NEW TECHNIQUES FOR WIND SCATTEROMETRY

Travis E. Oliphant
Department of Electrical and Computer Engineering
Master’s Degree, August 1996

ABSTRACT

Wind scatterometry is the determination of vector winds over the earth’s oceans using radar data from a satellite scatterometer. In this thesis I develop three techniques which can augment current methods of wind retrieval from scatterometer data. In the first part I investigate methods of obtaining reliable error covariance estimates on retrieved wind. After expanding the accepted statistical measurement model to incorporate geophysical model function uncertainty, I derive the Cramer-Rao bound for both point-wise and model-based retrieved winds and show that it can be used as an approximation to the covariance. In the second part I develop a method to eliminate wind aliases using a hypothesis testing procedure based on a likelihood ratio statistic. I then apply this method to both point-wise and model-based retrieved winds with significant success, showing that, especially with model-based wind fields, a single wind estimate can often be retrieved from the data alone. In the third part of this thesis, I develop three additional wind field models for use in model-based retrieval and compare them with previously used models.

COMMITTEE APPROVAL:

David G. Long, Committee Chairman

Richard L. Frost, Committee Member

Wynn C. Stirling, Committee Member

NEW TECHNIQUES FOR WIND SCATTEROMETRY

A Thesis
Submitted to the
Department of Electrical and Computer Engineering
Brigham Young University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

by
Travis E. Oliphant
August 1996

This thesis by Travis E. Oliphant is accepted in its present form by the Department of Electrical and Computer Engineering of Brigham Young University as satisfying the thesis requirement for the degree of Master of Science.

David G. Long, Committee Chairman

Richard L. Frost, Committee Member

Wynn C. Stirling, Committee Member

Brent E. Nelson, Department Chairman

Date

Wynn C. Stirling, Graduate Coordinator

DEDICATION

To my wife
Amy

ACKNOWLEDGMENTS

First and foremost I would like to acknowledge Dr. David Long for his guidance and support from before I began my graduate studies through completion of this thesis. His helpful evaluations and suggestions have been extremely beneficial. Dr. Wynn Stirling has also played an important role in helping me formulate some of the ideas for this thesis. In addition, I would like to thank Paul Johnson for countless insightful discussions that have broadened my understanding. I would also like to thank my parents Norman and Elizabeth Oliphant for their continual encouragement and support. Finally, reality compels me to acknowledge that without the sacrifices of my wife on behalf of me and our daughter, Amanda, this thesis would not exist. Amy, I love you.

Contents

Dedication **iii**

Acknowledgments **iv**

1 Introduction **1**
- 1.1 Background . . . 1
- 1.2 Current wind retrieval methods . . . 2
- 1.3 Summary of Contributions . . . 3

2 Wind Scatterometry Background **4**
- 2.1 Overview . . . 4
- 2.2 Scatterometers . . . 4
  - 2.2.1 Radar equation . . . 4
  - 2.2.2 ERS-1 and NSCAT instruments . . . 5
- 2.3 Geophysical model function . . . 7
- 2.4 Statistical Models of Measured Backscatter . . . 12
  - 2.4.1 Motivation for studying statistical models . . . 12
  - 2.4.2 Model explanation . . . 13
- 2.5 Point-wise Wind Retrieval . . . 16
  - 2.5.1 Method . . . 16
  - 2.5.2 Problems with point-wise retrieval . . . 18
- 2.6 Model-based wind retrieval . . . 19
  - 2.6.1 Method . . . 19
  - 2.6.2 Status of model-based retrieval . . . 20

3 Error estimation in Wind Retrieval **22**
- 3.1 Statistical Model Revisited . . . 22
  - 3.1.1 Uncertainty in the GMF . . . 22
  - 3.1.2 Effect of GMF uncertainty on overall measurement model . . . 24
  - 3.1.3 Effect on Wind Estimation . . . 30
  - 3.1.4 Final measurement model . . . 31
- 3.2 Application of new model . . . 32
- 3.3 Cramer-Rao Bound for Wind Scatterometry . . . 35
  - 3.3.1 Description of Cramer-Rao Bound . . . 35
  - 3.3.2 Derivation for point-wise retrieval . . . 36
  - 3.3.3 Covariance of model-based wind estimate . . . 40
  - 3.3.4 Point-wise Results . . . 45
  - 3.3.5 Model-based Results . . . 53
  - 3.3.6 Summary . . . 59

3.4 Direct Covariance Approximation 62
3.4.1 Method 62
3.4.2 Derivation 63
3.4.3 Comparisons with Cramer-Rao bound 65
3.5 Summary of Wind Covariance Estimates 71

4 Distinguishing Maxima in the Maximum-Likelihood Equation 72
4.1 Problem background 73
4.2 Setting up the hypothesis test 74
4.2.1 General Approach 74
4.2.2 Practical Approach 75
4.3 Determining the size of the test 76
4.3.1 Defining a related statistic 76
4.3.2 Finding an explicit formula for the size 78
4.4 Using estimate of wind instead of expected value. 81
4.5 Determining an upper-bound for the size 86
4.6 Extension to model-based 87
4.7 Application to point-wise wind-alias elimination 90
4.7.1 ERS-1 90
4.7.2 Simulated NSCAT 98
4.7.3 Discussion 98
4.8 Application to model-based alias elimination 98
4.8.1 Case 1: $F = I$ 100
4.8.2 Case 2: $F \neq I$ 103
4.9 Summary 103

5 Wind Field Modeling 104
5.1 Motivation 104
5.2 Wind Vector Space and Model Sub-Space 105
5.3 Karhunen-Loeve Basis Wind-Field Vectors 107
5.3.1 Determination of the K-L F matrix 107
5.3.2 10×10 K-L basis wind fields 109
5.4 Legendre Polynomial Basis Functions 114
5.5 Fourier-series wind model 116
5.6 Comparisons 117
5.7 Summary 118

6 Conclusions 130
6.1 Summary of Contributions 130
6.1.1 Error estimation in wind retrieval 130
6.1.2 Distinguishing maxima for wind alias elimination 131
6.1.3 Wind field modeling 131
6.2 Future Research 131

6.2.1 Error estimation in wind retrieval 132
6.2.2 Distinguishing maxima for wind alias elimination 132
6.2.3 Wind field modeling 133

A Approximating a Near-Gaussian Density 139

B Bias in the wind estimate 141
B.1 Approximating the bias 141
B.2 Biased Cramer-Rao bound 142
B.3 Examples 144
B.3.1 ERS-1 144
B.3.2 NSCAT 148
B.4 Conclusion 148

C Detailed algebra for wind alias elimination 152

D Relationship between model-based Fisher information matrix and point-wise Fisher information matrices when F has rank $2MN$. 155

E Calculating curl and divergence of model wind field by matrix multiplication 157
E.1 Derivation of the Approximation of *Maximum* Order Derivative Matrix . . 158
E.1.1 Interpolating the derivative estimator 158
E.1.2 Differentiating the interpolating function 161
E.2 Using the MODA Matrix with a Row-scanned matrix 163
E.3 Using MODA to find Γ 164

F MATLAB Code for calculating Legendre and Fourier F matrices 166
F.1 Legendre F matrix 166
F.2 Fourier F matrix 167

List of Tables

3.1 Central moments, μ_m, for $m = 2 \ldots 5$ for the expanded probability model of the measurements z given the wind w. 29
4.1 Objective function values and test results for wind-field aliases in Figures 4.17 and 4.18. $K_{pm} = 0.13$. 100
5.1 NRSS projection errors for four different models as a function of the number of model parameters. 127
5.2 RMS speed errors for four different models as a function of the number of model parameters. 128
5.3 RMS direction errors for four different models as a function of the number of model parameters. 129

List of Figures

2.1 Demonstration of scatterometer operation. 5
2.2 Measurement geometry for ERS-1 and NSCAT satellites. 6
2.3 Plots of σ° versus relative azimuth angle, χ for various incidence angles and speeds. The speeds chosen for plotting are 5, 15, 25, 35, and 45 m/s where σ° increases with increasing wind speed. The GMF is the SASS-2, Ku-band model function to be used with initial NSCAT data. 9
2.4 Plots of wind vectors that would support the measurements taken for a particular wind vector cell for ERS-1 and NSCAT. 11
2.5 Block diagram showing how wind vector affects final scatterometer measurement z and where noise is added. 13
2.6 Example likelihood functions used in wind retrieval for both ERS-1 and NSCAT. 17
3.1 Model for σ° measurements used in this thesis. 24
3.2 Probability of z given w compared with a true Gaussian of the same mean and variance. The two curves in (a) are nearly identical. 26
3.3 Comparison of actual density function to a Gaussian with a first order correction factor. 29
3.4 Distributions for wind retrieved using a Gaussian model and a third-moment correction model. 31
3.5 Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at near-track swath location (ERS-1). 46
3.6 Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at mid-track swath location (ERS-1). 47
3.7 Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at far-track swath location (ERS-1). 48
3.8 Locus of wind vectors that would give rise to measurements obtained with no noise when the true wind is 5 m/s at 120°. 50
3.9 ERS-1 dependence of wind estimate uncertainty on cell location within a swath. 51
3.10 ERS-1 dependence of wind estimate uncertainty on K_{pm} for far-swath. . . 52
3.11 Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at far-track swath location (NSCAT). 54
3.12 NSCAT dependence of wind estimate uncertainty on cell location within a swath. 55
3.13 NSCAT dependence of wind estimate uncertainty on K_{pm} for far-swath. . 55

3.14 The top plot shows RMS direction standard deviation as a function of location in the model region. The bottom plot shows RMS speed standard deviation as a function of location. The Cramer-Rao bound for 129 simulated wind fields mapped to the geometry of the ascending portion of ERS-1 orbit 4452 were used to compute the RMS values. 57
3.15 Concentration ellipses for both point-wise (solid) and 22-parameter model-based (dashed) wind estimates calculated using the Cramer-Rao bound for a particular ‘‘true’’ wind field and corresponding projected field. Note that the ellipses are plotted in rectangular, u, v space. The axis for all the subplots is given by the upper left-hand corner subplot. Cross-track is across the page (1 to 10), and along-track is down the page. 58
3.16 Concentration ellipses for both point-wise (solid) and 12-parameter model-based (dashed) wind estimates calculated using the Cramer-Rao bound for a particular ‘‘true’’ wind field and corresponding projected field. Note that the ellipses are plotted in rectangular, u, v space. The axis for all the subplots is given by the upper left-hand corner subplot. Cross-track is across the page (1 to 10), and along-track is down the page. 60
3.17 Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0$). 66
3.18 Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0$). 67
3.19 Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0.20$). 68
3.20 Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0.20$). 69
4.1 Dependence on wind velocity of α_0 with $K_{pm} = 0$ for the wind aliases of a mid-swath ERS-1 cell in the ascending portion of orbit 7220. The apparent peaks of the last 4 figures is actually the effect of numerical round-off error. The surfaces are essentially flat. 82
4.2 Dependence on wind velocity of α_0 with $K_{pm} = 0.20$ for the wind aliases of a mid-swath ERS-1 cell in the ascending portion of orbit 7220. 83
4.3 Dependence on wind velocity of α_0 with $K_{pm} = 0$ and $K_{pm} = 0.20$ for the wind aliases of a near-swath simulated NSCAT cell. 84
4.4 Comparison of the best Chernoff bound with the numerically calculated value of α_0 for ERS-1 data. 87
4.5 Comparison of the best Chernoff bound with the numerically calculated value of α_0 for simulated NSCAT data. 88

4.6 Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the ascending portion of ERS-1 revolution 4452 ($K_{pm} = 0$). 91
4.7 Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the descending portion of ERS-1 revolution 4459 ($K_{pm} = 0$). 91
4.8 Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the ascending portion of ERS-1 revolution 4452 ($K_{pm} = 0.20$). 92
4.9 Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the descending portion of ERS-1 revolution 4459 ($K_{pm} = 0.20$). 92
4.10 Performance of alias-elimination scheme on actual data taken from the ascending portion of ERS-1 revolution 4452 at 25 km resolution (assuming $K_{pm} = 0$). 93
4.11 Effectiveness of alias-elimination scheme over a portion of the swath for ERS-1 orbit 4452, ascending (assuming $K_{pm} = 0$). 95
4.12 Effectiveness of alias-elimination scheme over a portion of the swath for ERS-1 orbit 4452, ascending (assuming $K_{pm} = 0$). 96
4.13 Performance of alias-elimination scheme on actual data taken from the ascending portion of ERS-1 revolution 4452 at 25 km resolution (assuming $K_{pm} = 0.20$). ESA winds are those estimated by the European Space Agency which operates ERS-1. 97
4.14 Performance of alias-elimination scheme on actual data taken from the ascending portion of ERS-1 revolution 4452 at 25 km resolution assuming $K_{pm} = 0$ and a size threshold of 1e-5. 97
4.15 Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information from a simulated NSCAT orbit at 50 km resolution ($K_{pm} = 0$). 98
4.16 Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information from a simulated NSCAT orbit at 50 km resolution ($K_{pm} = 0.2$). 99
4.17 (a) All aliases for a wind region in the descending portion of ERS-1 orbit 4448. (b) Wind-field alias constructed from most likely point-wise aliases. (c) Wind-field alias constructed from the median filter solution initialized with the most likely field. (d) Wind-field alias constructed from point-wise solutions closest to 180° from wind vectors in (c). 101
4.18 Optimized model-based wind fields: (a) initialized with point-wise field in Figure 4.17(c), (b) initialized with field obtained by rotating all vectors in Figure 4.17(c) by 180°. 102
5.1 Logarithmic plot of the eigenvalues of the 10×10 autocorrelation matrix. (Note the different vertical scales.) 110
5.2 Wind field basis vectors 1--6. 111

5.3 Wind field basis vectors 6--12. 112
5.4 Wind field basis vectors 12--18. 113
5.5 First example of projecting a wind field onto PBC and Karhunen-Loeve models. Wind-field taken from ascending portion of ERS-1 revolution 4452. 119
5.6 First example of projecting a wind field onto Legendre and Fourier models. Wind-field taken from ascending portion of ERS-1 revolution 4452. 120
5.7 Second Example of projecting a wind field onto PBC and Karhunen-Loeve models. Wind-field taken from ascending portion of ERS-1 revolution 4459. 121
5.8 Second Example of projecting a wind field onto Legendre and Fourier models. Wind-field taken from ascending portion of ERS-1 revolution 4459. 122
5.9 NRSS projection error as a function of number of unknowns in model for several models. 123
5.10 RMS speed error as a function of number of unknowns in model for several models. 124
5.11 RMS direction error as a function of number of unknowns in model for several models. 125
B.1 Comparison of approximate analytic bias with simulated bias for ERS-1 at near swath. $K_{pm} = 0$. 145
B.2 Comparison of approximate analytic bias with simulated bias for ERS-1 at near swath. $K_{pm} = 0.17$. 146
B.3 Comparison of Cramer-Rao bound predictions on the standard deviations of wind speed and wind direction with simulations for ERS-1 at near swath. The standard deviation of wind speed is shown for a true wind speed of 25 m/s while the standard deviation of wind direction is shown for a true wind speed of 5 m/s. 147
B.4 Comparison of approximate analytic bias with simulated bias for NSCAT at near swath. $K_{pm} = 0$. 148
B.5 Comparison of approximate analytic bias with simulated bias for NSCAT at near swath. $K_{pm} = 0.17$. 149
B.6 Comparison of Cramer-Rao bound predictions on the standard deviations of wind speed and wind direction with simulations for NSCAT at near swath. The standard deviation of wind speed is shown for a true wind speed of 25 m/s while the standard deviation of wind direction is shown for a true wind speed of 5 m/s. 150

Chapter 1

INTRODUCTION

1.1 Background

Accurate short-term weather prediction is an important element of modern science. We commonly rely on the five-day forecast to plan our excursions and outdoor activities. More importantly, many industries including transportation and agriculture rely on weather prediction in order to make crucial decisions. Unfortunately, weather prediction is often hindered by lack of knowledge of current global weather conditions. Before the age of satellites for monitoring the earth, information about current conditions was limited to isolated weather stations. This means that over the oceans, where 3/4 of weather occurs, almost nothing was known about daily weather conditions.

Satellites have made it possible to understand current weather conditions on a global scale. An important piece of weather information for meteorology is accurate high-resolution near-surface winds over the ocean. In 1978 the experimental Seasat scatterometer (SASS) first demonstrated that a radar could accurately infer vector winds over the ocean's surface from space at an unprecedented spatial resolution and frequency (Davison and Harrison, 1990; Levy and Brown, 1991; Stoffelen and Cats, 1991).

A wind scatterometer such as the one on-board Seasat is an active instrument that sends pulses of microwave radiation to the earth and measures the return power. From this measurement it infers the normalized radar cross section, σ°, of the ocean surface. With a set of these noisy σ° measurements, the wind vector can be estimated using a relationship between surface wind over the ocean and σ°. This relationship is known as the Geophysical Model Function (GMF).

SASS operated at 14.6 GHz (Ku-band) with either horizontal or vertical polarization and nominally used two measurements of σ° to estimate the wind velocity over a 50km by 50km cell(Grantham et al., 1977; Johnson et al., 1980). A spacecraft power failure shortened the SASS mission, but the SASS successes prompted further plans for wind scatterometers.

In July 1991 the European Space Agency (ESA) launched its first wind scatterometer on-board the ERS-1 (European Remote Sensing-1) satellite. This wind scatterometer operates at 5.3 GHz (C-band) with vertical polarization and makes three measurements over each cell of a 25km by 25km grid from which wind velocity is estimated on a 50km by 50km grid (Attema, 1991). ESA launched a second identical scatterometer on-board ERS-2 on April 20, 1995, but technical difficulties made the wind data unavailable until the summer of 1996.

NASA has plans for two of its own wind scatterometers: NSCAT (NASA SCATterometer) to be launched in August, 1996 on-board a Japanese satellite, ADEOS (ADvanced Earth Observing System), and SeaWinds to be launched in 1999. NSCAT will take four measurements over each cell of a 25km by 25km grid and will nominally retrieve wind over a 50km by 50km grid, although 25km by 25km retrieval is possible (Naderi et al., 1991).

1.2 Current wind retrieval methods

There are two general methodologies employed for wind retrieval using scatterometer measurements: point-wise wind retrieval and model-based wind retrieval. Since there is a relationship between σ° measurements and the wind vector over a single patch of the ocean surface, wind may be estimated patch-by-patch from the σ° measurements over just that patch. Such wind retrieval is conveniently named point-wise retrieval and each patch is called a *wind vector cell.* As will be explained in more detail in Section 2.3, the nature of the geophysical model function which relates σ° to winds is such that a single wind vector estimate can rarely, if ever, be resolved from only σ° measurements (Long and Mendel, 1990a). Typically, a second step called ‘‘dealiasing’’ is used to select from among the several ambiguities a single wind vector in each cell. The dealiasing procedure relies on *ad hoc* considerations of how wind should behave from cell to cell and is consequently more difficult to analyze.

As a result of these difficulties the second approach to wind retrieval known as model-based retrieval has been suggested and developed (Long and Mendel, 1990b). This method uses a wind field model covering several wind vector cells. The σ° measurements are then used to estimate the parameters of the model. This method of wind retrieval has limitations. The chief problem is that estimating the parameters of the wind field model becomes a maximization problem with multiple local maxima in a multi-dimensional space. The task of finding all significant local maxima is difficult with the current wind field model.

One missing element in both methods of wind retrieval is that no information is reported about the possible error in the retrieved wind. While simulations have been used traditionally (Leotta and Long, 1989), no reliable procedure for reporting error on each wind estimate has been established and so error estimates on retrieved wind are not commonly reported.

1.3 Summary of Contributions

The contributions of this thesis can be divided into three categories each of which address some problem in current wind scatterometry.

The first major contribution of this thesis is to address the problem of error estimates on retrieved winds. As the statistical model used for the σ° measurements is the most important factor affecting the reliability of such error estimates, I pursue an investigation into this statistical model. Using this model, I develop a method of determining approximate error-bars on both point-wise and model-based retrieved winds. Part of this development involves the derivation of the Cramer-Rao bound for both point-wise and model-based wind retrieval.

The second major contribution of this thesis is to apply the statistical model used for wind retrieval to develop a decision-theory-based method of distinguishing among the maxima of a maximum-likelihood equation such as the one used in wind estimation. For both point-wise and model-based retrieval the likelihood function has multiple maxima, only a few of which are large enough to be believable. I apply decision theory to aid in determining which maxima actually correspond to wind estimates that statistically support the measurements.

While the previous two contributions address problems in both point-wise and model-based retrieval, the final general contribution is focused on model-based retrieval. I investigate alternative wind field models using a basis-field concept in order to improve the practical implementation of model-based retrieval by reducing the number of parameters to be estimated.

To demonstrate the applicability of these contributions, actual data from the ERS-1 scatterometer will be used as will simulated data from NSCAT. Nonetheless, the derivations in this thesis can be generally applied to any of the scatterometers previously mentioned.

Chapter 2

WIND SCATTEROMETRY BACKGROUND

2.1 Overview

In order to understand the contributions made by this thesis, it is important to review some fundamentals of wind scatterometry. As a result, this chapter is included to discuss in more detail scatterometers and how wind can be estimated from scatterometer data. This will be accomplished by first briefly describing a scatterometer and how a measurement of normalized radar cross section, σ°, is made. Second, the Geophysical Model Function which relates the ocean wind vector to σ° will be described. Third, a statistical model of the backscatter measurements will be presented which is useful in developing a method of point-wise wind retrieval using σ° data. Finally, model-based wind retrieval will be reviewed as an extension to point-wise retrieval.

2.2 Scatterometers

2.2.1 Radar equation

A scatterometer is an active instrument designed to measure the scattering properties of a target. Refer to Figure 2.1. The scatterometer transmits pulses of electromagnetic radiation and receives the energy scattered off the target. The scattering properties of the target are grouped into a single parameter, σ°, known as the differential scattering cross section. This parameter is a combination of the transmitting gain back towards the antenna and the fraction of intercepted power scattered instead of absorbed. The returned power is related to σ° and the transmitted power according to the radar equation (Ulaby et al., 1982):

$$P_r = \frac{P_t G^2 \lambda^2 A}{(4\pi)^3 R^4} \sigma^\circ, \tag{2.1}$$

where P_t is the power transmitted, G is the gain of the transmitting antenna, λ is the wavelength of the electromagnetic wave, A is the effective illuminated area, and R is the distance from the scatterometer to the target. All the parameters relating σ° to the power

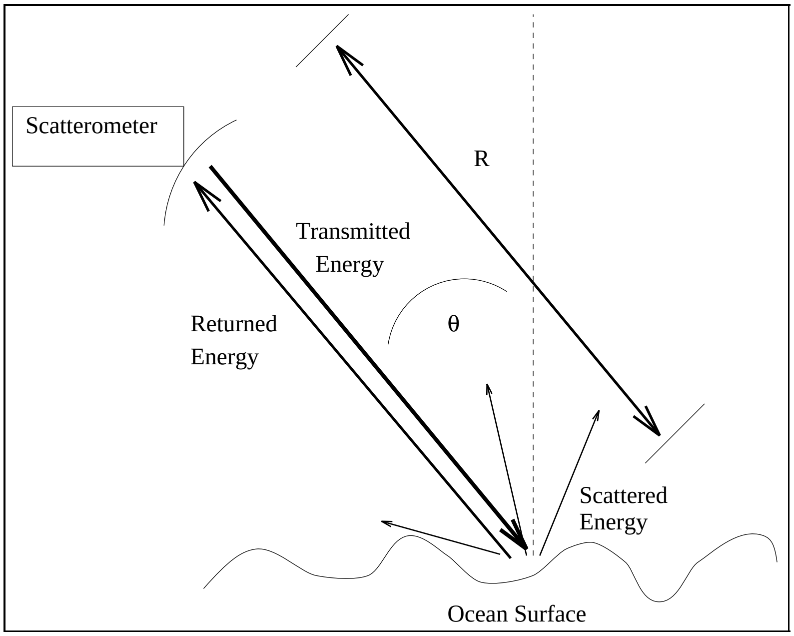


Figure 2.1: *Demonstration of scatterometer operation.*

received can be grouped into one constant C so that

$$P_r = C\sigma^\circ.$$

Note that knowing both the gain of the scatterometer antenna and the distance to the target are imperative to being able to infer σ° from the received power measurement. To make a measurement of σ°, the scatterometer transmits a pulse of radiation and makes a power measurement, P_s of received power. Due to the presence of background noise it then makes a noise-only power measurement, P_n, and determines the return power as $\hat{P}_r = P_s - P_n$. The measurement, z, of the radar scattering cross section is then inferred using the radar parameters, C.

$$z = \frac{\hat{P}_r}{C}$$

2.2.2 ERS-1 and NSCAT instruments

This thesis involves data from ERS-1 and simulated data from NSCAT so a brief description of these instruments is given. Both of these scatterometers are fan-beam scatterometers with illumination patterns shown in Figure 2.2. Both instruments coregister

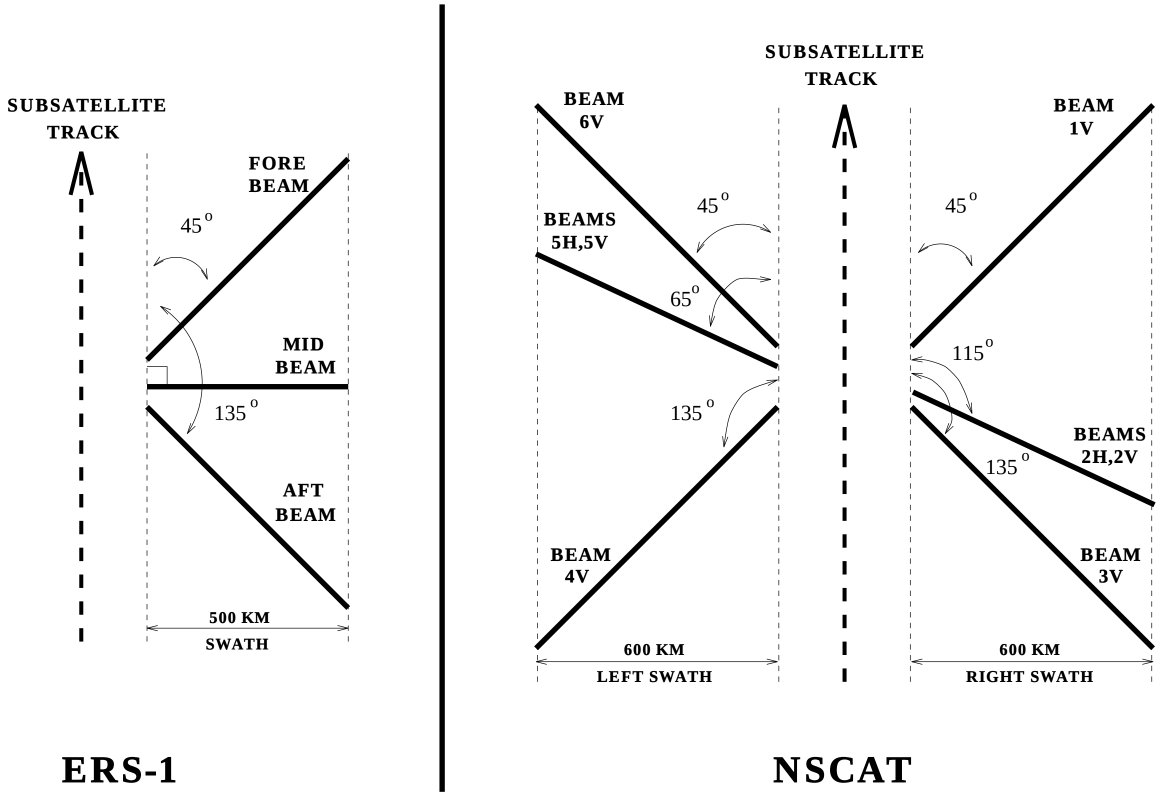


Figure 2.2: *Measurement geometry for ERS-1 and NSCAT satellites.*

the σ° measurements onto a rectangular grid of cells aligned with subsatellite along-track and cross-track directions, called wind vector cells, within which wind can be estimated. On the other hand, there are notable differences between the two instruments that affect wind retrieval beyond the obvious geometry differences. One of these is the way in which the two instruments construct a grid of wind vector cells.

The instrument on ERS-1 uses yaw-steering, programmed such that the effects of the earth's rotation is counteracted, to coregister the σ° measurements onto 19 nodes (Attema, 1991). Ten of these nodes correspond to center points of 50km by 50km cells that span the swath. The other 9 nodes are placed between these ten to generate 19 nodes separated by 25km. Three measurements are provided by ESA at each node, one for each of the beams. Wind can be retrieved at each of the 19 cross-track nodes using the three measurements resulting in 25km by 25km sampling of the wind field. While sampled at 25km by 25km the effective resolution is 50km by 50km (Davis, 1993).

The NSCAT instrument uses digital Doppler processing to colocate the σ° measurements onto 24 cross-track wind vector cells that span each 600km swath. Wind can then be retrieved at true 25km by 25km resolution using the σ° measurements from each cell (Naderi et al., 1991). In order to improve wind retrieval accuracy, the measurements are further grouped into 50km by 50km cells over which the wind vectors are estimated. This results in 12 cross-track wind vector cells on each side of the spacecraft.

There are other notable differences between the two instruments. ERS-1 operates at 5.6 GHz (C-band) and uses a different Geophysical Model Function than NSCAT which operates at 14 GHz (Ku-band). ERS-1 transmits only vertically polarized pulses from its three beams, while NSCAT transmits both vertical and horizontal polarization from its center beam. The result is that NSCAT nominally has four available measurements from which to retrieve wind over the 25km by 25km cells instead of three. A final important difference between the two instruments is level of transmit-power: ERS-1 transmits approximately 5kW peak power compared to about 100W peak power for NSCAT. As a result, the signal-to-noise ratio is higher for ERS-1 than for NSCAT.

2.3 Geophysical model function

With a general understanding of how σ° measurements are made, consider how these measurements are used to estimate (retrieve) wind velocity over the ocean. There is an indirect relationship between the wind blowing over the ocean surface and the radar scattering cross section, σ°, which the scatterometer measures. Wind creates waves on the ocean surface. The waves determine how much the ocean reflects electromagnetic energy. This correlation between wind velocity and σ° allows estimation of wind from σ°.

The reliability of the wind estimate based on σ° data is directly related to our understanding of the relationship between wind velocity and σ°. Several studies have been made on the connection between wind and waves and between wind and σ° (Donelan and W.J. Pierson, 1987; Jones et al., 1977). These studies show that the connection is a complex one. As a result, developing a theoretically based model of the dependence of σ° on wind velocity derived from first principles is extremely unwieldy. Consequently, empirical studies have been made to understand the relationship. Out of these studies have come operational Geophysical Model Functions (GMF) which relate wind velocity to σ°.

The GMF may be generally expressed as

$$\sigma^\circ = \mathcal{M}(\theta, U, \chi, f, p)$$

where U is the wind speed, χ is the azimuth angle between the wind direction and the radar azimuth angle, θ is the radar incidence angle, f is the frequency of the radar, and p is the polarization of the transmitted energy (horizontal or vertical). Usually the frequency, f, is fixed for a particular instrument. In addition, it is often desirable to express the dependence on the wind direction, ϕ, and radar azimuth angle, ψ, separately. As a result we often write

$$\sigma^\circ = \mathcal{M}(\theta, U, \psi - \phi, p) = \mathcal{M}(\theta, U, \phi, \psi, p).$$

The direction convention chosen for the azimuth angle and wind direction is arbitrary as long as it is consistent since only the difference is important to the GMF. We will assume the convention of measuring these angles in degrees clockwise from north.

Several functional forms have been proposed for the GMF. A common approach, however, is to simply store the empirically-derived data points in a multi-dimensional table and interpolate intermediate values (Naderi et al., 1991). This is the approach used in this thesis.

According to the relationship predicted by the GMF, σ° is a function of only wind speed, relative wind direction, and radar incidence angle. However, other factors such as sea temperature, local salinity, and long gravity waves also affect the relationship. Thus there is some uncertainty, or modeling error, in the GMF (Long, 1989).

Given a GMF, wind estimation (or retrieval) becomes an inversion problem. Measurements of σ° are taken at known values of ψ and θ. Values of U and ϕ are then selected as the wind estimate in order to be statistically consistent with these observations.

Examining the SASS-2 GMF

In this section the SASS-2 GMF (Wentz et al., 1984), derived for (Ku-band) from Seasat data, will be examined in order to illustrate its dependence on wind speed and direction. This will aid in understanding the problems associated with inverting the model function to obtain wind estimates.

Figure 2.3 shows plots of $\mathcal{M}$ versus relative azimuth angle, χ, for several values of wind speed and incidence angle. Notice the $\cos(2\chi)$ dependence on azimuth angle.

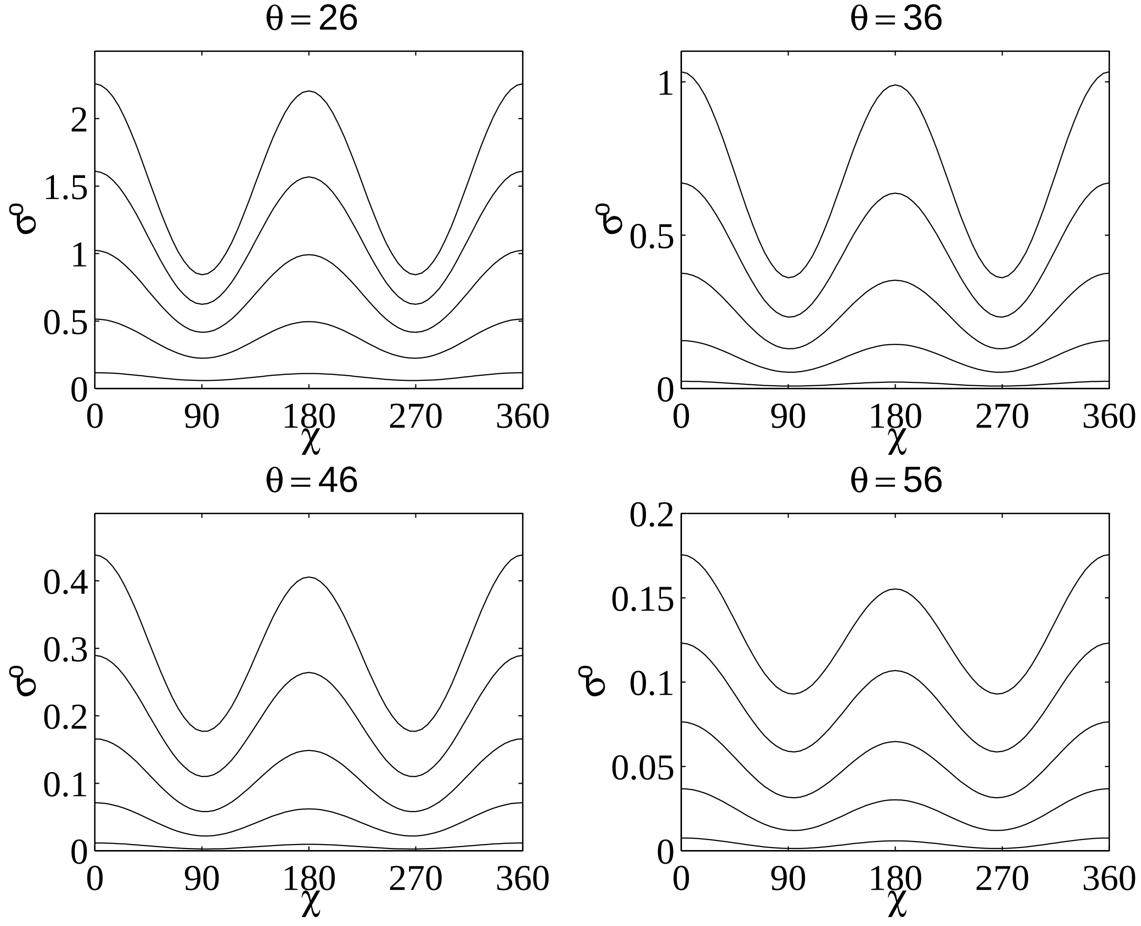


Figure 2.3: *Plots of σ° versus relative azimuth angle, χ for various incidence angles and speeds. The speeds chosen for plotting are 5, 15, 25, 35, and 45 m/s where σ° increases with increasing wind speed. The GMF is the SASS-2, Ku-band model function to be used with initial NSCAT data.*

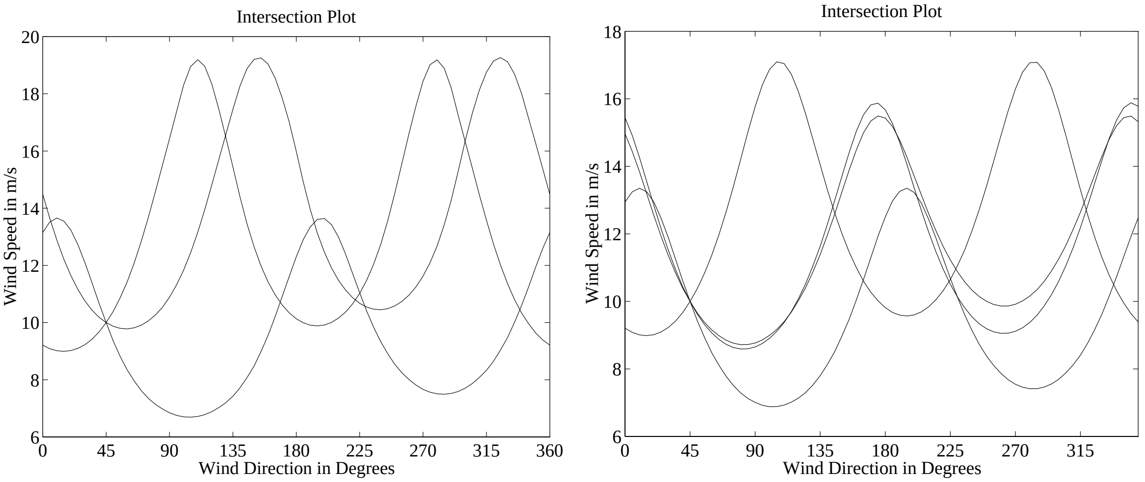
There is also a general trend of increasing σ° for increasing wind speed. In addition, σ° decreases as incidence angle increases.

Considering the monotonic trend of σ° for increasing wind speed, it is not surprising that wind speed estimation is fairly accurate using most estimation techniques. The double-cosine dependence on wind direction, however, makes it difficult to estimate wind direction from σ° data. In particular, notice the similarity of σ° when measured with a relative azimuth angle of 0° and 180°, especially at low incidence angles and low wind speeds. This upwind/downwind similarity makes it very difficult to uniquely resolve a single wind direction from noisy σ° measurements. As will be seen later, usually a set of at least two wind velocities, called aliases, must be returned over each region as possible wind solutions.

Figure 2.3 also shows that more than one wind speed and direction can give rise to the same value of σ° for fixed radar incidence and azimuth angles. Consider the upper-left plot in the figure, and notice that for any value of χ there will be one speed which gives rise to $\sigma^\circ = 0.5$. As a result, multiple measurements are needed to define a finite set of possible wind vectors.

Another way to visualize the model function as it relates to wind retrieval and see more clearly the need for multiple measurements is to plot the locus of wind velocities that give rise to a single σ° for fixed radar incidence and azimuth angles. In Figure 2.4(a) such a plot is shown for three sets of radar angles encountered with data from the ERS-1 scatterometer. Figure 2.4(b) shows the same type of plot for four sets of radar angles representative of data from the NSCAT scatterometer. These plots were generated by choosing a wind speed and direction and generating σ° values using the C-band GMF (CMODFDP) for ERS-1 (Freilich and Dunbar, 1993) and SASS-2 for NSCAT. For each radar incidence and azimuth angle, all the wind speeds and directions that would induce that σ° were plotted as a single curve. The true wind speed and direction chosen was 10 m/s and 45°.

As shown in Figure 2.2, the beams on ERS-1 are spaced about 45° apart with the fore and aft beams at the same incidence angle and the middle beam at a lower incidence angle. Figure 2.4(a) shows that according to the GMF all three of these measurements are needed to uniquely determine a wind velocity. This figure also shows that there is a near intersection point approximately 180° from the true wind direction due to the upwind/downwind similarity. In the presence of noise, it is almost always impossible to



(a) ERS-1

(b) NSCAT

Figure 2.4: *Plots of wind vectors that would support the measurements taken for a particular wind vector cell for ERS-1 and NSCAT.*

distinguish this alias from the true wind velocity as the effect of noise is to move the curves vertically.

The beams on NSCAT are similar to those on ERS-1 except for the center beam which is offset from $45°$ as shown in Figure 2.2. In addition, two polarizations are measured on the middle beam so that nominally four measurements are made over each 25km region. The plot in Figure 2.4(b) shows that the second polarization on the middle beam does not add significant extra information as to the true wind velocity. Notice that the $180°$ alias is also present for NSCAT measurements.

2.4 Statistical Models of Measured Backscatter

2.4.1 Motivation for studying statistical models

In general, wind estimation can be considered in the class of image reconstruction problems. The $\sigma°$ measurements are the observed ‘‘image’’ and the wind to be retrieved is the true ‘‘image’’. The $\sigma°$ measurements are corrupted by noise that is in general multiplicative, that is to say the statistics of the noise are dependent on the actual $\sigma°$ value. At each pixel, or cell, in a region the wind velocity is mapped to a set of $\sigma°$ values through a so-called point non-linearity, corresponding to the actual scattering properties of the wind-driven surface. Noise corrupts the true $\sigma°$ values during measurement. The job of wind retrieval is to restore the original wind ‘‘image’’ from the observed $\sigma°$ ‘‘image’’.

There are several properties of this image reconstruction problem that make it particularly difficult to perform. The most impeding difficulty is the non-linear mapping between wind velocity and $\sigma°$. Because this mapping is not only non-linear but non-unique, it is at best extremely difficult to resolve a single wind direction even with a high signal-to-noise ratio. At worst, six wind directions could reasonably give rise to the observations. A second difficulty is that the ‘‘image’’ to be restored has two values per pixel (wind speed and direction) and the observed ‘‘image’’ has many values per pixel corresponding to the different radar looks necessary to infer a finite set of wind velocities.

These difficulties could explain why the techniques of image restoration have not been widely applied to the wind retrieval problem. Instead, traditional estimation techniques have been most often used to retrieve wind from measured $\sigma°$, *e.g.* maximum-likelihood estimation. In fact, the working algorithm that JPL (Jet Propulsion Laboratory) will use for NSCAT retrieval is a maximum likelihood estimation procedure (Dunbar et al.,

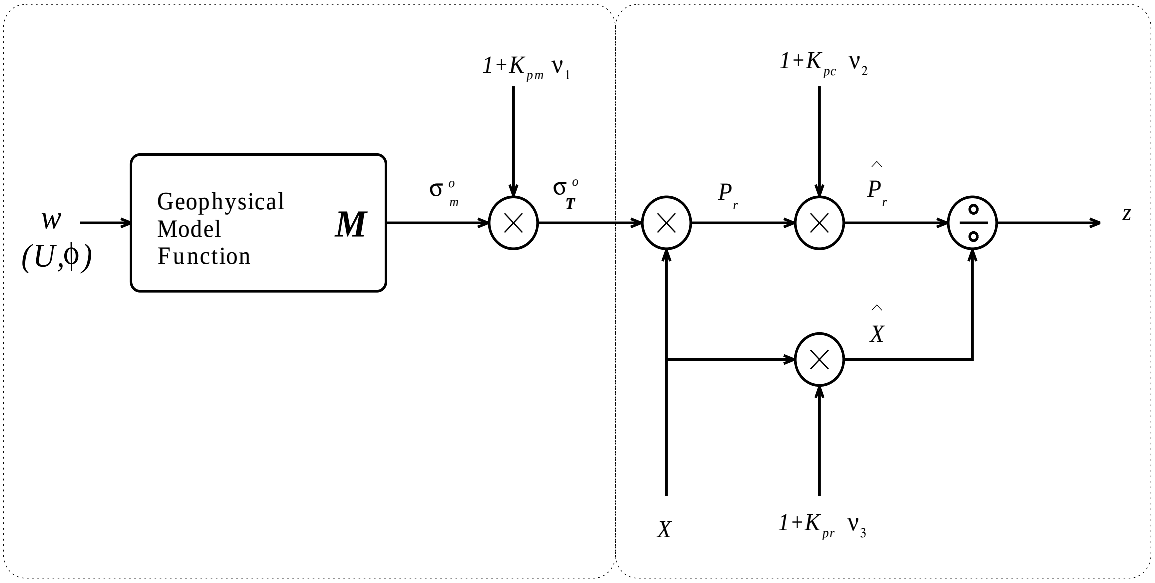


Figure 2.5: *Block diagram showing how wind vector affects final scatterometer measurement z and where noise is added.*

1988).

Since many wind estimation procedures, including those used in current wind retrieval techniques rely on a statistical model of the measurements, the next section presents a statistical model of σ°. In the following z represents a σ° measurement, σ_t° represents the true value of σ° that would have been measured in a noise free environment, and C represents the parameters in the radar equation as shown in Section 2.2.

2.4.2 Model explanation

Many kinds of estimation procedures are dependent on a model for the probability density of the observations conditioned on the object to be estimated. In this case we are concerned with the probability density function (PDF) of the measured value of normalized radar cross section z conditioned on the wind velocity $\mathbf{w}$. The components of wind velocity can be in either polar (U, ϕ) or rectangular (u, v) form[1]. This PDF is denoted $p_z(z|\mathbf{w})$. While $\mathbf{w}$ can be treated as either deterministic or random, it is traditionally treated as deterministic.

For completeness, this section presents a general model of the σ° measurements

---

[1] It is convention to measure ϕ clockwise from north so that $u = U\sin(\phi)$ and $v = U\cos(\phi)$.

introduced by Long (1989) and shown in Figure 2.5. This model accounts for most sources of variability in σ° given a fixed wind velocity. In practice a more simplistic statistical model is employed, so the assumptions leading to this PDF for the measurements are also outlined.

Assuming fixed incidence angle, azimuth angle, and polarization, the wind vector gives rise to a predicted normalized radar cross section of the ocean surface given by the GMF:

$$\sigma_m^\circ = \mathcal{M}(\theta, \mathbf{w}, \psi, p). \tag{2.2}$$

Although wind is a dominant factor, it is not the only factor affecting σ°. As a result, even with a fixed wind vector, there will still be some variability in the true value of σ°. This uncertainty can be modeled by defining a new random variable, σ_t°:

$$\sigma_t^\circ = \sigma_m^\circ(1 + K_{pm} v_1), \tag{2.3}$$

where v_1 is a zero-mean, unit-variance Gaussian random variable. Note that K_{pm} is in general a function of the wind vector. The true σ° of the surface, σ_t°, is scaled by the true radar equation parameter, C, to get a true power return, P_r. The measurement of the returned power made by the scatterometer is corrupted by instrument noise and background noise. This source of uncertainty can be modeled by representing the scatterometer estimate of the returned power as a random variable $\hat{P}_r$:

$$\hat{P}_r = P_r(1 + K_{pc} v_2) = C\sigma_t^\circ(1 + K_{pc} v_2). \tag{2.4}$$

Note again that K_{pc} is in general a function of P_r and v_2 is another zero-mean, unit-variance Gaussian random variable. The measured value of returned power is then divided by an estimate of the radar parameters, $\hat{C}$, in order to obtain a measurement of σ°. Uncertainty in the radar parameter estimate can be modeled by treating $\hat{C}$ as a random variable through the introduction of another zero-mean, unit-variance Gaussian random variable:

$$\hat{C} = C(1 + K_{pr} v_3). \tag{2.5}$$

Combining these expressions, the resulting σ° measurement can be written as

$$z = \frac{\hat{P}_r}{\hat{C}_r} = \frac{\mathcal{M}(\theta, \mathbf{w}, \psi, p)(1 + K_{pm} v_1)(1 + K_{pc} v_2)}{1 + K_{pr} v_3} \tag{2.6}$$

This is a formidable model, especially when it is recognized that K_{pm} can be a function of $\mathbf{w}$, and K_{pc} can be a function of $C\sigma_t^\circ$. Compounding the problem is that the only piece of this model that is well understood is K_{pc}. There has been much effort placed in understanding the noise model for the power measurements. It is typically given as (Naderi et al., 1991)

$$K_{pc}^2 = \alpha + \frac{\beta_1}{P_r/P_n} + \frac{\gamma_1}{P_r^2/P_n^2}, \tag{2.7}$$

where P_n is the measured background noise power, $P_r = C\sigma_t^\circ$ is the true return power from the ocean surface, and α, β_1 and γ_1 are values that depend on the instrument design and are reported with the satellite data. (For ERS-1, the signal to noise ratio is large enough that it is assumed that $\beta_1 \approx \gamma_1 \approx 0$.)

Since this is the only well understood source of noise, frequently it is assumed that it is the only source of noise. In other words it is assumed that $K_{pm} = 0$ and $K_{pr} = 0$. With these assumptions, the measurement model becomes (note that the dependencies of the GMF on radar incidence and azimuth angles, wind velocity, and radar polarization have been suppressed).

$$\begin{aligned} z &= \mathcal{M}(1 + K_{pc}v_2), && (2.8) \\ K_{pc} &= \sqrt{\alpha + \frac{\beta_1 P_n}{C\mathcal{M}} + \frac{\gamma_1 P_n^2}{C^2\mathcal{M}^2}}, \\ &= \sqrt{\alpha + \frac{\beta}{\mathcal{M}} + \frac{\gamma}{\mathcal{M}^2}}. && (2.9) \end{aligned}$$

With this model, it is apparent that z is a Gaussian random variable with mean $\mathcal{M}$ and variance $\varsigma_z = \alpha\mathcal{M}^2 + \beta\mathcal{M} + \gamma$ ($\beta = \gamma = 0$ for ERS-1). While the assumptions used to obtain this model may be strained, they do allow us to use a comfortable density function for $p_z(z|\mathbf{w})$.

$$\begin{aligned} p_z(z|\mathbf{w}) &= \frac{1}{\sqrt{2\pi\varsigma_z^2}} \exp\left[-\frac{(z - \mathcal{M}(\theta, \mathbf{w}, \psi, p))^2}{2\varsigma_z^2}\right], && (2.10) \\ \varsigma_z^2 &= \alpha\mathcal{M}^2(\theta, \mathbf{w}, \psi, p) + \beta\mathcal{M}(\theta, \mathbf{w}, \psi, p) + \gamma. && (2.11) \end{aligned}$$

This is the probability distribution on the measurements z that is commonly used to estimate wind from scatterometer measurements of σ°.

2.5 Point-wise Wind Retrieval

2.5.1 Method

As described above, at least two measurements at different azimuth angles are required to obtain a finite set of wind estimates. We denote each measurement, z_k, and the measurement vector, $\mathrm{z} = [z_1, \dots, z_K]$. Associated with each measurement is a set of radar information. In particular, values are given for the radar azimuth and incidence angles (ψ, θ), the K_{pc}-equation constants (α, β, γ), and (for NSCAT measurements) the electromagnetic polarization of the antenna ($p =$ horizontal/vertical). These values are also subscripted by the index, k, given to z_k. Since the commonly used model accounts only for communication noise, it is reasonable to assume that the elements of z are statistically independent. In addition, we simplify notation by writing $\mathcal{M}(\theta_k, U, \phi, \psi_k, p_k)$ as $\mathcal{M}_k$ with the dependence on the wind implied, and the dependence on the radar values expressed by the subscript, k. As a result, the joint distribution of z can be written as

$$
\begin{aligned}
p_{\mathbf{z}}(\mathbf{z}|\mathbf{w}) &= \prod_{k=1}^{K} \frac{1}{\varsigma_{z_k}\sqrt{2\pi}} \exp\left[\frac{-(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^2}\right], \qquad (2.12)\\
\varsigma_{z_k}^2 &= \alpha_k \mathcal{M}_k^2 + \beta_k \mathcal{M}_k + \gamma_k. \qquad (2.13)
\end{aligned}
$$

With this statistical model of the measurements given the wind, the wind is estimated given observations, z_0, by selecting the wind vector $\hat{\mathrm{w}} = (\hat{U}, \hat{\phi})$ that gives the maximum value of the density function. This is simply maximum likelihood estimation:

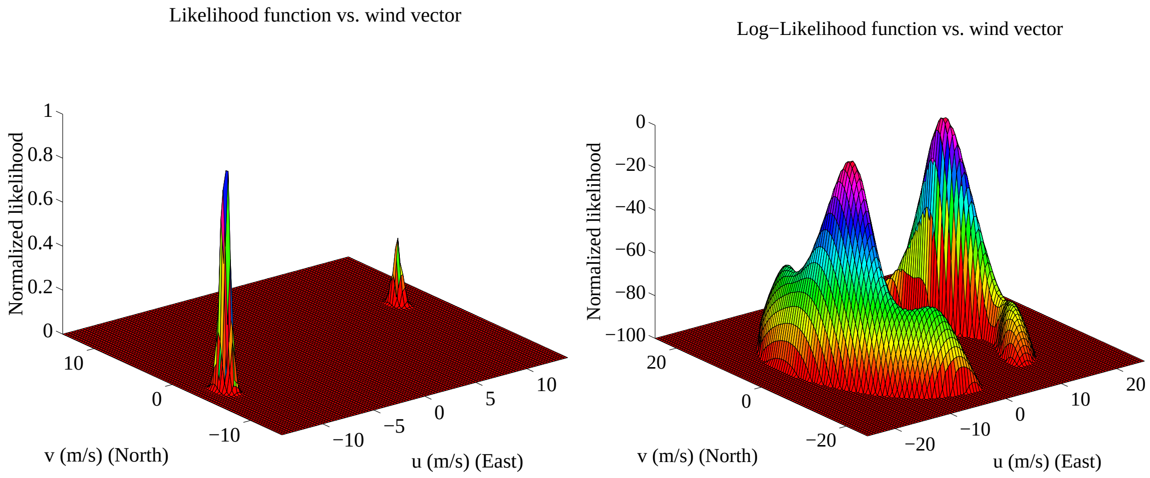
$$
\hat{\mathrm{w}} = \arg\max_{\mathbf{w}} p_{\mathbf{z}}(\mathbf{z}_0|\mathbf{w}) \qquad (2.14)
$$

An equivalent but more computationally tractable problem is to maximize the log-likelihood function over w:

$$
\begin{aligned}
\hat{\mathrm{w}} &= \arg\max_{\mathbf{w}} L(\mathbf{w}, \mathbf{z}_0), \qquad (2.15)\\
L(\mathbf{w}, \mathbf{z}_0) &= \log p_{\mathbf{z}}(\mathbf{z_0}|\mathbf{w})\\
&= -\sum_{k=1}^{K}\left\{\frac{[z_k - \mathcal{M}_k^2]}{2\varsigma_{z_k}^2} + \frac{1}{2}\log[\varsigma_{z_k}^2] + \frac{1}{2}\log(2\pi)\right\}. \qquad (2.16)
\end{aligned}
$$

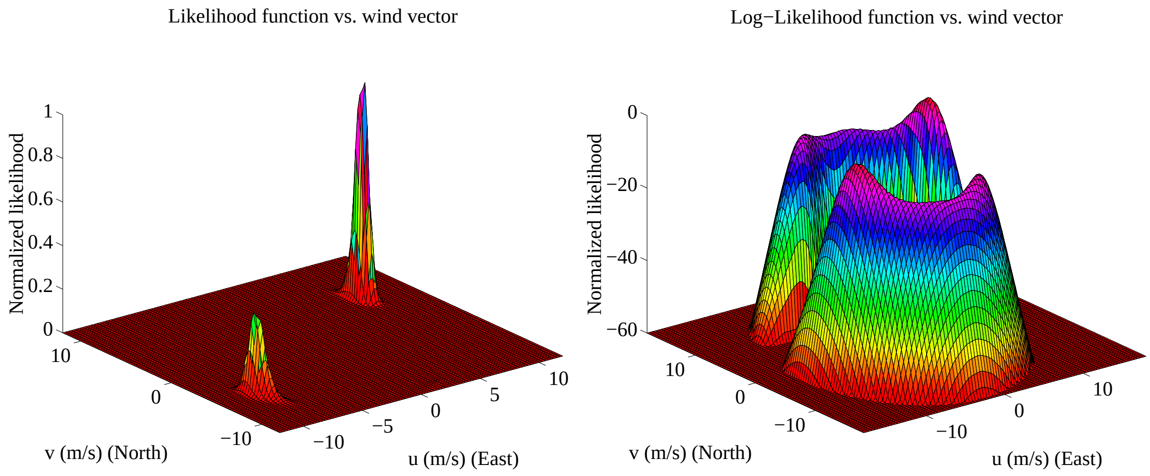
The empirical dependence of $\mathcal{M}_k$ on w requires that the optimization of $L(\mathrm{w}, \mathrm{z}_0)$ be done numerically.

