2003-05-14

Studies to Improve Estimation of the Electromagnetic Bias in Radar Altimetry

Justin DeWitt Smith
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Electrical and Computer Engineering Commons

BYU ScholarsArchive Citation
Smith, Justin DeWitt, "Studies to Improve Estimation of the Electromagnetic Bias in Radar Altimetry" (2003). All Theses and Dissertations. 75.
https://scholarsarchive.byu.edu/etd/75

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
STUDIES TO IMPROVE ESTIMATION OF THE ELECTROMAGNETIC BIAS IN RADAR ALTIMETRY

by

Justin DeWitt Smith

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 1999
This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

David V. Arnold, Chair

Date

Michael A. Jensen

Date

David G. Long
As chair of the candidate’s graduate committee, I have read the thesis of Justin DeWitt Smith 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

David V. Arnold
Chair, Graduate Committee

Accepted for the Department

Lee Swindlehurst
Graduate Coordinator

Accepted for the College

Douglas M. Chabries
Dean, College of Engineering and Technology
ABSTRACT

STUDIES TO IMPROVE ESTIMATION OF THE ELECTROMAGNETIC BIAS IN RADAR ALTIMETRY

Justin DeWitt Smith
Department of Electrical and Computer Engineering
Master of Science

In May of 2000 Jason-1, a joint project between NASA and the French space agency CNES, will be launched. Its mission is to continue the highly successful gathering of data which TOPEX/Poseidon has collected since August of 1992. The main goal of Jason-1 is to achieve higher accuracy in measuring the mean sea level (MSL). In order to do so, the electromagnetic (EM) bias must be estimated more accurately because it is the largest contributing error. This thesis presents two different studies which add to the knowledge and improve estimation of the EM bias, and thus assists Jason-1 in achieving its primary goal. Oceanographic data collected from two different experiments are analyzed; one in the Gulf of Mexico (GME) and the other in Bass Strait, Australia (BSE).

The first study is a spatial analysis of the backscattered power versus the phase of the wave. Its purpose is to determine why the normalized EM bias stops increasing and levels out at high wind speeds (about 11 m/s) and then decreases at higher wind speeds. Two possible causes are investigated. First, it could be due to
a shift in the backscatter power modulation to the forward or rear face of the wave crests. Second, it may be due to the backscatter power becoming more homogeneous throughout the wave profile. This study is novel because it uses the knowledge of the spatial distribution of both the backscatter and wave displacement for the study of the EM bias. Both contribute to the EM bias decrease, but the latter cause seems to be the dominant effect. This study is performed on GME data.

The second study uses two different nonparametric regression (NPR) techniques to estimate the EM bias. A recent study of satellite data from the TOPEX/Poseidon altimeter supports that the bias is modeled better using NPR regression. A traditional parametric fit is compared to two NPR techniques with GME data. The parametric fit is a variation of NASA’s equation used to estimate EM bias for their Geophysical Data Records (GDRs). The two NPR techniques used are the Nadaraya-Watson Regression (NWR) and Local Linear Regression (LLR) estimators. Two smoothing kernel functions are used with each NPR technique, namely the Gaussian and the Epanechnikov kernels. NPR methods essentially consist of statistically smoothing the measured EM bias over a gridded two-dimensional plane. Satellite and tower EM bias estimates are compared in the wind and significant wave height plane. Another recent study has shown that wave slope is strongly correlated to EM bias. With this knowledge, EM bias is estimated over several two-dimensional planes which include wave slope in attempt to reduce the residual bias. This portion of the study is performed on GME and BSE data. It is shown that a combination of slope, significant wave height, and wind speed used in conjunction with these NPR methods produces the best EM bias estimate for tower data.
ACKNOWLEDGMENTS

I would like to thank my parents who have always inspired me to do my best and to do what I like. I would also like to thank all of the professors who made learning fun for me. I am grateful for the friends which I have made in the Microwave Earth Remote Sensing (MERS) Lab. Finally, I would like to thank Dr. David Arnold for giving me the opportunity to perform this research.
Contents

Acknowledgments vii

List of Tables xiii

List of Figures xv

1 Introduction 1
   1.1 Description of Problems ........................................ 2
   1.2 Contributions .................................................... 2
   1.3 Overview ....................................................... 3

2 Background 5
   2.1 Electromagnetic Bias ............................................. 5
      2.1.1 Relationship of EM Bias to Tower Data ..................... 5
      2.1.2 Relationship of EM Bias to Satellite Altimetry Data .... 7
      2.1.3 Additional Information about EM Bias ..................... 9
   2.2 Experimental Data Sets ......................................... 10
      2.2.1 Gulf of Mexico Experiment and Data Set .................. 10
      2.2.2 Bass Strait Data Set .................................... 19
      2.2.3 TOPEX/Poseidon Data Set ................................ 20

3 Spatial Wave Analysis of the Electromagnetic Bias 23
   3.1 Possible Causes for EM bias decrease at high wind speeds .... 23
   3.2 Method and Results ............................................. 25
   3.3 Summary .................................................... 35
4 Traditional EM Bias Estimation
  4.1 Parametric Fitting to GME and BSE EM Bias .......................... 41
  4.2 Summary ........................................................................ 43

5 Nonparametric Estimation of Electromagnetic Bias .................. 49
  5.1 Limitations of Gridding without Smoothing ......................... 51
  5.2 Two-Dimensional Nonparametric Regression ....................... 52
    5.2.1 Nadaraya-Watson Regression Estimator ...................... 53
    5.2.2 Local Linear Regression Estimator .......................... 54
    5.2.3 Kernel Functions ............................................... 55
    5.2.4 Bandwidth Selection ......................................... 57
  5.3 Results ........................................................................... 58
    5.3.1 NWR EM Bias Estimation ................................. 58
    5.3.2 LLR EM Bias Estimation ............................... 59
    5.3.3 NWR Implementation Differences for Tower versus Satellite Data 68
    5.3.4 Comparison of Results to TOPEX/Poseidon Data ........ 69
    5.3.5 Residual Biases ............................................... 75
  5.4 Summary ........................................................................ 78

6 Nonparametric EM Bias Estimation with Wave Slope ............. 81
  6.1 Wave Slope ................................................................. 81
  6.2 Results of using Wave Slope with NPR Bias Estimates .......... 83
    6.2.1 Estimation with Wave Slope and Significant Wave Height . 86
    6.2.2 Estimation with Wave Slope and Wind Speed ............. 86
    6.2.3 Estimation with Wave Slope, Wind Speed and Significant Wave Height ............................................. 87
  6.3 Summary ........................................................................ 123

7 Conclusions ....................................................................... 129
  7.1 General Summary ......................................................... 129
  7.2 Contributions .............................................................. 130
7.3 Future Work .................................................. 131

A Two-dimensional Nonparametric Algorithm 133

Bibliography 141
List of Tables

2.1 Statistics of the available C- and Ku-band data for Gulf of Mexico Experiment. .................................................. 13
2.2 Statistics of the available C- and Ku-band data for the Bass Strait Experiment. .................................................. 19
4.1 Coefficients for Eq. (4.1) to estimate EM bias for Gulf of Mexico and Bass Strait Experiments .................................................. 42
5.1 Bandwidth Values for Gulf of Mexico and Bass Strait Experiments. .................................................. 58
5.2 Statistics of the residual biases and the correlation coefficient between measured and estimated EM bias for the Gulf of Mexico and Bass Straits Experiments. Includes the improvement of NPR methods over the parametric fit. .................................................. 78
6.1 Statistics of the residual biases for the Gulf of Mexico and Bass Straits Experiments after removing $s,H$ NPR estimates. ............... 99
6.2 Statistics of the residual biases for the Gulf of Mexico and Bass Straits Experiments after removing $s,U$ NPR estimates. ............... 111
6.3 Statistics of the residual biases for the Gulf of Mexico and Bass Straits Experiments after removing $s,U$ NPR estimates. ............... 122
6.4 Summary Comparison of NPR Estimates for the Gulf of Mexico Experiment. .................................................. 124
6.5 Summary Comparison of NPR Estimates for Bass Straits Experiment. .................................................. 125
6.6 LLR-Gaussian Results for GME with Varying Wave Slope Calculations .................................................. 127
List of Figures

2.1 Plot of normalized $\sigma^0$ versus wave displacement. MSL is assumed to be at 0 m. .................................................. 6
2.2 Representation of the EM bias shift in the backscatter received by a satellite. .................................................. 8
2.3 C-band EM bias dependence on $U$ and $H_{1/3}$ for GME. ............... 11
2.4 Ku-band EM bias dependence on $U$ and $H_{1/3}$ for GME. ............... 12
2.5 Geometry of the tower used in GME. The nadir facing scatterometers are located at the center of the figure. ...................... 14
2.6 Picture of platform B in Fig. 2.5. This picture was taken from platform C facing southwest. .................................. 15
2.7 Picture of platform C in Fig. 2.5. This picture was taken from platform B facing northeast. .................................. 15
2.8 Wind Speed as a function of time for the GME data set. .................. 16
2.9 Significant wave height, $H_{1/3}$, as a function of time for the GME data set. 17
2.10 EM bias calculated from Ku-band scatterometer as a function of time for the GME data set. ............................... 18
2.11 EM bias estimation for TOPEX/Poseidon data using the Nadaraya-Watson Regression (NWR) estimator. The isolines enclose 95% of the global data set. ........................................ 21
2.12 EM bias estimation for TOPEX/Poseidon data using the Local Linear Regression (LLR) estimator. The isolines enclose 95% of the global data set. ........................................ 22
3.1 C and Ku Band Normalized Bias versus Wind Speed. ................. 24
3.2 Example Wave Phase Plot. ........................................ 25
3.3 Thirty second wave displacement example. The asterisks (*) correspond to local crests and troughs found by use of the algorithm described in the text. ............................... 26
3.4 Typical wave phase versus $\bar{\sigma}$ plot for C and Ku-band averaged for 10 minutes of data. Average wind speed=4.2 m/s. The y-axis is average normalized power, $\bar{\sigma}$. ................................................. 28
3.5 Normalized C-band centroid $\bar{\sigma}_C$ plotted versus wind speed. .... 30
3.6 Normalized Ku-band centroid $\bar{\sigma}_{Ku}$ plotted versus wind speed. .... 31
3.7 Time plots of wind speed, wind direction, C-band centroids, Ku-band centroids. .......................................................... 32
3.8 Average of Ku-band $\phi$ versus $\bar{\sigma}$ plots for February. The solid line includes all data for the entire month and the dotted line is for all data with a $3.9^\circ \pm 0.5^\circ$ centroid shift. The y-axis is average normalized power, $\bar{\sigma}$. ......................................................... 33
3.9 Plot of Ku-band $\bar{\sigma}$ trough/$\bar{\sigma}$ crest versus wind speed. Note that the plot never goes below 1, indicating that the power reflected at the troughs is always greater than at the crests. ................................. 34
3.10 Normalized C-band centroid $\bar{\sigma}_C$ plotted versus wind direction. .... 36
3.11 Normalized Ku-band centroid $\bar{\sigma}_{Ku}$ plotted versus wind direction. ... 37
3.12 Areas of assumed wind and tower interference inferred from Figs. 3.10 and 3.11. .......................................................... 38
3.13 New normalized C-band centroid $\bar{\sigma}_C$ plotted versus wind direction. .. 39
3.14 New normalized Ku-band centroid $\bar{\sigma}_{Ku}$ plotted versus wind direction. 39
4.1 GME - C-band Parametric EM bias fit in centimeters with respect to $U$ and $H^{\frac{1}{3}}$. ......................................................... 44
4.2 GME - Ku-band Parametric EM bias fit in centimeters with respect to $U$ and $H^{\frac{1}{3}}$. ......................................................... 45
4.3 Cropped GME residual Ku-band bias in centimeters. ......................... 45
4.4 C-band Residual bias for GME after removing parametric fit. .............. 46
4.5 Ku-Band Residual bias for GME after removing parametric fit. ............ 47
4.6 BSE - Ku-band Parametric EM bias fit in centimeters with respect to $U$ and $H^{1/3}$. ......................................................... 48
4.7 Cropped BSE residual Ku-band bias in centimeters. ................. 48
5.1 Data densities in the wind speed and significant wave height plane.
Includes hourly average values of the data from the Gulf of Mexico
Experiment (above) and the Bass Strait Experiment (below). ........... 50
5.2 Averaged bias values over a grid of $dU = dH = 0.25$ for GME. Note
the contour pattern that emerges. Data used is from GME. ............ 51
5.3 Comparison of the Gaussian and Epanechnikov kernel functions with
$h_U = 2$ and $h_H = 0.32$. Both are centered at $(U, H) = (6, 1.5)$. ......... 56
5.4 GME: NWR-Gaussian bias contours in $U, H$ plane. ................. 60
5.5 GME: NWR-Epanechnikov bias contours in $U, H$ plane. .......... 61
5.6 GME: LLR-Gaussian bias contours in $U, H$ plane. ................ 62
5.7 GME: LLR-Epanechnikov bias contours in $U, H$ plane. .......... 63
5.8 BSE: NWR-Gaussian bias contours in $U, H$ plane. ................. 64
5.9 BSE: NWR-Epanechnikov bias contours in $U, H$ plane. .......... 65
5.10 BSE: LLR-Gaussian bias contours in $U, H$ plane. ............... 66
5.11 BSE: LLR-Epanechnikov bias contours in $U, H$ plane. .......... 67
5.12 TOPEX/Poseidon: NWR*-Gauss. bias contours in $U, H$ plane. .... 70
5.13 TOPEX/Poseidon: LLR-Epan. bias contours in $U, H$ plane. ...... 71
5.14 The EM bias as a function of footprint size. The surface roughness
parameter is varied $\lambda = 10, 30$, and 100. ............................. 72
5.15 Difference plot of $|\epsilon_{TOPEX}| - |\epsilon_{GME}|$ for LLR estimate. ....... 73
5.16 Difference plot of $|\epsilon_{TOPEX}| - |\epsilon_{BSE}|$ for LLR estimate. ....... 74
5.17 Residual biases for GME data. Subplots $a - d$ correspond to the re-
moval of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and
LLR-Epan. in the $U, H$ plane. ............................................. 76
5.18 Residual biases for BSE data. Subplots $a - d$ correspond to the removal
of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-
Epan. in the $U, H$ plane. ............................................. 77
6.1 Wave displacement frequency spectrum, $\Phi(f)$, and Slope spectrum, $S(f)$ for one-hour of wave displacement data. The average $U$ and $H_\frac{1}{3}$ are 2.2 m/s and 1.5 m respectively. .......................... 82
6.2 EM bias dependence on rms wave slope. .......................... 84
6.3 Normalized EM bias dependence on rms wave slope.  .......... 85
6.4 Data densities in the significant wave height and wave slope plane. Includes hourly average values of the data from the Gulf of Mexico Experiment (above) and the Bass Strait Experiment (below). ...... 88
6.5 GME: NWR-Gaussian bias contours in $s, H$ plane. ............ 89
6.6 GME: NWR-Epanechnikov bias contours in $s, H$ plane. ..... 90
6.7 GME: LLR-Gaussian bias contours in $s, H$ plane. .......... 91
6.8 GME: LLR-Epanechnikov bias contours in $s, H$ plane. .... 92
6.9 BSE: NWR-Gaussian bias contours in $s, H$ plane. ........... 93
6.10 BSE: NWR-Epanechnikov bias contours in $s, H$ plane. ...... 94
6.11 BSE: LLR-Gaussian bias contours in $s, H$ plane. .......... 95
6.12 BSE: LLR-Epanechnikov bias contours in $s, H$ plane. .... 96
6.13 Residual biases after removing $s, H$ estimates for GME data. ...... 97
6.14 Residual biases after removing $s, H$ estimates for BSE data. ...... 98
6.15 Data densities in the wind speed and wave slope plane. Includes hourly average values of the data from the Gulf of Mexico Experiment (above) and the Bass Strait Experiment (below). ............ 100
6.16 GME: NWR-Gaussian bias contours in $s, U$ plane. .......... 101
6.17 GME: NWR-Epanechnikov bias contours in $s, U$ plane. ..... 102
6.18 GME: LLR-Gaussian bias contours in $s, U$ plane. .......... 103
6.19 GME: LLR-Epanechnikov bias contours in $s, U$ plane. .... 104
6.20 BSE: NWR-Gaussian bias contours in $s, U$ plane. ........... 105
6.21 BSE: NWR-Epanechnikov bias contours in $s, U$ plane. ...... 106
6.22 BSE: LLR-Gaussian bias contours in $s, U$ plane. .......... 107
6.23 BSE: LLR-Epanechnikov bias contours in $s, U$ plane. ....... 108
6.24 Residual biases after removing $s, U$ estimates for GME data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan. 109

