

# Mitigation of Sea Ice Contamination in QuikSCAT Wind Retrieval

Weston J. Hullinger and David G. Long, *Fellow, IEEE*

**Abstract**—Scatterometers provide frequent estimates of near-surface wind vectors over the Earth’s oceans. However, in the polar oceans, the presence of sea ice in or near the measurement footprint can adversely affect scatterometer measurements, resulting in inaccurate wind estimates. Currently, such ice contamination is mitigated by discarding measurements within 50 km of the detected sea ice. This approach is imperfect and causes loss of coverage. We present a new algorithm that detects ice-contaminated measurements based on a metric called the ice contribution ratio (ICR), which measures the spatial ice contribution for each measurement. Determined by simulation, we threshold the ICR depending on local wind, ice backscatter, and cross-track location. Using ICR processing, wind is retrieved almost 40 km closer to sea ice than has been previously possible, while ensuring wind accuracy. Using ICR processing, retrieved wind distributions more closely resemble uncontaminated distributions than winds retrieved using previous methods. The algorithm is applied to QuikSCAT in this paper, but could be applied to other scatterometers such as the Oceansat-2 scatterometer.

**Index Terms**—QuikSCAT, remote sensing, scatterometry, sea ice, wind, wind retrieval.

## I. INTRODUCTION

**K**NOWLEDGE of polar winds is fundamental to understanding the Earth’s climate, atmospheric heat flow, ocean currents, and possibly sea ice formation. Satellite scatterometry enables daily wind vector observations over the ocean in these regions.

The SeaWinds scatterometer on the QuikSCAT satellite provided invaluable data regarding global climate from 1999 to 2009. The term QuikSCAT is also commonly used to refer to the scatterometer. Covering 90% of the Earth daily, QuikSCAT’s radar collected microwave backscatter measurements of the normalized radar cross section or  $\sigma^0$ .

Improved understanding of Arctic and Antarctic systems results from QuikSCAT’s extensive coverage of the polar regions. Measured  $\sigma^0$  have long been used to estimate ocean wind vectors using wind retrieval algorithms. However, wind retrieval accuracy is lost when ocean  $\sigma^0$  measurements are contaminated by signal contributions from land, icebergs, or sea ice.

Sea ice is a frequent contaminant of wind estimates in polar regions because of its constant movement and radar

signature which is similar to that of the ocean at high wind speeds. As the radar takes measurements, its antenna pattern can illuminate both ocean and ice simultaneously. The ice’s high radar signature may bias the  $\sigma^0$  measurements, resulting in erroneous wind retrieval. Previously, contaminated winds were eliminated by detecting the sea ice extent and then only estimating winds beyond a distance threshold of 50 km from the ice [1]. Our analysis suggests that this method can still result in ice contamination and also eliminates many valid estimates, making the loss of coverage a high price for wind accuracy. This inspires the need to find a better tradeoff between coverage and accuracy. Here we propose a new algorithm for detecting ice-contaminated QuikSCAT measurements in order to discard them before wind retrieval.

A number of algorithms have been developed for mapping the sea ice extent from scatterometer measurements, e.g., [2]–[4]. These apply maximum likelihood and Bayesian-based techniques to detect the location of the ice extent. The algorithm presented in this paper differs significantly from these by producing wind estimates that have both less ice contamination and higher coverage than obtained by simply detecting sea ice and moving away from it.

A successful land contamination mitigation technique was presented in 2009 by Owen and Long [5]. By measuring the ratio of how much of the antenna pattern overlaps land, the algorithm estimates wind error due to land for each  $\sigma^0$  measurement. Land-contaminated measurements that can be expected to result in high wind error are discarded prior to wind retrieval. The algorithm in this paper is inspired by this work but has significant modifications to deal with variability in sea ice movement from day to day, which can be as much as 80 km in a single day [6], [7].

In this paper, a metric for ice contamination detection is introduced, i.e., the ice contribution ratio (ICR). Measurement ICR values are used to identify contaminated  $\sigma^0$  measurements that are discarded prior to wind retrieval. Wind retrieval performed using the remaining measurements is thus ensured to be accurate. We find that this algorithm mitigates ice contamination more successfully than previous methods while retrieving winds almost 40 km closer to the ice edge.

Section II provides a relevant overview of the scatterometer on QuikSCAT, and Section III details the ICR algorithm. Therein, the ICR model is presented, ICR estimation is discussed, and threshold determination is explained. Section IV presents case studies for evaluation, while Section V validates the quality of the estimated winds on large data sets. Section VI concludes this paper.

Manuscript received February 15, 2012; revised September 22, 2012 and March 14, 2013; accepted April 5, 2013. Date of publication May 17, 2013; date of current version December 24, 2013.

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

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2013.2258400

## II. QUIKSCAT OVERVIEW

The QuikSCAT mission ended on November 23, 2009, but its measurements continue to provide valuable information about Earth's environment. Another similar scatterometer [8] started its mission aboard Oceansat-2 on November 4, 2009, enabling data continuation. While the study in this paper pertains to QuikSCAT, the ICR algorithm can be applied to Oceansat-2 data.

QuikSCAT transmits and receives pulses at 13.4 GHz with a dual polarized pencil-beam antenna system which utilizes a horizontally polarized inner beam and a vertically polarized outer beam pointed at  $46^\circ$  and  $54^\circ$  incident angles, respectively. The two beams rotate as the satellite moves along its track, covering a 1600-km-wide swath. Using this configuration, QuikSCAT collects forward- and aft-looking measurements for both polarizations.

An observed measurement  $\sigma_{\text{Obs}}^0$  can be modeled with noise according to

$$\sigma_{\text{Obs}}^0 = \sigma_{\text{True}}^0 (1 + K_p \eta) \quad (1)$$

where  $\eta \sim \mathcal{N}(0, 1)$  and  $K_p$  depends on measurement parameters such as the signal-to-noise ratio (SNR) and geometry [9]. The quantity  $\sigma_{\text{True}}^0$  is the noise-free measurement value.

Each  $\sigma_{\text{True}}^0$  is an average of a distributed  $\sigma^0$  on the Earth's surface weighted by the spatial response function  $R(\mathbf{v})$  over the footprint. The spatial response function is due to the antenna pattern's normalized projection onto the ground combined with signal processing [10]. The weighting  $R(\mathbf{v})$  averages the spatially distributed  $\sigma^0(\mathbf{v})$  according to

$$\sigma_{\text{True}}^0 = \frac{\int_{\text{footprint}} \sigma^0(\mathbf{v}) R(\mathbf{v}) d\mathbf{v}}{\int_{\text{footprint}} R(\mathbf{v}) d\mathbf{v}} \quad (2)$$

where  $\mathbf{v}$  is the position vector of the location on the ground.

The measurements that result from the full footprint spatial response are referred to as "egg" measurements. Egg measurements are used in the conventional 25-km-resolution L2B wind data product reported by the Jet Propulsion Laboratory (JPL) [1].

Each egg measurement is made up of several smaller "slice" measurements collected using range/Doppler filtering. Researchers have derived and tabulated the slice spatial response function for QuikSCAT, as seen in Fig. 1 [9], [10]. Using this response function and image reconstruction algorithms, researchers produce 2.5-km high-resolution (HR) wind products [11], [12] with the tradeoff of increased noise compared to conventional resolution wind products.

Both conventional and HR wind products are gridded on an along-track/cross-track grid. Each grid element, termed a wind vector cell (WVC), has 25 km resolution for conventional products and 2.5 km resolution for HR. The WVC index in the direction of flight is termed the along-track index, while that across the flight track is termed the cross-track index.

