Two-dimensional variational ambiguity removal in wind scatterometry

Jur Vogelzang, Ad Stoffelen, John de Vries
Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

Hans Bonekamp
EUMETSAT, Darmstadt, Germany

Corresponding author address: Jur Vogelzang, Royal Netherlands Meteorological Institute, Wilhelminalaan 10, 3732 GK De Bilt, The Netherlands.
E-mail: vogelzan@knmi.nl

ABSTRACT

Wind scatterometry is an important technique for retrieving high-resolution near-surface wind information in ocean areas where other techniques provide sparse coverage. Scatterometer winds are principally ambiguous in wind direction. In this paper a variational method for ambiguity removal, named 2DVAR, is presented. 2DVAR performs an incremental analysis based on the ambiguous scatterometer wind vector solutions and a model forecast, and selects the ambiguity closest to the analysis as solution. The correctness of the 2DVAR implementation is demonstrated to machine precision. The merits of 2DVAR are shown in a number of case studies and in a more general statistical comparison. SeaWinds observations at 25 km resolution are known to be noisy, especially in the nadir part of the swath, due to the observation geometry. It is shown that the noise is effectively suppressed by application of 2DVAR in combination with the Multi Solution Scheme (MSS). MSS retains the local wind vector probability density function after inversion, rather than only a limited number of ambiguous solutions. The effect of MSS is to increase the influence of the background. This can be decreased again by adjusting the parameters in the observation wind error model. A case study on an extratropical hurricane observed with SeaWinds shows that reliable wind estimates can be obtained wind speeds more than 40 m/s. ASCAT has a better observation geometry than SeaWinds and therefore the MSS is not needed, but a case study on a tropical cyclone shows that the 2DVAR error model should be tuned right to reproduce the correct circulation patterns.

1. Introduction

Wind scatterometry is a widely used technique for measuring global ocean surface winds from space. Current operational applications include assimilation into global models for numerical weather prediction like that of the European Centre for Medium range Weather Forecasting (ECMWF) (Hersbach, 2007) and detection of tropical and extratropical hurricanes for marine warning (Sienkiewicz et al., 2007). At this moment three scatterometers are operational: SCAT, ASCAT, and SeaWinds. SCAT and ASCAT are three-beam instruments operating at C-band with VV polarization and 45° separation in azimuth angle between the beams. SeaWinds is a rotating fan HH and VV polarized scatterometer at Ku-band. The advantage of SeaWinds over ASCAT is a larger swath width, but its disadvantages are increased rain sensitivity and swath-dependent measurement geometry, which is particularly unfavourable in the nadir part and the outer parts of the swath.

Scatterometer measurements are assumed to be described by a Geophysical Model Function (GMF), an empirical function that gives the radar cross section as a function of wind speed and direction at 10 m anemometer height, incidence and azimuth angles, radar frequency, and polarization. Numerical inversion of the GMF yields the scatterometer wind measurement. Due to the nature of the GMF, this procedure generally produces more than one solution. These multiple solutions are referred to as ambiguities. If the scatterometer observations are to be assimilated in a numerical weather prediction (NWP) model, the ambiguities and their a-priori probabilities can be fed into the variational data assimilation scheme of that model to be combined with other

observations (Stoffelen and Anderson, 1997). If, on the other hand, the scatterometer observations are intended as stand-alone information source, it is necessary to select the solution that is most likely the correct one. This is done in the ambiguity removal (AR) step. A number of ambiguity removal methods has been proposed. A simple method is the first rank method where the solution with the highest a-priori probability is selected. Another simple method is the closest-to-background method. This method requires the presence of a NWP wind field, and the solution closest to the model prediction is selected. The closest-to-background method runs into difficulties in cases where the model prediction is wrong.

Several more sophisticated AR schemes were evaluated before the launch of the first European Remote Sensing (ERS-1) satellite (Graham et al. 1989). A scheme called CREO (Cavanié and Lecomte, 1987) was selected and implemented by the European Space Agency (ESA). CREO constructs two antiparallel fields (since ERS scatterometer data generally contain two ambiguities per Wind Vector Cell, WVC) and selects the field with the most first rank selections. Stoffelen and Anderson (1995) showed that CREO has severe shortcomings. As an alternative, they proposed the PreScat scheme (Stoffelen and Anderson, 1997) which is based on the SLICE scheme (Offiler, 1987). PreScat uses a 5 by 5 WVC box filter to choose the wind direction at the centre as the most likely weighted average of the surrounding ambiguities. The method is initialised with the closest-to-background solution. The Jet Propulsion Laboratory of the National Aeronautics and Space Administration (NASA-JPL) developed an ambiguity removal method for SeaWinds, called Directional Interval Retrieval with Threshold Nudging (DIRTH). This method combines the backscatter measurements into likely wind direction

intervals (rather than solutions) by accounting for their probabilities. A NWP model guess and a constraint of wind direction continuity is subsequently used to construct a consistent wind field (Stiles et al., 2002).

The ambiguity removal problem can also be solved within a variational approach. This requires availability of a model prediction of the wind field (background). An analysis wind field is constructed from the observations and the background by minimizing a cost function which may contain constraints on smoothness, statistical consistency, physical consistency, etc. The solution closest to the analysis is selected (so such methods may as well be referred to as closest-to-analysis).

The VARscat algorithm was developed for processing scatterometer measurements (Roquet and Ratier, 1986; Leru, 1999) and to improve the operational scheme used at the Institute Français de Recherche pour l'Exploitation de la Mer (IFREMER) (Quilfen and Cavanié, 1991). It is a variational method minimising a heuristic cost function. Another variational method is the successive corrections ambiguity removal (SCAR) developed at the Danish National Meteorological Institute (DNMI). *Hoffman et al*. (2003) present a two-dimensional variational method with a cost function consisting of seven terms for filtering and dynamical consistency. The minimisation problem is not preconditioned: minimisation is done in position space in terms of the wind analysis increments. It is also possible to input σ_0 values, so inversion may be included in this method. It compares well to a median filter ambiguity removal technique when applied to data from the National Aeronautics and Space Administration (NASA) Scatterometer (NSCAT) as shown by *Henderson et al*. (2003).

The PreScat, VARScat, SCAR and 2DVAR methods were compared for ERS-1 data by *Stoffelen et al.* (2000). These authors concluded that on average VARScat had best performance. PreScat was generally best in the tropics, SCAR in the Northern Hemisphere during winter, and VARScat on the Southern Hemisphere during summer. Two tunable properties, i.e., activity and spatial scope, in each ambiguity removal scheme were compared and shown to strongly affect the results. It was recommended to make spatial scope geographically dependent and to tune activity of the AR schemes in order to improve its performance. Since tuning of 2DVAR is both methodically and practically relatively simple, the 2DVAR parameters were tuned in order to improve AR performance. Moreover, its formulation and implementation were reviewed thoroughly and some errors were corrected.

In this paper 2DVAR is presented, a two-dimensional variational ambiguity removal technique that takes the a-priori probabilities of the ambiguities and the known error characteristics of observations and background into account. The advantage of 2DVAR is that it leads not only to spatially consistent wind fields that satisfy basic physical laws, but also to statistically consistent fields in which a WVC solution with high a-priori probability is more likely to be selected than one with low probability. The main differences of 2DVAR with respect to other similar methods are:

- minimalisation is performed in spectral space, thus optimising all spatial scales simultaneously;
- the problem is preconditioned, so inversion of the background error correlation matrix is trivial;

- the a-priori probabilities of the ambiguities are properly taken into account into the cost function, and therefore the selection probability increases with a-priori probability.

As stated before, the observation geometry of SeaWinds changes along the swath. In the nadir part, this leads to broad minima when inverting the GMF. To account for this, the Multiple Solution Scheme (MSS) retains the local wind vector probability density function after inversion, rather than only a limited number of ambiguous solutions at the local minima (Portabella, 2002).

The aim of the paper is twofold: presentation of the 2DVAR method and investigation of its behaviour and sensitivity to changes in the underlying error model for different types of scatterometer. The 2DVAR method is described in section 2, but more details can be found in (*Vogelzang*, 2007). Section 3 describes two tests that prove the correctness of the current 2DVAR implementation: the single observation test and the edge analysis. The single observation test is a case involving only one observation. The minimalization problem can now be solved analytically, and the analytical solution is reproduced to machine accuracy. The edge analysis shows that the incremental analysis does not suffer from under- or overfitting. A dataset consisting of on month of SeaWinds data is analysed in section 4. The data were processed with the SeaWinds Data Processor (SDP) version 1.5 with and without the MSS and using as background the winds predicted by the National Centers for Enviromnental Prediction (NCEP) and the ECMWF. It is shown that the choice of background has little effect and that 2DVAR in combination with the MSS effectively reduces the noise known to be present in the SeaWinds measurements. The noise level is estimated quantitatively for each WVC by

extrapolation of the autocorrelation function. Section 5 contains three case studies to the effect of the parameters in the 2DVAR error model. The first study is of a cyclone with a strong front observed by SeaWinds in the southern Pacific. The second case is that of an extratropical hurricane with wind speeds over 40 m/s, also observed by SeaWinds in the northern Pacific. The last case study deals with an ASCAT observation of a tropical cyclone east of India. The paper ends with the conclusions in section 6.

2. 2DVAR

a. Definitions

The basic idea behind 2DVAR is to fit the scatterometer observations and a model prediction (background) in a weighted field (analysis), and to select that local ambiguous solution that lies closest to the analysis. Such a procedure, basically following the approach of Daley (1991), requires knowledge on the error characteristics of observations and background. The observation error has been treated by Stoffelen (1998). Moreover, Portabella and Stoffelen (2004) show how local scatterometer wind vector ambiguities can be assigned an a-priori probability based on their distance to the Geophysical Model Function (inversion residual). The error characteristics of the background are known and monitored on a routine base at NWP centres (e.g., NWPSAF web site).

