Tellus (2006), 58A, 523–537
Printed in Singapore. All rights reserved

TELLUS

A methodology for fixed observational network design: theory and application to a simulated global prediction system

By S. P. KHARE* and J. L. ANDERSON, *National Center for Atmospheric Research, Boulder, CO, 80307-3000, USA*

(Manuscript received 13 September 2005; in final form 30 April 2006)

ABSTRACT

A methodology called the retrospective design algorithm (RDA) for fixed observational network design is presented. The RDA makes use of archived ensemble forecasts to estimate the value of various spatial configurations of fixed observational networks. For large dimensional geophysical applications of interest, the RDA can test a variety of fixed networks without having to repeatedly integrate the prediction model equations, unlike observing system simulation experiments (OSSEs). The RDA does not require the adjoint of the prediction model or observation operator. As a proof of concept, the utility of the RDA is tested by making direct comparisons to OSSEs for a simple surface pressure network design problem in a simulated global atmospheric prediction system. A high degree of correspondence between information obtained from OSSEs and the RDA was found for a variety of verification regions and forecast lead times.

1. Introduction

A key problem in atmospheric prediction is the optimization of finite observational resources. This problem, called observing network design, is a major focus of THORPEX (THORPEX International Science Plan, Version III, Shapiro and Thorpe, 2004; see web site www.mmm.ucar.edu/uswrp/programs/thorpex.html). One important design problem is to find the optimal fixed spatial configuration of a suite of observing instruments that minimizes time-mean forecast errors. Two issues must be addressed. The first is a methodology for quantifying the impact of a particular fixed spatial configuration on forecast errors. The second is an optimization algorithm capable of searching for the optimal configuration. The impact of a particular spatial configuration of instruments on operational forecast errors could be determined via field experiments. Given the practical difficulties of field experiments, the importance of optimizing networks of observations using simulation techniques has long been recognized (Emanuel et al., 1995, 1997; Dabbert et al., 1996). The use of simulation techniques to quantify the impact of a particular fixed spatial configuration on forecast errors is the central subject of this paper.

Observing system simulation experiments (OSSEs) can be used to quantify the value of an observing network [Bishop et al., 2003 (hereafter B03); Arnold and Dey, 1986; Daley, 1991]. The number of different observing networks that can be tested is limited by computational expense (B03). Consider the problem depicted in Fig. 1 where one is trying to optimize the location of one fixed observation. For each location on the horizontal axis of Fig. 1, an OSSE can be used to compute the value of a cost function Φ_{OSSE}. If one's initial guess is far from optimal, an optimization algorithm may require many evaluations of Φ_{OSSE} to find the minimum. The key practical question is whether or not one can find 'suitable' approximations to the OSSE cost functions. A 'suitable' approximate cost function would be one that mimics the shape of the Φ_{OSSE} and is cheap to evaluate. Mimicking the shape of Φ_{OSSE} is desirable since it would allow one to determine the approximate location of the minimum *and* the relative value of placing observations in different locations.

A method for observing network design is derived and demonstrated in B03. B03 derive an expression for the infinite time covariance of a linear, steady dynamical system under the influence of a fixed observing network. This method can be used to design fixed observing networks if a suitable, linear steady basic state of the prediction model can be found. B03 apply the method to a global quasi-geostrophic model (Marshall and Molteni, 1993). It may not be possible to find relevant, linear, steady basic states of the prediction model (Anderson, 1992). As a result, additional frameworks for designing networks of routine observations are required. In this paper, a new methodology for efficiently approximating information derived from doing OSSEs is developed and demonstrated. The new methodology,

*Corresponding author.
e-mail: khare@ucar.edu
DOI: 10.1111/j.1600-0870.2006.00200.x

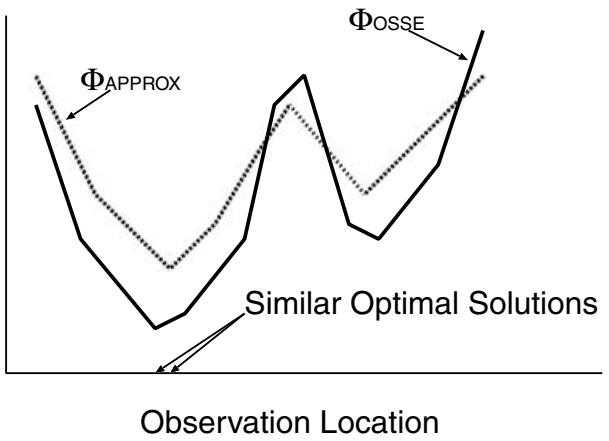

Fig. 1. Cost functions for a one-dimensional design problem. The horizontal axis corresponds to the location of a single observation. The goal is to find the optimal location on the horizontal of this single observation. The vertical axis corresponds to the cost function values obtained by running OSSEs (Φ_{OSSE}) (depicted by the solid line) and an approximate (Φ_{APPROX}) (depicted by the dashed line). In this contrived example, the Φ_{APPROX} cost function does a reasonable job of mimicking the shape of Φ_{OSSE}. The hope is that Φ_{APPROX} is much cheaper to evaluate than Φ_{OSSE}.

called the retrospective design algorithm (RDA), does not require finding a relevant linear steady basic state of the prediction model nor does it require development of the prediction model or observation operator linearizations. The RDA makes use of ensemble forecasts generated from an ensemble data assimilation system and does not require repeated integrations of the prediction model in testing various trial configurations of fixed observing networks.

The statistical and dynamical significance of the RDA is explored using perfect model experiments in a simulated global atmospheric prediction system. The RDA is tested by making *direct* comparisons to OSSEs. Results are shown for the problem of finding the optimal fixed location of a surface pressure observation in the mid-latitudes for a variety of verification regions and forecast lead times. This experimental design allows for clean interpretation of the impact of the dynamics. The results provide insight that will be crucial in applying the RDA to more general design problems in future work.

In section 2, a definition of the specific problem examined in this paper is provided. In section 3, the underlying theory behind the RDA is developed. Issues, such as computational scaling, are also discussed. In section 4, numerical experiments are presented and discussed. Section 5 provides a discussion and conclusions.

2. A definition of the problem of fixed observational network design

For simplicity and clarity, a particularly simple version of the fixed network design problem has been chosen as the focus of this paper. Let the true system state and forecast model dynamics be given by a time-discretized version of $d\mathbf{x}_t/dt = M(\mathbf{x}_t, t)$ (deterministic dynamics). The true state $\mathbf{x}_t$ is an n-dimensional vector. Let the current day, fixed observing network be summarized by a possibly nonlinear operator H^{current} which maps $\mathbf{x}_t$ onto the expected observations (assuming unbiased observation errors). The observation errors at a particular time t, $\varepsilon_t^{\mathrm{current}}$, generally include both representation and instrument errors (Cohn 1997). Observation values associated with the current day fixed network at time t are given by a B-dimensional vector $\mathbf{y}_t^{\mathrm{current}} = H^{\mathrm{current}}(\mathbf{x}_t) + \varepsilon_t^{\mathrm{current}}$. Assume that the statistical characteristics of $\varepsilon_t^{\mathrm{current}}$ (such as the error covariance) are known. Assume that observations are taken periodically. If H^{current} is linear, it is a $B \times n$ matrix.

Suppose that S observations will be added to H^{current}. The problem considered in this paper is to find the optimal fixed locations for the S additional observations. Let the index q denote the q^{th} spatial configuration of the additional observation locations. Let $H^{\mathrm{add},q}$ denote the (possibly nonlinear) observation operator associated with the q^{th} configuration of the additional observations. The observation values associated with a particular configuration of the additional observing network are denoted by the S-dimensional vector $\mathbf{y}^{\mathrm{add},q}{}_t = H^{\mathrm{add},q}(\mathbf{x}_t) + \varepsilon_t^{\mathrm{add}}$. The statistical characteristics of the S-dimensional observation errors $\varepsilon_t^{\mathrm{add}}$ (such as error covariance) are known and the same for every configuration q. If $H^{\mathrm{add},q}$ is linear, it is a $S \times n$ matrix.

A particular trial fixed network is defined as the combination of H^{current} and $H^{\mathrm{add},q}$, denoted by $H^{trial,q} = [H^{\mathrm{current}}; H^{\mathrm{add},q}]$. If both H^{current} and $H^{\mathrm{add},q}$ are linear, then $H^{trial,q}$ is a $(B + S) \times n$ matrix. Observation values associated with a particular configuration of the trial network are denoted by a $(B + S)$-dimensional vector $\mathbf{y}_t^{\mathrm{trial},q} = H^{\mathrm{trial},q}(\mathbf{x}_t) + \varepsilon_t^{\mathrm{trial}}$, with $\varepsilon_t^{\mathrm{trial}} = [\varepsilon_t^{\mathrm{current}}; \varepsilon_t^{\mathrm{add}}]$ and $\mathbf{y}_t^{trial,q} = [\mathbf{y}_t^{\mathrm{current}}; \mathbf{y}_t^{\mathrm{add},q}]$.

For every spatial configuration of $H^{\mathrm{trial},q}$, there is a value of an objective function Φ ($H^{\mathrm{trial},q}$). Φ ($H^{\mathrm{trial},q}$) is a function related to the uncertainty associated with estimating the system state. *The goal of the fixed network design problem* is to determine the configuration of $H^{\mathrm{trial},q}$ which minimizes the objective function $\Phi(H^{\mathrm{trial},q})$.

