TECHNICAL MEMORANDUM

511

The value of targeted observations Part I: Data denial experiments for the Atlantic and the Pacific

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

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

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

In the past, there has been a number of targeting experiments where extra observations were added in a small region to an existing operational system. Generally speaking, these experiments have shown that the impact of these 'targeted observations' is small, adding on average only a few hours of forecast skill. One of the weaknesses of these studies lies in that they did not provide any (reference) value of what the impact of removing all data over a larger region is, e.g. the ocean, and thus they were not able to interpret the impact of the targeted observations using these reference results. This work is the first of three companion papers that discuss the value of targeted observations: its focus is to examine the effect of removing all types of observations from most of the Pacific and Atlantic within the ECMWF operational system as of June 2005. In other words, this work discusses results from reference experiments necessary to properly interpret the impact of targeted observations.

This study not only looks at the mean impact but examines the skill of each individual forecasts in terms of root-mean-square-errors (RMSE) and an objective synoptic classification. Generally these two scores show similar characteristics.

In this work, experiments using both 4D- and 3D-Var data assimilation systems designed to provide a set of baseline experiments necessary to interpret the value of targeting experiments within these oceanic basins are discussed. These experiments indicate that if a 4D-VAR data assimilation system is used, the impact of removing observations remains confined to the denial region and does not progress very far down stream as the forecast period increases. For example, the Pacific denial experiment (covering a total period of six months) shows a degradation of the 48 hour forecasts over North America, while the effect in the medium range over Europe remains very small. The Atlantic denial experiment shows a downstream short range forecast degradation over Europe, although the magnitude of the short range impact is clearly smaller than when denying observations over the Pacific Ocean.

One of the key aspects discussed in this work is the sensitivity of the value of observations to using either a 4D- or a 3D-Var version of the ECMWF data-assimilation system. Results based on removing Pacific oceanic data for a winter season (three months in total) indicate that the impact of removing Pacific oceanic data is larger if a 3D-Var system is used. In particular, and contrary to the 4D-Var data denial experiment, a large number of 48-hour forecasts over North America are badly affected when Pacific data are denied in a 3D-Var data assimilation system. This indicates a much better ability of the ECMWF 4D-Var assimilation to cope in the presence of incomplete or deficient data coverage, as compared to its 3D-Var predecessor (operational up to November 1997, Courtier et al. 1994 and 1997).

1. Introduction

The optimization of the Global Observing System (GOS) is an important goal for the World Meteorological Organization (WMO), satellite and in situ data producers and National Weather Services in general. Since the quality of numerical forecasts heavily depend on initial conditions and therefore the full composite of the GOS, it is vital that all these components are assessed in the context of NWP. In particular, as targeting strategies are potentially a cost-effective way of designing the GOS, their value deserves full assessment too. Observing System Experiments (OSEs) are natural tools to perform this assessment.

This work follows previous OSEs' studies carried out at ECMWF. In 1997, a first serious attempt was made to conduct a comprehensive set of OSE's, denying the major components of the GOS from the then new three-dimensional variational assimilation system (Kelly, 1997). Towards the end of 1999 a more advanced version of the variational analysis (4D-VAR) had been developed and in addition there also had been significant changes to the GOS, following the launch of the first ATOVS instrument onboard NOAA satellites. A new set of OSE's was then performed (Bouttier and Kelly 2002, and Kelly et al 2004). The main conclusions from these studies were that satellite data played a much larger role in the GOS than before. For the first time, satellite data had more impact than radiosondes in the Northern Hemisphere, and as a consequence, in presence of satellite observations, the removal of large components of the conventional observing network had less impact on the accuracy of the forecast, particularly in summer.

More recently, an interesting impact study was carried out using vertical wind profile observations from an airborne Doppler lidar flying during the A-TREC experiment, showing that this limited dataset could still bring a small improvement skill of about 3% in the day 2-4 forecast over Europe(Weissmann and Cardinali 2006), although the number of cases was very limited.

This work aims at contributing to the research work that the international community has started about a decade ago to assess the value of targeted observations. In the past decade, there have been a number of field experiments to try to assess the impact of extra observations taken in some specific, case-dependent targeted areas which were identified using objective and subjective methods. These campaigns have been recently summarised in a review article, (Langland, 2005). The general conclusions of these campaigns were that in most cases the targeted observations tend to increase the forecast skill, but that on average the impact is quite small.

This work and its two companion papers, Buizza et al 2006 (hereafter BCKT06) and Cardinali et al 2006 (hereafter CBKT06), aim to address four major weaknesses that affected these earlier studies (see BCKT06 for a more detailed discussion):

- Poor matching between the target area identified by the method under investigation and the area actually sampled by the extra observations.
- Lack of a clean comparison between the impact of observations taken in objectively-defined and randomly selected areas
- Large variation of target areas and number of extra observations taken in the target area
- Low statistical significance of the results due to the limited number of cases and biased case selection.

These weaknesses are addressed by examining both 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. To be able to draw statistically significant conclusions, a very large number of cases (315 cases, selected to cover both summer and winter) have been considered. The data-assimilation and forecast experiments have been performed at the European Centre for Medium-Range Weather Forecasts (ECMWF) using the same resolution as the operational data assimilation system until the 1st of February 2006 (Mahfouf and Rabier 2000), that is:

- 4-dimensional variational assimilation,, 12-hour cycle window
- Analysis resolution: outer loop TL511L60, inner loops TL159L60
- Forecast resolution: TL511L60
- Model cycle 29r1 (operational from April 5 to June 5 2005)

This paper, first part of the three companion papers, focuses on 183 cases (winter 2003-04 and summer 2004) and discusses the value of observations taken over the Pacific Ocean to the quality of 2-day forecasts verified over North-America, and of observations taken over the Atlantic Ocean to the quality of 2-day forecasts verified over Europe. We also investigate the impact of Pacific observations on the medium-range forecasts over Europe, and discuss the sensitivity of the impact of observations to the data assimilation system used. The two key questions that are addressed in this paper are the following:

(i) what is the importance of data measured over the ocean, in a mean statistical sense?

(ii) What is the importance of data measured over the ocean, on specific meteorological cases?

The second companion paper (BCKT06) focuses on the same 183 cases, and discusses the value of observations taken in target areas, defined either objectively using singular vectors (SVs) or randomly, over the ocean for 2-day forecasts verified over downstream regions over land. The third companion paper (CBKT06) considers the whole dataset of cases (315 instead of the 183 cases discussed in this work), and investigates in details whether the impact of targeted observation taken on the Atlantic Ocean is larger in some specific situations, more particularly in situations of extra-tropical cyclone transitions.

This paper is organized as follows. Section 2 presents the methodology used to define the data-denial experiments, section 3 and 4 show the impact of the data denial on the 4D-Var operational system. Section 5 develops the relationship between objective and subjective assessment of individual forecasts. Section 6 presents the impact of the Pacific data denial in the less accurate 3D-Var ECMWF assimilation of finally the conclusions of this work are discussed in section 7.

2. General Methodology

Fig. 1 illustrates the experimental design used in the oceanic data denial and data injection experiments discussed in this paper and its two companions. In Fig. 1, T is the target area where observations should be taken to improve the d-day forecast inside the verification area Σ. In the first set of experiments presented in this paper, T is the ocean North of 30N (either Pacific or Atlantic). These experiments have been run following two set-ups: once to study the value of observations located in the Pacific Ocean for forecasts verified over North-America, and a second time to study the value 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)

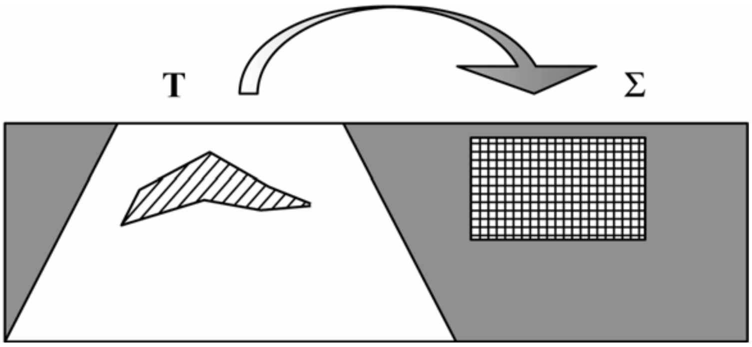


Figure 1. Schematic illustration of the concept of the 'value of targeted observations': the grey areas identify land, while the white region identify the ocean. T is the target area (e.g. identified using singular vectors) where observations should be taken to improve the d-day forecast inside the verification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

2.1 Experiment list

In this first part of the three companion papers, two types of data assimilation experiments that have been performed for both the Pacific-North American and the Atlantic-European regions (see BCKT06 for the other experiments) are discussed:

- SEAIN: all observations available over the ocean have been used
- SEAOUT: none of the observations available over the ocean have been used (due to the definition of the geographical definition of the ocean areas, some coastal observations have been excluded)

2.2 List of case studies

The results discussed in the first two parts of this work are based on experiments performed for summer 2004 and winter 2003-2004 More specifically, the two periods discussed in the first two parts of this work are:

- DJF04: from 1 December 2003 to 29 February 2004 (91 cases)
- JJA04: from 1 June to 31 August 2004 (92 cases)
- In the third part of this work (CBKT06), results from four types of data assimilation experiments (SEAIN, SEAOUT, SVOUT and RDOUT, see BCKT06 for a description of these experiments) run for the Atlantic-European region from winter 2003-04 and summer 2004 plus are accumulated to results obtained for four extra periods, characterised by particularly interesting atmospheric situations.

2.3 Number of observations

Table 1 indicates the number of observations available in the two ocean basins for a typical 12 hour window, and for five types of observations: Surface buoys (dribus), radiosondes temperatures at (500 hPa), 250 hPa temperature observations from aircraft, atmospheric motion winds (AMVs) and satellite radiances from AMSU-A channel 6. These data (which are denied in the experiments described below) are compared in percentage to the number of observations used in the full system for a longitude band (20-90 deg N).

Observation type	Total # observations in NH band (20°N to 90°N)	# observations removed from Atlantic ocean	% removed from Atlantic ocean	# observations removed from Pacific ocean	% removed from Pacific ocean
dribu	2280	897	39%	504	22%
temp(500 hPa)	3267	264	8%	82	3%
airep(250 hPa)	17886	1849	10%	1334	7%
AMVs (250 hPa)	3194	961	30%	547	17%
Amsua (aqua_channel 6)	10742	1516	14%	1829	17%

Table 1 Number of observations used in the analysis in a twelve hour period for some observation types (not all observation types used are shown). For reference column 2 show a selection of observations in the zonal latitude band (20°N to 90°N). Columns 3 to 6 show the number and percentages of these observation types removed in the Pacific or the Atlantic denial experiment.

Clearly, the relative number of satellite data (e.g. AMSUA) over the Pacific Ocean is larger than over the Atlantic, while it is the opposite for most conventional observations. In any case, the volume of data that is under scrutiny in this study (i.e. removed every 12 hours in the assimilation experiments) far exceeds the number of extra observations taken in past field campaigns used for targeting studies.

3. Winter (DJF 2004) results

The first part of this section discusses the impact of removing all oceanic data over the Pacific Ocean during winter 2003-2004 (December 2003 and January-February 2004, DJF04) on forecasts verified over 4 regions covering the Pacific Ocean, North-America, the Atlantic ocean and Europe. The second part discusses the impact of removing all Atlantic oceanic data during winter 2003-2004 over the same four regions. Summer results are discussed in Section 5.

3.1 DJF04 Pacific-North America data denial experiments

Fig. 2 shows the average 500 hPa geopotential root-mean-square forecast errors (RMSE) (as verified against the operational ECMWF analyses at a time) for experiments SEAIN (the reference experiment run using all oceanic data) and SEAOUT(Pacific), when all Pacific data are denied, verified over four regions: North Pacific, North America, North Atlantic and Europe. Results indicate that if no observations are taken over the Pacific Ocean, the forecast quality is reduced substantially over North Pacific and North America up to forecast day 8. Note that the signal (i.e. the negative impact of removing all oceanic data) reaches North America at forecast day 1, peaks at around forecast day 4-5, and then decays at about forecast day 8. Similarly, Fig. 2 shows that the signal reaches the Atlantic Ocean at about forecast day 3, and impacts the forecast skill over this region up to forecast day 10. By contrast, there is no clear indication of any major impact over Europe.

