GEOPHYSICAL MODELING ERROR IN WIND SCATTEROMETRY

Paul E. Johnson and David G. Long
Brigham Young University
459 CB, Provo, UT 84602
801-378-4884, FAX: 801-378-6586, e-mail: johnsonp@nscat.ee.byu.edu

Abstract--A scatterometer is a satellite-borne instrument designed to measure wind over the ocean. Scatterometer wind retrieval is based on the relationship between the wind over the ocean and the resulting scattering cross section of the surface; this relationship, termed the "geophysical model function," maps the wind speed, relative wind direction (relative to the antenna azimuth angle), antenna incidence angle, polarization and frequency band to the scattering cross section. The sea surface temperature, salinity, long waves, wind variability within a scatterometer footprint, etc., lend variability to the backscatter. A particular observation of the wind-dependent backscatter can be viewed as a random variable with mean given by the geophysical model function and variability due to unmodelled effects and measurement errors. Little is known about the variability due to unmodelled effects, or the statistics of this variability; this paper presents some considerations and simulations to estimate the magnitude of the model function error.

INTRODUCTION

The geophysical model function relates the wind over the ocean surface, along with parameters characterizing the way the radar looks at the surface, to the normalized radar cross section, σ^o, which is measured by the scatterometer. However, there are many unmodelled factors affecting the relationship between the wind and the radar cross section; these can be viewed as terms causing variability in the true value of the backscatter for given wind and satellite conditions. For example, CMOD4 doesn't account for temperature or salinity [1], which [2] suggests affect the backscatter. Understanding the magnitude and effect of this variability improves our understanding of the model function and the scatterometer measurement process.

In this paper, a measurement model is expressed, describing how the model function value of the backscatter is corrupted by thermal noise and unmodelled parameters; this leads to an equation for the variance of the model function. Then, simulated results demonstrate that this technique provides a means to estimate the model function error from scatterometer data. Data from the ERS-1 satellite is then examined to study the general behavior of the model function error.
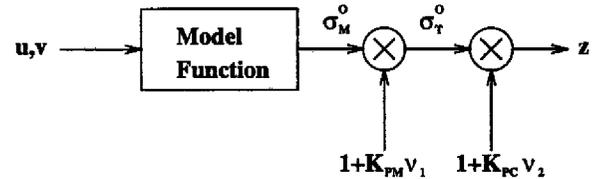
Figure 1: The model for scatterometer measurements. The wind is mapped through the model function to the model function backscatter; variability is introduced through K_{PM}, the effect of unmodelled parameters. The resulting "true" backscatter is corrupted by communication error (i.e., thermal noise) in the measurement process, which yields the measured value of the backscatter, z.

THE MEASUREMENT MODEL

Several sources introduce uncertainty into scatterometer measurements; in this paper we consider two: the communication error and the modeling error. The communication error, due to the thermal noise in the measurement process itself, is well understood [3]. Other causes of variability in the observed backscatter are collectively called "model function error."

Fig. 1 shows a block diagram of the measurement model. The model function maps the surface wind, along with the parameters of the scatterometer, to the model function backscatter, σ_M^o. This value is perturbed by unmodelled parameters to yield the true backscatter coefficient of the surface, σ_T^o. The measurement of the true backscatter, σ_T^o, is corrupted by thermal noise. The actual measurement, z, is modeled as

$$z = (1 + K_{PM}\nu_1)(1 + K_{PC}\nu_2)\sigma_M^o, \qquad (1)$$

where ν_1 and ν_2 are assumed to be independent, zero mean, unit variance, Gaussian random variables. K_{PM}^2 and K_{PC}^2 are the normalized variances for the modeling error and the communications error, respectively.

The expected value of the measurement, z, is σ_M^o, and the variance of z is:

$$\text{var}(z) = (K_{PM}^2 + K_{PC}^2 + K_{PM}^2 K_{PC}^2)\sigma_M^{o\,2}. \qquad (2)$$

To understand the effect of the modeling error, we examine the second moment of the measurement and

1

solve for K_{PM}^2:

$$K_{PM}^2 = \text{var}\left(\frac{z}{\sigma_M^o\sqrt{1+K_{PC}^2}}\right) - \frac{K_{PC}^2}{1+K_{PC}^2} \quad (3)$$

$$= \text{var}(d) - e \quad (4)$$

where d is a random variable based on the measurement, and e is the deterministic effect of the communication noise.