3. The retrospective design algorithm

3.1. Choices made when developing the retrospective design algorithm

Our goal has been to develop a computationally efficient algorithm. To help achieve our goal of computational efficiency, we have sought to develop an algorithm which does not require repeated integrations of the forecast model when evaluating $\Phi(H^{\mathrm{trial},q})$. We have sought to develop an algorithm which does not have to *repeatedly* account for the B observations in $H^{\mathrm{trial},q}$ which are not being optimized. So that the method can be readily applied to a variety of systems, we have sought to develop an

algorithm that does not require the tangent linear version of the forecast model or $H^{\text{trial},q}$, both of which may require intensive effort to develop if unavailable.

3.2. Notation and assumptions

Assume that an ensemble filter (EF) (e.g. Evensen, 1994) data assimilation system is being used. Let the time interval between observations be a constant $\triangle t_{\text{obs}} = t_{i+1} - t_i$ (i is an integer). Define a reference time t_{i-M} (M is an integer). Assume that an OSSE has been run from t_{i-M} to t_i using H^{current}. Assume that at each assimilation time $t_m \in [t_{i-M}, \ldots, t_i]$, an ensemble forecast has been generated and archived. Let the ensemble forecast initiated at t_m be denoted by $[ens_m]$. $[ens_m]$ is the collection of *forecast* ensembles at times $t_{m+1}, \ldots, t_{m+L}$ (L is an integer). The lead time of each archived ensemble forecast is $L\triangle t_{\text{obs}}$. The forecast ensembles $[ens_m]$ are obtained by propagating the posterior/updated ensemble valid for t_m forward to time t_{m+L}. Let the set of $M + 1$ stored ensemble forecasts be denoted by $[ens_{i-M}, \ldots, ens_i]$. Fig. 2 depicts a representative two dimensional system during an OSSE where the observing network is

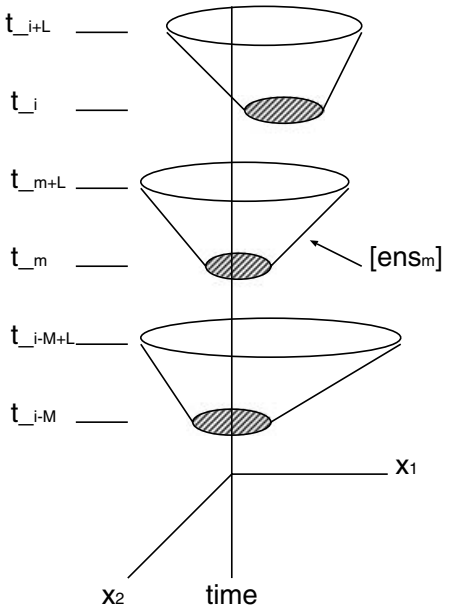

Fig. 2. A depiction of an OSSE where the observing network is H^{current} at each observing time. x_1 and x_2 are the system state variables and define a horizontal plane coming out of the page. The vertical axis is time increasing upwards. The shaded ovals represent the posterior/updated ensembles at times t_{i-M}, t_m, t_i. The surface area defined by each oval is parallel to the x_1, x_2 plane. The cones emanating from each posterior/updated ensemble represent the archived ensemble forecasts. The ovals at times $t_{i-M+L}, t_{m+L}, t_{i+L}$ represent the forecast ensembles initialized at times t_{i-M}, t_m, t_i respectively. The archived ensemble forecast $[ens_m]$ is depicted by the middle cone as indicated.

H^{current}. The shaded ovals represent the posterior/updated ensemble at times $[t_{i-M}, \ldots, t_m, \ldots, t_i]$. The cones emanating from each posterior/updated ensemble depict the archived ensemble forecasts.

3.3. Conditioning archived ensemble forecasts on the trial network $H^{\text{trial},q}$

An algorithm for conditioning the archived ensemble forecast $[ens_m]$ on the trial network $H^{\text{trial},q}$ will now be discussed. The conditioning is in the following sense. Using an ensemble Kalman filter (EnKF) based algorithm, $[ens_m]$ can be used to compute an estimate of the system state covariance at t_{m+L} *given* that the observing network at times $t_{m+1}, \ldots, t_{m+L}$ is $H^{\text{trial},q}$ (*not* just H^{current}). This covariance estimate is denoted by $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$, following the notation suggested by Bishop et al. (2001) for network design problems. $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ is an estimate of the posterior/updated 0-day forecast lead time system state covariance. The entire set of archived forecasts ($[ens_{i-M}, \ldots, ens_i]$) can be updated independently to obtain $M + 1$ estimates of covariance. An appropriate norm can be applied to all $M + 1$ estimates of covariance, which can then be averaged to assign a number to the objective function $\Phi(H^{\text{trial},q})$. Note that the archived forecasts $[ens_{i-M}, \ldots, ens_i]$ depicted in Fig. 2 need to be generated only once. As a result, many different configurations $H^{\text{trial},q}$ can be tested without having to repeatedly integrate the forecast model.

We now proceed with the technical description of the algorithm. Let $[ens_m]$ represent a sample of a prior distribution $\mathbf{p}(\mathbf{x}_{t_{m+1}}, \ldots, \mathbf{x}_{t_{m+L}})$. Let the network at times $t_{m+1}, \ldots, t_{m+L}$ be $H^{\text{trial},q}$. Given the values of observations at times $t_{m+1}, \ldots, t_{m+L}$ consistent with $H^{\text{trial},q}$ (denoted $\mathbf{y}^{\text{trial},q}$), the prior distribution can be updated using Bayes theorem,

$$\mathbf{p}\left(\mathbf{x}_{t_{m+1}}, \ldots, \mathbf{x}_{t_{m+L}}|\mathbf{y}^{\text{trial},q}\right) = \frac{\mathbf{p}\left(\mathbf{x}_{t_{m+1}}, \ldots, \mathbf{x}_{t_{m+L}}\right)\mathbf{p}\left(\mathbf{y}^{\text{trial},q}|\mathbf{x}_{t_{m+1}}, \ldots, \mathbf{x}_{t_{m+L}}\right)}{\text{Normalization}} \quad (1)$$

Observation values consistent with the trial network $H^{\text{trial},q}$ for times $t_{m+1}, \ldots, t_{m+L}$ must be specified. Let the values at times $t_{m+1}, \ldots, t_{m+L}$ be $\mathbf{y}^{\text{trial},q} = [H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+1}})^T, \ldots, H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+L}})^T]^T$ where $\overline{\mathbf{x}}^f_{t_{m+1}}$ and $\overline{\mathbf{x}}^f_{t_{m+L}}$ denote the ensemble mean of the archived forecasts valid for t_{m+1} and t_{m+L}, respectively.

Assume that the likelihood function $\mathbf{p}(\mathbf{y}^{\text{trial},q}|\mathbf{x}_{t_{m+1}}, \ldots, \mathbf{x}_{t_{m+L}})$ in eq. (1) is Gaussian with known variance and uncorrelated errors in both space and time. For the purposes of the update, let $\mathbf{p}(\mathbf{x}_{t_{m+1}}, \ldots, \mathbf{x}_{t_{m+L}})$ be approximated by a Gaussian. The covariance of the posterior/updated distribution in eq. (1) is then given by the Kalman filter (KF) posterior/updated covariance equation (Cohn, 1997). Using a Deterministic Ensemble Square Root Filter (DEnSRF) (Tippett et al. 2003), $[ens_m]$ can be transformed to give a posterior ensemble consistent with the

Kalman Filter (KF) update equations. The updated ensemble valid for t_{m+L} can then be used to compute a sample estimate of $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$. The KF updated covariance equation is independent of the specified observation values $\mathbf{y}^{\text{trial},q}$, but it is convenient to have $\mathbf{y}^{\text{trial},q} = [H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+1}})^T, \ldots, H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+L}})^T]^T$ as will be discussed below.

The above is an application of a DEnSRF in an augmented state space. In this paper, these updates are done using a particular DEnSRF, namely the ensemble adjustment Kalman filter (EAKF) of Anderson (2001, 2003). Note that for the DEnSRF algorithms considered in this paper, the normalization in (1) is not required. The reason is that, by assumption, each ensemble forecast is equally likely, or has weight $1/K$ (where K is the ensemble size). Under this assumption, the DEnSRF filter algorithm generates a set of equally likely updated ensemble members, whose covariance and mean are consistent with the KF updated equations. This update is done, irrespective of the number one assigns to the normalization. In this sense, the normalization is not required. Non-deterministic EnKFs (Houtekamer and Mitchell, 1998, 2001; Evensen, 1994; Evensen and van Leeuwen, 1996; Burgers et al., 1998) could also be used for this update but are not discussed in this paper. In this case, observation values have been specified over the entire interval from $t_{m+1}, \ldots, t_{m+L}$. $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ therefore corresponds to the posterior/updated 0-d forecast lead time covariance. A discussion of how this method extends to design problems for an arbitrary lead time is left to Appendix A.

