# AN IMPLEMENTATION OF FIELD-WISE WIND RETRIEVAL FOR SEAWINDS ON QUIKSCAT

by

Andrew S. Fletcher

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Electrical and Computer Engineering

Brigham Young University

August 2001

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Andrew S. Fletcher

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

| | |
|---|---|
| Date | Dr. David G. Long, Chair |
| Date | Dr. A. Lee Swindlehurst |
| Date | Dr. Karl F. Warnick |

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Andrew S. Fletcher in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Dr. David G. Long
Chair, Graduate Committee

Accepted for the Department

A. Lee Swindlehurst
Graduate Coordinator

Accepted for the College

Douglas M. Chabries
Dean, College of Engineering and Technology

ABSTRACT

AN IMPLEMENTATION OF FIELD-WISE WIND RETRIEVAL FOR SEAWINDS ON QUIKSCAT

Andrew S. Fletcher

Department of Electrical and Computer Engineering

Master of Science

Field-wise wind estimation (also known as model-based wind estimation) is a sophisticated technique to derive wind estimates from radar backscatter measurements. In contrast to the more traditional method known as point-wise wind retrieval, field-wise techniques estimate wind field model parameters. In this way, neighboring wind vectors are jointly estimated, ensuring consistency. This work presents an implementation of field-wise wind retrieval for the SeaWinds scatterometer on the QuikSCAT satellite.

Due to its sophistication, field-wise wind retrieval adds computational complexity and intensity. The tradeoffs necessary for practical implementation are examined and quantified. The Levenberg-Marquardt algorithm for minimizing the field-wise objective function is presented. As the objective function has several near-global local minima, several wind fields represent ambiguous wind field estimates. A deterministic method is proposed to ensure sufficient ambiguities are obtained. An improved method for selecting between ambiguous wind field estimates is also proposed.

With a large set of SeaWinds measurements and estimates available, the $\sigma^o$ measurement statistics are examined. The traditional noise model is evaluated for accuracy. A data-driven parameterization is proposed and shown to effectively estimate measurement bias and variance. The parameterized measurement model is used to generate Cramer-Rao bounds on estimator performance. Using the Cramer-Rao bound, field-wise and point-wise performances are compared.

# ACKNOWLEDGMENTS

First acknowledgments must go to David Long, who introduced me to the field of remote sensing. He had the vision to take me on as a research assistant when I was an inexperienced undergraduate. He has taught me many valuable principles of research that will serve as the foundation for my career.

I greatly value the associations that I obtained in the MERS lab. I grew intellectually by interacting with its talented students, and I feel that I have gained many life-long friends. Particular thanks to David Draper. He and I have worked together on much of the material in this thesis. His advice and listening ear were invaluable.

Finally, foremost thanks go to my wife Mary Beth. From the moment we met, she has been an inspiration to me. She has taken my dreams of graduate work to be her own, and has supported me every step of the way. I dedicate this work to her and to our daughter Erin.

# Contents

**Acknowledgments** **vi**

**List of Figures** **xiii**

**1 Introduction** **1**

1.1 Topic Motivation . . . . . 1

1.2 Scatterometer Wind Estimation . . . . . 2

1.3 Thesis Overview . . . . . 3

**2 Background** **5**

2.1 Scatterometry . . . . . 5

2.1.1 SeaWinds on QuikSCAT . . . . . 6

2.2 The Geophysical Model Function . . . . . 8

2.3 Estimating the Wind . . . . . 8

2.4 Point-wise Wind Retrieval . . . . . 12

2.4.1 The Point-wise Objective Function . . . . . 13

2.4.2 Point-wise Median Filter . . . . . 14

2.5 Field-wise Wind Retrieval . . . . . 15

2.5.1 The Wind Field Model . . . . . 15

2.5.2 The Field-wise Objective Function . . . . . 17

2.5.3 Field-wise Ambiguity Selection . . . . . 18

**3 Minimizing the Field-wise Objective Function** **21**

3.1 The Field-wise Objective Function . . . . . 21

3.2 The Levenberg-Marquardt Algorithm . . . . . 23

3.3 Parameter Choices . . . . . 26

3.4 Objective Function Minima . . . . . 28
3.4.1 Random Initialization Set . . . . . 29
3.4.2 Deterministic Initialization Set . . . . . 29
3.5 Results . . . . . 33

**4 Field-wise Ambiguity Selection** **39**
4.1 Median Filtering . . . . . 39
4.2 The Point-wise Median Filter . . . . . 40
4.3 The Field-wise Median Filter . . . . . 42
4.3.1 Combining Overlapping Regions . . . . . 43
4.3.2 Algorithm Summary . . . . . 44
4.3.3 Parameter Choices . . . . . 44
4.4 Results . . . . . 46
4.5 Spatial Frequency Enhancement . . . . . 48
4.5.1 Frequency Content of the Wind . . . . . 48
4.5.2 Increasing the Frequency Resolution . . . . . 49
4.5.3 High Model Order Results . . . . . 50

**5 Examination of the Backscatter Noise Model** **57**
5.1 Evaluating the Noise Model . . . . . 58
5.2 Data-driven Parameterization of Backscatter Measurements . . . . . 59
5.3 Estimator Performance Bounds . . . . . 64
5.3.1 Performance Comparisons using the C-R Bound . . . . . 66

**6 Conclusion** **71**
6.1 Summary of Contributions . . . . . 71
6.1.1 Objective Function Minimization . . . . . 71
6.1.2 Swath Estimate Selection to Span the Wind Spectrum . . . . 72
6.1.3 Measurement Noise Model and Estimator Performance . . . . 72
6.2 Future Research . . . . . 73

**A Field-wise Objective Function Derivatives** **75**

A.1 Field-wise Objective Functions . . . . . 75

A.1.1 The Squared Error (SE) Objective Function . . . . . 76

A.1.2 The Weighted Squared Error (WSE) Objective Function . . . 76

A.1.3 Maximum Likelihood (ML) Estimation . . . . . 77

A.1.4 The Reduced Maximum Likelihood (RML) Objective Function 78

A.2 Objective Function Gradients . . . . . 79

A.2.1 SE Objective Function Gradient . . . . . 79

A.2.2 WSE Objective Function Gradient . . . . . 81

A.2.3 RML Objective Function Gradient . . . . . 81

A.2.4 ML Objective Function Gradient . . . . . 82

A.3 Objective Function Hessian Matrices . . . . . 82

A.3.1 SE Objective Function Hessian . . . . . 82

A.3.2 The WSE Hessian Matrix . . . . . 86

**Bibliography** **88**

## List of Figures


2.1 The spacecraft and antenna geometries for SeaWinds on Quikscat. . . 6

2.2 A top view of the SeaWinds swath showing the two overlapping swaths from the inner and outer beam. . . . . 7

2.3 The Geophysical Model Function evaluated with $\theta$ of 46°. . . . . 9

2.4 A single backscatter measurement (with fixed geometries) is the range of a continuous domain of wind vector estimates projected through the GMF. . . . . 10

2.5 To resolve a unique wind vector estimates, several measurements at varying geometries are required. . . . . 11

2.6 In the presence of noise, no unique wind estimate can be determined. 12

2.7 The first 6 Karhuenen-Loeve basis fields. . . . . 17

2.8 The SeaWinds swath gridded into wind field regions as per the Richards method. . . . . 19

3.1 The eigenvalues of the wind autocorrelation matrix. . . . . 28

3.2 Top: Typical WSE objective function values for a region. Bottom: Difference between objective values of consecutive solutions. One obvious breakpoint occurs after solution 4. . . . . 31

3.3 The percentage of "pre-break" solutions against the number of initial searches. . . . . 32

3.4 Point-wise and field-wise products for QuikSCAT Rev 01289 beginning in the upper left corner at alongtrack 333 and crosstrack 35. . . . . 34

3.5 Point-wise and field-wise products for QuikSCAT Rev 01289 beginning at alongtrack 369 and crosstrack 35. . . . . 36

3.6 Point-wise and field-wise products for QuikSCAT Rev 03079 beginning at alongtrack 457 and crosstrack 59. . . . . 37
4.1 Typical point-wise wind field with each wind vector set to the first ambiguity. . . . . 41
4.2 Average modeling error as a function of region location. . . . . 45
4.3 Region overlap for field-wise estimation. . . . . 46
4.4 A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 3081 beginning at alongtrack 150. . . . . 51
4.5 A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 2176 beginning at alongtrack 415. . . . . 52
4.6 A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 04790 beginning at alongtrack 155. . . . 53
4.7 A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 3081 beginning at alongtrack 150. The field-wise product is performed using 26 model parameters. . . . . 54
4.8 A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 2176 beginning at alongtrack 415. The field-wise product is performed using 26 model parameters. . . . . 55
4.9 A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 04790 beginning at alongtrack 155. The field-wise product is performed using 26 model parameters. . . . . 56
5.1 A histogram of the random variable $y = z - \mathcal{M}(\hat{\mathbf{w}})$ normalized with (a) $\varsigma_1$ and (b) $\varsigma_2$. . . . . 59
5.2 The joint distribution of $z$ and $\mathcal{M}$. . . . . 60
5.3 The first and second central moments of $y|z_0 = z - \mathcal{M}|z_0$. . . . . 61
5.4 A piecewise polynomial fit to the first and second moments of $y = z - \mathcal{M}$. 62
5.5 A histogram of $y'$, the parameter normalized backscatter. . . . . 63
5.6 The point-wise Cramer-Rao bound with the traditional variance estimate and the parameterized variance estimate. An estimated variance is also presented. . . . . 67

5.7 The field-wise Cramer-Rao bound is compared to the parameterized point-wise Cramer-Rao bound. . . . . 70

# Chapter 1

# Introduction

For several decades, space-borne radar systems called scatterometers have been used to estimate near-surface ocean winds. By examining the electromagnetic power scattered and reflected off of the ocean's surface, an estimate for both the wind speed and direction may be derived. The traditional estimation method, known as point-wise wind retrieval, is generally effective, although limitations to the method have led to the development of more sophisticated methods. This thesis proposes an implementation of one such method, known as field-wise wind retrieval, for the Scatterometer SeaWinds on QuikSCAT, a NASA satellite launched in June of 1999.

## 1.1 Topic Motivation

Weather prediction and climate studies have been pursuits of great interest to the scientific community. In recent years, this interest has intensified particularly as the global community have become more aware of humanity's effects on the global ecosystem. Concerns about global warming and long-term weather patterns have increased the desire to thoroughly understand the geophysics of our planet.

Fortunately, modern technology has provided many new avenues for geophysical study. With the advent of space flight, global data is available to scientists on a scale unimaginable just decades before. Earth observing satellites have the capability to make daily high resolution measurements over much of the Earth's surface. These measurements can be made passively (*e.g.,* measuring the Earth's radiation or by photography) or actively (*e.g.,* radar systems).

Since its development, radar has demonstrated the capability to observe geophysical phenomena. Natural surfaces have distinctive signatures when illuminated by an electromagnetic wave. Water, ice, vegetation, and dry land each reflect and scatter waves in a measurable and distinctive way.

Radar systems also have great advantages for use with a satellite platform. Unlike satellite photography, radar is not limited by cloud cover or daylight, but can make measurements through virtually all weather conditions.

## 1.2 Scatterometer Wind Estimation

Remote sensing successfully gathers data from previously difficult sources. When restricted to on-site data collection, geophysical data has been limited to the more accessible sites. Remote sensing, however, allows measurements to be taken over the Antarctic, the Arctic, deserts, jungles, and the Earth's oceans. Ocean studies have particular significance, as they comprise a large majority of the Earth's surface.

The radar signature of the ocean surface is not constant across the globe. Indeed, the presence of ocean waves and ripples in the surface affect the scattering characteristics of the water. In this way, radar measurements can obtain information about the ocean roughness. As near-surface winds are a major contributor to ocean surface roughness, radar instruments can collect data to estimate wind vectors.

In 1978, the SeaSat Scatterometer (SASS) successfully demonstrated that near-surface ocean winds are estimable using a space-borne radar measurement. Since then, the relationship between ocean winds and radar backscatter has been explored and utilized, with several more satellites launched to estimate the winds. In 1999, NASA launched the instrument SeaWinds on QuikSCAT, which has proven to be a powerful instrument for estimating ocean wind vectors. A large portion of the world's oceans are measured daily, and wind estimates are available on a global scale. SeaWinds provides data to geophysical researchers with previously unattainable coverage and scale. Point-wise wind estimation has been successfully used with SeaWinds data; however, the point-wise technique has limitations.

## 1.3 Thesis Overview

Field-wise wind estimation ([1],[2]) has been proposed as a method for ocean wind retrieval which avoids some of the limitations of point-wise estimation. While the comparative advantages of this method over the traditional point-wise estimation have been presented [3], practical usage of field-wise is limited by its computational difficulty. This thesis examines several of the complicating issues involved in field-wise wind estimation, such as wind field model order, optimization routines, and ambiguity selection. A practical implementation is presented for SeaWinds data.

In Chapter 2, a brief summary of scatterometry and wind estimation practices are presented. The instrument SeaWinds is introduced, along with its unique characteristics. An empirical relationship between radar backscatter and near-surface wind vectors is presented. Point-wise wind retrieval, the traditional estimation method, is summarized. The concept of field-wise wind retrieval is also introduced, along with the Karhuenen-Loeve wind field model.

Chapter 3 examines the field-wise objective function, integral in obtaining wind field estimates. An algorithm to find the local minima of the objective function is presented. Chapter 3 also discusses the measures necessary to insure that sufficient estimates are located to remove ambiguity from the wind field.

In Chapter 4, the details of ambiguity selection for field-wise wind retrieval are presented. The chapter discusses combining separate wind field region estimates into a more comprehensive swath estimate. Chapter 4 also discusses the frequency resolution of field-wise wind estimates.

The noise model used in SeaWinds wind estimation is examined in Chapter 5. The traditional model is verified using point-wise SeaWinds estimates. A parameterization of the measurement bias variance is presented based upon SeaWinds measurements and wind estimates. The Cramer-Rao bound on estimator performance is presented for both point-wise and field-wise wind retrieval.

# Chapter 2

# Background

A brief summary of scatterometry and wind estimation is prerequisite to any discussion of field-wise wind retrieval. First, an introduction to general scatterometry is presented. Next, the relationship between backscatter measurements and the near-surface ocean winds is discussed. The estimation problem is then defined, including the statistical models for noise. Finally, several standard ambiguity selection methods are described.

## 2.1 Scatterometry

Any instrument designed to measure the scattering coefficient (of a surface or a volume) is known as a scatterometer [4]. The surface scattering coefficient can be simply understood as a surface roughness measurement, specifically the roughness on the same order as the electromagnetic wavelength. The standard measurement is the normalized radar cross section, denoted $\sigma^o$, which is obtained by the radar equation

$$\sigma^o = \frac{(4\pi)^3 R^4 L P_r}{P_t G^2 \lambda^2 A_{eff}}, \tag{2.1}$$

where $R$ is the range to the surface, $L$ is the system loss, $G$ is the gain of the antenna, $\lambda$ is the wavelength and $A_{eff}$ is the effective area of the antenna. The radar crossection is essentially a normalized ratio of the power received to the power transmitted. As rougher surfaces will scatter more power, $\sigma^o$ is proportional to surface roughness.

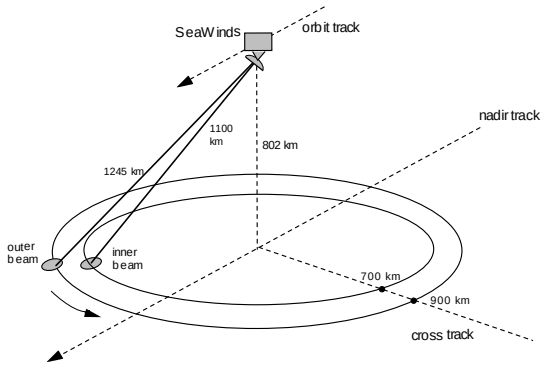


Figure 2.1: *The spacecraft and antenna geometries for SeaWinds on Quickscat.*

### 2.1.1 SeaWinds on QuikSCAT

In June of 1999, NASA launched the Ku-band scanning pencil-beam scatterometer SeaWinds. Launched on the platform QuikSCAT, SeaWinds is unique from previous scatterometers in that it employs a scanning pencil-beam as opposed to the previous fixed, fan beam antennas.

The scanning nature of the SeaWinds antenna has several effects. Unlike the NASA Scatterometer (NSCAT) and others, there is no nadir gap so that SeaWinds has a continuous swath 1800 km wide.

While SeaWinds has a wider swath, the improvement comes with a performance tradeoff. NSCAT and other scatterometers used fixed beams, and thus the variation in measurements from cell to cell was small. SeaWinds employs two scanning beams, both at Ku-band. On the swath edges, measurements are recorded only for the outer beam (see Fig. 2.2). This limitation reduces wind estimation effectiveness. Another tradeoff is experienced in examining the nadir region. While SeaWinds

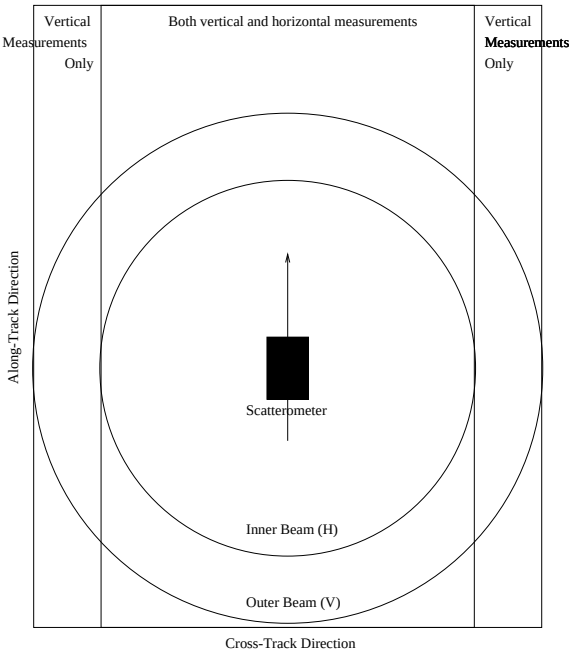


Figure 2.2: *A top view of the SeaWinds swath showing the two overlapping swaths from the inner and outer beam.*

records backscatter from the region, there is limited azimuthal variation between the measurements. As will be seen later in this chapter, azimuthal variation is necessary for wind estimation.

Due to its azimuth sampling, SeaWinds' performance varies greatly as a function of swath position. While performance can be good on either side of the swath center (known as the "sweet spot"), performance along the swath edges and in the nadir regions is significantly degraded.

SeaWinds is designed to generate a wind vector estimate for a grid of ocean cells. These grid points, designated wind vector cells (wvc's) are squares of side 25 km. Several backscatter measurements are observed in each cell, and in the estimation they are all assumed to be colocated (i.e. there is assumed to be no spatial variance within a wind vector cell).

## 2.2 The Geophysical Model Function

The ocean surface contains variations at many frequencies, with wavelengths varying from hundreds of meters (or even kilometers) to centimeters and fractions of centimeters. Microwave scatterometers observe roughness on the low end of this spectrum, with wavelengths on the order of two centimeters. Near-surface winds are the primary source for ocean surface roughness on the microwave wavelength. Thus, observed $\sigma^o$ measurements can be used to estimate the near-surface winds.

Accurate wind estimates from scatterometry require a known relationship between the observed $\sigma^o$ measurements and the wind vectors. Several theoretical scattering models have been explored, but a complete model has yet to be theoretically derived [5]. Instead, empirical data has been tabulated to form the Geophysical Model Function (GMF), represented by

$$\sigma^o = \mathcal{M}(S, \chi, \theta) \tag{2.2}$$

where $S$ is the near-surface wind speed, $\chi$ the wind direction relative to the instrument azimuth angle ($\psi - \phi$, where $\psi$ is the instrument azimuth angle and $\phi$ is the wind direction), and $\theta$ the instrument incidence angle.

The GMF variation with respect to both wind direction and wind speed introduces an important and complicating dimension to wind estimation. Figure 2.3 demonstrates the variation with respect to both $S$ and $\chi$. It is important to note that $\sigma^o$ increases with $S$ and varies in a nearly $\cos(2\chi)$ manner, with maxima at 180° (or downwind) and 0° (or upwind). The variation in $\chi$ is not exactly sinusoidal – there is an asymmetry referred to as the "upwind-downwind asymmetry" evident in larger $\sigma^o$ at 0° than at 180°. The variation in $\chi$ allows wind estimates to be a vector quantity, as both speed and direction can theoretically be resolved [5].

## 2.3 Estimating the Wind

The Geophysical Model Function is inherently non-invertible. Even with a fixed azimuth angle, the GMF maps the wind vector space of dimension 2 onto the backscatter space of dimension 1. Figure 2.4 shows the continuum of wind vectors

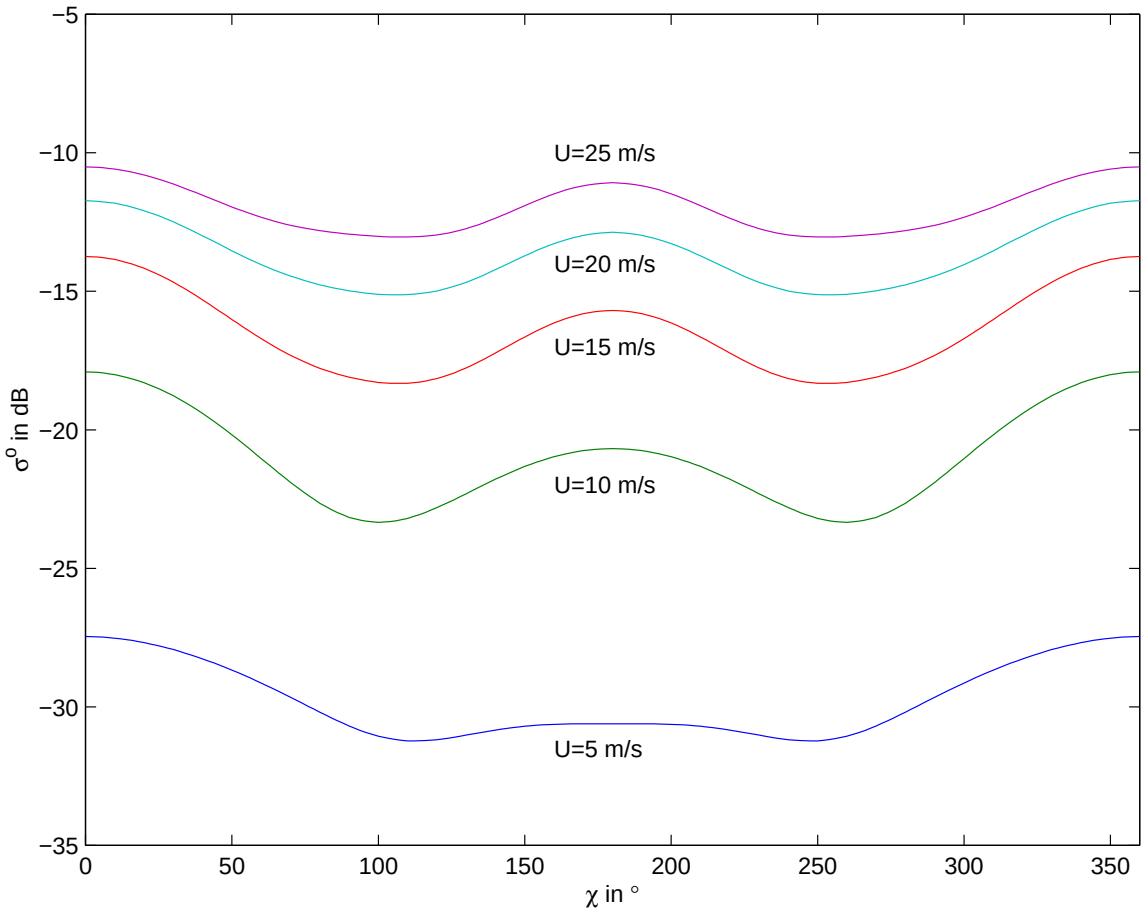


Figure 2.3: *The Geophysical Model Function evaluated with $\theta$ of 46°.*

that give rise to a single $\sigma^o$. For this reason, a wind estimate requires several $\sigma^o$ measurements.

In the noiseless case, several measurements can be used to generate a unique estimate. Figure 2.5, a simulation without noise, shows a clear intersection of each measurement at a single wind vector. Yet even without noise, several wind vectors provide "near-solutions." In the presence of noise, a unique solution cannot be determined; the "near-solutions" are indistinguishable from the correct estimate (see Fig. 2.6). Several wind vectors are classified as ambiguous estimates or aliases.

