Vol. 30

Remote Sensing Series

David Small

Generation of Digital Elevation Models through Spaceborne SAR Interferometry

Remote Sensing Laboratories
Department of Geography
University of Zurich, 1998

This volume was approved as a Ph.D. thesis by the Faculty of Philosophy II (Faculty of Sciences and Mathematics) of the University of Zürich in the winter semester 1997/98 on the basis of evaluations from Prof. Dr. Daniel Nüesch and Prof. Dr. Harold Haefner.

Die vorliegende Arbeit wurde von der Philosophischen Fakultät II (Mathematisch-naturwissenschaftliche Fakultät) der Universität Zürich im Wintersemester 1997/98 aufgrund der Gutachten von Prof. Dr. Daniel Nüesch und Prof. Dr. Harold Haefner als Inaugural-Dissertation angenommen.

Copies are available from: Remote Sensing Laboratories
Department of Geography
University of Zürich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland

FAX: +41 +1 635-6848
`http://www.geo.unizh.ch/`

Printed by the *Druckerei der Zentralstelle der Studentenschaft der Universität Zürich*

Leaving the tumultuous throng
To cut across the reflex of a star;
Image, that flying still before me, gleamed
Upon the glassy plain.

Yet still the solitary cliffs
Wheeled by me -- even as if the earth had rolled
With visible motion her diurnal round!

William Wordsworth (1770-1850)
- *The Prelude*

Preface

The idea of using the phase information within synthetic aperture radar imagery to resolve the terrain-geolocation height ambiguity is especially appealing in areas with topography as rugged as the Swiss Alps and the Canadian Rockies.

I was at once attracted to this "border region" between the domains of electrical engineers' radars and geographers' mapping systems. Prof. Dr. Nüesch's proposal of research on the subject of InSAR was both enticing and timely, as data from ESA's ERS-1 satellite became increasingly plentiful soon thereafter.

The Remote Sensing Laboratories at the University of Zürich's Department of Geography have a long tradition in the treatment of terrain-induced distortions found within remote sensing images.

This dissertation describes work done in the years 1992 to 1997 under the supervision of Prof. Dr. Daniel Nüesch and Prof. Dr. Harold Haefner.

I thank my RSL colleagues Arnold Barmettler, Stefan Füglistaler, Hilko Hoffman, Philipp Hirtz, Philippe Meuret, Dr. Erich Meier, Jens Piesbergen, Peter Pirchl, Adrian Schubert, Stefan Spörri, Oliver Stebler, Dr. Martin Suter, and Ruedi Wettstein for fertile discussions and contributing to a pleasant work environment. Collaborations in the early years with Dr. Wolfgang Goblirsch, Dr. Charles Werner and Dr. Urs Wegmüller will not soon be forgotten. Joachim Hagger, Jens Piesbergen and Bruno Weber deserve thanks for trying (and usually even succeeding) in keeping the computer system in check. Suggestions for improvements to ZIP from my RSL colleagues Stefan Biegger, Thomas Egli, Stefan Füglistaler, and Oliver Stebler yielded fruitful enhancements. Late night discussions at RSL with Drs. Francesco Holecz and Paolo Pasquali revealed novel problem solving approaches. I'm grateful to Dr. Urs Frei for his impromptu "crash course" in SAR geocoding once I arrived in Zürich, and to Dr. Christoph Graf for having introducing me to the world of map projections.

Collaborations on InSAR and SAR geocoding with the German Aerospace

Research Establishment were critical in providing both continual deadlines for software deliveries, as well as feedback on its effectiveness and requests for new “features”. Nico Adam, Dr. Richard Bamler, Dr. Gordon Davidson, Dr. Dirk Geudtner, Dr. Dieter Just, Detlev Kosmann, Walter Knöpfle, Achim Roth, Dr. Marcus Schwäbisch, and Gunter Schreier are all thanked for their contributions.

Many of the results presented here were obtained in the course of a contract with Matra-Cap Systèmes, France; their support is also acknowledged.

Steve Coulson, Gaute Solaas, and Jürg Lichtenegger all deserve thanks for making themselves available as contacts at ESA.

The Swiss Federal Office of Topography deserves thanks for allowing use of their *DHM25* elevation model as a reference. The German ERS Processing and Archiving Facility is thanked for providing German reference elevation models and precise orbit data. The support of the Swiss Meteorological Institute in providing meteorological measurements is gratefully acknowledged. The Swiss Federal Statistical Office is also thanked for making available their database of Swiss land use statistics. The provision of a tree height reference map by the Swiss Federal Institute for Forest, Snow, and Landscape Research is also appreciatively acknowledged.

I thank Prof. Dr. Harold Haefner, for his continual interest in the ongoing work, and for always having an open door.

I am especially indebted to Prof. Dr. Daniel Nüesch not only for his long standing encouragement and supervision, but also for his constant enthusiasm and support for the research and project work.

Finally, Marina has been there for me throughout, not in the least, by following me to the “old continent” and building a common future on its shores. Now, let a new dance begin...

Zürich, October 1997

David Small

Summary

The generation of digital elevation models (DEM's) through spaceborne SAR interferometry (InSAR) is a complex undertaking that works best when coherence, baseline (height sensitivity), and phase unwrapping considerations are balanced optimally.

Although it is a relatively simple matter to generate a flattened interferogram, and this has often been marketed as a "topographic contour map", the truth is that many non-trivial steps are necessary to move from this stage to the geocoded height model desired.

The generation of the DEM includes coregistration of the radar images, interferogram formation, flattening, resolution of the 2π ambiguity (phase unwrapping), unflattening, height calculation, and geocoding. Each of the above steps is described in some detail, and aspects of the SAR image acquisition process relevant to DEM generation are highlighted. Validation of the geocoded InSAR height models is carried out using DEM flattening, as well as forward and backward geocoding methodologies.

Repeat-pass InSAR test sites in Germany and Switzerland were chosen to be representative of flat and rugged temperate vegetated areas. Areal validation of height maps generated by ERS InSAR provides confidence in the technique. Height accuracies of 2.7 m were achieved over a predominantly coherent 12×13 km area near Bonn, Germany using three-day repeat ERS-1 data. Systematic biases in the height estimation were minimal over a 40×50 km standard ERS-1 quarter scene. In a more rugged 50×50 km area surrounding Bern, Switzerland height accuracies of approximately 30 m were achieved; however phase unwrapping was not successful in some areas, removing them from further consideration. Longer repeat-pass intervals, shorter baselines, lower coherence, or more rugged topography reduce the achievable accuracy.

Spectral-shift filtering dramatically decreases phase variance and is of critical importance for the large baselines that are optimal for the extraction of topography. Using ERS data, the optimum baseline for mapping purposes is dependent upon the slopes that can be expected within the scene. Given a relatively flat scene, with gently rolling topography, a baseline of 300-400m provides good height accuracy, while at

the same time not excessively sacrificing spatial resolution during the spectral shift filtering.

Consistency checks using multiple interferograms can be used to improve accuracy and assist in phase unwrapping. Phase unwrapping errors must be either manually corrected, or reduced through verification of consistency with data from additional interferograms.

Methods that increase the scope of information considered promise improved reliability. Improvement of the height estimate through combination of multiple tandem ERS-1/ERS-2 pairs is advisable - in severely sloped areas, combination of ascending/descending pairs is necessary to offer a more consistent ground resolution across the scene.

Given the right conditions, ERS interferometry can produce height models with respectable accuracy. The interferograms used must have a high enough coherence, which can be difficult to find in vegetated areas - but appears to peak in the winter. Another limitation is found in high slopes that challenge phase unwrapping. To some extent the conditions are contradictory - it is precisely in the scenes with good height sensitivity through the virtue of a large baseline that phase unwrapping becomes most difficult.

Present-day spaceborne InSAR systems suffer from lack of optimization to the task of height model generation. All existing systems are repeat pass, and suffer from unpredictable temporal decorrelation and baseline extent. Spatial decorrelation can be counteracted using spectral filtering, but decorrelation due to noise from low SNR signals, and due to temporal decorrelation between acquisitions must be taken as a given.

Height model generation using spaceborne repeat-pass InSAR can only be recommended given that baseline, slope, and coherence constraints are all satisfied. The baseline must be large enough to guarantee sufficient height sensitivity, the slopes must not exceed limits dictated by the ambiguity resolvable during phase unwrapping, and coherence must be maintained between acquisitions to allow reliable phase difference estimation. The weakness of ERS InSAR height derivation lies in hilly forested areas, where low coherences combine with topography to render height estimation problematic.

Future missions will incorporate multiple antennae, and may provide more accurate timing and orbit ephemeris information, eliminating the requirement for refinement of image and baseline geometries.

Single-pass systems that circumvent the problem of temporal decorrelation through simultaneous acquisition while maintaining an optimal baseline promise to realize the full potential of DEM generation using space-based InSAR.

Zusammenfassung

Die Erstellung digitaler Höhenmodelle (DHM) mittels satellitengestützter SAR Interferometrie (InSAR) ist ein komplexes Verfahren, das am besten funktioniert wenn Kohärenz, Basislinie (Höhenempfindlichkeit) und Phasenentflechtung optimal aufeinander abgestimmt sind.

Obwohl die Generierung eines abgeflachten Interferogramms relativ einfach ist, müssen einige nicht triviale Schritte unternommen werden, um das Interferogramm (fälschlicherweise oft als topografisches Höhenmodell vermarktet) in ein geokodiertes Höhenmodell umzuwandeln.

Die Erstellung eines DHM beinhaltet die Koregistrierung der Radarbilder, die Erstellung des Interferogramms, dessen Abflachung, die Auflösung der 2π-Mehrdeutigkeit (Phasenentflechtung), die Aufhebung der Abflachung, sowie Höhenkalkulationen und Geokodierung. Jeder dieser genannten Schritte werden beschrieben und Aspekte des SAR Bildaufnahmeprozesses, die für die DHM Herstellung relevant sind, hervorgehoben. Zur Validierung des geokodierten Höhenmodells werden die Methoden der DHM Abflachung, sowie der Vorwärts- und Rückwärtsgeokodierung hinzugezogen.

Wiederholt überflogene InSAR Testgelände in Deutschland und der Schweiz wurden so gewählt, dass sie Vegetationsflächen mit unterschiedlicher Topografie umfassten. Flächendeckende Genauigkeitskontrollen der durch ERS InSAR generierten DHM erzeugten Vertrauen in die Methode.

Über ein 12×13 km grosses Gebiet bei Bonn, welches im 3 Tage Rhythmus von ERS-1 überflogen wurde, konnten Höhengenauigkeiten von 2,7 m erzielt werden. Systematische Verschiebungen der Höhenschätzungen über eine 40×50 km grosse ERS-1 Viertelszene waren minimal. Über eine zerklüftetere 50×50 km grosse Fläche in der Region Bern konnten Höhengenauigkeiten von nur ungefähr 30 m erzielt werden. Allerdings war die Phasenentflechtung bei einigen Ausschnitten nicht erfolgreich. Faktoren wie längere Intervalle zwischen den Aufnahmen, kleinere Basislinien, tiefere Kohärenz und extrem zerklüftete Topo-

grafien reduzierten die erzielbaren Genauigkeiten.

Besonders bei grossen Basislinien, welche grosse Höhengenauigkeiten erlauben, führt das Filtern unter Berücksichtigung der Spektralverschiebungen zu einer deutlichen Verringerung der Phasenvarianz, und somit zu einer Verbesserung der Höhengenauigkeit. Die optimale Basislinie bei Kartierungen ist von den Hangneigungen abhängig, die in der Szene erwartet werden können. Bei relativ flachem Gelände mit sanften Hügeln erlauben Basislinien von 300 bis 400 m bei ERS-Daten gute Höhengenauigkeiten, ohne dass die räumliche Auflösung durch die Filterung übermässig beeinträchtigt wird.

Kontrollen der Phasen- bzw. Höhenübereinstimmungen mehrerer Interferogramme können benutzt werden, um Phasenentflechtung zu erleichtern und die Höhengenauigkeit zu verbessern. Fehler, die bei der Phasenentflechtung entstehen, müssen entweder manuell korrigiert oder durch weitere Kontrollen mittels Daten von zusätzlichen Interferogrammen reduziert werden.

Methoden, die den Umfang der berücksichtigten Informationen erweitern, versprechen eine erhöhte Zuverlässigkeit. Zur Verbesserung der Höhenschätzung ist die Kombination mehrerer Tandem ERS-1/ERS-2 Paare zu empfehlen. In gebirgigem Gelände ist die Kombination von Bildern absteigender/aufsteigender Satellitenbahnen notwendig, um eine konsistentere Bodenauflösung über die Szene zu gewährleisten.

Ausgehend davon, dass vorteilhafte Aufnahmebedingungen vorliegen, kann ERS Interferometrie Höhenmodelle mit respektabler Genauigkeit generieren. Die Interferogramme müssen ein Mindestmass an Kohärenz überschreiten, um eine Phasenschätzung zu erlauben. Vegetationsbedingte und saisonale Änderungen, sowie starke Hangneigungen üben einen negativen Einfluss auf die erzielbare Höhengenauigkeit aus. Stark bewachsene Gebiete weisen demnach eine geringere Kohärenz auf, wobei diese bei ruhigen Winterszenen vergleichsmässig hoch liegt. Starke Hangneigungen erschweren die Phasenentflechtung. Dieses führt zur paradoxen Situation, dass die Szenen mit der besten Höhenempfindlichkeit aufgrund ihrer grossen Basislinien die grössten Schwierigkeiten für die Phasenentflechtung darstellen.

Derzeit angewandte satellitengestützte InSAR-Systeme leiden unter der Tatsache, dass sie nicht zur Erstellung von Höhenmodellen konzipiert wurden und nicht dafür optimiert sind. Existierende Systeme erfordern Mehrfachüberflüge, da ihre Einzelantennen gleichzeitige Aufnahmen unmöglich macht. Weil die Aufnahmen zeitlich verschoben erfolgen müssen, sind Schwankungen der temporalen Dekorrelation

und Basislinienwerte nicht kontrollierbar.

Räumliche Dekorrelationen können durch Spektralfilterungen neutralisiert werden. Hingegen können Dekorrelationen, die auf tiefe SNR und/oder temporale Dekorrelationen zwischen Aufnahmen beruhen, nicht korrigiert werden.

Höhenmodellgenerierungen mittels satellitengestützter, multitemporaler InSAR-Überfluge können nur dann empfohlen werden, wenn die Basislinien-, Hangneigungs- und Kohärenz- Bedingungen erfüllt sind. Die Basislinie muss gross genug sein, um eine genügende Höhenempfindlichkeit zu gewährleisten. Die Hangneigungen dürfen die Grenzen, welche durch die Auflösung der Mehrdeutigkeit bei der Phasenentflechtung gegeben sind, nicht überschreiten. Die Kohärenz zwischen den zwei Aufnahmen muss eine zuverlässige Phasenschätzung erlauben.

Die Schwäche der ERS InSAR-Höhenableitung liegt in der Anwendung bei hügeligen, bewaldeten Gebieten, in denen tiefe Kohärenz in Kombination mit zerklüfteter Topografie zu Höhenbestimmungsproblemen führen.

Zukünftige Missionen müssen genauere Informationen bezüglich Zeit- und Umlaufbahnmessungen liefern, um die Notwendigkeit der Verfeinerung der Bild- und Basislinien-geometrien zu beseitigen.

Durch gleichzeitige Aufnahmen mittels Doppelantennen wird das Problem der temporalen Dekorrelation umgangen, während eine optimale Basislinie aufrechterhalten wird. Solche Systeme versprechen das volle Potential der DHM Herstellung mittels satellitengestütztem InSAR zu realisieren.

Table of Contents

Chapter 1

Introduction and Objectives

1.1 Introduction

Digital elevation models have a wide variety of applications, among them hydrological modelling, correction of terrain-induced distortions implicit in a sensor's acquisition geometry (terrain-geocoding), as well as radiometric correction, and aid in thematic interpretation (classification, snow cover mapping, biomass estimation).

The production of digital elevation models (DEM's) has long been a staple in the remote sensing business, and provides a core business, indeed, almost a *raison d'être* for the French SPOT series of satellites. Using this conventional technique, image pairs are processed using mature optical stereo methods to generate the elevation model [43]. In recent years, a technique known as *SAR Interferometry* (InSAR) has gained attention. First used in the 1960's to resolve ambiguity in measurements of Venus [72], and later in the 1970's [29] aimed at topographic mapping for the first time, the technique has undergone rapid development in the 1990's (due to widespread availability of ERS data), emerging as a potential competitor to conventional optical stereo methods.

Whereas photogrammetric stereo methods must await two cloud free acquisitions, SAR data acquisitions are less weather dependent. The influence of water vapour on C-band interferograms remains a topic of research. Optical stereo relies on high contrast within the scene, while SAR Interferometry has no such limitation. For high resolution DEM production, both optical stereo and airborne InSAR rely on very expensive airborne campaigns (although InSAR is not as dependent on the cloud-free constraint). SAR interferometry can potentially service large swaths, with a tandem or (better yet) a single-pass, dual-antenna configuration providing systematic large-area coverage.

Use of altimeters is restricted to low resolution applications. Ground surveys are prohibitively expensive (and sometimes politically impossible); cartographic data sources are expensive, and sometimes unavailable. Laser scanning LIDAR suffers from narrow swaths, and remains experimental.

The work that follows aims to evaluate the spaceborne repeat-pass InSAR technique for DEM generation.

In order to test the accuracy of ERS repeat-pass interferometry, test sites with

high quality reference DEM's are required. The InSAR data processing algorithm is described, and applied to data from the test sites. Results are presented, and conclusions are drawn.

The remainder of this chapter begins with a review of some basic concepts of Synthetic Aperture Radar (SAR) and Synthetic Aperture Radar Interferometry (InSAR), followed by an overview of the rest of the thesis.

1.2 Problem Definition and Objectives

Before a digital elevation model can be generated, a sequence of sub-problems [19][21][82] must each be carried out to a satisfactory conclusion. One begins with selection of the radar data acquisitions for processing, followed by co-registration into a common geometry, interferogram formation, removal of an ellipsoid or known reference-DEM geometry, and calculation of a coherence measure for each pixel in the scene.

The interferogram must then be "unwrapped" (the 2π ambiguity removed), the geometry refined, the height model generated, and the data "geocoded" into a map geometry common with other reference information available for the scene. Geocoding consists of solving equations describing the radar geometry (range sphere, Doppler cone) to transform each point within the scene into a geodetic reference geometry, which may then be transformed into a selected cartographic projection.

Each processing step will be covered in more detail in the chapters to come. The remainder of this introductory chapter reviews basic concepts of Synthetic Aperture Radar (SAR) and SAR Interferometry (InSAR), and outlines the structure of the chapters that follow.

After more detailed coverage of each step, InSAR-generated height models are compared with reference DEM's in a validation process, with results obtained from selected test sites. The validation results are discussed and evaluated. Elevation models are then applied to geometric and radiometric calibration of SAR images, and conclusions are drawn.

1.3 Definition of Key Terms

Examination of some common terms may help avoid later misunderstandings. Basic concepts of Synthetic Aperture Radar (SAR) and SAR Interferometry (InSAR) are therefore reviewed here.

1.3.1 Synthetic Aperture Radar

The geometry of satellite Synthetic Aperture Radar systems is illustrated in Figure 1-1. SAR data is imaged using a radar transmitting and receiving pulses perpendicular to its flight direction (side-looking) at an angle of incidence on the ground of θ (degrees). A series of processing steps produces 2D imagery aligned along iso-range lines in the

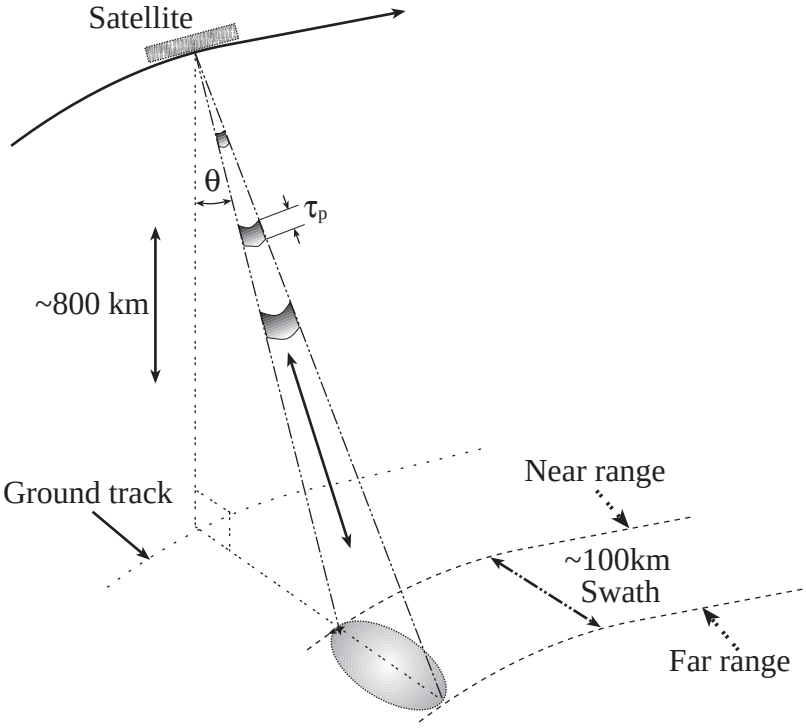


Figure 1-1: SAR imaging geometry (after [7])

range dimension, and ordered by iso-Doppler lines in the flight track (*azimuth*) dimension.

Range Compression (cross-track): The range resolution of a radar in the absence of pulse compression is determined by the length of the pulse τ_p (in seconds), as two objects must be separated by at least $c \cdot \tau_p/2$ in slant range to appear distinct, where c is the speed of light. Short pulses therefore provide the best resolution. However, peak power limitations at the transmitter restrict the minimum pulse length, making longer durations necessary.

One can resolve these conflicting requirements by marking different parts

of the transmitted signal with different carrier frequencies. Linear Frequency Modulation (FM), where the carrier frequency changes over the pulse duration ("chirped") allows combination of high resolution with low transmitted power requirements.

Given a chirp waveform $s(\tau)$, and a terrain reflectivity impulse response function $T(\tau)$, the echo returned is

$$E(\tau) = s(\tau) \otimes T(\tau), \quad \textbf{(1-1)}$$

where $\otimes$ is the convolution operator. The goal is then to recover the terrain reflectivity function $T(\tau)$ given the received echo $E(\tau)$. This requires convolution with the inverse function of a reconstructed transmitted range chirp (either analytical or a measured replica):

$$T(\tau) = s^{-1}(\tau) \otimes E(\tau). \quad \textbf{(1-2)}$$

The inverse function of a linear FM range chirp is a matched filter [7]:

$$T(\tau) = s^{*}(-\tau) \otimes E(\tau). \quad \textbf{(1-3)}$$

The correlation can be performed efficiently in the frequency domain. Given the range chirp spectrum $S(f) = FT\{s(\tau)\}$, the spectrum of the matched filter is then

$$H(f) = W(f) \times S^{*}(f), \quad \textbf{(1-4)}$$

where $W(f)$ is an amplitude weighting function used to balance a trade-off between optimum resolution and sidelobe interference. The range compressed terrain reflectivity data is then:

$$T(\tau) = FT^{-1}\{H(f) \cdot E(f)\} \quad \textbf{(1-5)}$$

The echoes have been autocorrelated with a reconstructed transmitted chirp to improve the resolution in the cross-track (range) direction. The slant range resolution R_r is no longer directly dependent on the pulse length, but on the range chirp bandwidth [7]:

$$R_r = c/(2B_r) \quad \textbf{(1-6)}$$

where B_r is the chirp bandwidth (see Appendix A for a description of variables). In the case of the ERS satellites [76], the achievable range resolution is improved to approximately 9.6m in slant range (~22m in ground range).

Azimuth Compression (along-track): In real-aperture radar systems, the footprint of the antenna beam on the ground places a limit on the radar along-track (azimuth) resolution: *small* footprints are therefore best, as resolution worsens with increasing range. A *large* antenna is needed for a *small* footprint - this is, however, impractical for both airborne and spaceborne satellite missions.

One can resolve this dilemma through a technique known as Synthetic Aperture Radar (SAR).

As with the sound of a passing train, the Doppler frequency of a given point *P*'s received microwave echoes change as the satellite approaches and recedes from *P*. Spatially scattered echo contributions (P_1, P_2, …, P_N) are each marked by their individual Doppler frequency. As seen previously with range compres-

sion, a matched filter (tuned this time to Doppler variations) significantly improves the achievable resolution. Azimuth compression step is actually more complicated, as a single point's echoes are distributed not linearly along the azimuth dimension, but over a near-quadratic curve broadly aligned along the azimuth dimension. Treatment of this "range migration" increases the complexity of the azimuth compression step.

Azimuth compression can also be considered as the "synthesis" of an antenna aperture along the radar's flight track, hence the name synthetic aperture radar. The attainable resolution approaches that achievable with an antenna extended along the length of the flight track.

Through mathematical derivation, it can be shown [11] that the achievable resolution A_r is equal to half the physical antenna length L_A:

$$A_r = L_A/2 \qquad \textbf{(1-7)}$$

The azimuth resolution is also independent of range, as wider footprints at far range are exactly compensated by longer Doppler integration times.

Figures 1-2 through 1-4 show examples of ERS-1 raw, range-compressed, and azimuth-compressed images respectively. All images were reduced in size through averaging for display of a complete 50x50km ERS-1 quarter scene. Even with such averaging, the size of a resolution element remains significantly larger than a single display pixel in the raw data (Figure 1-2), whereas the form of Lake Biel and Lake Murten become successively clearer in the range-compressed (Figure 1-3) and azimuth-compressed (Figure 1-4) images.

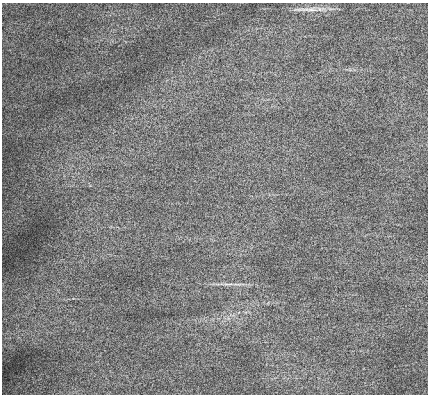


Figure 1-2: ERS-1 raw data - Bern, Switzerland - Nov. 24, 1991 [© ESA]

A processed SAR image (compressed in both range and azimuth) of Zürich, Switzerland is shown in Figure 1-5. Lake Zürich and the Greifensee are visible at the bottom centre and right. The smooth (and therefore dark) runways of Zürich-Kloten and Dübendorf airports are visible at top and centre-right, while bright targets within the city are visible near the centre of the image. The resolution achieved would

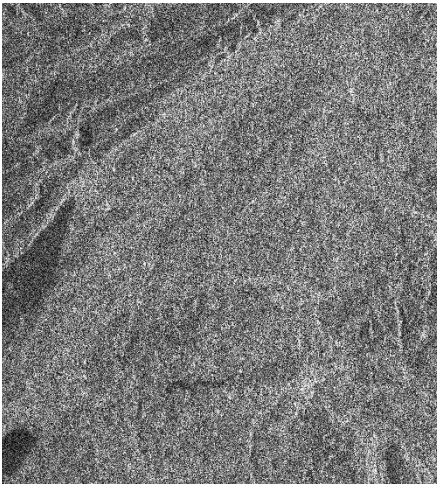


Figure 1-3: ERS-1 detected range-compressed data - Bern, Switzerland - Nov. 24, 1991 [© ESA]

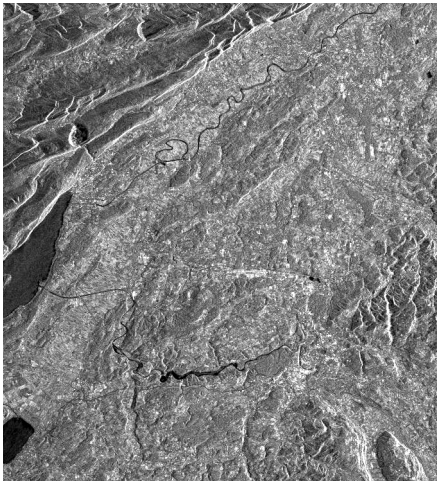


Figure 1-4: ERS-1 detected range and azimuth-compressed SLC data - Bern, Switzerland - Nov. 24, 1991 [© ESA]

have differed by orders of magnitude without range and azimuth compression.

However, the azimuth compression step is only possible when objects remain coherent over the length of the data acquisition (synthetic aperture length) - the phase information must be preserved. This is certainly not a problem for range compression, operating on the scale of the speed of light. The pulses used during azimuth compression are spread over a much longer time along the satellite's orbit (generally fractions of a second), but even ocean waves generally stay coherent over such a time scale, though before the launch of SEASAT this was subject to debate [7]. Due to their differing time scales, the range and azimuth dimensions are sometimes also referred to as the fast and slow-time axes, respectively.

Preservation of the phase information for compression purposes lead to the idea of using the phase for other applications. Phase *difference* information between two acquisitions can be used to calculate a height model.

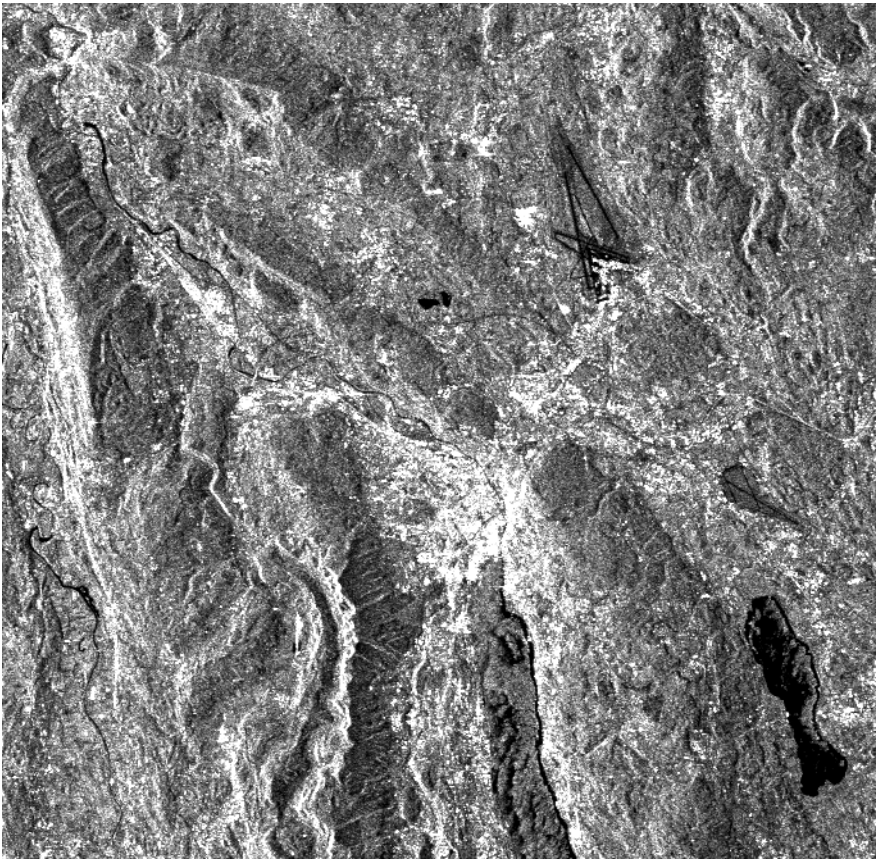


Figure 1-5: ERS-1 SAR amplitude image - Zürich, Switzerland - Nov. 4, 1995 [© ESA]

1.3.2 Synthetic Aperture Radar Interferometry

SAR imagery is two dimensional, as noted previously, ordered along iso-range and iso-Doppler lines. In the past, the 2-D nature of the data has been an obstacle to interpretation, as multiple points (at different elevations) can have the same range distance, causing ambiguity. A highly accurate elevation model

was required to transform images from their range-Doppler geometry into map reference systems, using a method known as SAR terrain geocoding [57].

Until relatively recently, conventional SAR image interpretation dealt only with the image amplitude. However, phase information is also present, and given that it is preserved by the SAR processor, the local phase information from multiple SAR acquisitions can be combined to resolve that 2-D ambiguity: local height variations can be extracted, and a 3-D model constructed.

The geometry of satellite SAR interferometry is illustrated in Figure 1-6. A SAR image is acquired by two antennae A_1 and A_2 at positions $\underset{\sim}{S}_1$ and $\underset{\sim}{S}_2$ respectively. The range distance to a point P within the image swath is R_1 from antenna A_1, and R_2 from antenna A_2.

Interferometers can be configured with a single or multiple antenna(e), and acquire their data either in a single or multiple pass(es).

Single Antenna, Repeat Pass: This is the cheapest system design, and also the most common. A single antenna both transmits and receives, acquiring first one image, then, in a repeat pass, the second. Difficulties with accurate aircraft guidance as well as motion compensation kept the configuration in the realm of spaceborne systems until the advent of recent improvements in aircraft avionics [31].

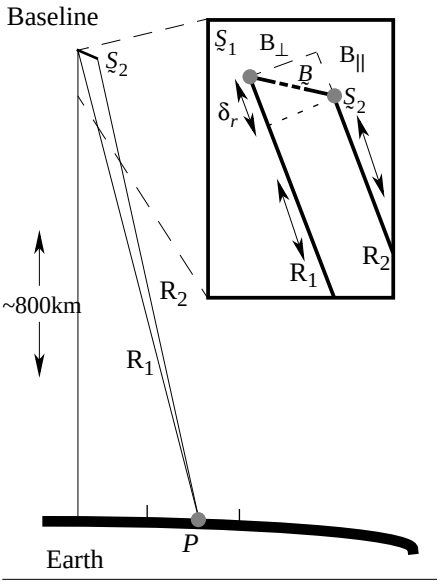


Figure 1-6: InSAR geometry - repeat pass

SAR interferometry has been performed using data from the following spaceborne SAR sensors: Seasat [44], SIR-B [15], the ESA ERS satellites [62], JERS-1 [73], SIR-C/X-SAR [5][61], Almaz [95], and RADARSAT-1 [20].

Dual Antennae, Single Pass: This system design consists of two antennae mounted on the same platform that view the scene simultaneously, a geometry

that has come to be widespread with airborne systems. To date, there have been no civilian satellite SAR systems. The absence of two-antennae civilian satellite SAR systems has meant that researchers have focused on repeat-pass InSAR techniques. NASA currently plans to launch a shuttle mission with two radar antennae on a topographic mapping mission (SRTM).

Dual Antennae, Repeat Pass: In this configuration, two satellites are used. The first satellite flies over the terrain, acquiring its data. After a repeat-pass interval, the second satellite acquires the scene. The ERS tandem configuration [9] was the first operational mission in this category. Logistical downlink interference prevented simultaneous acquisitions by the two satellites.

Difference in Distance: Radar sensors are devices that measure range. In SAR interferometry, the interferometric phase is directly related to the difference in range measurements made during the acquisitions.

Each pixel area within a SAR image processed to a single-look-complex stage contains phase information that is dependent on both the range R to the pixel and the scattering properties within the pixel:

$$\phi_1 = 2 \cdot \frac{2\pi}{\lambda} \cdot R_1 + \phi_{pixel,1} \quad \textbf{(1-8)}$$

$$\phi_2 = 2 \cdot \frac{2\pi}{\lambda} \cdot R_2 + \phi_{pixel,2} \quad \textbf{(1-9)}$$

where λ is the radar wavelength, ϕ_{pixel} describes the contribution to the phase from the scattering properties within the pixel, and the subscript 1 or 2 denotes the acquisition. The initial factor of two is necessary in repeat-pass configurations, where both transmit and receive paths are unique for each acquisition. In a single-pass configuration, the transmit path is often shared; in such cases, the initial factors of two in Equations 1-8 and 1-9 are eliminated.

Since the contribution of the pixel's scattering properties to the phase is generally not known, the phase values within a single image are ambiguous. However, given that the scatterers are not disturbed between acquisitions (no temporal decorrelation), and assuming for the moment no decorrelation due to baseline viewing angle, then

$$\phi_{pixel} = \phi_{pixel,1} = \phi_{pixel,2} \quad \textbf{(1-10)}$$

In that case, by forming the interferometric phase difference between two acquisitions,

$$\phi = \phi_2 - \phi_1 \quad \textbf{(1-11)}$$

the contribution from ϕ_{pixel} cancels out, and one is left with the geometrical contribution from the difference in ranges $\delta_r = R_2 - R_1$:

$$\phi = 2 \cdot \frac{2\pi}{\lambda} \cdot (R_2 - R_1) = 2 \cdot \frac{2\pi}{\lambda} \cdot \delta_r, \quad \textbf{(1-12)}$$

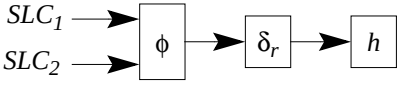


Figure 1-7: Height estimation through SAR interferometry

where ϕ is known as the *interferometric phase*, or the argument of the complex *interferogram*. It is directly related to the difference in distances from the satellite's two vantage points. As outlined in Figure 1-7, the generation of DEM's through SAR interferometry consists of first generating a good estimate of ϕ, then transforming that into the best possible estimate of δ_r, and finally transforming that estimate into height values on a map grid (geocoding).

The geolocation problem (transforming the data collected in SAR slant-range geometry into a map projection) can be considered as a solution of equations describing range spheres with radii R_1 and R_2 centred on A_1 and A_2 respectively, and their associated Doppler frequencies. The solution of these equations is detailed in Chapter 4.

Far-Field Approximation: To enable simplified models, the slant range "look" vectors for the two acquisitions may be assumed to be parallel (in the far field). The internal angle ξ in the Doppler "plane" between the baseline vector and the look vector is approximated [44] by:

$$\cos(\xi) \approx \delta_r / |\underset{\sim}{B}| , \qquad \textbf{(1-13)}$$

where the length of the baseline vector $\underset{\sim}{B}$ is $|\underset{\sim}{B}|$. The approximation has validity where the interferometric baseline is much smaller than the range distance $B \ll R_1$. In that case, inspection of Figure 1-8 shows that

$$B_\perp \approx R\Delta\theta , \qquad \textbf{(1-14)}$$

and for modelling purposes (e.g. sensitivity analysis) and/or restricted scene sizes, this is generally sufficiently accurate. However, for precise final calculations more exact models are required.

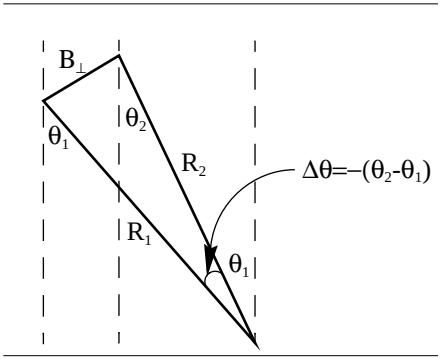


Figure 1-8: Relation between $\Delta\theta$ and $B_\perp$

The approximation is invalid for airborne geometries, and is also inaccurate for satellite geometries with large

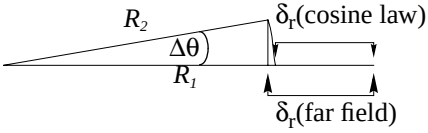


Figure 1-9: Far-field approximation

swaths. The difference is illustrated graphically in Figure 1-9.

Sensitivity of phase to height:

Adopting a simplified 2D geometry [44] with parallel orbits and a flat Earth (see Figure 1-10 for an illustration), the component of the interferometric baseline perpendicular to the line of sight is:

$$B_\perp = B_{\perp,x} + B_{\perp,y} = B_x \cos\theta + B_y \sin\theta \quad \textbf{(1-15)}$$

The interferometric phase of a point P is simply calculated as:

$$\phi = \frac{4\pi}{\lambda} \cdot (B_x \sin\theta - B_y \cos\theta), \quad \textbf{(1-16)}$$

where θ is the local incidence angle. Alternatively, one can rewrite Equation 1-16 in terms of H and R [34]:

$$\phi = \frac{4\pi}{\lambda} \cdot \left(B_x \sqrt{1 - \frac{(H-h)^2}{R^2}} - B_y \frac{H-h}{R} \right). \quad \textbf{(1-17)}$$

The sensitivity of the phase measurement to changes in height is then:

$$\frac{\partial\phi}{\partial h} = \frac{4\pi}{\lambda} \cdot \left(B_x \frac{\partial}{\partial h} \sqrt{1 - \frac{(H-h)^2}{R^2}} + B_y \frac{\partial}{\partial h} \frac{H-h}{R} \right), \quad \textbf{(1-18)}$$

simplifying to:

$$\frac{\partial\phi}{\partial h} = \frac{4\pi}{\lambda R} \cdot \left(B_x \cdot \frac{H-h}{\sqrt{R^2 - (H-h)^2}} + B_y \right).$$

Through trigonometry:

$$\frac{\partial\phi}{\partial h} = \frac{4\pi}{\lambda R \sin\theta} \cdot (B_x \cos\theta + B_y \sin\theta), \quad \textbf{(1-19)}$$

Applying Equation 1-15 one has a compact expression for the sensitivity of the interferometric phase to terrain height h:

$$\frac{\partial\phi}{\partial h} = \frac{4\pi}{\lambda} \cdot \frac{B_\perp}{R \sin\theta}. \quad \textbf{(1-20)}$$

Note that $\partial\phi/\partial h$ is a signed quantity, as is $B_\perp$. Their sign indicates the direction of change of phase given a positive change in height.

