

# An Advanced Ambiguity Selection Algorithm for SeaWinds

David W. Draper and David G. Long, *Senior Member, IEEE*

**Abstract**—SeaWinds on QuikSCAT, a spaceborne Ku-band scatterometer, estimates ocean winds via the relationship between the normalized radar backscatter and the vector wind. Scatterometer wind retrieval generates several possible wind vector solutions or ambiguities at each resolution cell, requiring a separate ambiguity selection step to give a unique solution. In processing SeaWinds on QuikSCAT data, the ambiguity selection is “nudged” or initialized using numerical weather prediction winds. We describe a sophisticated new ambiguity selection approach developed at Brigham Young University (BYU) that does not require nudging. The BYU method utilizes a low-order data-driven Karhunen–Loeve (KL) wind field model to promote self-consistency. Ambiguity selected winds from the BYU method and standard SeaWinds processing are compared over a set of 102 revs. A manual examination of the data suggests that the nonnudging BYU method selects a more self-consistent wind field in the absence of cyclonic storms. Over a set of cyclonic storm regions, BYU performs better in 9% of the cases and worse in 20% of the cases. Overall, the BYU algorithm selects 93% of the same ambiguities as the standard dataset. This comparison helps validate both nonnudging and nudging techniques and indicates that SeaWinds ambiguity selection can be generally accomplished without nudging.

**Index Terms**—Ambiguity selection, Karhunen–Loeve model, nudging, scatterometer, SeaWinds.

## I. INTRODUCTION

RETRIEVING ocean winds is the fundamental application of scatterometer data. Spaceborne scatterometers have proven utility over conventional *in situ* wind measurement techniques due to broad coverage and insensitivity to the time of day or cloud cover. The SeaWinds on QuikSCAT scatterometer, launched in 1999 by the National Aeronautics and Space Administration (NASA), provides global coverage of ocean surface winds on a daily basis. Scatterometers infer the near-surface ocean wind via the geophysical model function (GMF), which relates the vector wind to the normalized radar backscattering cross section ( $\sigma^0$ ) observed by the scatterometer. Because of symmetry in the GMF, the wind estimation process results in several possible wind vector solutions known as ambiguities. A separate ambiguity selection process is required to give a unique wind vector field [1].

In conventional SeaWinds pointwise ambiguity selection, NASA’s Jet Propulsion Laboratory (JPL) uses a method known

as *nudging* to initialize the ambiguity selection process and enhance self-consistency [2]. In nudging, the ambiguity at each wind vector cell (WVC) that most closely matches the flow of numeric weather prediction (NWP) fields is selected as an initial estimate. A modified *pointwise median filter* then iteratively selects the ambiguity at each WVC that best matches the directional flow of the surrounding  $7 \times 7$  WVC region until convergence is reached [3]. The main limitation to the nudging method is that it creates a dependence on the quality of outside information to select a unique solution from the ambiguity sets.

As an alternative to the nudged pointwise winds, Long [4] developed a fieldwise wind estimation technique. In fieldwise wind retrieval, estimates are made on a region-by-region basis using a low-order linear wind field model. The low-order model assumes an inherent correlation between neighboring wind vectors, which restricts the solution to a wind field satisfying the correlation constraints.

In addition to its fieldwise utility, the low-order wind model can also be used to make pointwise ambiguity selection. Gonzales and Long [5] demonstrated that some ambiguity selection errors in pointwise retrieved winds can be corrected by selecting the ambiguity that is closest to a least squares model fit to the JPL winds.

This paper describes a nonnudging pointwise ambiguity selection method for SeaWinds developed at Brigham Young University (BYU). The BYU method uses the low-order Karhunen–Loeve (KL) wind model to create an initial estimate of the overall wind flow which replaces nudging. A model-based correction technique then reselects ambiguities where the data are not self-consistent. The new method is self-contained and computationally efficient.

In Section II, we give an overview of the SeaWinds on QuikSCAT instrument and the KL wind model. In Section III, we describe the BYU technique in detail. Because we lack truth data, the ambiguity-selected JPL winds are used as a reference dataset to which the BYU method is compared in Section IV. We find that the BYU method generally selects the same ambiguities as traditional JPL pointwise ambiguity selection with more self-consistent performance in regions of low-frequency winds. The BYU method, however, is somewhat less able to correctly define fine-scale cyclonic flow than the JPL nudged method. The fact that both the BYU and JPL methods produce approximately the same result simultaneously helps validate both the nudged and nonnuded approaches. It also indicates that quality ambiguity selection can generally be performed without nudging.

Manuscript received May 24, 2002; revised November 28, 2002.

The authors are with the Microwave Earth Remote Sensing Lab, Brigham Young University, Provo, UT 84602 USA (e-mail: draperd@et.byu.edu; long@byu.edu).

Digital Object Identifier 10.1109/TGRS.2003.810228

## II. BACKGROUND

### A. SeaWinds on QuikSCAT

The SeaWinds scatterometer on QuikSCAT was launched in mid-1999 by NASA. The QuikSCAT satellite revolves in a near-polar orbit, covering 90% of the earth daily. SeaWinds' design enables swath coverage of 1800 km in cross track with no nadir gap. Measurements are obtained from an offset dual-feed pencil-beam antenna. The rotating antenna produces two beams that trace out a helical pattern on the surface. The inner (h-pol) beam measures the backscatter at  $46^\circ$  incidence. The outer (v-pol) beam operates at  $54^\circ$  incidence. The backscatter data are binned into (approximately)  $25 \times 25$  km resolution cells 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* (outer 8 WVCs on either side of the cross track), the instrument only receives measurements from the outer beam [6], [7].

The algorithm presented is customized to SeaWinds' geometry. Nevertheless, it can be adapted to other instruments with different sized swaths or WVC resolutions.

### B. Estimation of Pointwise Ambiguities

The scatterometer transmits a radar signal and measures the power scattered from the ocean surface. The returned power is used to calculate  $\sigma^\circ$ . Wind induces ocean-surface capillary waves to which microwave frequencies are particularly sensitive. The returned  $\sigma^\circ$  value is a function of the size and orientation of the waves, and thus a function of the wind [1].

Scatterometer wind retrieval requires multiple measurements from different azimuth angles to help reduce ambiguity in the solution. SeaWinds achieves the azimuthal diversity with fore and aft observations from the two beams. For each WVC, all backscatter measurements whose footprint center lies within the  $25 \times 25$  km cell are combined to create a wind vector estimate. Wind vector estimates are generated using a maximum-likelihood estimation (MLE) technique [1], [6], [8]. The MLE technique minimizes a weighted least squares objective function to find the "most likely" estimate of the wind given the measurements. The objective function is a measure of the error between the observed  $\sigma^\circ$  measurements and  $\sigma^\circ$  values generated by projecting the wind vector through the GMF for each observation. Due to symmetry in the GMF, the objective function generally has several local minima whose corresponding wind vector directions are typically  $90^\circ$  or  $180^\circ$  apart [9]. Each local minimum of the objective function corresponds to a possible correct solution or ambiguity. The ambiguities are ordered according to likelihood, where the most likely ambiguity is referred to as the "first" ambiguity, and the next most likely, the "second," and so forth. SeaWinds processing retains only the first four ambiguities. Because of noise, the "first" ambiguity is not always the closest to the true wind. Thus, ambiguity selection is required to produce a unique wind vector field.

