Objective analysis of marine winds with the benefit of the Radarsat-1 synthetic aperture radar: A nonlinear regression framework

Richard E. Danielson,1 Michael Dowd,2 and Harold Ritchie1,3

Received 22 June 2007; revised 29 October 2007; accepted 7 January 2008; published 15 May 2008.

[1] Surface wind analyses are constructed from spaceborne synthetic aperture radar (SAR) measurements along coastal regions of eastern and western North America and collocated operational marine wind forecasts. Each analysis minimizes the error sum of squares of the wind forecast, the SAR backscatter, and wind direction that is inferred from the SAR backscatter gradient. The relative importance of the SAR wind information is defined by its expected error covariances. Parameters that define these covariances are postulated for half the analyses by comparison with an independent set of buoy observations. The remaining analyses are found to compare better with buoy observations than conventional analysis approaches. It is suggested that SAR wind information generally be weighted strongly in an analysis and that an error covariance length scale of O[10 km] is appropriate.


1. Introduction

[2] The use of active microwave observations in marine wind analysis is predicated on a complex relationship between surface wind stress and the strength of the backscattered signal (the radar cross section). The basis for this relationship is Bragg scattering from gravity–capillary waves. These centimeter–scale ocean surface waves grow and decay quickly in tandem with the wind. Sustained forcing leads to the growth of longer wind waves that increase the radar cross section by tilting the short waves and making them more visible to a radar [Donelan and Pierson, 1987; Mourad et al., 2000]. Backscatter also depends on the relative direction of the wind-forcing, as well as on the microwave viewing incidence angle, frequency, and polarization [Vachon and Dobson, 2000]. Less well understood, perhaps, is the relevance of processes such as wave breaking [Plant, 2003; Mouche et al., 2005] and precipitation falling through the atmosphere [Katsaros et al., 2000] and impacting the ocean [Braun and Gade, 2006].

[3] Empirical models that simply relate wind speed and direction to radar cross section [e.g., Stoffelen and Anderson, 1993; Quilfen et al., 1998; Horstmann et al., 2003; Hersbach et al., 2007] have enabled satellites with scatterometers, as well as those with synthetic aperture radars (SARs), to fill significant observational gaps over the ocean. The calibration of empirical models has traditionally been performed using scatterometers because, at O[10-km] resolution, they afford both greater coverage and multiple views of the same scene from different angles. This latter capability often permits both wind speed and direction to be estimated directly from scatterometer radar cross section. SAR observations, on the other hand, afford only a single view of a given scene, but at much higher resolution (O[10-m]). They are therefore able to resolve coastal regions without being contaminated by land or sea ice, although more than just an empirical model is generally required to retrieve a wind vector from them.

[4] Numerous studies have demonstrated that if an independent estimate of wind direction is available, an empirical model can be used to estimate wind speed from SAR backscatter. Vachon and Dobson [1996] were perhaps the first to validate this using 16 ship observations that were exactly collocated (in time and space) with SAR observations. In their study, wind direction was estimated from linear patterns in backscatter at O[1-km] scales. Vachon and Dobson [2000] examined an extended data set and employed wind directions from collocated in situ observations. Differences between the in situ wind speeds and estimates derived from SAR data were found to have standard deviations of 2.4 m s⁻¹. Monaldo et al. [2001] used wind directions from a numerical model to obtain standard deviations of less than 2 m s⁻¹, and Monaldo et al. [2004] used scatterometer wind directions to show that further improvements in wind speed could be obtained. It follows that SAR wind speed is about as reliable as that obtained from higher resolution in situ and lower resolution scatterometer observations, when comparisons are made at a common resolution of O[10-km] [cf. Kent et al., 1998; Stoffelen, 1998].
Backscatter patterns are often found within SAR scenes. Those that have a robust physical interpretation hold the promise of providing an independent source of wind information that might be used to reduce the ambiguity in wind vector retrieval. Perhaps the most promising pattern has long been recognized as a proxy for wind direction: the linear pattern. Following the passage of a cold front over the Gulf Stream, Thompson et al. [1983] noted the correspondence between cloud streets in visual images and linear variations in SAR backscatter, which were interpreted as the surface expression of atmospheric roll convection. A similar synoptic situation was examined by Gerling [1986], who performed a spectral analysis of SAR subscenes to quantify the direction of O[1-km] linear features, which were taken to be indicative of the local wind direction (to within a 180° ambiguity). Alpers and Brümmer [1994] focused on two SAR scenes that showed evidence of boundary layer rolls. They used an empirical model and the direction of the observed backscatter variations, which were about 1 dB, to show that the corresponding surface wind variations associated with the rolls were about 1 m s⁻¹.

The feasibility of using linear patterns in any given SAR image as a proxy for wind direction was first examined using the European Remote Sensing (ERS-1 and 2) SAR instruments and a spectral approach, following Gerling [1986]. Vachon and Dobson [1996] examined variability at scales greater than 2.5 km and found 15 of 16 scenes for which wind direction could be estimated with an RMS difference (relative to ship observations) of 24°. Wackerman et al. [1996] focused on 7 out of 9 acquisitions for which backscatter variations could be found on scales between 3 km and 9 km, and for which the estimated wind direction differences had a standard deviation of 19°. Fetterer et al. [1998] searched 65 scenes for indications of variability at scales between 1.6 km and 6.4 km. They did not make a distinction according to whether scenes contained obvious evidence of boundary layer rolls and found an RMS difference of 37°. Lehner et al. [1998] focused on much shorter wavelengths of between 0.5 km and 1.5 km. They obtained an RMS difference of only 5° when considering 14 favorable scenes out of 20.

Fetterer et al. [1998] suggested that linear patterns in SAR scenes may not be caused by boundary layer rolls alone, thus calling into question the relevance of this initial physical interpretation. The frequency of rolls in the Gulf of Alaska and Bering Sea have been examined by Levy and Brown [1998] and Levy [2001]. They found that rolls of O [1 km] are present only about a third of the time and that wind streaks tend to develop more easily than boundary layer rolls in numerical simulations.

The preceding results suggest (and surface–based radar studies appear to confirm [e.g., Dankert and Horstmann, 2007]) that a wind direction proxy can be more robust if spatial scales of O[100 m] are considered. Koch [2000, 2004] also recognized this in studies of linear patterns at scales above 400 m. He employed a local gradient approach, which facilitated the attempt to mask non–wind related patterns (e.g., isolated regions with strong edges such as land, sea ice, and tidal features). The most frequent direction of the unmasked local gradients were then used as an indication of wind direction. Horstmann et al. [2002] and Koch and Feser [2006] employed this method and an empirical model to obtain wind speed. Differences between numerical model winds and their estimates were found to have standard deviations of, respectively, 21.6° and 17.6° for wind direction and 3.5 m s⁻¹ and 2.9 m s⁻¹ for wind speed.

The local gradient approach appears to be based on a robust physical interpretation of linear patterns in SAR scenes, but it employs close to the highest resolution that a satellite SAR can resolve. Atmospheric gravity waves tend to be found at slightly larger scales, but ocean gravity waves (particularly swell) are sometimes visible at this resolution [Dowd et al., 2001; Koch, 2004]. Koch and Feser [2006] did not analyze the wind field for 10 out of 80 dates on which SAR observations were available, owing to a lack of directional information even at the smallest scales. Because SAR scenes with poor wind direction information nevertheless contain useful wind speed information, some analyses employ the wind direction simulated by a numerical model instead [e.g., Monaldo et al., 2001]. Of course, this implies that a SAR wind direction is generally not informative, and for many scenes, this would be an invalid assumption.

