

Copyright

by

Kenneth Clinton Slatton

2001

**Adaptive Multiscale Estimation for Fusing Image
Data**

by

Kenneth Clinton Slatton, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2001

The Dissertation Committee for Kenneth Clinton Slatton certifies that this is the approved version of the following dissertation:

Adaptive Multiscale Estimation for Fusing Image Data

Committee:

Melba M. Crawford, Supervisor

Brian L. Evans, Co-Supervisor

Alan C. Bovik

Yunjin Kim

Hao Ling

Edward J. Powers

to my wife and my parents with love

Acknowledgments

I have a long list of people to thank for the support and guidance. First, I want to thank my parents William N. Slatton and Linda Matros. They inspired me to pursue an education and supported me in many ways during that education. They were by far my most important influences during the first twenty years of my life.

I also must thank my wonderful wife Jennifer for giving her love and support. It was not easy to be married to a graduate student. She sacrificed her standard of living and more importantly her time with me while I remained in school. She also brought a glorious new life, our son William H. Slatton, into the world during this time. For all of this, I will always be grateful and in awe of her.

I wish to thank my advisors, Prof. Melba M. Crawford and Prof. Brian L. Evans. Without their insightful guidance and words of encouragement this dissertation would never have been realized.

I also wish to thank other people that have served as teacher to me at different stages in my graduate education. Dr. Yunjin Kim was extremely generous with his time and helped me refine my research topic, which eventually led to my receiving a National Aeronautics and Space Administration Graduate Research Program Fellowship. Dr. Robert Treuhaft was also giving of his time when I struggled to understand certain radar scattering problems.

This work was supported by the National Aeronautics and Space Adminis-

tration, under the Topography and Surface Change Program (Grant NAG5-2954)
and the Graduate Student Research Fellowship Program (Grant NGT-50239).

KENNETH CLINTON SLATTON

The University of Texas at Austin

December 2001

Adaptive Multiscale Estimation for Fusing Image Data

Publication No. _____

Kenneth Clinton Slatton, Ph.D.

The University of Texas at Austin, 2001

Supervisors: Melba M. Crawford and Brian L. Evans

There is a critical need to accurately measure land surface topography over large areas around the world. Topographic data are required for a wide range of civilian, government, and military applications, including assessing the threat and impact of natural hazards such as flooding and planning military operations. Imaging radars have been used extensively to map terrain. They can operate in the microwave portion of the electromagnetic spectrum, which enables them to image during the day or night and under most weather conditions.

Interferometric synthetic aperture radar (INSAR) provides the best overall capability for measuring topography over areas of 10 km² and larger. For many application though, the resolution and accuracy of INSAR is insufficient. This is especially true if the surface is covered with vegetation. INSAR observations do not provide direct measurements of the true surface topography in vegetated areas, but instead yield a height that depends on the sensor characteristics, the surface el-

evation, and the vegetation. Laser altimeter (LIDAR) sensors can be used to obtain topographic measurements with an order of magnitude better resolution and accuracy than INSAR, but are generally restricted to areas of less than 10 km² because of limited coverage.

In this dissertation, I develop a data fusion framework for the statistically optimal combining of complementary data sets. I apply the framework to fusing INSAR and LIDAR data to produce improved estimates of topography. Neither INSAR nor LIDAR data strictly represents bare surface heights in the presence of vegetation, so prior to the data fusion, bare surface elevations and vegetation heights are estimated from the data by modeling the interactions between the incident energy from the sensors and the vegetation.

The transformed data sets are then combined to exploit the coverage of INSAR and resolution of LIDAR. The data fusion is performed using adaptive multiscale estimation to efficiently capture statistical correlation in the data across many scales. I extend a recently developed multiscale estimation method to allow adaptive estimation of non-stationary spatial processes.

The contributions of this work include (1) combining physical modeling with multiscale estimation to accommodate nonlinear measurement-state relationships, (2) extending multiscale estimation techniques to adaptively estimate non-stationary processes, and (3) improving estimates of ground elevations and vegetation heights for remote sensing applications.

Contents

Acknowledgments	v
Abstract	vii
List of Tables	xiii
List of Figures	xv
Chapter 1 Introduction	1
1.1 Remotely sensing topography	1
1.2 Adaptive multiscale estimation	6
1.3 Abbreviations and terminology used	6
1.4 Thesis statement and organization	7
Chapter 2 Characterizing Topographic Measurements	11
2.1 Introduction	11
2.2 The INSAR measurement	12
2.2.1 Obtaining height estimates	12
2.2.2 Random error sources	16
2.3 Estimating surface and vegetation heights with INSAR	19
2.3.1 Empirical approaches	20

2.3.2	Scattering model approach	21
2.3.3	Inverting the scattering model	27
2.3.4	Computational complexity	31
2.4	LIDAR measurements	31
2.5	Conclusions	34
Chapter 3 Data Fusion Via Multiscale Estimation		35
3.1	Introduction	35
3.2	Background	36
3.3	The Kalman Filter	39
3.3.1	Linear state model	39
3.3.2	Kalman filter algorithm	40
3.3.3	Kalman smoothing estimates	42
3.4	Multiresolution Kalman smoothing	43
3.4.1	MKS algorithm	48
3.4.2	Stochastic model identification	50
3.4.3	Computational complexity	52
3.5	Data fusion with the Kalman filter	52
3.5.1	Impact of additional data types	53
3.5.2	Error propagation	54
3.6	Conclusions	57
Chapter 4 Adaptive Multiscale Estimation		59
4.1	Introduction	59
4.2	Effects of modeling errors	60
4.3	Adaptive estimation approaches	65
4.4	Innovation–correlation method	70

4.4.1	Test for optimality	72
4.4.2	Estimation of model parameters	73
4.5	Multidimensional state model	75
4.5.1	Computational complexity	78
4.5.2	Simulations	78
4.6	Conclusions	92
Chapter 5 Fusing INSAR and LIDAR Data		93
5.1	Introduction	93
5.2	Physical modeling and MKS	94
5.2.1	Data description: Texas coastal environment	94
5.2.2	Site 1 results	98
5.2.3	Site 2 results	104
5.2.4	INSAR only case	106
5.3	Physical modeling and AMKS	107
5.3.1	Data description: Texas riparian environment	107
5.3.2	MKS results	107
5.3.3	AMKS Results	112
5.4	Fusing three data sets with MKS	121
5.4.1	Data description: Australian semi-arid environment	121
5.4.2	Results	125
5.5	Conclusions	130
Chapter 6 Conclusions		132
6.1	Summary	133
6.2	Future work	134
Bibliography		137

List of Tables

4.1	Mean squared error of the coarse observations in Fig. 4.10 relative to the true process values. H evaluated at $m = M$	82
4.2	Mean squared error of the MKS estimates relative to the true process values. H evaluated at $m = M$	83
4.3	Mean squared error of the AMKS estimates relative to the true process values. H evaluated at $m = M$	84
4.4	Mean squared error of the coarse observations, the MKS estimates, and the AMKS estimates relative to the true process values. $H = \{0 \text{ or } 1\}$ and evaluated at $m = M$	89
5.1	Mean squared error of the transformed INSAR data for the Elm Grove test site. H evaluated at $m = M$	104
5.2	Mean squared error after MKS for the Elm Grove test site. H evaluated at $m = M$	104
5.3	Mean squared error for the transformed INSAR data for the Melody Lane test site. H evaluated at $m = M$	105
5.4	Mean squared error after MKS for the Melody Lane test site. H evaluated at $m = M$	106

5.5	Mean squared error after MKS for the Melody Lane test site using only INSAR data. H evaluated at $m = M$	106
5.6	Mean squared error of the transformed INSAR data for the Riverside Park test site. H evaluated at $m = M$	112
5.7	Mean squared error after MKS for the Riverside test site. H evaluated at $m = M$	112
5.8	Mean squared error after AMKS for the Riverside test site. H evaluated at $m = M$	116

List of Figures

1.1	Characteristic interferometric synthetic aperture radar (INSAR) properties	3
1.2	The effect of vegetation on INSAR-derived topographic measurements.	3
1.3	Characteristic LIDAR properties	4
1.4	Overview of the data fusion framework	5
2.1	INSAR imaging geometry.	15
2.2	The effect of vegetation on the INSAR measurement. z_g is the true ground height, z_S is the nominal INSAR-derived surface height, and z_v is the true vegetation height. Neglecting noise, z_S will approximately equal z_g where there is no obscuring vegetation. Where vegetation is present, z_S will be between z_g and z_v	18
2.3	INSAR acquisition with dual-baseline system. Different complex image pairs (C_1C_2 or C_3C_4) are formed depending on the transmit/receive sequences used. The procedure yields two independent INSAR acquisitions.	19

2.4	Effects of vertically distributed scattering elements (representing vegetation) on the INSAR measurement. As the height of the scattering elements increases, (a) the differential path length $\Delta\rho$ increases, (b) the magnitude of the cross-correlation $C_1C_2^*$ decreases relative to the backscattered power $C_1C_1^*$ and $C_2C_2^*$, and (c) the backscattered power increases.	23
2.5	Sensitivity of $NICC$ to Δz_v and τ	26
2.6	Sensitivity of \mathcal{M} to baseline distance for different values of Δz_v	28
2.7	The modeled vegetation medium, which consists of a layer of randomly distributed scattering particles. Here, \hat{k} denotes the direction of the incident wave from the INSAR, and ρ_v denotes the path length in the vegetation medium.	30
2.8	Effect of vegetation on the LIDAR measurement. z_g is the true ground height, z_L is the nominal LIDAR-derived surface height, and z_v is the true vegetation height. Neglecting noise, z_L will approximately equal z_g where there is no obscuring vegetation. Where vegetation is present, z_L will approximately equal z_v	33
3.1	The quadtree data structure, where m denotes scale. The support of the tree at the finest scale is $2^M \times 2^M$, where $m \in \{0, \dots, M\}$	38
3.2	Block diagram of a Kalman filter, where $u \sim \mathcal{N}(0, 1)$, $Qu = w$, $Ru = v$	39
3.3	The recursive Kalman filter algorithm.	41
3.4	Block diagram of the overall data fusion process. $z(\cdot)$ quantities represent heights, $\sigma(\cdot)$ quantities represent standard deviations of measurement errors, and $P(\cdot)$ quantities represent final estimate error variances.	46

3.5	Simplified multiscale data structure (dyadic tree) used to present notation. B represents a backshift operator in scale. $s \wedge t$ represents a parent node to nodes s and t . α represents a forward shift operator in scale.	47
3.6	$\sqrt{P_k}$ for a simulated 1-D random process. No observations are present after $k = l = 300$. $\sqrt{P_k}$ converges to $\sqrt{Q/(1 - \Phi^2)}$, which is equal to 7.1421 in this example.	57
4.1	The quadtree data structure, as seen in Fig. 3.1, with observations shown at difference scales. The AMKS algorithm operates in scale just as the MKS algorithm does, but it has the capability to operate in the spatial dimensions to detect modeling errors and non-stationarities.	60
4.2	Simulated 1-D random process that is piecewise WSS, with $Q_2 = 4Q_1$ and $R_2 = 3R_1$	62
4.3	Adaptive \hat{x} and non-adaptive \hat{x}_{na} estimates of the process in Fig. 4.2. 63	
4.4	Estimate uncertainty and actual squared error for the adaptive and non-adaptive estimates in Fig. 4.3.	64
4.5	Simulated 1-D random process with large changes in the noise variances, $Q_2 = 200Q_1$ and $R_2 = 10R_1$	65
4.6	Autocorrelation of the innovations for the adaptive and non-adaptive estimates of the process in Fig. 4.2.	68
4.7	Autocorrelation of the innovations for the adaptive and non-adaptive estimates of the process in Fig. 4.5.	69

4.8	Graphical depiction of how information from the scale-wise filter is provided to the spatial filter. The current estimate is represented by the shaded pixel.	77
4.9	A composite multiscale process generated to simulate the <i>true</i> process at the finest scale in the multiscale pyramid. A rapidly varying process (region 2) is embedded in a slowly-varying process (region 1).	80
4.10	Simulated observations and measurement uncertainties. The region of support for the fine-scale data is a diagonal swath extending from the lower left corner to the upper right corner of that image.	81
4.11	Fused estimates and estimate uncertainty from the MKS algorithm operating on the process shown in Fig. 4.9.	83
4.12	Fused estimates and estimate uncertainty from the AMKS algorithm operating on the process shown in Fig. 4.9.	84
4.13	Fine-to-coarse process noise variances $Q(m, i, j)$ from the AMKS algorithm. The initial value for Q at $m = 9$ is uniform and approximately 1.0. At $m = 7$, the spatially estimated Q is incorporated back into the multiscale filter. Q is revised upward as the spatial filters pass over region 2.	85
4.14	A WSS multiscale process generated to simulate the <i>true</i> process at the finest scale in the multiscale pyramid. The initial Q in the MKS and AMKS filters is intentionally chosen to correspond to a smoother process.	87
4.15	Fine-to-coarse process noise variances $Q(m, i, j)$ from the AMKS algorithm.	88
4.16	Fused estimates and estimate uncertainty from the AMKS algorithm operating on the process shown in Fig. 4.14.	90

4.17	Transects from row 128 in the AMKS estimates. The vertical lines delineate the region of support of fine-resolution data in the central portion of the plot.	91
5.1	Polarimetric SAR image (L-band) of the test sites located on Bolivar, Texas. The boxes indicate portions of the image used in this analysis. The top subset is the Melody Lane site and the bottom subset is the Elm Grove site.	96
5.2	Photograph of trees at the Melody Lane site on Bolivar, Texas. The trees generally ranged between 6 m and 10 m in height.	97
5.3	Image lattices indicating the characteristic support of the data in the sparse acquisition scenario.	97
5.4	Transformed INSAR data (left column) and LIDAR data (right column) that serve as inputs to the multiscale estimation. Thatched areas indicate where LIDAR data were omitted.	99
5.5	MKS estimates of ground elevations and vegetation heights (left column) and the corresponding estimate uncertainties (right column).	101
5.6	Comparison of data fusion results and original data for Elm Grove site. Data extracted from row 38 in Fig. 5.5.	102
5.7	Comparison of data fusion results and original data for Melody Lane site. Data extracted from row 40 in Melody Lane output. . . .	105
5.8	Polarimetric SAR image (C-band) of Austin, Texas. A box indicates the location of the Riverside Park test site.	108
5.9	Photograph of trees in Riverside Park in Austin, Texas. The trees generally ranged from 7 m to 15 m.	109

5.10	MKS output for Riverside site, z_g . The region of support for the LIDAR data is a diagonal swath running from the lower left corner to the upper right corner of the LIDAR image.	113
5.11	MKS output for Riverside site, Δz_v . The region of support for the LIDAR data is a diagonal swath running from the lower left corner to the upper right corner of the LIDAR image.	114
5.12	AMKS output for Riverside site, z_g	117
5.13	Fine-to-coarse process noise variances $Q(m, i, j)$ at the Riverside from the AMKS algorithm, z_g . Q is revised downward in lower right half of the Q_5 image, corresponding to smoother terrain in that portion of the INSAR image in Fig. 5.12.	118
5.14	AMKS output for Riverside site, Δz_v	119
5.15	Fine-to-coarse process noise variances $Q(m, i, j)$ at the Riverside from the AMKS algorithm, Δz_v . Q is revised downward along a band running from the upper left to the lower right of the Q_5 image, corresponding to the smooth grassy area in that portion of the INSAR image in Fig. 5.14.	120
5.16	ERS-1 topographic image of the Finke River Gorge in Australia. The MKS algorithm is applied in the boxed region in the upper left portion of the image. The imaged area is approximately $50 \text{ km} \times 50 \text{ km}$	122
5.17	Relative coverage of the ERS and TOPSAR acquisitions.	123
5.18	Detailed views of the TOPSAR acquisitions of the Finke River. The arrows indicate the SAR look direction. Topography induced shadowing leads to data drop outs, which occur in different locations in the two images because of the different look directions. The displayed area is approximately $10 \text{ km} \times 10 \text{ km}$	124

5.19	MKS output for Finke site, z_g . 20 m ERS data cover the entire imaged area. 10 m TOPSAR data (upper left image) cover most of the imaged area, including the gorge. 5 m TOPSAR data (upper right image) cover only a thin swath, including part of the gorge. . .	127
5.20	Transects for the Finke site, z_g . The transects come from column 128 of the fused estimate image in Fig. 5.19.	128
5.21	Perspective view of the fused estimates in Fig. 5.19. Values at $n_1=128$ correspond to the transects in Fig. 5.20.	129

Chapter 1

Introduction

1.1 Remotely sensing topography

Many remote sensing applications require information that is difficult to obtain with a single sensor. Individual sensors are constrained to operate in certain acquisition geometries (altitude and viewing angle) and portions of the electromagnetic spectrum. This in turn limits the coverage and accuracy of the observations. Attempts have been made to combine (*fuse*) data from different sensors, but existing methods are often empirically based and ignore differences in the measurement types. Successful data fusion becomes especially difficult when the sensors involved employ dramatically different acquisition methods, wavelengths, and resolutions. Measuring terrestrial topography is one of the many applications where this problem arises.

Large areas ($>10 \text{ km}^2$) of the Earth's surface can be easily imaged with airborne radar sensors. Radar sensors belong to the class of *active* sensors because they supply their own incident energy to illuminate the target. This capability permits day or night operation. Also, radar can be used under most cloud cover and weather conditions because there is relatively little atmospheric attenuation in the

microwave portion of the electromagnetic spectrum.

When multiple radar images are acquired by an interferometric synthetic aperture radar (INSAR) sensor, topographic data can be obtained for the imaged area. Vertical height accuracies of 2-5 m can be obtained in non-vegetated regions with airborne INSAR data processed to $10\text{ m} \times 10\text{ m}$ terrain patches (pixels), as indicated in Fig. 1.1. Many times, the actual accuracy is much lower due to processing errors, atmospheric water vapor, and complex scattering from the surface and vegetation. The presence of vegetation can lead to errors in the computed surface topography because the incident microwave energy is scattered by the vegetation as well as the surface (see Fig. 1.2).

The constituent parts of the vegetation above the surface scatter the incident microwave energy. Those scattering contributions are coherently combined with scattered energy from the ground, which results in an incorrect computed height above the ground. The high variability of the vegetation also leads to reduced correlation between the INSAR image pairs used for interferometric processing, which results in greater measurement uncertainty. In forested areas, the magnitude of this error can be tens of meters, which is significant for many applications, such as hydrological modeling, particularly in low-relief terrain. A method is needed to distinguish surface elevations and vegetation heights. One approach is to develop a functional relationship relating the INSAR observations to terrain parameters using an electromagnetic scattering model [1]. Estimating the parameters from the observations is then equivalent to inverting the model. Unfortunately, baseline constraints and the sensitivity to the input parameters often lead to unacceptably large variances of the estimates. Treuhaft and Siqueira reported RMS values of about 8 m over forested terrain [1].

Laser altimeter (LIDAR) sensors can also be used to obtain topographic measurements. LIDAR sensors are active sensors, but can only collect data when

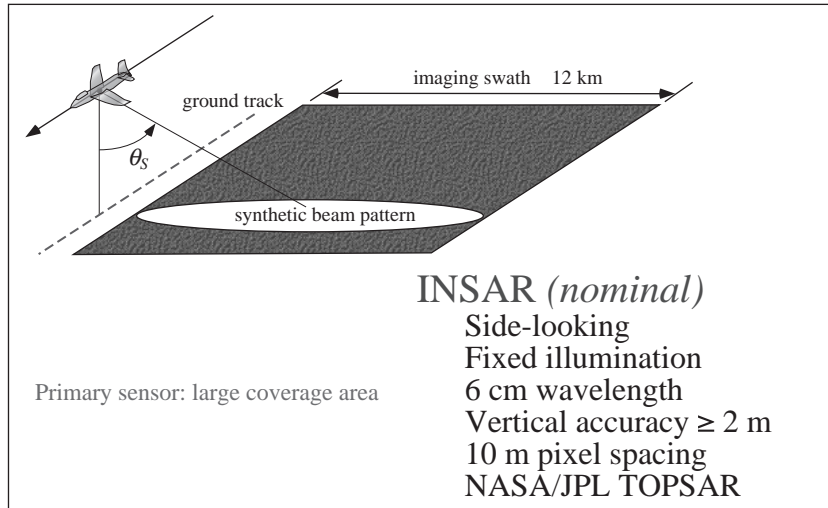


Figure 1.1: Characteristic interferometric synthetic aperture radar (INSAR) properties

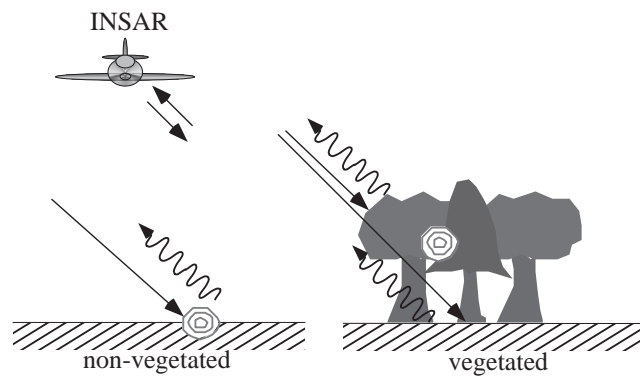


Figure 1.2: The effect of vegetation on INSAR-derived topographic measurements.

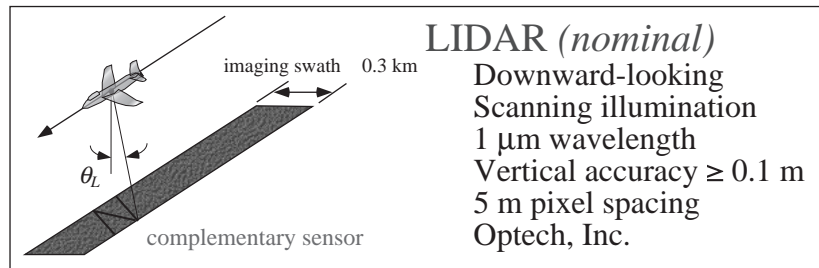


Figure 1.3: Characteristic LIDAR properties

clouds are absent and weather conditions are good since they operate in the optical portion of the spectrum. LIDAR sensors typically scan laser pulses through a small angle about the nadir direction, as depicted in Fig. 1.3. This imaging geometry, along with the small illumination footprints that are possible, allow LIDAR sensors to acquire very high resolution topographic data. Standard deviations in the measured heights of ± 10 cm are not uncommon, but imaging swaths are generally less than 1 km wide. Because of the small imaging swaths and sensitivity to cloud cover, it is not practical to use LIDAR exclusively to map large areas.

Improvements in the processing and modeling of radar data have reached a point of diminishing returns, and significant improvements in the measured topography are now most likely to be made by combining radar data with data from other sensors, such as LIDAR. In this dissertation, I present a novel framework for fusing image data from disparate sensors. The framework accommodates nonlinear relationships between the measurements and the estimated terrestrial parameters. It also allows for data collected at different resolutions, omissions in data coverage, and non-stationary terrain types.

The application of combining radar and laser topography measurements is used to demonstrate the framework. Many hydrological applications require large coverage areas and higher resolutions than INSAR can typically provide. These ap-

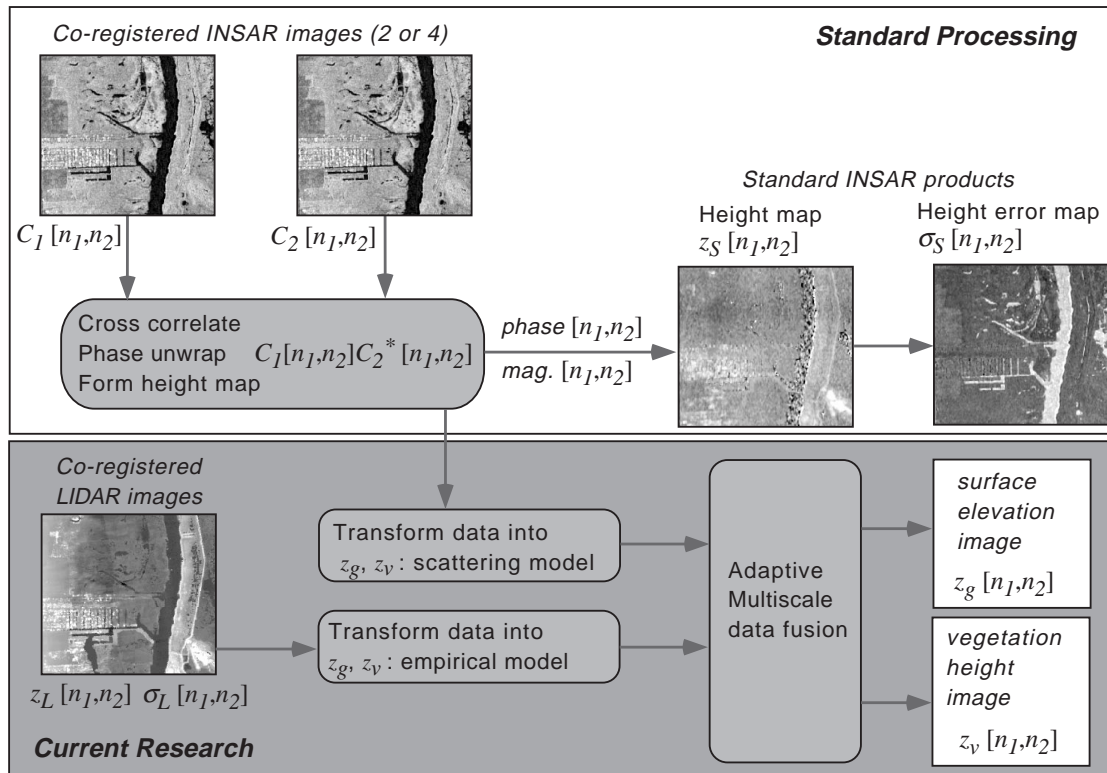


Figure 1.4: Overview of the data fusion framework

plications include monitoring wetland dynamics, finite element modeling of water catchments, and fault zones studies [2]. LIDAR data have sufficient resolution, but insufficient coverage to address these applications. These applications, as well as many others, motivate the fusion of INSAR and LIDAR data. Figure 1.4 depicts the overall data fusion framework employed in this work. I show that it is possible to improve the accuracy of the INSAR-derived estimates by combining LIDAR and INSAR data.

1.2 Adaptive multiscale estimation

The INSAR and LIDAR data are combined using a multiscale Kalman estimation approach, which provides the estimates and estimate errors at each pixel [3] [4]. I combine that data fusion method with physical modeling so that measurements that are not linearly related to the parameter estimates and have no closed-form inverse can be properly combined. However, this method is only optimal if the imaged area is wide sense stationary. In this context, stationarity implies a single terrain type with a single type of land cover. For example, a grass-covered plain with no trees or hills would yield wide sense stationary topographic measurements. If another type of terrain is present in the imagery, for example forested or hilly areas, the statistical moments of the topographic measurements will change, resulting non-stationary data. Clearly, the assumption of stationarity is a limiting one. So the data fusion method is extended to update the estimation algorithm when non-stationarities in the image data are detected. This is done by incorporating adaptive estimation techniques into the multiscale framework.

1.3 Abbreviations and terminology used

ACF: autocorrelation function

AMKS: adaptive multiscale Kalman smoothing

ALTM: airborne laser terrain mapping

CSR: center for space research

ERS: European remote sensing satellite

INSAR: interferometric synthetic aperture radar

JPL: jet propulsion laboratory

LIDAR: laser radar or laser altimeter

LMMSE: linear minimum mean squared error

MAP: maximum *a posteriori* probability

MKS: multiscale Kalman smoothing

ML: maximum likelihood

MMSE: minimum mean squared error

MSE: mean squared error

nadir: the direction directly below an observer

NASA: national aeronautics and space administration

NICC: normalized interferometric correlation coefficient

NLO: nonlinear optimization

RMS: root mean square

SAR: synthetic aperture radar

SQP: sequential quadratic programming

TOPSAR: topographic synthetic aperture radar

WSS: wide sense stationary

1.4 Thesis statement and organization

In this dissertation, I generalize a novel framework for combining data from multiple sensors and apply it to a particular problem in remote sensing involving the estimation of topography. In particular, I will defend the following thesis statement.

Statistically optimal estimates of bare-surface topography and vegetation heights are obtained by combining interferometric radar and laser altimeter measurements in an adaptive multiscale data-fusion framework that uses physical modeling to admit nonlinear measurement-state relationships.

The improved accuracy is locally sufficient for some applications, for which INSAR alone is insufficient. The ability to merge different data types also provides the capability to update data for a region without the potentially costly constraint of having to use the same sensor that first acquired the data.

The contributions of this work include (1) combining physical modeling with multiscale estimation to accommodate nonlinear measurement-state relationships, (2) extending multiscale estimation techniques to adaptively estimate non-stationary processes, and (3) improving estimates of ground elevations and vegetation heights for remote sensing applications.

Observations from different sensors must represent like quantities if they are to be properly combined, but different sensors usually do not measure the same phenomena. In this case, the observations must be transformed in some way during or prior to the fusion process. Combining data with a Kalman filter permits the fusion of quantities (state variables) that are not directly observed. If the observations are linearly related to the state variables, then the transformation can take place during the fusion process. If the observations are not linearly related to the state variables, then approximate techniques may be used to linearize the problem. These techniques require the observed process be primarily deterministic or slowly varying, and are therefore not useful for fusing topographic data. As discussed in Chapter 2, INSAR does not actually measure the height of the ground when vegetation is present. The same is true for LIDAR, although for different reasons. Therefore, to

apply Kalman based methods to the fusing of INSAR and LIDAR data, I transform each data set into estimates of bare surface elevations and vegetation heights. Using these transformations prior to the fusion allows the application of Kalman based fusion to stochastic problems with highly nonlinear measurement-state relations.

Multiscale Kalman smoothing (MKS) was developed in the early 1990s as method for multiresolution analysis and data fusion. The standard MKS algorithm allows the Kalman model parameters to vary in scale, but does not provide any mechanism for them to vary in space. Thus, there was an implicit wide sense stationary (WSS) assumption in the spatial dimensions. In general, remote sensing imagery will not be well approximated as a WSS process. An example would be an INSAR scene with forest in one half of the image and grassland in the other half. The first and second order statistical moments will be different in the two halves of the image, with the forested area having a larger mean height value and greater spatial variance. For the fused estimates to be strictly optimal, the process noise term in the Kalman filter must reflect the differences in variability in the two halves of the image. I generalize the linear state model so that I can employ an adaptive estimation technique developed for 1-D temporal estimation in the multiscale framework. The result is an adaptive MKS (AMKS) algorithm that adapts to non-stationarities in images.

The primary motivation of this research is to produce improved estimates of bare surface topography and vegetation heights for remote sensing applications. Initial estimates of bare surface topography and vegetation heights are obtained through modeling INSAR and LIDAR interactions with the terrain. Those estimates are then combined in a multiscale data fusion approach. Unlike previous multiscale data fusion methods, this estimator remains nearly optimal (in the mean squared error sense) even when the imaged terrain contains different terrain types. The mean squared errors of the final estimates are generally smaller than those obtained

by other methods. Yet, this method avoids making drastic assumptions about the vegetation medium that other methods often make.

This dissertation is organized in the following manner. I characterize INSAR and LIDAR data and methods for estimating bare-surface topography using INSAR data in Chapter 2. In Chapter 3, I present an existing multiscale estimation approach and some important properties of the Kalman filter. I am able to accommodate nonlinear measurements in the Kalman filter using results from Chapter 2. I extend the multiscale estimation framework in Chapter 4 to implement spatially adaptive multiscale estimation. In Chapter 5, I give results from the non-adaptive and adaptive methods and demonstrate fusion of more than two data sets. The algorithms are applied to three distinctly different terrain types to demonstrate the flexibility of the framework. Finally, I present conclusions in Chapter 6.

Chapter 2

Characterizing Topographic Measurements

2.1 Introduction

Topography of land surfaces can be calculated using observations from INSAR and LIDAR sensors. These nominal values can then be used, along with knowledge of how the incident energy interacts with the vegetation, to estimate bare surface elevations and vegetation heights independently from each sensor. The standard procedure for computing topography from INSAR data is presented in Section 2.2. In Section 2.3, methods for estimating bare surface elevations and vegetation heights from INSAR are discussed. LIDAR measurements are described in Section 2.4. An empirical algorithm for obtaining bare surface elevations and vegetation heights from LIDAR data is also discussed. Finally, conclusions are given in Section 2.5.

2.2 The INSAR measurement

Substantial signal processing must be applied to raw SAR data to form images with spatial fidelity. The sequence of operations is collectively referred to as *SAR processing* and is not addressed in this dissertation. Many references treat the subject in detail, including [5] [6] [7]. Discussion of SAR data in this dissertation assumes that the complex-valued images, denoted by C , have already been generated through SAR processing.

2.2.1 Obtaining height estimates

Interferometric processing requires two SAR observations. The two observations are either acquired by a single-antenna sensor passing over the target area twice in a technique known as *repeat-pass* INSAR, or by a dual-antenna sensor passing over once in a technique known as *single-pass* INSAR. Repeat-pass INSAR suffers from an effect known as temporal decorrelation. This phenomenon is caused by the scattering elements in the imaged area, such as tree branches, changing position or orientation on the scale of the SAR wavelength during the intervening time between acquisitions. Single-pass INSAR does not suffer from temporal decorrelation, but is usually constrained to small antenna separation distances (baselines). The difficulty in implementing baselines longer than a few meters on an aircraft owe to the INSAR requirement that the baseline remain very stable during acquisition.

The majority of INSAR data used in this dissertation are single-pass INSAR data acquired at C-band ($\lambda=5.6$ cm) by the NASA/JPL topographic SAR (TOP-SAR) sensor. Repeat-pass INSAR data from the European Remote Sensing Satellite (ERS-1) are also used. The source of INSAR data is indicated in each application.

Fig. 2.1 shows the INSAR imaging geometry. Single-pass INSAR sensors acquire two complex-valued SAR images C_1 and C_2 as the sensor passes over the

target. An interferometric phase difference ϕ is determined from the two images C_1 and C_2 , and is used to compute topographic heights. The relative phase is equal to the phase difference between C_1 and C_2 . The absolute phases Ψ_1 and Ψ_2 of C_1 and C_2 , respectively, are not known, so ϕ must be computed from the complex cross-correlation of C_1 and C_2 .

$$\begin{aligned}
C_1 &= A_1 e^{\Psi_1} \\
C_2 &= A_2 e^{\Psi_2} \\
C_1 C_2^* &= A_1 e^{\Psi_1} A_2 e^{-\Psi_2} \\
&= A_1 A_2 e^{\Psi_1 - \Psi_2} \\
\angle C_1 C_2^* &= \Psi_1 - \Psi_2 = \phi \pmod{2\pi}
\end{aligned} \tag{2.1}$$

The $\angle C_1 C_2^*$ phase image is modulo 2π , and is called the wrapped phase image or the interferogram. The modulo 2π image undergoes 2-D phase unwrapping, and a constant phase offset is added to obtain the absolute phase ϕ image [8]. Path length is related to the unwrapped phase through the propagation speed of light and the wavelength. Equation (2.2) describes ϕ as a function of the path length difference from the SAR antennas to each resolution cell on the ground [9] [10].

$$\phi = \frac{2\pi}{\lambda} \delta (\rho_{S1} - \rho_{S2}) \tag{2.2}$$

Here, λ is the wavelength, δ equals 1 or 2 depending on the SAR transceive mode, and ρ_{S1} and ρ_{S2} denote path lengths from the antennas to the ground, as indicated in Fig. 2.1.

