 2011 (track 381). Therefore, $d_{ij}$ is solved by InSAR phase unwrapping and the testing procedure described in Sec. 3.4.2; $H_{ij}$, $\phi_{ij,atmo}^1$ and $f_{\phi_{orbit}}^1$ are estimated during the SBAS InSAR processing; and $n_{ij}^1$ is assumed zero-mean. $\xi_{ij}^0$, $\eta_{ij}^0$ are not estimated because of the short perpendicular baselines chosen and the stability of the Envisat Doppler centroid (unlike ERS-2 in the Delft case, which was operating in its zero-gyro mode [91]).

Validation measurements

**Levelling in Delft:** Spirit levelling was performed shortly before or after most satellite overpasses, usually within 12 hours. For comparison with InSAR double differences, levelling double differences in radar LoS are derived from levelling height difference measurements $\Delta h_{ij}$ as

$$d_{ij,lev}^{0',1'} = \cos \theta_{inc} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \Delta h_{ij}^{0'} \\ \Delta h_{ij}^{1'} \end{bmatrix}$$

(3.9)

where $0'$ and $1'$ are the levelling measurement epochs closest to the InSAR master and slave times respectively, $i$ and $j$ the two levelling points (in this case, a bolt on the transponder platform and the apex of the corner reflector), and $\theta_{inc}$ the radar incidence angle. The levelling double differences are thus converted into the radar LoS assuming no horizontal deformation; a valid assumption since the reflectors and transponders were based 20-50 cm into the ground with no cause for horizontal deformation over 3-day intervals. Additionally, the meadow chosen was a secure site, meaning that extraneous sources of (horizontal) deformation could be ruled out. Campaign GPS measurements also confirmed this assumption.
Based on the closing errors of the levelling loops, the standard deviation is estimated to be 1.4 mm for a single height difference measurement converted to radar LoS, which propagates to a precision of 2 mm for double differences. This includes observational inaccuracies that may have crept in during the surveys due to, for example, windy or rainy conditions. The higher closing errors on such days are accounted for in the evaluation, see Sec. 3.4.2.

**GPS in Slovenia:** For the four installed GPS receivers (collocated with transponders), station coordinates are estimated in ITRF2008/IGS08 (see Appendix C) on a daily basis. The daily North-East-Up solutions are converted into double differences in the radar LoS for comparison with the corresponding transponder-InSAR double differences, as

\[
\begin{bmatrix}
\rho^0_i \\
\rho^0_j \\
\rho^1_i \\
\rho^1_j
\end{bmatrix}
\]

where the superscripts 0 and 1 are the GPS measurement dates corresponding to the InSAR master and slave dates respectively, \(i\) and \(j\) the two GPS receivers, and \(\rho\) the North-East-Up solutions converted into radar LoS. The formal standard deviations are estimated during daily processing to be below 0.5 mm in North and East directions, and 1.5–2 mm in the Up direction. These standard deviations are on the optimistic side because they only represent the internal accuracy during the processing. In particular, they ignore the effect of long periodic and systematic orbit errors. Instead of the formal standard deviations, the daily station repeatability (root mean square error of the daily station coordinates after fitting a linear trend) can be considered as an indicator of the quality of the GPS solution. Compared to the formal errors which tend to be too optimistic, the repeatabilities are overbounding the GPS errors and are on the pessimistic side. For the installed receivers in Slovenia, the repeatability for the North and East is 2–3 mm, and 5–7 mm for Up. More details on these GPS processing aspects can be found in [36, 37].

A moving average block filter of fortnightly length is applied to the GPS double differences converted to the radar LoS, in order to reduce the noise (due to carrier-phase multipath, different horizon masks, and so on) in the GPS time series that is not averaged out on a daily basis. Based on the conservative formal standard deviations and the pessimistic station repeatabilities, we estimate the standard deviation of the GPS measurements to be about 3 mm in the radar LoS direction, which translates to a double difference standard deviation of 4.3 mm.

**Estimation of a priori transponder precision**

The transponder SCR is estimated (Fig. 3.9) by taking the clutter to be the average intensity of four quadrant areas around the transponder that are not affected by the transponder sinc sidelobes in range and azimuth directions, and the signal to be the integration of the transponder intensity sinc pattern (in both range and azimuth) corrected for clutter [82, 105]. The average transponder SCR is calculated to be 14.3 dB. From this, us-
Figure 3.9: SCR estimation method for a transponder with phase-centre location marked in black. The grey area is used to estimate the clutter, and the white area to estimate the signal \([10, 82]\)

In Eq. (3.8), the phase standard deviation for a single SAR observation is estimated as 0.14 rad, which corresponds to a double difference LoS measurement standard deviation of 1.2 mm. We take this to be the \textit{a priori} transponder precision.

\textbf{Estimation of \textit{a posteriori} transponder precision}

The \textit{a priori} transponder double difference phase precision, derived from its SCR, can be further refined by comparing the transponder-InSAR deformation measurements with the corresponding validation measurements for both the experiments. Therefore, the \textit{a posteriori} transponder precision is estimated by using independent external validation measurements, levelling in the Delft case, and GPS in the Slovenia case.

\textbf{Functional and stochastic models:} If the stochastic vector \(y\) of \(m\) observations bears a known linear relationship with the vector \(x\) of \(n\) unknown parameters, we may write a model of observation equations as given in Eq. (2.2). We assume the null hypothesis that InSAR and the validation techniques (levelling or GPS) measure the same deformation signal \(d_{ij}^{01}\), with a possible offset \(b_{\text{offset}}\) between their time series owing to the bias introduced by the InSAR and validation measurement uncertainties at reference time. Therefore, the functional and stochastic models of our observations can be written as

\[
\begin{align*}
\mathbf{H}_0 : & \left[ \begin{array}{c}
\frac{\partial y_{\text{InSAR}}^{01}}{\partial x_{\text{InSAR}}^{01}} \\
\frac{\partial y_{\text{val}}^{01}}{\partial x_{\text{val}}^{01}}
\end{array} \right] = \left[ \begin{array}{c}
I \\
I
\end{array} \right] \left[ \begin{array}{c}
d_{ij}^{01} \\
b_{\text{offset}}
\end{array} \right] ; \\
Q_y &= \begin{bmatrix}
Q_{y,\text{InSAR}} & 0 \\
0 & Q_{y,\text{val}}
\end{bmatrix}
\end{align*}
\]

where \(H_0\) indicates the null hypothesis, \(I\) is the identity matrix, \(\frac{\partial y_{\text{InSAR}}^{01}}{\partial x_{\text{InSAR}}^{01}}\) and \(\frac{\partial y_{\text{val}}^{01}}{\partial x_{\text{val}}^{01}}\) the InSAR and validation double differences respectively, \(Q_y\) the overall VCM of all the observations, and \(Q_{y,\text{InSAR}}\) and \(Q_{y,\text{val}}\) respectively the \textit{a priori} VCMs of InSAR and validation measurements separately. \(Q_{y,\text{InSAR}}\) is derived from the \textit{a priori} transponder-InSAR double difference precision, and \(Q_{y,\text{val}}\) is derived from the estimated LoS-converted levelling and GPS double difference precisions of 2 mm and 4.3 mm respectively. \(Q_{y,\text{InSAR}}\) and \(Q_{y,\text{val}}\) also take into account the spatio-temporal covariances between observations.
Additionally, in case of levelling, measurements with a higher closing error are given proportionally lower weights. The best linear unbiased estimator (see Appendix A) of $\tilde{y}$ is given by [106]

$$\hat{y} = A(\tilde{A}^T \tilde{Q}_y^{-1} \tilde{A})^{-1} \tilde{A}^T \tilde{Q}_y^{-1} y.$$  \hfill (3.12)

**Hypothesis testing:** We use the Overall Model Test (OMT) [43] to determine if the null hypothesis is accepted (see Appendix A). The OMT test statistic is given by

$$T_{OMT} = \frac{\hat{e}^T \tilde{Q}_y^{-1} \hat{e}}{m - n}$$  \hfill (3.13)

where $\hat{e}$ is the vector of residuals $y - \hat{y}$, and the difference between the number of observations and unknowns, $m - n$, denotes the redundancy. $T_{OMT}$ should be close to 1 for $H_0$ to be accepted. Rejection of $H_0$ in the OMT can be due to the presence of large errors, anomalies or disturbances, or because the models (either functional or stochastic) fail to represent the measurements.

In Slovenia, the GPS receiver was located on the same baseplate as the transponder, and collocated InSAR and GPS measurements were made at the time of satellite overpass. In the Delft case, levelling was usually performed within 12 hours before or after the satellite overpass (which was at midnight). It could therefore be assumed that in both cases, InSAR and validation techniques were measuring the same deformation signal. Hence, our functional model is correct, and failure of the OMT can only be attributed to the presence of outliers and/or an incorrect stochastic model.

**Ambiguity resolution, outlier detection and variance component estimation:** We use the w-test [43] data snooping technique (see Appendix A) to check each observation for an outlier (i.e. a gross error). Since the w-test statistic is normally distributed with zero-mean and standard deviation equal to 1, we define an outlier to be an observation that yields a w-test statistic value larger than 2. The term $a_{ij}^1$ in Eq. (3.5) is also solved using the w-test; the value of $a_{ij}^1$ that gives $\phi_{ij}^1 \pm 2\pi a_{ij}^1$ the minimum w-test statistic is chosen to be the integer ambiguity. If a large number of outliers are detected, we do not remove them, but estimate a new value of InSAR standard deviation using the technique of least-squares variance component estimation (VCE) [107, 108]. We then check again for outliers using the new value of InSAR standard deviation, and remove these outliers, if any. The final \textit{a posteriori} InSAR double difference standard deviation is determined by applying VCE again on the subset of observations after outlier removal.

The following section summarizes the results obtained from the experiments by applying the above methodology.

### 3.5. Results of transponder validation

#### 3.5.1. Laboratory tests

Laboratory tests were performed on the transponder to measure its RCS. Transmitter and receiver horn antennas were aligned with respect to the transponder within its beamwidth,
3.5. Results of transponder validation

Figure 3.10: Example signals measured in the laboratory, from a corner reflector of \( L = 44 \text{ cm} \) (black), and from a transponder in Envisat/ERS (red) and Sentinel-1/Radarsat-2 (blue) bands.

at a distance of 8.7 m. A continuous-wave signal sweep between 5.1 and 5.7 GHz at -18 dBm was emitted from the transmitter-horn to simulate a satellite signal. The signal received from the transponder at the receiver-horn was corrected for the background signal, noise-filtered and visualized. The same setup and procedure was used with a trihedral corner reflector made of three isosceles right triangles, of isosceles side length \( L = 44 \text{ cm} \). Fig. 3.10 shows the received power from the corner reflector and the transponder (in its two bands of operation). The maximum RCS of a trihedral corner reflector is given by [102]

\[
\sigma_{\text{max}} = \frac{4\pi L^4}{3\lambda^2}
\]  

(3.14)

where \( \lambda \) is the radar wavelength. The maximum RCS of the corner reflector is therefore calculated to be about 17 dBm\(^2\). From Fig. 3.10, the transponder is about 15 dB stronger than the corner reflector, which gives us an indicative transponder RCS of 32 dBm\(^2\). This is equivalent to a corner reflector of \( L = 1 \text{ m} \).

3.5.2. Delft: ERS-2

The comparison between levelling and InSAR double differences in the radar LoS for the three device pairs is shown in Fig. 3.11.

\textit{A posteriori} transponder precision

For ERS-2 data, the transponder-InSAR double difference standard deviation in the radar LoS estimated by applying VCE on all the 57 observations is 6.7 mm. After detecting and removing 2 outliers, the estimate drops to 3.9 mm.
Figure 3.11: Comparison between InSAR and levelling double differences in the radar LoS. From top to bottom: device pairs T1-C2, T2-C1, T3-C1.
3.5. Results of transponder validation

Long-arc comparison

As can be seen in Fig. 3.5, transponder T1 and reflector C2 are of comparable distances from reflector C1 (~450 m). Assuming that the deformation in the vicinity of T1 and C2 was spatially correlated, i.e., T1 and C2 moved in the same way compared to C1, if the double differences between the transponder-reflector and reflector-reflector pairs match, we can infer that the transponder behaves like the corner reflector, though the converse may not necessarily be true. Fig. 3.12 shows this comparison.

It is observed that the transponder-reflector and reflector-reflector double differences follow similar trends. The standard deviation of the difference between the T1-C1 and C2-C1 measurements is 5.0 mm. Assuming that the corner reflector and transponder have comparable precision levels, the transponder double difference standard deviation is then 3.5 mm or better, since the difference also includes the relative motion between T1 and C2.

3.5.3. Slovenia: Envisat

The comparison between InSAR and GPS time series for the two InSAR tracks and three unit combinations is summarized in Figs. 3.13 and 3.14.

A posteriori transponder precision

Images from two Envisat tracks (108 and 381) were used in Slovenia, and hence the a posteriori transponder double difference precision is estimated separately for the two different LoS directions. VCE gives an estimate of 4.9 mm from 34 observations in track 108 and 4.6 mm from 25 observations in track 381. After detecting and removing 1 outlier in track 108, the transponder-InSAR double difference precision estimate drops to 1.8 mm.
Figure 3.13: Comparison between InSAR and GPS double differences in radar LoS for Envisat track 108. From top to bottom: device pairs T1-Ref, T2-Ref, T3-Ref.
3.5. Results of transponder validation

Figure 3.14: Comparison between InSAR and GPS double differences in radar LoS for Envisat track 381. From top to bottom: device pairs T1-Ref, T2-Ref, T3-Ref.
3.5.4. Wassenaar: Radarsat-2

In mid-2010, the experiment in Slovenia was set up with a twofold intention: to monitor a landslide area in Slovenia, and to validate the use of transponders. However, in October that year, Envisat entered into a phase of mission extension which involved a change in orbit control. As a result, the dataset over Slovenia was acquired with drifting perpendicular baselines, shown in Fig. 3.8. In April 2011, the opportunity of acquiring SAR data with 3-day repeat intervals presented itself, through the ERS-2 Ice-Phase Mission. Despite the zero-gyro mode of ERS-2, we decided to set up the experiment in Delft and use this chance to quickly acquire a long stack of SAR images for transponder validation.

The methodology proposed in Sec. 3.4.2 was developed to provide the most accurate estimate of transponder precision while still accounting for the inaccuracies in SAR, levelling and GPS data. Despite this, analysis of the results from the Delft and Slovenia experiments pointed out that several other factors could have contributed to lowering the transponder precision estimates. In case of the Delft experiment, levelling was performed within 12 hours of the actual satellite overpass, which was around midnight. From Fig. 3.11, there can be a variability of up to 10 mm over a 3-day period, possibly due to the swelling or shrinking of peat soil in the meadow. This means that the difference in InSAR and levelling measurement times may have contributed to an InSAR-levelling difference of up to 1–2 mm. Additionally, atmospheric delay differences may amount up to a millimetre or two, even for short (100 m) arcs. In the Slovenia trial, errors related to snow or frost on the GPS antenna during the winter months may have crept in. Moreover, relatively few images in each track (therefore lower number of observations for statistical analysis) may also have worsened the transponder precision estimate.

The values presented in Secs. 3.5.2 and 3.5.3 can therefore be interpreted as the upper limit of transponder precision estimates, given the information that was available via the two field experiments. Learning our lessons from these attempts, we decided to perform another validation experiment at a quiet piece of farmland in Wassenaar, the Netherlands, in 2013. Three transponders (called CAT1, CAT2 and CAT3) were installed firmly into the ground in Aug 2013. The conditions in Wassenaar were almost identical to those in Delft, and no significant deformation was expected in the area besides peat swelling/shrinking. To avoid this effect, the topsoil layer was removed at transponder locations, and the transponders were drilled into heavy concrete blocks dug 15 cm into the ground, as shown in Fig. 3.15. Images from two Radarsat-2 tracks were utilized: t102 and t202. The InSAR deformation estimates were compared with those from precise levelling, using the same methodology as in the Delft experiment, to quantify the transponder measurement precision for Radarsat-2.

The Wassenaar experiment was an improvement over the previous ones in Delft and Slovenia for several reasons:

1. More reliable Radarsat-2 data were available, without the problems of drifting baselines of Envisat or the zero-gyro mode of ERS-2.
2. The rigid installation of the transponders in Wassenaar on concrete blocks after removing the topsoil minimized the effect of possible swelling or shrinkage of the ambient peat in the interval between InSAR and levelling.
3.6. Summary and discussion

A more modern and precise levelling instrument was used to perform the validation measurements, and observations were repeated several times to increase the derived measurement precision (< 0.2 mm). Care was taken such that closing errors of less than 0.1 mm were achieved.

Fig. 3.16 shows the temporal variation in the amplitude characteristics at the transponder locations in Wassenaar, both before and after transponder installation. Radar brightness ($\beta_0$), i.e. the radar reflectivity per unit area in radar LoS direction, is used as the calibrated amplitude parameter; its definition and computation are explained in Sec. 5.3.3. In all three cases, a gain of the order of 10 dB and higher temporal stability is observed after transponder installation. The double-difference transponder precision was found to range between 0.8–1.0 mm in the radar LoS, by comparing transponder-InSAR double-differences with optical levelling using data from the period Aug 2013–Jul 2014. This is shown in Fig. 3.17 after conversion to vertical direction (horizontal components were not expected).

3.6. Summary and discussion

The applicability of active radar transponders for deformation monitoring has been demonstrated, both under controlled conditions (‘quiet’ meadows in Delft and Wassenaar) as well as in an operational setting (monitoring a landslide area in Slovenia). The empirical transponder-InSAR double difference phase measurement precisions in radar LoS derived from the three field experiments are summarized in Tab. 3.1. The Delft experiment (InSAR and levelling) and the Slovenia monitoring case (InSAR and GPS) show that the empirical standard deviation of transponder double difference phase measurements in the radar LoS for Envisat and ERS-2 is 1.8–4.6 mm after outlier removal. The use of Radarsat-2 data and better-constrained validation (levelling) measurements in Wassenaar yielded excellent precision estimates, of sub-millimetric level (0.8–1.0 mm).

We have estimated the a posteriori transponder precision by comparing transponder-InSAR measurements with ‘ground truth’ validation measurements, i.e. levelling and
Figure 3.16: $\beta_0$ time-series (t102 and t202) for the transponders deployed in Wassenaar. Transponder initialization signifies the time when the transponder was installed, or the first date when it was programmed for transponding. The dates when the transponder malfunctioned were due either to power supply failure, or logistical inconvenience in the timely update of the transponder internal calendar. Transponder testing dates are when tests were performed on the transponder, which might have influenced the signal. Transponder malfunction and testing dates are not considered in the analysis.
Figure 3.17: Comparison of InSAR double-differences with those of levelling for all combinations of transponders in Wassenaar. The transponders failed to operate during one SAR acquisition of each track (indicated with black lines) owing to a human error in programming their calendars.
Table 3.1: Empirical *a posteriori* precision ($\sigma$) of transponder-InSAR double difference phase measurements in radar LoS. T and C signify transponder and corner reflector, respectively. Standard deviations are in [mm].

<table>
<thead>
<tr>
<th>Location</th>
<th>Satellite</th>
<th>Device pair</th>
<th>#InSAR obs.</th>
<th>$\sigma_{\text{all obs.}}$</th>
<th>#outliers</th>
<th>$\sigma_{\text{outliers removed}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slovenia Env (t108)</td>
<td>T-T</td>
<td>34</td>
<td>4.9</td>
<td>1</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Slovenia Env (t381)</td>
<td>T-T</td>
<td>25</td>
<td>4.6</td>
<td>0</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>Delft</td>
<td>ERS-2</td>
<td>T-C</td>
<td>57</td>
<td>6.7</td>
<td>2</td>
<td>3.9</td>
</tr>
<tr>
<td>Wassenaar RS-2 (t102)</td>
<td>T-T</td>
<td>39</td>
<td>0.8</td>
<td>0</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Wassenaar RS-2 (t202)</td>
<td>T-T</td>
<td>39</td>
<td>1.0</td>
<td>0</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

GPS. The initial experiments under suboptimal conditions estimated the transponder double difference phase precision to be 1.8–4.6 mm for ERS/Envisat in the radar LoS, which is in the similar ballpark as the corresponding corner reflector precision of 1.5–2.6 mm estimated in [28]; the latter used large corner reflectors with $\sim$3 times the SCR of the transponders, and double the number of SAR images. To judge the level of significance of the estimated transponder double-difference phase precision, it can be compared to the maximum standard deviation of double-differences of distributed scatterers. This standard deviation follows from the uniform distribution of distributed scatterer phases in the interval $[-\pi, \pi]$ [109]:

$$
\sigma_{DS}^2 = \frac{1}{48} \left( \frac{\lambda}{\cos \theta_{\text{inc}}} \right)^2
$$

(3.15)

where $\lambda$ denotes the wavelength and $\theta_{\text{inc}}$ the incidence angle of the radar signal. For ERS/Envisat, this standard deviation is $\sim$9 mm. The estimated transponder double difference precisions are much lower than this maximum, which shows that these values could not have occurred from a uniformly random distribution.

The ‘proof of the pudding’, however, is the experiment in Wassenaar, which was performed accounting for the lessons learnt in the previous experiments. Much-improved estimates of transponder precision were obtained, i.e. 0.8–1.0 mm in the radar LoS direction. The results confirm that these low-cost transponders are phase-stable and can be used as compact and lightweight alternatives to corner reflectors for deformation monitoring.