The percentage of correct first ambiguities for a given swath location is known as instrument skill. For SeaWinds, each WVC in the inner beam region has at least two fore-looking observations (one for each beam) and at least two aft-looking observations, which provide sufficient measurement density and azimuthal diversity to afford a high instrument skill. On the

swath edges, poor instrument geometry from the outer beam measurements produces a much lower instrument skill. Also, the outer two cross-track positions on either side (1, 2, 75, and 76) are not always estimated in pointwise retrieval due to very poor instrument geometry. Thus, the effective wind swath width is 72 WVCs.

In regions of high instrument skill, the field of first ambiguities generally contains enough information to estimate the overall flow of the wind. Where the instrument skill is lower, additional information is needed to produce a self-consistent wind field. JPL uses a method known as thresholded nudging to account for the variation in instrument skill. Where the instrument skill is high, a smaller set of ambiguities is used in nudging. On the swath edges where instrument skill is low, all ambiguities may be used [2]. The thresholded nudging method reduces the impact of the nudging field in high instrument skill areas. Overall, the nudging process enhances the self-consistency of the ambiguity selected winds.

### C. KL Wind Field Model

Rather than using outside nudging data, the BYU ambiguity selection method utilizes the KL wind model to enhance self-consistency. The KL wind field model is a linear set of orthonormal basis wind fields derived from a sample set of SeaWinds on QuikSCAT winds [5]. A wind field can be approximated as a linear combination of the basis fields by a least squares fit. The KL model minimizes the basis restriction error given the autocorrelation of the wind [10].

In general, a KL basis set is formed by the eigenvalue decomposition of an autocorrelation matrix. With respect to wind fields, an autocorrelation matrix is estimated over a sample set of ambiguity selected wind fields by

$$\hat{R} = \frac{1}{M} \sum_{n=1}^M \mathbf{w}_n \mathbf{w}_n^T \quad (1)$$

where  $M$  is the number of  $N \times N$  regions examined, and  $\mathbf{w}_n$  is the standard vector form of the  $n$ th wind field. The *standard vector form* is created by stacking the column scanned U and V wind components into a  $2N^2$  length vector. The dimension of the empirical autocorrelation matrix  $\hat{R}$  is  $2N^2 \times 2N^2$ . The basis set is extracted by taking the eigenvalue ( $\Lambda$ ) decomposition of  $\hat{R}$  where

$$\hat{R} = S \Lambda S^T. \quad (2)$$

The diagonal elements of  $\Lambda$  are the eigenvalues, and the columns of  $S$  are the eigenvectors or basis fields of the KL model.

Eigenvalues are ordered from high to low. Because of the generally red spectrum of the wind [11], larger eigenvalues represent eigenvectors with lower spatial frequency. Restricting the basis suppresses high-frequency content due to noise and inconsistencies resulting from ambiguity selection errors. The basis matrix  $S$  is truncated to an appropriate number of vectors to give the restricted basis set  $F$ . A model fit to the wind field can be written as a linear combination of the restricted basis set, i.e.,

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

where  $\hat{\mathbf{x}}$  contains the coefficients for each parameter of the model [5].

### III. OVERVIEW OF BYU POINTWISE AMBIGUITY SELECTION

In this section, we describe the BYU algorithm in detail. To avoid nudging, we utilize the KL model to determine an initial ambiguity selection over the high instrument skill inner beam portion of the swath. The pointwise median filter is initialized from this initial selection. The swath edges, where the instrument skill is lower, are separately estimated by extrapolation of the inner beam wind flow. A correction routine locates and repairs inconsistencies in the selected wind field. The masking and repairing steps are repeated until the wind field meets convergence criteria. Fig. 1 outlines the steps involved in the BYU method. Section III-A describes the initial estimate. Section III-B gives an overview of the pointwise median filter. Section III-C describes ambiguity selection of the swath edges. Lastly, Section III-D details the repair process.

#### A. Initial Estimate

1) *Methodology*: The high instrument skill of the inner beam portion of the SeaWinds swath affords estimation of the main wind flow using the first ambiguities. Although we cannot calculate the actual instrument skill without truth data, a reasonable estimate is the average percentage of first ambiguities selected by the JPL method. On average, the JPL ambiguity selection method selects over 70% of first ambiguities in the inner beam portion of the swath. In the outer beam portion of the swath, only about 35% first ambiguities are selected. Because a vast majority of WVCs in the inner beam region have “correct” first ambiguities, the inner beam first ambiguity field can be utilized to initialize the ambiguity selection process.

Incorrect first ambiguities can be either isolated or clustered. In general, isolated incorrect first ambiguities can be corrected by simply applying the pointwise median filter to the field of first ambiguities. Small to large clusters of incorrect first ambiguities on the order of half the filter size or larger remain incorrect after median filtering. These errors can significantly affect the ambiguity-selected wind flow.

Rather than simply using the field of first ambiguities for an initial ambiguity selection, the BYU method replaces first ambiguities that oppose the main flow by second ambiguities. Since the JPL method selects over 90% first and second ambiguities in the inner beam region, a first/second ambiguity field yields better final solution. Further, where neither the first or second ambiguity produces a self-consistent solution, the ambiguity selection may be replaced by a third or fourth ambiguity to create an even better initial estimate. In this way, ambiguities are inserted into the swath based on priority. This priority-based ambiguity selection is the basic idea behind the BYU initial estimate.

In order to replace clusters of incorrect first ambiguities with second ambiguities, a low-order constraint is placed on the wind flow of the first ambiguity field. To do this, an initial low-order KL model fit is made to the first ambiguity field. High spatial frequency content caused by small clusters of incorrect first ambiguities are smoothed by the model fit. All WVCs are set to the nearest first or second ambiguity. This step replaces isolated vectors and small to medium clusters of incorrect first ambi-

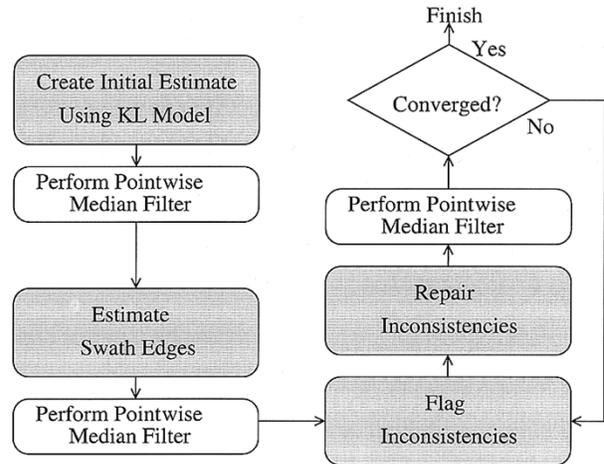


Fig. 1. Flowchart of the BYU ambiguity selection method.

guities with second ambiguities. The order of the KL model dictates the “smoothness” of the model fit, and thus influences the cluster size that is replaced by second ambiguities. Wind vector cells exhibiting large errors between the first/second ambiguity field and the model fit are flagged. Next, a second low-order model fit is performed to the first/second ambiguity field, weighting out flagged WVCs. Where neither first or second ambiguities are consistent with this model fit, the closest third or fourth ambiguity may be chosen.

