An Advanced Point-wise Ambiguity Selection Algorithm: Application to SeaWinds

David W. Draper and David G. Long
Brigham Young University, Microwave Earth Remote Sensing Laboratory
459 CB, Provo, UT 84602 801-378-4884, FAX: 801-378-6586 draperd@ee.byu.edu, long@ee.byu.edu

***Abstract*- The scatterometer wind estimation process results in several possible wind vectors (ambiguities) at each resolution cell. The current ambiguity selection technique applied to SeaWinds on QuikSCAT data requires outside data as part of the initialization. An advanced ambiguity selection algorithm known as BYU point-wise does not use nudging data; rather, it utilizes a low-order Karhunen Loeve (KL) wind field model to promote self-consistency. In application to a subset of SeaWinds data, BYU point-wise selects 93% of the same ambiguities as the JPL method. On a set of non-storm error regions, BYU point-wise performed subjectively better in 55% of regions and subjectively worse in only 11% of regions. In cyclonic-storm cases, BYU point-wise performed subjectively better in 11% of regions while performing worse in 23% of regions. Thus, BYU point-wise generally produces more consistent results in non-storm regions without the aid of external nudging data.**

I. Introduction

Scatterometers estimate the normalized backscatter cross section (σ^o) of the ocean by transmiting a signal and measuring the power in the return. Ku-band radar signals scatter mainly from wind-induced ocean capillary waves making σ^o a function of the wind. This relationship is known as the Geophysical Model Function (GMF). Due to noise and the symmetry of the GMF, the traditional point-wise estimation process results in several possible solutions at each wind vector cell known as *ambiguities*. An additional *ambiguity selection* step is required to select a unique solution to the wind.

To enhance self-consistency, NASA's Jet Propulsion Laboratory (JPL) *nudges* the initial point-wise wind estimate with numeric weather prediction wind fields. The *point-wise median filter* [3] then iteratively selects the ambiguity at each wvc matching the surrounding flow. Although nudging generally produces consistent wind fields, it also creates a dependence on outside information.

This paper describes a new self-contained point-wise ambiguity selection technique known as *BYU point-wise* which eliminates the use of nudging data. This method utilizes a low-order Karhunen-Loeve (KL) model to create an estimate of the overall wind flow used to initialize the point-wise ambiguity selection process. A correction algorithm then reestimates and corrects data that is not self-consistent. The algorithm is applied to SeaWinds on QuikSCAT data and results from a test data set are presented. We find that BYU point-wise generally selects the same ambiguities as traditional point-wise ambiguity selection with increased performance in regions of low frequency winds.

II. SeaWinds on QuikSCAT Geometry

BYU point-wise is customized to the geometry of SeaWinds on QuikSCAT. This Ku-band scatterometer, launched in mid-1999 and currently operated by NASA, is a dual pencil-beam instrument with a wide 1800 km swath. SeaWinds data is segmented into (approximately) 25 $\times$ 25 km resolution wind vector cells (wvcs) with a total swath size for one revolution (rev) of 76 wvcs in the cross track direction and 1624 wvcs in the along track direction. Along the *swath edges* (the outer 8 wvcs on either side of the cross track), the instrument only receives measurements from the outer beam. This results in low *instrument skill* for those wvcs i.e., a low percentage of correct "most likely" ambiguities. Also, we generally only deal with the inner 72 wvcs due low instrument skill on the outer wvcs.

III. KL Wind Field Model

The KL wind model is formed from the eigenvectors of a wind autocorrelation matrix $\hat{R}$ estimated over a sample set of point-wise ambiguity selected wind fields. The eigenvector matrix is truncated to appropriate number of vectors to give the restricted basis set F. Restricting the basis suppresses high frequency content such as noise and inconsistencies due to ambiguity selection errors. A model fit to the wind field can be written as a linear combination of the restricted basis set, i.e.

$$\mathbf{w}_{opt} = F\hat{\mathbf{x}}_{opt} \quad (1)$$

where $\hat{\mathbf{x}}_{opt}$ contains the coefficients for each parameter of the model.