In this chapter, we have provided a review of InSAR error sources and identified the primary contributors of inaccuracies in PS deformation estimates. Further, we have touched upon corner reflector measurement precision, based on prior experimentation. We have introduced low-cost radar transponders that can be used as more convenient alternatives to corner reflectors, and have estimated their measurement precision via three field experiments. All this information is utilized in constructing the VCM of InSAR measurements, which forms an input to the next step of geodetic network design for InSAR — combined FOD-THOD. We proceed to describe the combined FOD-THOD methodology in the following chapter.
4

Combined First- and Third-Order Design

*I know one thing: that I know nothing.*

Socrates

The focus of this chapter is to improve, extend or densify an existing network of InSAR measurements in an optimal way, if required, by introducing additional measurements, i.e. deploying devices such as corner reflectors or transponders. As explained in Ch. 2, this order of network design was classified as third-order design (THOD) in classical geodesy. First-order design (FOD) is not directly applicable to InSAR because the locations of the opportunistic InSAR observations cannot be designed. However, in the methodology presented in this chapter, FOD is used in THOD for constructing a criterion matrix [34], which is an idealized VCM of parameter estimates generated from the desired optimality criteria. We therefore term this design order for InSAR as combined FOD-THOD.

4.1. Overview

From the perspective of third parties (end-users) who make decisions based on land survey results, the InSAR technique is of little use unless they are guaranteed to have dense-enough measurements to estimate their final parameters of interest, e.g. ground displacements or deformation model parameters. Until now, they had no systematic method to determine *a priori* whether the naturally-occurring coherent targets have adequate spatial distribution to estimate the parameters of their interest. InSAR studies so far have performed network densification by deploying a pre-determined number of additional devices based on practical considerations, or ‘intuitively’ (with inter-device

Parts of this chapter have been published in the *IEEE Transactions on Geoscience and Remote Sensing* 53, 7 (2015) [38].
spacing based on implicit assumptions of atmospheric noise properties, and the range and variability of the expected deformation), e.g. [77, 88, 90, 110–117]. In this chapter, we formalize this ‘intuition’. We propose a method, that (1) provides a generic description of any deformation phenomenon, (2) determines whether the naturally-occurring InSAR measurements are adequate, (3) finds the minimum number of additional devices (if required) and (4) their optimal ground locations.

The method digests, as inputs, the expected locations and quality of existing coherent targets, the quality of the devices being deployed, and, if available, any prior knowledge of the deformation signal. At the core of the algorithm is a comparison of different VCMs of the parameters of interest, resulting from several possible device configurations, with the criterion matrix. The optimal device configuration is the one that achieves, with minimum devices, a quality closest (in the predefined metric of Eq. (2.4)) to that provided in the criterion matrix. More stringent user-requirements, reflected in the criterion matrix, would necessitate the deployment of a larger number of devices.

The outcome of the proposed method is a geodetically optimal network of the minimum number of devices that satisfies user-defined requirements on the quality of parameters that are of interest to them, even before acquiring a large stack of SAR data. The network is optimized with respect to precision, reliability and cost criteria. To determine the trade-off between geodetic location optimality and physical suitability, the designed device locations can be combined with other factors such as visibility by the satellite in terms of geometry and radiometry, land cover and location feasibility.

A schematic input-output representation of the proposed method is shown in Fig. 4.1. Devices are intuitively located in regions where coherent targets are absent or of poor quality. Therefore, some required inputs to the proposed method are the locations and quality of existing coherent targets, including all the dominant (and often correlated) InSAR error sources. The measurement quality of the deployed devices is an additional input; lower precision implies that more devices are required. The aspects of coherent target quality including that of corner reflectors and transponders have been detailed in Ch. 3. The expected PS locations can be estimated before acquiring and analyzing a stack of SAR images using several possible methods. PSI can be performed on previously acquired SAR data over the area of interest, preferably acquired with similar looking direction, resolution and wavelength as the SAR data intended to be utilized, to get an idea of the expected PS distribution. Likely PS candidates can be also be selected using the amplitude characteristics of a single SAR image, e.g. by amplitude or SCR thresholding [10, 12]. SAR sensor characteristics, optical data and measures of vegetation or urbanization can further be used to predict PS locations. Examples of such methods are given in [57–61]).

A key conceptual driver is the inclusion of all available a priori knowledge (APK) of deformation, i.e. the expected subsidence or uplift characteristics of the area. The proposed generic methodology is applicable at all levels of APK — from no available information at all, to a precise prediction of deformation. The case of no knowledge is trivial (each point could then deform differently from the next, necessitating a measurement at every pixel), and is therefore dismissed. The most relaxed case considered here is when statements on the smoothness, or spatial variability, of the expected deformation can be made. For example, if the source of deformation is very deep, we expect the spatial gra-
4.2. Methodology through synthetic experiments

The proposed methodology is first summarized (Figs. 4.2 and 4.3), and then discussed in detail for a situation in which a parametric model is available as APK (Sec. 4.2.3), followed by one in which only the smoothness of the deformation signal is assumed to be known (Sec. 4.2.4).

The InSAR observations $y$ are double-difference deformation measurements (in space...
with respect to a reference point, and in time between master and slave epochs) from coherent targets. The final parameters of interest $\hat{x}$ are estimated from $y$ via the mapping $M$ derived from the available APK of deformation. Examples of such a mapping are geophysical inversion, techniques such as least-squares estimation/prediction, finite element methods, interpolation, Monte Carlo inversion and so on. Any uncertainty in the APK can be accommodated into the stochastic parameters of unmodelled deformation, $Q_s$. More $Q_s$ terms can also be added, e.g. in case of multiple deformation regimes.

Considering the case of PS as coherent targets, the VCM of observations $Q_y$ can be estimated as in [12] using, in addition to $Q_s$, the stochastic parameters of InSAR error sources [6, 10, 27, 64, 81, 84] from prior knowledge, experiments or empirically from data. Dominant InSAR error sources are spatially correlated atmospheric errors ($Q_w$), and scattering and thermal noise ($Q_n$), and hence $Q_y = Q_w + Q_s + Q_n$ from Eq. (3.6). Further terms may be added for other noise sources such as orbital and topographic errors. $Q_y$ is then propagated to the VCM of estimated parameters, $Q_{\hat{x}}$.

Figure 4.3: Schematic diagram of the proposed method. Inputs are in grey boxes, and the final output in the green box. The operation of determining the VCM of InSAR observations is in the pink box, which is repeated for $p$ device configurations. The red arrows signify error propagation from the VCM of observations to that of the parameter estimates.
4.2. Methodology through synthetic experiments

4.2.1. Criterion matrix design

The criterion matrix $Q_{\hat{x},\text{ref}}$ can be specified entirely by the user. From the user’s point of view, the variance elements of the criterion matrix are easily provided. The covariance elements, however, are also dependent on correlated InSAR error sources, and are difficult to specify. A full criterion matrix is therefore constructed as the VCM of parameter estimates corresponding to an idealized network of observations that captures the expected deformation signal within bounds of the variance limits imposed by the user. This is the FOD aspect of the proposed combined FOD-THOD method. In other words, for constructing the criterion matrix, we design a network of pseudo-PS, assuming that the PS locations are not opportunistic and that we have full control on PS locations. In this step, we disregard any existing PS, and build the most efficient network of pseudo-PS that is adequate to estimate the parameters of interest within the specified precision limits. We also ensure redundancy in the this network using the reliability measures presented in Sec. 2.1.2. The VCM of parameters resulting from this ‘optimal’ first-order pseudo-PS network is then the criterion matrix. Building the criterion matrix in this way ensures that the user-specified parameter precision limits are incorporated, while also taking into account the effect of correlated InSAR error sources.

4.2.2. Sequential optimization by network construction

In the next step, an exhaustive search space of potential device locations is created, constrained to only include locations that would improve network reliability. Both the number and the locations of devices are varied within this search space to yield $p$ different configurations. $Q_y$ is propagated to $Q_{\hat{x}}$ for the PS-only case, and likewise $Q_{\hat{x}_1}, Q_{\hat{x}_2}, ..., Q_{\hat{x}_p}$ for the $p$ device configurations, which are all compared with $Q_{\hat{x},\text{ref}}$ using the distance metric of Eq. (2.4). The $k$th configuration that yields $Q_{\hat{x}_k}$ closest to $Q_{\hat{x},\text{ref}}$ is selected as the optimal configuration. If $Q_{\hat{x}}$ from the PS-only case already satisfies the desired criteria, then no additional device needs to be deployed.

It is noteworthy that the actual InSAR observations are not required in the design process; only their (expected) locations and quality, propagated to the quality of estimated parameters, are used. In other words, we apply our analysis in the parameter space, not in the observation space. Hence, accurate quality description and error propagation are crucial.

Two synthetic examples are presented below, which differ in terms of (1) the available APK of deformation, (2) the parameters of interest, (3) the choice of deployed devices, and (4) the optimality criteria. In all examples, PS quality as well as that of the deployed devices are specified as double-differences in the vertical direction.

4.2.3. Case 1: Parametric deformation model

Expected deformation

We start with a straightforward case, where the expected ground deformation is well-known, and the estimation problem can be cast in a parametric model, e.g. of source parameters or geometric coefficients. Some real-world examples are given in [119–126]. We select the simple Mogi model [125] for demonstration; the method can be easily extended to more complex deformation models.
The ground uplift $Z$ at any point $(X, Y)$ in a $5 \times 5$ km area is expected to follow the Mogi model as:

$$Z(X, Y) = \frac{(1 - \nu) \Delta V}{\pi d^2} \frac{1}{1 + \left(\frac{R_h}{d}\right)^2}$$  \hspace{1cm} (4.1)$$

with Poisson's ratio $\nu = 0.25$, source volume change $\Delta V = 50000 \text{ m}^3$ per year, and source depth $d = 2 \text{ km}$. $R_h$ is the horizontal distance between $(X, Y)$ and the centre of the source projected to the ground surface $(X_0, Y_0)$, given by $R_h = \sqrt{(X - X_0)^2 + (Y - Y_0)^2}$. The expected cumulative uplift after 10 years (assuming linear behaviour in time) with $(X_0, Y_0) = (2, 2) \text{ km}$ is shown in Fig. 4.4a, along with the locations of existing PS.

The objective is to estimate the four Mogi model parameters $X_0, Y_0, \Delta V$ and $d$ within the user-defined precision thresholds summarized in Tab. 4.1. Transponders may be deployed, if needed.

**Table 4.1: Maximum tolerable standard deviations (1 $\sigma$) of parameter estimates.**

<table>
<thead>
<tr>
<th>#</th>
<th>Mogi parameter</th>
<th>Desired precision of estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_0$</td>
<td>$\sigma_{x_0} \leq 300 \text{ m}$</td>
</tr>
<tr>
<td>2</td>
<td>$Y_0$</td>
<td>$\sigma_{y_0} \leq 300 \text{ m}$</td>
</tr>
<tr>
<td>3</td>
<td>$\Delta V$</td>
<td>$\sigma_{\Delta V} \leq 16700 \text{ m}^3$ per year</td>
</tr>
<tr>
<td>4</td>
<td>$d$</td>
<td>$\sigma_{d} \leq 300 \text{ m}$</td>
</tr>
</tbody>
</table>

**VCM of observations**

The VCM of observations $Q_y$ is estimated using Eq. (3.6). We approximate $Q_w$ by propagating from a single SLC with a realistic atmospheric variogram [127] of range = 10 km and sill variance = 25 mm$^2$ [6, 81]. $Q_s$ can be specified with some geophysical knowledge of the area; in this example we assume that there is no unmodelled deformation, i.e. $Q_s = 0$. All other error sources are included in the $Q_n$ term, as uncorrelated noise of standard deviation = 2 mm. From [29], we use 3 mm as the double-difference measurement precision of each transponder device to be deployed.

**Precision and reliability**

Parameter estimation is performed and $\hat{Q}_x$ is determined using *best linear unbiased estimation* (BLUE, see Appendix A). From Sec. 2.1.2, 0.7 is used as the reliability threshold.

**Criterion matrix**

$Q_{\hat{x}, \text{ref}}$ (Fig. 4.6a) is constructed as the VCM of estimates arising from an ideal network of observations that captures the anticipated deformation within the desired precision and reliability thresholds. Building this ideal network can be considered as FOD, since we do not account for the actual PS locations. A sparse regular grid of pseudo-observations (ideal locations of PS) is first created. A square geometry is chosen because it performs
only slightly worse than stronger configurations such as hexagonal, but is easier to implement \[128\]. \( \hat{Q}_{\hat{x}, \text{grid}_q} \) corresponding to this grid is estimated using BLUE error propagation; the diagonal elements yield parameter variances which are compared with the thresholds of Tab. 4.1. The relative redundancy \( \bar{r}_{\text{grid}_j} \) is also determined. The grid is then progressively densified. \( Q_{\hat{x}, \text{grid}_{q}} \) and \( \bar{r}_{\text{grid}_{q}} \) are estimated for \( q \) increasing densities of pseudo-observations. The optimal pseudo-observation grid is at that sampling density \( j \) where the precision criteria of all parameters (Fig. 4.4c), as well as the redundancy criterion of \( \bar{r}_{\text{grid}_j} \geq 0.7 \) (Fig. 4.4d), are met; \( j = 1.1 \) pseudo-observations per km\(^2\) for this example. This optimal grid is illustrated in Fig. 4.4b, whose VCM of estimates is the criterion matrix shown in Fig. 4.6a. The datum of the criterion matrix is identical to that of the desired VCM of estimates.

**Search space**

A search space of possible device locations is defined, which can be limited by infeasible locations, e.g. geophysical features such as water bodies, high ambient radar clutter, political or privacy considerations. In this example, we assume no such restriction. To ensure that any possible device location contributes to the geometrical strength of the network, a redundancy map is constructed (Fig. 4.5a). This map shows the redundancy numbers \( r_l \) using Eq. (2.6) for each location in the area, accounting for the locations of existing PS. Locations with \( r_l \leq 0.95 \) \[51\] are used to create an exhaustive search space,
4. Combined First- and Third-Order Design

Figure 4.5: (a) Redundancy map; the colours indicate redundancy numbers. (b) Device search space.

Figure 4.6: VCMs (4 × 4) indicating the quality of the four estimated Mogi parameters, with the order and units given in Tab. 4.1, for three cases: (a) Corresponding to Fig. 4.4b (i.e. the criterion matrix $Q_{\text{ref}}$); $Q_{\text{ref}}$ has 0 distance to itself. (b) Corresponding to the PS-only case of Fig. 4.4a (i.e. $Q_{x}$); $Q_{x}$ has a distance of 1.15 to $Q_{\text{ref}}$. (c) Corresponding to PS and the optimal device configuration of Fig. 4.7c (i.e. $Q_{x}$); $Q_{x}$ has a distance of 0.27 to $Q_{\text{ref}}$ and is therefore closer to $Q_{\text{ref}}$ than $Q_{x}$ is. In fact, $Q_{x}$ is closer to $Q_{\text{ref}}$ than any other configuration of two devices.

Optimal number and locations of devices

The VCM of estimates $Q_{x}$ corresponding to the PS-only case of Fig. 4.4 is shown in Fig. 4.6b. Optimal locations are determined sequentially; the location of one device is first found and fixed, before the location of the next is determined. For each added device, the $k$th location that yields $Q_{x_k}$ closest to $Q_{\text{ref}}$ using the distance metric of Eq. (2.4) is selected as the optimal location. The minimum number of devices (two) is determined from Fig. 4.7a; the PS-only case (zero devices) and the case of one optimally located device yields parameter estimates that exceed at least one of the desired precision thresholds of Tab. 4.1. Fig. 4.7b shows the ranking of locations for a single device; smaller distances from the criterion matrix imply better locations. The optimal configuration of two devices is shown in Fig. 4.7c, and the corresponding VCM of estimates in Fig. 4.6c. Note that the designed network is based on parameters of interest to the user (in this case the four Mogi parameters), irrespective of which is the most suitable model to describe the deformation. If the same deformation signal is parameterized using another model, with different user-defined thresholds on the quality of the associated parameters, a different network configuration may result.
4.2. Methodology through synthetic experiments

Figure 4.7: (a) A plot showing that a minimum of 2 devices are required to satisfy all the precision thresholds of Tab. 4.1. (b) Ranking of locations for a single device; darker colours imply better locations. (c) Optimal device locations.

4.2.4. Case 2: Smoothness-constrained deformation behaviour

Expected deformation

In this case, the only available information is an assumption on the expected smoothness or spatial variability of deformation. We can express this in terms of the sill variance and the range of an isotropic variogram [127] (as shown in Fig. 4.8a) without loss of generality; in case of known directions of anisotropy, anisotropic variograms may be used. A 3 × 3 km area including the PS locations is shown in Fig. 4.8b (without the background image, because the APK cannot be visualized).

The unknown parameters in this case are surface displacements at locations of interest, which could be a particular set of localities, buildings of historical importance, and so on. Here, the locations of interest are spread over the entire area at a specified spatial resolution depicted as the grey grid in Fig. 4.8b. The objective is to install corner reflectors, if required, to determine the surface displacements at these 256 locations of interest (grey grid points). The user-defined precision threshold of determining displacements is $\sigma_{\text{req}} \leq 4$ mm. Different weights can be given to these points of interest based on their monitoring priority; here, all grid points have equal weights.

VCM of observations

$Q_y$ is constructed using Eq. (3.6) as in Sec. 4.2.3. The only difference in this case is that $Q_s$ is derived from the variogram APK of Fig. 4.8a. From [28], we use 2 mm as the double-difference measurement precision of each corner reflector.
Figure 4.8: (a) APK of deformation expressed as a spatial isotropic variogram with range = 20 km and sill variance = 50 mm$^2$. (b) PS locations. The grey grid denotes points of interest where surface displacements are to be determined.

**Precision and reliability**

To determine the quality of interpolated surface displacements at the 256 points of interest, we use *best linear unbiased prediction* (BLUP) [43], in the special case of no available functional model (see Appendix A). The prediction quality, given by the VCM of the prediction error $\hat{Q}_\epsilon$, is used for network design, playing the role of $Q_x$ in the previous example. From Eq. (2.6), reliability cannot be used as an optimization criterion when the functional model is unknown. The device search space therefore cannot be refined based on a redundancy map, and consists of a dense grid of possible device locations spread throughout the area.

**Criterion matrix**

The $256 \times 256$ criterion matrix of prediction quality, $Q_{\epsilon_0,\text{ref}}$, is constructed as in Sec. 4.2.3 by gradual densification of a sparse uniform grid of measurements until the user-defined precision criterion is achieved. The VCM of this optimal pseudo-observation grid shown in Fig. 4.9a is the criterion matrix.

**Optimal number and locations of devices**

Following the same procedure as in Sec. 4.2.3, we find that placing one to four devices in any configuration causes the displacement at some of 256 points of interest to have a prediction error exceeding the requirement of 4 mm. The ranking of locations for one device is shown in Fig. 4.9c. When five devices are placed at the locations shown in Fig. 4.9b, the VCM of prediction error is closest to the criterion matrix and the displacement predicted at all points of interest fall within the precision requirement, making this the optimal network of devices.

In this section, the proposed network design method has been demonstrated for two cases: where a parametric (functional) model is available as APK, and where only smoothness (stochastic) information is available. In reality, there is quite often some amount of both functional and stochastic knowledge available. In the following section, we design a network for the case where rough parametric APK is available, along with some knowledge of the expected stochastic deviation from it. We also use this example
4.3. Optimality demonstration

In this section, we consider a case where the expected deformation is expressed as a parametric model. However, it is known that this model is non-ideal, and the possible deviation from the model is expressed stochastically. To demonstrate that the designed configuration is optimal \textit{a posteriori}, we use a Monte Carlo approach where deformation signal reconstructions derived from several different stochastic realizations of the observations are compared to the true deformation.

A vertical cross-section of deformation (uplift) along a 10 km profile is used as the functional model, for faster processing of a large number of Monte Carlo simulation runs, as well as better visualization of the resulting signal reconstructions. This uplift (green curve in Fig. 4.10) at any point $X$ is given by an isotropic Gaussian function

$$Z(X) = Z_{\text{max}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(4.2)

where $Z_{\text{max}} = 30$ mm, $\mu = 5$ km and $\sigma = 1$ km denote the peak (maximum deformation), mean (centre) and standard deviation (extent) of the Gaussian curve, respectively. The deviation from this model, i.e. the stochastic unmodelled deformation, is based on a variogram of sill variance = 4 mm$^2$ and range = 10 km. An instance of this is added to the functional model, and we treat this as the true ground deformation, shown as the red curve in Fig. 4.10 along with the existing PS. The objective is to predict the surface displacement at every location along the profile with a standard deviation of $\sigma_{\text{req}} \leq 4$ mm.

Transponders of 3 mm precision [29] may be deployed, if required.

Standard BLUP is used to predict displacements along the 10 km profile, $\hat{y}_0$, and to determine the VCM of prediction error, $Q_{\epsilon_0}$ (see Appendix A). We follow the remaining steps of Sec. 4.2, and the intermediate results are shown in Figs. 4.11, 4.12 and 4.15.

Cost optimization is performed based on Figs. 4.13a-c. The requirement of $\sigma_{\text{req}} \leq 4$ mm at every location along the profile is satisfied using a minimum of three devices (red line in Fig. 4.13b), with their optimal locations shown in Fig. 4.14c. Figs. 4.13a and b can aid

Figure 4.9: (a) Pseudo-observation grid (FOD) with optimal density (2.8 samples per km$^2$). (b) Optimal locations of 5 devices. (c) Ranking of best locations for a single device; darker colours imply better locations.

to confirm that the optimal network designed \textit{a priori}, i.e. before acquiring any actual measurements, is also optimal \textit{a posteriori}. 

4.3. Optimality demonstration

In this section, we consider a case where the expected deformation is expressed as a parametric model. However, it is known that this model is non-ideal, and the possible deviation from the model is expressed stochastically. To demonstrate that the designed configuration is optimal \textit{a posteriori}, we use a Monte Carlo approach where deformation signal reconstructions derived from several different stochastic realizations of the observations are compared to the true deformation.

