TECHNICAL MEMORANDUM

512

The value of targeted observations Part II: The value of observations taken in singular vectors-based target areas

Roberto Buizza, Carla Cardinali, Graeme Kelly and Jean-Noël Thépaut

Research Department

February 2007

To be submitted for publication in Q.J.Roy.Meteor.Soc.

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications.html

library@ecmwf.int

Abstract

Data-assimilation experiments have been run in seven different configurations for two seasons to assess the value of observations taken in target regions identified either using singular vectors (SVs) or randomly, and located over the Pacific or the Atlantic oceans. The value has been measured by the relative forecast error reduction in downstream areas, specifically a North-American region for targeted observations taken in the Pacific Ocean, and a European region for targeted observations taken in the Atlantic Ocean.

Overall, results have indicated (i) that observations taken in SV-target areas are more valuable than observations taken in randomly selected areas, (ii) that it is important that the daily set of singular vectors are used to compute the target areas, and (iii) that the value of targeted observations depends on the region, the season and the baseline observing system. *If the baseline observing system is data void over the ocean, then the average value of observations taken in SV-target areas is very high*. Considering for example winter 2004, SV-targeted observations over the Pacific (Atlantic) reduce the day-2 forecasts error of 500 hPa geopotential height forecasts in the verification region by 27.5% (19.1%), compared to 15.7% (14.9%) for observations taken in random areas. By contrast, *if the baseline observing system is data rich over the ocean, then the average value of observations taken in SV-target areas is rather small*. Considering for example winter 2004, it has been estimated that adding SV-targeted observations over the Pacific (Atlantic) would reduce, on average, the day-2 forecasts error in the verification region by 4.0% (2.0%), compared to 0.5% (1.7%) for observations in random areas. These average results have been confirmed by single-case investigations, and by a careful examination of time series of forecast errors.

These results indicate that more accurate assimilation systems that can exploit the potential value of localized observations are needed to increase the return on investment in targeting field experiments.

1. The value of targeted observations

In the past decade, field experiments have been organized to assess the impact of extra observations taken in some specific, case-dependent target areas identified using objective and subjective methods, on (mainly short-range) forecast accuracy. These campaigns, organized following a proposal at a workshop in 1995 (Snyder 1996), included, for example, in 1997 FASTEX (Fronts and Atlantic Storm-Track Experiment, Joly et al 1999), in 1998 NORPEX (NORth-Pacific Experiment, Langland et al 1999) and CALJET (California Land-falling JETs experiment, Ralph et al 1998), in 1999 and 2000 the Winter Storm Reconnaissance programs (WSR99 and WSR00, Szunyogh et al 2000) and in 2003 NATReC (Petersen et al 2003).

The general conclusions of these campaigns (see, e.g., Langland 2005 for a review of issues in targeted observing) were that in most of the cases the impact of the targeted observations was positive, and that on average the impact was small, with maximum error reductions in some specific cases for variables such as mean-sea-level-pressure of about 10-15%. For example, considering observations targeted using singular vectors in 5-cases of the FASTEX campaign, Montani et al (1999) concluded that the forecast impact was more evident if the verification area coincided with the region for which the singular vectors were optimized, and that in this case, short-range (up to day 2) prediction errors were reduced, on average, by 15%. Considering 70 cases of FASTEX, NORPEX, CALJET, WSR99 and WSR00 for which observations were targeted in areas identified using the Ensemble Transformed Kalman Filter technique (ETKF, Bishop et al 2001), Toth et al (2002) concluded that mean-sea-level-pressure forecast errors were reduced in ~65% of the cases, on average by ~10% in the verification region for which targeted observations were taken. In particular, Szunyogh et al (2000), who discussed 12 cases of the WSR00 campaign, concluded that the overall quality of the forecasts with the targeted data were improved in ~60% of the cases, during which an average forecast error reduction of 19% was detected for mean-sea-level-pressure forecasts in the 24-48 hour forecast range.

These studies were affected, by design and execution, by four major weaknesses:

- *Poor matching between the target area identified by the method under investigation and the area actually sampled by the extra observations*
  During field campaigns, it is often extremely difficult if not impossible to sample the target areas identified 'a-priori' by the objective method under testing, and thus sometimes these objectively-defined target areas were only partially sampled (e.g. because aircrafts could not fly the whole identified area), while at other times they were not sampled at all, but extra observations were taken in other regions (Cardinali and Buizza 2003).

- *Absence of a clean comparison between the impact of observations taken in objectively-defined target and randomly selected areas*
  To our knowledge, no study has yet compared the effect of observations taken in areas defined by an objective method, and observations taken in random areas, or in fixed areas, of similar size.

- *Large variation of target areas and number of extra observations taken in the target area*
  During these field campaigns, both the size of the target area and the number of extra observations varied substantially between the different cases. Thus, it was difficult to conclude what is the impact of daily taking extra N observation covering a target area of area K.

- *Low statistical significance due to the limited number of cases, and biased case selection*
  Most, if not all, of the previous works were based on few cases (say up to ~50), most of the times selected because they presented some potentially interesting developments. Since this selection procedure may introduce a bias in the data-set any statistically significant conclusion on the potential average impact of targeted observations could be drawn from these works.

This study aims to address these four major weaknesses by comparing data-injection and data-denial experiments designed so that for each case the targeted areas have the same size and coincide with the area identified by the objective targeting method under investigation. The impact of observations taken in objectively-defined target areas is also compared with observations taken in random or fixed areas with a similar size to the target areas. To be able to draw statistically significant conclusions with state-of-the-art systems, a very large number of cases (183, selected to cover different months of both the warm and the cold season) have been considered, and the European Centre for Medium-Range Weather Forecasts (ECMWF) data-assimilation and forecast experiments have been performed with the same horizontal resolution as the operational data-assimilation system the 21st of November 2000 and the 1st of February 2006 (Mahfouf and Rabier 2000) have been used:

- 4-dimensional variational assimilation system, with a 12-hour cycling

- Analysis resolution: outer loop $T_L511L60$, inner loops $T_L159L60$

- Forecast resolution: $T_L511L60$

- Model cycle 29r1 (operational between the 5^{th} of April and the 5th of June 2005)

The results of this study are discussed in three companion papers. The first of them (Kelly et al, 2007, hereafter KTBC07), investigated the value of observations taken over the Pacific ocean for 2-day forecasts verified over North-America, and of observations taken over the Atlantic ocean for 2-day forecasts verified over Europe. KTBC07 also compared the value of Pacific observations for medium-range forecasts verified over Europe, and discussed the sensitivity of the value of observations to the data-assimilation system. The three key conclusions of KTBC07 were that, on average, (i) ocean data are essential to have small day-2 forecast errors over the down-stream land, especially for the Pacific-North American region, (ii) the removal of observations taken over the Pacific has a very small impact on medium-range forecasts verified over Europe, and (iii) results strongly depend on the data-assimilation system used to assimilate the observations. On this latter point, KTBC07 showed that forecasts started from 3D-Var analyses are worse than forecasts started from 4D-Var analyses, and concluded that given the current observing system, the value of observations taken over the oceans is higher if a 3-dimensional variational assimilation system is used instead of a 4-dimensional system. In other words, a 4-dimensional variational system is more capable to propagate the information from data rich to data void areas, thus allowing for data gaps in a system otherwise overwhelmed by satellite observations.

This second paper focuses on 183 cases (winter 2003-04 and summer 2004), and discusses the value of observations taken in singular vector (SV) target areas over the ocean for 2-day forecasts verified over downstream regions over land. More specifically, the following questions are addressed:

- What is the 'value' of observations taken in a target region identified using SVs compared to the value of observations taken in a randomly chosen region? Is it important that SV-target regions are identified using the most recent analysis and forecast, or can an SV-average target region be used?

- Does the 'value' of observations depend on the region? What is the value of observations taken in SV-target regions of the Pacific Ocean for 2-day forecasts verified over North-America? How does it compare to the value of observations taken in SV-target regions of the Atlantic Ocean for 2-day forecasts verified over Europe?

The third companion paper (Cardinali et al 2007, hereafter CBKST07) considers the 183 cases discussed in the first two papers plus 100 extra cases, and investigates in details whether in some specific situations, such as the ones linked to extra-tropical cyclone transitions, the value of targeted observations can be different. Furthermore, CBKST07 presents a detailed analysis of the information content of targeted observations.

This paper is organized as follows. Section 2 presents the methodology used to define the data-injection and data-denial experiments, section 3 shows some examples of SV-target and random areas and section 4 discusses some statistics on the number of observations contained in each area. The main results of this work are presented in sections 5 and 6, and sub-sections, and finally the key conclusions of this work are discussed in section 7.

2. Methodology

Figure 1 illustrates the key concepts used in this paper to assess the value of targeted observations: T is the target area where observations should be taken to improve the day-d forecast inside the verification area Σ. Two sets of experiments have been run: one set was used to study the value of observations located in the Pacific Ocean for forecasts verified over North-America, and a second set was used to study the value of observations located in the Atlantic Ocean for forecasts verified over Europe. For both sets of experiments, the ocean is defined as the area between 30°N and 80°N latitude, while the two verification regions Σ have been defined as follows:

- European verification area: Σ_{EUR}=(10°W:25°E;35°N:60°N)

- North-American verification area: Σ_{NAM}=(125°W:90°W;35°N:60°N)

For each verification area, three types of target areas are considered:

- *SV-target areas*: for each initial date and time (d,h), the SV-target area (which has always the same size) is defined using the 10 leading SVs growing for 2 days [i.e. between (d,h) and (d+2,h)] with maximum final time total energy norm inside the verification region

- *SV-av-target* areas: for each initial date and time (d,h), the SV-av-target area is defined using the mean taken over the whole period under investigation of all the (d,h) SV-target areas (with the same size as the SV-target area)

- *Random areas*: for each initial date and time (d,h), the random area has an ellipsoidal shape centred randomly over the ocean and with the same geographical size as the SV-target area.

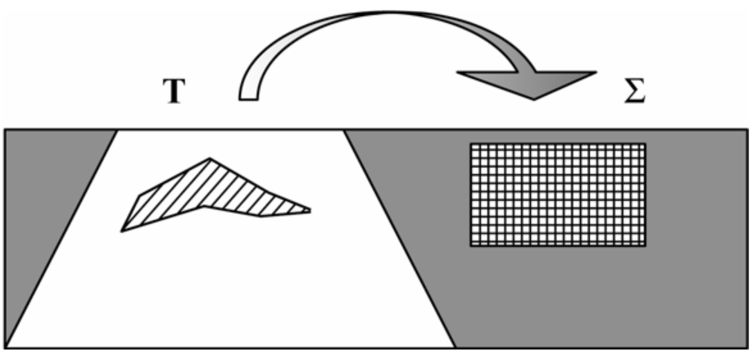

Figure 1. Schematic illustration of the concept of the 'value of targeted observations': the grey areas identify land, while the white region identify the ocean. T is the target area (e.g. identified using singular vectors) where observations should be taken to improve the d-day forecast inside the verification area Σ.

2.1 The concept of the 'value of targeted observations'

Consider, for example, experiments run to investigate the value of observations taken in the Pacific Ocean for day-2 forecasts over North America. Denote by REF a reference forecast experiment, for example the forecast started from the analysis generated using none of the observations available over the Pacific Ocean, and by X an experiment generated using the observations taken inside SV-target regions over the Pacific Ocean. Denote by $f_X(d)$ the day-d forecasts of field f given by the X-experiment, and by $sc_V\left[f_X(d)\right]$ a metric used to compute the forecast error inside region V: for example, set $sc_V\left[f_X(d)\right] \equiv rmse_{NA}\left[Z500(d2)\right]$, i.e. define sc(..) as the root-mean-square-error (*rmse*), and $f_X(d)$ as the 500 hPa geopotential height (Z500) day-2 forecast over North America. Define the relative forecast error $v\left(f_X, d, sc_V\right)$ as the relative difference between the forecast errors of experiment X and the reference REF, computed using the accuracy measure sc(..) inside the region V:

$$v\left(f_X, d, sc_V\right) = \frac{sc_V\left[f_X(d)\right] - sc_V\left[f_{REF}(d)\right]}{sc_V\left[f_{REF}(d)\right]}$$

For example, considering the 500 hPa geopotential height and the root-mean-square error, the relative forecast error is computed as

$$v(f_X, d2, sc_V) \equiv v(Z500_X, d2, rmse_{NA}) = \frac{rmse_{NA}[Z500_X(d2)] - rmse_{NA}[Z500_{REF}(d2)]}{rmse_{NA}[Z500_{REF}(d2)]}$$

Suppose that the X-experiment has been run by removing all Pacific Ocean observations apart for the ones located in SV-target regions, and that the reference REF has been run without any observations. Then, the relative forecast error can be used as a measure of the ‘value of observations taken in the SV-target region if no observations where available in the Pacific Ocean’. The value measured considering the root-mean-square-error of the 500 hPa geopotential height has been translated into gains in forecast hours $g\left(Z500_X, d2, rmse_{NA}\right)$ by normalizing it by the difference between the RMS errors at forecast days 3 and 1 (the multiplication by 48 transforms the normalized gain in hours):

$$g(Z500_X, d2, rmse_{NA}) \equiv 48 * \frac{rmse_{NA}[Z500_X(d2)] - rmse_{NA}[Z500_{REF}(d2)]}{rmse_{NA}[Z500_{REF}(d3)] - rmse_{NA}[Z500_{REF}(d1)]}$$

Similar definitions have been applied for to the Atlantic-European experiments, with the root-mean-square error measured considering the European verification region instead of the American one.