6.25 Residual biases after removing $s, U$ estimates for BSE data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan. 110

6.26 GME: NWR-Gaussian normalized bias contours in $s, U$ plane. 112

6.27 GME: NWR-Epanechnikov normalized bias contours in $s, H$ plane. 113

6.28 GME: LLR-Gaussian normalized bias contours in $s, H$ plane. 114

6.29 GME: LLR-Epanechnikov normalized bias contours in $s, H$ plane. 115

6.30 BSE: NWR-Gaussian normalized bias contours in $s, H$ plane. 116

6.31 BSE: NWR-Epanechnikov normalized bias contours in $s, H$ plane. 117

6.32 BSE: LLR-Gaussian normalized bias contours in $s, H$ plane. 118

6.33 BSE: LLR-Epanechnikov normalized bias contours in $s, H$ plane. 119

6.34 Residual biases after converting normalized bias to EM bias and removing the $s, U$ estimates for GME data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan. 120

6.35 Residual biases after removing $s, U$ estimates for BSE data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan. 121

6.36 Plot of Improvement in LLR-Gaussian as a function of the upper cutoff frequency used to compute wave slope. 126
Chapter 1

Introduction

In 1979, NASA and Jet Propulsion Laboratory (JPL) began planning TOPEX, an experiment which would use a satellite altimeter to measure the height of the world’s oceans [1]. The French space agency Centre National d’Etudes Spatiales (CNES) was planning a similar oceanographic mission at the same time called Poseidon [2]. In the early 1980’s, they pooled their resources to form a single mission called TOPEX/Poseidon. TOPEX/Poseidon has surpassed its expectations and Jason-1, the first of a projected series of satellites to succeed TOPEX/Poseidon, is expected to do the same [3]. Jason-1 is scheduled to be launched in May of 2000 [1].

The goal of Jason-1 is to obtain an RMS sea surface height estimate within 2.5 cm of the actual height. The current level of accuracy achieved with TOPEX/Poseidon is within 4.3 cm [3]. The improved accuracy will allow oceanographers and climate scientists to improve models which predict long term weather patterns and ocean circulation. This type of knowledge has already allowed communities along the western coast of the Americas to prepare for the effects of El Niño and has saved many lives and millions of dollars [1],[2]. In order for Jason-1 to accurately estimate the sea surface height, altimetry and orbital errors must be minimized. Currently, the largest of these errors is the electromagnetic (EM) bias [4] which results from wave troughs reflecting more power than wave crests. As a result mean sea level estimates are negatively biased toward the wave troughs. The EM bias, ϵ, and normalized EM bias, β, are further explained in Chapter 2.
1.1 Description of Problems

The magnitude of normalized EM bias increases with increasing wind speed. However, evidence has been found in both tower and satellite data that it begins to level off at high wind speeds. Arnold and Melville et al. showed that for Ku-band, the bias measured from a tower levels off and then decreases at wind speeds above 10 m/s [5], [6]. Hevizi et al. observed this behavior in satellite data near a wind speed of 12 m/s [7].

The first study presented in this work is an attempt to explain this phenomenon. It is expected that the cause is due to changes in the short wave modulation over the phase of the wave. Two possible short wave modulation causes are investigated using the Gulf of Mexico data set, and a conclusion is drawn as to which has the strongest effect on causing the EM bias to level off at high wind speeds.

Methods for bias estimation have usually modeled the bias as a function in terms of wind speed, $U$, and significant wave height, $H_{\text{sw}}^{1/3}$. Residual biases after removing these estimates have left considerable room for improvement. Because of this, a new approach for EM bias estimation recently suggested by Gaspar is nonparametric regression (NPR) [8], [9].

The second study, and main focus of this work, is to evaluate the improvement of bias estimation using two NPR estimation techniques: Nadaraya-Watson Regression (NWR) and Local Linear Regression (LLR). Melville provided strong evidence that the EM bias can also be modeled accurately using wave slope [10]. Of key interest is the NPR bias estimation improvement using wave slope as one of the input parameters. Several comparisons are presented to compare estimates of the EM bias.

1.2 Contributions

This is the first application of NPR bias estimation data collected from a tower on the ocean. While these methods have been performed on satellite data by Gaspar, this is the first NPR comparison of tower and satellite EM bias estimations. The use of wave slope to improve EM bias estimation is also new. Using wave slope as
an input to NPR bias estimation has not been performed previously. The knowledge which is gained from this new data analysis is the major contribution of this thesis.

Another significant contribution is the use of a spatial wave analysis technique to answer why the bias levels off at high wind speeds. Until now, this type of analysis has not been applied to a study of the electromagnetic bias.

1.3 Overview

In this thesis, methods for improving the estimation of the electromagnetic bias in radar altimetry are presented. Data sets from the Gulf of Mexico [1990] and the Bass Straits [1993] experiments are used. The data is analyzed to provide ways to achieve the estimation accuracy required for the Jason-1 altimeter in May 2000.

Chapter 2 presents the Gulf of Mexico and Bass Strait tower data sets which are used throughout the rest of the thesis. Some TOPEX/Poseidon data is also given for later comparisons.

Chapter 3 presents a relationship between EM bias and short wave modulation. This is done by performing a spatial analysis of the distribution of reflected power versus wave phase for the Gulf of Mexico Data. The steps to perform this method are explained, and the results answer why the bias rolls off at high wind speeds.

Chapters 4-6 are dedicated to estimation of the EM bias using several techniques. Chapter 4 discusses the use of a traditional parametric estimation of the bias for comparison purposes. A technique similar to the one used to derive NASA’s Geophysical Data Records (GDRs) is used. It is a parametric fit to EM bias using wind speed, $U$, and significant wave height, $H_{\frac{1}{3}}$. Then in Chapter 5, the nonparametric techniques are presented and applied to tower data with the same variables. These results are then compared to similar results for satellite data. Chapter 6 discusses the wave slope parameter and shows the improvement factor of using this parameter in combination with $U$ and $H_{\frac{1}{3}}$.

Chapter 7 concludes this thesis and summarizes the results. The main contributions are repeated in light of the results presented throughout the thesis. Future work in EM bias estimation using NPR estimation techniques is then suggested.
Chapter 2

Background

In this chapter a description of the electromagnetic (EM) bias is given, and then the experimental data sets from three experiments are presented. The primary data set is from the Gulf of Mexico Experiment, herein referred to as the GME data set. Other data sets which are briefly discussed in this chapter are from the Bass Strait Experiment (BSE), and the TOPEX/Poseidon satellite.

2.1 Electromagnetic Bias

The EM bias is defined as the difference between the mean reflective surface and the true mean sea level (MSL)\cite{5}. Due to the wave troughs reflecting more power than wave crests, MSL estimates are biased toward the troughs. Figure 2.1 shows a typical plot of relative backscatter versus wave displacement. It shows that the maximum amount of backscatter is received at negative displacement values (troughs), and the minimum occurs at the highest displacement values (crests). This roughly linear distribution of power versus displacement causes the bias to occur.

2.1.1 Relationship of EM Bias to Tower Data

For tower based data the EM bias is estimated as the power-weighted mean, or centroid, of wave displacement, $\eta$,

$$\epsilon = \frac{\sum \sigma^0 \eta}{\sum \sigma^0} \quad (2.1)$$

where $\sigma^0$ is the backscattered power measurement. This equation assumes that crests and troughs have positive and negative displacement values respectively, that true
Figure 2.1: Plot of normalized $\sigma^o$ versus wave displacement. MSL is assumed to be at 0 m.

MSL lies at 0 m, and that $\sigma^o$ and $\eta$ are sampled at the same rate. The displacement data, $\eta$, can be obtained in several ways from a tower. Two methods used in practice are with the use of a wire wave gauge on the surface, and through calculation of the Doppler of the backscattered power. If the power were uniformly reflected from each part of the wave, this equation would give 0 m, the true mean sea level. Since more power is reflected from wave troughs (see Fig. 2.1), the estimated MSL is biased below the true mean. The resulting $\epsilon$ is a direct measurement of the EM bias.
2.1.2 Relationship of EM Bias to Satellite Altimetry Data

A satellite is obviously much further from the surface than tower-based equipment, and estimating MSL is a more involved process. Direct estimates cannot be taken of the EM bias (as with the tower experiment), but it is still included in the MSL estimate and must be removed. The continuous equivalent for the tower EM bias equation (2.1) is

\[ \epsilon = \frac{\int_{-\infty}^{\infty} \eta \sigma^2(\eta)p(\eta) \, d\eta}{\int_{-\infty}^{\infty} \sigma^2(\eta)p(\eta) \, d\eta} \]  

where \( p(\eta) \) is the surface height density function. This equation also represents the centroid of the function \( \sigma^2(\eta)p(\eta) \).

The following illustrates how the EM bias is manifest in the backscatter received by a space-born altimeter. Brown states that the returned impulse response from the ocean’s rough surface is given by the convolution of the backscatter density function and the average flat surface impulse response [11]:

\[ P_I(t) = q(z) * P_{FS}(t) \]  
\[ P_I(\tau) = \frac{c}{2} \int_{-\infty}^{\infty} q\left(\frac{c}{2}(\tau - \hat{\tau})\right) P_{FS}(\hat{\tau}) \, d\hat{\tau}. \]

where \( z \) has been changed to its time equivalent \( \frac{c}{2} \hat{\tau} \), \( \tau = t - \frac{2}{c} h \) so that the returned impulse is centered at \( \tau = 0 \), and \( h \) is the height of the altimeter above the ocean surface. Brown calls \( q(z) \) the ‘specular point’ or wave height density function but it will be shown that the backscatter density is more appropriate in order to include the effect of the EM bias. For this analysis, \( P_{FS} \) will be assumed to be a unit step function, \( u(\tau) \), which is equivalent to assuming that the altimeter sends out a uniform plane wave. The reader is referred to [11] for more information about \( P_{FS} \). Brown argues that the equivalent width of the specular point density is small relative to the time scale over which \( P_{FS} \) exhibits appreciable variation, allowing \( P_I \) to be expressed as

\[ P_I(\tau) \approx \frac{c P_{FS}(\tau)}{2} \int_{0}^{\infty} q\left(\frac{c}{2}(\tau - \hat{\tau})\right) \, d\hat{\tau} \]

\[ P_I(\tau) \approx \frac{c}{2} \int_{0}^{\infty} q\left(\frac{c}{2}(\tau - \hat{\tau})\right) \, d\hat{\tau} \, \tau \geq 0. \]
Brown makes the assumption that $q(z)$ is a Gaussian which leads to an impulse response (and cumulative density) of

\[ P_I(\tau) \approx \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{c\tau}{2\sqrt{2} \sigma} \right) \right], \quad \tau \geq 0, \tag{2.8} \]
where $\sigma_\eta$ is the standard deviation of the surface height, and $\text{erf}(\cdot)$ denotes the error function. The problem with the purely Gaussian assumption is that it neglects the effect of EM bias. The true backscatter density is $q(z) = \sigma^2(\eta) p(\eta)$ where $p(\eta)$ is assumed to be Gaussian and $\sigma^2(\eta)$ represents the relative amount of power received at different surface heights. The function $\sigma^2(\eta)$ is odd about the origin and weights the Gaussian such that power reflected by wave troughs is more likely to return to the satellite. As indicated in Eq. (2.2) this causes the centroid of $q(z)$ to shift an amount equal to the EM bias. Since the impulse response is the cumulative density function, its rising edge shifts in time by $2\epsilon/c$. This time shift is shown in Fig. 2.2. The solid line represents the Gaussian density and the dotted line represents a Gaussian multiplied by the linear relationship

$$\frac{\sigma^2(\eta)}{\sigma^2} = \sigma^2(\tau) = m \frac{c\tau}{2\sigma_\eta} + 1. \tag{2.9}$$

where the slope $m$ is equal to $\epsilon/\sigma_\eta$. This linear relationship is derived from the data shown in Fig. 2.1. In this example $\sigma_\eta = .3$, and $m = .17$. Multiplying by the Gaussian and solving for the new impulse response yields

$$P_I(\tau) = -\frac{m}{\sqrt{2\pi}} e^{-\frac{x^2}{2(2\sigma_\eta/c)^2}} + \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{c\tau}{2\sqrt{2\sigma_\eta/c}} \right) \right]. \tag{2.10}$$

This equation is the dotted line in the bottom plot of Fig. 2.2. The shift is equal to the time it takes for the impulse to travel the extra distance to the shifted centroid. As stated previously, the new centroid is equal to the EM bias. Since the shifted Gaussian density function results in a shifted rising edge of the impulse response, satellite measurements of the MSL include the effect of EM bias.

### 2.1.3 Additional Information about EM Bias

The EM bias was first observed by Yaplee et al. [12], and since has been studied by several investigators. The reader is referred to the references [5]-[18] for further information regarding past EM bias studies. Typical values of EM bias measured from a tower range from 1 to 10 cm. Often normalized bias, which is defined as a
percentage of significant wave height,

\[ \beta = \frac{\epsilon}{H^{1/3}}(\%) \]  \hspace{1cm} (2.11)

is used in analysis because of the strong linear dependence of EM bias on \( H^{1/3} \) (see Figs. 2.3 and 2.4) and because the resulting bias is dimensionless [5]. Typical values of normalized bias lie between 1% and 5% of \( H^{1/3} \). Both \( \epsilon \) and \( \beta \) are used in this thesis.

2.2 Experimental Data Sets

The following information for GME and BSE will be used in later chapters and provides sufficient background information about the processed data sets. The processed data sets for TOPEX/Poseidon are used for comparison purposes and were provided by Phillip Gaspar, the head of the Space Oceanography Division of Collecte Localisation Satellites (CLS) in France. One of the data sets was recently published in the Journal of Geophysical Research [8], and the other has yet to be published [9].

2.2.1 Gulf of Mexico Experiment and Data Set

Arnold et al. conducted an experiment from a Shell Offshore oil production platform on the Gulf of Mexico for a six month period from December 1, 1989 to May 31, 1990 [5] [6]. Direct measurements of several oceanic parameters were taken during this period from nadir looking C- and Ku-band scatterometers mounted 18 meters above the mean sea level, corresponding to a 1.6 m footprint size. Figure 2.5 shows where the radar system was located between two of the major oil platforms. Figures 2.6 and 2.7 show pictures of the platforms which were located northeast and southwest of the scatterometers. The raw data collected from the scatterometers was saved in ten-minute records on floptical disk at a sample rate of 8 Hz, while other direct measurements taken by the anemometer, like wind speed and direction, were saved once every 10 minutes.

Raw data and processed GME data were reanalyzed for this thesis. Raw data from this experiment was first extracted from the floptical disks to which they were
Figure 2.3: C-band EM bias dependence on $U$ and $H^1_3$ for GME.
Figure 2.4: Ku-band EM bias dependence on $U$ and $H_s^4$ for GME.
originally saved. Raw data files 13157-13647, 15711-16597, and 17600-18093 were corrupted and therefore unusable. The available raw data for the entire experiment was reprocessed, but due to the loss of 8.2% of the data, it was decided to use a summary file of previously processed data when possible. Thus, the loss of corrupted data mainly affected the spatial wave analysis (explained in Chapter 3), and the amount of processed wave slope values (Chapter 6). This is because both of these analyses require the raw wave displacement data.