The model is applied to wind data using the weighted MAP least squares estimate described by [2],

$$\hat{F}^{\dagger}_{opt} = (F^TWF + F^T\hat{R}^{-1}F)^{-1}F^TW \quad (2)$$

where W is a weighting matrix with diagonal elements such that "1" corresponds to valid data cells and "0" corresponds to non-data cells or cells that are ignored. The model parameters are

$$\hat{\mathbf{x}}_{opt} = \hat{F}^{\dagger}_{opt}\mathbf{w} \quad (3)$$

where, $\mathbf{w}$ is the standard vector form of the observed wind field and $\hat{\mathbf{x}}_{opt}$ are the parameters of the model fit field. This model fit optimally interpolates values for cells that have been weighted out.

IV. Overview of BYU Point-wise

The BYU point-wise ambiguity selection algorithm has three main steps: (1) estimation of high instrument skill regions, (2) estimation of low instrument skill regions, and (3) reestimation of inconsistencies. In estimating high instrument skill regions (the inner-beam portion of SeaWinds swath), an initial low resolution wind field is determined by modeling the major flow produced by the first and second ambiguities. The point-wise median filter is initialized from this low order wind field. Low instrument skill regions (the swath edges) are separately estimated by extrapolation from the inner-beam portion. A reestimation routine then locates, masks out, and repairs inconsistencies in the selected wind field. The masking and repairing steps are repeated until the wind field meets certain convergence criteria. The following sections describe each of these steps in more detail.

0-7803-7031-7/01/$10.00 (C) 2001 IEEE

A. Initial Estimate: The Swath Center

An initial step estimates the inner-beam portion of the swath (a region of high instrument skill). In order to enhance self-consistency of the estimated wind, a low order wind field is formed using the KL model. This estimate is created by (1) making an initial least-squares model fit to the first ("most likely") ambiguity field, (2) setting all wvcs to the closest first or second ambiguity to the model fit, (3) masking out cells inconsistent with the model fit, and (4) making a second model fit, weighting out inconsistent cells. The following steps outline how this is applied to SeaWinds data.

1. *Choose First Ambiguities:* Select an initial field of all first ambiguities.
2. *Segment Swath:* Divide the swath into 60 × 60 overlapping (by 75%) regions excluding the swath edges.
3. *Decimate:* Decimate each 60 × 60 region into 9 interleaving 20 × 20 regions to allow the use of a small model.
4. *1st Model Fit:* Estimate each 20 × 20 decimated region with a model fit to the data, ignoring non-data and rain cells.
5. *Choose Closest Ambiguities:* Set each cell to the closest first or second ambiguity to the model fit.
6. *Flag Poor Cells:* Flag all cells exhibiting large errors from the first model fit.
7. *2nd Model Fit:* Reestimate the region with a least squares fit, ignoring non-data cells and flagged cells.
8. *Recreate Original Region:* Recreate the 60 × 60 region by interleaving 20 × 20 second model fits and median filtering.
9. *Save:* Save the center 30 along track rows.
10. *Overlap:* Overlap and average all saved regions.
11. *Point-wise median filter:* Perform the pointwise median filter on the entire swath (excluding swath edges).

B. Estimation of Low Instrument Skill Regions: The Swath Edges

The initial estimate includes only the inner portion of the swath. Due to low instrument skill in the outer-beam region, estimation of swath edges must be done separately. We make no assumption about the correctness of the first ambiguities on the swath edges. Instead, we use the wind flow of the inner-beam region to infer a solution for the outer-beam region. This is done by extrapolating values for the outside cells via the KL model. The following iterative algorithm accomplishes this.

1. *Segment Swath Edges:* Divide swath edges into 16 × 16 wvc overlapping (by 50%) regions with 9 center-swath cross-track rows, and 7 unestimated rows.
2. *Model Fit/Extrapolation:* Estimate each 16 × 16 region with a model fit using only the already estimated cells giving an estimate of the outer 7 cross track rows.
3. *Choose Closest Ambiguities:* Construct a new 16 × 16 field from the closest ambiguity to the model fit.
4. *Insert Selected Cells:* Insert the selected cells if the rms error between the closest ambiguity field and the model fit falls beneath a threshold. This threshold is relaxed for each pass until all wvcs have a unique vector selected.
5. *Point-wise Median Filter:* Perform the point-wise median filter on the entire swath.

