UNCERTAINTIES IN OCEANIC MICROWAVE REMOTE SENSING: THE RADAR FOOTPRINT, THE WIND-BACKSCATTER RELATIONSHIP, AND THE MEASUREMENT PROBABILITY DENSITY FUNCTION

by

Paul E. Johnson

A dissertation submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

Brigham Young University

August 1999

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a dissertation submitted by

Paul E. Johnson

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

Date	David G. Long, Chair
Date	David V. Arnold
Date	Wynn C. Stirling
Date	Richard L. Frost
Date	Michael A. Jensen

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the dissertation of Paul E. Johnson 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 G. Long
Chair, Graduate Committee

Accepted for the Department

Michael D. Rice
Graduate Coordinator

Accepted for the College

Douglas M. Chabries
Dean, College of Engineering and Technology

ABSTRACT

UNCERTAINTIES IN OCEANIC MICROWAVE REMOTE SENSING: THE RADAR FOOTPRINT, THE WIND-BACKSCATTER RELATIONSHIP, AND THE MEASUREMENT PROBABILITY DENSITY FUNCTION

Paul E. Johnson

Electrical and Computer Engineering

Doctor of Philosophy

Oceanic microwave remote sensing provides the data necessary for the estimation of significant geophysical parameters such as the near-surface vector wind. To obtain accurate estimates, a precise understandingof the measurements is critical. This work clarifies and quantifies specific uncertainties in the scattered power measured by an active radar instrument.

While there are many sources of uncertainty in remote sensing measurements, this work concentrates on three significant, yet largely unstudied effects. With a theoretical derivation of the backscatter from an ocean-like surface, results from this dissertation demonstrate that the backscatter decays with surface roughness with two distinct modes of behavior, affected by the size of the footprint. A technique is developed and scatterometer data analyzed to quantify the variability of spaceborne backscatter measurements for given wind conditions; the impact on wind retrieval is

described in terms of bias and the Cramer-Rao lower bound. The probability density function of modified periodogram averages (a spectral estimation technique) is derived in generality and for the specific case of power estimates made by the NASA scatterometer. The impact on wind retrieval is quantified.

ACKNOWLEDGMENTS

I am grateful for the experience offered by an education at Brigham Young University. It is a place where extremely talented instructors enlighten minds by bringing valuable information to life and inspire the soul with their love of God. The time I have been able to study under professors like Dr. David Long, my doctoral advisor, and Dr. David Arnold, my masters advisor, will be a treasure to me throughout my life. My life has been enriched by friends at BYU with whom I could discuss both the aspects of science and the application of religion—I express my thanks to Travis Oliphant for many hours of discussion and insight.

My dear wife, Melissa, is an incredible driving force in my life. This dissertation would not have been, except for her support. Because of her, I want to achieve my greatest potential; and with her, I will.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Since radars were first observed to be affected by weather patterns in World War II, electromagnetic systems have been used to measure geophysical systems. Both active and passive systems, using various parts of the spectrum, have been used to monitor rainfall, deforestation, wind, ice mass, and many other crucial variables. Such parameters provide valuable data about our world and how it functions.

Oceanic microwave remote sensing is the field of study and application in which oceans are observed from a distance, say from aircraft or spacecraft, using microwave energy. The oceans, covering over 70% of the earth's surface, govern our weather and climate. In particular, the boundary interaction at the surface between the ocean and the atmosphere experiences tremendous fluxes of heat, momentum, carbon and oxygen. On a global scale, these drive weather patterns, climate changes, atmospheric and oceanic compositions; the impact on humankind ranges from the short-term daily weather to the long-term viability of life on earth.

Recognition of the importance of understanding the fluxes at the ocean surface has brought together several distinct areas of study. This has traditionally been the domain of oceanography and meteorology. Physicists, chemists and biologists have increased the scientific rigor of geophysical fluid dynamics, atmospheric and oceanic chemical compositions, and the sensitivity of lifeforms to minor changes in such dynamics and compositions. More recently, engineers have entered the field to provide advanced measurement techniques, technology, and system understanding.

The most common active radar systems used in ocean remote sensing are altimeters and scatterometers. While using the same basic radar principles, the two

types of instruments are designed for very different purposes. Altimeters are nadir looking instruments (i.e., they are pointed straight down, normal to the ocean surface) and use the time of flight of the radar signal to make their primary measurement: the distance between the surface and the instrument. Altimeters can also be used as nadir looking scatterometers. Scatterometers are pointed at an angle away from nadir (commonly 20° to 60°), and use the amount of backscattered power to estimate the normalized radar cross section, from which other parameters are estimated.

Remote sensing radars measure the reflective properties of a surface. The normalized radar cross section (NRCS), also called backscatter (σ^o), is calculated using the radar equation [1]:

$$\sigma^o = \frac{(4\pi)^3 R^4 L P_s}{P_t G^2 \lambda^2 A}, \tag{1.1}$$

where R is the slant range to the surface, P_t is the transmitted power, P_s is the received backscattered power, L represents known system losses, G is antenna gain, A is the effective illuminated area, and λ is the wavelength of the transmitted radiation. In practice, a backscatter power measurement is actually the sum of the backscattered power and noise due to instrumentation and the natural emissivity of the earth and atmosphere. The noise power is estimated and subtracted from the signal-plus-noise measurement to estimate P_s in order to compute the σ^o.

This work focuses on uncertainties in the scattered power measured by an active radar instrument over the ocean. This requires some incursions into the realms of non-engineering disciplines. For example, the ocean surface roughness is forced by the near-surface wind, modulated by long waves (swell), governed by the fluid dynamics of salt water of a particular temperature and composition, and complicated by the physics of a rotating planet. On the other end of the measurement process is the estimation of meaningful and usable parameters for inclusion in the body of scientific knowledge: global wind vector maps to drive climate models and numerical weather prediction, ocean surface height for developing sea-floor topographic maps, etc.

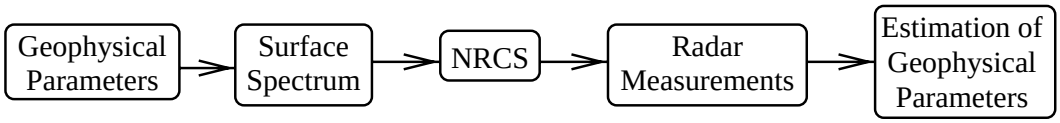


Figure 1.1: *Microwave remote sensing of the oceans provides the necessary data to estimate significant geophysical parameters, such as the vector wind. Such estimates are necessarily based on the measurements made by radars of the normalized radar cross section of the surface. The NRCS is determined by the shape and electromagnetic properties of the surface; for random surfaces like the ocean, it is determined in terms of the statistics of the surface (i.e., the surface spectrum). The surface spectrum is forced by the geophysical parameters. Current applications require significant approximations to model the processes leading to the parameter estimation. Even knowing the exact spectrum, the NRCS depends also on the relationship of the radar footprint to that spectrum. But the relationship between the geophysical parameters and the resultant spectrum is beyond the current knowledge of physical oceanography; approximate models are used to relate the most significant geophysical parameter, the vector wind, to the NRCS, ignoring all other influences. The measurements themselves introduce noise while modifying the statistics of the measurements through the analog processing.*

Between the forces that shape the ocean surface and the estimated geophysical parameters is a vast engineering realm to make appropriate, accurate measurements. It is the uncertainty in this realm which provides the thrust of this research. The relationship between the geophysical parameters of interest and the estimates of those parameters based on microwave remote sensing is graphically displayed in Fig. 1.1. The geophysical parameters (wind, rain, dielectric properties of the water, etc.) impart a particular statistical structure to the surface, described in terms of the surface power spectrum. The spectrum and water dielectric properties, along with the measurement geometry and electromagnetic wavelength, determine the scattering characteristics of the surface, quantified by the normalized radar cross section, NRCS, or backscatter coefficient, σ^o. The NRCS is related to measurements of the power scattered from the surface. And finally, estimates of significant geophysical parameters are made from the measurements.

Each step in this process is prone to error and uncertainty. While the surface spectrum is dictated by geophysical parameters, neither the precise relationship governing the spectrum nor all the parameters involved are known. Oceanographers have developed tunable models of the spectrum, providing several adjustable terms to establish realistic ocean spectra. Using a particular spectrum model, and including assumptions on the dielectric properties of the water and on the minimum radius of curvature of the surface, electromagnetic scattering theory provides the necessary theory to derive the NRCS. Still, identifying all of the surface and measurement parameters (e.g., spectrum parameters, wind forcing, swell, dielectric characteristics, radar altitude, incidence angle) that determine the theoretical scattering behavior, and the nature of their effects, is well beyond the current state of science. In order to avoid many of these uncertainties and still provide techniques for estimating at least some of the geophysical parameters (for example, the wind), a common approach has been to develop empirical models relating the vector wind (the dominant force in driving the surface spectrum) and the observation parameters to the NRCS. While this approach bypasses the gap caused by the lack of scientific understanding of the spectrum development, it is based on empirical measurements and statistical regression; empirical models do not incorporate non-wind factors of the spectrum and their accuracy is not understood. Whether the NRCS is developed from an exact knowledge of the spectrum or from an empirical model based on the wind, its value must be measured by the scatterometer, which involves considerable signal processing dependent on the radar design and the introduction of thermal noise. From the power measurements, statistical signal processing is employed to estimate the wind (and possibly other geophysical parameters), inverting the process under the constraint of some optimality criterion with the appropriate model.

A thorough exploration of all the issues involved in oceanic microwave remote sensing is well beyond the scope of a single Doctoral dissertation. Further, because several subtopics have been developed in detail by other researchers, this work has emphasized several important issues which affect the accuracy of NRCS measurements. In particular, I have focused on three related points of interest that have been

largely unstudied. I begin with an examination of the theoretical backscatter power from a surface with an assumed spectral shape to observe the dependence of the backscatter on typical spectrum parameters and especially on the size of the radar footprint. I develop a method to estimate the variability of empirical model functions relating the vector wind to the backscatter, based solely on the scatterometer measurements to quantify the uncertainty in the backscatter caused by unmodelled parameters. The digital processing involved in scatterometry modifies the statistics of the measurements, though the nature and impact of the modification have not been previously developed. I analyze the digital system to derive the probability density function of the measurements, in sufficient generality to be applicable in a wide variety of applications. The results are used to quantify the effects on the accuracy of wind retrieval. These areas of oceanic remote sensing measurement uncertainty and the contributions of this dissertation are described below.

1.1 Finite Footprint Effect on Backscatter

Empirical data collected from altimeters at different altitudes reveals intriguing differences: comparison of measurements from towers and from aircraft shows a distinct bias in the amount of power returned to the instrument. Scientists have struggled for a theoretical justification of this bias. The basis of electromagnetic scattering is rooted in Maxwell's equations and analysis techniques have been widely studied and applied. The Kirchoff approximation (also called the physical optics approximation) is a common simplification of scattering theory applied to random, rough surfaces. This approximation has been applied to surfaces in generality and to a large number of specific spectra. I derive the physical optics scattering from a simple, two-dimensional model of the ocean surface and examine the sensitivity of the backscatter to tunable parameters of the spectrum, as related to the radar footprint. While the surface roughness provides the dominant effect in determining the scattering coherency, the backscatter is quite sensitive to the power parameter of the power-law spectrum of the surface, particularly for values of the parameter less

than four. I find that using a finite footprint size causes the backscatter to experience two modes of asymptotic behavior as the surface roughness increases; that is, for a given surface roughness, the backscatter decays at different rates depending on the scatterometer footprint size. This bimodal behavior is magnified for directional surfaces because relevant surface information in the correlation function is stretched, thus requiring a larger footprint to capture it. In application, this implies that different scattering models are required for scattering from above and below the altitude threshold; use of a tower-based model with power measurements from a spaceborne radar (or vice-versa) will introduce errors.

1.2 Geophysical Modeling Error

Near-surface winds control energy, moisture and gas fluxes between the atmosphere and ocean; these fluxes are of vital importance as they drive weather, climate and global changes. Further, the wind is the dominant driving force of the surface roughness and spectrum, which affect the scatterometer measurements. Space-borne scatterometers provide unparalleled observation of the oceans on a global scale, generating invaluable data about the ocean state and oceanic and atmospheric circulation patterns.

The ability to infer wind from backscatter measurements presupposes a known relationship between the wind and the surface scattering characteristics. However, this relationship is not completely known; current geophysical model functions approximate the backscatter as a function of the electromagnetic frequency and polarization, the wind speed and azimuth direction relative to the radar antenna and the incidence angle, based on simple empirical techniques. Not included in the model function are temperature, salinity, and other factors. Because empirical estimates of the wind–backscatter relationship do not incorporate such non-wind influences they suffer from errors due to unmodelled parameters. Actual measurements exhibit variability due to these unmodelled effects, resulting in reduced wind retrieval accuracy. I have developed a simple method for estimating this variability from just scatterometer measurements and have explored the variability to determine its impact on wind

estimation and its dependence on additional measurable parameters. In this way, systematic factors in the backscatter, such as temperature and latitude, can be identified, helping researchers to focus their efforts on parameters which will improve the models. Further, this variability serves as an indirect measurement of these unmodelled parameters and can be incorporated as error bars on wind estimates.

1.3 The PDF of Scatterometer-Derived Wind Estimates

Spaceborne scatterometers, such as the recent NASA scatterometer, NSCAT, frequently use Doppler filtering to improve the resolution of the surface. The simple method of periodogram averaging is commonly used to estimate the power scattered from each resolution element of the surface, as described by a frequency band of the Doppler shift. Such processing modifies the statistics of the scattered signal. Estimation of geophysical parameters such as the vector wind requires accurate statistics of the power measurement. Wind estimation from scatterometer measurements, such as provided by NSCAT, relies on several NRCS measurements with different geometries. An objective function is formed and optimized based on the wind most likely to have generated the set of measurements; maximum-likelihood estimation and weighted least squares are common choices of estimators currently employed. However, estimators require, at least implicitly, a distribution for the measurements. While the Gaussian distribution has been assumed previously, I show that power estimates based on periodograms are not distributed Gaussian.

The NSCAT data collection includes spectrum estimates using Welch's technique. Welch's method for spectral estimation of averaging modified periodograms has been widely used for decades in many applications, though the statistics of such an estimate have not been developed. The pdf of a power estimate is derived here for an estimate based on an arbitrary number of frequency bins, overlapping data segments, amount of overlap and type of data window, given a correlated Gaussian input sequence. The pdf's of several cases are plotted and found to be distinctly non-Gaussian (the asymptotic result of averaging frequency bins and/or data segments), using the Kullback-Leibler distance as a measure. For limited numbers of frequency

bins or data segments, the precise pdf is considerably skewed and will be important in applications such as maximum-likelihood tests.

I also apply the pdf derivation to the specific case of NSCAT. Working from first principles and the design of the NSCAT signal processor I derive the distribution of the NSCAT measurements as a function of the surface NRCS, the signal to noise ratio and the cell number. The resulting distribution is skewed relative to the traditional Gaussian model. Simple compass simulations are used to compare the accuracy of winds estimated using the actual and Gaussian model distributions in order to quantify the bias caused to wind retrieval.

1.4 Dissertation Preview

Inherent in the wind measurement process are significant points of uncertainty. The electromagnetic backscatter depends critically on the unknown spectrum of the random surface. The surface spectrum is driven principally by the wind, but other effects impact the spectrum. The measurement process itself provides an estimate of the power; while the noise contributed by the radar is well understood, the digital processing involved in power estimation modifies the statistics of the estimate. In this dissertation I explore these fundamental issues. Background for each sub-topic is given in the appropriate chapter.

In Chapter 2 I develop the physical optics scattering from a random, perfectly conducting surface with a power-law spectrum as the theoretical version of scattering from the ocean. I quantify the sensitivity of the backscatter to various parameters of the spectrum and radar; in particular, I derive theoretical evidence for two asymptotic regimes of scattering based on the footprint size, suggesting differences in the scattering behavior for radars with different footprint size, such as those carried on aircraft compared to spacecraft.

In Chapter 3 I analyze a simple model which describes variability of the NRCS for given wind conditions and develop a method for quantifying that variability based solely on scatterometer measurements. I also process the NSCAT data set to estimate the variability constrained to narrow bins of wind speed, measurement

incidence angle, latitude, and time. I examine the impact of the variability on wind estimation and find that, on average, the wind estimate is not dramatically affected by the variability (this is expected because the empirical model functions are considered highly accurate), but the error bars, as quantified by the Cramer-Rao lower bound, are significantly increased by non-wind factors affecting the backscatter.

The effect of the digital processing on the signal power estimate is considered in Chapters 4 and 5. In Chapter 4 I derive the probability density function of a power estimate based on averaged, modified periodograms. The description is widely applicable in many engineering endeavors where Welch's spectral estimation technique is applied. In Chapter 5 I consider the resulting probability density function for power estimates based on the digital processing of NSCAT.

Finally, in Chapter 6, I describe some overall conclusions of the research. My contributions are identified within the larger context of the subject. I also describe avenues of research to extend this study.

Chapter 2

The Finite Radar Footprint

Microwave remote sensing is fundamentally dependent on electromagnetic scattering theory. A thorough understanding of the dominant causes of the theoretical backscatter is vital for a complete knowledge of the relationship between geophysical parameters and NRCS measurements. Several geophysical parameters contribute to the nature of the surface spectrum. However, the precise relationship is beyond the current state of science and that study in the realm of oceanography is outside the scope of this engineering dissertation.

In this chapter, the Kirchoff approximation (physical optics scattering) for the backscatter is expressed for general models of the two-dimensional surface correlation function from a perfectly conducting random surface. While the mathematics of the Kirchoff approximation have been well developed, the application to specific spectra has received limited attention and the identification of the critical parameters of a generalized spectrum and the measurement system have not been fully explored. Particularly intriguing is the consistent discrepancy between backscatter measurements made at low altitudes from towers and those made from aircraft and spacecraft [2]. Specifically addressed here, the identification of distinct modes of backscatter behavior are identified, based on the size of the radar footprint and the parameters of the measurement.

2.1 Overview

The basic model of the ocean surface suggested by physical oceanography and empirical studies describes an essentially perfectly conducting surface at

microwave frequencies with a spectrum based on a power-law decay from a low wavenumber limit. Frequently a separable spectrum is assumed with some form of directionality independent of the wavenumber magnitude. Energy is injected into the surface spectrum as short, capillary waves on the order of a few centimeters long with capillary action for the restoring force. As the wind continues to blow, non-linear interactions transfer the energy to longer and longer waves. These longer waves, restored by both gravity and capillary action, can sustain a larger amount of power in the spectrum. As a result, such a spectrum is commonly referred to as a red spectrum.

In Section 2.2, the well-known physical optics integral is considered, using asymptotic approximations as the surface roughness increases. The physical optics, or Kirchoff, approximation computes the NRCS of a random surface through use of the correlation function, assuming the limitations of physical optics scattering, but no other constraints on the surface roughness. For the special, simplistic, case of a separable spectrum with a low wave-number limit, the correlation function is derived explicitly and the asymptotic NRCS computed using the dominant behavior of the correlation function. While the applicability of this method is limited to the special spectrum described at nadir incidence, closed form solutions are derived which describe the various backscatter regimes. A recently developed technique for computing the asymptotic backscatter for an arbitrary, two-dimensional, power spectral density is also outlined. This method has wide applicability, but requires time-consuming numerical integrations with concerns about numerical stability and without the benefit of closed form solutions.

Section 2.3 analyzes the backscatter at nadir incidence from the simplistic spectrum, based on the traditional asymptotic approximation, as changes are made to the surface roughness, the spectrum power parameter and the surface directionality. The dominant effect on the backscatter, which is based on the coherence of the scattering, is due to the surface roughness. The power parameter of the ocean surface, p, has been argued to have a value of four and has been measured between 3.5 and 4.5. For $p < 4$ the backscatter depends strongly on the particular value of p, while for $p > 4$ the backscatter is only slightly affected. The surface's directionality

stretches the correlation function such that the effect of the finite footprint becomes obvious. At low (though asymptotically significant) roughnesses, the backscatter converges to one mode of behavior, or rate of decay. At higher roughnesses, where the surface is rough enough that all of the relevant surface variations are visible within the footprint, the backscatter converges to a final, steeper, rate of decay.

Section 2.4 applies the second backscatter computation method to a specific example of a more realistic surface spectrum. Similar trends of two modes of asymptotic decay are demonstrated, with the mode significantly affected by the footprint size.

Section 2.5 considers the effect of the footprint size on the electromagnetic bias. The EM bias is a difference between the actual mean sea level and the sea level identified by altimeters. The bias has been noted to have different values for altimeters mounted on platforms at different altitudes. Here, the scattering theory developed in the initial sections of this chapter are applied to show the effect of a finite footprint on the EM bias.

2.2 Physical Optics Scattering

In this section, the backscattering coefficient from a perfectly conducting rough surface having a Gaussian height distribution (a good approximation for the ocean surface [3]) with variance h^2 and surface height power spectral density $S(k, \theta)$ is determined. The well-known physical optics backscatter is developed with two asymptotic approximations as the surface roughness becomes large. Because the physical optics approximation uses it, the correlation function of the surface for typical ocean states is considered.

The physical optics approximation, also called the Kirchoff approximation, establishes a constraint on the variance of the random process: the radius of curvature at each point on the surface must be large compared to the electromagnetic wavelength [4]. Mathematically, this constraint is that $\rho_c \kappa \cos\theta_i \gg 1$, where ρ_c is the radius of curvature, κ is the electromagnetic wavenumber and θ_i is the incidence angle [5]. As the wavenumber of the surface increases, two effects occur: the radius of curvature

decreases and the amplitude of the surface variations decrease. A high frequency filter has been proposed to eliminate the frequency components of the surface which cause a small radius of curvature [6]. However, if the amplitude of the waves becomes less than the electromagnetic wavelength (so the scattering is coherent) before the radius of curvature requirement is violated, then the physical optics approximation is still applicable. This is the case for typical waves [7].

In the Kirchoff approximation, the backscattering coefficient is [8]

$$\sigma^o = \left(\frac{\kappa^2}{\pi \cos^2 \theta_i}\right) \int e^{i\vec{k}_b \cdot \vec{x}} \exp\{-\lambda[1 - C(\vec{x})]\}\, d\vec{x}. \tag{2.1}$$

where $\vec{\kappa}$ is the wavevector of the incident field, $\vec{k_b} = 2\vec{\kappa}_H$ is the Bragg wavenumber (twice the horizontal component of the electromagnetic wavenumber, $\vec{\kappa}_H$), and $\lambda = 4h^2\kappa^2 \cos^2 \theta_i$ provides a measure of the surface roughness. The correlation function, $C(\vec{x})$ is the inverse Fourier transform of the normalized surface height power spectral density $S(\vec{k})$, and is discussed in detail below.

At nadir incidence, $\theta_i = 0$, this reduces to

$$\sigma^o = \frac{\kappa^2}{\pi} \int \exp\{-\lambda[1 - C(\vec{x})]\}\, d\vec{x}. \tag{2.2}$$

2.2.1 Asymptotic Backscatter Approximations

Here, two asymptotic approximations of the physical optics backscatter are developed. First, a traditional asymptotic expansion of Eq. (2.2) is performed for backscatter from a nadir surface. The development uses realistic, closed forms for the correlation function which allow the backscatter to be determined explicitly. This simple, clear form allows analysis of the backscatter sensitivity to various parameters. Second, an alternate asymptotic expansion of the physical optics backscatter is described using multiple convolutions of the surface spectrum. While this form of the backscatter does not immediately reveal the backscatter dependencies, it achieves comparable numerical results and can be applied to any spectrum (satisfying the physical optics constraints) and can be applied at arbitrary incidence angles.

The asymptotic approximation of an integral of Laplace type, such as Eq. (2.2), is found by approximating $1 - C(\vec{x})$, about it's local minima [9]. Therefore, the

physical optics approximation at nadir incidence can be evaluated using the dominant behavior of the correlation function. Consider the class of correlation functions with dominant behavior described as

$$C_{dom}(\rho,\phi) = 1 - Ak_0^{\epsilon}\rho^{\epsilon}(a\cos 2\phi + b), \tag{2.3}$$

where the coefficients A, ϵ, a and b are functions of the spectral parameters, and the argument $\vec{x}$ has been expressed in polar coordinates ρ and ϕ. In Section 2.2.2, this equation for the dominant behavior of the correlation function is derived for a special, though representative, case of ocean spectra. If the integration is constrained to a finite range of a circular footprint of radius ρ_0, the asymptotic backscatter for normal incidence can be written as

$$\begin{aligned}
\sigma^o &= \frac{\kappa^2}{\pi}\int_0^{\rho_0}\int_0^{2\pi} \rho\ d\rho\ d\phi\ e^{-\lambda[1-C_{dom}(\rho,\phi)]} && (2.4)\\
&= \frac{\kappa^2}{\pi}\int_0^{\rho_0}\int_0^{2\pi} \rho\ d\rho\ d\phi\ e^{-\lambda A k_0^{\epsilon}\rho^{\epsilon}(a\cos 2\phi + b)} && (2.5)\\
&= 2\kappa^2\int_0^{\rho_0} d\rho\ \rho e^{-\lambda A b k_0^{\epsilon}\rho^{\epsilon}} I_0\left(\lambda A a k_0^{\epsilon}\rho^{\epsilon}\right) && (2.6)\\
&= \frac{2\kappa^2}{\epsilon k_0^2}\left(\lambda A\right)^{-\frac{2}{\epsilon}} b^{-\frac{2}{\epsilon}}\int_0^{u_0} du\ u^{\frac{2}{\epsilon}-1}e^{-u} I_0\left(\frac{a}{b}u\right) && (2.7)
\end{aligned}$$

where $u_0 = \lambda A b k_0^{\epsilon}\rho_0^{\epsilon}$. Note that if the limit on the integration extends to infinity, as is commonly assumed in asymptotic analysis (representing either infinite roughness or infinite footprint size), the integral is available in closed form

$$\sigma^o = \frac{2\kappa^2}{\epsilon k_0^2}\left(\lambda A\right)^{-\frac{2}{\epsilon}}\frac{b^{-\frac{2}{\epsilon}+1}}{\sqrt{b^2-a^2}} \tag{2.8}$$

which corresponds to the result in [10].

One additional approximation to Eq. (2.7), which does not capture the behavior of the backscatter at very small roughness values, stems from approximating the modified Bessel function using its large argument behavior,

$$I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}} \qquad \text{as } x \to \infty. \tag{2.9}$$

With this approximation, the integral may be represented as

$$\sigma^o = \frac{2\kappa^2}{\epsilon k_0^2}\left(\lambda A\right)^{-\frac{2}{\epsilon}}\frac{(b-a)^{\frac{1}{2}-\frac{2}{\epsilon}}}{\sqrt{2\pi a}}\gamma\left(\frac{2}{\epsilon}-\frac{1}{2}, \left(1-\frac{a}{b}\right)u_0\right) \tag{2.10}$$

where the incomplete Gamma function is defined as [11]

$$\gamma(a,x) = \int_0^x e^{-t} t^{a-1}\, dt. \tag{2.11}$$

Warnick showed that an alternate asymptotic approximation of the physical optics integral for an arbitrary one-dimensional surface as the roughness parameter λ increases to infinity can be expressed in terms of multiple convolutions of the spectrum (or as the Fourier transform of the correlation function raised to a power) [12]. A similar development can be made in two (and higher) dimensions, as well. Expanding the roughness-dependent term of Eq. (2.1) about the origin yields

$$\begin{aligned} \exp\{-\lambda[1-C(\vec{x})]\} &\approx \{1-[1-C(\vec{x})]\}^{\lambda} && (2.12)\\ &= C(\vec{x})^{\lambda}. && (2.13) \end{aligned}$$

Equation (2.1) is then

$$\sigma^o = \left(\frac{\kappa^2}{\pi \cos^2\theta_i}\right) \int e^{i\vec{k}_b\cdot\vec{x}} C(\vec{x})^{\lambda}\, d\vec{x}. \tag{2.14}$$

The backscatter is thus proportional to the spectrum, convolved with itself λ times and evaluated at the Bragg wavenumber,

$$\sigma^o = \frac{4\pi\kappa^2}{\cos^2\theta_i} S^{(\lambda)}(\vec{k_b}). \tag{2.15}$$

At nadir incidence, Eq. (2.15) reduces to

$$\sigma^o = 4\pi\kappa^2 S^{(\lambda)}(0,0). \tag{2.16}$$

Note that the finite footprint effect can be observed in this formulation of the backscatter by windowing the correlation function according to the size of the footprint.

Two powerful methods of computing the asymptotic backscatter from a rough surface have been developed from the physical optics approximation. The first method, Eq. (2.10), approximates the correlation function with only its dominant behavior in order to find the nadir backscatter in closed form. The second method, Eqs. (2.15) and (2.16), use multiple convolutions of the surface spectrum to compute the asymptotic backscatter for arbitrary correlation functions and at arbitrary incidence angles.

Figure (2.1) plots the physical optics backscatter, Eq. (2.2), the asymptotic approximation using just the dominant term of the correlation function, Eq. (2.10), and the method of multiple convolutions, Eq. (2.16). The figure shows the strong agreement between the methods, with some minor differences at low roughness where the backscatter is transitioning between the purely specular scattering (with a constant value of $\kappa^2\rho_0^2$) to it's first asymptotic slope. It should be noted that while the physical optics and convolution methods are more general, they require special care in the numerical integrations of the correlation function. Further, Eq. (2.7) (not shown on the plot) also provides the specular scattering behavior at extremely low roughness values with a fairly benign integration and using only the dominant behavior of the backscatter. The vertical line shows the approximate transition point between two rates of asymptotic decay. For this plot, the footprint size was $\rho_0 = 1$ m, the electromagnetic wavenumber was $\kappa = 200$ rad/m, the power parameter was chosen as $p = 4.9$, the low-wavenumber limit was $k_0 = 0.1$ rad/m, and the directionality parameter was $\theta_0 = \frac{\pi}{32}$ rad. Note that more detailed definitions of these parameters are provided in the following section.

2.2.2 The Correlation Function

The correlation function is the inverse Fourier transform of the power spectral density normalized by the surface height variance [13],

$$C(\vec{x}) = \int e^{-i\vec{k}\cdot\vec{x}} S(\vec{k})\, d\vec{k}. \tag{2.17}$$

The power spectral density of the ocean surface indicates the relationship between adjacent points on the surface, and is frequently approximated as being separable

$$S = K(k)\Theta(\theta). \tag{2.18}$$

Empirical models of the ocean spectrum ([14] through [20]) include a low wavenumber where the function is a maximum, a power-law decay to the capillary waves, and an angular distribution function, $\Theta(\theta)$. A common representation of the power

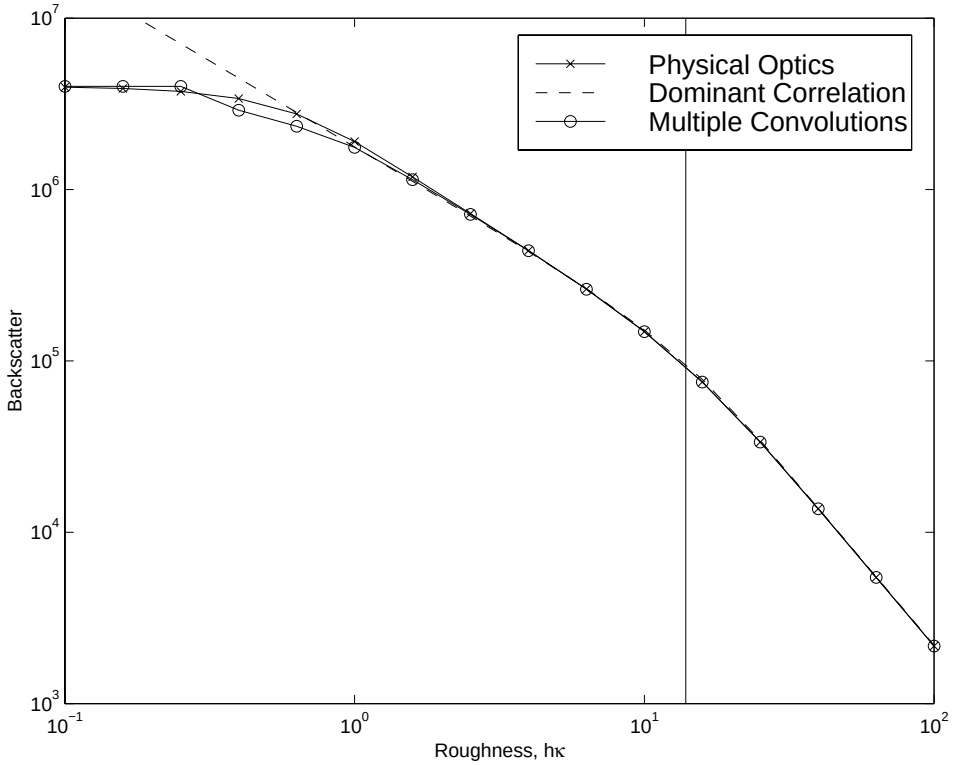


Figure 2.1: *Comparison between the methods of computing the backscatter. The physical optics integral and the method of multiple convolutions show three distinct modes of decay. At extremely low roughnesses, specular scattering dominates the process and the backscatter is constant at* $\kappa^2\rho_0^2$. *At moderate roughness values, the backscatter decays at a moderate rate—this is the lower roughness regime of the backscatter as computed using just the dominant behavior of the correlation function and the asymptotic approximation on the modified Bessel function. At high roughness values, the backscatter decays at a constant, steeper rate.*

spectral density incorporating these features is

$$S = S_0 \exp\left(-\frac{k_0}{k}\right) k^{-p} \cos^{2q}\theta \tag{2.19}$$

where S_0 serves as a normalizing constant, k is the wavenumber, k_0 describes the peak of the spectrum (the peak is actually at $k = k_0/p$), p is a power-law parameter, and θ is the angular direction of the waves. Figure 2.2 displays a typical spectrum based on Eq. (2.19), using $p = 4.5$, $k_0 = 0.1$ and $q = 4$. Ref. [21] derives the correlation function for the special case of $p = 4$ and $q = 1$, and describes the involved derivation

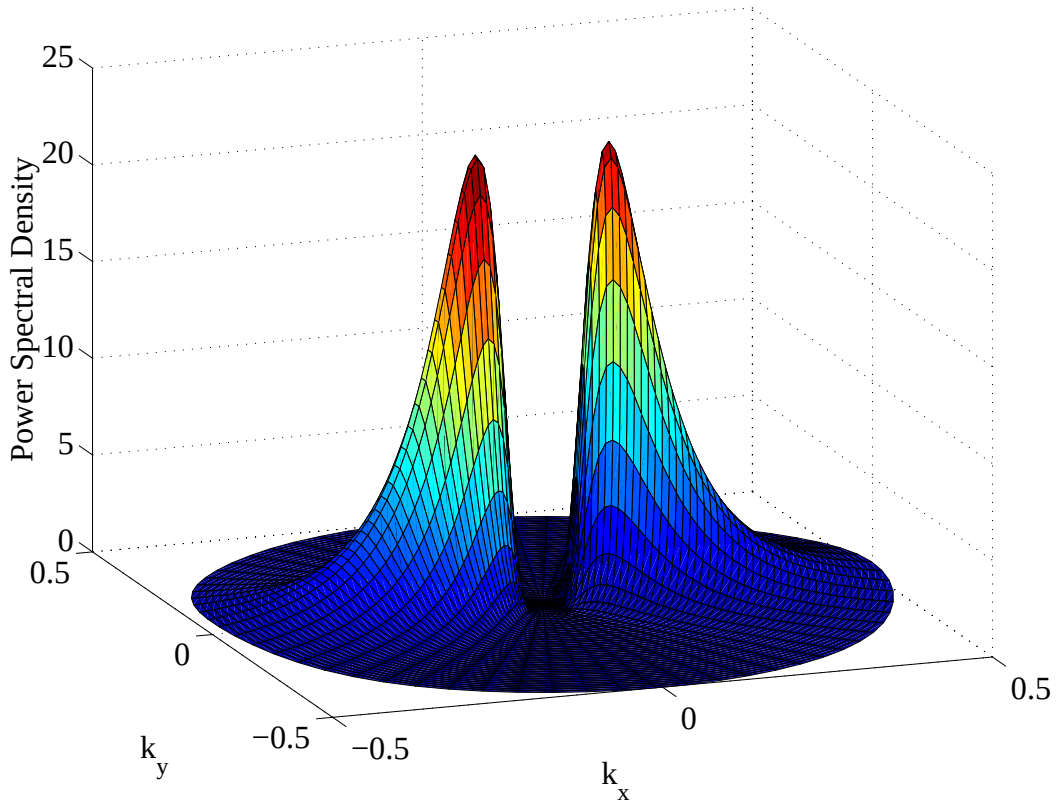


Figure 2.2: *The power spectral density for a surface with smooth transitions at the low wave number and on the directionality.*

necessary for higher values of q based on integration by parts. Eq. (2.19) provides a highly tunable and quite realistic spectrum for ocean-like surfaces, though analytically it proves difficult to incorporate.

A simple approximation for the power spectral density applies hard limits for a low wavenumber limit corresponding to the spectral peak and allows a general angular distribution function $\Theta(\theta)$. While this function allows expressions such as $\cos^{2q}\theta$, it also allows simpler approximations such as rectangular or triangular functions. Above the low wavenumber limit (k_0) the spectrum falls as k^{-p} where the power parameter, p, is in the range between 3 and 5, corresponding to theoretical and empirical estimates. The surface model assumed here, incorporating these critical parameters, can be expressed for its power spectral density as

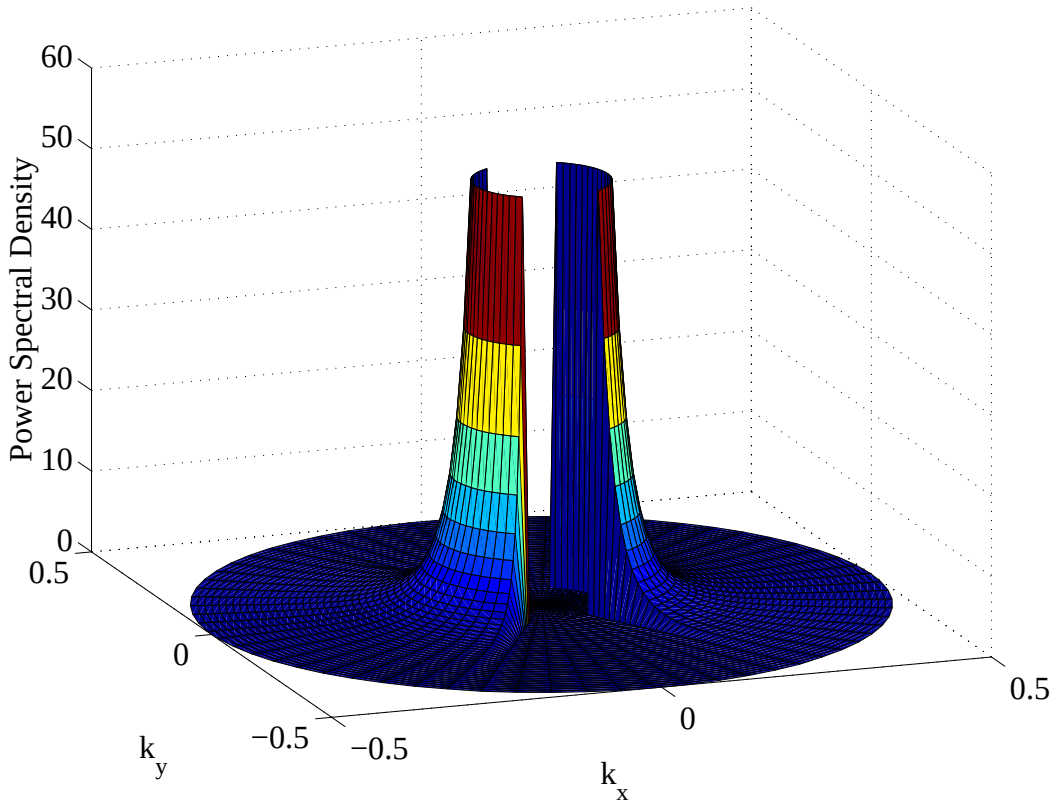


Figure 2.3: *The power spectral density for a surface with hard limits on the low wave number and on the directionality.*

$$S = S_0 k^{-p} \Theta(\theta) \qquad \text{for } k \geq k_0. \tag{2.20}$$

Figure 2.3 shows a three dimensional plot of a typical power spectral density for a directional surface based on this model, using hard limits for the angular distribution, $\Theta(\theta)$ at $\pm\theta_0/2$, with $p = 4.5$, $k_0 = 0.1$ and $\theta_0 = \pi/3$.

For the simplistic spectrum of Eq. (2.20), the inverse Fourier transform can be computed for the correlation function. Because of symmetry in the spectrum, the cosine form of the transform can be used, requiring integration only over the half plane, $k_x > 0$. Using polar coordinates for both wavenumber space (k_x and k_y transform to k and θ) and real space (x and y transform to ρ and ϕ), the correlation coefficient can be computed as

$$C(\rho,\phi) = S_0 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta \Theta(\theta) \int_{k_0}^{\infty} dk \;\; k^{-p+1} \cos[k\rho|\cos(\theta-\phi)|], \tag{2.21}$$

where k is the wavenumber of a surface wave and S_0 is a constant that normalizes the correlation coefficient to one when ρ is zero. Thus, the normalization constant is easily found to be

$$S_0 = \frac{p-2}{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \Theta(\theta)\; d\theta} k_0^{p-2}. \tag{2.22}$$

The integral over k can be found in terms of a hypergeometric function and expressed as a summation, yielding the correlation function as:

$$\begin{aligned} C(\rho,\phi) &= 1 - \frac{\sqrt{\pi},\;\left(2-\frac{p}{2}\right)}{2,\;\left(\frac{p-1}{2}\right)} k_0^{p-2}\rho^{p-2} T(p-2,\phi) \\ &\quad + \sum_{r=1}^{\infty} \left(\frac{p-2}{p-2-2r}\right) \frac{(-1)^r}{(2r)!} k_0^{2r}\rho^{2r} T(2r,\phi), \end{aligned} \tag{2.23}$$

where T depends on the particular choice of the angular distribution function $\Theta(\theta)$:

$$T(\nu,\phi) = \frac{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta \;\; \Theta(\theta)|\cos(\theta-\phi)|^{\nu}}{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta \;\; \Theta(\theta)} \tag{2.24}$$

Note in passing that the correlation function for a composite surface, one defined by multiple power parameters, p_i, and directionalities, θ_i, of a separable spectrum maintains the same form as Eq. (2.23). In Appendix A the correlation function for a composite surface is derived as the weighted sum of simple correlation functions parameterized by wavenumber transition points, power parameters and directionality limits.

For even integers, and the typical directionality functions, the identity for the expansion for a cosine raised to an integer power (see Appendix B) can be used to express the function $T(2r,\phi)$. For other values of the argument, $T(p-2,\phi)$ can be approximated as the linear interpolation between two values of even integers.

Assuming hard limits for the directionality (at $\pm\frac{\theta_0}{2}$) the term $T(2r,\phi)$ can be expressed as

$$T(2r,\phi) = 2^{-2r}\frac{(2r)!}{(r!)^2} + 2^{-2r+1}\sum_{k=0}^{r-1}\binom{2r}{k}\frac{\sin\theta_0(r-k)}{\theta_0(r-k)}\cos 2\phi(k-r) \tag{2.25}$$

For the special case $\nu = 2$, and assuming a directionality with hard limits,

$$T(2,\phi) \;=\; \frac{1}{2}\left(\frac{\sin\theta_0}{\theta_0}\cos 2\phi + 1\right). \tag{2.26}$$

Similarly, assuming a directionality of $\Theta(\theta) = \cos^{2q}(\theta)$,

$$T(2r,\phi) = 2^{-2r}\frac{(2r)!}{(r!)^2} + 2^{-2r+1}\sum_{k=\max(0,r-q)}^{r-1}\binom{2r}{k}\frac{(q!)^2\cos 2\phi(k-r)}{(q+k-r)!\,(q-k+r)!} \tag{2.27}$$

And the special case of $\nu = 2$ yields

$$T(2,\phi) \;=\; \frac{1}{2}\left(\frac{q}{q+1}\cos 2\phi + 1\right). \tag{2.28}$$

For values of ν around 2, regardless of the particular shape of $\Theta(\theta)$, a very good approximation for the term is $T(\nu,\phi) \approx a\cos 2\phi + b$. Thus, the particular choice of the form of the directionality is not critical; the important parameter is the effective width of the directionality (i.e., value of θ_0 or q). Unidirectional surfaces ($\theta_0 = 0$ or $q \to \infty$) yield the maximum value of 0.5 (equal to the b coefficient), while spreading the waves out reduces the value of a to a minimum of zero for isotropic surfaces ($\theta_0 = 0$ or $q = 0$).

Asymptotic evaluation of the backscatter integrals only requires the dominant behavior of the correlation coefficient. This dominant behavior is determined by the value of p, based on the lowest power of ρ (since the integral will add significant contribution only near $\rho = 0$). Note that in *this* asymptotic analysis, only the dominant term is used. Near $p = 4$, the ρ^2 and ρ^{p-2} terms are of comparable order and a better, though more complicated, approximation would use both terms. In avoiding this added complication, convergence of the backscatter to the asymptotic solution for surfaces with p near four requires higher roughnesses values.

With the approximation on $T(\nu,\phi)$, the dominant behavior of the correlation function can be written (for $p \neq 4$) as

$$C_{dom}(\rho,\phi) = 1 - Ak_0^{\epsilon}\rho^{\epsilon}(a\cos 2\phi + b) \tag{2.29}$$

where

$$\epsilon = \begin{cases} p-2 & \text{for } p < 4 \\ 2 & \text{for } p > 4 \end{cases} \tag{2.30}$$

and

$$A = \begin{cases} \frac{\sqrt{\pi}\Gamma\left(2-\frac{p}{2}\right)}{2\Gamma\left(\frac{p-1}{2}\right)} & \text{for } p < 4 \\ \frac{1}{2}\left(\frac{p-2}{p-4}\right) & \text{for } p > 4 \end{cases} . \tag{2.31}$$

2.3 Nadir Backscatter with Traditional Asymptotics

Equation (2.10) provides the asymptotic limit of the backscatter at nadir, assuming a finite radar footprint. Recalling that $\lambda = 4h^2\kappa^2$ (where h is the surface wave height and κ is the electromagnetic wavenumber) the fundamental parameters of the spectrum and scatterometer which affect the asymptotic backscatter are seen to be the surface wave height, h, the low wavenumber cutoff, k_0, the power parameter of the spectral decay, p, the surface directionality, θ_0, the electromagnetic wavenumber, κ and the footprint size ρ_0. In this section plots of Eq. (2.10) are displayed with various sets of parameters to clarify the impact of the finite footprint.

The general behavior of the backscatter, that exploited by scatterometers, is its strong dependence on the surface roughness which can be defined as $h\kappa$, the surface wave height times the electromagnetic wavenumber. Figure 2.4 plots the backscatter, σ^0, against roughness on a logarithmic plot at values of $p = 3.1$ and $p = 4.9$ for $\theta_0 = \pi/32$ using Eqs. (2.7) and (2.10). Note that Eq. (2.7) (the solid line plotted for each value of p), using the modified Bessel function, includes the low roughness limit; if λ is small, the backscatter is constant at a value of $\kappa^2\rho_0^2$. Equation (2.10) does not show this limit because of the asymptotic approximation on the Bessel function. Figure 2.4 reveals two modes, or slopes of decay, depending on the surface roughness. The final asymptotic slope obviously depends on p as is readily apparent in Eq. (2.8) where (for infinitely large footprints) the backscatter is proportional to the roughness raised to $-\frac{4}{\epsilon}$. Because ϵ depends on p, the asymptotic slope of the backscatter is distinctly different for the two regions of p, changing considerably for $p < 4$ but remaining constant for $p > 4$.

The transition point between the intermediate slope and the final asymptotic slope can be identified by examining the incomplete gamma function of Eq. (2.10). The incomplete gamma function, $\gamma(x, a)$, demonstrates two very different

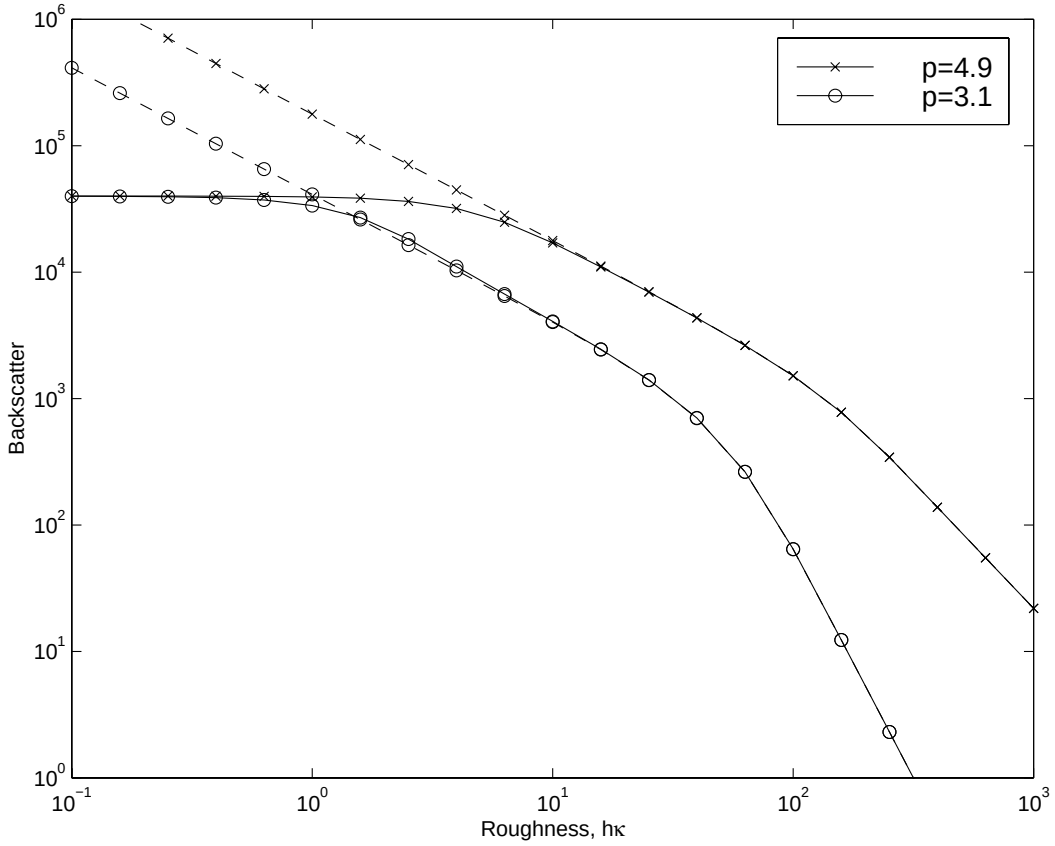


Figure 2.4: *The general behavior of the backscatter is determined by the surface roughness. The solid line for each value of p is based on Eq. (2.7), while the dashed lines are based on Eq. (2.10). For this plot,* $\theta_0 = \pi/32$ *(5.6°),* $\kappa = 200$*,* $k_0 = 0.1$ *and h is varied to establish the surface roughness.*

behaviors, with the function becoming very insensitive to x above $a = 2x$, where it approaches , (x). This can be seen by taking the derivative of $\gamma(x, a)$ using Leibniz' Rule:

$$\frac{\partial}{\partial a}\gamma(x,a) = \frac{\partial}{\partial a}\int_0^a dt\, e^{-t}t^{x-1} \tag{2.32}$$

$$= e^{-a}a^{x-1}. \tag{2.33}$$

The transition point occurs when Eq. (2.33) approaches it's asymptotic value of zero. Note that scaling by $1/$, (x), Eq. (2.33) becomes a Gamma probability density function in a, with mean and variance both equal to x—-which is negligible beyond

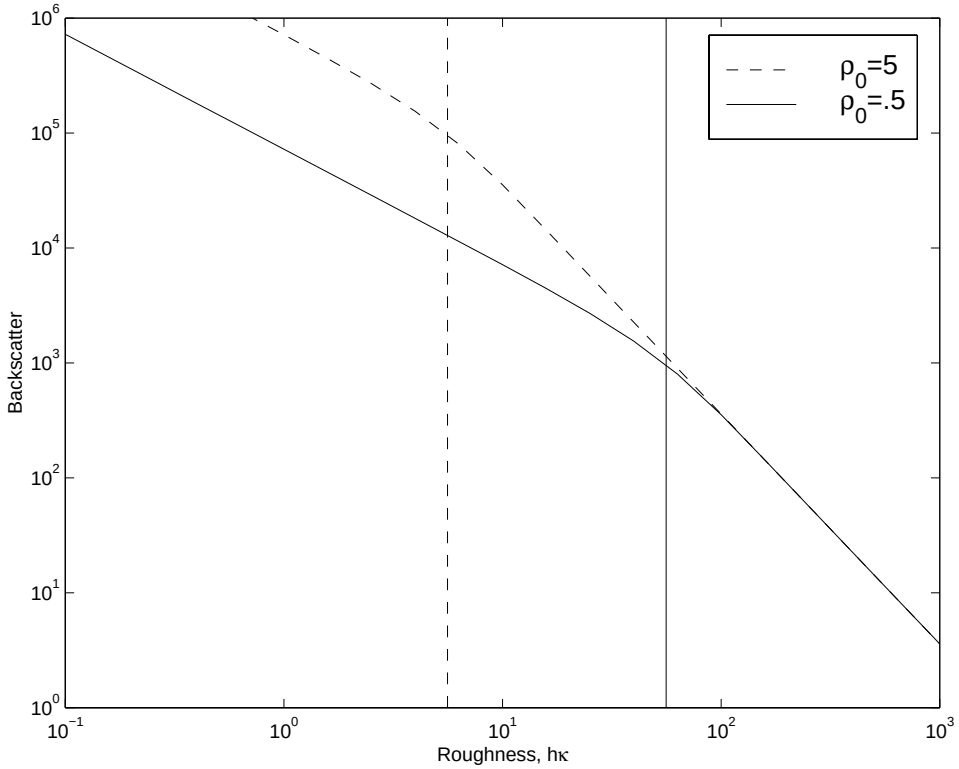


Figure 2.5: *The asymptotic backscatter is plotted against surface roughness for two footprint sizes. The vertical lines indicate the knees of the two plots as determined from the knee of the incomplete gamma function. Values chosen for this plot reasonably represent the ocean surface:* $p = 4.5$, $k_0 = 0.1$, *and* $\theta_0 = \pi/4$. *The plot used* $\kappa = 200$, *for microwaves, and the footprint sizes are not unrealistic of tower-mounted and aircraft-mounted altimeters.*

about twice the mean. Recognizing the crudeness of this approximation, it provides a reasonably accurate estimate for the point where the incomplete Gamma function changes behavior. The incomplete Gamma function $\gamma(x, a)$, then, is essentially equal to the Gamma function , (x) for all values of $a > 2x$.

