# Wind Bias from Sub-optimal Estimation Due to Geophysical Modeling Error


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



## ABSTRACT

*Inherent in all empirical estimates of the geophysical model function (which relates the wind to the normalized radar cross section, NRCS, of the ocean surface) is uncertainty caused by parameters not included in the empirical model. This can be thought of as variability in the NRCS for given wind conditions. When the estimated variability is included in the maximum likelihood wind retrieval algorithm, complicated interdependencies arise between the estimated wind and the estimated variability.*

*Techniques for simplifying these interdependencies results in sub-optimal wind estimates. The model function variability can be ignored, or the log-variance term in the maximum likelihood estimation could be eliminated. In this paper, we consider the impact of these suggestions on NSCAT compass simulations. As expected, each method results in biased wind estimates. However, the biases are small, particularly when compared to the added complexity of the full solution.*


## INTRODUCTION

Scatterometry is based on a geophysical model function which relates the vector wind to the backscatter measurement [1]. 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 [2]. Empirical estimates of this function are based on aircraft scatterometer missions, refined with the growing body of satellite-borne scatterometer data [3]. These empirical models are believed to be accurate on average, but the variability and sensitivity to non-wind factors are not known.

A simple model which describes the basic measurement process is depicted in Fig. 1 [4]. The geophysical model function is modeled as the combination of an empirical model function, perturbed by a random, multiplicative term. 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 unmodeled parameters, along with a zero-mean unit-variance random variable, $\nu$, to yield the true backscatter coefficient of the surface (NRCS). The scatterometer, in attempting to measure the true backscatter, introduces communications, or radiometric, noise based on the temperature of the antenna. This term is quite well understood from first principles and is typically included

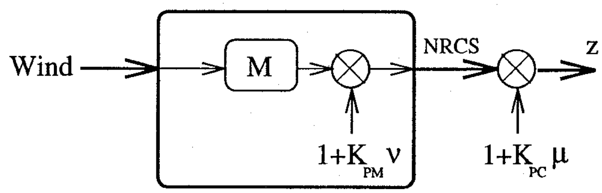


**Figure 1:** *The model for scatterometer measurements of the normalized radar cross section. The wind is mapped to the scatterometer measurement with uncertainty in the measurement caused by communcation noise ($K_{PC}$ term) and variability in the model function ($K_{PM}$ term).*

in radar systems as a multiplicative term, where $K_{PC}$ is the normalized standard deviation of the measurements from NRCS and $\mu$ is a zero-mean, unit-variance gaussian random variable [5]. A simple technique has been developed to estimate $K_{PM}$ directly from scatterometer measurements [4]. In this paper, the effect of geophysical modeling error on wind estimation is explored.

## WIND RETRIEVAL AND MEASUREMENT NOISE

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.

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 maximizes the probability of the measurements:

$$\hat{w} = \arg\max_{w} \frac{p(w)}{p(\vec{z})} p(\vec{z}|w). \quad (1)$$

The probability of the measurements, $p(\vec{z})$, does not change the maximization over the wind, and 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 reasonably

0-7803-4403-0/98/$10.00 © 1998 IEEE

independent, but model variability ($K_{PM}$) probably introduces some correlation between the measurements. For simplicity, we ignore this dependence and express the pdf as

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

where $\mathcal{M}_k$ is the empirical model function value based on the given wind and the $k$th measurement conditions, and the variance of the measurements is $\delta_k^2 = K_{PM}^2 + K_{PC}^2 + K_{PM}^2 K_{PC}^2$.

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

$$L(w,z) = -\sum_{k=1}^{K} \left\{ \frac{[(z_k - \mathcal{M}_k)^2]}{2\delta_k^2} + \frac{1}{2}\log\left[\delta_k^2\right] + \frac{1}{2}\log\left[2\pi\right] \right\} \quad (3)$$

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 method.

Estimation of $K_{PM}$ from scatterometer measurements requires, in place of the true wind, the estimated wind. This likelihood function displays a dependence of the wind retrieval on $K_{PM}$, resulting in a complex relationship between the two; setting $K_{PM} = 0$ in the retrieval simplifies the estimator, but reduces the optimality. Removing the log variance term from the log-likelihood function (reducing the problem to weighted least squares) dramatically reduces the impact of $K_{PM}$, but also changes the optimality conditions of the 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 of a simulated $K_{PM}$ value, and the resulting NRCS was measured with a noisy ($K_{PC}$) scatterometer and the wind estimated from such measurements. Figs. 2 and 3 display the average results of 100 000 simulations by plotting the retrieved speed and direction errors (simulated minus retrieved) for three scenarios. The solid line (labelled 'a') is for simulation and retrieval with $K_{PM} = 0$. The dashed line (labelled '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 (labelled '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). The plots show that with the unrealistic case of $K_{PM} = 0$ (there are always