Although this initial estimate corrects small to medium clusters of incorrect vectors (depending on the KL model order), large regions where the first ambiguity field is incorrect must be repaired separately with a more sophisticated routine (see Section III-D).

2) *Detailed Description of the Initial Estimate: Application to SeaWinds Swath*: The large swath size prohibits applying a KL model to the entire swath due to computational concerns. As a result, the swath is divided into  $60 \times 60$  WVC sections overlapping by 75% in the along-track direction. The  $60 \times 60$  region size is chosen because it spans the entire inner beam portion of the cross track. A square region size is chosen for historical reasons [4], [5]. Since the model fit requires the inversion of a matrix that has on the order of  $N^4$  elements where  $N$  is the width of a region, each  $60 \times 60$  section is additionally decimated into nine interleaved fields of  $20 \times 20$  WVCs. This is equivalent to increasing the measurement spacing from 25 km to 75 km for each field. Segmentation and decimation allow the use of a reasonably sized model to reduce the computational expense of the problem.

For each  $20 \times 20$  decimated region, a model fit is made to the first ambiguity field. Nonoccean WVCs and WVCs that contain significant rain contamination as determined by the L2B rain flag [12] are weighted out of the model fit. The model fit is computed via a weighted regularized least squares estimate

$$\hat{F}_{\text{opt}}^{\dagger} = (F^T W F + \Lambda_F^{-1})^{-1} F^T W \quad (4)$$

where  $\Lambda_F$  contains the eigenvalues of the basis vectors in  $F$ , and  $W$  is a weighting matrix with diagonal elements of “1” corresponding to valid data cells and “0” corresponding to nondata

cells or cells that are to be ignored. The coefficients for the basis fields  $\hat{\mathbf{x}}_{\text{opt}}$  are

$$\hat{\mathbf{x}}_{\text{opt}} = \hat{F}_{\text{opt}}^\dagger \mathbf{w} \quad (5)$$

where  $\mathbf{w}$  is the standard vector form of the wind field. The model fit field  $\hat{\mathbf{w}}_{\text{opt}}$  is constructed by the equation

$$\hat{\mathbf{w}}_{\text{opt}} = F \hat{\mathbf{x}}_{\text{opt}}. \quad (6)$$

This model fit is regularized by the eigenvalues to give a realistic solution to points that have been weighted out.

After a model fit is made to the first ambiguity field, all WVCs in the  $20 \times 20$  region are set to the nearest first or second ambiguity to the model fit. The directional and vector error between the new ambiguity-selected field and the model fit are computed for each WVC over the region. Cells where the directional error exceeds  $45^\circ$  or the vector error is greater than the average wind speed of the region are flagged as poor. A second higher order model fit is then made in which vectors flagged as poor are weighted out. The second model fit interpolates new values for the flagged cells.

After second model fits are made to all nine fields corresponding to a  $60 \times 60$  region, the entire  $60 \times 60$  region is reconstructed by interleaving the second model fit fields. The  $u$  and  $v$  components of the  $60 \times 60$  reconstructed field are then median filtered to ensure consistency among the interleaved fields. To rebuild the entire swath, the center 30 along-track rows are saved. All along-track sections are likewise estimated, and each section is overlapped and window averaged. A simple triangular averaging window is used in our implementation, although the shape of the window is not critical. This creates a low-resolution initial wind field close to the wind flow dictated by the first and second ambiguities. Lastly, each WVC is set to the nearest ambiguity to the low-resolution wind field. All ambiguities are included in the selection process. The result is an ambiguity-selected field dictated mainly by the flow of the first and second ambiguities. The steps involved with creating the initial estimate are shown in Fig. 2.

### B. Pointwise Median Filter

The initial ambiguity-selected field replaces the nudging data used in traditional pointwise estimation. Next, the pointwise median filter is employed to insure self-consistency among each selected ambiguity and its neighbors. The pointwise median filter selects the ambiguity that minimizes the directional error between it and the surrounding initial wind vectors, i.e.,

$$\hat{n} = \arg \min_n \sum_{k=i-3}^{i+3} \sum_{l=j-3}^{j+3} |\phi_{ij}^n - \Phi_{kl}|_{\{[0^\circ, 180^\circ]\}} \quad (7)$$

where  $\Phi_{kl}$  are directions of the surrounding wind vectors, and  $\phi_{ij}^n$  is the direction of the  $n$ th ambiguity at WVC  $i, j$ . The new chosen ambiguity,  $\hat{n}$  replaces the initial estimate for the first iteration and replaces the previous selected ambiguity on each successive iteration. The pointwise median filter is iterated until convergence is reached. This is the same pointwise median filter implemented by JPL, although JPL initializes the filter with NWP model fields.

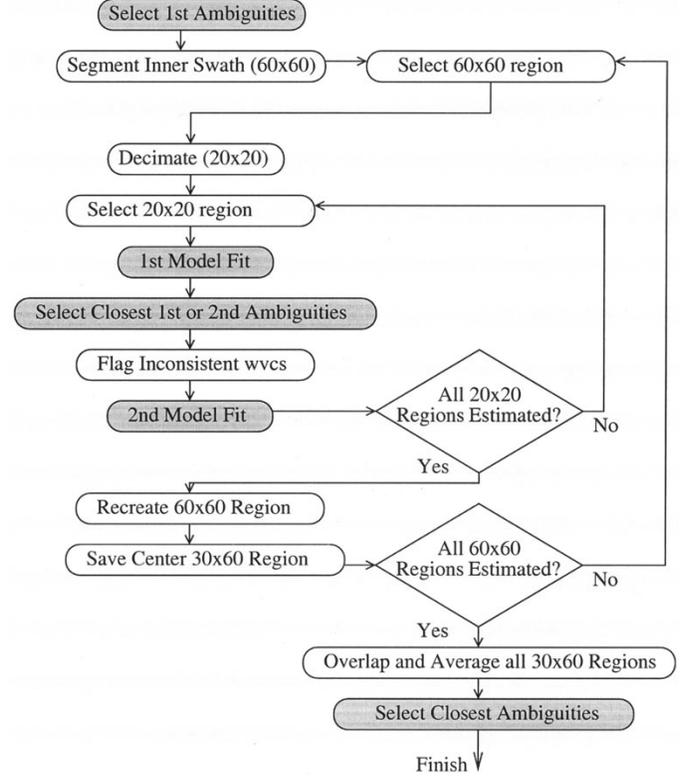


Fig. 2. Flowchart summarizing the initial estimate. The shaded bubbles show the most important steps.

### C. Estimating Swath Edges

The initial selected ambiguity field only includes the inner beam portion of the swath. Due to low instrument skill in the outer beam region, ambiguity selection of swath edges must be performed 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 by extrapolating values for the outer cells via the KL model.

