Tellus (1999), 51A, 469–488
Printed in UK. All rights reserved

Copyright © Munksgaard, 1999
TELLUS
ISSN 0280–6495

Variational assimilation of time sequences of surface observations with serially correlated errors

By HEIKKI JÄRVINEN, ERIK ANDERSSON* and FRANÇOIS BOUTTIER, *ECMWF, Shinfield Park, RG2 9AX Reading, Berkshire, UK*

(Manuscript received 15 December 1998; in final form 30 April 1999)

ABSTRACT

Assimilation of observations from frequently reporting surface stations with a four-dimensional variational assimilation system (4D-Var) is described. A model for the serial observation error correlation is applied to observed time sequences of surface pressure observations, whereby the relative weight of the mean information over the temporal variations is decreased in the assimilation. Variational quality control is performed jointly for each time sequence of observations so as to either keep or reject all observations belonging to a time sequence. The operational practice at ECMWF has previously been to use just one pressure datum from each station within each 6-h assimilation time window. The increase of observational information used in these assimilation experiments results in a small but systematic increase in the short-range forecast accuracy. The r.m.s. of the analysis increments is decreased in the experiments, which means there is an improved consistency between the background and the observations. A study of a rapidly developing small-scale synoptic system (the Irish Christmas Storm in 1997) showed that both the background and the analysis became more accurate when more frequent observations were assimilated. Single-observation experiments showed that a surface pressure time-sequence of data from a single surface station can intensify the analysis of a mid-latitude baroclinic system, that was underestimated in the background, when used in a 6-h 4D-Var. The method to assimilate time sequences presented in this paper has been implemented into the ECMWF operational 4D-Var assimilation system.

1. Introduction

Surface pressure tendency is one of the most accurately observed atmospheric quantities. Measurements are unaffected by instrument calibration problems, assuming any instrument bias remains constant during the interval between subsequent surface pressure measurements. Such observations are widely available and they have been used in subjective synoptic analysis and short-range forecasting providing information about the evolution and movement of synoptic disturbances. It has, however, proven to be difficult to make effective use of these observations in numerical data assimilation and weather prediction. The difficulty is due to the static nature of three-dimensional data assimilation schemes, such as Optimum Interpolation (Lorenc, 1981) or three-dimensional variational assimilation (3D-Var) (Parrish and Derber, 1992; Courtier et al., 1998), which do not take the time-dimension fully into account. The recent introduction of four-dimensional variational assimilation (4D-Var) in operational NWP (Rabier et al., 1999) has presented the possibility to use more frequent (asynoptic) observations at their appropriate time. In this paper we develop a technique to make use of the temporal variation in time sequences of surface observations, in 4D-Var.

* Corresponding author.
e-mail: e.andersson@ecmwf.int

It is conceivable to construct an observation operator for surface pressure tendency, also in a static scheme, like 3D-Var. Such an operator could be based on the continuity equation and involve the integration of the model divergence profile through the column of air above the observation location, thereby obtaining the model's instantaneous rate-of-change of surface pressure at the observation point. Pressure tendency observations could in such a scheme be used as static measurements of vertically integrated divergence, provided that a dynamical constraint is applied to prevent the information to project on high frequency gravity mode oscillations. There is however a point of concern behind this idea, which explains why it has to our knowledge not been attempted. The reporting practice of surface pressure tendency at synoptic stations does not provide observations of instantaneous surface pressure change but rather observations of pressure change during the last 3 h. The absolute value of the change is a difference between the barometer reading at observing time and the reading 3 h earlier. This value is supplemented in SYNOP reports with a qualitative description of the nature of the change as it appears on a barograph paper, for instance increasing then steady. The observation operator for 3D-Var suggested above, does not therefore correspond to the actual available surface pressure tendency quantity.

Because of the reporting practice, the surface pressure tendency observations are, in fact, redundant whenever the previous surface pressure observation itself is available. In other words, surface pressure tendency information can be recovered from the time sequences of surface pressure observations in an assimilation scheme that allows the use of time series of data. There are 2 possibilities: either using the pressure difference between consecutive pressure observations, or using the sequence of pressure observations itself. The latter approach has been taken in this paper.

In contrast to static assimilation schemes, 4D-Var compares the observations to a model trajectory (Rabier et al., 1998a) which extends over the assimilation time window, currently 6 h at ECMWF. There are 2 related benefits for the use of observations. First, the observations are compared with the model counterparts at the appropriate time. Second, it should be possible to update the model trajectory by using the dynamic information contained in the time sequences of observations. (The OI/3D-Var implementation of FGAT (First Guess at Appropriate Time) also compares the observations at appropriate time but can only improve the assimilation of the mean observed state.) The operational implementation of 4D-Var at ECMWF (Rabier et al., 1999) is a temporal extension of 3D-Var (Courtier et al., 1998; Rabier et al., 1998b; Andersson et al., 1998). The observation operators (Courtier et al., 1998; Andersson et al., 1994; Stoffelen and Anderson 1997; Phalippou 1996) and the background term (Derber and Bouttier 1999) are the same as in 3D-Var. Therefore, it ought to be possible to assimilate observed surface pressure variations in 4D-Var by applying the existing observation operator to time sequences of surface pressure observations. This idea can be extended to other variables as well, for instance to winds, temperatures and humidities. Difficulties may arise, however, if the observed time series is affected by, e.g., local thermal circulation systems not resolved by the assimilating model.

Daley (1992) foresaw that 4D-Var may require even greater knowledge of relevant observation error statistics than earlier data assimilation schemes. In his study of serially correlated observation errors in a Kalman Filter framework Daley found that the existence of serial correlation reduces the information content in the observations, and reduces the analysis accuracy (compared to serially un-correlated observations). However, the effect is small for observation types characterized by low sampling frequency (e.g., radiosonde observations), and it is also less important if the advecting wind velocity is high. In this paper we use observations with a 1-h sampling frequency near the surface where the advecting wind velocity is often small. In accordance with Daley's theoretical analysis we found it necessary to generalize the 4D-Var formulation and incorporate serial correlation of observation error, in order to make effective use of these frequent observations.

Our work on pressure time-sequences was strongly motivated by Bengtsson's (1980) theoretical paper. Long before the time of 4D-Var Bengtsson studied the optimal analysis response to pressure tendency data in a baroclinic system. With an ingenious yet simple assimilation system, built around a quasi-geostrophic model, he was able to demonstrate that it is possible to update

the flow at upper levels using surface observations only. By isolating the fastest growing modes of the quasi-geostrophic model and correcting their growth over an assimilation period he simulated almost exactly the functionality of 4D-Var. In so doing he predicted several of the most relevant results presented in this paper. In 4D-Var the effective covariance matrix of background error is modified by the fastest growing modes of the atmosphere (Thépaut et al., 1993; 1996), which, as will be demonstrated, enables flow-dependent correction of the upper levels in response to surface pressure time-sequences. This effect is most important in the vicinity of storms.

The outline of the paper is as follows: First, in Section 2, the availability of observations from frequently reporting surface stations is discussed and the method for selecting (or screening) observations for use in the analysis on an hourly or 6-hourly basis is described. The results from initial experimentation using more frequent observations is described in Section 3. A modified cost function calculation allowing for serial correlation of observation error is developed in Section 4 and results of its filtering properties are presented. The benefits and problems with this approach are discussed. An option to perform the variational quality control (VarQC) by considering all observations from each time sequence jointly is introduced. A performance evaluation based on a series of assimilation and forecast experiments is presented in Section 5 together with a single-observation experiment illustrating the impact of one surface pressure time-sequence on a rapidly developing baroclinic system. A summary and conclusions are given in Section 6.

2. Observation selection

There are many surface stations around the world that report more than once every 6 h. Most assimilation systems, however, use only one report in each assimilation cycle, the length of which is usually 6 h. This has been the case in ECMWF's OI and 3D-Var systems as well as in its first operational 4D-Var system. In this section we outline the observation selection algorithm applied so far within 3D and 4D-Var (see Järvinen and Undén 1997 for details), and demonstrate that a large resource of additional observations from frequently reporting stations is available, that could potentially be used in 4D-Var.

In both 3D and 4D-Var an initial 6-h model integration is carried out, starting from the background fields, to obtain what is called the high resolution trajectory (Rabier et al., 1998a). All the observational data within the assimilation time window are presented to this model integration and their model counterparts are calculated by applying the observation operators to the high resolution trajectory at appropriate time. The difference between an observed value and its model counterpart is loosely called departure from the background or background departure, in the following. In Kalman filter terminology it is also called the innovation vector. These departures are subjected to the background quality control (BgQC, Andersson and Järvinen 1999), which rejects outliers. From the subset of good quality observations some redundant observations are removed so that a unique set of observations is retained for the assimilation.

In 3D-Var multiple reports from the same station are considered redundant. Preference is then given to the report closest to the centre of the assimilation time window. This observation screening configuration is called 3D-screening. In 4D-Var, the trajectory is available hourly and the observations are organized into 1-h time slots, accordingly. Observations are selected within each time slot independently. Preference is in this case given to the observation closest to the centre of the time slot. This configuration is called 4D-screening. The assimilation system has been prepared such that either 3D or 4D-screening can be performed in 4D-Var, in order to enable comparison between 3D and 4D-Var results using similar sets of observations. The first ECMWF operational implementation of 4D-Var used 3D-screening.

The choice of screening method primarily affects the observation types that are reported frequently, e.g., synoptic stations on land (SYNOP) and on sea (SHIP) as well as drifting buoys (DRIBU). The global number of SYNOP surface pressure observations used with 4D-screening is roughly twice the number in a corresponding assimilation using 3D-screening, in the same period. The difference is mainly due to the observations made outside the main (synoptic) observing times. The number of DRIBU observations increases typically by a

factor of 3 with 4D-screening, and the number of TOVS reports increases by about 20%, as the satellite orbits overlap in the polar areas and the hourly thinning procedure (to a density of approximately $1/(250\,\mathrm{km})^2$) retains more reports there (compared to 6 hourly thinning in 3D-screening). The number of observations of other types remains largely unchanged.

