QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY
Q. J. R. Meteorol. Soc. **133**: 697–710 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/qj.61

RMetS
ROYAL METEOROLOGICAL SOCIETY

The properties of sensitive area predictions based on the ensemble transform Kalman filter (ETKF)

G. N. Petersen,[a]* S. J. Majumdar[b] and A. J. Thorpe[a]
[a] *University of Reading, UK,*
[b] *University of Miami, US*

ABSTRACT: The spatial characteristics of ensemble transform Kalman filter (ETKF) sensitive area predictions (SAPs) are explored using ensemble forecasts from the European Centre for Medium-Range Weather Forecasts for the period of the 2003 North Atlantic THORPEX Regional Campaign. The ensemble size necessary for a robust sensitive area prediction is found to be surprisingly small: a 10-member ensemble is capable of replicating approximately the same sensitive area structure as a 50-member ensemble. This result is corroborated by the fact that the leading eigenvector of the ensemble perturbations explains over 70% of the ensemble variance and possesses a nearly identical spatial structure regardless of the ensemble size. The structures of the SAPs were found to vary with the lead-time between the ensemble initialization and the adaptive observing time, indicating the necessity of using as recent an ensemble as possible in ensemble-based sensitive area predictions. The ETKF SAPs exhibit similar structures at different levels in the atmosphere and there is no indication of a vertical tilt. A relationship is found between the SAPs and the zonal wind, horizontal temperature gradient and the Eady index, indicating that the ETKF identifies regions with significant gradients in the mass-momentum field as regions of large initial error or large error growth. Copyright © 2007 Royal Meteorological Society

KEY WORDS adaptive observations; forecast uncertainty; THORPEX

Received 1 September 2006; Revised 25 January 2007; Accepted 2 February 2007

1. Introduction

Errors in numerical weather prediction arise due to uncertainties in the model, including those due to uncertainties in the representation of unresolved scales, and in the initial state, namely the analysis. Analysis errors evolve nonlinearly into forecast errors, with errors in unstable regions often growing faster than errors in stable regions. Consequently, large forecast errors can arise due to small analysis errors that grow rapidly during the forecast or large analysis errors that may or may not grow rapidly during the forecast.

In order to produce a model analysis that most accurately describes the observed reality, a data assimilation scheme combines a first guess field, usually a short-range forecast, with observations. Historically, the operational observational network has included routine observations from land-based synoptic stations, radiosondes released at synoptic times and satellite-borne sensors. These observations cannot be varied at will, and have not been designed to be collected in meteorologically important locations. Currently, the routine observational network may be complemented by adaptive observations, including an increased density or frequency of observations already sampled routinely and new observation types, such as reports from commercial aircraft, dropsondes deployed by research aircraft, unmanned aircraft systems and satellite-based observations targeted toward specific regions. The deployment of these adaptive observations in selected locations can be informed by an automated adaptive observing strategy that is based on numerical models. The purpose of this paper is to assess the robustness and structure of locations selected by one such strategy.

The challenge of optimal adaptive sampling of satellite data via data thinning will become even more relevant over the next decade as the quantity of observations will increase by several orders of magnitude. It is also presently thought that additional *in situ* observations over the oceans may provide a worthy complement to the satellite data, particularly in cloudy regions in which low-level satellite data are limited. It has been suggested that such regions, in which baroclinicity is often high, are sensitive areas in which analysis uncertainty may evolve into large short-range forecast errors (Majumdar *et al*., 2002a; Thorpe and Petersen, 2006).

The reduction of analysis and subsequent forecast errors via assimilation of adaptive observations is a primary goal of THORPEX (THORPEX: A global atmospheric programme, is a World Weather Research Programme. For further information see http://www.wmo.int/thorpex.). To reduce forecast errors effectively within some prescribed verification region, usually of societal

* Correspondence to: G. N. Petersen, School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK.
E-mail: g.n.petersen@uea.ac.uk

relevance, at some future verification time, an adaptive observing technique searches for a 'target' region, in which the assimilation of adaptive observations at earlier analysis times will lead to the greatest forecast improvement. The technique should therefore account for the three-dimensional location of the observations, their spatial and temporal density, the types of variables assimilated, their errors and any associated correlations, the operational data assimilation scheme, and the forward evolution of errors from the analysis time(s) to the verification time.

Several mathematical techniques for identifying the locations of sensitive regions have been developed and tested in a hierarchy of models. Such techniques either focus on initial condition uncertainty (Lorenz and Emanuel, 1998) or the sensitivity of a forecast to initial conditions (e.g. Rabier *et al*., 1996; Pu *et al*., 1997; Palmer *et al*., 1998; Bishop and Toth, 1999); while others combine data assimilation with error propagation and growth (e.g. Baker and Daley, 2000; Bishop *et al*., 2001, Gelaro *et al*., 2002; Leutbecher, 2003). Langland (2005) reviews the context in which the various techniques have been applied.

The ensemble transform Kalman filter (ETKF, Bishop *et al*., 2001) combines error covariance information from ensemble forecasts with error statistics associated with the routine and adaptive observational networks to predict the reduction in forecast error variance within a prescribed forecast verification region, due to the assimilation of any given set of adaptive observations. The sensitive area predictions (SAPs) made by the ETKF are commonly summarized in a composite 'summary map' that exhibits the predicted vertically averaged reduction in forecast error variance within the verification region, as a function of the location of the adaptive observations.

Since 2001, the ETKF has been used operationally at the National Centers for Environmental Prediction (NCEP) during annual Winter Storm Reconnaissance (WSR) Programs (Szunyogh *et al*., 2000, 2002), in which the aim is to improve 1–4 day forecasts over the continental United States and Alaska. Using ensembles produced at NCEP and the European Centre for Medium-Range Weather Forecasts (ECMWF), the ETKF has demonstrated the ability to discriminate between good and poor locations for adaptive sampling in spite of its limitations (Majumdar *et al*., 2001, 2002b). The ETKF was also used along with other sensitive area prediction (SAP) methods during the North Atlantic THORPEX field Campaign (A-TReC) that took place between October and December 2003. The primary aim of the campaign was to test real-time quasi-operational adaptive sampling of observations over the North Atlantic. In total 21 cases were targeted, some on multiple days, resulting in 40 targeted events (Mansfield *et al*., 2005). The SAP methods successfully identified sensitive areas but on average the forecast improvement was small (Petersen and Thorpe, 2007; Fourrie *et al*., 2006). This was due to the forecasts in general being quite good during the campaign, which results in less opportunity for improvements when adding observations (Petersen and Thorpe, 2007; Morss and Emanuel, 2002).

The SAPs are the primary product by which adaptive observations are deployed. They possess several limitations, given their theoretical assumptions and the fact that a single summary map is usually supplied as guidance. In order to identify priorities for improving SAPs for future applications, it is necessary to understand the robustness and characteristics of the currently operational SAPs. While a similar study has been performed by Buizza and Montani (1999) using targeted total energy singular vectors (TESVs), no such analysis has taken place using the ETKF. Several properties of the ETKF SAPs require investigation. Firstly, it is well known that only a limited number (currently less than 100) of ensemble forecasts can be produced daily at any operational forecast centre, due to computational restrictions. The extent to which adaptive sampling guidance depends on ensemble size may affect future decisions on ensemble construction, and will give observing programme planners an estimate of the extent to which the reliability of the ETKF may be compromised by ensemble size. It is also well known that a larger ensemble will alleviate spurious long-distance correlations and hence spurious adaptive sampling guidance. Secondly, for practical reasons a lead-time between the initialization of the ensemble and the analysis time, when targeting would occur, is necessary. The lead-time may be as much as 2–3 days, by which time nonlinearity may have compromised the assumptions of the ETKF (Gilmour *et al*., 2001). The effect of this extended lead-time requires investigation in respect of the concern that the assumptions within the ETKF are more severely violated as the lead-time increases. Thirdly, the vertical structure of ETKF sensitivity area predictions has been unexplored to date. Although analysis-time total-energy singular vectors exhibit a highly baroclinic structure tilting against the vertical wind shear (Buizza and Palmer 1995), it is unclear whether the ETKF follows suit. An improved understanding of the vertical structure of the SAP would raise new questions about the current method of targeting a vertically averaged area and identify potentially important locations at specific levels in which observations would help reduce forecast errors. Finally, the synoptic and dynamic relevance of the primary sensitive areas has not been addressed significantly in previous papers. A partial conclusion was offered by Majumdar *et al*. (2002a), who found that the sensitive areas selected by the ETKF and total-energy singular vectors for 1–2 day forecasts of winter weather sometimes corresponded to baroclinic zones. Szunyogh *et al*. (2002) speculated that the sensitive regions selected by the ETKF are likely to be connected dynamically to the verification areas by downstream baroclinic development. However, the relation of the sensitive regions to the properties of the ensemble forecasts that they are based on has not been investigated.