The backscatter then should become essentially independent of the footprint size when the surface roughness parameter is greater than

$$\lambda_0 \approx \left(\frac{4 - \epsilon}{\epsilon A (b - a) k_0^{\epsilon}} \right) \rho_0^{-\epsilon}. \tag{2.34}$$

The important observation of this chapter is the effect of a finite footprint on the backscatter. Frequency and altitude of altimeters affect the behavior of the

backscatter as described by [22–24]. As seen in Eq. (2.10), a finite footprint will modify the backscatter in some way that depends on the surface roughness. Figure 2.5 plots the normalized asymptotic backscatter of Eq. (2.10) against the roughness for two different footprint sizes, $\rho_0 = 0.5$ and $\rho_0 = 5$, holding other surface parameters constant. While both footprint sizes display the same general behavior of following a constant slope up to some roughness value and then decaying at a steeper slope, the larger footprint (the dashed line) reaches the steeper, final slope at a lower roughness value. The vertical lines indicate the λ_0 values of Eq. (2.34) for each of the two footprint sizes. Note that between the λ_0 values, the slopes of the backscatter for the two footprint sizes are different; outside this range of roughnesses, the slopes of the two cases are identical. For a particular sea state, at a roughness of, say, $h\kappa = 5$ on this scale, a radar system with a 1 m footprint yields a shallower slope than a system with a 10 m footprint.

To describe the cause of the finite footprint effect, consider Fig. 2.6. A surface has been established with the proper power spectral density, and directionality of $\theta_0 = \pi/8$. The top plot shows the surface with a roughness of one, while the bottom plot shows the same surface (with the same scale) with a roughness of ten. The surface variance in the top plot causes incoherent scattering in one direction, but the other direction's variance is small so the scattering is coherent: the surface looks nearly unidirectional and the backscatter is limited by a finite footprint to the intermediate range. The bottom plot, with the greater roughness, causes incoherent scattering in both directions: the rougher surface converges to the final asymptotic result with a smaller footprint. Mathematically, the effect of the finite footprint is found in the integrand of the backscatter equation, $e^{-\lambda[1-C_{dom}(\rho,\phi)]}$. The integral requires information from the correlation function out to the critical values beyond which the integrand contribution is negligible. These values are commonly chosen where the integrand is e^{-1} [12] or where

$$C_{dom}(\rho, \phi) = 1 - \frac{1}{\lambda}. \tag{2.35}$$

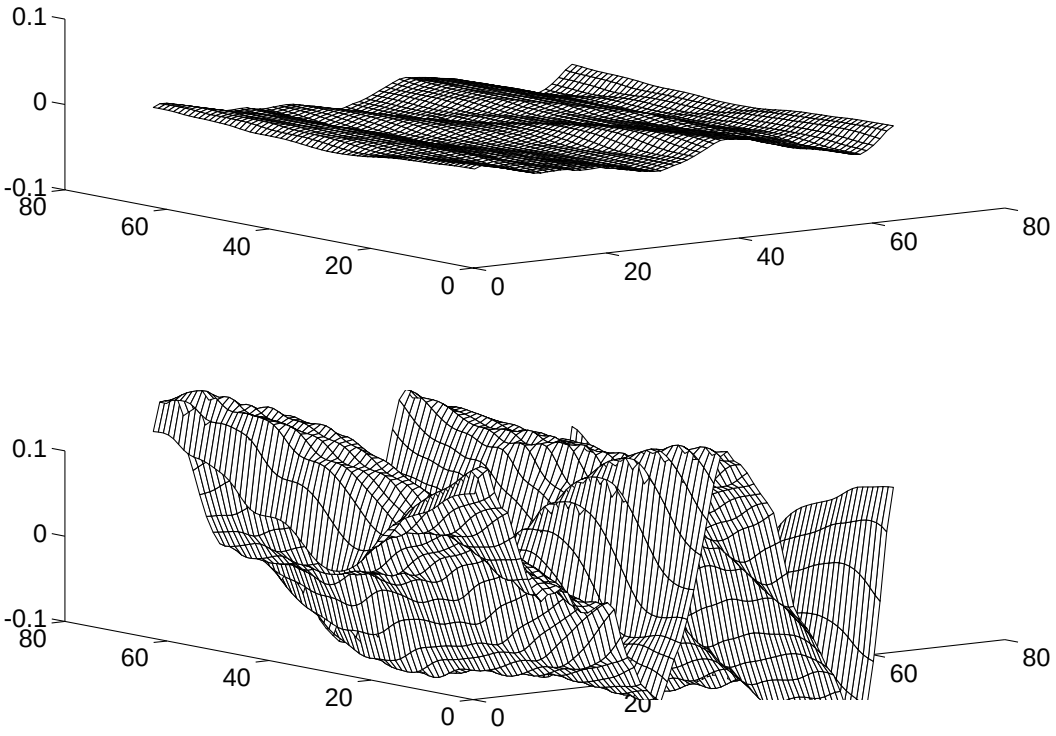


Figure 2.6: *A directional random surface with a roughness of (top)* $h\kappa = 1$ *and (bottom)* $h\kappa = 10$.

Figure 2.7 displays contour plots of where the correlation function equals $1 - 1/\lambda$, using $p = 4.9$ and $k_0 = 1$. Reducing the surface directionality, θ_0, tends to stretch the correlation function in one direction (perpendicular to the dominant direction of the surface waves), and as the roughness increases, the size of the critical contour decreases. If part of the contour is not contained in the footprint, the resulting backscatter will have a different behavior than if all of the contour (i.e., significant information of the correlation function) is within the footprint. Thus, while a given footprint size will not contain a low-roughness contour completely, it will contain a high-roughness contour. The asymptotic backscatter will converge first to a solution limited by the finite footprint, and eventually with increasing roughness will converge to a solution based on the entire contour.

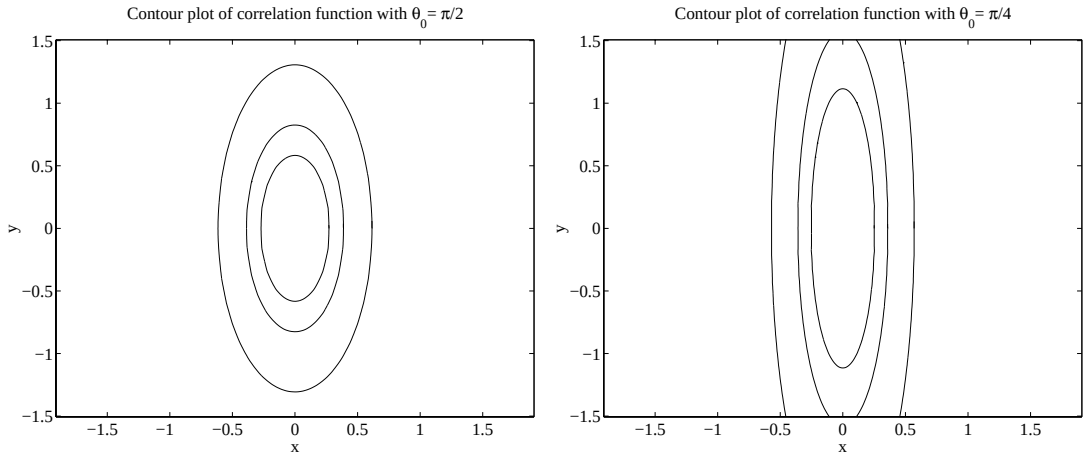


Figure 2.7: *Contour plot where the correlation function equals 0.5, 0.8 and 0.9, for surfaces with* $\theta_0 = \frac{\pi}{2}$ *(left) and* $\theta_0 = \frac{\pi}{4}$ *(right).*

From Eq. (2.34), the roughness value where the behavior changes from footprint-limited to the final asymptotic mode depends on the size of the footprint, the low-wavenumber cutoff, the spectrum power-law decay and the directionality of the surface as described by the coefficient, a. Figure 2.8 displays the backscatter, showing the effect of a more narrow spectrum. The backscatter from the more narrow spectrum ($\theta_0 = \pi/128$) requires a much larger roughness to converge to the final asymptotic behavior.

One of the advantages of using Eq. (2.10) to compute the asymptotic backscatter is to be able to directly identify the sensitivity of the backscatter to parameters such as the surface roughness. In particular, the slopes of the two modes of asymptotic behavior are readily determined. The slope of the decay is the partial derivative of the log of σ^o with respect to the log of λ (note that the slope is defined in terms of the roughness as described by $\lambda = 4h^2\kappa^2$):

$$\alpha = \frac{\partial}{\partial(\log\lambda)} \log\sigma^o \tag{2.36}$$

$$= -\frac{2}{\epsilon} + (N\lambda)^d \frac{e^{-N\lambda}}{\gamma(d, N\lambda)} \tag{2.37}$$

where $N = (b-a)A(k_0\rho_0)^\epsilon$ and $d = 2/\epsilon - 1/2$. As previously pointed out, the

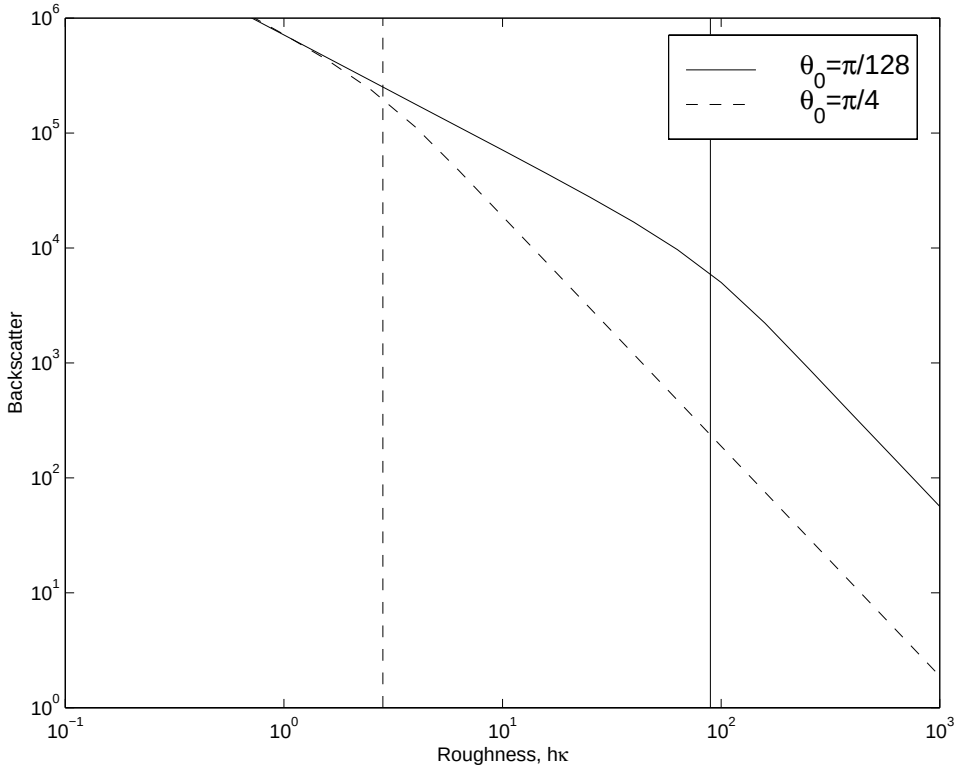


Figure 2.8: *The asymptotic backscatter is plotted against surface roughness for two directionality values. The wider directionality, $\theta_0 = \pi/4$ (45°), quickly converges to the final asymptotic behavior, while the narrower directionality, $\theta_0 = \pi/128$ (1.5°), requires a higher roughness.*

incomplete gamma function experiences two modes of behavior, with a transition point approximately given by Eq. (2.34). The two asymptotic backscatter slopes can be found by taking the limits of α as λ goes to zero and to infinity (the term on the right has limits of d and 0):

$$\alpha = \begin{cases} -\frac{1}{2} & \text{for } \lambda < \lambda_0 \\ -\frac{2}{\epsilon} & \text{for } \lambda > \lambda_0 \end{cases} . \tag{2.38}$$

2.4 Asymptotic Backscatter Using Multiple Convolutions

While using the unrealistic spectrum with hard limits on the wavenumber allows the backscatter to be analyzed in terms of fairly simple functions, the

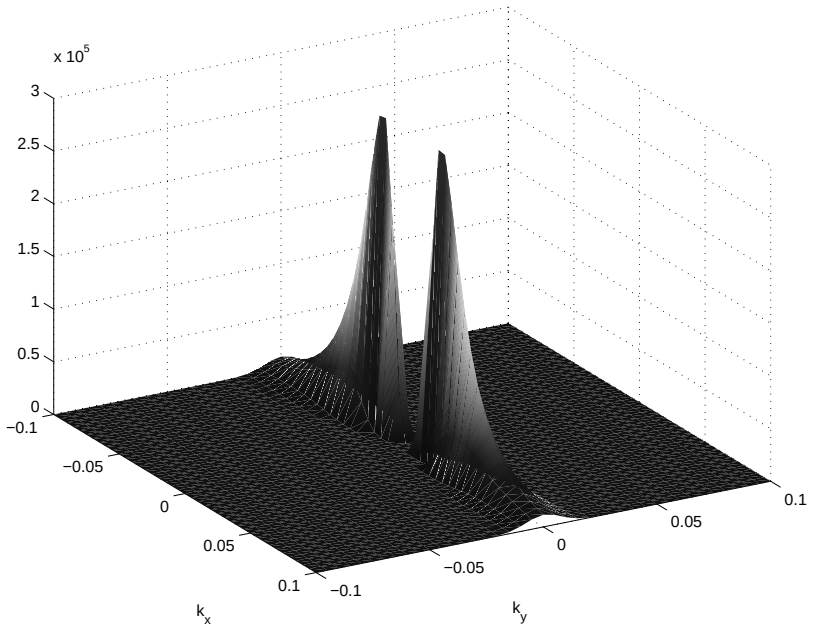


Figure 2.9: *A more realistic spectrum, without discontinuities in k.*

method for computing the backscatter using multiple convolutions allows the asymptotic backscatter to be computed using arbitrary correlation functions (constrained by the physical optics approximation) and also applies at off-nadir incidence angles. In this section, we consider a more realistic spectrum using the form supplied by Eq. (2.19).

Figure (2.9) displays the spectrum that will be considered in this section. Note the smooth, low-wavenumber rise, as well as the smooth angular transitions. Here we use $p = 4.5$, $k_0 = 0.1$ (corresponding to a peak at $k_{peak} = k_0/p = 0.022$) and $q = 87$ (comparable to $\theta_0 = \frac{\pi}{12}$).

The backscatter was computed for the spectrum described in Fig. (2.9), using the multiple convolution method. Figure (2.10) shows the resulting backscatter computed using three footprint sizes. Recall that a perfectly smooth surface would result in $\sigma^o = \kappa^2\rho_0^2$. For 10, 20 and 30 m footprints, this corresponds to 4×10^6, 1.6×10^7, and 3.6×10^7, respectively. As the surface roughness increases, the three curves merge because the extremities of the correlation function no longer contribute

to the backscatter integral—all pertinent information is contained within the smaller footprint size. Also shown in Fig. (2.10) is the sensitivity of the backscatter to roughness. The backscatter slope, α, defined as the partial derivative of the log of the backscatter with respect to the log of λ [Eq. (2.37)] begins near zero where all the footprint sizes experience specular scattering. As the roughness increases, the scattering experiences first one mode of decay ($\alpha \approx -\frac{1}{2}$) and then the final asymptotic mode ($\alpha \approx -\frac{2}{\epsilon}$). Each footprint size transitions at different roughness values. The vertical, dashed lines, indicate the transition points as computed using Eq. (2.34) where the backscatter slope changes from its first asymptotic slope (experienced only briefly) to its final asymptotic slope for the three footprint sizes.

2.5 Electromagnetic Bias

One application of the analysis performed in this chapter to is to examine the effect of the footprint on the electromagnetic bias. The EM bias, ϵ_B, is an error present in radar altimeter data in which the apparent mean sea height is biased toward the troughs of the waves [7]. The empirically determined bias has been found to behave differently for altimeters mounted on towers, as compared to altimeters mounted on aircraft and satellites. One plausible explanation of this observation is the difference in footprint sizes. Tower-mounted altimeters typically have small footprints, while high-altitude altimeters have large footprints. In this section I explore the hypothesis that the footprint size has an effect on the EM bias.

The EM bias can be defined mathematically as the ratio of first two moments of the backscatter coefficient profile at a given surface displacement. That is, if the surface is displaced a distance η from its mean level, then σ^o_η describes the amount of backscatter from points with that displacement. The bias is then

$$\epsilon_B = \frac{\int d\eta \ \eta \sigma^o_\eta p_\eta(\eta)}{\int d\eta \ \sigma^o_\eta p_\eta(\eta)} \tag{2.39}$$

where $p_\eta(\eta)$ is the probability density function of the surface displacement.

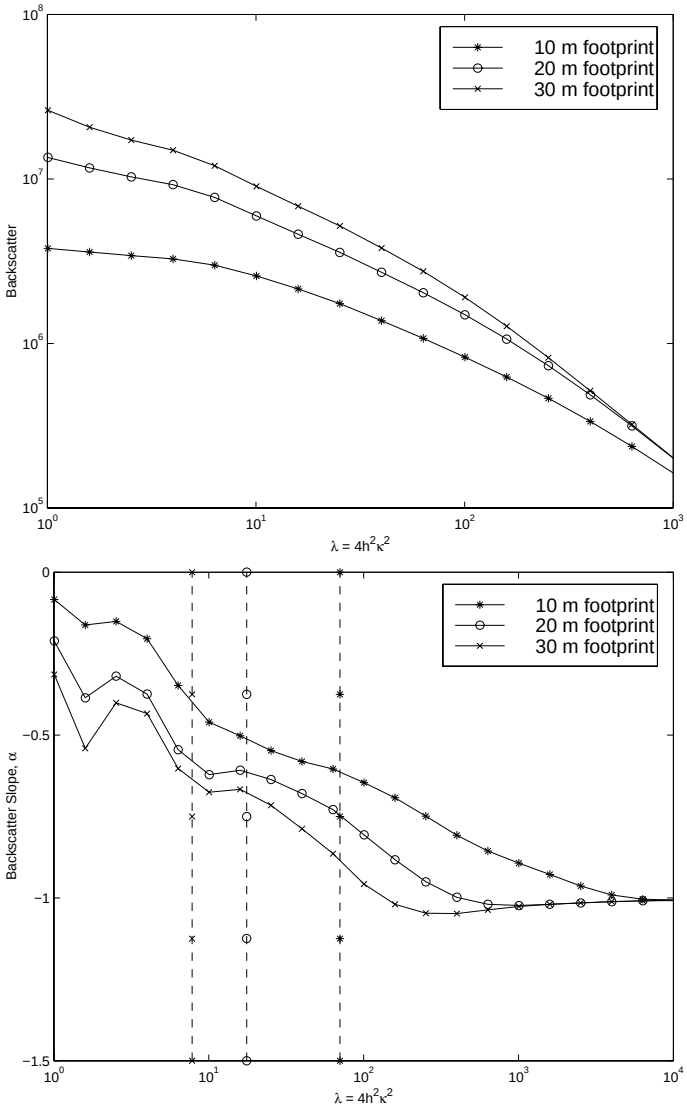


Figure 2.10: *The backscatter from a smooth spectrum (plotted on the top) also demonstrates bimodal behavior. The slope of the backscatter with respect to* λ *is plotted (on the bottom) for the three footprint sizes. The slope converges first to a value of about -0.5, and then, as the roughness becomes asymptotically large, converges to a slope around -1.*

In Eqs. (2.10) and (2.38) the backscatter was found to have a behavior of the form

$$\sigma^o = K\left(4h^2\kappa^2\right)^{\alpha} \tag{2.40}$$

where K is a proportionality constant, h is the RMS surface displacement, κ is the electromagnetic wave number, and α is the rate of decay. The plots in the previous section showed that using the more realistic spectrum with multiple convolutions to compute the backscatter yields comparable results.

Arnold described the RMS surface displacement in terms of short-wave modulation, σ_m, with a modulation parameter, m, [7]

$$h = \sigma_m\left(1 + m\frac{\eta}{\sqrt{\overline{\eta^2}}}\right) \tag{2.41}$$

where σ_m is the RMS wave height of the short waves riding on top of the long waves and m is the modulation parameter. The backscatter using this model can be approximated for small values of the modulation parameter, m, as

$$\sigma_\eta^o \approx K\left(4\kappa^2\right)^{\alpha}\sigma_m^{2\alpha}\left(1 + 2\alpha m\frac{\eta}{\sqrt{\overline{\eta^2}}}\right). \tag{2.42}$$

The EM bias, then, is

$$\epsilon_B = \frac{\int d\eta\ \left(\eta + 2\alpha m\frac{\eta^2}{\sqrt{\overline{\eta^2}}}\right)p_\eta(\eta)}{\int d\eta\ \left(1 + 2\alpha m\frac{\eta}{\sqrt{\overline{\eta^2}}}\right)p_\eta(\eta)} \tag{2.43}$$

$$= \frac{2\alpha m}{\sqrt{\overline{\eta^2}}}\frac{\int d\eta\ \eta^2 p_\eta(\eta)}{\int d\eta\ p_\eta(\eta)} \tag{2.44}$$

$$= 2\alpha m\sqrt{\overline{\eta^2}}. \tag{2.45}$$

Equation (2.37) defines the slope of the backscatter, α, based on the approximate spectrum (using a hard limit for the low-wavenumber cutoff, k_0). The EM bias is then

$$\epsilon_B = 2m\sqrt{\overline{\eta^2}}\left[-\frac{2}{\epsilon} + (N\lambda)^d\frac{e^{-N\lambda}}{\gamma\left(d, N\lambda\right)}\right]. \tag{2.46}$$

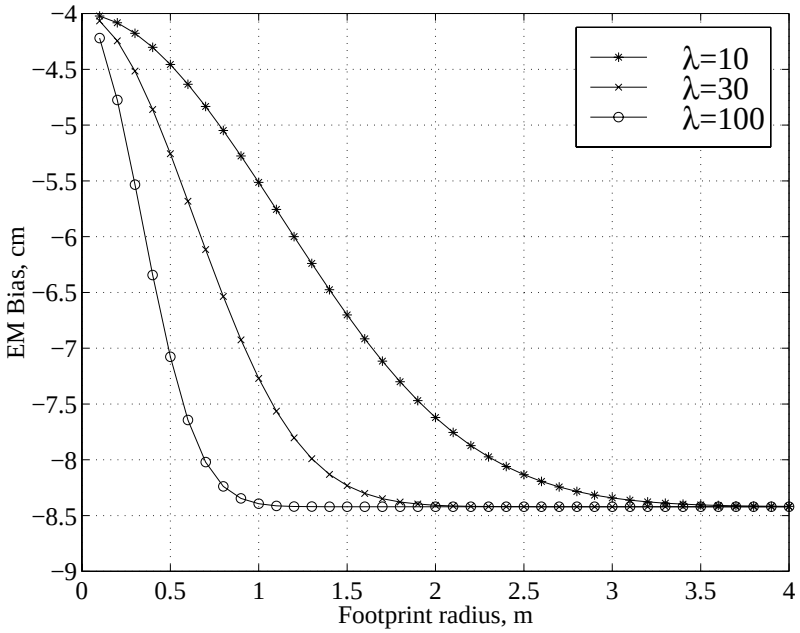


Figure 2.11: *The EM bias as a function of footprint size. Using Eq. (2.45), the bias is plotted for the case of* $p = 3.9$, $\theta_0 = \frac{\pi}{16}$, $k_0 = 1$, $m = 0.1$ *and* $\sqrt{\overline{\eta^2}} = 0.4$. *The roughness parameter is varied* $\lambda = 10$, *30, and 100.*

Using reasonable values of $p = 3.9$, $\theta_0 = \frac{\pi}{16}$, $k_0 = 1$, $m = 0.1$ and $\sqrt{\overline{\eta^2}} = 0.4$, Fig. 2.11 plots the EM bias as a function of the footprint size for the three cases of surface roughness, $\lambda = 4h^2\kappa^2 = 10$, 30, and 100. The theory of this chapter predicts that the bias is relatively small for small footprints, but as the footprint size increases, the bias moves to a more negative value. Some unpublished studies comparing satellite altimeter data to tower-mounted altimeter data suggest that the bias observed by towers, with footprints of $\rho_0 = 0.8$ m, is roughly 75% smaller than that observed by satellites, with footprints of hundreds of kilometers. The plot shows a good correspondence with this, showing the bias for a 0.8 m footprint at -6.5 cm with $\lambda = 30$, compared to an infinitely large footprint with a bias of -8.4 cm.

2.6 Conclusions

In this chapter, the physical optics backscatter was considered, and two methods for the asymptotic backscatter at nadir incidence were derived. The figures presented in this chapter describe the long understood dependence on the surface roughness, and also display the impact of correctly identifying the power parameter of the surface power spectral density. The most interesting result seen from the equations derived here is the effect of a finite radar footprint.

Surfaces with small roughnesses cause specular scattering, which means that the backscatter at nadir incidence is at a maximum. At large roughness values, the scattering becomes more diffuse and the backscatter power decreases. The power parameter of the ocean surface, p, has been argued to have a value of four and has been measured between 3.5 and 4.5. The effect of this uncertainty has been unclear because the sensitivity of the backscatter to the parameter p was unknown. Examining the equations developed in this chapter, however, it is clear that changing from $p = 3.5$ to $p = 4$ will cause the slope of the backscatter to change considerably, while further increasing p beyond 4 will not change the backscatter slope appreciably.

The bimodal effect of the backscatter caused by the finite footprint is particularly interesting. The transition point between the two backscatter modes is determined where the footprint size completely contains all of the relevant information of the correlation function. The points of the correlation function that are important are near the contour $C = 1 - \frac{1}{\lambda}$; equivalently, Eq. 2.34 describes the size of the footprint based on the surface parameters, p, k_0, θ_0 (which determines a and b), and the surface roughness. If the surface roughness is smaller than some threshold, recall Eq. (2.38), then the asymptotic slope of the backscatter is $-\frac{1}{2}$, but if the surface roughness is larger than the threshold, the asymptotic slope is $-\frac{2}{\epsilon}$, depending only on p. Of particular note is the effect of θ_0. Surface wave directionality tends to stretch the correlation function such that the effect of the footprint size becomes readily apparent. This suggests that different footprint sizes, typical for different platform heights will yield consistently biased backscatter measurements.

With a theoretical description of the backscatter dependence on the footprint size, the impact on altimetry has been qualitatively described. In particular, the EM bias was shown to be more negative for altimeters with large footprints than for altimeters with small footprints. It would be realistic, then, to observe smaller values of the bias in tower-mounted data than from satellite data.

While the causal factors of the surface spectrum are not completely understood, this study identifies the sensitivity of the backscatter to the significant parameters of a generalized spectrum. Recognizing that the backscatter behavior depends strongly on the relationship between the radar footprint and the correlation function provides the insight that not only will altimeters at different platform heights yield biased results, but also that scatterometer measurements (off nadir) using non-symmetric footprints will be complicated by the relationship between the geometries of the footprint and the directional spectrum generated by the wind.

Chapter 3

Uncertainties in the Wind-Backscatter Relationship

In Chapter 2 I developed the NRCS from a rough surface with an ocean-like spectrum. Unfortunately, understanding of the relationship between geophysical parameters and the surface spectrum is still in its infancy due to the complicated dynamical processes involved [25–27]. Because of the substantial theoretical difficulties in determining the spectrum from geophysical parameters (recall Fig. 1.1), a practical approach is necessary to determine the NRCS based on significant geophysical parameters (termed the geophysical model function).

Recognizing both the importance of wind as a geophysical parameter and its dominant role in determining the backscatter, experimentally based approximations of the geophysical model function have been used to relate the vector wind to the backscatter measurement [28–31]. Such empirical estimates are based on aircraft scatterometer missions, refined with the growing body of satellite-borne scatterometer data [32–34]. These empirical models are generally accurate in the mean, but because additional geophysical parameters affect the backscatter, there is some variability to the backscatter for given wind conditions [35, 36]. This variability and the sensitivity to non-wind factors are not well understood. This chapter examines this variability and develops a novel method for estimating it for a given model function solely from scatterometer measurements; the magnitude of the variability is reported for a commonly used Ku-band model functions .

In 1978, the satellite SEASAT carried the scatterometer SASS in orbit; although the power supply failed after only three months of data collection, the value of spaceborne scatterometers in wind estimation was demonstrated [37, 38]. The first

European Remote Sensing (ERS-1) satellite used an Active Microwave Instrument (AMI) to collect data from 1991 through 1996, when it was replaced by the similarly designed ERS-2. The NASA Scatterometer, NSCAT, created at JPL was launched on the ADEOS spacecraft in August of 1996 [29], and collected data through June 1997 until ADEOS failed due to mechanical problems. JPL will launch the next generation scatterometer, QuikScat, in May 1999. This growing body of space-borne scatterometer data with colocated *in situ* measurements (ships, buoys, etc.) has resulted in a very strong confidence in the ability of an empirical model function to provide the mean backscatter.

3.1 Overview

3.1.1 The Empirical Model Function

The normalized radar cross section, or backscatter, σ^o, is physically generated by conditions of the scattering surface; this can be described by a geophysical model function [26]. A thorough understanding of the parameters affecting the backscatter is beyond the current state of science and empirical models have been developed as useful approximations. Early experiments suggested a strong correlation between the wind speed over the ocean and the backscatter coefficient (the roughness of the ocean surface is related to the energy in the wind). A simple empirical model was suggested by [39],

$$\sigma^o = aU^\alpha. \tag{3.1}$$

Later experiments [40] showed that the constants, a and α of this model are functions of the relative azimuth angle of the wind, χ, with $a = a_0 q$ and $\alpha = \alpha_0 + \delta$ where a_0 and α_0 are the average values over all wind directions and q and δ contain the χ dependence. Experimentally, q varies by up to ± 0.5 about one and δ varies by up to ± 0.2 about zero. Since δ is small, a better backscatter model, including azimuth dependence, is

$$\sigma^o = a_0 q U^{\alpha_0+\delta} \tag{3.2}$$

$$\approx a_0 q U^{\alpha_0}\left(1+\delta \ln U\right) \tag{3.3}$$

$$= a_0 U^{\alpha_0}\left(q + 2.3 q\delta \log_{10} U\right). \tag{3.4}$$

The angular dependent terms, q and $q\delta$ can be expressed in terms of a truncated Fourier series based on the relative wind direction, χ:

$$q \approx 1 + a_1 \cos\chi + a_2 \cos 2\chi \tag{3.5}$$

$$2.3q\delta \approx \alpha_1 \cos\chi + \alpha_2 \cos 2\chi \tag{3.6}$$

where $q\delta$ has no bias because it varies about zero, and there is an assumed symmetry about the wind direction resulting in no sine terms in the expansion. The current state-of-the-art model for backscatter, σ^o_M where the M subscript identifies this as the empirical model function, incorporating wind speed and direction, is [32]

$$\sigma^o_M = A_0 + A_1 \cos\chi + A_2 \cos 2\chi, \tag{3.7}$$

with

$$A_0 = a_0 U^{\alpha_0} \tag{3.8}$$

$$A_1 = A_0(a_1 + \alpha_1 \log_{10} U) \tag{3.9}$$

$$A_2 = A_0(a_2 + \alpha_2 \log_{10} U) \tag{3.10}$$

where U is the wind speed, χ is the relative azimuth direction and the six model parameters are determined by regression analysis, using this model to fit the *in situ* measured winds to the satellite measured backscatter values, along with considerable *ad hoc* tuning to a particular instrument. Values for these parameters are maintained in C-band and Ku-band tables at H-pol and V-pol, as functions of incidence angle. Usually, $A_2 \gg A_1$, conveying the dominance of the double cosine behavior with wind direction. Figure 3.1 displays some typical examples of the empirical model function backscatter as a function of the wind speed and azimuth angle.

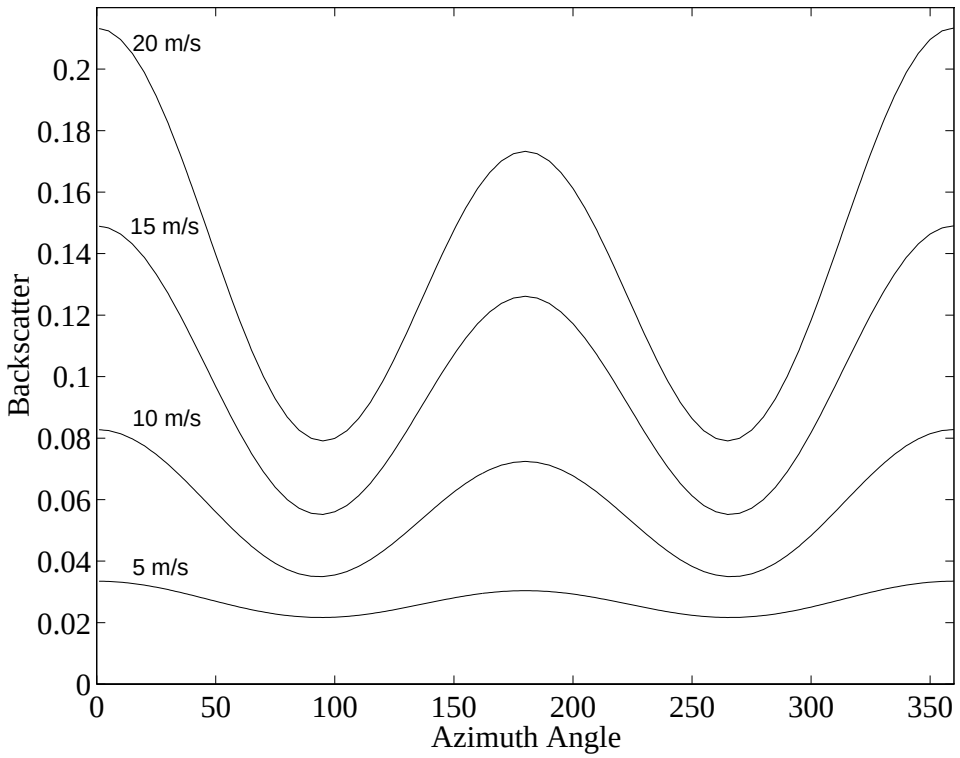


Figure 3.1: *The empirical model function relates the wind speed, relative azimuth direction, frequency band, polarization and incidence angle to the backscatter. This plot, made with the C-band model function at an incidence angle of 35°, shows the double cosine behavior with azimuth angle and the increasing return with wind speed.*

3.1.2 Chapter Preview

While empirical model functions have become increasingly more developed, it has become clear that other parameters, besides just the observation geometry and the vector wind, affect the backscatter. These other parameters can be viewed as variability in the backscatter about the mean provided by the empirical model function. This can be seen by observing the variability in the measurement, z, for given wind conditions, and recognizing that there is more variability than can be accounted for by the communication noise, K_{PC}. Determining, from first principles, the other parameters that affect the backscatter would yield great scientific value; such an analysis is beyond the current state of the art in geophysics. However, simple

estimates of the magnitude of this model function uncertainty can be determined solely from the scatterometer measurements.

In the next section, a simple model for incorporating model function variability in the measurement process is developed in a multiplicative form, comparable to the way communication noise is applied to the backscatter measurements. In Section 3.3, a technique for estimating the variability based on this simple model is described. Section 3.4 reports data-driven estimates of the variability based on NSCAT data and explores its dependence on several parameters. Specifically in Section 3.4, the model function variability is estimated for narrow bins of latitude, season, wind speed, and incidence angle. The model function variability is found to be highly correlated with the latitude, wind speed and measurement incidence angle, and slightly correlated with time. Section 3.5 identifies how the variability affects the wind estimation process; we find through simulation that methods of suboptimal estimation, modifying the way variability is incorporated into the wind estimation, has minimal impact on estimated wind. Model function variability, while not significantly affecting wind estimates, affects the confidence in the wind estimates and the ranking of the wind ambiguities.

3.2 The Multiplicative Noise Model

The backscatter measurements made by a scatterometer are noisy, indirect measurements of the wind. In particular, consider two significant sources of variability in the measurements for given wind conditions: variability in the empirical model function, and thermal noise in the communication channel. A simple model for scatterometer measurements of the backscatter is displayed in Fig. 3.2. The geophysical model function maps the surface wind to the normalized radar cross section (NRCS, σ^o), that is, the GMF translates the wind to some surface shape. The scatterometer attempts to measure this value, but introduces communications, or radiometric, noise based on the temperature of the antenna.

The variability in the backscatter measurements, z, given the true wind, appears to be reasonably represented as a multiplicative term. Figure 3.3 provides

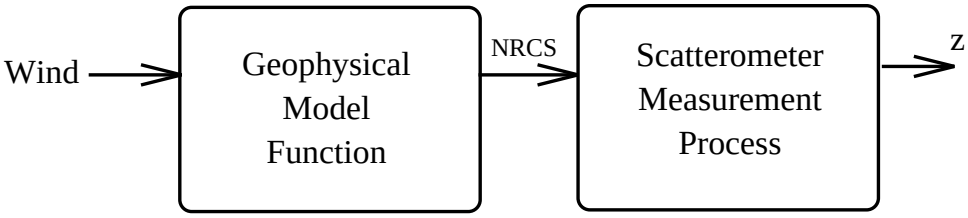


Figure 3.2: *The model for scatterometer measurements of the normalized radar cross section. The wind is mapped through the true model function to yield the NRCS. The resulting "true" backscatter is corrupted by communication error (i.e., thermal noise) in the measurement process, which yields the measured value of the backscatter, z.*

a comparison between the measurements and the backscatter predicted by the empirical model function based on the estimated wind and the satellite measurement geometry. This figure displays a compilation of more than a day's worth of NSCAT measurements. There is a clear, linear, relationship between the measurements and the model function backscatter. The scatter plot splits into two regimes at high values of the backscatter; this split occurs at high wind speeds (greater than about 15 m/s) and low incidence angles (less than 20°).

The nature of the measurement variability is more easily seen in Fig. 3.4, where histograms of the measured backscatter given the retrieval information are plotted along the line indicating the means of the histograms. Again the bifurcation of the measurements into two regimes at high backscatter values is evident.

Figure 3.5 displays the standard deviation of z given the retrieval information. This figure reveals an affine relationship (for backscatter values below the splitting of the backscatter), with a very small y-intercept value. This suggests a multiplicative noise model of the form $z = \sigma^o_M (1 + \nu) + \delta$, where ν and δ are relatively small random variables. I now consider models for the two components of the overall noise: the model function variability and the communication noise.

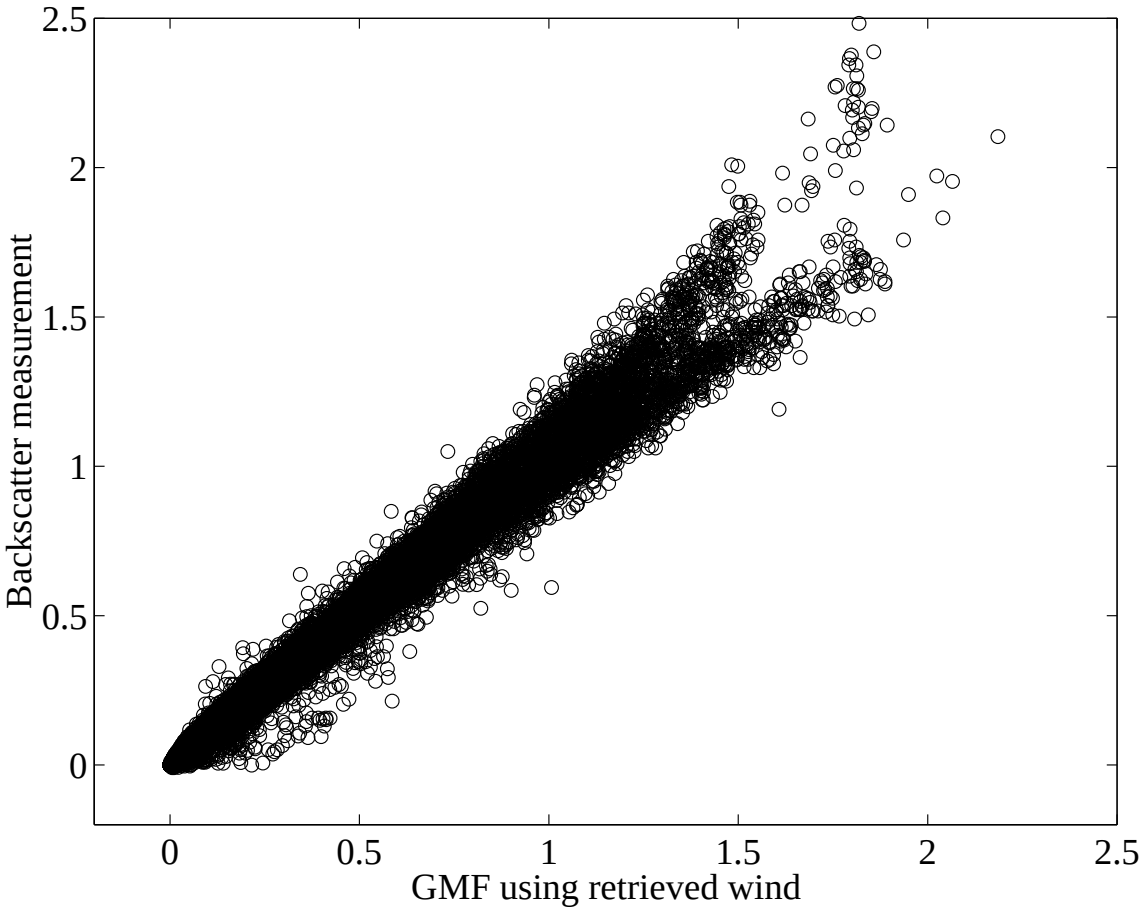


Figure 3.3: *A scatter plot of the backscatter measurements against the actual surface backscatter predicted by the empirical model function and the retrieved wind.*

3.2.1 Communication Noise

The second source of noise mentioned above, that due to the communications channel for the measurement process, is well understood from first principles. Fischer found the noise variance of a scatterometer with an analog processor to depend on the signal (i.e., true backscatter) itself [41], or

$$z = \sigma_T^o \left(1 + K_{PC}\mu\right), \tag{3.11}$$

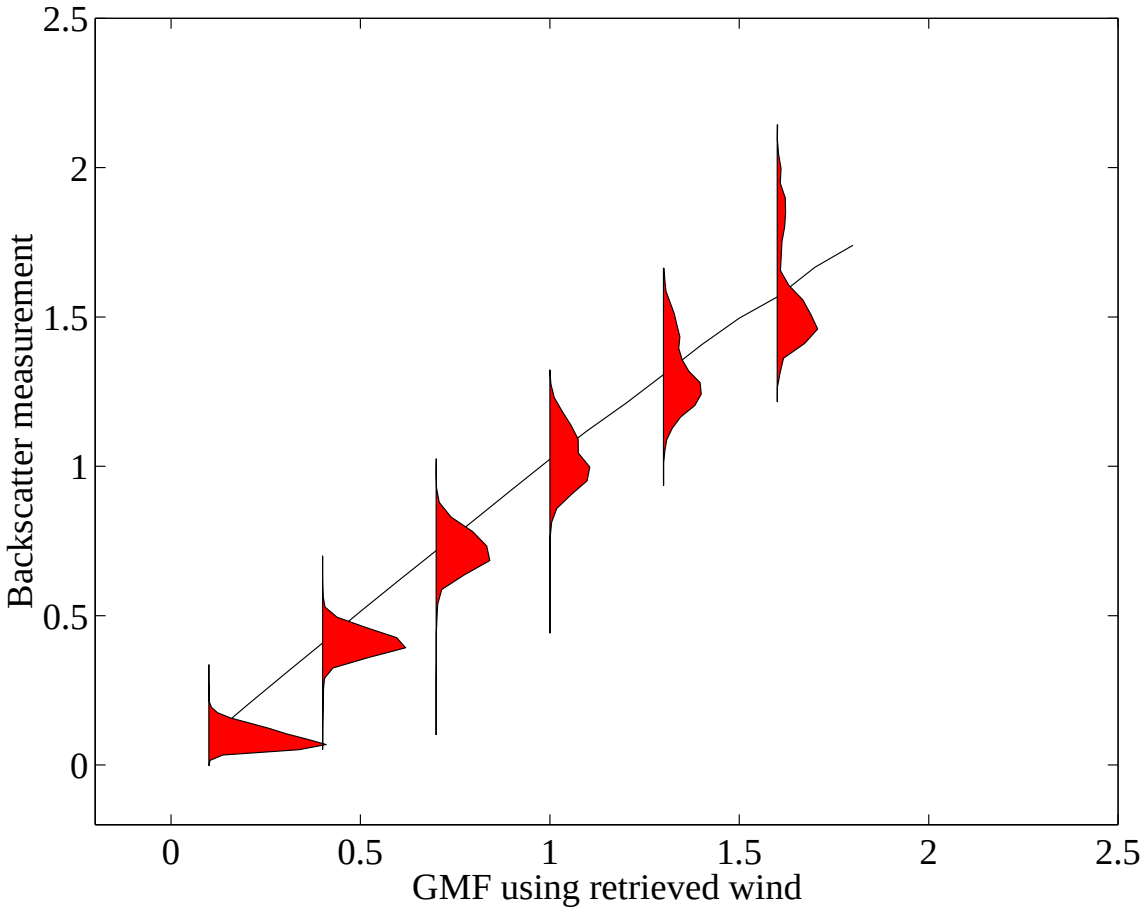


Figure 3.4: *The mean of the measurements, given all retrieval information, is plotted, revealing that the mean measurement is (essentially) an unbiased estimate of the backscatter based on the retrieved wind. The (scaled) histograms describe the variability of the measurements for given retrieval conditions.*

where μ is a standard normal random variable and K_{PC}^2 represents the normalized variance of z. This normalized variance depends on the signal to noise ratio:

$$K_{PC}^2 = A + \frac{B}{\mathrm{SNR}} + \frac{C}{\mathrm{SNR}^2}, \tag{3.12}$$

where A, B and C depend on the time-bandwidth product of the measurement. Note that in the case of infinite signal to noise ratio, the normalized variance is a constant, independent of the measurement.

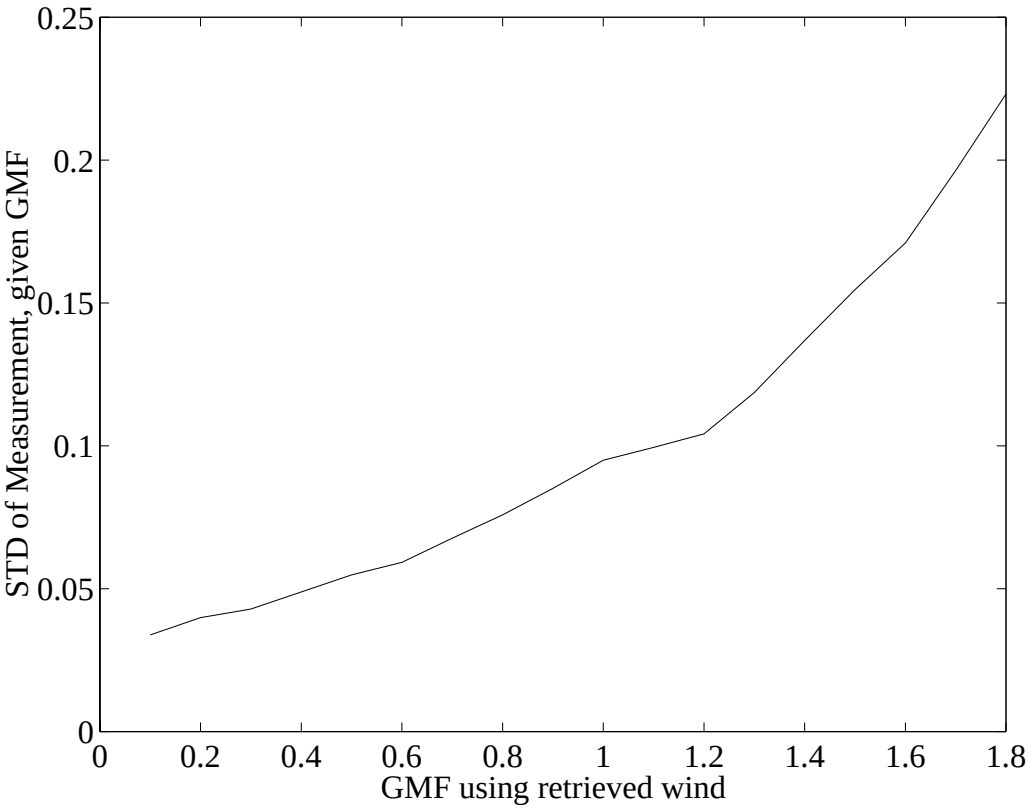


Figure 3.5: *The standard deviation of the measurements given the model function backscatter.*

Similarly, the variance of the measurements based on the digital processor used by NSCAT is found to have the same form for K_{PC}^2, though with more complicated expressions for the constants, A, B and C [42]. Also, the normalized variance in the form of K_{PC}^2 has been established for other digital processors with arbitrary modulation schemes [43].

3.2.2 Model Function Variability

The other source of variability to consider is the variability of the true backscatter, NRCS, given the average wind over the region. Unfortunately, the physical relationship between the wind and backscatter is not completely known. Here I describe some simple justifications of a multiplicative model for this variability.

One probable source of model function variability is due to non-uniform wind over the entire scatterometer footprint. Consider a 25 km square region (the size of an NSCAT resolution cell) with mean wind speed U_0. Because wind is turbulent it exhibits spatial and temporal fluctuations about this mean value. The variation in the wind over the region results in backscatter variability. In particular, consider the surface with N subregions of equal area, each with a constant wind speed, and consider the ith region as having wind speed $U_i = U_0 + u_i$ where the average wind speed, U_0, is modified by some additive, zero mean noise (for simplicity consider the u_i to be iid). The backscatter is then (using the model of Eq. (3.1) with $\alpha = 2$, again for simplicity):

$$\begin{aligned} \sigma^o &= \frac{a}{N}\sum_{i=1}^{N}(U_0 + u_i)^2 && (3.13)\\ &= aU_0^2 + \frac{a}{N}\sum_{i=1}^{N}u_i^2 + \frac{2aU_0}{N}\sum_{i=1}^{N}u_i && (3.14)\\ &= \sigma_M^o + a\left(\frac{N}{N-1}\right)s^2 + 2aU_0 m && (3.15) \end{aligned}$$

where $\sigma_M^o = aU_0^2$, s^2 and m are the sample variance and mean of u_i. In the limit of infinite subregions ($N \to \infty$), the true backscatter is related to the model function prediction, σ_M^o, as

$$\sigma_T^o = \sigma_M^o + a\mathrm{var}(U). \tag{3.16}$$

Determining the variance of U for a 25 km wind cell is non-trivial because of the lack of high resolution wind data. A rough estimate of the variance of U can be formed by considering a larger region, say 150 km (6 by 6 cells of size 25 km). Based on one half-rev of NSCAT data and using the first ambiguity of retrieved wind (though wind speed is fairly similar for all ambiguities), the sample mean and sample standard deviation of the wind speed were computed for over 15 000 regions. In Fig. 3.6, representative data points are plotted (roughly 2 000); also plotted is the 1st order polynomial fit to the data in a least-squares sense. While there is considerable spread in the scatter plot, the slight slope suggests the possibility that the standard deviation can be approximately written as being proportional to the mean wind speed.