The processed data came from a summary file from Arnold which provides 99% of the original one-hour processed Ku-band data and 96% of the C-band data. The summary file includes 1269 hours of Ku-band data and 1223 hours of C-band data of Arnold’s original 1280 hours of usable data. The statistics of the available processed data are reproduced in Table 2.1. These statistics are within a few percent of the values published by Arnold [1992]. The slight differences are due to missing data from the summary file.

For a detailed explanation of the experiment and data collection techniques the reader is referred to [5] and [6]. In this thesis, it suffices to know that the data was sampled at 8 Hz (parameters were written to disk every 0.125 seconds). One hour averages were computed for wind speed (referred to as $U$), wind direction, significant wave height (which is equal to four times the standard deviation of the wave displacement and referred to as $H_{\frac{1}{3}}$), and EM bias. Plots of these parameters during the length of the experiment are shown in Figs. 2.8 - 2.10. The wind speed plot

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Stdv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{\frac{1}{3}}$ (m)</td>
<td>0.60</td>
<td>3.18</td>
<td>1.44</td>
<td>0.47</td>
</tr>
<tr>
<td>$U$ (m/s)</td>
<td>0.1</td>
<td>16.9</td>
<td>6.9</td>
<td>2.9</td>
</tr>
<tr>
<td>$\epsilon_C$ (cm)</td>
<td>-0.35</td>
<td>-19.9</td>
<td>-5.6</td>
<td>3.3</td>
</tr>
<tr>
<td>$\beta_C$ (%)</td>
<td>-0.55</td>
<td>-6.33</td>
<td>-3.57</td>
<td>1.11</td>
</tr>
<tr>
<td>$\epsilon_{Ku}$ (cm)</td>
<td>-1.03</td>
<td>-13.8</td>
<td>-5.5</td>
<td>2.3</td>
</tr>
<tr>
<td>$\beta_{Ku}$ (%)</td>
<td>-1.62</td>
<td>-5.25</td>
<td>-3.71</td>
<td>0.63</td>
</tr>
</tbody>
</table>
Figure 2.5: Geometry of the tower used in GME. The nadir facing scatterometers are located at the center of the figure.
Figure 2.6: Picture of platform B in Fig. 2.5. This picture was taken from platform C facing southwest.

Figure 2.7: Picture of platform C in Fig. 2.5. This picture was taken from platform B facing northeast.
Figure 2.8: Wind Speed as a function of time for the GME data set.
Figure 2.9: Significant wave height, $H_{\text{m}}^{1/3}$, as a function of time for the GME data set.
Figure 2.10: EM bias calculated from Ku-band scatterometer as a function of time for the GME data set.
and significant wave height plot show when storms passed through. This time series also shows where the data was edited for tower interference. Particularly noticeable are the five storm surges during the first 25 days of the experiment. A similar time plot of the Ku-band bias shows how strongly the bias is correlated with wind speed.

Wave displacement was measured with an infrared wave gauge and was also calculated from the scatterometer Doppler. Due to the wave gauge not being functional during the entire experiment, corrections to account for differences between the two were calculated during the period when the wave gauge was functional, and applied to the entire experiment.

2.2.2 Bass Strait Data Set

Ken Melville, who currently works at Scripps Institute of Oceanography, provided data from the BSE. This experiment was performed from June 16th to September 26, 1992 at an offshore oil platform in Bass Strait, Australia. The Bass Strait is located between mainland Australia and Tasmania. Data was collected from 2 Ku-band scatterometers located at 15 and 25 m above sea level. The data from the 15 m scatterometer is used in this thesis to compare to the 18 m GME data; it had a footprint size of approximately 1.7 m. The BSE data was also edited for tower interference and 1539 hours of usable data were collected. Similar to GME, the BSE data was stored as one-hour averages. The statistics for the data are shown in Table 2.2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Stdv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H^\frac{1}{3}$ (m)</td>
<td>0.54</td>
<td>4.0</td>
<td>1.5</td>
<td>0.55</td>
</tr>
<tr>
<td>$\bar{U}$ (m/s)</td>
<td>0.3</td>
<td>17.25</td>
<td>6.1</td>
<td>3.3</td>
</tr>
<tr>
<td>$\epsilon_C$ (cm)</td>
<td>-0.48</td>
<td>-14.4</td>
<td>0.48</td>
<td>4.8</td>
</tr>
<tr>
<td>$\beta_C$ ($%H^\frac{1}{3}$)</td>
<td>-0.74</td>
<td>-5.27</td>
<td>-3.09</td>
<td>1.0</td>
</tr>
<tr>
<td>$\epsilon_{Ku}$ (cm)</td>
<td>-0.42</td>
<td>-16.9</td>
<td>-5.07</td>
<td>3.11</td>
</tr>
<tr>
<td>$\beta_{Ku}$ ($%H^\frac{1}{3}$)</td>
<td>-0.69</td>
<td>-5.68</td>
<td>-3.17</td>
<td>1.12</td>
</tr>
</tbody>
</table>
2.2.3 TOPEX/Poseidon Data Set

Gaspar [8] [9] provided a limited subset of processed TOPEX/Poseidon data. Two plots are provided in which he performs two different nonparametric regressions (NPR) to estimate the EM bias from $U$ and $H^{1/3}$. Since this data was collected from a satellite which illuminates a large area of the open ocean, the range of wind speeds and significant wave heights is larger than for the tower measurements. Wind speed ranges from 0 to 20.75 m/s and $H^{1/3}$ ranges from 0 to 11.75 m.

The first plot (and data set) is the same one from Fig. 3 of his June 1998 JGR article (See Fig. 2.11 in this thesis) [8]. The NPR technique used to generate this plot is a modified version of the Nadaraya-Watson Regression (NWR) estimator presented in this thesis. The differences are explained in Chapter 5. The second plot (and data set) is an unpublished result in which Gaspar has performed Local Linear Regression (LLR) estimation of the EM bias (see Fig. 2.12) [9]. Similar EM bias estimates for tower data will be computed and compared to subplots of Figs. 2.11 and 2.12.
Figure 2.11: EM bias estimation for TOPEX/Poseidon data using the Nadaraya-Watson Regression (NWR) estimator. The isolines enclose 95% of the global data set.
Figure 2.12: EM bias estimation for TOPEX/Poseidon data using the Local Linear Regression (LLR) estimator. The isolines enclose 95% of the global data set.
Chapter 3

Spatial Wave Analysis of the Electromagnetic Bias

As stated in the introduction, both tower and satellite data measurements of the electromagnetic bias level off at high wind speeds. Arnold [1992] showed that for Ku-band, the normalized bias measured from a tower levels off and even decreases at wind speeds above 10 m/s [5]. At C-band the effect is less pronounced and occurs at a higher wind speed (See Fig. 3.1). Hevizi observed this behavior in satellite data near a wind speed of 12 m/s [7].

It is believed that the cause for this is a change in the short-wave modulation over the wave profile at high wind speeds. It is well known that short-wave modulation has an effect on the EM bias [14]. However, it is not understood how the short-wave modulation changes at high wind speeds. The method of spatial wave analysis (SWA) provides direct measurements of the backscattered power versus the wave profile. Interpretation of the results provides insight about the short-wave modulation and its effect on the EM bias at high wind speeds.

3.1 Possible Causes for EM bias decrease at high wind speeds

Two possible reasons for the bias leveling off at high wind speeds are investigated in this chapter. The first possibility is that a significant amount of backscattered power is shifting from the trough to either the forward or rear face of the wave crests. A shift to the forward side of the wave crest could occur if the crests start breaking at high wind speeds, resulting in a rougher surface on the front side of the cresting wave than on the back. This would affect the scattering properties of the wave profile, resulting in an electromagnetic bias decrease at high wind speeds.
Figure 3.1: C- and Ku-Band Normalized Bias versus Wind Speed for GME. At 11 m/s the Ku-band normalized EM bias rolls off and begins to decrease. C-band normalized EM bias levels off near 12-13 m/s.
A second possibility is that at high wind speeds the wave profile begins to become more uniformly uneven. Physically, this means that the waves are becoming more choppy and the short wave modulation is distributing more equally over the wave. This would cause the backscattered power at the crests to increase, and the power at the trough to decrease. Since bias is due mainly to the non-uniform backscatter from the wave profile, this would result in a decrease in the bias.

### 3.2 Method and Results

Spatial wave analysis (SWA) refers to the analysis of the characteristics of the backscattered power over the phase, $\phi$, of the physical ocean wave. For this analysis, the definition of ‘one wave phase’ is from the peak of a previously passing wave to the next passing peak. The purpose for dividing the wave in this manner is so that the trough is located at $\phi = 0^\circ$, positive values of $\phi$ correspond to the front or rising side of a wave, and conversely negative values correspond to the back or falling side of the wave.
Figure 3.3: Thirty second wave displacement example. The asterisks (*) correspond to local crests and troughs found by use of the algorithm described in the text.

previous wave. The limits on $\phi$ are $[-180^\circ \leq \phi \leq 180^\circ]$ where $|\phi| = 180^\circ$ correspond to wave peaks (See Fig. 3.2).

The proposed causes for the bias decrease at high wind speeds can be attributed to changes in short wave modulation over the wave profile. A natural approach is to characterize the backscattered power, or $\sigma^o$, with respect to $\phi$. This is accomplished by computing $\bar{\sigma^o}$ over discrete bins of $\phi$, where $\bar{\sigma^o}$ is the normalized average $\sigma^o$ in each bin.

Wave displacement data from GME is used to establish $\phi$. The data collected during the month of February contains several storm surges which provide a wide range of wave displacement data and high wind speeds. It was sampled and saved to disk at 8 Hz in ten-minute segments (4934 sampled points per record). It should be noted that the equipment to measure and estimate wave displacement did not detect wave direction, only the rising and falling of the wave within the area illuminated by the scatterometer with respect to time (See Fig. 3.3). With this understood, two equal length vectors are available with wave displacement and $\sigma^o$ values sampled at
The steps used to estimate $\bar{\sigma}^\circ$ versus $\phi$ are explained as follows:

1. Find all points where the slope of the wave displacement changes from positive to negative. Place the indices in a new vector $\vec{m}$. This is equivalent to finding all the points where the derivative is equal to zero in a continuous waveform.

2. Remove local minimum and maximum points from $\vec{m}$ that do not correspond to actual wave crests and troughs. This is done choosing only the displacement values that correspond to highest maximums and lowest minimums. This step also insures that $\vec{m}$ is of the form $[\text{min}, \text{max}, \text{min}, \text{max}, \ldots]$ (See Fig. 3.3).

3. Wave phase is broken into 25 bins between $[-180^\circ, 180^\circ]$ of equal size. The vector $\vec{m}$ then is used to index $\sigma^\circ$ and power is distributed into respective bins. For example, the index points $m(i)$ and $m(i + 2)$ correspond to the endpoints of one wave phase, and $m(i + 1)$ corresponds exactly to the trough (assuming $m(i)$ is a maximum). If there are exactly 12 points between $m(i)$ and $m(i + 1)$ then the power from each point in the $\sigma^\circ$ vector is placed into respective $\phi$ bins for the falling side of the wave. The same is done for the rising side of the wave. The power scattered at $m(i + 1)$ always falls into the 13th $\phi$ bin, because it corresponds to the trough.

4. The number of points between $m(i)$ and $m(i + 2)$ is seldom exactly 25, and the number of data points over the falling portion of the wave is not usually the same as the rising portion. Because of this, the power is unequally distributed into each wave phase bin. To account for this, the total amount of power which falls into each $\phi$ bin is summed and divided by $N$, the number of $\sigma^\circ$ values that contributed to it; then each power measurement is divided by the maximum power. The resulting average normalized power for that binned section of wave phase is called $\bar{\sigma}^\circ$.

5. Then $\bar{\sigma}^\circ$ versus $\phi$ data is generated for each ten minute segment of wave displacement data, and the centroid of that power is computed.
Figure 3.4: Typical wave phase versus $\bar{\sigma}^\phi$ plot for C and Ku-band averaged for 10 minutes of data. Average wind speed=4.2 m/s. The y-axis is average normalized power, $\bar{\sigma}^\phi$.

These centroid values are computed in degrees of $\phi$ and indicate where the centroid of power is backscattered with respect to the phase of the wave. C- and Ku-band $\bar{\sigma}^\phi$ data is used to generate ten-minute centroid values for each band for the entire month of February. Since the rest of the data for the experiment is available in the same ten-minute records, these centroid values can be compared to wind speed and bias values.

Typical C- and Ku-band $\bar{\sigma}^\phi$ versus $\phi$ plots are shown in Fig. 3.4. The average wind speed during this ten minute segment of data is 4.2 m/s, and one notes that the centroids shift slightly towards the front side of the wave. As expected, more power is reflected from the trough of the wave ($0^\circ$) than from the crests. The average $\bar{\sigma}^\phi$ values in this plot are 0.71 and 0.75 for C- and Ku-band. Since the peak power is normalized to 1.0, this means that roughly 29% to 25% more power is reflected from the trough than from mean sea level (MSL). The average $\bar{\sigma}^\phi$ values at the crests are 0.47 and 0.61 for C- and Ku-band respectively implying that 24% to 14% less power is backscattered from the crests than MSL. These differences are the known cause of the electromagnetic bias.
The centroid of the power exhibits a small shifting to the right and to the left of the wave trough. The shift is larger at Ku-band than C-band, indicating that more short wave modulation probably exists between 2.1 cm ($\lambda_{Ku}$) and 6 cm ($\lambda_C$) than above 6 cm. The absolute value of the mean shift in the two cases is 2.4° and 3.9°. Numerically, this means that if the peak to peak ocean wavelength is 60 m then the centroid power shift is 0.39 m to 0.65 m away from the trough.

A histogram of centroid phase shifts versus wind speed for C- and Ku-band data is shown in Fig. 3.5 and 3.6. It must be noted that the y-axis scale on both plots only displays ±20° of the wave phase. The magnitude of the centroid shifts at C- and Ku-band is small. The centroid values are strongly bifurcated at Ku-band; they are rarely equal to 0°. For C-band the centroid values are positive more often than negative. The centroid appears to be affected by changes in wind direction in the presence of high wind speeds (see Fig. 3.7).

The average Ku-band power versus wave phase for the entire month is shown in Fig. 3.8. A sample waveform is used with this power distribution to test the effect of a 3.9° centroid shift. EM bias is given in the previous chapter as,

$$\epsilon = \frac{\sum \sigma^0 \eta}{\sum \sigma^0}$$

(3.1)

where $\eta$ is the wave displacement. Assuming a sinusoidal wave displacement vector, $\eta = -\cos(x) : x \in [-180, 180]$ (as in Fig. 3.2), the calculated bias is -9.68 cm. An average of 300 power versus phase distributions with a centroid shift within ±0.5° of 3.9° yields a bias of -8.93 cm. Thus, the average centroid shift for Ku-band changes the bias by 0.75 cm.

A plot of Ku-band ($\sigma^0_{\text{strough}}$)/($\sigma^0_{\text{crest}}$) versus wind speed shows a trend similar to Ku-band EM bias versus wind speed (see Fig. 3.9). The negative slope at wind speeds above 11 m/s indicate that the power scattered from the crest is getting nearer to the power scattered from the trough. This means that the wave profile is becoming more uniformly uneven.

Scatterplots of the centroids versus wind direction reveal that SWA may possibly be used to confirm wind and tower interference. In Fig. 2.5 the geometry of the
Figure 3.5: Normalized C-band centroid $\overline{\sigma}_C$ plotted versus wind speed.
Figure 3.6: Normalized Ku-band centroid $\overline{\sigma^c_{Ku}}$ plotted versus wind speed.
Figure 3.7: Time plots of wind speed, wind direction, C-band centroids, Ku-band centroids.
Figure 3.8: Average of Ku-band $\phi$ versus $\sigma^0$ plots for February. The solid line includes all data for the entire month and the dotted line is for all data with a $3.9^\circ \pm 0.5^\circ$ centroid shift. The y-axis is average normalized power, $\overline{\sigma^0}$. 