In practice, the phase is computed using INSAR system parameters as shown

in (2.3) because they are more precisely known than the path lengths ρ_{S1} and ρ_{S2} :

$$\phi \approx \frac{-2\pi}{\lambda} b \sin(\theta_S - \alpha) \quad (2.3)$$

Here, b is the electrical baseline length, θ_S is the incidence angle, and α is the baseline angle from horizontal. The electrical, or effective, baseline length equals δ times the physical baseline length. The expression in (2.3) is a first-order approximation for ϕ , which is commonly used for clarity. Terms involving the square of the baseline have been neglected because $b \ll \rho_{S1}, \rho_{S2}$.

Using (2.3) to compute θ_S , one can then solve for the topographic height z_S and horizontal location y_S of each pixel using (2.4).

$$\begin{aligned} z_S &= h_S - \rho_{S1} \cos \theta_S \\ y_S &= \rho_{S1} \sin \theta_S \end{aligned} \quad (2.4)$$

Here, h_S is the INSAR altitude, and ρ_{S1} is the path length from the reference antenna to the resolution cell. Equations (2.2)-(2.4) assume that the SAR illumination is perpendicular to its velocity vector. The INSAR data used in this dissertation were previously processed to yield z_S images. The raw TOPSAR data were processed at the Jet Propulsion Laboratory (JPL) and the ERS complex images were processed at the Center for Space Research (CSR).

In addition to the correlation phase of the image pairs $\angle C_1 C_2^*$ used for computing z_S , the correlation coefficient

$$\gamma = \frac{|\langle C_1 C_2^* \rangle|}{\sqrt{\langle |C_1|^2 \rangle} \sqrt{\langle |C_2|^2 \rangle}} \quad (2.5)$$

is used to compute the uncertainty in the calculated value of z_S [10]. Here, $\langle \rangle$ denotes an ensemble average. Values of γ close to unity represent small standard deviations of the measured phase, which correspond to low uncertainty in z_S .

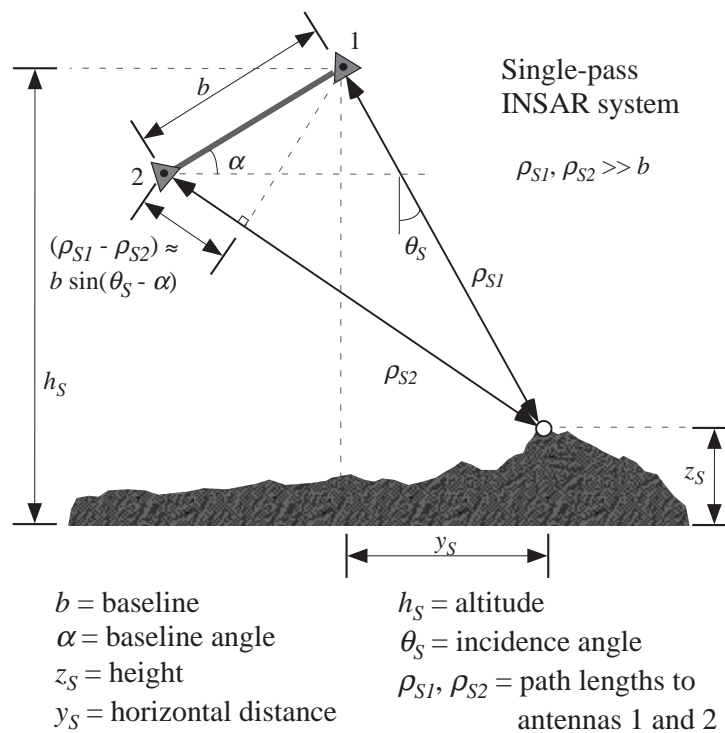


Figure 2.1: INSAR imaging geometry.

Both systematic and random sources of decorrelation exist. Most potential sources for systematic decorrelation occur during acquisition or INSAR processing. These include registration errors, uncompensated motion of the SAR platform during acquisition, and multipath at the SAR receiving antenna. These sources are best dealt with during INSAR processing and are not addressed in this dissertation. Modeling the sources of random errors leads to a calculation for the uncertainty in z_S in the following section.

2.2.2 Random error sources

The most commonly modeled sources of decorrelation due to random errors are temporal, spatial, and thermal [11]. As stated earlier, temporal decorrelation is due to random changes in position or orientation of scattering elements between acquisitions. It does not affect single-pass INSAR data. Spatial decorrelation is inherent to the INSAR measurement because each antenna views the target from a slightly different angle, resulting in slightly displaced impulse response functions from the target. Spatial decorrelation is computed as a function of baseline length and incidence angle and does not account for different types of scattering at the target. Thermal decorrelation arises from the thermal noise in the SAR instrument. These sources are assumed to be independent, thus their total contribution ρ_{total} is the product of their individual contributions.

The phase standard deviation σ_ϕ , sometimes called the phase error, will depend on ρ_{total} and the number of independent phase measurements. When the complex images are formed, σ_ϕ is usually too large for successful registration and unwrapping. The most common procedure to reduce the phase error, at the expense of spatial resolution, is to average adjacent pixels in the original complex image [5] [9]. This procedure is known as coherent multilook processing, or simply coherent

multilooking. The original complex image is referred to as a single-look complex (SLC) image, and the averaged image is referred to as the multilook image. The phase error is reduced as N_{look} increases, where N_{look} is number of resolution cells averaged together.

The phase error was computed as a function of ρ_{total} and N_{look} in [9] via Monte Carlo simulations. The reduction is nonlinear but monotonic in N_{look} with most of the reduction realized by $N_{look} = 4$ [12]. Thus, for $N_{look} \geq 4$, the Cramer-Rao bound is typically used to estimate σ_ϕ from the γ images using [10]

$$\sigma_\phi = \frac{1}{\sqrt{2N_{look}}} \frac{\sqrt{1 - \gamma^2}}{\gamma} \quad (2.6)$$

We can then compute height uncertainty σ_h from phase uncertainty σ_ϕ using an expression derived in [12], which is obtained by differentiating (2.4) with respect to ϕ , retaining all higher order terms, and assuming zero ground slope:

$$\sigma_h = \frac{\lambda \rho_{S1}}{2\pi b} \frac{\sin \theta_S}{\cos(\alpha - \theta_S)} \sigma_\phi \quad (2.7)$$

This equation serves as a lower bound on the height standard deviation due to phase noise. In the treatment of Kalman filtering in Chapter 3, the measurement error variance of the INSAR heights z_S is taken to be the square of σ_h .

Because σ_h only accounts for temporal, spatial, and thermal height errors, any vegetation that is present introduces additional height errors on the order of the vegetation height. This phenomenon occurs because the received signal contains scattering contributions from the vegetation above the ground as well as from the ground itself. The computed elevation z_S will generally be greater than the elevation of the bare ground z_g and less than the elevation of the vegetation canopy z_v (see Fig. 2.2). However, the correlation γ usually decreases, relative to the backscattered power, over diffuse targets, such as vegetation. This effect is known as speckle decorrelation. Thus, the actual height error is dependent on the vegetation height,

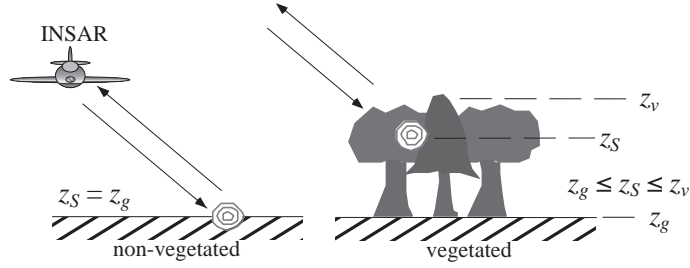


Figure 2.2: The effect of vegetation on the INSAR measurement. z_g is the true ground height, z_S is the nominal INSAR-derived surface height, and z_v is the true vegetation height. Neglecting noise, z_S will approximately equal z_g where there is no obscuring vegetation. Where vegetation is present, z_S will be between z_g and z_v .

density, and structure. This source of phase error is ignored in standard INSAR processing.

A single INSAR observation is not sufficient to resolve this ambiguity in z_S due to vegetation. An additional observation is needed to separate z_g and z_v . The TOPSAR sensor is capable of transmitting and receiving with each of its two C-band antennas and sequencing the transmit and receive events to obtain two different INSAR acquisitions. The procedure yields different relative path lengths corresponding to two different electrical baselines, as shown in Fig. 2.3. The result is two different complex image pairs C_1C_2 and C_3C_4 . The cross-transceive mode yields image pair C_1C_2 and $\delta = 1$ in (2.2). The co-transceive mode yields image pair C_3C_4 and $\delta = 2$ in (2.2).

The two effective baselines are quite similar, $b_1=2.5$ m and $b_2=5.0$ m. Thus, the two INSAR images will be highly correlated globally. However, the slight difference in viewing angle yields small, random differences in the coherent scattering contributions in each resolution cell in the SAR image. Corresponding pixels in the two INSAR scenes will have similar values, but random differences will remain, which allows the images to be used as independent observations [1]. This method

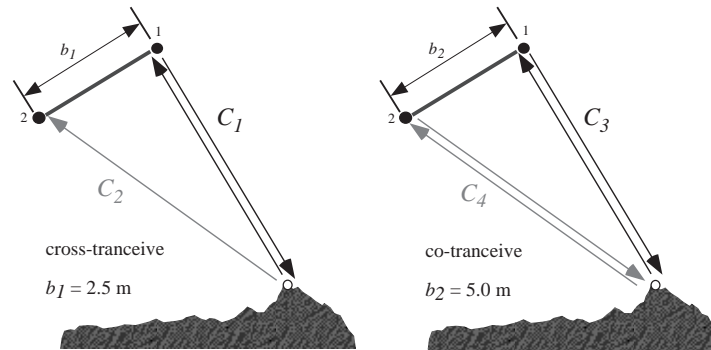


Figure 2.3: INSAR acquisition with dual-baseline system. Different complex image pairs (C_1C_2 or C_3C_4) are formed depending on the transmit/receive sequences used. The procedure yields two independent INSAR acquisitions.

of data collection is referred to as *dual baseline* because two INSAR images are acquired in a single-pass.

2.3 Estimating surface and vegetation heights with INSAR

A scattering model that relates INSAR measurements to the true surface and vegetation heights is now presented. It is shown that the minimal set of terrain parameters that must be estimated requires at least two INSAR observations. Two nominal INSAR acquisitions would work in principle, but the temporal decorrelation would limit their use. The dual-baseline acquisition described in Section 2.2 is used because it provides two INSAR observations without temporal decorrelation. The scattering model is inverted numerically to obtain estimates of z_g and $\Delta z_v (=z_v-z_g)$.

2.3.1 Empirical approaches

Recent attempts to account for the effects of vegetation based on regression and empirical modeling have achieved limited success when applied to repeat-pass INSAR data [13] [14]. In [13], interferometric phase, correlation, and single-polarization backscatter data were classified into general land cover types using INSAR data from the first European Remote Sensing (ERS-1) satellite. The results were affected by temporal decorrelation caused by the changes in the scene between acquisitions. Accurate determination of the vegetation heights suffered because temporal decorrelation can mimic low correlation due to scattering from vegetation. Also, empirical methods are necessarily particular to the data used to develop the relationship between the measurements and the estimated parameters. For example, the work in [14] assumed that the density of the vegetation cover was known and constant. This is an unrealistic assumption, and can be avoided if a formal scattering model is used, rather than empirical relationships, to relate the INSAR observations to key terrain parameters.

Another approach that has been investigated for improved INSAR height determination is the use of multiple frequencies [15]. In principle, this approach could be applied to the problem of estimating bare-surface and vegetation heights because as the wavelength of the SAR increases, it will penetrate deeper into the vegetation layer. This occurs because electromagnetic waves do not scatter significantly when they encounter objects that are much smaller than the wavelength. Thus short wavelengths, e.g. X-band ($\lambda = 2$ cm) will scatter strongly off of the tops of the vegetation canopy, while longer wavelengths, e.g. P-band ($\lambda = 65$ cm) will penetrate into the canopy. The potential exists to image an area with both high and low frequencies and subtract the computed heights from each to arrive at a vegetation height. The height of the bare surface would be the height computed from the

low-frequency SAR. However, this technique alone is not sufficient to address the problem because, as seen in (2.7) the height error is directly proportional to wavelength. The resulting order of magnitude increase in σ_h from X-band to P-band will lead to errors on the same scale as the vegetation signal that is being estimated.

2.3.2 Scattering model approach

Backscattered SAR signals from vegetated terrain have been successfully simulated using a variety of scattering models. Coherent wave models, which compute the approximate scattered electric field from layers of discrete, randomly oriented scatterers over a ground surface, have been particularly successful [16] [17] [18] [19]. These models solve for a real-valued measure of the scattered power, but they are coherent because the phase is retained internally for the calculation of coherent interactions among the scattered waves. These models are used to calculate the normalized SAR backscattering cross-section, or backscattering coefficient σ^o , for different polarizations. Other approaches, based on radiative transfer theory, have also been used to simulate the SAR response to vegetation covered terrain [20], but they do not account for coherent surface-vegetation interactions [21]. The approach used in this dissertation is based on a coherent wave model.

These nonlinear models typically require detailed knowledge of the statistical and electrical properties of the scatterers to compute relatively few output parameters, such as the strength of the scattered field at the principal linear polarizations. For example, a scattering model that solves for components of σ^o from the dominant scattering mechanisms of the ground and vegetation requires information on the size, amount, orientation, and dielectric constant of the constituent vegetation elements and roughness and dielectric constant of the ground. In a typical implementation, at least nine input parameters must be specified. In addition to the

large number of inputs that must be specified, the solutions are not unique. Different combinations of inputs can lead to very similar SAR responses. Attempts to invert this type of model directly will therefore be ill posed.

In addition, we need a model that is formulated to calculate INSAR observations rather than SAR observations. In an effort to separate bare-surface and vegetation heights from the height of the phase scattering center, [22] investigated what phenomena would affect the INSAR data and contain useful information for separating bare-surface and vegetation heights. It was found that INSAR observations are sensitive to vertical distributions of scatterers above the surface, especially to the depth and density of the layer of scatterers. Vegetation can be represented by such a layer of scatterers, just as in the coherent SAR models.

Fig. 2.4 shows the effect of vertically distributed scatterers (vegetation) on the INSAR observations. As the height of the scattering layer increases, (a) the differential path length $\Delta\rho$ increases, (b) the magnitude of the cross-correlation $|C_1 C_2^*|$ decreases relative to the backscattered power at each antenna $C_1 C_1^*$ and $C_2 C_2^*$, and (c) the backscattered power increases. Thus, it is possible to relate INSAR data to the height of the vegetation layer, which leads to estimating the height of the vegetation over the bare surface.

A coherent wave model was re-parameterized in [1] to reduce the number of inputs to three key terrain parameters: ground elevation z_g (m), vegetation height Δz_v (m), and extinction coefficient of the vegetation τ (Np/m), which is related to density. The model \mathcal{M} relates the terrain parameters $\{z_g, \Delta z_v, \tau\}$ to the normalized interferometric cross correlation (*NICC*).

The scattering model, which is derived in [1], is

$$NICC \approx \mathcal{M}(\mathbf{x}) = \frac{2\tau A_r \exp(j\phi)}{\left(\exp\left(\frac{2\tau\Delta z_v}{\cos\theta_S}\right) - 1\right) \cos\theta_S} \int_0^{\Delta z_v} \exp(j\alpha_z \zeta) \exp\left(\frac{2\tau\zeta}{\cos\theta_S}\right) d\zeta \quad (2.8)$$

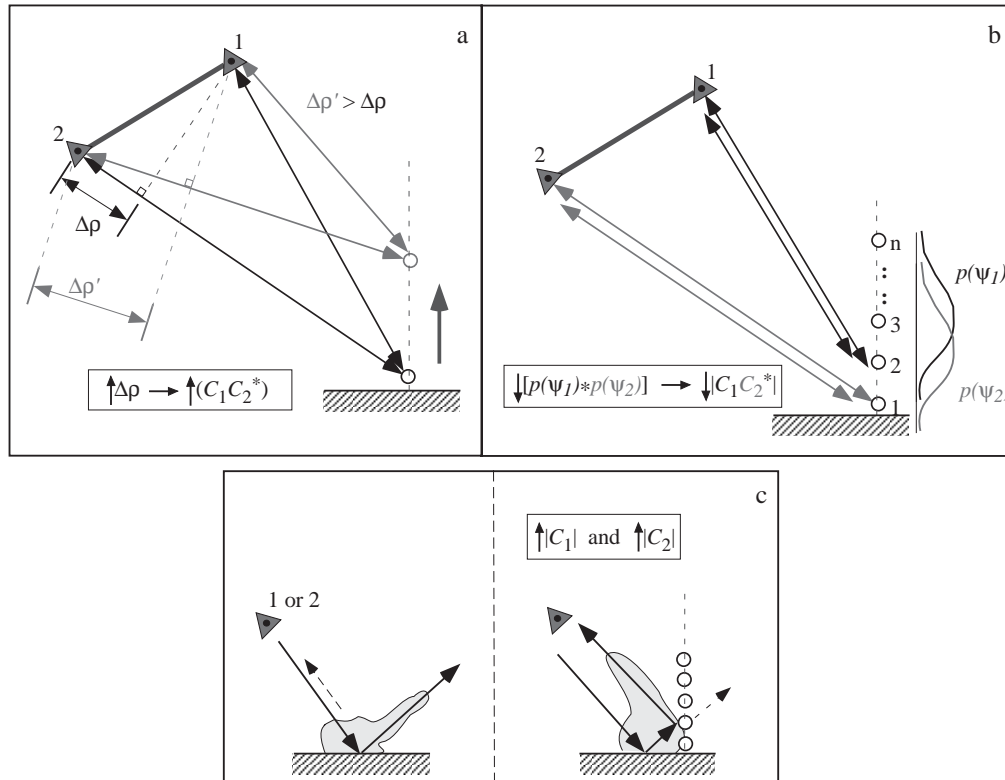


Figure 2.4: Effects of vertically distributed scattering elements (representing vegetation) on the INSAR measurement. As the height of the scattering elements increases, (a) the differential path length $\Delta\rho$ increases, (b) the magnitude of the cross-correlation $C_1 C_2^*$ decreases relative to the backscattered power $C_1 C_1^*$ and $C_2 C_2^*$, and (c) the backscattered power increases.

where α_z is the derivative of interferometric phase with respect to the vertical direction, ζ is a generalized height variable for integration, and A_r is a normalized correlation amplitude factor that represents the contribution to the magnitude of *NICC* (decorrelation effect) from the surface. It accounts for the effects of spatial decorrelation inherent in all INSAR measurements. The phase contribution from the ground ϕ is related to z_g through (2.3) and (2.4).

Equation 2.8 derives from considering a collection of surface and volume scatterers in a single SAR resolution cell. The scattered field from a single scatterer, multiplied by the probability of a scatterer at that location, is integrated over the volume of the vegetation layer yielding an expected value of the scattered field due to a single scatterer in the layer. The sum of all such integrals over particles in the volume yields the total volume scattering contribution. A similar procedure is followed for the integral over the surface patch in the resolution cell to yield the surface scattering contribution. In the derivation in [1], the summations are replaced by aggregate terms such as the density of scatterers in the layer times the contribution of a single scatterer. The probability density functions are written in terms of SAR parameters including range (perpendicular to flight direction) and azimuth (parallel to the flight direction) resolution functions. Using the coherent wave approach of [19], the general expressions for the scattered field received at each antenna are replaced by scattering amplitude functions and phasor notation. The resulting expressions for C_1 and C_2 are multiplied together, after taking the complex conjugate of C_2 , and normalized to arrive at (2.8).

From (2.8), the integral over vegetation height has a simple analytical solution $\int e^{\beta u} du = e^{\beta u} / \beta$, but is left in integral form to emphasize the contribution of the vegetation to the phase. The factor that multiplies the integral consists of an amplitude and phase term. The amplitude describes the power of the scattered wave as a function of free-space distance from the antennas to the resolution cell,

the distance traveled within the vegetation medium $\Delta z_v / \cos(\theta_S)$, and the effective density of the vegetation medium τ . The phase term $\exp(j\phi)$ describes the absolute phase associated with the distance from antenna 1 to the intersection of the resolution cell and the ground plane. The integral over the vegetation describes the propagation in the vegetation effective medium, with terms for the phase delay due to the path length in the medium and attenuation due to the permittivity (via extinction coefficient) of the effective medium.

By rewriting the model after evaluating the integral and rearranging terms,

$$NICC = A_r e^{j\phi} \left[\frac{2\tau}{j\alpha_z \cos \theta_S + 2\tau} \right] \left[\frac{e^{j\alpha_z \Delta z_v} e^{\frac{2\tau \Delta z_v}{\cos \theta_S}} - 1}{e^{\frac{2\tau \Delta z_v}{\cos \theta_S}} - 1} \right] \quad (2.9)$$

We can now look at the limiting case of no vegetation. Taking the limit of (2.15) as $\Delta z_v \rightarrow 0$ directly yields an indeterminate $0/0$ form. However, since all functions of Δz_v are continuous and differentiable, L'Hôpital's Rule can be used to determine the limit. Taking the partial derivatives with respect to Δz_v of the numerator and denominator:

$$A_r e^{j\phi} \left[\frac{2\tau}{j\alpha_z \cos \theta_S + 2\tau} \right] \left[\frac{\left(j\alpha_z + \frac{2\tau}{\cos \theta_S} \right) e^{\left(j\alpha_z + \frac{2\tau}{\cos \theta_S} \right) \Delta z_v}}{\left(\frac{2\tau}{\cos \theta_S} \right) e^{\left(\frac{2\tau}{\cos \theta_S} \right) \Delta z_v}} \right] \quad (2.10)$$

By L'Hôpital's Rule, the limit of (2.10) as $\Delta z_v \rightarrow 0$ equals the limit of $NICC$:

$$\lim_{\Delta z_v \rightarrow 0} NICC = A_r e^{j\phi} \quad (2.11)$$

The result is a normalized equivalent to the standard INSAR result, which disregards vegetation effects. Here, ϕ is the phase $\angle C_1 C_2^*$ after phase unwrapping and absolute referencing. $A_r = \gamma$, normalized by the zero-baseline correlation coefficient. In Fig. 2.5, $NICC$ is plotted as a function of Δz_v for a set of τ values while holding z_g constant to examine its behavior as vegetation height increases. $NICC$ is slowly varying and monotonic over the range of physically reasonable values of

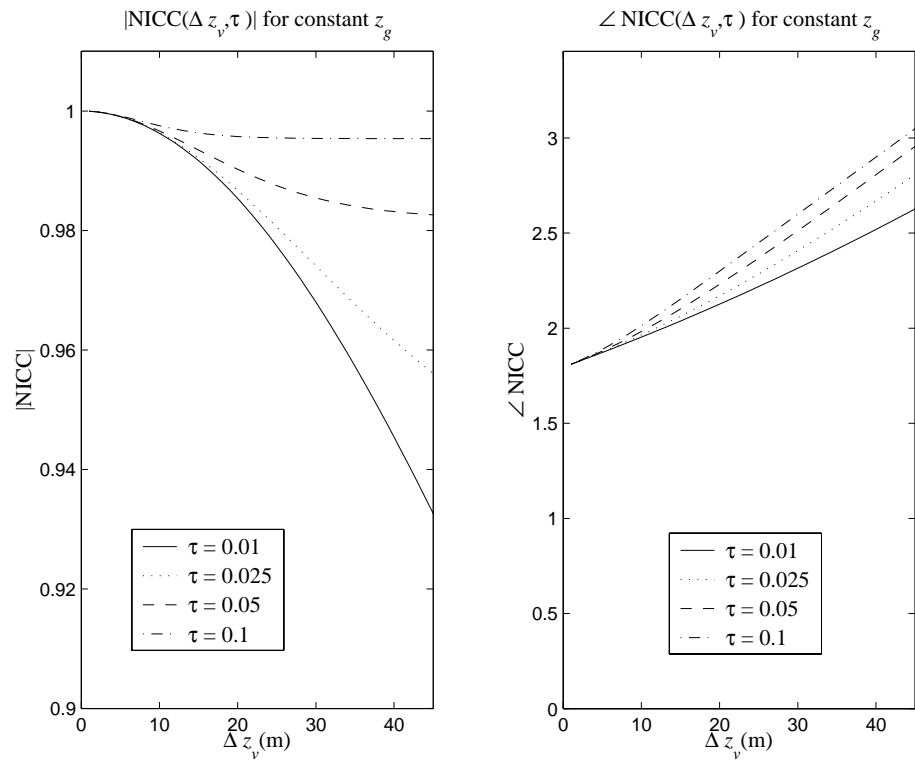


Figure 2.5: Sensitivity of NICC to Δz_v and τ .

Δz_v and τ . As the extinction coefficient τ increases, the vegetation appears more opaque to the INSAR. This is the reason $\angle NICC$ rises faster as a function of Δz_v for larger values of τ . The trends in Fig. 2.5 agree with the expected behavior of $NICC$ shown in Fig. 2.4.

Before INSAR data from the two TOPSAR baselines are used in Section 2.3.3 to estimate terrain parameters, it is important to demonstrate the sensitivity of $NICC$ to the differences in TOPSAR baselines $b_1 = 2.5$ m and $b_2 = 5.0$ m. Figure 2.6 shows how $NICC$ varies for baselines between 1 and 5 meters. The magnitude of $NICC$ ranges between about 0.98 and 0.92. This is comparable to the range in $|NICC|$ values as a function of Δz_v in Fig. 2.5. The phase of $NICC$ varies as the ratio of baseline length to wavelength. With $1 \leq b \leq 5$ and $\lambda=0.056$ m, $\angle NICC$ will range through several cycles of 2π radians. Thus, the unwrapped $\angle NICC$ data will be quite sensitive to changes in baseline length on the order of 1 m or larger. It is apparent that even for the relatively small baseline differences available on a single pass of the TOPSAR sensor, $NICC$ data acquired at the two baselines will comprise two independent measurements of the terrain.

\mathcal{M} is shown to be approximately equal to the $NICC$ because it assumes zero ground slopes, a homogeneous vegetation layer (per resolution cell), and uniformly distributed orientations of the scattering elements. Non-zero ground slopes and preferential orientations of the vegetation are considered in [1], but they require the addition of fully polarimetric SAR data, which are not generally available coincident with TOPSAR data.

2.3.3 Inverting the scattering model

The $NICC$ is computed from two complex SAR images for each baseline. Thus, dual-baseline INSAR provides two complex-valued (four real-valued) observations,

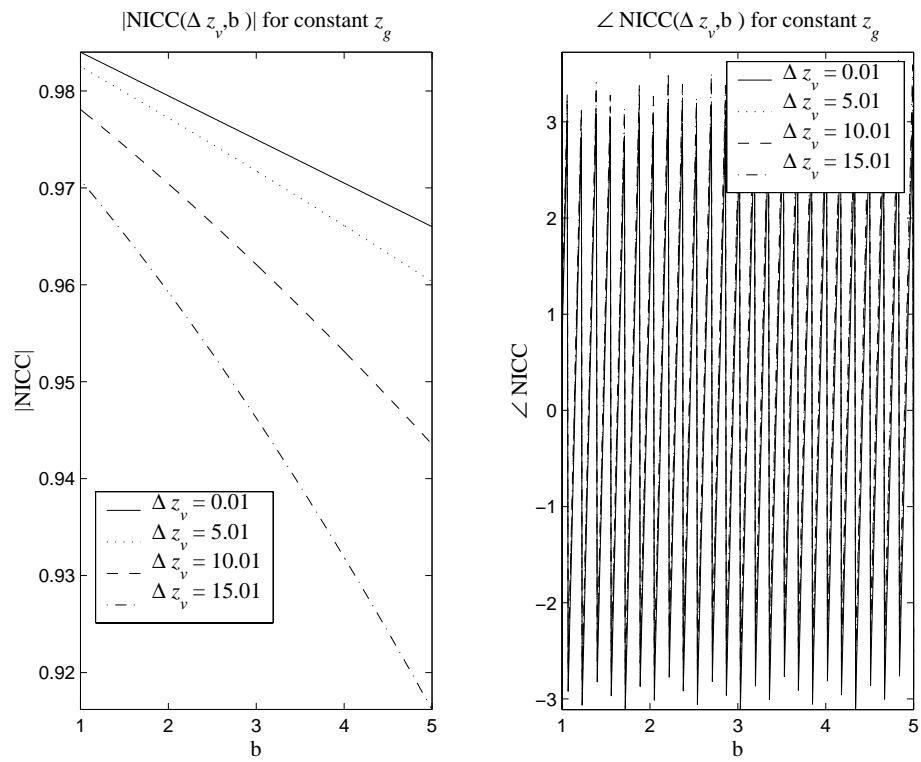


Figure 2.6: Sensitivity of \mathcal{M} to baseline distance for different values of Δz_v .

which are sufficient to solve for the three real-valued terrain parameters at each pixel numerically. The model \mathcal{M} for a single baseline relates the terrain parameters \mathbf{x} to the $NICC$, where $\mathcal{M}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$. The observations for both baselines are represented by the vector \mathbf{d}

$$\mathbf{d} = \begin{bmatrix} |NICC|_{b1} \\ \angle NICC_{b1} \\ |NICC|_{b2} \\ \angle NICC_{b2} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} z_g \\ \Delta z_v \\ \tau \end{bmatrix} \quad (2.12)$$

The elements of the vector \mathbf{M} represent the magnitude and phase components of $\mathcal{M}(\mathbf{x})$ for each baseline.

$$\mathbf{M} = \begin{bmatrix} |\mathcal{M}(\mathbf{x})|_{b1} \\ \angle \mathcal{M}(\mathbf{x})_{b1} \\ |\mathcal{M}(\mathbf{x})|_{b2} \\ \angle \mathcal{M}(\mathbf{x})_{b2} \end{bmatrix} \quad (2.13)$$

$$\mathbf{d} = \mathbf{M} \quad (2.14)$$

The $|NICC|$ data for each baseline are obtained from correlation magnitude data supplied by the TOPSAR processor and then normalized by zero-baseline magnitudes to correct for the decorrelation effect of large incidence angles [1]. The $|NICC|$ for zero baseline is nominally equal to unity. A zero-baseline $|NICC|$ that is less than unity signifies an uncompensated source of decorrelation. Thus, normalizing by the zero-baseline $|NICC|$ removes the effects of uncompensated sources of decorrelation. The $\angle NICC$ data are computed by scaling the z_S data for the two non-zero baselines by the derivative of interferometric phase with respect to the vertical. All INSAR data are ground range projected [5] before computing $NICC$ images.

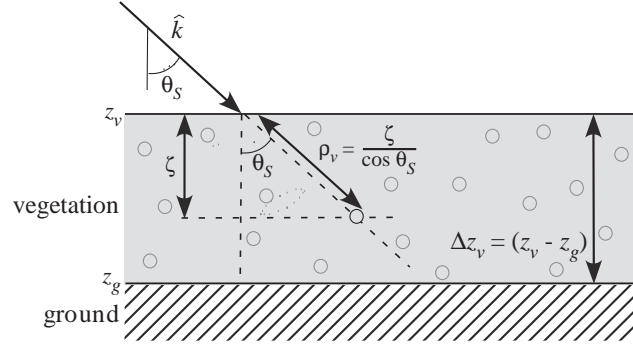


Figure 2.7: The modeled vegetation medium, which consists of a layer of randomly distributed scattering particles. Here, \hat{k} denotes the direction of the incident wave from the INSAR, and ρ_v denotes the path length in the vegetation medium.

Even after parameterizing the coherent scattering model to obtain (2.8), there is no closed-form inverse. Therefore a numerical inversion approach is required. The numerical inversion in [1] is reformulated as a constrained nonlinear optimization problem to solve for z_g and Δz_v . The objective function is the square of the L_2 -norm between the actual observations and the modeled observations, given by

$$\min \|\mathbf{M} - \mathbf{d}\|_2^2 \quad \text{subject to } \mathbf{x} \in \mathcal{X} \subset \mathbb{R}^3 \quad \{0 < \theta_S < \frac{\pi}{2}\}. \quad (2.15)$$

The objective function and inequality constraints in (2.15) are real-valued and twice differentiable on the feasible region denoted by the set \mathcal{X} , so the optimization is implemented as a sequential quadratic programming problem (SQP) [23]. The search space is smooth because the scattering model is highly simplified. However, this technique could be used to provide good starting guesses when inverting more sophisticated scattering models that yield less well-characterized search spaces. The value of σ_h is scaled by the ratio $\text{Var}(z_g)/\text{Var}(z_S)$ and $\text{Var}(\Delta z_v)/\text{Var}(z_S)$ to obtain the uncertainty associated with the transformed INSAR data z_g and Δz_v .

2.3.4 Computational complexity

The nonlinear constrained optimization is performed for each pixel. Over the range of physically reasonable values of the vegetation extinction coefficient τ [24], the objective function is a weak function of τ , which leads to very low-gradient regions in the search space. As a result, on the order of 100 SQP iterations can be required to converge to a solution for some pixels. This nonlinear optimization is the most computationally expensive step in the data fusion framework. Because the physical modeling is implemented before the actual data fusion however, other methods of obtaining initial estimates of the terrain parameters from the INSAR data could be used. As mentioned in Section 2.3.1, other methods exist that determine bare-surface elevations indirectly, thus avoiding complex scattering models and inverse problems that require nonlinear optimization techniques. The scattering model approach was used in this work because it is theoretically based, but it is not fundamental to the overall data fusion framework.

2.4 LIDAR measurements

The LIDAR data used in this work were acquired by an Optech airborne laser terrain mapping (ALTM) sensor, operating at a near-infrared wavelength ($\lambda=1047$ nm). The LIDAR is a nadir-looking sensor that scans a narrow swath about the airplane's ground track. The footprint size, spacing, and swath width can be adjusted by varying the airplane's altitude and the laser's repetition rate. The data used here came from two different acquisitions. Each acquisition had a footprint size of 15 cm, spacing of 1-2 m, and swath width of 600 m. The time-tagged laser pulses were gridded into regular image lattices of LIDAR heights z_L . The LIDAR data acquired over the Texas coastal site were gridded to postings of $10 \text{ m} \times 10 \text{ m}$ with about

15 pulse returns per pixel. The LIDAR data over Austin, Texas were acquired at a higher sampling rate and gridded to $1\text{ m} \times 1\text{ m}$ with 20 pulse returns per pixel. Because multiple pulses correspond to each pixel in the LIDAR images, standard deviations σ_L were calculated at each pixel.

LIDAR heights were compared to kinematic global positioning system (GPS) data collected along roads in the study area. The GPS data were acquired by a Trimble 4000SSi system, which has an accuracy of approximately $\pm 0.10\text{ m}$. The LIDAR data agree with the GPS bare surface heights to within $\pm 0.15\text{ m}$. *In situ* measurements of tree heights confirmed that the LIDAR heights are generally within $\pm 1\text{ m}$ of the vegetation heights. Essentially, LIDAR measures vegetation heights everywhere and measures surface elevations where vegetation is negligible. For this work, it is assumed the LIDAR measures z_v directly because the short wavelengths of the optical portion of the spectrum do not penetrate the foliage. Because of the small footprint, however, some penetration is possible among multiple pulses, and pulses may even reach the ground through gaps in the canopy (see Fig. 2.8). Hence, some processing is required to compute z_g at every pixel in the LIDAR data.