Using a DEnSRF, a formal expression for $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ is given by eq. B.1 of Appendix B. Tippett et al. 2003 show that numerically equivalent computations of $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ will be obtained for any version of a DEnSRF when no covariance localization is used. Application of the ETKF formalism (Bishop et al., 2001) to condition $[ens_m]$ on $H^{\text{trial},q}$ yields an equivalent expression to eq. B.1. As shown in Bishop et al. (2001), when applied to a linear system, $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ will be equivalent to what would be obtained via a sequential in time filtering process. The motivation for using a DEnSRF in computing $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ is therefore related to dynamical linearity. Again, if the dynamics are linear from t_m to t_{m+L}, the sample estimate of $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ is equivalent to what would be obtained via a sequential in time filtering process where the observation values at times t_{m+k} are given by $\mathbf{y}^{\text{trial},q}_{m+k} = [H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+k}})^T]^T (k = 1, \ldots, L)$ (Bishop et al. 2001). This motivates our specification of $\mathbf{y}^{\text{trial},q} = [H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+1}})^T, \ldots, H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+L}})^T]^T$. For linear dynamics, sample estimates of $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ which are equivalent to a sequential in time filtering process can be obtained *without* requiring the repeated integration of the model equations.

When applying the RDA to a non-linear system, the 'equivalence' for linear dynamics should be kept in mind when specifying the length of the archived ensemble forecast ($L\triangle t$). Our suggestion is that L can be chosen to correspond to the time scale where the dynamics are nearly linear. This would ensure that sample estimates of $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ correspond to output from a sequential in time filtering process. Rough estimates of the linear dynamical time scale will be used in the numerical results in this paper. A detailed discussion of this point can be found in Khare (2004).

3.4. The expression for the objective function $\Phi(H^{\text{trial},q})$

The procedure outlined above can be followed for all $M + 1$ archived forecasts resulting in $M + 1$ sample covariance estimates conditioned on $H^{\text{trial},q}$. For a mean squared error norm (used in this paper), the trace of $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ is used and is denoted by $Tr(\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q}))$. The objective function used in this paper is then,

$$\Phi(H^{\text{trial},q}) = \sum_{j=-M}^{0} Tr\left(\mathbf{P}\left(\mathbf{x}_{t_{i+j+L}}|H^{\text{trial},q}\right)\right)/(M+1) \tag{2}$$

Note that if the verification region consists of a subset of state variables, the trace only includes state variables within the verification region. Information theory based norms could also be used (Schneider and Griffies, 1999; Majda and Timofeyev, 2000; Kleeman, 2002; DelSole, 2004). Error estimates associated with the evaluation of $\Phi(H^{\text{trial},q})$ can be obtained by computing the standard error of the mean.

3.5. Computational cost of evaluating $\Phi(H^{\text{trial},q})$ using the RDA

Assume that an OSSE with $H^{\text{trial},q}$ has been run and the ensemble forecasts $[ens_{i-M}, \ldots, ens_i]$ have been archived. In computing eq. (2) no further model integrations are required. The archived ensemble forecasts can first be updated with the B observations not being optimized. This pre-processing is done without approximation when observations are assimilated sequentially, given the assumption of independent observational errors. To be consistent with eq. (1), the B observations not being optimized must, of course, influence the entire augmented state vector $\mathbf{z}_m = [\mathbf{x}^T_{t_{m+1}}, \ldots, \mathbf{x}^T_{t_{m+L}}]^T$ (consisting of L states of dimension, n), that is, a given observation must influence present, past and future states. Updating each $[ens_m]$ with observations from one time level requires L assimilations of B observations onto a state vector of size n. Updating each $[ens_m]$ with observations from all time levels therefore requires L^2 assimilations of B observations onto a state vector of size n.

The trial network $H^{\text{trial},q}$ consists of S observations being optimized. For $[ens_m]$, one is only interested in the covariance at t_{m+L}. The S observations at a particular time need only influence present and future states. Therefore the cost of evaluating $\Phi(H^{\text{trial},q})$ is the cost of doing $L(L + 1)/2$ assimilations of S observations onto a state vector of size n for all $M + 1$ archived forecasts. The RDA can be combined with heuristic optimization algorithms such as simulated annealing. When applying

heuristic optimization algorithms to complicated nonlinear objective functions, many evaluations of the objective function may be required. Eliminating the need for repeated integrations of the prediction model in evaluating $\Phi(H^{\text{trial},q})$ is important to reducing the computational cost. Note that in computing eq. (2) using the RDA, all our conditions discussed in section 3.1 have been satisfied.

4. Application of the retrospective design algorithm to a simulated global prediction system

This section begins with a description of the model (Section 4.1) and design problem to be studied (Section 4.2). In Section 4.3, results obtained by doing OSSEs will be discussed in detail. Finally, in Section 4.4, the results obtained using the RDA will be compared to the results obtained from OSSEs.

4.1. Model description

The prognostic variables in the atmospheric general circulation model (AGCM) are zonal wind U, meridional wind V, temperature T and surface pressure PS. Our intention was to use a simplified model whose behaviour is qualitatively related to the observed atmosphere. The model that has been used in our numerical experiments is the dynamical core used in the GFDL global atmospheric model (Anderson et al., 2005, hereafter A05), using a forcing and dissipation suggested by Held and Suarez (1994). The forcing is a Newtonian relaxation towards a zonally symmetric state without any daily or seasonal cycle. The form of the forcing is equivalent in the northern and southern hemispheres. In both hemispheres, the forcing drives a pole to equator temperature gradient. A simple Rayleigh damping is used near the surface for dissipation. The model has no topography, landmasses or parametrizations of subgrid scale processes. The forcing used in this model can be thought of as a replacement for detailed radiative, turbulence and moist convective parametrizations (Held and Suarez, 1994). It is emphasized that the forcing and dissipation in the model is zonally symmetric.

Numerical integration of this model is done using a B-grid discretization and a vertical discretization described in Simmons and Burridge (1981). The model consists of 30 by 60 equally spaced points in latitude and longitude, respectively with five vertical levels. This discretization results in a total of 28 200 prognostic state variables. A model time step of 1 hour was used. This is approximately the minimum horizontal and vertical resolution required to generate baroclinic instability with a time-mean climatology that is somewhat similar to the observed atmosphere (A05). A random sample of the model's climatology was obtained by making a 100-yr integration of the AGCM from a state of rest. The temperatures at the highest level of this model equilibrates very slowly compared to all the other variables (A05). It may take several years of integration for the highest level temperatures to equilibrate, while it appears all other state variables have equilibrated (A05). As a result, a 100-yr integration was used to obtain a state considered to be a random sample of the model's climatology. This state was integrated an additional 4000 d. The zonal mean, time mean zonal wind over the 4000 d is plotted in Fig. 3. Snapshots of the mid-level (level 3) temperature and surface pressure fields at a particular time are shown in Figs 4 and 5 respectively. Eastward-moving wave-packets in the temperature and pressure fields are observed in the mid-latitudes.

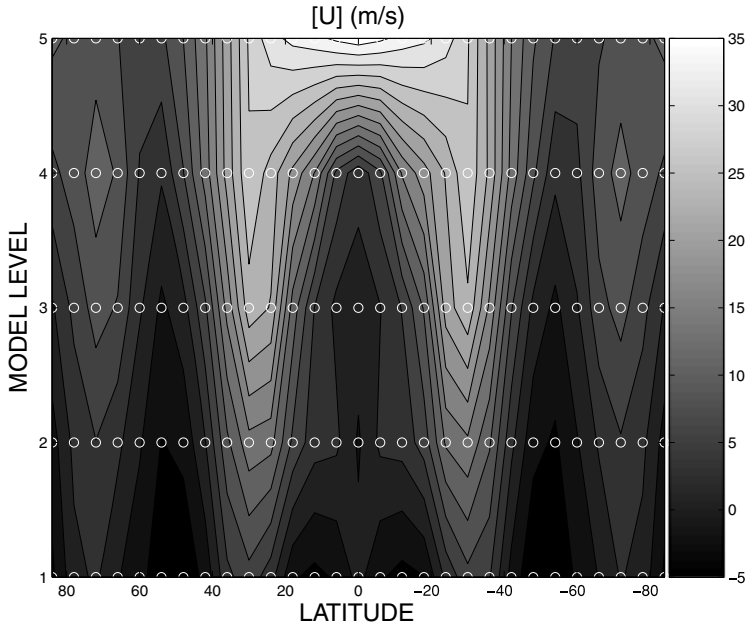

Fig. 3. Zonally averaged, time mean zonal wind. Contours are drawn every 2.5 m s^{-1} from −5 to 35 m s^{-1}. White circles indicate model grid points for the zonal wind U.

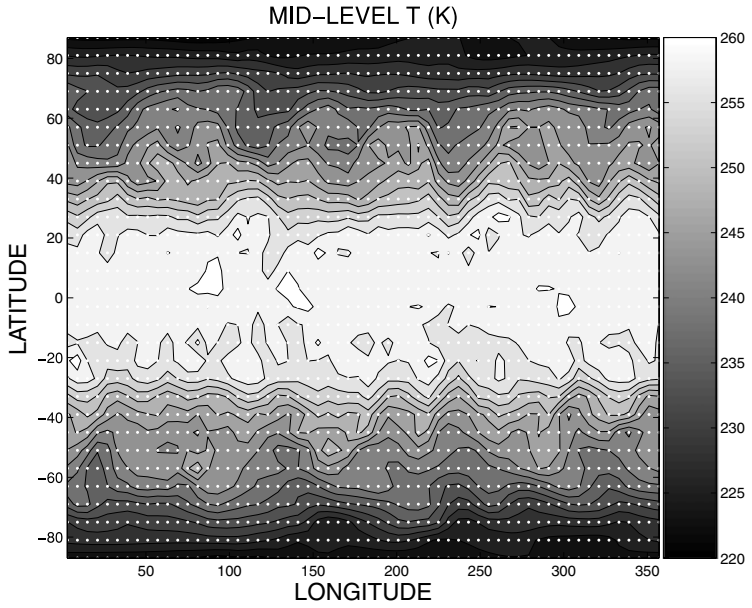

Fig. 4. Snapshot of mid-level temperature. Contours are drawn every 2.5 K from 220 to 260 K. White dots indicate horizontal locations of T grid points.