Level 2 scatterometer observations are processed onto a regular grid called swath. Ambiguity removal is usually done on a limited set of scatterometer observations. In 2DVAR such a set, referred to as a batch, measures 1900 km in the across track direction and 2200 km along track. The equidistant two dimensional batch grid follows the Earth's surface. Moreover, the batch grid is extended in both directions to have the observations

surrounded by empty grid cells, resulting in a total size of 3200 km in both directions. Note that 2DVAR operates in a local coordinate scheme with its x-axis perpendicular to the satellite moving direction and its y-axis parallel to it. In this section we will consider the wind vector components perpendicular and parallel to the satellite moving direction, t and l, respectively. The local rotation angle of the 2DVAR batch grid can be found with sufficient precision from the known positions of the WVC's (Vogelzang, 2006b).

b. Formulation of the problem

Suppose that inversion and quality control resulted in a set of possible scatterometer wind solutions (ambiguities) at all WVC grid points stored in a state vector $\mathbf{v}_o^k$ with ambiguity index k. Suppose also that the background information is contained in a state vector $\mathbf{x}_b$. The conditional probability that $\mathbf{x}$ expresses the true state of the surface wind field given the ambiguities $\mathbf{v}_o^k$ equals $P(\mathbf{x}|\mathbf{v}_o^k)$. This can be related to the background as (Lorenc, 1986)

$$P(\mathbf{x}|\mathbf{v}_o^k) \propto P(\mathbf{v}_o^k|\mathbf{x})P(\mathbf{x}|\mathbf{x}_b), \quad (1)$$

where $P(\mathbf{v}_o^k|\mathbf{x})$ is the conditional probability that the ambiguous scatterometer wind solutions $\mathbf{v}_o^k$ are observed given the state vector $\mathbf{x}$ and $P(\mathbf{x}|\mathbf{x}_b)$ is the conditional probability that $\mathbf{x}$ represents the surface wind field given $\mathbf{x}_b$. The most likely state vector $\mathbf{x}$ is the analysis. It is found by maximizing (1), or, equivalently, minimizing the cost function J given by

$$J(\mathbf{v}_o^k, \mathbf{x}, \mathbf{x}_b) = -2\ln P(\mathbf{v}_o^k|\mathbf{x}) - 2\ln P(\mathbf{x}|\mathbf{x}_b). \quad (2)$$

To increase the computational efficiency of 2DVAR, analysis increments $\delta\mathbf{x}$ are used rather than the state vector $\mathbf{x}$ itself (incremental formulation),

$$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b \,,\; \delta\mathbf{v}^k = \mathbf{v}_o^k - \mathbf{x}_b \,. \quad (3)$$

For each scatterometer observation the background field is assumed to be known at the same position and time, if necessary from interpolation. The cost function (2) can be rewritten as

$$J(\delta\mathbf{v}^k, \delta\mathbf{x}) = J_o(\delta\mathbf{v}^k, \delta\mathbf{x}) + J_b(\delta\mathbf{x}), \quad (4)$$

with J_o the observational term and J_b the background term.

c. Definition of the cost function

In terms of the analysis increments on the batch grid in the spatial domain, the observation cost function reads (Stoffelen and Anderson, 1997)

$$J_o = \sum_{i,j=1}^{N_1,N_2} \left[\sum_{k=1}^{M_{ij}} \left(\frac{\left(\delta t_{ij} - \delta t_{ij,k}^{(o)}\right)^2}{\sigma_t^2} + \frac{\left(\delta l_{ij} - \delta l_{ij,k}^{(o)}\right)^2}{\sigma_l^2} - 2\ln P_k \right)^{-\lambda} \right]^{-1/\lambda} , \quad (5)$$

with (i, j) the indices of the batch grid cell, N_1 and N_2 the number of batch grid cells parallel and perpendicular to the satellite moving direction, respectively, and M_{ij} the number of ambiguities in cell (i, j). Further, δt_{ij} and δl_{ij} stand for the components of the analysis increments at cell (i, j) perpendicular and parallel to the satellite moving direction, respectively. Similarly, $\delta t_{ij,k}^{(o)}$ and $\delta l_{ij,k}^{(o)}$ stand for the observation increments of the ambiguity with index k. In (5) σ_t and σ_l stand for the expected standard deviation of the error in the scatterometer wind components. Both for SeaWinds and ASCAT

$\sigma_t = \sigma_l = 1.8$ m/s. The inversion and quality control procedures give P_k, the a-priori probability of ambiguity number k being the correct solution (see also section 2e). The parameter λ is an empirical parameter that gives optimal separation between multiple solutions for $\lambda = 4$.

Assuming that the errors in the background wind field are Gaussian, the background term of the cost function can be written in state vector representation as the quadratic form

$$J_b(\delta\mathbf{x}) = (\delta\mathbf{x})^T \mathbf{B}_{\delta\mathbf{x}}^{-1}(\delta\mathbf{x}) + C, \quad (6)$$

with $\mathbf{B}_{\delta\mathbf{x}}$ the matrix of background wind error covariances, the subscript indicating that it is defined in the spatial domain. The superscript T indicates the transpose of a vector or matrix. Note that the transpose suffices here since $\delta\mathbf{x}$ is a real vector. In the general case the Hermitian conjugate (complex conjugate of the transpose) should be taken. The constant C may be neglected during minimalization. Evaluation of (6) requires inversion of $\mathbf{B}_{\delta\mathbf{x}}$, which may be time consuming since it is not diagonal. It is more efficient to transform the background cost function to the spatial frequency domain with a Fourier transformation $\mathbf{F}$ and to express it in terms of stream function and velocity potential increments using an inverse Helmholz transformation $\mathbf{H}^{-1}$. This leads to

$$J_b = (\delta\boldsymbol{\xi})^{*T} \mathbf{B}_{\delta\boldsymbol{\xi}}^{-1}(\delta\boldsymbol{\xi}), \quad (7)$$

with $\delta\boldsymbol{\xi} = \mathbf{H}^{-1}\mathbf{F}\delta\mathbf{x}$. In this representation, the background error correlation matrix is (approximately) block-diagonal and can be factorized into error variances $\boldsymbol{\Sigma}$ and error correlations $\mathbf{P}$ by $\mathbf{B}_{\delta\xi} = \boldsymbol{\Sigma}^{*T}\mathbf{P}\boldsymbol{\Sigma}$ with

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_\chi & 0 \\ 0 & \sigma_\psi \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} \rho_{\chi\chi} & 0 \\ 0 & \rho_{\psi\psi} \end{pmatrix}. \text{ (8)}$$

Since $\boldsymbol{\Sigma}$ and $\mathbf{P}$ are real, it is possible to condition the problem by writing $\mathbf{B}_{\delta\xi} = \boldsymbol{\Sigma}^T \mathbf{P}^{1/2} \mathbf{P}^{1/2} \boldsymbol{\Sigma}$ and defining the preconditioned state vector $\delta\boldsymbol{\zeta} = \boldsymbol{\Sigma}^{-1} \mathbf{P}^{-1/2} \delta\boldsymbol{\xi}$. The relation between $\delta\boldsymbol{\zeta}$ and $\delta\mathbf{x}$ is now given by the conditioning transformation

$$\delta\boldsymbol{\zeta} = \mathbf{Z}\delta\mathbf{x} = \boldsymbol{\Sigma}^{-1} \mathbf{P}^{-1/2} \mathbf{H}^{-1} \mathbf{F} \delta\mathbf{x}, \text{ (9a)}$$

with inverse transformation

$$\delta\mathbf{x} = \mathbf{Z}^{-1}\delta\boldsymbol{\zeta} = \mathbf{F}^{-1} \mathbf{H} \mathbf{P}^{1/2} \boldsymbol{\Sigma} \delta\boldsymbol{\zeta}. \text{ (9b)}$$

This reduces the background error correlation matrix in terms of the conditioned state vector $\delta\boldsymbol{\zeta}$ to the identity matrix, and the background part of the cost function simply reads

$$J_b = (\delta\boldsymbol{\zeta})^{*T} (\delta\boldsymbol{\zeta}) = \left|\delta\boldsymbol{\zeta}\right|^2. \text{ (10)}$$

d. Minimalization and gradient

The minimalization is done with routine LBFGS, a limited-memory quasi Newton routine written by J. Nocedal (Liu and Nocedal, 1989). This routine proves to be fast and accurate. In 2DVAR good results are obtained with an initial step size of $30 J_t \left|\nabla J_t\right|^{-1}$, with J_t the total initial cost function and ∇J_t its gradient with respect to the control vector components. A typical batch requires less than 100 function evaluations to converge. See Vogelzang (2007) for detailed information. Note that 2DVAR uses the same minimization algorithm as Hoffman et al. (2003).

The minimalisation procedure starts with $\delta\boldsymbol{\xi} = 0$, so the initial analysis equals the background, and uses the gradient of the cost function with respect to the control variables. For the background part the gradient simply reads $\nabla J_b = 2\delta\boldsymbol{\xi}$. For the observation part the state vector is transformed to the spatial domain using the inverse conditioning transformation (9b). The gradient is evaluated in the spatial domain by differentiating (5) to δt and δl, packing the derivatives into a state vector, and transforming this state vector back to the spectral domain using the adjoint (complex conjugate of the transpose) of the inverse conditioning transformation, see, e.g., (Errico, 1997; Giering and Kaminski, 1998).

At this point it should be remarked that the control vector used in the actual minimalisation is not necessarily equal to the state vector (Hoffman et al., 2003). In the spatial domain the state vector $\delta\mathbf{x}$ is real with $2N_1N_2$ components and equals the control vector. In the spectral domain, the state vector $\delta\boldsymbol{\xi}$ is complex and has twice as much components. However, only components with nonnegative spatial frequency are independent (Press et al., 1988). The number of independent components remains $2N_1N_2$, but an additional packing/unpacking transformation is needed to go from state vectors to control vectors and vice versa in the spectral domain.

e. Error model

The stream function and the velocity potential are not observable quantities, but their error variances and error correlations can be derived from the wind field, either from theory or from measurements (or a combination of the two). The background error

correlations in the velocity potential and stream function in the spatial domain, $\rho_{\chi\chi}$ and $\rho_{\psi\psi}$, are modeled as Gaussian functions following Daley (1991),

$$\rho_{\psi\psi}(r) = (1-\nu^2)\sigma_\psi^2 L_\psi^2 e^{-r^2/R_\psi^2} \quad , \text{(11a)}$$

$$\rho_{\chi\chi}(r) = \nu^2\sigma_\chi^2 L_\chi^2 e^{-r^2/R_\chi^2} \quad , \text{(11b)}$$

where σ_ψ and σ_χ stand for the standard deviation of the error in ψ and χ, respectively, ν^2 for the ratio of the rotational and the divergent contribution to the wind field, and R_ψ and R_χ for the correlation lengths, the length scales that determine the extent of the error correlations. These parameters have a physical meaning and can not be varied arbitrarily. Moreover, they are not independent, since the impact on the analysis is determined by the ratio of the error standard deviation and the correlation length (de Vries and Stoffelen, 2000). The scaling parameters L_ψ and L_χ in (11) are defined as

$$L_\psi^2 = -\left.\frac{2f_\psi(r)}{\nabla^2 f_\psi(r)}\right|_{r=0}, \quad L_\chi^2 = -\left.\frac{2f_\chi(r)}{\nabla^2 f_\chi(r)}\right|_{r=0} \quad . \text{(12)}$$

Equation (12) holds for any form of the error correlation function. For the Gaussian form (11) one readily finds $L_\psi^2 = \frac{1}{2}R_\psi^2$ and $L_\chi^2 = \frac{1}{2}R_\chi^2$.