3. Initial experimentation with frequent observations

In this section, we present results from the initial assimilation and forecast experiments using asynoptic observations from frequently reporting stations, i.e., 4D-screening in 4D-Var. An example will show that these additional observations improved the analysis of rapidly developing small-scale synoptic systems. With up to 6 observations over 6 h from many surface stations, however, the assimilation became more sensitive to isolated biased observations which, as we will demonstrate, had a detrimental impact in the Southern Hemisphere.

3.1. *A rapidly developing small-scale synoptic system*

The additional asynoptic observations introduce information on the temporal variation of surface pressure, temperature, humidity and wind. These variations are usually strongest in association with relatively small- scale, rapidly moving, synoptic systems, such as intense baroclinic developments and tropical cyclones. The analysis of such systems is especially difficult if the development takes place over the relatively data sparse oceans. One such case took place on Christmas Eve 1997 when the western coast of Ireland was hit by a vicious winter storm. The operational ECMWF assimilation did not capture this development well – the analysed low was not deep enough. The analysis for 12 UTC 24 December 1997 is shown in Fig. 1 for the North-Eastern Atlantic area. The analysed surface pressure at the centre of the storm is 983.5 hPa. This is about 5 to 10 hPa higher than a subjective surface analysis would suggest (A. Persson, personal communication) – supported by an observation from Valentia observatory on the south western tip of Ireland (Fig. 1) which reported 977.8 hPa.

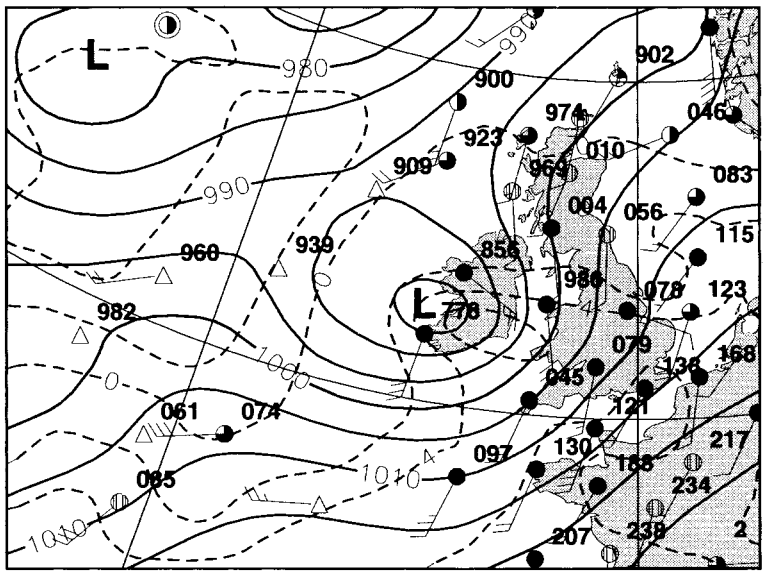

Fig. 1. Operational ECMWF analysis of surface pressure (solid isolines) and of 850 hPa temperature (dashed isolines) for North Eastern Atlantic at 12 UTC 24 December 1997. Selected surface observations of 10-m wind, surface pressure and total cloudiness are plotted with standard notations. SYNOP(DRIBU) observations are marked with a station ring (triangle).

The limited spatial resolution of the analysis increments is an obvious deficiency of a global assimilation system when analysing rapidly developing small-scale features (Lönnberg 1988; Rabier et al., 1998b; Derber and Bouttier 1999). The 4D-Var minimization problem is solved using the incremental minimization method (Courtier et al., 1994) at operational T213/T63 resolution. The analysis increments are computed at T63 resolution and added to the full T213 model fields (as explained in detail by Courtier et al., 1998). Early results from experiments using 4D-Var at T213/T106 resolution suggest that the analysis for the Irish Christmas Eve storm is greatly improved by enhanced analysis resolution. At the moment this is a costly option, however, and the relevant question is whether the T213/T63 analysis could be improved by using more frequent observations.

The operational 4D-Var assimilation system at the time applied 3D-screening of observations and there was a large number of asynoptic SYNOP/SHIP and DRIBU (drifting buoy) reports in the area of the developing storm that remained unused as they were considered redundant. Timing information of relevance to the development of the storm may therefore have been lost. Poor background fields have also caused more BgQC rejections for the wrong reason – not because observations were wrong but because the background field was inaccurate (Hollingsworth et al., 1986).

A data assimilation experiment was run to study the effect of additional asynoptic observations on this case. The assimilation system used in the experiment is identical to the operational 4D-Var except that 4D-screening of observations was applied from 00 UTC 23 December 1997 onwards, i.e., seven analysis cycles before the storm hit the west coast of Ireland. Table 1 shows the number of SYNOP/SHIP and DRIBU surface pressure observations available in the area of the storm over a period of 3 days as well as the number actually used in the case of 3D and 4D-screening, respectively. Table 1 shows that 4D-screening uses the majority of observations, whilst 3D-screening makes use of less than one third in this particular area. There is also a considerable increase in the number of used ten-metre wind observations (not shown). Note also that not nearly all DRIBU observations have been used even with 4D-screening, because they often report even more frequently than once per hour.

The additional surface observations do improve the analysis and the forecast of the storm, at T213/T63 resolution. The effect of these observations can be seen in Fig. 2 where the difference in surface pressure between the experiment and operations is displayed. Negative values (dashed isolines) indicate lower surface pressure in the experiment. The experiment's surface pressure background field at the time of the storm is about 3.2 hPa lower to the west of Ireland and some 2.0 hPa higher to the south-east of Ireland (Fig. 2a). The analysed surface pressure in the experiment is about 3.2 hPa lower just west of Ireland, and there is a small area of about 1.0 hPa higher pressure around the eastern part of Ireland (Fig. 2b). The conclusion is that the additional asynoptic observations have improved the analysis locally. It is not only the analysis that is improved but also the background is more accurate, although the depression is still too far to the west in the background (compared with Fig. 1). The 4D-Var analysis with the enhanced sets of observations manages to correct it eastward. The analysis difference extends upwards through the troposphere with a westward tilt and weakens so that at the 300 hPa level the difference is approximately one quarter of the surface value in terms of geopotential height (not shown).

Another visualization of the improved background is given in Fig. 3, where a time series of the observed surface pressure at a selected location

Table 1. *Number of available and used surface pressure observations in the area of the Christmas Eve storm (47 to 57°N, 25 to 5°W) during the period from 23 to 25 December 1997 for 3D and 4D-screening, respectively*

Observation type	Available	Used with 3D-screening	Used with 4D-screening
SYNOP land	2324	402 (17%)	2175 (94%)
SYNOP ship	635	192 (30%)	547 (86%)
DRIBU, drifting buoys	524	63 (12%)	316 (60%)

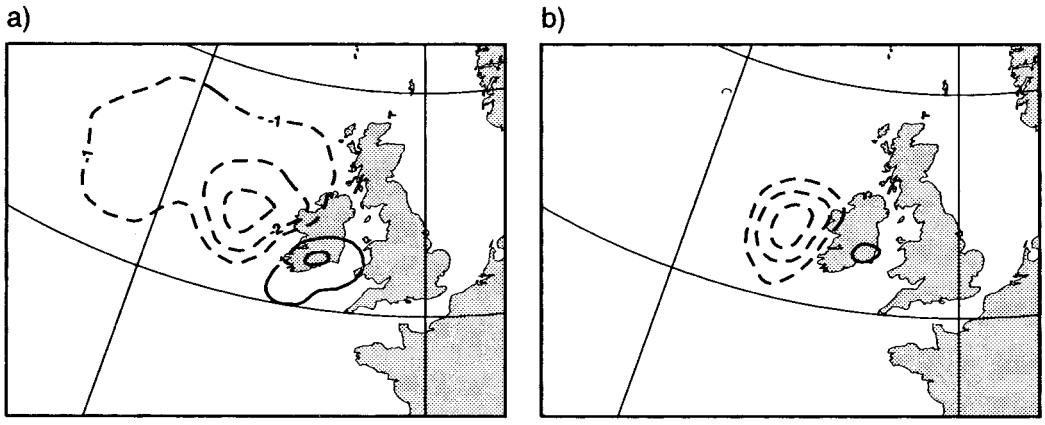

Fig. 2. Surface pressure difference between the experiment and the operational fields as seen in the background (a) and in the analysis (b), respectively, for 12 UTC 24 December 1997. Negative values (dashed isolines) indicate lower surface pressure in the experiment. Contour interval is 1 hPa. The area is the same as in Fig. 1.

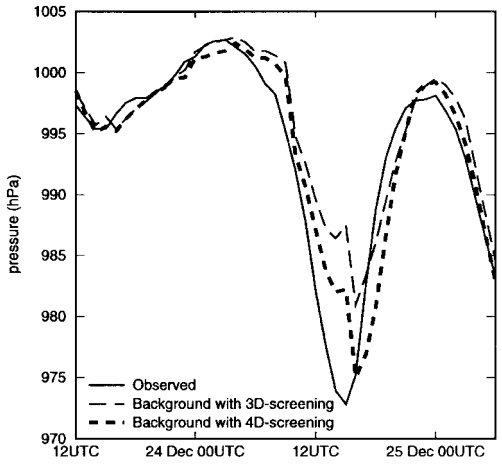

Fig. 3. Observed surface pressure at Belmullet in Ireland (54.14°N, 10.00°W) around the time of the Christmas Eve storm in 1997 (solid line), background surface pressure of the ECMWF operational 4D-Var assimilation system using 3D-screening (long dashed line) and the background surface pressure from a 4D-Var assimilation experiment using 4D-screening (thick dashed line).

(Belmullet) on the Irish West coast (solid line) is plotted together with the background from the operational system (long dashed line) and from the experiment (thick dashed line). The operational background field at the time of deepest depression is about 8 hPa too weak, whereas the experimental background is much closer, only about 2 hPa too weak, both occurring some 4 h later than observed, though. Most of the time, the background follows the observed surface pressure more closely in the experiment than in operations.

The conclusion of these experiments is that by using more of the frequently available observations (at operational resolution T213/T63) it is possible to improve the analysis and short range forecast, i.e., background, of a rapidly developing synoptic system. The surface analysis has become more realistic. The analysis of the studied case can also be improved by increasing the resolution of the analysis increments, as suggested by T213/T106 results (not shown).