2.2 Experiments’ description

Figure 2 shows a schematic of the 6 types of data-assimilations that have been performed:

- SEAIN: all observations available over the ocean have been used
- SEAOUT: none of the observations available over the ocean have been used
- SVIN: only observations available in the SV-target area have been used
- SVOUT: all observations apart for the ones in the SV-target area have been used
- RDIN: only observations available in randomly defined area have been used
- RDOUT: all observations apart for the ones in randomly defined area have been used

In experiments SVIN and SVOUT, at each data-assimilation cycle observations are either injected (in SVIN) or removed (in SVOUT) in the same target area defined using SVs (noting that SVs are updated every 12 hours). Similarly, in experiments RDIN and RDOUT observations are injected or removed in the same random area.

For a limited subset of cases, the following experiment has also performed:

- SVavIN: as SVIN but always using only the observations available in the SV-average target area, defined by averaging the SV-areas computed for each day of the period under investigation

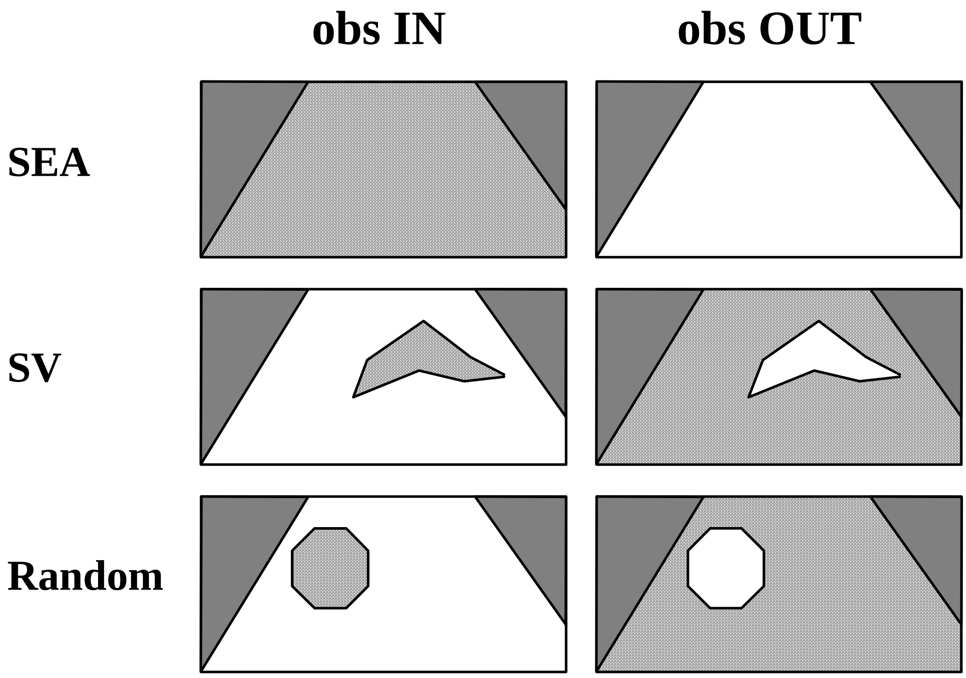

Figure 2. Schematic of the data-assimilation experiments: the full-grey triangles identify land; the white areas with black dots identify the areas where observations have been used, while the white areas identify the areas where observations are not used.

2.3 Definition of the SV-target areas

Singular vectors (Buizza and Palmer 1995) identify the perturbations growing during a finite time interval, called the optimization time interval, with the largest amplification rate measured using a defined norm. Following their successful use in the ECMWF Ensemble Prediction System (EPS, Molteni et al 1996), in January-February 1997 targeted SVs were used for the first time to define regions where extra observations to be taken during the FASTEX (Fronts and Atlantic Storm-Track Experiment) field campaign (Buizza and Montani 1999, Montani et al 1999).

In this work, the optimization time interval has been set to 2 days, and the total energy metric has been used to measure growth (see Palmer et al 1998, for a discussion of the SVs sensitivity to the choice of the metric). Following Buizza and Montani (1999), once the 10-leading targeted SVs with maximum energy over verification area Σ have been computed (see Appendix A for a detailed description of the methodology used to compute the SVs), the SV-target area has been defined using the weighted-average, vertically-integrated total energy function :

$$f_{\Sigma}(s) = \frac{1}{10}\sum_{j=1}^{10}\frac{\sigma_j}{\sigma_1} f_{j,\Sigma}(s)$$

$$f_{j,\Sigma}(s) = \int TE_{j,\Sigma}(s,p)dp$$

where σ_j is the singular value of the j-th SV, $TE_{j,\Sigma}(s,p)$ is the total energy of the j-th SV at the grid point s, with the grid points defined on a regular 2° degree latitude-longitude grid (this coarse resolution reflects the fact that SVs have been computed with a T63L40 resolution), and p is the pressure level.

To localize the SV-target areas only over the ocean, the following step functions have been used:

- LSM(s)=1 if s is a sea grid point, and LSM(s)=0 if s is a land grid point
- OPAC(s)=1 if the grid point has latitude between 30°N and 80°N and longitude between 140°E and 230°E, and OPAC(s)=0 otherwise
- OATL(s)=1 if the grid point has latitude between 30°N and 80°N and longitude between 80°W and 10°W, and OATL(s)=0 otherwise

Using these functions, 'Pacific ocean' refers to the area defined by the grid points for which $LSM(s)\cdot O_{PAC}(s)\equiv 1$, and 'Atlantic ocean' refers to the area defined by the grid points s for which $LSM(s)\cdot O_{ATL}(s)\equiv 1$

The SV-target areas over the Pacific have been defined as the 100 grid points over the ocean with the largest value of the product of the function $f_{NAM}(s)$, i.e. function $f(s)$ computed using SVs with maximum total energy at final time over the North-American region Σ_{NAM}, and masks LSM and O_{PAC}:

$$t_{NAM}(s) = f_{NAM}(s)\cdot LSM(s)\cdot O_{PAC}(s)$$

Similarly, the SV-target areas over the Atlantic have been defined as the 100 grid points (thus, each target area had an area of 20°-degrees-squared) over the ocean with the largest value of the product of the function computed using SVs with maximum total energy at final time over the European region Σ_{EUR}, and masks LSM and O_{ATL}:

$$t_{EUR}(s) = f_{EUR}(s)\cdot LSM(s)\cdot O_{ATL}(s)$$

The SVs and the corresponding target areas have been computed every 12-hours, with a configuration very similar to the one used in the ECMWF ensemble prediction system but with a higher resolution: T63L40 resolution, 48-hour optimization time interval, total energy norm and (dry) simplified physics (Buizza 1994b).

2.4 Definition of the random areas

The random areas have been defined to be approximately circular, with the centre of the circle randomly selected inside the ocean region of interest, and large enough to contain the same number of grid points as the SV-target areas. More precisely, once the centre of the random circle has been selected, the random area is defined as the same number of grid points as the SV-target area (i.e. 100) located inside a circle centred at the randomly selected centre, and characterized by a unitary value of the land-sea mask $LSM(s)$and the ocean masks $O_\Sigma(S)$. Thus, by definition, the random areas are perfectly circular only if the circle is completely inside the ocean region and does not include any land-grid point. This choice to work with a circular geometry may be a weakness of this study, but it has been thought that such a choice would be more justifiable than choosing other elongated, e.g. ellipsoidal, more ad-hoc geometries. Compared to any other 2D-shape, a randomly-positioned circle has the advantage that it is defined by only two parameters (i.e. the coordinates of its centre).

2.5 List of case studies

The results discussed in this work are based on experiments performed for the most recent winter and summer seasons (since these experiments started in September 2004, the last summer and winter periods available were summer 2004 and winter 2003-2004). This selection had two advantages: firstly, it made it easier to compare results from these data-assimilation experiments with operational analysis and forecasts since the same model and configuration were used, and secondly, it guaranteed that that both the model and the observation systems used in the experiments were the state of the art. More precisely, these two periods are:

- DJF04: from 1 December 2003 to 29 February 2004 (91 cases)
- JJA04: from 1 June to 31 August 2004 (92 cases)

Given the fact that all but SEAIN experiments have been run twice, the total number of T_L511L60 12-hourly-cycling 4D-Var assimilation days run to produce all the data discussed in this work amount to 2379 days (13 assimilation experiments run for 183 days), i.e. 6.5 years!

3. Examples of SV-target and random areas

Figure 3 shows two examples of SV-target and random areas used in the Pacific-North American SVIN and RDIN experiments started at 12 UTC of the 1st and the 10th of December 2003, and the SV-average target area used in the SVavIN experiment. In the first case, the two areas do not overlap, while in the second case the two areas have some grid points in common over the western Pacific. Figure 4 shows the SV-target and random areas used in the corresponding Atlantic-European SVIN and RDIN experiments started at 12 UTC of the 1st and the 10th of December 2003. In this case, since the Atlantic Ocean basin is smaller, the chance that the two areas have some grid points in common is higher: indeed, the two areas have some grid points in common in both cases.

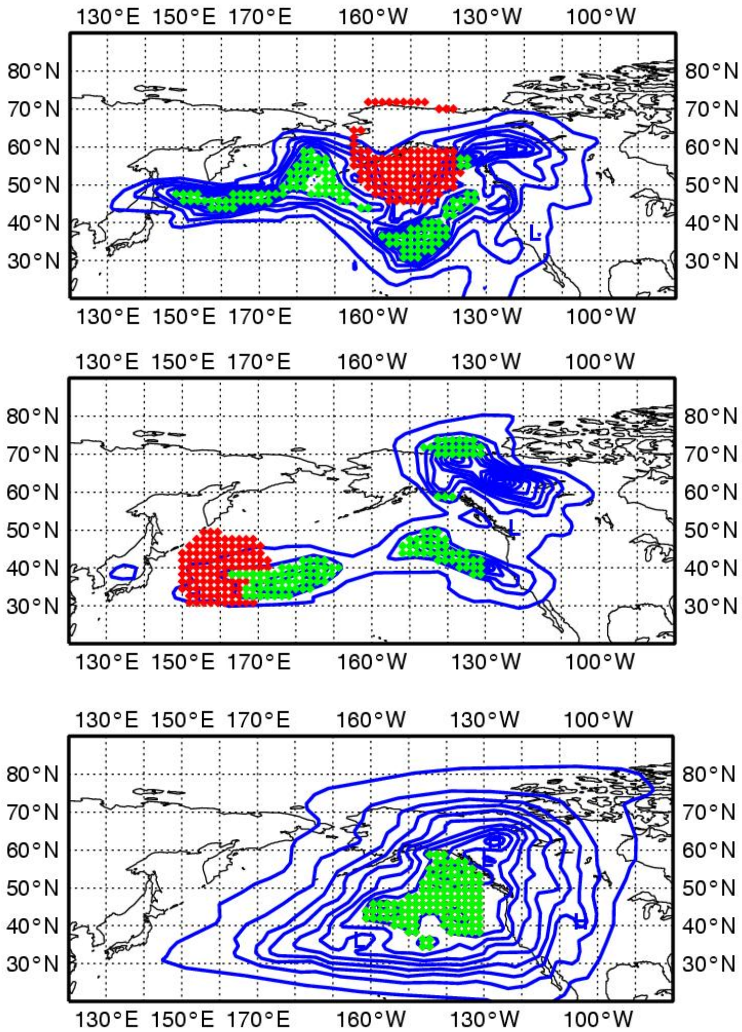

Figure 3. SV-target (green symbols) and random (red symbols) areas for the Pacific-North American region for the 12 UTC of 1 December (top panel) and 10 December (middle panel) 2003, and SV-average target area for winter 2004 (bottom panel). In each panel, the contour isolines show the weighted-average, vertically-integrated total energy of the 10 leading SVs.

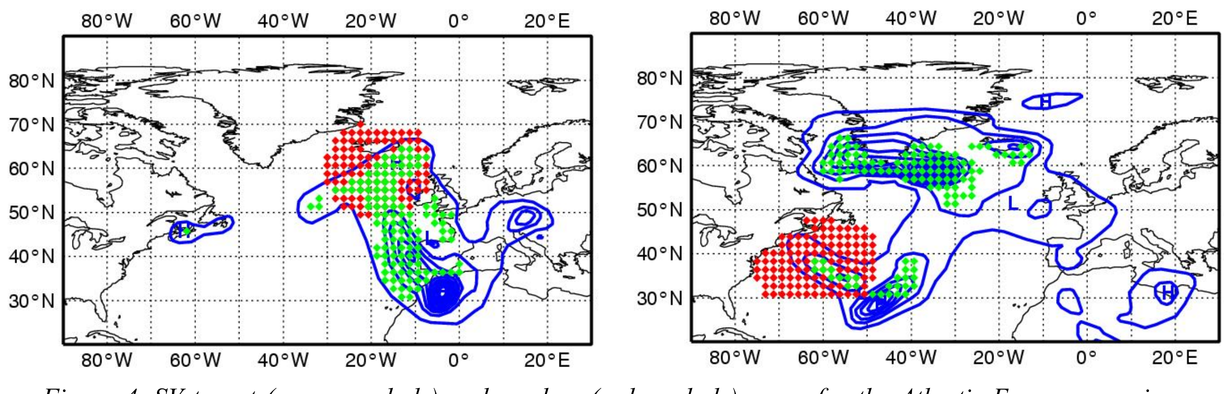

Figure 4. SV-target (green symbols) and random (red symbols) areas for the Atlantic-European region for the 12 UTC of 1 December (left panel) and 10 December (right panel) 2003. In each panel, the contour isolines show the weighted-average, vertically-integrated total energy of the 10 leading SVs.

