Idealized Adaptive Observation Strategies for Improving Numerical Weather Prediction

REBECCA E. MORSS AND KERRY A. EMANUEL

Program in Atmospheres, Oceans, and Climate, Massachusetts Institute of Technology, Cambridge, Massachusetts

CHRIS SNYDER

Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 19 July 1999, in final form 26 May 2000)

ABSTRACT

Adaptive sampling uses information about individual atmospheric situations to identify regions where additional observations are likely to improve weather forecasts of interest. The observation network could be adapted for a wide range of forecasting goals, and it could be adapted either by allocating existing observations differently or by adding observations from programmable platforms to the existing network. In this study, observing strategies are explored in a simulated idealized system with a three-dimensional quasigeostrophic model and a realistic data assimilation scheme. Using simple error norms, idealized adaptive observations are compared to nonadaptive observations for a range of observation densities.

The results presented show that in this simulated system, the influence of both adaptive and nonadaptive observations depends strongly on the observation density. For sparse observation networks, the simple adaptive strategies tested are beneficial: adaptive observations can, on average, reduce analysis and forecast errors more than the same number of nonadaptive observations, and they can reduce errors by a given amount using fewer observational resources. In contrast, for dense observation networks it is much more difficult to benefit from adapting observations, at least for the data assimilation method used here. The results suggest that the adaptive strategies tested are most effective when the observations are adapted regularly and frequently, giving the data assimilation system as many opportunities as possible to reduce errors as they evolve. They also indicate that ensemble-based estimates of initial condition errors may be useful for adaptive observations. Further study is needed to understand the extent to which the results from this idealized study apply to more complex, more realistic systems.

1. Introduction

For several decades, it has been known that numerical weather forecasts are sensitive to small changes in initial conditions. This means that even if the atmosphere could be modeled perfectly, errors in the initial conditions would grow, leading to forecast errors and forecast failures. The initial conditions for operational numerical weather prediction (NWP) models, called analyses, are based on a combination of short-range numerical forecasts with observations. The observations, where available, constrain the atmospheric state in the forecast model to be as close as possible to the true atmospheric state.

The current observation network is made up of three types of observation platforms. Fixed platforms, such as rawinsonde stations, observe at preselected times and locations, generally near population centers and thus over land. In the data gaps left by the fixed network, there are observations from platforms of opportunity, such as planes and ships, and from remote sensing platforms. At any given time, however, these latter two types of platforms are usually at locations that were selected for reasons other than weather prediction. Observations over the oceans often also have limited areal coverage or vertical resolution.

Because the observation network is inhomogeneous, in any forecast situation there are likely to be regions where information about the initial conditions is particularly important, but where there are insufficient observations. If we could identify these important regions, it might be possible to deploy programmable observation platforms, or reallocate existing platforms, to take data in them and improve the initial conditions. These adaptive (also called targeted) observations could both reduce global analysis errors and help predict important weather phenomena on various temporal and spatial scales.

Corresponding author address: Dr. Rebecca Morss, MMM/ASP, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.
E-mail: morss@ucar.edu

Forecasts will probably benefit little from observations in regions where the initial conditions are already quite accurate or where errors will have only a small effect on forecasts of interest. Thus, effective adaptive observation strategies are likely to incorporate information both about probable analysis errors and about rapidly amplifying forecast errors. In addition to subjectively identifying important features (Snyder 1996), several types of objective adaptive observation strategies have been proposed, including estimating errors in the initial conditions using ensemble spread (e.g., Lorenz and Emanuel 1998), using adjoint techniques to calculate optimal perturbations (singular vectors) or sensitivities of forecasts to small changes in the initial conditions (e.g., Bergot et al. 1999; Gelaro et al. 1999; Palmer et al. 1998), and combining the error probability and error growth criteria in a subspace of ensemble perturbations (e.g., the ensemble transform technique; Bishop and Toth 1999). Berliner et al. (1999) provide mathematical formulations for several aspects of adaptive observation problems.

Current work on adaptive observations includes several approaches, ranging from taking observations to improve weather forecasts in real time to idealized studies with simplified models. In the real atmosphere, adaptive observations have been tested for midlatitude forecasts in several recent field experiments and operational programs, including the Fronts and Atlantic Storm Track Experiment (Emanuel and Langland 1998; Gelaro et al. 1999; Langland et al. 1999a; Szunyogh et al. 1999), the North Pacific Experiment (Langland et al. 1999b; Langland 1999; and Toth et al. 1999), and the Winter Storm Reconnaissance Program (Szunyogh et al. 2000; Toth et al. 2000), and for tropical cyclone forecasts (Langland 1999). Although the results have been encouraging, the influence of the real-world adaptive observations has been mixed. The limited information available in the real world also makes it difficult to interpret the results in detail and to obtain enough good cases to allow one to draw firm conclusions.

To eliminate many of the practical constraints in testing real-world observations, observing system simulation experiments (OSSEs) use a forecast model and a data assimilation system to simulate observations and an analysis and forecast cycle. Using a simplified forecast model for OSSEs both significantly reduces the computational cost and makes understanding the results easier. Thus, several adaptive observations studies have been performed with one-dimensional models (e.g., Lorenz and Emanuel 1998; Hansen and Smith 2000). Testing observation strategies with low-order models is an important first step. Because the models used in these studies have few degrees of freedom and do not require a complex three-dimensional data assimilation system, however, the results may only apply to real NWP in a limited sense.