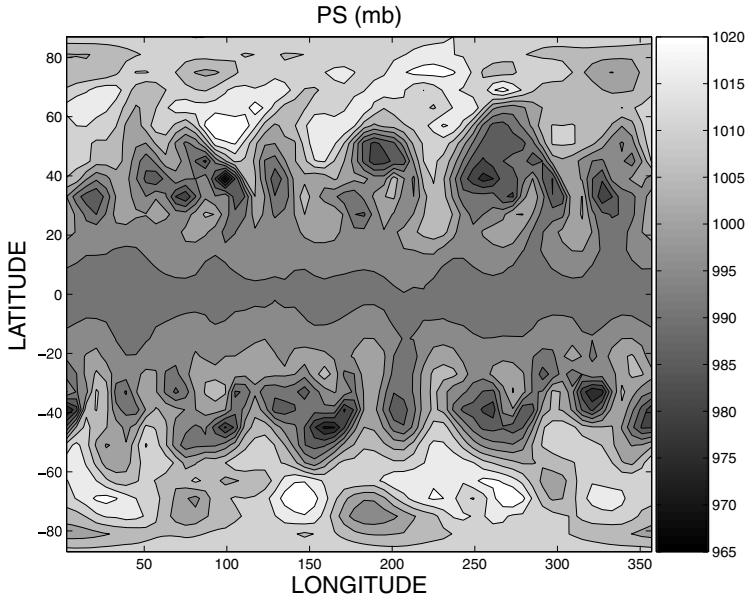

Fig. 5. Snapshot of surface pressure *PS*. Contours are drawn every 5mb from 965 to 1020mb. The horizontal locations of the *PS* grid points are the same as Fig. 5.

4.2. The design problem

In a spatially extended model like an AGCM, any number of experimental designs are possible. In making this initial comparison between the RDA and OSSEs, we have chosen a simple experimental design which affords a relatively clean interpretation of the cost functions being compared. While our chosen experimental design involves a number of specific parameter choices, some additional commentary regarding the sensitivity of our results to these parameter choices will be provided in Section 4.5. We have chosen a design problem where the influence of the model dynamics is likely to be strong in the following sense. The equator to pole temperature gradient specified in the AGCM's forcing drives a jet stream that is strongly peaked at roughly 33° latitude (see Fig. 3). The problem of choice is to determine optimal placement of observations along the 33°N latitude circle for a variety of forecast lead times and verification region sizes. Given the strong zonal flow at this latitude, it is likely that the downstream influence of observations can be more easily ascertained than had we chosen a latitude with a weak zonal flow.

Having constrained the problem to a latitude circle with a strong zonal flow, a variety of experiments are still possible. In a practical setting, the type of problem will depend on the type

of instruments being optimized (radar, radiosonde, GPS) and the choice of verification region. For simplicity, this paper focuses on the design of *PS* observations to reduce errors in *PS* for a variety of forecast lead times and verification region sizes. This choice has both theoretical and practical motivation. For mid-latitude large-scale atmospheric dynamics, *PS* is strongly correlated to winds and temperatures. A number of perfect model (Anderson et al., 2005) and real forecast model (Whitaker et al., 2003) ensemble data assimilation studies show that a high degree of information about the entire atmospheric state can be obtained from observations of just *PS*. Given the relative ease with which *PS* stations can be designed and implemented, considerable interest in the design of *PS* networks is developing within the NWP (especially meso-scale) community (Shapiro and Thorpe, 2004).

Our choice of problem can now be stated. The goal is to find the optimal placement of a single *PS* observation along the latitude circle at 33°N to reduce errors in *PS* along the same latitude circle (for a variety of different longitudinal verification regions). The latitude circle is composed of 60 equally spaced locations (i.e. 6° longitude separation). The first *PS* grid point is located at (33, 6) degrees, while the final grid point is at (33, 360). The locations of the *PS* observations are the same as the *PS* in the B-grid numerics.

Before finding the optimal placement of a single *PS* observation along the 33N latitude circle, the current-day observing network H^{current} must be specified. H^{current} has been specified so that when a long OSSE is run with H^{current}, the uncertainty in *PS* along the 33° latitude circle is less than climatology, but not so small that the placement of an additional observation is insignificant. To simplify the interpretation of the results, the H^{current} has been designed to be zonally symmetric. The rationale is that zonal asymmetries in H^{current} will introduce zonal asymmetries in the cost functions, which in turn complicates the interpretation of the results.

The locations of *PS* observed by the H^{current} are depicted in Fig. 6. All 60 *PS* grid points are observed at each latitude where observations are taken. H^{current} does not consist of any other observations. The observed latitudes are [3, 9, 15, 21, 45, 51, 57, 63, 69, 75, 81, 87] degrees, leaving a data void or 'gap' along the latitude of interest. The total number of *PS* observations at each observing time is 720. As will be discussed in Section 4.3, the chosen observational standard deviation of the *PS* observations in H^{current} has been tuned through trial and error to reduce the mean standard deviations in *PS* along the 33° latitude circle to roughly half the climatological standard deviation. To emphasize the influence of the additional *PS* observations, its observational standard deviation has been set to 0.1mb.

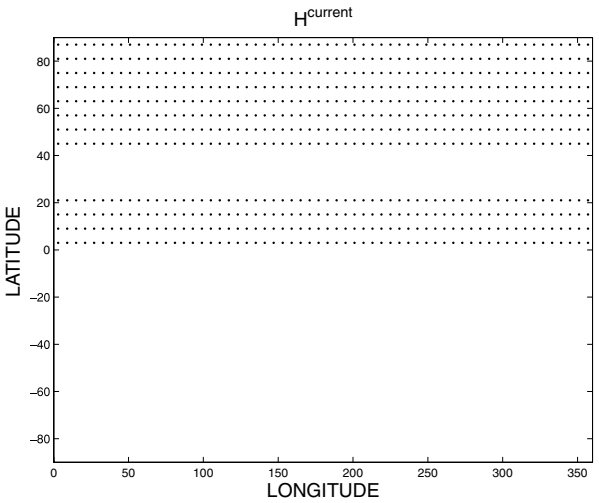

Fig. 6. H^{current}—the current day fixed network. *PS* observations are taken at locations indicated by the dots every 12 h. The observational standard deviation for each *PS* observation is 7mb.

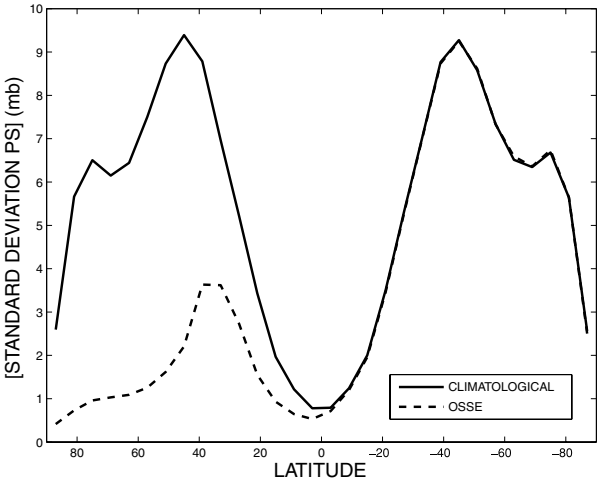

Fig. 7. Zonally averaged, time mean *PS* standard deviation for climatology versus latitude is depicted by the solid line. The same quantity for output from an OSSE where the observing network is H^{current} depicted in Fig. 6 is depicted by the dashed line.

4.3. *OSSE results*

Current day operational systems are limited to roughly 10^2 ensemble members. To mimic this restriction, an ensemble size of $K = 20$ was used for all the OSSE results. An initial ensemble was generated by adding random numbers from a normal distribution with mean 0 and standard deviation 10^{-7} to all model state variables of a random sample of the climatology designated to be the 'true' state. The 20-member ensemble and the initial 'true' state were integrated for additional 10 yr. The true state and ensemble were then considered a random sample of the climatological distribution (our experience with this model suggests that such a random sample could have been achieved using a shorter integration). The ensemble was integrated a further 200 d. The time mean, zonal mean *PS* standard deviation is given by the solid line in Fig. 7.

A05 run a set of numerical experiments to measure the global RMS T error growth in all five levels of this AGCM. A05 find error doubling times of approximately 4 d which is roughly twice that of real forecast models (A05, Simmons et al., 1995; Simmons and Hollingsworth, 2002). Generally, error doubling times in atmospheric models are strongly dependent on the resolved scales (Lorenz, 1995). This motivates our choice of $\triangle t_{\text{obs}} = 12$ (twice the 6-hour interval used in many operational systems).

