Influence of the assimilation scheme on the efficiency of adaptive observations

By THIERRY BERGOT*
Centre National de Recherches Météorologiques, France

(Received 15 May 2000; revised 11 September 2000)

SUMMARY

The influence of assimilation schemes on the efficiency of adaptive observations is investigated, with three- and four-dimensional variational data assimilation (3D-Var and 4D-Var) systems, for 20 cases of cyclogenesis during the French and Atlantic Storm-Track Experiment (FASTEX). During FASTEX, adaptive observations were made in such a way that the forecast of synoptic-scale systems characterized by a weak predictability was improved.

This study shows that, on average, adaptive observations significantly influence the forecast for both assimilation schemes. However, with 3D-Var, the inclusion of adaptive observations leads to a blending of improvements and worsenings, resulting in a very weak mean improvement. In contrast, with 4D-Var the forecasts are, on average, significantly improved at the verification time (a mean improvement of about 10% and a maximum improvement of more than 50%) without producing large variability at medium-range.

Results confirm that the positive impact of 4D-Var primarily occurs as a result of a change in the initial conditions at the most unstable directions. This positive impact is particularly obvious when adaptive observations are mainly located at the border of the sensitive area. In such cases, the known deficiencies of 3D-Var may be detrimental. This study also demonstrates that the impact of sampling of the sensitive area depends on the assimilation scheme.

KEYWORDS: Adaptive observations Data assimilation Predictability

1. INTRODUCTION

The forecast of some meteorological events (rapid cyclogenesis, for example) remains a difficult prediction problem, even at short range (12 to 48 hours). This is a consequence of the chaotic behaviour of the atmospheric flow. In these situations, small but fast-growing initial errors lead to significant uncertainties in forecasts, and can lead to severe forecast failures. These forecast failures sometimes have dramatic consequences in terms of societal impact (strong winds, rainfall, etc.). Unfortunately, it is very difficult to control fully the fast-growing errors with the current observational network and current assimilation systems, particularly over regions with sparse data, such as the oceans. Moreover, the lack of predictability is not uniform over the globe but is strongly flow-dependent whilst, in contrast, the observations used for producing initial conditions are mainly routine observations at fixed locations. In this context, the development of adaptive-observation strategies has become an important area of research during recent years. In the so-called 'adaptive-observation' or 'targeted-observation' strategies, the underlying idea is that the quality of the forecast of weather features can be improved by adding observations in predicted sensitive areas. The strategy is adaptive in the sense that the location of these measurements may vary from day to day. One of the main scientific objectives of recent field experiments, such as FASTEX† (Joly et al. 1999); NORPEN‡ (Langland et al. 1999) or WSRP§ (Szunyogh et al. 2000) was to provide targeted observations in order to control the development of forecast errors and, therefore, to try to improve the forecasting of systems characterized by a weak predictability.

* Corresponding address: Météo-France, CNRM/GMME, 42, avenue G. Coriolis, F-31057 Toulouse cedex, France. e-mail: Thierry.Bergot@meteo.fr
‡ Fronts and Atlantic Storm-Track EXperiment.
§ NORth Pacific EXperiment.
§ Winter Storm Reconnaissance Program.
During the years preceding these experiments, different ways of finding the optimal locations for adaptive observations have been suggested. One kind of strategy is based on adjoint products and use the linear-adjoint versions of a nonlinear forecast model. These adjoint techniques include the computation of singular vectors (SVs) (Buizza and Palmer 1995; Gelaro et al. 1999; Bergot et al. 1999) or the computation of sensitivity fields (Rabier et al. 1996; Langland et al. 1996; Bergot et al. 1996). This adjoint approach is based on the idea that, while only a small part of the analysis error may project onto an unstable subspace, the growth of this component of the analysis error dominates the forecast error. Another kind of targeting strategy used during previous field experiments is based on the ensemble transform technique (Bishop and Toth 1999). This technique uses the operational ensemble forecasts from the National Centers for Environmental Prediction (NCEP) or from the European Centre for Medium-Range Weather Forecasts (ECMWF). The main goal is to find the locations of observations that will minimize a given norm of the forecast-error variance inside a verifying area. Additional details on the methodology of adaptive observations can also be found in the papers by Snyder (1996) or Emanuel and Langland (1998).

The results from these field experiments demonstrate that the inclusion of adaptive observations can lead, in some cases, to a significant improvement in forecast skill. However, a systematic study of all targeting missions carried out during FASTEX shows that a large case-to-case variability still exists, with cases of strong improvements to the forecast, but also cases of strong worsening, depending on many factors (Bergot 1999). Among these, one can cite, for instance, the sampling of the sensitive area; a partial survey of the sensitive area can lead, with three-dimensional variational data assimilation (3D-Var), to a deterioration in the forecast quality (Aberson 2000; Bergot et al. 1999). This point is very important for the feasibility of the application of adaptive observations, and is currently being studied (Berliner et al. 1999; Baker and Daley 2000; Doerenbecher and Bergot 2000). However, studies have also shown that the impact of additional observations was primarily due to corrections of the initial error that project onto an unstable subspace of weak dimension (Gelaro et al. 1999; Gelaro et al. 2000; Bergot et al. 1999; Hello et al. 2000). This unstable subspace can be described by the first SVs of the forecast model. Therefore, if one wants adaptive observations to improve the forecast significantly, the assimilation scheme should be able to minimize the errors that project onto this unstable subspace. Consequently, the success of adaptive observations will also strongly depend on the assimilation scheme used. A theoretical study with a very simplified and idealized model suggested that adaptive observations can be more efficient if dynamical assimilation covariances are specified (Fischer et al. 1998).

The main goal of the work described in this paper is to investigate the efficiency of adaptive observations obtained in real time during FASTEX with different up-to-date assimilation systems, and to assess and understand the effect of the assimilation scheme on this efficiency. In order to explore this idea in a real context, a systematic survey of the forecast impact of the FASTEX targeted observations are presented in this article for both 3D-Var and 4D-Var assimilation schemes.