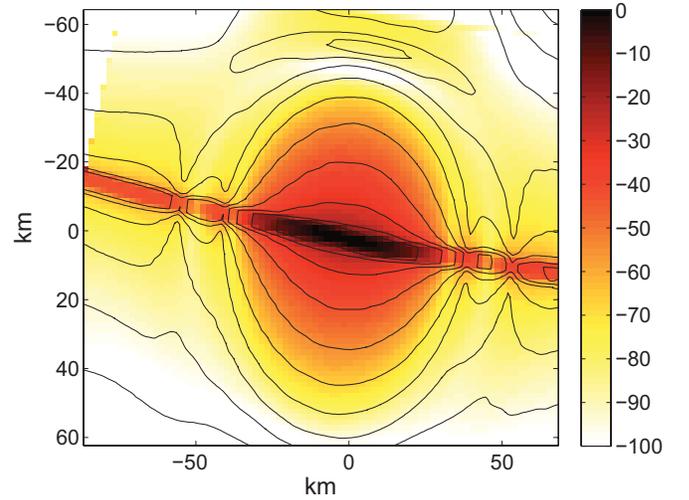


Fig. 1. Sample slice spatial response function for QuikSCAT in decibels. Contours are spaced 10 dB apart, with the major axis in the near horizontal direction. Notice that the function reduces much more slowly in the direction of the major axis. This nonsymmetry results in varying spatial contribution to the measurements depending on the antenna rotation angle [10].

Before HR wind retrieval is performed, the  $\sigma^0$  measurements are averaged into WVCs using the AVE algorithm. The AVE algorithm enhances the data resolution from a single pass by applying irregular sampling theory [13]. Scatterometer wind retrieval requires  $\sigma^0$  from several different azimuth angles (azimuth diversity) at each WVC to estimate the wind speed and direction. For QuikSCAT, four possible azimuth angles (v-pol fore, v-pol aft, h-pol fore, and h-pol aft) are collected for most of the swath, see [1]. When applying the AVE algorithm, each azimuth type is treated separately, and each WVC contains an average value for each azimuth angle.

Wind retrieval relies on a geophysical model function (GMF) [14]–[16] relating the near-surface vector wind and the radar backscatter. When the wind is low, the ocean surface is smooth and most microwave energy is reflected away from the antenna. When the wind speed is higher, the surface is rougher, resulting in more Bragg scattering back to the radar. Using these backscatter measurements, wind vectors are estimated using the Gaussian noise model in (1) and a maximum likelihood estimator [12], [17]. Alternatively, a manifold approach to wind retrieval has been demonstrated [18].

Studying winds next to the ice edge requires a map of sea ice location. Although other ice edge products could be used, for convenience and consistency with the JPL L2B processing we employ the daily QuikSCAT ice extent product produced by the Remund–Long (RL) algorithm [2], [19] which uses polar images generated by the scatterometer image reconstruction (SIR) algorithm. To produce high-resolution polar images, the SIR algorithm [20] applies irregular sampling theory to the plenitude of measurements taken by QuikSCAT over multiple revolutions. Then the RL algorithm uses an iterative maximum likelihood classifier technique to produce sea ice extent maps. Both the QuikSCAT SIR images and sea ice extent products are available from [www.scp.byu.edu](http://www.scp.byu.edu). The RL ice extent products were dilated by 50 km to mask winds in producing the subsequent day's standard JPL L2B product [1].

### III. ICE CONTRIBUTION RATIO ALGORITHM

This section details the ICR algorithm. First, the ICR is developed as a  $\sigma^0$  measurement model parameter. To estimate the ICR for a given measurement, a Bayesian approach is used to calculate the prior and posterior probabilities of sea ice. The algorithm thresholds the ICR in order to bound estimated wind error.

#### A. ICR Model

The integral in the numerator of (2) can be divided into two partitions corresponding to the disjoint regions of the footprint, which are ice and ocean. Assuming the surface backscatter of each region separately is approximately constant over the integration regions, ice and ocean  $\sigma^0$  factor out of the corresponding integrals, and (2) becomes

$$\begin{aligned}\sigma_{\text{True}}^0 &= \frac{\sigma_i^0 \int_{\text{ice}} R(\mathbf{v}) d\mathbf{v}}{\int_{\text{footprint}} R(\mathbf{v}) d\mathbf{v}} + \frac{\sigma_o^0 \int_{\text{ocean}} R(\mathbf{v}) d\mathbf{v}}{\int_{\text{footprint}} R(\mathbf{v}) d\mathbf{v}} \\ &= \sigma_i^0 \text{ICR} + \sigma_o^0 (1 - \text{ICR})\end{aligned}\quad (3)$$

where the factored backscatter is  $\sigma_i^0$  for ice and  $\sigma_o^0$  for ocean and

$$\text{ICR} = \frac{\int_{\text{ice}} R(\mathbf{v}) d\mathbf{v}}{\int_{\text{footprint}} R(\mathbf{v}) d\mathbf{v}}\quad (4)$$

is defined as the ICR.

The ICR for a given measurement is the fraction of the spatial response over sea ice and ranges from 0 to 1. In ICR processing, this quantity is calculated for each measurement. If the ICR is greater than a threshold, then the measurement is discarded as ‘‘ice contaminated.’’ The ICR thresholds are generated by simulation as described in Section III-E. First, however, a method for estimating the ICR for a given measurement is described.

#### B. ICR Estimation

To quantify the amount of ice contribution in each measurement, the ICR expression in (4) is discretized into the summations

$$\text{ICR} \approx \frac{\sum_{\text{ice}} R[n]}{\sum_{\text{footprint}} R[n]} = \frac{\sum_{\text{footprint}} I[n]R[n]}{\sum_{\text{footprint}} R[n]}\quad (5)$$

where the indicator function  $I[n]$  takes the value of 1 if location  $n$  corresponds to sea ice and 0 if it corresponds to ocean. To help account for the uncertainty in the movement of the ice edge from day to day,  $I[n]$  is treated as a binary random variable. As such,  $I[n]$  has a binary probability mass function (PMF), with the probability of ocean for  $I[n] = 0$  and the probability of sea ice for  $I[n] = 1$ . Using this PMF, an estimate for the ICR is chosen by taking the expected value of (5). This allows us to change  $I[n]$  into the probability

$P_n(\text{ice}|\sigma_{\text{Obs}}^0)$  of ice occurring in the  $n$ th index so that

$$\text{ICR} \approx \frac{\sum_{\text{footprint}} P_n(\text{ice}|\sigma_{\text{Obs}}^0)R[n]}{\sum_{\text{footprint}} R[n]}\quad (6)$$

This expression suggests a Bayesian approach to calculating the ICR for each  $\sigma^0$  measurement, which is how we choose to make this calculation. As such,  $P_n(\text{ice}|\sigma_{\text{Obs}}^0)$  is referred to as the posterior probability of ice.

The posterior probability of ice, given the observed  $\sigma^0$  measurements, can be calculated using

$$\begin{aligned}P_n(\text{ice}|\sigma_{\text{Obs}}^0) &= \frac{P_n(\text{ice})P_n(\sigma_{\text{Obs}}^0|\text{ice})}{P_n(\text{ice})P_n(\sigma_{\text{Obs}}^0|\text{ice}) + P_n(\text{ocean})P_n(\sigma_{\text{Obs}}^0|\text{ocean})}\end{aligned}\quad (7)$$

where  $P_n(\text{ice})$  is the prior probability of ice and  $P_n(\text{ocean}) = 1 - P_n(\text{ice})$  is the prior probability of ocean. The quantities  $P_n(\sigma_{\text{Obs}}^0|\text{ice})$  and  $P_n(\sigma_{\text{Obs}}^0|\text{ocean})$  are the observation probabilities, which are the probabilities of making the measured observations given that the  $n$ th pixel is ice or ocean, respectively. Both the prior and observation probabilities are needed to calculate the posterior.