This function typically has several local maxima due to the symmetry inherent in the GMF. Figure 2.6(a) shows a representative normalized likelihood function from an



(a) Likelihood function, ERS-1

(b) Log-likelihood function, ERS-1



(c) Likelihood function, NSCAT

(d) Log-likelihood function, NSCAT

Figure 2.6: *Example likelihood functions used in wind retrieval for both ERS-1 and NSCAT.*

actual ERS-1 wind vector cell. Notice the presence of two dominant peaks which indicate possible wind solutions. Figure 2.6(b) shows the log-likelihood function for the same wind vector cell. This plot makes evident that the likelihood function has more than two local maxima, even though the other maxima are not as significant as the first two. Similar plots for simulated NSCAT measurements are shown in Figures 2.6(c) and 2.6(d). Notice that the smaller signal-to-noise ratio for NSCAT measurements widens the peaks.

In typical wind retrieval all of these local maxima are saved for each wind vector cell as possible wind solutions. Each of these possible solutions is called an ambiguity, or alias. As is evident from the figures, the set of wind solutions have similar speeds but widely varying directions, and some are more likely solutions than others.

A second step called dealiasing, or ambiguity removal, must be performed to choose a single wind field. This step takes into account that wind from cell to cell is correlated. One particular method that is used is an iterative median-filter algorithm described in detail by Shaffer (Shaffer et al., 1991). Briefly, all cells are initialized to the wind estimates giving the highest likelihood values. On each pass, the wind vector in the center of a moving window is replaced by the ambiguity closest to the median of the window. This is repeated until no changes occur.

2.5.2 Problems with point-wise retrieval

Although point-wise retrieval is a useful method for wind field estimation over the ocean, it does have short-comings. One of the problems is the unjustified technique of keeping all local maxima of the likelihood function. It seems reasonable to keep local maxima that are comparable in magnitude to the global maximum, but to keep maxima that are much smaller than the global maximum seems to ignore the basic philosophy of maximum likelihood estimation: choose the wind vector that maximizes the probability of observing the measurements. Passing wind estimates on to the dealiasing step that give rise to significantly lower maxima ignores the fact that such a wind vector would only with very low probability have given rise to the observed measurements. This complicates the dealiasing step and decreases the reliability of the retrieved wind.

Another short-coming of point-wise retrieval is the lack of any derived error-bars for the wind estimates. Part of this problem is due to the lack of understanding about how to model uncertainty in the GMF which relates wind to σ°. Since the wind estimate depends directly on the σ° measurements, knowledge about the distribution of the wind estimate depends on a reliable model of the σ° measurement distribution. Even though little is understood about all noise sources in the σ° measurements, approximations can be made in order to give more specific information as to the accuracy of wind estimates.

A final shortcoming of point-wise wind retrieval is the dealiasing step, which is largely *ad hoc* and difficult to analyze. In addition, it can be error prone over large regions due to selection of the wrong alias. In order to remove this complicated dealiasing step, model-based wind retrieval was developed. This technique will be described in the next section.

2.6 Model-based wind retrieval

2.6.1 Method

To overcome some of the problems with the dealiasing step of point-wise wind retrieval, model-based wind retrieval was introduced by Long and Mendel (1990a). In model-based estimation, the wind vectors over a large region are estimated at the same time. These wind vectors are combined into a single wind field:

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_l \\ \vdots \\ \mathbf{w}_{MN} \end{bmatrix}, \tag{2.17}$$

where l is the lexicographical index into an $M \times N$ matrix. This wind-field matrix is made up of M cross-track rows with N along-track wind vectors. The wind vector in the lth lexicographical cell of this matrix is denoted, $\mathbf{w}_l = (U_l, \phi_l) = (u_l, v_l)$.

The σ° measurements over the region from which the estimate is to be made are grouped into a large measurement vector:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_l \\ \vdots \\ \mathbf{z}_{MN} \end{bmatrix}. \tag{2.18}$$

In this block vector, $\mathbf{z}_l$ is a variable length vector representing the measurements taken over the lth lexicographical cell in the region. It is sometimes convenient to represent the individual elements of each measurement vector and so we establish notation. The kth measurement in the lth vector of Z will be denoted $Z_{k,l}$. Associated with each measurement, $Z_{k,l}$, is a set of radar information. Values are given for the radar azimuth and incidence angles (ψ, θ), the K_{pc} equation constants (α, β, γ), and (for NSCAT measurements) the electromagnetic polarization of the antenna ($p =$ horizontal/vertical). These values will also be subscripted by the index, k, l, given to $Z_{k,l}$.

The joint probability density function of the large measurement vector is constructed from the marginal distributions of the measurements over each cell by assuming independence. As a result,

$$\begin{aligned} p_{\mathbf{Z}}(\mathbf{Z}|\mathbf{W}) &= \prod_{l=1}^{MN} \prod_{k=1}^{K(l)} \frac{1}{\sqrt{2\pi}\varsigma_{Z_{k,l}}} \exp\left[\frac{-\left(Z_{k,l} - \mathcal{M}_{k,l}\right)^2}{2\varsigma_{Z_{k,l}}^2}\right], \\ \varsigma_{Z_{k,l}} &= \alpha_{k,l}\mathcal{M}_{k,l}^2 + \beta_{k,l}\mathcal{M}_{k,l} + \gamma_{k,l}, \end{aligned} \tag{2.19}$$

where $K(l)$ represents the number of measurements in $\mathbf{z}_l$. In addition, to simplify notation we have written $\mathcal{M}(\theta_{k,l}, U_l, \phi_l, \psi_{k,l}, p_{k,l})$ as $\mathcal{M}_{k,l}$.

The key element of model-based estimation is the introduction of a wind model to introduce correlation in the wind field. This model can be written generally as $\mathbf{W} = \mathrm{g}(\mathbf{X})$, where g is a mapping from a dimensionally smaller space of model parameters, $\mathbf{X}$, to the $2MN$ dimension space of sampled wind fields.

Instead of directly estimating the wind using Eq. (2.19), the model parameters are first estimated using maximum likelihood

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X}} \left\{\log p_{\mathbf{Z}}\left[\mathbf{Z}|\mathbf{g}(\mathbf{X})\right]\right\}. \tag{2.20}$$

Then the wind estimate is determined as $\hat{\mathbf{W}} = \mathrm{g}(\hat{\mathbf{X}})$. Note that the same problem of multiple solutions exists due to the symmetry of the GMF. This time, however, many fewer wind fields need to be pieced together to cover a satellite track so the task of dealiasing becomes much simpler.

2.6.2 Status of model-based retrieval

The model in current use is a linear model derived from first principles but which assumes a low-order polynomial curl and divergence field along with a low-order polynomial for the required boundary conditions (Long and Mendel, 1990a). Without going into all the details, the model can be represented by the linear equation.

$$\mathbf{W} = F\mathbf{X} \tag{2.21}$$

The matrix F is $2MN \times M_x$ where $M \times N$ is the wind field region size and M_x is the number of parameters in the model. The number of parameters can be varied according to the model order desired. In this model, the elements of $\mathbf{W}$ are defined to be the rectangular, (u, v) components of the wind vectors in the region. In addition all of the u components are lexicographically ordered into the wind-field vector prior to row-scanning all of the v components. Specifically, $\mathbf{W}$ is defined as

$$\mathbf{W} = \begin{bmatrix} u_1 \\ \vdots \\ u_l \\ \vdots \\ u_{MN} \\ v_1 \\ \vdots \\ v_l \\ \vdots \\ v_{MN} \end{bmatrix}, \tag{2.22}$$

where l is the lexicographical index into the $M \times N$ region.

The current limitation of model-based retrieval is the lack of a suitable algorithm for finding all of the important local maxima expected from the log-likelihood function. Part of this problem is due to the large number of model parameters to be estimated. The other part of this problem is the existence of unimportant local maxima inhibiting the search. Another missing element in model-based retrieval is the same as for point-wise retrieval: a field-by-field estimate of the error in retrieved winds.

Chapter 3

ERROR ESTIMATION IN WIND RETRIEVAL

3.1 Statistical Model Revisited

The most important factor in determining a reliable error estimate for retrieved winds is a realistic model of the noise in the measurements. In Section 2.4.2 a commonly used distribution was presented. As discussed in that section this statistical model makes two assumptions that while simplifying the estimation procedure leave doubts about its accuracy in determining error estimates for wind retrieval. The first simplifying assumption often made is that the GMF is completely accurate in relating wind velocities to σ°. The second assumption is that the radar parameter, C, is known exactly. Both of these are only approximations that need improvement in order to report reliable covariance estimates of retrieved wind. The purpose of this first section is to ameliorate the problem by tighting the first assumption. In particular, this section examines the effect on the probability density function of the measurements if random error in the GMF is allowed.

In order to perform this investigation, the general model for the measurements introduced by Long (1989) and given by Eq. (2.6) is used. As uncertainty in the radar parameter, C, is not considered here, K_{pr} is still assumed to be zero. As part of the investigation on the effect of keeping a non-zero value for K_{pm}, a more detailed discussion of uncertainty in the GMF is presented.

3.1.1 Uncertainty in the GMF

The Geophysical Model Function can be most generally thought of as an empirically derived table indexed by four independent variables: incidence angle, relative azimuth angle between the wind direction and the radar antenna, wind speed, and polarization. There is a fifth independent variable, frequency. Since current scatterometers operate at a single frequency a separate table is constructed for each desired operating frequency (14 GHz for NSCAT and 5.6 GHz for ERS-1). The complexity involved in deriving the GMF makes assessment of GMF uncertainty difficult.

As a result, we will assume a simple model for GMF uncertainty and determine the effect on the overall σ° measurement model. Regardless of the experiments used to derive the GMF, it will always maintain a certain degree of inaccuracy since wind velocity is not the only geophysical parameter affecting σ° of the ocean surface. Other factors can feasibly modify σ° of the ocean surface: long-waves, distance from land, sea-surface temperature, and salinity are a few parameters that can also influence σ°. The current GMF ignores all of these factors and assumes that wind velocity is the only source of σ° changes.

To account for the inaccuracy of this presumption one method is to treat the true σ° of the ocean surface as a random variable, σ_t°, with mean given by the GMF and variance given by a table similar to the GMF. In principle this table could be constructed empirically using the data collected to form the GMF. It could be constructed using the additional variability of σ° measurements for the same wind velocity not accounted for by instrument noise and background radiation. In general, this would be a function of the same parameters as the GMF. To obtain a general idea of how uncertainty in σ_t° translates to uncertainty in wind estimates, we deal only with the variance of σ_t° and assume σ_t° is a Gaussian random variable.

Since multiple measurements must be used to estimate wind velocity, it is more realistic to consider σ_t° as a random process with dependence on radar measurement angles and polarization. This allows consideration of the correlation between σ_t° from the different sets of radar angles and polarization that are used to retrieve wind. However, no prior correlation information on σ_t° is available, and it is not the purpose of this thesis to construct such information; therefore, we assume that the σ_t° random process is completely uncorrelated. Given the prior Gaussian assumption, we effectively assume that different realizations of σ_t° are independent for the purposes of estimating the wind. As a result, we only need to consider σ_t° as a random variable.

To clarify the description given in the preceding paragraphs, let $\mathcal{M}(\theta, \psi - \phi, U, p)$ represent the GMF for a given frequency, incidence angle, relative azimuth angle, wind speed, and polarization. Associated with the GMF is a hypothetical variance of the GMF denoted, $\mathcal{V}(\theta, \psi - \phi, U, p)$ which accounts for the model function uncertainty. Then, for a particular set of θ, χ, U, and p, σ_t° is modeled as a Gaussian random variable with mean $\mathcal{M}$ and variance $\mathcal{V}$.

It is customary in scatterometry to define the normalized standard deviation, K_p of a random variable. To distinguish it from other noise sources, the K_p of σ_t° is called

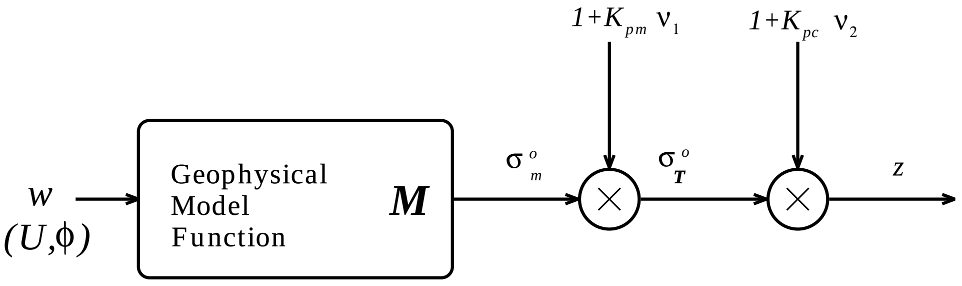


Figure 3.1: *Model for σ° measurements used in this thesis.*

K_{pm}, and defined as

$$K_{pm} = \frac{\sqrt{\mathcal{V}}}{\mathcal{M}}. \tag{3.1}$$

Then σ°_t can be written as

$$\sigma^\circ_t = \mathcal{M}(1 + K_{pm} v_1), \tag{3.2}$$

where v_1 is a zero-mean unit variance Gaussian random variable. Note that K_{pm} is generally a function of $\theta, \psi - \phi, U,$ and p, as well as frequency. Using this notation, the probability density function of σ°_t (the true normalized radar cross section of the ocean surface) can be written for a fixed wind velocity, w, as

$$p_{\sigma^\circ_t}(\sigma^\circ_t|\mathrm{w}) = \frac{1}{\sqrt{2\pi\mathcal{M}^2 K^2_{pm}}} \exp\left[-\frac{(\sigma^\circ_t - \mathcal{M})^2}{2\mathcal{M}^2 K^2_{pm}}\right]. \tag{3.3}$$

3.1.2 Effect of GMF uncertainty on overall measurement model

With a model for uncertainty in the GMF, the effect on the overall measurement model can be explored. Consider Figure 3.1 which is a simplified block-diagram similar to Figure 2.5 but showing only communication noise and GMF uncertainty. From this figure we see that

$$z = \mathcal{M}(1 + K_{pm} v_1)(1 + K_{pc} v_2) = \sigma^\circ_t(1 + K_{pc} v_2), \tag{3.4}$$

where K_{pc} is given by Eq. (2.9) with $\mathcal{M}$ replaced by σ°_t:

$$K_{pc} = \sqrt{\alpha + \frac{\beta}{\sigma^\circ_t} + \frac{\gamma}{\sigma^{\circ 2}_t}}. \tag{3.5}$$

Now z is the product of two normal random variables which are not completely independent since K_{pc} depends on σ_t°. As a result, the density function of z is difficult to derive directly. However, using conditional probability we can formulate an integral expression for the density function of z:

$$\begin{aligned} p_z(z|\mathbf{w}) &= \int_{\sigma_t^\circ} p_{z,\sigma_t^\circ}\left(z, \sigma_t^\circ|\mathbf{w}\right) d\sigma_t^\circ\,, \\ &= \int_{\sigma_t^\circ} p_{z|\sigma_t^\circ}\left(z|\sigma_t^\circ\right) p_{\sigma_t^\circ}\left(\sigma_t^\circ|\mathbf{w}\right) \\ p_z(z|\mathbf{w}) &= \int_{\sigma_t^\circ} \frac{\exp\left[-\frac{(z-\sigma_t^\circ)^2}{2(\alpha\sigma_t^{\circ 2}+\beta\sigma_t^\circ+\gamma)}\right]}{\sqrt{2\pi(\alpha\sigma_t^{\circ 2}+\beta\sigma_t^\circ+\gamma)}} \frac{\exp\left[-\frac{(\sigma_t^\circ-\mathcal{M})^2}{2\mathcal{M}^2K_{pm}^2}\right]}{\sqrt{2\pi\mathcal{M}^2K_{pm}^2}} d\sigma_t^\circ \end{aligned} \tag{3.6}$$

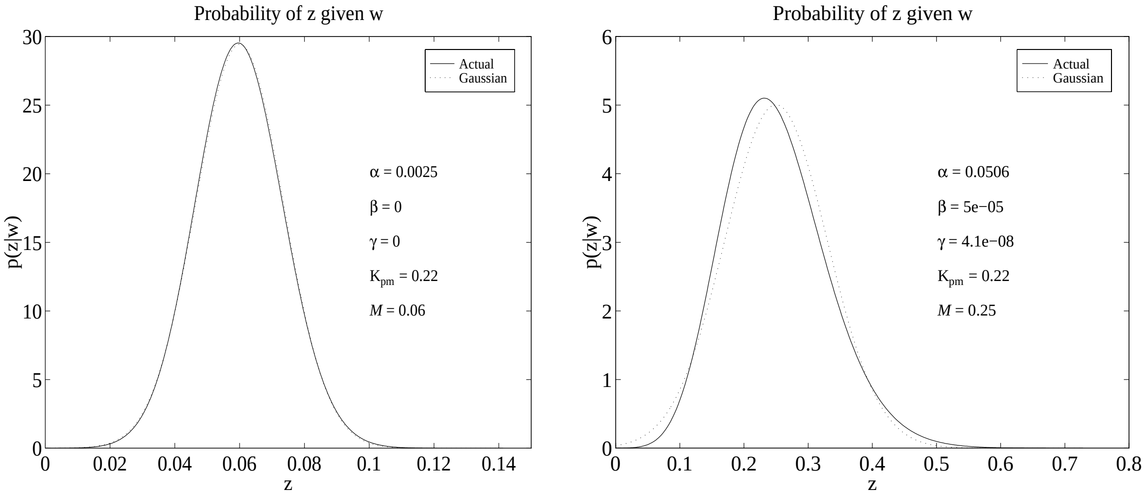
It is evident that this density function is not Gaussian. To see its shape, representative plots are shown in Figure 3.2 for particular values of $\mathcal{M}$, K_{pm}, α, β, and γ. For all but K_{pm} these values were chosen to correspond to actual ERS-1 data and expected NSCAT data. The value for K_{pm} was chosen based on research done by P. Johnson (1996) in estimating K_{pm} for C-band (ERS-1). For comparison, a plot of a true Gaussian density with the same mean and variance is shown as well. Although not exactly Gaussian, the distribution is near-Gaussian. When α is small, as for ERS-1 data, the distribution is almost exactly Gaussian. When α is larger, as for NSCAT data, the distribution is skewed to the left.

While there is not a simple closed form expression for Eq. (3.6), we can obtain a useful closed expression for $p(z|\mathbf{w})$ if we make one very reasonable assumption. Note that K_{pc} is only dependent on σ_t° when β and γ are non-zero. In addition, β and γ are usually very close to zero and much smaller than α. In fact, simply setting them to zero has no visually perceivable effect on the density shown in Figure 3.2(b). As a result, K_{pc} can be made independent of σ_t° with little effect on the model by using the mean of σ_t° instead of σ_t° in the equation for K_{pc}:

$$K_{pc} = \sqrt{\alpha + \frac{\beta}{\mathcal{M}} + \frac{\gamma}{\mathcal{M}^2}}. \tag{3.7}$$

Now, K_{pc} is no longer a function of σ_t° so that z is the product of two *independent* normal random variables.

While this assumption still does not make the density function easy to express, it does allow straight-forward computation of the central moments. This ultimately allows



(a) ERS-1 parameters

(b) NSCAT parameters

Figure 3.2: *Probability of z given* **w** *compared with a true Gaussian of the same mean and variance. The two curves in (a) are nearly identical.*

an expression for $p_z(z|\mathbf{w})$ to be written using an expansion in Hermite polynomials whose coefficients are determined by these central moments. This is a useful expansion technique for approximating near-Gaussian distributions and is outlined in Appendix A.

The result is

$$\begin{aligned} p(z|\mathbf{w}) &= \frac{e^{-\frac{(z-\mu)^2}{2\mu_2}}}{\sqrt{2\pi\mu_2}}\left[1+\sum_{k=3}^{\infty} C_k H_k\left(\frac{z-\mu}{\sqrt{2\mu_2}}\right)\right], && (3.8)\\ C_k &= \sum_{n=0}^{\lfloor\frac{k-3}{2}\rfloor} \frac{(-1)^n\left(\mu_{k-2n}-\mu^N_{k-2n}\right)}{n!(k-2n)!\mu_2^{k/2-n}2^{k/2+n}}. && (3.9)\end{aligned}$$

In these equations, μ_m is the mth central moment of $p_z(z|\mathbf{w})$, μ is the mean of $p_z(z|\mathbf{w})$, and μ^N_m is the mth moment of a zero-mean Gaussian with variance μ_2. It is well known that $\mu^N_m = 0$ for m odd and $\mu^N_m = (m-1)!!\mu_2^m$ for m even. We now derive an expression for μ_m assuming K_{pc} is independent of σ°_t, but first we need μ.

$$\begin{aligned} \mu &= E[z] = E[\mathcal{M}(1+K_{pm}v_1)(1+K_{pc}v_2)],\\ &= \mathcal{M}E[1+K_{pm}v_1+K_{pc}v_2+K_{pm}K_{pc}v_1v_2]\\ &= \mathcal{M}(1+K_{pm}E[v_1]+K_{pc}E[v_2]+K_{pm}K_{pc}E[v_1]E[v_2])\\ &= \mathcal{M}. && (3.10)\end{aligned}$$

Now the central moments can be computed using iterated conditional expectations and the binomial expansion:

$$\begin{aligned}
\mu_m &= E\left[(z-\mathcal{M})^m\right], \\
&= E\left\{E\left[(z-\sigma_t^\circ+\sigma_t^\circ+\mathcal{M})^m|\sigma_t^\circ\right]\right\}, \\
&= E\left\{E\left[\sum_{k=0}^{m}\binom{m}{k}(z-\sigma_t^\circ)^k|\sigma_t^\circ\right](\sigma_t^\circ-\mathcal{M})^{m-k}\right\}, \\
&= \sum_{\substack{k=0\\ k \text{ even}}}^{m}\binom{m}{k}(k-1)!!K_{pc}^{k}E\left[\sigma_t^{\circ k}(\sigma_t^\circ-\mathcal{M})^{m-k}\right], \\
&= \sum_{\substack{k=0\\ k \text{ even}}}^{m}\binom{m}{k}(k-1)!!K_{pc}^{k}E\left[(\sigma_t^\circ-\mathcal{M}+\mathcal{M})^{k}(\sigma_t^\circ-\mathcal{M})^{m-k}\right], \\
&= \sum_{\substack{k=0\\ k \text{ even}}}^{m}\binom{m}{k}(k-1)!!K_{pc}^{k}E\left[\sum_{l=0}^{k}\binom{k}{l}\mathcal{M}^{l}(\sigma_t^\circ-\mathcal{M})^{k-l}(\sigma_t^\circ-\mathcal{M})^{m-k}\right], \\
&= \sum_{\substack{k=0\\ k \text{ even}}}^{m}\sum_{l=0}^{k}\binom{m}{k}\binom{k}{l}\mathcal{M}^{l}(k-1)!!K_{pc}^{k}E\left[(\sigma_t^\circ-\mathcal{M})^{m-l}\right], \\
&= \sum_{\substack{k=0\\ k \text{ even}}}^{m}\sum_{\substack{l=0\\ m\text{-}l \text{ even}}}^{k}\binom{m}{k}\binom{k}{l}\mathcal{M}^{l}(k-1)!!K_{pc}^{k}(m-l-1)!!K_{pm}^{m-l}\mathcal{M}^{m-l}, \\
&= \mathcal{M}^{m}\sum_{\substack{k=0\\ k \text{ even}}}^{m}\sum_{\substack{l=0\\ m\text{-}l \text{ even}}}^{k}\binom{m}{k}\binom{k}{l}(k-1)!!(m-l-1)!!K_{pc}^{k}K_{pm}^{m-l}.
\end{aligned} \tag{3.11}$$

In these equations,

$$n!! = \begin{cases} 1\times 3\times\cdots\times n & n\geq 1 \ (n \text{ odd}) \\ 1 & n=-1. \end{cases} \tag{3.12}$$

This compact expression for the central moments can be rewritten for odd m and even m:

$$\begin{aligned}
\mu_{m \text{ odd}} &= \mathcal{M}^{m}\sum_{k=0}^{\frac{m-1}{2}}\sum_{l=0}^{k-1}\binom{m}{2k}\binom{2k}{2l+1}(2k-1)!!(m-2l-2)!!K_{pc}^{2k}K_{pm}^{m-2l-1}, \\
\mu_{m \text{ even}} &= \mathcal{M}^{m}\sum_{k=0}^{\frac{m}{2}}\sum_{l=0}^{k}\binom{m}{2k}\binom{2k}{2l}(2k-1)!!(m-2l-1)!!K_{pc}^{2k}K_{pm}^{m-2l}.
\end{aligned} \tag{3.13}$$

While it is useful to have a complete expression for the probability distribution $p_z(z|\mathbf{w})$ given by Eq. (3.8), usually only the first few moments are needed to approximate the true distribution accurately. As a result, the first four central moments are explicitly

m	μ_m
2	$\mathcal{M}^2(K_{pc}^2 + K_{pm}^2 + K_{pc}^2 K_{pm}^2)$
3	$6\mathcal{M}^3 K_{pc}^2 K_{pm}^2$
4	$3\mathcal{M}^4 \left(K_{pc}^4 + K_{pm}^4 + 3K_{pc}^4 K_{pm}^4 + 6K_{pc}^4 K_{pm}^2 + 6K_{pc}^2 K_{pm}^4 + 2K_{pc}^2 K_{pm}^2\right)$
5	$60\mathcal{M}^5 K_{pc}^2 K_{pm}^2 \left(K_{pc}^2 + K_{pm}^2 + 3K_{pc}^2 K_{pm}^2\right)$

Table 3.1: *Central moments, μ_m, for $m = 2 \ldots 5$ for the expanded probability model of the measurements z given the wind $\mathbf{w}$.*

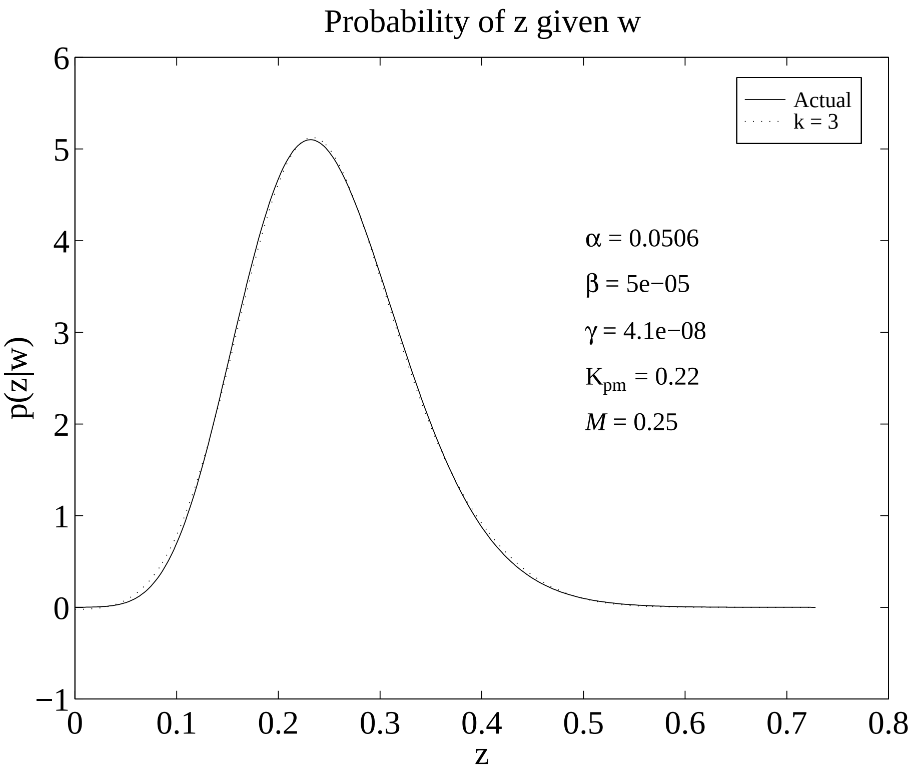


Figure 3.3: *Comparison of actual density function to a Gaussian with a first order correction factor.*

listed in Table 3.1. Figure 3.3 shows how well the probability density function in Eq. (3.8) fits the actual density function when using only a third-moment correction factor.

Given the preceding analysis, which describes in detail the density function of the measurements z, a question that begs response arises: ‘‘How should we apply this distribution to wind estimation?’’ It seems overzealous to use the exact density function given by Eqs. (3.6) or (3.8). Not only are these equations unwieldy but they are obtained by simplistically assuming Gaussian statistics for σ_t°. As a result, it is reasonable to use only the the first two moments of the resulting distribution and assume that the measurements are Gaussian-distributed with mean $\mathcal{M}$ and variance $\mu_2 = \mathcal{M}^2(K_{pc}^2 + K_{pm}^2 + K_{pc}^2 K_{pm}^2)$.

For ERS-1 data, where K_{pc} $(= \sqrt{\alpha})$ is small, the exact distribution is so close to Gaussian that it is reasonable to assume it is Gaussian. For NSCAT data where K_{pc} is not as small, the distribution has a high enough third moment to distinguish itself visually from a true Gaussian. In the next section we investigate the effect of this difference on wind estimation.

3.1.3 Effect on Wind Estimation

The effect on wind estimation for NSCAT data of using a Gaussian model for the measurements when they are not truly Gaussian can be explored through simulations. The wind is retrieved using two different probability models to observe if the different retrieval methods have any effect on the speed and direction distribution of the estimates.

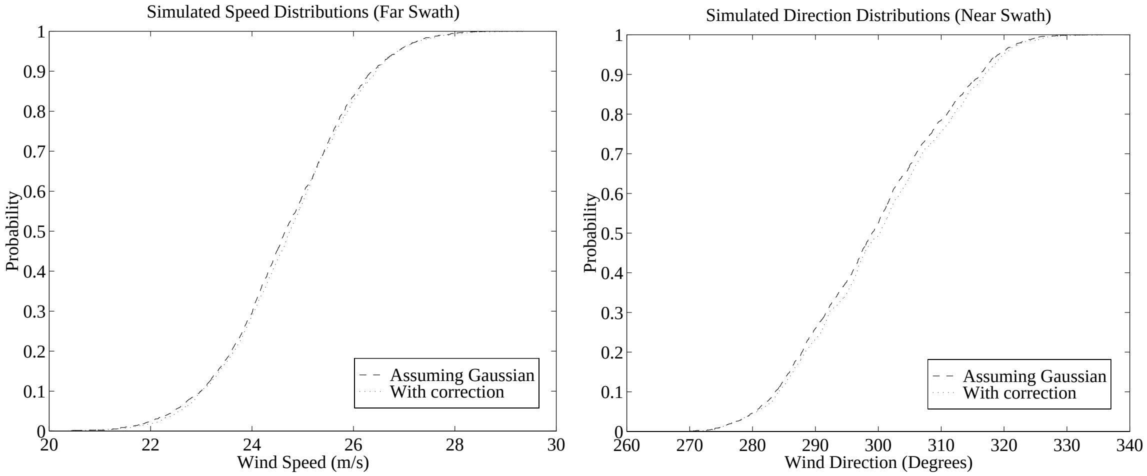
The first probability model used in retrieving wind is a strictly Gaussian model with variance $\mu_2 = \mathcal{M}^2(K_{pc}^2 + K_{pm}^2 + K_{pc}^2 K_{pm}^2)$. The second is a near-Gaussian model with a third moment correction using (3.8):

$$p(z|\mathbf{w}) \approx \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{(z-\mathcal{M})^2}{2\mu_2}} \left[1 + \frac{\mathcal{M}^3 K_{pc}^2 K_{pm}^2}{2\sqrt{2}\mu_2^{3/2}} H_3\left(\frac{z-\mathcal{M}}{\sqrt{2\mu_2}}\right)\right] \tag{3.14}$$

where $H_3(x) = 8x^3 - 12x$. In both of these models, the multiple measurements used to retrieve the wind are assumed independent.

For the simulations in this sections, we chose a true wind vector and then constructed the measurements according to the multiplicative noise model given in (3.4). A presumably large value of K_{pm} $(= 0.20)$ was chosen and values of K_{pc} were computed based on noise information supplied with the simulated NSCAT data.

Three cell locations in an NSCAT swath were chosen: one near the subsatellite track, one mid-way in the swath, and one far from the subsatellite track. In each cell, eighteen simulations were performed with 2000 retrievals per simulation. The eighteen different true wind vectors chosen for the simulations comprised the product space of three speeds (5,15,25 m/s) and six directions $(0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ)$. With the data from each simulation, two sample distributions were constructed, one for wind retrieved using the Gaussian probability model and one for wind retrieved using the corrected model.



(a) Speed Distribution

(b) Direction Distribution

Figure 3.4: *Distributions for wind retrieved using a Gaussian model and a third-moment correction model.*

In every case tested, the two distributions matched exceptionally well. In fact, Figures 3.4(a) and 3.4(b) present the worst cases encountered for the speed and direction distributions respectively. These figures were chosen based on visual inspection of all of the calculated distributions.

Under worst-case scenarios, which these figures represent, the effect of assuming a Gaussian model to retrieve the wind is to introduce a small bias into the retrieval. When compared to the overall variance of the simulated distributions, this bias is quite small. As a result, we conclude this investigation by stating that there is almost no important effect on wind retrieval of using a Gaussian model instead of the true model derived by assuming independent, Gaussian noise in the GMF.

3.1.4 Final measurement model

Since the effect of the non-Gaussian PDF on wind estimation is small, and it is much easier to deal with Gaussian statistics, the measurement model used in the rest of this paper is Gaussian. In addition, we assume that the multiple measurements needed for wind estimation are independent[1]. As a result, the joint density of the measurements used

---

[1] This assumption is probably not a good one for uncertainty due to geophysical modeling error. Further research needs to be done in this area to determine the effect on wind retrieval of correlation of GMF

for wind retrieval is

$$
\begin{aligned}
p_{\mathbf{z}}(\mathbf{z}|\mathbf{w}) &= \prod_{k=1}^{K} \frac{1}{\varsigma_{z_k}\sqrt{2\pi}} e^{\frac{-(z_k-\mathcal{M}_k)^2}{2\varsigma_{z_k}^2}}, && (3.15)\\
\varsigma_{z_k}^2 &= \mathcal{M}_k^2(K_{pc}^2 + K_{pm}^2 + K_{pc}^2 K_{pm}^2), && (3.16)\\
&= \epsilon_k \mathcal{M}_k^2 + (\beta_k \mathcal{M}_k + \gamma_k)\left(1 + K_{pm}^2\right), && (3.17)\\
\epsilon_k &= \alpha_k + K_{pm}^2 + \alpha_k K_{pm}^2. && (3.18)
\end{aligned}
$$

For model-based retrieval we write the density function of $p_{\mathbf{Z}}(\mathbf{Z}|\mathbf{W}) = p_{\mathbf{Z}}(\mathbf{Z}|\mathbf{X})$:

$$
\begin{aligned}
p_{\mathbf{Z}}(\mathbf{Z}|\mathbf{X}) &= \prod_{l=1}^{MN} \prod_{k=1}^{K(l)} \frac{1}{\sqrt{2\pi}\varsigma_{Z_{k,l}}} e^{\frac{-\left(Z_{k,l}-\mathcal{M}_{k,l}\right)^2}{2\varsigma_{Z_{k,l}}^2}}, && (3.19)\\
\varsigma_{Z_{k,l}}^2 &= \epsilon_{k,l} \mathcal{M}_{k,l}^2 + (\beta_{k,l} \mathcal{M}_{k,l} + \gamma_{k,l})\left(1 + K_{pm}^2\right), && (3.20)\\
\epsilon_{k,l} &= \alpha_{k,l} + K_{pm}^2 + \alpha_{k,l} K_{pm}^2. && (3.21)
\end{aligned}
$$

As only rough estimates of K_{pm}^2 are available, it is treated as a constant in this thesis. Notice that we assume $\mathbf{W} = F\mathbf{X} \equiv \mathbf{W}_F$ in the above noise model for wind field retrieval. As a result, the estimation model does not account for possible wind field modeling error. The effect of this omission on the results presented in this thesis is an area that should be addressed in future research.

The form of the probability density functions for both point-wise and model-based retrieval are essentially identical. As a result, statistical techniques applied to point-wise retrieval can be almost directly applied to model-based retrieval. The only modifications are due to the difference in what is being estimated: in point-wise retrieval a single wind vector is retrieved, while in model-based retrieval, the model parameters for a wind field are estimated from which the wind-field estimate is derived.

3.2 Application of new model

In the previous section an enhanced model of σ° measurements was presented by incorporating uncertainty in the GMF which relates σ° to wind velocity. Including this uncertainty gives more confidence in the measurement model from which error estimates on retrieved wind can be inferred.

The following two sections use the model developed in the previous section to discuss two different approaches for determining a measure of uncertainty in both

---

modeling error among the measurements used to retrieve wind point-wise and model-based retrieved winds. Both approaches attempt to approximate the covariance matrix of the wind estimate. The first approach is to assume that the wind estimator is efficient and use the Cramer-Rao bound as the covariance of the wind estimator. The second approach is to expand the implicit function that relates the σ° measurements and the wind estimate in a Taylor series and to use this series to approximate the covariance of the wind estimate. In both cases, we assume that the estimate is unbiased---taking advantage of this asymptotic property of maximum likelihood estimators. Although unbiasedness is not guaranteed for a finite number of measurements, simulations suggest that the maximum likelihood estimate is sufficiently unbiased for our purposes (Leotta and Long, 1989; however, see Appendix B).

With the mean and covariance of the wind estimate determined, a confidence interval can be developed using a Gaussian assumption with concentration ellipses. Two difficulties must be dealt with, however, in order to make sense of the results that follow.

Both methods of approximation give the mean and covariance of the wind estimate as a function of the true wind. Since the true wind is unknown during wind retrieval, one wonders if such a result is useful in reporting uncertainty in actual wind estimates. The best solution would be to obtain a set estimate of the wind in which we estimate a region for which we are $x\%$ sure that the true wind lies. This is a difficult problem. Alternatively, we could take a conservative approach in determining a confidence region for the wind estimate by varying the true wind around the calculated wind estimate and using the largest calculated covariance as the covariance for the concentration ellipse. In this thesis, however, a simpler approach is used. The covariance matrix of the wind estimate is calculated assuming the estimate is the true wind. This technique allows simple reporting of wind retrieval uncertainty for a particular wind estimate.

The other important difficulty with making covariance estimates of retrieved winds is the inherent multiple-solution problem of wind estimation. In general, after maximizing the likelihood function there is not a single wind estimate of the true wind. This makes it difficult to talk about the statistics of *the* wind estimate until after dealiasing has been accomplished. Trying to account for this dealiasing step in a complete statistical development of *the* wind estimate is formidable.

However, useful results are possible if we ignore the dealiasing step and focus attention on the wind aliases themselves[2]. One of the aliases has an expected value over

---

[2] A formal definition of wind aliases is deferred to Section 3.4.1.

the measurements which is the true wind. As a result, we determine the covariance of all of the aliases. Then, when one wind vector is chosen in the dealiasing step, a covariance will be associated with it. This covariance can be reported along with the retrieved wind as an estimate of wind retrieval error. While this does not convey the complete error picture since dealiasing error is ignored, it does provide a quantitative measure of how sensitive each alias is to the noisy measurements.

Most of the time treating each wind alias as random vector dependent on the measurements makes simple mathematical sense as the wind aliases usually correspond to well-separated maxima of the likelihood function. If the peaks are not well separated due to high noise levels (which can occur in point-wise wind retrieval at low wind speeds) it is usually because the measurements are not providing enough information to make a prediction of the wind direction.

With the understanding that we are investigating the statistics of each wind alias separately, we proceed to develop methods of approximating the covariance of these wind estimates.

3.3 Cramer-Rao Bound for Wind Scatterometry

In the previous section a model of σ° measurements was constructed that attempts to incorporate uncertainty in the GMF which maps wind velocity to σ°. With this uncertainty included we can with more confidence proceed to approximate the uncertainty in the wind estimate made from the σ° measurements. The first technique used is to apply the Cramer-Rao (C-R) bound as an approximation to the covariance matrix. Since the maximum likelihood estimator used in wind retrieval is asymptotically efficient, we expect that for a sufficiently large number of measurements the C-R bound will be equal to the covariance matrix of the wind estimates.

We first describe the C-R bound and give its general form. A derivation of the C-R bound for point-wise retrieval is then presented and used to obtain the model-based C-R bound. Finally selected results of the bound applied to both point-wise and model-based retrieval are presented for ERS-1 and NSCAT measurement geometries. Some of these results were presented previously by Oliphant and Long (1996a).

3.3.1 Description of Cramer-Rao Bound

The Cramer-Rao bound gives the lower bound on the error covariance matrix of any unbiased estimator. It depends on the probability model of the measurements and answers the question: ‘‘Given the uncertainty in the measurements, what is the least amount of uncertainty in any unbiased estimate made from the measurements?’’ This lower bound is the inverse of the Fisher information matrix, J, defined for the wind estimation problem as

$$J = E\left\{\left[\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial \mathbf{w}}\right]^T \frac{\partial L(\mathbf{w},\mathbf{z})}{\partial \mathbf{w}}\right\}, \tag{3.22}$$

where $\mathbf{w}$ can represent either the point-wise or model-based wind vector and expectation is taken over the measurements, $\mathbf{z}$. Here L is the log-likelihood function for either point-wise or model-based retrieval, and the derivative with respect to the wind vector is a row operator.

To describe the Cramer-Rao bound in more detail, let $\hat{\mathbf{w}}$ be an estimate of true wind $\mathbf{w}$ with

$$E[\hat{\mathbf{w}}] = \mathbf{w}, \tag{3.23}$$

then

$$C = E[(\hat{\mathrm{w}} - \mathrm{w})(\hat{\mathrm{w}} - \mathrm{w})^T] \geq J^{-1}. \tag{3.24}$$

3.3.2 Derivation for point-wise retrieval

For point-wise retrieval w is a 2-element vector and J is a 2×2 matrix. The diagonal elements of the inverse of J give the lower bound on the variance of the individual estimators of the elements of w assuming the estimator is unbiased[3]. This inverse can be expressed simply in terms of the elements of **J** resulting in an explicit expression for the lower bound on the covariance matrix of the point-wise estimate.

$$C \geq J^{-1} = \frac{1}{J_{11}J_{22} - J_{12}J_{21}} \begin{bmatrix} J_{22} & -J_{21} \\ -J_{12} & J_{11} \end{bmatrix}, \tag{3.25}$$

where $J_{12} = J_{21}$. We now derive an expression for J_{ij}.

For clarity, the ij element of the Fisher information matrix for point-wise retrieval is given as

$$J_{ij} = E\left[\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_i}\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_j}\right], \tag{3.26}$$

where the expectation operator is performed over the measurements z. Note that we can identify either $w_1 = U$ and $w_2 = \phi$ or $w_1 = u$ and $w_2 = v$.

Recalling the equation for the likelihood function $p_{\mathbf{z}}(\mathrm{z}|\mathrm{w})$ given by Eq. (3.15) we write the log-likelihood function as

$$L(\mathbf{w},\mathbf{z}) = -\sum_{k=1}^{K}\left[\frac{(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^2} + \frac{1}{2}\log\left(2\pi\varsigma_{z_k}^2\right)\right]. \tag{3.27}$$

For reference, the partials of the log-likelihood function can be explicitly written for use in the Fisher information matrix:

$$\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_1} = \sum_{k=1}^{K}\frac{\partial \mathcal{M}_k}{\partial w_1}\frac{z_k - \mathcal{M}_k}{\varsigma_{z_k}^2} + \left[\frac{(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^4} - \frac{1}{2\varsigma_{z_k}^2}\right]\frac{\partial \varsigma_{z_k}^2}{\partial w_1}, \tag{3.28}$$

$$\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_2} = \sum_{k=1}^{K}\frac{\partial \mathcal{M}_k}{\partial w_2}\frac{z_k - \mathcal{M}_k}{\varsigma_{z_k}^2} + \left[\frac{(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^4} - \frac{1}{2\varsigma_{z_k}^2}\right]\frac{\partial \varsigma_{z_k}^2}{\partial w_2}, \tag{3.29}$$

---

[3] Simulations and use of a maximum likelihood estimate of $\mathbf{w}$ suggest that $\hat{\mathbf{w}}$ is generally unbiased, (Later results suggest that at certain wind directions the estimate is somewhat biased. See Appendix B.)

where, using Eq. (3.17) and assuming K_{pm} is constant,

$$\begin{aligned}
\frac{\partial \varsigma_{z_k}^2}{\partial w_1} &= 2(\alpha_k + K_{pm}^2 + \alpha_k K_{pm}^2)\mathcal{M}_k \frac{\partial \mathcal{M}_k}{\partial w_1} + \beta_k (1 + K_{pm}^2) \frac{\partial \mathcal{M}_k}{\partial w_1}, \\
&= \left[2\epsilon_k \mathcal{M}_k + \beta_k (1 + K_{pm}^2) \right] \frac{\partial \mathcal{M}_k}{\partial w_1}, \qquad (3.30) \\
\frac{\partial \varsigma_{z_k}^2}{\partial w_2} &= 2(\alpha_k + K_{pm}^2 + \alpha_k K_{pm}^2)\mathcal{M}_k \frac{\partial \mathcal{M}_k}{\partial w_2} + \beta_k (1 + K_{pm}^2) \frac{\partial \mathcal{M}_k}{\partial w_2}, \\
&= \left[2\epsilon_k \mathcal{M}_k + \beta_k (1 + K_{pm}^2) \right] \frac{\partial \mathcal{M}_k}{\partial w_2}. \qquad (3.31)
\end{aligned}$$

The ij element of the Fisher matrix can be expanded as

$$\begin{aligned}
J_{ij} &= E\left\{ \sum_{k=1}^{K} \frac{\partial \mathcal{M}_k}{\partial w_i} \frac{z_k - \mathcal{M}_k}{\varsigma_{z_k}^2} + \left[\frac{(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^4} - \frac{1}{2\varsigma_{z_k}^2} \right] \frac{\partial \varsigma_{z_k}^2}{\partial w_i} \right\} \\
&\quad \times \left\{ \sum_{k=1}^{K} \frac{\partial \mathcal{M}_k}{\partial w_j} \frac{z_k - \mathcal{M}_k}{\varsigma_{z_k}^2} + \left[\frac{(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^4} - \frac{1}{2\varsigma_{z_k}^2} \right] \frac{\partial \varsigma_{z_k}^2}{\partial w_j} \right\}, \qquad (3.32) \\
&= \sum_{k=1}^{K} \sum_{l=1}^{K} E\left\{ \frac{\partial \mathcal{M}_k}{\partial w_i} \frac{z_k - \mathcal{M}_k}{\varsigma_{z_k}^2} + \left[\frac{(z_k - \mathcal{M}_k)^2}{2\varsigma_{z_k}^4} - \frac{1}{2\varsigma_{z_k}^2} \right] \frac{\partial \varsigma_{z_k}^2}{\partial w_i} \right\} \\
&\quad \times \left\{ \frac{\partial \mathcal{M}_l}{\partial w_j} \frac{z_l - \mathcal{M}_l}{\varsigma_{z_l}^2} + \left[\frac{(z_l - \mathcal{M}_l)^2}{2\varsigma_{z_l}^4} - \frac{1}{2\varsigma_{z_l}^2} \right] \frac{\partial \varsigma_{z_l}^2}{\partial w_j} \right\}. \qquad (3.33)
\end{aligned}$$

This expression can be simplified by using the fact that the measurements, z_k, are assumed independent. As a result,

$$E[g(z_k)h(z_l)] = E[g(z_k)]E[h(z_l)] \qquad k \neq l, \qquad (3.34)$$

for functions g and h. Using this it is clear that

$$\begin{aligned}
E[(z_k - \mathcal{M}_k)(z_l - \mathcal{M}_l)] &= \varsigma_{z_k}^2 \delta_{kl}, \\
E[(z_k - \mathcal{M}_k)] &= 0, \\
E[(z_l - \mathcal{M}_l)(z_k - \mathcal{M}_k)^2] &= 0, \\
E[(z_k - \mathcal{M}_k)^2(z_l - \mathcal{M}_l)^2] &= \varsigma_{z_k}^2 \varsigma_{z_l}^2 + 2\varsigma_{z_k}^4 \delta_{kl},
\end{aligned}$$

Making these substitutions allows simplification of the result:

$$\begin{aligned}
J_{ij} &= \sum_{k=1}^{K}\sum_{l=1}^{K}\left\{\frac{\partial\mathcal{M}_k}{\partial w_i}\frac{\varsigma_{z_k}^2\delta_{kl}}{\varsigma_{z_k}^2\varsigma_{z_l}^2}\frac{\partial\mathcal{M}_l}{\partial w_j}\right. \\
&\quad\left.+\frac{\partial\varsigma_{z_k}^2}{\partial w_i}\left[\frac{\varsigma_{z_k}^2\varsigma_{z_l}^2+2\varsigma_{z_k}^4\delta_{kl}}{4\varsigma_{z_k}^4\varsigma_{z_l}^4}-\frac{\varsigma_{z_k}^2}{4\varsigma_{z_k}^4\varsigma_{z_l}^2}-\frac{\varsigma_{z_l}^2}{4\varsigma_{z_k}^2\varsigma_{z_l}^4}+\frac{1}{4\varsigma_{z_k}^2\varsigma_{z_l}^2}\right]\frac{\partial\varsigma_{z_l}^2}{\partial w_j}\right\}, \\
&= \sum_{k=1}^{K}\sum_{l=1}^{K}\left[\frac{\partial\mathcal{M}_k}{\partial w_i}\frac{\delta_{kl}}{\varsigma_{z_l}^2}\frac{\partial\mathcal{M}_l}{\partial w_j}+\frac{\partial\varsigma_{z_k}^2}{\partial w_i}\frac{\delta_{kl}}{2\varsigma_{z_l}^4}\frac{\partial\varsigma_{z_l}^2}{\partial w_j}\right], \\
&= \sum_{k=1}^{K}\left[\frac{\partial\mathcal{M}_k}{\partial w_i}\frac{1}{\varsigma_{z_k}^2}\frac{\partial\mathcal{M}_k}{\partial w_j}+\frac{\partial\varsigma_{z_k}^2}{\partial w_i}\frac{1}{2\varsigma_{z_k}^4}\frac{\partial\varsigma_{z_k}^2}{\partial w_j}\right].
\end{aligned} \tag{3.35}$$

This expression can be further simplified by substituting from Eqs. (3.30) and (3.31)

$$\begin{aligned}
J_{ij} &= \sum_{k=1}^{K}\left[\frac{1}{\epsilon_k\mathcal{M}_k^2+(\beta_k\mathcal{M}_k+\gamma_k)(1+K_{pm}^2)}\right. \\
&\quad\left.+\frac{(2\epsilon_k\mathcal{M}_k+\beta_k(1+K_{pm}^2))^2}{2(\epsilon_k\mathcal{M}_k^2+(\beta_k\mathcal{M}_k+\gamma_k)(1+K_{pm}^2))^2}\right]\frac{\partial\mathcal{M}_k}{\partial w_i}\frac{\partial\mathcal{M}_k}{\partial w_j}.
\end{aligned} \tag{3.36}$$

It is interesting to note that when $\beta_k = \gamma_k = 0$ (the ERS-1 noise model) this expression simplifies even further to

$$J_{ij} = \sum_{k=1}^{K}\left[\frac{1}{\epsilon_k M_k^2}+\frac{2}{\mathcal{M}_k^2}\right]\frac{\partial\mathcal{M}_k}{\partial w_i}\frac{\partial\mathcal{M}_k}{\partial w_j}. \tag{3.37}$$

It is also useful to express the Fisher information matrix using matrix notation. To this end, note that the ij element of J has the general form

$$J_{ij} = \sum_{k=1}^{K} a_k b_{k,i} b_{k,j}. \tag{3.38}$$

Thus, J can be written as

$$J = G^T A G, \tag{3.39}$$

where A is a $K \times K$ diagonal matrix with elements

$$A_{kk} = \frac{1}{\epsilon_k\mathcal{M}_k^2+(\beta_k\mathcal{M}_k+\gamma_k)(1+K_{pm}^2)}+\frac{(2\epsilon_k\mathcal{M}_k+\beta_k(1+K_{pm}^2))^2}{2(\epsilon_k\mathcal{M}_k^2+(\beta_k\mathcal{M}_k+\gamma_k)(1+K_{pm}^2))^2}, \tag{3.40}$$

and G is a $K \times 2$ (row) derivative matrix:

$$G = \frac{\partial \mathcal{M}}{\partial \mathbf{w}} = \left\{ \frac{\partial \mathcal{M}_i}{\partial w_j} \right\}. \tag{3.41}$$

We can verify Equation (3.35) by comparing it with the derivation given by Scharf (1991) for a general multivariate normal model. In this model, $\mathbf{X}$ is a random sample consisting of M independent, N-dimensional, normal random vectors. The mean of each of the M random vectors is the $N \times 1$ vector $\mathbf{m}(\theta) \equiv \mathbf{m}$ where θ is the parameter vector. Similarly, $\mathbf{R}(\theta) \equiv \mathbf{R}$ is the $N \times N$ covariance matrix of each of the M random vectors. The joint distribution of the random sample can be written as

$$f_\theta(\mathbf{x}) = (2\pi)^{-MN/2} |\mathbf{R}|^{-M/2} \exp \left[-\frac{1}{2} \sum_{i=1}^{M} (\mathbf{x}_i - \mathbf{m})^T \mathbf{R}^{-1} (\mathbf{x}_i - \mathbf{m}) \right]. \tag{3.42}$$

When this is the joint distribution, the elements of the Fisher information matrix can be expressed are

$$J_{ij} = \frac{M}{2} \mathrm{tr} \left(\mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \theta_i} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \theta_j} \right) + M \frac{\partial \mathbf{m}^T}{\partial \theta_i} \mathbf{R}^{-1} \frac{\partial \mathbf{m}}{\partial \theta_j}. \tag{3.43}$$

In our case, the random sample $\mathbf{z}$ is the one measurement at K different azimuth and incidence looks by the scatterometer. As a result, $M = 1$. Using the measurement model, two identifications can be made:

$$\mathbf{m} = \begin{bmatrix} \mathcal{M}_1 \\ \vdots \\ \mathcal{M}_K \end{bmatrix} \tag{3.44}$$

$$\mathbf{R} = \begin{bmatrix} \varsigma_{z_1}^2 & & 0 \\ & \ddots & \\ 0 & & \varsigma_{z_K}^2 \end{bmatrix} \tag{3.45}$$

With these identifications, it can be shown that Equation (3.35) gives the same value for elements of the Fisher information matrix as Equation (3.43) taken from Scharf (1991).