After a short description of the FASTEX experiment, the strategy used for this study is described (section 2). In section 3, the ability of FASTEX upstream observations to influence the forecast is tested for both 3D-Var and 4D-Var, without taking into account the real evolution of the atmosphere (i.e. in the context of a perfect model). Section 4 focuses on the ability of adaptive observations to improve the forecasts. A discussion of the results is presented in section 5. And finally, section 6 is devoted to conclusions and perspectives.
TABLE 1. SUMMARY OF THE FASTEX FLIGHTS BY THE LEARJET-36 AND THE GULFSTREAM-IV AIRCRAFT DURING THE INTENSIVE OBSERVING PERIODS (IOPs)

<table>
<thead>
<tr>
<th>IOP</th>
<th>Initial time</th>
<th>Number of data</th>
<th>Final time</th>
<th>Optimization time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Learjet-36 flights</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOP3</td>
<td>12 UTC 13 January 1997</td>
<td>14</td>
<td>12 UTC 15 January 1997</td>
<td>48</td>
</tr>
<tr>
<td>IOP5</td>
<td>12 UTC 20 January 1997</td>
<td>16</td>
<td>00 UTC 22 January 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP9</td>
<td>00 UTC 01 February 1997</td>
<td>10</td>
<td>00 UTC 03 February 1997</td>
<td>48</td>
</tr>
<tr>
<td>IOP11</td>
<td>12 UTC 04 February 1997</td>
<td>15</td>
<td>00 UTC 06 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP13</td>
<td>12 UTC 10 February 1997</td>
<td>19</td>
<td>12 UTC 12 February 1997</td>
<td>48</td>
</tr>
<tr>
<td>IOP14</td>
<td>12 UTC 12 February 1997</td>
<td>20</td>
<td>00 UTC 14 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP15</td>
<td>12 UTC 14 February 1997</td>
<td>17</td>
<td>00 UTC 16 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP16</td>
<td>12 UTC 16 February 1997</td>
<td>18</td>
<td>18 UTC 17 February 1997</td>
<td>30</td>
</tr>
<tr>
<td>IOP17</td>
<td>00 UTC 18 February 1997</td>
<td>13</td>
<td>12 UTC 19 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP18</td>
<td>00 UTC 22 February 1997</td>
<td>18</td>
<td>12 UTC 23 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>FLOP6</td>
<td>00 UTC 23 February 1997</td>
<td>10</td>
<td>00 UTC 25 February 1997</td>
<td>48</td>
</tr>
<tr>
<td>IOP19</td>
<td>00 UTC 26 February 1997</td>
<td>16</td>
<td>12 UTC 27 February 1997</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Gulfstream-IV flights</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOP9</td>
<td>12 UTC 01 February 1997</td>
<td>26</td>
<td>00 UTC 03 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP10</td>
<td>18 UTC 03 February 1997</td>
<td>32</td>
<td>00 UTC 05 February 1997</td>
<td>30</td>
</tr>
<tr>
<td>IOP12</td>
<td>12 UTC 08 February 1997</td>
<td>16</td>
<td>00 UTC 10 February 1997</td>
<td>36</td>
</tr>
<tr>
<td>IOP15</td>
<td>06 UTC 15 February 1997</td>
<td>28</td>
<td>00 UTC 16 February 1997</td>
<td>18</td>
</tr>
<tr>
<td>IOP17</td>
<td>18 UTC 17 February 1997</td>
<td>20</td>
<td>12 UTC 19 February 1997</td>
<td>42</td>
</tr>
<tr>
<td>IOP18</td>
<td>18 UTC 18 February 1997</td>
<td>52</td>
<td>12 UTC 19 February 1997</td>
<td>18</td>
</tr>
</tbody>
</table>

The initial time corresponds to the observation time and the final time corresponds to the verification time used for the computation of adjoint-based products. The last column indicates the optimization time in the adjoint computations.

2. METHODOLOGY

(a) Datasets

The FASTEX field experiment has been described in detail by Joly et al. (1999), and only some information concerning the data used in this study is summarized here. The FASTEX adaptive observing platforms included the Learjet-36 (hereafter L36), commissioned by the National Center for Atmospheric Research (NCAR) and based at St Johns (Newfoundland), and the Gulfstream-IV (hereafter GIV), a National Oceanographic and Atmospheric Administration (NOAA) aircraft operated jointly by NOAA, NRL*, CNRS† and Météo-France, and based at Shannon (Ireland). Both aircraft were equipped for global positioning system dropsondes‡.

The L36 missions were clearly dedicated to objective targeting purposes, primarily using ensemble-transform targets, but also adjoint targets. Twelve L36 flights were carried out for twelve different Intensive Observing Periods (IOPs), and the total number of dropsondes available is 186 (a mean number of soundings of about 16 per flight), see Table 1. For the GIV, both so-called ‘objective’ purposes (adjoint-based targets) and ‘subjective’ purposes (cyclogenesis precursors such as potential-vorticity anomalies, for example) were often combined. Nine GIV flights were carried out for seven IOPs (see Table 1). The total number of available dropsondes is 232 and the mean number of soundings is about 25 per flight.

* Naval Research Laboratory.
† Centre National de Recherches Scientifiques.
‡ see http://www.meteo.fr/cnrm/fastex/
Only two flights (and IOPs) were performed in January because of the meteorological situation (Greenland ridge and blocking until the beginning of February). The large-scale flow during the second period was characterized by a strong zonal flow and intense baroclinic developments. This explains the large number of flights during February.

(b) The assimilation and forecast model

The model used was the French ARPEGE/IFS operational forecast model (Courtier et al. 1991), developed in cooperation with the ECMWF, with a horizontal resolution of T95C3.5 (stretching factor of 3.5) and 27 vertical levels. A lower resolution than the operational one was chosen because of the very large number of experiments required for completing the present study, and also in order to retain consistency with a previous study (Bergot 1999).