An empirical method for removing vegetation heights in LIDAR data from [25] was used to transform LIDAR data over the study areas into estimates of z_g and $\Delta z_v (= z_v - z_g)$. An empirical approach was used because no widely applicable LIDAR scattering models for vegetation exist. Slatton, Crawford, and Evans [26] found σ_L to be an effective indicator of vegetation in LIDAR data. The approach taken in [25] is to threshold σ_L to isolate pixels that contain negligible vegetation. Heights are computed for the thresholded pixels using interpolation. The risk of interpolating across significant topographic features is small except in the uncommon case of very dense vegetation covering a surface with high topographic relief. Estimates of z_g and Δz_v are obtained from LIDAR processed at CSR, and used as inputs to the data fusion algorithm. Equation (2.16) summarizes this process math-

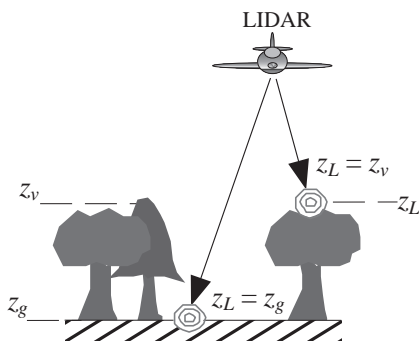


Figure 2.8: Effect of vegetation on the LIDAR measurement. z_g is the true ground height, z_L is the nominal LIDAR-derived surface height, and z_v is the true vegetation height. Neglecting noise, z_L will approximately equal z_g where there is no obscuring vegetation. Where vegetation is present, z_L will approximately equal z_v .

ematically, where N is the set of all pixels in the LIDAR image and (n_1, n_2) denote pixel coordinates. The measurement uncertainty in the LIDAR data is taken to be $\sigma_L(n_1, n_2)$.

$$z_v = z_L \quad \forall (n_1, n_2) \in N$$

$$z_g = \begin{cases} z_L & \forall (n_1, n_2) \in N | \sigma_L(n_1, n_2) < \text{threshold} \\ \text{linear interpolation, otherwise} \end{cases} \quad (2.16)$$

All operations on the LIDAR data are non-iterative. The most computationally demanding step is the gridding of the raw, irregularly-spaced LIDAR points into a uniformly-spaced LIDAR image, requiring $\mathcal{O}(N)$ operations where N is the total number of raw LIDAR points. Under nominal acquisition conditions, most LIDAR data gridded to 5 m resolution will map approximately 100 raw points to a single image pixel. The reduction of data by this factor reduces the number of all subsequent operations on the LIDAR imagery by the same factor.

2.5 Conclusions

I describe fundamental INSAR measurements and geometry and present an electromagnetic scattering model that relates a critical set of terrain parameters to the INSAR observations. I give the dependence of the scattering model on the terrain parameters and a method of numerically inverting the model to obtain estimates of the terrain parameters. I describe LIDAR measurements and an empirical method for estimating terrain parameters from the LIDAR data.

Chapter 3

Data Fusion Via Multiscale Estimation

3.1 Introduction

Chapter 2 discusses methods for transforming INSAR and LIDAR data into estimates of z_g and Δz_v . After that transformation, the corresponding quantities can be fused. This chapter presents a multiscale estimation approach to data fusion that is based on the Kalman filter. The basic multiscale Kalman smoothing (MKS) algorithm was developed in [3]. Its application has been limited to problems in which the observations were linearly related to the state variables or cases where a nonlinear relation had a closed-form inverse. In the present application, no closed-form inverse between the original observations and state variables exists. Combining the physical modeling of Chapter 2 with the MKS algorithm represents an extension of the existing approach because it expands the domain of problems that can be addressed with the MKS method. In Section 3.2 background material is presented on the development of the MKS algorithm. In Section 3.3 fundamental aspects of

Kalman filtering are discussed. The MKS algorithm is described in Section 3.4, and aspects of the data fusion problem are discussed in Section 3.5. Finally, Section 3.6 gives conclusions.

3.2 Background

The Kalman filter represents one type of estimator. There are several other estimation methods, including maximum likelihood (ML), maximum a posteriori (MAP), and minimum mean squared error (MMSE) [27]. ML and MAP estimators both require the computation of conditional probability distribution functions (pdfs). Both ML and MAP seek the most likely value of the random process given a measurement. So both estimators find the best estimate with respect to the *current* observation. The MMSE estimator produces an estimate that is globally optimal as it seeks to minimize the variance of the ensemble of observations. For a given observation, the ML or MAP estimates may have a smaller squared error than the MMSE estimate, but over many observations, the MMSE will have the smallest MSE.

The linear MMSE (LMMSE) estimate trades off between global optimality and computational complexity. Unlike the other methods given above, LMMSE does not require calculation of conditional densities. Instead, it relies on second-order moments of the process and observations. Thus the LMMSE can be implemented as set of linear equations that can be solved with non-iterative techniques. If the process or the observations are not normally distributed, i.e. they are not completely described by their first two moments (mean and variance), the LMMSE will represent a suboptimal estimator. However, even when this is the case, the performance is often close to optimal.

Normality of the data is assumed throughout this dissertation. Terrain reflectivity is typically assumed to be a white, Gaussian process [28]. The resulting

interferometric phase is distributed approximately Gaussian. Similarly, the topography observed by the LIDAR is assumed Gaussian, and the resulting estimates of ground and vegetation heights are approximately Gaussian. Techniques exist for transforming non-normally distributed data so that they are nearly normal [29]. The choice of transformation is largely data dependent though and would necessitate additional processing steps.

Multiscale, or hierarchical, signal modeling has received much attention over the last ten years [4] [30] [31] [32] [33]. The work has been motivated primarily by the need to develop stochastic models that capture the true multiscale character of many natural processes and a need to combine signals and process measurements of differing resolutions (data fusion). Approaches to hierarchical data modeling have included methods based on fine-to-coarse transformations of spatial models and methods that directly model processes on multiscale data structures, such as dyadic trees or quadtrees [33]. In this dissertation, I use the latter approach with quadtrees to combine INSAR and LIDAR data of disparate resolutions.

Some multiresolution approaches to data fusion have been reported, but they either assume a linear relationship between the measurements and the state variables or linearize the process equations [3] [30]. A physical model was used in [4] for the solution of a multiscale estimation problem, but only to determine measurement variances in the stochastic data model. The data were still considered direct measurements of the state variables. Schneider, Fieguth, Karl, and Willsky recently solved a nonlinear problem (image segmentation) with the linear multiscale Kalman filter approach [32]. In that work, however, a closed-form relationship exists between the state variables and observations. Thus, it is possible to estimate the state using linear estimation by defining a pseudo-observation that is a linear function of one of the state variables. Microwave scattering interactions with vegetation are

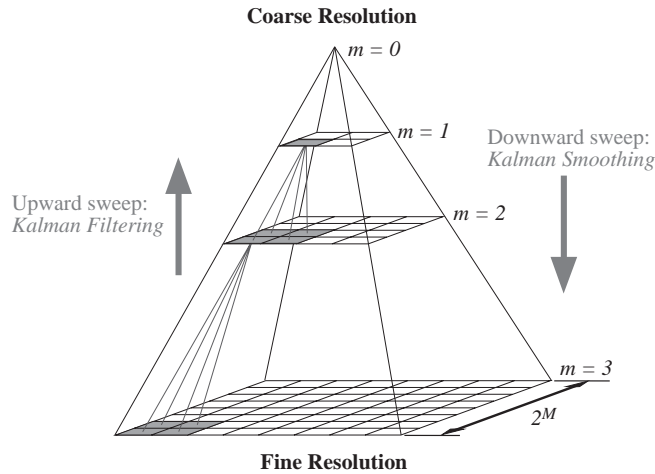


Figure 3.1: The quadtree data structure, where m denotes scale. The support of the tree at the finest scale is $2^M \times 2^M$, where $m \in \{0, \dots, M\}$.

extremely complex. Even though I employ the simplest scattering model that still retains the essential physical relationships between surface and vegetation heights and the INSAR observations, no closed-form inverse relation exists. Therefore, decoupling the nonlinearities from the multiscale linear estimation is required.

Fig. 3.1 shows a multiscale pyramid, or quadtree, which is used in the data fusion. Wavelet decompositions have also been applied to multiresolution analysis, but they do not readily accommodate irregularly spaced data or indirect measurements of the process being estimated [30]. In this work, the more general problem of combining sparse and dissimilar data types at multiple scales must be solved. For this, a state-space approach is natural, which leads to a Kalman filter formulation. In order to understand the multiscale Kalman filter, some background on Kalman filter theory is needed.

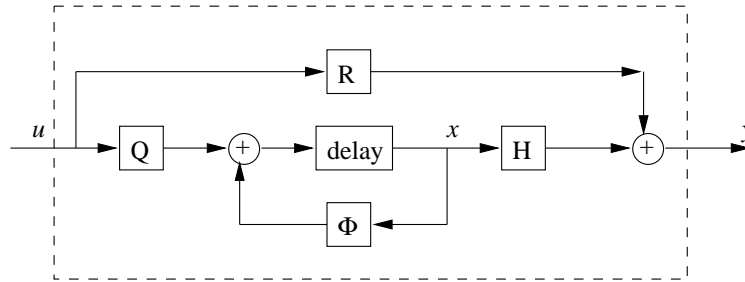


Figure 3.2: Block diagram of a Kalman filter, where $u \sim \mathcal{N}(0, 1)$, $Qu = w$, $Ru = v$.

3.3 The Kalman Filter

The Kalman filter can be presented as a recursive filter or as a linear estimator. The former approach leads to formulations of a frequency response and transfer function. The latter leads to a set of simultaneous linear equations. The linear estimator is the most natural representation because the input, output, and state variables are explicit. Fig. 3.2 depicts the basic operations of the Kalman filter.

Wiener and Kalman filters are both LMMSE estimators. However, the Kalman filter is particularly suited to situations where the process to be estimated is not observed directly. This allows uncertainty in the knowledge process and the measurements to be treated explicitly. The Kalman filter balances the variance of the process with the variance of the observations in a linear combination to arrive at an estimate. As a result, it handles sparse or missing data in a natural manner.

3.3.1 Linear state model

The set of linear equations that comprise the Kalman filter include a state relation

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \mathbf{w}_k \quad (3.1)$$

where

$\mathbf{x}_k = (n \times 1)$ state vector at step k

$\Phi_k = (n \times n)$ state transition matrix

$\mathbf{w}_k = (n \times 1)$ process noise

and a measurement relation

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (3.2)$$

where

$\mathbf{y}_k = (m \times 1)$ measurement vector at step k

$\mathbf{H}_k = (m \times n)$ state-measurement relation

$\mathbf{v}_k = (m \times 1)$ measurement noise

3.3.2 Kalman filter algorithm

The recursive algorithm that produces the Kalman filter estimates is given in Fig. 3.3.

In Fig. 3.3, the process noise \mathbf{w} is assumed to be white (uncorrelated for nonzero lag) with covariance \mathbf{Q} :

$$E[\mathbf{w}_k \mathbf{w}_i^T] = \begin{cases} \mathbf{Q}_k, & i = k \\ 0, & i \neq k \end{cases} \quad (3.3)$$

The measurement noise \mathbf{v} is assumed to be white (uncorrelated for nonzero lag) with covariance \mathbf{R} :

$$E[\mathbf{v}_k \mathbf{v}_i^T] = \begin{cases} \mathbf{R}_k, & i = k \\ 0, & i \neq k \end{cases} \quad (3.4)$$

The process and measurement noise are not cross-correlated.

$$E[\mathbf{w}_k \mathbf{v}_i^T] = 0 \quad \forall i, k \quad (3.5)$$

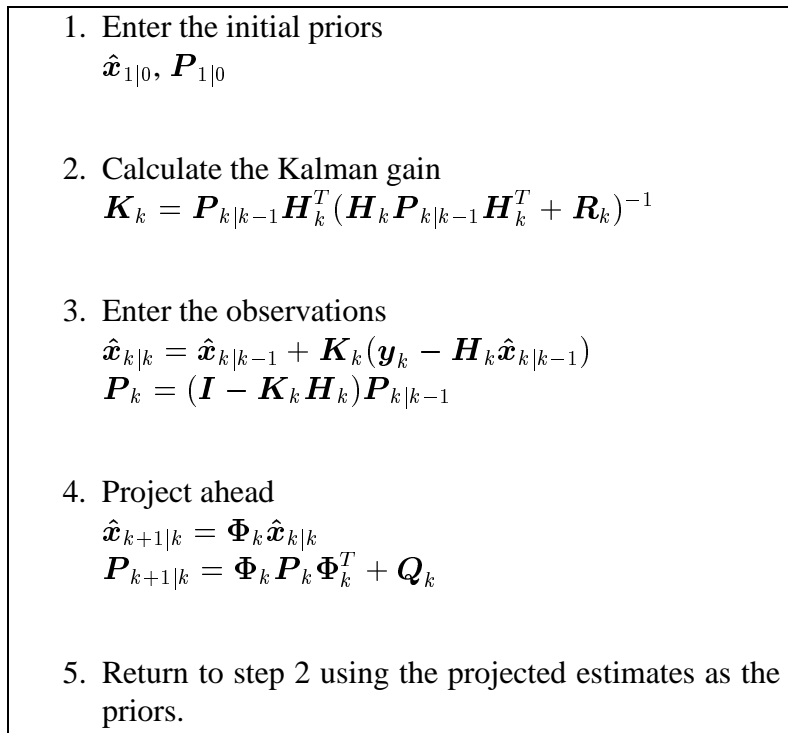


Figure 3.3: The recursive Kalman filter algorithm.

The conditions on \boldsymbol{w} and \boldsymbol{v} given by (3.3)–(3.5) are necessary conditions for the Kalman filter to be the optimal LMMSE estimator. It is also necessary that Φ , \boldsymbol{Q} , \boldsymbol{H} , and \boldsymbol{R} be known. In practice, these parameters will not be known exactly, but they are often known approximately. The performance of the filter when these parameters are not perfectly known or the noise terms are not exactly white is often quite good, although not strictly optimal.

3.3.3 Kalman smoothing estimates

The Kalman filter has been widely used to estimate the values of a set of state variables given a model for the evolution of the state variables and their relationship to the observations. The filter minimizes the trace of the error covariance matrix \boldsymbol{P} , which is equivalent to minimizing the mean squared error (MMSE) between the state \boldsymbol{x} and the estimate $\hat{\boldsymbol{x}}$, for each sample of the independent variable [34]. In most applications, the independent variable is either a temporal or spatial coordinate. For the k^{th} sample, where E is the expectation operator, the error covariance matrix is

$$\boldsymbol{P}(k|k) = \boldsymbol{P}_k = E[\boldsymbol{e}_k \boldsymbol{e}_k^T] = E[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^T] \quad (3.6)$$

The k^{th} estimate $\hat{\boldsymbol{x}}_k$ is conditioned on the observation at sample k (if present) and the previous estimate of the state $\hat{\boldsymbol{x}}_{k-1}$. The Kalman formulation naturally accommodates missing data because previous estimated state values are used in conjunction with observations to determine the current state. The observations and the state model are weighted according to their variances.

Kalman smoothing is accomplished if a return sweep through the data is performed. Smoothing utilizes observations subsequent to the current sample to improve the estimated state, in the sense of reduced variance [34]. The error covari-

ance of the smoothed estimates is

$$\begin{aligned}\mathbf{P}^s(k|N) &= \mathbf{P}(k|k) + \mathbf{J}(k)[\mathbf{P}^s(k+1|N) - \mathbf{P}(k+1|k)]\mathbf{J}^T(k) \\ \mathbf{J}(k) &= \mathbf{P}(k|k)\Phi^T(k+1, k)\mathbf{P}^{-1}(k+1|k)\end{aligned}\quad (3.7)$$

where N is the total number of samples to estimate, $k = N - 1, N - 2, \dots, 0$, and Φ is the state transition matrix. The formulation in (3.7) is sometimes referred to as Rauch-Tung-Striebel smoothing or fixed-interval smoothing. More recently, Kalman smoothing has been adapted for multiscale data representations [3] [30] [4]. The derivation of multiscale Kalman smoothing (MKS) is a generalization of classical Kalman smoothing, with the addition of a merge step to accommodate the change in support of the data at each scale. Because of the underlying Kalman formulation, this method can accommodate irregularly spaced and sparse observations. In addition to the estimates, it provides the corresponding uncertainty (error variance) for every estimate, which is essential for quantitative analysis. Unlike the Wiener filter, which also provides MMSE estimates, the Kalman formulation explicitly separates the measurement and state relations, which makes it better suited to data fusion [34].

3.4 Multiresolution Kalman smoothing

To present the multiresolution version of the Kalman filter, it is helpful to amend notation slightly. In discussions of 1-D processes, the argument k is used to denote the recursion index of the filter. Multiscale processes are modeled on tree data structures, and s will denote the nodes of the tree and replace k as the index of recursion. For a 2-D process, multiresolution Kalman smoothing begins with a fine-to-coarse sweep up the quadtree that is analogous to Kalman filtering with an added merge step. This is followed by a coarse-to-fine sweep down the quadtree that

corresponds to Kalman smoothing. Using the scalar form for clarity, the coarse-to-fine linear dynamic model is given as

$$\begin{aligned} x(s) &= \Phi(s)x(Bs) + \Gamma(s)w(s) \quad \forall s \in \mathcal{S}, s \neq 1 \\ y(s) &= H(s)x(s) + v(s) \quad \forall s \in \mathcal{T} \subseteq \mathcal{S} \end{aligned} \quad (3.8)$$

where x is the state variable, and y represents the observations. The stochastic forcing function w is a Gaussian white noise process with identity variance, and the measurement error v is a Gaussian white noise process with scale dependent variance $R(s)$. \mathcal{S} represents the set of all nodes on the quadtree, and \mathcal{T} denotes those nodes at which an observation is available. s is the node index on the tree, and $s = 1$ denotes the root node. B is a backshift operator in scale, such that Bs is one scale coarser than s . Φ is the coarse-to-fine state transition operator, Γ is the coarse-to-fine stochastic detail scaling function, H is the measurement-state model, and R represents the measurement variance of the observations. A complete description of the MKS algorithm can be found in [3] [4]. This algorithm is noniterative and has constant computational complexity per pixel with $\mathcal{O}(\mathcal{S}_M)$ operations, where \mathcal{S}_M is the number of nodes at the finest scale $m = M$.

A significant limitation of this basic Kalman approach is the requirement that observations be linearly related to the state processes. This is often not the case, and is certainly not valid when INSAR and LIDAR measurements are used to estimate ground elevations and vegetation heights. The *linearized Kalman filter* does accommodate nonlinearities in the measurement-state relation, but does so by estimating small deviations about nominal state values. This approach can be adequate for slowly-varying processes with dynamics that are well enough understood to specify a meaningful deterministic nominal state. However, topography and vegetation heights can change abruptly and in general must be modeled statistically. The *extended Kalman filter* allows the linearization to take place about the esti-

mated process instead of a fixed deterministic process, but as a result, estimates can diverge. To avoid these problems, we estimate the state variables from the data by modeling the scattering interactions. In the case of INSAR, we invert a microwave scattering model, and for LIDAR, an empirical model is employed. Modeling provides greater insight into the nature of the state processes and accommodates the nonlinear measurement-state relations.

Fig. 3.4 depicts the overall framework. It shows the physical modeling of the INSAR and LIDAR data as separate from the data fusion, which occurs in the multiscale estimation step. The data fusion method is general, with the primary requirement being that the observations are linearly related to the state variables. If new data types or new applications are to be investigated, then one need only replace the details of the physical modeling with some other operation that maps the data to some linear function of the state variables to be estimated.

In the multiscale case, any filtering operation will need to incorporate data at various scales and return a final result at the finest scale to be useful. Thus the Kalman filtering must necessarily be followed by Kalman smoothing to return to the finest scale. The total algorithm will be referred to as the multiscale Kalman smoothing (MKS) algorithm. The MKS algorithm was first presented in [3], and was applied to remotely sensed oceanographic data in [4]. We present a summary of the algorithm here. Fig. 3.5 shows a simple multiscale data structure to present the notation. The bold notation is dropped in the multiscale expressions for clarity, but it is understood that the quantities that were matrices and vectors in the 1-D case remain so in the multiscale case.

The process noise variance must be allowed to vary with scale in order to describe self-similar stochastic processes, such as topography. We assume $w(s) \sim \mathcal{N}(0, 1)$, thus $E[w(s)w^T(t)] = I\delta_{s,t}$. So Γ^2 represents the variance of the stochastic detail that is incorporated as the resolution increases. The initial priors are specified

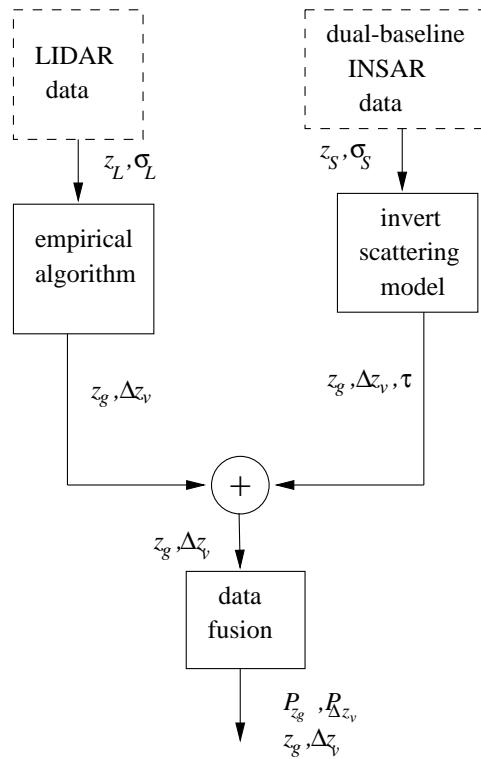


Figure 3.4: Block diagram of the overall data fusion process. $z(\cdot)$ quantities represent heights, $\sigma(\cdot)$ quantities represent standard deviations of measurement errors, and $P(\cdot)$ quantities represent final estimate error variances.

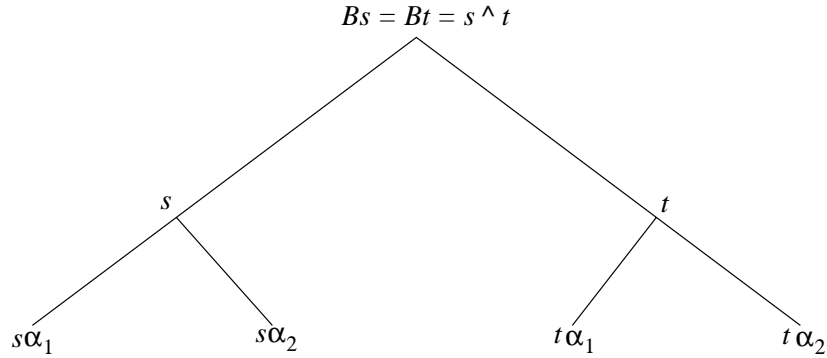


Figure 3.5: Simplified multiscale data structure (dyadic tree) used to present notation. B represents a backshift operator in scale. $s \wedge t$ represents a parent node to nodes s and t . α represents a forward shift operator in scale.

for the process model that evolves downward (3.8). Assuming a zero-mean process [4], the priors for the state mean and covariance are

$$\begin{aligned}\hat{x}_{1|0} &= E[x(1)] = 0 \\ P_{1|0} &= E[x(1)x^T(1)] = P_s(1)\end{aligned}\tag{3.9}$$

where $s = 1$ represents the root node, and $P_s(s)$ denotes the covariance of the state at node s . These priors are used with the recursion in (3.8) to generate a realization of a multiscale stochastic process. The resulting estimates of the state and state covariance at the leaf nodes are then used as the initial priors in the upward Kalman filter. In general, little is known about the process to be estimated, so $P_s(1)$ is chosen to be some large number, such as 10^5 . The *a priori* process model is now completely specified. The corresponding upward model can be specified, which the

Kalman filter will track:

$$\begin{aligned} x(Bs) &= F(s)x(s) + \bar{w}(s) \\ y(s) &= H(s)x(s) + v(s) \end{aligned} \tag{3.10}$$

$$F(s) = P_s(Bs)\Phi^T(s)P_s^{-1}(s) \tag{3.11}$$

$$\begin{aligned} E[\bar{w}(s)\bar{w}^T(s)] &= P_s(Bs)[I - \Phi^T(s)P_s^{-1}(s)\Phi(s)P_s(Bs)] \\ &= Q(s) \end{aligned} \tag{3.12}$$

Here, Q is the process noise covariance in the upward model. It is the multiscale analog to Q in the temporal Kalman filter in Section 3.3. F is the fine-to-coarse state transition operator. The MKS algorithm now proceeds with initialization, the upward sweep, and finally the downward sweep, as described in Section 3.4.1.

3.4.1 MKS algorithm

Initialization

At the leaf nodes, enter the prior values

$$\hat{x}(s|s+) = 0 \tag{3.13}$$

$$P(s|s+) = P_s \tag{3.14}$$

Upward sweep

The upward sweep is equivalent to a Kalman filter operating in scale with an additional merge step that reduces the support of the estimates at each scale. Having defined the initial priors at the leaf nodes, the algorithm proceeds from the base of

the quadtree up to the root node

$$\hat{x}(s|s) = \hat{x}(s|s+) + K(s)[y(s) - H(s)\hat{x}(s|s)] \quad (3.15)$$

$$P(s|s) = [I - K(s)H(s)]P(s|s+) \quad (3.16)$$

$$K(s) = P(s|s+)H^T(s)V^{-1}(s) \quad (3.17)$$

$$V(s) = H(s)P(s|s+)H^T(s) + R(s) \quad (3.18)$$

The project-ahead step is applied at all scales above the base. For $i = \{1, 2, 3, 4\}$, we have

$$\hat{x}(s|s\alpha_i) = F(s\alpha_i)\hat{x}(s\alpha_i|s\alpha_i) \quad (3.19)$$

$$P(s|s\alpha_i) = F(s\alpha_i)P(s\alpha_i|s\alpha_i)F^T(s\alpha_i) + Q(s\alpha_i) \quad (3.20)$$

The four project-ahead estimates from each 2×2 group of nodes at the previous scale must be merged into a single value at the current scale

$$\hat{x}(s|s+) = P(s|s+) \sum_{i=1}^q P^{-1}(s|s\alpha_i)\hat{x}(s|s\alpha_i) \quad (3.21)$$

$$P(s|s+) = \left[(1 - q)P_s^{-1}(s) + \sum_{i=1}^q P^{-1}(s|s\alpha_i) \right]^{-1} \quad (3.22)$$

For the quadtree, $q = 4$.

Downward sweep

The upward estimate at the root node is used as the initial condition $\hat{x}^s(0) = \hat{x}(0|0)$ to start the downward sweep, where the superscript s refers to a smoothed quantity.

The downward sweep then proceeds down the tree

$$\hat{x}^s(s) = \hat{x}(s|s) + J(s)[\hat{x}^s(Bs) - \hat{x}(Bs|s)] \quad (3.23)$$

$$P^s(s) = P(s|s) + J(s)[P^s(Bs) - P(Bs|s)]J^T(s) \quad (3.24)$$

$$J(s) = P(s|s)F^T(s)P^{-1}(Bs|s) \quad (3.25)$$

3.4.2 Stochastic model identification

Before (3.8) can be applied to data, a general form for the stochastic model of the state must be chosen that approximates the scale-to-scale evolution of the state process. Once a general form is selected, specific values for the model parameters must be computed. In this dissertation, a fractional Brownian motion model is assumed, and the power spectrum of the stochastic model is matched to that of the data.

Many natural processes, such as topography, can exhibit approximately self-similar statistics across resolution scales. Statistical self similarity implies

$$x(\xi s) \stackrel{\mathcal{D}}{=} \xi^\rho x(s) \quad , 0 < \rho < 1 \quad , \xi > 0 \quad (3.26)$$

where ρ is the Hurst exponent, and $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution [35]. This characteristic can be closely approximated with $1/f^\mu$ stochastic models, where μ determines the rate of change of the process variability with spatial scale [36]. These models are closely related to fractional Brownian motion, which is characterized by zero initial state and Gaussian statistics [4] [34]. For scalar processes, the coarse-to-fine process noise is given by a white, Gaussian random variable with zero mean and variance Γ^2 , where Γ may be a function of scale. In the fractional Brownian motion model, the coarse-to-fine state transition is unity ($\Phi(s) = 1$), which indicates that the unforced state is perfectly correlated through scale. All of the variation is provided by the stochastic detail function

$$\Gamma = \Gamma_0 2^{(1-\mu)m/2} \quad (3.27)$$

where m is the scale, Γ_0^2 is the variance at the root node, and μ governs the rate at which the variance changes with scale. With $\mu > 1$, the variance decreases monotonically with increasing resolution.

Because a $1/f^\mu$ process implies a power spectrum that is proportional to $1/|f|^\mu$, μ and Γ_0 can be chosen so that the power spectrum of the multiscale data model matches the sample power spectrum of the observations. The power spectrum of the INSAR data is computed at the finest scale at which INSAR data are available. The power spectrum of a realization of the stochastic data model in (3.8) is computed at the same scale using empirically selected values of μ and Γ_0 that produce a close match.

Power spectra of discrete image data can only represent signal energy over a finite range of spatial frequencies, so a 2-D Hamming window is applied to the data prior to computing the power spectra to reduce aliasing. Integrability of the power spectrum is maintained at high frequencies because of the bound on f due to discrete image data. Physical processes, such as topography, diverge from the $1/f$ model at very low frequencies, which leads to finite values for the power spectrum near $f = 0$.

Because the INSAR and LIDAR data are transformed into estimates of the state variables by the physical modeling, H is simply a selection matrix. It is a binary indicator function for the presence of an observation at a particular scale and pixel location.

The measurement error variance R is determined from sensor parameters and from the data. In the case of INSAR data, the Cramer-Rao bound for the phase noise is used to determine a spatially-varying $R(s)$. For the LIDAR data, a spatially-varying $R(s)$ is determined from the variability of individual LIDAR pulse returns. These error sources were discussed in Chapter 2.

$$R(s) = \begin{cases} R_{INSAR} & \forall s \in \mathcal{T}_S \subseteq \mathcal{T} \\ R_{LIDAR} & \forall s \in \mathcal{T}_L \subseteq \mathcal{T} \end{cases} \quad (3.28)$$

Here, \mathcal{T}_S is the set of nodes where SAR data are available, and \mathcal{T}_L is the set of

nodes where LIDAR data are available.

The state variable $x(s)$ is z_g when estimating surface elevations and Δz_v when estimating vegetation heights. The methods used to obtain initial estimates of z_g and Δz_v from the INSAR and LIDAR observations were described in Chapter 2. All of the parameters in (3.8) are thus determined.

3.4.3 Computational complexity

This algorithm is non-iterative and the number of operations grows linearly with the number of leaf nodes. To see this, let S_M be the number of leaf nodes. There are $\lfloor \frac{4}{3}S_M \rfloor$ nodes on the tree. Thus, the total number of nodes is linear in S_M . Since all operations are local (nearest neighbor children and one parent), the complexity per node is constant. So total complexity is linear in S_M .

3.5 Data fusion with the Kalman filter

The INSAR data is the primary measurement because of its wide coverage area. We wish to combine LIDAR data with the INSAR data to improve the estimates. In the MKS algorithm, data from different sensors are incorporated in the Kalman observation variable y . At the finest scale, y will represent LIDAR data, while at a coarser scale, it will represent INSAR data. When combining data from different sensors, we need to know whether or not adding new data will always help; i.e., whether or not it will always reduce the MSE, so that we avoid deteriorating the estimate quality by fusing data. Section 3.5.1 addresses this issue. It is also of interest to know how the estimate error will propagate where no observations are available. This is a common situation at the finest scale, where the data are often sparse, or there is incomplete coverage. Section 3.5.2 describes the behavior of the

estimate error when observations are absent.

3.5.1 Impact of additional data types

I determine the impact of adding another data type (LIDAR) in terms of the estimate uncertainty. Ground truth measurements conducted at the test sites demonstrated that the LIDAR data are closer to the actual surface and vegetation heights than the INSAR data once all relative 0^{th} -order (offsets) and 1^{st} -order (tilts) biases between the INSAR and LIDAR have been removed. LIDAR data also exhibit smaller measurement error variance than the INSAR data, i.e. $\mathbf{R}_L < \mathbf{R}_S$.

The Kalman gain determines how the observations are weighted relative to the state model in determining the estimate.

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (3.29)$$

Consider two limiting cases, where the measurement is perfect $\mathbf{R}_k = 0$, and where the measurement is completely untrustworthy, or there is no measurement $\mathbf{R}_k = \infty$.

$$\text{for } \mathbf{R}_k = 0 : \quad \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + 0)^{-1} = 1 \quad (3.30)$$

$$\mathbf{P}_k = (\mathbf{I} - 1 \mathbf{H}_k) \mathbf{P}_{k|k-1} = 0$$

$$\text{for } \mathbf{R}_k = \infty : \quad \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \infty)^{-1} = 0 \quad (3.31)$$

$$\mathbf{P}_k = (\mathbf{I} - 0 \mathbf{H}_k) \mathbf{P}_{k|k-1} = \mathbf{P}_k > 0$$

Smaller \mathbf{R} reduces the estimate uncertainty \mathbf{P}_k , thus adding the LIDAR will always improve the final estimate at sample k .

3.5.2 Error propagation

In Section 3.5.1 it was proven that under the stated conditions adding LIDAR data at sample k will improve the estimate at sample k . Now we wish to determine what effects the presence of measurement might have on neighboring estimates where data is no longer available. Specifically, we wish to know how the estimate uncertainty deteriorates as the Kalman filter moves across samples where data are missing due to sparse coverage or data drop outs. The 1-D Kalman filter is examined here because it results in a simple closed-form solution. We will show that the error uncertainty remains bounded. Bounded estimate uncertainty was proven for the multiscale case in [37], but in that case, no explicit relation to the Kalman model parameters was obtained.

Let us denote the recursion steps by the sequence $k = \{0, 1, \dots, l, l + 1, \dots\}$, where l is the last step at which an observation is available. For any step $k > l$, estimates are determined solely from the state model, the posterior estimate $\hat{\mathbf{x}}$, and estimate error covariance \mathbf{P} at $k = l$. As described in Section 3.5.1, $\mathbf{R}_k = \infty \rightarrow \mathbf{K}_k = 0$ at nodes $k > l$. The estimate error variance \mathbf{P}_k varies with k as:

$$\mathbf{P}_{l+1} = (\mathbf{I} - \mathbf{K}_{l+1}\mathbf{H}_{l+1})\mathbf{P}_{l+1}^- \quad (3.32)$$

$$\mathbf{P}_{l+1} = \mathbf{P}_{l+1}^- = \Phi_l \mathbf{P}_l \Phi_l^T + \mathbf{Q}_l$$

$$\mathbf{P}_{l+2} = \mathbf{P}_{l+2}^- = \Phi_{l+1} \mathbf{P}_{l+1} \Phi_{l+1}^T + \mathbf{Q}_{l+1}$$

$$\mathbf{P}_{l+2} = \Phi_{l+1} [\Phi_l \mathbf{P}_l \Phi_l^T + \mathbf{Q}_l] \Phi_{l+1}^T + \mathbf{Q}_{l+1} \quad (3.33)$$

Here, \mathbf{P}_k^- is used as compact notation for $\mathbf{P}_{k|k-1}$. Let d be the number of steps beyond step l , and assume Φ is constant. Then \mathbf{P}_{l+d} becomes

$$\mathbf{P}_{l+d} = \Phi^d \mathbf{P}_l (\Phi^T)^d + \Phi^{d-1} \mathbf{Q}_l (\Phi^T)^{d-1} + \dots + \Phi^l \mathbf{Q}_{l+d-1} (\Phi^T)^l$$

The functional form is made apparent by considering the scalar case:

$$P_k = \Phi^{2d} P_l + \sum_{i=1}^d \Phi^{2(i-1)} Q_{k-i} \quad \forall k > l \quad (3.34)$$

where $d = k - l$. P_k varies exponentially with d , thus it varies exponentially with k once no more observations are present. It is now shown that P_k remains bounded as k increases.