Ambiguity Height: The phase can only be measured modulo 2π. One can use Equation 1-20 to approximate the local relation between height changes and the corresponding interferometric phase increment:

$$\Delta\phi \approx \frac{\partial\phi}{\partial h} \cdot \Delta h \quad \textbf{(1-21)}$$

Note that both Equations 1-20 and 1-21 are approximations: higher accuracy

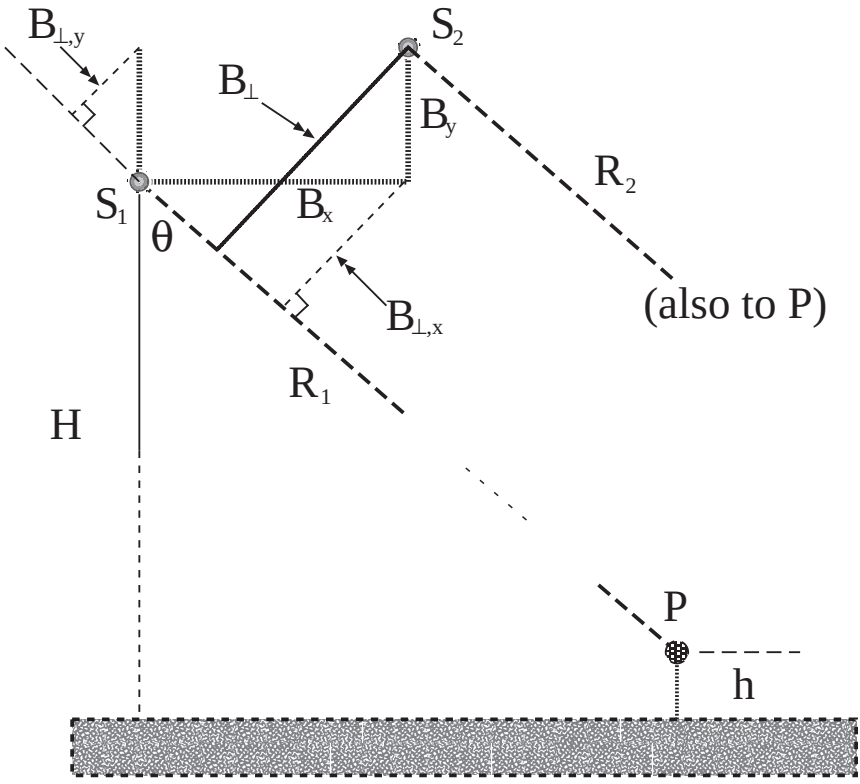


Figure 1-10: Simplified parallel-orbit geometry (after [44])

requires consideration of non-parallel look vectors as well as non-parallel orbits. However, the approximations suffice for general modelling of the sensitivity of the interferometric phase to changes in height.

The height interval corresponding to a "full fringe" ($\Delta\phi = 2\pi$) is known as the ambiguity height $\Delta h_{2\pi}$ (some refer to it as the "altitude of ambiguity"). Following from Equations 1-20 and 1-21, it is calculated as:

$$\Delta h_{2\pi} = \frac{\lambda}{2} \cdot \frac{R \sin\theta}{B_\perp} \qquad \textbf{(1-22)}$$

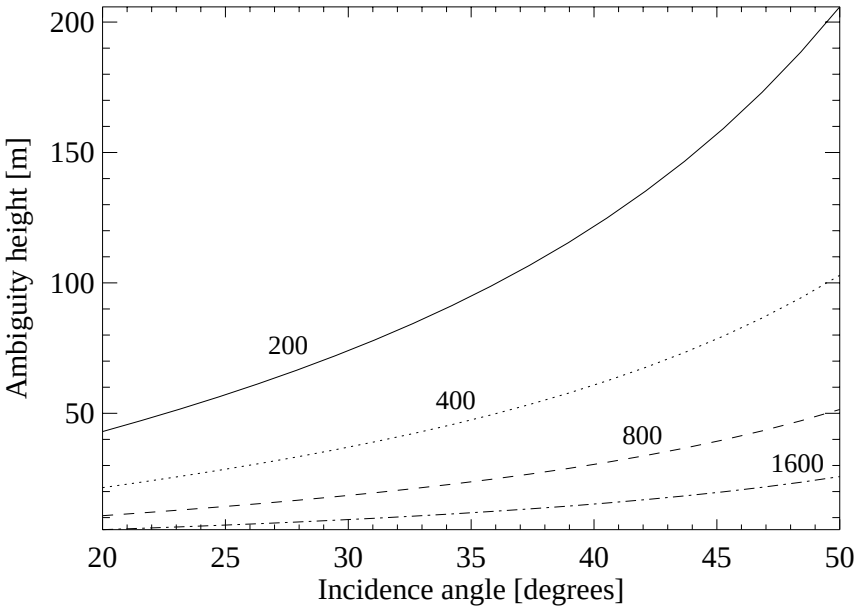
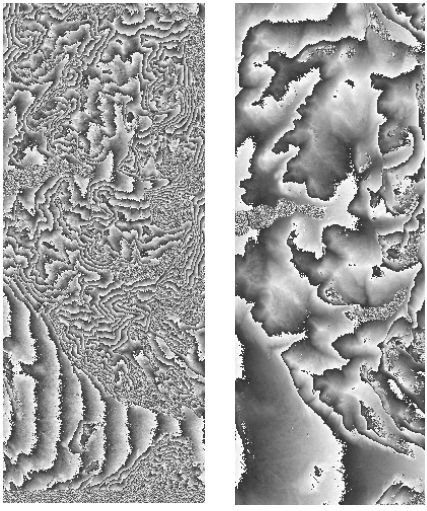


Figure 1-11: Plot of ambiguity altitude versus incidence angle for $B_\perp = 200m$ (solid), $B_\perp = 400m$ (dot), $B_\perp = 800m$ (dash), $B_\perp = 1600m$ (dash-dot)

For C-band satellite systems at an altitude of 785km, Figure 1-11 plots the ambiguity height vs. incidence angle for four perpendicular baseline values. Note that ERS nominal incidence angles range from 20° to 26°. The magnitude of the ambiguity height decreases as the baseline increases. Equation 1-20 tells us that larger baselines cause the interferometric phase to be increasingly sensitive to height changes. However, Figure 1-11 shows that at large baselines, the height ambiguity decreases: even small height changes could exceed $\Delta h_{2\pi}$, causing ambiguity that must be resolved through a process known as phase unwrapping.

Figure 1-12 illustrates the dependence of ambiguity height on the radar wavelength. The two SIR-C flattened interferograms were calculated using data acquired over an area surrounding the Sihlsee in Switzerland during the SIR-C/X-SAR mission of October 1994.

They show that the interferometric phase is approximately four times more sensitive to height changes at C-band than at L-band.



(a) (b)

Figure 1-12: Dependence of ambiguity height on radar wavelength - Oct. 1994 SIR-C flattened interferograms - Sihlsee, Switzerland: (a) C-band, (b) L-band

Coherence: The consistency of a phase measurement can be quantified with this measure. The phase values measured from microwave backscatter are composite values formed by complex summation of the contributions from individual scatterers within a single resolution cell (see Figure 1-13).Values vary between zero (completely inconsistent, random), and one (no variation from the composite value).

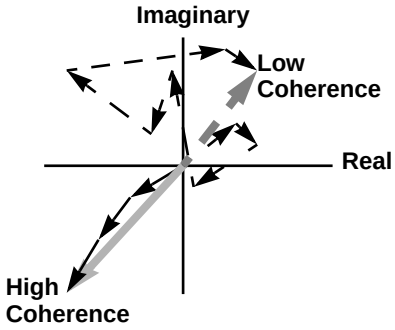


Figure 1-13: Coherence and phase measurement

Differential InSAR: Given more than a single pair of interferograms (i.e. more than a single interferometric phase value), the differential InSAR technique becomes possible. The multiple interferograms can come about either through combination of three or more acquisitions (double difference) or of two plus a synthetic interferogram calculated from a reference digital elevation model (DEM elimination). Movements of the terrain between acquisitions (e.g. an earthquake) can then be made with accuracies in the realm of centimetres.

1.4 Overview

Chapter 2 describes the reference systems used to specify satellite orbits and their relationships to one another (baselines).

Chapter 3 describes the processes leading up to interferogram and height model formation. Image co-registration (both coarse and fine), resampling, filtering of shifted image spectra, interferogram formation and flattening, and calculation of the interferometric coherence are all covered. An adaptive smoothing filter is then outlined, the phase unwrapping process is reviewed, and the transformation of phase differences to height values is described.

Chapter 4 deals with the geocoding of the interferometric terrain information: both forward and backward approaches are covered.

Chapter 5 demonstrates an application of DEM's: their use in image simulation of SAR images, and discusses some aspects of the relationship between radar and map geometries.

Chapter 6 treats the evaluation of the height accuracy estimates, and their error sources.

Chapter 7 concludes with a summary of the results achieved, and provides recommendations for further work.

Chapter 2

Orbits and Baselines

2.1 Introduction

This chapter describes the reference systems used to specify satellite orbits and their relationship to one another (orbital baseline) [32].

2.2 Reference Systems

2.2.1 Inertial

An inertial frame is useful for calculations involving Newtonian mechanics. It is fixed to the stars. Comparisons to points fixed on the Earth's surface can only be made through a transformation incorporating rotation about the Earth's axis using the Julian date and time.

Although the reference frame is generally already Earth-centred, the rotational position of the Earth (with respect to a reference date) must be calculated. This transformation produces coordinates in an Earth-centred rotating (ECR) reference system.

2.2.2 Earth Centred Rotating (*ECR*)

This reference frame is fixed to the Earth's surface, rotating and moving with the Earth as it moves around the sun. Not being an inertial frame, it does not lend itself to use in orbital mechanics. However, its close relationship to geographical coordinate systems makes it a necessary intermediary in referring satellite orbit information to locations on the Earth's surface.

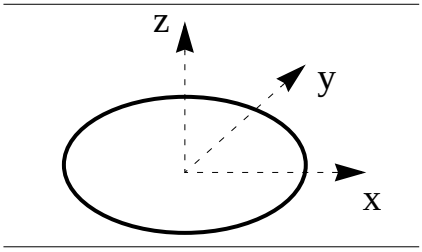


Figure 2-1: Earth-Centred-Rotating (*ECR*) coordinate system

Figure 2-1 shows the orientation of the Cartesian Earth-centred-rotating axes with respect to a geocentric oblate ellipsoid of the Earth:

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1 \qquad \textbf{(2-1)}$$

where a and b are the lengths of the semi-major and semi-minor axes respectively. Note that the z-axis is aligned with the Earth's poles, while the x and y axes are in the plane of the equator. The

x-axis intersects the prime meridian. While a geocentric ellipsoid has been chosen here for simplicity, details concerning non-geocentric ellipsoids are found in Chapter 4.

2.2.3 Tangential Cross-track Normal (*TCN*)

The geometry of an interferometric orbit pair is simplified if expressed in the "*TCN*" coordinate system, named for its *tangential*, *cross-track*, and *normal* components respectively.

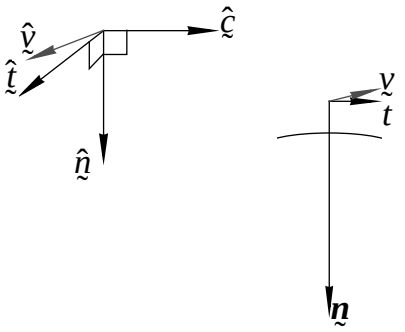


Figure 2-2: *TCN* coordinate system

The orthonormal basis vectors $\hat{\underset{\sim}{t}}$, $\hat{\underset{\sim}{c}}$, and $\hat{\underset{\sim}{n}}$ that define the *TCN* coordinate system are calculated as follows:

- The *normal* basis vector $\hat{\underset{\sim}{n}}$ is calculated from the mid-scene spacecraft position state vector:

$$\hat{\underset{\sim}{n}} = \underset{\sim}{n}/|\underset{\sim}{n}| \qquad \textbf{(2-2)}$$

  where $\underset{\sim}{n}$ is the vector pointing from the spacecraft to the Earth's centre.
- The *cross-track* basis vector $\hat{\underset{\sim}{c}}$ is formed from the cross-product of the normal basis vector and the spacecraft velocity state vector:

$$\hat{\underset{\sim}{c}} = (\hat{\underset{\sim}{n}} \times \underset{\sim}{v})/|\hat{\underset{\sim}{n}} \times \underset{\sim}{v}| \,. \qquad \textbf{(2-3)}$$

  Note that the velocity vector $\underset{\sim}{v}$ is *not* necessarily perpendicular to $\hat{\underset{\sim}{n}}$.
- The *tangential* vector $\hat{\underset{\sim}{t}}$ completes the orthonormal basis:

$$\hat{\underset{\sim}{t}} = \hat{\underset{\sim}{c}} \times \hat{\underset{\sim}{n}} \,. \qquad \textbf{(2-4)}$$

The orthonormal *TCN* coordinate system is useful for reducing the (by its nature) 3-D orbital baseline problem down to a simplified 2-D model.

2.3 Orbit Models

The orbital motion of a celestial body around another can be described in a number of ways. The appropriate manner of description is dependent upon the application at hand.

State Vectors: Measurements of satellite orbits are generally delivered to users in the form of a sequence of state vectors:

$$\begin{bmatrix} \underset{\sim}{S}^1 \\ \dots \\ \underset{\sim}{S}^N \end{bmatrix} = \begin{bmatrix} S_x^1 & S_y^1 & S_z^1 \\ \dots & \dots & \dots \\ S_x^N & S_y^N & S_z^N \end{bmatrix}, \qquad \textbf{(2-5)}$$

retrieved from an orbit model constructed on the basis of position and

velocity measurements performed from stations scattered across the globe. The state vectors can be delivered in inertial or *ECR* coordinates. Since SAR interferometry applications reference their data to locations on the Earth's surface, inertial vectors must be converted to *ECR* coordinates at this stage.

The user is then free to transform the available state vectors into an appropriate type of local orbit model. Three such models are described below.

2.3.1 Kepler Elements

Orbit propagators are used to extrapolate from a given set of state vectors to predict future orbital motion of a satellite. Such applications often prefer to describe orbits in terms of the Kepler elements: semi-major axis, inclination, eccentricity, argument of pericentre, longitude of node, and mean anomaly. A similar, related, set of parameters are the equinoctial elements.

While such a framework is necessary for orbit propagators, SAR scenes are typically limited to a few hundred kilometres in length. For modelling within such a small local area, simpler descriptions suffice.

2.3.2 Polynomial Fit

Given the state vectors described in Equation 2-5, one can fit a polynomial function to each dimension of the spacecraft position $\underset{\sim}{S}$ as a function of azimuth time t. The number of coefficients M of the polynomial must be less than or equal to the number of available state vectors. The spacecraft position is retrieved from the polynomial coefficients:

$$\underset{\sim}{S} = \begin{bmatrix} S_x \\ S_y \\ S_z \end{bmatrix} = \sum_{i=0}^{M-1} \begin{bmatrix} a_i \cdot t^i \\ b_i \cdot t^i \\ c_i \cdot t^i \end{bmatrix}, \qquad \textbf{(2-6)}$$

where a, b, and c denote the coefficients of the x, y, and z polynomial respectively. The sets of coefficients a, b, and c are calculated given a sufficient number of state vectors using a minimum least squares optimization technique.

The polynomial model offers a simple description of local orbit behaviour. Optimization techniques exist for refinement of such a model using tiepoints (see Chapter 4). However, other models more appropriate to orbit modelling can be preferable for some applications.

2.3.3 Spline Model

Cubic spline models set constraints on the first and second derivatives of a polynomial, allowing continuous interpolation using only neighbourhood points.

An orbit model can be computed efficiently once using a set of state vectors, and used thereafter for computationally "cheap" interpolation. Given a dense set of state vectors, simulation of the global movements of a satellite over multiple orbits is well suited to this computation-

ally efficient model. However, improvement of the orbit model using ground control points and linear or non-linear least squares is not directly supported by splines: a polynomial model is used in these cases.

2.4 Orbital Baseline

The relationship between two satellite orbits (baseline) may also be described in a variety of ways. The following subsections describe baseline notation systems, and methods for the calculation of the baseline.

2.4.1 Baseline Components

TCN: The normal (vertical), cross-track (horizontal), and tangential components of the baseline are oriented along the axes defined by the *TCN* coordinate system (see Section 2.2.3). Such notation can be useful in reducing the three-dimensional nature of the baseline to two dimensions (with the T-component set to zero) for modelling purposes.

Line-of-Sight Component: To calculate a basis vector in the line-of-sight (slant range) direction, one can approximate the Earth's radius r_E at the scene centre, and then use the cosine law to calculate the incidence angle θ:

$$\theta = \operatorname{acos}\left(\frac{|\underset{\sim}{S}_1|^2 + |\underset{\sim}{R}_1|^2 - r_E^{\,2}}{2|\underset{\sim}{S}_1||\underset{\sim}{R}_1|}\right) \qquad \textbf{(2-7)}$$

Noting that $R_1 = |\underset{\sim}{R}_1|$, the slant-range basis vector $\underset{\sim}{R}_1/|\underset{\sim}{R}_1|$ is then:

$$\underset{\sim}{R}_1/|\underset{\sim}{R}_1| = \cos(\theta)\hat{\underset{\sim}{n}} + \sin(\theta)\hat{\underset{\sim}{c}} \qquad \textbf{(2-8)}$$

and the component of the baseline in that direction is

$$B_{||} = \underset{\sim}{B} \cdot \underset{\sim}{R}_1/|\underset{\sim}{R}_1| \qquad \textbf{(2-9)}$$

Normal to Line-of-Sight Component: The normal to line-of-sight basis vector $\underset{\sim}{\chi}$ is then calculated via:

$$\underset{\sim}{\chi} = -\sin(\theta)\hat{\underset{\sim}{n}} + \cos(\theta)\hat{\underset{\sim}{c}} \qquad \textbf{(2-10)}$$

and the component of the baseline in that direction is

$$B_{\perp} = \underset{\sim}{B} \cdot \underset{\sim}{\chi} \qquad \textbf{(2-11)}$$

As the incidence angle θ varies across the swath (and also with terrain), a rough mean value is used here. This baseline description is therefore only useful for modelling purposes (e.g. sensitivity analyses).

2.4.2 Baseline Calculation

Calculation of the baseline between two defined orbits requires knowledge of the relationship between points in the second orbit and those in the first. A number of methods for its estimation have been developed over the years. Five are described below.

Fringe Frequency: Early studies of SAR interferometry did not have high quality orbital state vector data available. They used the fringe frequency of

the interferogram in a flat area to estimate the normal to line-of-sight component of the baseline [62].

$$B_{\perp} = \frac{\lambda}{4\pi} \cdot \frac{R_1 \cdot \tan\theta \cdot \Delta\phi}{\Delta R} \qquad \textbf{(2-12)}$$

where $\Delta\phi$ is the change in phase, and ΔR the change in slant range distance. Image-to-image co-registration parameters (range offset and trend) may also be used to make such an estimate of the baseline.

The fringe-frequency method makes a far-field approximation, and is therefore only useful for extremely small range swaths.

Orbit-to-Orbit - Closest Approach: One can approach baseline estimation as an optimization problem; many techniques exist for solving such problems. The numerical calculation technique known as Brent's method provides a computationally efficient method for solving minimization problems [67].

A simple method of determining the correspondence between two orbits is to calculate the point of closest approach between them, assuming that it is then when the spacecraft imaged the same point again. The objective function F is the magnitude of the hypothetical baseline $\underset{\sim}{B}_H$:

$$F = \left|\underset{\sim}{B}_H\right| \qquad \textbf{(2-13)}$$

The minimum distance is defined by the minimum baseline length.

Orbit-to-Orbit - Zero Dot Product: This procedure is similar to the preceding "closest approach" method. The dot product between the hypothetical baseline vector relating the orbits and the tangential vector (from the local definition of *TCN*) in the first orbit. The objective function F is

$$F = \left|\underset{\sim}{B}_H \cdot \underset{\sim}{t}\right| \qquad \textbf{(2-14)}$$

instead of "minimum distance". As with the "closest approach" algorithm, The numerical technique known as "Brent's method" [67] is used to iterate to the minimum absolute dot product. By definition, the tangential component of the baseline is then zero.

Tiepoints + unwrapped phase: An iterative non-linear least squares fit can be used to adjust the baseline model and phase constant [84].

A model is constructed for the baseline $\underset{\sim}{B}_i$ with a constant component in the normal direction B_n, and a cross-track component varying linearly in azimuth,

$$B_{c,i} = B_c + \alpha_c t_i, \qquad \textbf{(2-15)}$$

where B_c is the cross-track component of the baseline at mid-scene, and α_c governs the linear variation in azimuth (caused by non parallel orbits), with t_i representing the orbit time at which the point was acquired. The baseline model is then

$$\underset{\sim}{B}_i = B_{n,i}\hat{\underset{\sim}{n}} + (B_c + \alpha_c t_i)\hat{\underset{\sim}{c}} \qquad \textbf{(2-16)}$$

Selected tiepoints are used to refine estimates for α_c , B_c, and B_n.

Image Simulation: All four previous baseline estimation methods provide baseline models useful in restricted contexts, from sensitivity analyses to height model generation within restricted areas.

However, to handle the general case, a 2D paradigm for baseline calculations is insufficient: the 3D nature of the baseline needs to be modelled [78].

Figure 2-3 shows the difference (exaggerated) between 2D and 3D baseline models. Two non-parallel orbits are shown at the left. A zero-Doppler SAR is assumed for simplicity, acquiring a swath perpendicular to the flight track.

The orbit-to-orbit "closest approach" and "zero tangential component" baseline estimation methods do not take into account variations of the baseline across the swath.

An image simulation algorithm (see Chapter 3) is able to correctly model the effect, producing a 3D baseline model.

Although two zero-Doppler acquisitions are shown in Figure 2-3, in practice, one could have both non-parallel orbits as well as variable Doppler centroids during the image acquisitions. One could even choose a specific Doppler centroid during the second acquisition to compensate for the expected azimuth convergence of the respective orbits. In that case, the Doppler spectra would exactly overlap, as with parallel orbits acquired at the same Doppler centroid, and there would be no 3D baseline.

2.4.3 Critical Baseline

SAR interferometry is only possible when the ground reflectivity spectra acquired by the two (or more) antennae overlap. When the perpendicular component of the baseline increases beyond a limit known as the "critical baseline", no phase information is preserved, coherence is lost, and interferometry is not possible. For the ESA ERS satellites, the *critical baseline* $B_{\perp,cr}$ is calculated as:

$$B_{\perp,cr} = \frac{\lambda R \tan\theta}{2R_r} \approx 1.1km\,, \qquad \textbf{(2-17)}$$

where R_r is the slant range resolution (9.64 m), λ is the wavelength (5.6 cm), R is the slant range (850 km), and θ is the nominal incidence angle (~23°). See Section 3.3.1 for a detailed development of this limit.

The critical baseline can be significantly reduced by surface slopes that influence the local incidence angle.

2.4.4 Parallelism of ERS Orbits

Given the effect of non-parallel orbits seen in Figure 2-3, one might ask: how large is this effect for the ERS tandem geometry?

To answer this question, the most highly accurate global ERS orbit product (precise orbit: ESA "PRC" product) was used to compute the baseline between

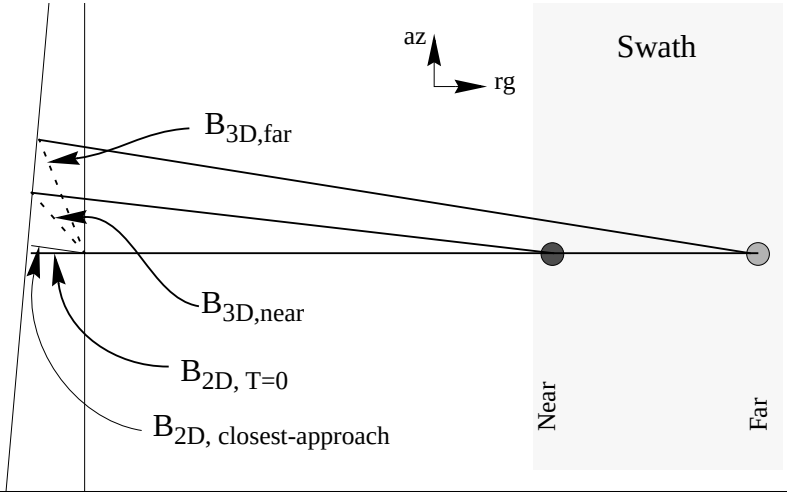


Figure 2-3: Baseline Geometries: 2D and 3D for non-parallel orbits

the ERS-1 and ERS-2 over the day October 22, 1995. The ground tracks of the orbits over the 24 hour period are shown in Figure 2-4. The latitude and longitude values displayed are geocentric, and were computed based on the WGS84 ellipsoid. The satellites made no orbital manoeuvres during the day.

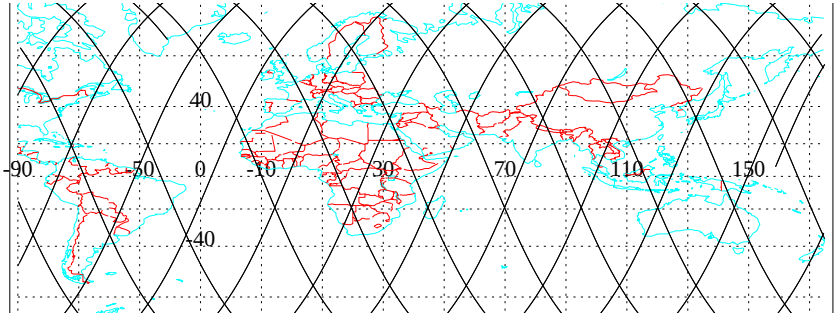


Figure 2-4: ERS-1 ground track over October 22-23, 1995

The computed baseline components associated with the ground track are shown in Figure 2-5. Note the variation in the vertical (normal) and horizontal (cross-track) components. As expected from experience with ERS interferograms, the normal component varies slowly, while the cross-track component changes significantly as the satellites traverse their sun-synchronous near-polar orbits. Orbital convergence causes the cross-track (horizontal) baseline component to be largest near the equator, and smallest near the poles.

The instantaneous angle between the two orbits was also calculated over the 24 hour period, and is plotted in Figure 2-6. Note how the angle between the orbits remains relatively constant, on the order of a thousandth of one degree.

The extent of azimuth convergence can be higher with other interferometer configurations (e.g. SIR-B) [15].

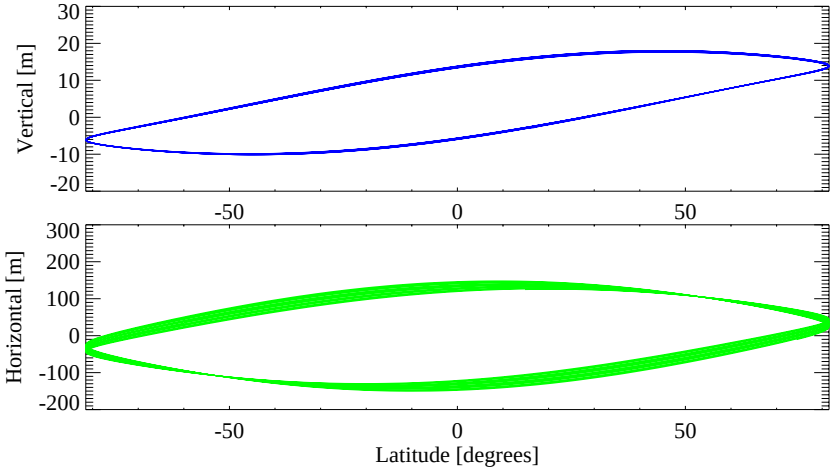


Figure 2-5: ERS tandem baseline components: variation over October 22-23, 1995

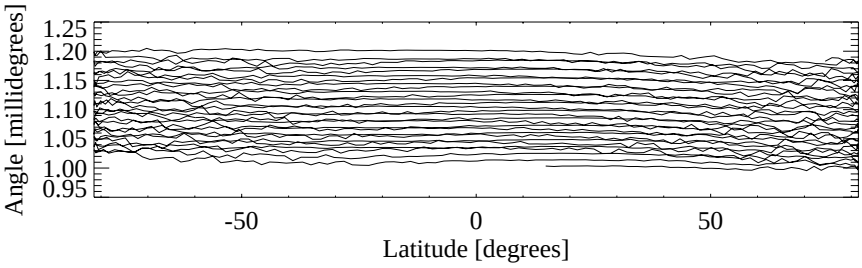


Figure 2-6: ERS tandem baseline angle: variation over October 22-23, 1995 -- minimal deviation from parallel orbits

Chapter 3

Interferogram Processing

3.1 Introduction

This chapter describes the processes leading up to interferogram formation, together with associated flattening and coherence calculation steps.

Figure 3-1 provides an overview of the interferogram processing steps. Candidate pair(s) of acquisitions are first selected, and the SAR signal data is processed to the single-look-complex (SLC) stage. Given that no co-registration is performed at the SAR processing stage (a valid assumption for most general purpose SAR processors), the SAR images must be registered into a common geometry before interferogram formation. Progressively finer registration models are calculated, and the designated "slave" image is resampled into the geometry of the "master".

The interferogram may now be calculated unabated, followed by removal of a modelled "expected" phase difference using a reference elevation model available for the area and/or a simple Earth ellipsoid model.

A coherence image can be calculated at this point and used to highlight areas subject to temporal decorrelation between acquisitions. Areas with a poor phase signal are improved using adaptive filtering, and the 2π ambiguity in the phase difference measurement is "unwrapped".

This chapter discusses each of the above steps in more detail, covering all steps shown in Figure 3-1 up to and including phase unwrapping.

3.2 Registration

3.2.1 Orbital Coarse Estimate

Azimuth Offset: Figure 3-2 shows the time reference frames for the two scenes.

The azimuth offset is calculated as follows. The baseline is calculated for the very first range line from the first scene (t_{1a}). The time in the second orbit's reference frame t_2 is retrieved from the baseline calculation and compared to the time of the second scene's first range line (both in the second scene's time reference frame). The azimuth pixel offset c_a between the two scenes in pixels is then:

$$c_a = PRF \times (t_2 - t_{2a}) \qquad \textbf{(3-1)}$$

where PRF is the pulse repetition frequency of the radar (the data are assumed here to be single look).

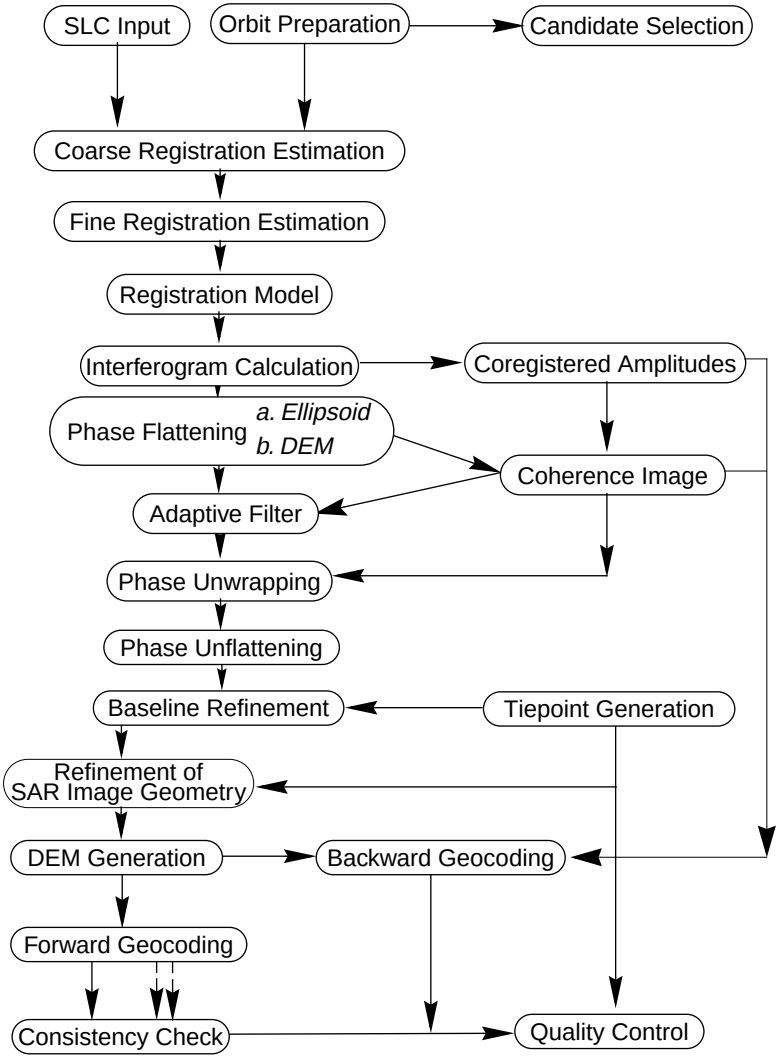


Figure 3-1: InSAR processing steps

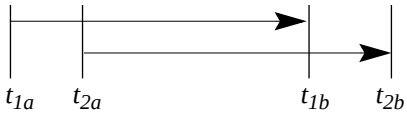


Figure 3-2: Azimuth coarse offset geometry

Range Offset: The range offset between the two images is calculated as follows. First, the position of the satellite at mid scene during the first orbit $\underset{\sim}{S}_1$ (x,y,z) is computed. The baseline at that point is calculated, and used to determine the position of the second satellite $\underset{\sim}{S}_2$ (x,y,z) at the mid point of the first scene:

$$\underset{\sim}{S}_2 = \underset{\sim}{S}_1 + \underset{\sim}{B} \qquad \textbf{(3-2)}$$

Assuming a simplified spherical Earth with radius r_E (for rough estimation purposes only), at that latitude, the Earth-centred Cartesian coordinates of the scene centre may be calculated from the latitude ζ and longitude Π:

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = r_E \times \begin{bmatrix} \cos\zeta \times \cos\Pi \\ \cos\zeta \times \sin\Pi \\ \sin\zeta \end{bmatrix} \qquad \textbf{(3-3)}$$

where ζ is the scene centre latitude and Π is the scene centre longitude. The slant range "look vectors" $\underset{\sim}{R}_1$ and $\underset{\sim}{R}_2$ for both orbits are calculated by subtracting the respective spacecraft positions from the target position vector $\underset{\sim}{T}$ (x,y,z):

$$\begin{aligned} \underset{\sim}{R}_1 &= \underset{\sim}{T} - \underset{\sim}{S}_1 \\ \underset{\sim}{R}_2 &= \underset{\sim}{T} - \underset{\sim}{S}_2 \end{aligned} \qquad \textbf{(3-4)}$$

Finally, the coarse range pixel offset c_r is calculated via:

$$c_r = \frac{|\underset{\sim}{R}_2| - |\underset{\sim}{R}_1|}{R_s} \qquad \textbf{(3-5)}$$

where R_s is the slant range pixel spacing.

3.2.2 Subpixel Accuracy

Correlation of the SAR image amplitude can be used to refine the registration model calculated from the orbit data.

This section describes how the initial offset estimates are refined using interferometric fringe visibility to achieve registration with subpixel accuracy.

A set of image "chips" distributed across the image are selected, and the initial offsets calculated previously are used as the departure point for the calculation of a set of 32×32 pixel interferograms with half pixel offsets. Since the interferometric fringes are most visible at the optimal registration, a fringe visibility criterion [15] is used to select the correct registration. Both the range and azimuth spectra are zero-padded to oversample the image chips by a factor of two prior to formation of miniature interferograms.

The exact offsets are interpolated in the neighbourhood of the maximum of the visibility function.

Figure 3-3 shows the trends in the offset estimates across the Bonn scene pair acquired on the 14th and 17th of March, 1992. Both the mean range and azimuth offsets have been subtracted from each vector to highlight the trend across the image. The lengths of the vectors are highly exaggerated.

Points along the grid with no offset vector shown had a fringe signal-to-noise ratio [45] (“fringe SNR”) below the set threshold, and were not used to model the offset trend. In those cases, no offset choice stood out high enough above the rest to warrant a claim that fringes had been found. Figure 3-4 shows the offset trend for the same scene pair, this time with a higher “fringe SNR” threshold set. Note the dominant influence of the range position on the range and azimuth offsets.

The lower threshold results in more data points, but there is also a larger variance. The higher threshold reduces the number of points, but their quality is higher and more uniform.

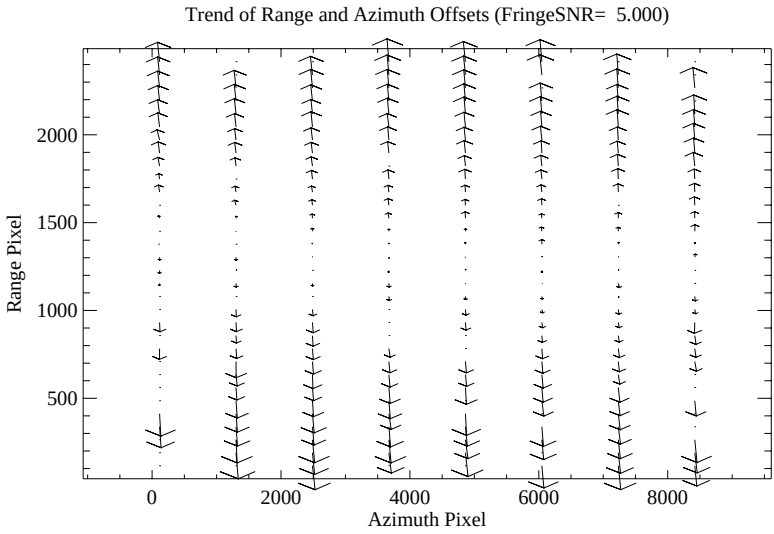


Figure 3-3: Illustration of image-to-image chip offset estimates: SNR=5

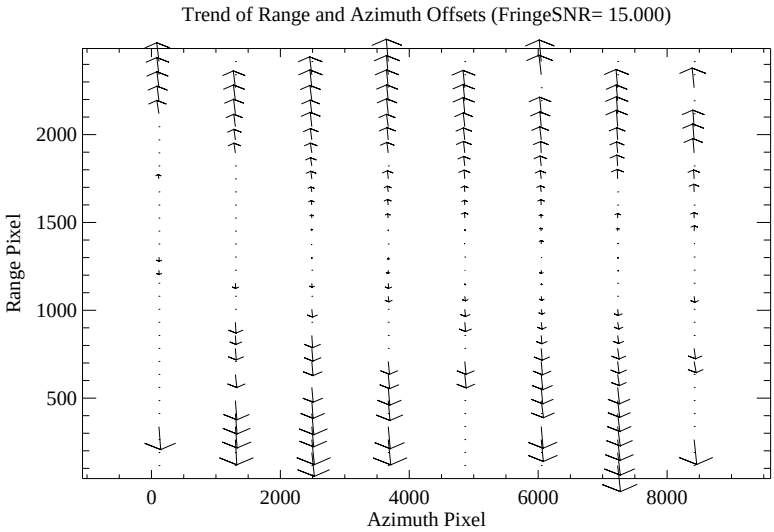


Figure 3-4: Illustration of image-to-image chip offset estimates: SNR=15

Once the offset estimates are available, a least squares fit is used to provide a model across the image of the offset dependency on range and azimuth location within the image. This model is used to resample the “slave” image into the geometry of the “master”.