Two variational assimilation schemes were used: an incremental 3D-Var scheme and an incremental 4D-Var scheme (both used the so-called incremental approach (Courtier et al. 1994)). In this study the resolution of the increments was T63 with 27 vertical levels. The 3D-Var scheme was the assimilation scheme used operationally at Météo-France until June 2000 (Thépaut et al. 1998). The 4D-Var assimilation scheme was the temporal extension of the 3D-Var over a six-hour window; the principle of the scheme is similar to the operational ECMWF system (Rabier et al. 2000). This assimilation scheme has been operational at Météo-France since June 2000. The 4D-Var scheme minimizes a cost function measuring the distance between a model trajectory and the available information (background field and observations) over a six-hour assimilation window. The 4D-Var minimization is organized in three outer loops (i.e. three minimization sequences of the linear low-resolution model) and three updates of the high-resolution trajectory (Gauthier and Thépaut 2000). During the first two outer loops, the minimization is done almost adiabatically (with very simplified dry physics) with 25 iterations (inner loops), and during the last outer loop the minimization is performed with more elaborate linearized physics (Janiskova et al. 1999) and 20 iterations. The covariances used at the start of each assimilation (i.e. three hours before the analysis time) are isotropic, and similar to the 3D-Var ones. However, these covariances are affected by the dynamics during the six-hour assimilation period (see Rabier et al. (2000) for more details).

(c) Methodology of the tests

The observational network was significantly upgraded during FASTEX (with for example ships in the middle of the Atlantic Ocean and increased soundings in USA and Canadian coastal areas). A reference cycle of analyses without any additional FASTEX data has been performed with a 3D-Var scheme (Bergot 1999) and this allows the effectiveness of adaptive observations with respect to the conventional observation system to be tested. This analysis cycle covers the entire FASTEX period (January–February 1997), and leads to the construction of a background field uninfluenced by FASTEX data.

Strictly speaking, the effect of the analysis can only be identified precisely in a single assimilation cycle run from identical background fields and observations. For this study, the background fields for both 3D-Var and 4D-Var were provided from the reference reanalyses, and the observations contain the conventional and/or targeted ones. For each case studied, four analyses and forecasts have been generated (two for the 3D-Var and two for the 4D-Var). All these experiments used the same background field, and only differed because of the assimilation scheme and/or the observations used. The
first experiments that only used the conventional observations (without the FASTEX data) are hereafter referred to as WFD3 for 3D-Var and WFD4 for 4D-Var. The second experiments that used the targeted observations (in addition to the conventional ones) are hereafter referred to as TD3 for 3D-Var and TD4 for 4D-Var.

As in the paper by Bergot (1999), the effect of FASTEX adaptive observations was assessed in two different ways. In a first stage, the ability of adaptive observations to influence the analysis and forecasts was tested without taking into account the real evolution of the atmosphere. At this stage, the impact of the adaptive observations was defined as the difference between two forecasts (without taking into account the real evolution of the atmosphere). The evolution of the impact signal in time and space allows testing of the efficiency of the targeted data to produce diverging forecasts, thereby assessing the crucial role of these observations. In a second stage, the ability of the targeted data to improve the forecasts was investigated, and the error was defined as the difference between the forecast and a verifying analysis. This allows an assessment of the improvement to the forecast caused by the targeted observations to be made.

To help in quantifying the efficiency of targeted observations, a forecast verifying area was defined, with boundaries at 0° and 30°W, and at 35°N and 65°N. This verifying area was constructed so as to contain the trajectories of the FASTEX lows studied, and to reflect the influence of the targeted observations over the west European coast for synoptic features. The score used to quantify the influence of targeted observations was the kinetic energy near the top of the boundary layer, at 850 hPa. This score was chosen because it is directly related to wind intensity and is, therefore, meaningful in terms of socio-economic consequences.

(d) SV computation

SVs are intimately associated with the notion of predictability (Lorenz 1965; Farrel 1990). Specifically, the SVs are very helpful in estimating the evolution of the initial errors during a forecast. It seems that the SVs can be useful tools for diagnosing and understanding the influence of the assimilation scheme. SVs have been systematically computed for the 20 cases studied, and the principle of the SV calculation is briefly discussed here.

SVs represent the linear perturbations with maximum growth measured in terms of specified norms and over a specified finite time interval. For two given scalar products $\langle \cdot, \cdot \rangle_{E_0}$ at initial time, and $\langle \cdot, \cdot \rangle_{E_1}$ at final time, the norm of the perturbation at final time is:

$$\langle \delta X(t_1) ; \delta X(t_1) \rangle_{E_1} = \langle L_E^* L \delta X(t_0) ; \delta X(t_0) \rangle_{E_0}$$

(1)

where the operator $L_E^*$ is the adjoint of the operator $L$ following the norms $E_0$ and $E_1$, and $X$ represents the state vector of the model.

The square root of the eigenvalues of the matrix $L_E^* L$ are called the singular values, and the eigenvectors are the SVs. The calculation of the SVs and the singular values uses the iterative Lanczos algorithm. A so-called local projection operator is applied at final time to identify perturbations whose norm is maximized over the verifying area (0–30°W, 35–65°N).

For this study, computations have been made with an energy-based scalar product at the initial time ($E_0$). This energy scalar product is the same as the scalar product employed to study weather predictability (at the NRL or ECMWF, for example). At the final time, the scalar product used is the kinetic energy around 850 hPa ($E_1$). This norm was chosen because it is directly related to wind intensity and is, therefore, meaningful in term of socio-economic consequences.
3. IMPACT OF FASTEX TARGETED DATA

(a) Mean impact

As previously explained, the goal of this section is to assess the ability of targeted data to influence the forecasts and to produce diverging forecasts. The impact function at time $t$, $\delta I(t)$, is defined as the difference between two forecasts without targeted data (WFD) and with targeted data (TD):

$$\delta I(t) = X_{WFD}(t) - X_{TD}(t)$$

where $X$ is the state vector of the model. To measure the divergence between two forecasts, one focuses on the kinetic energy norm of $\delta I(t)$ at 850 hPa inside the verifying area (0–30°W, 35–65°N).

Figure 1 shows the mean impact of adaptive observations for 3D-Var and 4D-Var as a function of the forecast time. The results with 3D-Var are qualitatively similar to those obtained by Bergot (1999) from mean-sea-level pressure diagnostics.

It is important to remember that the impacts of data from the GIV and L36 flights have different characteristics, on average. Moreover, the influence of the assimilation scheme on the impact signal is also different for the L36 and GIV flights.