This completes the derivation of the lower bound on the covariance of the point-wise estimate of the wind. We now extend this result to model-based wind retrieval and determine an estimate of the covariance of model-based wind estimates.

3.3.3 Covariance of model-based wind estimate

Before directly applying the Cramer-Rao bound to model-based wind estimates, we first take a closer look at the error covariance matrix for model-based retrieval.

$$C_{\mathbf{W}}^{e} = E\left[\hat{\mathbf{W}} - \mathbf{W}\right]\left[\hat{\mathbf{W}} - \mathbf{W}\right]^{T}. \tag{3.46}$$

Here, $\mathbf{W}$ is the true wind-field vector and $\hat{\mathbf{W}}$ is the estimate calculated by estimating the model-parameters, $\hat{\mathbf{X}}$, and mapping them through the model:

$$\hat{\mathbf{W}} = F\hat{\mathbf{X}}. \tag{3.47}$$

This error covariance matrix may be rewritten as

$$\begin{aligned} C_{\mathbf{W}}^{e} &= E\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}} + E\hat{\mathbf{W}} - \mathbf{W}\right]\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}} + E\hat{\mathbf{W}} - \mathbf{W}\right]^{T} \\ &= E\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}}\right]\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}}\right]^{T} + E\left[E\hat{\mathbf{W}} - \mathbf{W}\right]\left[E\hat{\mathbf{W}} - \mathbf{W}\right]^{T} \\ &= C_{\mathbf{W}} + C_{\mathbf{W}}^{F}, \end{aligned} \tag{3.48}$$

where

$$C_{\mathbf{W}} \equiv E\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}}\right]\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}}\right]^{T}, \tag{3.49}$$

$$C_{\mathbf{W}}^{F} \equiv E\left[E\hat{\mathbf{W}} - \mathbf{W}\right]\left[E\hat{\mathbf{W}} - \mathbf{W}\right]^{T}. \tag{3.50}$$

When written in this way it is evident that the total error in the model-based wind estimate is the sum of the covariance matrix of the wind estimate, $C_{\mathbf{W}}$, and the modeling error, $C_{\mathbf{W}}^{F}$. Assume $\hat{\mathbf{X}}$ is an unbiased estimate of the true model parameters:

$$E\hat{\mathbf{X}} = \mathbf{X} = F^{\dagger}\mathbf{W}, \tag{3.51}$$

and

$$E\hat{\mathbf{W}} = F\mathbf{X} = FF^{\dagger}\mathbf{W} = \mathbf{W}_F. \tag{3.52}$$

Consequently, we can rewrite the modeling error as

$$\begin{aligned} C_{\mathbf{W}}^{F} &= \left(FF^{\dagger} - I\right)\mathbf{W}\mathbf{W}^{T}\left(FF^{\dagger} - I\right)^{T}, \\ &= \left(FF^{\dagger} - I\right)\mathbf{W}\mathbf{W}^{T}\left(FF^{\dagger} - I\right). \end{aligned} \tag{3.53}$$

This error represents how accurately the model portrays real wind and is one of the considerations in developing a wind model. Since this section does not deal directly with wind modeling, we focus only on approximating $C_{\mathbf{W}}$ using the Cramer-Rao bound and recognize that the total error in model-based-retrieved winds includes some modeling error as well.

In order to use the Cramer-Rao bound to approximate the covariance of the wind estimate, it is important to recognize that the bound only directly applies to the parameter being estimated. In model-based wind retrieval the estimated object is a parameter vector, $\mathbf{X}$. As a result, to find an approximation to the covariance of the wind estimate, we first approximate the covariance of the model parameters using the Cramer-Rao bound and then use $\hat{\mathbf{W}} = F\hat{\mathbf{X}}$ to estimate the covariance of the wind-field vector:

$$\begin{aligned} C_{\mathbf{W}} &= E\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}}\right]\left[\hat{\mathbf{W}} - E\hat{\mathbf{W}}\right]^T, \\ &= E\left[F\hat{\mathbf{X}} - FE\hat{\mathbf{X}}\right]\left[F\hat{\mathbf{X}} - FE\hat{\mathbf{X}}\right]^T, \\ &= F\left[\hat{\mathbf{X}} - E\hat{\mathbf{X}}\right]\left[\hat{\mathbf{X}} - E\hat{\mathbf{X}}\right]^T F^T, \\ &= FC_{\mathbf{X}}F^T. \end{aligned} \tag{3.54}$$

Since $\hat{\mathbf{X}}$ is estimated using maximum likelihood, it is asymptotically unbiased with covariance equal to the Cramer-Rao bound. With the large number of measurements over the $M \times N$ region used to estimate $\hat{\mathbf{X}}$ it is reasonable to assume that $\hat{\mathbf{X}}$ is unbiased and its covariance can be well-approximated by the Cramer-Rao bound:

$$C_{\mathbf{X}} \approx J_{\mathbf{X}}^{-1}, \tag{3.55}$$

where $J_{\mathbf{X}}$ is the Fisher information matrix of the model parameters:

$$J_{\mathbf{X}} = E\left\{\left[\frac{\partial L(\mathbf{X},\mathbf{Z})}{\partial \mathbf{X}}\right]^T \frac{\partial L(\mathbf{X},\mathbf{Z})}{\partial \mathbf{X}}\right\}. \tag{3.56}$$

With this expression for the covariance on the model parameters, the covariance on the wind-field estimate can be well-approximated as

$$C_{\mathbf{W}} \approx FJ_{\mathbf{X}}^{-1}F^T. \tag{3.57}$$

To calculate the elements of $J_{\mathbf{X}}$ we recognize that the likelihood function for estimating $\mathbf{X}$ from $\mathbf{Z}$ is essentially identical to that used for point-wise retrieval in estimating w from z. Consequently, we can simply re-write Eq. (3.35) in model-based notation as the Fisher matrix for the estimated model parameters:

$$(J_{\mathbf{X}})_{ij} = \sum_{l=1}^{MN}\sum_{k=1}^{K(l)}\left[\frac{\partial \mathcal{M}_{k,l}}{\partial X_i}\frac{1}{\varsigma^2_{Z_{k,l}}}\frac{\partial \mathcal{M}_{k,l}}{\partial X_j} + \frac{\partial \varsigma^2_{Z_{k,l}}}{\partial X_i}\frac{1}{2\varsigma^4_{Z_{k,l}}}\frac{\partial \varsigma^2_{Z_{k,l}}}{\partial X_j}\right]. \tag{3.58}$$

The only difference between this equation and the corresponding one derived for the point-wise case is that the partial derivatives are taken with respect to the model parameters X_i and the double indexing specifies not only which measurement over a single cell but also which cell over a large region.

This expression can be further simplified as in the point-wise case by substituting from Eqs. (3.30) and (3.31)

$$(J_{\mathbf{X}})_{ij} = \sum_{l=1}^{MN} \sum_{k=1}^{K(l)} \left[\frac{1}{\epsilon_{k,l}\mathcal{M}_{k,l}^2 + (\beta_{k,l}\mathcal{M}_{k,l} + \gamma_{k,l})(1 + K_{pm}^2)} + \right. \left. \frac{(2\epsilon_{k,l}\mathcal{M}_{k,l} + \beta_{k,l}(1 + K_{pm}^2))^2}{2(\epsilon_{k,l}\mathcal{M}_{k,l}^2 + (\beta_{k,l}\mathcal{M}_{k,l} + \gamma_{k,l})(1 + K_{pm}^2))^2} \right] \frac{\partial \mathcal{M}_{k,l}}{\partial X_i} \frac{\partial \mathcal{M}_{k,l}}{\partial X_j} \tag{3.59}$$

It is again interesting to write the expression for $\beta = \gamma = 0$ (the ERS-1 noise model) which simplifies it even further to

$$(J_{\mathbf{X}})_{ij} = \sum_{l=1}^{MN} \sum_{k=1}^{K(l)} \left[\frac{1}{\epsilon_{k,l}\mathcal{M}_{k,l}^2} + \frac{2}{\mathcal{M}_{k,l}^2} \right] \frac{\partial \mathcal{M}_{k,l}}{\partial X_i} \frac{\partial \mathcal{M}_{k,l}}{\partial X_j}. \tag{3.60}$$

As with the point-wise expression, $J_{\mathbf{X}}$ can be written using matrix multiplication:

$$J_{\mathbf{X}} = G^T A G, \tag{3.61}$$

where A is an $L \times L$ diagonal matrix, with $L = \sum_l^{MN} K(l)$, having elements

$$A_{l'l'} = \frac{1}{\epsilon_{l'}\mathcal{M}_{l'}^2 + (\beta_{l'}\mathcal{M}_{l'} + \gamma_{l'})(1 + K_{pm}^2)} + \frac{(2\epsilon_{l'}\mathcal{M}_{l'} + \beta_{l'}(1 + K_{pm}^2))^2}{2(\epsilon_{l'}\mathcal{M}_{l'}^2 + (\beta_{l'}\mathcal{M}_{l'} + \gamma_{l'})(1 + K_{pm}^2))^2}. \tag{3.62}$$

Here l' is a new index that subsumes k and l. In addition, G is an $L \times M_x$ (row) derivative matrix:

$$G = \frac{\partial \mathcal{M}}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial \mathcal{M}_1}{\partial \mathbf{X}} \\ \vdots \\ \frac{\partial \mathcal{M}_{MN}}{\partial \mathbf{X}} \end{bmatrix}, \tag{3.63}$$

where $\partial \mathcal{M}_l / \partial \mathbf{X}$ is a $K(l) \times M_x$ sub-matrix containing the derivatives of the model function with respect to the model parameters for all the measurements in cell l.

These partial derivatives with respect to the model parameters can be calculated using the chain rule. First, note that $\mathcal{M}_l$ is only a function of the wind components in cell, l. Then, define w_l^p to be the wind vector in the lth lexicographical cell written in polar form (with ϕ measured in degrees clockwise from North). Similarly, we define $\mathbf{w}_l^r$ to be the wind vector written using rectangular coordinates. Then,

$$\frac{\partial \mathcal{M}_l}{\partial \mathbf{X}} = \frac{\partial \mathcal{M}_l}{\partial \mathbf{w}_l^p} \frac{\partial \mathbf{w}_l^p}{\partial \mathbf{w}_l^r} \frac{\partial \mathbf{w}_l^r}{\partial \mathbf{X}}, \tag{3.64}$$

where for clarity,

$$\frac{\partial \mathcal{M}_l}{\partial \mathbf{w}_l^p} = \begin{bmatrix} \frac{\partial \mathcal{M}_{1,l}}{\partial U_l} & \frac{\partial \mathcal{M}_{1,l}}{\partial \phi_l} \\ \vdots & \vdots \\ \frac{\partial \mathcal{M}_{K(l),l}}{\partial U_l} & \frac{\partial \mathcal{M}_{K(l),l}}{\partial \phi_l} \end{bmatrix}, \tag{3.65}$$

$$\frac{\partial \mathbf{w}_l^p}{\partial \mathbf{w}_l^r} = \begin{bmatrix} \frac{\partial U_l}{\partial u_l} & \frac{\partial U_l}{\partial v_l} \\ \frac{\partial \phi_l}{\partial u_l} & \frac{\partial \phi_l}{\partial v_l} \end{bmatrix} = \begin{bmatrix} \sin\left(\frac{\phi_l \pi}{180}\right) & \cos\left(\frac{\phi_l \pi}{180}\right) \\ \frac{180}{\pi U_l} \cos\left(\frac{\phi_l \pi}{180}\right) & -\frac{180}{\pi U_l} \sin\left(\frac{\phi_l \pi}{180}\right) \end{bmatrix}, \tag{3.66}$$

$$\frac{\partial \mathbf{w}_l^r}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial u_l}{\partial X_1} & \cdots & \frac{\partial u_l}{\partial X_{M_x}} \\ \frac{\partial v_l}{\partial X_1} & \cdots & \frac{\partial v_l}{\partial X_{M_x}} \end{bmatrix} = \begin{bmatrix} F_l \\ F_{l+MN} \end{bmatrix}, \tag{3.67}$$

where F_l is defined to be the lth row of F. Note that the last derivative is written assuming the common rectangular-component ordering of the individual wind vectors into the wind-field vector, $\mathbf{W}$, given by Eq. (2.22).

With this expression for the partial derivatives, the expression for G given by Eq. (3.63) can be rewritten as

$$G = \begin{bmatrix} \frac{\partial \mathcal{M}_1}{\partial \mathbf{w}_1^p} \frac{\partial \mathbf{w}_1^p}{\partial \mathbf{w}_1^r} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \frac{\partial \mathcal{M}_{MN}}{\partial \mathbf{w}_{MN}^p} \frac{\partial \mathbf{w}_{MN}^p}{\partial \mathbf{w}_{MN}^r} \end{bmatrix} \begin{bmatrix} F_1 \\ F_{1+MN} \\ \vdots \\ F_{MN} \\ F_{2MN} \end{bmatrix}, \tag{3.68}$$

$$= T \Pi F, \tag{3.69}$$

where T is the $L \times 2MN$ block-diagonal matrix of derivatives:

$$T = \begin{bmatrix} \frac{\partial \mathcal{M}_1}{\partial \mathbf{w}_1^p} \frac{\partial \mathbf{w}_1^p}{\partial \mathbf{w}_1^r} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \frac{\partial \mathcal{M}_{MN}}{\partial \mathbf{w}_{MN}^p} \frac{\partial \mathbf{w}_{MN}^p}{\partial \mathbf{w}_{MN}^r} \end{bmatrix}, \tag{3.70}$$

and Π is a permutation matrix calculated to interleave the rows of F as necessary. In particular, it is the $2MN \times 2MN$ identity matrix with row k permuted according to the map:

$$k \mapsto \begin{cases} 2k-1 & k \leq MN, \\ k-MN+1 & k > MN. \end{cases} \tag{3.71}$$

With these definitions, the Fisher information matrix for the model parameters can be written in matrix notation as

$$\begin{aligned} J_{\mathbf{X}} &= F^T \Pi^T T^T A T \Pi F, & (3.72) \\ &= F^T Q F, & (3.73) \end{aligned}$$

where $Q = \Pi^T T^T A T \Pi$ is $2MN \times 2MN$ and dependent on the GMF and its derivates calculated at $\mathbf{W}_F = F\mathbf{X}$.

With the Fisher information matrix of the model-parameters calculated, the lower bound on the covariance of the wind-field vector can be calculated as shown previously:

$$C_{\mathbf{W}} = F J_{\mathbf{X}}^{-1} F^T = F(F^T Q F)^{-1} F^T. \tag{3.74}$$

The rank of this covariance matrix is at most the rank of F which is less than $2MN$. This means that certain linear combinations of the individual wind-vector estimates comprising $\mathbf{W}$ are completely correlated. Most of the time, however, interest is only in the MN sub-matrices of $C_{\mathbf{W}}$ along the main diagonal which specify the covariance matrices of the individual wind vectors in the $M \times N$ region.

As the current wind field model is written in terms of the rectangular components of the wind, the covariance matrix is that of the rectangular components of the wind vector. In order to write the covariance for the wind-field vector in polar form, at least two simple methods are possible.

In the first approach realizations of a random vector, $\tilde{\mathbf{W}}_F$, with covariance matrix given by $C_{\mathbf{W}}$ and mean given by $\mathbf{W}_F$ corresponding to the cell in the region are generated as outlined. This random vector is then mapped to the polar version and the sample covariance calculated.

An alternative approach is to use a first-order Taylor series for calculating the wind-field vector in polar form, $\mathbf{W}_F^p$ in terms of the wind-field vector in rectangular form, $\mathbf{W}_F^r$. Then, the covariance of $\mathbf{W}_F^p$ can be calculated from $C_{\mathbf{W}}$ using

$$C_{\mathbf{W}^p} = \frac{\partial \mathbf{W}^p}{\partial \mathbf{W}^r} C_{\mathbf{W}} \frac{\partial \mathbf{W}^p}{\partial \mathbf{W}^r}^T, \tag{3.75}$$

where

$$\frac{\partial \mathbf{W}^p}{\partial \mathbf{W}^r} = \begin{bmatrix} \frac{\partial \mathbf{w}_1^p}{\partial \mathbf{w}_1^r} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \frac{\partial \mathbf{w}_{MN}^p}{\partial \mathbf{w}_{MN}^r} \end{bmatrix}. \tag{3.76}$$

Of course this approximation will be quite poor if the covariance is very large.

Using these expressions, the lower bound on the covariance of the model-based estimate can be calculated for the entire wind field or for each wind vector cell in the region in both (U, ϕ) and (u, v) coordinates. It should be noted again that since the wind-field model is not perfect, modeling error is also introduced into the model-based wind estimate. As a result the total error is the sum of the covariance calculated here and the modeling error shown earlier. Some idea of the magnitude of this bias can be obtained by fitting the model to point-wise retrieved winds and measuring the fit error (Long, 1989).

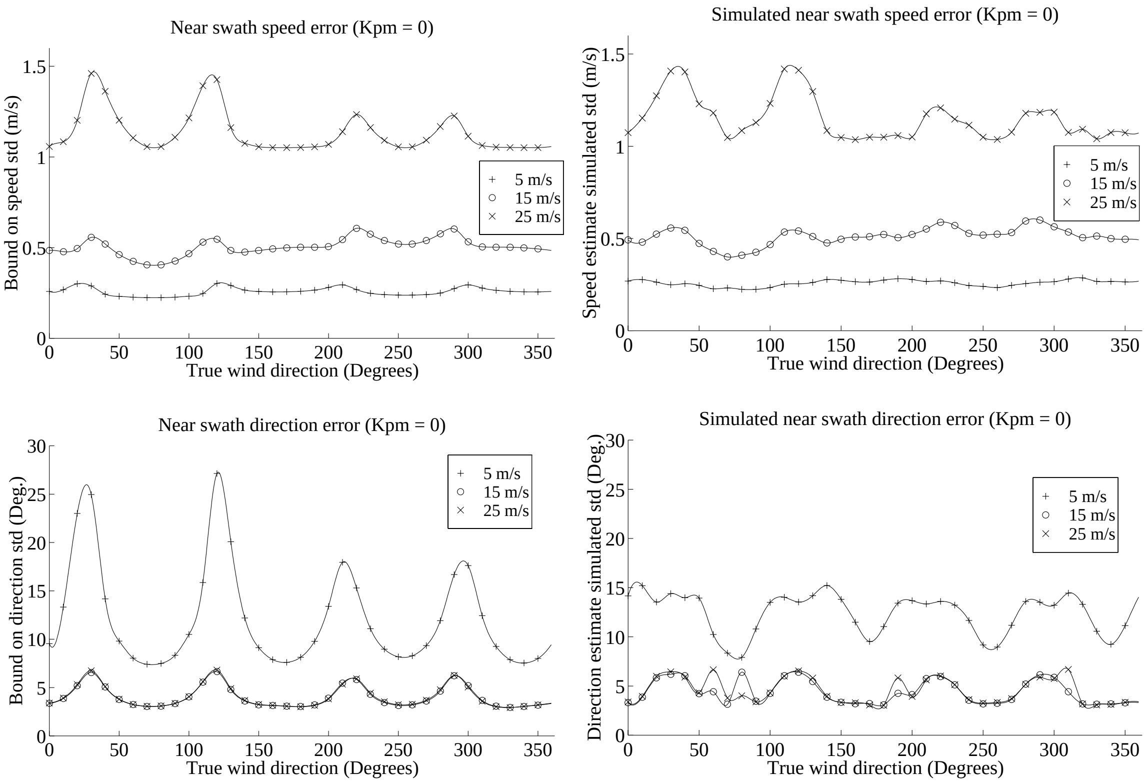
3.3.4 Point-wise Results

There are two key results with respect to the point-wise Cramer-Rao bound. The first result is that the wind estimate is nearly efficient for most true wind velocities so that the Cramer-Rao bound is a reasonable estimate of the covariance. The second result is that the covariance of the wind estimate grows particularly large for certain true wind directions. In the following we will demonstrate these two results for both ERS-1 and NSCAT geometries.

ERS-1

As a review of ERS-1 wind estimation geometry, for each wind vector cell three measurements are taken at azimuth angles 45° apart with the center beam at a lower incidence angle. At 25km resolution, there are 19 cross-track cells each with a different set of incidence angles for the measurement

In order to evaluate the performance of the wind retrieval estimator, we chose three cross-track locations which can be labeled far, mid, and near with respect to the subsatellite track. We evaluated the Cramer-Rao bound on the standard deviation of the speed and direction estimates at these locations for a variety of true wind speeds and directions. For these same true wind velocities we also performed a Monte Carlo simulation wherein σ° measurements are simulated according to the assumed model and



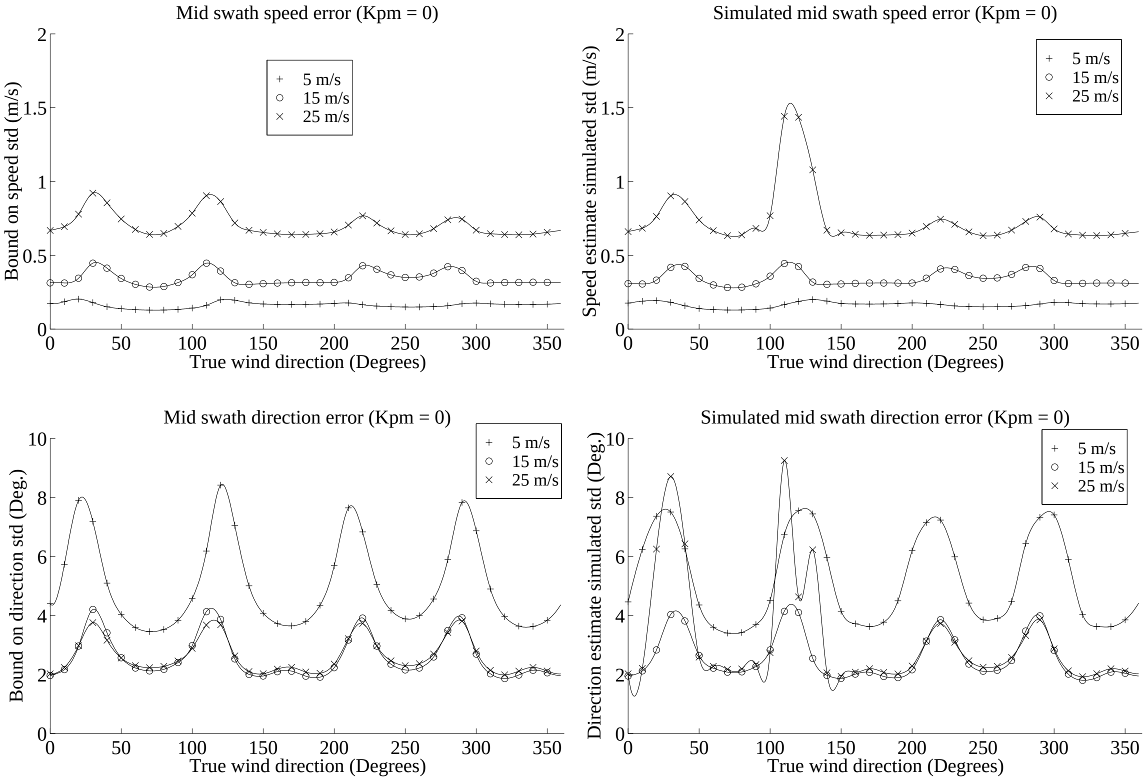
(a) C-R bound

(b) Simulations

Figure 3.5: *Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at near-track swath location (ERS-1).*

wind is retrieved using the maximum likelihood algorithm. The sample standard deviation for both the speed and direction was then computed so that it can be compared with the lower bound. Figures 3.5 to 3.7 show the results.

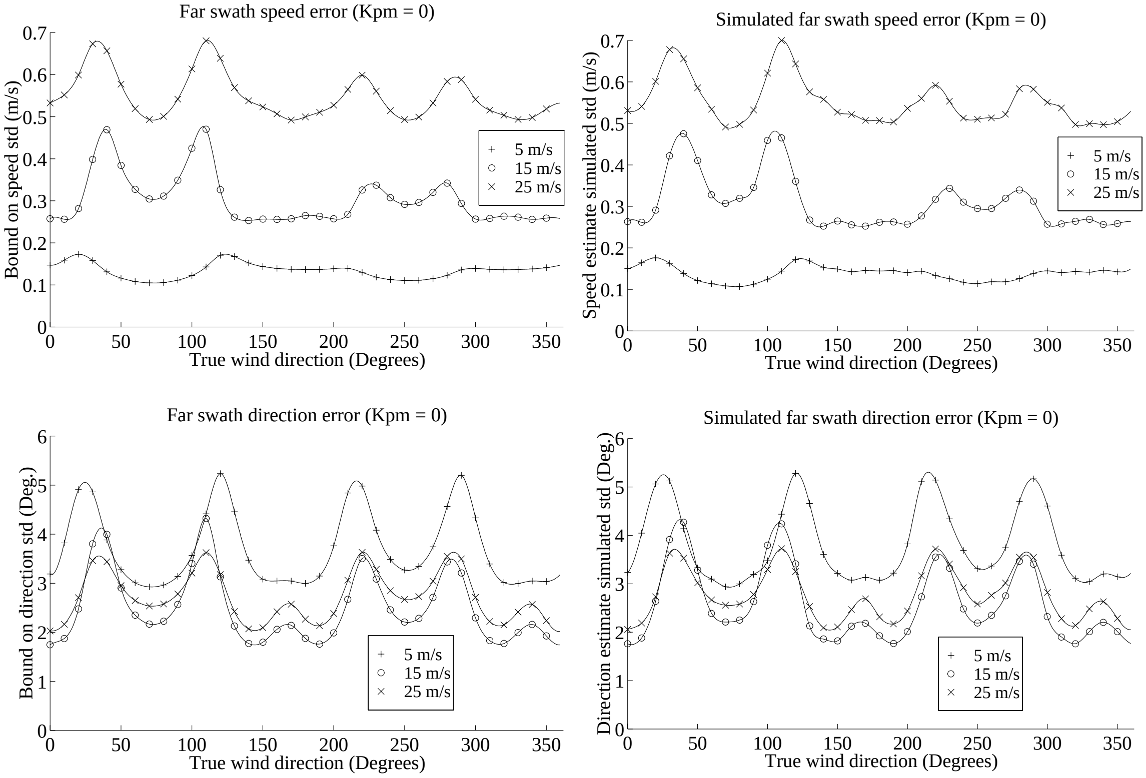
These figures all show that simulations are in good agreement with the Cramer-Rao bound indicating that the wind estimator is nearly efficient, especially at far swath. There is some discrepancy between simulations and the bound at mid-swath for 25 m/s winds which can be attributed to small sample size in the simulation-based variance estimate. The most notable discrepancy between the Cramer-Rao bound and the simulations, however, is that the significant peaks of large predicted uncertainty are ‘‘washed’’ out by the simulations. This is most notable at low wind speeds (5 m/s) for near swath where the Cramer-Rao bound predicts an extremely high standard deviation in the wind direction.



(a) C-R bound

(b) Simulations

Figure 3.6: *Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at mid-track swath location (ERS-1).*



(a) C-R bound

(b) Simulations

Figure 3.7: *Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at far-track swath location (ERS-1).*

One explanation for the discrepancy is premature termination of the algorithm for finding the maximum of the log-likelihood function. The reason the Cramer-Rao bound predicts a large variance in the estimate at certain places is that the log-likelihood function has small second derivatives and is extremely flat in a large neighborhood near the true local maximum. As a result, the wind retrieval algorithm may more easily select a near-maximum as the solution instead of the true maximum which can be some distance away.

Additional simulation studies show that this ‘‘washing’’ out effect in the simulations occurs whenever a large peak variance is predicted by the Cramer-Rao bound. All of this effect cannot be attributed to premature termination of the maximization algorithm. Another explanation is that since the wind estimator is outperforming the best unbiased estimate at certain true wind vectors, it must be a biased estimate at these wind vectors. This leads us to wonder whether or not the Cramer-Rao bound even applies. However, the figures show that for most values of true wind, the Cramer-Rao bound agrees well with simulations. As a result we conclude that the estimate is unbiased at most true wind vectors with a covariance approximately equal to the unbiased Cramer-Rao bound. In Appendix B we describe how to approximate the bias and then use the biased Cramer-Rao bound to approximate the covariance for these wind vectors where the unbiased Cramer-Rao bound approximation appears to over-predict the variance calculated from simulations.

Besides showing that the wind estimator is nearly efficient, Figures 3.5 to 3.7 also reveal interesting facts about the dependence of wind uncertainty on the true wind direction. The most prominent features on the plots are the large peaks at certain wind directions. These large peaks occur when the true direction is aligned with the azimuth direction of the fore or aft beam, with larger peaks occuring when the wind is pointing along the direction of the beam.

This result can be understood by recalling that the fore and aft beams are at the same incidence angle and separated by 90° in azimuth angle. In addition, upwind and crosswind relative azimuth angles give maxima and minima respectively in the GMF for a given wind speed. Consequently, for a fixed wind speed if the wind is aligned with the aft beam then aft will see a maximum σ° return while the fore beam will see a minimum.

The situation is reversed when we fix σ° and consider a plot of wind speeds that would give rise to that σ° measurement as a function of wind direction. This is illustrated in Figure 3.8 where the σ° is fixed and a plot of wind speeds and wind directions that could generate that σ° is graphed for each of the three beams. The true wind is 5 m/s and aligned

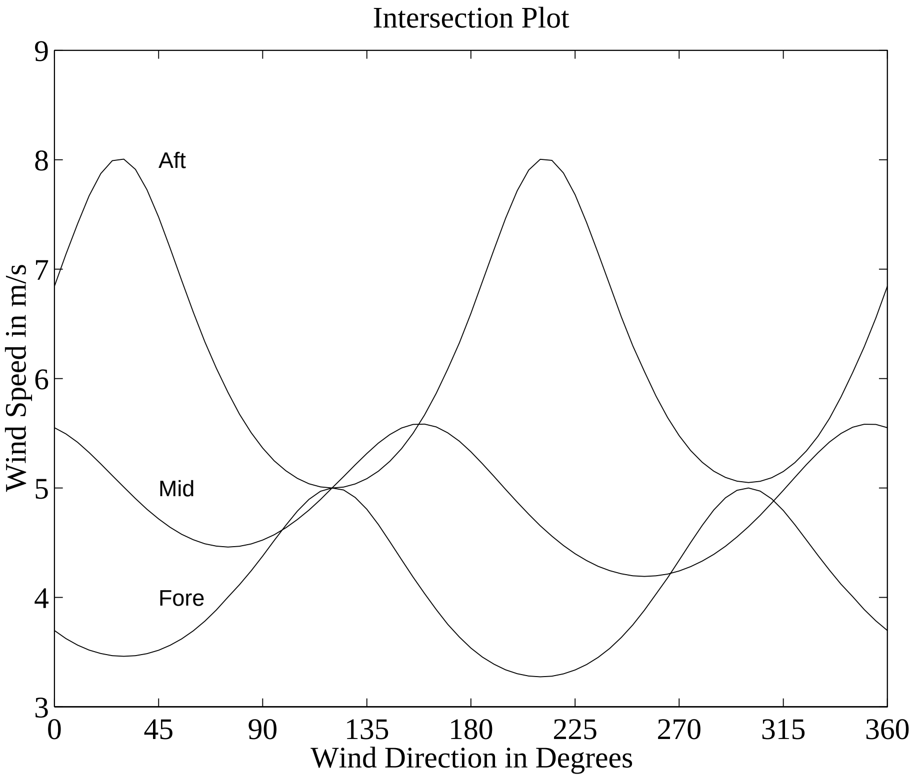


Figure 3.8: *Locus of wind vectors that would give rise to measurements obtained with no noise when the true wind is 5 m/s at 120°.*

with the aft beam at 120°.

According to the aft beam, a 120° wind direction gives a maximum σ° return for each wind speed. Thus when σ° is fixed, a smaller wind speed is necessary at 120° to generate the fixed σ° than for other wind directions. As a result, there is minimum for the aft beam at 120° in Figure 3.8. The reverse is true for the fore beam.

This reasoning only emphasizes the fact that the wind vector (given by the intersection of the three curves in the figure) is at point where the fore and aft beam curves are tangent. It is clear that adding even a small amount of noise to this problem (which has the effect of shifting the curves up and down) can wildly shift the direction value of the intersection point. As a result, the prediction of large variance at certain wind directions made by the Cramer-Rao bound can be interpreted physically as due to the GMF.

Other interesting trends are evident from Figures 3.5 to 3.7. One of these is the general result that wind estimate uncertainty generally decreases from near swath to

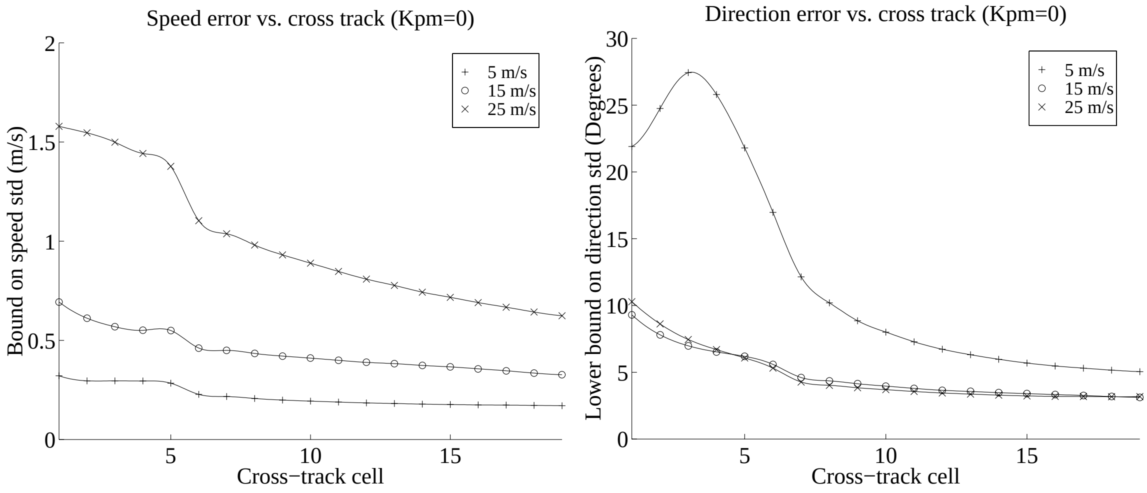


Figure 3.9: *ERS-1 dependence of wind estimate uncertainty on cell location within a swath.*

far swath, especially in the direction estimate. This is shown more clearly in Figure 3.9 which shows the lower bound on the standard deviation of both wind speed and wind direction estimates plotted against cross-track location for several wind speeds. The true wind direction corresponds to one of the peaks in the previous Figures.

This decreasing trend from near to far swath can be understood by noting that azimuth modulation, i.e. upwind/crosswind ratio, increases for increasing incidence angle, and far swath is measured with a greater incidence angle than near swath. As a result, there is more direction information in the measurements at far swath than at near swath.

So far, all of the results we have presented for the Cramer-Rao bound have assumed that there is no error in the GMF (i.e. $K_{pm} = 0$). As we have already discussed that this is unrealistic, we would like to observe the effect that K_{pm} has on the estimates of standard deviation. To do this we examine the worst-case wind direction (aligned with the aft beam) for a far-track swath. Refer to Figure 3.10. Note that the uncertainty in the wind estimate is quite sensitive to K_{pm} especially at low wind speeds. This can be understood in light of the large transmit power (and therefore small K_{pc}) of the scatterometer on ERS-1. Since very little noise is assumed in the first place, increasing it changes the estimation situation considerably.

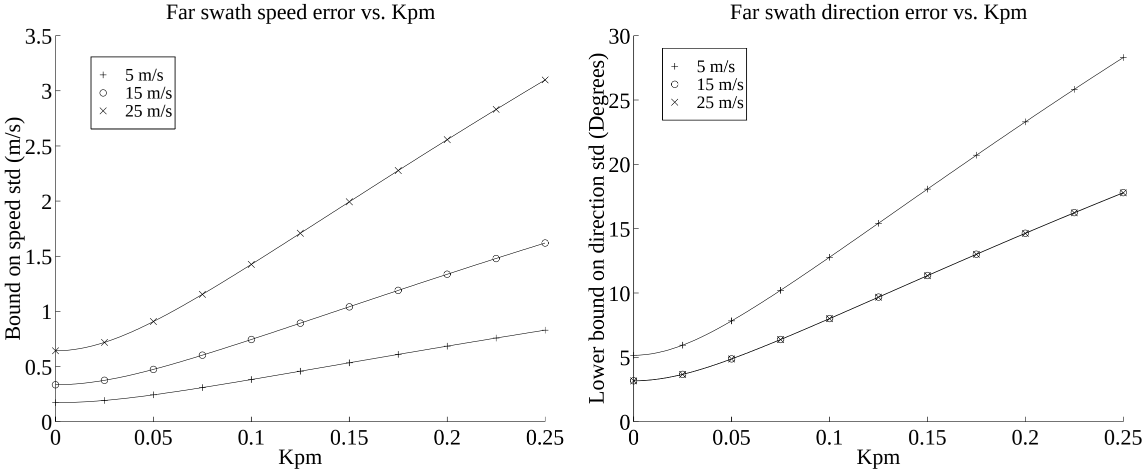


Figure 3.10: *ERS-1 dependence of wind estimate uncertainty on K_{pm} for far-swath.*

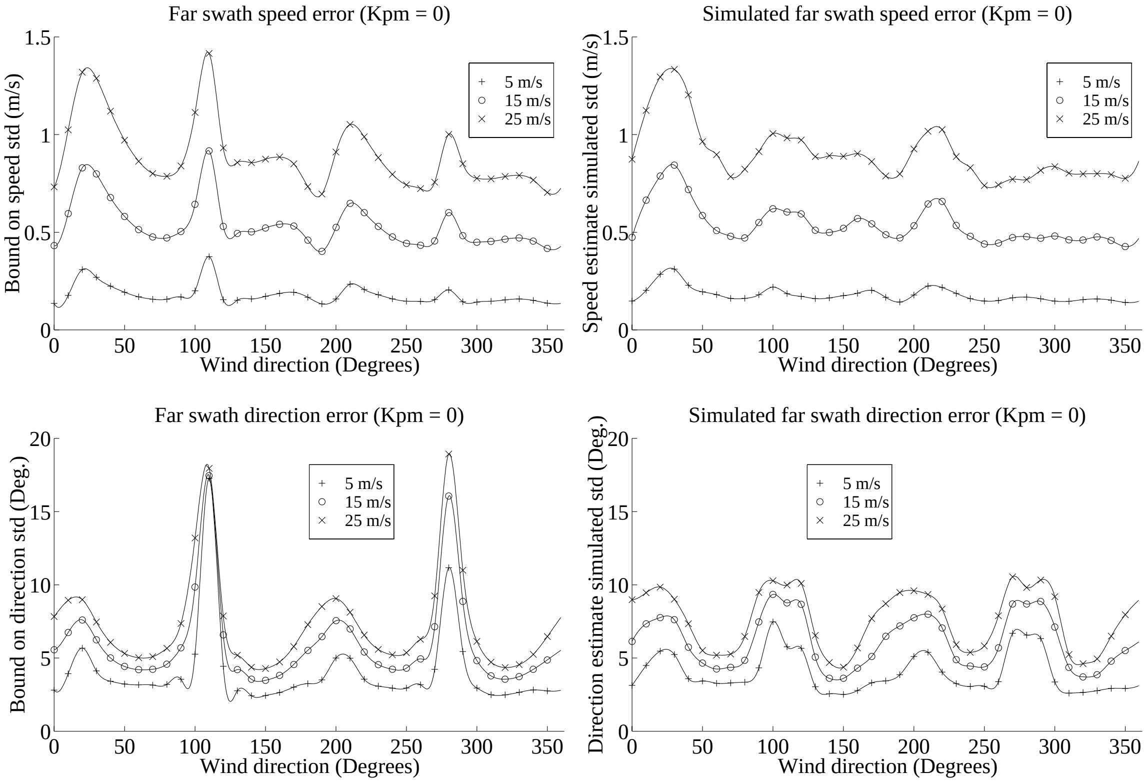
NSCAT

In order to understand estimates of retrieved wind uncertainty given by the Cramer-Rao bound for NSCAT, it is helpful to briefly review NSCAT measurement geometry. At 50 km resolution there are left and right swaths each containing 12 cross-track cells in which wind is estimated. There are nominally sixteen measurements for each cell (four from each beam). For each swath, two beams are separated by 90° in azimuth while a third beam is separated by 20° from one of the first two. The third beam is also at a lower incidence angle and transmits/receives both vertical and horizontal polarization.

In the following figures (3.11 to 3.13) we repeat several of the same figures that we constructed for the ERS-1 geometry but with the NSCAT geometry. Several of the same observations can be made as were made for ERS-1. For example the same ‘‘smoothing’’ of the predicted variance peak by the simulations can be seen. There are some notable differences, however. One difference is that due to the asymmetrical distribution of the beams, there is one very large variance peak when the true direction is aligned with the fore or aft beam depending on which one the center beam is closest to (different for left and right swaths). Another notable difference is that wind direction uncertainty decreases for decreasing wind speed. This is in contrast to ERS-1 and is due to the increased azimuth modulation of the Ku-band (14 GHz) GMF for low wind speeds. A final important difference is that NSCAT is less sensitive to K_{pm} than ERS-1 suggesting that with realistic values of K_{pm}, NSCAT may retrieve wind more accurately.

3.3.5 Model-based Results

This section is included in order to apply the model-based Cramer-Rao bound in approximating the covariance of model-based wind estimates. Using representative wind fields, the dependency on cell location of the covariance of model-based wind estimates is shown. In addition, some insight into the relationship between model-based covariance and point-wise covariance is drawn using an example wind field. This is followed by a demonstration of how the number of parameters used in the wind-field model can affect the covariance of the model-based estimates. ERS-1 geometry is assumed but the results are qualitatively similar for NSCAT geometry. In this section it is assumed that K_{pm} is zero as non-zero K_{pm} only increases the overall uncertainty without changing the general behavior of the results. In addition, Long’s wind field model discussed in Section 2.6.2 is


Far swath speed error (Kpm = 0)

Simulated far swath speed error (Kpm = 0)

Far swath direction error (Kpm = 0)

Simulated far swath direction error (Kpm = 0)

(a) C-R bound

(b) Simulations

Figure 3.11: *Comparison of standard deviation for wind speed and direction estimates between those computed with Cramer-Rao bound and simulated values at far-track swath location (NSCAT).*

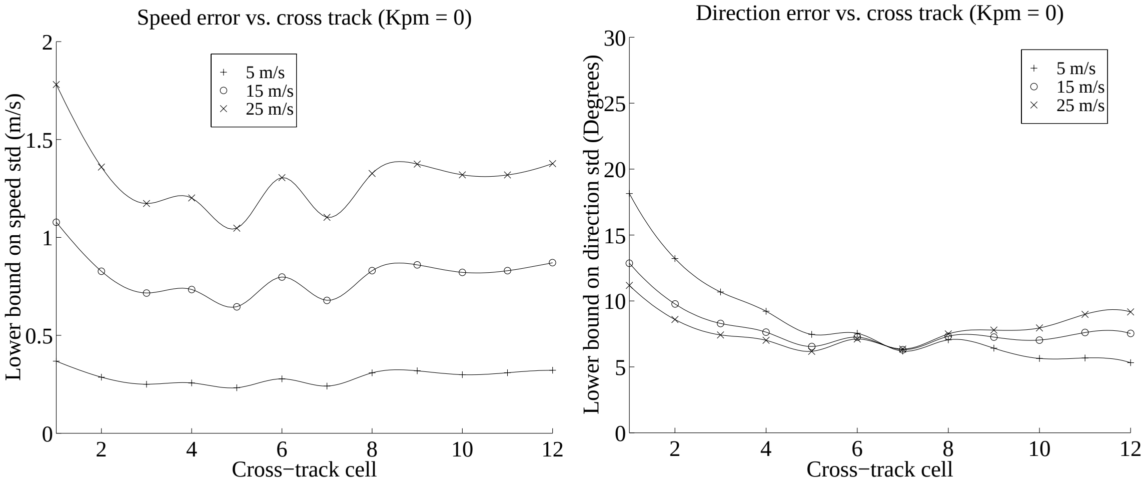


Figure 3.12: *NSCAT dependence of wind estimate uncertainty on cell location within a swath.*

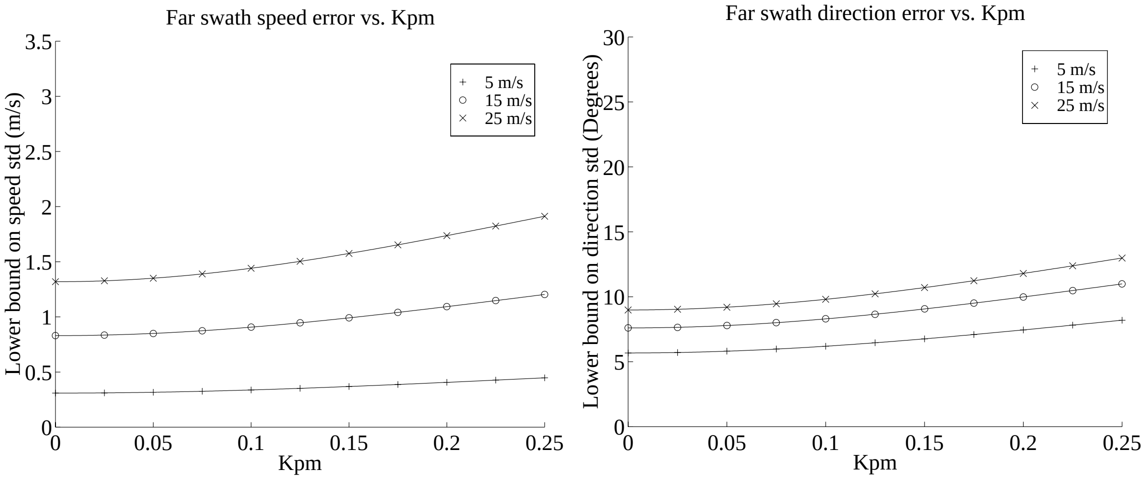


Figure 3.13: *NSCAT dependence of wind estimate uncertainty on K_{pm} for far-swath.*

used throughout.

Figure 3.14 shows how the variance of model-based direction and speed estimates depends on cell location in the region. The data for this figure was calculated by averaging the speed and direction computed using the model-based Cramer-Rao bound. The ascending portion of ERS-1 orbit 4452 was used to obtain the radar angles and noise parameters. A simulated wind field mapped to this orbit was used to generate the true model parameters needed by the Cramer-Rao bound. Notice that both speed and direction standard deviations are on average lower in the center of the region than along the borders.

It is also of interest to compare the covariance of model-based wind estimates to that of point-wise wind estimates. To demonstrate some of the relationships for a particular wind field using ERS-1 geometry, Figure 3.15 is included. To produce this figure a simulated wind field was selected as the true wind field. Using the point-wise Cramer-Rao bound, the covariance of the wind estimate in rectangular coordinates was calculated for each cell location in the region. In addition, the model-based Cramer-Rao bound was calculated for the projected wind field, $\mathbf{W}_F = FF^\dagger\mathbf{W}$, with F created using Long's model with 22 unknown parameters[4]. Assuming jointly Gaussian distributions for the u, v estimates, each covariance was used to calculate a 90% concentration ellipse for the wind vector estimate (Scharf, 1991, p. 225). The concentration ellipses are centered around the simulated wind field for the point-wise estimates, but for the model-based estimates they are centered around the projected wind field.

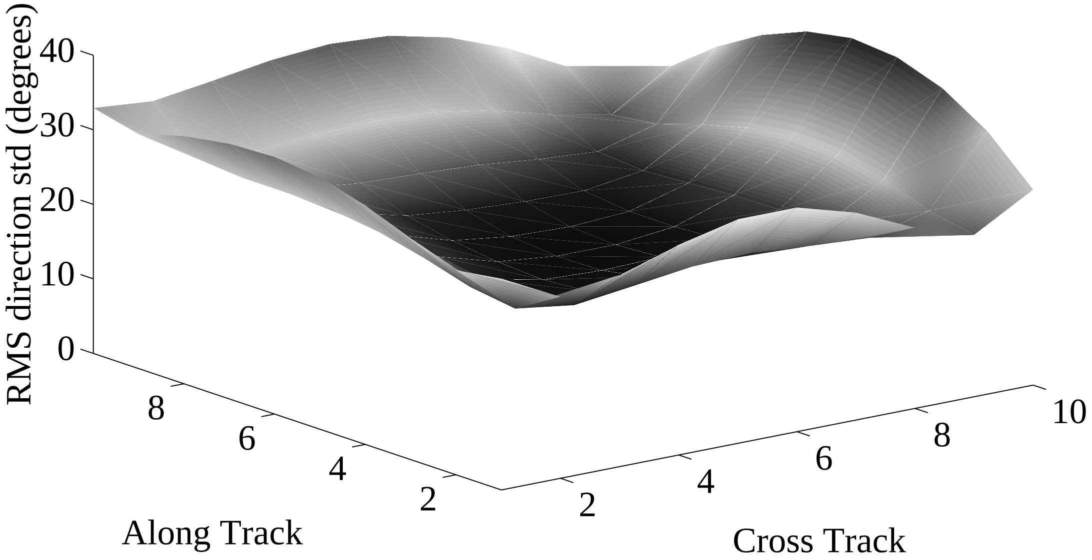
From this figure it is clear that point-wise estimates improve with increasing cross-track direction (increasing incidence angle). In addition, this figure shows that model-based wind estimates have concentration ellipses which are elongated with respect to point-wise wind estimates. As a result, the model-based speed estimates have lower variance than point-wise estimates [5]. On the other hand, this elongation also means that the model-based direction estimates have higher variance than the point-wise estimates at nearly all but the first few cross-track locations.

Figure 3.15 was generated using a model with 22 parameters. However, the particular wind field used as a true wind in this example can be well-modeled by fewer than 22 parameters. The model parameters which do little to reduce modeling error for

---

[4] Second-order curl and divergence parameterized boundary conditions model with 10 boundary condition parameters

[5] Recall, however, that modeling error also contributes to total model-based retrieval error



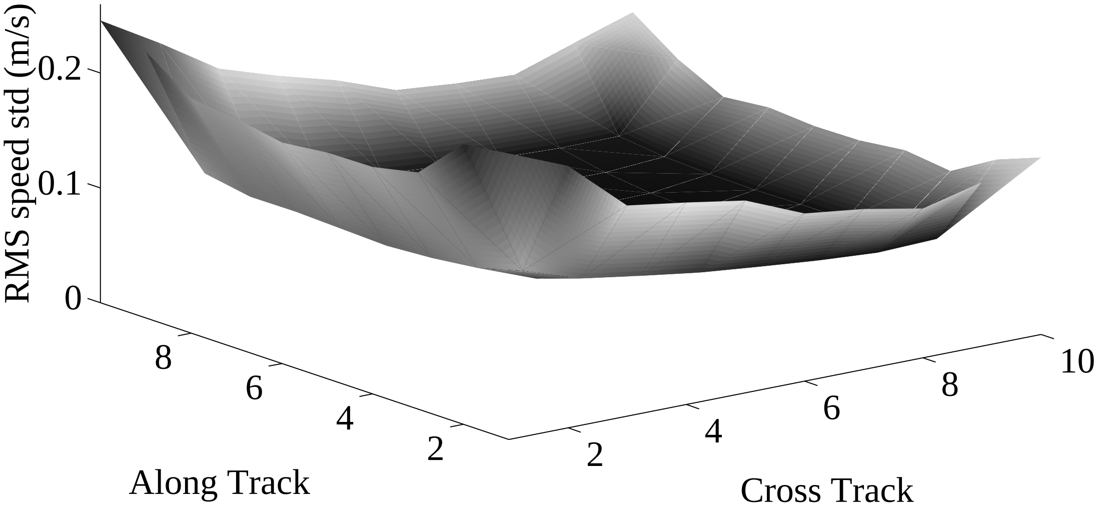


Figure 3.14: *The top plot shows RMS direction standard deviation as a function of location in the model region. The bottom plot shows RMS speed standard deviation as a function of location. The Cramer-Rao bound for 129 simulated wind fields mapped to the geometry of the ascending portion of ERS-1 orbit 4452 were used to compute the RMS values.*

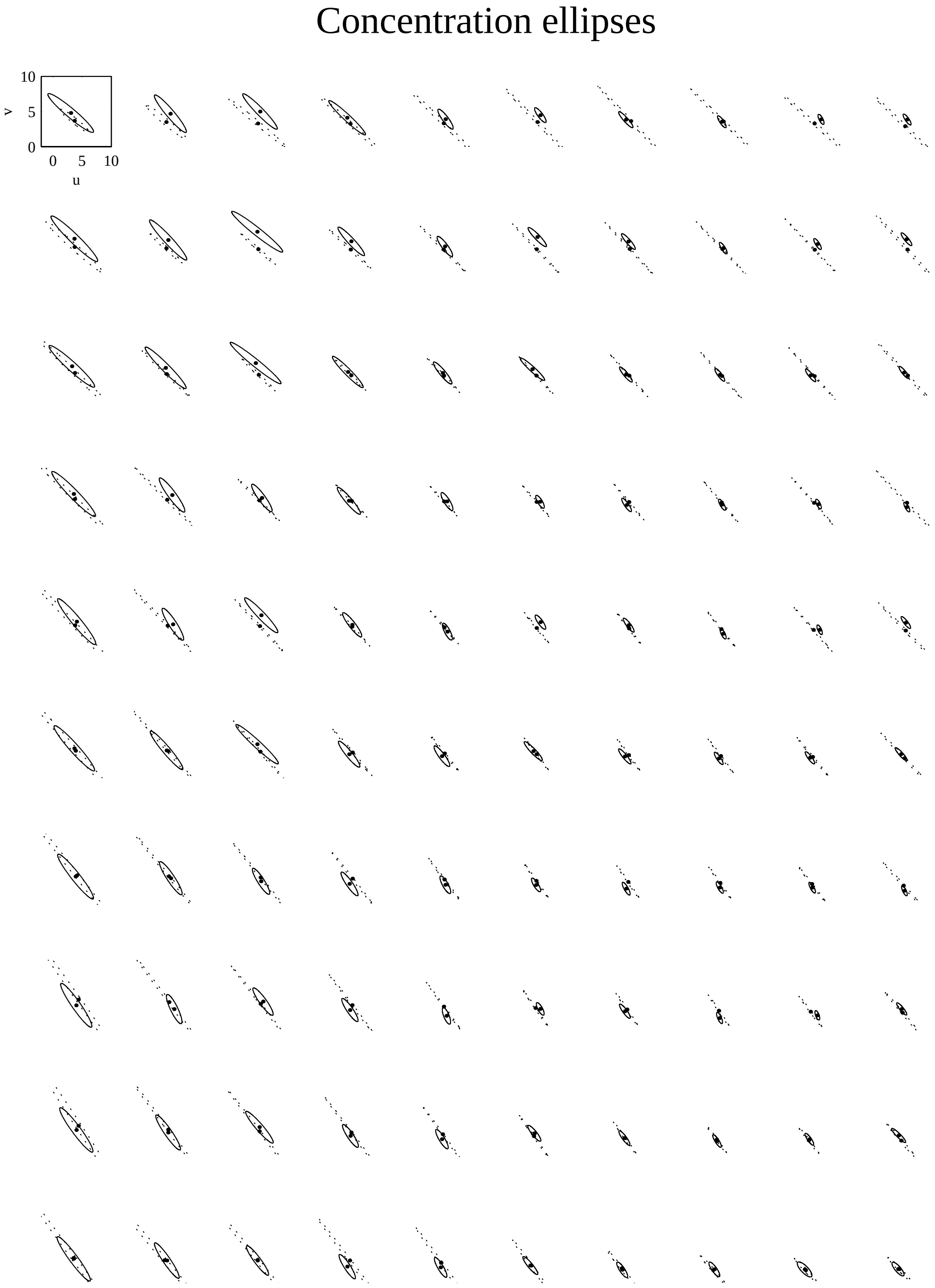


Figure 3.15: *Concentration ellipses for both point-wise (solid) and 22-parameter model-based (dashed) wind estimates calculated using the Cramer-Rao bound for a particular "true" wind field and corresponding projected field. Note that the ellipses are plotted in rectangular, u, v space. The axis for all the subplots is given by the upper left-hand corner subplot. Cross-track is across the page (1 to 10), and along-track is down the page.*

this particular wind field serve as nuisance parameters in model-based estimation. The uncertainty in estimating these unimportant parameters increases the overall uncertainty in the model-based retrieved wind. Improved model-based estimation could be obtained by using fewer parameters in the model, corresponding to the most-important columns of F for this particular wind field.

