2002-01-01

Evaluation of a compound probability model with tower-mounted scatterometer data

David G. Long
david_long@byu.edu

Benjamin E. Barrowes

Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Electrical and Computer Engineering Commons

Original Publication Citation

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/facpub/550

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Evaluation of a Compound Probability Model With Tower-Mounted Scatterometer Data

Benjamin E. Barrowes and David G. Long, Senior Member, IEEE

Abstract—Six months of data from the YSCAT94 experiment conducted at the CCIW WAVES research platform on Lake Ontario, Canada, are analyzed to evaluate a compound probability model. YSCAT was an ultrawideband small footprint (\(\leq 1\) m) microwave scatterometer that operated at frequencies of 2–18 GHz, incidence angles from 0° to 60°, both h-pol and v-pol, and which tracked the wind using simultaneous weather measurements. The probability distribution function of the measured instantaneous backscattered amplitude (\(p(\alpha)\)) is compared to theoretical distributions developed from the composite model and a simple wave spectrum. Model parameters of the resulting Rayleigh/generalized lognormal distribution probability density function (pdf) (\(C, \alpha_1, \alpha_2\)) are derived directly from the data and are found to demonstrate relationships with wind speed, incidence angle, and radar frequency.

Index Terms—Generalized lognormal distribution, microwave scatterometer, sea surface scattering, small-footprint scatterometer.

I. INTRODUCTION

The principal application for scatterometers is ocean microwave anemometry, i.e., wind speed estimation over bodies of water through radar cross section measurements. The relationship between the environmental parameters of the air–sea interface and the observed radar cross section (\(\sigma^2\)) is referred to as the geophysical model function (GMF). Understanding the GMF is central in interpreting scatterometer data. However, it remains poorly understood due to the complexity of the air–sea interface. The normalized radar cross section \(\sigma^2\) of the sea surface is dependent on many parameters including incidence angle, microwave frequency, transmit and receive polarizations, wind direction, long wave field, salinity of the water, water temperature, air temperature, and other factors [1]. Tower-mounted scatterometers such as YSCAT are deployed in an effort to better describe the geophysical model function and aid in the understanding of the relationship between environmental parameters and radar backscatter. This paper focuses on the probability distribution of the instantaneous amplitude of electromagnetic backscatter (\(p(\alpha)\)) from a wind-roughened water surface.

A brief summary of the YSCAT instrument and the YSCAT94 experiment is provided in Section II. In Section III, \(p(\alpha)\) is modeled by a conditional probability [see (2)] following the development in [2]. The distribution for \(p(\sigma^2)\) is discussed in Section III-A-1, while \(p(\sigma^2)\) is derived for both the h-pol and v-pol cases in Section III-A-2. In Section IV the resulting distribution for \(p(\alpha)\), referred to as a Rayleigh/generalized lognormal distribution (R/gln), is calculated numerically using these distributions for \(p(\alpha)\) and \(p(\sigma^2)\) and then compared to YSCAT94 data sorted according to frequency, polarization, wind direction (upwind or downwind), incidence angle, and wind speed. This is followed by a summary and conclusion.

II. YSCAT INSTRUMENT

YSCAT is a tower mounted microwave scatterometer designed to collect normalized radar cross section (\(\sigma^2\)) measurements of the sea surface under varying radar and environmental parameters [3]. For this study, YSCAT gathered data at frequencies of 2 GHz (S-band), 3 GHz (S-band), 5 GHz (C-band), 10 GHz (X-band), and 14 GHz (K_a-band) and at incidence angles of 0° (nadir), 10°, 20°, 25°, 30°, 40°, 50°, and 60°  analyzed. YSCAT’s antenna was specially designed to provide a fixed beamwidth of approximately five degrees over most of its 2–18 GHz operating bandwidth. Mounted 10 m above the water surface of Lake Ontario, Canada, during the YSCAT94 experiment, the antenna footprint diameter was approximately 1 m for midrange incidence angles. YSCAT could transmit and receive at both horizontal and vertical polarizations and tracked the wind direction with the aid of simultaneous weather data acquired at the site. In this paper, “upwind” and “downwind” include ±20° of the wind direction. In situ measurements of wind speed, wind direction, rainfall, and water temperature measurements were also recorded. For a more detailed description of YSCAT, the reader is referred to [3] and [4]. A summary of YSCAT’s RF parameters is provided in Table I.

