SELECTION OF OPTIMUM MEDIAN-FILTER-BASED AMBIGUITY REMOVAL ALGORITHM PARAMETERS FOR NSCAT

Scott Shaffer, R. Scott Dunbar, S. Vincent Hsiao, and David G. Long

Jet Propulsion Laboratory, California Institute of Technology

4800 Oak Grove Dr., Pasadena, CA 91109

Abstract

The NASA Scatterometer, NSCAT, is an active spaceborne radar designed to measure the normalized radar backscatter coefficient (σ_o) of the ocean surface. These measurements can, in turn, be used to infer the surface vector wind over the ocean using a geophysical model function. Because of the nature of the model function, several ambiguous wind vectors result. A median-filter-based ambiguity removal algorithm will be used by the NSCAT ground data processor to select the "best" wind vector from the set of ambiguous wind vectors. This process is commonly known as "dealiasing" or ambiguity removal. The algorithm incorporates a number of selectable parameters such as window size, mode, and likelihood weighting which can be adjusted to optimize algorithm performance. This paper describes the baseline NSCAT ambiguity removal algorithm and the method used to select the set of optimum parameter values. An extensive simulation of the NSCAT instrument and ground data processor provides a means of testing the resulting "tuned" algorithm. This simulation generates the ambiguous wind field vectors expected from the instrument as it orbits over a set of realistic mesoscale wind fields. The ambiguous wind field is then dealiased using the median-based ambiguity removal algorithm. Performance is measured by comparison of the unambiguous wind fields with the "true" wind fields. Results have shown that the median-filter-based ambiguity removal algorithm satisfies NSCAT mission requirements.

1. Introduction

The feasibility of spaceborne scatterometers to make estimates of both wind speed and direction over the ocean was demonstrated by the Seasat scatterometer (Grantham *et al.*, 1982). A scatterometer measures the normalized radar backscatter coefficient (σ_o) of the ocean surface. These measurements can then be used to estimate vector surface wind using a geophysical model function and a wind retrieval algorithm. The NASA scatterometer, NSCAT, is scheduled to fly in the mid 1990's (Martin *et al.*, 1986). The σ_o data will be processed to vector wind estimates at 50 km resolution using a point-wise Maximum Likelihood (ML) technique (Chi and Li, 1988). Because of the harmonic dependence of σ_o on wind direction, though, the retrieval technique is unable to uniquely resolve the wind vector. A set of 2 to 6 possible wind vectors known as ambiguities are determined (Chi *et al.*, 1986). Associated with each ambiguity is a likelihood value which may be used to order the ambiguities. For NSCAT, a median-based ambiguity removal algorithm will be used to select one wind vector from the set of ambiguities that, when successful, is the closest vector to the true wind for each wind resolution cell.

This paper describes the NSCAT ambiguity removal algorithm and the set of adjustable parameters which permit algorithm tuning. Descriptions of the mesoscale wind fields and instrument simulation used in the tuning process are also provided. Methods to evaluate algorithm performance are then developed. The tuning procedure is outlined. Finally, the optimum algorithm parameters and corresponding performance are given.

2. Algorithm Description

The median-based ambiguity removal algorithm was chosen for use by NSCAT because of its simplicity, tunability, and overall performance. It is an extension of the vector median computation given in the appendix, with additional weighting parameters introduced to permit algorithm tuning and performance enhancement. The steps it follows to select the ambiguities are listed below:

1. The algorithm constructs a two dimensional array of wind vectors of sufficient size to contain an entire swath.
2. The array is initialized using the the "Most Likely" wind vectors generated by the ML retrieval algorithm.
3. For each wind vector cell (WVC) in the array, determine the ambiguity, k, at the point (i,j) which minimizes the error function for one of two possible modes of operation:

Mode

$$(0)\quad E_{ij}^k = \frac{1}{(L_{ij}^k)^P} \sum_{m=i-l}^{i+l} \sum_{n=j-l}^{j+l} W_{mn} \cos^{-1}\left[\frac{\mathbf{A}_{ij}^k \cdot \mathbf{U}_{mn}}{|\mathbf{A}_{ij}^k||\mathbf{U}_{mn}|}\right]$$

$$(1)\quad E_{ij}^k = \frac{1}{(L_{ij}^k)^P} \sum_{m=i-l}^{i+l} \sum_{n=j-l}^{j+l} W_{mn} \, |\mathbf{A}_{ij}^k - \mathbf{U}_{mn}|$$

where

$\mathbf{A}_{ij}^k$ = k^{th} ambiguity vector at point (i,j)

$\mathbf{U}_{mn}$ = $(m,n)^{\text{th}}$ vector in the array

W_{mn} = location weight for the $(m,n)^{\text{th}}$ vector

L_{ij}^k = retrieval algorithm likelihood associated with the k^{th} ambiguity at (i,j)

P = likelihood weighting factor

N = algorithm window size

l = $(N-1)/2$ index bound

Wind vector cells not containing data are ignored.

4. Replace $\mathbf{U}_{ij}$ with the vector $\mathbf{A}_{ij}^k$ corresponding to the lowest error function.

5. Repeat steps 3 and 4 until convergence.

The four adjustable parameters which permit algorithm tuning are: mode, window size (N), likelihood weight (P), and location weight (W_{mn}). The mode controls whether wind speed is included in the calculation of the error function (E_{ij}^k). For Mode 0, the vectors A_{ij}^k and U_{mn} are in the direction of the wind ambiguities and have unit length. The mode 1 vectors are in the same direction as mode 0, but their magnitudes corresponds to the ambiguity wind speeds. N can range between 3 and 11 WVCs on a side. W_{mn} controls the relative contribution of each vector in the window. P determines the advantage given to vectors with a higher retrieval algorithm likelihood.

3. Algorithm Simulation

The goal of ambiguity removal tuning is to select the set of parameters which maximizes the algorithm performance. An extensive simulation of the NSCAT instrument and the ground data processor provides a useful data set to use for algorithm tuning and performance assessment. The instrument simulation includes σ_o modeling error, measurement geometry determination error, etc. The baseline antenna polarization mix was used. The ground processing simulation includes σ_o recovery, grouping and wind retrieval using the ML technique. The resulting swath of ambiguous wind vectors has a 50 km resolution and extends 600 km in the cross track direction.

Twelve mesoscale wind fields originally generated by the European Center for Medium-range Weather Forecasting were used to simulate wind structures over the ocean. The fields were constructed using both SASS and conventional data. Additional small scale variability was added (Bevan and Freilich, 1987). The files were selected to span a wide range of wind features including sharp fronts and small-scale cyclones with wind speeds between 0 and 30^+ m/s (Long *et al.*, 1989).

4. Algorithm Performance

Two performance metrics were used to select the optimum set of ambiguity removal parameters, the algorithm skill and the 12 by 12 clumpiness metric. The algorithm skill is the percentage of wind vectors selected by the median-based ambiguity removal algorithm which are also the closest vectors to the true wind direction for a given mesoscale wind file. Clumpiness is the tendency of ambiguity selection errors to occur in the proximity of other ambiguity errors. The 12 by 12 clumpiness metric is a means of estimating the clumpiness of ambiguity errors in a given wind file. It is defined as the percentage of 12 by 12 contiguous WVC regions having greater than 85% successful selection. The higher this percentage for a given algorithm skill, the less clumpy the file is. NSCAT science requirements state that the algorithm skill must be greater than 96.0% and the clumpiness metric be greater than 98.0% in regions of the swath where the true wind speed is between 3 and 30 m/s (Long, 1988). To conform with these requirements, both performance measurements are only calculated in this wind speed region.

5. Tuning Procedure

Performance of the median-based ambiguity removal algorithm depends on the features of the wind field. The twelve wind files were divided into two separate groups, a test group and a withheld group. The test group was used to tune the ambiguity removal algorithm while the withheld group was used to determine the actual performance of the algorithm. Averaging the performance over the six wind files in the test group provides a reasonable measure for ranking the sets of input parameters.