Noise is introduced into the estimation in two places. First, the GMF is not a perfect model for $\sigma^o$. While near-surface winds are the primary source of backscatter,

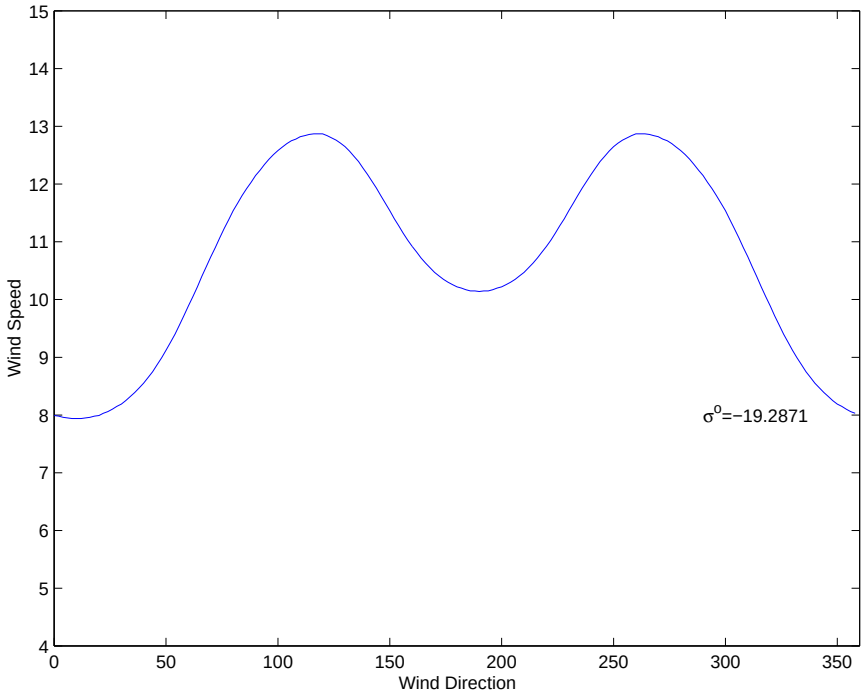


Figure 2.4: *A single backscatter measurement (with fixed geometries) is the range of a continuous domain of wind vector estimates projected through the GMF.*

the observed $\sigma^o$ is also a function of unknown and inestimable variables (i.e. water salinity, long waves, rain contamination, etc.). Thus, for a given wind vector the observed $\sigma^o$ will randomly vary from the predicted $\sigma^o$. This variation is classified as "modeling error" or "modeling noise" and is modeled as a Gaussian random variable with variance denoted $K_{pm}^2$ [6]. Thus the "true" backscatter is a random variable

$$\sigma^o = (1 + K_{pm} v_1) \mathcal{M}(\mathbf{w}), \tag{2.3}$$

where $v_1$ is a unit variance Gaussian random variable, and $\mathcal{M}(\mathbf{w})$ refers to the GMF evaluated at the wind vector $\mathbf{w}$. (Note that for convenience, the dependence upon the measurement geometries $\theta$ and $\psi$ have been omitted.)

Further noise is introduced by the scatterometer during measurement. The standard noise model [7] is that the observed measurement, $z$, is a realization of a

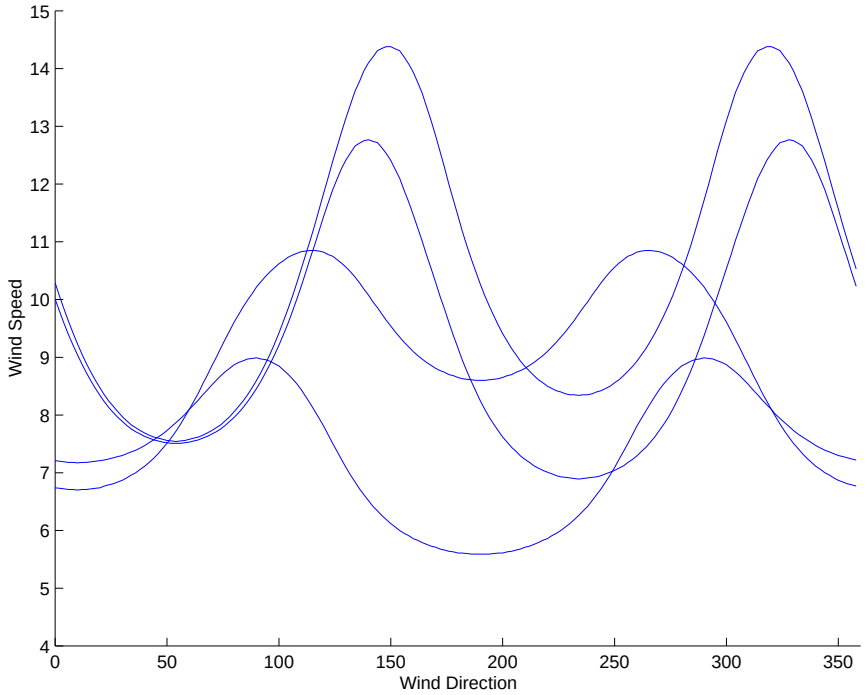


Figure 2.5: *To resolve a unique wind vector estimates, several measurements at varying geometries are required. Each curve represents the possible solution for a given noisy $\sigma^o$ measurement.*

random variable with mean $\sigma^o$:

$$z = \sigma^o + n, \tag{2.4}$$

where $n$ is a zero-mean Gaussian random variable. The variance of $n$ is given by

$$Var[n] = \alpha(\sigma^o)^2 + \beta\sigma^o + \gamma \tag{2.5}$$

where $\alpha$, $\beta$, and $\gamma$ are noise parameters which are dependent upon the instrument but independent of the wind vector [7]. Equation (2.4) is often expressed

$$z = (1 + K_{pc}v_2)\sigma^o, \tag{2.6}$$

where $v_2$ is a zero-mean, unit variance Gaussian random variable, and $K_{pc}$ is

$$K_{pc} = \sqrt{\alpha + \frac{\beta}{\sigma^o} + \frac{\gamma}{(\sigma^o)^2}}. \tag{2.7}$$

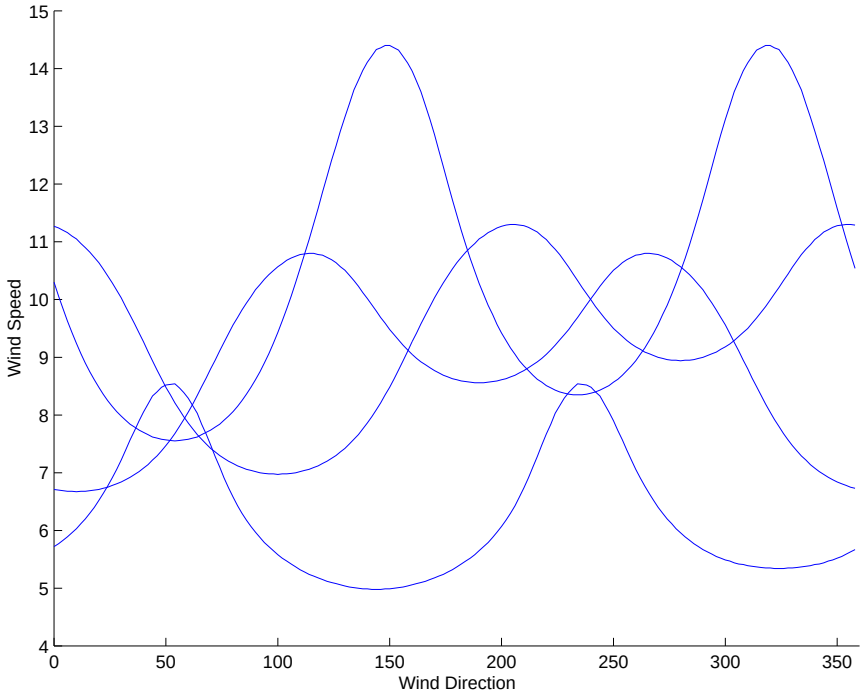


Figure 2.6: *In the presence of noise, no unique wind estimate can be determined.*

Combining both the modeling noise and the instrument noise into one model completes the statistical model for the estimation:

$$z = (1 + K_{pc}v_2)(1 + K_{pm}v_1)\mathcal{M}(\mathbf{w}). \tag{2.8}$$

Thus the observed measurement is a realization of the random variable $z$. While instrument noise ($K_{pc}$) is a relatively well understood quantity [5], modeling error ($K_{pm}$) is not [8]. Therefore, $K_{pm}$ is in general be assumed to be 0 in this thesis.

## 2.4 Point-wise Wind Retrieval

Point-wise wind estimation calculates a wind vector estimate for each wvc. In the estimation phase, the estimate at each wvc is calculated independently – there is no implicit correlation from cell to cell. A correlation is assumed only after the estimate has been performed to select between the ambiguous estimates. Point-wise wind retrieval is the primary method employed by the SeaWinds science team.

### 2.4.1 The Point-wise Objective Function

Under the preceding noise model, the observed measurements are a random variable parametrized by the true wind vector $\mathbf{w}$. Thus we may compute the probability density function

$$p(z|\mathbf{w}) = \frac{1}{\sqrt{2\pi\varsigma}} \exp\left\{-\frac{(z-\mathcal{M}(\mathbf{w}))^2}{2\varsigma}\right\} \tag{2.9}$$

$$\varsigma = \alpha(\mathcal{M}(\mathbf{w}))^2 + \beta\mathcal{M}(\mathbf{w}) + \gamma. \tag{2.10}$$

If we assume each measurement is independent, we may compute the joint distribution of the $K$ measurements at a given wind vector cell as the product of the individual pdf's:

$$p(\mathbf{z}|\mathbf{w}) = \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi\varsigma_k}} \exp\left\{-\frac{(z_k-\mathcal{M}_k(\mathbf{w}))^2}{2\varsigma_k}\right\}. \tag{2.11}$$

We may formulate a maximum likelihood estimate as the maximum of the conditional pdf given in Eq. (2.11):

$$\hat{\mathbf{w}} = \arg\max_{\mathbf{w}} p(\mathbf{z}|\mathbf{w}). \tag{2.12}$$

This arg max operation is equivalent to minimizing the negative log-likelihood function

$$\begin{aligned}
\hat{\mathbf{w}} &= \arg\max_{\mathbf{w}} p(\mathbf{z}|\mathbf{w}) \\
&= \arg\min_{\mathbf{w}} \{-\ln(p(\mathbf{z}|\mathbf{w}))\} \\
&= \arg\min_{\mathbf{w}} \sum_{k=1}^{K} \left\{\frac{1}{2}\ln(2\pi\varsigma_k) + \frac{(z-\mathcal{M}(\mathbf{w}))^2}{2\varsigma_k}\right\} \\
&= \arg\min_{\mathbf{w}} \sum_{k=1}^{K} \left\{\ln(2\pi\varsigma_k) + \frac{(z-\mathcal{M}(\mathbf{w}))^2}{\varsigma_k}\right\}.
\end{aligned} \tag{2.13}$$

Equation (2.13) can be interpreted as a cost function or objective function, the variance-normalized squared error between the observations and the forward projection of the wind vector. Calculating the objective function minimum is equivalent to maximizing the conditional pdf of Eq. (2.11). Equation (2.13) is referred to as the point-wise objective function.

As mentioned above, the shape of the GMF leaves several ambiguous estimates of the wind vector. For this reason, instead of simply choosing the global minimum of the objective function, each of the near-global local minima is calculated. The result is a set of ambiguous estimates, known as ambiguities or aliases.

### 2.4.2 Point-wise Median Filter

To generate an estimate for the entire swath, a single wind vector estimate must be chosen from among the ambiguities at each wind vector cell. Ambiguity selection allows a consistent swath to be chosen by assuming a correlation between wvc's. Although there are numerous potential selection algorithms, each is somewhat ad hoc and subject to error. The most common algorithm is called the point-wise median filter (PWMF). The PWMF assumes that adjacent wind vectors are highly correlated, and the difference between them should be small. Simply stated, the PWMF iteratively selects the ambiguity that minimizes the error between itself and its neighboring cells. The selected ambiguity at cell $(i,j)$ is

$$\hat{\mathbf{w}} = \arg\min_{k} \sum_{n=j-3}^{j+3} \sum_{m=i-3}^{i+3} ||\mathbf{w}_{mn}^{k} - U_{mn}||, \tag{2.14}$$

where $U_{mn}$ represents the wind vector at cell $(m,n)$ and $k$ is the index of ambiguities.

The median filter must be initialized, and a common first selection is the global minimum (or the first ambiguity) of the objective function for each wvc in the swath. The median filter is then successively implemented until there are no further changes (or a sufficiently small number of changes are reached).

Point-wise wind retrieval has several advantages. First, the objective function is relatively cheap both to calculate and to minimize, and thus the ambiguities are easily obtained. Second, in the "sweet spot" there is almost always an ambiguity that is relatively close to the true surface wind. Point-wise wind retrieval's largest shortcoming, however, is in the ambiguity selection step. As the PWMF is heavily biased toward its initialization, if a large number of the first ambiguities are incorrect the swath estimate can have significant error. For this reason, several "nudging" schemes initialize the point-wise field using weather information independent of the

scatterometer. Such an approach, while reasonably effective, potentially introduces errors inherent in the weather model. Point-wise estimation is also degraded in the nadir region and the swath edges where there is insufficient variation in the measurements to provide a reliable estimate.

## 2.5 Field-wise Wind Retrieval

Field-wise wind retrieval applies a correlation condition on the wind vector cells during the estimation process. By estimating an entire region of wvc's simultaneously, field-wise wind retrieval can simplify the ambiguity removal step problematic in point-wise wind retrieval. Furthermore, field-wise wind retrieval can compensate for cells with noisy, corrupt, or insufficient measurements. By using good data from neighboring wvc's, problem regions can be estimated with a higher degree of reliability. In this way, swath edge and nadir performance may be improved as compared to point-wise estimation.

### 2.5.1 The Wind Field Model

To describe the cell-wise correlation of the wind field, wind field models are employed. For this reason, field-wise estimation is also known as model based estimation. Several wind field models have been proposed, including a wind divergence, curl and boundary conditions model [1], a Legendre model, and a Karhuenen-Loeve (KL) model [9]. The KL model is used exclusively in this thesis.

We represent the wind in the region in the following way. Let $u$ and $v$ represent the rectangular components of a wind vector. Thus,

$$u = S\cos(\phi), \tag{2.15}$$

$$v = S\sin(\phi). \tag{2.16}$$

where $S$ is the wind speed and $\phi$ is the wind direction. Let $\mathbf{u}$ and $\mathbf{v}$ be the rectangular components of each cell in a region vectorized with cross-track direction varying more

rapidly. The vector $\mathbf{w}$ then represents the entire wind field:

$$\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}. \tag{2.17}$$

For a region with $M$ cross-track cells and $N$ along-track cells, $\mathbf{w}$ has dimension $2MN$. The purpose of a wind field model is to represent the wind field with fewer than $2MN$ dimensions. For any linear model, $\mathbf{w}$ can be represented in the matrix vector equation

$$\mathbf{w} = F\mathbf{x} \tag{2.18}$$

where $\mathbf{x}$ a vector of model parameters and $F$ is a "tall" matrix representing the transform.

The KL transform is chosen as a wind field model because it is known to minimize basis restriction error [10]. The KL transform matrix $F$ is obtained by the eigen decomposition of the autocorrelation matrix $R_{ww}$. The columns of $F$ are the eigenvectors associated with the largest eignenvalues of the autocorrelation matrix of $\mathbf{w}$ $R_{ww}$.

The autocorrelation matrix is estimated using point-wise retrieved winds to generate a large sample of $\mathbf{w}$ vectors, and a sample autocorrelation is computed

$$R_{ww} = E\left\{\mathbf{w}\mathbf{w}^T\right\} \tag{2.19}$$

$$\hat{R}_{ww} = \sum_{k=1}^{K} \frac{\mathbf{w}_k \mathbf{w}_k^T}{K}. \tag{2.20}$$

The first six eigenvectors of $R_{ww}$ are shown in Fig. 2.7. Note that the first two bases represent constant fields in the along-track and cross-track directions. These represent the fields with the highest amount of energy, and thus are the primary bases in the model. With bases ordered according to descending eigenvalues, the KL model can be shown to optimally maximize energy compaction [10]; thus truncating the KL model generates optimally small modeling error.

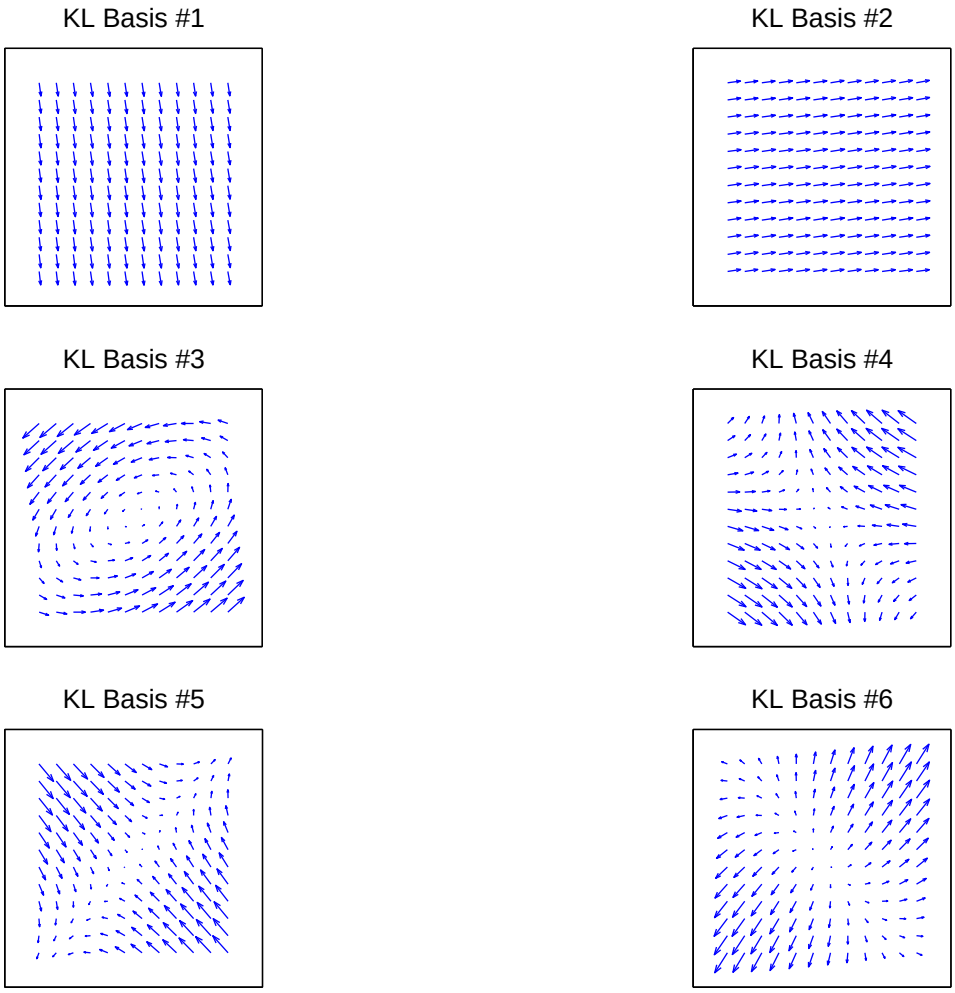


Figure 2.7: *The first 6 Karhuenen-Loeve basis fields.*

### 2.5.2 The Field-wise Objective Function

The field-wise estimation model is a direct extension of the point-wise estimator. Field-wise wind retrieval assumes that each measurement is an independent Gaussian random variable. Thus the field-wise objective function is given by

$$p_{\mathbf{z}}(\mathbf{z}|\mathbf{x}) = \prod_{i=1}^{N} \prod_{j=1}^{M} \prod_{k=1}^{K_{ij}} \frac{1}{\sqrt{2\pi\varsigma_{ijk}^2}} \exp\left\{ -\frac{1}{2} \left( \frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}} \right)^2 \right\}. \tag{2.21}$$

where $i$ and $j$ index the region by along-track and cross-track, respectively, $k$ indexes the measurement at a given cell, and $K_{ij}$ is the number of measurements at a given

cell. Note the field-wise dependence on $\mathbf{x}$ instead of $\mathbf{w}$. Recall that $\mathbf{w} = F\mathbf{x}$, so $\mathbf{x}$ completely parameterizes the wind field estimate. $\mathcal{M}_{ijk}(\mathbf{w})$ represents the GMF evaluated at the wind vector estimate of cell $(i, j)$ for measurement $k$. This notation can be simplified by converting from $\mathbf{w}$ dependance to $\mathbf{x}$ dependance and dropping the $F$ dependance

$$\mathcal{M}_{ijk}(\mathbf{w}) = \mathcal{M}_{ijk}(F\mathbf{x}) = \mathcal{M}_{ijk}(\mathbf{x}).$$

### 2.5.3 Field-wise Ambiguity Selection

As in point-wise wind retrieval, each objective function gives rise to a global solution as well as a near-global local solution. Thus instead of purely maximum likelihood estimation, each near-global solution must be found to generate a set of ambiguous wind field region estimates. As in the point-wise case, to generate a consistent swath estimate, a single ambiguity must be selected for each region.

Richards proposed a field-wise ambiguity selection algorithm for SeaWinds [3]. In his algorithm, the wind field is initialized to the global solution (first ambiguity) and then correlational constraints are iteratively imposed between regions.

In the Richards method, the SeaWinds swath is divided into overlapping regions (see Fig. 2.8), and wind field ambiguity sets are calculated at each region. When combining selected ambiguities at each region into a single swath estimate, swath portions with overlapping estimates are averaged, with weighting determined by location within the region. (As the wind field models more accurately the vectors in the region center, these vectors are given greater weight in the averaging process.)

The correlation constraint between regions is imposed by low-pass filtering the combined swath. At each region, the ambiguity closest to this averaged field is selected, and the process is repeated. The iterations continue until sufficiently few regions are changed; this swath estimate is the final field-wise product.

While developed before SeaWinds' launch, the Richards algorithm demonstrated relatively successful ambiguity selection for simulated data. Estimate performance in nadir and swath edges was improved from point-wise wind retrieval, although it was mildly degraded in the "sweet spot." The Richards algorithm's most

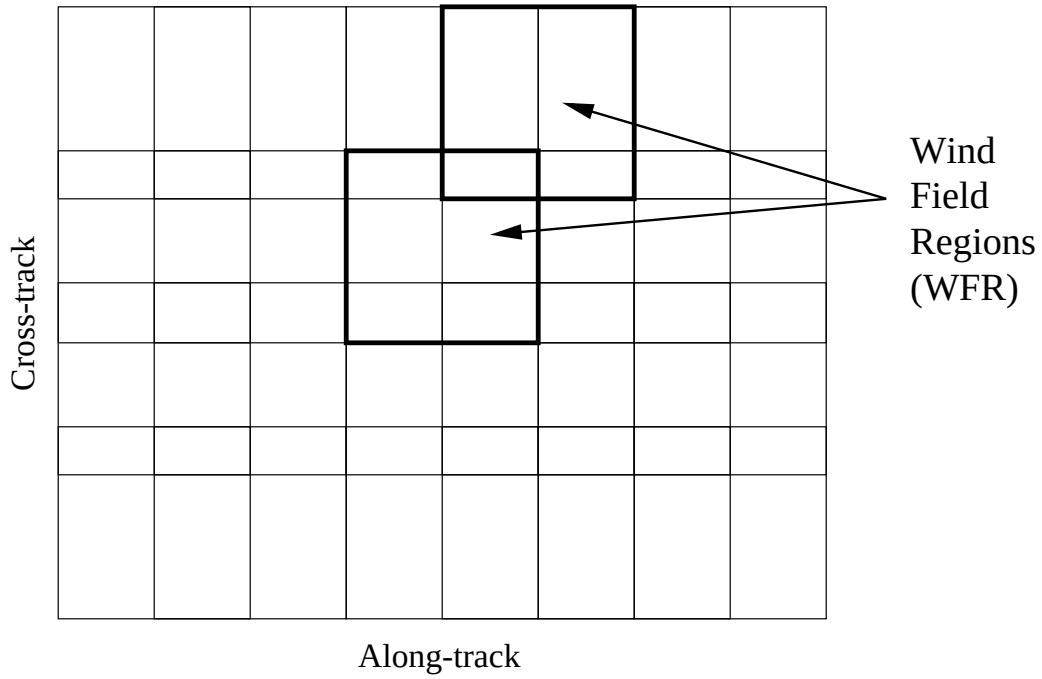


Figure 2.8: *The SeaWinds swath gridded into wind field regions as per the Richards method.*

significant drawback is the ambiguity selection stage. Although it enjoys good success, the field-wise ambiguity selection algorithm is highly dependent upon the initialization choice, with the final estimate heavily biased toward its initialization.

# Chapter 3

# Minimizing the Field-wise Objective Function

For field-wise wind retrieval, the objective function provides an error metric between an estimated wind field and the observed radar backscatter ($\sigma^o$). To generate the "best" estimate, the objective function is minimized, thus finding the wind field estimate "closest" to the observed $\sigma^o$ measurements.

Due to the nearly symmetric nature of the Geophysical Model Function (GMF), the objective function often has several local minima with objective function values very close to the global minimum. As with point-wise wind retrieval, it is important to locate all these near-global local minima. Thus several ambiguous solutions are retained and a consistent field is generated through the process of "ambiguity selection."

This chapter presents a detailed description of the field-wise objective function. A minimization algorithm is presented, as well as a methodology for locating the necessary local minima.

## 3.1 The Field-wise Objective Function

In Chapter 2, the field-wise objective function is derived as a maximum likelihood estimation problem. Alternatively, we may consider the objective function as a score function or error metric between a wind estimate and the observed measurements.

Given a field-wise wind estimate $\mathbf{x}$, we perform a forward projection using the GMF,

$$\hat{z}_{ijk} = \mathcal{M}_{ijk}(\mathbf{x}). \tag{3.1}$$

The quantity $\hat{z}_{ijk}$ is the backscatter value expected for the $k^{th}$ measurement of cell $(i,j)$ if $\mathbf{x}$ were the true value of the wind field model parameter vector. Thus the objective function can be composed of any error metric between the observed measurement $z$ and the forward projection $\mathcal{M}(\mathbf{x})$.

Perhaps the simplest and most common objective function uses the squared error metric. This metric can be described as the Euclidean distance or the $\mathcal{L}_2$ norm, which gives rise to the "least squares" solution. Thus, the objective function can be written as

$$J_{SE}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}(z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))^2, \tag{3.2}$$

where $K_{ij}$ is the number of measurements per cross-track cell, $ij$.