The ensemble data assimilation method used was the EAKF (Anderson, 2001, 2003). Observations were assimilated sequentially. A heuristic adjustment to the sequential EAKF algorithm called covariance localization (CL) was used to help alleviate problems due to sampling errors. Suppose that the EAKF is updating state variable x_α with an observation of a physically unrelated state variable x_β. For small ensemble sizes, the observation of x_β may be spuriously correlated to x_α in some instances. Consequently, updating the ensemble of state variable x_α with an observation of x_β may degrade the quality of the analysis (Hamill et al., 2001). To address this problem, covariances between the prior state variables and observation variables are multiplied by a fifth-order piecewise rational function with compact support which decreases as a function of horizontal distance along the surface of the sphere (Houtekamer and Mitchell, 1998; Gaspari and Cohn, 1999). The function is characterized by c, the half-width of the correlation function (Gaspari and Cohn, 1999). A half-width parameter of $c = 0.3$ radians was used for all the OSSE results. Trial and error reveals that significantly higher and lower values of c yield lower quality assimilations. This value of $c = 0.3$ is roughly similar to values used in other studies with ensemble filters in AGCMs (Hamill et al., 2001; Houtekamer and Mitchell, 2001, A05;). To give a feel for the length scales involves with this parameter choice, Fig. 8 depicts the localization function centered at grid point 29 as a function of distance along 33°N latitude circle.

Starting with the $K = 20$ member sample of climatology and the true state, an OSSE with H^{current} was performed. All *PS* observations were obtained by adding a random error, sampled from a Gaussian distribution with a standard deviation of 7mb, to the true value of *PS*. An OSSE has been run for the equivalent of 300 d. The time mean (over the last 200 d), zonal mean, assimilation time standard deviation of *PS* versus latitude is given by the dashed line in Fig. 7. For the latitude of interest at 33N, Fig. 7 shows that the uncertainty has been reduced by roughly a factor of two compared to climatology. The uncertainty values along the latitude band of interest are less than climatology but opportunity exists for decreasing the uncertainty by the placement of one-additional observation. Animations of the posterior *PS* standard deviation reveal somewhat coherent packets of uncertainty being advected eastward along the 33° degree latitude circle. For the southern hemisphere, Fig. 7 shows that the *PS* standard deviation is nearly unchanged for the OSSE compared to climatology. Given a localization half width of $c = 0.3$ radians, observations of northern hemisphere *PS* will only influence state variables within roughly 34°. Therefore, the agreement for the southern hemisphere in Fig. 7 is partially due to the localization which is limiting the impact of the northern hemisphere observations on southern hemisphere state variables. However, all 240 *PS* observations in H^{current} at latitudes [3, 9, 15, 21] will have *some* influence on a subset of southern hemisphere state variables if there exists any non-zero correlation between the observed variables and the southern hemisphere state variables being updated. The fact that Fig. 7 reveals such clear cut agreement in the southern hemisphere suggests that the northern and southern hemisphere dynamics are weakly correlated.

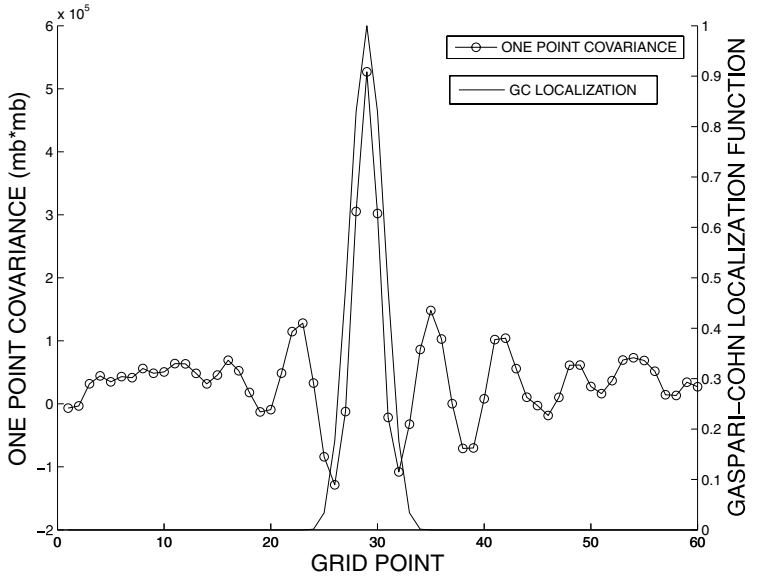

Fig. 8. In this plot, the reference point is grid point 29. The solid line intersecting the circles depicts the one point covariance along the 33°N latitude circle between the *PS* at the reference point and the *PS* at all the other grid points. The covariances are computed from the AGCMs climatology. The solid line is a plot of the Gaspari-Cohn localization function used in the OSSEs centered around the reference grid point. The half-width of the localization function is roughly five grid points.

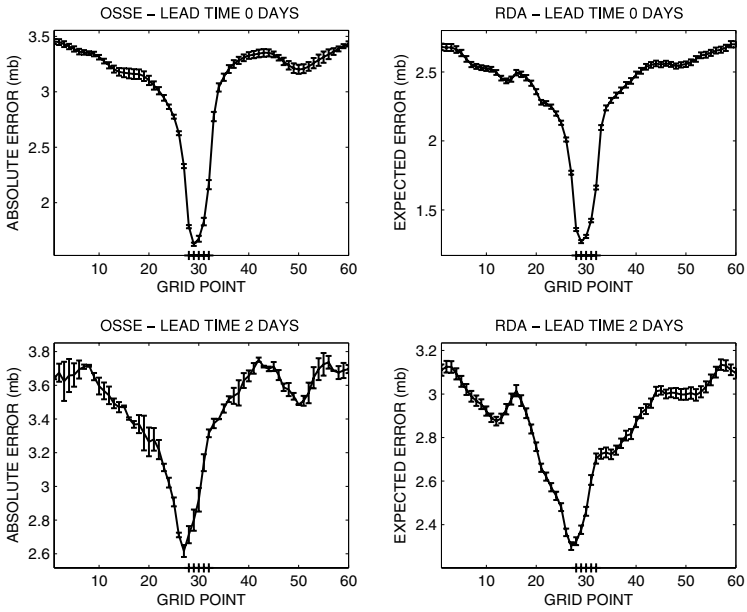

Fig. 9. Comparison of OSSEs versus the RDA for a small verification region. On each panel, the horizontal axis indicates the grid point along the 33° latitude circle where a single *PS* observation was tested. The *PS* grid points which comprise the verification region are indicated by the +. For all panels, error bars were estimated by calculating the standard error of the mean over 1000 samples. The upper (lower) left-hand panels are the mean absolute errors (mb) for the verification region obtained from OSSEs for a 0 (2) d forecast lead time. The upper (lower) right panels depict the mean expected errors (mb) for the verification region obtained using the RDA for a 0 (2) day forecast lead time. Note that the climatological time mean zonal mean standard deviation at 33°N is roughly 7.5mb.

4.3.1. Results for a localized verification region. The *PS* grid points along the 33N latitude circle at longitudes [6, 12, . . . , 360] will be called grid points 1, 2, . . . , 60. For the results in Fig. 9 the verification region consists of five consecutive *PS* grid points at latitude 33N and longitudes [168, 174, 180, 186, 192, 198], that is grid points [28, 29, 30, 31, 32]. The directions towards grid point 1 is upstream (west) and towards grid point 60 is downstream (east). The verification region in the upper left-hand panel of Fig. 9 is indicated by the crosses on the horizontal axis.

To obtain the results in the upper left panel of Fig. 9, 60 independent OSSEs were run. The OSSEs all begin with the same initial ensemble and truth, but are independent in the sense that the location of the single additional *PS* observation is different for all 60 OSSEs that were run. In each case, the observing network consists of H^{current} plus one additional *PS* observation located at one of the 60 grid points (i.e. $H^{\text{trial},q}$ with $B + S = 721$ *PS* observations). The upper left panel of Fig. 9 is a plot of the mean 0-d forecast errors versus the grid point location of the one additional *PS* observation. The errors are joined by a solid line for aesthetic purposes only. For all 60 independent experiments, an 11 000-d assimilation was run. Results from the last 10 000 d were used to compute statistics. Errors were collected every 10 d. The error doubling time of the AGCM is roughly 4 d (A05), so estimates of error every 10 d were assumed statistically independent.

The errors were computed in the following way. For a given point in the verification region, the error is the absolute difference between the ensemble mean and truth in *PS*. The errors were averaged over the verification region for all 1000 independent times. The vertical axis for the upper left panel of Fig. 9 is the mean over the 1000 statistically independent assimilation times. Ensemble means are commonly used to evaluate forecast/analysis errors in ensemble data assimilation studies (Whitaker and Hamill, 2002; Snyder and Zhang, 2003; Anderson, 2001; Houtekamer and Mitchell, 1998, 2001). The error bars on the upper left panel of Fig. 9 were computed using the standard errors of the mean. The population standard deviations for the average error in the verification region were computed using the 1000 independent error estimates.

Nearby the verification region, the upper left panel of Fig. 9 has a statistically significant upstream shift. Note that even though the additional *PS* observation is highly accurate (standard deviation of 0.1mb), the numerical values on the vertical axis represent the average error over the verification region. When the verification region size was shrunk to just one grid point, the minimum error dropped to roughly 0.1mb. The subtle, yet statistically significant dip in the cost function for locations nearby grid point 52 is interesting but not understood. The analogous cost function for 2-d forecast errors is depicted in the lower left panel of Fig. 9. The upstream shift in the cost function is not surprising given the strong zonal flow. To determine, whether or not the magnitude of the upstream shift is physically reasonable, one can look at the distance traveled by a particle being advected by the zonal/time mean zonal wind at 33N. Assume that the advection speed is characterized by the mid-level wind in Fig. 3 at 33°N (roughly 20 m/s). A particle traveling along the 33°N latitude circle would cover roughly 7% of the circumference, corresponding to roughly four grid points. Therefore, the magnitude of the upstream shift on the lower left panel of Fig. 9 is reasonable.