Figure 3.16 shows how using fewer parameters in the model reduces the extent of the model-based concentration ellipses while not significantly affecting the projection error. In this figure, a 12-parameter model was used corresponding to the most important columns of the 22-parameter F for this particular wind field[6]. Notice that the direction uncertainty is substantially reduced with little increase in projection error.

One conclusion that can be drawn from this last example is that model-based estimation often has two conflicting requirements for selecting a model. The first is to reduce modeling error which is generally done by increasing the number of columns of F. On the other hand, increasing the columns of F can increase the uncertainty in the model-based estimate, particularly in wind direction[7]. An ideal model for the wind would therefore have the number of parameters which would minimize some cost functional based on both modeling and estimation error. This ideal model would probably be wind field dependent.

3.3.6 Summary

In this section we have derived the Cramer-Rao bound on the covariance of wind estimates for both point-wise and model-based estimators. In addition, some calculated results were used to show typical wind errors obtained in wind retrieval as well as to compare the ERS-1 and NSCAT scatterometers. In fact, this demonstrates one of the most useful aspects of the Cramer-Rao bound as it applies to scatterometers: its ability to rapidly compare scatterometer designs in terms of wind retrieval accuracy.

One limitation of the Cramer-Rao bound is that it gives us the best performance possible, but does not give us any direct information about the performance of the particular estimator we are using. As a result of this limitation, in the next section we demonstrate

---

[6] These were precisely those columns remaining in Long's 12-parameter model constructed with constant curl and divergence and 10 boundary condition parameters.

[7] It should be noted that increasing the number of columns of F does not always increase uncertainty. When F is 200×200, then model-based uncertainty is exactly equal to point-wise uncertainty. (See Appendix D).

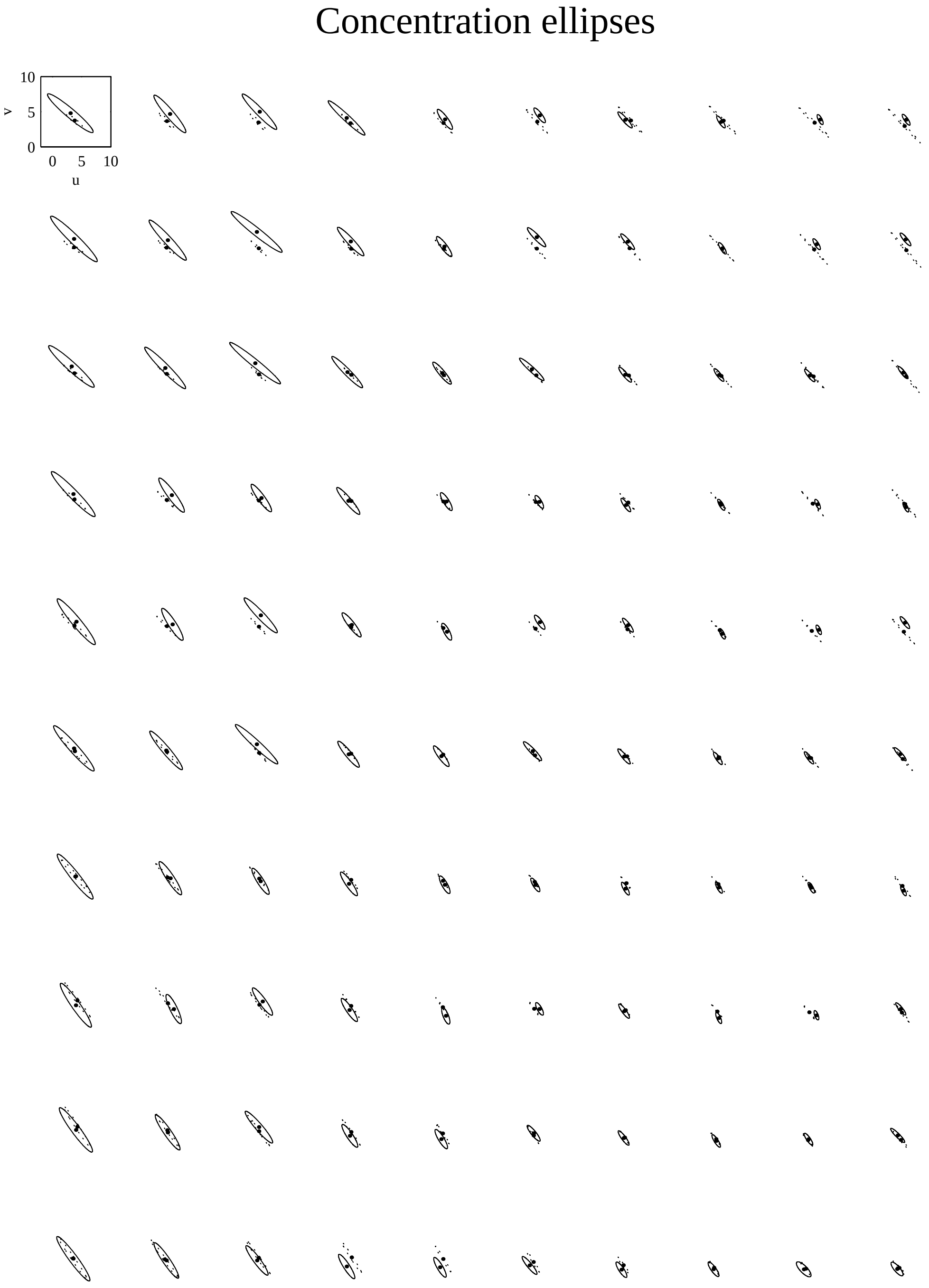


Figure 3.16: *Concentration ellipses for both point-wise (solid) and 12-parameter model-based (dashed) wind estimates calculated using the Cramer-Rao bound for a particular "true" wind field and corresponding projected field. Note that the ellipses are plotted in rectangular, u, v space. The axis for all the subplots is given by the upper left-hand corner subplot. Cross-track is across the page (1 to 10), and along-track is down the page.*

an alternative approach which attempts to approximate the covariance of the wind retrieval using a direct approach.

3.4 Direct Covariance Approximation

In this section we adapt the work of Fessler (1995; 1996) to finding an approximation of the covariance of each wind estimate. This technique applies to both model-based and point-wise wind estimates, but we will limit a detailed description to the point-wise case in order to facilitate explanation of the technique.

Fessler has described how the mean and covariance of implicitly defined estimators can be approximated using a Taylor series and the implicit function theorem (Fessler, 1995; Fessler, 1996). We apply this technique to approximating the covariance of the wind estimate obtained using maximum likelihood.

3.4.1 Method

The general approach is to recognize that the wind estimate is a function of the measurements of σ°. As a result, uncertainty in the measurements translates directly to uncertainty in the estimate. The function relating σ° measurements to wind estimate is implicitly defined. We can find derivatives of this function and thereby construct a Taylor-series approximation to the implicit function. Then, how the covariance of the wind estimate depends on the covariance of the σ° measurements can be determined directly.

The wind estimate for both point-wise and model-based retrieval can be written as

$$\hat{\mathbf{w}} = \arg\max_{\mathbf{w}} \left[L(\mathbf{w}, \mathbf{z}) \right], \tag{3.77}$$

where $L(\mathbf{w}, \mathbf{z})$ is the log-likelihood function given by Eq. (3.27) for point-wise retrieval.

Since all local maxima are considered as wind estimates, the above equation defines a set of wind estimates $\{\hat{\mathbf{w}}_n\}$. Each wind estimate is called a wind alias. At each wind alias, the implicit function that relates the wind estimate to the measurements is implicitly defined by solutions to the following (column) vector equation:

$$\mathbf{0} = \frac{\partial L(\mathbf{w}, \mathbf{z})^T}{\partial \mathbf{w}} |_{\mathbf{w}=\hat{\mathbf{w}}}, \tag{3.78}$$

This vector equation consists of p equations, where p is the number of elements in w ($p = 2$ for point-wise estimation, $p = 2MN$ for model-based estimation). Each row of this vector equation is listed explicitly for the point-wise case in Eqs. (3.28) and (3.29). As long as the Jacobian of this set of equations is non-zero at the point $(\mathbf{z}, \hat{\mathbf{w}}_n)$, then the general implicit function theorem states that these equations define a unique smooth function in the neighborhood of $(\mathbf{z}, \hat{\mathbf{w}}_n)$:

$$\hat{\mathbf{w}}_n = \mathbf{h}_n(\mathbf{z}), \tag{3.79}$$

and the derivatives may be computed by implicit differentiation (Marsden and Tromba, 1988, p. 287).

We can use this theorem and the unique smooth function it guarantees to formally define what we mean by a wind-alias random vector. Specifically, a given true wind, $\mathbf{w}$, gives rise to measurements, $\mathbf{z}$, which we model as random variables with mean $\bar{\mathbf{z}} = [\mathcal{M}_k(\mathbf{w})]$. Given this mean value there is a set of functions $\{\mathbf{h}_n(\mathbf{z})\}$ valid for $\mathbf{z}$ near $\bar{\mathbf{z}}$. Each function defines a wind-alias random vector, $\hat{\mathbf{w}}_n$ (since $\mathbf{z}$ is a random vector). Notice that the implicit function theorem is not specific about what it means to be *near* $\bar{\mathbf{z}}$. Clearly the utility of this definition for the wind-alias random vector breaks down if the variance of $\mathbf{z}$ is larger than the neighborhood of validity of the function $\mathbf{h}_n(\mathbf{z})$. We will not investigate this difficulty in this thesis, but only mention that this can become a problem in low-wind speed conditions where the peaks of the maximum likelihood equation are not well isolated[8].

With the notion of a wind-alias random vector defined we will cease distinguishing aliases by a subscript and simply denote a particular wind alias as $\hat{\mathbf{w}} = \mathbf{h}(\mathbf{z})$. Given the functional form of $\mathbf{h}$ we could theoretically determine the statistics of $\hat{\mathbf{w}}$ from the statistics of $\mathbf{z}$. Even though such a form is unavailable, we can expand $\mathbf{h}(\mathbf{z})$ in a Taylor series around $\bar{\mathbf{z}}$, and estimate the covariance of $\hat{\mathbf{w}}$ using the series expansion. The derivatives of $\mathbf{h}$ can be calculated using implicit differentiation. For simplicity in notation, in what follows we denote the derivative of a function, $f(x)$, evaluated at a point, x_0, as $\partial f(x_0)/\partial x$.

3.4.2 Derivation

The first-order Taylor formula allows us to write:

$$\hat{\mathbf{w}} = \mathbf{h}(\mathbf{z}) \approx \mathbf{h}(\bar{\mathbf{z}}) + D(\mathbf{z} - \bar{\mathbf{z}}), \tag{3.80}$$

---

[8] In such situations the measurements to not provide enough information to accurately predict the direction. This is consistent with the observation that wind direction retrieval is not accurate at low wind speeds.

where $D = \partial \mathbf{h}(\bar{\mathbf{z}})/\partial \mathbf{z}$ is a (row) derivative matrix with elements

$$D_{ij} = \frac{\partial h_i}{\partial z_j}. \tag{3.81}$$

From this approximation for $\hat{\mathbf{w}}$, estimates of the mean and covariance of $\hat{\mathbf{w}}$ can be obtained:

$$\begin{aligned} E\left(\hat{\mathbf{w}}\right) &= \mathbf{h}\left(\bar{\mathbf{z}}\right), && (3.82) \\ C_{\hat{\mathbf{w}}} &= E\left[(\hat{\mathbf{w}} - E\hat{\mathbf{w}})(\hat{\mathbf{w}} - E\hat{\mathbf{w}})^T\right], \\ &= DE\left[(\mathbf{z} - \bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}})^T\right] D^T, \\ &= DC_{\mathbf{z}}D^T. && (3.83) \end{aligned}$$

Note that this expression depends on the true wind value through the mean and covariance of the measurements since in point-wise notation,

$$\begin{aligned} \bar{\mathbf{z}} &= \left[\mathcal{M}_1 \cdots \mathcal{M}_K\right]^T, && (3.84) \\ C_{\mathbf{z}} &= \begin{bmatrix} \varsigma_{z_1}^2 & & 0 \\ & \ddots & \\ 0 & & \varsigma_{z_K}^2 \end{bmatrix}, && (3.85) \end{aligned}$$

and ς_{z_k} is given by Eq. (3.17).

The expression for the wind estimate covariance also depends on the derivative of $\mathbf{h}$ with respect to the measurements evaluated at the measurement mean. This derivative can be calculated using implicit differentiation. Using the chain rule, Eq. (3.78) can be differentiated with respect to the measurements $\mathbf{z}$ and evaluated at $\bar{\mathbf{z}}$:

$$\begin{aligned} \mathbf{0} &= \frac{\partial}{\partial \mathbf{z}}\left[\frac{\partial L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial \mathbf{w}}\right]^T + \frac{\partial}{\partial \mathbf{w}}\left[\frac{\partial L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial \mathbf{w}}\right]^T \frac{\partial}{\partial \mathbf{z}}\mathbf{h}(\bar{\mathbf{z}}), \\ \mathbf{0} &= D^{11} + D^{20}D, && (3.86) \end{aligned}$$

where

$$\begin{aligned} D^{11} &= \frac{\partial}{\partial \mathbf{z}}\left[\frac{\partial L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial \mathbf{w}}\right]^T = \left\{\frac{\partial^2 L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial z_j \partial w_i}\right\}, && (3.87) \\ D^{20} &= \frac{\partial}{\partial \mathbf{w}}\left[\frac{\partial L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial \mathbf{w}}\right]^T = \left\{\frac{\partial^2 L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial w_i \partial w_j}\right\}. && (3.88) \end{aligned}$$

Solving for D, we obtain

$$D = -(D^{20})^{-1}D^{11}. \tag{3.89}$$

For evaluating the derivatives at $(\bar{\mathbf{z}}, \mathbf{h}(\bar{\mathbf{z}}))$, there are at least two options for determining $\mathbf{h}(\bar{\mathbf{z}})$. The first option is to perform wind retrieval using noise-free ''measurements'', $\bar{\mathbf{z}}$, calculated from the true wind. An alternative is to recognize from Eq. (3.82) that $\mathbf{h}(\bar{\mathbf{z}})$ is a very good approximation to the expected value of $\hat{\mathbf{w}}$, which is the true wind, so that $\mathbf{h}(\bar{\mathbf{z}}) \approx \mathbf{w}$. The latter method has the advantage of avoiding the maximization required for the first option. As a result, this is the method used to obtain the results presented below.

For the point-wise case, the derivatives can be explicitly written using Eqs. (3.28) and (3.29) and substituting $\bar{z}_k = \mathcal{M}_k$:

$$D_{ij}^{11} = \frac{1}{\varsigma_{z_j}^2}\frac{\partial \mathcal{M}_j}{\partial w_i}, \tag{3.90}$$

$$D_{ij}^{20} = \sum_{k=1}^{K}\left[-\frac{\partial \mathcal{M}_k}{\partial w_i}\frac{1}{\varsigma_{z_k}^2}\frac{\partial \mathcal{M}_k}{\partial w_j} + \frac{\partial \varsigma_{z_k}^2}{\partial w_i}\frac{1}{2\varsigma_{z_k}^4}\frac{\partial \varsigma_{z_k}^2}{\partial w_j} - \frac{1}{2\varsigma_{z_k}^2}\frac{\partial^2 \varsigma_{z_k}^2}{\partial w_i \partial w_j}\right], \tag{3.91}$$

where for clarity

$$\frac{\partial^2 \varsigma_{z_k}^2}{\partial w_i \partial w_j} = \left[2\epsilon_k \mathcal{M}_k + \beta_k(1+K_{pm}^2)\right]\frac{\partial^2 \mathcal{M}_k}{\partial w_i \partial w_j} + 2\epsilon_k \frac{\partial \mathcal{M}_k}{\partial w_i}\frac{\partial \mathcal{M}_k}{\partial w_j}, \tag{3.92}$$

$$\frac{\partial \varsigma_{z_k}^2}{\partial w_i} = \left[2\epsilon_k \mathcal{M}_k + \beta_k(1+K_{pm}^2)\right]\frac{\partial \mathcal{M}_k}{\partial w_i}. \tag{3.93}$$

With the details described, we write the resulting covariance estimate for the retrieved wind by combining Eqs. (3.83) and (3.89):

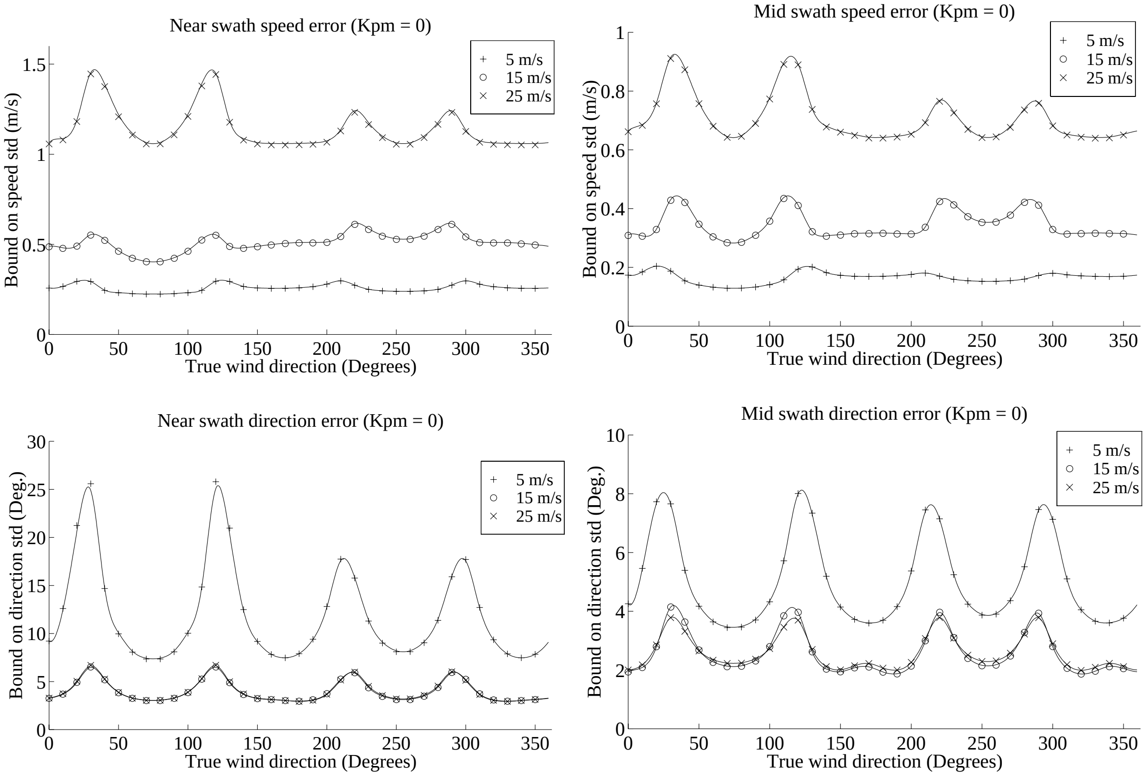
$$C_{\hat{\mathbf{w}}} = (D^{20})^{-1} D^{11} C_{\mathbf{z}} (D^{11})^T (D^{20})^{-1}. \tag{3.94}$$

Note that we have made use of the symmetry of D^{20}.

Eq. (3.94) gives an estimate of the covariance matrix for either point-wise or model based retrieval. In the following we present some calculations for point-wise retrieval using both ERS-1 and NSCAT geometries and compare them with Cramer-Rao calculations shown previously.

3.4.3 Comparisons with Cramer-Rao bound

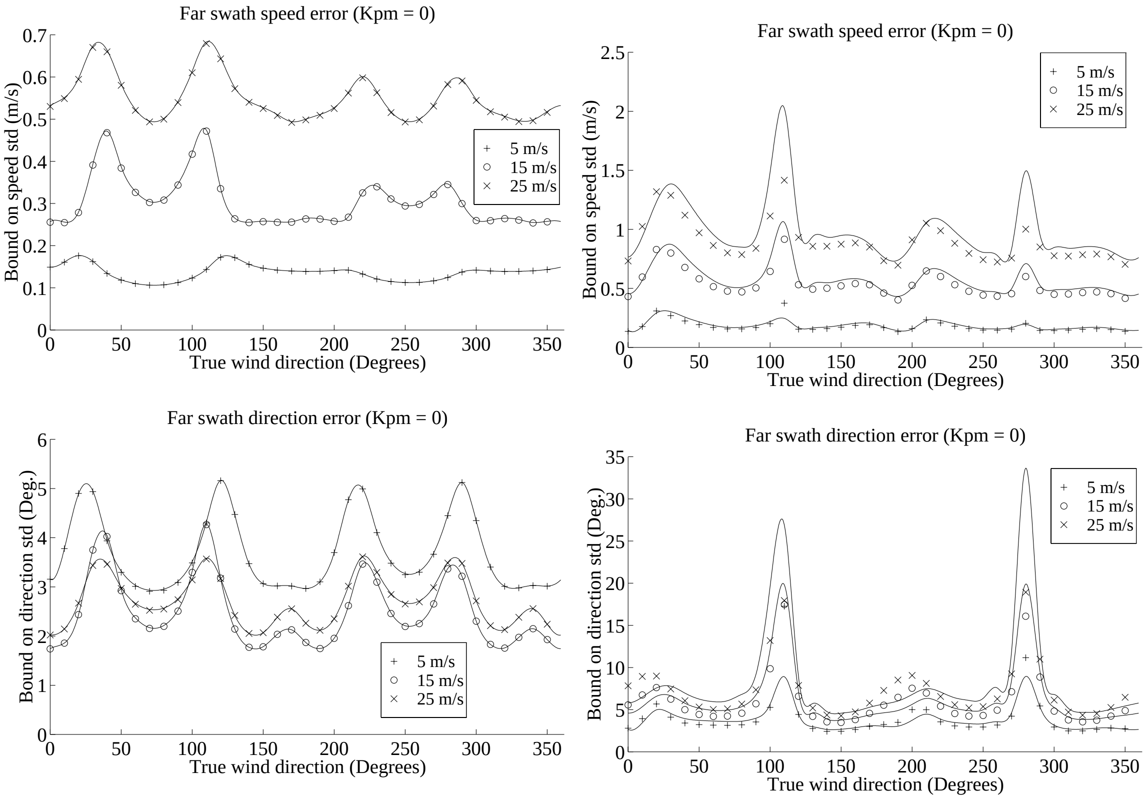
Figures 3.17 to 3.20 show comparisons between the speed and direction standard deviations obtained using the Cramer-Rao bound and the direct covariance approximation developed in the previous section. In Figures 3.17 and 3.18 $K_{pm} = 0$ is used, while



(a) ERS-1 near swath

(b) ERS-1 mid swath

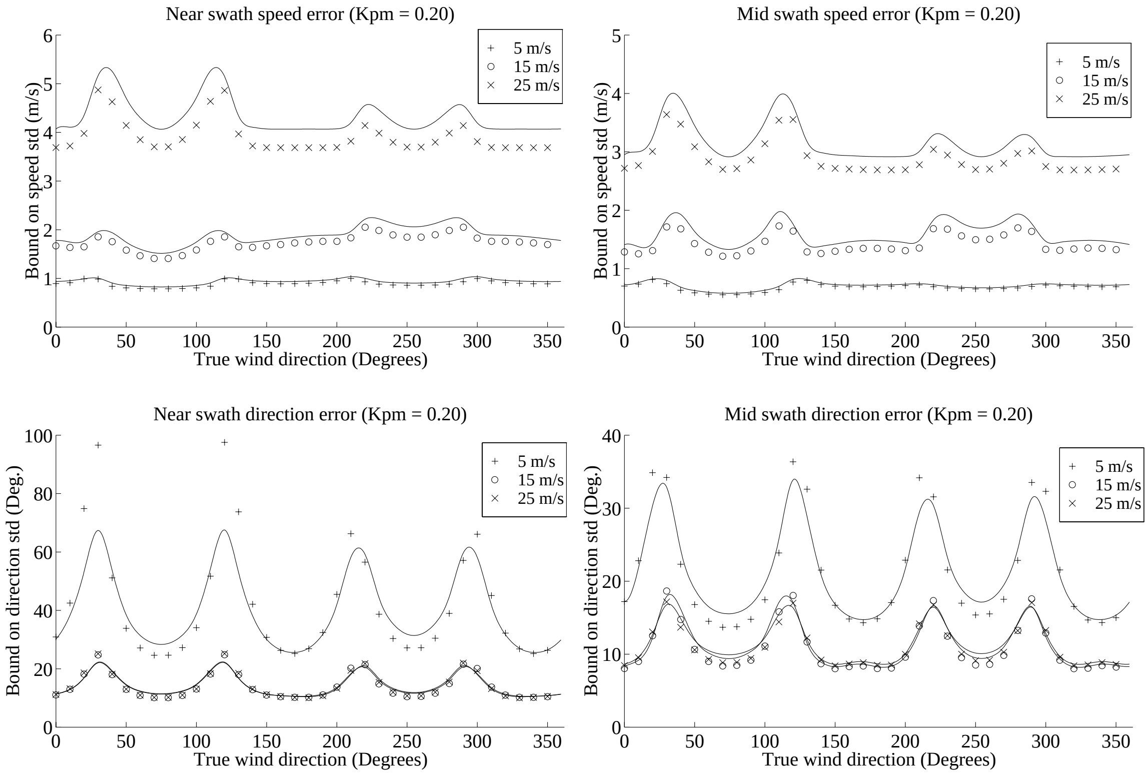
Figure 3.17: *Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0$).*



(a) ERS-1 far swath

(b) NSCAT far swath

Figure 3.18: *Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0$).*



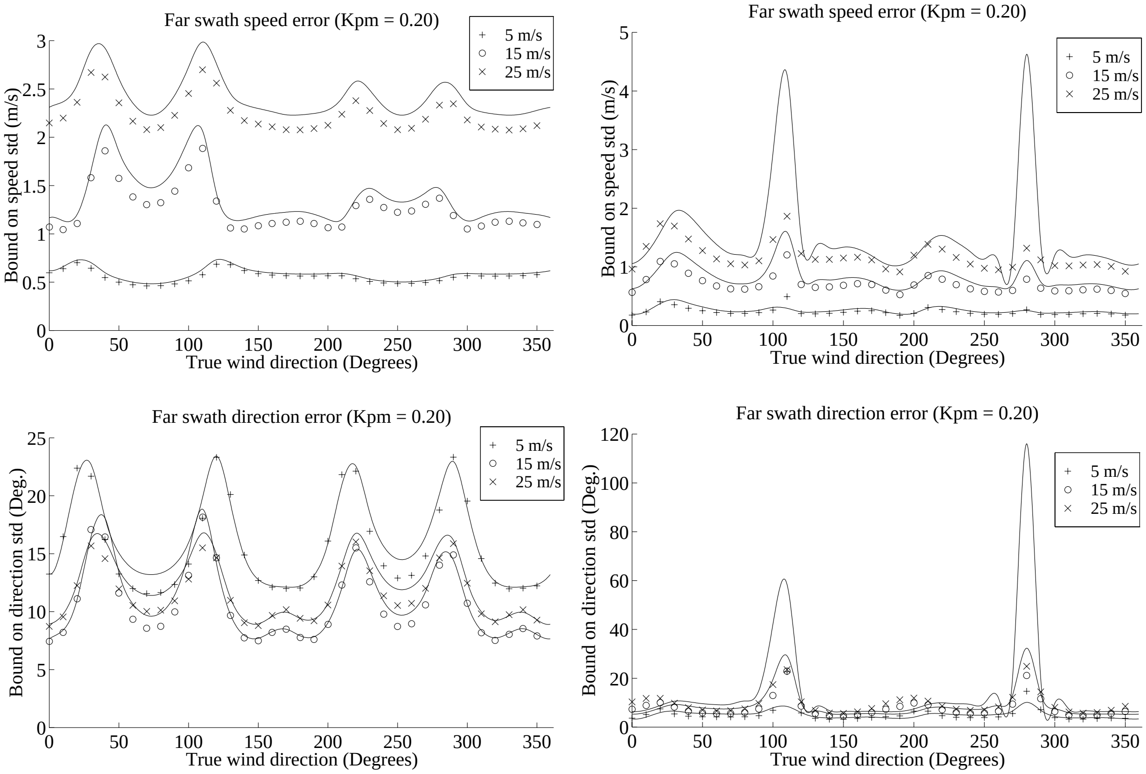
(a) ERS-1 near swath

(b) ERS-1 mid swath

Figure 3.19: *Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0.20$).*

in Figures 3.19 and 3.20 $K_{pm} = 0.20$ is used. In order to display representative wind retrieval conditions, the radar parameters and noise information were selected from three wind vector cells spanning the ERS-1 swath. In addition, results are shown using simulated information from one far-swath NSCAT cell.

From these figures some general conclusions can be drawn. One clear conclusion is that for ERS-1 at low K_{pm} values, the Cramer-Rao bound and the direct covariance approximation return nearly identical values. The results from the Cramer-Rao bound section show that the estimator is very nearly efficient under these low noise conditions. As a result, it is clear that when K_{pc} and K_{pm} are both low the direct covariance approximation is excellent. In other words, the implicit function $\mathbf{h}$ is well-approximated by a linear function under low-noise conditions. This is an intuitive result.



(a) ERS-1 far swath

(b) NSCAT far swath

Figure 3.20: *Comparison between direct approximation of covariance and Cramer-Rao bound. The individual marks are the Cramer-Rao bound while the lines are cubic spline fits to the direct approximation data ($K_{pm} = 0.20$).*

For NSCAT when $K_{pm} = 0$, the results are not quite as good. Figure 3.18(b) shows that the direct covariance approximation results in standard deviations which are less than the Cramer-Rao bound at certain wind velocities and greater than the bound at others. This problem of beating the lower bound is best seen in direction standard deviation estimates. An explanation for this effect is that at certain wind directions, especially where the Cramer-Rao bound predicts a high variance, the estimator is actually biased, so that the Cramer-Rao bound does not apply. Therefore, the actual covariance can beat the unbiased Cramer-Rao bound. However, since at these locations, the results still do not agree with simulations, it can also be said that at certain wind directions, the implicit function $\mathbf{h}$ is not well-approximated by a linear function, thus the covariance approximation is poor.

This problem is also evident in wind direction standard-deviation approximations for ERS-1 at near and mid swath when K_{pm} is large, especially at a wind speed of 5 m/s. It should be noted, however, that for ERS-1 when K_{pm} is large wind speed variance approximations have the expected relationship with the lower bound. Figures 3.19 and 3.20(a) show that the wind speed standard deviation computed using the direct approach are consistently bounded below by the Cramer-Rao bound. Such consistency leads to the conclusion that the implicit function for wind speed is well approximated by a line but the estimator is not efficient for large K_{pm} values. As a result, for ERS-1 data with $K_{pm} \approx 0.20$, more accurate wind speed variance estimates can be obtained using the direct covariance approach than using the Cramer-Rao bound. However, to obtain more accurate estimates of the wind direction variance, the direct covariance approximation presented here would need to be extended until the problem of under-prediction of the unbiased Cramer-Rao bound is removed.

The approximation can be improved by fitting the implicit function, $\mathbf{h}$, with a second- or higher-order polynomial through expanding $\mathbf{h}$ in a higher-order Taylor series. However, this dramatically increases the complexity of the problem. Not only must higher derivates of $\mathbf{h}$ be computed but the covariance becomes a function of higher-order moments of the measurement vector $\mathbf{z}$.

As a result the direct covariance approximation as presented in this section is only useful as a second-order correction to the wind speed variance predicted by the Cramer-Rao bound under conditions of large K_{pm}. Before it is used with confidence, however, it should be extended to use higher-order derivatives so that the magnitude of the error incurred in approximating the implicit function, $\mathbf{h}$, by a truncated Taylor series can be evaluated.

3.5 Summary of Wind Covariance Estimates

This chapter has discussed the problem of calculating the covariance of retrieved wind for both point-wise and model-based approaches. In order to present a more realistic picture of errors in wind estimates, the noise model of the σ° measurements was first expanded to include uncertainty in the GMF. The resulting noise model has an increased variance according to the fixed value of K_{pm} assumed.

With this noise model, the Cramer-Rao bound was derived as an estimate to the covariance of winds retrieved using both point-wise and model-based approaches. Comparison of the bound to simulated point-wise retrieval shows that the Cramer-Rao bound is a good approximation to the covariance of point-wise retrieved winds. The accuracy of the approximation decreases for increasing noise levels and at certain wind directions where the predicted variance is high. For, model-based retrieval, the asymptotic property of maximum likelihood estimators justifies use of the bound to approximate the model-based covariance.

Results show that uncertainty in wind retrieval is highly sensitive to the direction of the true wind. In particular, high retrieval uncertainty is expected when the true wind aligns with either the fore or aft beam. In addition, wind retrieval uncertainty is adversely affected by large values of K_{pm} particularly for the ERS-1 instrument. Uncertainty of NSCAT-retrieved winds is less influenced by K_{pm}. The ability to compare scatterometer designs easily with respect to wind retrieval uncertainty has implications for using the Cramer-Rao bound as a wind-scatterometer design tool.

Since the covariance wind estimate is not guaranteed to be the Cramer-Rao bound, another method of obtaining covariance estimates was briefly developed and applied to point-wise wind retrieval. The results showed that a much more complicated extension would need to be implemented in order to reliably obtain covariance estimates more accurate than the Cramer-Rao bound.

Chapter 4

DISTINGUISHING MAXIMA IN THE MAXIMUM-LIKELIHOOD EQUATION

In the previous chapter, some solutions to the problem of estimating the uncertainty in the retrieved wind for both model-based and point-wise wind estimates were presented. Now we turn our attention to a different problem common to both point-wise and model-based wind retrieval. Both methods obtain a wind estimate by maximizing an objective function (usually maximum likelihood) which has multiple local maxima. The usual maximum likelihood procedure is to use the global maximum as the maximum likelihood estimate. In this case, however, two or more of the maxima are of comparable magnitude, reducing the justification for selecting only the global maxima as the wind estimate. As a result, historically all of the local maxima have been kept as possible wind solutions. A second step, known as dealiasing, is used to select a unique wind vector. However, accepting all of the local maxima as possible wind solutions may result in the dealiasing step choosing a solution with a low likelihood value. Choosing a wind vector with a low likelihood value means that according to the noise model for the measurements, only with very low probability would the observed measurements have occured if the true wind were the selected wind vector. The problem for wind retrieval is to distinguish among the local maxima of the likelihood function and determine which have sufficiently high likelihood to be considered by the dealiasing step as a legitimate possible wind solution.

In this chapter a decision-theoretic solution to this problem is presented. The approach used is to use hypothesis-testing procedures on the likelihood-ratio statistic. In particular, the size of a test fashioned to eliminate the alias from future consideration is determined. This size represents the probability of throwing out the alias when it corresponds to the true wind solution. As a result, this size can be used to determine whether or not the wind alias could be eliminated.

The outline of this chapter proceeds as follows. First a more detailed description of the problem is given. This is followed by development of a practical decision rule for both point-wise and model-based retrieval. The technique is then applied to point-wise retrieved wind fields to demonstrate its effectiveness on ERS-1 and NSCAT data (Oliphant and Long, 1996b). Finally, an example of applying the technique to model-based retrieved winds is presented.

4.1 Problem background

In both point-wise and model-based wind estimation, a maximum likelihood estimate is used to determine the wind. For point-wise wind estimation the log-likelihood function is maximized over the wind vector to choose wind estimates. For model-based retrieval an analogous log-likelihood function is maximized over the model parameters to choose the the wind estimate. In both retrieval methods, a problem surfaces in that more than one wind vector or set of model parameters gives rise to a significant maximum in the log-likelihood function. This was demonstrated in Chapter 2 for point-wise wind retrieval.

As a result of this estimation ambiguity, all maxima of the likelihood function have been traditionally kept as possible wind solutions. However, keeping all wind estimates which are only *local* maxima of the likelihood function does not seem to adhere to the general philosophy of using the maximum likelihood estimate: pick the wind velocity which gives the largest likelihood that the measurements would have been observed. It is hard to statistically justify retaining a wind velocity estimate that gives a distinctly lower likelihood value than the maximum value. The problem, however, has been how to determine what ‘‘distinctly lower’’ means. Where does one place such a threshold in a statistically justifiable way?

One answer is to form a decision rule that uses the likelihood function values to distinguish among the possible wind solutions. For each wind alias in question, a likelihood ratio statistic is formed with the most likely wind alias. The observed ratio of the likelihood values is then a realization of that statistic. The probability that such a realization would have been observed if the wind alias in question were the true wind is evaluated. The decision whether or not this particular wind alias is legitimate can be made on the basis of such a probability. In the next section this idea is developed in terms of standard hypothesis testing where it is found that this probability is the size of a likelihood-ratio decision rule fashioned to eliminate the alias.

4.2 Setting up the hypothesis test

4.2.1 General Approach

To establish notation, we form a set of possible wind estimates corresponding to maxima in the maximum likelihood equation. For point-wise retrieval these are wind vectors, w, while for model-based retrieval these are projected wind fields, $\mathbf{W}_F = F\mathbf{X}$. For the sake of convenience we describe the development of the hypothesis test in terms of point-wise retrieved wind vectors which can be generalized to model-based estimates in a straightforward manner.

For point-wise retrieval we have a set of realizations of wind estimate alias random variables $\left\{\hat{\mathrm{w}}_0^{(n)}\right\}$ containing generally two to six possible estimates and ordered so that $\hat{\mathrm{w}}_0^{(1)}$ corresponds to the wind estimate giving the largest likelihood value[1]. We know that only one of these estimates corresponds to the true wind solution. In other words $E[\hat{\mathrm{w}}^{(n)}] = \mathrm{w}$ for a particular value of n.

The decision problem is deciding whether or not a particular wind alias deserves to be considered a possible wind candidate in future dealiasing procedures. In other words, for each n the decision is whether or not to eliminate $\hat{\mathrm{w}}_0^{(n)}$ from future dealiasing steps. Since $\hat{\mathrm{w}}_0^{(1)}$ represents the global maximum of the likelihood function it is always retained as a possible wind solution. For each $n > 1$, the decision can be expressed in terms of a hypothesis test. The simple null hypothesis is

$$H_0 : \mathrm{w} = E[\hat{\mathrm{w}}^{(n)}], \tag{4.1}$$

leaving a composite alternative hypothesis

$$H_1 : \mathrm{w} \in S - \left\{E[\hat{\mathrm{w}}^{(n)}]\right\}, \tag{4.2}$$

where S is a set of expected values of all the wind aliases.

In general, the decision based on the observed vector of K σ° values, $\mathbf{z}$, is $H_{\phi(\mathbf{z})}$ where $\phi(\mathbf{z})$ is a decision rule:

$$\phi(\mathbf{z}) = \begin{cases} 0 & \mathbf{z} \in R_1, \\ 1 & \mathbf{z} \in \widetilde{R}_1, \end{cases} \tag{4.3}$$

with R_1 a subset of $\mathbf{R}^K$ with complement $\widetilde{R}_1$. As a review, the size of this test is defined as $\alpha_0 = \Pr[\text{reject } H_0 \mid H_0]$ and the power of this test is defined as $\beta(\mathrm{w}^{(k)}) = \Pr[\text{reject } H_0$

---

[1] Note that we have distinguished the realization from the random variable by using the subscript 0.

$| H_1]$ (Scharf, 1991). Ideally, R_1 is determined so that the decision rule is uniformly most powerful for each size. This means that R_1 is chosen so that the probability of eliminating the alias when it should be eliminated is maximized for every $\mathrm{w} \in H_1$, while keeping the probability of throwing out the correct alias to a specified small number. This is a difficult problem in the given context, and it is not even guaranteed that such a decision rule exists. As a result we turn to a more practical decision rule based on the likelihood ratio.

4.2.2 Practical Approach

As an alternative to the general decision rule, we develop a practical decision rule for each wind alias ($n > 1$). This rule allows elimination of improbable aliases (though perhaps not optimally in the sense of Neyman-Pearson) while still keeping the probability of throwing out the correct alias to a small number, α_0. Using the probability density function given in Eq. (3.15) (or in Eq. (3.19) for model-based retrieval) form the log-likelihood-ratio statistic,

$$\Lambda(\mathbf{z}) = \log\left(\frac{p_{\mathbf{z}}(\mathbf{z}|\hat{\mathbf{w}}_0^{(n)})}{p_{\mathbf{z}}(\mathbf{z}|\hat{\mathbf{w}}_0^{(1)})}\right), \tag{4.4}$$

so that $\Lambda(\mathbf{z}_0) \in (-\infty, 0)$. Then form a threshold decision rule based on the observed measurements $\mathbf{z_0}$:

$$\phi(\mathbf{z_0}) = \begin{cases} 0 & \Lambda(\mathbf{z}_0) > \kappa, \\ 1 & \Lambda(\mathbf{z}_0) \leq \kappa. \end{cases} \tag{4.5}$$

At this juncture one point regarding acceptance of H_0 should be clarified. Even though the decision may be in favor of H_0 for a particular wind alias, this does not imply that this wind alias actually corresponds to the true wind. This would be true only if the power of the test is very high for each wind vector in the alternate hypothesis set (so that the probability of choosing H_0 when the true wind is in the alternative set is very low). This is unlikely given that $\hat{\mathrm{w}}_0^{(1)}$ has a larger likelihood value than the wind alias in question. In recognition of this caveat, it is evident that this test cannot eliminate the dealiasing problem completely. The real gain for wind dealiasing from this test comes when the measurements statistically corroborate acceptance of H_1 and therefore rejection of H_0. This allows confident elimination of the alias in question from future dealiasing steps.

Standard hypothesis testing procedure is to choose the size, α_0, so that the probability of rejecting H_0 when it should be accepted is satisfactorily small, *i.e.*, an acceptably low probability of eliminating the alias corresponding to the correct wind is chosen. Then for each hypothesis test (one for each wind alias) the resulting threshold $\kappa \in (-\infty, 0)$ is calculated and the observed value of the statistic, Λ, is compared against this threshold to determine which hypothesis to accept. In this case, however, the purpose of the test is to eliminate wind aliases whenever it is statistically justifiable. As a result, we turn the standard hypothesis testing paradigm around a bit and start by choosing the decision threshold to be equal to the observed likelihood statistic, $\kappa_0 = \Lambda(\mathbf{z}_0)$. This guarantees that the test eliminates the alias. Then the size, α_0, of the test is computed. If the size is acceptably small (*i.e.*, the probability of eliminating the alias when it is the true wind is low), the decision rule is implemented and the alias is eliminated from future consideration. If the size is too large the κ_0-decision rule is ignored and the wind vector is kept as a possible wind solution.

4.3 Determining the size of the test

4.3.1 Defining a related statistic

In order to apply the alias-elimination methodology outlined in the previous section, a formula for the size, α_0, as a function of κ_0 is needed[2]:

$$\alpha_0 = \text{Prob}[\,\Lambda(\mathbf{z}) \leq \kappa_0 \,|\, \mathbf{w} = E[\hat{\mathbf{w}}^{(n)}]] = F_\Lambda(\kappa_0). \tag{4.6}$$

To calculate α_0, the GMF needs to be evaluated at different wind velocity vectors so we establish a simplified notation. For a given measurement, z_k, represented in the likelihood function we denote the value of the GMF evaluated at wind velocity vector $\hat{\mathbf{w}}_0^{(n)}$ (with the radar information associated with z_k) as $\mathcal{M}_{k,n}$, *i.e.*, if $\hat{\mathbf{w}}_0^{(n)} = (U^{(n)}, \phi^{(n)})$ then

$$\mathcal{M}_{k,n} = \mathcal{M}(\theta_k, U^{(n)}, \phi^{(n)}, \psi_k, p_k). \tag{4.7}$$

The value of the GMF evaluated at a general wind vector, $\mathbf{w} = (U, \phi)$, using radar information associated with z_k is

$$\mathcal{M}_k = \mathcal{M}(\theta_k, U, \phi, \psi_k, p_k). \tag{4.8}$$

---

[2] Perhaps we should define $\kappa_0^{(n)}$ and $\alpha_0^{(n)}$ for each $n > 1$ in order to emphasize that the test is separately applied for each wind alias and that the sizes of each test are different, but for notational simplicity we do not.

For model-based retrieval we extend this notation so that

$$\mathcal{M}_{k,l,n} = \mathcal{M}(\theta_{k,l}, U_l^{(n)}, \phi_l^{(n)}, \psi_{k,l}, p_{k,l}). \tag{4.9}$$

Using this notation and Eq. (3.15) we can write $\Lambda(\mathbf{z})$ as

$$\Lambda(\mathbf{z}) = -\sum_{k=1}^{K} \left[\log\left(\frac{\varsigma_{k,n}}{\varsigma_{k,1}}\right) + \frac{(z_k - M_{k,n})^2}{2\varsigma_{k,n}^2} - \frac{(z_k - M_{k,1})^2}{2\varsigma_{k,1}^2} \right], \tag{4.10}$$

where

$$\varsigma_{k,n}^2 = \epsilon_k \mathcal{M}_{k,n}^2 + (\beta_k \mathcal{M}_{k,n} + \gamma_k)(1 + K_{pm}^2). \tag{4.11}$$

with $\epsilon_k = \alpha_k + K_{pm}^2 + \alpha_k K_{pm}^2$.

With foresight we define several terms useful in simplifying this expression:

$$\begin{aligned}
a &= \sum_{k=1}^{K} \log\left(\frac{\varsigma_{k,n}}{\varsigma_{k,1}}\right), && (4.12)\\
x_k &= \frac{z_k - \mathcal{M}_k}{\varsigma_k}, && (4.13)\\
b &= \sum_{k=1}^{K} \frac{(\mathcal{M}_{k,1} - \mathcal{M}_{k,n})^2}{2\left(\varsigma_{k,1}^2 - \varsigma_{k,n}^2\right)}, && (4.14)\\
Y &= -\Lambda(\mathbf{z}) - a + b, && (4.15)\\
c_k &= \frac{\varsigma_k^2}{2}\left(\frac{1}{\varsigma_{k,n}^2} - \frac{1}{\varsigma_{k,1}^2}\right), && (4.16)\\
d_k &= \frac{\varsigma_{k,n}^2(\mathcal{M}_k - \mathcal{M}_{k,1}) - \varsigma_{k,1}^2(\mathcal{M}_k - \mathcal{M}_{k,n})}{\varsigma_k(\varsigma_{k,1}^2 - \varsigma_{k,n}^2)}. && (4.17)
\end{aligned}$$

With these definitions, it is shown in Appendix C that Eq. (4.10) can be written as

$$\begin{aligned}
-\Lambda(\mathbf{z}) - a + b &= \sum_{k=1}^{K} c_k (x_k - d_k)^2, \\
Y &= \sum_{k=1}^{K} c_k (x_k - d_k)^2. && (4.18)
\end{aligned}$$

Given a particular value of the wind vector, $\mathbf{w}$, each x_k is a zero-mean, unit-variance Gaussian random variable. Thus, α_0 can be computed as the probability that Y, a random variable which is related to Λ, is greater than or equal to $y_0 = -\kappa_0 - a + b$:

$$\begin{aligned}
\alpha_0(\kappa_0, \mathbf{w}) &= \text{Prob}[\ \Lambda(\mathbf{z}) \le \kappa_0 \mid \mathbf{w} = E[\hat{\mathbf{w}}^{(n)}]], && (4.19)\\
&= \text{Prob}[\ Y \ge -\kappa_0 - a + b \mid \mathbf{w} = E[\hat{\mathbf{w}}^{(n)}]], && (4.20)\\
&= 1 - F_Y(-\kappa_0 - a + b) = \int_{y_0}^{\infty} f_Y(y)\,dy\,. && (4.21)
\end{aligned}$$

The density function of Y can, in theory, be derived from the zero-mean, unit-variance normal density functions of x_k since Y is a quadratic function of each x_k. Also, since each c_k and d_k is parameterized by the wind vector, Y is also parameterized by the wind vector which we have indicated by writing α_0 explicitly as a function of w.

4.3.2 Finding an explicit formula for the size

In order to find α_0 for a given κ_0, the density function of Y is needed. Since each of the measurements z_k is assumed independent in the measurement model, Y is the sum of independent random variables.

$$Y = \sum_{k=1}^{K} R_k, \tag{4.22}$$

where $R_k = c_k(x_k - d_k)^2$. As a result, $f_Y(y)$ can be written as the inverse Laplace transform of the product of the moment generating functions of R_k:

$$\begin{aligned} f_Y(y) &= \frac{1}{2\pi}\int_C e^{-sy}\Phi_Y(s)ds\,, && (4.23)\\ &= \frac{1}{2\pi}\int_C e^{sy}\prod_{k=1}^{K}\Phi_{R_k}(s;c_k,d_k)ds\,, && (4.24)\end{aligned}$$

where C is a vertical line in the region of convergence of $\Phi_Y(s)$. In this expression, $\Phi_{R_k}(s;c_k,d_k)$ is the moment generating function of R_k:

$$\Phi_{R_k}(s) = \int_{-\infty}^{\infty} e^{sr} f_{R_k}(r;c_k,d_k). \tag{4.25}$$

Thus, to calculate the density of Y the density of each $f_{R_k}(r;c_k,d_k)$ must be calculated and then used to calculate $\Phi_{R_k}(s)$.

Calculating the density of f_{R_k}

Given a zero-mean, unit-variance random variable, x, we want to calculate the density of $R = c(x-d)^2$, which we denote, $f_R(r;c,d)$. This is accomplished using the fundamental theorem of statistics which states that to find $f_R(r)$ for a specific r, find the real roots of $r = g(x)$ denoting them x_m. In the region where no real roots exist $f_R(r) = 0$, otherwise,

$$f_R(r) = \sum_m \frac{f_x(x_m)}{|g'(x_m)|}. \tag{4.26}$$

Applying this to finding the density of R, we see that $r = g(x) = c(x-d)^2$ has two real roots provided $rc > 0$:

$$x_1 = d + \sqrt{\frac{r}{c}}, \tag{4.27}$$

$$x_2 = d - \sqrt{\frac{r}{c}}. \tag{4.28}$$

As a result, $|g'(x_1)| = |g'(x_2)| = 2\sqrt{cr}$. Therefore,

$$f_R(r;c,d) = \frac{1}{\sqrt{4cr}}\left[f_x(d+\sqrt{r/c}) + f_x(d-\sqrt{r/c})\right]U(rc). \tag{4.29}$$

In this equation, $U(x)$ is the unit step function defined to be zero for $x < 0$ and one for $x \geq 0$.

By substituting in the density function for a zero-mean, unit-variance random variable, the density for R can be rewritten as

$$f_R(r;c,d) = \frac{1}{\sqrt{2\pi}\sqrt{4cr}}\left[e^{-\frac{1}{2}\left(d+\sqrt{r/c}\right)^2} + e^{-\frac{1}{2}\left(d-\sqrt{r/c}\right)^2}\right]U(rc). \tag{4.30}$$

This equation can be simplified using hyperbolic cosines to

$$f_R(r;c,d) = \frac{1}{\sqrt{2\pi}\sqrt{cr}}e^{-\frac{d^2}{2}}e^{-\frac{r}{2c}}\cosh\left(d\sqrt{\frac{r}{c}}\right)U(rc). \tag{4.31}$$

This density function is right-sided if c is positive or left-sided if c is negative. This density can be recognized as that of a scaled non-central chi-square random variable.