While simple, the squared error objective function fails to incorporate known information about the variance of each measurement. As presented in Chapter 2, each measurement has a variance estimated to be

$$\varsigma_{ijk}(\mathbf{x}) \;\; = \;\; \alpha(\mathcal{M}_{ijk}(\mathbf{x}))^2 + \beta\mathcal{M}_{ijk}(\mathbf{x}) + \gamma. \tag{3.3}$$

If each measurement is scaled by its variance, then the objective function becomes

$$J_{RML}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left(\frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}(\mathbf{x})}\right)^2. \tag{3.4}$$

This objective function is almost the same result that arises in maximum likelihood estimation. If we take the negative log-likelihood of Eq. (2.21), the result is

$$J_{ML}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left\{\left(\frac{Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}(\mathbf{x})}\right)^2 + \ln\varsigma_{ijk}^2(\mathbf{x})\right\}. \tag{3.5}$$

While it is not technically a maximum likelihood estimator, Eq. (3.4) is preferable to Eq. (3.5) due to the $\ln\varsigma_{ijk}^2(\mathbf{x})$ term.

We may understand the preference for Eq. (3.4) by examining the scale of each the two terms in the summation of Eq. (3.5). The first term is the square of a zero-mean, unit-variance Gaussian random variable, thus a $\chi^2$ random variable with one degree of freedom, which has mean value of 1. As $\varsigma^2$ is on the order of $10^{-5}$, $\ln\varsigma^2$ is on the order of -11.5, so summed over all of the measurements, the $\ln\varsigma^2$ term

dominates. The parameters $\alpha, \beta$, and $\gamma$ are only approximations which are nearly constant from measurement to measurement, so the dominant term is not nearly as relevant to the error metric as the smaller $\chi^2$.

Minimization of the field-wise objective function is greatly facilitated by further simplifying the objective function of Eq. (3.4). Most minimization routines (and all effective routines for the complicated field-wise minimization problem) require an evaluation of the first and sometimes the second partial derivatives of the objective function with respect to the estimate. The dependence of $\varsigma$ on $\mathbf{x}$ significantly complicates these partials.

To simplify the model, we set

$$\varsigma_{ijk} \quad = \quad \alpha z_{ijk}^2 + \beta z_{ijk} + \gamma. \tag{3.6}$$

In words, the variance of the random variable $z$ is estimated using the observed realization, or the measured $\sigma^o$ rather than the estimated parameter value.

## 3.2 The Levenberg-Marquardt Algorithm

Function minimization (or optimization) arises in many engineering contexts. In creating a minimization algorithm, the classic tradeoff is between fast local convergence and global convergence. On the two extremes are gradient descent, a slow stable method, and Newton's method, which shows rapid local convergence though it is often unstable. The Levenberg-Marquardt algorithm is a hybrid between gradient descent and Newton's method [11]. By combining the two, Levenberg-Marquardt incorporates the fast local convergence of Newton's method with the global stability of gradient descent. A brief summary of all three algorithms follows.

Gradient descent is a simple minimization technique. Since the gradient of a function points in the direction of steepest increase, the negative gradient points in the direction of steepest decrease. Thus an iterative minimization algorithm can be described by

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha J'(\mathbf{x}^k), \tag{3.7}$$

where $J'(\mathbf{x}^k)$ represents the function gradient evaluated at $\mathbf{x}^k$. By careful selection of the step size $\alpha$, gradient descent guarantees convergence to a local minimum [12]. The convergence, however, can be extremely slow, which necessitates the development of faster algorithms.

Newton's method is an example of a minimization algorithm with fast local convergence. That is, when close to a local minimum, Newton's method iterates to the local minimum in only a few iterations. The objective function near some point $\mathbf{x}_o$ can be represented by a Taylor expansion

$$J(\mathbf{x}) = c + (\mathbf{x} - \mathbf{x}_o)^T J'(\mathbf{x}_o) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_o)^T J''(\mathbf{x}_o)(\mathbf{x} - \mathbf{x}_o) + h.o.t. \tag{3.8}$$

where $J''$ represents the Hessian matrix (the matrix of second derivatives). If the function is exactly quadratic (i.e. the higher order terms are non-existent), then

$$J'(\mathbf{x}) = J'(\mathbf{x}_o) + J''(\mathbf{x}_o)(\mathbf{x} - \mathbf{x}_o), \tag{3.9}$$

and

$$\mathbf{x} = -(J''(\mathbf{x}_o))^{-1} J'(\mathbf{x}_o) + \mathbf{x}_o \tag{3.10}$$

is a critical point, where $J'(\mathbf{x}) = 0$. If the function is exactly quadratic, such a calculation locates the local minimum in a single iteration. Since many functions are approximately quadratic near a local minimum, Eq. (3.10) suggests an iterative method of successive Newton steps:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - (J''(\mathbf{x}^k))^{-1} J'(\mathbf{x}^k). \tag{3.11}$$

When close to a local minimum, such an algorithm converges very quickly as the function is well represented by its quadratic approximation. However, when far from a minimum point, the function may not be well represented by a quadratic, and often Newton's method does not converge at all.

The Levenberg-Marquardt algorithm combines gradient descent and Newton's method using a clever hybrid scheme. The algorithm is specifically designed for use with the Weighted Squared Error (WSE) objective function, i.e. a constant

variance $\chi^2$ objective function. Using the simplified variance of Eq. (3.6) the field-wise objective function fits this criterion, and is given by

$$J_{WSE}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left(\frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2. \tag{3.12}$$

The WSE Hessian is derived to be

$$\frac{\partial^2 J_{WSE}}{\partial \mathbf{x}_p \partial \mathbf{x}_q} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left[\frac{(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))}{\varsigma_{ijk}^2}\frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p \partial \mathbf{x}_q} - \frac{\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_q}}{\varsigma_{ijk}^2}\right]. \tag{3.13}$$

(The complete derivation for the derivatives of the field-wise objective function is provided in Appendix A.)

A first derivative approximation ($\hat{J}''$) is made to the Hessian matrix by keeping only the second term in the summation of Eq. (A.59),

$$\frac{\partial^2 \hat{J}_{WSE}}{\partial \mathbf{x}_p \partial \mathbf{x}_q} = 2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \frac{\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_q}}{\varsigma_{ijk}^2}. \tag{3.14}$$

The first term from Eq.(A.59) can be dropped by noting that the second derivative information is multiplied by the term $\frac{(Z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))}{\varsigma_{ijk}^2}$ which for a successful model is a zero mean, unit-variance random variable. This term tends to cancel out over the entire summation [11].

The value of making this approximation is two-fold. First, second derivatives of the model-function need not be calculated – a valuable savings in computation time. Second, each term along the diagonal of the Hessian approximation is non-negative. This second fact makes possible the Levenberg-Marquardt algorithm, as will become apparent.

Let $A = \hat{J}''(\mathbf{x}^k) + \lambda I$, the Hessian approximation with a constant value $\lambda$ added to the main diagonal. Let $\mathbf{b} = -J'(\mathbf{x}^k)$. A single step in the Levenberg-Marquardt algorithm is then

$$\mathbf{x}^{k+1} = \mathbf{x}^k + A^{-1}\mathbf{b}. \tag{3.15}$$

As $\lambda \to 0$, $A \to \hat{J}''(\mathbf{x}^k)$ and Eq. (3.15) is a full (approximate) Newton step. As $\lambda$ gets large, $A$ is diagonally dominant, and $A^{-1} \approx \frac{1}{\lambda} I$. In this case, Eq. (3.15) becomes Eq. (3.7) with $\alpha = \frac{1}{\lambda}$. Here we have used the fact that the main diagonal of $A$ is non-negative. By varying $\lambda$, Eq. (3.15) is either a gradient descent step or an (approximate) Newton step.

The Levenberg-Marquardt algorithm begins with an initial guess $\mathbf{x}^0$ and a small value of $\lambda$ (e.g. $\lambda = .001$). The algorithm is iterated by Eq. (3.15). If $J(\mathbf{x}^{k+1}) < J(\mathbf{x}^k)$, the step is accepted, and $\lambda$ is reduced, say $\lambda = .1\lambda$. If $J(\mathbf{x}^{k+1}) > J(\mathbf{x}^k)$, the step is rejected, and $\lambda$ is increased, say $\lambda = 10\lambda$. The algorithm is stopped when $J(\mathbf{x}^{k+1}) - J(\mathbf{x}^k)$ is less than the stopping criterion (0.01 for this case).

The choice of stopping criteria may seem odd. Theoretically, $J'(\mathbf{x}^k)$ should be effectively $\mathbf{0}$ for a local minimum. In practice, optimizing to nearly $\mathbf{0}$ wastes computation time with little gain. For some cases, however, the optimization finds several "minima" within a hyper-ball which are similar but not identical solutions. To compensate, after the Levenberg-Marquardt algorithm is completed, a full Newton-method is iterated using the complete Hessian matrix and stopping when the gradient is sufficiently close to $\mathbf{0}$. As the Newton-iteration presumably starts near a local minimum, it converges rapidly to the optimal solution. This second iteration is generally successful in "tightening" the solutions around a true local minimum with acceptable computational cost.

### 3.3 Parameter Choices

Despite a fast-converging algorithm, minimizing the field-wise objective function is highly intensive computationally and therefore time-consuming. The minimization cost greatly increases as a function of region size and number of model parameter. Thus, small regions with only a few model parameters are preferred computationally. With such a preference comes a significant tradeoff: larger regions provide a greater cell-wise correlation, and more model parameters provide higher frequency resolution.

To understand the tradeoffs involved in KL model truncation and wind field region size, Draper [13] analyzed the frequency content of the KL basis fields. Using

the Discrete Fourier Transform (DFT), Draper analyzed the natural truncation points within the KL model.

Table 3.1: *Comparison of frequency bin groupings between different models. The last model parameter that used only the first N bins is listed.*

| Nth Bin | SeaWinds 12x12 | SeaWinds 12x12 | SeaWinds 16x16 | SeaWinds 24x24 |
|---|---|---|---|---|
| 1 | 2 | 2 | 2 | 2 |
| 2 | 6 | 6 | 6 | 6 |
| 3 | 12 | 12 | 12 | 12 |
| 4 | - | 26 | 26 | 26 |
| 5 | - | 36 | 36 | 38 |
| 6 | - | 50 | 53 | 57 |
| 7 | - | 65 | 70 | 78 |

To investigate the natural truncation points, Draper examined the number of bases necessary to represent each DFT bin. He found an orderly representation, i.e. the first two bases represented the first bin, and the next four bases the second bin (see Table 3.1). In this way, we may see that truncating the KL model is in many ways equivalent to imposing a low frequency constraint on the retrieved wind fields.

The natural truncation points are also observable in the eigenvalues of the wind autocorrelation. For the 12×12 model, Fig. 3.1 shows the eigenvalues in descending order. Note that sharp drop-offs in eigenvalues correspond with those observed in Table 3.1.

Previous implementations of field-wise wind retrieval have used 24×24 sized regions with 22 model parameters [3]. Such a large region and high truncation point makes objective function minimization prohibitive, and for this reason, a smaller wind field region is chosen. The wind field region size is chosen for the current implementation to be 12×12. Truncating the 12×12 at 6 model parameters is equivalent in frequency content to 22 model parameters for a 24×24 region.

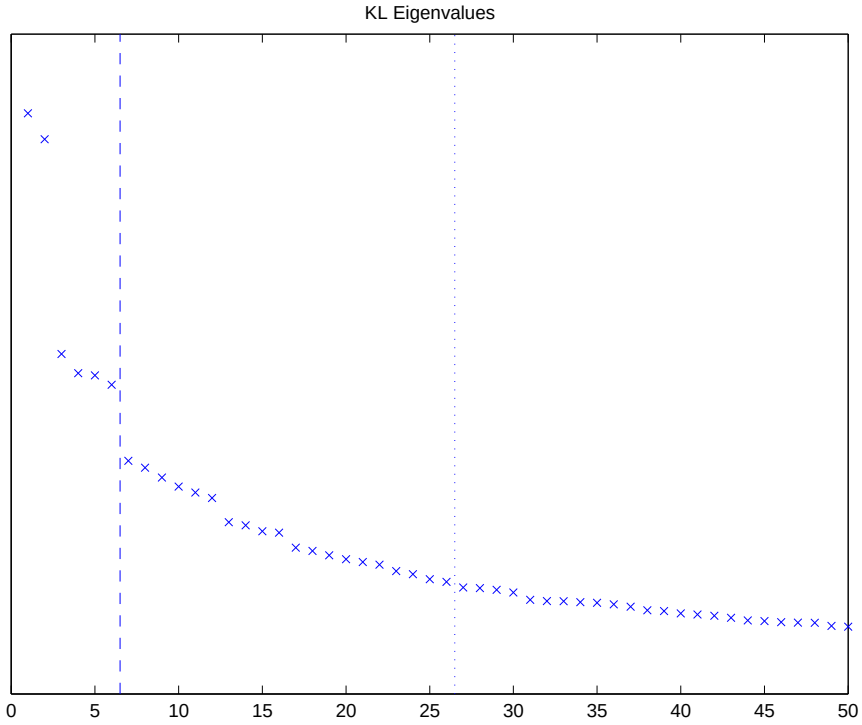


Figure 3.1: *The eigenvalues of the wind autocorrelation matrix. Note the discrete drop-offs after the* $2^{nd}, 6^{th}, 12^{th}, 16^{th}$, *and* $26^{th}$ *eigenvalues. Dotted lines after 6 and 26 parameters represent chosen cutoffs for wind estimation.*

### 3.4 Objective Function Minima

As with point-wise wind retrieval, minimization of the field-wise objective function does not produce a unique wind field estimate. The shape of the GMF causes several local minima of similar likelihood. These can be considered ambiguous estimates. Instead of generating a unique estimate for each region, a set of ambiguities is estimated. A separate process, known as "ambiguity selection" chooses one ambiguity for each region to create a consistent swath estimate.

Generating a solution set requires minimizations from multiple initial vectors; the initial vectors form the initialization set. Optimization is performed starting at each vector in the initialization set, and the resulting minima make up the solution set.

The objective function may have many local minima, some near-global and some not. While it is not necessary to include all local minima in the solution set, selection of a consistent wind field generally requires identification of most of the near-global minima. (Recall that local minima with values much greater than the global minimum are wind field estimates that are much less likely, based on the observed measurements and thus less important.) Ideally, only the "true" solution needs to be located, but it is not difficult to identify these. Therefore the optimization needs to assure that the best ambiguity is within the solution set.

### 3.4.1 Random Initialization Set

In the past, the initialization set was randomly generated [3], [14]. The idea is to select a large number of initialization values to span the wind field space. In this approach, the standard deviation ($\sigma$) of each model parameter is calculated, and each vector element is a realization of a uniform random variable over the range $[-2\sigma, 2\sigma]$. Thus, each vector in the initialization set is randomly generated, with independent elements. Random generation requires a large initialization set to have confidence in finding the near-global minima.

### 3.4.2 Deterministic Initialization Set

Alternatively, a deterministic initialization set can be chosen to best span the wind field space. A deterministic set can utilize the correlation between the model parameters, and assures that each subspace of the reasonable solution space is being searched for local minima.

The initialization set needs to be representative of the range of observed wind fields. To generate the set, three assumptions are employed. First, the initialization set need only span the common solution space. There is little reason to search the wind field space in where the regions are known to be highly unlikely. Secondly, we may assume that initial vectors that are very close together optimize to the same local minimum.

The third assumption follows from observing the comparable wind speeds of field-wise and point-wise ambiguities. Ambiguous estimates often agree on the wind speed but vary as to the estimated wind direction. Good wind speed estimates are much simpler to obtain than accurate representations of the wind direction. The point-wise ambiguity set provides a reliable estimate of the wind speed, which may be used as *a priori* information when generating a field-wise estimate. The *a priori* information can be used to reduce the dimension of the wind field space to be searched. The initial vectors are all scaled to match the estimated region speed. Therefore, the initialization set need only be composed of unit-norm vectors.

A sample set is obtained by creating a least squares model fit to point-wise wind fields in over 50,000 regions of actual SeaWinds data. To employ the first assumption, all vectors with parameters outside the range of $[-2\sigma, 2\sigma]$ are excluded. All of the remaining vectors in the sample set are normalized to 1 m/s Root Mean Square (RMS) wind speed. (Prior to optimization, each vector is multiplied by the average point-wise wind speed for the region.) Next, all of the redundant solutions (those within 0.125 m/s RMS difference) are eliminated. These three steps reduce the sample space set to 165 (normalized) vectors.

However, 165 optimizations are still too costly to be computed for each region; the initialization set must be a reduction of this sample. If we assume that all relevant minima are located by optimizing these 165 vectors, any further reduction of the initialization set may come at the cost of missing local minima. To quantify this tradeoff, 1422 randomly selected regions are optimized using all 165 vectors as the initialization set. By examining the resulting minima and the initial vectors that "find" them, an initialization set may be chosen to locate the most important minima with an acceptable computational load.

On average, each region contains ten local minima. When examining the objective function values for each solution, there is always at least one clear break in the objective function values, as shown in Fig. 3.2. In this case, there are four minima that we may classify as near-global or "pre-break" and five minima that we may classify as "post-break." "Post-break" solutions are generally less realistic, having

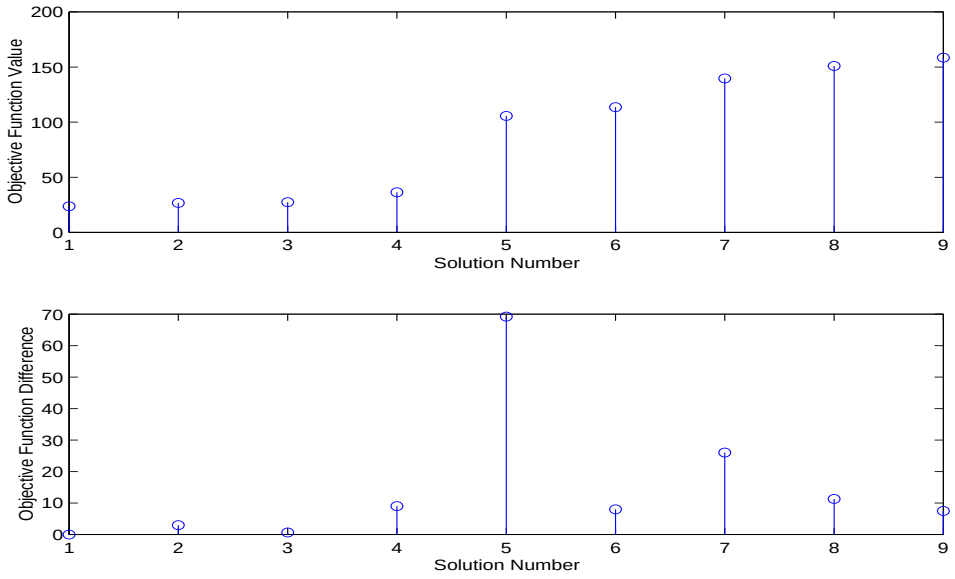


Figure 3.2: *Top: Typical WSE objective function values for a region. Bottom: Difference between objective values of consecutive solutions. One obvious breakpoint occurs after solution 4.*

objective function values significantly higher than the global minimum. Thus they are estimates with significantly greater disparity from the observed measurements. While legitimate local minima of the field-wise objective function, they are less likely to contain the correct wind field. Therefore, in determining an initialization set, only the "pre-break" solutions are considered essential.

The initialization set is determined from the sample set of 165 vectors in an iterative manner. For each vector in the sample set, the percentage of regions in which the vector optimized to a "pre-break" solution is calculated. The vector with the highest percentage is placed in the initialization set. For the remaining vectors, we calculate the percentage of regions in which the vector optimizes to a "pre-break" solution not previously "found." Again, the vector with the highest percentage is added to the initialization set. This process is repeated indefinitely until an arbitrary percentage of "pre-break" solutions is obtained.

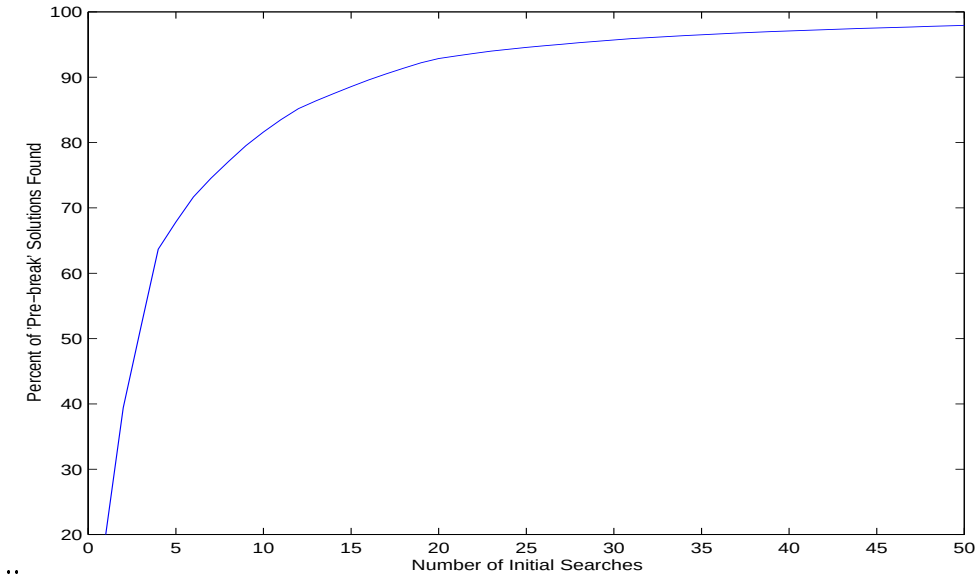


Figure 3.3: *The percentage of "pre-break" solutions against the number of initial searches.*

Figure 3.3 demonstrates the tradeoff between initialization set size and near-global minima "found." Increasing the size of the initialization set yields ever diminishing returns. While failing to locate a near-global minimum can have significant consequences in the ambiguity selection process, the computational cost of an optimization forces this possibility. Requiring 94% of "pre-break" solutions, necessitates using only 23 initial vectors in the initialization, which results in acceptable computation times.

Concluding the discussion on the deterministic initialization set, it is important to note the limited use of "post-break" and "pre-break" distinctions. When estimating the wind, all local minima are retained for the solution set, regardless of their objective function. "Pre-break" solutions are only considered when selecting the initialization set. Thus, there is no objective function threshold for an ambiguity.

## 3.5 Results

Scatterometric wind estimation is the only source for near-surface wind estimates on a global scale. Thus, wind retrieval is performed completely in the absence of "truth" data. Without independent data to verify results, analysis of wind retrieval is limited to either simulated wind retrieval or anecdotal evidence. Without truth data, wind fields and swath estimates are judged on their believability, self-consistency, and simply what appears to be correct. For this reason, analysis of field-wise ambiguities can be performed mostly by comparison with point-wise wind retrieval.

Compared with point-wise retrieved regions, field-wise ambiguities have several dominant characteristics. The truncated KL model imposes a low-frequency criterion on the wind estimate. This minimizes the noise inherent in the estimate, though if the wind contains higher frequency features these are also suppressed.

One goal of field-wise wind retrieval is the simplification of the ambiguity selection step of wind estimation. While ambiguity removal is still necessary for field-wise wind retrieval and will be discussed in the next chapter, the dimension of field-wise ambiguity removal is greatly reduced from point-wise. For a given $12 \times 12$ region, field-wise estimation generates between two and ten ambiguities. Each ambiguity represents a relatively reasonable wind field. In comparison, at each of the 144 wind vector cells in the region, point-wise wind retrieval generates between two and four ambiguities. Thus there are often thousands of combinations of ambiguities, many of which are unreasonable wind fields.

As discussed in Chapter 2, the nadir region has minimal azimuthal variation, which limits the effectiveness of wind-retrieval. Minimal azimuthal sampling creates a situation analogous to an ill-conditioned matrix. Small perturbations in the backscatter measurements can significantly modify wind estimates. The point-wise result in the nadir region is often a noisy field. In simulations (see [3]), field-wise wind retrieval improves estimate performance in the nadir region and on the swath edges. Figure 3.4 shows typical results from the nadir region for both the point-wise and field-wise products. Note that the field-wise ambiguities are lower frequency, which appears to remove much of the noise.