The swath edges, including the outer beam WVCs and the outer edge of the inner beam, are divided into  $16 \times 16$  WVC regions overlapping by 50% in the along-track direction. These regions contain nine cross-track rows assigned unique ambiguities by the previous steps and seven cross-track rows of unselected outer beam WVCs. The choice of  $N = 16$  is arbitrary, but provides a good compromise between computational efficiency and inclusion of inner beam data. The outer cross-track row is not included because wind retrieval is not performed there. A low-order model fit is made to each  $16 \times 16$  region using only the inner beam wind vectors. This yields an estimate of the outer seven cross-track rows via interpolation of the model fit. A new  $16 \times 16$  field is constructed from the closest ambiguity to the model fit. If the rms error between the closest alias field and the model fit falls beneath a threshold, the new ambiguity selections are inserted into the swath. This threshold is set at 1.5 m/s and is relaxed (increased by 1.5 m/s) for each pass until all WVCs have a unique vector selected. Thus, the best solutions are inserted first and used to influence subsequent estimates. The swath is again passed through the pointwise median filter after estimating the swath edges. The estimation of the swath edges is summarized in Fig. 3.

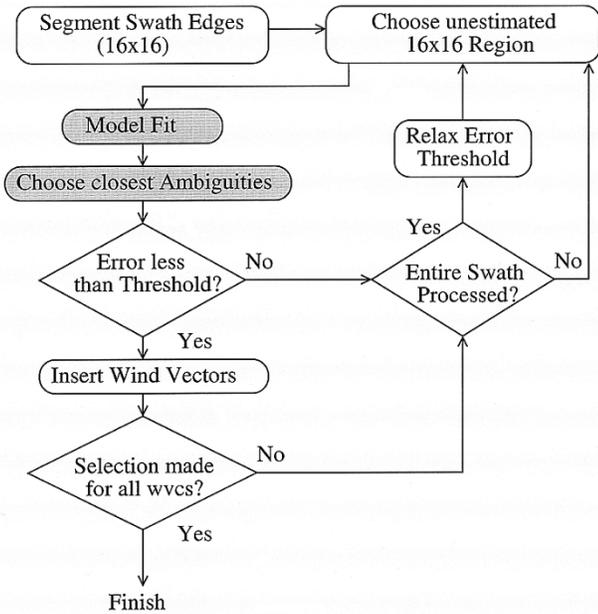


Fig. 3. Flowchart summarizing the estimation of swath edges. The shaded bubbles show the crucial steps.

#### D. Repairing Inconsistencies

The previous steps yield a unique ambiguity at each WVC. Although this selected ambiguity field could be the final product, due to possible large clusters of incorrect first ambiguities in the initial estimate, there still may be areas of significant ambiguity selection errors. Clusters of ambiguity selection errors are generally characterized by  $90^\circ$  to  $180^\circ$  shifts in the selected wind flow. The transition between a cluster of ambiguity selection errors and the correctly selected wind is typically sharp, and it results in unnatural and inconsistent wind flow. In order to correct such possible errors, we develop an inconsistency flag followed by a correction algorithm.

1) *Inconsistency Flag*: Wind vector cells neighboring the sharp transitions from the correctly selected winds to the ambiguity selection errors can be identified by evaluating the consistency of the wind flow [13]. In order to flag inconsistent WVCs, several image processing tools are used.

A true median filter is applied to the  $u$  and  $v$  components of the ambiguity-selected wind (the value at each WVC is replaced by the median of the values around it within a  $3 \times 3$  WVC region). This technique reduces the noise in the wind field while preserving edges caused by inconsistent wind flow. The median wind field is then filtered with a  $3 \times 3$  WVC averaging filter (the components of each cell are replaced by the average of those around it). This technique reduces noise, but smooths edges. The median field and the averaged field are compared. Cells are flagged as *inconsistent* where the normalized vector difference between the averaged and median field is large (greater than 0.25). Fig. 4 shows an example of the steps involved with flagging ambiguity selection error edges.

The WVCs flagged as inconsistent indicate the location of ambiguity selection error edges. In order to repair areas of ambiguity selection error, we flag WVCs isolated by the inconsistency flag and other features such as low wind speed areas, the swath edge, or land. The “filling in” of such isolated regions is

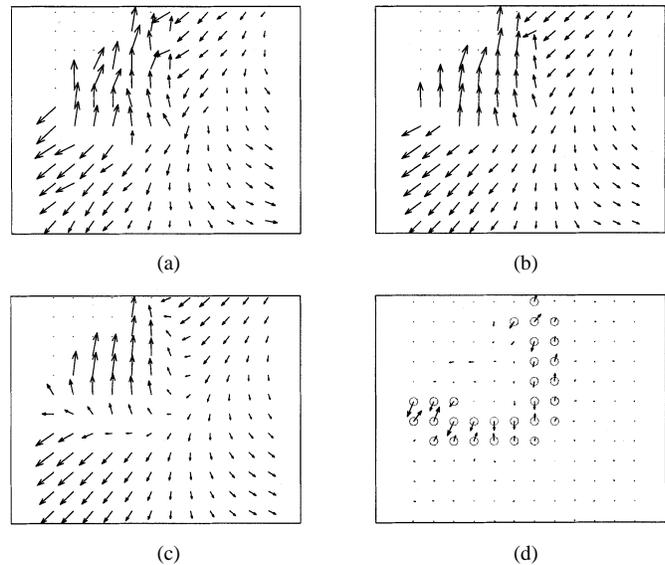


Fig. 4. Example of the steps involved with flagging ambiguity selection error edges. (a) Ambiguity-selected wind, (b) median-filtered wind, (c) average-filtered wind, and (d) difference field. Where the errors are large, the WVCs are flagged as inconsistent (circled).

accomplished through modified dilation and erosion techniques (see the Appendix for details).

2) *Selecting a Consistent Wind Field*: All isolated ambiguity selection error regions are reestimated through interpolation using the KL model. The swath is divided into  $72 \times 72$  WVC sections overlapping by 50% in the along-track direction. This region size is chosen because it includes the entire cross track except the outer two cells on either side that do not always contain estimated ambiguities. Each region is decimated into nine  $24 \times 24$  WVC smaller interleaving regions to reduce computational expense. This region size is larger than the  $20 \times 20$  WVC region size used earlier because the outer beam portion of the swath is now included in the estimation process. Each region is model fit using a truncated  $24 \times 24$  KL model, weighting out flagged cells. Flagged cells are reestimated through interpolation, and the ambiguities closest to the model fit are selected. Each  $72 \times 72$  WVC region is then reconstructed, and the center 36 along-track rows are kept. The outer along-track rows are discarded because they are more likely to contain modeling errors. The 50% overlap provides an estimate of the entire swath from the individual pieces. The swath is reconstructed from the  $72 \times 36$  pieces, and pointwise median filter ambiguity selection is performed on the entire swath.

3) *Iteration of the Repair Process*: For each iteration of the repair process, the output field is compared to the input field. When the number of changing cells falls beneath a threshold (set at 20) or a maximum number of iterations is reached, the algorithm stops. Most swaths converge in about ten iterations. The repair process is summarized in Fig. 5.