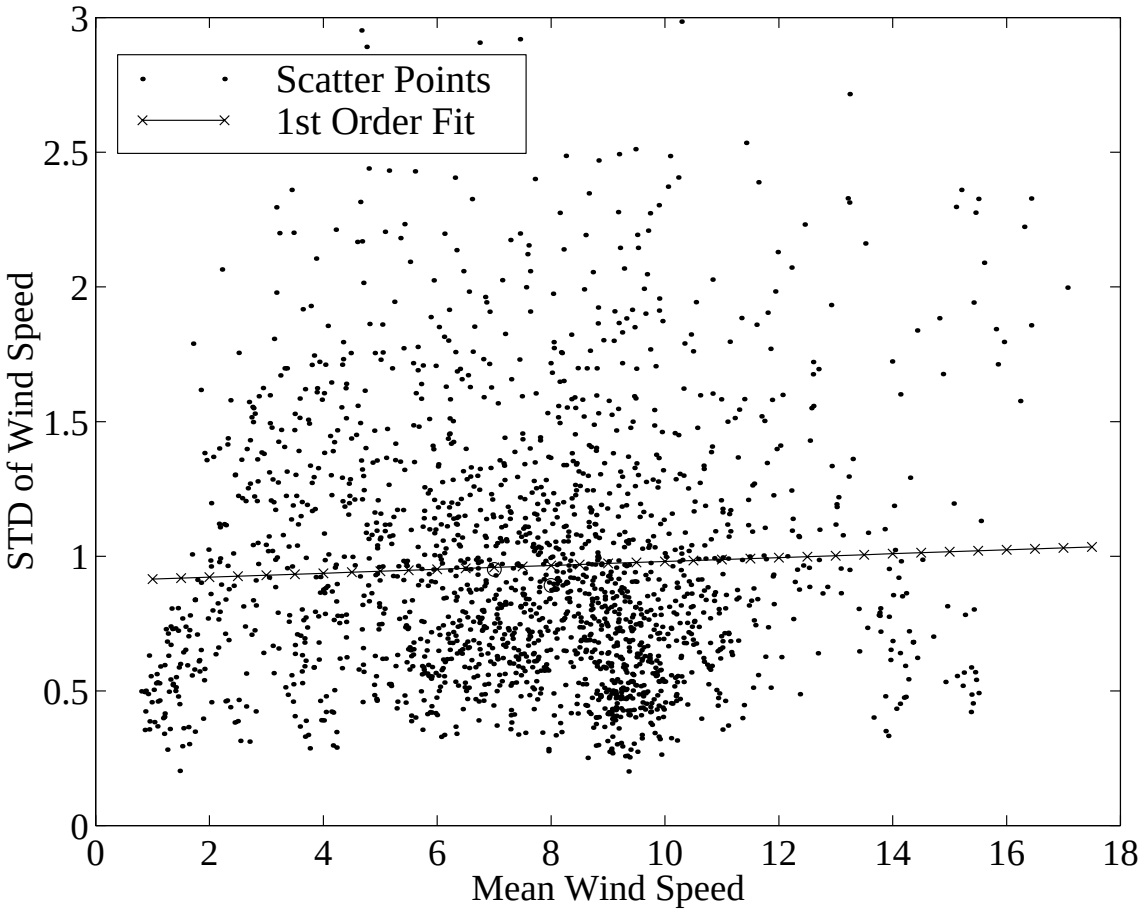


Figure 3.6: *The standard deviation of wind speeds vs the mean speed based on a region of 6 by 6 25 km cells.*

Thus, the true backscatter based on a region of non-uniform wind speeds is related to the model function backscatter as

$$\sigma_T^o \approx \sigma_M^o + \alpha a U^2 \tag{3.17}$$

$$= \sigma_M^o (1 + \alpha) . \tag{3.18}$$

The large and non-uniform spread of data points shown in Figure 3.6 shows that while the standard deviation of the wind speed has a small slope, there is some dependence on its mean. This suggests that the variability of the wind is actually a function of the backscatter (i.e., of the incidence angle, wind speed, etc.), just as the

communication noise, K_{PC}, is with finite SNR. Thus, the backscatter measurement could be written in any form where σ^o_M is modified by a mean-dependent random noise term.

Traditionally, radar systems involving fading statistics [44] model the noise as a multiplicative term because the communication noise in the limit of infinite signal to noise ratio is multiplicative with K_P independent of the backscatter [41]. Further, [45] provides an example of using a multiplicative model noise in the C-band model, CMOD4, used with ERS-2; this was justified in terms of an additive noise term in log-space, where the objective function is more linear and thus more suited to maximum-likelihood estimation.

One additional use of K_{PM} is in the general design of scatterometers. Considerable effort is expended in optimizing the integration time and modulation scheme to achieve a small value for K_{PC}. By introducing K_{PM} in a comparable way, the designers can recognize the futility of reducing K_{PC} far below the value of K_{PM}; if K_{PM} is much larger than K_{PC}, there is little point in working to reduce K_{PC}.

With these considerations in mind, the measurement model, including both communication noise and model function variability, is

$$z = \sigma^o_M \left(1 + K_{PC}\mu\right)\left(1 + K_{PM}\nu\right), \tag{3.19}$$

where z is the measured backscatter, σ^o_M is the model function prediction of the backscatter, based on incident conditions and wind conditions, μ and ν are independent, identically distributed, zero-mean, unit variance random variables, K_{PC} is the normalized standard deviation of the communication noise and K_{PM} is the normalized standard deviation of the model function uncertainty.

3.3 Model Variability Estimation Technique

As previously noted, the geophysical model function is not completely understood. Empirical estimates of the GMF, described in the previous section, describe the relationship between the wind and the backscatter on average, but a particular

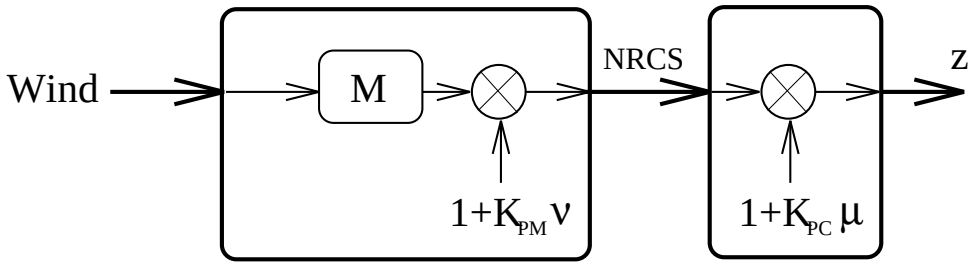


Figure 3.7: *Lacking a complete understanding of the geophysical model function, the relationship between the wind and the surface backscatter is modeled using the empirical model function and a model variability term, K_{PM}, included analogously to K_{PC} as a multiplicative term. The box on the left represents the impact of the model variability, while the box on the right shows the variability in the measurements caused by the communication noise.*

realization can be considerably modified by unmodelled parameters (e.g., swell, salinity, etc.) The simple model of Eq. (3.19) which describes the basic measurement process, including variability about the empirical model function, is depicted in Fig. 3.7 [36, 46]. The empirical model function, M, maps the surface wind, along with the parameters of the scatterometer, to the model function backscatter. This value is perturbed by unmodelled parameters, along with a zero-mean unit-variance normal random variable, ν, to yield the true backscatter coefficient of the surface. The statistics of this model are analyzed in this section and an original method developed for estimating K_{PM}. This method will be examined through simulated data and used to estimate K_{PM} for the NSCAT model function, NSCAT-1.

The expected value of each measurement, $\mathrm{E}(z)$, is σ^o_M, and the variability is observed by looking at higher order statistics. The variance of the measurement, z, is

$$\begin{aligned} \mathrm{var}(z) &= {\sigma^o_M}^2 \mathrm{var}(1 + K_{PM}\nu + K_{PC}\mu + K_{PM}K_{PC}\nu\mu) && (3.20) \\ &= {\sigma^o_M}^2 (K^2_{PM} + K^2_{PC} + K^2_{PM}K^2_{PC}) && (3.21) \end{aligned}$$

so

$$K^2_{PM} = \frac{\mathrm{var}(z)}{{\sigma^o_M}^2(1 + K^2_{PC})} - \frac{K^2_{PC}}{1 + K^2_{PC}}. \qquad (3.22)$$

For a narrow range of observation geometries and wind conditions, σ_M^o and K_{PC} are nearly constants, $\overline{\sigma_M^o}$ and $\overline{K_{PC}}$. Defining a sample variance in terms of these small data set averages,

$$SV = \text{var}\left(\frac{z}{\overline{\sigma_M^o}\sqrt{1+\left(\overline{K_{PC}}\right)^2}}\right) \tag{3.23}$$

and a sample mean of the corresponding data

$$SM = \frac{\left(\overline{K_{PC}}\right)^2}{1+\left(\overline{K_{PC}}\right)^2} \tag{3.24}$$

an estimate of K_{PM}^2 for this data segment is obtained by subtracting the two. This simplistic approach results in an estimate of K_{PM} dominated by variations in the measurement, z, but because the variability of σ_M^o and K_{PC} are ignored, several data segments are used to generate SV, SM and the corresponding estimate of K_{PM}. The average of these estimates is used as the final estimate:

$$\widehat{K_{PM}^2} = \overline{(SV - SM)}. \tag{3.25}$$

The variance of a sample variance is given as [13]

$$\text{var}(SV) = \frac{2\sigma^4}{n-1} \tag{3.26}$$

where σ^2 is the variance of the underlying random variable, and SV provides an unbiased estimate of the variance. The variance of SM, for a narrow band of wind and measurement conditions, is negligibly small. Thus, from SV and SM an estimate of K_{PM}^2 and its variance are computed; several such estimates can be averaged to improve the estimate.

One caveat that must be recognized is that neither the true wind or the real value of K_{PM} are known when analyzing scatterometer data. The estimated wind is a random quantity (since it is computed from the estimated wind); thus, by inheritance, σ_M^o is also random (recall Fig. 3.7). The mean of σ_M^o is based on the assumed value of K_{PM}, since this will alter the estimates of the wind. To ameliorate the impact of this circular effect, an iterative procedure was developed in which

1. A value of K_{PM} is assumed.

2. The wind is estimated based on the assumed value.

3. The mean and variance of $\widehat{K_{PM}^2}$ are estimated as described above.

4. This mean is now used as the assumed value in a new wind estimation.

This procedure is repeated until the estimated value of K_{PM} converges.

This iterative procedure works well to converge to a solution, but the solution is biased because of the variance of the retrieval process. Using the retrieved (rather than the true) winds results in an underestimate of K_{PM}^2. The variance of the retrieval process causes a variance in σ_M^o; in Eq. (3.22), such a variance would reduce the estimate of K_{PM}^2. Admittedly, this is a simplistic analysis describing the effect of the variance of the retrieval process on the estimate of K_{PM}, but the *ad hoc* correction factor that follows only requires the general effect of using estimated rather than true winds. To correct for this effect, a simple quadratic function is used and the fundamental theorem of statistics invoked to compute the derived distribution of the corrected function. Extensive simulations show this to be a reliable method of estimating the model function variability.

To examine the effectiveness of the estimation procedure described above, first consider a simulation. In the simulation, a wind field based on typical winds reported by numerical weather prediction models is used, along with the geometry and noise levels of actual satellite observations. Introducing a reasonable value of K_{PM} in the simulation, known terms include the following:

- The true (simulated) wind and measurement geometry
- The backscatter from the empirical model function, σ_M^o
- The model function standard deviation, K_{PM}
- The communication noise standard deviation, K_{PC}
- The simulated measurement, z, (generated according to Monte Carlo techniques.)

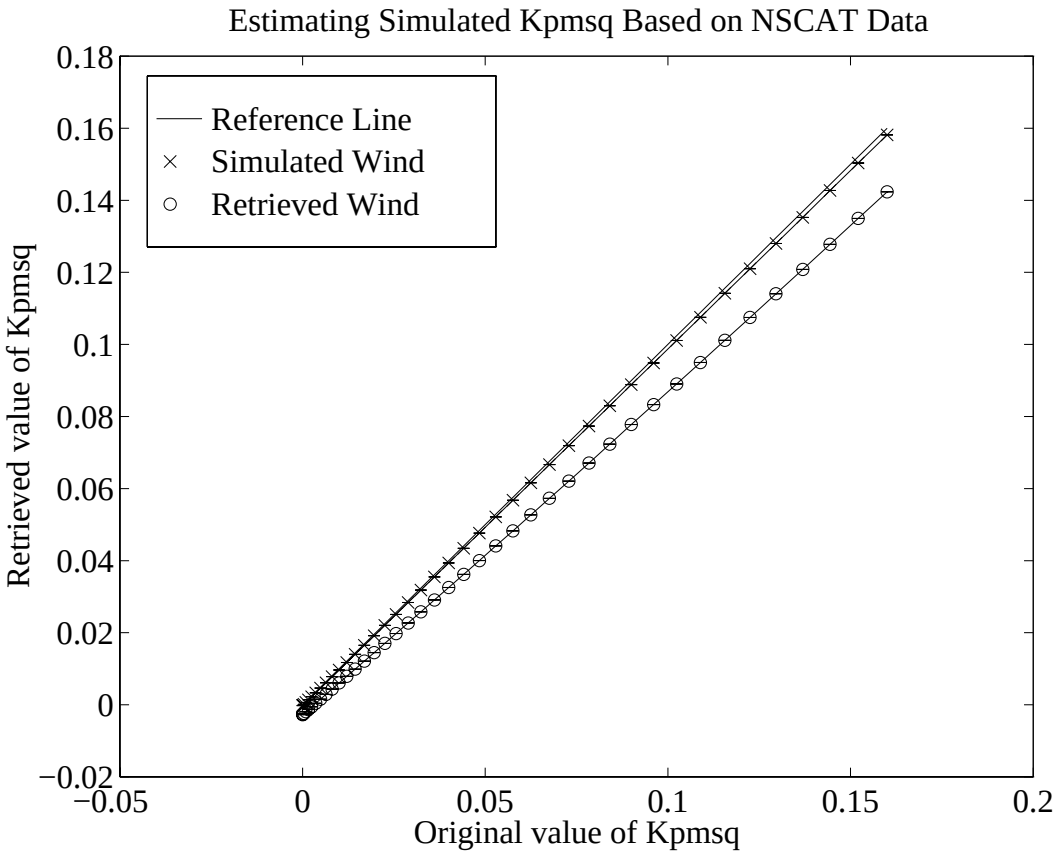


Figure 3.8: *Simulation based on NSCAT data.*

The estimation procedure is then applied to estimate K_{PM}, and the accuracy of the method is quantified. Note that here there is no need for the recursive estimation of K_{PM}^2, since the correct value to use in the wind estimation is known.

Although NSCAT processing at JPL assumes a nominal value of $K_{PM} = 0.17$ [36], there is no reported analysis of K_{PM} for NSCAT. There is also no description of the sensitivity of K_{PM} to various parameters. Having removed JPL's assumed value, I simulate my own values and attempt estimation of K_{PM}. Figure 3.8 shows the NSCAT simulation results, with a reference line, a curve based on the simulated wind and a curve based on the retrieved wind. A simple quadratic correction compensates well for the slight bias and curvature in the estimation procedure, allowing the model function variance to be estimated based on retrieved wind.

Defining x to be K_{PM}^2 estimated with the retrieved wind, and

$$y = \sqrt{ax^2 + bx + c} \tag{3.27}$$

to be the corrected value of the estimate of K_{PM} and invoking the fundamental theorem of statistics, the mean and variance of y can be approximated in terms of the mean, μ_x, and variance, σ_x^2, of x.

The mean of y is

$$\begin{aligned} \mu_y &= \int f_y(y) y \; dy \\ &= \int f_x(x) y \; dx \\ &= \int f_x(x) \sqrt{ax^2 + bx + c} \; dx \\ &\approx y_0 + \frac{a\sigma_x^2}{2y_0} - \frac{(2a\mu_x + b)^2 \sigma_x^2}{y_0^3} \end{aligned} \tag{3.28}$$

where $y_0 = \sqrt{a\mu_x^2 + b\mu_x + c}$ and y is expanded about $x = \mu_x$. This approximation is based on the assumption that the distribution of x is symmetric and highly localized about μ_x. Experience has shown that this is generally an accurate description of K_{PM}^2, except at lower values of the mean. Although the distribution of K_{PM}^2 is always unimodal, the necessity that $K_{PM}^2 > 0$ skews the distribution for smaller values, reducing the accuracy of the approximation. While this is not a particularly satisfying way to implement the correction factor, it appears to work well when moderate values are being estimated. The estimate of the variance of y follows the same procedure to yield

$$\sigma_y^2 \approx (2a\mu_x + b)^2 \left(\frac{a\sigma_x^4}{y_0^4} + \frac{2\sigma_x^2}{y_0^2} \right) - (2a\mu_x + b)^4 \frac{\sigma_x^4}{y_0^6} - \frac{a^2 \sigma_x^4}{4y_0^2}. \tag{3.29}$$

3.4 Data-Driven Estimates of K_{PM}

With a statistical method for estimating the model function variability, and the necessary correction factor determined, in this section I estimate K_{PM} for the Ku-band geophysical model function (NSCAT1) and identify which parameters affect it. Specifically, I examine the effects of wind speed, incidence angle, latitude and time.

Having binned the data according to 20° latitude bins, 5° incidence angle bins, and 2 m/s wind speed bins, K_{PM} was estimated based on 40 half-revs of NSCAT data for each week examined. Note that examination of ascending and descending portions of the data resulted in comparable results for K_{PM}; the results presented here are based on descending data. Further recognize that a typical week results in approximately 100 revs; 40 revs provided a substantial sample of the week's data while reducing the computation load of using the full week of data.

The estimate of K_{PM} depends on the wind speed. Different regions of the world, though, experience different wind speed distributions. Equatorial regions, for example typically have lower wind speeds than upper latitude regions. Further, these distributions change with seasons. Some bins, therefore, have little or no data. Figures C.1 through C.7 (in Appendix C) display histograms of the wind speed for each week of data (based on the 40 descending revs). Histograms for each week are provided for six latitude bands, -60 to -40, -40 to -20, -20 to 0, 0 to 20, 20 to 40, and 40 to 60. The plots are arranged such that the southern hemisphere plots are on the left and the northern hemisphere on the right, with the equatorial regions at the bottom and the high latitudes at the top. The histograms show that the southern hemisphere, particularly at high latitudes, have many more valid data points (due to less land cover), and that the distributions are distinctly non-Rayleigh (the theoretical distribution of wind speed) and that they vary from week to week. However, combining the worldwide wind speeds for any particular week in a single histogram yields a very precise Rayleigh distribution. Of particular interest is the fact that there is a clear upper limit for the wind speed which is considerably higher for higher latitudes. Thus, estimates of K_{PM} for higher wind speeds and lower latitudes will be severely restricted by the limited amount of data.

Binning the data as described above results in a four-dimensional array for K_{PM}. That is, K_{PM} is estimated as a function of latitude, incidence angle, wind speed and week of the mission. This requires several plots to display the estimates. For clarity of presentation, these plots are archived in Appendix C.

Figures C.8 through C.14 display plots of K_{PM} vs. wind speed, for the sampled weeks, with a separate line for each incidence angle. The estimate of K_{PM} is higher for higher incidence angles. Note the strong correlation of the estimate with wind speed, K_{PM} decreases with increasing wind speed. The estimation procedure was hard-limited to a maximum value of 0.6, as evident in the plots. Further, the less reliable estimates at high wind speeds and low latitudes is clear.

To provide another view of the data, Figs. C.15 through C.22 display the estimates of K_{PM} as functions of time through the mission, with a separate line for each wind speed and a separate plot for each incidence angle. Note that there is very little temporal dependence for the estimates, indicating that K_{PM} is not strongly dependent on seasonal variations.

As one additional view of the data, Figs. C.23 through C.31 display the estimates of K_{PM} as functions of time through the mission, with a separate line for each incidence angle and a separate plot for each wind speed bin. Temporal variations that appear in these plots seem to be results of regions with insufficient data.

3.5 Impact on Wind Retrieval

As described earlier, empirical model functions yield unique backscatter values for given wind and measurement conditions, but the mapping is many-to-one. To estimate the wind, maximum likelihood estimations (MLE) is used to identify the wind most likely to have generated a set of measurements, as described in Appendix D.

The technique developed above to estimate, K_{PM} from scatterometer measurements uses, in place of the true wind, the estimated wind. The log-likelihood function used to retrieve the wind (see Appendix D), displays a dependence on K_{PM} through the variance of the measurements. This results in a complex relationship between the estimates of wind and K_{PM}. The iterative process outlined in Section 3.3 is unwieldy for researchers interested in only the wind.

Here the impact of K_{PM} on wind retrieval is examined [47]. If the impact is minimal, the complex relationship can be ameliorated by setting K_{PM} to be a

constant (say, 0) for the purposes of wind retrieval. Also considered is the possibility of removing the log variance term from the log-likelihood function (reducing the problem to weighted least squares). This will dramatically reduce the dependence of the likelihood function on K_{PM} (again simplifying the computation), but will also alter the optimality conditions of the maximum-likelihood estimator.

Compass simulations were performed in which a wind vector generated a simulated ocean surface via the empirical model function. This surface was corrupted by random noise with a simulated K_{PM} value, and the resulting NRCS was measured with a noisy scatterometer (K_{PC} and other measurement parameters based on typical NSCAT values) and the wind estimated from such measurements. Figures 3.9, 3.10, and 3.11 display the average results of 100 000 simulations by plotting the retrieved speed and direction errors (simulated minus closest retrieved ambiguity) for three scenarios. The solid line (labeled 'a') is for simulation and retrieval with $K_{PM} = 0$. The dashed line (labeled 'b') uses $K_{PM} = 0.2$ in the simulation, but $K_{PM} = 0$ in the retrieval. That is, 'b' is comparable to real-world estimation where there is uncertainty in the model function, but it is not accounted for in the retrieval process. Finally, the dash-dot line (labeled 'c') simulates surfaces with modeling error, and uses it in the retrieval process (assuming that it is known exactly rather than having to estimate it).

These plots show that with the unrealistic case of $K_{PM} = 0$ (there are always unknown and unmeasured parameters that are not incorporated in empirical model functions) the wind speed estimate is asymptotically unbiased (essentially); while including the model uncertainty yields biased wind speed estimates. Even if K_{PM} is known exactly and accounted for in the estimation, the retrieved speed is biased high. The result is similar for wind simulated at 10 m/s. The retrieved direction errors are modified by K_{PM}, but in non-predictable ways. The same behavior is evident at mid and far swath, though is less pronounced.

While not dramatically biasing the wind estimates, including K_{PM} in the estimation procedure does have a significant impact on the shape of the log-likelihood function. Specifically, it changes the relative magnitudes of the ambiguities, sometimes completely suppressing lower-valued ambiguities and sometimes shifting the

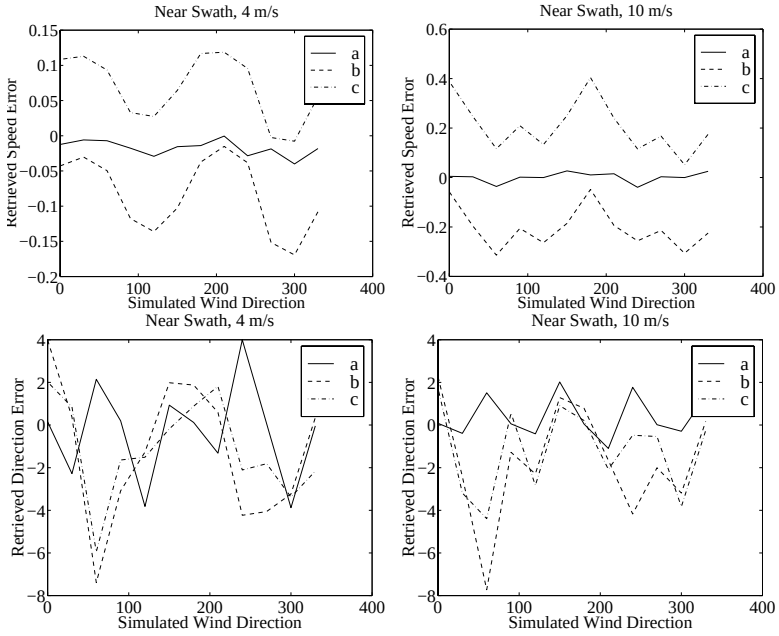


Figure 3.9: *Impact of K_{PM} on simulated wind estimation at near swath.*

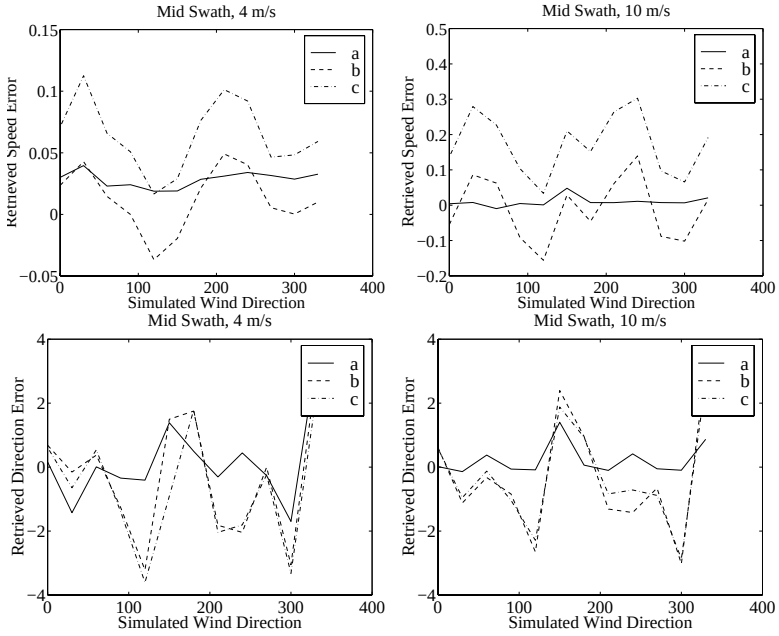


Figure 3.10: *Impact of K_{PM} on simulated wind estimation at mid swath.*

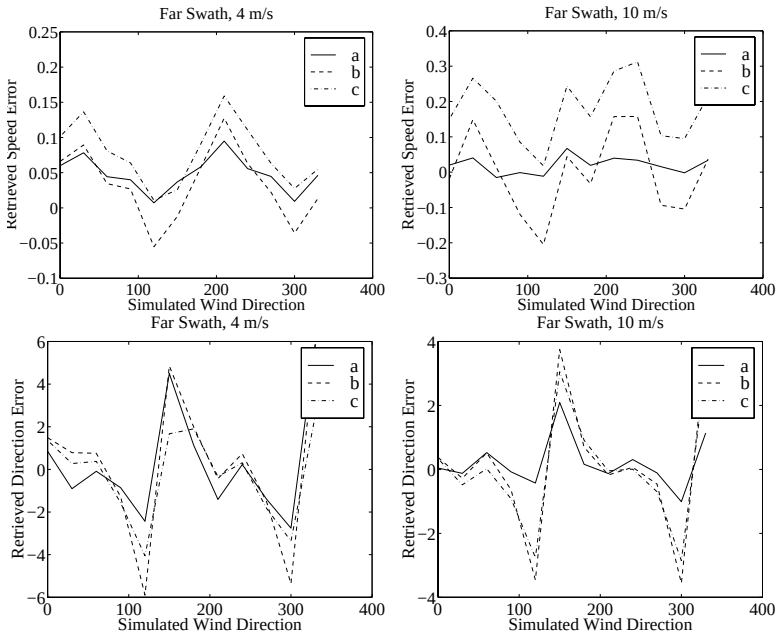


Figure 3.11: *Impact of K_{PM} on simulated wind estimation at far swath.*

ranking of the ambiguities. Recall that all local peaks are kept as possible wind solutions, but the ranking is useful in some situations.

For a typical measurement geometry, a compass simulation was performed based on a wind speed of 8 m/s and various wind directions; the log-likelihood function, normalized to have a maximum of 1, was plotted assuming four values of K_{PM} in Fig. 3.12. Twelve subfigures show the normalized log-likelihood function for simulated directions of 0 through 330°. Note that with $K_{PM} = 0$ in the estimation process (the solid line), the ML estimate of wind direction, that is, the first ambiguity, corresponds to the simulated direction as expected.

Compass simulations were also performed with the log variance term removed from the estimator, with the results displayed in Figs. 3.13, 3.14, and 3.15 for near, mid, and far swath locations, respectively. Again, if K_{PM} is non-zero, there is a small bias in the speed estimates (on the order of 2%). Note that there is now much less dependence on whether or not K_{PM} is included in the wind retrieval.

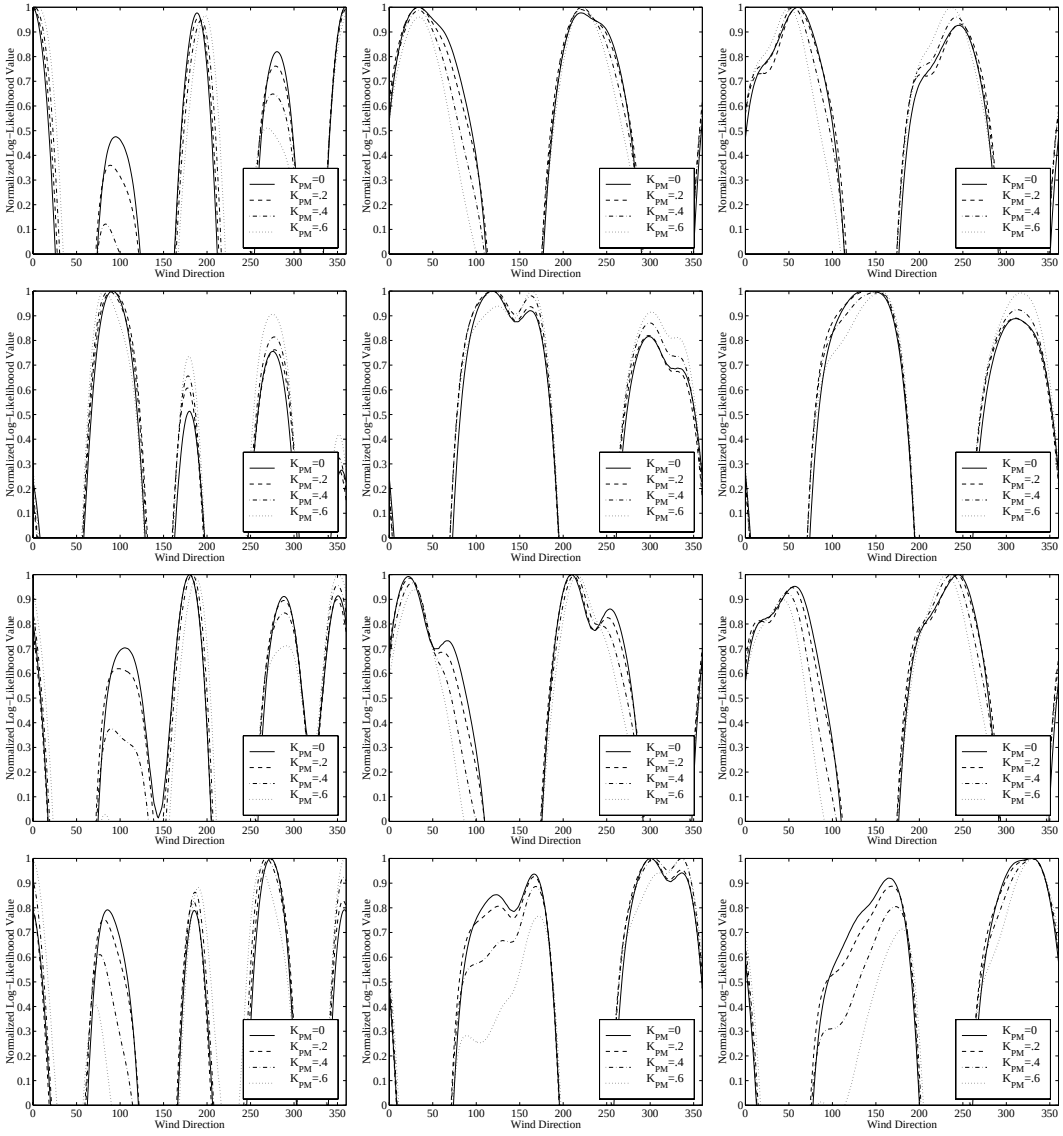


Figure 3.12: *Impact of K_{PM} on the log-likelihood function (normalized to have a maximum of 1).*

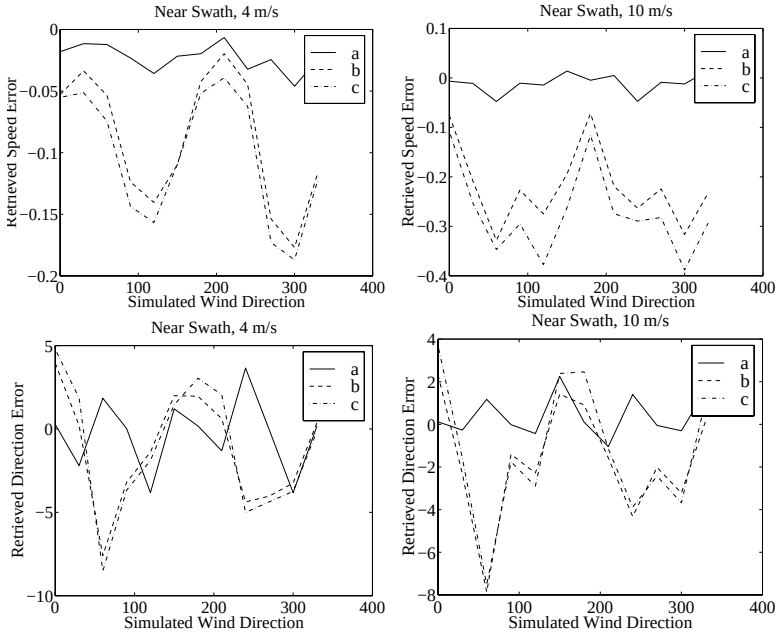


Figure 3.13: *Impact of removing the log variance term on simulated wind estimation at near swath.*

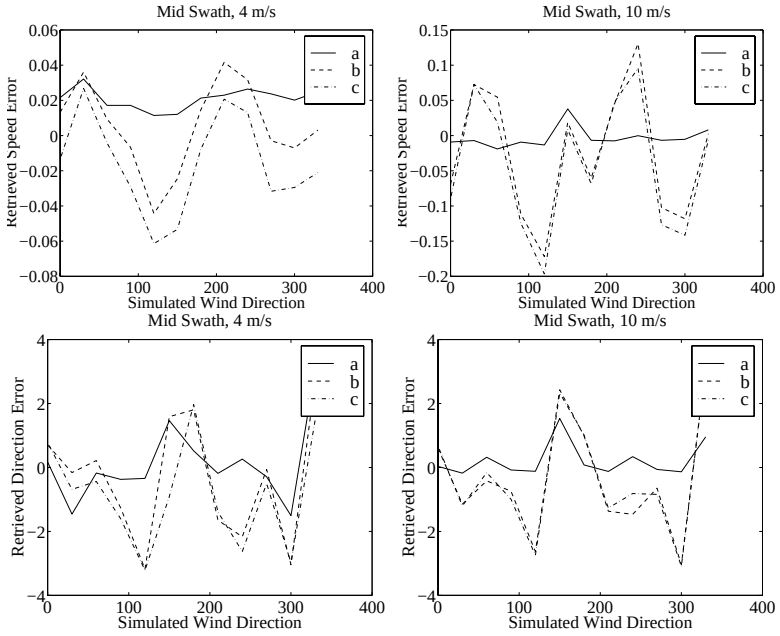


Figure 3.14: *Impact of removing the log variance term on simulated wind estimation at mid swath.*

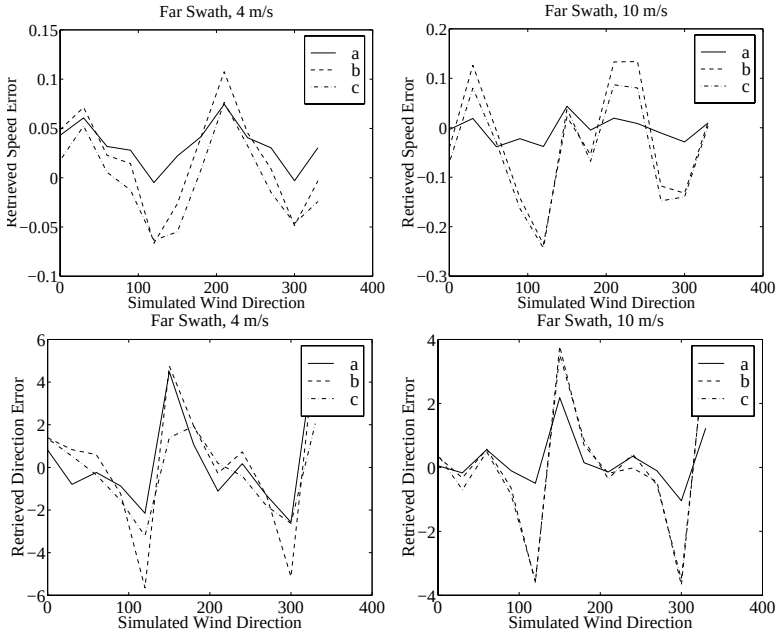


Figure 3.15: *Impact of removing the log variance term on simulated wind estimation at far swath.*

The changes caused by model function variability on the log-likelihood function, and on the amount of error in a particular realization of the wind retrieval process, can be quantified with the Cramer-Rao (CR) lower bound. The CR bound states the minimum limit of the error covariance matrix of an unbiased estimator [48]. While earlier plots indicate that wind estimation including K_{PM} is slightly biased, the CR bound provides a reasonable measure of the error covariance of wind estimates.

The CR bound has been developed in detail for wind estimation [46]. The error covariance matrix, C, of an unbiased estimator is bounded by the inverse of the Fischer information matrix, J:

$$\mathbf{C} = E\left[\hat{\theta}-\theta\right]\left[\hat{\theta}-\theta\right]^T \tag{3.30}$$

$$\geq \mathbf{J}^{-1} \tag{3.31}$$

$$= \frac{1}{J_{11}J_{22}-J_{12}J_{21}}\begin{bmatrix} J_{22} & -J_{21} \\ -J_{12} & J_{11} \end{bmatrix}. \tag{3.32}$$

The parameter vector, θ, can be defined either in terms of rectangular or polar wind coordinates; I use the polar form: θ is the vector of wind speed and wind direction. The Fischer information matrix describes the variation of the log-likelihood function with respect to the two-dimension, vector wind (or, in general, to the parameter being estimated).

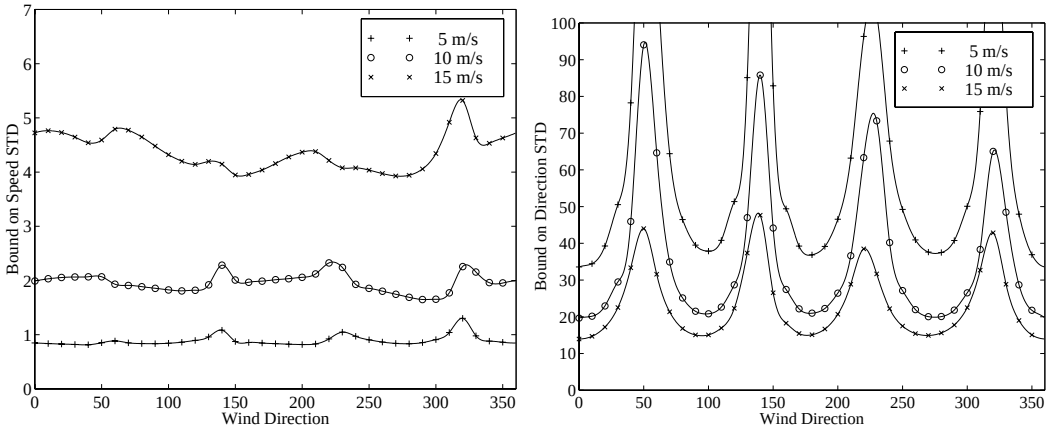
Oliphant [46] showed that for a set of K independent backscatter measurement, assumed to have a Gaussian distribution with mean $\mathbf{m}$, a vector of the NRCS measurements, and covariance $\mathbf{R}$, a matrix with diagonal entries of the variance of the K measurements and off-diagonal terms equal to zero, then the elements of the Fischer information matrix can be expressed as

$$J_{ij} = \frac{1}{2}\mathrm{tr}\left(\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \theta_i}\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \theta_j}\right) + \frac{\partial \mathbf{m}^T}{\partial \theta_i}\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \theta_j}. \tag{3.33}$$

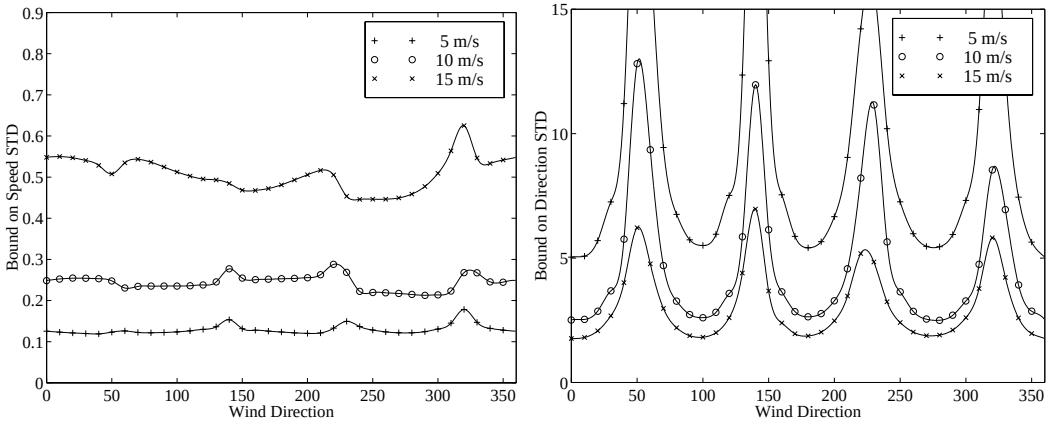
It should be noted that the CR bound does not depend on actual scatterometer measurements, but only on the model function and typical noise parameters for the instrument at each observation geometry. Because the CR bound depends so strongly on the noise, an increase in the model variability results in a substantial increase in the error bounds.

Figure 3.16 displays the CR bound for a typical NSCAT measurement, using 4 measurements in a 25 km near-swath cell. If there is no model function variability (the bottom plots), the bounds on the standard deviations of the wind speed and wind direction are really quite small. The large peaks correspond to winds that align with one of the scatterometer antennas [46]. In those directions, the assumption of negligible bias in the wind estimates fails. While the CR bound has been extended to include biased estimators [49], Fig. 3.16 clearly displays the general increase in the error covariance with increasing K_{PM}. The top figures show the CR bound on speed and direction. When there is substantial model function variability, $K_{PM} = 0.3$, the bounds are much higher than the case shown in the bottom two plots of no variability.

To further demonstrate the large increase in the std of wind estimates caused by K_{PM}, Fig. 3.17 plots the CR bounds for wind speed and direction as functions of K_{PM} value. Note that the same measurement geometry and noise values



(a) $K_{PM} = 0.3$ (b) $K_{PM} = 0.3$



(c) $K_{PM} = 0$ (d) $K_{PM} = 0$

Figure 3.16: *Plots of the Cramer-Rao lower bound of the standard deviation of the errors in wind speed and wind direction. The bottom plots, with* $K_{PM} = 0$*, yield much smaller bounds for the errors than the top plots, with* $K_{PM} = 0.3$*. Note that this is quite a large value of* K_{PM}*, resulting in the large difference in the vertical scales used in the figures—*$K_{PM} = 0.3$ *results in approximately 7 times larger standard deviations in both speed and direction.*

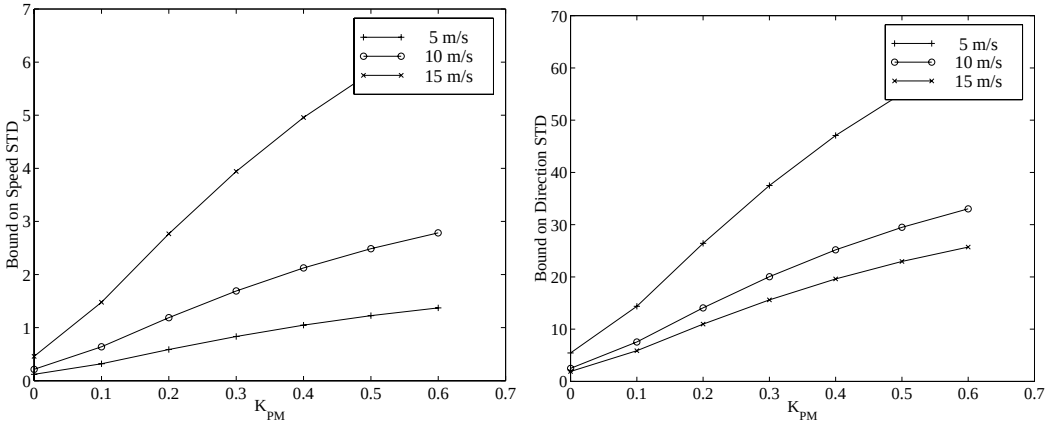


Figure 3.17: *The Cramer-Rao lower bound of the standard deviation of the errors in wind speed and wind direction are highly correlated with the model function variability, K_{PM}. Increasing K_{PM} substantially increases the bound on the errors. This plot was made based on the geometry of a near-swath cell, with the wind direction at* 280°.

as in Fig. 3.16 were chosen; the wind direction was set to 280° in order to avoid bias effects.

It was previously noted that the value of K_{PM} had little impact on the mean of the closest ambiguity. However, because the log-likelihood function is modified by K_{PM}, the possibility of changing the rankings of the ambiguities exists. In many applications, all the ambiguities are kept as possible solutions and used to identify large scale wind fields (see, for example, Appendix H and [50]). The ambiguity skill describes how frequently an ambiguity is closest to the true wind. The value of K_{PM} has the potential of dramatically reducing the ambiguity skill.

In order to quantify the effect of K_{PM} on the ambiguity skill, a simulation was performed in which a wind vector was observed using measurement geometry and noise figures typical of NSCAT at mid swath. The value of K_{PM} was varied and several wind estimates were identified with the log-likelihood function. The highest peak of the log-likelihood function was noted as the first ambiguity. The peak closest to the correct solution was also noted. The first ambiguity skill was computed as the percentage of cases in which the first ambiguity corresponded to the peak closest to

the true wind. Winds were simulated using a true wind speed of 8 m/s and azimuth directions of 0, 30, 60 and 90 deg. Comparable skill levels were observed at other wind speeds.

Figure 3.18 displays the results of the simulation using 1000 noise realizations for each data point. The plot shows the percentage of runs in which the highest valued peak of the log-likelihood function corresponded to the correct ambiguity. That is, the vertical axis displays the first ambiguity skill. If the measurements are not corrupted by K_{PM} (i.e., the NRCS is precisely determined by the wind alone) then the first ambiguity skill is high, nearly 90% at 0 and 90 degrees, and even higher at intermediate wind directions. However, in the realistic case where K_{PM} is not zero, the log-likelihood function is modified such that the highest peak is not closest to the simulated wind. For large values of K_{PM}, the first ambiguity skill is reduced for all wind directions. Recalling the plots of the log-likelihood function in Fig. 3.12, the peaks in the function for winds at 0 and 90 degrees are sharper than those at intermediate angles. This makes the first ambiguity skill at these wind directions more sensitive to changes in the noise level than other wind directions.

3.6 Conclusions

Recognizing the difficulties in deriving an accurate theoretical surface spectrum from geophysical parameters, empirical approximations of the relationship between the wind and NRCS have been developed, bypassing the need for immediate knowledge of the surface spectrum. Such an approach provides a reasonable and practical way to estimate the wind from backscatter measurements. Unmodelled geophysical parameters that affect the backscatter introduce uncertainty in the wind estimation procedure.

In this chapter, I developed a technique to estimate the model variability, K_{PM}, from scatterometer data alone. Binning the estimates provides a view of which parameters change the backscatter from its expected value. I binned the estimates of K_{PM} according to latitude, wind speed, incidence angle and time in the NSCAT mission, and computed estimates of K_{PM}, which were commonly on the order of 0.3

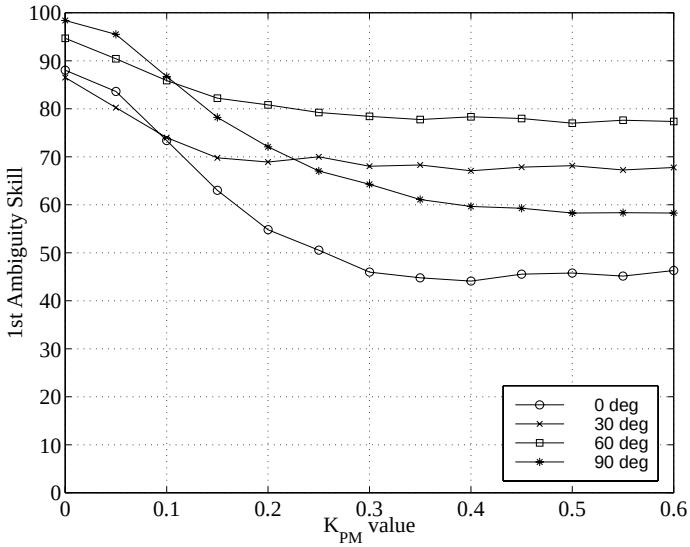


Figure 3.18: *K_{PM} alters the shape of the log-likelihood function and changes the relative amplitudes of the peaks. This can cause low ranking ambiguities (when $K_{PM} = 0$) to be selected as the MLE when K_{PM} is large.*

and sometimes as high as 0.6. This is extremely high, particularly when compared with the communication noise, K_{PC}.

The impact on wind retrieval was also explored. While the estimates of K_{PM} are often quite large, the retrieved wind was shown through compass simulations to be only slightly biased, on the order of a few percent. However, because K_{PM} alters the log-likelihood function substantially, the error possible in an individual retrieval is larger than previously believed and the relative ranking of the ambiguities is often changed by model function variability. The CR bound on the error covariance matrix was computed, and found to be very sensitive to model function variability. A possibly more significant problem caused by modeling error is the sensitivity of the ambiguity removal skill to the value of K_{PM}.

Chapter 4

The Probability Density of Spectral Estimates Based on Modified Periodogram Averages

Accurate power measurements are a critical component of microwave remote sensing. The quality of estimates of the geophysical parameters is directly related to the quality of the measurements. The observed backscatter power is a function of the surface. Because the surface is random, the power measurement is actually an estimate of the power based on a particular sampling time and bandwidth; it is an estimate with some probability distribution, traditionally assumed to be Gaussian. Using the correct distribution can be critical in many estimation theoretic contexts.

The NSCAT power measurement is based on Welch's estimate of a signal spectrum. Welch's technique of averaging modified periodograms for spectral estimation has been widely applied for over 30 years, [51], but without a thorough statistical analysis of the technique. Indeed, only two papers extend beyond the mean and variance of the spectral estimates. The first, [52], describes the probability distributions for discrete Fourier spectra based on a single periodogram, for data both smoothed and unsmoothed, but doesn't describe the correlation between frequency bins. The second, [53], derives the joint density functions for two frequency bins, including windowed data and averaging over non-overlapping data periodogram estimates.

This chapter analyzes the statistics of spectral estimates using Welch's method by deriving the pdf of the sum of frequency bins in great generality. The analysis is readily applicable to many situations involving modified periodograms

for spectral estimation. In Chapter 5, the specific case of NSCAT measurements is considered to describe the correct pdf for those measurements and the impact of using approximate pdfs for geophysical parameter estimates [54].

4.1 Overview

For a data stream, $x[j]$, Welch's modified periodogram averaging provides a technique for estimating the spectrum [51]. The sequence is segmented into K overlapping sequences, each of length L, such that the ith sequence is defined by $x[i,n] = x\left[n + i(1-r)L\right]$, where the indices are $i = 0, \ldots, K-1$ and $n = 0, \ldots, L-1$; r is the percentage of overlap. The transform, or periodogram, of this sequence is labeled $X[i,k]$. The data sequence is windowed to minimize spectral leakage. The window can be applied through multiplication in the time-domain ($w[n]$), or through circular convolution in the frequency-domain ($W[k]$). The estimate of the power spectral density of each segment, $P[i,k]$, is

$$P[i,k] = \frac{1}{LU}\left|X[i,k] * W[k]\right|^2 \tag{4.1}$$

where U is a scale factor dependent on the window [51]. To obtain the power in a particular frequency band, several bins of the modified periodogram may be summed. Further, the variance of the estimate is reduced by averaging over i, all of the data segments.