Theorem 1. *In the absence of measurements, the Kalman filter error variance P_k remains bounded as $k \rightarrow \infty$.*

Proof: 1. *Without loss of generality, we can assume the following conditions on the Kalman filter hold:*

$$\begin{aligned} 0 < \Phi < 1 \\ 0 < P_l < \infty \\ 0 < Q < \infty \end{aligned} \quad (3.35)$$

P_l and Q are variances, so they are finite and positive. Naturally occurring and stable stochastic processes satisfy $0 < \Phi < 1$. From (3.34),

$$P_k = \Phi^{2(k-1)} P_l + \sum_{i=1}^{k-1} \Phi^{2(i-1)} Q_{k-i} \quad (3.36)$$

$$\text{Let} \quad (3.37)$$

$$a_k = \Phi^{2(k-1)} P_l \quad (3.38)$$

$$b_k = \sum_{i=1}^{k-1} \Phi^{2(i-1)} Q_{k-i} \quad (3.39)$$

$$\text{Then} \quad (3.40)$$

$$P_k = a_k + b_k \quad (3.41)$$

$$\begin{aligned} a_k &= \Phi^{2(k-1)} P_l \\ \lim_{k \rightarrow \infty} a_k &= P_l \lim_{k \rightarrow \infty} \Phi^{2k-2} = \frac{P_l}{\Phi^2} \lim_{k \rightarrow \infty} \Phi^{2k} \\ \lim_{k \rightarrow \infty} a_k &= \frac{P_l}{\Phi^2} (0) = 0 \end{aligned} \quad (3.42)$$

$$b_k = \sum_{i=1}^{k-1} \Phi^{2(i-1)} Q_{k-i} \quad (3.43)$$

The dependence of P_k on k is made clearer if Q is assumed to be constant.

$$b_k = Q \sum_{i=1}^{k-1} \Phi^{2(i-1)} \quad (3.44)$$

Substitute $j = i - 1$. We obtain a geometric series in Φ^2 . For $|\Phi| < 1$, the series converges to a finite limit.

$$\lim_{k \rightarrow \infty} b_k = Q \lim_{k \rightarrow \infty} \sum_{j=0}^{k-1} \Phi^{2j} \rightarrow \frac{Q}{1 - \Phi^2} \quad (3.45)$$

So, for P_k , we have

$$\lim_{k \rightarrow \infty} P_k = \lim_{k \rightarrow \infty} a_k + \lim_{k \rightarrow \infty} b_k \quad (3.46)$$

$$\lim_{k \rightarrow \infty} P_k = \frac{Q}{1 - \Phi^2} \quad (3.47)$$

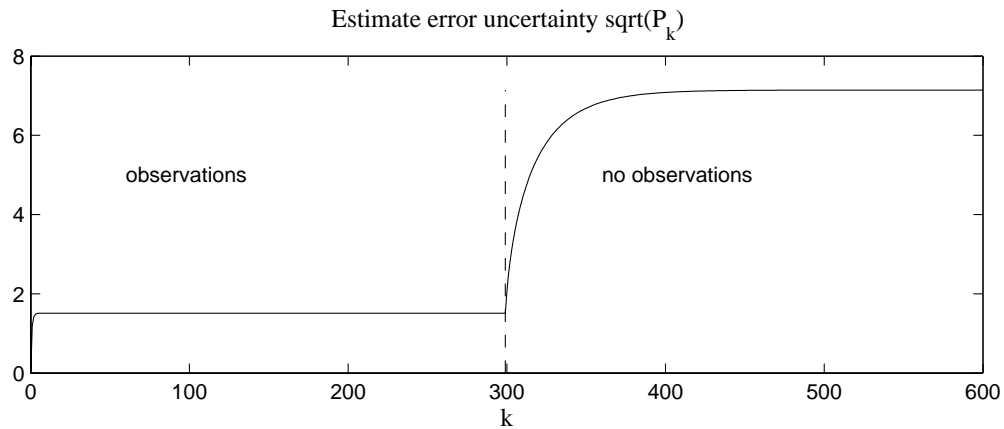


Figure 3.6: $\sqrt{P_k}$ for a simulated 1-D random process. No observations are present after $k = l = 300$. $\sqrt{P_k}$ converges to $\sqrt{Q/(1 - \Phi^2)}$, which is equal to 7.1421 in this example.

P_k converges to a finite limit as $k \rightarrow \infty$ and thus remains bounded.

Q.E.D ■

The estimate uncertainty $\sqrt{P_k}$ is plotted in Fig. 3.6 to show how it converges to its upper bound. In this example, $Q = 2$, $\Phi = 0.9802$, and $\sqrt{P_k}$ converges to a value of 7.1421.

3.6 Conclusions

This chapter presents the basic MKS algorithm. It also presents classical 1-D Kalman filter theory to illustrate Kalman filtering and smoothing algorithms and the assumptions on correlations that are required for optimal filter performance. These assumptions play an important role in the adaptive estimation case, as will be seen in the following chapter. Finally, I provide a criterion for deciding whether a new data source is worth incorporating and prove that the estimate error remains

bounded in the absence of data.

Chapter 4

Adaptive Multiscale Estimation

4.1 Introduction

The MKS algorithm and some important properties of Kalman filtering of multiple data sets were presented in Chapter 3. In this chapter, the algorithm is generalized to allow locally adaptive updating of the Kalman model parameters. By employing a multidimensional process model, the Kalman filter can operate in space as well as scale, as shown in Fig. 4.1. The filter detects suboptimal performance in the Kalman filter that arises from incorrect initial parameter values or non-stationarities in the data. Detection occurs when the filter operates in the spatial dimension. That information is propagated as the filter operates in scale resulting in a mechanism to determine spatially-varying process noise.

The impact of modeling errors is discussed in Section 4.2. In Section 4.3 common approaches to adaptive estimation are described, and in Section 4.4 the particular adaptive method used in this dissertation is presented. The incorporation of the adaptive estimation into the multidimensional algorithm is discussed in Section 4.5. The result is an adaptive MKS (AMKS) algorithm. Finally, in Section 4.6

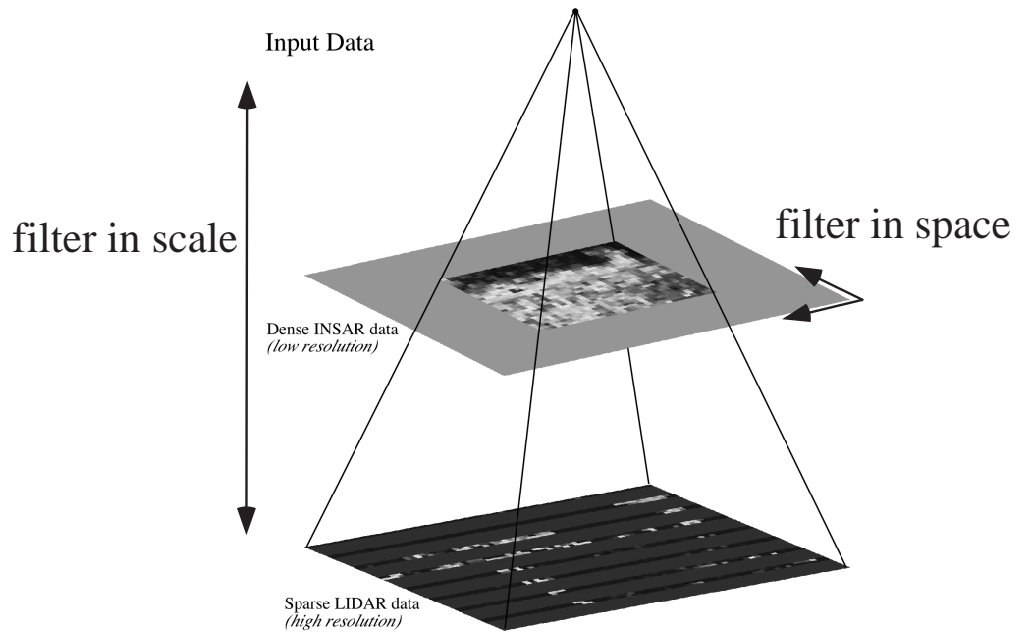


Figure 4.1: The quadtree data structure, as seen in Fig. 3.1, with observations shown at difference scales. The AMKS algorithm operates in scale just as the MKS algorithm does, but it has the capability to operate in the spatial dimensions to detect modeling errors and non-stationarities.

conclusions are presented. Throughout this chapter, the bold notation for vectors and matrices is dropped for clearer exposition. All vector quantities introduced in Chapter 3 remain vector quantities unless stated otherwise.

4.2 Effects of modeling errors

If any of the Kalman model parameters Φ , Q , H , or R is incorrect, or if either noise process w or v is not white, the Kalman estimates will be suboptimal. Consider a 1-D signal that is piecewise wide sense stationary (WSS) in two regions, as shown in Fig. 4.2. The true values of Q and R are different in the two regions, and there is a

step change between the regions. Thus, the signal changes both its first and second order statistical moments at the point of discontinuity. A non-adaptive Kalman filter is applied to the data that uses the correct model values in region 1 and continues using the same values in region 2, yielding estimates \hat{x}_{na} . A second filter also tracks the process, but it uses the correct model values in both regions, yielding estimates \hat{x} . The second filter represents a perfectly adaptive filter.

The effects of non-adaptive estimation are manifested in region 2 (Fig. 4.3). Both \hat{x} and \hat{x}_{na} estimates are optimal in region 1, but only \hat{x} is optimal in region 2. In the detailed view, it can be seen that \hat{x} tracks the true process more closely than \hat{x}_{na} , especially just after the discontinuity. In spite of an apparently large change in the process from region 1 to region 2, it can be seen that both estimates track the process reasonably well.

One of the strengths of the Kalman filter is its ability to provide reasonable estimates in spite of some modeling errors. This is possible because of the $(y_k - H_k \hat{x}_{k+1|k})$ term that appears in the filter algorithm. This term represents the difference between the current observation and the projected estimate, in other words the projection error. It allows the filter to track large-magnitude changes in the state process, and plays an important role in the adaptive algorithm used in this dissertation.

In Fig. 4.4, it is seen that the calculated estimate uncertainty \sqrt{P} is smaller for the non-adaptive filter. However, the *true error*, represented by the difference between the simulated state process and the estimates, is smaller for the adaptive filter. The non-adaptive filter had a MSE of 0.280 in this simulation, and the adaptive filter had a MSE of 0.266, corresponding to a 5% improvement. Using the non-adaptive filter is detrimental in two ways. The estimates are less accurate in the mean square sense, and the calculated estimate uncertainty is incorrect, which may lead to erroneous interpretation of the results.

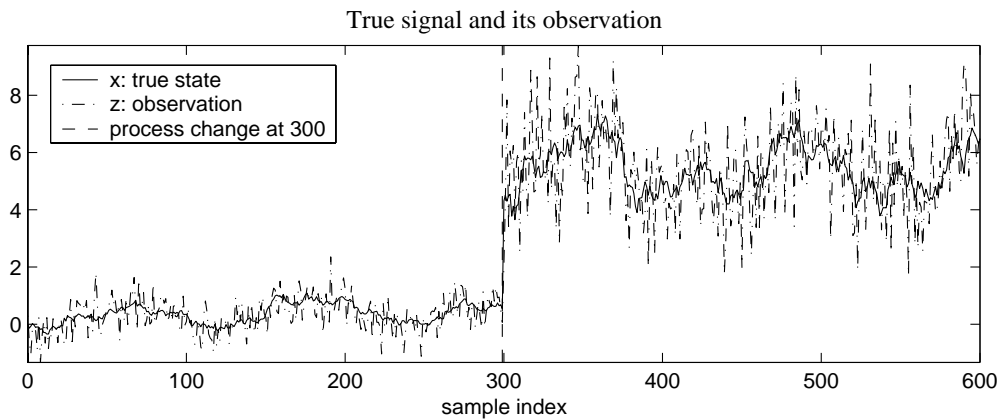


Figure 4.2: Simulated 1-D random process that is piecewise WSS, with $Q_2 = 4Q_1$ and $R_2 = 3R_1$

For the simulation in Fig. 4.2, $Q_2 = 4Q_1$ and $R_2 = 3R_1$. In Fig. 4.5, a much larger change in the noise variances in region 2 is simulated, where $Q_2 = 200Q_1$ and $R_2 = 10R_1$. The step change is omitted to obtain clearer plots of the signals. The non-adaptive filter had a MSE of 24.59 in this simulation, and the adaptive filter had a MSE of 13.50, corresponding to a 45% improvement.

An important lesson learned from this 1-D simulation is that even a perfectly adaptive Kalman filter may only slightly improve the estimates obtained from a standard Kalman filter. The magnitude of that improvement, and its importance to the application, will depend on the variability of state process and the magnitude of the modeling errors. An adaptive capability may be warranted, even if it provides a relatively small improvement in the mean squared error, if it provides greater understanding of the process under study.

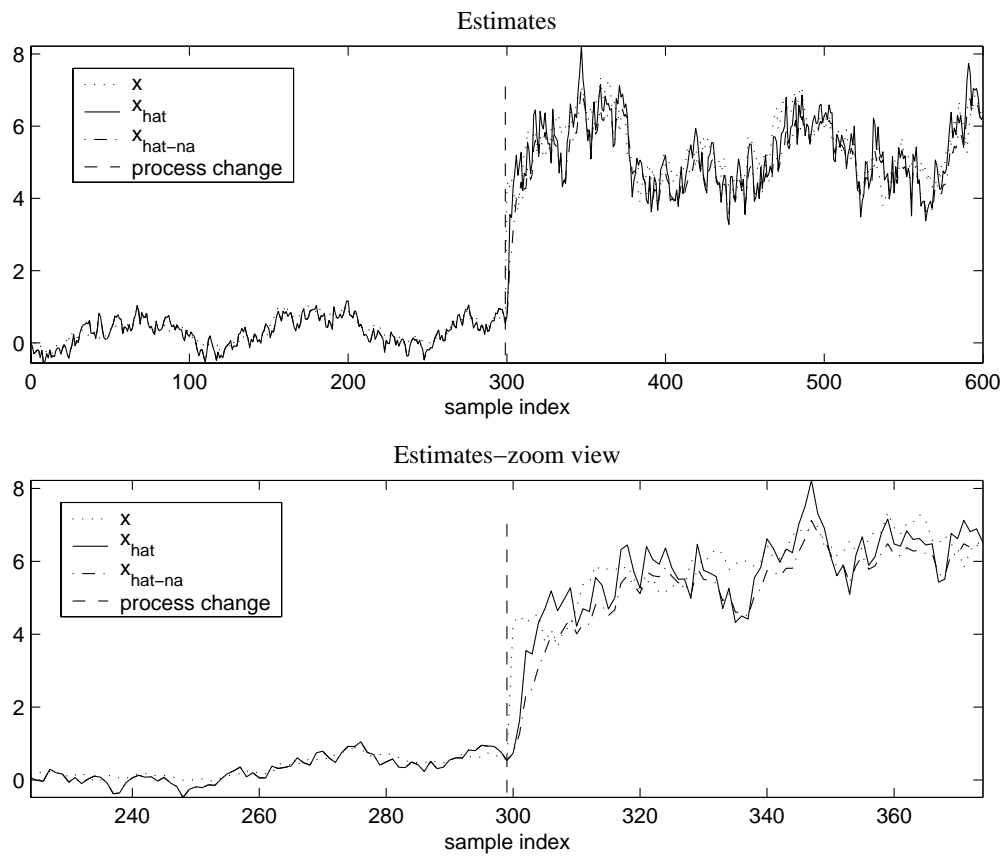


Figure 4.3: Adaptive \hat{x} and non-adaptive \hat{x}_{na} estimates of the process in Fig. 4.2.

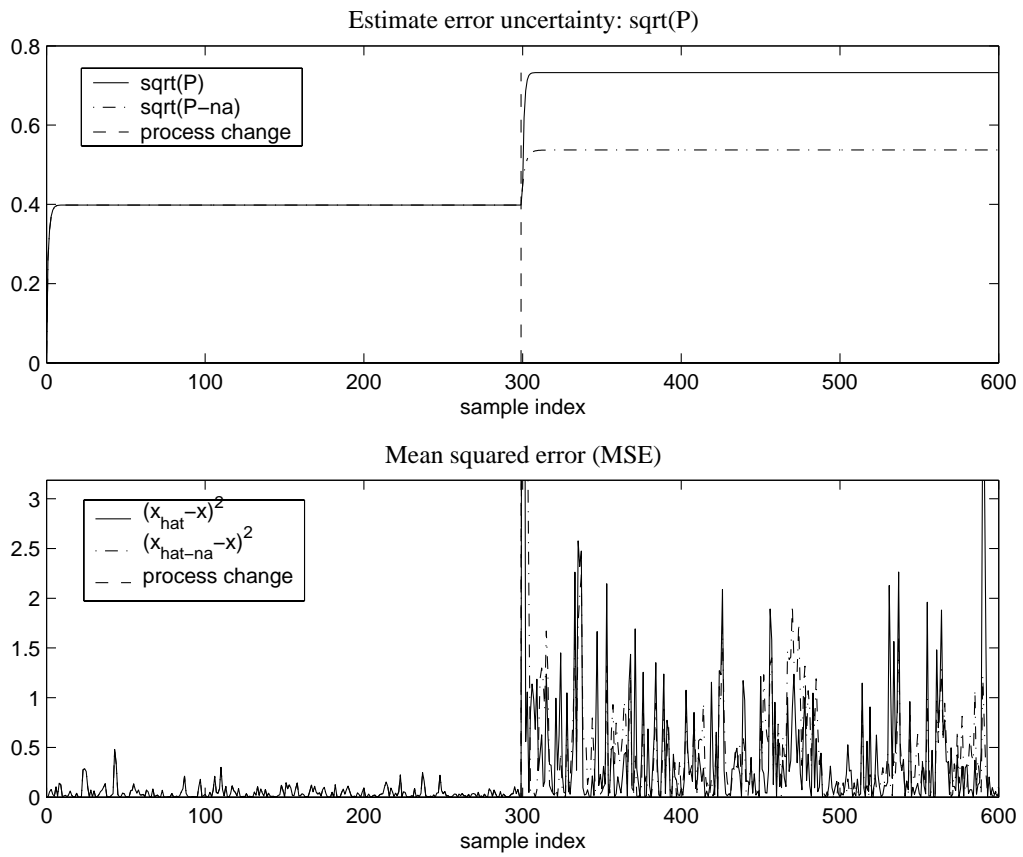


Figure 4.4: Estimate uncertainty and actual squared error for the adaptive and non-adaptive estimates in Fig. 4.3.

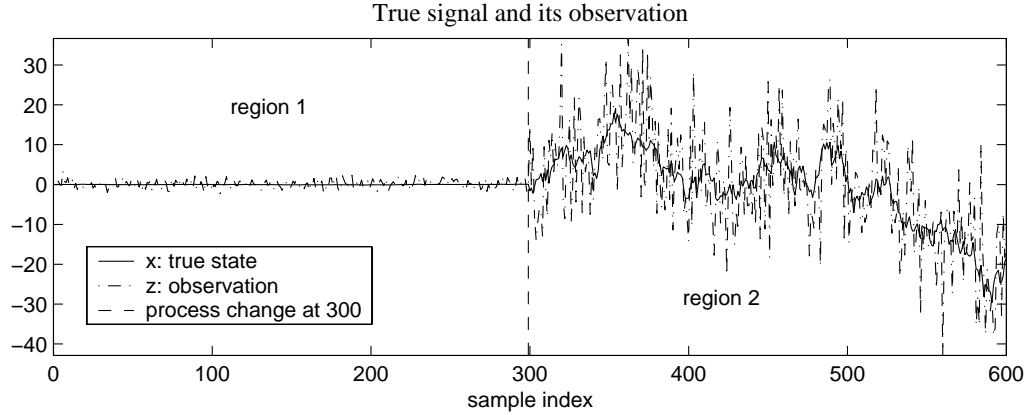


Figure 4.5: Simulated 1-D random process with large changes in the noise variances, $Q_2 = 200Q_1$ and $R_2 = 10R_1$.

4.3 Adaptive estimation approaches

Several approaches to adaptive estimation have been reported in the literature [38]. The most common are Bayesian methods [39], maximum likelihood methods [40], correlation methods [41], covariance matching methods [42], and multi-model methods [43]. Let α denote the set of unknown Kalman filter parameters $\{\Phi, H, Q, R\}$. Let Y^k denote the set of observations $\{y_1, \dots, y_k\}$.

Bayesian estimation seeks to compute an *a posteriori* probability density $p(\alpha|Y^k)$ and then forms the estimate as the conditional expectation

$$\hat{x}_{k|k} = E[x_k|Y^k] = \int_A \hat{x}(\alpha)_{k|k} p(\alpha|Y^k) d\alpha \quad (4.1)$$

where, A represents the region of support of the density function [44]. When A represents a large space (continuous) or set (discrete), the computational complexity of the integration can be prohibitive. Also, the conditional probability $p(\alpha|Y^k)$ is

determined from Bayes Rule as

$$p(\alpha|Y^k) = \frac{p(Y^k|\alpha)p(\alpha)}{p(Y^k)} \quad (4.2)$$

which is sensitive to the accurate determination of the prior probabilities $p(\alpha)$. This requires substantial *a priori* knowledge that is not generally available.

Maximum likelihood methods seek the maximum of a likelihood function $L(x_k, \alpha)$ representing the joint density function of the process x_k [45].

$$L(x_k, \alpha) = p(x_k, \alpha|Y^k) \quad (4.3)$$

Iterative nonlinear optimization techniques are often required to locate the maximum. Slow rates of convergence, discontinuities in the search space, and local maxima can lead to difficulties in the implementation of maximum likelihood methods [38].

In correlation methods, the autocorrelation of the observations or of the prediction errors in the Kalman filter is calculated. The prediction errors in the Kalman filter are also known as the innovations. They are given by

$$\nu_k = y_k - H_k \hat{x}_{k+1|k} \quad (4.4)$$

The system parameters are related to the observed correlations to solve for unknown parameters. Correlation methods assume that the random processes are ergodic so that the sample variance can be used to estimate the autocorrelations. It is usually preferred to use the innovations rather than the observations because they comprise a statistically independent sequence under optimal filter operation. This is the innovation–correlation method, and it is discussed further in Section 4.4. The sequence ν_k is tested to determine whether it is actually white by calculating the autocorrelation function. A non-white sequence implies that the model parameters may be incorrect or the noise may be correlated. Non-iterative algorithms are available

to update K , R , and Q when this occurs [41]. Figure 4.6 shows the autocorrelation function (ACF) of ν_k for the adaptive and non-adaptive filters in Fig. 4.3.

The ACF of an ideal white noise process is characterized by a large value at zero lag and much smaller magnitude values for all other lags, which indicates no significant correlation among the elements of the signal. The ACFs in Fig. 4.6 both appear consistent with a white noise process, but the ACF of the non-adaptive filter is shown to be non-white using a 95% confidence interval test. Thus, it is possible to discern the presence of errors in Q and R by calculating $\text{ACF}(\nu(k))$. The confidence interval test is discussed further in Section 4.4. The ACFs for the simulation in Fig. 4.5 are shown in Fig. 4.7. In this case, the ACF of the non-adaptive filter is clearly non-white, and is shown to be non-white by the confidence interval test. The ACF of the adaptive filter is considered to be approximately white since fewer than 5% of the ACF samples exceed the confidence test limit.

The covariance matching method tries to make the covariance of the innovation sequence equal to its theoretical value. However, this method may fail to converge when Q is not known.

Multi-model methods have become the most common adaptive estimation approach in recent years. They are appealing because they consist of a set of Kalman filters, thus they are non-iterative. In cases where the process dynamics are well known, and the parameters can take on only a small set of values, this approach may be the best. This approach is often used to estimate trajectories of orbiting satellites that obey well-known dynamical equations of motion in a small number of different circumstances. However, the dynamics of geological, erosional, and biological processes that generate topography and vegetation do not reduce to simple analytical governing equations. Thus, no deterministic models exist that describe surface and vegetation heights under general circumstances. Therefore, the multi-model approach is not appropriate for the estimation of topographic data in

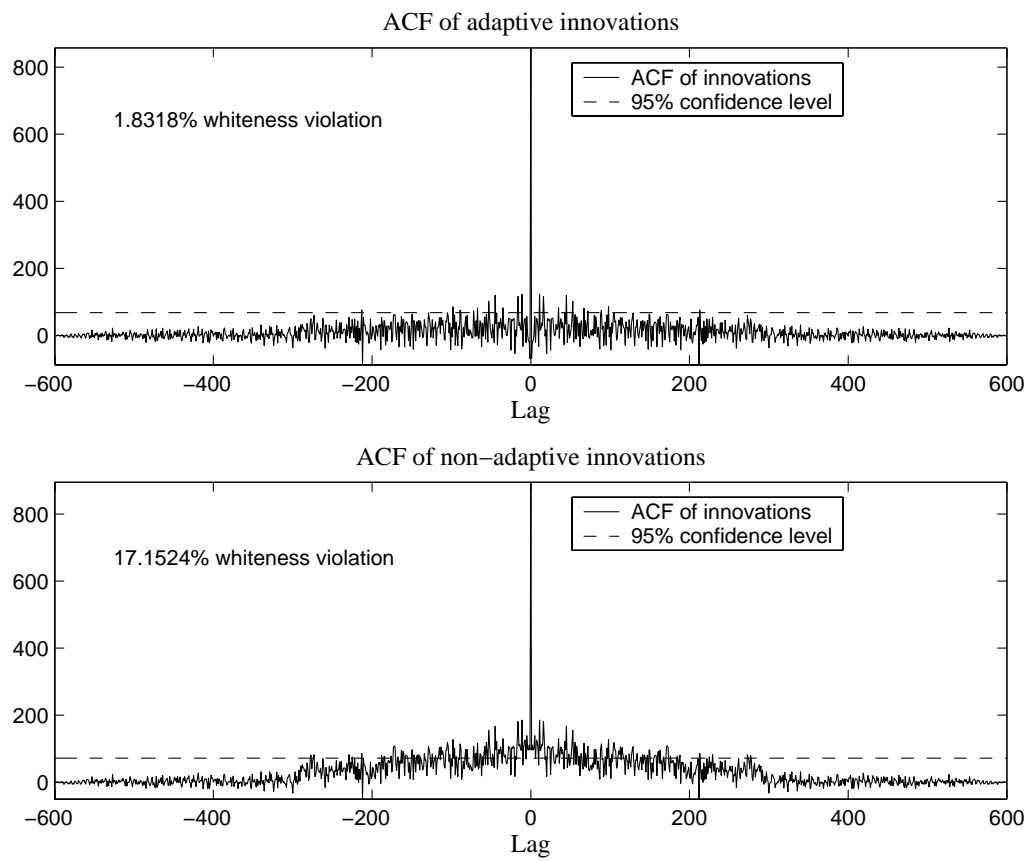


Figure 4.6: Autocorrelation of the innovations for the adaptive and non-adaptive estimates of the process in Fig. 4.2.

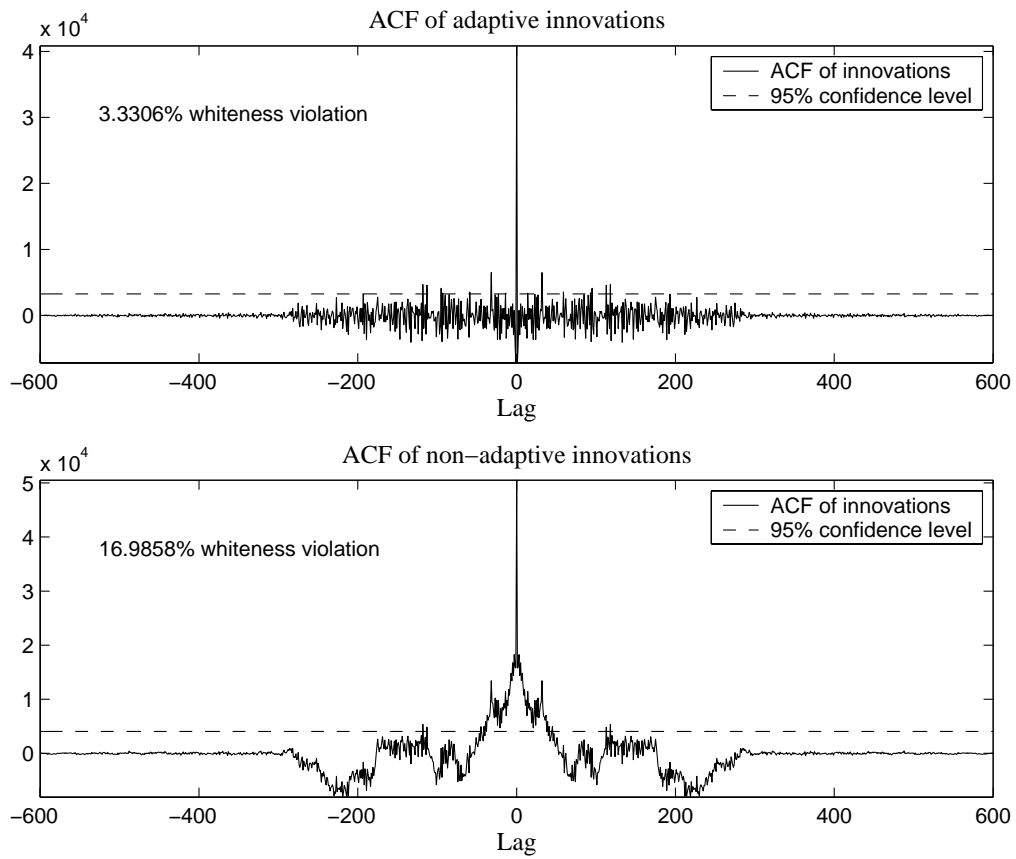


Figure 4.7: Autocorrelation of the innovations for the adaptive and non-adaptive estimates of the process in Fig. 4.5.

general.

4.4 Innovation–correlation method

The innovation–correlation method is theoretically based, and is computationally attractive because it is noniterative. It is also well suited to stochastic problems in which the precise dynamics of the process are not known. The measurement residuals (innovations) comprise a zero-mean, white, Gaussian sequence when the model parameters are correct. Therefore, testing the innovation sequence for non-white behavior by computing the autocorrelation function indicates whether the model contains errors or not.

Mehra showed that the process noise variance in a Kalman filter could be correctly estimated using the autocorrelation function (ACF) of the innovation sequence to obtain the optimal filter [38] for the stationary case. Recently, this method has been extended to examine the innovation sequence locally, and thus update the estimate of the process noise locally for non-stationary processes [46]. In this dissertation, the method of [46] is implemented within a multiresolution framework.

The innovation sequence is a function of the process noise variance Q , which is estimated from the local autocorrelation function. This method produces asymptotically normal, unbiased, and consistent estimates of Q . The method of [46] uses a batch of recently computed estimates to update the process noise variance. The innovations are the prediction error, and thus represent the “innovative” information provided to the estimate via the observations:

$$\nu_k = y_k - H_k \hat{x}_{k|k-1} = H_k e_k + v_k \quad (4.5)$$

Here, $e_k = x_k - \hat{x}_{k|k-1}$ is the error between the true state and the project-ahead (*a priori*) estimate. Mehra [41] showed that when the model parameters are correct

and the noise processes are white, the innovations comprise a Gaussian white noise sequence.

Theorem 2. *The innovations ν_k comprise a white noise sequence when the filter is operating optimally.*

Proof: 2. By (3.3)–(3.5),

$$E[\nu_k \nu_j^T] = E[(He_k + v_k)(He_j + v_j)^T] \quad (4.6)$$

Since v_k is uncorrelated with e_j and v_j for $k \neq j$.

$$= E[He_k(He_j + v_j)^T] \quad (4.7)$$

Recall $y_k - H_k \hat{x}_{k|k-1} = H_k e_k + v_k$

$$= E[He_k(y_j + H \hat{x}_{j|j-1})^T] \quad (4.8)$$

For $k \neq j$, e_k is uncorrelated with y_j . Also, $\hat{x}_{j|j-1}$ depends only on y_l , where $l < j$.

$$E[\nu_k \nu_j^T] = 0, \quad \text{for } k > j \quad (4.9)$$

Similarly, $E[\nu_k \nu_j^T] = 0$, for $k < j$. Thus,

$$E[\nu_k \nu_j^T] = 0, \quad \forall k \neq j. \quad (4.10)$$

Q.E.D ■

Since e_k and v_k are assumed to be distributed Gaussian, ν_k is also distributed Gaussian since it is a linear combination of e_k and v_k .

If model errors exist, however, the assumptions about uncorrelated processes will be violated, thus $E[\nu_k \nu_j^T] \neq 0$ in general. It is therefore possible to detect

the presence of model errors by determining if ν_k forms a white sequence. If the innovations are not white, the sequence $ACF(\nu_k)$ is used to update Q :

$$ACF(\nu_k) = E[\nu_k \nu_{k-j}^T] = HE[e_k e_{k-j}^T]H^T + HE[e_k \nu_{k-j}^T] \quad \text{for } k \neq j \quad (4.11)$$

$E[e_k e_{k-j}^T]$ can be related to Q , which leads to an algorithm for updating Q . For $k \neq j$

$$E[e_k e_{k-j}^T] = [\Phi(I - K_{ss}H(k))]^k P_{ss} \quad (4.12)$$

$$E[e_k \nu_{k-j}^T] = -[\Phi(I - K_{ss}H(k))]^{k-1} \Phi K_{ss} R \quad (4.13)$$

$$\begin{aligned} P_{ss} &= \Phi(I - K_{ss}H(k))P_{ss}(I - H^T(k)K_{ss}^T)\Phi^T + \\ &\quad \Phi K_{ss} R(k)K_{ss}^T \Phi^T + Q \end{aligned} \quad (4.14)$$

where the subscript ss refers to steady state. For $k = j$,

$$E[\nu_k \nu_{k-j}^T] = H(k)P_{ss}H^T(k) + R(k) \quad (4.15)$$

The covariance matrix for the innovations $E[\nu_k \nu_{k-j}^T]$ is denoted by $C(k)$

$$C(k) = \begin{cases} H(k)[\Phi(I - K_{ss}H(k))]^{k-1} \\ \quad \cdot \Phi(P_{ss}H^T(k) - K_{ss}C_0) & \text{for } k \neq 0 \\ H(k)P_{ss}H^T(k) + R(k) & \text{for } k = 0 \end{cases} \quad (4.16)$$

where k denotes the lag of the autocorrelation function. For scalar, zero-mean processes, $C(k)$ is the autocorrelation function.