#### IV. COMPARISON OF BYU AND JPL AMBIGUITY SELECTION

Without an extensive truth dataset, assessing the quality of the BYU ambiguity selection is difficult [5]. As a result, we use the JPL level 2B (L2B) nudged winds as a reference dataset.

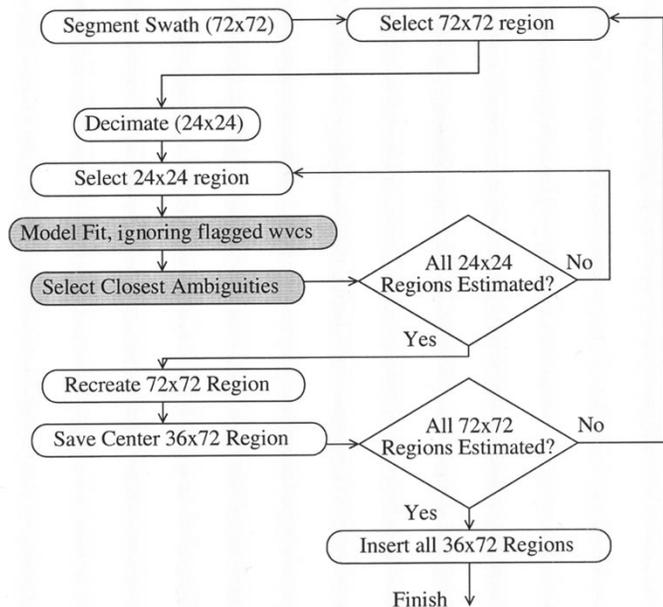


Fig. 5. Flowchart summarizing the repair process. The shaded bubbles show the most important steps.

We perform three comparisons of BYU to standard JPL processing on a set of 102 revs. Ambiguity selection is performed on revs 1000–1050 and revs 6000–6050 of QuikSCAT data. In Section IV-A, we perform a direct comparison of the ambiguities selected by both techniques. In Section IV-B, we perform a quality assurance analysis with a model-based technique [13]. In Section IV-C, we present statistics based on manual inspection of the data.

A. Direct Comparison of Selected Ambiguities

To quantify the similarity of the BYU ambiguity-selected winds to the JPL reference dataset, we compare the percentage of each ambiguity selected by both techniques over the test dataset in Table I. The BYU algorithm selects slightly fewer 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%), which is consistent with our assumption 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, we compare the percentage of ambiguities similarly selected by both the BYU and JPL algorithms and the percent of each ambiguity changed in Table II. From Table II, the ambiguity selection is the same for both BYU and JPL in 93% of WVCs. Thus, the BYU algorithm gives generally the same wind vector solution as the JPL product, which simultaneously helps validate both BYU and JPL techniques. This result is significant because the BYU method was independently developed and not tuned against the JPL reference data. It suggests that the SeaWinds noisy ambiguity sets contain a sufficient percentage of correct first ambiguities to allow self-contained ambiguity selection that is very close to a nudged solution.

Next, we compare the percentage of similar ambiguities selected as a function of cross-track position and wind speed. To

TABLE I  
PERCENTAGES OF AMBIGUITIES CHOSEN BY THE JPL L2B PRODUCT AND THE BYU ALGORITHM

| Category                  | JPL L2B product | BYU method |
|---------------------------|-----------------|------------|
| 1 <sup>st</sup> Ambiguity | 65.47%          | 64.80%     |
| 2 <sup>nd</sup> Ambiguity | 20.58%          | 20.76%     |
| 3 <sup>rd</sup> Ambiguity | 8.63%           | 8.89%      |
| 4 <sup>th</sup> Ambiguity | 5.32%           | 5.55%      |

TABLE II  
PERCENTAGE OF AMBIGUITIES CHOSEN THE SAME AND CHANGED FROM FIRST TO SECOND OR FROM FIRST OR SECOND TO THIRD OR FOURTH FOR THE SAMPLE SET OF BYU AND JPL AMBIGUITY-SELECTED WINDS

| Category   | Percentage |
|--|------------|
| Same ambiguity selected for JPL and BYU  | 93.03%     |
| JPL selected 1 <sup>st</sup> ambiguity and<br>BYU selected 2 <sup>nd</sup> ambiguity                                       | 1.41%      |
| JPL selected 2 <sup>nd</sup> ambiguity and<br>BYU selected 1 <sup>st</sup> ambiguity                                       | 1.13%      |
| JPL selected 1 <sup>st</sup> or 2 <sup>nd</sup> ambiguity and<br>BYU selected 3 <sup>rd</sup> or 4 <sup>th</sup> ambiguity | 2.23%      |
| JPL selected 3 <sup>rd</sup> or 4 <sup>th</sup> ambiguity and<br>BYU selected 1 <sup>st</sup> or 2 <sup>nd</sup> ambiguity | 1.74%      |

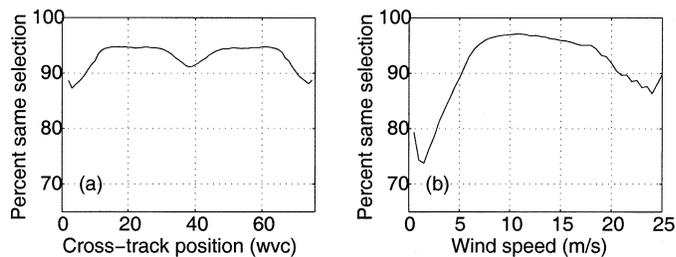


Fig. 6. Percentage of individual WVCs selected the same for both JPL and BYU methods per (a) cross-track position and (b) average wind speed of the ambiguities.

assign a unique wind speed to a WVC, we average wind speeds of all ambiguities. Since all ambiguities at a WVC generally have similar wind speeds, the averaging does not significantly affect the results. Fig. 6 displays the percentage of ambiguities chosen the same per cross-track position and wind speed. Fig. 6 indicates that the ambiguity selection differs the most along swath edges and at nadir. This is intuitive, because more ambiguities are generally produced in these regions, allowing for a higher probability of a different selection. Also, low and high wind speed data (which tend to be the noisiest) have a higher percentage of changed WVCs.

B. Quality Assurance Assessment

In this section, we present a quality assurance (QA) assessment of the ambiguity-selected data based on the self-consis-

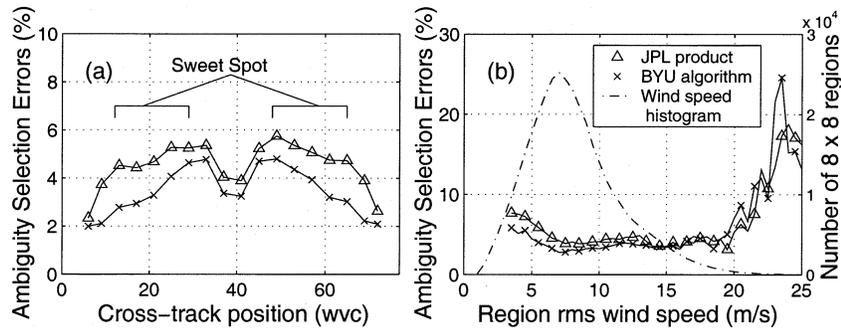


Fig. 7. Percentage of  $8 \times 8$  regions flagged as containing possible ambiguity selection errors by the QA analysis method (a) per cross-track position and (b) rms wind speed for both JPL and BYU methods (left axes). A histogram of wind speeds is shown in (b), indicating the number of regions inspected per rms wind speed bin (right axis).