This paper focuses on assessing the robustness and structure of ETKF sensitive area predictions for 40

DOI: 10.1002/qj

independent targeted 30–78 h forecasts of potentially significant weather in verification regions selected during the A-TReC. The following questions are addressed: (1) How large does the ensemble need to be in order for the prediction of sensitive areas to be useful? (2) How robust are the sensitive area calculations with increasing time interval between the initial time of the ensemble forecast and the targeting (analysis) time? (3) How do the sensitive areas vary vertically, and is there a tilt with height? (4) What are the primary physical variables that correspond most closely to the target areas?

Section 2 gives a brief introduction to the ETKF, followed by a description of the issues to be addressed and evaluation metrics in Section 3. The robustness and structure of the ETKF sensitive area predictions are presented and discussed in Section 4, followed by concluding remarks in Section 5.

2. Ensemble Transform Kalman filter (ETKF)

The Kalman filter is a commonly known procedure to estimate the variance of the state of the atmosphere, at a time t, given observations up to and including those at analysis time, t_a, where $t_a \leq t$. This can be done by estimating the error covariance matrix of the atmospheric state. Thus, the Kalman filter could be used to predict the reduction in forecast error variance at a future verification time, t_v, due to the assimilation of all possible combinations of observations prior to and at the analysis time. However, the unfeasibly high computational cost renders it necessary to approximate the Kalman filter. One increasingly popular form is the ensemble Kalman filter, which uses perturbations from an ensemble forecast to approximate the error covariance (Evensen, 2003). The ETKF is one of a family of ensemble Kalman filters whose primary practical application to date has been adaptive sampling. In this section, we describe its application in the context of the A-TReC and this paper. For a detailed theoretical background of the ETKF, the reader is referred to Bishop *et al*. (2001).

In this paper, the ETKF uses output from operational ECMWF ensembles (Buizza *et al*., 2003) initialized at time $t_i (< t_a)$ to predict the reduction in forecast error variance at verification time $t_v (> t_a)$, based on adaptive observations to be assimilated at the analysis time t_a. The resolution of the ensemble output is 2.5°, and the state variables and all observations considered here comprize horizontal winds and temperature at 200, 500 and 850 hPa. The essence of the ETKF is the 'transformation' (via matrix multiplication) of ensemble perturbations, defined by the deviation of each ensemble member from the ensemble mean, into new perturbations based on the configuration and error statistics of the observational network. First, the ETKF is used to transform the original ensemble perturbations, initialized at t_i, such that they represent a theoretically consistent estimate of analysis error covariance based on routine observations collected at time t_a. The routine observations comprize horizontal wind and temperature at all radiosonde locations plus a crude approximation of satellite-based temperature over the oceans at the same atmospheric levels. The new estimate of forecast error variance, assuming that these routine observations have been assimilated at time t_a, is then computed from the transformed ensemble perturbations and stored.

In order to predict the influence of assimilating an additional set of adaptive observations at t_a on forecast error variance at the later verification time t_v, serial assimilation theory (Bishop *et al*., 2001) is used to rapidly transform the ensemble perturbations again, to yield the reduction in forecast error variance (or 'signal variance') due to the addition of these adaptive observations. The signal variance uses a perturbation total energy norm, defined as

$$\sigma^2 = 0.5(u'^2 + v'^2) + \frac{c_p}{T_r} T'^2, \tag{1}$$

where u', v' and T' are ensemble perturbations of the horizontal wind components and the temperature, c_p is the specific heat and $T_r = 300$ K is a reference temperature. (Note that this definition of the total energy norm deviates numerically from the usual definition: $\sigma^2 = 0.5(u'^2 + v'^2 + \frac{c_p}{T_r} T'^2)$. However, Equation (1) is applied in all past ETKF calculations and since the larger emphasis on the temperature perturbations does not change the pattern of the summary map but results in slightly higher signal variance values, it is retained here. Note also that we use the term σ^2 for the signal variance instead of σ as in earlier ETKF papers. This is done to be consistent with the accustomed notation for variance). The procedure to compute signal variance is repeated for several hundred hypothetical 'test-probes' of horizontal wind and temperature, centred at each model grid point (x, y) in the domain where targeting may occur. Each test-probe comprizes 9 adaptive observations in a 3×3 formation, one at each adjacent grid point, 2.5° apart. The additional observations are assumed to crudely represent dropsondes descending through the troposphere, and are thus deemed to be 'vertically aligned'. The SAP represents the averaged forecast signal variance within the verification region, denoted by $\overline{\sigma^2}(x, y)$, which is plotted as a function of the central location of all the test-probes in a composite 'summary map'. The location (x_m, y_m) that maximizes $\overline{\sigma^2}(x, y)$, and hence possesses the maximum likelihood of significantly reducing forecast errors, is deemed optimal for adaptive sampling. The efficiency of the computations enables several forecast scenarios (and summary maps) to be assessed daily. Further details on the construction of summary maps and the routine and adaptive observational networks are provided in Majumdar *et al*. (2002b, 2006). While the ETKF prediction of reduction of forecast error variance is compromised due to non-linearity, inaccuracies and inconsistency in the ETKF and operational data assimilation schemes, as well as differences between the forward model used in the ETKF and in operations, Majumdar *et al*. (2001, 2002b)

DOI: 10.1002/qj

demonstrated that the ETKF is generally able to predict the reduction in NCEP forecast error variance due to the deployment of adaptive dropsonde observations. The present version of the ETKF uses neither horizontal weightings to account for the different sizes of grid boxes, nor vertical weightings to account for the atmospheric density at different levels. Such weightings will be considered in future versions of the ETKF.

Figure 1 shows a typical summary map of $\overline{\sigma^2}(x, y)$, aimed at improving a 30 hour forecast over western Europe. The ETKF predicts that a large reduction in forecast error within the verification region is most likely if adaptive observations were to be collected at 45 °N 70 °W. This would be the primary target area. The collection of supplementary observations in the upper-level jet region upstream of 30 °W is also predicted to produce a likely significant influence on the same forecast. Most of the upper-level jet region is therefore a secondary target area.

3. The experimental framework

The robustness and the structure of the ETKF sensitive area predictions (SAPs) are addressed in the following manner:

1. The largest ensemble can be expected to give the best SAP, due to the superior ETKF estimation of error covariance and reduction of spurious long-distance correlations. Can a smaller ensemble explain the structure of the variance sufficiently to replicate the SAP? In order to evaluate how the SAPs vary with ensemble size, the size of the ensemble applied in the ETKF is varied from 5 to 50 and the summary maps compared.
2. How robust are the predictions with respect to the lead-time, namely the time interval between the ensemble initialization, t_i, and the targeting analysis time, t_a? The lead-time is a necessity for deciding upon a verification time and region, to compute the SAPs, to decide if and where to target, and to issue instructions and prepare observation platforms for targeting. In order to assess the dependence of the SAPs on the lead-time, SAPs with lead-times varying from 30 to 90 hours are compared.
3. How representative is the standard vertically integrated SAP, $\overline{\sigma^2}(x, y)$ for the predicted sensitivity at the different levels in the atmosphere considered? Is there a vertical tilt in the sensitive area predictions? In order to investigate this, two methods are used. Firstly, the signal variance at each pressure level within the verification region, $\sigma_k^2(x, y)$ where $k \in [850, 500, 200]$ hPa, is compared with their vertical average, $\overline{\sigma^2}(x, y)$. Secondly, the SAPs resulting from the assumption that targeted observations are deployed and assimilated on one pressure level instead of all three, $\overline{\sigma^2}^k(x, y)$ where $k \in [850, 500, 200]$ hPa, are compared with $\sigma^2(x, y)$.
4. How are SAPs related to the ensemble-based forecast error variance and do the SAPs correspond to areas of increased baroclinicity or wind shear? The relationship between the SAPs and the meteorological fields near the maximum predicted sensitivity is explored.

Since the primary consideration in adaptive sampling is the region to deploy extra observations, the comparison of the summary maps from different simulations is mainly

