

Assimilating ambiguous QuikScat scatterometer observations in HIRLAM 3-D-Var at the Norwegian Meteorological Institute

By FRANK THOMAS TVETER, *Norwegian Meteorological Institute, PO Box 43 Blindern, N-0313 Oslo, Norway*

(Manuscript received 3 September 2004; in final form 28 July 2005)

ABSTRACT

The Norwegian Meteorological Institute has implemented ambiguous QuikScat observations in the HIRLAM three-dimensional variational data assimilation system, by approximating the method that will achieve the best mean squared error verification results according to the Bayesian decision theory. A four-month long impact study of this implementation during the 2003/2004 winter showed that assimilating QuikScat observations had a small positive effect on the average mean sea level pressure forecasting skill.

1. Introduction

Placed on board the QuikScat polar orbiting satellite, the Sea-Winds scatterometer instrument covers large ocean areas giving valuable wind observations for operational meteorology (Atlas et al., 2001; Liu, 2002). The scatterometer is a radar, emitting microwave pulses towards the sea surface and measuring the backscattered signal from different directions as the satellite passes by the target area. The backscatter properties of the sea surface depend on the small capillary sea waves. These waves in turn depend on the near surface wind. An empirical relationship can be found between the backscatter measurements and the 10-m wind speed and wind direction relative to the antenna-viewing angle. It turns out that several wind vector solutions are typically in agreement with a set of backscatter measurements. A scatterometer observation is therefore usually presented as a set of wind vector solutions, called ambiguities. The NOAA/NESDIS (National Oceanic and Atmospheric Administration/National Environmental Satellite, Data and Information Service) provides QuikScat observations operationally on BUFR (Binary Universal Form for the Representation of meteorological data) format to the user community as backscatter measurements and the corresponding ambiguous wind vectors on 25-km resolution. In the context of the EUMETSAT/NWP SAF (European Organisation for the Exploitation of Meteorological Satellites/Numerical Weather Prediction Satellite Application Facility), the KNMI (the Koninklijk Nederlands Meteorologisch Instituut) provides software for averaging the 25-km resolution QuikScat backscatter measurements to 100 km, and retrieving the corresponding wind vector ambiguities (de Vries, 2004). By averaging the backscatter measurements, the retrieved winds become less ambiguous, and therefore more suitable for assimilation purposes.

Different techniques have been developed for assimilating scatterometer data in numerical weather prediction systems. Stoffelen and Anderson (Stoffelen and Anderson, 1997) developed a preprocessor that selected the ambiguities that gave the most spatially consistent wind field. These ambiguities were in turn assimilated in the ECMWF (European Center for Medium-Range Weather Forecasting) model as observations with Gaussian error characteristics. Portabella and Stoffelen (Portabella and Stoffelen, 2004) assimilated QuikScat observations by introducing a non-quadratic cost function to a (two-dimensional) variational analysis, similar to a method used, for instance, by Isaksen and Janssen (Isaksen and Janssen, 2004). They based their cost function on the principle that the resulting ‘analysis’ should be the ‘most probable state of the atmosphere’, as described by Lorenc (Lorenc, 1988). This principle had been suggested earlier by Lorenc (Lorenc, 1986) as one of two possible data assimilation principles, where the second principle was to use the ‘expected state of the atmosphere’ as the ‘analysis’. Ingleby and Lorenc (Ingleby and Lorenc, 1993) used Bayesian decision theory in a discussion of the two assimilation principles, identifying the two different ‘loss functions’ for which they are optimal. They did not elaborate on this difference. It is known from the Bayesian decision theory that the loss function for which the ‘expected state of the atmosphere’ is optimal, is a

Corresponding author.
e-mail: f.t.tveter@met.no

representation of the root mean squared error (RMSE) verification method. These results are re-visited in Section 2. The paper further shows how an ambiguous QuikScat observation can be represented as a ‘super-observation’ if we use the ‘expected state of the atmosphere’ as the ‘analysis’. The implementation of the 100-km resolution ambiguous QuikScat observations in HIRLAM 3-D-Var (3-dimensional variational analysis) (Gustafsson et al., 2001) at the Norwegian Meteorological Institute is presented in Section 3. Examples of super-observations are shown in Section 4, along with a comparison of the RMSE verification score from a forecast trial that uses QuikScat observations according to this implementation and scores from a trial that do not use these observations. Forecasts of two storms in the verification period are also compared.

2. Bayesian decision theory applied to numerical weather prediction

The *analysis* vector, $\mathbf{X}_a$, is in this context that the state of the atmosphere that is used as a starting point for the integration of the future state of the atmosphere in a forecast model. An analysis is determined using observations and a short forecast, also known as the first guess or background. If the first guess, $\mathbf{X}_b$, and the observations, $\mathbf{Y}$, have Gaussian error characteristics and no bias, and if the observation operator $\mathbf{H}$ is linear, the ‘best’ analysis (Lorenc, 1986) is given by

$$\mathbf{X}_a = \mathbf{X}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{Y} - \mathbf{H}\mathbf{X}_b), \quad (1)$$

where $\mathbf{R}$ is the observation error covariance matrix and $\mathbf{B}$ is the background error covariance matrix. This analysis can be derived by seeking the ‘most probable state of the atmosphere’ according to the available model for the probability density of the true state of the atmosphere given the observations and first guess, $P[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b]$. Note here that we are considering $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{X}_b$ to be random variables.

MSE verification methods are commonly used to verify numerical weather prediction systems. These methods typically calculate

$$\text{MSE penalty} = \frac{1}{m} \sum_{i=1}^m (Y_i - H_i(\mathbf{X}_f))^2, \quad (2)$$

where Y_i is an independent reference observation (of a given type), $H_i(\mathbf{X}_f)$ is the model forecast for the reference observation and m is the number of reference observations. The MSE penalty is calculated separately for different reference observation types, and a numerical weather prediction system is considered to be ‘better’ if it gives an overall lower MSE penalty for all reference observation types. Note that these results are not changed if we compare the square root of the MSE penalty function, also referred to as RMSE verification.

For short forecasts, the MSE verification methods provide a measure of the quality of the analysis. In the Bayesian decision theory, the MSE verification method for short forecasts, linear

reference observation operators and un-specified reference observation types, can be represented by the quadratic *loss* function, $l[\mathbf{X}, \mathbf{X}_a] = (\mathbf{X} - \mathbf{X}_a)^T(\mathbf{X} - \mathbf{X}_a)$, where $\mathbf{X}$ represents the unknown true state of the atmosphere. The corresponding *Bayes risk* is given by

$$\begin{aligned} B(\mathbf{X}_a) &= E[l[\mathbf{X}, \mathbf{X}_a]] \\ &= E[E[(\mathbf{X} - \mathbf{X}_a)^T(\mathbf{X} - \mathbf{X}_a)|\mathbf{Y} \cap \mathbf{X}_b]] \\ &= E[\text{Var}[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b]] \\ &\quad + E[(E[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b] - \mathbf{X}_a)^T(E[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b] - \mathbf{X}_a)] \end{aligned}$$

A small Bayes risk corresponds to a good MSE verification score. An important result from the Bayesian decision theory (Ingleby and Lorenc, 1993; Rice, 1995) is that the $\mathbf{X}_a^{**}$ function that minimizes this particular Bayes risk, is given by

$$\mathbf{X}_a^{**} = E[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b], \quad (3)$$

which is ‘the expected state of the atmosphere’.

3. Assimilating ambiguous observations in HIRLAM 3-D-Var

3.1. The HIRLAM 3-D-Var cost function

The analysis algorithm used in HIRLAM 3-D-Var, is designed to find the ‘most probable state of the atmosphere’ from a model for the probability density function for the atmosphere, $P[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b]$, (Lorenc, 1986). This is done by minimizing a cost function, $J[\mathbf{X}]$, given by

$$J[\mathbf{X}] \propto -\log[P[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b]].$$

We see that the atmospheric state $\mathbf{X}$ which minimizes $J[\mathbf{X}]$, maximizes $P[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b]$. If Bayes rule is applied and it is assumed that the observation errors are independent of the first guess errors, and further that the observation errors are independent of each other, the cost function can be written as

$$\begin{aligned} J[\mathbf{X}] &= J_b[\mathbf{X}] + J_o[\mathbf{X}] \\ J_b[\mathbf{X}] &\propto -\log[P[\mathbf{X}|\mathbf{X}_b]] \\ J_o[\mathbf{X}] &\propto -\log[P[\mathbf{Y}|\mathbf{X} \cap \mathbf{X}_b]] \\ &= \sum_{i=1}^r J_{o,i}[\mathbf{X}], \end{aligned}$$

where the index i denotes a single observation, and r is the number of observations. Adding a new observation denoted by $r + 1$, is then simply done by adding a contribution $J_{o,r+1}[\mathbf{X}] = -\log[P[\mathbf{Y}_{r+1}|\mathbf{X} \cap \mathbf{X}_b]]$ to the cost function. The cost function is minimized by a general mathematical algorithm, and the starting point for the minimization is the first guess, $\mathbf{X}_b$. The HIRLAM 3-D-Var minimization algorithm (conjugate gradients) works best when the cost function is quadratic and does not change the shape from one iteration to the next.

3.2. The Gaussian analysis

Let us assume that we have a set of observations with Gaussian error characteristics, $\mathbf{Y}$, and another set of independent non-Gaussian observations, $\tilde{\mathbf{Y}}$. Using Bayes rule, we may write the probability density function for the state of the atmosphere as

$$P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \mathbf{Y} \cap \mathbf{X}_b] = \frac{P[\tilde{\mathbf{Y}}|\mathbf{X} \cap \mathbf{Y} \cap \mathbf{X}_b]P[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b]}{P[\tilde{\mathbf{Y}}|\mathbf{Y} \cap \mathbf{X}_b]}. \quad (4)$$

The probability density function for observations with linear observation operators and Gaussian error characteristics may be modelled according to

$$P[\mathbf{Y}|\mathbf{X} \cap \mathbf{X}_b] \propto e^{-J_o[\mathbf{X}]}$$

$$J_o[\mathbf{X}] = \frac{1}{2} (\mathbf{Y} - \mathbf{H}\mathbf{X})^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{H}\mathbf{X}).$$

In HIRLAM 3-D-Var, it is assumed that the background error is Gaussian, giving

$$P[\mathbf{X}|\mathbf{X}_b] \propto e^{-J_b[\mathbf{X}]}$$

$$J_b[\mathbf{X}] = \frac{1}{2} (\mathbf{X} - \mathbf{X}_b)^T \mathbf{B}^{-1} (\mathbf{X} - \mathbf{X}_b).$$

Next, let us define a ‘Gaussian analysis’ by

$$\tilde{\mathbf{X}}_b = \mathbf{X}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{Y} - \mathbf{H}\mathbf{X}_b)$$

$$\tilde{\mathbf{B}}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}.$$

By comparing with eq. (1), we observe that $\tilde{\mathbf{X}}_b$ is the analysis which we would have if we only assimilated the first guess, $\mathbf{X}_b$, and the Gaussian observations, $\mathbf{Y}$.

It can be shown, by using Bayes rule and some algebra, that

$$P[\mathbf{X}|\mathbf{Y} \cap \mathbf{X}_b] = \frac{P[\mathbf{Y}|\mathbf{X} \cap \mathbf{X}_b]P[\mathbf{X}|\mathbf{X}_b]}{P[\mathbf{Y}|\mathbf{X}_b]} \\ = P[\mathbf{X}|\tilde{\mathbf{X}}_b]$$

$$P[\mathbf{X}|\tilde{\mathbf{X}}_b] \propto e^{-\tilde{J}_b[\mathbf{X}]}$$

$$\tilde{J}_b[\mathbf{X}] = \frac{1}{2} (\mathbf{X} - \tilde{\mathbf{X}}_b)^T \tilde{\mathbf{B}}^{-1} (\mathbf{X} - \tilde{\mathbf{X}}_b).$$

Further, since our observation errors are independent of each other and the first guess error, we have that

$$P[\tilde{\mathbf{Y}}|\mathbf{Y} \cap \mathbf{X} \cap \mathbf{X}_b] = \frac{P[\tilde{\mathbf{Y}} \cap \mathbf{Y} \cap \mathbf{X}_b|\mathbf{X}]P[\mathbf{X}]}{P[\mathbf{Y} \cap \mathbf{X} \cap \mathbf{X}_b]} \\ = \frac{P[\tilde{\mathbf{Y}}|\mathbf{X}]P[\mathbf{Y} \cap \mathbf{X}_b|\mathbf{X}]P[\mathbf{X}]}{P[\mathbf{Y} \cap \mathbf{X}_b|\mathbf{X}]P[\mathbf{X}]} \\ = P[\tilde{\mathbf{Y}}|\mathbf{X}].$$

We may now rewrite eq. (4), using the Gaussian analysis,

$$P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \mathbf{Y} \cap \mathbf{X}_b] = \frac{P[\tilde{\mathbf{Y}}|\mathbf{X}]P[\mathbf{X}|\tilde{\mathbf{X}}_b]}{P[\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}_b]} \\ = P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \tilde{\mathbf{X}}_b], \quad (5)$$

where we see that the normalization term now has been written as $P[\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}_b]$. We observe that the probability density function $P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \mathbf{Y} \cap \mathbf{X}_b]$ is equivalent to $P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \tilde{\mathbf{X}}_b]$, where the ef-

fects of the Gaussian observations, $\mathbf{Y}$, have been included in the Gaussian analysis, $\tilde{\mathbf{X}}_b$.

3.3. Assimilating a single ambiguous observation

An ambiguous scatterometer observation consists of n ambiguities, $\tilde{\mathbf{Y}} = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n\}$. For QuikScat, n is 2, 3 or 4, but in principle it may be any positive integer. The probability density function associated with a single ambiguous observation can be modelled according to

$$P[\tilde{\mathbf{Y}}|\mathbf{X}] \propto \hat{\gamma} + \sum_{k=1}^n \hat{q}_k P[\mathbf{Y}_k|\mathbf{X}]$$

$$P[\mathbf{Y}_k|\mathbf{X}] \propto e^{-\tilde{J}_k[\mathbf{X}]}$$

$$\tilde{J}_k[\mathbf{X}] = \frac{1}{2} (\mathbf{Y}_k - \mathbf{H}\mathbf{X})^T \tilde{\mathbf{R}}^{-1} (\mathbf{Y}_k - \mathbf{H}\mathbf{X}),$$

where $\hat{\gamma}$ represents the a priori probability for ‘gross error’ (within a given range), $\hat{q}_k$ represents the a priori probability that ambiguity k is ‘correct’, $\mathbf{Y}_k$ is the value of the k th wind vector ambiguity, $\mathbf{H}$ is the (linearized) observation operator (which calculates the 10-m wind in the target area from the model state) and $\tilde{\mathbf{R}}$ is the observation error covariance matrix for the ‘correct’ ambiguity. The a priori probability that ambiguity k is correct is available from the retrieval algorithm that estimates the ambiguities from the backscatter measurements. These a priori probabilities are based on how well the simulated backscatter (derived for each ambiguity), agrees with the actual backscatter observation.

We assume that the observation error is independent of the first guess error. Using the Gaussian analysis and our non-Gaussian scatterometer observation along with the Bayes rule, we may write the probability density function for the true atmospheric state as

$$P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \tilde{\mathbf{X}}_b] = \frac{P[\tilde{\mathbf{Y}}|\mathbf{X}]P[\mathbf{X}|\tilde{\mathbf{X}}_b]}{P[\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}_b]} \\ P[\tilde{\mathbf{Y}}|\tilde{\mathbf{X}}_b] = \int_{-\infty}^{\infty} P[\tilde{\mathbf{Y}}|\mathbf{X}]P[\mathbf{X}|\tilde{\mathbf{X}}_b] d\mathbf{X} \\ = \hat{\gamma} + \sum_{k=1}^n \hat{q}_k P[\mathbf{Y}_k|\tilde{\mathbf{X}}_b] \\ P[\mathbf{Y}_k|\tilde{\mathbf{X}}_b] \propto e^{-\tilde{J}_k} \\ \tilde{J}_k = \frac{1}{2} (\mathbf{Y}_k - \mathbf{H}\tilde{\mathbf{X}}_b)^T (\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1} (\mathbf{Y}_k - \mathbf{H}\tilde{\mathbf{X}}_b). \quad (6)$$

We may further write,

$$P[\mathbf{X}|\tilde{\mathbf{Y}} \cap \tilde{\mathbf{X}}_b] = \gamma[\tilde{\mathbf{X}}_b]P[\mathbf{X}|\tilde{\mathbf{X}}_b] \\ + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] \frac{P[\mathbf{Y}_k|\mathbf{X}]P[\mathbf{X}|\tilde{\mathbf{X}}_b]}{P[\mathbf{Y}_k|\tilde{\mathbf{X}}_b]} \\ \gamma[\tilde{\mathbf{X}}_b] = \frac{\hat{\gamma}}{\hat{\gamma} + \sum_{i=1}^n \hat{q}_i P[\mathbf{Y}_i|\tilde{\mathbf{X}}_b]} \\ q_k[\tilde{\mathbf{X}}_b] = \hat{q}_k \frac{P[\mathbf{Y}_k|\tilde{\mathbf{X}}_b]}{\hat{\gamma} + \sum_{i=1}^n \hat{q}_i P[\mathbf{Y}_i|\tilde{\mathbf{X}}_b]}. \quad (7)$$

Note that $\gamma[\tilde{\mathbf{X}}_b] + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] = 1$. Using eq. (3), it gives

$$\begin{aligned} \mathbf{X}_a^{**} &= E[\mathbf{X}|\mathbf{Y} \cap \tilde{\mathbf{X}}_b] \\ &= \int_{-\infty}^{\infty} \mathbf{X} P[\mathbf{X}|\mathbf{Y} \cap \tilde{\mathbf{X}}_b] d\mathbf{X} \\ &= \gamma[\tilde{\mathbf{X}}_b] \tilde{\mathbf{X}}_b + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] \mathbf{X}_k \\ \mathbf{X}_k &= \tilde{\mathbf{X}}_b + \tilde{\mathbf{B}}\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1}(\mathbf{Y}_k - \mathbf{H}\tilde{\mathbf{X}}_b). \end{aligned} \quad (8)$$

This is the ‘expected state of the atmosphere’ analysis when we are assimilating one ambiguous QuikScat observation. Note how that the Gaussian analysis in eq. (8) can be considered to be another ambiguity with a probability given by gross error.

3.4. The super-observation approach

Let us assume that we know the a posteriori probabilities that each ambiguity is correct, $q_k[\tilde{\mathbf{X}}_b]$. We may then define a superobservation according to

$$\tilde{\mathbf{Y}}_s[\tilde{\mathbf{X}}_b] = \gamma[\tilde{\mathbf{X}}_b] \mathbf{H}\tilde{\mathbf{X}}_b + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] \mathbf{Y}_k. \quad (9)$$

If we assimilate the superobservation as a conventional wind observation with Gaussian error characteristics given by $\tilde{\mathbf{R}}$ and in accordance to eq. (1), we get

$$\begin{aligned} \mathbf{X}_a &= \tilde{\mathbf{X}}_b + \tilde{\mathbf{B}}\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1}(\tilde{\mathbf{Y}}_s[\tilde{\mathbf{X}}_b] - \mathbf{H}\tilde{\mathbf{X}}_b) \\ &= \left(\gamma[\tilde{\mathbf{X}}_b] + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] \right) \tilde{\mathbf{X}}_b \\ &\quad + \tilde{\mathbf{B}}\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1} \left(\gamma[\tilde{\mathbf{X}}_b] \mathbf{H}\tilde{\mathbf{X}}_b \right. \\ &\quad \left. + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] \mathbf{Y}_k - \left(\gamma[\tilde{\mathbf{X}}_b] + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] \right) \mathbf{H}\tilde{\mathbf{X}}_b \right) \\ &= \gamma[\tilde{\mathbf{X}}_b] \tilde{\mathbf{X}}_b \\ &\quad + \sum_{k=1}^n q_k[\tilde{\mathbf{X}}_b] (\tilde{\mathbf{X}}_b + \tilde{\mathbf{B}}\mathbf{H}^T(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T + \tilde{\mathbf{R}})^{-1}(\mathbf{Y}_k - \mathbf{H}\tilde{\mathbf{X}}_b)) \end{aligned}$$

which is identical to eq. (8), i.e. the ‘expected state of the atmosphere’ analysis when we are assimilating one ambiguous QuikScat observation.

We could say that our ambiguous QuikScat observation contains less information than a conventional wind observation, since additional information ($\tilde{\mathbf{X}}_b$) is needed to estimate the superobservation, which can be assimilated in the same way as a conventional wind observation. As a consequence, the error in the superobservation is correlated with the error in the additional information used. However, assimilating the superobservation ($\tilde{\mathbf{Y}}_s$) together with the additional information ($\tilde{\mathbf{X}}_b$) will still give an optimal analysis, according to the theory above. If we instead had calculated the superobservations ($\tilde{\mathbf{Y}}_s$) using the first guess ($\tilde{\mathbf{X}}_b$), and then assimilated the superobservation ($\tilde{\mathbf{Y}}_s$) and first guess ($\tilde{\mathbf{X}}_b$) together with Gaussian observations ($\mathbf{Y}$), the

first guess would have been given a greater weight relative to the Gaussian observations due to the correlation between the superobservation error and the first guess error. By using the Gaussian analysis to calculate the superobservations, we do not upset the balance between the first guess and the Gaussian observations. Also, the quality of the superobservation depends on the quality of the additional information used. The Gaussian analysis should in theory have a smaller error than the first guess, and should therefore yield higher quality superobservations.

The a priori probabilities, $\hat{q}_k$, may peak in a range of values, while the a posteriori probabilities, $q_k[\tilde{\mathbf{X}}_b]$, calculated according to eq. (7) usually are $\sim 0\%$ or $\sim 100\%$. This is illustrated in Fig. 1, which shows an example of the probability distribution for a priori probabilities, and the corresponding a posteriori probability

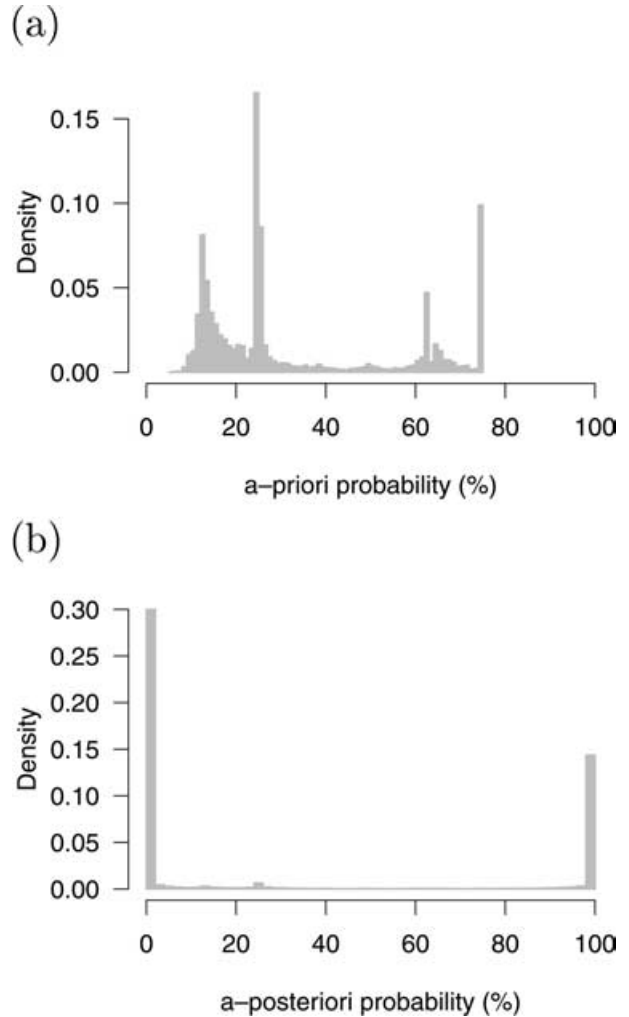


Fig. 1. (a) Statistical distribution of 36 000 (25-km resolution) a priori probabilities (all ambiguities). (b) Corresponding distribution of the a posteriori probabilities. The NCEP model wind that comes with the QuikScat observations was used as a first guess, and no conventional observations were used.

distribution. It follows that the superobservation in most cases is almost identical to one of the ambiguities. Note that when the wind speed is very low, the ambiguities are relatively close to each other compared to the Gaussian analysis error, and the a posteriori probabilities will in these cases remain close to the a priori probabilities. We could say that the Gaussian analysis contains little additional information when the wind speed is very low. It should also be noted that the superobservation combines two sources of information, namely the QuikScat ambiguous observation and a Gaussian analysis based on a weather prediction model. If the Gaussian analysis wind field is changed, by for instance changing the forecast model, a superobservation in an area where there is a large change in the Gaussian analysis may align itself with a completely different ambiguity. Obviously, the most accurate Gaussian analysis field will yield the ‘best’ superobservation.

As other ambiguous observations are added to our problem, the analysis becomes increasingly complex since the QuikScat observations themselves have correlated errors. A correct and complex analysis would have ensured that observations with correlated errors received lesser weight than similar observations with uncorrelated errors. We approximate this error correlation effect by reducing the weight of the superobservations (by increasing $\tilde{\mathbf{R}}$), and otherwise treating the superobservations as having independent errors. In HIRLAM 3-D-Var, the following cost function contribution is added for each QuikScat observation,

$$J_{o,k}[\mathbf{X}] = \frac{1}{2}(\tilde{\mathbf{Y}}_{s,k}[\tilde{\mathbf{X}}_b] - \mathbf{H}_k\mathbf{X})^T \tilde{\mathbf{R}}^{-1}(\tilde{\mathbf{Y}}_{s,k}[\tilde{\mathbf{X}}_b] - \mathbf{H}_k\mathbf{X}), \quad (10)$$

where the index k refers to superobservation number k . The (reduced weight) observation error covariance used to calculate the cost function contribution in eq. (10) is $(2.5 \text{ m s}^{-1})^2$ on the diagonal, and zero otherwise.

The superobservation is initially calculated using the first guess and an a priori gross error of $\hat{\gamma} = 0.1$. If the resulting superobservation is closest to the first guess (in wind vector space), the observation is rejected, otherwise the superobservation is re-calculated, this time using zero gross error. Figure 2 shows the distribution of the difference between the resulting superobservation and the first guess wind speed for a four-month long period (using 400 000 QuikScat observations). The average bias between the superobservation and the first guess wind speed is 0.8 m s^{-1} . Note the large ‘tail’ where the superobservations have higher wind speed than the first guess. The large ‘tail’ is interpreted as the effects of rain contamination (which gives a stronger radar backscatter resulting in an overestimation of the wind speed). The superobservations with wind speeds more than 2.0 m s^{-1} greater than the first guess wind speed are discarded, so that the resulting bias comes down to 0.07 m s^{-1} . Note that this is a very crude and restrictive quality control method that will cause rejection of some ‘correct’ observations with potentially high information content. About 25% of the observations are

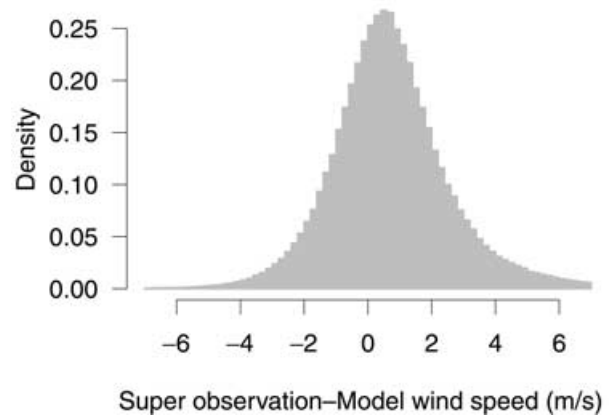


Fig. 2. Statistical distribution of the difference between superobservation wind speed and model wind speed, for 400 000 (100-km resolution) QuikScat observations (used in the impact experiment).

discarded on average due to the gross error check and the bias-correction procedure. About 1% of the QuikScat observations are discarded in the 100-km thinning procedure.

There are always 80 iterations in the minimization algorithm. The QuikScat observations are ignored for the first 20 iterations. At iteration number 20, it is assumed that the control vector has approached the Gaussian analysis (i.e. the analysis we would have if we ignored all QuikScat observations). The superobservations are then calculated according to eq. (10) (using the control vector) and their cost contributions are from now on added to the global cost function (no other changes are made to the assimilation system).

The Norwegian Meteorological Institute uses a fixed Gaussian innovation covariance matrix $(\mathbf{H}\mathbf{B}\mathbf{H}^T + \tilde{\mathbf{R}})$, with $(2.0 \text{ m s}^{-1})^2$ on the diagonal and zero otherwise, to calculate the a posteriori probabilities (in eqs 6 and 7). Note that $\tilde{\mathbf{B}}$ (which is never calculated explicitly) is implicitly defined by the global cost function (before the contributions from the superobservations are added). The superobservations are re-calculated every five iterations using the current state of the control vector. The matrix $\mathbf{H}\mathbf{B}\mathbf{H}^T + \tilde{\mathbf{R}}$ should ideally be re-estimated every time the superobservations are calculated so that it reflects the covariance in the differences between the control vector (Gaussian analysis) and the ‘correct ambiguity’. It turns out that the a posteriori probabilities are not too sensitive to $\mathbf{H}\mathbf{B}\mathbf{H}^T + \tilde{\mathbf{R}}$. There are for instance almost no visible changes in plots of superobservations, if the diagonal elements in $\mathbf{H}\mathbf{B}\mathbf{H}^T + \tilde{\mathbf{R}}$ are increased from $(2.0 \text{ m s}^{-1})^2$ to $(2.5 \text{ m s}^{-1})^2$. Since the a posteriori probabilities are not sensitive to $\mathbf{H}\mathbf{B}\mathbf{H}^T + \tilde{\mathbf{R}}$, this matrix is not changed during the minimization.

Variational quality control (Andersson and Järvinen, 1999) is activated from iteration number 40. In this setup, the superobservations will influence on how the variational quality control rejects observations from other data sources. However, variational

quality control has not been applied to the Gaussian analysis before it is used to calculate the QuikScat superobservations in the HIRLAM 3-D-Var implementation, and it is therefore not clear if the quality of these superobservations is better than those calculated from the first guess. In the end, it turns out that there is no significant difference in the impact from superobservations calculated using the Gaussian analysis compared to using the first guess (which is how superobservations are calculated when they are assimilated operationally). With the introduction of ‘outer loops’ (Courtier et al., 1994) in future releases of HIRLAM, a Gaussian analysis that has undergone variational quality control will be available. However, if one uses this Gaussian analysis to calculate the superobservation, the variational quality control for the conventional data in the first loop will not take the QuikScat observations into consideration. It is not clear at this point how QuikScat observations best should be combined with variational quality control in an outer loop implementation of HIRLAM 3-D-Var. Note that the current HIRLAM 3-D-Var implementation of QuikScat observations violates the assumption in the mathematical minimization algorithm that the cost function is quadratic, but apparently not enough to cause any serious problems.

4. Results

4.1. Examples

The superobservation approach has been applied to the 25-km resolution ambiguous QuikScat observations from NOAA/NESDIS in a now-casting product at the Norwegian Meteorological Institute. The 25-km resolution QuikScat observations in the NOAA/NESDIS BUFR product come with a NCEP (National Center for Environmental Prediction) model wind and an ambiguity selection. The QuikScat superobservations are calculated independently of each other, using the NCEP model wind as the first guess and no other Gaussian observations. In this setup, the Gaussian analysis, $\hat{\mathbf{X}}_b$, is equal to the first guess, $\mathbf{X}_b$. The gross error term is set to zero. Observations flagged as rain-contaminated by the NOAA/NESDIS QuikScat processing software or by a MLE rain-contamination check designed by KNMI, are rejected. Superobservations that are closer to the first guess than any of the ambiguities are also removed.

Figure 3 shows the NCEP model first guess wind for a weak low-pressure centre in the Atlantic Ocean, the ambiguities selected by NOAA/NESDIS and the corresponding 25-km resolution QuikScat superobservations. The QuikScat superobservations and the NOAA/NESDIS selected ambiguities are in close agreement, and suggest a more easterly location of the low-pressure centre. Figure 4 shows the NOAA/NESDIS selected ambiguities and the 25-km resolution superobservations for another case in the North Atlantic Ocean. Note that the NOAA/NESDIS selected ambiguities are more spatially consistent than the superobservations. This is what we would expect

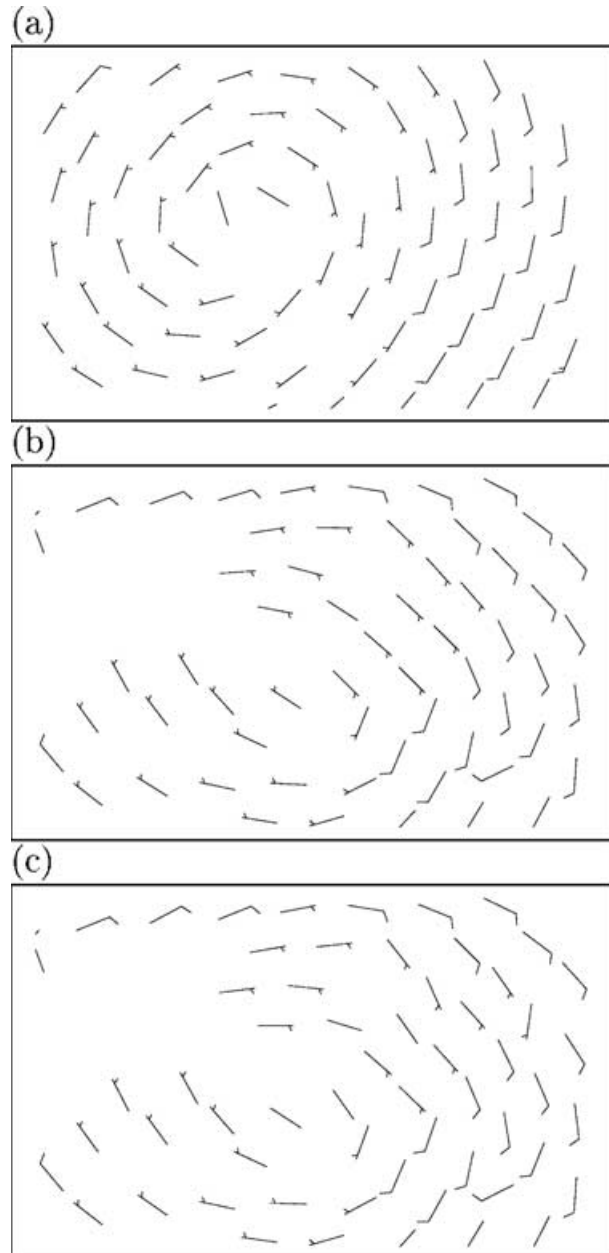


Fig. 3. (a) NCEP first guess model 10-m winds showing a small storm in the Atlantic Ocean (20N, 50W), at 29 December 2004, 09 UTC. (b) NOAA/NESDIS selected ambiguities (25-km resolution). (c) QuikScat superobservations (25-km resolution), where the NCEP model winds were used as first guess in the analysis.

since the superobservation analysis is done separately for each observation, while the procedure behind the NOAA/NESDIS selection takes spatial consistency into consideration in the analysis so that the selected neighbouring winds have similar wind direction. Figure 5 shows another case where the superobservations provide a wind field in agreement with independent ship observations. The corresponding NOAA/NESDIS selected

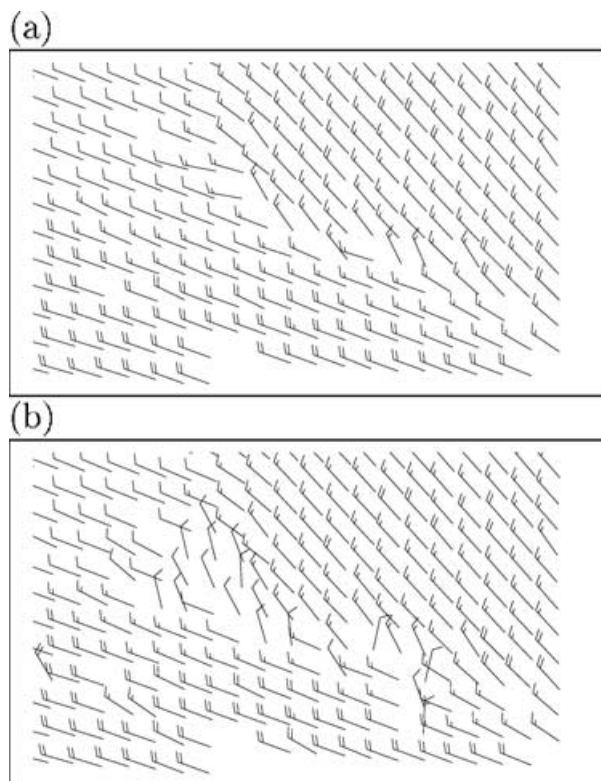


Fig. 4. (a) NOAA/NESDIS selected ambiguities (25-km resolution) in the North Atlantic Ocean (45N, 20W), at 28 December 2004, 06 UTC. (b) QuikScat superobservations (25-km resolution), where the NCEP model winds were used as first guess in the analysis.

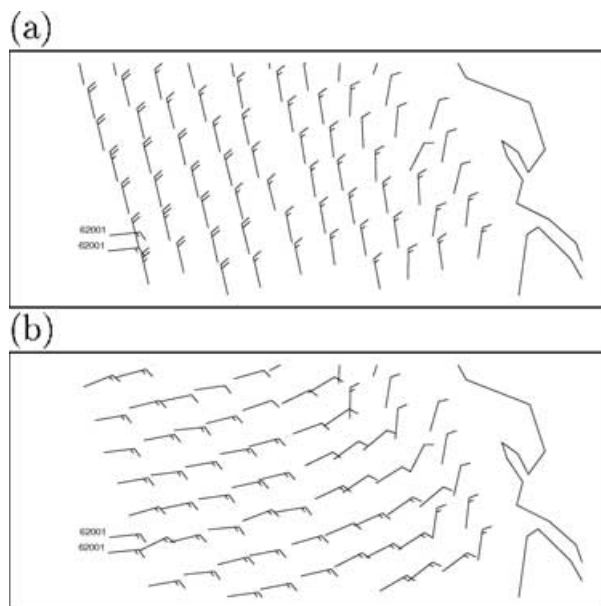


Fig. 5. (a) NOAA/NESDIS selected ambiguities (25-km resolution) close to France (45N, 5W), at 7 March 2005, 05 UTC. Two wind observations made by a ship are also shown. (b) QuikScat superobservations (25-km resolution), where the NCEP model winds were used as first guess in the analysis.

ambiguities have here supported each other in such a way that the resulting selection is spatially consistent, but completely wrong.

4.2. Verification results

An impact study was performed over four months from November 2003 until February 2004. Version 6.2.1 of the HIRLAM forecast model was used. The model used digital-filtering initialization, semi-Lagrangian advection scheme and physical parametrizations according to 'Savijärvi radiation', 'STRACO condensation', 'CBR turbulence scheme' and 'ISBA surface scheme' (Undén et al., 2002). Frames were taken from the ECMWF global model as lateral boundaries. The resolution was 20 km. Version 6.2.1 of the HIRVDA (HIRLAM Variational Data Assimilation) was used to assimilate observations in a 3-hr assimilation cycle with first guess at appropriate time (FGAT) enabled. The reference trial assimilated QuikScat observations along with ATOVS (Advanced TIROS Operational Vertical Sounder data) and conventional observations. The experiment trial was identical to the reference trial, except that it did not use the QuikScat observations. There was no observation cut-off (QuikScat data typically has a 2-hr time delay), and typically 500 QuikScat observations (100-km resolution) were available for an assimilation cycle. Note that the geometry of the QuikScat polar orbit causes the observation density to increase with latitude.

Figure 6 shows the geographical distribution of the standard deviation in the MSLP (mean sea level pressure) analysis between the experiment and reference trials, for the model area. Note that there are no visible differences over central Europe, where there are many conventional observations (and no QuikScat observations). Figure 7 shows the corresponding

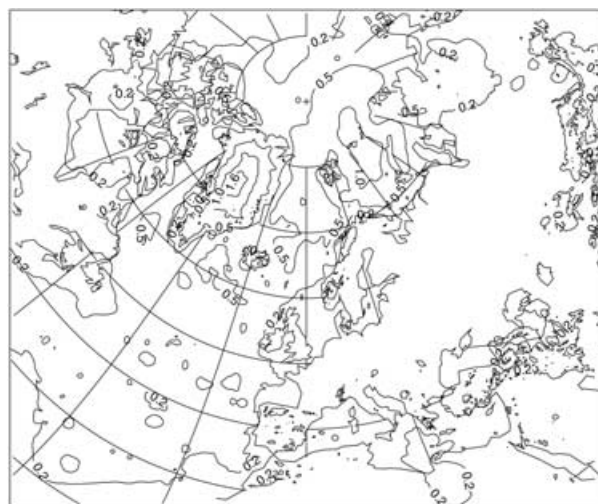


Fig. 6. Geographical distribution of the standard deviation in MSLP (analysis) between the reference and experiment trials.

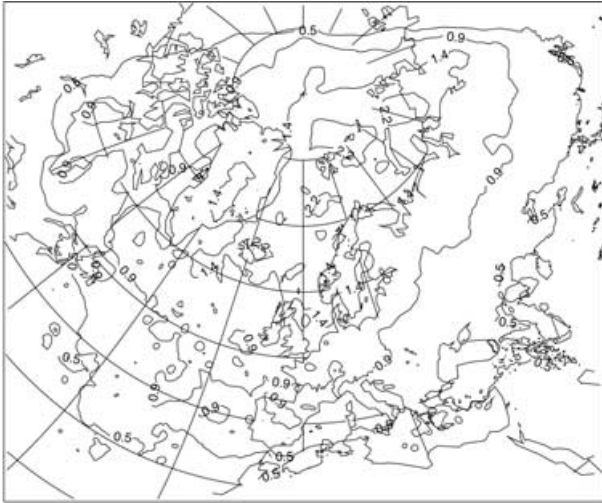


Fig. 7. Geographical distribution of the standard deviation in MSLP (48-hr forecast time) between the reference and experiment trials.

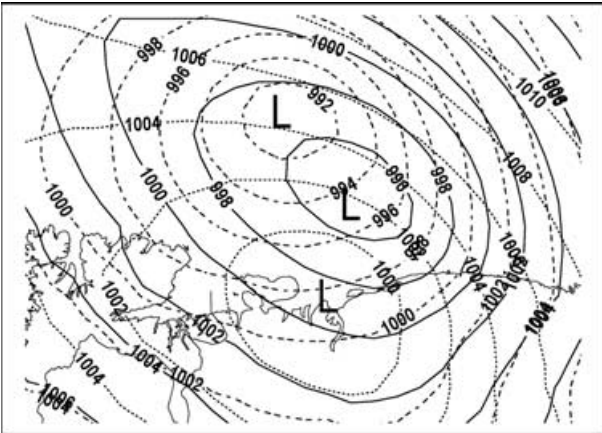


Fig. 8. An example of 24-hr forecasts at 15 February 2004, 00 UTC, in the reference (dotted) and experiment (dashed, no QuikScat) trials compared with the experiment analysis (solid) at the corresponding time. The reference analysis at the corresponding time was in close agreement with the experiment analysis.

standard deviation in the MSLP 48-hr forecast. The differences in central Europe originate from perturbations in weather systems that are advected inland from the ocean during the forecast period.

Figure 8 shows an example of a storm that moved in the Barents Sea along the coast of Norway, before turning inland at the Kola Peninsula. The figure shows 24-hr forecast from the reference and experiment (no QuikScat) trials, and the experiment analysis at the relevant time, just before the storm made landfall. We observe that the low in the experiment analysis is located in between the low in the experiment forecast and the low in the reference forecast. The relevant reference analysis was not too different from the experiment analysis (and in better agreement with the reference forecast, as we could expect).

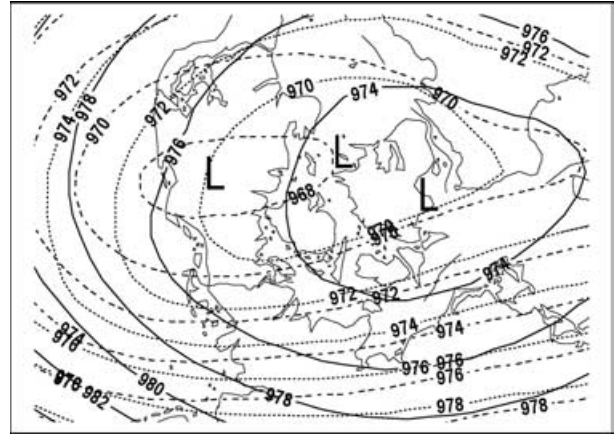


Fig. 9. An example of 24-hr forecasts at 14 January 2004, 00 UTC, in the reference (dotted) and experiment (dashed, no QuikScat) trials compared with the experiment analysis (solid) at the corresponding time. The reference analysis at the corresponding time was in close agreement with the experiment analysis.

Figure 9 shows an example of a storm that passed over Denmark on 14 January 2004. We observe that the experiment analysis low in this case is located closer to the reference trial than the experiment (no QuikScat) trial. Note that the pressure in a storm centre usually increases when it makes landfall.

The RMSE verification method used in this paper is based on the following penalty function,

$$\text{Rms}[t] = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{r,i} - \mathbf{H}_{r,i} \mathbf{X}[t])^2},$$

where m is the number of new independent verification observations (as suggested by the European working group on limited area modelling, for details on observations in the EWGLAM list, see for instance Vedel and Huang, 2004), i denotes each new independent observation (which has not yet been used in the assimilation), $\mathbf{H}_{r,i}$ is the forward operator for each new independent observation and $\mathbf{X}[t]$ is a $+t$ -hr forecast for the state of the atmosphere which is valid for the observation time. Only forecasts that start at 00 hr and observations at 00, 06, 12 and 18 hr are used in the verification. The different parameters (pressure, wind speed, etc.) are verified separately. The Rms is calculated and compared for the two trials. Note that the trial that gives lowest Rms penalty (wrt exactly the same, independent observations), is better according to this verification method. Figure 10 shows the standard deviation (Std) and bias (Bias) for MSLP for both the reference (with QuikScat) and experiment (no QuikScat) for the whole trial period (four months). Note that $\text{Rms}[t] = \sqrt{\text{Std}[t]^2 + \text{Bias}[t]^2}$. We observe that the trial where QuikScat observations were assimilated has a lower error (RMSE) than the trial without QuikScat observations. Figure 11 gives the same impression for the 10-m wind (FF10). The QuikScat observations made no significant impact on

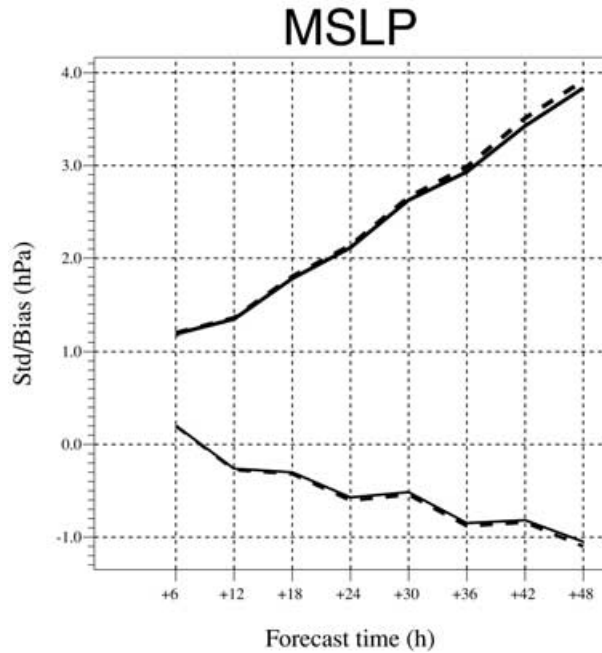


Fig. 10. Comparison of the MSLP verification for two trials, one with QuikScat observations and one without (dashed). The thick (upper) curves show standard deviation, while the thin (lower) curves show the bias (forecast minus observed). The verification period was 5 November 2003 to 29 February 2004.

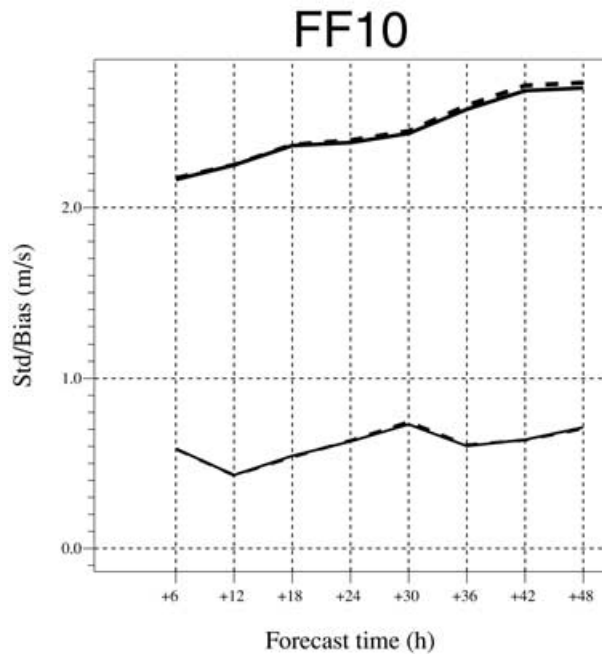


Fig. 11. Comparison of the 10-m wind speed verification for two trials, one with QuikScat observations and one without (dashed). The verification period was 5 November 2003 to 29 February 2004.

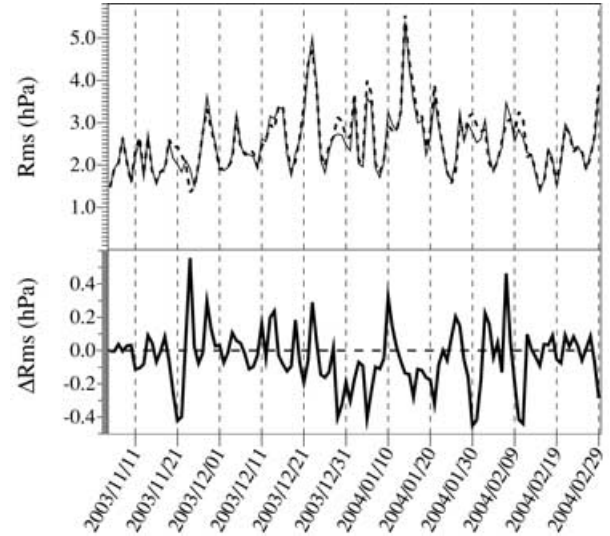


Fig. 12. Daily contribution to the MSLP verification for two trials, one with QuikScat observations and one without (dashed). The upper panel shows daily Rms for the two trials. The lower panel shows the difference between the two trials (a negative Δ Rms indicates that the QuikScat observations had a positive effect on the verification results).

the other surface parameters or parameters higher up in the atmosphere.

By studying variations in verification statistics for sub-periods of the four-month long verification period, we get an impression of how much we may expect the impact of assimilating QuikScat observations to vary. Figure 12 shows the daily contribution to the Rms ($\sum_t \text{Rms}[t]$) for the MSLP, and it is clear that the positive verification results accumulated during a period from the middle of December until the middle of February.

Since $\hat{q}_k$ is the a priori probability for each ambiguity, and q_k is the corresponding a posteriori probability (where we have taken the Gaussian analysis into consideration), we should have that $\hat{q}_k \simeq E[\hat{q}_k] = E[E[q_k]] \simeq \bar{q}_k$. We may use this relationship to check if we are using an innovation model which is consistent with the data. Table 1 gives the average a priori and a posteriori values for the probability that each ambiguity is correct, for a typical assimilation cycle. There is a reasonable agreement between the average a priori and a posteriori probabilities, indicating that the innovation model is acceptable.

Table 1. Average a priori and a posteriori ambiguity probabilities (%) for a typical assimilation cycle (100-km resolution)

n	$\bar{q}_1$	$\bar{q}_2$	$\bar{q}_3$	$\bar{q}_4$	$\bar{q}_1$	$\bar{q}_2$	$\bar{q}_3$	$\bar{q}_4$
4	79.16	13.57	4.24	3.03	75.22	10.65	3.49	2.47
3	91.51	8.32	0.17		88.55	5.20	0.00	
2	97.90	2.10			96.25	0.07		

5. Conclusions

When the purpose of the analysis is to find the ‘expected state of the atmosphere’, the ambiguous QuikScat observations may be presented to the analysis system as superobservations with Gaussian error characteristics. This superobservation approach has been implemented at the Norwegian Meteorological Institute in HIRLAM 3-D-Var. Assimilating QuikScat observations according to this approach gave an overall small positive impact in MSLP and FF10 for a four-month period during the 2003/2004 Winter.

6. Acknowledgments

This work was sponsored by the Norwegian Space Centre. The author is also grateful for the support from Dr. Ole Vignes, John Bjørnar Bremnes and suggestions contributed by Lars-Anders Breivik. The perspectives and comments from the two anonymous reviewers are also greatly appreciated.

References

- Andersson, E. and Järvinen, H. 1999. Variational Quality Control. *Q. J. R. Meteorol. Soc.* **125**, 697–722.
- Atlas, R., Hoffman, R. N., Leidner, S. M., Sienkiewicz, J., Yu, T.-W. and co-authors 2001. The effects of marine winds from scatterometer data on weather analysis and forecasting. *Bull. Am. Meteorol. Soc.* **82**, 1965–1990.
- Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4-D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367–1388.
- de Vries, J. 2004. QuikSCAT data processor user manual v1.3b. *KNMI*, NWPSAF-KN-UD-001 02/2004. Available at http://www.knmi.nl/scatterometer/sw_proc/pdf/.
- Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M. and co-authors 2001. Three-dimensional variational data assimilation for a limited area model, Part I: General formulation and the background error constraint. *Tellus* **53A**, 425–446.
- Liu, W. T. 2002. Progress in scatterometer application. *J. Oceanogr.* **58**, 121–136.
- Ingleby, B. and Lorenc, A. 1993. Bayesian quality control using multivariate normal distributions. *Q. J. R. Meteorol. Soc.* **119**, 1195–1225.
- Isaksen, L. and Janssen, P. A. E. M. 2004. Impact of ERS scatterometer winds in ECMWF’s assimilation system. *Q. J. R. Meteorol. Soc.* **130**, 1793–1814.
- Lorenc, A. C. 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **112**, 1177–1194.
- Lorenc, A. C. 1988. Optimal nonlinear objective analysis. *Q. J. R. Meteorol. Soc.* **114**, 205–240.
- Portabella, M. and Stoffelen, A. 2004. A probabilistic approach for SeaWinds data assimilation. *Q. J. R. Meteorol. Soc.* **130**, 127–159.
- Rice, J. A. 1995. *Mathematical Statistics and Data Analysis* 2nd Edition, Duxbury Press, California, 584 pp.
- Stoffelen, A. and Anderson, D. 1997. Ambiguity removal and assimilation of scatterometer data. *Q. J. R. Meteorol. Soc.* **123**, 491–518.
- Undén, P., Rontu, L., Järvinen, H., Lynch, D., Calvo, J. and co-authors 2002. HIRLAM-5 Scientific Documentation. Available at <http://hirlam.knmi.nl>
- Vedel, H. and Huang, X.-Y. 2004. Impact of ground based GPS data on NWP forecasts. *J. Meteorol. Soc. Japan* **82**, 459–472.