A vertical cross-section of deformation (uplift) along a 10 km profile is used as the functional model, for faster processing of a large number of Monte Carlo simulation runs, as well as better visualization of the resulting signal reconstructions. This uplift (green curve in Fig. 4.10) at any point $X$ is given by an isotropic Gaussian function

$$Z(X) = Z_{\text{max}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(4.2)

where $Z_{\text{max}} = 30$ mm, $\mu = 5$ km and $\sigma = 1$ km denote the peak (maximum deformation), mean (centre) and standard deviation (extent) of the Gaussian curve, respectively. The deviation from this model, i.e. the stochastic unmodelled deformation, is based on a variogram of sill variance = 4 mm$^2$ and range = 10 km. An instance of this is added to the functional model, and we treat this as the true ground deformation, shown as the red curve in Fig. 4.10 along with the existing PS. The objective is to predict the surface displacement at every location along the profile with a standard deviation of $\sigma_{\text{req}} \leq 4$ mm. Transponders of 3 mm precision [29] may be deployed, if required.

Standard BLUP is used to predict displacements along the 10 km profile, $\hat{y}_0$, and to determine the VCM of prediction error, $Q_{\epsilon_0}$ (see Appendix A). We follow the remaining steps of Sec. 4.2, and the intermediate results are shown in Figs. 4.11, 4.12 and 4.15.

Cost optimization is performed based on Figs. 4.13a-c. The requirement of $\sigma_{\text{req}} \leq 4$ mm at every location along the profile is satisfied using a minimum of three devices (red line in Fig. 4.13b), with their optimal locations shown in Fig. 4.14c. Figs. 4.13a and b can aid
Figure 4.10: True deformation (red curve), and PS locations. The green curve is the Gaussian functional model, based on Eq. (4.2).

Figure 4.11: Determining the optimal sampling based on (a) the precision criterion of $\sigma_{\text{req}} \leq 4$ mm, and (b) the relative redundancy criterion of 0.7. The latter sets denser sampling (1 per km) in this case. (c) The resulting pseudo-observations (FOD), whose VCM of prediction error is the criterion matrix.
in decision-making: e.g., if it is acceptable for the error in 6% of the predicted displacements to be marginally higher than 4 mm, two devices are adequate. Fig. 4.13d shows the ranking of locations for a single device.

To demonstrate optimality, we compare the designed device locations with three other locations chosen intuitively within the largest PS gap (blue squares in Fig. 4.14b): at the peak, at where the curve bends away, and at a point in between. $10^4$ different noise realizations of the observations are created, using $Q_y$ estimated as in Sec. 4.2.3, also accounting for the unmodelled deformation. For the three cases of (a) PS only, (b) PS and device locations selected intuitively, and (c) PS and the designed device locations (Figs. 4.14a-c respectively), the deformation signal is reconstructed using BLUP (see Appendix A) for each of the $10^4$ observation realizations, shown as the green curves in Figs. 4.14d-f. The black curve is the mean of the green curves, which is compared with the ground-truth red curve.

We define two performance metrics to quantify the reconstruction quality in Figs. 4.14d-f; these metrics are empirical, and unrelated to the distance metric used in the design process, for independent validation. The first is the root-mean-squared error (RMSE) between the reconstruction curves and the true deformation. The second is the norm in the metric of the criterion matrix, calculated per reconstruction as $\hat{\epsilon}_0^T Q_{\epsilon_0,\text{ref}}^{-1} \hat{\epsilon}_0$, where $\hat{\epsilon}_0$ is the vector of residuals with respect to the true deformation, i.e. $\hat{\epsilon}_0 = y_0 - \hat{y}_0$. The mean of these norms gives an indication of how well the reconstructions perform with respect to the desired quality of the parameter estimates, i.e. the criterion matrix. It is observed from Figs. 4.14d-f through both performance metrics that (1) deploying three devices yields better reconstruction quality than the PS-only case, and (2) the locations designed by the proposed method lead to better signal reconstruction than those chosen intuitively.

In the next section, the proposed method is applied on a real-world case of ground subsidence in the Netherlands.
Figure 4.13: (a) A plot showing that 3 devices are required so that no point along the profile has a prediction error greater than 4 mm (b) Variation of prediction error (standard deviation) with number of devices. At 3 devices, the worst prediction error (red) is less than the threshold of 4 mm (green). The blue line shows the average prediction error of all points in the profile. (c) Distance of prediction error VCM from criterion matrix vs. number of devices; minimum distance is at 3 devices. (d) Ranking of optimal locations for a single device; darker colours imply better locations.
4.3. Optimality demonstration

Figure 4.14: (a) PS only. (b) PS and device locations selected intuitively. (c) PS and optimal device locations. The signal reconstructions from (a), (b) and (c), shown in (d), (e) and (f) respectively, based on $10^4$ different realizations of observation sets, are the green curves, their average the black curves, and true deformation the red curves. Based on both the performance metrics (mean RMSE and norm in the metric of $Q_{\epsilon_0,\text{ref}}$), configuration (c) gives the best signal reconstruction.

Figure 4.15: (a) Criterion matrix $Q_{\epsilon_0,\text{ref}}$ based on the pseudo-observations of Fig. 4.11c, and VCMs of prediction error at all the points in the profile corresponding to (b) PS only, (c) device locations selected intuitively, and (d) the optimal device locations. From the distance metrics, (d) is closer to (a) than (b) and (c) are. The colourbars are shown equal for comparison; the maximum variance in (b) is 534 mm².
4. Combined First- and Third-Order Design

Figure 4.16: (a) PS deformation rate map of the Roswinkel gas field in radar LoS direction using archived SAR data (ERS descending track, 1992–2003, 71 SAR images). (b) APK of deformation, with PS locations derived from (a).

4.4. Case study: Roswinkel subsidence

Gas was extracted from the Roswinkel field in the Netherlands from 1980–2005, during which subsidence was measured through nine levelling campaigns, and also through InSAR (Fig. 4.16a). The objective here is to design an optimal network of corner reflectors, if required, for a future InSAR monitoring requirement, e.g. if gas extraction were to resume. Reservoir engineers [129] have deduced the approximate shape of the gas reservoir (Fig. 4.17), which can be approximated by a Gaussian model [126] as:

\[
Z(X, Y) = Z_{\text{max}} e^{-0.5r^2},
\]

where

\[
r^2 = \left(\frac{(X - X_0) \sin \alpha + (Y - Y_0) \cos \alpha}{a}\right)^2 + \left(\frac{(X - X_0) \cos \alpha - (Y - Y_0) \sin \alpha}{b}\right)^2.
\]

\[(X_0, Y_0) = (9, 11) \text{ km} \]

are the coordinates of the bowl centre in X and Y directions, \(a = 6.5 \text{ km} \) and \(b = 3.1 \text{ km} \) the lengths of the semi-major and semi-minor axes respectively, \(\alpha = 59^\circ \) the bearing angle of the semi-major axis (with respect to +Y-axis), \(Z_{\text{max}} = 96 \text{ mm} \) the peak cumulative subsidence, and \(\delta = 3 \) the bowl shape parameter. This APK of anisotropic deformation and the locations of existing PS (from archived SAR data) are shown in Fig. 4.16. The Gaussian bowl parameters have been chosen accounting for uncertainties in our knowledge of the local subsurface and deformation mechanisms. We have therefore built a worst-case scenario of the expected subsiding area and magnitude in order to design a conservative network of devices; it is highly unlikely that the deformation will exceed this in future.

The aim is to estimate the Gaussian bowl parameters from the InSAR data within the user-defined precision thresholds of Tab. 4.2. We follow the steps described in the BLUE case of Sec. 4.2.3; the intermediate results are shown in Figs. 4.18 and 4.19.

Three corner reflectors are required to satisfy all the precision criteria (Fig. 4.20a); their optimal locations are shown in Fig. 4.20b. Using less or more reflectors could depend on the available budget, or on changes to the precision requirements of Tab. 4.2. Fig. 4.21 shows the optimal locations for one, two or four corner reflectors.
4.4. Case study: Roswinkel subsidence

Figure 4.17: A map showing the shape of the Roswinkel gas reservoir (pink) and the zone where ground deformation is expected (translucent white), based on geological modelling, reservoir simulation and geomechanical modelling, as well as an inverse procedure that used geodetic information on the subsidence, predictive models, and prior knowledge on the subsurface characteristics to constrain the uncertainties of the reservoir and geological parameters [129].

Figure 4.18: Determining the optimum sampling based on (a) the precision criteria of Tab. 4.2, and (b) the relative redundancy criterion of 0.7. The former sets denser sampling (0.7 per km²) in this case. (c) The resulting pseudo-observation grid, whose VCM of estimates is the 7 × 7 criterion matrix.
Figure 4.19: (a) Redundancy map. (b) Device search space. (c) Ranking of best locations for a single corner reflector; darker colours imply better locations.

Figure 4.20: (a) A plot showing the minimum number of corner reflectors (i.e. 3) required to satisfy the thresholds in Tab. 4.2 (b) Optimal corner reflector locations.

Figure 4.21: Optimal locations for (a) 1, (b) 2, or (c) 4 corner reflectors.
4.5. Summary and discussion

We have proposed a combined FOD-THOD method for InSAR that determines the minimum number and optimal locations of ‘artificial’ coherent targets (e.g. corner reflectors or transponders) to augment an opportunistic but sparse ‘natural’ coherent target network. The design can be performed before any actual InSAR acquisitions are available, given the expected locations of coherent scatterers, their quality, accurate error propagation from measurements to estimates, and the available a priori knowledge of the deformation characteristics, if any. This knowledge can vary from the trivial case of no prior information, to the case where merely the smoothness of expected deformation is known, to the case where a large amount of information is available and educated assumptions of deformation models are possible.

The design method is oriented towards the practical needs of the end-user of InSAR-derived deformation estimates. Therefore, the primary objective is not to find the best possible model to describe the deformation and its mechanisms, but to achieve closeness to user-defined criteria of deformation monitoring quality. The difference is that the former would entail minimization of the estimation error, which would yield the trivial result that deploying more devices would always reduce the estimation error further. In contrast, our approach gives the minimum number of additional measurements required to satisfy user-needs, without over-designing the network.

At its core, the algorithm compares full VCMs of the estimated parameters to a criterion matrix, which is based on the desired quality of parameter estimates. Variances and covariances of individual parameters can thus be tuned according to user requirements. In case only variance thresholds are specified by the user, the full (and more realistic) criterion matrix is constructed from these, accounting for the effect of spatially correlated InSAR error sources and unmodelled deformation on the quality of parameter estimates.

An implementational consideration is the effect of a highly non-linear relationship between the observations and estimated parameters. In such a case, model linearization could produce artefacts which impede correct error propagation from observations to estimates. An alternative way to solve such a highly non-linear problem is through Monte Carlo error propagation, at the cost of computational time and resources.

### Table 4.2: Maximum tolerable standard deviations (1σ) of parameter estimates.

<table>
<thead>
<tr>
<th>#</th>
<th>Gaussian parameter</th>
<th>Desired precision of estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_0$</td>
<td>$\sigma_{x_0} \leq 100$ m</td>
</tr>
<tr>
<td>2</td>
<td>$Y_0$</td>
<td>$\sigma_{y_0} \leq 100$ m</td>
</tr>
<tr>
<td>3</td>
<td>$a$</td>
<td>$\sigma_a \leq 200$ m</td>
</tr>
<tr>
<td>4</td>
<td>$b$</td>
<td>$\sigma_b \leq 200$ m</td>
</tr>
<tr>
<td>5</td>
<td>$\alpha$</td>
<td>$\sigma_{\alpha} \leq 1^\circ$</td>
</tr>
<tr>
<td>6</td>
<td>$\delta$</td>
<td>$\sigma_{\delta} \leq 0.11$</td>
</tr>
<tr>
<td>7</td>
<td>$Z_{\text{max}}$</td>
<td>$\sigma_{Z_{\text{max}}} \leq 4$ mm</td>
</tr>
</tbody>
</table>
Zero-Order Design

If the only tool you have is a hammer, 
you treat everything as if it were a nail. 
Abraham Maslow

InSAR is used to estimate double-difference scatterer displacements (in space and in time) based on complex-valued radar reflections stemming from these scatterers. When used in a time-series approach, the double-difference displacements can be generalized to, e.g., single-difference (space-only) velocity estimates, or displacement behaviour parameterized via polynomial coefficients [6, 9]. Whatever the parameterization is, a key property is that the estimates are inherently relative. If we focus on changes in elevation (or in the radar LoS), the estimates are always relative to a reference point with a particular location (or, equivalently, the average of scatterers within a reference area), and a reference epoch. The value assigned to the InSAR reference point can be arbitrary; often, for convenient interpretation of the displacement estimates, a ‘conventional’ value is chosen, e.g. zero displacement or velocity for a reference point situated in an area assumed stable. In these cases the value is deterministic and has zero standard deviation. When InSAR studies use a point or area ‘assumed to be stable’ as reference, the interpretation of the deformation results are subject to this assumption.

Although effectively all geodetic positioning techniques are relative, some applications require ‘absolute’ InSAR displacement estimates, i.e., estimates which can be expressed in a well-defined terrestrial reference frame (TRF), and therefore related to the results of other techniques. Whenever two or more sets of estimates stem from specific reference frames or datums, there is a need for datum connection [130]. Here we propose a method for datum connection, to merge the positioning results of InSAR surveys and studies referring to a global TRF with associated noise variance-covariance matrices. In this method, the value for the InSAR reference point displacement is obtained from other techniques, e.g. GNSS, in a well-defined TRF. The displacement of the InSAR reference point in this TRF is then stochastic, and the errors can be propagated. In the
remainder of this chapter, we will use the term ‘absolute’ deformation to mean displacement expressed in a well-defined TRF.

Ground deformation can be represented as a vector displacement field, which is dependent on both the physical displacement and the reference frame or datum. The relevant geodetic theory on reference frames and datums is summarized in Appendix C. InSAR deformation estimates have a local datum, since they are always with respect to a reference point/area within the SAR image. To represent all the InSAR-derived deformation estimates in a TRF, the local datum should be connected to a global one. For one-dimensional deformation monitoring such as by InSAR (in $z$-direction, or equivalently, in radar LoS direction), the 14 datum parameters of translation, rotation, scale and their respective time derivatives reduce to only the translation parameter and its derivative. The estimation of these two parameters is the focus of this chapter.

In our framework of geodetic network design, we term the transformation of InSAR deformation estimates from a specific local datum to a global TRF as zero-order design (ZOD). This chapter describes (1) the mathematical approach for connecting InSAR to a TRF; (2) experimental results with collocated InSAR-GNSS measurements, and (3) a feasibility study on a national scale, extensible to a European and a global scale. We have performed this study at selected permanent GNSS stations in the Netherlands using Radarsat-2 data.

Datum connection for InSAR is described in Sec. 5.1. Sec. 5.2 compares two datum connection scenarios and shows the benefit of InSAR-GNSS measurement collocation. The experimental setup and results of the datum connection method, in particular at Ijmuiden where InSAR and GNSS measurements are collocated at a tide-gauge, are discussed in Sec. 5.3. ‘Absolute’ displacement and height results are presented in Sec. 5.3.3, along with a description of the various factors considered for precise datum connection. A collocation feasibility map of all the permanent GNSS stations in the Netherlands is presented in Sec. 5.3.5, towards implementing the proposed approach on a countrywide scale. Sec. 5.4 concludes the chapter, with an overall analysis and discussion.

5.1. Datum connection for InSAR

In time-series InSAR, particularly in the PSI approach, differential deformation parameters are estimated along arcs joining pairs of scatterers, at two or more epochs in time. These spatio-temporal differences for a single arc are referred to as double differences, and are the first physically meaningful parameters in the estimation of scatterer displacements or velocities. By solving for all identified arcs, and integrating the estimates, we can join them in a common datum. If we only focus on changes in elevation (or in the radar LoS), this datum is defined by a basis — a reference point (or the average of several/all points), and a reference epoch.

If only single-differences in the temporal sense are considered, the satellite state vectors and the derived satellite orbit models during the SAR image acquisition can be used to transform from the radar coordinates of InSAR measurements to a TRF, using additional information from the Doppler centroid (defining the squint angle), the local earth radius (from the ellipsoid) and the topographic height [6]. However, orbit models currently do not have adequate precision to reach geodetic precision requirements. This necessitates the use of double differences, i.e. a reference point with respect to which
5.1. Datum connection for InSAR

height differences can be precisely estimated.

Mathematically, the InSAR-derived double-difference deformation at location \( p_1(x_1, y_1) \) and (slave) time \( t_1 \), with respect to reference location \( p_0(x_0, y_0) \) and reference (master) time \( t_0 \), is given by

\[
\Delta p \Delta t h = [h(p_1, t_1) - h(p_1, t_0)] - [h(p_0, t_1) - h(p_0, t_0)] \tag{5.1}
\]

where \( h \) here denotes the height in any arbitrary datum. \( h(p_0, t_1) \) can be rewritten as

\[
h(p_0, t_1) = h(p_0, t_0) + v(p_0)(t_1 - t_0) \tag{5.2}
\]

where \( v(p_0) \) is the velocity of the reference point, assuming linear displacement between master and slave acquisitions. The quantity of our interest is \([h(p_1, t_1) - h(p_1, t_0)]\), which is the ‘absolute’ deformation of point \( p \) between epochs \( t_1 \) and \( t_0 \). Using Eq. (5.2) in Eq. (5.1) and rearranging, we get

\[
[h(p_1, t_1) - h(p_1, t_0)] = \Delta p \Delta t h + v(p_0)(t_1 - t_0). \tag{5.3}
\]

The ‘absolute’ deformation can therefore be estimated by knowing the double difference value \( \Delta p \Delta t h \) which is obtained from InSAR, as well as the velocity of the reference point \( v(p_0) \). InSAR is blind to the velocity of the reference point. Large-scale motion affecting the entire area of interest including the reference point in the same way will not be detected at all, and local motion affecting only the reference point would manifest itself in all the other deformation estimates. In order to determine \( v(p_0) \), two approaches may be considered:

1. a **geophysical approach**, where the velocity of the reference point is assumed from geophysical knowledge. In such a case, \( v(p_0) \) is then ‘known’ (e.g. by using a displacement model, or assuming the local stability of well-founded buildings), and this value is deterministic with zero standard deviation. Such an approach is subjective, and quality assessment or control difficult; error margins should therefore be chosen conservatively.

2. a **geodetic approach**, where \( v(p_0) \) is estimated by connecting the InSAR reference point to additional (collocated) geodetic measurements, e.g. a GNSS antenna. The reference point displacement and height (in a standard TRF) are then stochastic, and the associated errors can be propagated. All the other InSAR-derived displacements can subsequently be defined in the same TRF.

We choose the geodetic approach for ZOD, and proceed to describe it further. As an aside, note that the InSAR-derived deformation in Eq. (5.1) has been converted from the radar LoS direction to vertical, since heights and height-changes are in the local vertical direction. Such a conversion requires InSAR data from at least two different looking directions of the satellite, or a well-known model of the horizontal deformation. In the absence of either, since InSAR is more sensitive to vertical deformation than to horizontal, the latter is often ignored, resulting in an error in the estimated vertical component depending on the magnitude and orientation of horizontal deformation [131]. If ‘absolute’ LoS deformation is desired, Eq. (5.3) is still applicable. \( v(p_0) \) is then estimated in the LoS direction; this is possible e.g. using GNSS, which is capable of precise 3D positioning.
5. Zero-Order Design

5.1.1. S-transformations

InSAR phase observations are stochastic and are frequently assumed to be normally distributed \([6, 64]\). Consequently, the estimated displacements are stochastic as well, with unknown expectation values to be estimated from the sample measurement values. The precision of these displacements can be described completely by their VCM. Considering PSI, the displacement of a certain PS not only has a variance, but also covariances in the space-time domain — in space with nearby PS (owing to e.g. atmospheric errors), and in time with nearby points in the PS time-series (due to e.g. orbital errors). Therefore, PS that are closer to the reference point have smaller variances than those farther away, making the VCM of PS displacements depend on the choice of the reference point.

The theory of similarity transformations or S-transformations \([33, 35, 69, 70]\) can be used to transform deformation estimates from one datum to another, and to give the relation between VCMs arising from the choice of different reference points. If \(y_1\) is the \(m\)-vector of displacements defined in a reference system or datum \(D_1\), and \(y_2\) the vector of the same observations but defined in another datum \(D_2\), the transformation from \(y_1\) to \(y_2\) can be represented by a linear transformation \(S\) as

\[
y_2 = S y_1,
\]

and the associated quality of observations, denoted by the VCMs \(Q_{y_1}\) and \(Q_{y_2}\) respectively, can be propagated as

\[
Q_{y_2} = S Q_{y_1} S^T.
\]

The generic form of \(S\), called the the S-transformation matrix, is given by \([33, 35, 69, 70]\)

\[
S = I - H(D_2^T H)^{-1} D_2^T
\]

with \(I\) being an \(m \times m\) identity matrix. \(D_i\) is a matrix that defines the datum in each case \(i\), i.e.

\[
D_1^T y_{1 \bot} = 0 \text{ and } D_2^T y_{2 \bot} = 0
\]

for the two datums respectively.

For the one-dimensional deformation monitoring case of InSAR, \(D_i\) is a binary vector with the value of unity at the location(s) of the reference point(s). As an example, if the displacement at the \(k\)th location is assigned as the transformed reference point, then \(D_2 = [0 \ 0 \ \ldots \ 1 \ \ldots \ 0]^T\), with the value of unity at the \(k\)th position. The vector \(H\) is derived from the inner constraint of the measurement network, according to which the average displacement is set to zero \([33, 35, 69]\); such a constraint is necessary to set the datum of a free network. In other words, \(H\) is the most generic case of \(D_i\) where the average of all displacements is taken as reference, i.e. \(H = [1 \ 1 \ \ldots \ 1 \ \ldots \ 1]^T\).