3.2. Isolated biased observations

Before deciding on the final configuration of 4D-Var for operational implementation (Rabier et al., 1999), the impact of using more frequent observations was tested in assimilation and forecast experiments. The resulting forecast scores were neutral or slightly positive in the Northern Hemisphere and clearly negative in the Southern Hemisphere. The poorer Southern Hemisphere forecast scores turned out to be due to biased time sequences of surface pressure observations from isolated stations. The decision was taken then to continue 4D-Var experimentation with the same selection of observation as in 3D-Var, i.e., applying the 3D-screening in 4D-Var, until the problems were solved. The experiments are described in more detail in this section.

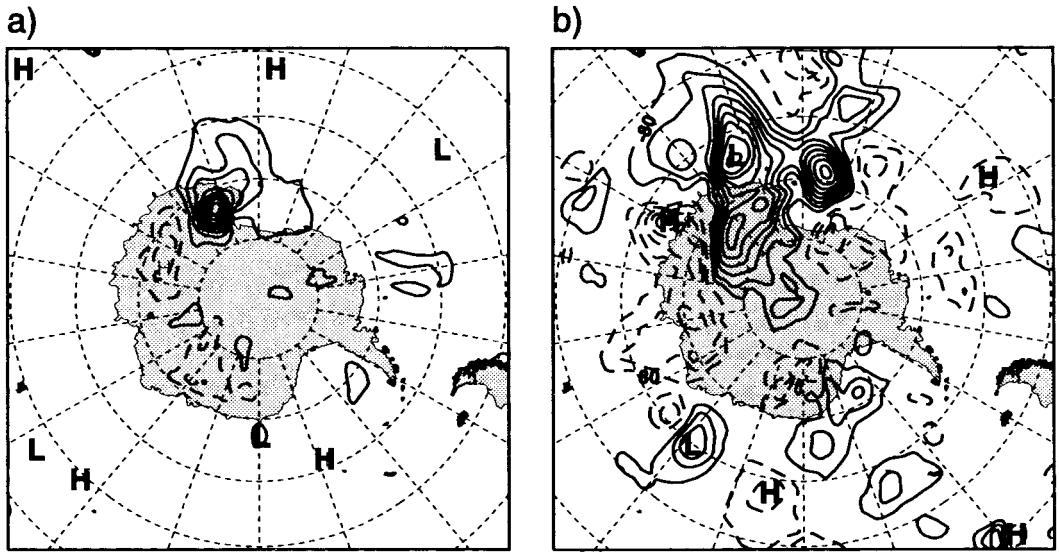

Fig. 4. 850 hPa geopotential difference between two 4D-Var experiments: 4D3S minus 4D4S, in the Antarctic region: (a) is for the analysis on 12 UTC 3 February 1997 and (b) for the subsequent 48 h forecast. The contour interval is 80 $m^2 s^{-2}$ The maximum difference in (b) is roughly 700 $m^2 s^{-2}$.

3.2.1. Forecast results. The impact of using more frequent observations (4D-screening) was studied in a 2 week test period in January 1996 with 3 different assimilation systems: 3D-Var with 3D-screening (3D3S), 4D-Var applying 3D-screening (4D3S) and 4D-Var applying 4D-screening (4D4S). In the Northern Hemisphere both 4D-Var systems produced equally good forecast scores. In the medium range both 4D-Var systems scored 6 to 12 h better than 3D-Var both at 1000 and 500 hPa (not shown). The Southern Hemisphere scores were however best for the 4D3S system, i.e., for the one without the additional observations. There the inclusion of asynoptic observations deteriorated the 4D-Var forecast performance to the level of, or even below, the 3D-Var (not shown). Another set of experiments was performed for a period in February 1997. After 1 week of assimilation it became clear that 4D4S again performed systematically worse than 4D3S.

3.2.2. Investigation of the forecast failures over the Antarctic region. The forecast scores in the 4D-Var experiments applying 4D-screening are particularly poor over the Antarctic region. The analysis difference between 4D3S and 4D4S in terms of 850 hPa geopotential for 12 UTC 3 February 1997 is shown in Fig. 4a. This is the 3rd day of the experiment. The difference pattern is a very localized one. In the subsequent forecast the pattern drifts to a synoptically active region. An interaction with travelling synoptic waves takes place leading to propagation and an increase of the difference. The 48-h forecast 850 geopotential difference is shown in Fig. 4b. The associated hemispheric forecast scores are poor throughout the forecast range, most notably after 2 days (not shown).

The localized pattern of analysis differences suggests that a few stations at the edge of the Antarctic plateau may have caused this large analysis difference. Investigations into the observation departure statistics revealed certain stations in this area with significant biases against the background for all times of the day. When time sequences (of up to 6 observations) were used, large analysis increments were produced in the vicinity of those stations. For isolated observations, particularly in the Southern Hemisphere, there were no mechanisms to prevent these unrealistic increments to appear and to develop into forecast errors. In the Northern Hemisphere, in contrast, there are usually neighbouring observa-

tions which constrain the analysis and the effect of any biased observations tends to be less pronounced.

3.2.3. Discussion. The bias in the Antarctic case is partly due to a mismatch between the true and model orographies. Extrapolation of the model atmosphere into the model orography is needed to calculate the model counterpart of the observed quantity. The subterranean extrapolation is performed applying a standard atmospheric lapse rate for temperature and humidity which results in a systematic difference in some cases. The correct way of handling these particular SYNOP stations could involve a bias correction scheme, or improvements of the observation operator itself. The latter possibility is outside the scope of this paper, and a bias correction scheme (for selected stations) has also not been pursued in this context as it would be logistically difficult and it cannot be assumed that the biases remain steady in time.

Another possible way of addressing the bias problem is to allow for serial correlation of the observation errors. The effect of such correlation would be to draw the analysis more towards the temporal information contained in the time sequences, than towards the time-mean observed value, providing the model for temporal observation error correlation is suitably chosen. This idea will be further developed in the next section.

In summary, the results of the initial experimentation using more frequent observations are positive with respect to the 4D-Var analysis of small scale synoptic systems, but it is prone, in its plain formulation, to be adversely affected by biased stations in data sparse areas.

4. Serial observation error correlation: formulation

In this section, the concept of serially correlated observation errors will be developed in an effort to reduce the negative impact from biases in time sequences. This involves changes to the way the observation term of the cost function is calculated for these observations. Changes will also be introduced to the variational quality control so that it can be applied simultaneously to all observations of a time sequence. A proof that the "forced adjoint technique" of, e.g., LeDimet and Talagrand (1986) applies also in the case of serial correlations is given in the Section 8.

4.1. Observation cost function

In previous presentations (Rabier et al., 1998a) the 4D-Var cost function has been written as a sum of separate cost function contributions, one per time slot. With serial correlation of observation error, this formulation must be generalized.

4.1.1. Standard expression. The standard expression for the 3D/4D-Var observations cost function (Lorenc, 1986) is:

$$J_o = \frac{1}{2}(\boldsymbol{y} - H\boldsymbol{x})^T \boldsymbol{R}^{-1}(\boldsymbol{y} - H\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{y}$ is the array of observations, with error covariance matrix $\boldsymbol{R}$, $\boldsymbol{x}$ is the model state and H the observation operator. The matrix $\boldsymbol{R}$ contains measurement error plus error of representativity. The expression degenerates in the case of uncorrelated observations to a sum of individual J_{oi} contributions (one for each observed datum), i.e.,

$$J_o = \sum_i J_{oi} = \sum_i \frac{1}{2}\left(\frac{y_i - H_i\boldsymbol{x}}{\sigma_{oi}}\right)^2 \tag{2}$$

with σ_{oi} the observation error standard deviation and i an index to each observed datum. In practice on a parallel computer, the contribution from each observation is stored in memory and then summed up in a predetermined sequence, in order to get reproducible results irrespective of how the observations have been distributed across the processors.

4.1.2. Serially correlated observation error. Assume $\boldsymbol{z}$ is the vector of n correlated observations with elements $(y_i - H_i x_i)/\sigma_{oi}$, i.e., the normalized departures.

$$\boldsymbol{z} = \begin{bmatrix} \dfrac{y_1 - H_1 x_1}{\sigma_{o1}} \\ \cdots \\ \dfrac{y_n - H_n x_n}{\sigma_{on}} \end{bmatrix}.$$

The observation cost function is then given by

$$J_o = \frac{1}{2}\boldsymbol{z}^T \boldsymbol{O}^{-1}\boldsymbol{z}. \tag{3}$$

The gradient with respect to x of eq. (3) is

$$\nabla_x J_o = -\boldsymbol{H}^* \boldsymbol{S}\boldsymbol{O}^{-1}\boldsymbol{z}, \tag{4}$$

where $\boldsymbol{O}$ is the $n \times n$ matrix of observation error correlations. $\boldsymbol{S}$ is a diagonal matrix with elements $1/\sigma_{oi}$, and $\boldsymbol{H}^*$ is the adjoint of the tangent linear observation operator $\boldsymbol{H}$. The calculation of $\boldsymbol{O}^{-1}\boldsymbol{z}$ is performed by solving the linear system of equations:

$$\boldsymbol{O}\tilde{\boldsymbol{z}} = \boldsymbol{z} \tag{5}$$

for $\tilde{\boldsymbol{z}}$ using the Choleski decomposition. $\tilde{\boldsymbol{z}} = \boldsymbol{O}^{-1}\boldsymbol{z}$ is sometimes called the effective departure. Once we have $\tilde{\boldsymbol{z}}$, J_o becomes $J_o = (1/2)\boldsymbol{z}^T\tilde{\boldsymbol{z}}$ and $\nabla_x J_o = -\boldsymbol{H}^*\boldsymbol{S}\tilde{\boldsymbol{z}}$. The array $\tilde{\boldsymbol{z}}$ is calculated in the forward branch of the observation operators. For efficiency, it is saved temporarily in the observation array (one number per observed datum) to be re-used (rather than recalculated) in the adjoint.