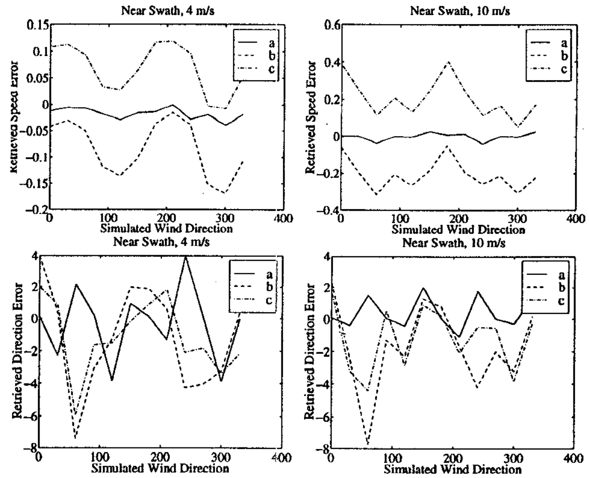


**Figure 2:** *Impact of $K_{PM}$ on simulated wind estimation at near swath. See text for legend definition.*

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

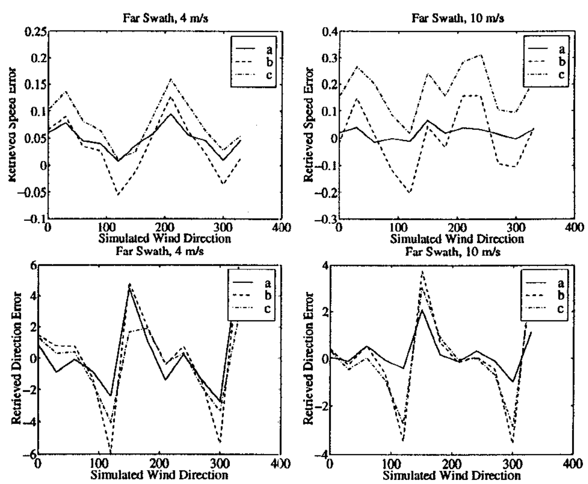


**Figure 3:** *Impact of $K_{PM}$ on simulated wind estimation at far swath. See text for legend definition.*

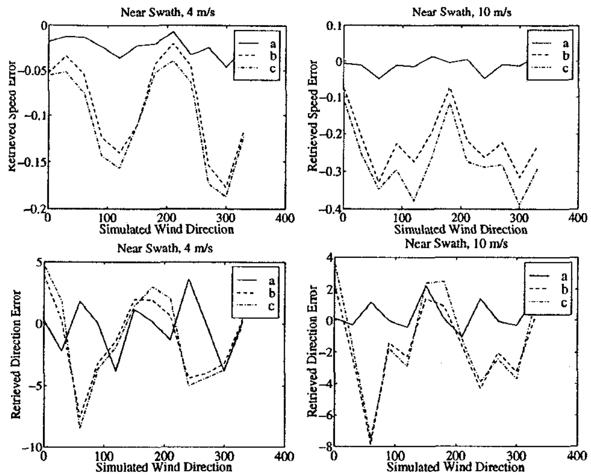


**Figure 4**: *Impact of removing the log variance term on simulated wind estimation at near swath.*

non-predictable ways. Also note that the same behavior is evident at far swath, though is less pronounced.

Similar compass simulations were performed with the log variance term removed from the estimator, with the results displayed in Figs. 4 and 5 for near 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.

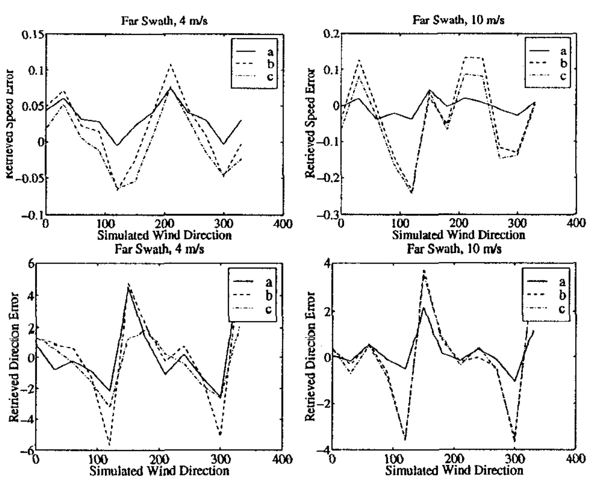


**Figure 5**: *Impact of removing the log variance term on simulated wind estimation at far swath.*

## CONCLUSIONS

All parameters affecting the backscatter are not known at the current state of science, and are certainly not all measured with scatterometers; thus all empirical model functions will inherently have model variability. Having previously developed a technique to estimate this variability, in this paper we examine the effect on wind retrieval of $K_{PM}$ and, in particular, the bias caused by suboptimal ways of estimating the wind without $K_{PM}$.

Simulations demonstrate that non-zero values of $K_{PM}$ cause a slight bias in the wind speed retrieval, on the order of 2%, compared to scattering from surfaces with $K_{PM} = 0$. If $K_{PM}$ is known exactly and incorporated in the retrieval process, the speed is still biased. Considering the very small bias caused by model variability, and noting the inability of the MLE to reduce that bias even if $K_{PM}$ is known exactly, it seems an unnecessary complication to the wind retrieval process to assume anything other than $K_{PM} = 0$.

Removing the log variance term, to reduce the estimate to a weighted least squares estimate, similarly introduces a small bias in the wind speed estimation, which is slightly amplified by model variability. Because of the small magnitude of the bias, this term can be dropped for computational simplicity, despite the very slight bias caused by removing the term. We note that removing the log variance term also greatly reduces bias caused by an incorrect value of $K_{PM}$ in the retrieval.

## REFERENCES

[1] 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, pp. 5771–5792, April 1992.

[2] 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, pp. 850–866, June 1991.

[3] F. J. Wentz, "A model function for ocean radar cross sections at 14.6 GHz," *Journal of Geophysical Research*, vol. 89, pp. 3689–3704, May 1984.

[4] 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.

[5] 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.