4.4.1 Test for optimality

The whiteness of $\nu(k)$ is determined from the estimates of $C(k)$, given by $\hat{C}(k)$:

$$\hat{C}(k) = \frac{1}{N_b} \sum_{j=k}^{N_b-1} \nu(j) \nu^T(j-k) \quad k = 0, 1, \dots, N_{lag} \quad (4.17)$$

The data are processed in batches of N_b trailing samples. For a particular batch, N_{lag} samples of $\hat{C}(k)$ are calculated ($N_{lag} < N_b$). The *ACF* of the innovations is calculated and the 95% confidence interval test for a Gaussian random variable is used to test for whiteness [46]:

$$\text{for } C_{95\%} = 1.96\sigma_c, \quad P\{-C_{95\%} \leq C \leq C_{95\%}\} = 0.95 \quad (4.18)$$

where σ_c is the standard deviation of $C(k)$. If more than 5% of the *ACF* samples, not including the zero lag sample, fall outside the range $(-C_{95\%}, C_{95\%})$, then $\nu(k)$ is considered non-white.

4.4.2 Estimation of model parameters

Once $\hat{C}(k)$ is computed in (4.17), estimates of $P_{ss}H^T$ are obtained for each k using (4.16). The resulting N_{lag} recursive equations can be organized into a matrix calculation:

$$\begin{bmatrix} H(k)\Phi \\ H(k)\Phi^2 \\ \vdots \\ H(k)\Phi^{N_{lag}} \end{bmatrix} [P_{ss}H^T] - \begin{bmatrix} \hat{C}_1 + H(k)\Phi K_{ss}\hat{C}_0 \\ \hat{C}_2 + H(k)\Phi K_{ss}\hat{C}_1 + H(k)\Phi^2 K_{ss}\hat{C}_0 \\ \vdots \\ \hat{C}_{N_{lag}} + H(k)\Phi K_{ss}\hat{C}_{N_{lag}-1} + \dots + H(k)\Phi^{N_{lag}} K_{ss}\hat{C}_0 \end{bmatrix} = 0 \quad (4.19)$$

Using the approach of [38], the following expression is used to solve for \hat{Q} . The argument k is dropped because the estimate is made for each batch of N_b samples.

For $j = 1, \dots, N_{lag}$

$$\begin{aligned} \sum_{i=0}^{j-1} H \Phi^i \hat{Q} (\Phi^T)^{i-j} H^T &= \widehat{H P_{ss}} (\Phi^T)^{-j} H^T - H \Phi^j \widehat{P_{ss} H^T} \\ &\quad - \sum_{i=0}^{j-1} H \Phi^i B (\Phi^T)^{i-j} H^T \end{aligned} \quad (4.20)$$

where

$$B = \Phi (K_{ss} \hat{C}_0 K_{ss}^T - \widehat{P_{ss} H^T} K_{ss}^T - K_{ss} \widehat{H P_{ss}}) \Phi^T \quad (4.21)$$

$\widehat{P_{ss} H^T}$ is obtained by solving (4.19). Since $P_{ss} = P_{ss}^T$, $(\widehat{P_{ss} H^T})^T = \widehat{H P_{ss}}$. Equations (4.20) and (4.21) yield a set of linear equations that can be used to solve for each element in the \hat{Q} matrix. For the scalar case, \hat{Q} is substituted for Q in subsequent Kalman filter calculations:

$$\hat{Q} = \frac{\widehat{H P_{ss}} \Phi^{-j} H - H \Phi^j \widehat{P_{ss} H^T} - \sum_{i=0}^{j-1} \Phi^i B H^2}{\sum_{i=0}^{j-1} \Phi^{2i-j} H^2} \quad (4.22)$$

The innovation–correlation method was presented by [41] [38] with the assumption that the process noise is WSS over the region of support of the data. It was subsequently extended by [46] to accommodate process noise that is only piecewise WSS. The primary extension was to implement the algorithm over a finite number of estimates (batches) rather than over the entire set of estimates. Thus K , R , and Q can be locally updated. This extension to piecewise WSS processes is important for processing remote sensing images since images of terrain are not WSS in general. For example, the first and second moments of ground elevations and vegetation heights are different in a flat prairie and a hilly forest. An image containing both types of terrain therefore contains non-stationary data.

In this dissertation, I implement the method of [46] in a multiscale framework so that the MKS model parameters presented in Chapter 3 are not constrained to be uniform at a given scale. Since the multiscale pyramid depth only grows as

\log_2 of the image width, it is impractical to update Q in scale directly using this method. For example, a 1024×1024 image corresponds to pyramid depth of 10. Instead, at a given scale in the pyramid where complete (dense) INSAR observations reside, Kalman filters are applied in the spatial dimensions to update Q locally. That information is then used to update Q in the multiscale filter. This method is non-iterative.

4.5 Multidimensional state model

Combining the multiscale Kalman filter in Chapter 3 and the adaptive estimation of Q in Section 4.4 can be viewed in a general multidimensional state-space framework. A linear multidimensional state model, known as the Fornasini-Marchesini Form II (FM-II) state model, was developed in [47]. This representation can admit process dynamics in multiple dimensions. The original model considers bases of uniform support, such as image row and column. However, it can be extended to admit a dimension of non-uniform support, such as scale, if a merge step is incorporated, as in the MKS algorithm. In this way, a process model for $x(s)$ can be written that accounts for process dynamics in scale and spatial dimensions. The node index s is written in terms of these dimensions $s = \{m, i, j\}$, where m represents scale (level of the quadtree) and (i, j) are the image pixel coordinates. The fine-to-coarse state model is given by

$$\begin{aligned}
x(m, i, j) &= \Phi(m \mid m+1) \frac{1}{4} \sum_{l=0}^1 \sum_{k=0}^1 x(m+1, 2i-l, 2j-k) & (4.23) \\
&+ \Phi(i \mid i-1) x(m, i-1, j) + \Phi(j \mid j-1) x(m, i, j-1) \\
&+ \Gamma(m \mid m+1) \frac{1}{4} \sum_{l=0}^1 \sum_{k=0}^1 w(m+1, 2i-l, 2j-k) \\
&+ \Gamma(i \mid i-1) w(m, i-1, j) + \Gamma(j \mid j-1) w(m, i, j-1)
\end{aligned}$$

where $\{(m+1, 2i-1, 2j), (m+1, 2i, 2j), (m+1, 2i-1, 2j-1), (m+1, 2i, 2j-1), (m, i-1, j), (m, i, j-1)\}$ is the set of nodes from which the priors in the upward filter are derived.

Formulating the adaptive estimation algorithm under this generalized process model can lead to difficulties in determining regions of support for the autocorrelation function. In this dissertation, the Kalman filtering is restricted to a sequence of 1-D filter operations, and all process dynamics are assumed separable. This simplification is used to make the integration of the spatially-adaptive estimation into the MKS framework tractable, but it is not a fundamental limitation of the FM-II model.

Fig. 4.8 graphically depicts how the information from the set of priors is provided to the current estimate in the spatial filter. The filter starts at the base of the quadtree ($m = M$) using the standard multiscale procedure presented in Chapter 3. Upon reaching a level that contains dense observations ($m = M - 1$ in this example), two 1-D Kalman filters operate along the rows and columns of the image. The *a priori* estimates of the spatial filters incorporate the *a priori* multiscale estimates at $m = M - 1$. The fine-to-coarse process noise variance $Q(M - 1, i, j)$ is updated by the row-wise filter if non-white innovations are detected. $Q(M - 1, i, j)$ may again be updated as the column-wise filter operates on the columns. The $Q(M - 1, i, j)$ arrays from the row and column filtering operations are averaged together to obtain a final estimate of the spatially varying $\hat{Q}(M - 1, i, j)$.

The process noise is updated by each spatial filter in blocks of N_b pixels, where the value of N_b is chosen to balance the coarseness of the adaptive updating and variance of the *ACF* estimates. Taking N as the total number of pixels in a row or column of the image, N_b satisfies $N_b \leq N/2$ so at least one complete batch of updated parameters may be incorporated into the final estimates in each row and column. Yet, if N_b is chosen too small, the estimates of the *ACF*(ν) over N_b pixels

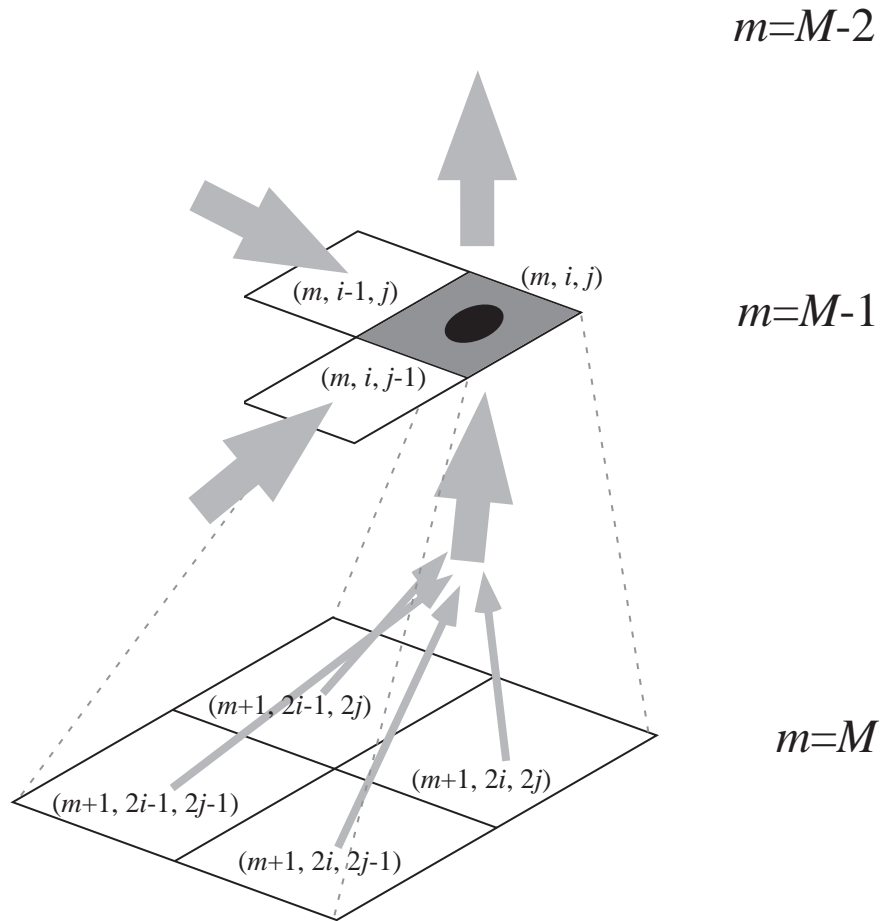


Figure 4.8: Graphical depiction of how information from the scale-wise filter is provided to the spatial filter. The current estimate is represented by the shaded pixel.

are unreliable because the variance of the sample ACF is inversely proportional to N_b . In this work, $N_b = N/4$

The spatial portion of the Kalman filter is applied only at levels in the quadtree where dense observations are present so that the innovation sequences have uniform support. The $\hat{Q}(M - 1, i, j)$ values from the spatially-adaptive filtering are used in the scale-wise component of the filter at the next coarser level ($m = M - 2$) and propagated up to the root node. The filter proceeds with the standard multiscale algorithm up the remainder of the quadtree and back down the quadtree.

4.5.1 Computational complexity

This approach to the adaptive estimation of the process noise is non-iterative. A set of N_{lag} simultaneous linear equations is solved for each batch of N_b samples. For example, with N as the number of samples in each row of an image, Q is updated up to $N/N_b - 1$ times per row, assuming the innovations are nonwhite throughout the image. For images between 128×128 and 512×512 , the AMKS algorithm required roughly 15% more computation time than the standard MKS algorithm.

4.5.2 Simulations

Images of natural terrain are very complicated and only approximately satisfy assumptions of statistical behavior, such as white process noise and measurement noise. Simulation results are presented here to investigate the adaptive multiscale Kalman smoother characteristics. Two different simulations are considered. The first simulates two different processes in a single image, thus no single value for Q can be correct for the entire image. One process is slowly varying and the other is rapidly varying to simulate an image of two different terrain types. The second

simulation consists of a single stochastic process, but the initial estimate for the process noise is chosen too low. In both cases, the adaptive algorithm is able to detect and react to resulting errors in the Kalman model.

Figure 4.9 depicts the finest scale realization of a composite multiscale stochastic process. A rapidly varying process with rectangular support (region 2) is embedded in a slowly varying process (region 1). Region 1 represents minimal topography or vegetation, e.g. grassland, and region 2 represents significant topography or tall vegetation, e.g. forest. Each process was generated with a multiscale linear process model (3.8), assuming fractional Brownian motion dynamics (3.27). For process 1, $\Gamma_0 = 2$ and $\mu = 3$. For process 2, $\Gamma_0 = 4$ and $\mu = 1$. Thus, process 2 has four times the variance at the root node, and its power decreases more slowly with spatial frequency.

Figure 4.10 shows observations of the true process taken at two different scales. These simulated observations are representative of INSAR acquired at a coarse resolution and LIDAR acquired at the finest resolution. The finest scale ($m = M$) observations cover only a diagonal swath through the process, running from the lower left corner to the upper right corner. This swath is used to simulate incomplete coverage often associated with LIDAR coverage in an INSAR scene.

Each “sensor” actually observes a different process, one at scale $m = M$ and the other at scale $m = M - 1$. Each sensor has its own associated signal-dependent measurement noise. The coarse observations have a larger measurement error variance because that is typical for INSAR and LIDAR systems. The measurement error of both sensors increases where the variance of the observed process increases. This is also representative of actual INSAR and LIDAR observations.

Results from the basic MKS algorithm are shown in Fig. 4.11, and results from the AMKS algorithm are shown in Fig. 4.12. There is little apparent difference in the fused estimate images, but Fig. 4.13 indicates that the fine-to-coarse process

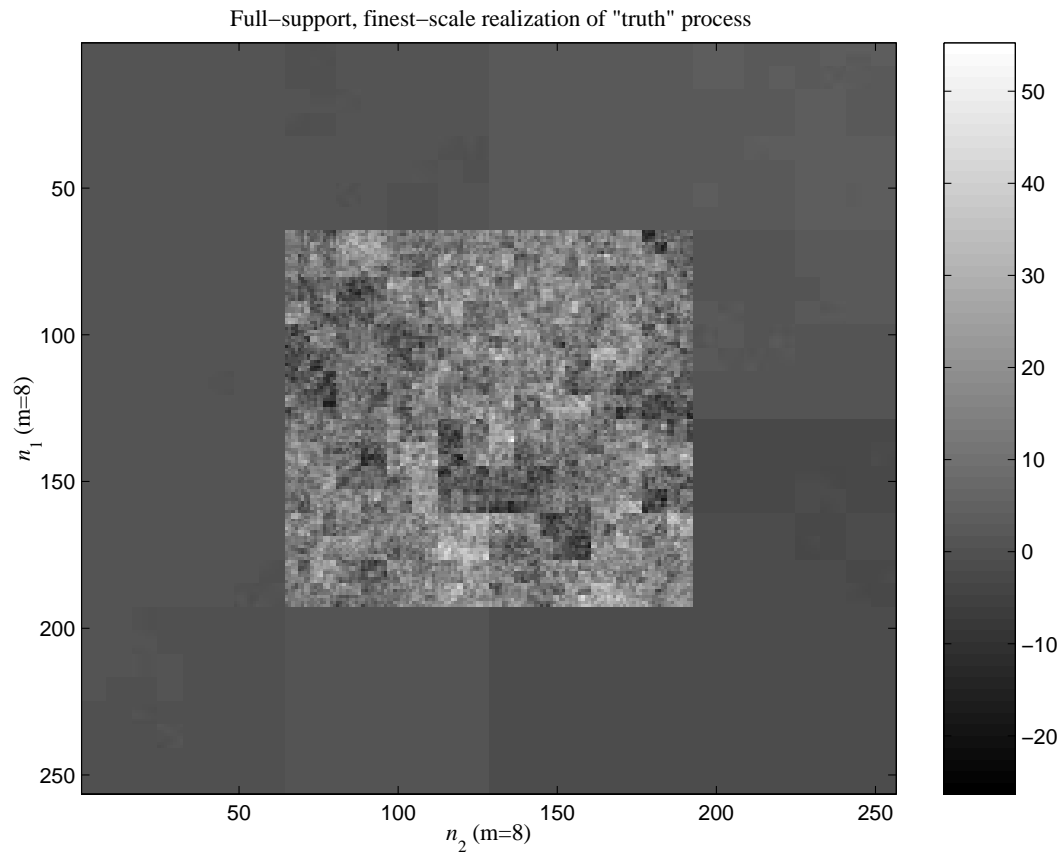


Figure 4.9: A composite multiscale process generated to simulate the *true* process at the finest scale in the multiscale pyramid. A rapidly varying process (region 2) is embedded in a slowly-varying process (region 1).

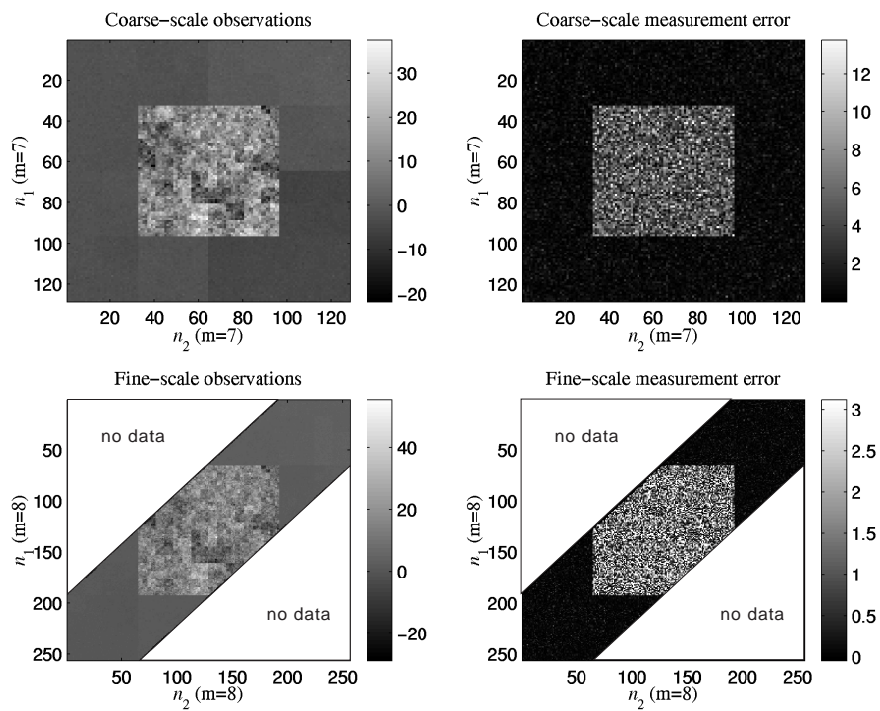


Figure 4.10: Simulated observations and measurement uncertainties. The region of support for the fine-scale data is a diagonal swath extending from the lower left corner to the upper right corner of that image.

Table 4.1: Mean squared error of the coarse observations in Fig. 4.10 relative to the true process values. H evaluated at $m = M$.

height estimates	
$H=\{0 \text{ or } 1\}$	18.5 m ²
$H=\{1\}$	22.3 m ²
$H=\{0\}$	15.4 m ²

noise variance was locally updated by the AMKS algorithm. The Q image one level above the dense observations (Q_7 in this example) indicates that the algorithm revised Q upward after passing over region 2. This is the expected result because the true process noise in region 2 is greater than the initial value of Q used at the finer scales. The improvement in the estimated process noise variance is confirmed by examining the mean differences between the true Q and the updated Q and between the true Q and the Q used in the MKS algorithm. For this simulation, $\overline{Q_{true} - Q_{AMKS}}$ is 53% smaller than $\overline{Q_{true} - Q_{MKS}}$.

Mean squared error results are given for “raw” coarse observations (Table 4.1), MKS fused estimates (Table 4.2), and the AMKS fused estimates (Table 4.3). In Table 4.1, error is defined using the coarse observations because observations at the finest scale often have incomplete coverage. The coarse observations are upsampled by a factor of 2 so they can be differenced with the true process at the finest scale. A dramatic decrease in MSE occurs when the MKS algorithm is applied. This is expected since even a suboptimal Kalman filter can produce estimates with much smaller error variance than the original observations, as shown in Figs. 4.2 and 4.3. The MSE obtained with the AMKS algorithm is smaller than the MKS result, albeit only by a small amount. This is consistent with the perfectly adaptive estimation simulated in Fig. 4.3, where a small error in Q led to a modest improvement of 5% in the MSE of the adaptive filter.

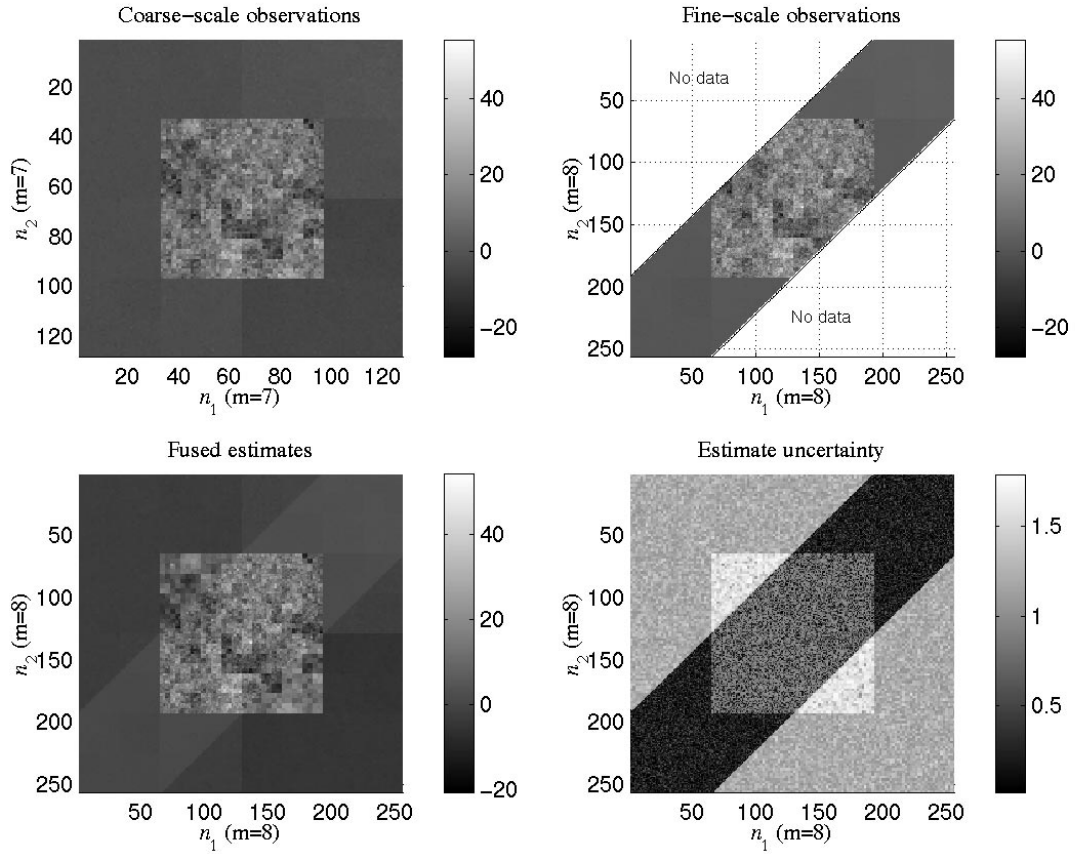


Figure 4.11: Fused estimates and estimate uncertainty from the MKS algorithm operating on the process shown in Fig. 4.9.

Table 4.2: Mean squared error of the MKS estimates relative to the true process values. H evaluated at $m = M$.

bare surface elevation	
$H = \{0 \text{ or } 1\}$	9.16 m ²
$H = \{1\}$	2.12 m ²
$H = \{0\}$	14.7 m ²

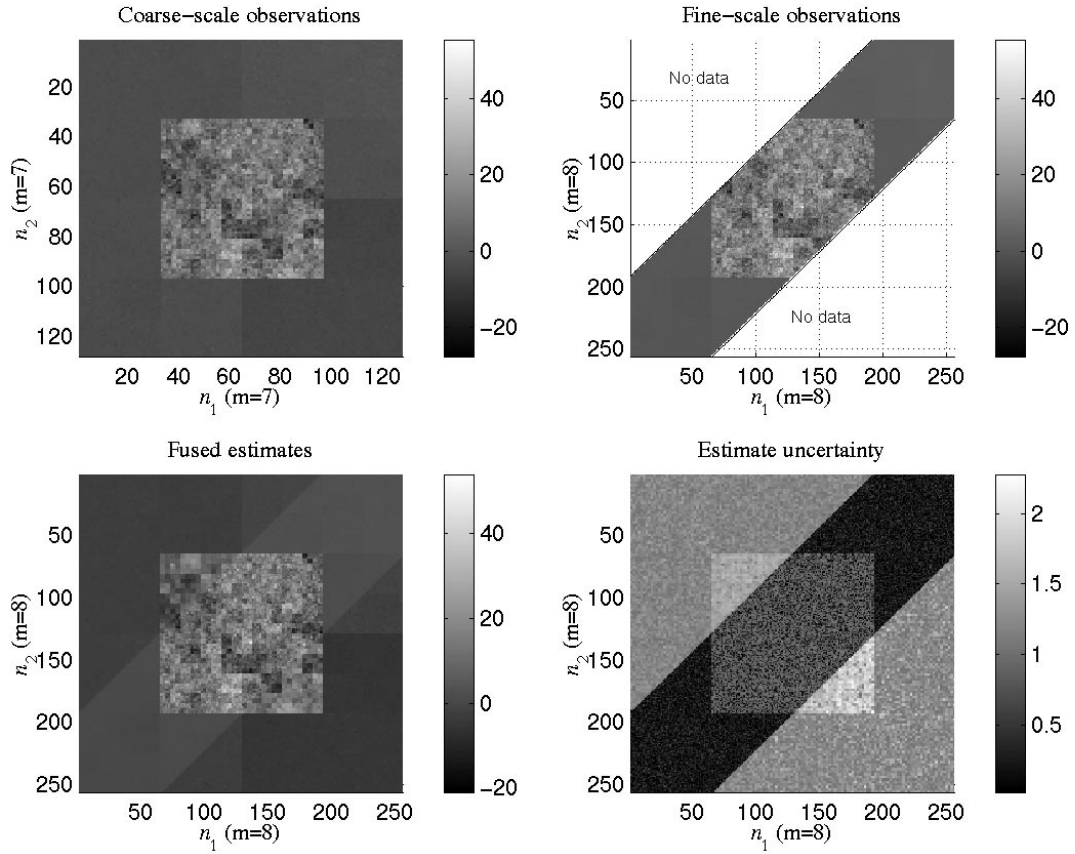


Figure 4.12: Fused estimates and estimate uncertainty from the AMKS algorithm operating on the process shown in Fig. 4.9.

Table 4.3: Mean squared error of the AMKS estimates relative to the true process values. H evaluated at $m = M$.

bare surface elevation	
$H = \{0 \text{ or } 1\}$	9.13 m ²
$H = \{1\}$	2.12 m ²
$H = \{0\}$	14.7 m ²

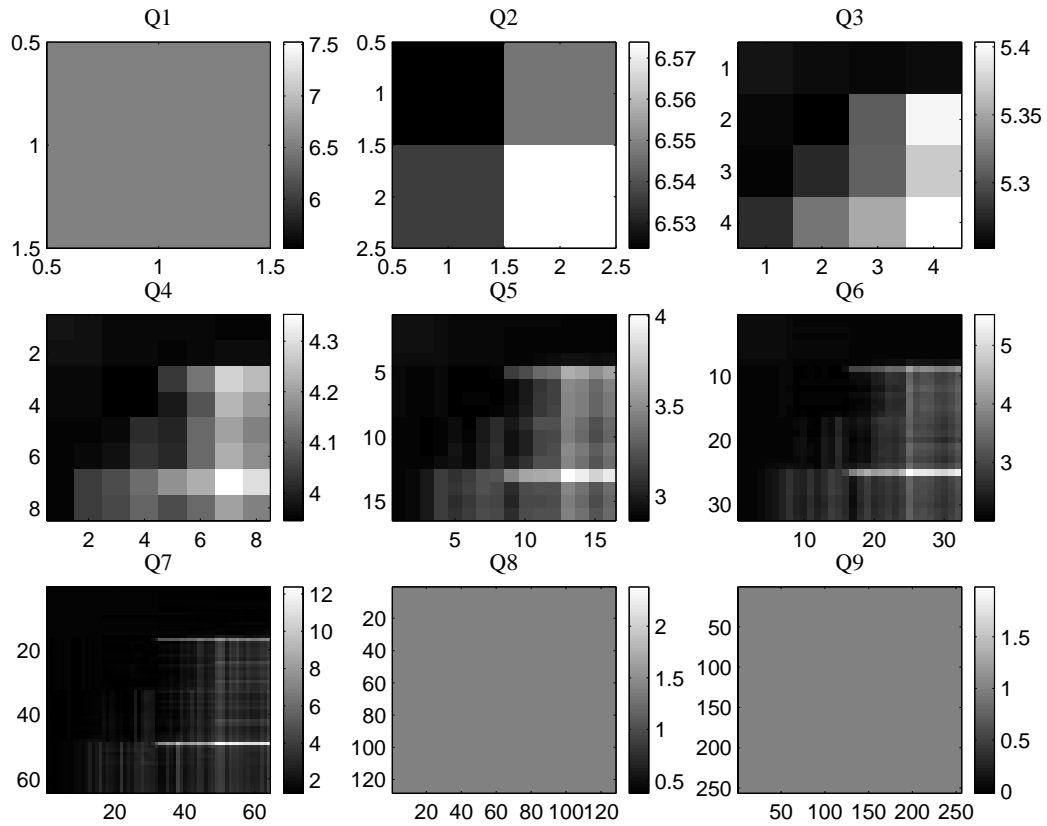


Figure 4.13: Fine-to-coarse process noise variances $Q(m, i, j)$ from the AMKS algorithm. The initial value for Q at $m = 9$ is uniform and approximately 1.0. At $m = 7$, the spatially estimated Q is incorporated back into the multiscale filter. Q is revised upward as the spatial filters pass over region 2.

The impact of adaptively filtering the observations depends on both the data and the size of the set of samples N_b used to test for whiteness in the innovations. The impact generally increases when there are large errors in the initial Q or large discontinuities in the process. In the current example, the simulated processes were made to have similar magnitudes and variability to imaged vegetation heights in meters. As a result, the difference in the true Q values in region 1 and region 2 was relatively small. The impact also tends to increase when the window size N_b decreases.

A simpler example is constructed to further study the behavior of the AMKS algorithm. A 2-D simulation of a WSS process is shown in Fig. 4.14. Here, the variability is made larger than would be encountered in a typical INSAR image so that large errors in the initial Q can be tested. There is no explicit non-stationarity in this process, but the Γ_0 and μ are intentionally chosen to yield an underestimated fine-to-coarse process noise variance Q . The MKS algorithm is applied to the simulated data using the underestimated values for Q . The AMKS algorithm is then applied to the data. It starts off with the same Q as the MKS model. Upon reaching the scale where the dense observations are present ($m = M - 1$), the spatial filters are applied to the image. Fig. 4.15 shows the Q arrays at each stage in the upward Kalman filter. The values of Q significantly increase as the spatial filters operate along the rows and columns. The general trend is for Q to increase towards the lower right corner of the images.

In Fig. 4.16, the estimates and estimate uncertainty from the AMKS algorithm are shown. The estimate image is similar to that obtained in the MKS result, but the estimate uncertainty image clearly shows the effects of the updated Q . The uncertainty is larger in the lower right corner of the image. This can be explained from the Kalman filter equations in Fig. 3.3. When Q is revised upward, the error covariance will also be revised upward, which in turn will increase the uncertainty

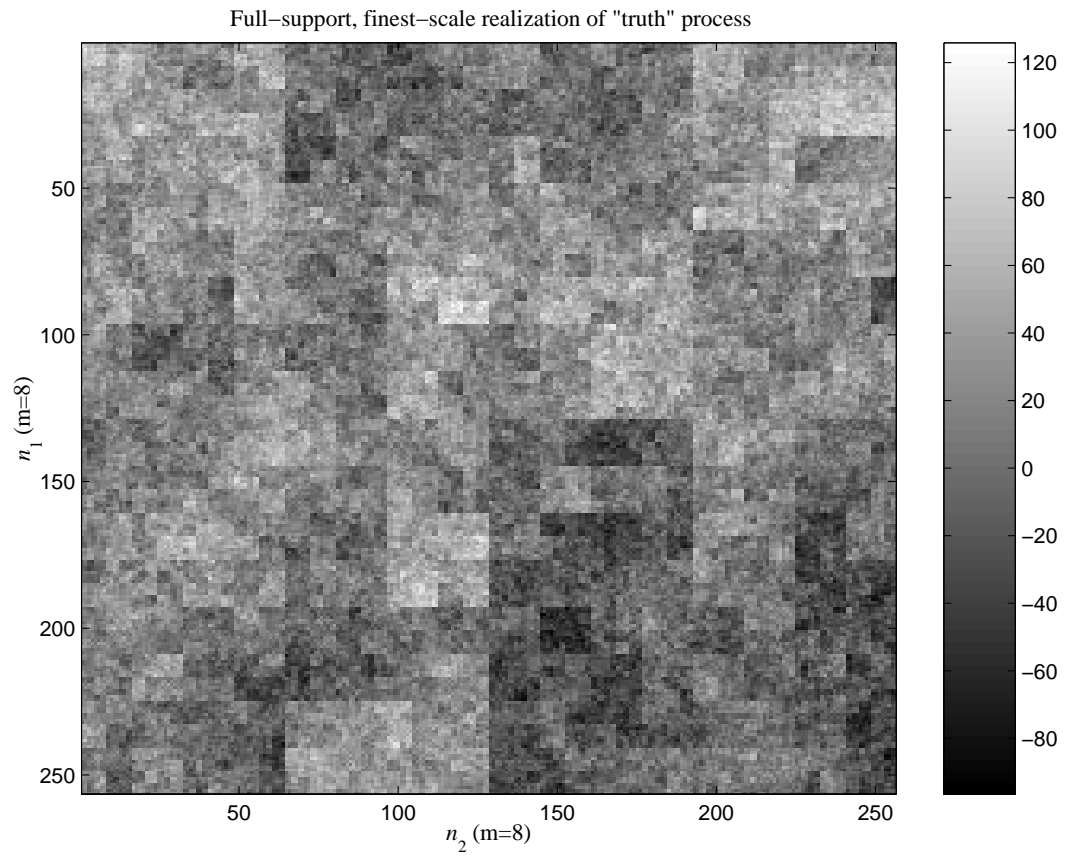


Figure 4.14: A WSS multiscale process generated to simulate the *true* process at the finest scale in the multiscale pyramid. The initial Q in the MKS and AMKS filters is intentionally chosen to correspond to a smoother process.

