# An Algorithm to Assess the Accuracy of NASA Scatterometer Data

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

*Abstract*— A simple wind field model can be used to evaluate the accuracy of pointwise ambiguity removal for NASA Scatterometer (NSCAT) data. Errors in pointwise ambiguity removal result in large model-fit errors when the pointwise wind estimates are assimilated into the model. By thresholding the error, regions containing ambiguity removal error can be identified. For many of these regions, the ambiguity selection can be improved using the model-fit field. We have developed a new automated algorithm for evaluating the quality of the pointwise ambiguity selection and for correcting the ambiguity selection. This paper presents this correction algorithm, which is generally applicable to other scatterometers, and the results for NSCAT data.

## INTRODUCTION

Scatterometers do not directly measure the wind; rather the speed and direction of the wind are inferred from the normalized radar cross section ($\sigma^o$) measurements of the ocean surface. The wind is related to $\sigma^o$ via a geophysical model function. However, there is a continuum of possible wind vectors for any particular $\sigma^o$ measurement. Several measurements reduce the set of possible wind vectors to between two and six, each with similar speeds but widely varying directions. An ambiguity removal algorithm must be employed to determine the correct direction.

Another method to determine wind measurements is model based wind retrieval [1]. The wind field model provides a description of the near-surface wind field over the scatterometer measurement swath and is optimized for scatterometer wind retrieval. Wind field models are based on the spatial correlation between wind vectors. The swath is sectioned into rectangular regions and the wind is extracted over the entire region instead of by individual resolution elements. The model relates the components of the wind vector field over this region to a set of model parameters [1] [2]. The models may be either data-driven or dynamics-based.

While [1] used a simple dynamics-driven model, in this paper we adopt a data driven model. We use the Karhunen-Loeve (KL) model since it is known to minimize the basis restriction error [3]. The KL model is derived from the eigenvectors of the autocorrelation matrix of the sampled wind field. Using standard eigenvalue/eigenvector decomposition methods, the KL model is formed as the lower subset of the sorted eigenvectors of the sample autocorrelation matrix. In this paper, the model matrix was subjectively chosen as the first 22 eigenvectors for the tradeoff between modeling error and the ability to locate regions with ambiguity removal errors. We note, however, that there is little performance difference in the algorithm when using between 20 and 30 eigenvectors.

## IDENTIFYING AND CORRECTING POSSIBLE AMBIGUITY REMOVAL ERRORS

The wind field model can be used to improve the pointwise wind product by identifying and correcting ambiguity removal errors. The quality of the fit of the estimated point-wise wind to a simple wind field model over a small area provides a measure of possible ambiguity errors; large errors in the fit suggest possible ambiguity removal errors while small errors suggest a realistic wind field. Areas with errors can be corrected by choosing the alias closest in direction to the model-fit.

A number of wind fields were manually examined to identify ambiguity removal errors. After segmenting the data into small regions, the model was fit in the least-squares sense to the wind fields over each region. Several statistics were calculated for each of these regions. When any of these values are large, the wind field is not realistic and is flagged as containing possible ambiguity removal errors [4].

There are several considerations for the model-fit and the data produced by NSCAT for the implementation of this algorithm. First, noisy scatterometer $\sigma^o$ measurements produce noisy retrieved winds at times. The error in the model-fit can be high even though the ambiguity removal has been done correctly. Second, at low wind speeds, the wind is highly variable which is complicated by the low signal to noise ratio. As a result, the estimate error in low wind speed regions is larger than in high wind speed regions–even for perfect ambiguity removal. These errors can be larger than the thresholds and will cause these regions to be incorrectly identified.

This method locates the boundaries of the regions that have possible ambiguity removal errors. Once the regions with possible ambiguity removal errors are identified, it is natural to try to find a means of correcting these errors. To this end, a technique for correcting the errors has been developed.