4.1.3. Incremental formulation, 6-h 4D-Var over several time slots. In the incremental formulation (Courtier et al., 1994), the normalized departure, $\boldsymbol{z}$, is computed from increments δx at low resolution (currently T63), using the tangent-linear observation operators, $\boldsymbol{H}$

$$\boldsymbol{z} = (\boldsymbol{d} - \boldsymbol{H}\delta\boldsymbol{x})\boldsymbol{S} \tag{6}$$

with the innovation vector $\boldsymbol{d}$ computed from the high resolution trajectory $\boldsymbol{x}^{HR}$ using the full non-linear observation operators, H

$$\boldsymbol{d} = \boldsymbol{y} - H\boldsymbol{x}^{HR}. \tag{7}$$

Integration of the tangent linear model, $\boldsymbol{M}$, gives $\delta\boldsymbol{x}_t$ for each (hourly) time slot as follows

$$\delta\boldsymbol{x}_t = \boldsymbol{M}_t\delta\boldsymbol{x}_{t-1}. \tag{8}$$

In practice $\boldsymbol{H}$ has been split in a horizontal and a vertical part, $\boldsymbol{H} = \boldsymbol{H}_v\boldsymbol{H}_h$. The horizontal interpolation of model profiles to observation locations $\boldsymbol{H}_h$ is performed once per time slot, and the results $\boldsymbol{H}_{h,t}\delta\boldsymbol{x}_t$ for all time slots $t = [0, T]$ are stored in memory. In operational 6-h 4D-Var $T = 6$. Introducing the symbol $\overline{\delta\boldsymbol{x}_t} = \boldsymbol{H}_{h,t}\delta\boldsymbol{x}_t$ for brevity, we concatenate all time slots to form the vector $\overline{\delta\boldsymbol{x}}$:

$$\overline{\delta\boldsymbol{x}} = \begin{bmatrix} \overline{\delta\boldsymbol{x}_0} \\ \overline{\delta\boldsymbol{x}_1} \\ \cdots \\ \overline{\delta\boldsymbol{x}_T} \end{bmatrix}.$$

The vertical part of the operator, $\boldsymbol{H}_v$, can then be applied to all observations at once, irrespective of time slot. This leads to longer vector-loops, and higher efficiency on computers with vector processors. The normalized departures are thus computed

$$\boldsymbol{z} = [\boldsymbol{d} - \boldsymbol{H}_v\overline{\delta\boldsymbol{x}}]\boldsymbol{S}. \tag{9}$$

In the case of non-zero observation error correlation $\tilde{\boldsymbol{z}}$ is calculated solving eq. (5), otherwise $\tilde{\boldsymbol{z}} = \boldsymbol{z}$, and $J_o = (1/2)\boldsymbol{z}^T\tilde{\boldsymbol{z}}$ as before.

We have seen that the forward vertical operators are applied once only, across all time slots, and that the horizontal operators are applied once per time slot. The same applies to the adjoint, but in the opposite order. The adjoint of the vertical operators $\boldsymbol{H}_v^*$ are applied first, to get the gradient of observation cost function with respect to $\overline{\delta\boldsymbol{x}}$, $\nabla_{\overline{\delta x}} J_o$:

$$\nabla_{\overline{\delta x}} J_o = -\boldsymbol{H}_v^*\boldsymbol{S}\tilde{\boldsymbol{z}}. \tag{10}$$

The gradient is then organized by time slot, i.e.,

$$\nabla_{\overline{\delta x}} J_o = \begin{bmatrix} \nabla_{\overline{\delta x_0}} J_o \\ \nabla_{\overline{\delta x_1}} J_o \\ \cdots \\ \nabla_{\overline{\delta x_T}} J_o \end{bmatrix},$$

and the adjoint of the horizontal interpolation, $\boldsymbol{H}_h^*$ is applied once per time slot, to form $\nabla_{\delta x} J_o$

$$\nabla_{\delta x} J_o = \begin{bmatrix} \boldsymbol{H}_{h,0}^* \nabla_{\overline{\delta x_0}} J_o \\ \boldsymbol{H}_{h,1}^* \nabla_{\overline{\delta x_1}} J_o \\ \cdots \\ \boldsymbol{H}_{h,T}^* \nabla_{\overline{\delta x_T}} J_o \end{bmatrix}.$$

The gradient is transported backwards in time, recursively applying the adjoint of the model (the adjoint of eq. (8))

$$\nabla_{\delta x_{t-1}} J_o = \boldsymbol{M}_t^* \nabla_{\delta x_t} J_0 \tag{11}$$

which finally yields $\nabla_{\delta x_0} J_o$, the gradient at initial time.

4.1.4. Introducing serial correlation. When we introduce serial correlation of observation errors, the application of $\boldsymbol{O}$ in eq. (5) will involve observations from several time slots. The normalized departures are calculated using eq. (9), as before. Once the effective departures $\tilde{\boldsymbol{z}}$ are stored in the observation array, the adjoint calculations can

proceed as normal, without any modification, using eq. (10) and eq. (11).

With serial correlation the gradient for one time slot will depend on observations not just from the same time slot, but from all time slots. The forcing of the adjoint model (at a given time) thus depends on observations from all time slots: past, present and future. This may at first seem counter-intuitive, but it is formally correct (see Section 8). A more straight-forward interpretation is that the serial correlations filter, in time, the "signal" from the observations. The unfiltered signal is $\boldsymbol{z}$ and the filtered signal is $\tilde{\boldsymbol{z}}$.

4.2. Specifying serial correlations

We do not know of a practical method to obtain estimates of the serial correlation of observation error. Statistics of background departures are dominated by correlations in the background error. For now we have adopted a simple ad hoc Gaussian model for the serial correlation of SYNOP/SHIP and DRIBU surface pressure and height observations. The correlation r between 2 observations at times t_1 and t_2 is thus:

$$r = a \exp\left[\frac{-(t_1 - t_2)^2}{\tau^2}\right] + (1 - a)\delta_{t_1 - t_2}, \qquad (12)$$

with an e-folding time τ of 6 h and $a = 0.3$. δ is the Kronecker delta.

An eigenvalue decomposition of a 6×6 correlation matrix of observations 1 h apart, using eq. (12) (Table 2, first entry), showed that the effect of the serial correlation is to reduce the weight given to the mean of the 6 observations (equivalent to increasing the observation error standard deviation for the mean by a factor 1.50), and to increase the weight given to higher modes (see Subsection 4.8 of Daley, 1991). The first eigenvector corresponds to the 6-h mean, the second to the mean 6-h tendency and the third to a pressure drop followed by a rise (or vice versa) within the 6-h period. The different entries in the table show results for variations in the parameter values τ and a, and in the functional form for r. Form A uses eq. (12) and B uses an exponential $r = a \exp(-|t_1 - t_2|/\tau) + (1 - a)\delta_{t_1 - t_2}$ for the correlation function. We concluded that the form of the correlation function is not too important. Even a constant correlation of 30% between all 6 observations (4th entry in Table 2) would give the desired effect.

Table 2. *Eigenvector decomposition of a 6×6 correlation matrix with observations 1 h apart*

Form	Parameter values	Ev1	Ev2	Ev3
A	$\tau = 6$, $a = 0.3$	1.50	0.96	0.84
A	$\tau = 6$, $a = 0.6$	1.88	0.92	0.65
A	$\tau = 24$, $a = 0.3$	1.58	0.85	0.84
A	$\tau = \infty$, $a = 0.3$	1.58	0.85	0.84
B	$\tau = 6$, $a = 0.3$	1.43	0.98	0.89
B	$\tau = 24$, $a = 0.3$	1.54	0.89	0.85

The table shows the eigenvalue for the first three eigenvectors (Ev) for different parameter values τ and a. Form A uses eq. (12) and B uses an exponential form (see main text) for the correlation function. The eigenvalues for Ev4, 5 and 6 are within 2% of those for Ev3.

4.3. Joint VarQC of time sequence of observations

The variational quality control (VarQC) for the time sequences of SYNOP/SHIP and DRIBU surface pressure and height observations was modified such that the check is applied jointly for all observations in the time sequences. One joint quality control decision is computed which then applies equally to all data in the time sequence. This is similar to the VarQC formulation for u and v-component winds, which are also checked jointly. The theory and application of VarQC have been described by Ingleby and Lorenc (1993) and Andersson and Järvinen (1999).

4.3.1. Formulation. The probability density function (p.d.f.) with VarQC p^{QC} is expressed as a sum of 2 terms: one representing good observations, modelled by a Gaussian N and one representing observations with gross errors, often modelled by a flat p.d.f. F. A is the a-priori probability of gross error.

$$p^{\mathrm{QC}} = (1 - A)N + AF. \qquad (13)$$

For n (maximum 6 in this case) independent observations, we multiply the p.d.f.s for each observation, to obtain the combined p.d.f.:

$$p^{\mathrm{QC}} = \prod_{i=1}^{n} [(1 - A_i)N_i + A_i F_i]. \qquad (14)$$

For consistency with the assumption of serial error correlation, we shall assume that also the gross errors are not independent. If we assume that one

error source affects all observations in the time sequence we may write the a-priori probability that all observations in the time sequence are correct, $1 - A_{ts}$:

$$1 - A_{ts} = \prod_{i=1}^{n} (1 - A_i), \tag{15}$$

we can then write:

$$p^{QC} = (1 - A_{ts})N^n + A_{ts} \prod_{i=1}^{n} F_i, \tag{16}$$

where N^n is an n-dimensional Gaussian with correlation matrix $\boldsymbol{O}$ and observation errors as in Subsection 4.1, and F is a flat distribution (representing the observations with gross errors). The VarQC-modified cost function is $J_o^{QC} = -\ln p^{QC}$.

The a posteriori probability of gross error P_{ts} is then, by construction, the same for all observations in the time sequence and given by

$$P_{ts} = \frac{\gamma_{ts}}{\gamma_{ts} + \exp[-J_o]}, \tag{17}$$

where γ_{ts} is

$$\gamma_{ts} = \frac{A_{ts}(\sqrt{2\pi})^n}{(1 - A_{ts})\left(\prod_{i=1}^{n} 2d_i\right)}, \tag{18}$$

and d_i is the width of the flat p.d.f.

The gradient of the cost-function with VarQC, expressed with respect to the normal cost function without VarQC, is finally:

$$\nabla J_o^{QC} = \nabla J_o (1 - P_{ts}). \tag{19}$$

Eq. (19) shows that time sequences that are found likely to be incorrect ($P_{ts} \approx 1$) are given reduced weight in the analysis. Conversely, time sequences that are found to be correct ($P_{ts} \approx 0$) are given the weight they would have had using purely Gaussian observation error p.d.f. The value of P_{ts} is re-calculated every iteration of the main 4D-Var minimisation, hence observations may gradually regain (or lose) influence on the analysis during the course of minimisation, due to neighbouring observations for example.