In this chapter, the statistical properties of this method of spectral estimation are explored for the case where $x[j]$ is a wide-sense stationary, Gaussian sequence, with distribution $\mathrm{N}(0, \mathbf{R})$. The process is formulated in vector space notation, from which the resulting power estimate is shown to be a quadratic form in the data vector $\mathbf{x}$. The pdf corresponding to this quadratic form is given in generality and several representative examples are plotted. Simulations are reported in which sample pdf's of the power estimates based on averaging modified periodograms match the theoretical functions derived here. Finally, I compute the Kullback-Leibler distance of this density function from a Gaussian.

4.2 Vector Space Analysis of Welch's Spectrum Estimation

To derive the pdf of a spectral estimate of the power in a frequency band of a data stream, it is useful to consider the vector space formulation of the estimate. For the real, length N, data vector $\mathbf{x} = [x_0, x_1, \cdots, x_{N-1}]^T$, distributed Gaussian with zero-mean and covariance matrix $\mathbf{R}$, a data window matrix $\Omega = \mathrm{diag}(w[n])$ is commonly used to filter the data. Welch's method segments $\mathbf{x}$ into data segments, possibly overlapping, of length L. For an arbitrary data segment, i, the periodogram estimate of the power spectral density in frequency bin k can be written as

$$\begin{aligned} P[i,k] &= \mathbf{x}^T \Upsilon[i,k] \mathbf{x} && (4.2) \\ &= \mathbf{x}^T \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{LU}\Omega\left(\mathbf{C}_k + \mathbf{S}_k\right)\Omega & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{x} && (4.3) \end{aligned}$$

where the sizes of the $\mathbf{0}$ matrices are adjusted to pick which data segment is being used and the matrices $\mathbf{C}_k$ and $\mathbf{S}_k$ are from the cosine and sine transform kernels. The (m,n)th element of $\mathbf{C}_k + \mathbf{S}_k$ is

$$\left(\mathbf{C}_k + \mathbf{S}_k\right)_{m,n} = \cos\left(2\pi \frac{k}{Z}\left(n - m\right)\right), \quad (4.4)$$

where Z is the length of the zero-padded data vector (i.e., if $Z = L$, each data segment has not been padded).

Averaging the power from the data segments and summing over several frequency bins yields a quadratic form in the normal random vector $\mathbf{x}$:

$$\begin{aligned} P &= \frac{1}{K}\sum_{i=0}^{K-1}\sum_{k=k_1}^{k_2} P[i,k] && (4.5) \\ &= \mathbf{x}^T \Upsilon \mathbf{x} && (4.6) \end{aligned}$$

where the matrix Υ depends on the length of the data segment, the amount of zero-padding, the data window, the frequency bins of interest, the number of periodograms to average, and the amount of overlap of the data segments:

$$\Upsilon = \frac{1}{K}\sum_{i=0}^{K-1}\sum_{k=k_1}^{k_2} \Upsilon[i,k]. \quad (4.7)$$

Formation of the matrix Υ can be facilitated for arbitrary overlap by partitioning the fundamental matrix of Eq. (4.3), $\frac{1}{LU}\Omega\left(\mathbf{C}_k + \mathbf{S}_k\right)\Omega$, such that overlapping partitions from different data segments can be added. For example, with non-overlapping data ($r = 0$), partitioning is unnecessary and Υ is a block diagonal matrix which is formed as the Kronecker product

$$\Upsilon = \mathbf{I}_K \otimes \left[\Omega\left(\frac{1}{KLU}\sum_{k=k_1}^{k_2}\mathbf{C}_k + \mathbf{S}_k\right)\Omega\right]. \tag{4.8}$$

For 50% overlap ($r = 0.5$), the fundamental matrix can be partitioned into four submatrices, (each of $\frac{L}{2}$ elements square).

The moment-generating function of a quadratic form in a zero-mean Gaussian vector, such as Eq. (4.6), can be expressed as [48, p. 65]

$$M(t) = \prod_{i=1}^{D}(1 - t2\eta_i)^{-\frac{\nu_i}{2}} \tag{4.9}$$

where the η_i are the D distinct, non-zero eigenvalues of $\mathbf{R}\Upsilon$. Each eigenvalue has multiplicity ν_i.

The mean and variance of the power estimate, P, are respectively,

$$\mu_P = \sum_{i=1}^{D}\eta_i\nu_i \tag{4.10}$$

$$\sigma_P^2 = \sum_{i=1}^{D}2\eta_i^2\nu_i. \tag{4.11}$$

Defining some convenient variables,

$$d_i = -\frac{1}{2\eta_i} \tag{4.12}$$

$$g = \prod_{i=1}^{D}(2\eta_i)^{-\frac{\nu_i}{2}} \tag{4.13}$$

the pdf corresponding to Eq. (4.9) is (see Appendix)

$$f(y) = \frac{g}{,\ \left(\frac{J}{2}\right)}y^{\frac{J}{2}-1}\Phi_2\left(\frac{\nu_1}{2},\cdots,\frac{\nu_D}{2};\frac{J}{2};d_1y,d_2y,\cdots,d_Dy\right) \tag{4.14}$$

where Φ_2 is the generalized hypergeometric function and J is the total number of non-zero eigenvalues.

While Eq. (4.14) provides a completely general solution for the pdf, calculation of the hypergeometric function is computationally intensive. However, if all of the non-zero eigenvalues have even multiplicities, $M(s)$ can be expanded with a partial fraction expansion. This yields a simple and practical density function (see Appendix F):

$$f_P(p) \;=\; g\sum_{i=1}^{D}\sum_{j=1}^{h_i} A_{ij}\frac{1}{(j-1)!}p^{j-1}e^{d_i p}\,U(p) \tag{4.15}$$

where $U(p)$ is the unit step function which ensures that the power estimate is non-negative. The coefficients of the partial fraction expansion, A_{ij}, are defined by Eq. (F.9). Even when these eigenvalues are distinct, they are typically in near-pairs for commonly used windows. Clustering the non-zero eigenvalues into groups with even multiplicities to use with Eq. (4.15) provides an accurate numerical approximation for practical computation of the density function.

Note in passing that for complex data, $\mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I$, with the assumption that the vectors $\mathbf{x}_R$ and $\mathbf{x}_I$ are independent and distributed N(0,$\mathbf{R}$) so that $\mathbf{x}$ is complex circular N(0,2$\mathbf{R}$), the moment-generating function for the power estimate is the square of Eq. (4.9):

$$M(t) = \prod_{i=1}^{D}(1 - t2\eta_i)^{-\nu_i}\,. \tag{4.16}$$

This ensures integer powers, so the partial fraction expansion always yields the exact pdf without grouping eigenvalues.

In practice, the cumulative distribution function is frequently desired; this is easily found by integration of Eq. (4.15), resulting in a sum of incomplete Gamma functions:

$$F_P(p) = g\sum_{i=1}^{D}\sum_{j=1}^{h_i} A_{ij}\frac{(-d_i)^{-j}}{(j-1)!}\gamma\,(j, -d_i p)\,U(p) \tag{4.17}$$

where the incomplete Gamma function is defined as [11]:

$$\gamma\,(a,x) = \int_0^x e^{-t}t^{a-1}\,dt. \tag{4.18}$$

4.3 The PDF's for Special Cases

In this section, Eq. (4.15) is developed for several special cases. In particular, note that the pdf of the power in a frequency band based on averaging modified periodograms requires computation of the eigenvalues of $\mathbf{R}\Upsilon$, where $\mathbf{R}$ is the covariance matrix of the Gaussian sequence, $\mathbf{x}$. For simplicity in the examples presented throughout most of this section, the signal $\mathbf{x}$ will be assumed to be white, $\mathbf{R} = \sigma^2\mathbf{I}$; I present one example of a colored signal. For white noise, the eigenvalues of $\mathbf{R}\Upsilon$, labeled η are simply σ^2 times the eigenvalues of Υ, labeled λ; that is, $\eta_i = \sigma^2\lambda_i$. I also discuss the distance between a Gaussian density and the derived theoretical density.

I will consider, as examples, two useful and interesting examples for $w[n]$: the rectangular window and the Hann window. For a rectangular data window, $w[n] = 1$ so the window matrix is the identity, $\Omega = \mathbf{I}$ and $U = 1$. The Hann window is defined as $w[n] = \sin^2\left(\pi\frac{n}{L}\right)$ and $U = \frac{3}{8}$.

First, the well known result for a single frequency bin based on a single data segment is developed for rectangular and Hann data windows. The more complicated case of averaging non-overlapping data segments with multiple frequency bins is then found. The case of 50% overlapping data segments and multiple bins is considered. I also include an example of spectral estimation for a colored sequence. The section concludes with an examination of the Kullback-Leibler distance as a measure of how different the actual densities are from a Gaussian with the same first and second moments.

4.3.1 One Frequency Bin, and a Single Data Segment

As a simple example, consider the case of a single frequency bin, ($k_1 = k_2 = k$), and a single data segment, ($K = 1$). The examples we present use even data segments, $L = 2q$, with no zero-padding of the segments, though the theory of the previous section includes these possibilities.

For the rectangular window, the non-zero eigenvalues of Υ are easily found to be

$$\lambda(\Upsilon) \;=\; \begin{cases} 1 & \text{for } k = 0, L/2 \\ \frac{1}{2},\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{4.19}$$

The resulting density function is Gamma:

$$f_P(p) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} p^{-\frac{1}{2}} \exp\left(-\frac{p}{2\sigma^2}\right)\ U(p) & \text{for } k = 0, \frac{L}{2} \\ \frac{1}{\sigma^2} \exp\left(-\frac{p}{\sigma^2}\right)\ U(p) & \text{otherwise} \end{cases} \tag{4.20}$$

Note that for $k = 0, \frac{L}{2}$, with $\sigma^2 = 1$, the distribution is a Chi-square distribution with one degree of freedom; for other frequency bins, the distribution is exponential, a well-known result.

Similarly, for the Hann window, the non-zero eigenvalues of Υ are

$$\lambda(\Upsilon) \;=\; \begin{cases} 1 & k = 0, L/2 \\ \frac{5}{12},\ \frac{7}{12} & k = \pm 1, L/2 \pm 1 \\ \frac{1}{2},\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{4.21}$$

The pdf is the same as for the rectangular data window except at $k = \pm 1, \frac{L}{2} \pm 1$, where the solution is in terms of the modified Bessel function:

$$f_P(p) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} p^{-\frac{1}{2}} \exp\left(-\frac{p}{2\sigma^2}\right)\ U(p) & \text{for } k = 0, \frac{L}{2} \\ \frac{1}{2\sigma^2}\sqrt{\frac{36}{35}} e^{-\frac{18}{35\sigma^2}p} I_0\left(\frac{3}{35\sigma^2}p\right)\ U(p) & \text{for } k = \pm 1, \frac{L}{2} \pm 1 \\ \frac{1}{\sigma^2} \exp\left(-\frac{p}{\sigma^2}\right)\ U(p) & \text{for otherwise} \end{cases} \tag{4.22}$$

Figure 4.1 displays the pdf's of the power in a single frequency bin, based on a single periodogram using a rectangular data window and a Hann window. The two data windows yield identical pdf's, except at $k = \pm 1, \frac{L}{2} \pm 1$ where the Hann window uses a Bessel function. For the plot, $\sigma^2 = 1$ is used. For all of the plots, the mean of P is 1; i.e., the estimate is an unbiased estimate of the signal power.

4.3.2 Averaging Non-Overlapping Data Segments

As described in Section 4.2, the case of non-overlapping data segments, Υ is formed as a Kronecker product. The eigenvalues of a Kronecker product are all the products of the eigenvalues of the two matrices. Applied to the problem at hand,

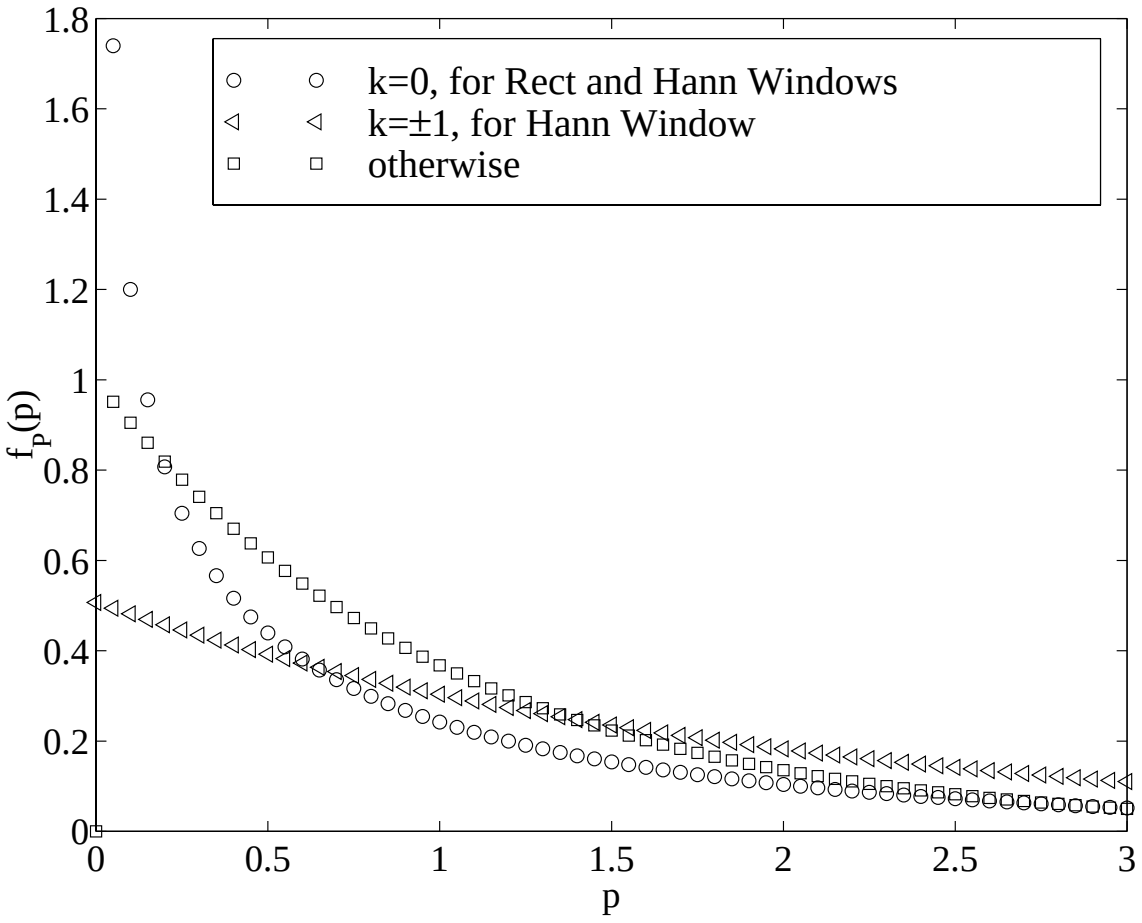


Figure 4.1: *The pdf of the power in a frequency bin estimated through a single periodogram estimate.*

averaging K segments increases the multiplicity of each eigenvalue for a single data segment by a factor of K.

For the rectangular data window with no overlap, each combination of b bins and K data segments can be characterized as having $D = 1$ distinct non-zero eigenvalue with multiplicity $\nu = 2Kb$ (as long as the frequency bins do not span across $k = 0$ or $\frac{L}{2}$ where the value of the eigenvalue is different). Summing several frequency bins yields the single non-zero eigenvalue $D = 1$ with multiplicity $\nu = 2b$; averaging over independent data segments increases the multiplicity of this eigenvalue to $\nu = 2Kb$. Further, the value of this eigenvalue is $\lambda = \frac{b}{\nu} = \frac{1}{2K}$. The pdf of a power

estimate based on b frequency bins, averaging K non-overlapping periodograms, with a rectangular data window is a Gamma distribution with $\alpha = Kb$ and $\beta = \frac{\sigma^2}{K}$:

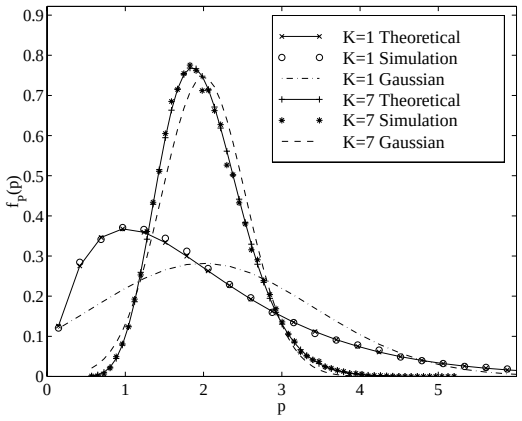
$$f_P(p) = \left(\frac{K}{\sigma^2}\right)^{Kb} \frac{1}{,\ (Kb)} p^{Kb-1} e^{-p\frac{K}{\sigma^2}}\ U(p). \tag{4.23}$$

The Hann window with no overlap, introduces some distinctions. Specifically, while the Kronecker product increases all the multiplicities by a factor of K, the window introduces correlation between the frequency bins such that summing multiple frequency bins yields more distinct non-zero eigenvalues. Again avoiding frequency bins $k = 0, \pm 1, \frac{L}{2}$ and $\frac{L}{2} \pm 1$, the number of distinct non-zero eigenvalues is $D = b$, each with multiplicity $\nu = 2K$; that is, h_i of Eq. (4.15) is K. Note that for a single bin, this is identical to the case of the rectangular data window.

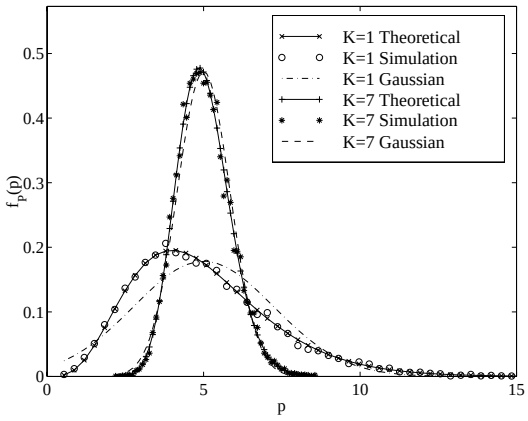
Some examples of the density function for rectangular and Hann windows are displayed in Fig. 4.2 for non-overlapping data segments. Here I have assumed the data stream, $\mathbf{x}$, is white Gaussian noise, with variance $\sigma^2 = 1$ and summed 2 and 5 frequency bins. The solid lines display the theoretical density, while the broken lines are the corresponding Gaussian densities with the same mean and variance and the circles and asterisks display the results of a Monte Carlo simulation in which Welch's method was applied to over 13 000 random data segments of length $L = 2^{17}$ to estimate the pdf of the power spectrum for each case. While not an exhaustive set, these plots demonstrate the general behavior of the pdf of the power. In every case, the Monte Carlo simulation corresponds well to the theoretical distribution. Further, the mode of the distribution occurs at a value of p considerably less than the mean—the tail on the left side of the distribution is much lower than that on the right side. As the number of frequency bins added together increases and/or as the number of data segments increases, the density becomes less skewed, and closer to a Gaussian.

4.3.3 PDF for 50% Overlapping Data Segments

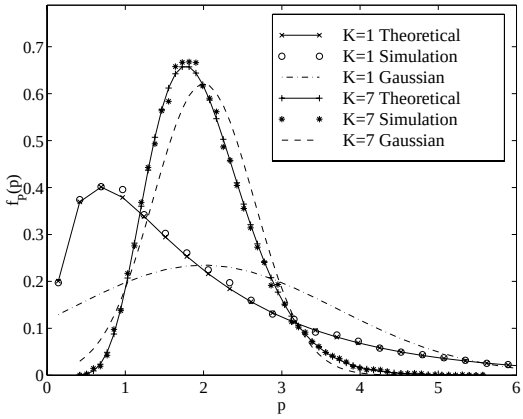
Having considered the cases of a single data segment and K non-overlapping segments, consider now the case of overlapping segments. In Section 4.2 is a description of how overlapping data segments can be readily analyzed by partitioning



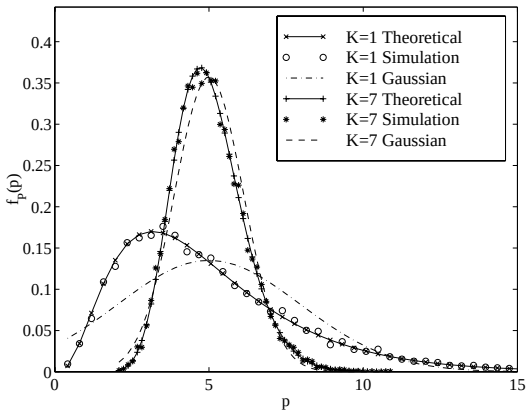
(a) Rectangular data window, 2 frequency bins



(b) Rectangular data window, 5 frequency bins



(c) Hann data window, 2 frequency bins



(d) Hann data window, 5 frequency bins

Figure 4.2: *Probability density functions for non-overlapping data segments based on the sum of 2 and 5 frequency bins (adjacent frequency bins were chosen well away from $k = 0, \frac{L}{2}$). The solid lines plot the theoretical pdf's based on 1 and 7 data segments, the dashed lines indicate the corresponding Gaussian densities with identical first and second moments, and the circles and asterisks are the results of Monte Carlo simulations.*

the matrix for each periodogram and summing the corresponding submatrices. With this method, it is straightforward to construct such a matrix for any overlap and numerically evaluate the eigenvalues.

Examples of the densities based on 50% overlapping data segments, summing 2 and 5 frequency bins, are displayed in Fig. 4.3. As before, the data stream, $\mathbf{x}$, is white Gaussian noise, with variance $\sigma^2 = 1$. The dashed lines, indicating the corresponding Gaussian density, are quite different from the theoretical pdf. The general form of the pdf of the power is shown to include considerable skew, with the right tail being much higher. The plot also shows the agreement of Monte Carlo simulation results with the theoretical density function as circles and asterisks.

4.3.4 A Correlated Data Segment

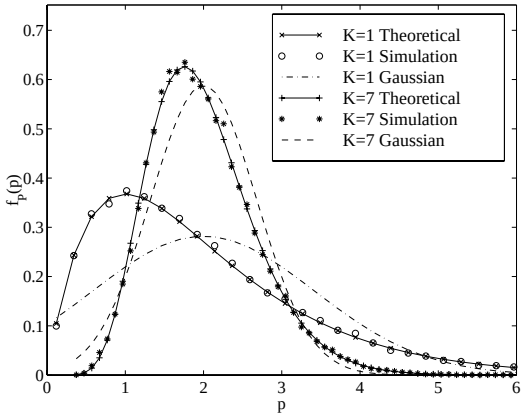
While the previous examples described the pdf of power estimates from Gaussian, white data sequences, the derivation allowed a more general, correlated sequence with arbitrary autocorrelation matrix, $\mathbf{R}$. Here I consider an example of this more general case. One constraint on the autocorrelation matrix, due to the fact that the data sequence is assumed stationary, is that $\mathbf{R}$ is Toeplitz. A common type of sequence results from a first-order Markov process with covariance function [55]

$$r(n) = \rho^{|n|}, \qquad |\rho| < 1, \forall n. \tag{4.24}$$

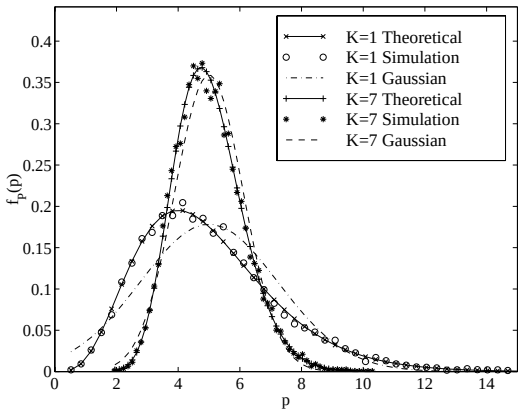
For an L element data segment, $\mathbf{x}$, this results in a symmetric autocorrelation matrix with a value for the qth subdiagonal as ρ^q.

Equation (4.15) provides the probability density function of the power in a frequency bin based on the eigenvalues of $\mathbf{R}\Upsilon$. The basic form of the pdf will be the same for correlated sequences as it was for white sequences; the only difference will be that sequence correlation may modify the eigenvalues, changing the details of the pdf.

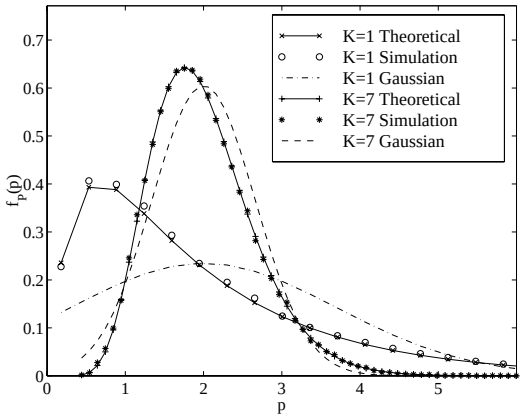
In Fig. 4.4 the pdf of the estimated power spectrum is plotted. The x-axis is the normalized frequency, plotted from 0 to π (the spectrum is symmetric). The pdf is displayed at four frequencies, with the mean of each estimate (indicated by



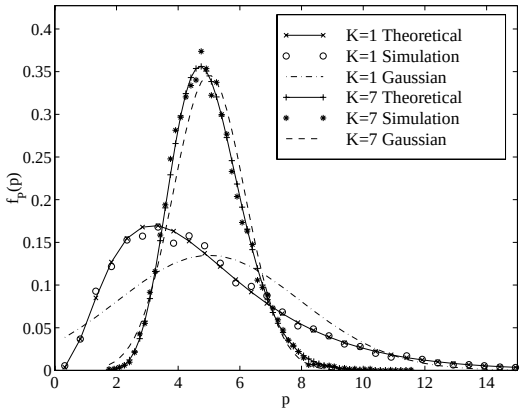
(a) Rectangular data window, 2 frequency bins



(b) Rectangular data window, 5 frequency bins



(c) Hann data window, 2 frequency bins



(d) Hann data window, 5 frequency bins

Figure 4.3: *Probability density functions for 50% overlapping data segments based on the sum of 2 and 5 frequency bins (adjacent frequency bins were chosen well away from $k = 0, \frac{L}{2}$). The solid lines indicate the theoretical pdf's, the dashed lines indicate the corresponding Gaussian densities, and the circles and asterisks are the results of Monte Carlo simulations.*

small circles) corresponding well with the assumed spectrum of a first-order Markov process with correlation coefficient, $\rho = 0.3$. For this plot, each pdf was based on a single frequency bin estimated with $K = 5$ data segments, each of length $L = 64$, with 50% overlap, (the data window choice makes no effect for a single frequency bin).

4.3.5 Relative Entropy

The plots of probability density functions presented to this point suggest a wide range of behavior depending on the particular parameters of how many frequency bins, data segments, amount of overlap and type of data window used. The pdf asymptotically approaches a Gaussian in the limit as more frequency bins and/or data segments are used.

The "distance" between two probability density functions, f and g, can be quantified with the concept of relative entropy, or Kullback-Leibler distance [56]:

$$D\left(f\|g\right) = \int f(x) \log \frac{f(x)}{g(x)}\, dx. \tag{4.25}$$

Because the asymptotic behavior of the pdf is Gaussian, and because it is not uncommon to compute the first two moments of the density and assume a Gaussian, it is significant to consider the distance between the theoretical pdf and a Gaussian with the same mean and variance.

For the special case where $\mathbf{R}\Upsilon$ has only a single non-zero eigenvalue, η, with multiplicity ν, the resulting pdf is Gamma with parameters $\alpha = \frac{\nu}{2}$ and $\beta = \eta$. This case corresponds to an uncorrelated data stream for which the power is estimated using a rectangular data window and non-overlapping data; this is a useful case because the relative entropy between Gamma and Gaussian densities with mean $\alpha\beta$ and variance $\alpha\beta^2$ can be written in closed form:

$$D\left(f\|g\right) \;=\; \log\left(\frac{\sqrt{2\pi\alpha}}{,\left(\alpha\right)}\right) - \alpha + \frac{1}{2} - \left(\alpha - 1\right)\gamma + \left(\alpha - 1\right)\sum_{j=1}^{\alpha-1}\frac{1}{j}. \tag{4.26}$$

In more general cases, the relative entropy can be computed numerically. For several situations, $D(f\|g)$ is plotted in Fig. 4.5. Data points are calculated for

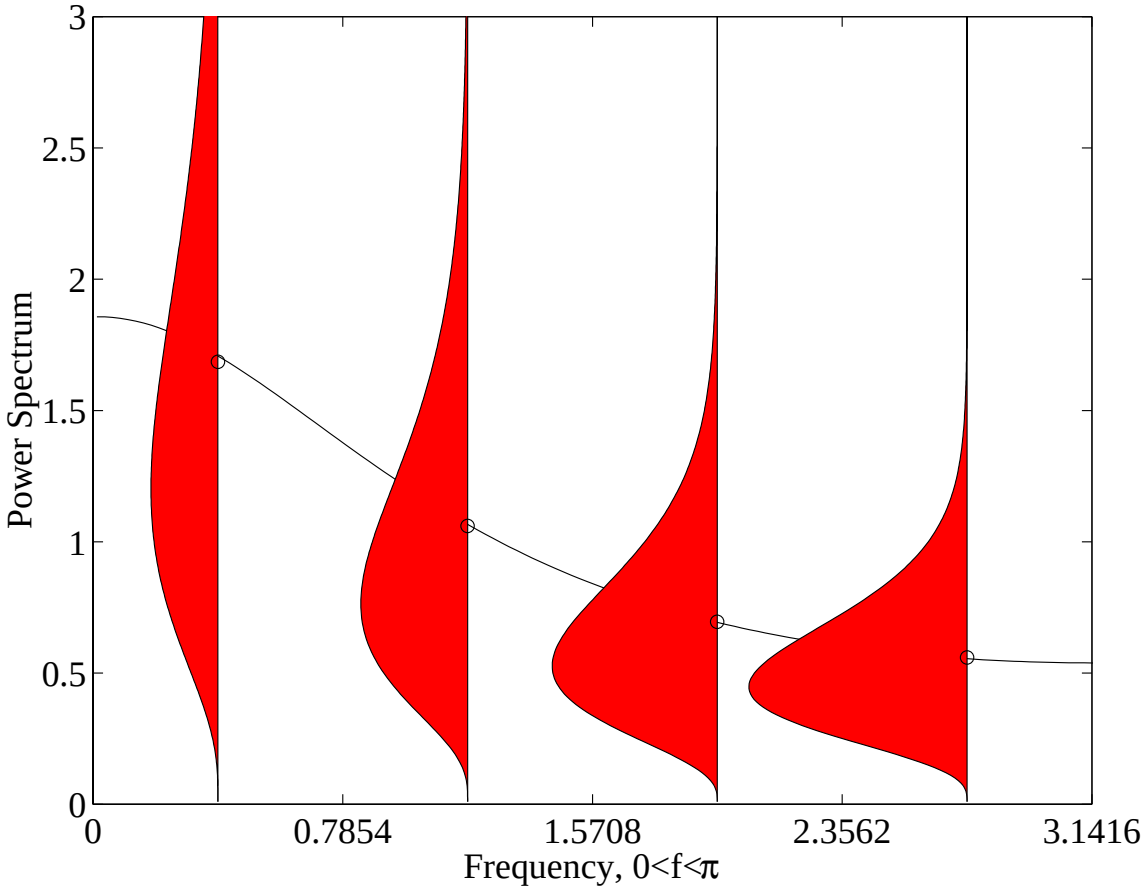


Figure 4.4: *For a correlated data sequence, the power spectrum (the Fourier transform of the autocorrelation function) varies with frequency. The solid line represents the spectrum of a first-order Markov process with correlation coefficient,* $\rho = 0.3$. *The filled figures describe the pdf of the estimated spectrum at the four selected frequencies. Note that each pdf is skewed toward low power, while the mean of the pdf (indicated by the small circles) correspond to the spectrum established by the Markov process.*

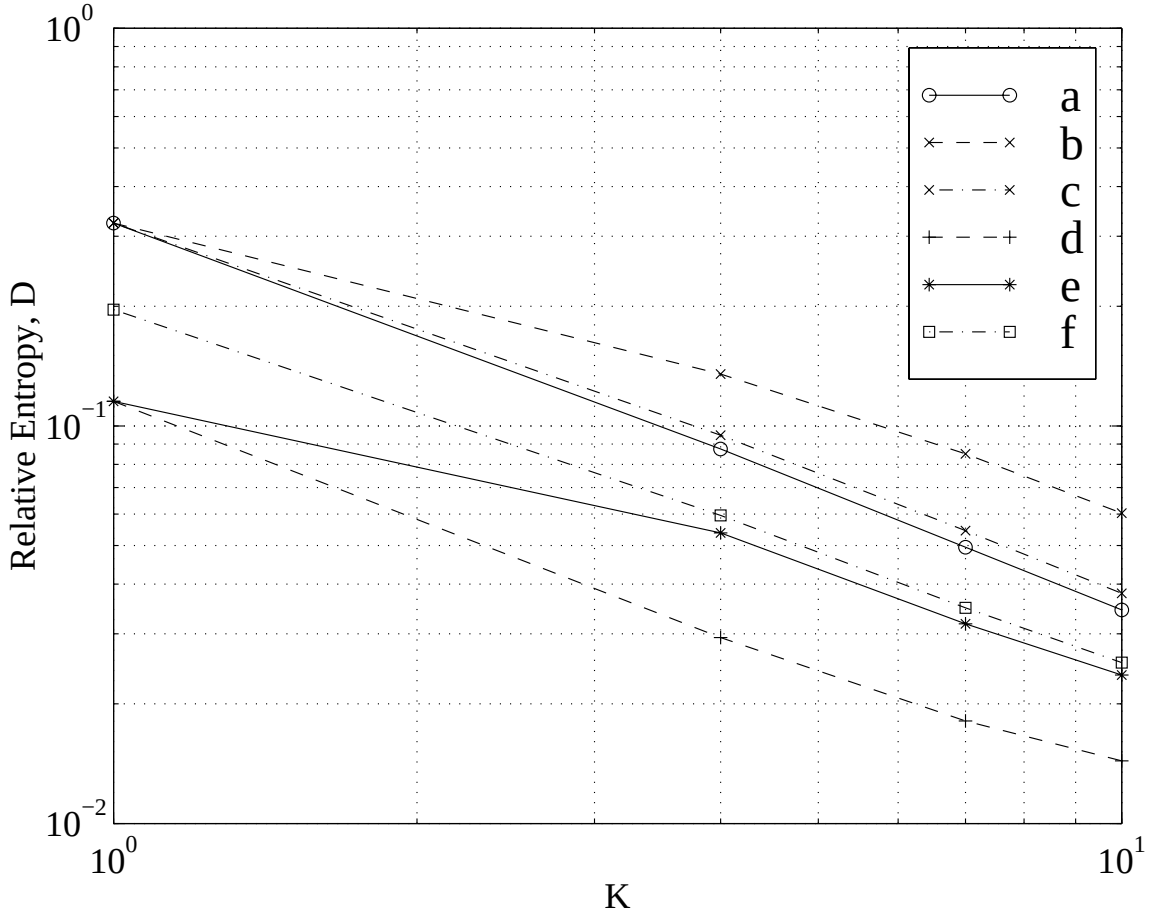


Figure 4.5: *The relative entropy, or Kullback-Leibler distance, between a Gamma distribution and a Gaussian distribution with the same mean and variance as a function of K, the number of data segments used in the periodogram estimate. (a) 1 bin, non-overlapping segments, rect window, (b) 1 bin, overlapping segments, rect window, (c) 1 bin, overlapping segments, Hann window, (d) 3 bins, non-overlapping segments, rect window, (e) 3 bins, overlapping segments, rect window, (f) 3 bins, non-overlapping segments, Hann window.*

$K = 1, 4, 7$ and 10 data segments (with linear interpolation between the data points). In every case there is a nearly exponential decay as the pdf's asymptotically approach Gaussian as the number of data segments and/or frequency bins increases.

4.4 Conclusions

Welch's modified periodogram averaging has served as a simple, common technique for spectral estimation for three decades. However, the statistical structure of the estimate has never been fully reported; in this chapter I present the probability density function of a spectral estimate of a frequency band based on Welch's method with Gaussian input. The probability density function of an estimate will be useful in many situations. For example, maximum-likelihood estimations approximating the distribution as Gaussian overestimate the mode because of the skewed structure of the correct pdf.

The critical component of the analysis requires the eigenvalues of a matrix, $\mathbf{R}\Upsilon$, which can be computed numerically. Several examples of the pdf have been plotted for illustration purposes. Simulations were presented in which the sample pdf's of the spectral estimates were compared to the theoretical functions derived here. Finally, the Kullback-Leibler distance between the correct pdf and the corresponding Gaussian density was plotted for several representative cases, displaying the monotonically converging behavior of the pdf to a Gaussian as the frequency bins and/or the number of data segments is increased.

Chapter 5

The Probability Distribution of NSCAT Measurements

Estimation of geophysical parameters based on their modulation of NRCS requires accurate measurements of the power scattered from the surface. Inherent in such measurements is considerable uncertainty. While the signal transmitted from the scatterometer is deterministic, the received signal is random based on the scattering from the random surface. The scattered power is accurately described as a white, Gaussian signal. Resolution of the antenna footprint into ocean cells is typically done by Doppler filtering. Further, an estimate of the noise in the receiver must be subtracted from the noisy measurement. These complications modify the statistics of the backscatter measurements. For lack of a better choice, and with only a sample mean and variance of the measurements, the pdfs traditionally have been assumed Gaussian.

However, the pdf of the sum of frequency bins, based on Welch's method with K overlapping data segments, and an arbitrary data window, is explained in the previous chapter, and found to be clearly non-Gaussian [54]. In this chapter I apply the theory to the specific case of NSCAT processing.

Wind retrieval with NSCAT involves Doppler filtering to resolve ocean cells; the signal-plus-noise measurement is made as the sum of a range of frequency bins of the estimated power spectrum [29]. Welch's method for periodogram estimation [51] is used with 50% overlapping data segments and a Hann window to minimize spectral leakage [42]. The signal power for a cell is then computed by subtracting a noise-only power estimate from the signal-plus-noise estimate. This signal power, used in the wind estimation, has been assumed to have a Gaussian distribution. However,

Welch's periodogram estimate is not Gaussian [54]. In this Chapter, the probability density function of NSCAT power estimates is found, and the effect on wind retrieval is considered. The development is made with sufficient generality to readily extend to other applications of periodogram estimation.

5.1 Introduction

The pdfs of a noise-only measurement and of a signal-plus-noise measurement, appropriate for NSCAT processing, only require simple scale changes from basic pdfs; the pdf of the signal-only measurement is then found as the correlation of these two pdfs. In this chapter I describe the scale changes necessary based on the number of overlapping data segments and frequency bins used (these vary throughout the data set). I compare these density function with Gaussian densities of the same mean and variance. I also report compass simulation results which describe the impact of using the incorrect pdf on wind retrieval (see also [57]).

5.2 Conditional PDF of NSCAT Measurements

For NSCAT, two power estimates are made, a signal-plus-noise measurement and a noise-only measurement. The power in the signal is estimated as the difference between these two power estimates: $P = P_1 - P_2$. Because P_1 and P_2 are independent random variables, distributed as described in Chapter 4, the pdf of the signal-only power is the convolution of the first pdf with that of the second pdf with a negative argument.

NSCAT employs both time and frequency filtering to resolve a wind cell. Depending on the cell location within the swath, the estimate of the signal-plus-noise, P_1, is based on either 2, 3 or 7 overlapping data segments ($L = 256$) and from 5 to 20 frequency bins, and is improved by averaging 25 pulses [42].

A single pulse would result in a pdf as given above by Eq. (4.15) and repeated here for reference (with the subscript changed to 'basic' because this basic pdf will be scaled for given signal power):

Table 5.1: *Eigenvalues of the basic signal-plus-noise pdfs for a single pulse with typical NSCAT parameters: $K = 2$, 3 and 7, $b = 5$, $L = 256$, $r = 0.5$, and a Hann window.*

$K = 2$		$K = 3$		$K = 7$	
Eigenvalues	Mult	Eigenvalues	Mult	Eigenvalues	Mult
0.5932	2	0.3968	2	0.1688	14
0.5860	2	0.3935	2	0.1216	14
0.4252	2	0.3901	2	0.0877	4
0.4052	2	0.2890	2	0.0727	2
0.2578	2	0.2795	2	0.0608	2
0.1495	2	0.2721	2	0.0488	2
0.0613	2	0.1916	2	0.0374	2
0.0185	2	0.1378	2	0.0226	4
0.0030	2	0.0849	2	0.0078	6
0.0002	2	0.0413	2	0.0003	20
		0.0173	2		
		0.0051	2		
		0.0009	2		
		0.0001	4		

$$f_{\text{basic}}(p) \;=\; g \sum_{i=1}^{D} \sum_{j=1}^{h_i} A_{ij} \frac{1}{(j-1)!} p^{j-1} e^{d_i p}\, U(p). \tag{5.1}$$

The eigenvalues and their multiplicities, for a single pulse are based on $K = 2$, 3 or 7 data segments of length $L = 256$ and $b = 5$ to 20 frequency bins, $r = 0.5$ for 50% overlap and NSCAT applies a Hann data window to minimize spectral leakage. Equation (5.1) then provides the pdf of the generic signal-plus-noise, with mean equal to b. For example, Table 5.1 lists the eigenvalues corresponding to the basic signal-plus-noise for the case of 5 frequency bins and a single scatterometer pulse. The corresponding pdfs are plotted in Fig. 5.1, revealing the familiar shape seen in Chapter 4 and each case having a mean equal to the number of frequency bins, 5.

Because NSCAT averages 25 signal-plus-noise pulses, it is useful to define the normalized pdf based on the self-convolution of 25 pulses (i.e., the pulses are assumed i.i.d. so that the average power can be computed by convolution) and normalized by b frequency bins, so the normalized distribution has a mean of 1:

$$f_{\text{norm}}(p) = 25b f_{\text{basic}}^{(25)}\left(\frac{p}{25b}\right), \tag{5.2}$$

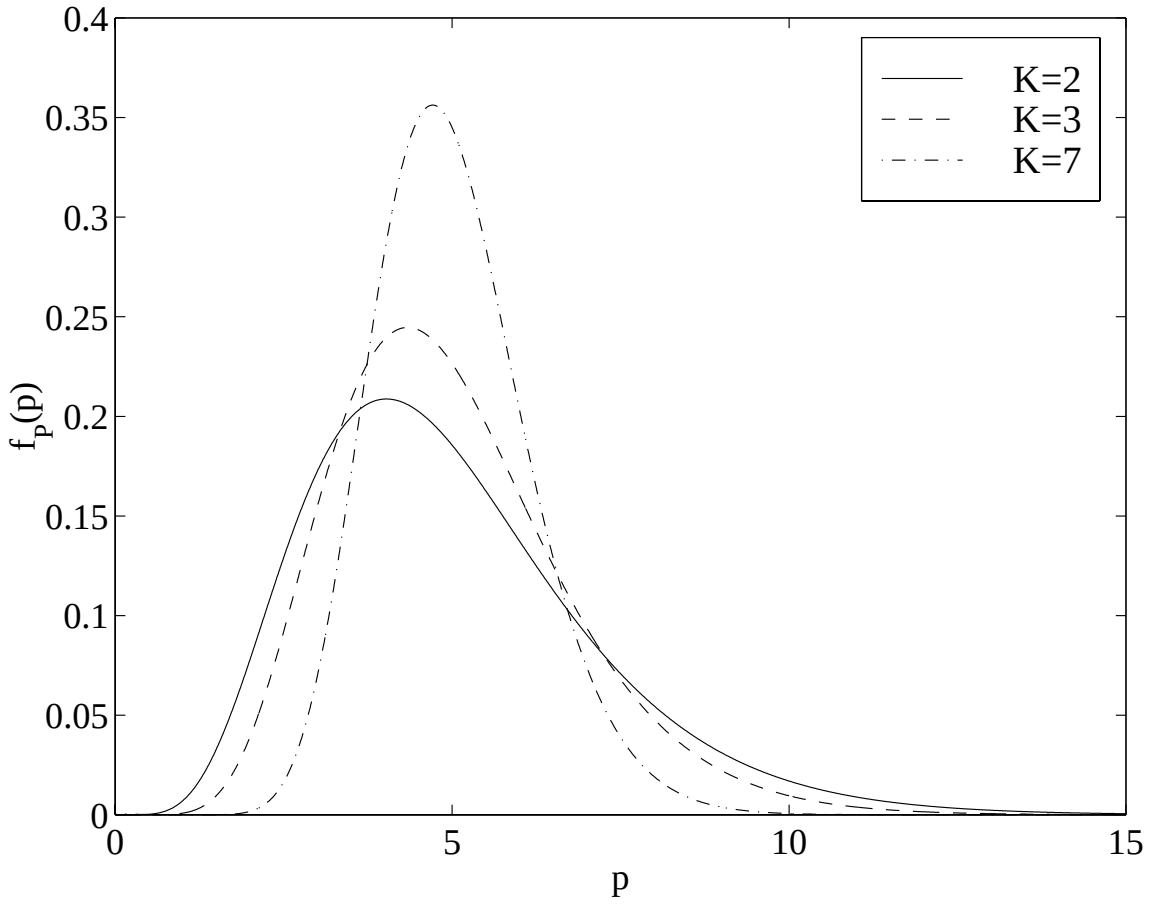


Figure 5.1: *Generic signal-plus-noise pdfs for a single pulse with typical NSCAT parameters*

where the superscript in parenthesis indicates self-convolution. The pdf of the signal-plus-noise, given signal and noise powers, P_S and P_N, is

$$f_{\mathrm{SN}}(p|P_S, P_N) \;=\; \frac{25b}{P_S + P_N} f^{(25)}_{\mathrm{basic}} \left(\frac{p}{25b} \left(P_S + P_N \right) \right). \tag{5.3}$$

The noise power estimate, P_2, averages 4 'pulses' and sums more than 200 frequency bins. For this estimate, the scatterometer transmits nothing, but takes measurements according to the same timing as during the signal-plus-noise measurement.

The pdf of the signal-only power, $f_S(p)$, is then the convolution of $f_{SN}(p)$ with the pdf of the noise estimate with a negative argument: $f_S(p) = f_{SN}(p) *$

$f_N(-p)$. Because the noise estimate sums 200 frequency bins and 4 pulses, the noise is essentially Gaussian with a variance which becomes negligibly small. The pdf $f_N(p)$ can be well approximated with a delta function at the noise power. The signal-only pdf is then simply

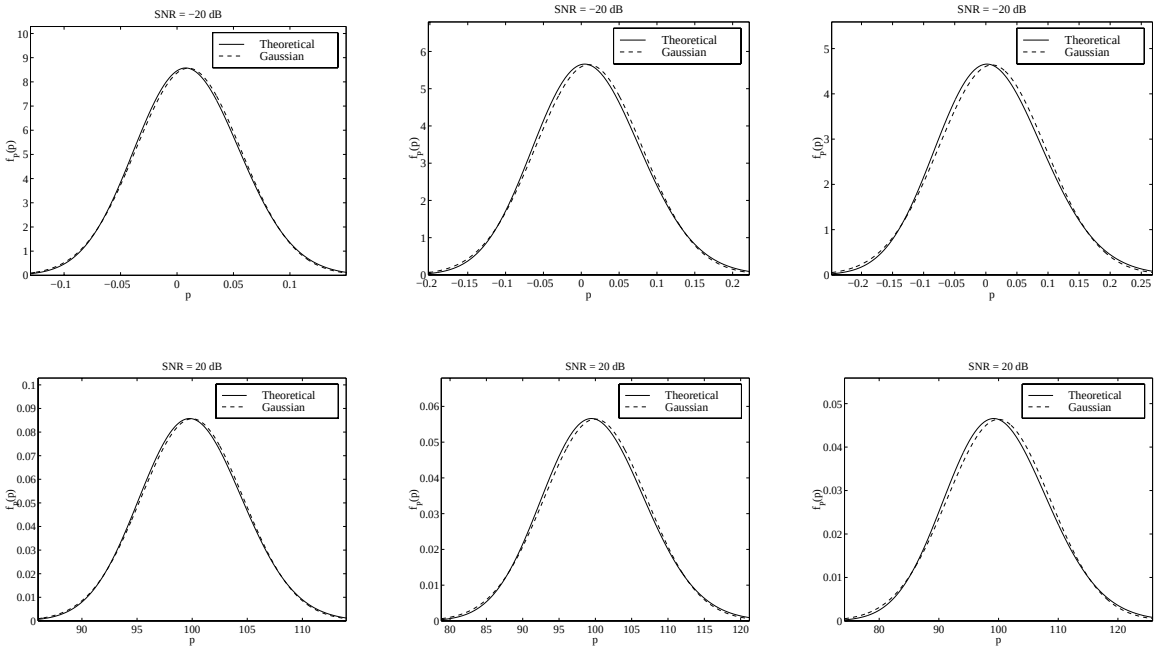
$$f_S(p|P_S,P_N) \approx \frac{25b}{P_S+P_N} f^{(25)}_{\text{basic}}\left(\frac{p}{25b}(P_S+P_N)+P_N\right) \quad (5.4)$$

$$= \frac{25b}{P_N(1+SNR)} f^{(25)}_{\text{basic}}\left[P_N\left(\frac{p}{25b}(1+SNR)+1\right)\right]. \quad (5.5)$$

Figure 5.2 displays several examples of the density functions of power estimates. The near, mid and far swath cells use 7, 3 and 2 overlapping data segments, respectively, and 5 frequency bins. Note that the SNR simply scales the power estimate for each case. The dashed line indicates a Gaussian density with the same mean and variance. It is clear that the actual distribution is skewed toward low σ^o values. For near swath cells, the true distribution is more Gaussian.

Two useful results related to the density function are the normalized bias and the probability of negative measurements. Defining the normalized bias as the difference between the mode and the mean, normalized by the mean of the distribution, a simple measure is obtained of the distance between a Gaussian and the correct density. Also of interest is the probability of a negative power estimate. This can be computed as the integral of the pdf from $-\infty$ to 0, or equivalently as the cumulative distribution function evaluated at zero $F(0)$. These measures are displayed in Fig. 5.3 as functions of the SNR for the three representative cell cases.

The concern is that because the actual density is skewed, the wind estimation may be biased. A compass simulation shows that this bias is small, though non-zero. Using a particular measurement geometry and noise parameters (taken from an NSCAT L1.7 file) simulated backscatter measurements for a given wind vector cell are used to estimate the wind in a traditional compass simulation. The simulated measurements were then biased, according to Fig. 5.3, and the wind is estimated with this set of biased measurements. Note that a more precise method would be to use the correct pdf, rather than simply a shifted Gaussian distribution; but to obtain a reasonably accurate result with minimal code modification for this



(d) Near Swath (e) Mid Swath (f) Far Swath

Figure 5.2: *Some sample pdfs for the signal-only power estimate based on NSCAT processing. Because there is less averaging for far swath cells (fewer data segments and fewer frequency bins used), the skew increases. For the different parts of the swath, the SNR scales the fundamental pdf.*

preliminary study, this approach was chosen. Figure 5.4 displays the results of this compass simulation for a cell in the far swath (2 overlapping data segments and 5 frequency bins). The plots show the difference between the wind estimates made without the correcting bias and the wind estimates with the correcting bias for both speed and direction. As the simulated wind speed increases, the difference between the two estimates increases, while the direction difference is minimal.

5.3 Impact of Wind Distribution

In the previous section, the conditional probability density function of the signal power estimate was derived for NSCAT processing as Eq. (5.5). In terms of wind retrieval, this is the conditional pdf which, by adjusting the wind estimate, results in the MLE wind (see Appendix D).

The pdf of the measurements, then, is the integral of this conditional pdf multiplied by the prior pdf of the signal power:

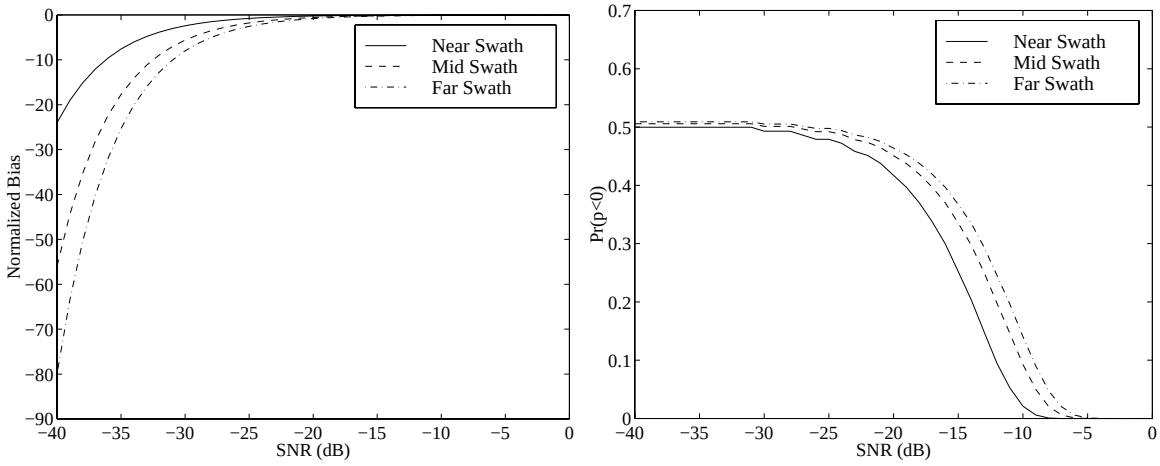
$$f_S(p) = \int_{-\infty}^{\infty} f_S(p|P_S) f(P_S)\, dP_S. \tag{5.6}$$

The distribution of the signal power is not known. It will depend on the distribution of the wind and on the model function relating the wind to the backscatter. In this section I develop the backscatter measurements, z, based on prior distributions of the wind and the empirical model function.