An important consideration in making the corrections is that some regions are poorly modeled by the wind field

0-7803-4403-0/98/$10.00 © 1998 IEEE

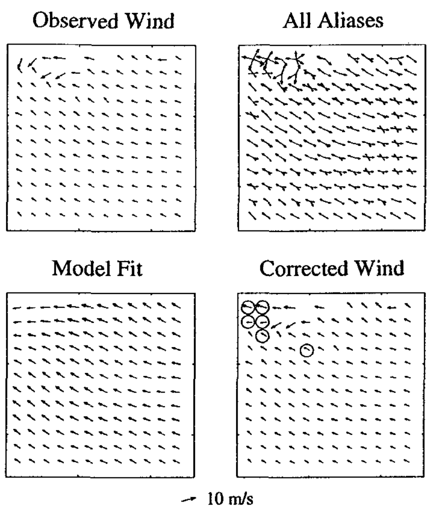


**Figure 1**: The plots for a corrected wind field. The circled vectors are those that were changed according to the algorithm.

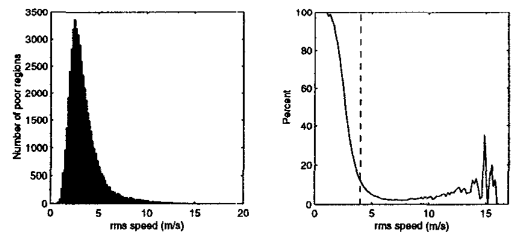


**Figure 2**: (left) Histogram of the rms speed for all regions classified as "poor" in the nine month NSCAT mission. (right) the percent of the total regions which are classified as "poor" at each rms wind speed bin. The vertical dashed line is at 4 m/s.

model resulting in a poor model-fit for the reasons given above. So, regions with considerable numbers of possible ambiguity removal errors are not considered candidates for the correction algorithm. Large numbers of possible ambiguity removal errors are typically low wind speed regions or regions with significant areas of ambiguity removal errors in which the model-fit should not be used as a means of correction.

Thus, for regions identified as having ambiguity removal errors, the correction technique proceeds as follows: determine the number of possible ambiguity removal errors by identifying those that exceed the thresholds; if this number is greater than a given threshold, do not correct any of the vectors for this region; otherwise, choose the alias closest in direction to the model-fit as the corrected wind.

Fig. 1 demonstrates the use of the correction algorithm. As can be seen, the observed wind product contains several ambiguity removal errors. The algorithm chooses the alias that is closest in direction to the model-fit field, producing a subjectively better wind field.

## RESULTS FOR NSCAT

This algorithm was tested on the full nine month NSCAT data set. To be considered candidates for the correction technique in this implementation, 20% or fewer of the vectors in the region can be identified as possible errors. Regions not considered candidates are classified as "poor". Of the 408,069 regions examined, approximately 82% were considered candidates for the correction algorithm. However, only 4% of the individual vectors for these regions were identified as possible errors. Only 10% of these were corrected using this model-based algorithm. For the remaining, the alias closest in direction to the model-fit was the original vector. This implies that only 0.4% of the vectors for these regions were actually correctable. Thus, NSCAT ambiguity removal is almost completely effective for non-poor regions.

Much can be said about the remaining poor regions (18% of the data for the NSCAT mission). Fig. 2 summarizes key statistics for this portion of the data. The majority of these regions (approximately 74%) have rms speeds lower than 4 m/s. The scatterometer does not perform well at such low wind speeds and ambiguity removal algorithms have difficulty distinguishing the correct wind vector at these speeds.

"Poor" regions with rms winds speeds in excess of 4 m/s all contain significant ambiguity removal errors. This represents only 5% of the total data for the NSCAT mission. However, not every wind vector in these regions is in error, a fact verified by a subjective analysis of these regions. Combining this result with the effectiveness of NSCAT for non-poor regions, we conclude that the skill of NSCAT is 95% or better for regions with rms winds speeds greater than 4 m/s.

The performance of NSCAT was also evaluated as a function of time. From Fig. 3, it is clear that the accuracy of NSCAT declines towards the end of the mission. This is most likely a seasonal effect. To see this more clearly, the performance of NSCAT was evaluated over several latitude bands in the Pacific Ocean as described in Fig. 4. The statistics are these bands are described in Fig. 5. The expected variation of wind speed with latitude is clearly evident. There is a strong correlation between the ambiguity removal performance and the rms wind speed, with reduced overall ambiguity removal performance (i.e., more poor regions) at lower wind speeds. Thus, the wind speed distribution in each band affects the ambiguity re-

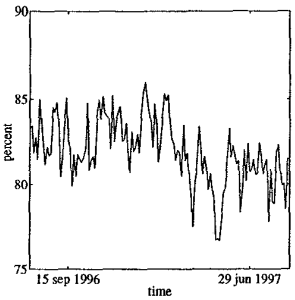


**Figure 3:** The percent of non-poor regions versus time over the nine month NSCAT mission. Each point represents the average computed over approximately two days.