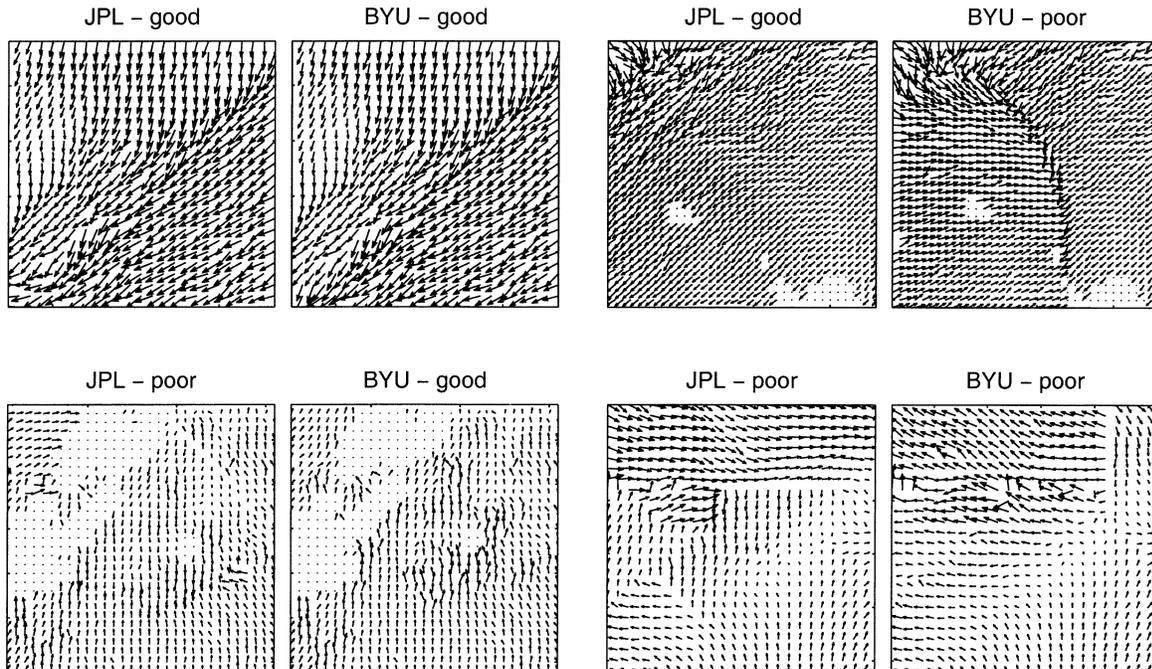


Fig. 8. Examples of the various classifications of regions *without* cyclonic features when comparing the JPL and BYU ambiguity selection routines.

tency of the winds (see [13]). This method compares the selected wind to the low-order KL model fit over each  $8 \times 8$  region in the swath. Where the wind field exceeds certain empirically determined error thresholds from the model fit and other criteria, the  $8 \times 8$  WVC area is identified as a possible ambiguity selection error. The use of this method is somewhat incestuous, because KL models are used in both the BYU ambiguity selection and QA techniques. However, because the BYU and JPL winds are derived from the same noisy ambiguity sets, the performance of the QA method should be consistent for both BYU winds and the JPL reference winds. The QA analysis is only valid for wind speeds greater than 3.5 m/s.

The QA method is applied to both JPL and BYU ambiguity-selected test sets. Of JPL ambiguity-selected winds, 4.5% of  $8 \times 8$  WVC regions inspected by the QA method are determined to contain possible ambiguity selection errors. Of the BYU wind vector selections, about 3.4% of all regions are determined to contain possible ambiguity selection errors. These numbers suggest that the overall consistency of the BYU method is somewhat better than JPL.

Next, we examine the percentage of QA-determined ambiguity selection errors as a function of cross-track position and region rms wind speed (see Fig. 7). The region rms wind speed is defined as

$$\left( \frac{1}{n} \sum_{i,j} U_{i,j}^2 \right)^{1/2} \quad (8)$$

where  $U_{i,j}$  is the wind speed at cell  $\{i, j\}$  of the region, and  $n$  is the number of valid wind data WVCs in the region (WVCs over ocean). The BYU method performs particularly better in the "sweet spot" (off-nadir region characterized by high instrument skill) and at low to moderate wind speeds (4–10 m/s). Many of the "sweet spot" cases occur in rain-corrupted areas where thresholded nudging fails to select ambiguities consistent with the overall flow of the nonrain-corrupted wind [13]. At moderately high wind speeds (10–18 m/s) both methods perform about the same. At extreme winds (18+ m/s) the JPL method performs better. Extreme winds, however, only occur a small percentage of the time.

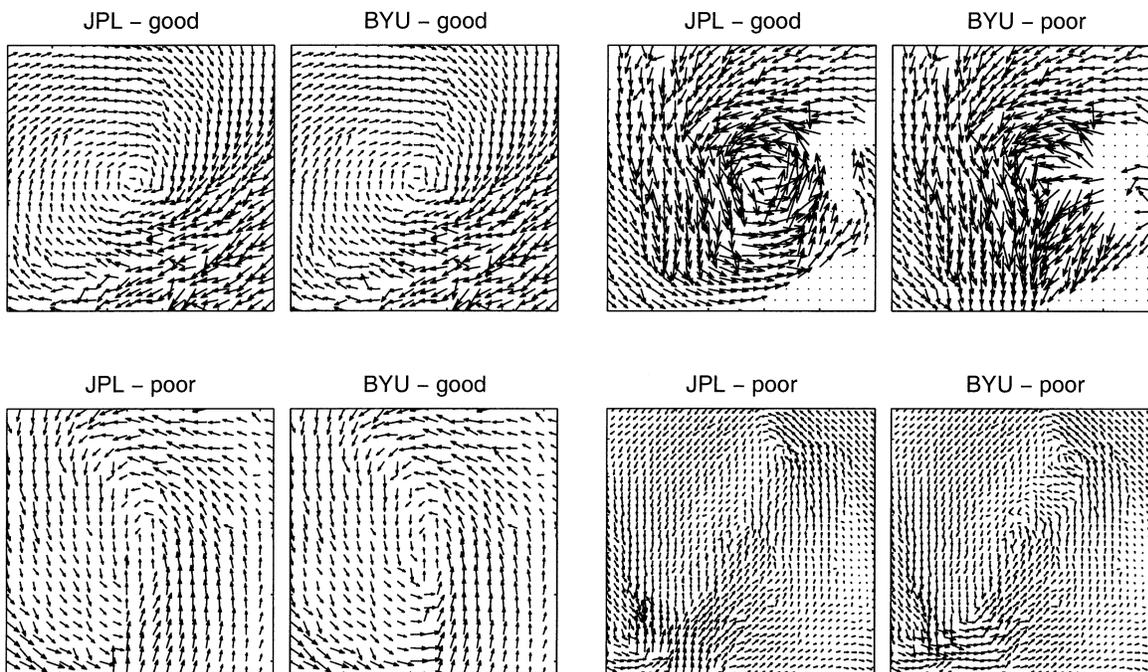


Fig. 9. Examples of the various classifications of regions with cyclonic features when comparing the JPL and BYU ambiguity selection routines.