The goal of any analysis approach is to obtain the best estimate of the true wind field. If both SAR and numerical model data are available (as they are in this study), then one step toward a more formal approach for combining them is to estimate the error statistics associated with their wind information and use these to guide the composition of an analysis. Qualitatively speaking, errors are related to the confidence given to these two independent sources of wind information. Quantitative aspects of this approach are well established in the context of nonlinear regression [Seber and Wild, 1989], Bayesian analysis [Lorenc, 1986], and data assimilation [Daley, 1991]. In the simplest application of these techniques, it is often assumed that the error statistics have a Gaussian distribution.

The differences in wind speed and direction noted above, and those summarized by Beal et al. [2004], have broadly characterized the errors in SAR wind information. Few studies have employed such a characterization in the analysis of the surface wind, but Portabella et al. [2002] is one exception (others include Perrie et al. [2002] and Kerbaol [2006]). Their study combines wind information from SAR backscatter and numerical model forecasts. This is accomplished by minimizing a cost function that involves the weighted differences between their analysis and the wind information, where weights are inversely proportional to the postulated SAR and numerical model errors. In this
2.1. SAR Backscatter

context, most analysis methods implicitly assume that the various sources of wind speed and direction information are either exact and have no error, or have infinite error and are not made use of.

[12] Nonlinear regression provides a formal context within which to make use of errors in SAR backscatter and in the wind direction inferred from its gradient [Seber and Wild, 1989]. The cost function to be minimized is the generalized error sum of squares, which permits the spatial and cross–covariance of errors to be considered in the search for the desired analysis. Although scatterometer observations of O[10-km] resolution appear to have errors with almost no spatial covariance [Stoffelen and Anderson, 1997a], SAR observations have much finer resolution. Thus spatially coherent SAR errors may exist, owing both to coherent small–scale variability in the features being imaged, as well as to the imaging process itself. If such a constraint applies, it would be important to accommodate this in the construction of an analysis. One question that then arises is how much of an improvement might be obtained by considering spatial error covariance as opposed to considering SAR wind direction information.

[13] This study examines a range of error postulates for SAR backscatter and its wind direction information, with the objective of obtaining the best estimate of the true wind field. Nonlinear regression is employed to combine SAR wind information with numerical wind forecasts and the resulting analyses are then compared with in situ observations. The general form of the error postulates is given in section 3, along with an example of an analysis. Several analyses are performed in section 4 to identify a reasonable error parameterization. In situ observations are used to evaluate the nonlinear regression results as well as other analysis approaches. A discussion and summary are found in section 5. The next section describes the SAR and numerical model data that are combined to create the analyses and the buoy data that is used for validation.

2. Data

2.1. SAR Backscatter

[14] The Radarsat-1 satellite has been in near–continuous operation since November 1995. It carries a horizontally polarized C-band (5.6-cm) SAR in a polar orbit about 800 km above the surface. We focus on 484 SAR scenes that were acquired between June 2004 and March 2006 over the east and west coasts of North America. The coverage of these acquisitions is shown in Figure 1. All scenes employed a ScanSAR beam mode whose individual beam modes (Wide-1 and Wide-2) have been radiometrically calibrated since November 1999 [RSI, 2000]. As shown in the next section, however, the dependence of radar cross section on incidence angle also requires a residual adjustment [Monaldo et al., 2001; Beal et al., 2004]. Radar cross section is obtained at a resolution of 50 m covering swaths of greater than 300 km. The corresponding incidence angles vary from 20° (near range) to almost 40° (far range).

[15] The 50-m SAR data is masked and smoothed to facilitate processing. The resolution of each SAR acquisition is first reduced to 100 m, 200 m, and 400 m by three passes of the smoothing operator described by Koch [2004]. The 400-m data is used to mask regions that appear to be contaminated within the ScanSAR swath and along its edges. (These regions are identified visually by a strong rangeward gradient oriented along the satellite track.) Radar cross section is then masked over land, sea ice, and within regions of precipitation (as depicted by a 1-h accumulation in the forecast model). Also, because empirical models of the backscatter–wind relationship are expected to have a restricted range of validity [Shankaranarayanan and Donelan, 2001; Hersbach et al., 2007], SAR observations are masked where the strongest possible wind is less than 2 m s \(^{-1}\) (assuming a wind direction toward the satellite) and the weakest possible wind is greater than 40 m s \(^{-1}\) (assuming a wind direction parallel to the satellite track, as this yields a conservative estimate of this range of validity in terms of backscatter). The remaining 400-m backscatter pixels are considered to be appropriate for estimating wind–related SAR errors. These are subsequently smoothed to 12.8-km resolution, which is roughly equivalent to the resolution of the forecast model described below.

2.2. SAR Backscatter Gradient

[16] A more sophisticated sequence of masking and smoothing procedures is employed, following Koch [2004], to estimate wind direction from the SAR backscatter gradient. Directional estimates are calculated at three different resolutions (starting with the 100-m, 200-m, and 400-m pixels). Isolated regions of strong backscatter contrast are taken to be boundaries of land, sea ice, or tidal features and these are incorporated into the existing land mask before gradients are computed. Gradients are then weighted by how similar they are to their (unmasked) neighbors. The weighted gradient direction that is most frequent, within a

---

**Figure 1.** The location of 484 Radarsat-1 ScanSAR (Wide-1/Wide-2 beam mode) scenes (overlapping squares) that were acquired between June 2004 and March 2006 over the (a) west and (b) east coasts of North America. Included are 35 collocated buoy platforms (dots).
domain of about \((50 \text{ km})^2\) centered on any given location, is
taken to be perpendicular to the local best estimate of wind
direction (with a \(180^\circ\) ambiguity). The three estimates of
direction are derived wherever the smoothed SAR back-
scatter is valid at 12.8-km resolution.

[17] The 100-m, 200-m, and 400-m gradient directions are
sensitive to linear patterns at scales between 0.4–2 km,
0.8–4 km, and 1.6–8 km, respectively [Koch and Feser, 2006].
Although these directional estimates are often similar,
we assume that one is more sensitive to the wind-
related patterns of interest. Whereas Horstmann et al. [2002] chose the estimate whose directions best matched
numerical model data, Koch and Feser [2006] made the
choice subjectively. Koch [2004] suggested that the 100-m
estimate is appropriate if the ocean is sheltered from swell
and Koch and Feser [2006] confirmed that the 400-m
estimate is rarely selected. We opt for perhaps the simplest
objective choice, which is to take the smoothest of the three
(as defined by the squared difference in wind direction
components, averaged over all valid parts of the SAR
scene). This is certainly not ideal for sharp changes in wind
direction (i.e., fronts), but it conveniently minimizes the
occurrence of nearly perpendicular directions at adjacent
locations. (This effect might be caused by a sensitivity to
oceanic and atmospheric gravity waves that happen to be
oriented roughly at right angles to the prevailing wind.)

This happens, the three estimates often differ and the
smoothest estimate often seems appropriate. As in Koch and Feser [2006], the higher resolution estimates are thus
chosen more frequently: of 484 scenes, 70% use the 100-m
gradient directions, 20% use the 200-m directions, and 10%
use the 400-m directions.

2.3. Numerical Model Winds

[18] The other source of wind information that contrib-
utes to our analyses is provided by an operational data
assimilation and forecast system: the Global Environmental
Multi-scale (GEM hereafter) regional system [Laroche et al., 1999; Mailhot et al., 2006]. The forecast component of
GEM produces, every 12 h, a series of hourly wind
forecasts, from which we consider the earliest pair of
forecasts from 6 h onward that brackets the SAR acquisition
time. This is done to allow the forecast model to spin up
[Mailhot et al., 2006], but also because none of the buoy
observations that we use for validation have been assimil-
ated into the model initial conditions. The resolution of the
forecasts is 15 km. Boundary layer processes that operate at
scales of \(\text{O}[100 \text{ m}]\), such as those discussed in the Intro-
duction, are not explicitly resolved in this model. Instead,
unresolved impacts are parameterized [Belair et al., 2005].