Calculating the moment generating function of R

Since we now have the density function of R, calculating $\Phi_R(s;c,d)$ is a straightforward task:

$$\Phi_R(s;c,d) = E\{e^{sR}\} = \int_{-\infty}^{\infty} e^{sr} f_R(r;c,d)dr\,. \tag{4.32}$$

Substituting in the density function we obtain,

$$\Phi_R(s;c>0,d) = \int_0^{\infty} \frac{e^{sr}}{\sqrt{2\pi}\sqrt{4cr}}\left[e^{-\frac{1}{2}\left(d+\sqrt{r/c}\right)^2} + e^{-\frac{1}{2}\left(d-\sqrt{r/c}\right)^2}\right]dr\,, \tag{4.33}$$

$$\Phi_R(s;c<0,d) = \int_{-\infty}^{0} \frac{e^{sr}}{\sqrt{2\pi}\sqrt{4cr}}\left[e^{-\frac{1}{2}\left(d+\sqrt{r/c}\right)^2} + e^{-\frac{1}{2}\left(d-\sqrt{r/c}\right)^2}\right]dr\,. \tag{4.34}$$

With a change of variables, $u = \sqrt{r/c}$, these equations become

$$\Phi_R(s; c > 0, d) = \int_0^\infty \frac{ce^{scu^2}}{|c|\sqrt{2\pi}} \left[e^{-\frac{1}{2}(d+u)^2} + e^{-\frac{1}{2}(d-u)^2} \right] du \,, \tag{4.35}$$

$$\Phi_R(s; c < 0, d) = -\int_0^\infty \frac{ce^{scu^2}}{|c|\sqrt{2\pi}} \left[e^{-\frac{1}{2}(d+u)^2} + e^{-\frac{1}{2}(d-u)^2} \right] du \,. \tag{4.36}$$

This can be combined into one expression for both positive and negative c:

$$\Phi_R(s; c, d) = \int_0^\infty \frac{e^{scu^2}}{\sqrt{2\pi}} \left[e^{-\frac{1}{2}(d+u)^2} + e^{-\frac{1}{2}(d-u)^2} \right] du \,. \tag{4.37}$$

Each part of this integral can be evaluated separately by completing the square in the exponents. The final result is

$$\Phi_R(s; c, d) = \frac{1}{\sqrt{1-2cs}} \exp\left(\frac{d^2 cs}{1-2cs} \right), \qquad \text{Re}\{s\} c < \frac{1}{2}. \tag{4.38}$$

Obtaining an explicit formula for α_0

With the moment generating function of each R_k calculated, the moment generating function for Y can be written as

$$\Phi_Y(s) = \prod_{k=1}^{K} \frac{1}{\sqrt{1-2c_k s}} \exp\left(\frac{d_k^2 c_k s}{1-2c_k s} \right), \tag{4.39}$$

where the region of convergence is given by[3]:

$$\text{Re}\{s\} \in \left[-\frac{1}{2|\min(0, \{c_k\})|}, \frac{1}{2\max(0, \{c_k\})} \right] \tag{4.40}$$

Note that the imaginary axis is always in the region of convergence so that the Fourier transform of $f_Y(y)$ exists and is equal to $\Phi_Y(s)$ evaluated along the imaginary axis. As a result, we can write the density of $f_Y(y)$, and hence α_0, using Fourier transforms:

$$\alpha_0(\kappa_0, \mathbf{w}) = \frac{1}{2\pi} \int_{y_0}^\infty \int_{-\infty}^\infty e^{-j\nu y} \Phi_Y(j\nu) d\nu \, dy \,. \tag{4.41}$$

Interchanging the order of integration,

$$\alpha_0(\kappa_0, \mathbf{w}) = \frac{1}{2\pi} \int_{-\infty}^\infty d\nu \, \Phi_Y(j\nu) \int_{y_0}^\infty dy \, e^{-j\nu y}. \tag{4.42}$$

---

[3] $\pm 1/0 \equiv \pm\infty$

The inner integral can be recognized as the Fourier transform of a shifted unit-step function, $U(y-y_0)$, with ν as the transform variable. As a result,

$$\begin{aligned}\alpha_0(\kappa_0, \mathbf{w}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\nu\, \Phi_Y(j\nu)e^{-j\nu y_0}\left[\frac{1}{j\nu}+\pi\delta(\nu)\right], \\ &= \frac{1}{2}\Phi_Y(0)+\frac{1}{2j\pi}\int_{-\infty}^{\infty} d\nu\, \frac{1}{\nu}\Phi(j\nu)e^{-j\nu y_0}. \qquad (4.43)\end{aligned}$$

We can simplify this a little further by first noting that $\Phi_Y(0)=1$ and then defining $G(\nu)=1/\nu e^{-j\nu y_0}\Phi_Y(j\nu)$. Since $f_Y(y)$ is a real function, $\Phi_Y(-j\nu)=\Phi_Y^*(j\nu)$ and $G(-\nu)=-G^*(\nu)$. Consequently, the real part of the integrand is antisymmetric while the imaginary part of the integrand is symmetric. Hence, α_0 can be written as

$$\alpha_0(\kappa_0, \mathbf{w}) = \frac{1}{2}+\frac{1}{\pi}\int_0^{\infty} d\nu\, \mathrm{Im}\{\frac{1}{\nu}e^{-j\nu y_0}\Phi_Y(j\nu)\}. \qquad (4.44)$$

Explicitly writing the imaginary part we can express this equation as

$$\begin{aligned}\alpha_0(\kappa_0, \mathbf{w}) &= \frac{1}{2}+\frac{1}{\pi}\int_0^{\infty} d\nu\, \frac{\sin[\theta(\nu)]}{\nu\rho(\nu)}, & (4.45)\\ \theta(\nu) &= -\nu y_0+\sum_{k=1}^{K}\left[\frac{\nu d_k^2 c_k}{1+4c_k^2\nu^2}+\frac{1}{2}\tan^{-1}(2c_k\nu)\right], & (4.46)\\ \rho(\nu) &= \prod_{k=1}^{K}(1+4c_k^2\nu^2)^{1/4}\exp\left(\frac{2\nu^2c_k^2d_k^2}{1+4c_k^2\nu^2}\right). & (4.47)\end{aligned}$$

This expression can be trivially modified to match a similar expression for $1-\alpha_0$ given in the literature (Imhof, 1961).

4.4 Using estimate of wind instead of expected value.

In order to determine α_0 we need to know $E[\hat{\mathbf{w}}_n]$ (where the expectation is done over the measurements, $\mathbf{z}$) so that each c_k and d_k can be calculated. Unfortunately, given only one realization of each wind alias calculated by wind retrieval, the expected values are unknown. In its place we can use the realization of $\hat{\mathbf{w}}_n$ determined by the wind retrieval procedure. This is our best approximation to $E[\hat{\mathbf{w}}_n]$. To illustrate the sensitivity of α_0 to this substitution Figures 4.1 and 4.2 show α_0 as a function of wind velocity. For these figures, actual scatterometer data from a particular ERS-1 cell was used to calculate α_0 for several values of wind velocity around each wind alias using Eq. (4.44). Figure 4.3 shows a similar plot using simulated NSCAT data.

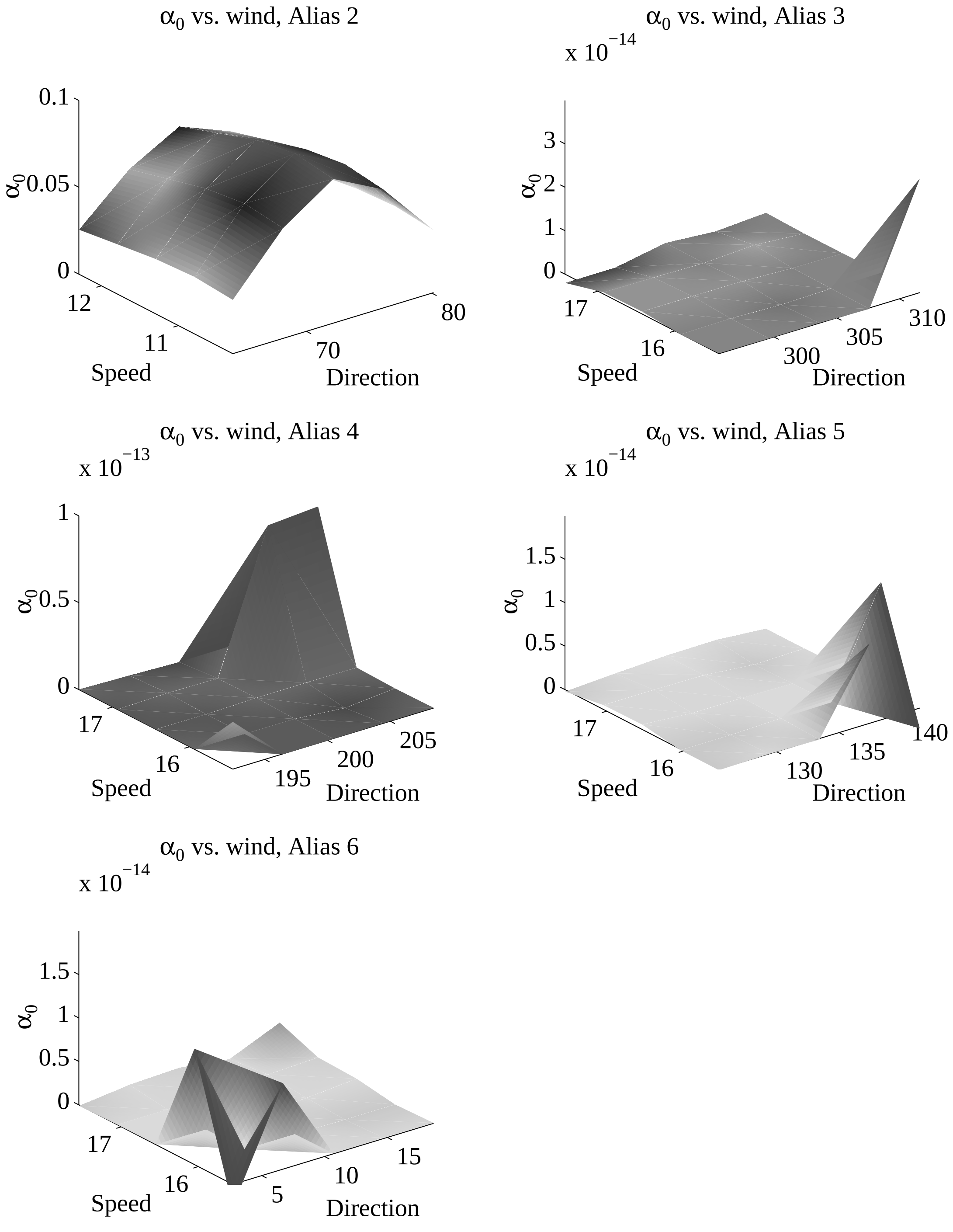
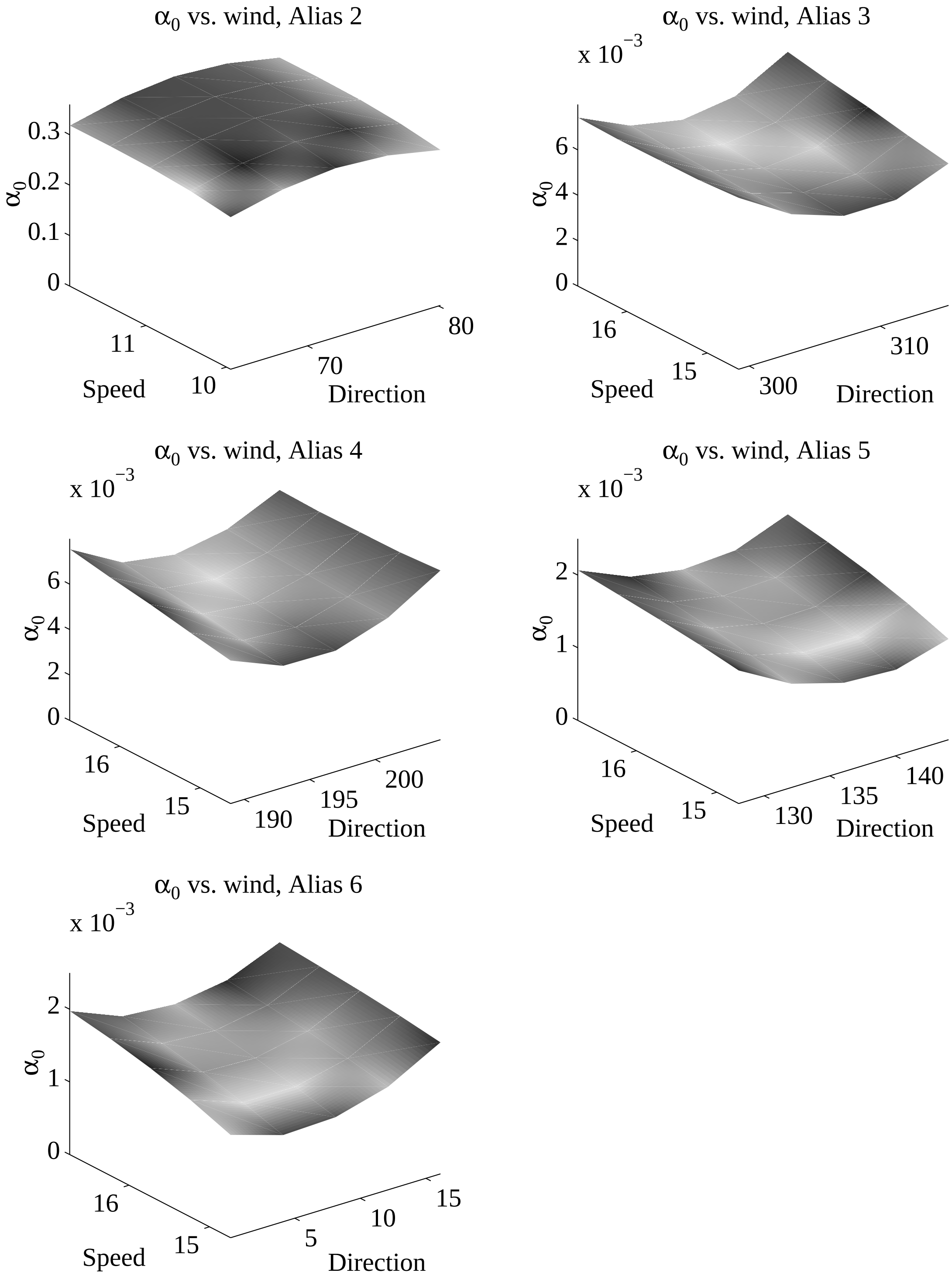


Figure 4.1: *Dependence on wind velocity of α_0 with $K_{pm} = 0$ for the wind aliases of a mid-swath ERS-1 cell in the ascending portion of orbit 7220. The apparent peaks of the last 4 figures is actually the effect of numerical round-off error. The surfaces are essentially flat.*



α_0 vs. wind, Alias 2

α_0 vs. wind, Alias 3

α_0 vs. wind, Alias 4

α_0 vs. wind, Alias 5

α_0 vs. wind, Alias 6

Figure 4.2: *Dependence on wind velocity of α_0 with $K_{pm} = 0.20$ for the wind aliases of a mid-swath ERS-1 cell in the ascending portion of orbit 7220.*

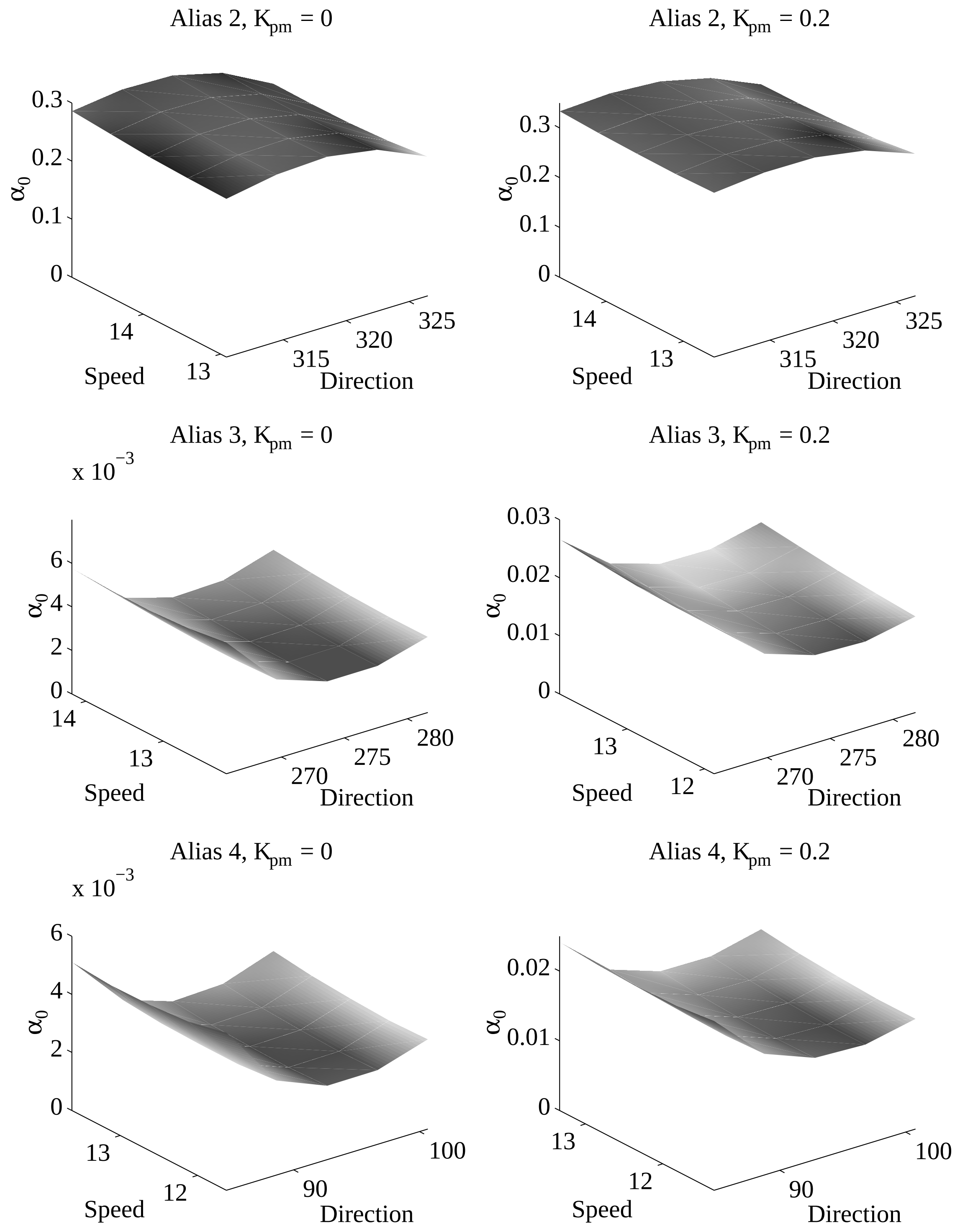


Figure 4.3: *Dependence on wind velocity of* α_0 *with* $K_{pm} = 0$ *and* $K_{pm} = 0.20$ *for the wind aliases of a near-swath simulated NSCAT cell.*

These figures show that α_0 changes as a function of the true wind vector. However, the change is small enough that it can be accounted for in practical alias elimination by choosing a more conservative threshold for what value of α_0 is considered low enough to eliminate the alias. Thus, the true wind vector need not be known.

Since in practice, α_0 is calculated for each alias with $\mathbf{w} = \hat{\mathbf{w}}_0^{(n)}$, $\mathcal{M}_k$ can be replaced with $\mathcal{M}_{k,n}$ in the expressions for c_k and d_k. As a result, c_k and d_k simplify to [refer to Eqs. (4.16) and (4.17)]:

$$\begin{aligned} c_k &= \frac{1}{2}\left(1 - \frac{\varsigma_{k,n}^2}{\varsigma_{k,1}^2}\right), && (4.48) \\ d_k &= \frac{\varsigma_{k,n}(\mathcal{M}_{k,n} - \mathcal{M}_{k,1})}{\varsigma_{k,1}^2 - \varsigma_{k,n}^2}. && (4.49) \end{aligned}$$

For completeness in calculating α_0, y_0 in Eq. (4.44) is

$$\begin{aligned} y_0 &= \log p_{\mathbf{z}}(\mathbf{z}_0|\hat{\mathbf{w}}_0^{(1)}) - \log p_{\mathbf{z}}(\mathbf{z}_0|\hat{\mathbf{w}}_0^{(n)}) - a + b, && (4.50) \\ a &= \sum_{k=1}^{K} \log\left(\frac{\varsigma_{k,n}}{\varsigma_{k,1}}\right), && (4.51) \\ b &= \sum_{k=1}^{K} \frac{(\mathcal{M}_{k,1} - \mathcal{M}_{k,n})^2}{2\left(\varsigma_{k,1}^2 - \varsigma_{k,n}^2\right)}. && (4.52) \end{aligned}$$

If the retrieved value of the wind alias is used, the test for wind alias elimination essentially becomes a calculation of the probability that an observed log-likelihood statistic value would be so small if the alias in question were really the true wind. Since there does not seem to be a simple closed-form expression for α_0, we must perform a numerical integration for each alias in order to decide to eliminate it or not. This can add significant processing time to the wind retrieval process. Note, however, that an exact calculation of α_0 is not necessarily needed. Consider the values of α_0 in the final four plots of Figure 4.1. These values are small enough that a relatively tight upper-bound on α_0 could suffice in order to determine that the true α_0 is small enough. To be useful, this upper bound should be both easy to compute and tight enough that it can be used effectively to eliminate aliases. In the next section, such a bound is described.

4.5 Determining an upper-bound for the size

The Markov inequality states that if the density of a random variable x is zero for $x < 0$, then for any $\epsilon > 0$,

$$\text{Pr}[x \geq \epsilon] \leq \frac{Ex}{\epsilon}. \tag{4.53}$$

With our random variable Y, form a new random variable $x = e^{sY}$, where s is an arbitrary real number. Since an exponential can never be negative for real values of Y, the density of x is zero for $x < 0$. In addition, if y_0 is any real number, then $e^{sy_0} > 0$. Thus, the conditions of the Markov inequality are satisfied with $x = e^{sY}$ and $\epsilon = e^{sy_0}$ so

$$\text{Pr}[e^{sY} \geq e^{sy_0}] \leq \frac{E\{e^{sY}\}}{e^{sy_0}}. \tag{4.54}$$

If we require that $s > 0$ then we can immediately see that

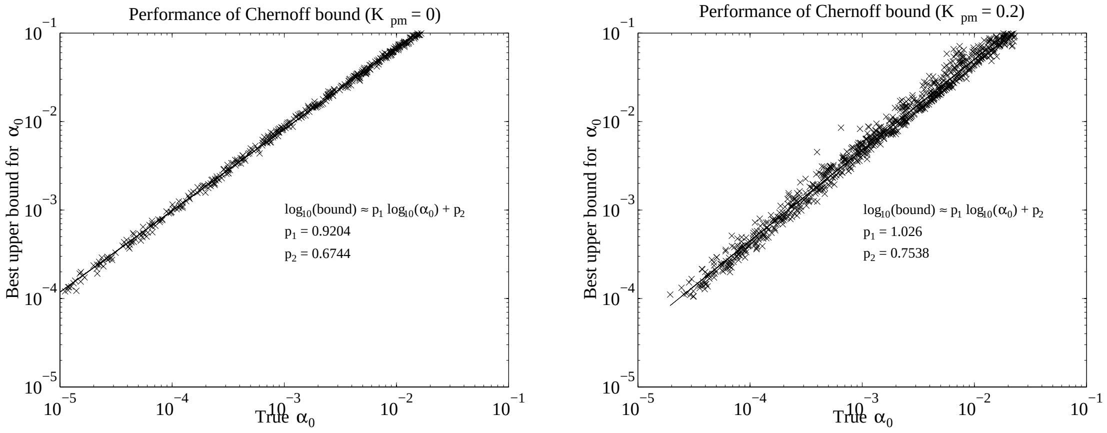
$$\begin{aligned} \alpha_0 &= \text{Pr}[Y \geq y_0] \\ &= \text{Pr}[e^{sY} \geq e^{sy_0}] \\ &\leq e^{-sy_0}\Phi_Y(s) \qquad s > 0. \end{aligned} \tag{4.55}$$

The above equation is one version of the Chernoff bound. Since it holds for all real, positive s in the region of convergence of Φ_Y, we can calculate a bound at any such value of s and use that as the bound for α_0.

The tightest Chernoff bound is found by finding the real, positive value of s in the region of convergence that minimizes $e^{-sy_0}\Phi_Y(s)$. However, since the bound is also valid for any value of $s > 0$ in the region of convergence, a single value for the bound could be used for all tests. Experiments using scatterometer data show that the commonly used bound given by $s = 1$ is tight enough to eliminate many aliases. Choosing $s = 1$, in the Chernoff bound, the bound for α_0 becomes

$$\alpha_0 \leq e^{-y_0}\Phi_Y(1) = e^{-y_0} \prod_{k=1}^{K} \frac{1}{\sqrt{1-2c_k}} \exp\left(\frac{d_k^2 c_k}{1-2c_k}\right). \tag{4.56}$$

To evaluate how the Chernoff bound compares with the actual value of α_0 computed using numerical integration, Figures 4.4 and 4.5 are given. Figure 4.4 presents two scatter plots showing how the best Chernoff bound compares to the true size for two different values of K_{pm}. In Figure 4.5 a similar plot shows the results for wind estimates



(a) $K_{pm} = 0$

(b) $K_{pm} = 0.2$

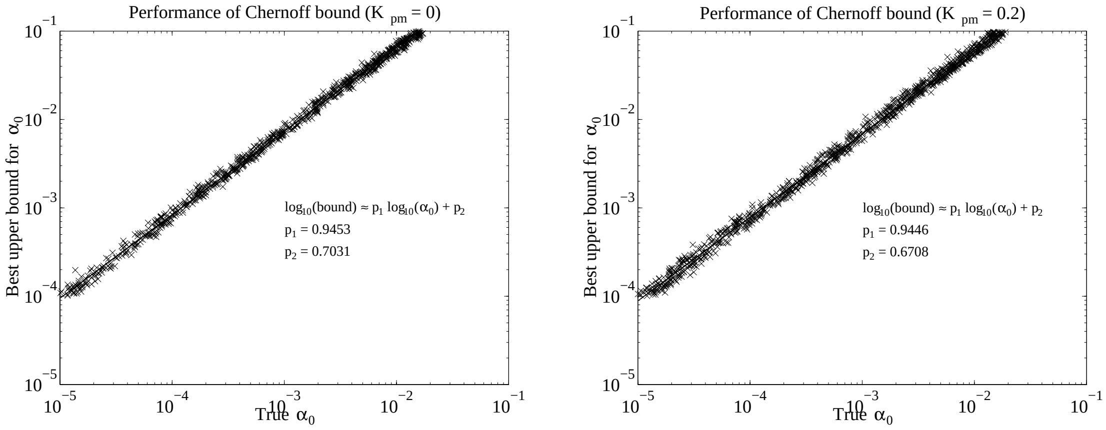
Figure 4.4: *Comparison of the best Chernoff bound with the numerically calculated value of α_0 for ERS-1 data.*

using NSCAT data. Also shown in these plots is an approximate linear fit in log-log space to the data. These plots were constructed by calculating both the best Chernoff bound and the actual value of α_0 for all aliases in the wind vector cells of the ascending portion of an actual ERS-1 and a simulated NSCAT orbit. These plots suggest that the bound is a good one. Moreover, the information on the plots can be used in interpreting the computed Chernoff bound in an alias-elimination algorithm.

One convenient way to implement the upper bound is to first calculate the $s = 1$ bound on α_0 for all aliases. For those aliases which cannot be justifiably eliminated based on the upper bound, the integration can be performed to calculate the true value of α_0. Since the upper bound is simple to calculate and is often enough to support removing the alias, significant processing time can be saved by avoiding the integration.

4.6 Extension to model-based

Extending the above alias elimination procedure to model-based processing is straightforward since the measurement model is the same. Instead of eliminating wind-vector aliases we are trying to eliminate wind-field aliases. We set up the same hypothesis test and evaluate the size of the test assuming a decision rule that uses the observed likelihood ratio. The resulting equation for α_0 and the upper bound on α_0 are exactly the



(a) $K_{pm} = 0$

(b) $K_{pm} = 0.2$

Figure 4.5: *Comparison of the best Chernoff bound with the numerically calculated value of* α_0 *for simulated NSCAT data.*

same as given by Eqs. (4.44) and (4.55). The only difference is that the moment generating function and y_0 are written to include all of the measurements made over the entire region.

We are given model-based estimates, $\hat{\mathbf{W}}_0^{(n)} = F\hat{\mathbf{X}}_0^{(n)}$, representing maxima in the model-based maximum likelihood equation. We would like to determine which, if any, of these aliases can be eliminated from future processing. The procedure is precisely the same as for point-wise estimates, just substitute $\hat{\mathbf{W}}_0^{(n)}$ for $\hat{\mathrm{w}}_0^{(n)}$. There are some notational differences, however, that need explanation. We need to express the GMF evaluated for the kth measurement of the lth wind vector in the nth wind-field alias. This is denoted

$$\mathcal{M}_{k,l,n} = \mathcal{M}(\theta_{k,l}, U_l^{(n)}, \phi_l^{(n)}, \psi_{k,l}, p_{k,l}). \tag{4.57}$$

The variance of the kth measurement of the lth wind vector in the nth wind-field alias is denoted similarly:

$$\varsigma_{k,l,n}^2 = \epsilon_{k,l}\mathcal{M}_{k,l,n}^2 + (\beta_{k,l}\mathcal{M}_{k,l,n} + \gamma_{k,l})(1 + K_{pm}^2). \tag{4.58}$$

With this notation, the model-based moment-generating function for use in testing the nth wind-field alias can be written:

$$\Phi_Y(s) = \prod_{l=1}^{MN}\prod_{k=1}^{K(l)} \frac{1}{\sqrt{1-2c_{k,l}s}} \exp\left(\frac{d_{k,l}^2 c_{k,l} s}{1-2c_{k,l}s}\right), \tag{4.59}$$

where the region of convergence is again given by

$$\mathrm{Re}\{s\} \in \left[-\frac{1}{2|\min(0,\{c_{k,l}\})|}, \frac{1}{2\max(0,\{c_{k,l}\})}\right]. \tag{4.60}$$

In addition,

$$c_{k,l} = \frac{1}{2}\left(1 - \frac{\varsigma_{k,l,n}^2}{\varsigma_{k,l,1}^2}\right), \tag{4.61}$$

$$d_{k,l} = \frac{\varsigma_{k,l,n}(M_{k,l,n} - M_{k,l,1})}{\varsigma_{k,l,1}^2 - \varsigma_{k,l,n}^2}. \tag{4.62}$$

The expression for y_0 is

$$y_0 = \log p(\mathbf{Z}_0|\mathbf{X}_0^{(1)}) - \log p(\mathbf{Z}_0|\mathbf{X}_0^{(n)}) - a + b, \tag{4.63}$$

$$a = \sum_{l=1}^{MN}\sum_{k=1}^{K(l)} \log\left(\frac{\varsigma_{k,l,n}}{\varsigma_{k,l,1}}\right), \tag{4.64}$$

$$b = \sum_{l=1}^{MN}\sum_{k=1}^{K(l)} \frac{(M_{k,l,n} - M_{k,l,1})^2}{2\left(\varsigma_{k,l,1}^2 - \varsigma_{k,l,n}^2\right)}, \tag{4.65}$$

where $\mathbf{Z}_0$ is the vector of observed measurements over the region.

4.7 Application to point-wise wind-alias elimination

In order to evaluate the performance of the statistical test developed in this chapter on eliminating wind-vector aliases, both ERS-1 and NSCAT is used for point-wise retrieval.

4.7.1 ERS-1

In this part we evaluate the technique as it applies to ERS-1 data. We consider both simulated and real data.

Simulated Data

As a first check of the procedure for eliminating wind-vector aliases, several simulations were performed. The true wind used in these simulations came from the European Center for Medium-range Weather Forecasting (ECMWF). These wind fields are described by Long (1989). The noise and radar information from part of a true ERS-1 orbit are used to generate simulated σ° measurements with probability density given by the noise model of Eq. (3.15). The number of aliases in each wind vector cell when applying the alias-elimination scheme is compared to the number of aliases per wind vector cell when the scheme is not applied.

The results of these simulations are shown in Figures 4.6 and 4.7 assuming $K_{pm} = 0$. These figures show that the alias-elimination scheme can be effective in simplifying the dealiasing step by reducing the number of aliases to two in most of the cells. Particularly interesting is that the alias problem is completely eliminated for several wind vector cells in both simulated swaths. This indicates that if the signal to noise ratio is large enough, a single wind estimate can be obtained.

The results presented above indicate that this alias-elimination scheme is remarkably successful in removing wind aliases, even to the point of eliminating all aliases at certain cells. However, in performing the simulations we assumed somewhat unrealistically that $K_{pm} = 0$. To see the effect that K_{pm} has on simulations, we perform the simulations again only this time letting $K_{pm} = 0.20$. The results of this simulation are shown in Figures 4.8 and 4.9. Notice that alias elimination is not as dramatic as when K_{pm} is zero.

The performance of the alias-elimination scheme on simulated data motivates

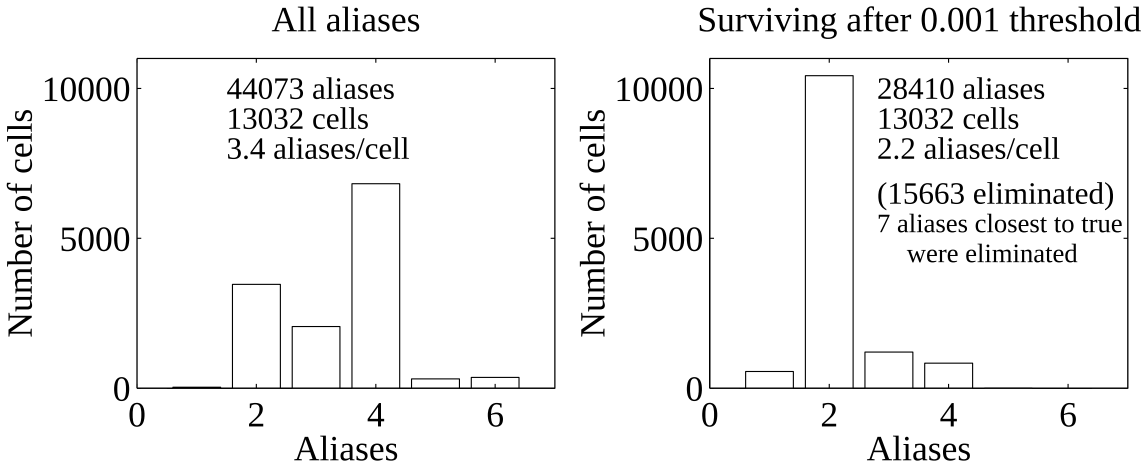


Figure 4.6: *Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the ascending portion of ERS-1 revolution 4452 ($K_{pm} = 0$).*

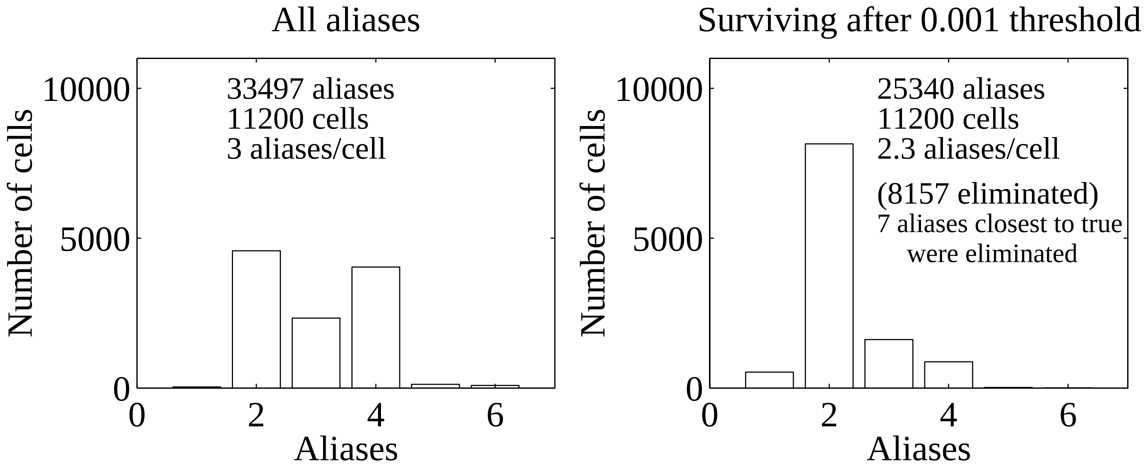


Figure 4.7: *Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the descending portion of ERS-1 revolution 4459 ($K_{pm} = 0$).*

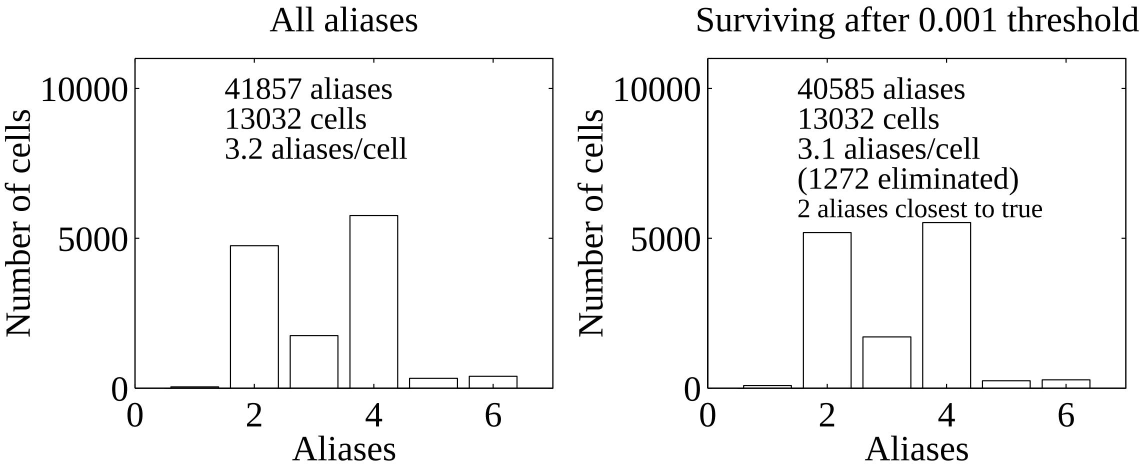


Figure 4.8: *Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the ascending portion of ERS-1 revolution 4452 ($K_{pm} = 0.20$).*

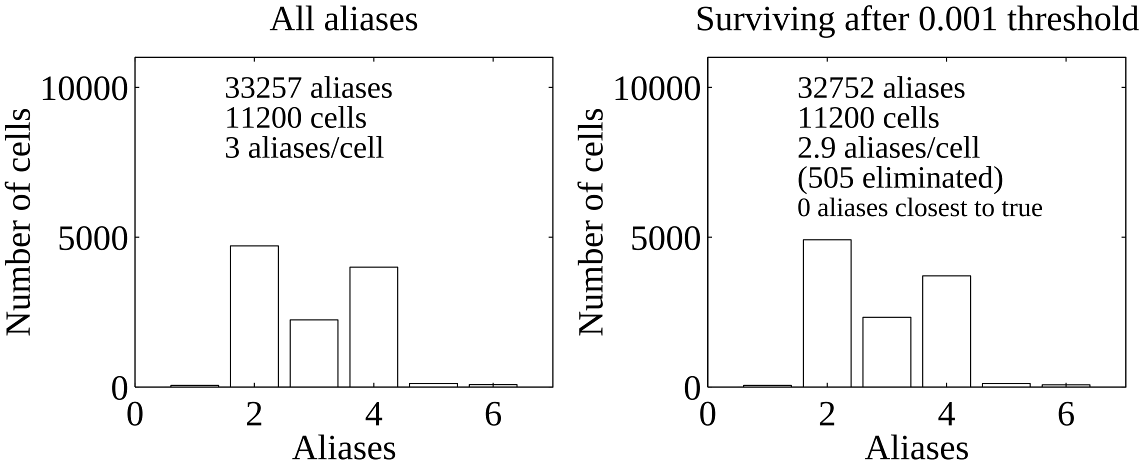


Figure 4.9: *Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information are from the descending portion of ERS-1 revolution 4459 ($K_{pm} = 0.20$).*

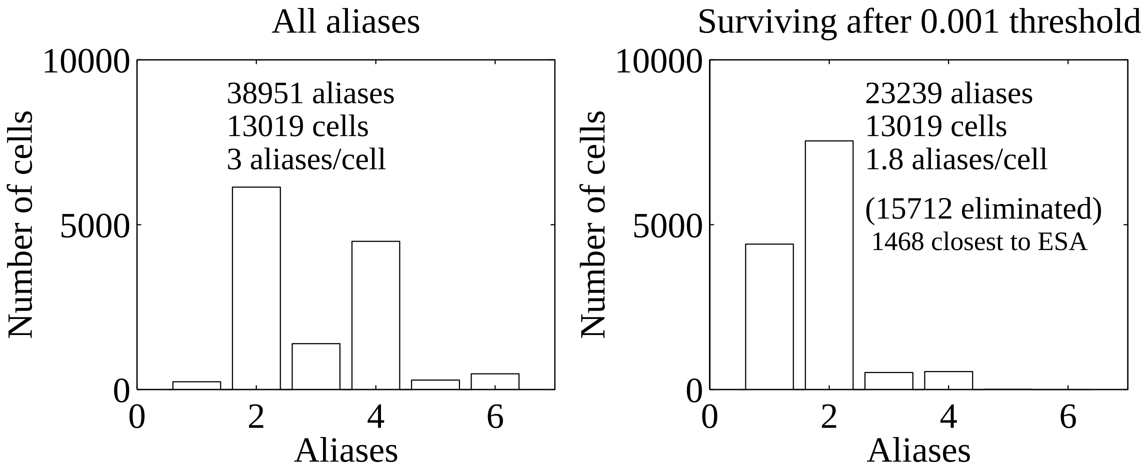


Figure 4.10: *Performance of alias-elimination scheme on actual data taken from the ascending portion of ERS-1 revolution 4452 at 25 km resolution (assuming $K_{pm} = 0$).*

looking at how effective the alias elimination scheme can be on real data. The question looms as to what value to pick for K_{pm} when applying the technique to real data. This is a difficult question for two main reasons. First of all, we have used a simplistic method of introducing GMF modeling error into the measurements by assuming that all measurements used to retrieve the wind are corrupted independently by Gaussian noise. The second reason is that even assuming the independent model is accurate, only rough estimates of K_{pm} are available (Johnson et al., 1996). As was evident in simulation, if we assume that K_{pm} is near 0.20 as estimated by Johnson (1996), then the effectiveness of alias reduction is drastically diminished. As a result of these difficulties we perform alias-elimination on actual data assuming $K_{pm} = 0$, recognizing that caution should be exercised in interpreting the results, especially when the elimination process completely removes all but the most likely wind alias.

Actual Data

In Figure 4.10 the overall results of applying the alias-elimination scheme to actual wind retrieved with the ERS-1 scatterometer assuming $K_{pm} = 0$ is presented. Note that nearly all of the wind vector cells return only one or two wind aliases when the acceptable test size limit is 0.001.

To get a more visual representation of the utility of this elimination process, Figures 4.11 and 4.12 present portions of the swath which generated the results in Figure

4.10. In these figures the right plot shows an along-track segment of the swath with only those wind aliases which whose hypothesis test had a size greater than 0.001. The left plot contains all of the wind aliases for the swath.

It is evident from these figures that the dealiasing step is greatly simplified when hypothesis testing to remove wind aliases is performed. Another interesting feature of these plots is that over many regions only one consistent estimate of the wind field is returned. Admittedly, there are some cases where after hypothesis testing adjacent wind vector cells each return only one wind vector cell which are inconsistent with each other. This results in an unbelievable wind pattern. For an example of this, see the left plot in Figure 4.11. Look at along-track 194, cross-track 17, and compare with along-track 195, cross-track 18. Apparently, the true wind was thrown out as an improbable alias in cell (195,18). Since this behavior occurs at a frequency that is inconsistent with the accepted hypothesis test size (.001), it is evident that the model is in error (*e.g.*, $K_{pm} \neq 0$). On the other hand, for this value of K_{pm}, the elimination scheme was successful at determining a single consistent wind field for numerous regions. This is rather surprising if K_{pm} is truly 0.20 as assumed in Figure 4.13 where there is a dramatic loss in the number of wind aliases eliminated by the scheme.

These results give anecdotal evidence to the idea that GMF modeling error cannot be ignored, yet assuming large, independent noise sources in the measurement process to account for modeling error may be overly cautious. Perhaps a better approach would be to assume some covariance in the GMF modeling error and re-derive the hypothesis test based on such a model.

In the mean time, the hypothesis test derived in this thesis can still be used to eliminate aliases effectively if we assume $K_{pm} = 0$ and then use a more conservative acceptable size on the test for alias elimination (*e.g.*, 1e-5 instead of 1e-3). Alternatively, we can forego an attempt to eliminate all of the aliases and always return at least two wind vectors.

Another possible approach that deserves more attention than can be granted in this thesis is to use measurements from neighboring cells to estimate a single, low-resolution wind vector. This could reduce the noise levels (even for large K_{pm}) so that the hypothesis test would have greater power to eliminate aliases. The wind alias at high-resolution could then be selected as the one closest to the low-resolution wind vector.

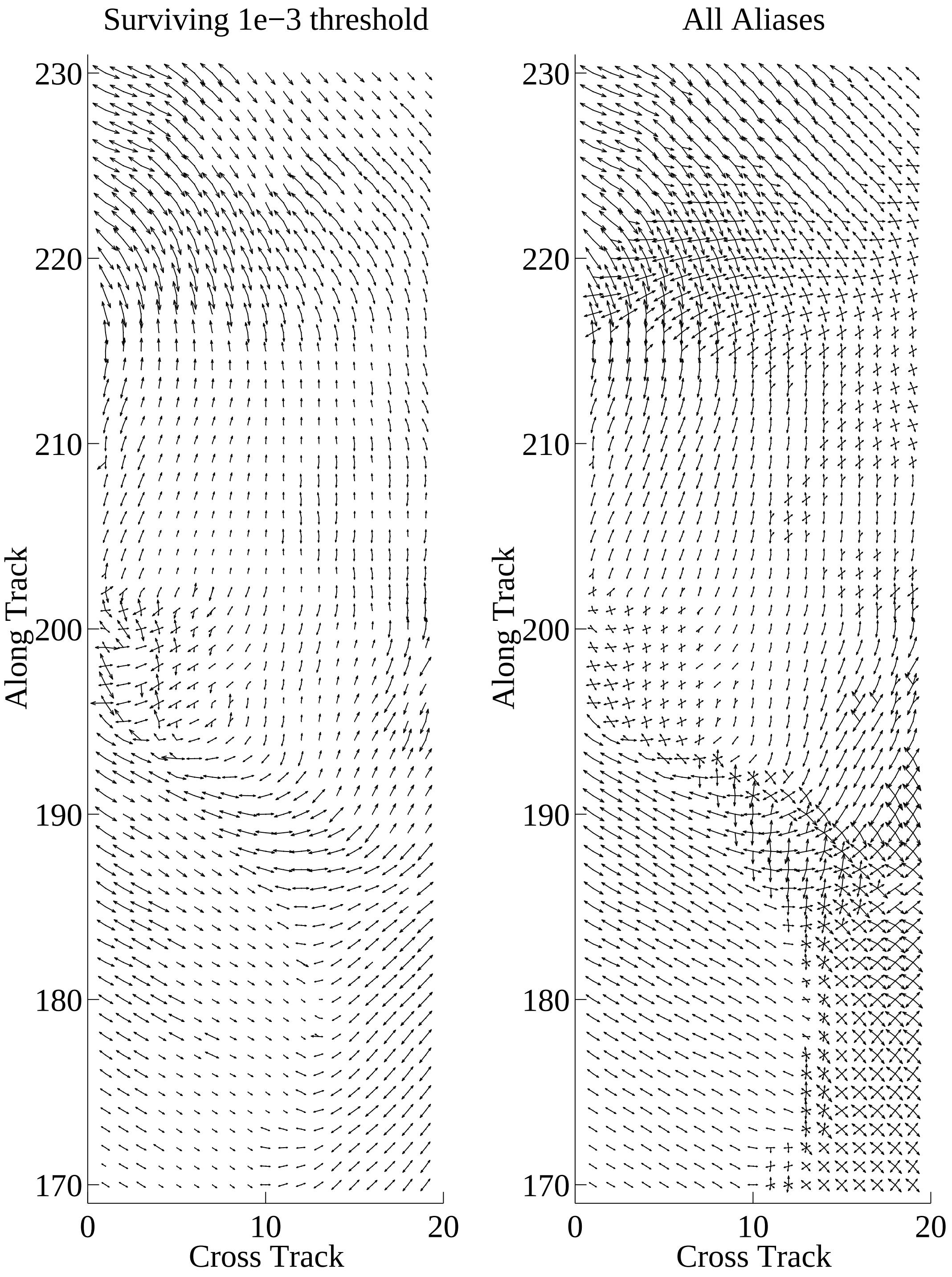


Figure 4.11: *Effectiveness of alias-elimination scheme over a portion of the swath for ERS-1 orbit 4452, ascending (assuming* $K_{pm} = 0$*).*

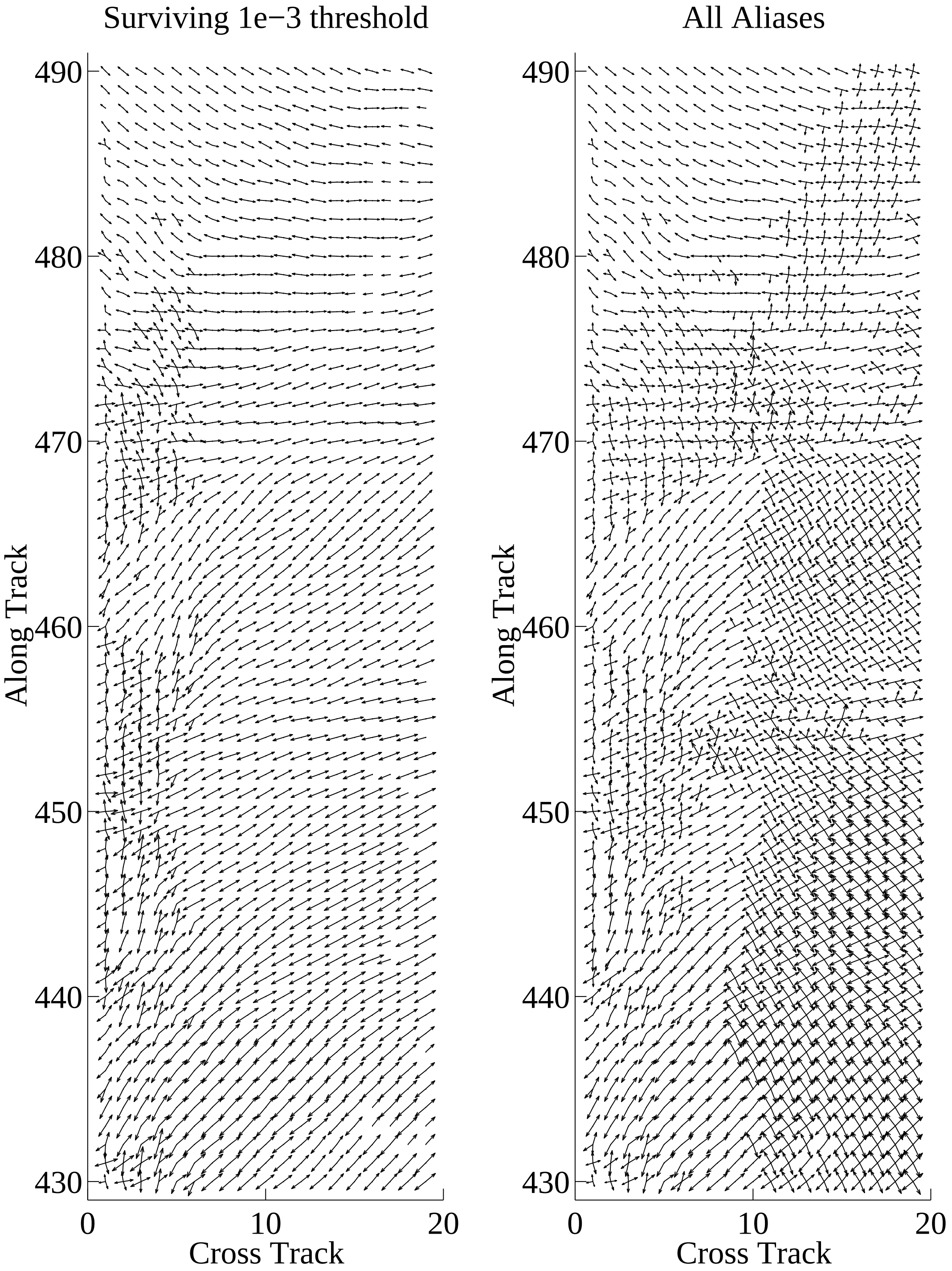


Figure 4.12: *Effectiveness of alias-elimination scheme over a portion of the swath for ERS-1 orbit 4452, ascending (assuming $K_{pm} = 0$).*

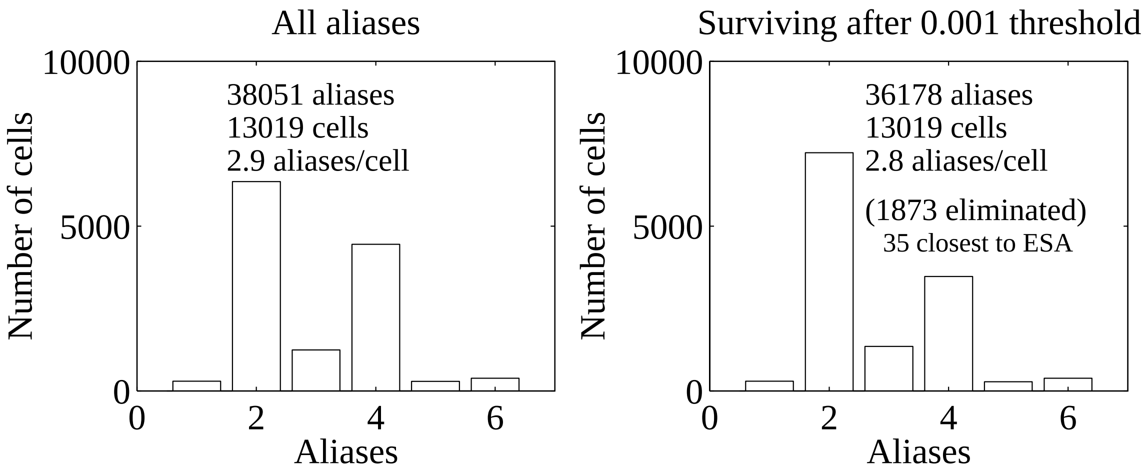


Figure 4.13: *Performance of alias-elimination scheme on actual data taken from the ascending portion of ERS-1 revolution 4452 at 25 km resolution (assuming* $K_{pm} = 0.20$*). ESA winds are those estimated by the European Space Agency which operates ERS-1.*

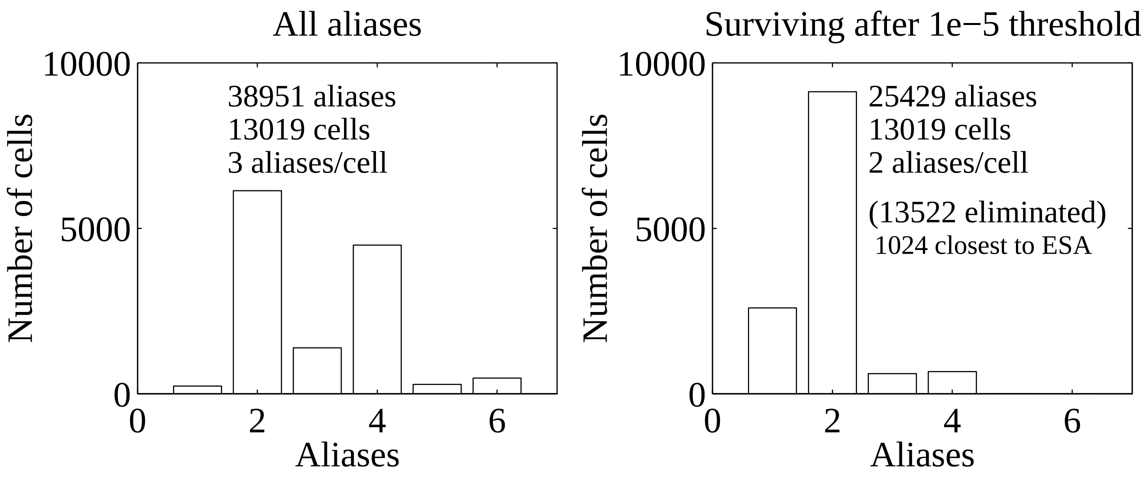


Figure 4.14: *Performance of alias-elimination scheme on actual data taken from the ascending portion of ERS-1 revolution 4452 at 25 km resolution assuming* $K_{pm} = 0$ *and a size threshold of 1e-5.*

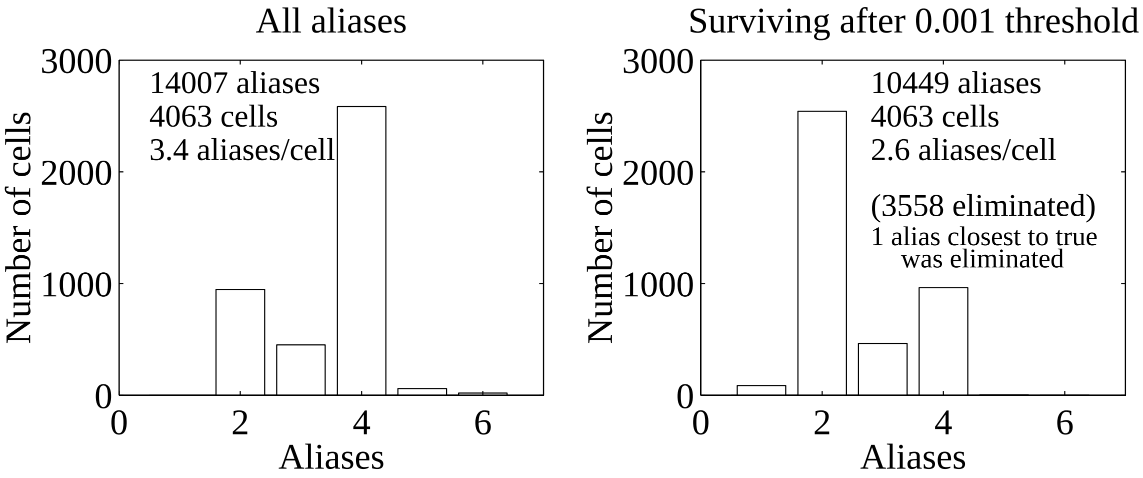


Figure 4.15: *Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information from a simulated NSCAT orbit at 50 km resolution* ($K_{pm} = 0$).