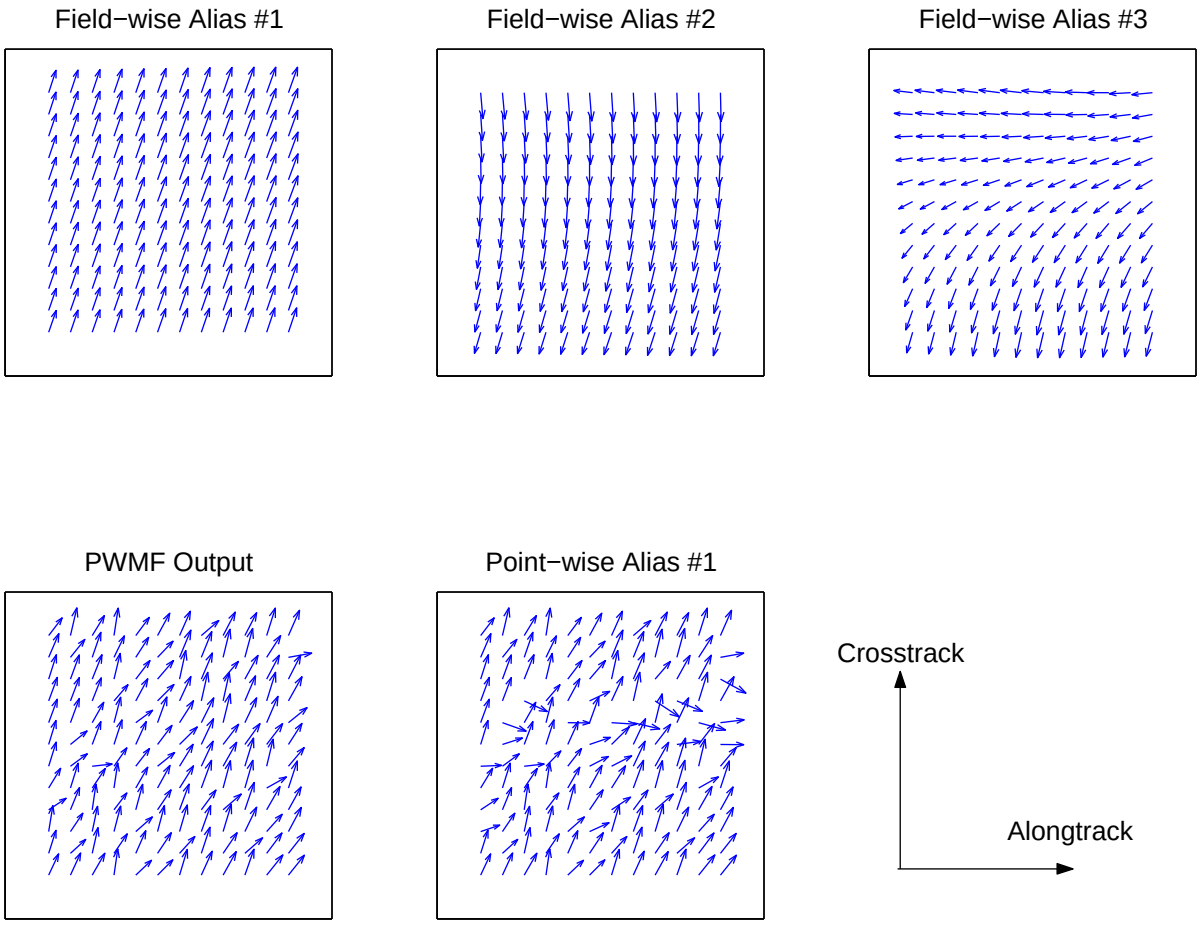


Figure 3.4: *Point-wise and field-wise products for QuikSCAT Rev 01289 beginning in the upper left corner at alongtrack 333 and crosstrack 35.*

Aside from the nadir region and the swath edges, field-wise wind retrieval performance is observably better in the presence of rain corruption. Rain affects surface roughness, thus adding an element that is difficult to remove for wind retrieval. The point-wise product is generally characterized by high speed vectors close to perpendicular with the satellite track. Figure 3.5 shows a point-wise region with an apparent rain artifact. The field-wise product contains no obvious artifact while maintaining several believable wind fields. In this case, the second ambiguity is quite close to the general wind flow shown in the point-wise product, but it shows no rain contamination.

Aside from the increased computation, the largest drawback of field-wise wind estimation is its limited modeling of wind features. With a model order low enough to permit practical computation, the frequency constraint limits the ability to model fronts and storms, permitting only estimation of general wind flow. Figure 3.6 is an example of a potentially troubling region. Although the front observed in the point-wise product may not be accurate, field-wise wind retrieval succeeds only in modeling the general flow for the region. To model the high frequency front requires a higher order model, which is computationally costly. This limitation is discussed further in Chapter 4 and a potential remedy is proposed.

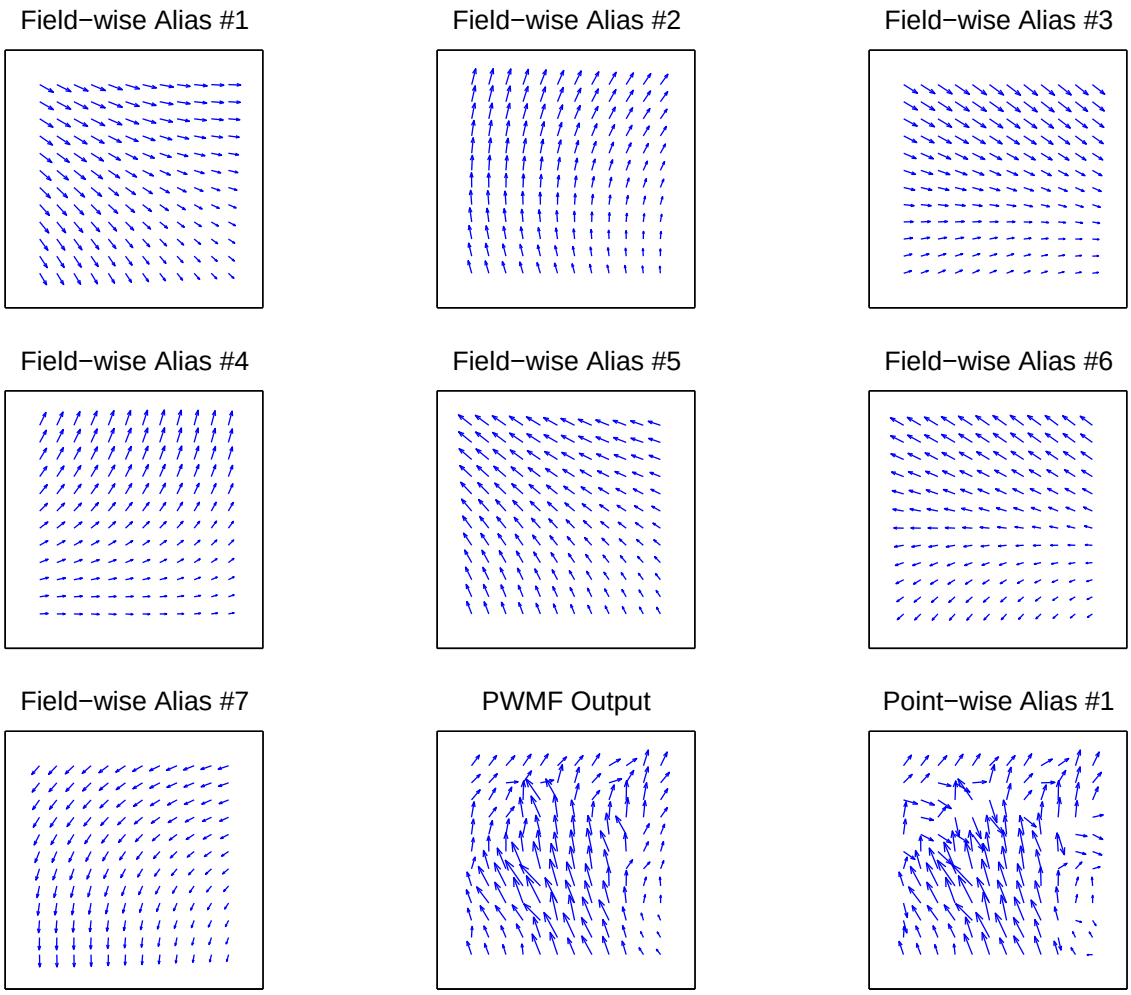


Figure 3.5: *Point-wise and field-wise products for QuikSCAT Rev 01289 beginning at alongtrack 369 and crosstrack 35. The high speed vectors in the point-wise product are characteristic of rain contamination. Note that the field-wise does not demonstrate the same artifact, while the second field-wise alias maintains a similar wind flow to the PWMF product.*

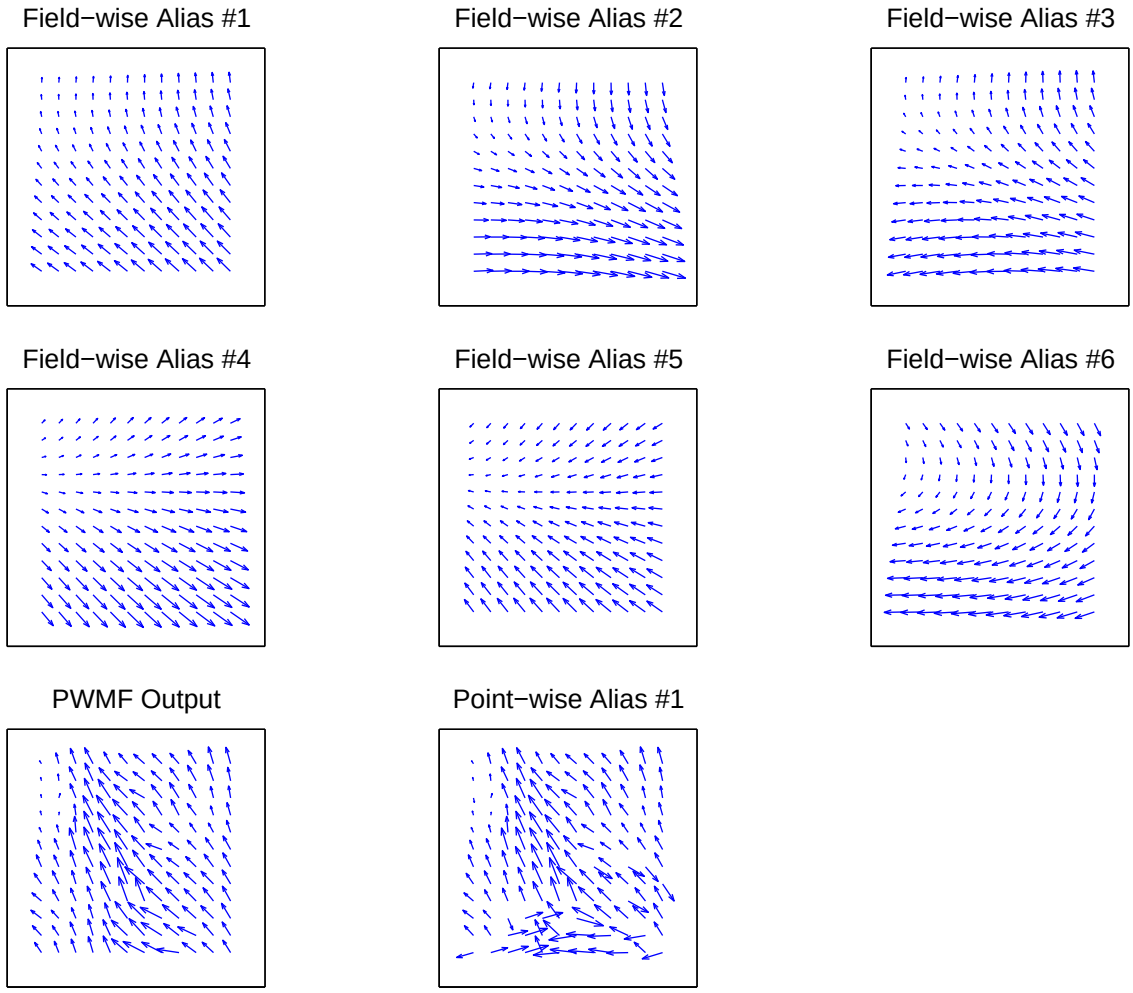


Figure 3.6: *Point-wise and field-wise products for QuikSCAT Rev 03079 beginning at alongtrack 457 and crosstrack 59. The wind feature shown in the point-wise product is essentially eliminated in the field-wise retrieval (Field-wise Alias #1). While not necessarily evidence of field-wise retrieval failure, in the absence of truth data the inability to model features is a troubling result.*

# Chapter 4

# Field-wise Ambiguity Selection

In scatterometer wind vector estimation, a necessary step is ambiguity selection. Due to the symmetry inherent in the GMF and the sufficiently high noise level in the measurements, several wind fields appear nearly equal in their likelihood. For this reason, after the estimation stage in the wind retrieval process, a set of ambiguous solutions is generated and the "correct" estimate must be selected from this set.

Several ambiguity selection algorithms have been proposed for both point-wise and field-wise wind retrieval, but as yet no algorithm has demonstrated any degree of optimality. Ambiguity selection is very difficult to analyze statistically, and as a result, each algorithm can only be classified as *ad hoc*. In the absence of truth data (which is only available for simulated data and introduces its own complications), ambiguity selection performance is often characterized simply by what subjectively "looks good" or appears to give consistent results.

The algorithm implemented in this thesis is based largely on the Richards algorithm [3], with several significant modifications. The modifications are detailed in this chapter, and are such that the revised algorithm is titled the Field-wise Median Filter (FWMF).

## 4.1 Median Filtering

A median filter is an image processing technique used to improve image quality while maintaining edge features. While spatial averaging of an image can be shown to be nearly optimal in its statistical performance (see [10]), an averaging filter generally

blurs an image by smoothing edges. While statistically optimal, edges are essential features in images; thus median filters are preferred.

A median filter performs no average, but instead chooses a median value from within a window around an image pixel. The median filtered image $v$ is computed from the image $y$ at pixel $(m, n)$ by [10]

$$v(m,n) = median\{y(m-k, n-l), (k,l) \in W\} \tag{4.1}$$

where $W$ is a chosen window. The effect of the median filter is to reduce image noise while maintaining image edges. The filter is especially adept when in the presence of binary noise, i.e. an isolated pixel is highly variant from its neighboring pixels.

## 4.2 The Point-wise Median Filter

Before considering the field-wise median filter, it is valuable to re-examine the point-wise median filter (PWMF). The PWMF is introduced in Chapter 2, and is the most common point-wise ambiguity selection method.

The PWMF is based on the image processing principles described above, but is a bit of a misnomer. Instead of choosing among the median of the surrounding values, the surrounding values suggest a choice from within a discrete set of estimates (the ambiguity set). The PWMF is not a spatial median filter as described in Section 4.1, but due to its discrete nature is of the same class. The distinction is important, for the resulting selected field is composed entirely of ambiguities that are local minima of the objective function. Thus each wind vector estimate locally minimizes the measurement errors.

Recall that the spatial median filter of Section 4.1 maintains image edges while removing binary noise. The PWMF has similar performance, and as such is a reasonable algorithm choice. Edge preservation is highly desirable in wind estimation since features such as fronts and storms are of great interest. Furthermore, ambiguity selection errors are somewhat analogous to binary noise in images. Consider Fig. 4.1, which shows the first point-wise ambiguity at each wind vector cell. While the overall wind flow is apparent (a majority of the vectors point to the right), several isolated

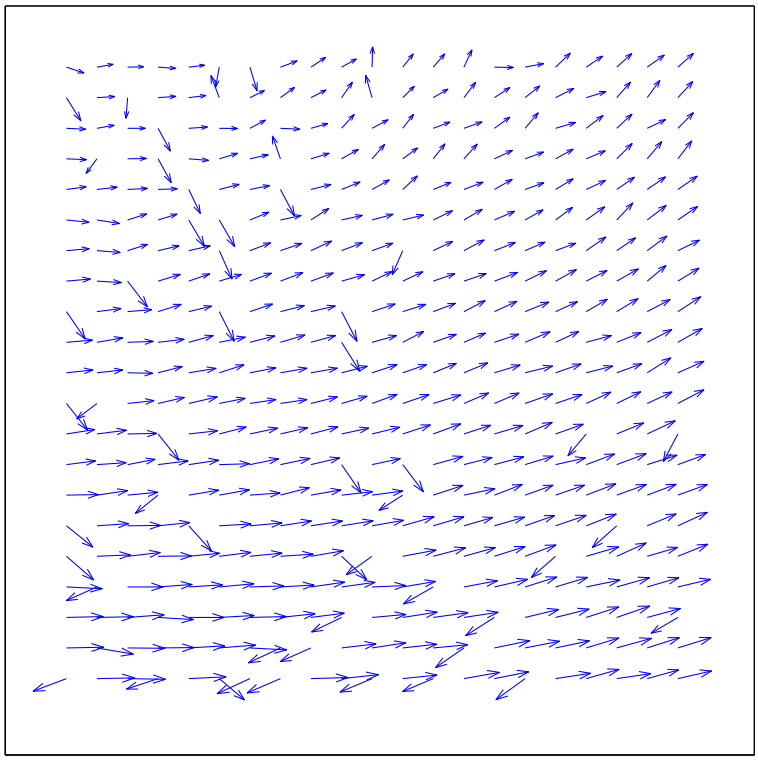

Figure 4.1: *Typical point-wise wind field with each wind vector set to the first ambiguity. Note that the field contains several obvious ambiguity selection errors – individual wind vectors at odds with the neighboring wind vectors.*

wind vectors point in entirely different directions. The PWMF proves to be generally proficient at removing such obvious selection errors.

As described in Chapter 2, the PWMF chooses the ambiguity that minimizes the error with the surrounding cells. Thus the point-wise median filter is given by

$$\begin{aligned} \hat{\mathbf{w}} &= \arg\min_{k} \sum_{n=j-l}^{j+l} \sum_{m=i-3}^{i+3} a_{m'n'} ||\mathbf{w}_{ij}^{k} - U_{mn}||, \qquad (4.2) \\ m' &= m - i, \\ n' &= n - i, \end{aligned}$$

where $a_{m'n'}$ is a weight applied to each neighboring wvc and the window size is $(2l+1) \times (2l+1)$. A standard window size is $7 \times 7$.

The PWMF is relatively cheap computationally, a distinct advantage of the method. It is also relatively successful at selecting a self-consistent swath estimate. The PWMF is generally successful in removing binary noise, but does not reduce the noise inherent in the estimates. In other words, even if the PWMF successfully selects the wind vector closest to the "true" wind, the ambiguity selection does not reduce the error between the closest ambiguity and the true wind. This is, in fact, a characteristic of all ambiguity selection schemes, not merely the PWMF. The process of ambiguity selection is a choice from among a discrete set of solutions, and thus any noise reduction is inherently limited to discrete noise by the parameters of ambiguity selection.

The other major drawback of the PWMF is its bias toward initial conditions. The final selected swath is highly dependent on the initial field. For highly inconsistent initial fields, the final result from the PWMF often retains the initial inconsistency. More specifically, the PWMF corrects individual ambiguity selections, but generally does not alter the overall wind flow from the initalized field. Thus if the initial wind flow in a region is incorrect, the PWMF may retain the error. To improve performance, a nudging field is often used; the swath is initialized to the ambiguity closest to an independent weather model. While nudging generally improves perceived performance, dependence on outside data is problematic and non-ideal. Errors in the nudging field also can introduce inconsistencies. The user community would prefer that nudging not be used. Nevertheless, nudging is effective in improving ambiguity selection.

### 4.3 The Field-wise Median Filter

As in point-wise estimation, field-wise wind retrieval requires an ambiguity selection step. Each wind field region has several ambiguous estimates. To generate a swath estimate, one estimate must be selected for each region. This section presents a method for field-wise ambiguity selection, entitled the Field-wise Median Filter.

In point-wise wind estimation, no cell-wise correlation is assumed during the estimation phase. Correlation is applied only during the ambiguity selection phase

when the PWMF applies an *ad hoc* band limitation. The error metric from Eq. (4.2) gives preference to the alias that is most similar to its neighbors, thus favoring low frequencies.

Such an approach makes little sense for field-wise ambiguity selection. Field-wise ambiguities are already band-limited during the estimation process. By using a truncated wind field model, field-wise wind estimates are limited only to those frequencies contained in the model bases. For this reason, field-wise ambiguity selection must use an alternative consistency criterion.

The main criterion for ambiguity selection is overall swath consistency. For point-wise, this is achieved by choosing neighboring estimates that vary as little as possible with their neighboring estimates. For field-wise, wind field regions are colocated; that is wind field estimates are chosen to overlap by several wvc's. The overlapping sections of the wind field estimate can be used to quantify the variation between neighboring wind field estimates. If neighboring ambiguities have very similar estimates for the overlapping wvc's, then they may be considered "close" or consistent estimates.

The field-wise median filter follows a principle similar to the point-wise median filter. An initial field is chosen, and each ambiguity at each wind field region is compared with the neighboring selections. The ambiguity "closest" to its neighbors is selected. Thus, the selected ambiguity at a region is a function of its neighboring regions. Ambiguities are selected from a discrete set. By following similar methods to the PWMF, the field-wise median filter also has similar results. Specifically, the field-wise median filter maintains image edges and removes binary errors.

### 4.3.1 Combining Overlapping Regions

Overlapping regions help to make ambiguity selection more feasible, but generate a separate issue – how to combine the selected estimates into a single swath estimate. A simple average between the coincident estimates is a straightforward solution, but a weighted average is preferable. To recognize the preference of weighted averaging, note that the KL modeling accuracy is a function of location within the

region. This can be demonstrated by computing the average error in computing a model fit.

A model fit to the region vector $\mathbf{w}$ is computed

$$\hat{\mathbf{w}} = FF^T\mathbf{w}, \tag{4.3}$$

where $F$ is the truncated KL model. In general a model fit is calculated $FF^\dagger\mathbf{w}$, where $F^\dagger = (F^TF)^{-1}F^T$. For the KL model, however, the basis vectors form an orthonormal set, so $F^TF = I$ where $I$ is the identity matrix and $F^\dagger = F^T$. Computing model fits to several swaths of point-wise retrieved data, we may note the error variation across the region (see Fig. 4.3.1). Note that the wind vectors in a circle about the center of the region appear to contribute the most to a model fit. We may conclude that the KL model gives preference to those locations within the region. Such a conclusion suggests that the observed errors of Fig. 4.3.1 are logical weights when averaging wind regions. Thus, to average cell estimates each estimate is weighted according to region position by the inverse of the average modeling error.

### 4.3.2 Algorithm Summary

To perform the field-wise median filter, each region is initialized to its first ambiguity (the global minimum of the field-wise objective function). For each region in the swath, all of the neighboring regions are combined. As there is overlap between regions, this creates an averaged wind field coincident with the given region. Each ambiguity for the given region is then compared to the averaged wind field, and the ambiguity with the smallest error is selected. After the process has been performed for every region in the swath, the process is repeated iteratively until the swath estimate stabilizes (i.e., each region remains with the same selected ambiguity).

### 4.3.3 Parameter Choices

As mentioned in Chapter 3, the estimated regions are $12 \times 12$ wvc in size. In his implementation, Richards used $24 \times 24$ wvc regions and overlapped by 50% in the alongtrack direction and either 25% or 33% in the crosstrack direction (depending

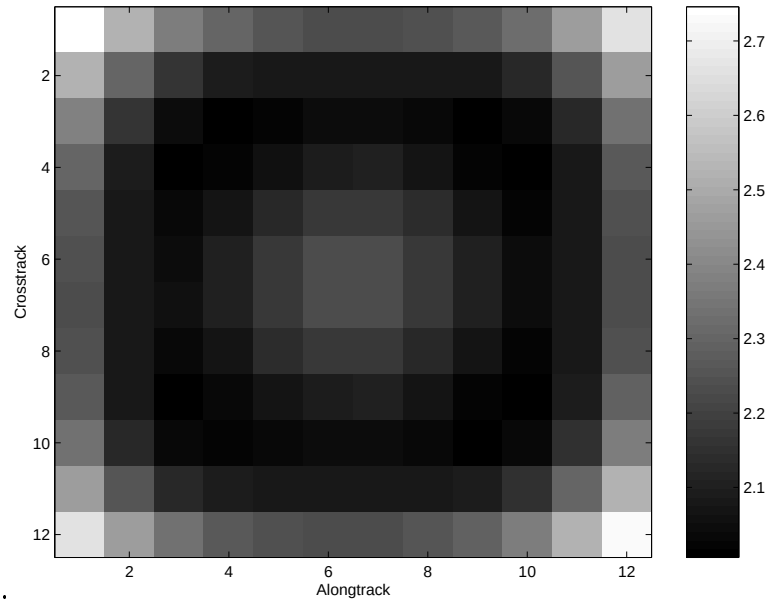


Figure 4.2: *Average modeling error as a function of region location. The region edges, especially the corners are much less accurate than the region's center.*

upon swath position). The minimal overlap reduced the number of coincident estimates between regions which are essential to performance of the field-wise median filter. When the neighboring regions are combined for comparison many cells are the result of averaging the less-accurate region edges, and few from the region center. Such an overlapping scheme degrades the apparent performance of the field-wise median filter.

In this thesis, the overlap is set at 67% in both the alongtrack and crosstrack directions. Thus, in performing the field-wise median filter, the field averaged from neighboring fields contains many region estimates, and thus reduces spurious comparisons. The resulting swath estimate in general is significantly more consistent than fields with less overlap.

Overlapping regions and averaged fields come with performance tradeoffs. While high levels of overlap facilitate ambiguity selection, heavily averaged estimates may be considered sub-optimal. In the final swath estimate, each wind vector, rather

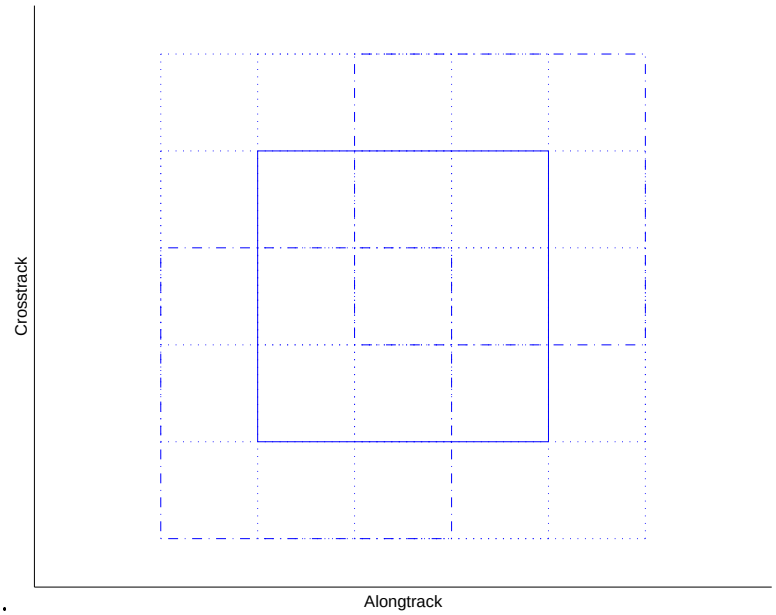


Figure 4.3: *Region overlap for field-wise estimation. Region to be selected outlined by solid line.*

than being the result of a direct estimation, is an average result from several estimates. In addition, greater overlap increases the number of wind regions that must be estimated, thus increasing computational costs.