The “slave” image can be transformed into the geometry of the first using a variety of resampling techniques. These include bilinear, biquadratic, cubic-spline, frequency domain multiplication with a complex exponential (equivalent to a rigid shift), and convolution with a finite impulse response “FIR” resampling kernel.

The implementation of a resampling technique for the single look complex “SLC” imagery requires that attention be paid as well to the effect of the resampling kernel on the phase [60]. Since the azimuth spectra are generally shifted by f_{Dc}, a repositioning back to zero central frequency (through multiplication with a complex exponential in the spatial domain) makes it possible to use standard resampling kernels.

3.3 Interferogram

3.3.1 Calculation

This section describes the calculation of the interferogram itself.

Given co-registered single look complex values $\underset{\sim}{c}_1$ and $\underset{\sim}{c}_2$, the interferogram $\underset{\sim}{G}$ is calculated as:

$$\underset{\sim}{G} = \sum_{i}^{N} \underset{\sim}{c}_1(i)^* \underset{\sim}{c}_2(i) . \qquad \textbf{(3-6)}$$

Multiple looks (N) are taken at this point to reduce data volume. The interferometric phase ϕ_G is expressed simply as:

$$\phi_G = \operatorname{atan}\left(\frac{\Im\{\underset{\sim}{G}\}}{\Re\{\underset{\sim}{G}\}}\right). \qquad \textbf{(3-7)}$$

Note that ϕ_G can only be known modulo 2π from the above equation. Further processing steps, known as phase unwrapping (see Section 3.6) are necessary to resolve this inherent ambiguity.

Oversampling: Multiplication of the two complex SLC values in the spatial domain (Equation 3-6) is equivalent to a convolution of their spectra in the frequency domain. Figure 3-5 illustrates how the convolution of such roughly rectangular spectra $\underset{\sim}{C}_1(f)$ and $\underset{\sim}{C}_2(f)$ results in a spectrum for the interferogram,

$$\underset{\sim}{G}(f) = \underset{\sim}{C}_1(f) \otimes \underset{\sim}{C}_2(f) , \qquad \textbf{(3-8)}$$

where $\underset{\sim}{G}(f) = FT\{\underset{\sim}{G}\}$ that is shaped like a triangle, and that has twice the bandwidth of each input SLC image.

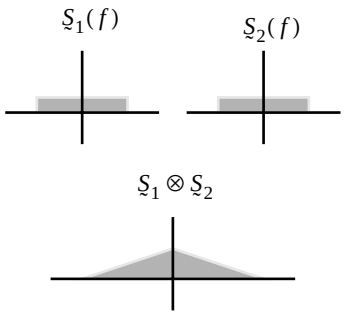


Figure 3-5: Convolution of rectangular spectra: interferogram has triangular spectrum with double bandwidth

The increased bandwidth requires a higher sampling frequency than is necessary for a single SLC image. If the sampling frequency is not increased beyond the interferogram's Nyquist sampling frequency (Shannon sampling theorem), frequencies will be aliased on top of one another, resulting in a "noisier" interferogram, with lower overall coherence. Although Figure 3-5 illustrates the relation for 1-D spectra, the principle applies to both the range and azimuth dimensions of the interferogram.

Example: Figure 3-21(c) shows the interferogram phase values for an ERS-1 image pair acquired in March 1992. Interferogram phase values (between 0

and 2π) are shown in grey scale. The orbits were descending passes; near range is at the right, far range on the left. The associated amplitude images for the two acquisitions are shown in Figure 3-21(a) and (b) respectively.

Fringe Frequency: Adopting the simplified 2D geometry [44] with parallel orbits and a flat Earth shown in Figure 1-10, one can calculate the expected rate of change of the interferometric phase by taking the derivative of the phase expression ϕ (Equation 1-17) with respect to slant range distance R

$$\frac{\partial\phi}{\partial R} = \frac{4\pi}{\lambda}\cdot\left(B_x\frac{\partial}{\partial R}\sqrt{1-\frac{\Gamma^2}{R^2}} - B_y\frac{\partial}{\partial R}\frac{\Gamma}{R}\right), \quad \textbf{(3-9)}$$

where $\Gamma = H - h$. Equation 3-9 reduces after some simplification to:

$$\frac{\partial\phi}{\partial R} = \frac{4\pi}{\lambda}\cdot\frac{\Gamma}{R^2}\cdot\left(B_x\frac{\Gamma}{\sqrt{R^2-\Gamma^2}} + B_y\right). \quad \textbf{(3-10)}$$

Through trigonometry one has

$$\frac{\partial\phi}{\partial R} = \frac{4\pi}{\lambda R}\cdot\cos\theta\cdot\left(\frac{B_x}{\tan\theta} + B_y\right), \quad \textbf{(3-11)}$$

which can be rewritten as

$$\frac{\partial\phi}{\partial R} = \frac{4\pi}{\lambda R}\cdot\frac{\cos\theta}{\sin\theta}(B_x\cos\theta + B_y\sin\theta) \quad \textbf{(3-12)}$$

reducing to

$$\frac{\partial\phi}{\partial R} = \frac{4\pi}{\lambda}\cdot\frac{B_\perp}{R\tan\theta}, \quad \textbf{(3-13)}$$

giving a compact expression (similar to Equation 1-20) expressing the sensitivity of the interferometric phase to changes in the range distance.

One can use Equation 3-13 to estimate the local relation between changes in the range distance ΔR and the interferometric phase increment $\Delta\phi$:

$$\Delta\phi \approx \frac{\partial\phi}{\partial R}\cdot\Delta R. \quad \textbf{(3-14)}$$

A full fringe in the original interferogram is therefore spread over a slant range distance

$$\Delta R_{2\pi} = \frac{\lambda}{2}\cdot\frac{R\tan\theta}{B_\perp} \quad \textbf{(3-15)}$$

where $\Delta R_{2\pi}$ is called the slant range fringe distance. The above relation describes the fringe frequency in the range direction (modelled using a simplified flat Earth).

The slant range fringe distance is plotted in Figure 3-6 versus the incidence angle for four baseline lengths. Note that ERS nominal incidence angles range from 20° to 26°.

Rescaling the range fringe frequency from space (Equation 3-15) to time, the fringe frequency f_{fringe} becomes

$$f_{fringe} = \frac{c/2}{\Delta R_{2\pi}} = f_0\cdot\frac{B_\perp}{R\tan\theta}. \quad \textbf{(3-16)}$$

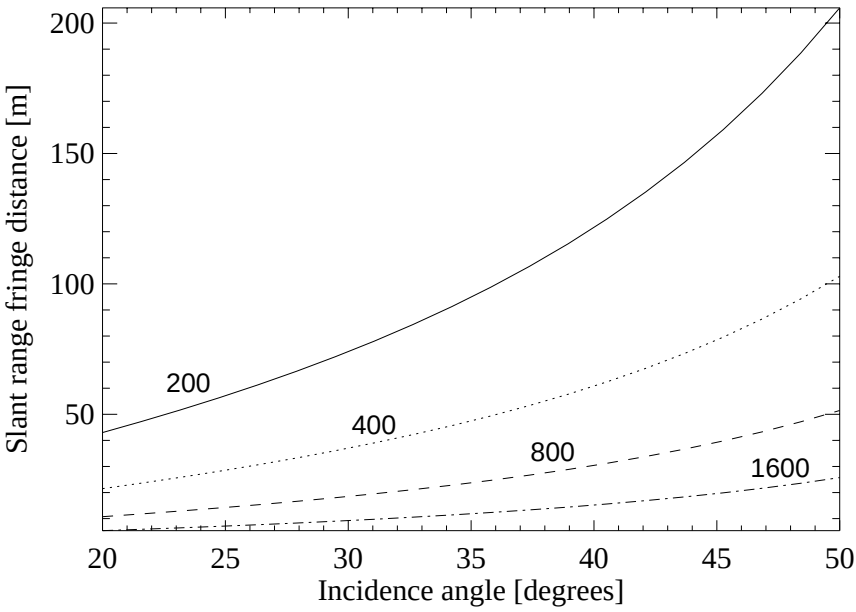


Figure 3-6: Plot of slant range fringe distance vs. incidence angle for $B_\perp = 200m$ (solid), $B_\perp = 400m$ (dot), $B_\perp = 800m$ (dash), $B_\perp = 1600m$ (dash-dot)

Shifted Range Spectra: The two (or more) antenna locations used for image acquisition induce not only the interferometric phase difference (outlined in Section 1.3.2) but also shift the portions of the ground reflectivity spectrum that are acquired from each vantage point [17].

Expressed another way, the same "harmonic ground structure transforms into different echo frequency components" [2]. The extent of the spectral shift in the slant range plane can be approximated [17] as:

$$\Delta f_r \approx f_0 \cdot \frac{\Delta\theta}{\tan(\theta - \alpha)} . \qquad \textbf{(3-17)}$$

where f_0 is the radar carrier frequency, and α is the local ground slope.

Inspection of Figure 1-8 shows that the far-field approximation allows one to approximate $\Delta\theta$ by:

$$\Delta\theta \approx \frac{B_\perp}{R} . \qquad \textbf{(3-18)}$$

Combining the above two equations, one has:

$$\Delta f_r \approx f_0 \cdot \frac{B_\perp}{R \cdot \tan(\theta - \alpha)}, \qquad \textbf{(3-19)}$$

where α is the ground slope, relating the perpendicular component of the baseline to the shift in the slant-range spectrum Δf_r.

This shift in the range spectrum is equal to the range fringe-frequency in the unflattened "raw" interferogram (Equation 3-16).

The range frequency shift is plotted versus incidence angle for different baselines in Figure 3-7. Note that the shift becomes smaller with decreasing baseline and increasing incidence angle.

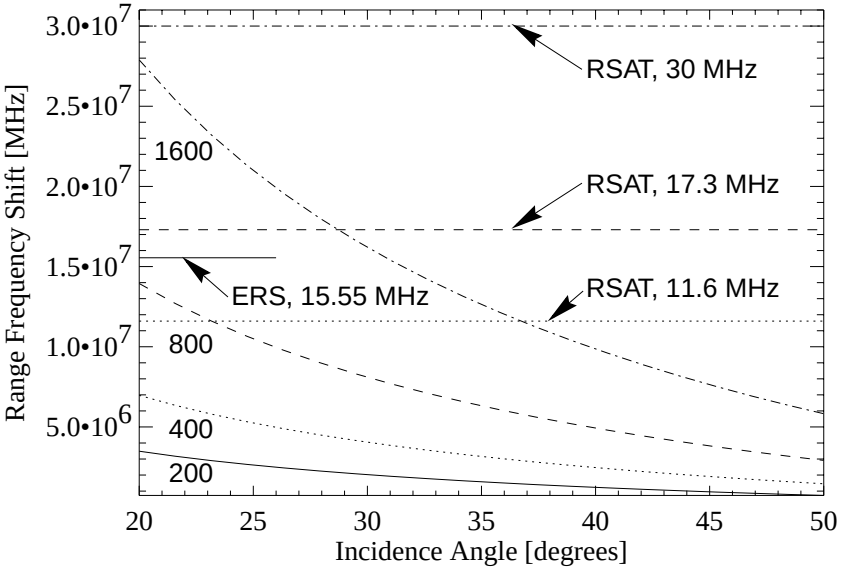


Figure 3-7: Plot of range frequency shift vs. incidence angle for $B_\perp = 200m$ (solid), $B_\perp = 400m$ (dot), $B_\perp = 800m$ (dash), $B_\perp = 1600m$ (dash-dot)

Critical Baseline: At the point where the shift exceeds the system range bandwidth itself (displayed in Figure 3-7 for ERS as well as RADARSAT's three chirp modes), the "common band" of the ground reflectivity spectra observed during both acquisitions reduces to zero.

The critical baseline occurs when the spectrum is shifted its full length B_r, i.e. $|\Delta f_r| = B_r$. At that point there is no longer overlap in the spectra of the two acquisitions, and all coherence is lost.

Relating Equations (1-6) and (3-19), one has:

$$B_{\perp,cr} \approx \frac{\lambda R \tan\theta}{2R_r}, \qquad \textbf{(3-20)}$$

resulting in $B_{\perp,cr} \approx 1.1km$ for the ERS satellites, with R_r being the slant range resolution (10.2 m), λ the wavelength (5.6 cm), R the slant range (850 km), and θ the nominal incidence angle (~23°). Note that (as illustrated in Figure 3-7) the extent of the frequency shift varies with incidence angle, and the critical baseline therefore varies across the swath. In addition, surface slopes of α=0° have been assumed in the above; the critical baseline can be significantly reduced by surface slopes that influence the local incidence angle.

Shifted Azimuth Spectra: Just as the range spectra become shifted due to variable viewing angles of the terrain, differing squint angles can cause shifted azimuth spectra.

Parameter	Symbol	ERS Value
Range Bandwidth	B_r	15.55 MHz
Range Sampling Rate	f_{Rs}	18.96 MHz
Azimuth Processed Bandwidth	B_a	≈1410 Hz
Azimuth Sampling Rate (Pulse Repetition Frequency)	f_{As} or PRF	≈1680 Hz

Table 3-1: Range and azimuth bandwidths (taken from ESA image product headers)

Processing of the ERS-1/2 tandem data (see Table 6-2) revealed differing Doppler centroids for the ERS-1 and ERS-2 satellites. The mean difference Δf_a was estimated to be approximately 280 Hz.

The difference was observed between ERS-1/2 data acquired over Bern, Switzerland in June-August 1995, as well as Egyptian (November 1995) and Moroccan (June 1995) tandem pairs. Values read within the image header annotation generally agreed with centroids calculated through spectral summation. Figure 3-8 shows the Doppler spectra for the Bern data. The ERS-1 data behave consistently (note overlap),

and the ERS-2 data are also consistent (with each other). However, the ERS-1 and ERS-2 centroid estimates clearly fall into two separate populations.

Doppler spectra from the ERS-1 commissioning phase (Bern, November 1991), as well as from the 3-day repeat ERS-1 first ice phase (Bonn, Germany - March 14, 17, 20, 29, 1992) were also reviewed. In these cases, the ERS-1 spectra never exhibited any such shift.

As a consequence of this ERS-1 / ERS-2 shift, an extra azimuth filter applied during interferogram generation is required to fully capture the scene's potential coherence (at the cost of poorer azimuth resolution).

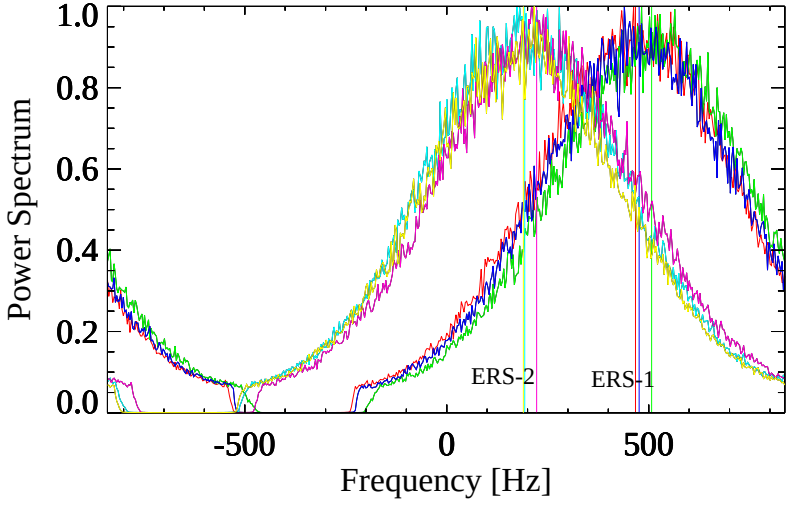


Figure 3-8: ERS-1/2 Doppler centroid estimates - Bern, Switzerland - June - August, 1995

3.3.2 Associated SAR Image Intensities

The N-look single-scene intensities I_1 and I_2,

$$I_1 = \sum_i^N \varsigma_1(i)\varsigma_1(i)^* \text{ and} \tag{3-21}$$

$$I_2 = \sum_i^N \varsigma_2(i)\varsigma_2(i)^*, \tag{3-22}$$

are calculated from the single look complex values $\underset{\sim}{c}_1$ and $\underset{\sim}{c}_2$, and stored for future reference.

The interferogram can be normalized by the two single-scene intensities to counteract the influence of bright point targets that might otherwise dominate the calculation of the "fringe SNR". The normalized interferogram $\hat{\underset{\sim}{G}}$

$$\hat{\underset{\sim}{G}} = \frac{\underset{\sim}{G}}{I_1 \cdot I_2} \qquad \textbf{(3-23)}$$

has magnitude equal to the *N*-look coherence. The file size is reduced at this stage, as (for ERS) *N*=5 azimuth looks are taken. An example of such coregistered intensity images is shown in Figure 3-21(a-b).

3.3.3 Interferogram Flattening

In preparation for the phase-unwrapping step to come, the expected phase (calculated using a system model) is removed, producing a "flattened" interferogram $\hat{\underset{\sim}{F}}$ that is easier to unwrap.

One early method (still commonly used) simply deramps the interferogram, removing the fringe frequency as calculated using Equation 3-13. The interferogram itself can be used to estimate the frequency to be removed (neglecting terrain influences and Earth curvature).

Later methods [84] used a simple spherical Earth and 2-D interferometric baseline model to estimate the expected Earth phase. Earth curvature was only accounted for imperfectly, and the 3-D nature of the interferometric baseline (given non-parallel orbits) was not modelled.

For the results shown here, a "synthetic interferogram" is generated using orbit data and an ellipsoid model of the Earth's surface (possibly with available elevation model data covering the area of interest). This technique models Earth curvature realistically, and captures the 3-D nature of the interferometric baseline when repeat-pass orbits are non-parallel.

Given the height and position of points within the scene, the positions of antennae A_1 and A_2 at $\underset{\sim}{S}_1$ and $\underset{\sim}{S}_2$ may be calculated through geolocation. Given the positions of the antennae, the *difference in their distances* δ_r to a common target in the scene is also known through $\delta_r = R_2 - R_1$. Knowledge of δ_r and use of Equation 1-12 provides an estimate of the interferometric phase difference.

The following subsections describe how δ_r is calculated given (a) a reference DEM covering the area, or (b) an ellipsoidal Earth model.

Digital Elevation Model: This subsection describes the use of a reference digital elevation model (provided independently of the InSAR data at hand) to flatten an interferogram. Establishing a relation between the elevation model in map geometry and the SAR data (in radar slant range geometry) requires a geolocation step. Therefore,

some of the discussion here preludes more detailed coverage in the chapter on geocoding (see Chapter 4).

The DEM is presented in a geocoded map reference geometry. The discussion below assumes a map geometry (Northing, Easting), although geographic coordinates (latitude and longitude) could be substituted.

For each DEM location $\underset{\sim}{P}_d$ (in the same Cartesian coordinates as the orbital state vectors), a backward geolocation process is carried out for both the first and second acquisitions.

Looping through all DEM locations, such geolocation allows one to calculate the spacecraft position $\underset{\sim}{S}$ for each DEM location. Knowledge of both the DEM location and the associated spacecraft position for both first and second images enables calculation of the range from the antenna to the DEM position:

$$R_1 = |\underset{\sim}{R}_1| = |\underset{\sim}{S}_1 - \underset{\sim}{P}_d| , \qquad \textbf{(3-24)}$$

$$R_2 = |\underset{\sim}{R}_2| = |\underset{\sim}{S}_2 - \underset{\sim}{P}_d| . \qquad \textbf{(3-25)}$$

Once both R_1 and R_2 are known, calculation of the expected phase difference ϕ_d at location $\underset{\sim}{P}$ is a simple matter.

$$\phi_d(N,E) = \frac{4\pi}{\lambda}[R_2(N,E) - R_1(N,E)] \qquad \textbf{(3-26)}$$

Geolocation allows one to build up expected phase differences for all DEM locations, stored in $\phi_d(N,E)$.

The phase differences are however required in the radar slant-range geometry. Since knowledge of the spacecraft location $\underset{\sim}{S}$ provides the local range and azimuth coordinates (r, a) the expected phase difference can be assigned to a 2-D slant range geometry:

$$\phi_d(r(N,E),a(N,E)) = \phi_d(N,E) \qquad \textbf{(3-27)}$$

Since the values of r and a do no generally correspond to the raster grid locations in the slant range imagery, a regridding step is necessary. Triangulation or the nearest-neighbour method are appropriate depending upon the relation between the pixel spacings used in slant range and the reference DEM. In the nearest-neighbour method, after looping through the whole DEM file, slant range image locations with no value at the end of this process are assigned an expected phase difference based on interpolation of neighbouring values.

One now has a slant-range geometry synthetic interferogram $\phi_d(r, a)$, an example of which is shown in Figure 3-22(c).

Ellipsoid Model: In principle, an ellipsoid could be modelled using the method outlined above, by simply setting all heights to a constant value. In practice however, the interest in having as fast a processor as possible demands that a more computationally efficient method for the ellipsoid flattening case be employed.

Similarly to the DEM method described above, the expected phase

from the ellipsoid ϕ_e can be calculated from the difference in range distances to a point on the ellipsoid:

$$\phi_e = \frac{4\pi}{\lambda} \cdot (|\underset{\sim}{S}_2 - \underset{\sim}{P}_e| - |\underset{\sim}{S}_1 - \underset{\sim}{P}_e|) \qquad \textbf{(3-28)}$$

Unlike the general DEM case described above, for a simpler ellipsoid model, the behaviour of the phase across range and azimuth can be described by a simple polynomial function:

$$\phi_e = c_0 + c_1 \cdot r + c_2 \cdot a + c_3 \cdot r \cdot a + c_4 \cdot r^2 + c_5 \cdot a^2 \qquad \textbf{(3-29)}$$

where r and a are indices representing range and azimuth respectively.

Construction of such a polynomial requires at least six sets of $\{\phi_e, r, a\}$ and $\{\underset{\sim}{S}_1, \underset{\sim}{S}_2, \underset{\sim}{P}\}$, with more than six preferred to form an over-determined system of equations. Given a large enough set of $\{\phi_e, r, a\}$ the coefficients are easily retrieved using a least squares fit. Calculation of even 50 such sets of values can be done significantly faster than the DEM method outlined on the previous page.

Speedy construction of the polynomial best fitting the scene area under study requires a slightly modified geolocation step relating the ellipsoid (in Earth-centred Cartesian space) to the range and azimuth coordinates of the SAR imagery. Given certain range and azimuth image coordinates $\{\underset{\sim}{S}_1(i), r_1(i), a_1(i)\}$, one uses a forward-geocoding method (see Section 4.5) to calculate the corresponding point on the ellipsoid $\underset{\sim}{P}_e(i)$. Using the solved-for point $\underset{\sim}{P}_e(i)$ as a reference, a backward-geocoding approach is then used to solve for the corresponding positions in the slant range geometry of the second acquisition $\{\underset{\sim}{S}_2(i), r_2(i), a_2(i)\}$. The method is illustrated in Figure 3-9.

This method guarantees that all points are within the image area under study. The set of points $\{\underset{\sim}{S}_1(i), r_1(i), a_1(i)\}$ is best distributed across the length and breadth of the scene, to ensure a representative polynomial model for $\phi_e(r, a)$.

Since the polynomial in Equation 3-29 encapsulates the same information as would a synthetic interferogram based on an ellipsoid, calculation of an explicit data file containing a synthetic interferogram is not required. The polynomial is used to internally calculate the expected local phase difference within the program performing the flattening.

Example: The result of such ellipsoid flattening is displayed Figure 3-10 for an interferometric pair acquired over Zürich, Switzerland in Nov. 1995. One notes a clear correspondence between the fringes visible and elevations. Lake Zürich, the Greifensee, and forested areas appear as areas with “noisy” highly variable phase values. The local mountains Üetliberg, Zürichberg, and Hönggerberg are all now distinct, as are

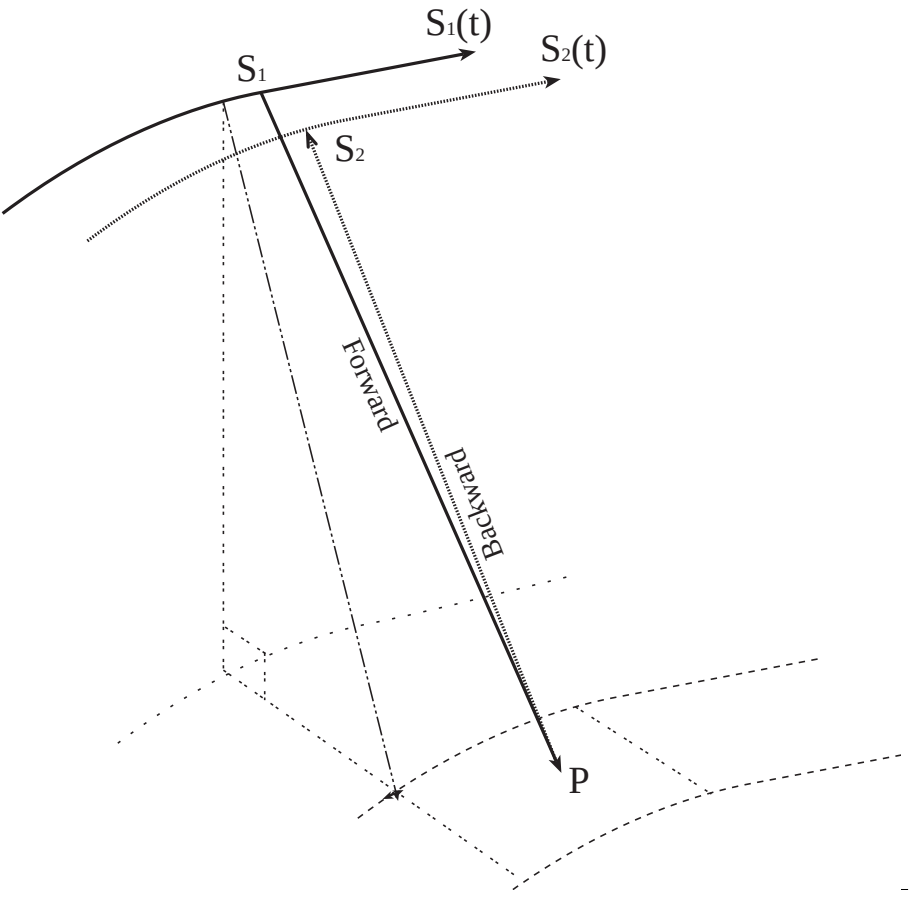


Figure 3-9: Ellipsoid flattening - Phase retrieval through forward and backward geocoding

the relatively flat Limmat and Glatt valleys.

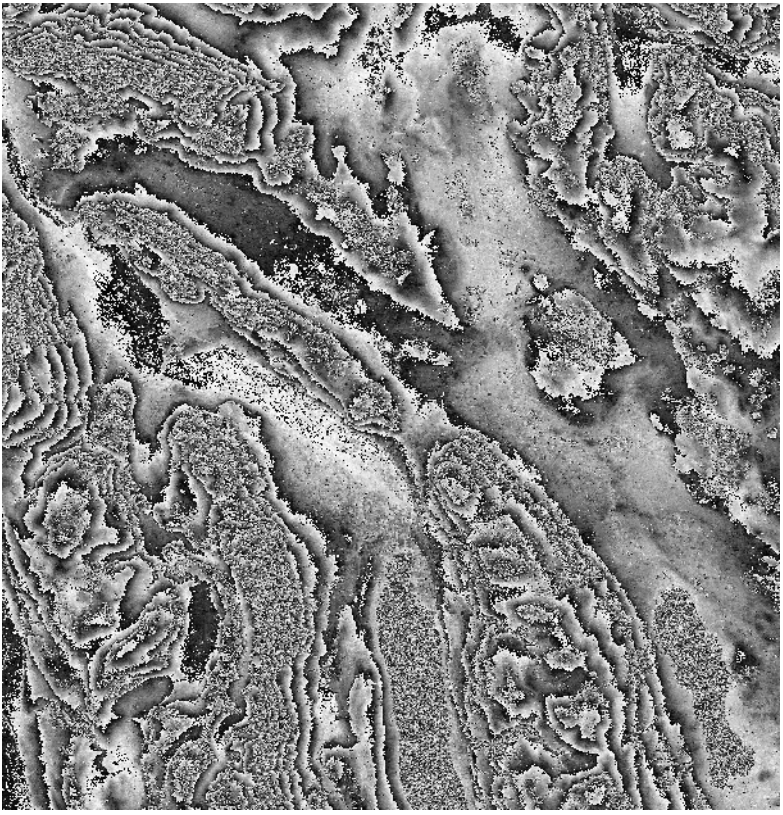


Figure 3-10: ERS-1 / ERS-2 Ellipsoid-flattened Interferogram - Zürich, Switzerland - Nov. 4-5, 1995 [© ESA / RSL]

Flattening: Given that the expected phase $\phi_m(r, a)$ has been calculated (either from the ellipsoid or DEM),

$$\phi_m = \begin{cases} \phi_d \\ \phi_e \end{cases} \quad \textbf{(3-30)}$$

the modelled fringes are subtracted from the raw (unflattened) interferogram

phase. Performed for each pixel i in the image, this transforms the *unflattened* interferogram $\underset{\sim}{G}$ into a *flattened* interferogram $\underset{\sim}{F}$:

$$\underset{\sim}{F}(r,a) = \underset{\sim}{G}(r,a) \cdot e^{-j\phi_m(r,a)} . \qquad \textbf{(3-31)}$$

Example: For the Bonn interferogram shown in Figure 3-21(c), the corresponding ellipsoid-generated synthetic interferogram is shown in Figure 3-22(a), with the ellipsoid-flattened phase displayed in Figure 3-22(b). The DEM-generated synthetic interferogram is shown in Figure 3-22(c), with the DEM-flattened phase shown in Figure 3-22(d).

3.4 Coherence Calculation

Given two co-registered N-look complex SAR images $\underset{\sim}{G}_1$ and $\underset{\sim}{G}_2$, one calculates the interferometric coherence γ as a ratio between coherent and incoherent vector summations:

$$\gamma = \frac{\left| \sum_{k=1}^{M} \underset{\sim}{F}_k \right|}{\sqrt{\sum_{k=1}^{M} \underset{\sim}{G}_{1,k}} \sqrt{\sum_{k=1}^{M} \underset{\sim}{G}_{2,k}}} . \qquad \textbf{(3-32)}$$

Using ERS data, one can set N=5 (all in azimuth), and M=9 (3 each in range and azimuth). To avoid a biased estimate of the coherence [88], special attention must be paid to the window size used. For the results shown here, unless otherwise indicated, the coherence estimate is formed using a windowwith a length of 15 single-look pixels (azimuth), and a width of 3 pixels (range), for a total of 45 “looks”.

Note that Equation 3-32 is equivalent to direct summation over an $M \times N$ window, since:

$$\sum_{i=1}^{MN} \underset{\sim}{g}_{1,i}^{*} \underset{\sim}{g}_{2,i} \quad , \qquad \textbf{(3-33)}$$

$$= \sum_{k=1}^{M} \left(\sum_{i=1}^{N} \underset{\sim}{g}_{1,i}^{*} \underset{\sim}{g}_{2,i} \right)$$

$$= \sum_{k=1}^{M} \underset{\sim}{G}_k$$

and:

$$\sum_{i=1}^{MN} \underset{\sim}{g}_{1,i} \underset{\sim}{g}_{1,i}^{*} \quad . \qquad \textbf{(3-34)}$$

$$= \sum_{k=1}^{M} \left(\sum_{i=1}^{N} \underset{\sim}{g}_{1,i} \underset{\sim}{g}_{1,i}^{*} \right)$$

$$= \sum_{k=1}^{M} \underset{\sim}{G}_{1,k}$$

The two step calculation results in a significant reduction in the computational cost of the coherence calculation. For the 15×3 case, the number of complex addition operations required per term is reduced by a factor of 3.2 (from 45 to 9+5=14). Given that the N-look interferogram and intensity files are

already calculated, the savings factor is increased to 5 (N in general).

Note that the coherence is calculated from the flattened interferogram $\hat{\underset{\sim}{F}}$, and not the original $\hat{\underset{\sim}{G}}$. This is absolutely necessary when working with large baselines to avoid biasing from the range phase trend, which otherwise causes severe underestimation of γ. Given that one already has a high resolution elevation model, biases introduced to the coherence estimation by local slopes can be removed by first compensating for the phase difference expected due to the topography. This is particularly important when large window sizes (M and N) are used, as otherwise terrain-induced loss of coherence can be mistaken for true temporal decorrelation [88].

Figure 3-11 shows the effect of such flattening on the coherence. ERS tandem coherence data from a scene surrounding Bern, Switzerland were used. A large window size (7×35) was chosen to exaggerate the effect. Note the improvement in coherence when slope-induced decorrelation is removed. InSAR-slopes can, however, over-estimate the coherence, when truly noisy pseudo-“slopes” are flattened even in the absence of local topography.

Example: Figure 3-12 shows a sample coherence image for the Zürich ERS tandem image pair acquired on Nov. 4 and 5, 1995. Lake Zürich, the Limmat river, and the Greifensee are seen to have

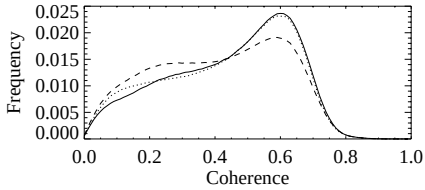


Figure 3-11: Coherence histograms for (a) ellipsoid flattened [dotted], (b) flattened with InSAR slopes [dashed], (c) flattened with reference DEM [solid]

low coherence (black), as do forested areas (e.g. eastern side of the Üetliberg). The rail yard approaches to Zürich’s main railway station exhibit high coherence due to strong double-bounce high-SNR scattering, while the urban and rural landscapes surrounding the city have characteristically mid to high coherence values.

Coherence images provide a new channel for thematic interpretation, providing information complementary to the SAR image magnitude.

3.4.1 Spectral-Shift Filter

The recoverable coherence is improved through the use of a spectral-shift filtering step [17].

The improvement is illustrated in Figure 3-13 using an image pair acquired by ERS-1 in March, 1992

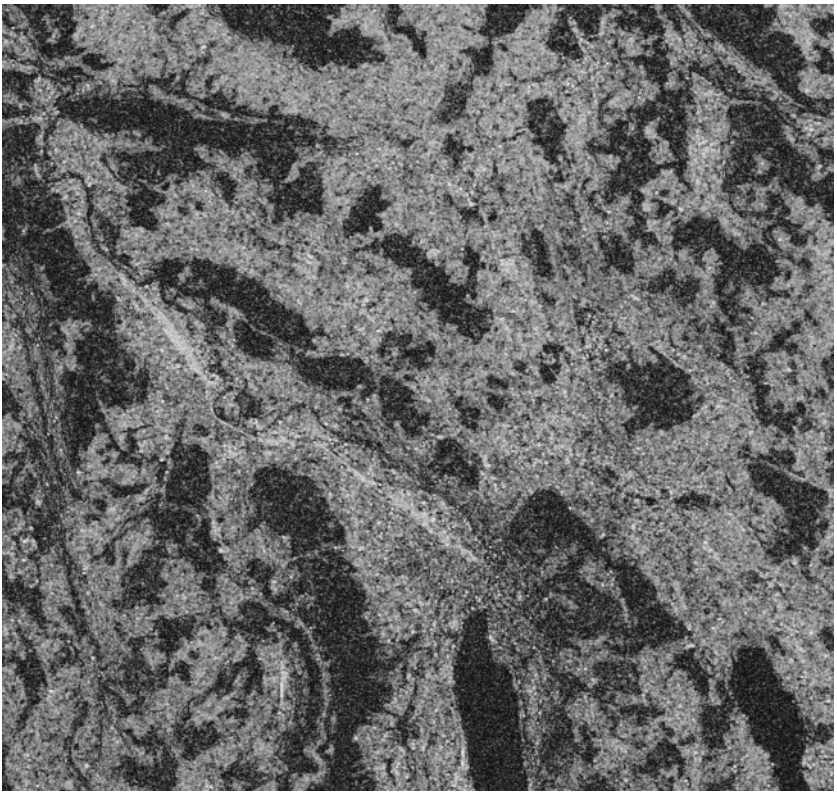


Figure 3-12: ERS-1 / ERS-2 Coherence Image - Zürich, Switzerland; Nov. 4-5, 1995; Black=0, White=1 [© ESA / RSL]

(3-day repeat cycle). The length of the perpendicular baseline was approximately 420m. Both images use the same coherence scale, from black at 0 to white at 1. Note the significant improvement in the coherence contrast achieved through filtering.

For the northeastern portion of the scene, Figure 3-14 highlights those areas where the coherence estimate increases (in the majority of pixels) or decreases

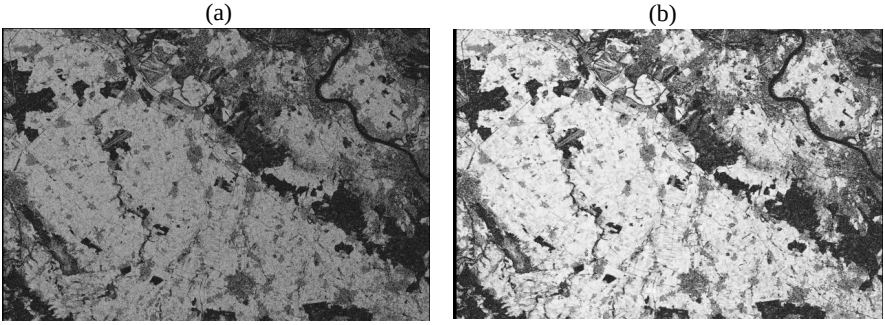


Figure 3-13: ERS-1 coherence improvement from range filtering - Bonn, Germany; Black=0, White=1; Image extent 30×40km - (a) Without filtering - (b) With filtering [© ESA / RSL]

(comparatively rarely, predominantly in low coherence areas).

Assuming no temporal decorrelation in the repeat-pass interval, the theoretically achievable coherence without any filtering γ_H, representing the InSAR system's response, is [2]:

$$\gamma_H = \left(1 - \frac{|\Delta f_r|}{B_r}\right) \cdot \left(1 - \frac{|\Delta f_a|}{B_a}\right) \qquad \textbf{(3-35)}$$

3.4.2 Decorrelation Sources

The interferometric phase difference signal is influenced by decorrelation from the following sources:

- spatial decorrelation, caused by spectral misalignment of the ground reflectivity spectra acquired from the displaced antennae
- additive thermal noise
- temporal decorrelation

The total observed coherence is modelled [97] as:

$$\gamma = \gamma_H \cdot \gamma_{noise} \cdot \gamma_{temporal} \qquad \textbf{(3-36)}$$

with γ_H representing the systemic spatial decorrelation, γ_{noise} the additive thermal noise, and $\gamma_{temporal}$ the scene decorrelation that took place between the two acquisitions.

In designing an InSAR system for the purposes of height model generation, one is interested in maximizing all three terms, to produce a system with homogeneously high coherence. Through the spectral shift filtering described in Section 3.3.1 one can achieve a γ_H approaching unity, though the approximations inherent in describing the terrain as a linear time invariant system cause some residual decorrelation.

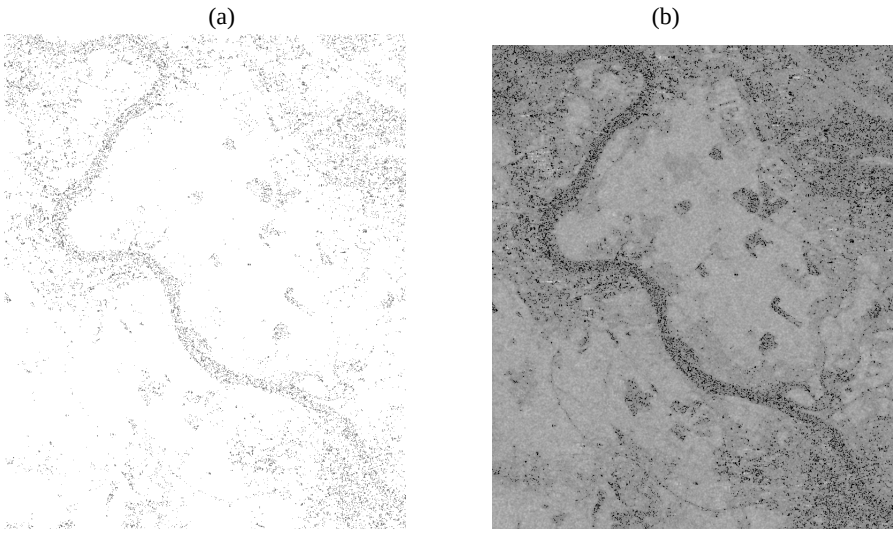


Figure 3-14: Improvement in estimation of ERS-1 coherence through range filtering - Rhine river in an area close to Bonn, Germany - (a) Non-white pixels have lower coherence through filtering - (b) Non-black pixels have higher coherence through filtering

With γ_H maximised, one is left with the temporal and noise contributions. The temporal term could be eliminated through simultaneous acquisitions (as in the airborne case), while the noise term is controllable if one can guarantee a high SNR for the echoes returned. Existing satellite InSAR systems were not designed with height model generation in mind, so the decorrelation values achievable are sub-optimal. The next section describes measures that can be taken to counteract decorrelation effects that are present due to system-design compromises.