C. Reestimating Inconsistent Regions

Errors in ambiguity selection are generally evidenced by inconsistencies in the selected wind flow [1]. In order to correct possible ambiguity selection errors errors, we develop an inconsistency/ambiguity selection error flag followed by a correction algorithm.

Inconsistency Flag To locate possible inconsistencies in the wind flow, we median filter and then average filter the swath. Both of these techniques remove noise, but the average filter smoothes edges, while the median filter preserves edges. Cells that exhibit substantial vector error between the two types of filtered fields are flagged as inconsistent. This "edge detection" scheme locates ambiguity selection errors and wind flow inconsistencies caused by natural phenomena such as fronts.

Ambiguity Selection Error Flag Ambiguity selection errors are generally not limited to isolated points, but groups of connected wvcs. After detecting the location of inconsistencies, an additional step is performed which locates all regions isolated by the cells flagged as inconsistent, and either (1) low wind speed cells, (2) the swath edge, or (3) non-data points (land/ice). Because these cells are isolated by the inconsistency flag, they can be confidently identified as entire regions of ambiguity selection errors and are additionally flagged as possible ambiguity selection errors. These flagged cells are reestimated by the repair process.

Repair Process All regions flagged as inconsistent are reestimated through interpolation using the KL model by the following routine.

1. *Segment Swath:* Divide the swath into 72 × 72 wvc sections overlapping (by 50%) in the along track direction (excluding the outer 2 cross-track cells on either side).
2. *Decimate:* Decimate each region into nine 24 × 24 wvc smaller interleaving regions.
3. *Model Fit:* Model fit the region with a 24 × 24 model, ignoring cells flagged as ambiguity selection errors.
4. *Reestimate Flagged Cells:* Reestimate flagged cells through extrapolation, and select the closest ambiguity to the model fit.
5. *Recreate original region:* Reconstruct each 72 × 72 wvc region and save the center 36 along track rows.
6. *Reconstruct Swath:* Reconstruct the entire swath from the 72 × 36 pieces.
7. *Point-wise Median Filter:* Perform the point-wise median filter on the entire swath.
8. *Iterate Repair Process:* Iterate the repair process until the number of cells changed falls beneath a threshold.

V. Preliminary Results

The BYU point-wise ambiguity selection algorithm is evaluated on revs 1000-1050 of JPL L2B data. Direct comparison of the ambiguities chosen by the BYU algorithm and the JPL product is summarized in Table I.

TABLE I

PERCENTAGE OF AMBIGUITIES CHOSEN IN THE JPL LEVEL 2B (L2B) WIND PRODUCT AND THE BYU POINT-WISE AMBIGUITY SELECTION ALGORITHM.

Type of Ambiguity	JPL L2B Product	BYU Point-wise
1st Ambiguity	66.13%	65.64%
2nd Ambiguity	20.19%	20.21%
3rd Ambiguity	8.41%	8.61%
4th Ambiguity	5.26%	5.54%

Table I shows that the BYU algorithm selects slightly less first ambiguities than the JPL product and slightly more of the other ambiguities. Like the JPL product, the BYU algorithm selects a majority of first and second ambiguities (about 85%). This suggests that the overall flow of the wind is dictated mainly by the first two ambiguities. The other ambiguities are chosen in approximately the same proportion as the JPL product.

Further, BYU point-wise selects over 93% of the same ambiguities as the JPL product in the test data set. Thus, the BYU algorithm gives generally the same result as the JPL product, independently validating JPL's technique.

A. Comparison of Problem Areas

Here, we provide a subjective comparison of the JPL and BYU results on ambiguity selection error and storm regions. The sample data set for both BYU and JPL products was manually examined for possible ambiguity selection errors and cyclonic storms. These areas range in size from a few cells up to hundreds of cells. It is estimated that the ambiguity selection errors cover up to 5% of the data. The total number of ambiguity selection error regions and cyclonic storm regions are 196 and 61 respectively. Each area containing either an ambiguity selection error or a storm is identified as "good" if there subjectively appears to be no errors in the selected field, or "poor" if there appears to be errors present. Table II summarizes statistics for the study.