We proceed to apply this theory to InSAR-GNSS datum connection via two practical scenarios in the following section. Further treatment on S-transformations can be found in \([33, 35, 69, 70]\).
5.2. Practical scenarios

In order to connect InSAR to the GNSS datum using the geodetic approach, it is required that the InSAR reference point and the corresponding GNSS benchmark reflect the same deformation signal. Two scenarios are presented here. The first one is opportunistic, where we make use of the existing PS within a reference area (i.e. a certain radius from the GNSS measurement), assuming that these PS measure the same deformation signal as the GNSS antenna (Sec. 5.2.1). In the second scenario, we deliberately collocate a PS with a GNSS antenna, to eliminate any assumptions on the requirement of both techniques measuring identical signals at the reference point (Sec. 5.2.2).

5.2.1. Reference PS around a GNSS antenna

Consider a scenario where InSAR analysis of a certain area yields a PS deformation map in the vertical direction as shown in Fig. 5.1a. The deformation at each PS point is with respect to a reference point; for simplicity, we consider the cumulative deformation at a single epoch with respect to the InSAR master time epoch. This map also contains a GNSS antenna, the position of which is marked with a cross. The associated spatial VCM of deformation estimates shown in Fig. 5.1e, with the PS sorted in the order of increasing distance from the GNSS antenna.

There are a few PS in the vicinity of the GNSS antenna. In the absence of any prior information regarding possible relative deformation between the GNSS antenna and its surroundings, we assume that the deformation given by average of the PS within a small radius around the GNSS antenna is equal to that measured by GNSS; this is the most simple approach, which can be improved, e.g., by deterministic or stochastic interpola-
With this assumption, we proceed to connect the InSAR survey results to the GNSS datum as follows:

1. The InSAR survey results $y_{\text{initial}}$ with the initial reference point are first transformed to another datum $D_{\text{trans}}$ defined by the reference area marked with the circle (Fig. 5.1b). Using the notation of Eqs. (5.5), (5.6) and (5.7),

$$S_{\text{trans}} = I - H(D_{\text{trans}}^T H)^{-1} D_{\text{trans}}^T \quad (5.8)$$

$$y_{\text{trans}} = S_{\text{trans}} y_{\text{initial}} \quad (5.9)$$

with the unity values in $D_{\text{trans}}$ corresponding to the PS within the reference area. The quality of deformation estimates is propagated to the VCM shown in Fig. 5.1f by

$$Q_{y_{\text{trans}}} = S_{\text{trans}} Q_{y_{\text{initial}}} S_{\text{trans}}^T. \quad (5.10)$$

2. The transformed InSAR survey results $y_{\text{trans}}$ are then connected to a standard TRF, by adding the GNSS measurement to all the PS and propagating the errors of both techniques, i.e.

$$y_{\text{TRF}} = y_{\text{trans}} + H y_{\text{GNSS}} \quad (5.11)$$

where $y_{\text{GNSS}}$ is the GNSS measurement which can be linked to a standard TRF. In this example, $y_{\text{GNSS}}$ is a subsidence value of 13.5 mm with respect to the GNSS measurement acquired at the InSAR master time epoch. Linear least-squares error propagation yields

$$Q_{y_{\text{TRF}}} = Q_{y_{\text{trans}}} + HQ_{y_{\text{GNNS}}} H^T \quad (5.12)$$

where $Q_{y_{\text{GNNS}}}$ contains the variance of the GNSS measurement, taken to be 16 mm$^2$ in this example. Figs. 5.1c and 5.1g show the PS map connected to the GNSS datum and the associated quality, respectively. This PS map shows the ‘absolute’ deformation of every point, including the reference points, since the InSAR master time epoch.

It can be shown analytically that the double-difference information content in all the datums (Figs. 5.1e-g) is identical, irrespective of the GNSS measurements and their quality. Let $y_{\text{initial}}$ and $\hat{y}_{\text{initial}}$ denote respectively the original double-difference InSAR survey results with respect to the initial reference point, and the results transformed back to the initial reference point after GNSS datum connection. To show that $\hat{y}_{\text{initial}} = y_{\text{initial}}$, we make use of the following properties of S-transformations for any two datums $D_1$ and $D_2$ [69]:

1. $S_1 y_{\text{initial}} = y_{\text{initial}}$: measurements transformed to their own datum remain the same.
2. $S_1 S_1 = S_1$: transforming to the same datum twice is equivalent to doing so only once.
3. $S_1 S_2 = S_1$: transforming to a datum via another datum is equivalent to directly transforming to the final datum.
Using Eq. (5.4) and the equations in Sec. 5.2, the back-transformed set of measurements are given by

\[ \hat{y}_{\text{initial}} = S_{\text{initial}} y_{\text{TRF}} \]

\[ = S_{\text{initial}} (S_{\text{trans}} y_{\text{initial}} + H y_{\text{GNSS}}) \]  

(5.13)

(5.14)

From properties (1) and (3), \( S_{\text{initial}} S_{\text{trans}} y_{\text{initial}} = y_{\text{initial}} \). By expanding \( S_{\text{initial}} \) using Eq. (5.6),

\[ \hat{y}_{\text{initial}} = y_{\text{initial}} + [H - H(D_{\text{initial}}^T H)^{-1} (D_{\text{initial}}^T H)] y_{\text{GNSS}} \]  

(5.15)

\[ = y_{\text{initial}} \]  

(5.16)

which also implies that

\[ Q_{\hat{y}_{\text{initial}}} = Q_{y_{\text{initial}}} \].  

(5.17)

This is visualized in Figs. 5.1d and 5.1h, where transforming \( y_{\text{TRF}} \) back to the initial reference system gives results identical to Figs. 5.1a and 5.1e. In other words, by connecting PS displacements to a TRF via GNSS, the relative displacements with respect to the reference point remain exactly the same, with an identical VCM. The added value is that these spatially-relative displacements can be converted into ‘absolute’ height changes with respect to a TRF. The VCM of these ‘absolute’ heights then also includes the quality of GNSS measurements.

Some advantages of using the PS in a reference area around the GNSS antenna are that the approach is opportunistic, i.e. does not require additional hardware installation, and that averaging the PS in the reference area can cancel out some noise effects such as scattering or thermal noise. However, there are also some limitations. Firstly, it is not guaranteed to have PS in the vicinity of a GNSS antenna, especially in highly decorrelating non-urbanized areas. Moreover, the (strong) assumption that the PS points near a GNSS antenna exhibit the same displacement as that measured by GNSS does not always hold in practice, because of the presence of local variations, e.g. [71,132]. Sometimes, even two different parts of the same building may deform differently [26, 133–138]. Also, the apparent deformation of a nearby radar scatterer is not necessarily the relevant deformation measured by the GNSS antenna. PS can often result from multiple-bounce building-ground (dihedral or trihedral) reflections, meaning that even if the building on which the GNSS antenna is mounted happens to be a PS, the deformation estimate may include a component from the differential building-ground motion because of swelling/compaction in the surrounding soil [71, 139–141]. Therefore, in order to perform rigorous datum connection between InSAR and GNSS, it is important to know the location of the effective scattering phase centre of the PS reference, and ensure that there is no relative deformation between this phase centre and the GNSS antenna. Such a scenario is presented in the following subsection.

### 5.2.2. Reference PS collocated with a GNSS antenna

In this scenario, a stable radar scatterer such as an active transponder [29, 142] is physically connected to a GNSS antenna, as shown in Figs. 5.2 and 5.3. By doing so, it is ensured that the InSAR reference point and the GNSS antenna experience the same deformation, owing to the rigid physical connection between them.
Transponders are small, compact and lightweight devices, with phase stability comparable to that of corner reflectors, as shown in Chapter 2. They are currently usable with C-band radar satellites (e.g. Radarsat-2, Sentinel-1). They can be configured for both ascending and descending passes of multiple satellites using the same rigid installation setup. Their small size, low mass and convenient shape makes it easy to mechanically attach them to GNSS antenna installations. They also pose less risk than corner reflectors in terms of introducing multipath effects to the GNSS measurements. They are less sensitive to environmental effects (e.g. wind, rain, snow) than corner reflectors, and therefore require lower maintenance effort.

Consider the cumulative deformation map at a single epoch, as described in the previous example. This is shown again in Fig. 5.4a, with the associated VCM of deformation estimates in Fig. 5.4e. The difference here is that the GNSS antenna has a collocated PS in the form of a transponder, with double-difference measurement precision taken to be 1 mm (a reasonable estimate, as shown in Sec. 3.5.4).

We arrive at the ‘absolute’ deformation map of the area in a TRF, as explained in
Sec. 5.2.1. The datum connection procedure uses Eqs. (5.8) to (5.12), setting the unity value in $D_{\text{trans}}$ at the location corresponding to the transponder. The intermediate results of datum transformation to the transponder as reference point are shown in Figs. 5.4b and 5.4f, and those of datum connection to GNSS in Figs. 5.4c and 5.4g. In this example too, Figs. 5.4d and 5.4h confirm that the double-difference information is preserved during datum transformation and connection, irrespective of transponder and GNSS measurements and quality.

There are several advantages of this scenario. Firstly, and most importantly, the location of the phase centre of the reference PS (i.e. the transponder) is known [142] to be within the transponder box, ruling out the effects introduced if the reference PS phase centre is unknown (e.g. local deformation variations and multiple-bounce reflections as described in Sec. 5.2.1). Note that there should ideally be no other dominant radar scatterer (or PS) within the resolution cell of the transponder, since otherwise the effective phase centre (vector sum of the phasors resulting from the transponder and the existing scatterer) would once again be unknown. This can be ensured by choosing the candidate GNSS antenna that is not already a PS and does not have bright ambient clutter, or by installing a portable collocated transponder-GNSS unit (e.g. I2GPS [36, 37], Fig. 5.3) in a decorrelating area. Further advantages of this scenario are that the sturdy connection between the transponder and the GNSS antenna ensures that they both experience the same deformation; no assumptions are required. Additionally, this approach can perform datum connection between InSAR survey results stemming from different satellites and looking directions, simply by programming the transponder accordingly; this is in contrast with the previous opportunistic scenario where the PS distribution and characterisation in the reference area could differ among various satellites and tracks. A limitation of the collocation approach is that it relies on a single-point connection between
InSAR and GNSS datums, i.e. there is less redundancy in case of transponder hardware or power supply failure. Multiple collocated transponder-GNSS installations can be a way to remedy this.

The simplified simulated scenarios described in this section have been provided as a proof of concept; they can be extended to time-series deformation incorporating temporal variance-covariance information, as well as to multi-track and multi-GNSS cases. In the following section, we study the case of IJmuiden in the Netherlands, where a transponder has been permanently attached to a GPS station since 2012. We also explore similar setups in Eijsden and Vlissingen, and finally determine a set of feasible GPS stations across the country, that can be used with InSAR for datum connection.

5.3. InSAR-GNSS datum connection in the Netherlands

5.3.1. Experiment objectives

To explore the feasibility and the practical aspects of the proposed datum connection approach, an experiment has been set up, involving different permanent GPS stations across the Netherlands. The main aims of this experiment are

1. to determine the feasibility of the transponder-GNSS measurement collocation approach for datum connection.
2. to gain experience in the practical considerations of collocating a PS with a GNSS measurement.
3. to acquire a long time-series of SAR data after transponder-GNSS collocation, perform InSAR deformation analysis, connect the resulting estimates to a standard height datum such as the NAP, and analyse this ‘absolute’ deformation map.

5.3.2. GNSS stations and SAR data

Permanent GNSS (GPS) stations are spread throughout the Netherlands; we consider the Active GPS Reference System (AGRS) and TU Delft stations, some of which are International GNSS Service (IGS) and EUREF Permanent Network (EPN) stations, as well as the Netherlands Positioning Service (NETPOS) stations. The Netherlands is covered by three Radarsat-2 tracks (descending orbit): t102, t202 and t302. Fig. 5.6 shows the SAR coverage and the locations of the Dutch permanent GPS stations. Compact active transponders (CATs) are collocated with GPS measurements at three different permanent GPS stations, and Radarsat-2 SAR times series have been built up since 2012. The three stations of IJmuiden (IJMU), Eijsden (EIJS) and Vlissingen (VLIS), shown in Fig. 5.5, were chosen because

1. each site is visible in at least one of the three available Radarsat-2 tracks covering the Netherlands (IJMU became visible in two tracks along the course of the experiment). The stations are spread out over three corners of the country.
2. all the three sites are also tide-gauge stations, for possible future comparison of ‘absolute’ InSAR deformation estimates with the measured changes in sea-level.
3. the three GPS stations had varying levels of radar backscatter prior to transponder installation. VLIS had high levels of ambient clutter around the GPS station (ships
and several metallic structures), while EIJS had very low clutter (grassy fields). IJMU had intermediate clutter levels.

4. the sites are easily accessible by road.

Three more transponders (CAT1, CAT2 and CAT3) were deployed at an open field in Wassenaar (WASS) for controlled experimentation (see Sec. 3.5.4).

![Figure 5.5: The permanent GPS stations of IJMU, EIJS and VLIS, from left to right respectively.](image)

In IJMU and VLIS, the transponders are directly attached to the mast of the GPS antenna via a rigid and sturdy structure, as shown in Fig. 5.2. The transponder in EIJS was mounted on the roof of the building housing the tide-gauge, because in this case, the GPS mast is not directly connected to the building, and it was logistically inconvenient to attach the transponder in correct orientation to this mast.

5.3.3. Datum connection in IJmuiden

A transponder was deployed in May 2012 at the permanent GPS station of IJMU. Fig. 5.2 shows the mechanical construction of the transponder-GPS setup.

**Calibrated amplitude behaviour in SAR images**

SAR images from Radarsat-2 track t202 were acquired for a few months in 2012. In November 2012, it was decided for logistical reasons to change the position of the transponder, rotating it to the opposite side of the GPS mast. From August 2013, SAR images from Radarsat-2 track t102 were also made available.

SAR amplitudes contain information on the power of the received radar signal from ground targets. In addition to the interaction of the transmitted radar signal with the target (i.e. the RCS), the received radar signal is also affected by factors such as the system loss or the antenna gain and effective aperture, which introduce a radiometric bias [143]. For correct comparison of targets imaged with different sensors or under different operating conditions, it is important to first perform radiometric calibration, to account for all the contributions in the radiometric values not due to the target characteristics [144]. To do this, SAR image pixels are converted to quantities that represent the RCS, such as the backscatter coefficient ($\sigma_0$, the RCS per unit area in the ground-range), radar brightness ($\beta_0$, the RCS per unit area in the slant-range or radar LoS direction) or $\gamma_0$ (the RCS per unit area of the incident wavefront, i.e. perpendicular to slant-range). Their relationship is illustrated in Fig. 5.7. It has been shown in [143–145] that for quantitative use of the information present in SAR amplitude imagery, $\beta_0$ or the *radar brightness* is
Figure 5.6: Location of Dutch permanent GPS sites and Radarsat-2 tracks. Images from track t102 (west) and t202 (centre) have been merged together to form one single long image each. The longer track t302 (east) consists of three images.
the most preferred parameter, because of its independence on sea-level geoid models typically used to approximate the local incidence angles.

Radarsat-2 SLC images are available without radiometric calibration, but are accompanied by scaling look-up tables (LUTs) that allow for $\sigma_0$, $\beta_0$ and $\gamma_0$ calibrations. To convert the complex digital value $N_{jk}$ at range $j$ and azimuth $k$ to its $\beta_0$ value in [dB], we use

$$\beta_0 = 10 \log_{10} \frac{|N_{jk}|^2}{G_j^2}$$

where $G_j$ is the range-dependent $\beta_0$ LUT gain value [143, 144]. Fig. 5.8 shows the temporal performance of $\beta_0$ at IJMU for both the Radarsat-2 tracks; observe the higher temporal stability after transponder installation.

**InSAR processing**

InSAR deformation analysis of the area around IJMU was performed using the time-series of SAR images in Radarsat-2 track t202. The Delft implementation of the PSI algorithm, DePSI [10, 12], was used for this analysis. The SAR acquisition dated 15 August 2013 was selected as the InSAR master. Atmospheric Phase Screen (APS) estimation was performed using the entire stack of images since 2010, but PS selection was done using only the images acquired after the transponder was rotated to its final position in Nov 2012. The resulting deformation estimates are shown in Fig. 5.9. The reference point in this case (marked with a white square) is the default chosen by DePSI, based on the temporal ensemble coherence estimator, which is a metric to describe the deviation between the deformation time series and the deformation model estimated [12]. The pair of PS points (arc) that has the highest ensemble coherence is selected, and of the pair, the PS closer to the centre of the image is chosen as reference point by default. The de-
Figure 5.8: $\beta_0$ time-series for IJMU (t102 and t202). CAT initialization signifies the time when the transponder was installed, or the first date when it was programmed for transponding. CAT rotation indicates the time when the transponder was aligned from an initial position to the final one, along the axis of the GPS antenna pole. The dates when the CAT malfunctioned were due either to power supply failure, or logistical inconvenience in the timely update of the transponder internal calendar. CAT testing dates are when tests were performed on the transponder, which might have influenced the signal. CAT malfunction and testing dates are not considered in the analysis.

formation at all other PS points are relative to this reference point, i.e. the deformation of the reference point is effectively assumed to be zero.

Consider the deformation maps that would have resulted if PS #2 or 3 had been in the arc with highest ensemble coherence, and were therefore chosen as reference points by DePSI. The resulting scenarios are depicted in Figs. 5.10 and 5.11. While the maps in Figs. 5.9, 5.10 and 5.11 contain identical information in the double-difference sense, it is impossible to gain insight into the physical deformation being experienced by the PS points without prior knowledge of reference point motion. In other words, any motion in a reference point that has been assumed stable reflects in the deformation time-series of all other PS points.

In order to determine the ‘absolute’ deformation of every PS point in a TRF, we first refer the PS deformation map to the transponder installed in IJMU, i.e. use PS #1 as the reference point. The GPS-derived deformation at this point, described in the following subsection, is then used as external information on the actual motion of this reference point.

**GPS data and processing**

The IJMU permanent GPS station belongs to AGRS, a GNSS network operated by the Dutch cadastre which is used to link RD-NAP and ETRS89 (see Appendix C). The GPS antenna is situated on the roof of a building that houses a tide-gauge. The observations of the GPS antenna are stored in hourly and daily data files at intervals of 10 and 30 seconds respectively. These data files, in Solution Independent Exchange (SINEX) format
Figure 5.9: PS deformation (in radar LoS) rate map of the IJmuiden area. The reference point selected by DePSI is marked with a white square. Three example PS are also marked, numbered 1, 2 and 3, along with their respective deformation time-series. PS #1 is the transponder installed at IJMU.

Figure 5.10: PS deformation (in radar LoS) rate map of the IJmuiden area referred to PS #2. The deformation time-series of the three example PS are also shown.
Figure 5.11: PS deformation (in radar LoS) rate map of the IJmuiden area referred to PS #3. The deformation time-series of the three example PS are also shown.

[146, 147], are then processed with Bernese 5.0 [148], the de facto standard software for monitoring permanent GNSS networks. SINEX enables standardized analysis and interpretation of GPS results by preserving information on the hardware (receiver, antenna), occupancy, a priori information and the correspondence between hardware, solution and input files. The SINEX files are processed in the high-precision static mode using the AGRSCLUS strategy, according to the guidelines of the EUREF commission [149, 150].

AGRSCLUS uses several reference frame sites operated by the IGS as shown in Fig. 5.12. BRUS was decommissioned in 2012, and MORP was introduced as a new reference site, causing a small offset in the IJMU time series [149, 150]. Additionally, the antenna of IJMU GPS receiver was replaced in 2014, causing another offset in the time series. A longer time-series of observations with the new antenna are required to reliably estimate the offset and for adequate testing [150]. Therefore, we utilize data only from the older GPS antenna in our analysis, i.e. a time-series spanning eight years (2006-2014).

The goal of GPS processing for our application is to precisely determine the vertical position of the antenna in a standard TRF, yielding a time-series of height changes at that position caused by the physical motion of the GPS antenna. But even a stationary antenna (not undergoing any physical motion) is subject to certain displacements which are not accounted for in a standard TRF. These displacements are very similar for antennas at short distances from each other, and are largely eliminated in a relative setup over short baselines. However, corrections for these displacements are necessary when the baselines involved are up to several hundreds of km (as in our case, see Fig. 5.12), and are applied in accordance with the conventions specified by the International Earth Rotation and Reference Systems Service (IERS) [151]. These include models describing
the displacements of the GNSS antenna due to various geophysical effects, such as [151]

1. tidal motions (usually near diurnal and semi-diurnal frequencies) and other accurately modelled displacements of reference points (usually at longer periods), e.g. solid earth tides, ocean tidal loading, atmospheric pressure loading, rotational deformation due to polar motion, and

2. other displacements of reference points such as due to environmental conditions. Models used here are e.g. reference values for local temperature, antenna phase centre offsets and variations, and so on.

Non-tidal motions associated with changing environmental loads (very broad spectral content) are not included in the corrections, as they normally change very little over typical integration spans and are less accurately modelled [151]. Tab. 5.1 lists some sources of GNSS site displacements and range effects, along with their approximate magnitudes in the vertical direction [150–152]. The resulting GNSS solutions, (after correction) reflect the physical motion of the station in a standard TRF, even up to the cm- to mm-level.

The corrections specified by the IERS should ideally also be applied to InSAR before datum connection, to be consistent with international conventions. However, given the differential nature of InSAR and the areal extent of a typical SAR image, the effect of these phenomena are expected to largely cancel out; some effects may also leak into the APS and hence be filtered out. In case of wide-area InSAR processing (over several hundreds of km), the effects of these conventional models would need to be evaluated prior to datum connection, and if found to be significant, the deformation estimates corrected for.

The GPS solutions of the IJMU permanent station referred to the NAP are shown in Fig. 5.13. These are weekly solutions obtained through processing in static mode; the rationale behind using weekly solutions is explained in Sec. 5.3.3. The associated error
Table 5.1: Approximate magnitudes of GNSS site displacements and range effects, corrected for according to the IERS conventions [150–152].