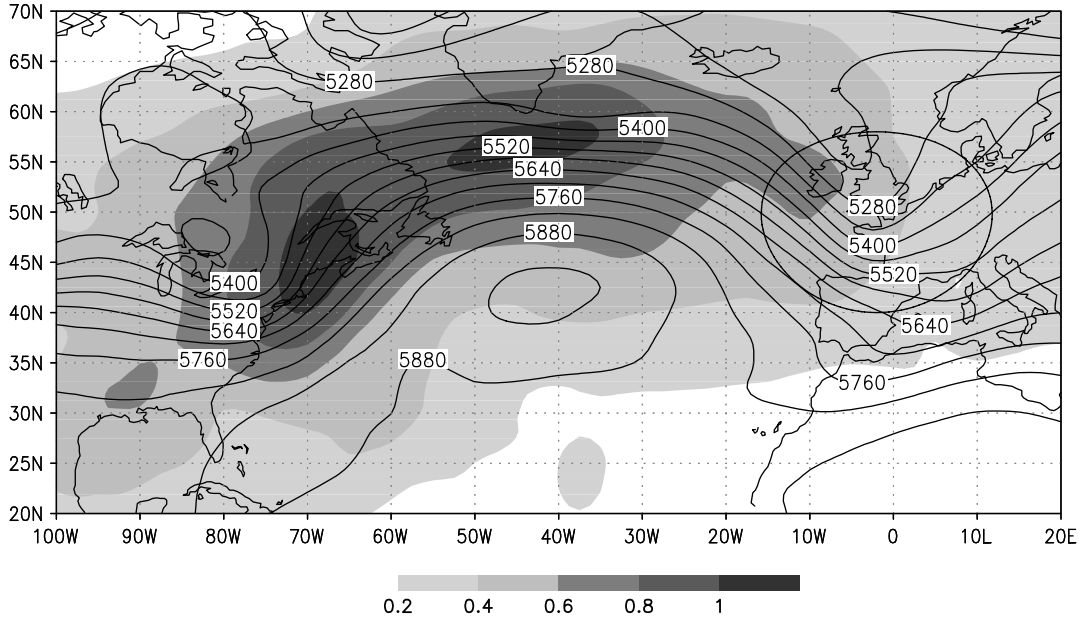

Figure 1. A summary map for A-TReC 7.2 showing the average signal variance of the total energy norm $\overline{\sigma^2}(x, y)$ (J kg^{-1}, shaded) and the control forecast of the 500 hPa geopotential height (m, contours). $t_i = 27/10$ 2003 00 UTC, $t_a = 29/10$ 2003 18 UTC and $t_v = 31/10$ 2003 00 UTC, resulting in a lead-time of 66 hours and an aim of improving a 30 hours forecast. The ellipse represents the verification area. The shading interval is 0.1 J kg^{-1} and the contour interval is 60 m.

DOI: 10.1002/qj

based on geographical overlap and spatial correlation. The geographical overlap is calculated in a similar manner to Leutbecher *et al.* (2004), but here it is based on the number of grid points in the sensitive region. If S_j and S_k are sensitive areas containing the same number of grid points with highest average signal variance, then the geographical overlap, O_{jk}, is defined as the ratio of the number of grid points of their intersection to the number of grid points in question

$$O_{jk} = \frac{\text{grid points}(S_j \cap S_k)}{\text{grid points}(S_j)}. \quad (2)$$

A second metric is defined to compare the magnitudes of the forecast signal variance within the verification region between two summary maps. For example, the relative contributions of this signal variance at the three pressure levels, denoted by $\sigma^2_{200}(x, y)$, $\sigma^2_{500}(x, y)$, and $\underline{\sigma^2_{850}}(x, y)$, to the vertically averaged signal variance $\overline{\sigma^2}(x, y)$ will be evaluated to address question 3 above. If the primary target regions are located similarly in the summary maps it can be sufficient to compare the magnitudes at the location $(x_m, \underline{y_m})$ of the maximum standard average signal variance $\overline{\sigma^2}(x, y)$. If $\sigma^2_j(x_m, y_m)$ and $\sigma^2_k(x_m, y_m)$ are the values of two estimates of signal variance at (x_m, y_m) (e.g. at different levels), the normalized difference between the values can then be defined as

$$\Delta\sigma^2 = \frac{\sigma^2_j(x_m, y_m) - \sigma^2_k(x_m, y_m)}{\sigma^2_k(x_m, y_m)}. \quad (3)$$

As no geographical weighting is included in the ETKF, no weighting is included in the metrics. During the A-TReC field campaign, 40 events were targeted, each event having a unique combination of targeting times, verification times and verification regions. However, since there are events that share a targeting time and therefore the initial time of the ensemble, the number of ensemble forecasts is 38.

The standard ETKF predictions in this paper are computed using the same setup as during the field campaign, with a lead-time $t_a - t_i = 66$ hours and a verification time t_v varying between 30 and 78 hours after targeting, most frequently 42 hours. Twenty-five ECMWF ensemble perturbations, defined as deviations of the first 25 ensemble members from the mean of the whole ECMWF ensemble of 50 members (Buizza *et al.*, 2003), are used, and this configuration is denoted by **epert25**. The horizontal resolution in all calculations is 2.5°.

The hypothetical 'test-probes' of adaptive observations that are used to produce the ETKF summary map extend over the whole North Atlantic region: 20 °N–70 °N, 100 °W–20 °E, see Fig. 1. While the verification region is usually over Europe, a few events have their verification region at the east coast of North America. In order to account for target areas upstream of the verification regions, the calculations for these events are performed for a North America/North Atlantic region, 20 °N–70 °N 160 °W–40 °W.

4. Results

4.1. The effect of the ensemble size

In order to investigate the variation of the SAPs with ensemble size, in addition to **epert25** defined above, four additional sets of ETKF calculations are made for the same A-TReC cases:

epert49: The whole ensemble is applied, resulting in 49 perturbations. (In order to avoid linear dependence between the ensemble perturbations, one perturbation is removed from the ensemble.)
epert24: An ensemble of 25 members is applied, resulting in 24 perturbations.
epert9: An ensemble of 10 members is applied, resulting in 9 perturbations.
epert4: An ensemble of 5 members is applied, resulting in 4 perturbations.

To obtain subsamples from the full ensemble, the first 25, 10 and 5 members of the whole ensemble are selected, and therefore the larger ensembles always contain the smaller ones.

As the **epert49** calculations are run with the largest ensemble, the summary maps resulting from these calculations are assumed to be the best and the results from the other calculations are compared to these. Figure 2 shows box plots of the geographical overlap between **epert49** and each of **epert25**, **epert24**, **epert9** and **epert4**, respectively. Each box plot summarizes the geographical overlap computed using Equation (2) for all events. The median, 25% and 75% quartiles are depicted, together with the smallest and largest outliers.

A comparison of Figs 2(a) and (b) shows that there is little difference between the structure of $\overline{\sigma^2}(x, y)$ based on 25 perturbations from the 50-member ensemble (**epert25**) and that based on 24 perturbations from 25-member ensemble (**epert24**). The discussion of **epert25** will therefore be omitted.

Figure 2(b) shows that in more than 75% of the events, the primary target area for **epert24** is in the same location as that for **epert49**. In fact, with the exception of one event, there is a geographical overlap of 0.8 or more when the 10 or 20 grid points with highest signal variance are considered. Even when secondary target areas are considered by including the 50 best grid points that cover approximately 3×10^6 geographical km^2, there is an 0.9 geographical overlap in more than 75% of the events with a median of 0.97. The only significant outlier is A-TReC 18.1. This is because the primary target area identified in the summary map based on **epert49** is not found in calculations with the smaller ensembles (The **epert49** primary target region is a quasi-stationary upper level cyclone over the western North Atlantic. It is doubtful that the prediction is accurate but rather due