3.5 Adaptive Filtering

Phase unwrapping becomes very difficult in areas where phase estimates are ambiguous (e.g. forest).

3.5.1 Cramer-Rao Bound

Spatial resolution can be sacrificed to obtain an improved estimate of the phase (with reduced variation) [44]. By averaging adjacent pixels, the standard devi-

ation of the phase is reduced by $1/\sqrt{N}$, where N is the number of pixels averaged.

The Cramer-Rao bound provides an upper bound on the standard deviation of the differential phase as a function of both coherence and the number of pixels averaged:

$$\sigma_\phi = \frac{1}{\sqrt{2N}} \cdot \frac{\sqrt{1-\gamma^2}}{\gamma} \qquad \textbf{(3-37)}$$

Figure 3-15 illustrates the dependence of the standard deviation of the phase on the number of looks taken and the coherence value.

One can invert Equation 3-37 to determine how many pixels must be averaged to bring the phase standard deviation below a given level:

$$N = \frac{\left(\frac{1}{\gamma^2} - 1\right)}{2\sigma_\phi^2} \qquad \textbf{(3-38)}$$

A filter can be applied to the flattened interferogram, responding to the local coherence by choosing the appropriate size of the area for averaging. Figure 3-16 illustrates the dependence of the number of “looks” on the coherence and phase noise.

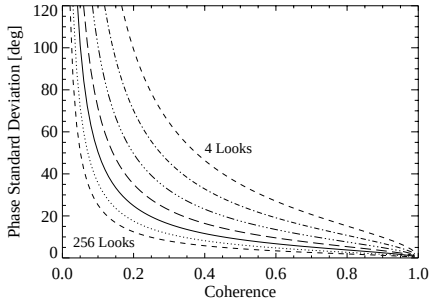


Figure 3-15: Cramer-Rao Bound: σ_ϕ as a function of coherence and estimation area

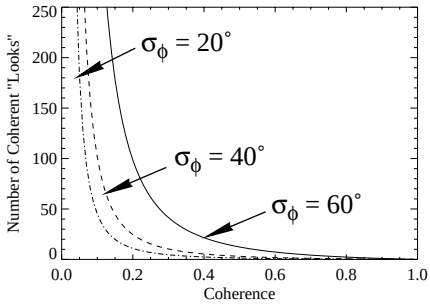


Figure 3-16: Cramer-Rao bound: estimation area as a function of coherence and σ_ϕ

3.5.2 Smoothing Filter

To expedite phase unwrapping by reducing phase noise, the phase estimate should be based upon a larger number of looks. Simple averaging is however not a good approach, as the local slope is ignored.

Instead, a bandpass filter weighting function is applied to a neighbourhood spectrum of $\hat{\underset{\sim}{F}}$, producing a filtered

interferogram $\hat{\underset{\sim}{F}}'$. Areas with a high coherence (and fringe SNR) generally allow unambiguous determination of the local peak frequency: a band-pass filter *BP* centred on the peak is used to reduce phase variation. Given $\omega = FT(\hat{\underset{\sim}{F}})$:

$$\hat{\underset{\sim}{F}}' = FT^{-1}(BP\{\omega\}) \quad \textbf{(3-39)}$$

where *FT* is a neighbourhood Fourier Transform, and FT^{-1} is its inverse.

In areas with lower coherence (and fringe SNR), reliable determination of the "peak" for use in the band-pass filter is not possible. In this case, phase variation can be reduced by using the magnitude of the local spectrum itself as a filter [25]:

$$\hat{\underset{\sim}{F}}' = FT^{-1}(|\omega|^{\alpha}\omega), \quad \textbf{(3-40)}$$

with α governing the filter's strength. This approach acknowledges that the local peak cannot be reliably found in low coherence areas, yet amplifies the strongest ground-reflectivity harmonics.

The adaptive response to local fringe SNR conditions as well as the local slope means that phase variation is reduced while at the same time preserving phase slopes. Adaptive smoothing allows the phase unwrapping step to proceed unencumbered by high phase noise.

Figure 3-17 shows the ellipsoid-flattened interferogram after adaptive filtering. In comparison to Figure 3-10, note the clearer definition in the fringes.

3.6 Phase Unwrapping

As mentioned in Section 3.3.1, the phase of the interferogram can only be known modulo 2π (see Figure 3-18). This section describes methods that have been developed to resolve this 2π ambiguity. In the discussion that follows, one distinguishes between wrapped phase values ψ and unwrapped phase values ϕ, where, adopting the notation of [23], $\psi(i, j) = W\{\phi(i, j)\}$, and

$$W\{\phi(i, j)\} = \phi(i, j) \bmod 2\pi - \pi. \quad \textbf{(3-41)}$$

Given that the maximum interval between any two neighbouring pixels is $\pm\pi$

$$|\phi(i, j) - \phi(i+1, j)| < \pi \quad \textbf{(3-42)}$$

and

$$|\phi(i, j) - \phi(i, j+1)| < \pi. \quad \textbf{(3-43)}$$

then the gradient of the true phase is equal to that of the wrapped phase:

$$\nabla\phi(i, j) = \hat{\nabla}\psi(i, j). \quad \textbf{(3-44)}$$

Knowledge of the gradient field $\nabla\phi(i, j)$ enables recovery of the phase through integration along an arbitrary path [1].

A filtered flattened interferogram $\underset{\sim}{F}$ is generally used as input:

$$\psi(i, j) = \operatorname{atan}\left(\frac{\Im\{\underset{\sim}{F}\}}{\Re\{\underset{\sim}{F}\}}\right). \quad \textbf{(3-45)}$$

Removal of systematic phase trends (flattening) produces an interferogram with fewer slopes, reducing potential ambiguity, while filtering ensures a relatively smooth phase gradient. However,

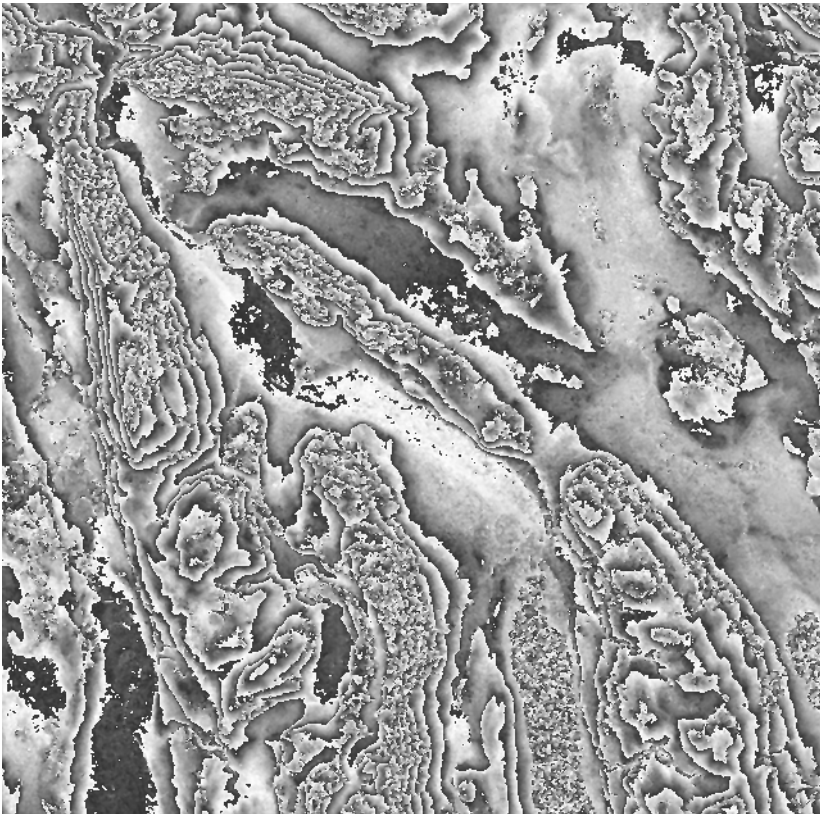


Figure 3-17: ERS-1 / ERS-2 adaptive filtered ellipsoid-flattened interferogram - Zürich, Switzerland- Nov. 4-5, 1995 [© ESA / RSL]

even after filtering, the assumption that all differences do not exceed the π bound can be violated if the terrain slope exceeds half the ambiguity height (undersampling), or if low coherence introduces sufficient phase noise.

3.6.1 Branch-Cut Method

The branch-cut method was one of the first applied to radar interferograms [26]. One first localizes pixel-to-pixel phase differences greater than π, connecting such ambiguous areas with cut-lines.

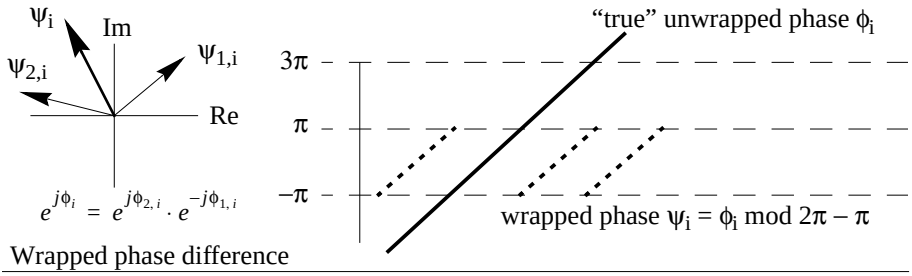


Figure 3-18: Wrapped vs. unwrapped phase

Starting at a given "seed" point, the phase is then integrated across the image, taking care not to cross any cut lines. Unreachable areas are marked with a "no data" value. The following subsections describe each of these steps in more detail.

The first task is the localization of areas with neighbourhood phase differences greater than π. Figure 3-19 illustrates a path integral around an image location (*r*,*a*) within an interferogram. Defining a "charge" function *Q* as shown in Figure 3-20,

$$Q_i = \begin{cases} -1 & \Delta\psi_i < -\pi \\ 0 \quad \text{where} & -\pi < \Delta\psi_i < \pi \\ 1 & \Delta\psi_i > \pi \end{cases} \qquad \textbf{(3-46)}$$

the "charge" *Q* notes violations of the assumptions 3-42 and 3-43.

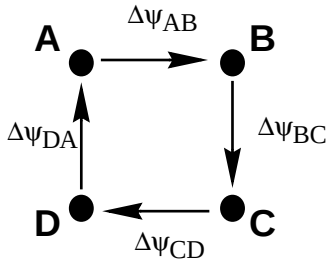


Figure 3-19: Residue Calculation

The sum of charges accumulated over the neighbourhood path ABCDA,

$$Q = Q_{AB} + Q_{BC} + Q_{CD} + Q_{DA} \qquad \textbf{(3-47)}$$

should be zero, or the path integral around the point will not be conservative. In that case, integrating the first derivative of the wrapped phase does not return the true phase, but adds a spurious

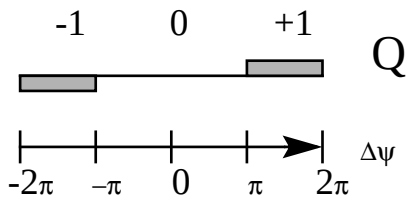


Figure 3-20: "Charge" from phase difference

$\pm 2\pi$. Points with $Q > 0$ are known as positive residues, while negative residues have $Q < 0$.

By connecting positive and negative residues with "branch cuts" [26][65], and not allowing any integration paths to cross such cut lines, one ensures that the wrapped phase gradient used is conservative ($\hat{\nabla}\psi = \nabla\phi$).

Phase integration is performed by successively expanding a wavefront of "visited" pixels. Pixels neighbouring the current "visited" set are considered for expansion in the next round if they are not blocked by a cut line.

The solution produced using such integration is not unique, as it depends on the positioning of the branch-cuts. Although the branch-cut method is one of the oldest applied to the InSAR problem, the intense research activity devoted to the phase unwrapping problem in recent years has not yet produced a robust rival that consistently supplies better results when using the same input data.

3.6.2 Least Squares Estimation

Least squares estimation (LSE) phase unwrapping methods [23] have emerged in recent years and undergone successive refinement [1][68]. They do not use the path-following approach associated with the branch-cut method, but instead apply least-squares optimization to arrive at a unique solution.

The wrapped phase gradient is considered to be a sum of the unwrapped phase gradient and a noise term $\underset{\sim}{\eta}$ [1]:

$$\hat{\nabla}\psi = \nabla\phi + \underset{\sim}{\eta} \qquad \textbf{(3-48)}$$

One then estimates the unwrapped phase by minimizing the noise through least squares solution of:

$$\sum_i \sum_j |\nabla\phi(i, j) - W\{\nabla\psi(i, j)\}|^2 \qquad \textbf{(3-49)}$$

The objective function can be minimized quickly using Fourier transforms [23], producing a unique solution (not path-dependent). However, residues continue to cause problems by warping the unwrapped phase field. A weighted least squares method [23] deals with low-coherence areas common in repeat-pass InSAR data sets through setting zero weight to unreliable phase values.

However, the presence of non-zero slopes creates bias [1], causing unacceptable systematic underestimation of the slope, even at coherence values as

high as 0.8. The weighting procedure must connect positive with negative residues to avoid such bias [1], driving LSE towards equivalence with the branch-cut method described earlier.

3.6.3 DEM Flattening

If a high quality reference DEM exists for the study area, the methodology described in Section 3.3.3 can be used to "flatten" the interferogram to the point where residual fringes all lie within the interval [-π, +π]. In areas with light rolling topography, a coarse DEM might be used in this way to flatten an interferogram to within one fringe, which could be used to refine the original DEM.

Unfortunately, DEM flattening does not help in areas with strong slopes. In such cases, a coarse DEM provides only marginally better flattening than a simple ellipsoid.

However, the technique is useful for height model validation (as shown in Section 6.3.3). When height model generation is not the goal (e.g. differential interferometry), and a high quality reference DEM already exists for the study area, the approach is unequalled in its ability to highlight coherent phase shifts (caused for example by tectonic plate movements [52]).

3.6.4 Multi-baseline Flattening

A variation of the DEM flattening method makes use of interferograms with baselines of varying height sensitivity.

Through the use of multiple interferograms, the extent of the 2π ambiguity can be reduced by using pairs with perpendicular baselines of varying lengths. Interferograms with large ambiguity heights are more easily unwrapped, and can be used to flatten interferograms with larger, more sensitive, baselines.

Interferograms with smaller baselines have larger ambiguity heights, making them less likely to be undersampled. A small baseline interferogram can be unwrapped and used to flatten a co-registered interferogram acquired with a larger baseline and a smaller ambiguity height.

Alternatively, if the perpendicular components of the baselines are near integer multiples of one another, the large baseline interferogram can be directly used to flatten the smaller baseline interferogram [50]. Unfortunately, such an approach also increases the phase noise [94], rendering the approach impractical in many cases.

3.6.5 Multi-baseline Unwrapping

Also known as "maximum likelihood unwrapping" [13], this method is a modification of the path-following branch-cut methodology, and considers a set of N co-registered wrapped interferograms with phase values $\{\phi_1, \phi_2, \ldots, \phi_N\}$ acquired with different baselines.

During phase integration, the conditional probability for height increments Δh is calculated given the set of local phase increments $\{\Delta\phi_1, \Delta\phi_2, \ldots, \Delta\phi_N\}$.

Phase to height conversion must be integrated within the unwrapping step, and can be applied within patches to allow local simplification of the calculation [13]. A confidence or "reliability" measure can be derived from the integral of the non-peak portion of the probability density function (pdf).

The "reliability" of all points on the expanding "visited" unwrapped integration front is evaluated, and the point on the front with the highest value (above a threshold) is used for expansion.

As the method makes use of more information (extra interferograms) its unwrapped phase estimates are more reliable.

3.6.6 Unwrapping Conclusions

The branch-cut and least-squares methods produce roughly comparable results when given equivalent input data. Differences appear in methods that use additional input data (e.g. rough DEM or multiple interferograms).

In the single-interferogram case, both the branch-cut and least squares methods have many similarities. Although the branch-cut output is dependent on integration path taken, and therefore non-unique, the least squares method must also resort to similar cut-lines to avoid inherent slope biases. Given low coherence areas and therefore an absence of unambiguous phase values, it may be unrealistic to expect an ambiguity-free unwrapped phase field given only a single interferogram.

For this reason, DEM generation using spaceborne repeat-pass data requires use of multiple interferograms to increase the reliability of the unwrapped phase estimates.

3.7 Phase Unflattening

Once unwrapped, the absolute phase ϕ_{UF} may be "unflattened" (the reverse of the flattening process described in Section 3.3.3) via:

$$\phi_{UG}(r, a) = \phi_{UF}(r, a) + \phi_m(r, a) \qquad \textbf{(3-50)}$$

where $\phi_{UG}(r, a)$ is introduced here as the unwrapped absolute phase.

Although the phase unwrapping step resolves the 2π ambiguity, determination of the 2π *multiple* is beyond its scope. During phase integration, the "seed" point is assumed to have an unwrapped phase in the modulo 2π interval. Tie-points are generally therefore used (a single point suffices) to localize the 2π interval through addition of a phase constant ϕ_c.

Given knowledge of the phase constant, the range difference between the two acquisitions is directly related to the unflattened unwrapped phase:

$$\delta_r = \frac{\lambda}{4\pi} \cdot (\phi_{UG} + \phi_c) . \qquad \textbf{(3-51)}$$

As first described in Chapter 1, height model generation through InSAR consists of first estimating the interferometric phase ϕ, unwrapping it, transforming it to an estimate of difference in distance δ_r, and using this geometric measure to generate a height model.
Now that δ_r has been estimated, height calculation and terrain geocoding steps may now be carried out unimpeded.

3.8 Visualization of Interferogram Products

Figures 3-21 to 3-23 illustrate various stages in the processing of the interferogram. ERS-1 data acquired from an area close to Bonn, Germany on March 14 and 29, 1992 is used throughout. All images show the same subscene in slant-range geometry. All processing was done using the ZIP software described in Appendix C.

Figures 3-21(a) and 3-21(b) present the intensity images corresponding to the March 14 and March 29 acquisitions. Figure 3-21(c) shows the phase of the raw interferogram $\underset{\sim}{G}$ (note the fringe lines roughly parallel to the azimuth direction), while Figure 3-21(d) displays the corresponding coherence image.

Figure 3-22(a) shows the modelled ellipsoid phase ϕ_e, next to (b) the flattened interferogram that results when it is subtracted from the raw interferogram. Similarly, Figure 3-22(c) presents the modelled DEM phase ϕ_d, next to (d) the DEM-flattened interferogram that results when ϕ_d is subtracted from the raw interferogram.
Figure 3-23(a) illustrates a shaded relief representation of the reference DEM (after being transformed into slant-range geometry), while (b) shows the same slant-range DEM as a greyscale. Figure 3-23(c) presents an ellipsoid-flattened version (i.e. $\phi_d - \phi_e$) of the DEM synthetic interferogram, next to (d) the ellipsoid-flattened version of the true interferogram (identical to Figure 3-22(b)).

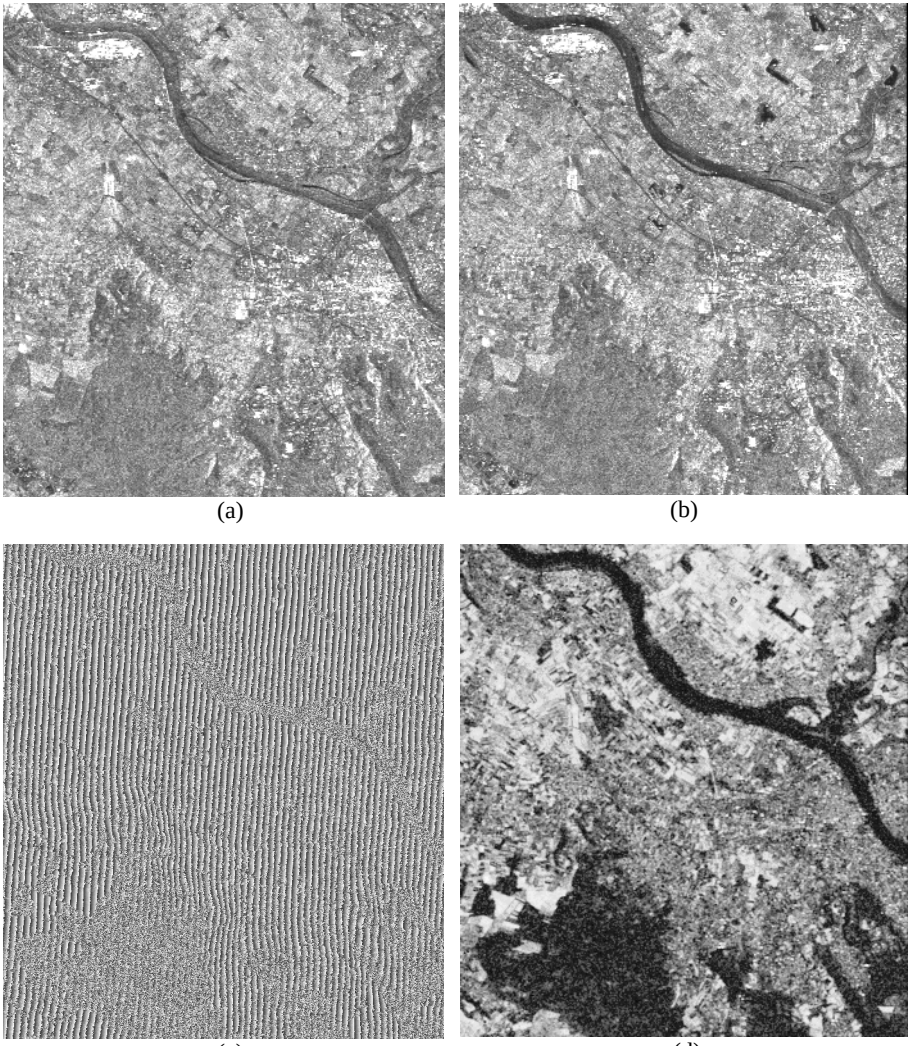


Figure 3-21: ERS-1 InSAR products - Bonn, Germany - March 14-29, 1992
(a) Backscatter March 14, 1992, (b) Backscatter March 29, 1992,
(c) Raw interferogram, (d) Coherence

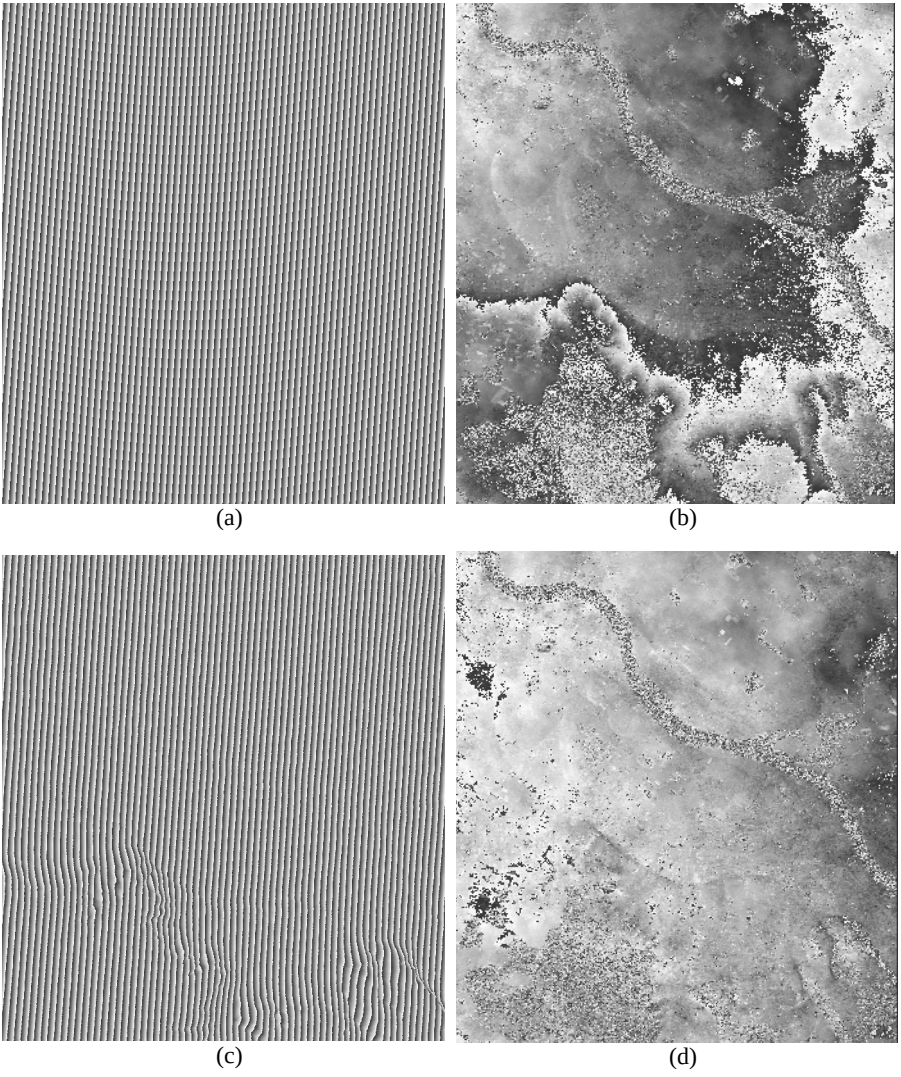


Figure 3-22: ERS-1 InSAR products - Bonn, Germany - March 14-29, 1992
(a) Synthetic-ellipsoid phase, (b) Ellipsoid-flattened ERS interferogram, (c) Synthetic-DEM phase, (d) DEM-flattened interferogram

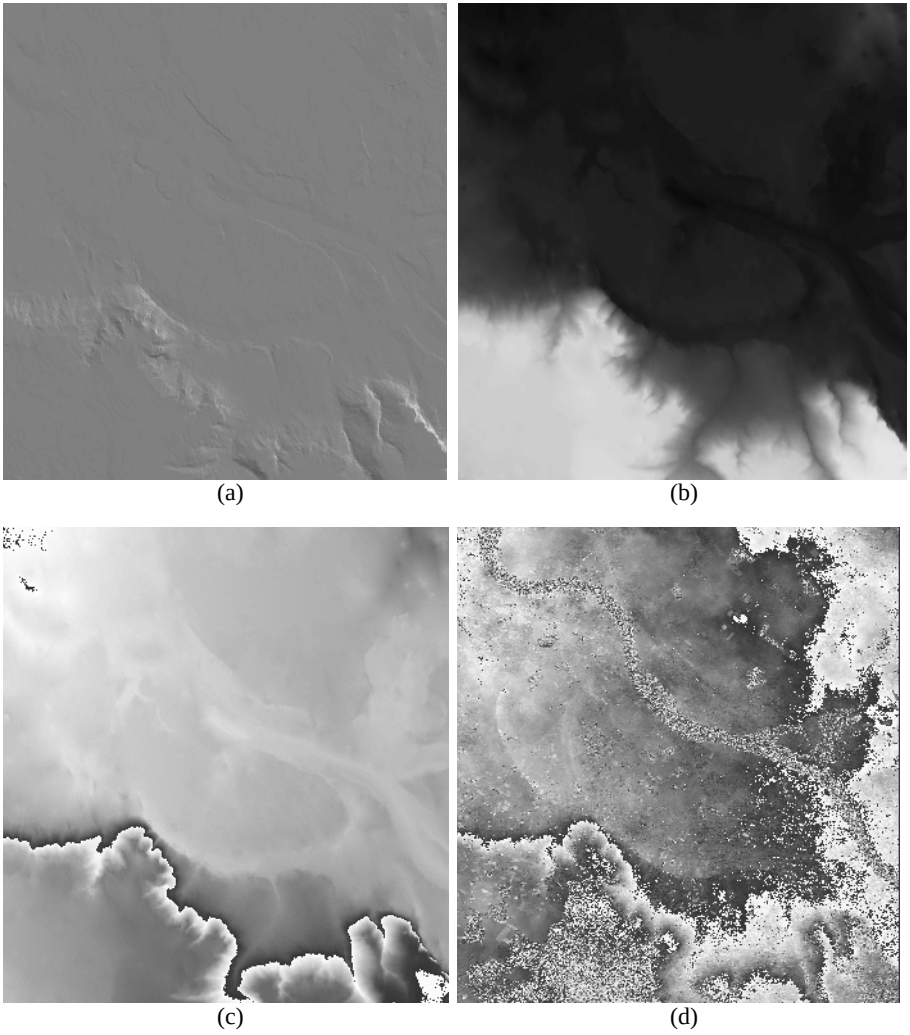


Figure 3-23: ERS-1 InSAR products - Bonn, Germany - March 14-29, 1992
(a) Shaded relief model of reference DEM, (b) Reference DEM,
(c) Ellipsoid-flattened synthetic interferogram, (d) Ellipsoid-flattened ERS interferogram

Chapter 4

Geocoding and Height Model Generation

4.1 Introduction

As outlined in previous chapters, the generation of height models through SAR interferometry can be broadly broken down into the following steps (see Figure 1-7):

- estimation of phase difference ϕ
- transformation to difference in distance δ_r
- transformation of difference in distance to height values h

Estimation of first ϕ and then δ_r was covered in Chapter 3. This chapter describes the generation of the height model together with the geocoding process, the transformation from radar slant range vs. azimuth coordinate system into a map projection system.

4.2 Cartographic Transformations

The following sections describe some of the principles involved in specifying the position of a point on the Earth's surface.

4.2.1 Map Projections

Height models are generally presented in a specific map projection: some general properties are therefore reviewed here.

As illustrated in Figure 4-1, map projections were invented to represent the three dimensional round Earth on a flat two dimensional surface [87].

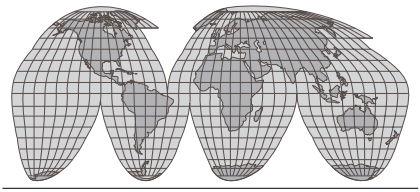


Figure 4-1: Map projections - transforming 3D information into a 2D plane

Common systems project locations geometrically onto a cylinder, cone, or tangent plane, which can then be "unfolded" onto a flat surface. The two dimensional map coordinates are generally labelled "Northing" (vertical) and "Easting" (horizontal).

4.2.2 Oblate Ellipsoid

The rotation of the Earth causes its radius to expand slightly at the equator in comparison to the poles: geographers therefore often use an ellipsoid rather than a sphere to model its shape. More specifically, one uses an oblate ellipsoid, as two of its three axes are equal in length.

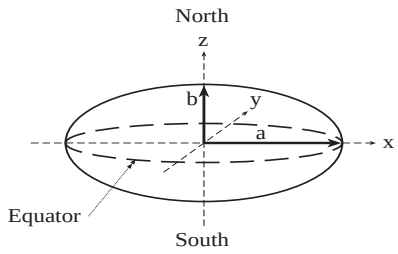


Figure 4-2: The Earth as an oblate ellipsoid

Figure 4-2 illustrates such an oblate ellipsoid oriented along the Cartesian axes x, y, z, with the lengths of the ellipsoids x and y axes equal in length. Geocentric ellipsoids are centred at $\{x, y, z\} = \{0, 0, 0\}$. The extent of the equatorial bulge is exaggerated for emphasis. The lengths of the radii of the semi-major (x,y) and semi-minor (z) axes are labelled a and b respectively. Given the latitude, longitude, and height of a point related to such an ellipsoid, one is able to calculate the point's Cartesian coordinates as [28]:

$$\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} (r_E + h)\cos\zeta\cos\Pi \\ (r_E + h)\cos\zeta\sin\Pi \\ [(1-e^2)r_E + h]\sin\zeta \end{bmatrix}, \qquad \textbf{(4-1)}$$

where ζ is latitude, Π is longitude, $e = (a^2 - b^2)/a^2$ is the ellipsoid's eccentricity, and the local Earth radius r_E is:

$$r_E = \frac{a}{\sqrt{1-(e\sin\zeta)^2}}. \qquad \textbf{(4-2)}$$

4.2.3 Datum Shifts

Early surveyors were always concerned with a geographically constrained area of the world (e.g. their country, empire, etc.). Initially, they developed map projections based geometrically on spheres or ellipsoids whose parameters best described the curvature of the Earth in their region.

In more recent times, mapping and surveying have developed into a global enterprise, and demand for interoperability between different map projection systems has arisen. Interchange between different map systems is governed by a *datum shift*, whereby the reference ellipsoid of a given local map projection *A* is shifted (possibly also rotated and/or scaled) to correspond to that of another datum *B*. Mathematically, such a trans-

formation is known as a three dimensional Helmert transformation [14]:

$$P_B = \begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = M \cdot D \cdot P_A + \Delta P \qquad \textbf{(4-3)}$$

$$P_B = M \cdot \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \cdot \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \qquad \textbf{(4-4)}$$

The scaling is performed by the scalar operator $M = (1 + m)$, while the ellipsoid is displaced by the translation vector ΔP and the rotation matrix D rotates the ellipsoid around each of the coordinate axes by α, β, and γ respectively. The elements of the rotation matrix D are:

$$d_{11} = \cos\beta\cos\gamma$$

$$d_{12} = \cos\alpha\sin\gamma + \sin\alpha\sin\beta\cos\gamma$$

$$d_{13} = \sin\alpha\sin\gamma - \cos\alpha\sin\beta\cos\gamma$$

$$d_{21} = -\cos\beta\sin\gamma$$

$$d_{22} = \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma$$

$$d_{23} = \sin\alpha\cos\gamma + \cos\alpha\sin\beta\sin\gamma$$

$$d_{31} = \sin\beta$$

$$d_{32} = -\sin\alpha\cos\beta$$

$$d_{33} = \cos\alpha\cos\beta$$

Through translation, rotation, and scaling, one is able to translate points from local reference system *A* to reference system *B*. A datum shift with only translation terms (no rotation or scaling) is known as a three parameter datum shift. A datum shift that incorporates three rotation and a single scaling parameter in addition to the three translation parameters is known as a full seven parameter datum shift.

A datum shift is necessary if one wishes to compare surveys, for example, acquired in the Italian reference system to those in, say, the Swiss mapping system, as their reference ellipsoids differ.

Combinatorial analysis dictates that as the number of different mapping systems grow, the number of one-to-one translation systems between them quickly grows out of control. One therefore usually uses an intermediate "global" standard reference system as an intermediary. Given that one has n reference systems, one then only requires transformation definitions for all reference systems to and from the standard $(n - 1)$ rather than the

$$\binom{n}{2},$$

(i.e. "n choose 2") that would otherwise have been required.

The process for translating between any two map reference systems *A* and *B* is illustrated in Figure 4-3. For example, one might have a set of height values stored in *A*'s projection system, and require their corresponding global Cartesian coordinate values for comparison with satellite orbit ephemeris data.

Using geometric equations specific to the projection system *A* (could be cylindrical, conic, planar, etc.), one transforms the *A*-system map coordinates (Northing, Easting, height) into geographic coordinates (latitude, longitude, height). Once these have been obtained, one can calculate the equivalent Cartesian coordinates in *A*'s local datum, and through a datum shift operation such as the one presented in Equation 4-4 produce Cartesian coordinates in a global reference frame (typically WGS84). These can be directly compared with Earth-Centred-Rotating (ECR) satellite ephemeris data.

In the example described above, provision of the point's equivalent global Cartesian coordinates solved the problem. In other cases, the availability of the global Cartesian coordinates is used as a starting point to transform the point into a second map projection system *B*, through a process opposite to that described above for *A*.

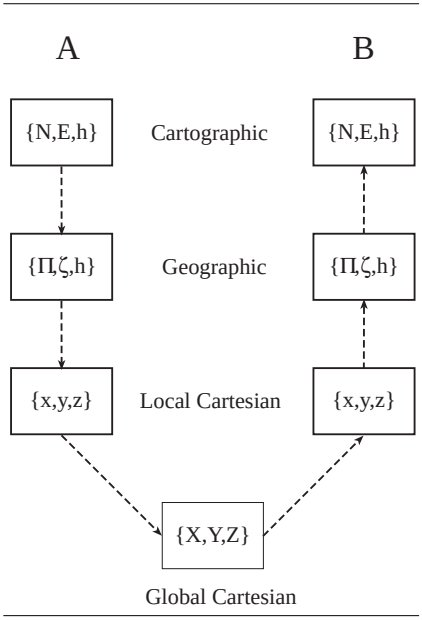


Figure 4-3: Cartographic transformations

4.3 Backward Geocoding

SAR geocoding denotes the process whereby SAR images are transformed from their native range-Doppler radar geometry into a terrestrial mapping system [8]. Using a digital elevation model of the terrain, or if a DEM is not available, an ellipsoid Earth model, the slant or ground range SAR images are resampled into a map geometry. Geocoding performed with a DEM is known as *terrain-geocoding* [57]. When an ellipsoid model is employed, the term *ellipsoid-geocoding* is used.

Both ellipsoid and terrain geocoding have become operational [77] in recent years since the launch of the ERS-1 satellite, and the establishment of processing centres across Europe. The following sections briefly review the backward geocoding process which will be used later for validation of interferometric height models. Although ellipsoid geo-

coding can also be implemented using a backward geocoding technique, it is not covered here.

The term backward geocoding denotes the fact that the process originates in map geometry, transforming backward to radar geometry. Within forward geocoding methodologies (see Section 4.5) one originates in radar "image" geometry and transforms forward into the reference "object" map geometry.

4.3.1 Process

The process of backward geocoding is illustrated in Figure 4-4. One iterates in map geometry through the reference DEM, and calculates the corresponding range Doppler coordinates. Once these are known, the range Doppler image can be sampled at that location (using an appropriate resampling kernel) and stored in map geometry. Iteration through the DEM can produce a completely resampled radar image in map geometry.

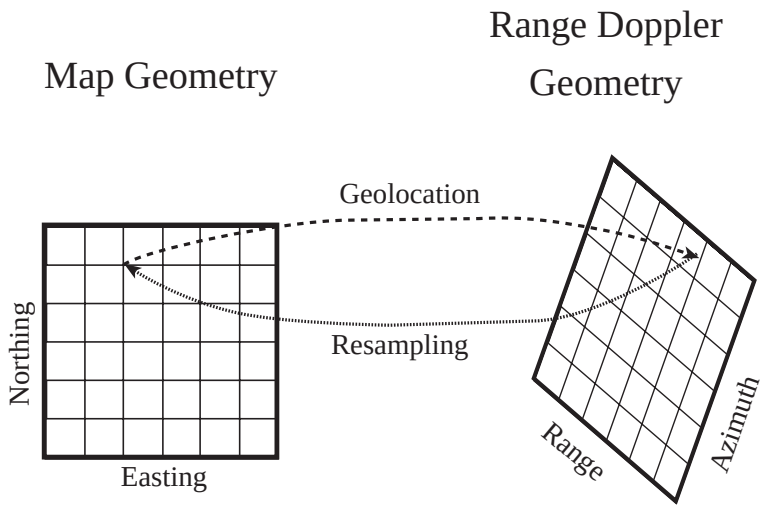


Figure 4-4: Backward geocoding methodology

4.3.2 Geolocation

For each point $\underset{\sim}{P}_{map}$ in the reference DEM with a valid height, the cartographic transformations outlined in the previous section can be used to transform the point into global Cartesian coordinate space $\underset{\sim}{P}_d$, allowing direct comparison with the orbit ephemeris data.

The spatial relationship between the radar range Doppler geometry and a Cartesian reference frame with origin at the Earth's geocentre is governed by range and Doppler equations [8].