## 4.4 Results

As mentioned previously, a discussion of results for wind retrieval is almost completely limited to anecdotal evidence. As such, the results for the field-wise median filter are limited to discussing observed features of the estimated swaths.

The aim of field-wise ambiguity selection is two-fold; it must select a consistent swath estimate from among the region ambiguities while maintaining the desirable characteristics of field-wise estimation. Specifically, the selected swath estimate must continue to improve noise performance and successfully estimate general wind flow, even in the presence of contamination. The ambiguity selection must generate a consistent swath without diminishing the performance of individual ambiguities.

In general, the field-wise median filter is quite successful at maintaining the characteristic performance of the individual ambiguities. Of primary consideration is noise reduction, while maintaining overall wind flow. Figure 4.4 shows the field-wise and point-wise products in a region with a cyclone feature. Observe that the general wind flow is nearly identical between the two, while the field-wise product is noticeably noise reduced.

Figure 4.5 demonstrates field-wise performance amid rain contamination. The center of the shown swath has the signature common to rain contamination; specifically wind vector estimates perpendicular to the path of the satellite. Once again the field-wise performance appears superior to point-wise, as the swath estimate is not so glaringly inconsistent. The reader is reminded, however, that since truth data is not known the apparent improved performance is not necessarily a reflection of the actual near-surface wind.

Field-wise ambiguity selection is on a significantly reduced scale from point-wise ambiguity selection. Indeed, a stated advantage of field-wise wind retrieval is the simplification of the ambiguity selection step. While such a claim is accurate, it underscores a significant shortcoming of field-wise wind retrieval. An incorrect selection in point-wise estimation results in increased noise, often of the discrete variety. An incorrect selection for field-wise wind retrieval fails in a more detrimental manner. As the region is larger than one single cell, the ambiguity selection error is apparent over a much larger area. A field-wise ambiguity selection error has a broad spatial effect.

Figure 4.6 shows an example of a field-wise ambiguity selection error. Note that the boxed region reveals the algorithm's attempt to reconcile inconsistent estimates above and below. The resulting wind field is not very realistic, nor is the overall wind flow.

Errors in the wind field are not always the fault of the selection algorithm. The point-wise product in Fig. 4.6 shows equal difficulty reconciling the ambiguities for the region in question. While the error is not so glaring to the eye, adjacent estimates point in opposite directions. In most cases, field-wise trouble spots are coincident with

point-wise difficulties. Likely there are insufficient ambiguities to resolve the region in a consistent manner due to deficiencies in the data, not in the estimation process.

Another source for error is in the estimation process. While all near-global ambiguities should be obtained, as discussed in Chapter 3, this is not practical. An ambiguity selection error may be due to an incomplete solution set. That is, the "true" wind field is not among the ambiguities so the ambiguity selection algorithm is forced to select erroneously. Further minimization effort in trouble spots may yield improved results. This is a topic for further research.

Generalizing, the field-wise median filter is generally successful in selecting a consistent swath estimate from the ambiguities. The algorithm maintains the desirable qualities of field-wise estimation, namely improved performance in trouble spots such as rain contamination and nadir regions. Selection errors generally occur in regions which also contain point-wise errors, from which we may conclude that the data are inconsistent in such an area.

## 4.5 Spatial Frequency Enhancement

An advantage of field-wise wind retrieval is its "low-pass" characteristic. By using a truncated wind field model, a low frequency constraint is placed on the estimated winds. While the constraint demonstrably reduces the noise, the low-pass estimator can be criticized for "filtering out" signal as well. The field-wise product presented in this chapter is quite successful in displaying the large-scale wind flow, but the frequency resolution is quite limited. This chapter discusses the necessary frequency range to model the wind and a method for field-wise estimation that spans the range.

### 4.5.1 Frequency Content of the Wind

As discussed previously, both point-wise and field-wise wind retrieval assume that the wind field is, in general, low frequency. Point-wise wind retrieval applies the assumption during the ambiguity selection stage, and does so in an ad hoc manner. Using the PWMF, frequency is minimized over a small window. For field-wise wind

retrieval, the low frequency constraint can be quantified more precisely by examining the frequency content of the model bases.

The low frequency constraints are applied because the wind is assumed to have a red spectrum. In other words, on the scale at which the wind is estimated, the wind power is assumed to be located in the lower frequency ranges, with almost no power in the higher frequencies. Noise is generally modeled as a white spectrum for any problem. That is to say that the power spectral density is flat across all frequencies.

Draper [13] used the difference between the wind and noise spectra to quantify the extent of the wind spectrum. He examined the DFT of the point-wise retrieved winds for SeaWinds. He concluded that the majority of the wind energy is contained in the first 1/4-1/3 of the DFT bins. At higher frequencies, the spectrum is virtually flat, suggesting majority contribution from noise. The spatial frequency observation suggests a method for choosing KL model truncation points. For any given region size, the model should be chosen to span between 1/3 and 1/4 of the DFT bins.

For a 12 size region, to span 1/3 of the frequency bins, 12 model parameters should be included. To span 1/4 of the frequency bins, 26 model parameters should be chosen. This analysis suggests that 6 model parameters (the number used for the results in chapters 3 and 4) are insufficient to span the entire wind spectrum.

### 4.5.2 Increasing the Frequency Resolution

When choosing the number of model parameters to include in field-wise estimation, computational load is a significant consideration. To perform field-wise estimation, each valid region in the swath must be optimized several times to obtain a comprehensive solution set. For six model parameters, 23 minimizations are performed at each region in the swath.

A higher model order increases the computational load in two ways. First, the minimization algorithm is significantly more complicated as a function of model size. Each iteration grows in cost on the order of the square of the model size. In addition, while the algorithm successfully minimizes the higher order objective function, the convergence is significantly slower requiring more iterations. Also consider that the

number of initialization vectors needed to span the reasonable solution space grows as a function of model order. Thus, increasing the number of model parameters requires more optimizations that are significantly more intensive to perform.

The computational reality leads to an unsatisfactory tradeoff. To span the entire wind spectrum requires a model order between 12 and 26, but computational requirements practically limit the global search to an order of six.

A logical compromise is a two-step estimator. Starting with no prior knowledge requires a global search of the kind detailed in Chapter 3. As the global search necessitates many optimizations per region, a low model order must be used to facilitate the computation. The estimated regions are then ambiguity selected using the field-wise median filter. Finally, the results of the field-wise median filter are used as the initial vectors in a higher order search.

As the general wind flow is already known before the higher order estimation, the minimization is reduced from global to local, necessitating only a single optimization. Given that the power spectrum of the wind is denser in the low frequencies, it is hypothesized that essentially one higher order alias is to be expected close to the field-wise selected wind.

### 4.5.3 High Model Order Results

As only one optimization is necessary per region, the second optimization is performed with 26 model parameters. In general, the resulting wind estimates appear quite reasonable. The higher order wind fields maintain the same characteristics as the lower order estimates, while including a larger range of frequencies. Figures 4.7, 4.8, and 4.9 show resulting wind estimates using 26 model parameters. The regions are the same as are shown in Figs. 4.4, 4.5, and 4.6.

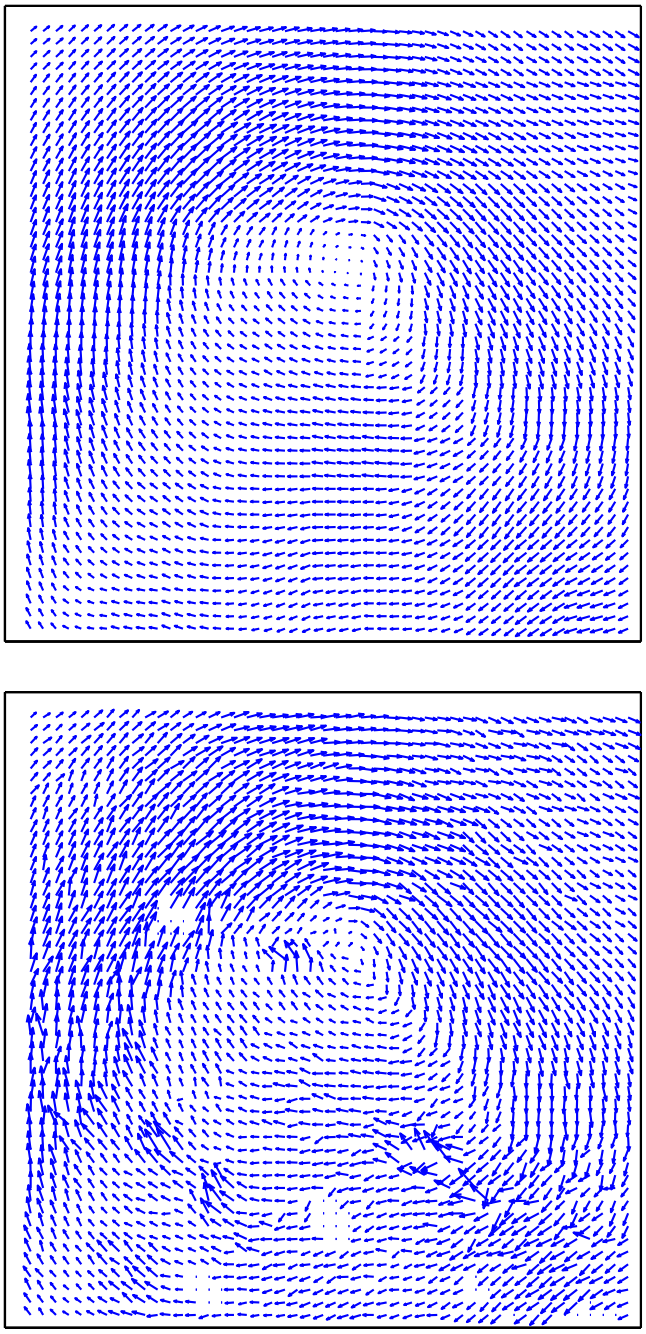

Figure 4.4: *A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 3081 beginning at alongtrack 150. Note that the field-wise product demonstrates a significant reduction in noise. The rain artifacts apparent in the point-wise product are not evident in the field-wise product.*

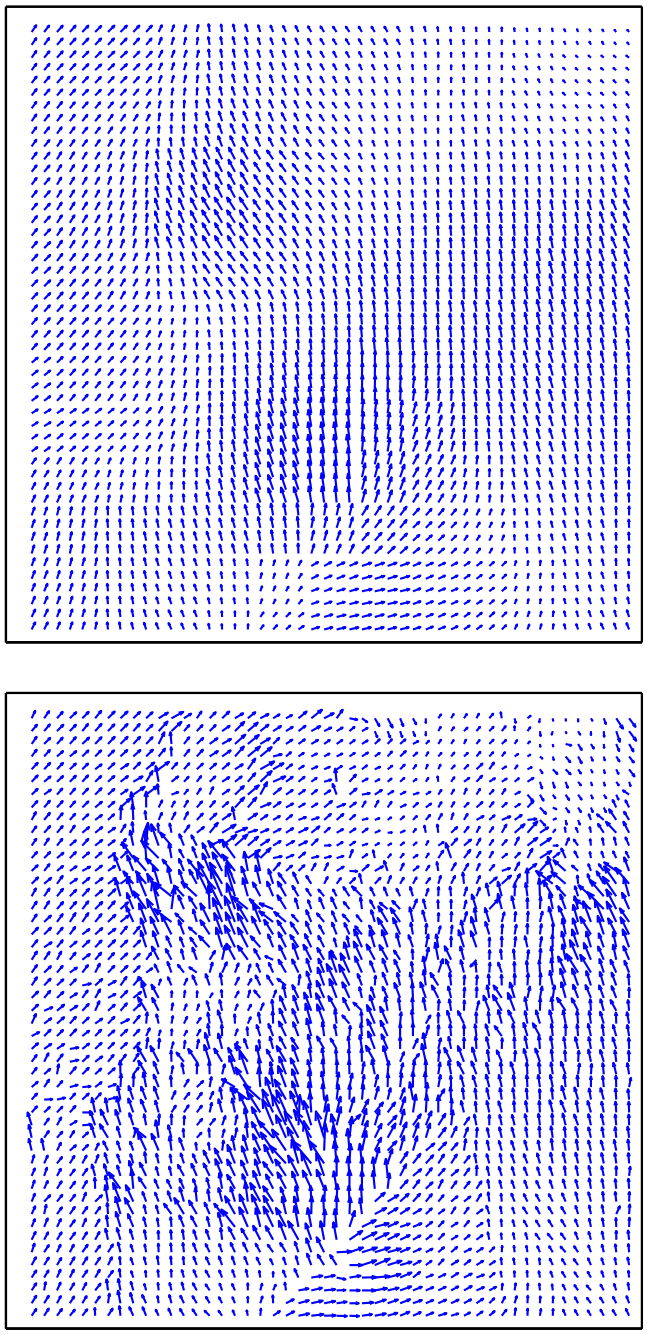

Figure 4.5: *A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 2176 beginning at alongtrack 415. The point-wise product shows many cell with rain contamination which greatly distort the overall wind flow. The field-wise product successfully models the wind flow without showing the signature of rain contamination.*

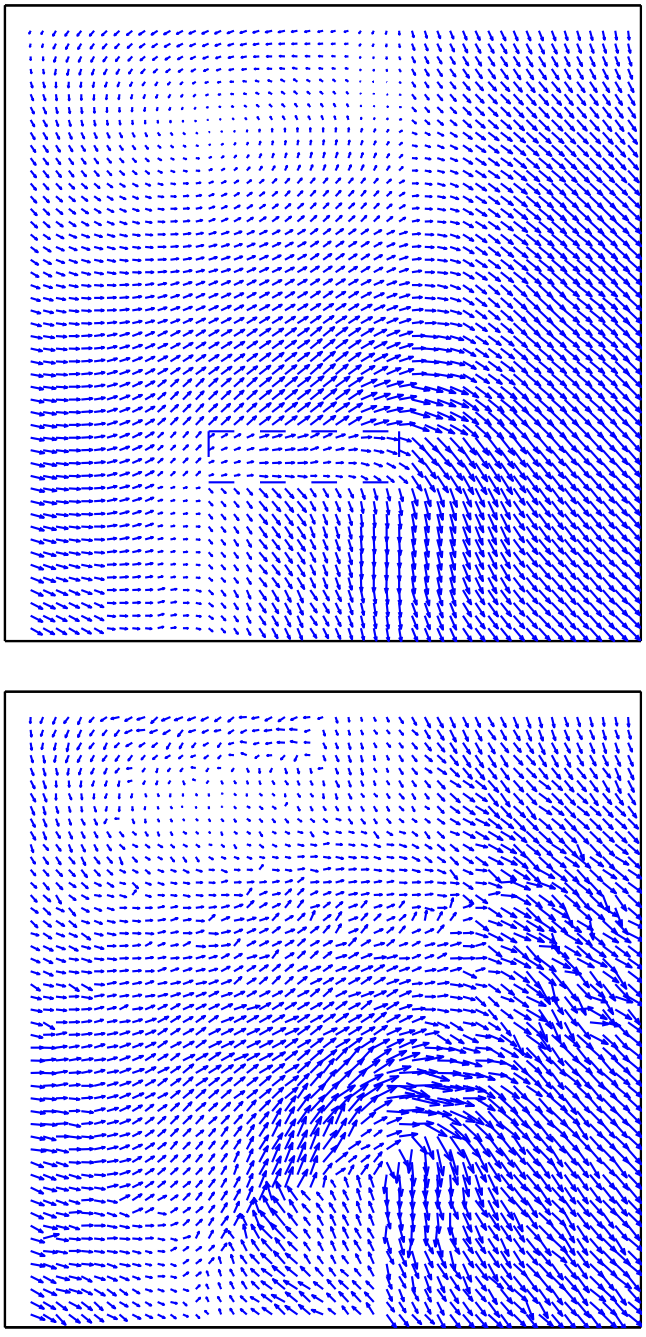

Figure 4.6: *A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 04790 beginning at alongtrack 155. Note the ambiguity selection error in the field-wise product outlined by the dotted line.*

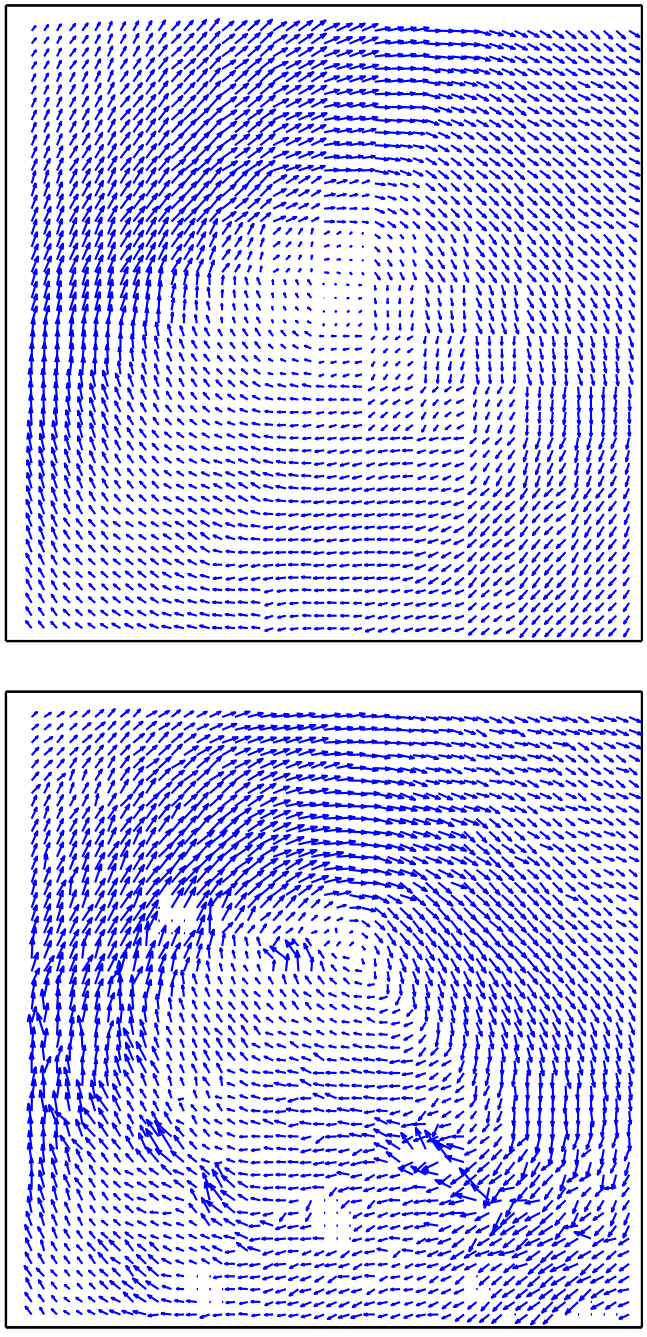

Figure 4.7: *A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 3081 beginning at alongtrack 150. The field-wise product is performed using 26 model parameters.*

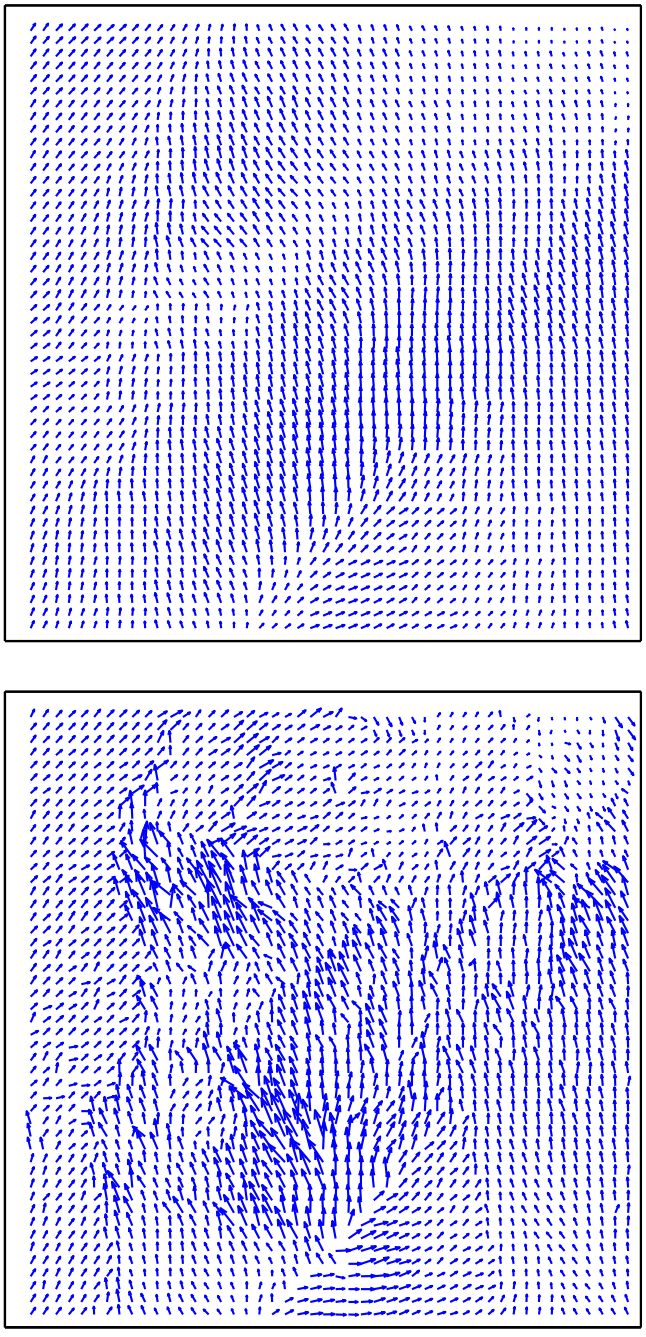

Figure 4.8: *A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 2176 beginning at alongtrack 415. The field-wise product is performed using 26 model parameters.*

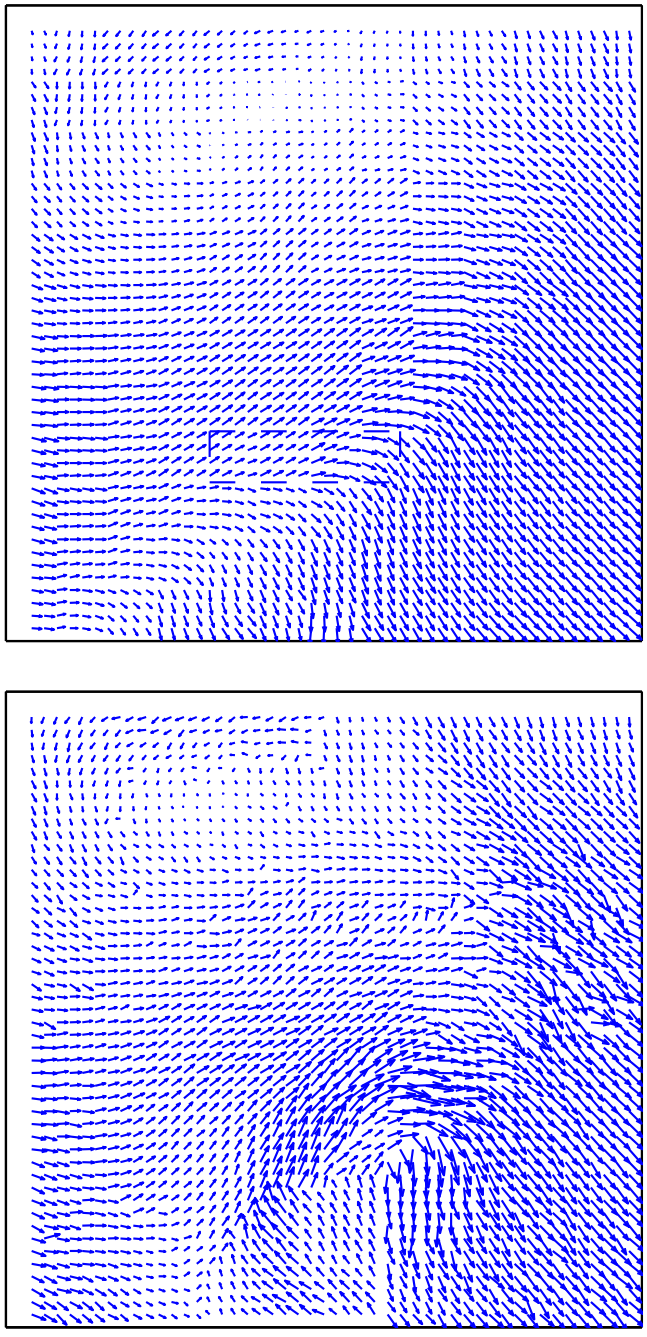

Figure 4.9: *A comparison between the point-wise and field-wise ambiguity selected products for QuikSCAT Rev 04790 beginning at alongtrack 155. The field-wise product is performed using 26 model parameters.*

# Chapter 5

## Examination of the Backscatter Noise Model

The wind retrieval objective function, both for the point-wise and the field-wise case, utilizes a model for the $\sigma^o$ measurements. The model, presented in Chapter 2 and defined in Eq. (2.8), contains representations for measurement noise ($K_{pc}$) and for modeling noise ($K_{pm}$). The measurement noise, also known as communication noise, can be derived from first principles [8] and is modeled as quadratic in SNR. The modeling noise for the wind estimation problem is less understood. Johnson [8] developed a model similar in form to that used for $K_{pc}$ and attempted an iterative estimation of the value $K_{pm}$. While the iterative approach demonstrated surprisingly large values for $K_{pm}$, the lack of a closed form representation results in the parameter being disregarded in wind estimation implementations.