4.7.2 Simulated NSCAT

We expect the results for NSCAT data to be qualitatively similar to those for ERS-1. Since we have only limited simulated NSCAT data available, only two simulation results for NSCAT are given. The result in Figure 4.15 assumes $K_{pm} = 0$, while that of Figure 4.16 assumes $K_{pm} = 0.20$.

4.7.3 Discussion

In this section we have seen that applying the elimination scheme to point-wise retrieved wind fields can greatly ameliorate the dealiasing problem. When the signal to noise ratio is high enough, the scheme can even select a single wind vector reliably. This property of being able to select a single wind vector is even more pronounced when the test is applied to model-based processing as demonstrated in the next section.

4.8 Application to model-based alias elimination

This section is included to give examples that demonstrate how the test developed in this Chapter can be applied to model-based retrieval. The demonstration is with ERS-1 data. The test is applied using the same region of ERS-1 data under

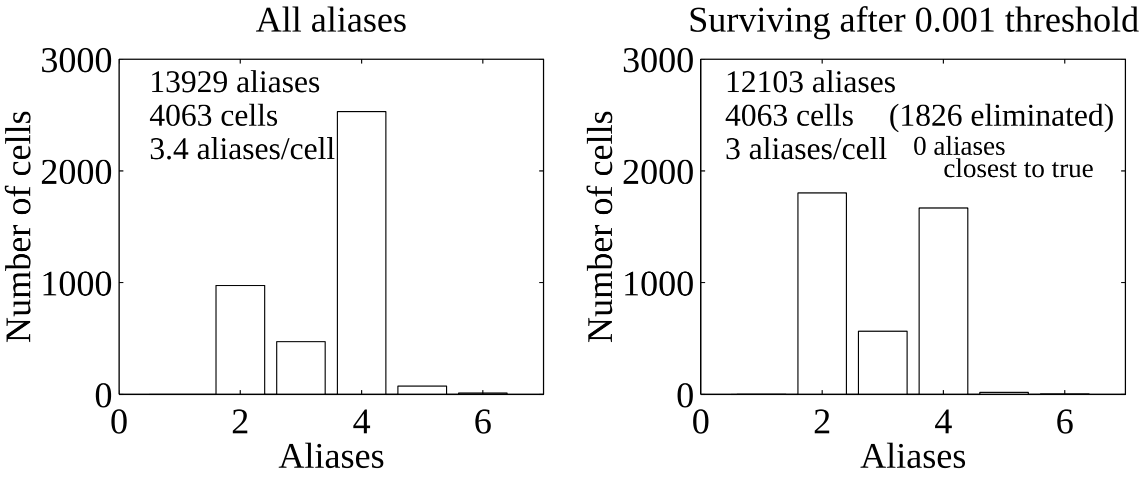


Figure 4.16: *Performance of alias-elimination scheme on simulated data. Retrieval geometry and noise information from a simulated NSCAT orbit at 50 km resolution* ($K_{pm} = 0.2$).

Wind Field	Objective Function	α_0 bound ($s = 1$)	α_0 bound (best s)	α_0
Case 1: $F = I$				
Most likely	-2410.9			
First alias	-2406.5	0.107	.095 ($s = 1.28$)	0.015
Second alias	-2391.6	6.50e-5	4.08e-5 ($s = 1.27$)	3.45e-6
Case 2: $F \neq I$				
First alias	-2372.0			
Second alias	-2359.2	1.71e-3	1.51e-3 ($s = 0.88$)	1.55e-4

Table 4.1: *Objective function values and test results for wind-field aliases in Figures 4.17 and 4.18.* $K_{pm} = 0.13$.

two scenarios. In the first scenario we assume $F = I$, while in the second we use the 22-parameter F matrix derived by Long (1989).

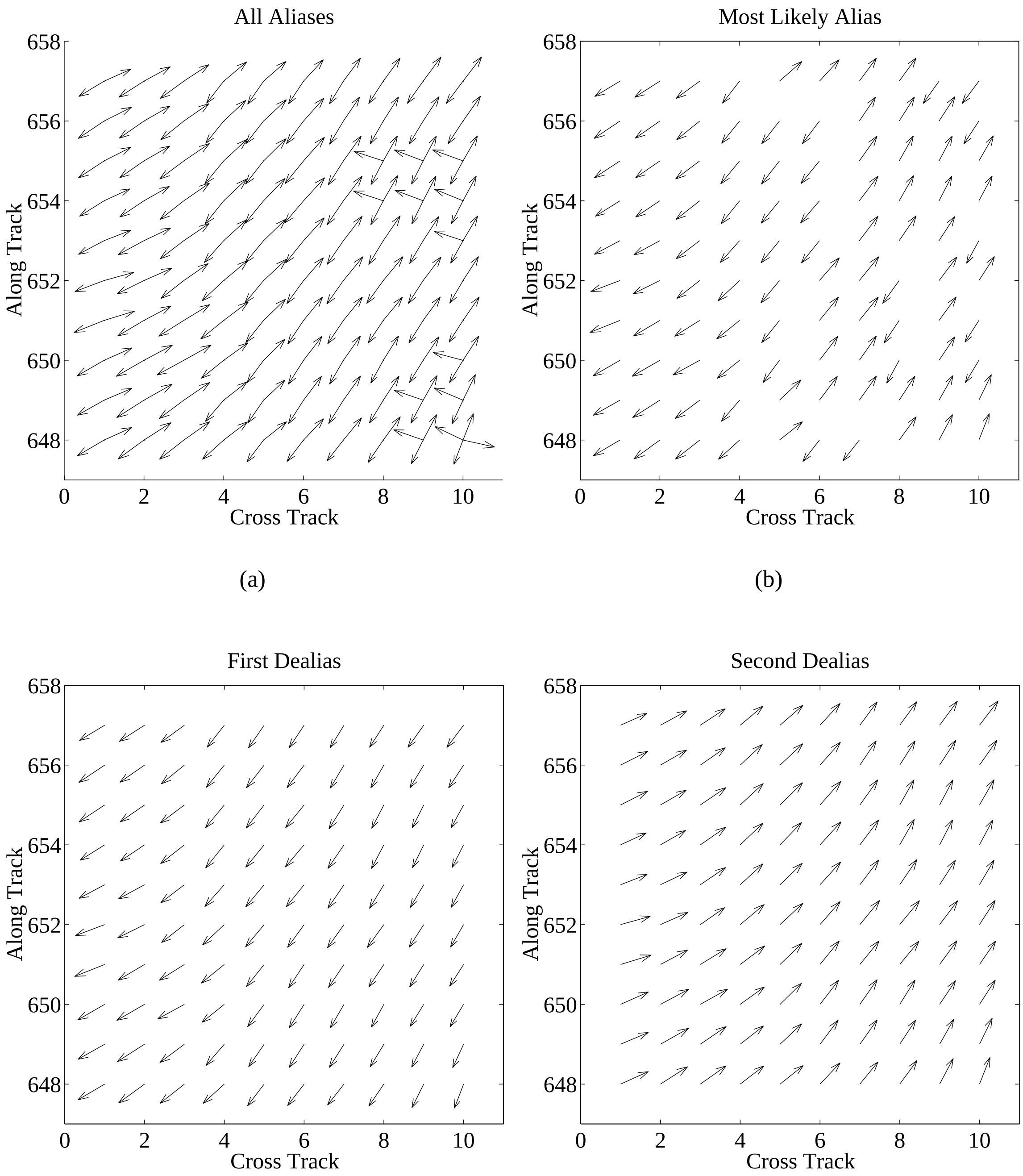
4.8.1 Case 1: $F = I$

The first case considered is model-based retrieval in which the model is the identity matrix, $F = I$. This is nothing more than point-wise retrieval performed over an entire wind field simultaneously with a 200-variable likelihood function. As a result, this large objective function has maxima at all possible combinations of point-wise solutions. To apply the test we compare wind fields giving maxima to the field-likelihood function; each field is found by combining point-wise wind vector solutions found with $K_{pm} = 0.13$. Figure 4.17 shows the selected wind fields along with a plot of all the aliases obtained using point-wise retrieval in a particular 10×10 region.

Table 4.1[4] shows the sizes of the decision rules which eliminate the fields in Figures 4.17(c) and 4.17(d). These values show that it is statistically reasonable to eliminate the second wind-field alias. The most likely wind-field alias (based on the measurements) can be eliminated due to its irregularity, leaving only one realistic wind field. This shows that when all of the measurements over an entire region are combined together, one wind field can often be determined (even with large K_{pm} values).

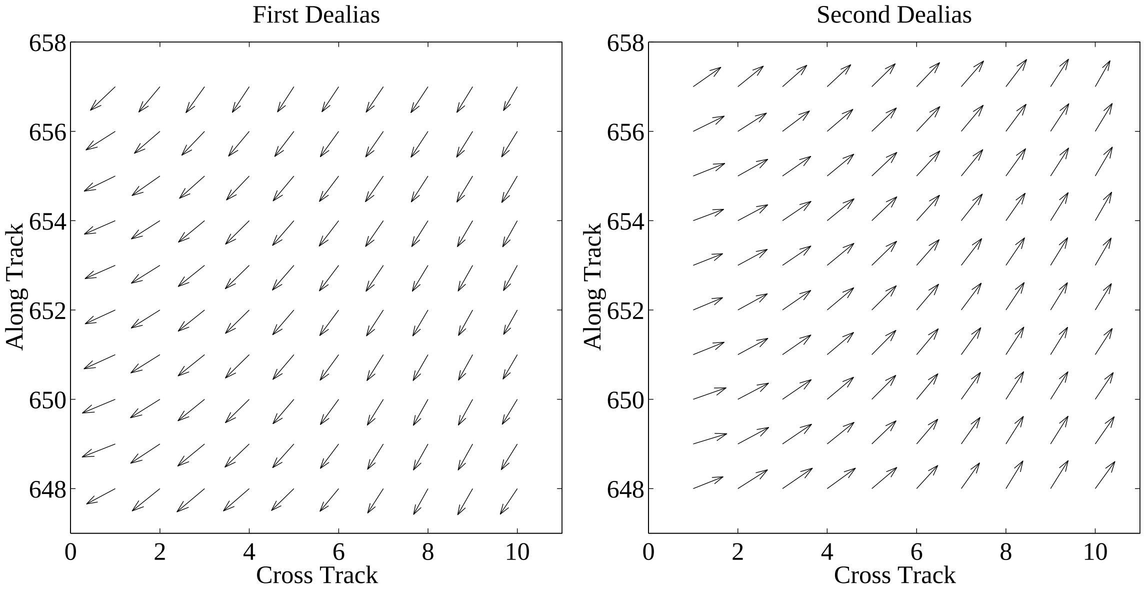
---

[4] Objective function values shown in this table are the values of the negative log-likelihood function with all constants removed.



(a) (b) (c) (d)

Figure 4.17: *(a) All aliases for a wind region in the descending portion of ERS-1 orbit 4448. (b) Wind-field alias constructed from most likely point-wise aliases. (c) Wind-field alias constructed from the median filter solution initialized with the most likely field. (d) Wind-field alias constructed from point-wise solutions closest to 180° from wind vectors in (c).*



(a) (b)

Figure 4.18: *Optimized model-based wind fields: (a) initialized with point-wise field in Figure 4.17(c), (b) initialized with field obtained by rotating all vectors in Figure 4.17(c) by 180°.*

4.8.2 Case 2: $F \neq I$

In the first case no model for the wind field was assumed so that $F = I$. This made the problem a straightforward extension of the point-wise case by avoiding discussion of wind field modeling error. The test can still be applied to model-based retrieved wind where $F \neq I$. However, in interpreting the size calculation, it is important to recall that calculation of the size assumes the estimated wind-field alias is the true wind. If F represents an approximate wind model, then this is not necessarily true. This is the same problem seen in point-wise estimation of using the estimated wind-vector alias instead of the true expected value of the wind-vector alias in calculating α_0. It is compounded here by the additional modeling error possible. As a result, test results should be interpreted more conservatively for model-based retrieval. With this in mind, we give an example of applying the alias elimination scheme to model-based estimation using Long's 22-parameter F matrix. Figure 4.18 shows the wind-field aliases in question and Table 4.1 shows the results of the test. Again, we can choose the first wind-field alias as the solution with confidence.

4.9 Summary

In this Chapter a method of alias elimination which can be used with both point-wise and model-based wind retrieval was developed by applying traditional decision theory. The result is both an integral expression and an upper bound for the size of a decision rule that always throws out the alias under consideration. Applying this decision rule to both ERS-1 and simulated NSCAT data shows that it can be used to effectively eliminate wind aliases. In fact, quite often a single wind field can be selected by applying the rule to wind-fields. In order to apply the results presented here to current point-wise wind retrieval, we suggest using a conservative acceptable size threshold (1e-5) and a small value of K_{pm} (0 to 0.05).

Chapter 5

WIND FIELD MODELING

While the previous two chapters have presented improvements to wind scatterometry that apply to both point-wise and model-based wind retrieval, this chapter is devoted to discussing improvements to model-based wind retrieval. Specifically, in this chapter we discuss alternative wind models to the model developed by Long (1993).

5.1 Motivation

Recall that model-based wind retrieval relies on a model of a discrete wind field, $\mathbf{W} = \mathrm{g}(\mathbf{X})$, where $\mathbf{X}$ is a vector of parameters for the model. To be useful in wind estimation, $\mathbf{X}$ should have fewer elements than $\mathbf{W}$. In addition, the model should ideally fit all observable wind fields. This last idealization is a formidable one given the complexity of wind fields. Long succeeded in deriving a realistic model based on physical principles by using the fact that any vector field can be written in terms of its curl, its divergence, and its boundary conditions. Long discretized this basic relationship and wrote the equation in terms of boundary conditions on the pressure field. In order to reduce the number of parameters in $\mathbf{X}$, he assumed that the curl and divergence of the wind field could both be modeled as low-order bivariate polynomials. This allowed him to write a linear equation for the discrete wind field:

$$\mathbf{W} = F\mathbf{X}. \tag{5.1}$$

The parameter vector $\mathbf{X}$ contains the coefficients of the bivariate polynomials modeling the curl and divergence as well as the boundary conditions on the pressure field. The details of the F matrix will not be given here but can be found in Long (1989). This model is called the Non-parameterized Boundary condition (NB) model. This is in contrast to the Parameterized Boundary Conditions (PBC) model in which the boundary conditions on the pressure field are parameterized according to a polynomial or Fourier model.

Both of these models derived by Long are based on the assumption that the underlying curl and divergence of the wind field can be modeled as a low order polynomial.

A limitation of this model is that this assumption is not directly based on physics other than the known fact that wind has a red spectrum on average (Freilich and Chelton, 1986).

Another short-coming of Long's model is an apparent inconsistency in the model's definition. This inconsistency can be seen in comparing curl and divergence calculated from estimated wind fields using first difference approximations with the assumed polynomial curl and divergence calculated from the estimated parameters. The comparison shows that the two calculations do not return comparable results. Given these difficulties, it seems reasonable to investigate other possible wind field models for use in model-based retrieval.

The purpose of this chapter is to introduce three such models. These models do not seek to be ideal wind models based on physics, but are intended to be practical models that can be tuned to observations. Each model is linear of the form, $\mathbf{W} = F\mathbf{X}$, similar to Long's model. The difference is the definition of the F matrix.

In the following we adopt a slight change to the convention introduced by Long in ordering the elements of the wind field into the vector, $\mathbf{W}$. Namely, the wind-field vector, $\mathbf{W}$, is defined in terms of the rectangular components of each wind velocity:

$$\mathbf{W} = \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix}, \tag{5.2}$$

where $\mathbf{U}$ is a lexicographically-ordered version of the image formed by taking the east, rectangular component of each wind vector in the field. $\mathbf{V}$ is defined similarly as the lexicographically-ordered version of the north-rectangular-component image. The lexicographic ordering is done by varying cross-track most rapidly. Long ordered the wind vectors by varying along-track most rapidly. The difference is insignificant provided it is consistent.

5.2 Wind Vector Space and Model Sub-Space

The central idea to wind modeling in this chapter is to change the paradigm used in looking for a model. Instead of seeking a physics-based model, we use a signal-processing approach. Consider the wind field vector, $\mathbf{W}$, describing the wind at every cell in an $M \times N$ region as an element in a $2MN$-dimensional vector space. Define a subset of the elements of this vector space consisting of observable winds, or ''physical'' winds. It is clear that such winds occupy a proper subset of this space as a randomly chosen wind-field vector does not ‘‘look’’ like real wind. In order to approximate this subset in a simple and useful manner we use a vector sub-space spanned by a truncated set of M_x basis vectors. As a result, winds in the sub-space can be written as $\mathbf{W}_F = F\mathbf{X}$. The columns of F comprise the basis set whose span approximates the subset containing observable winds.

Given a true wind-field vector, $\mathbf{W}$, the closest vector in the sub-space spanned by the columns of F can be found provided a definition of distance is given. For this purpose, the definition of distance on the wind-field vector space is the standard Euclidean distance:

$$d^2(\mathbf{W}_F, \mathbf{W}) = (\mathbf{W} - \mathbf{W}_F)^T(\mathbf{W} - \mathbf{W}_F). \tag{5.3}$$

With this definition for the distance operator, finding the $\mathbf{X}$ vector that minimizes the distance between $\mathbf{W}$ and $\mathbf{W}_F = F\mathbf{X}$ for a given F is straightforward. This is a common problem in signal-processing applications and the result can be found in many texts (*e.g.* page 365 of Scharf). The result is

$$\mathbf{X} = (F^TF)^{-1}F^T\mathbf{W} \equiv F^\dagger\mathbf{W}, \tag{5.4}$$

where we have defined $F^\dagger$ as the pseudo-inverse of F. Thus we can write the projection of the wind field onto the space spanned by the columns of F as

$$\mathbf{W}_F = FF^\dagger\mathbf{W}. \tag{5.5}$$

In comparing models it is convenient to use a distance metric normalized by the length of the wind-field vector and so we define the normalized mean square error of a projected wind-field vector as

$$\begin{aligned} E_F^2 &= \frac{d(\mathbf{W}, \mathbf{W}_F)}{d(\mathbf{W}, \mathbf{0})}, && (5.6) \\ &= \frac{(\mathbf{W} - \mathbf{W}_F)^T(\mathbf{W} - \mathbf{W}_F)}{\mathbf{W}^T\mathbf{W}}, \\ &= \frac{\mathbf{W}^T(I - FF^\dagger)^T(I - FF^\dagger)\mathbf{W}}{\mathbf{W}^T\mathbf{W}}, && (5.7) \\ & && (5.8) \end{aligned}$$

where I is a $2MN \times 2MN$ identity matrix. If the columns of F are orthonormal then this expression can be simplified further since $F^\dagger = F^T$:

$$E_F^2 = \frac{\mathbf{W}^T(I - FF^T)\mathbf{W}}{\mathbf{W}^T\mathbf{W}}. \tag{5.9}$$

With the concept of a model sub-space developed, wind field modeling can now be stated as determining the fewest number of basis vectors that reduces the normalized error of the projected wind field to an acceptable value. In the next three sections we develop three wind models that attempt to do this. First we describe a data-driven model where the basis vectors of the Karhunen-Loeve transform are selected as the basis vectors for the model. The second model assumes that the wind's rectangular components can be written using a low-order polynomial. The third model assumes that these rectangular components can be approximated well with a low-order Fourier series. The last section of this chapter compares these models with the (PBC) model developed by Long (1989).

5.3 Karhunen-Loeve Basis Wind-Field Vectors

Given a wind-field vector space with elements defined in terms of its rectangular components we seek a truncated basis, F, with as few basis vectors as possible while keeping the average error between the true wind field and the model wind field, $\mathbf{W}_F$, small. An alternative way to state the problem is that we seek a model matrix, F, with M_x basis vectors that minimizes the average error between $\mathbf{W}$ and $\mathbf{W}_F$. It is a well-known property of the Karhunen-Loeve transform that for any value of M_x the matrix F which minimizes this basis-restriction error contains columns equal to the largest M_x eigenvectors of the autocorrelation matrix of $\mathbf{W}$. These are the Karhunen-Loeve (K-L) basis vectors of $\mathbf{W}$. In the following a general description detailing how the columns of F can be calculated from a data-set of wind fields is described.

5.3.1 Determination of the K-L F matrix

For a given $M \times N$ region size, the columns of F are the M_x eigenvectors corresponding to the largest eigenvalues of the autocorrelation matrix, R:

$$R = E\left[\mathbf{W}\mathbf{W}^T\right] \tag{5.10}$$

Since there is not a functional form for R, the autocorrelation matrix must be estimated. We do this by calculating the sample average of $\mathbf{W}\mathbf{W}^T$.

Without true wind fields to use in estimating the autocorrelation matrix, alternate wind fields must be employed. One possibility is to use numerical weather prediction wind fields such as supplied by the European Center for Mid-range Weather Forecasting (ECMWF). These fields are described by Long (1989). One concern with using these fields to develop the K-L model is that since the K-L basis vectors are completely data driven, the result will be strongly influenced by the model used by ECMWF. Another concern is that these wind fields are at a resolution of approximately 200 km by 200 km rather than the typical 25 km or 50 km resolution at which wind is retrieved with scatterometers.

As a result of these concerns, we use point-wise-retrieved wind fields with the ambiguities removed using a median-filter to estimate the autocorrelation matrix. Even though these fields have ambiguity removal error and estimation uncertainty, they represent the best high-resolution data set of true wind fields available.

The estimate of R is found by computing the sample average of the correlation matrix for L point-wise-retrieved wind fields:

$$R \approx \frac{1}{L} \sum_{n=1}^{L} \hat{\mathbf{W}}_n \hat{\mathbf{W}}_n^T. \tag{5.11}$$

With an estimate of R, the model matrix, F, can be found by solving the eigenvector-eigenvalue equation:

$$RF = F\Lambda, \tag{5.12}$$

where Λ is a diagonal matrix of the largest M_x eigenvalues of R:

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{M_x} \end{bmatrix}, \tag{5.13}$$

with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{M_x}$. Since R is symmetric, a solution exists such that $F^T F = I$, where I is an $M_x \times M_x$ identity matrix. We choose this solution as the model matrix, F, of order M_x. Note that since F consists of orthonormalized eigenvectors, projections are particularly easy to compute:

$$\mathbf{W}_F = FF^T\mathbf{W}. \tag{5.14}$$

For each column of F, there is an associated eigenvalue with a useful interpretation. This eigenvalue can be interpreted as the average energy of that column vector in the data set used to generate R. To see this, note

$$\begin{aligned}
\Lambda &= F^T R F, && (5.15)\\
&= F^T \frac{1}{L}\left(\sum_{n=1}^{L} \hat{\mathbf{W}}_n \hat{\mathbf{W}}_n^T\right) F, && (5.16)\\
&= \frac{1}{L}\sum_{n=1}^{L} F^T \hat{\mathbf{W}}_n (F^T \hat{\mathbf{W}}_n)^T, && (5.17)\\
&= \frac{1}{L}\sum_{n=1}^{L} \mathbf{X}_n \mathbf{X}_n^T. && (5.18)
\end{aligned}$$

Thus, if $X_{k,n}$ represents the coefficient of the kth basis vector for the nth wind field in the data set then

$$\lambda_k = \frac{1}{L}\sum_{n=1}^{L} X_{k,n}^2. \qquad (5.19)$$

The larger the eigenvalue, the more important the eigenvector as a basis vector for the space spanned by the data set. This further justifies using the first M_x ordered eigenvectors computed from our best estimate of ‘‘real’’ wind as basis vectors of the wind-field.

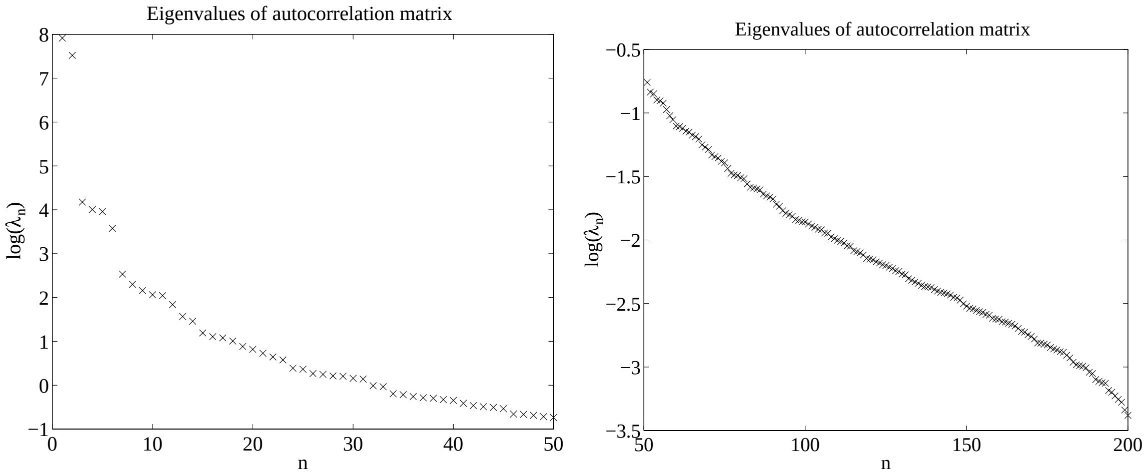
In the next section we apply the previous development to compute the F matrix applicable to a 10×10 region using point-wise wind fields from the ERS-1 scatterometer.

5.3.2 10×10 K-L basis wind fields

For ERS-1 the swath is 500 km wide with 19 wind vector cells at 25 km resolution. As a result, a 10×10 model region will cover the swath with two regions. It is therefore, a useful region size for ERS-1 model-based retrieval.

In this section we explore the K-L basis wind-field vectors computed using ERS-1 point-wise data by plotting several of the vectors. These basis vectors were derived using point-wise retrieval and a median-filter algorithm. The autocorrelation matrix was calculated using Eq. (5.11) and more than 200,000 wind fields retrieved over the Pacific Ocean for 113 ERS-1 revolutions coming from the last week of June 1992, the last week of September 1992, and the first part of October 1992. All eigenvectors and corresponding eigenvalues of R were then found. To form the F matrix for a particular truncated basis order, M_x eigenvectors corresponding to the largest eigenvalues are selected and placed in the columns of F.

In order to determine how many basis wind-field vectors to use in the model matrix F, it is helpful to look at a plot of the eigenvalues of the autocorrelation matrix.



(a) Vectors 1 to 50

(b) Vectors 51 to 200

Figure 5.1: *Logarithmic plot of the eigenvalues of the* 10×10 *autocorrelation matrix. (Note the different vertical scales.)*

Figure 5.1 shows such a plot. Notice that on the logarithmic scale natural breaks are evident which signal important groups of basis vectors for the wind field. These groups and the values of the eigenvalues can be used as a guide for determining how many basis vectors to use in the F matrix.

To give a general idea of what the wind field basis vectors look like, Figures 5.2 to 5.4 show the first eighteen basis wind-field vectors re-mapped into a wind field for plotting. It is interesting that the K-L basis selection procedure pulls out common wind field features as basis wind-field vectors. For example, the first two basis fields simply describe a mean wind field. Notice that this mean is the most dominant feature in most wind fields as evidenced by the large break between $n = 2$ and $n = 3$ in the eigenvalue plot in Figure 5.1(a).

Basis fields 3 to 6 are the next important group for wind fields according to the eigenvalues. Interestingly these correspond to a cyclone (3), two basis fields describing col points (4 and 5), and a first-order divergent field. These are physically reasonable fields. The relative values of the eigenvalues also suggests that many wind fields can be modeled reasonably with just these six basis fields. The next 12 basis vectors show higher order variability that can occur within the region.

While the K-L basis vectors are theoretically the best ones to choose for

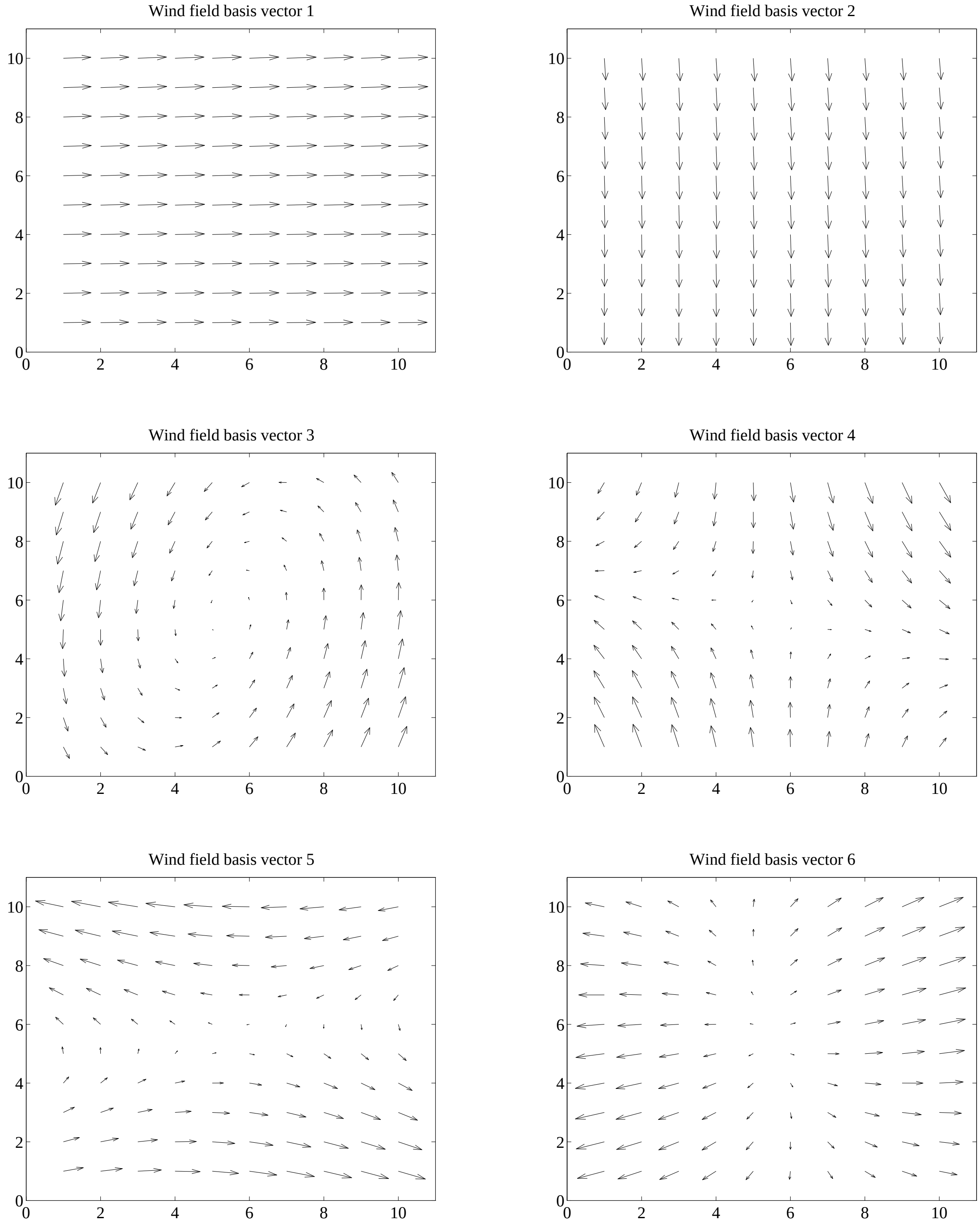


Figure 5.2: *Wind field basis vectors 1–6.*

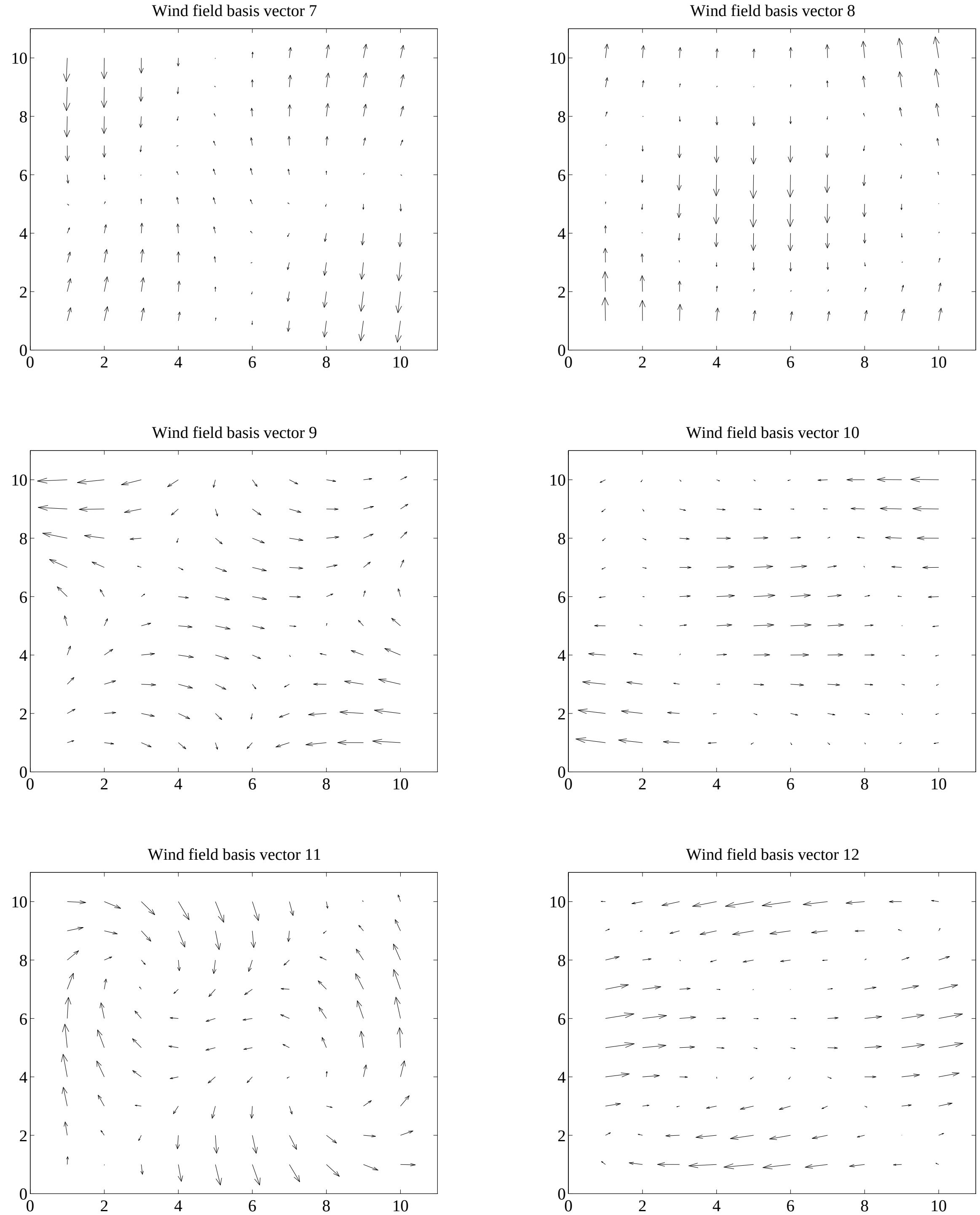


Figure 5.3: *Wind field basis vectors 6–12.*

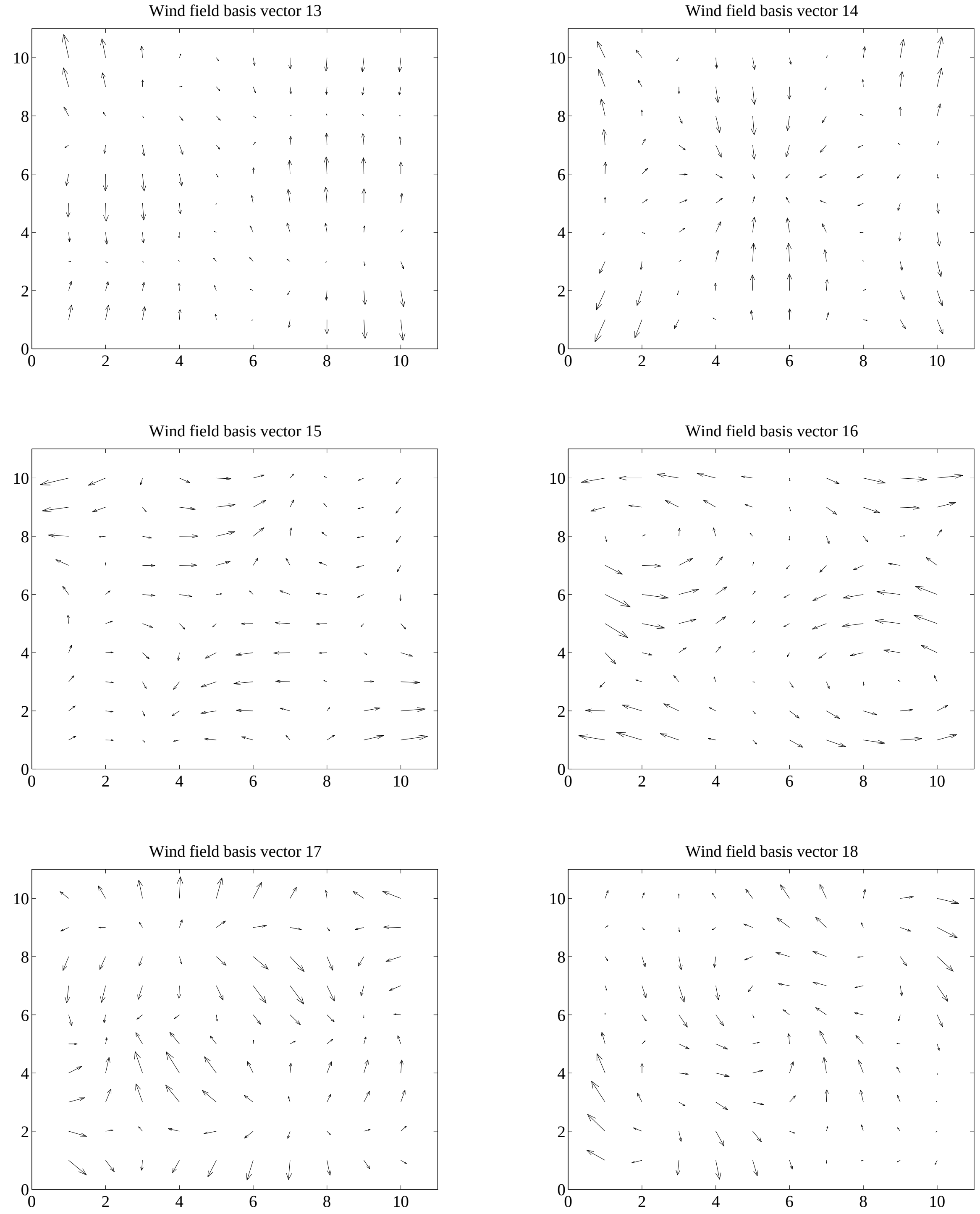


Figure 5.4: *Wind field basis vectors 12–18.*

minimizing the average error in the fewest number of unknowns, it has several disadvantages that encourage investigation of other wind-field basis vectors. One of these disadvantages is that the basis vectors must be recomputed if the region size is changed. A related disadvantage is that since we do not have true wind fields from which to estimate the autocorrelation matrix, our result will be tainted by the uncertainty in the data set used to estimate R. While averaging instills confidence in the first several basis vectors, it is difficult to determine at what point the basis vectors are dominated by noise in the sample data. Because of these limitations it is useful to discuss other possible models for wind fields.

5.4 Legendre Polynomial Basis Functions

Both the NB and PBC models of the wind field use a low-order bi-variate polynomial to model the curl and divergence of a wind field (Long, 1989). If the curl and divergence are modeled as low order polynomials, then the wind field itself will exhibit low order polynomial behavior. Recognizing the success of the NB and PBC models in modeling wind fields, we postulate that each rectangular component of the wind itself can be modeled as a low order polynomial. In order to avoid numerical difficulties in inverting matrices constructed from ordinary polynomials, we use orthogonal, Legendre polynomials as basis functions for the u and v component images of the wind field. In this section, a derivation of the F matrix useful for representing the wind field with Legendre, orthogonal polynomials is given.

In Long's wind field model, the curl and divergence are represented as low-order polynomials and a basis set consisting of $\{x^n y^n\}$ is used to represent them (1993). These non-orthogonal polynomials are sufficient for low order polynomials; however, for polynomial orders of four or more, these polynomials give an F matrix that is badly conditioned and prone to error in numerical pseudo-inverse computation. Since higher-order polynomials may need to be used when modeling the wind itself as polynomials, the evenly-weighted, orthogonal, Legendre polynomials are used as basis functions.

Legendre polynomials are defined on the interval $[-1, 1]$ according to the following sum:

$$\phi_n(x) = \frac{1}{2^n} \sum_{m=0}^{n/2} (-1)^m \binom{n}{m} \binom{2n-2m}{n} x^{n-2m} \qquad x \in [-1, 1]. \tag{5.20}$$

For an $M \times N$ region, the sample row i must be mapped from integers in $[1, M]$ to the interval of definition of the polynomials $[-1, 1]$ The column index, j must be similarly mapped from integers in $[1, N]$ to $[-1, 1]$. This is accomplished by the sampling function:

$$x_{i,M} = \frac{2(i-1)}{M-1} - 1 \qquad k \in [1, M]. \tag{5.21}$$

Using this function the components of the wind field at each grid location of an $M \times N$ region can be expressed using Legendre polynomials as

$$U_{i,j} = \sum_{m=0}^{M_u} \sum_{n=0}^{M_u - m} u_{m,n} \phi(x_{i,M}) \phi(x_{j,N}), \tag{5.22}$$

$$V_{i,j} = \sum_{m=0}^{M_v} \sum_{n=0}^{M_v - m} v_{m,n} \phi(x_{i,M}) \phi(x_{j,N}), \tag{5.23}$$

so that the indices in the sum satisfy $m + n \leq M_u$ and $m + n \leq M_v$ respectively.

When the matrices U and V are row-scanned into column vectors $\overline{U}$ and $\overline{V}$, the relations can be rewritten as

$$\overline{U} = \sum_{m=0}^{M_u} \sum_{n=0}^{M_u} u_{m,n} Q_{m,n}, \tag{5.24}$$

$$\overline{V} = \sum_{m=0}^{M_v} \sum_{n=0}^{M_v} v_{m,n} Q_{m,n}, \tag{5.25}$$

where the $Q_{m,n}$ are MN element (column) vectors whose kth element, $q_{m,n}^{k}$, is given by

$$q_{m,n}^{k} = \phi_m(x_{\text{row}(k),M}) \phi_n(x_{\text{col}(k),N}), \tag{5.26}$$

$$\text{row}(k) = \left\lfloor \frac{k-1}{N} \right\rfloor + 1, \tag{5.27}$$

$$\text{col}(k) = \text{mod}(k-1, N) + 1, \tag{5.28}$$

where $\lfloor x \rfloor$ is defined as the greatest integer less than or equal to x.

In order to convert the above expressions for $\overline{U}$ and $\overline{V}$ into one matrix expression $\mathbf{W} = F\mathbf{X}$, it is necessary to construct a $MN \times (M_u + 2)(M_u + 1)/2$ matrix Q_u and a $MN \times (M_v + 2)(M_v + 1)/2$ matrix Q_v from the $Q_{m,n}$ vectors. From these matrices the $2MN \times (M_u + 2)(M_u + 1)/2 + (M_v + 2)(M_v + 1)/2$ matrix F can be constructed. The details are

$$
\begin{aligned}
W &= \begin{bmatrix} \overline{U} \\ \overline{V} \end{bmatrix}, & (5.29)\\
Q_u &= \left[Q_{0,0}|Q_{0,1}|\cdots|Q_{m,n}|\cdots|Q_{M_u,0}\right], & (5.30)\\
Q_v &= \left[Q_{0,0}|Q_{0,1}|\cdots|Q_{m,n}|\cdots|Q_{M_v,0}\right], & (5.31)\\
F &= \begin{bmatrix} Q_u & 0 \\ 0 & Q_v \end{bmatrix}. & (5.32)
\end{aligned}
$$

The order of the columns of Q_u and Q_v is important only in determining where the parameters $u_{m,n}$ and $v_{m,n}$ go in the X vector. Using the above equations, the F matrix can be easily constructed for any model order M_u, M_v and any region size $M \times N$. For a MATLAB implementation of the above description see Appendix F.

5.5 Fourier-series wind model

Investigators have noted that the rectangular components of mesoscale wind have average power spectral densities that follow an inverse power law where the power is approximately 2 (Freilich and Chelton, 1986). With this in mind, it is reasonable to develop a wind field model that uses Fourier basis functions.

The development of the Fourier F matrix is similar to the development of the Legendre F matrix. We write the (i, j) elements of th U and V images as

$$
\begin{aligned}
U_{ij} &= u^c_{0,0} + \sum_{m=0}^{M_u} \sum_{n=0}^{M_u-m} \left[u^s_{m,n} \sin\left(\frac{mi\pi}{M} + \frac{nj\pi}{N}\right) + u^c_{m,n} \cos\left(\frac{mi\pi}{M} + \frac{nj\pi}{N}\right)\right] & (5.33)\\
V_{ij} &= v^c_{0,0} + \sum_{m=0}^{M_v} \sum_{n=0}^{M_v-m} \left[v^s_{m,n} \sin\left(\frac{mi\pi}{M} + \frac{nj\pi}{N}\right) + v^c_{m,n} \cos\left(\frac{mi\pi}{M} + \frac{nj\pi}{N}\right)\right], & (5.34)
\end{aligned}
$$

where the indices for the summations also satisfy $0 < m + n \leq M_u$ and $0 < m + n \leq M_v$ respectively.

When these are row-scanned into column vectors, $\overline{U}$ and $\overline{V}$, we can write

$$
\begin{aligned}
\overline{U} &= \sum_{m=0}^{M_u} \sum_{n=0}^{M_u-m} u^c_{m,n} Q^c_{m,n} + \sum_{m=0}^{M_u} \sum_{n=0}^{M_u-m} u^s_{m,n} Q^s_{m,n}, & (5.35)\\
\overline{V} &= \sum_{m=0}^{M_v} \sum_{n=0}^{M_v-m} v^c_{m,n} Q^c_{m,n} + \sum_{m=0}^{M_v} \sum_{n=0}^{M_v-m} v^s_{m,n} Q^s_{m,n}, & (5.36)
\end{aligned}
$$

where $Q^c_{m,n}$, and $Q^s_{m,n}$ are MN element (column) vectors whose kth elements ${}_kq^c_{m,n}$, and ${}_kq^s_{m,n}$ are given by

$$
\begin{aligned}
{}_kq^c_{m,n} &= \cos\left(\frac{m\text{row}(k)\pi}{M} + \frac{n\text{col}(k)\pi}{N}\right), && (5.37)\\
{}_kq^s_{m,n} &= \sin\left(\frac{m\text{row}(k)\pi}{M} + \frac{n\text{col}(k)\pi}{N}\right), && (5.38)\\
\text{row}(k) &= \left\lfloor \frac{k-1}{N} \right\rfloor + 1, && (5.39)\\
\text{col}(k) &= \text{mod}(k-1, N) + 1. && (5.40)
\end{aligned}
$$

Note that $Q^s_{0,0}$ is a vector of all zeros.

In order to convert the above expressions for $\overline{U}$ and $\overline{V}$ into one matrix expression $W = FX$, it is necessary to construct a $MN \times (M_u+1)(M_u+2)-1$ matrix Q_u and a $MN \times (M_v+1)(M_v+2)-1$ matrix Q_v from the Q^s_m and Q^c_m vectors. From these matrices the $2MN \times (M_u+1)(M_u+2)+(M_v+1)(M_u+2)-2$ matrix, F, can be constructed. The details are

$$
\begin{aligned}
W &= \begin{bmatrix} \overline{U} \\ \overline{V} \end{bmatrix}, && (5.41)\\
Q_u &= \left[Q^c_{0,0}|Q^c_{0,1}|Q^s_{0,1}|\cdots|Q^c_{m,n}Q^s_{m,n}|\cdots|Q^c_{M_u,0}|Q^s_{M_u,0}\right], && (5.42)\\
Q_v &= \left[Q^c_{0,0}|Q^c_{0,1}|Q^s_{0,1}|\cdots|Q^c_{m,n}Q^s_{m,n}|\cdots|Q^c_{M_v,0}|Q^s_{M_v,0}\right], && (5.43)\\
F &= \begin{bmatrix} Q_u & 0 \\ 0 & Q_v \end{bmatrix}. && (5.44)
\end{aligned}
$$

The order of the columns of Q_u and Q_v determines where the parameters $u^c_{m,n}$, $u^s_{m,n}$, $v^c_{m,n}$ and $v^s_{m,n}$ go in the X vector. Using the above equations, the F matrix can be easily constructed for any model order M_u, M_v and any region size $M \times N$. For a MATLAB implementation of the above description see Appendix F.

5.6 Comparisons

In the following model comparisons, three types of fit error will be reported: NRSS vector error, RMS speed error, and RMS direction error. NRSS vector error is a descriptive acronym for the normalized wind-field vector error defined in Eq. (5.7). RMS speed error and RMS direction error are the square-root of the mean-squared speed and direction error respectively.

To compare the models we first show two examples of ‘‘true’’ wind fields projected onto the sub-spaces spanned by the different model matrices. Each true wind field is taken from ERS-1 point-wise-retrieved winds, and the projections are calculated using Eq. (5.5) and the different F matrices.

The first true wind field is shown in Figure 5.5 along with projected wind fields using the PBC and two Karhunen-Loeve (K-L) models. The order of the PBC model was selected to have 22 unknown model parameters (corresponding to 22 basis vectors). The two K-L models were selected to have 22 and 6 basis vectors.

Figure 5.6 shows two Legendre and two Fourier model fits to this same true wind field. The model orders for the first Legendre model were selected to be $M_u = 3, M_v = 3$ to give 20 basis vectors. The model orders for the second Legendre model were selected to be $M_u = 1, M_v = 1$ to give 6 basis vectors. The model orders for the two Fourier models were selected to be $M_u = 2, M_v = 2$, and $M_u = 1, M_v = 1$ to give 22 and 10 basis vectors respectively.

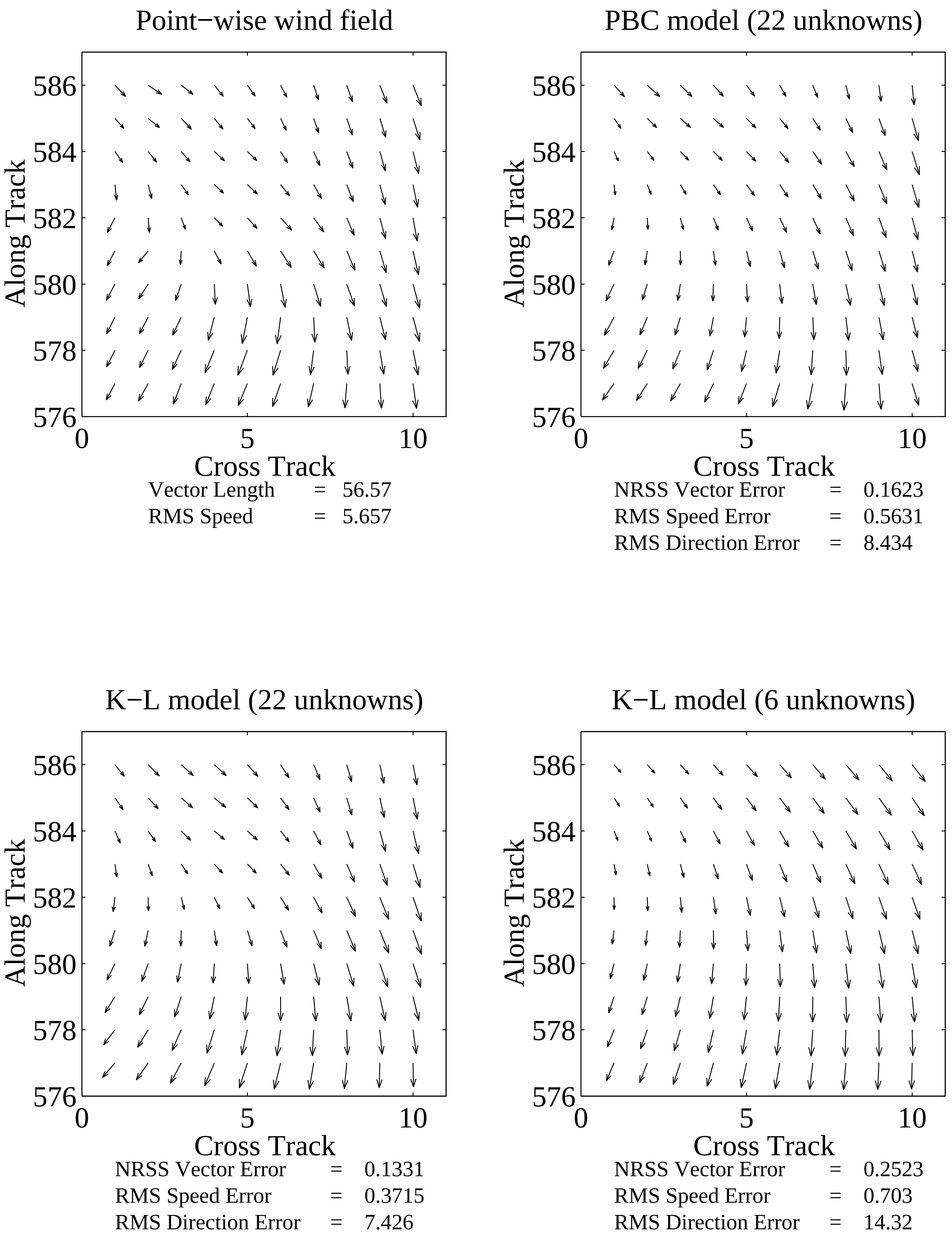
The second true wind field and its projections are shown in Figures 5.7 and 5.8. From these figures it is apparent that the projection error of all four models is similar for these particular wind fields.

In order to further evaluate the different wind models suggested above we calculate the normalized fit error (NRSS) to several thousand simulated wind fields. The simulated wind fields were generated using European Center for Medium-Range Weather Forecasting (ECMWF) numerical weather predication winds as described by Long (1989, p. 215). The average NRSS error was computed using several different model orders for each model. Figure 5.9 and Table 5.1 shows the average results for 19,197 simulated wind fields.