[19] The GEM wind forecasts have been interpolated
vertically from the lowest model level to 10 m above the
surface (the level at which empirical backscatter–wind
relationships are generally calibrated [Hersbach et al.,
2007]). We then interpolate the winds both to the transit
time and location of the SAR observations (and note that
time and location of the SAR observations (and note that
errors in these data are increased slightly by doing so).
GEM winds are also rotated into the local SAR coordinate,
which is oriented along the satellite track, and analyzed
winds are subsequently rotated back for comparison with in
situ observations.

2.4. Buoy Winds

[20] The buoy winds used to gauge the relative success of
different analysis methods are taken from 32 platforms
operated by Environment Canada and the United States
National Oceanographic and Atmospheric Administration,
as well as three others (L, M, and N) operated by the Gulf
of Maine Ocean Observing System (Figure 1). A total of
831 buoy observations are found to be collocated with our
484 SAR acquisitions (Figure 2). Wind speed has a mean of
7.5 m s\(^{-1}\) and a maximum of 17.8 m s\(^{-1}\). The direction
and zonal component have a westerly preference, as expected,
while the distribution of the meridional component is more
symmetric.

[21] Buoy observations are selected by allowing a max-
imum separation between collocated buoy and SAR obser-
vation of 50 km and 30 min. These observations occur
throughout the 60-min period centered on SAR acquisition,
but the most frequent are within only a few minutes of exact
collocation (not shown). Each observation is a 10-min
average at a height of about 3–5 m above the surface.
Although buoy wind speed can be adjusted to the 10-m
reference level by considering stability [Walmsley, 1988],
this has the effect of increasing variance in the analysis–
buoy differences (not shown). Thus as in Fetterer et al. [1998],
no adjustment of buoy observations to the 10-m
reference height is made.

[22] Moored buoys provide some of the most reliable in
situ marine wind observations [Thomas et al., 2005], but at
the resolution of the analyses performed in this study, they
may suffer most from errors of representation. In other
words, if the true wind field is taken to be at 12.8-km
resolution, then the wind variance that the buoys resolve on
all scales smaller than 12.8 km can be considered an error
variance that may be comparable to (or even larger than)
that of either the SAR or GEM data. Stoffelen [1998] illustrates that such errors complicate a direct calibration of scatterometer winds by linear regression [see also Kent et al., 1998]. We indicate in the Appendix that buoy observations may nevertheless be useful to distinguish between the various analyses that we construct in section 4. We return to the estimation of SAR errors using independent and erroneous buoy observations in section 5 and explore next a statistical approach for combining SAR and GEM wind information.

3. Methodology

[23] A nonlinear regression framework provides an estimate the true wind field from SAR and GEM data and determines the proper composition of an analysis according to the expected errors in these data. Our quantification of expected errors begins with a definition of the errors themselves. If \( \mathbf{Y} \) is a vector containing both the SAR and GEM data and \( \mathbf{x}^* \) is the true wind field, the corresponding SAR and GEM errors (\( \mathbf{E} \)) are related by an equation of the form \( \mathbf{Y} = \mathbf{C}(\mathbf{x}^*) + \mathbf{E} \). Specifically, we have

\[
\begin{pmatrix}
  y_1 \\
  d_1
\end{pmatrix} = \begin{pmatrix}
  \mathbf{c}(\mathbf{x}^*) \\
  \mathbf{d}(\mathbf{x}^*)
\end{pmatrix} + \begin{pmatrix}
  \mathbf{e}_1 \\
  \mathbf{e}_2
\end{pmatrix}.
\]

(1)

Each row is an equation for the quantities on the left-hand side (lhs): the SAR radar cross section (\( y_1 \), which includes two preprocessing steps given below), the derived SAR wind direction (\( d_1 \)), and the GEM wind vector (\( x_2 \)). Each column vector has dimension 5N (for a SAR scene with N valid observations at 12.8-km resolution).

[24] Our postulation of SAR errors (\( \mathbf{e}_1 \)) in the next section is facilitated here by an appropriate definition of the various terms in (1). Stoffelen and Anderson [1997a] and Stoffelen [1998] found that scatterometer and wind errors are easier to quantify when backscatter is expressed in decibels and wind vectors are expressed as components (rather than as wind speed and direction). Thus the units of the first row are decibels and we define the third row in terms of cross-track (\( u \)) and along-track (\( v \)) components. All wind directions in the second row are doubled (modulo 360°) because \( \mathbf{d}_1 \), with its 180° ambiguity, is then uniquely defined [Koch, 2004]. (For example, doubling either 90° or 270° maps both onto 180°.) Also, these terms are expressed as unit vectors owing to the 360° ambiguity in any scalar direction (whether it is doubled or not). Below, we also introduce two SAR preprocessing steps: a normalization that reduces the strong incidence angle dependence of each term in the first row, and a minor correction of \( y_1 \) to reduce a wind speed bias at near and far range [Monaldo et al., 2001; Beal et al., 2004].

[25] The functions \( \mathbf{c} \) and \( \mathbf{d} \) on the right-hand side (rhs) of (1) operate on the true wind at each analysis location. The function \( \mathbf{d} \) provides unit vectors that point at twice the direction of the wind. An issue that continues to be explored [cf. Thompson et al., 1998; Mouche et al., 2005] is the most appropriate choice for each of the three components that make up the function \( \mathbf{c} \). This defines our empirical relationship between the radar cross section and the wind. Although a proper examination is beyond our present scope, it seems reasonable to choose the first component to be the C-band model (CMOD hereafter) that was developed by Hersbach et al. [2007] for the vertically polarized ERS-2 scatterometer. The second component is the correction for CMOD that was proposed by Vachon and Dobson [2000] for adapting vertically polarized backscatter—wind relationships to the horizontally polarized Radarsat-1 SAR. The third component is a normalization to reduce the incidence—angle dependence of backscatter [cf. Wackerman et al., 1996].

[26] The three components of \( \mathbf{c} \) can be written as a function of a given wind vector \([u, v]^T\) that is viewed at an incidence angle \( \theta \) (which effectively locates the wind vector relative to the satellite). If \( n(\theta) \) is an incidence—angle normalization, our estimate of radar cross section is then obtained at this location according to

\[
\mathbf{c}(\theta) = n(\theta)10 \log_{10} \left( b_0(V, \theta) [1 + b_1(V, \theta) \cos(\phi)] + b_2(V, \theta) \cos(2\phi) \right)^{1.6} \left( \frac{1 + \tan^2(\theta)}{1 + 2 \tan^2(\theta)} \right)^{0.2}.
\]

(2)

Here, wind speed (\( V \)) and wind direction (\( \phi \)) are implicit functions of \( u \) and \( v \) and each of the coefficients \( b_0 - b_2 \) is a function of wind speed and incidence angle. Terms involving wind dependence on the rhs (\( b_0[\cdot]^2 \)) are given by Hersbach et al. [2007] and the squared ratio that depends only on incidence angle (\( \cdot \)) is the polarization ratio of Vachon and Dobson [2000]. He et al. [2005] found that the sinusoidal dependence on wind direction can yield (for some directions) a weak sensitivity of CMOD to changes in direction. Sensitivity to changes in wind speed is often stronger [Portabella et al., 2002], but in general, CMOD exhibits smooth (nonlinear) behavior over a range of environmental and imaging conditions [e.g., Beal et al., 2004, Figure 1].