TABLE III

TOTAL NUMBERS AND PERCENTAGES OF REGIONS SUBJECTIVELY RATED “GOOD” OR “POOR” IN BOTH BYU AND JPL DATASETS FOR  $38 \times 38$  WVC AREAS WITHOUT AND WITH CYCLONIC FEATURES. PERCENTAGES ARE CALCULATED FOR REGIONS WITH AND WITHOUT CYCLONES SEPARATELY

| JPL  | Without Cyclone |            | With Cyclone |            |
|------|-----------------|------------|--------------|------------|
|      | BYU Good        | BYU Poor   | BYU Good     | BYU Poor   |
| Good | 6154 (94.8%)    | 63 (1.0%)  | 91 (56.2%)   | 32 (19.7%) |
| Poor | 147 (2.3%)      | 125 (1.9%) | 15 (9.3%)    | 24 (14.8%) |

C. Analysis of Storm and Nonstorm Regions

This section provides a subjective comparison of BYU ambiguity selection with the JPL selection in regions with and without cyclonic features. For each revolution of test data, the swath is segmented into  $38 \times 38$  WVC sections. Because cyclones represent areas of high spatial frequency, regions containing cyclonic features are analyzed separately. A  $38 \times 38$  region is rated “good” if there is a visually consistent wind flow. If the region contains an area of visually inconsistent flow, it is rated “poor.” These ratings are applied to each  $38 \times 38$  WVC region of the 102 swath test dataset. Examples of each type of region are found in Figs. 8 and 9. Table III summarizes statistics on the regions for the study.

For noncyclonic areas, both JPL and BYU methods produce very visually consistent flow in about 95% of the test dataset. The BYU algorithm on average produces somewhat more non-cyclonic “good” regions. From visual observation, JPL poor areas often occur where the first ambiguity field is corrupted by rain from which the JPL algorithm cannot recover. This is consistent with results found in [13]. Ambiguity selection errors in the BYU data are often associated with large regions of incor-

rect first ambiguities or fine-scale wind features. An interesting observation is that in some cases, large areas of incorrect first ambiguities occur near the southern end of the swath at moderately high wind speeds, creating ambiguity selection errors in the BYU data, but not in the nudged JPL winds.

For cyclonic cases, the BYU method performs the same as the JPL winds in about 71% of the cases. Of the remaining cases, BYU winds are subjectively better in about 9% of the cases and worse in about 20% of the cases. Thus, the BYU method is somewhat less able to produce realistic cyclonic features. Although the JPL nudging technique is not perfect in storm regions, it provides an initial guess of the placement of cyclonic rotational features. These features may not be clearly evident in the noisy first ambiguity field. Additionally, in the BYU method, cyclonic storms are sometimes overly smoothed by the low-pass effect of the KL model, resulting in poor ambiguity selection. Of the poorly retrieved storms for the BYU algorithm, a higher percentage occur on the swath edges than in the inner swath. Eighteen storms are centered on the swath edge in the test dataset. Of these storms, 11 (61%) are subjectively identified as “poor,” a much higher percentage than the overall 34.5% of BYU storm cases that are identified as “poor.” The higher rate of poorly retrieved storms on the swath edge is probably due to the fact that solutions for the swath edge are selected via extrapolation of the inner swath. Thus, small-scale features located in the other swath region are more likely to be missed by the algorithm.

The analysis in this section suggests that the BYU method is somewhat less effective than JPL in creating visually consistent cyclonic storm regions, but is generally better in areas lacking cyclonic features. The BYU method could be improved by detecting and separately processing storm regions with a specialized algorithm or by mitigating the effects of rain in storm regions. Further work is planned to accomplish this.

## V. SUMMARY AND CONCLUSION

BYU pointwise ambiguity selection uses a data-driven model rather than a nudging field to produce self-consistent wind fields for scatterometer wind ambiguity selection. A correction routine locates and corrects further large inconsistencies in the wind. The algorithm is applied to a set of test revs. The algorithm generates the same basic wind flow as the current JPL product without using the nudging field.

Largest differences between the JPL and BYU selections occur in low and high wind speed regions, at nadir and along swath edges. Our QA analysis demonstrates that low to moderate wind speed data (between 4 and 10 m/s) are generally more self-consistent for the BYU method, while higher wind speed data are somewhat less self-consistent. Also, there is increased performance in the sweet spot for the BYU method, especially in rain-corrupted regions.

From a manual inspection of the JPL and BYU ambiguity selected winds, we conclude that BYU produces fewer possible ambiguity selection errors in regions without cyclonic storms. Ambiguity selection errors in the BYU algorithm are generally associated with storms, extreme winds, and large areas of incorrect first ambiguities. However, these cases are rare. Overall, the BYU method selects 93% of the same ambiguities as the JPL. This result indicates that ambiguity selection can be generally accomplished without the use of a nudging field. The BYU ambiguity selection method is well suited for operational ambiguity selection, since the NWP nudging field is not required and the algorithm is computationally efficient.

## APPENDIX

Dilation and erosion are morphological operations that expand or contract features of a binary image [10]. Morphological operations involve a structuring element, similar to a convolution kernel. One of the simplest structuring elements is a “nearest neighbor” element. The “nearest neighbor” element is shown in Fig. 10. If the “nearest neighbor” element is used, a dilation step turns on a pixel where any neighboring pixel is turned on. Erosion turns off a pixel where any neighboring pixel is off. A general technique of filling bounded regions is to dilate for several iterations, and then erode for several iterations.

Morphological operations are a subclass of cellular automata. A cellular automaton is an array of identically programmed cells that interact with each other. For each cell, there is a state (in the binary case, ON or OFF), a neighborhood, and a set of rules on how the state changes. Morphological operations are binary, but a multivalued operation is needed in the BYU method of locating isolated regions of ambiguity selection errors. Thus, we modify the dilation and erosion techniques to better suit our application by defining a cellular automaton over the swath.

For each WVC, we define four states. State 1 is assigned to all WVCs whose median-filtered wind vector is less than 3 m/s or are nondata WVCs. Note that the outside cross-track row is assigned state 1 because retrieval is not performed there. State 2 is assigned to all WVCs that are flagged as “inconsistent” (see Section III-D). These are the edges of the regions of ambiguity selection error. Because states 1 and 2 are defined by the characteristics of the selected wind flow and not by the states of the surrounding cells, they never change during the dilation and erosion steps.

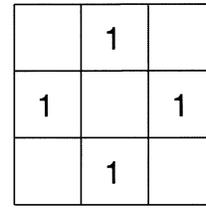


Fig. 10. “Nearest neighbor” structuring element. A value of “1” indicates a nearest neighbor. The center pixel is the origin.