A range "sphere" centred on the spacecraft position $\underset{\sim}{S}_1$ with radius equal to the range R_1 satisfies the equation:

$$\left|\underset{\sim}{S}_1 - \underset{\sim}{P}_d\right| = R_1 , \qquad \textbf{(4-5)}$$

while the Doppler relation can be written as:

$$\frac{2}{\lambda} \times \frac{\underset{\sim}{S}_1 - \underset{\sim}{P}_d}{\left|\underset{\sim}{S}_1 - \underset{\sim}{P}_d\right|} \cdot (\underset{\sim}{v}_p - \underset{\sim}{v}_s) = f_{proc} , \qquad \textbf{(4-6)}$$

where λ is the radar wavelength, $\underset{\sim}{v}_p$ is the velocity of the point on the Earth, $\underset{\sim}{v}_s$ is the velocity of the spacecraft, and f_{proc} is the processed Doppler frequency. Since $\underset{\sim}{P}_d$ is known and

$$\underset{\sim}{v}_p = 0 \qquad \textbf{(4-7)}$$

within an ECR coordinate system (or can be calculated from the height and the Earth rotation rate in an inertial Cartesian system), and both λ and f_{ref} are well defined, the remaining unknowns are $\underset{\sim}{S}_1$ and $\underset{\sim}{v}_s$.

Since the spacecraft orbit is often known quite accurately (see Chapter 2) one needs only to iterate along the orbit model (e.g. Equation 2-6) searching for a point satisfying the Doppler condition [56]. That point provides the spacecraft position and azimuth image coordinate a. Once these are known, Equation 4-5 provides the corresponding range distance, and range image coordinate r.

In a SAR geocoder, r and a are then used to resample the backscatter value at (r, a) corresponding to the map geometry position $\underset{\sim}{P}_{map}$.

Backward geocoding can be used to geocode a slant-range interferometric height model into map geometry for validation. It can also be used to geocode other interferometric image products (e.g. backscatter, coherence, differential phase) for intercomparison with other map geometry data. However, it is not useful for operational height model generation using InSAR, as it presumes that a high resolution elevation already exists in map geometry!

Nevertheless the technique is helpful for validation purposes, and can serve as a reference for evaluating the accuracy of other geocoding methods.

Although forward geocoding must be used for operational DEM generation, other InSAR applications are more satisfactorily served by the backward method, as one can generate a DEM of an area *once* with the forward method,

and then use the backward method later successively to terrain geocode all types of slant-range information. Using the backward method, one is not restricted to height information from a single InSAR pair: the reference elevation model used could be an amalgam formed from several acquisitions, even a mosaic formed through combination of DEM's from different InSAR sensors.

4.4 Ancillary Data Products

If one takes some care during the terrain geocoding process to iterate through the DEM first from near to far range, and then in a second pass from far to near range, one can generate useful ancillary data products that indicate areas in a radar shadow and/or radar layover condition. Figure 4-5 illustrates the geometry of both the layover and shadow conditions.

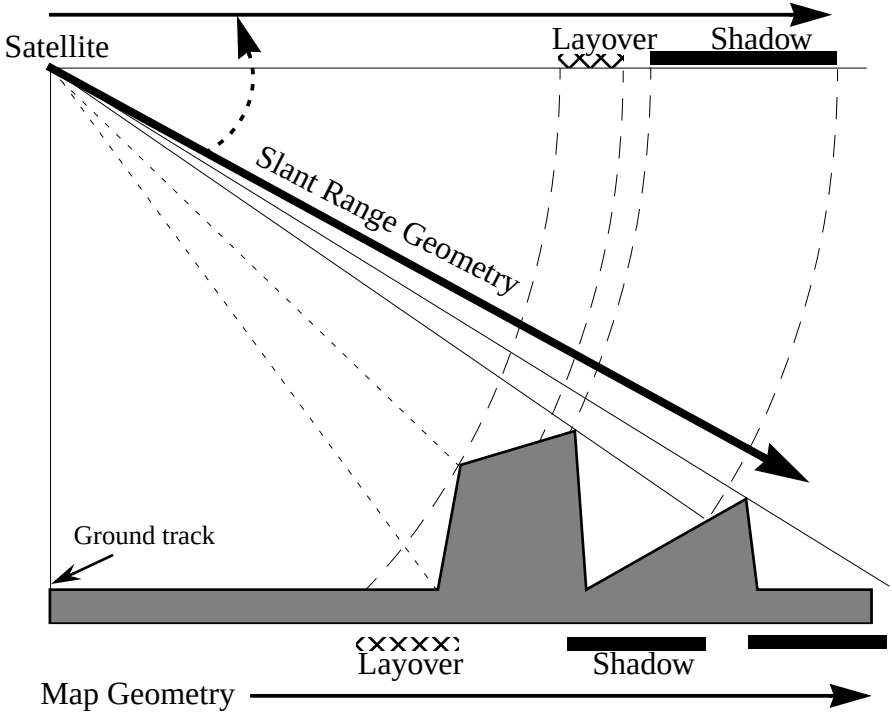


Figure 4-5: SAR layover and shadow geometry: slant range and map geometries

4.4.1 Radar Shadow

If one keeps a running tab of the lowest off-nadir angle for each azimuth line during the near-to-far iteration through the DEM, the off-nadir angle should be monotonically increasing, except in areas with radar shadow, where it will decrease briefly (see Figure 4-5), before resuming it climb [56]. This characteristic can be used to mark DEM locations within radar shadow, a property that is also important for image simulation (see Chapter 4).

4.4.2 Radar Layover

In areas with steep slopes, the peak of a mountain may be closer to the radar satellite than an adjacent valley floor situated closer to the satellite's ground track. The effect causes concentration of a large map geometry area within a small slant range region, and is strongest at steep incidence angles.

If one iterates through the DEM from near to far range, keeping a record within each azimuth line a (see Figure 4-4), the slant range R_1 should usually be increasing; areas where this is not the case can be flagged as layover [56]. Unfortunately this method alone does not capture areas sweeping back ("recovering") from layover but still within its bounds. To capture those areas, a second pass through the DEM is required, this time from far to near range: areas where the slant range increases (rather than decreases) are also marked as layover. One distinguishes between *active* map geometry areas that cause layover and *passive* areas whose backscatter is overlaid with contributions from adjacent layover action.

4.4.3 Local Incidence Angle

The local incidence angle is also easily calculated for each point in the DEM by first computing the normal vector $\underset{\sim}{N}$ of the local terrain using a neighbourhood planar estimate. The component of the local surface plane in the azimuth plane defined by normal vector $\underset{\sim}{A} = \underset{\sim}{S}_1 \times \underset{\sim}{P}_d$ is:

$$\underset{\sim}{N}_A = \underset{\sim}{N} \times \underset{\sim}{A}, \qquad \textbf{(4-8)}$$

and the local angle of incidence θ is then:

$$\theta = \frac{\pi}{2} - \operatorname{acos}\left(\frac{\underset{\sim}{R}_1 \cdot \underset{\sim}{N}_A}{|\underset{\sim}{R}_1||\underset{\sim}{N}_A|}\right) \qquad \textbf{(4-9)}$$

4.5 Forward Geocoding

The terrain geocoding methodology described in Section 4.3 can be applied to any single-scene SAR data, provided that a high resolution DEM is available for the area.

In the case of operational height model generation using InSAR, no pre-existing DEM can be assumed: backward geocoding cannot be employed - one must use a forward approach. Operational height model generation should be distinguished here from InSAR

height model validation: in that case a reference height model is available, and backward geocoding can be a valid means of enabling comparison.

4.5.1 Range Doppler Range (RDR)

In the SAR interferometry case, at least two SAR images are available, and in addition to Equations 4-5 and 4-6, similar equations are also available for the second scene [24]:

$$\left|\underset{\sim}{S}_2 - \underset{\sim}{P}_d\right| = R_2 , \qquad \textbf{(4-10)}$$

and

$$\frac{2}{\lambda} \times \frac{\underset{\sim}{S}_2 - \underset{\sim}{P}_d}{\left|\underset{\sim}{S}_2 - \underset{\sim}{P}_d\right|} \cdot (\underset{\sim}{v}_p - \underset{\sim}{v}_s) = f_{proc,2} . \qquad \textbf{(4-11)}$$

Through knowledge of the difference in range distance δ_r for each point in the two acquisitions, one can rewrite Equation 4-10 as

$$\left|\underset{\sim}{S}_2 - \underset{\sim}{P}_d\right| = R_1 + \delta_r , \qquad \textbf{(4-12)}$$

making use of the information gained through the interferometric phase. Equations 4-5, 4-6, and 4-12 form a set of three equations in three unknowns $\{P_x, P_y, P_z\}$. Although they are non-linear, iterative numerical methods can be used to solve for $\underset{\sim}{P}_d$.

An initial starting point $\{P_{x_0}, P_{y_0}, P_{z_0}\}$ within the scene is selected and then an iterative method (e.g. Newton-Raphson [67]) is used to solve the set of simultaneous non-linear equations. That solution then serves as the starting point for the solution of the equations defining the next point.

A set of irregularly-gridded points $\{P_x, P_y, P_z\}$ in WGS84 global Cartesian geometry is the result. These are transformed easily into a map geometry via path *B* shown in Figure 4-3, resulting in map coordinates $\{P_N, P_E, P_h\}$. As the points do not generally coincide with a regular grid in any map geometry, a regridding step (e.g. triangulation) follows the solution of the equations (see Figure 4-6).

In its solution of three simultaneous geolocation equations, the RDR approach does not distinguish between geocoding and height model generation - both position and height are produced simultaneously.

4.5.2 Slant Range Height Model

In spite of the simplicity of the RDR approach described above, it can be beneficial to first calculate the height model in slant range geometry (see Section 4.8), and in a later step forward geocode to map geometry.

In such a case, the second acquisition's range sphere equation (Equation 4-12) is replaced with an equation describing an oblate ellipsoid inflated by the height value determined interferometrically.

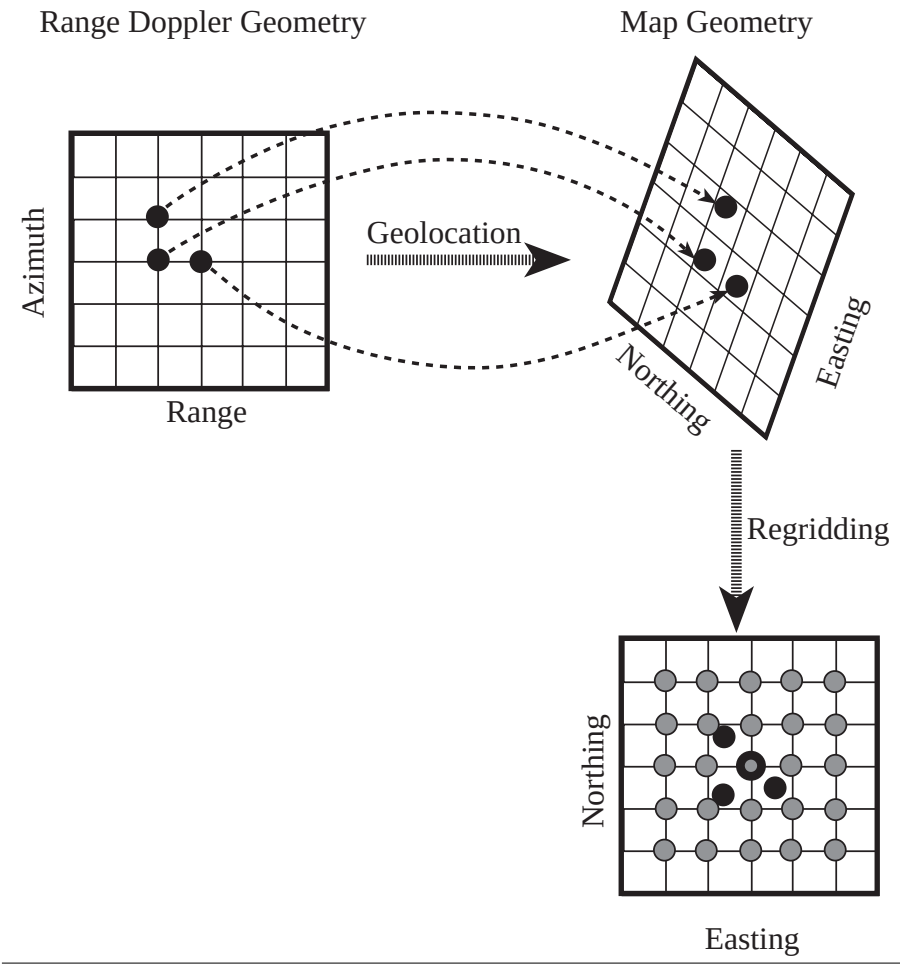


Figure 4-6: Forward geocoding methodology

An inflated oblate ellipsoid is described using the equation

$$\frac{(P_x - x_0)^2 + (P_y - y_0)^2}{(a+h)^2} + \frac{(P_z - z_0)^2}{(b+h)^2} = 1 \quad , \qquad \textbf{(4-13)}$$

where h is the height estimate determined interferometrically, a and b are respectively the semi-major and semi-minor axes of the oblate ellipsoid.

Equation 4-13 is the simplified version appropriate for a three parameter datum shift transformation - translation only: (x_0, y_0, z_0). Although it shares the same principle, the more general seven parameter datum shift transformation (Equation 4-4) is somewhat more complicated, as it adds a scale and three rotation parameters.

The ellipsoid equation 4-13 deviates minimally from the form of a height h above a reference ellipsoid (Equation 4-1) with axes a and b. For height values of h=1000m, numerical simulation showed maximum deviation to be less than 2mm (at mid latitudes).

Equations 4-5, 4-6, and 4-13 form a set of three equations in three unknowns $\{P_x, P_y, P_z\}$. As in the RDR method, an initial starting point $(P_{x_0}, P_{y_0}, P_{z_0})$ is selected within the scene and an iterative method is used to solve the set of simultaneous non-linear equations. That solution then serves as the starting point for the solution of the equations defining the next point. As with RDR, a transformation from global Cartesian to map geometry $\{P_N, P_E, P_h\}$ and a regridding step follow.

4.6 Comparison of Backward vs. Forward

Both backward and forward geocoding methodologies that directly support full seven parameter datum shift transformations were implemented and used to calculate the results that follow. Both implementations geocode an ERS quarter scene in less than an hour on a contemporary UNIX workstation. Provided that the reference and interferometric height model are of comparable accuracies, the quality of the geocoding from the two methods is nearly identical.

An interferometric height model from an interferogram calculated using ERS-1 three day repeat data acquired over Bonn, Germany in March, 1992 was geocoded using both the forward and backward methodologies [81]. Figure 4-7 shows the difference between the forward and backward geocoded height models as a gray scale, with white/black saturation at ±5 m.

Height differences appear at gravel pits not considered in the reference DEM: differing height references result in planimetric shifts. Noisy InSAR height estimates within forested areas are also geolocated differently. A kidney-shaped mask marks a forested area

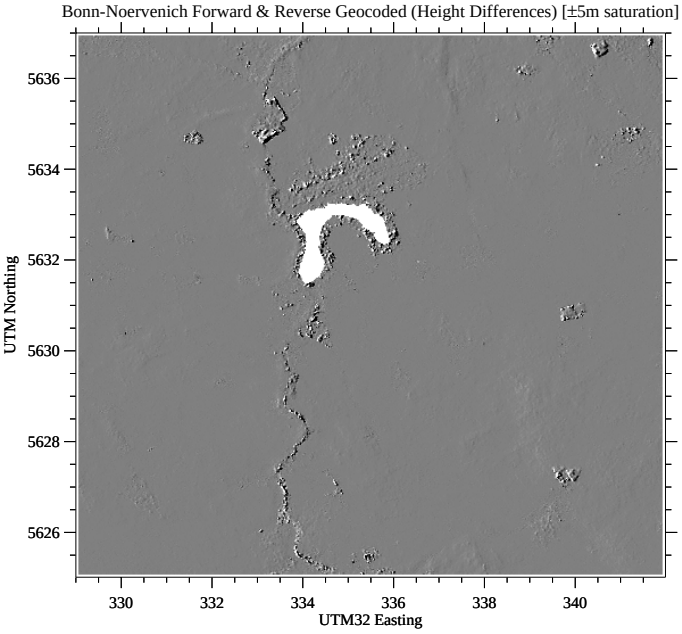


Figure 4-7: Comparison of interferometric height models geocoded using backward and forward methodologies - Bonn, Germany - March, 1992 -grayscale height difference from -5m to +5m - white in forest was inaccessible to phase unwrapping - otherwise, saturation at white=+5m, black=-5m

that was not penetrated by the phase-unwrapping algorithm. Otherwise, there are only marginal differences between the final height models (in map geometry).

4.7 Refinement of Geometry

The accuracy of the ERS orbits delivered by the German Processing and Archiving Facility (D-PAF) is approximately 30cm in the along and cross-track directions, with a slightly higher radial accuracy (8cm) [49]. However, phase to

height conversion requires a higher standard.

The geometry must therefore be refined using tiepoints. Two refinements are performed using two types of ground control points (GCP's). Height tiepoints are measured in areas of uniform unwrapped phase, while position tiepoints (Northing, Easting) are measured in parts of the image that can be localized with high geometric accuracy (e.g. bridges, road crossings) [82].

Height tiepoints distributed well across range and azimuth are used to refine the baseline model. An iterative non-linear least squares fit algorithm is used to estimate a set of parameters, including the absolute phase constant, the cross-track baseline component and its trend along the azimuth dimension. The radial orbit component is usually left at its nominal value, as its estimate is more accurate.

Before such refinement height estimates have errors on the order of kilometres (open loop). Once the phase constant (only) has been localized, height errors are typically reduced by an order of magnitude. Further refinement of the cross-track baseline component (and trend in azimuth) reduces the height estimation error to still lower values.

Position tiepoints are used to refine the SAR image acquisition geometry, in a manner similar to that employed in conventional GTC production (terrain-geocoding). As with the baseline geometry above, several approaches are also possible here. In conventional ERS GTC production, the azimuth starting point, pixel spacing, near range value, and range pixel spacing are all refined based on tiepoints well distributed across the scene. One may also use the tiepoints to refine the orbit itself. Given a polynomial description of the orbit, one typically refines the constant and linear terms, leaving higher order terms at their nominal values.

Some InSAR processors use GCP's to remove systematic artefacts resulting from inaccurate Earth flattening. As implemented in the processor described here, GCP's are not used for any such purpose, as the expected ellipsoid phase used during flattening is reintroduced after phase unwrapping - see Equation 3-50. There is no need for correction of the flattening using GCP's.

It is the baseline and the imaging geometries that are refined using the GCP sets. Based on the refined output, the interferometric height of each point in the image with a valid unwrapped phase may then be calculated [82]. The following sections describe that process.

4.8 Phase to Height Conversion Methods

Although a flattened interferogram in many ways resembles a "topographic contour map" the truth is that many non-trivial steps are necessary to move from

this stage to the geocoded height model desired. Resolving the 2π ambiguity (phase unwrapping) can in some cases be an insurmountable task (or at least require manual operator intervention).

The following sections describe various methods [81] for conversion of the interferometric phase to a height model.

4.8.1 Height Direct from Unwrapped Flattened Phase

The strong prima facie resemblance between flattened interferometric phase and height contour maps suggests that conversion of phase[1] to height should be a simple matter.

Using the concept of ambiguity height $\Delta h_{2\pi}$, one can construct a simple relation:

$$h(r, a) = h_0 + \Delta\phi_F(r, a) \cdot \frac{\Delta h_{2\pi}}{2\pi}, \qquad \textbf{(4-14)}$$

where h_0 is a reference height (GCP) within the image with phase $\phi_{F,0}$, and $\Delta\phi_F(r, a) = \phi_F(r, a) - \phi_{F,0}$.

Unfortunately, robust phase to height conversion is not quite so simple. The concept of ambiguity height itself is only valid within restricted contexts. As discussed in Chapter 1, a far-field approximation is necessary, as is an assumption of parallel orbits.

The far-field approximation distorts the true effect of the interferometric baseline, requiring adjustment using tiepoints to correct the bias introduced [84]. And the fact that satellite orbits are not truly parallel causes iso-range pixels adjacent in azimuth to have slightly different baselines, leading to disparate phase to height sensitivities (and ambiguity heights).

The method also suffers from the fact that the ambiguity height is not itself independent of height.

Correction of the biases introduced is not a matter of improved reflattening using tiepoints, as such measures can neither counteract the height nor the azimuth[2] dependence of the ambiguity height.

Despite these shortcomings, the method's simplicity continues to make it popular; many InSAR processors continue to use the approach.

4.8.2 Empirical Look-up Table

One can extend the ellipsoid flattening procedure described in Section 3.3.3 to produce not one ellipsoid phase difference value ϕ_e per pixel, but multiple values $\{\phi_1(h_1), \phi_2(h_2), \ldots, \phi_3(h_N)\}$ for a set of N reference heights spread across the expected range for the scene (e.g. $\{0, 2000, 4000\}$ metres etc.) [78].

One now has a set of points where both the height and corresponding expected phase are known. A polyno-

1. N.B. unwrapped

2. due to converging orbits

mial function can be used to model the phase to height relationship:

$$h = \sum_{i=1}^{M} c_i \cdot \phi_e^{i-1}, \qquad \textbf{(4-15)}$$

where the set of phases and heights are used to solve for the series of polynomial coefficients $\{c_1, \ldots, c_M\}$:

$$\begin{bmatrix} c_1 \\ \ldots \\ c_M \end{bmatrix} = \begin{bmatrix} 1 & \phi_1 & \ldots & \phi_1^{M-1} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & \phi_N & \ldots & \phi_N^{M-1} \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ \ldots \\ h_N \end{bmatrix}. \qquad \textbf{(4-16)}$$

Although one could calculate such coefficients for every range and azimuth coordinate in the radar image, inversion of the matrix above is time consuming, so (as with the ellipsoid flattening described in Section 3.3.3) a faster solution is to compute the coefficients for a set of points on a coarse grid distributed across both range and azimuth. Once these are determined, the coefficients $\{c_1(r, a), \ldots, c_M(r, a)\}$ corresponding to an arbitrary (range, azimuth) location can be calculated either through a 2D polynomial fit to the results from the coarse grid, or through interpolation between coarse grid values. The height of an arbitrary pixel is then calculated as

$$h(r, a) \qquad \textbf{(4-17)}$$

$$= \sum_{i=1}^{M} c_i(r, a) \cdot [\phi_{UG}(r, a) - \phi_c]^{i-1}$$

where ϕ_c is the phase constant introduced in Section 3.7 caused by ambiguity remaining after phase unwrapping.

The empirical look-up table approach requires a single phase constant and captures the true 3D baseline (see Figure 2-3) through image simulation. Given highly precise knowledge of the baseline, and at the cost of increased complexity, orbit pairs with significant divergence can be accurately treated. However, tiepoint-based adjustment remains necessary in non-ideal cases.

4.8.3 Baseline Rotation

Another method of phase to height conversion calculates the internal angle between the baseline and the local look vector, rotates the baseline to form the look vector, and then calculates the corresponding height value. Baseline rotation methods are in some ways simpler than those that use image simulation, as they do not capture the full 3D baseline.

Far Field Approximation: As introduced in Section 1.3.2, to enable simplified models, the slant range "look" vectors for the two acquisitions may be assumed to be parallel (in the far field). The internal angle ξ in the azimuth plane between the baseline vector and the local look vector (see Figure 4-8) is then approximated by: $\cos(\xi) \approx \delta_r / |\underset{\sim}{B}|$.

Cosine Law: The angle can be calculated more exactly. For each point in the

image, given knowledge of the difference in slant range distance between the two acquisitions δ_r, the slant range distance during the second acquisition is $R_2 = R_1 + \delta_r$. Since all sides of the triangle formed by R_1, R_2, and $|\underset{\sim}{B}|$ are known, the cosine law allows calculation of the internal angle ξ:

$$\cos(\xi) = \frac{|\underset{\sim}{B}|^2 + |\underset{\sim}{R}_1|^2 - |\underset{\sim}{R}_2|^2}{2|\underset{\sim}{B}||\underset{\sim}{R}_1|} \qquad \textbf{(4-18)}$$

Height calculation: Rotating the baseline vector down by this angle forms the local look vector, and the height above the reference ellipsoid can be calculated [82].

Defining the ground range difference $g_{d,i}$ as the difference between the ground range at pixel i (g_i) and the scene centre ground range (g_c), Figure 4-8 shows that $g_{d,i}$ may be calculated as:

$$g_{d,i} = \sqrt{R_{1,i}^2 - H^2} - \sqrt{R_{1,c}^2 - H^2} \qquad \textbf{(4-19)}$$

$$g_{d,i} = R_{1,i} \cdot \sin(\theta_i + \beta) - R_{1,c} \cdot \sin(\theta_c + \beta) \quad , \qquad \textbf{(4-20)}$$

where θ_c is the incidence angle at mid-scene, and θ_i is the local incidence angle.

The interferometrically derived height h_i may be estimated by rotating the baseline vector. After subsequent rotation of the look vector from the $\{\underset{\sim}{t}, \underset{\sim}{c}, \underset{\sim}{n}\}$ to the $\{\underset{\sim}{t}, \underset{\sim}{c}', \underset{\sim}{n}'\}$ coordinate system, $\hat{\underset{\sim}{n}}'$ is perpendicular to the tangent plane), and the estimated map heights h_i are:

$$h_i \cong H - R_{1,n',i} + [\sqrt{r_E^2 + g_{d,i}^2} - r_E] \qquad \textbf{(4-21)}$$

A height is computed for every point in the scene with the exception of those marked during the phase unwrapping step as lacking reliable phase information.

Baseline rotation methods differ from those operating on the flattened phase, as they do not integrate across the scene. They operate on the unwrapped “raw” interferogram phase, and treat each pixel independently of the others. An estimate of the phase constant ϕ_c is required as the initial condition. The far-field approximation, if taken, results in systematic height errors (tens of metres). Use of cosine law (rather than the far-field approximation) more accurately reproduces the true baseline geometry. In cases with larger orbital divergences, the systematic error increases due to the 2D baseline geometry used. These errors can be mitigated through tiepoint-based baseline tweaking [84].

4.9 Baseline Refinement

Unfortunately, even the best available standard ERS orbit state vectors are not accurate enough to allow a straightforward (i.e. no feedback - “open loop”) transformation from differential phase to topographic heights. The heights of a few known points in the image are

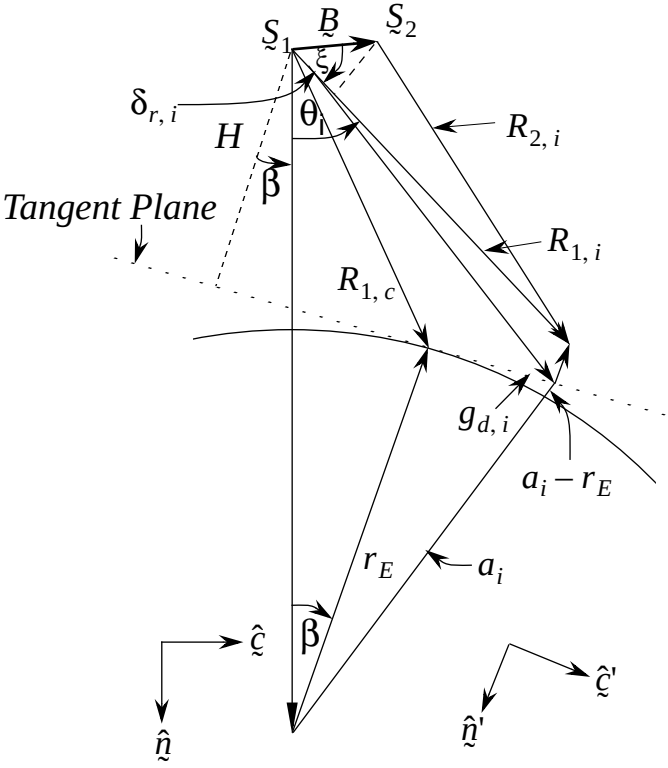


Figure 4-8: Baseline rotation geometry (see Appendix A for description of variables)

required to calibrate the transformation from interferometric phase to topographic height.

The phase unwrapping procedure provides the interferometric phase ϕ_i up to an unknown constant ϕ_c. The slant range difference $\delta_{r,i}$ is then

$$\delta_{r,i} = |\underset{\sim}{R}_2| - |\underset{\sim}{R}_1| = \frac{\lambda}{4\pi}(\phi_i + \phi_c) . \qquad \textbf{(4-22)}$$

Making a simplifying assumption that the tangential component of both the

baseline and the look vector can be neglected, the difference in distance becomes

$$\delta_{r,i} = R_1(\sqrt{f} - 1) \quad \textbf{(4-23)}$$

where

$$f = 1 + \frac{B_c^{\,2}}{R_1^2} + \frac{B_n^2}{R_1^2} - \frac{2}{R_1}[B_n \cos\theta + B_c \sin\theta] \quad \textbf{(4-24)}$$

The corresponding interferometric phase may be modelled geometrically as

$$\phi_i = \frac{4\pi}{\lambda}\left(\sqrt{|\underset{\sim}{R}_1|^2 + |\underset{\sim}{B}|^2 - (2\underset{\sim}{R}_1 \cdot \underset{\sim}{B}_i)} - |\underset{\sim}{R}_1|\right) - \phi_c \quad \textbf{(4-25)}$$

Using a series approximation[1] from Equation 4-25, one can optionally construct a linear model for the phase difference

$$\phi_i = \frac{4\pi}{\lambda} \cdot \frac{\underset{\sim}{R}_1 \cdot \underset{\sim}{B}_i}{|\underset{\sim}{R}_1|} - \phi_c\,. \quad \textbf{(4-26)}$$

Using expressions for the baseline and look vector (Equations 2-10 and 2-11) one then has:

$$\phi_i - \frac{4\pi}{\lambda}\hat{R}_{n,i} B_{n,i} \quad \textbf{(4-27)}$$

$$= \left[\frac{4\pi}{\lambda} \cdot \hat{R}_{c,i} \;\; \frac{4\pi}{\lambda} \cdot \hat{R}_{c,i} t_i \;\; -1.0\right] \begin{bmatrix} B_c \\ \alpha_c \\ \phi_c \end{bmatrix}$$

4.9.1 Refinement

Three tiepoints allow one to create a system of equations that can be inverted to solve for the model parameters $\{B_c, \alpha_c, B_n, \phi_c\}$. The three tiepoints used to solve the system of equations should be well-distributed across the scene to avoid singularities. Improved accuracy can be obtained by using more tiepoints, and performing a non-linear least squares fit using the iterative Levenberg-Marquadt algorithm [67]. The algorithm requires that the first derivatives of parameters to be fit be known.

The sensitivity of the interferometric phase to changes in each of the parameters $\{B_c, \alpha_c, B_n, \phi_c\}$ can be calculated by taking partial derivatives of Equation 4-25 with respect to each

1. equivalent to the far field approximation

parameter. After some derivation, the sensitivities are:

$$\frac{\partial \phi_i}{\partial \phi_c} = -1.0 \qquad \textbf{(4-28)}$$

$$\frac{\partial \phi_i}{\partial B_c} = -\frac{4\pi}{\lambda} \cdot \frac{1}{\sqrt{f_i}} \cdot \left[\frac{\widehat{B_c}}{R_{1,i}} - \sin\theta_i \right] \qquad \textbf{(4-29)}$$

$$\frac{\partial \phi_i}{\partial \alpha_c} = t_i \cdot \frac{\partial \phi_i}{\partial B_c} \qquad \textbf{(4-30)}$$

$$\frac{\partial \phi_i}{\partial B_n} = -\frac{4\pi}{\lambda} \cdot \frac{1}{\sqrt{f_i}} \cdot \left[\frac{B_n}{R_{1,i}} - \cos\theta_i \right] \qquad \textbf{(4-31)}$$

4.9.2 Baseline Quality

Five baseline models were tested for their accuracy using 18 tiepoints distributed across the scene. Tiepoints for an interferometric pair acquired over an area near Bonn, Germany were taken from topographic maps produced by the German *Landesvermessungsamt Nordrhein-Westfalen* (1981).

The five models were:

- Open-loop solution using baselines calculated straight from the orbit state vectors in the SAR data product header.
- Non-linear least squares fit solution for ϕ_c alone; baseline $\{B_c, \alpha_c, B_n\}$ from state vectors.
- Linear solution to equation Equation 4-26 from three tiepoints.
- Non-linear least squares fit to $\{B_c, \alpha_c, \phi_c\}$; B_n from state vectors.
- Non-linear least squares fit to 4 parameters $\{B_c, \alpha_c, B_n, \phi_c\}$.

Tiepoints are required for all but the first model.

For each model, Table 4-1 reports the mean square error between the estimated height and that read from the map for the Bonn interferometric pair from March, 1992. Multiparameter fits perform significantly better than the open-loop or one-parameter (ϕ_c) models. In the four parameter case, ϕ_c and B_n are not independent, causing singularity.

For all but the singular four parameter least squares fit, baseline values wander less than a metre from the orbital state vectors.

	Open-loop	Phase Constant	Linear Approx.	Non-linear 3 param.	Non-linear 4 param.
Parameters Fit	none	ϕ_c	B_c, α_c, ϕ_c	B_c, α_c, ϕ_c	B_c, α_c, B_n, ϕ_c
# tiepoints	0	18	3	18	18
B_c [m]	-158.95	-158.95	-158.18	-158.63	-156.94
α_c [m/s]	.1961	.1961	.1867	.1795	.1793
B_n [m]	-45.54	-45.54	-45.54	-45.54	-41.40
$\lambda\phi_c/4\pi$ [m]	0.0	-0.055	.6274	.4499	4.9176
RMS Error [m]	59	54	16	8.2	8.2

Table 4-1: Quality of baseline models

Chapter 5

SAR Image Simulation

5.1 Introduction

Many of the techniques developed for the geocoding of InSAR information are also useful for SAR image simulation. Such image simulation can be used to correct a radiometric signal σ^0 for influences from topography. However, it is also useful as a way to illustrate the relationship between the radar acquisition geometry and the map geometries that are more familiar to geographers. This chapter uses image simulation techniques to illustrate that relationship, and to demonstrate an application of the digital elevation models produced via InSAR: the geometric and radiometric calibration of SAR images [79].

Multi-mode intercomparison: The RADARSAT satellite has a more flexible design than ERS-1, allowing a choice of incidence angles, as well as three different chirp bandwidths. However, the greater variety of modes made possible by the increased flexibility has a price. Intercomparison of images collected at different times is complicated by the variable incidence angles (and backscatter properties). Given an elevation model, an image simulation model can be used to reproduce the expected backscatter pattern. Each image can be calibrated using the simulated backscatter modelled using its geometry. Intercomparison of multi-mode data sets becomes more practical.

5.2 Image Simulation

The SAR backscatter return within a single pixel is a sum of the contributions from all terrain locations within the confines of a volume bound by range and Doppler frequency equations. Within a single Doppler plane, Figure 4-5 makes clear that some radar ranges may have no backscatter contributed (radar shadow), while others may be composed of contributions from a very large area in map geometry (e.g. layover). In the latter case, multiple map (or DEM) geometry pixels contribute to a single slant range radar geometry pixel. This can be described as a *many-to-one* transformation, as many map-geometry locations contribute to the backscatter at a single radar geometry coordinate.

The SAR image simulator implementation described here proceeds sequentially through the DEM area

under study, first initializing ground area to zero for all range / azimuth locations:

$$\forall r, a \;\; A(r, a) = 0 . \qquad \textbf{(5-1)}$$

One forms facets from four adjacent pixels, and computes the local incidence angle at each point, as well as the local illuminated area; both of these values are then transformed into the SAR ground range geometry. The local incidence angle value helps improve the local (radar geometry) estimate of the *mean* local incidence angle. After confirming that the current map geometry pixel is not in radar shadow, one knows that it provides an illuminated area that contributes to the backscatter somewhere in the SAR image geometry. The area estimate is added to that already contributed from other map-geometry pixels at that radar image coordinate.

$$A(r, a) = A(r, a) + A_{local} \qquad \textbf{(5-2)}$$

where A_{local} is the surface area of the local map geometry pixel. Note that this requires either an image blocking procedure or a large memory size, as random access is required to the running sum of local illuminated area $A(r, a)$ (in SAR geometry).

Depending upon the relation between the DEM and radar resolutions, the interpolation method used to resample from map to radar image geometry can become important. Bilinear interpolation uses the appropriate weighting to distribute the contribution of the local illuminated area into four adjacent radar geometry pixels. Oversampling may also be necessary if the DEM resolution is coarser than that of the SAR image.

5.3 Map vs. Radar Geometry

Figure 5-1 illustrates how more than one pixel within the DEM can contribute to the backscatter at a single location in range Doppler radar geometry. Although both the map and range Doppler grids are spaced at the same nominal pixel resolution, radar foreshortening and layover effects cause more than one map geometry pixel to contribute to the backscatter within a single radar geometry grid cell. Particularly in layover areas, even tens of pixels can contribute to a single slant range location. In short, one can say that in such cases a *many-to-one* transformation takes place between map and radar geometry, as *many* map geometry pixels contribute to the content of *one* radar geometry pixel. Note that this transformation applies to all radar information - being a part of the image acquisition process itself, it applies to backscatter, interferometric phase, and coherence.

Both radiometric calibration and height model generation using InSAR need to take this into account. In the height model generation case, the effect restricts the achievable resolution in highly sloped areas. Using spaceborne data, one must counteract this using combination of ascending and descending data sets [83].

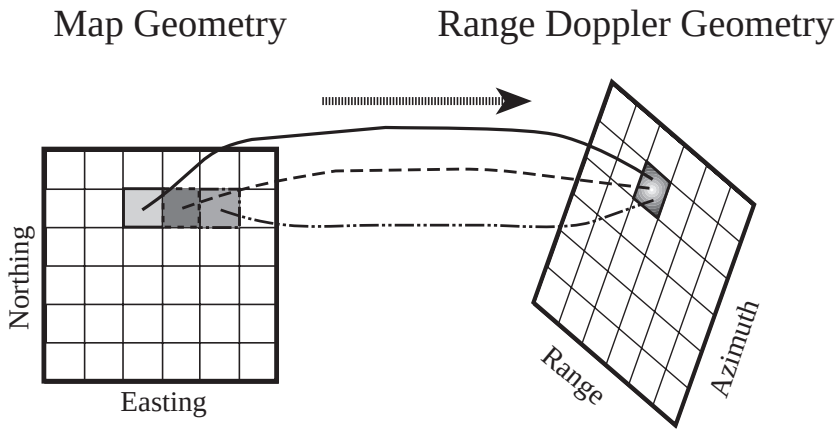


Figure 5-1: Relationship between radar and map geometries

Within the following few pages, radiometric *normalization* is used to illustrate the effect, as the radar brightness measured by SAR sensors is more directly related to the simulated illuminated area.

5.4 Radiometric Normalization

The amplitude values retrieved from the SAR are related to the local area illuminated by the radar beam. A terrain-based SAR image simulation relates the backscatter to the local area:

$$\sigma^0_{simulated} \sim A(r, a)\,. \qquad \textbf{(5-3)}$$

Area estimates from terrain facets [4] calculated using the reference DEM were summed sequentially, and output in the SAR ground range geometry. Given the local area estimates, the backscatter coefficient may then be calculated as

$$\sigma^0(\theta) = \frac{\bar{I}}{K} \cdot \frac{A_0}{A(r, a)} \qquad \textbf{(5-4)}$$

where σ^0 is the backscatter coefficient, $\bar{I}$ the pixel intensity, K a calibration constant, θ the local incidence angle, $A(r, a)$ the local ground scattering area, and A_0 the reference ground scattering area.

The local area estimates provided by the SAR image simulation are used to normalize each ground-range pixel with respect to a reference area.

Radiometric correction often also requires consideration of the antenna pattern variation over the swath. For certain SAR geometries (e.g. SIR-C, wide swath ScanSAR) the local terrain height should also be considered during this correction. However, for single-beam RADARSAT scenes, the high altitude of the satellite together with the relatively small range of incidence angles and the “flatness” of the antenna pattern within a single beam swath reduces the magnitude of such radiometric errors to one or two dB. They are therefore neglected here.

In addition to the local illuminated area, truly robust radiometric calibration of the radar backscatter should incorporate correction for local incidence angle and antenna pattern distribution effects.

5.5 RADARSAT Example

Two data sets provided through the Canadian Space Agency’s RADARSAT Application Development and Research Opportunity (ADRO) programme were used to illustrate the method. The test area is situated in the area surrounding Zürich, Switzerland (see Figure 5-2). A DEM with an original 25 metre grid size, the *DHM25* height model from the Swiss Federal Office of Topography, was used.

Images were geometrically and radiometrically calibrated using the methods described above.