For the GIV-flight observations (Fig. 1(b)), the influence of the assimilation scheme is weak, up to 48 h (with similar mean impacts and similar standard deviations for both 3D-Var and 4D-Var). At 30–36 h (close to the verifying time, see Table 1), the targeted observations have the same non-negligible impact for both assimilation schemes. The magnitude of this mean impact is about 50% of the mean forecast error at that time. However, with 3D-Var the adaptive observations clearly affect the medium-range forecast. This impact at medium range is significantly reduced with the 4D-Var scheme by a factor of two (there is a reduction of the mean impact and of the standard deviation). This result demonstrates that, with 4D-Var, adaptive observations produce less variability in the medium-range forecasts.

For data from the L36 flights (Fig. 1(a)), with 3D-Var the shape of the signal turns out to be mixed, with no clear range of maximum impact. With 4D-Var the L36 impact exhibits a targeted behaviour with a peak at 60 h. However, this peak is far from the mean verifying time (about 40 h, see Table 1). This discrepancy (of about 20 h) is relatively large when one considers that, in contrast, the peak and the mean verifying times are quite similar for the GIV flights. For these L36 flights, the assimilation scheme seems to influence significantly the impact signal, with a reduction of the mean impact of about 25% up to 48 h. The impact at 72 h is also significantly reduced thanks to the use of 4D-Var, as with the results using data from the GIV flights.

The different behaviour of the impact signal might be a consequence either of differences between the targeting methods used for planning the flights, and/or of differences in the sampling of the sensitive areas. This point is studied in the following sections. Nevertheless, this first result suggests that the assimilation scheme plays an important role in the efficiency of adaptive observations to influence the forecasts.

(b) Detailed study of the impact at verification time

In this section, one focuses on the impact signal at the verification time (defined in Table 1). In order to study the dependence of the impact at verification time on the modifications of the initial conditions, the projection of the initial impact on the unstable subspace defined by $i$ SVs is examined. This projection, $P_i$, is defined by the sum of the
Figure 1. Time evolution of the impact signal for (a) the Learjet-36 and (b) the Gulfstream-IV targeted flights. The mean impact is plotted with a solid line for 4D-Var and with a dash-dotted line for 3D-Var. The vertical bars represent the standard deviations.
energy dot product between the SV $V_k(t_0)$ and the impact $\delta I(t_0)$ at initial time:

$$P_i = \left( \sum_{k=1}^{i} (V_k(t_0); \delta I(t_0))_{E_0}^2 \right)^{1/2} \text{ (J kg}^{-1})$$

(i) **L36 flights.** Figure 2(a) shows the impact of the L36 targeted data at the verification time. The projections of the modification to the initial conditions due to the inclusion of the L36 targeted data onto the unstable subspace, $P_i$, (defined by two SVs and ten SVs, respectively) are plotted in Figs. 2(b) and (c).

The case of the L36 flight for FLOP6 is noteworthy. For this case, the adjoint-based target area was not sampled (Fig. 3(a)), and the data were in a ‘null area’ from the adjoint–target point of view. Figure 2(a) shows that the impact is small at the verification time with 3D-Var (about 4.3% of the forecast error), and even negligible with 4D-Var (about 1.2% of the forecast error). The influence of the assimilation scheme is weak on the unstable subspace defined by these first two SVs, and one can just observe small differences on the ten-SV unstable subspace. This result suggests that data in a null area have a weak impact on the forecast for both assimilation schemes.

As previously mentioned, the impact of the L36 data is significantly weaker with 4D-Var than with 3D-Var. One finds a similar result at the verification time for several IOPs (IOP3, IOP11, IOP14, IOP17, for the most significant cases). The case of the L36 flights for IOP17 is typical of the influence of the assimilation scheme on the impact signal, and has been studied in detail. The impact of the targeted data at the verification time is significantly reduced with 4D-Var (by a factor of about three). This reduction is consistent with the reduction of the impact at the initial time in the direction of the first two SVs (a reduction by a factor of about three). A detailed study of the geographical location of the data (Fig. 3(b)) shows that these targeted data are located close to the western border of the the sensitive area defined by the first two SVs (behind the stronger maximum), and that only a small portion of the SV-based sensitive area is sampled. This is not specific to this IOP, since for many others only a few observations were made in the core of the SV-based targets. Therefore, the SV-targets were not often sampled in detail. Data located at the upstream border of the SV-based target can explain the late time of the maximum of the impact (60 h for the maximum of the mean impact for a verification time between 36–48 h). This is indisputably the case for at least IOP14, IOP15 and IOP17.

These examples suggest that the influence of the assimilation scheme strongly depends on the sampling of the sensitive area.

(ii) **GIV flights.** Figure 4(a) shows the impact of the GIV targeted data at the verification time, and the projections $P_i$ defined by two and ten SVs are plotted in Figs. 4(b) and (c). The first case studied is given by the GIV flight for IOP9. This IOP is characterized by a weak singular value (less than 3). In this case, the GIV data have weak impacts at the verification time for both assimilation schemes (about 2.7% of the forecast error with 3D-Var, and about 4.3% with 4D-Var), and the influence of the assimilation scheme is weak. The projections of the analysis changes at the initial time onto the first two SVs or ten SVs unstable subspaces are similar with 3D-Var and 4D-Var. One can note in Fig. 5(a) that the geographical area defined by the first two SVs is relatively well sampled. This example suggests that cases of weak amplification rates can lead to very weak impact even if the SV targets are well sampled.