Figure 3.9: Plot of Ku-band $\sigma^0_{\text{trough}}/\sigma^0_{\text{crest}}$ versus wind speed. Note that the plot never goes below 1, indicating that the power reflected at the troughs is always greater than at the crests.
Gulf of Mexico tower and equipment location shows the directions of possible wave interference to be between $[40^\circ, 95^\circ]$ and $[180^\circ, 275^\circ]$. The main wind interference occurs between $[135^\circ, 180^\circ]$. Figures 3.10 and 3.11 show the scatterplots with the indicated directions added with vertical dotted lines. The data for C- and Ku-band strongly indicates the $[135^\circ, 180^\circ]$ cutoffs. This is due to wind interference causing the anemometer to get very few readings. Between $[40^\circ, 95^\circ]$ the computed centroid shifts are very similar to those between $[0^\circ, 135^\circ]$ for C-band. But for Ku-band this region is where a major centroid shift occurs. Because the trends continue outside of this region for Ku-band, tower interference at these angles does not appear to affect the data much. Between $[180^\circ, 275^\circ]$ the centroid values are more sparse and have a noticeably larger variance for both C- and Ku-band data. When waves came from these directions, they passed through two towers and the waves were much more affected than those coming from $[40^\circ, 95^\circ]$.

Assuming that interference patterns can be inferred from this data, wind directions between $[0^\circ, 50^\circ]$, $[80^\circ, 135^\circ]$, and $[275^\circ, 360^\circ]$ are good. Fig. 3.12 shows these wind direction cutoffs as dashed lines. Using this mask, 885 of 2695 total points of data were removed.

The resulting plots of centroid versus wind speed using the interference pattern explained are shown in Figs. 3.13 and 3.14. They are not significantly different from Figs. 3.13 and 3.14 indicating that the results presented previously do not need to be modified. They are included in this thesis for the benefit of the reader. A similar replotting of Fig. 3.9 indicated no change in the general trend (because it was virtually identical, it is not included here).

### 3.3 Summary

The results indicate that at high wind speeds the wave profile begins to become more uniformly uneven, supporting the second hypothesis. As the wind speeds increase beyond 11 m/s the amount of backscatter from the crests increases. This is seen by the change in slope in Fig. 3.9 at 11 m/s. The trend is very similar to Ku-band bias versus wind speed (See Fig. 3.1). In support of the first hypothesis,
Figure 3.10: Normalized C-band centroid $\bar{\sigma}_C$ plotted versus wind direction.
Figure 3.11: Normalized Ku-band centroid $\bar{\phi}_{Ku}$ plotted versus wind direction.
Figure 3.12: Areas of assumed wind and tower interference inferred from Figs. 3.10 and 3.11.
Figure 3.13: New normalized C-band centroid $\overline{\sigma}_C$ plotted versus wind direction.

Figure 3.14: New normalized Ku-band centroid $\overline{\sigma}_{Ku}$ plotted versus wind direction.
the centroid does shift from the front to the back side of the cresting wave numerous times throughout the month of February (see Fig. 3.7). A 3.9° Ku-band centroid shift affects the magnitude of the bias by 0.75 cm. This is approximately 13% of the mean Ku-band EM bias. This centroid shift, however, appears to be more correlated to wind direction than to high wind speeds. Therefore, while it does affect the bias, it is not the cause of the normalized bias leveling out at high wind speeds.

In addition to answering the question about why the normalized EM bias levels off at high wind speeds, SWA also indicated areas where the wind and wave data were corrupted from tower interference. Areas where the waves passed through physical obstructions from the tower are indicated by both positive and negative values of centroid shifts, resulting in a higher amount of variance. Non-affected areas have strictly positive or negative shifts. Removing the assumed areas of interference does not affect the centroid versus wind speed trends or the plot of \( \frac{\sigma^2_{\text{trough}}}{\sigma^2_{\text{crest}}} \) versus wind speed. Thus, interference does not affect the results.
Chapter 4

Traditional EM Bias Estimation

Traditionally, wind speed, $U$, and significant wave height, $H_{\frac{\lambda}{3}}$, are used with an equation such as Eq. (4.1) to estimate the EM bias [16]. This equation is trained with direct measurements of $U$, $H_{\frac{\lambda}{3}}$, and $\epsilon$ in order to obtain appropriate coefficients, $a_i$:

$$
\epsilon = H_{\frac{\lambda}{3}} \left[ a_1 + a_2 H_{\frac{\lambda}{3}} + a_3 U + a_4 H_{\frac{\lambda}{3}}^2 + a_5 U^2 + a_6 H_{\frac{\lambda}{3}} U \right] \tag{4.1}
$$

This equation is a modification of the NASA equation used to generate the highly successful TOPEX Geophysical Data Records (GDRs). The difference is that the parameter $\rho$ has been removed which Gaspar shows does not significantly improve the EM bias estimate [16].

4.1 Parametric Fitting to GME and BSE EM Bias

In order for Eq. (4.1) to accurately estimate the EM bias, appropriate coefficients ($a_i$) must be obtained by training Eq. (4.1) with measurements of $\epsilon$, $U$, and $H_{\frac{\lambda}{3}}$. The training method results in a residual bias with minimum variance and zero mean (See Table 5.2). This method is an iterative minimum least squares problem. First, a good guess for the coefficients is provided to the algorithm, then the bias is estimated and the residual is computed. The coefficients are altered until the minimum variance zero-mean resulting residual bias remains. Since the purpose of this chapter is to provide a comparison for the upcoming chapters, no more will be said of this training method. Applying this minimization problem to the GME data resulted in the coefficients shown in Table 4.1. Note that the largest magnitude coefficient in
Table 4.1: Coefficients obtained for GME and BSE data.

<table>
<thead>
<tr>
<th>Term in eq.</th>
<th>Coeff.</th>
<th>GME C-band</th>
<th>GME Ku-band</th>
<th>BSE Ku-band</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>$a_1$</td>
<td>0.6560</td>
<td>-1.0005</td>
<td>0.5041</td>
</tr>
<tr>
<td>$H^{1}_{3}$</td>
<td>$a_2$</td>
<td>-3.1693</td>
<td>-1.5996</td>
<td>-1.8083</td>
</tr>
<tr>
<td>$U$</td>
<td>$a_3$</td>
<td>-0.1493</td>
<td>-0.2937</td>
<td>-0.4339</td>
</tr>
<tr>
<td>$H^{12}_{3}$</td>
<td>$a_4$</td>
<td>0.8010</td>
<td>0.2631</td>
<td>0.3359</td>
</tr>
<tr>
<td>$U^2$</td>
<td>$a_5$</td>
<td>0.0024</td>
<td>0.0066</td>
<td>0.0146</td>
</tr>
<tr>
<td>$H^{1}_{3}U$</td>
<td>$a_6$</td>
<td>-0.0560</td>
<td>0.0589</td>
<td>0.0118</td>
</tr>
</tbody>
</table>

both cases is $a_2$ indicating that $H^{1}_{3}$ is most strongly correlated to normalized bias. In both results $a_5$ and $a_6$ are very small indicating a very small dependence on $U^2$ and $H^{1}_{3}U$. The first term, $a_1$, is an offset which makes the mean of the residual bias as close to zero as possible.

A useful parameter to quantify the goodness of an EM bias estimate using to an actual measurements is Pearson’s correlation coefficient. It is given by Freund [19] as,

$$
\hat{\rho} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}.
$$  \hspace{1cm} (4.2)

where $x_i$ and $y_i$ correspond to the measured data points and the estimated data points.

Figures 4.1 and 4.2 show the GME EM bias estimate in centimeters. The correlation coefficients given by Eq. (4.2) for this estimate are 0.9760 and 0.9732 for C and Ku-band respectively. The residual bias is defined as

$$
\epsilon_{\text{residual}} = \epsilon_m - \epsilon_{\text{est}}
$$  \hspace{1cm} (4.3)

where $\epsilon_m$ is the measured EM bias for the described data set, and $\epsilon_{\text{est}}$ is the estimate. These residual biases are plotted for the entire GME data set in Figs. 4.4 and 4.5.
The days of the experiment when there is no data accounts for times when tower interference corrupted the data and it was removed. Figure 4.3 then shows the same Ku-band residual bias with the no-data regions removed with a y-axis of ±5 cm. This method is used to plot residual biases in subsequent chapters. The x-axis in this plot is essentially time, but is left unlabeled to remind the reader of the gaps due to tower interference. The residual EM biases for C and Ku-band have standard deviations, \( \sigma_{RES} \), of 0.61 and 0.71 cm respectively.

Training Eq. (4.1) with the Ku-band BSE data yielded the coefficients shown in Table 4.1. Using these coefficients the EM bias was calculated and plotted in the two-dimensional plane of \( U \) and \( H^\frac{1}{3} \). Figure 4.6 shows the BSE EM bias estimate in centimeters. The correlation coefficient for this Ku-band EM bias estimate is 0.9586. The standard deviation of the residual bias is 0.8855 cm. Figure 4.7 shows the residual bias after removing the parametric fit estimate.

4.2 Summary

The traditional parametric estimation of the bias using a second order Taylor’s Series expansion on the variables \( U \) and \( H^\frac{1}{3} \) resulted in a fair EM bias estimate. The residual biases are not completely random. The fluctuations in the remaining bias suggest a correlation to some other oceanic parameter. Assuming the nonparametric method results in a better estimate, it may reduce the residual bias even more. In addition, wave slope may be correlated to the residual bias. In Chapters 5 and 6 the results presented here are compared to nonparametric regression (NPR) methods of EM bias estimation.
Figure 4.1: GME - C-band Parametric EM bias fit in centimeters with respect to $U$ and $H^1/3$. 
Figure 4.2: GME - Ku-band Parametric EM bias fit in centimeters with respect to $U$ and $H^{\frac{1}{3}}$.

Figure 4.3: Cropped Ku-Band Residual bias for GME after removing estimated bias from the parametric fit. In this plot the time periods where there is no data have been removed, and the data points are connected with a solid line. This is the format which the residual biases in subsequent chapters will be presented.
Figure 4.4: C-band Residual bias for the length of the GME after removing parametric fit. The parametric estimate was removed from the measured C-band bias and the residual is shown in centimeters.
Figure 4.5: Ku-Band Residual bias for length of the GME after removing parametric fit. The parametric estimate was removed from the measured Ku-band bias and the residual is shown in centimeters.
Figure 4.6: BSE - Ku-band Parametric EM bias fit in centimeters with respect to $U$ and $H^{1/3}$.

Figure 4.7: Cropped Ku-Band Residual bias for BSE after removing estimated bias from the parametric fit. In this plot the time periods where there is no data have been removed, and the data points are connected with a solid line. This is the format which the residual biases in subsequent chapters will be presented.
Chapter 5

Nonparametric Estimation of Electromagnetic Bias

In a recent publication by Gaspar and Florens [8], the EM bias for TOPEX/Poseidon data is estimated using a nonparametric regression (NPR) technique referred to as Nadaraya-Watson regression (NWR) (See Fig. 2.11). In their paper, Gaspar and Florens argue that the traditional parametric estimation models of the bias are not true least squares approximations of the bias, but are only best match fitting functions to the densest data region. In fact, traditional estimation models are poor estimators for less dense data regions [8]. This was the first application of a NPR technique to the field of EM bias estimation. Simonoff states that using NPR estimators allows alternative (and perhaps unexpected) structure of the data to come through, which might be constrained by a parametric fit [20]. Since the optimal form for a parametric fit to EM bias is unknown, it is expected that the NPR method performs better than existing parametric fits.

In this chapter the derivation for NWR and a higher order NPR estimator called Local Linear Regression (LLR) is presented. A known problem with NWR is that it is biased poorly at natural boundary regions in the data set [21]. Gaspar and Florens also indicate that with a large data set, such as TOPEX/Poseidon, the NWR estimate is biased where the data density has an abrupt change [9]. LLR does not have this problem. Bowman states that LLR’s behavior near the edges of the region over which the data has been collected is superior to that of the NWR [22]. Gaspar recently provided me with his results from performing LLR on EM bias measured from TOPEX/Poseidon data [9]. Both NWR and LLR are applied to Ku-band data collected from the Gulf of Mexico (GME) and Bass Strait, Australia
Figure 5.1: Data densities in the wind speed and significant wave height plane. Includes hourly average values of the data from the Gulf of Mexico Experiment (above) and the Bass Strait Experiment (below).
Figure 5.2: Averaged bias values over a grid of \(dU = dH = 0.25\) for GME. Note the contour pattern that emerges. Data used is from GME.

(BSE) experiments, and compared to Gaspar’s results. The EM bias in both cases is estimated over a two-dimensional space of \(U\) and \(H_{\frac{1}{3}}\).

5.1 Limitations of Gridding without Smoothing

When a two-dimensional histogram is plotted of wind speed and significant wave height, one notices where all the data points lie (See Fig. 5.1). Thus, data density, as referred to in this thesis, is where the data lies in this plane. In the data sets each of these points has an EM bias value associated with it. If a grid of \(dU = dH_{\frac{1}{3}} = 0.25\) is chosen and the average value of the Ku-band EM bias is taken at each grid point, the contour plot shown in Fig. 5.2 results. Contour lines emerge in the densest data region (see Fig. 5.1), but they are not smooth. The result would not be a good estimator for EM bias. It would be more desirable to obtain a smoothed plot which estimates values of the bias over the entire range of possible wind speeds and wave heights. This is the purpose for turning to nonparametric techniques.
Two-Dimensional Nonparametric Regression

Nonparametric regression (NPR) is a method to statistically smooth a data set such that a valid estimate for one variable is available over a chosen grid spacing of different variable(s). Data points which don’t lie exactly on the grid spacing can easily be interpolated.

It is desired to estimate the electromagnetic bias, \( \epsilon \), given \( U \) and \( H \). To simplify the notation, let \( H^{1/3} \) be noted by simply \( H \). For both regression estimators, consider the data triples \((U_1, H_1, \epsilon_1), (U_2, H_2, \epsilon_2), \ldots, (U_n, H_n, \epsilon_n)\), which form an independent and identically distributed sample from a population \((U, H, \epsilon)\). Letting \( \bar{\epsilon} \) equal \((U, H)\) in general; \( \bar{\epsilon}_o \) equal \((U_o, H_o)\), an arbitrary point on the chosen grid space; and \( \bar{X}_i \) equal the sample \((U_i, H_i)\) from measured wind and significant wave height from either GME or BSE, it is desired to estimate the regression function,

\[
\hat{\epsilon}(\bar{x}) = E[\epsilon|\bar{x}=(U,H)]
\]  

(5.1)

where \( \hat{\epsilon} \) is the EM bias estimate. For the two regression estimators presented in this paper, this equation is expanded about the point \( \bar{\epsilon}_o = (U_o, H_o) \) using a Taylor’s series,

\[
\hat{\epsilon}(\bar{x}) \approx a_0 + a_1(U - U_o) + a_2(H - H_o)
\]  

(5.2)

where \( a_0 \) is \( \hat{\epsilon}(\bar{\epsilon}_o) \), one point of the desired EM bias estimate, and the coefficients \( a_1 \) and \( a_2 \) are the partial derivatives of \( \hat{\epsilon}(\bar{\epsilon}_o) \) with respect to \( U \) and \( H \) respectively. Then the NPR estimate of the EM bias, \( \tilde{\epsilon} \), is found by solving the minimum least squares problem,

\[
\min_{a_0,a_1,a_2} \sum_{i=1}^{n} \left( \epsilon_i - a_0 - a_1(U - U_o) - a_2(H - H_o) \right)^2 K_h(X_i - x_o)
\]  

(5.3)

for all \( a_i \) where \( K \) is a two-dimensional kernel smoothing function, with bandwidth parameter \( h \). Both \( K \) and \( h \) will be explained later. From this equation it can be seen that the estimate, \( \hat{\epsilon}(\bar{x}) \), is being subtracted from the measured value of EM bias, \( \epsilon_i \), and being squared. This is the squared error term which is being multiplied by \( K_h \) and minimized. It is important to note that this is not the same as parametric fitting where \( a_i \)'s are solved for and then plugged back into the equation to estimate the EM
bias (i.e. Eq (5.2)), but \( a_0 \) is the EM bias estimate at a point on the chosen grid spacing and is a smoothed version of the data near it. It will be shown that estimates for the partial derivatives of EM bias with respect to \( U \) and \( H \) are also provided with the LLR solution.