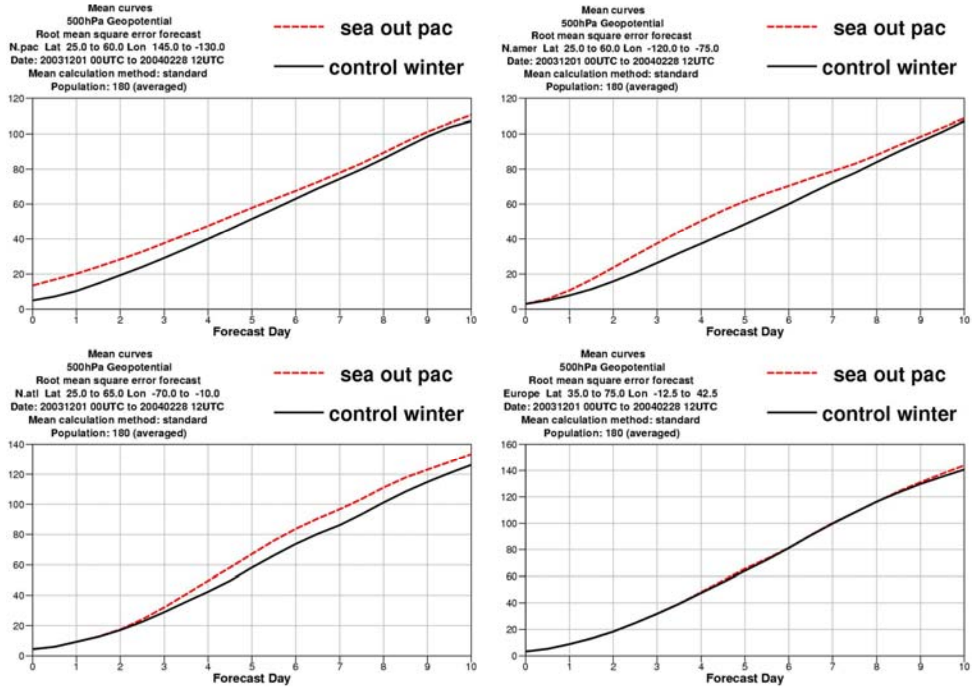


Figure 2. Winter Pacific forecasts: Verification of mean 500hPa Geopotential RMSE for up to day ten for SEAOUT in red and SEAIN in black: Both experiments are verified using ECMWF Operational analysis. Verification regions: top left North Pacific, top right North America, bottom left North Atlantic and Bottom right the European region.

Fig. 3 shows the time series of the 500 hPa geopotential RMSE for the SEAOUT(Pacific) and SEAIN forecasts over North America (at day-2) and Europe (at day-6). Almost all 2-day forecasts verified over North America (Fig. 3, upper panel) are affected by the Pacific data denial, while there are only very few cases for which the forecasts verified over Europe differ (Fig. 3, bottom panel). In other words, Figs. 2 and 3 show that removing all data over the Pacific ocean not only has on average an almost neutral impact on forecasts verified over Europe, but also does not cause any major ‘forecast bust’.

Fig. 3 also shows that removing observations over the Pacific causes on average an increase of the 2-day 500 hPa forecast RMSE over North America more than 10 geopotential meters (which would correspond roughly to a 33% increase of the error) and about 30% of forecast error differences exceed 15 geopotential meters.

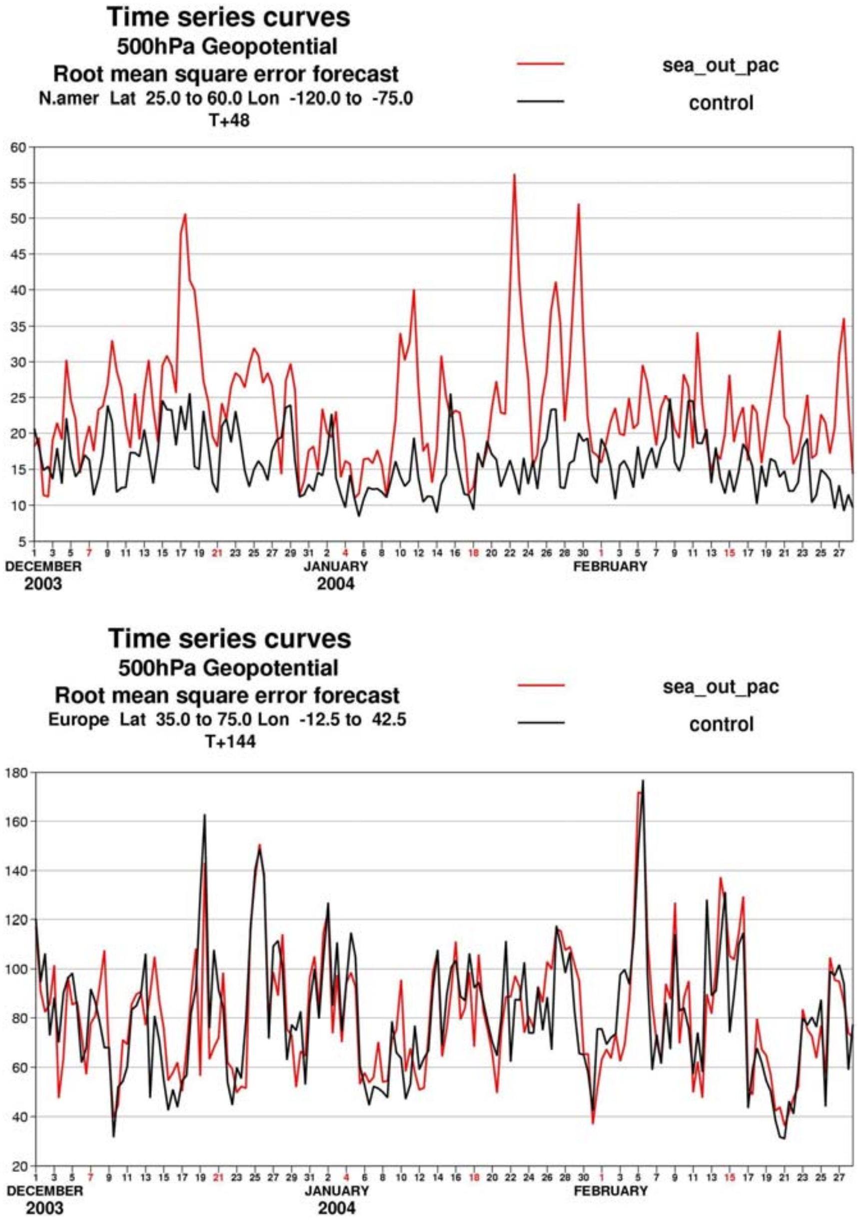


Figure 3. Winter Pacific forecasts: Verification of time series 500hPa Geopotential RMSE for day two (North America upper panel) and day six (Europe lower panel) for SEAOUT in red and SEAIN in black. Both experiments are verified using ECMWF Operational analysis.

In order to evaluate synoptically the consequence of an a priori extremely “large” error (say 20m at day 2 over a limited area) Fig. 4 compares the 2-day forecasts started at 12 UTC of the 29th of January from respectively the SEAIN and SEAOUT(Pacific) experiments for one of the cases for which the impact of the data removal increased the RMSE by more than 20m over North America. The difference in RMSE is linked to a different development of cyclonic circulations over the great lakes and Arizona: compared to the analysis (Fig. 4, top-left panel), SEAIN (Fig. 4, bottom-right panel) shows a better propagation of the cyclonic circulation over the Great Lakes and a slightly more accurate development of the cyclonic circulation over Arizona than SEAOUT(Pacific) (Fig. 4, top-right panel). However, these maps show that, despite the large difference in RMSE , the SEAIN and SEAOUT(Pacific) forecasts are rather similar synoptically, and related mainly to the development of few localised systems.

To further explore and understand the propagation of the impact of the removal of the data over the Pacific Ocean, the normalized RMSE of the forecast differences at day 1, 2, 5 and 6 between SEAIN and SEAOUT(Pacific), averaged over three months, is shown on Fig. 5 (all forecasts are validated using the ECMWF operational analysis). Fig. 5 confirms the average results displayed on Fig. 2. Indeed, the largest impact of Pacific data denial is at day 1: from then onwards the signal decreases in intensity and spreads out, reaching the Atlantic Ocean at forecast day 3. Fig. 5 also confirms that Pacific data has almost no impact on average over European region (one can notice a small and localised detrimental impact over Scandinavia, but as confirmed by Fig.2 right-bottom panel and by other statistical tests, this impact is not significant).

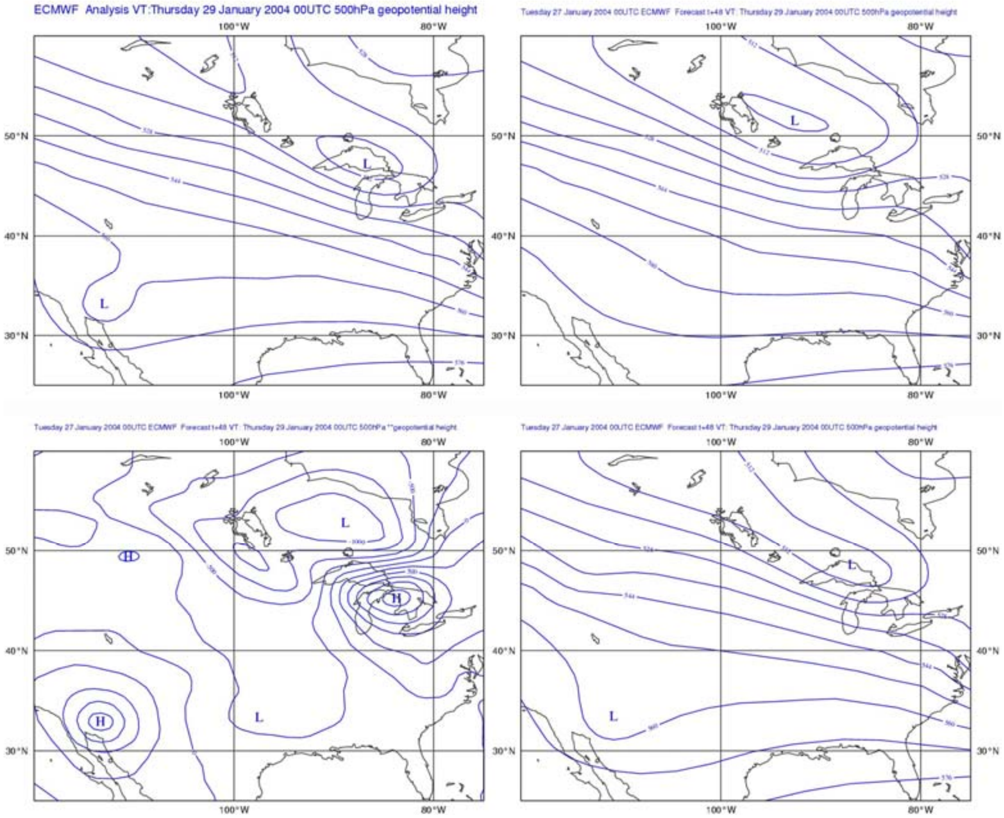
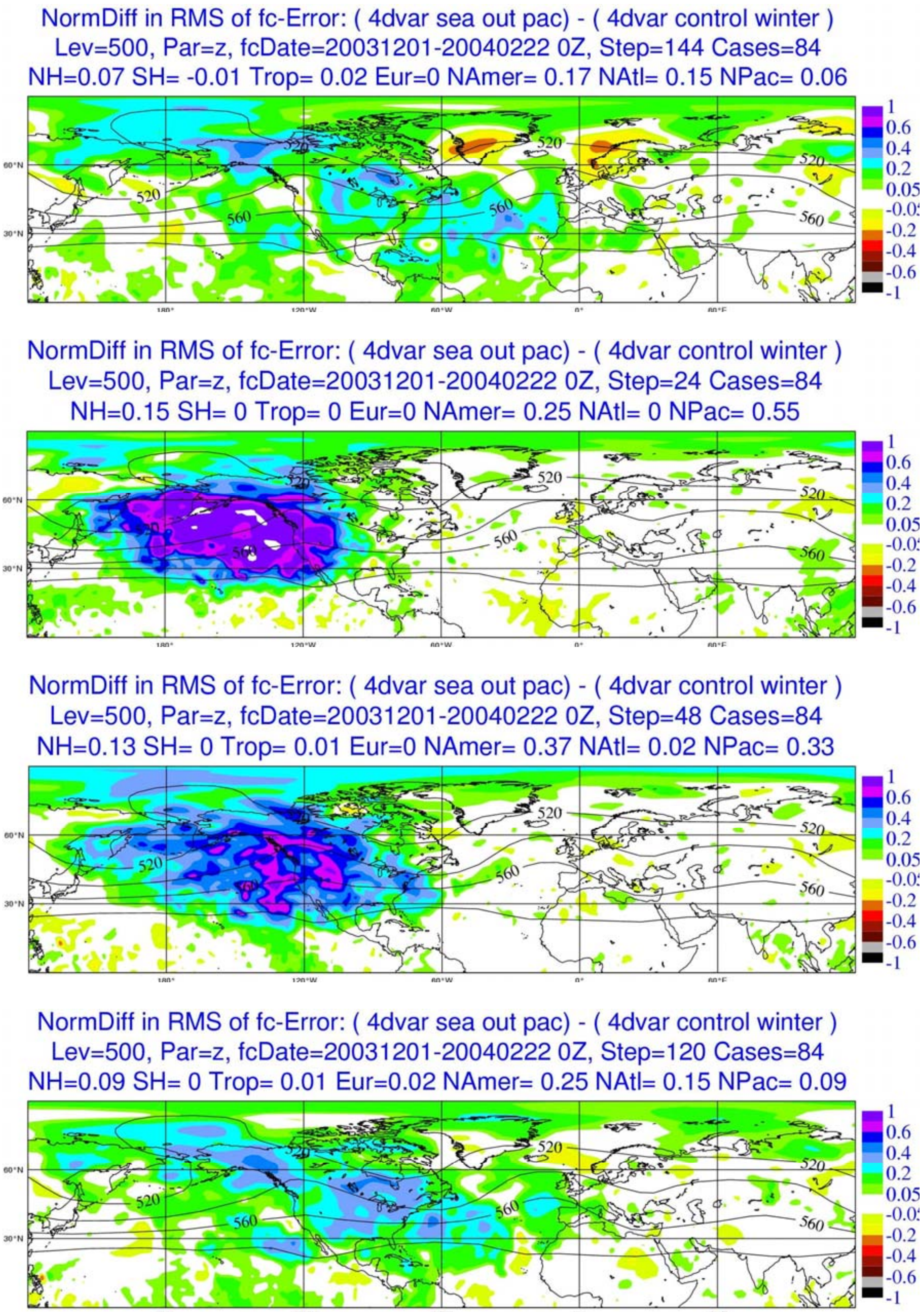