The YSCAT94 experiment consists of data collected by the YSCAT instrument when it was deployed for a period of six months, from June to November, 1994 on the WAVES research platform operated by the Canada Centre for Inland Waters (CCIW) about 1.1 km from the western shore of Lake Ontario. Water depth at this site is about 12 m, and the annual variation in water depth is less that 0.5 m. There are no significant tides, seiches, or associated currents, and other random currents are typically less than 10 cm/s. The CCIW tower was designed to minimize both wind and wave disruption [1] and therefore no effort is made to account for turbulence on the wind and the waves caused by the tower. Prevailing winds in this area were...
TABLE I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2-18 GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center Frequency</td>
<td></td>
</tr>
<tr>
<td>Peak Output Power</td>
<td>23 dBm</td>
</tr>
<tr>
<td>Transmit Polarization</td>
<td>V or H</td>
</tr>
<tr>
<td>Two-Way Antenna Beam width</td>
<td>5-10°</td>
</tr>
<tr>
<td>Receive Polarization</td>
<td>V or H</td>
</tr>
<tr>
<td>Polarization Isolation</td>
<td>15-20 dB</td>
</tr>
<tr>
<td>Dynamic Range</td>
<td>60 dB</td>
</tr>
<tr>
<td>Baseband Signal Bandwidth</td>
<td>900 Hz</td>
</tr>
</tbody>
</table>

TABLE II

<table>
<thead>
<tr>
<th>Polarization</th>
<th>Frequency (GHz)</th>
<th>Wind Speed (m/s)</th>
<th>Count of Valid YSCAT94 σ² Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>2</td>
<td>5</td>
<td>0-5</td>
</tr>
<tr>
<td>H</td>
<td>3</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>H</td>
<td>5</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>H</td>
<td>10</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>V</td>
<td>2</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>V</td>
<td>3</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>V</td>
<td>5</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>V</td>
<td>10</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>V</td>
<td>14</td>
<td>7.5-10</td>
<td>&gt; 10</td>
</tr>
</tbody>
</table>

Barrowes and Long: Evaluation of a Compound Probability Model

III. YSCAT94 Backscatter Distributions

The most common model for sea scattered radar return at moderate incidence angles is the composite model. The composite model assumes that the sea surface is composed of small independent patches each of which has a normalized radar cross section \( \sigma^2 \) given by small perturbation theory (SPT) as \[1] \[5\]

\[
\sigma^2 = 16\pi k_m^4 |g_{pp}(\theta_i)|^2 \Psi(2k_m \sin \theta_i; 0)
\]  (1)

where \( \sigma^2 \) is the normalized radar cross section, \( \theta_i \) is the incidence angle, \( g_{pp}(\theta_i) \) is a polarization dependent reflection coefficient with \( pp \) being either \( hh \) or \( vv \), \( k_m \) is the microwave wavenumber, and \( \Psi \) is the wave height spectral density evaluated at the Bragg wavelength \( k_b = 2k_m \sin \theta_i \).

These patches of relatively small waves (on the order of centimeters), are modulated, or tilted, by larger waves with wavelengths typically on the order of meters. Consequently, the radar cross section distribution depends on the distribution of the long wave field. Following Gotwols and Thompson [2], the compound probability model is reviewed and extended in the remainder of this section.