The background error correlation model is readily Fourier transformed, either numerically or analytically, to the spectral domain where it remains Gaussian. The default values for the parameters were found by de Vries and Stoffelen (2000) and are listed in table 1. These values were obtained after the intercomparison of 2DVAR with other variational methods mentioned before. The correlation length in the Tropics is higher than in the Extratropics to account for the general large scale convective

circulation structures around the equator (but with exception of tropical cyclones, see section 5e). In the Extratropics the circulation is more rotational as reflected in the smaller value for ν^2.

The a-priori probability P_k of ambiguity *k* being the correct solution in (6) follows from the inversion step as (Portabella and Stoffelen, 2004)

$$P_k = Ne^{-R_{MLE}^2/\overline{R}^2} \quad , \quad (13)$$

where R_{MLE} is the distance from the ambiguity to the scatterometer measurement in observation space and $\overline{R}$ an empirical constant. The normalization N guarantees that the sum over the a-priori probabilities of all ambiguities in a WVC equals one. The probability model (13) can be extended by including gross error probabilities (GEP's) with a constant probability distribution P_{GE} over a finite domain such that (Lorenc, 1986)

Deleted: $$$ referentie $$$

$$P_k = P_{GE} + N'e^{-R_{MLE}^2/\overline{R}^2} \quad , \quad (14)$$

with the normalization of the Gaussian part adapted to ensure that the sum over the a-priori probabilities of all ambiguities in a WVC remains one. The gross error probability imposes a minimum value to the a-priori probability which implies that from a certain threshold the magnitude of R_{MLE} no longer matters. The default value of P_{GE} is 0.0075.

f. Variational Quality Control

The results of 2DVAR can be used to assess the quality of the ambiguity removal. In the present 2DVAR formulation the Variational Quality Control (VQC) flag is set for each WVC where the contribution of J_o exceeds the threshold value of 12. This happens

for WVC's where the a-priori probability of each ambiguity equals the gross error probability.

3. Tests

In this section two tests are described that demonstrate the correctness of the current 2DVAR implementation. The first one is the single observation test. It involves only one observation and the minimization problem can be solved analytically. This test shows that all normalizations are correct. The second test is the edge analysis, showing that the analysis increments do not suffer from severe over- or underfitting.

a. Single Observation Test

In case there is exactly one observation, the 2DVAR problem can be solved analytically. Suppose that at some 2DVAR batch grid point (i, j) there is one observation (t_o, l_o). Starting with zero background increment and zero analysis increment field, the only contribution to the cost function and its gradient originates from this observation. From (6) this contribution reads

$$J_o = \frac{t_o^2 + l_o^2}{\sigma_o^2} \quad , \quad (15)$$

$$\frac{\partial J_o}{\partial t_{ij}} = \frac{2t_o}{\sigma_o^2} \quad , \quad \frac{\partial J_o}{\partial l_{ij}} = \frac{2l_o}{\sigma_o^2} \quad . \quad (16)$$

with $\sigma_o = \sigma_t = \sigma_l$. Now the 2DVAR problem reduces to an optimal interpolation problem (Daley, 1991) with solution

$$J_t^{\text{final}} = \frac{\sigma_b^2 \sigma_o^2}{(\sigma_b^2 + \sigma_o^2)^2} J_t^{\text{initial}} \quad , \text{(17)}$$

where $\sigma_b = \sigma_\chi = \sigma_\psi$ is the standard deviation of the background error. At the solution, the gradient of the total cost function should be zero, since the total cost function is minimal there, so

$$\nabla J_b = -\nabla J_o \quad . \text{(18)}$$

The background potential field can now be retrieved. The analysis wind at the observation point satisfies

$$(t,l) = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}(t_o, l_o) \quad . \text{(19)}$$

Note that the background error model parameters ν, R_χ, and R_ψ do not enter (19) explicitly. However, they do appear in the full expression for the analysis wind field (Vogelzang, 2007).

Figure 1 shows the resulting wind fields for (t_o, l_o) equal to (0,1) m s^{-1} and ν equal to zero (purely rotational) or one (purely divergent). The observation is located in the centre of the grid, x and y equal to 1600 km .The correlation lengths R_ψ and R_χ are both equal to 300 km. The standard deviation of both the observation error and the background error was set equal to 1.8 m s^{-1}. The wind speed at x and y equal to 1600 km should equal half of the observed speed. This is satisfied with an accuracy better than $2 \cdot 10^{-5}$. The circulation patterns in figure 1 form the basic building blocks from which 2DVAR constructs its analysis increments. The single observation test proved to be of crucial importance in fixing all normalizations and getting the definition of the control vector right.

b. Edge analysis

The 2DVAR batch grid should be large enough for the analysis increments to go to zero at the edges. Figure 2 shows the extreme (minimum and maximum) values of the analysis increments perpendicular and parallel to the satellite direction. Since the correlation length of the background error differs for the Tropics and the Extratropics, see table 1, the curves in figure 2 have been separated accordingly. The curves in figure 2 were obtained with SDP version 1.5 without applying the MSS at 25 km resolution with the ECMWF wind field as background, using all SeaWinds data from December 2004 (see section 4).

The batch grid has a width of 32 points. With a cell size of 100 km, the batch grid is 3200 km wide. The width of the free zone around the observations is defined as 5 cells, or 500 km. Since the SeaWinds swath is 1800 km wide, there are 4 cells remaining. These are inserted at the right hand side of the batch grid as an additional free zone. Moreover, SDP version 1.5 does not process the outer swath, so there is an extra strip of 200 km without observations at each side. Therefore the region with observations across the batch grid, marked with the vertical black dashed lines in figure 2, extends from cell 7 ($x = 600$ km) to cell 22 ($x = 2100$ km).

Figure 2 shows that the analysis increments in the Extratropics (dotted curves) are generally a bit higher than in the Tropics (dashed curves). It also shows that the analysis increments go to zero at the edges, faster in the Extratropics than in the Tropics. This is according to expectation, since the background error correlation length is 300 km in the Extratropics and 600 km in the Tropics. The free zone should be two or three times larger

than the background error correlation length, so it is large enough in the Extratropics, but a bit tight in the Tropics.

However, the most important conclusion from figure 2 is that there are no clear signs of over- or underfitting: the curves in figure 2 are not too smooth and neither too wildly varying. Moreover, they approach zero at the edge of the batch grid at a rate that agrees with the value of the background error correlation length. Application of the MSS and/or using the NCEP wind field as background yields similar results and the same conclusion. In the next section the edge analysis is revisited.

4. Statistical validation

a. Introduction

In this section 2DVAR is tested using some statistical methods using SeaWinds data from December 2004. The dataset contained all orbits that started in this period, orbits 28388 up to and including 28829. Some orbits turned out to be incomplete and were rejected for further processing, leaving a dataset of 434 full orbits. The NOAA BUFR files were processed with SDP version 1.5 using as background the NCEP model wind field (which is available in the NOAA BUFR product) or the ECMWF wind field. The NCEP model wind field is a 24 hour forecast of the 1000 mb wind. The ECMWF wind field is a 3 – 9 hour forecast of the wind speed at 10 m anemometer height, and is therefore expected to compare better with the scatterometer winds which are also at 10 m.

Processing was done with and without application of the MSS. To investigate how the analysis increments behave at the edges of the batch grid, the final 2DVAR

analysis increments perpendicular and parallel to the satellite direction, δt and δl, respectively, were dumped during processing.

As SeaWinds is a rotating fan-beam scatterometer, its observation geometry varies across the swath. At 25 km resolution there are 76 WVC's, and the swath is divided in three parts: the outer swath (WVC 1-10 and 67-76), the mid swath or "sweet" swath (WVC 11-30 and 47-66), and the nadir swath (WVC 31-46).

b. Wind field comparison

The datasets described above were intercompared by calculating the statistics of the differences in wind components for the zonal components, u, and the meridional components, υ. Tables 2a and 2b show the standard deviations of the differences in zonal wind component, σ_u, and meridional wind component, σ_υ, respectively. Model wind vectors were only considered when the associated scatterometer wind vectors were valid in order to prevent contamination of the NCEP and ECMWF model wind vector comparison by land pixels.

The largest differences occur between the selected wind without MSS and the NCEP model ($\sigma_u \approx 2.5$ m/s; $\sigma_\upsilon \approx 2.2$ m/s). It makes little difference whether the selected winds were obtained using the NCEP winds or the ECMWF winds as background. Note that the differences between the selected winds and the ECMWF model ($\sigma_u \approx 1.9$ m/s; $\sigma_\upsilon \approx 1.8$ m/s) are smaller than those between the selected winds and the NCEP model ($\sigma_u \approx 2.5$ m/s; $\sigma_\upsilon \approx 2.4$ m/s). This means that the scatterometer winds are closer to the ECMWF model than to the NCEP model. As the ECMWF First Guess wind is better than

the NCEP + 24 hour 1000 mb wind, it is concluded that the scatterometer observations yield useful information.

This is corroborated by the results for the MSS winds. Now the 2DVAR method has more freedom in selecting the optimal wind vector, and the influence of the background field becomes slightly more important. The difference between the MSS wind field and the NCEP model is rather large when using the ECMWF winds as background ($\sigma_u \approx 2.2$ m/s; $\sigma_\upsilon \approx 2.1$ m/s) and somewhat smaller when using the NCEP wind as background ($\sigma_u \approx 2.1$ m/s; $\sigma_\upsilon \approx 2.0$ m/s). Compared to the ECMWF winds, the differences are smaller: $\sigma_u \approx \sigma_\upsilon \approx 1.5$ m/s with NCEP background; $\sigma_u \approx 1.5$ m/s and $\sigma_\upsilon \approx 1.4$ m/s with ECMWF background.

Note that the difference between the selected winds with NCEP background and ECMWF background is small, but slightly larger for MSS ($\sigma_u \approx \sigma_\upsilon \approx 0.6$ m/s). This is consistent with the notion that without MSS the ambiguity removal method has limited choice between the (four at most) solutions found by the inversion algorithm. The details of the background field appear rather unimportant.

Using the NCEP field as background, application of MSS reduces the standard deviation of the error with respect to the ECMWF model with $\sqrt{(1.79)^2 - (1.51)^2} = 0.96$ m/s for u and with $\sqrt{(1.75)^2 - (1.50)^2} = 0.90$ m/s for υ. This once more shows the added value of scatterometer information derived with MSS. Note that the selected winds (no MSS) based on the NCEP wind field actually add variance to the ECMWF-NCEP difference, thereby degrading the wind field.