5. Experimentation with the modified assimilation system

In this section, we investigate to what extent the modified system can alleviate the detrimental impact of isolated biased time sequences, identified as the main problem in the initial set of experiments. For ease of implementation we apply the changes equally, to all stations globally. The performance of the modified assimilation system is assessed in assimilation and forecast experiments.

5.1. *Bias correction properties of the modified system*

We have seen that the standard 4D-Var using 4D-screening (4D4S) is more sensitive to biased isolated stations than 4D3S is. The importance of the 2 modifications developed in the preceding section in reducing the negative impact of biased time sequences was investigated in a set of three 3-day assimilations. In a first experiment serial observation error correlation was activated separately (4D4S_SC). In a second experiment joint VarQC was activated separately (4D4S_JVQC), and in a third experiment both were activated simultaneously (4D4S_SC_JVQC).

Mean 850 hPa geopotential analysis differences against 4D3S over the 3 days (1–3 February 1997), for the Antarctic region, are displayed in Fig. 5, such that in the areas of positive values (solid isolines) the experiment has higher values of 850 hPa geopotential than 4D3S. Fig. 5a shows 4D4S minus 4D3S (which is rather similar to Fig. 4a). In the next panel (Fig. 5b) we see that the experiment 4D4S_SC brings only a small improvement compared to Fig. 5a. The effect from the VarQC modification alone (4D4S_JVQC) also leaves large mean analysis differences (Fig. 5c), but 4D4S_JVQC is slightly better than 4D4S_SC. The combined modification (joint VarQC and serial correlation of observation errors, experiment 4D4S_SC_JVQC) gives the best result (Fig. 5d). The combined effect is larger than the sum of the effects from the 2 separate modifications. This is because joint VarQC in the last experiment is applied to the filtered time sequences, which results in more rejections of biased observations than in the other experiments. The combination of modifications introduced in Section 4 can thus be thought of as a conservative and targeted bias correction scheme, removing a large part of the impact on the analysis of the biased time sequences of observations.

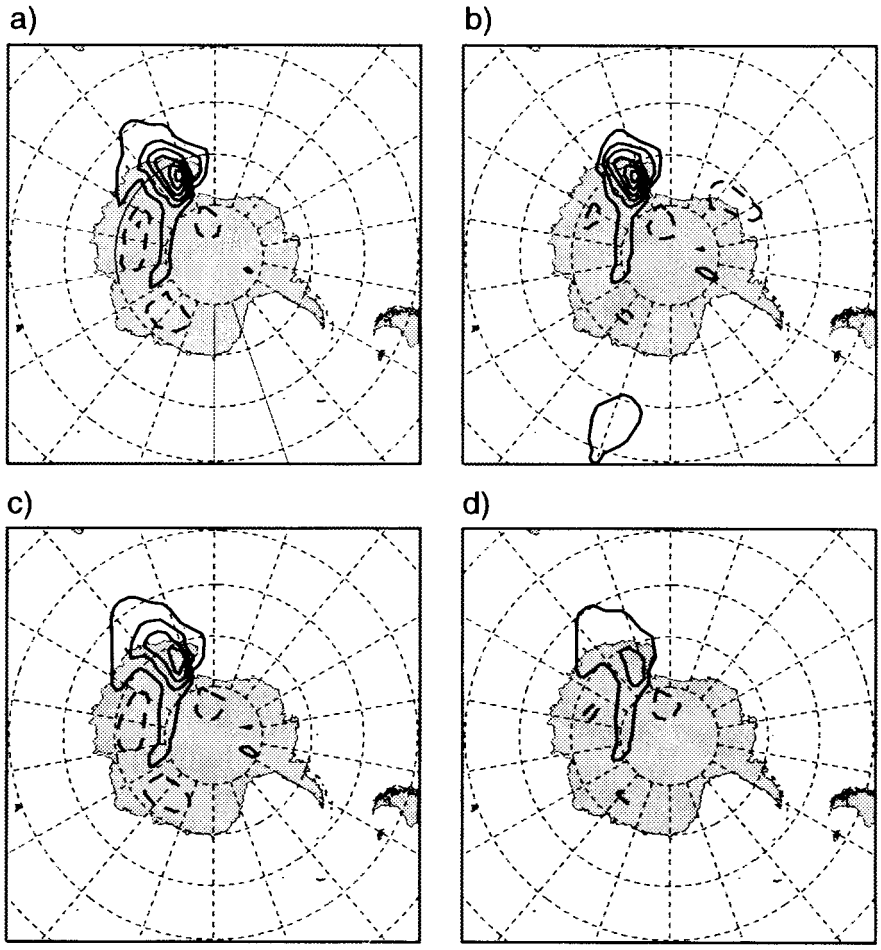

Fig. 5. Mean 850 hPa geopotential analysis differences between 4D3S and four different 4D-screening experiments: 4D4S minus 4D3S (a), 4D4S_SC minus 4D3S (b), 4D4S_JVQC minus 4D3S (c) and 4D4S_SC_JVQC minus 4D3S (d), respectively, during the period of 1–3 February 1997. The area is the same as in Fig. 4. The contour interval is 80 $m^2 s^{-2}$.

5.2. *Departure statistics*

We now turn to a more general evaluation of the modified 4D4S system, in terms of quality of the background and analysis with respect to observations. The mean and standard deviation of background and analysis departures are given in Table 3 for SYNOP/SHIP and DRIBU surface pressure and ten-metre wind observations over the North Atlantic for a 2-week period in May 1997. The statistics have been calculated for the unmodified system (4D4S) and for the modified system (4D4S_SC_JVQC), the results for the latter experiment are under the heading "exp" in Table 3. In collecting the statistics we have considered only those observations that have been used (i.e., passed all quality control) by both assimilation systems, in order to have truly comparable results.

There have been several significant improvements in the departure statistics. In the modified system there has been a decrease in the standard deviation of background departures (Table 3), which indicates an improvement in the quality of the background. The decrease is statistically significant (F-test) for surface pressure observations. There has also been a statistically significant

Table 3. *Mean and standard deviation of background and analysis departures for* SYNOP/SHIP *and* DRIBU *surface pressure (p_s) and 10-m wind components (u, v) over part of the North Atlantic (30–75°N, 10–60°W) in 4D4S (cntrl) and in the modified system 4D4S_SC_JVQC (exp) for period 15–28 May 1997*

		Background departure				Analysis departure			
		mean		std-deviation		mean		std-deviation	
Variable	Sample	cntrl	exp	cntrl	exp	cntrl	exp	cntrl	exp
DRIBU p_s	2148	8.63	7.44(1)	87.3	79.9(4)	8.53	6.78	67.9	64.9(2)
DRIBU u	499	−0.01	0.03	2.18	2.04	−0.05	−0.02	1.85	1.65(3)
DRIBU v	449	−0.12	−0.14	2.03	1.98	−0.05	−0.05	1.74	1.59(2)
SYNOP p_s	4333	−8.70	−13.01(1)	134.4	130.5(2)	−14.6	−16.4(1)	114.5	114.8
SYNOP u	4002	−0.09	−0.10	3.16	3.13	−0.06	−0.07	3.13	2.82(4)
SYNOP v	4002	−0.15	−0.17	3.11	3.10	−0.10	−0.10	2.80	2.80

A decrease in the mean departures in the modified system at a statistical significance level of 0.5% is denoted by "(1)". A decrease in the standard deviation of departures at statistical significance levels of 5%, 1% and 0.1% are denoted by "(2)", "(3)" and "(4)", respectively. Units are Pa for p_s and are ms^{-1} for u and v.

decrease in the standard deviation of analysis departures, especially for ten-metre u-component wind. There has been a consistent and statistically significant change (t-test for matched pairs) in the mean of background and analysis departures in surface pressure (Table 3), to slightly lower values. This implies that the modifications have increased the mean surface pressure slightly in this North Atlantic area, in the background as well as in the analysis. This is not surprising as the serial observation error correlation has been applied to the surface pressure observations and it is an indication that the observed mean departure has been assimilated slightly less in the experiment.

5.3. *The Irish Christmas storm and single-observation experiments*

The assimilation 4D4S was superior to 4D3S in analysing the Irish Christmas storm, already in the initial set of experiments. The same experiment as in Subsection 3.1 was repeated using the modifications of Section 4. We found that the surface pressure analysis had further deepened to the southwest of Ireland in this new experiment for 12 UTC 24 December 1997 (not shown), which means a further improvement over the standard 4D4S. We studied the Irish storm case further by carrying out single-observation experiments as described in the following.

Single-observation experiments have been used in several previous studies of 4D-Var to illustrate the effect of the dynamics on the 4D-Var analysis increments (Thépaut et al., 1993; 1996, Rabier et al., 1999). Here we use the same technique to study the 4D-Var response to a time sequence of surface pressure data. We have chosen a time sequence from the SYNOP station Malin Head (03980, at 55.37°N, 7.33°W) at the northern tip of Ireland. In the period when the storm was approaching Ireland from the southwest Malin Head reported a rapid pressure fall of 20.1 hPa in 5 h, from 997.1 hPa at 10 UTC to 977.0 hPa at 15 UTC. The background in the same period showed a pressure fall of 15.8 hPa, i.e., an underestimate of 4.3 hPa, indicative of a less intense storm in the background than in reality.