A. Compound Probability Model

The compound probability model, originally proposed by Valenzuela and Laing [6], considers the aforementioned two scales of waves separately. According to the model, the radar cross section \( \sigma^2 \) of the sea surface depends on both the waves which are on the order of or smaller than the radar footprint (the Bragg waves) and on the underlying tilt imposed from waves with wavelengths much larger than the radar footprint (gravity waves). For the former case of shorter wavelength waves, \( \sigma^2 \) is considered constant but the instantaneous amplitude of the return varies, yielding the conditional probability \( p(\sigma^2 | \sigma^2) \). For the latter case of longer wavelength waves due to incidence angle and hydrodynamic modulation by long wavelength waves, \( \sigma^2 \) is allowed to vary with probability \( p(\sigma^2) \). The amplitude distribution may then be expressed as the conditional probability

\[
p(\sigma^2) = \int_0^\infty p(\sigma^2 | \sigma^2) p(\sigma^2) d\sigma^2.
\]  (2)

The probability of measuring a given backscatter amplitude \( \sigma^2 \) can be calculated by considering distributions on the orders of both scales.

1) Distribution of \( p(\sigma^2) \): When the scatterometer footprint is large, the scattered fields should be normally distributed via the central limit theorem. In this case, the amplitude \( \sigma^2 \) of the radar return should be Rayleigh distributed [7]. On the other hand, when the scatterometer footprint is on the order of the intermediate to large sized waves, this assumption is less valid, but should still hold if the footprint encompasses several

westerly, which provided fetches from 1.1–2 km. Due to this short fetch, waves with periods of four seconds were common, while waves with periods of 8 s or more were rare [4].

The data from the YSCAT94 experiment analyzed in this paper consist of one minute backscatter amplitude \( \sigma^2 \) records measured at a 2 kHz sampling rate. The 2 Hz power measurements are averaged to yield a data rate of 10 Hz. This integration time of 100 ms is long compared to the coherence time of the waves on the scale of interest and therefore may effect our estimated backscatter distributions, especially at low amplitudes where spikes may dominate the otherwise small average. Radar cross sections \( \sigma^2 \) were calculated from these records, and were binned according to frequency, polarization, wind direction (upwind or downwind), incidence angle, and wind speed. The vast majority of one minute backscatter amplitude records were measured during steady wind conditions leading to stable \( \sigma^2 \) values. However, one minute data records with means far removed from the aggregate mean of that data bin were generally found to have fluctuating wind speed measurements and unstable \( \sigma^2 \) values and were therefore discarded. Data collected during or after rainy periods and data corrupted by equipment failures or other sources of error (e.g., ships, birds) were also removed. Table II summarizes the resulting number of \( \sigma^2 \) measurements in each data bin after these data records were discarded. To reduce errors introduced by receiver gain fluctuation due to temperature changes, distance variations from the water surface, or other factors, each one minute data record was first normalized by dividing by the mean of that individual record. Subsequently, all one minute data records in that bin were multiplied by the aggregate mean of all one minute data records in that bin.
Phillip’s spectrum, Gotwols and Thompson return in log space as a function of wave slope and zero mean and noting that

\[ \sigma_x^c = 2a_1 s_x + a_2 s_x^2 \]  

(3)

where \( p \) is either \( v \) or \( h \) for \( v \)-pol and \( h \)-pol respectively. From (3) we can find an expression for \( p(\sigma_x^c) \) by applying the transformation law for probability density

\[ p(\sigma_x^c) = \frac{p(s_x)}{|d\sigma_x^c/ds_x|}. \]  

(4)

Assuming that the wave slope distribution \( p(s_x) \) is a normal distribution with variance \( \sigma_x^2 \) and zero mean and noting that

\[ |d\sigma_x^c/ds_x| = C(a_1 + 2a_2 s_x)^{a_1/2 + a_2 s_x^2} \]

(5)

and from (3)

\[ s_x = \frac{-a_1 \pm \sqrt{a_1^2 + 4a_2 \ln \left( \frac{\sigma_x^2}{C} \right)}}{2a_2} \]  