### 5.2.1 Nadaraya-Watson Regression Estimator

The Nadaraya-Watson Regression (NWR) estimator \([21]\) is obtained by solving Eq. (5.3) when \( a_1 = a_2 = 0 \). Under these conditions the solution to the minimization problem yields,

\[
\hat{\epsilon}_{NW}(\bar{x}) = \sum_{i=1}^{n} w_i \epsilon_i \tag{5.4}
\]

where

\[
w_i = \frac{K_h \left( \bar{x} - \bar{X}_i \right)}{\sum_{i=1}^{n} K_h \left( \bar{x} - \bar{X}_i \right)} \tag{5.5}
\]

In these equations, \( \bar{X}_i \) and \( \bar{x} \) are as defined previously.

Applying the two-dimensional NWR estimator to EM bias estimation amounts to convolving a two dimensional kernel function, \( K_h \), with bias values over the \( U \) and \( H \) plane. NWR is classified by Fan and Gijbels as a locally weighted average kernel estimator \([21]\).

If \( \hat{\epsilon}_n \) is defined as the column vector of all measured EM biases, \( \epsilon_i \), corresponding to the two column matrix pairs of \( \bar{X}_n = [\bar{U} \bar{H}]_n \), the Nadaraya-Watson estimator can be expressed in matrix form as

\[
\hat{\epsilon}_{NW}(\bar{x}) = \bar{W}_n^T \hat{\epsilon}_n \tag{5.6}
\]

where the row vector of weights, \( \bar{W}_n^T \), is defined as

\[
\bar{W}_n^T = \frac{K_h \left( \bar{x} - \bar{X}_n \right)}{\sum_{i=1}^{n} K_h \left( \bar{x} - \bar{X}_n \right)} \tag{5.7}
\]

This algorithm was coded in \textit{Matlab} and tested several times to ensure that it worked correctly, and then applied to data from GME and BSE. The two tower EM bias
results are compared to each other and then compared to the TOPEX/Poseidon NWR bias estimate.

5.2.2 Local Linear Regression Estimator

The Local Linear Regression (LLR) estimator estimates the EM bias by minimizing Eq. (5.3) with all the terms. Since all the terms are used, it is desirable to use matrix notation. In order to present the solution a few definitions must be made in addition to the ones previously defined. Let the following matrices be defined as

\[
\tilde{A} = \begin{pmatrix} a_0 & a_1 & a_2 \end{pmatrix}^T, 
\]

\[
\mathcal{X} = \begin{pmatrix} 1 & U - U_1 & H - H_1 \\ 1 & U - U_2 & H - H_2 \\ \vdots & \vdots & \vdots \\ 1 & U - U_n & H - H_n \end{pmatrix}, 
\]

and

\[
\mathcal{W} = \text{diag}\{K_h(\tilde{x} - \tilde{X}_n)\} 
\]

where \(K_h\) is the two-dimensional kernel function with bandwidth parameters, \(h\), which will be explained in the next section. With these matrices defined, the LLR solution to Eq. (5.3) yields

\[
\tilde{A} = (\mathcal{X}^T \mathcal{W} \mathcal{X})^{-1} \mathcal{X}^T \mathcal{W} \tilde{e}_n 
\]

where \(\tilde{e}_n\) is the column vector of EM bias measurements as in NWR. The matrix \(\tilde{A}\) gives one grid point of the EM bias estimate, \(a_0 = \tilde{e}_0\), and \(a_1\) and \(a_2\) are the partial derivatives of \(\tilde{e}\) with respect to \(U\) and \(H\) as explained by Eq. (5.2). This algorithm was also coded and tested, then NWR and LLR were combined into one Matlab process called \(\text{ffnp}\) for which the code is included in Appendix A.
5.2.3 Kernel Functions

The kernel function acts the same for both NPR estimators. Its purpose is to smooth the EM bias estimation over the $U$ and $H^{\frac{1}{3}}$ plane. The normalizing term in both Eqs. (5.7) and (5.11) serves the purpose of insuring that the weighting values, $w_i$ satisfy $\sum_i^n w_i = 1$. The kernel function itself is also assumed to satisfy $\int K(\vec{t}) \, d\vec{t} = 1$, but does not have to because of its presence in both the numerator and the denominator of both of the solutions.

The two dimensional Gaussian and Epanechnikov kernel functions are used in the analysis of the Gulf of Mexico data. The simplified two dimensional form of the Gaussian is

$$K_h \left( \vec{x} - \vec{X}_i \right) = K_g \left( \frac{x - [U,H]_i}{h} \right) = C_g \left( \exp \left[ - \left( \frac{(U - U_i)^2}{2h_U^2} + \frac{(H - H_i)^2}{2h_H^2} \right) \right] \right)$$

(5.12)

where $C_g$ is a constant to normalize the Kernel function such that $\int K(\vec{t}) \, d\vec{t} = 1$, and $h_X$ are bandwidth parameters which are a function of the standard deviation of the data. The two dimensional Epanechnikov is similar with a normalizing variable $C_e$

$$K_h \left( \vec{x} - \vec{X}_i \right) = K_e \left( \frac{x - [U,H]_i}{h} \right) = C_e \left( 1 - \left( \frac{U - U_i}{h_U} \right)^2 - \left( \frac{H - H_i}{h_H} \right)^2 \right)_+$$

(5.13)

where the $()_+$ means to use only positive values. This kernel function is equal to zero for negative values. This means that if the grid point $[U,H]$ is close enough to $[U,H]_i$ for the kernel to be positive then it is included in the weighting of the estimation at that grid point, otherwise it is not. One notes from Eqs. (5.7) and (5.11) that the normalizing constant, $C_x$, cancels out. Thus, the amplitude of the kernel does not effect the nonparametric estimation method, while its overall shape does.

The Epanechnikov kernel is actually the first three terms of the Taylor’s series expansion of the Gaussian kernel. For this reason, it makes sense that only the
positive values of the Epanechnikov kernel are used (considering the shape of a two-dimensional Gaussian). In addition, negative weighting would have an adverse effect on the outcome.

Graphically, data points nearest to the evaluation point $\bar{x}$ are weighted more heavily while points further away are weighted less (see Fig. 5.3). With the Epanechnikov kernel, no weights are assigned outside the bandlimited region of the kernel. The Epanechnikov kernel is limited to estimating EM bias values within a boundary of where the actual data lies in the $(U, H)$ plane. But this is not a setback because bias estimates for values far outside of the actual data region are not meaningful. For example, low values of wind speed are seldom or never associated with high values of significant wave height, and vice versa.
5.2.4 Bandwidth Selection

The bandwidth is closely tied to the kernel function; its function is to define the two-dimensional area over which the bias values are weighted in $(U, H^\frac{1}{3})$ space. As might be expected, in the Gaussian kernel the bandwidths are the standard deviations of the two-dimensional kernel function. Thus, 99% of the Gaussian weighting lies within the elliptical region defined by $3(h_U, h_H)$. The Epanechnikov kernel is zero outside of the ellipse defined by its bandwidth values while the Gaussian continues to have finite values over a long tail. Because of these characteristics, the Gaussian appears to estimate areas where there is not data. The function of the bandwidth parameter is best understood through the use of Fig. 5.3. Both kernel functions are shown centered at $\bar{x} = (6, 1.5)$, with the same bandwidth values ($h_U = 2$ and $h_H = 0.32$). These bandwidths were chosen for graphical analysis purposes and are larger than the actual bandwidths which are used in the NPR estimates. Even though the bandwidths are the same, the area which will be smoothed by the Gaussian is larger than the Epanechnikov. Due to this difference, typically the bandwidth used for the Epanechnikov kernel is larger in order to achieve the same amount of smoothing.

Simonoff [20] states that the asymptotically optimal bandwidth for the kernel estimator of $\hat{f}$ satisfies

$$h_o = \left[ \frac{R(K)}{\sigma_k^4 R(f^n)} \right]^{1/5} n^{-1/5}, \quad (5.14)$$

where

$$R(g) = \int g(u)^2 \, du, \quad (5.15)$$

and $f$ is the reference density. This is optimal because it leads to an asymptotic minimum mean integrated squared error (MISE) [20]. Unfortunately, the density function of EM bias is not well defined. In his paper, Gaspar assumes that $f$ is roughly Gaussian [8]. Under this assumption, Simonoff [20] provides bandwidth values which are proportional to the standard deviation of the data,

$$h_x^g = 1.059 \sigma_x n^{-1/5} \quad (5.16)$$
Table 5.1: Bandwidth Values for Gulf of Mexico and Bass Strait Experiments.

<table>
<thead>
<tr>
<th>Kernel Function</th>
<th>GME</th>
<th>BSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( h_U )</td>
<td>( h_H )</td>
</tr>
<tr>
<td>Gaussian</td>
<td>0.74</td>
<td>0.12</td>
</tr>
<tr>
<td>Epanechnikov</td>
<td>1.63</td>
<td>0.26</td>
</tr>
</tbody>
</table>

and in order to achieve the same MISE with the Epanechnikov kernel, the bandwidth is

\[
h^*_x = 2.347 \sigma_x n^{-1/5}
\]

(5.17)

where \( x \) corresponds to \( U \) or \( H \), \( \sigma_x \) is the standard deviation of the respective data set, and \( n \) is the number of total data points. Thus, the bandwidth is proportional to the standard deviations of \( U \) and \( H \), and inversely proportional to the amount of data points available.

For the GME data set the standard deviations of the wind speed and significant wave height were \( \sigma_U = 2.9 \) m/s and \( \sigma_H = 0.47 \) m, for the 1269 data points. For the 1539 data points of the BSE they were \( \sigma_U = 3.3 \) m/s and \( \sigma_H = 0.55 \) m. Using these values in Eqs. (5.16) and (5.17) yields the bandwidth values shown in Table 5.1 which are used with the NPR estimation techniques.

5.3 Results

The results using two estimation techniques for GME and BSE are presented in this section. NWR and LLR have similar results in both cases.

5.3.1 NWR EM Bias Estimation

Using the same grid spacing used in the TOPEX data facilitates comparisons between tower and satellite data. Thus, spacings of \( dU = dH = .25 \) m were chosen, and the EM bias values were estimated at each point on the grid using Eqs. (5.7) and (5.11). For GME data, NWR results are shown in Figs. 5.4 and 5.5. For BSE
data, similar plots are shown in Figs. 5.8 and 5.9. Comparison of these figures show that the Gaussian kernel contour lines extend to the edges of the plot, while the Epanechnikov kernel does not. The region of greatest interest is where the data lies, outside of which has been grayed out. Within this region the results of the two kernel functions are nearly identical for both GME plots and both BSE plots.

5.3.2 LLR EM Bias Estimation

Using the LLR technique very similar results were obtained. Figures 5.6, 5.7, 5.10, and 5.11 show the results of this technique for GME and BSE data using the same grid spacing.

The following pages contain the resulting figures. (The chapter Summary follows.)
Figure 5.4: GME: NWR-Gaussian bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.5: GME: NWR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.6: GME: LLR-Gaussian bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.7: GME: LLR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.8: BSE: NWR-Gaussian bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.9: BSE: NWR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.10: BSE: LLR-Gaussian bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
Figure 5.11: BSE: LLR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $U, H$ plane.
5.3.3 NWR Implementation Differences for Tower versus Satellite Data

The NWR technique used by Gaspar [8] is not identical to the NWR method used to obtain tower EM bias estimates. This is due to the fact that satellite EM bias estimates are more difficult to obtain.

In Chapter 2 it was shown that the EM bias for tower data, in situ measurements provide direct observations of the EM bias,

\[ y = \epsilon(U, H) + e_t \]  \hspace{1cm} (5.18)

where \( y \) is the processed EM bias measurement which includes noise, \( e_t \). Satellite data, on the other hand, is not a direct measurement but a difference estimate from two crossover points:

\[ y_1 - y_2 = \epsilon(U, H)_1 - \epsilon(U, H)_2 + e_s \]  \hspace{1cm} (5.19)

where \( y_i \) is a sea surface height measurement at a crossover point, and the 1 and 2 indices refer to measurements made on the ascending/descending arcs. Retrieving the EM bias estimate from Eq. (5.19) is a more difficult task than retrieving it from Eq. (5.18). Due to the higher number of unknowns in the linear system of equations that results, Gaspar uses a good estimate of the bias in the densest data region as an imposed value constraint. Then a solution is computed and the value of the bias at (0,0) was subtracted out as a mean value, assuming that no bias can occur at this physical condition. This method leaves room for an unknown bias in the final solution. Gaspar states that this unknown bias may be ±1 cm [9]. It is important to note that a ±1 cm shift is equal to a ±0.67% of the mean significant wave height value for both experiments (GME and BSE). That leaves 1.3% of room for error.

A second difference is due to the vast amount of data available from TOPEX/Poseidon. Gaspar states that using each of the 100 repeat cycles of the global data set available would require the inversion of a hundred \([(n - 1) \times (n - 1)]\) matrices, where \( n \) is typically between 6000 and 7000 [8]. His approach requires much less computer time to evaluate. Gaspar takes the mean of several iterations of the NWR estimator with a subset of 500 points of the total TOPEX data set each time. Because of this,
the TOPEX NWR algorithm for estimating EM bias is more *smoothed* than the tower NWR estimate. These explanations may account for some of the major differences in the EM bias estimates for the NWR technique.

### 5.3.4 Comparison of Results to TOPEX/Poseidon Data

Subplots with the same axis limits as the tower data are plotted for TOPEX/Poseidon data in Figs. 5.12 and 5.13. These two plots are very different from each other. For tower EM bias estimates Figs. 5.4 - 5.11 are all very similar to each other. The explanations outlined in the previous section possibly account for these differences. The LLR method was applied to TOPEX/Poseidon data more recently than NWR and a more efficient method has been used [9]. The results are much more similar to tower data than the NWR method.

Certain patterns in the satellite and tower EM bias estimates are similar. For example, the contour lines decrease from left to right, and at a wind speed of about 12 m/s the contour lines in the TOPEX contour begin to slope upward from left to right. The tower data exhibits this pattern at about 10.5 m/s. Because the bulk of the data for the tower experiments is near a boundary region for the TOPEX data, the NWR estimate is more likely to be biased in this region. Therefore, the NWR comparison is poor. Gaspar and Florens recognized this problem and have since reanalyzed the TOPEX/Poseidon data using LLR [9]. The mentioned similarities between tower and satellite data using the LLR technique are much more pronounced. More importantly, the actual EM bias values estimated using the LLR technique are closer.

There are some differences in the two contour plots as well. The first noticeable difference is the apparent slope of the contour lines from left to right on the plot. The tower results show an almost constant bias value for significant wave height values at low wind speeds; the contour lines are nearly horizontal. The contour lines at low wind speeds for the LLR TOPEX data have a slope of approximately $-1/8 \frac{m}{m/s}$.