Figure 4. In this case there is a large synoptic difference in two 48hr European forecasts between SEAIN and SEAOUT. The top left panel is the verifying Operational analysis, top right the SEAOUT forecast and bottom right SEAIN 48 hour .The bottom left panel is the difference between the SEAIN and SEAOUT


NormDiff in RMS of fc-Error: (4dvar sea out pac) - (4dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040222 0Z, Step=144 Cases=84
NH=0.07 SH= -0.01 Trop= 0.02 Eur=0 NAmer= 0.17 NAtl= 0.15 NPac= 0.06

NormDiff in RMS of fc-Error: (4dvar sea out pac) - (4dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040222 0Z, Step=24 Cases=84
NH=0.15 SH= 0 Trop= 0 Eur=0 NAmer= 0.25 NAtl= 0 NPac= 0.55

NormDiff in RMS of fc-Error: (4dvar sea out pac) - (4dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040222 0Z, Step=48 Cases=84
NH=0.13 SH= 0 Trop= 0.01 Eur=0 NAmer= 0.37 NAtl= 0.02 NPac= 0.33

NormDiff in RMS of fc-Error: (4dvar sea out pac) - (4dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040222 0Z, Step=120 Cases=84
NH=0.09 SH= 0 Trop= 0.01 Eur=0.02 NAmer= 0.25 NAtl= 0.15 NPac= 0.09

Figure 5. Winter Pacific: Normalized RMSE differences between SEAIN forecast and SEAOUT. Blue-purple show the negative impact and yellow- black positive impact of SEAOUT. Panels 1 to 4 show forecasts errors for day 1,2,5 and 7.

3.2 DJF04 Atlantic-Europe data denial experiments

Fig. 6 shows the 500 hPa geopotential RMSE of the forecasts from the SEAIN (the reference experiment run using all oceanic data) and SEAOUT(Atlantic) experiment removing all Atlantic data, averaged over three months and verified over four regions: North Atlantic, Europe, North Pacific and North America.

Generally speaking, the impact over Europe of removing Atlantic data is smaller than the impact over North America of removing Pacific data (Fig.2): At day 3 or 4, SEAOUT(Atlantic) exhibits a loss of forecast skill of about 12 hours against SEAIN, to be compared with 24 hours in the case of SEAOUT(Pacific). As expected, removing data over the Atlantic Ocean does not affect on average North Pacific or North America, and this at all forecast ranges.

Fig. 7 shows the time series of RMSE of the forecasts from SEAIN and SEAOUT(Atlantic) experiments. Compared to the top panel of Fig. 3, the top panel of Fig. 7 shows that there are slightly fewer periods for which the difference between SEAOUT and SEAIN is very large (say in excess of 15 m), typically 14% of the cases. As expected, the bottom panel of Fig. 7 also confirms that the non impact of denying all data over the Atlantic Ocean is systematic in the medium range over North America (no forecast bust is identified in the time series).

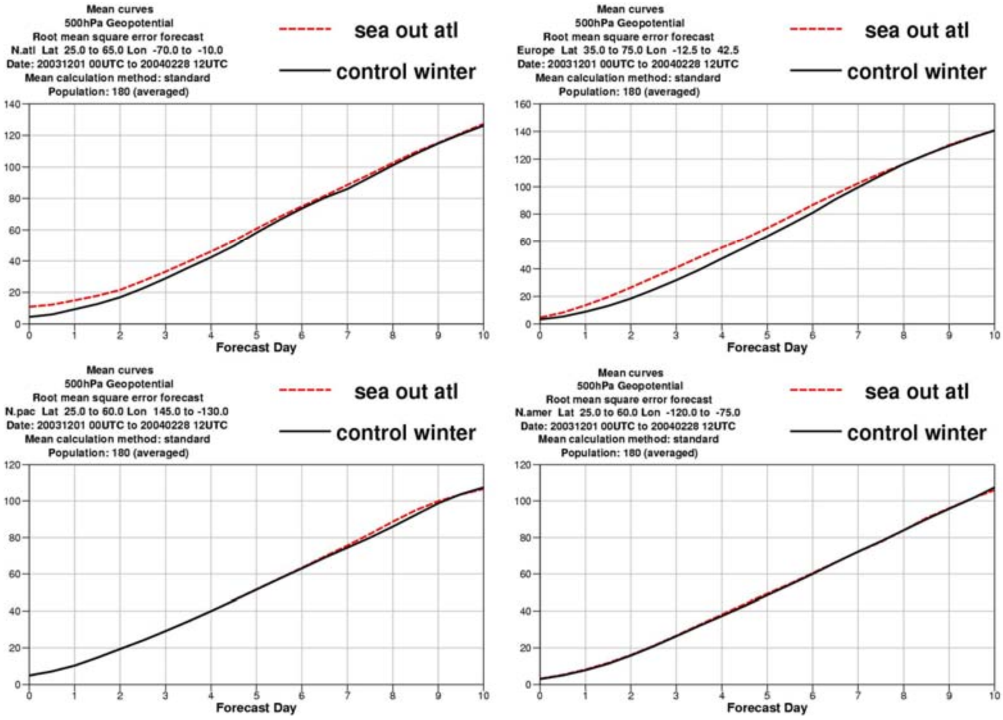


Figure 6. Winter Atlantic forecasts: Verification of mean 500hPa Geopotential RMSE for up to day ten for SEAOUT in red and SEAIN in black: Both experiments are verified using ECMWF Operational analysis. Verification regions: top left North Atlantic, top right Europe, bottom left North Pacific and Bottom right the North America.

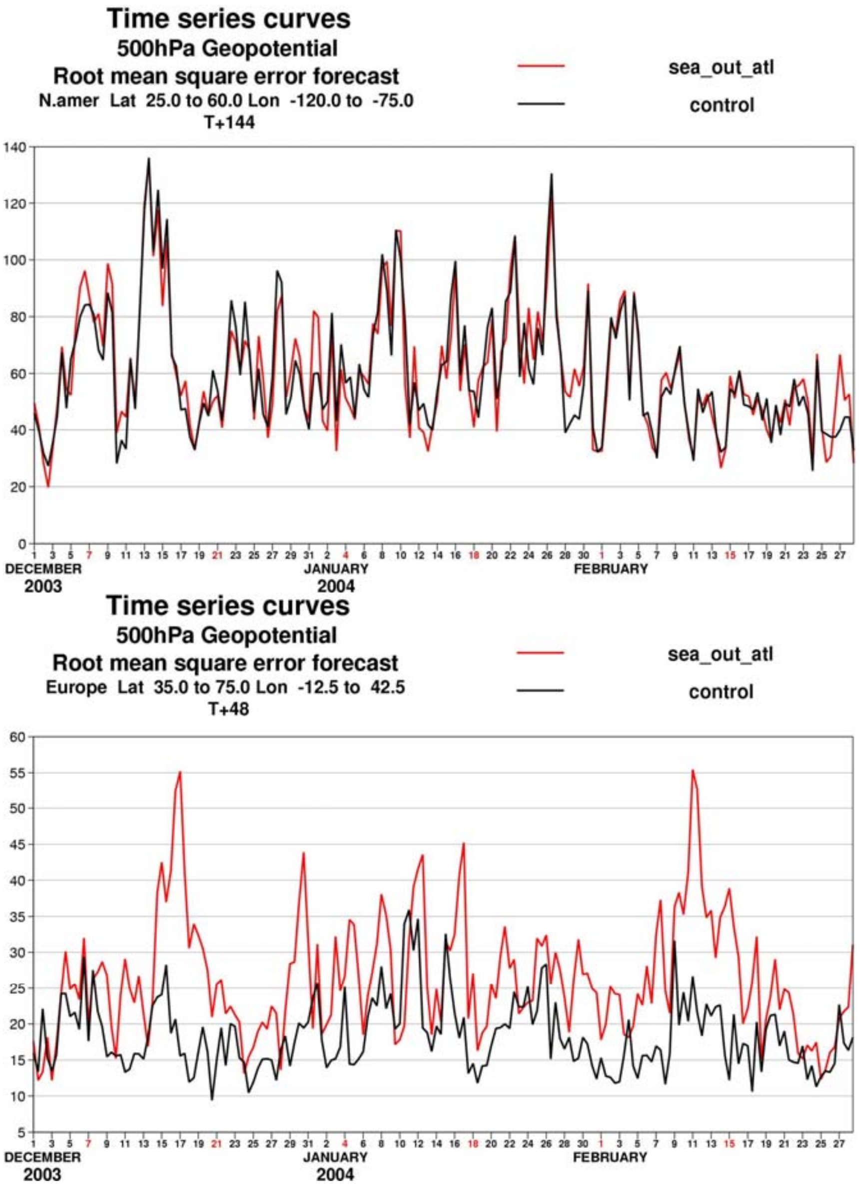


Figure 7. Winter Atlantic forecasts: Verification of time series 500hPa Geopotential RMSE for day two (Europe upper panel) and day six (North America lower panel) for SEAOUT in red and SEAIN in black. Both experiments are verified using ECMWF Operational analysis.

Fig. 8 displays the same diagnostic as Fig.5, but for SEAIN and SEAOUT (Atlantic) experiments. Overall, the normalised RMSE of the forecast differences tends to be smaller, confirming the smaller forecast impact of data over the Atlantic Ocean. This is not surprising: The Atlantic basin is a little more than half the size of the Pacific and the impact of denying observations over the Atlantic is quickly compensated by the presence of the well observed American continent in its vicinity.