(6)

it can be shown that for \( a_2 = 0 \), [2]

\[ p(\sigma_x^c) = \frac{1}{a_1 \sigma_x^c \sqrt{2\pi}} \exp \left\{ -\frac{(\ln \sigma_x^c - \ln C)^2}{2a_1^2 \sigma_x^c} \right\} \]  

(7)

\[ p(a) = \int_0^\infty \frac{p(a|\sigma_x^c)}{p(\sigma_x^c)} d\sigma_x^c \]  

(8)

\[ = \int_0^\infty \frac{2a_2^2 \sigma_x^c}{p(a|\sigma_x^c)} \exp \left\{ \frac{a_1^2}{4a_2^2} \right\} \exp \left\{ \pm \frac{a_1 \sqrt{a_1^2 + 4a_2 \ln \frac{\sigma_x^c}{C}}}{4\sigma_x^2 a_2^2} \right\} \left( \frac{1}{2\sigma_x^c} \right) \]  

(9)
and for the more general case of \( a_1, a_2 \neq 0 \) \([10]\)

\[
p(\sigma^c | s_x) = \frac{1}{\sigma_c (a_1 + 2a_2 s_x)} C e^{a_1 s_x + a_2 s_x^2} \sqrt{2\pi} \cdot \exp \left\{ -\frac{(s_x)^2}{2\sigma_x^2} \right\} \tag{8}\]

with \( s_x \) as defined by (6) and for \( s_x \to N(0, \sigma_x) \). The pdf in (7) is referred to as the lognormal distribution. The pdf described by (8), was first identified in [11] as the generalized lognormal distribution and is the focus of the next section. From (2), \( p(\sigma^c) \) may be modeled by (7) for \( a_2 = 0 \) and (8) for \( a_1, a_2 \neq 0 \).

For the more general case of \( a_1, a_2 \neq 0 \), (2) becomes (9), shown at the bottom of the previous page, where polarization dependence has been dropped for notational convenience. Using (3) and (4), (9) may also be written in terms of \( s_x \) as

\[
p(\sigma) = \int_0^\infty p(\sigma^c | s_x) \cdot p(s_x) \cdot ds_x = \int_0^\infty 2aC e^{a_1 s_x + a_2 s_x^2} e^{-s_x^2 C e^{a_1 s_x + a_2 s_x^2} \frac{2\pi}{\sigma^2}} ds_x, \tag{10}\]

Equation (10) provides a model which describes the pdf of the radar backscatter amplitude from the sea surface at midrange incidence angles. More is said about this distribution in Section III-C.

### B. Generalized Lognormal Distribution

The generalized lognormal distribution is defined in terms of \( s_x \) by (8) or in terms of \( \sigma^c \) [using (3)] as in (11), shown at the bottom of the next page. Four parameters, \( a_1, a_2, C, \) and \( \sigma_x \), dictate the distribution function produced by (11). However, the distribution can be degenerate (i.e., it has nonunique solutions) when all four parameters are left free. In order to contrast \( a_1, a_2, \) and \( C \) values for different data bins, this problem of degeneracy is alleviated by fixing \( \sigma_x \) to be the representative slope standard deviation (\( \sigma_x \)) of 0.0914. This value of the slope standard deviation was chosen from the \( \sigma_x \) population calculated from wire wave gauge data [10] recorded at the CCIW site. The absolute values of \( a_1, a_2, \) and \( C \) are dependent on this fixed \( \sigma_x \). However, the relative values of these three parameters is comparable to other studies because the shape of the generalized lognormal distribution is dependent only upon the relative \( a_1 \) and \( a_2 \) [11].

To aid in the comparison of YSCAT94 data with (10), the mean and variance of the generalized lognormal distribution

---

**Fig. 2.** Aggregate amplitude mean (\( \mu \)) and variance (\( \sigma^2 \)) of YSCAT94 amplitude measurements according to Bragg wavelength (\( \lambda \)) for a fixed wind speed of 4 m/s \([a] \) and \([b] \) and for a fixed wave slope \([c] \) and \([d] \). Circles indicate v-pol, upwind values. Pluses indicate h-pol, upwind values.