The results in tables 2a-b were obtained for those wind vectors for which the VQC flag (see section 2f) was not set. The number of valid vectors may therefore differ

slightly for each of the sets, between 18.8 million and 19.2 million, a variation of 2% that may influence the statistics. However, inspection of all histograms of the wind differences revealed that their distributions are well behaved without significant outliers.

c. Noise estimation

Figure 3 shows the autocorrelation in the zonal wind component *u* at 25 km resolution obtained from the ECMWF field and from the SDP results with and without MSS. The left hand panel shows the full curves; the right-hand panel an enlargement at short distances. The autocorrelation at zero distance equals 1 by definition. The SDP result without MSS (black curve) shows a clear discontinuity at short distances, while the SDP result with MSS (blue curve) and the ECMWF result (red curve) approach 1 continuously. This discontinuity is caused by an uncorrelated noise component adding only variance. The size of the discontinuity can be estimated by extrapolating the curve to zero distance (dashed black curve). The extrapolated curve crosses the y-axis at $1-a$ rather than 1. It is easy to show that the standard deviation of the noise, σ_n, satisfies (Vogelzang, 2006a)

$$\sigma_n = \sigma_s \sqrt{a}\,. \quad (20)$$

where σ_s is the standard deviation of the total signal.

Figure 4 shows the standard deviation of the noise for the SDP wind components at 25 km and 50 km resolution processed without MSS. At coarser resolutions the noise level reduces and the extrapolation distance increases, leading to larger uncertainties in the noise estimate. The extrapolation may even overshoot the autocorrelation, leading to an extrapolated autocorrelation larger than 1 at *x=0* and, hence, a negative noise variance

estimate. This happens at 50 km resolution in the mid swath and at 100 km resolution all over the swath. Such points have been excluded from figure 4.

Figure 4 shows that the noise level decreases as the resolution becomes coarser. At 100 km resolution the noise estimates are invalid, indicating negligible noise contribution. At 25 km resolution the standard deviation of the noise may exceed 1 m/s for υ and 1.5 m/s for *u*. These figures agree well with the overall reduction in standard deviation of the difference between the scatterometer winds and the ECMWF background when switching on MSS as presented in the previous section. When MSS is applied the noise component disappears and no valid noise estimates are obtained.

d. Analysis statistics

Some statistics of the analysis increments were already presented in section 3.2 to show that 2DVAR exhibits no signs of severe overfitting and that the free edge is sufficiently large in the Extratropics but rather tight in the Tropics. Figure 5 shows the results for the average absolute value of the analysis increment across the 2DVAR batch grid. This figure shows that in all cases the average absolute analysis increment is smaller when applying MSS (dashed curves). This is no surprise, since 2DVAR is expected to find a solution with reasonable probability not too far from the background in this case. Without MSS (dotted curve) the ambiguities are further away from each other, leading to larger analysis increments. A second reason lies in the more noisy character of the scatterometer winds without MSS.

Figure 5 also shows that the average absolute analysis increments in the Tropics (left panels) do not approach zero properly, again indicating that the free edge is rather

tight. In the Extratropics (right panels) the free edge is sufficiently large. Note also that the average absolute analysis increments are higher with the NCEP wind field as background (lower panels) than those with the ECMWF wind field as background (upper panels). When using the NCEP background in the Extratropics, δt shows some mild signs of overfitting around $x = 400$ km and $x = 2400$ km.

5. Case studies

In the previous sections it was shown that 2DVAR performs well in a statistical sense. In combination with MSS it suppresses the noise component in SeaWinds data at high resolution. In this section three cases will be studied in more detail in order to support the conclusions drawn from the statistical analyses in the previous section and to gain more insight in the role of the parameters in the error model.

The first case is a SeaWinds observation of a cyclone with a strong front in the southern Pacific on August 6, 2006. There is a clear mismatch in the position of the cyclone and the front between the observations and the NCEP background. The second case is a SeaWinds observation of a severe hurricane over the northern Pacific on December 30, 2004 with wind speeds over 40 m/s. The third case is an ASCAT observation of a tropical cyclone in the Indian Ocean which is not present in the ECMWF background. The effect of changing 2DVAR settings will be illustrated.

a. Case Pacific cyclone and front

Figure 6 shows the NOAA result for SeaWinds measurements recorded on August 6, 2006 in the Pacific Ocean off the coast of Chile. A deep low pressure area

located approximately at 80° W and 45° S is accompanied with an extended frontal area on its northern and northeastern side. The front has an irregular shape around 75° W in figure 6 where a large number of cells have their rain flag set (orange arrows). To the north of the front a few erroneous wind vectors can be seen. This shape is not present in the NCEP model field shown in figure 7. Moreover, the NCEP model locates the front more to the south (the grid point 80° W, 30° S is a suitable reference), and the centre of the cyclone more to the west.

The SDP wind field without MSS is shown in figure 8. The wind field is noisy and the eastern part of the front is not very clearly visible because many points there are flagged as rain points and therefore rejected for further processing by SDP. The VQC flag (see section 2f) is set in a number of WVC's along the front and near the centre of the cyclone (purple arrows). The location of the cyclone agrees with the NOAA result, figure 6, while the location of the front agrees with the NCEP background, figure 7. Note the strong convergence in the region 30° S-35° S, 75° W-80° W. This structure seems not realistic.

Figure 9 shows the result when MSS is applied. The wind field is now smooth, also north of the front, because the noise has been filtered out. The front appears smoother and extends more to the east. No WVC's are flagged in the frontal zone. Southwest of the front line some wavy structures appear in the wind field. The convergence in the region 30° S-35° S, 75° W-80° W has disappeared, and the centre of the cyclone has moved slightly to the west, indicating larger influence from the background.

The influence of the background increases with decreasing background error standard deviation. Figure 10 is obtained with a background error standard deviation of 1 m/s (left) and 3 m/s (right), while the standard value is 2 m/s, see table 1. In the left hand panel of figure 10, a few WVC's along the front now have their VQC flag set, and the centre of the low moved slightly further to the west, more in agreement-with the background field in figure 7. In the right hand panel the observations have more weight. The front is more clearly defined and the centre of the low lies more to the east.

Deleted: a

Deleted: (

Deleted: A

Deleted: Figure 10b is obtained with a background error standard deviation

Deleted: of 3 m/s. Now

Deleted: than in figure 10a,

The influence of the background increases with increasing background error correlation length. Figure 11 is obtained with a background correlation length of 250 km (left) and 350 km (right), whereas the standard value for the Extratropics is 300 km, see table 1). In the left hand panel of figure 11 the influence of the background is smaller than in the right hand panel; the front is more clearly defined and the centre of the low is more to the east. Note the similarity between the left hand panel of figure 10 and the right hand panel of figure 11, and between the right hand panel of figure 10 and the left hand panel of figure 11.

Deleted: s

Deleted: a and 11b are

Deleted: , respectively (

Deleted: a

Deleted: figure 11b

Deleted: s

Deleted: a

Deleted: b

Deleted: s

Deleted: b

Deleted: a

b Case Pacific hurricane

On December 30, 2004, a strong hurricane raging over the northern Pacific was observed by SeaWinds. Figure 12 shows the SDP result with ECMWF background and MSS applied to reduce the noise. WVC's flagged as contaminated by rain have been left out of figure 12, while VQC flagged wind vectors are drawn in purple. Observations and background give the same position for the centre of the hurricane, but the observed wind speeds around the centre are higher than the modelled ones, especially to the south of the

centre where the VQC flags are set in figure 12. This is no surprise: it is well known that NWP models tend to underestimate the strongest winds in hurricanes.

Figures 13 shows the centre of the hurricane. Now the VQC flagged wind vectors are in black, while the other arrow colours indicate the wind speed range. The left hand panel of figure 13 shows the same results as figure 12, but the right hand panel is obtained without using the Gross Error Probabilities (GEP). With GEP (standard value 0.0075 for all WVC's) the wind directions south of the centre are obviously wrong. Without GEP the black arrows fit well in the overall circulation pattern. The wind speeds south of the centre of the hurricane exceed 40 ms^{-1}.

The GEP's effectively impose a maximum on the observation part of the cost function. In cases where ambiguities and background are far apart, the a-priori probabilities becomes constant and 2DVAR can only distinguish the ambiguities by their distance to the background. As a result, 2DVAR starts behaving like closest-to-background and will select the ambiguity with its direction closest to the background (in MSS all ambiguities in a WVC have similar speeds). This is what happens in the left hand panel of figure 13. When the GEP's are turned off, the a-priori probabilities influence 2DVAR's selection process, resulting in selections that may deviate more from the background as in the right hand panel of figure 13. 2DVAR now selects ambiguities with a high a-priori probability which fit better into the general circulation pattern and have a somewhat higher speed. Note that the VQC flag is set more often, because of the large difference with the background.

Deleted: a-b

Deleted: has b

Deleted: een neglected and the

Deleted: Figure 13a

Deleted: figure 13b

Deleted: them, the circulation pattern shows up much better

Deleted: Moreover, 2DVAR selects higher wind speeds over 40 m/s

Formatted: Superscript

Deleted: a

Deleted: b

Deleted: However,

Deleted: for these selections is still set

This case shows that the VQC flag does not necessarily imply that the selection is wrong. It may also indicate errors in the background. Anyhow, such cases should be handled with care.

c. Case Indian cyclone

The last case considered here is an ASCAT observation of a tropical cyclone on August 31, 2007, east of India. Standard processing (no MSS, 2DVAR settings from table 1, using the ECMWF wind as background and 25 km resolution) clearly shows the cyclone, see the upper left panel of figure 14. Some WVC's near the centre of the cyclone have their MLE flags set due to rainfall and are shown in orange. The cyclone is on the southern hemisphere, so the circulation direction is clockwise. Closer inspection of the upper left panel of figure 14 reveals that the wind direction is off by about 180° in an area north to northeast of the centre. This suggest a problem in 2DVAR.

The background wind field is shown in the upper right panel of figure 14. The reason for the failure of 2DVAR is obvious: the cyclone is not present in the background. Instead, the background shows a frontal structure. This is rather surprising, since ASCAT data were already assimilated in the ECMWF model in August 2007. The most likely explanation is that earlier ASCAT observations missed the cyclone because spatial coverage near the equator is about 75% per day, considering both ascending and descending orbits. Note that the ECMWF background is the most recent prediction before the scatterometer data of figure 14 were assimilated. Predictions after inclusion of the data of figure 14 do show the cyclone, though at a slightly different position (Hersbach, 2007, private communication).

Application of MSS (lower left panel of figure 14) does not improve the wind field. Because 2DVAR has more solutions to choose from, it is likely that a solution with lower a-priori probability close to the background is preferred over one with slightly higher a-priori probability further away. As a result, the influence of the background increases. This leads to a compromise between the observations and the background in which the cyclone is visible with the right circulation direction everywhere, but too weak and at a wrong position.