#### C. Prior Contribution and Generation

The first step to calculating the posterior is to generate  $P_n(\text{ice})$ , which is the prior probability of sea ice for the  $n$ th pixel. To construct this prior, we use the sea ice maps created by the RL algorithm; however, other ice detection algorithms could also be used. The RL algorithm is used for convenience, being used in the standard L2B product and having a fine grid independent of WVC.

To generate the priors, daily binary RL ice maps are averaged over a time window including previous, current, and subsequent days. The use of subsequent days makes the algorithm noncausal, but improves performance. Using multiple days enables detection of ice that has disconnected from the main ice extent (disjoint ice). Using a noncausal prior forgoes near-real-time processing, but future research hopes to develop a prior generation method that is based only on previous and current days. For example, a prior could be generated similarly with previous days and then dilated depending on the time of year. The results presented in this paper are based on a time window spanning 23 days.

It is informative to consider the construction of a prior with disjoint ice regions. During the summer melt, portions of ice may be isolated from the main ice extent and may not be mapped by the RL algorithm. Such ice typically melts within a short time, but it still causes ice contamination during the melt period. An example of this is seen in Fig. 2, which is an HR wind estimate image contaminated by ice. The RL edge is also shown for comparison. Since undetected ice exists for only a short time, the multiple day prior results in a moderate probability of ice for the region as seen in Fig. 3. This enables the ICR algorithm to detect areas of ice contamination even if the RL algorithm does not detect ice for that particular day.

In order to minimize the probability that disjoint ice is lost, a low but nonzero prior is assigned to locations that otherwise

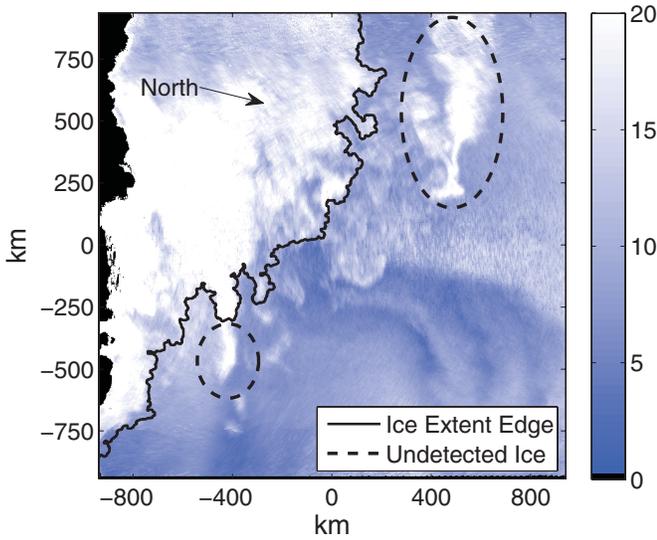


Fig. 2. Image of wind speed estimates (m/s) for WVCs south of Africa on December 15, 2004, for orbit revolution 28590. The axes represent distances from  $61.1^\circ$  south and  $33.4^\circ$  east. In this image, wind speed is retrieved for all WVCs without regard to sea ice location. Ice-covered and ice-contaminated areas result in apparent high winds. The dashed lines show areas of undetected ice, disjoint from the main ice extent. The RL edge is shown as a solid line.

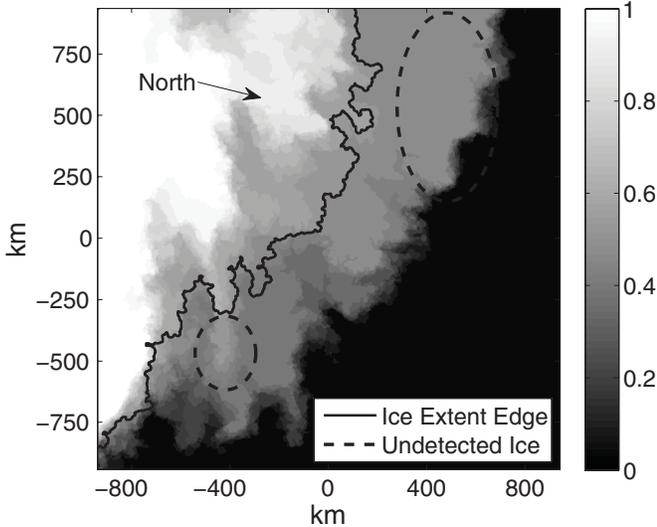


Fig. 3. Image of prior probability of ice made by averaging 23 days of RL ice maps for the area shown in Fig. 2.

would be 0. Similarly, a probability of 100% is lowered to avoid forcing ice classification.

#### D. Observation Probability

The posterior calculation in (7) requires observation probabilities  $P_n(\sigma_{\text{Obs}}^0|\text{ice})$  and  $P_n(\sigma_{\text{Obs}}^0|\text{ocean})$ . These are estimated using a year-long training set of  $\sigma^0$  observations from 2004, where ice and ocean are defined by daily RL ice maps. In each region, the four observation types (v-pol, h-pol, fore, and aft) are binned into four-dimensional histograms.

Figs. 4 and 5 show horizontally polarized distributions of  $\sigma^0$  for ocean and sea ice throughout the year 2004. The histograms change throughout the year because of changing weather and dialectic properties of the sea ice. Therefore,

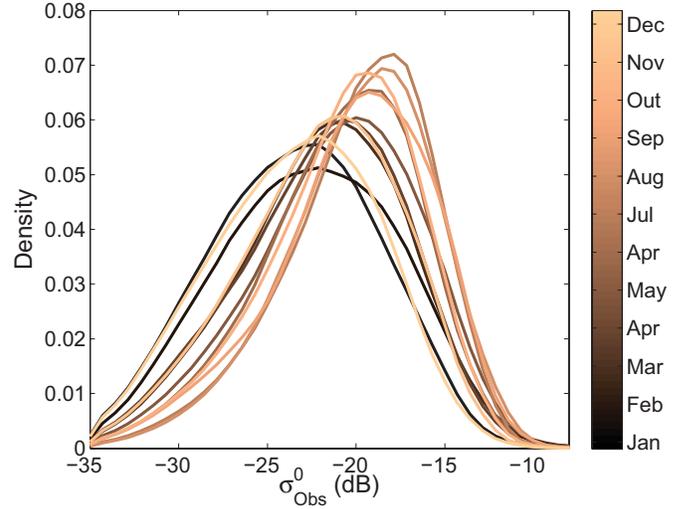


Fig. 4. Observational probability distribution of horizontally polarized  $\sigma_{\text{Obs}}^0$  for the Antarctic ocean throughout 2004. Note that the distribution changes throughout the year.

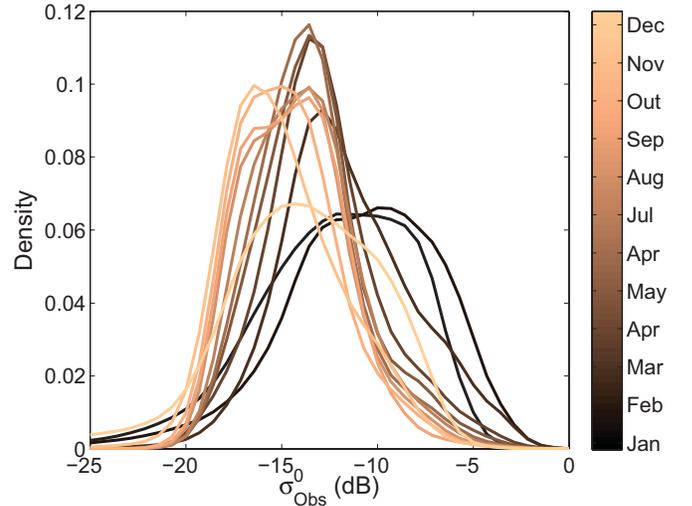


Fig. 5. Observational probability distribution of horizontally polarized  $\sigma_{\text{Obs}}^0$  for Antarctic sea ice throughout 2004. Note that the distribution changes throughout the year.

the histograms are made on a monthly basis and are normalized to obtain the observation probability estimates.