DOI: 10.1002/qj

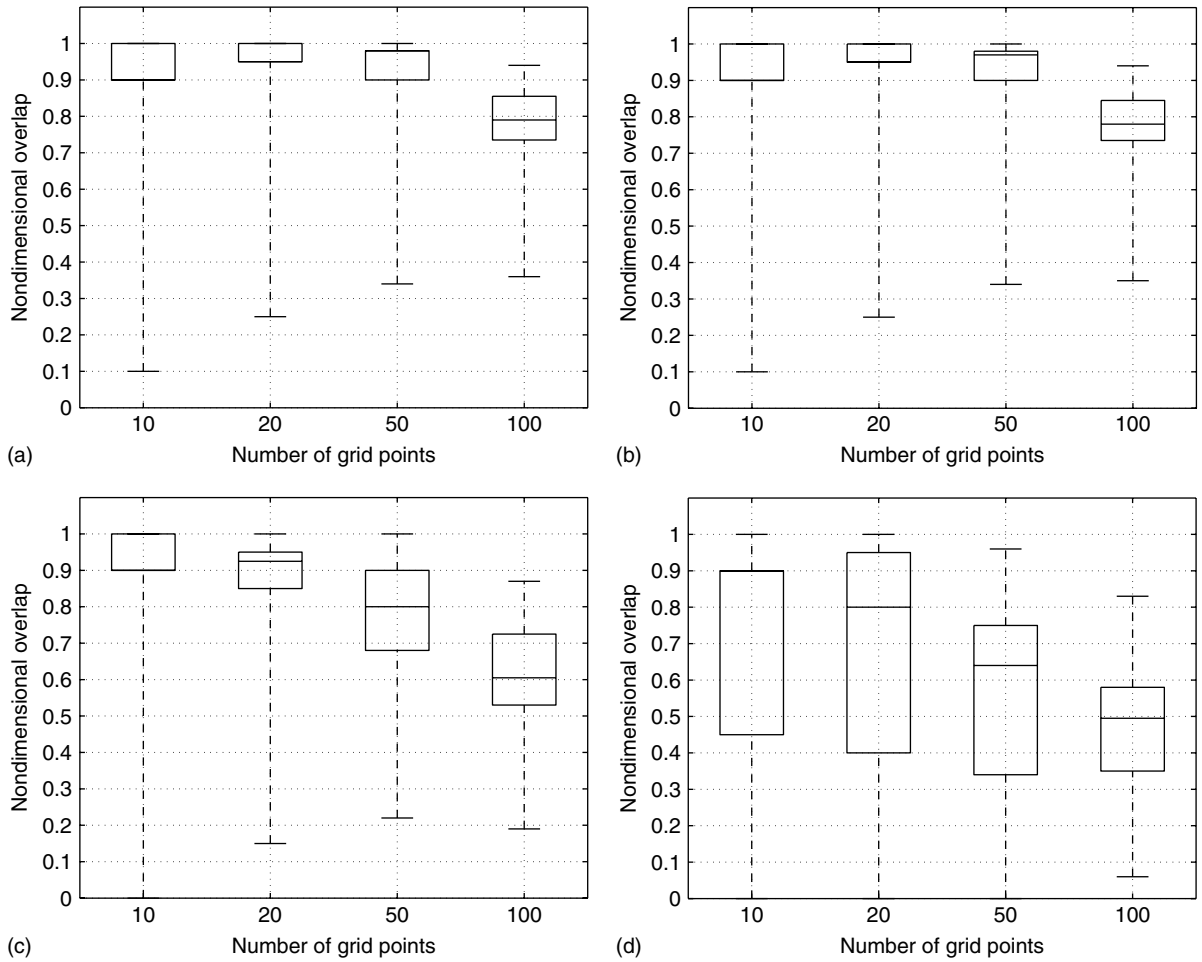

Figure 2. Box plots of the geographical overlap between the SAPs computed using 49 ensemble perturbations and (a) 25 perturbations from an ensemble mean over 50 members, (b) 24 perturbations, (c) 9 perturbations and (d) 4 perturbations. The overlap is shown for the 10, 20, 50 and 100 grid points in which hypothetical observations produce the highest vertically averaged signal variance within the verification region $\overline{\sigma^2}(x, y)$. The box plots show the median, and the 25% and 75% quartiles. The whiskers extend to the outliers farthest away from the median. Each box plot contains 40 data points, one for each targeted event.

to spurious long-distance correlations in the ETKF.). The geographical overlap was also calculated for the last 25 ensemble members of the whole ensemble and similar overlap found with **epert49**.

When the ensemble is reduced to 10 members, the geographical overlap still remains large (Fig. 2(c)). More outliers exist, but the primary target area is still the same in over 75% of the events. For target regions comprising 100 grid points, the overlap exceeds 0.5 in most events.

It isn't until the ensemble size has been reduced down to 5 members that the prediction of the primary sensitive area becomes more crude and the structure is markedly different (Fig. 2(d)). However, even with such a small ensemble the median is still above 0.6 for the primary and secondary target regions, although the quartiles are considerably wider than for the larger ensembles.

An ensemble of 10 members seems to be sufficient to find not only the same primary target area as the full ensemble but also most of the structure of the summary map. To justify the latter statement, Fig. 3(a) shows the total variance for the different ensembles normalized by the total variance of the largest ensemble. The error variance of the two larger ensembles is of comparable size, with **epert24** explaining over 95% of the variance of the entire ensemble **epert49**. As the ensemble size decreases, the total variance drops off, indicating that the smaller ensembles sample less of the error covariance space explained by the full ensemble. Since a 5-member ensemble only captures 58% of the variance of the 50-member ensemble, important information on error structures may be missing.

The eigenspectrum of an ensemble contains information on how many eigenvectors are necessary to replicate the ensemble variance. By computing the eigenvalues of the ensemble covariance matrix $\mathbf{X}\mathbf{X}^T$, in which $\mathbf{X}$ lists the total energy ensemble perturbations at all locations and T denotes the matrix transpose, an eigenspectrum normalized by the total variance of the ensemble can be produced (Fig. 3b). For all ensemble sizes, the leading eigenvector explains at least 70% of the ensemble variance, with the second eigenvector explaining roughly 10% of the ensemble variance.

It is of interest to know if the structures that explain most of the variance are similar, irrespective of the

DOI: 10.1002/qj

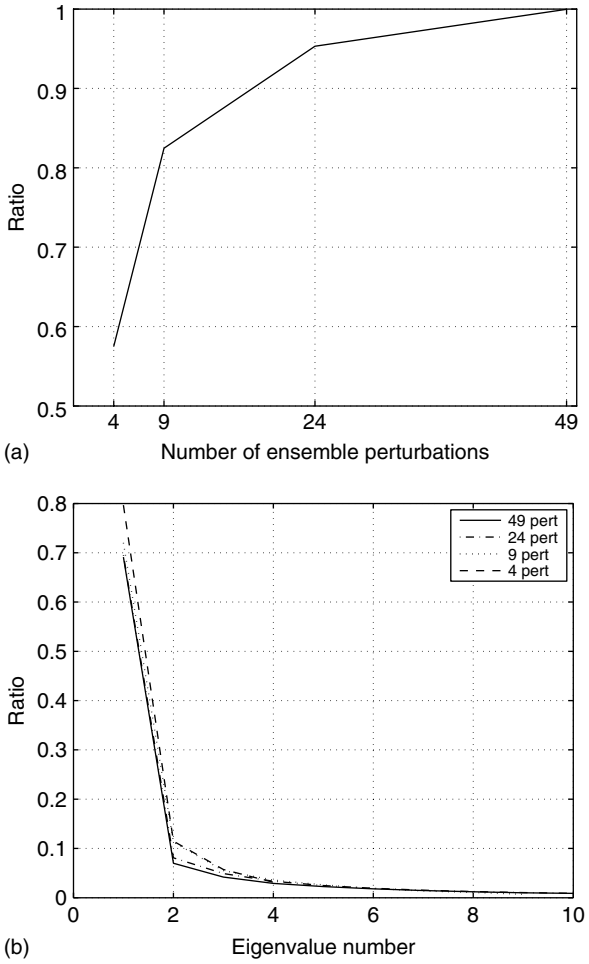

Figure 3. (a) The total variance of the total energy ensemble perturbations and (b) the eigenspectrum of integrated total energy perturbations for the different ensemble sizes. The values in (a) are normalized with the total variance of **epert49** and in (b) with the total variance of each ensemble.

ensemble size. The spatial correlation between each eigenvector $\mathbf{e_j}$ of $\mathbf{X}\mathbf{X}^T$ for ensembles **epert24**, **epert9** and **epert4** and the corresponding j'th eigenvector of the 50-member ensemble is computed for all ensembles used during A-TReC. A two-tailed t-test is used to determine whether the correlation is significant. If the N grid points were all independent, the degrees of freedom would be equal to $N - 2$. However, as there are large regions where there is a dependency between neighbouring grid points, the degrees of freedom are reduced significantly. To take this dependency into account, 16 degrees of freedom are used reducing the degrees of freedom to 1/64 of the total number of grid points (S. Keeley 2006, pers. comm.). Furthermore, it can be argued that the spatial correlation has to be larger than about 0.75 for the two patterns to be regarded as showing the same structures (e.g. Wallace and Gutzler, 1981).

Table I shows the percentage of events that have spatial correlation of over 0.75 that are significant at the 95% level. The table shows clearly that the structure associated with the leading eigenvector is the same irrespective of the ensemble size. In the case of **epert24** the structures

Table I. Hit rates for the first four eigenvectors: The percentage of ensemble forecasts where the total energy structure associated with one of the first 4 eigenvectors has a spatial correlation of over 0.75 at a 95% significance level with the structure associated with the same eigenvector from the 50 member ensemble.

	epert24	**epert9**	**epert4**
Eigenvector 1	100	100	89
Eigenvector 2	61	24	0
Eigenvector 3	42	8	0
Eigenvector 4	42	8	0

associated with eigenvectors 2–4 are often the same as for **epert49**, while the smaller ensembles have only the first structure in common. However, since the leading eigenvector explains almost all of the variance in the ensemble, most of the information about the ensemble spread is contained in the first structure. Thus, even an ensemble of 5 members is in most cases capable of replicating the spatial structure of the ensemble spread although the total variance is much lower, indicating that there is a poorer sampling of the variance space.

There is no systematic order of the ECMWF ensemble members. Thus, the use of the first 5 and 10 members of the whole ensemble in the sub-ensembles should not influence the results. To verify this, the spatial correlation calculations are also carried out for different combinations of 10- and 5- member ensembles. The results are similar with the first structure being replicated in all events for 10 members and most events for 5 members. This indicates that the hit rates in Table I are not dependent on the choice of the ensemble members. The calculations are also conducted for different lead-times, down to 18 hours, with almost identical results. However, the norm chosen might be important, e.g. the eigenspectrum for temperature perturbations does not fall off as rapidly as the total energy eigenspectrum. Approximately the first 10 eigenvectors are required to explain 75% of the temperature variance in the whole ensemble, while the leading 2 eigenvectors are sufficient for the total energy norm. In order to ensure that these results are not specific to the period originally chosen, the eigenspectra for the ensemble forecasts for the same calendar dates but from 2004 are also calculated. The weather situation in the North Atlantic region differed from that during the A-TReC but similar hitrates are found. This suggests that the results are not special to the period although they might be dependent on the perturbation method and the model of the ensemble. A direct counterpart to these results was produced by Buizza and Montani (1999), who demonstrated that only small differences existed in target areas when the number of total-energy singular vectors was increased from 4 to 10.

4.2. The impact of the length of the ensemble forecast

The time needed to make decisions regarding which area, if any, to target as well as to activate the appropriate observational platforms limits how short the forecast

DOI: 10.1002/qj

range can be, i.e. the lead-time between the initialization of the ensemble forecast and the targeting time. Using different lead-times while keeping the time interval between targeting and verification constant requires ensemble forecasts initialized at different times, and thus a different geographical spread and magnitude is expected.