In the first experiment (MH_TS) the whole time sequence of 6 surface pressure observations was used. In the second experiment (MH_12) only the observation at 12 UTC was used. This mimics the observation selection of 4D and 3D-screening, respectively. Given the Malin Head data values, the MH_12 analysis should reduce the surface pressure (at 12 UTC) compared to the background, by up to 1.7 hPa. The MH_TS analysis, on the other hand, should produce a more rapid pressure fall. The resulting surface pressure analysis increments at initial and end time (9 and 15 UTC) are shown for both experiments in Fig. 6. Upper panels show MH_TS and lower panels show MH_12. We can see that both analyses have lowered the surface pressure. MH_TS has a sharp gradient in the increment (at initial time) to the southwest (i.e., upstream) of the observation, whereas the MH_12 increment has a very weak

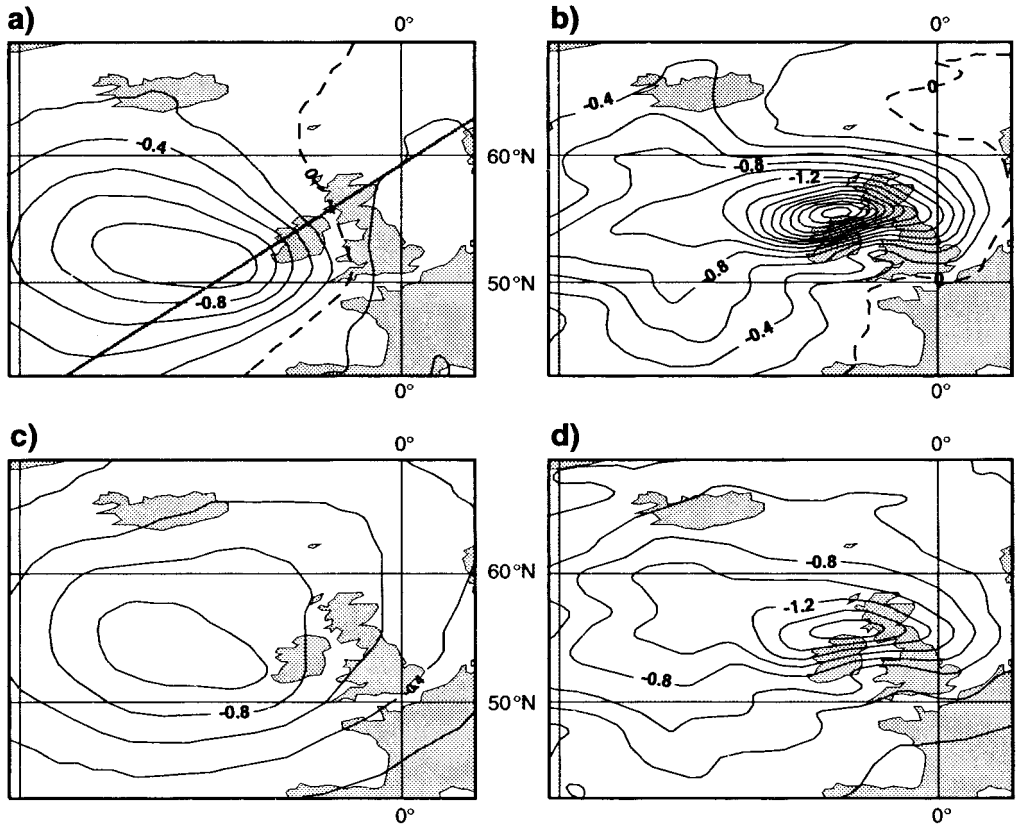

Fig. 6. Surface-pressure analysis increments (hPa) from 4D-Var single-observation experiments at 12 UTC 24 December 1997. Panels (a) and (b) show MH_TS (time sequence) at 9 and 15 UTC, respectively, and (c) and (d) show MH_12 (one observation at 12 UTC, see main text) also at 9 and 15 UTC, respectively. The observation (Malin Head) is near the northern tip of Ireland. The contour interval is 0.2 hPa, with the zero line dashed. The line plotted in (a) indicates the orientation of the cross-sections in Fig. 7.

gradient over Ireland, reducing the pressure in a more wide spread area than MH_TS. The analysis increment produced by MH_TS intensifies rapidly (from −1.1 to −3.1 hPa, in the centre) over 6 h, whereas the MH_12 increment intensifies more modestly (from −1.1 to −1.7 hPa). From the MH_TS result we can conclude that 4D-Var successfully produced an increment at initial time which intensifies rapidly during the 6 h of assimilation, so as to reduce the mis-fit to the observed time sequence of surface pressure at Malin Head. The other experiment (MH_12) which lacks temporal information from the observations, has produced an increment which is more steady in time.

The vertical structure of the background and analysis increments at initial time (9 UTC) is shown in Fig. 7, for both experiments. The panels to the left show potential temperature and those to the right show vertical velocity. The background is the same for both experiments. In Fig. 7b we see that the negative surface pressure increments at the surface are associated with deep negative temperature increments in the troposphere, which change sign at the tropopause. By comparison with panel a) we see that the main increments are within the cold air mass, and that the change in sign corresponds to the location of the actual tropopause in the background. Furthermore, the analysis has also warmed the warm air mass to the south of the cross-section, on the southern and eastern side of the cold front (not shown), thereby strengthening the frontal temperature gra-

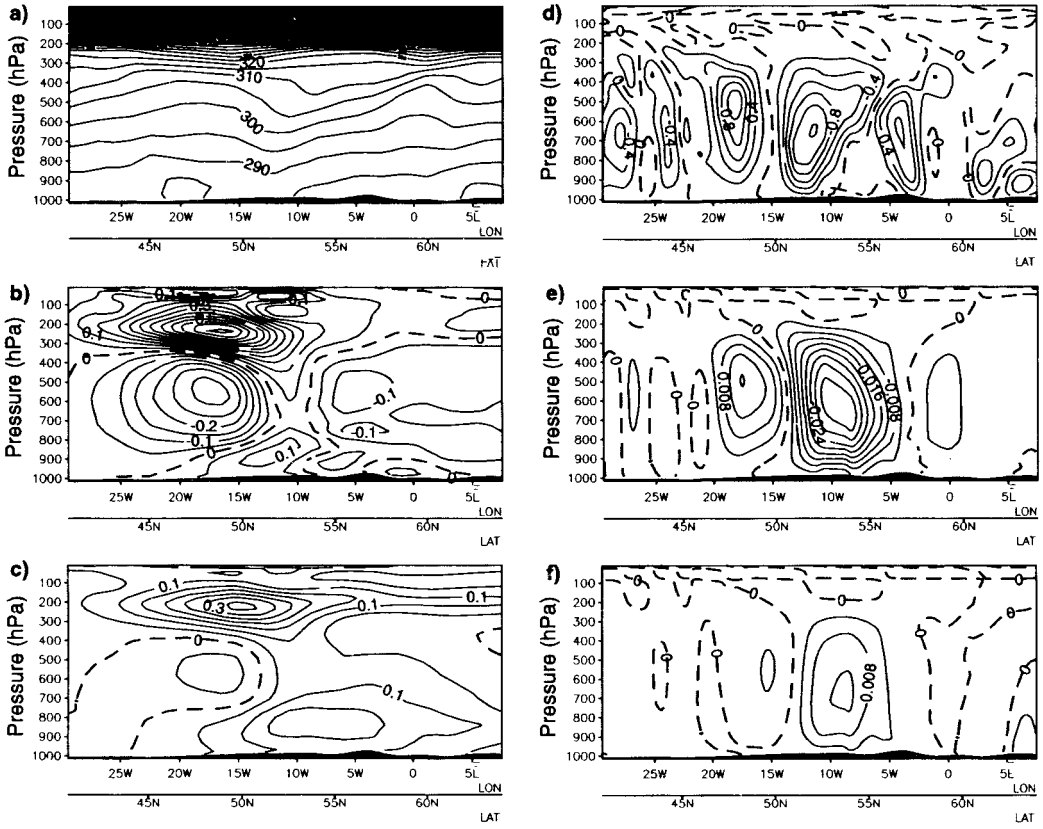

Fig. 7. Results of 4D-Var single-observation analysis experiments at 12 UTC 24 December 1997, using surface pressure observations from Malin Head (SYNOP 03980, at 55.37°N, 7.33°W), only. The panels to the left show potential temperature (K) and the panels to the right show vertical velocity (Pa s^{-1}): (a) and (d) show the background. Remaining panels show analysis increments: (b) and (e) using the full time sequence of hourly observations (i.e., 6 surface pressure observations) with serial correlation, and (c) and (f) using the Malin Head observation at 12 UTC, only. The increments refer to the initial time of the 4D-Var assimilation window, i.e., 9 UTC. The contour intervals are 5 K (a), 0.05 K (b) and (c), 0.2 Pa s^{-1} (d) and 0.004 Pa s^{-1} (e) and (f), with zero lines dashed. The location of the cross-section is as indicated in Fig. 6a.

dient. Experiment MH_12 (Fig. 7c) shows a similar structure in temperature increments, with less amplitude. In terms of vertical velocity, we see by comparing panels (e) with (d) that both the upward (negative) and downward (positive) motion associated with the approaching low has been intensified by the analysis increments of MH_TS. Again, MH_12 shows a similar structure (panel f) to MH_TS, but far weaker.

Note that the maximum temperature increments, created by the time sequence of surface pressure observations at Malin Head, are located ten degrees to the west (upstream) of the station. In a static scheme (e.g., 3D-Var) temperature increments would have their maximum directly above the station location. Their vertical structure would follow the shape of the specified T-ps cross-correlations, i.e., a local positive maximum in the lower troposphere and second positive maximum in the lower stratosphere (see Rabier et al., 1998b, their Fig. 9b and the discussion in Section 5 of Andersson et al., 1998). The actual increments in panel (b) are therefore far from what would be expected from the background term alone, which is a demonstration of the strong dynamical influence on the 4D-Var increments, in this case.

These single-observation experiments have shown that a pressure time-sequence from a single surface station can intensify the analysis of a mid-latitude baroclinic system, that was underestimated in the background, when used in a 6-h 4D-Var system.

5.4. *Data assimilation impact*

4D-Var with 4D-screening, serial correlation and joint VarQC has been tested in data assimilation and forecast experiments in 3 separate periods in May, November and December 1997, totalling 39 days, see Table 4. It has been compared to 4D3S (3D-screening) and in 2 of the periods standard 4D-screening (4D4S) was run, too.

The impact on analyses is largest in the Antarctic region, in the North Atlantic, the North-East Pacific, over Southern Africa and South America. This can be seen from the r.m.s. of analysis difference between 4D4S_SC_JVQC and its 4D3S control, as shown in Fig. 8. The additional surface observations have had largest impact in these otherwise relatively data sparse areas. With 4D-screening we use more TOVS data in the polar regions where orbits overlap, which may also have contributed to the differences at high latitudes.