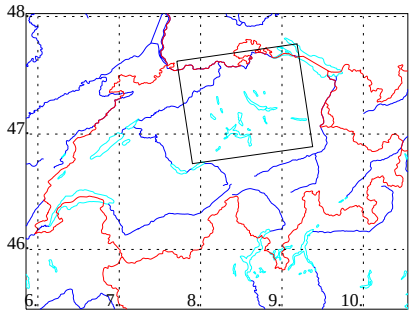


Figure 5-2: Overview of RADARSAT test site Zürich

For the ascending scene, the original ground range SAR image is shown in Figure 5-3, which may be compared to the concomitant image simulation in Figure 5-4. The mean local incidence angle image (in SAR ground range geometry) calculated as a by-product of the image simulation is shown in Figure 5-5.

Note the strong layover and radar shadow effects in the lower right corner. Comparisons between the image simulations and their corresponding true RADARSAT images show that the strong returns from layover are modelled well. Shadowed areas are also reproduced realistically.

After radiometric correction through normalization for the local pixel area, the calibrated image (see Figure 5-6) appears “flatter”, with most topography-

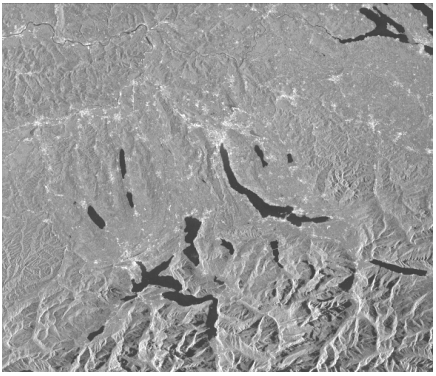


Figure 5-3: RADARSAT standard beam 7 ground range image

Zürich, Switzerland; August 30, 1996, orbit 4290, ascending, right-looking, $\theta = 47°$, resolution: 25×25m, displayed pixel spacing: 192×192m, image area 112×99km

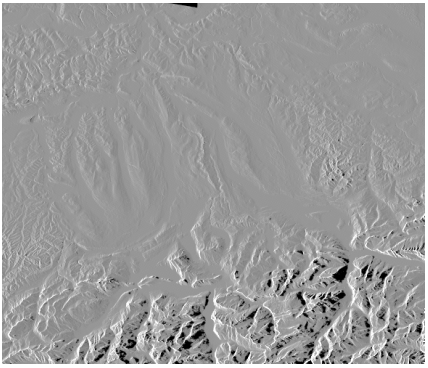


Figure 5-4: Simulation of RADARSAT standard beam 7 ground range image

Calculated using DHM25, courtesy Swiss Federal Office of Topography
Zürich, Switzerland; August 30, 1996, orbit 4290, ascending, right-looking, $\theta = 47°$, displayed pixel spacing: 192×192m

induced distortions (exception: radar shadow) substantially reduced. In the case of radar shadow, there is no signal to normalize, and no improvement is possible. However even extreme layover can be handled using the multifaceted image simulation, as an accurate normalization factor may be calculated from the reference DEM.

Radiometric normalization amplifies thematic information that relate to radar backscatter differences (low backscatter from water bodies, strong double-bounce returns from urban areas). The difference in backscatter values before and after radiometric correction can be compared in terrain-geocoded maps shown in Figure 5-7 (no correction) and Figure 5-8 (with correction).

Intercomparison between mixed mode data (differing incidence angles, ascending/descending, etc.) is made easier without the distraction of topography-induced distortions.

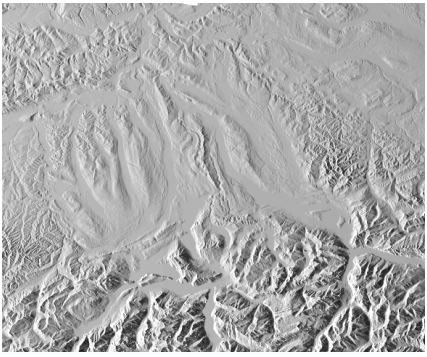


Figure 5-5: Local incidence angles for RADARSAT standard beam 7 ground range image

Calculated using DHM25, courtesy Swiss Federal Office of Topography
Zürich, Switzerland; August 30, 1996, orbit 4290, ascending, right-looking, $\theta = 47°$, displayed pixel spacing: 192×192m, image area 112×99km

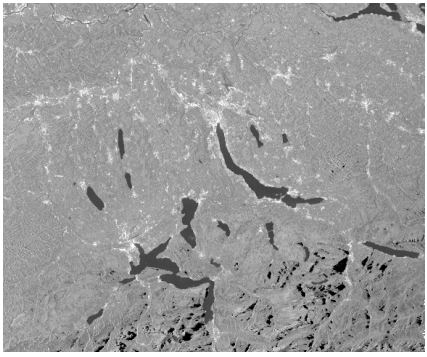


Figure 5-6: RADARSAT standard beam 7 terrain-geocoded radiometrically corrected image

Zürich, Switzerland; August 30, 1996, orbit 4290, ascending, right-looking, $\theta = 47°$, resolution: 25×25m, displayed pixel spacing: 200×200m

5.6 Discussion

SAR image simulation is an effective tool for illustrating the *many-to-one* nature of the relationship between map and radar geometry in steeply sloped areas. The loss of resolution can only be counteracted through integration with data from other viewing angles (e.g. combination of ascending / descending scenes) [37].

Image simulation is also useful for mission planning, more important than ever with RADARSAT's wide palette of available modes.

SAR image simulation is useful for mission planning as an aid in estimating the amount of layover and/or shadow to be expected, for automatic tiepointing, but also for radiometric calibration.

Radiometric normalization for the local illuminated pixel area removes distractions from topography-induced distortions. Even layover areas can be satisfactorily normalized - unfortunately, with no improvement to their poor local resolution in map geometry. However, no radiometric improvement is to be found in radar shadow areas, as there is no signal to normalize in such cases.

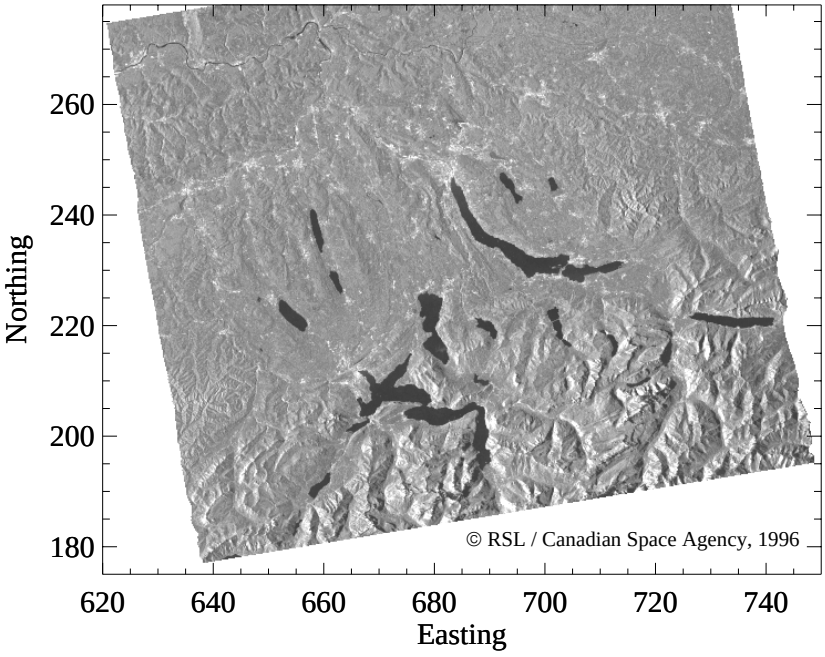


Figure 5-7: Terrain-geocoded RADARSAT standard beam 7 ground range image
Zürich, Switzerland; August 30, 1996, orbit 4290, ascending, right-looking, $\theta = 47°$, resolution: 25×25m, displayed pixel spacing: 200×200m

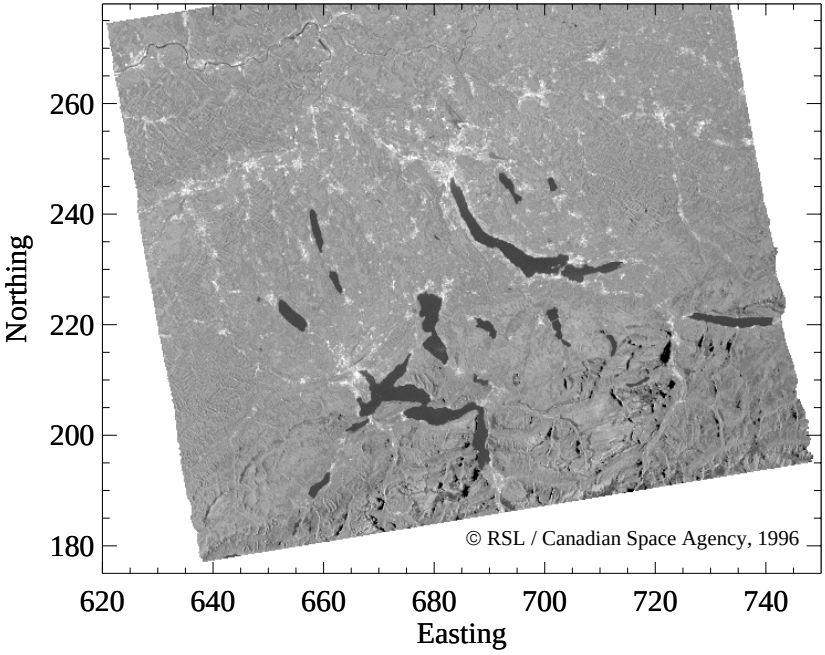


Figure 5-8: Radiometrically corrected RADARSAT standard beam 7 ground range image

Zürich, Switzerland; August 30, 1996, orbit 4290, ascending, right-looking, $\theta = 47°$, resolution: 25×25m, displayed pixel spacing: 200×200m

Chapter 6

Height Model Validation

6.1 Introduction

Many papers have been published in recent years establishing improved techniques for the calculation of interferograms (the subject matter of Chapter 3). The phase to height conversion and geocoding steps (Chapter 4) have received increasing attention as well. However, the final stage of analysis, the validation of the height models generated through InSAR, has been at times relatively neglected [41]. This was often understandably a consequence of results from intermediate processing steps being available for publication before that of the "end of the processing line" DEM product.

For example, thematic interpretation of interferometric coherence images requires no significant computation beyond interferogram formation (N.B. no phase unwrapping), and has been the subject of many papers.

The work presented here aims to redress the relative lack of results evaluating the final product of the DEM generation process: the geocoded DEM itself.

This chapter evaluates the accuracy of the height accuracy estimates, and investigates their error sources.

6.2 Height Model Validation Methods

The accuracy achievable using the ERS InSAR system for height model generation was estimated using three different approaches, summarized in Table 6-1.

Comparison via a synthetic interferogram avoids the need for phase unwrapping, but requires a high quality reference elevation model, and is therefore not a "blind" test of end-to-end system performance.

Backward geocoding enables direct comparison of height models in the final map geometry, but it is also not a "blind" test of end-to-end InSAR elevation model generation, as the quality of the geocoding is non-representative.

Only a "blind" forward geocoding from slant range into map geometry, with reference heights introduced exclusively at the final height comparison, enables a true end-to-end validation of the InSAR height model generation process.

The relative success of the forward geocoding is tested through overlay of the ERS intensity images in map geometry, geocoded using the forward and backward methodologies. Areas where features do not overlap indicate disa-

Method	Geometry	Advantages	Disadvantages
Synthetic Interferogram	Slant Range	• No phase unwrapping	• Comparison not in map coordinates • Not “blind”
Backward Geocoding	Map Coordinates	• Supplies quick feedback on coherence, local incidence angle, and height model error interdependencies • Allows use of refined height model assembled from heterogeneous sources	• Uses reference height values • Not “blind”
Forward Geocoding	Map Coordinates	• “Blind” geocoding • Reference height model is not “touched” until final comparison	• Geolocation of areas with poor phase information will be erroneous, and may not register well with ground truth information

Table 6-1: Height model validation methods

greement in the height models (resulting in planimetric shifts during geocoding).

Validation of the height models is carried out principally through calculation of the difference between the forward geocoded height model and the reference model. Root mean square (RMS) difference values together with histograms of the distribution of the height differences provide quality measures.

The transformation between map and radar geometry undergone during the acquisition of SAR imagery can include system inherent *many-to-one* conversions and is therefore not always reversible. As illustrated in Chapter 5, the local resolution in foreshortened and layover areas can deviate very significantly from nominal values. The influence of the local slope (and also coherence) on the height accuracy is therefore studied.

Finally, consistency checks are performed between several forward-geocoded height models, both to counteract differential (e.g. atmospheric) effects and to increase the height model accuracy.

Two test sites were selected for detailed study. The first site, surrounding Bern (Switzerland), has a large height variation and steep slopes. The second

site, west of Bonn (Germany), was used as an example of an area with moderate topography.

6.3 Test Site Bern

6.3.1 Introduction

The first test site, situated close to Bern, Switzerland, provides an example with moderate to steeply sloped topography, with varying baselines (from the ERS-1/ERS-2 tandem mission) distributed seasonally, as well as 1:25’000 maps, and a high-quality reference DEM.

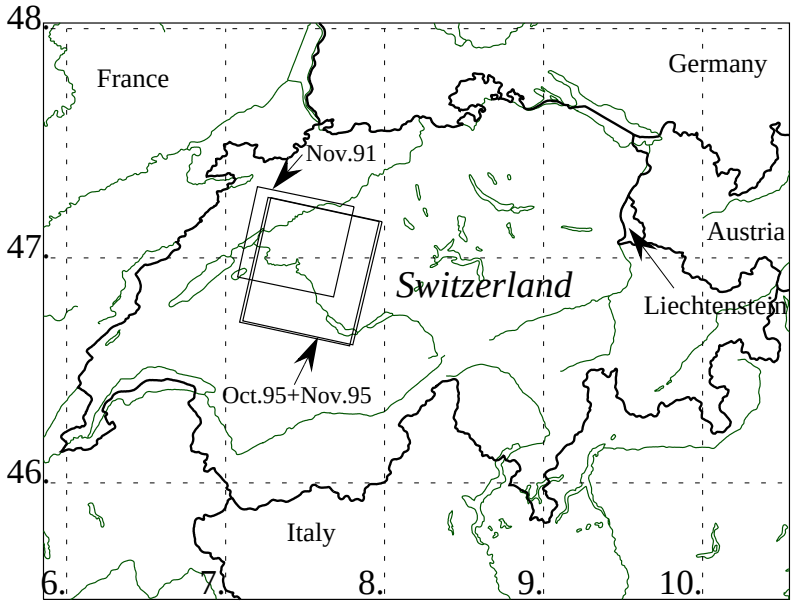


Figure 6-1: Bern, Switzerland - Geographic location of ERS SAR data - Nov. 1991, Oct. 1995, Nov. 1995

Phase	Date	Orbit	Data	$B_\perp$
ERS-1 Commissioning	24.11.91	1868	ERS-1 SLC	58 m
	27.11.91	1911		
ERS-1/2 Tandem	04.06.95	20322	ERS-1 SLC	118
	05.06.95	649	ERS-2 SLC	
	09.07.95	20823	ERS-1 SLC	23
	10.07.95	1150	ERS-2 SLC	
	13.08.95	21324	ERS-1 SLC	43
	14.08.95	1651	ERS-2 SLC	
	18.09.95	2152	ERS-2 SLC	
	22.10.95	22326	ERS-1 SLC	-105
	23.10.95	2653	ERS-2 SLC	
	26.11.95	22827	ERS-1 SLC	-136
	27.11.95	3154	ERS-2 SLC	
	31.12.95	23328	ERS-1 SLC	205
	1.1.96	3655	ERS-2 SLC	

Table 6-2: Bern test site - ERS data - $B_\perp$ = Component of baseline perpendicular to line of sight

Location: Figure 6-1 shows the geographical location of the Bern test site. The area contains both deciduous and coniferous forests, arable land, meadows, pastures, and some densely populated areas. Isolated bare rock and alpine pastures are present near the southern edge of the images. Terrain height ranges from 400m to 2400m within the Oct. 1995 / Nov. 1995 frames shown in Figure 6-1.

ERS Data: Table 6-2 lists the ERS-1 and ERS-2 data available for the Bern test site. All data were acquired during descending passes, and were delivered in standard ESA SLC quarter-scene format. A single pair from 1991, and a set of six tandem pairs were acquired from 1995.

Two one-day repeat pairs (October and November) with above average coherence and height sensitivity were selected for more detailed study. The more sensitive 205m baseline of the 95.12.31 / 96.01.01 pair was unfortunately unavailable, as melting and heavy precipitation destroyed almost all coherence in the 24 hour interval between acquisitions.

Inspection of Figure 6-1 shows that the Oct. 1995 / Nov. 1995 frames do not completely overlap with the Nov. 1991 quarter scene boundaries. However, the common area is sufficiently large to allow comparisons.

Reference Data: Table 6-3 lists the ancillary reference data available for the Bern test site. The Swiss reference digital elevation model *DHM25* has an advertised accuracy within the test area of approximately 3m. ERS precise orbit "PRC" products were acquired from the D-PAF to assist in baseline estimation.

A shaded relief model calculated from the DHM25 reference height model is shown in Figure 6-2. The Jura mountain range is visible in the northwest; the Central Alps of the Bernese Oberland begin in the southeast.

Data	Source
Bern reference DEM "*DHM25*" • Pixel spacing 25m • Heights quantized to integer decimetres	Swiss Federal Office of Topography
Meteorological data • hourly measurements of temperature, humidity, and precipitation	Swiss Meteorological Institute
ERS-1/2 precise orbits	D-PAF
Land use statistics 1979/85 • Pixel spacing 100m	Swiss Federal Statistical Office
Tree height reference • Pixel spacing 500m	Swiss Federal Institute for Forest, Snow, and Landscape Research

Table 6-3: Bern test site - Ancillary data

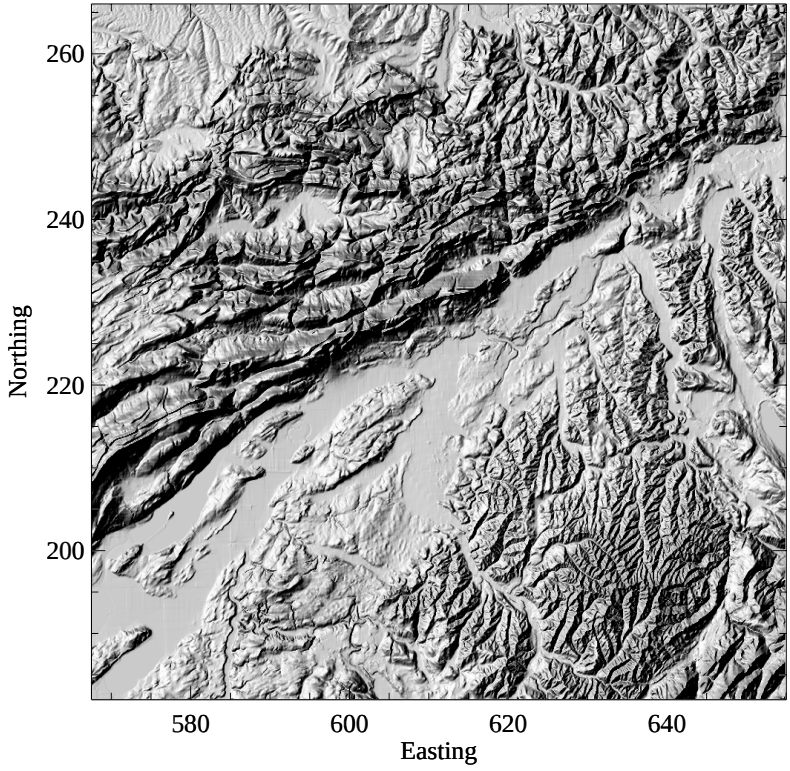


Figure 6-2: Bern - Shaded relief calculated from reference digital elevation model *DHM25* - data courtesy: © Swiss Federal Office of Topography

Hourly weather station measurements acquired from the Swiss Meteorological Institute were also available for the Bern test site.

6.3.2 Data Selection

Bern was chosen as a test site to investigate the influence of a variety of slopes as well as vegetation states on the DEM generation process. Within the tandem data inventory (see Table 6-2) the

Oct. 1995 and Nov. 1995 pairs were selected as having the best combination of simultaneously high height sensitivity and coherence.

A three-day repeat pair from Nov. 1991 was selected together with two ERS tandem pairs (Oct. 1995 and Nov. 1995 - each with a 1-day interval) for more detailed study. All three pairs exhibited above-average repeat-pass coherence for the Bern study area. Figure 6-3 shows the slant range coherence images for the two pairs. Mean coherence (including water bodies) was approximately 0.5 for both pairs.

Global coherence histograms are shown in Figure 6-4. The November pair generally shows a slightly higher coherence; however, the October pair shows stronger coherence values within some forested areas.

Meteorological data from the Bern-Liebefeld weather station, provided courtesy of the Swiss Meteorological Institute (the German name is *Schweizerische Meteorologische Anstalt* "*SMA*"), are shown in Figure 6-5. For both 1995 interferograms, no precipitation occurred in the tandem interval between the ERS-1/ERS-2 acquisitions, which took place at approximately 11:22am local time. Both humidity and temperature were comparatively stable in the interval between the two November acquisitions: this may account for the slightly higher global coherence evident in Figures 6-3 and 6-4.

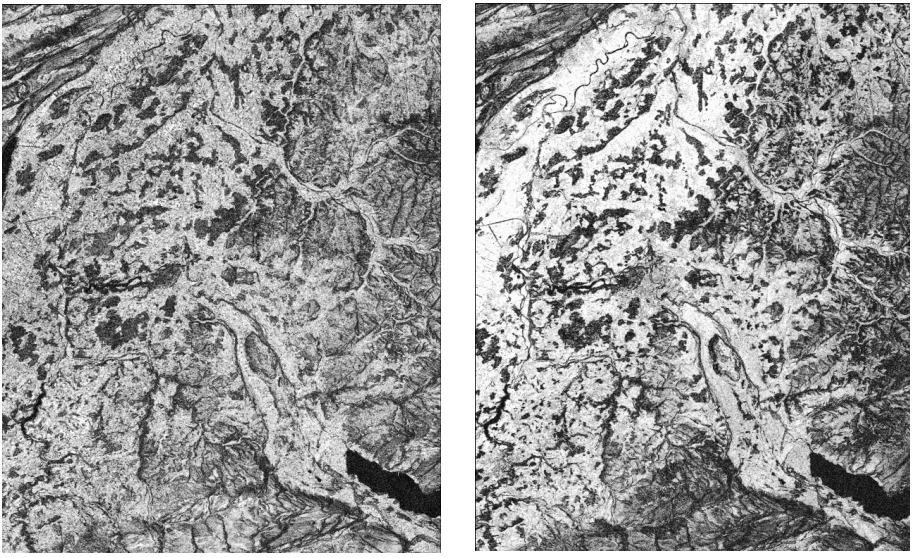
Results from other investigations often select a small sub-scene to showcase the optimal results achievable using the ERS tandem InSAR configuration [44]. However, since systematic error sources often appear only over larger areas, results for the full SLC quarter scene processed are shown here, with selected extracts presenting more detail where required.

The extreme topography in the Bern scene (in comparison to the Bonn area) and numerous hilly forest areas make phase unwrapping more challenging. Larger height errors are the result. Some areas are rendered inaccessible to phase unwrapping.

6.3.3 DEM Flattening

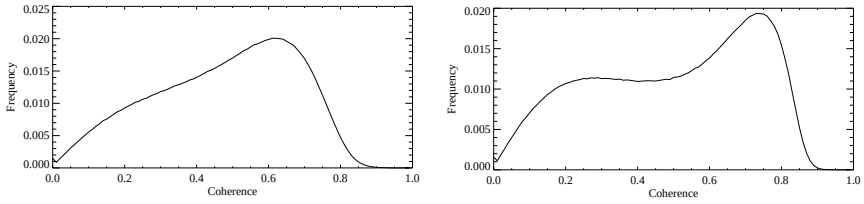
DEM flattening was performed on the interferograms to investigate the consistency of their phase behaviours, and the potential accuracy achievable (assuming a successful phase unwrapping algorithm). Figure 6-6 shows the Bern tandem interferograms after flattening with the DHM25 reference elevation model. One sees that both interferograms are truly flat. Systematic global phase trends are no longer apparent.

Lakes, rivers, and layover regions exhibit random phase values (e.g. Lake Thun in the southeast). Most other areas show remarkably flat phase values. An exception is found in forest stands within the scene (discussed below). Small residual misregistrations between



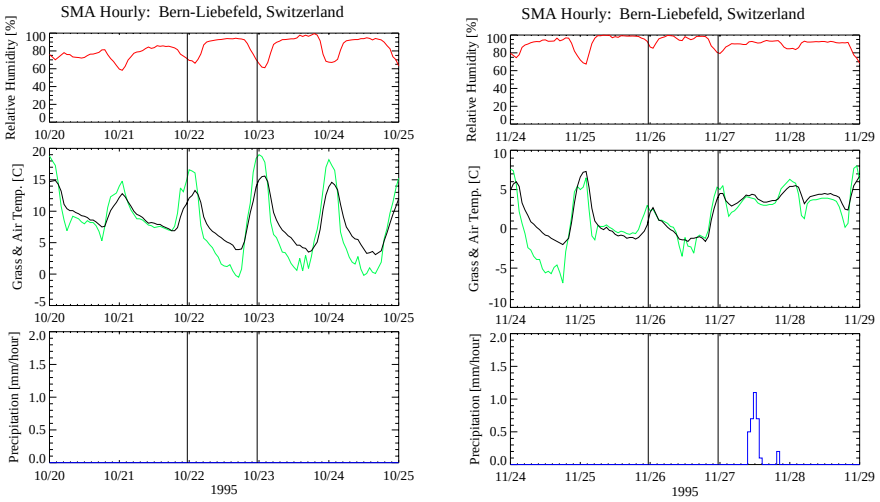
(a) (b)

Figure 6-3: ERS 1-day repeat tandem coherence - Bern, Switzerland - black=0, white=1 (a) Oct. 1995, (b) Nov. 1995



(a) (b)

Figure 6-4: ERS tandem coherence global quarter scene histograms - Bern, Switzerland (a) Oct. 1995, (b) Nov. 1995



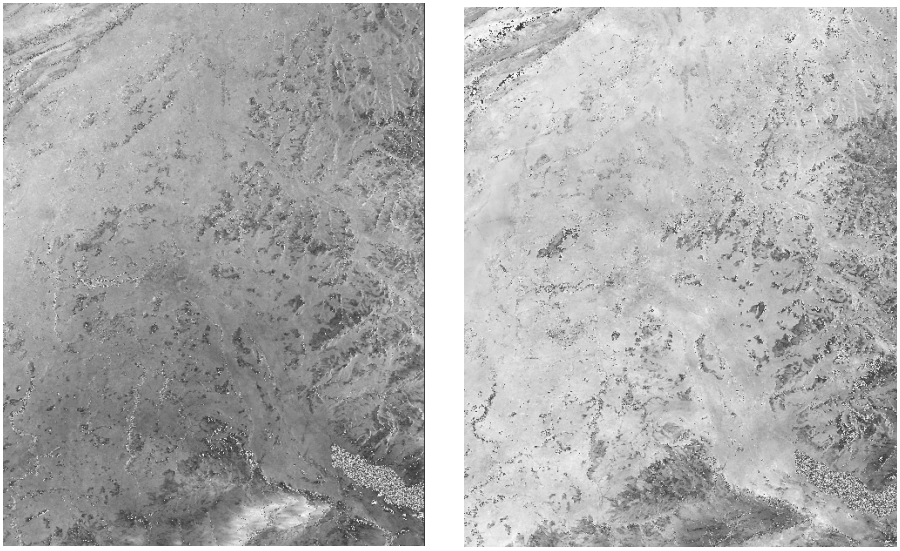
(a) (b)

Figure 6-5: Meteorological data from Swiss Meteorological Institute - Days are centred at noon, vertical lines indicate acquisition times - Bern, Switzerland (a) Oct. 22-23, 1995, (b) Nov. 26-27, 1995

the DEM and the slant range geometry can also result in local phase ramps.

Figure 6-7 shows an extract from the Nov. 1995 DEM-flattened interferogram displayed in Figure 6-6 juxtaposed with a scanned section from a Swiss topographic map. Note how the forest stands are clearly distinguishable in the phase image. The behaviour is consistent for the three Bern interferograms (Nov. 1991, Oct. 1995, and Nov. 1995). The phase difference can be explained by the fact that the reference height model *DHM25* (produced through digitizing map contour lines derived through aerial stereo photography) is based on ground height measurements, while the C-band interferometric height estimates are based upon radar returns echoed from the tree canopy.

Again considering the full DEM-flattened interferograms (Figure 6-6), a phase trend is visible in the south of the image, southwest of the city of Thun, where the terrain climbs quickly from 500m to 2200m. The trend might be



(a) (b)

Figure 6-6: Bern, Switzerland - DHM25-flattened interferogram phase - black=0, white=2π (a) Oct. 1995, (b) Nov. 1995

caused by differing atmospheric path lengths, or by forest stands on the mountain, but it could also be due to a systematic height-dependent bias in the DEM-flattening process. No clear conclusion is possible at this time.

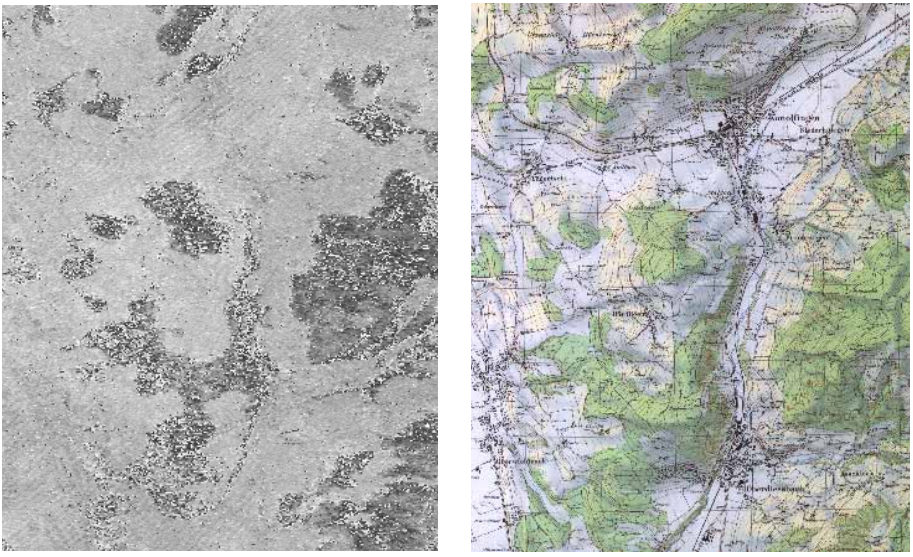
Surface vs. Elevation: The phase residuals caused by the differing references (ground level in the reference *DHM25*, tree branches in digital surface model *DSM*) were studied in more depth, using the effect of tree heights for illustration.

A test area within the Bern tandem October 1995 DEM-flattened interferogram was selected for detailed measurements. Using a 100m map of Swiss land use statistics from 1979/85 (resampled to 25m) to demarcate forest/non-forest classes, together with an erosion operator on the class map to reduce boundary effects, phase differences were translated via the local ambiguity height to a mean height difference of 12m. Local mean phase measurement coherency

was 0.5 for the forest class, and 0.8 for the non-forest class.

This measurement was compared with a 500m tree height raster produced by the Swiss Federal Institute for Forest, Snow, and Landscape Research. These tree height reference measurements were made using aerial photography acquired in 1993. They quantify the mean tree height as approximately 22m within the test area.

The interferometric measurement of forest phase difference is noisy, as evidenced by the lower coherence and histograms of the phase distribution. Its value is also biased at 12m vs. the reference value of 22m. Such a bias is expected due to the higher degree of trunk-scattering in the winter. Summer acquisitions would provide purer canopy returns, but coherence values would sink further still.



(a) (b)

Figure 6-7: Differential phase from forest stands (a) Detail from DEM flattened interferogram - black=0, white=2π - Bern, Nov. 1995 (Figure 6-6) (b) Extract of cartographic data from the National map of Switzerland 1:25’000, courtesy: © Swiss Federal Office of Topography

Due to the noisiness of the measurement, it is not to be recommended as a method for producing maps of tree heights. Comparison with the reference does however help improve understanding of the winter backscatter mechanisms. This is helpful in cataloguing and quantifying the sources of bias that can occur during DEM production.

DEM Improvement: The phase differences seen in Figures 6-6 and 6-7 could be used to improve the reference elevation model by shifting its reference to that used by the ERS sensor [50]. Such an approach avoids some of the errors introduced by phase unwrapping.

However, the scheme has weaknesses when applied to height model validation. Validation usually presumes that one has an independent reference for comparison. If the reference is itself used during processing, it becomes difficult to uncompromisingly validate the end-product against the reference, as the end-product has been tainted through contact. Use of a high-quality DEM colours the “blindness” of the validation process, as the InSAR height estimate loses independence from the reference value.

Use of a low resolution DEM is helpful for scenes with lightly rolling topography, and is less compromising to the blindness of the test. However, phase unwrapping is relatively easy in such conditions, and the Bern scene contains stronger slopes - removal of low frequency topography would be of only marginal benefit.

6.3.4 Incidence Angle

Backward geocoding may be used to quickly validate the interferometric height models, though the forward geocoding methodology gives a more authentic “blind” end-to-end system test of the accuracy of an InSAR-based DEM generator.

Results from such backward-geocoded difference maps are not shown here, as the forward-geocoded results suffice. However, backward geocoding was used for investigation of the influence of local incidence angle (θ) and coherence on height estimation. A consistent geolocation is guaranteed and local incidence angle values are globally available.

The backward geocoding technique was used to generate geocoded local incidence angle and layover maps (see Section 4.4) to analyse their influence on height model accuracy. Figure 6-8 displays a histogram showing the frequency of different local incidence angle values, as distributed over the Oct.1995 scene. As expected, the most frequent values are in the 20-23° range (the scene is a near range quarter), but one sees that the area’s plentiful sloped areas widen the distribution. The distribution within flatter test sites (e.g. the Bonn scene discussed in Section 6.4) is much narrower.

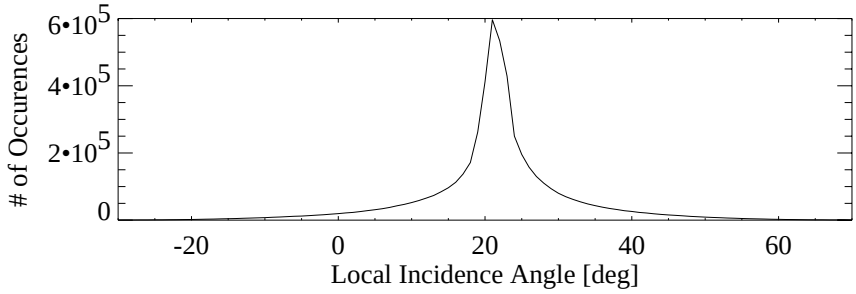


Figure 6-8: Histogram of local incidence angle - Bern, Switzerland - Oct. 1995

Backward geocoding was used to produce a local incidence angle map of the October 1995 scene. The local complement of the incidence angle is shown as a grayscale in Figure 6-9.

6.3.5 Layover

Backward geocoding also allows convenient calculation of layover and shadow maps, as explained in Section 4.4. The layover map corresponding to the Oct. 95 acquisition geometry is shown in Figure 6-10. The maps shown in Figures 6-9 and 6-10 were used to study the effect of local incidence angle (LIA) and layover on coherence and on the accuracy of height estimates.

Figure 6-11 shows the distributions of local incidence angle and of coherence within the layover areas marked in Figure 6-10. As expected, the distribution is centred on an incidence angle of zero degrees, but areas adjacent to the steep slope (but still in layover) widen the distribution. There are even a few examples of areas in layover that exhibit slopes pointing in the opposite direction ($\theta > 23°$). Figure 6-11(b) shows the distribution of coherence within layover areas. Extremely low values are observed, as the wide variety of phase values throughout the layover cell result in a phase return with an essentially uniform distribution. The resulting coherence distribution approaches the lower limit possible [88].

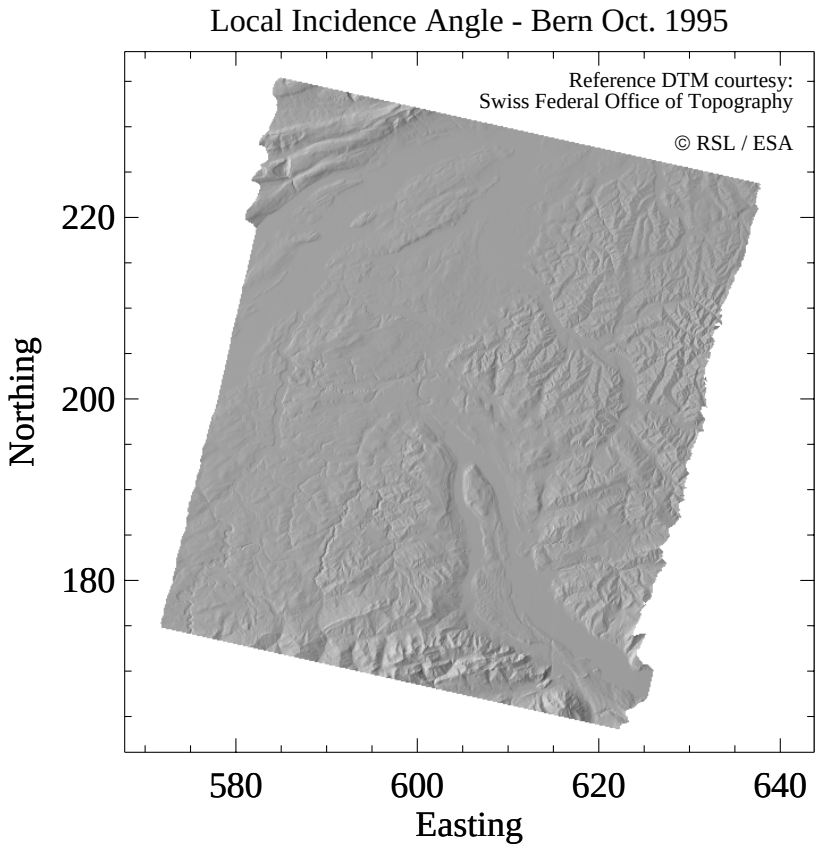


Figure 6-9: Bern, Switzerland - Oct. 1995 - Local incidence angle complement; black=0, white=180° – *DHM25* © Swiss Federal Office of Topography

6.3.6 Coherence

Figure 6-12(a) shows the dependence of coherence on local incidence angle. Coherence is generally highest in flat areas, with an incidence angle of 20-23°. Coherence drops off quickly as one moves from 20° to 0° approaching layo-

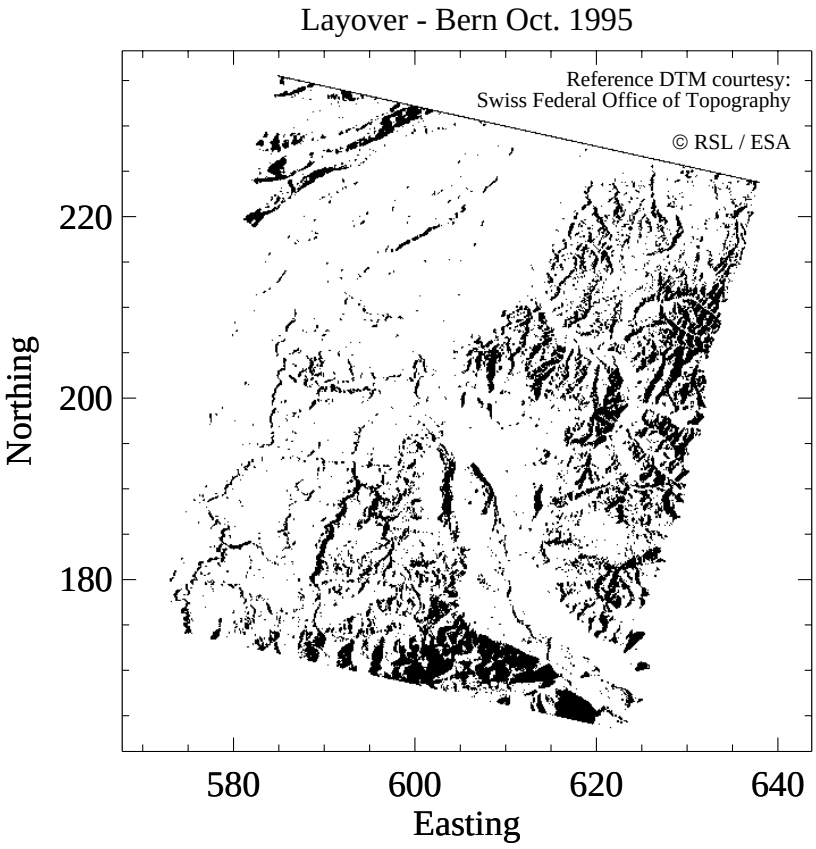


Figure 6-10: Bern, Switzerland - Oct. 1995 - Layover regions marked in black

ver, while in the other direction, for slopes on the *other* side of the mountain, coherence declines more slowly, as the local pixel size (measured in map geometry) becomes smaller. Figure 6-12(b) shows the dependence of the height error (calculated via a back-geocoded comparison with the reference DEM) on the local incidence angle. The trend is oppo-

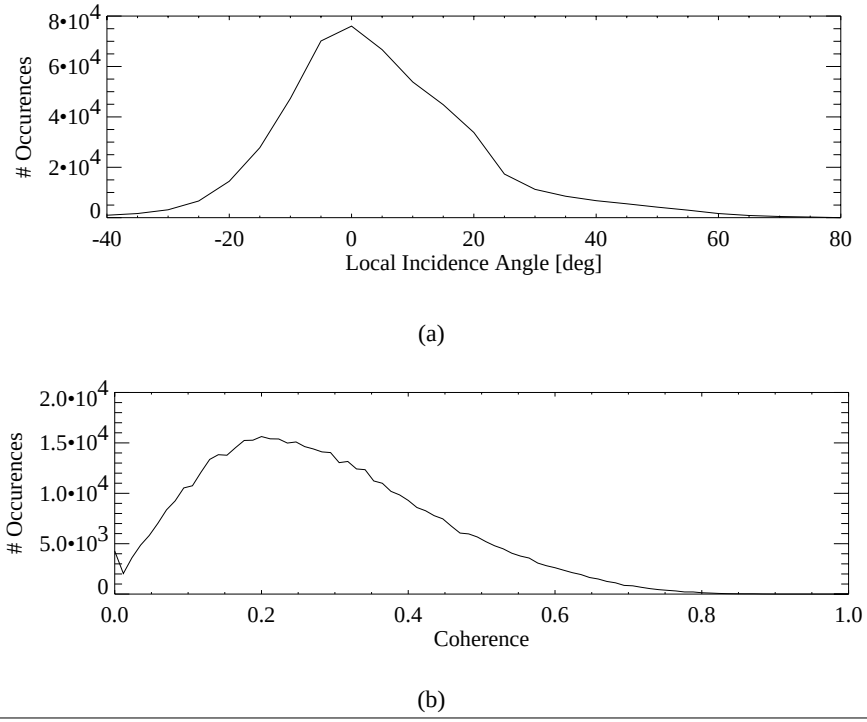


Figure 6-11: Layover - Histograms of local incidence angle and coherence in layover areas - Bern, Switzerland - Oct. 1995 (a) Local incidence angle, (b) Coherence

site that of coherence: as expected, height errors are smallest in flat areas, increasing in regions with stronger slopes. The highest accuracies are achieved in areas with incidence angles in the range 20-23°. Error levels are high - this is a single scene with global phase unwrapping errors still present.