In order to gain some insight into how much the structure of the SAPs can be altered by the use of different lead-times, the ETKF sensitive area predictions for the A-TReC events are recalculated using lead-times in the interval 30 to 90 hours and the summary maps compared. Here, as in the remainder of the paper, the A-TReC ensemble size, **epert25**, is applied. Figure 4 shows the percentage of events that have a geographical overlap of at least 0.7 between forecasts of different lead-times when considering the 50 grid points with highest summary map value. The difference in lead-time is shown on the x-axis while the shorter lead-time itself is shown on the y-axis. Thus for example, the percentage of events with at least 0.7 overlap between the summary map based on 30 hours lead-time and 42 hours (+12 hours) is 87%. The figure shows that there is usually a good agreement between SAPs when the difference in the forecast range is 12 hours. As the forecast range difference increases there is a reduction in the overlap and the likelihood for different structures of the secondary target regions is high when comparing forecasts with a difference in lead-time of more than 36 hours. There is a decrease in geographical overlap for lead-time difference of 12 hours when the forecast range is increased. However, this might be coincidental, as a similar pattern is not seen for larger lead-time differences.

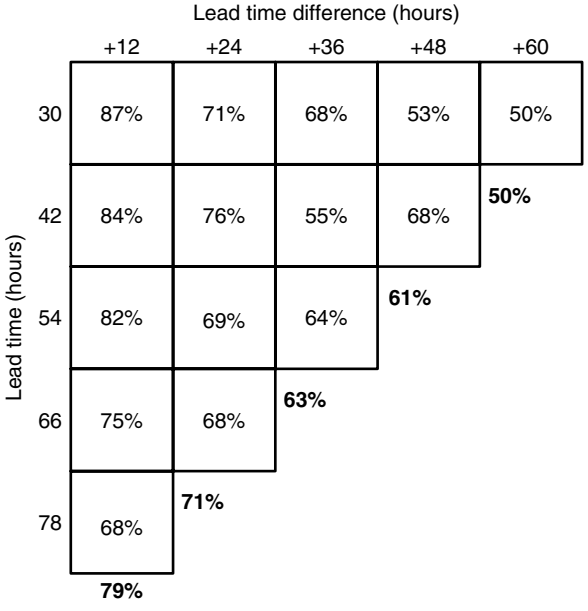

Figure 4. The percentage of the events that have a geographical overlap of at least 0.7 for the 50 grid points with highest summary map value for SAPs based on different lead-times. x-axis: The lead-time difference and y-axis: The shorter lead-time (hours). The values below the columns represent the average overlap for each lead-time difference.

It is important to note that even though the geographical overlap is in general quite high for small changes in the lead-time, this is dependent on the difference between the ensemble forecasts initialized at the different times. If there is a large discrepancy between the geographical distribution of variance in the respective ensembles, the SAPs computed will likely be in different locations.

4.3. The vertical structure of the sensitive areas

The standard ETKF summary map is a vertical average of the total energy signal variance at 850, 500 and 200 hPa within the verification region, produced assuming hypothetical targeted observations at the same three levels, vertically aligned. Thus, no information is given about the vertical structure of the sensitive areas, even though some types of observations, such as AMDAR (Aircraft Meteorological Data Relay) aboard commercial aircraft, are collected at single levels. Here, the vertical structure of the sensitive areas is examined using two different approaches. Firstly, we investigate the geographical similarity between a summary map for which the signal variance within the verification region is computed at one pressure level and the standard summary map. This is done by comparing $\sigma_k^2(x, y)$ for $k \in [850, 500, 200]$ hPa with the vertically integrated signal variance $\overline{\sigma^2}(x, y)$. Secondly, the constraint of assuming vertically aligned targeted observations in the standard summary map is investigated. In other words, would a summary map based on targeting at a single horizontal level be different from the standard summary map? In order to examine this, the quantity $\overline{\sigma^2}^k(x, y)$ based on observing at individual levels $k \in [850, 500, 200]$ hPa is compared with $\overline{\sigma^2}(x, y)$.

The geographical overlap between $\overline{\sigma^2}(x, y)$ and $\sigma_k^2(x, y)$ at each verification level is shown in the box plots in Figure 5, for varying sizes of target regions. The overlap exceeds 0.7 for 75% of the events at all 3 levels, except at 850 hPa for the smallest target area. The maximum SAP at each level is thus in the same location or close to (Within less than about 5° distance.) the primary target area of the standard summary map. In the events where the absolute maxima do not line up vertically, there is still a local maximum in the primary target region. In those instances there are usually two maximum SAP regions at all three levels with the absolute maximum region varying between the two. This is detectable from the box plots, as the geographical overlap increases when the area considered is enlarged and especially the overlap of the outliers increases. This also indicates that there is no systematic tilting with height in the SAPs as the overlap would then remain similar irrespective of the size of the sensitive area considered. The same conclusion is reached by comparing the locations of the maximum sensitivity at the different levels (not shown).

The fractional difference between the summary map maxima with verification norm at each individual level ($\sigma_{200}^2(x, y)$, $\sigma_{500}^2(x, y)$, and $\sigma_{850}^2(x, y)$) and $\overline{\sigma^2}(x, y)$ at the location of maximum averaged sensitivity, (x_m, y_m), is

DOI: 10.1002/qj

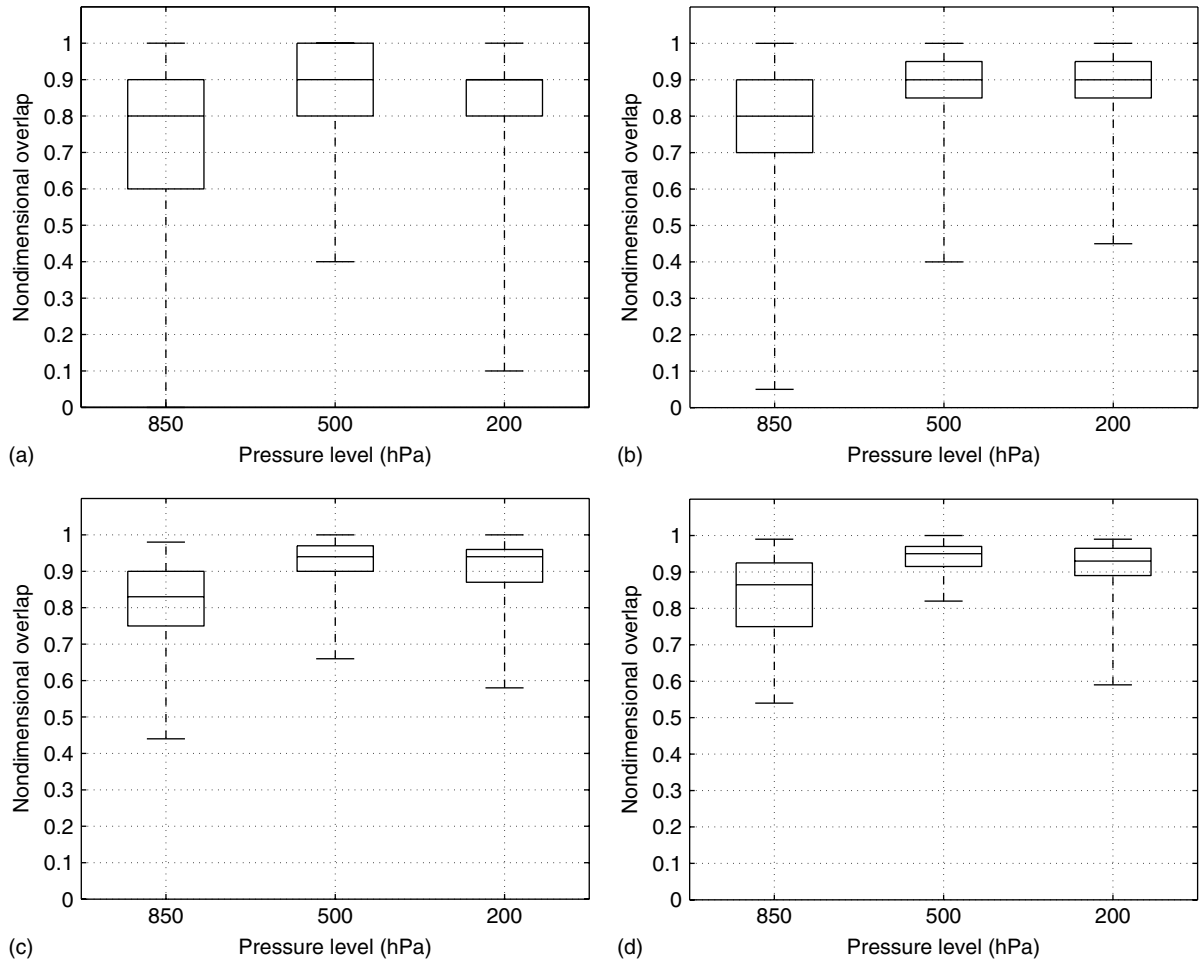

Figure 5. Box plots of the geographical overlap between $\overline{\sigma^2}(x, y)$ of the total energy norm and $\sigma^2_{850}(x, y)$, $\sigma^2_{500}(x, y)$ and $\sigma^2_{200}(x, y)$, respectively. The overlap is shown for (a) 10, (b) 20, (c) 50 and (d) 100 grid points. The box plot shows the 25% and 75% quartile as well as the median. The whiskers extend to the outliers farthest away from the median. Each box plot contains 40 data points, one for each targeted event.

calculated using Equation (3). This fractional difference gives an indication of whether the summary map is dominated by the $\sigma^2(x, y)$ at one level or if all three levels contribute equally. The fractional difference shows that $\sigma^2_{850}(x_m, y_m)$ has the lowest values, about 40% lower than $\overline{\sigma^2}(x, y)$ (not shown). On the other hand, $\sigma^2_{200}(x_m, y_m)$ is the dominant factor in the average, about 40% higher values than in the standard summary map. Similar results are found when the SAPs over a larger area are compared.