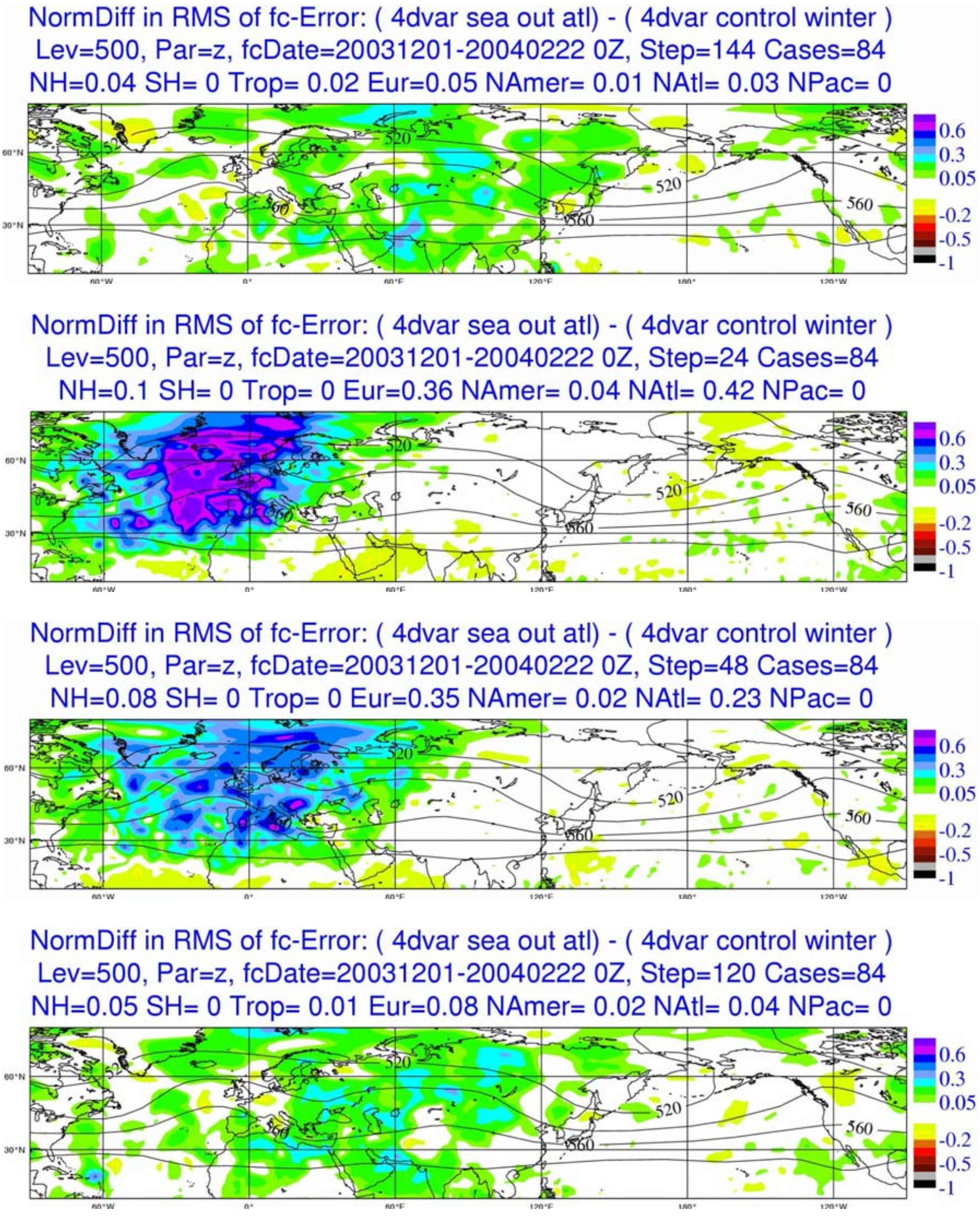


Figure 8. Winter Atlantic: Normalized RMSE differences between SEAIN forecast and SEAOUT. Blue-purple show the negative impact and yellow- black positive impact of the SEAOUT. Panels 1 to 4 show forecasts errors for day 1,2,5 and 7.

4. Summer (JJA 2004) results

This section discusses the impact of the denial experiments preformed now for the summer period. The same diagnostics as in Section 3 are presented.

4.1 JJA04 Pacific-North American data-denial experiments

Fig. 9 shows the average 500 hPa geopotential RMSE of the forecasts (as verified against the operational ECMWF analyses at the time) for experiments SEAIN (the reference experiment run using all oceanic data) and SEAOUT(Pacific), when all Pacific data are denied, verified over four regions: North Pacific, North America, North Atlantic and Europe. We notice that compared to winter (see Fig.2), the overall magnitude of the impact of the denial experiment in summer is reduced (e.g a factor of 2 at day 4 over North America) and that actually no 48 h forecast error differences exceed 15 m. (contrary to Fig 3) . However, the loss of skill remains similar (about 24 hour at day 4). Note also this time that removing data over the Pacific Ocean has a small but detectable impact over Europe at day 6-7 (however tests have been run and concluded negatively about the statistical significance of this impact).

Fig. 10 shows the times series over the summer period, of 2-day (resp. 6-day) 500 hPa forecast RMSE computed for North America (resp. Europe) (see Fig. 3 for the corresponding time series for winter 2003-04). Fig. 10 exhibits a large number of busts over North America at day 2. However, the time series of the day 6 forecast error over Europe shows that the RMSE differences between the two experiments is small (4 cases with large differences have been noted).

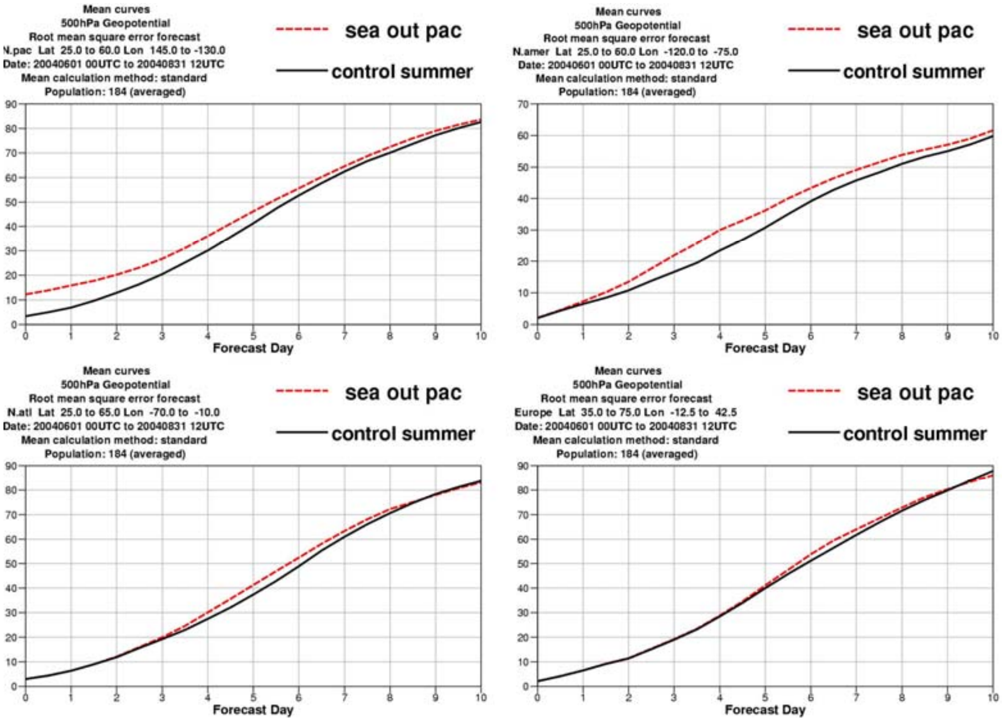


Figure 9. Summer Pacific forecasts: Verification of mean 500hPa Geopotential RMSE for up to day ten for SEAOUT in red and SEAIN in black.: Both experiments are verified using ECMWF Operational analysis. Verification regions: top left North Pacific, top right North America, bottom left North Atlantic and Bottom right the European region.

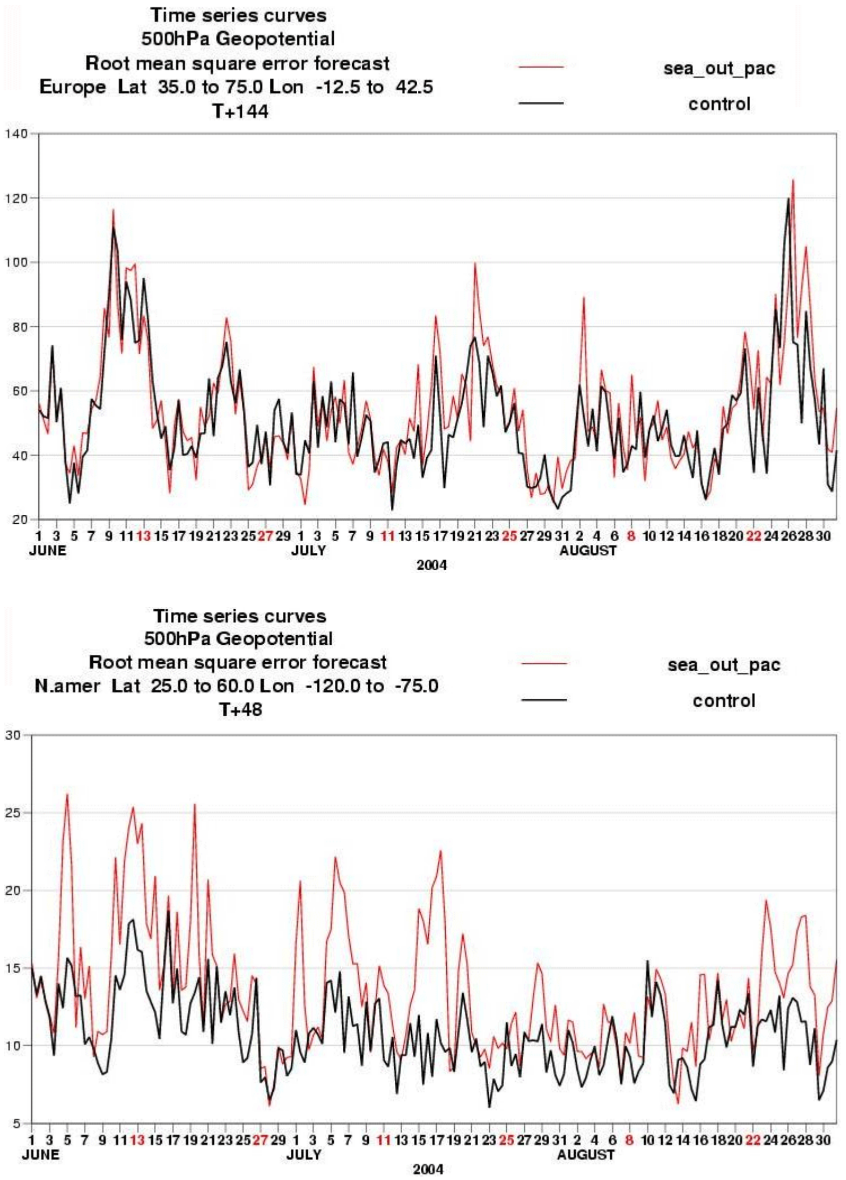


Figure 10. Summer Pacific forecasts: Verification of time series 500hPa Geopotential RMSE for day two (North America upper panel) and day six (Europe lower panel) for SEAOUT in red and SEAIN in black . Both experiments are verified using ECMWF Operational analysis.

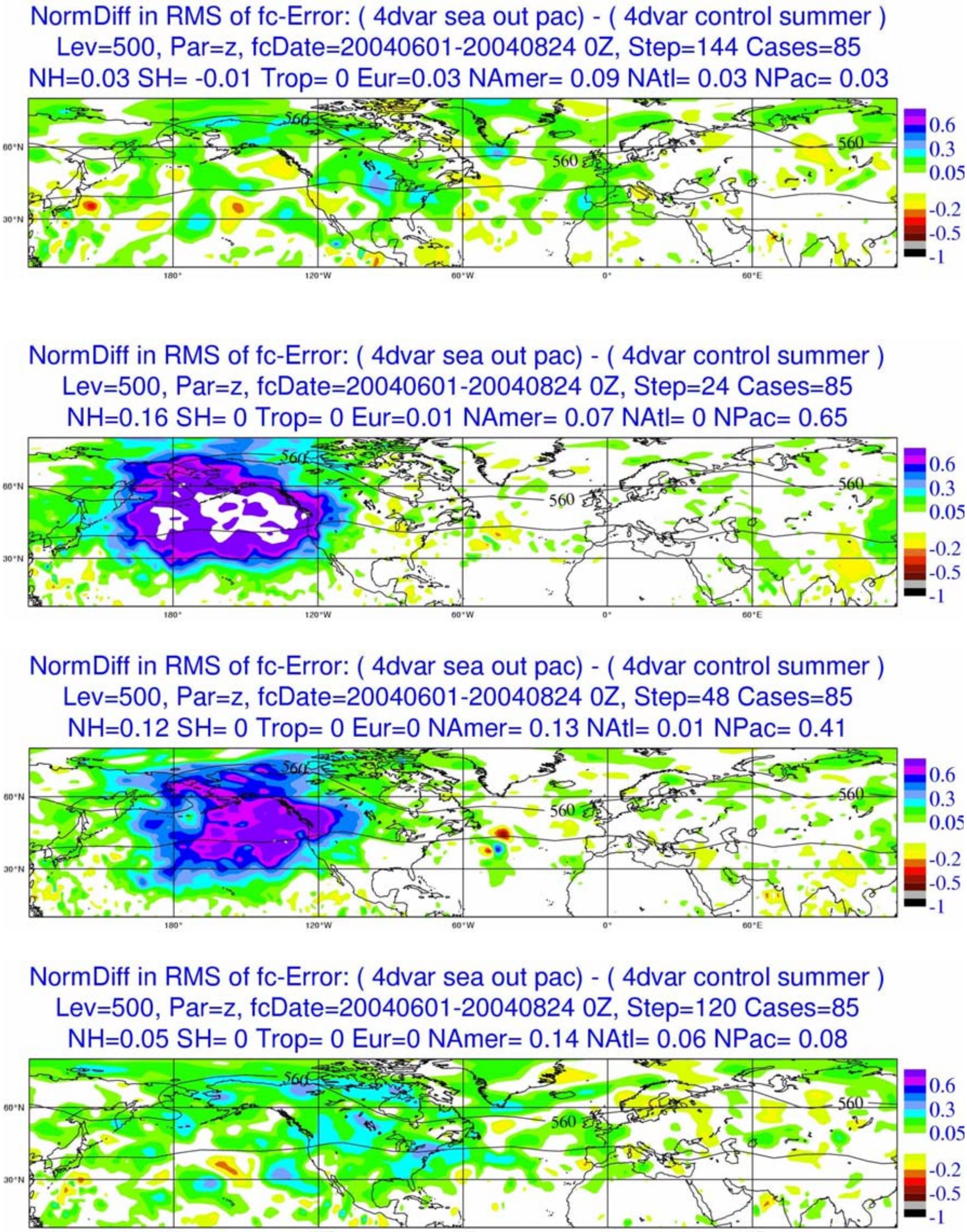


Figure 11. Summer Pacific: Normalized RMSE differences between SEAIN forecast and SEAOUT. Blue-purple show the negative impact and yellow- black positive impact of SEAOUT. Panels 1 to 4 show forecasts errors for day 1,2,5 and 7.