[27] The normalization \( n(\theta) \) to reduce incidence—angle dependence is the third component of \( \mathbf{c} \) and we also apply this as a preprocessing step to SAR backscatter (\( y_1 \)). Normalized backscatter happens to be easier to visualize because it correlates well with wind speed, but our primary justification for reducing incidence—angle dependence is that it permits greater flexibility in our SAR error postulates (e.g., spatial error covariance can be isotropic). Figure 3a illustrates the typical dependence of SAR backscatter on incidence angle, with higher values at near range. This is an average profile for all 484 SAR scenes. A quadratic function that fits this profile (thin line) is used to normalize backscatter to a reference value of −13 dB, which corresponds to an incidence angle of about 30° (the horizontal line is the average profile after normalization and an individual SAR scene after normalization is shown in Figure 4. For all scenes, the quantity \( n(\theta) = -13(0.015\theta^2 - 1.69\theta + 23.95)^{-1} \) (with \( \theta \) in degrees) thus multiplies the terms on the first row of (1).

[28] It is necessary to apply one other preprocessing step to SAR backscatter. Monaldo et al. [2001] and Beal et al. [2004] have employed the SAR—speed, GEM—direction type of analysis approach to reveal an apparent bias in the Radarsat-1 backscatter and wind speed. This may be related to an analogue— to—digital saturation bias that Radarsat-1 is known to have [Vachon et al., 1997] and is not reproduced by CMOD. The dotted line in Figure 3b confirms this trend
Figure 3. Along-track averages of (a) radar cross section (dB) and (b) wind speed (m s\(^{-1}\)) for 484 Radarsat-1 ScanSAR acquisitions as a function of incidence angle (at 0.5° intervals). The dotted lines exclude, and dot-dashed lines include, an adjustment for bias in radar cross section (see text). The thin line in Figure 3a is a quadratic fit to the adjusted radar cross section. The thick solid lines are (a) normalized and adjusted radar cross section and (b) GEM wind speed. The SAR wind speed profiles in Figure 3b are calculated using the SAR–speed GEM–direction method.

for our 484 SAR scenes. The corresponding GEM wind speed (solid line) shows no such trend. Backscatter appears to be biased low at low incidence (near range) and high at high incidence (far range). We thus adjust the radar cross section of each SAR scene to bring the two wind speed estimates into better agreement (dot-dashed line). It suffices to multiply \( y \) by a function that depends linearly on incidence angle \( [1 + 0.0045 (\theta - 32)] (\theta \text{ in degrees}) \). This is a backscatter correction of only about 5% (in dB) at the lateral edges of an acquisition.

[29] The preceding definition of terms in (1) facilitates a parameterization of the SAR and GEM error statistics. Notably, SAR error bias is reduced and the GEM errors are more Gaussian in character [Stoffelen, 1998]. The SAR errors \( e \) refer to a normalized backscatter profile and include errors in \( c \) and in the SAR observations themselves (both radiometric and geometric, with the latter including errors in incidence angle). The errors \( e_d \) can be attributed directly to the SAR wind direction estimate \( (d) \).

### 3.1. Analysis Cost Function

[30] The true wind field \( x^* \) is unknown, which implies that the SAR and GEM errors \( (E) \) are not directly observable. However, it is still possible to postulate the statistics of these errors and obtain an estimate of the true winds [Portabella et al., 2002]. The postulates can then be checked by verifying that their wind estimates are reasonable. We first presume to know two statistics of \( E \): a mean that is approximately zero and a covariance matrix \( \Sigma \) (the expected value of \( EE^T \)). The diagonal elements of \( \Sigma \) are variances and the off–diagonal elements are spatial and cross–covariances.

[31] It is typical to give less weight in an analysis to the sources of wind information that are expected to have larger errors. This is done in nonlinear regression [Seber and Wild, 1989, Chapter 2] by minimizing the generalized error sum of squares

\[
J(x) = |Y - C(x)|^T \Sigma^{-1} |Y - C(x)| = \|Y - C(x)\|_\Sigma^{-1}, \tag{3}
\]

where weights are determined by the inverse error covariance matrix. The least squares estimator that minimizes \( J \) is our desired wind analysis \( \hat{x} \). It is well known that when \( C \) is linear, \( \hat{x} \) is the minimum variance estimate of the true wind field [e.g., Duncan and Horn, 1972]. Assuming that the statistics of \( E \) also have a Gaussian distribution, \( \hat{x} \) is the maximum likelihood estimate [Lorenc, 1986]. When \( C \) is nonlinear, it is necessary to iteratively evaluate \( C(x) \) as \( J \) is being minimized [Seber and Wild, 1989]. Dowd et al. [2001] employed this method to combine ocean wave information from buoy observations and SAR image spectra. We use the same approach to combine the SAR and GEM wind information.

[32] It is convenient to simplify \( \Sigma \) by defining its \((5N)^2\) elements in terms of only a couple of key parameters. (We then estimate the value of these parameters in the next section.) As in Portabella et al. [2002], we focus on the error variance of the SAR backscatter, but also consider a length scale for its spatial covariance, as well as the error variance of the SAR directional estimates. Our first simplification of \( \Sigma \) is to neglect the cross–covariance between the three types of errors. Outer blocks of \( \Sigma \) are thus assigned zero values and the remaining diagonal blocks are defined by three error covariance matrices \( R \), \( D \), and \( B \) for the SAR backscatter, SAR directions, and GEM winds, respectively. The error sum of squares (3) can then be written

\[
J(x) = \|y - c(x)\|_R^{-1} + \|d - d(x)\|_D^{-1} + \|x - x\|_B^{-1}, \tag{4}
\]
Figure 4. Example of a surface wind analysis at 12.8-km resolution near the Queen Charlotte Islands at 1509 UTC 12 December 2004 (Radosat-1 was descending, so incidence angle and range increase to the left). Shown are (a) SAR observations (shaded at 1-dB intervals with a scale on the right), (b) SAR observations normalized to reduce incidence angle dependence (see text), (c) GEM model winds interpolated to the SAR grid (every fifth wind is shown, with a half barb being 2.5 m s\(^{-1}\) and a full barb being 5 m s\(^{-1}\)), and the analyses of (d) SAR radar cross section and (e) wind. Collocated buoy winds are labeled in Figure 4c and included in Figure 4e, along with the analysis–minus–GEM wind speed difference (shaded at 1-m s\(^{-1}\) intervals with a scale above). Note that the northwest buoy is collocated with a gridbox on the perimeter of the analysis, but is actually located within 50 km to the west.
where each term on the rhs has the form of (3). We refer to (4) as our analysis cost function.

3.1.1. SAR Errors

[33] The error statistics of Radarsat-1 SAR wind information are not well known, although errors in SAR and scatterometer radar cross section may be similar. Scatterometer errors have traditionally been expressed as a percentage of the backscatter itself (in physical or linear units, as opposed to in dB) [e.g., Stoffelen and Anderson, 1993]. More recently, Stoffelen and Anderson [1997b] found that this led to a strong dependence of ERS-1 scatterometer errors on wind direction. Instead of $y$, they considered backscatter in linear units to the power of 0.625, as this reduced the apparent dependence to one on only wind speed and incidence angle. Such a parameterization is employed operationally [Hersbach et al., 2007]. Stoffelen and Anderson [1997a] also found improved wind estimates using a cost function that involved the log of this transformed variable.

[1997a] also found improved wind estimates using a cost function (4) as our analysis cost function.

More recently, Stoffelen and Anderson [1997b] found that this led to a strong dependence of ERS-1 scatterometer errors on wind direction. Instead of $y$, they considered backscatter in linear units to the power of 0.625, as this reduced the apparent dependence to one on only wind speed and incidence angle. Such a parameterization is employed operationally [Hersbach et al., 2007]. Stoffelen and Anderson [1997a] also found improved wind estimates using a cost function that involved the log of this transformed variable.