Now we examine how the summary map is constrained, due to the standard assumption that the targeted observations are vertically aligned. By computing separate ETKF summary maps for targeted observations to be assimilated at each individual level, the ETKF may identify areas that are sensitive at one level and less sensitive at the other levels. Figure 6 shows the geographical overlap of the standard summary map, assuming vertically aligned observations, and summary maps based on targeting at only one horizontal level, $\overline{\sigma^2}^k(x, y)$ where $k \in [850, 500, 200]$. Only two sizes of SAP regions are shown: 10 and 50 grid points. The location of the primary target region varies from one level to the next, see Fig. 6(a). $\overline{\sigma^2}^{500}(x, y)$ has the greatest overlap with the standard summary map, in over 75% of the events the overlap is 0.5 or more, while there is considerably less frequency of such overlaps for $\overline{\sigma^2}^{850}(x, y)$ and $\overline{\sigma^2}^{200}(x, y)$. However, when considering a larger area of maximum SAP, there is less difference between the levels. The overlap in about 75% of the events is over 0.5 for all levels, see Fig. 6(b). This is because even though the location of the absolute maximum varies, it is not shifted much and the horizontal structure of the SAP is often similar. There is no systematic tilting with height (not shown).

The results of this section indicate that the summary map represents the ETKF sensitive area predictions adequately. There is not a significant loss of information when the vertical integration is done to produce $\overline{\sigma^2}(x, y)$ presented in the summary map. As there is no evidence of tilting of the SAPs with height and the spatial structure is similar at all three levels, the fact that the magnitude of $\sigma^2(x, y)$ increases with height has limited influence on the results presented in the summary map. Furthermore, assuming only horizontal observational deployment results in some differences in the location of the

DOI: 10.1002/qj

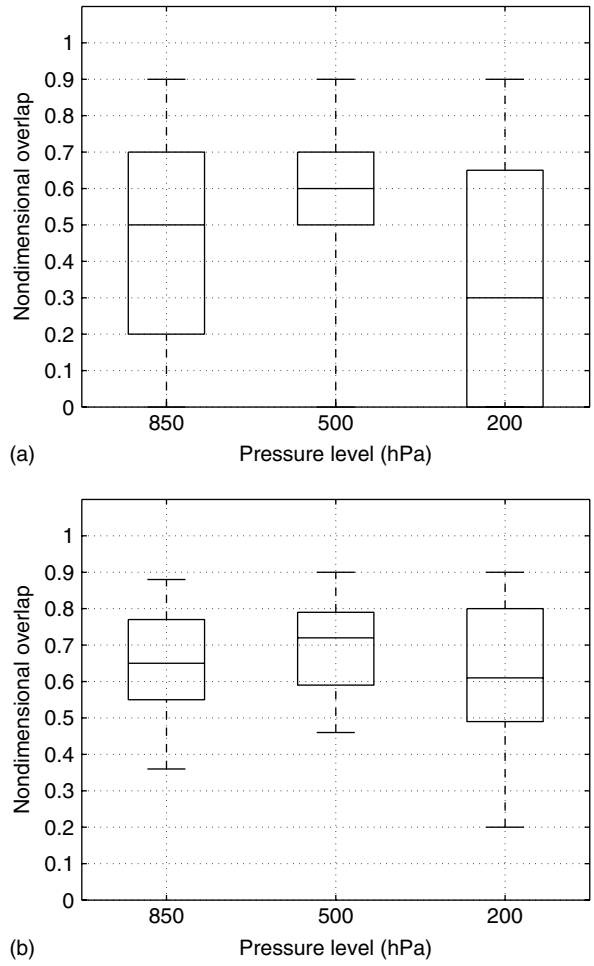

Figure 6. Box plots of the geographical overlap between $\overline{\sigma^2}(x, y)$ and $\overline{\sigma^2}^k(x, y)$ where targeting occurs only at level k, $k \in [850, 500, 200]$ hPa. The overlap is shown for (a) 10 and (b) 50 grid points. The box plot shows the 25% and 75% quartiles and the median. The whiskers extend to the outliers farthest away from the median. Each box plot contains 40 data points, one for each targeted event.

maximum values but little changes in the overall structure of the summary map. This means that since there is little vertical structure in the SAPs, the assumption of vertically aligned observations has little impact on the summary map.

4.4. The relationship between the sensitive area and the meteorological fields

In this section, the relationship between the meteorological fields and the sensitive areas is explored. The data from the ECMWF ensemble that the ETKF calculations are based on are, as mentioned earlier, u, v and T at three levels in the atmosphere. Figure 7 shows an example of a sensitive area prediction overlaid with corresponding ensemble forecasts of different meteorological fields. There are qualitative similarities between the SAP and some of the meteorological fields: the ensemble mean zonal wind, vorticity and the Eady index as well as the horizontal gradient of ensemble mean temperature. The ensemble perturbation total energy norm is a measure of the forecast uncertainty at targeting time and has a pattern quite similar to the SAP. The Eady index is defined as

$$\sigma_E = 0.31 f \left| \frac{d\mathbf{v}}{dz} \right| N^{-1}, \tag{4}$$

where f is the Coriolis parameter, $\mathbf{v}$ the horizontal wind vector and N the Brunt–Väisälä frequency. The Eady index is an expression of the growth rate of the most unstable Eady mode (Hoskins and Valdes, 1990) and thus a measure of the baroclinic instability. Buizza and Palmer (1995) showed that the Eady index is a qualitative diagnosis of the principal regions where singular vectors are located. The Eady index is here calculated for the 500 hPa level, using the 850 and 200 hPa wind and temperature to estimate the static stability and vertical wind shear. This crude estimate of the Eady index is based only on information available to the ETKF for the sensitive area predictions.

The spatial correlation between the SAPs and a meteorological field is calculated in an area surrounding the maximum ETKF sensitivity. To estimate the relationship in a dynamically meaningful area the area is bounded by $+/-30^\circ$ latitude and longitude. A two-tailed t-test shows that for 10 degrees of freedom (corresponding to 1/64th of the total number of grid points in the bounded region), the correlation has to be larger than 0.58 in order to be significant on a 95% level. The spatial correlation emphasizes what was suggested by Fig. 7. There is some significant correlation between the SAPs and 500 hPa ensemble mean zonal wind and horizontal gradient of ensemble mean temperature as well as the Eady index, see Table II. The spatial correlation is significant in 40%, 37.5% and 40% of the events, respectively, while for other variables (such as vorticity) there is much less frequent significant correlation. It is noteworthy that 37.5% of the events had a significant correlation with two or three of the above mentioned fields. The ETKF sensitive areas are thus often areas of predicted strong upper tropospheric zonal wind, i.e. the jet stream, and baroclinicity at low- and mid-atmospheric levels. There is a significant correlation between the ensemble spread and the SAPs in roughly half or more of the events, see Table III. This is to be expected as the spread indicates the uncertainty in the ensemble forecast and such high uncertainty at the analysis time may lead to high uncertainty at the verification time. Hence, the assimilation of observations in such regions would be expected to modify the forecast uncertainty significantly. For a larger geographical area the spatial correlation might not be as high because the ensemble spread may be large in other regions that are not dynamically relevant to the verification region. The ensemble spread thus identifies the uncertainty in the prediction of temperature in highly baroclinic regions as well as regions of high zonal wind. When the situation is such that additional observations in these regions can affect a later forecast for a verification area, the ETKF identifies these regions as sensitive. The ETKF also identifies locations of relatively low ensemble spread, from which the spread amplifies into the verification region.