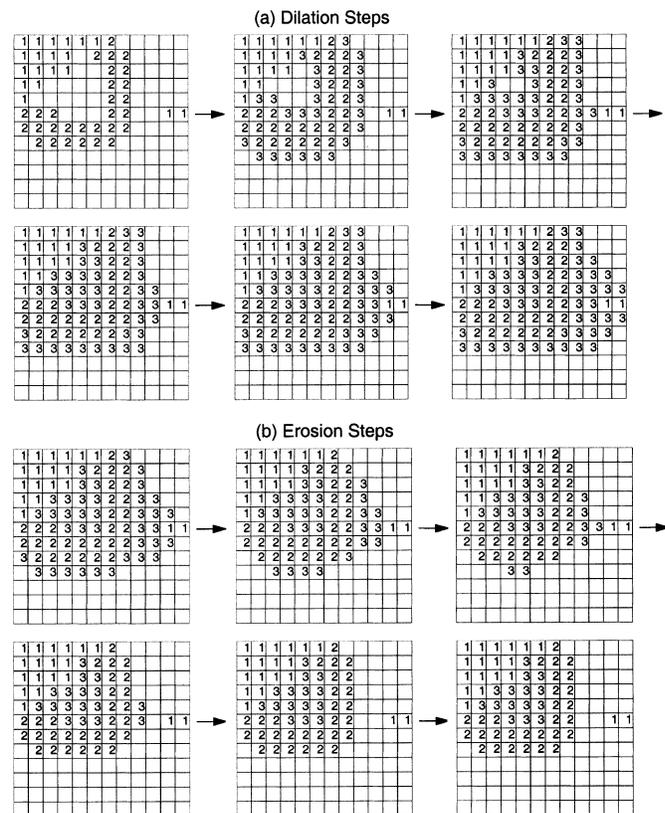


Fig. 11. (a) Dilation and (b) erosion steps, demonstrating how an isolated region is “filled in.” The initial state matrix is derived from the example in Fig. 4. To enhance readability, state 0 cells are left blank.

Cells not assigned state 1 or 2 are initially assigned state 0. Through dilation and erosion, state 3 is assigned to all isolated regions of ambiguity selection error. We redefine dilation as the changing from state 0 to state 3, and erosion as the changing from state 3 to state 0. Our neighborhood is all “nearest neighbors.”

During the dilation step, the rules for states change are as follows: States 1 and 2 do not change. State 0 changes to state 3 when the neighborhood contains at least one state 2 cell, or contains a state 3 cell accompanied by at least one other state 1 or 3 cell. After iterating, these rules allow the inconsistent edges to dilate until they come in contact with WVCs of state 1, 2, or 3, filling isolated regions. The dilation step is iterated 20 times.

During the erosion step, the rules on state changes are modified. State 3 changes to state 0 when the neighborhood contains one state 0 cell and no state 2 cells, or contains at least two state 0 cells. Again, cells of state 1 or 2 never change. Thus, the non-isolated cells erode away, leaving only the isolated regions. The erosion step is iterated 40 times. All WVCs with nonzero state

are flagged as isolated regions of ambiguity selection error. We demonstrate the dilation and erosion steps in Fig. 11.

#### REFERENCES

- [1] 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," *Proc. IEEE*, vol. 79, pp. 850–866, June 1991.
- [2] B. W. Stiles, B. D. Pollard, and R. S. Dunbar, "Direction interval retrieval with thresholded nudging: A method for improving the accuracy of QuikSCAT winds," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, pp. 79–89, Jan. 2002.
- [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. Geosci. Remote Sensing*, vol. 29, pp. 167–174, Jan. 1991.
- [4] D. G. Long, "Wind field model-based estimation of SEASAT scatterometer winds," *J. Geophys. Res.*, vol. 98, pp. 12 651–14 688, 1993.
- [5] A. E. Gonzales and D. G. Long, "An assessment of NSCAT ambiguity removal," *J. Geophys. Res.*, vol. 104, pp. 11 449–11 457, 1999.
- [6] R. S. Dunbar, S. V. Hsiao, Y. Kim, K. S. Pak, B. H. Weiss, and A. Zhang, "Science algorithm specification for SeaWinds on QuikSCAT and SeaWinds on ADEOS-II," Jet Propulsion Lab., Pasadena, CA, 2001.
- [7] M. W. Spencer, C. Wu, and D. G. Long, "Tradeoffs in the design of a spaceborne scanning pencil beam scatterometer: Application to SeaWinds," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 115–126, Jan. 1997.
- [8] C. Chi and F. K. Li, "Comparative study of several wind estimation algorithms for spaceborne scatterometers," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, pp. 115–121, Mar. 1988.
- [9] D. G. Long and J. M. Mendel, "Identifiability in wind estimation from scatterometer measurements," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 268–276, Mar. 1991.
- [10] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [11] M. H. Freilich and D. B. Chelton, "Wavenumber spectra of pacific winds measured by the seasat scatterometer," *J. Phys. Oceanogr.*, vol. 16, pp. 741–757, 1986.
- [12] J. N. Huddleston and B. W. Stiles, "A multidimensional histogram rain flagging technique for seawinds on QuikSCAT," in *Proc. IGARSS*, July 24–28, 2000, pp. 1232–1234.
- [13] D. W. Draper and D. G. Long, "An assessment of SeaWinds on QuikSCAT wind retrieval," *J. Geophys. Res.*, vol. 107, pp. (5)1–(5)14, 2002.



Mr. Draper is a member of Tau Beta Pi.

**David W. Draper** received the B.S. degree in electrical engineering from Brigham Young University (BYU), Provo, UT, in 2000. He is currently pursuing the Ph.D. degree in electrical engineering at BYU.

He joined the Microwave Earth Remote Sensing Research Group at BYU in 1999. His current research interests are remote sensing of ocean winds and rain using satellite scatterometer data. He has developed algorithms to identify typical errors in scatterometer wind estimation and is using wind and rain models to improve the wind estimation technique.



**David G. Long** (S'80–SM'98) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1989.

He is currently a Professor in the Electrical and Computer Engineering Department, Brigham Young University, Provo, UT, where he teaches upper division and graduate courses in communications, microwave remote sensing, radar, and signal processing. From 1983 to 1990, he was with the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL), Pasadena, CA, where he developed advanced radar remote sensing systems. While at JPL, he was the Senior Project Engineer on the NASA Scatterometer (NSCAT) project. NSCAT was flown aboard the Japanese Advanced Earth Observing System (ADEOS) from 1996 to 1997. He was also the Experiment Manager on SCANSAT, the precursor to SeaWinds. He is the Principle Investigator on several NASA-sponsored interdisciplinary research projects in remote sensing, including innovative radar systems, spaceborne scatterometry of the ocean and land, and modeling of atmospheric dynamics. He is a member of the SeaWinds Science Working Team. He has numerous publications in signal processing and radar scatterometry. His research interests include microwave remote sensing, radar theory, space-based sensing, estimation theory, computer graphics, signal processing, and mesoscale atmospheric dynamics.