The mean difference between the magnitude of the LLR curves for satellite and tower data is 0.94 cm and 1.3 cm for GME and BSE respectively. This means
Figure 5.12: TOPEX/Poseidon: NWR*-Gauss. bias contours in $U, H$ plane. The data is plotted with grid limits and shading from previously shown GME tower data. The EM bias estimates are in centimeters. (* indicates that this NWR is not equivalent to the NWR method outlined in this thesis. The differences are explained in Section 5.3.3.)
Figure 5.13: TOPEX/Poseidon: LLR-Epan. bias contours in $U, H$ plane. The data is plotted with grid limits and shading from previously shown GME tower data. The EM bias estimates are in centimeters.
that the satellite bias is slightly larger in magnitude than the tower data on average. A recent theoretical study by Johnson suggests that EM bias scattered from a large footprint, such as a satellite, is larger in magnitude than the scatter from a smaller footprint [23].

In order to better understand Johnson’s results some fundamental parameters of the ocean surface spectrum must be defined. He defines \( \lambda \) to be a measure of surface roughness equal to \( 4h^2\kappa^2 \) where \( h^2 \) is the surface wave height variance and \( \kappa \) is the electromagnetic wave number. In addition \( k_o \) is the low wave number cutoff; \( p \) is a parameter defining rate of spectral decay; and \( \theta_0 \) is the surface directionality.

Figure 5.14 shows EM bias plotted versus footprint radius for three different surface roughnesses (\( \lambda = 10, 30, 100 \)). Values of \( p = 3.9, \theta_0 = \frac{\pi}{16} \), and \( k_o = 1 \) are used. The footprint size of the GME and BSE Ku band scatterometers were approximately 1.6 m and 1.7 m, corresponding roughly to a 0.8 m radius. The footprint radius of
Figure 5.15: Difference plot of $|e_{TOPEX}| - |e_{GME}|$ for LLR estimate. The magnitude of the difference decreases at higher significant wave heights (a rough surface). This supports a recent theoretical study.
Figure 5.16: Difference plot of $|\epsilon_{TOPEX}| - |\epsilon_{BSE}|$ for LLR estimate. The magnitude of the difference decreases at high wind speeds and significant wave heights (a rough surface). This supports a recent theoretical study.
the satellite is several orders of magnitude larger in comparison. From this plot, one notices that the EM bias reaches a limit beyond a footprint radius of 3 m. Thus, the mean difference of 0.94 cm or 1.3 cm between satellite and tower EM bias is possibly explained by the difference in footprint sizes. The larger magnitude difference may be accounted for by BSE’s slightly larger footprint radius. The plot suggests that for a typically rough surface ($\lambda = 30$), the difference between the EM bias for a 0.8 m radius footprint and an infinitely large one is roughly 2 cm.

Figure 5.14 also shows that as the surface becomes rougher for a given footprint radius (0.8 m), the difference between the magnitude of the estimated EM bias for the two footprint sizes decreases. At high wind speeds and large significant wave heights (a rough surface), the difference between the EM bias estimates would be the smallest. This theory is confirmed more strongly with BSE data than GME. (See Figs. 5.15 and 5.16).

### 5.3.5 Residual Biases

The estimated bias for several different NPR techniques was removed from the measured bias. The remaining residual bias was then observed for remaining correlation with other variables. The residual biases are plotted in Figs. 5.17 and 5.18. The x-axis on these plots corresponds to time, but has been left unlabeled due to the intervals of removed data points from tower interference. Refer to Fig. 2.10 for an actual time plot of the bias over the length of the GME.

From these figures and Table 5.2 one notes that the NPR techniques worked better for GME than for BSE. The average of the residual standard deviations, $\sigma_{RES}$, in both cases was 0.5484 cm and 0.8183 cm. The value of $\sigma_{RES}$ for the LLR technique are slightly lower than NWR in both experiments. An interesting observation is that while $\sigma_{RES}$ decreases from NWR to LLR, the correlation coefficient, $\hat{\rho}$, sometimes decreases simultaneously. This is counter-intuitive. This can be explained by the larger magnitude residual mean, maximum, and/or minimum residual bias. The correlation coefficient is a function of more than just the standard deviation of the residual. For this reason, both goodness measures of the EM bias estimate are used.
Figure 5.17: Residual biases for GME data. Subplots \(a - d\) correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan. in the \(U, H\) plane.
Figure 5.18: Residual biases for BSE data. Subplots a-d correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan. in the $U, H$ plane.
Table 5.2: Statistics of the residual biases and the correlation coefficient between measured and estimated EM bias for the GME and BSE. The table includes the improvement of NPR estimators over traditional parametric fitting.

<table>
<thead>
<tr>
<th>Method</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>$\sigma_{RES}$</th>
<th>$\hat{\rho}$ (corr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOM Exp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Param. Fit</td>
<td>-2.4311</td>
<td>2.2048</td>
<td>0.0000</td>
<td>0.6131</td>
<td>0.9651</td>
</tr>
<tr>
<td>NWR - Gauss.</td>
<td>-2.0073</td>
<td>2.3079</td>
<td>-0.0057</td>
<td>0.5584</td>
<td>0.9808</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-1.9353</td>
<td>2.3200</td>
<td>-0.0033</td>
<td>0.5551</td>
<td>0.9804</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-1.7849</td>
<td>2.4410</td>
<td>0.0085</td>
<td>0.5373</td>
<td>0.9784</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-1.7819</td>
<td>2.4416</td>
<td>0.0085</td>
<td>0.5428</td>
<td>0.9776</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>$\sigma_{RES}$</th>
<th>$\hat{\rho}$ (corr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOM Exp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Param. Fit</td>
<td>-3.1507</td>
<td>3.2322</td>
<td>0.0000</td>
<td>0.8855</td>
<td>0.9586</td>
</tr>
<tr>
<td>NWR - Gauss.</td>
<td>-2.9670</td>
<td>3.3971</td>
<td>-0.009</td>
<td>0.8272</td>
<td>0.9755</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-2.9680</td>
<td>3.4408</td>
<td>-0.0012</td>
<td>0.8243</td>
<td>0.9751</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-2.7563</td>
<td>3.4369</td>
<td>-0.0208</td>
<td>0.8101</td>
<td>0.9737</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-2.8066</td>
<td>3.4460</td>
<td>-0.0246</td>
<td>0.8116</td>
<td>0.9735</td>
</tr>
</tbody>
</table>

5.4 Summary

Satellite and tower LLR EM bias estimates are very similar. The differences between the two can possibly be explained with a recent theoretical study by Johnson [23]. The magnitude of the satellite EM bias is larger than that of both GME and BSE. This difference was shown to be due to differences in footprint size. In addition, for BSE the magnitude of the difference between the two decreases as the surface becomes rougher. This is also in support of the recent theoretical work.

The NPR techniques performed better than the parametric fit given in the previous chapter for both GME and BSE. For GME the parametric fit leads to $\sigma_{RES}$ of 0.6131 cm while the average for the four NPR techniques is 0.5484 cm. The correlation coefficient for Ku-band GME data using the parametric fit is 0.9651 while the average for the four NPR estimates is 0.9793. For BSE the $\sigma_{RES}$ values are larger, indicating
that the bias is not estimated as well as for GME. However, the improvement factors are very similar. Using NPR techniques, $\sigma_{RES}$ reduces by approximately 0.06 cm (a factor of 10%) and the correlation coefficient increases approximately 1.5% for both data sets. This is an improvement over the parametric fit.
Chapter 6

Nonparametric EM Bias Estimation with Wave Slope

A recent study by Melville shows that EM bias is strongly correlated to wave slope [10]. In this chapter the derivation of this parameter is given and applied to the GME data set. Since wave slope was already processed for the BSE data set, it can be used to compare the results obtained for GME. Then estimations of the EM bias are produced using the NPR technique outlined in the previous chapter. Wave slope is paired with $H_{\frac{1}{3}}$ and $U$ and then it is determined which combination of variables performs a better NWR or LLR estimation of the bias.

6.1 Wave Slope

The method outlined by Cox and Munk is used to derive characteristic wave slopes from the available raw GME data [24]. First, the deep water dispersion relation is used to estimate the wave number, $k$,

$$k \approx \frac{(2\pi f)^2}{g}$$

(6.1)

where $g = 9.8 \text{ m/s}^2$ is the gravitational constant. This assumption is valid because the water depth for GME was 40 m and for BSE it was 57 m deep. Then the wave displacement spectrum, $\Phi(f)$, is obtained by computing the power spectral density of the wave displacement vector, $\eta$. A typical one-hour average of $\Phi(f)$ is shown in Fig. 6.1. The average $U$ and $H_{\frac{1}{3}}$ for this example are 2.2 m/s and 1.5 m respectively, which are typical values for GME. Using Eq. (6.1) and the wave displacement frequency spectrum, the slope spectrum is adapted from Cox and Munk [24] as

$$S(f_i) = \frac{(2\pi f_i)^4}{g^2} \Phi(f_i).$$

(6.2)
Figure 6.1: Wave displacement frequency spectrum, $\Phi(f)$, and Slope spectrum, $S(f)$ for one-hour of wave displacement data. The average $U$ and $H^{1/3}$ are 2.2 m/s and 1.5 m respectively.

The rms slope, $s$, is then calculated using a discrete method,

$$ s = \left[ \frac{f_s}{N_{FFT}} \sum_{i=1}^{N} \frac{(2\pi f_i)^4}{g^2} \Phi(f_i) \right]^{1/2} \tag{6.3} $$

where $N_{FFT}$ corresponds to the number of points in the FFT and $f_s$ is the sample frequency. Because the derivation given by Cox and Munk estimates the mean squared wave slope, square-rooting the sum gives the rms wave slope. One rms wave slope value is calculated for each ten minute record of wave displacement data, and six are averaged to obtain the one-hour slope average corresponding to the same one-hour averages of $U$, $H^{1/3}$, and EM bias. Again, this requires reprocessing the raw data. The number of available one-hour records was 1175 of the original 1280. As mentioned in Chapter 2, this 8.2% loss of data corresponds to the corrupted data which was unavailable.
The BSE data provided by Kendall Melville was already processed and wave slope values had already been calculated. Melville states that one of the main difficulties in determining rms wave slope is the appropriate choice of a cut-off frequency [10]. He chooses a cutoff of 0.8 Hz because only slope contribution from waves comparable in scale or longer than the diameter of the footprint of the scatterometer are desired for the analysis. This cut-off can be seen naturally in both spectrums of Fig. 6.1 for the GME data. For both GME and BSE data the slope was estimated with a lower cutoff of 0 Hz. When the two slope estimates were compared, the GME slopes were nearly a factor of two greater than the BSE slopes. There is one main difference in the method used to obtain the slope estimates for these two experiments – the constant \((f_s/N_{\text{FFT}})\) in Eq. (6.3). Melville uses \(1/N\) where \(N\) is equal to \((f_N N_{\text{FFT}})/f_s\). Since the final result is square rooted, this means that if the wave conditions were identical for both experiments, BSE slopes should be \((1/f_N)^{1/2} = 1.12\) larger than GME slopes. Instead, the GME slopes are approximately a factor of 2 larger than the BSE slopes. This means that the wave conditions were not identical for the two experiments. Since the estimate of EM bias using the NPR methods will be unaltered by doing so, the GME slope data was divided by two such that the comparison would allow the grid to be on the same axis. Figure 6.2 shows the dependence of EM bias on this calculated wave slope for GME and BSE. For comparison, Fig. 6.3 shows the dependence of normalized EM bias on wave slope. One notes that normalized EM bias is linearly correlated to the wave slope for each experiment. The BSE has 1535 total one-hour averages while GME has 1175. Because BSE has 360 more data points than GME, the limits of the slope versus bias are more defined.

6.2 Results of using Wave Slope with NPR Bias Estimates

Since the derived NPR techniques are two-dimensional and because this is the first use of rms slope with these techniques, three different EM bias estimates are calculated in this section. The first EM bias estimate uses rms slope \((s)\), and \(H\). Then \(s\) and \(U\) are used. Finally, all three parameters are used by estimating normalized bias, \(\beta\), against \(s\) and \(U\).
Figure 6.2: EM bias dependence on rms wave slope for GME (above) and BSE (below). BSE has 364 more data points than GME.
Figure 6.3: Normalized EM bias dependence on rms wave slope for GME (above) and BSE (below). BSE has 364 more data points than GME.
6.2.1 Estimation with Wave Slope and Significant Wave Height

An appropriate grid spacing for \( s \) needs to be chosen before the NPR techniques can be implemented. Figure 6.4 shows the data densities for slope versus significant wave height for GME and BSE. The chosen spacing for \( s \) is \( ds = 0.0025 \). This may seem small, but the limits for \( s \) are \([0.025, 0.15]\), and that makes 51 equally spaced grid points. The new grid spacing for \( H \frac{1}{3} \) is \( dH = 0.125 \). This is smaller than the previously chosen spacing because it is not being compared to satellite data. Figures 6.5-6.12 show the results for GME and BSE data. The residual biases are shown in Figs. 6.13 and 6.14 for GME and BSE respectively. The blip in LLR-Epan occurred at \( H \frac{1}{3} = 2.66, s = 0.1019 \). The measured EM bias is -10.3 cm and the estimated is -4.0 cm. Looking at Fig. 6.8 reveals that this estimate must lie on the boundary where it is falling off towards zero, because the nearest EM bias contour is -9.5 cm which would result in a smaller residual. This is a problem with the Epanechnikov kernel which the Gaussian does not exhibit.

The statistics of the residual bias for these estimates are shown in Table 6.1. It can be seen that this method is a considerable improvement over the non-slope technique for BSE, while it is a smaller improvement for GME. For BSE the standard deviation in the residual, \( \sigma_{RES} \), is reduced by a factor of 2.1, while for GME it is only reduced by a factor of 1.2. This is a strong improvement over the NPR technique using \( U, H \) in the previous chapter. In general values of \( \sigma_{RES} \) are reduced by 0.021 cm from NWR to LLR, except for the LLR-Epanechnikov method. The correlation coefficients improve by 0.5-2.0% for both experiments. One notes that for BSE the correlation coefficient increases from NWR to LLR because both the magnitude of the mean and the standard deviation decrease. In the \( \epsilon(s, H) \) plane, the best method for GME and BSE is LLR-Gauss, based on \( \sigma_{RES} \) as a primary and \( \rho \) as a secondary indicator.

6.2.2 Estimation with Wave Slope and Wind Speed

Figure 6.15 shows the data densities for slope versus wind speed for GME and BSE. The grid spacing for wind speed is \( dU = 0.25 \). The wave slope grid spacing is the same as previously, \( ds = 0.0025 \). Using these grid spacings, the EM bias was
estimated in $s, U$ space. Figures 6.16-6.23 show the results for GME and BSE data. This EM bias estimate was then removed from the actual bias values in order to observe the residual bias; these are shown in Figs. 6.24 and 6.25 for GME and BSE respectively.

The statistics of the residual biases are shown in Table 6.2. The values of $\sigma_{\text{RES}}$ did not decrease using only $s$ and $U$, in fact, they increased significantly. The results show that using $s$ with $U$ is not as good as using $H_{\frac{1}{3}}$ and $U$. Not using $H_{\frac{1}{3}}$ is unwise since it is so strongly related to EM bias.

### 6.2.3 Estimation with Wave Slope, Wind Speed and Significant Wave Height

In order to apply the two-dimensional NPR estimate outlined in this thesis to EM bias in three dimensions, it is necessary to estimate normalized bias, $\beta$, in the $s, U$ plane. Because the data density clearly does not change (See Fig. 6.15), it is not replotted. The normalized EM bias estimate is multiplied by $H_{\frac{1}{3}}$ to obtain the actual EM bias estimate for comparison to the other methods outlined in this chapter.