Over one hundred different input parameter configurations for the algorithm were evaluated, spanning most degrees of algorithm freedom; mode, window size, location weight, and likelihood weight. Each configuration was used to select the unambiguous wind vectors for the six mesoscale wind files in the test group. In addition to calculating the algorithm skill and clumpiness metrics, spatial plots of the ambiguity errors and distributions of the errors as a function of wind speed and swath location were made.

6. Tuning Results

The variation in ambiguity removal performance observed in this study indicates that algorithm tuning can have a significant effect. One of the most prominent features is the dependence on algorithm mode. In every case where two configurations differ only by mode, the configuration using mode 1 has a higher algorithm skill; hence, the vector error function performs better than the direction error function. The median-based algorithm also exhibits a strong dependence on window size. The algorithm performance peaks at 7 by 7, though the difference between 5, 7, and 9 is quite small. The algorithm skill drops substantially when the window size is reduced to 3 by 3 WVCs. The results also display a slight dependence upon the likelihood weighting. The algorithm skill for both modes 0 and 1 increases as the likelihood weight increases between 0.0 and 2.0. The skill for mode 1 drops as the likelihood weight increases from 2.0 to 3.0. Mode 0 increases slightly over the same range, but the increases is insignificant. Figure 1 summarizes this observed relationship between mode, window size, likelihood weight, and the average algorithm skill for uniformly weighted configurations over the test group of wind files. A comparison has also been made using the 12 by 12 metric with similar results (see Figure 2).

With the exception of configurations using mode 0 or a window size $N = 3$, the algorithm performance is quite good and fairly uniform, making the selection of a single, "optimum" configuration difficult. Based on Figures 1, 2, and a subjective assessment of algorithm performance, the optimum set of ambiguity removal algorithm parameters has been chosen as follows:

Mode:	1 (Vector error function)
Window Size:	7 by 7 WVC
Window Shape:	Square, Uniform weight
Likelihood Weight:	2 (Advantage to "most likely"

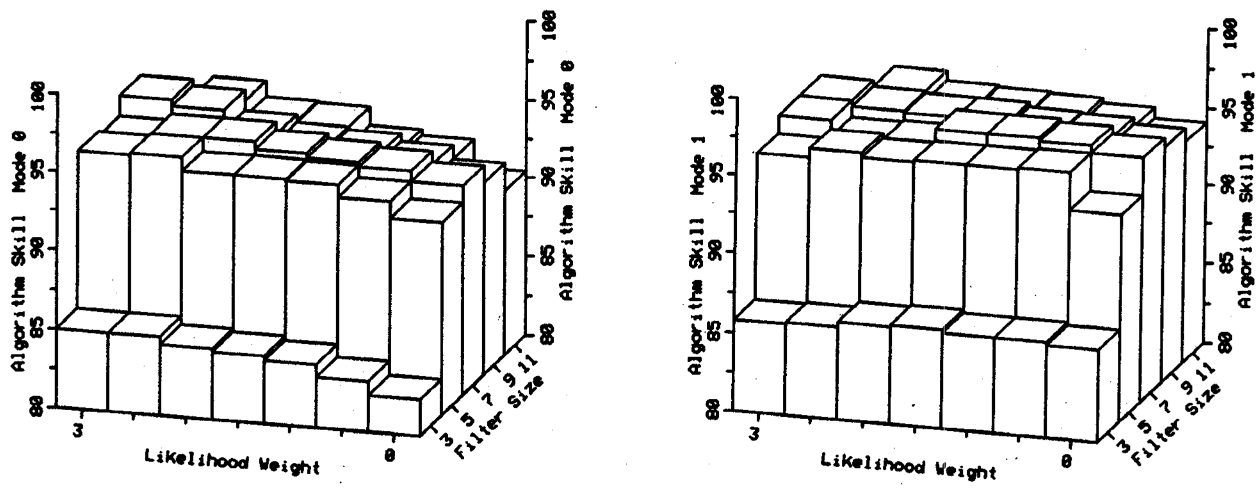

Fig. 1. Histogram of algorithm skill verses filter size and likelihood weight for both mode 0 (left) and mode 1 (right).