Normalized 500 hPa geopotential RMSE forecast differences between SEAIN and SEAOUT(Pacific) are shown in Fig. 11. The signal propagation during summer is qualitatively similar to the propagation in winter (Fig. 5) except at day 7, when a hint of degradation over Europe with SEAOUT (Pacific) can be observed.

4.2 JJA04 Atlantic-European data denial experiments

Fig. 12 shows the average 500 hPa geopotential RMSE for the forecasts (as verified against the operational ECMWF analyses at a time) for experiments SEAIN (the reference experiment run using all oceanic data) and SEAOUT(Atlantic), when all Atlantic data are denied, verified over the four same regions. Qualitatively, results are relatively similar to the one obtained from SEAOUT (Pacific) Quantitatively however, the loss of skill of experiment SEAOUT (Atlantic) is only typically 12 hours, at day four, against 24 hours for SEAOUT (Pacific).

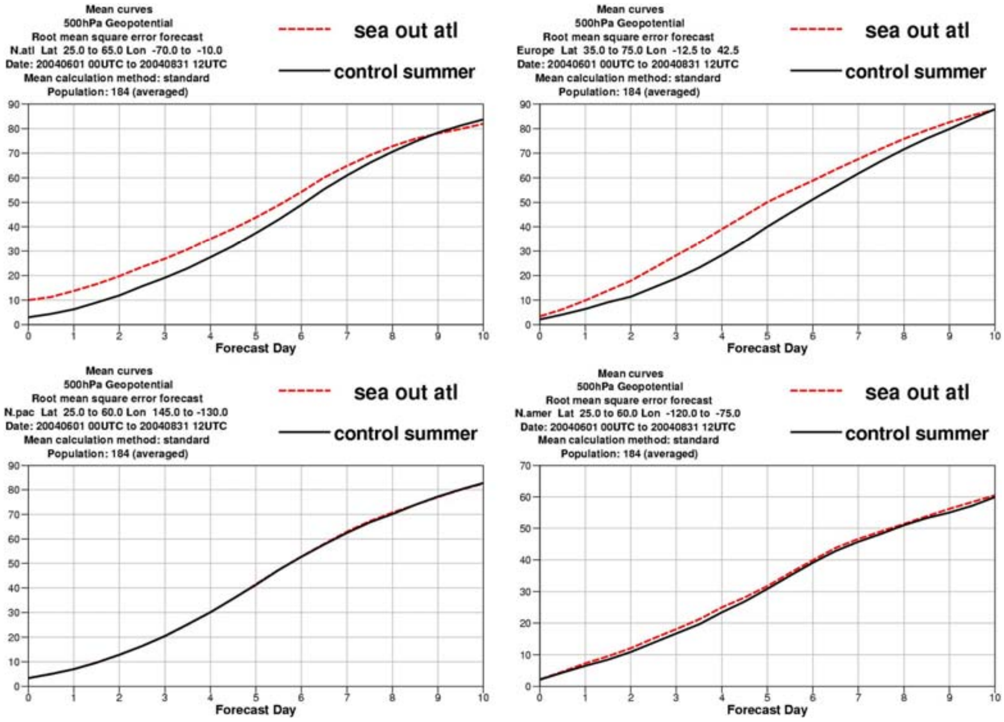


Figure 12. Summer Atlantic forecasts: Verification of mean 500hPa Geopotential RMSE for up to day ten for SEAOUT in red and SEAIN in black: Both experiments are verified using ECMWF Operational analysis. Verification regions: top left North Atlantic, top right Europe, bottom left North Pacific and Bottom right the North America.

Fig. 13 shows the time series of the 500 hPa geopotential RMSE of the forecasts from the SEAIN and SEAOUT(Atlantic) forecasts verified at day 2 over Europe and day 7 over North America, and Fig. 14 shows the normalized 500 hPa geopotential RMSE of the forecasts differences, averaged over the summer period, between the two experiments. Results are again qualitatively similar to the ones obtained with SEAOUT (Pacific). Conclusions are similar, i.e. the impact of denying observations remains confined to the immediate downstream region and does not propagate further.

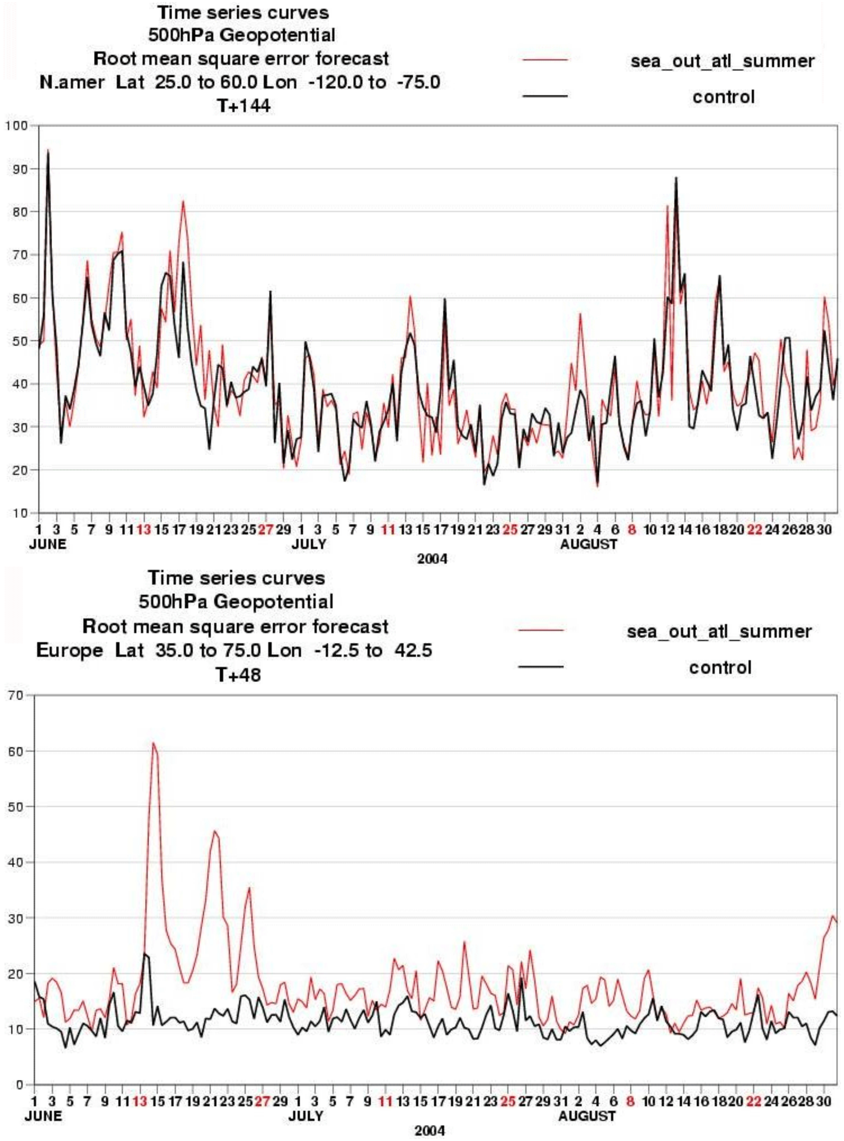


Figure 13. Summer Atlantic forecasts: Verification of time series 500hPa Geopotential RMSE for day two (Europe upper panel) and day six (North America lower panel) for SEAOUT in red and SEAIN in black. Both experiments are verified using ECMWF Operational analysis.

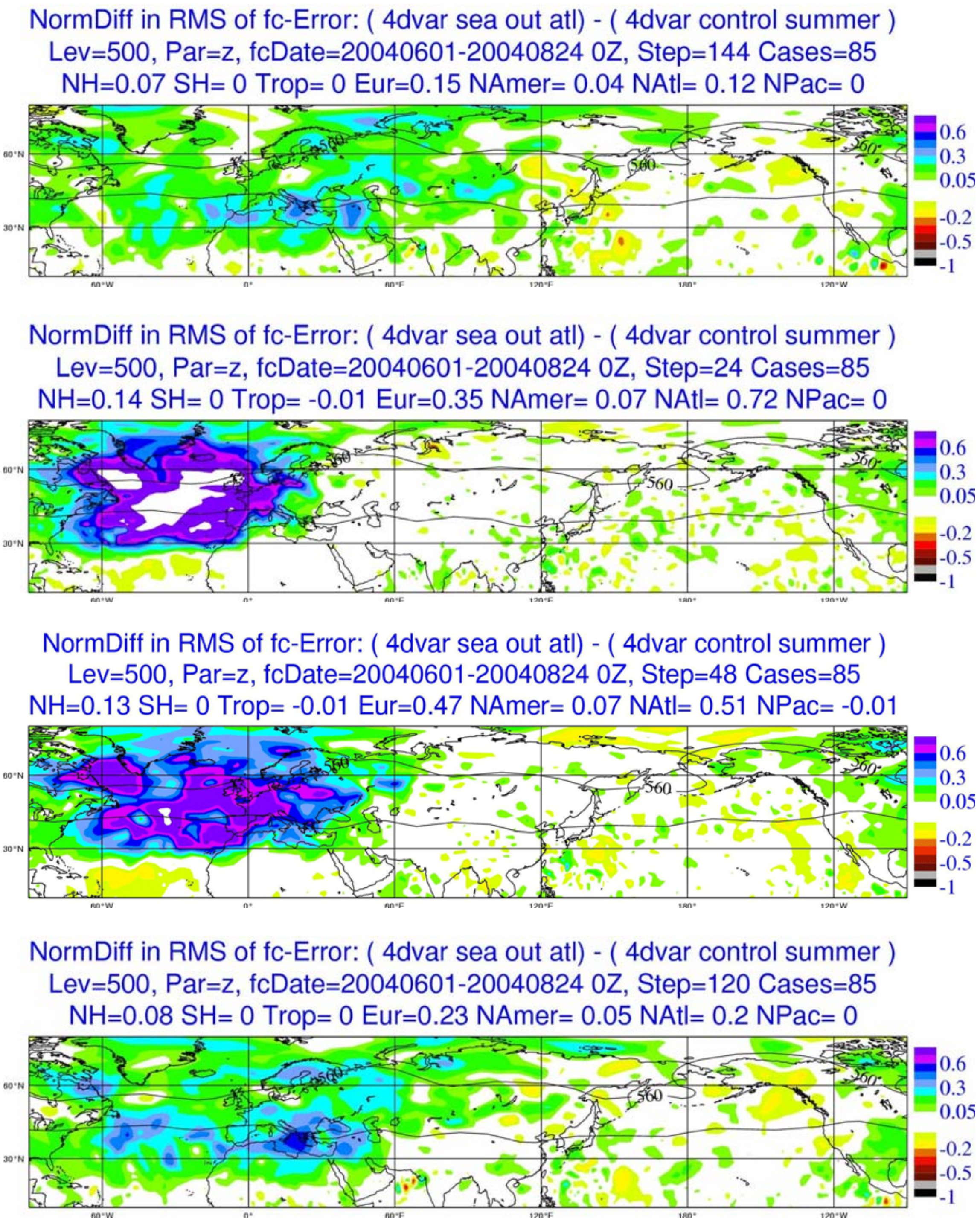


Figure 14. Summer Atlantic: Normalized RMSE differences between SEAIN forecast and SEAOUT. Blue-purple show the negative impact and yellow- black positive impact of SEAOUT. Panels 1 to 4 show forecast errors for day 1,2,5 and 7.