For simplicity in the field-wise implementation presented in this thesis, the noise model of Eq. (2.8) is reduced in two manners. First, as mentioned in Chapter 2, $K_{pm}$ is assumed to be 0. The second reduction, represented by Eq. (3.6), bases the communication noise on the observed backscatter, rather than the "true" backscatter. By such an assumption, the noise variance is no longer a function of the estimated wind, simplifying the objective function considerably.

This chapter examines the validity of the noise model in both its complete and reduced form. A separate noise parameterization is presented, which is based solely on data from the SeaWinds instrument.

## 5.1 Evaluating the Noise Model

With months of SeaWinds data available as a sample set, a verification of the noise model is possible. To develop the experiment, recall that the observed backscatter is modeled as a random variable

$$z = (1 + K_{pc} v_2)\sigma^o, \tag{5.1}$$

where $\sigma^o$ represents the "true" backscatter. Using the assumption that $K_{pm}$ is zero, the "true" backscatter is the forward projection of the wind vector $\mathbf{w}$ (*i.e.*, $\sigma^o = \mathcal{M}(\mathbf{w})$). While the true wind is unknown to generate this model, we can use the estimated wind vector to determine the measurement variance.

Assuming that the point-wise estimated winds are equivalent to the true winds in the mean, we may generate a histogram to demonstrate the accuracy of the noise model. The histogram permits a comparison of the variance estimate of Eq. (3.6) with the more traditional Eq. (3.3).

The experiment is performed as follows. The quantity $y = z - \mathcal{M}(\hat{\mathbf{w}})$ is a zero mean random variable. The variance is represented by either

$$\varsigma_1^2 = \alpha \mathcal{M}(\hat{\mathbf{w}})^2 + \beta \mathcal{M}(\hat{\mathbf{w}}) + \gamma, \tag{5.2}$$

or

$$\varsigma_2^2 = \alpha z^2 + \beta z + \gamma. \tag{5.3}$$

The quantities $\frac{y}{\varsigma_1}$ and $\frac{y}{\varsigma_2}$ are calculated for 50 revs of SeaWinds data and a histogram is created.

The resulting histograms, shown in Fig. 5.1, have an interesting symmetry. At first glance, $\varsigma_1$ and $\varsigma_2$ appear to invert the distribution about $y = 0$. This interpretation is incorrect. Note that both distributions are calculated with the same values of $y$. Given that both $\varsigma_1$ and $\varsigma_2$ are always positive values, the normalization does not affect the sign of the random variable.

The symmetry may be instead interpreted in terms of variance bias. The random variable $y$ is calculated as the difference between $z$ and $\mathcal{M}(\hat{\mathbf{w}})$, which are

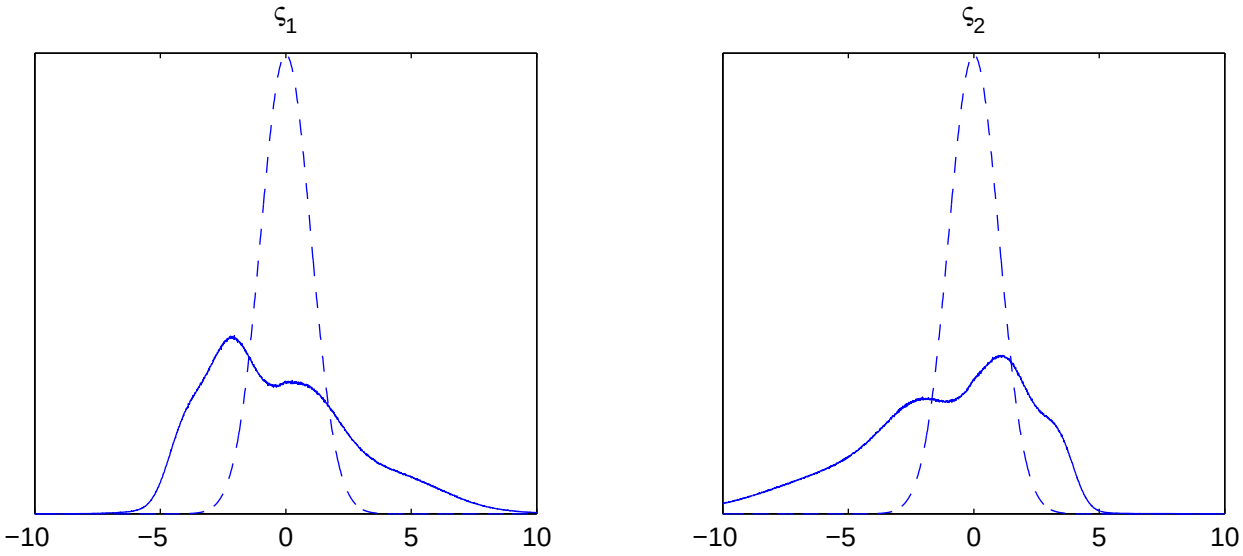


Figure 5.1: *A histogram of the random variable* $y = z - \mathcal{M}(\hat{\mathbf{w}})$ *normalized with (a)* $\varsigma_1$ *and (b)* $\varsigma_2$*. Overlaying both figures is a unit normal distribution for comparison.*

used to calculate $\varsigma_2$ and $\varsigma_1$ respectively. When $y > 0$, the value of $z$ is larger, and the $\varsigma_2$ normalization is aggressive, causing a small tail, while the $\varsigma_1$ normalization is weak, allowing a longer tail. For $y < 0$ the performance is reversed.

## 5.2 Data-driven Parameterization of Backscatter Measurements

Aside from a discussion of the symmetries, Fig. 5.1 demonstrates some deficiencies in the current $\sigma^o$ distribution. With $K_{pm}$ assumed to be 0, the calculated value of $K_{pc}$ is insufficient to normalize the backscatter vectors to unit variance. Furthermore, the Gaussian assumption appears an oversimplification of the problem.

While the non-Gaussian nature of the measurements is intrinsic, the large sample of measurements and corresponding wind estimates can be used to parameterize the variance of $y$. By calculating a joint histogram of $z$ and $\mathcal{M}(\hat{\mathbf{w}})$ (shortened hereafter to $\mathcal{M}$) we may approximate $p(z, \mathcal{M})$, the joint probability density function. The joint pdf allows data-driven estimates for all of the desired moments $z$, $\mathcal{M}$, and $y$.

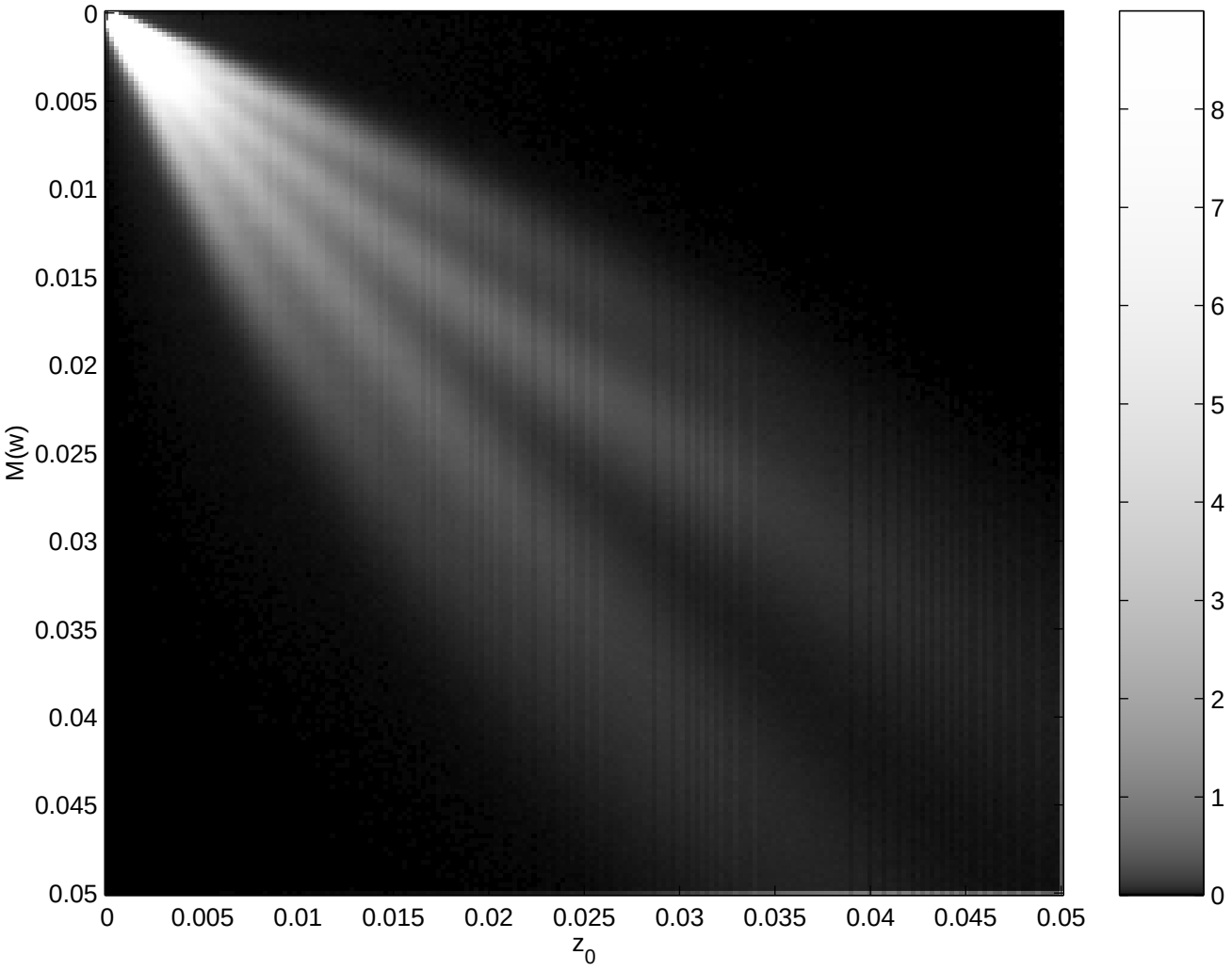


Figure 5.2: *The joint distribution of $z$ and $\mathcal{M}$*

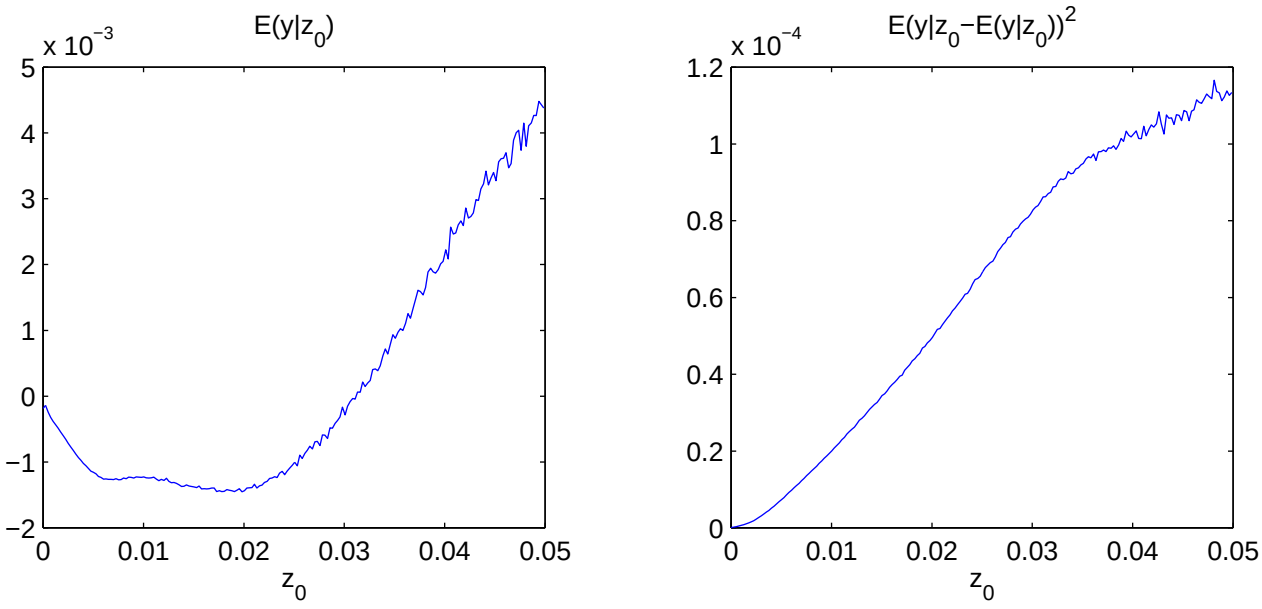


Figure 5.3: *The first and second central moments of* $y|z_0 = z - \mathcal{M}|z_0$. *In words, for a given observed measurement* ($z_0$) *value, the figures show the measurement bias and variance.*

The moments of $y$ vary both as a function of $z$ and $\mathcal{M}$, but it is desirable to parameterize in $z$, as $z$ is the observed backscatter, and is known prior to estimation. Thus, for a given $z_0$, we may calculate the central moments of $y$,

$$E(y|z_0) = E(z_0 - \mathcal{M}|z_o) = z_0 - E(\mathcal{M}|z_0) = B(z_0), \tag{5.4}$$

$$\begin{aligned} E(y|z_0 - E(y|z_0))^2 &= E(z_0 - \mathcal{M}|z_0 - B(z_0))^2 \\ &= E(\mathcal{M}|z_0)^2 - (z_0 - B(z_0))^2. \end{aligned} \tag{5.5}$$

A plot of each moment appears in Fig. 5.3.

It is of interest to note that $y$ is not zero mean. The first moment has a bias which is negative over the first half of the range of values for $z$. The larger values of $z$, where the bias is positive, are much less common comprising less than 5% of the observed backscatter measurements. Thus for most backscatter, $\mathcal{M}$ is greater than $z$. This is verified by the histograms taken in Fig. 5.1. In each case, there are more values of $y$ less than zero (although this is difficult to ascertain from the figures).

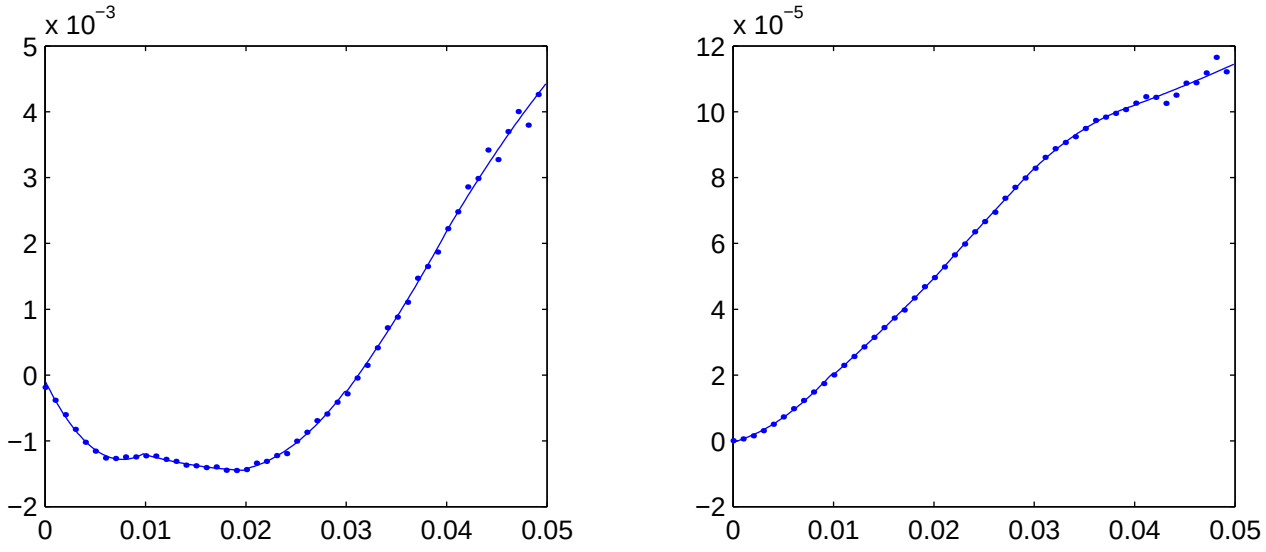


Figure 5.4: *A piecewise polynomial fit to the first and second moments of* $y = z - \mathcal{M}$. *The solid line is the parameterization and the dots represent measured data points.*

As the dependence on $z$ is quantifiable, both the bias and the variance may be parameterized. Both moments are too complicated to be represented by a single polynomial fit without a prohibitively high the polynomial order. Instead, the data is segmented into smaller pieces, and each piece is fit with a second order polynomial. The resulting piece-wise polynomial fit is shown in Fig. 5.4.

Using the parameterization, a histogram is computed for the random variable $y'$

$$\begin{aligned} y' &= \frac{(z - \mathcal{M} - b)}{\varsigma}, \\ \varsigma &= \alpha_{vp} z^2 + \beta_{vp} + \gamma_{vp}, \\ b &= \alpha_{bp} z^2 + \beta_{bp} + \gamma_{bp}, \end{aligned} \tag{5.6}$$

where $\alpha$, $\beta$, and $\gamma$ are the piece-wise polynomial coefficients for the variance and the bias. The resulting histogram is shown in Fig. 5.5.

While it is clear from Fig. 5.5 that the measurements are not normally distributed, the first two moments can be represented quite accurately. Thus the error

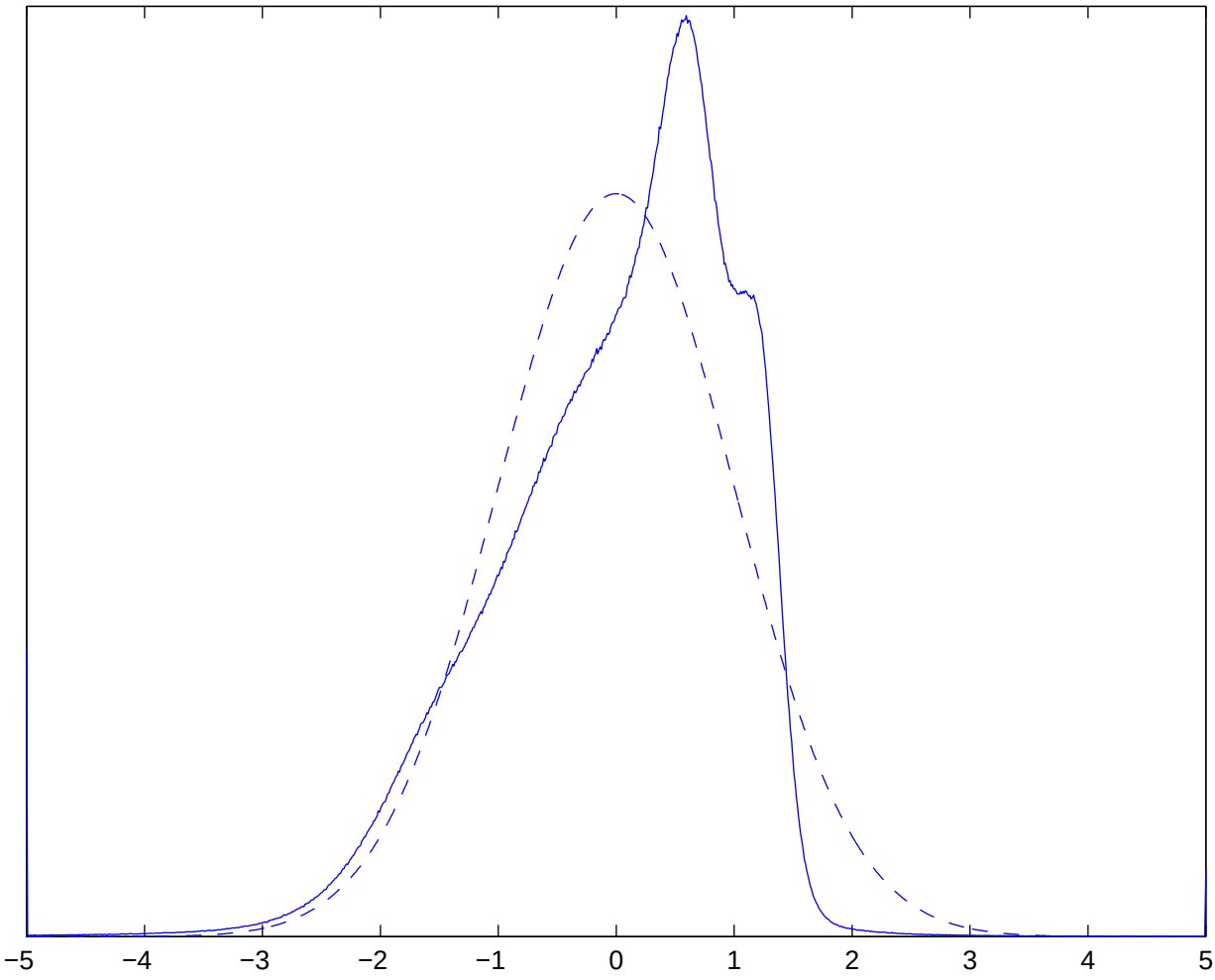


Figure 5.5: *A histogram of $y'$, the parameter normalized backscatter. The parameterization of the bias successfully normalizes the distribution, i.e. $y'$ is zero-mean and unit variance. The unit Gaussian distribution is included for reference.*

between the forward projection $\mathcal{M}(\hat{\mathbf{w}})$ and the observed backscatter $z$ can be normalized to zero mean and unit variance to a much greater degree of accuracy.

### 5.3 Estimator Performance Bounds

The empirical noise model adds legitimacy to several statistical analyses. Perhaps the most valuable analysis comes in the realization of bounds on the error covariance. Oliphant and Long [6] analyzed the accuracy of the wind estimates based upon the Cramer-Rao (C-R) bound. Using the parameterized bias and variance models, these results can be refined. Given the statistical accuracy of the first and second moments, greater confidence may be placed upon the results.

The C-R bound may be expressed

$$R(\mathbf{w}) = E_{\mathbf{z}}\{[\hat{\mathbf{w}} - \mathbf{w}][\hat{\mathbf{w}} - \mathbf{w}]^T\} \tag{5.7}$$

$$\geq J^{-1}(\mathbf{w}), \tag{5.8}$$

where $J(\mathbf{w})$ is the Fisher information matrix (FIM). The FIM is defined as

$$J(\mathbf{w}) = E\left\{\left[\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial \mathbf{w}}\right]^T \frac{\partial L(\mathbf{w},\mathbf{z})}{\partial \mathbf{w}}\right\}, \tag{5.9}$$

where

$$L(\mathbf{w},\mathbf{z}) = \ln p(\mathbf{z}|\mathbf{w}). \tag{5.10}$$

We may write the distribution of the backscatter measurements as

$$p(\mathbf{z}|\mathbf{w}) = \frac{1}{(2\pi)^{K/2}\sqrt{\sum_{k=1}^{K}\varsigma_k^2}} \exp\left\{-\frac{1}{2}\sum_{k=1}^{K}\frac{(\mathbf{z}_k - \mathcal{M}_k(\mathbf{w}) - b_k)^2}{\varsigma_k^2}\right\}. \tag{5.11}$$

Before proceeding, we should note the limitations of this model. As demonstrated in Fig. 5.5, the measurements are not normally distributed. The first two moments are known, however, and thus the Gaussian model in Eq. (5.11) is accurate only in mean and variance.

Furthermore, both $b$ and $\varsigma$ are estimates parameterized by the realization of $z$. In essence, the variance and mean of the random variable are functions its realization.

While this is not good statistical practice, it is included here to simplify the expression of $J(\mathbf{w})$. The same assumption is made in the variance estimate of Eq. (3.6).

Using Eq. (5.11),

$$\frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_i} = \sum_{k=1}^{K} \frac{(\mathbf{z}_k - \mathcal{M}_k(\mathbf{w}) - b_k)}{\varsigma_k^2} \frac{\partial \mathcal{M}_k(\mathbf{w})}{\partial w_i}, \tag{5.12}$$

and

$$(J)_{ij} = E_{\mathbf{z}} \frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_i} \frac{\partial L(\mathbf{w},\mathbf{z})}{\partial w_j} \tag{5.13}$$

$$= E_{\mathbf{z}} \sum_{k=1}^{K} \sum_{l=1}^{K} \frac{(\mathbf{z}_k - \mathcal{M}_k(\mathbf{w}) - b_k)}{\varsigma_k^2} \frac{\partial \mathcal{M}_k(\mathbf{w})}{\partial w_i} \times \frac{(\mathbf{z}_l - \mathcal{M}(\mathbf{w}) - b_l)}{\varsigma_l^2} \frac{\partial \mathcal{M}_l(\mathbf{w})}{\partial w_j}. \tag{5.14}$$

Assuming the measurements are independent, the elements of the FIM simplify to

$$(J)_{ij} = \sum_{k=1}^{K} \frac{E_{\mathbf{z}}(\mathbf{z}_k - \mathcal{M}_k(\mathbf{w}) - b_k)^2}{\varsigma_k^4} \frac{\partial \mathcal{M}_k(\mathbf{w})}{\partial w_i} \frac{\partial \mathcal{M}_k(\mathbf{w})}{\partial w_j} \tag{5.15}$$

$$= \sum_{k=1}^{K} \frac{1}{\varsigma_k^2} \frac{\partial \mathcal{M}_k(\mathbf{w})}{\partial w_i} \frac{\partial \mathcal{M}_k(\mathbf{w})}{\partial w_j}. \tag{5.16}$$

(Note that in this equation the vector $\mathbf{w}$ may be composed of either the speed and the direction of the wind vector, or the rectangular components of the wind field, $u$ and $v$.)

