Wind Field Models and Model Order Selection for Wind Estimation

Charles G. Brown, Paul E. Johnson, Stephen L. Richards, and David G. Long
Brigham Young University, MERS Laboratory
459 CB, Provo, UT 84602
801-378-4884, FAX: 801-378-6586 browncg@@salt.ee.byu.edu

Abstract—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 for many resolution elements, simultaneously, using a wind field model that constrains the spatial variability of the wind.

In this paper several wind field models are presented for use in field-wise wind estimation. Model accuracy, as a function of the number of model parameters, is reported for each model. This accuracy is evaluated using NSCAT JPL nudged L2.0 data.

In order to reduce the computational load, automated classification schemes are developed to select the optimal number of model parameters necessary for a given wind field. Classification is performed through hypothesis testing on raw NSCAT data and point-wise estimates.

INTRODUCTION

Radar backscatter data from NSCAT (NASA Scatterometer) are related to the near-surface ocean wind through a geophysical model function. However, due to the nature of the geophysical model function, each set of measurements yields an ambiguous set of estimates of the corresponding near-surface ocean wind in traditional point-wise estimation. An additional step is required to identify a unique solution from the very large number of possible fields. Field-wise estimation provides wind estimates for many resolution elements simultaneously by estimating the parameters of a wind field model, involving local or global optimizations on the model parameters. The accuracy of field-wise estimation increases with the number of model parameters. However, as the number of model parameters increases so does the computational expense of the optimizations.

In this paper we present several simple wind field models and evaluate the average accuracy of each model as a function of the number of model parameters. We also present two wind field classification algorithms. The first is based on radar backscatter data (σ^0), while the second uses the ambiguous point-wise estimates. The algorithms are used to classify which wind fields can be modeled by low-order models and which fields require higher-order models. This approach decreases the average number of model parameters without significantly increasing the average modeling error. Throughout this work model accuracy is evaluated relative to NSCAT JPL nudged data, though we recognize the occasional failure of nudged data to identify the correct wind ambiguity. These failures can adversely affect the performance of our algorithm beyond that expected from simulated wind.

WIND FIELD MODELS

We consider linear wind field models applied to square regions, 600 km on each side. They are described by a matrix F in the equation $\mathbf{W}_M = F\mathbf{X}$, where $\mathbf{X}$ is a column vector of model parameters and $\mathbf{W}_M = \left(\mathbf{U}^T\ \mathbf{V}^T\right)^T$, with $\mathbf{U}$ and $\mathbf{V}$ defined as column vectors containing the rectangular components of the wind vector cells. The model order is determined by the number of parameters in $\mathbf{X}$ or, equivalently, the number of columns of F that are used. We evaluate several models in this section: Fourier and Legendre basis models [1], an orthogonal basis model derived using the Karhunen-Loeve transform [2], and the parameterized boundary conditions (PBC) model [3].

The matrix F of the Fourier model is constructed by sampling two-dimensional Fourier basis functions, while that of the Legendre model is created by sampling two-dimensional Legendre polynomials. The Karhunen-Loeve (KL) model, in contrast, is data driven. To create F, NSCAT JPL nudged L2.0 data was used to estimate the autocorrelation matrix for 600 km square regions, and the orthonormalized eigenvectors of the autocorrelation matrix were column-scanned to form the columns of the model matrix F. Finally, the columns of F were sorted in descending order according to the corresponding eigenvalues. The PBC model describes the field in terms of the divergence and vorticity of the pressure field along the region boundary [3].

In order to evaluate each of the models and compare their performance, the models were fit to regions of NSCAT JPL nudged data under a least-squared error constraint for each of the four models and for a range of model orders. The model based wind, $\hat{\mathbf{W}}$ is computed according to the equation $\hat{\mathbf{W}} = FF^{\dagger}\mathbf{W}$, where $F^{\dagger}$ is the pseudo-inverse of F, and $\mathbf{W}$ is the nudged wind field. The average errors for fits with the four models are displayed in Fig. 1 as functions of the number of model parameters used. The speed, direction, vector and normalized RMS errors between the model fit field $\hat{\mathbf{W}}$ and the nudged field $\mathbf{W}$ show similar