In Fig. 8, the one-point climatological covariance for *PS* along the 33N latitude circle is plotted. The reference point in Fig. 8 is grid point 29. The positive and negative values of correlation away from the reference point is likely due to the baroclinic waves present in the mid-latitudes. Fig. 8 provides some physical

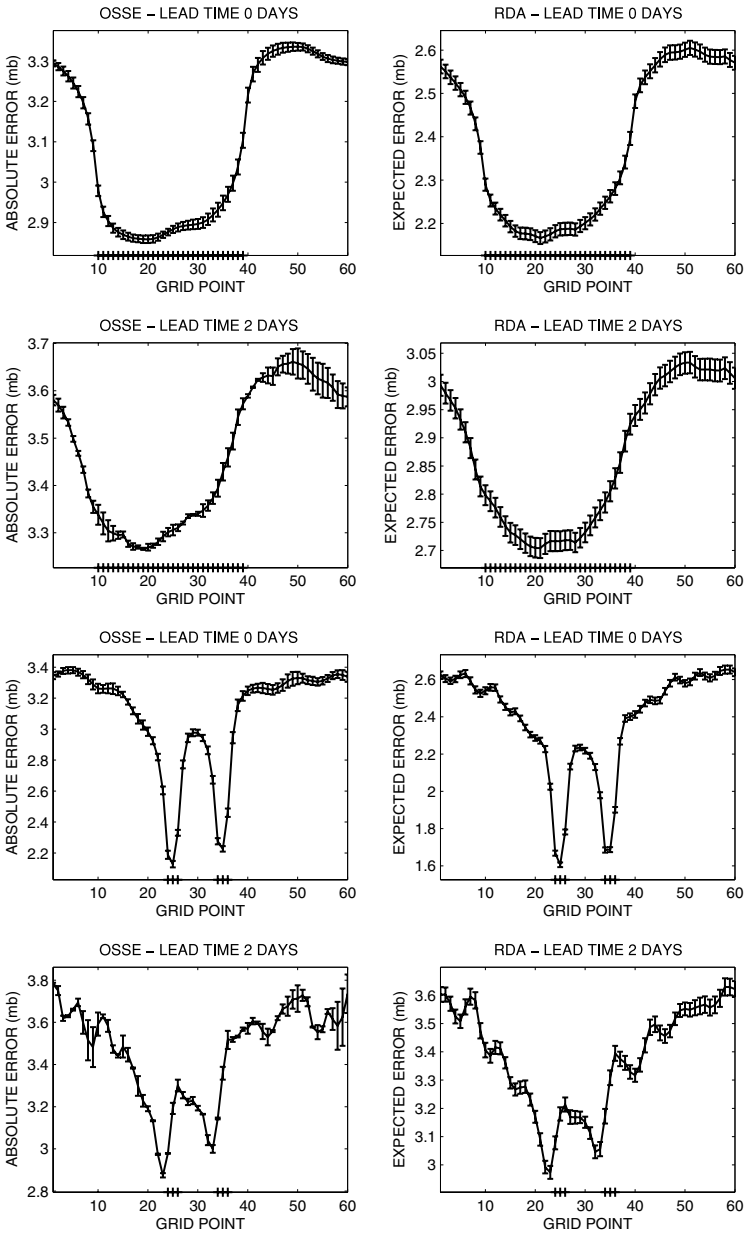

Fig. 10. As in Fig. 9 but for a large verification region. Note that the climatological time mean zonal mean standard deviation at 33°N is roughly 7.5mb.

Fig. 11. As in Fig. 9 but for a discontinuous verification region. Note that the climatological time mean zonal mean standard deviation at 33°N is roughly 7.5mb.

basis for understanding the increase in errors away from the verification region depicted in Fig. 9. Generally, as revealed by Fig. 8, climatological correlations decay quickly away from a given grid point suggesting that a given observation is more likely to influence nearby grid points. This helps explain the somewhat sharp increase in errors when observations are placed at locations outside the verification region, which has been been observed for the results in Fig. 9 and the subsequent results in Figs 10 and 11.

For the OSSE results in Fig. 9 and subsequent figures, there are occasionally large differences in the magnitude of the error bars between neighboring grid points. This is attributed to the fact that the population standard deviations were inaccurately estimated using just 1000 samples. Obtaining very accurate estimates of the population standard deviations would require very long integrations and was computationally infeasible.

4.3.2. Results for a large verification region. The verification region consists of half the *PS* grid points along 33N. The cost function for the OSSE assimilation time errors is depicted in the upper left panel of Fig. 10. The upstream character of the cost function is exhibited by comparing the differences in the error when the observation is located at the western and eastern boundaries of the verification region. The lower left panel of Fig. 10 shows the analogous cost function for the 2-d forecast problem. Only a slight upstream 'shift' in the cost function is apparent.

4.3.3. Results for a discontinuous verification region. The verification region now consists of grid points [24, 25, 26, 34, 35, 36]. The OSSE cost function for assimilation time errors is given by the upper left panel of Fig. 11. It is interesting to see that the western region local minima yields lower errors. This is attributable to the strong zonal flow. The analogous result for the 2-d forecast problem is depicted in the lower left panel of Fig. 11. Once again, there is a significant upstream shift in the shape nearby the verification region.

4.4. RDA results

The results for the RDA make use of 1000 archived ensemble forecasts obtained from a single OSSE with H^{current}. The initial times of each archived forecast are separated by 10 d. It was assumed that applying the RDA to this set of ensemble forecasts resulted in 1000 independent estimates of expected error. In Khare (2004) the RDA is applied to the Lorenz 1996 model (Lorenz, 1995). There, the use of covariance localization in computing $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ reduces noise in the estimates of $\Phi(H^{\text{trial},q})$. An issue when using covariance localization in this context is that observed variables and state variables being updated are not necessarily valid for the same time. One approach for handling this type of problem is the group filter localization technique. The group filter localization technique and details concerning how the RDA results were generated are left to Appendix C.

The chosen length of the archived ensemble forecasts is $L\triangle t_{\text{obs}} = 4$ d. Given that $\triangle t_{\text{obs}} = 12$ h, the value of L defined in Section 3.3 is 8. Note that 4 d is roughly equivalent to empirical estimates of the AGCM's error doubling time. The rationale for this choice of $L = 8$ is that the dynamics remain weakly nonlinear up to 4 d. When covariance estimates $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ are computed using covariance localization, the result is not necessarily equivalent to a sequential in time filtering process even when the model dynamics are linear. The reason is that a covariance localization for covariance between state variables at different times required by the RDA can not be easily related to the covariance localization for covariance between simultaneous variables used in a sequential in time filtering procedure (Khare and Anderson, 2006).

4.4.1. Results for a localized verification region. The upper right panel of Fig. 9 is the 0-d forecast error results for the localized verification region analogous to the upper left panel for OSSEs. The expected errors on the vertical axis of the upper right panel of Fig. 9 are obtained in the following way. The RDA processes the 1000 archived forecasts to generate 1000 independent estimates of standard deviation for all state variables in the verification region (*PS* at grid points $28 \rightarrow 32$). For each archived forecast, the error is given by the mean standard deviation over the verification region. The vertical axis is the mean over the 1000 cases. The error bars on Fig. 9 are given by the standard error of the mean. Nearby the verification region, the RDA does a respectable job of mimicking the upstream character of the results generated by the OSSEs (upper left-hand panel of Fig. 9). Subtle differences are evident.

Note that for all comparisons between OSSEs and the RDA, results are plotted on different scales. Recall that the output from the RDA would be equivalent to a sequential in time filtering process if the model dynamics were linear (assuming no covariance localization). For linear model dynamics the expected errors obtained by the RDA should be equivalent to the expected errors obtained form OSSEs. Given that the AGCM's dynamics are non-linear, there is no a priori reason why the numerical values of the cost functions in the upper panels of Fig. 9 should match. Note that the values of the expected errors for the RDA are typically smaller than the mean absolute errors for the OSSEs. It is possible that when applied to other nonlinear systems or design problems this result would not hold. For the results in the upper panels of Fig. 9 and all subsequent comparisons, the difference between maximum and minimum values of the cost functions are roughly equivalent. Therefore, if the RDA is capable of mimicking the *shape* of the OSSE cost function, it serves as a useful method for assessing the relative value of placing the *PS* observation in different locations for the problems examined in this paper.

The analogous plot for the 2-d forecast error problem is given by lower right panel of Fig. 9. The RDA shows the impact of the advection in assessing the relative value the additional *PS* observation.

4.4.2. Results for a large verification region. The cost function for the RDA assimilation time errors is depicted in the upper right panel of Fig. 10. The RDA does a nice job in mimicking the general shape of the OSSE cost function. The RDA cost function has a similar upstream character. The analogous result for the 2-d forecast problem is depicted in the lower right-hand panel of Fig. 10.

4.4.3. Results for a discontinuous verification region. The analogous result for the RDA assimilation time errors is given in the upper right panel of Fig. 11. The RDA again does a reasonable job of mimicking the information in the OSSE cost function. The first local minima yields lower expected error. The analogous result for the 2-d forecast problem is depicted in the lower right panel of Fig. 11. As for the OSSE results, there is a significant upstream shift nearby the verification region.