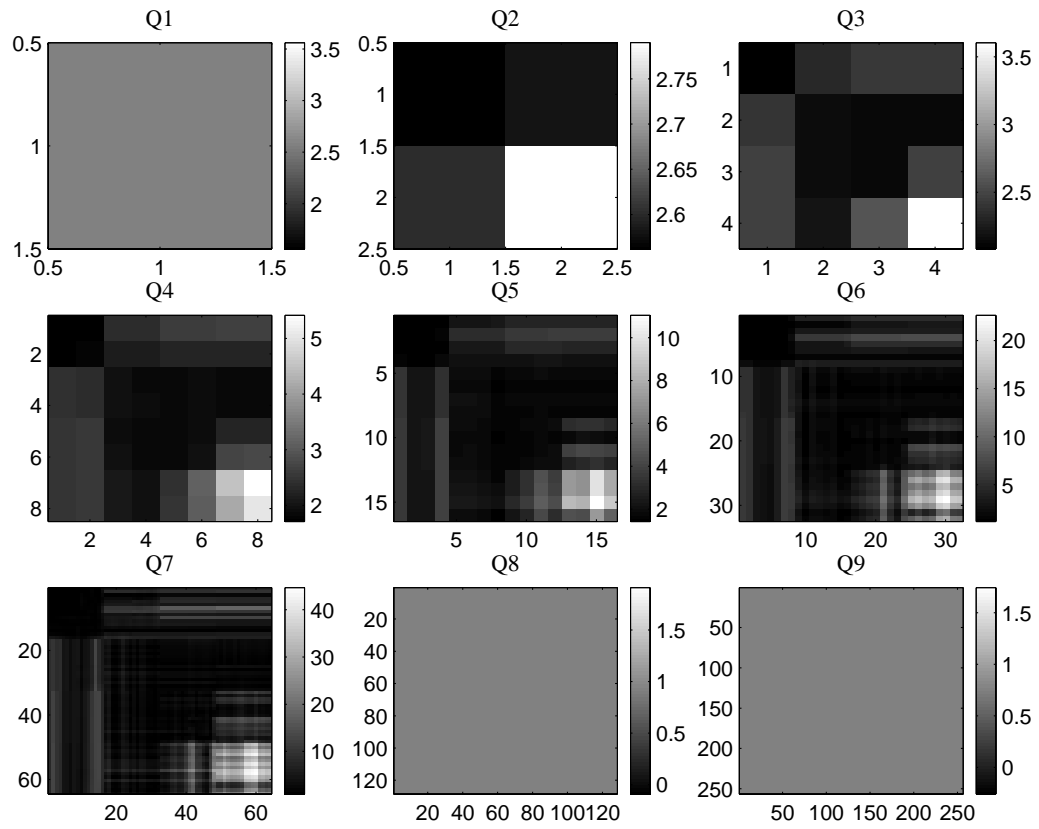


Figure 4.15: Fine-to-coarse process noise variances $Q(m, i, j)$ from the AMKS algorithm.

Table 4.4: Mean squared error of the coarse observations, the MKS estimates, and the AMKS estimates relative to the true process values. $H = \{0 \text{ or } 1\}$ and evaluated at $m = M$.

bare surface elevation	
prior to data fusion	1894.m ²
after MKS	419.7 m ²
after AMKS	406.5 m ²

of the Kalman smoothed estimates in (3.7).

Transects of the AMKS estimates, coarse observations, and true state process are shown in Fig. 4.17. In the region where fine-resolution data are available, the AMKS estimates correspond relatively well to the true process. Where only coarse data are available, the AMKS estimates tend to over smooth. This is expected since the initial Q was chosen too small, which causes the Kalman filter to smooth the observations too much. In the MKS algorithm, there is no mechanism to correct for this. However, in the AMKS results, the updated Q approaches the true Q for large values of column number and row number in Fig. 4.16. The AMKS algorithm thus reduces the over smoothing in the right side of the plot.

Table 4.4 lists the global MSE results before data fusion, after MKS, and after AMKS. The AMKS algorithm achieved a reduction in MSE of approximately 3% over that of the MKS algorithm. In addition to the reduction in MSE, The difference between the average value of the true Q used to generate the process and the updated Q computed in the AMKS algorithm was 12% less than that of the Q used in the MKS algorithm.

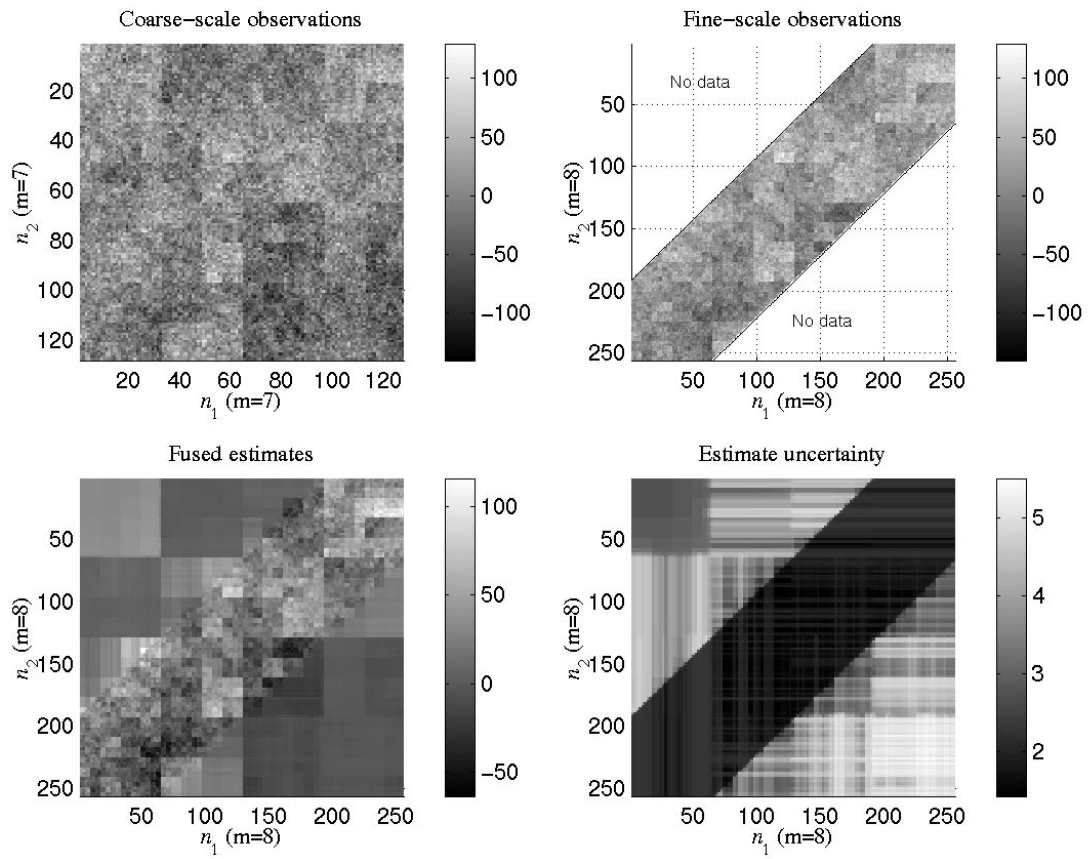


Figure 4.16: Fused estimates and estimate uncertainty from the AMKS algorithm operating on the process shown in Fig. 4.14.

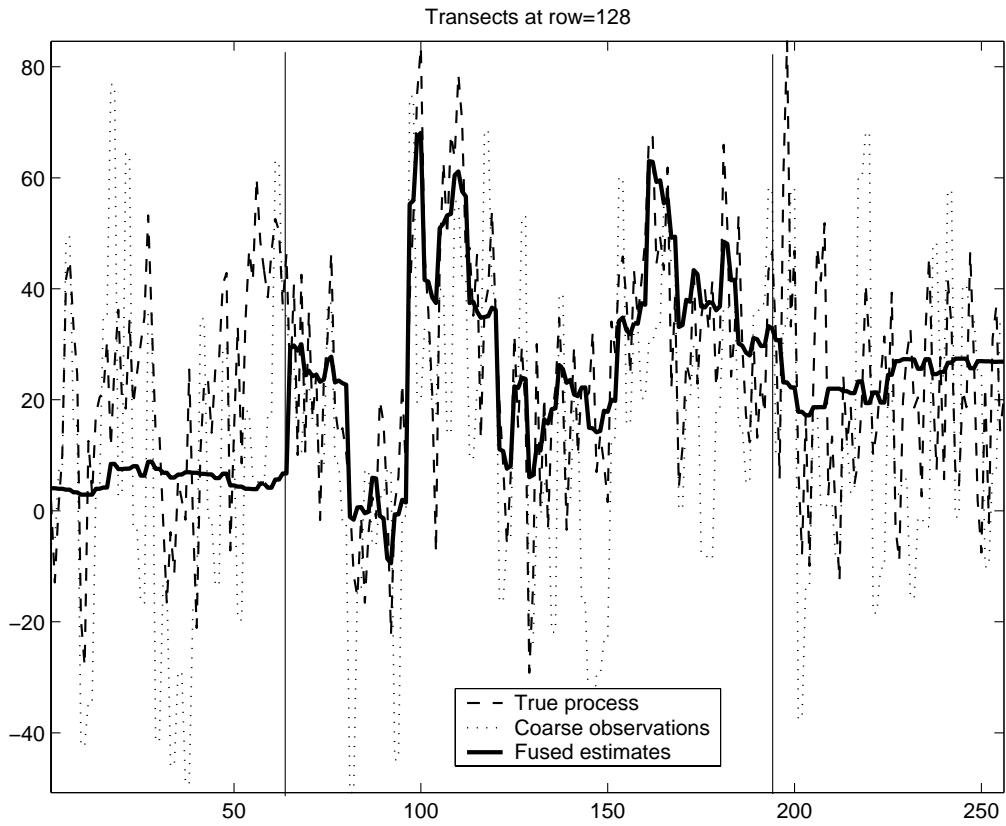


Figure 4.17: Transects from row 128 in the AMKS estimates. The vertical lines delineate the region of support of fine-resolution data in the central portion of the plot.

4.6 Conclusions

In this chapter, a new linear process model was presented that allows the formulation of a separable Kalman filter that operates in scale and spatial dimensions. Using a separable formulation allowed the straight forward incorporation of existing adaptive estimation techniques. This is an important feature because it will also facilitate the incorporation of any of the other adaptive estimation techniques mentioned in Section 4.3.

The resulting adaptive MKS (AMKS) algorithm represents an improvement over the MKS algorithm because it provides a mechanism to detect and adapt to spatial variations in the process noise variance. The AMKS algorithm will generally have a greater impact on the final values of Q and the estimates when the errors in the initial value of Q or any local differences in the variance of the true process (non-stationarities) are large. As was seen in the 1-D simulations in Section 4.2, even a perfectly adaptive filter will have minimal impact if the initial errors in Q are less than an order of magnitude.

When the full multiscale multidimensional case is considered along with fusion of multiple data sets, analytically deriving the mutual effects of the relative values of Q and R on the adaptive estimation is difficult. But the general trend holds that if errors in Q are large, the AMKS algorithm can impact the final estimates significantly. In addition to the effects of adaptive estimation on the final estimates, the AMKS algorithm is a useful tool for providing insight into the local variability of the process being estimated. Thus, it can be used to discover where errors in the MKS model might be present. The standard MKS algorithm and the new AMKS algorithm are applied to several sets of remote sensing data in the following chapter.

Chapter 5

Fusing INSAR and LIDAR Data

5.1 Introduction

Herein, the estimation framework is applied to remotely sensed data over coastal and riparian environments in Texas and a semi-arid environment in Australia. The physical modeling of INSAR and LIDAR discussed in Chapter 2 was performed on all of the Texas data to account for the vegetation effects. The MKS algorithm from Chapter 3 is applied to results from the physical modeling in Section 5.2 over a Texas coastal site. Estimates from fused INSAR and LIDAR are examined. The smoothing effect of the MKS algorithm on a single data type is also examined. In Section 5.3, results from the AMKS algorithm in Chapter 4 are compared with results from the MKS algorithm. The two algorithms are applied to a riparian environment in Texas. A third environment, a semi-arid region in Australia, is examined in Section 5.4 to further demonstrate the applicability of the framework to various environments. While LIDAR data were not available over the Australian site, coarse-resolution space-borne INSAR data were available, along with two sets of airborne INSAR data. The three different INSAR data sets, acquired at differ-

ent resolutions and times, are fused to obtain improved topography estimates. The results are evaluated in Section 5.5.

5.2 Physical modeling and MKS

Some of the issues associated with analyzing real data sets should be mentioned before considering results. Several steps must be taken to prepare the data, which have already undergone nominal processing into distributed data products, for analysis. The first step is to ground-range project the INSAR data so that it is free of the gross spatial distortion that accompanies imaging radar observations. The airborne INSAR data from the two baselines are not calibrated relative to each other, so relative biases must be removed in the correlation and height data. Also, a small number of large-magnitude outliers that are common in INSAR data must be removed. The LIDAR data are gridded from an irregularly-spaced sequence of time-tagged values into a uniform image lattice. Data sets acquired in vegetated areas must undergo the transformation into z_g and Δz_v values. Measurement uncertainty images are generated for each topographic image. Finally, all images must be co-registered so that common features in the images will have corresponding pixel coordinates.

5.2.1 Data description: Texas coastal environment

The Texas coastal test sites are located on Bolivar Peninsula, Texas, where there is overlapping INSAR and LIDAR coverage, and ground truth data have been acquired. The peninsula contains tree stands, grasslands, and bare fields, making it a good location to assess vegetation removal methods. The NASA/JPL TOPSAR sensor imaged the test site in a single pass. The imaging swath was 12 km in width and roughly 50 km in length. A portion of a corresponding polarimetric SAR scene

is shown in Fig. 5.1. Figure 5.2 shows some of the trees in the test site.

Data from approximately 10 overlapping flight lines from the ALTM sensor had to be combined to cover the region shown in Fig. 5.1. The 18-look TOPSAR data have a posting of $10\text{ m} \times 10\text{ m}$ after ground range projection. The LIDAR data were originally gridded to $10\text{ m} \times 10\text{ m}$ (higher resolution LIDAR products were not available for this area). Dense LIDAR coverage over large areas is not typical, so some of the LIDAR data were withheld from the data fusion to produce a more realistic acquisition scenario, as shown in Fig. 5.3. The INSAR data were averaged to a posting of $20\text{ m} \times 20\text{ m}$ to obtain a representative data set of sparse LIDAR coverage at the highest resolution and dense INSAR coverage at the next coarser resolution. Less than one quarter of the available LIDAR data were used in the data fusion. In the particular example, two adjacent rows are retained from every nine rows of the original LIDAR data, and the remaining seven rows are omitted. This sampling pattern represents a scaled version of the narrow and widely-spaced swaths that may be acquired by a spaceborne LIDAR system, such as the NASA Geoscience Laser Altimeter System.

Two sets of data are investigated, the Elm Grove and Melody Lane locations in Fig. 5.1. The sites are located at different incidence angles, which allows the evaluation of the estimation results with respect to location in the INSAR imagery. As noted in Subsection 2.2.2, the measurement uncertainty σ_h of the INSAR data is a function of incidence angle, and is generally larger in the far range. The Elm Grove site contains a large stand of trees, but its location in the far range of the INSAR data results in noisy estimates. Results for the Elm Grove site are presented in Subsection 5.2.2, and results for the Melody Lane site are contained in Subsection 5.2.3. Results based on only the INSAR data are presented in Subsection 5.2.4 to demonstrate that the smoothing effects of the MKS algorithm can improve the estimates even when no LIDAR data are available.

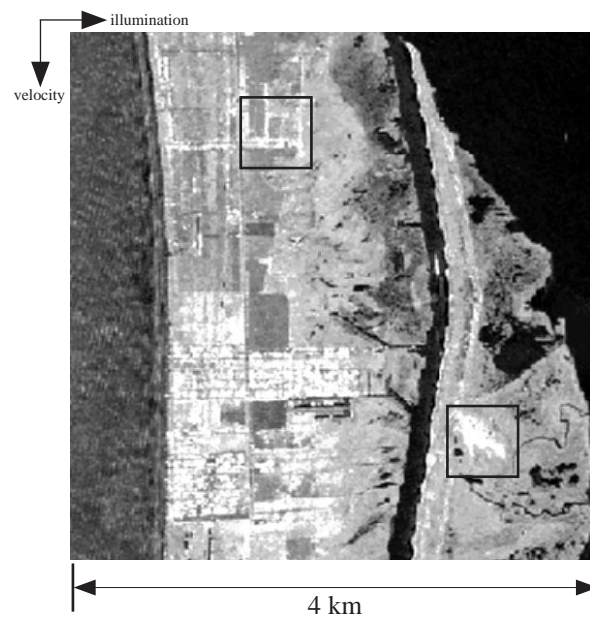


Figure 5.1: Polarimetric SAR image (L-band) of the test sites located on Bolivar, Texas. The boxes indicate portions of the image used in this analysis. The top subset is the Melody Lane site and the bottom subset is the Elm Grove site.



Figure 5.2: Photograph of trees at the Melody Lane site on Bolivar, Texas. The trees generally ranged between 6 m and 10 m in height.

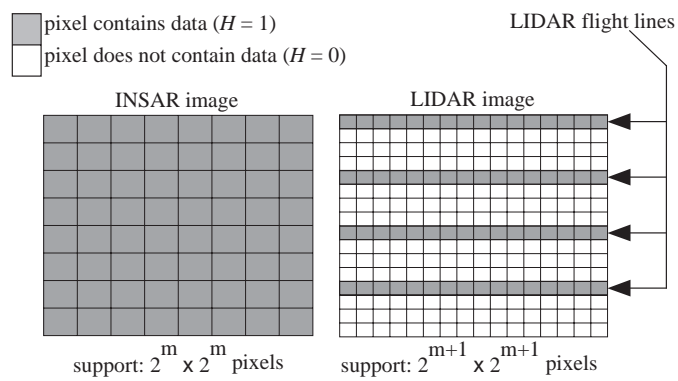


Figure 5.3: Image lattices indicating the characteristic support of the data in the sparse acquisition scenario.

5.2.2 Site 1 results

Estimates of the state variables, determined by inverting the INSAR scattering model and transforming the LIDAR data, serve as the observations in the MKS algorithm. Fig. 5.4 shows the input data for the MKS algorithm. The upper left image shows the estimated ground surface obtained from the INSAR inversion. Each pixel represents a $20\text{ m} \times 20\text{ m}$ area on the ground. The estimates exhibit a roughly uniform variability across the image, but adjacent pixels can differ by as much as 8 m, which is more than the actual topography. Thus, the INSAR inversion results comprise a noisy set of observations.

The estimated vegetation heights obtained from the INSAR inversion are shown in the lower left image. The bright oval feature is the grove of trees. The surrounding grassy area is 5 to 10 m lower. The right column shows the corresponding topography and vegetation heights derived from the LIDAR data. Each pixel in the LIDAR images covers a $10\text{ m} \times 10\text{ m}$ area on the ground. Several horizontal swaths of LIDAR data are withheld to mimic a sparse coverage, as indicated in Fig. 5.3. These pixels are shown with a thatched pattern in Fig. 5.4. The thin strips of LIDAR data appear homogeneous over larger areas because of the low noise floor of the LIDAR data. Comparison of the INSAR derived estimates for z_g and Δz_v with the LIDAR derived estimates in Fig. 5.4 indicates no significant bias in the INSAR inversion. The INSAR estimates are noisier than the LIDAR estimates, but they exhibit similar local average values.

The multiscale smoothing improves the estimates (in the MSE sense) of the ground surface and vegetation heights. Fig. 5.5 shows the estimates in the left column and estimate uncertainties in the right column, computed as the square root of the smoothed estimate variances. These are the estimate variances calculated in the Kalman filter. A horizontal banding is apparent in the uncertainty images that

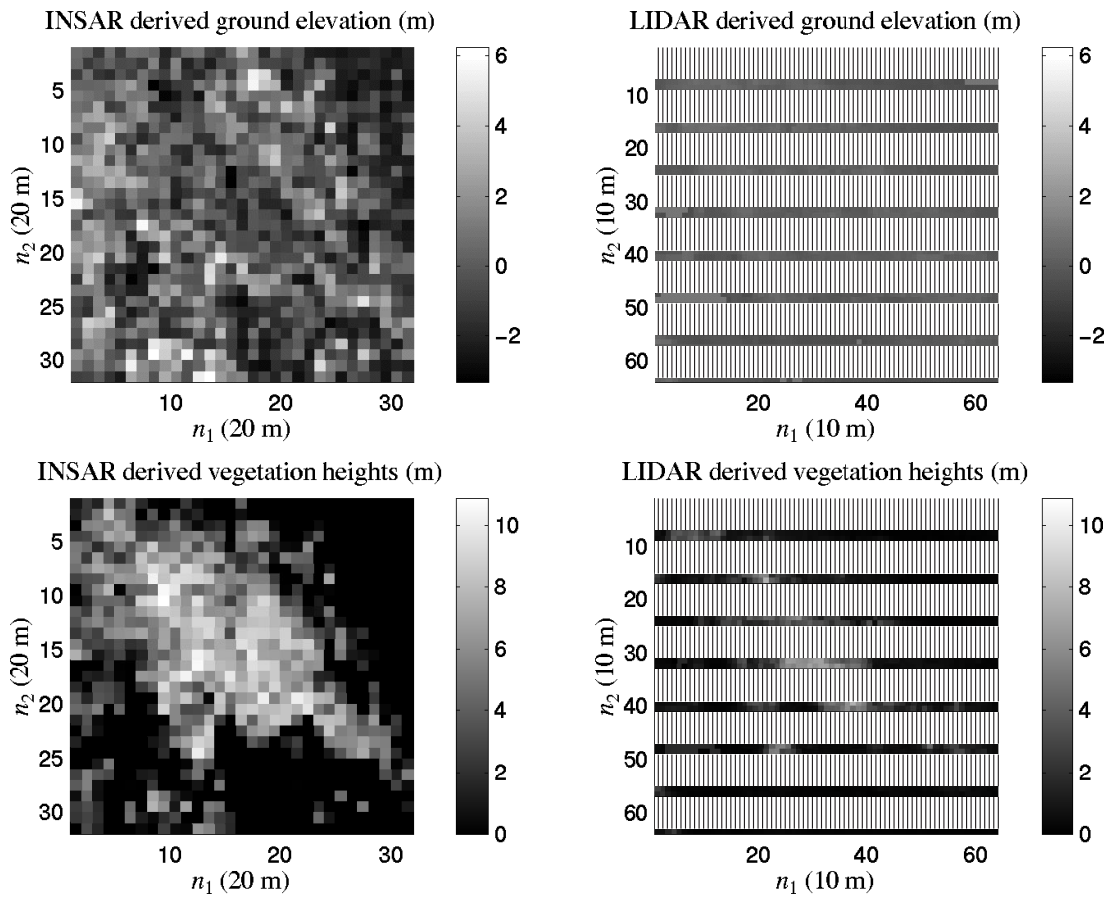


Figure 5.4: Transformed INSAR data (left column) and LIDAR data (right column) that serve as inputs to the multiscale estimation. Thatched areas indicate where LIDAR data were omitted.

corresponds to the locations where LIDAR data were omitted. This pattern occurs because the estimate uncertainty is greatest where the LIDAR data are absent, which is due to the smaller measurement error of the LIDAR data relative to the INSAR data.

The uncertainty is also large where there is significant vegetation because the measurement errors of both INSAR and LIDAR increase over highly variable targets. The uncertainty is greater for estimates of vegetation height than ground elevation because there is generally greater variability in the actual vegetation heights. This is reflected in the stochastic detail scaling function Γ . By matching the power spectra of the stochastic model and the data, as discussed in Section 3.4.2, values of $\Gamma_0=0.9$ and $\mu=1.5$ were obtained for the bare surface estimates, while values of $\Gamma_0=3.0$ and $\mu=2.0$ were obtained for the vegetation height estimates.

Fig. 5.6 shows the MKS results with the original INSAR and LIDAR data for a single row in the Elm Grove output images. The transects in Fig. 5.6 correspond to row 38 in the estimate images in Fig. 5.5. This is one of the rows that contains LIDAR data. A similar plot is shown in next section for a row where LIDAR were omitted. In Fig. 5.5, the MKS estimates of vegetation height do not track the LIDAR data as closely as initially expected. The reason for this is a relatively large measurement uncertainty for the LIDAR data. The uncertainty for the LIDAR was calculated as the standard deviation of heights in each grid cell (pixel). All values in the grid cell were weighted equally. Thus, LIDAR returns near the edge of the cell affected the calculated measurement error as much as returns near the center. If the returns were weighted inversely to their distance from the cell center, a smaller uncertainty would generally result, and the MKS estimates would track the LIDAR more closely.

In general, the expected trends are confirmed in Fig. 5.6. The LIDAR closely approximates the tops of the vegetation, while the INSAR penetrates into the vege-

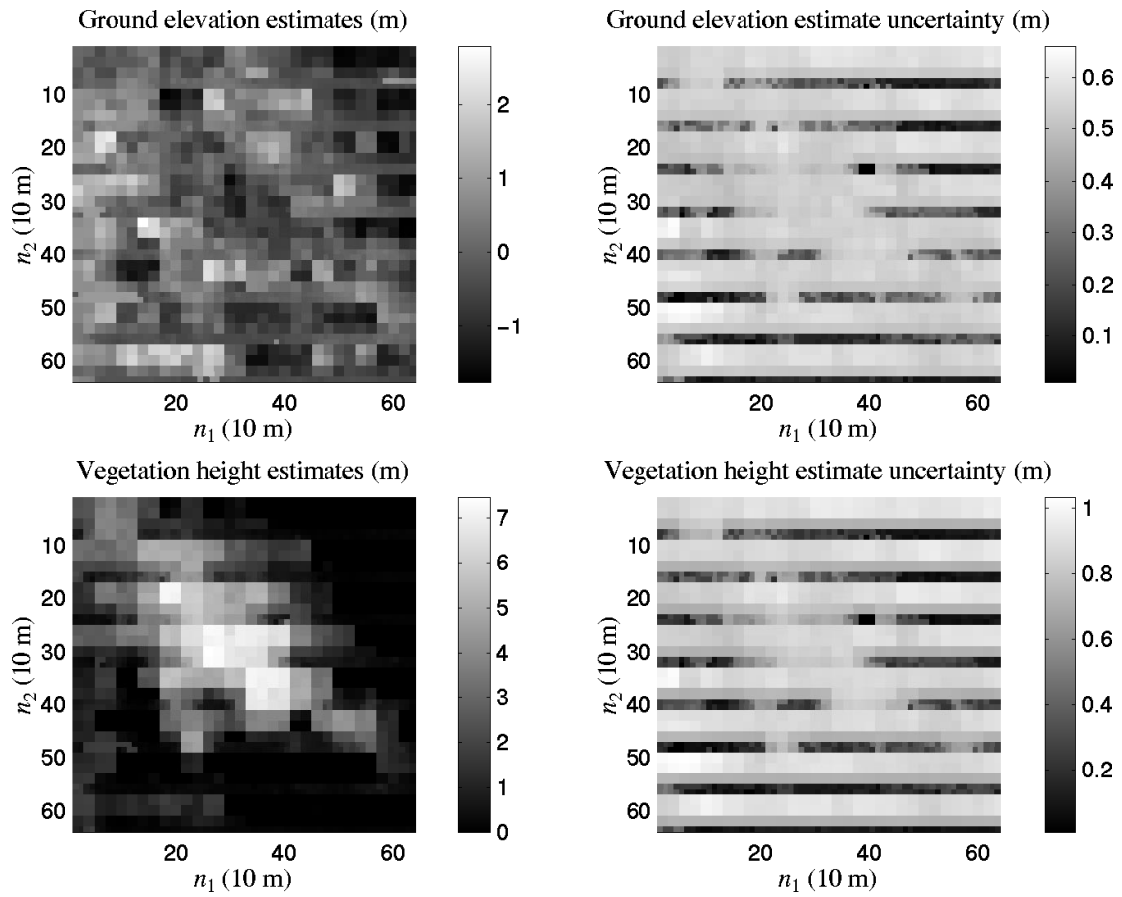


Figure 5.5: MKS estimates of ground elevations and vegetation heights (left column) and the corresponding estimate uncertainties (right column).

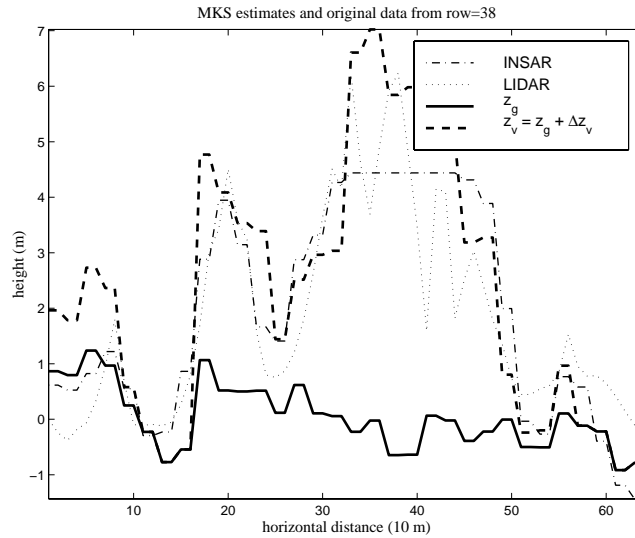


Figure 5.6: Comparison of data fusion results and original data for Elm Grove site. Data extracted from row 38 in Fig. 5.5.

tation. The lower resolution INSAR data fails to capture rapid changes in the vegetation heights that are apparent in the LIDAR. The MKS ground surface estimates \hat{z}_g are within one meter of the lowest points in the LIDAR (ground shots), whereas the vegetation height estimates $\Delta\hat{z}_v$ approximate the LIDAR heights. The $\Delta\hat{z}_v$ values track rapid changes in the vegetation better than the INSAR, and in some cases capture the full vegetation height. With MSE values of less than 1 m for z_g and less than 5 m for Δz_v , reliable estimates of the ground surface and the vegetation heights are obtained that retain much of the vertical and horizontal resolution of the LIDAR.

Local biases and large variances in the INSAR data can result in some negative Δz_v values in the final MKS estimates. The $\angle NICC$ phase values from the two TOPSAR baselines are not jointly calibrated, so the z_S data must be corrected for vertical offset and range tilt prior to any processing. This is accomplished using

known elevations in the near and far range of the INSAR imagery. Range tilt (in the direction of illumination) is a common artifact in INSAR data resulting from absolute phase ambiguity [10]. The range tilt correction removes only global bias between the transformed INSAR and LIDAR heights, however. Small (generally < 2 m) local biases due to data processing artifacts, such as incorrect motion compensation, may remain. In addition, surface scattering and non-zero ground slopes and SAR squint angles are not modeled. The simplified modeling can lead to large variances in the transformed INSAR data, which may cause some of the MKS estimates of vegetation height to be negative. Pixels with $\Delta \hat{z}_v < 0$ are assigned a value of zero in the final MKS estimates to maintain physically reasonable results.

The full (dense) LIDAR data are used for evaluating the results. Because the LIDAR data closely approximate ground truth, we define the estimate error to be the difference between the estimates and the corresponding transformed LIDAR observations. To see the effects of the MKS, we compute the error between the LIDAR and transformed INSAR estimates, as well as between the LIDAR and MKS estimates. The Kalman filter provides optimal estimates in the MSE sense. From the error definition, the mean squared error (MSE) can be computed for the whole image, where LIDAR was retained, and where it was omitted. Table 5.1 gives the MSE results obtained without any multiscale estimation. These results are obtained using just the transformed INSAR and LIDAR data. Table 5.2 shows the estimates and estimate errors after running the MKS algorithm. After the MKS data fusion, the percent reduction in global MSE is 90% for the bare surface elevations and 87% for vegetation heights.

Table 5.1: Mean squared error of the transformed INSAR data for the Elm Grove test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	5.56 m ²	$H=\{0 \text{ or } 1\}$	22.7 m ²
$H=\{1\}$	5.76 m ²	$H=\{1\}$	23.8 m ²
$H=\{0\}$	5.49 m ²	$H=\{0\}$	22.4 m ²

Table 5.2: Mean squared error after MKS for the Elm Grove test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	0.566 m ²	$H=\{0 \text{ or } 1\}$	3.00 m ²
$H=\{1\}$	0.064 m ²	$H=\{1\}$	0.482 m ²
$H=\{0\}$	0.734 m ²	$H=\{0\}$	3.84 m ²

5.2.3 Site 2 results

The key differences between the Melody Lane and Elm Grove sites are that the Elm Grove site has a larger expanse of trees and is located farther out in range in the SAR image, which implies larger σ_h . This results in larger MSE values for the Elm Grove site. The MSE improvement from multiscale estimation is more pronounced in the Elm Grove site because the larger height uncertainties require more smoothing. Also, incorporating LIDAR information in the MKS has a larger impact when there is more tall vegetation.

Fig. 5.6 shows data from a row where the LIDAR data were retained. Analogous data from a row where the LIDAR data were withheld are shown in Fig. 5.7 for the Melody Lane site. LIDAR data are shown for comparison purposes, but were not used in the estimation of these pixels. The \hat{z}_g and $\Delta\hat{z}_v$ estimates still approximately correspond to the LIDAR data because information from nearby LIDAR affects the estimates through the merging of pixels in the Kalman filtering step. Ta-

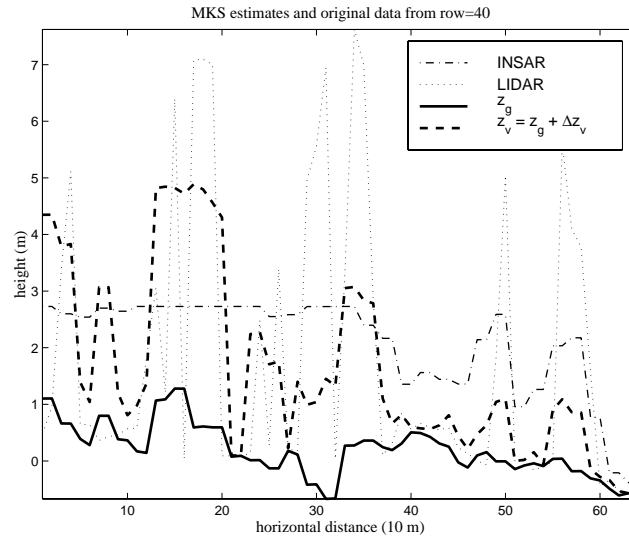


Figure 5.7: Comparison of data fusion results and original data for Melody Lane site. Data extracted from row 40 in Melody Lane output.

Table 5.3 gives the MSE results obtained without any multiscale estimation. Table 5.4 shows the estimates and estimate errors after running the MKS algorithm. After the MKS data fusion, the percent reduction in global MSE is 88% for the bare surface elevations and 66% for vegetation heights.

Table 5.3: Mean squared error for the transformed INSAR data for the Melody Lane test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	1.37 m ²	$H=\{0 \text{ or } 1\}$	4.79 m ²
$H=\{1\}$	1.20 m ²	$H=\{1\}$	4.93 m ²
$H=\{0\}$	1.43 m ²	$H=\{0\}$	4.75 m ²

Table 5.4: Mean squared error after MKS for the Melody Lane test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	0.167 m ²	$H=\{0 \text{ or } 1\}$	1.61 m ²
$H=\{1\}$	0.040 m ²	$H=\{1\}$	0.997 m ²
$H=\{0\}$	0.210 m ²	$H=\{0\}$	1.81 m ²

Table 5.5: Mean squared error after MKS for the Melody Lane test site using only INSAR data. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{1\}$	0.262 m ²	$H=\{1\}$	2.73 m ²

5.2.4 INSAR only case

The utility of the multiscale estimation was also investigated as an optimal smoother for a single data type. Table 5.5 shows the global estimated errors from the MKS algorithm when no LIDAR data are used for the Melody Lane site. The global MSE using LIDAR (Table 5.4) is 57% and 70% less than for the INSAR-only case for bare surface and vegetation, respectively. Notice, however, that the INSAR-only case is itself better than simply inverting the INSAR scattering model (Table 5.3) by 85% and 43% for bare surface and vegetation, respectively. We conclude that combining physical modeling with multiscale Kalman smoothing is useful even when operating on a single data type. This result is not an arbitrary smoothing of the INSAR data. Recall that the Kalman-based smoothing is affected by sensor characteristics, such as the measurement error.