DOI: 10.1002/qj

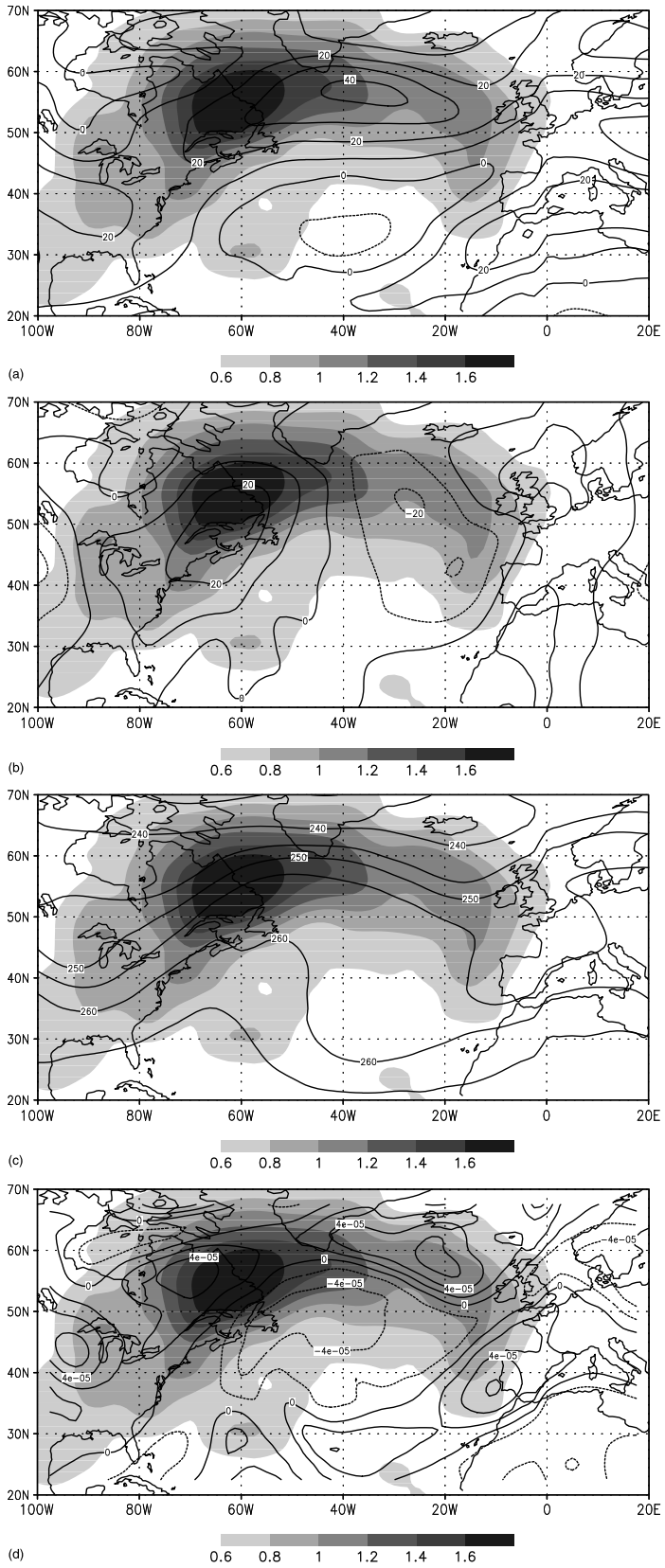

Figure 7. A-TReC 7.1, t_i = 26/10 2003 00 UTC, t_a = 28/10 2003 18 UTC and t_v = 31/10 2003 00 UTC. The sensitive area prediction is shaded. In contours, ensemble mean (a) zonal wind component (m s^{-1}), (b) meridional wind component (m s^{-1}), (c) temperature (K) and (d) vorticity (s^{-1}), horizontal gradient of ensemble mean (e) temperature (K m^{-1}) and (f) vorticity (m^{-1} s^{-1}), (g) ensemble mean Eady index (day^{-1}) and (h) ensemble perturbation total energy norm (J kg^{-1}).

DOI: 10.1002/qj

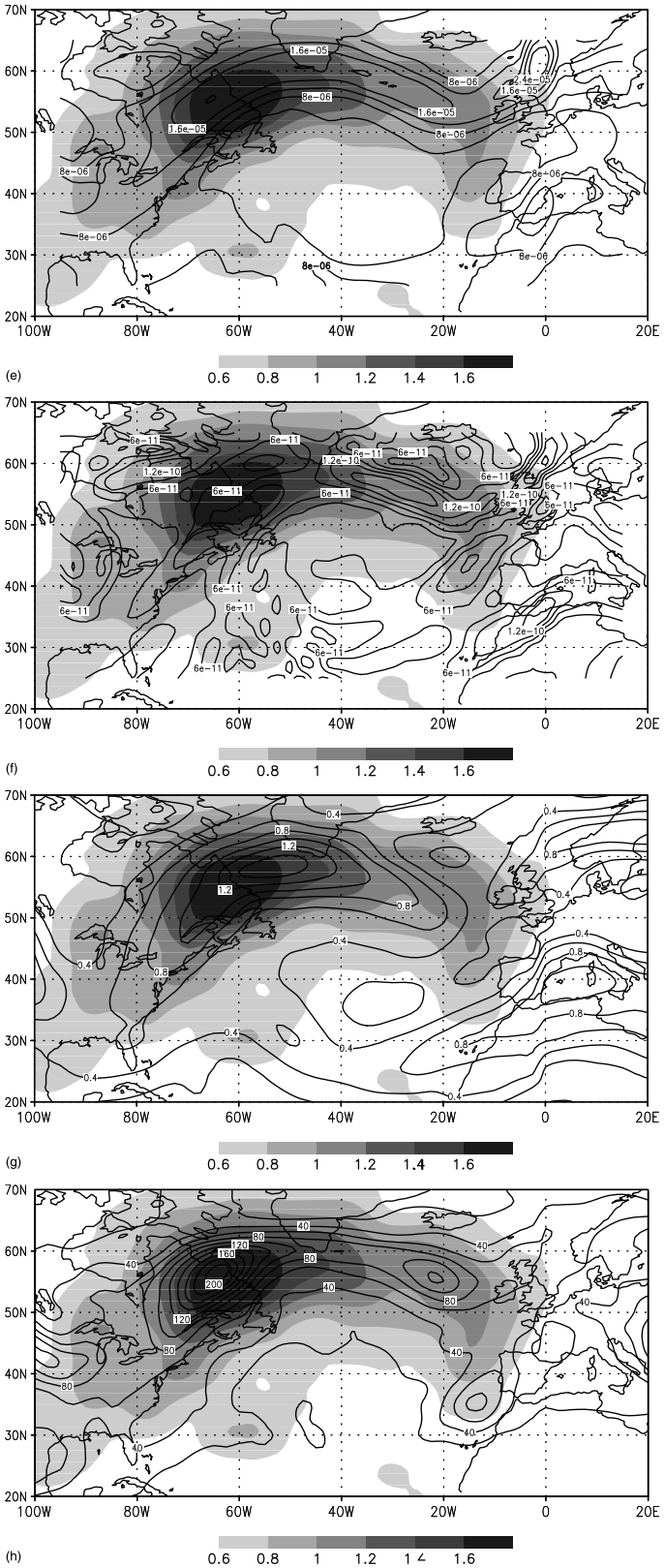

Figure 7. (*Continued*).

5. Conclusions and Discussion

The purpose of this paper is to investigate the characteristics and quantify the robustness of sensitive area predictions based on the ETKF using the ECMWF ensemble and a total energy norm.

The primary and secondary target regions found by applying the ETKF to an ensemble of 10 members are comparable to those based on 50 ensemble members. This is due to the 10 ensemble members sampling over 80% of the phase space of the larger ensemble. The leading

DOI: 10.1002/qj

Table II. The percentage of events, out of the total 40.0 events, with spatial correlation of 0.58 or more between the SAP and the meteorological field in question.

Variable	850 hPa	500 hPa	200 hPa
Ensemble mean			
u	27.5	40.0	25.0
v	0.0	10.0	10.0
T	15.0	2.5	5.0
Vorticity	2.5	2.5	2.5
Eady index		40	
Horizontal gradients of ensemble mean			
u	0.0	5.0	0.0
v	5.0	15.0	2.5
T	27.5	37.5	10.0
Vorticity	0.0	2.5	0.0

Table III. The percentage of events, out of the total 40 events, with spatial correlation of 0.58 or more between the SAP and the ensemble spread.

Variable	850 hPa	500 hPa	200 hPa
U	50.0	52.5	42.5
V	60.0	72.5	62.5
T	52.5	47.5	52.5
Total energy	82.5	85.0	87.5

eigenvector of the covariance matrix of the ensemble perturbations explains about 90% of the variance in the ensemble, and its associated structure is nearly identical for ensembles with more than 5 members. These results are independent of the period studied and the forecast range, although they may be dependent on the type of ensemble and the norm chosen. The findings here are also consistent with those of Buizza and Palmer (1998), who showed that a small ensemble can sample the error variance sufficiently during the first 5 days of the forecast, as long as the initial perturbations are based on sufficiently many singular vectors. An analogy can also be drawn between our results and those of Buizza and Montani (1999), who demonstrated that the sensitive area prediction from 4 TESVs is similar to the corresponding prediction using 10 TESVs. One may therefore conclude that only minor modifications to the synoptic-scale structure of the SAPs are achieved when extra ensemble members or targeted singular vectors are added. However, further research is required to investigate whether this conclusion holds when additional models are introduced, and different scales (such as the mesoscale) are investigated.