The comparison of the winter and summer results indicate that, on average, the impact of ocean data denial is smaller in summer, due to the weaker summer circulation and the fact that, in the Northern Hemisphere, errors grow less rapidly in summer than in winter (this is reflected also in the fact that singular vectors grow less rapidly in summer than in winter, see Buizza and Palmer 1995).

5. Synoptic assessment of Atlantic data denial experiments.

An interesting question is whether the impact of a given component of Observing System as measured by standard scores such as RMSE of the forecasts over a given area, can be translated in tangible synoptic meteorological features.

It is of course valid to express the impact of a measurement program in terms of an average percentage reduction of the RMSE of the forecasts, or in terms of average gains in predictability of 'x' hours at forecast day 'y', but very rarely these average verifications are compared to improvement in terms of improvement in the description of the synoptic scale flow. This section focuses on this issue, and shows how the average impact of oceanic data denial translates in terms of synoptic evaluation. This discussion is restricted here to the 48 h forecasts over Europe (based on SEAIN and SEAOUT (Atlantic) experiments), as it already required considerable effort to manually classify each pair of 48 hour forecasts. In the future a further study may be done with the North American forecasts.

First, SEAIN and SEAOUT (Atlantic) forecasts have been subjectively compared, and each case has been classified accordingly to their difference into four different categories: no difference, small, medium and large differences. Forecasts with no synoptic difference have been assigned a value of 0, while forecasts with small, medium and large differences have been assigned, respectively, a value of 2.5, 5 or 15 units. The numerical values were chosen to convert the synoptic classification to numerical values that could be plotted alongside the timeseries of RMSE error of 48 hour forecasts differences, see Figs. 15 and 16. Table 2 summarizes this classification. These time series of 'synoptic difference value' have then been compared to time series of 'RMSE of the forecast differences', to assess whether cases with a large synoptic value (i.e. with a large synoptic difference) corresponded to cases with a large difference in RMSE of the forecasts. The RMSE forecast verification grid (0.5 latitude x 0.5 longitude) was used to represent the small synoptic scales. This is finer than the standard ECMWF verification grid.

Subjective analysis of synoptic difference	**Assigned "synoptic difference value"**
Unable to distinguish	0
Small difference	2.5
Medium difference	5
Large difference	15

Table 2 Categories used to classify synoptic-scale differences, and 'synoptic difference values' assigned to each category. These numerical values were subjectively chosen by comparison with RMSE of forecast differences to enable both forecast skill values to be graphically displayed (see Figs. 15 and 16).

Fig. 15 shows the 2-day RMSE of the forecasts from the SEAIN and SEAOUT (Atlantic) over Europe, the difference between these RMSE of the forecasts, and the 'synoptic difference value' defined above. Note that there is a rather good agreement between cases with a large RMSE of the forecast difference and cases with a large 'synoptic difference value'. Similar results have been obtained for winter (Fig. 16). This shows a good consistency between an objective scoring system and the subjective and visual analysis that has been performed on each individual case.

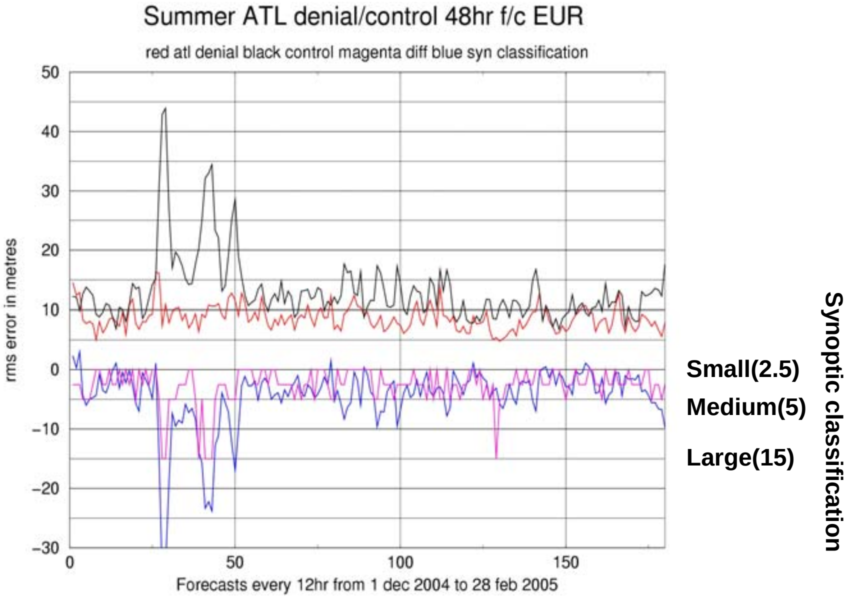


Figure 15. Comparison of Synoptic classification of the Summer Atlantic forecasts and RMSE errors and RMSE differences: Upper panel time series 500hPa Geopotential RMSE for up to day two for Europe SEAIN in red and SEAOUT in black with both experiments are verified using ECMWF Operational analysis. Bottom panel the difference of the two forecasts (SEAIN and SEAOUT) for the synoptic classification index (magenta) and the RMSE difference (blue).

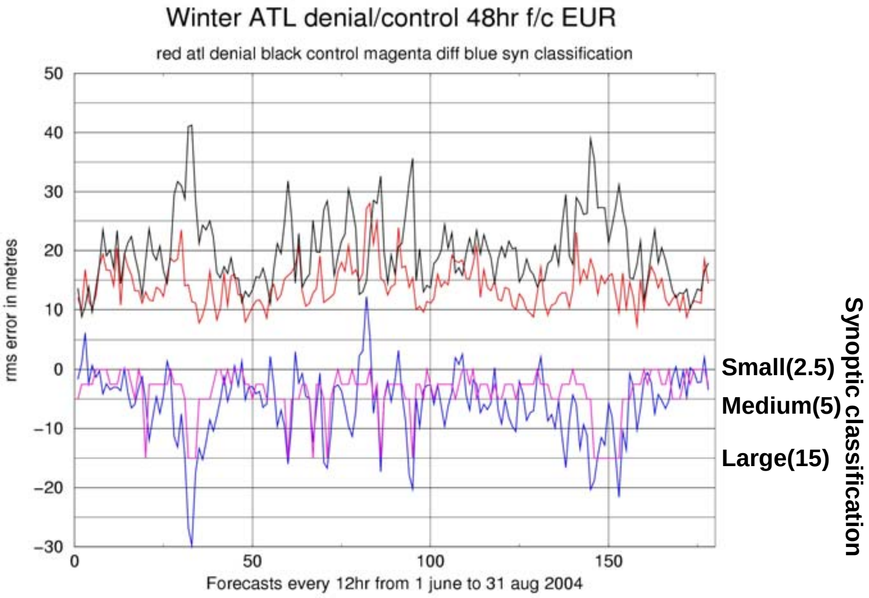


Figure 16. Comparison of Synoptic classification of the Winter Atlantic forecasts and RMSE errors and RMSE differences: Upper panel time series 500hPa Geopotential RMSE for up to day two for Europe SEAIN in red and SEAOUT in black with both experiments are verified using ECMWF Operational analysis. Bottom panel the difference of the two forecasts (SEAIN and SEAOUT) for the synoptic classification index (magenta) and the RMSE difference (blue).

Table 3 lists the total number of cases for which either a large RMSE of the forecast differences or a large ‘synoptic difference value’, or both, have been detected during summer 2004 and winter 2003-04. Table 3 confirms the good consistency between the two diagnostics and that overall there are 10 summer (20 winter) cases characterized by either a large RMSE or a large synoptic difference, out of the 180 (178) cases considered These results suggest that, subjectively, it is only when forecast errors differ by more than 15m that the forecast differences become visible and detectable when looking at synoptic maps. It should however be mentioned that this visual evaluation is intrinsically subjective and restricted to 500 hPa geopotential maps. A full synoptic evaluation, looking at weather parameters remains to be done to provide a full comparison between synoptic evaluation and objective scoring.

Period	**No of cases (Two per day)**	**RMSE> 15m**	**RMSE > 15m and Large Synoptic difference**	**Large Synoptic difference**	**No of cases with either RMSE >15m or large synoptic difference**
Summer	180	8	5	7	10
Winter	178	14	10	16	20

Table 3 Total number of cases (column 1) and number of cases with: 2-day forecasts verified over Europe characterized by: RMSE of the forecast differences larger than 15m (column 2), with RMSE of forecast differences larger than 15m and 'synoptic difference values' equal to 15 units (column 3), with 'synoptic difference values' equal to 15 units (column 4), and with either 'synoptic difference values' equal to 15 units or RMSE of forecast differences larger than 15m (column 5)..

Fig. 17 shows the case with the largest synoptic difference in the 2-day forecast over Europe. A small low pressure system is located south of UK (see verifying analysis SEAIN, top left panel). The SEAIN 2-day forecast (top right panel) evolves this low slightly too slowly, whereas the SEAOUT (Atlantic) forecast (bottom right panel) intensifies the low too much north over the UK. Both forecasts differ from the analysis, and their difference (bottom left panel) has an RMSE over Europe of 16 m.

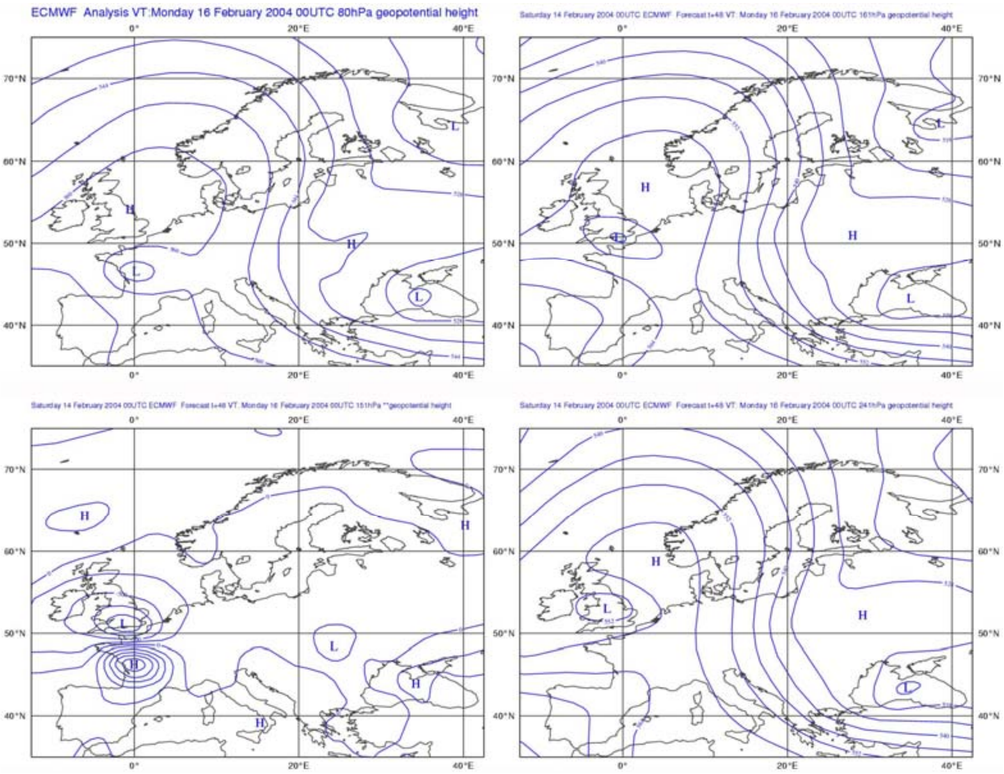


Figure 17. In this case there is a large 'synoptic difference' in two 48hr European forecasts between SEAIN and SEAOUT. The top left panel is the verifying Operational analysis, top right the SEAIN forecast with a 48 hour forecast error RMSE error of 16 m and bottom right SEAOUT 48 hour forecast error of 24 m. The bottom left panel is the difference between the SEAIN and SEAOUT forecasts.

Fig. 18 shows the case characterised by the SEAIN 2-day forecast with the largest RMSE over Europe (27 m). In this case, the error of the SEAOUT(Atlantic) forecast is almost half (15 m). From a synoptic point of view, the forecast difference between the SEAIN and the SEAOUT(Atlantic) forecasts has been classified as ‘large’ because of the poor positioning of the low-pressure system located over the North Sea and because of the large differences in the atmospheric flow over the Adriatic Sea. These two figures are just examples aiming at providing an illustration of the correspondence between a forecast error difference of 15m or higher, and the synoptic difference that can be seen comparing two maps over a region such as Europe.