Averaging the C-R bound for each rectangular component as a function of crosstrack position, we may compare the bound as derived in [6] with the new distribution presented here. To help compare the bounds, we also estimate the estimator variance by correlating the selected wind alias with the ECMWF nudging field. While ECMWF winds cannot be considered truth data, in the mean they provide a statistical measure of the estimate moments. Figure 5.6 shows the bound under each noise model as well as the variance computed against the ECMWF fields.

From Fig. 5.6 it is evident that the parameterized estimate for the C-R bound is much closer to the observed variance of the estimate. Neither, however, approaches the variance estimate computed with ECMWF fields. To explain the significant

separation between the bound and the ECMWF estimated variance, we note three caveats. First, the ECMWF fields are considered "true" wind fields when generating the estimated variance. The ECMWF is a weather model, and its inaccuracies, as well as inaccuracies in the scatterometer derived estimate, are included in the variance. We may consider the curve in Fig. 5.6 as an upper limit on the estimator variance. Secondly, the C-R bound requires evaluating the GMF at the true wind. As the true wind is unavailable, the selected ambiguity is used in its place. While necessary given the absence of truth data, this practice likely introduces inaccuracies in the C-R bound estimate. Finally, it should be noted that for an estimator with ambiguous solutions at low SNRs, the C-R bound is less accurate [15].

### 5.3.1 Performance Comparisons using the C-R Bound

The C-R bound can be computed for the field-wise model parameters by the same derivation as the point-wise model. The distribution for field-wise is

$$\begin{aligned} p(\mathbf{z}|\mathbf{x}) &= \frac{1}{(2\pi)^{MNK}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\varsigma_{ijk}^2}} \times \\ &\quad \exp\left\{-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\frac{(\mathbf{z}_{ijk}-\mathcal{M}_{ijk}(\mathbf{x})-b_{ijk})^2}{\varsigma_{ijk}}\right\}, \end{aligned} \tag{5.17}$$

and the elements of the FIM are given by

$$(J)_{pq} = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\frac{1}{\varsigma_{ijk}^2}\frac{\partial\mathcal{M}}{\partial\mathbf{x}_p}\frac{\partial\mathcal{M}}{\partial\mathbf{x}_q}. \tag{5.18}$$

Before the point-wise bounds and the field-wise bounds can be compared, we must examine several issues. First of all, the field-wise C-R bound is performed in the parameter space. In other words, the variance of each parameter estimate is bounded, rather than each individual wind vector.

The second issue is examined by Oliphant [9]. Oliphant writes the error covariance matrix as a sum of two matrices, the covariance of the wind estimate and

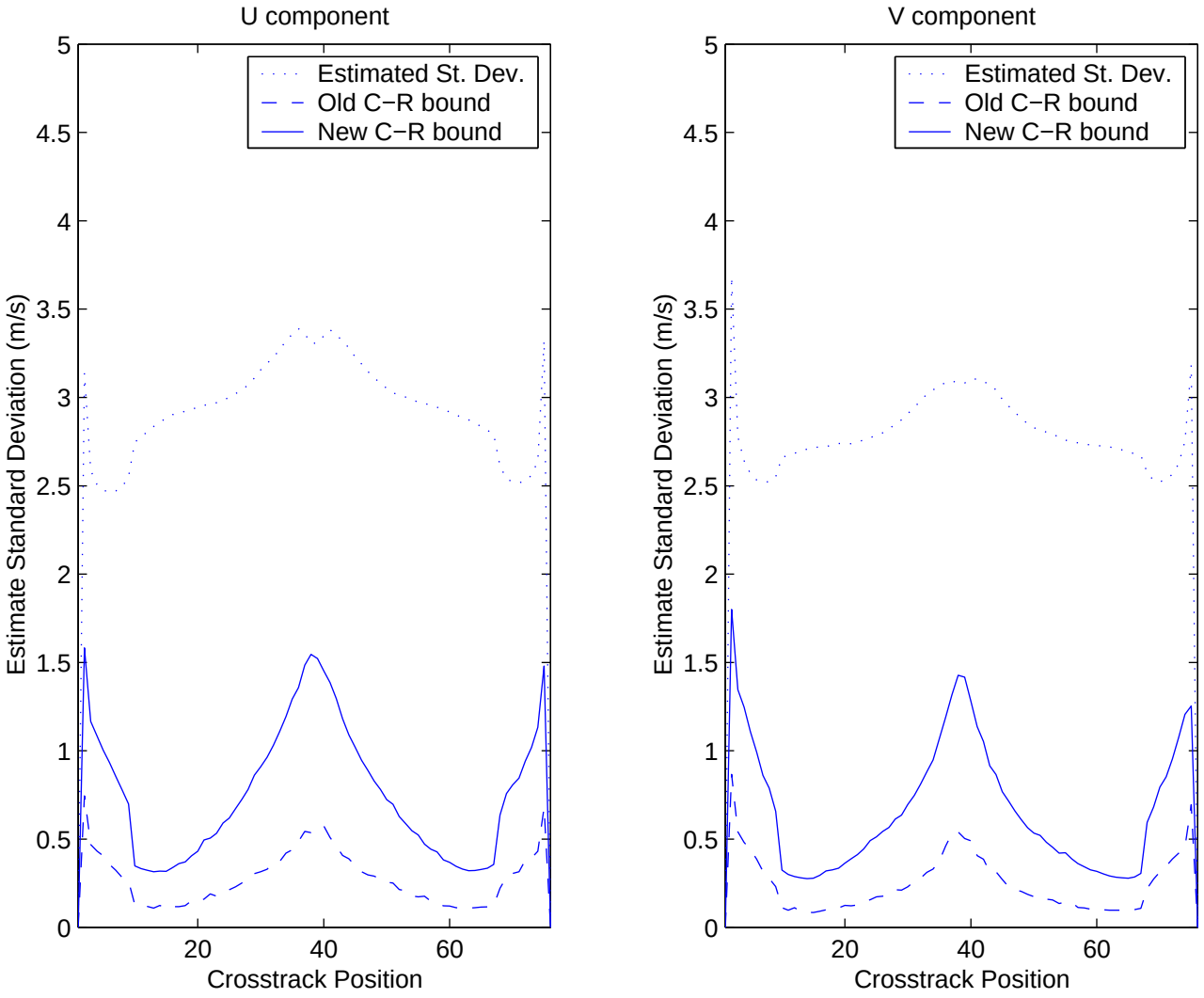


Figure 5.6: *The point-wise Cramer-Rao bound with the traditional variance estimate and the parameterized variance estimate. An estimated variance based on ECMWF nudging fields is also presented.*

the modeling error matrix:

$$\begin{aligned} C_{\mathbf{w}}^{e} &= E[\hat{\mathbf{w}} - \mathbf{w}]^T[\hat{\mathbf{w}} - \mathbf{w}] \\ &= E[\hat{\mathbf{w}} - E\hat{\mathbf{w}} + E\hat{\mathbf{w}} - \mathbf{w}]^T[\hat{\mathbf{w}} - E\hat{\mathbf{w}} + E\hat{\mathbf{w}} - \mathbf{w}] \\ &= E[\hat{\mathbf{w}} - E\hat{\mathbf{w}}]^T[\hat{\mathbf{w}} - E\hat{\mathbf{w}}] - E[E\hat{\mathbf{w}} - \mathbf{w}]^T[E\hat{\mathbf{w}} - \mathbf{w}] \\ &= C_{\mathbf{w}} + C_{\mathbf{w}}^{F}. \end{aligned} \tag{5.19}$$

Assuming that the estimator is unbiased, (*i.e.*, $E\hat{\mathbf{x}} = \mathbf{x} = F^T\mathbf{w}$), the C-R bound places a bound only upon the matrix $C_{\mathbf{w}}$. Before comparing the point-wise C-R bound with the field-wise result, the modeling error matrix $C_{\mathbf{w}}^{F}$ must also be included.

For the unbiased estimator, the error covariance matrix can be written [9]

$$\begin{aligned} C_{\mathbf{w}}^{F} &= E[E\hat{\mathbf{w}} - \mathbf{w}]^T[E\hat{\mathbf{w}} - \mathbf{w}] \\ &= E[FE\hat{\mathbf{x}} - \mathbf{w}]^T[FE\hat{\mathbf{x}} - \mathbf{w}] \\ &= E[FF^T\mathbf{w} - \mathbf{w}]^T[FF^T\mathbf{w} - \mathbf{w}] \\ &= (FF^T - I)E\mathbf{w}\mathbf{w}^T(FF^T - I). \end{aligned} \tag{5.20}$$

For the KL model, we may further simplify this expression for $C_{\mathbf{w}}^{F}$. Recall that the KL model is computed by taking the eigen decomposition of the wind autocorrelation matrix, $E\mathbf{w}\mathbf{w}^T$. The model matrix $F$ is composed of the eigenvectors associated with the largest eigenvalues. Thus,

$$E\mathbf{w}\mathbf{w}^T = \begin{bmatrix} F & G \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} F^T \\ G^T \end{bmatrix} \tag{5.21}$$

$$= F\Sigma_1 F^T + G\Sigma_2 G^T, \tag{5.22}$$

where $F$ is the transform matrix, $G$ contains the eigenvectors not included in the model, and $\Sigma_1$ and $\Sigma_2$ represent the eigenvalues of $E\mathbf{w}\mathbf{w}^T$. (Note that $F^TG = 0$ as

the eigenvectors are mutually orthogonal.) Using Eq. (5.22), Eq. (5.20) becomes

$$\begin{aligned} C_{\mathbf{w}}^{F} &= (FF^T - I)(F\Sigma_1 F^T + G\Sigma_2 G^T)(FF^T - I) \\ &= FF^T F\Sigma_1 F^T FF^T - F\Sigma_1 F^T FF^T - FF^T F\Sigma_1 F^T + F\Sigma_1 F^T \\ &+ G\Sigma_2 G^T \\ &= 2F\Sigma_1 F^T - 2F\Sigma_1 F^T + G\Sigma_2 G^T \\ &= G\Sigma_2 G^T. \end{aligned} \tag{5.23}$$

To convert the field-wise C-R bound into a bound on individual wind vectors, we rewrite the bound equation and note the following:

$$E[\hat{\mathbf{x}} - \mathbf{x}][\hat{\mathbf{x}} - \mathbf{x}]^T \geq J^{-1}(\mathbf{x}) \tag{5.24}$$

$$FE[\hat{\mathbf{x}} - \mathbf{x}][\hat{\mathbf{x}} - \mathbf{x}]^T F^T \geq FJ^{-1}(\mathbf{x})F^T \tag{5.25}$$

$$E[\hat{\mathbf{w}} - \mathbf{w}][\hat{\mathbf{w}} - \mathbf{w}]^T \geq FJ^{-1}(\mathbf{x})F^T \tag{5.26}$$

$$C_{\mathbf{w}} \geq FJ^{-1}(\mathbf{x})F^T. \tag{5.27}$$

Including the modeling error,

$$C_{\mathbf{w}}^{e} \geq FJ^{-1}(\mathbf{x})F^T + G\Sigma_2 G^T. \tag{5.28}$$

This result is now directly comparable to the point-wise C-R bound. The average C-R bound in the field-wise swath is compared to the point-wise C-R bound in Fig. 5.7. Note that field-wise has significantly improved performance at nadir as well as along the swath edges, even accounting for modeling error. In general, field-wise C-R bound suggests superior performance to point-wise wind retrieval.

It is interesting to note that the field-wise C-R bound is comparatively flat across the swath. Nadir performance is not significantly worse than sweet-spot performance. Such a result suggests that field-wise is successful in incorporating measurements from neighboring wind vector cells to compensate for minimal data at nadir. The C-R bound quantifies claims of improvement in areas with limited measurements.

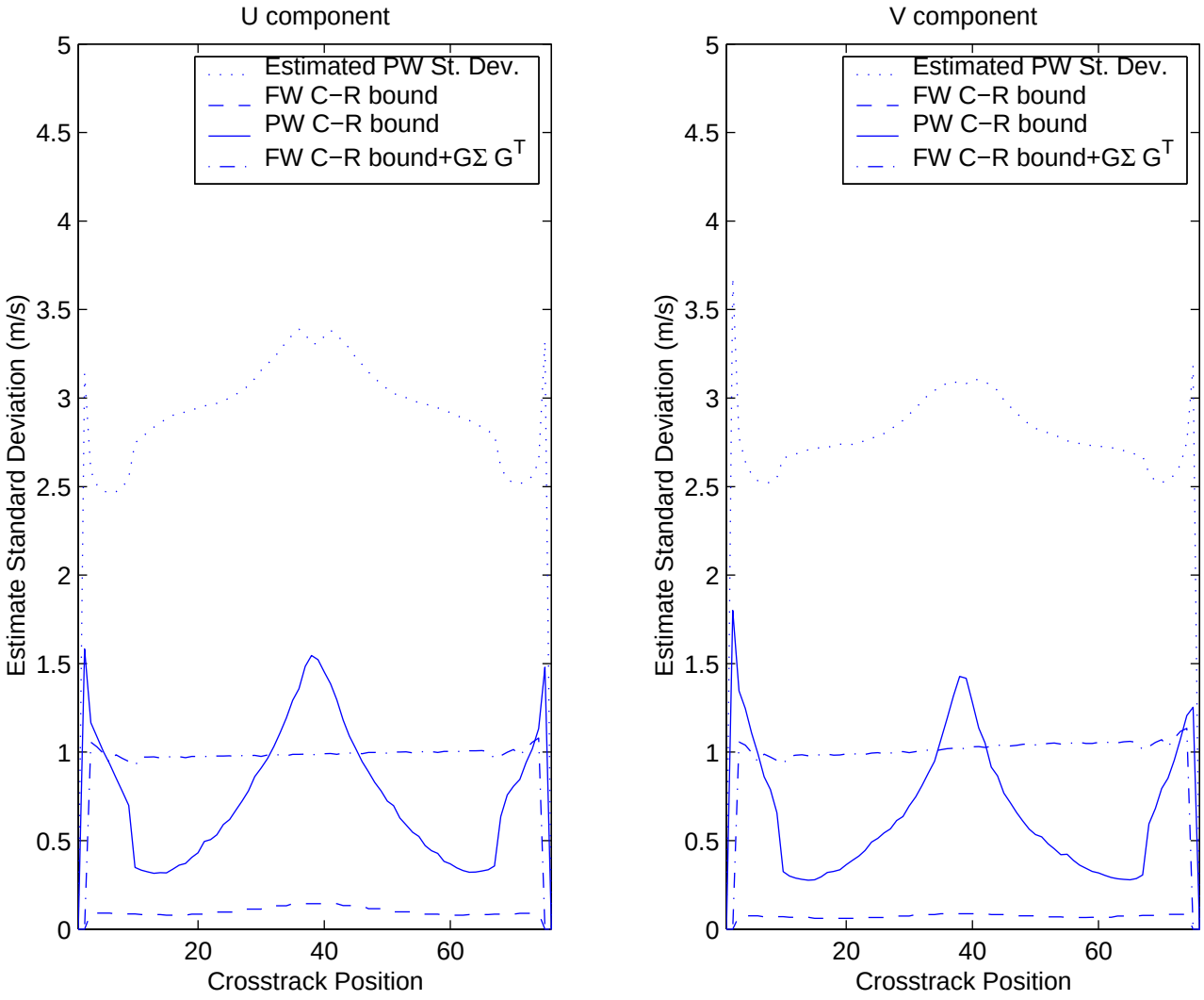


Figure 5.7: *The field-wise Cramer-Rao bound is compared to the parameterized point-wise Cramer-Rao bound. The field-wise bound is displayed twice, with and without accounting for modeling error. For comparison, the estimated point-wise variance using the ECMWF is also included.*

# Chapter 6

# Conclusion

This thesis has presented an implementation of field-wise wind estimation from actual scatterometer data and considered a number of key issues described below. The results are directly applicable to SeaWinds data, and can be implemented on a large scale with reasonable computational cost.

## 6.1 Summary of Contributions

The contributions of this thesis can be divided into two areas. First, an improved algorithm has been presented to minimize the objective function and and assure a sufficiently complete ambiguity set. Second, an improved ambiguity selection algorithm has been presented. The selection method includes a two-step estimator to verify that the wind spectrum can be completely represented. In addition to the field-wise estimation process, the measurement noise model has been examined. By creating a more accurate model, the effectiveness of both field-wise and point-wise wind estimates may be quantified. Each of these contributions is considered in greater detail below.

### 6.1.1 Objective Function Minimization

In Chapter 3, several objective functions were examined and compared. A weighted least squares objective function with a simplified noise model was presented for use in field-wise wind estimation. The Levenberg-Marquardt minimization algorithm was applied to the field-wise objective function. The gradient and Hessian matrices were derived for use within the algorithm.

Chapter 3 also discussed generating a sufficiently complete set of local minima for the objective function. The tradeoff between ambiguity set completeness and computational load was quantified, and a set of initial vectors was presented to maximize performance within reasonable computational parameters.

### 6.1.2 Swath Estimate Selection to Span the Wind Spectrum

A field-wise ambiguity selection algorithm was presented in Chapter 4. Modified from an algorithm presented by Richards [3], the swath estimate was selected using a field-wise median filtering approach. The modeling accuracy across the wind field region was presented, and the resulting modeling errors were used to combine adjacent region ambiguities into a consistent wind field.

The spectrum of model-based wind estimation was examined in Chapter 4 and compared with the estimated spectrum of the near-surface ocean winds. A method was presented to generate model-based estimates that completely span the wind spectrum. This was accomplished using a two-step estimator. The initial estimator used six model parameters to obtain a general wind flow. The second estimate was performed on a higher order (26 model parameters) to enhance the frequency resolution of the estimates.

### 6.1.3 Measurement Noise Model and Estimator Performance

The traditional models for scatterometer wind estimation were examined in Chapter 5. The measurement variance as used to weight the objective function was quantified, by examining a large set of SeaWinds data. It was found that the standard noise model underestimates the measurement variance. The existence of a bias in the point-wise estimator was verified and quantified.

A parameterization of the measurements was presented as a function of radar cross-section. Using a piece-wise polynomial fit, both the measurement bias and variance can be estimated from the backscatter. In this way, the distribution of the measurements as a function of wind vector may be known with greater accuracy.

Using the more accurate parameterized measurement distribution, the estimator performance may be known with greater accuracy. The Cramer-Rao bound was re-derived in Chapter 5 for both point-wise and field-wise estimation with the parameterized noise model. The bounds were compared, with field-wise retrieval showing distinctly lower estimate variance, particularly at nadir and in the swath edges.

## 6.2 Future Research

The field-wise estimation algorithm has been applied to a set of 55 revolutions of SeaWinds data. This set represents four days of measurements. While such an implementation is sufficient for developmental purposes, an evaluation over a range of revolutions should be performed and the results analyzed. Such an implementation could provide greater insight into the ambiguity selection process. Point-wise quality assessment algorithms [16] [17] could be adapted to the field-wise selection process.

While the implementation of field-wise estimation presented in this thesis has an acceptable computational cost, point-wise estimation is still much cheaper to perform. As field-wise and point-wise estimates have similar performance in many regions, future research could develop a hybrid scheme, using field-wise estimation only in regions contaminated by rain or demonstrating noisy estimates.

The parameterization of the noise model opens several areas of research. With a more accurate Cramer-Rao bound, the variance of the wind estimates may be examined on a more detailed basis. For example, the estimator variance may be computed as a function of along-track position, backscatter power, or azimuth angle. Estimator performance may also be examined as function of time, to analyze the stationarity of the instrument. The Cramer-Rao bound may also have value in quality assessment algorithms. Wind estimates with unusually high variance bounds may be flagged as potential errors. The bound may also be used in developing hybrid algorithms between point-wise and field-wise, flagging areas where field-wise estimation might be warranted.

# Appendix A

# Field-wise Objective Function Derivatives

In order to minimize the field-wise objective function, first derivative information is necessary. For the Levenberg-Marquardt algorithm implemented in this thesis, first and second derivative information is necessary. This appendix presents the four standard field-wise objective functions (Squared Error, Weighted Squared Error, Maximum Likelihood, and Reduced Maximum Likelihood) and derives the gradient for each and Hessian for the Squared Error and Weighted Squared Error Objective functions.

## A.1 Field-wise Objective Functions

To mathematically represent the measurements in a region, let $z$ be a three-dimensional array containing the observed values ($\sigma^o$) for each measurement. The first two dimensions index the measurement location (along-track and cross-track position) of a wind field region. The third dimension indexes the measurement number in each cell. Thus $z_{ijk}$ is the $k^{th}$ measurement of the $ij^{th}$ cell in the region. (Note that the number of measurements may vary at each swath location).

To represent the forward projection of one wind vector measurement, let $\mathcal{M}(U, \psi - \phi, \theta)$ be the GMF, where $\phi$ is the wind direction, $U$ is the wind speed, and $\psi$ is the instrument azimuth angle. As $\mathbf{x}$ parameterizes every wind vector in the region, the forward projection of the $k^{th}$ measurement of the $ij^{th}$ cell can be denoted $\mathcal{M}_{ijk}(\mathbf{x})$ where the measurement geometry is implied by $k$.

### A.1.1 The Squared Error (SE) Objective Function

Perhaps the simplest and most common objective function is the squared error objective function. This error metric can be described as the Euclidean distance or the $\mathcal{L}_2$ norm, which gives rise to the "least squares" solution. Thus, the objective function can be written as

$$J_{SE}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}(z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))^2, \tag{A.1}$$

where $K_{ij}$ is the number of measurements per cross-track cell, $ij$.

### A.1.2 The Weighted Squared Error (WSE) Objective Function

While certainly the simplest option, the squared error objective function fails to make use of all available information, and, as a result, can be overly sensitive to noise. The measuring instrument introduces noise that has been well studied. The noise is represented by a zero-mean, Gaussian random variable $\nu$, with variance $(K_{PC}^2)$ given by

$$\varsigma^2 = \alpha(\sigma_T^o)^2 + \beta\sigma_T^o + \gamma. \tag{A.2}$$

The parameters $\alpha, \beta$, and $\gamma$ are functions of the instrument design and signal to noise ratio (SNR), and $\sigma_T^o$ is the "true" $\sigma^o$ measurement (i.e. the $\sigma^o$ that would be observed in the absence of measurement noise). Thus, $\sigma^o$ is a realization of the random variable equation

$$\sigma^o = \sigma_T^o + \nu. \tag{A.3}$$

Using the variance estimate from Eq. (A.2), instrument noise can be accounted for, by dividing each term in the squared error sum by the measurement variance. Thus,

$$J_{WSE}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left(\frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2, \tag{A.4}$$

represents an objective function that can be classified as a "weighted squared error." It may be valuable to note that minimizing this objective function can be considered a maximum likelihood estimator, assuming that the variance of each measurement is

constant with respect to $\mathbf{x}$. This assumption will be examined in greater detail in the following section. $J_{WSE}(\mathbf{x})$ is also a quantity known as a "chi-square" ($\chi^2$).

### A.1.3 Maximum Likelihood (ML) Estimation

In the preceding section, the weighted squared error objective function was casually mentioned to be a maximum likelihood (ML) estimator given a constant measurement variance. The ML estimator is explicitly derived in this section.

The ML estimator calculates the model parameters most likely to give rise to the observed measurements. For a given $\mathbf{x}$, the estimator evaluates the probability that the observed measurements $z$ would occur. The estimated quantity $\hat{\mathbf{x}}_{ML}$ is the $\mathbf{x}$ that maximizes this probability. Thus,

$$\hat{\mathbf{x}}_{ML} = \arg\max_{\mathbf{x}} p_z(z|\mathbf{x}). \tag{A.5}$$

If the measurements are assumed to be independently Gaussian, with variance $\varsigma^2$ defined above, then

$$p_z(z|\mathbf{x}) = \prod_{i=1}^{N}\prod_{j=1}^{M}\prod_{k=1}^{K_{ij}} \frac{1}{\sqrt{2\pi\varsigma_{ijk}^2}} \exp\left\{-\frac{1}{2}\left(\frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2\right\}. \tag{A.6}$$

Computing the maximum of $p_z(z|\mathbf{x})$ is equivalent to computing the minimum of the negative log-likelihood function $\mathcal{L}(\mathbf{x}) = -\ln p_z(z|\mathbf{x})$, which is

$$\mathcal{L}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left\{\frac{1}{2}\ln 2\pi + \frac{1}{2}\ln\varsigma_{ijk}^2 + \frac{1}{2}\left(\frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2\right\}. \tag{A.7}$$

Note that the first two terms in the sum are constant with respect to $\mathbf{x}$, so they may be disregarded when calculating the arg min. The common scale factor of $\frac{1}{2}$ may also be ignored. Therefore the weighted squared error objective function is

$$\begin{aligned}
\hat{\mathbf{x}}_{ML} &= \arg\min_{\mathbf{x}} p_z(z|\mathbf{x}) \\
&= \arg\min_{\mathbf{x}} \left\{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left(\frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}}\right)^2\right\} \\
&= \arg\min_{\mathbf{x}}\{J_{SE}(\mathbf{x})\}.
\end{aligned} \tag{A.8}$$

Before declaring the weighted squared error a maximum likelihood estimator, the constant variance assumption needs to be examined in greater detail. Recall from Eq. (A.2) that $\varsigma^2$ depends upon the value of $\sigma_T^o$. Also recall that in computing $p_z(z|\mathbf{x})$, we estimate the probability of the observed measurements under the assumption that the true wind field is represented by $\mathbf{x}$. Under this assumption, $\sigma_T^o = \mathcal{M}(\mathbf{x})$. Thus, $\varsigma^2$ is a function of $\mathbf{x}$:

$$\varsigma_{ijk}^2(\mathbf{x}) = \alpha \mathcal{M}_{ijk}^2(\mathbf{x}) + \beta \mathcal{M}_{ijk}(\mathbf{x}) + \gamma. \tag{A.9}$$

This dependence on $\mathbf{x}$ changes the simplification of $\mathcal{L}(\mathbf{x})$; the $\frac{1}{2}\ln\varsigma^2$ term must be retained the minimization. Thus the objective function for maximum likelihood estimation is

$$J_{ML}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left\{ \left( \frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}(\mathbf{x})} \right)^2 + \ln \varsigma_{ijk}^2(\mathbf{x}) \right\}. \tag{A.10}$$

While both Eqs. (A.4) and (A.10) can be said to represent objective functions of maximum likelihood estimation, the constant variance assumption in Eq. (A.4) is inconsistent with the probability model $p_z(z|\mathbf{x})$. Therefore, for the duration of this paper "maximum likelihood" will refer exclusively to Eq. (A.10).