Another interesting case is the first GIV flight for IOP17 (18 UTC 17 February). This flight had mainly a targeting goal (sampling of adjoint-based targets) in the early stage of
Figure 2. (a) The impact of the L36 targeted data. The other panels show the impact at the verification time of the projection $P_t$ (see text) of the analysis changes onto an unstable subspace defined by (b) the first two and (c) the first ten singular vectors. The left-hand bars are obtained with 3D-Var and the right-hand bars with 4D-Var.
Figure 3. The Learjet-36 flights for intensive observing periods (a) FLOP6 and (b) IOP17 (see Table 1), with the temperature of the first singular vector at 700 hPa superimposed. The circled crosses indicate the dropsonde positions.

the development of the cyclone. The singular values are one of the largest over the whole FASTEX experiment. As for the L36 flight for this IOP, the impact of the targeted data is clearly reduced with 4D-Var (by a factor two). This reduction is not very pronounced in the direction of the first two SVs, but is particularly clear when the ten first SVs are considered. This can be explained by the good sampling of the target defined by the first two SVs (the sensitive area defined operationally during FASTEX is based on these first two SVs). Nevertheless, the following SVs have significant amplification rate (more than five) and are not well sampled by the GIV data; in particular, the southern border of the sensitive area is only partially sampled, as emphasized by Langland et al. (1999).
Figure 4. As Fig. 2, but for the Gulfstream-IV targeted data.
A comparison of Figs. 2(a) and (c), or Figs. 4(a) and (c), illustrates the overall analogy between the modification of the initial conditions in an unstable subspace of low dimension (defined here by the first ten SVs), and the modification of the impact at verification time. This result supports the idea that the growth of the initial difference between two forecasts is dominated by a relatively small number of unstable structures. These structures seem to be the most efficient in controlling the divergence between two forecasts. Similar results have been obtained by Gelaro et al. (2000) on NORPEX cases. With similar information (background field and observations), the assimilation scheme can produce significantly different initial conditions in these unstable directions.
Consequently, the assimilation scheme can play an important role in the effectiveness of adaptive observations.

4. IMPROVEMENT OF THE FORECASTS

(a) Mean improvement

The purpose of using adaptive observations is to improve the forecasts by assimilating properly positioned data. One way to test whether this goal has been reached is to quantify the quality of the forecasts. Unfortunately, verification of the forecasts was not a major objective of FASTEX. However, numerous radiosondes were added in the North Atlantic region during FASTEX (FASTEX ship soundings every hour and a half, conventional soundings every three hours), and the best available description of the atmosphere is given by the operational analyses.

To measure the quality of the forecasts, one focuses on the 850 hPa kinetic-energy norm of the errors $\delta E(t)$, inside the verifying area ($0-30^\circ W$, $35-65^\circ N$). This error is defined as the difference between the forecast and the verifying analysis.

The reduction of the forecast errors for both 3D-Var and 4D-Var, thanks to the use of FASTEX adaptive observations, is shown in Fig. 6(a) for L36 and Fig. 6(b) for GIV targeted flights. Positive values indicate forecast improvement (i.e. a reduction of the kinetic-energy norm of the error) due to the inclusion of targeted data.

With 3D-Var, adaptive observations have a very weak influence on the forecast error, on average, with only a very small mean improvement for the 12–36 h forecasts that include the GIV data (a reduction of about 3% of the error at 24 h, see Table 2). The inclusion of the L36 data with the 3D-Var assimilation scheme shows a mean systematic worsening of the forecasts (an increase in the error of about 2% at 36 h, see Table 2). Generally speaking, the inclusion of FASTEX adaptive observations with 3D-Var leads to a mixture of improvement and worsening, as previously shown by Bergot (1999) in terms of the mean-sea-level pressure root-mean-square scores. From an operational point of view, it is disappointing to see that, in some cases, targeted data assimilated with 3D-Var produce a large, but negative, impact and that adding observations in sensitive areas can greatly increase the chance of worsening the forecasts.

In contrast, with 4D-Var the assimilation of FASTEX adaptive observations leads to a systematic mean improvement of the forecasts, at all forecast ranges, and for both aircraft (Fig. 6). The mean reduction of the error due to the inclusion of GIV data is about 12% at 24 h (see Table 2). In agreement with the discussion in the previous section, the GIV and L36 flights exhibit different behaviours in terms of the forecast improvement. However, the impact signal (Fig. 1), and forecast improvement (Fig. 6) exhibit similar behaviours with 4D-Var: a maximum improvement and maximum impact at 36 h for the GIV data, versus a maximum improvement and maximum impact at 60 h for the L36 data. Another remark is that the spread of the forecast improvement is clearly reduced with 4D-Var after 48 h. This result suggests that, with 4D-Var, the inclusion of adaptive observations leads to a significant improvement of the short-term forecast without producing an important variability of the medium-range forecast errors. Similar results have been obtained by Rabier et al. (2000) with the ECMWF 4D-Var system.

(b) Improvement at verification time

In this section, the forecast error at verification time (defined in Table 1) is studied. In order to interpret the response of the analysis and forecast to the assimilation of targeted data, the forecast error at verification time is projected onto the evolved unstable subspace defined by the evolved and normalized SVs. This projection, $Q_I$, is defined by
Figure 6. The time evolution of the improvements in the forecast for (a) the Learjet-36 and (b) the Gulfstream-IV targeted flights. The mean improvement is plotted with a solid line for 4D-Var and with a dash-dotted line for 3D-Var. The vertical bars represent the standard deviations.
TABLE 2. The time evolution of the mean relative improvement (%) of the forecasts from initial data produced by 3D-Var and 4D-Var with the Learjet-36 and Gulfstream-IV data included

<table>
<thead>
<tr>
<th>Forecast time (h)</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
<th>36</th>
<th>42</th>
<th>48</th>
<th>54</th>
<th>60</th>
<th>66</th>
<th>72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learjet-36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-Var</td>
<td>-0.4</td>
<td>-1.0</td>
<td>-1.1</td>
<td>-2.2</td>
<td>-2.5</td>
<td>-1.3</td>
<td>-0.2</td>
<td>+0.2</td>
<td>-0.8</td>
<td>-1.2</td>
<td>-0.4</td>
</tr>
<tr>
<td>4D-Var</td>
<td>+0.6</td>
<td>+0.6</td>
<td>+1.2</td>
<td>+0.6</td>
<td>+1.5</td>
<td>+2.7</td>
<td>+5.1</td>
<td>+6.1</td>
<td>+6.4</td>
<td>+1.1</td>
<td>-0.2</td>
</tr>
<tr>
<td>Gulfstream-IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-Var</td>
<td>-0.5</td>
<td>+1.0</td>
<td>+2.9</td>
<td>+1.1</td>
<td>+1.5</td>
<td>-0.5</td>
<td>-0.5</td>
<td>+2.3</td>
<td>-0.2</td>
<td>+0.9</td>
<td>-2.0</td>
</tr>
<tr>
<td>4D-Var</td>
<td>+5.0</td>
<td>+12.4</td>
<td>+11.6</td>
<td>+8.0</td>
<td>+7.8</td>
<td>+4.4</td>
<td>+1.3</td>
<td>+3.4</td>
<td>+1.4</td>
<td>+2.8</td>
<td>+1.7</td>
</tr>
</tbody>
</table>

the sum of the kinetic energy dot product between the SVs $V_k(t_1)$ and the forecast error at verification time $\delta E(t_1)$:

$$Q_i = \left( \sum_{k=1}^{i} (V_k(t_1); \delta E(t_1))^2 \right)^{1/2}$$  (4)

If the linear hypothesis is valid and if the model is perfect, this projection $Q_i$ explains that part of the forecast error that results from the analysis error onto the unstable subspace defined by $i$ SVs.