4. Number of observations contained in the SV-target and random areas

It is interesting to compare the percentage of observations located over the two Oceans: Table 1 shows, for example, that the Pacific ocean (as defined in section 2.3) contains on average 17.4% of AMSU-A channel-6 observations, and 7.4% of 250 hPa temperature observations from aircrafts, while the Atlantic ocean contains, respectively, 13.8% and 9.8% observations. The fact that the number of AMSU-A observations is smaller for the Atlantic Ocean is due to the fact that this ocean basin is 20° degree narrower in longitude (see section 2.3).

Experiment	Total # AMSU-A ch-6 obs [#(…)/#(ocean obs)]	Total # aircraft obs [#(…)/#(ocean obs)]
Control (North of 20°N)	493450	901620
Pacific ocean	85960 [17.4% of control]	66900 [7.4% of control]
Pacific SV-target area	6800 [7.9% of Pac ocean]	14550 [21.7% of Pac ocean]
Pacific random area	6770 [7.8% of Pac ocean]	13390 [20.0% of Pac ocean]
Atlantic whole ocean	65830 [13.4% of control]	88190 [9.8% of control]
Atlantic SV-target area	8410 [12.7% of Atl ocean]	13810 [15.6% of Atl ocean]
Atlantic random area	6910 [10.5% of Atl ocean]	13300 [15.1% of Atl ocean]

Table 1. Average total number of AMSU-A channel-6 (second column) and of aircraft 250 hPa temperature observations (third column), and relative number of observations contained in the ocean target-areas (in square brackets) for the Pacific-North American and the Atlantic-European experiments.

It is also important to verify that the number of observations contained inside an SV-target or a random area is, on average, very close. Table 1 shows that, on average and for each ocean basin, the percentage of ocean observations contained in the SV-target or random areas are very close; in other words, on average, the same amount of observations are injected or removed in the SVIN/SVOUT or RDIN/RDOUT experiments.

5. Northern Hemisphere winter results (DJF04)

This section discusses results from the winter 2003-2004 (December 2003 and January-February 2004, DJF04) experiments, organized in the following way: first, section 5.1 and sub-sections summarize the results of the data-injection experiments (SVIN and RDIN) run for both the Pacific-North-American and the Atlantic-European regions, then section 5.2 compares results obtained using SV-average and SV-target regions (SVavIN and SVIN) for the Pacific-North American region, and finally section 5.3 and subsections summarize the results of the data-denial experiments (SVOUT and RDOUT) for both regions. For each region, experiments are compared to two references: SEAIN, the experiment run using all ocean data, and SEAOUT, the experiment run without data over the ocean.

5.1 DJF04 data-injection experiments: SV-target versus random areas

5.1.1 DJF04 Pacific-North America data-injection experiments

The top panel of Fig. 5 shows the average score of experiments SVIN, RDIN, SEAIN and SEAOUT run for the Pacific-North American region between forecast day 1 and 4. The comparison of SEAIN and SEAOUT scores indicates that if no observations are taken over the Pacific Ocean, the forecast quality is reduced dramatically, and the comparison of SVIN and RDIN scores indicates that taking observations in the SV-target ocean regions has a larger, positive impact than if observations are taken randomly in ocean regions with a similar size.

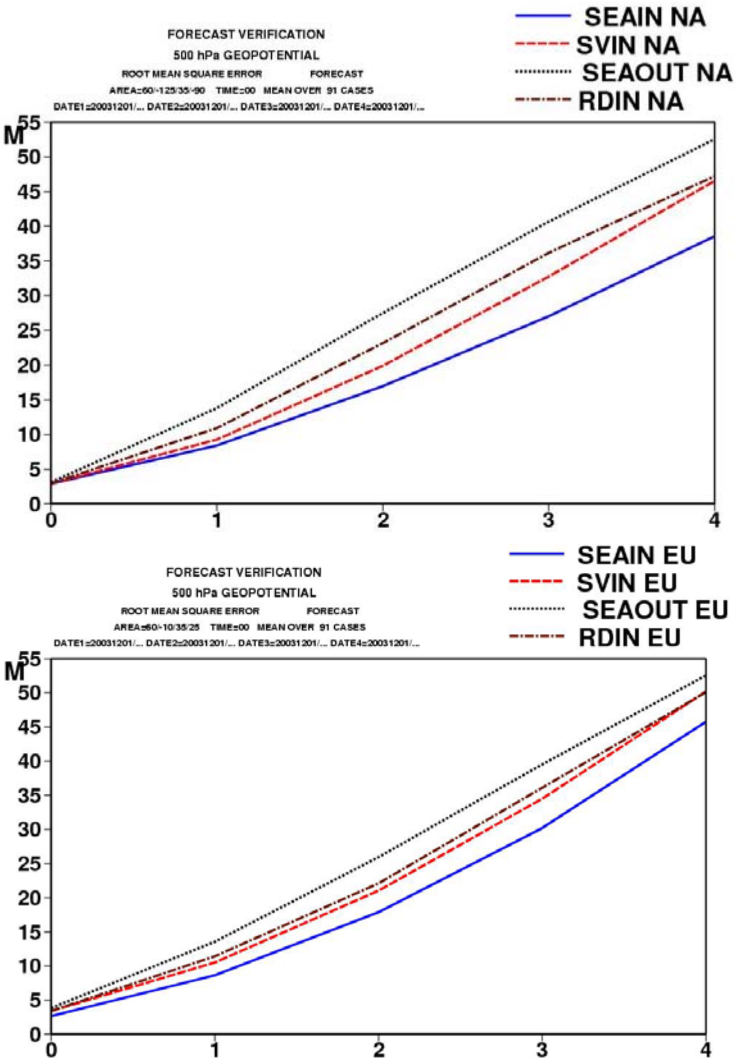

Figure 5. DJF04-average root-mean-square error of 500 hPa geopotential height forecasts of data-injection experiment SEAIN (control, blue line), SEAOUT (dotted green line), SVIN (red line) and RDIN (chain-dashed brown line) run for the Pacific-North American region (top panel) and the Atlantic-European region (bottom panel).

Figure 6 and 7 show the 'value' (as defined in section 2.1) of observations considering the 1000 and the 500 hPa geopotential height fields, and the 850 hPa temperature and wind-vector winds. Considering, for example, the 500 hPa geopotential height day-2 forecast, Fig. 6 shows that taking observations over the whole Pacific Ocean reduces, on average, the forecast error by 38.3%, taking observations only in SV-target areas by 27.5%, while taking observations in random areas by 15.7%. In other words, using a small number of observations taken in SV-target areas (e.g. 7.9% of all AMSU-A channel-6 observations available over the Pacific, see section 4 and Table 1) is enough to reduce the forecast error by 27.5%, i.e. to recover 71.8% (=0.275/0.383) of the error reduction induced by all the Pacific observations. Furthermore, Fig. 6 shows that observations taken in the SV-target areas have a higher 'value' (27.5% compared to 15.7%, i.e. almost double) than the observations taken in random areas of equal size. Similar considerations can be drawn by considering the 1000 hPa geopotential height, and the 850 hPa temperature and wind (Fig. 7).

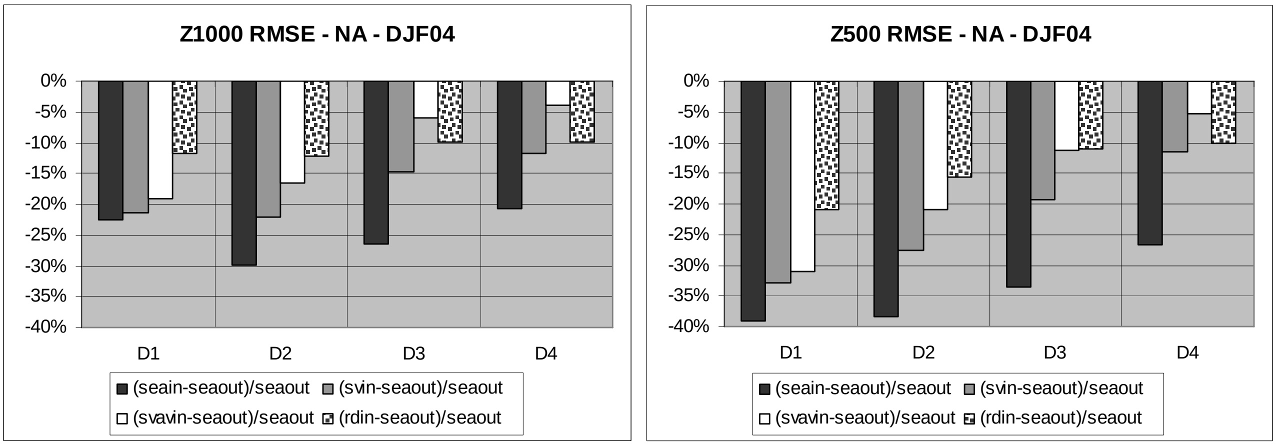

Figure 6. DJF04 Pacific-North America data-injection experiments.

Value $v(f_X, f_{SEAOUT}, d, rmse_{NA}) = \dfrac{rmse_{NA}[f_X(d)] - rmse_{NA}[f_{SEAOUT}(d)]}{rmse_{NA}[f_{SEAOUT}(d)]}$ *of Pacific observations (CON, black bars), and of targeted observations taken in SV-target (SVIN, grey bars), SV-climatological targets (SVavIN, white bars) and random (RDIN, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 over North America for the 1000 hPa (left panel) and the 500 hPa (right panel) geopotential height.*

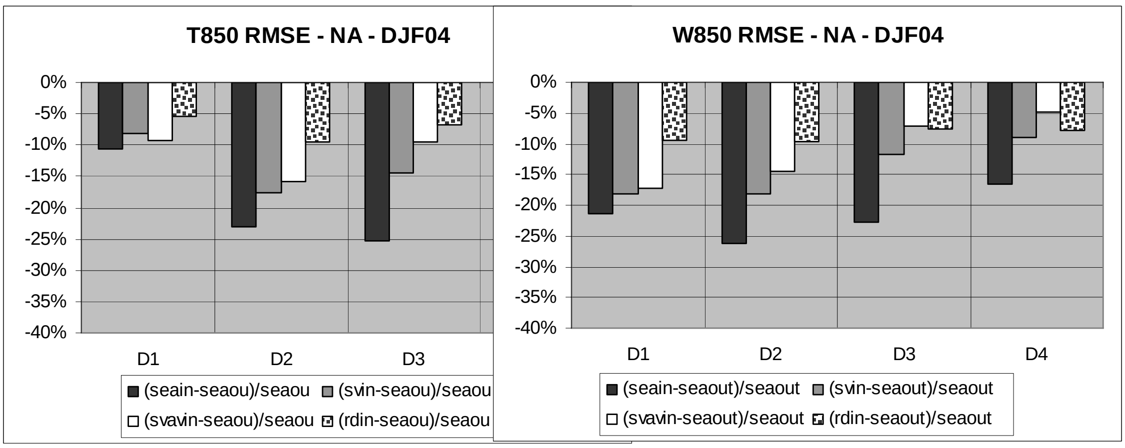

Figure 7. DJF04 Pacific-North American data-injection experiments.

Value $v(f_X, f_{SEAOUT}, d, rmse_{NA}) = \dfrac{rmse_{NA}[f_X(d)] - rmse_{NA}[f_{SEAOUT}(d)]}{rmse_{NA}[f_{SEAOU}T(d)]}$ *of Pacific observations (CON, black bars), and of targeted observations taken in SV-target (SVIN, grey bars), SV-climatological targets (SVavIN, white bars) and random (RDIN, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 over North America for the 850 hPa temperature (left panel) and wind (right panel).*

Considering again the top panel of Fig. 5, it is interesting to point out that the difference between the effects of observations taken in the SV-target or the random areas decreases after forecast day 3. For the 500 hPa geopotential height, the differences between these two sets of experiments are statistically significant at the 0.1% level at forecast days 1 and 2, at 0.2% level at forecast day 3 and 0.5% at forecast day 4. This decreasing difference can be explained by the fact that the SVs have been optimized to grow over a 48-hour time interval inside the North-American region: after 48-hours, the SVs loose optimality because they propagate and move outside the verification region.

The top panel of Fig. 8 shows the time-series of the 48-hour forecast error of SEAIN, SVIN and RDIN. It is interesting to note that apart for two periods (28-29 Dec and 21 Feb), SVIN either outperforms or has a similar error to RDIN, and that for many cases SVIN has a similar error to SEAIN. This figure confirms the average results, and it shows that for this season, the impact of removing/re-injecting observations in the Pacific Ocean can be detected in many cases (this can be seen, e.g., by comparing RDIN and SEAIN: RDIN is significantly worse in at least 6 periods).