### A.1.4 The Reduced Maximum Likelihood (RML) Objective Function

While a theoretically sound objective function, in practice Eq. (A.10) presents some difficulties. Examining the scale of the terms in the summation reveals one reason. The first term is the square of a zero-mean, unit-variance Gaussian random variable, thus a $\chi^2$ random variable with one degree of freedom, which has mean value of 1. $\varsigma^2$ is on the order of $10^{-5}$, so $\ln\varsigma^2$ is on the order of -11.5, so summed over all of the measurements, the $\ln\varsigma^2$ dominates. The parameters $\alpha$, $\beta$, and $\gamma$ are only rough approximations, though, so the dominant term is not as accurate as the $\chi^2$ term.

For this reason, the final objective function analyzed here is the reduced maximum likelihood objective function:

$$J_{RML}(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left( \frac{z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})}{\varsigma_{ijk}(\mathbf{x})} \right)^2. \tag{A.11}$$

## A.2 Objective Function Gradients

As mentioned before, to be useful as an estimation tool, the objective function minima must be obtained. Many minimization routines require the calculation of the objective function gradient. Below, the gradient is analytically derived for the four cited objective functions.

### A.2.1 SE Objective Function Gradient

Evaluation of the gradient requires a straightforward application of the chain rule, differentiating with respect to each model parameter. With respect to the $p^{th}$ model parameter, the partial derivative of Eq. (A.1) is:

$$\frac{\partial}{\partial \mathbf{x}_p} J_{SE}(\mathbf{x}) = -2 \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{K_{ij}} (z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})) \frac{\partial}{\partial \mathbf{x}_p} \mathcal{M}_{ijk}(\mathbf{x}) \tag{A.12}$$

where

$$\frac{\partial}{\partial \mathbf{x}_p} \mathcal{M}_{ijk}(\mathbf{x}) = \frac{\partial \mathcal{M}(u_{ij}, v_{ij})}{\partial u_{ij}} \frac{\partial u_{ij}}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}(u_{ij}, v_{ij})}{\partial v_{ij}} \frac{\partial v_{ij}}{\partial \mathbf{x}_p}. \tag{A.13}$$

The terms $u_{ij}$ and $v_{ij}$ represent the rectangular components of the wind field at the $ij_{th}$ wvc. Note that these may be represented in terms of the $F$ matrix representing any linear wind field model:

$$u_{ij} = F_l^T \mathbf{x}, \tag{A.14}$$

$$v_{ij} = F_{l+MN}^T \mathbf{x}, \tag{A.15}$$

$$l = N(i-1) + j, \tag{A.16}$$

where $l$ is the index into the column scanned representation of the wind region and $F_l^T$ is the $l^{th}$ row of the wind field model transform matrix. Therefore,

$$\frac{\partial u_{ij}}{\partial \mathbf{X}_p} = F_{l,p} \tag{A.17}$$

$$\frac{\partial v_{ij}}{\partial \mathbf{X}_p} = F_{l+MN,p}. \tag{A.18}$$

The Geophysical Model Function is an empirically derived table of values with no closed form solution. The table has three dimensions: wind speed $U$, relative azimuth $\chi$ (instrument azimuth $\psi$ - wind direction $\phi$), and incidence angle $\theta$. In order to evaluate the function, an interpolation routine must be used. The IMSL bspline function is used, interpolating in all three directions. Through this function, partial derivatives can be easily obtained with respect to wind speed and relative azimuth, i.e. $\frac{\partial \mathcal{M}}{\partial s}$ and $\frac{\partial \mathcal{M}}{\partial \chi}$. These are related to the rectangular components by

$$s = \sqrt{u^2 + v^2}, \tag{A.19}$$

$$\chi = \psi - \phi, \tag{A.20}$$

$$\phi = \tan^{-1}(\frac{v}{u}). \tag{A.21}$$

Thus,

$$\frac{\partial \mathcal{M}}{\partial v} = \frac{\partial \mathcal{M}}{\partial s}\frac{\partial s}{\partial v} + \frac{\partial \mathcal{M}}{\partial \phi}\frac{\partial \phi}{\partial v} \tag{A.22}$$

$$\frac{\partial \mathcal{M}}{\partial u} = \frac{\partial \mathcal{M}}{\partial s}\frac{\partial s}{\partial u} + \frac{\partial \mathcal{M}}{\partial \phi}\frac{\partial \phi}{\partial u} \tag{A.23}$$

$$\frac{\partial s}{\partial u} = \frac{u}{\sqrt{u^2 + v^2}} \tag{A.24}$$

$$\frac{\partial s}{\partial v} = \frac{v}{\sqrt{u^2 + v^2}} \tag{A.25}$$

$$\frac{\partial \mathcal{M}}{\partial \phi} = \frac{\partial \mathcal{M}}{\partial \chi}\frac{\partial \chi}{\partial \phi} = -\frac{\partial \mathcal{M}}{\partial \chi}. \tag{A.26}$$

The partials of $\phi$ with respect to $u$ and $v$ require more careful attention. The inverse tangent with only one argument, defined on the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, has a well known derivative:

$$\frac{d}{dx}\tan^{-1}(x) = \frac{1}{1 + x^2}. \tag{A.27}$$

For purposes in wind retrieval, the four quadrant inverse tangent (defined on the interval $[-\pi, \pi]$, and denoted $\tan_4^{-1}$) is necessary. This can be defined in the following

way:

$$\tan_4^{-1}(\frac{v}{u}) = \begin{cases} \tan^{-1}(\frac{v}{u}) & \text{Quadrants I and IV,} \\ \tan^{-1}(\frac{v}{u}) + \pi & \text{Quadrant II,} \\ \tan^{-1}(\frac{v}{u}) - \pi & \text{Quadrant III.} \end{cases} \tag{A.28}$$

Thus the partial derivatives will be the same in all quadrants, i.e.

$$\begin{aligned} \frac{\partial \phi}{\partial u} = \frac{\partial}{\partial u} \tan_4^{-1}(\frac{v}{u}) &= \frac{1}{1+\frac{v^2}{u^2}} \frac{-v}{u^2} \\ &= \frac{-v}{u^2+v^2}. \end{aligned} \tag{A.29}$$

$$\begin{aligned} \frac{\partial \phi}{\partial v} = \frac{\partial}{\partial v} \tan_4^{-1}(\frac{v}{u}) &= \frac{1}{1+\frac{v^2}{u^2}} \frac{1}{u} \\ &= \frac{u}{u^2+v^2}. \end{aligned} \tag{A.30}$$

Therefore,

$$\frac{\partial \mathcal{M}}{\partial v} = \frac{\partial \mathcal{M}}{\partial s} \frac{v}{\sqrt{u^2+v^2}} + \frac{\partial \mathcal{M}}{\partial \phi} \frac{u}{u^2+v^2}, \tag{A.31}$$

$$\frac{\partial \mathcal{M}}{\partial u} = \frac{\partial \mathcal{M}}{\partial s} \frac{u}{\sqrt{u^2+v^2}} - \frac{\partial \mathcal{M}}{\partial \phi} \frac{v}{u^2+v^2}. \tag{A.32}$$

### A.2.2 WSE Objective Function Gradient

The WSE objective function differs from the SE objective function by only the $\varsigma$ term which is constant with respect to $\mathbf{x}$. Thus, the gradient differs from Eq. (A.12) by the same term:

$$\frac{\partial}{\partial \mathbf{x}_p} J_{WSE}(\mathbf{x}) = -2 \sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{k=1}^{K_{ij}} \left[ \frac{(z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))}{\varsigma_{ijk}^2} \right] \frac{\partial}{\partial \mathbf{x}_p} \mathcal{M}_{ijk}(\mathbf{x}). \tag{A.33}$$

### A.2.3 RML Objective Function Gradient

The RML objective function differs from the WSE objective function only in that $\varsigma^2$ depends upon $\mathbf{x}$. Computation of the gradient requires use of the derivative

quotient rule:

$$\frac{\partial}{\partial \mathbf{x}_p} J_{RML}(\mathbf{x}) = -\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left\{ \frac{2(z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial}{\partial \mathbf{x}_p}\mathcal{M}_{ijk}(\mathbf{x}_p)}{\varsigma_{ijk}^2} + \frac{(z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x}))^2 \frac{\partial \varsigma_{ijk}^2}{\partial \mathbf{x}_p}}{(\varsigma_{ijk}^2)^2} \right\}, \tag{A.34}$$

where

$$\frac{\partial \varsigma_{ijk}^2}{\partial \mathbf{x}_p} = 2\alpha \mathcal{M}_{ijk}(\mathbf{x}) \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p} + \beta \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}. \tag{A.35}$$

### A.2.4 ML Objective Function Gradient

Differentiating Eq. (A.10) requires only the addition of one term to Eq. (A.34):

$$\frac{\partial}{\partial \mathbf{x}_p} J_{ML}(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}_p} J_{RML}(\mathbf{x}) + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \frac{1}{\varsigma_{ijk}^2(\mathbf{x})} \frac{\partial \varsigma_{ijk}^2(\mathbf{x})}{\partial \mathbf{x}_p}. \tag{A.36}$$

## A.3 Objective Function Hessian Matrices

Several minimization algorithms for the objective function require a realization of the Hessian matrix, or the matrix of double partials. Although the derivation is involved, like the gradient, it is a straightforward implementation of the chain rule.

### A.3.1 SE Objective Function Hessian

To completely specify the derivation, it is sufficient to derive expressions for the following:

$$\frac{\partial^2 J_{SE}}{\partial \mathbf{x}_p^2} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left[ (z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})) \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p^2} - \left( \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p} \right)^2 \right], \tag{A.37}$$

and

$$\frac{\partial^2 J_{SE}}{\partial \mathbf{x}_p \partial \mathbf{x}_q} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}} \left[ (z_{ijk} - \mathcal{M}_{ijk}(\mathbf{x})) \frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p \partial \mathbf{x}_q} - \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p} \frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_q} \right], \tag{A.38}$$

where $p \neq q$. By the chain rule,

$$\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_p^2} &= \frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial u}\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}}{\partial v}\frac{\partial v}{\partial \mathbf{x}_p}\right) \\
&= \frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial u}\right)\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}}{\partial u}\frac{\partial^2 u}{\partial \mathbf{x}_p^2} + \frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial v}\right)\frac{\partial v}{\partial \mathbf{x}_p} + \frac{\partial \mathcal{M}}{\partial v}\frac{\partial^2 v}{\partial \mathbf{x}_p^2} \\
&= \frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial u}\right)\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial v}\right)\frac{\partial v}{\partial \mathbf{x}_p}, \qquad \text{(A.39)} \\
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_q \partial \mathbf{x}_p} &= \frac{\partial}{\partial \mathbf{x}_q}\left(\frac{\partial \mathcal{M}}{\partial u}\right)\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial}{\partial \mathbf{x}_q}\left(\frac{\partial \mathcal{M}}{\partial v}\right)\frac{\partial v}{\partial \mathbf{x}_p}. \qquad \text{(A.40)}
\end{aligned}$$

Note that the double partials, $\frac{\partial^2 v}{\partial \mathbf{x}_p^2}$ and $\frac{\partial^2 u}{\partial \mathbf{x}_p^2}$ are both 0. The mixed partials of the model function are further developed as

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial u}\right) &= \frac{\partial^2 \mathcal{M}}{\partial u^2}\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial u \partial v}\frac{\partial v}{\partial \mathbf{x}_p}, \qquad \text{(A.41)} \\
\frac{\partial}{\partial \mathbf{x}_p}\left(\frac{\partial \mathcal{M}}{\partial v}\right) &= \frac{\partial^2 \mathcal{M}}{\partial u \partial v}\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial v^2}\frac{\partial v}{\partial \mathbf{x}_p}, \qquad \text{(A.42)}
\end{aligned}$$

and thus, the above simplifies to

$$\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_p^2} &= \frac{\partial^2 \mathcal{M}}{\partial u^2}\left(\frac{\partial u}{\partial \mathbf{x}_p}\right)^2 + 2\frac{\partial^2 \mathcal{M}}{\partial u \partial v}\frac{\partial v}{\partial \mathbf{x}_p}\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial v^2}\left(\frac{\partial v}{\partial \mathbf{x}_p}\right)^2, \qquad \text{(A.43)} \\
\frac{\partial^2 \mathcal{M}}{\partial \mathbf{x}_q \partial \mathbf{x}_p} &= \frac{\partial^2 \mathcal{M}}{\partial u^2}\frac{\partial u}{\partial \mathbf{x}_q}\frac{\partial u}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial u \partial v}\frac{\partial v}{\partial \mathbf{x}_q}\frac{\partial u}{\partial \mathbf{x}_p} \\
&\quad + \frac{\partial^2 \mathcal{M}}{\partial u \partial v}\frac{\partial u}{\partial \mathbf{x}_q}\frac{\partial v}{\partial \mathbf{x}_p} + \frac{\partial^2 \mathcal{M}}{\partial v^2}\frac{\partial v}{\partial \mathbf{x}_q}\frac{\partial v}{\partial \mathbf{x}_p}. \qquad \text{(A.44)}
\end{aligned}$$

When using the bspline version of the Geophysical Model Function, the double partials $\frac{\partial^2 \mathcal{M}}{\partial u^2}, \frac{\partial^2 \mathcal{M}}{\partial v^2}$, and $\frac{\partial^2 \mathcal{M}}{\partial u \partial v}$ are not directly available, as the model function is splined with respect to $s$ and $\chi$. Thus, expressions for these partials must also be derived to

implement the Hessian matrix.

$$\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial u^2} &= \frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial u}\right) \\
&= \frac{\partial}{\partial u}\left(\frac{u}{\sqrt{u^2+v^2}}\frac{\partial \mathcal{M}}{\partial s} - \frac{v}{u^2+v^2}\frac{\partial \mathcal{M}}{\partial \phi}\right) \\
&= \frac{u}{\sqrt{u^2+v^2}}\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial s}\right) + \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad -\frac{v}{(u^2+v^2)}\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) + \frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} \qquad \text{(A.45)} \\
\frac{\partial^2 \mathcal{M}}{\partial v^2} &= \frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial v}\right) \\
&= \frac{\partial}{\partial v}\left(\frac{v}{\sqrt{u^2+v^2}}\frac{\partial \mathcal{M}}{\partial s} + \frac{u}{u^2+v^2}\frac{\partial \mathcal{M}}{\partial \phi}\right) \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial s}\right) + \frac{u^2}{(u^2+v^2)^{3/2}}\frac{\partial \mathcal{M}}{\partial s} \\
&\quad -\frac{2uv}{(u^2+v^2)^2}\frac{\partial \mathcal{M}}{\partial \phi} + \frac{u}{u^2+v^2}\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right), \qquad \text{(A.46)}
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial s}\right) &= \frac{\partial^2 \mathcal{M}}{\partial s^2}\frac{\partial s}{\partial u} + \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial \phi}{\partial u} \\
&= \frac{u}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial s^2} - \frac{v}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}, \qquad \text{(A.47)} \\
\frac{\partial}{\partial u}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) &= \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial s}{\partial u} + \frac{\partial^2 \mathcal{M}}{\partial \phi^2}\frac{\partial \phi}{\partial u} \\
&= \frac{u}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} - \frac{v}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2}, \qquad \text{(A.48)} \\
\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial s}\right) &= \frac{\partial^2 \mathcal{M}}{\partial s^2}\frac{\partial s}{\partial v} + \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial \phi}{\partial v} \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial s^2} + \frac{u}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}, \qquad \text{(A.49)} \\
\frac{\partial}{\partial v}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) &= \frac{\partial^2 \mathcal{M}}{\partial \phi \partial s}\frac{\partial s}{\partial v} + \frac{\partial^2 \mathcal{M}}{\partial \phi^2}\frac{\partial \phi}{\partial v} \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} + \frac{u}{u^2+v^2}\frac{\partial^2 \mathcal{M}}{\partial \phi^2}. \qquad \text{(A.50)}
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\partial^2\mathcal{M}}{\partial u^2} &= \frac{u^2}{u^2+v^2}\frac{\partial^2\mathcal{M}}{\partial s^2} - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} + \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial\mathcal{M}}{\partial s} \\
&\quad -\frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} + \frac{v^2}{(u^2+v^2)^2}\frac{\partial^2\mathcal{M}}{\partial\phi^2} + \frac{2uv}{(u^2+v^2)^2}\frac{\partial\mathcal{M}}{\partial\phi} \\
&= \frac{u^2}{u^2+v^2}\frac{\partial^2\mathcal{M}}{\partial s^2} + \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial\mathcal{M}}{\partial s} - \frac{2uv}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} \\
&\quad +\frac{v^2}{(u^2+v^2)^2}\frac{\partial^2\mathcal{M}}{\partial\phi^2} + \frac{2uv}{(u^2+v^2)^2}\frac{\partial\mathcal{M}}{\partial\phi}, \qquad \text{(A.51)} \\
\frac{\partial^2\mathcal{M}}{\partial v^2} &= \frac{v^2}{u^2+v^2}\frac{\partial^2\mathcal{M}}{\partial s^2} + \frac{2uv}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} + \frac{u^2}{(u^2+v^2)^{3/2}}\frac{\partial\mathcal{M}}{\partial s} \\
&\quad -\frac{2uv}{(u^2+v^2)^2}\frac{\partial\mathcal{M}}{\partial\phi} + \frac{u^2}{(u^2+v^2)^2}\frac{\partial^2\mathcal{M}}{\partial\phi^2}. \qquad \text{(A.52)}
\end{aligned}$$

The mixed partial is found to be

$$\begin{aligned}
\frac{\partial^2\mathcal{M}}{\partial u\partial v} &= \frac{\partial}{\partial u}\left(\frac{v}{\sqrt{u^2+v^2}}\frac{\partial\mathcal{M}}{\partial s} + \frac{u}{u^2+v^2}\frac{\partial\mathcal{M}}{\partial\phi}\right) \\
&= \frac{v}{\sqrt{u^2+v^2}}\frac{\partial}{\partial u}\left(\frac{\partial\mathcal{M}}{\partial s}\right) - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial\mathcal{M}}{\partial s} \\
&\quad +\frac{u}{u^2+v^2}\frac{\partial}{\partial u}\left(\frac{\partial\mathcal{M}}{\partial\phi}\right) + \frac{v^2-u^2}{(u^2+v^2)^2}\frac{\partial\mathcal{M}}{\partial\phi} \\
&= \frac{uv}{u^2+v^2}\frac{\partial^2\mathcal{M}}{\partial s^2} - \frac{v^2}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial\mathcal{M}}{\partial s} \\
&\quad -\frac{2u^2}{(u^2+v^2)^2}\frac{\partial\mathcal{M}}{\partial\phi} + \frac{u^2}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} - \frac{uv}{(u^2+v^2)^2}\frac{\partial^2\mathcal{M}}{\partial\phi^2} \\
&= \frac{uv}{u^2+v^2}\frac{\partial^2\mathcal{M}}{\partial s^2} + \frac{u^2-v^2}{(u^2+v^2)^{3/2}}\frac{\partial^2\mathcal{M}}{\partial\phi\partial s} - \frac{uv}{(u^2+v^2)^{3/2}}\frac{\partial\mathcal{M}}{\partial s} \\
&\quad +\frac{v^2-u^2}{(u^2+v^2)^2}\frac{\partial\mathcal{M}}{\partial\phi} - \frac{uv}{(u^2+v^2)^2}\frac{\partial^2\mathcal{M}}{\partial\phi^2}. \qquad \text{(A.53)}
\end{aligned}$$

Obtaining $\frac{\partial^2\mathcal{M}}{\partial\phi^2}$ and $\frac{\partial^2\mathcal{M}}{\partial\phi\partial s}$ from the model function requires only a simple application of the chain rule. Recall that

$$\begin{aligned}
\chi &= \psi - \phi, \qquad \text{(A.54)} \\
\frac{\partial\mathcal{M}}{\partial\phi} &= -\frac{\partial\mathcal{M}}{\partial\chi}. \qquad \text{(A.55)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\partial^2 \mathcal{M}}{\partial \phi^2} &= \frac{\partial}{\partial \phi}\left(\frac{\partial \mathcal{M}}{\partial \phi}\right) \\
&= \frac{\partial}{\partial \phi}\left(-\frac{\partial \mathcal{M}}{\partial \chi}\right) \\
&= -\frac{\partial^2 \mathcal{M}}{\partial \chi^2}\frac{\partial \chi}{\partial \phi} \\
&= \frac{\partial^2 \mathcal{M}}{\partial \chi^2}, \qquad \text{(A.56)} \\
\frac{\partial^2 \mathcal{M}}{\partial \phi \partial s} &= \frac{\partial}{\partial \phi}\left(\frac{\partial \mathcal{M}}{\partial s}\right) \\
&= \frac{\partial^2 \mathcal{M}}{\partial \chi \partial s}\frac{\partial \chi}{\partial \phi} \\
&= -\frac{\partial^2 \mathcal{M}}{\partial \chi \partial s}. \qquad \text{(A.57)}
\end{aligned}$$

### A.3.2 The WSE Hessian Matrix

As with the gradient of $J_{WSE}$, the Hessian requires only the addition of the constant $\varsigma^2$ term to the $J_{SE}$ Hessian matrix

(A.58)

$$\frac{\partial^2 J_{WSE}}{\partial \mathbf{x}_p^2} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left[\frac{(z_{ijk}-\mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p^2}}{\varsigma_{ijk}^2} - \left(\frac{\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}}{\varsigma_{ijk}^2}\right)^2\right],$$

(A.59)

$$\frac{\partial^2 J_{WSE}}{\partial \mathbf{x}_p \partial \mathbf{x}_q} = -2\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K_{ij}}\left[\frac{(z_{ijk}-\mathcal{M}_{ijk}(\mathbf{x}))\frac{\partial^2 \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p \partial \mathbf{x}_q}}{\varsigma_{ijk}^2} - \frac{\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_p}\frac{\partial \mathcal{M}_{ijk}(\mathbf{x})}{\partial \mathbf{x}_q}}{\varsigma_{ijk}^2}\right].$$

## Bibliography


[1] David G. Long, "Wind field model-based estimation of Seasat scatterometer winds," *Journal of Geophysical Research*, vol. 98, no. C8, pp. 14651–14668, August 1993.

[2] David G. Long and Jerry 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, May 1990.

[3] Stephen L. Richards, "A field-wise wind retrieval algorithm for SeaWinds," M.S. thesis, Brigham Young University, August 1999.

[4] Fawwaz T. Ulaby, Richard K. Moore, and Adrian K. Fung, *Microwave Remote Sensing*, vol. 2, Artech House Inc., Norwood, MA, 1981.

[5] Firouz M. Naderi, Michael H. Freilich, and David G. 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, June 1991.

[6] Travis E. Oliphant and David G. Long, "Accuracy of scatterometer-derived winds using the Cramer-Rao bound," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 6, November 1999.

[7] Chong-Yung Chi and Fuk K. Li, "A comparative study of several wind estimation algorithms for spaceborne scatterometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 2, 1988.

[8] Paul E. Johnson, *Uncertainties in Oceanic Microwave Remote Sensing: The Radar Footprint, The Wind-Backscatter Relationship, and the Measurement Probability Density Function*, Ph.D. thesis, Brigham Young University, 1999.

[9] Travis E. Oliphant, "New techniques for wind scatterometry," M.S. thesis, Brigham Young University, August 1996.

[10] Anil K. Jain, *Fundamental of Digital Image Processing*, Prentice Hall, New Jersey, 1989.

[11] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical Recipies in Fortran*, Cambridge University Press, 1992.

[12] Todd K. Moon and Wynn C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice Hall, 2000.

[13] David W. Draper, "Spatial frequency analysis and truncation issues for the QSCAT KL model," BYU MERS Internal Report, 2000.

[14] Charles G. Brown, "A field-wise wind retrieval algorithm for the NASA scatterometer," M.S. thesis, Brigham Young University, September 1998.

[15] Michael Zatman and Steven T. Smith, "Resolution and ambiguity bounds for interferometric-like systems," in *Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems, and Computers*, 1998, vol. 2, pp. 1466–1470.

[16] Amy E. Gonzales and David G. Long, "An assessment of NSCAT ambiguity removal," *Journal of Geophysical Research*, vol. 104, no. C5, pp. 11449–11457, May 15 1999.

[17] Amy E. Gonzales, "An assessment of the NASA scatterometer ambiguity removal technique," M.S. thesis, Brigham Young University, May 1998.