The radar equation, Eq. (1.1), states that the backscatter is linearly related to the power measurement, $\sigma^o = XP_S$. Transforming the pdf of the measurements to backscatter space, and assuming the noise power is deterministic, yields:

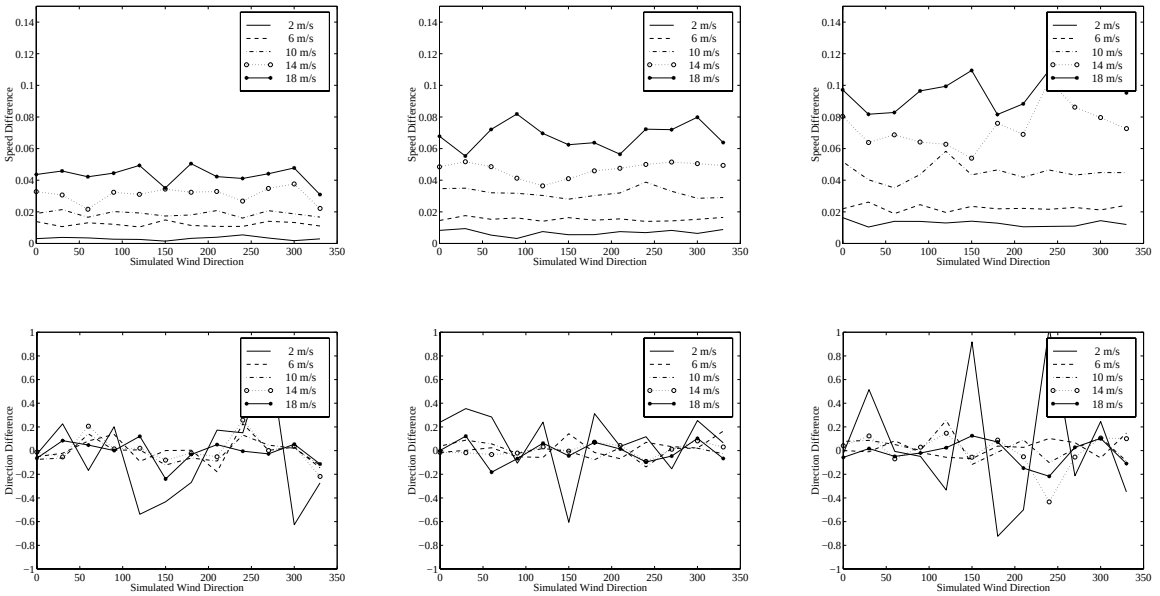
$$f_Z(z) \;=\; \int_{-\infty}^{\infty} \frac{25b}{X\,(P_S + P_N)} f_{\text{basic}}^{(25)} \left(\frac{z}{25b}\,(P_S + P_N) + P_N \right) f(P_S)\, dP_S. \tag{5.7}$$

The pdf of the signal power is related simply to the pdf of the backscatter (as a function of the wind). For Fig. 5.5 I established a set of wind vectors with realistic statistics (plotted on the left), with a Rayleigh distribution for the wind speed and a uniform distribution for the direction. Using a typical NSCAT measurement geometry for a single beam (at near-swath) I computed the resulting backscatter according



(a) (b)

Figure 5.3: *Left: The normalized bias, defined as the difference in the modes of the theoretical and Gaussian densities, normalized by the mean, depends on the SNR and has similar behavior for all cells. Right: The probability of negative power estimates as a function of the signal to noise ratio.*



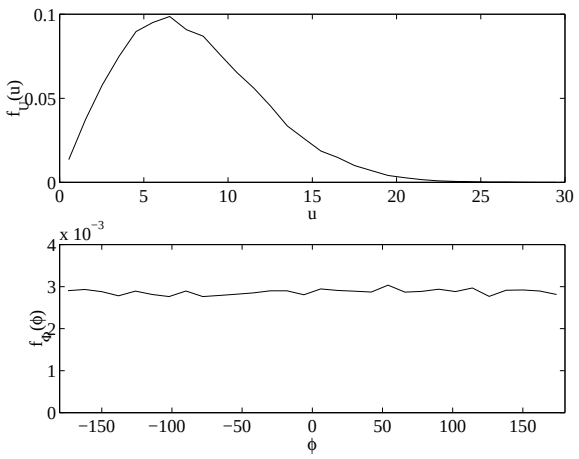
(d) Near Swath (e) Mid Swath (f) Far Swath

Figure 5.4: *Impact on wind retrieval with bias error in the pdf. For the three parts of the swath, results from compass simulations are displayed showing the difference between the wind retrieved using traditional methods and that retrieved assuming a biased Gaussian distribution for the scatterometer measurements. The latter more closely approximates the correct distribution.*

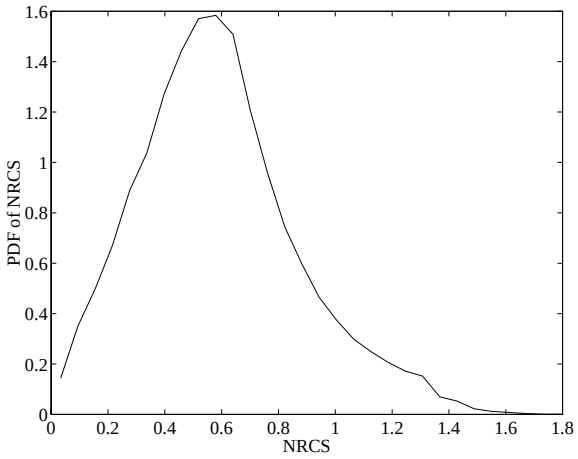
to the empirical model function. The plots on the right display the sample pdf of the backscatter. The evidence is growing for a limit on the wind speed, below which the NRCS and wind speed are only slightly correlated; the bottom plots represent this case where winds below a threshold of 3 m/s do not contribute to the backscatter.

5.4 Conclusion

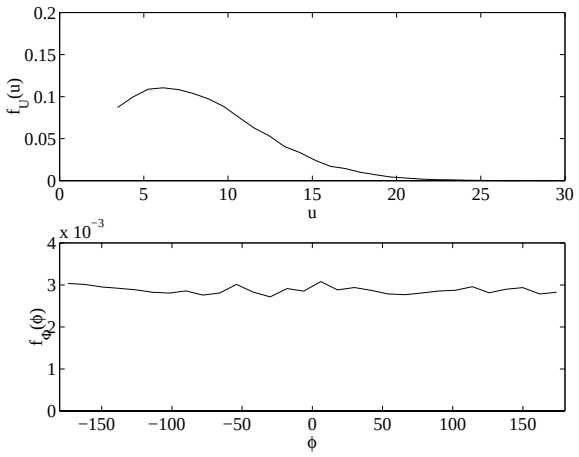
The probability structure of NSCAT signal estimates is not biased, but shows considerable skew. Initial observations suggest that this has negligible impact on the estimated wind direction, though the wind speed bias should be investigated further. The skew of the probability density function of the power measurements causes current wind retrieval techniques (which assume measurements of Gaussian random variables) to be biased high. This effect is ameliorated to some extent by models like NSCAT1, which are tuned to the retrieved wind, as well as the considerable averaging resulting in nearly Gaussian densities (see Chapter 4).



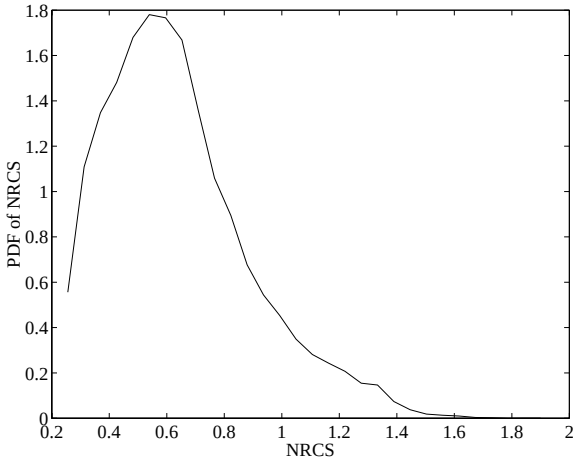
(a) PDF of Simulated Wind ($U0 = 0$)



(b) PDF of Backscatter ($U0 = 0$)



(c) PDF of Simulated Wind ($U0 = 3$)



(d) PDF of Backscatter ($U0 = 3$)

Figure 5.5: *Sample estimate of the pdf of the model function backscatter.*

Chapter 6

Conclusion

Oceanic microwave remote sensing is a valuable tool for monitoring critical climatological and hydrological processes. Active radar systems provide measurements of the EM power scattered from the surface, which serve as indirect measurements of significant geophysical parameters. While such systems have been in use for decades, significant gaps in our understanding of the processes involved still persist. In this dissertation, I have clarified and quantified uncertainties in oceanic microwave remote sensing.

The general diagram of the translation of the geophysical parameters to their estimation via microwave remote sensing is displayed in Fig. 1.1. The geophysical parameters (e.g., wind, rain, swell) develop ocean waves with a particular surface spectrum. The NRCS is determined by the surface spectrum, the radar observation parameters and the ocean dielectric characteristics. The radar measures the scattered power, which is linearly related to the NRCS. From the measurements, the geophysical parameters such as wind can be estimated. In this research I explored three significant and previously unstudied sources of error and uncertainty in the radar measurements.

6.1 Contributions

Several techniques for evaluating the theoretical backscatter from a random surface have been developed over the last several decades, and continue to find a significant place in current literature. I applied the theory of physical optics scattering, also called the Kirchoff approximation, to nadir scattering from ocean-like

spectra. I performed a sensitivity analysis of the backscatter and identified a particularly interesting change in the scattering characteristics depending on whether the radar footprint was larger or smaller than a computable threshold. I derived the footprint size threshold in terms of the significant parameters of the spectrum and radar system. This work has been presented at several international technical conferences, and will be submitted for publication in *The Journal of Geophysical Research.*

Because the precise relationship between all of the geophysical parameters and the NRCS is not understood, empirical estimates are commonly used to provide reasonable estimates of the vector wind from the backscatter measurements. However, inherent in all empirical estimates of the geophysical model function is uncertainty caused by unmodelled parameters. While an intense search for the complete geophysical model function continues in the scientific community, current applications require a quantitative understanding of the variability of empirical model functions. My research provides such a quantitative analysis. I developed a simple model incorporating model variability as multiplicative noise in the true NRCS, distinguishing it from the NRCS predicted by the model. From the noise model, I developed an algorithm to estimate, from scatterometer measurements alone, the variability of the backscatter for given wind conditions. Separating the data into distinct bins I was able to estimate the variability as a function of incidence angle, speed, latitude, and time. I performed the data analysis of the nine month NSCAT mission to quantify the variability. I also investigated the impact of this variability on wind estimates and found that, because empirical model functions provide good estimates of the mean NRCS for wind conditions, the average vector wind estimates have only a very small bias. But the covariance of the wind estimates, as quantified by the Cramer-Rao lower bound, is substantially increased by the variability of the model functions. This work has also been presented at several international conferences and will be submitted for journal publication in *IEEE Transactions on Geoscience and Remote Sensing.*

The measurements made by the radar provide substantial signal processing. The statistical changes to the random signal were previously misunderstood. I derived the probability density function of a power estimate for a correlated Gaussian

signal, based on Welch's modified periodogram averaging technique, and find it to be skewed from the Gaussian pdf previously assumed. Beyond the very general derivation of the pdf, I derived the correct pdf for the NSCAT signal processing design. Including the estimate of the signal power from the difference of a signal-plus-noise measurement and a noise-only measurement, I derived the pdf of NSCAT power measurements based on the SNR, the number of frequency bins summed for the surface resolution element, and the number of overlapping data segments used (each of these change across the swath). I demonstrated that because NSCAT includes considerable averaging and because empirical model functions used to estimate the wind have been tuned to the assumed Gaussian distribution, there is only a small bias in the retrieved wind vector. This work has been presented at several international technical conferences. A manuscript describing the general development of the pdf of spectral estimates has been accepted for publication in *IEEE Transactions on Signal Processing* [54].

Here, I itemize my specific contributions to understanding and quantifying sources of error and uncertainty in the radar measurements:

- The Finite Radar Footprint
  - Derivation of the traditional asymptotic expansion of the physical optics scattering integral for nadir backscatter.
  - Extension of the multiple-convolution of the spectrum method for computing the asymptotic backscatter to two dimensions.
    * Arbitrary incidence angle
    * Arbitrary spectral form
  - Derivation of the correlation function from the general class of spectra with:
    * a hard low-wavenumber limit
    * a power-law decay in wavenumber
    * an arbitrary angular dependence.

- Extension of the correlation function to composite surfaces characterized as a sum of spectra with differing parameters
- Closed form solution of the asymptotic backscatter at nadir incidence, using the traditional asymptotics method, allowing a sensitivity analysis.
    * Several representative plots are produced demonstrating the backscatter behavior
    * The backscatter decreases with surface roughness
    * The decay experiences distinct rates:
        · Very low roughnesses experience specular scattering, the backscatter is constant
        · Intermediate asymptotic roughnesses achieve a moderate, constant rate of decay
        · Final asymptotic roughnesses converge to a large, constant rate of decay
    * The transition point between the moderate asymptotic rate of decay and the final asymptotic rate is derived.
        · The transition point is at higher roughness values with large p values
        · The transition point is at higher roughness values with more unidirectional (less isotropic) surfaces
        · The transition point is at higher roughness values with smaller k_0 values.
        · The transition point is at higher roughness values with smaller footprint sizes.
- The physical mechanism of the transition between moderate and final asymptotic rates is described
    * Directional surfaces with very low roughnesses do not have sufficient heights in the cross-wave directions to be recognizable to the radar

using a finite footprint, so the surface looks unidirectional (which has a shallow rate of decay). Very rough, directional surfaces have sufficient height to scatter the radar signal incoherently in both the along-wave and cross-wave directions.

    * In terms of the correlation function, if the footprint size contains enough of the correlation function, then the final asymptotic backscatter slope is achieved, but if the footprint is small (relative to the width of the correlation function) then the long distance correlations will not be observed in the backscatter—the backscatter will have a smaller rate of decay.

    * Highly directional surfaces stretch the correlation function, pulling portions of it outside a finite footprint.

  - The multiple-convolutions method was applied to a more realistic spectrum (smooth low-wavenumber transition) to show that the two modes of asymptotic decay are revealed with similar footprint-size dependence

  - The electromagnetic bias identified in altimeter data was shown to be affected by the the size of the footprint. A theoretical explanation qualitatively showed that smaller footprints should have less EM bias than larger footprints.

- Uncertainties in the Wind-Backscatter Relationship

  - Development of a multiplicative model for the backscatter variability under given wind conditions

  - Arguments leading to a multiplicative model.

  - Development of an iterative technique to estimate the model function variability based solely on the model function and scatterometer data.

    * Simulations showing the convergence of the technique

    * Determination of a quadratic correction factor to remove the bias

- Analysis of the NSCAT data set to determine the model function variability, K_{PM}, estimated as a function of four parameters:
  - Retrieved wind speed in 2 m/s bins. Variability decreases with increasing wind speed.
  - Incidence angle in 5 degree bins. Variability increases with increasing incidence angle.
  - Time in approximately 3 day bins (40 revolutions) sampled approximately monthly. Estimates of K_{PM} show minimal variation with temporal variations.
  - Latitude in 30 degree bins. Variability increases with increasing magnitude of latitude (i.e., the variability is smaller at the equator than at the poles).
- Identification of the impact of K_{PM} on wind estimation
  - Simulations were performed to estimate the wind from typical scatterometer measurements with and without model function variability
    - Any variability of the model function biases (slightly) the retrieved wind speed and direction estimated using a log-likelihood estimator, as commonly done for scatterometer data
    - Precise knowledge of the value of K_{PM} does not improve the wind estimate bias.
  - Compass simulations were performed which show that changing the value of K_{PM} does not dramatically change the locations of the local peaks of the log-likelihood function, but the shape and relative magnitudes of the peaks are modified. Thus, while K_{PM} does not predictably change the bias of the possible solutions (all the local peaks), it can change the ranking of the peaks and their curvatures (that is, the variance of the ambiguities).
  - Weighted-least squares estimation was considered as a means of reducing the sensitivity of the wind estimation to model function variability.

    * The variance of the wind estimates was demonstrated to be dramatically affected by model function variability.
        · The Cramer-Rao lower bound for the error covariance matrix was used to compute the bounds on estimated wind speeds and directions, based on the NSCAT1 model function, using various K_{PM} values.
        · Even moderate values of K_{PM} produce substantial increases in the variance bounds on the speed and direction.
    * Compass simulations were performed to show the reduction in ambiguity removal skill caused by K_{PM}. Because the log-likelihood function is modified by larger values of K_{PM}, the first ambiguity skill is dramatically reduced.

- The pdf of Scatterometer Measurements
    – Derivation of the pdf of spectral estimates based on Welch's modified periodogram averages
        * Welch's method written in vector notation.
        * Derivation of the moment generating function, transformation to probability density function in terms of the generalized hypergeometric function and a more practical solution using a partial fraction expansion.
    – Explicit derivation (and plots) of the pdf for several representative cases with different data windows, amounts of overlap and number of frequency bins to use in the averaging, and including a case with colored noise.
    – Computation of the relative entropy between the exact pdf and a Gaussian density function (commonly used as the standard).
    – Application to NSCAT power measurements

* Derivation of the pdf of the signal-only power estimates, as the difference between the signal-plus-noise and and noise-only measurements, using data windows, overlap and numbers of bins appropriate for NSCAT.
* Computation of the normalized bias of power measurements and the probability of negative measurements.
* Wind estimation was performed using pdf's biased according to the exact pdf. The results showed a minimal bias to the wind speed and direction estimates.

6.2 Future Work

The field of oceanic microwave remote sensing includes several areas which need additional study and clarification, some of which is underway by various researchers.

Though quite far beyond the current state of oceanography, an understanding of the relationship between the geophysical parameters and the surface spectrum would be invaluable. As an intermediate step, an understanding of which parameters, along with wind speed and direction, dominate the spectrum development. The variability of empirical model functions, K_{PM}, could be used as a guide. Binning K_{PM} with various geophysical parameters (sea surface temperature, etc.) could provide a view of the sensitivity of the backscatter to those parameters.

The sensitivity of the NRCS, based on a given spectrum, was performed at nadir incidence and assuming infinite conductivity; relaxing these constraints would provide greater understanding of the scattering mechanisms and the expected returns. The work of scattering slightly off-nadir (less than about 20°) was begun by Andrew Greenwood in his Master's Thesis, [58]. Inclusion of the finite footprint effect in his work would involve substantial geometry and bookkeeping, though should prove straightforward. Moving off-nadir would stretch the footprint, making it oval instead of round, and the relative orientations of the footprint and the directional spectrum would become critical. At angles just slightly off-nadir the eccentricity of the footprint

may be sufficiently small that it could still be approximated reasonably well by a circular footprint, reducing the complexity of this extension. However, extending to larger incidence angles would quickly reduce the quality of such an approximation. At larger angles, the geometry becomes representative of scatterometers and, therefore, of substantial interest. The dominant scattering mechanism at larger angles is Bragg scattering (based on constructive interference) rather than specular reflections, and as Ref. [58] showed, asymptotic analysis can include additional terms. The use of the multiple convolutions technique will probably prove useful in off-nadir analysis, particularly as identification of more realistic spectra are found.

The analysis in this dissertation of the pdf of spectral estimates assumed only that the initial data sequence be a (possibly correlated) Gaussian stochastic process. While this is a very reasonable form for scattering from the ocean surface (the surface includes a tremendous number of scattering centers), other applications may be based on other statistics. Development of the moment-generating function of a quadratic form in a non-Gaussian vector would be useful, though probably daunting. Application of the pdf to other scatterometers, such as ERS-1/2 and SeaWinds, would be useful, to verify that sufficient averaging is being used to allow the Gaussian approximation in wind retrieval.

In the dissertation I computed and plotted the probability of negative backscatter measurements. Scatterometer data could be analyzed to identify areas with many negative measurements, and possibly identify a correlation between the frequency of negative measurements and the lowest wind speed identifiable by the spaceborne scatterometer. Specifically, some have suggested that the simple empirical models of the wind-backscatter relationship, with the backscatter proportional to the wind speed raised to some power, is only valid above some low wind speed threshold. Correlation of a large frequency of negative measurements, beyond that predicted using the correct pdf, with areas of low wind speed, could help in identifying a minimum wind speed for the empirical models.

Appendix A

Composite models of the Ocean Surface

The spectrum presented in Chapter 2 is quite simplistic. Here I derive the correlation function for a composite spectrum. Consider the power-law spectrum defined as

$$S(k,\theta) = S_0 \begin{cases} 0 & \text{for } k \le k_0 \\ k^{-p_0}\Theta(\theta,\theta_0) & \text{for } k_0 \le k < k_1 \\ \beta k^{-p_1}\Theta(\theta,\theta_1) & \text{for } k \ge k_1 \end{cases} \tag{A.1}$$

where β provides a scaling factor to adjust the relative amount of power in each section of the spectrum and $\Theta(\theta,\theta_i)$ indicates the possibility of different directionality functions for the two regions. Here I will use hard limit cutoffs at $\pm\frac{\theta_i}{2}$, though comparable results could be obtained for a $\cos^{q_i}$ form of directionality.

The correlation function is then expressed as

$$\begin{aligned} C(\rho,\phi) \quad = \quad & S_0 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta\Theta(\theta,\theta_0) \int_{k_0}^{\infty} dk \;\; k^{-p_0+1}\cos[k\rho|\cos(\theta-\phi)|] \\ - \quad & S_0 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta\Theta(\theta,\theta_0) \int_{k_1}^{\infty} dk \;\; k^{-p_0+1}\cos[k\rho|\cos(\theta-\phi)|] \\ + \quad & S_0\beta \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta\Theta(\theta,\theta_1) \int_{k_1}^{\infty} dk \;\; k^{-p_1+1}\cos[k\rho|\cos(\theta-\phi)|] \end{aligned} \tag{A.2}$$

The scale factor S_0 must be such that the correlation function is 1 at $\rho = 0$:

$$\begin{aligned} S_0 \quad &= \quad \left(\frac{\theta_0}{p_0-2}k_0^{-p_0+2} - \frac{\theta_0}{p_0-2}k_1^{-p_0+2} + \beta\frac{\theta_1}{p_1-2}k_1^{-p_1+2}\right)^{-1} \\ &= \quad \left[S_{000}^{-1} - S_{100}^{-1} + \beta S_{111}^{-1}\right]^{-1} \end{aligned} \tag{A.3}$$

where the subscripts of S_{lmn} are binary indicators of k_l, p_m and θ_n and S_{lmn} is the appropriate scaling factor of the power spectral density function associated with l,

m, and n. Thus, the correlation function for a composite surface model is a linear combination of correlation functions based on the parameters of the composite model:

$$\begin{aligned} C(\rho,\phi) &= \frac{S_0}{S_{000}}C(\rho,\phi;k_0,p_0,\theta_0) - \frac{S_0}{S_{100}}C(\rho,\phi;k_1,p_0,\theta_0) \\ &\quad +\beta\frac{S_0}{S_{111}}C(\rho,\phi;k_1,p_1,\theta_1) \end{aligned} \tag{A.4}$$

which can be written in the same form as Eq. (2.17), with

$$C = 1 - A_0\rho^{p_0-2} - A_1\rho^{p_1-2} + \sum_{r=1}^{\infty} B_r\rho^{2r}, \tag{A.5}$$

with A_0, A_1, and B_1 all of the form $a\cos 2\phi + b$.

Note that the dominant term of the correlation function of a composite surface has the same form, $C \approx 1 - A(a\cos 2\phi + b)\rho^{\epsilon}$, as that found for the basic surface. From this we conclude that the backscatter from such a surface, using a finite footprint, will follow the same asymptotic behavior exhibited in Eq. (2.7) resulting in two modes of asymptotic behavior.

Appendix B

Expansion of a Cosine to a Power

The function, $f(x) = \cos^n(x)$ can be written as the sum of cosine terms; in this appendix, this expansion is derived. The cosine function can be expanded with Euler's identity

$$\begin{aligned} f(x) &= \cos^n(x) && \text{(B.1)} \\ &= \left(\frac{e^x + e^{-x}}{2}\right)^n && \text{(B.2)} \end{aligned}$$

and the binomial approximation applied [59]:

$$\begin{aligned} f(x) &= 2^{-n} \sum_{k=0}^{n} \binom{n}{k} e^{kx} e^{-x(n-k)} && \text{(B.3)} \\ &= 2^{-n} \sum_{k=0}^{n} \binom{n}{k} e^{-x(n-2k)} && \text{(B.4)} \\ &= 2^{-n} \sum_{k=0}^{n} \binom{n}{k} \left[\cos\left(-x(n-2k)\right) + j \sin\left(-x(n-2k)\right)\right]. && \text{(B.5)} \end{aligned}$$

Recognizing the symmetry of the series, $f(x)$ can be written (for an even integer, $n = 2r$):

$$f(x) = 2^{-2r} \binom{2r}{r} + 2^{-2r+1} \sum_{k=0}^{r-1} \binom{2r}{k} \cos\left(2x(r-k)\right). \qquad \text{(B.6)}$$

If $n = 2r + 1$ is odd, the dc term is zero:

$$f(x) = 2^{-2r} \sum_{k=0}^{r} \binom{2r+1}{k} \cos\left(x(2r-2k+1)\right). \qquad \text{(B.7)}$$

Appendix C

Estimates of K_{PM} for the NSCAT Mission

This Appendix serves to archive the substantial number of plots used to display the estimates of the model function variability, K_{PM}, based on NSCAT data.

Figures C.1 through C.7 display histograms of the wind speed for each week of data (based on the 40 descending revs). Histograms for each week are provided for six latitude bands, -60 to -40, -40 to -20, -20 to 0, 0 to 20, 20 to 40, and 40 to 60. The plots are arranged such that the southern hemisphere plots are on the left and the northern hemisphere on the right, with the equatorial regions at the bottom and the high latitudes at the top.

Figures C.8 through C.14 display plots of K_{PM} vs. wind speed, for the sampled weeks, with a separate line for each incidence angle. The higher incidence angles (far swath) produce higher estimates of K_{PM}. The curves correspond to six incidence angle ranges of (20, 25), (25, 30), (30, 35), (35, 40), (40, 45) and (45, 50) degrees. The six subfigures correspond to the six latitude bands as described above.

Figures. C.15 through C.22 display the estimates of K_{PM} as functions of time through the mission, with a separate line for each wind speed and a separate plot for each incidence angle bin. The estimate of K_{PM} is smaller for higher wind speeds. The speeds bins plotted are 2 m/s wide, centered at 3, 7, 11, 15 and 19 m/s.

Figures C.23 through C.31 display the estimates of K_{PM} as functions of time through the mission, with a separate line for each incidence angle and a separate plot for each wind speed bin. Again, the curves are for the incidence angle ranges of (20, 25), (25, 30), (30, 35), (35, 40), (40, 45) and (45, 50) degrees.

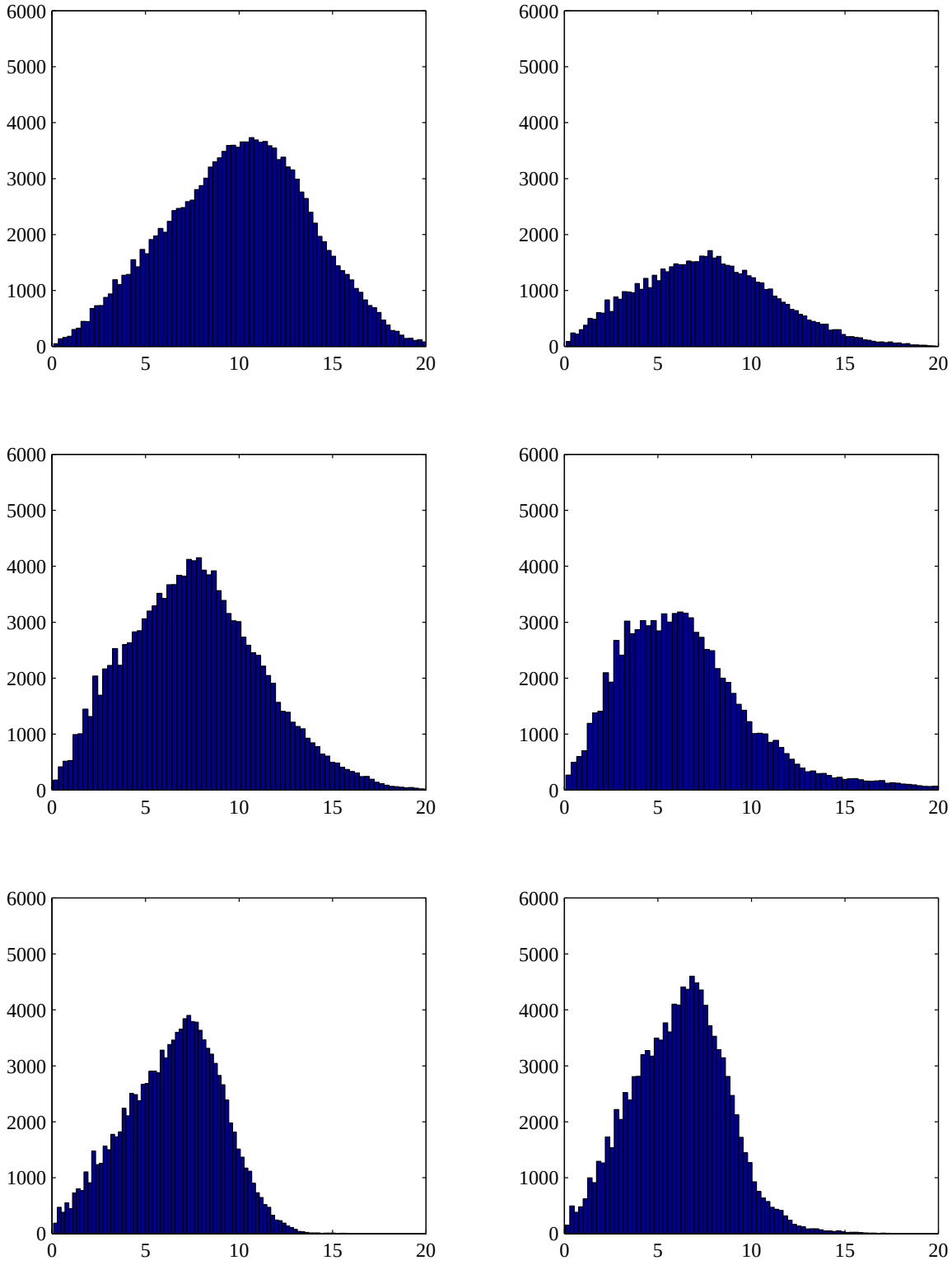


Figure C.1: *Wind speed histograms for week S3 (Sep 26). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

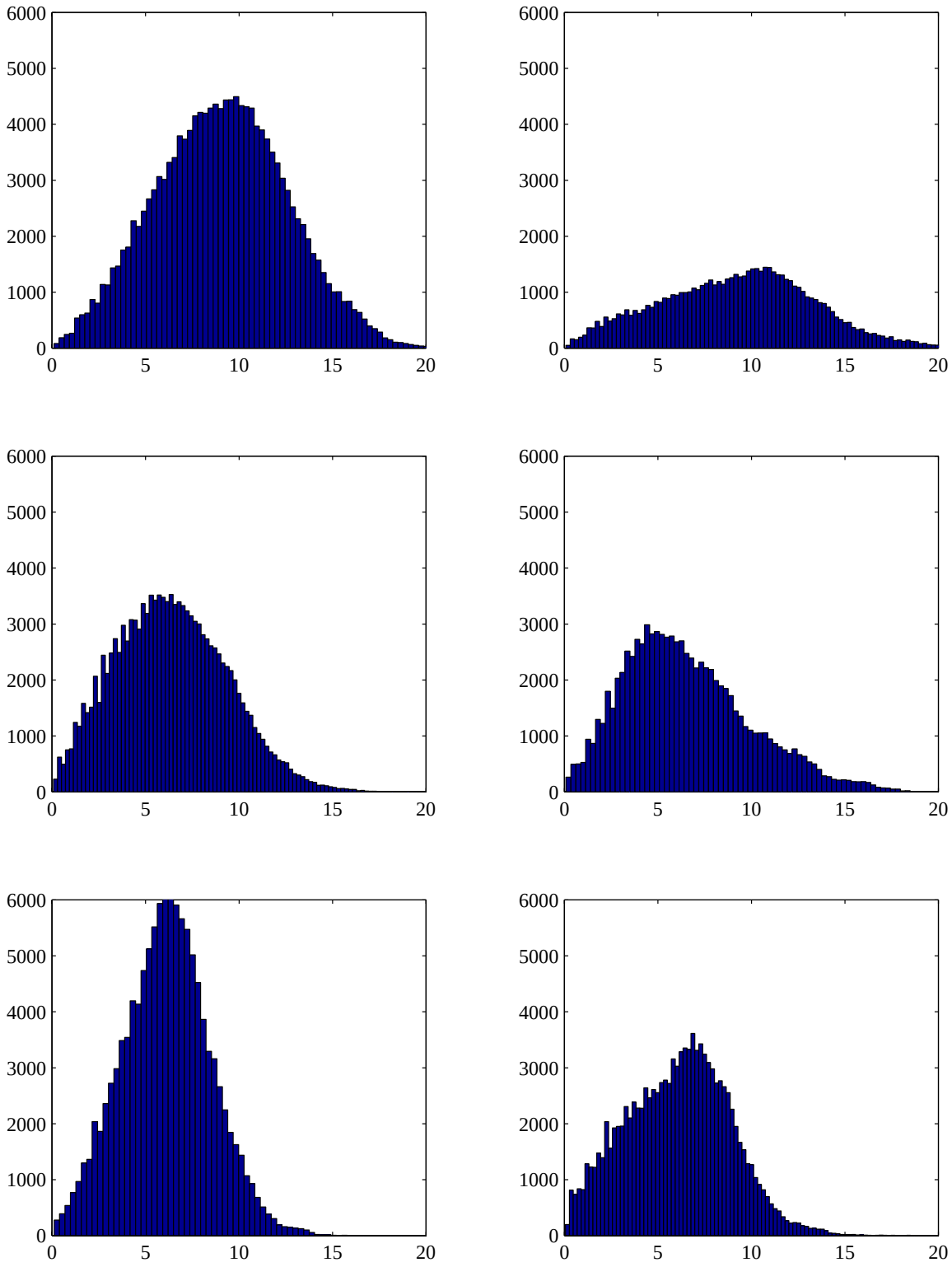


Figure C.2: *Wind speed histograms for week S9 (Oct 29). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

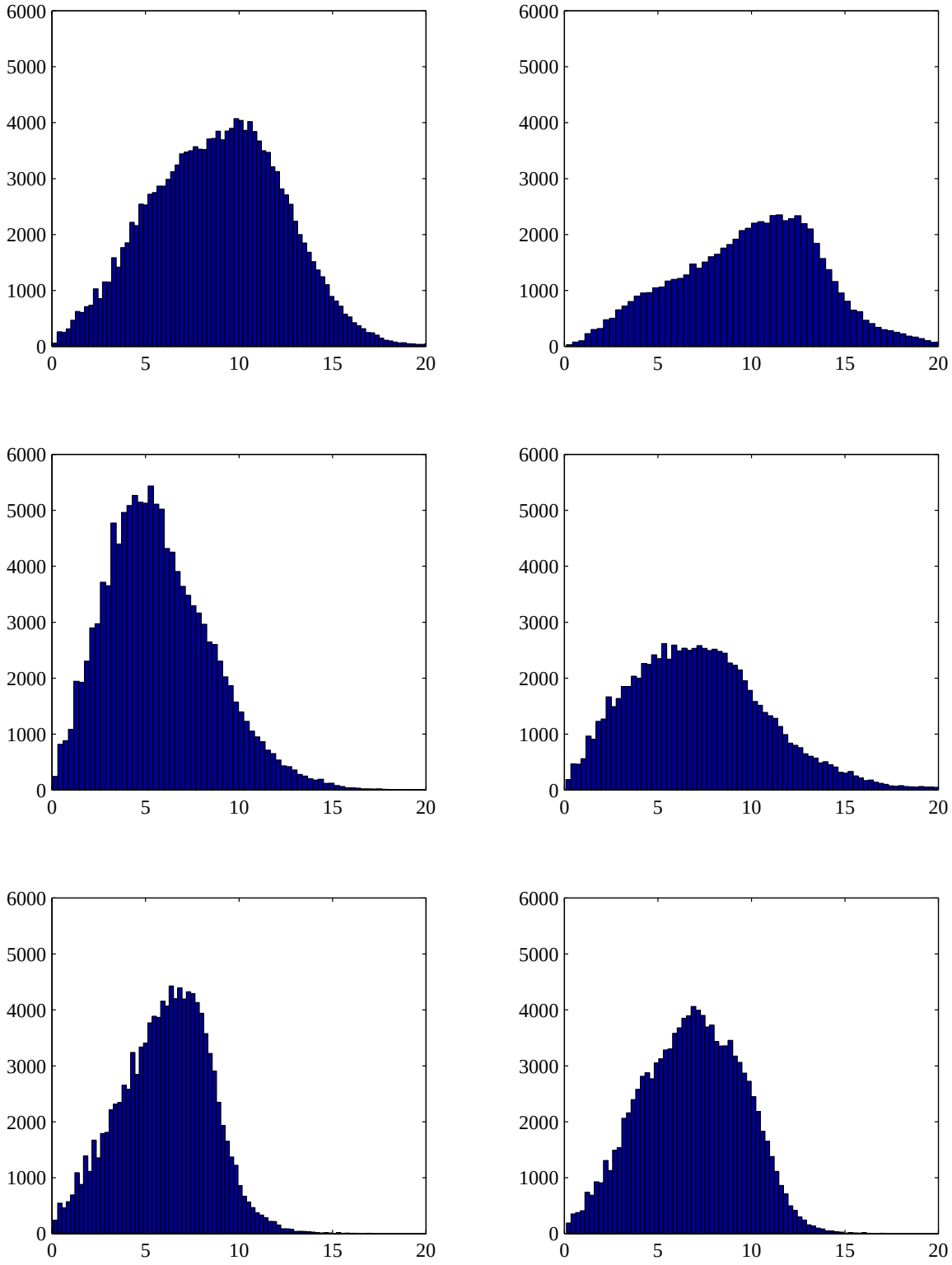


Figure C.3: *Wind speed histograms for week S15 (Dec 11). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

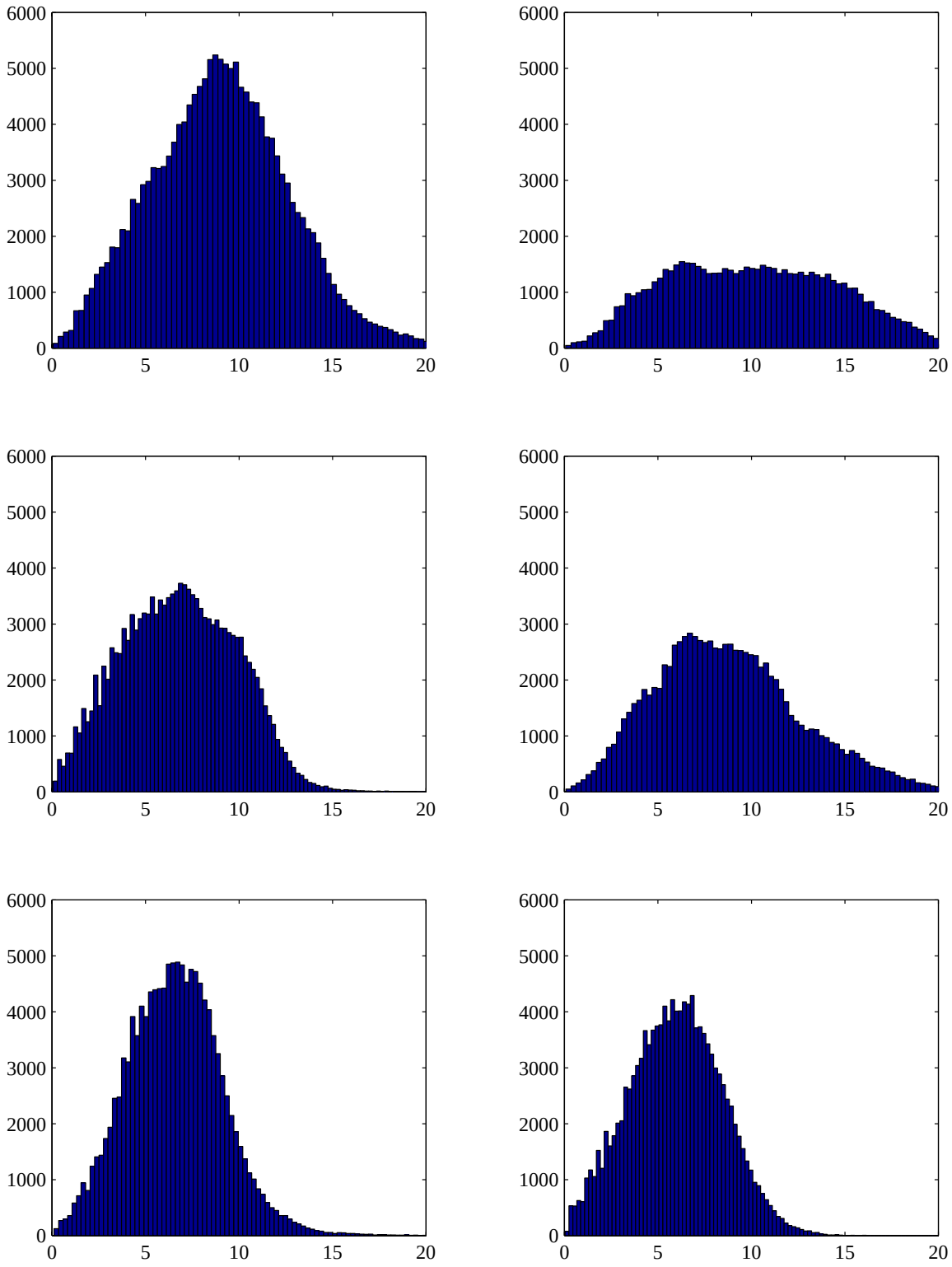


Figure C.4: *Wind speed histograms for week S21 (Jan 22). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

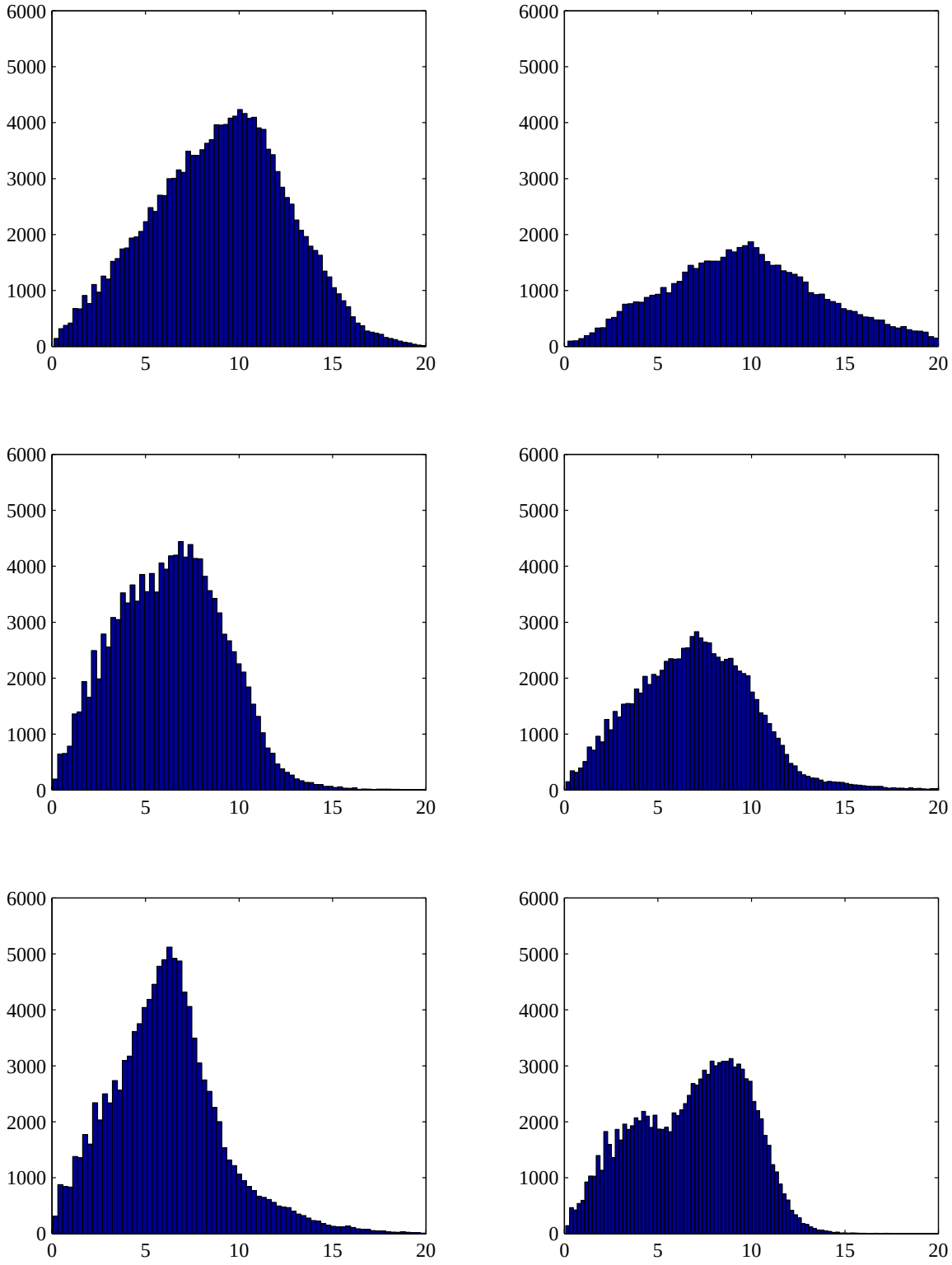


Figure C.5: *Wind speed histograms for week S27 (Mar 05). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

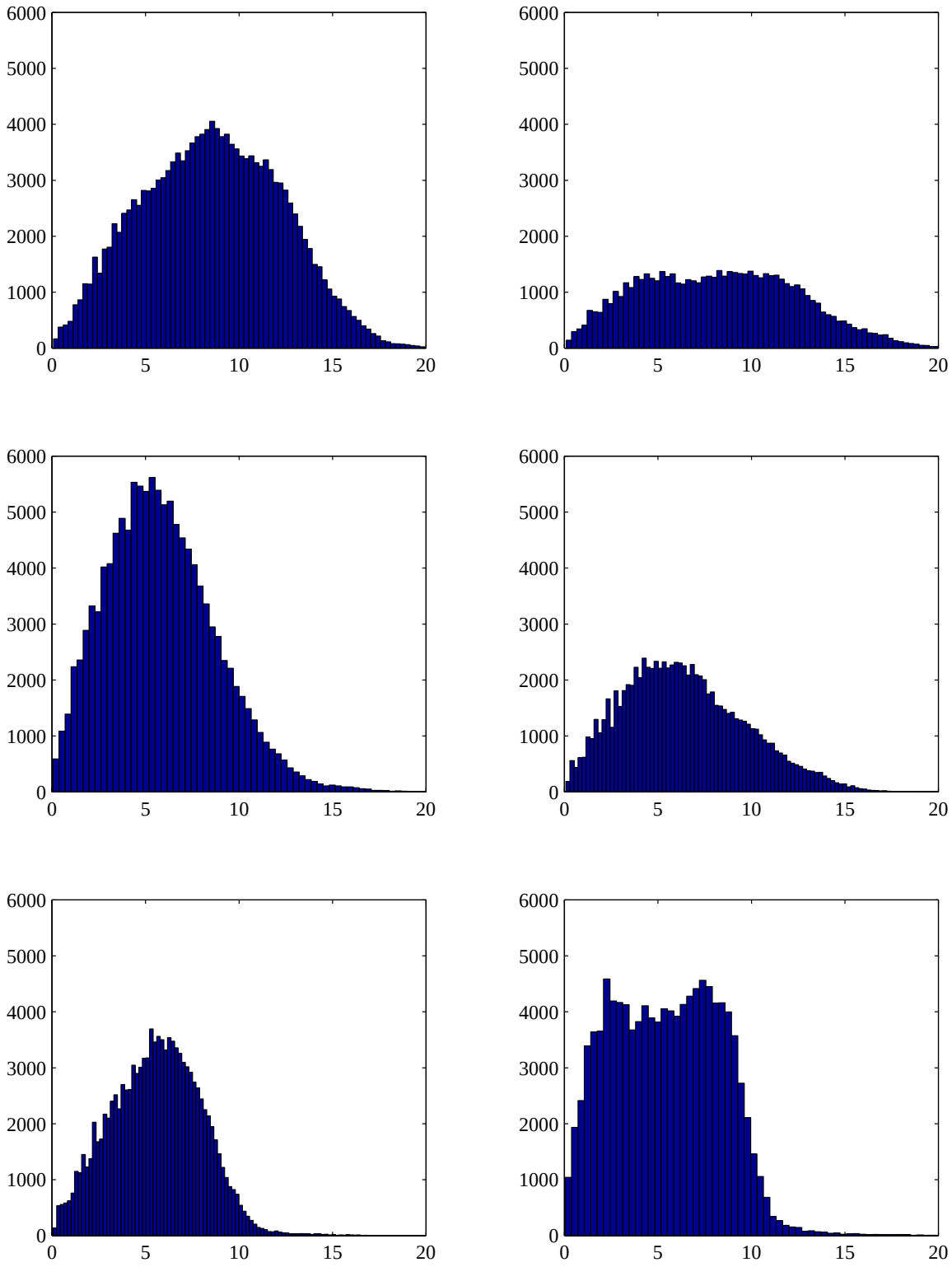


Figure C.6: *Wind speed histograms for week S33 (Apr 16). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

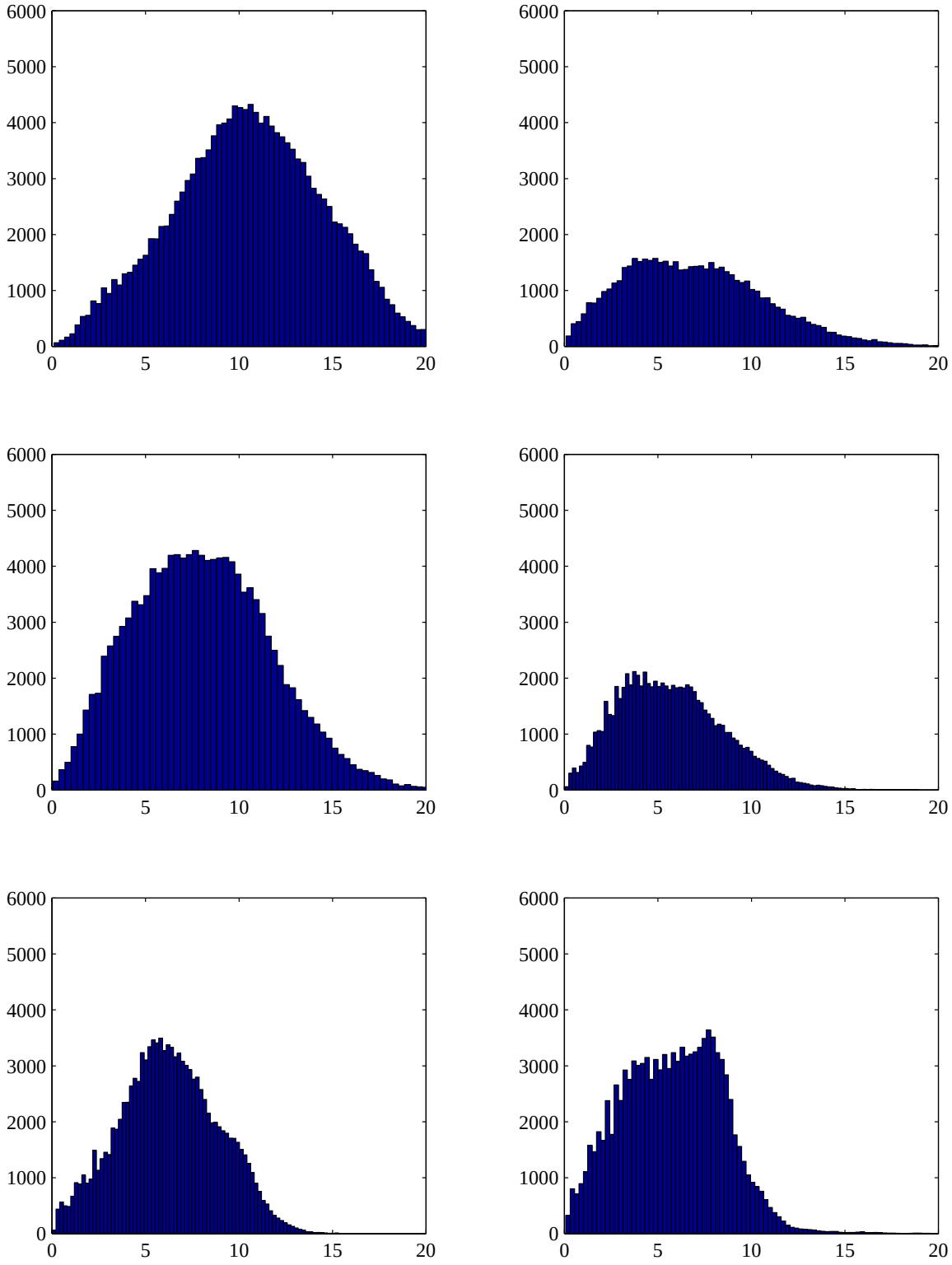


Figure C.7: *Wind speed histograms for week S39 (May 27). Plots on the left are for the southern hemisphere, plots on the right are for the northern hemisphere; near equatorial regions at the bottom, high latitudes are at the top.*