Figure 6-13 shows the dependence of successful height estimation on coherence and local incidence angle, with *successful* defined in this context as being within 10m of the reference. Note that these estimates were only performed on areas that had been unwrapped. The influences of the two factors are there-

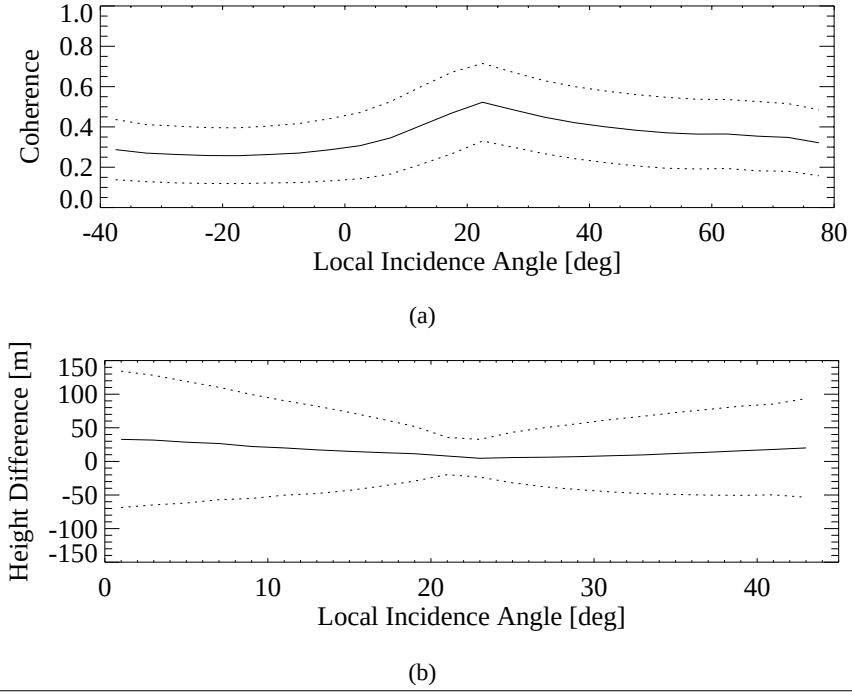


Figure 6-12: Bern, Switzerland - Oct. 1995 (a) Mean coherence ±standard deviation vs. local incidence angle; (b) Mean height error ±standard deviation vs. local incidence angle

fore understated. Note also that each percentage value was calculated within a pixel population satisfying coherence or local incidence angle constraints, i.e. there is no remaining bias for the global distributions of coherence and local incidence angle as shown in Figure 6-8).

Interpreting Figure 6-13(a), one sees that phase unwrapping becomes (as expected) more successful with increasing coherence, and flatter terrain. Slopes facing the radar (boundary layover and beyond) do not unwrap well, while slopes facing *away* present less of a problem. Areas with incidence angles less than zero contain contributions from spatially disparate scatterers that cannot be considered a point source. As a con-

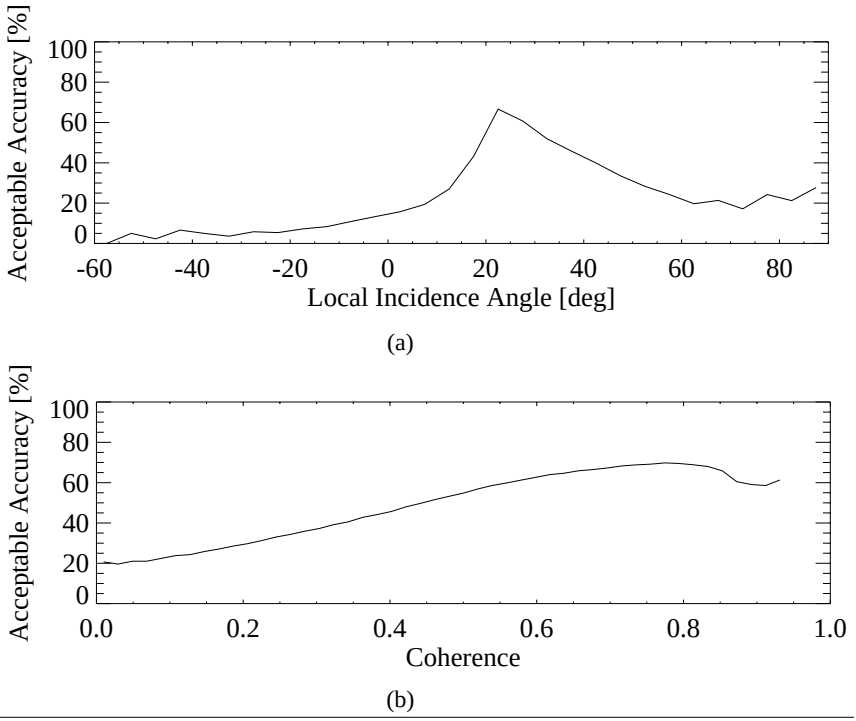


Figure 6-13: Bern, Switzerland - Oct.1995 - Success of height estimation as a function of (a) Local incidence angle, (b) Coherence

sequence, their phase values are noisy, coherence is low, and height accuracy becomes unacceptable.

Figure 6-13(b) confirms that height accuracy generally improves with increasing coherence. The small dip seen for high coherence values is based on a small number of samples, and may be due to the chance placement of a global phase unwrapping error.

6.3.7 Height Differences

The distributions of the height differences are shown in Figure 6-14. Note that phase unwrapping errors produce “islands” of incorrect height values, corrupting the height model. Consistency

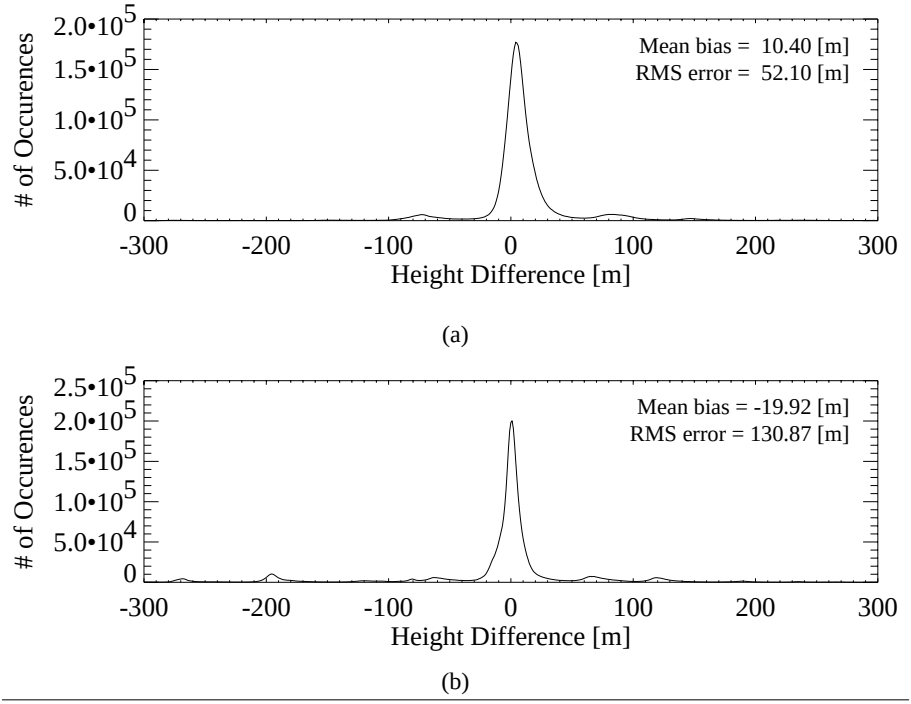
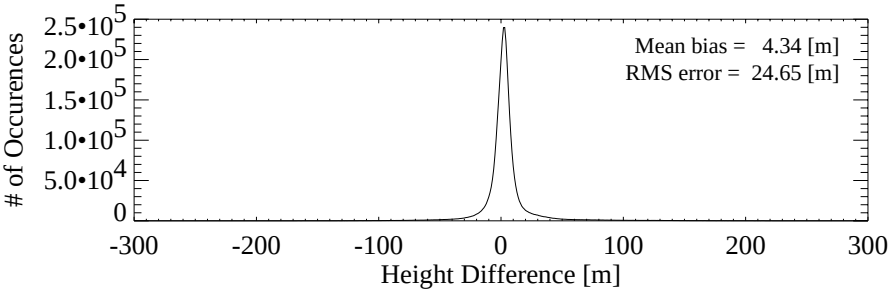


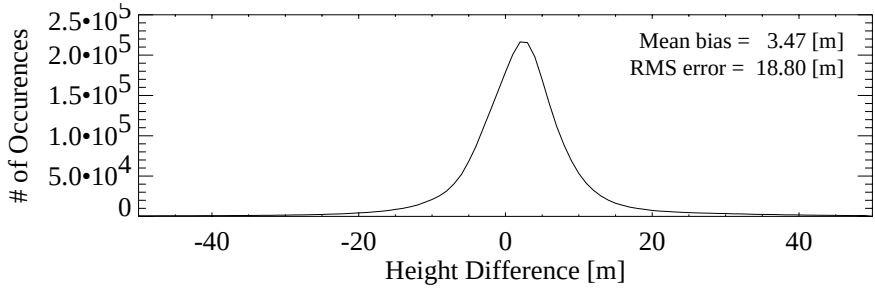
Figure 6-14: Histogram of “blind” forward geocoded height differences within all unwrapped areas - Bern, Switzerland (a) Oct. 1995, (b) Nov.1995

checking using more than a single forward-geocoded height value (e.g. Oct. 95 and Nov. 95) can substantially reduce this error source. The resulting height difference distributions are shown in Figure 6-15. Figure 6-15(a) shows the height difference histogram for all unwrapped areas. If one restricts the area under study to parts of the image with a local incidence angle in the range from 14°–30°, the distribution of height differences improves slightly - see Figure 6-15(b).

Figure 6-16 shows the dependence of the height accuracy achieved in the final end-product height model on the local incidence angle. For local incidence angles ranging from 0 to 45°, the figure shows the mean height error, together with its standard deviation. Note the



(a)



(b)

Figure 6-15: Histogram of “blind” forward geocoded height differences - Bern, Switzerland - Combination DEM’s from Oct. 1995 and Nov. 1995 - (a) All areas, (b) Areas where $14° < \theta < 30°$

improvement in comparison to results from a single interferogram, e.g. Figure 6-12(b). The final height accuracies are respectable: the remaining problem is their lack of omnipresence. Efforts to increase the robustness of the phase unwrapping step are aimed at remedying this problem.

The requirement for reliable phase unwrapping can be the weak link in the production of accurate InSAR height maps. In the absence of completely reliable phase unwrapping, consistency checking using multiple interferograms helps improve the achievable forward geocoded “blind” height accuracy (although it also reduces coverage).

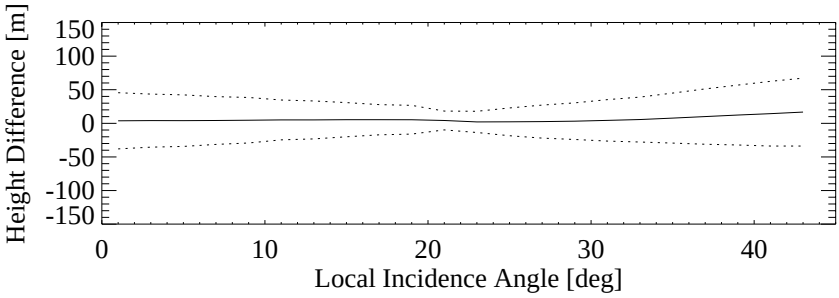


Figure 6-16: Height difference (mean ± standard deviation) vs. local incidence angle - Bern, Switzerland - Combination of two forward geocoded DEM's, Oct. 1995 + Nov. 1995

6.4 Test Site Bonn

The second test site, in the area west of Bonn, Germany, provides an example of a vegetated scene with moderate topography. A variety of baselines and repeat pass intervals are available.

Neither the Bern nor the Bonn test site was chosen to represent the "optimal" potential results from ERS InSAR height model generation. Indeed, the fact that both sites are heavily vegetated significantly affects their repeat-pass coherence values. Results from the sites are instead meant to be representative of the potential of spaceborne repeat-pass C-band InSAR in vegetated areas with hilly (Bonn) or hilly-to-mountainous (Bern) topography. High coherence desert conditions are not studied here - a study of *best case* ERS InSAR would require an arid test site.

6.4.1 Introduction

The Bonn test site provides an example of a vegetated scene with moderate topography. A variety of baselines and repeat pass intervals are available, as the data were acquired during the three-day repeat cycle "first ERS-1 ice phase" orbit configuration [12].

Location: Figure 6-17 shows the location of the Bonn test site. The Rhine river crosses through the northeast corner of the scene. One finds arable land, mixed deciduous and coniferous forests,

Figure 6-17: Bonn, Germany - Geographic location of ERS-1 data - March 14, 1992

open-pit coal mines, gravel pits, and densely populated areas within the region.

ERS Data: Table 6-4 lists the input data available for the study. ERS data were provided in single look complex (SLC) quarter-scene form by ESA. Four ERS-1 acquisitions were available - the associated baselines are listed in Table 6-5. The two interferograms 14.03.92 / 17.03.92 and 14.03.92 / 29.03.92 pro-

Test site	Date	Orbit	Data
Bonn	14.03.92	3459	ERS-1 SLC
	17.03.92	3502	
	20.03.92	3545	
	29.03.92	3674	

Table 6-4: Bonn test site - ERS data

Primary	Secondary	$B_\perp$ [m]
14.03.92	17.03.92	-420
	20.03.92	-921
	29.03.92	130
17.03.92	20.03.92	-502
	29.03.92	550
20.03.92	29.03.92	1052

Table 6-5: Bonn test site - Baselines; $B_\perp$ = Component of baseline perpendicular to line of sight

Data	Source
Bonn Reference DEM • Pixel spacing 12.5m • Heights quantized to integer metres	D-PAF (Permission granted by D-PAF for use of DEM in geocoding research)
ERS-1 Precise Orbits	D-PAF

Table 6-6: Bonn test site - Ancillary data

vide a good variety of baselines and repeat-pass temporal intervals. They were selected for more detailed study.

Reference Data: The ancillary data also available to assist in evaluation of the Bonn test site are listed in Table 6-6. Precise orbit products were ordered from the D-PAF for use in geocoding and baseline estimation. A reference DEM (originally produced by digitizing map sheet contours) was provided by the GEOS group at the D-PAF.

6.4.2 Coherence

As an orientation aid, Figure 6-18 shows the full quarter-scene coherence map corresponding to the March 14-17 pair. The Bonn-Cologne airport is situated in the northeastern corner. In general, high coherence was preserved over the three day interval. As expected, lower coherence is observed on the Rhine river in the northeast, as well as within forested areas both in the southwest corner, and scattered in the east.

6.4.3 Refinement of Geometry

The geometry was refined using tie-points selected from 1:25’000 and 1:50’000 topographic maps produced by the *Landesvermessungsamt Nordrhein-Westfalen*. For baseline refinement, tie-points were selected in flat areas with stable phase values that did not vary within their neighbourhood, resulting in low demands on planimetric accuracy (also within the SAR image). Refinement of the image geometry was performed (as during conventional SAR terrain geocoding) through the use of intensity features (e.g. bridges) that could be localized accurately both in the SAR images and the reference maps.

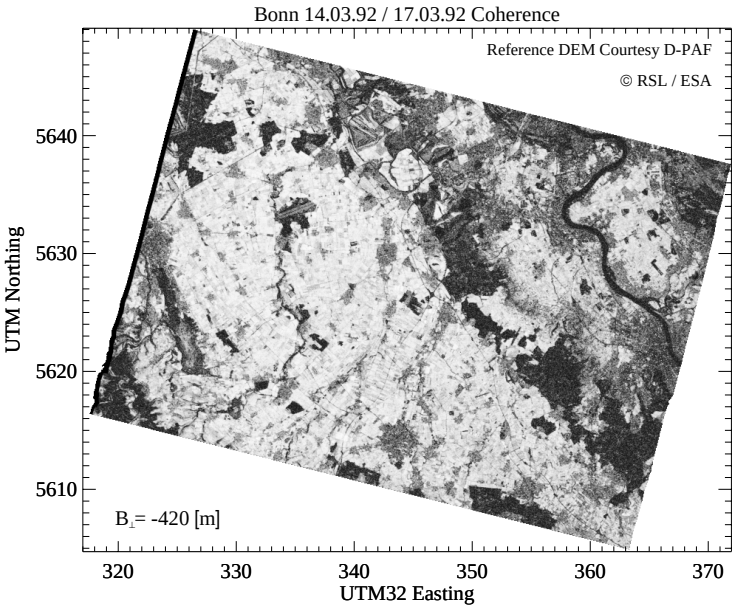


Figure 6-18: ERS-1 coherence map calculated using ERS-1 data from 14 and 17 March 1992. Coherence is shown in grayscale, with black=0, white=1.

6.4.4 DEM Flattening

DEM flattening was used to investigate the potential height accuracy achievable given a correctly unwrapped interferogram. Figure 6-19 shows a 15-day repeat interferogram flattened using an ellipsoid model (without the use of a reference elevation model). One sees both topographic fringes as well as coherent phase changes that took place during the 15 day period between acquisitions (e.g. individual “jumping” fields).

Figure 6-20 shows the DEM-flattened version of the same interferogram. With the exception of areas where the reference elevation model was no longer current (some mining pits, gravel pits), topographic fringes are no longer visible. The fringes that remain are due to coherent phase changes over the 15-day interval - in small areas due to variable scattering depths within fields, but in other areas (middle, top right) likely due to changes in atmospheric path length. Note that the relatively long time inter-

Figure 6-19: Ellipsoid-flattened ERS-1 interferogram - Bonn, Germany 14.03.92 - 29.03.92

val of 15 days increases the interferogram's susceptibility to such differential effects. See Figure 3-22 for a magnified view of the northeast corner.

6.4.5 Backward Geocoding

The interferogram was filtered, the phase was unwrapped, the baseline geometry refined, and the DEM was calculated and geocoded. Figure 6-21 shows the derived height map for an area close to the town of Nörvenich within the March 1992 Bonn scene. Height is shown as a grayscale cycle of length 120m. Note that there is terrain variation of approximately 80m within the region, and (compare with Figure 6-18) relatively high coherence values.

The interferometric heights were compared point-by-point with the reference elevations, and an elevation-difference map was produced. The result is shown qualitatively in Figure 6-22, with the height difference coded as grayscale

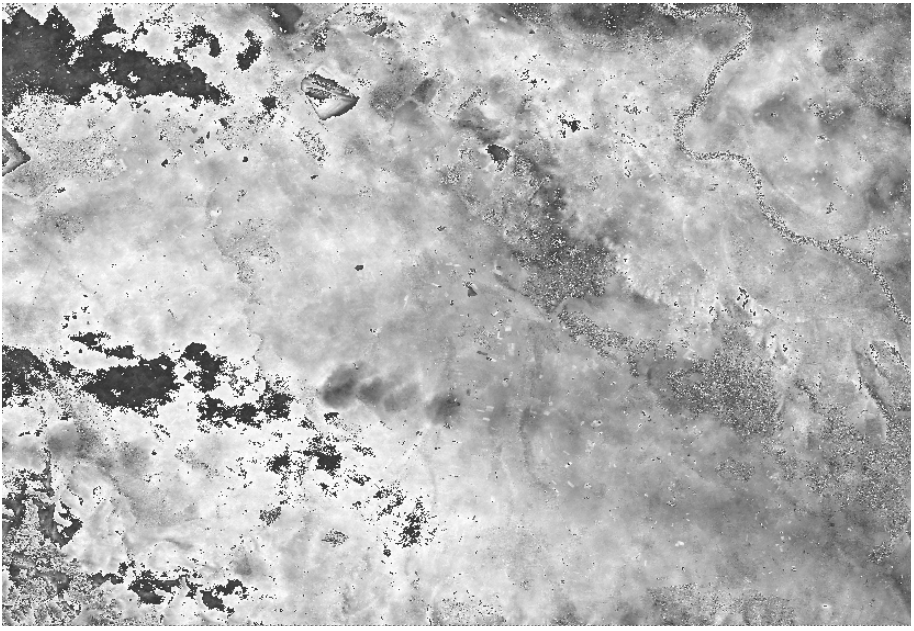


Figure 6-20: DEM-flattened ERS-1 interferogram - Bonn, Germany 14.03.92 - 29.03.92

with saturation maxima at ±10m. Black and dark gray values indicate negative differences (InSAR height < reference). White and light gray values indicate positive differences (InSAR height > reference). Areas with close to zero height difference have a middle-gray value.

One sees that the interferometric height estimates agree remarkably well with those from the reference DEM. Gravel pits are visible at (334,5635), (340, 5631), and (340.5, 5636.5) as negative height difference values. The reference DEM did not register these hypsographic features, causing a significant difference when compared to the height calculated interferometrically.

A one dimensional profile comparing the interferometric height values to the reference DEM is shown in Figure 6-23. Note the generally good agreement over the 150m of height variation along the Easting profile. Figure 6-24 shows a profile at greater scale, accentuating the

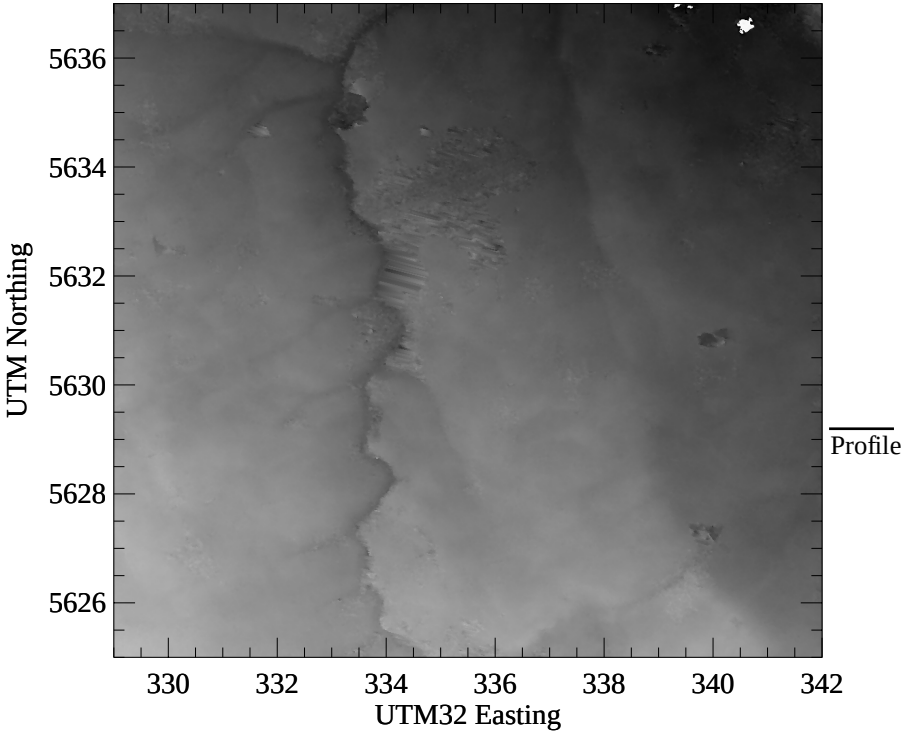


Figure 6-21: Bonn-Nörvenich - ERS-1 InSAR DEM from 14 and 17 March 1992. Height is shown as a grayscale cycle of length 120m; height cycle increasing from black to white - Profile shown in Figure 6-24

behaviour in sloped regions. At this level of detail, the one metre quantization of the reference DEM becomes visible.

6.4.6 Quantitative Results

A more quantitative representation of the height difference distribution is presented in Figure 6-25. Histograms show the frequency of each height difference value within the Bonn-Nörvenich area

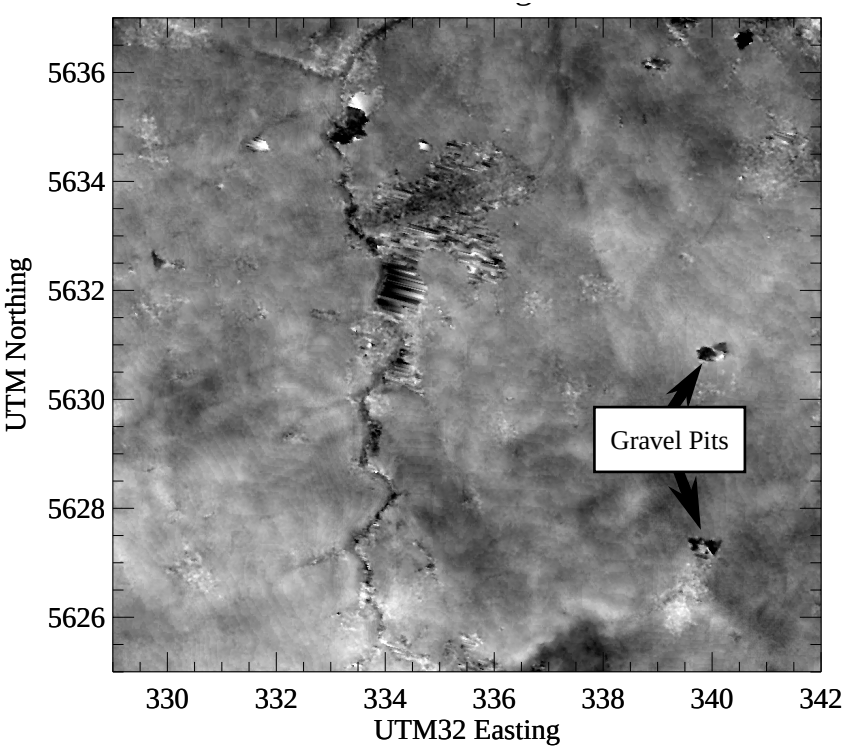


Figure 6-22: Bonn-Nörvenich - Difference between ERS-1 InSAR DEM (14 and 17 March 1992) and reference DEM. Height differences are shown as grayscale with saturation maxima at ±10m. Negative differences tend toward black, positive differences toward white, negligible height differences appear as a middle gray.

displayed in Figure 6-22. Note that all pixels within the region (of all coherences) were used to compute the histograms, and that no significant systematic height bias is visible. A global root mean square (RMS) error of 2.7m was calculated over the 12x13km area for the interferogram calculated from the March 14/17 1992 pair. Considering the fact that the reference DEM is quantized to

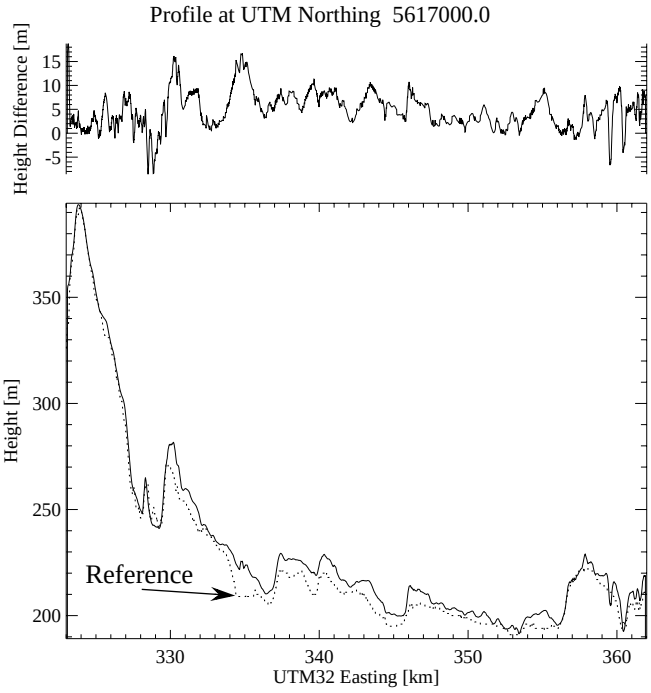


Figure 6-23: Height profile - Bonn 14/17 March 1992 height estimate vs. reference height

1m intervals, this is encouraging for the future of ERS digital elevation model extraction. As seen in Figure 6-26, the height estimates computed using the March 14/29 1992 interferogram were less accurate. The longer time interval both reduced mean coherence over the scene and introduced more coherent phase shifts that distorted the height estimates. The shorter baseline and longer time interval result in an RMS error of about 12m over the Bonn-Nörvenich region.

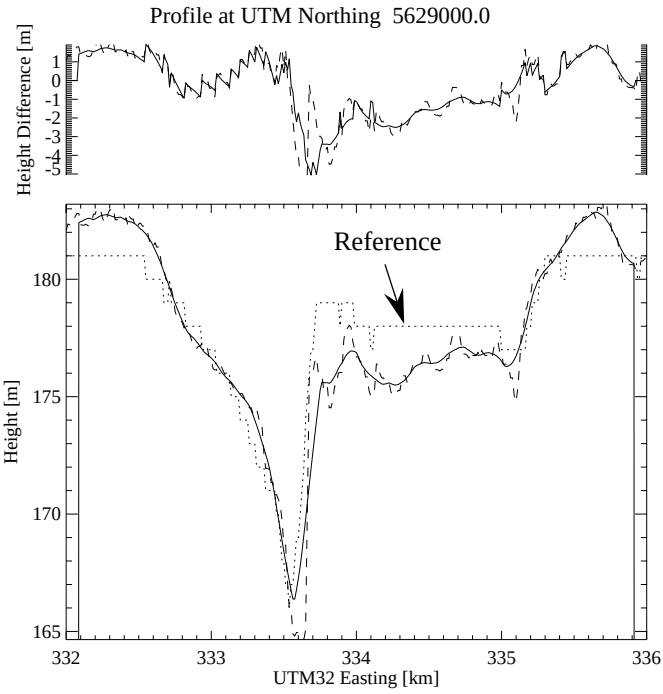


Figure 6-24: Height profile - Bonn-Nörvenich 14/17 March 1992 InSAR height estimate (dotted line) vs. reference height (dashed=raw, solid=filtered)

Two significant factors concerning the Bonn scene are the low mean slope, and the large (highly sensitive) interferometric baseline. Only in such areas with low mean slopes does such a sensitive baseline not impede phase unwrapping. One is able to form accurate phase and height estimates by taking more samples.

6.4.7 Forward Geocoding

The interferometric height model from the Bonn scene was geocoded using both the forward and backward methodologies. For this case, the forward and backward geocoding methodologies were shown to produce results with only minimal differences (see Section 4.6). Planimetric geocoding differences occurred

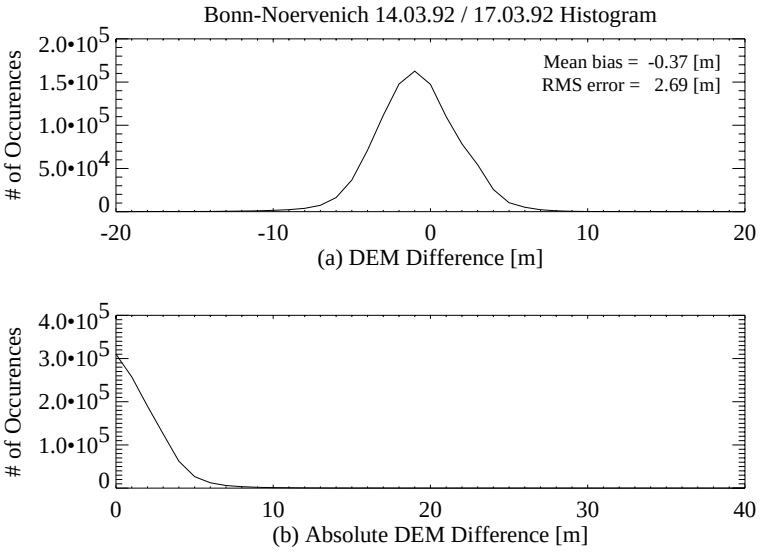


Figure 6-25: Bonn-Nörvenich 14/17 March 1992 - Histogram of height differences between interferometrically calculated heights and reference elevations for all pixels within the Bonn-Nörvenich region; (a) Height difference, (b) Absolute value of height difference

where the interferometric height estimates differ (e.g. gravel pits, forest).

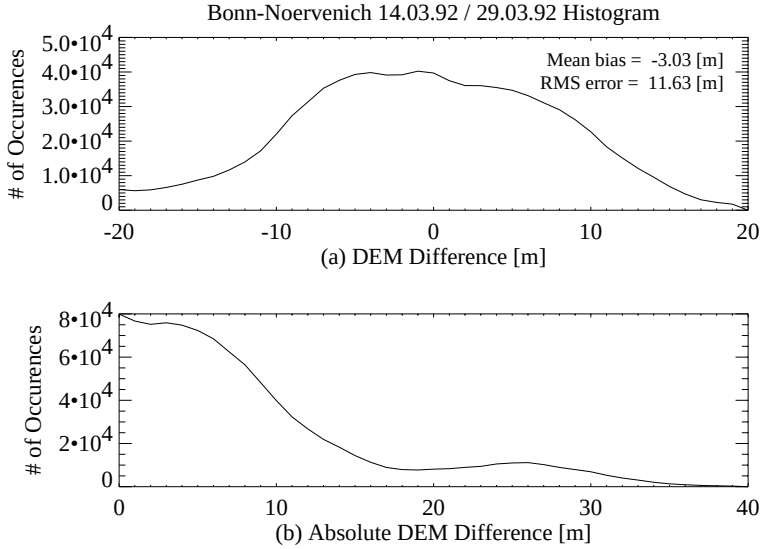


Figure 6-26: Bonn-Nörvenich 14/29 March 1992 - Histogram of height differences between interferometrically calculated heights and reference elevations for all pixels within region; (a) Height difference, (b) Absolute value of height difference

Chapter 7

Conclusions and Outlook

Tis a thing impossible, to frame
Conceptions equal to the soul's desires;
And the most difficult of tasks to *keep*
Heights which the soul is competent to gain.
William Wordsworth, *The Excursion*

The following pages describe general conclusions based on the results presented in the preceding chapters, and provide an outlook toward future developments.

Complexity: Although it is a relatively simple matter to generate a flattened interferogram, and this has often been marketed as a "topographic contour map" the truth is that many non-trivial steps are necessary to move from this stage to the geocoded height model desired by most potential customers. Although a flattened interferogram does offer *prima facie* similarity to a contour map, resolving the 2π ambiguity can in some cases be insurmountable (or require manual operator intervention). In comparison to height model generation, thematic interpretation of coherence images is much less computationally intensive, and can be carried out at a comparatively earlier processing stage. Neither a phase unwrapping nor a forward geocoding step are necessary. The complexity of the height model generation process has hindered its operationalization (despite some premature announcements) until the present.

Accuracies: Test sites were chosen in Germany and Switzerland to be representative of flat and rugged temperate vegetated areas.

Areal validation of height maps generated by repeat-pass ERS InSAR provides confidence in the InSAR technique. RMS accuracies of 2.7m were achieved over a predominantly coherent 12×13km area near Bonn, Germany using three-day repeat ERS-1 data. Systematic biases in the height estimation were minimal over a 40×50 km ERS-1 quarter scene.

The height model calculated from a fifteen-day repeat interferogram over the same area showed RMS errors of 11.6m - the smaller baseline increased sensitivity to coherent phase anomalies, and temporal decorrelation increased. Further improvement in the achievable height accuracy requires improved understanding of the regional phase shifts (e.g. atmospheric influences) that can take place between repeat-pass acquisitions.

Validation Methodologies: DEM flattening using a high quality elevation model is useful for investigating the extent of non-topography-induced phase behaviour while avoiding the necessity of phase unwrapping. Backward geocoding techniques are useful for guaranteeing correct geolocation also in the absence of infallible phase unwrapping. Only forward geocoding allows a "blind" test of end-to-end system performance. Forward and backward geocoding methodologies were shown to have essentially equivalent accuracies given height information of comparable quality.

Although backward geocoding is excluded from use in InSAR height model *production*, it does provide one of many *markets* for the height models produced via SAR interferometry. Good reference height models are required for an accurate terrain correction geolocation of SAR imagery.

Baseline: The baseline between two overpasses of a satellite is not generally known accurately before the data acquisition. The size of the baseline that does occur is partially due to chance.

InSAR based height estimates improve in accuracy with longer baselines, worsening where slopes become extreme, as well as in areas of low coherence (e.g. forest). Spectral-shift filtering dramatically decreases phase variance and is of critical importance for the large baselines that are optimal for the extraction of topography. The weakness of ERS-1 InSAR height derivation lies in hilly forested areas, where low coherences combine with topography to render height estimation problematic.

The optimum baseline for mapping purposes is dependent upon the slopes that can be expected within the scene. Given a relatively flat scene, with gently rolling topography, a baseline of 300-400m provides good height accuracy while at the same time not excessively sacrificing spatial resolution during the spectral shift filtering. Scenes with more significant slopes must make do with smaller baselines and less accuracy, lest they fall prey to unsuccessful phase unwrapping. Consistency checks using multiple interferograms can be used to improve accuracy and assist in phase unwrapping.

Phase Unwrapping: Phase unwrapping errors must be either manually corrected, or reduced through consistency checks with data from additional interferograms. Methods that increase the scope of information considered promise improved reliability. Improvement of the height estimate through combination of multiple tandem ERS-1/ERS-2 pairs is advisable - in severely sloped areas, combination of ascending/descending pairs is necessary to offer a more consistent ground resolution across the scene.