[34] Portabella et al. [2002] postulated the ERS-2 SAR error variance (diagonal elements of $\mathbf{R}$) to be the square of about 8% of the observed backscatter (in linear units). For backscatter $y$, (in dB units), this suggests an additive error parameterization. However, because we choose to weight by the normalized backscatter, errors do not vary as much with range (Figure 3a) and a multiplicative parameterization (in dB) seems reasonable. Our use of normalized backscatter renders the spatial covariance (off–diagonal elements of $\mathbf{R}$) isotropic and we assume this falls off exponentially with distance according to

$$R(i,j) = W_R^2 y_i y_j \exp \left[ -d(i,j)^2 / 2L_R^2 \right].$$

Here, $i$ and $j$ denote locations on the SAR scene, $W_R$ is a weight that scales the SAR error variance, $y$ is the normalized backscatter, $d$ is the distance between locations $i$ and $j$, and $L_R$ is a SAR covariance length scale. The two parameters that vary are $W_R$ (from a fraction of a percent to about 30%) and $L_R$ (from 0 km to about 150 km). As for the GEM errors below, we distinguish between spatial auto-correlation at zero and nonzero lag, and weight the variance terms more relative to the covariance terms by including the factor 0.8 in $W_R$ for $i \neq j$.

3.1.2. SAR Direction Errors

[35] Koch and Feser [2006] postulated that wind speed is stronger, SAR estimates of wind direction are more similar to those of a numerical model. Although this is subject to the caveat that there are errors in the model winds [cf. Kent et al., 1998; Stoffelen, 1998], it suggests that the error variance for SAR direction should be weighted by the inverse of the normalized backscatter (as the inverse magnitude in dB roughly covaries with wind speed after normalization). On the other hand, an argument can be made that all SAR wind information (as derived from both backscatter and its gradient) should be weighted similarly relative to the GEM information. In that case, error variance should be weighted by the normalized backscatter, as in (5). Both weighting schemes are examined in the next section, but for simplicity, we begin with homogeneous (unweighted) SAR direction error variance according to

$$\mathbf{D}(i,i) = W_D^2.$$

Here, $W_D$ is the same for both along- and cross-track unit–vector errors. This unitless parameter varies from 0.5 to infinite [i.e., no quality is attributed to the directional information and the term weighted by $\mathbf{D}^{-1}$ in (4) becomes zero]. Also, we do not attempt to postulate spatial error covariance for direction (although its length scale may be similar to $L_R$ above) and thus set the off–diagonal elements of $\mathbf{D}$ to zero.

3.1.3. GEM Errors

[36] The error covariance of a numerical model has long been a central quantity in atmospheric data assimilation [e.g., Hollingsworth and Lonnberg, 1986]. Various methods have been used to estimate matrices such as $\mathbf{B}$ for different model variables and dimensions. The “NMC” method [Parrish and Derber, 1992] uses the difference between 24-h and 48-h forecasts, and recently, Monte Carlo approaches have become popular [Houtekamer et al., 2005]. A fixed estimate of $\mathbf{B}$ is employed in this study because this constrains the number of parameters to be specified. (Spatially dependent observation and background covariances are not easily separable.) However, because this defines the weight given to the GEM winds, more realistic error covariance might be expected to yield a better analysis. (A lower resolution model might also be used to produce a better analysis if its errors are well known.)

[37] Our parameterization of the GEM error covariance matrix follows that of Daley [1991, Chapter 5]. We assume that the GEM wind component errors are homogeneous and that the corresponding velocity potential and stream function errors are isotropic, have the same length scale, and are not correlated. The blocks of the $\mathbf{B}$ matrix are then

$$B(u_i, u_j) = A \left[ \cos^2 \phi \left( 1 - \nu^2 \frac{d^2}{L_B^2} \right) + \sin^2 \phi \left( 1 - \frac{d^2}{L_B^2} + \nu^2 \frac{d^2}{L_B^2} \right) \right].$$

$$B(v_i, v_j) = A \left[ \sin^2 \phi \left( 1 - \nu^2 \frac{d^2}{L_B^2} \right) + \cos^2 \phi \left( 1 - \frac{d^2}{L_B^2} + \nu^2 \frac{d^2}{L_B^2} \right) \right].$$

$$B(u_i, v_j) = A \cos \phi \sin \phi \frac{d^2}{L_B^2} \left( 1 - 2\nu^2 \right) = B(v_i, u_j)$$

$$A = W_B^2 \exp \left[ -d^2 / 2L_B^2 \right],$$

where $\phi$ is the angle defined by the two locations $i$ and $j$ (relative to the cross–track direction). The variable $\nu$ is 0.15. This defines the expected ratio of divergent to total kinetic energy and corresponds to a constant Ekman turning angle of about 23° [Polavarapu, 1995]. The remaining variables have been estimated by comparison with buoy observations, following Hollingsworth and Lonnberg [1986]. The length scale $L_B$ is taken to be 350 km based on a fitted correlation function (not shown). The GEM error covariance weight $W_B$ is taken to be $\sqrt{3}$ m s$^{-1}$ for both wind components. On the basis of the GEM–buoy autocorrelation, which reveals enhanced uncertainty at zero lag, we weight the variance terms slightly more relative to the covariance terms by including a factor of 0.95 in $W_B$ for $i \neq j$. These approximations seem consistent with Buehner et al. [2005], wherein for errors in the near–surface horizontal wind vector components, the extratropical variances are approximately 4 m$^2$ s$^{-2}$ and the de-correlation length scale is a few hundred kilometers.

[38] Once the error covariance matrices have been specified, the desired wind analysis (8) that minimizes the cost function (4) can be sought. It is standard practice to perform...
the minimization iteratively by a sequence of linear regressions. At each stage, \( e(x) \) and \( d(x) \) are re-evaluated, along with their derivatives with respect to \( x \). The derivatives provide the local gradient of \( J \) and an iterative search along the lines of steepest descent yields the analysis. Convergence is assumed to occur when the magnitude of the cost function gradient is three orders of magnitude smaller than that of the GEM forecast, which is our first guess for \( x \). We note that for a 12.8-km analysis, most of the computational time (10–15 min) is required not to minimize (4), but initially to reduce the resolution from 25 m to 100 m and then to calculate the SAR estimates of wind direction.

3.2. Analysis Example

[39] It is instructive to analyze an individual SAR scene before examining composite results for the 484 acquisitions. Here, we illustrate that SAR and GEM data can be combined using nonlinear regression and that buoy observations can be used for validation. The example chosen is an offshore wind event of moderate strength centered over the Queen Charlotte Islands. (RadarSat-1 was descending over the west coast of British Columbia, with incidence angle and range increasing to the west). Strong range dependence is apparent in the raw backscatter that was acquired on 12 December 2004 (Figure 4a). This range dependence is reduced by normalization and wind speed information is more apparent (Figure 4b). An increase in both backscatter and wind speed to the south are evident in both this SAR scene and the corresponding GEM winds (Figure 4c). However, the increase in backscatter is mostly to the southeast, and the GEM wind forecast does not resolve a possible wind shadow near buoy 46208.

[40] An estimate of wind direction from the SAR gradient calculation [Koch, 2004] indicates flow around the southern tip of the islands (Figure 4b). (All three initial estimates, using 100-m, 200-m, and 400-m gradients, reveal the same large-scale pattern, but the directions are smoothest using the 200-m gradients.) This orientation of the flow is locally more consistent with the buoy observations than the GEM wind directions, which suggests that the SAR directional information should be weighted strongly.