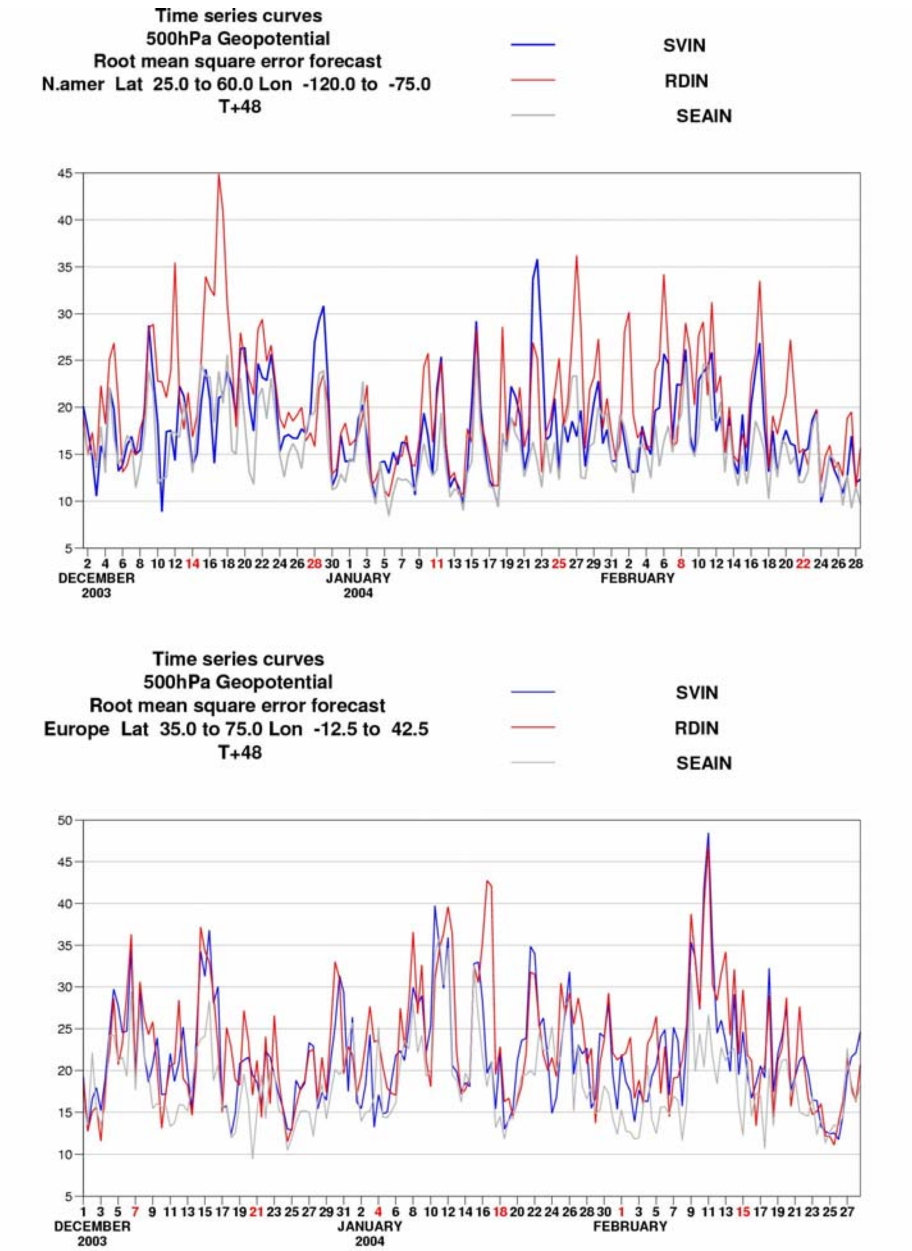

Figure 8. DJF04 time series of 48-hour forecast error of experiment SEAIN (grey line), and Pacific injection experiments SVIN (blue line) and RDIN (red line) verified over North America (top panel), and of experiment SEAIN, and Atlantic injection experiments SVIN and RDIN verified over Europe (bottom panel).

5.1.2 DJF04 Atlantic-Europe data-injection experiments

Results for the Atlantic-European region (Fig. 5, bottom panel) are qualitatively similar to the ones of the Pacific-North American region (Fig. 5, top panel) but quantitatively different. A possible explanation of these results is the fact that the Atlantic Ocean is smaller than the Pacific Ocean: in fact, the Atlantic oceanic region used in these experiments is 20° degree smaller in longitude than the Pacific oceanic area (see section 2.3). By consequence, the upstream area data-rich of good quality observations (North-America for the Atlantic Ocean, and China-Japan for the Pacific Ocean) is closer to the verification region.

Figures 9 and 10 show the 'value' of observations taken in the different regions. Considering, as before, the 500 hPa geopotential height day-2 forecast, Fig. 9 shows that taking observations over the whole Atlantic Ocean reduces, on average, the forecast error by 31.2%, taking observations only in SV-target areas by 19.1%, while taking observations in random areas by 14.9%. In other words, using a small number of observations taken in SV-target areas (e.g. 12.7% of all AMSU-A channel-6 observations available over the Atlantic, see section 4 and Table 1) is enough to reduce the forecast error by 19.1%, i.e. to recover 61.1% (=0.191/0.312) of the error reduction induced by all the Pacific data. Figure 9 also confirms another conclusion drawn from the Pacific-North American experiments: observations taken in the SV-target areas have a higher 'value' (19% compared to 15%) than the observations taken in random areas of equal size. Similar considerations can be drawn by considering the 1000 hPa geopotential height, and the 850 hPa temperature and wind (Fig. 10).

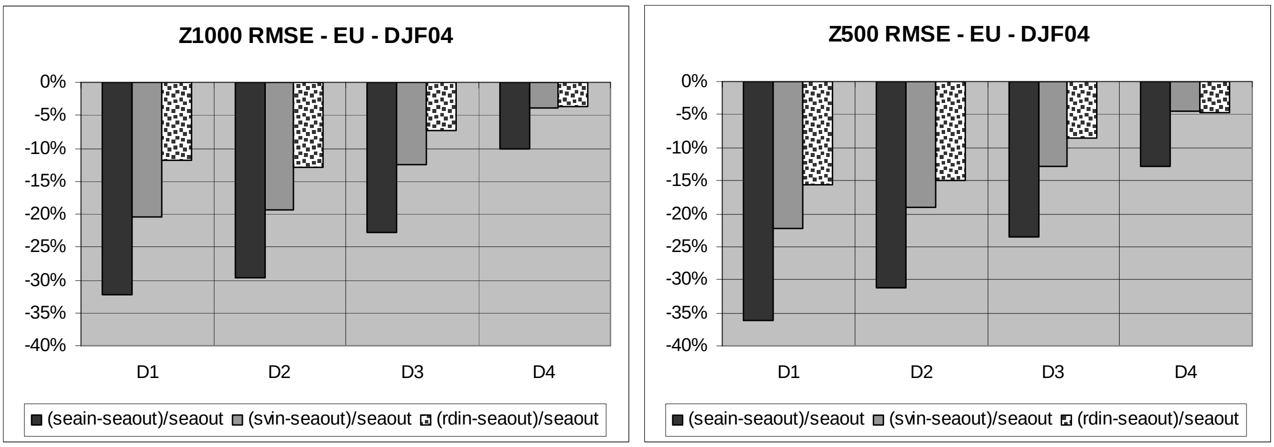

Figure 9. DJF04 Atlantic-European data-injection experiments. Value $v(f_X, f_{SEAOUT}, d, rmse_{EU}) = \frac{rmse_{EU}[f_X(d)] - rmse_{EU}[f_{SEAOUT}(d)]}{rmse_{EU}[f_{SEAOUT}(d)]}$ *of Atlantic observations (CON, black bars), and of targeted observations taken in SV-target (SVIN, grey bars) and random (RDIN, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 over Europe for the 1000 hPa (left panel) and the 500 hPa (right panel) geopotential height.*

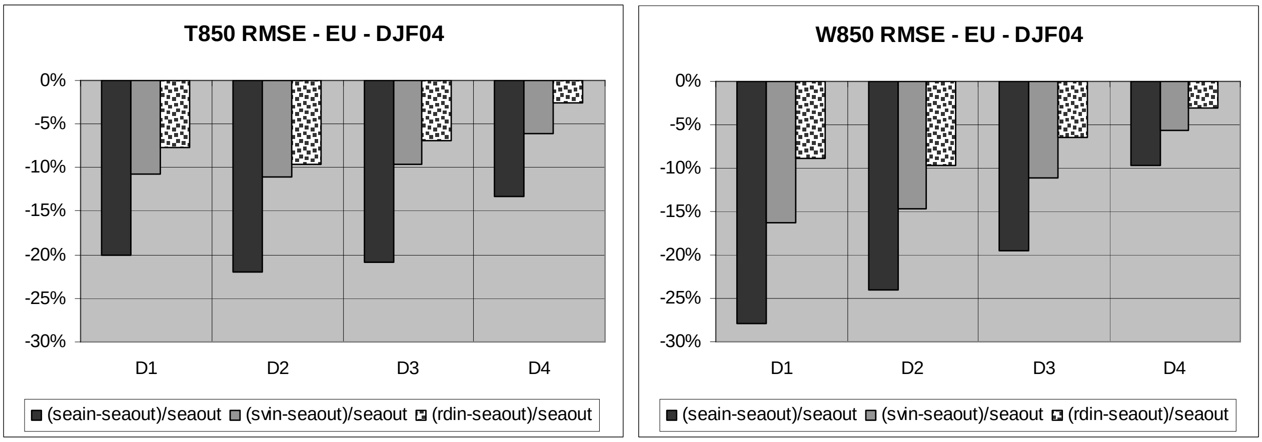

Figure 10. DJF04 Atlantic-European data-injection experiments. Value $v(f_X, f_{SEAOUT}, d, rmse_{EU}) = \frac{rmse_{EU}[f_X(d)] - rmse_{EU}[f_{SEAOUT}(d)]}{rmse_{EU}[f_{SEAOUT}(d)]}$ *of Atlantic observations (CON, black bars), and of targeted observations taken in SV-target (SVIN, grey bars) and random (RDIN, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 over Europe for the 850 hPa temperature (left panel) and wind (right panel).*

Note that, as it was the case for the Pacific-North American experiments, the difference between the effect of observations taken in the SV-target or random areas decreases after forecast day 3, again linked to the fact that the SVs have been optimized to grow over a 48-hour time interval inside the European region. Considering, for example, the 500 hPa geopotential height, differences between the SVIN and RDIN experiments are statistically significant at the 1.0% level at forecast day 1, at the 10.0% level at forecast day 2, at 5.0% level at forecast day 3, and are not statistically significant at forecast day 4.

The bottom panel of Fig. 8 shows the time-series of the 48-hour forecast error of SEAIN, SVIN and RDIN: only for three cases (15 Dec, 10 and 20 Jan) SVIN performs slightly worse than RDIN. This figure confirms the average results, in particular the fact that the difference between SVIN and RDIN is smaller than for the Pacific/North American region (RDIN is significantly worse in at least 7 periods).

5.2 DJF04 data-injection experiments: SV- versus SV-average-target areas

To assess whether it is important that the SVs of the day rather than SVs computed for another day or for an average ‘climatological’ field are used to define the SV-target area, SVavIN experiments have been run for the Pacific-North American region. Figures 6 and 7 show that the value of observations taken in the SVavIN-target areas is between the value of the observations taken in the SVIN-target and the random areas. Considering, again, the 500 hPa geopotential height day-2 forecast, Fig. 6 shows that taking observations in SVav-target areas reduces the forecast error by 20.9%, compared to 27.5% for observations taken in the SV-target areas and 15.7% for the observations taken in random areas. Table 2 lists these results, expressed both in terms of normalized differences and in corresponding hours of forecast gains (see section 2.1 for its definition). Table 2 indicates that the forecast error can be reduced more if the target regions are computed daily, i.e. they are computed for the daily atmospheric flow. Considering the 500 hPa geopotential height, the differences between SVIN and SVavIN experiments are not statistically significant at forecast day 1, while they are statistically significant at the 0.1% level at forecast day 2 and 3, and at the 1.0% level at forecast day

4. (SVavIN experiments were not performed for the Atlantic-European regions simply because of lack of computer resources, but it: it is expected that such a comparison would lead to qualitatively similar results.)

Value of targeted observations measured using day-2 Z500 fcs	Verified over North America	
	Winter (DJF04)	Summer (JJA04)
SEAOUT	27.49 m	16.57 m
SEAIN	16.96 m	11.96 m
SVIN	19.93 m	13.90 m
RDIN	23.18 m	14.61 m
SVOUT	18.06 m	12.20 m
RDOUT	17.11 m	11.81 m
SVavIN	21.75 m	--
(SEAIN-SEAOUT) / SEAOUT	38.3% (18.8h)	27.8% (13.5h)
(SVIN-SEAOUT) / SEAOUT	27.5% (13.5h)	16.0% (7.8h)
(RDIN-SEAOUT) / SEAOUT	15.7% (7.7h)	11.8% (5.7h)
(SVavIN-SEAOUT) / SEAOUT	20.9% (10.3h)	--
(SEAIN-SVOUT) / SEAOUT= [(SEAIN-SEAOUT) / SEAOUT]-[(SVOUT-SEAOUT) / SEAOUT]	4.0% (2.0h)	1.3% (0.7h)
(SEAIN-RDOUT) / SEAOUT= [(SEAIN-SEAOUT) / SEAOUT]-[(RDOUT-SEAOUT) / SEAOUT]	0.5% (0.3h)	-1.1% (-0.4h)
(SVOUT-SEAIN) / SEAIN	6.5%	2%
(RDOUT-SEAIN) / SEAIN	0.9%	-1.5%

Table 2. Average value of observations taken in different target areas, measured considering the root-mean-square-errors (RMSE) of the day-2 500 hPa geopotential height forecasts over North America. The first seven rows list the RMSE of the different forecasts (m). The other rows show the differences, in %, between normalized-RMSEs, with the normalization done using the RMSE of the SEAOUT experiment for all but the last two rows (with values in brackets are expressed in terms of hours of predictability gains)where the RMSE of SEAIN has been used.