Finally, the posterior is calculated using the prior and observation probabilities according to (7), where the measured  $\sigma_{\text{Obs}}^0$  is taken from the AVE-processed  $\sigma^0$  field. The posterior assigns high probabilities over regions of disjoint ice, as seen in Fig. 6. With this posterior in place, the ICR can be calculated for each measurement.

The ICR algorithm eliminates measurements that have an ICR above a location-specific threshold to ensure that measurements are not ice contaminated. These ICR thresholds are determined through simulation as described in the next section.

#### E. Threshold Determination

ICR processing discards a  $\sigma^0$  measurement if its ICR is greater than an allowable ICR threshold in order to bound the estimated wind speed error. Unfortunately, wind error

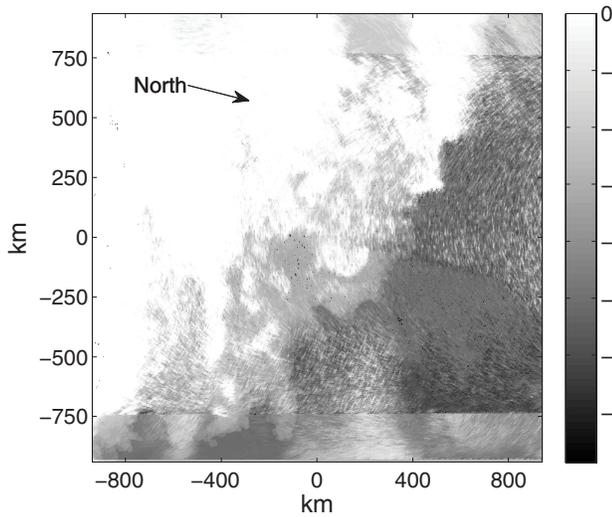


Fig. 6. Image of the values of the posterior probability of sea ice in decibels for each pixel in the area shown in Fig. 2. Notice the high probability of ice over the contaminated regions of Fig. 2.

cannot be expressed analytically because of the complexities of the GMF. However, the wind error can be analyzed using Monte Carlo simulations, which enable an evaluation of how the wind error is affected by different values of ICR, wind speed, direction,  $\sigma_i^0$ , and cross track. The objective of the simulations is to determine which values of ICR result in excessive RMS wind speed error for a given wind,  $\sigma_i^0$ , and cross-track location. In the ICR algorithm, we use the wind direction that causes the highest wind error in order to be conservative [5].

The simulation uses the noisy measurement ICR model obtained by combining (1) and (3)

$$\sigma_{\text{Obs}}^0 = \left( \sigma_i^0 \text{ICR} + \sigma_o^0 (1 - \text{ICR}) \right) (1 + K_p \eta). \quad (8)$$

The wind GMF provides a mapping between the winds and  $\sigma_o^0$ . For the simulation, we choose discrete values of the model parameters over various ranges. For  $\sigma_i^0$ , the values range from  $-19$  to  $-3$  dB in order to exhaust nominal values as seen in Fig. 5. Wind speeds range between 3 and 30 m/s to account for observed winds as well as the range over which QuikSCAT can accurately retrieve winds. The ICR ranges from 0 to 0.8192 to exhaust its possible values. The simulation is repeated for each cross-track swath location to account for the scatterometer's wind retrieval performance, which varies across the swath.

The simulation is performed for each combination of simulation parameters. First, it creates a wind field with uniform wind speed and direction from the simulation parameters. This wind field is projected through the GMF to obtain the  $\sigma_o^0$  values for the slice measurements. The ocean backscatter  $\sigma_o^0$  combines with the ICR and  $\sigma_i^0$  parameters in (8) to simulate ice contamination, after which Monte Carlo noise is added to obtain  $\sigma_{\text{Obs}}^0$  for each slice accordingly.

The next step is to process simulated observations with the AVE algorithm, followed by wind retrieval to obtain winds. For each set of parameters (ICR,  $\sigma_i^0$ , wind speed, and cross-track location), we simulate 1500 HR WVCs [5]. The retrieved

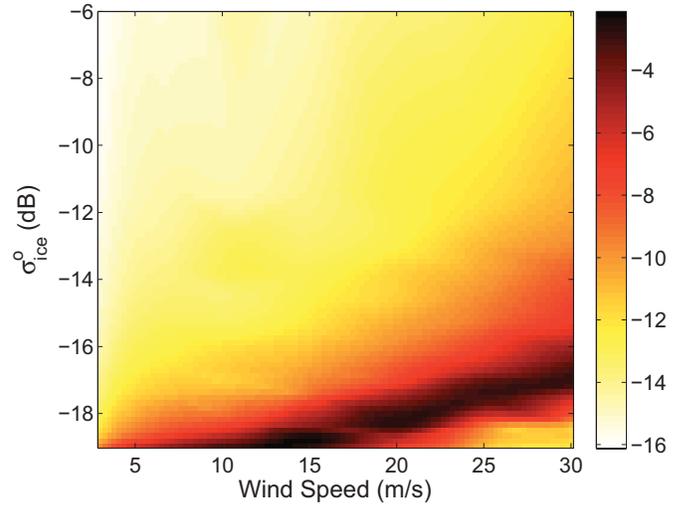


Fig. 7. Plot of the ICR threshold in decibels as a function of wind speed and ice backscatter. The value shown is the average taken across all the cross-track locations.

wind for each WVC is compared to the original true wind to calculate the RMS wind speed error.

The next step is to determine which values of ICR produce acceptable RMS wind errors. To do this, acceptable RMS wind errors must be chosen and used to compute a threshold on the ICR value. Because the nominal wind error without ice varies with wind speed and cross-track position, a single wind error threshold would not be appropriate. Instead, we adopt a speed- and swath-location-dependent threshold scheme. To aid in selecting the thresholds, a relative RMS error metric is defined as

$$\epsilon_{\text{rel}} = \frac{\epsilon_{\text{Ice}} - \epsilon_{\text{Ice Free}}}{\epsilon_{\text{Ice Free}}} \quad (9)$$

where  $\epsilon_{\text{Ice}}$  is the RMS wind speed error for ice-contaminated winds and  $\epsilon_{\text{Ice Free}}$  is the RMS wind speed error for ice-free (uncontaminated) winds.  $\epsilon_{\text{rel}}$  measures the percentage of the RMS wind error that is due to ice contamination. The ice-free wind error is obtained through simulation with the same wind and cross-track parameter as  $\epsilon_{\text{Ice}}$  but with ICR set to zero.

We arbitrarily choose to threshold the relative error at 10% for high wind speeds. If the relative error results in an RMS error less than 2 m/s, then the threshold is set to 2 m/s RMS error. This is chosen to comply with QuikSCAT mission objectives which require RMS wind speed error to be below 2 m/s for wind speeds from 2 to 20 m/s [21]. The error thresholds correspond to the maximum tolerable RMS wind speed error due to ice contamination for each set of  $\sigma_i^0$ , wind speed, and cross-track location. The corresponding maximum ICR that results in less than the RMS wind error threshold is termed the ICR threshold.