EQUATIONS TO ESTIMATE K_{PM}

The model function backscatter, σ_M^o, and the communication error, K_{PC}, depend on several parameters, including wind speed, wind direction, and radar incidence angle. Equation (3) further requires knowledge of the variance of the measurements for a given set of these parameters. Assuming that K_{PM} is a constant (at least over a sufficiently small range of the parameters) allows us to average over measurements with similar sets of the parameters to yield an estimate of the average value of K_{PM}^2, and the variance of this estimate. The data is subdivided to yield M estimates of $\delta = \text{var}(d)$ and of $\epsilon = \text{E}(e)$:

$$\begin{pmatrix} d_{1,1} & e_{1,1} \\ d_{1,2} & e_{1,2} \\ \vdots & \vdots \\ d_{1,N_0} & e_{1,N_0} \\ \cdots\cdots\cdots\cdots \\ \vdots & \vdots \\ \cdots\cdots\cdots\cdots \\ d_{M,1} & e_{M,1} \\ d_{M,2} & e_{M,2} \\ \vdots & \vdots \\ d_{M,N_0} & e_{M,N_0} \end{pmatrix} \Rightarrow \begin{pmatrix} \widehat{\delta_1} & \widehat{\epsilon_1} \\ \widehat{\delta_2} & \widehat{\epsilon_2} \\ \vdots & \vdots \\ \widehat{\delta_M} & \widehat{\epsilon_M} \end{pmatrix} . \quad (5)$$

These estimates produce M estimates of the model function error squared, and the variance of this estimate:

$$\begin{pmatrix} \widehat{K_{PM,1}^2} = \widehat{\delta_1} - \widehat{\epsilon_1} \\ \widehat{K_{PM,2}^2} = \widehat{\delta_1} - \widehat{\epsilon_1} \\ \vdots \\ \widehat{K_{PM,M}^2} = \widehat{\delta_1} - \widehat{\epsilon_1} \end{pmatrix} \Rightarrow \begin{pmatrix} \text{E}(\widehat{K_{PM}^2}) \\ \text{var}(\widehat{K_{PM}^2}) \end{pmatrix} . \quad (6)$$

Here, $\text{E}(\widehat{K_{PM}^2})$ is simply the sample mean of the M estimates. This estimate, though, is a function of random variables, so it is also a random variable with an associated variance. If we assume d is distributed normally, then the variance of the sample variances, $\widehat{\delta_m}$, depends on the true variance and the number of samples used to

estimate the variance:

$$\text{var}(\widehat{\delta_m}) = \frac{2\delta_m^2}{N_m - 1} \quad (7)$$

This fact allows us to obtain the variance of the estimate of K_{PM}^2

$$\text{var}(\widehat{K_{PM}^2}) = \frac{2}{N}\sum_{m=1}^{M}\frac{N_m}{N_m - 1}\delta_m^2 \quad (8)$$

We don't know the true variance δ_m, but choosing N_m large enough so that the estimate $\widehat{\delta_m}$ is sufficiently close to δ_m results in negligible error.

We now have estimates of the mean and variance of K_{PM}^2, so to obtain the mean and variance of K_{PM}, the function of a random variable, the principles of derived distributions and the fundamental theorem of statistics can be used [4]. Approximating with asymptotic integration, the expected value of the estimate of the model function error is

$$\text{E}(\widehat{K_{PM}}) \approx \sqrt{\text{E}(\widehat{K_{PM}^2}) - \frac{\text{var}(\widehat{K_{PM}^2})}{8}\text{E}(\widehat{K_{PM}^2})^{-\frac{3}{2}}} \quad (9)$$

and the variance of this estimate is

$$\text{var}(\widehat{K_{PM}}) \approx \frac{\text{var}(\widehat{K_{PM}^2})}{4\text{E}(\widehat{K_{PM}^2})} - \frac{\text{var}(\widehat{K_{PM}^2})^2}{64\text{E}(\widehat{K_{PM}^2})^3} \quad (10)$$

SIMULATIONS TO ESTIMATE K_{PM}

A simulated wind field, along with simulated ERS-1 measurements for several revolutions provides a test case for the estimation procedure. Values of K_{PM} are introduced in the simulation to add uncertainty about the backscatter and (9) is used to see its ability to estimate the value of K_{PM}.

Recalling Eq. (3), we see that knowledge of σ_M^o, the model function backscatter is required. However, the model function backscatter due to the true wind is much different from that due to the retrieved wind. Fig. 2 compares the terms z/σ_M^o where the model function value is generated from true and then the retrieved winds. Normalizing the measurements by the backscatter that results from the retrieved wind, yields a much smaller variance than when the measurements are normalized by the backscatter based on the true wind. This difference in variance biases the estimate of the model function error.