<table>
<thead>
<tr>
<th>Effect</th>
<th>Magnitude in vertical (radial) direction [cm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solid earth tides</td>
<td>30</td>
</tr>
<tr>
<td>Polar tides</td>
<td>2.5</td>
</tr>
<tr>
<td>Ocean loading</td>
<td>&lt;5</td>
</tr>
<tr>
<td>Atmospheric loading</td>
<td>&lt;3 (maximum), &lt;0.15 (typical)</td>
</tr>
<tr>
<td>Satellite antenna offsets</td>
<td>100</td>
</tr>
<tr>
<td>Antenna phase centre variation</td>
<td>4</td>
</tr>
<tr>
<td>Phase wind-up</td>
<td>10</td>
</tr>
<tr>
<td>Higher-order ionospheric delays and bending</td>
<td>5</td>
</tr>
<tr>
<td>Instrumental delays</td>
<td>6</td>
</tr>
<tr>
<td>Multipath effect on carrier phase</td>
<td>1</td>
</tr>
</tbody>
</table>

bars are propagated from the information contained in the SINEX files, scaled based on empirical variance information available for the station.

The reference position of IJMU in metres is known at epoch 2010.0 to be \((X, Y, Z) = (3882053.2800, 309346.2071, 5034330.2291)\) in ETRS89/ETRF2000, which is converted to the RD-NAP system using standard transformation routines [150, 153]; Fig. 5.13 therefore shows the GPS solutions (in Up-direction only) as NAP heights. A linear deformation trend is estimated for the time span between 2006-2014, i.e. for solutions obtained from the old antenna prior to antenna replacement in 2014. This GPS-derived linear trend is used as \(v(p_0)\) in Eq. (5.3) for ‘absolute’ InSAR displacements. Additionally, if ‘absolute’ InSAR heights are also desired, \(h(p_1, t_1)\) can be obtained by determining \(h(p_1, t_0)\) using either a precise DEM, or the InSAR height estimates which are relative to the reference transponder with known reference position. The following subsection describes datum connection in more detail.

**Datum connection**

As mentioned earlier, the objective of datum connection is to make the interpretation of InSAR deformation estimates independent of any postulation on reference point stability. To do this, an InSAR persistent scatterer is collocated with a GPS antenna and used as the InSAR reference point. The motion of this reference point is determined using the GPS solutions obtained from the antenna.

A rigorous approach is *epoch-wise datum connection*, i.e. to use the best estimate of the instantaneous position of the GPS antenna at the time of each SAR acquisition. Temporal smoothing of the GPS solutions may be allowed considering assumptions on the ‘real’ physical behaviour of the antenna position. The requirement for this approach, however, is that instantaneous GPS position estimates be available with at least the precision achievable with InSAR, i.e. mm-level precision. GPS solutions with the highest precision are achieved using *static* processing, where the computed position is based on
GPS observations collected over a length of time, e.g. an hour, a day or a week. This processing strategy eliminates many of the temporal error sources, e.g. due to the atmosphere (estimated), multipath effects (averaged out) and loading mismodelling (averaged out). The precision is therefore a function of the observation duration. Most authors in literature specify the achieved precision for different observation durations, often using different metrics, making intercomparison difficult; a summary is provided in [152]. To reach mm-level precision, five or more days of GPS observations are recommended, with a minimum requirement of two days [150]. Instantaneous positions can be computed through the *kinematic* processing strategy, using 10- or 30-second interval GPS observations, e.g. for fast-moving objects such as co-/post-seismic motions or buoys at sea. However, these positions have significantly larger error bars, and are more susceptible to systematic errors and carrier-phase multipath effects [150, 152]. As a ballpark figure, 30-second kinematic solutions at IJMU yield standard deviations of about 2 cm [150]. Therefore, for precise geodetic datum connection, static GPS processing is preferred.

A consequence of static processing is that the actual movement of the GPS antenna within the averaging period also smooths out, including any sudden anomalous behaviour that may occur around the time of the SAR acquisition. Epoch-wise datum connection therefore requires the implicit (strong) assumption that the instantaneous InSAR deformation estimate at SAR acquisition time is the same as the averaged GPS solution. Since weekly GPS solutions have error bars comparable to those achievable with InSAR, (as shown in Fig. 5.13), we choose to use them, assuming that it is unlikely for a sturdy permanent GPS station to exhibit anomalous deformation behaviour exceeding the precision limits of both measurement techniques, i.e. >5 mm, within the time span.
of a week.

To summarize, there is a trade-off between the precision of GPS observations/solutions and the assumption of deformation signal smoothness; the longer the averaging period, the higher the resulting GPS precision, but the stronger the assumption on deformation signal smoothness. In addition to imposing this assumption, epoch-wise datum connection is unable to incorporate GPS solutions from time epochs other than those covering SAR acquisition times, i.e. the weekly solutions in Fig. 5.13 corresponding to weeks without SAR acquisitions are discarded.

As an alternative approach, we introduce *model-based datum connection*, where a deformation model is estimated through all available weekly GPS solutions, even those not coinciding with SAR acquisition times. This model can then be attributed as the deformation of the InSAR reference point, if the following requirements are met:

1. The GPS time series should be long enough to reflect the true deformation trend. Referring to Fig. 5.13, the displacement rate reflected by the GPS solutions between December 2010 and December 2011, or between August 2009 and December 2010 are very different from the rate reflected by the entire time series. While it is clear that a longer time series will yield a better estimate of the deformation rate, care must be taken that the GPS time-series spans longer than the period of the expected deformation with the longest wavelength. Higher order models may also be estimated, if the physical cause for the non-linearity can be explained.

2. Similar assumptions on the model of expected deformation should be applied both during InSAR and GPS processing. In other words, if the model estimated through the GPS solutions is a linear deformation rate, the assumption that the InSAR reference point also deforms linearly (without any anomalous motion) should hold. This is ensured by performing InSAR *reference point noise removal*, correcting for anomalous motion affecting only the reference point and not the other PS, as described in the following paragraph.

Model-based datum connection is performed in the area of IJmuiden by first estimating a linear deformation model through the weekly GPS solutions, as shown in Fig. 5.13. Based on this model, the vertical position of the IJMU GPS station is found to subside at the rate of 1.9 mm/y. The PSI results of Fig. 5.9 are then transformed such that PS #1, i.e. the transponder collocated with the GPS antenna, is the InSAR reference point; all other deformation estimates are referred to this transponder. The resulting displacement rate map is shown in Fig. 5.14, with the reference point having a postulated zero-displacement time series. In reality, however, the reference point is also a scatterer, with thermal noise and a certain unknown coherence per epoch, which we term as the *reference point noise* (RPN). If not accounted for, the RPN manifests in every arc, i.e. the time series of every other PS contains the noise of the reference point in addition to its own noise. We isolate and remove the RPN as follows [154, 155]:

1. **Temporal deramping**: The linear displacement rate of each PS is subtracted from the corresponding displacement time series. The residues contain noise and any unmodelled deformation.
2. **RPN estimation**: The temporally-deramped PS are averaged per epoch, and then contain the components present in all the PS at that epoch. We attribute these components to the RPN, shown in Fig. 5.15(c) for all the SAR acquisition epochs.

3. **RPN removal and reramping**: For each PS, the RPN is subtracted from the deformation for all epochs, and the corresponding linear displacement rate is added back to each time series. This RPN-corrected time series is smoother than the original time series, parameterized by a smaller average gradient (see Fig. 5.14), computed by differencing the time series between epochs. Figs. 5.15(a) and (b) show that RPN correction reduces the component of temporal variability of the deformation estimates that appears in all the PS.

RPN removal therefore corrects for anomalous motion that affects only the reference point and not the other PS. The remaining motion at the InSAR reference point can then be modelled by the linear deformation derived from the collocated GPS antenna, as shown in the time series of the reference point, PS #1, in Fig. 5.16, with the deformation rates of all other PS affected accordingly.

Fig. 5.16 shows the ‘absolute’ vertical deformation rate map of the IJmuiden area, referred to the NAP. In the nomenclature of Eq. (5.3), the term $v(p_0)$ has been derived from Fig. 5.13. The deformation time-series of each PS in this map denotes the temporal evolution of the NAP heights of the corresponding radar scatterer.

Upon comparing the ‘absolute’ deformation in Fig. 5.16 with the default DePSI results of Fig. 5.9, we observe significantly differing PS deformation rates; in fact, the transpon-
Figure 5.15: Plots showing the spatio-temporal variation of deformation (a) before, and (b) after RPN mitigation. Note that RPN correction reduces noise in the temporal direction, which appears as vertical bands in (a). (c) shows the estimated RPN per epoch.

der (i.e. PS #1) shows opposite directions of deformation. We stress here that the information content in both cases is identical in the double-difference sense. However, Fig. 5.16 also contains information in the form of single-differences (in time only), because of external GNSS information available at the reference point. Note that the DePSI default criterion of having a low-noise reference point is not compromised by choosing a different reference point for datum connection, because of the RPN correction performed.

Fig. 5.17 shows the PS heights $h(p_1, t_1)$ estimated using DePSI and referred to the RD-NAP; the term $h(p_1, t_0)$ of Eq. (5.3) has been derived from the reference GPS position. These ‘absolute’ maps should be viewed with the consideration that the location of the phase centre of the transponder is known to be within the enclosing box, but not more precisely. We have accounted for the measured height difference of 50.6 cm between the GPS antenna and the base of the transponder during datum connection. While this does not affect the deformation rate maps, the estimates of ‘absolute’ heights in the NAP should consider this uncertainty, in addition to the inherent uncertainty in InSAR height estimation.

Note that in the IJmuiden example, we have chosen to project the LoS InSAR deformation estimates to the vertical direction (assuming no horizontal deformation) and use only the Up-component of the GPS solutions for datum connection, so that the resulting deformation map can be defined as changes in (NAP) heights. Such an approach is valid only if the magnitude of horizontal deformation can either be neglected, accurately
Figure 5.16: ‘Absolute’ vertical deformation rate map of the IJmuiden area, referred to the NAP. The deformation time-series of the three example PS are also shown, as changes in NAP heights.

Figure 5.17: The locations of the PS in the IJmuiden area, in the RD-NAP system (right). The PS heights are ‘absolute’, with 0 corresponding to the NAP (left, photograph taken from [156]).
Figure 5.18: (a) The monthly mean sea level in the period 2006-2013, and (b) the annual mean sea level in the period 1872-2013, as measured at the IJmuiden tide-gauge station, referred to the NAP [157, 158]. Such information, in conjunction with a coastal deformation map also referred to the NAP, can be used for flood-risk mapping.

modelled, or measured using satellite data of different looking directions [131]. If not, it is also possible to perform datum connection in the radar LoS direction, by using the 3D GPS solutions also projected to the radar LoS. Then the deformation results can be described as changes in \((X, Y, Z)\) coordinates in an earth-centred TRF, instead of (NAP) height changes.

The ‘absolute’ estimates in Figs. 5.16 and 5.17 can directly be compared with any other information that is in the RD-NAP system, such as historical levelling data or sea-level changes. Fig. 5.18 shows the changes in the sea level as measured by the tide-gauge at IJMU, both in the recent past and over the last century, referred to the NAP. The generation of a flood risk map based on the effect of sea-level changes on the deforming coastal lands is in itself a detailed topic of research, and is beyond the scope of this thesis.

5.3.4. Eijsden and Vlissingen

In Sec. 5.3.2, we introduced the three GNSS stations selected in the Netherlands for our experiment, and we have described the results of datum connection for the case of IJMU in Sec. 5.3.3. In case of EIJ, the transponder was not mechanically attached to the GNSS antenna, owing to logistical difficulties. Instead, the transponder was mounted on the roof of the tide-gauge building adjacent to the GNSS mast, shown in Fig. 5.5. Since it could not be ensured that the transponder and the GNSS antenna were experiencing identical deformation signals, it was decided to not use this experiment site for actual
5.3. InSAR-GNSS datum connection in the Netherlands

datum connection, but for a feasibility analysis only. The case of VLIS was also used as a feasibility test-site only, as the high ambient radar clutter led to no significant improvement in the SCR, and therefore to the phase stability. Moreover, the transponder at VLIS was initially plagued by power supply issues. The time series of radar brightness $\beta_0$ at EIJS and VLIS are depicted in Fig. 5.19. During the period that the transponder was operational, EIJS in particular (low ambient clutter) shows much smaller variability and much higher average values of $\beta_0$.

In summary, of the three GNSS sites initially selected for transponder deployment, IJMU and EIJS were found to be suitable locations for transponder collocation from the point of view of ambient clutter, whereas VLIS showed no significant improvement in radar brightness or amplitude stability with transponder deployment. Based on these experiences, we explore in the following section the feasibility of datum connection using collocated InSAR-GNSS measurements at a countrywide level in the Netherlands.

5.3.5. Feasibility study of transponder collocation at GNSS stations

It was discussed in Sec. 5.2.2 that it is necessary to collocate a transponder and a GNSS antenna at a location where there is no phase interference from other radar scatterers. The effective phase centre of the transponder pixel is then known to be within the transponder. In case ambient bright scatterers are present, the effective phase value of the pixel would be the superposition of the transponder phase and the phases from the other scatterers, including contributions from possible deformation of the other scat-
ters (local variations, building-ground reflections, etc.), which would not be desirable. The objective is therefore to determine candidate GNSS stations in the Netherlands that are not already PS, and do not contain strong scatterers in the neighbouring pixels.

52 permanent GPS stations were investigated in this study, with locations as shown in Fig. 5.6. The three Radarsat-2 tracks over the Netherlands (t102, t202 and t302) were processed for the period of August 2013 to June 2014 (14 SAR images). Over smaller areas, data from a longer period of time (June 2010 to June 2014, 61 SAR images) was processed; in Fig. 5.6, these include the upper half of the overlapping area between t102 and t202, as well as around the GNSS stations of VLIS and EIJS. The automated SAR processing entailed coregistration and resampling of each slave image to a common grid with that of the master image. Radar coordinates (line and pixel numbers) were computed from the coordinates of the GNSS stations at antenna height. Using the line and pixel numbers, crops of 300 \times 300 m around the GNSS stations were generated. The range coordinates were corrected for atmospheric delays, assuming the same Zenith Total Delay (ZTD) for every day and all stations. Geolocation quality has been verified at the locations of IJMU, EIJS and WASS with operating transponders; the quality is of the order of 1 pixel (i.e. a few metres).

To determine the suitability of a GNSS site for transponder installation, the following metrics computed from the resampled SAR images at all the GNSS station locations (as well as at WASS) are considered:

- **Radar brightness** ($\beta_0$), which denotes the radar reflectivity per unit area in slant range as explained in Sec. 5.3.3, and is calculated using Eq. (5.18). A site with low $\beta_0$ prior to transponder installation is preferred. The empirical threshold on $\beta_0$ is based on the lessons learnt from transponder deployment in IJMU, EIJS and VLIS; IJMU was the GNSS station with the maximum $\beta_0$ that still showed an improvement when a transponder was installed.

- **Signal-to-clutter ratio** (SCR) is defined as in Eq. (3.7), and is computed as shown in Fig. 3.9. The relation between phase standard deviation and SCR is given in Eq. (3.8). Adopting a threshold on the noise level in the phase observations of 0.5 rad, a threshold for the average SCR over all images of 2 dB is obtained; this can be used to detect if a pixel is a candidate PS [10, 12] before transponder installation. Disadvantages of this metric are that the clutter is likely to be overestimated in urban areas (because neighbouring pixels may also contain bright scatterers), and significantly more computational resources are required than for $\beta_0$.

The desirable attributes towards location suitability are summarized in Tab. 5.2. Visual inspection refers to the study of the radar backscatter of GNSS stations by eye, using mean reflectivity maps (average radar brightness) and animations of changes in amplitude characteristics at and around GNSS stations. Using results both from visual inspection and from thresholding the metrics, we produce a transponder collocation feasibility map of all permanent GNSS stations considered in the Netherlands, shown in Fig. 5.20. Tabs. 5.4 and 5.3 provide information on the decision made at each GNSS station. Note that these results are valid only for the transponders used in this study; more powerful transponders with higher RCS may still be usable at some of the stations that have been marked as unsuitable. Fig. 5.21 shows the $\beta_0$ characteristics around a few example
Table 5.2: Desirable attributes towards location suitability. $\beta_{0,\text{IJMU}}$ refers to average radar brightness of IJMU before transponder deployment.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Before transponder installation</th>
<th>After transponder installation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>Not bright, low clutter</td>
<td>Bright, low clutter, consistent in time</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>$\beta_0 \leq \beta_{0,\text{IJMU}}$</td>
<td>$\beta_0$ as high as possible</td>
</tr>
<tr>
<td>SCR</td>
<td>SCR $\leq$ 2 dB</td>
<td>SCR as high as possible</td>
</tr>
</tbody>
</table>

GNSS stations; the $\beta_0$ values are averages over the entire data stack (prior to transponder installation in case of IJMU, EIJS and VLIS).

The feasibility map indicates that it is not a trivial exercise to find suitable GNSS stations for transponder collocation. Many GNSS sites are unfavourable for transponder collocation, owing to the fact that permanent GNSS antennas are usually mounted on buildings which are bright radar scatterers, and in urban areas surrounded by strong levels of clutter. There are several regions of the world where permanent GNSS stations are not as densely distributed as in the Netherlands. In such regions, if the sparse GNSS stations are found to be unfavourable for transponder collocations, a portable collocated transponder-GNSS installation (I2GPS [36, 37]) would need to be deployed at locations with low background clutter for datum connection (see Figs. 5.3 and 3.6). The position and deformation rate of the GNSS antenna of the I2GPS unit can be determined in a TRF with respect to a reference permanent GNSS station, and used for subsequent datum connection with InSAR via the collocated transponder.

5.4. Summary and discussion

In this chapter, we have performed ZOD for InSAR by making use of permanent GNSS stations to convert the local datum of InSAR survey results to a standard geodetic one. An opportunistic approach for datum connection is to utilize the coherent InSAR scatterers that happen to occur in the vicinity of GNSS stations as reference points, assuming they show the same displacement behavior as that measured by GNSS. However, this (harsh) assumption does not always hold in practice, because of local variations in displacements. Also, coherent InSAR scatterers can often result from multiple-bounce reflections, and could contain a component from ambient effects (e.g. swelling/compaction in the surrounding soil) which would not be measured by a well-founded GNSS antenna. Furthermore, it is not guaranteed to have coherent InSAR scatterers in the vicinity of a GNSS station. In the proposed method, we therefore collocate phase-stable radar transponders with GNSS stations. This way, no assumptions are required regarding the relative motion between the transponder and the GNSS antenna, since their collocation via a sturdy mechanical connection ensures that they both experience the same movement. We thus obtain a network of radar beacons with independently observed position time series in a TRF.

As these reference GNSS stations are often located spatially dense enough to have one or more stations in a typical radar image (170 $\times$ 250 km for Sentinel-1), there is an...
Figure 5.20: Map of the Netherlands, depicting the suitability of GNSS stations for collocation with transponders of RCS $\sim 32$ dBm$^2$, using Radarsat-2 data in descending mode. Location suitability was determined using the amplitude characteristics of the area before transponder installation, as summarized in Tabs. 5.4 and 5.3.
Table 5.3: Decision of GPS station suitability for collocation with transponders of RCS ~ 32 dBm², using different Radarsat-2 tracks. Red (cross) = not suitable, green (tick) = suitable, yellow (?) = can be considered (August 2013 - June 2014, 14 images).

<table>
<thead>
<tr>
<th>Station</th>
<th>t102</th>
<th>t202</th>
<th>t302</th>
<th>Visual check</th>
<th>Overall location suitability</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHOL</td>
<td>✔</td>
<td>✔</td>
<td>❌</td>
<td>✔</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>MDBG</td>
<td>✔️</td>
<td>❌</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>RILL</td>
<td>✔️</td>
<td>❌</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>SASS</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>SDYK</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>WILLL</td>
<td>✔️</td>
<td>❌</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>ADR2</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>ADRN</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>APEL</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>?</td>
</tr>
<tr>
<td>ARNH</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>BRD2</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>BRDA</td>
<td>❌</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>CABW</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>EHN</td>
<td>✖️</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
<td>?</td>
</tr>
<tr>
<td>JOUR</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>KOS1</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>KOSG</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>LELY</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>LITH</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>LVRD</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>STAV</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>TERS</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>TSCM</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>TIEL</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>UTR2</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>UTRE</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>VILE</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
<td>✔️</td>
</tr>
<tr>
<td>ASSN</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
</tr>
<tr>
<td>DZYL</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
</tr>
<tr>
<td>EMMN</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✖️</td>
</tr>
<tr>
<td>HOOOG</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>SCHI</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>TAPL</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>WSRT</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>EBRG</td>
<td>✔️</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>OLDE</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>ZWO2</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>ZWOL</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>BOXM</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>KERK</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>MSTR</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>ROER</td>
<td>✖️</td>
<td>✖️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Table 5.4: Decision of GPS station suitability for collocation with transponders of RCS \( \sim 32 \text{ dBm}^2 \), using different Radarsat-2 tracks. Red (cross) = not suitable, green (tick) = suitable, yellow (?) = can be considered (June 2010 - June 2014, 61 images).

<table>
<thead>
<tr>
<th>Station</th>
<th>t102</th>
<th>t202</th>
<th>t302</th>
<th>Visual check</th>
<th>Overall location suitability</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK2</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>X</td>
<td>Only starting from March 2012 (t102).</td>
</tr>
<tr>
<td>ALKM</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>Only starting from March 2012 (t102).</td>
</tr>
<tr>
<td>AMST</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>Only starting from March 2012 (t102).</td>
</tr>
<tr>
<td>DELF</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>Some strong reflections from a building nearby that could interfere with a transponder.</td>
</tr>
<tr>
<td>DLF1</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>Some strong reflections from a building nearby that could interfere with a transponder; only good with t202.</td>
</tr>
<tr>
<td>IJMU</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>There is some overlap in t202 with another scatterer, t102 visually seems to be better. Confirmed good site, despite high SCR before transponder installation.</td>
</tr>
<tr>
<td>RDAM</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>ZOET</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>EUS</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>Confirmed good site.</td>
</tr>
<tr>
<td>VLIS</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
<td>Too many scatterers nearby.</td>
</tr>
</tbody>
</table>

Figure 5.21: \( \beta_0 \) characteristics in 300 × 300 m areas around some example GNSS stations, before transponder installation wherever applicable: (a) IJMU, (b) EJS, (c) VLIS and (d) DELF. Based on the ambient clutter, (a) and (b) are examples of good locations, whereas (c) and (d) are not suitable for transponder installation.
‘absolute’ InSAR reference point in each scene, connected to a TRF such as the International/European Terrestrial Reference Frame (ITRF/ETRF). The available VCMs of the GNSS stations can be used when transforming relative InSAR displacement estimates to displacements with respect to the reference ellipsoid. Subsequently, the (time series of) vertical positions can then be referred to an international or national height system using the local state-of-the-art geoid.