Thus the simulations determine an ICR threshold for each  $\sigma_i^0$ , wind speed, and cross-track location. To illustrate typical threshold values, Fig. 7 shows the threshold, averaged over all cross-track locations, for each wind speed and  $\sigma_i^0$ . Low ICR thresholds mean that even small portions of ice in a measurement significantly impact the winds, while large ICR thresholds suggest that the wind retrieval is less sensitive to

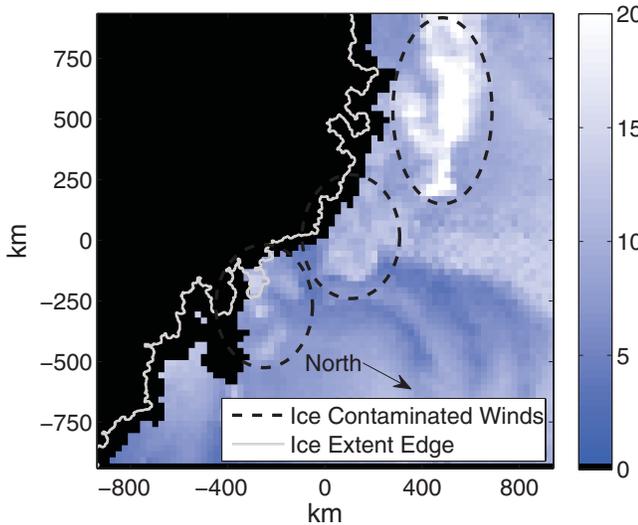


Fig. 8. Image of conventional L2B wind speeds (m/s) produced by discarding all measurements within 50 km from the RL sea ice edge (shown in gray). The region is the same as shown in Fig. 2. Wind is not retrieved in the black regions. Note that there is still significant ice contamination, especially in the upper right part of the region.

ice. Typically, low ICR thresholds are associated with low wind speeds and high  $\sigma_i^0$ . This is intuitive because ice with high  $\sigma_i^0$  in a low-wind (low  $\sigma_o^0$ ) environment biases the overall  $\sigma_{Obs}^0$  upward away from  $\sigma_o^0$ . Inversely, as the wind becomes higher, the ICR threshold becomes less restrictive depending on  $\sigma_i^0$ . This occurs when  $\sigma_i^0$  and  $\sigma_o^0$  are comparable, so that  $\sigma_{Obs}^0$  is only minimally affected by ice contamination. Also note that, at extremely high wind speeds and low ice brightness, the ice can bias the wind estimate to be low, thus creating a pocket in the lower right-hand corner of the Fig. 5.

When choosing an ICR threshold to apply on real data,  $\sigma_i^0$  and wind speed must be estimated. The wind is estimated by finding nearby WVCs in the L2B product that have the lowest posterior ice probability and keep the minimum wind speed in order to be conservative. The  $\sigma_i^0$  estimate employs the AVE algorithm's  $\sigma^0$  values that were used to create the posterior. The maximum  $\sigma^0$  is found in nearby WVCs that have the highest posterior ice probability. In this case, the maximum value is the most conservative. Using these  $\sigma^0$  and wind estimates, an appropriate ICR threshold is chosen for a given measurement.

#### F. ICR Algorithm Summary

A summary of the ICR algorithm is given in the following steps.

- 1) Precompute ICR thresholds by simulation.
- 2) Precompute observation probability distributions from a training set.
- 3) Compute prior probability of ice using RL ice maps in a temporal window.
- 4) For each revolution, use the AVE algorithm to create WVCs in the region of interest with the four azimuth types. These are used to evaluate the observational probability distributions.

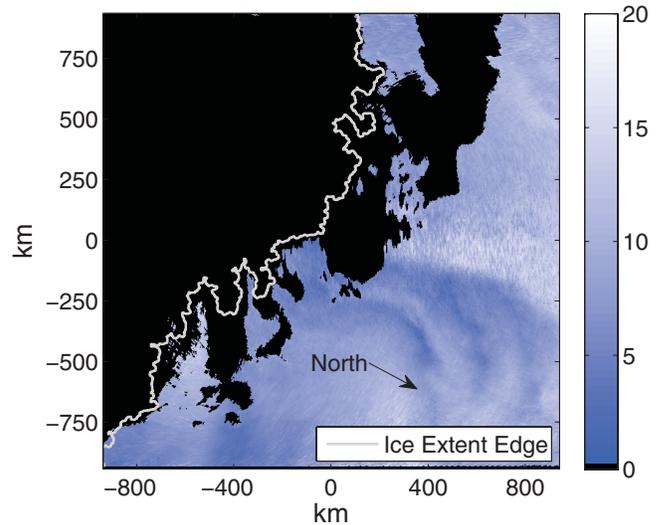


Fig. 9. Image of high-resolution wind speeds (m/s) produced using ICR processing for the area shown in Fig. 2. Wind is not retrieved in the black regions. The RL ice edge is shown for comparison.

- 5) Calculate the posterior probability of ice using products of steps 2 through 4 and (7).
- 6) Calculate the ICR for each slice measurement in the region of interest using (6).
- 7) Discard slice measurements with ICR above the ICR thresholds computed in step 1.
- 8) Using remaining  $\sigma^0$  measurements, apply the AVE algorithm and wind retrieval to make HR or conventional wind fields.

This results in winds whose RMS error due to ice contamination is less than the error thresholds. We note that every ICR wind product employs the same thresholds and observation probabilities, enabling them to be computed in preprocessing. The remainder of the steps can be executed in real time.

#### IV. CASE STUDIES

Two case studies are considered here to demonstrate that ICR processing mitigates ice contamination and retrieves more WVCs than the L2B product. Case study 1 is from orbit revolution 28 590 during the 2004 Antarctic melting period on December 15 south of Africa. Case study 2 is from orbit revolution 6367 in the 2000 Antarctic winter on October 8 east of the Drake Passage.

Case study 1 demonstrates the ice contamination mitigation capability of the ICR processing. As shown in Fig. 8, L2B winds are highly contaminated while HR ICR-processed winds in Fig. 9 exclude ice-contaminated regions. Agreeing with the HR product, the conventional 25-km ICR product is contamination-free as shown in Fig. 10. ICR processing effectively removes 573 contaminated conventional WVCs from the L2B product while also adding 100 WVCs not retrieved by the L2B product.

ICR processing retrieves more uncontaminated WVCs than the L2B product, as demonstrated by Case study 2. Fig. 11 shows wind speeds retrieved for all WVCs without regard to sea ice location. To illustrate winds retrieved by different

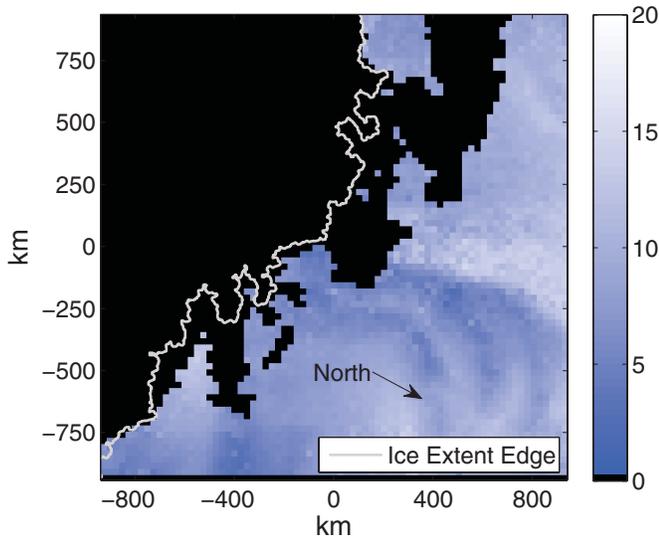


Fig. 10. Image of conventional wind speeds (m/s) produced using ICR processing for the area shown in Fig. 2. Wind is not retrieved in black regions. The RL ice edge is shown for comparison.