0-7803-3836-7/97/$10.00

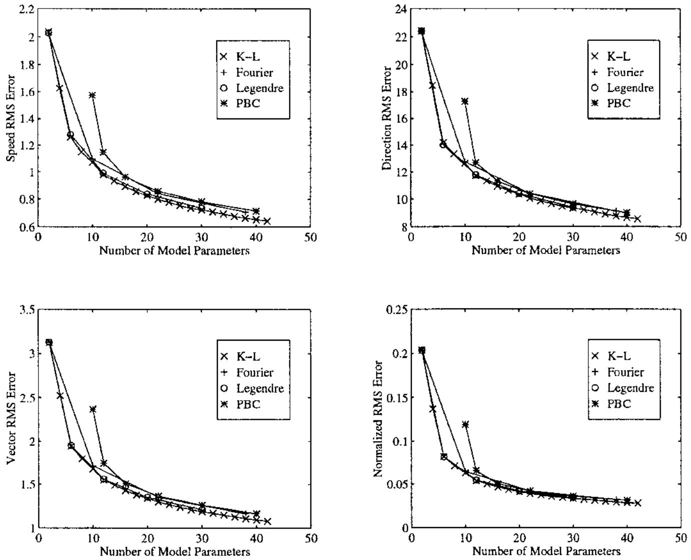

Figure 1: *Average speed, direction, vector, and normalized rms error as a function of number of model parameters. These errors represent the difference between the nudged wind field,* $\mathbf{W}$*, and the least-squares model fit,* $\hat{\mathbf{W}}$.

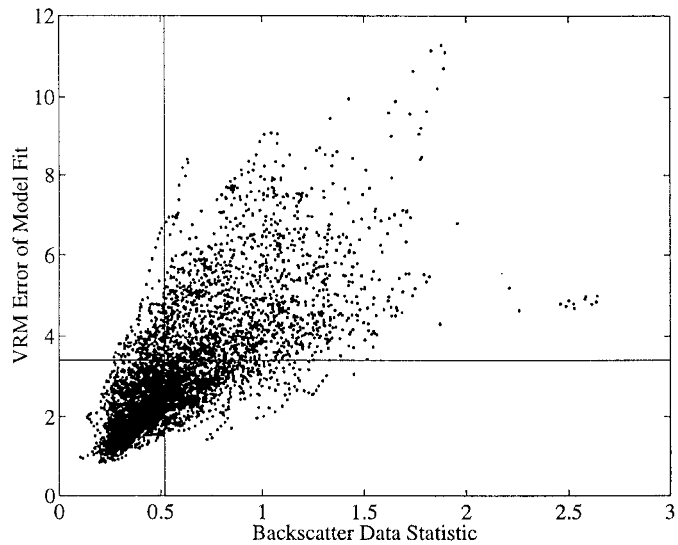

Figure 2: *In general a field that is poorly fit by a 2 parameter KL model has a large value of the statistic, y, defined for the BC algorithm as the standard deviation of the normalized backscatter.*

trends. Specifically that increasing the model order improves model accuracy, and that in general the choice of a particular model is not critical. Unfortunately, as the model order increases so does the computation required to estimate the wind using the model. The compromise for each field-wise estimate, then, is to minimize the number of model parameters while minimizing the model error. We propose to use different model orders for each optimization depending on the wind in the region.

MODEL ORDER SELECTION

Wind field classification algorithms can be used to select models with a minimal number of parameters while keeping the error in an acceptable range. The result increases the computational efficiency of field-wise estimation without significantly increasing the modeling error. In this section we describe two simple wind field classification algorithms that test the hypothesis that a field is poorly modeled by a low-order model. Lacking a clearly superior model using few parameters, the two algorithms described here are developed using the KL model, though adaptation to other models is straightforward.

Both classification algorithms divide wind fields into two disjoint classes based on some computable statistic of the data. Either the region is well modeled by a low-order model (designated θ_0), or it is poorly modeled by the low-order model (θ_1). Comparing a statistic, y, to a threshold, ν, provides the basis for the binary hypothesis test:

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

The Backscatter Classification Algorithm

The backscatter classification (BC) algorithm relates the variability in the wind field to the variability in the corresponding σ^0 data, that is, the statistic for BC is computed directly from the σ^0 measurements. Selecting only smooth wind fields, so the dominant backscatter dependence is due to incidence angle, an average backscatter is computed for each beam as a function of the cross track cell. In the BC algorithm the σ^0 measurements are first normalized with respect to these averages. The statistic, y, for the BC algorithm is defined as the standard deviation of the σ^0 values of all the beams normalized by the average backscatter values.

Fig. 2 displays the relationship between the standard deviation of the normalized backscatter, y, and the vector rms (VRMS) error of the 2 parameter KL model fit to over 5000 NSCAT JPL nudged fields. The strong correlation between the statistic and the VRMS error of the 2 parameter KL model fit can be exploited to estimate the range of the VRMS error given y.