The lead-time of the ensemble influences the structure of the sensitive area predictions. A different lead-time implies a different ensemble and hence a different structure of the geographical distribution of the variance. This can potentially lead to different sensitive area predictions. The drop in overlap of the sensitive area predictions from about 80% to 60% when the lead-time difference increases from 12 to 36 hours indicates the necessity of using as recent an ensemble as possible when deciding on a target region. Equivalently, the optimality of the targeted deployment of resources that require a long lead-time (such as manned aircraft) is expected to be compromised significantly.

The predictions of the location of the sensitive areas do not vary much with height, when either the target or the verification levels are considered individually. The structures are generally stacked vertically with little or no tilting. The strength of the local maxima may vary between levels, causing the location of the maximum SAP to vary. The location of the primary target region when considering one targeting level is often slightly different from the one in the standard summary map. However, when the area of interest is enlarged the geographical overlap increases indicating that, even though the primary target region might differ, the secondary is similar. Thus, the overall structure is similar. Therefore, little information is lost by either vertically integrating the signal variance or assuming vertically aligned targeted observations and the summary map is a good representation of the ETKF sensitive area prediction.

The ETKF sensitive areas are often found to be in regions in which strong zonal winds are expected, particularly at mid-tropospheric levels in areas near the jet stream. Furthermore, the sensitive regions are also associated with areas of large mid-tropospheric baroclinic instability, through the Eady index, and low- to mid-level baroclinic regions due to the frequent spatial correlation with the horizontal temperature gradient. This is not entirely consistent with the common view that the ETKF targets are often found in the upper troposphere and in more developed stages of baroclinic waves, while the TESV targets are often in the middle and lower troposphere and in early stages of baroclinic wave development (Langland, 2005).

The ETKF target areas do not exhibit a tilt in the vertical, unlike total-energy singular vectors (Buizza and Palmer, 1995). To interpret this, it is important to note that the method used to construct the TESV SAP (a weighted average of the leading TESVs) is very different from that used to construct the ETKF SAP. The ETKF is based on the observation type, the data assimilation scheme used to assimilate the observations, and the propagation of error variance from the observing time to the verification time. Because of this, there are several reasons why the ETKF SAP may look very different from the tilted structure of a typical growing baroclinic wave. Firstly, the resolution of the ETKF SAP is coarse, since the 'test-probe' in this study has a 5° latitudinal and longitudinal extent. Secondly, the assimilation of an observation using an ensemble Kalman filter is able to influence a state estimate at a remote location, if a significant covariance exists between errors in the relevant variables at the two locations. A large error covariance is expected to exist between variables at all levels in the broad vicinity of the baroclinic zone, given that the amplitude and phase of the low-level

DOI: 10.1002/qj

baroclinic structure and the associated flow aloft is likely to be vastly different in all ensemble members at a +66 hour lead-time. Hence, the distribution of analysis error variance, and therefore reduction of analysis error variance due to the targeted observation, is likely to be rather broad. Thirdly, the propagation of this reduction of error variance into the verification region is likely to be similar for adjacent observation locations, given the breadth of the error variance distribution. Thus, the ETKF SAP is likely to be much broader and more diffuse than the TESV SAP, and hence less obviously tilted.

The main conclusions of the study are that the ETKF sensitive area predictions have little vertical structure on the synoptic scale, they can vary significantly according to the lead-time, and that they are robust to changes in the ensemble size beyond 10 members. However, it requires emphasis that the computations were performed using an ensemble from a single model. The conclusions are likely to be modified, and the reliability of the ETKF targets improved, as ensembles from the THORPEX Interactive Grand Global Ensemble are included. For example, the spectrum of error variance is likely to become significantly broader. Other benefits from the larger, multi-model ensemble include the reduction of spurious remote targets, and a more accurate quantitative prediction of analysis and forecast error variance based on different configurations of observations. Other modifications to the ETKF include the use of higher-resolution ensemble outputs and a shorter lead-time (if practical circumstances permit for certain applications), which are expected to lead to finer-scale, more focused targets that may exhibit a richer vertical structure than found in this paper. Other issues that require exploring are the extent to which non-linearity compromises the performance of the ETKF, and the sensitivity to observing and verification of particular variables.

Acknowledgements

The authors would like to thank David S. Richardson for valuable discussions and advice. G. N. Petersen was funded by Vaisala and the UK Met Office through a THORPEX Postdoctoral Research Fellowship. S. J. Majumdar gratefully acknowledges funding from NOAA THORPEX.

References

Baker NL, Daley R. 2000. Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.* **126**: 1431–1454.

Bishop CH, Toth Z. 1999. Ensemble transformation and adaptive observations. *J. Atmos. Sci.* **56**: 1748–1765.

Bishop CH, Etherton BJ, Majumdar SJ. 2001. Adaptive sampling with the ensemble Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **129**: 420–436.

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

Buizza R, Palmer TN. 1995. The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.* **56**: 1434–1456.

Buizza R, Palmer TN. 1998. Impact of ensemble size on ensemble prediction. *Mon. Weather Rev.* **126**: 2503–2517.

Buizza R, Richardson DS, Palmer TN. 2003. Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensemble. *Q. J. R. Meteorol. Soc.* **129**: 1269–1288.

Evensen G. 2003. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics* **53**: 343–367.

Fourrié N, Marchal D, Rabier F, Chapnik B, Desroziers G. 2006. Impact study of the 2003 North Atlantic THORPEX Regional Campaign. *Q. J. R. Meteorol. Soc.* **132**: 275–295.

Gelaro R, Rosmond T, Daley R. 2002. Singular Vector Calculations with an Analysis Error Variance Metric. *Mon. Weather Rev.* **130**: 1166–1186.

Gilmour I, Smith LA, Buizza R. 2001. Linear regime duration: Is 24 hours a long time in synoptic weather forecasting? *J. Atmos. Sci.* **58**: 3525–3539.

Hoskins BJ, Valdes PJ. 1990. On the existence of storm tracks. *J. Atmos. Sci.* **47**: 1854–1864.

Langland RH. 2005. Issues in targeting observing. *Q. J. R. Meteorol. Soc.* **131**: 3409–3425.

Leutbecher M. 2003. A reduced rank estimate of forecast error variance changes due to intermittent modifications of the observing network. *J. Atmos. Sci.* **60**: 729–742.

Leutbecher M, Anderson E, Beljaars A, Hortal M, Janssen P. 2004. Planning of adaptive observations during the Atlantic THORPEX regional Campaign 2003. *ECMWF Newsletter* **102**: 16–25. Available from www.ecmwf.int.

Lorenz EN, Emanuel KA. 1998. Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.* **55**: 399–414.

Majumdar SJ, Aberson SD, Bishop CH, Buizza R, Peng MS, Reynolds CA. 2006. A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Weather Rev.* **134**: 2354–2372.

Majumdar SJ, Bishop CH, Etherton BJ, Szunyogh I, Toth Z. 2001. Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations? *Q. J. R. Meteorol. Soc.* **127**: 2803–2820.

Majumdar SJ, Bishop CH, Buizza R, Gelaro R. 2002a.. A comparison of Ensemble Transform Kalman Filter targeting guidance with ECMWF and NRL total energy singular vector guidance. *Q. J. R. Meteorol. Soc.* **128**: 2527–2549.

Majumdar SJ, Bishop CH, Etherton BJ, Toth Z. 2002b.. Adaptive sampling with the ensemble Kalman filter. Part II: Field program implementation. *Mon. Weather Rev.* **130**: 1356–1369.

Mansfield D, Richardson D, Truscott B. 2005. An overview of the Atlantic THORPEX Regional campaign. Pp 91–94. in Proceedings of the First THORPEX international science symposium, 6–10 December 2004, Montreal, Canada. Available from the WMO secretariat.

Morss RE, Emanuel KA. 2002. Influence of added observations on analysis and forecast error: Results from idealized systems. *Mon. Weather Rev.* **128**: 285–321.

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

Petersen GN, Thorpe AJ. 2007. The impact on weather forecasts of targeted observations during A-TReC. *Q. J. R. Meteorol. Soc.* in press.

Pu ZX, Kalnay E, Sela J, Szunyogh I. 1997. Sensitivity of forecast error to initial conditions with a quasi-inverse linear method. *Mon. Weather Rev.* **125**: 2479–2503.

Rabier F, Klinker E, Courtier P, Hollingsworth A. 1996. Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.* **122**: 121–150.

Szunyogh I, Toth Z, Morss RE, Majumdar SJ, Etherton BJ, Bishop CH. 2000. The effect of targeted dropsonde observations during the 1999 Winter Storm Reconnaissance program. *Mon. Weather Rev.* **128**: 3520–3537.

Szunyogh I, Toth Z, Zimin AV, Majumdar SJ, Persson A. 2002. Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Mon. Weather Rev.* **130**: 1144–1165.

Thorpe AJ, Petersen GN. 2006. 'Predictability and targeted observations'. Pp 561–583 in *Predictability of Weather and Climate*. Ed. Palmer T, Hagedorn R (eds). Cambridge University Press: Cambridge, UK.

Wallace JM, Gutzler DS. 1981. Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Weather Rev.* **109**: 784–812.

DOI: 10.1002/qj