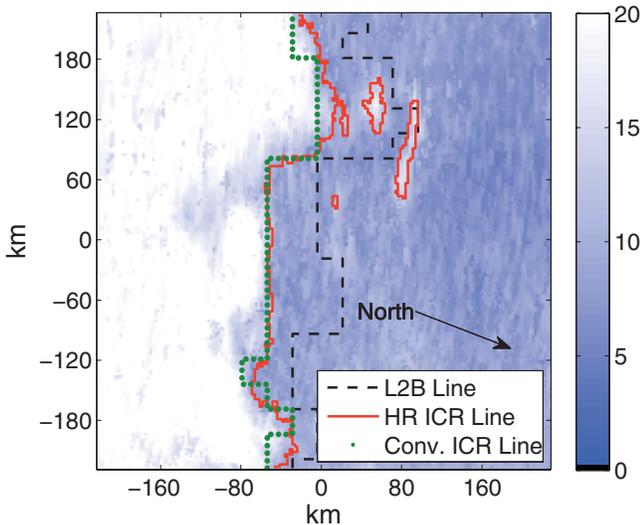


Fig. 11. Image of high-resolution wind speeds (m/s) retrieved without regard to ice conditions east of the Drake Passage with lines representing the boundary of retrievable winds using L2B product and ICR processed products. The data is taken from October 8, 2000, orbit revolution 6367. The axes represent distances from  $58.0^\circ$  south and  $40.3^\circ$  west. Note that ICR processing increases the number of retrieved WVCs compared to the L2B product.

methods, the three superimposed lines represent the boundary of retrieved winds in the L2B, HR ICR, and conventional ICR products. To the left of these lines are ice-contaminated winds, while to the right are uncontaminated winds. The ICR lines are much closer to the sea ice than the L2B line, resulting in 59 more conventional WVCs in this small region. Furthermore, 308 more conventional WVCs are retrieved over the swath, exemplifying that ICR processing can retrieve winds closer to ice than the conventional 50-km distance.

At first glance, it may appear from Fig. 11 that HR is more conservative than the conventional ICR product. However, this is not the case, as shown in Section V. The same measurements are used in both products, but the conventional resolution

product appears nearer to the ice because the WVCs include a larger area.

These case studies demonstrate that the ICR algorithm both excludes ice contamination and increases the number of uncontaminated WVCs. The large set validation that follows further demonstrates these points.

## V. LARGE SET VALIDATION

This section presents performance analysis for the ICR algorithm for several large time series in the Northern and Southern Hemispheres for 2000, 2004, and 2008. For this study, the observation probabilities  $P_n(\sigma_{\text{Obs}}^0|\text{ice})$  and  $P_n(\sigma_{\text{Obs}}^0|\text{ocean})$  are estimated using the 2004 Antarctic time series as a training set, while those of 2000 and 2008 are the withheld sets. The success in the Arctic, using an Antarctic training set, demonstrates that the algorithm is robust in its training set dependence.

The 2000 and 2004 Antarctic time series have a combined 553 revolutions sampled arbitrarily throughout the year. The Arctic 2000 and 2004 time series have a combined 333 revolutions selected arbitrarily within regions and times for which the metrics can be easily measured. The 2008 time series contains the much larger sampling of 5133 revolutions for the Antarctic and 3593 for the Arctic. Validation is performed on all these time series.

To aid the analysis, two metrics are defined to measure ice proximity and wind speed error. The wind speed error metric requires local uncontaminated winds which are discussed. The metrics are used to measure the algorithm's success in the 2000 and 2004 time series. Finally, wind distributions are compared as a validation for the 2008 time series.

### A. Validation Metrics

Two metrics measure the ICR algorithm's success. The first is the standoff distance (SOD), and the second is relative RMS wind speed error.

The SOD measures the distance from the sea ice edge that wind can be retrieved by a given algorithm. To calculate this, the ICR-determined WVCs that are closest to the ice edge are identified. Then the SOD is defined as

$$\text{SOD} = \text{mean}(d_1, d_2, \dots, d_N) \quad (10)$$

where  $d_i$  is the shortest distance to the ice edge from the center of the  $i$ th WVC, as illustrated in Fig. 12.

The relative RMS wind speed error  $\epsilon_{\text{rel}}$  is similar to the metric defined in (9), except that  $\epsilon_{\text{Ice}}$  is the RMS error of ICR-determined winds near the ice edge whereas  $\epsilon_{\text{Ice Free}}$  is the RMS wind speed error for ice-free (uncontaminated) winds. To calculate the error, we use a wind product from the National Centers for Environmental Prediction (NCEP) for comparison [22]. Similar results are obtained when using wind products from the European Center for Medium Scale Weather Forecasting.

### B. Local Ice-Free Winds

To approximate  $\epsilon_{\text{Ice Free}}$ , we investigate winds found a short distance away from the ICR-determined winds. In execution,

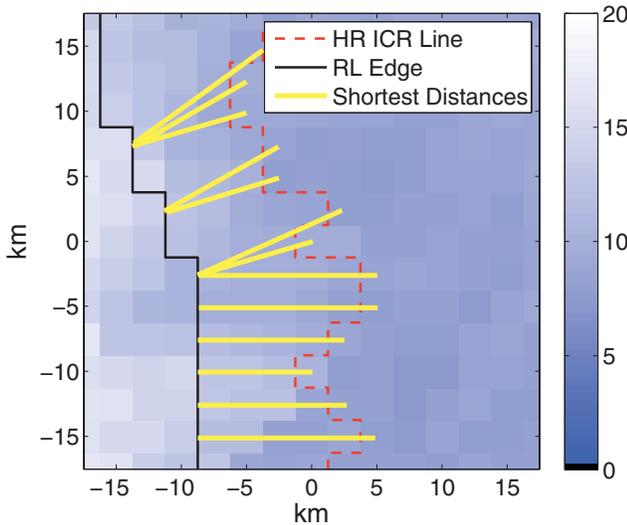


Fig. 12. Illustration of the ice edge and HR ICR line superimposed on a wind speed field (m/s). The wind speeds in this image are retrieved for all WVCs without regard to sea ice location. The yellow lines are the shortest distance from the center of each WVC along the HR ICR line to the ice edge line. The SOD is defined as the mean of these distances.

we take care that  $\epsilon_{\text{Ice Free}}$  winds are sufficiently far from sea ice to guarantee no ice contamination. Over short distances, wind statistics are assumed to be approximately constant, as will be investigated. This is important because we want any major differences between  $\epsilon_{\text{Ice}}$  and  $\epsilon_{\text{Ice Free}}$  to be only due to ice contamination.

In order to evaluate how the RMS wind error varies over short distances, we measure the RMS wind speed error as a function of the distance from the RL ice edge. In using the RL edge, one must be careful to exclude free-floating ice in the analysis since the RL maps only show ice connected to the main body. Fig. 13 shows RMS speed errors as a function of distance from the ice edge for the four time series. The initial drop in RMS error is due to the drop in ice contamination as one moves away from the edge. After this, the error reaches a floor related to normal wind retrieval skill, which validates our assumption of locally constant wind statistics. The smaller change in error when farther than 100 km out is attributed to changing wind statistics over larger distance. The variation from year to year is the result of variations in large-scale wind statistics and the areas of undetected ice patches. Based on these results, we cautiously average the RMS error between 100 and 200 km to use as  $\epsilon_{\text{Ice Free}}$  in (9).

### C. Results

The  $\epsilon_{\text{rel}}$  and SOD are measured for the four time series from the HR ICR winds, conventional ICR winds, and L2B winds as shown in Table I. The ICR SOD is consistently lower than the L2B SOD in all four time series, indicating that ICR processing retrieves winds closer to the ice edge than the L2B product. In the Antarctic ocean, HR and conventional ICR winds are retrieved almost 40 km closer, while they are about 15 km closer in the Arctic ocean. Thus the greatest increase in SOD performance is seen in the Antarctic ocean.