**Fig. 3.** \( |\alpha_1| \) values versus significant wave slope. \((a) 20^\circ \), \((b) 25^\circ \), \((c) 30^\circ \), and \((d) 40^\circ \) incidence angle. Data is for v-pol, downwind case. The thin solid line corresponds to 2 GHz, dashed line corresponds to 3 GHz, dotted line corresponds to 5 GHz, dash/dot line corresponds to 10 GHz, and the bold solid line corresponds to 14 GHz. Circles indicate bins with only one or two minutes of data.
Fig. 4. $a_2$ values versus wind speed and significant wave slope. (a), (b) 20°, (c), (d) 30°, and (e) and (f) 40° incidence angle. Data is for v-pol, downwind case. The thin solid line corresponds to 2 GHz, the dashed line corresponds to 3 GHz, the dotted line corresponds to 5 GHz, dash/dot line corresponds to 10 GHz, and the bold solid line corresponds to 14 GHz. Asterisks correspond to bins with no data, and circles indicate bins with only one or two minutes of data.

Equation (12) may be written as (14), shown at the bottom of the next page, and thus the mean of the generalized lognormal distribution is given by

$$\mu_{gln} = \frac{C \exp \left( \frac{a_1^2 \sigma^2}{1 - 2a_2\sigma^2} \right)}{\sqrt{1 - 2a_2\sigma^2}}.$$  \hfill (15)
produced in this manner may be fit to histograms generated from YSCAT94 10 Hz sampled raw data records [10] for each data case using the Kullback–Leibler distance between two probability distributions [12]

\[ p(f||g) = \sum_{n=-\infty}^{\infty} \left| f(x_n) \log \frac{f(x_n)}{g(x_n)} \right|. \]  

(17)

By minimizing (17), \( C, a_1, \) and \( a_2 \) values are determined for each YSCAT94 case according to frequency, polarization, wind direction, incidence angle, wind speed, and estimated significant wave slope. An example fit of the R/gln is shown in Fig. 1.

\[ \mu_{\text{gln}} = \lim_{a \to \infty} \int_0^{a} \frac{\exp \left( \frac{a_1^2}{4\sigma_z^2} \right) \exp \left( \pm a_1 \sqrt{\frac{a_1^2 + 4a_2 \ln \frac{\sigma_z}{C}}{4\sigma_z^2a_2^2}} \right) \left( \frac{\sigma_z}{C} \right)^{(1/2\sigma_z^2)}}{\sigma_z \sigma^2 \sqrt{\pi} \sqrt{a_1^2 + 4a_2 \ln \frac{\sigma_z}{C}}} \sigma^2 \, d\sigma^2. \]  

(12)

\[ \mu_{\text{gln}} = \lim_{a \to \infty} \int_0^{a} C \exp \left( \frac{u^2(2\sigma_z a_2 - 1) + u(2a_1) - (a_1^2 + 2a_2 \sigma_z^2 a_2^2)}{8\sigma_z^2a_2^2} \right) \frac{\sigma_z a_2 \sqrt{2\pi}}{} \, du. \]  

(14)
The fit produced by the R/gln is excellent even in the tails while a best fit Weibull distribution does not model the high amplitude portion of \( p(\alpha) \) as accurately. \( C, a_1, \) and \( a_2 \) results calculated in this manner for the more than one thousand data cases are discussed in Section IV.

IV. RESULTS

Fitting the R/gln distribution to the YSCAT94 data results in the distribution parameters which describe the behavior of the instantaneous amplitude backscatter distributions \( p(\alpha) \) of the YSCAT94 data set. Some statements about trends visible in the R/gln parameters, and specific examples from a few select cases are presented here. For a more exhaustive report of these results, the reader is referred to [10].