5.3 DJF04 data-denial experiments: SV-target versus random areas

In this section, the SEAIN experiment (i.e. analysis obtained with a data-rich observational system) is used as a reference rather than SEAOUT, and the value of observations located in SV-target or random areas is assessed using data-denial experiments. Using SEAIN as a reference makes it possible to address the following question: if the ocean is well observed, would it make any difference if observations are removed in some specific areas? This is a different question to the one that targeting experiments aim to address, namely: if the ocean is well observed, would it make any difference if extra observations are taken in some specific areas? But since in this latter case the reference would be even data-richer than SEAIN, in the hypothesis that the characteristics (type, quality, content of information) of future extra observations is similar to the characteristics of the observations removed, these results should provide an upper bound of the impact that taking extra oceanic targeted observations can have on forecasts verified over downstream regions.

5.3.1 DJF04 Pacific-North America data-denial experiments

The top panel of Fig. 11 shows the average score of experiments SVOUT, RDOUT, to be compared with the two reference experiments SEAIN and SEAOUT, run for the Pacific-North American region between forecast day 1 and 4. Results indicate that RDOUT and SEAIN have almost identical scores, indicating that removing data in the random areas has no detectable impact on forecast error over North America. By contrast, the comparison of SVOUT and SEAIN indicates that removing data from the SV-target areas has a small but detectable impact.

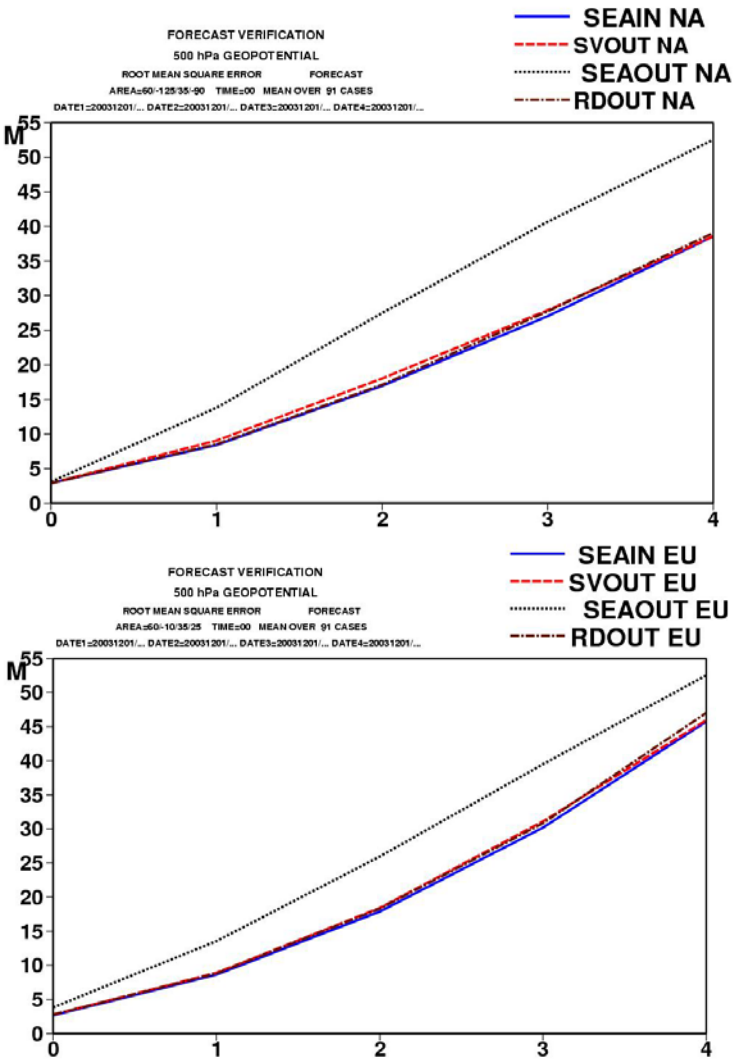

Figure 11. DJF04-average root-mean-square error of 500 hPa geopotential height forecasts of data-denial experiment SEAIN (control, blue line), SEAOUT (dotted green line), SVOUT (red line) and RDOUT (chain-dashed brown line) of the Pacific-North American (top panel) and Atlantic-European (bottom panel) experiments.

The top panel of Fig. 12 shows the 'value' of observations removed from the SV-target and the random areas measured considering the 500 hPa geopotential height fields (results for the other three variables considered in the previous sections are very similar, and are not shown). If we consider, for example, the day-2 forecasts, Fig. 12 shows that taking observations over the whole Pacific Ocean or taking observation over the whole Pacific Ocean apart for random areas reduces the forecast error by a very similar same amount, 38.3% and 37.8% (Table 2). By contrast, Fig. 12 shows that taking observation over the whole Pacific Ocean apart for SV-target areas reduces the forecast error only by 34.3%. Thus, adding observations in SV-target areas (i.e. comparing experiments SEAIN and SVOUT) has a small, but detectable value, since the forecast error is further reduced by 4%, from 34.3% to 38.3%. The differences between SEAIN and SVOUT are statistically significant at the 0.1% level at forecast days 1 and 2, at 5.0% level at forecast day 3 and are not statistically significant at forecast day 4. The differences between SVOUT and RDOUT are statistically significant at 0.1% level at forecast days 1 and 2, and are not statistically significant afterwards. As it was the case for the corresponding data-injection experiments, these data-denial experiments confirm that observations taken in SV-target areas have a higher 'value' than observations taken in random areas, and provide an estimate of the sensitivity of the value of these observations to the baseline observing system that is used as reference.

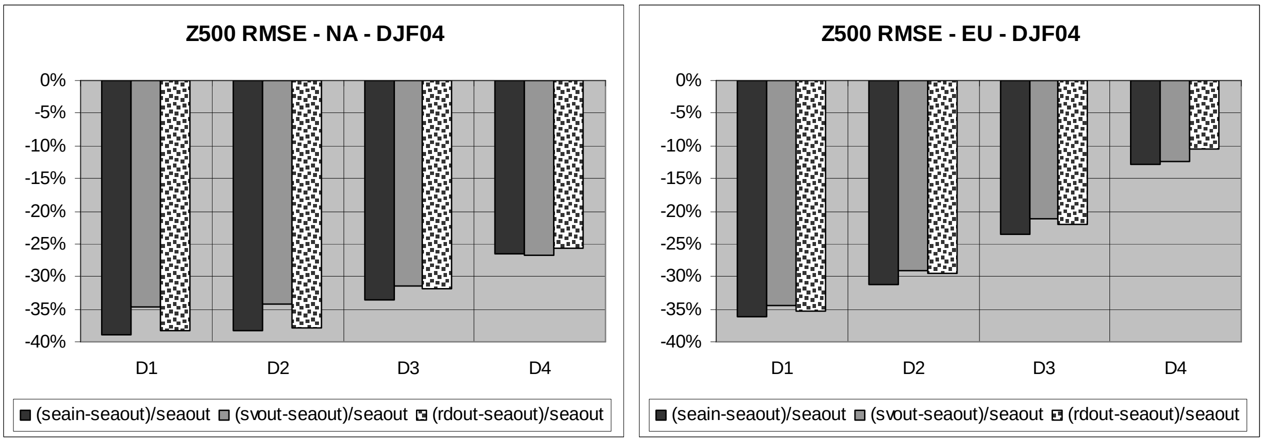

Figure 12. DJF04 data-denial experiments. Value of ocean observations (CON, black bars), and of observations removed in SV-target (SVOUT, grey bars) and random (RDOUT, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 for the 500 hPa geopotential height of the Pacific-North American (left panel) and the Atlantic-European (right panel) experiments.

The following conclusions can be drawn from Table 2, which lists a summary of the data-denial results considering the 500 hPa geopotential height day-2 forecasts:

- The comparison of experiments SVIN, RDIN and SEAOUT (section 5.1.1) indicates that *if no observations are taken in the Pacific ocean*, then observation taken in the SV-target areas can reduce the forecast error by 27.5%, while taking observations in random areas reduces it by 15.7%.

- The comparison of experiments SVOUT, RDOUT, SEAIN and SEAOUT indicates that *if observations are already taken in the whole Pacific ocean apart for the SV-target areas*, then observation taken in the SV-target areas can reduce the forecast error by 4%, while observations taken in the random areas can reduce the forecast error by 0.5%.

5.3.2 DJF04 Atlantic-Europe data-denial experiments

The bottom panel of Fig. 11 shows the average score of experiments SVOUT, RDOUT, to be compared with the two reference experiments SEAIN and SEAOUT, run for the Atlantic-European region between forecast day 1 and 4, and the bottom panel of Fig. 12 shows the 'value' of observations removed from the SV-target and the random areas measured considering the 500 hPa geopotential height fields. If we consider, for example, the day-2 forecasts, Fig. 12 shows that taking observations over the whole Atlantic Ocean reduces the forecast error by 31.2%, taking observation over the whole Atlantic Ocean apart for random areas reduces the forecast error by 29.5%, while taking observation over the whole Atlantic Ocean apart for SV-target areas reduces the forecast error by 29.2%. Thus, adding observations in SV-target areas (i.e.

comparing experiments SEAIN and SVOUT) has a slightly larger impact than adding observations in random areas, since the forecast error is further reduced by 2.0% (from 29.2% to 31.2%) compared to 1.7% (from 29.5% to 31.2%). The differences between SVOUT and SEAIN are statistically significant at the 5% level at forecast days 1, 2 and 3, and are not statistically significant at forecast day 4, but the differences between SVOUT and RDOUT are not statistically significant. Compared to the Pacific-North America region, these data-denial experiment show that observations taken in SV-target areas have only a slightly higher 'value' than observations taken in random areas, but the difference is not statistically significant.

The following conclusions can be drawn from Table 3, which lists a summary of the data-denial results considering the 500 hPa geopotential height day-2 forecasts:

- The comparison of experiments SVIN, RDIN and SEAOUT (section 5.1.2) have shown that if no observations are taken in the Atlantic ocean, then observation taken in the SV-target areas can reduce the forecast error by 19.1%, while taking observations in random areas reduces it by 14.9%.

- The comparison of experiments SVOUT, RDOUT and SEAOUT have shown that if observations are already taken in the whole Atlantic ocean apart for the SV-target areas, then observation taken in the SV-target areas can reduce the forecast error by 2.0%, while observations taken in the random areas can reduce the forecast error by 1.7%.

Value of targeted observations measured using day-2 Z500 fcs	**Verified over Europe**	
	Winter (DJF04)	**Summer (JJA04)**
SEAOUT	25.98 m	17.96 m
SEAIN	17.87 m	10.25 m
SVIN	21.02 m	12.82 m
RDIN	22.10 m	13.04 m
SVOUT	18.40 m	10.62 m
RDOUT	18.33 m	10.40 m
SVavIN	--	--
(SEAIN-SEAOUT) / SEAOUT	31.2% (15.0h)	42.9% (21.7h)
(SVIN-SEAOUT) / SEAOUT	19.1% (9.2h)	28.6% (14.5h)
(RDIN-SEAOUT) / SEAOUT	14.9% (7.2h)	27.3% (13.8h)
(SVavIN-SEAOUT) / SEAOUT	--	--
(SEAIN-SVOUT) / SEAOUT= [(SEAIN-SEAOUT) / SEAOUT]- [(SVOUT-SEAOUT) / SEAOUT]	2.0% (1.0h)	2.1% (1.0h)
(SEAIN-RDOUT) / SEAOUT= [(SEAIN-SEAOUT) / SEAOUT]- [(RDOUT-SEAOUT) / SEAOUT]	1.7% (0.8h)	0.8% (0.4h)
(SVOUT-SEAIN) / SEAIN	3.0%	3.6%
(RDOUT-SEAIN) / SEAIN	2.6%	1.6%

Table 3 Average value of observations taken in different target areas, measured considering the root-mean-square-errors (RMSE) of the day-2 500 hPa geopotential height forecasts over Europe. The first seven rows list the RMSE of the different forecasts (m). The other rows show the differences, in %, between normalized-RMSEs, with the normalization done using the RMSE of the SEAOUT experiment for all but the last two rows (with values in brackets are expressed in terms of hours of predictability gains)where the RMSE of SEAIN has been used.

6. Northern Hemisphere summer results (JJA04)

This section mirrors section 5, but discusses the results obtained for summer 2004.

6.1 JJA04 data-injection experiments: SV-target versus random areas

6.1.1 JJA04 Pacific-North America data-injection experiments

The top panel of Fig. 13 shows the average score of experiments SVIN, RDIN, SEAIN and SEAOUT run for the Pacific-North American region between forecast day 1 and 4. Overall, results confirm the conclusions of the winter experiments (Fig. 5, top panel), but quantitatively the differences between the four experiments are smaller for this period.