Table 4. *List of data assimilation experiments testing the impact of 4D-screening (4D4S) on its own, or together with serial correlation (SC) and joint VarQC*

Period	From	To	No. days
1	15 May 1997	31 May 1997	17
2	11 Nov 1997	24 Nov 1997	14
3	28 Nov 1997	5 Dec 1997	8

The controls are 4D-Var with 3D-screening (4D3S).

The use of time sequences of surface pressure observations improves the accuracy of the background in the assimilation, as previously shown in Fig. 3 and Table 3. If the background has improved significantly then analysis increments should be smaller in the experiment than in the control. Experiments with 4D-screening (alone) have systematically shown a small reduction in 1000 hPa height increments in many areas but have shown a clear deterioration in the Antarctic, as discussed in Subsection 3.2. With the introduction of serial correlation and joint VarQC the situation has improved quite clearly. Fig. 9 shows that the 1000 hPa analysis increments generally are smaller in many areas in both hemispheres.

5.5. *Forecast impact*

The forecast impact is small. It tends to be positive especially at short range: +24 to +96 h,

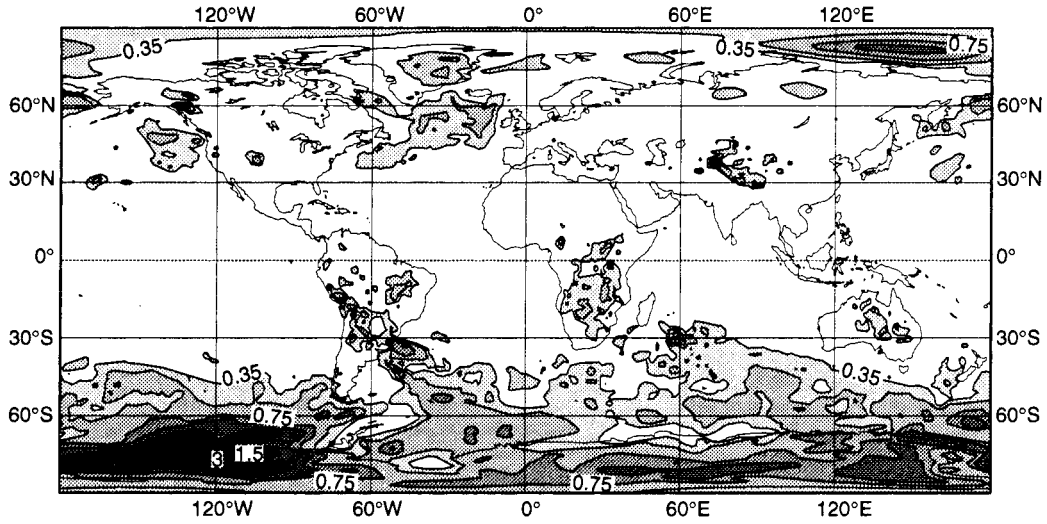

Fig. 8. 1000 hPa geopotential height r.m.s. of analysis difference between the experiment (4D-screening + serial correlation + joint VarQC) and its control (3D-screening), for the period 00 UTC 11 November 1997 to 18 UTC 24 November 1997. The contours are 0.35, 0.5, 0.75, 1.0, 1.5, 2.0 and 3.0 decametres.

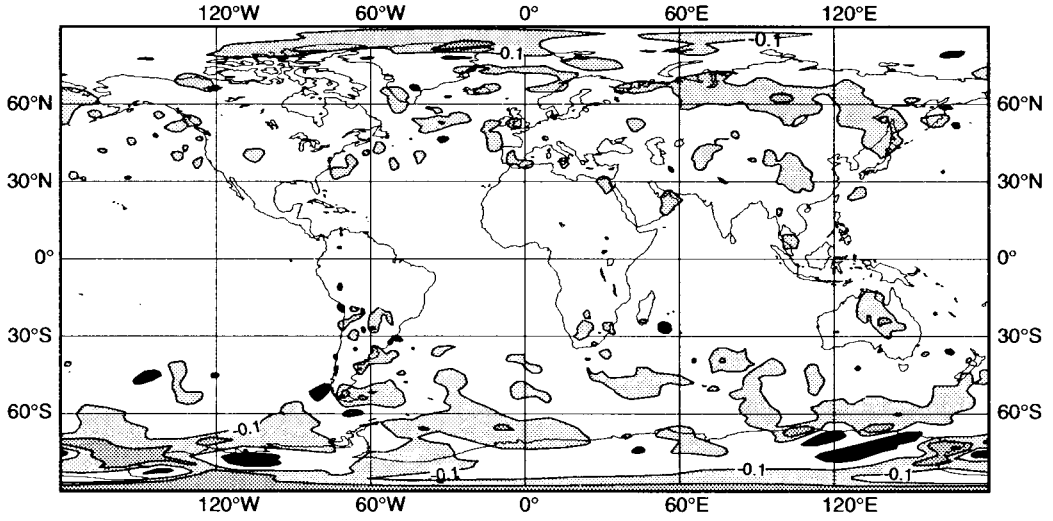

Fig. 9. 1000 hPa geopotential height difference between r.m.s. of analysis increments in the experiment (4D-screening + serial correlation + joint VarQC) and r.m.s. of analysis increments in the control (3D- screening), for the period 00 UTC 11 November 1997 to 18 UTC 24 November 1997. The contours are ± 0.1, ± 0.25 and ± 0.50 dm. Light (dark) shading indicates negative (positive) values, and there is no shading in the interval from -0.1 to 0.1 dm.

in both hemispheres and Europe. Fig. 10 shows the Northern Hemisphere scores only – Southern Hemisphere and Europe are similar. To test the significance of the difference in these forecast scores we have applied the student *t*-test. The results are summarized in Table 5.

It shows that the forecast results are significantly better for Europe at +48 h, Northern Hemisphere at +48, +72 and +96 h, and Southern Hemisphere at +48 h. Of the tested scores (+48, +72, +96 and +120 h at 1000 and 500 hPa) only Europe at 500 hPa +120 h was significantly worse. Although the forecast impact is small, it is systematic enough to be significant. There was no discernible impact on Tropical scores, and no significant impact beyond day 5.

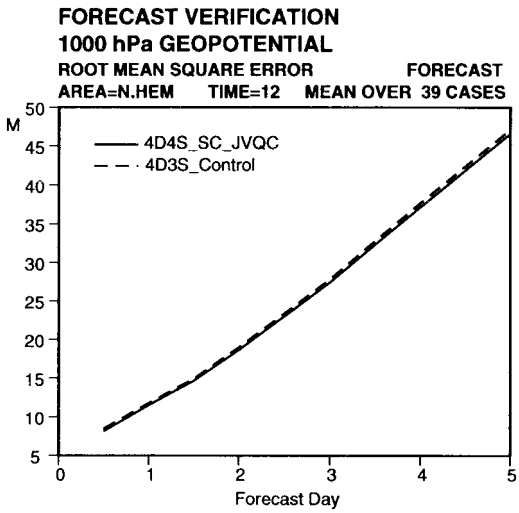

Fig. 10. Average 1000 hPa geopotential height r.m.s. of forecast error (m) for the experiment (4D-screening + serial correlation + joint VarQC) full line, and its control, dashed, for Northern Hemisphere. There are 39 cases in all – see Table 4 for the experiment definitions and periods.

6. Conclusions

Frequently reporting synoptic stations and drifting buoys constitute a resource of observational information which is not always fully utilized in operational numerical data assimilation. In this paper we have presented a method to assimilate these observations in the context of 4D-Var. It allows the assimilation of a large number of surface observations which were neither used in ECMWF's OI and 3D-Var systems, nor in the first operational implementation of 4D-Var.

In the ECMWF assimilation system, there is an

Table 5. *Summary of significance test applied to the difference in forecast scores between the experiments (4D4S_SC_JVQC) and their controls (4D3S) for 39 cases in 3 separate periods (Table 4)*

Area	Level (hPa)	Fc. Step (h)	Verdict	Significance level (%)
Europe	1000	48	BETTER	5.0
	500	48	BETTER	0.1
		72	BETTER	5.0
		120	WORSE	5.0
N. Hem	1000	48	BETTER	0.1
		72	BETTER	1.0
		96	BETTER	5.0
	500	48	BETTER	2.0
		72	BETTER	5.0
S. Hem	1000	48	BETTER	2.0
	500	48	BETTER	5.0

The significance test was carried out for the r.m.s. of forecast error for steps +48, +72, +96 and +120 h, for geopotential height at 1000 and 500 hPa, for Europe and the two hemispheres. Only those scores that are significant at the 5% level (or higher) have been included.

observation screening stage which selects a unique set of good quality observations for use in the 3D or 4D-Var assimilation. The observation screening can be performed either 6-hourly for 3D/4D-Var, or hourly for 4D-Var (sometimes called 3D- and 4D-screening, respectively). The first operational implementation of 4D-Var at ECMWF applied 3D-screening, which enabled an easier comparison with 3D-Var results as the number of observations used in both systems was very nearly the same (Rabier et al., 1999). In 4D-screening time sequences of up to 6 observations per station can be retained for use in the analysis. Globally, 4D-screening retains approximately twice as many SYNOP/SHIP observations and 3 times as many DRIBU observations as used operationally. Also TOVS data over the poles, where the orbits overlap, increase in numbers, somewhat.

With up to 6 observations from each station, 4D-Var becomes vulnerable to isolated observations that exhibit biases relative to the model background fields. This was found to be the case especially over the Antarctic region where some synoptic stations over the high orography have large biases against the model surface pressure. It caused a deterioration of Southern Hemisphere forecast performance. We have addressed the bias problem by introducing serial correlation of observation error and by modifying the variational quality control of the time sequences of SYNOP/SHIP and DRIBU surface pressure and height observations. Serial correlation of observation error of a Gaussian form shifts the emphasis from the mean observed value to the temporal variation in the time sequences. It has not been possible to support the choice of correlation model with statistics. On the other hand, it was demonstrated that there is not a strong sensitivity to the choice of the actual form of the correlation function. The variational quality control was modified to check all observations from the time sequences for each station jointly, so that it rejects or accepts the entire time sequence. These 2 developments successfully solve the initial difficulties with biased isolated stations, and acts as a conservative targeted bias correction scheme removing most of the detrimental effect of the biased time sequences. A more straightforward bias correction scheme of selected stations was not tried in this context as difficulties with its maintenance were foreseen.