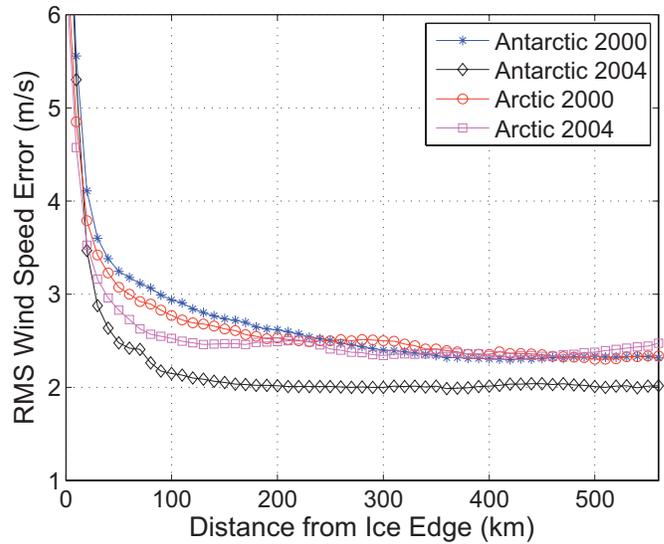


Fig. 13. RMS wind speed error as a function of distance from the RL ice edge for 2000 and 2004 time series in the Antarctic and Arctic using NCEP for comparison. Sets contain more than 8 million WVCs.

TABLE I  
STANDOFF DISTANCE (SOD) AND RELATIVE RMS WIND SPEED ERROR  $\epsilon_{\text{REL}}$  RESULTS FOR ICR PROCESSED WINDS NEXT TO THE ICE EDGE

Metric	Set	HR	Conventional	L2B
SOD (km)	Antarctic 2000	22.7	22.8	57.9
	Antarctic 2004	22.9	23.1	64.7
	Arctic 2000	22.1	22.2	38.7
	Arctic 2004	22.0	21.8	37.2
$\epsilon_{\text{rel}}$ (%)	Antarctic 2000	21.7	18.1	31.6
	Antarctic 2004	30.4	27.3	34.3
	Arctic 2000	16.5	20.6	31.2
	Arctic 2004	21.9	26.7	40.0

The metrics were measured using 886 orbit revolutions. For the L2B case, revolutions 44 721 and 46 852 were excluded because of severe and obvious ice contamination, enabling a fairer comparison between the relative errors. Conventional ICR and L2B products combine more than 100 000 WVCs to measure each metric. HR ICR SOD uses approximately 0.5 million WVCs, while  $\epsilon_{\text{rel}}$  uses approximately 1 million WVCs.

With an average measured SOD of 22.5 km, HR and conventional ICR winds have remarkably little deviation. The consistency of the SOD measurements between years demonstrates the algorithm's effectiveness.

The  $\epsilon_{\text{rel}}$  metric is lower for ICR products than for the L2B product, suggesting that the WVCs near the ice edge have less ice contamination in ICR winds than in the L2B winds. Unfortunately,  $\epsilon_{\text{rel}}$  exceeds the 10% target set in Section III-E and has more variability than the SOD. In part, this can be attributed to upsampling and error of the NCEP winds used for comparison. Another possible origin is that the relative error is larger than 10% for low wind speeds as described in Section III-E. Still, ICR products improve upon the L2B product with lower  $\epsilon_{\text{rel}}$  in all cases. In order to further demonstrate the algorithm's effectiveness, wind distributions are analyzed.

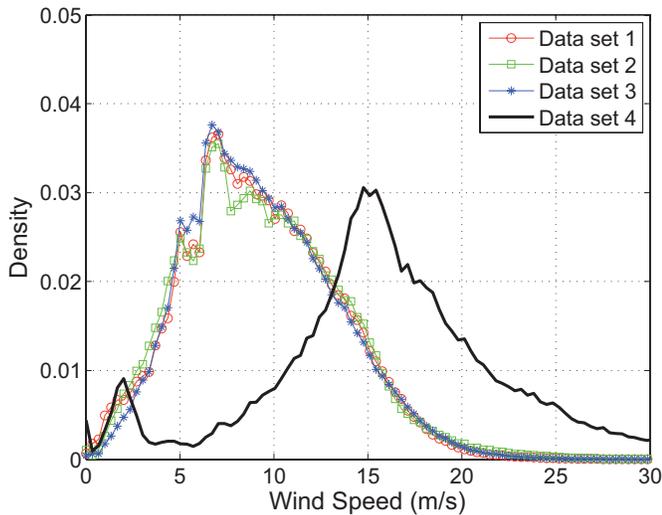


Fig. 14. Wind speed distributions in the Antarctic Ocean taken from 5133 QuikSCAT ascending passes during 2008. The distributions come from conventional WVCs from four different sets: 1) WVCs retrieved by ICR processing and agreed upon by the L2B product; 2) WVCs retrieved by L2B product and agreed upon by ICR processing; 3) WVCs retrieved by ICR processing but considered ice contaminated by the L2B product; and 4) WVCs retrieved by the L2B product but considered ice contaminated by ICR processing. There are more than 4 million WVCs in the combined datasets.

#### D. Wind Distributions

The probability distributions of wind can be useful for validating the removal of ice contamination because wind statistics are approximately constant over the open ocean. Thus, wind statistics should be about the same near the ice edge as they are far from edge. If ice contamination is present, wind speeds will appear higher than they otherwise would be.

Wind distributions are used to demonstrate the effectiveness of ICR processing in 2008 for Antarctic and Arctic winds. Fig. 14 shows wind speed distributions for the Antarctic ocean and Fig. 15 for the Arctic ocean, where each figure contains distributions consisting of the following:

- 1) Conventional WVCs retrieved using ICR processing, for which the L2B product also retrieves winds.
- 2) Conventional WVCs retrieved in L2B product, for which ICR processing also retrieves winds.
- 3) Conventional WVCs retrieved using ICR processing, for which the L2B product does not retrieve winds.
- 4) Conventional WVCs retrieved in L2B product, for which ICR processing does not retrieve winds.

Distributions 1 and 2 are from WVCs for which both the L2B and ICR processed products agree on performing wind retrieval within 100 km of the ice edge. Although the WVC locations are the same for these distributions, different slices may be used for wind retrieval since ICR processing may discard some. Note that sets 1 and 2 agree well in Fig. 14 and also separately agree in Fig. 15. These distributions are likely ice-free since both the ICR and L2B products agree.

Distribution 3 closely matches 1 and 2 in both figures. This demonstrates that ICR processing results in WVCs that are uncontaminated by ice.

Demonstrating contamination, distribution 4 has a mode that occurs much higher than the uncontaminated distributions as

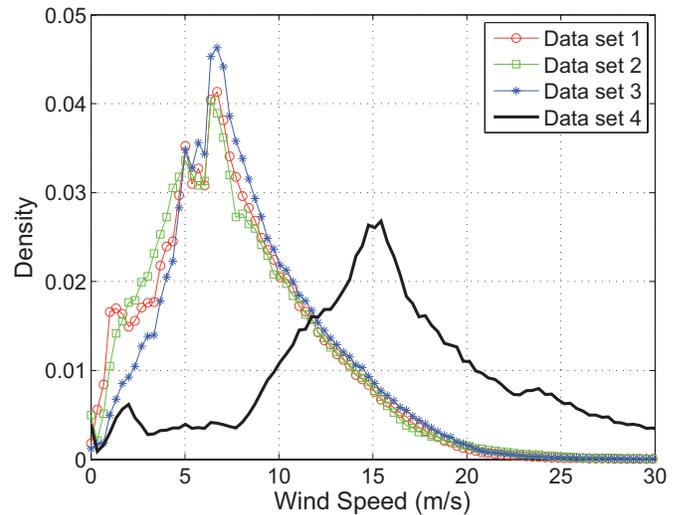


Fig. 15. Wind speed distributions in the Arctic Ocean taken from 3593 QuikSCAT descending passes during 2008. The explanations for the four distribution datasets are given in Fig. 14. There are more than 6 million WVCs in the combined datasets.

well as a high tail in the high wind speeds. This is expected for ice-contaminated winds because ice has a radar signature similar to that of ocean at high wind speeds. Because of the large difference, we conclude that the L2B product's distribution is contaminated by ice, which the ICR processing mitigates. These results agree with those found in Table I in that the ICR algorithm produces WVCs that are less contaminated by ice than the L2B product.