The right circulation direction can be obtained by increasing the influence of the observations relative to that of the background. This can be achieved by increasing the error variance of the background or decreasing that of the observations. However, unrealistically large or small values have to be adopted in order to remove the directional discrepancy. Another possibility is to decrease the background error correlation length. This seems more plausible, since the standard 2DVAR value for the background error correlation length in the Tropics is 600 km. Such a large value performs well in fitting the large scale circulation patterns generally found in the Tropics, but fails for small rotational structures. This is shown in the lower right panel of figure 14, which was obtained without MSS and using a reduced background error correlation length of 300 km, the value it attains in the Extratropics. This removes the error in wind direction.

d. Resume

The relative balance between observations and background in 2DVAR is controlled by the error model. In the present implementation it contains the parameters listed in table 3. The effect of changing the parameter value is also shown. It should be

kept in mind that the parameters have a physical meaning, so their values can not be changed arbitrarily, and that they are not independent. This is shown in particular by the case studies of the ASCAT observation of a tropical cyclone, where good results were obtained by reducing the background error correlation length to a value consistent with the size of the observed structure. The case of the extratropical hurricane observed with SeaWinds on December 30, 2004, shows that the gross error probability is better switched off when MSS is applied. This particular case also raises some questions regarding the interpretation of the variational quality control (VQC) flag. In the present version of 2DVAR it is set when the observation cost in a WVC exceeds a threshold value of 12. This does not necessarily mean that the selected solution is incorrect – it may very well be that the control state is incorrect because of mislocated structures in the background or that the winds are very high locally (very small structures).

The MSS offers 2DVAR more ambiguities to choose from. In case of broad minima in the distance between the observation and the GMF, as is the case for SeaWinds in the nadir part of the swath, 2DVAR is able to select a solution with reasonable a-priori probability that is spatially consistent with the neighbouring WVCs. Also, the influence of the background on the final selection increases, as is clearly shown in the case studies.

6. Conclusions

In this paper a new method for ambiguity removal named 2DVAR is presented. It is a generic method applicable to SeaWinds and ASCAT (or ERS) data. 2DVAR constructs an incremental analysis from the background and the observations, taking the

a-priori probabilities of the latter into account. The minimalization problem is fully conditioned and solved in the spectral representation of stream function and wind potential. The selected ambiguous solution is the one closest to the analysis. The present implementation satisfies the single observation test, a nontrivial case with analytical solution, and shows no clear signs of under- or oversampling.

2DVAR in combination with MSS proves effective in removing the noise in SeaWinds data at 25 km resolution. Especially in the nadir part of the SeaWinds swath MSS allows 2DVAR to choose from more solutions with comparable a-priori probability. As a result, MSS here increases the influence of neighbouring WVCs and the background. The latter is not always a desirable property, notably when the background is in error or when little noise is present. ASCAT data at 25 km resolution contains little noise due to the more favourable observation geometry and low rain sensitivity and is therefore being processed without MSS. The noise level can be estimated using a simple and robust method based on extrapolation of the autocorrelation to zero distance. For Seawinds at 25 km resolution the noise standard deviation exceeds 1.5 m/s in the zonal wind component u and 1.0 m/s in the meridional component v.

The influence of the background can be controlled by the parameters of the error model in 2DVAR. It is decreased by increasing the background error variance, decreasing the observation error variance, increasing the background error correlation length or decreasing the gross error probabilities. These parameters are not independent and have a physical meaning, so they can not be varied arbitrarily. This is shown by the ASCAT observation of a tropical cyclone which is well reproduced when the background error correlation length is adapted to the size of the cyclone. Another example is the SeaWinds

observation of an extratropical hurricane where winds exceeding 40 m/s are reproduced well when limiting the background influence by switching off the gross error probabilities off. Further experimentation with 2DVAR to optimise these settings is recommended. Moreover, multiple-loop iteration may be required to fit small-scale tropical cyclones in a rather large-scale background.

The 2DVAR software is being developed and freely distributed in the context of the NWP Satellite Application Facility (SAF). 2DVAR provides a simplified framework to test improvements to a more complete 3D- or 4D-Var data assimilation of ambiguous scatterometer data. 2DVAR may further be used to process winds from the forthcoming Indian and Chinese scatterometers, for instance to aid in marine and coastal warnings.

The current 2DVAR implementation is rather rigid with respect to the size and dimension of the grid on which the analysis increments are calculated. This can be improved by implementation of a mixed-radix Fourier transform.

Acknowledgments.

The authors wish to thank their collegues from KNMI's scatterometry group, Jos de Kloe, Marcos Portabella, Anton Verhoef, and Jeroen Verspeek, for their interest in this work, stimulating discussions, and helpful advice in software issues. This work has been funded by EUMETSAT in the context of the NWP SAF part of the Satellite Application Facility Network.

References

Cavanié, A. and P. Lecomte, 1987: *Vol. 1 – Study of a method to dealias winds from ERS-1 data. Vol 2 – Wind retrieval and dealiasing subroutines.* ESA contract 6874/87/CP-I(sc) report, ESA publications division, Noordwijk, The Netherlands.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, UK.

Errico, R.M., 1997: What is an adjoint model? *Bull. Am. Met. Soc.* 2577-2591.

Giering, R. and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Trans. Math. Software.* **24**, 437-474.

Graham, R., D. Anderson, A. Hollingsworth, and H. Böttger, 1989: *Evaluation of ERS-1 wind extraction and ambiguity removal algorithms: meteorological and statistical evaluation.* ECMWF Report, ECMWF, Reading, UK.

Henderson, J.M., R.N. Hoffman, S.M. Leidner, R. Atlas, E. Brin, and J.V. Ardizzone, 2003: A comparison of a two-dimensional varational analysis method and a median filter for NSCAT ambiguity removal. *J. Geophys. Res.,* **C108**, 3176-3188.

Hersbach, H., 2007: The preparation of the assimilation of ASCAT scatterometer data at ECMWF. 2007 Joint EUMETSAT/AMS conference, Sept 24-28, 2007, Amsterdam, Netherlands.

Hoffmann, R.N., S.M. Leidner, J.M. Henderson, R. Atlas, J.V. Ardizzone, and S.C. Bloom, 2003: A two-dimensional varational analysis method for NSCAT ambiguity removal: methodology, sensitivity, and tuning. *J. Atm. Ocean. Tech.*, **20**, 585-605.

KNMI scat web site, www.knmi.nl/scatterometer/

Leru, M., 1999: *Inversion des measures radars diffusiomètriques d'ERS-1 et ERS-2: Etude d'une nouvelle approche basée sur une méthode variationelle.* IFREMER technical report 99-05.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Q. J. Royal Meteorol. Soc.*, **112**, 1177-1194.

NWP SAF web site, www.metoffice.gov.uk/research/interproj/nwpsaf/

Offiler, D., 1987: *ERS-1 wind retrieval algorithms.* UKMO. 19 Branch Memorandum No 86, UKMO, Bracknell, UK.

Portabella, M., 2002: *Wind field retrieval from satellite radar systems.* Thesis, University of Barcelona, Spain.

Portabella, M., and A. Stoffelen, 2004 : A probabilistic approach for SeaWinds data assimilation. *Q. J. Royal. Meteorol. Soc.*, **130**, 1-26.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, 1988: *Numerical Recipes.* Cambridge University Press, Cambridge, UK.

Quilfen, Y. and A. Cavanié, 1991: A high precision wind algorithm for the ERS-1 scatterometer and its validation. *Proc. IGARSS 1991,* Espoo, Finland, 873-876.

Roquet, H. and A. Ratier, 1988: Towards direct variational assimilation of scatterometer backscatter measurements into numerical weather prediction models. *Proc. IGARSS 1988,* Edinburgh, Scotland, 257-260.

Sienkiewics, J.M., J.M. Von Ahn, G.M. McFadden, and M. Stewart, 2007: Hurricane force extratropical cyclones as detected by QuikSCAT. 2007 Joint EUMETSAT/AMS conference, Sept 24-28, 2007, Amsterdam, Netherlands.

Stiles, B.W., B.D. Pollard, and R.S. Dunbar, 2002: Direction interval retrieval with thresholded nudging: A method for improving the accuracy of QuikSCAT winds. *IEEE Trans. Geosci.Remote Sens.* **40**, 79-89.

Stoffelen, A., 1998: *Scatterometry*. Thesis, University of Utrecht, The Netherlands.

Stoffelen, A. and D. Anderson, 1995: *The ECMWF contribution to the characterization, interpretation, calibration and validation of ERS-1 scatterometer backscatter measurements and their use in numerical weather prediction models.* ESA contract 9097/90/NL/BI report, ECMWF, Reading, UK.

Stoffelen, A. and D. Anderson, 1997: Ambiguity removal and assimilation of scatterometer data. *Q. J. R.. Meteorol. Soc.,* **123**, 491-518.

Stoffelen, A., S. de Haan, Y. Quilfen, and H. Schyberg, 2000: *ERS scatterometer Ambiguity removal scheme comparison.* Technical Report OSISAF, EUMETSAT, Darmstadt, Germany. (obtainable from www.knmi.nl/publications/fulltexts/safosi_w_acomparison.pdf)

Vogelzang, J., 2006a: *On the quality of high resolution wind fields.* Technical Report NWPSAF_KN_TR_002, EUMETSAT, Darmstadt, Germany. (obtainable from www.metoffice.gov.uk/research/interproj/nwpsaf/)

Vogelzang, J., 2006b: *The orientation of SeaWinds wind vector cells.* Technical Report NWPSAF_KN_TR_003, EUMETSAT, Darmstadt, Germany. (obtainable from www.metoffice.gov.uk/research/interproj/nwpsaf/)

Vogelzang. J., 2007: *Two dimensional variational ambiguity removal.* Technical Report NWPSAF_KN_TR_004, EUMETSAT, Darmstadt, Germany. (obtainable from www.metoffice.gov.uk/research/interproj/nwpsaf/)

de Vries, J.C.W. and A.C.M. Stoffelen, 2000: 2D variational ambiguity removal. *Technical Report 226*, Royal Netherlands Meteorological Institute (KNMI), de Bilt, the Netherlands, 66 pp.

List of Figures

FIG. 1. Results of the single observation test for ν=0 (left) and ν=1 (right).

FIG. 2. Extreme values of the analysis increments as a function of the distance across the batch grid. Batch grid cells containing observations are located between the black vertical dashed lines.

FIG. 3. Autocorrelation of the zonal wind component u from the ECMWF model and from SDP with and without MSS.

FIG. 4. Standard deviation of the noise in the zonal and meridional wind components u and v obtained by SDP at 25 km and 50 km resolution.

FIG. 5. Average value of the analysis increment components perpendicular and parallel to the satellite direction as a function of distance across the 2VAR batch grid. Upper panels: ECMWF background; lower panels: NCEP background; left panels: Tropics; right panels: Extratropics.

FIG. 6. Original NOAA wind field.

FIG. 7. NCEP model field.

FIG. 8. SDP without MSS.

FIG. 9. SDP with MSS.

FIG. 10. SDP with MSS. Left: background error standard deviation of 1 m/s; right: background error standard deviation of 3 m/s.

FIG. 11. SDP with MSS. Left: 250 km background error correlation length; right: 350 km background error correlation length.

FIG 12. Hurricane wind field in the northern Pacific obtained with SDP with MSS.

FIG. 13. Centre of the hurricane on December 30, 2004. The arrow colors indicate the wind speed range (yellow: 0-10 m/s, green: 10-20 m/s; cyan: 20-30 m/s; blue: 30-40 m/s) and the setting of the variational quality control flag (black). Left: with Gross Error Probabilities; Right: without.

FIG. 14. Tropical cyclone east of India on August 31, 2007 observed with ASCAT at 25 km resolution. Upper left: standard processing result without MSS; upper right: ECMWF background; lower left: result with MSS; lower right: result without MSS and with 300 km background error correlation length.

Deleted: FIG. 10a.SDP with MSS and a background error standard deviation of 1 m/s.¶
¶
FIG. 10b. SDP with MSS and a background error standard deviation of 3 m/s.¶

Deleted: a

Deleted: .¶
¶
FIG. 11b. SDP with

Deleted: FIG. 13a. Centre of the hurricane on December 30, 2004. The arrow colors indicate the wind speed range (yellow: 0-10 m/s, green: 10-20 m/s; cyan: 20-30 m/s; blue: 30-40 m/s; violet: more than 40 m/s).¶
¶
FIG 13b. As figure 13a, but without Gross Error Probabilities.¶

Deleted: FIG. 14. Tropical cyclone east of India on August 31, 2007 observed with ASCAT at 25 km resolution. Upper left: standard processing result without MSS; upper right: ECMWF background; lower left: standard processing with MSS; lower right: processing with 300 km background error correlation length.¶

TABLE 1. Default 2DVAR parameters for the background error correlation model

Zone	Latitude	$R_\psi = R_\chi$	$\sigma_\psi = \sigma_\chi$	ν^2
Northern Hemisphere	> +20°	300 km	2 m^2/s	0.2
Tropics	(-20°,+20°)	600 km	2 m^2/s	0.6
Southern Hemisphere	<-20°	300 km	2 m^2/s	0.2

TABLE 2a. Standard deviation of the differences in the zonal wind component. Rows and columns labeled "Model" pertain to model wind, "No MSS" to scatterometer wind without MSS, and "MSS" to scatterometer winds with MSS.

σ_u (m/s)		NCEP			ECMWF		
		Model	No MSS	MSS	Model	No MSS	MSS
NCEP	Model	--	2.53	2.14	1.79	2.54	2.22
	No MSS	2.53	--	1.21	1.93	0.59	1.27
	MSS	2.14	1.21	--	1.52	1.29	0.60
ECMWF	Model	1.79	1.93	1.52	--	1.94	1.48
	No MSS	2.54	0.59	1.29	1.94	--	1.22
	MSS	2.22	1.27	0.60	1.48	1.22	--

TABLE 2b. As table 2a, but for the differences in the meridional wind component.

σ_v (m/s)		NCEP			ECMWF		
		Model	No MSS	MSS	Model	No MSS	MSS
NCEP	Model	--	2.22	1.95	1.75	2.22	2.07
	Sel	2.22	--	1.06	1.76	0.43	1.09
	MSS	1.95	1.06	--	1.50	1.08	0.55
ECMWF	Model	1.75	1.76	1.50	--	1.75	1.41
	No MSS	2.22	0.43	1.08	1.75	--	1.06
	MSS	2.07	1.09	0.55	1.41	1.06	--

TABLE 3. 2DVAR error model parameters and their effect after decreasing its value.

Parameter name	Symbol	Effect of decreasing the parameter value
Observation error standard deviation	$\varepsilon_t, \varepsilon_l$	Increases influence of observations relative to background
Background error standard deviation	$\varepsilon_\chi, \varepsilon_\psi$	Increases influence of background relative to observations
Background error correlation length	R_χ, R_ψ	Increases influence of observations relative to background
Gross Error Probabilities	P_{GE}	Increases influence of observations relative to background

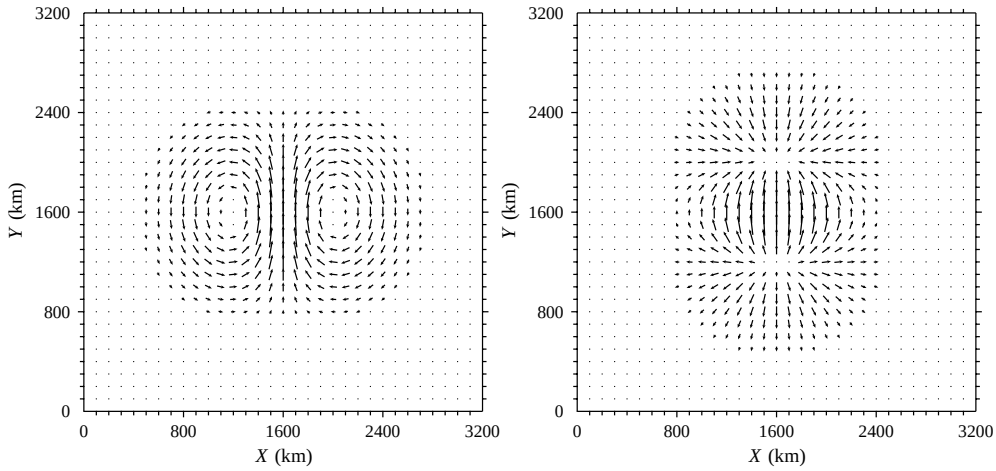
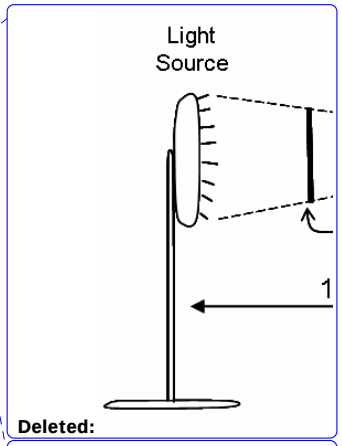


FIG. 1. Enter the caption for your figure here. Repeat as necessary for each of your figures.



Deleted:

Deleted: ¶

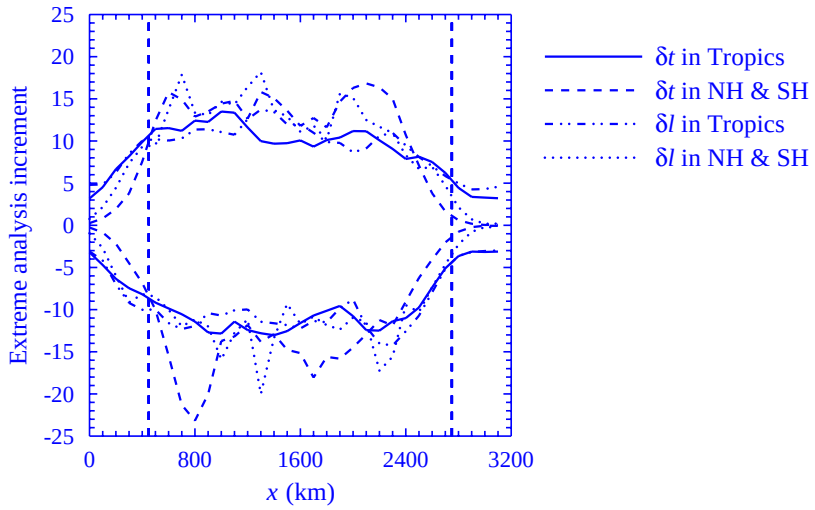


FIG. 2. Extreme values of the analysis increments as a function of the distance across the batch grid. Batch grid cells containing observations are located between the black vertical dashed lines.

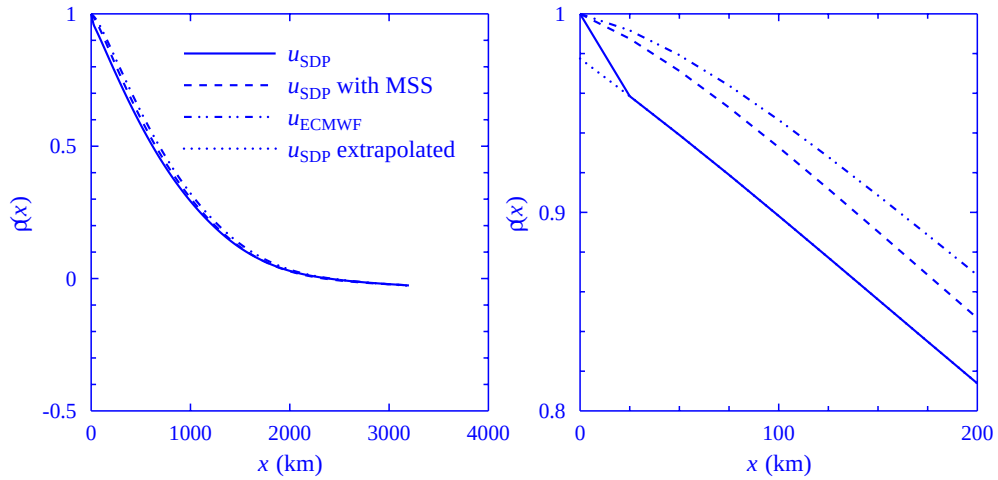


FIG. 3. Autocorrelation of the zonal wind component u from the ECMWF model and from SDP with and without MSS.

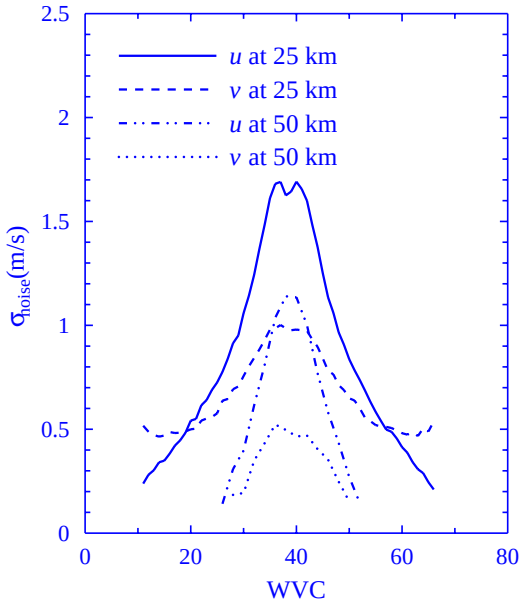


FIG. 4. Standard deviation of the noise in the zonal and meridional wind components u and v obtained by SDP at 25 km and 50 km resolution.

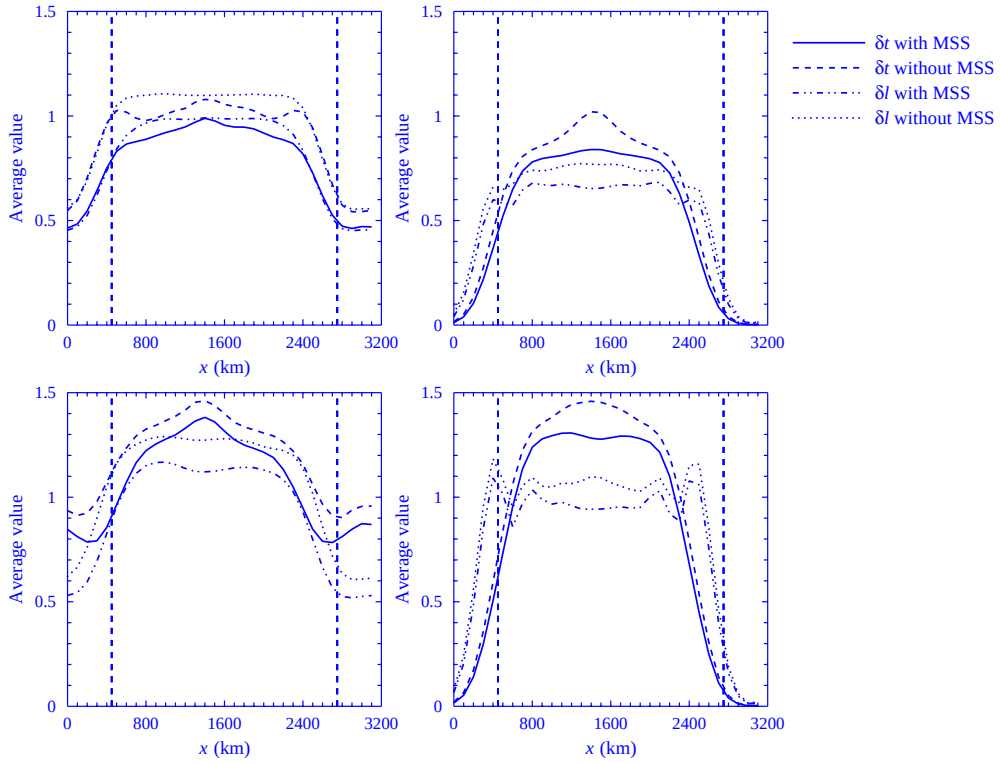


FIG. 5. Average value of the analysis increment components perpendicular and parallel to the satellite direction as a function of distance across the 2VAR batch grid. Upper panels: ECMWF background; lower panels: NCEP background; left panels: Tropics; right panels: Extratropics.

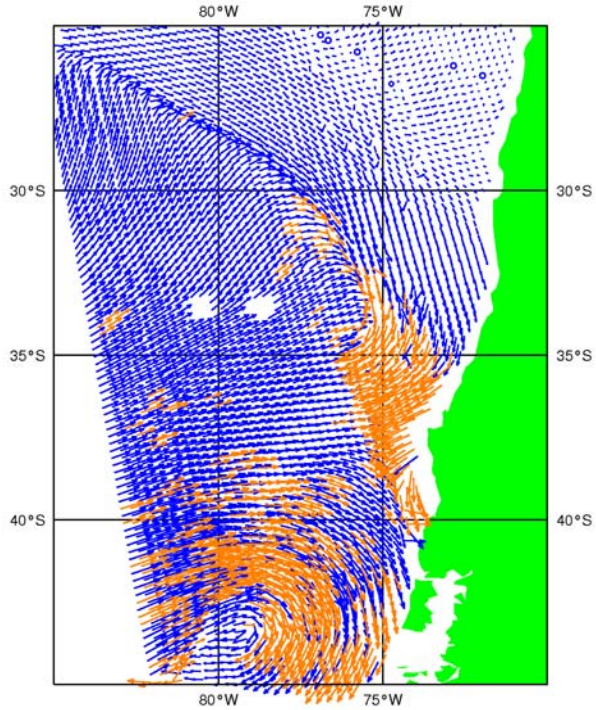


Fig. 3. Original NCEP1 wind field

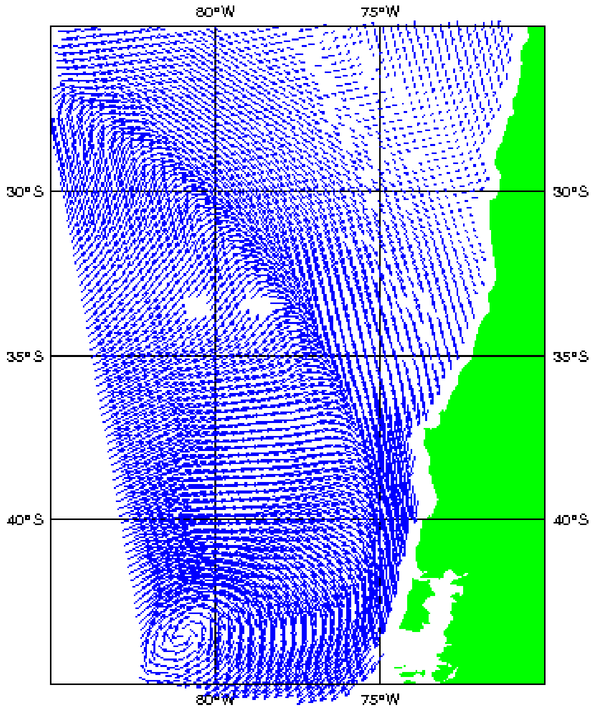


Fig. 7: NCEP model field.

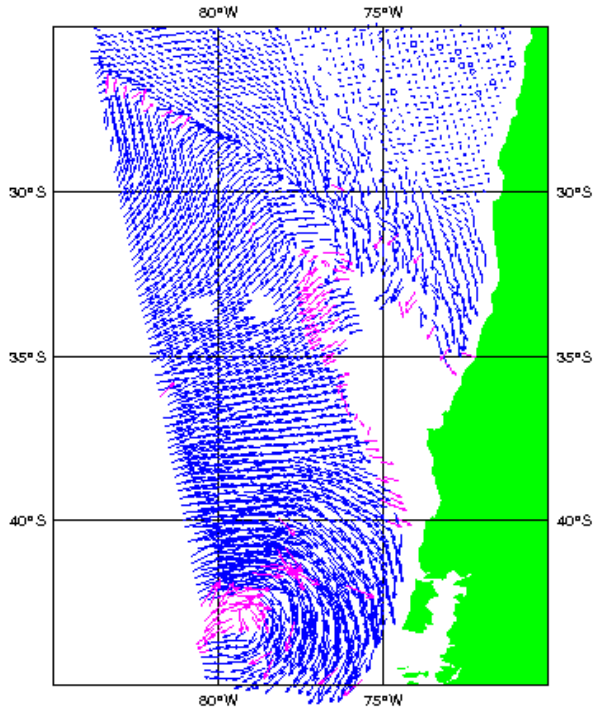
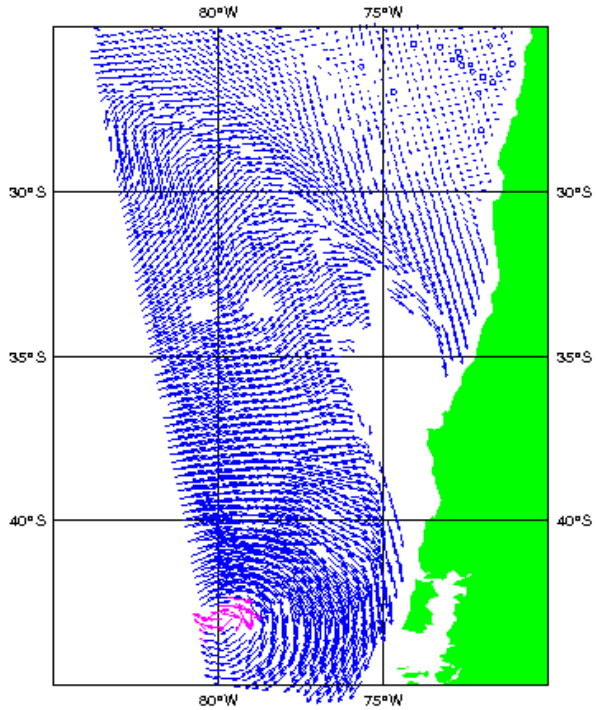


Fig. 3. SST without AVSA.


80°W
75°W
30° S
35° S
40° S

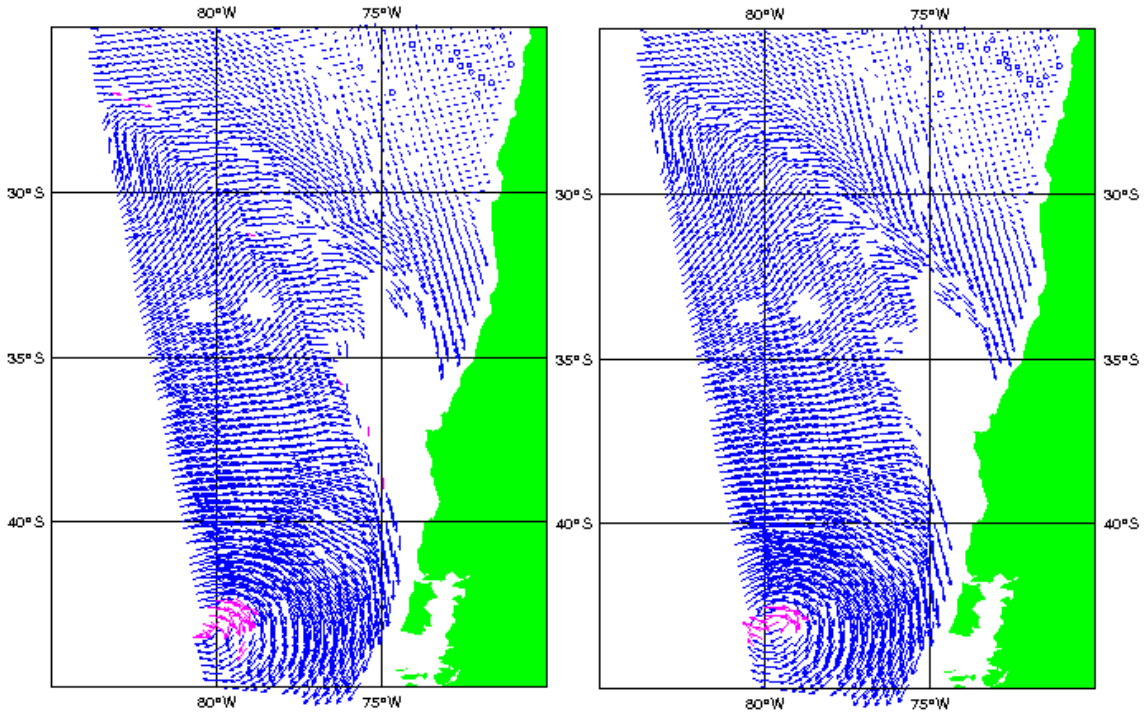


FIG. 10. SDP with MSS. Left: background error standard deviation of 1 m/s; right: background error standard deviation of 3 m/s.

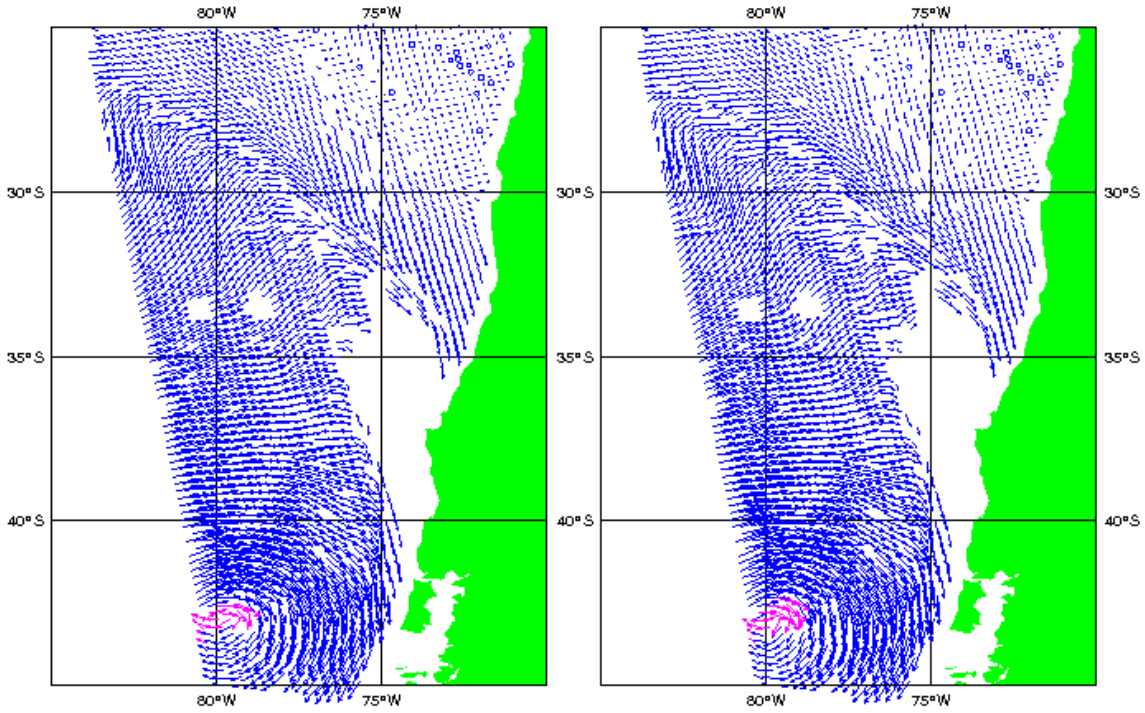


FIG. 11. SDP with MSS. Left: 250 km background error correlation length; right: 350 km background error correlation length.

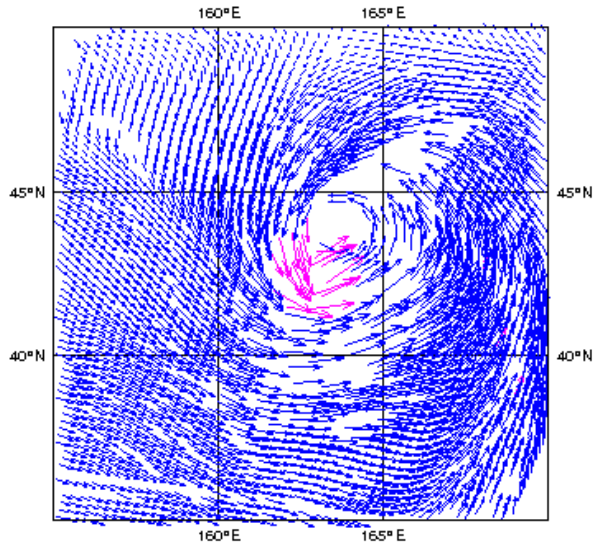


FIG 12. Hurricane wind field in the northern Pacific obtained with SDP with MSS.

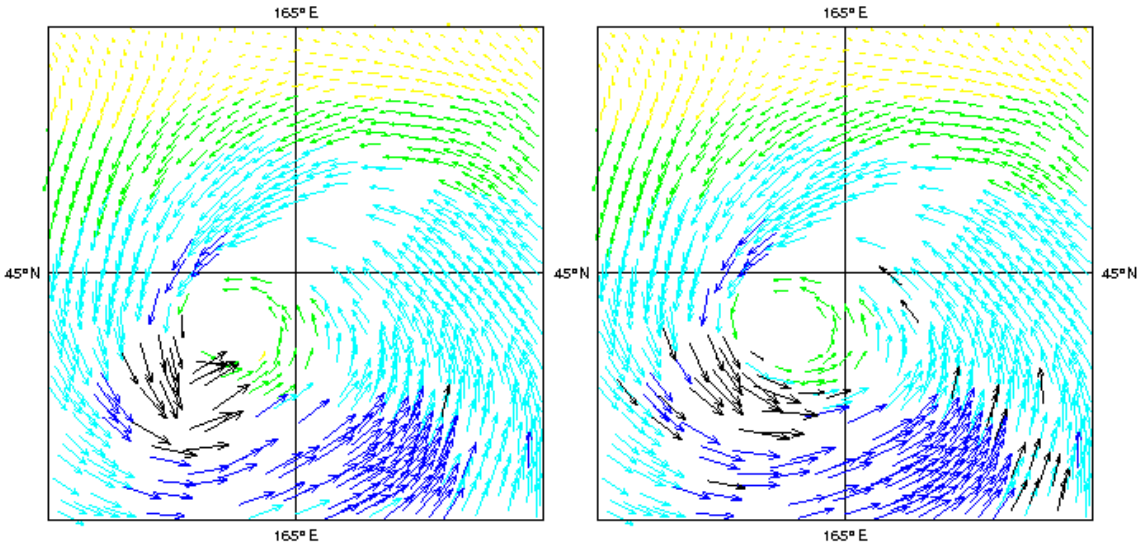


FIG. 13. Centre of the hurricane on December 30, 2004. The arrow colors indicate the wind speed range (yellow: 0-10 m/s, green: 10-20 m/s; cyan: 20-30 m/s; blue: 30-40 m/s) and the setting of the variational quality control flag (black). Left: with Gross Error Probabilities; right: without.

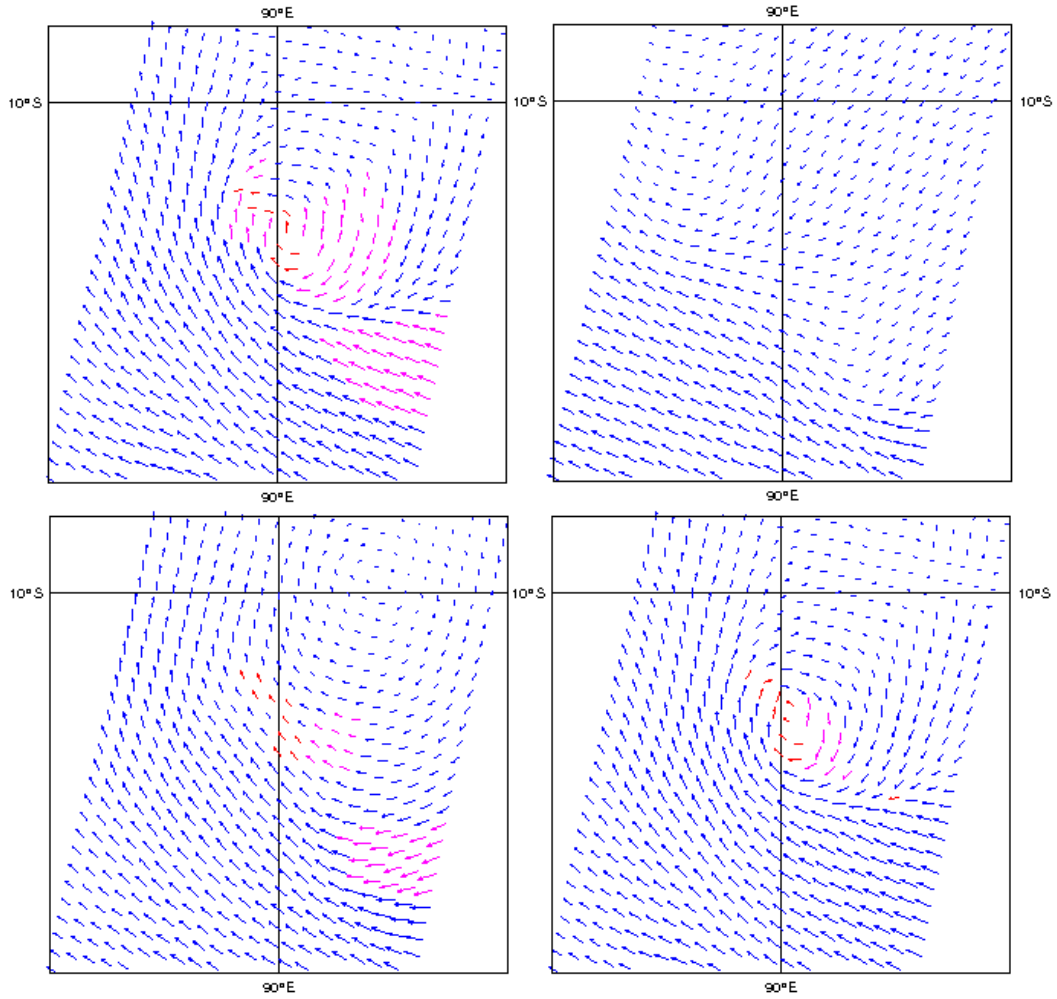


FIG. 14. Tropical cyclone east of India on August 31, 2007 observed with ASCAT at 25 km resolution. Upper left: standard processing result without MSS; upper right: ECMWF background; lower left: result with MSS; lower right: result without MSS and with 300 km background error correlation length.