(i) L36 flights. Figure 7(a) shows the improvement of the forecast at the verification time due to the inclusion of L36 targeted data, for both 3D-Var and 4D-Var. The difference of $Q_i$ between the reference (WFD) and targeted (TD) forecasts are plotted in Fig. 7(b) for two SVs and in Fig. 7(c) for ten SVs.

Some cases show a weak influence of the assimilation scheme on the efficiency of adaptive observations. The first example is given by the L36 flight for IOP9, which also represents the best improvement of the forecast obtained with the L36 data, for both 3D-Var and 4D-Var. One observes an unambiguous improvement of the forecast, and a reduction of the forecast error of about 35%. The projections onto the unstable subspace (two SVs or ten SVs) are similar for both 3D-Var and 4D-Var. Szunyogh et al. (1999) have also obtained an unambiguous forecast improvement for this case with the NCEP analysis–forecast model. Therefore, this result seems very robust and does not depend strongly on the assimilation scheme or the forecast model used. The IOP9 case is clearly a success, in the sense of targeted observing strategy, for different assimilation and forecast models. The second largest reduction of the forecast error is given by the assimilation of targeted data for IOP18 (+22% for 3D-Var and +27% for 4D-Var). For this case too, the projection of the errors onto the unstable subspace is relatively similar for both 3D-Var and 4D-Var, and the influence of the assimilation scheme is relatively weak.

When data are located in a null area (far away from the sensitive area), as in the FLOP6 case (Fig. 3(a)), the influence of the assimilation scheme is weak for both the impact signal and the improvement of the forecast. The inclusion of these data have a negligible influence on the forecast error with both assimilation schemes (less than 1%), and the projection of the error onto the unstable directions is also similar for both assimilation schemes.

Another interesting case is the L36 flight for IOP17. As previously mentioned, numerous targeted data lie at the upstream border of the SV-based target (Fig. 3(b)) and the impact signal is clearly reduced with 4D-Var (Fig. 2(a)). With 3D-Var, the
Figure 7. Improvement of the forecast due to the inclusion of the Learjet-36 targeted data. (a) The improvement at the verification time, and the difference of $Q^*_t$ between reference and targeted forecasts (see text) for (b) an unstable subspace defined by the first two singular vectors and (c) for an unstable subspace defined by the first ten singular vectors. The left-hand bars are obtained with the 3D-Var assimilation scheme and the right-hand bars with 4D-Var.
forecast is worsened by about $-26\%$, while with 4D-Var the inclusion of targeted data improves the forecast by about $8\%$. More than $76\%$ of the forecast error with 3D-Var is explained by the error of the projection onto the ten-SV unstable subspace. This result suggests that with 3D-Var the inclusion of targeted data that only sample a part of the sensitive area may introduce analysis errors in the unstable subspace. In these cases, the inaccuracies of the assimilation processes can be detrimental by introducing small errors in the sensitive area which, by definition, will amplify and worsen the forecast. This problem is significantly reduced with 4D-Var. The same behavior is also obtained for other IOPs, such as IOP11. This point is studied in detail in section 5.

In the case of IOP15, the inclusion of targeted data clearly worsens the forecast for both assimilation schemes, even though the worsening is weaker with 4D-Var (worsening by about $-77\%$ for the 3D-Var and $-40\%$ for the 4D-Var). Szunyogh et al. (1999) also obtained significant forecast worsening for this case with the NCEP analysis--forecast model. Therefore, it seems that the problem is not related to model errors, but rather to initial conditions. For this case too, the errors of the projection onto the ten-SV unstable subspace explain $70\%$ of the forecast error with 3D-Var and $61\%$ with 4D-Var. A detailed study of the flight shows that the data are principally located in the border of the adjoint-based sensitive area.

(ii) *GIV flights.* Figure 8(a) shows the improvement of the forecasts at the verification time due to the inclusion of the GIV targeted data. The difference of $Q_i$ between the reference and targeted forecasts is plotted on Fig. 8(b) for two SVs and on Fig. 8(c) for ten SVs.

The first GIV flights for IOP17 (18 UTC 17 February) and IOP18 (12 UTC 22 February) mainly have a targeting goal (sampling of adjoint-based targets). The inclusion of these data clearly significantly reduces the kinetic-energy error of the forecasts with 3D-Var (about $+19\%$ for IOP17 and $+35\%$ for IOP18). This improvement is still larger with 4D-Var (about $+25\%$ for IOP17 and $+51\%$ for IOP18). For these IOPs, the influence of the assimilation scheme is particularly clear on the unstable subspace defined by ten SVs.

The second GIV flight of IOP17 (18 UTC 18 February) mainly has a subjective goal (sampling of cyclogenesis precursors, see Joly et al. (1999)). Many dropsondes (52) were launched, however the improvement is weak for both 3D-Var and 4D-Var. These data do not correct errors in the unstable subspace, for both assimilation schemes. It seems that this flight is a typical example of a useless flight, in terms of efficiency.

One case corresponds to a worsening of the forecasts for both assimilation schemes: IOP12. The worsening is of about $-31\%$ with 3D-Var and about $-16\%$ with 4D-Var. For this IOP, strong instabilities are present. The most unstable sensitive area is close to the north-east coast of the USA (Fig. 5(b)), and the GIV data are located downstream of this area. The influence of the assimilation scheme on the errors in the most unstable directions is weak (Fig. 8), despite a strong influence on the impact (Fig. 4 shows a strong increase in the impact in the most unstable directions with 4D-Var). It is difficult to know exactly the reason for this unusual behaviour, but it seems that it can be explained by a very partial and downstream sampling of the sensitive area. This point is studied in section 5.