4.5. Further discussion

In this initial investigation of the RDA's ability to mimic information derived from OSSEs, we have focused on a simplified experimental design with specific parameter choices. A few comments regarding the sensitivity and robustness of these results to alternative parameter choices will now be provided. What if H^{current} included zero observations? When H^{current} is empty, the archived ensemble forecasts used by the RDA would simply be evolving samples of the model's climatology. While we have not tested this particular case using the AGCM, we note that in

Khare (2004), comparisons between the RDA and OSSEs for this particular case were made using a 40-variable Lorenz 1996 model. A high degree of agreement between the RDA and OSSEs was obtained. What if H^{current} included observations at all the *PS* grid points in the northern hemisphere and we wanted to determine the optimal placement of one additional observation along the 33N latitude circle? Assuming that the most important factor in the RDA's ability to approximate OSSEs is its ability to ascertain the downstream influence of observations, and based on the results of this paper, our feeling is that the RDA would remain an effective way of approximating OSSEs in this case. In fact, if H^{current} included observations at all *PS* grid points in the northern hemisphere which resulted in a smaller time mean zonal mean standard deviation in *PS* along the 33N latitude circle, the RDA may be even more effective as the implicit linearity requirement of the RDA would more likely be satisfied. Finally, we note that for the experiments in this section, the discrepancy between the observational standard deviations in H^{current} and the additional observations is large (7 and 0.1mb, respectively). As discussed above, we chose a standard deviation of 7mb for the observations in H^{current} so that the time mean zonal mean standard deviation in *PS* along 33°N was roughly half of climatology. We chose a standard deviation of 0.1mb for the additional observation so that its influence can be easily ascertained. Had we chosen a larger more realistic observation error, the 'signal' due to the additional observation would have been less distinct. More samples (archived ensemble forecasts) and longer OSSEs would have been required to make meaningful comparisons between cost functions. The key limiting factor in applying the RDA is the implicit linearity assumption regarding the evolution of the archived ensemble forecasts. Our choice of 0.1mb was merely a practical choice which has allowed us to make meaningful cost function comparisons while reducing computational expense.

5. Discussion and conclusions

A methodology for fixed observational network design has been presented. The methodology, called the retrospective design algorithm (RDA), uses output from an ensemble data assimilation system to evaluate the objective function. The objective function relates uncertainty in estimating the system state to the spatial configuration of a fixed observational network. The RDA does not require repeated integrations of the model equations in successive evaluations of the objective function, in contrast to observing system simulation experiments. The RDA is therefore a computationally efficient means of computing the objective function. The RDA is a methodology for computing an objective function which assesses the average influence of a trial observing network on an archive of independent ensemble forecasts. The novel contribution of this paper is the recognition that such an average can be efficiently used for fixed network design and the first proof of this concept in a simulated global prediction system.

The utility of the RDA has been tested by making direct comparisons to OSSEs using experiments in an atmospheric general circulation model. The problem of finding the optimal placement of a single surface pressure observation in the mid-latitudes for small, large and discontinuous verification regions for 0-2 d forecast lead times has been examined. A high degree of correspondence in the shapes of the objective functions computed via the RDA and OSSEs was found. While the results in this paper are indeed encouraging, further research is required to assess the RDA's utility in designing networks with many observations. The results in this paper provide a foundation of understanding which will be useful in such future work.

The theoretical developments and numerical experiments in this paper have focused on the problem of finding the optimal placement of an additional fixed observation to the current day observing network. The RDA can be applied to a wide variety of problem setups. For example, the RDA could be applied to the problem of finding the optimal re-configuration of a subset of today's fixed global observing network. A detailed and thorough discussion of applying the methods of this paper to more general design problems can be found in Khare (2004).

Application of the RDA does not require finding a suitable linear basic state of the prediction model or code for the prediction model and observation operator linearizations. However, our suggestion is that the lead time of the archived ensemble forecasts used in implementing the RDA should not dramatically exceed a reasonable estimate of the prediction model's linear dynamical time scale. It is in this sense that linearity comes in when using the RDA. A challenge for future researchers is to develop algorithms which are significantly more computationally efficient than OSSEs which do not implicitly or explicitly appeal to linearity.

Our initial investigation into the utility and application of the RDA leads us to a number of interesting research questions. What are the most suitable types of optimization algorithms to use in conjunction with the RDA? When applying the RDA to a given design problem, what is the impact of the chosen norm? Can the RDA be combined with the economic framework of Morss et al. (2004) to yield a practical approach to evaluating public investment in observing systems? Finally, can the methods of Bishop et al. (2003) and the RDA be combined to create a methodology more useful than either method independently? We hope that questions of this nature will be addressed in future studies.

Can the RDA be applied to ensemble forecasts generated from operational centers to help answer any operationally relevant design questions? Given that related methods have already been applied in the context of adaptive observations to operational ensemble forecasts (Majumdar et al., 2001, 2002), the RDA is certainly a promising method for designing fixed observing networks in operational settings. This results in this paper provide motivation for archiving ensemble forecasts generated over

many seasons at operational centres. As the use of ensemble data assimilation systems becomes more widespread in geophysical sciences, the potential for using the RDA to answer questions of optimal design for synoptic-scale, climate, oceanic and mesoscale prediction problems increases.

6. Acknowledgments

This work was completed while the first author was a PhD student in the Atmospheric and Oceanic Sciences programme at Princeton University. The authors would like to acknowledge members of GFDL and the AOS programme at Princeton for helpful discussions leading up to this work. The authors would also like to thank Tomoko Matsuo and Alain Caya for helpful comments on earlier versions of this paper.

Appendix A. Extension to forecast lead time problems

When conditioning $[ens_m]$ on $H^{\text{trial},q}$ in section 3.3, observation values are specified at times $t_{m+1}, \ldots, t_{m+L}$, therefore the $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ corresponds to posterior/updated 0-d forecast lead time covariance. Estimates of covariance can also be computed for a desired forecast lead time. When updating $[ens_m]$ for a forecast lead time of $(L-l)\ \triangle t_{\text{obs}}$, assuming that $L > l$, observations should only be specified from $t_{m+1} \rightarrow t_{m+l}$ in the interval $t_{m+1}, \ldots, t_{m+l}, \ldots, t_{m+L}$. Therefore, the specified observation values in eq. (1) should be set to $\mathbf{y}^{\text{trial},q} = [H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+1}})^T, \ldots, H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+l}})^T]^T$.

Appendix B. Computing $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ using a Deterministic ensemble square root filter

Here, the problem of computing $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ from $[ens_m]$ with a DEnSRF is examined in detail. Let $[\mathbf{x}^f_{t_{m+j},\kappa}]$ denote the ensemble forecast for t_{m+j}, where $\kappa = 1, \ldots, K$ with ensemble size K. The index $j = 1, \ldots, L$ where L indicates the length of the archived ensemble forecast. Superscript f denotes forecast. The matrix of perturbations $\mathbf{X}^f_{m+j}$ is an $n \times K$ matrix whose κ^{th} column is $[\mathbf{x}^f_{t_{m+j},\kappa} - \overline{\mathbf{x}^f_{t_{m+j},\kappa}}]/\sqrt{K-1}$, where $\overline{\mathbf{x}^f_{t_{m+j},\kappa}}$ is the archived ensemble forecast mean valid at t_{m+j}. Applying the formalism for the DEnSRF (Tippett et al. 2003) to the method described in Section 3.3, the expression for the updated covariance for t_{m+L} is given by,

$$\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q}) = \mathbf{X}^f_{m+L}\mathbf{S}_{m+1}\ldots\mathbf{S}_{m+L}{\mathbf{S}_{m+L}}^T\ldots{\mathbf{S}_{m+1}}^T{\mathbf{X}^f_{m+L}}^T \tag{B1}$$

where the $\mathbf{S}$ matrices are obtained by solving the equation $\mathbf{S}_{m+1+\alpha}{\mathbf{S}_{m+1+\alpha}}^T = (\mathbf{I} - \mathbf{V}_{m+1+\alpha}\mathbf{D}^{-1}_{m+1+\alpha}{\mathbf{V}_{m+1+\alpha}}^T)$ where $\mathbf{D}_{m+1+\alpha} = {\mathbf{V}_{m+1+\alpha}}^T\mathbf{V}_{m+1+\alpha} + \mathbf{R}$ and $\mathbf{V}_{m+1+\alpha} = (H^{\text{trial},q}(\mathbf{X}_{m+1+\alpha}\mathbf{S}_m, \ldots, \mathbf{S}_{m+\alpha}))^T$. $\mathbf{R}$ is the $(B+S)\times(B+S)$-dimensional error covariance associated with $H^{\text{trial},q}$. The index $\alpha = 0, \ldots, L$. Note that in eq. B.1, the dots between $\mathbf{S}_{m+1}$ and $\mathbf{S}_{m+L}$ denotes the product of matrices. For example, if $L = 3$, then $\mathbf{S}_{m+1}\ldots\mathbf{S}_{m+L}$ denotes $\mathbf{S}_{m+1}\mathbf{S}_{m+2}\mathbf{S}_{m+3}$. $\mathbf{S}_m$ is a $K \times K$ identity matrix. $H^{\text{trial},q}(\mathbf{X}_{m+1+\alpha}\mathbf{S}_m, \ldots, \mathbf{S}_{m+\alpha})$ is defined for nonlinear observation operators using the method described in Houtekamer and Mitchell (2001). When using this update procedure, the ensemble mean of the archived ensemble forecast at each particular time t_{m+j} does not change since the observation values have been set to $\mathbf{y}^{\text{trial},q} = [H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+1}})^T, \ldots, H^{\text{trial},q}(\overline{\mathbf{x}}^f_{t_{m+L}})^T]^T$.