Coherence: Experience with the Bern test site observing the seasonal variation of one day repeat coherence suggests that the winter season appears to be optimum.

The minimum coherence useful for mapping is dependent on the slopes within the scene, the required accuracy, and the scene's baseline. However, as a rule of thumb, requiring that the mean coherence be above 0.5 can be a useful discriminator.

Geometry: If used during height calculation, the far-field approximation present within many models of InSAR geometry sets limitations on the swath width that can be considered. The greatest benefit of the approximation is that the linearization it provides makes it possible to develop frequency domain filtering techniques to counteract the shifted ground reflectivity spectra (also known as baseline decorrelation) caused by the differing viewing angles from the displaced flight tracks. Use of the far-field approximation is also worthwhile in sensitivity analysis and computation of nominal fringe frequency and ambiguity height values.

In highly sloped areas, height model generation requires consideration of the many-to-one nature of the SAR acquisition geometry. Combination of ascending and descending data sets is required to guarantee a minimum resolution. Radiometric normalization that ignores the many-to-one nature of the acquisition geometry, using instead a simple neighbourhood planar approximation, is doomed to failure in areas wherever slopes lose gentleness.

Conditions: Given the right conditions, repeat-pass ERS interferometry can produce height models with respectable accuracy. The interferograms used must have a minimum repeat-pass coherence, which can be difficult to find in vegetated areas - but appears to peak in the winter. Another limitation is found in high slopes that challenge phase unwrapping. To some extent the conditions are contradictory - it is precisely in the scenes with good height sensitivity through the virtue of a large baseline that phase unwrapping becomes most difficult.

Advanced techniques increase the robustness of InSAR processing by making use of more information (e.g. from multiple interferograms) to improve the reliability of the phase unwrapping step.

Existing spaceborne InSAR systems suffer from lack of optimization to the task of height model generation. All existing systems are repeat pass, and suffer from unpredictable temporal decorrelation and baseline extent. Spatial decorrelation can be counteracted using spectral filtering, but decorrelation due to noise from low SNR signals, and due to temporal decorrelation between acquisitions must be accepted.

Spaceborne interferometers have one key advantage over airborne systems: they require no motion compensation. Within the space of a single data acquisition, the satellite orbit need not consider motion induced by atmospheric drag or the platform roll effects that plague airborne systems.

Height model generation using spaceborne repeat-pass InSAR can only be recommended given that baseline, slope (local incidence angle), and coherence (temporal decorrelation) constraints are all satisfied. The baseline must be large enough to guarantee sufficient height sensitivity, the slopes must not exceed limits dictated by the ambiguity resolvable during phase unwrapping, and coherence must be maintained between acquisitions to allow reliable phase difference estimation.

Outlook: The Shuttle Radar Topographic Mapper (SRTM) mission being planned by NASA and Germany for launch in the year 2000 will map the non-polar regions of the Earth using two C-band antennae displaced using a fixed boom extending from the space shuttle cargo bay. Other future configurations being studied include satellite pairs that fix their baseline using a tether connection, as well as non-tethered tandem satellites that synchronize their transmit/receive windows. L-band repeat pass configurations are also being considered - temporal decorrelation acts slower at that wavelength. Future missions may provide more accurate timing and orbit ephemeris information, eliminating the requirement for refinement of image and baseline geometries.

Double-antennae single-pass systems that circumvent the problem of temporal decorrelation while maintaining an optimal baseline promise to realize the full potential of DEM generation using space-based InSAR.

References

[1] Bamler R., Davidson G., and Adam N., *On the nature of noise in 2-D phase unwrapping*, Proc. of European Symposium on Satellite Remote Sensing III, Taormina, Italy, Sept. 23-26, 1996.

[2] Bamler R. and Just D., *Phase Statistics and Decorrelation in SAR Interferograms*, Proc. of IEEE-IGARSS'93, pp. 980-984.

[3] Barmettler A., Pasquali P., Small D., and Nüesch D., *Cross-Compatibility of ERS-SLC Products*, Proc. of ESA-FRINGE'96 Workshop, Zürich, Switzerland, Sept. 30-Oct. 2, 1996, Volume 2, pp. 129-138.

[4] Beaudoin A., Stussi N., Troufleau D., Desbois N., Piet L., Deshayes M., *On the use of ERS-1 SAR data over hilly terrain: Necessity of radiometric corrections for thematic applications*, Proc. of IEEE-IGARSS'95, Florence, Italy, July 10-14, 1995, pp. 2179-2182.

[5] Coltelli M., Dutra L., Fornaro G., Franceschetti G., Lanari R., Migliaccio M., Moreira J., Papathanassiou K., Puglisi G., Riccio D., Schwäbisch M., *SIR-C/X-SAR Interferometry over Mt. Etna: DEM Generation, Accuracy Assessment and Data Interpretation*, DLR Forschungsbericht 95-48, 1996.

[6] Coulson, S.N., *SAR Interferometry with ERS-1*, Earth Observation Quarterly, Number 40, April 1993, pp. 20-23.

[7] Curlander J.C. and McDonough R.N., Synthetic Aperture Radar: Systems & Signal Processing, John Wiley and Sons, New York, 1991.

[8] Curlander J.C., *Location of Spaceborne SAR Imagery*, IEEE Trans. on Geoscience and Remote Sensing, Volume GE-20, Number 3, July 1982, pp. 359-364.

[9] Duchossois G., Kohlhammer G., Martin P., *Completion of the ERS Tandem Mission*, Earth Observation Quarterly, Number 52, June 1996.

[10] Durand, J.M., *Theoretical Feasability of SAR Interferometry*, Proc. of IEEE-IGARSS'87, pp. 1541-1544.

[11] Elachi C., Spaceborne Radar Remote Sensing: Applications and Techniques, IEEE Press, New York, 1988.

[12] ESA Earthnet ESRIN, ERS-1 System, Report SP-1146, European Space Agency Publications Division, 1992.

[13] Ferreti A., Guarnieri A.M., Prati C., and Rocca F., *Multi-baseline interferometric techniques and applications*, Proc. of ESA-FRINGE 1996, Zürich, Switzerland, Sept. 30-Oct. 2, 1996, pp. 243-252.

[14] Frei U., Graf K. Ch., and Meier E. *Cartographic Reference Systems*, chapter in SAR Geocoding: Data and Systems, ed. G. Schreier, Herbert Wichmann Verlag GmbH, 1993.

[15] Gabriel A.K. and Goldstein R.M, *Crossed Orbit Interferometry: Theory and Experimental Results from SIR-B*, Intl. J. Remote Sensing, Volume 9, Number 5, 1988, pp. 857-872.

[16] Gabriel A.K., Goldstein R.M., and Zebker H.A., *Mapping Small Elevation Changes Over Large Areas: Differential Radar Interferometry*, Journal of Geophysical Research, Volume 94, Number B7, July 10, 1989, pp. 9183-9191.

[17] Gatelli, F., Monti Guarnieri A., Parizzi F., Pasquali P., Prati C., and Rocca F., *The Wavenumber Shift in SAR Interferometry*, IEEE Trans. on Geoscience and Remote Sensing, Volume 32, Number 4, July 1994, pp. 855-865.

[18] Gatelli, F., Monti Guarnieri A., Parizzi F., Pasquali P., Prati C., and Rocca F., *Developments in ERS-1 SAR Interferometry: Use of the Spectral Shift*, Proc. of Symposium "From Optics to Radar: SPOT and ERS Applications", Cépaduès-Éditions, May 1993, pp. 491-500.

[19] Gens, R. and Genderen J.L., *Review Article - SAR interferometry -- issues, techniques, applications*, Intl. Journal of Remote Sensing, Volume 17, Number 10, July 10, 1996, pp. 1803-1835.

[20] Geudtner D., Vachon P.W., Mattar K.E., and Gray A.L., *RADARSAT Repeat-pass SAR Interferometry: Results Over an Arctic Test Site*, Proc. of Geomatics in the Era of RADARSAT, Ottawa, Canada, May 1997.

[21] Geudtner D., *Die interferometrische Verarbeitung von SAR-Daten des ERS-1*, DLR Forschungsbericht 95-28, Germany, 1995.

[22] Geudtner D. and Schwäbisch M., *An algorithm for precise reconstruction of InSAR imaging geometry*, Proc. of EUSAR'96, Königswinter, Germany, 1996, pp. 249-252.

[23] Ghiglia D.C. and Romero L.A., *Robust Two-Dimensional Weighted and Unweighted Phase Unwrapping that uses Fourier Transforms and Iterative Methods*, J. Opt. Soc. Am., Vol. A11, pp. 107-117, 1994.

[24] Goblirsch W., Werner C., Nüesch D., and Faller N., *Accuracy of interferometric elevation models generated from DoSAR airborne data*, Proc. of IEEE-IGARSS'95, Florence, Italy, 1995.

[25] Goldstein R.A. and Werner C.L., *Radar Ice Motion Interferometry*, Proc. of ERS Symposium, Florence, Italy, 1997.

[26] Goldstein R.A., Zebker H.A., and Werner C.L, *Satellite Radar Interferometry: Two-dimensional Phase Unwrapping*, Radio Science, Volume 23, Number 4, July-August 1988, pp. 713-720.

[27] Graf, K.Ch., Suter M., Hagger J. and Nüesch D., *Computer Graphics and Remote Sensing - A Synthesis for Environmental Planning and Civil Engineering*, Proc. of Eurographics'94, Oslo, Norway, Sept. 1994, Volume 13, Number 3, pp. 11-22, p. 519.

[28] Graf K.C., *Verwendung Geodätischer Abbildungen bei der Geocodierung von Satelliten-Bildern*, Remote Sensing Series, Volume 13, Department of Geography, University of Zürich, Zürich, Switzerland, 1988.

[29] Graham, L.C., *Synthetic Interferometer Radar for Topographic Mapping*, Proc. of the IEEE, Volume 62, Number 6, June 1974, pp. 763-768.

[30] Gray L., Mattar K.E., and Farris-Manning P.J., *Airborne SAR Interferometry for Terrain Elevation*, Proc. of IEEE-IGARSS'92, pp. 1589-1591.

[31] Gray L. and Farris-Manning P.J., *Two-Pass Interferometry with Airborne Synthetic Aperture Radar*, IEEE Trans. on Geoscience and Remote Sensing, Volume 31, Number 1, January 1993, pp. 180-191.

[32] Gruber T., Massman F.-H., Reigber C., *ERS D-PAF Altimeter and Orbit Global Products Manual*, ERS-D-GPM-31200, Dec. 1996.

[33] Guarnieri A.M., Parizzi F., Pasquali P., Prati C., and Rocca F., *SAR Interferometry Experiments with ERS-1*, Proc. of IEEE-IGARSS'93, pp. 991-993.

[34] Hartl P. and Xia Y., *Besonderheiten der Datenverarbeitung bei der SAR Interferometrie*, Zeitschrift für Photogrammetrie und Fernerkundung, 1993, pp. 214-222.

[35] Hartl P. and Thiel, K.-H., *Fields of Experiments in ERS-1 SAR Interferometry in Bonn and Naples*, Proc. of Symposium "From Optics to Radar: SPOT and ERS Applications", Cépaduès-Éditions, May 1993, pp. 525-534.

[36] Hirosawa H. and Kobayashi N., *Terrain Height Measurement by Synthetic Aperture Radar with an Interferometer*, Intl. J. Remote Sensing, Volume 7, Number 3, 1986, pp. 339-348.

[37] Haefner H., Holecz F., Meier E., Nüesch D., and Piesbergen J., *Capabilities and Limitations of ERS-1 SAR Data for Snowcover Determination in Mountainous Regions*,

Proc. of Second ERS-1 Symposium, Hamburg, Germany, October 11-14, 1993, pp. 971-976.

[38] Holecz F., Meier E., and Nüesch D. *Postprocessing of Relief Induced Radiometric Distorted Spaceborne SAR Imagery*, chapter in SAR Geocoding: Data and Systems, ed. G. Schreier, Herbert Wichmann Verlag GmbH, 1993.

[39] Holecz F., Meier E., Piesbergen J., and Nüesch D., *Topographic Effects on Radar Cross Section*, Proc. of CEOS SAR Calibration Workshop, September 20-24, 1993, ESA-ESTEC, ESA WPP-048, pp. 23-28.

[40] Holecz F., Meier E., Piesbergen J., Nüesch D., and Moreira J., *Rigorous Derivation of Backscattering Coefficient*, Proc. of CEOS SAR Calibration Workshop, September 28-30, 1994, Ann Arbor, Michigan.

[41] Hutchins, M., *ERS-1 Off-Line Cal/Val Status Report*, DRA/CIS(CSC2)/5/14-CR-001, Issue 1, United Kingdom, January 12, 1994.

[42] Kenyi L.W., Raggam H., Kubista E., *Feasibility of Atmospheric Effect (sic) in Interferometric Data and its Interpretation*, ESA-ESTEC Final Report, April 1996.

[43] Konecny G., *Current status and future possibilities for topographic mapping from space*, EARSeL Advances in Remote Sensing, Volume 4, Number 2, Oct. 1995, pp. 1-18.

[44] Li F.K. and Goldstein R.M., *Studies of Multibaseline Spaceborne Interferometric Synthetic Aperture Radars*, IEEE Trans. on Geoscience and Remote Sensing, Volume 28, Number 1, January 1990, pp. 88-97.

[45] Lin Q., Vesecky J.F., and Zebker H.A., *New Approaches in Interferometric SAR Data Processing*, IEEE Trans. on Geoscience and Remote Sensing, Volume 30, Number 3, May 1992, pp. 560-567.

[46] Madsen S.N., H.A. Zebker, and J. Martin, *Topographic Mapping Using Radar Interferometry: Processing Techniques*, IEEE Trans. on Geoscience and Remote Sensing, Volume 31, Number 1, January 1993, pp. 246-255.

[47] Madsen S.N. and Zebker H.A., *Automated Absolute Phase Retrieval in Across-track Interferometry*, Proc. of IEEE-IGARSS'92, Houston, USA, pp. 1582-1584.

[48] Marquadt, D. W., *An Algorithm for Least Squares Estimation of Non-linear Parameters*, Journal for the Society of Industrial and Applied Math, 11, 1963, pp. 431-441.

[49] Massmann, F.-H., Reigber C., König R., Raimondo J.C., and Rajasenan C., *Operational ERS-1 Orbit Determination at D-PAF*, Proc. of First ERS-1 Symposium, Cannes, France, November 4-6, 1992, pp. 451-455.

[50] Massonnet D., Vadon H., and Rossi M., *Reduction of the Need for Phase Unwrapping in SAR Interferometry*, IEEE Trans. on Geoscience and Remote Sensing, Volume 34, Number 2, March 1996, pp. 489-497.

[51] Massonnet, D. Feigl K., Rossi M., and Adragna F., *Radar interferometric mapping of deformation in the year after the Landers earthquake*, Nature, Volume 369, May 19, 1994, pp. 227-230.

[52] Massonnet D., M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigi, and T. Rabaute, *The displacement field of the Landers earthquake mapped by radar interferometry*, Nature, Volume 364, July 8, 1993, pp. 138-142.

[53] Massonnet D. and Rabaute T., *Radar Interferometry: Limits and Potential*, IEEE Trans. on Geoscience and Remote Sensing, Volume 31, Number 2, March 1993, pp. 455-464.

[54] Massonnet D., *Perspectives de l'Interférométrie par Radar*, Bulletin de la Société Française de Photogrammétrie et Télédétection, Number 127, 1993, pp. 3-15.

[55] Massonnet D., *Geoscientific Applications at CNES*, chapter in SAR Geocoding: Data and Systems, ed. G. Schreier, Herbert Wichmann Verlag GmbH, 1993.

[56] Meier E., *Geometrische Korrektur von Bildern orbitgestützter SAR-Systeme*, Remote Sensing Series, Volume 15, Department of Geography, University of Zürich, Zürich, Switzerland, 1989.

[57] Meier E., Frei U., and Nüesch D., *Precise Terrain Corrected Geocoded Images*, chapter in SAR Geocoding: Data and Systems, ed. G. Schreier, Herbert Wichmann Verlag GmbH, 1993.

[58] Moccia A. and Vetrella S., *A Tethered Interferometric Synthetic Aperture Radar (SAR) for a Topographic Mission*, IEEE Trans. on Geoscience and Remote Sensing, Volume 30, Number 1, Jan. 1992, pp. 103-109.

[59] Oppenheim A.V., Willsky A.S., Young I.T., Signals and Systems, Prentice-Hall, New Jersey, 1983.

[60] Oppenheim A.V. and Schafer R.W., Discrete-Time Signal Processing, Prentice-Hall, New Jersey, 1989.

[61] Pasquali P., Holecz F., Small D., and Michel T., *Calibration and Classification of SIR-C Polarimetric and Interferometric SAR Data in Areas with Slope Variations*, Proc. of IEEE-IGARSS'97, Singapore, pp. 448-450.

[62] Prati C., Rocca F., Guarnieri A.M., and Pasquali P., *Report on ERS-1 SAR Interferometric Techniques and Applications*, ESA Contract Report, June 1994.

[63] Prati C. and Rocca F., *Limits to the Resolution of Elevation Maps from Stereo SAR Images*, Intl. J. Remote Sensing, Volume 11, Number 12, 1990, pp. 2215-2235.

[64] Prati C., Rocca F., Guarnieri A.M., and Damonti E., *Seismic Migration For SAR Focusing: Interferometrical Applications*, IEEE Trans. on Geoscience and Remote Sensing, Volume 28, Number 4, July 1990, pp. 627-640.

[65] Prati C., Giani M. and Leuratti N., *SAR Interferometry: A 2-D Phase Unwrapping Technique Based on Phase and Absolute Values Informations*, Proc. of IEEE-IGARSS'90, College-Park, USA, pp. 2043-2046.

[66] Prati C., Rocca F., and Guarnieri A.M., *Effects of Speckle and Additive Noise on the Altimetric Resolution of Interferometric SAR (ISAR) Surveys*, Proc. of IGARSS'89, Vancouver, Canada, pp. 2469-2472.

[67] Press W.H., Teukolsky S.A., Vetterling W.T., and Flannery B.P., Numerical Recipes in C: The Art of Scientific Computing, 2nd Edition, Cambridge University Press, 1992.

[68] Pritt M.D., *Phase unwrapping by means of multigrid techniques for interferometric SAR*, IEEE Trans. on Geoscience and Remote Sensing, Volume 34, Number 3, May 1996, pp. 728-738.

[69] Raney K., Freeman A., Hawkins B., Bamler R., *A Plea for Radar Brightness*, Proc. of IEEE-IGARSS'94, Pasadena, USA, pp. 1090-1092.

[70] Rocca F. and Prati C., *Innovative Applications of Repeated Satellite SAR Surveys*, Proc. of European ISY'92 Conference, Abstracts Book II, Munich, April 1992, pp. 52-53.

[71] Rodríguez E. and Martin J.M., *Theory and Design of Interferometric Synthetic Aperture Radars*, IEE Proceedings-F, Volume 139, Number 2, April 1992, pp. 147-159.

[72] Rogers A.E.E. and Ingalls R.P., *Venus: Mapping the Surface Reflectivity by Radar Interferometry*, Science, August 22, 1969, pp. 797-799.

[73] Rossi M., Rogron B., and Massonet, D., *JERS-1 SAR: Image Quality Assessment, Interferometric Capabilities Validation*, Proc. of CEOS SAR Calibration Workshop, Ann Arbor, Michigan, USA, Sept. 28-30, 1994, pp. 156-167.

[74] Roth A. and Schreier G., *The D-PAF ERS-1 Geocoding System*, Intl. J. Remote Sensing, Volume 13, Number 9, 1992, pp. 1619-1625.

[75] Roth A., Hügel T., Kosmann D., Matschke M., and Schreier G., *Experiences with ERS-1 SAR Geopositional Accuracy*, Proc. of IEEE-IGARSS'93, Tokyo, Japan, pp. 1450-1452.

[76] Schreier G., *Geometrical Properties of SAR Images*, Chapter in Geocoding: Data and Systems, ed. G. Schreier, Herbert Wichmann Verlag GmbH, 1993.

[77] Schreier G., Kosmann D., Roth A., *The D-PAF ERS-1 SAR Geocoding System GEOS*, Chapter in Geocoding: Data and Systems, ed. G. Schreier, Herbert Wichmann Verlag GmbH, 1993.

[78] Schwäbisch M., *Die SAR-Interferometrie zur Erzeugung digitaler Geländemodelle*, DLR Forschungsbericht 95-25, Germany, August 1995.

[79] Small D., Holecz F., Meier E., Nüesch D., and Barmettler A., *Geometric and Radiometric Calibration of RADARSAT Images*, Proc. of Geomatics in the Era of RADARSAT, Ottawa, Canada, May 1997.

[80] Small D. and Nüesch D., *Validation of Height Models from ERS Interferometry*, Proc. of ESA-FRINGE 1996, Zürich, Switzerland, Sept. 30-Oct. 2, 1996, Vol. 2, pp. 43-53.

[81] Small D., Pasquali P., and Füglistaler S., *A Comparison of Phase to Height Conversion Methods for SAR Interferometry*, Proc. of IEEE-IGARSS'96, Lincoln, Nebraska, USA, May 27-31, 1996, pp. 342-344.

[82] Small D., Werner C., and Nüesch D., *Geocoding and Validation of ERS-1 InSAR-derived Digital Elevation Models*, EARSeL Advances in Remote Sensing, Volume 4, Number 2, Oct. 1995, pp. 26-39, pp. I-II.

[83] Small D., Meier E.,Wegmüller U., and Nüesch D., *Applications of Geocoded ERS-1 InSAR-derived Digital Terrain Information*, Proc. of CEOS SAR Calibration Workshop, Ann Arbor, Michigan, USA, September 28-30, 1994, pp. 184-190.

[84] Small D., Werner C., and Nüesch D., *Baseline Modelling for ERS-1 SAR Interferometry*, Proc. of IEEE-IGARSS'93, pp. 1204-1206.

[85] Small D., Werner C., and Nüesch D., *Registration of ERS-1 SLC Products for Interferometry*, Proc. of Fourth GEOSAR Workshop, Loipersdorf, Austria, May 26-28, 1993, pp. 63-66.

[86] Small D. and Nüesch D., ERS-1 InSAR Concept Study, Report to DLR, Aug. 1992.

[87] Snyder J.P., Flattening the Earth: Two Thousand Years of Map Projections, University of Chicago Press, 1993.

[88] Stebler O., Pasquali P., Small D., Holecz F., and Nüesch D., *Analysis of ERS-SAR Tandem Time-Series Using Coherence and Backscattering Coefficient*, Proc. of ESA-FRINGE 1996, Zürich, Switzerland, Sept. 30-Oct. 2, 1996, Vol. 2, pp. 71-81.

[89] Villasenor J. and Zebker H.A., *Temporal Decorrelation in Repeat-Pass Interferometry*, Proc. of IEEE-IGARSS'92, pp. 941-943.

[90] Wegmüller U. and Werner C., *Analyses of Interferometric Land Surface Signatures*, Proc. of PIERS'94, July 1994.

[91] Wegmüller U., Werner C., Small D., and Nüesch D., *Analysis of Land Surface Signatures Using SAR Interferometry*, RSL Report to ESA-ESTEC, Contract 134 516, June 1994.

[92] Werner C., Wegmüller U., Small, D., and Rosen, P., *Applications of Interferometrically Derived Terrain Slopes: Normalization of SAR Backscatter and the Interferometric Correlation Coeffient*, Proc. of Second ERS-1 Symposium, Hamburg, Germany, October 1993, pp. 205-210.

[93] Werner C., Small D., Rosen P., Hensley S., and Zebker H., *Techniques and Applications of SAR Interferometry for ERS-1*, Proc. of Symposium "From Optics to Radar: SPOT and ERS Applications", Cépaduès-Éditions, May 1993, pp. 483-490.

[94] Xu W., *Comments on "Reduction of the Need for Phase Unwrapping in Radar Interferometry"*, IEEE-Geoscience and Remote Sensing, Volume 35, Number 4, July 1997, p. 1073.

[95] Zakharov A.I., *Ice Fields Movement in the Antarctic Area from Almaz-1 Repeated Orbits SAR Interferometry*, Proc. of CEOS SAR Calibration Workshop, Ann Arbor, Michigan, USA, Sept. 28-30, 1994, pp. 191-200.

[96] Zebker H.A., Werner C., Rosen P., and Hensley S., *Accuracy of Topographic Maps Derived from ERS-1 Interferometric Radar*, IEEE-Geoscience and Remote Sensing, Volume 32, Number 4, July 1994, pp. 823-836.

[97] Zebker H.A. and Villasenor J., *Decorrelation in Interferometric Radar Echoes*, IEEE Trans. on Geoscience and Remote Sensing, Volume 30, Number 5, September 1992, pp. 950-959.

[98] Zebker H.A. and Goldstein R.M., *Topographic Mapping from Interferometric Synthetic Aperture Radar Observations*, Journal of Geophysical Research, Volume 91, Number B5, April 10, 1986, pp. 4993-4999.

List of Figures

List of Tables

Appendix A

List of Variables

Parameter	Symbol	Units
Azimuth resolution	A_r	m
Azimuth pixel spacing (sampling)	A_s	m
Baseline vector between two acquisition locations	$\underset{\sim}{B}$	m
Length of component of baseline vector parallel to line of sight	$B_{\parallel}$	m
Length of component of baseline vector perpendicular to line of sight	$B_{\perp}$	m
Critical baseline	$B_{\perp, cr}$	m
Cross-track component of baseline	B_c	m
Instantaneous cross-track component of baseline at azimuth position i	$B_{c, i}$	m
Normal component of baseline	B_n	m
Azimuth processed bandwidth	B_a	Hz
Range chirp bandwidth	B_r	MHz
Speed of light	c	m/s

Parameter	Symbol	Units
Basis vector for “cross-track” axis of TCN coordinate system	$\hat{\underset{\sim}{c}}$	-
Coarse azimuth offset	c_a	pixels
Coarse range offset	c_r	pixels
Radar carrier frequency	f_0	GHz
Azimuth Sampling Rate (also PRF)	f_{As}	Hz
Range Sampling Rate	f_{Rs}	MHz
Shift in azimuth frequency	Δf_a	Hz
Shift in range frequency	Δf_r	MHz
Complex flattened interferogram	$\underset{\sim}{F}$	
Complex normalized flattened interferogram	$\hat{\underset{\sim}{F}}$	
Complex filtered flattened interferogram	$\underset{\sim}{F}'$	
Complex raw interferogram	$\underset{\sim}{G}$	
Complex normalized raw interferogram	$\hat{\underset{\sim}{G}}$	
Measured height value	h	m
Ambiguity height: change in elevation required to cause a full fringe of phase rotation	$\Delta h_{2\pi}$	m
Basis vector for “normal” axis of TCN coordinate system	$\hat{\underset{\sim}{n}}$	-

Parameter	Symbol	Units
Reference height of satellite in simplified 2-D geometry	H	m
Position vector of point on Earth's surface	$\underset{\sim}{P}$	m
Position vector of point on DEM	$\underset{\sim}{P}_d$	m
Position vector of point on ellipsoid	$\underset{\sim}{P}_e$	m
Pulse Repetition Frequency	PRF	Hz
Local radius of Earth	r_E	m
Range distance from spacecraft to point on Earth's surface	R	m
Range distance from antenna 1 to point on Earth's surface	R_1	m
Range distance from antenna 1 to point on Earth's surface	R_2	m
Slant range resolution	R_r	m
Slant range pixel spacing (sampling)	R_s	m
Slant range distance per nominal "flat Earth" fringe	$\Delta R_{2\pi}$	m
Spacecraft position vector, antenna 1	$\underset{\sim}{S}_1$	m
Spacecraft position vector, antenna 2	$\underset{\sim}{S}_2$	m
Basis vector for "tangential" axis of TCN coordinate system	$\hat{\underset{\sim}{t}}$	-

Parameter	Symbol	Units
Instantaneous spacecraft velocity vector	$\underset{\sim}{v}_s$	m/s
Instantaneous velocity vector of point on Earth's surface	$\underset{\sim}{v}_p$	m/s
Ground slope	α	radians
Change in cross-track baseline per unit time caused by orbital azimuth convergence	α_c	m/s
Difference in range distances between two acquisitions	δ_r	m
Interferometric phase	ϕ	radians
Phase difference modeled using DEM	ϕ_d	radians
Phase difference modeled using ellipsoid	ϕ_e	radians
Phase difference modeled synthetically using either DEM or ellipsoid	ϕ_m	radians
Interferometric coherence	γ	-
Coherence caused by system response	γ_H	-
Radar wavelength	λ	cm
Incidence angle between line-of-sight vector and Earth surface normal	θ	radians
Nominal change in incidence angle between acquisitions	$\Delta\theta$	radians
Backscatter coefficient	σ^0	

Parameter	Symbol	Units
Pulse length	τ_p	s

Appendix B

Glossary

This glossary contains short explanations of remote sensing terms and acronyms that are used in this dissertation.

Coherence

is a measure used in SAR interferometry to quantify the amount of phase noise present. The quantity ranges between zero and one, with zero indicating a completely uncorrelated random phase, and a maximum of one resulting when all phase contributions are identical.

DEM

stands for Digital Elevation Model. Elevations are ground level heights above a given reference, especially that of the ocean.

D-PAF

is an acronym for the German (*Deutsches*) Processing and Archiving Facility, one of the four major facilities constructed in Europe for the processing of ERS data.

DSM

is an acronym for Digital Surface Model. Surface height values are stored rather than the ground-level elevation measurements found in a DEM. For example, tree heights are considered in a DSM, but ignored in a DEM.

DTM

is an acronym for Digital Terrain Model. DTM and DEM are often used synonymously. DTM is the more general expression since it can, in addition to elevation information, also include description of terrain break lines and other topographical features.

ECR

is an acronym for Earth Centred Rotating. The reference frame is not fixed to the stars, but rather to the rotating Earth.

ERS

is an acronym for European Remote Sensing Satellite. At C-band, VV polarization, the two ESA satellites ERS-1 (launched July 1991) and ERS-2 (launched April 1995) have pioneered operational active microwave remote sensing from space.

ESA

is an acronym for European Space Agency, an organization with a mandate

from a group of European states to further the exploration and exploitation of space.

FIR

is an acronym for Finite Impulse Response. Within the field of signal processing, an FIR filter has a kernel of finite length.

GEOS

is the name of the operational SAR geocoding software developed by a consortium of the D-PAF, the University of Zürich's Remote Sensing Laboratories, and the Joanneum Research Centre in Graz, Austria. The software system has been used to operationally produce terrain-geocoded ERS products since 1992.

GEC

is an acronym for Geocoded Ellipsoid Corrected. An ellipsoid model is used to transform from range Doppler geometry into the chosen map reference system.

GIS

is an acronym for Geographic Information System. Spatial information is stored and organized in a computer.

GTC

is an acronym for Geocoded Terrain Corrected. A digital elevation model is used to transform from range Doppler geometry into the chosen map reference system. Terrain correction enables overlay of multi-temporal SAR images acquired with heterogeneous geometries, and is also a prerequisite for combination of SAR data with geographically-tagged information from other sources.

Height Field

A height field is a matrix of numerical values representing the vertical height above some reference surface.

Hypsography

is that branch of geography that deals with the measurement and mapping of the topography of the Earth above sea level.

Layover

In areas with steep slopes, the peak of a mountain may be closer to the radar satellite than a valley floor closer to the radar sensor's ground track. In such cases, distance from the ground track does not increase with slant range, as in flat or less steep terrain. Distance from the ground track instead *decreases* with increasing slant range, before reversing again, reassuming its positive correlation with slant range. The effect causes concentration of a large *map geometry* area within a small *slant range* region, and is strongest at steep incidence angles.

RSAT-1
see RADARSAT-1.

RADARSAT-1
is the first Earth-observation satellite launched by Canada, in November 1995. At C-band HH polarization, it uses electronic beam steering to acquire swaths from a variety of incidence angles, and optionally interleaving (ScanSAR) to form a wider swath.

RGB
is an abbreviation for Red, Green and Blue, the additive primary colours.

SAR
is an acronym for Synthetic Aperture Radar. SAR is an active microwave remote sensing system that makes use of signal processing techniques to “synthesize” a pseudo-antenna along the length of the radar platform’s flight track. The generated images show the Earth’s reflective properties at microwave wavelengths.

Shadow
Radar shadow occurs when terrain obstructs a portion of the Earth’s surface from a line of sight connection to the radar sensor’s flight track. Such areas are not illuminated by the microwave pulses emitted by the radar, and therefore do not produce an echo. After range and azimuth compression, their range-Doppler coordinates are empty. Note that radar shadow differs from optical shadows in that the illumination originates from the sensor itself, and not from the sun as in electro optical imagery.

SNR
is an abbreviation for signal to noise ratio, the ratio between the pure signal being estimated and the noise in the channel being measured.

SRTM
is an acronym for Shuttle Radar Topography Mission. Working together with the German Space Agency, NASA plans to launch a week-long shuttle mission near the end of the century, using pairs of C-band and X-band antennae (one of each on a boom) to map non-polar areas of the globe.

Tandem Mission
One often refers to the use of both the ERS-1 and ERS-2 satellites together in tandem as a “tandem mission”. During 1995 and 1996 they acquired much of the globe with a repeat-pass interval of just one day. In comparison to the previously sparsely scattered set of available InSAR pairs (from earlier ERS-1 3-day repeat orbits), the tandem mission significantly increased the available coverage.

TIN

is an acronym for Triangular Irregular Network. By means of a TIN DTM's can be described both accurately and concisely.

WGS84

is an acronym for World Geodetic System, 1984. It describes an oblate ellipsoid with origin at the centre of the Earth, and has become the most widely used geocentric global datum.

Appendix C

Hardware / Software

Hardware

Computers

All software has been developed on SUN SPARCstations running the UNIX operating system Solaris. The development platform was equipped with 112 MBytes of memory.

Software

ZIP

The InSAR geocoding software ZIP is coded mainly in the programming languages C / C++, with some visualizations done using IDL. Written mainly by David Small of RSL.

FrameMaker

FrameMaker is a document publishing program with graphic and equation editing capabilities. Commercial software from Adobe.

XV

XV is an interactive image manipulation program for the X Window System. It can operate on most current image file formats and can cope with 8-bit as well as with 24-bit frame buffers. XV was developed by J. Bradley and is shareware.

Remote Sensing Series

Vol. 5 ***R. Schoch:*** **Land-Cover Studies and Crop Acreage Estimates from Aerial Photography and Satellite Imagery – A Case Study in the Region of Ta'izz-Turbah, Yemen Arab Republic**. (246 pp, 1982, sFr.40.–)

Vol. 6 ***E. Meier:*** **Unterscheidung verschiedener Schneetypen im digitalen Multispektralscannerbild**. (96 pp, 1982, sFr. 25.–)

Vol. 7 ***D. Nüesch:*** **Augmentation of Landsat MSS Data by Seasat SAR Imagery for Agricultural Inventories**. (Reprint of: NASA-Rep.IT-E2-04233), (106 pp, 1982, sFr. 30.–)

Vol. 8 ***U. Frei:*** **Geometrische Korrekturen von NOAA-AVHRR-Daten**. (87 pp, 1984, sFr. 25.–)

Vol. 9 ***H. Haefner und F. Hugentobler:*** **Assessment and Monitoring of Abandoned Agricultural Land in the Swiss Alps**. (21 pp, 2 maps, 1985, sFr. 20.–)

Vol. 10 ***M. Keller:*** **Ausaperungskartierung mit Landsat-MSS Daten zur Erfassung ökologischer Einflussgrössen im Gebirge**. (111 pp, 1987, sFr. 36.–)

Vol. 11 ***M. Baumgartner:*** **Schneeschmelz-Abflussimulationen basierend auf Schneeflächenbestimmungen mit digitalen Landsat-MSS und NOAA-AVHRR-Daten**. (201 pp, 1987, sFr. 35.–)

Vol. 12 ***H. Maurer und E. Meier:*** **Digitale Waldschadenkartierung der Nadelhölzer**. (34 pp, 1987, sFr. 20.–)

Vol. 13 ***K. Ch. Graf:*** **Verwendung geodätischer Abbildungen bei der Geocodierung von Satellitenbildern**. (116 pp, 1988, sFr. 30.–)

Vol. 14 ***H. H. Schiesser:*** **Fernerkundung von Hagelschäden mittels Wetterradar untersucht an Ackerkulturen**. (199 pp, 1988, sFr. 35.–)

Vol. 15 ***E. Meier:*** **Geometrische Korrektur von Bildern orbitgestützter SAR-Systeme**. (137 pp, 1989, sFr. 35.–)

Vol. 16 ***P. Meyer:*** **Segmentierung und symbolische Beschreibung als Grundlage zur Klassifikation landwirtschaftlicher Kulturen in einem Multispektralscanner-Datensatz**. (151 pp, 1990, sFr. 40.–)

Vol. 17 ***K. Itten, P. Meyer, T. Kellenberger, C. Affentranger, I. Leiss:*** **Mapping of Swiss Forests with NOAA-AVHRR**. (62 pp, 1991, sFr. 30.–)

Vol. 18 ***K. Itten, P. Meyer, T. Kellenberger, R. Leu, St. Sandmeier, P. Bitter and K. Seidel:*** **Correction of the Impact of Topography and Atmosphere on Landsat-TM Forest Mapping of Alpine Regions**. (50 pp, 1992, sFr. 40.–)

Vol. 19 ***B. Oester:*** **Erfassen der Waldschaden-Entwicklung anhand von grossmassstäblichen Infrarot-Farbluftbildern**. (163 pp, 1991, sFr. 40.–)

Vol. 20 ***U.J. Burkart:*** **Simulation und Prognose des Schmelzwasserabflusses sowie Bestimmung des regionalen Schneewasseräquivalents unter Zuhilfenahme von Satellitenbilddaten – dargestellt am Beispiel des Einzugsgebietes Rhein-Felsberg und Teilgebieten**. (126 pp, 1991, sFr. 40.–)

Vol. 21 ***H. Haefner, A.H. Schumann:*** **Remote Sensing Applications in Hydrology and Water Resources Management**. (39 pp, 1992, sFr. 30.–)

Vol. 22 ***U. Frei:*** **Compilation of Cartographic and Spaceborne Remote Sensing Data for Thematic/Topographic Mapping**. (102 pp, 1993, sFr. 40.–)

Vol. 23 ***F. Holecz:*** **Postprocessing von SAR-Satellitenbilddaten**. (149 pp, 1993, sFr. 40.–)

Vol. 24 ***A.M. Darvishsefat:*** **Einsatz und Fusion von multisensoralen Satellitenbilddaten zur Erfassung von Waldinventuren**. (146 pp, 1995, sFr. 40.–)

Vol. 25 ***K. Ch. Graf:*** **Realistic Landscape Rendering using Remote Sensing Images, Digital Terrain Models and 3D Objects**. (120 pp, 1995, sFr. 40.–)

Vol. 26 ***St. Sandmeier:*** **A Physically-Based Radiometric Correction Model. Correction of Atmospheric and Illumination Effects in Optical Satellite Data of Rugged Terrain**. (144 pp, 1995, sFr. 40.–)

Vol. 27 ***W. Brüsch:*** **Das Snowmelt Runoff Model ETH (SRM-ETH) als universelles Simulations- und Prognosesystem von Schneeschmelz-Abflussmengen**. (166 pp, 1996, sFr. 45.–)

Vol. 28 ***T. Kellenberger:*** **Erfassung der Waldfläche in der Schweiz mit multispektralen Satellitenbilddaten. Grundlagen, Methodenentwicklung und Anwendungen**. (290 pp, 1996, sFr. 35.–)

Vol. 29 ***M. Suter:*** **Aspekte der interaktiven real-time 3D-Landschaftsvisualisierung**. (150 pp, 1997, sFr. 40.–)

Vol. 30 ***D. Small:*** **Generation of Digital Elevation Models through Spaceborne SAR Interferometry**. (168 pp, 1998, sFr. 40.–)

Order form:

Vol. / Title:.....

Name: .....

Address:.....

Please send to:

Remote Sensing Laboratories, *Department of Geography*
University of Zurich-Irchel, Winterthurerstrasse 190
CH-8057 Zurich, *Switzerland*
Phone 0041-1-635 51 31, Fax 0041-1-635 68 48