Figures 8(a) and (c), or Figs. 7(a) and (c), show similarities. One observes that a large fraction of the forecast error is explained by the projection of these errors onto the leading SVs (evolved at verification time). If the linear hypothesis is valid and if the model is perfect, this result implies that the initial errors that project onto an unstable subspace of small dimension are strongly correlated with the forecast errors at the
Figure 8. As Fig. 7, but for the Gulfstream-IV targeted data.
verification time. Consequently, the growth of the initial errors seems dominated by a relatively small number of unstable structures. The assimilation scheme plays an important role in the control of the errors in the unstable directions. With similar information (background field and observations), different assimilation schemes can produce significantly different initial conditions in these unstable directions. Consequently, the quality of the assimilation scheme is very important for the efficiency of adaptive observations. Moreover, the sampling of the sensitive area is also of extreme importance.

5. What does the influence of assimilation scheme depend on?

The influence of the assimilation scheme can depend on several factors. The first factor is the quality of the background field. Bergot (1999) has shown that an improvement in the forecast (due to the inclusion of FASTEX targeted observations) mainly occurs when the background field is quite bad. It would be helpful to test the influence of the assimilation scheme in these cases. The second factor is the dynamically unstable properties of the atmosphere. As shown by Hello et al. (2000), the construction of optimal initial conditions should be different in neutral or weakly unstable cases. One way of assessing this atmospheric property is to examine the first singular value that represents the strongest amplification rate of perturbations (under the linear hypothesis), which can be an estimation of the predictability of the atmosphere. And finally, the third factor is the sampling of the sensitive area. Aberson (2000) and Bergot et al. (1999) suggested that a partial sampling of the sensitive area can lead to a worsening of the forecast.

(a) Influence of the background-field quality

The goal of the analysis is to correct the background field in order to produce the best (global) initial conditions, and the best (global) forecast. Therefore, the quality of the background field plays an important role in the efficiency of adaptive observations. The scatter plots of the improvement of the forecast with both 3D-Var and 4D-Var as function of the background-field errors (not shown), lead to the same conclusions as the previous study with 3D-Var by Bergot (1999). For both assimilation schemes, the larger the forecast error from the background field, the more the forecast can be improved. Therefore, this result does not depend strongly on the assimilation scheme used, and seems robust. It seems that the impact of the assimilation system on the efficiency of adaptive observations does not depend strongly on the quality of the background field.

(b) Influence of the linear amplification rate

Numerous studies have shown that the growth of the initial difference between two forecasts, or of the initial errors, is dominated by a relatively small number of unstable structures (Gelaro et al. 1999; Gelaro et al. 2000; Bergot et al. 1999; Hello et al. 2000). These structures seem to be the most efficient ones by controlling the divergence between two forecasts. These unstable directions can be characterized by linear amplification rates, or singular values.

Figure 9(a) displays the difference in the impact (see section 3) between 3D-Var and 4D-Var as a function of the amplification rate of the first SV (the first singular value). The impact is mostly weaker with 4D-Var than with 3D-Var when the first singular value is stronger than 4. In weakly unstable cases (singular values smaller than 4) the spread is strong, and the influence of the assimilation scheme on the ability of adaptive observations to produce diverging forecasts is not clear.

Figure 9(b) compares the difference in forecast improvement between 3D-Var and 4D-Var, due to the inclusion of adaptive observations, as a function of the first singular
value. This figure shows that, when strong instabilities appear, the improvement of the forecast is larger with 4D-Var than with 3D-Var.

A comparison of Figs. 9(a) and (b) suggests that 3D-Var can lead to too strong an impact on the most unstable directions in the case of strong instabilities. In such cases, the improvement of the forecast is clearly larger with 4D-Var.
Figure 10. Cross-sections of the increments of temperature (K) for the Gulfstream-IV flight for IOP12 (a) with 3D-Var and (b) with 4D-Var (contour interval 0.3).

(c) Influence of the sampling strategy

The sampling of the sensitive area is probably one of the major problems to solve for operational feasibility of the targeted-observation concept. The results shown in section 4 suggest that the optimal sampling of a sensitive area strongly depends on the assimilation scheme used. Bishop et al. (2000) has obtained similar results in an idealized context, with an ensemble Kalman filter. A major difference between 4D-Var and 3D-Var is the influence of the dynamic evolution on the covariances during the assimilation period. While the covariances are isotropic with 3D-Var, they are affected by dynamic evolution during the six-hour assimilation window with 4D-Var. The 4D-Var structure functions are vertically tilted in a way that is consistent with the meteorological situation. However, this effect is rather weak over a six-hour assimilation window, as clearly illustrated by Rabier et al. (2000) with a single-observation experiment.
Figure 11. The influence of the sampling of the sensitive area on the improvement of the forecast (a) with 3D-Var and (b) with 4D-Var (○ small improvement or degradation, × degradation smaller than $-100$ J kg$^{-1}$, + improvement larger than $100$ J kg$^{-1}$).

As previously discussed in section 4, the GIV flight for IOP12 is an interesting case to study. The targeted data lie at the downstream border of the SV-based target (Fig. 5(b)). The increments for this case are plotted in Fig. 10(a) for 3D-Var and in Fig. 10(b) for 4D-Var. One remarks that the increments are slightly more vertically tilted with 4D-Var than with 3D-Var. This result is consistent with the single-observation
experiment of Rabier et al. (2000). The difference between the increments with 3D-Var and 4D-Var is weak. Nevertheless, a part of the difference is located in a strongly sensitive area, and can amplify and influence the forecast. This seems to be the case, and the 4D-Var forecasts are better than the 3D-Var ones. However, there still remain forecast errors in 4D-Var. This seems to indicate that 4D-Var can still benefit from better specification of the covariances (with a Kalman filter, for example).