5.3 Physical modeling and AMKS

In this section, INSAR and LIDAR data are transformed through the physical modeling and fused, as in Section 5.2. The standard and adaptive MKS algorithms are both applied to the data, and the results are compared. Standard multiscale estimation may be sufficient for small areas where the terrain and land cover are fairly simple, but adaptive estimation has the potential to handle larger areas with more spatial complexity.

5.3.1 Data description: Texas riparian environment

Figure 5.8 shows a polarimetric SAR image (C-band) of the central portion of Austin, Texas. The image was acquired by the NASA/JPL TOPSAR system, with a resolution of 10 m. The Riverside Park area is used as a test site. The Riverside Park site contains flat grassy areas and natural tree-covered areas as well. It is one of the largest contiguous areas in the city that is virtually free of buildings. Some of the grassy areas contain tall light poles and small buildings associated with recreational baseball fields, but areas that were free of obstructions were available. Figure 5.9 shows a photograph of the Riverside Park test site.

5.3.2 MKS results

The MKS algorithm was applied to the fusion of INSAR and LIDAR imagery acquired over the Riverside test site. In this case, the available LIDAR data were gridded to a very high spatial resolution of 1.25 m. The INSAR imagery has a pixel spacing of 10 m after ground range projecting, placing it three levels above the base of the quadtree data structure. As with the Texas coastal data, the INSAR data are from the NASA/JPL TOPSAR sensor, and the LIDAR data are from the Optech

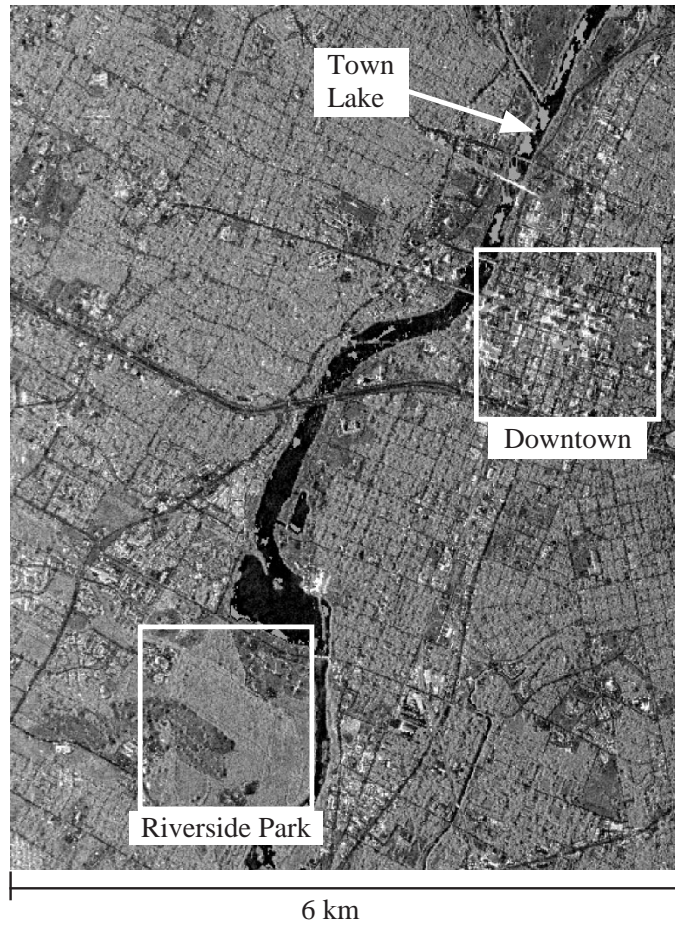


Figure 5.8: Polarimetric SAR image (C-band) of Austin, Texas. A box indicates the location of the Riverside Park test site.



Figure 5.9: Photograph of trees in Riverside Park in Austin, Texas. The trees generally ranged from 7 m to 15 m.

ALTM sensor. The LIDAR selection is different for this site than for the Bolivar sites. A single diagonal swath running from the lower left to the upper right of the imaged area represents a contiguous LIDAR acquisition over a region of interest, instead of the sparse grid pattern used for the Bolivar test sites.

Following the steps applied to the Bolivar test site, the INSAR scattering model is inverted to provide estimates of ground elevations z_g and vegetation heights Δz_v . These quantities are also derived from the LIDAR data using the empirical modeling. These data then serve as the inputs to the MKS algorithm.

The estimation of z_g are shown Fig. 5.10. In the image of INSAR ground elevations, the z_g estimates from the INSAR inversion are shown. The values range from roughly -2 m to 2 m. The area is quite flat, but a slight raised area is evident in upper right portion of the image. The LIDAR ground elevation image shows a swath of LIDAR data. The LIDAR image has very little visible noise. A slight raised area is again visible in the upper right portion of the image. This raised area in the INSAR estimates also corresponds to the location of trees and reveals a small bias in the INSAR inversion. However, it also corresponds to a change in surface heights, as indicated in the LIDAR image. This small local bias is most likely due to the non-zero surface slopes in this portion of the image. The scattering model used in this work does not account for non-zero surface slopes.

The fused estimates of z_g reflect the high resolution of the LIDAR in the diagonal swath. Outside that swath, estimates are coarse, but smoother than the INSAR derived elevations. The estimate uncertainty varies with the signal, but the dominant effect in this case is the relative measurement error of the sensors. Thus the uncertainty image is dominated by the swath of small values corresponding to the LIDAR data. This effect is compounded by the fact that the INSAR data is located three levels above the LIDAR data in the quadtree. Table 5.6 provides MSE results using the original data. As in Section 5.2, the coarse data is used because it

will generally have complete and dense coverage over the region of interest.

Results for the estimation of Δz_v are shown Fig. 5.11. In the image of INSAR derived vegetation heights, the heights range from 0 m to about 7 m. The bright area in the upper right portion of the image is a tree-covered area. An area of smaller trees is present along the left side of the image. The remaining portions of the image are grassy areas. The LIDAR derived vegetation heights reveal the same pattern in detail. Notice that the highest values in the LIDAR image are about 9 m, indicating that the INSAR generally underestimates the tree heights due to penetration into the canopy. The fused estimates of Δz_v preserve the detail of the LIDAR data. The estimated tree heights just outside the swath of LIDAR are generally between the LIDAR and INSAR derived tree heights.

The effect of the LIDAR data on the final estimates decreases with distance from the LIDAR data. The rate at which it decreases depends on the spatial arrangement of the LIDAR data, the relative resolutions of the LIDAR and INSAR data, and the relative measurement uncertainties of the data. However, as shown in Section 3.5.2 for the 1-D case, the estimate uncertainty will remain bounded. As in Fig. 5.10, the estimate uncertainty is dominated by the sensor measurement errors, which cause it to appear as a diagonal swath with only slight target-dependent variations visible.

Table 5.7 provides the MSE results for MKS filtered data. notice that the MSE is smaller than that of the inversion results, even where LIDAR data are absent ($H=0$). This happens for two reasons. There is some sharpening of the INSAR estimates close to the pixels containing LIDAR data. This is a result of the merging of estimates in the upward sweep of the MKS algorithm. However, the primary reason that the MSE improves for $H=0$ after applying MKS is the smoothing of the noisy INSAR estimates by the Kalman filter.

The MSE results for z_g and Δz_v for $H=0$ in Table 5.7 are larger than those

Table 5.6: Mean squared error of the transformed INSAR data for the Riverside Park test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	1.29 m ²	$H=\{0 \text{ or } 1\}$	9.33 m ²
$H=\{1\}$	1.42 m ²	$H=\{1\}$	8.03 m ²
$H=\{0\}$	1.19 m ²	$H=\{0\}$	10.4 m ²

Table 5.7: Mean squared error after MKS for the Riverside test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	0.217 m ²	$H=\{0 \text{ or } 1\}$	5.56 m ²
$H=\{1\}$	1.00e ⁻⁵ m ²	$H=\{1\}$	5.78e ⁻³ m ²
$H=\{0\}$	0.388 m ²	$H=\{0\}$	9.94 m ²

in Table 5.4 even though the vegetation heights are roughly the same in the two data sets. This is partly due to the different extents of vegetation in the two scenes. The primary cause, however, is that the INSAR data reside at one scale above the LIDAR in the Bolivar data ($m = M - 1$) and at three scales above the LIDAR in the Riverside data ($m = M - 3$). The INSAR data are projected down three scales in the Riverside data, compared to only one scale in the Bolivar data. More process noise is therefore injected into the INSAR estimates in the Riverside data because process noise is added to the estimates each time they are projected down a level in the quadtree.

5.3.3 AMKS Results

The AMKS algorithm was also applied to the Riverside test site. AMKS results for the estimation of bare-surface elevations z_g are shown Fig. 5.12. This figure appears very similar to the MKS result in Fig. 5.10. Looking at the INSAR derived

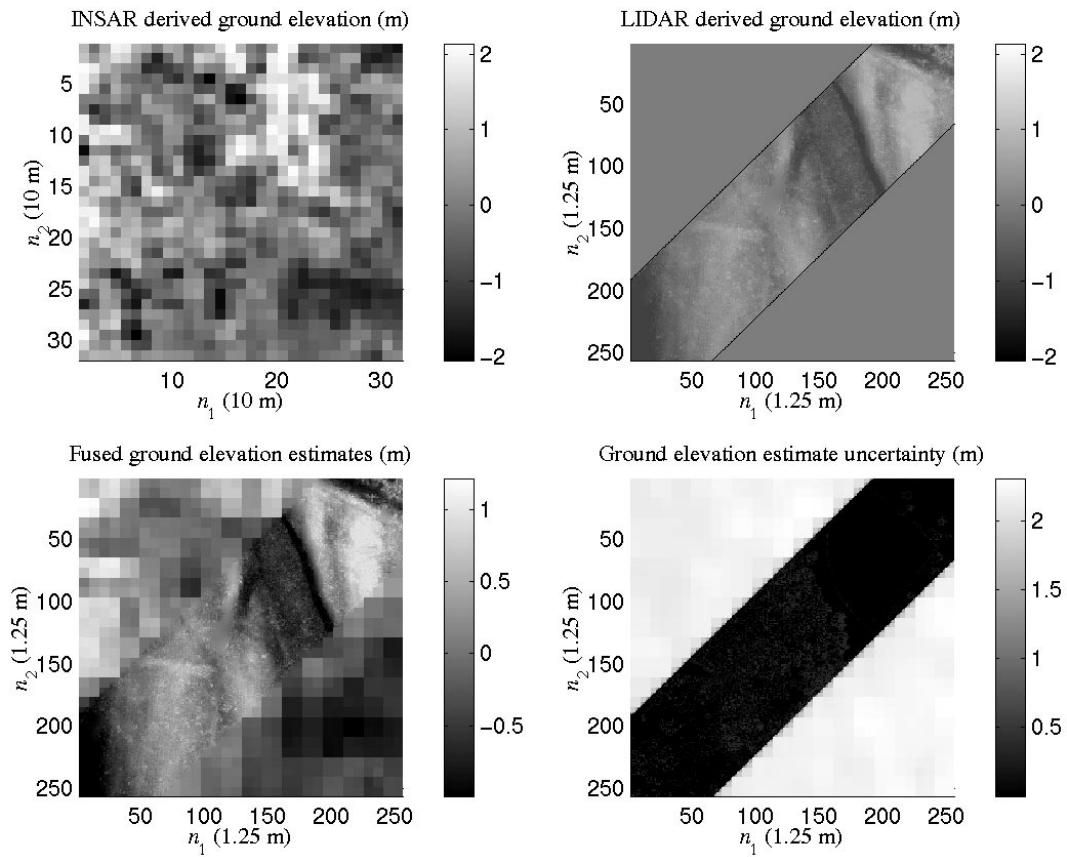


Figure 5.10: MKS output for Riverside site, z_g . The region of support for the LIDAR data is a diagonal swath running from the lower left corner to the upper right corner of the LIDAR image.

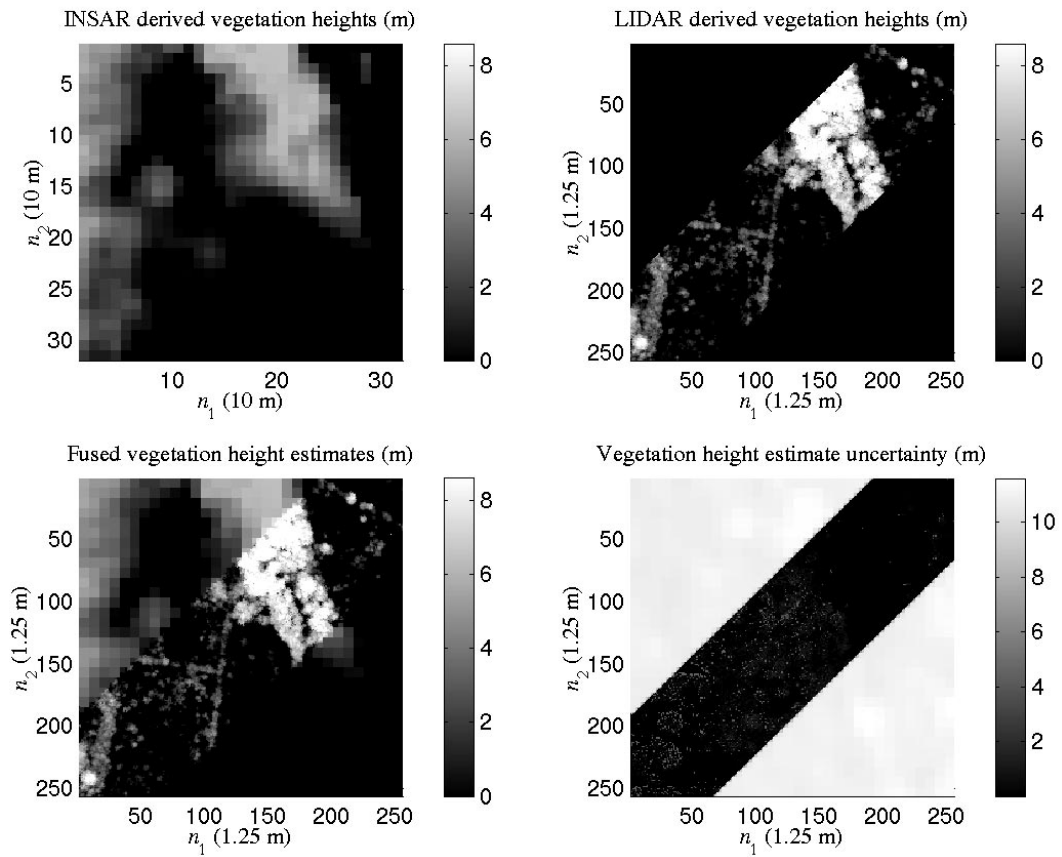


Figure 5.11: MKS output for Riverside site, Δz_v . The region of support for the LIDAR data is a diagonal swath running from the lower left corner to the upper right corner of the LIDAR image.

estimates of z_g , they exhibit greater variability in the upper left half of the image and decreased variability in the lower right half of the image.

The locally updated estimates of process noise variance Q at each scale are shown in Fig. 5.13. Initially, the Q arrays are spatially uniform at the base of the quadtree ($m = 9$), just as in the MKS case. In fact, the initial value for Q is exactly the same as it was in the MKS case. At the scale where INSAR observations are present, values of Q are updated for groups of N_b pixels in the rows and columns following the any detection of non-white innovations. The Q arrays remain uniform until the scale above the INSAR observations ($m = 5$). At that level, information from the spatial filter is made available to the scale-wise filter.

In Fig. 5.13 it is apparent that where Q was updated, it was generally revised downward from its MKS value. This indicates that the MKS estimate for Q was too large. Q is not updated in the first N_b rows and columns because of the latency of N_b pixels in the adaptive estimation algorithm. Thus, the upper left block of N_b pixels at $m = 5$ remains unchanged, where $N_b = 4$. Notice that in the lower right half of the image at $m = 5$ the locations of the downward revisions in Q correspond to the area of decreased variability (lower right half) in the INSAR image in Fig. 5.12.

Results for the estimation of vegetation heights Δz_v are shown Fig. 5.14. This figure appears very similar to the MKS result in Fig. 5.11. The INSAR derived estimates of Δz_v exhibit areas of trees (bright areas) in the upper right portion and along the left side of the image. In between these trees is a grassy area with lower spatial variability.

The locally updated estimates of process noise variance at each scale are shown in Fig. 5.15. Again, we see that Q is locally revised downward starting at scale $m = 5$, indicating that the MKS Q was too large. The region where Q is revised downward corresponds to the band of short grass adjacent to the band of trees visible in Fig. 5.14.

Table 5.8: Mean squared error after AMKS for the Riverside test site. H evaluated at $m = M$.

bare surface elevation		vegetation height	
$H=\{0 \text{ or } 1\}$	0.216 m ²	$H=\{0 \text{ or } 1\}$	5.55 m ²
$H=\{1\}$	1.00e ⁻⁵ m ²	$H=\{1\}$	5.78e ⁻³ m ²
$H=\{0\}$	0.386 m ²	$H=\{0\}$	9.92 m ²

Table 5.8 provides the MSE results for AMKS filtered data. We see that AMKS does improve MSE, but only slightly. A 0.5% reduction in global MSE for z_g and a 0.2% reduction in global MSE for Δz_v . This small reduction in MSE for relatively small model errors is consistent with the simulation results from Section 4.5.2. A contributing factor is the fact the spatial filter must operate on a dense data set to obtain meaningful ACF estimates used to test for optimality. Since the dense data are usually only present at coarse resolutions, the spatial filter is forced to operate on low-resolution data. Thus, model errors that are detectable in the high-resolution data, may not be detected in the coarse data. The magnitudes of non-stationarities may also be reduced as the resolution decreases, leading to smaller changes in the updated process noise variance. This is particularly true in the Riverside Park data, where the scale of the INSAR data ($m = M - 3$) is eight times coarser than the LIDAR data at the base ($m = M$).

As was stated for the 2-D simulation results in Chapter 4, reduction of MSE is not the sole motivation for implementing an adaptive data fusion algorithm. For example, by studying the updated Q images at each scale, it is possible to discern how the target variability changes with location. In both Fig. 5.13 and Fig. 5.15, Q was revised lower where there ground was smoothest and the vegetation was shortest. This provides insight into the nature of the terrain and vegetation, which can be used to guide subsequent algorithm development and interpretation.

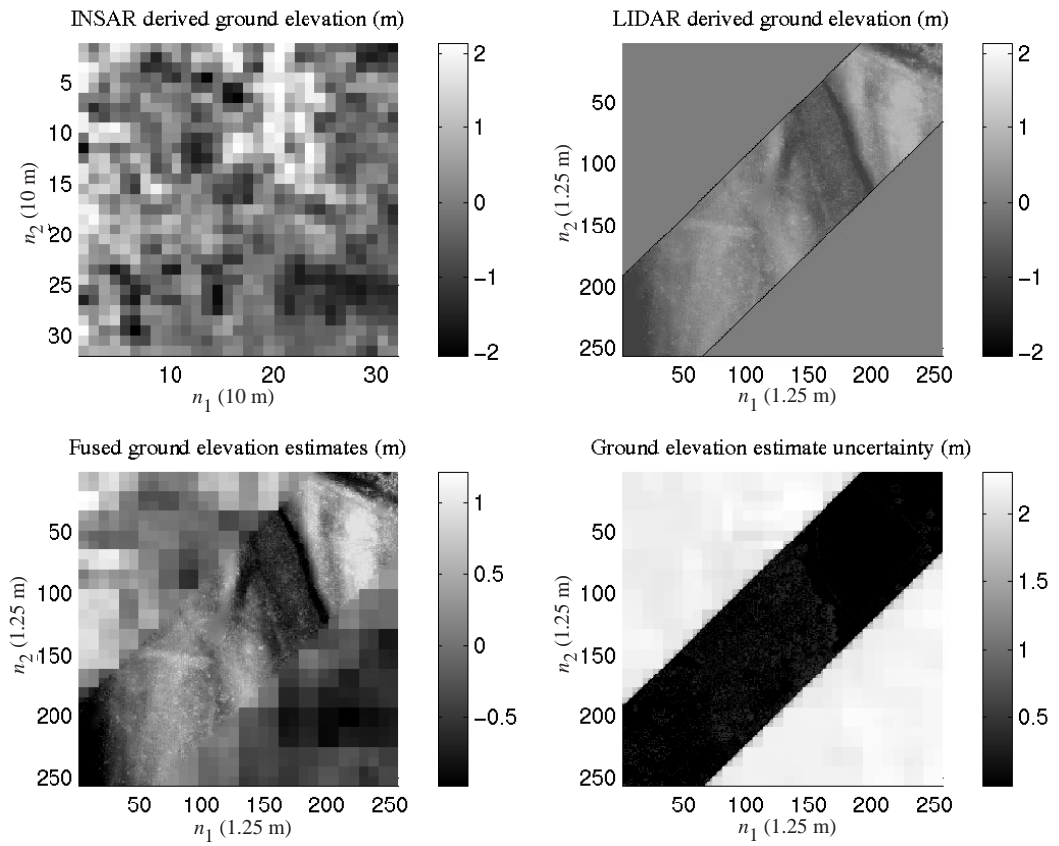


Figure 5.12: AMKS output for Riverside site, z_g .

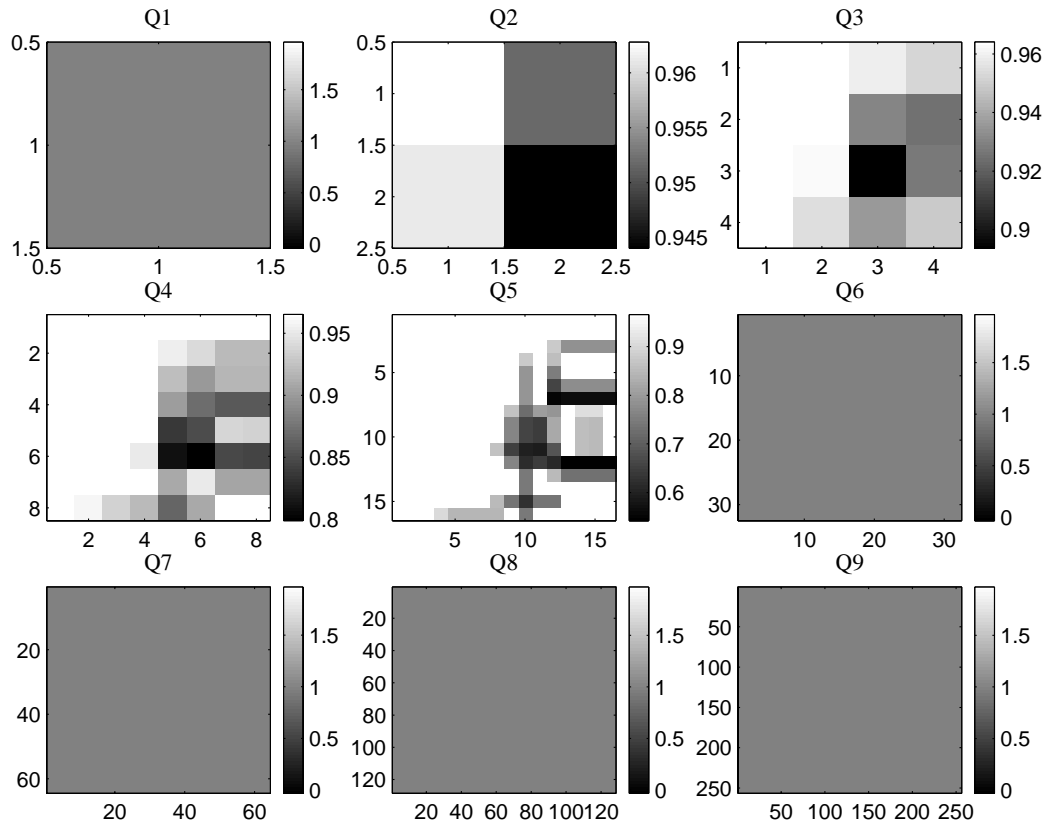


Figure 5.13: Fine-to-coarse process noise variances $Q(m, i, j)$ at the Riverside from the AMKS algorithm, z_g . Q is revised downward in lower right half of the Q_5 image, corresponding to smoother terrain in that portion of the INSAR image in Fig. 5.12.

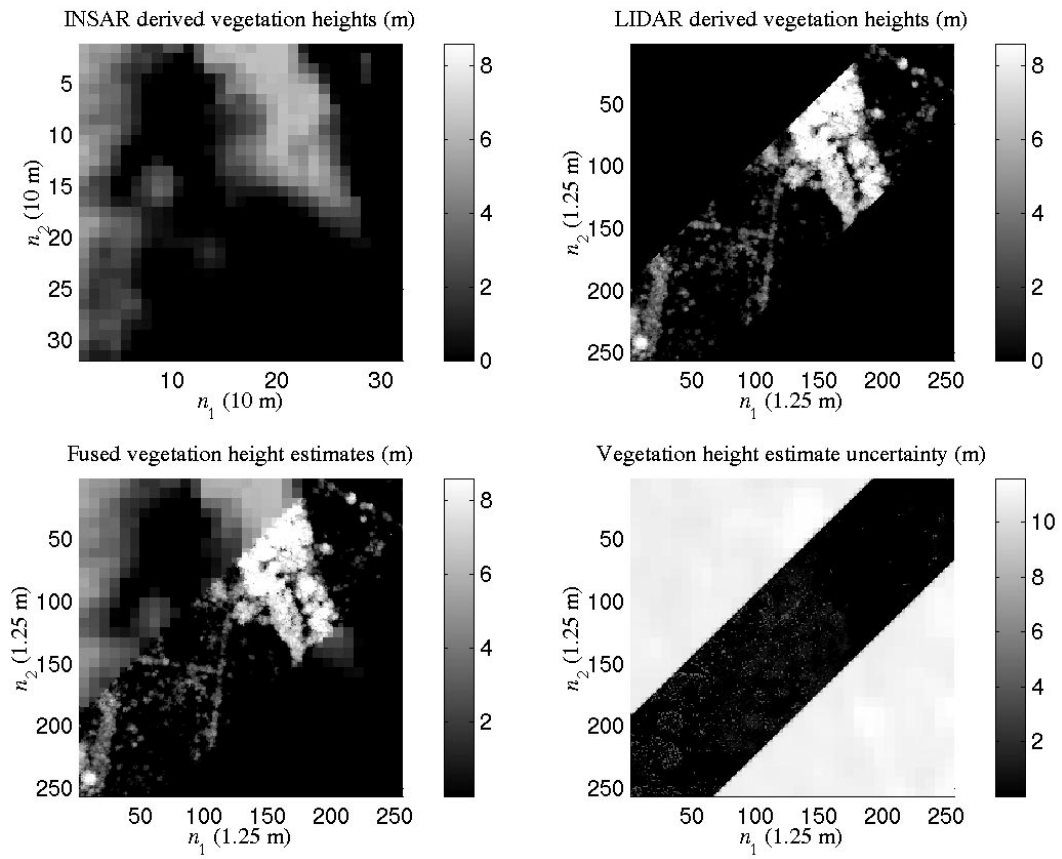


Figure 5.14: AMKS output for Riverside site, Δz_v .

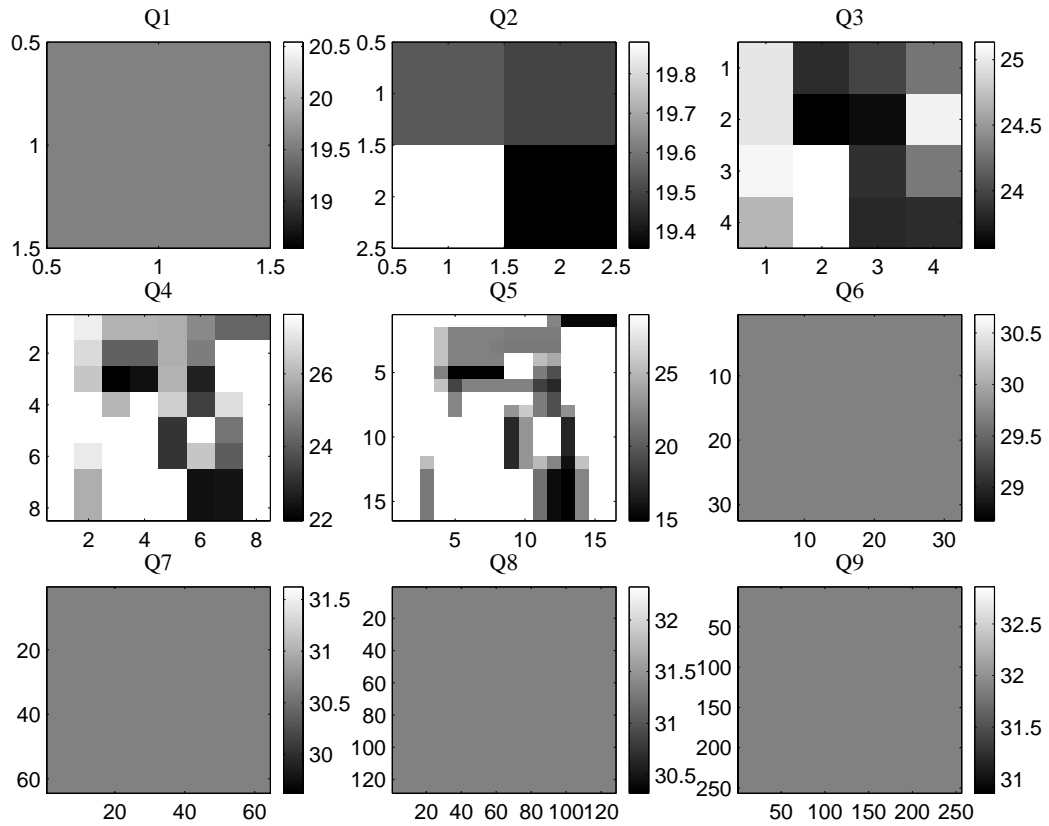


Figure 5.15: Fine-to-coarse process noise variances $Q(m, i, j)$ at the Riverside from the AMKS algorithm, Δz_v . Q is revised downward along a band running from the upper left to the lower right of the $Q5$ image, corresponding to the smooth grassy area in that portion of the INSAR image in Fig. 5.14.

5.4 Fusing three data sets with MKS

Finally, we investigate the fusion of three data types to show that the method is not restricted to fusing only two data sets. In this case, coarse-resolution spaceborne INSAR data from the ERS satellites are combined with higher resolution TOPSAR acquisitions. The ERS data consist of one INSAR image pair, thus only one baseline. Therefore, it is not possible to transform the ERS data into estimates of z_g and Δz_v using the approach described in Chapter 2. So all three INSAR data sets represent the nominal INSAR heights z_S . Given that z_S data are used in this case, a test site over a semi-arid region with minimal vegetation cover was chosen. Because there is little vegetation, we can assume $z_S \approx z_g$. Figure 5.16 shows a SAR backscattered power image of the area, acquired by the ERS-1 satellite at C-band. The MKS algorithm is applied in the boxed region in the upper left portion of the image. That is an interesting portion of the image because there is a rapid change in topography from a plateau down to the Missionary Plain. River gorges associated with the Finke River and small tributaries are present there.

5.4.1 Data description: Australian semi-arid environment

The Finke River gorge is located in central Australia. This ancient river system flows between the MacDonnell and James mountain ranges, crossing a sandstone plain. The area is not heavily vegetated except some isolate areas near river gorges. Fig. 5.17 shows the coverage of the TOPSAR in relation to the coverage of the ERS data. A TOPSAR line acquired in 1996 covers the same area as a line acquired in 2000, but they were acquired from opposite directions. Because of the steep topography in the area, there are areas on the back sides of hills where shadowing prevents SAR acquisition. Shadowing is a common phenomenon in SAR images. Fig. 5.18 shows detailed views of the 1996 and 2000 TOPSAR acquisitions. Notice

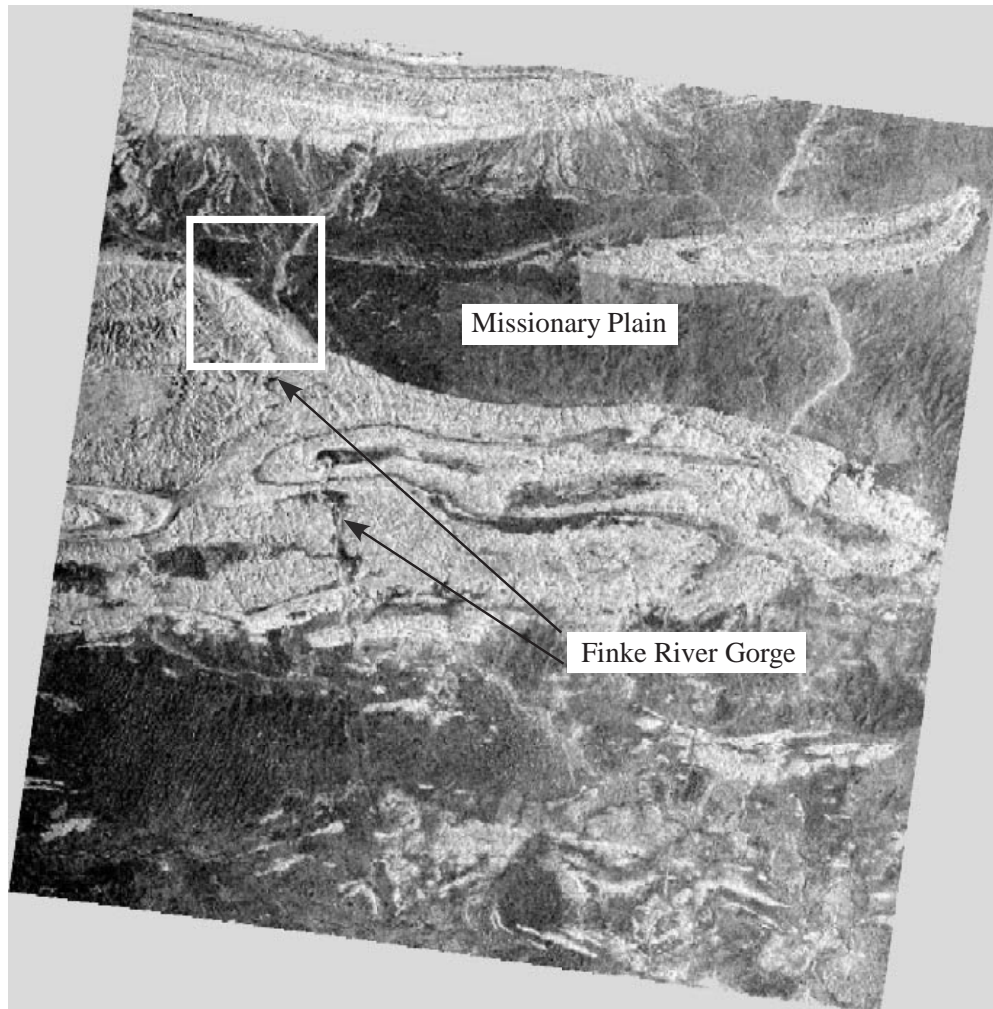


Figure 5.16: ERS-1 topographic image of the Finke River Gorge in Australia. The MKS algorithm is applied in the boxed region in the upper left portion of the image. The imaged area is approximately $50 \text{ km} \times 50 \text{ km}$.