For most areas containing possible ambiguity selection errors, the BYU algorithm performs better. This is generally true where the JPL's first estimate is corrupted by rain or noise from which the JPL algorithm cannot recover. From manual inspection, areas where both fail usually appear to not be repairable. Areas where only the BYU algorithm fails generally contain fine scale detail or extreme winds. These areas are often smoothed over by the KL model's low-pass effect. In order to correct such areas, they must be detected and processed separately. An important result of this analysis is that only 7% of cyclonic storm regions are identified as poor in both methods simultaneously. This indicates that storm ambiguity selection can be improved by combining the JPL and BYU algorithms. Some examples of manually inspected ambiguity selection error and cyclonic storm regions are given in Fig. 1.

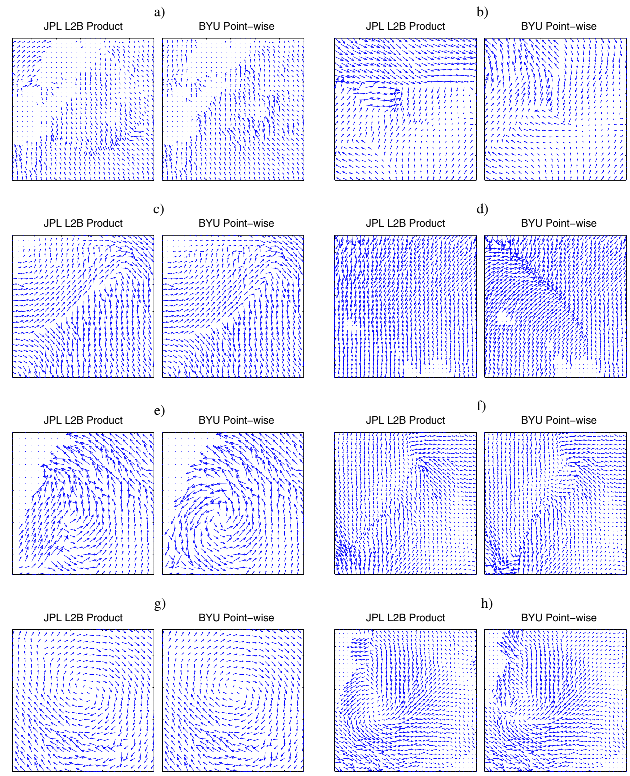

Fig. 1. Examples of the various classifications of regions when comparing the JPL and BYU dealiasing routines. The BYU rendered an area of possible ambiguity selection error a) more consistently b) differently and equally non-consistent c) the same d) less consistently than the JPL product. The BYU algorithm rendered a cyclonic storm e) more consistently f) equally non-consistent g) equally consistent h) less consistently than the JPL product.

VI. Conclusions

BYU point-wise uses a data-driven model rather than a nudging field to produce self-consistent wind fields. In addition, a correction routine locates and corrects further inconsistencies in the wind. BYU point-wise generates the same basic wind flow as the current JPL product and produces less areas of possible ambiguity selection errors in non-storm regions. Only 7% of cyclonic storm regions contain significant errors in both BYU and JPL data sets. These results demonstrate that ambiguity selection can be effectively performed without the use of nudging data. In addition, ambiguity selection can be significantly improved by combining nudging and non-nudging techniques.

TABLE II

Comparison of ambiguity selection errors and storm regions in the JPL and BYU data sets.

	Ambiguity Selection Errors ($\sim$ 5% of data)		Cyclonic Storms	
	BYU		BYU	
JPL	Good	Poor	Good	Poor
Good	0%	11%	59%	23%
Poor	55%	34%	11%	7%

References

[1] A. E. Gonzales and D. G. Long, "An Assessment of NSCAT Ambiguity Removal," *Journal of Geophysical Research*, Vol. 104, No. C5, pp. 11449-11457, 1999.

[2] S. L. Richards, "A Field-wise Wind Retrieval Algorithm for SeaWinds," Masters Thesis, Brigham Young University, 1999.

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