An application of the proposed approach is to tie overlapping or non-overlapping radar datasets together, yielding datasets comparable over large distances, even across oceans and continents. Additionally, region-wide deformation can be linked to sea-level changes via collocated InSAR-GNSS measurements, thereby contributing valuable information, e.g. towards flood risk assessment. The collocated transponder-GNSS approach can perform datum connection between InSAR survey results stemming from different satellites and looking directions, simply by programming the transponder accordingly. The regular, standardized and frequent acquisitions of the Sentinel-1 mission will also make the connection of all future radar observations to a TRF possible. A further benefit of this approach that has not been assessed in this chapter is the correction of residual orbit errors in different radar datasets. Also, phase ambiguities can be estimated in an absolute sense, as the connection via the GNSS station yields exact geometric ranges.

To make an operational ‘absolute’ countrywide deformation (rate) map, it is important that transponder-GNSS collocations are performed at multiple GNSS sites in the country. At least one collocation is required per SAR image, preferably more to allow for redundancy. The feasibility map of Fig. 5.20 is helpful to make choices on where best to collocate transponders. In case a SAR image does not contain a permanent GNSS station favourable to transponder collocation, a portable collocated transponder-GNSS device can be deployed instead.

The primary causes of ground deformation in the Netherlands are soil compaction, tectonics and isostasy, with neo-tectonics being the least-known component [161]. Studies of long-wavelength tectonics and isostasy are based on geological studies over very long time periods, as well as prediction models. The time scales over which we measure deformation with InSAR, however, are much shorter; we measure quicker changes over time, meaning that such models cannot be constrained very well with InSAR time-series of a few (tens of) years. Soil compaction, on the other hand, can cause significant short-term surface deformation in the Netherlands.

The ‘absolute’ deformation map in our example case study (see Fig. 5.16) shows that during the considered time period (2012-2014), the (sandy) coastal regions in the IJmuiden area have been mostly stable, while some regions further inland show subsidence rates of 5–10 mm/y with sporadic points subsiding up to 15 mm/y. These localized rates are in tune with those suggested by prognosis models of the expected displacement rates due to soil compaction, which have been built based on a geostatistical interpretation of data from shallow boreholes [159, 160]. Fig. 5.22 shows such a prognosis model; the IJmuiden area is marked with a blue square. To make this model, a 3D model of the subsurface is constructed by Kriging the lithology (sand, clay or peat) determined by each borehole over the entire area, accounting also for models on the spatial distribution of each lithology. The expected settlement and compaction is then calculated, after introducing a number of generalizations. This prognosis is grossly approximate in
Figure 5.22: Predicted surface displacement rates caused by the Holocene layer for the period 2000-2050, based on geological information, assuming the same weather conditions and water table level [159, 160]. The Ijmuiden area of interest is marked with a blue square; the expected rates in this area are similar to those measured using InSAR.

a geodetic sense, both in terms of displacement rates and their spatial distribution and resolution, but provides us with a general idea of the levels of subsidence one could expect from soil compaction. An operational ‘absolute’ InSAR deformation rate map of the Netherlands, incorporating GPS measurements from across the country as well as connected to historical levelling data via the NAP, can contribute valuable insights towards constraining such deformation prognosis models.
Conclusions

The best way to predict the future is to create it.
Abraham Lincoln

In this final chapter, we summarize the detailed conclusions and recommendations from previous chapters, in order to compose the big picture of geodetic network design as applied to InSAR. The general problem statement was given in Ch. 1 as

How can an optimal geodetic network be designed for InSAR, towards the goals of measurement densification in decorrelating areas, and integration of the deformation estimates into a standard terrestrial reference frame?

This study points out that the spatial distribution of InSAR measurements can be improved by utilizing the parameter space instead of the observation space for designing a network of additional measurements. A methodology has been proposed where the design of this network is optimized towards precision, reliability and economy. For any arbitrary network configuration, the quality of InSAR measurements (both naturally occurring and deliberately deployed) is propagated to the quality of the estimated parameters that are of interest to the user. If the existing InSAR measurements are found to be too sparse for the monitoring application at hand, the most efficient device deployment configuration with a parameter estimation quality that satisfies user-requirements is selected. Examples of such devices are corner reflectors and transponders. Also as part of this research, low-cost transponders developed over the recent years were validated for geodetic applications; they were found to measure deformation with precision comparable to corner reflectors. Additionally, owing to their favourable physical characteristics, transponders play a major role in the methodology proposed in this study towards connecting InSAR deformation estimates to a standard terrestrial reference frame via collocated GNSS measurements.

The general problem statement has been cast in the classical framework of geodetic network design, which involves four design orders. Not all of these orders can directly
be applied to InSAR, because the technique utilizes existing radar scatterers whose locations and quality are not under our control (Ch. 2):

- **FOD** (first-order design, finding the optimum measurement locations) is not directly applicable to InSAR, since the locations of naturally-occurring coherent targets cannot be designed. In the extreme case where there are no coherent targets in the area of interest, FOD can be used to determine the optimal spatial sampling required. In the proposed methodology, concepts of FOD are used in THOD to translate user-requirements into a criterion matrix (Ch. 4 and [38]).

- **SOD** (second-order design, deciding which observations to make and with what precision) is also not entirely applicable to InSAR. Unlike geodetic techniques such as levelling where observations may be repeated until the desired precision is achieved, the precision of InSAR measurements cannot be altered. But another important difference between InSAR and other geodetic techniques is that it is not straightforward to determine the precision of InSAR measurements *a priori*, owing to the dependence of precision on the physical properties of the radar scatterer and surrounding clutter. SOD for InSAR therefore involves the quality description of InSAR measurements. The precision of deployed devices is also key to optimizing their numbers; the more precise the measurement, the less the number of devices required. Corner reflectors have been studied extensively in the past; here, we have estimated the achievable precision of active radar transponders via several field experiments with data from different satellites (Ch. 3 and [29]).

- **THOD** (third-order design, improving an existing network) is the design order that is mainly applicable to InSAR, for improving the spatial sampling. If the existing InSAR measurements are found to be too sparse, information on their (probable) locations and quality, the quality of the devices to be deployed, user-defined monitoring requirements and any available knowledge of the expected deformation is translated to the optimal configuration (number and locations) of devices (Ch. 4 and [38]).

- **ZOD** (zero-order design, defining a reference frame or datum) is applied to InSAR in scenarios where knowing the relative displacements of radar scatterers with respect to another (reference) scatterer alone is not adequate. By using a transponder (collocated with a GNSS measurement) as the InSAR reference point, it is possible to determine the physical deformation experienced by the reference scatterer in a standard TRF. The interpretation of InSAR deformation estimates is then independent of any postulation on reference scatterer stability; the estimates can be compared with other datasets also defined in the same TRF, e.g. changes in sea level (Ch. 5).

Geodetic network design for InSAR therefore consists of three main steps: SOD, combined FOD-THOD and ZOD in chronological order (Ch. 2). A concise summary of the proposed methodology has been depicted in Fig. 1.1. We present below the main contributions of this research in relation to the three specific research questions listed in Ch. 1, followed by recommendations for further studies.
6.1. Contributions

6.1.1. Transponder viability

Specific research question 1:

Passive devices such as corner reflectors have been used in the past for InSAR measurement network densification. However, they suffer from drawbacks related to their large size and weight, conspicuousness, and loss of reliability due to geometric variations as well as material and maintenance-related degradation over several years of deployment. As an alternative, low-cost active radar transponders have been developed recently, which are smaller, lighter, and less conspicuous. Are such transponders a viable alternative to corner reflectors in terms of the measurement precision achieved?

1. The low-cost active radar transponders tested in this study were found to be viable alternatives to corner reflectors in terms of phase stability and the geodetic measurement precision obtained. This was deduced through field experiments conducted in the Netherlands and in Slovenia, and by comparing transponder-derived InSAR deformation estimates using ERS-2, Envisat and Radarsat-2 to those from levelling and GPS (Ch. 3).

2. The experiments using Radarsat-2 data and precise levelling have indicated a double-difference precision range of 0.8 – 1.0 mm for transponder measurements in radar LoS. Transponder measurements resulting from ERS-2 and Envisat data, both in suboptimal conditions (zero-gyro mode and drifting baselines, respectively), when compared with levelling and GPS respectively, have yielded double-difference precision ranges of the order of 1.8 – 4.6 mm in radar LoS (Ch. 3).

6.1.2. Optimal densification of InSAR measurement networks

Specific research question 2:

Prior to device deployment, are the existing InSAR measurements adequately distributed for the monitoring application at hand? If not, what is the minimum number of additional devices (corner reflectors or transponders) required to be deployed? From a geodetic point of view, what are the optimal ground locations of these devices?

1. InSAR studies so far have located a predetermined number of devices based on practical deployment considerations, or ‘intuitively’. Devices are usually deployed in regions where coherent targets are absent, with inter-measurement spacing based on implicit assumptions of atmospheric noise properties, and the expected deformation extent. In this work, we have formalized this ‘intuition’ via a geodetic network design framework. First, the need for additional InSAR measurements is assessed, based on monitoring requirements defined by the user. Then, if needed, both the optimal number and locations of additional measurements are determined (Ch. 4).
2. The theoretical geodetic framework for designing an optimal network, used in the past e.g. for levelling and GNSS networks, has been tailored to the case of InSAR network densification, considering precision, reliability and economy as the three optimality criteria. The final precision requirements of the parameters of interest are specified by the user, which are translated into a criterion matrix. In case a functional model of the expected deformation is specified, the reliability criterion is used to define a search space of possible device locations, in order to limit the search space to only those locations where gross errors or outliers can be effectively detected and resisted. The quality of parameter estimates obtained from various different device configurations within the search space are compared to the criterion matrix using a distance metric; lower values of this metric indicate configurations that are closer to user-requirements. This design method is therefore oriented towards the practical needs of the end-user of InSAR-derived deformation estimates. The primary objective is not to find the best possible model to describe the deformation and its mechanisms, but to achieve closeness to user-defined requirements on deformation monitoring quality. The difference is that the former would entail minimization of the estimation error, which would yield the trivial result that deploying more devices would always reduce the estimation error further. In contrast, the proposed approach gives the minimum number of additional measurements required to satisfy user-needs without over-designing the network, thereby accounting for the criterion of economy (Ch. 4).

6.1.3. InSAR-GNSS datum connection

Specific research question 3:

How can InSAR displacement estimates be connected to a standard terrestrial reference frame?

1. InSAR displacement estimates have been connected to a standard TRF via GNSS. A geodetic approach has been proposed towards InSAR-GNSS datum connection, where the deformation rate of the InSAR reference point is estimated by connecting it to (collocated) GNSS measurements. All other InSAR deformation estimates which are relative to this reference point are thus defined in the same TRF as the GNSS measurements (Ch. 5).

2. Collocation of the InSAR reference point and the GNSS measurement has been ensured by mechanically attaching transponders to permanent GNSS stations. The sturdy physical connection eliminates the need for any assumptions on the relative motion between the phase centre of the InSAR reference point (transponder) and the GNSS measurement. This setup has been implemented at three permanent GPS stations in the Netherlands: IJmuiden, Eijsden and Vlissingen (Ch. 5).

3. The study of InSAR-GNSS datum connection at IJmuiden has shown that the instantaneous GPS observations (even averaging over several minutes or hours) yields position solutions that have precisions of a few cm in the vertical direction, which is an order of magnitude worse than the transponder deformation measurement
precision. To improve precision, the GPS observations are averaged over several days to a week, assuming the GPS station to move predictably within the averaging period. As a consequence, any unexpected variation in the deformation signal smooths out; a datum-connection inaccuracy caused by a SAR acquisition coinciding with such a variation should be accounted for, in the overall error budget. There is thus a trade-off between the precision of GPS observations/solutions and the assumption of deformation signal smoothness (e.g. a linear model); the longer the averaging period, the higher the resulting GPS precision, but the stronger the assumption on deformation signal smoothness. We have chosen to use weekly GPS solutions; the rationale is that it is unlikely that a sturdy permanent GPS station would exhibit anomalous deformation behaviour exceeding the precision limits of both measurement techniques, i.e. >5 mm, within the time span of a week. Moreover, we reduce the noise contribution of the InSAR reference transponder in the deformation time series of all other InSAR measurement points, thereby correcting for possible anomalous motion experienced by the reference transponder that makes its behaviour deviate from the assumed smooth model (Ch. 5).

4. Experiences with transponder-GPS collocation in IJmuiden, Eijsden and Vlissingen, as well as analysis of radar amplitude and clutter characteristics at all permanent GNSS sites in the Netherlands have shown that it is not a trivial exercise to find suitable GNSS stations for transponder collocation. Because most permanent GNSS receivers are mounted on buildings and are situated in urban areas, about 70 – 80% of the stations were found to be unsuitable for transponder collocation, owing to strong radar backscatter prior to transponder installation and high levels of ambient clutter.

6.2. Recommendations

1. Ground deformation can be caused by several factors, such as structural instabilities and shallow/deep subsurface motion. If it is desired to monitor the ground deformation arising from deep subsurface motion such as that caused by fluid injection or extraction, there may be significant contamination e.g. from shallow subsurface motion due to groundwater level variations. When deploying corner reflectors or transponders, care must be taken to mount them firmly to prevent any autonomous motion, using foundations of adequate depth as to capture the deformation signal of interest.

2. Transponders may be designed such that more of their electronic capabilities are utilized. For instance, a delay line may be implemented in order to virtually locate a transponder at a different part of the image that has low clutter, to improve its apparent SCR. Additionally, messages may be coded to the signal returning to the satellite, with information about ambient conditions and transponder health.

3. Determining the distribution of existing InSAR measurements prior to device deployment can be performed using an archived stack of SAR data, a single SAR amplitude image, external information (e.g. optical data, vegetation indices, degree
6. Conclusions

of urbanization), or a combination of these. The inherent variability in the predicted locations of existing InSAR measurements (see Appendix B for examples on the effect of different assumptions on PS density) can affect the designed optimal number and locations of additional devices. When in doubt, the worst-case scenario of the distribution of existing InSAR measurements should be used. Likewise, conservative assumptions should be made on the expected deformation, so that network densification provides a fail-safe measurement distribution, especially in high-risk applications.

4. The methodology proposed in Ch. 4 can be adapted to further optimize the search space of potential device locations. Clearly, the larger the search space, the longer the computational time. When a functional model is available as a priori knowledge, reliability has been used as a criterion for streamlining the search space. If a finer location resolution is desired, or if no prior functional model is available to constrain the search space, a two-tier approach can be implemented. The algorithm can first be run using a coarse search space, and in a second iteration, the algorithm repeated using a finer search space within the area that has been determined to be optimal in the first iteration. Alternatively, stochastic search or evolutionary genetic algorithms may also be investigated.

5. In Ch. 4, we have only considered ground deformation in the vertical direction, neglecting the horizontal component. In case SAR data from different looking directions of a satellite are available, combined FOD-THOD can be performed such that optimal locations are determined based on the decomposed vertical and horizontal deformation as separate parameters to be estimated.

6. The combined FOD-THOD scheme can be extended in the temporal domain as well. The examples in Ch. 4 have assumed the ground deformation to behave linearly in time, and compared only VCMs denoting spatial behavior. By comparing 3D spatio-temporal VCMs, the locations of added devices can also be optimized accounting for the temporal behaviour of ground deformation.

7. Ch. 4 contains simple simulated and real-world examples as proofs of concept for the proposed algorithm, with the intention of comparing the designed optimal network of devices with intuitive placement. The algorithm should further be demonstrated on more complex deformation patterns (e.g. a functional model that includes geological faults) and mechanisms (i.e. with more intricate models of subsurface processes that could complicate the relationship between surface deformation and the final parameters of interest). The methodology can also be extended to include subsurface parameters instead of surface deformation as a priori knowledge, using data assimilation techniques.

8. Methodologies derived from the covariance comparison framework of Ch. 4 have applications that are not limited to geodetic network design; they can be used for any network design/densification problem where the quality of estimates can be propagated from the quality of observations. The framework can also be extended for use in intermediate steps of InSAR processing and post-processing. For example, large PS/DS data can be downsampled more effectively than regular grid or
quadtree approaches, by retaining higher measurement density at locations found to be more crucial with respect to the expected deformation. The PS candidate selection step of a PSI processing chain such as DePSI [10, 12] can also be improved likewise; PS candidates that are in crucial locations with respect to the anticipated deformation can be forcefully retained, so that there is some information available in crucial areas, even if of lower quality.

9. More precise instantaneous GNSS position estimates would enable epoch-wise InSAR-GNSS datum connection, further minimizing the assumptions currently required, e.g. that the motion of the GNSS receiver and the collocated transponder follows a particular model.

10. Connecting InSAR to absolute gravimetric measurements should be investigated. Absolute gravimetry could serve as an alternative to GNSS for vertical datum realization, since it does not suffer from the uncertainties in a TRF, caused e.g. by the motion of the earth’s centre of mass or slight changes in the orientation of the TRF. Linear trends in vertical motion can be determined with an accuracy of 1–2 mm/y using GNSS; absolute gravimetry can provide higher accuracies depending on the location [162]. However, careful analysis is required to convert the gravity changes measured by gravimetry into vertical motion. In addition, other signals that may cause changes in gravity need to be modelled and corrected for. Absolute gravimetry can also complement GNSS. Gravimetry would help in providing a physical understanding of the nature of the observed motion, and can be used to separate the effects of mass redistribution (physics) from those of changes in elevation (geometry); this is particularly important in the context of comparison with sea-level changes [163].

11. A possible future scenario in the era of Sentinel-1 would be that there is at least one case of transponder-GNSS collocation within every 170 × 250 km Sentinel-1 SAR image. Over such large areas, it would be necessary to perform the same corrections (specified by the IERS) as applied to GNSS prior to datum connection; further research is required to determine if some of these effects leak into the APS and are filtered out, to prevent over-correction. After datum connection, all the InSAR deformation estimates arising from these SAR images would be defined in a standard TRF. These estimates could be stitched together to produce country- or continent-level deformation maps, which could then directly be compared to other datasets also referred to the same TRF, such as historical levelling data and changes in the sea level. Such a comparison would provide valuable information towards flood risk assessment, especially in low-lying coastal countries such as the Netherlands.
References


[47] M. Kavouras, *On the detection of outliers and the determination of reliability in geodetic networks* (Department of Surveying Engineering, University of New Brunswick, Canada, 1982).


References


Estimation, Prediction and Hypothesis Testing

Measurements are stochastic quantities that are affected by random and/or systematic errors. A systematic methodology to estimate unknown parameters from uncertain measurements and propagate the associated errors has been developed at the TU Delft over the last few decades. This appendix presents a brief summary of the aspects of the methodology that are relevant to this thesis: estimation, prediction and hypothesis testing in linear(ized) models.

A.1. Best linear unbiased estimation (BLUE)

BLUE is an estimation technique that has the optimal properties of unbiasedness and minimum variance. The BLUE estimator is equal to the Maximum Likelihood (ML) estimator for normally-distributed data [43, 106], and has the same structure as the weighted least-squares estimator [43].

For the linear system in 2.2, the BLUE parameter estimator $\hat{x}$ and its VCM $Q_x$ are given by [43, 106]

$$\hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} y$$ and $$Q_x = (A^T Q_y^{-1} A)^{-1}$$

respectively. Eq. (A.2) is used in Secs. 4.2.3 and 4.4.

A non-linear system such as in Eq. (2.1) can be linearized before applying BLUE. Preliminary values of the unknown parameters ($x_0$) are derived from the APK of deformation, and parameter increments added to these $x_0$ are estimated. The first-order Taylor series approximations for each function form the design matrix $A$, which is the Jacobian.
matrix of the observation equations [41, 43]

\[ A = \partial_x^T F(x_0) = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(x_0) & \cdots & \frac{\partial}{\partial x_q} f_1(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_p(x_0) & \cdots & \frac{\partial}{\partial x_q} f_p(x_0) \end{bmatrix}. \] (A.3)

The linearized approximation of Eq. (2.1) is then

\[ E\{\Delta y\} = A \Delta x; \quad D\{\Delta y\} = Q_y \] (A.4)

where \( \Delta y = y - F(x_0) \) and \( \Delta x = x - x_0 \). This can be solved iteratively, e.g. using the Gauss-Newton method. The BLUE parameter estimator \( \hat{x} \) becomes

\[ \hat{x} = x_0 + (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} \Delta y \] (A.5)

with \( Q_x \) as in Eq. (A.2).