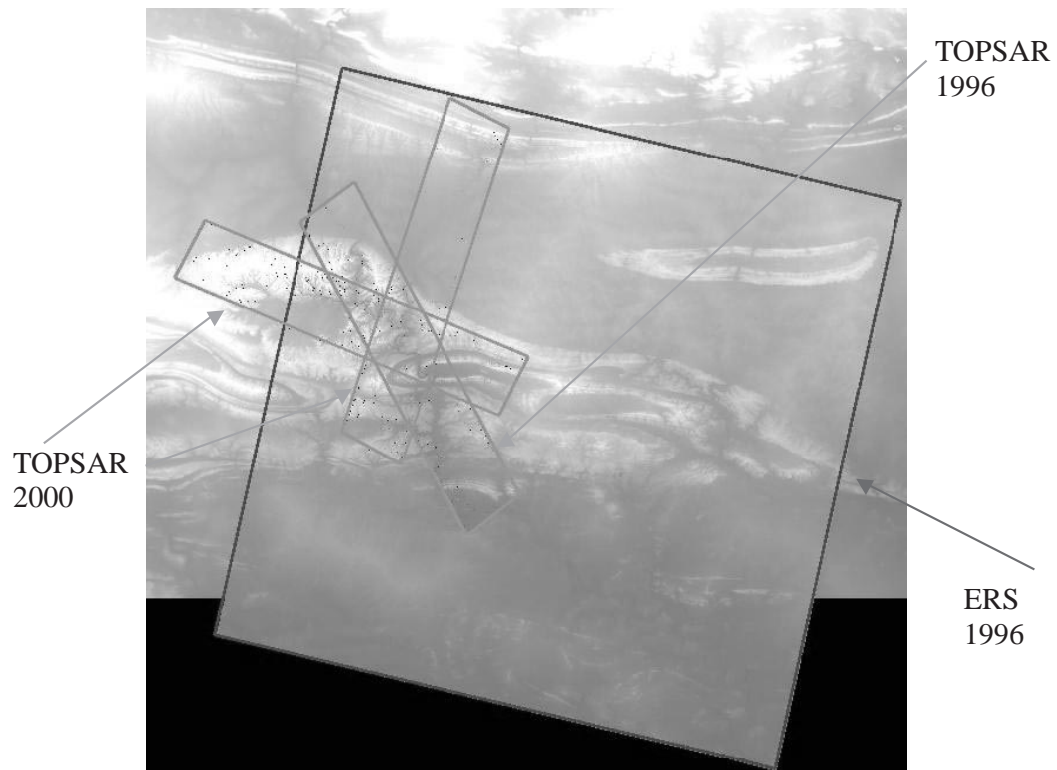
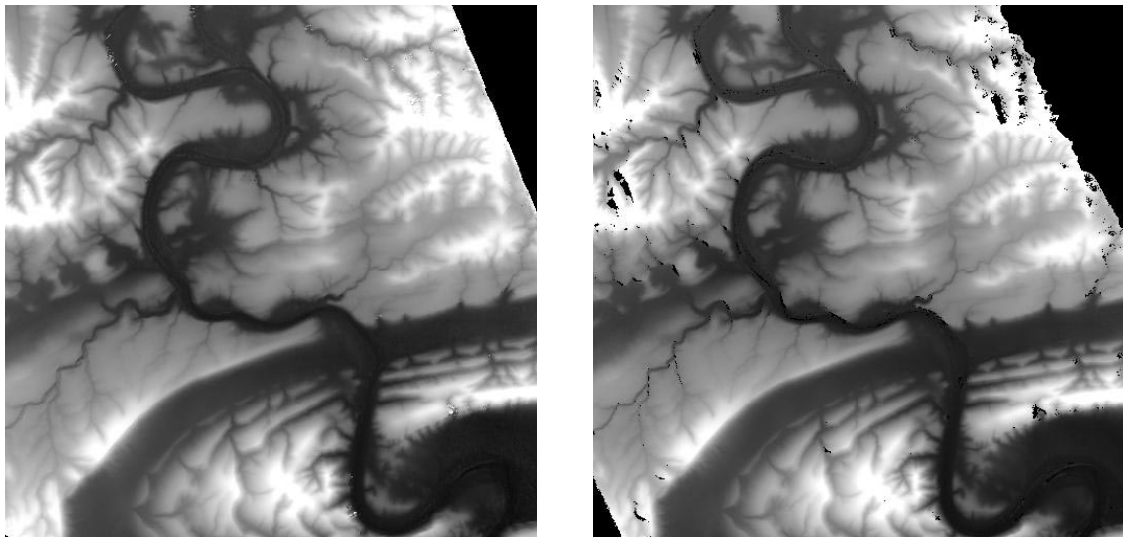


Figure 5.17: Relative coverage of the ERS and TOPSAR acquisitions.

that the shadow regions are manifested as data dropouts and appear in different locations in the two images due to differences in the look directions. The data fusion fills in these data dropouts during the coarse-to-fine Kalman smoothing step.

A pair of SAR images was acquired by the ERS-1 and ERS-2 satellites while operating in the one-day tandem phase of their missions over the test site in 1996. The data were processed to a spatial resolution of 20 m. Within the nearly square area imaged by the ERS satellites, the TOPSAR sensor acquired long swaths of data in 1996 and in 2000. The 1996 acquisition was processed to a resolution of 10 m, using a relatively high number of multilooks to reduce noise. The 2000 acquisition



Finke River 1996 ↗

Finke River 2000 ↙

Figure 5.18: Detailed views of the TOPSAR acquisitions of the Finke River. The arrows indicate the SAR look direction. Topography induced shadowing leads to data drop outs, which occur in different locations in the two images because of the different look directions. The displayed area is approximately $10 \text{ km} \times 10 \text{ km}$.

imaged the area from the opposite side and was processed to a resolution of 5 m, using fewer multilooks.

5.4.2 Results

The MKS estimates from the three fused INSAR images are shown in Fig. 5.19. Portions of the two TOPSAR images were omitted to simulate the limited coverage of airborne INSAR relative to the full ERS scene. In Fig. 5.19, the 10 m TOPSAR data are restricted to a horizontal swath that includes the gorge. The 5 m TOPSAR data are restricted to a smaller swath covering the central part of the gorge. ERS data cover the entire image.

The image of estimate uncertainty shows that the coarsest data yield the highest uncertainty. The 10 m TOPSAR data yields intermediate values of uncertainty, and the 5 m TOPSAR yields the lowest uncertainty. It is interesting to note that these patterns in the estimate uncertainty derive from both the actual INSAR height error σ_h and the resolutions of the different data sets. Recall from (2.7) that σ_h is a characteristic of the processed INSAR data. The 10 m TOPSAR data have the smallest σ_h , but they yield higher estimate uncertainty values than the 5 m TOPSAR data since they reside at a coarser scale in the quadtree. The smaller σ_h of the 10 m TOPSAR data is due to the large number of multilooks used in the processing of the data. Eighteen looks were used in the 10 m data versus only nine in the 5 m data.

In Fig. 5.20, a transect through the fused estimates clearly shows how the estimate \hat{x}^s tracks the best data available, where best refers to minimum variance. Where only ERS data are present, the estimates track them, but deviations are apparent because of the relatively high measurement error of the ERS data. Where the 5 m TOPSAR data are available, the estimates track the observations very closely.

Fig. 5.21 shows a perspective view of the fused estimates in Fig. 5.19. The higher resolution of the 5 m TOPSAR data is apparent in the central gorge, relative to the terrain on either side of the gorge.

Examination of Fig. 5.20 reveals that the filter smooths the ERS data noticeably where only ERS data are available. When the Kalman model parameters are correct, this smoothing represents the minimum variance estimate of the signal in the presence of the measurement noise R .

Tests were performed on the Finke River data set where the value of the process noise variance Q used in the Kalman filter was too small, approximately correct, and too large. Q was determined by comparing the power spectra of the data model and the observations.

When Q used in the filter is too small, the filter over smooths, resulting in a loss of information about the true signal value. When Q is too large, the filter erroneously considers the data to be more reliable. In such a case, the final estimates will track the coarse data very closely in areas where only coarse data are available. However, that does not represent greater retention of information about the *true signal*. Rather, it represents greater retention of the observed signal and fails to account for the full uncertainty in those observations. In the case of IN-SAR observations, that uncertainty is calculated from the $|NICC|$ data, and it can be significantly greater than the apparent spatial variability of the $\angle NICC$ data indicates.

In the optimal case, it is possible for regions of the fused estimate image where only coarse data are available to appear blocky. This Kalman filter estimate represents the best estimate of the true signal relative to a global MSE criterion and the available stochastic data model. The final estimates could display a different characteristic behavior if additional information is made available to refine the data model. For example, if a surface is known to be very smooth, e.g. large-scale

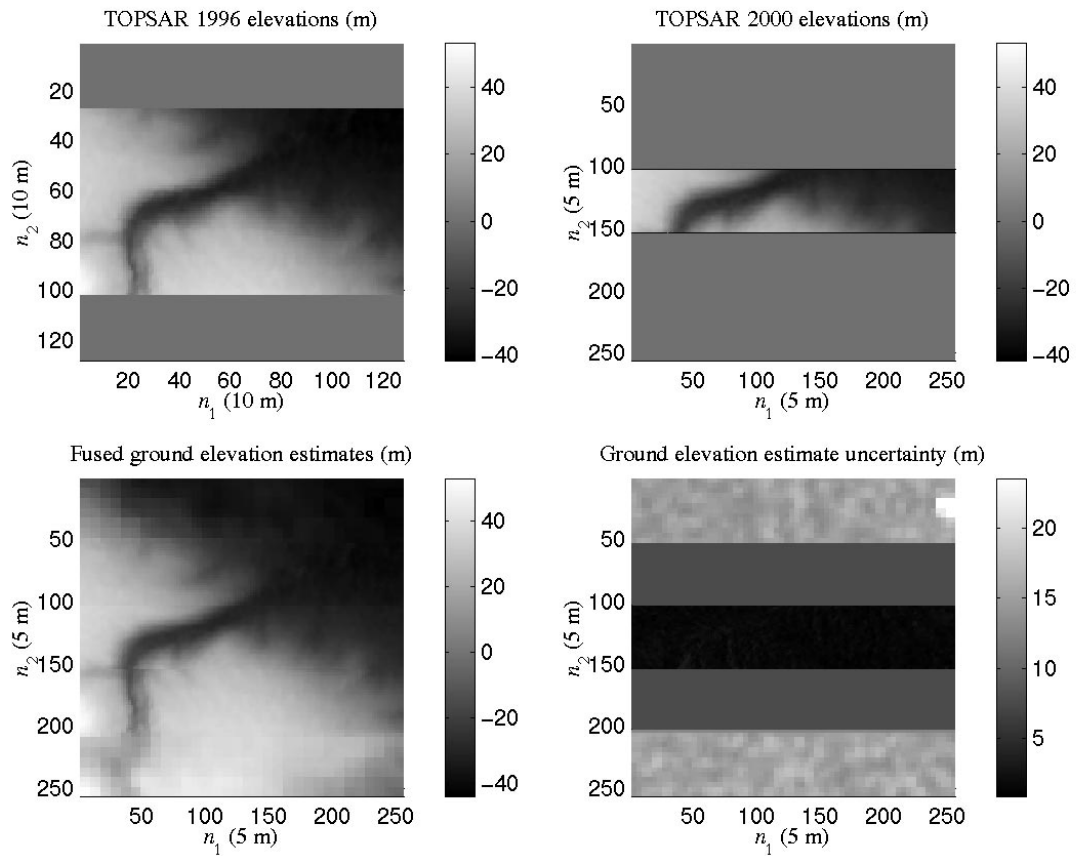


Figure 5.19: MKS output for Finke site, z_g . 20 m ERS data cover the entire imaged area. 10 m TOPSAR data (upper left image) cover most of the imaged area, including the gorge. 5 m TOPSAR data (upper right image) cover only a thin swath, including part of the gorge.

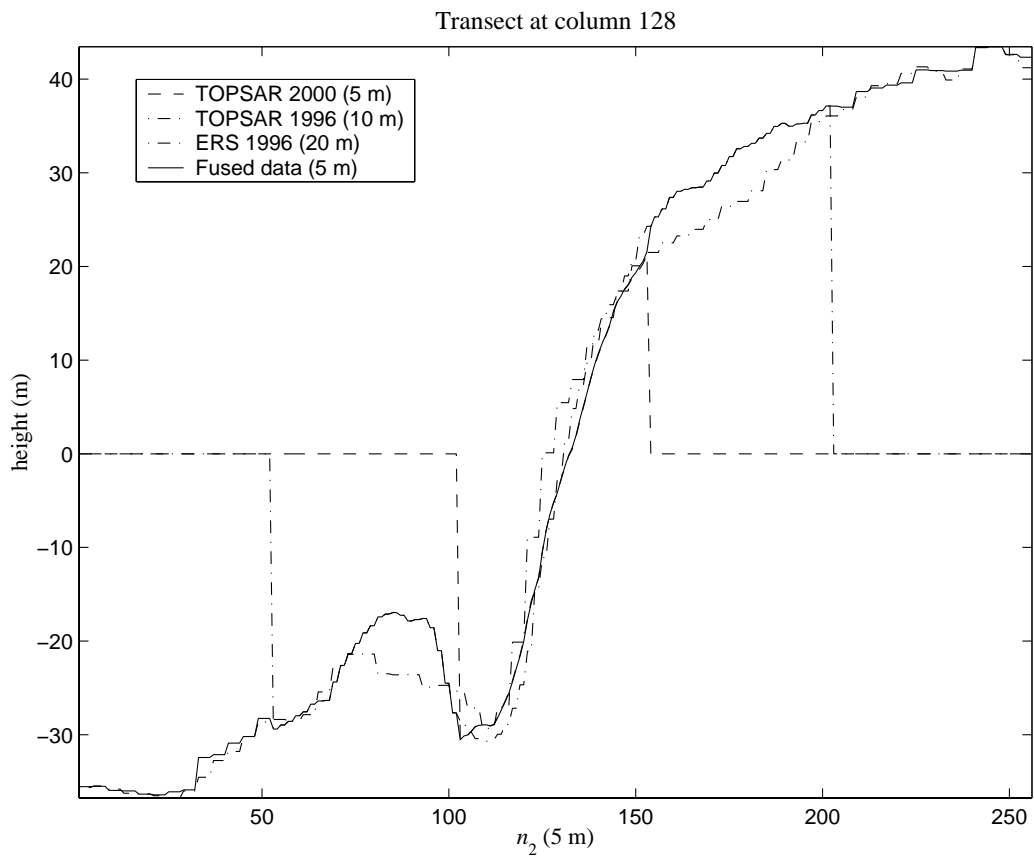


Figure 5.20: Transects for the Finke site, z_g . The transects come from column 128 of the fused estimate image in Fig. 5.19.

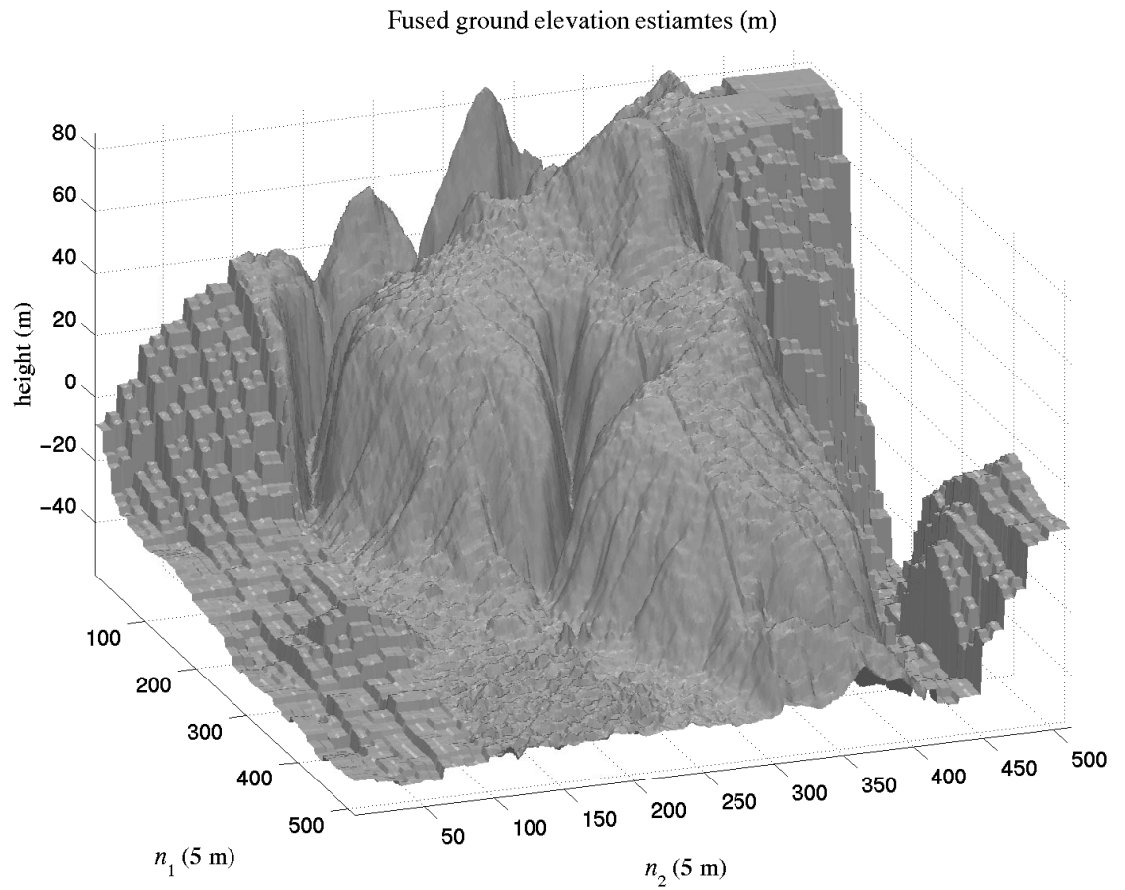


Figure 5.21: Perspective view of the fused estimates in Fig. 5.19. Values at $n_1=128$ correspond to the transects in Fig. 5.20.

ice sheets, the state vector in the Kalman filter could be augmented to estimate gradients between neighboring pixels. In such a case, the blocky appearance of the fused estimates could be reduced in a manner that was consistent with the physical process being observed.

5.5 Conclusions

In this chapter, three distinctly different cases were analyzed. First, physical modeling was combined with the standard MKS algorithm to allow the highly nonlinear problem of bare-surface and vegetation height estimation to be treated with the Kalman-based data fusion technique. Second, adaptive and nonadaptive fusion results were presented and compared. It was shown that the adaptive estimation leads to slightly smaller global error than the non-adaptive estimation. It is worth noting that resolution of the data used to update the process noise variance will effect the degree to which the final estimates can be improved. The impact of the adaptive estimation would probably be greater if complete coverage was available at the finest scale. It was also demonstrated that the data fusion could be applied to more than two types of topographic data sets, thus expanding the class of topographic applications that can be addressed.

The data sets examined in this dissertation represent the typical availability of data in remote sensing of topography. A high-resolution data set is available, but is sparse or incomplete, and a coarse data set is available that is complete. It is also typical that the high-resolution data has a smaller measurement uncertainty. In this arrangement, the final fused estimates can resemble the high-resolution data simply having been spliced into the coarse data. The possibility of splicing the data sets might be regarded as simple alternative the Kalman-based fusion. However, splicing will not result in optimal estimates in general.

The Kalman-based fusion represents a significant improvement over a brute-force fusion method of simply splicing two data sets together. In particular, the splicing operation can be viewed as a special case of the Kalman-based fusion where measurement uncertainties near zero are used. A more realistic approach is to assign physically meaningful finite measurement errors to each sensor. This incorporates more information into the fusion process. In cases, where a data set is noisy or corrupted in some way, splicing it into the estimates is a poor approach and will yield suboptimal estimates. In addition, filtering upward on the quadtree involves merging of information. This makes it possible to sharpen the INSAR data in the vicinity of the LIDAR data. Finally, the Kalman-based approach provides extensive information about stochastic process in scale and space that may be used to guide subsequent acquisitions of data.

Chapter 6

Conclusions

I have shown that combining physical modeling with multiscale estimation can significantly improve estimates of bare surface topography and vegetation heights obtained from INSAR. The estimates can be further improved if LIDAR is combined with the INSAR data. Multiscale estimation is a natural approach for fusing data with different resolutions. By including physical modeling and target-dependent measurement variances, the approach is made general for proper fusion of dissimilar data types.

The multiscale method was also extended to allow spatially adaptive estimation. The adaptive algorithm yields better error characteristics than the non-adaptive multiscale filter and accommodates non-stationarities in the image data. Statistically optimal estimates of the state variables, in the mean squared sense, are obtained, conditioned on the physical modeling and observations. I applied the estimation framework to remotely sensed data over coastal and riparian environments in Texas and a semi-arid environment in Australia to demonstrate its ability to accommodate several different remote sensing scenarios.

6.1 Summary

Key contributions of this work include (1) combining physical modeling with multiscale estimation to accommodate nonlinear measurement-state relationships, (2) extending adaptive estimation theory to spatially-varying multiscale processes, and (3) improving estimates of ground elevations and vegetation heights for remote sensing applications. The approach used in this work of combining physical modeling with multiscale estimation to solve for physical parameters is not limited to any particular data type or model. Different scattering models or data types can be investigated and compared.

Many applications have been addressed with *data fusion*, but care is not always taken to combine the data in an appropriate manner. Sensors that operate at different resolutions, wavelengths, and use different acquisition geometries will not strictly measure the same process, even when the sensors are intended to measure the same phenomena. This is the case with INSAR and LIDAR sensors. Transforming the observations into a set of common parameters before combining them is an essential step. In this dissertation, I transformed both INSAR and LIDAR according to physically-based models prior to the data fusion. This allows the application of Kalman-based fusion to data with nonlinear measurement-state relations and no closed-form inverse.

The original MKS algorithm can be adaptive in *scale* if the coarse-to-fine process noise variance $\Gamma^2(m)$ varies with m . This occurs when the data exhibit a $1/f^\mu$ power spectrum and the data model is matched to that spectrum. However, no mechanism for making the algorithm adaptive in spatial dimensions has been reported. Such a mechanism was presented in this dissertation. The spatially adaptive algorithm remains optimal even when the observed process is non-stationary.

This research was motivated in large part by the topographic measurement

application. The goal was to develop a theoretically-based optimal method that produced globally better topographic estimates than those obtained without it. I have shown that in the INSAR smoothing case (one data type) the estimates are better than the original INSAR measurements. In the non-adaptive data fusion cases, the estimates are consistently better than those obtained by the coarse sensor alone and had greater coverage than the fine sensor. The adaptive AMKS algorithm provided improved estimates relative to the MKS algorithm. The degree of improvement was data dependent and was diminished because the adaptive filter was confined to coarse levels in the quadtree. However, it was also shown to be a useful tool to locate errors in the MKS model parameters and relate them to different terrain or land cover types.

6.2 Future work

Recommendations for future work include (1) making the spatial filter a 2-D reduced update Kalman filter rather than two separable 1-D filters, (2) investigating non-iterative methods to convert INSAR into z_g and Δz_v to avoid the need for two observations (dual-baseline or two passes) and to reduce computational complexity, and (3) applying multi-model (filter bank) approaches to the non-stationary problem instead of updating the parameters of a single model.

The AMKS algorithm developed in this dissertation assumed that the multidimensional Kalman filter could be considered separable. This made the incorporation of adaptive estimation techniques straight forward, but it ignores spatial correlation among neighboring pixels in front of and beside the current estimate. Since natural images often exhibit significant correlation in all directions, employing a Kalman filter with a 2-D region of support for the spatial filter would treat the spatial correlation more generally. Such filters do exist, and are generally termed

reduced update Kalman filters [48]. They are still causal filters, but the region of support can extend over portions of adjacent rows and columns to incorporate 2-D correlations. Implementing the adaptive estimation would be more involved if this type of spatial filter was used, but is mathematically tractable. The first step would be to derive a generalized version of the innovation–correlation approach that used a 2-D sequence of innovations.

The most computationally demanding step in the fusion process is the transforming of INSAR data into bare-surface and vegetation heights. The method used in this dissertation was to numerically invert an INSAR scattering model. Nonlinear constrained optimization techniques were required to accomplish the inversion, leading to an iterative solution at each pixel. As pointed out in Chapter 2, other methods exist for estimating surface and vegetation heights from INSAR data that do not require iterative methods. These methods are not based on formal scattering theory and are only approximate, but they should be investigated for their potential to reduce the overall computational burden. Specifically, the method of [14] could be extended to different terrain types.

Finally, other methods for adapting the Kalman model parameters should be investigated. The innovation–correlation method was used in this dissertation because it is theoretically based and has only moderate computational complexity. The other approaches listed in Chapter 4 should be investigated also. In particular, the multiple model approach should be investigated. The primary disadvantage of this method is that it assumes that the observed process will follow one of a small set of known and constant dynamic states, or possibly a combination of such dynamic states. Each state is represented by a set of Kalman model parameters, comprised of a bank of filters. If a non-stationarity is detected in the process, a new model can be substituted for the old one. In general, highly random processes like topography, will not be well characterized by a small set of discrete dynamical

states. However, if reasonably good approximations to archetype terrains, such as forest and grassland, are determined, the multiple model approach could lead to an efficient adaptive estimation algorithm because there would no longer be a need to solve a large set of linear equations to update the model. Auxiliary data from optical sensors could be used to delineate changes in terrain or land cover, and aid in the determination of spatially varying Q values.

Bibliography

- [1] R. N. Treuhaft and P. R. Siqueira, “Vertical structure of vegetated land surfaces from interferometric and polarimetric radar,” *Radio Science*, vol. 35, pp. 141–177, Jan. 2000.
- [2] H. A. Zebker, T. G. Farr, R. P. Salazar, and T. H. Dixon, “Mapping the world’s topography using radar interferometry: The TOPSAT mission,” *Proc. IEEE*, vol. 82, pp. 1774–1786, Dec. 1994.
- [3] K. C. Chou, A. S. Willsky, and A. Benveniste, “Multiscale recursive estimation, data fusion, and regularization,” *IEEE Trans. Automatic Control*, vol. 39, pp. 464–478, Mar. 1994.
- [4] P. W. Fieguth, W. C. Karl, A. S. Willsky, and C. Wunsch, “Multiresolution optimal interpolation and statistical analysis of TOPEX/POSEIDON satellite altimetry,” *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 280–292, Mar. 1995.
- [5] J. Curlander and R. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*. New York, NY: Wiley, 1991.
- [6] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, vol. 2. Norwood, MA: Artech House, 1986.

- [7] J. P. Fitch, *Synthetic Aperture Radar*. New York, NY: Springer-Verlag, 1988.
- [8] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Science*, vol. 23, pp. 713–720, July 1988.
- [9] F. Li and R. M. Goldstein, "Studies of multi-baseline spaceborne interferometric synthetic aperture radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 88–97, Jan. 1990.
- [10] S. N. Madsen, J. M. Martin, and H. A. Zebker, "Analysis and evaluation of the NASA/JPL TOPSAR across-track interferometric SAR system," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 383–391, Mar. 1995.
- [11] H. A. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 950–959, Sept. 1992.
- [12] E. Rodriguez and J. M. Martin, "Theory and design of interferometric synthetic aperture radars," *IEE Proc.-F*, vol. 139, pp. 147–159, Apr. 1992.
- [13] U. Wegmuller and C. Werner, "Retrieval of vegetation parameters with SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 18–24, Jan. 1997.
- [14] J. O. Hagberg, L. M. Ulander, and J. Askne, "Repeat-pass SAR interferometry over forested terrain," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 331–340, Mar. 1995.
- [15] R. Lanari, G. Fornaro, D. Riccio, M. Migliaccio, K. P. Papathanassiou, J. R. Moreira, M. Schwäbisch, L. Dutra, G. Puglisi, G. Franceschetti, and

- M. Coltelli, "Generation of digital elevation models by using SIR-C/X-SAR multifrequency two-pass interferometry: The Etna Case Study," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 1097–1114, Sept. 1996.
- [16] K. C. Slatton, M. M. Crawford, J. Gibeaut, and R. Gutierrez, "Modeling SAR backscattering response to coastal inundation," in *Proc. IEEE Int. Geosci. Remote Sensing Symp.*, vol. 5, (Hamburg, Germany), pp. 2446–2448, July 1999.
- [17] S. S. Saatchi, D. M. L. Vine, and R. H. Lang, "Microwave backscattering and emission model for grass canopies," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 177–186, Jan. 1994.
- [18] R. H. Lang and J. S. Sidhu, "Electromagnetic backscattering from a layer of vegetation: A discrete approach," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-21, pp. 62–71, Jan. 1983.
- [19] R. H. Lang, "Electromagnetic backscattering from a sparse distribution of lossy dielectric scatterers," *Radio Science*, vol. 16, pp. 15–30, Jan. 1981.
- [20] L. Tsang, M. C. Kubacsi, and J. A. Kong, "Radiative transfer theory for active remote sensing of a layer of small ellipsoidal scatterers," *Radio Science*, vol. 16, pp. 321–329, May 1981.
- [21] S. S. Saatchi and K. C. McDonald, "Coherent effects in microwave backscattering models for forest canopies," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1032–1044, July 1997.
- [22] R. N. Treuhaft, S. N. Madsen, M. Moghaddam, and J. J. van Zyl, "Vegetation characteristics and underlying topography from interferometric radar," *Radio Science*, vol. 31, pp. 1449–1485, Nov. 1996.

- [23] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York, NY: Wiley, 2nd ed., 1993.
- [24] K. Sarabandi and Y. Lin, "Simulation of interferometric SAR response for characterizing the scattering phase center statistics of forest canopies," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 115–125, Jan. 2000.
- [25] A. Neuenschwander, M. Crawford, C. Weed, and R. Gutierrez, "Extraction of digital elevation models for airborne laser terrain mapping data," in *Proc. IEEE Int. Geosci. Remote Sensing Symp.*, vol. 5, (Honolulu, HI), pp. 2305–2307, July 2000.
- [26] K. C. Slatton, M. M. Crawford, and B. L. Evans, "Combining interferometric radar and laser altimeter data to improve estimates of topography," in *Proc. IEEE Int. Geosci. Remote Sensing Symp.*, vol. 3, (Honolulu, HI), pp. 960–962, July 2000.
- [27] E. W. Kamen and J. K. Su, *Introduction to Optimal Estimation*. London, UK: Springer-Verlag, 1999.
- [28] D. Just and R. Bamler, "Phase statistics of interferograms with applications to synthetic aperture radar," *Applied Optics*, vol. 33, no. 20, pp. 4361–4368, 1994.
- [29] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*. Upper Saddle River, NJ: Prentice Hall, 4th ed., 1998.
- [30] M. Daniel and A. Willsky, "A multiresolution methodology for signal-level fusion and data assimilation with applications to remote sensing," *Proc. IEEE*, vol. 85, pp. 164–180, Jan. 1997.

- [31] J. L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis: the Multiscale Approach*. Cambridge, UK: Cambridge University Press, 1998.
- [32] M. K. Schneider, P. W. Fieguth, W. C. Karl, and A. S. Willsky, "Multiscale methods for the segmentation and reconstruction of signals and images," *IEEE Trans. Image Proc.*, vol. 9, pp. 456–468, Mar. 2000.
- [33] J. Laferte, P. Perez, and F. Heitz, "Discrete Markov image modeling and inference on the quadtree," *IEEE Trans. Image Proc.*, vol. 9, no. 3, pp. 390–404, 2000.
- [34] R. Brown and P. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*. New York, NY: Wiley, 3rd ed., 1997.
- [35] M. Daniel and A. Willsky, "The modeling and estimation of statistically self-similar processes in a multiresolution framework," *IEEE Trans. Information Theory*, vol. 45, pp. 955–970, Apr. 1999.
- [36] D. Evans, T. Farr, and J. van Zyl, "Estimates of surface roughness derived from synthetic aperture radar (SAR) data," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no. 2, pp. 382–389, 1992.
- [37] K. C. Chou, *A Stochastic Modeling Approach to Multiscale Signal Processing*. Cambridge, Massachusetts: Ph.D. Dissertation, Massachusetts Institute of Technology, 1991.
- [38] R. K. Mehra, "Approaches to adaptive filtering," *IEEE Trans. Automatic Control*, pp. 693–698, Oct. 1972.
- [39] D. T. Magill, "Optimal adaptive estimation of sampled stochastic processes," *IEEE Trans. Automatic Control*, vol. AC-10, pp. 434–439, Oct. 1965.

- [40] R. L. Kashyap, "Maximum likelihood identification of stochastic linear systems," *IEEE Trans. Automatic Control*, vol. AC-15, pp. 25–34, Feb. 1970.
- [41] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Trans. Automatic Control*, vol. AC-15, pp. 175–184, Apr. 1970.
- [42] J. C. Shellenbarger, "Estimation of covariance parameters for an adaptive Kalman filter," *Proc. National Electronics Conf.*, vol. 22, pp. 698–702, 1966.
- [43] P. D. Hanlon and P. S. Maybeck, "Multiple-model adaptive estimation using a residual correlation Kalman filter bank," *IEEE Trans. Aerospace and Electronic Systems*, vol. 36, pp. 393–406, Apr. 2000.
- [44] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York, NY: Wiley, 1973.
- [45] G. Casella and R. L. Berger, *Statistical Inference*. Belmont, CA: Duxbury Press, 1990.
- [46] G. Noriega and S. Pasupathy, "Adaptive estimation of noise covariance matrices in real-time preprocessing of geophysical data," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1146–1159, Sep. 1997.
- [47] E. Fornasini and G. Marchesini, "Doubly-indexed dynamical systems: State-space models and structural properties," *Mathematical Systems Theory*, vol. 9, pp. 59–72, 1978.
- [48] J. W. Woods and C. W. Radewan, "Kalman filtering in two dimensions," *IEEE Trans. Information Theory*, vol. IT-23, pp. 473–482, July 1977.

Vita

Clint Slatton was born in Huntsville, Alabama in 1970. From there his family moved to El Paso, Texas, where he remained through high school.

He came to the University of Texas at Austin in the fall of 1988. He graduated third in his class with B.S. in aerospace engineering in 1993. He went on to obtain a M.S. in aerospace engineering in 1997. During this time he received a summer scholarship from the California Institute of Technology and interned several times with the American Airlines Turbine Engine Facility and the Jet Propulsion Laboratory.

In 1997, he made a decision to move into the field of electrical engineering. He obtained a M.S. in electrical engineering in 1997. During his time in electrical engineering, he received several full-support and supplemental fellowships from The University of Texas at Austin, The Texas Space Grant Consortium, and The National Aeronautics and Space Administration. He is very grateful to those institutions for their support.

He received his Ph.D.E.E degree in the area of Telecommunications and Information Systems Engineering from The Department of Electrical and Computer Engineering at the University of Texas at Austin, Austin, Texas. His current research interests include multiresolution data fusion, digital signal processing for synthetic aperture radar, and inverse problems in radar scattering.

Permanent Address: 3109 Jazz Street

Round Rock, Texas, 78664

This dissertation was typeset with $\text{\LaTeX} 2_{\epsilon}$ ¹ by the author.

¹ $\text{\LaTeX} 2_{\epsilon}$ is an extension of \LaTeX . \LaTeX is a collection of macros for \TeX . \TeX is a trademark of the American Mathematical Society. The macros used in formatting this dissertation were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin.