Tellus (1995), 47A, 62–79
Printed in Belgium – all rights reserved

Copyright © Munksgaard, 1995

TELLUS
ISSN 0280-6495

Preparations for the assimilation of ERS-1 surface wind observations into numerical weather prediction models

By MARIKEN HOMLEID* and LARS-ANDERS BREIVIK, *The Norwegian Meteorological Institute, P.O. Box 43-Blindern, N-0313 Oslo, Norway*

(Manuscript received 24 March 1993; in final form 9 March 1994)

ABSTRACT

This paper presents an investigation of how to use spatially dense and correlated observations in numerical weather prediction (NWP) models. As an example, we study the error characteristics of ERS-1 scatterometer winds, by comparison to collocated 10 m winds from the NWP model of The Norwegian Meteorological Institute. We attempt to determine standard deviations and horizontal correlation of the scatterometer wind errors. The potential of ERS-1 observations to improve the fields of geopotential height and wind in NWP models, is discussed within the context of statistical interpolation. The sensitivity of a three-dimensional, multivariate analysis scheme to the specification of the error structure of surface wind observations is studied within the same context. The results indicate that the analysis result is very sensitive to misspecification of the horizontal correlation when the observations are highly correlated and dense.

1. Introduction

The ERS-1 satellite was launched by ESA (the European Space Agency), 17 of July 1991. ERS-1 is the first European Remote Sensing satellite, and it carries several active radar instruments. Their mission is to sense oceanographic and meteorological parameters for experimental purposes and as a demonstration of the capability of the instruments in operational meteorology and oceanography on a global scale. Among the instruments is the AMI (Advanced Microwave Instrument) scatterometer, measuring radar backscatter, from which wind fields over the sea can be derived. The satellite covers ocean areas where few conventional observations are available. These wind fields are of interest as input to numerical weather prediction (NWP) models. The quality of the wind vectors produced from ERS-1 data has been validated and improved by ESA's geophysical analysis team consisting of scientists from several European meteorological centres. Some centres have also started assimilating the data into NWP models. Breivik et al. (1993) demonstrate a slightly improved forecast skill of DNMI's limited area model by inclusion of the data. In general the results depend on the observation quality as well as the capability of the analysis/forecast system to exploit the information.

To obtain positive impact of the data, it is important to utilize the data properly. The present paper reports on idealized experiments that have been performed to shed some light on questions such as: the potential of surface wind observations to improve the analysed fields; the optimal data density for assimilation; sensitivity to the specification of statistical parameters for the observation errors.

The experiments have been performed within the context of statistical interpolation (SI). Besides giving analysed values with estimates of the standard deviations of the analysis errors, the SI scheme also provides estimates of true standard deviations of the analysis errors when the applied statistical parameters deviate from the optimal ones in a known way. The equations and the experimental design are given in Section 3. SI is still the most commonly used method for analysing

* Corresponding author.

observations for NWP models. It is closely related to other methods such as variational analysis and Kalman filtering (Lorenc, 1986), making the results of the present study generally applicable. The analysis in the limited area model (LAM50) at the Norwegian Meteorological Institute (DNMI) is performed by successive corrections (Bratseth, 1986; Grønås and Midtbø, 1987). The analysis weights are constructed in such a way that the results converge towards the SI results. We will discuss whether the results of the sensitivity studies within the SI context, presented in Section 6, are applicable also for the analysis in LAM50. Seaman (1983) performs sensitivity studies with both SI and successive corrections and compares the results. He studies the sensitivity to departures of statistical properties of observational and background field errors from their true values, univariately and in two dimensions. In the present study we will concentrate on the sensitivity of three-dimensional, multivariate analysis results to the misspecification of statistical properties for correlated errors in surface wind observations.

A fundamental requirement for these studies is a thorough knowledge about the properties and error characteristics of the processed wind fields. In Section 2, the physical principles of the measurements and the data processing are discussed.

The quality of the wind vectors obtained so far, as validated by ESA's geophysical analysis team, are summarized in Section 4. Further studies of covariance structure of the observation errors have been performed by comparing the ERS-1 wind observations received from ESA to collocated surface winds from LAM50 at DNMI. These results are also presented in Section 4. In Section 5 the potential of ERS-1 data to improve wind fields in NWP models is discussed. The results of the sensitivity studies are presented in Section 6. We end up with summary and discussions in Section 7.

2. Surface winds from the ERS-1 scatterometer

ERS-1 is a polar orbiting satellite operating at about 785 km height. The mean period of one cycle is 100 min. The wind scatterometer is single-sided with 3 antennae pointing to the right hand side relative to the satellite flight direction, 45° ahead, 90° to the right and 45° back. The antennae cover a swath of 500 km width. Each of the 3 antennae produces a beam of radar energy at 5.3 GHz. Some energy is backscattered due to in-phase reflections from wind induced capillary/gravity-waves on the sea surface. The return signal is measured for different incidence angles, θ, from 18° to 57° nadir. The level of backscattering is termed Normalised Radar Cross-Section (NRCS) or σ^0. σ^0 increases with surface wind speed, decreases with incidence angle and depends on the radar beam angle relative to the wind direction as well. σ^0 might also be a function of other geophysical parameters, e.g., sea surface temperature and sea state, but these parameters have not yet been shown to have a significant effect on the accuracy of the retrieved wind vectors. Three values of σ^0 are obtained at each node on a grid of 25 km resolution. The radiometric resolution is 50 km.

A scatterometer model function, relating the measured σ^0 to wind speed, wind direction relative to antenna look angle 10 m above sea surface and incidence angle, has been fitted to empirical data. Using this transfer function in a minimisation when combining the measurements from the three antennae, wind speed and direction are found. This retrieval procedure generally converges to two ambigous wind vectors. The solutions are ranked in increasing order of the residual. The two most probable solutions have in general similar wind speed, but differ by roughly 180°. Choosing the "correct" solution is the task of the ambiguity removal or dealiasing. In ESA's operational suite the solution is chosen on the basis of wind direction consistency of nearby cells and the average value of the retrieval residual or the correlation with a forecast wind field from the ECMWF NWP model.

The processing and validation of the data is an iterative procedure. When the calculated wind fields show clear deficiencies, certain parts of the processing algorithms are subject to improvements. The results have improved significantly since the calibration period in 1991 by slightly changing the form of the empirical model function and by tuning the coefficients of this function. Also the wind retrieval procedure has been optimised. A more thorough description of the scatterometer and the wind retrieval procedure is found in e.g., Offiler (1990) and Stoffelen (1993).

3. Idealized experiments with horizontally correlated surface winds within the context of statistical interpolation

Multivariate, 3-dimensional statistical interpolation (SI) has been applied to study the sensitivity of the analysis result to the specification of covariances of the observation errors. In this section the SI equations are presented. The notation closely follows Daley (1991, Chapter 4), where also all model assumptions and derivation of the equations are given. The variables to be analysed are geopotential height, z, and wind vectors decomposed on a polar-stereographic analysis grid as u and v. Let $\boldsymbol{f}_{\mathrm{O}}$ be a vector consisting of K observations, $f_{\mathrm{O}k}$, $k=1, \ldots, K$. $\boldsymbol{f}_{\mathrm{B}}$, with elements $f_{\mathrm{B}k}$, $k=1, \ldots, K$ is a vector with corresponding background values, e.g., 6-h forecasts from the NWP model. The SI scheme gives linear estimates of the variables, $f_{\mathrm{A}i}$, at the analysis grid point, i, as weighted means of observations and background field:

$$f_{\mathrm{A}i}=f_{\mathrm{B}i}+\boldsymbol{W}_i^{\mathrm{T}}(\boldsymbol{f}_{\mathrm{O}}-\boldsymbol{f}_{\mathrm{B}}) \tag{3.1}$$

with expected error variance at grid point i,

$$E_{\mathrm{A}}^2=E_{\mathrm{B}}^2-2\boldsymbol{W}_i^{\mathrm{T}}\boldsymbol{B}_i+\boldsymbol{W}_i^{\mathrm{T}}(B+O)\,\boldsymbol{W}_i. \tag{3.2}$$

E_{B}^2 is the expected variance of the background error at grid point i.

$\boldsymbol{B}_i$ is a vector whose elements are background error covariances between all observation positions and the analysis grid point i.

O is the observation error covariance matrix with covariances between all observations.

B is the corresponding background error covariance matrix.

According to SI theory, optimal weights are

$$\boldsymbol{W}_i=(B+O)^{-1}\,\boldsymbol{B}_i, \tag{3.3}$$

implying minimum variance for the analysis errors:

$$E_{\mathrm{A}}^2=E_{\mathrm{B}}^2-\boldsymbol{B}_i^{\mathrm{T}}(B+O)^{-1}\,\boldsymbol{B}_i. \tag{3.4}$$

We will now derive an expression for the true variance estimated when the parameters made use of in the analysis equation deviate from the optimal values in a known way. The background errors and the observation errors are supposed to be multinormally distributed. Optimal values for SI are B, O, $\boldsymbol{B}_i$, $\boldsymbol{W}_i$, ... as before. All values from the erroneous model which we apply for SI are marked with a star, e.g., B^*, O^*. The weights calculated will then be:

$$\boldsymbol{W}_i^*=(B^*+O^*)^{-1}\,\boldsymbol{B}_i^*. \tag{3.5}$$

Inserting these weights in (3.2) gives the true variance:

$$E_{\mathrm{A}}^{*2}=E_{\mathrm{B}}^2-2\boldsymbol{W}_i^{*\mathrm{T}}\boldsymbol{B}_i+\boldsymbol{W}_i^{*\mathrm{T}}(B+O)\,\boldsymbol{W}_i^*. \tag{3.6}$$

In Section 6, we will see how the estimates of the true standard deviations of the analysis errors for z, u and v deviate from the optimal values when the covariances of the observation errors vary in a known way. In all experiments, the statistical parameters describing the covariances of the background errors are the same as those used in the analysis scheme at ECMWF (Lønnberg and Shaw, 1987). Parameters defining the degree of geostrophic coupling and divergence are taken from the same source.

The experimental analysis is performed on a grid with horizontal resolution 50 km; there are 61×49 grid points covering a 3000 km × 2400 km area. Results at 1000 and 850 hPa will be presented.

The experiments include only observations of surface winds from ERS-1. These observations are processed on a nearly regular grid with about 50 km between each observation. We made some simplifying assumptions concerning observational configuration: the observations are available in a regular grid, 50 km × 50 km; the observations are valid for the same pressure level, 1000 hPa.

4. Properties of surface winds from the ERS-1 scatterometer

Before utilizing the wind observations in an analysis scheme, the covariance structure of the observation errors must be specified. Results from validation obtained so far (Breivik et al., 1993) indicate that the standard deviations of the differences between observed and model first guess wind speed are about 2 m/s. Uncertainty in dealiasing gives rise to a 180° alias problem. The ESA real time processed wind directions are about

180° wrong in less than 5% of the cases. The alias problems most often occur in areas of light winds and in areas with high temporal variability, e.g., near low pressure centres and surface fronts. In many of these cases the wind speed is acceptable.

As will be illustrated in Section 6 the results of the analysis depend on good estimates of the statistical parameters describing the covariances, including the degree of horizontal correlation. Thus, we have put some effort into estimating these parameters. As the conventional observation network is rather sparse in ocean areas covered by ERS-1, the best source available for characterizing the covariance structure of the observation errors is analysed or short time forecasted surface winds from a NWP model. We have used 10 m winds from LAM50, which will be called background winds in the following. The ERS-1 observations give more small-scale information than LAM50, details which are impossible to verify in this comparison. The term "observation error" includes this small scale variance, which is generally referred to as representativeness error. Since both the observation errors and the background errors are horizontally correlated, the separation of the statistical parameters for each error type is not a trivial problem. In order to clarify and simplify the problem we adopt notation and introduce some commonly made assumptions for wind error covariances, following Daley (1991, Chapter 5).

Let $u(\mathbf{r}_1)$, $v(\mathbf{r}_1)$ and $u(\mathbf{r}_2)$, $v(\mathbf{r}_2)$ be the horizontal wind components on a polar-stereographic grid at two positions, $\mathbf{r}_1$ and $\mathbf{r}_2$, and let $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ be the vector between the positions. $\mathbf{r}$ can be described by r, the distance between the positions, and ϕ, the angle $\mathbf{r}$ makes with the east–west axis. When assuming homogeneity, the covariances between the errors of the u- and v-component of the background winds are only functions of r and ϕ and can be defined as $C_{uu,\mathrm{B}}(r,\phi)$, $C_{uv,\mathrm{B}}(r,\phi)$ and $C_{vv,\mathrm{B}}(r,\phi)$. The wind vectors can also be decomposed into a longitudinal component, l, along $\mathbf{r}$, and a transverse component, t, perpendicular to $\mathbf{r}$:

$$l=u\cos\phi+v\sin\phi, \qquad t=-u\sin\phi+v\cos\phi.$$

Under an assumption of isotropy, the covariances of the errors of l and t are only functions of r, not of ϕ, and are defined by $C_{ll,\mathrm{B}}(r)$, $C_{lt,\mathrm{B}}(r)$ and $C_{tt,\mathrm{B}}(r)$. The errors of the velocity potential and the stream function are assumed to be non-correlated (Hollingsworth and Lønnberg, 1986), implying that $C_{lt,\mathrm{B}}(r)=0$.

Variances of the errors of u, v, l and t are defined as $E^2_{u,\mathrm{B}}=C_{uu,\mathrm{B}}(\mathbf{r}=0)$, $E^2_{v,\mathrm{B}}=C_{vv,\mathrm{B}}(\mathbf{r}=0)$, $E^2_{l,\mathrm{B}}=C_{ll,\mathrm{B}}(\mathbf{r}=0)$ and $E^2_{t,\mathrm{B}}=C_{tt,\mathrm{B}}(\mathbf{r}=0)$.

The correlation functions, ρ, are defined through the formulas

$$\begin{aligned}
C_{uu,\mathrm{B}}(r,\phi)&=E^2_{u,\mathrm{B}}(r,\phi)\,\rho_{uu,\mathrm{B}}(r,\phi),\\
C_{uv,\mathrm{B}}(r,\phi)&=E_{u,\mathrm{B}}E_{v,\mathrm{B}}\,\rho_{uv,\mathrm{B}}(r,\phi),\\
C_{vv,\mathrm{B}}(r,\phi)&=E^2_{v,\mathrm{B}}\,\rho_{vv,\mathrm{B}}(r,\phi),\\
C_{ll,\mathrm{B}}(r)&=E^2_{l,\mathrm{B}}\,\rho_{ll,\mathrm{B}}(r),\\
C_{tt,\mathrm{B}}(r)&=E^2_{t,\mathrm{B}}\,\rho_{tt,\mathrm{B}}(r).
\end{aligned}$$

With the above assumptions the variances are identical:

$$E^2_{u,\mathrm{B}}=E^2_{v,\mathrm{B}}=E^2_{l,\mathrm{B}}=E^2_{t,\mathrm{B}},$$

and the correlation functions are related:

$$\begin{aligned}
\rho_{uu,\mathrm{B}}(r,\phi)&=\rho_{ll,\mathrm{B}}(r)\cos^2\phi+\rho_{tt,\mathrm{B}}(r)\sin^2\phi,\\
\rho_{uv,\mathrm{B}}(r,\phi)&=\rho_{vu,\mathrm{B}}(r,\phi)\\
&=(\rho_{ll,\mathrm{B}}(r)-\rho_{tt,\mathrm{B}}(r))\sin\phi\cos\phi,\\
\rho_{vv,\mathrm{B}}(r,\phi)&=\rho_{ll,\mathrm{B}}(r)\sin^2\phi+\rho_{tt,\mathrm{B}}(r)\cos^2\phi.
\end{aligned} \tag{4.1}$$

The same notation is used for the covariances of the wind observation errors:

$E_{u,\mathrm{O}}$, $E_{v,\mathrm{O}}$, $E_{l,\mathrm{O}}$ and $E_{t,\mathrm{O}}$-standard deviations of the errors of the u-, v-, longitudinal and transverse component of the observed wind;

$\rho_{uu,\mathrm{O}}(r,\phi)$, $\rho_{uv,\mathrm{O}}(r,\phi)$, $\rho_{vv,\mathrm{O}}(r,\phi)$, $\rho_{ll,\mathrm{O}}(r,\phi)$, $\rho_{lt,\mathrm{O}}(r,\phi)$, $\rho_{tt,\mathrm{O}}(r,\phi)$-correlation between these components as a function of distance, r, and direction, ϕ.

Furthermore we define statistical parameters describing the covariance structure of the *deviation* between observations and background winds:

$E_{l,\mathrm{D}}$ and $E_{t,\mathrm{D}}$-standard deviations of the longitudinal and transverse components;

$\rho_{ll,\mathrm{D}}(r,\phi)$, $\rho_{lt,\mathrm{D}}(r,\phi)$, $\rho_{tt,\mathrm{D}}(r,\phi)$-correlation between these components as a function of distance, r, and direction, ϕ.

Is it possible to extract characteristics of the observation errors if we assume the covariance structure of the errors of the background winds to be known? From observations and collocated background winds we have estimated standard deviations and correlation functions for transverse and longitudinal components of the deviations between observations and background winds. When assuming independence between observation error and background error, the standard deviations and correlations are related through the formulas:

$$
\begin{aligned}
&E_{l,\mathrm{O}}^2 = E_{l,\mathrm{D}}^2 - E_{l,\mathrm{B}}^2, \qquad E_{t,\mathrm{O}}^2 = E_{t,\mathrm{D}}^2 - E_{t,\mathrm{B}}^2,\\
&C_{ll,\mathrm{O}}(r,\phi) = C_{ll,\mathrm{D}}(r,\phi) - C_{ll,\mathrm{B}}(r),\\
&C_{lt,\mathrm{O}}(r,\phi) = C_{lt,\mathrm{D}}(r,\phi),\\
&C_{tt,\mathrm{O}}(r,\phi) = C_{tt,\mathrm{D}}(r,\phi) - C_{tt,\mathrm{B}}(r), \qquad (4.2)\\
&\rho_{ll,\mathrm{O}}(r,\phi) = C_{ll,\mathrm{O}}(r,\phi)/E_{l,\mathrm{O}}^2,\\
&\rho_{lt,\mathrm{O}}(r,\phi) = C_{lt,\mathrm{O}}(r,\phi)/E_{l,\mathrm{O}}E_{t,\mathrm{O}},\\
&\rho_{tt,\mathrm{O}}(r,\phi) = C_{tt,\mathrm{O}}(r,\phi)/E_{t,\mathrm{O}}^2.
\end{aligned}
$$

Suppose the covariance structure of the background winds is known. Then these formulae relate the standard deviation and the correlations of the observation errors with the standard deviation and the correlation of the deviations. Before trying to separate the statistics for the observation errors from the statistics for the background error, the properties of deviations between observations and collocated background winds are studied.

The collocation data for these studies consists of surface winds deduced from ERS-1 observations, processed by ESA in real time, from August, September and October 1993. The observations are collocated with forecasted 10 m winds from LAM50. The background values have been interpolated in time by cubic splines from 3, 6 and 9 hour forecasts to fit the observation time. Observations with direction deviating more than 90° from the background wind direction are excluded from the calculations. To avoid observations over sea ice, observations are also excluded when the sea surface temperature is less than 2°. Collocations are grouped in 6-h periods, centred around the analysis times 0, 6, 12 and 18 GMT. The deviations between observations and background winds have been decomposed into longitudinal and transverse components before estimating correlation functions and standard deviations. The calculations have been performed for data sets of varying size, from data sets consisting of only one pass to data sets consisting of all passes close to, e.g., 00 UTC for one month. To study possible anisotropies, the calculations have also been stratified based on the mid beam incidence angle of the instrument, on across-swath scan line and on wind direction.

Results from 6 different data sets consisting of observations close to 00 and 12 UTC for the months of August, September and October 1993 will be presented. The six data sets are called AUG00, AUG12, SEP00, SEP12, OCT00 and OCT12. When performing calculations on data sets of this size, we find that the estimates for standard deviations of longitudinal and transverse components are very close. They are ranging between 1.7 and 3.0. The standard deviations based on the observations with lowest incidence angle seem to be slightly higher than the others, indicating that the technique for retrieving wind fields from backscattered radar signal has difficulty when the incidence angle is small.

Empirical correlations based on averages in 50 km bins have also been calculated for different data sets. We obtained a wide range of curves, but found no significant dependencies on wind direction, scan line or incidence angle. We make the assumption that $\rho_{ll,\mathrm{D}}$ and $\rho_{tt,\mathrm{D}}$ are independent of ϕ. Fig. 1 presents empirical correlation curves based on all observations in each of the six data sets, AUG00, AUG12, SEP00, SEP12, OCT00 and OCT12, (a) for the longitudinal component, (b) for the transverse component. The correlations between the transverse components are slightly lower than the correlations of the longitudinal components. For background errors, the correlations of the transverse component are significantly lower than of the longitudinal component when the flow is close to non-divergent (see, e.g., Daley, 1991, Fig. 5.2). The cross correlations (not shown) are very close to 0, indicating that the longitudinal and transverse components of the deviations can be taken as independent.

When assuming that the longitudinal and the transverse components have the same standard deviations, that the correlation functions are independent of ϕ, and that there is no correlation between l and t, formulas (4.2) simplify to:

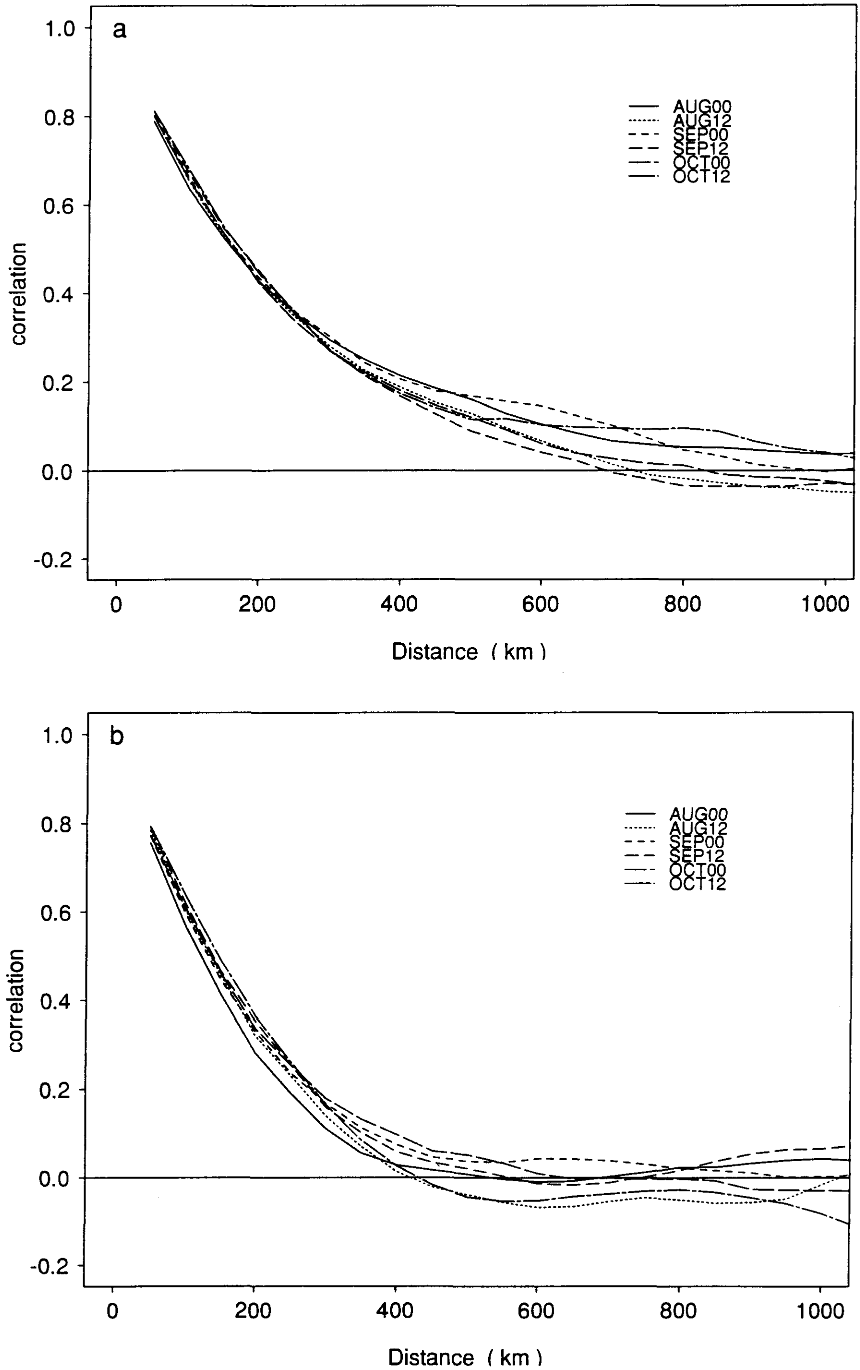

Fig. 1. Empirical correlation of longitudinal (a) and transversal (b) components of differences between observed and background wind as a function of distance between observations. One curve for each of the data sets AUG00, ..., OCT12, defined in the text.

$$
\begin{aligned}
&E_{l,\mathrm{O}}^2 = E_{t,\mathrm{O}}^2 = E_\mathrm{O}^2 = E_\mathrm{D}^2 - E_\mathrm{B}^2,\\
&C_{ll,\mathrm{O}}(r) = C_{ll,\mathrm{D}}(r) - C_{ll,\mathrm{B}}(r),\\
&C_{lt,\mathrm{O}}(r) = 0,\\
&C_{tt,\mathrm{O}}(r) = C_{tt,\mathrm{D}}(r) - C_{tt,\mathrm{B}}(r),\\
&\rho_{ll,\mathrm{O}}(r) = C_{ll,\mathrm{O}}(r)/E_\mathrm{O}^2,\\
&\rho_{lt,\mathrm{O}}(r,\phi) = 0,\\
&\rho_{tt,\mathrm{O}}(r) = C_{tt,\mathrm{O}}(r)/E_\mathrm{O}^2.
\end{aligned}
\tag{4.3}
$$

To extract values for the statistical parameters of the observation error, values of the standard deviations and correlation functions of the background error must be specified. The statistical parameters applied in analysis schemes to describe the background error are mean values deduced from empirical studies. The values are known to vary both in time and space. We will see how the background error, fixed within certain limits, suggest a range of variation for the observation errors, both for their standard deviations and their horizontal correlation.

For the deviations, we found $E_\mathrm{D} \in [1.7\ \mathrm{m/s}, 3.0\ \mathrm{m/s}]$. Major meteorological centres apply values for E_B of approximately 2 m/s in their analysis. Using $E_\mathrm{B} = 2$ m/s in (4.3) leads to values close to zero for the standard deviation of the observation errors. $E_\mathrm{B} = E_\mathrm{O}$ implies that $E_\mathrm{O} \in [1.2\ \mathrm{m/s}, 2.1\ \mathrm{m/s}]$. $E_\mathrm{B} = 1$ m/s implies that $E_\mathrm{O} \in [1.4\ \mathrm{m/s}, 2.8\ \mathrm{m/s}]$. In Section 5, we will see how sensitive the analysis result is to the variation of the standard deviation of the observation error within this range.

Based on (4.3), empirical correlations for the longitudinal and transverse components of the observation error have been calculated. Bessel functions are applied for describing the correlation of the background errors, Bessel functions truncated at radial mode 5 as applied in the analysis scheme in LAM50 and as documented in Lønnberg and Shaw (1987). Curves based on calculations in each of the 6 data sets with $E_\mathrm{B} = E_\mathrm{O}$ are presented in Fig. 2 as dotted lines, (a) for the longitudinal component, (b) for the transverse component. Corresponding curves with other values for E_B have also been produced. With $E_\mathrm{B} >$ 1.3 m/s, the resulting correlation curves are no longer monotonically decreasing, as can be seen in Fig. 2. The reason for this might be that the error correlations of the background errors are not well described by the applied Bessel functions. However, for values of $r < 150$ km or when $E_\mathrm{B} <$ 1.3 m/s, the curves for the longitudinal and transverse components are so close that we assume that $\rho_{ll,\mathrm{O}}(r) = \rho_{tt,\mathrm{O}}(r) = \rho_\mathrm{O}(r)$. This implies from (4.1) that

$$
\begin{aligned}
\rho_{uu,\mathrm{O}}(r,\phi) &= \rho_{ll,\mathrm{O}}(r)\cos^2\phi + \rho_{tt,\mathrm{O}}(r)\sin^2\phi\\
&= \rho_\mathrm{O}(r),\\
\rho_{uv,\mathrm{O}}(r,\phi) &= \rho_{vu,\mathrm{O}}(r,\phi)\\
&= (\rho_{ll,\mathrm{O}}(r) - \rho_{tt,\mathrm{O}}(r))\sin\phi\cos\phi = 0,\\
\rho_{vv,\mathrm{O}}(r,\phi) &= \rho_\mathrm{O}(r).
\end{aligned}
\tag{4.4}
$$

The correlation functions of the u- and v- component of the wind observation error are identical and their cross correlation equals 0.

As the empirical correlations seem to vary quite a lot, we have not put too much effort on fitting parametric functions to the empirical values. We have evaluated 3 classes of parametric functions by plotting them together with the empirical correlations. The candidates have been

Gaussian functions:

$G(r, R) = \exp(-0.5(r/R)^2)$;

spherical functions:

$S(r, R) = 1 - 1.5r/R + 0.5(r/R)^3$;

and 3rd-order autoregressive functions, which are used as structure functions for height correlations in the Canadian analysis scheme, Mitchell et al. (1990),

$$
\begin{aligned}
&\mathrm{TOAR}(r, R) = (1+\alpha)^{-1}\,(f(r,R) + \alpha f(r, nR)),\\
&f(r,R) = (1 + r/R + (r/R)^2/3)\exp(-r/R)\\
&\quad \text{where} \quad \alpha = 0.2 \quad \text{and} \quad n = 3.
\end{aligned}
$$

The empirical curves in Fig. 2 indicate that the correlation at 0 separation is less than one, e.g., 0.75. That means that the observation errors also have an uncorrelated part. Fig. 2 also presents 0.75 TOAR(45), 0.75 G(100) and 0.75 S(400). It seems like the decrease in correlation as a function of distance is well described by third order autoregressive functions for $r < 150$ km. The Gaussian function is too smooth close to zero while the spherical function is too sharp. In Section 6 we will

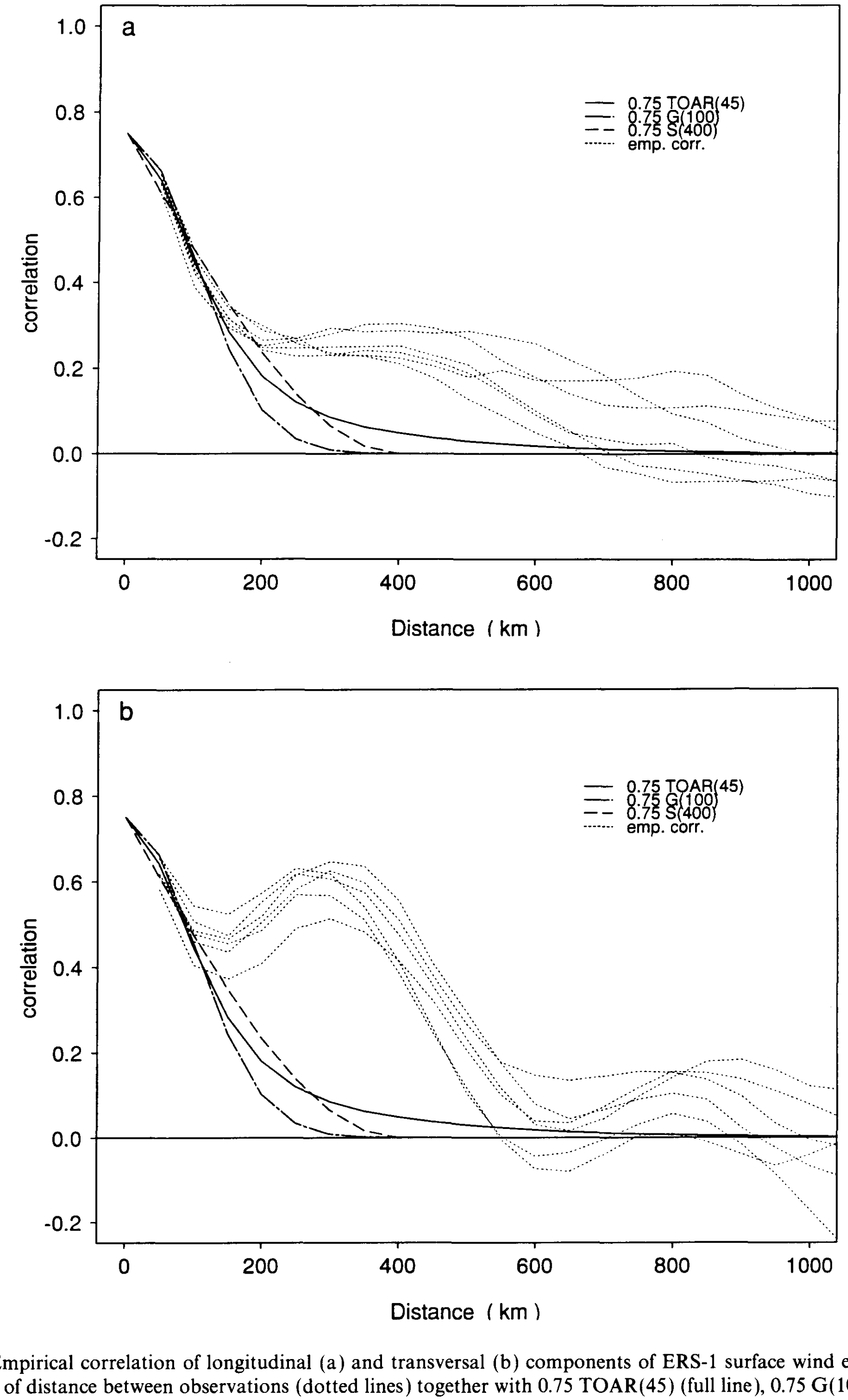

Fig. 2. Empirical correlation of longitudinal (a) and transversal (b) components of ERS-1 surface wind errors as a function of distance between observations (dotted lines) together with 0.75 TOAR(45) (full line), 0.75 G(100) (dash-dot line) and 0.75 S(400) (dashed line).

study the sensitivity of the analysis scheme to the shape and radius of the structure function and to the size of the uncorrelated part of the observation error.

We summarize the deduced properties of the observation errors, which will be underlying assumptions in the experiments presented in the next sections.

The correlation between the observation errors of the u-components of two wind vectors is identical with the correlation between the errors of the v-components, and depends only on the distance between the two wind observations.

The errors in u- and v-components are uncorrelated for all pairs of observations,

$$\rho_{uv,\mathrm{O}}(r) = 0.$$

The standard deviations of the observation errors are the same for all observations.

5. Potential of surface winds from the ERS-1 scatterometer

The scatterometer on ERS-1 provides a detailed description of the surface wind over the sea. To evaluate how important such information can be for the forecasts, case studies comparing forecasts based on fields analysed with ERS-1 data of varying resolution, and with no ERS-1 data, have to be performed. In this section the question is approached in a different way. What are the improvements of the analysed fields on the average? When assuming a fixed observation configuration and certain values for the statistical parameters describing the covariances of the observation and background errors, the SI scheme gives estimates of the standard deviations of the errors in the analysed values which are independent of the observations themselves. These estimates depend on the chosen statistical parameters. Underlying the experiments to be presented in this section are the simplifying assumptions made in Section 3. The calculations have been performed for three different sets of values describing the covariances of the observation errors:

(1) $E_\mathrm{O} = 1.5 \cdot E_\mathrm{B} = 2.9$ m/s, $\rho_\mathrm{O} = 0.75$ TOAR(80),
(2) $E_\mathrm{O} = 1.0 \cdot E_\mathrm{B} = 1.9$ m/s, $\rho_\mathrm{O} = 0.75$ TOAR(60),
(3) $E_\mathrm{O} = 0.75 \cdot E_\mathrm{B} = 1.4$ m/s, $\rho_\mathrm{O} = 0.75$ TOAR(40).

More relevant than the absolute values of the standard deviations of the analysis errors are the relative values of the standard deviations when calculated for different observation configurations. These numbers can be used as guidance for determining optimal observation density. Four different data sets have been used, with only the distance between neighbouring observations in the grid varying. The distances are 50, 100, 150 and 200 km, in grid A, B, C and D, respectively, as presented in Table 1. Standard deviations of the analysis errors of the geopotential height, z, and the u- and v-component of the wind vector have been calculated at all grid points within the analysis area for the 4 data sets and for the 3 sets of statistical parameters. Figs. 3a, 4a and 5a show normalized standard deviations estimated for z, u and v at 1000 hPa. The standard deviations are normalized with respect to the standard deviation of the background errors, $E_{z,\mathrm{B}} = 9.6$ m and $E_{u,\mathrm{B}} = E_{v,\mathrm{B}} = 1.9$ m/s at 1000 hPa. In this case $E_\mathrm{O} = 2.9$ m/s and $\rho_\mathrm{O} = 0.75$ TOAR(80), i.e., observation error set (1). The observations used are those of grid A with 50 km between each observation as indicated in Fig. 3a. The figures illustrate the area which has been influenced by observations. Outside this area the standard deviations of the analysis errors are equal to the standard deviation of the background error which, when normalized,

Table 1. *Observation configurations*

Distance between observations	Observation grid points	No. of observations
A. 50 km	(33, 34, 35, ..., 60, 61) × (21, 22, 23, ..., 28, 29)	$29 \times 9 = 261$
B. 100 km	(33, 35, 37, ..., 59, 61) × (21, 23, 25, 27, 29)	$15 \times 5 = 75$
C. 150 km	(33, 36, 39, ..., 57, 60) × (21, 24, 27, 30)	$10 \times 4 = 40$
D. 200 km	(33, 37, 41, ..., 57, 61) × (21, 25, 29)	$8 \times 3 = 24$

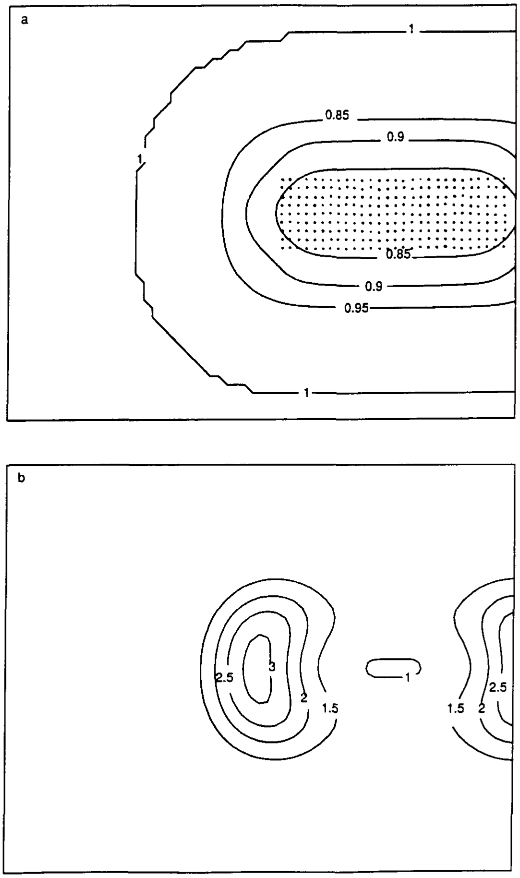

Fig. 3. (a) Optimal values of standard deviations of the analysis errors for z at 1000 hPa normalized with respect to $E_{z,\mathrm{B}} = 9.6$ m when utilizing observations from grid A. $E_\mathrm{O} = 1.5 \cdot E_\mathrm{B} = 2.9$ m/s, $\rho_\mathrm{O} = 0.75$ TOAR(80). (b) True values of standard deviations of the analysis errors for z at 1000 hPa normalized with respect to $E_{z,\mathrm{B}} = 9.6$ m when applying $\rho_\mathrm{O}^* = 0$ instead of $\rho_\mathrm{O} = 0.75$ TOAR(80). $E_\mathrm{O}^* = E_\mathrm{O} = 2.9$ m/s. Observations are from grid A.

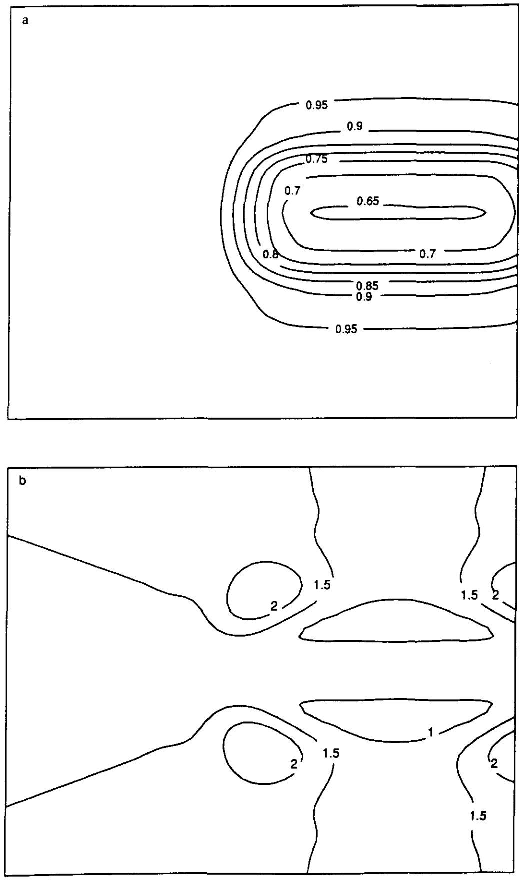

Fig. 4. As in Fig. 3 but for u. The values are normalized with respect to $E_{u,\mathrm{B}} = 1.9$ m/s.

equals 1. The strongest reductions in the standard deviations are found in the centre of the observation area, as expected. The minimum values are 0.80 for z, 0.64 for u and v. To simplify the presentation of results, mean values of the standard deviations within the observation area have been calculated. For this case, with observations from grid A, the mean values are 0.820 for z, 0.675 for u and 0.644 for v. The difference between the values for u and v is due to the orientation of the satellite pass in the $x - y$ grid. Corresponding numbers for the 4 observation configurations are presented in Table 2.

When utilizing observations with 50 km resolution from ERS-1 from one orbit (grid A), the mean values of the standard deviations estimated within the observation area become 0.820 for z, 0.675 for u and 0.644 for v. As the covariances specified for observation and background errors are supposed to be optimal, these numbers are the lowest we can get with observations of that quality. The numbers in Table 2 show that the mean values of $E_{z,\mathrm{A}}$, $E_{u,\mathrm{A}}$ and $E_{v,\mathrm{A}}$ do not increase dramatically when reducing the number of observations. Even when utilizing observations on a 200 km × 200 km grid, $E_{u,\mathrm{A}}$ and $E_{v,\mathrm{A}}$ are reduced by nearly 30%.

As stated before, the numbers in Table 2 depend highly on the specified covariances. The results

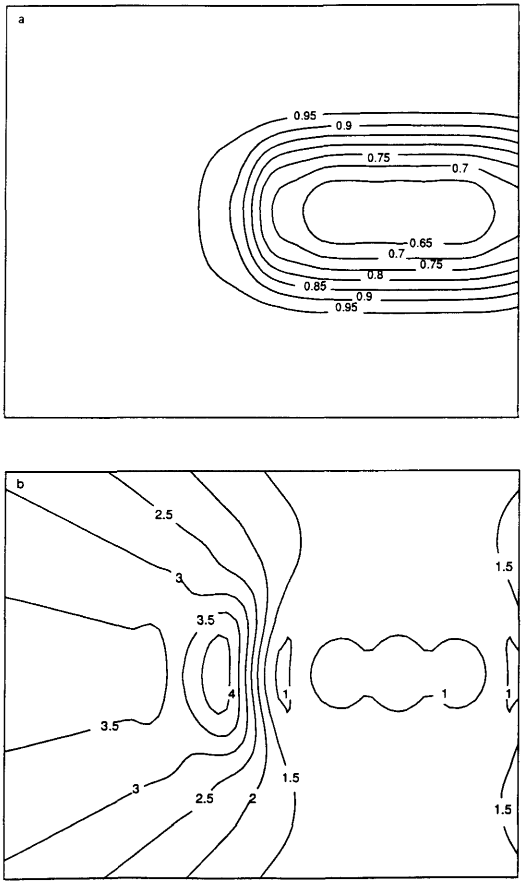

Fig. 5. As in Fig. 3 but for v. The values are normalized with respect to $E_{v,B} = 1.9$ m/s.

Table 2. *Mean values within the observation area for standard deviations estimated for the analysis errors of z, u and v at 1000 hPa with observation configurations A, B, C and D, and covariance structure for the observation errors given by* $E_O = 1.5 \cdot E_B = 2.9$ *m/s and* $\rho_O = 0.75$ *TOAR(80)*

Standard deviations of the analysis error estimated for			
Grid	z	u	v
A	0.820	0.675	0.644
B	0.826	0.692	0.662
C	0.833	0.711	0.684
D	0.840	0.738	0.707

Table 3. *As in Table 2 but for* $E_O = E_B = 1.9$ *m/s and* $\rho_O = 0.75$ *TOAR(60)*

Standard deviations of the analysis error estimated for			
Grid	z	u	v
A	0.734	0.530	0.492
B	0.740	0.544	0.510
C	0.747	0.564	0.534
D	0.757	0.597	0.562

from experiments with equal quality of ERS-1 observations and background, $E_O = E_B = 1.9$ m/s and $\rho_O = 0.75$ TOAR(60), are presented in Table 3. As expected, observations of higher quality imply a greater reduction in the standard deviations estimated for the analysis errors, about 50% within the observation area for u and v when utilizing all observations. When assuming that the observations are of higher quality than the background, $E_O = 0.75 \cdot E_B = 1.4$ m/s and $\rho_O = 0.75$ TOAR(40), the corresponding reduction in standard deviations is more than 60%, see Table 4. The conclusion concerning observation density does not change: there is no large difference between utilizing all the observations or only the observations on a 200 km × 200 km grid.

6. Sensitivity studies

What happens when the statistical parameters applied in our scheme for SI deviate from their optimal values? What are the consequences of assuming that observations have uncorrelated errors when, in reality, the errors are correlated? Comparing true standard deviations of analysed

Table 4. *As in Table 2 but for* $E_O = 0.75 \cdot E_B = 1.4$ *m/s and* $\rho_O = 0.75$ *TOAR(40)*

Standard deviations of the analysis error estimated for			
Grid	z	u	v
A	0.662	0.393	0.355
B	0.668	0.408	0.375
C	0.674	0.434	0.408
D	0.691	0.477	0.447

variables in an erroneous model with corresponding optimal values can indicate answers to these questions. In Subsections 6.1, 6.2 and 6.3, we will study the sensitivity to misspecifications of standard deviations, degree of horizontal correlation and structure function for the horizontal correlation for various optimal models for the covariance of the observation errors. The sensitivity to misspecified covariances of the observation errors depends on the quality of the observations. With high quality observations we find low sensitivity to misspecifications. The results presented in the following concern mainly the case of low quality observations. In Subsection 6.1 we will also pose the question of sensitivity in another way: What errors follow from applying a certain model for the observation error covariances if this model deviates from the optimal one? When presenting standard deviations of the analysis error, the standard deviations will always be normalized with respect to the standard deviations of the background error.

6.1. Sensitivity to misspecification of standard deviations and degree of horizontal correlation for observation errors

The sensitivity to misspecified covariances is also dependent on observation configuration. Obviously, and as will be demonstrated later in this section, the analysed fields are more sensitive to the specification of statistical parameters when the observation density is high. As an introductory example, results are presented from experiments using observations with a 50 km spacing (grid A). The observation errors are supposed to be highly correlated, $\rho_O = 0.75\,\mathrm{TOAR}(80)$, and with standard deviation $E_O = 1.5 \cdot E_B = 2.9$ m/s. In this situation the correlation between nearby observations is 0.7. Figs. 3a, 4a and 5a give optimal standard deviations of the analysis errors of z, u and v at each grid point when analysing all observations in grid A. When erroneously assuming too low, or no, horizontal correlation between observation errors, each observation will get too much weight in the analysis. The implied errors in the analysed fields are reflected in the true standard deviations for the analysis errors calculated by (3.6). Figs. 3b, 4b and 5b give true standard deviations for the analysis errors of z, u and v, when applying $\rho_O^* = 0$ instead of $\rho_O = 0.75\,\mathrm{TOAR}(80)$. The standard deviation of the observation error is assumed to be correct: $E_O^* = E_O = 2.9$ m/s. The figures show that the true standard deviations are close to or larger than 1 for all analysis grid points. This means that the background fields are degraded by the analysis. The degradation is not severe within the observation area, but just outside there are areas with large analysis errors.

Before commenting on this behaviour of the SI scheme we will present results from experiments to explore systematically the sensitivity to misspecification of both standard deviations and degree of horizontal correlation of the observation errors. The experiment described above has been repeated for all combinations of $E_O^* = b^* E_B^*$ where $b^* = 1, 1.2, ..., 2.2$ and $\rho_O^* = a^*\,\mathrm{TOAR}(80)$ where $a^* = 0, 0.1, 0.2, ..., 0.9$. The figures to be presented are based on use of observations from grid B with 100 km resolution and grid D with 200 km resolution, while results from grid A and grid C with 50 km and 150 km resolution are only commented on. Fig. 6 gives true standard deviations of the analysis errors of z, u and v at 1000 hPa averaged over the observation area. The figures illustrate that degradation of the background fields occurs only when the degree of horizontal correlation is assumed to be zero or close to zero *and* the standard deviations of the observation error is assumed to be low. The figures also suggest that assuming too low horizontal correlation can be compensated for by assuming a higher value than the optimal one for the standard deviations.

The true standard deviations are also studied at a grid point where maximum values are expected. Fig. 7 corresponds to Fig. 6, but gives results for grid point (28, 25) where extreme errors in analysed values of z and v are expected. We find that degradation of the background fields are caused by a wider range of values for a^* and b^*, and the errors when assuming no horizontal correlation and low standard deviation are bigger. The same is true for the analysis errors of u at grid point (32, 16) where this component has its largest errors (not shown).

When comparing results from experiments with varying observation density, it is evident that the sensitivity to misspecification of statistical parameters increases with increasing observation density. When calculating true standard deviations with observations from grid A, the most extreme parameter choices; horizontal correlation, $a^* = 0$ or 1, in combination with $E_O^* < 1.5 \cdot E_B^*$, implied

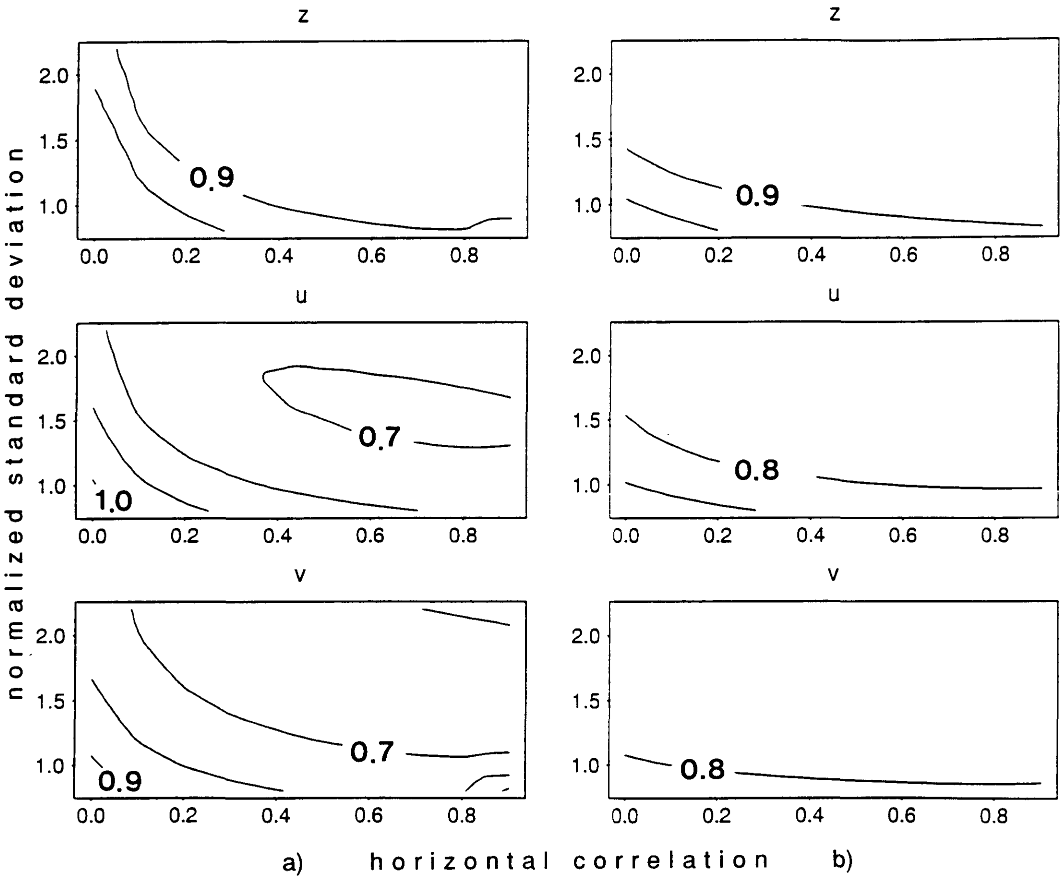

Fig. 6. Mean value for grid points in the observation area of true standard deviations of analysis errors of z, u and v at 1000 hPa, normalized with respect to E_B, as a function of a^* and $b^* = E_O^*/E_B^*$ (applied values for degree of horizontal correlation and normalized standard deviation of observation errors). Optimal values are $a = 0.75$ and $b = 1.5$. The left column, *a*, are based on results when observations are from grid B, the right, *b*, when observations are from grid D.

illconditioned covariance matrices. The results presented in the introductory example indicate to what degree the assumption of no horizontal correlation between observation errors, which in reality are highly correlated, can cause degradation of the background fields through the analysis. In this case we also found that it is difficult to partly compensate for the assumption of no horizontal correlation by increasing the value of the standard deviation. The result will frequently be an analysis with true standard deviations larger than 1. Seaman (1983) discusses the sensitivity of SI to departures of the assumed properties of observational and background field errors from their true values in the univariate case. He does also find large analysis errors in grid points close to observation dense areas when assuming that horizontally correlated observation errors are not correlated. When extrapolating to grid points outside the observation area, the SI weight given to each observation is very different depending on whether the observations are correlated or not. The same is not true when interpolating to grid points within the observation area.

The question of sensitivity can also be posed in another way: How far from optimal are the analysed fields when giving high quality satellite data too low weight through the parameters

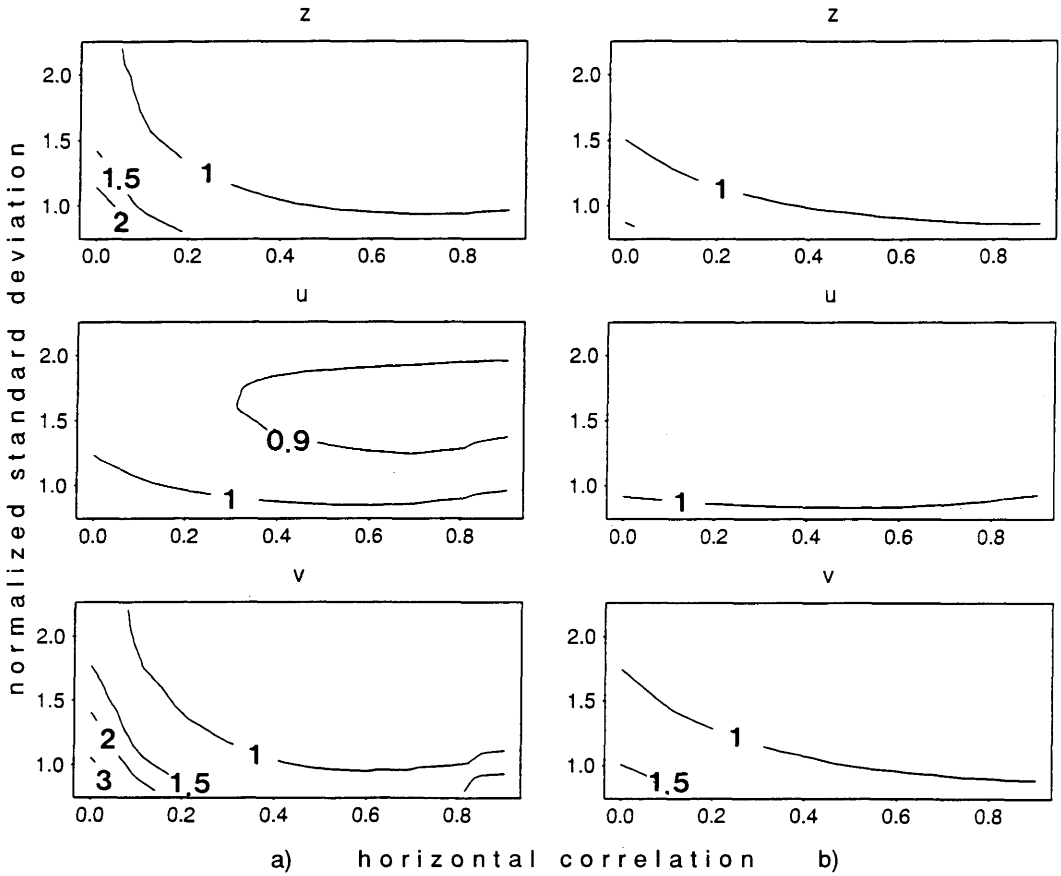

Fig. 7. True standard deviations of analysis errors of z, u and v at 1000 hPa at grid point (28, 25), normalized with respect to E_B, as a function of a^* and $b^* = E_O^*/E_B^*$ (applied values for degree of horizontal correlation and normalized standard deviation of observation errors). Optimal values are $a = 0.75$ and $b = 1.5$. The left column, a, are based on results when observations are from grid B, the right, b, when observations are from grid D.

defining the covariances of the observation errors? The answer depends on how far the specified parameters are from the optimal ones, on the density of the observations and on the analysis grid point. To study the problem, we will fix the statistical parameters for modelling the covariances of the observation errors in the SI scheme, E_O^* and ρ_O^*, and study deviations between true and optimal standard deviations of the analysis errors when the optimal parameters, E_O and ρ_O, vary. Let $E_O^* = 1.5 \cdot E_B^* = 2.9$ m/s and $\rho_O^* = 0.75\,\mathrm{TOAR}(80)$. This parameter choice would imply too low weight to the observations if their quality were higher, $E_O <$ 2.9 m/s, and/or their errors were less correlated, $a < 0.75$ in $\rho_O = a\,\mathrm{TOAR}(80)$. To get an indication of where in the analysis grid the largest deviations between true and optimal standard deviations of the analysis errors occur, we studied a case with $E_O = E_B$ and $a = 0.5$.

Observations from grid A, with 50 km resolution, were analysed. The largest differences between true and optimal standard deviations of the analysis errors for z, u and v were 0.04 for z, 0.08 for u and v and were found in the centre of the observation area for all components. Results from grid point (47, 25), in the centre of the observation area, will be presented to demonstrate the deviations between true and optimal analysis results as a function of the optimal values for a and $b = E_O/E_B$. Fig. 8 gives optimal values, true values

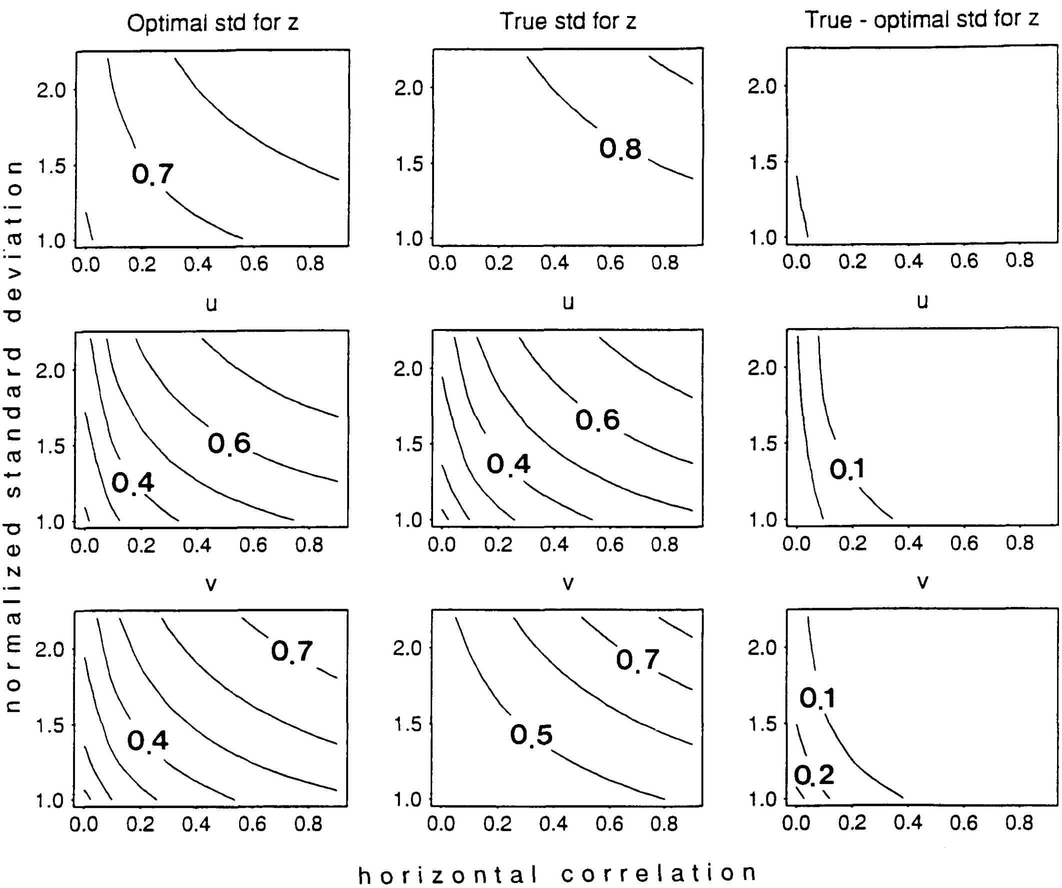

Fig. 8. Optimal values, true values and their differences of normalized standard deviations of the analysis errors of z, u and v at grid point (47, 25) at 1000 hPa as a function of the optimal values of a and $b = E_O/E_B$ when applying $a^* = 0.75$ and $b^* = 1.5$. Observations are from grid A.

and their difference for the standard deviations of analysis errors of z, u and v when analysing observations from grid A as a function of a and $b = E_O/E_B$. The figures illustrate how the analysis quality increases with increasing quality of the observations and decreasing horizontal correlation between the observation errors. The deviations between true and optimal values of the standard deviations become rather big compared to the optimal values only when the optimal values of a **and** $b = E_O/E_B$ are very low. When a is larger than 0.4, the deviations between true and optimal standard deviations of the analysis errors are found to be very small for all values of b.

When the observation density is reduced, the sensitivity to misspecification of the degree of horizontal correlation becomes negligible (not shown).

6.2. *Sensitivity to the structure of the horizontal correlation function*

In this section, we will study the sensitivity of the analysis result to the structure of the horizontal correlation function and find out whether a choice of a simpler structure function, e.g., spherical or gaussian, degrades the analysis result. We will assume that $E_O = 1.5 \cdot E_B = 2.9$ m/s and $\rho_O = 0.75$ TOAR(80). The sensitivity to the shape of the correlation function increases with increasing observation density. We will comment only on results from the most extreme case with observations in a $50 \text{ km} \times 50 \text{ km}$ grid. Fig. 2 presents

TOAR(45) together with S(400) and G(100). In this section we are considering S(800) and G(200) which are close to TOAR(80). The spherical function is sharper close to zero than the autoregressive, implying lower horizontal correlation between nearby observations. As it is the behaviour of the structure function for short distances which is the most important, it is not unexpected that the standard deviations estimated for the analysis errors of z, u and v are lower with S(800) than with TOAR(80). The minimum value estimated for v is about 0.1 lower with S(800). The differences are less for u and negligible for z. But what are the true standard deviations estimated when erroneously applying S(800) instead of TOAR(80)? The largest differences between true and optimal standard deviations estimated for v in this case are found in the centre of the observation area. They are negligible, less than 0.02. The mean differences between true and optimal standard deviations within the observation area for z, u and v are given in Table 5. All the values are found to be very small.

The experiment was repeated with the erroneous assumption of gaussian correlation function instead of autoregressive. The gaussian function is smoother close to zero, implying higher horizontal correlations between nearby observations, and thereby less weight to each observation. The standard deviations estimated for the analysis errors of z, u and v when assuming G(200) are higher than when assuming TOAR(80). But the true standard deviations estimated are also in this case very close to the optimal ones. Table 5 summarizes the mean differences within the observation area for z, u and v. The differences at 850 hPa were in both cases less than half as big as those at 1000 hPa.

Table 5. *Mean values within the observation area for deviation between true and optimal standard deviations estimated for z, u and v at 1000 hPa when applying S(800) or G(200) instead of TOAR(80) as correlation function when analysing wind observations in a 50 km × 50 km grid.* $E_O = 1.5 \cdot E_B = 2.9$ *m/s*

Correlation function	z	u	v
S(800)	0.005	0.014	0.019
G(200)	0.010	0.011	0.050

6.3. Sensitivity to the influence radius of the observation error correlation

This section is devoted to the problem of how sensitive the analysis results are to the specification of influence radius for the horizontal correlation of observation errors. The study is performed for observations in grid A. The optimal model for the observation error covariances is still defined by $E_O = 1.5 \cdot E_B = 2.9$ m/s and $\rho_O = 0.75$ TOAR(80).

To illustrate the sensitivity, experiments with four different influence radii have been performed. In all cases it is supposed that E_O is correct and that the structure of the correlation function is a TOAR. Only the modelled radii vary, from 0 (which means that the observation errors are not horizontally correlated), 40 km (implying that the observation errors are supposed to be less correlated than what is true), 80 km (which describes the correct model) to 120 km (implying that the observation errors are supposed to be more correlated than what is true).

The results from calculations of optimal and true standard deviations for errors in the analysed values of z, u and v have been studied at 1000 and 850 hPa. Maximum values for deviations between true and optimal standard deviations, for the analysis errors of z, u and v in 1000 hPa when assuming an influence radius other than 80 km, are found to occur outside the observation area.

The mean values of the true standard deviations within the observation area and the maximum deviations between true and optimal standard deviations are summarized in two tables, Table 6 for 1000 hPa and Table 7 for 850 hPa. Again we see that assuming no horizontal correlation, $r = 0$,

Table 6. *Results at 1000 hPa when varying the correlation function influence radius. Wind observations are on a 50 km × 50 km grid;* $E_O = 1.5 \cdot E_B = 2.9$ *m/s,* $\rho_O = 0.75$ *TOAR(r)*

	Mean of true standard deviations within observation area			Maximum deviation between true and optimal standard deviations		
r (km)	z	u	v	z	u	v
0	1.61	1.03	0.97	2.27	1.53	3.36
40	0.84	0.71	0.67	0.03	0.03	0.05
80 (optimal)	**0.82**	**0.68**	**0.64**	**0.00**	**0.00**	**0.00**
120	0.82	0.69	0.67	0.01	0.02	0.08

Table 7. *As in Table 6 but at 850 hPa*

	Mean of true standard deviations within observation area			Maximum deviation between true and optimal standard deviations		
r (km)	z	u	v	z	u	v
0	1.36	1.02	0.98	1.43	0.92	2.21
40	0.92	0.86	0.84	0.02	0.01	0.02
80 (optimal)	**0.91**	**0.85**	**0.84**	**0.00**	**0.00**	**0.00**
120	0.91	0.85	0.84	0.01	0.01	0.04

implies large deviations between true and optimal standard deviations for the analysis error. The results are not very sensitive to the choice of 40 km or 120 km influence radius instead of 80 km.

7. Summary and discussions

To be able to properly utilize observations with varying, and only partly known, error characteristics, it is important to know the sensitivity of the analysis result to misspecifications of the covariances of the observation errors. Experiments within the SI context can illustrate the average behaviour of the analysis scheme.

When comparing ERS-1 surface wind observations to 10 m winds from LAM50, we found standard deviations of the u- and v-components of the differences ranging between 1.7 m/s and 3.0 m/s. When analysing all observations from one satellite pass assuming that $E_O = E_B$ and $\rho_O = 0.75\,\mathrm{TOAR}(60)$, the standard deviations of the background errors will be reduced by 25% for z and 50% for u and v.

Sensitivity studies have been performed to examine how far from optimal the results will be when applying incorrect parameters to describe the observation error covariances. The conclusions are based upon the assumptions made in Section 3. The sensitivity depends on the degree of correlation of the observation errors, implying that the results also depend on observation configuration. A serious degradation of the background fields is only found when assimilating low quality observations on a grid of 50 km resolution, and when assuming no horizontal correlation **and** a low standard deviation of the observation errors. In this case too much weight is given to the observations. The largest errors occur close to, but outside, the observation area.

In the LAM50 analysis at DNMI we apply Bratseth's successive corrections scheme instead of SI. As it is not obvious that the convergence of the results with this scheme to SI results is fast over observation-dense areas (Seaman, 1988), it seems worthwhile to consider the different behaviour of SI and successive correction schemes in this respect. Seaman (1983) compares results obtained with SI and successive corrections. He finds that the SI results are better at the fringe of data-dense areas, assuming that the error covariances are correctly specified, while the results from successive corrections are less sensitive to misspecifications of the error covariances. Thus, the problem of extreme errors just outside the observation area due to misspecification of degree of horizontal correlation of the errors might be avoided with the successive correction scheme.

Inside the observation area the error made by assuming no, or very small, horizontal correlation can be compensated for by assuming higher standard deviations for the observation errors than what is correct. However, outside the observation area in some positions the assumption of high standard deviations does not compensate for the assumption of no horizontal correlation.

In light of the above results it seems reasonable to thin the observations if they have strongly correlated errors, at least when the correlation structure is unknown. The fact that the information loss through elimination of observations with correlated errors is relatively small, supports this suggestion.

8. Acknowledgements

This study is part of the first author's dr.scient studies, supported by The Research Council of Norway/Division NAVF. Nils Lid Hjort, Department of Mathematics and Statistics, University of Oslo, was her dr.scient supervisor. Peter Lønnberg, Finnish Meteorological Institute, has given invaluable support through discussions in the early stage of the work and careful reading and commenting on draft versions. We are also grateful to Thor Erik Nordeng, DNMI/ECMWF, for his comments on the paper.

REFERENCES

Bratseth, A. M. 1986. Statistical interpolation by means of successive corrections. *Tellus* **38A**, 439–447.

Breivik, L.-A., Haugse, B. and Homleid, M. 1993. Results from assimilation of ERS-1 scatterometer wind information in an operational limited area model. In: *Proceedings of the 2nd ERS-1 Symposium: "Space at the service of our environment"*. ESA Publication Division, ESTEC, Noordwijk, The Netherlands.

Daley, R. 1991. *Atmospheric data analysis.* Cambridge: Cambridge University Press, 98–187.

Grønås, S. and Midtbø, K. H. 1987. Operational multivariate analysis by successive corrections. *Proceedings: Collection of papers presented at the WMO/IUGG NWP Symposium*, Tokyo, 4–8 August 1986.

Hollingsworth, A. and Lønnberg, P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* **38A**, 111–136.

Lorenc, A. 1986. Analysis methods for numerical weather prediction. *Quart. J. R. Meteor. Soc.* **114**, 205–240.

Lønnberg, P. and Shaw, D. (ed.). 1987. Research manual 1. ECMWF Data Assimilation. Scientific Documentation, *Meteorological Bulletin*, ECMWF Reading.

Mitchell, H. L., Charette, C., Chouinard, C. and Brasnett, B. 1990. Revised interpolation statistics for the Canadian data assimilation procedure: their derivation and application. *Mon. Wea. Rev.* **118**, 1591–1614.

Offiler, D. 1990. Wind fields and surface fluxes. In: *Microwave remote sensing for oceanographic and marine weather-forecast models* (ed. R. A. Vaughan). Dordrecht/The Netherlands. Kluwer Academic Publishers, 355–374.

Seaman, R. S. 1983. Objective analysis accuracies of statistical interpolation and successive correction schemes. *Aust. Met. Mag.* **31**, 225–240.

Seaman, R. S. 1988. Some real data tests of the interpolation accuracy of Bratseth's successive correction method. *Tellus* **40A**, 173–176.

Stoffelen, A. 1993. From measurement to model: ERS-1 scatterometer data assimilation. *ECMWF Seminar Proceedings*, 6–10 September 1993, Reading, UK.