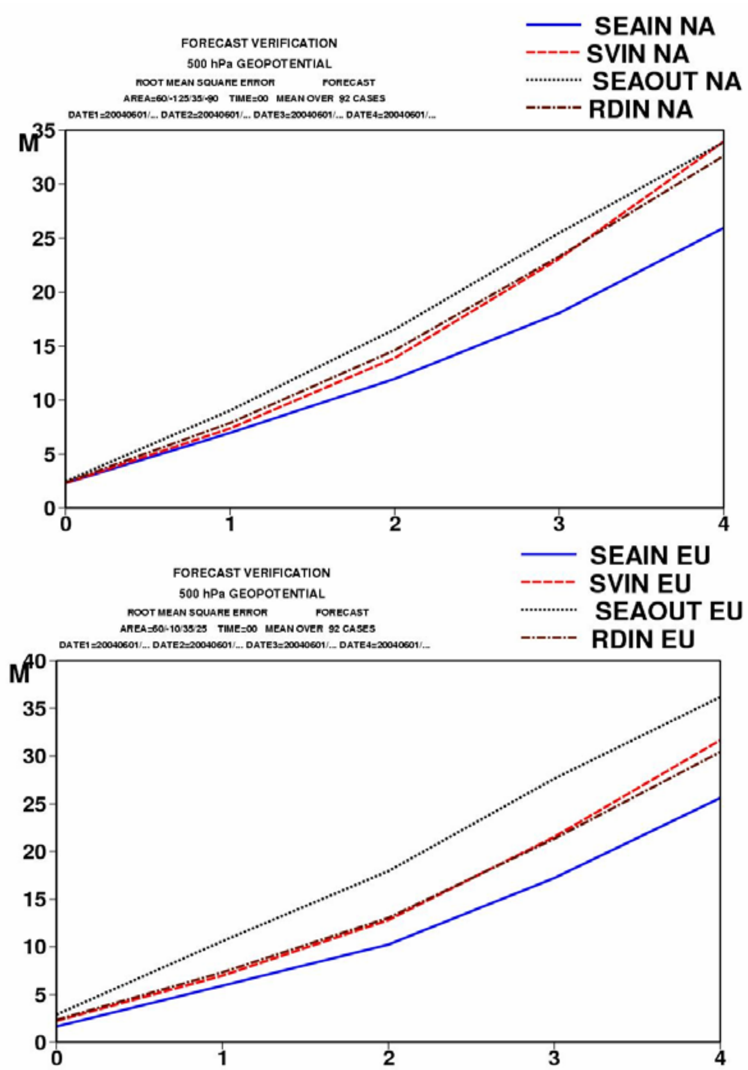

Figure 13. JJA04-average root-mean-square error of 500 hPa geopotential height forecasts of data-injection experiment SEAIN (control, blue line), SEAOUT (dotted green line), SVIN (red line) and RDIN (chain-dashed brown line) run for the Pacific-North American (top panel) and the Atlantic-European (bottom panel) regions.

The top panel of Fig. 14 shows the 'value' of observations defined considering the 500 hPa geopotential height fields. Considering the day-2 forecast, Fig. 14 shows that taking observations over the whole Pacific ocean reduces, on average, the forecast error by 27.8% (it was 38.3% in winter 2004, see Table 2), taking observations only in SV-target areas by 16.1 % (it was 27.5% in winter 2004), while taking observations in random areas by 11.8% (it was 15.7% in winter 2004). Thus, using observations taken only in SV-target areas is enough to recover 57.8% (=0.161/0.278) of the error reduction induced by all the Pacific data. This value is quantitatively smaller but not dramatically different to the error reduction obtained in winter 2004 (71.8%, see section 5.1.1), and is similar to the value obtained considering the Atlantic-European region in winter 2004 (61.1%, see section 5.1.2).

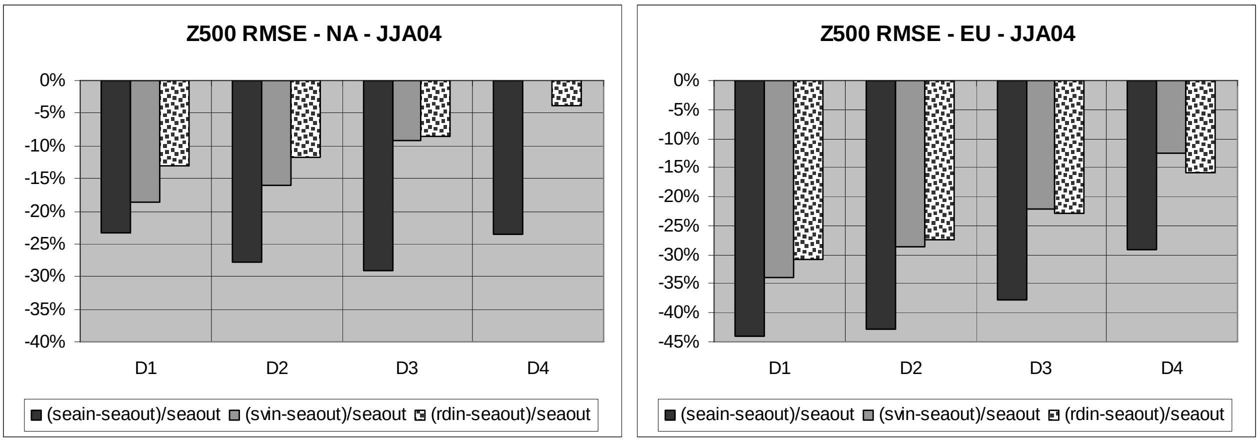

Figure 14. JJA04 data-injection experiments. Value of ocean observations (CON, black bars), and of targeted observations taken in SV-target (SVIN, grey bars) and random (RDIN, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 over Europe for the 500 hPa geopotential height for the Pacific-North American (left panel) and the Atlantic European (right panel) experiments.

Considering the SVIN and RDIN experiments, their difference is statistically significant at the 0.1% level at forecast days 1 and at the 5% level at forecast day 2, but it is not statistically significant afterwards. Thus, differences are smaller than in winter: considering the day-2 forecast for the 500 hPa geopotential height (see Table 2), the extra value of SV-targeted observation is 4.3% in summer (16.1% minus 11.8%), while it was 11.8% in winter (27.5% minus 15.7%).

The top panel of Fig. 15 shows the time-series of the 48-hour forecast error of SEAIN, SVIN and RDIN. Apart for four periods (19 June, 15 and 19 July, and 21 Aug), SVIN either outperforms or has a similar error to RDIN. This figure confirms that compared to winter (Fig. 8, top panel), during summer 2004 the impact of removing/re-injecting observations in the Pacific Ocean is smaller and can be detected in few cases (this can be seen, e.g., by comparing RDIN and SEAIN: RDIN is significantly worse in 3 periods).

6.1.2 JJA04 Atlantic-Europe data-injection experiments

The bottom panels of Figs. 13-14 show the average score of experiments SVIN, RDIN, SEAIN and SEAOUT for the Atlantic-European region between forecast day 1 and 4, and their value. Overall, results confirm the conclusions of the winter experiments (Fig. 5), but quantitatively the differences between the four experiments are smaller. Considering the day-2 forecast, Fig. 14 shows that taking observations over the whole Atlantic ocean reduces, on average, the forecast error by 42.9% (it was 31.2% in winter 2004, see Table 3), taking observations only in SV-target areas by 28.6 % (it was 19.1% in winter 2004), while taking observations in random areas by 27.3% (it was 14.9% in winter 2004). Thus, using observations taken only in SV-target areas is enough to recover 66.7% (=0.286/0.429) of the error reduction induced by all the

Atlantic data. This value is quantitatively smaller but not dramatically different to the error reduction obtained in winter 2004 (61.1%, see section 5.1.2).

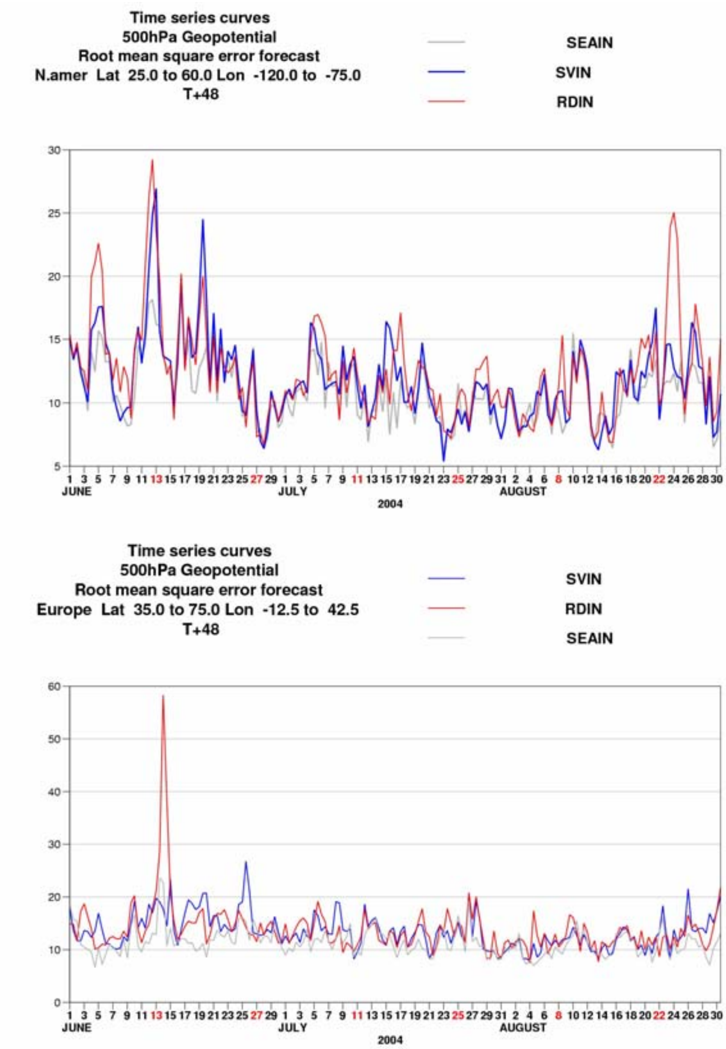

Figure 15. JJA04 time series of 48-hour forecast error of experiment SEAIN (grey line), and Pacific injection experiments SVIN (blue line) and RDIN (red line) verified over North America (top panel), and of experiment SEAIN, and Atlantic injection experiments SVIN and RDIN verified over Europe (bottom panel).

Considering the SVIN and RDIN experiments, their difference is statistically significant at the 10% level only at forecast day 2, and their differences are smaller than in winter and less statistically significant: considering the day-2 forecast for the 500 hPa geopotential height (see Table 2), the extra value of SV-targeted observation is 4.3% in summer (16.1% minus 11.8%), while it was 11.8% in winter (27.5% minus 15.7%).

The bottom panel of Fig. 15 shows the time-series of the 48-hour forecast error of SEAIN, SVIN and RDIN. Apart for three periods (17-19 and 25 June, and 8 July), SVIN either outperforms or has a similar error to RDIN. This figure confirms that during summer 2004 the impact of removing/re-injecting observations in the Atlantic Ocean is smaller and can be detected in few cases (this can be seen, e.g., by comparing RDIN and SEAIN: RDIN is significantly worse in only 1 period).

6.2 JJA04 data-denial experiments: SV-target versus random areas

This section is organized as section 5.3: section 6.2.1 summarizes the results of the JJA04 Pacific/North American data-denial experiments, and section 6.2.2 summarizes the results of the JJA04 Atlantic/European data-denial experiments.

6.2.1 JJA04 Pacific-North America data-denial experiments

The top panel of Fig. 16 shows the average score of experiments SVOUT, RDOUT, to be compared with the two reference experiments SEAIN and SEAOUT, run for the Pacific-North American region between forecast day 1 and 4. Results indicate that RDOUT, SVOUT and SEAIN have almost identical scores. The top panel of Fig. 17 shows the 'value' of observations removed from the SV-target and the random areas measured considering the 500 hPa geopotential height fields (results for the other three variables considered in the previous sections are very similar, and are not shown). Considering the day-2 forecasts, Fig. 17 shows that taking observations over the whole Pacific ocean reduces the forecast error by 27.8%, taking observation over the whole Pacific ocean apart for random areas reduces the forecast error by 28.7% (the difference is small and not statistically significant), while taking observation over the whole Pacific ocean apart for SV-target areas reduces the forecast error by 26.3%. Thus, for the Pacific Ocean also in summer 2004 adding observations in SV-target areas (i.e. comparing experiments SEAIN and SVOUT) has a small, but detectable value, since the forecast error is further reduced by 1.5%, from 26.3% to 27.8%. By contrast, adding data in the random areas leads to a small increase in forecast error of 2.4%, from 28.7% to 26.3%. The differences between SVOUT and RDOUT are statistically significant at the 10.0% level at forecast days 1 and at the 0.1% level at forecast day, but are not statistically significant afterwards.