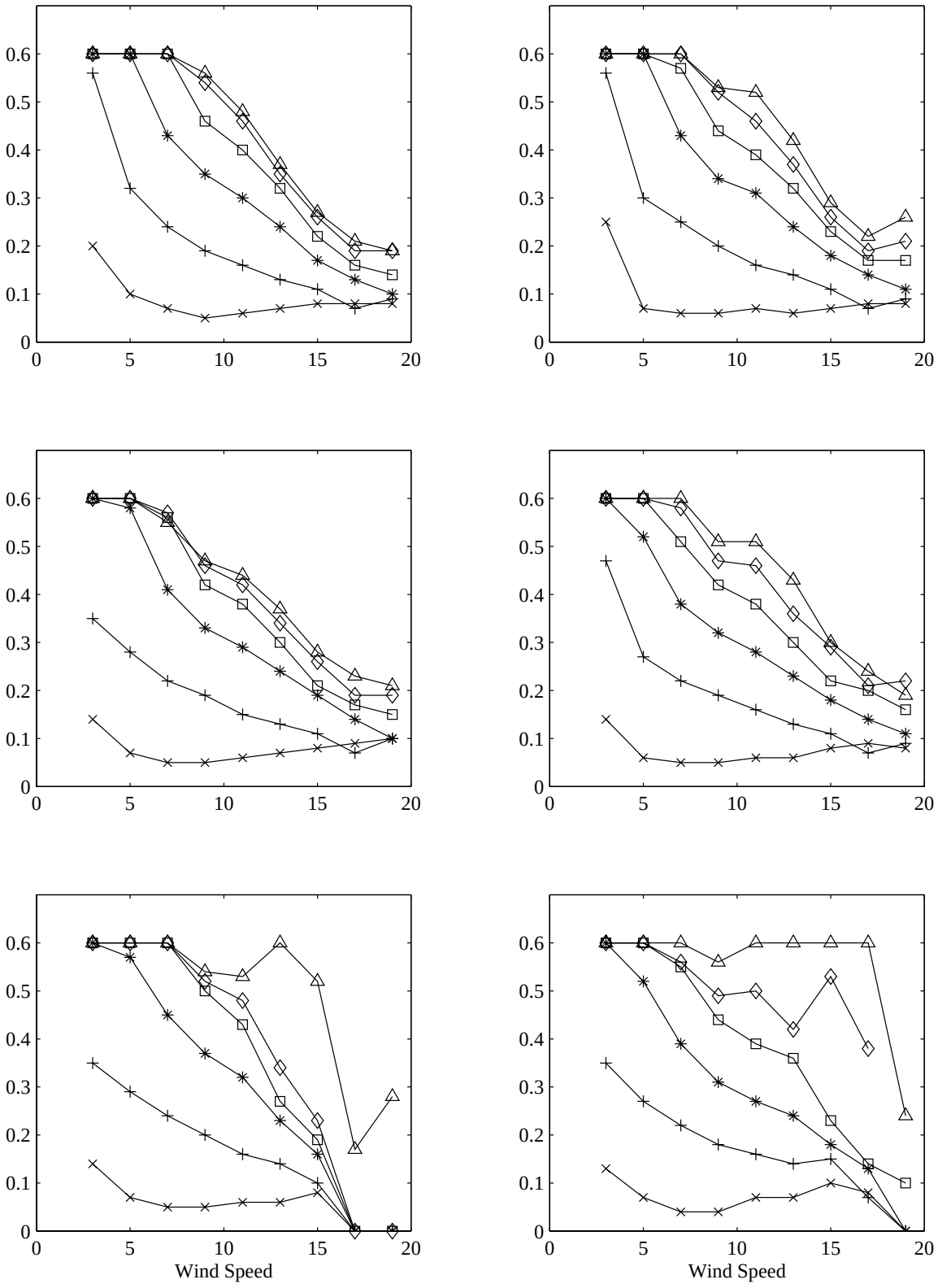


Figure C.8: *K_{PM} estimates for week S3 (Sep 26). Each line represents a separate incidence angle bin.*

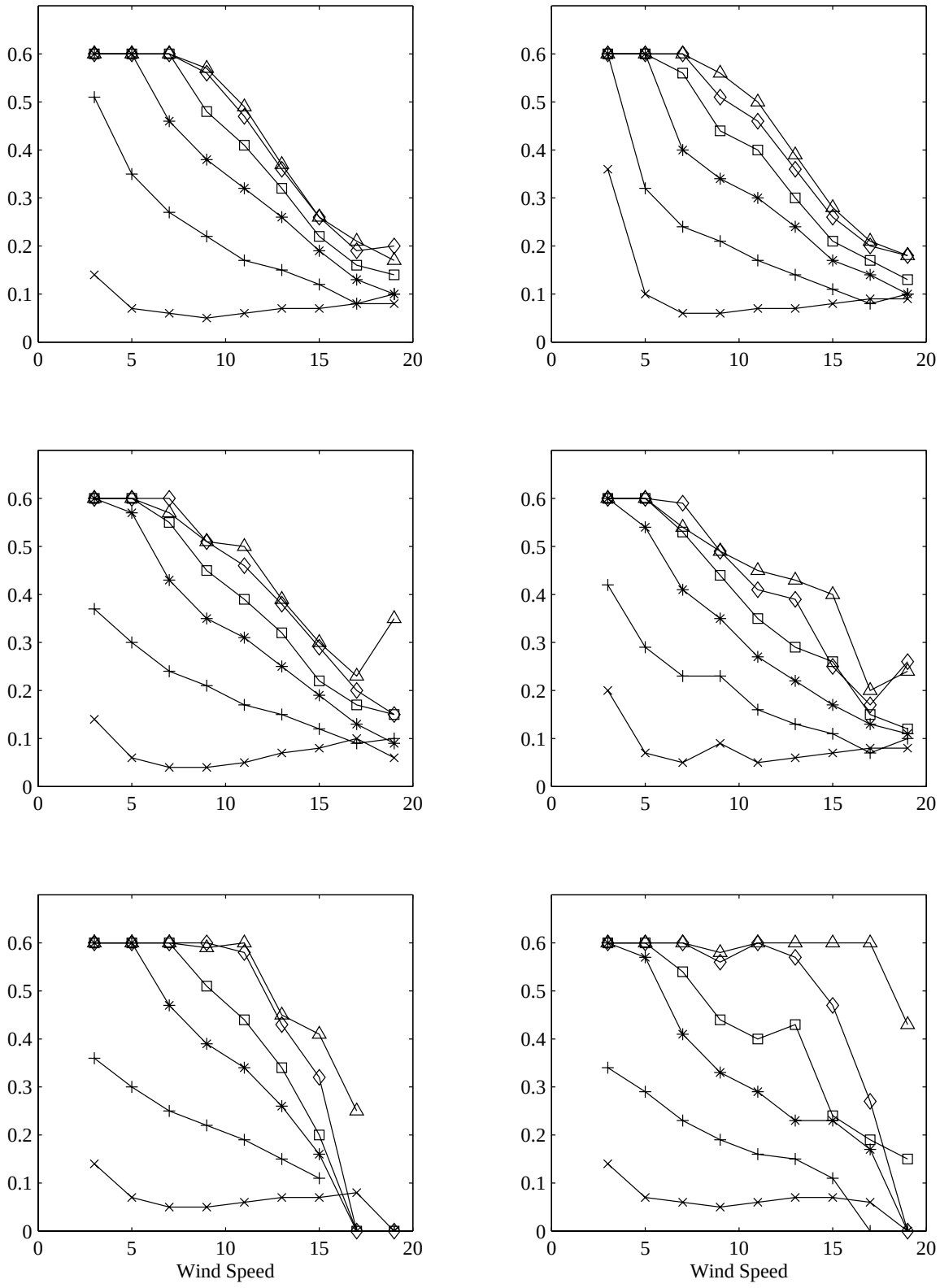


Figure C.9: *K_{PM} estimates for week S9 (Oct 29). Each line represents a separate incidence angle bin.*

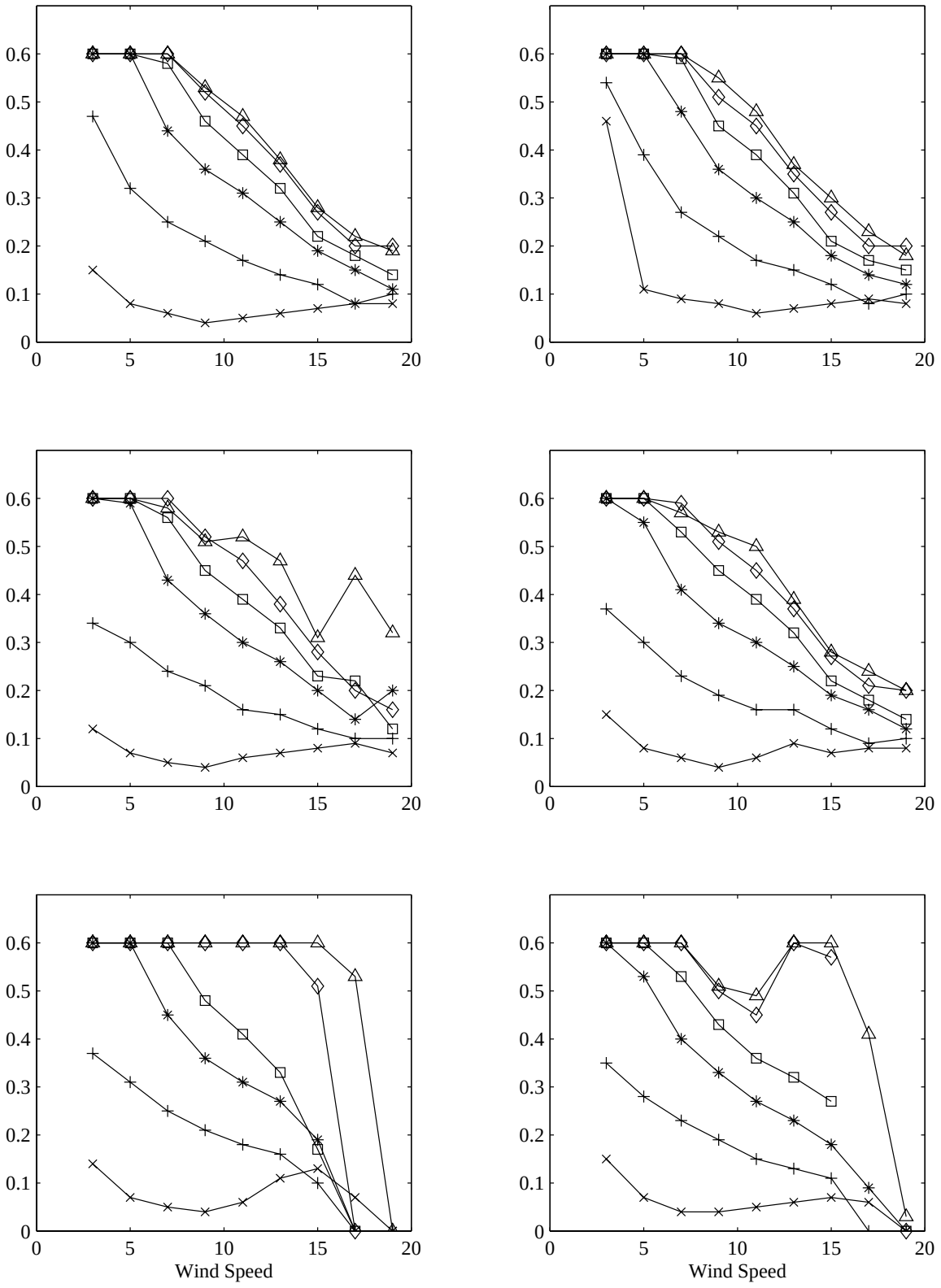


Figure C.10: *K_{PM} estimates for week S15 (Dec 11). Each line represents a separate incidence angle bin.*

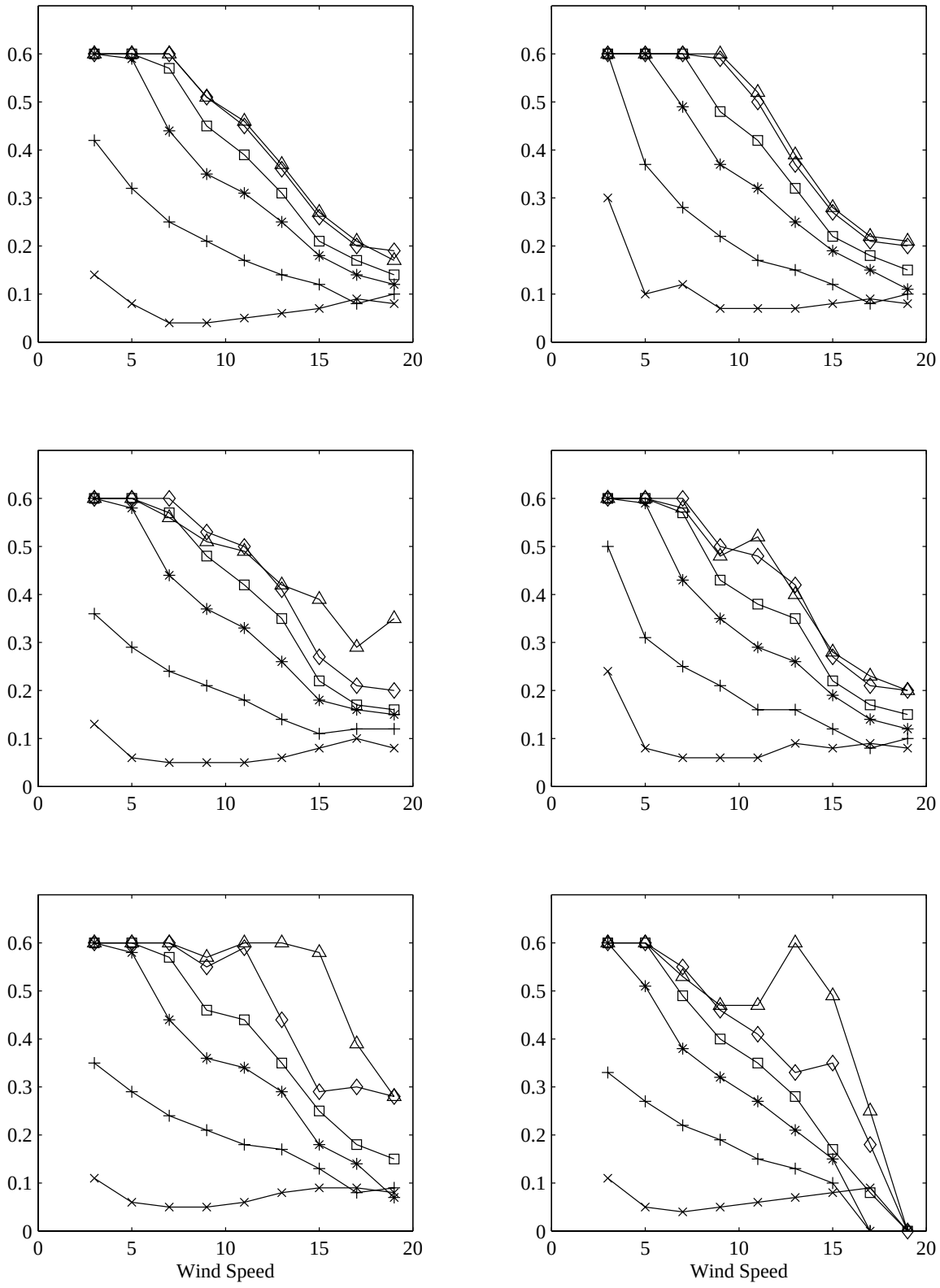


Figure C.11: *K_{PM} estimates for week S21 (Jan 22). Each line represents a separate incidence angle bin.*

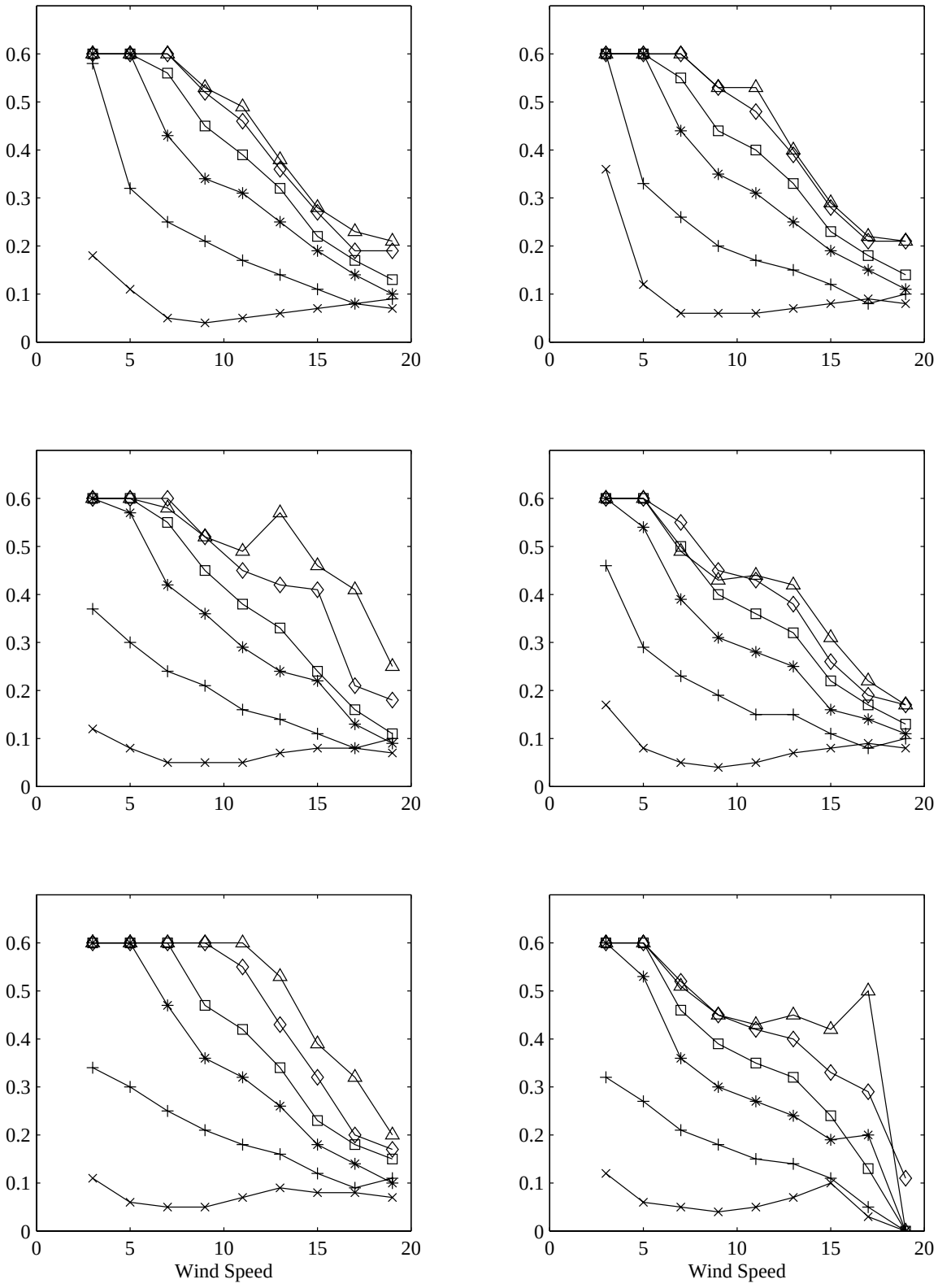


Figure C.12: *K_{PM} estimates for week S27 (Mar 05). Each line represents a separate incidence angle bin.*

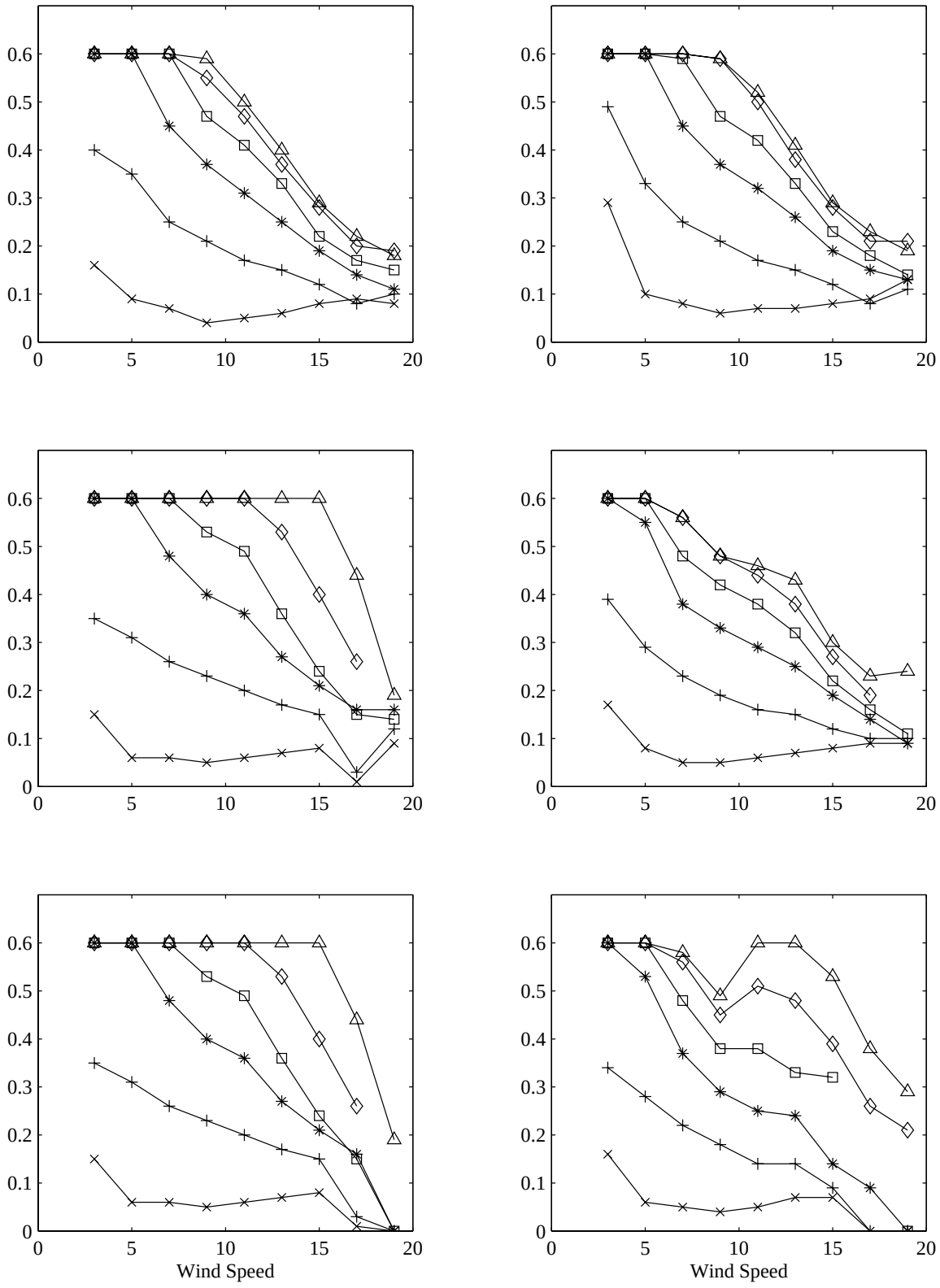


Figure C.13: *K_{PM} estimates for week S33 (Apr 16). Each line represents a separate incidence angle bin.*

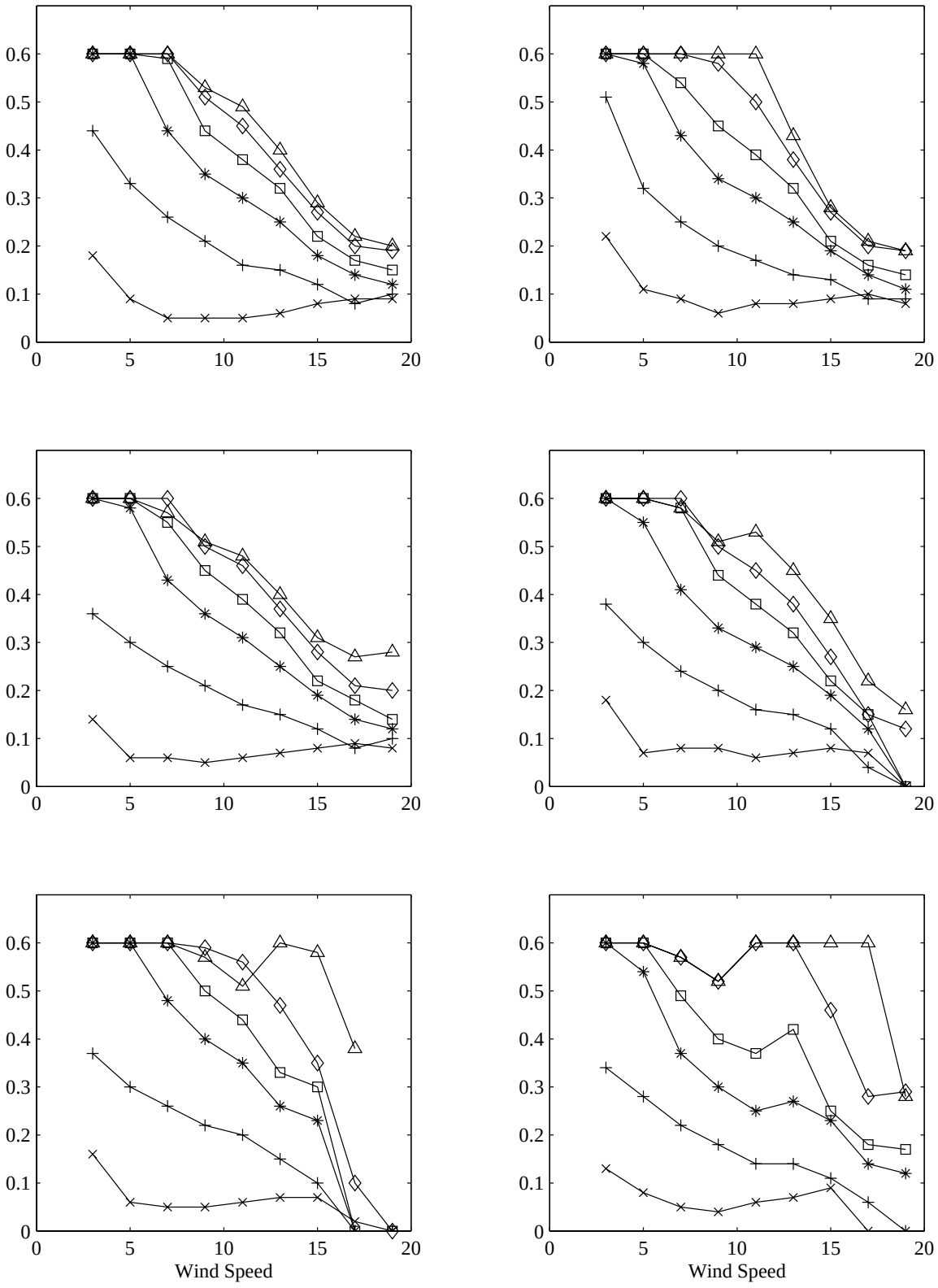


Figure C.14: *K_{PM} estimates for week S39 (May 27). Each line represents a separate incidence angle bin.*

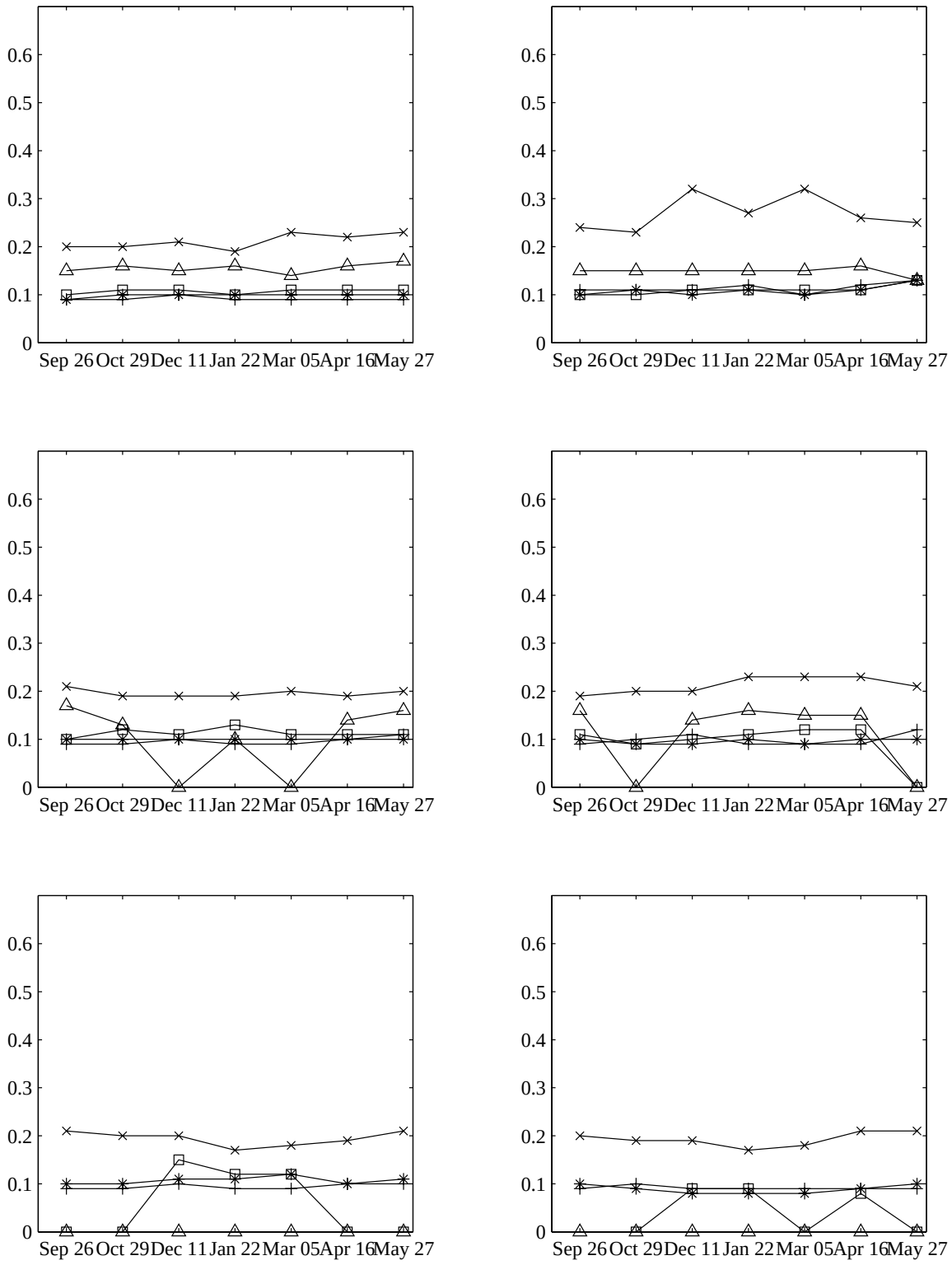


Figure C.15: *K_{PM} estimates for incidence angles between 15 and 20°. Each line represents a separate speed bin.*

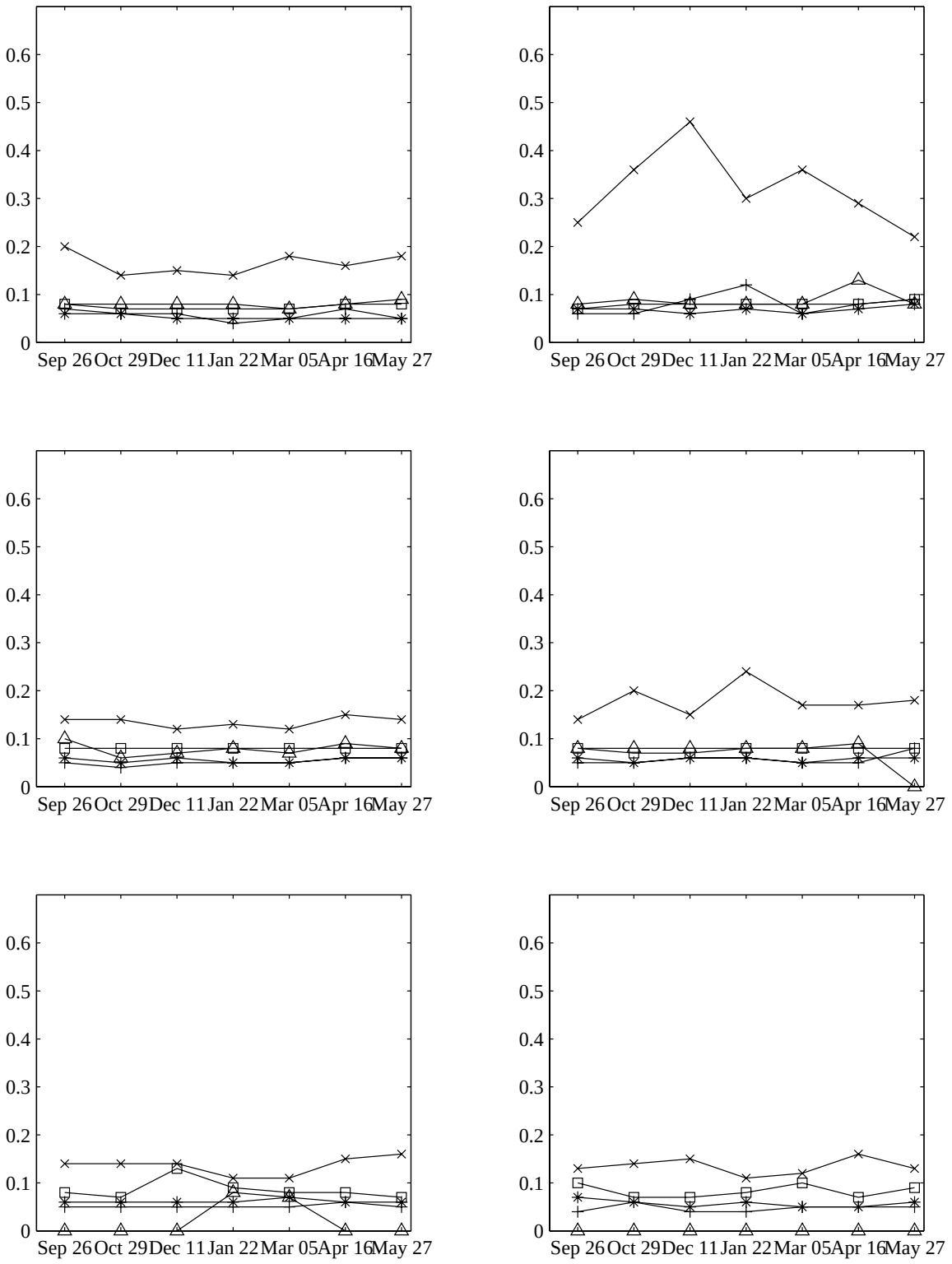


Figure C.16: K_{PM} *estimates for incidence angles between 20 and 25°. Each line represents a separate speed bin.*

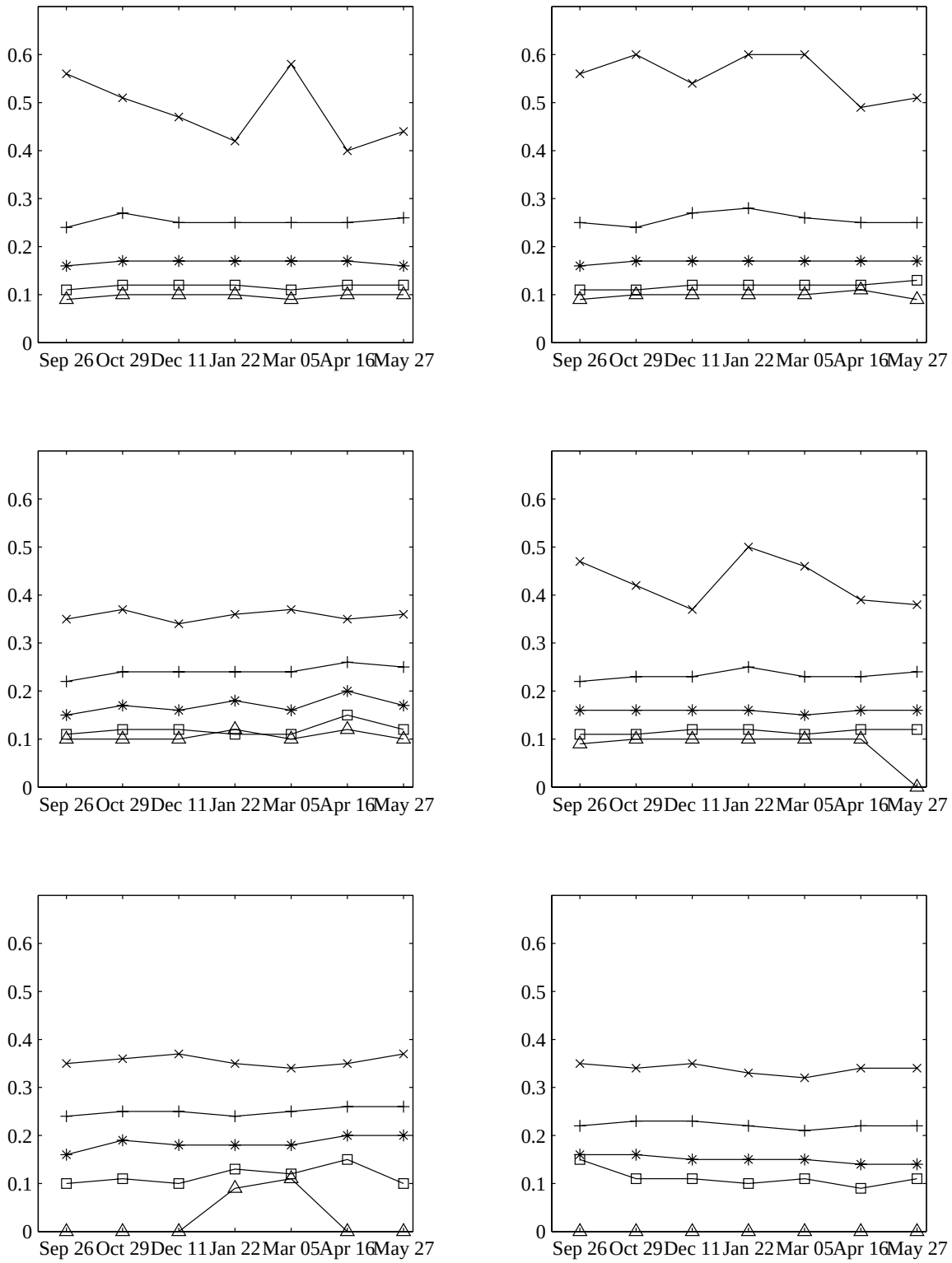


Figure C.17: *K_{PM} estimates for incidence angles between 25 and 30°. Each line represents a separate speed bin.*

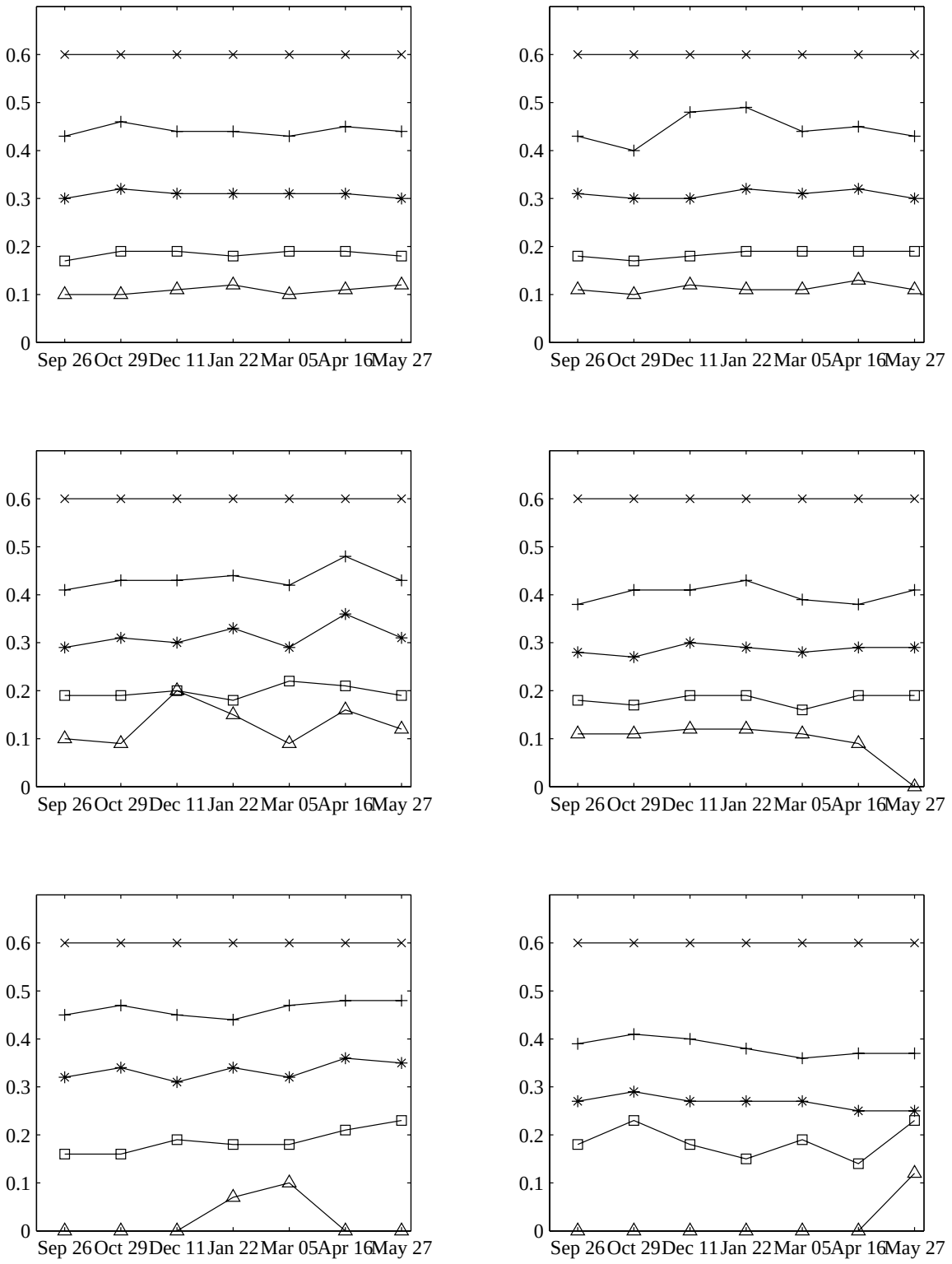


Figure C.18: K_{PM} *estimates for incidence angles between 30 and 35°. Each line represents a separate speed bin.*

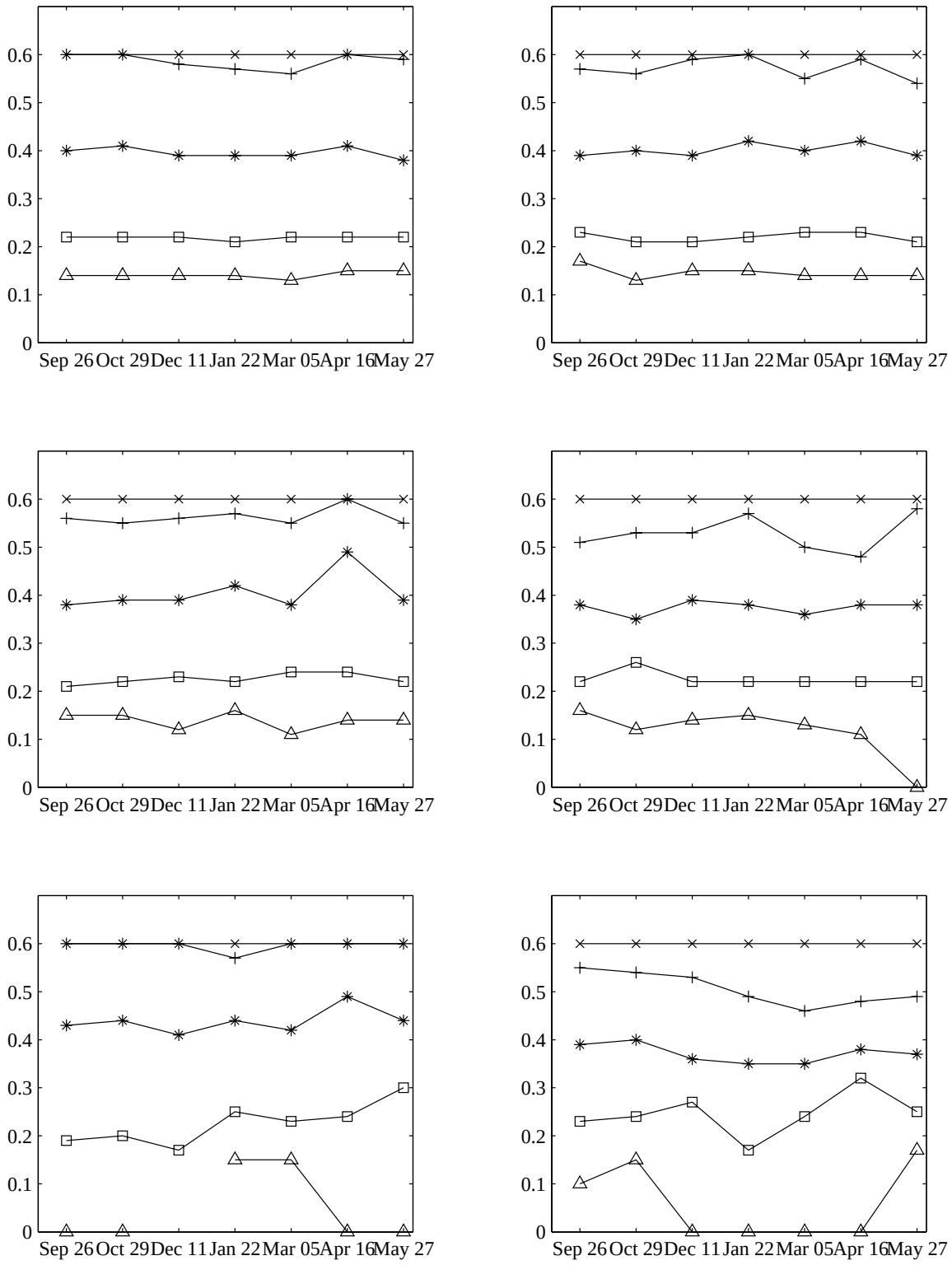


Figure C.19: *K_{PM} estimates for incidence angles between 35 and 40°. Each line represents a separate speed bin.*

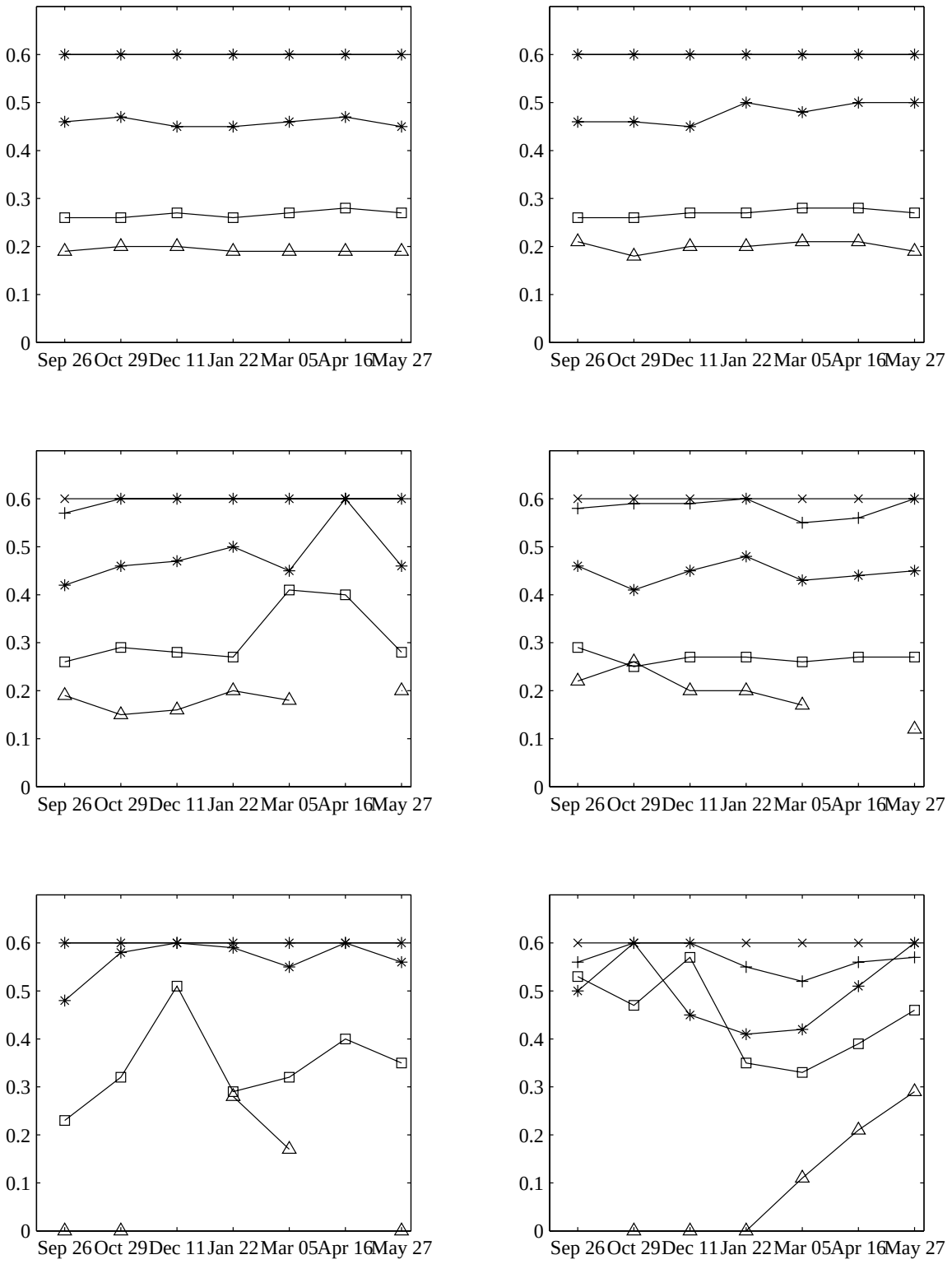


Figure C.20: K_{PM} *estimates for incidence angles between 40 and 45°. Each line represents a separate speed bin.*