After converting the estimated normalized bias, $\beta$, to EM bias the residual bias is computed and plotted in Figs. 6.34 and 6.35. The statistics of this residual bias are shown in Table 6.3. Once again, this method performs better with BSE than GME. In fact, the improvement for GME isn’t as good as the $s, H_{\frac{1}{3}}$ NPR methods (without using $U$). A possible explanation for this is given in the chapter summary. The values of $\sigma_{\text{RES}}$ for BSE decrease by a factor of 3.0 over non-slope NPR techniques, while the improvement for GME is a factor of 1.2. The same type of trend is observed in the correlation coefficients.
Figure 6.4: Data densities in the significant wave height and wave slope plane. Includes hourly average values of the data from the Gulf of Mexico Experiment (above) and the Bass Strait Experiment (below).
Figure 6.5: GME: NWR-Gaussian bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.6: GME: NWR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.7: GME: LLR-Gaussian bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.8: GME: LLR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s$, $H$ plane.
Figure 6.9: BSE: NWR-Gaussian bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.10: BSE: NWR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.11: BSE: LLR-Gaussian bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.12: BSE: LLR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s, H$ plane.
Figure 6.13: Residual biases after removing $s, H$ estimates for GME data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss, NWR-Epan, LLR-Gauss, and LLR-Epan.
Figure 6.14: Residual biases after removing $s, H$ estimates for BSE data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan.
Table 6.1: Statistics of the residual biases for the Gulf of Mexico and Bass Straits Experiments after removing $s,H$ NPR estimates.

<table>
<thead>
<tr>
<th>Method</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>$\sigma_{RES}$</th>
<th>$\hat{\rho}$ (corr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWR - Gauss.</td>
<td>-2.2482</td>
<td>1.5467</td>
<td>-0.0038</td>
<td>0.4582</td>
<td>0.9860</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-2.2159</td>
<td>1.5759</td>
<td>-0.0028</td>
<td>0.4586</td>
<td>0.9856</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-2.1211</td>
<td>1.5441</td>
<td>-0.0066</td>
<td>0.4366</td>
<td>0.9854</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-2.1772</td>
<td>6.3311</td>
<td>0.0022</td>
<td>0.4767</td>
<td>0.9822</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>$\sigma_{RES}$</th>
<th>$\hat{\rho}$ (corr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWR - Gauss.</td>
<td>-1.7483</td>
<td>1.4876</td>
<td>-0.0040</td>
<td>0.3961</td>
<td>0.9929</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-1.7900</td>
<td>1.5362</td>
<td>-0.0035</td>
<td>0.3965</td>
<td>0.9928</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-1.6359</td>
<td>1.4591</td>
<td>0.0031</td>
<td>0.3745</td>
<td>0.9932</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-1.6419</td>
<td>1.5319</td>
<td>0.0029</td>
<td>0.3757</td>
<td>0.9932</td>
</tr>
</tbody>
</table>
Figure 6.15: Data densities in the wind speed and wave slope plane. Includes hourly average values of the data from the Gulf of Mexico Experiment (above) and the Bass Strait Experiment (below).
Figure 6.16: GME: NWR-Gaussian bias estimates. EM bias contours are in centimeters in the $s,U$ plane.
Figure 6.17: GME: NWR-Epanechnikov bias estimates. EM bias contours are in centimeters in the \( s, U \) plane.
Figure 6.18: GME: LLR-Gaussian bias estimates. EM bias contours are in centimeters in the $s, U$ plane.
Figure 6.19: GME: LLR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s, U$ plane.
Figure 6.20: BSE: NWR-Gaussian bias estimates. EM bias contours are in centimeters in the $s,U$ plane.
Figure 6.21: BSE: NWR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s, U$ plane.
Figure 6.22: BSE: LLR-Gaussian bias estimates. EM bias contours are in centimeters in the $s, U$ plane.
Figure 6.23: BSE: LLR-Epanechnikov bias estimates. EM bias contours are in centimeters in the $s, U$ plane.
Figure 6.24: Residual biases after removing $s, U$ estimates for GME data. Subplots a - d correspond to the removal of EM bias using NWR-Gauss, NWR-Epan, LLR-Gauss, and LLR-Epan.
Figure 6.25: Residual biases after removing $s, U$ estimates for BSE data. Subplots a - d correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan.
Table 6.2: Statistics of the residual biases for the Gulf of Mexico and Bass Straits Experiments after removing $s,U$ NPR estimates.

<table>
<thead>
<tr>
<th>Method</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>$\sigma_{RES}$</th>
<th>$\hat{\rho}$ (corr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf of Mexico Exp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NWR - Gauss.</td>
<td>-4.3867</td>
<td>1.9045</td>
<td>-0.0262</td>
<td>0.8006</td>
<td>.9700</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-4.3881</td>
<td>1.8942</td>
<td>-0.0232</td>
<td>0.7944</td>
<td>.9695</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-4.3308</td>
<td>2.1393</td>
<td>0.0303</td>
<td>0.7309</td>
<td>.9664</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-4.3380</td>
<td>2.1274</td>
<td>0.0319</td>
<td>0.7353</td>
<td>.9658</td>
</tr>
<tr>
<td>Bass Straits Exp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NWR - Gauss.</td>
<td>-3.7813</td>
<td>2.1380</td>
<td>-0.0181</td>
<td>0.9120</td>
<td>0.9738</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-3.8307</td>
<td>2.1486</td>
<td>-0.0159</td>
<td>0.9085</td>
<td>0.9731</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-3.9131</td>
<td>2.5034</td>
<td>0.0310</td>
<td>0.8388</td>
<td>0.9714</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-3.9605</td>
<td>2.4774</td>
<td>0.0267</td>
<td>0.8433</td>
<td>0.9711</td>
</tr>
</tbody>
</table>
Figure 6.26: GME: NWR-Gaussian normalized bias estimates. Normalized EM bias contours are a negative percent of $H^2_4$ in the $s, H$ plane.
Figure 6.27: GME: NWR-Epanechnikov normalized bias estimates. Normalized EM bias contours are a negative percent of $H^3_3$ in the $s, H$ plane.
Figure 6.28: GME: LLR-Gaussian normalized bias estimates. Normalized EM bias contours are a negative percent of $H^{\frac{1}{3}}$in the $s, H$ plane.
Figure 6.29: GME: LLR-Epanechnikov normalized bias estimates. Normalized EM bias contours are a negative percent of $H_{3}^{1}$ in the s, $H$ plane.
Figure 6.30: BSE: NWR-Gaussian bias estimates. EM normalized bias contours are in centimeters in the $s, H$ plane.
Figure 6.31: BSE: NWR-Epanechnikov bias estimates. EM normalized bias contours are in centimeters in the $s, H$ plane.
Figure 6.32: BSE: LLR-Gaussian bias estimates. EM normalized bias contours are in centimeters in the $s, H$ plane.
Figure 6.33: BSE: LLR-Epanechnikov bias estimates. EM normalized bias contours are in centimeters in the $s, H$ plane.
Figure 6.34: Residual biases after converting normalized bias to EM bias and removing the $s,U$ estimates for GME data. Subplots $a$ - $d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan.
Figure 6.35: Residual biases after removing $s, U$ estimates for BSE data. Subplots $a - d$ correspond to the removal of EM bias using NWR-Gauss., NWR-Epan., LLR-Gauss., and LLR-Epan.
Table 6.3: Statistics of the residual biases for the Gulf of Mexico and Bass Straits Experiments after removing $s,U$ NPR estimates.

<table>
<thead>
<tr>
<th>Method</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>$\sigma_{RES}$</th>
<th>$\hat{\rho}$ (corr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf of Mexico Exp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NWR - Gauss.</td>
<td>-1.8910</td>
<td>1.9755</td>
<td>0.0153</td>
<td>0.4718</td>
<td>0.9822</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-1.9155</td>
<td>1.9854</td>
<td>0.0169</td>
<td>0.4734</td>
<td>0.9820</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-1.8843</td>
<td>1.9507</td>
<td>0.0217</td>
<td>0.4608</td>
<td>0.9825</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-1.8987</td>
<td>1.9684</td>
<td>0.0239</td>
<td>0.4631</td>
<td>0.9823</td>
</tr>
<tr>
<td>Bass Straits Exp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NWR - Gauss.</td>
<td>-1.0160</td>
<td>1.0424</td>
<td>0.0068</td>
<td>0.2788</td>
<td>0.9960</td>
</tr>
<tr>
<td>NWR - Epan.</td>
<td>-1.0172</td>
<td>1.0504</td>
<td>0.0061</td>
<td>0.2776</td>
<td>0.9961</td>
</tr>
<tr>
<td>LLR - Gauss.</td>
<td>-1.0247</td>
<td>0.9690</td>
<td>-0.0128</td>
<td>0.2675</td>
<td>0.9963</td>
</tr>
<tr>
<td>LLR - Epan.</td>
<td>-1.0265</td>
<td>0.9761</td>
<td>-0.0103</td>
<td>0.2686</td>
<td>0.9962</td>
</tr>
</tbody>
</table>
6.3 Summary

Tables 6.4 and 6.5 provide a summary of all the EM bias estimates presented in this thesis. The first column is an explanation of the NPR method used - $\epsilon(x, y)$ indicates that EM bias was estimated while $\beta(x, y)$ indicates that EM bias was calculated from an estimate of normalized bias. A star in the second column indicates whether wave slope was used. The third and forth columns are the correlation coefficient between the estimate and the measured EM bias, $\rho$, and the standard deviation of the residual bias, $\sigma_{RES}$. Table 6.4 includes the parametric fit comparison data from Chapter 4. Both tables are ordered from the largest to the smallest value of $\sigma_{RES}$. One notes immediately that for both GME and BSE the first four bias estimates using $s$ and $U$ without $H_{1/3}$ were the poorest - resulting in EM bias estimates that were worse than the parametric fit in Chapter 4. It is necessary to include $H_{1/3}$ in any EM bias estimate because it is so strongly correlated to EM bias (regardless of the use of wave slope). The LLR method is slightly better than the NWR method in most cases. In general there is a greater improvement factor between different two-dimensional planes than between different NPR estimates and kernel functions. These improvement jumps are greater for BSE than from GME.

The results show that for GME the lowest standard deviation in the residual occurs with the LLR-Gauss estimate in the $(s, H)$ plane and the highest correlation coefficient is with the NWR-Gauss, in the same plane. With these estimates $\sigma_{RES}$ decreases 0.18 cm and $\hat{\rho}$ increases 2.1% from the traditional parametric fit. The best method for BSE is the normalized EM bias estimate using LLR-Gauss, in the $(s, U)$ plane. It is noted that the same method (LLR-Gauss) in different planes is the best EM bias estimate for GME and BSE. For GME using all three parameters was not as good as using only two $(s, H)$. For BSE, however, all three parameters produce the best results. These results indicate that $s$ and $H$ are most strongly correlated to the bias and thus, the most valuable parameters for EM bias estimation. However, $U$ does provide a small amount of improvement beyond using just $s$ and $H$.

The use of wave slope resulted in a larger improvement factor with BSE than for GME. One reason may be due to the way in which the slopes were calculated.
Table 6.4: Summary statistics of all GME NPR estimates. The first column is an explanation of the method used (All are NPR except the one labeled Param. Fit) and the \((x,y)\) plane in which EM bias estimates were computed. The second column indicates whether or not the wave slope parameter was used. The third and forth columns are the correlation coefficient between the estimate and the measured EM bias and the standard deviation of the residual bias.

<table>
<thead>
<tr>
<th>Method</th>
<th>Est((x,y))</th>
<th>Slope?</th>
<th>(\hat{\rho}) (corr.)</th>
<th>(\sigma_{RES})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWR-Gauss.</td>
<td>(\epsilon(s,U))</td>
<td>*</td>
<td>0.9700</td>
<td>0.8006</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>(\epsilon(s,U))</td>
<td>*</td>
<td>0.9695</td>
<td>0.7944</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>(\epsilon(s,U))</td>
<td>*</td>
<td>0.9658</td>
<td>0.7353</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>(\epsilon(s,U))</td>
<td>*</td>
<td>0.9664</td>
<td>0.7309</td>
</tr>
<tr>
<td>Param. Fit</td>
<td>(\epsilon(U,H))</td>
<td></td>
<td>0.9651</td>
<td>0.6131</td>
</tr>
<tr>
<td>NWR-Gauss.</td>
<td>(\epsilon(U,H))</td>
<td></td>
<td>0.9808</td>
<td>0.5584</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>(\epsilon(U,H))</td>
<td></td>
<td>0.9804</td>
<td>0.5551</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>(\epsilon(U,H))</td>
<td></td>
<td>0.9776</td>
<td>0.5428</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>(\epsilon(U,H))</td>
<td></td>
<td>0.9784</td>
<td>0.5373</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>(\epsilon(s,H))</td>
<td>*</td>
<td>0.9822</td>
<td>0.4767</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>(\beta(s,U))</td>
<td>*</td>
<td>0.9820</td>
<td>0.4734</td>
</tr>
<tr>
<td>NWR-Gauss.</td>
<td>(\beta(s,U))</td>
<td>*</td>
<td>0.9822</td>
<td>0.4718</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>(\beta(s,U))</td>
<td>*</td>
<td>0.9823</td>
<td>0.4631</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>(\beta(s,U))</td>
<td>*</td>
<td>0.9825</td>
<td>0.4608</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>(\epsilon(s,H))</td>
<td>*</td>
<td>0.9856</td>
<td>0.4586</td>
</tr>
<tr>
<td>NWR-Gauss.</td>
<td>(\epsilon(s,H))</td>
<td>*</td>
<td><strong>0.9860</strong></td>
<td>0.4582</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>(\epsilon(s,H))</td>
<td>*</td>
<td>0.9854</td>
<td><strong>0.4366</strong></td>
</tr>
</tbody>
</table>

The algorithm to compute wave slope may be optimal for BSE and not for GME. The same frequency cutoffs were used in GME as in BSE (0 to 0.8 Hz). These cutoffs were chosen by Melville specifically for BSE, but perhaps they are not optimal cutoff limits for GME due to different physical wave conditions. With this assumption, wave slopes were recomputed with several different combinations of frequency cutoffs. Then the NPR method which performed the best (LLR-Gauss. \(\epsilon(s,H)\)) was implemented with each different slope estimate. The grid spacing was chosen to be small to minimize the interpolation error when estimating EM biases with the new solutions \((ds = 0.0025\)
Table 6.5: Summary statistics of all BSE NPR estimates. The first column is an explanation of the NPR method used. The second column indicates whether or not the wave slope parameter was used. The third and forth columns are the correlation coefficient between the estimate and the measured EM bias and the standard deviation of the residual bias.