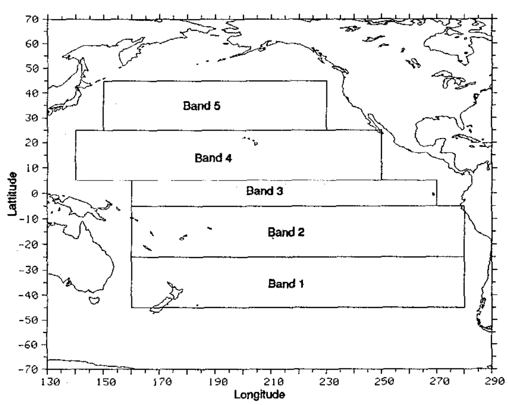


**Figure 4:** Geographical latitude bands in the Pacific.

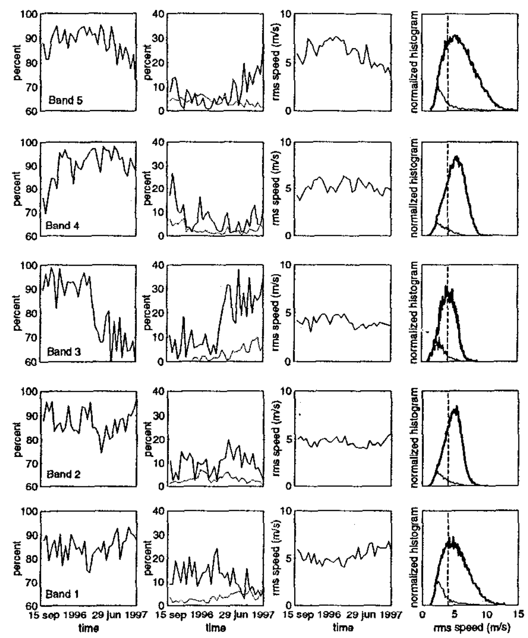


**Figure 5:** (left) Percentage of non-poor regions as a function of time over the NSCAT mission. (left, middle) Percentage of poor regions with an rms wind speed greater than (solid) and less than (dotted) 4 m/s. (right, middle) Average regional rms wind speed as a function of time. (right) Normalized histograms of (bold) all regions and (light) those classified as poor by the QA algorithm. The vertical dashed line is at 4 m/s.

moval performance and seasonal changes in the wind speed distribution results in temporal variations in the ambiguity removal performance. In particular, increased storm activity in the Northern Hemisphere results in increased wind speed with improved ambiguity removal during the winter months in Bands 4 and 5. Similarly, the number of poor regions increases during the Southern Hemisphere summer due to a decrease in the rms wind speed. Because of its low rms wind speed, Equatorial Band 3 is the most sensitive to changes in the mean rms wind speed with a significant drop in the percent of non-poor regions corresponding to a small drop in the rms wind speed at the start of 1997.

## CONCLUSIONS

The detection algorithm works very well in identifying regions with possible selection errors. Once the errors are detected, they can be corrected by choosing the point-wise alias closest the the model-fit. The correction algorithm consistently produces a subjectively more realistic wind field. This technique provides a quick way in which to measure the accuracy of NSCAT ambiguity removal using only NSCAT data. Further research to optimize this technique, such as finding a theoretical basis for determining when the regions are modeled well, is in progress.

## REFERENCES

[1] D.G. Long, "Wind Field Model-Based Estimation of SEASAT Scatterometer Winds," *Journal of Geophysical Research*, pp. 14,651-14688, Vol. 98, No. C8, 1993.

[2] Oliphant, T.E., "New Techniques for Wind Scatterometry," *M.S. Thesis*, 168 pp., Brigham Young University, Aug 1996.

[3] Anil K. Jain, *Fundamentals of Digital Image Processing.* Prentice Hall, 1989.

[4] Amy E. Gonzales and D.G. Long, "A Quality Assurance Algorithm for NASA Scatterometer Ambiguity Removal," *Proc. IGARSS '97*, pp. 246-248, 1997.