In addition, RMS speed and direction error were calculated to give a general idea of the possible error in speed and direction for a particular wind vector cell when using a specific wind model. The results of these calculations for the same simulated wind fields are shown in Figures 5.10 and 5.11 and Tables 5.2 and 5.3

5.7 Summary

In this chapter, three wind field models were developed as alternatives to the parameterized boundary condition model introduced by Long. It was found that all of the models have similar projection error for a given number of unknowns, although the



Point−wise wind field

Vector Length = 56.57
RMS Speed = 5.657

PBC model (22 unknowns)

NRSS Vector Error = 0.1623
RMS Speed Error = 0.5631
RMS Direction Error = 8.434

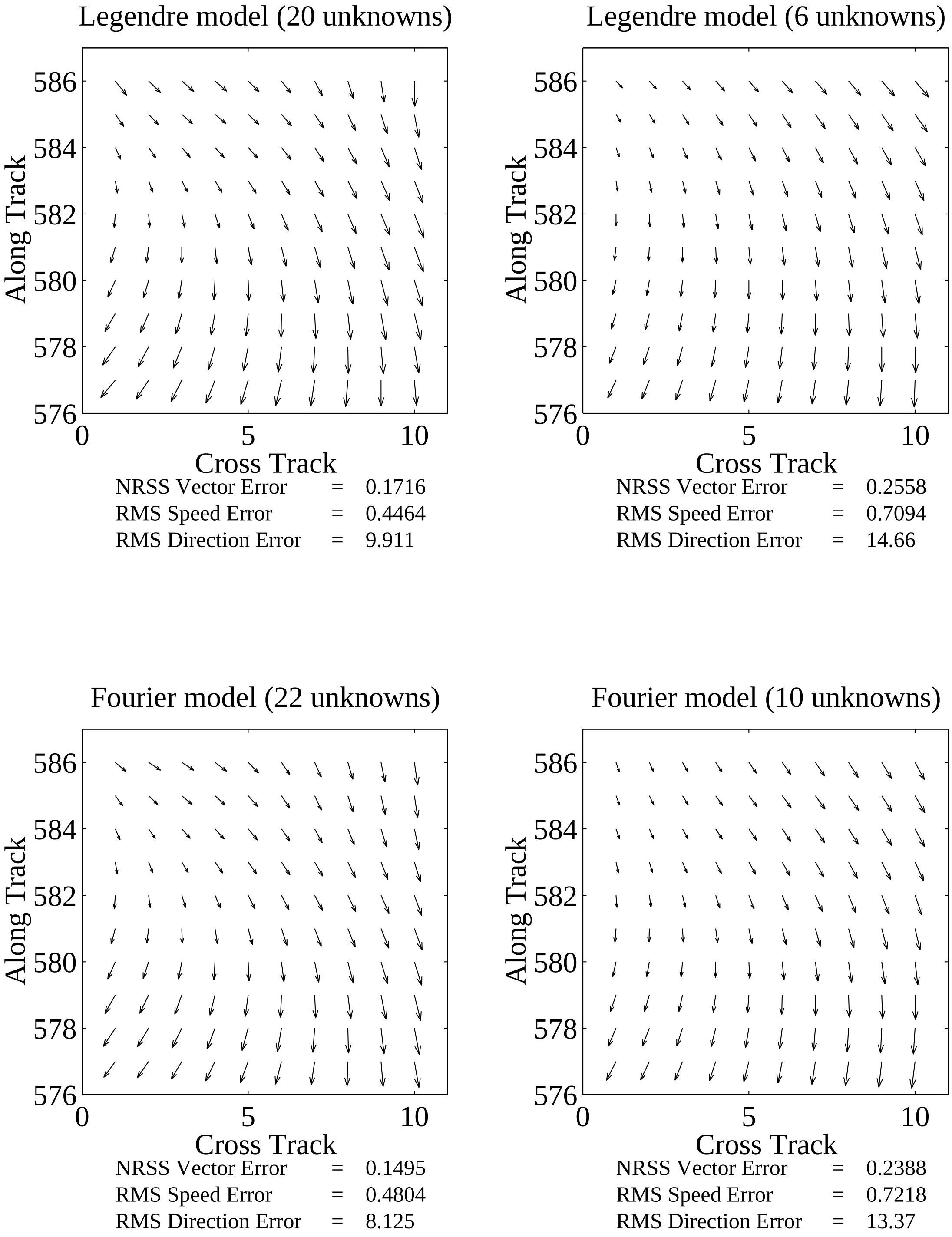
K−L model (22 unknowns)

NRSS Vector Error = 0.1331
RMS Speed Error = 0.3715
RMS Direction Error = 7.426

K−L model (6 unknowns)

NRSS Vector Error = 0.2523
RMS Speed Error = 0.703
RMS Direction Error = 14.32

Figure 5.5: *First example of projecting a wind field onto PBC and Karhunen-Loeve models. Wind-field taken from ascending portion of ERS-1 revolution 4452.*



Legendre model (20 unknowns)

NRSS Vector Error = 0.1716
RMS Speed Error = 0.4464
RMS Direction Error = 9.911

Legendre model (6 unknowns)

NRSS Vector Error = 0.2558
RMS Speed Error = 0.7094
RMS Direction Error = 14.66

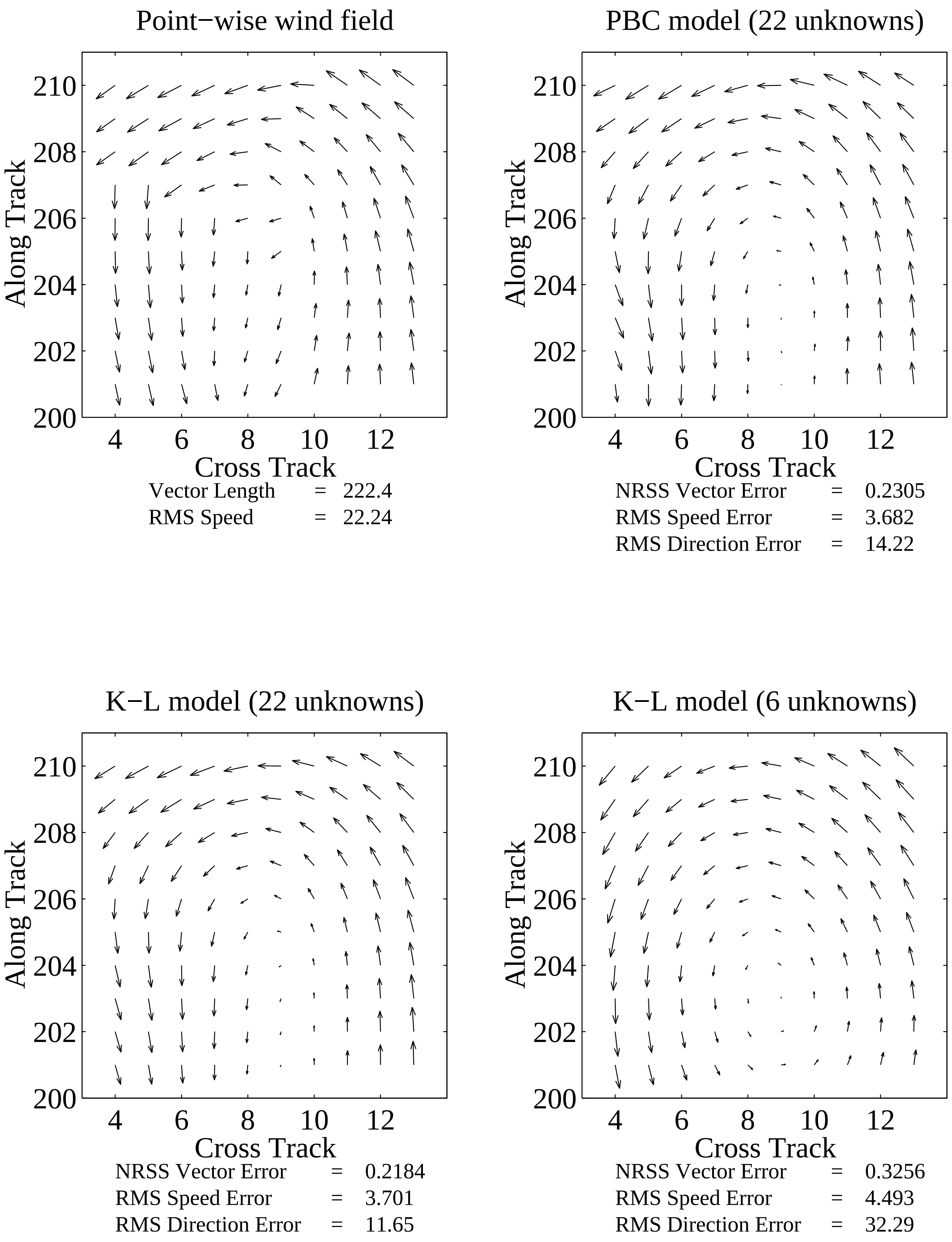
Fourier model (22 unknowns)

NRSS Vector Error = 0.1495
RMS Speed Error = 0.4804
RMS Direction Error = 8.125

Fourier model (10 unknowns)

NRSS Vector Error = 0.2388
RMS Speed Error = 0.7218
RMS Direction Error = 13.37

Figure 5.6: *First example of projecting a wind field onto Legendre and Fourier models. Wind-field taken from ascending portion of ERS-1 revolution 4452.*



Point−wise wind field

Vector Length = 222.4
RMS Speed = 22.24

PBC model (22 unknowns)

NRSS Vector Error = 0.2305
RMS Speed Error = 3.682
RMS Direction Error = 14.22

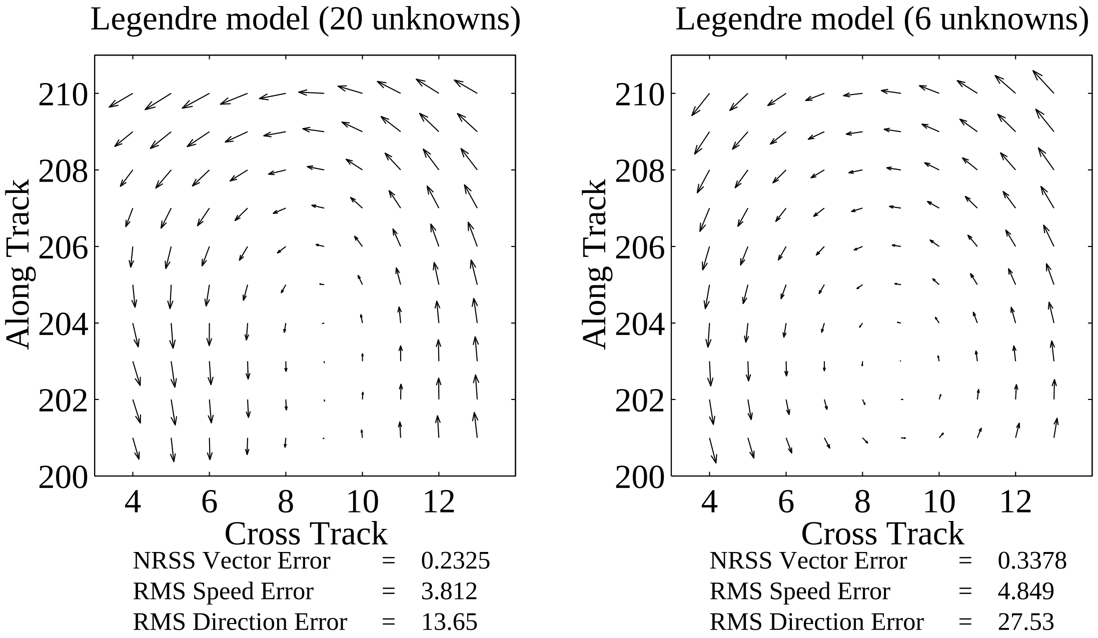
K−L model (22 unknowns)

NRSS Vector Error = 0.2184
RMS Speed Error = 3.701
RMS Direction Error = 11.65

K−L model (6 unknowns)

NRSS Vector Error = 0.3256
RMS Speed Error = 4.493
RMS Direction Error = 32.29

Figure 5.7: *Second Example of projecting a wind field onto PBC and Karhunen-Loeve models. Wind-field taken from ascending portion of ERS-1 revolution 4459.*



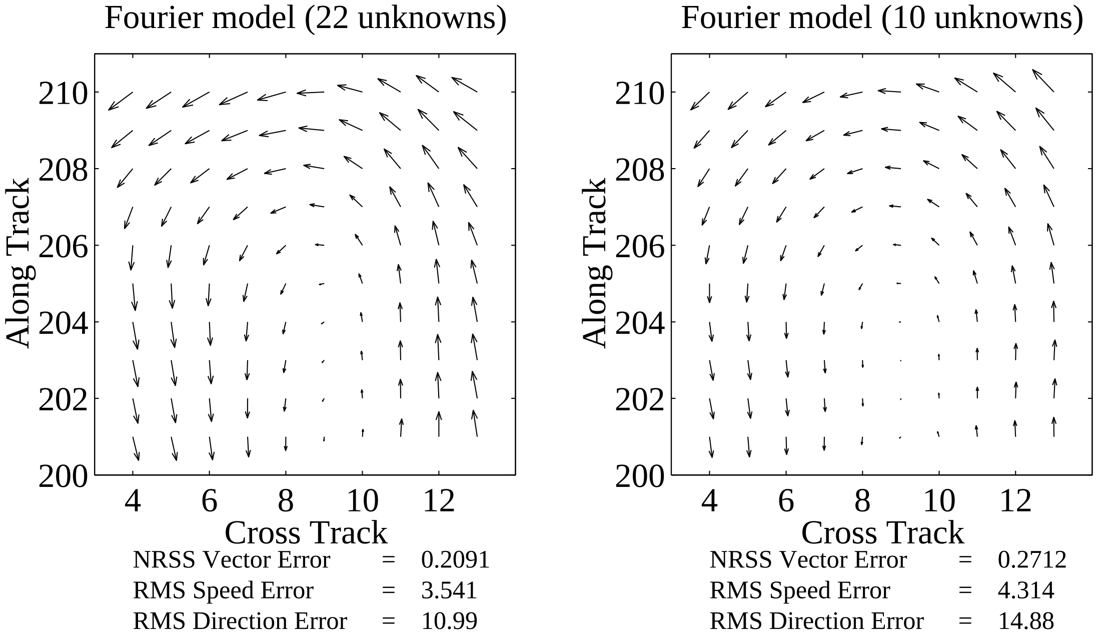


Figure 5.8: *Second Example of projecting a wind field onto Legendre and Fourier models. Wind-field taken from ascending portion of ERS-1 revolution 4459.*

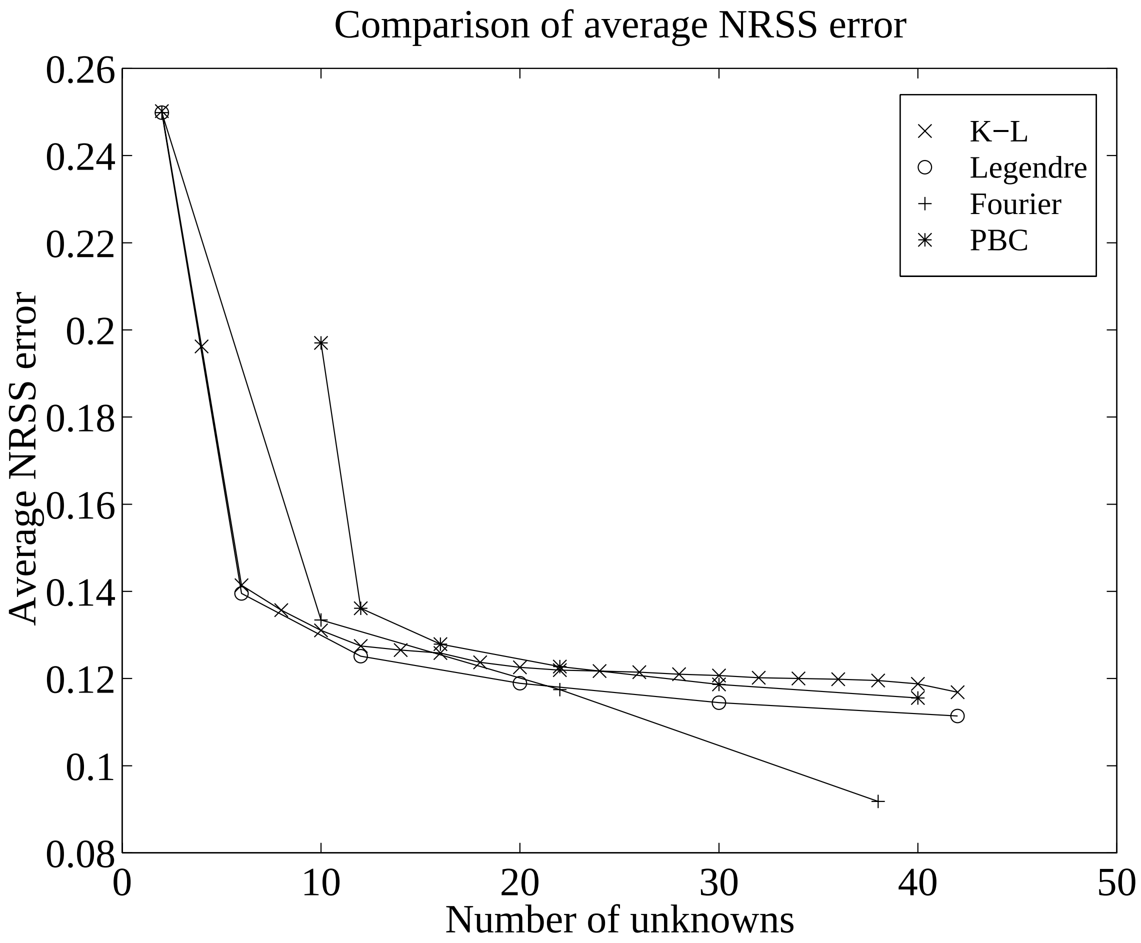


Figure 5.9: *NRSS projection error as a function of number of unknowns in model for several models.*

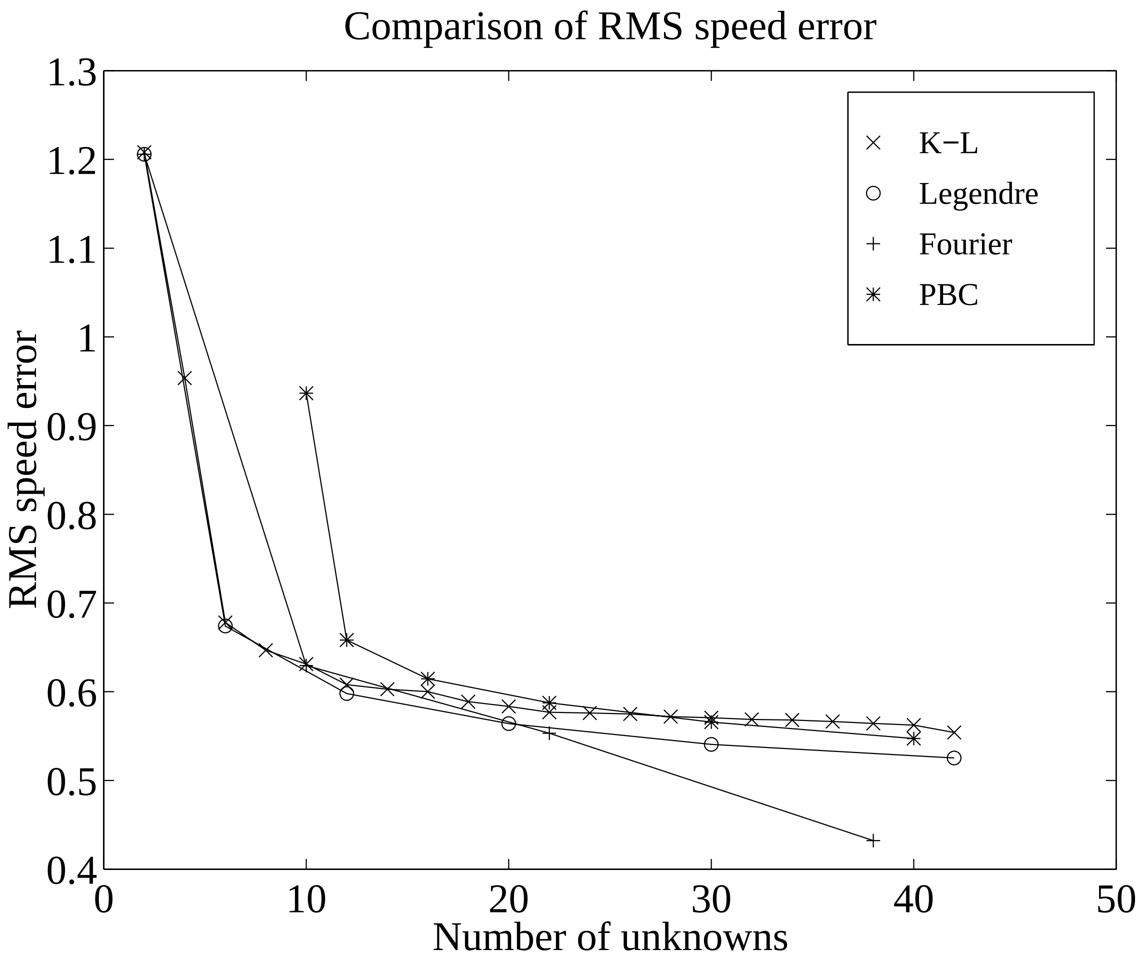


Figure 5.10: *RMS speed error as a function of number of unknowns in model for several models.*

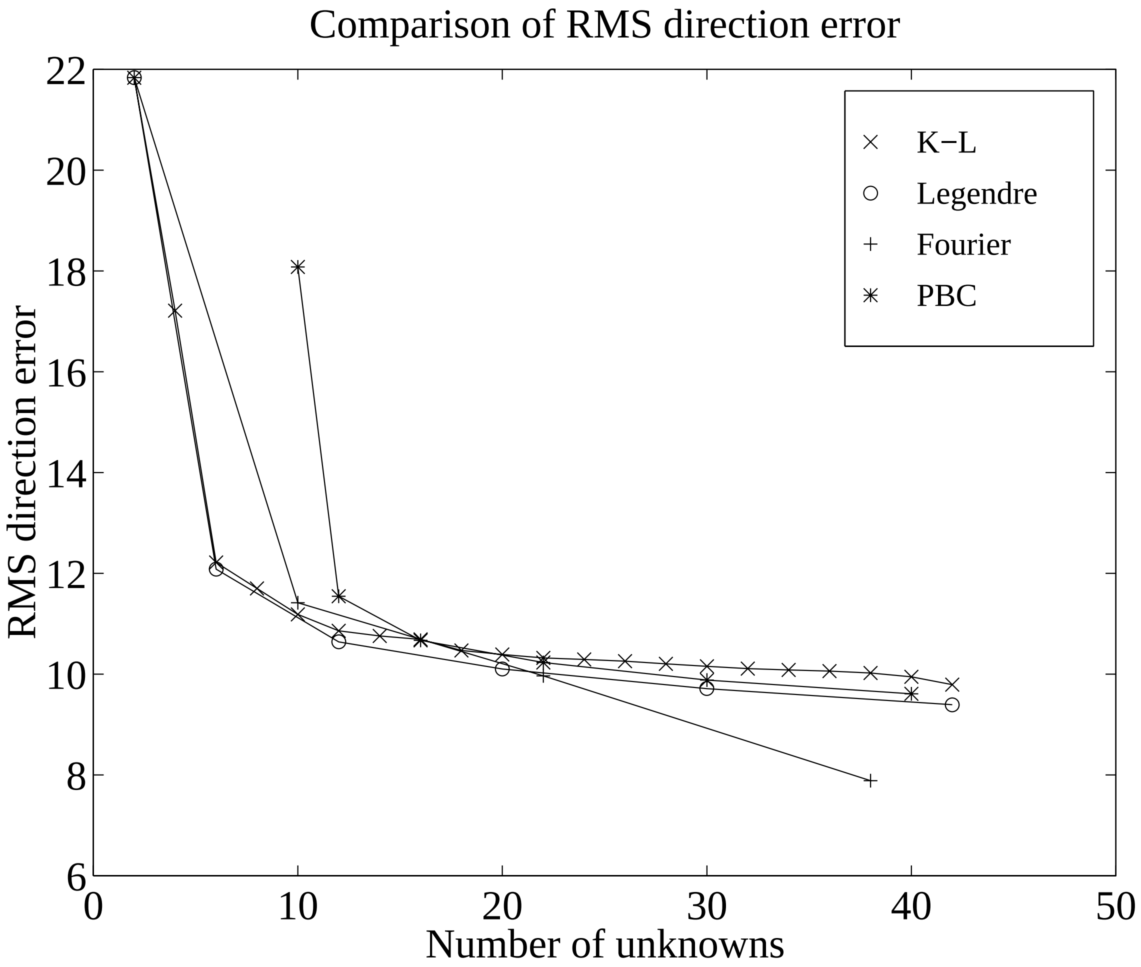


Figure 5.11: *RMS direction error as a function of number of unknowns in model for several models.*

Fourier model performed better on simulated wind fields. Due to the orthonormality of the Karhunen-Loeve basis and its flexibility in model-order, it is recommended that this basis set be used in model-based retrieval.

Average NRSS Errors for Projected Wind Fields				
Unknowns	K-L	Legendre	Fourier	PBC
2	0.2502	0.2499	0.2499	
4	0.1962			
6	0.1414	0.1395		
8	0.1357			
10	0.1311		0.1334	0.1970
12	0.1275	0.1251		0.1361
14	0.1265			
16	0.1259			0.1279
18	0.1237			
20	0.1226	0.1189		
22	0.1219		0.1175	0.1228
24	0.1217			
26	0.1215			
28	0.1210			
30	0.1207	0.1145		0.1186
32	0.1202			
34	0.1200			
36	0.1198			
38	0.1196		0.0918	
40	0.1188			0.1155
42	0.1169	0.1114		

Table 5.1: *NRSS projection errors for four different models as a function of the number of model parameters.*

RMS Speed Errors for Projected Wind Fields				
Unknowns	K-L	Legendre	Fourier	PBC
2	1.2083	1.2059	1.2059	
4	0.9535			
6	0.6780	0.6739		
8	0.6468			
10	0.6312		0.6294	0.9365
12	0.6079	0.5979		0.6583
14	0.6028			
16	0.6000			0.6146
18	0.5887			
20	0.5835	0.5640		
22	0.5769		0.5532	0.5874
24	0.5760			
26	0.5749			
28	0.5720			
30	0.5706	0.5406		0.5658
32	0.5688			
34	0.5681			
36	0.5664			
38	0.5643		0.4321	
40	0.5623			0.5473
42	0.5542	0.5252		

Table 5.2: *RMS speed errors for four different models as a function of the number of model parameters.*

RMS Direction Errors for Projected Wind Fields				
Unknowns	K-L	Legendre	Fourier	PBC
2	21.8350	21.8353	21.8353	
4	17.2108			
6	12.2157	12.0838		
8	11.6990			
10	11.1860		11.4178	18.0792
12	10.8591	10.6407		11.5436
14	10.7594			
16	10.6881			10.6700
18	10.4715			
20	10.3890	10.1015		
22	10.3211		9.9649	10.2303
24	10.2930			
26	10.2584			
28	10.2060			
30	10.1535	9.7140		9.8828
32	10.1108			
34	10.0838			
36	10.0594			
38	10.0246		7.8874	
40	9.9455			9.6103
42	9.7928	9.3950		

Table 5.3: *RMS direction errors for four different models as a function of the number of model parameters.*

Chapter 6

CONCLUSIONS

6.1 Summary of Contributions

In this thesis I have developed three areas which aid wind estimation over the ocean using scatterometer data. These three general contributions are (1) a method for estimating error in both point-wise and model-based retrieved winds, (2) a method for eliminating improbable wind aliases, and (3) alternate wind models for use in model-based wind retrieval. A more complete summary of each of these main areas follows.

6.1.1 Error estimation in wind retrieval

In Chapter 3 two approaches to estimate error in point-wise and model-based retrieved winds were developed. As part of this development, a measurement noise model which includes uncertainty in the Geophysical Model Function (GMF) was developed. It was found that the resulting noise probability density function could be considered Gaussian with little effect on wind retrieval for both NSCAT and ERS-1 scatterometers.

In addition, the Cramer-Rao bound was derived for both point-wise and model-based retrieval processes as an approximation to the covariance of retrieved winds. Using this approximation it was shown that the covariance of point-wise retrieved winds is sensitive to the true wind direction. In particular, if the true wind direction is aligned with either the fore or aft beam, then estimation uncertainty dramatically increases. Moreover, comparison of the Cramer-Rao bound with simulations show that the maximum likelihood estimator for point-wise wind retrieval is nearly efficient for most true wind vectors. Furthermore, comparing the estimated covariance of model-based and point-wise retrieved winds shows that model-based wind estimates can have greater variance than point-wise estimates at some cell locations in a region. This is true especially if the model contains unimportant basis vectors whose parameters cannot be estimated accurately for the particular region under consideration, suggesting that a one-size-fits-all linear model may not be adequate for accurate wind retrieval.

Also in Chapter 3 a direct covariance approximation method using the implicit function theorem was described and implemented for point-wise retrieved winds. It was found that the approximation compares favorably with the Cramer-Rao bound approximation for many true wind vectors. However, when the estimation uncertainty grows, the direct covariance approximation differs drastically from the Cramer-Rao bound. As a result, it was suggested that to make use of the direct covariance approximation, more terms in the Taylor expansion would need to be kept at considerable cost in complexity. Therefore, the Cramer-Rao bound is suggested as the better approximation to the covariance of wind estimates.

6.1.2 Distinguishing maxima for wind alias elimination

In Chapter 4, a method for eliminating wind aliases for both point-wise and model-based retrieval was suggested based on a likelihood-ratio decision rule. It was found that the approach can ameliorate the dealiasing problem in point-wise retrieval and often eliminate it altogether with model-based retrieval. Even if the approach does not dismiss the need for dealiasing, it allows a dealiasing algorithm to work with only those winds that are statistically consistent with the measurements.

6.1.3 Wind field modeling

Finally, in Chapter 5 several additional linear models were proposed for use in model-based retrieval. All of these models are based on modeling the wind directly instead of through its curl and divergence. One of these models is a Karhunen-Loeve basis set for 25km wind fields. Comparing the modeling error introduced by each of these models with Long’s (PBC) model for varying model-orders shows that all models give similar error. The Fourier model had less modeling error on simulated wind fields although this is probably a result of the averaging done to obtain these fields. However, it is recommended that the Karhunen-Loeve model be used in model-based retrieval since it admits the greatest flexibility in model-order selection and has orthonormal columns.

6.2 Future Research

There are several possible directions for future research which come out of the essential contributions of this thesis. As a result this section will discuss future research as it pertains to each of the three basic areas.

6.2.1 Error estimation in wind retrieval

One area for future research is improving the model used for σ° measurements and evaluating its effect on wind retrieval. It is evident that geophysical modeling error must be included to obtain believable estimates of retrieved-wind uncertainty. In this thesis it was assumed that uncorrelated Gaussian noise was added to each measurement used in wind retrieval. Research could be done which investigates the effect on wind retrieval of correlation in the geophysical modeling error. Most importantly in this regard, the effect on wind estimate uncertainty should be investigated assuming some covariance in the noise added due to the geophysical model function error. In particular, the Cramer-Rao and/or the direct covariance estimate should be re-derived with covariance in the measurements due to modeling error included.

Another extension derived from the Chapter on wind retrieval uncertainty is to investigate the Cramer-Rao bound on model-based estimates in more depth. In particular, the effect of wind field modeling error on the estimates of $\mathbf{X}$ should be investigated to determine the effect on the uncertainty on $\hat{\mathbf{X}}$. An additional extension to the Chapter on wind retrieval uncertainty is to calculate an error bound on the direct covariance estimate by using Taylor's second-order remainder formula. This would involve using implicit differentiation to calculate the second derivatives of the wind estimate function. This would also be useful in expanding the discussion of bias contained in Appendix B. Of course, an ideal extension to the work on wind retrieval uncertainty would be the derivation of a region-estimate of the wind.

6.2.2 Distinguishing maxima for wind alias elimination

There are several possible extensions to the chapter which develops the statistical test to distinguish among the maxima that correspond to wind aliases. One area of future research would be to evaluate in more depth the possibility of applying the decision rule to model-based retrieved winds in order to distinguish a single wind-field among the several that give maxima to the log-likelihood function. As shown in Chapter 4, application to model-based winds is straightforward and has the potential to uniquely identify a single wind field.

Another possible extension to this chapter is to examine the effect of covariance in the geophysical modeling error as described before. Essentially, this would involve a re-derivation of the distribution of the likelihood-ratio statistic assuming covariance in the σ° measurements. Such a derivation would necessarily use a slightly different approach than the one used in this thesis, since independence was immediately exploited in calculating the moment generating function. Nonetheless through a linear transformation, the likelihood-ratio statistic could be brought to a similar form as used in this thesis (Imhof, 1961). As a result, the expression for α_0 would be essentially the same (with different definitions for c_k, d_k, and y_0).

A third possible extension to the work in Chapter 4 is the possibility of using the test to remove the ambiguity problem completely in some cases by using multi-scale wind retrieval. There is a definite upwind/downwind asymmetry in the GMF which indicates that theoretically it should be possible to resolve a single wind direction if the noise is reduced to a low enough level. Another way to see this is to recall that without noise there is only one wind velocity that exactly agrees with three or more measurements according to the GMF. The effect of noise can be reduced by reducing the resolution of wind retrieval. In other words, more measurements are used for each estimated wind vector and the wind is retrieved on a 100km by 100km or larger grid. Theoretically, as more measurements are gathered, the true wind direction could be distinguished by using the statistical test developed in this thesis. Then, the wind alias closest to this direction could be selected for wind retrieved on a grid of finer resolution. One limit to this technique is correlation length in the wind field at the desired resolution. If the winds at this resolution change a great deal over the larger region, then averaging the measurements from different regions would ‘‘wash out’’ the true wind direction. Nonetheless, preliminary results suggest that this technique could work.

6.2.3 Wind field modeling

There are also several possible areas of future research with respect to wind field modeling. One possible extension involves the Karhunen-Loeve basis. In this thesis all types of wind fields from both north and south of the equator in the Pacific Ocean were used as data from which the basis vectors were estimated. One disadvantage of using such a broad data set, is that specific phenomenological features such as fronts are buried in the noise. As a result, a truncated basis may not model certain features such as fronts, well. If a data set could be generated that includes such specific phenomenological features as fronts or cyclones then a K-L basis could be estimated from this data set. Then, model-based retrieval could be performed by first examining the σ° measurements to determine which model would ‘‘fit’’ the data and then using that model. This could result in significant reduction of the number of basis wind-field vectors that need to be used. Such a procedure could also be implemented with the K-L basis vectors already presented by simply using models of different orders as the different available models.

Another extension to the area of wind field modeling is to examine wind field models that directly model speed and direction instead of the rectangular components of the wind vector. This may result in more useful models. However, such research would probably have less impact than a model-based approach that used varying models as described above.

A final path of future research is to compare models with respect to the model-based Cramer-Rao bound uncertainty derived in Chapter 3. Since this is a function of the wind field model used, some wind field models may perform better than others. This estimation precision criteria along with a model accuracy criteria could be used to derive a future ‘‘best’’ model.

BIBLIOGRAPHY

Abel, J., ‘‘A bound on mean-square estimate error.’’ *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1675--1680, 1993.

Attema, E. P., ‘‘The Active Microwave Instrument On-Board the ERS-1 Satellite.’’ *Proceedings of the IEEE*, vol. 79, no. 6, pp. 791--799, 1991.

Davis, W. B., *Enhanced Resolution Imaging from Remotely Sensed Microwave Data*. Master’s thesis, Brigham Young University, 1993.

Davison, J. and D. Harrison, ‘‘Comparison of Seasat Scatterometer Winds with Tropical Pacific Observations.’’ *Journal of Geophysical Research*, vol. 95, no. C3, pp. 3403--3410, 1990.

Donelan, M. and J. W.J. Pierson, ‘‘Radar Scattering and Equilibrium Ranges in Wind-Generated Waves with Application to Scatterometry.’’ *Journal of Geophysical Research*, vol. 92, no. C5, pp. 4971--5029, 1987.

Dunbar, R., S. Hsiao, and B. Lambrigtsen, ‘‘Science Algorithm Specification for the NASA Scatterometer Project.’’ Vols. 1 & 2, D-5610, NASA Jet Propulsion Laboratory Report 622-107, Pasadena, CA 91103, 1988.

Fessler, J. A., ‘‘Moments of Implicitly Defined Estimators (e.g. ML and MAP): Applications to Transmission Tomography.’’ In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing,* vol. 3662, pp. 2291--2294, IEEE, 1995.

Fessler, J. A., ‘‘Mean and Variance of Implicitly Defined Biased Estimators (such as Penalized Maximum Likelihood): Applications to Tomography.’’ *IEEE Transactions on Image Processing*, vol. 5, no. 3, pp. 493--506, 1996.

Freilich, M. and D. Chelton, ‘‘Wavenumber Spectra of Pacific Winds Measured by the Seasat Scatterometer.’’ *Journal of Physical Oceanography*, vol. 16, no. 4, pp. 741--757, 1986.

Freilich, M. and R. Dunbar, ‘‘A Preliminary C-band Scatterometer Model Function for the ERS-1 AMI Instrument.’’ In *Proceedings of the First ERS-1 Symposium*, pp. 79--84, Cannes, France, ESA, 1993, SP-359.

Grantham, W., E. Bracalente, W. Jones, and J. Johnson, ‘‘The Seasat-A Sattelite Scatterometer.’’ *IEEE Journal of Oceanic Engineering*, vol. OE-2, no. 2, pp. 200--206, 1977.

Imhof, J., ‘‘Computing the distribution of quadratic forms in normal variables.’’ *Biometrika*, vol. 48, no. 3 and 4, pp. 419--426, 1961.

Johnson, J., J. L.A. Williams, E. Bracalente, F. Beck, and W. Grantham, ‘‘Seasat-A Satellite Scatterometer Instrument Evaluation.’’ *IEEE Journal of Oceanic Engineering*, vol. OE-5, no. 2, pp. 138--144, 1980.

Johnson, P. E., D. G. Long, and T. E. Oliphant, ‘‘Geophysical Modeling Error in Wind Scatterometry.’’ In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, pp. 1721--1723, IEEE, 1996.

Jones, W., L. Schroeder, and J. Mitchell, ‘‘Aircraft Measurements of the Microwave Scattering Signature of the Ocean.’’ *IEEE Journal of Oceanic Engineering*, vol. OE-2, no. 1, pp. 52--61, 1977.

Leotta, D. F. and D. G. Long, ‘‘Probability Distribution of Wind Retrieval Error for the NASA Scatterometer.’’ In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, pp. 1466--1469, IEEE, 1989.

Levy, G. and R. A. Brown, ‘‘Southern Hemisphere Synoptic Weather from a Satellite Scatterometer.’’ *Monthly Weather Review*, vol. 119, no. 12, pp. 2803--2813, 1991.

Long, D. G., *Model-Based Estimation of Wind Fields Over the Oceans from Wind Scatterometer Measurements*. Ph.D. thesis, University of Southern California, 1989.

Long, D. G., ‘‘Wind Field Model-Based Estimation of Seasat Scatterometer Winds.’’ *Journal of Geophysical Research*, vol. 98, no. C8, pp. 14651--14668, 1993.

Long, D. G. and J. M. Mendel, ‘‘Model-Based Estimation of Wind Fields Over the Ocean From Wind Scatterometer Measurements, Part I: Development of the Wind Field

Model.'' *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 349--360, 1990a.

Long, D. G. and J. M. Mendel, ''Model-Based Estimation of Wind Fields Over the Ocean From Wind Scatterometer Measurements, Part II: Model Parameter Estimation.'' *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 361--373, 1990b.

Marsden, J. E. and A. J. Tromba, *Vector Calculus*. W.H. Freeman and Company, 3rd edn., 1988.

Naderi, F., M. Freilich, and D. Long, ''Spaceborne Radar Measurement of Wind Velocity Over the Ocean---An Overview of the NSCAT Scatterometer System.'' *Proceedings of the IEEE*, vol. 79, no. 6, pp. 850--866, 1991.

Oliphant, T. E. and D. G. Long, ''Cramer-Rao Bound for Wind Estimation from Scatterometer Measurement.'' In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, pp. 1724--1726, IEEE, 1996a.

Oliphant, T. E. and D. G. Long, ''Development of a Statistical Method for Eliminating Improbable Wind Aliases in Scatterometer Wind Retrieval.'' In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, pp. 1715--1717, IEEE, 1996b.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 3rd edn., 1991.

Scharf, L. L., *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley, Reading, Massachusetts, 1991.

Shaffer, S. J., R. S. Dunbar, S. V. Hsiao, and D. G. Long, ''A Median-Filter-Based Ambiguity Removal Algorithm for NSCAT.'' *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, no. 1, pp. 167--174, 1991.

Stoffelen, A. C. M. and G. J. Cats, ''The Impact of Seasat-A Scatterometer Data on High-Resolution Analyses and Forecasts: The Devlopment of the QE II Storm.'' *Monthly Weather Review*, vol. 119, no. 12, pp. 2794--2802, 1991.

Ulaby, F. T., R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, vol. I. Artech House, Inc., 685 Canton Street Norwood, MA 02062, 1982.

Wentz, F., S. Peteherych, and L. Thomas, ‘‘A Model Function for Ocean Radar Cross Sections at 14.6 GHz.’’ *Journal of Geophysical Research*, vol. 89, no. C3, pp. 3689--3704, 1984.

Appendix A

APPROXIMATING A NEAR-GAUSSIAN DENSITY

Often it is convenient to approximate a near-Gaussian density function using known functions. The approach outlined in this appendix gives more detail for the technique given on page 217 of Papoulis (1991). The technique is to expand the error between the real density and the approximating Gaussian density in Hermite polynomials. The coefficients can then be determined in terms of the moments of the real density function.

Suppose the density to be approximated is denoted as $f(x)$ with central moments

$$\mu_m = E\left[(x-\mu)^n\right] \qquad m > 1, \tag{A.1}$$

where μ is the mean of the density function.

The Gaussian density which matches the mean and variance of this distribution is

$$f_N(x) = \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{(x-\mu)^2}{2\mu_2}}. \tag{A.2}$$

The error between the true density function and this Gaussian distribution can be expanded in Hermite polynomials:

$$\begin{aligned} \epsilon(x) &= f(x) - f_N(x), && \text{(A.3)} \\ &= \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{(x-\mu)^2}{2\mu_2}} \sum_{k=3}^{\infty} C_k H_k\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right). && \text{(A.4)} \\ & && \text{(A.5)} \end{aligned}$$

We use (unnormalized) Hermite polynomials defined as

$$\begin{aligned} H_k(x) &= \sum_{n=0}^{\lfloor \frac{k}{2} \rfloor} D_{k,n} x^{k-2n}, && \text{(A.6)} \\ D_{k,n} &= \frac{(-1)^n k! 2^{k-2n}}{n!(k-2n)!}. && \text{(A.7)} \end{aligned}$$

To find the coefficients, C_k, in this expansion, multiply both sides of Eq. (A.4) by $H_n\left((x-\mu)/\sqrt{2\mu_2}\right)$ and integrate over x:

$$\begin{aligned}\int_{-\infty}^{\infty} dx\,\epsilon(x) H_n\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right) &= \int_{-\infty}^{\infty} dx\,\frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{(x-\mu)^2}{2\mu_2}} \sum_{k=3}^{\infty} C_k H_k\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right) H_n\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right), \\ &= \sum_{k=3}^{\infty} \frac{C_k}{\sqrt{\pi}} \int_{-\infty}^{\infty} d\alpha\, e^{-\alpha^2} H_k(\alpha) H_n(\alpha), \qquad \text{(A.8)}\end{aligned}$$

where we have made the substitution $\alpha = (x-\mu)/\sqrt{2\mu_2}$. Using the orthogonality relation for Hermite polynomials,

$$\int_{-\infty}^{\infty} d\alpha\, e^{-\alpha^2} H_k(\alpha) H_n(\alpha) = 2^n n! \sqrt{\pi}\,\delta_{k,n}, \qquad \text{(A.9)}$$

we see that

$$\int_{-\infty}^{\infty} dx\,\epsilon(x) H_n\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right) = 2^n n! C_n. \qquad \text{(A.10)}$$

Consequently,

$$\begin{aligned} C_k &= \frac{1}{2^k k!} \int_{-\infty}^{\infty} dx\,\epsilon(x) H_k\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right), \qquad \text{(A.11)} \\ &= \frac{1}{2^k k!} \int_{-\infty}^{\infty} dx \sum_{n=0}^{\lfloor \frac{k}{2} \rfloor} \frac{D_{k,n}}{(2\mu_2)^{\frac{k}{2}-n}} \int_{-\infty}^{\infty} dx\,(x-\mu)^{k-2n}\left[f(x) - f_N(x)\right], \qquad \text{(A.12)} \\ &= \frac{1}{2^k k!} \sum_{n=0}^{\lfloor \frac{k}{2} \rfloor} \frac{D_{k,n}}{(2\mu_2)^{\frac{k}{2}-n}} \left(\mu_{k-2n} - \mu_{k-2n}^N\right). \qquad \text{(A.13)}\end{aligned}$$

Note that $\mu_0 = \mu_0^N = 1$, $\mu_1 = \mu_1^N = 0$, and $\mu_2 = \mu_2^N$. As a result, the sum over n can be rewritten with upper limit $k/2 - 2$ when k is even and $(k-1)/2 - 1$ when k is odd. This can be written for both odd and even k as $\lfloor (k-3)/2 \rfloor$. As a result, we can write

$$C_k = \sum_{n=0}^{\lfloor \frac{k-3}{2} \rfloor} \frac{(-1)^n \left(\mu_{k-2n} - \mu_{k-2n}^N\right)}{n!(k-2n)!\mu_2^{k/2-n} 2^{k/2+n}}. \qquad \text{(A.14)}$$

With this expression we can write an expansion for the original probability density, $f(x)$:

$$\begin{aligned} f(x) &= \epsilon(x) + f_N(x), \\ &= \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{(x-\mu)^2}{2\mu_2}} \left[1 + \sum_{k=3}^{\infty} C_k H_k\left(\frac{x-\mu}{\sqrt{2\mu_2}}\right)\right]. \qquad \text{(A.15)}\end{aligned}$$

Appendix B

BIAS IN THE WIND ESTIMATE

This appendix extends the results of Chapter 3 to include a discussion of the bias in the wind estimate. While the techniques discussed applies to both point-wise and model-based estimates, implementation and notation will focus exclusively on the point-wise wind estimate.

In Section 3.3 we assumed that the wind estimate is unbiased or constant-biased[1]. Comparing the predicted covariance with simulations showed that this is a reasonable assumption for most wind vectors. At low incidence angles (near-swath) and low wind speeds, when the true wind direction is aligned with the fore and/or aft beams, the unbiased Cramer-Rao bound seriously over-predicts the simulated variance (especially for the wind direction estimate). This leads us to hypothesize that under these conditions the wind estimate is biased. In this appendix we derive a method to approximate the wind bias and then apply the Cramer-Rao bound for biased estimators to approximate the covariance. This results in an approximation to the covariance of retrieved wind that more closely follows simulations.

B.1 Approximating the bias

To approximate the bias in the wind estimate, we use the approach discussed in Section 3.4 and adapted from Fessler (1995; 1996). As a review, in this approach the implicit function relating the measurements to the wind estimate is expanded in a Taylor series about the mean of the measurements. In Section 3.4 we used a first-order expansion to directly approximate the covariance of the wind estimate. In this section, we use the same first-order expansion to approximate the mean (and therefore the bias) of the wind estimate.

In particular, we rewrite the equation and its derivation for Eq. (3.82):

$$\hat{\mathbf{w}} = \mathbf{h}(\mathbf{z}) \approx \mathbf{h}(\bar{\mathbf{z}}) + D(\mathbf{z} - \bar{\mathbf{z}}), \tag{B.1}$$

---

[1] Although not explicitly stated in the text, the unbiased Cramer-Rao bound also applies to constant-biased wind estimates.

where $D = \partial \mathbf{h}(\bar{\mathbf{z}})/\partial \mathbf{z}$ is a (row) derivative matrix with elements

$$D_{ij} = \frac{\partial h_i}{\partial z_j}. \tag{B.2}$$

From this approximation for $\hat{\mathbf{w}}$ and estimate of the mean of $\hat{\mathbf{w}}$ can be obtained:

$$\begin{aligned} E\left(\hat{\mathbf{w}}\right) &= \mathbf{h}\left(\bar{\mathbf{z}}\right), \\ &= \mathbf{h}\left[\mathcal{M}(\mathbf{w})\right]. \end{aligned} \tag{B.3}$$

This equation implies that an approximation to the wind estimate bias can be obtained by applying the wind retrieval maximization algorithm to noise-free ‘‘measurements’’ obtained using the GMF on the true wind $\mathbf{w}$. (In this context, $\mathcal{M}(\mathbf{w})$ refers to the vector of σ_m° values needed for wind retrieval.) The wind bias is the difference between the expected value of the estimate and the true wind $\mathbf{w}$. As a result, the bias can be approximated as:

$$B(\mathbf{w}) \equiv E\left(\hat{\mathbf{w}}\right) - \mathbf{w}, \tag{B.4}$$

$$\approx \mathbf{h}\left[\mathcal{M}(\mathbf{w})\right] - \mathbf{w}. \tag{B.5}$$

B.2 Biased Cramer-Rao bound

There is a more general version of the Cramer-Rao bound which admits the possibility of bias in the wind estimate. This bound on the covariance can be expressed as (Abel, 1993):

$$C \geq \frac{\partial E\left(\hat{\mathbf{w}}\right)}{\partial \mathbf{w}} J^{-1} \left[\frac{\partial E\left(\hat{\mathbf{w}}\right)}{\partial \mathbf{w}}\right]^T, \tag{B.6}$$

$$\geq \Gamma J^{-1} \Gamma^T, \tag{B.7}$$

where

$$\Gamma = \frac{\partial E\left(\hat{\mathbf{w}}\right)}{\partial \mathbf{w}}, \tag{B.8}$$

so that

$$\Gamma_{ij} = \frac{\partial E\left(\hat{w}_i\right)}{\partial w_j}. \tag{B.9}$$

Because of the bias, the total mean square error includes this covariance as well as the bias:

$$\text{MSE} \geq B(\mathbf{w})B(\mathbf{w})^T + \Gamma J^{-1}\Gamma^T. \tag{B.10}$$

Γ can be approximated using the approximation to the bias derived in the previous section and the chain rule:

$$\begin{aligned}
\Gamma(\mathbf{w}) &= \frac{\partial E\left(\hat{\mathbf{w}}\right)}{\partial \mathbf{w}}, \\
&\approx \frac{\partial \mathbf{h}\left[\mathcal{M}(\mathbf{w})\right]}{\partial \mathbf{w}}, && \text{(B.11)} \\
&\approx \frac{\partial \mathbf{h}(\mathbf{z})}{\partial \mathbf{z}} \frac{\partial \mathcal{M}(\mathbf{w})}{\partial \mathbf{w}}, && \text{(B.12)} \\
&\approx DG. && \text{(B.13)}
\end{aligned}$$

In this expression we have used notation as in Chapter 3 where

$$D \equiv \frac{\partial \mathbf{h}(\mathbf{z})}{\partial \mathbf{z}}, \tag{B.14}$$

and

$$G = \frac{\partial \mathcal{M}(\mathbf{w})}{\partial \mathbf{w}}. \tag{B.15}$$

All derivatives are evaluated at the point, $(\bar{\mathbf{z}}, \mathbf{h}(\bar{\mathbf{z}})) = (\mathcal{M}(\mathbf{w}), \mathbf{h}(\mathcal{M}(\mathbf{w})))$ as in Chapter 3. In that chapter, however, we made the approximation that $\mathbf{h}(\bar{\mathbf{z}}) = \mathbf{w}$ (unbiased) to calculate the derivative matrix D. As a result, the expression for the second derivatives of the likelihood function given there are not quite correct. As these are necessary to compute D, they will be given here in their complete form. From Chapter 3 we have

$$D = -(D^{20})^{-1}D^{11}, \tag{B.16}$$

where

$$\begin{aligned}
D^{11} &= \frac{\partial}{\partial \mathbf{z}}\left[\frac{\partial L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial \mathbf{w}}\right]^T = \left\{\frac{\partial^2 L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial z_j \partial w_i}\right\}, && \text{(B.17)} \\
D^{20} &= \frac{\partial}{\partial \mathbf{w}}\left[\frac{\partial L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial \mathbf{w}}\right]^T = \left\{\frac{\partial^2 L(\mathbf{h}(\bar{\mathbf{z}}), \bar{\mathbf{z}})}{\partial w_i \partial w_j}\right\}. && \text{(B.18)}
\end{aligned}$$

Defining $\mathcal{M}_k(\mathbf{h}(\bar{\mathbf{z}})) = \widetilde{\mathcal{M}}_k$, and recalling that $\bar{z}_k = \mathcal{M}_k(\mathbf{w}) \equiv \mathcal{M}_k$, we can write the elements of D^{11} and D^{20}:

$$\begin{aligned} D^{11}_{ij} &= \frac{1}{\varsigma^2_{z_j}} \frac{\partial \widetilde{\mathcal{M}}_j}{\partial w_i} + \frac{\mathcal{M}_j - \widetilde{\mathcal{M}}_j}{\varsigma^4_{z_j}} \frac{\partial \varsigma^2_{z_j}}{\partial w_i}, & \text{(B.19)} \\ D^{20}_{ij} &= \sum_{k=1}^{K} \Bigg[-\frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_i} \frac{1}{\varsigma^2_{z_k}} \frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_j} + \frac{\partial \varsigma^2_{z_k}}{\partial w_i} \frac{1}{2\varsigma^4_{z_k}} \frac{\partial \varsigma^2_{z_k}}{\partial w_j} - \frac{1}{2\varsigma^2_{z_k}} \frac{\partial^2 \varsigma^2_{z_k}}{\partial w_i \partial w_j} \\ &\quad - \frac{\mathcal{M}_k - \widetilde{\mathcal{M}}_k}{\varsigma^4_{z_k}} \left(\frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_i} \frac{\partial \varsigma^2_{z_k}}{\partial w_j} + \frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_j} \frac{\partial \varsigma^2_{z_k}}{\partial w_i} \right) + \frac{\mathcal{M}_k - \widetilde{\mathcal{M}}_k}{\varsigma^2_{z_k}} \frac{\partial^2 \widetilde{\mathcal{M}}_k}{\partial w_j \partial w_i} \\ &\quad - \frac{\partial \varsigma^2_{z_k}}{\partial w_i} \frac{(\mathcal{M}_k - \widetilde{\mathcal{M}}_k)^2}{\varsigma^6_{z_k}} \frac{\partial \varsigma^2_{z_k}}{\partial w_j} + \frac{(\mathcal{M}_k - \widetilde{\mathcal{M}}_k)^2}{2\varsigma^4_{z_k}} \frac{\partial^2 \varsigma^2_{z_k}}{\partial w_j \partial w_i} \Bigg], & \text{(B.20)} \end{aligned}$$

where for clarity

$$\begin{aligned} \frac{\partial^2 \varsigma^2_{z_k}}{\partial w_i \partial w_j} &= \left[2\epsilon \widetilde{\mathcal{M}}_k + \beta(1 + K^2_{pm}) \right] \frac{\partial^2 \widetilde{\mathcal{M}}_k}{\partial w_i \partial w_j} + 2\epsilon \frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_i} \frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_j}, & \text{(B.21)} \\ \frac{\partial \varsigma^2_{z_k}}{\partial w_i} &= \left[2\epsilon \widetilde{\mathcal{M}}_k + \beta(1 + K^2_{pm}) \right] \frac{\partial \widetilde{\mathcal{M}}_k}{\partial w_i}. & \text{(B.22)} \end{aligned}$$

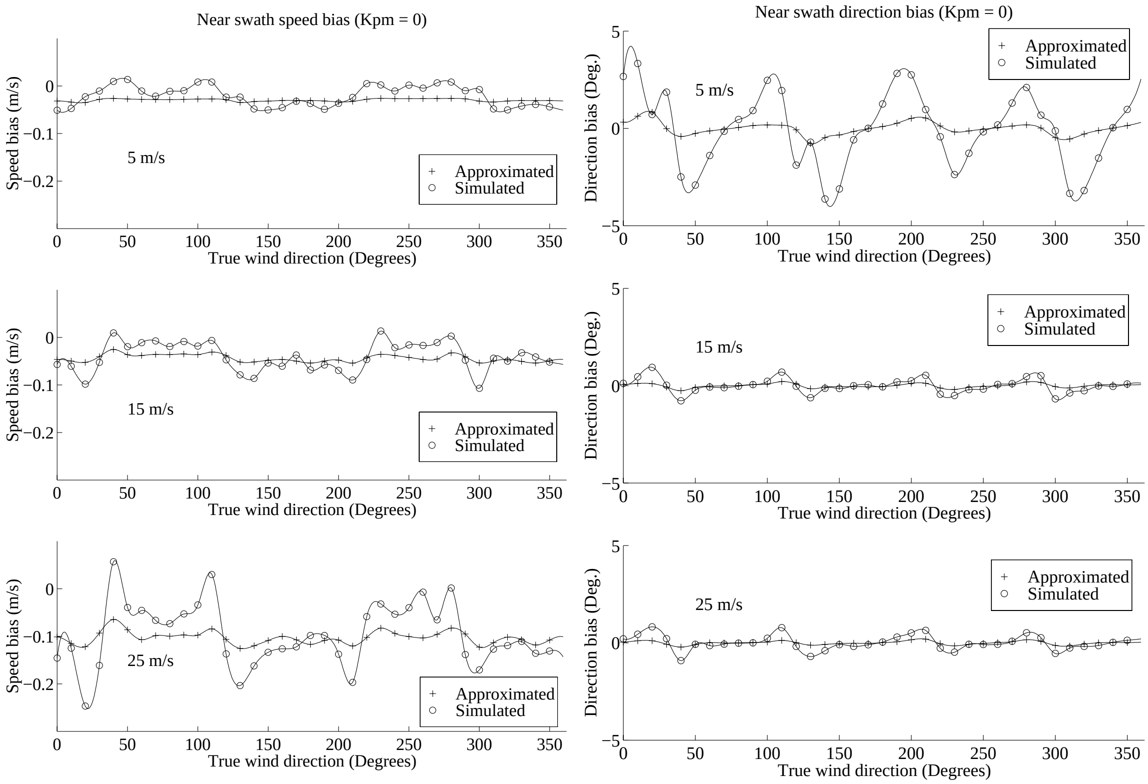
Note that the notation emphasizes that the partials of $\mathcal{M}_k$ are evaluated at $\mathbf{h}(\bar{\mathbf{z}})$.