<table>
<thead>
<tr>
<th>Method</th>
<th>Est(x,y)</th>
<th>Slope?</th>
<th>$\hat{\rho}$ (corr.)</th>
<th>$\sigma_{RES}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NWR-Gauss.</td>
<td>$\epsilon(s, U)$</td>
<td>*</td>
<td>0.9738</td>
<td>0.9120</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>$\epsilon(s, U)$</td>
<td>*</td>
<td>0.9731</td>
<td>0.9085</td>
</tr>
<tr>
<td>Param. Fit.</td>
<td>$\epsilon(U, H)$</td>
<td></td>
<td>0.9586</td>
<td>0.8855</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>$\epsilon(s, U)$</td>
<td>*</td>
<td>0.9711</td>
<td>0.8433</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>$\epsilon(s, U)$</td>
<td>*</td>
<td>0.9714</td>
<td>0.8388</td>
</tr>
<tr>
<td>NWR-Gauss.</td>
<td>$\epsilon(U, H)$</td>
<td></td>
<td>0.9755</td>
<td>0.8272</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>$\epsilon(U, H)$</td>
<td></td>
<td>0.9751</td>
<td>0.8243</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>$\epsilon(U, H)$</td>
<td></td>
<td>0.9735</td>
<td>0.8116</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>$\epsilon(U, H)$</td>
<td></td>
<td>0.9737</td>
<td>0.8101</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>$\epsilon(s, H)$</td>
<td>*</td>
<td>0.9928</td>
<td>0.3965</td>
</tr>
<tr>
<td>NWR-Gauss.</td>
<td>$\epsilon(s, H)$</td>
<td>*</td>
<td>0.9929</td>
<td>0.3961</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>$\epsilon(s, H)$</td>
<td>*</td>
<td>0.9932</td>
<td>0.3757</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>$\epsilon(s, H)$</td>
<td>*</td>
<td>0.9932</td>
<td>0.3745</td>
</tr>
<tr>
<td>NWR-Gauss.</td>
<td>$\beta(s, U)$</td>
<td>*</td>
<td>0.9960</td>
<td>0.2788</td>
</tr>
<tr>
<td>NWR-Epan.</td>
<td>$\beta(s, U)$</td>
<td>*</td>
<td>0.9961</td>
<td>0.2776</td>
</tr>
<tr>
<td>LLR-Epan.</td>
<td>$\beta(s, U)$</td>
<td>*</td>
<td>0.9962</td>
<td>0.2686</td>
</tr>
<tr>
<td>LLR-Gauss.</td>
<td>$\beta(s, U)$</td>
<td>*</td>
<td>0.9963</td>
<td>0.2675</td>
</tr>
</tbody>
</table>

and $dH = 0.025$). The results are shown in Table 6.6 and plotted in Fig. 6.36. The lower cutoff frequencies 0.0 Hz and 0.2 Hz are noted by solid and dashed lines respectively. Values of $\sigma_{RES}$ and $\hat{\rho}$ are plotted against the upper cutoff frequency. It is noted that lower and upper cutoff frequencies of 0.0 Hz and 0.4 Hz give the best results, but the improvement factor is small. The large difference in the overall improvement for BSE and GME is still unexplained.
Figure 6.36: Plot of Improvement in LLR-Gaussian as a function of the upper cutoff frequency used to compute wave slope. The upper plot shows the change in residual standard deviation, $\sigma_{RES}$, and the lower plot shows change in correlation coefficient, $\hat{\rho}$. The solid and dashed lines correspond to lower cutoff frequencies of 0.0 and 0.2 Hz. This is a plot of the values in Table 6.6.
Table 6.6: LLR-Gaussian results for GME with varying wave slope calculations. This table shows the correlation coefficient and the standard deviation of the residual bias for different slope estimates. The slope estimates are calculated using the specified upper and lower frequency cutoffs in the table in Eq. (6.3). The NPR method used is LLR-Gauss. $\epsilon(s,H)$, the best method for GME from Table 6.4.

<table>
<thead>
<tr>
<th>Cutoff Fr. (Hz)</th>
<th>Performance</th>
<th>lower</th>
<th>upper</th>
<th>$\hat{\rho}$ (corr.)</th>
<th>$\sigma_{RES}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td></td>
<td>0.3</td>
<td>0.4</td>
<td>0.98301</td>
<td>0.403</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4</td>
<td>0.4</td>
<td>0.98546</td>
<td>0.4075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.6</td>
<td>0.98491</td>
<td>0.4151</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.6</td>
<td>0.7</td>
<td>0.98405</td>
<td>0.4266</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
<td>0.8</td>
<td>0.98361</td>
<td>0.4323</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.9</td>
<td>0.98340</td>
<td>0.4350</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>1.0</td>
<td>0.98330</td>
<td>0.4363</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>1.1</td>
<td>0.98324</td>
<td>0.4371</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1</td>
<td>0.3</td>
<td>0.98312</td>
<td>0.4387</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2</td>
<td>0.4</td>
<td>0.98491</td>
<td>0.4149</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.6</td>
<td>0.98432</td>
<td>0.4229</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.6</td>
<td>0.7</td>
<td>0.98366</td>
<td>0.4316</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
<td>0.8</td>
<td>0.98332</td>
<td>0.4360</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8</td>
<td>0.9</td>
<td>0.98316</td>
<td>0.4381</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9</td>
<td>1.0</td>
<td>0.98307</td>
<td>0.4393</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>1.1</td>
<td>0.98301</td>
<td>0.4400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1</td>
<td>0.4</td>
<td>0.98295</td>
<td>0.4407</td>
</tr>
</tbody>
</table>
Chapter 7

Conclusions

This chapter summarizes the main results of this thesis. The specific contributions to the field of EM bias estimation for radar altimetry are presented. Finally, possible applications of the results are given.

7.1 General Summary

This thesis presents the results of several data analysis techniques which add to the knowledge of EM bias and help improve EM bias estimation techniques. In Chapter 3 the answer to why the bias rolls off at high wind speeds was given. The results support the hypothesis that the wave profile becomes more uniformly uneven at high wind speeds. The centroid of power versus the phase of the wave does move from the center of the wave phase during the experiment causing the EM bias to vary, but the shift is uncorrelated to high wind speeds.

Tower and satellite data were shown to be very similar. The differences which exist in the magnitude of the estimates were compared to a recent theoretical study. The data from both tower experiments supports the theory that a large footprint size results in a larger magnitude EM bias. In support of the same theory, it was also shown that as the ocean surface became more rough the difference between the two EM biases decreases.

Several methods for estimating the EM bias were presented. A traditional parametric method is compared to two NPR techniques, NWR and LLR. In both cases there is an improvement over using a parametric fitting equation, except when $H_3$ is not used. The amount of improvement is shown in Tables 6.4 and 6.5. In both
cases the best estimate includes wave slope. The best method for GME was shown to be LLR-Gauss, in the $s, H$ plane. For BSE it was shown to be the normalized EM bias estimate LLR-Gauss, in the $s, U$ plane. These NPR estimates reduce the standard deviation in the leftover bias and are also more highly correlated to the original measurement of EM bias.

### 7.2 Contributions

This thesis has makes three primary contributions to the study of EM bias. The first is the knowledge provided from the spatial analysis of the wave profile and the backscattered power. This study provides knowledge about the characteristics of the backscattered power versus the phase of the wave over the length of the Gulf of Mexico tower experiment. Two EM bias altering phenomenons were observed. This study provides unique information as it is the first time that spatial wave analysis (SWA) has been used to study EM bias.

The second contribution of this thesis is the application of nonparametric regression NPR techniques to two experimental data sets from the Gulf of Mexico (GME) and Bass Straits (BSE). It was shown that using the nonparametric regression (NPR) techniques, NWR and LLR, considerable improvement is made in estimation of the EM bias. Using wave slope as an input to NWR and LLR in addition to wind speed and significant wave height provided the best EM bias estimates.

An additional contribution provided by this thesis is the comparison of tower to satellite measured EM bias. This study compares NWR and LLR estimated EM biases for TOPEX/Poseidon to tower measurements from GME and BSE. The LLR EM bias estimates for tower and satellite data provided a good comparison. It showed that the EM bias measured from a satellite is larger in magnitude than that measured from a tower. This is a confirmation to a recent theoretical study predicting the differences in EM bias based on the difference in footprint size of the radars.
7.3 Future Work

Upon completing this thesis, many ideas for future work come to mind. The confirmation of recent theoretical work by Johnson may provide a means for computing the expected difference in EM bias measured from a tower and satellite. Also, the knowledge gained through the spatial analysis may be further studied to better understand the EM bias.

Because it has been shown that nonparametric estimation can provide significant improvement over parametric fitting, it may provide the improved performance factor required by Jason-1 in May of 2000. More work must be done to implement NPR estimation to satellite data. The work provided in this thesis and that which has been done by others provides a solid foundation upon which more research may be done. Nonparametric regression estimation techniques share one problem with parametric fitting equations – they need to be trained with a measured EM bias estimate. The inherent problem with using a measurement as truth data is that errors specific to the system which collected the data are included. This probably accounts for some of the differences between GME and BSE. Perhaps the answer for Jason-1 lies in the possibility of a continually updated truth data set provided by an identical system on a tower in the ocean (like a ground station). Despite the method used to calibrate satellite measurements, it has been shown that nonparametric regression techniques, which utilize wave slope in addition to $U$ and $H_{\frac{1}{3}}$, are more successful than parametric fitting equations for EM bias estimation.
Appendix A

Two-dimensional Nonparametric Algorithm

In Chapter 5 the non-parametric regression algorithms were put forth. The Matlab code used to perform NWR and LLR with different kernel function is shown here. The function is called _np, short for fast and furious nonparametrics. The purpose for incorporating all the NP algorithms in one function is to simplify its use. I make the Matlab code available so that it may be used, improved, and added upon.

```
function [zzgrid, xxvec, yyvec, dzz] = np(x, y, z, t, k, dx, dy, xys0); 
% [est, xvec, yvec] = np(x, y, z, type, kernel, dx, dy, [x0 x1 y0 y1]);

% Non-parametric 2-D Regression Estimation. If you are familiar with ffgrid,
% this is the same concept with smoothing involved
%
% Inputs: Must input at least the five variables: x,y,z,type,kernel.
% X,Y,Z -> Randomly located Z values over an X,Y plane.
% i.e. Corresponding points of f(X(i),Y(i)) = Z(i) for all i.
%
% TYPE chooses the type of non-parametric technique
% to be used for regression estimation. Valid inputs are
% 0 - Nadaraya-Watson Estimator
% 1 - Local Linear Regression
% (These are explained by Simonoff: Smoothing Methods in Statistics)
% (More may be added later...)
%
% KERNEL chooses the kernel function to be used for the 2-D smoothing
% parameter. Valid inputs are:
% 0 - gaussian 
% 1 - epanechnikov 
% 2 - bivweight 
% 3 - triweight
% * the _+ means that only values > 0 are used.
%
% DX,DY -> spacing on X,Y grid over which estimated Z values will be returned. If omitted, default is 50 bins in each direction.
% If only dx is inputted, then it is assumed that DY=DX.
% If dx or dy is negative, then the variable is taken as the number of bins rather than a grid resolution.

% The vector containing the limits [x0 x1 y0 y1] can be padded with NaNs if only certain limits are desired. If only x0 and y1 are defined, the other limits will be chosen according to min/max of
```

133
% Output: The vectors XVEC and YVEC define the grid over which Z values
% have been estimated with EST.
% If no output arguments are given, FF_NP will output a plot with
% two subplots -> a 3D surface of the data and a contour plot.
% Requires kernel.m. Tested under Matlab 5.2.1.1420
% Written by Justin Smith @ Brigham Young University
% Thanks to Oyvind Breivik from University of Bergen, NORWAY
% for the original ffgird, which was used as a skeleton for
% this function.

DX = -.50; % default value
x = x(:);
y = y(:);
z = z(:);
xyz = NaN*ones(1,7);

if (nargin < 6)
    help ff_np
    disp('NOT ENOUGH INPUTS: Refer to above help.');
    return;
end
if (nargin==5)
dx = DX;
dy = DX;
xyz0 = floor(min(x));
end
if (nargin == 6 & length(dx) > 1) % This means dx and dy are omitted and dx
    dx = dx; % is actually xyz0.
end
if (nargin == 6 & length(dy) > 1) % This means dy is omitted, dx=dy, and dy
    dy = dy; % is actually xyz0.
end
if (nargin == 6 & length(dx) == 1) % This means that dx is dx, dx=dy, and there
    dy = dx; % is no xyz0.
end
if (real(dx) == 0)
dx = DX + dx;
end
pad = imag(dx);
dx = real(dx);

nxyz = length(xyz0);
xyz(1:nxyz) = xyz0;

if (isreal(xyz(4)))
    xyz(4) = ceil(max(y));
end
if (isreal(xyz(3)))
    xyz(3) = floor(min(y));
end
if (isreal(xyz(2)))
    xyz(2) = ceil(max(x));
end
if (isreal(xyz(1)))
    xyz(1) = floor(min(x));
end
x0 = xyz(1); x1 = xyz(2); y0 = xyz(3); y1 = xyz(4);
z0 = floor(min(z)); z1 = ceil(max(z));
if (dx < 0)
    dx = (x1 - x0)/abs(dx);
end
if (dy < 0)
    dy = (y1 - y0)/abs(dy);
end
xvec = [x0:dx:x1];
yvec = [y0:dy:y1];
[X,Y]=meshgrid(xvec,yvec);
xd1=[ones(size(x)) x y];

nx = length(xvec);
yy = length(yvec);
if t>0
    est=zeros(ny,nx,2+t);
else
    est=zeros(ny,nx);
end
n=length(x);
std_x=std(x);
std_y=std(y);
multipliers=1.089*[1.2.214;2.623;2.978;1.74];
C=multipliers(k+1); % choose the correct smoothing parameter for the kernel
hx=C*std_x*n.^(-1/5);
hy=C*std_y*n.^(-1/5);
for j=1:nx
    for i=1:ny
        w=kernel(x-X(i,j),y-Y(i,j),hx,hy,k);
        % Nadaraya Watson Estimator
        if t==0
            est(i,j)=sum(z.*w);
        else
            % Local Linear Regression estimator
            if t==1
                x2d=[zeros(size(x)) ones(size(x)) X(i,j) ones(size(x)) Y(i,j)];
                XD=xd1=x2d;
                if length(find(w))>0
                    est(i,j,:)=pinv(XD'*[v.*XD(:,1) v.*XD(:,2) v.*XD(:,3)]*XD'*[v.*w];
                else
                    est(i,j,:)=0;
                end
            % Nadaraya Watson Estimator used for all other t inputs for now.
            else
                est(i,j)=sum(z.*w);
            end
        end
    end
    est=est(:,:,1);
    if t==1
        dz2=zeros(size(est(:,:,2:3)));
        dz2=est(:,:,2:3);
    end
    Nil = (est == 0); % Empty grid points are set to default
    est(Nil) = pad;
end
if (nargout == 0) % no output, then plot
    figure
    subplot(1,2,1)
    surf(xvec, yvec, est);  
    dum=ceil((z1-z0)/h);
    axis([x0 x1 y0 y1 z0-dum z1+dum]);
This function calls the function *kernel* which is another *Matlab* code which was written to facilitate the NP estimates. This program implements the two-dimensional kernel function. The code is included below:

```matlab
function out=kernel(X1, X2, h1, h2, Ktype);
% function out=Kernel(X1, X2, h1, h2, Ktype)
%
% This evaluates the kernel function:
%  K = \(2\pi h_1 h_2\)^{-1} \exp\left(-\frac{X_1^2}{2h_1^2}\right) \exp\left(-\frac{X_2^2}{2h_2^2}\right);
% valid Ktypes so far:
% 0 --- 'gaussian'
% 1 --- 'epanechnikov'
% 2 --- 'biweight'
% 3 --- 'triweight'
%
% This function is needed by ff_np. Normally X1=x-Xi, X2=y-Yi for vectors
% x, y, Xi, Yi the same size (where Xi and Yi are one value).
% See ff_np.m for more information. For more information about Kernel
% functions see Smoothing Methods in Statistics by Simonoff, [1996].
% Justin Smith

K=0;
if Ktype==0, K=exp(-(X1).^2./(2*h1^2))*exp(-(X2).^2./(2*h2^2));
elif Ktype==1, K=1-(X1./h1).^2-(X2./h2).^2;
elif Ktype==2, K=(1-(X1./h1).^2).^2*(1-(X2./h2).^2);;
elif Ktype==3, K=15*(X1./h1).^2.*(X2./h2).^2*exp(-(X1./h1).^2-(X2./h2).^2);;
end
```

136
K(find(K<0))=0;
elseif Ktype==2,
    K=(1-(X1/h1).^2-(X2/h2).^2);
    K(find(K<0))=0;
    K=K.*2;
elseif Ktype==3,
    K=(1-(X1/h1).^2-(X2/h2).^2);
    K(find(K<0))=0;
    K=K.*3;
else
    disp('Bad Ktype - see help on Kernel Function');
end
if length(find(K))~=0
    K=K./sum(K);
end
out=K;

Other algorithms which were used for the data analysis in this thesis are not included here but include the algorithm for computing the wave slope for GME and the algorithms for the spatial analysis of Chapter 3.
Bibliography


[16] Phillipp Gaspar, Françoise Ogor, Pierre-Yves Le Traon, and Ouan-Zan Zanife, “Estimating the sea state bias of the TOPEX and POSEIDON altimeters from