Each R/gln parameter displays general trends when tabulated versus environmental parameters. For example, the mean backscatter amplitude for each data case exhibits a general trend of increasing with wind speed and significant wave slope, though the increase is more gradual with significant wave slope. In general, the higher the frequency and the higher the incidence angle (i.e., the smaller the Bragg wavelength \( \Delta \)), the lower the mean amplitude. The same (very) general statements may be given for the variance of the amplitude measurements for each data case, although for both the mean and variance a decreasing trend can be observed for 0° (nadir) scattering, especially at higher frequencies. Fig. 2 shows the amplitude mean and variance for different Bragg wavelengths.

In general, \( C \) values display trends reversed from those of the amplitude mean and variance. \( C \) values tend to decrease according to a log relationship with wind speed and significant wave slope with the exception of 0° (nadir) cases, which display a slight tendency to increase. It should be remembered, however, that the model under consideration has no theoretical justification at very low incidence angles and therefore model parameters in this regime should be viewed accordingly.

R/gln \( a_1 \) values tend to stay in the same range of 5–25 for different frequencies and incidence angles. A slight upward trend can be seen in many cases when plotted versus significant wave...
values also tend to decrease as incidence angle increases, as can be seen in Fig. 3. R/gl n 2 values display a decreasing trend when considered versus wind speed and significant wave slope. These 2 values tend to be lower for smaller Bragg wavelengths (higher frequencies and increasing incidence angles) as illustrated in Fig. 4. A typical progression of the backscatter distributions for incidence angles of 0°–60° is shown graphically in Figs. 5–7. For the case of Bragg scattering from the ocean surface, increased Bragg wavelength—Wind speed sensitivity,” IEEE Trans. Geosci. Remote Sensing, vol. 34, pp. 656–666, Mar. 1996.


V. Conclusion

Following Gotwols and Thompson [2], the probability distribution function for the amplitude of the backscatter was calculated based on conditional probabilities (2). For YSCAT94 data, p(σ 2) was assumed to be Rayleigh distributed, and the distribution for p(σ 2) was theoretically shown to be the generalized lognormal distribution for midincidence angles. This model was based on a second degree polynomial in log space which approximated the normalized radar cross section (σ 2) predicted by the composite model of a sea containing waves generated by a simple Phillips power spectrum. The resulting Rayleigh/generalized lognormal distribution derived from (2) was fit to YSCAT94 empirical amplitude data distributions. The goodness of the fit demonstrates validity of this model especially for midrange incidence angle backscatter. Trends in the Rayleigh/generalized lognormal distribution parameters a1, a2, and C were identified. C displayed a very distinct trend with an inverse relationship to the mean: data cases with greater means corresponded to smaller values for C. The a1 and a2 parameters also displayed trends when considered versus wind speed and significant wave slope.

Acknowledgment

The authors wish to thank M. Donelan for permitting them to use wave staff data from his experiments and the Canada Centre for Inland Waters (CCIW), Burlington, ON, Canada, for permitting them to deploy YSCAT at the WAVES research platform in 1994.

Benjamin E. Barrowes received the B.S. and M.S. degrees in electrical engineering from Brigham Young University, Provo, UT, both in 1999, and is currently pursuing the Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge. His research interests include wind–wave interaction and electromagnetic wave scattering from the sea surface and from random media.

David G. Long (M’79–SM’98) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1989. From 1983 to 1990, he worked for NASA’s Jet Propulsion Laboratory, Pasadena, CA, where he developed advanced radar remote sensing systems. While at JPL, he was Project Engineer on the NASA Scatterometer (NSCAT) Project, which flew from 1996 to 1997. He also managed the SCANSCAT project, the precursor to SeaWinds that was first flown in 1999. He is currently a Professor in the Electrical and Computer Engineering Department, Brigham Young University, Provo, UT, where he teaches upper division and graduate courses in communications, microwave remote sensing, radar, and signal processing. He is the Principle Investigator on several NASA-sponsored research projects in remote sensing. He has numerous publications in signal processing and radar scatterometry. His research interests include microwave remote sensing, radar theory, space-based sensing, estimation theory, signal processing, and mesoscale atmospheric dynamics.