The initial experimentation showed that the enhanced observation frequency improved the short range forecast of rapidly developing small scale synoptic systems. A case study of the Irish Christmas Eve Storm in 1997 showed that the use of time sequences improved the background of the assimilation. Single-observation experiments showed that a pressure time-sequence from a single surface station can intensify the analysis of a mid-latitude baroclinic system, that was underestimated in the background, when used in a 6-h 4D-Var. The actual analysis increments in the studied case were far from what would be expected from the background term alone, which is a demonstration of the strong dynamical influence on the 4D-Var increments.

We have shown that the accuracy of the background 1000 hPa height fields have improved in the assimilation experiments using time sequences of observations. A relevant measure for the relative accuracy of the background field is the r.m.s. of the analysis increments, when comparing 2 assimilation systems using the same observations. The more consistent the background is with the observations, the smaller is the r.m.s. of analysis increments. The improvement is largest over mid-latitudes and polar areas.

The forecast impact is small, and yet significantly positive at 1000 hPa for Europe at +48 h, Northern Hemisphere at +48, +72 and +96 h and Southern

Hemisphere at +48 h, for example. The performance for Europe +120 h, at 500 hPa was however slightly (and significantly) worse. The proposed method for using time sequences of observations was, based on these results, introduced operationally in ECMWF 4D-Var in June 1998. The extra computational cost is small at approximately 4% of 4D-Var.

7. Acknowledgements

We are grateful to our colleagues Anthony Hollingsworth, Florence Rabier, Adrian Simmons and Per Undén for valuable discussions and comments. Michael Fisher provided the significance test of paired forecasts used in Table 5 and Anders Persson provided subjective analyses of the Irish Christmas Eve storm. Mats Hamrud advised us on how to implement serial correlations efficiently and Jan Haseler's work facilitated the experimentation. Jocelyn Williams skilfully improved the figures. The helpful comments by 2 anonymous reviewers are gratefully acknowledged.

8. Appendix A

Generalization of 4D-Var with serial observation error correlation

A crucial ingredient in the practical implementation of a 4D-Var assimilation is the computation of the gradient of the observation cost-function J_o using a forced adjoint model integration, as demonstrated in, e.g., Le Dimet and Talagrand (1986) with observation errors not correlated in time. However, it is not obvious that this technique remains valid when serial error correlations are introduced. This result is proven below. It implies that observation serial error correlations can be introduced into a 4D-Var code with minimal effort.

As in Subsection 4.1, the notation follows Ide et al. (1998). For the sake of simplicity we shall assume that there are only 2 observation times, identified by subscripts 1 and 2 (the algebra is trivially extended to an arbitrary number of times). The model states at these times are linked to the control variable $\boldsymbol{x}$ by the forecast operators M_1 and M_2, i.e., $\boldsymbol{x}_1 = M_1\boldsymbol{x}$ and $\boldsymbol{x}_2 = M_2M_1\boldsymbol{x}$. We denote the observation departures as follows:

$$\boldsymbol{e}_1 = (H_1M_1\boldsymbol{x} - \boldsymbol{y}_1),$$
$$\boldsymbol{e}_2 = (H_2M_2M_1\boldsymbol{x} - \boldsymbol{y}_2).$$

Denoting the linearized versions of H and M by $\boldsymbol{H}$ and $\boldsymbol{M}$, the differentials of the departures with respect to the control variable $\boldsymbol{x}$ are:

$$\delta\boldsymbol{e}_1 = \boldsymbol{H}_1\boldsymbol{M}_1\delta\boldsymbol{x},$$
$$\delta\boldsymbol{e}_2 = \boldsymbol{H}_2\boldsymbol{M}_2\boldsymbol{M}_1\delta\boldsymbol{x}.$$

In the uncorrelated case, it is a classic result that the observation cost-function, its differential and its gradient with respect to $\boldsymbol{x}$ are, respectively:

$$J_o(\boldsymbol{x}) = \boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{R}_1^{-1}\boldsymbol{e}_1 + \boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{R}_2^{-1}\boldsymbol{e}_2,$$
$$\delta J_o = 2(\delta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{R}_1^{-1}\boldsymbol{e}_1 + \delta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{M}_2^{\mathrm{T}}\boldsymbol{H}_2^{\mathrm{T}}\boldsymbol{R}_2^{-1}\boldsymbol{e}_2),$$
$$\nabla J_o = 2\boldsymbol{M}_1^{\mathrm{T}}(\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{R}_1^{-1}\boldsymbol{e}_1 + \boldsymbol{M}_2^{\mathrm{T}}\boldsymbol{H}_2^{\mathrm{T}}\boldsymbol{R}_2^{-1}\boldsymbol{e}_2$$

where $\boldsymbol{R}_i$ is the observation error covariance matrix for time slot i. The evaluation of ∇J_o entails the computation of $\boldsymbol{H}_2^{\mathrm{T}}\boldsymbol{R}_2^{-1}\boldsymbol{e}_2$ and $\boldsymbol{H}_1^{\mathrm{T}}\boldsymbol{R}_1^{-1}\boldsymbol{e}_1$, and their multiplication by the adjoint time-stepping operators $\boldsymbol{M}_2^{\mathrm{T}}$ and then $\boldsymbol{M}_1^{\mathrm{T}}$, which is called the "forced adjoint model".

If there are cross-correlations between the observation errors in both time slots, J_o is defined using the inverse of the observation error covariance matrix $\boldsymbol{R}$ between the slots. The inverse of $\boldsymbol{R}$ is written as a block matrix according to the subspaces defined by the observation times:

$$\boldsymbol{R}^{-1} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^{\mathrm{T}} & \boldsymbol{A}_2 \end{bmatrix},$$

and the cost function is now:

$$J_o\boldsymbol{x} = [\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_2^{\mathrm{T}}]\begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^{\mathrm{T}} & \boldsymbol{A}_2 \end{bmatrix}\begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{bmatrix}.$$

Its differential is equal to

$$\delta J_o = 2[\delta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{H}_1^{\mathrm{T}} \quad \delta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{M}_2^{\mathrm{T}}\boldsymbol{H}_2^{\mathrm{T}}] \times \begin{bmatrix} A_1 & A_{12} \\ A_{12}^{\mathrm{T}} & A_2 \end{bmatrix}\begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{bmatrix},$$

and the gradient can be factorized as follows:

$$\nabla J_o = 2\boldsymbol{M}_1^{\mathrm{T}}\left\{\boldsymbol{H}_1^{\mathrm{T}}[A_1 A_{12}]\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} + M_2^{\mathrm{T}}H_2^{\mathrm{T}}[A_{12}^{\mathrm{T}} A_2]\begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{bmatrix}\right\},$$

which shows that, despite the increased algebraic complexity introduced in the cost-function, ∇J_o is still computed using a forced adjoint model integration, provided that the product between $\boldsymbol{R}^{-1}$

and the departures has been suitably precomputed, stored and provided line by line to the corresponding observation time slots of the forced adjoint run, as in Subsection 4.1. The incremental formulation described in Subsection 4.1.3 only changes the expression of the departures, otherwise the algebra remains the same. The variational quality control does not change the algorithm either, because it is merely a scaling of ∇J_o.

REFERENCES

Andersson, E., Pailleux, J., Thépaut, J-N., Eyre, J., McNally, A. P., Kelly, G. and Courtier, P. 1994. Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.* **120**, 627–653.

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Brankovic, C., Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Thépaut, J-N. and Viterbo, P. 1998. The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part III: Experimental results. *Q. J. R. Meteorol. Soc.* **124**, 1831–1860.

Andersson, E. and Järvinen, H. 1999. Variational quality control. *Q. J. R. Meteorol. Soc.* **125**, 697–722.

Bengtsson, L. 1980. On the use of a time sequence of surface pressures in four-dimensional data assimilation. *Tellus* **32**, 189–197.

Courtier, P., Thépaut, J-N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367–1387.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. 1998. The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.* **124**, 1783–1808.

Daley, R. 1992. The effect of serially correlated observation and model error on atmospheric data assimilation. *Mon. Wea. Rev.* **120**, 164–177.

Derber, J. and Bouttier, F. 1999. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221.

Hollingsworth, A., Shaw, D., Lönnberg, P., Illari, L., Arpe, K. and Simmons, A. 1986. Monitoring of observation and analysis quality by a data-assimilation system. *Mon. Wea. Rev.* **114**, 1225–1242.

Ide, K., Courtier P., Ghil, M. and Lorenc, A. 1997. Unified notations for data assimilation: operational, sequential and variational. *J. Met. Soc. Japan* **75**, 181–189.

Ingleby, N. B. and Lorenc, A. C. 1993. Bayesian quality control using multivariate normal distributions. *Q. J. R. Meteorol. Soc.* **119**, 1195–1225.

Järvinen, H. and Undén, P. 1997. *Observation screening and first guess quality control in the ECMWF 3D-Var data assimilation system.* ECMWF Tech. Memo. 236. Available from ECMWF.

Le Dimet, F.-X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* **38A**, 97–110.

Lorenc, A. C. 1981. A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.* **109**, 701–721.

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

Lönnberg, P. 1988. Developments in the ECMWF analysis system. Proc. ECMWF seminar on *Data assimilation and the use of satellite data*, Reading, 5–9 November 1988, Vol. 1, 75–120. Published by ECMWF, Shinfield Park, Reading, RG2 9AX, UK.

Parrish, D. F. and Derber, J. C. 1992. The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763.

Phalippou, L. 1996. Variational retrieval of humidity profile, wind speed and cloud liquid-water path with the SSM/I: Potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **122**, 327–355.

Rabier, F., Thépaut, J-N. and Courtier, P. 1998a. Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Q. J. R. Meteorol. Soc.* **124**, 1861–1887.

Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F. 1998b. The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. *Q. J. R. Meteorol. Soc.* **124**, 1809–1829.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J-F. and Simmons, A. 1999. The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, submitted.

Thépaut, J.-N., Hoffman, R. N. and Courtier, P. 1993. Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.* **121**, 3393–3414.

Thépaut, J.-N., Courtier, P., Belaud, G. and Lemaître, G. 1996. Dynamical structure functions in four-dimensional variational assimilation: A case study. *Q. J. R. Meteorol. Soc.* **122**, 535–561.

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