As it was the case in winter, compared to the corresponding data-injection experiments (SVIN and RDIN), the SVOUT and RDOUT data-denial experiments confirm that observations taken in SV-target areas have a higher 'value' than observations taken in random areas, but indicate that the value of these observations is strongly dependent on the baseline observing system. In summary, considering for example the 500 hPa geopotential height (see Table 2):

- The comparison of experiments SVIN with SEAOUT has shown that if no observations are taken in the Pacific ocean, then observation taken in the SV-target areas can reduce the forecast error by 16.0% (it was 27.5% in winter 2004).
- The comparison of experiment SVOUT with SEAIN has shown that if observations are already taken in the whole Pacific ocean apart for the SV-target areas, then observation taken in the SV-target areas can reduce the forecast error by 1.3% (it was 4.0% in winter 2004).

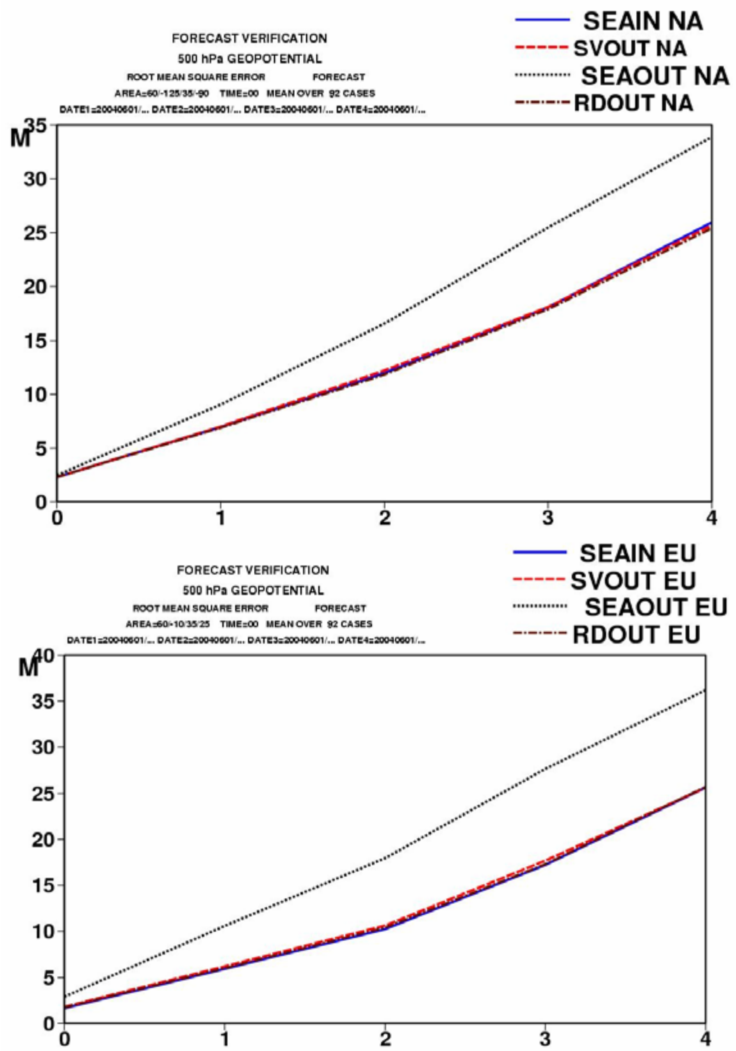

Figure 16. JJA04-average root-mean-square error of 500 hPa geopotential height forecasts of data-denial experiment SEAIN (control, blue line), SEAOUT (dotted green line), SVOUT (red line) and RDOUT (chain-dashed brown line) of the Pacific-North American (top panel) and the Atlantic European (bottom panel) experiments.

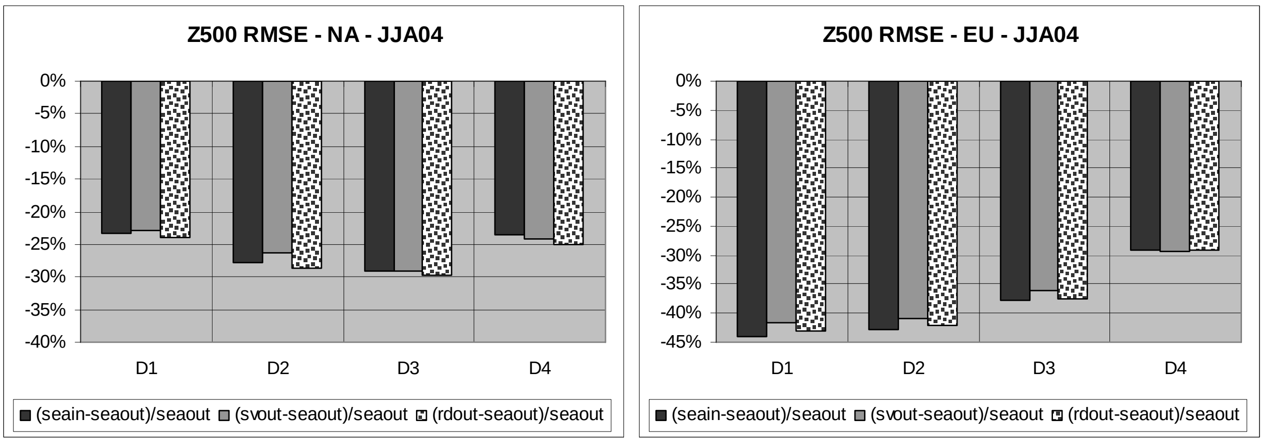

Figure 17. JJA04 data-denial experiments. Value of ocean observations (CON, black bars), and of observations removed in SV-target (SVOUT, grey bars) and random (RDOUT, dotted bars) areas, computed using SEAOUT as reference for different variables and forecast days 1 to 4 for the 500 hPa geopotential height of the Pacific-North American (left panel) and the Atlantic European (right panel) experiments.

6.2.2 JJA04 Atlantic-European data-denial experiments

The bottom panel of Fig. 16 shows the average score of experiments SVOUT, RDOUT, to be compared with the two reference experiments SEAIN and SEAOUT, run for the Atlantic-European region between forecast day 1 and 4. Results indicate that RDOUT and SEAIN have almost identical scores, while SVOUT has slightly higher RMS values. The bottom panel of Fig. 17 shows the 'value' of observations removed from the SV-target and the random areas measured considering the 500 hPa geopotential height fields. Considering the day-2 forecasts, Fig. 17 shows that taking observations over the whole Pacific ocean reduces the forecast error by 42.9%, taking observation over the whole Pacific ocean apart for random areas reduces the forecast error by 42.1% (the difference is small and not statistically significant), while taking observation over the whole Pacific ocean apart for SV-target areas reduces the forecast error by 40.8%. Thus, for the Atlantic Ocean, also in summer 2004 adding observations in SV-target areas (i.e. comparing experiments SEAIN and SVOUT) has a small, but detectable value, since the forecast error is further reduced by 2.1%, from 40.8% to 42.9%. By contrast, adding data in the random areas leads to a small increase in forecast error of 0.8%, from 42.1% to 42.9%. The differences between SVOUT and RDOUT are statistically significant at the 10.0% level at forecast days 1 and at the 5% level at forecast day 3, but are not statistically significant at forecast days 2 and 4.

Compared to the corresponding data-injection experiments (SVIN and RDIN), the SVOUT and RDOUT data-denial experiments confirm that observations taken in SV-target areas have a higher 'value' than observations taken in random areas, but indicate that the value of these observations is strongly dependent on the baseline observing system. In summary, considering for example the 500 hPa geopotential height (see Table 3):

- The comparison of experiments SVIN with SEAOUT has shown that if no observations are taken in the Pacific ocean, then observation taken in the SV-target areas can reduce the forecast error by 28.6% (it was 19.1% in winter 2004).

- The comparison of experiment SVOUT with SEAIN has shown that if observations are already taken in the whole Pacific ocean apart for the SV-target areas, then observation taken in the SV-target areas can reduce the forecast error by 2.1% (it was 2.0% in winter 2004).

7. Conclusions

This paper is the second of three companion papers that summarize the results of a large number of data-assimilation experiments performed to address some fundamental questions related to the value of targeted adaptive observations:

- What is the 'value' of observations taken in target regions identified using SVs compared to the value of observations taken in randomly chosen regions? Is it important that SV-target regions are identified using the most recent analysis and forecast, or can an SV-average-target regions be used?

- Does the 'value' of observations depend on the region? What is the value of observations taken in SV-target regions of the Pacific Ocean for 2-day forecasts verified over North-America? How does it compare to the value of observations taken in SV-target regions of the Atlantic Ocean for 2-day forecasts verified over Europe?

These questions have been addressed considering 4-dimensional variational data-assimilation experiments performed at T_L511L60 for two seasons, winter 2003/2004 (DJF04, 91 cases) and summer 2004 (JJA04, 92 cases). Experiments were designed to overcome three key weaknesses that, in our view, affected previous similar studies: (i) experiments were designed to guarantee a complete matching between the target area identified by the method under investigation and the area actually sampled by the targeted observations, (ii) target areas were forced to have the same size, and (iii) a large number of case studies has been considered, to increase the statistical significance of the conclusions.

The following three general conclusions can be drawn from all these experiments:

1) Observations taken in SV-target areas are more valuable than observations taken in random areas, but the difference depend on the region, on the season and on the baseline observing system used as a reference

2) It is important that the daily set of singular vectors are used to compute the target areas: experiments run for the Pacific/North American region for winter 2004 indicated that observations taken in a fixed target area identified considering the average of all the daily SV-target areas would have a smaller value than observations taken in the daily SV-target areas (while SVIN experiments have an average a 27.5% smaller error than SEAOUT, SVavIN have a 20.9% smaller error, see Table 2). However, this has to be mitigated against the cost of deploying a daily targeting strategy.

3) The value of observations taken in SV-target areas defined using SVs optimized for a 2-day time period starts decreasing after forecast day 3, due to the fact that after forecast day 3 the impact of the targeted observations moves outside the verification region

More extensively, the following considerations can be drawn from this study.

Considering the sensitivity of the value of observations to the region, results based on winter 2003-2004 have indicated that the value of observations taken in Pacific SV-target areas for day-2 forecasts verified over North America is higher than the value of observations taken in Atlantic SV-target areas for day-2 forecasts verified over Europe, but the reverse is true for summer 2004. During winter, the difference could be linked to the fact that due to the stronger (compared to summer) jet stream, the positive effect of North American observations propagate further during winter than during summer. The fact that the reverse is true during summer may be linked the fact that during summer 2004 extra-tropical transitions affected the Atlantic Ocean, causing an increased sensitivity to observations. This aspect will be further discussed in CBKST07, which focuses on the impact of targeted observations taken in the Atlantic Ocean during cases of extra-tropical transitions. It is interesting to point out that for both seasons the difference between injecting observations in SV-target or random areas is larger for the Pacific/North American region. A possible explanation of these results is the fact that the Atlantic Ocean is 20° degree smaller in longitude smaller than the Pacific Ocean (see section 2.3). By consequence, the chance that the SV-target and the random areas overlap is higher over the Atlantic Ocean.

Considering the sensitivity of the value of observations to the season and to the baseline observing system, experiments have indicated that:

- *If the baseline observing system is data void (i.e. no observations) over the ocean*, then the average value of observations taken in SV-target areas is very high. Considering the value measured using the root-mean-square-error of 500 hPa geopotential height forecasts (Tables 2 and 3), results indicate that SV-targeted observations are capable to reduce the day-2 average forecast error in the verification region by:
    - 27.5% in DJF04 for SV-targeted Pacific obs and fcs verified over N. America, which corresponds to 13.5 hours of forecast gains;
    - 19.1% in DJF04 for SV-targeted Atlantic obs and fcs verified over Europe, which corresponds to 9.2 hours of forecast gains;
    - 16.0% in JJA04 for SV-targeted Pacific obs and fcs verified over N. America, which corresponds to 7.8 hours of forecast gains;
    - 28.6% in JJA04 for SV-targeted Atlantic obs and fcs verified over Europe, which corresponds to 9.2 hours of forecast gains.
- If the baseline observing system is data rich (i.e. with all observations) over the ocean, then the average value of observations taken in SV-target areas is very small. Considering the value measured using the root-mean-square-error of 500 hPa geopotential height forecasts, results indicate that removing SV-targeted observations increases the day-2 forecasts error in the verification region by:
    - 4.0% in DJF04 for SV-targeted Pacific obs and fcs verified over N. America, which corresponds to 2 hours of forecast gains;
    - 2.0% in DJF04 for SV-targeted Atlantic obs and fcs verified over Europe, which corresponds to 1 hour of forecast gains;
    - 1.3% in JJA04 for SV-targeted Pacific obs and fcs verified over N. America, which corresponds to 0.7 hours of forecast gains;
    - 2.0% in JJA04 for SV-targeted Atlantic obs and fcs verified over Europe, which corresponds to 1.0 hour of forecast gains.

The data-denial experiments do not replicate precisely the impact that adding extra observations taken in targeted regions may have on forecast accuracy, but in our view they can be used to estimate the potential average impact that they may have. In the hypothesis that the characteristics (type, quality, content of information) of future extra observations is similar to the characteristics of the observations removed, the data-denial experiments provide an upper bound of the expected average impact that extra observations may have. These data-denial experiments indicate that, on average, if extra observations are taken in SV-target regions over the ocean with a size of 20-degree-squared and assimilated using the ECMWF 4-dimensional variational data-assimilation systems with a T_L511L60 resolution, one should only expect a small impact on the forecast accuracy measured in a downstream area, for both the Pacific-North American and the Atlantic-European regions.