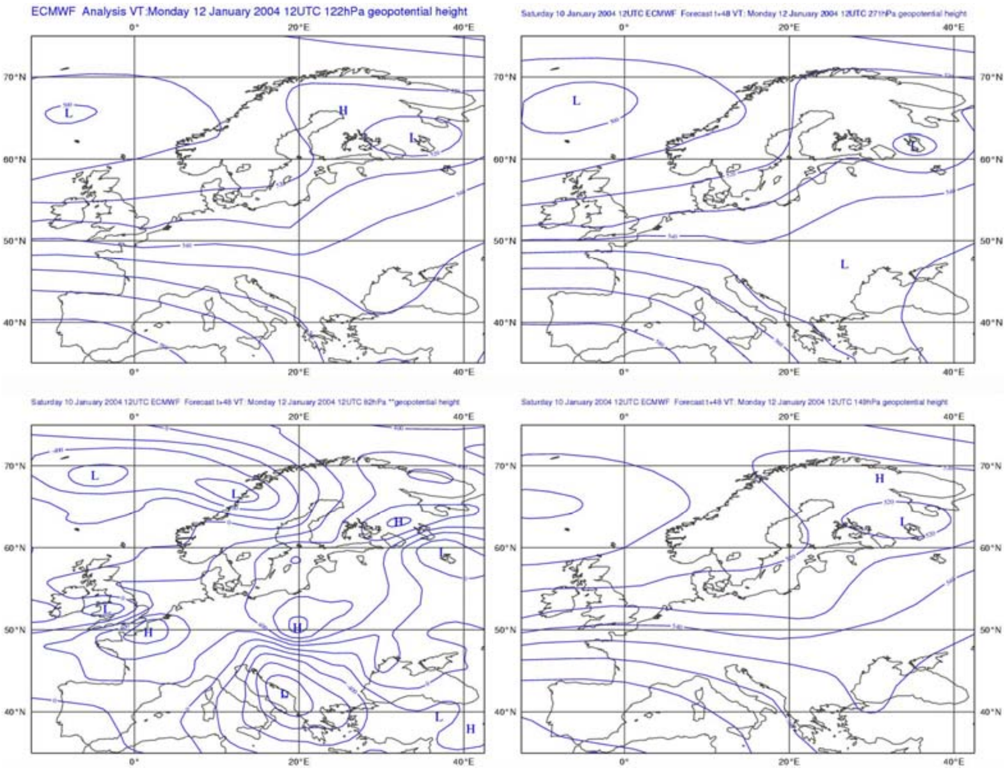


Figure 18. In this case there is a large RMSE difference in two 48hr European forecasts between SEAIN and SEAOUT. The top left panel is the verifying Operational analysis, top right the SEAIN forecast with a 48 hour forecast error RMSE error of 27 m. and bottom right SEAOUT 48 hour forecast error of 15 m. The bottom left panel is the difference between the SEAIN and SEAOUT forecasts.

Considering the questions posed at the beginning of this section, this analysis indicates that there is a reasonably good agreement between forecast cases characterised by a large difference in the RMSE and cases with a large synoptic difference, and that the value of RMSE differences needs to exceed about 15 meters to correspond to a large disagreement between synoptic maps.

6. Comparison of 3D and 4D-Var data ECMWF assimilation systems.

To assess the sensitivity of the results discussed above to the data-assimilation system, an additional pair of 3D-Var experiments (SEAIN and SEAOUT) has been run for the winter period at the same horizontal and vertical resolution than operational ECMWF 4D-Var system used in this study. 4D-Var has several advantages over 3D-Var, the two main ones being a) a better use of asynoptic data and 2) a longer assimilation window (12 hour versus 6 hour).

These experiments have only been run only for the Pacific data denial study (SEAOUT (PAC)), because the largest impact was obtained from the original 4D-VAR data denial experiment and also because of computer resources limitation.

6.1 Experimental results

Fig. 19 shows the average 500 hPa geopotential forecast RMSE scores from the SEAIN and SEAOUT experiments started from both the 3D- and the 4D-Var analyses (and verified over the four areas mentioned previously). Fig. 19 clearly shows that forecasts started from 4D-Var analyses are more accurate than corresponding forecasts started from 3D-Var analyses (with or without data over the Pacific Ocean). Part of this signal is due to a longer assimilation window used in 4D-Var. Fig. 19 also shows that the average difference between SEAIN and SEAOUT is larger for the 3D-Var experiment. This is confirmed on Fig. 20 by the corresponding time series of the 2-day (respectively 6-day) 500 hPa geopotential RMSE of the forecasts over North America (resp. Europe). Fig. 20 shows that removing oceanic data is more likely to create forecast 'busts' if a 3D-Var assimilation system is used instead of 4D-Var. In other words, 4D-Var has not only the effect of producing, on average, more accurate forecasts, but it also shows its power to cope with the lack of data over a large area within the assimilation procedure.

Fig. 21 shows the normalized 500 hPa geopotential RMSE difference between the SEAIN and SEAOUT 3D-Var forecast experiments at forecast day 1, 2, 5 and 6. Compared to the signal propagation detected in the 4D-Var experiments (Fig. 5), the 3D-Var signal extends furthest from the Pacific Ocean, and have a sizeable impact also at forecast day 7 over Europe and Asia.

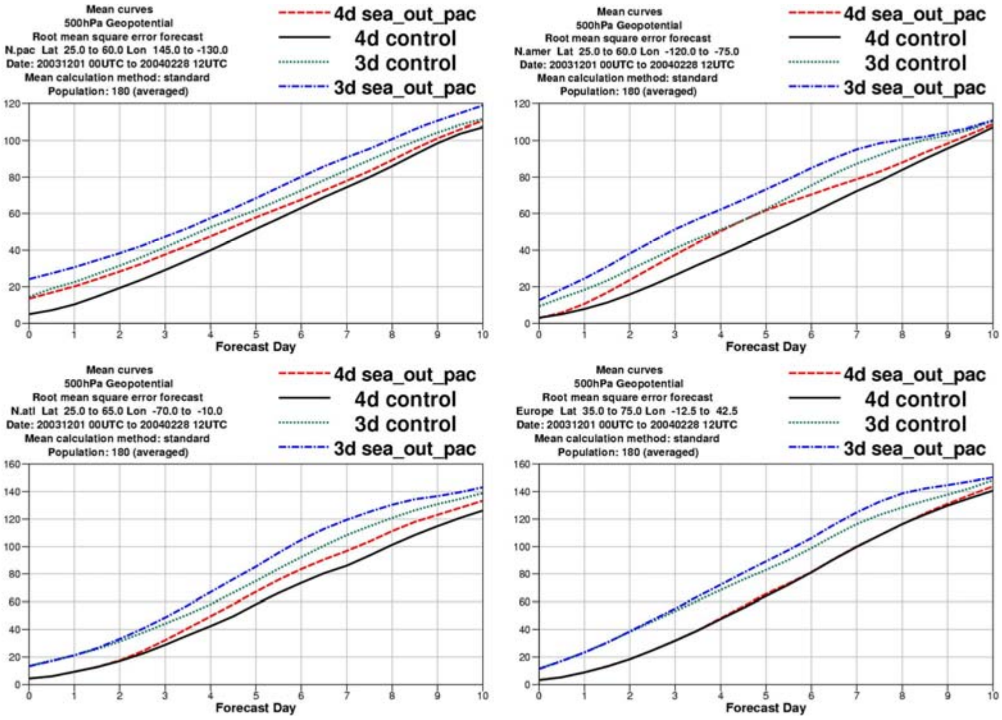


Figure 19. Comparison 4D-Var and 3D-Var for the Winter Pacific forecasts: Verification of mean 500hPa Geopotential RMSE for up to day ten for SEAOUT (4D-Var) in red, SEAOUT (3D-Var) in blue, SEAIN (4D-Var) in black and SEAIN(3D-Var) in green.: Both experiments are verified using ECMWF Operational analysis. Verification regions: top left North Pacific, top right North America, bottom left North Atlantic and Bottom right the European region.

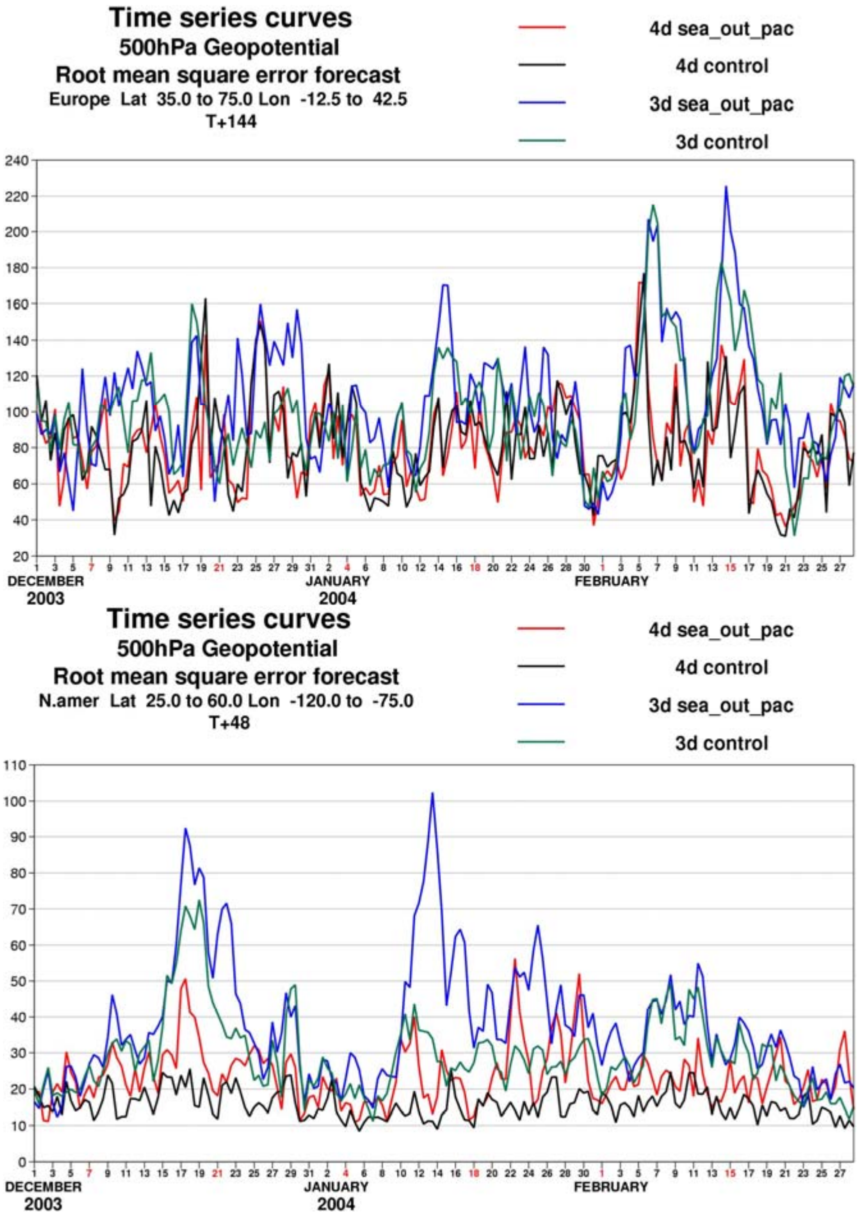
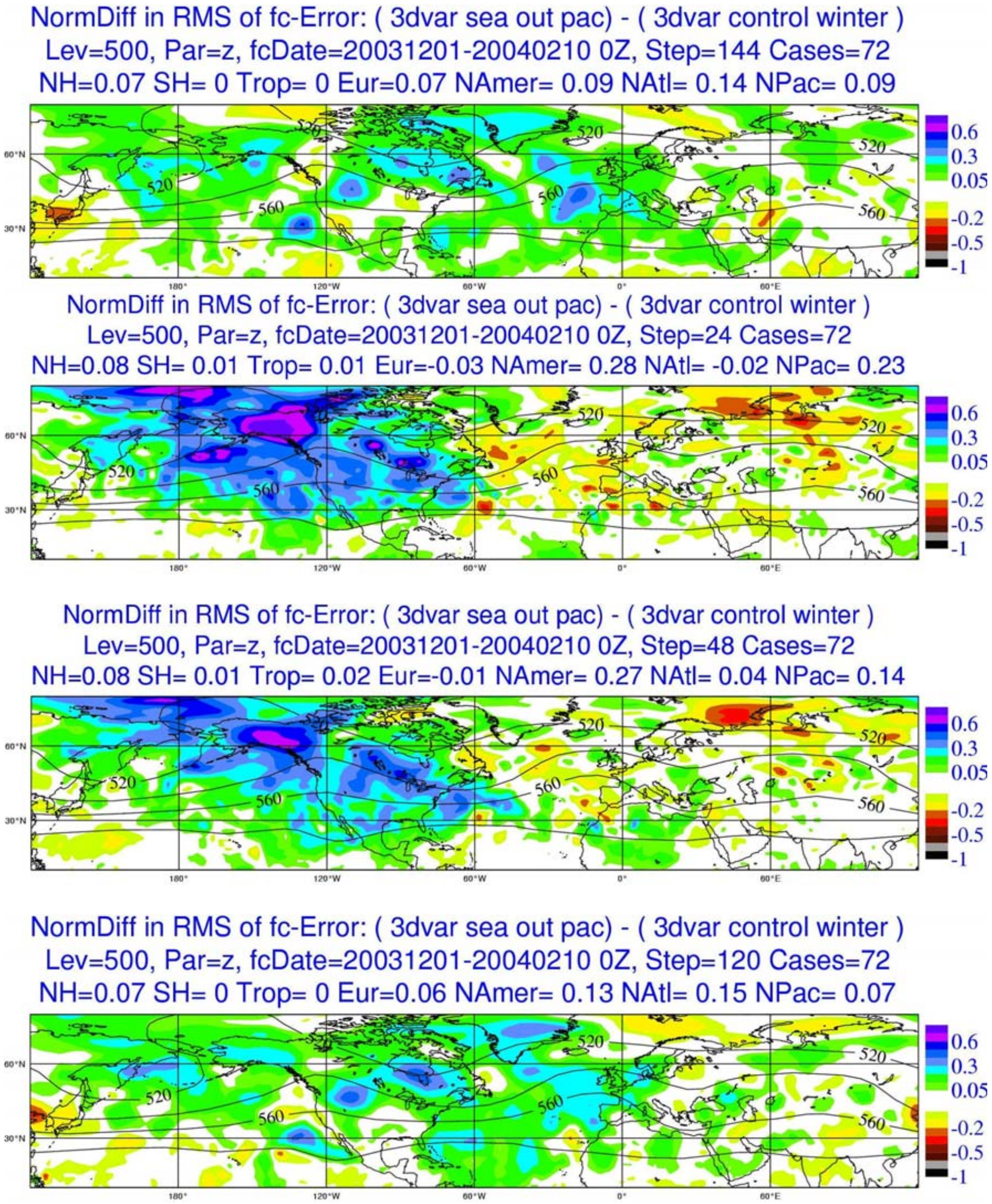