[41] We choose the free parameters of our SAR error covariance matrices to be: \( W_R = 3.5\% \) (of the local SAR backscatter), \( L_R = 35 \) km, and \( W_D = 0.5 \). These values are selected by a subjective comparison of the buoy observations and the resulting analyses. After tuning the covariance parameters, the analyzed wind directions (Figure 4e) are consistent with the buoys. The wind shadow to the west and strong increase in wind speed to the southeast of the Queen Charlotte Islands are resolved in the analysis as well. It follows that in this case, wind information can be extracted from a SAR acquisition and used to modify a short-term numerical forecast so that better agreement is obtained with in situ observations. A similar approach is employed in the next section to construct large groups of analyses.

4. Results

[42] The general form of the SAR and GEM wind information errors is given in the preceding section. Here, we focus on the SAR error parameters \( W_R, L_R \), and \( W_D \). Our specification of these is examined in terms of similarities between the resulting analyzed winds and buoy observations. It is first convenient to divide the 831 buoy collocations into two groups. This permits the SAR error parameters to be tuned for one group and then analysis–buoy differences to be validated using an independent set of collocations. Collocations are arbitrarily divided into 415 before, and 416 after, the end of June 2005. Both groups are reasonably well distributed over the preceding 13 months and subsequent 9 months, with no significant differences in wind speed or direction.

4.1. SAR Error Parameter Search

[43] The former group of 415 buoy collocations is first used as a training data set to search for reasonable values of \( W_R, L_R \), and \( W_D \). Here, we employ the analysis–buoy differences as a proxy for analysis errors. (Our goal is to minimize the latter, but we only have observations of the former.) Although the buoys themselves have errors, a simple statistical model for the collocations (given in the Appendix) indicates that if the analysis errors, buoy errors, and true wind field are indeed independent (i.e., their covariance is approximately zero) and the buoy error variance is not large, then analysis–buoy variance is a good proxy for analysis error variance. We thus seek to identify better SAR error parameters by minimizing the analysis–buoy variance.

[44] Figure 5 illustrates the dependence of analysis–buoy standard deviation on our three SAR error parameters. A wind vector can be described equivalently in terms of speed and direction or zonal and meridional components, so all four measures are shown. It is evident that minima in one component (unshaded values) does not necessarily correspond to minima in the others, although the correspondence generally suffices to determine a single choice of \( W_R \), \( L_R \), and \( W_D \). The circled values are the global minima for each wind component. There is consensus among the four components that \( L_R \) is between about 15 km and 35 km. An estimate of this length scale for SAR has not previously been given. Scatterometer errors, by comparison, do not appear to have any spatial error covariance [Stoffelen and Anderson, 1997a]. If we choose \( L_R = 0 \) here, analyses do compare better with the buoys than the GEM winds, at least in terms of the speed and zonal wind components (lighter shading). However, all wind components indicate that \( L_R \approx 15 \) km would be a more appropriate choice.

[45] The values of \( W_R \) that minimize the analysis–buoy standard deviation range between 0.5% to 3.5% of the SAR backscatter. (We employ \( W_R = 1.5\% \) below.) This represents a smaller error than that given by Portabella et al. [2002] and implies that SAR backscatter should generally be weighted strongly in the analysis. Analysis–buoy standard deviation varies most as a function of \( W_R \), and this makes the choice of this parameter relatively easy. Standard deviation varies least with \( W_D \), however, and this is a more difficult parameter to estimate. Analyses that are constructed without any wind streak information (\( W_D \) infinite) are similar in a statistical sense to those that weight this information strongly. We have experimented with different parameterizations of these errors (i.e., weighting by normalized backscatter or its inverse), but find no discernable difference in the magnitude or variation of standard deviation. There are many possible explanations for this, but one is that errors in the wind streak information may vary consid-
erably from scene to scene, as previous studies have sug-
gested [e.g., Koch and Feser, 2006]. Despite this caveat,
there appears to be some benefit to weighting this informa-
tion strongly. The values of $WD$ that minimize the analysis–
buoy standard deviation are not prominent, but range be-
tween 0.5 to 3.5, so we employ $WD = 3.5$ below.

4.2. Analysis Validation

The latter group of 416 buoy collocations (Figures 6a–
6d) can be used as a validation data set to compare our analy-
yses with other estimates of the true wind field. We refer to the
others as the SAR–speed GEM–direction, GEM–speed
GEM–direction, and SAR–speed SAR–direction wind esti-
mates. As with the nonlinear regression approach, each of
these alternative approaches makes some assumptions about
the errors in SAR and GEM data. The first assigns infinite error
to all sources of wind information except the SAR wind speed
and GEM wind direction, which are assumed to have no errors
(i.e., CMOD is used to determine wind speed, given backscat-
ter and the GEM wind direction). The second approach

<table>
<thead>
<tr>
<th>$W_D$ (%) of SAR Obs</th>
<th>$W_R$ (km)</th>
<th>$W_D = 0.5$</th>
<th>$W_D = 1.5$</th>
<th>$W_D = 3.5$</th>
<th>$W_D = 7.5$</th>
<th>$W_D = \infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.5</td>
<td>0.9</td>
<td>0.303 4.15</td>
<td>0.31 4.12</td>
<td>0.31 4.26</td>
<td>0.31 4.48</td>
<td>0.31 4.51</td>
</tr>
<tr>
<td>7.5</td>
<td>1.7</td>
<td>0.295 0.76</td>
<td>0.29 0.69</td>
<td>0.29 0.59</td>
<td>0.29 0.49</td>
<td>0.29 0.46</td>
</tr>
<tr>
<td>3.5</td>
<td>1.1</td>
<td>0.295 0.76</td>
<td>0.29 0.69</td>
<td>0.29 0.59</td>
<td>0.29 0.49</td>
<td>0.29 0.46</td>
</tr>
<tr>
<td>1.5</td>
<td>0.6</td>
<td>0.295 0.76</td>
<td>0.29 0.69</td>
<td>0.29 0.59</td>
<td>0.29 0.49</td>
<td>0.29 0.46</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.295 0.76</td>
<td>0.29 0.69</td>
<td>0.29 0.59</td>
<td>0.29 0.49</td>
<td>0.29 0.46</td>
</tr>
<tr>
<td>1e-4</td>
<td>0.3</td>
<td>0.295 0.76</td>
<td>0.29 0.69</td>
<td>0.29 0.59</td>
<td>0.29 0.49</td>
<td>0.29 0.46</td>
</tr>
</tbody>
</table>

Figure 5. Standard deviation of differences between analyzed winds and buoy observations using the first
group of 415 collocations for training (see text). Wind speed (m s$^{-1}$) is the far–left column, meteorological
direction (degrees) is the middle–left column, zonal wind (m s$^{-1}$) is the middle–right column, and meridional
wind (m s$^{-1}$) is the far–right column. Standard deviation is shown as a function of the SAR error variance
weight $W_R$ (ordinates), the SAR error covariance length scale $L_R$ (abscissae), and the SAR directional error
variance weight $W_D$ (rows). Parameter values increase exponentially (as a percentage of normalized SAR
backscatter, in km, and without units, respectively). Dark shading denotes values that are larger than the
corresponding GEM–buoy standard deviations. Light shading denotes values that are smaller. Unshaded
values are the local minima for each panel and those that are circled are the global minima.

Figure 5. Standard deviation of differences between analyzed winds and buoy observations using the first
group of 415 collocations for training (see text). Wind speed (m s$^{-1}$) is the far–left column, meteorological
direction (degrees) is the middle–left column, zonal wind (m s$^{-1}$) is the middle–right column, and meridional
wind (m s$^{-1}$) is the far–right column. Standard deviation is shown as a function of the SAR error variance
weight $W_R$ (ordinates), the SAR error covariance length scale $L_R$ (abscissae), and the SAR directional error
variance weight $W_D$ (rows). Parameter values increase exponentially (as a percentage of normalized SAR
backscatter, in km, and without units, respectively). Dark shading denotes values that are larger than the
corresponding GEM–buoy standard deviations. Light shading denotes values that are smaller. Unshaded
values are the local minima for each panel and those that are circled are the global minima.
assumes no errors in the GEM winds, but assigns infinite errors to (and makes no use of) SAR wind information. Conversely, the third approach makes no use of the GEM wind information. Instead, we choose one of the two SAR wind streak directions (i.e., the one that is within 90° of the validating buoy wind direction) and use CMOD to determine a SAR wind speed, as in the work of Koch and Feser [2006]. The error assumptions made by these alternative approaches can be seen as special cases of our more comprehensive approach. For this comparison, our regression analyses employ $WR = 1.5\%$, $LR = 15$ km, and $WD = 3.5$.