In satellite data, the true wind is not known. The wind is retrieved with maximum likelihood estimation [5] and de-aliased using a median filter based approach [6]; an estimate of the model function backscatter is obtained from the geophysical model function using the retrieved

2
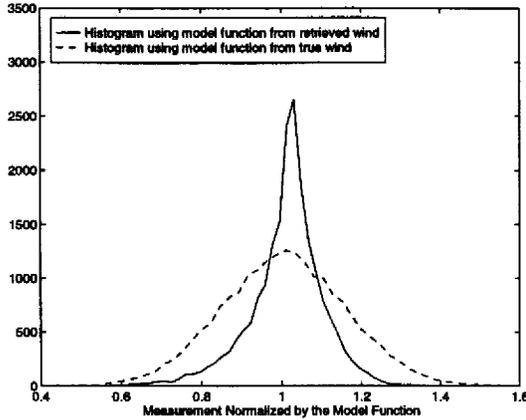
Figure 2: The backscatter measurement normalized by the model function based on the retrieved wind, $\sigma^o_{M,retrieved}$, yields a much smaller variance than that based on the true winds of the simulation, $\sigma^o_{M,true}$. This causes the estimate of the model function error, K_{PM}, to be low when the retrieved winds are used.
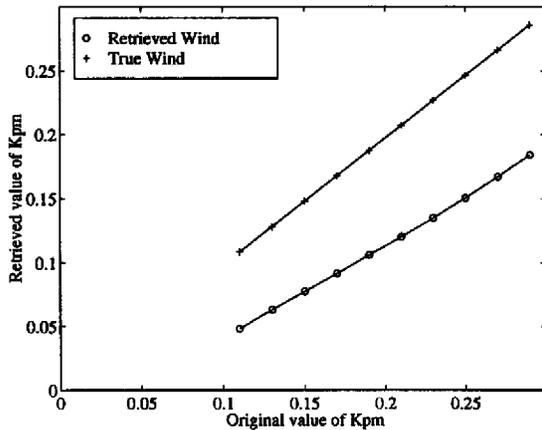
Figure 3: The estimate of K_{PM} based on the true wind provides a good reconstruction of the true value of K_{PM}, while that based on the retrieved wind (Median Filter De-aliasing) is consistently low.

wind. The estimate of the model function error based on the retrieved wind is consistently low, as shown in Fig. 3, while the estimate of K_{PM} using the true wind yields almost exact reconstructions of the simulated values of the model function error. This is due to the difference in the model function when driven by true wind as opposed to retrieved wind.

Recognizing the consistent bias of the estimate using retrieved rather than true winds, we can perform regression analysis to estimate a correction; this will allow K_{PM} to be estimated directly from the retrieved wind. Unfortunately, simple regression analysis assumes that

the regressor variable (in this case, $\widehat{K_{PM}}$ found using the retrieved wind) are deterministic; if there are errors in the regressor variable, more complicated techniques need to be employed. We assume for simplicity that because the variance of the estimates found with the retrieved wind are small in practice, this effect can be neglected. Fitting a quadratic equation to the data displayed in Fig. 3 suggests that a simple functional form relates the estimate of the K_{PM} based on the retrieved wind, to that resulting from the true wind:

$$K_{PM} = -0.966\widehat{K_{PM}}^2 + 1.567\widehat{K_{PM}} + 0.035. \quad (11)$$

Simulating additional wind fields and estimating the model function error with (9), and then using the correction of (11) results in accurate estimation of the value of K_{PM} used in the simulation.

Equation (11) implies that the corrected value of the model function error is a function of that using retrieved winds---which was a random variable. Thus, the mean and variance of the corrected value needs to be determined again, using derived distributions. These can be approximated using the first and second moments of the estimates $\widehat{K_{PM}}$ and $\widehat{K^2_{PM}}$.

ESTIMATES BASED ON ERS-1 DATA

Binning the ERS-1 data according to various parameters reveals the behavior of the model function error. In this section, rough estimates of K_{PM} are found, appropriately adjusted for use when the backscatter is based on the retrieved wind using (11), and its derived distribution. Representative values of the model function error are found, and its sensitivity to incidence angle and wind speed are observed.

Fig. 4 plots the value of K_{PM}, against the incidence angle, with error bars indicating one standard deviation from the mean. For low incidence angles, meaning near-swath, the model function error is very high, nearly 0.3. The value decreases as the incidence angle is increased. At about 45 degrees, the value of K_{PM} reaches a minimum of about 0.175, and then increases as the incidence angle continues to increase. This seems rather odd behavior, but could be linked to the fact that the model function is based on Bragg scattering which is an approximation generally valid from 20 to 60 degrees incidence angles. Near the edges of this range, the approximation is less valid.