**A.2. Best linear unbiased prediction (BLUP)**

Prediction implies that a function of an observable random vector \( y \) is used to guess the outcome of another random but unobservable vector \( y_0 \). It is also known in different fields of application as least-squares collocation, Kriging interpolation or objective analysis [164]. BLUP is a prediction technique; BLUE becomes BLUP when the parameter vector is random instead of deterministic [43]. For the following general linear(ized) system of equations

\[ \begin{bmatrix} y \\ y_0 \end{bmatrix} = \begin{bmatrix} A \\ A_0 \end{bmatrix} x + \begin{bmatrix} e \\ e_0 \end{bmatrix} \] (A.6)

where the design matrices \( A \) and \( A_0 \) are of orders \( m \times n \) and \( m_0 \times n \) respectively, \( x \) is the non-random parameter vector, and \( [e^T, e_0^T]^T \) is a random vector with expectation and dispersion given by [164]

\[ E\left\{ \begin{bmatrix} e \\ e_0 \end{bmatrix} \right\} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad D\left\{ \begin{bmatrix} e \\ e_0 \end{bmatrix} \right\} = D\left\{ \begin{bmatrix} y \\ y_0 \end{bmatrix} \right\} = \begin{bmatrix} Q_y & Q_{yy_0} \\ Q_{y_0y} & Q_{y_0} \end{bmatrix}, \] (A.7)

the BLUP estimator-predictor pair \( \hat{x} \) and \( \hat{y}_0 \) are given by [164]:

\[ \hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} y, \] (A.8)

\[ \hat{y}_0 = A_0 \hat{x} + Q_{y_0y} Q_y^{-1} (y - A \hat{x}). \] (A.9)

The vector \( \hat{y}_0 \), in our context, contains the predicted displacements at the points of interest. The prediction quality is determined by the VCM of the prediction error \( \hat{e}_0 = y_0 - \hat{y}_0 \), rather than that of the predictor \( \hat{y}_0 \) itself [164]. This error VCM is given by

\[ Q_{\hat{e}_0} = Q_{y_0} - Q_{y_0y} Q_y^{-1} Q_{yy_0} + (A_0 - Q_{y_0y} Q_y^{-1} A) Q_{\hat{x}} (A_0 - Q_{y_0y} Q_y^{-1} A)^T \] (A.10)
with $Q_\hat{\beta}$ from (A.2). Eqs. (A.9) and (A.10) are used in Sec. 4.3.

In the special case of unknown functional model, (2.1) reduces to only the stochastic part. Therefore, without the design matrices $A$ and $A_0$, $\hat{y}_0$ and $Q_\hat{e}_0$ from (A.9) and (A.10) become

$$\hat{y}_0 = Q_{y0}yQ_y^{-1}y,$$

$$Q_\hat{e}_0 = Q_{y0} - Q_{y0}Q_y^{-1}Q_yy_0.$$ (A.11) (A.12)

Eq. (A.12) is used in Sec. 4.2.4. It is observed from (A.2) and (A.10) that the observations $y$ are not required for either the precision description of the BLUE estimates $Q_\hat{\beta}$ or of the BLUP prediction error $Q_\hat{e}_0$.

The quality description of stochastic variables consists of precision and reliability, which are independent components; a high measurement precision does not guarantee reliable estimation of the unknown parameters, and vice versa. Precision and reliability together describe the accuracy [12, 43]. The precision of the measurements and the estimators of the unknown parameters are denoted by their respective VCMs $Q_y$ and $Q_\hat{\beta}$.

The reliability describes the sensitivity of the estimators to model errors, and is ensured by hypothesis testing.

A.3. Hypothesis testing in linear(ized) models

After estimation and/or prediction is performed, the validity of the mathematical model is tested for errors. This is done by the comparison of hypotheses. A statistical hypothesis is an assertion or conjecture about the probability distribution of one or more random variables, for which a random sample is available (e.g. through measurements) [43]. The null hypothesis $H_0$ is the default or nominal state, assuming no errors in the model. An alternative hypothesis $H_a$ assumes a certain error in the model, i.e. a departure from the nominal state. These hypotheses can be symbolically denoted as:

$$H_0 : \quad E[y] = Ax; \quad D[y] = Q_y$$

$$H_a : \quad E[y] = Ax + C_q\nabla; \quad D[y] = Q_y$$ (A.13) (A.14)

where $\nabla$ contains $q$ additional unknown parameters, and $C_q$ describes the functional relationship between these parameters and $y$. The stochastic model $Q_y$ is assumed to be correct, and the same under both the hypotheses. Acceptance or rejection of $H_0$ is based on the test statistic $T_q$, derived from the Generalized Likelihood Ratio Test as [43, 165]

$$T_q = \hat{\beta}_0^TQ_y^{-1}\hat{\beta}_0 - \hat{\beta}_a^TQ_y^{-1}\hat{\beta}_a$$ (A.15)

with the subscripts 0 and $a$ corresponding to the null and alternative hypotheses respectively. The test statistic can be used to decide whether $H_a$ is a significant extension of $H_0$. $T_q$ has a $\chi^2$-distribution with $q$ degrees of freedom, i.e.

$$H_0 : \quad T_q \sim \chi^2(q, 0)$$

$$H_a : \quad T_q \sim \chi^2(q, \lambda)$$ (A.16) (A.17)
A. Estimation, Prediction and Hypothesis Testing

Figure A.1: The concept of hypothesis testing, where \( H_0 \) is compared to \( H_a \). The test statistic \( T_q \) in relation to the critical value \( k_\alpha \) determines the acceptance or rejection of \( H_0 \). \( k_\alpha \) is based on the chosen level of significance \( \alpha \), which is the probability of type-I errors (i.e. incorrect rejection of \( H_0 \)). The probability of type-II errors (i.e. incorrect acceptance of \( H_0 \)) is denoted by \( \beta \). The power of the test (i.e. the probability that \( H_0 \) is correctly rejected) \( \gamma = 1 - \beta \). Figure taken from [12].

where \( \lambda \) is the non-centrality parameter given by

\[
\lambda = \nabla^T C_q^T R^T Q_y^{-1} R C_q \nabla
\]  

(A.18)

with the redundancy matrix \( R \) as defined in Eq. (2.6). The critical value \( k_\alpha \) determines whether \( H_0 \) should be rejected, i.e.

\[
\text{reject } H_0 \text{ if } T_q > k_\alpha
\]  

(A.19)

as illustrated in Fig. A.1. It can be observed that to lower the probability of type-I and type-II errors, both \( \alpha \) and \( \beta \) should be minimized. But a smaller value of \( \alpha \) implies a larger value of \( \beta \), and vice versa. According to the Neyman-Pearson principle [166], which states that from all tests with the same probability of type-I errors, the one for which the chance of type-II errors is as small as possible should be used, it follows that the power of the test \( \gamma = 1 - \beta \) should be as large as possible while fixing \( \alpha \). Typical values for \( \alpha \) are chosen in the range of 0.001–0.05. By setting \( \gamma \) to its maximum value \( \gamma_0 \), the lower bound of the non-centrality parameter can be computed from the relation [10]

\[
\lambda_0 = \lambda(\alpha, q, \gamma = \gamma_0).
\]  

(A.20)

The link between \( \lambda_0 \) and reliability has been shown in Eq. (2.5); the size of model error that can just be detected with a fixed probability is given by the minimum detectable bias.

In hypothesis testing, \( H_0 \) is first tested to detect whether there are any errors in the model, using the Overall Model Test (OMT) [12, 43, 165]. OMT is a general test of the discrepancies between observed data and assumed model; there are no assumptions on the specific kind of errors. The number of errors is taken to be equal to the redundancy, i.e. \( q = m - n \), and therefore the redundancy of the model under \( H_a \) is zero. The test statistic of Eq. (A.15) then reduces to

\[
T_{q=m-n} = \hat{\epsilon}_0^T Q_y^{-1} \hat{\epsilon}_0
\]  

(A.21)
from which Eq. (3.13) is derived.

If $H_0$ is rejected, the source of the most significant error is identified by testing $H_a$. Most often, a test on outliers in the observations is performed using the w-test, which is a one-dimensional test (i.e. $q = 1$). The w-test is performed by checking each observation for a large error; the observation with the highest $T_q$ (above $k_a$) is the most significant error. For diagonal $Q_y$, the w-test statistic for the $i$th observation simplifies to [12, 43, 165]

\[
T_{q=1} = \frac{\hat{\sigma}_0^2}{\sigma_i^2}. \tag{A.22}
\]

The most significant error is removed from the functional model, and the hypothesis testing procedure is repeated until the OMT detects no more errors.
Towards Repeatability, Reliability and Robustness in InSAR

A number of implicit assumptions are required in order to perform successful InSAR data processing and parameter estimation. It is important to know which information to use, how arbitrary they are, which assumptions to make, and how important or sensitive these assumptions are. These assumptions are not only limited to the well-known phase unwrapping conditions and paradigms, but also on a priori assumptions on atmospheric behaviour, (micro)topographic variability, direction of the deformation vector, and spatiotemporal behaviour of the deformation signal.

Using the test case of Harlingen in the Netherlands, we demonstrate how changes in these assumptions affect the estimated parameters. Owing to gas extraction and salt mining, areas around the town of Franeker in the region around Harlingen have experienced significant subsidence as measured by periodic levelling campaigns. ERS data over this region is processed using different PSI settings, which differ in their a priori choice of deformation model, network design, persistent scatterer selection method, atmosphere filter length and shape, noise filter length and detrending. The influence of a posteriori filtering is also studied. The recommendation from this case-study is that the only solution to improve the reliability of InSAR results is to make these assumptions explicit and falsifiable.

B.1. Introduction

As more and more excellent radar data are becoming available, InSAR techniques are being routinely applied on these data to measure ground deformation. The coloured deformation maps produced by (time-series) InSAR often seem to be uniquely

This appendix has been published in the proceedings of the 8th International Workshop on Advances in the Science and Applications of SAR Interferometry (FRINGE), Frascati, Italy, 2011 [167].

125
B. Towards Repeatability, Reliability and Robustness in InSAR

Interpretable, indubitable, and a result of objective data processing. Unfortunately, this proposition does not always withstand critical assessment.

Several comparative studies have been performed in the past for InSAR verification and validation, such as PSIC4 [168], the Terrafirma validation experiment [169–171], and many other standalone experiments. However, a common feature of these studies has been that success in a particular case study does not necessarily imply success in another, i.e., generic feasibility cannot be guaranteed. This could be because of different radar data, locations, deformation and meteorological phenomena, processing methods, a priori information, and so on. Different methodologies, therefore, appear to produce significantly different results, which is unacceptable from a scientific point of view. Additionally, the current quality description of time-series InSAR deformation estimates is inadequate in terms of reliability, idealization precision or a full VCM. There is thus a lack of transparency for end-users.

There are two main considerations time-series InSAR techniques to note. Firstly, there is a clear need for a match between the spatiotemporal sampling and extent of data in relation to the spatiotemporal variability of the deformation phenomenon. The availability of coherent scatterers is obviously not guaranteed in all cases. Secondly, and the main subject of this appendix, there are a number of implicit assumptions one needs to make in order to perform successful interferometric data processing and parameter estimation, many of which are location-specific. The InSAR results can depend on the data (number of images, their quality, temporal separation, homogeneity and extent), scatterer properties (number, quality, spatial separation, homogeneity and extent), and signal (temporal and spatial smoothness, extent, abrupt changes in space and time). We will discuss further about a priori information or assumptions in the following section.

B.2. Making assumptions

InSAR combines pairs of radar observations, and the interferometric phase observation per resolution cell (pixel) contains contributions from different factors, as given in Eq. (3.2). This parameter estimation problem is inherently underdetermined and ill-posed. This means that it can only be solved by introducing a priori information. Some examples of a priori information introduced at various processing stages are depicted in Fig. B.1. However, there are some important questions to answer while introducing a priori information. What information do we add? How arbitrary is it? Which assumptions do we make? How important/sensitive are these assumptions?

B.3. The Harlingen case study

Gas and salt extraction in the region of Harlingen in northwestern Friesland, the Netherlands (Fig. B.2), have resulted in a cumulative subsidence of up to 32 cm (1988 to 2006), confirmed by periodic levelling campaigns [172] and PSI performed on this area.

First, the area was processed using the default set of parameters shown in Fig. B.3. The results that were obtained with this setting are shown in Fig. B.4. A subsidence rate of up to 2 cm/y was detected. To study the influence of assumptions on PSI, 15 different
B.4. Results

The number of PS points accepted in a particular setting was plotted against processing setting, as shown in Fig. B.5. It is observed that up to 30% of PS can go undetected with a change in assumption.
Deformation parameterisation:  
Detrending:  
PS selection:  
Grid size 1st order points:  
Network design:  
Temporal filter:  
APS spatial filter:  

<table>
<thead>
<tr>
<th>Default</th>
<th>Linear</th>
<th>Quadratic</th>
<th>Linear-periodic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold</td>
<td>Each grid cell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150 m</td>
<td>200 m</td>
<td>300 m</td>
<td>500 m</td>
</tr>
<tr>
<td>Spider</td>
<td>Delonay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>1 year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exponential</td>
<td>Gaussian</td>
<td>Spherical</td>
<td></td>
</tr>
</tbody>
</table>

Figure B.3: Tested parameters. The parameters in green are the default parameters, and in each setting, one parameter was changed with respect to the default set.

Fig. B.6 shows a PS acceptance map after *a posteriori* filtering, which indicates the number of settings in which a particular PS point was detected. Though a large number of points are yellow (detected by all settings), there are still a significant number of red points (detected by only one setting).

Some examples of differences in deformation estimates obtained using different settings is depicted in Fig. B.8, for deformation parameterisation model and network design parameters. It may be observed at first glance that in most of the study area, the differences are limited. However, there are a few points with differences that can be significant in relation to the reliability question. The maximum difference that we obtained between any two settings amounted to about 5 cm of cumulative deformation, as shown in Fig. B.7.

**B.5. Conclusions**

Our study of the Harlingen area indicates that *a priori* assumptions on expected behaviour do matter. The maximum ‘error’ on the estimated parameter, cumulative deformation, is several centimetres. Depending on location, this could have significant effect on, for instance, interpolation.

InSAR remains an ill-posed, underdetermined problem, mainly due to phase ambiguity, but also due to overlapping spectral characteristics of the estimated geophysical signals (e.g. APS, deformation). It is a very challenging task to obtain estimates with a full VCM. There are currently many assumptions that remain implicit or unexpressed, which make the results not falsifiable.

*A priori* information can play an important role in PSI parameter estimation, depending on the data, deformation signal or method used. In other words, a change in assumptions may lead to a significant change in results (density of PS points, signal magnitude). In order to make PSI results falsifiable, we need to make our assumptions explicit. The influence of these assumptions on the results must also be part of the final product.
### Table B.1: Fifteen different PSI processing settings

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
<th>Default setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Default setting</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Detrending</td>
<td>No detrending</td>
</tr>
<tr>
<td>3</td>
<td>Defo. parameterisation = quadratic</td>
<td>Linear</td>
</tr>
<tr>
<td>4</td>
<td>Defo. parameterisation = linear + periodic</td>
<td>Linear</td>
</tr>
<tr>
<td>5</td>
<td>Grid size (first order points) = 150 m</td>
<td>200 m</td>
</tr>
<tr>
<td>6</td>
<td>Grid size (first order points) = 300 m</td>
<td>200 m</td>
</tr>
<tr>
<td>7</td>
<td>Grid size (first order points) = 500 m</td>
<td>200 m</td>
</tr>
<tr>
<td>8</td>
<td>PS selection method = each grid</td>
<td>Threshold</td>
</tr>
<tr>
<td>9</td>
<td>Noise filter length = 1 year</td>
<td>6 months</td>
</tr>
<tr>
<td>10</td>
<td>Noise filter length = 3 months</td>
<td>6 months</td>
</tr>
<tr>
<td>11</td>
<td>Network design = Delaunay</td>
<td>Spider</td>
</tr>
<tr>
<td>12</td>
<td>APS spatial filter = 3 months</td>
<td>1 year</td>
</tr>
<tr>
<td>13</td>
<td>APS spatial filter = 6 months</td>
<td>1 year</td>
</tr>
<tr>
<td>14</td>
<td>APS spatial filter shape = exponential</td>
<td>Gaussian</td>
</tr>
<tr>
<td>15</td>
<td>APS spatial filter shape = spherical</td>
<td>Gaussian</td>
</tr>
</tbody>
</table>

Figure B.4: Deformation map of Harlingen obtained using the default set of parameters, after *a posteriori* filtering.
Table B.2: \textit{A posteriori} filter thresholds.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence</td>
<td>0.2</td>
</tr>
<tr>
<td>Height residual</td>
<td>30 m</td>
</tr>
<tr>
<td>Spatiotemporal consistency</td>
<td>0.008</td>
</tr>
<tr>
<td>Deformation rate</td>
<td>0.025 m/y</td>
</tr>
</tbody>
</table>

Figure B.5: Distribution of PS with setting. Top: without \textit{a posteriori} filtering. Bottom: with \textit{a posteriori} filtering.
Figure B.6: PS acceptance map, after *a posteriori* filtering. A PS point of a particular colour indicates the number of settings in which that point was detected.

Figure B.7: Maximum difference between two settings (cumulative deformation). Note that in the encircled region of interest (salt mining), the differences amount up to 5 cm.
Figure B.8: Some examples of differences between PSI assumptions/settings. The colourbar corresponds to the difference plots. Although the differences appear limited, they can be significant in relation to the reliability question.
Coordinate Systems, Datums, Reference Systems and Frames

A consistent description of the position and displacements of points on the earth’s surface requires a mathematical framework, or a coordinate system. A coordinate system uses one or more numbers, coordinates, to uniquely determine the position of a point in Euclidean space [173, 174]. Coordinate systems can be assigned to a reference system, which models geodetic observations as a function of unknown parameters of interest. A reference system therefore defines constants, conventions, models, and parameters, which serve as the necessary basis for the mathematical representation of geometric and physical quantities [175]. Different coordinate systems assigned to the same reference system can be transformed by a purely mathematical operation.

Coordinate systems can be related to a geodetic datum, which specifies how they are linked to the earth [31, 70]. A datum consists of the basic (minimum) parameters that define and describe the origin, orientation and scale of the coordinate axes. It thus involves a reference surface against which position measurements are made, with an associated model of the shape of the earth for computing positions. Modern geodetic datums range from flat-earth models used for plane surveying to complex models used for international applications which completely describe the size, shape, orientation, gravity field, and angular velocity of the Earth. Datum transformations are transformations between coordinates of two different reference systems, as illustrated schematically in Fig. C.1. Unlike purely mathematical coordinate conversions, datum transformation parameters can be stochastic, in which case the errors need to be propagated.

Several different datums have been used historically [173, 176]. In the late 19th- and early 20th centuries, many countries developed their own independent national coordinate systems, based on astronomical observations. The ED50 (European Datum 1950) was aimed at linking the various European datums and create a European reference system. The satellite era heralded the development of world-wide reference systems (e.g. the World Geodetic Systems WGS60 and WGS72), with the most recent version (WGS84) based on GNSS in 1987. These reference systems were well-aligned to the centre of mass.
Figure C.1: The difference between coordinate conversions (horizontal operations) and datum transformations (vertical operation between reference systems A and B). \((x, y)\) are the coordinates of a point in a certain coordinate system, \(H\) and \(h\) the associated ellipsoidal and orthometric heights respectively (see Fig. C.2), \((\phi, \lambda)\) the geodetic latitude and longitude of the point, and \((X, Y, Z)\) the Cartesian coordinates in a certain reference system. Figure taken from [173].

Figure C.2: The relationship between orthometric height \(H\), ellipsoidal height \(h\) and geoid height \(N\).

and rotation axis of the earth at the time of the reference epoch. Later the International Terrestrial Reference System (ITRS, a geocentric global reference system), and the European Terrestrial Reference System (ETRS, the standard reference system for Europe) were established.

The ITRS is a global spatial reference system co-rotating with the earth in its diurnal motion in space. Positions of points at the earth’s surface have coordinates which experience only small changes in time, owing to geophysical effects such as tectonic or tidal deformations. Reference frames are (different) realizations of the same reference system, both physically (by a solid materialization of points) and mathematically (by the determination of parameters e.g., geometric coordinates), that help fix the datum [175]. ITRS is realized through a number of International Terrestrial Reference Frames (ITRF) [151, 173]. Seven parameters are required to fix an ITRF at a given epoch; their time derivatives are added to define the ITRF time evolution. The selection of these fourteen parameters, called datum definition, establishes the ITRF origin, scale, orientation and their time evolution. ITRF can be used to describe plate tectonics, regional subsidence or deformation in a global context, or to represent the earth when measuring its rotation in space. Each ITRF contains positions and velocities of points
(stations) located on earth in every continent and tectonic plate. These are computed from space-geodetic techniques such as Very Long Baseline Interferometry (VLBI), Lunar Laser Ranging (LLR), Satellite Laser Ranging (SLR), Global Navigation Satellite Systems (GNSS) or Doppler Orbitography and Radiopositioning Integrated by Satellite system (DORIS) [151, 173]. Different datums are specified through time because their realizations, or estimates of the datum, change over time because of the addition of new stations and improvements in survey methods. There have been 12 realizations of the ITRS so far: ITRF88, ITRF89, ITRF90, ITRF91, ITRF92, ITRF93, ITRF94, ITRF96, ITRF97, ITRF2000, ITRF2005 and ITRF2008, with ITRF2013 currently under preparation [177]. The numbers in the ITRF designation specify the last year of data that were used. ITRF is the most accurate terrestrial reference frame until now, and is directly linked to WGS84 which is used by GPS.

ITRS plays an important role in studying global geodynamics. However, it is less suited for use as a European georeferencing system because the entire continent exhibits similar velocities of a few centimetres per year. Therefore, the ETRS89 has been defined as the standard European reference system, better suited for mm-level land surveying and precision mapping applications, and less dependent on the date on which a set of coordinates was acquired [173, 178]. ETRS89 is coincident with ITRS at the epoch 1989.0. It is fixed to the stable part of the Eurasian Plate, making station velocities very small. ETRS89 is accessed through the EUREF Permanent GNSS Network [178] or through one of many national and commercial GNSS networks in Europe. For each realization of the ITRS (i.e. ITRFxx), the corresponding frame in ETRS89 can be computed, called as ETRFxx. The ETRS/ETRF (referred to the Eurasian plate) are directly related to the ITRS/ITRF.

Figure C.3: The NAP monument (left) and an NAP benchmark (right) in Amsterdam. Photographs taken from [156, 179].