The fact that in summer the value of observations taken in SV-target is smaller, and the fact that the difference between the value of observations taken in the SV-target or random areas is also smaller, is possibly due to the SVs' characteristics (Buizza and Palmer 1995): in winter, the SVs' amplification rate spectrum is steeper, which makes it easier to separate between the leading 10 from the others, in particular from the directions spanned by the random area. Furthermore, in winter the SVs are more localized in the storm track region, again making their location more 'different' from the location identified by the random areas. Finally, it is worth reminding that the SVs have been computed with a simplified, dry tangent forward and adjoint physics, which may make their computation less accurate in summer, a period during which moist processes play a bigger role than in winter.

These values could be compared to the reduction of the t+48h forecast error of the ECMWF high-resolution forecast between 1995 and 2005: over North America, the root-mean-square was reduced from about 25 to 16 meters, i.e. by ~36%, while over Europe the root-mean-square error was reduced from about 24 to 15 meters, i.e. by ~37% (Adrian Simmons, personal communication). In other words, developments of the observation network, and of the ECMWF data assimilation and forecasting system lead to a forecast error reduction of about 3.6-3.7% per annum. Thus, the average impact of SV-targeted observations in the case of a data rich baseline observing system over the ocean is comparable to the annual forecast error reduction of the ECMWF high-resolution forecast.

These considerations suggest three possible ways to continue this type of studies:

- The first possible extension would be to investigate whether using moist SVs (Ehrendorfer et al 1999, Coutinho et al 2004) would increase the value of observations taken in SV-target areas, and increase the difference between the value of observations taken in SV-target and random areas in the summer. Such a study will have to face the fact that moist processes make the singular vectors grow faster, and this may lead to an earlier break-down of the validity of the linear approximation (Gilmour et al 2001), thus making it impossible to work with a 48-hour optimisation time interval.

- The second possible extension would be to investigate the sensitivity of the value of SV-targeted observations to the size of the area covered by the observations. The results discussed in this work are based on areas covering an area of 20°-degree-squared, which are rather large areas, possibly too large to be targeted during field campaigns. During this work, only few experiments have been run using smaller target areas: results from these experiments (not shown) indicate that the impact is very small if the area is halved. It could be interesting to assess whether this is true for a larger number of cases.

- The third possible extension would be to investigate the sensitivity of the value of targeted observations taken over land from higher quality observation platforms, capable to provide more accurate data both in cloud-free and cloud-covered areas. High quality data in the right place could have a non-negligible impact; however, such experiments, if they were to show too little impact, could be criticized as being unrepresentative on the grounds that observation removal is taking place over regions where error growth characteristics are rather different than those over the oceanic storm track (A Simmons, personal communication).

This work proposes a framework that could be applied to study the value of other objective targeting methodologies, and to investigate their sensitivity to the data-assimilation system used to assimilate the extra observations. Thus, two other possible extensions of this work would be to apply this framework to investigate whether other objective methodologies could outperform the singular vector one used in this work, and to apply it to investigate the value of different types of observations (e.g. of data taken using different observing platforms, or to study the impact of using more observations in target areas and less observations elsewhere).

Acknowledgements

The authors would like to thank Erik Anderson, Philippe Bougeault and Adrian Simmons for their very useful comments to an earlier version of this paper.

Appendix A: Singular vectors definition

Lorenz (1965) was the first author to indicate that perturbation growth in realistic models is related to the eigenvalues and eigenvectors of the operator product of the tangent forward linear and adjoint operators, i.e. to the singular values and singular vectors of the tangent forward operator. Subsequently, Farrell (1982), studying the growth of perturbations in baroclinic flows, showed that, although the long time asymptotic behavior is dominated by discrete exponentially growing normal modes when they exist, physically realistic perturbations could present for finite time intervals amplification rates greater than the most unstable normal mode amplification rate. Since 1992, singular vectors (Buizza and Palmer 1995) have been used in the ECMWF Ensemble Prediction System (EPS) to simulate the effect of initial uncertainties (Molteni et al 1996). The SVs used in this work has been computed in a similar way as it is done in the ECMWF EPS.

Let χ be the state vector that describes the state of the atmosphere, of which the evolution equations can be formally written as

(A.1) $$\frac{\partial \chi}{\partial \mathrm{t}} = A(\chi)$$

Denote by $\chi(t)$ an integration of these equations from t_0 to t. The time evolution of a small perturbation x around the time evolving trajectory $\chi(t)$ can be described, in a first approximation, by the linearised model equations

(A.2) $$\frac{\partial \mathrm{x}}{\partial \mathrm{t}} = A_l x,$$

where $A_l = \left.\frac{\partial A(x)}{\partial x}\right|_{\chi(t)}$ is the tangent operator computed at the trajectory point $\chi(t)$.

Denote by $L(t,t_0)$ be the integral forward propagator of the dynamical equations linearized about a non-linear trajectory $\chi(t)$

(A.3) $$\mathrm{x(t)} = \mathrm{L(t,t_0)}x(t_0),$$

that maps a perturbation x at initial time t_0 to the optimization time t. Denote by $(..;..)_E$ the inner product between two vectors x and y

(A.4) $$(x;y)_E = \langle x; Ey \rangle,$$

defined by the matrix E. Denote by L^{*E} the adjoint of L with respect to the inner product $(..;..)_E$,

(A.5) $$(L^{*E}x;y)_E = (x;Ly)_E$$

Note that the adjoint of L with respect to the inner product defined by E can be written in terms of the adjoint L^* defined with respect to the canonical Euclidean scalar product,

(A.6) $$\mathrm{L}^{*\mathrm{E}} = E^{-1}L^{*}E.$$

The squared norm of a perturbation x at time t can be computed as

(A.7) $\|x(t)\|_E^2 = (x(t_0); L^{*E} Lx(t_0))_E$

From these equations, it follows that the problem of finding the phase space directions x for which $\|x(t)\|_E^2 / \|x(t_0)\|_E^2$ is maximum can be reduced to the computation of the eigenvectors $v_I(t_0)$ with the largest eigenvalues σ_i^2, i.e. to the solution of the eigenvalue problem

(A.8) $$\begin{aligned} K \cdot v_i(t_0) &= \sigma_i^2 v_i(t_0) \\ K &= E^{-1/2} \cdot L^* \cdot E \cdot L \cdot E^{-1/2} \end{aligned}$$

The square roots of the eigenvalues, σ_i, are called the singular values and the eigenvectors $v_i(t_0)$ the (right) singular vectors of L with respect to the inner product E (see, e.g., Noble and Daniel 1977). The singular vectors with largest singular values identify the directions characterized by maximum growth. The time interval t-t_0 is called optimization time interval.

To compute SVs with final-time energy maximized inside a specific region, a projection operator is required (Buizza 1994a). Denote by x_g the grid point representation of the state vector x, by S the spectral-to-grid point transformation operator, $x_g=Sx$, and by Gx_g the multiplication of the vector x_g, defined in grid point space, by the function $h(s)$:

(A.9) $$\begin{aligned} h(s) &= 1 \forall s \in \Sigma \\ h(s) &= 0 \forall s \notin \Sigma \end{aligned},$$

where s defines the coordinate of a grid point, and Σ is a geographical region. The application of the local projection operator T defined as

(A.10) $T = S^{-1}GS$,

to any vector x sets the vector x to zero for all grid points outside the geographical region Σ.

The SVs with maximum final time norm inside a specific region Σ can be computed by solving the following eigenvalue problem

(A.11) $K = E^{-1/2} \cdot T \cdot L^* \cdot E \cdot L \cdot E^{-1/2}$

which is defined by the metric, the tangent forward and adjoint operators, the optimization time interval and the projection operator.

References

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the Ensemble Transformed Kalman Filter. Part I: Theoretical aspects. *Mon. Wea. Rev*., **129**, 430-436.

Buizza, R., 1994a: Localization of optimal perturbations using a projection operator. *Q. J. R.. Meteorol. Soc.*, **120**, 1647-1682.

Buizza, R., 1994b: Sensitivity of Optimal Unstable Structures. *Q. J. R. Meteorol. Soc*., **120**, 429-451.

Buizza, R., and T.N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. J. Atmos. Sci., **52**, 1434-1456.

Buizza, R., and A. Montani, 1999: Targeting observations using singular vectors. *J. Atmos. Sci.*, **56**, 2965-2985.

Cardinali, C., and R. Buizza, 2003: Forecast skill of targeted observations: a singular-vector-based diagnostic. *J. Atmos. Sci*., **60**, 1927-1940.

Cardinali, C., R. Buizza, G. Kelly, M. Shapiro, and J.-N. Thépaut, 2007: The value of targeted observations Part III: weather regimes influence. *Q. J. R. Meteorol. Soc*., submitted. Also published as ECMWF Technical Memorandum number 513.

Coutinho, M. M., B. J. Hoskins, and R. Buizza, 2004: The influence of physical processes on extratropical singular vectors. *J. Atmos. Sci*., **61**, 195-209.

Ehrendorfer, M., R. M. Errico, and K,. D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci*., **56**, 1627-1648.

Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663-1686.

Gilmour, I., L. A. Smith, and R. Buizza, 2001: On the duration of the linear regime: is 24 hours a long time in weather forecasting?. *J. Atmos. Sci*., **58**, 3525-3539.

Joly, A. K. A. Browning, P. Bessemoulin, J.-P. Cammas, G. Caniaux, J.-P. Chalon, S. A. Clough, R. Dirks, K. A. Emanuel, L. Eymard, F. Lalaurette, R. Gall, T. D. Hewson, P. H. Hildebrand, D. Jorgensen, R. H. Langland, Y. Lemaitre, P. Marcard, J. A. Moore, P. O. G. Persson, F. Roux, M. A. Shapiro, C. Snyder, Z. Toth, R. M. Wakimoto, 1999: Overview of the field phase of the Fronts and Atlantic Storm-Track Experiment (FASTEX) project. *Q. J. R. Meteorol. Soc*., **125**, 3131-3163.

Kelly, G, J.-N. Thépaut, R. Buizza, and C. Cardinali, 2007: The value of targeted observations Part I: data denial experiments for the Atlantic and the Pacific. *Q. J. R. Meteorol. Soc*., submitted. Also available as ECMWF RD Technical Memorandum number 511.

Langland, R. H., 2005: Issues in targeted observing. *Q. J. R. Meteorol. Soc*., **613**, 3409-3425.

Lorenz, C. E., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321-333.

Mahfouf, J. F. and F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: experimental results with improved physics. *Q. J. R. Meteorol. Soc.*, **126**, 1171-1190.

Majumdar, S., Bishop, C., Buizza, R., and Gelaro, R., 2002: A comparison of PSU-NCEP Ensemble Transform Kalman Filter targeting guidance with ECMWF and NRL Singular Vector guidance. *Q. J. R. Meteorol. Soc.*, **128**, 2527-2549.

Majumdar, S J, Aberson, S D, Bishop, C H, Buizza, R, Peng, M, and Reynolds, C, 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.*, in press.

Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T., 1996: The new ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.*, **122**, 73-119.

Montani, A., Thorpe, A. J., Buizza, R., and Unden, P., 1999: Forecast skill of the ECMWF model using targeted observations during FASTEX. *Q. J. R. Meteorol. Soc.*, **125**, 3219-3240.

Noble, B., and Daniel, J. W., 1977: *Applied linear algebra*, Prenctice-Hall, Inc., pp. 477.

Palmer, T. N., Gelaro, R., Barkmeijer, J., and Buizza, R., 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633-653.

Petersen, G. N., Dumelow, R., and Thorpe, A. J., 2006: Impact of NATReC observations on Global model forecasts. *Geoph. Res. Abs.*, **8**, 08999.

Ralph, F. M., O. Persson, D. Reynolds, P. Neiman, W. Nuss, J. Schmidt, D. Jorgensen, C. King, A. White, J. Bao, W. Neff, D. Kinsmill, D. Miller, Z. Toth, and J. WIlczak, 1998: The use of tropospheric profiling in CALJET. *Proceedings of the 4th Symposium Tropospheric Profiling: Needs and Technology*, 20-25 September, Snowmass, CO, p. 258-260.

Snyder, C., 1996: Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Amer. Meteorol. Soc.*, **77**, 953-961.

Szunyogh, I., Z. Toth, S. Majumdar, R. Morss, B. Etherton, and C. Bishop, 2000: The effect of targeted observations during the 1999 Winter Storm Reconnaissance program. *Mon. Wea. Rev.*, **128**, 3520-3537.

Szunyogh, I., Z. Toth, A. Zimin, S. Majumdar, and A. Persson, 2002: On the propagation of the effect of targeted observations: the 2000 Winter Storm Reconnaissance program. *Mon. Wea. Rev.*, **130**, 1144-1165.

Toth, Z., I. Szunyogh, C. Bishop, S. Majumdar, R. Morss, J. Moskaitis, D. Reynolds, D. Weinbrenner, D. Michaud, N. Surgi, M. Ralph, J. Parrish, J. Talbot, J. Pavone, and S. Lord, 2002: Adaptive observations at NCEP: past, present and future. Proceedings of the *Symposium on Observations, Data Assimilation, and Probabilistic Prediction*, 13-17 January 2002, Orlando, FL.