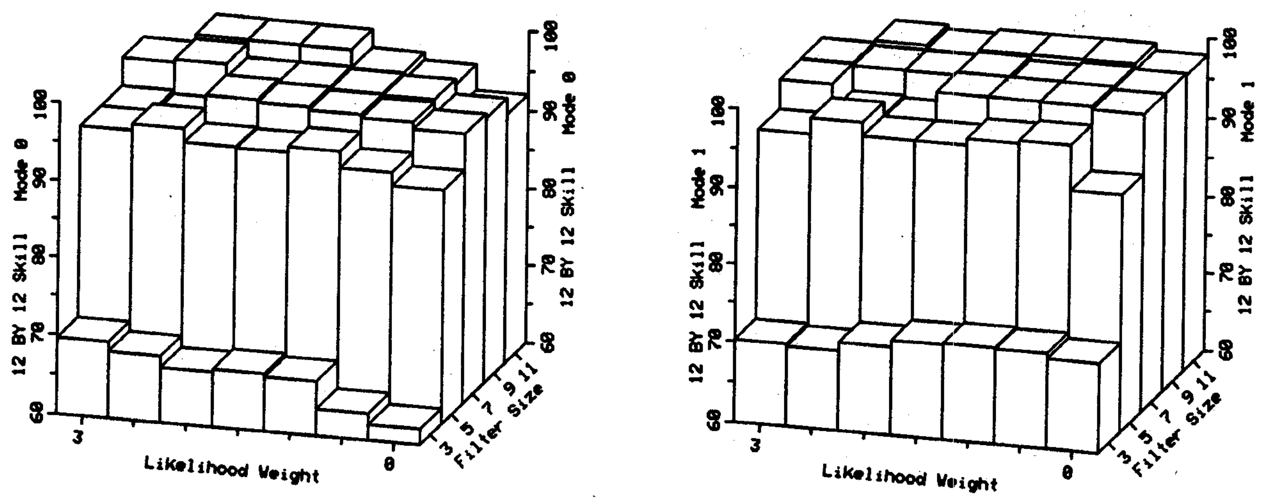

Fig. 2. Histogram of the 12 by 12 clumpiness metric verses filter size and likelihood weight for both mode 0 (left) and mode 1 (right).

With these values, the median-based ambiguity removal algorithm achieves an algorithm skill of 96.7% and a 12 by 12 metric of 98.7% over the test group.

Although several configurations utilizing non-uniform location weighting were also tested, it is clear that any improvement in performance due to location weighting will be small. Detailed analysis of non-uniform configuration performance was complicated by the large number of possible weighting gradients.

7. Algorithm Evaluation

To validate the selection of an optimum set of algorithm parameters. the test was repeated using the withheld group of wind fields. Although the average performance of all algorithm configurations was slightly lower for the withheld data set, the drop was not significant. The optimum parameter configuration selected above ranked 4^{th} using the withheld group of wind fields. The difference in algorithm skill, though, between it and the highest ranked configuration for the withheld group was only .3%. The standard deviation in the algorithm performance averaged over the 6 withheld wind fields is approximately 2%, making this difference insignificant. When both wind field sets were combined, the optimum set of algorithm parameters ranked the highest. The performance of both metrics for the optimum algorithm configuration is summarized in Table 1.