The model function error is also quite sensitive to speed. Fig. 5 plots the estimate of K_{PM} for two incidence angle bins, [17, 37] degrees and [37, 57] degrees, against wind speed. Other incidence angle bins follow similar trends. Of course the incidence angle dependence is apparent here, showing that K_{PM} is higher for low incidence angles. Further, the model function

3
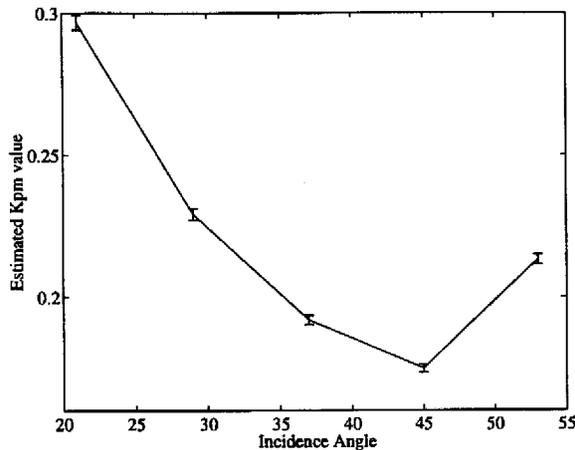
Figure 4: The estimate of K_{PM} varies between about 0.175 and 0.3, depending on the incidence angle.
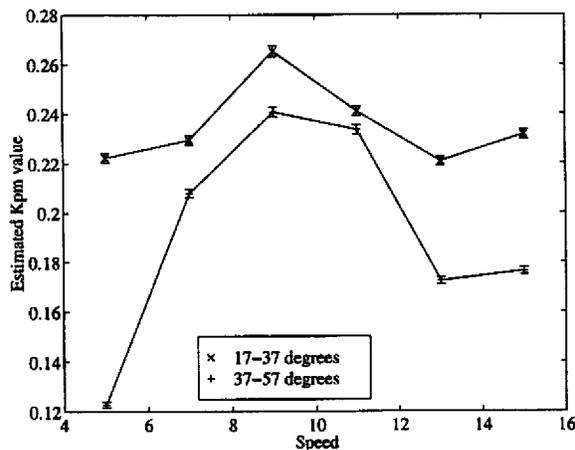
Figure 5: The estimate of K_{PM} depends on the wind speed. This plot was produced by binning K_{PM} values estimated for ERS-1 satellite data observed at low incidence angles and high incidence angles with different wind speeds.

error appears to be lower at extreme wind speeds (both low and high), and higher for wind speeds near the mean wind speed.

DISCUSSION

Unmodelled effects in the geophysical model function and the wind retrieval process contribute variability to the backscatter of the ocean surface. In this paper, we have found an expression for the model function error based on a simple model.

Simulations show that if the true surface wind is known, then the value of K_{PM} can be accurately es-

timated. Using retrieved wind, instead, the estimated model function error is consistently less than the actual value of K_{PM}. The correction function, found from simulations using independent, Gaussian random variables to introduce both communication noise and model function error, permits accurate estimates of K_{PM} based on the retrieved wind.

Examining ERS-1 data indicates the general behavior of the model function error. Low incidence angles have high model function errors, moderate incidence angles have low K_{PM}, and high incidence angles experience moderate values. K_{PM} is also sensitive to wind speed, though more so for lower incidence angles.

These results indicate that the model function error is appreciable, particularly when compared to the communication noise inherent in the scatterometer. These uncertainties limit the confidence that can be placed in the geophysical model, and need to be further understood to enhance the wind retrieval process.

This work is being continued to observe the impact of the model function error on the confidence intervals of retrieved wind from satellite data.

REFERENCES

[1] P. LeComte, "CMOD4 model description," Tech. Rep. ER-TN-ESA-GP-1120, European Space Agency, March 1993.

[2] M. Donelan and W. Pierson, "Radar scattering and equilibrium ranges in wind-generated waves with application to scatterometry," *IEEE Journal of Geophysical Research*, vol. 92, pp. 4971--5029, 1987.

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

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

[5] S. J. Shaffer, R. S. Dunbar, S. V. Hsiao, and D. G. Long, "A median-filter-based ambiguity removal algorithm for NSCAT," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, pp. 167--174, Jan 1991.

[6] C.-Y. Chi and F. Li, "A comparative study of several wind estimation algorithms for spaceborne scatterometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-26, pp. 115--121, March 1988.