Note that the ETKF (Bishop et al., 2001) is a particular version of a DEnSRF. Tippett et al. 2003 discuss how numerically equivalent computations of posterior covariance will be obtained for any version of DEnSRF when no covariance localization is being used. Applying the ETKF formalism to condition $[ens_m]$ on $H^{\text{trial},q}$ will yield an expression equivalent to eq. B.1 where the $\mathbf{S}$ matrices in eq. B.1 are the $\mathbf{T}$ matrices given by eq. 18b of Bishop et al. (2001).

Appendix C. Group filter localization details

To understand how the group filter methodology is used to obtain results for the RDA, it is first necessary to understand how it is applied to an ensemble filtering scheme. For some time t_i, let the prior ensemble be $[\mathbf{x}^1_{t_i}, \ldots, \mathbf{x}^K_{t_i}]$ (K is the ensemble size). Without loss of generality, independent observations can be assimilated sequentially (Anderson, 2003), therefore, it suffices to understand how the group filter is applied to assimilate one observation. Let the true model state be $\mathbf{x}_{t_i}$. The expected value of the observation is some function h onto the true model state, namely $h(\mathbf{x}_{t_i})$. The actual observation value is the expected observation plus some error and is denoted by y. As shown in Anderson (2003), individual elements of the state vector $\mathbf{x}_{t_i}$ can be updated independently using information from the observation (without approximation). It suffices to examine the update of the first state variable x_1 with y. The prior ensemble for state variable x_1 is given by $[x^1_{1,t_i}, \ldots, x^K_{1,t_i}]$. An ensemble of observation value estimates are obtained by applying the observation operator h on to the prior ensemble given by, $[h(\mathbf{x}^1_{t_i}), \ldots, h(\mathbf{x}^N_{t_i})]$. Increments to the individual ensemble members in $[x^1_{1,t_i}, \ldots, x^N_{1,t_i}]$ are proportional to the sample covariance between $[x^1_{1,t_i}, \ldots, x^N_{1,t_i}]$ and $[h(\mathbf{x}^1_{t_i}), \ldots, h(\mathbf{x}^N_{t_i})]$ divided by the sample variance for x_1, denoted by $\beta_{x_{1,y}} = \sigma_{x_1,y}/\sigma_{x_1,x_1}$ (Anderson, 2003). Due to limited ensemble size, estimates of $\beta_{x_{1,y}}$ may be contaminated by sampling error. The idea behind localization is to multiply $\beta_{x_{1,y}}$ by a factor between [0, 1] to minimize the expected error in the mean increment.

Using Monte Carlo methods, the group filter methodology deterministically solves for the factor multiplying $\beta_{x_{1,y}}$. The group filter methodology entails running M groups of K-member EAKF assimilations *simultaneously*. When running M groups, M estimates of $\beta_{x_{1,y}}$ are available and denoted by $\beta^1_{x_{1,y}}, \ldots, \beta^M_{x_{1,y}}$. The idea is to solve for a confidence factor α, between [0, 1], which multiplies the M estimates of $\beta_{x_{1,y}}$. This confidence factor

is, in effect, the localization. The confidence factor is the value of alpha that minimizes the following quantity,

$$\sqrt{\sum_{j=1}^{K}\sum_{l=1,l\neq j}^{K}\left(\alpha\beta^{l}_{x_1,y}-\beta^{j}_{x_1,y}\right)^2} \tag{C1}$$

If α is found to be negative, it is set to zero. α is chosen to minimize the expected RMS difference between the increment in a state variable and the increment that would be used if the correct regression factor were used. An advantage in using this method to compute $\mathbf{P}(\mathbf{x}_{t_{m+L}}|H^{\text{trial},q})$ is that the group filter naturally generalizes to the case where observed variables and updated state variables are not valid at the same time.

The archived ensemble forecasts used for the RDA result were obtained by running a long OSSE with $H^{\text{current}} \cdot M = 4$ groups of $K = 20$ ensemble members were run. Retrospective forecasts are generated for all $M = 4$ groups and archived.

The RDA algorithm has been implemented using the group filter localization methodology. This makes use of all $M = 4$ groups of archived ensemble forecasts. All $M = 4$ groups of archived forecasts were updated. To compute the objective function $\Phi(H^{\text{trial},q})$ (eq. (2)) the sample covariance is computed *only* using the first group of posterior/updated $K = 20$ ensemble members. The rationale for doing so is that the results for the OSSEs were computed using a $K = 20$ ensemble. This effectively gives the OSSEs and the RDA an equal number of ensemble members when evaluating the cost function.

References

Anderson, J. L. 1992. Barotropic stationary states and persistent anomalies in the atmosphere. *J, Atmos. Sci.* **49**, 1709–1722.

Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.

Anderson, J. L. 2003. A local least squares framework for ensemble filtering. *Mon. Wea. Rev.* **131**, 634–642.

Anderson, J. L., Wyman, B., Zhang, S. and Hoar, T. 2005. Assimilation of surface pressure observations using an ensemble filter in an idealized global atmospheric prediction system. *J, Atmos. Sci.* **62**, 2925–2938.

Arnold, C. and Dey, C., 1986. Observation system simulation experiments: Past, present and future. *Bull. Am. Meteor. Soc.* **67**, 687–695.

Bishop, C. H., Etherton, B. J. and Majumdar, S. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part 1. Theoretical aspects. *Mon. Wea. Rev.* **129**, 420–436.

Bishop, C. H., Reynolds, C. A. and Tippett, M. K. 2003. Optimization of the Fixed Global Observing Network in a Simple Model. *J. Atmos. Sci.* **60**, 1471–1489.

Burgers, G., van Leeuwen, P. J. and Evensen, G. 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.* **126**, 1719–1724.

Cohn, S. E. 1997. An introduction to estimation theory. *J Meteorol Soc Japan.* **75**, 257–288.

Dabbert, W. F. and co-authors, 1996. Research opportunities from merging atmospheric observing and modeling capabilities. *Bull. Am. Meteor. Soc.* **77**, 305–323.

Daley, R. 1991. *Atmospheric Data Analysis.* Vol. 2. Cambridge University Press, 457 pp.

DelSole, T. 2004. Predictability and information theory, Part I: Measures of predictability. *J, Atmos. Sci.* **61**, 2425–2440.

Emanuel, K. A., Raymond, D., Betts, A., Bosart, L., Bretherton, C. and co-authors, 1995. Report of the first prospectus development team of the US Weather Research Program to NOAA and NSF. *Bull. Am. Meteor. Soc.* **76**, 1194—1208.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys Res.* **99**, 10 143–10 162.

Evensen, G. and van Leeuwen, P. J. 1996. Assimilation of geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model. *Mon. Wea. Rev.* **128**, 1187–1193.

Held, I. and Suarez, M. J. 1994. A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Am. Meteor. Soc.* **75**, 1825–1830.

Houtekamer, P. L. and Mitchell, H. L. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.* **126**, 796–811.

Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.

Khare, S. P. 2004. Observing network design for improved prediction of geophysical fluid flows - analysis of ensemble methods. PhD thesis, Princeton University, 195 pp.

Khare, S. P. and Anderson, J. L. 2006. An examination of ensemble filter based adaptive observation methodologies. *Tellus A* **58**, 179–195.

Kleeman, R., 2002. Measuring dynamical prediction utility using relative entropy. *J, Atmos. Sci.* **59**, 2057–2072.

Lorenz, E. N. 1995. Predictability: A Problem Partly Solve. *Proc. Seminar on Predictability.* **1**, ECMWF, Reading, Berkshire, UK, 1–18.

Majda, A. J. and Timofeyev, I. 2000. Remarkable statistical behavior for truncated Burgers-Hopf dynamics. *Proc. National Acad. Sci.* **97**, 12 413–12 417.

Majumdar, S. J., Bishop, C. H., Etherton, B. J., Szunyogh, I. and 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, S. J., Bishop, C. H., Buizza, R. and Gelaro, R. 2002. A comparison of ensemble transform Kalman filter targeting guidance with ECMWF and NRL total energy singular vector guidance? *Q. J. R. Meteorol. Soc.* **128**, 1–23.

Marshall, J. and Molteni, F. 1993. Toward a dynamical understanding of planetary-scale flow regimes. *J, Atmos. Sci.* **50**, 1792–1818.

Morss, R. E., Miller, K. and Vasil, M. 2004. A systematic economic approach to evaluating public investment in observations for weather forecasting. *Mon. Wea. Rev.* **86**, 181–191.

Schneider, T. and Griffies, S. M. 1999. A conceptual framework for predictability studies. *J. Climate.* **12**, 3133–3155.

Simmons, A. J. and Burridge, D. M. 1981. An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.* **109**, 758–766.

Simmons, A. J., Mureau, R. and Petroliagas, T. 1995. Error growth and predictability estimates for the ECMWF forecasting system. *Q. J. R. Meteorol Soc.* **121**, 1739–1771.

Simmons, A. J. and Hollingsworth, A. 2002. Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol Soc.* **128**, 647–677.

Snyder, C. and Zhang, F. 2003. Assimilation of simulated doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.* **131**, 1663–1667.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2002. Ensemble square-root filters. *Mon. Wea. Rev.* **131**, 1485–1490.

Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1931–1924.

Whitaker, J. S., Compo, G. P., Wei, X. and Hamill, T. M. 2003. Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.* **132**, 1190–1200.