## VI. CONCLUSION

In summary, to quantify the ice contamination of each  $\sigma^0$  measurement, the ICR was defined. Thresholds on the ICR value were determined by limiting the error in Monte Carlo-simulated winds. Then, to mitigate ice-contaminated winds, the ICR algorithm removed ice-contaminated slices prior to wind retrieval.

The algorithm's wind estimation validation employs NCEP and L2B winds, and ice proximity validation uses the RL ice maps. Unfortunately, datasets to enable the validation of  $\sigma^0$  contamination due to sea ice are unavailable. While ice extent detection has been well studied, this is the first dataset that mitigates  $\sigma^0$  and wind contamination due to the presence of sea ice. Various datasets can be used to identify the ice edge, but unfortunately no sensors are able to directly measure the  $\sigma^0$  contamination due to sea ice that would be observed by the scatterometer. Possible future research could include studying the validity of ICR model on C-band data.

ICR processing retrieves winds almost 40 km closer to the sea ice edge than previous methods while improving accuracy according to NCEP. Section V-D shows that ICR processing eliminates measurements that estimate winds to be on average 8 m/s different than normal winds. Case study 1 exemplifies the algorithm's ability to mitigate ice contamination, and case study 2 further exemplifies the success of the algorithm, retrieving winds closer to the ice than the L2B product.

To continue polar wind coverage, ICR processing can be applied to the Oceansat-2 scatterometer [8], which is currently

operational (as of May 1, 2013). Furthermore, it is anticipated that a similar algorithm can be applied to the Advanced Scatterometer aboard the Metop-2 satellite.

## REFERENCES

- [1] T. Lungu and P. Callahan, "QuikSCAT science data product user's manual: Overview and geophysical data products," *D-18053-Rev. A, Version*, vol. 3, pp. 1–91, Jan. 2006.
- [2] Q. Remund and D. Long, "Sea ice mapping algorithm for QuikSCAT and SeaWinds," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 3, Jul. 1998, pp. 1686–1688.
- [3] H. Anderson and D. Long, "Sea ice mapping method for SeaWinds," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 647–657, Mar. 2005.
- [4] M. Belmonte Rivas and A. Stoffelen, "New Bayesian algorithm for sea ice detection with QuikSCAT," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 1894–1901, Jun. 2011.
- [5] M. Owen and D. Long, "Land-contamination compensation for QuikSCAT near-coastal wind retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 839–850, Mar. 2009.
- [6] M. Jeffries and A. G. Union, *Antarctic Sea Ice: Physical Processes, Interactions, and Variability*. Washington, DC, USA: American Geophysical Union, 1998.
- [7] J. Allen and D. Long, "Microwave observations of daily Antarctic sea-ice edge expansion and contraction rates," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 54–58, Jan. 2006.
- [8] K. Padia and S.-D. Team, *Oceansat-2 Scatterometer Algorithms for Sigma-0, Processing and Products Format*, Dept. Advanced Image Processing Group, Space Applications Centre, Gujarat, India, Tech. Rep., 2010.
- [9] M. Spencer, C. Wu, and D. Long, "Improved resolution backscatter measurements with the SeaWinds pencil-beam scatterometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 1, pp. 89–104, Jan. 2000.
- [10] I. Ashcraft and D. Long, "The spatial response function of SeaWinds backscatter measurements," *Proc. SPIE*, vol. 5151, pp. 609–618, Aug. 2003.
- [11] D. Long, "High resolution wind retrieval from SeaWinds," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 2, Dec. 2002, pp. 751–753.
- [12] B. Williams, M. Owen, and D. Long, "The ultra high resolution QuikSCAT product," in *Proc. Radar Conf.*, May 2009, pp. 1–6.
- [13] D. Long, P. Hardin, and P. Whiting, "Resolution enhancement of spaceborne scatterometer data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 3, pp. 700–715, May 1993.
- [14] M. Freilich and R. Dunbar, "Derivation of satellite wind model functions using operational surface wind analyses: An altimeter example," *J. Geophys. Res.*, vol. 98, no. C8, pp. 14633–14649, 1993.
- [15] F. Wentz and D. Smith, "A model function for the ocean-normalized radar cross section at 14 GHz derived from NSCAT observations," *J. Geophys. Res.*, vol. 104, no. C5, pp. 11499–11514, 1999.
- [16] F. Naderi, M. Freilich, and D. Long, "Spaceborne radar measurement of wind velocity over the ocean—an overview of the NSCAT scatterometer system," *Proc. IEEE*, vol. 79, no. 6, pp. 850–866, Dec. 2002.
- [17] C. Chi and F. Li, "A comparative study of several wind estimation algorithms for spaceborne scatterometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 2, pp. 115–121, Mar. 1988.
- [18] A. Stoffelen and M. Portabella, "On Bayesian scatterometer wind inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1523–1533, Jun. 2006.
- [19] Q. Remund and D. Long, "Iterative estimation of Antarctic sea ice extent using SeaWinds data," in *Proc. Geosci. Remote Sens. Symp.*, vol. 2, Jul. 2000, pp. 491–493.
- [20] D. Early and D. Long, "Image reconstruction and enhanced resolution imaging from irregular samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 291–302, Feb. 2001.
- [21] *SeaWinds on ADEOS-II Level 2B Ocean Wind Vectors in 25km Swath*, D-24902 ed., Physical Oceanography DAAC, Jet Propulsion Laboratory, Pasadena, CA, USA, Oct. 2002.
- [22] (Mar 2011). *National Centers for Environmental Prediction* [Online]. Available: <http://www.ncep.noaa.gov>



**Weston J. Hullinger** received the M.S. degree in electrical engineering from Brigham Young University (BYU), Provo, UT, USA, in April 2012.

He was a Researcher with the BYU Microwave Earth Remote Sensing Laboratory from 2009 to 2012. His current research interests include Bayesian ice contamination mitigation for scatterometer measurements and ocean wind statistics. He is currently with imSAR, LLC in Utah.



**David G. Long** (S'80–SM'98–F'08) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1989.

He was with NASA's Jet Propulsion Laboratory (JPL), Pasadena, CA, USA, from 1983 to 1990, where he developed advanced radar remote sensing systems. While at JPL, he was the Project Engineer on the NASA Scatterometer (NSCAT) project which flew from 1996 to 1997. He also managed the SCANSAT project, the precursor to SeaWinds which was launched in 1999 on QuikSCAT and 2002 on ADEOS-II. He is currently a Professor with the Electrical and Computer Engineering Department, Brigham Young University, Provo, UT, USA, where he teaches upper division and graduate courses in communications, microwave remote sensing, radar, and signal processing, and is the Director of the BYU Center for Remote Sensing. He is the Principle Investigator on several NASA-sponsored research projects in remote sensing. He has authored or co-authored over 400 publications in various areas including signal processing, radar scatterometry, and synthetic aperture radar. His current research interests include microwave remote sensing, radar theory, space-based sensing, estimation theory, signal processing, and mesoscale atmospheric dynamics.

Dr. Long was the recipient of the NASA Certificate of Recognition several times. He is an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.