B.3 Examples

This section presents some examples of the wind estimate bias and the biased Cramer-Rao bound. Both ERS-1 and NSCAT examples are presented. The examples show that accounting for the bias in the manner described above improves the agreement between predicted variance and simulation, especially for the unrealistically large values of covariance predicted at certain true wind directions for near swath, low wind speeds and large K_{pm} values. The discrepancy between predictions and simulations found in Chapter 3 is not completely eliminated with this technique but it is improved.

B.3.1 ERS-1

Two examples are chosen for presentation. Since the most marked difference between simulations and the unbiased Cramer-Rao bound occurs at near swath for low wind speeds, both of these examples are at near swath. The first example uses $K_{pm} = 0$ while the second uses $K_{pm} = 0.17$. The wind cell is the same as that used to generate Figure 3.5. Figure B.1 compares the approximate bias in the wind estimate as predicted by

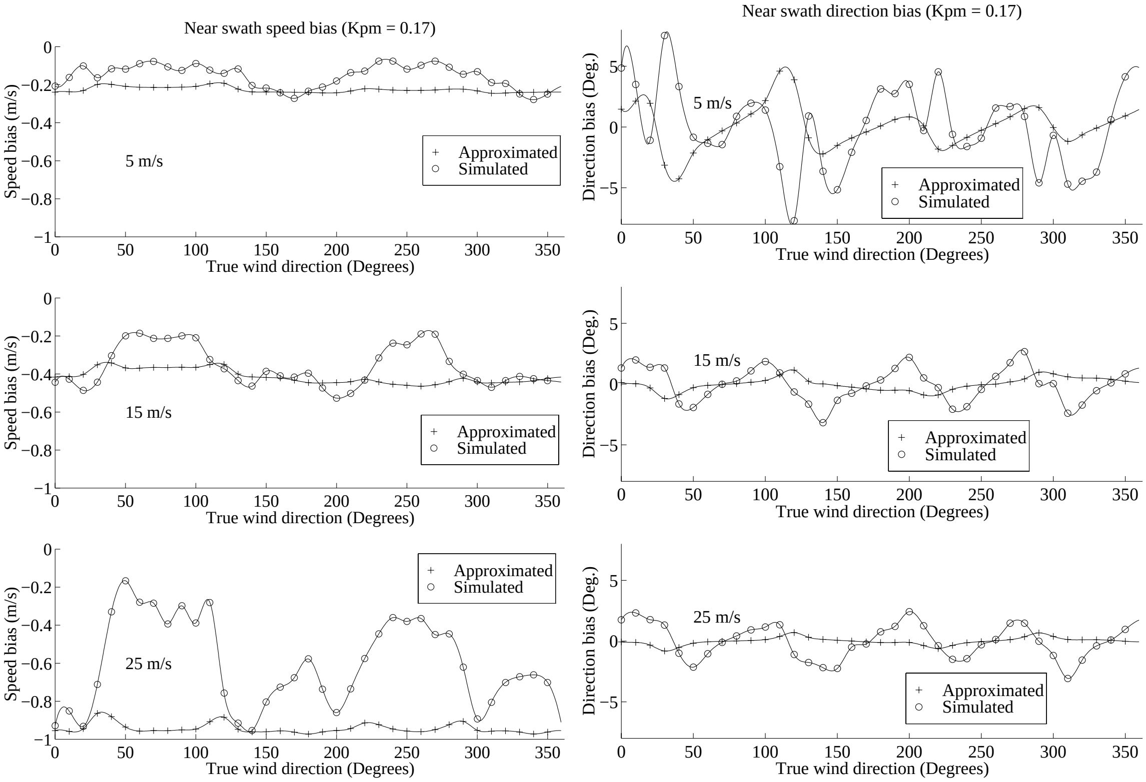


(a) Speed Bias

(b) Direction Bias

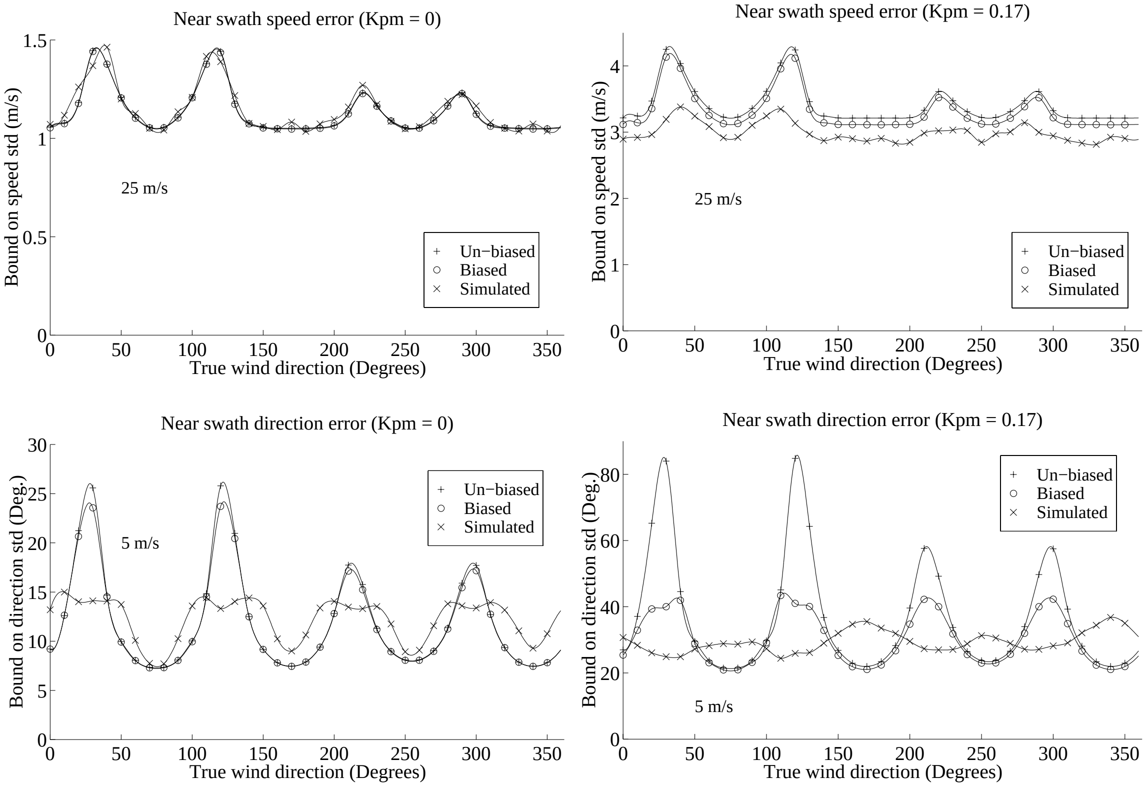
Figure B.1: *Comparison of approximate analytic bias with simulated bias for ERS-1 at near swath.* $K_{pm} = 0$.

Eq. (B.5) to the bias estimated with a simulation of $N = 2000$ retrievals. Note that for wind directions where the unbiased Cramer-Rao bound predicted large variance in the wind direction estimate, the derivative of the bias with respect to the wind direction becomes negative. This has the effect of decreasing the variance bound of the wind direction estimate as shown in Figure B.3(a) for a true wind speed of 5 m/s. Also shown in that figure is a comparison of the unbiased and biased Cramer-Rao bound with simulations as a function of true wind direction. Figures B.2 and B.3(b) show similar plots assuming $K_{pm} = 0.17$. Notice that the variance reduction effect of the unbiased Cramer-Rao bound is even more pronounced when the unbiased variance bound is especially high. The discrepancy with simulations on the wind direction standard deviation is still not completely understood but may be due to the need for higher order terms in the Taylor-series expansion of $\mathbf{h}(\mathbf{z})$.



(a) Speed Bias (b) Direction Bias

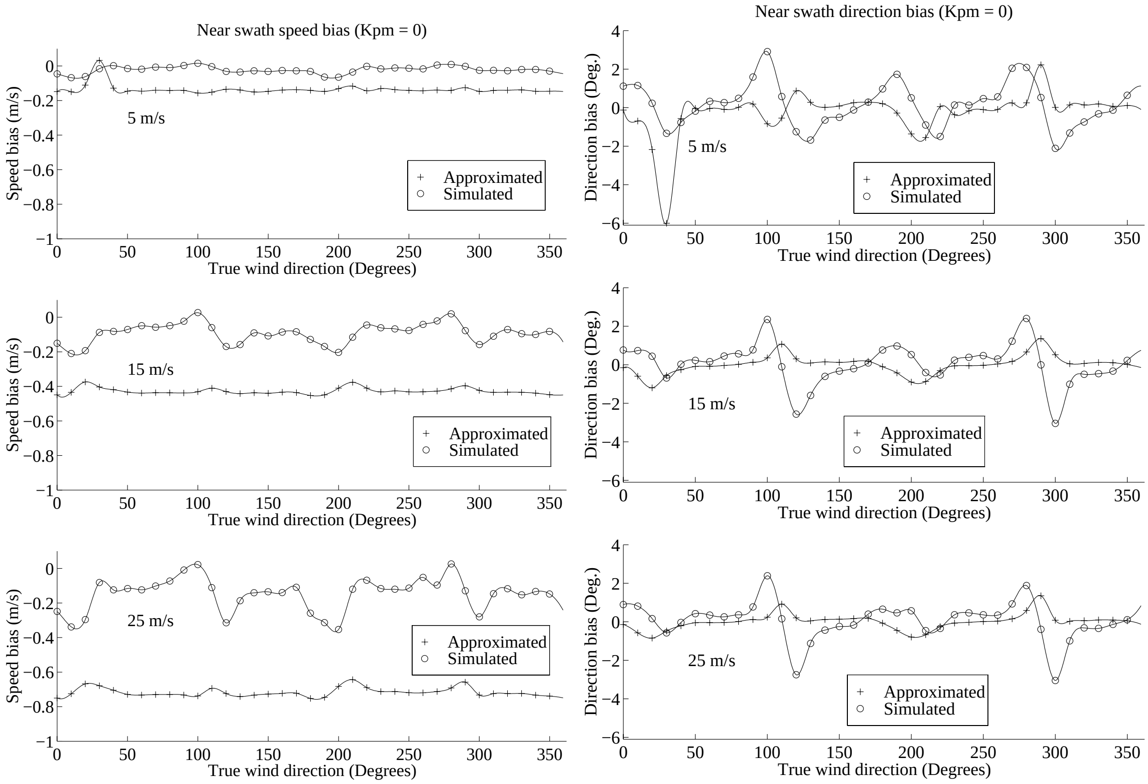
Figure B.2: *Comparison of approximate analytic bias with simulated bias for ERS-1 at near swath.* $K_{pm} = 0.17$*.*



(a) $K_{pm} = 0$

(b) $K_{pm} = 0.17$

Figure B.3: *Comparison of Cramer-Rao bound predictions on the standard deviations of wind speed and wind direction with simulations for ERS-1 at near swath. The standard deviation of wind speed is shown for a true wind speed of 25 m/s while the standard deviation of wind direction is shown for a true wind speed of 5 m/s.*



(a) Speed Bias

(b) Direction Bias

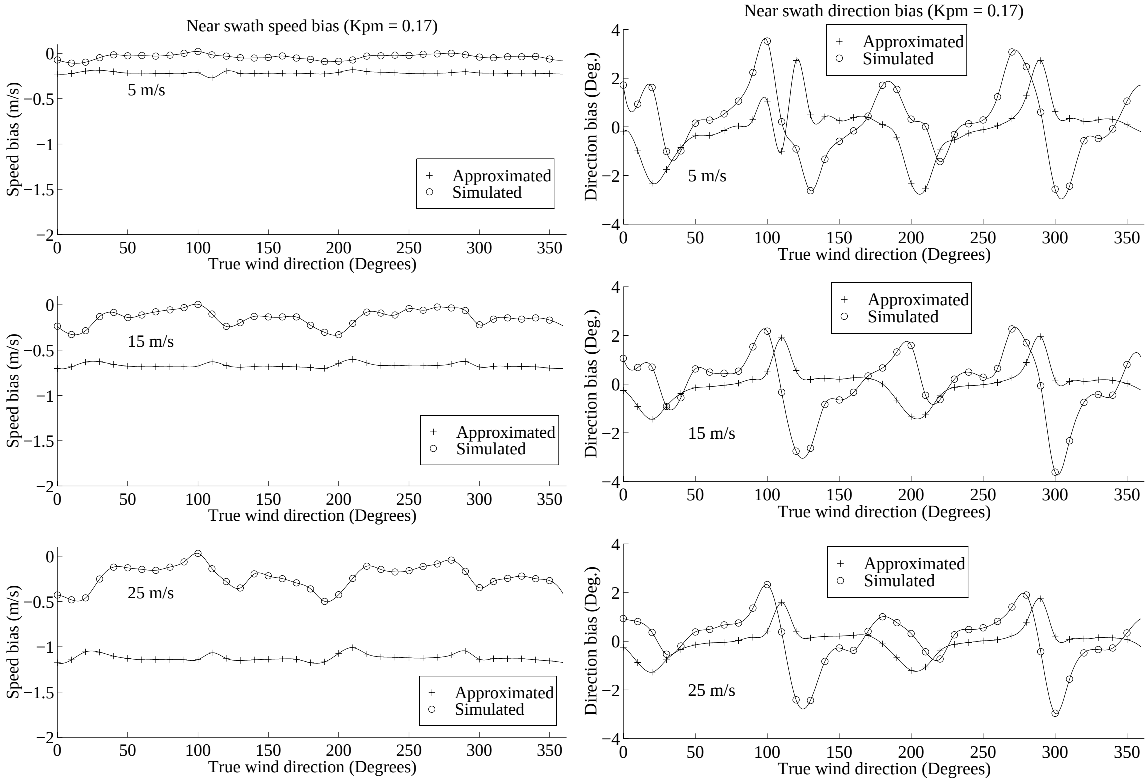
Figure B.4: *Comparison of approximate analytic bias with simulated bias for NSCAT at near swath.* $K_{pm} = 0$.

B.3.2 NSCAT

Two examples are presented for NSCAT geometry at far swath. One example assumes $K_{pm} = 0$, and the other assumes $K_{pm} = 0.17$. Plots similar to those shown for ERS-1 data are presented. Again, we conclude that the biased Cramer-Rao bound improves the agreement with simulations in situations where the unbiased Cramer-Rao bound predicts unusually high variance. However, it still does not resolve all of the discrepancies.

B.4 Conclusion

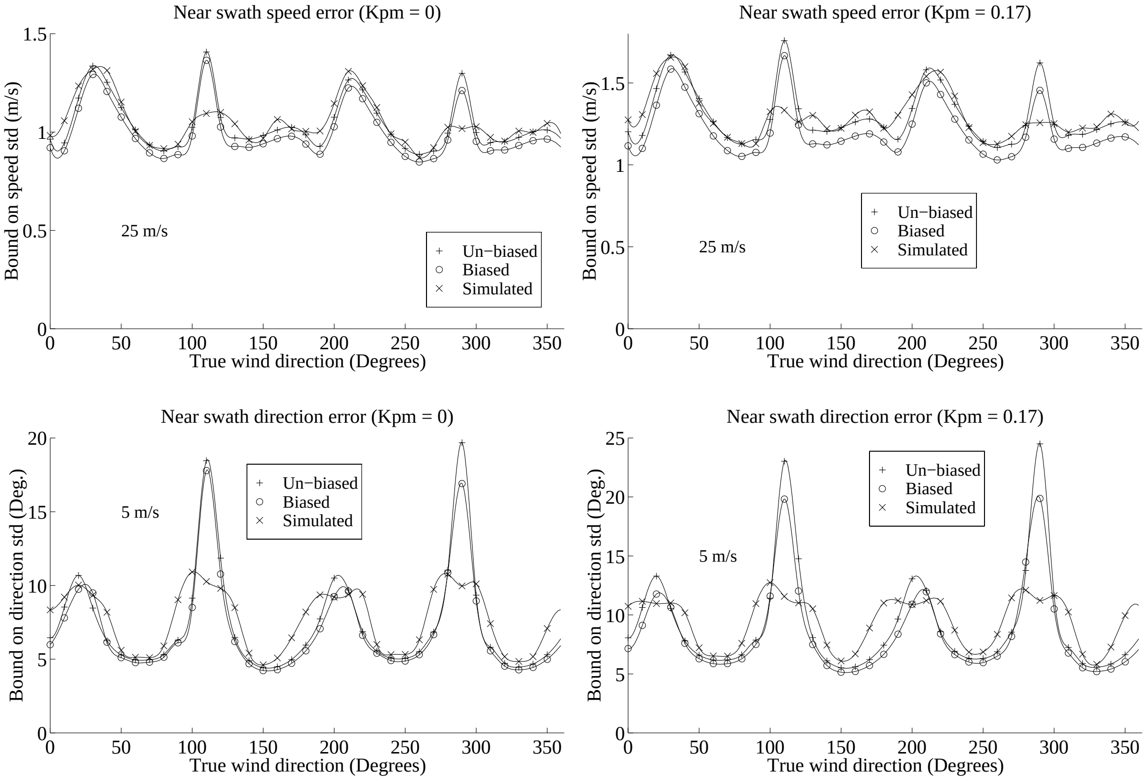
In this appendix we have derived an approximation to the wind retrieval bias along with an approximation to the gradient of the bias. The gradient is useful in



(a) Speed Bias

(b) Direction Bias

Figure B.5: *Comparison of approximate analytic bias with simulated bias for NSCAT at near swath.* $K_{pm} = 0.17$.



(a) $K_{pm} = 0$ (b) $K_{pm} = 0.17$

Figure B.6: *Comparison of Cramer-Rao bound predictions on the standard deviations of wind speed and wind direction with simulations for NSCAT at near swath. The standard deviation of wind speed is shown for a true wind speed of 25 m/s while the standard deviation of wind direction is shown for a true wind speed of 5 m/s.*

calculating the biased Cramer-Rao bound. The biased Cramer-Rao bound helps explain the discrepancies between the unbiased Cramer-Rao bound and simulations for certain wind directions where the unbiased Cramer-Rao bound predicts high variance. Although the approximations derived here do not completely resolve all the discrepancies, they suggest that with more accurate approximations to the bias (obtainable by higher order expansions in the Taylor series used for $\mathbf{h}$), a more accurate bound could be realized.

Appendix C

DETAILED ALGEBRA FOR WIND ALIAS ELIMINATION

This appendix is provided to describe the missing steps in Chapter 4 in going from Eq. (4.10):

$$\Lambda(\mathbf{z}) = -\sum_{k=1}^{K}\left[\log\left(\frac{\varsigma_{k,n}}{\varsigma_{k,1}}\right) + \frac{(z_k - M_{k,n})^2}{2\varsigma_{k,n}^2} - \frac{(z_k - M_{k,1})^2}{2\varsigma_{k,1}^2}\right], \tag{C.1}$$

to Eq. (4.18):

$$Y = \sum_{k=1}^{K} c_k(x_k - d_k)^2, \tag{C.2}$$

where

$$\begin{aligned}
a &= \sum_{k=1}^{K}\log\left(\frac{\varsigma_{k,n}}{\varsigma_{k,1}}\right), && \text{(C.3)}\\
x_k &= \frac{(z_k) - \mathcal{M}_k}{\varsigma_k}, && \text{(C.4)}\\
b &= \sum_{k=1}^{K}\frac{(M_{k,1} - M_{k,n})^2}{2\left(\varsigma_{k,1}^2 - \varsigma_{k,n}^2\right)}, && \text{(C.5)}\\
Y &= -\Lambda(\mathbf{z}) - a + b, && \text{(C.6)}\\
c_k &= \frac{\varsigma_k^2}{2}\left(\frac{1}{\varsigma_{k,n}^2} - \frac{1}{\varsigma_{k,1}^2}\right), && \text{(C.7)}\\
d_k &= \frac{\varsigma_{k,n}^2(\mathcal{M}_k - \mathcal{M}_{k,1}) - \varsigma_{k,1}^2(\mathcal{M}_k - \mathcal{M}_{k,n})}{\varsigma_k(\varsigma_{k,1}^2 - \varsigma_{k,n}^2)}. && \text{(C.8)}
\end{aligned}$$

Beginning with,

$$\Lambda(\mathbf{z}) = -\sum_{k=1}^{K}\left[\log\left(\frac{\varsigma_{k,n}}{\varsigma_{k,1}}\right) + \frac{(z_k - M_{k,n})^2}{2\varsigma_{k,n}^2} - \frac{(z_k - M_{k,1})^2}{2\varsigma_{k,1}^2}\right], \tag{C.9}$$

it is straightforward to obtain,

$$-\Lambda(\mathbf{z}) - a = \sum_{k=1}^{K}\left[\frac{(\varsigma_k x_k + \mathcal{M}_k - \mathcal{M}_{k,n})^2}{2\varsigma_{k,n}^2} - \frac{(\varsigma_k x_k + \mathcal{M}_k - M_{k,1})^2}{2\varsigma_{k,1}^2}\right]. \tag{C.10}$$

Each term in the sum has the form

$$\frac{(px+q)^2}{r} - \frac{(px+s)^2}{t}. \tag{C.11}$$

Expanding and collecting in x we obtain,

$$\left(\frac{p^2}{r} - \frac{p^2}{t}\right) x^2 - 2\left(\frac{ps}{t} - \frac{pq}{r}\right) x + \frac{q^2}{r} - \frac{s^2}{t}, \tag{C.12}$$

$$cx^2 - 2dcx + \frac{q^2}{r} - \frac{s^2}{t}, \tag{C.13}$$

where

$$\begin{aligned}
c &= \left(\frac{p^2}{r} - \frac{p^2}{t}\right), \qquad &(\text{C.14})\\
d &= \frac{1}{c}\left(\frac{ps}{t} - \frac{pq}{r}\right), \qquad &(\text{C.15})\\
&= \frac{rt}{p^2(t-r)}\frac{psr-pqt}{rt}, \qquad &(\text{C.16})\\
&= \frac{rs-tq}{p(t-r)}. \qquad &(\text{C.17})
\end{aligned}$$

By completing the square this expression becomes,

$$c(x-d)^2 + f, \tag{C.18}$$

where

$$\begin{aligned}
f &= -cd^2 + \frac{q^2t - s^2r}{rt}, \qquad &(\text{C.19})\\
&= -\frac{p^2(t-r)(rs-tq)^2}{rtp^2(t-r)^2} + \frac{q^2t-s^2r}{rt}, \qquad &(\text{C.20})\\
&= \frac{(t-r)(q^2t-s^2r)-(rs-tq)^2}{rt(t-r)}, \qquad &(\text{C.21})\\
&= \frac{q^2t^2 - q^2rt - s^2rt + s^2r^2 - s^2r^2 + 2rtqs - q^2t^2}{rt(t-r)}, \qquad &(\text{C.22})\\
&= -\frac{rt(s-q)^2}{rt(t-r)}, \qquad &(\text{C.23})\\
&= -\frac{(q-s)^2}{(t-r)}. \qquad &(\text{C.24})
\end{aligned}$$

Thus, Eq. (C.10) can be written as

$$-\Lambda(\mathbf{z}) - a = \sum_{k=1}^{K} \left[c_k(x_k - d_k)^2 + f_k\right], \tag{C.25}$$

where

$$c_k = \frac{\varsigma_k^2}{2}\left(\frac{1}{\varsigma_{k,n}^2} - \frac{1}{\varsigma_{k,1}^2}\right), \tag{C.26}$$

$$d_k = \frac{\varsigma_{k,n}^2(\mathcal{M}_k - \mathcal{M}_{k,1}) - \varsigma_{k,1}^2(\mathcal{M}_k - \mathcal{M}_{k,n})}{\varsigma_k(\varsigma_{k,1}^2 - \varsigma_{k,n}^2)}, \tag{C.27}$$

$$f_k = -\frac{(\mathcal{M}_{k,1} - \mathcal{M}_{k,n})^2}{2\left(\varsigma_{k,1}^2 - \varsigma_{k,n}^2\right)}. \tag{C.28}$$

Consequently, defining

$$b = -\sum_{k=1}^{K} f_k, \tag{C.29}$$

we obtain

$$-\Lambda(\mathbf{z}) - a + b = \sum_{k=1} c_k(x_k - d_k)^2, \tag{C.30}$$

$$Y = \sum_{k=1} c_k(x_k - d_k)^2 \tag{C.31}$$

as desired.

Appendix D

RELATIONSHIP BETWEEN MODEL-BASED FISHER INFORMATION MATRIX AND POINT-WISE FISHER INFORMATION MATRICES WHEN F HAS RANK $2MN$.

In this appendix, it is shown that the model-based Fisher information can be constructed from the Fisher information matrices of the individual wind vectors in the region when the model matrix, F has rank $2MN$, where $M \times N$ is the region size.

Assume that the wind-field vector is formed by row-ordering a wind field.

$$\mathbf{W} = \begin{bmatrix} \mathrm{w}_1 \\ \vdots \\ \mathrm{w}_l \\ \vdots \\ \mathrm{w}_{MN} \end{bmatrix}. \tag{D.1}$$

For a linear model, $\mathbf{W} = F\mathbf{X}$, when F has rank $2MN$ then the model does not restrict the wind field in anyway. Consequently, the definition of $J_{\mathbf{W}}$ is

$$J_{\mathbf{W}} = E\left\{ \left[\frac{\partial L(\mathbf{W},\mathbf{Z})}{\partial \mathbf{W}} \right]^T \frac{\partial L(\mathbf{W},\mathbf{Z})}{\partial \mathbf{W}} \right\}, \tag{D.2}$$

where

$$L(\mathbf{W},\mathbf{Z}) = -\sum_{l=1}^{MN} \sum_{k=1}^{K(l)} \left[\frac{\left(Z_{k,l} - \mathcal{M}^2_{k,l}\right)}{2\varsigma^2_{Z_{k,l}}} + \frac{1}{2}\log\left(2\pi\varsigma^2_{Z_{k,l}}\right) \right]. \tag{D.3}$$

$J_{\mathbf{W}}$ is a $2MN \times 2MN$ block-matrix composed of $M \times N$, 2×2 blocks. The block in the mth row and nth column is

$$J_{\mathrm{w}_{m,n}} = E\left\{ \left[\frac{\partial L(\mathbf{W},\mathbf{Z})}{\partial \mathrm{w}_m} \right]^T \frac{\partial L(\mathbf{W},\mathbf{Z})}{\partial \mathrm{w}_n} \right\}. \tag{D.4}$$

The dependence of the log-likelihood equation on w_m occurs only in $\mathcal{M}_{k,m}$. As a result,

$$\frac{\partial L(\mathbf{W},\mathbf{Z})}{\partial \mathrm{w}_m} = \sum_{k=1}^{K(m)} \frac{\partial \mathcal{M}_{k,m}}{\partial \mathrm{w}_m} \frac{Z_{k,m} - \mathcal{M}_{k,m}}{\varsigma^2_{Z_{k,m}}} + \left[\frac{(Z_{k,m} - \mathcal{M}_{k,m})^2}{2\varsigma^4_{Z_{k,m}}} - \frac{1}{2\varsigma^2_{Z_{k,m}}} \right] \frac{\partial \varsigma^2_{Z_{k,m}}}{\partial \mathrm{w}_m} \tag{D.5}$$

Consequently, we write

$$\begin{aligned} J_{\mathbf{w}_{m,n}} &= E\left\{\sum_{k=1}^{K(m)} \frac{\partial \mathcal{M}_{k,m}}{\partial \mathbf{w}_m} \frac{Z_{k,m} - \mathcal{M}_{k,m}}{\varsigma^2_{Z_{k,m}}} + \left[\frac{(Z_{k,m} - \mathcal{M}_{k,m})^2}{2\varsigma^4_{Z_{k,m}}} - \frac{1}{2\varsigma^2_{Z_{k,m}}}\right] \frac{\partial \varsigma^2_{Z_{k,m}}}{\partial \mathbf{w}_m}\right\}^T \\ &\times \left\{\sum_{k=1}^{K(n)} \frac{\partial \mathcal{M}_{k,n}}{\partial \mathbf{w}_n} \frac{Z_{k,n} - \mathcal{M}_{k,n}}{\varsigma^2_{Z_{k,n}}} + \left[\frac{(Z_{k,n} - \mathcal{M}_{k,n})^2}{2\varsigma^4_{Z_{k,n}}} - \frac{1}{2\varsigma^2_{Z_{k,n}}}\right] \frac{\partial \varsigma^2_{Z_{k,n}}}{\partial \mathbf{w}_n}\right\}, \end{aligned} \tag{D.6}$$

where expectation is taken over the measurement vector $\mathbf{Z}$. If $m \neq n$, the independence assumption of the measurements implies that the expectation of the product can be written as the product of the expectations. Performing the expectations is then straightforward:

$$\begin{aligned} J_{\mathbf{w}_{m,n}} &= \left\{\sum_{k=1}^{K(m)} \frac{\partial \mathcal{M}_{k,m}}{\partial \mathbf{w}_m} \frac{\mathcal{M}_{k,m} - \mathcal{M}_{k,m}}{\varsigma^2_{Z_{k,m}}} + \left[\frac{\varsigma^2_{Z_{k,m}}}{2\varsigma^4_{Z_{k,m}}} - \frac{1}{2\varsigma^2_{Z_{k,m}}}\right] \frac{\partial \varsigma^2_{Z_{k,m}}}{\partial \mathbf{w}_m}\right\}^T \\ &\times \left\{\sum_{k=1}^{K(n)} \frac{\partial \mathcal{M}_{k,n}}{\partial \mathbf{w}_n} \frac{\mathcal{M}_{k,n} - \mathcal{M}_{k,n}}{\varsigma^2_{Z_{k,n}}} + \left[\frac{\varsigma^2_{Z_{k,n}}}{2\varsigma^4_{Z_{k,n}}} - \frac{1}{2\varsigma^2_{Z_{k,n}}}\right] \frac{\partial \varsigma^2_{Z_{k,n}}}{\partial \mathbf{w}_n}\right\}, \qquad \text{(D.7)} \\ &= \mathbf{0} \qquad m \neq n. \qquad \text{(D.8)} \end{aligned}$$

Thus, only blocks where $m = n$ survive the expectation so $J_{\mathbf{W}}$ can be written as a block-diagonal matrix where each block is 2×2:

$$J_{\mathbf{W}} = \begin{bmatrix} J_{\mathbf{w}_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_{\mathbf{w}_{MN}} \end{bmatrix}. \tag{D.9}$$

Appendix E

CALCULATING CURL AND DIVERGENCE OF MODEL WIND FIELD BY MATRIX MULTIPLICATION

In spectral analysis of the curl and divergence of ERS1-retrieved winds, a discrepancy between the spectra of the curl and divergence retrieved using two different methods is observed. The first method uses matrix multiplication between the $\mathbf{X}$ vector and a matrix designed to selected the curl coefficients in $\mathbf{X}$ to get the assumed low order bivariate polynomial value of the curl at each point. A similar matrix is used to obtain the divergence. The second method, uses first differences on the model-based wind itself to approximate the curl and divergence.

For both the curl and divergence spectra, the two methods produce spectra that agree well until wavenumbers are high enough that the waves they represent fit inside one of the 10×10 regions. Inside this region the two methods produce divergent spectra. This suggests that the two methods for computing curl and divergence are not equivalent. Since the first difference approximation is believable, the model formulation which dictates how curl and divergence are extracted from $\mathbf{X}$ is incorrect. The matrices used to extract curl and divergence from $\mathbf{X}$ have zero entries for locations corresponding to boundary conditions in the $\mathbf{X}$ vector, apparently indicating that the boundary conditions have an effect on the curl (and divergence) of the field, a fact not previously considered.

As the model-based wind retrieval method still does an effective job of retrieving many wind fields, it is desirable to find a matrix that can be multiplied by the $\mathbf{X}$ vector to obtain the correct curl and divergence. The approach to this problem used in this appendix is to find a numerical derivative that uses all of the available data in a region to find the curl and divergence. This numerical derivative is given in terms of a matrix that can be multiplied by a vector (or a matrix) representing a sampled function with the result a vector (or a matrix) that is the derivative of the function at each sampled value. The derivation of this matrix is considered in the following.

E.1 Derivation of the Approximation of *Maximum* Order Derivative Matrix

The Problem: Given an $M \times N$ matrix U with elements u_{ij} representing uniform samples at $x = is$ and $y = jt$ of a function $f(x, y)$, find matrices that can be multiplied by U to approximate, using all available data, $\partial f/\partial i$ and $\partial f/\partial j$.

For the remainder of this discussion the function $f(is, jt)$ is written as $f(i, j)$. The problem is solved by first finding the correct matrix for $\partial f/\partial i$ and then making a slight modification to find the correct matrix for $\partial f/\partial j$.

This problem can be solved using two seemingly different approaches, yet each returns the same solution. This first approach is to interpolate (extrapolate) a derivative estimator function using Lagrange interpolating polynomials. The second approach is to calculate the derivative of the interpolating polynomial of the function itself.

E.1.1 Interpolating the derivative estimator

One approach to finding a numerical approximation to the derivative with respect to i is to define the following derivative estimator function:

$$\widehat{g}_{mn}(x) = \frac{f(m + x, n) - f(m, n)}{x}. \tag{E.1}$$

When x is an appropriate integer (between 1 and the size, M, of the matrix), this definition can be written with matrix notation as,

$$\widehat{g}_{mn}(k) = g_k = \frac{u_{m+k,n} - u_{m,n}}{k}. \tag{E.2}$$

Letting k run from $1 - m$ to $M - m$ (skipping $k = 0$) in column n of U, $M - 1$ samples of the derivative estimator function can be constructed for each column. Then, noting that

$$\widehat{g}_{mn}(0) \equiv \lim_{k \to 0} \widehat{g}_{mn}(k) = \left.\frac{\partial f}{\partial i}\right|_{mn}, \tag{E.3}$$

interpolation or extrapolation of the constructed samples of $\widehat{g}_{mn}(x)$ to $x = 0$ will provide an approximation to the derivative. The method of interpolation used is Lagrange polynomial interpolation.

In an $M \times N$ matrix U there is enough data for $M - 1$ samples of the function $\widehat{g}_{mn}(x)$. As mentioned before, these occur at $x = k = 1 - m \ldots - 1, 1 \ldots M - m$. The

Lagrange interpolating polynomial (as a function of x) that fits these sample points is given by

$$P_{mn}(x) = \sum_{\substack{k=1-m \\ k\neq 0}}^{M-m} g_k \prod_{\substack{l=1-m \\ l\neq k,0}}^{M-m} \frac{x-l}{k-l}. \tag{E.4}$$

Therefore, $\widehat{g}_{mn}(0)$ can be approximated by $P_{mn}(0)$:

$$\begin{aligned}
P_{mn}(0) &= \sum_{\substack{k=1-m \\ k\neq 0}}^{M-m} g_k \prod_{\substack{l=1-m \\ l\neq k,0}}^{M-m} \frac{-l}{k-l} \\
&= \sum_{k=1-m}^{-1} g_k \left(\prod_{\substack{l=1-m \\ l\neq k}}^{-1} \frac{-l}{k-l} \right) \left(\prod_{l=1}^{M-m} \frac{-l}{k-l} \right) + \sum_{k=1}^{M-m} g_k \left(\prod_{l=1-m}^{-1} \frac{-l}{k-l} \right) \left(\prod_{\substack{l=1 \\ l\neq k}}^{M-m} \frac{-l}{k-l} \right) \\
&= \sum_{k=1}^{m-1} g_{-k} \left(\prod_{\substack{l=1 \\ l\neq k}}^{m-1} \frac{l}{l-k} \right) \left(\prod_{l=1}^{M-m} \frac{l}{k+l} \right) + \sum_{k=1}^{M-m} g_k \left(\prod_{l=1}^{m-1} \frac{l}{k+l} \right) \left(\prod_{\substack{l=1 \\ l\neq k}}^{M-m} \frac{l}{l-k} \right) \\
&= \sum_{k=1}^{m-1} g_{-k} \left((-1)^{k-1} \frac{(m-1)!}{(m-1-k)!} \right) \left(\frac{(M-m)!}{(M-m+k)!} \right) \\
&\quad + \sum_{k=1}^{M-m} g_k \left(\frac{(m-1)!}{(m-1+k)!} \right) \left((-1)^{k-1} \frac{(M-m)!}{(M-m-k)!} \right).
\end{aligned}$$

Substituting Equation (E.2) into this result gives $P_{mn}(0) \approx \left. \partial f / \partial i \right|_{mn}$ as a linear combination of elements in the nth column of U:

$$\begin{aligned}
P_{mn}(0) &= \sum_{k=1}^{m-1} \left(\frac{u_{m-k,n} - u_{m,n}}{k} \right) \left((-1)^k \frac{(m-1)!}{(m-1-k)!} \right) \left(\frac{(M-m)!}{(M-m+k)!} \right) \\
&\quad - \sum_{k=1}^{M-m} \left(\frac{u_{m+k,n} - u_{m,n}}{k} \right) \left(\frac{(m-1)!}{(m-1+k)!} \right) \left((-1)^k \frac{(M-m)!}{(M-m-k)!} \right).
\end{aligned}$$

Since the above approximation for $\left. \partial f / \partial i \right|_{mn}$ is a linear combination of the elements in the nth column of U, it can be expressed as a matrix equation relating the matrix U to its row-derivative at every point by,

$$\widehat{D}_{U,\text{rows}} = D_i(M) U$$

where $D_i(M)$ is a $M \times M$ matrix whose elements $d^i_{m,n}$ are defined by

$$d^i_{m,n} = \begin{cases} \displaystyle\sum_{k=1}^{M-m} \frac{(-1)^k (M-m)!(m-1)!}{k(M-m-k)!(m-1+k)!} - \sum_{k=1}^{m-1} \frac{(-1)^k (M-m)!(m-1)!}{k(M-m+k)!(m-1-k)!} & m = n, \\ \displaystyle\frac{(-1)^{m-n}(M-m)!(m-1)!}{(m-n)(M-n)!(n-1)!} & m \neq n. \end{cases}$$

Letting $l = k + m$ in the first diagonal summation and $l = m - k$ in the second this can be rewritten as

$$d^i_{m,n} = \begin{cases} \displaystyle\sum_{\substack{l=1 \\ l \neq m}}^{M} \frac{(-1)^{l-m}(M-m)!(m-1)!}{(l-m)(l-1)!(M-l)!} & m = n, \\ \displaystyle\frac{(-1)^{m-n}(M-m)!(m-1)!}{(m-n)(M-n)!(n-1)!} & m \neq n. \end{cases} \tag{E.5}$$

This is the desired result as it gives a matrix for computing a *maximum* order approximation to $\partial f/\partial i|_{mn}$ for each m and n in the matrix U.

This idea can be easily extended to determine a matrix expression for finding $\partial f/\partial j|_{mn}$ for each m and n in U by realizing that,

$$\widehat{D}_{U,\text{cols}} = \left(\widehat{D}_{U^T,\text{rows}}\right)^T = \left(D_i(N)U^T\right)^T = UD_i^T(N) = UD_j(N). \tag{E.6}$$

Notice that if U is $M \times N$, then $D_j(N)$ is an $N \times N$ version of the D_i^T matrix.

For clarity some examples of $D_i(M)$ are given below.

$$D_i(4) = \begin{bmatrix} -\frac{11}{6} & 3 & -3/2 & 1/3 \\ -1/3 & -1/2 & 1 & -1/6 \\ 1/6 & -1 & 1/2 & 1/3 \\ -1/3 & 3/2 & -3 & \frac{11}{6} \end{bmatrix},$$

$$D_i(5) = \begin{bmatrix} -\frac{25}{12} & 4 & -3 & 4/3 & -1/4 \\ -1/4 & -5/6 & 3/2 & -1/2 & 1/12 \\ 1/12 & -2/3 & 0 & 2/3 & -1/12 \\ -1/12 & 1/2 & -3/2 & 5/6 & 1/4 \\ 1/4 & -4/3 & 3 & -4 & \frac{25}{12} \end{bmatrix},$$

$$D_i(6) = \begin{bmatrix} -\frac{137}{60} & 5 & -5 & 10/3 & -5/4 & 1/5 \\ -1/5 & -\frac{13}{12} & 2 & -1 & 1/3 & -1/20 \\ 1/20 & -1/2 & -1/3 & 1 & -1/4 & 1/30 \\ -1/30 & 1/4 & -1 & 1/3 & 1/2 & -1/20 \\ 1/20 & -1/3 & 1 & -2 & \frac{13}{12} & 1/5 \\ -1/5 & 5/4 & -10/3 & 5 & -5 & \frac{137}{60} \end{bmatrix}.$$

E.1.2 Differentiating the interpolating function

A second approach can be used to derive the same $D_i(M)$ matrix. This approach has the advantage of returning a simpler expression for the diagonal elements. This approach uses the samples of $f(x, y)$ contained in a given column of matrix U ($u_{k,n}$ where $k = 1 \dots M$) directly to construct a Lagrange interpolating polynomial $Q_n(x)$ for $f(x, n)$. Then $\partial f/\partial i|_{mn}$ is approximated by taking the derivative of $Q_n(x)$ directly and evaluating $Q_n'(x)$ at $x = m$.

The Lagrange interpolating polynomial for $f(x, n)$ using the M elements of the nth column of U is given by

$$Q_n(x) = \sum_{k=1}^{M} u_{k,n} L_k(x), \tag{E.7}$$

with

$$L_k(x) = \prod_{\substack{l=1 \\ l \neq k}}^{M} \frac{x - l}{k - l}.$$

The derivative of this polynomial evaluated at $x = m$, $Q_n'(m) (\approx \partial f/\partial i|_{mn})$, can be written as

$$Q_n'(m) = \sum_{k=1}^{M} u_{k,n} \; L_k'(x)|_{x=m} \,. \tag{E.8}$$

Multiple application of the product rule reveals that

$$L_k'(x) = \left(\prod_{\substack{l=1 \\ l \neq k}}^{M} \frac{1}{k - l} \right) \left(\sum_{\substack{l=1 \\ l \neq k}}^{M} \prod_{\substack{o=1 \\ o \neq k,l}}^{M} (x - o) \right).$$

Equation (E.8) can be rewritten as a matrix equation to relate the matrix U to its row-derivative at each point,

$$\widehat{D}_{U,\text{rows}} = D_i(M)U, \tag{E.9}$$

where this time, the elements of $D_i(M)$ are given by $d^i_{m,n} = L'_n(m)$. To get a better understanding of this matrix, $L'_n(m)$ is evaluated below for $m = n$ and $m \neq n$.

Evaluating $L'_n(m)$ for $m = n$ gives

$$\begin{aligned}
L'_m(m) &= \left(\prod_{\substack{l=1\\l\neq m}}^{M} \frac{1}{m-l}\right)\left(\sum_{\substack{l=1\\l\neq m}}^{M}\left[\frac{1}{(m-l)}\prod_{\substack{o=1\\o\neq m}}^{M}(m-o)\right]\right)\\
&= \frac{(-1)^{m-M}}{(m-1)!(M-m)!}\sum_{\substack{l=1\\l\neq m}}^{M}\frac{(-1)^{M-m}(m-1)!(M-m)!}{(m-l)}\\
&= \sum_{\substack{l=1\\l\neq m}}^{M}\frac{1}{m-l}.
\end{aligned} \tag{E.10}$$

Recognizing that for $m \neq n$,

$$\prod_{\substack{o=1\\o\neq n,l}}^{M}(m-o) = \begin{cases} \displaystyle\prod_{\substack{o=1\\o\neq n,m}}^{M}(m-o) & l = m,\\ 0 & l \neq m,\end{cases}$$

allows simplifying $L'_n(m)$ for $m \neq n$:

$$\begin{aligned}
L'_n(m) &= \left(\prod_{\substack{l=1\\l\neq n}}^{M}\frac{1}{n-l}\right)\left(\sum_{\substack{l=1\\l\neq n}}^{M}\prod_{\substack{o=1\\o\neq n,l}}^{M}(m-o)\right)\\
&= \left(\prod_{\substack{l=1\\l\neq n}}^{M}\frac{1}{n-l}\right)\left(\prod_{\substack{o=1\\o\neq n,m}}^{M}(m-o)\right)\\
&= \left(\frac{(-1)^{n-M}}{(n-1)!(M-n)!}\right)\left(\frac{(-1)^{M-m}(m-1)!(M-m)!}{(m-n)}\right)\\
&= \frac{(-1)^{n-m}(m-1)!(M-m)!}{(m-n)(n-1)!(M-n)!}\\
&= \frac{(-1)^{m-n}(M-m)!(m-1)!}{(m-n)(M-n)!(n-1)!}.
\end{aligned} \tag{E.11}$$

Using Equations (E.10) and (E.11), the elements of $D_i(M)$, $d^i_{m,n} = L'_n(m)$ can be explicitly written as

$$d^i_{m,n} = \begin{cases} \displaystyle\sum_{\substack{l=1 \\ l \neq m}}^{M} \frac{1}{m-l} & m = n, \\ \dfrac{(-1)^{m-n}(M-m)!(m-1)!}{(m-n)(M-n)!(n-1)!} & m \neq n. \end{cases} \tag{E.12}$$

The off-diagonal elements can immediately be seen to correspond to those given in the earlier expression for $D_i(M)$. Empirical evidence suggests that the diagonal elements are the same as well although a proof is not immediately obvious. We expect the two matrices to be the same since the two methods both use a Lagrange polynomial interpolator. The second derivation, however, gives a much simpler expression for the diagonal elements.

E.2 Using the MODA Matrix with a Row-scanned matrix

In the previous section it was shown that an approximation of *maximum* order for $\partial f/\partial i|_{mn}$, where samples $f(i,j)$ are elements $u_{i,j}$ of the $M \times N$ matrix U, can be found with the following equation:

$$\widehat{D}_{U,\text{rows}} = D_i(M)U. \tag{E.13}$$

It was also shown that a similar expression can be used to find an approximation of *maximum* order for $\partial f/\partial j|_{mn}$:

$$\widehat{D}_{U,cols} = UD_i^T(N). \tag{E.14}$$

Let $\overline{X}$ be a row-scanned version of X. It can be shown that the matrix equation

$$A = BC,$$

where A is $M \times N$, B is $M \times P$, and C is $P \times N$, can be written as either the matrix-vector equation

$$\overline{A} = (B \otimes I_{N\times N})\overline{C}$$

or the matrix-vector equation

$$\overline{A} = (I_{M\times M} \otimes C^T)\overline{B}.$$

Here $\otimes$ represents the Kronecker product for matrices. Using this fact, Equations (E.13) and (E.14) can be written as

$$\overline{\widehat{D}}_{U,rows} = D_x\overline{U}, \tag{E.15}$$

$$\overline{\widehat{D}}_{U,cols} = D_y\overline{U}, \tag{E.16}$$

where

$$D_x = D_i(M) \otimes I_{N\times N},$$

$$D_y = I_{M\times M} \otimes D_i(N).$$

E.3 Using MODA to find Γ

Given a wind field over an MxN region represented by a vector $\mathbf{X}$, with $\mathbf{W} = F\mathbf{X}$ and

$$\mathbf{W} = \begin{bmatrix} \overline{U} \\ \overline{V} \end{bmatrix},$$

the MODA matrix can be used to find the curl and divergence over the region from the $\mathbf{X}$ vector by multiplication by appropriate Γ matrices. Using the MODA matrix is appropriate since the wind field represented by $\mathbf{X}$ is a well-filtered field with no high-frequency components.

Represent the row-scanned curl of the wind field in a region by $\overline{C}$, and represent the row-scanned divergence by $\overline{D}$. They are related to the row-scanned wind field components $\overline{U}$ and $\overline{V}$ through the MODA matrix:

$$\begin{aligned} \overline{C} &= -D_y\overline{U} + D_x\overline{V}, \\ \overline{D} &= D_x\overline{U} + D_y\overline{V}. \end{aligned}$$

These equations can be rewritten as a matrix equation:

$$\begin{bmatrix} \overline{C} \\ \overline{D} \end{bmatrix} = \begin{bmatrix} -D_y & D_x \\ D_x & D_y \end{bmatrix} \begin{bmatrix} \overline{U} \\ \overline{V} \end{bmatrix}. \tag{E.17}$$

Since $\mathbf{W} = F\mathbf{X}$, we can rewrite Equation (E.17) as

$$\mathbf{H} = \Gamma\mathbf{X}. \tag{E.18}$$

In this equation,

$$\mathbf{H} = \begin{bmatrix} \overline{C} \\ \overline{D} \end{bmatrix},$$

and

$$\Gamma = \begin{bmatrix} -D_y & D_x \\ D_x & D_y \end{bmatrix} F. \tag{E.19}$$

Equation (E.19) gives an appropriate Γ matrix to use to obtain the curl and divergence from the X parameters of the wind field.

Appendix F

MATLAB CODE FOR CALCULATING LEGENDRE AND FOURIER F MATRICES

F.1 Legendre F matrix

This MATLAB function takes as arguments the size of the region and the model order in both u and v and returns the appropriate F matrix using a Legendre polynomial model.

```
function Fr = legebas(M,N,Mu,Mv)

% function Fr = legebas(M,N,Mu,Mv)
%
%   Returns a Legendre-polynomial Fr matrix for computing
%     the wind from the X parameters in the
%     model-based wind retrieval technique.  MxN is size of region,
%     Mu is u-component model order and Mv is v-component model order.
%

%
% Calculate Yu and Yv such that U = Yu*X and V = Yv*X
%

Nu = (Mu+1)*(Mu+2)/2;
Nv = (Mv+1)*(Mv+2)/2;

Yu = [];
Yv = [];
Qmn = zeros(M*N,1);

for m = 0:Mu
   for n = 0:Mu-m,
      for k=1:M*N,
          row = floor((k-1)/N)+1;
          col = rem(k-1,N)+1;
          Qmn(k) = lege(row,M,m)*lege(col,N,n);
      end
      Yu = [Yu Qmn];
   end
end

for m = 0:Mv
   for n=0:Mv-m
      for k=1:M*N,
          row = floor((k-1)/N)+1;
          col = rem(k-1,N)+1;
          Qmn(k) = lege(row,M,m)*lege(col,N,n);
      end
      Yv = [Yv Qmn];
   end
end
```

```
Yu = [Yu zeros(N^2,Nv)];
Yv = [zeros(N^2,Nu) Yv];

Fr = [Yu;Yv];

++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
function phi = lege(k,N,n)

% Computes the Legendre polynomial of order n evaluated at k
%  where k lies in interval [1,N].

phi = 0;
for l = 0:floor(n/2),
    phi = phi + (-1)^l*binom(n,l)*binom(2*n-2*l,n)*(2*(k-1)/(N-1)-1)^(n-2*l);
end
phi = phi/(2^n);
```

F.2 Fourier F matrix

This MATLAB function takes as arguments the size of the region and the model order in both u and v and returns the appropriate F matrix using a Fourier-series model.

```
function Fr = fourierbas(M,N,Mu,Mv)

% function Fr = fourierbas(M,N,Mu,Mv)
%
%   Returns a Fourier Fr matrix for computing
%     the wind from the X parameters in the
%     model-based wind retrieval technique.  MxN is size of region,
%     Mu is u-component model order and Mv is v-component model order.
%

%
% Calculate Yu and Yv such that U = Yu*X and V = Yv*X
%

Nu = (Mu+1)*(Mu+2) - 1;
Nv = (Mv+1)*(Mv+2) - 1;

Yu = ones(M*N,1);
Yv = ones(M*N,1);
Qcm = zeros(M*N,1);
Qsm = zeros(M*N,1);

for m = 0:Mu
    for n = 0:Mu-m
    if (n+m)>0,
      for k=1:M*N,
          row = floor((k-1)/N)+1;
          col = rem(k-1,N)+1;
          Qcm(k) = cos(m*pi*row/M + n*pi*col/N);
          Qsm(k) = sin(m*pi*row/M + n*pi*col/N);
      end
      Yu = [Yu Qcm Qsm];
    end
   end
end

for m = 0:Mv
```

```
   for n = 0:Mu-m
    if (n+m)>0,
      for k=1:M*N,
          row = floor((k-1)/N)+1;
          col = rem(k-1,N)+1;
          Qcm(k) = cos(m*pi*row/M + n*pi*col/N);
          Qsm(k) = sin(m*pi*row/M + n*pi*col/N);
      end
      Yv = [Yv Qcm Qsm];
    end
   end
end

Yu = [Yu zeros(M*N,Nv)];
Yv = [zeros(M*N,Nu) Yv];

Fr = [Yu;Yv];
```