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 2 parameter KL model. The definition of "well" modeled, and the choice of the threshold, depends on the particular application. If, for example, the wind field class θ_0 is defined as wind fields that have a 2 parameter KL model fit VRMS error less than 3.4 m/s (below the horizontal line of Fig. 2), then a threshold ν is selected from the normalized backscatter standard deviation. This selection requires a compromise between the probabilities of correctly classifying a wind field and the probabilities of misclassifying it. With the choice of $\nu = 0.52$ (selected to declare half the wind fields in θ_0 and half in θ_1), for example, the probability of correctly classifying a θ_1 wind field is 86% and the probability of incorrectly classifying a θ_0 wind field 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, a modest increase from the 1.2 m/s error when 40 parameters are used (see Fig. 1).

The Point-wise Classification Algorithm

The point-wise (PC) algorithm classifies wind fields according to patterns in the point-wise ambiguity field, and relies on the statistics of the field directions only, regardless of the wind speed. The point-wise ambiguities are pre-processed to remove spurious ambiguities according to [4] with a probability of removing the correct ambiguity of 10^{-4}. The statistic, y, is defined for the PC algorithm as the minimum DRMS error between a mean wind field and the closest point-wise ambiguity field (minimized over all possible directions of the mean wind fields).

Fig. 3 reveals the correlation between y and the DRMS error in the 2 parameter KL model fit to over 3500 nudged wind fields. The correlation allows estimation of the range of model fit DRMS error given y. If the wind field class θ_0 is defined as wind fields that have a 2 parameter KL model fit DRMS error less than 20°, then the probability of correctly classifying a θ_1 wind field is 93% for $\nu = 14°$. The probability of incorrectly classifying a θ_0 wind field is about 26%. Again, the choice of thresholds would depend on the application. The choice of $\nu = 14°$ divides the declared regions in two nearly equal sized classes, and the average DRMS error of the θ_0 fields is 12°. This is a very modest increase from the 8° DRMS error resulting from the use of 40 model parameters.

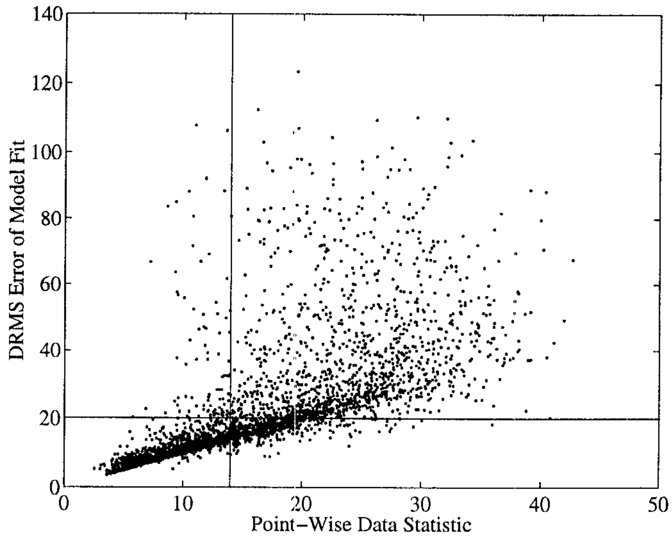

Figure 3: *The DRMS error in a 2 parameter KL model fit to NSCAT JPL nudged wind fields is highly correlated with the PC algorithm statistic, y.*

CONCLUSIONS

Field-wise wind estimation profoundly reduces the number of ambiguities and reduces the computational load of scatterometer wind estimation. Examination of modeling error with four typical models reveals only minor differences. However, the average quality of the model fit is strongly influenced by the number of model parameters used. Increasing the number of model parameters increases modeling accuracy; however, it also increases computational expense. Classification algorithms, such as the BC and PC algorithms 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.

REFERENCES

[1] T. E. Oliphant, "New Techniques for Wind Scatterometry," Master's Thesis, Brigham Young University, 1996.

[2] J. Gunther and D. G. Long, "Models for the Near-Surface Oceanic Vorticity and Divergence," IGARSS, pp. 951-953, August 1994.

[3] 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 Trans. Geosci. Remote Sensing, vol. 28, no. 3, pp. 349-360, May 1990.

[4] T. E. Oliphant and D. G. Long, "Development of Statistical Methods for Eliminating Improbable Wind Aliases in Scatterometer Wind Retrieval," IGARSS, pp. 1715-1717, May 1996.