Table 1
Median-Based
Ambiguity Removal Algorithm
Metric Performance

Wind Files	Algorithm Skill	12 by 12 Metric
Test Set	96.7%	98.69%
Withheld Set	96.0%	98.07%
Combined	96.4%	98.38%

8. Conclusion

The median-based ambiguity removal algorithm is a very effective method of selecting the closest vector to the true wind field from a set of ambiguities. It is a simple algorithm to implement and performs well over a large variety of wind fields. An optimum configuration for the ambiguity removal algorithm has been selected. The configuration consists of a 7 by 7 square filter with all locations within the filter weighted equally and an advantage given to ambiguities selected as most likely wind vector by the wind retrieval algorithm. The algorithm skill is 96.7%, with 98.69% of the 12 by 12 contiguous regions having greater than 85% successful selection.

Appendix

The median of a set of N values is defined for N (odd) as the $[(N+1)/2]^{th}$ largest number, i.e. there is an equal number of values greater and lesser in magnitude than the median. Because the selection of a median is determined only by the order of numbers, the median is not affected by extremely large or small values in the data. Therefore, the advantage of a median filter is that it eliminates impulses while preserving edges. On the contrary, a linear filter such as a mean filter or a Hamming window filter smooths (not eliminates) impulses and smears edges.

It is difficult to define the median for circular data such as directions because the order of numbers cannot be specified. One method developed by Mardia defines the median of directional data to be the direction α, such that an equal number of data exist in the half circles $(\alpha, \alpha+180°)$ and $(\alpha-180°, \alpha)$ (Mardia, 1972). This definition is not only cumbersome to work, but also gives multiple solutions. Mardia also pointed out that the circular mean deviation is a minimum when measured from the median direction. i.e. for circular data $a(1), a(2), \ldots\ldots, a(N)$, the median is the value a(m) which minimize the function:

$$S(m) = (1/N)\sum_{i=1}^{N} | a(m) - a(i) | \quad 1 \leq m \leq N.$$

Here the absolute values are defined as positive angles between 0° and 180°. For multi-dimensional data such as vectors and complex numbers, it is even more difficult to define the median. The minimum deviation can be extended to define the median of multidimensional data. For vector data, $\mathbf{V}(i)$, the median is defined as the vector $\mathbf{V}(m)$ which minimizes $S(m)$, where

$$S(m) = \sum_{i=1}^{N} | \mathbf{V}(m) - \mathbf{V}(i) | \quad \text{and} \quad 1 \leq m \leq N.$$

Here the absolute value represents the magnitude of the vectors. Similar to the median of non-circular and circular data, the vector median defined here is not affected by the data of extreme values. The NSCAT median-filter-based ambiguity removal algorithm is a generalization of this technique.

Acknowledgments

The research described in this paper was performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

[1] Bevan, R., and M. Freilich, Wind Fields for Big Sim, JPL NSCAT/MHF 87-15, October 1, 1987

[2] Chi, C-Y., R. Aroian and F. Li, "Simulation Studies for the NASA Scatterometer on NROSS," *Proceedings of the International Geoscience and Remote Sensing Symposium,* pp 1673-1676, Sept. 1986.

[3] Chi, C-Y., and F. K. Li, "A Comparative Study of Several Wind Estimation Algorithms for Spaceborne Scatterometers," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. GE-26, No. 2, pp.115-121, March 1988.

[4] Grantham, W. L., *et al.*, "Performance Evaluation of an Operational Spaceborne Scatterometer," *IEEE Transactions on Geoscience and Remote Sensing,* Vol. GE-20, No. 3, pp. 250-254, July 1982.

[5] Long, D., "NASA Scatterometer System Functional Requirements", JPL Internal Document D-2866, Rev. A, March, 1988

[6] Long, D., *et al.*, "Results of a Study on Polarization Mix Selection for the NSCAT Scatterometer," To be presented at the International Geoscience and Remote Sensing Symposium, July, 1989.

[7] Mardia, K. V., *Statistics of Directional Data,* Academic Press, 1972.

[8] Martin, B. D., *et al.*, "An Overview of the NSCAT/N-ROSS Program," *Proceedings of a Workshop on ERS-1 Wind and Wave Calibration,* 2-6 June, 1986, pp. 143-149 (ESA Sp-262), 1986.