Figure 20. Comparison 4D-Var and 3D-Var for the Winter Pacific forecasts: Verification of time series 500hPa Geopotential RMSE for SEAOUT (4D-Var) in red, SEAOUT (3D-Var) in blue, SEAIN (4D-Var) in black and SEAIN(3D-Var) in green.: Both experiments are verified using ECMWF Operational analysis. Verification regions: top left North Pacific, top is day 2 forecasts for North America, bottom day 7 forecasts for the European region.



NormDiff in RMS of fc-Error: (3dvar sea out pac) - (3dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040210 0Z, Step=144 Cases=72
NH=0.07 SH= 0 Trop= 0 Eur=0.07 NAmer= 0.09 NAtl= 0.14 NPac= 0.09

NormDiff in RMS of fc-Error: (3dvar sea out pac) - (3dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040210 0Z, Step=24 Cases=72
NH=0.08 SH= 0.01 Trop= 0.01 Eur=-0.03 NAmer= 0.28 NAtl= -0.02 NPac= 0.23

NormDiff in RMS of fc-Error: (3dvar sea out pac) - (3dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040210 0Z, Step=48 Cases=72
NH=0.08 SH= 0.01 Trop= 0.02 Eur=-0.01 NAmer= 0.27 NAtl= 0.04 NPac= 0.14

NormDiff in RMS of fc-Error: (3dvar sea out pac) - (3dvar control winter)
Lev=500, Par=z, fcDate=20031201-20040210 0Z, Step=120 Cases=72
NH=0.07 SH= 0 Trop= 0 Eur=0.06 NAmer= 0.13 NAtl= 0.15 NPac= 0.07

Figure 21. 3DVAR mean RMSE error. Winter Pacific: Normalized RMSE differences between SEAIN forecast and SEAOUT. Blue-purple show the negative impact and yellow- black positive impact of SEAOUT. Panels 1 to 4 show forecast errors for day 1,2,5 and 7.

6.2 Interpretation of the sensitivity of the added value of the observations to the data assimilation system

The fact that denying observations in 3D-Var has a larger impact than in 4D-Var should be interpreted with care. Fig. 22 displays an idealistic representation of the gain in analysis accuracy as a function of the number of observations (all units and numbers are arbitrary here: on the x-axis, more than 200 observations tend towards infinity, and on the y-axis, one indicates an arbitrary limit in terms of analysis improvement towards which the contribution from any additional observation would be infinitesimal). 4D-Var and 3D-Var should give the same results in terms of analysis accuracy when there are no observations and when there are infinitely many observations (the assimilation system does not matter anymore). Our experience, backed up by many impact studies performed in the past, suggests that the analysis improvement in case of 4D-Var is likely to follow the blue curve, while a less performing system (e.g. 3D-Var) will follow the red curve.

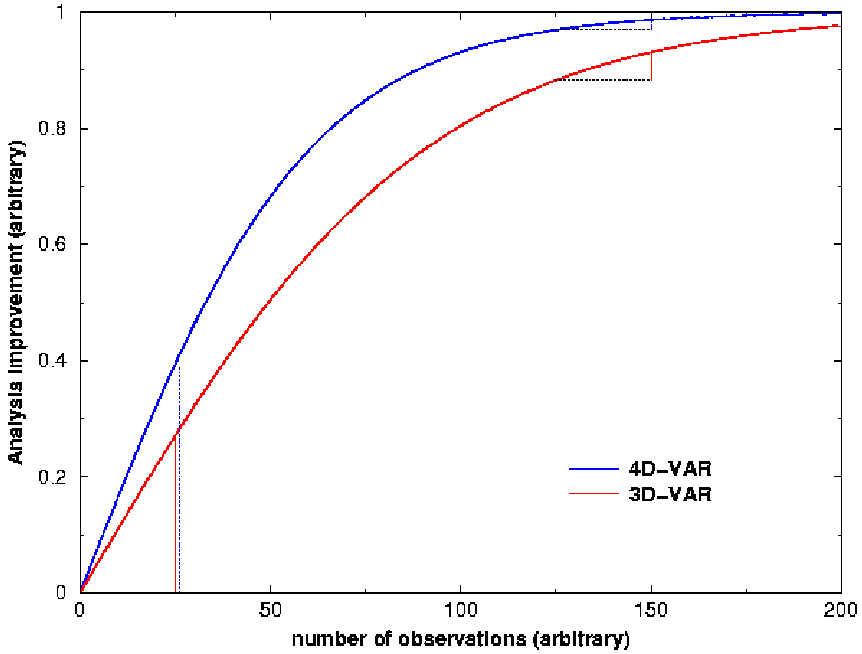


Figure 22: Idealistic representation of the analysis improvement as a function of the number of assimilated observations (numbers and units are arbitrary). Blue: 4D-VAR, RED (3D-Var). The vertical full red bars represent the improvement of the 3D-VAR analysis when one goes from zero to 25 observations, and from 125 to 150 observations. The vertical dotted blue bars represent the same quantities for 4D-VAR.

For example, going from 0 to 25 observations will lead to a .26 improvement in 3D-Var (red vertical bar), against .39 (blue dotted vertical bar) in 4D-Var. It has indeed been demonstrated and documented in Thépaut (2006) that in the sole presence of surface pressure observations, 4D-Var was outstandingly better than 3D-Var, indicating a much sharper curve in terms of analysis improvement as a function of number of observations in a data void context. The experiments reported here are performed in a system overwhelmed by observations (in particular satellite data), which probably corresponds to a regime of around 150 observations in our idealistic context. Although 4D-Var is clearly better than 3D-Var in our experimental framework, the incremental gain (loss) achieved by adding (denying) an overall limited number of observations will be higher in 3D-Var than in 4D-Var (red and blue vertical bars around 150 observations: increment of .017 for 4D-Var against .048 for 3D-Var).

7. Conclusions

This paper is the first of three companion papers that summarize the results of a large number of data-assimilation experiments performed to address some key questions linked to adaptive observation targeting, specifically the importance of data over the ocean in a mean statistical sense and on individual cases.

These questions have been addressed considering both 3 and 4-dimensional variational data-assimilation experiments performed at fairly high resolution (T511/T159 outer/inner loops) using the ECMWF system that was operational in June 2005 for two periods, winter 2003/2004 (DJF04, 91 days) and summer 2004 (JJA04, 90 days).

The key conclusions that can be drawn from the experiments discussed in this work are the following:

(i) With regard to the 4D-Var data denial experiments, the Pacific oceanic data is more important in terms of 2 day forecast impact over North America than the Atlantic oceanic data in terms of 2 day forecast impact over Europe, although both denial experiments show a degradation downstream. The comparison of 4D- and 3D-Var experiments have indicated that removing oceanic data has a larger impact on downstream forecast accuracy if a 3D-Var assimilation system is used.

(ii) In the 4D-Var system the influence of observations remains fairly local and mainly affects the immediate downstream region throughout the whole 10 day forecast range. In particular, on average there is little impact of removing observations in the Pacific Ocean on medium-range forecasts over Europe. Similarly, removing observations in the Atlantic Ocean does not affect forecasts for the Pacific/North American region. Again, the comparison of 4D- and 3D-Var experiments indicates that the impact of removing oceanic observation is less local if a 3D-Var assimilation system is used.

(iii) On a few, selected cases the impact of removing oceanic data can be very large and detectable on 500 hPa geopotential synoptic maps. In particular, a detailed synoptic evaluation performed for the Atlantic/European region has indicated that for large synoptic differences between the 48 hour forecasts of SEAIN (the experiment run using all oceanic data) and SEAOUT (the experiment run removing all oceanic data over the Atlantic), the forecast RMSE differences could exceed 15 geopotential meters on a small number of data denial forecasts (16% in winter and 14% in summer). The Pacific SEAOUT experiments have a much larger effect than the Atlantic ones. The impact remains relatively small in Summer.

These results provide a framework to test the concept of targeting: the differences between the oceanic denial experiment (SEAOUT) and the control analyses (SEAIN) will be used as a reference in the second part of this work (BCKT06), where the value of observations taken in small, targeted oceanic regions is discussed. One of the interesting outcomes of this study is the local nature of data denial when 4D-Var is used to assimilate the data. The propagation of the analysis error down-stream reduces rapidly, and little effect on any European forecasts was detected from the Pacific denial experiments. This result is an indication of the superiority of 4D-Var over the previous 3D-Var assimilation system, and a proof of its robustness and capacity to compensate from the lack of accurate observations.

8. References

Bouttier, F. and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Q. J. R. Meteorol. Soc.* **127**, 1469-1488.

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

Buizza, R., C. Cardinali, G. Kelly, and J.-N. Thepaut, 2006: The value of targeted observations Part II: the value of observations taken in singular vectors-based target areas. *Q. J. R. Meteorol. Soc.*, submitted.

Cardinali, C., Buizza, R., Kelly, G and J.-N. Thepaut, 2006: The value of targeted observations Part III: influence of different weather regimes. *Q. J. R. Meteorol. Soc.*, submitted.

Courtier, P., J.N. Thépaut , and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367-1388.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabierand, and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var) I: Formulation. *Q. J. R. Meteorol. Soc.*, **124**, 1783-1807.

Kelly, G. 1997:Influence of observations on the operational ECMWF system. *Bull. of WMO,* **46**, 336-341.

Kelly, G., A. McNally, J.-N. Thépaut and M. Szyndel, 2004. Observing System Experiments of all main data types in the ECMWF operational system. In proceedings of the *Third WMO Workshop on the Impact of Various Observing Systems on NWP* Alpbach, Austria, 9 - 12 March 2004 .

Langland, R. H., 2006: Issues in targeted observing. *Q. J. R. Meteorol. Soc.*, **131**, 3409-3426.

Thépaut, J.-N., 2006: Assimilating only surface pressure observations in 3D and 4DVAR. In: Proceedings of the ECMWF/GEO Workshop on Atmospheric Re-analysis, ECMWF, 19-22 June 2006.

Weissmann,M and C. Cardinali,2006: The impact of airborne Doppler lider observations on ECMWF forecasts. *ECMWF Technical Memorandum No. 505*, November 2006.