In the Netherlands, the Dutch Triangulation System (RD, Rijksdriehoeksstelsel in Dutch) is a commonly used reference system, in combination with the Amsterdam Ordnance Datum (NAP, Normaal Amsterdams Peil in Dutch) which defines the height datum (Fig. C.3). The NAP is also the datum for the European Vertical Reference System (EVRS), a height reference system defined as the zero-tidal equipotential surface for which the Earth gravity field potential is constant [180, 181]. From 2000 onwards, the RD-NAP system is directly linked to ETRS89 through a standard transformation procedure [153, 173].
Nomenclature

Symbols and operators

\((X_0, Y_0)\) Coordinates of the centre of the deforming area

\(\alpha\) Bearing angle of the semi-major axis with respect to +Y-axis (clear from context)

\(\alpha\) Level of significance of a test, probability of type-I errors (clear from context)

\(\beta\) Probability of type-II errors

\(\beta_0\) Radar brightness, the RCS per unit area in the slant-range

\(\delta\) Bowl shape parameter

\(\eta\) Subpixel position in range

\(\gamma\) Power of the test

\(\gamma_0\) Maximum value of the power of the test (clear from context)

\(\gamma_0\) RCS per unit area of the incident wavefront (clear from context)

\(\hat{\epsilon}\) Vector of residuals \((y - \hat{y})\)

\(\hat{x}\) Estimator of \(x\)

\(\hat{y}\) Estimator of \(y\)

\(\hat{y}_0\) Predicted (interpolated) values of \(y\) at points of interest

\(\hat{x}\) Estimate of \(x\)

\(\lambda\) Non-centrality parameter (clear from context)

\(\lambda\) Radar wavelength (clear from context)

\(\lambda_0\) Lower bound of the non-centrality parameter

\(\lambda_i(Q_1, Q_2)\) \(i\)th eigenvalue of \(Q_1 Q_2^{-1}\)

\(\mathbb{R}^m\) Real-valued observation space

\(\mathbb{R}^n\) Real-valued parameter space

\(\mu\) Mean of the Gaussian curve
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nabla \hat{x}$</td>
<td>Maximum effect of undetected gross errors on the estimator $\hat{x}$</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Poisson's ratio</td>
</tr>
<tr>
<td>$\phi_{\text{defo}}$</td>
<td>Phase component from deformation</td>
</tr>
<tr>
<td>$\phi_{\text{flat}}$</td>
<td>Phase component from a reference surface (ellipsoid)</td>
</tr>
<tr>
<td>$\phi_{\text{noise}}$</td>
<td>Phase component due to noise</td>
</tr>
<tr>
<td>$\phi_{\text{orb}}$</td>
<td>Phase component from errors in orbit determination</td>
</tr>
<tr>
<td>$\phi_{\text{range}}$</td>
<td>Range-dependent phase</td>
</tr>
<tr>
<td>$\phi_{\text{scat}}$</td>
<td>Phase component from changes in scattering within a resolution cell</td>
</tr>
<tr>
<td>$\phi_{\text{topo}}$</td>
<td>Phase component from uncompensated topography</td>
</tr>
<tr>
<td>$\phi_{ij}^{01}, \phi_{ij}^{1}$</td>
<td>Double-difference interferometric phase between points $i$ and $j$ and time epochs $t_0$ and $t_1$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>North-East-Up GPS solutions converted to radar LoS</td>
</tr>
<tr>
<td>$\rho_a$</td>
<td>Azimuth resolution</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>Slant-range resolution</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>$\sigma_0$</td>
<td>Backscatter coefficient, the RCS per unit area in the ground-range</td>
</tr>
<tr>
<td>$\sigma_\phi^2$</td>
<td>Phase variance</td>
</tr>
<tr>
<td>$\sigma_{\text{req}}$</td>
<td>Required standard deviation</td>
</tr>
<tr>
<td>$\theta_{\text{inc}}$</td>
<td>Incidence angle</td>
</tr>
<tr>
<td>$\Delta V$</td>
<td>Source volume change</td>
</tr>
<tr>
<td>$\Delta_p \Delta t h$</td>
<td>Double-difference (in space and in time) height measured by InSAR</td>
</tr>
<tr>
<td>$T_q$</td>
<td>Test statistic</td>
</tr>
<tr>
<td>$y$</td>
<td>Vector of observations</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Subpixel position in azimuth</td>
</tr>
<tr>
<td>$A$</td>
<td>Design matrix</td>
</tr>
<tr>
<td>$a$</td>
<td>Length of the semi-major axis of an ellipse</td>
</tr>
<tr>
<td>$a_{ij}$</td>
<td>Integer phase ambiguity between points $i$ and $j$</td>
</tr>
<tr>
<td>$b$</td>
<td>Length of the semi-minor axis of an ellipse</td>
</tr>
<tr>
<td>$B^\perp$</td>
<td>Perpendicular baseline</td>
</tr>
</tbody>
</table>
Nomenclature

\( c \)  
Clutter amplitude within a pixel

\( d \)  
Source depth

\( d(.) \)  
Distance metric operator

\( D[.] \)  
Dispersion operator

\( d_{i,j}^{01}, d_{i,j}^1 \)  
Double-difference deformation in radar LoS between points \( i \) and \( j \) and time epochs \( t_0 \) and \( t_1 \)

\( D_i \)  
Datum matrix for case \( i \)

\( E[.]. \)  
Expectation operator

\( F(.) \)  
(Non-linear) vector function from observation space to parameter space

\( f_{\theta \text{orbit}} \)  
(Residual) orbital trend

\( f_{DC} \)  
Doppler centroid frequency

\( G \)  
\( \beta_0 \) LUT gain value

\( H \)  
Inner constraint matrix

\( h \)  
Height

\( H_0 \)  
Null hypothesis

\( H_a \)  
Alternative hypothesis

\( H_{i,j} \)  
(Residual) topographic height between points \( i \) and \( j \)

\( I \)  
Identity matrix

\( k_\alpha \)  
Critical value determining the acceptance or rejection of \( H_0 \)

\( L \)  
Isosceles side length of a trihedral corner reflector

\( m \)  
Number of observations

\( n \)  
Number of estimated parameters

\( n_{i,j} \)  
Measurement noise and processing-induced errors between points \( i \) and \( j \)

\( N_{j,k} \)  
Complex digital value of a SAR image at range \( j \) and azimuth \( k \)

\( p_0(x_0, y_0) \)  
Location of the reference point

\( p_1(x_1, y_1) \)  
Location of a measurement point

\( P^\perp_A \)  
Orthogonal projection of \( y \) onto \( \hat{e} \) in the metric of \( Q_y \)

\( Q_{\hat{e}_0, \text{ref}} \)  
Criterion matrix of prediction quality

\( Q_{\hat{e}_0} \)  
VCM of prediction error
\( Q_{\text{ref}} \)  Criterion matrix of estimation quality
\( Q_\hat{x} \)  VCM of parameter estimates
\( Q_n \)  VCM of noise
\( Q_s \)  VCM of unmodelled deformation
\( Q_w \)  VCM of residual atmospheric error
\( Q_y \)  VCM of observations
\( R \)  Redundancy matrix
\( R_h \)  Horizontal distance between the point under consideration and \((X_0, Y_0)\)
\( r_i \)  Redundancy numbers
\( R_i \)  Range at point \( i \)
\( S \)  S-transformation matrix
\( s \)  Signal (pixel amplitude)
\( t_0 \)  Master (reference) time epoch
\( t_1 \)  Slave time epoch
\( T_{\text{OMT}} \)  Overall model test statistic
\( v \)  Velocity
\( x \)  Vector of parameters to be estimated
\( Z_{\text{max}} \)  Deformation corresponding to the peak of a Gaussian curve
\( (\cdot)_{ii} \)  \((i, i)\)th element
\( \Delta h_{ij} \)  Height difference between points \( i \) and \( j \)
\( \{\cdot\}^T \)  Matrix transpose operator

**Abbreviations and acronyms**

3D  Three-dimensional
AGRS  Active GPS Reference System
APK  *A Priori* Knowledge
APS  Atmospheric Phase Screen
BLUE  Best Linear Unbiased Estimation
BLUP  Best Linear Unbiased Prediction
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>Compact Active Transponder</td>
</tr>
<tr>
<td>DEM</td>
<td>Digital Elevation Model</td>
</tr>
<tr>
<td>DePSI</td>
<td>Delft PSI implementation</td>
</tr>
<tr>
<td>Doris</td>
<td>Delft Object-oriented Radar Interferometric Software</td>
</tr>
<tr>
<td>DS</td>
<td>Distributed Scatterer(s)</td>
</tr>
<tr>
<td>EIJS</td>
<td>Eijsden permanent GPS station</td>
</tr>
<tr>
<td>EPN</td>
<td>EUREF Permanent Network</td>
</tr>
<tr>
<td>ERS</td>
<td>European Remote Sensing satellite</td>
</tr>
<tr>
<td>ETRF</td>
<td>European Terrestrial Reference Frame</td>
</tr>
<tr>
<td>ETRS</td>
<td>European Terrestrial Reference System</td>
</tr>
<tr>
<td>EUREF</td>
<td>International Association of Geodesy (IAG) Reference Frame Subcommission for Europe</td>
</tr>
<tr>
<td>EVRS</td>
<td>European Vertical Reference System</td>
</tr>
<tr>
<td>FOD</td>
<td>First-Order Design</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System(s)</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>I2GPS</td>
<td>Integrated Interferometry and GNSS for Precision Survey</td>
</tr>
<tr>
<td>IERS</td>
<td>International Earth Rotation and Reference Systems Service</td>
</tr>
<tr>
<td>IGS</td>
<td>International GNSS Service</td>
</tr>
<tr>
<td>IJMU</td>
<td>IJmuiden permanent GPS station</td>
</tr>
<tr>
<td>InSAR</td>
<td>Interferometric Synthetic Aperture Radar</td>
</tr>
<tr>
<td>ITRF</td>
<td>International Terrestrial Reference Frame</td>
</tr>
<tr>
<td>ITRS</td>
<td>International Terrestrial Reference System</td>
</tr>
<tr>
<td>LoS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>LUT</td>
<td>Look-Up Table</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>NAP</td>
<td><em>Normaal Amsterdams Peil</em>, the Amsterdam Ordnance Datum</td>
</tr>
<tr>
<td>NETPOS</td>
<td>Netherlands Positioning Service</td>
</tr>
<tr>
<td>OMT</td>
<td>Overall Model Test</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>PS</td>
<td>Persistent Scatterer(s)</td>
</tr>
<tr>
<td>PSI</td>
<td>Persistent Scatterer Interferometry</td>
</tr>
<tr>
<td>PSIC4</td>
<td>PSI Codes Cross-Comparison and Certification</td>
</tr>
<tr>
<td>Radar</td>
<td>Radio detection and ranging</td>
</tr>
<tr>
<td>RCS</td>
<td>Radar Cross-Section</td>
</tr>
<tr>
<td>RD</td>
<td><em>Rijksdriehoeksstelsel</em>, the Dutch Triangulation System</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root-Mean-Squared Error</td>
</tr>
<tr>
<td>RPN</td>
<td>Reference Point Noise</td>
</tr>
<tr>
<td>SAR</td>
<td>Synthetic Aperture Radar</td>
</tr>
<tr>
<td>SBAS</td>
<td>Small Baseline Subset</td>
</tr>
<tr>
<td>SCR</td>
<td>Signal-to-Clutter Ratio</td>
</tr>
<tr>
<td>SINEX</td>
<td>Solution Independent Exchange</td>
</tr>
<tr>
<td>SLC</td>
<td>Single-Look Complex</td>
</tr>
<tr>
<td>SOD</td>
<td>Second-Order Design</td>
</tr>
<tr>
<td>SRTM</td>
<td>Shuttle Radar Topography Mission</td>
</tr>
<tr>
<td>StaMPS</td>
<td>Stanford Method for Persistent Scatterers</td>
</tr>
<tr>
<td>THOD</td>
<td>Third-Order Design</td>
</tr>
<tr>
<td>TRF</td>
<td>Terrestrial Reference Frame</td>
</tr>
<tr>
<td>VCE</td>
<td>Variance Component Estimation</td>
</tr>
<tr>
<td>VCM</td>
<td>Variance-Covariance Matrix</td>
</tr>
<tr>
<td>VLIS</td>
<td>Vlissingen permanent GPS station</td>
</tr>
<tr>
<td>WASS</td>
<td>Wassenaar test GPS station</td>
</tr>
<tr>
<td>WGS</td>
<td>World Geodetic System</td>
</tr>
<tr>
<td>ZOD</td>
<td>Zero-Order Design</td>
</tr>
<tr>
<td>ZTD</td>
<td>Zenith Total Delay</td>
</tr>
</tbody>
</table>
About the Author

Pooja Mahapatra was born on 18 January 1986 in Bangalore, India. She started her Bachelor of Engineering (BE) in Electronics and Communication Engineering in 2003 at MS Ramaiah Institute of Technology in Bangalore, graduating in 2007; her Bachelor thesis was based on a research project on audio signal processing at the Indian Institute of Science, Bangalore. During this period, she won a competition and was sponsored by the Indian Space Research Organisation (ISRO) to attend an international space conference in Valencia, Spain. This motivated her to apply to a European Master’s programme in space science and technology (SpaceMaster). She was awarded a scholarship by the European Commission to pursue this course.

She moved to Europe in the summer of 2007, and did coursework in Würzburg (Germany) and Kiruna (Sweden). In Germany, she also had the chance to work on a practical project on small-satellite design at the Fraunhofer Ernst-Mach-Institute (EMI) in Freiburg. Her Master thesis work (‘A Prototype System for Autonomous Rover-Based Planetary Geology’) was performed at the European Space Agency (ESA/ESTEC) in Noordwijk (the Netherlands). She graduated in 2009 with Master of Science (MSc) degrees from Luleå University of Technology (Sweden) and University Paul Sabatier Toulouse III (France).

Her growing interest in space applications, the geosciences and Dutch life led her to start her PhD on InSAR and geodesy at the Delft University of Technology in 2010. She worked on various projects as part of her PhD, most notably the Dutch national programme CATO2 (CO₂ Capture, Transport and Storage in the Netherlands) and the EU FP7 project I2GPS (Integrated Interferometry and GNSS for Precision Survey). This thesis describes her PhD work. Since 2014, she has been a Senior InSAR Engineer at SkyGeo, a company in Delft specializing in InSAR R&D and services.
List of Publications

Peer-reviewed journal articles

1. P. Mahapatra, S. Samiei-Esfahany, R. Hanssen, 


Conference proceedings

1. P. Mahapatra, H. van der Marel, F. van Leijen, S. Samiei-Esfahany, R. Klees, R. Hanssen, *Connecting InSAR to a global geodetic datum: Towards absolute scatterer displacements*, accepted for the 9th International Workshop on Advances in the Science and Applications of SAR Interferometry and Sentinel-1 InSAR Workshop (FRINGE 2015), Frascati, Italy.


A priori knowledge, 4, 46, 128

Backscatter coefficient, 75

Coherence, 20
Coherent target, 22
Coordinate system, 133
Corner reflector, 25, 35, 44
Criterion matrix, 6, 11, 12, 15, 46, 49, 50, 52, 54

Datum, 66, 133
Datum connection, 65, 82
Datum matrix, 68
Datum transformation, 68, 133
Decorrelation, 21, 23
Distance metric, 11, 15, 46, 49, 52
Distributed scatterers, 3, 20, 24
Double difference, 20, 22, 65–67

Economy, 13
Estimation, 34, 50, 119
ETRF, 135
ETRS, 134
EVRS, 135
External reliability, 12

First-order design, 5, 6, 9, 13, 14, 45, 98

Gauss-Markov model, 10
Gaussian subsidence bowl, 60
Geocoding, 16
Geodesy, 2
GPS processing, 32, 80

Hypothesis testing, 121

I2GPS, 8, 16, 29, 72
Inner constraint matrix, 68
InSAR, 2, 20, 65, 66
InSAR-GNSS collocation, 71, 75, 89
Interferogram, 3, 20
Internal reliability, 12
ITRF, 134
ITRS, 134

Local datum, 66

Mogi model, 50

NAP, 85, 87, 135
Network construction, 15, 49
Network design, 9, 13
Network reduction, 15

Optimality criteria, 10
Overall model test, 34, 122

Persistent Scatterer Interferometry, 3, 22, 127
Persistent scatterers, 3, 22
Phase stability, 27, 44
Point scatterers, 20
Precision, 10, 24, 32, 33, 35, 37, 41, 50, 54, 83, 121
Prediction, 54, 55, 120

Radar brightness, 41, 75, 89, 90
Radar cross-section, 26, 34, 75
Redundancy numbers, 12, 51, 54, 57, 62
Reference frame, 7, 16, 65, 80, 134
Reference point, 65–67, 77
Reference point noise, 84
Reference system, 133
Reliability, 12, 50, 54, 121, 122
Roswinkel subsidence, 60

S-transformation, 16, 68
Search space, 51, 57, 62
Second-order design, 5, 6, 9, 13, 14, 19, 98
Sequential optimization, 15, 49
Signal-to-clutter ratio, 24, 90
SINEX, 80
Single-look complex image, 20
Small baseline subset, 3
SqueeSAR, 3
Subsidence prognosis, 95

Third-order design, 5, 6, 9, 13, 14, 45, 98
Transponder, 25, 41, 71, 75
Type-I and II errors, 122

Variance-covariance matrix, 10, 68, 119

W-test, 34, 123

Zero-order design, 5, 6, 9, 13, 16, 66, 98
Propositions
accompanying the dissertation
Geodetic Network Design for InSAR
by
Pooja Mahapatra

1. Expressing the precision of double-difference InSAR displacement estimates at the sub-millimeter level is meaningless. (Ch.3)

2. The estimation of deformation from InSAR and the design of an optimal geodetic network for InSAR are both impossible without any assumptions on the spatio-temporal behaviour of the ground deformation. (Ch. 4)

3. In practice, deformation estimates need to be defined in a standard Terrestrial Reference Frame only if the effect of deformation is to be related with other physical phenomena such as sea level changes. (Ch. 5)

4. Corner reflectors and transponders should also be used as reassurance to the public, who are often incredulous about millimetre-level deformation being monitored from several hundred kilometres away in space, even though this is effectively a ‘placebo’ effect. (Chs. 3 and 4)

5. In developing countries where a significant fraction of the population faces more severe problems than those caused by millimetre-level ground deformation, InSAR will remain a technique that solves ‘luxury problems’ until it can be used operationally to predict disasters and help in planning preemptive evacuation.

6. Scientific communication meant for the public should state facts, objective interpretations and remedies; not opinions.

7. The difference between an engineer and a scientist lies in how they tackle problems: a scientist finds the cause, an engineer finds a workaround.

8. The large pool of easily-accessible information on the internet stifles the creation of original ideas and encourages ‘quick-fix’ application of existing knowledge.

9. The best measure of success is happiness. (from Richard Branson)

10. The faculty premises at the TU Delft should remain open at all times, including overnight. Providing 24/7 access to workspaces increases the research throughput of any institution.

11. Regular long-distance train journeys contribute positively to scientific research.

These propositions are regarded as opposable and defendable, and have been approved as such by the supervisor prof. dr. ir. R. F. Hanssen.
**Stellingen**

behoorende bij het proefschrift

**Geodetic Network Design for InSAR**

door

**Pooja Mahapatra**

1. Het op sub-millimeter niveau beschrijven van de precisie van ‘double-difference’ InSAR schattingen is zinloos. (H. 3)

2. Zowel het schatten van deformaties als het ontwerpen van een optimaal geodetisch netwerk voor InSAR is onmogelijk zonder aannames over het spatio-temporele gedrag van de bodemdeformatie. (H. 4)

3. In de praktijk hoeven deformatieschattingen alleen gedefinieerd te worden in een standaard Terrestrisch Referentie Frame wanneer deze gerelateerd dienen te worden aan andere fysische fenomenen, zoals zeespiegelveranderingen. (H. 5)

4. Hoekreflectoren en transponders zouden ook gebruikt moeten worden in relatie tot de publieke acceptatie van InSAR metingen, aangezien het publiek vaak weinig fiducie heeft in millimeters gemeten vanuit de ruimte, ook al is dit feitelijk een ‘placebo-effect’. (H.3 en 4)

5. In ontwikkelingslanden, waar een significant deel van de bevolking ernstigere problemen ondervindt dan die veroorzaakt door bodembeweging op millimeter-niveau, zal InSAR een techniek blijven die ‘luxeproblemen’ oplost, totdat deze operationeel gebruikt kan worden om rampen te voorspellen en kan helpen in de planning van preventieve evacuaties.


7. Het verschil tussen een ingenieur en een wetenschapper uit zich in de manier waarop zij problemen aanpakken: een wetenschapper zoekt een oorzaak, een ingenieur zoekt een oplossing.

8. Het grote aanbod van eenvoudig toegankelijke informatie op het internet beperkt de ontwikkeling van nieuwe ideeën en stimuleert het ‘simpelweg’ toepassen van bestaande kennis.

9. De beste maat voor succes is tevredenheid. (naar Richard Branson)

10. De faculteitsgebouwen van de TU Delft zouden altijd toegankelijk moeten zijn, ook ‘s nachts. Het verlenen van 24/7 toegang tot werkplekken vergroot de onderzoekspromotiviteit van ieder instituut.

11. Frequent lange-afstandsreizen per trein zijn vruchtbaar voor wetenschappelijk onderzoek.

Deze stellingen worden opponeerbaar en verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotor prof. dr. ir. R. F. Hanssen.
You are cordially invited to the public defence of my PhD thesis 'Geodetic network design for InSAR: Application to ground deformation monitoring' on Friday 22 May 2015 from 10:00 - 11:00 in the Senaatzaal at the Aula of the TU Delft.

I will give a brief introduction to the thesis at 9:30.

A reception will follow immediately after the ceremony.

Pooja S. Mahapatra