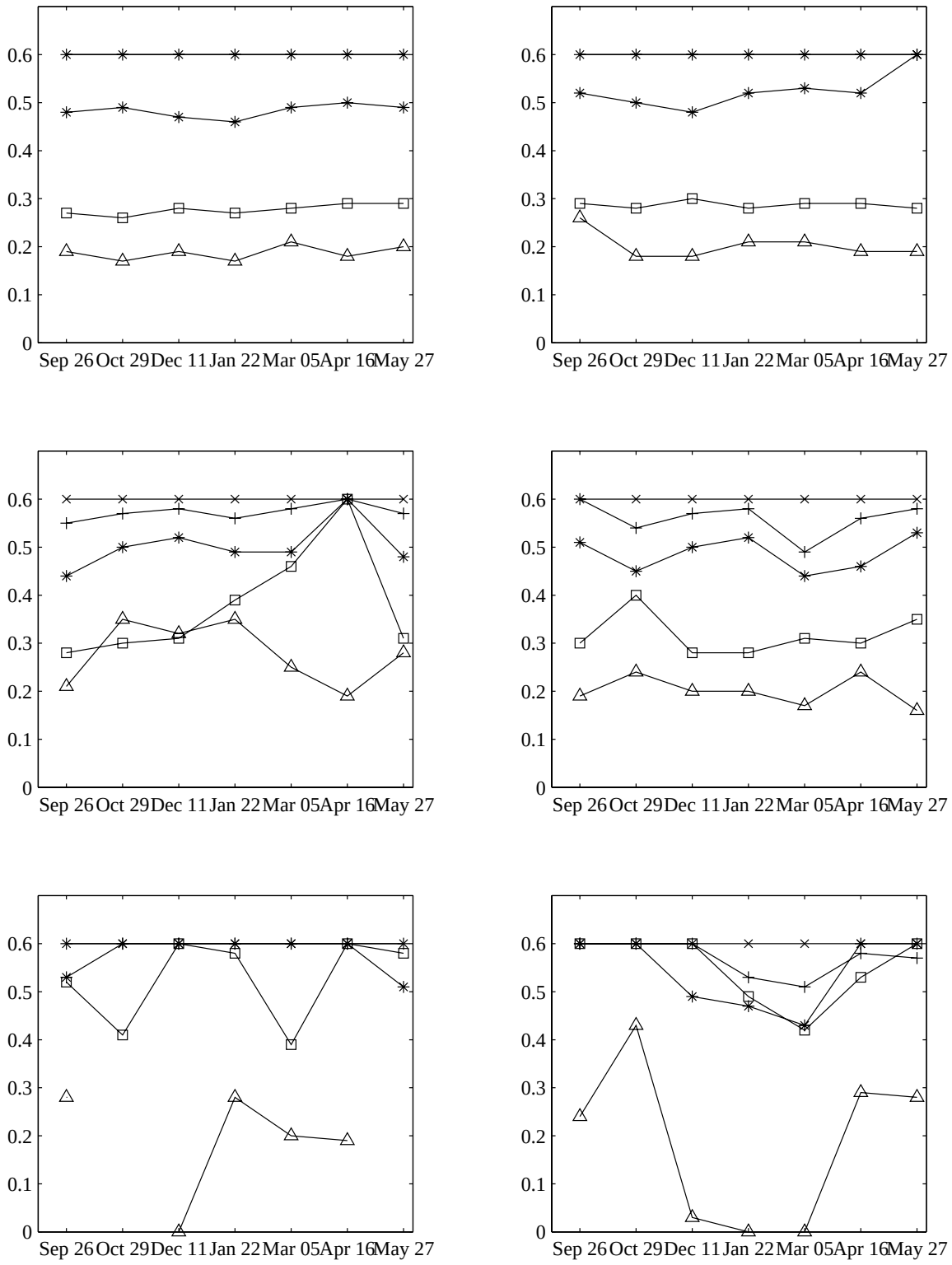


Figure C.21: *K_{PM} estimates for incidence angles between 45 and 50°. Each line represents a separate speed bin.*

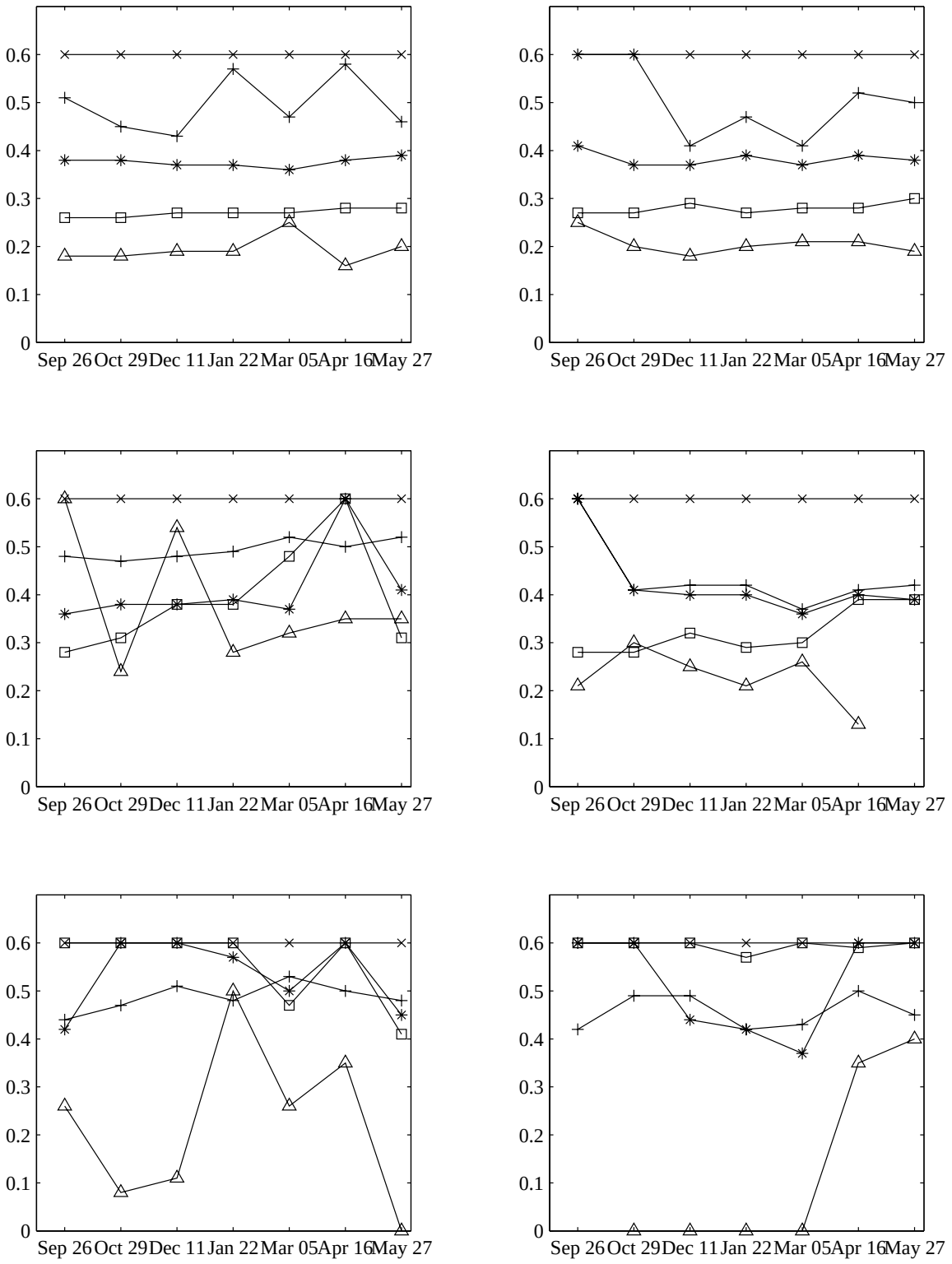


Figure C.22: K_{PM} *estimates for incidence angles between 50 and 55°. Each line represents a separate speed bin.*

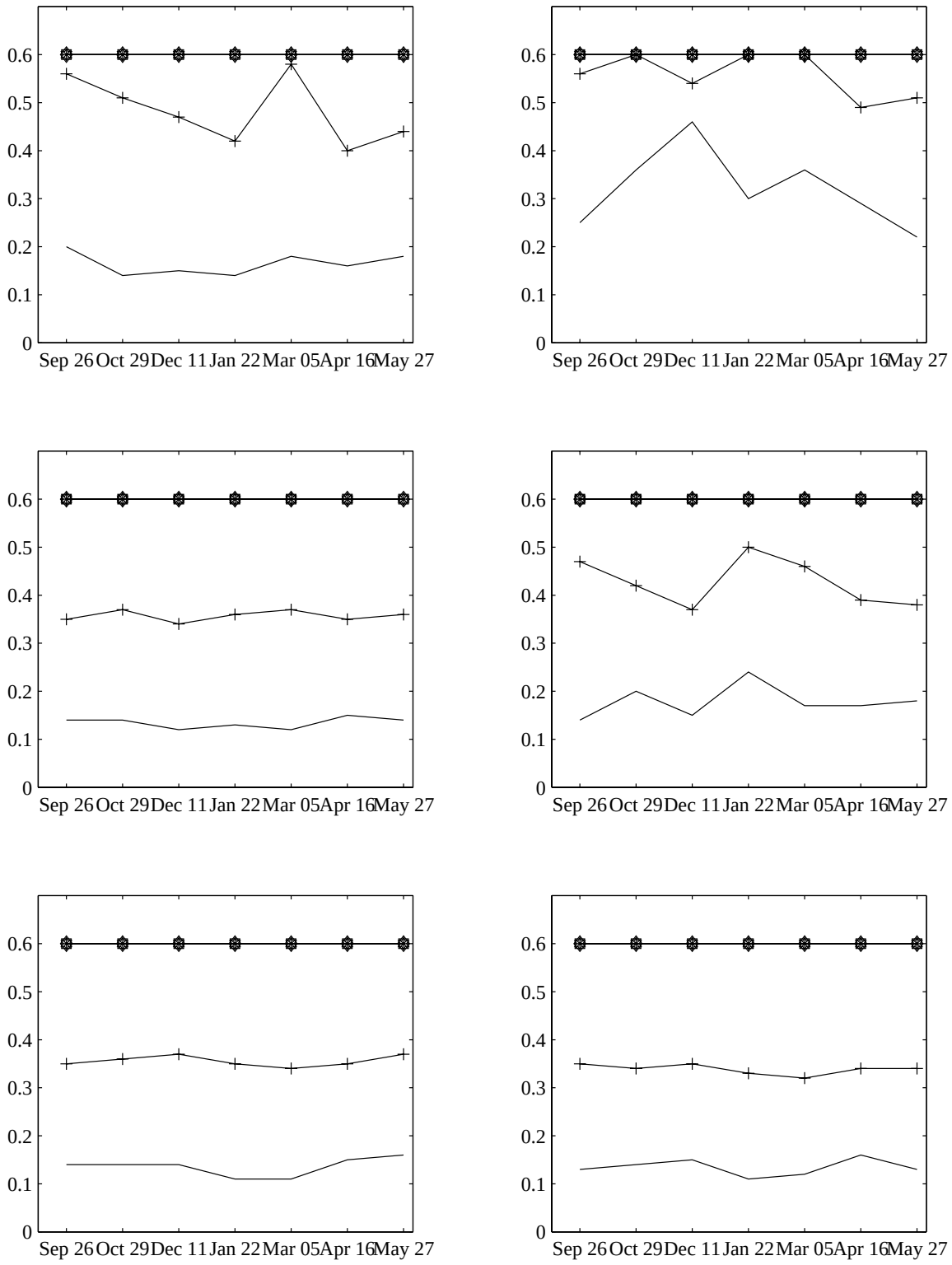


Figure C.23: K_{PM} *estimates for wind speeds of 3 m/s. Each line represents a separate incidence angle.*

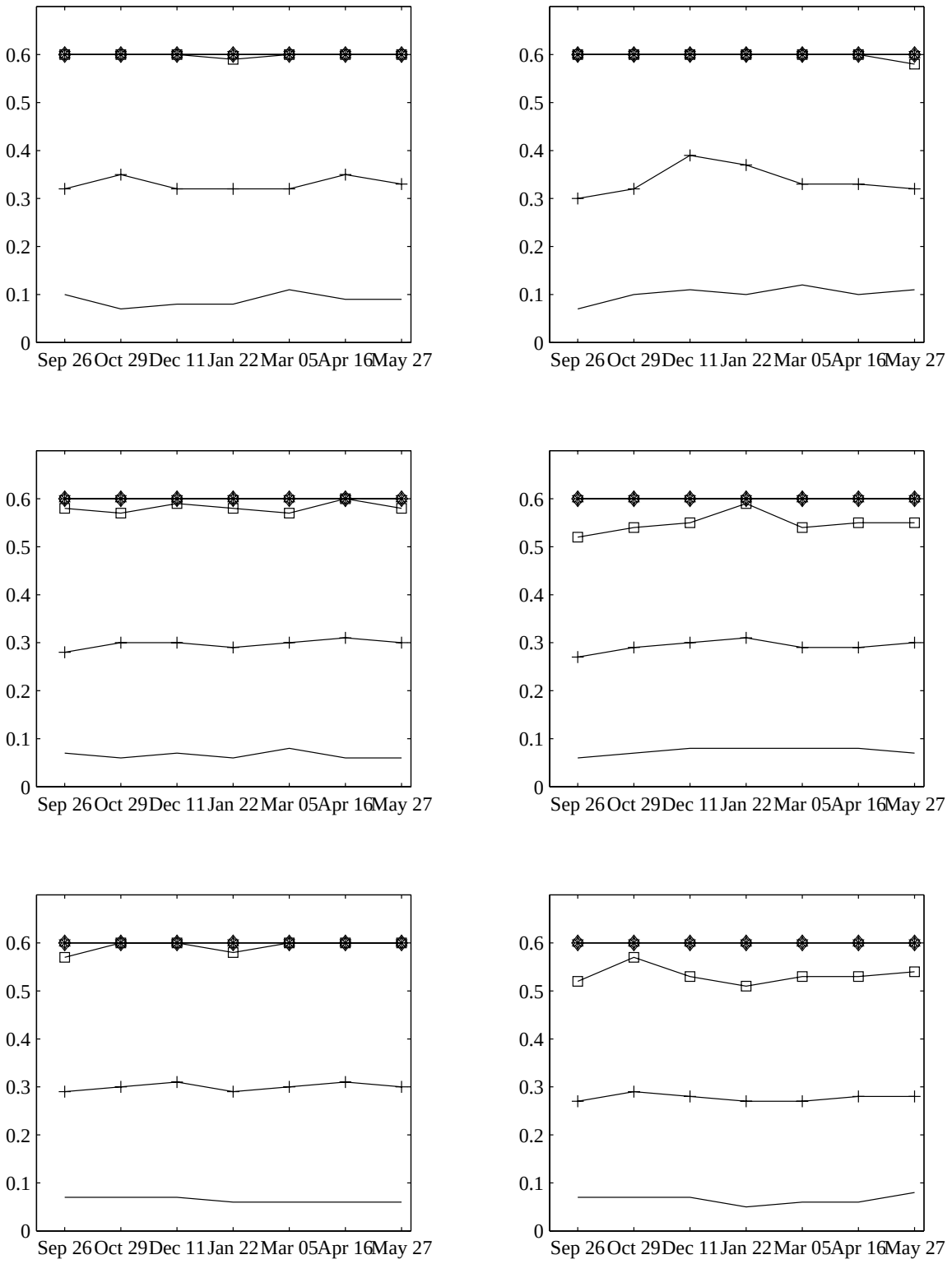


Figure C.24: K_{PM} *estimates for wind speeds of 5 m/s. Each line represents a separate incidence angle.*

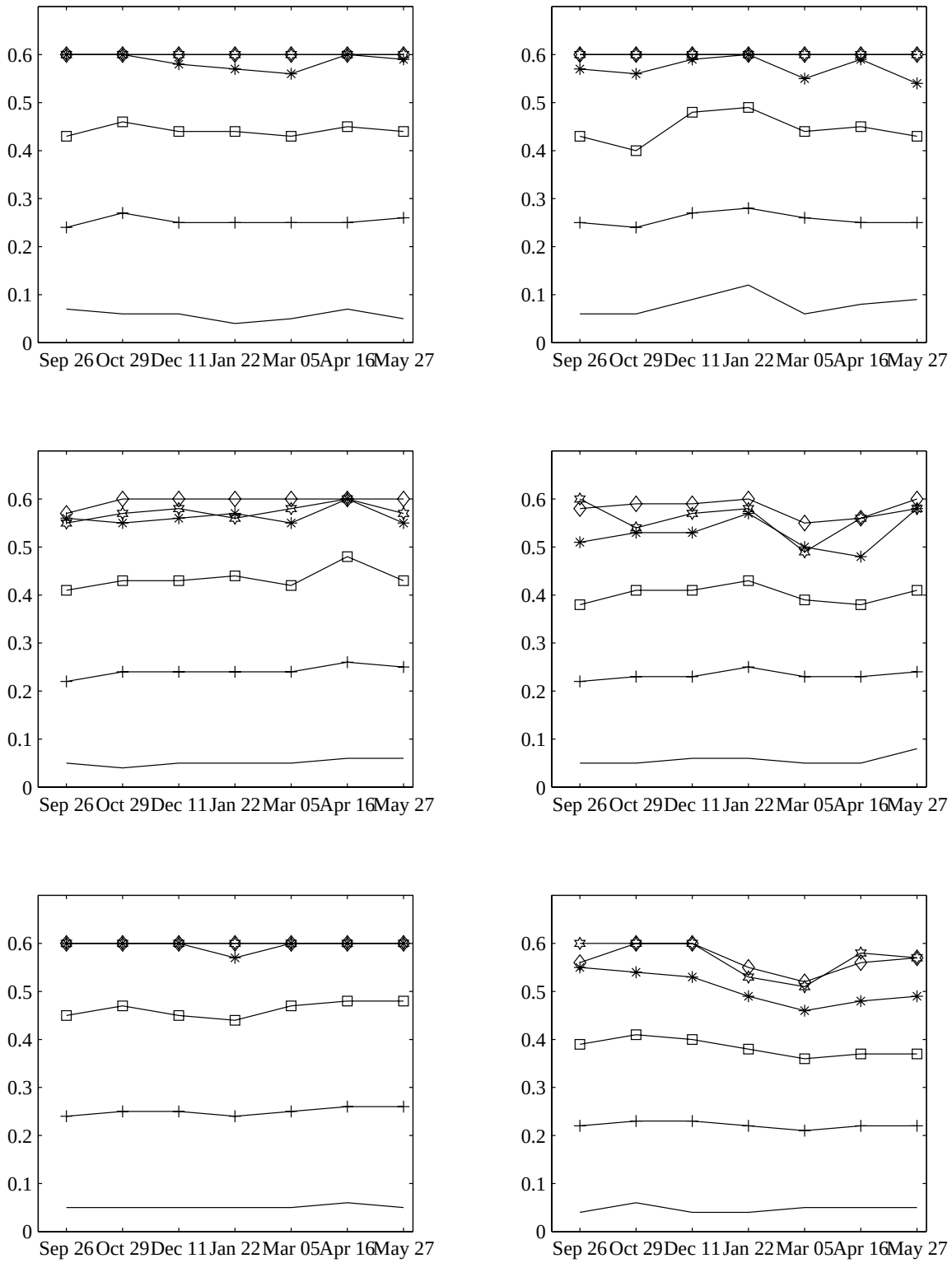


Figure C.25: *K_{PM} estimates for wind speeds of 7 m/s. Each line represents a separate incidence angle.*

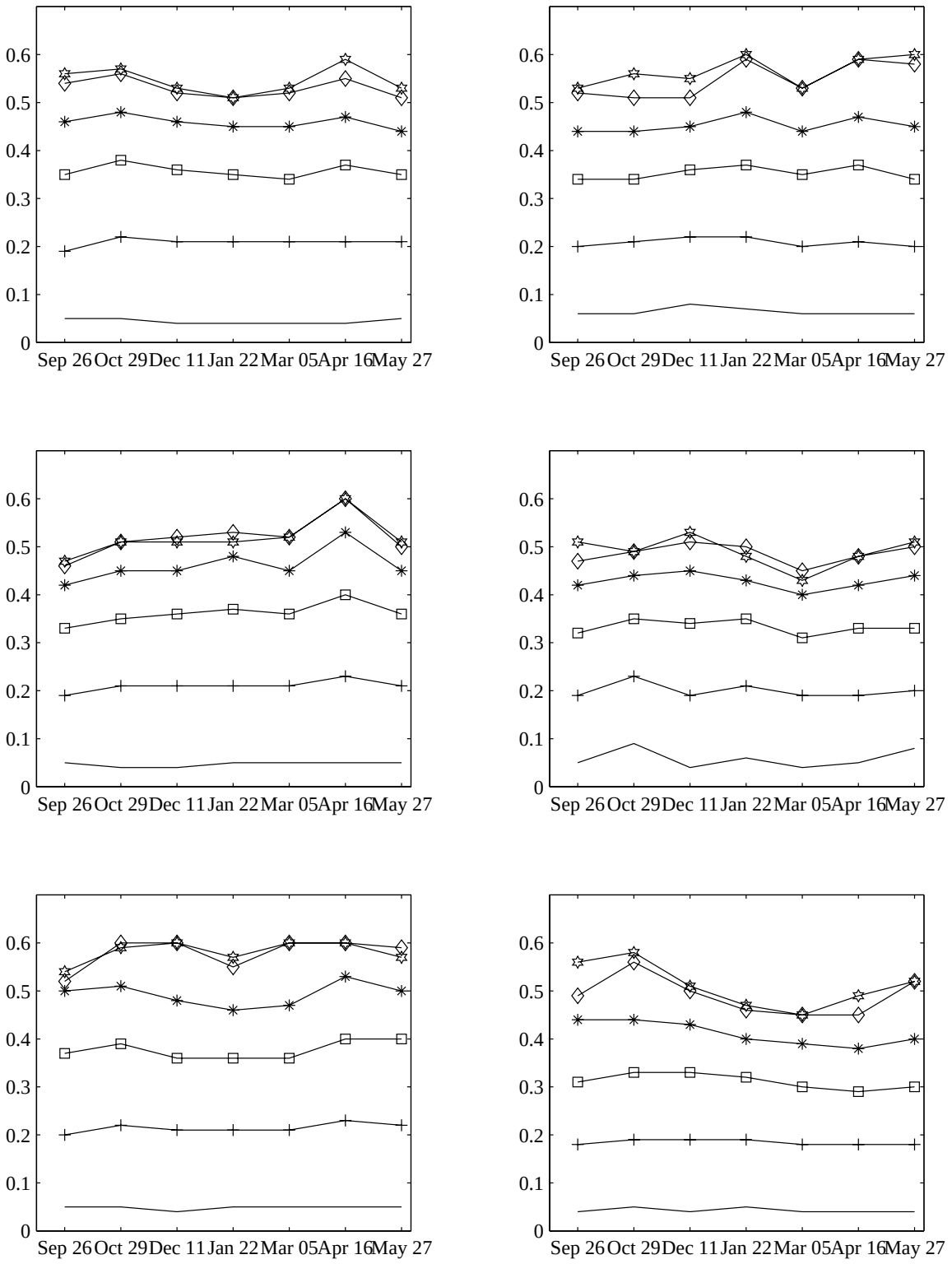


Figure C.26: *K_{PM} estimates for wind speeds of 9 m/s. Each line represents a separate incidence angle.*

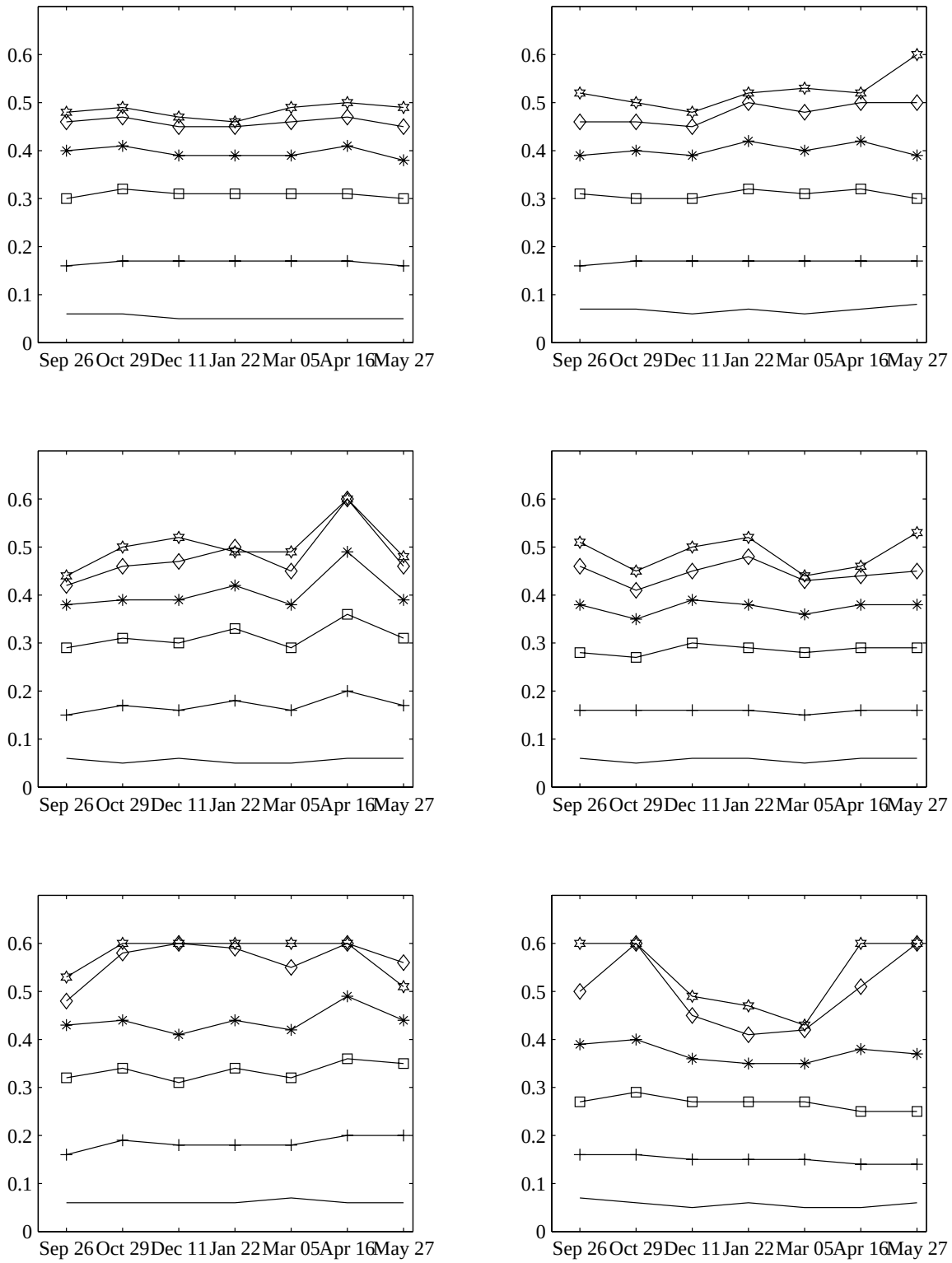


Figure C.27: *K_{PM} estimates for wind speeds of 11 m/s. Each line represents a separate incidence angle.*

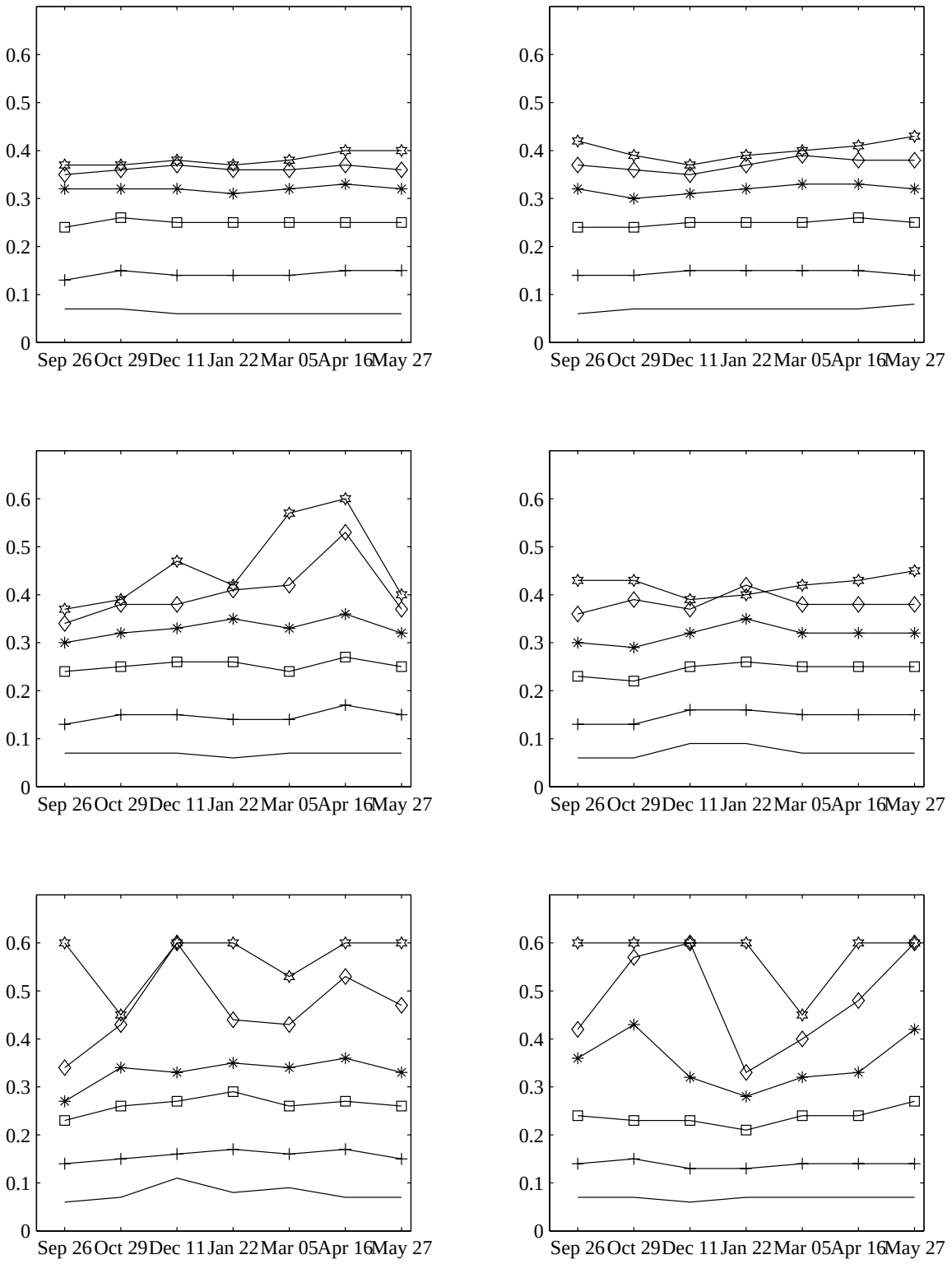


Figure C.28: K_{PM} *estimates for wind speeds of 13 m/s. Each line represents a separate incidence angle.*

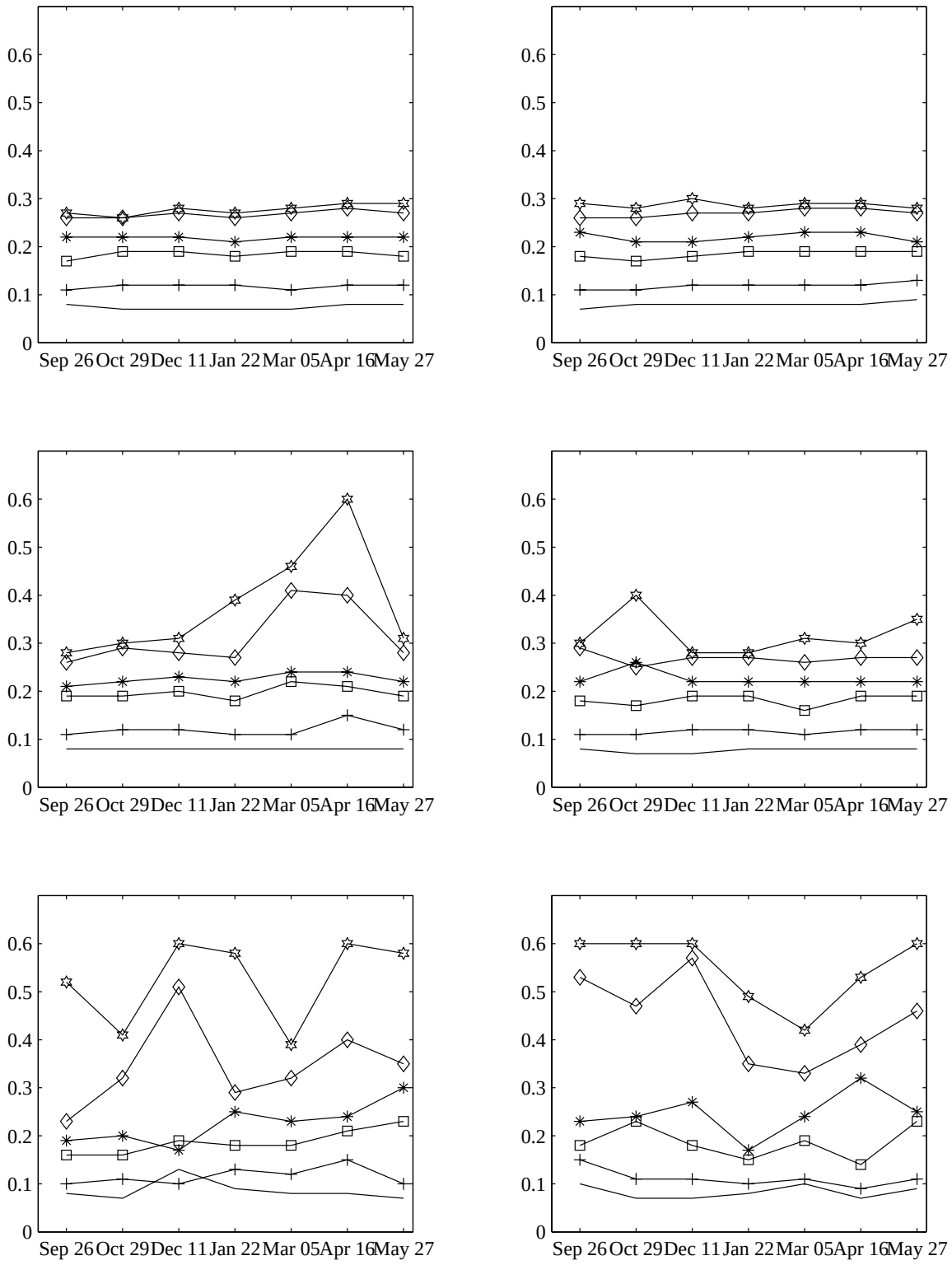


Figure C.29: K_{PM} *estimates for wind speeds of 15 m/s. Each line represents a separate incidence angle.*

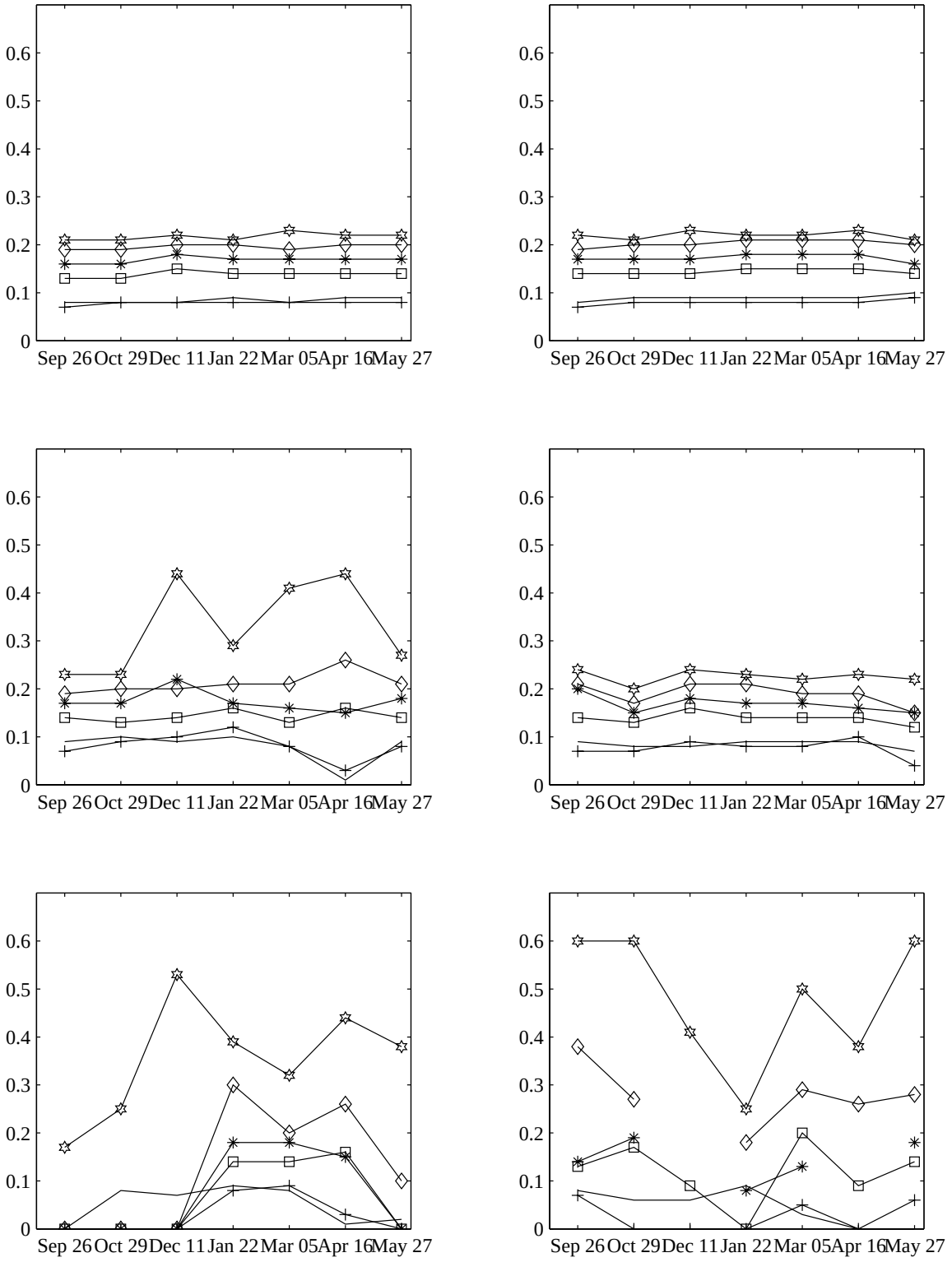


Figure C.30: *K_{PM} estimates for wind speeds of 17 m/s. Each line represents a separate incidence angle.*

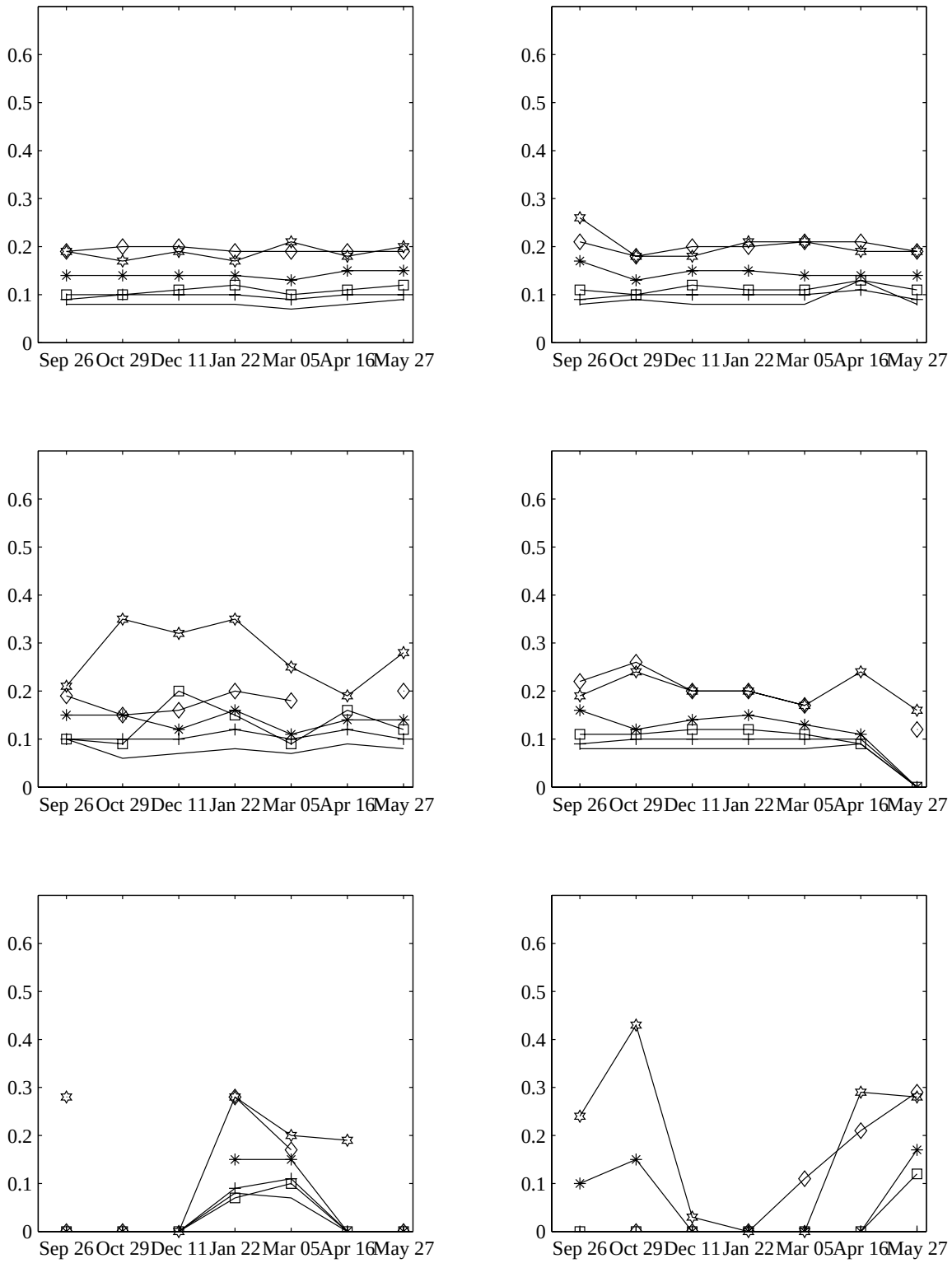


Figure C.31: K_{PM} *estimates for wind speeds of 19 m/s. Each line represents a separate incidence angle.*

Appendix D

Maximum Likelihood Estimation of the Wind

To estimate the wind, multiple measurements must be made of each wind cell with different measurement conditions and estimation theory employed to identify the wind most likely to have produced all of the measurements. This is displayed graphically in Fig. D.1.

Wind retrieval, as employed in NSCAT processing is based on the techniques of maximum likelihood estimation (MLE). The wind estimate, $\hat{w}$, is selected as the most probable wind, given the measurements; this can be inverted with Bayes' rule to be interpreted as the wind which maximized the probability of the measurements:

$$\hat{w} = \arg\max_{w} \; p(w|\vec{z}) \tag{D.1}$$

$$= \arg\max_{w} \; \frac{p(w)}{p(\vec{z})} p(\vec{z}|w). \tag{D.2}$$

The probability of the measurements, $p(\vec{z})$, does not change the maximization over the wind; without *a priori* knowledge, the probability of the wind, $p(w)$, is typically taken to be uniform, leaving just $p(\vec{z}|w)$. Because of the large amount of averaging involved in each measurement, the Central Limit Theorem is invoked to assume a Gaussian distribution for the measurements given the wind. It is further assumed, with strained credibility, that the measurements of a wind cell are independent; the communication noise (K_{PC}), is probably reasonably independent, but model variability (K_{PM}) probably introduces some correlation between the measurements. The pdf

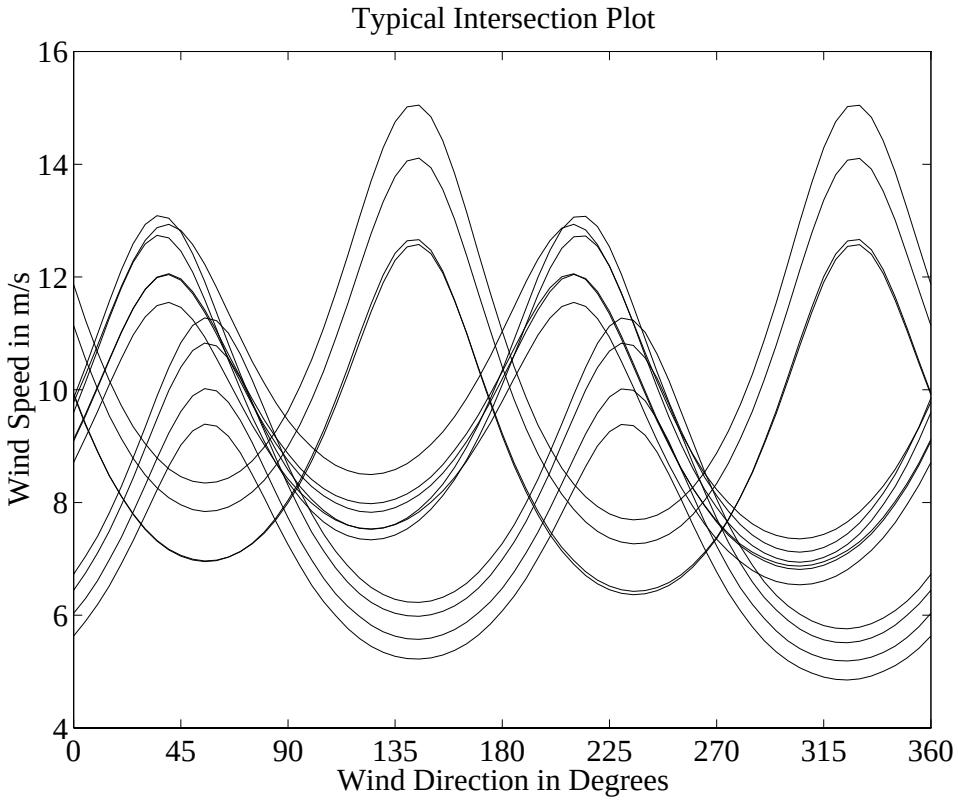


Figure D.1: *A single backscatter measurement produces a continuum of possible wind speeds and directions in a given cell. By making additional measurements of the same location, but with different measurement conditions, wind vectors can be identified as more likely to have caused all of the many measurements. In this case, the 13 measurements (one curve for each) were most likely caused by a wind approximately 8 m/s, with a relative azimuth angle of about 270 degrees. If there had been no noise in the measurements, (including no variability in the model function), the curves would have exactly intersected at a unique wind vector; measurement noise and model variability effectively bias particular measurement realizations. If there is enough noise in the measurements, other wind vectors (say 9 m/s and 90 degrees) can be erroneously retrieved as the wind.*

can then be expressed as

$$p(\vec{z}|w) = \prod_{k=1}^{K} \frac{1}{\delta_k \sqrt{2\pi}} \exp\left[\frac{-(z_k - \mathcal{M}_k)^2}{2\delta_k^2}\right] \tag{D.3}$$

where $\mathcal{M}_k$ is the empirical model function value based on the given wind and the kth measurement conditions, and δ_k^2 is the variance of the measurements—a simple evaluation of our assumed model yields

$$\delta_k^2 = K_{PM}^2 + K_{PC}^2 + K_{PM}^2 K_{PC}^2. \tag{D.4}$$

The independent Gaussian density is commonly employed in MLE because the natural log of the pdf can be equivalently maximized; defining the log-likelihood function, $L(w, z)$

$$\begin{aligned} L(w,z) &= \log p(\vec{z}|w) & \text{(D.5)} \\ &= -\sum_{k=1}^{K} \left\{ \frac{[(z_k - \mathcal{M}_k)^2]}{2\delta_k^2} + \frac{1}{2}\log\left[\delta_k^2\right] + \frac{1}{2}\log\left[2\pi\right] \right\} & \text{(D.6)} \end{aligned}$$

results in a simple function to be maximized. Obviously the final term, $\frac{1}{2}\log(2\pi)$, will not modify the maximization. The first term is the weighted-least squares estimate, an alternate estimation procedure.

Appendix E

Derivation of the Moment-Generating Function

Following [48], the moment-generating function for a quadratic form, $P = \mathbf{x}^T \Upsilon \mathbf{x}$, where $\mathbf{x}$ is a random vector and Υ is an arbitrary matrix, is:

$$M(t) = \mathrm{E}\left(e^{tP}\right) \tag{E.1}$$

$$= \int d\mathbf{x}\, e^{t\mathbf{x}^T\Upsilon\mathbf{x}} f(\mathbf{x}). \tag{E.2}$$

With the assumption that the data stream, $\mathbf{x}$, is normally distributed, $N(0, \mathbf{R})$, the moment-generating function becomes

$$M(t) = \int d\mathbf{x}\, e^{t\mathbf{x}^T\Upsilon\mathbf{x}} (2\pi)^{-\frac{N}{2}} |\mathbf{R}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{R}^{-1}\mathbf{x}\right\} \tag{E.3}$$

$$= \int d\mathbf{x}\ (2\pi)^{-\frac{N}{2}} |\mathbf{R}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{x}^T\left[\mathbf{R}^{-1} - t2\Upsilon\right]\mathbf{x}\right\} \tag{E.4}$$

$$= \int d\mathbf{x}\ (2\pi)^{-\frac{N}{2}} |\mathbf{R}|^{-\frac{1}{2}} \frac{|\mathbf{K}|^{-\frac{1}{2}}}{|\mathbf{K}|^{-\frac{1}{2}}} \exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{K}^{-1}\mathbf{x}\right\} \tag{E.5}$$

$$= |\mathbf{R}\mathbf{K}^{-1}|^{-\frac{1}{2}} \tag{E.6}$$

$$= |\mathbf{I} - t2\mathbf{R}\Upsilon)|^{-\frac{1}{2}} \tag{E.7}$$

$$= \prod_{n=1}^{N} (1 - t2\eta_n)^{-\frac{1}{2}} \tag{E.8}$$

$$= \prod_{n=1}^{D} (1 - t2\eta_n)^{-\frac{\nu_n}{2}} \tag{E.9}$$

where $\mathbf{K}^{-1} = \mathbf{I} - t2\mathbf{R}\Upsilon$ and the η_n are eigenvalues of $\mathbf{R}\Upsilon$. D indicates the number of *distinct* non-zero eigenvalues; the nth distinct eigenvalue has multiplicity ν_n. For the simple case of a *white* Gaussian vector, $\mathbf{x}$, the covariance is $\mathbf{R} = \sigma^2\mathbf{I}$; the eigenvalues of $\mathbf{R}\Upsilon$ are simply σ^2 times the eigenvalues of Υ.

The first two moments of the distribution are easily determined from the moment-generating function. Recognizing that $M(0) = 1$ (as is required for a valid moment-generating function), the first moment of P, the mean, is

$$\begin{aligned} m_1 &= M'(0) && \text{(E.10)} \\ &= \frac{d}{dt} \prod_{n=1}^{D} (1 - t2\eta_n)^{-\frac{\nu_n}{2}} \Bigg|_{t=0} && \text{(E.11)} \\ &= M(t) \sum_{n=1}^{D} \eta_n \nu_n \left(1 - t2\eta_n\right) \Bigg|_{t=0} && \text{(E.12)} \\ &= \sum_{n=1}^{D} \eta_n \nu_n && \text{(E.13)} \end{aligned}$$

and the second moment of P is

$$\begin{aligned} m_2 &= M''(0) && \text{(E.14)} \\ &= M(t) \sum_{n=1}^{D} \left(1 - t2\eta_n\right)^{-2} \eta_n^2 2\nu_n + M(t) \left(\sum_{n=1}^{D} \left(1 - t2\eta_n\right)^{-1} \eta_n \nu_n \right)^2 \Bigg|_{t=0} && \text{(E.15)} \\ &= \sum_{n=1}^{D} \eta_n^2 2\nu_n + \left(\sum_{n=1}^{D} \eta_n \nu_n \right)^2 && \text{(E.16)} \end{aligned}$$

so the variance of P is

$$\begin{aligned} \sigma_P^2 &= m_2 - m_1^2 && \text{(E.17)} \\ &= \sum_{n=1}^{D} \eta_n^2 2\nu_n. && \text{(E.18)} \end{aligned}$$

Higher order moments can be similarly computed.

Appendix F

The PDF from the Moment-Generating Function

The probability density function is the inverse Laplace transform of the moment-generating function with the negative of the argument; that is, the kernel of the moment-generating function is e^{tx}, while that of the Laplace transform is e^{-sx} [13]. In this section, the pdf is found from the moment-generating function in generality, and in the simpler case when all eigenvalue multiplicities are even.