Figure 6. Frequency distributions as in Figure 2, but using the latter group of 416 buoy collocations for validation (see text). Shown are wind speed, meteorological direction, and zonal and meridional components (grey), as well as differences relative to the buoys (black). Distributions are (a–d) buoys, and analyses based on (e–h) nonlinear regression, (i–l) SAR wind speed and GEM wind direction, m–p) GEM wind speed and direction, and (q–t) SAR wind speed and direction. The mean and standard deviation values in each panel refer to (a–d) buoys and (e–t) differences relative to the buoys.

[47] Frequency distributions for the four analysis approaches (grey) and their differences relative to the buoys (black) are shown in Figures 6e–6t. The standard deviation of analysis–buoy differences can again be employed as a proxy for analysis errors. According to this measure, the four approaches are arranged with generally decreasing analysis errors from bottom to top. Relative to the SAR wind direction (Figure 6r), the GEM wind direction is generally a better estimate (Figures 6j and 6n). In turn, the GEM wind direction and SAR backscatter define, according to (2), a wind speed estimate (Figure 6i) that is more accurate than that of GEM (Figure 6m). This is also true of the zonal wind component (Figures 6k and 6o), but not of the meridional component (Figures 6l and 6p).

[48] It is evident from the SAR–speed GEM–direction winds that this combination of data compares well with buoy observations, which implies that errors in both wind estimates are not large. However, a further reduction in these differences is obtained by weighting the SAR and
GEM wind information according to their expected error covariance. Discrepancies with the buoy observations are thus minimized using nonlinear regression (Figures 6e–6h).

Our comparison of approaches can be re-examined by switching the training and validation data sets. The latter group of 416 collocations is instead used to perform a SAR error parameter search, as in Figure 5 (not shown). This confirms that $W_R = 1.5\%$ and $L_R = 15$ km are appropriate, but the SAR directional information can be weighted slightly more, so we choose $W_D = 1.5$. The standard deviation of analysis–buoy differences can then be calculated, as in Figures 6e–6t, but for the earlier group of 415 collocations. These values are given in Table 1. The nonlinear regression approach again minimizes the discrepancies with each buoy wind component, except that in this comparison, the GEM meridional wind is more consistent with the buoys. (The GEM wind speed is also more consistent than the SAR–speed GEM–direction estimate, but otherwise, the relative rankings are unchanged.) As an aside, another gauge for the nonlinear regression wind speed is the SAR wind speed that is computed using the buoy wind direction [cf. Vachon and Dobson, 2000]. The SAR–speed buoy–direction standard deviation values of 1.71 m s$^{-1}$ and 1.88 m s$^{-1}$ are also larger than the non-linear regression values of 1.60 m s$^{-1}$ (Figure 6e) and 1.55 m s$^{-1}$ (Table 1), respectively.

Finally, it is instructive to compare our wind component differences with those of previous studies. Kerbaol [2006] employs error postulates for the numerical model winds and also reports an improvement in wind speed relative to 908 buoy collocations (i.e., an RMS value of 1.84 m s$^{-1}$ versus the corresponding SAR–speed GEM–direction value of 2.16 m s$^{-1}$). This can be compared to the SAR–speed GEM–direction wind speed standard deviation values of 1.72 m s$^{-1}$ (Figure 6i) and 1.99 m s$^{-1}$ (Table 1), which are roughly equivalent to the value of 1.76 m s$^{-1}$ given by Monaldo et al. [2001].

The SAR–speed SAR–direction approach yields wind direction standard deviation values of 35.16° (Figure 6r) and 35.77° (Table 1), which are larger than the values of 21.6° and 17.6° given by Horstmann et al. [2002] and Koch and Feser [2006], respectively. This can be attributed in part to our inclusion of the subset of SAR scenes that may have little directional information (e.g., Koch and Feser [2006] omit 10 of 80 dates) and to our objective choice of the smoothest directional estimate. (Standard deviation values are about the same when using the 200-m gradient directions instead. They are slightly larger when using the 100-m or 400-m gradient directions.) It is clear from Koch and Feser [2006] and Figure 4b that the high–resolution SAR backscatter gradient contains information about wind direction, although it may be necessary to gauge directional errors on a scene–by–scene basis [cf. Koch, 2004]. For some scenes, it may also be appropriate to mask linear patterns that have the same orientation as ocean surface waves [cf. Dowd et al., 2001].

### Table 1. Standard Deviation of Analysis–Buoy Differences

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Speed, m s$^{-1}$</th>
<th>Dir, degree</th>
<th>U, m s$^{-1}$</th>
<th>V, m s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonlinear regression</td>
<td>1.55</td>
<td>29.52</td>
<td>2.11</td>
<td>2.52</td>
</tr>
<tr>
<td>SAR speed GEM direction</td>
<td>1.99</td>
<td>29.94</td>
<td>2.17</td>
<td>2.97</td>
</tr>
<tr>
<td>GEM speed GEM direction</td>
<td>1.85</td>
<td>29.94</td>
<td>2.36</td>
<td>2.40</td>
</tr>
<tr>
<td>SAR speed SAR direction</td>
<td>2.42</td>
<td>35.77</td>
<td>3.17</td>
<td>4.19</td>
</tr>
</tbody>
</table>

*As in Figures 6e–6t, but using the earlier group of 415 buoy collocations for validation (see text).*

### 5. Conclusions

Postulates of SAR wind information error statistics have been employed in a nonlinear regression approach to surface marine wind analysis. This has enabled the construction of analyses with minimum error variance using SAR acquisitions and Global Environmental Multiscale (GEM) model forecasts. Two SAR–wind relationships have been considered. The first is the European Remote Sensing (ERS) C-Band model with the Radarsat-1 polarization correction of Vachon and Dobson [2000]. This provides an estimate of SAR backscatter given wind speed and direction. The second is based on the proposition that SAR acquisitions resolve coherent wind streak patterns at $O[100$-m] scales that can be used to determine wind direction. This 180°-ambiguous wind direction is derived from the SAR backscatter gradient [Koch, 2004].

The feasibility of using linear patterns in SAR images as a proxy for wind direction has traditionally been attributed to the presence of boundary layer rolls [cf. Etting and Brown, 1993]. Wind streaks should also act as a proxy for wind direction and these may also be more prevalent [Drobinski and Foster, 2003; Koch and Feser, 2006]. A SAR is one of few spaceborne instruments capable of resolving structures with scales of $O[100$ m]. Although a complete physical description of the patterns found at this scale remains to be clarified, our results provide further support that for wind estimation, there is some benefit to considering this information. Moreover, subsequent improvements in both wind direction estimation and in backscatter–wind relationships should be relatively easy to absorb into our analysis framework.

A total of 484 SAR scenes have been examined for coastal regions of eastern and western North America. Each objective analysis has employed good quality SAR and GEM data at 12.8-km and 15-km resolution, respectively. These data were first transformed so that SAR backscatter error bias would be reduced [Monaldo et al., 2001; Beal et al., 2004] and wind errors would be more Gaussian in character [Stoffelen, 1998]. Analyses were then constructed by nonlinear regression, in which the weighted sum of squares of the differences between analyses and data were minimized. The weights were defined by the expected error covariance of the SAR and GEM data [Seber and Wild, 1989].