In order to study the relationship between the sampling and the influence of the assimilation, the data have been classified according to their location with respect to the adjoint-based target. Following the maximum value (max) within the target area identified by the first two SVs, three classes of observations are defined:

- data in a null area—data in a location where the sensitivity is less than 0.01 × max;
- data at the border—data located where the sensitivity is less than 0.65 × max and larger than 0.01 × max;
- data in the core area—data at a place where the sensitivity is larger than 0.65 × max.

Figure 11 plots the improvement of the forecasts with 3D-Var and 4D-Var, as a function of the number of observations in the core and at the border of the sensitive area. When almost all the data are located in the border of the sensitive area (GIV data for IOP12—thirteen in the border region and one in the core; or L36 data for IOP15—seven in the border region and one in the core), the inclusion of these data leads to a worsening of the forecast for both assimilation schemes. However, the worsening is weaker with 4D-Var than with 3D-Var. It seems that both assimilation schemes cannot fully control the most unstable initial error in these cases (strong impact and worsening of the forecast). The positive influence of 4D-Var is clearer when more than three observations are present in the core of the sensitive area. In these sampling cases (L36 data for IOP17—ten in the border region and three in the core; L36 data for IOP13—eleven in the border region and five in the core; or L36 data for IOP16—eight in the border region and seven in the core), the inclusion of adaptive observations leads to a systematic worsening of the forecast with 3D-Var. With 4D-Var, this type of sampling has a small impact on the forecast quality. Finally, when the sensitive area is well sampled, the forecast is clearly improved with both 3D-Var or 4D-Var.

These results clearly demonstrate that the sampling of the sensitive area depends on the assimilation scheme used. If the observations are made at the border of the sensitive area, a approximate interpolation of these data can lead to indirect, but significant, errors in the unstable directions. The inaccuracies of the assimilation processes could, therefore, be detrimental through the introduction of small errors in the unstable directions, which by definition will amplify and worsen the forecast. This problem seems to be significantly reduced with 4D-Var. A precise assimilation scheme is doubtlessly necessary whenever the sensitive area is only partially sampled.

6. CONCLUSION

The concept of adaptive observations has been tested under operational conditions during FASTEX. The efficiency of adaptive observations is tested here with two data assimilation schemes, an incremental 3D-Var and an incremental 4D-Var. This allows the testing of the influence of the assimilation scheme on the efficiency of adaptive observations.

In a first step, the effect of the model deficiencies was not addressed, and the impact was defined as the difference between forecast experiments with and without targeted
FASTEX flights. A study of the impact suggests that the FASTEX adaptive observations significantly influence the forecasts, on average, for both assimilation schemes. For the GIV data, the magnitude of the mean impact of the targeted data at short range represents about 50% of the mean forecast error for both 3D-Var and 4D-Var. In contrast, the 4D-Var scheme significantly reduces the impact of the L36 data, at short range. This reduction is particularly clear when only a small portion of the SV-based target is sampled. Another interesting finding is that the impact at medium range is clearly reduced with 4D-Var, for both L36 and GIV data.

In a second stage, the improvement of the forecast quality is assessed, and forecasts started from analyses with and without targeted data are compared with analysed fields. With 3D-Var, the inclusion of targeted data leads to a blending of strong improvements and strong worsening, resulting in a weak improvement on average. In contrast, the assimilation of the FASTEX targeted observations with 4D-Var shows a mean improvement in the forecasts at all forecast ranges, for both aircraft. This result demonstrates that an accurate assimilation scheme is crucial for properly assimilating targeted observations. Another interesting result is that with 4D-Var, the inclusion of targeted data leads to a significant improvement of the short-term forecast, without producing a large variability of the medium-range forecast (in contrast to 3D-Var).

The positive influence of 4D-Var is particularly obvious when targeted observations are located at the border of the sensitive area only. In this case, 3D-Var can increase the initial errors in the most unstable directions, and consequently worsen the forecast. In the case of poor sampling of the sensitive area, the inaccuracies of the assimilation scheme are detrimental. A detailed sampling of the sensitive area is absolutely necessary with an approximate assimilation system, as previously demonstrated by Bergot et al. (1999). This constraint is reduced with 4D-Var.

Results presented here add to the evidence that the initial errors in an unstable subspace of small dimension play a key role on the forecast errors. The efficiency of adaptive observations clearly depends on the degree of sophistication of the assimilation scheme used to interpolate these data and to produce initial conditions. One possibility for making better use of adaptive observations is to develop for them a specific ‘targeted’ component of the assimilation scheme, by using SVs to define covariance structures (see, for example, the work of Hello et al. (2000)). This ‘targeted type’ of assimilation scheme seems to be essential for fully benefiting from adaptive observations. The current assimilation schemes are not able to produce efficient initial fields (i.e. with no error in the most unstable directions), and the assimilation errors can conceal the positive impact of the targeted data. However, the results presented here show that 4D-Var can exploit the dynamical processes for interpolating the information coming from adaptive observations more appropriately and more accurately than 3D-Var.

This work clearly demonstrates that the sampling of the sensitive area, and the assimilation scheme, are intimately dependent. In this way, it seems useful to include the assimilation scheme inside the definition of the adaptive observations. This would allow the definition of an optimal sampling of the target area associated with the assimilation scheme used subsequently to produce the initial conditions (Baker and Daley 2000; Doerenbecher and Bergot 2000). The definition of the sampling strategy strongly depends on the assimilation scheme used. An optimal sampling strategy of a sensitive area for a given assimilation scheme is probably not optimal for another assimilation scheme. One of the major conclusion of this study is that an optimal and universal targeting method, working for all assimilation systems, will probably never exist.
ACKNOWLEDGEMENTS

I am particularly grateful to Philippe Lopez, Florence Rabier, and one of the reviewers for helpful comments on the first version of this article. I also acknowledge the assimilation team from Météo-France, particularly Jean-Noël Thépaut for his friendly support during the experiments. I address special thanks to my colleagues from the GMME/RECYF* team for daily scientific discussion and technical support.

FASTEX was supported by the Programme Atmosphère et Océan à Moyenne Échelle of the Institut National des Sciences de l'Univers under contract 97/01, and by the European Commission under contract ENV4-CT96-0322.

REFERENCES


* Groupe de Modélisation a Moyenne Echelle/Recherches et Experiences sur les CYclogenases et les Frontogeneses.