Defining $s = -t$, the Laplace transform of the pdf of the random variable $P = \mathbf{x}^T \Upsilon \mathbf{x}$, which is a central quadratic form in $\mathbf{x}$,where $\mathbf{x}$ is distributed $\mathrm{N}(0, \mathbf{R})$, can be written as (with g and d_i defined in the text)

$$M(s) \;=\; \frac{g}{s^{\frac{J}{2}}} \prod_{i=1}^{D} \left(1 - \frac{d_i}{s}\right)^{-\frac{\nu_i}{2}}. \tag{F.1}$$

The inverse Laplace transform of Eq. (F.1) is [60],

$$f(y) = \frac{g}{, \left(\frac{J}{2}\right)} y^{\frac{J}{2}-1} \Phi_2 \left(\frac{\nu_1}{2}, \cdots, \frac{\nu_D}{2}; \frac{J}{2}; d_1 y, d_2 y, \cdots, d_D y\right). \tag{F.2}$$

The generalized hypergeometric function, Φ_2, is defined with a sum over D dimensional space:

$$\Phi_2 \left(b_1, \cdots, b_D; c; x_1, \cdots, x_D\right) = \sum \frac{(b_1)_{m_1} \cdots (b_D)_{m_D}}{(c)_{m_1+\cdots+m_D}\, m_1! \cdots m_D!} x_1^{m_1} \cdots x_D^{m_D} \tag{F.3}$$

where the summation Σ is a D-fold summation with m_1 through m_D each running from 0 to ∞, and the Pochhammer symbol is defined as

$$(a)_n = \frac{, (a+n)}{, (a)}. \tag{F.4}$$

While Eq. (F.2) provides a general solution, calculation of the generalized hypergeometric function is computationally restrictive. We therefore seek special cases to provide more practical solutions.

If all of the non-zero eigenvalues have an even number of multiplicities, the pdf can be developed as a partial fraction expansion [61], this approach was demonstrated by [62]. Defining $h_i = \frac{\nu_i}{2}$, Eq. (F.1) can be written as

$$M(s) = g \sum_{i=1}^{D} \sum_{j=1}^{h_i} \frac{A_{ij}}{(s-d_i)^j} \tag{F.5}$$

where

$$A_{ij} = \frac{1}{(h_i - j)!} \left[\frac{d^{h_i - j}}{ds^{h_i - j}} \left[(s-d_i)^{h_i} \frac{M(s)}{g} \right] \right] \Bigg|_{s=d_i}. \tag{F.6}$$

The derivatives required in A_{ij} can be be written as

$$D_i^{(m)}(s) = \frac{d^m}{ds^m} \prod_{k=1, k\neq i}^{D} (s-d_k)^{-h_k}, \qquad m \geq 0 \tag{F.7}$$

where the higher order derivatives can be computed recursively from $D_i^{(0)}(s)$

$$D_i^{(m+1)}(s) = \sum_{n=0}^{m} \begin{pmatrix} m \\ n \end{pmatrix} D_i^{(n)}(s) \sum_{k=1, k\neq i}^{D} h_k \frac{(m-n)!}{(d_k - s)^{m-n+1}}, \qquad m \geq 0. \tag{F.8}$$

Thus, the coefficient A_{ij} can be written as

$$A_{ij} = \frac{1}{h_i - j} \sum_{n=0}^{h_i - j - 1} \frac{D_i^{(n)}(d_i)}{n!} \sum_{k=1, k\neq i}^{D} \frac{h_k}{(d_k - d_i)^{h_i - j - n}}, \quad i = 1 \ldots D, j = 1 \ldots h_i. \tag{F.9}$$

Taking the inverse Laplace transform of Eq. (F.5), the pdf of a power estimate is

$$f_P(p) = g \sum_{i=1}^{D} \sum_{j=1}^{h_i} A_{ij} \frac{1}{(j-1)!} p^{j-1} e^{d_i p} U(p) \tag{F.10}$$

where $U(p)$ is the unit step function.

Appendix G

Useful Facts and Properties of the Gamma Distribution

The probability density function of a gamma distributed random variable, G, is

$$f_G(g) = \frac{1}{\beta^\alpha, (\alpha)} g^{\alpha-1} e^{-g/\beta} U(g) \tag{G.1}$$

with $\alpha \geq 0$ and $\beta > 0$. The mean and variance of the Gamma distribution are, respectively:

$$\mu = \alpha\beta \tag{G.2}$$

$$\sigma^2 = \alpha\beta^2. \tag{G.3}$$

These can be determined from the moment-generating function,

$$M_G(t) = \mathrm{E}\left(e^{tg}\right) \tag{G.4}$$

$$= (1 - \beta t)^{-\alpha} \tag{G.5}$$

Some special cases of the Gamma distribution are known by other names:

$\alpha = 1$	Exponential distribution
$\alpha = \frac{\nu}{2}$, $\beta = 2$	Chi-square distribution, with ν degrees of freedom
α is an integer	Erlang

G.1 A Gamma RV by Squaring a Zero-Mean Gaussian RV

One common way to generate a Gamma rv is to square a Gaussian rv. To find the distribution of a function of a random variable, $y = g(x)$, we have

$$\sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|} \tag{G.6}$$

where we use the solutions $x_i = g^{-1}(y)$. Applying this technique to the special case of $y = x^2$ where x is distributed $\mathrm{N}(0, \sigma^2)$ yields no solutions for $y < 0$ and two solutions for $y > 0$, $x_1 = \sqrt{y}$ and $x_2 = -\sqrt{y}$:

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right] U(y) && \text{(G.7)} \\ &= \frac{1}{\sqrt{y}} f_X(\sqrt{y}) U(y) && \text{(G.8)} \\ &= \frac{1}{\sqrt{2\pi y}\sigma} e^{-\frac{y}{2\sigma^2}} U(y) && \text{(G.9)} \\ &= \frac{1}{\beta^\alpha,\ (\alpha)} y^{\alpha-1} e^{-y/\beta} U(y) && \text{(G.10)} \end{aligned}$$

where $\alpha = 1/2$ and $\beta = 2\sigma^2$. So we have a Gamma distribution (or a Chi-square with one degree of freedom).

G.2 Scaling a Gamma RV Yields a Gamma RV

Scaling a Gamma rv by a constant yields a new Gamma with the same α parameter, but β scaled by the same constant. Let $S = aG$, where G is Gamma distributed then

$$\begin{aligned} f_S(s) &= \frac{1}{|a|} f_G(s/a) && \text{(G.11)} \\ &= \frac{1}{|a|\beta^\alpha,\ (\alpha)} \left(\frac{s}{a}\right)^{\alpha-1} e^{-s/(a\beta)} U(s/a) && \text{(G.12)} \\ &= \frac{1}{(a\beta)^\alpha,\ (\alpha)} (s)^{\alpha-1} e^{-s/(a\beta)} U(s) && \text{(G.13)} \end{aligned}$$

G.3 The Sum of Independent Gamma RV's

What is the pdf of the sum of Gamma rv's, $Z = X + Y$? Here, assume that X and Y are independent, so $f_Z(z)$ is simply the convolution of two Gamma distributions. Convolution of distributions is equivalent to multiplication of moment generating functions:

$$\begin{aligned} M_Z(t) &= M_X(t) M_Y(t) && \text{(G.14)} \\ &= (1 - \beta_X)^{-\alpha_X} (1 - \beta_Y)^{-\alpha_Y} && \text{(G.15)} \end{aligned}$$

If X and Y have the same β parameter, then $M_Z(t) = (1-\beta)^{-(\alpha_X+\alpha_Y)}$, so Z is distributed Gamma with $\alpha_Z = \alpha_X + \alpha_Y$ and $\beta_Z = \beta_X = \beta_Y = \beta$. The pdf for the sum of Gamma rv's with $\beta_X \neq \beta_Y$, but identical α can be computed with a convolution, yielding a rather messy result involving Bessel functions:

$$f_Z(z) = \frac{e^{\frac{-z}{2\,\beta_1} - \frac{z}{2\,\beta_2}}\; z^{-1+3\,\alpha} \left(\frac{z}{\beta_1} - \frac{z}{\beta_2}\right)^{\frac{1}{2}-\alpha} \sqrt{\pi}\, \mathrm{I}_{-\frac{1}{2}+\alpha}\left(\frac{z}{2\,\beta_1} - \frac{z}{2\,\beta_2}\right)}{(\beta_1\,\beta_2)^{\alpha}\,,\,(\alpha)}. \tag{G.16}$$

G.4 Kullback-Leibler Distance Between a Gamma Distribution and a Gaussian Distribution

Because the mean and variance of a periodogram estimate are easily computed, the distribution of the estimate is typically assumed to be Gaussian, with this mean and variance. Here we have shown that for several cases the distribution is actually Gamma. How far is this distribution from the Gaussian typically assumed? A standard measure is the Kullback Leibler distance, or relative entropy, where the distance between two densities is defined as [56, p. 231]

$$D\,(f\|g) = \int f(x) \log \frac{f(x)}{g(x)}\, dx. \tag{G.17}$$

While not a true metric, relative entropy has the property of non-negativity, $D(f\|g) \geq 0$, with equality if and only if the distributions are identical, $f = g$.

Define $f(x)$ to be the Gamma distribution,

$$f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha},\,(\alpha)} e^{-\frac{x}{\beta}}\; U(x) \tag{G.18}$$

and $g(x)$ to be the Gaussian distribution, with mean, $\alpha\beta$ and variance $\alpha\beta^2$ (identical to the Gamma distribution),

$$g(x) = \frac{1}{\sqrt{2\pi\alpha\beta^2}} \exp\left[-\frac{(x-\alpha\beta)^2}{2\alpha\beta^2}\right]. \tag{G.19}$$

The natural logarithm of the ratio of these distributions is

$$\log\left(\frac{f(x)}{g(x)}\right) \;=\; (\alpha-1)\log\left(\frac{x}{\beta}\right) + \log\left(\frac{\sqrt{2\pi\alpha}}{,\,(\alpha)}\right) + \frac{x^2}{2\alpha\beta^2} - \frac{2x}{\beta} + \frac{\alpha}{2}. \tag{G.20}$$

The relative entropy, D, computed as the integral of this expression multiplied by $f(x)$, can be thought of as an expectation with respect to $f(x)$:

$$D\left(f\|g\right) = \mathrm{E}\log\left(\frac{f(x)}{g(x)}\right) \tag{G.21}$$

$$= (\alpha-1)\,\mathrm{E}\log\left(\frac{x}{\beta}\right) + \log\left(\frac{\sqrt{2\pi\alpha}}{, (\alpha)}\right) + \frac{1}{2} - \alpha. \tag{G.22}$$

The expectation of the log of x/β is

$$\mathrm{E}\log\left(\frac{x}{\beta}\right) = \int_0^\infty f(x)\log\left(\frac{x}{\beta}\right)\,dx \tag{G.23}$$

$$= \int_0^\infty \frac{x^{\alpha-1}}{\beta^\alpha, (\alpha)} e^{-\frac{x}{\beta}} \log\frac{x}{\beta}\,dx \tag{G.24}$$

$$= \frac{1}{, (\alpha)}\int_0^\infty u^{\alpha-1}e^{-u}\log u\,du \tag{G.25}$$

$$= \psi(\alpha) \tag{G.26}$$

$$= \frac{, '(\alpha)}{, (\alpha)} \tag{G.27}$$

$$= -\gamma + \sum_{j=1}^{\alpha-1}\frac{1}{j} \tag{G.28}$$

where $\psi(\alpha)$ is the digamma function (the derivative of the log of the Gamma function), which can be expressed for integer arguments, α, in terms of the Euler gamma constant, $\gamma \approx 0.57721$ and a finite sum (see Gradsteyn and Ryzhik, p. 604). Thus, the relative entropy between a Gamma distribution and a Gaussian distribution is

$$D\left(f\|g\right) = (\alpha-1)\left(-\gamma + \sum_{j=1}^{\alpha-1}\frac{1}{j}\right) + \log\left(\frac{\sqrt{2\pi\alpha}}{, (\alpha)}\right) + \frac{1}{2} - \alpha. \tag{G.29}$$

Appendix H

Binary Classification of Wind Fields through Hypothesis Testing on Scatterometer Measurements

H.1 Abstract

[1]*Scatterometers are radars specially designed to near-surface wind over the ocean from space. Traditional scatterometer wind estimation inverts the model function relationship between the wind and backscatter at each resolution element, yielding a set of ambiguities due to the many-to-one mapping of the model function. Field-wise wind estimation dramatically reduces the number of ambiguities by estimating the wind at many resolution elements, simultaneously, using a wind model that constrains the spatial variability of the wind. However, the appropriate choice of the model order needed for a particular wind field is not known* a priori. *The approximate model order is valuable because of the implicit trade-off between the computational complexity of high-order models and the imprecise model fit of low-order models. In this paper, a simple binary classification of wind fields is proposed which identifies whether or not a region will be well modeled by a low-order wind model. The raw scatterometer measurements provide data about the wind that can be exploited through hypothesis testing to identify the appropriate model order to use in field-wise wind estimation. Improved processing algorithms lead to better use of the data.*

[1] *Published and presented [63]*

H.2 Introduction

Naval radar operators during World War II observed considerably more noise in their radar returns during stormy weather; with this simple beginning, scatterometry was born [1]. Scatterometers are high frequency radars designed to infer the physical state of a system based on measuring the backscatter from that system. In particular, the last 20 years have seen the use of several space borne scatterometers to estimate near-surface ocean winds with considerable success [29]. The estimation procedure is not unique; that is, several wind vectors (as many as six) are typically found that could have produced the measurements.

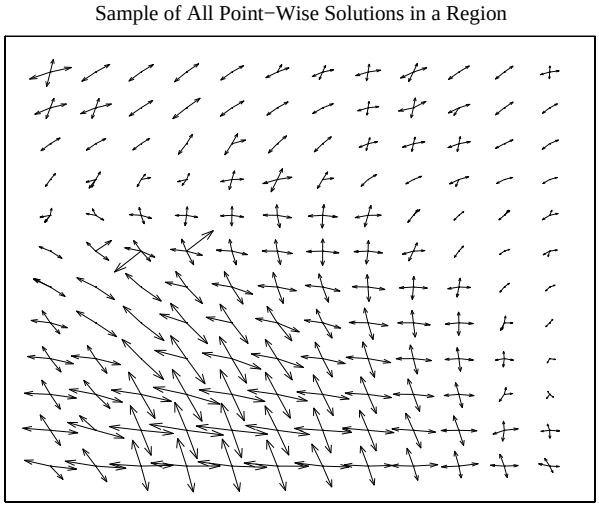


Figure H.1: *Constructing a unique wind field from the multiple point-wise estimates is a daunting task. Each resolution element can have as many as six point-wise estimates. Determining the optimal field for a 12 by 12 region would require comparing as many as* 6^{144} *fields.*

For example, Fig. H.1 displays all of the possible wind vectors based on point-wise wind estimation, throughout a region; correctly identifying a unique wind field from the many possible combinations is not a well defined process, involving considerable computational resources and often prone to error. To ameliorate the

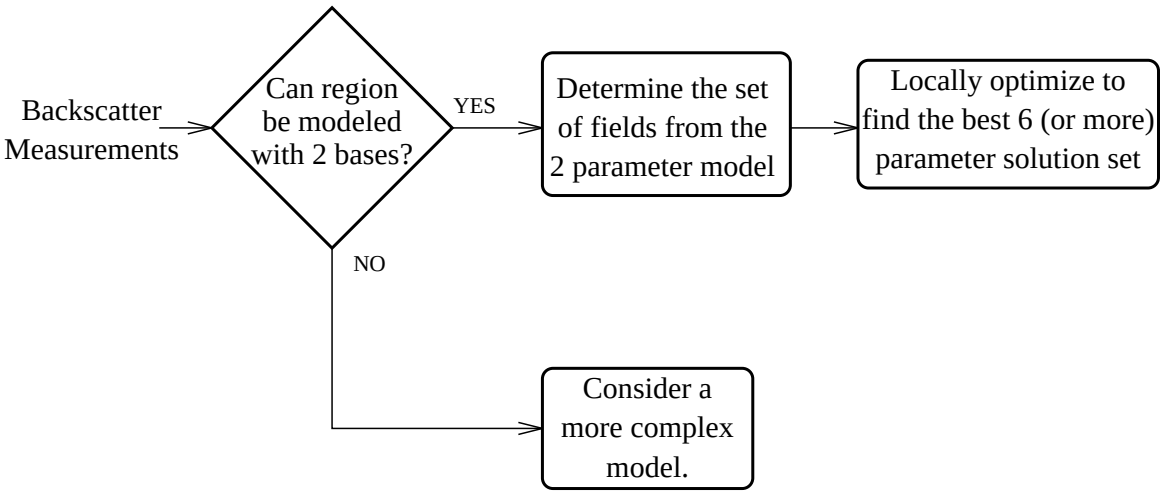


Figure H.2: *Flow diagram for hypothesis testing. A statistic of the backscatter measurements for a region is computed, and a hypothesis test performed on this statistic. If the region is identified by the test as likely to be well modeled by a mean wind field, a set of fields is determined by globally optimizing that two-dimensional space; these solutions serve as initial values in a local optimization in a higher-dimensional space to more accurately estimate the wind field. If the hypothesis test reveals that the field will not be (probably) adequately fit by two bases, more work will need to be done. Experience suggests that a slight majority of wind fields are adequately fit by a mean wind field; those that are not can be used to develop models specifically designed for more difficult wind fields. Such models could include low wind speed models and nonlinear models for fronts and cyclones. Additional hypothesis tests could be cascaded after this one: if the region is not fit by two bases, is it a low wind speed region or a front or a cyclone?*

problem of selecting unique wind vectors, field-wise estimation has been introduced in which an assumed model for the spatial correlation of wind vectors constrains the possible estimates [46, 64]. Flexible models which span a wide range of wind fields require many parameters—searching a high-dimensional space for wind field estimation is computationally prohibitive [65]. On the other hand, models with only a few parameters are easily searched, but represent a limited range of wind fields [30]. In this paper a simple algorithm is examined to identify, directly from the backscatter measurements, whether or not a particular region can be well modeled by a simple,

low-order model. This approach decreases the average number of model parameters without significantly increasing the average modeling error.

Due to the nature of the wind estimation objective function, simple gradient search techniques can not guarantee convergence to the global optimum; the wind model vector space must be searched exhaustively. While exhaustively searching even a 6 dimensional vector space is prohibitive, a two dimensional search is fairly straightforward. An exhaustive search involves selecting a two dimensional (i.e., mean) wind field and computing the probability of measuring the actual scatterometer measurements given that wind field and identifying the wind field (or set of fields) that maximizes that probability. Figure H.2 indicates a simple block structure of a hypothesis test to identify from the measurements whether or not a field is well modeled by a mean wind field. If a field is well modeled, there is no need to waste additional resources trying to estimates higher order parameters; a mean wind field can be estimated, and if a better estimate is required, a local optimization in a higher dimensional space can be performed. If the field is not well modeled by a mean wind field, then substantially more work needs to be done with this field—perhaps a six parameter model, or a low-order non-linear model tuned to specific wind phenomena would work. The methodology presented here could easily be extended to additional hypotheses to continue reducing the set of difficult wind fields.

In the next section a statistic on the measurement is identified which has a strong correlation to the error of the model fit; it is this statistic that can be used in a hypothesis test. Hypothesis testing is then briefly described and specifically applied to wind field classes and the backscatter statistic. Finally, some conclusions are presented to put this work in perspective.

H.3 Statistics on the Measurements

Initial examination of the backscatter field over a region reveals little relationship to the underlying wind field. Figure H.3 shows a gray scale image of the average backscatter measurements for the region displayed in Fig. H.1. While there is some correspondence between the magnitude of the backscatter and the wind speed,

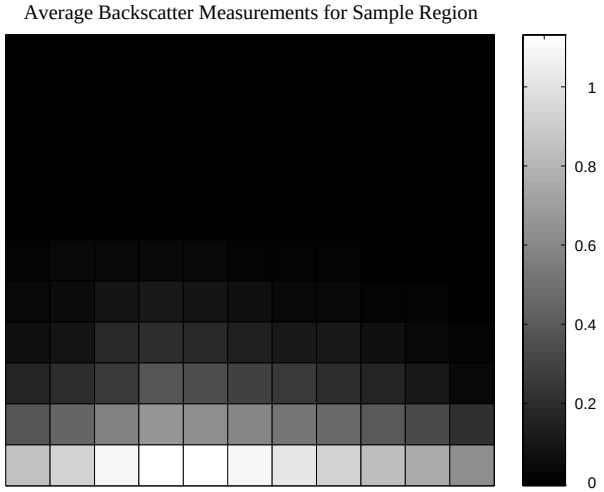


Figure H.3: *The average backscatter measurements for the wind field region displayed in Fig. H.1. While there is some relationship between the backscatter and the wind speed and complexity, the relationship becomes unclear for cells along the top because the backscatter is dramatically reduced with the increased incidence angle.*

the cells at the top of the image (which have much larger incidence angles) have negligible backscatter values, because of the strong incidence angle dependence of the measurements [32]. In order to observe the variability of the backscatter caused by wind variations, the incidence angle dependence must be removed. I selected over 2000 regions from NSCAT (NASA Scatterometer) data in which a two parameter model (the mean wind field) fit the field selected by Jet Propulsion Laboratories (JPL) very well, according to the normalized-vector RMS error. Each of these regions had fairly constant along track backscatter values and a strong cross-track dependence for the backscatter. Averaging over all the regions, and over the along-track cells to yield the cross-track dependence of the backscatter for each beam (the beams have different relative azimuths and possibly different calibration errors) provides the average backscatter for each beam as a function of the cross track cell. The results are shown in Fig. H.4 where the curves show the cross-track dependence of the backscatter for each beam. Now we can examine the normalized backscatter of a region, where we normalize by dividing each backscatter measurement by the averages described by

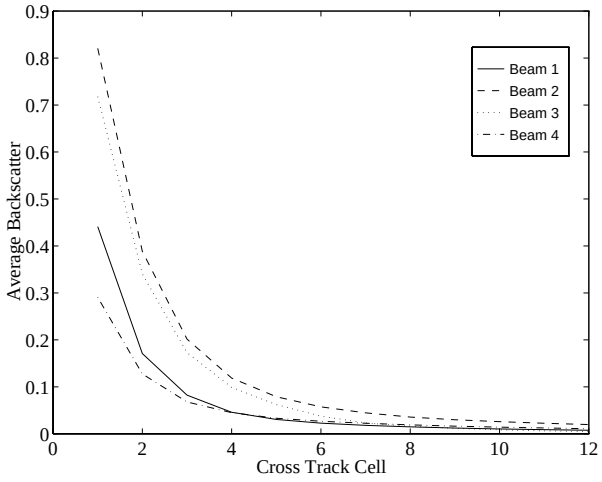


Figure H.4: *The average backscatter for very smooth wind fields, i.e., those which a two parameter model fits the JPL nudged wind with less than 0.1 NRMS error.*

Fig. H.4. Any patterns or variations in the normalized backscatter should then be due to the underlying wind pattern.

The statistics of the normalized backscatter seem to have some relationship with the quality of fit of the wind model to the wind field. Examining more than 5000 regions, the RMS of the standard deviation of the normalized backscatter from each beam is highly correlated with the VRMS (Vector RMS) error between the mean wind field estimate and the JPL estimate of the wind; this standard deviation is selected as the statistic for use in the hypothesis test. Figure H.5 shows the relationship between the statistic and the VRMS error for the 5000 wind fields. By setting a VRMS error criterion, the field can be classified as *good* or *bad* based on whether the field exceeds the criterion for quality.

H.4 Hypothesis Testing

Wind field classification algorithms can be used to select models with a minimal number of parameters while keeping the error within an acceptable range. The result increases the computational efficiency of field-wise estimation without significantly increasing the modeling error. In this section a simple classification

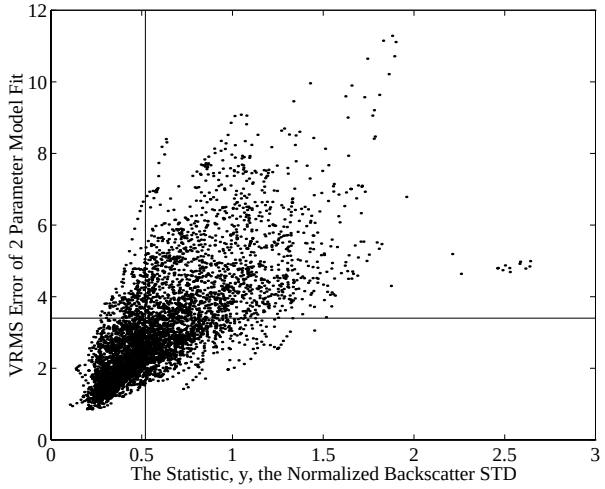


Figure H.5: *In general, the fields that are poorly fit by a mean wind field (these have a high VRMS error) have a large value of the measurement statistic (the std of the normalized backscatter).*

algorithm is described which tests the hypothesis that a field is poorly modeled by a low-order model—specifically, by a mean wind field. Referring to Fig. H.2, a VRMS error threshold is selected to identify fields with error greater than the threshold as a *bad* field (since it is poorly fit with two bases) and fields with less error as *good*. Either the region is adequately modeled by a low-order model (designated good, θ_0), or it is poorly modeled by the low-order model (bad, θ_1). Comparing the statistic, y, to a threshold, ν, provides the basis for the binary hypothesis test [48]:

$$\text{Wind Class Declaration} = \begin{cases} \theta_1 & \text{if} \quad y > \nu \\ \theta_0 & \text{if} \quad y \leq \nu. \end{cases}$$

The statistic, y, is defined as the standard deviation of the σ^0 values of all the beams normalized by the average backscatter values to remove the cross track dependence.

The choice of a threshold for the VRMS error of model fit identifies a field as being either well (θ_0) or poorly (θ_1) modeled by a mean wind field. The definition of "well" modeled, and the choice of the threshold, depends on the particular application. The horizontal line in Fig. H.5 illustrates, for a given application, the

separation of the wind fields into two classes—θ_0 if the VRMS error is below the line and θ_1 if the error is above the line. Having identified the fields as θ_0 or θ_1, the empirical probability density functions of good and bad wind fields can be computed as functions of the statistic. Figure H.6 plots these densities for a few values of the VRMS error threshold.

The density functions of good and bad wind fields displayed in Fig. H.6 provide the probabilistic measures necessary for a hypothesis test. By setting a threshold on the measurement statistic, we can test the hypothesis that a wind field will be bad. If our measurement is less than the threshold, the field is declared good, θ_0; if the measurement is above the threshold, the field is declared bad, θ_1. The vertical line in Fig. H.5 illustrates this point by identifying a threshold on the statistic; the two classes of wind fields (separated by the horizontal line) and the two classes of declared wind fields (distinguished by the vertical line) define four distinct regions in the figure, which can be characterized through two numbers: the probability of false alarm and the probability of detection.

The probability of false alarm is the probability that we incorrectly identify a good field as bad; this probability is computed as the area under the pdf of good fields above the threshold. The probability of detection is the probability that we correctly identify a bad field as bad; this is computed as the area under the pdf of bad fields above the threshold. Optimally we want a low probability of false alarm and a high probability of detection; adjustment of the hypothesis threshold requires a trade-off between these quality measures.

In this case, the probability of detection is critical. If we miss detection of a bad wind field, we will try to use the two-parameter model on a field that contains more features than a simple mean flow. On the other hand, the probability of false alarm is not so important. If we classify a good wind field as bad, we will look at it more closely and use a more involved model—this more complicated model will work just fine and the wind will be estimated with a little more trouble. Of course if we set our threshold too low, we classify everything as bad and don't gain any savings in computation from the classification. Figure H.7 displays characteristic curves for our

four VRMS thresholds that identify the quality of the fit by displaying the probability of detection against the probability of false alarm as the statistic threshold is adjusted.

If, for example, the wind field class θ_0 is defined as wind fields that have a 2 parameter model fit VRMS error less than 3.4 m/s (below the horizontal line of Fig. H.5), and with the choice of $\nu = 0.52$ (selected to declare half the wind fields in θ_0 and half in θ_1), the probability of correctly classifying a θ_1 wind field (probability of detection) is 86% and the probability of incorrectly classifying a θ_0 wind field (probability of false alarm) is 32%. With these thresholds (rather arbitrarily chosen) 50% of the wind fields are declared to be well modeled by just 2 parameters—in fact, the average VRMS error of these fits is 2.2 m/s. For comparison, if two parameters had been used for all the regions, the average VRMS error would have been 3.2 m/s, and if 40 parameters had been used, the average VRMS error would have been 1.2 m/s. Thus for a moderate increase in modeling error, the number of required parameters was reduced from 40 to two in half the regions—with a significant computational saving.

H.5 Conclusions

Field-wise wind estimation profoundly reduces the number of ambiguities and reduces the computational load of scatterometer wind estimation. Increasing the number of model parameters increases modeling accuracy; however, it also increases computational expense. Classification algorithms, such as that presented here, can be used to decrease the average number of model parameters without significantly increasing the average modeling error. Identifying, *a priori*, fields that will be well modeled by a low-order model conserves computing resources for more difficult fields. Further, fields that are classified as poorly modeled by a mean wind field can be used to develop improved models specific to certain features like fronts and cyclones. Low-order models can be developed for these cases without the need of using the generic model which would require many parameters to model unusual fields.

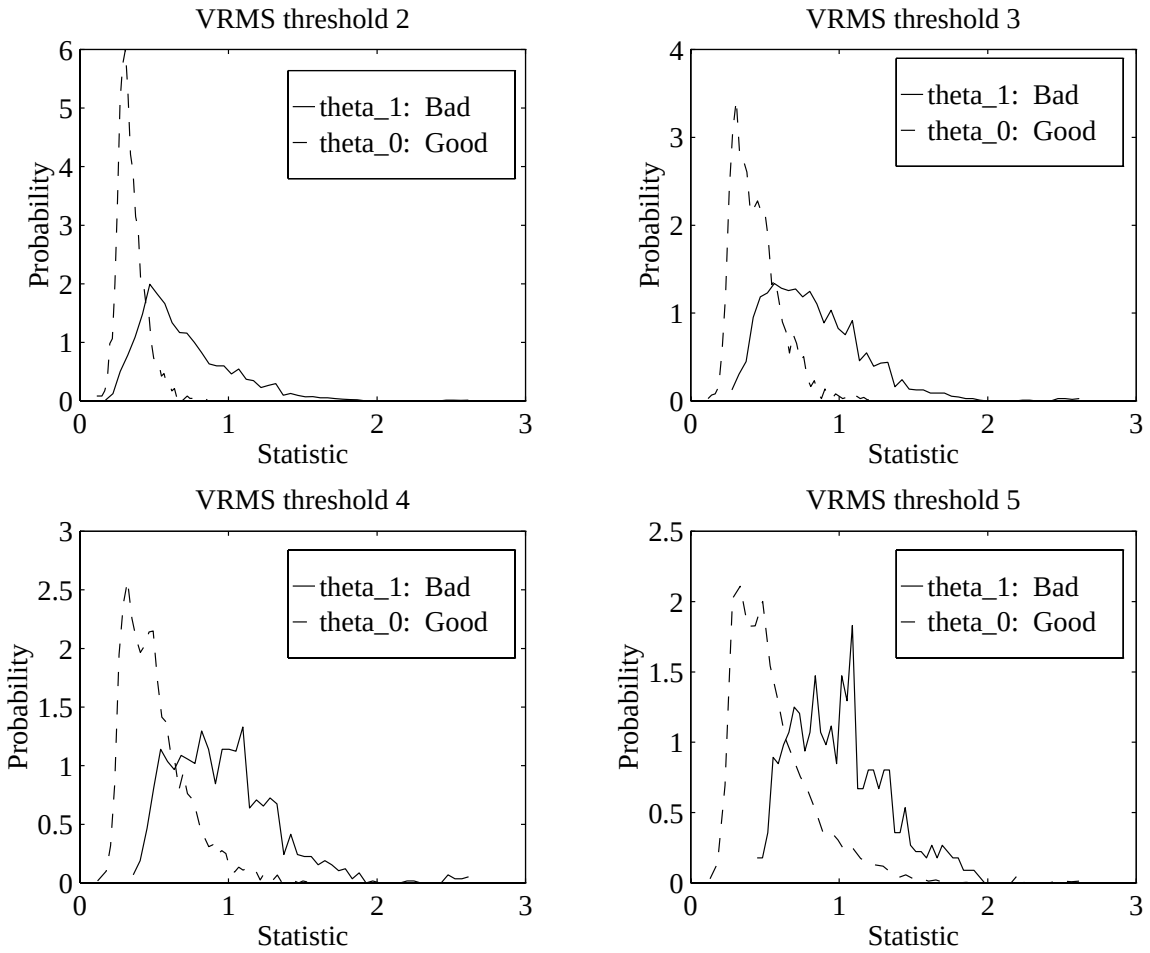


Figure H.6: *Empirically derived probability density functions of good and bad wind fields as functions of the statistic of the scatterometer measurements. As the VRMS error threshold is increased, the empirical density function becomes more erratic because there is a much smaller data set with which to estimate the density.*

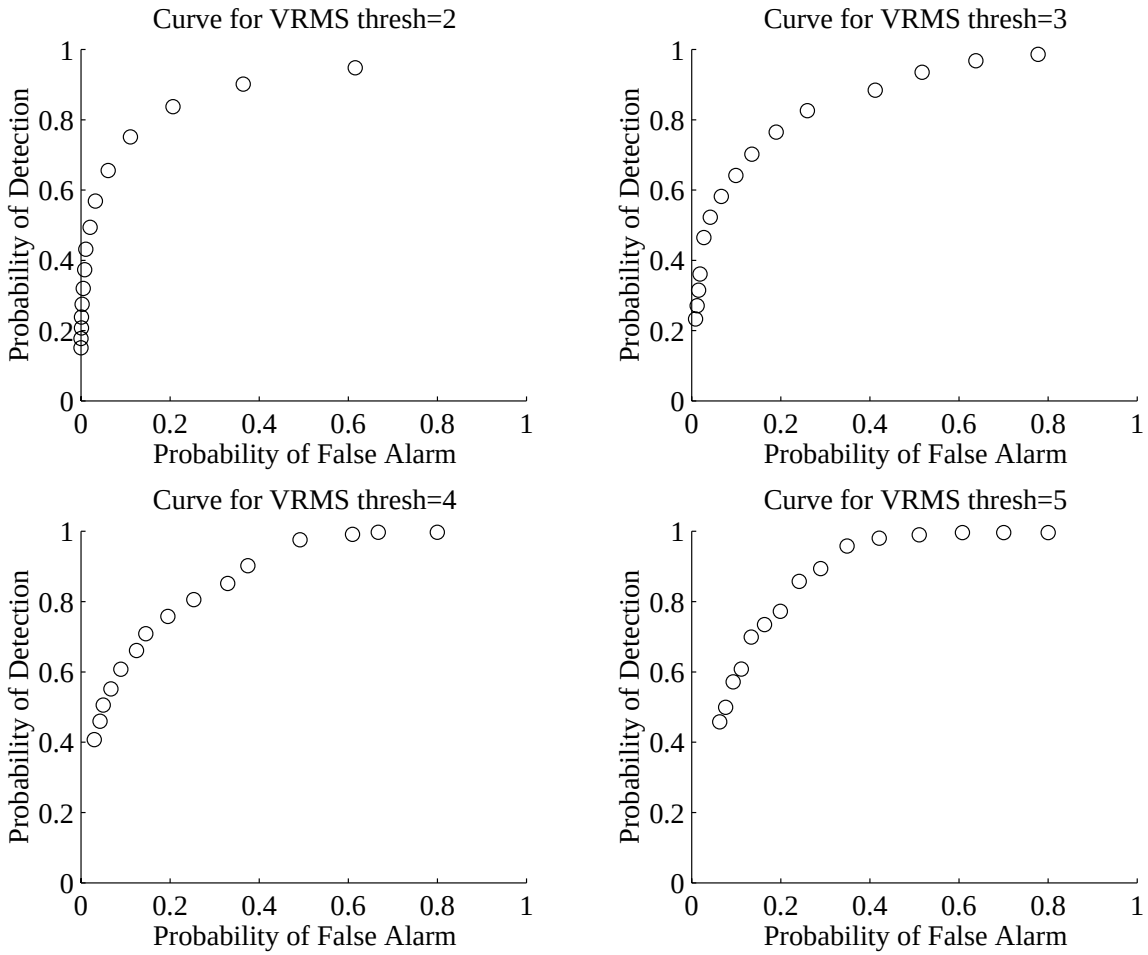


Figure H.7: *Characteristic curves based on adjusting the threshold on the measurement statistic (for each of four VRMS error thresholds identifying the quality of the fit). The goal is to choose a threshold on the statistic to yield a low probability of false alarm while maintaining a high probability of detection.*

Appendix I

Glossary of Mathematical Symbols

k, k_1, k_2	Frequency bins of the spectral estimate.
p	Power parameter of the ocean surface spectrum, p. 12.
z	Measured value of the backscatter, corrupted by noise.
A_{ij}	Coefficients used in the partial fraction expansion.
C, $C(\rho, \phi)$, $C(x, y)$	The correlation function in two-dimensional real space. The inverse Fourier Transform of the surface power spectrum.
$I_0(x)$	The zeroth order modified Bessel function.
K_{PM}	The normalized standard deviation of the model function, normalized by the mean value.
NRCS	The normalized radar cross section, equivalently expressed by the symbol for the backscatter coefficient, σ^o (see below), p. 2.
NSCAT	The NASA scatterometer. Developed at the Jet Propulsion Laboratories and launched in August, 1996, the satellite power supply failed in June 1997.
$\mathbf{R}$	The autocorrelation matrix of a data vector.
S, $S(k, \theta)$	The surface power spectrum, in two-dimensional wavenumber space. Computed as the Fourier Transform of the correlation function, C.

U Wind speed.

γ The incomplete Gamma function.

μ Standard normal random variable.

ν Standard normal random variable.

σ^o The backscatter coefficient, see also NRCS, pp. 2, 14.

σ^o_M The NRCS value predicted by the empirical model function, given the wind and measurement conditions.

σ^o_T The true value of the NRCS as developed by the geophysical parameters.

Υ Central matrix of the quadratic form yielding the power estimate of a band of frequencies based on Welch's modified periodogram averaging for a data vector, $\mathbf{x}$.

Φ_2 The generalized hypergeometric function.

R Range from radar to scattering target (p. 2).

L Losses in the radar system. (p. 2).

P_s Power scattered from the surface. (p. 2).

P_t Power transmitted by the radar (p. 2).

G Gain of the radar system (p. 2).

λ 1: Electromagnetic wavelength (p. 2). 2: Surface roughness, the large parameter in asymptotic analysis defined as $4h^2\kappa^2$ (p. 14).

A Area of target (p. 2).

$\vec{\kappa}$ Electromagnetic wavenumber, $2\pi/\lambda$, (p. 14).

κ Magnitude of the electromagnetic wavenumber (p. 14).

θ_i Incidence angle, measured from normal (p. 14).

$\mathbf{k}_b$ Bragg wavenumber (p. 14).

$\vec{\kappa}_H$ Horizontal component of the electromagnetic wavenumber (p. 14).

h Height of the ocean surface defined as the standard deviation of the surface (p. 14).

ρ_c The radius of curvature of the ocean surface (p. 13).

$K(k)$ The wavenumber magnitude dependence of the separable version of the ocean power spectrum (see also $\Theta(\theta)$) (p. 17).

$\Theta(\theta)$ The angular dependence of the separable version of the ocean power spectrum (see also $K(k)$) (p. 17).

k_0 Low wavenumber limit (or peak) of the ocean surface spectrum (p. 18).

q Directionality parameter of a separable ocean sectrum, $\cos^{2q}(\theta)$, see also q (p. 18).

θ_0 Directionality parameter of a separable ocean spectrum with hard limits at $\pm\theta_0$, see also $\Theta(\theta)$, (p. 21).

k_x, k_y Components of the ocean surface wavenumbers (p. 20).

S_0 Normalizing constant for the ocean spectrum (p. 18).

, (x) The Gamma function.

ρ Polar magnitude coordinate in real space (p. 20).

ϕ Polar angular coordinate in real space (p. 20).

$T(\nu, \phi)$ Function containing the angular dependence of the surface correlation function (p. 21).

a Coefficient in the approximation for $T(\nu, \phi)$ (p. 22).

b Coefficient in the approximation for $T(\nu, \phi)$ (p. 22).

ϵ Coefficient in the expression for the dominant terms of the correlation function (p. 22).

A Coefficient in the expression for the dominant terms of the correlation function (p. 22).

λ_0 Roughness limit describing the transition point between two types of asymptotic behaviors caused by the finite radar footprint (p. 25).

C_{dom} The dominant terms of the correlation function (p. 22).

χ Relative azimuth angle between the wind direction and radar azimuth angle (p. 39).

A_0, A_1, A_2 Coefficients in empirical model functions relating the wind to the backscatter (p. 39).

z Bacscatter measurement (p. 43).

K_{PC} Normalized standard deviation of the communication noise (p. 43).

A, B, C Coefficients in the computation of K_{PC} (p. 44).

SNR Signal to Noise Ratio (p. 44).

SV Sample variance (p. 50).

SM Sample mean (p. 50).

$x[j]$, $x[i,n]$ Data sequence, (p. 68).

L Length of data segment (p. 68).

K Number of data segments (p. 68).

r Amount of overlap in Welch's spectral estimation technique (p. 68).

$w[n]$ Data window in time domain (p. 68).

$W[k]$ Data window in frequency domain (p. 68).

k, k_1, k_2 Frequency bin (p. 68).

$P[i,k]$ Power estimate in the kth frequency bin, based on the ith data segment (p. 69).

U Normalization factor dependent on the data window (p. 68).

$X[i,k]$ Fourier Transform of the ith data segment (p. 68).

$\mathbf{x}$ Data vector (p. 69).

N Length of entire data vector, $\mathbf{x}$ (p. 69).

Ω Matrix form of the data window (p. 69).

$\Upsilon[i,k]$, Υ Central matrix in the quadratic form used to estimate the spectrum (p. 69).

C_k, S_k Matrices describing the cosine and sine effects in the Fourier Transform (p. 69).

Z The length of the zero-padded data segments (p. 69).

$M(t)$ The moment-generating function (p. 70).

η_i Eigenvalues of $\mathbf{R}\Upsilon$ (p. 70).

ν_i Multiplicity of the eigenvalue, η_i (p. 70).

D The number of distinct non-zero eigenvalues (p. 70).

μ_P Mean of the power estimate (p. 70).

σ_P^2 Variance of the power estimate (p. 70).

d_i Convenient coefficent based on the eigenvalues, η_i (p. 70).

g Convenient coefficient based on the eigenvalues, η_i (p. 70).

J The total number of non-zero eigenvalues, (p. 70).

λ_i Eigenvalues of Υ (p. 72).

b Number of frequency bins (p. 74).

ρ Parameter for the Markov process (p. 77).

$D(f||g)$ Relative entropy (Kullback-Leibler distance) between two probability density functions, f and g (p. 79).

α Parameter of the Gamma distribution, see also β (p. 79).

β Parameter of the Gamma distribution, see also β (p. 79).

Bibliography

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing*, vol. 1, Artech House Inc., Norwood, MA, 1981.

[2] G. P. de Loor and P. Hoogeboom, "A discrepancy between ground-based and airborne radar backscatter measurements", *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-20, pp. 134–136, 1982.

[3] B. Kinsman, *Wind Waves*, Prentice-Hall, Englewood Cliffs, New Jersey, 1965.

[4] P. Beckmann and A. Spizzachino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Macmillan, New York, 1963.

[5] G. R. Valenzuela, "Theories for the interaction of electromagnetic and oceanic waves–a review", *Boundary Layer Meteorology*, vol. 13, pp. 61–85, 1978.

[6] G. L. Tyler, "Wavelength dependence in radio-wave scattering and specular-point theory", *Radio Science*, vol. 11, pp. 83–91, 1976.

[7] D. V. Arnold, *Electromagnetic Bias in Radar Altimetry at Microwave Frequencies*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1992.

[8] D. Holliday, G. ST-Cyr, and N. E. Woods, "A radar ocean imaging model for small to moderate incidence angles", *Int. J. Remote Sensing*, vol. 7, no. 12, pp. 1809–1934, 1986.

[9] N. Bleistein and R. A. Handelsman, *Asymptotic expansion of integrals*, Dover, 1986.

[10] P. E. Johnson, "Backscatter sensitivity to the parameters of a two-dimensional scattering model", Master's thesis, Brigham Young University, 1994.

[11] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 1994.

[12] K. F. Warnick and D. V. Arnold, "Generalization of the geometrical optics scattering limit for a rough conducting surface", *JOSA A*, vol. 15, no. 9, pp. 2355–2361, September 1998.

[13] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Inc., New York, 1991.

[14] M. L. Banner, I. S. F. Jones, and J. C. Trinder, "Wavenumber spectra of short gravity waves", *Journal of Fluid Mechanics*, vol. 198, pp. 321–344, 1989.

[15] N. E. Huang, C. Tung, and S. R. Long, *The Sea: Ocean Engineering Science*, vol. 9, chapter 6 Wave Spectra, pp. 197–237, John Wiley and Sons, 1990.

[16] B. Jahne and K. S. Riemer, "Two-dimensional wave number spectra of small-scale water surface waves", *Journal of Geophysical Research*, vol. 95, pp. 11,531–11,546, 1990.

[17] O. M. Phillips, "The equilibrium range in the spectrum of wind-generated waves", *Journal of Fluid Mechanics*, vol. 4, pp. 426–434, 1958.

[18] J. W. Johnson and D. E. Weissman, "Two-frequency microwave resonance measurements from an aircraft: A qualitative estimate of the directional ocean surface spectrum", *Radio Science*, vol. 19, no. 3, pp. 841–854, 1984.

[19] D. L. Schuler, "Remote sensing of directional gravity wave spectra and surface currents using a microwave dual-frequency scatterometer", *Radio Science*, vol. 13, no. 2, pp. 321–331, Mar–Apr 1978.

[20] O. H. Shemdin, H. M. Tran, and S. C. Wu, "Directional measurement of short ocean waves with stereophotography", *Journal of Geophysical Research*, vol. 93, pp. 13,891–13,901, 1988.

[21] W. T. Shaw, A. J. Dougan, and R. J. A. Tough, "Analytical expressions for correlation functions and Kirchoff integrals for Gaussian surfaces with ocean-like spectra", *IEEE Transactions on Antennas and Propogation*, vol. 44, no. 11, pp. 1454–1463, November 1996.

[22] D. V. Arnold, W. K. Melville, R. H. Stewart, J. A. Kong, W. C. Keller, and E. Lamarre, "Measurements of electromagnetic bias at Ku and C bands", *Journal of Geophysical Research*, vol. 100, pp. 969–980, January 1995.

[23] L. G. Hevizi, E. J. Walsh, R. E. McIntosh, D. Vandermark, D. E. Hines, R. N. Swift, and J. Scott, "Electromagnetic bias in sea surface range measurements at frequencies of the TOPEX/POSEIDON satellite", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 2, pp. 376–388, March 1993.

[24] E. J. Walsh and et. al., "Frequency dependence of the electromagnetic bias in radar altimeter sea surface range measurements", *Journal of Geophysical Research*, vol. 96, pp. 20 571–20 583, 1991.

[25] J. Pedlosky, *Geophysical Fluid Dynamics*, Springer-Verlag, New York, 1987.

[26] M. R. Keller, W. C. Keller, and W. J. Plant, "A wave tank study of the dependence of X band cross sections on wind speed and water temperature", *Journal of Geophysical Research*, vol. 97, no. C4, pp. 5771–5792, April 1992.

[27] W. J. Pierson, "Examples of, reasons for, and consequences of the poor quality of wind data from ships for the marine boundary layer: Implications for remote sensing", *Journal of Geophysical Research*, vol. 95, no. C8, pp. 13313–13339, August 1990.

[28] M. I. Skolnik, Ed., *Radar Handbook*, McGraw-Hill, New York, 2nd edition, 1990.

[29] F. M. Naderi, M. H. Freilich, and D. G. Long, "Spaceborne radar measurement of wind velocity over the ocean—an overview of the NSCAT scatterometer system", *Proceedings of the IEEE*, vol. 79, no. 6, pp. 850–866, June 1991.

[30] D. G. Long and J. M. Mendel, "Model-based estimation of wind fields over the ocean from wind scatterometer measurements, part i: Development of the wind field model", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 349–360, May 1990.

[31] D. G. Long, "Current progress in KU band model functions", NSCAT Science Team Report, `http://mers.byu.edu/Papers/Mers_reports.html`, 1994.

[32] F. J. Wentz, "A model function for ocean radar cross sections at 14.6 GHz", *Journal of Geophysical Research*, vol. 89, no. C3, pp. 3689–3704, May 1984.

[33] W. T. Liu, "The effects of the variations in sea surface temperature and atmospheric stability in the estimation of average wind speed by Seasat-SASS", *Journal of Physical Oceanography*, vol. 14, no. 2, pp. 392–401, February 1994.

[34] R. N. Hoffman and J. F. Louis, "The influence of atmospheric stratification on scatterometer winds", *Journal of Geophysical Research*, vol. 95, no. C6, pp. 9723–9730, June 1990.

[35] P. E. Johnson and D. G. Long, "Backscatter variability observed in C-band and Ku-band scatterometer data", in *IGARSS '97*. International Geoscience and Remote Sensing Symposium, 1997, pp. 1856–1858.

[36] P. E. Johnson, D. G. Long, and T. E. Oliphant, "Geophysical modeling error in wind scatterometry", in *IGARSS '96*. International Geoscience and Remote Sensing Symposium, 1996, pp. 1721–1723.

[37] W. Timothy Liu and W. G. Large, "Determination of surface stress by Seasat-SASS: A case study with JASIN data", *Journal of Physical Oceanography*, vol. 11, no. 12, pp. 1603–1611, December 1981.

[38] E. Kalnay and R. Atlas, "Global analysis of ocean surface wind and wind stress using a general circulation model and Seasat scatterometer winds", *Journal of Geophysical Research*, vol. 91, no. C2, pp. 2233–2240, February 1986.

[39] N. W. Guinard, Jr. J. T. Ransone, and J. C. Daley, "Variations of the NRCS of the sea with increasing roughness", *Journal of Geophysical Research*, vol. 76, pp. 1525, 1971.

[40] W. L. Jones and L. C. Schroeder, "Radar backscatter from the ocean: dependence on surface friction velocity", *Boundary Layer Meteorology*, vol. 13, pp. 133–149, 1978.

[41] R. E. Fischer, "Standard deviation of scatterometer measurements from space", *IEEE Transactions on Geoscience Electronics*, vol. GE-10, no. 2, pp. 106–113, 1972.

[42] D. G. Long, C-Y Chi, and F. K. Li, "The design of an onboard digital doppler processor for a spaceborne scatterometer", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 6, pp. 869–878, November 1988.

[43] D. G. Long and M. W. Spencer, "Radar backscatter measurement accuracy for a spaceborne pencil-beam wind scatterometer with transmit modulation", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 102–114, January 1997.

[44] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing*, vol. 2, Artech House Inc., Norwood, MA, 1981.

[45] A. C. M. Stoffelen and D. Anderson, "Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4", *Journal of Geophysical Research*, vol. 102, no. C3, pp. 5767–5780, March 1997.

[46] T. E. Oliphant, "New techniques for wind scatterometry", Master's thesis, Brigham Young University, August 1996.

[47] P. E. Johnson and D. G. Long, "Wind bias from sub-optimal estimation due to geophysical modeling error", in *IGARSS '98*. International Geoscience and Remote Sensing Symposium, 1998.

[48] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*, Addison-Wesley Publishing Company, 1991.

[49] J. Abel, "A bound on mean-square estimate error", *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1675–1680, May 1993.

[50] C. G. Brown, P. E. Johnson, and D. G. Long, "Wind field models and model order selection for wind estimation", in *IGARSS '97*. International Geoscience and Remote Sensing Symposium, 1997, pp. 1847–1849.

[51] P. D. Welch, "The use of the fast Fourier transform for the estimation of power spectra", *IEEE Trans. Audio Electroacoustics*, vol. AU-15, pp. 70–73, June 1967.

[52] T. S. Durrani and J. M. Nightingale, "Probability distributions for discrete fourier spectra", *Proceedings of the IEE*, vol. 120, no. 2, pp. 299–311, February 1973.

[53] T. S. Durrani, "Joint density functions for digital spectra", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-22, no. 5, pp. 314–320, October 1974.

[54] P. E. Johnson and D. G. Long, "The probability density of Welch's modified periodogram estimates", *IEEE Transactions on Signal Processing*, 1999, In Press.

[55] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.

[56] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.

[57] P. E. Johnson and D. G. Long, "The probability distribution of NSCAT measurements", in *IGARSS '98*. International Geoscience and Remote Sensing Symposium, 1998, pp. 968–970.

[58] A. D. Greenwood, "Azimuth modulation of the radar backscatter at near-normal incidence", Master's thesis, Brigham Young University, 1995.

[59] W. H. Beyer, *CRC Standard Mathematical Tables*, 28 edition, 1987.

[60] G. E. Roberts and H. Kaufman, *Table of Laplace Transforms*, W. B. Saunders Company, Philadelphia and London, 1966.

[61] A. V. Oppenheim and A. S. Willsky, *Signals and Systems*, Prentice Hall, 1983.

[62] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables", *Biometrika*, vol. 48, no. 3 and 4, pp. 419–426, 1961.

[63] P. E. Johnson, "Binary classification of wind fields through hypothesis testing on scatterometer measurements", in *RMNSGC Symposium, '97*. Rocky Mountain NASA Space Grant Consortium, June 1997.

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

[65] J. Gunther and D. G. Long, "Models for the near-surface oceanic vorticity and divergence", in *IGARSS '94*. International Geoscience and Remote Sensing Symposium, August 1994.