[55] This study has focused on a rudimentary parameterization of the SAR error covariance. Error covariance for SAR backscatter and wind direction inferred from its gradient were expressed in terms of three variables. Backscatter error variance was weighted by $W_R$, as a percentage of the backscatter itself, and spatial covariance was assumed to fall off exponentially with a length scale $L_R$. The SAR
directional error variance was defined by a unitless parameter $W_D$. (Its spatial covariance was ignored.) We did not focus on the GEM error covariances, which were fixed according to Daley [1991], and we assumed negligible cross–covariance between the SAR and GEM errors. A range in each of the three SAR error parameters was employed to identify the best combination. This was found by minimizing analysis–buoy differences (a proxy for analysis errors). Of the 831 analysis and buoy collocations that were available, one half was used to tune the SAR error parameters. The other half was used to compare nonlinear regression analyses and more conventional SAR and GEM wind analyses.

It was found to be only somewhat beneficial to include the SAR directional information in the analyses ($W_D \approx 3.5$). It was easier to identify backscatter error parameters ($W_R$ and $L_R$) that minimized the analysis–buoy differences. Backscatter appeared to require a small error (or strong weight) of only a few percent of the backscatter itself. An error covariance length scale of between 15 km and 35 km was found to be appropriate. This seems consistent with scatterometer errors, which have a resolution of just lower than 35 km and are assumed to have little spatial covariance [Stoffelen and Anderson, 1997a]. We expect subsequent studies to combine SAR and model wind information at increasingly higher resolution, and when this happens, spatial covariance will become an increasingly important constraint as the effective degrees of freedom in determining the analysis should remain roughly constant.

Comparisons were also made with conventional methods of combining the SAR and GEM data. It is perhaps not surprising that the winds based on the SAR data alone were least similar to the buoy observations that were used for validation. The GEM winds compared well with the buoys, and sometimes better than the combination of SAR–derived wind speed and GEM wind direction. However, the nonlinear regression wind estimates were generally most consistent with the buoys. In particular, the analysis–buoy wind speed difference has a standard deviation of about 1.6 m s$^{-1}$. This agreement would deteriorate, however, by not considering spatial error covariance ($L_R = 0$) or by postulating larger SAR backscatter errors (e.g., $W_R = 15.5\%$). Appropriate wind improvements could thus be identified, but were found to be sensitive to the choice of SAR error statistics.

Surface wind information that is detected by synthetic aperture radar is often underutilized when the marine environment is assessed [Friedman et al., 2001]. A formal context for objective analysis has been outlined in this study using a nonlinear regression framework. The demonstrated improvements that result from this approach can be attributed to the guiding principle that information should be weighted more if it is expected to have smaller errors. Our specification of these errors, including their spatial and cross–covariance, has been somewhat rudimentary. This includes our GEM error specification, which affects the quality of the resulting analyses and perhaps even our estimation of the SAR error parameters of interest. A more appropriate context to examine this is in the assimilation step of the GEM forecast system [Laroche et al., 1999; Choisnard and Laroche, 2008], where the GEM wind errors are time varying (i.e., scene–dependent) and are presumably better represented. Such an approach would also permit an exploration of the impact of assimilating SAR observations on forecasts, particularly for coastal marine regions that are otherwise poorly observed.

Appendix A: Validation Using Noisy Buoy Observations

Moored buoys are known to suffer from errors, including the errors of representativeness that arise when the true wind analysis is taken to be at a resolution of O[10 km]. Despite the enhanced variance that buoys resolve, they may still be useful to identify groups of analyses that are closer to the true wind field in a statistical sense. This study considers pairs of analysis groups that differ only in terms of their error postulates. If each group is compared to the same group of buoy observations, then because the buoy errors are constant, any relative change in analysis–buoy differences may be attributed to analysis errors. In principle, this allows better error postulates to be identified. Here, we quantify the assumptions that are implicit in this approach to validation.

It is instructive to consider an analyzed wind component (i.e., one of either the zonal, meridional, speed, or direction component) and its collocated buoy observation in terms of the error representation given by Stoffelen [1998]. These two wind estimates are

$$
\begin{align*}
a &= s_a(t + e_a) \\
b &= s_b(t + e_b),
\end{align*}
$$

where $t$ is the true wind, $s_a$ and $s_b$ are unknown constants that permit a first–order calibration of $a$ and $b$, and $e_a$ and $e_b$ are the analysis and buoy errors, respectively. A statistical measure of proximity to the true wind is the variance of differences between the analyses and the buoys. If our collocations provide a good measure of this quantity and (A1) applies, the individual contributions are

$$
\begin{align*}
\text{Var}(a - b) &= s_a^2 \text{Var}(e_a) + 2s_a(s_a - s_b) \text{Cov}(t, e_a) \\
&\quad - 2s_a s_b \text{Cov}(e_a, e_b) + s_b^2 \text{Var}(e_b) \\
&\quad - 2s_b(s_a - s_b) \text{Cov}(t, e_b) + (s_a - s_b)^2 \text{Var}(t),
\end{align*}
$$

where $\text{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2$ and $\text{Cov}(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle$ are the variance and covariance, respectively, and $\langle \cdot \rangle$ denotes the expectation operator.

The analysis and buoy wind estimates are independent, so we can assume that $e_a$ and $e_b$ are uncorrelated with each other and with $t$. We can also assume that the analyzed and buoy winds are essentially unbiased (i.e., $\langle e_a \rangle = \langle e_b \rangle = 0$) owing to the small differences in their means (Figure 6). It follows that the second, third, and fifth term on the rhs of (A2) are negligible.

The fourth and sixth terms on the rhs of (A2) may be large, but the buoy and true winds do not vary from one group of analyses to another. Although $s_a$ and $s_b$ are unknown, it is reasonable to assume that they are both within about 10% of unity [Stoffelen, 1998]. If $s_a$ has only weak dependence on our SAR error postulates, then changes in the fourth and sixth term are also both negligible. It follows that for any pair of analysis groups, the change in
(A2) is mainly defined by the first term on the rhs, and specifically, by \( \varepsilon_i \). More generally, if variations in the first term on the rhs (our signal) are not dominated by the other rhs terms, including buoyy errors (our noise), then better error postulates should be identifiable among groups of analyses. This provides a justification for our selection in section 4 of the analyses whose variance (A2) is smaller.

Acknowledgments. V. Kerbal and an anonymous reviewer are thanked for comments that led to improvements in our presentation. This research has been sponsored by the Canadian Space Agency. The SAR scenes have been obtained from the Gatineau office of MacDonald Dettwiler and Associates Ltd. and constitute many of the scenes acquired by the Integrated Satellite Tracking of Oil Polluters (ISTOP) program of Environment Canada. John Wolfe at the Canadian Centre for Remote Sensing kindly provided code to calculate radar cross section. Luc Fillion and Mark Buchner provided assistance with the cost function minimization. Bridget Thomas is thanked for providing algorithms to vertically adjust buoyy wind observations. These buoyy data were obtained from the Pacific–Yukon regional archives of Environment Canada, the United States National Oceanographic Data Center (www.nodc.noaa.gov), and the Gulf of Maine Ocean Observing System (www.gomoos.org).

References

Alpers, W., and B. Brümmer (1994), Atmospheric boundary layer winds observed by the synthetic aperture radar aboard the ERS-1 satellite, J. Geophys. Res., 99(C6), 12,613–12,621.


---

R. E. Danielson and H. Ritchie, Department of Oceanography, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia, Canada B3H 4J1. (rick@phys.ocean.dal.ca)

M. Dowd, Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, Canada.