We use an approach that bridges the gap between the full numerical weather prediction system and the idealized simple model experiments described above: an idealized OSSE setup with a fully three-dimensional, yet simplified, quasigeostrophic forecast model and a realistic data assimilation system. The basic framework for the experiments follows that used by Lorenz and Emanuel (1998) and previous observing system simulation studies (e.g., Jastrow and Halem 1970). With repeatable experiments, simulated observations, and perfect knowledge, we avoid many of the logistics that make improving real-time forecasts difficult; we are limited neither by the number of possible cases nor by our ability to sample those cases well. Later, if desired, some of the idealizations can be relaxed and the importance of the real-world constraints evaluated.

The goal of this study is not to evaluate specific adaptive observation strategies. Instead, we have chosen to address more general aspects of observation networks. The understanding gained can help us develop improved observing and forecasting strategies; it can also help us interpret results from other studies of observing systems. Only two simplified adaptive strategies, based only on error in the initial conditions, are tested, along with two nonadaptive strategies. Most results presented are also aggregated over a large number of cases, in terms of domain- and time-averaged error; the improvement (or degradation) that occurs in any individual forecast is a complex issue that is addressed in a separate work (Morss and Emanuel 2000).

Sections 2 and 3 describe the quasigeostrophic model and three-dimensional variational data assimilation system and why they were chosen. Throughout the study, we consider how the simplified model and the data assimilation system used might affect the results. Section 4 explains the OSSE setup, and section 5 explains how the observation strategies were selected and implemented. In section 6, results are presented from networks with different densities of fixed observations. We do not focus on the details of the fixed observation results but rather use them to develop a framework for studying further changes in observation networks.

In section 7, we compare global adaptive observations to global nonadaptive observations. First, we evaluate whether an idealized adaptive strategy can improve analyses and forecasts on a statistically significant basis, in a fully three-dimensional dynamical system with a realistic data assimilation system. Then, we explore how one might best allocate a limited amount of observational resources in space and time. Section 8 discusses two major limitations of the idealized strategy: imperfect knowledge of errors in the initial conditions, and targeting lead time. In section 9, we compare observing strategies when, instead of being implemented globally, they are added to a preexisting fixed observing network. Finally, we summarize and discuss whether and how the results might extend to more complex forecasting systems.

2. Quasigeostrophic model

The quasigeostrophic (QG) model used is a gridpoint beta-plane channel model, periodic in longitude, developed at the National Center for Atmospheric Research and described in Rotunno and Bao (1996). This multilevel QG model was selected because its large-scale three-dimensional dynamical behavior is similar to that of the real atmosphere, yet its relative simplicity makes a large number of runs computationally feasible. Only an overview of the QG model is given here; for further details, see Rotunno and Bao (1996) and Morss (1999).

The version of the QG model used in this study has constant stratification and a tropopause with fixed height (but varying temperature) at the upper boundary. It is forced by relaxing potential vorticity in the interior and potential temperature at the upper and lower boundaries to a specified zonal mean state; the model has no orography or seasonal cycle. The zonal mean reference state is a Hoskins–West jet (Hoskins and West 1979), with a zonal wind (u) maximum at the tropopause, $u = 0$ at the channel walls, and a corresponding meridional temperature gradient. The maximum zonal wind speed of the reference-state jet is scaled to 60 m s^{-1}, and the relaxation timescale is 20 days. Dissipation includes fourth-order horizontal diffusion and, at the lower boundary, Ekman pumping.

The dimensional domain in this study has a circumference of 16 000 km, a channel width of 8000 km (approximately 70° latitude), and a depth (H) of 9 km. The Rossby radius of deformation ($R_d = NH/f$, where the Brunt–Väisälä frequency $N = 0.011\,293$ s^{-1} and $f = 1 \times 10^{-4}$ s^{-1}) is approximately 1000 km.

The standard-resolution runs have 250-km horizontal grid spacing and 5 vertical levels, with an error doubling time of approximately 2–3.5 days. With 125-km horizontal grid spacing and 8 vertical levels, the error doubling time is 1–2 days. The results in sections 6 and 7 have been tested for the higher model resolution, and they are qualitatively similar to those shown (Morss 1999). It is not computationally feasible, however, to run all experiments for the higher resolution. Therefore, to compensate for the somewhat slow error growth at the standard resolution, some experiments have also been performed with different observation frequencies (sections 6b and 7b).

3. Data assimilation system

The data assimilation system controls how observations are incorporated into the forecast model. Thus, it is a key aspect of any data impact study using a relatively complex forecast model. For the initial evaluation of observation strategies, we implemented a three-dimensional variational (3DVAR) data assimilation scheme. The 3DVAR used here is based on the Spectral Statistical-Interpolation analysis system currently operational at the National Centers for Environmental Prediction in the United States (Parrish and Derber 1992, hereafter PD92) and is similar to the systems operational at several other weather prediction centers. The data assimilation system is described only briefly here; for further detail, see the appendix and Morss (1999).

Although more sophisticated data assimilation systems are currently being developed and implemented, 3DVAR was selected because it has been well tested, it is similar to operational schemes, it is easily adaptable, and it is computationally less intensive than more complex schemes. Relative to more sophisticated schemes, 3DVAR is advantageous because the way it incorporates the observational information is easier to understand. It also serves as a baseline data assimilation system, with which future data assimilation systems can be compared (e.g., Hamill and Snyder 2000).

3DVAR produces an analysis field at each data assimilation time by combining a weighted short-term model forecast (also called the first guess or the background field) with weighted observations. How much weight the analysis gives to the background field relative to the observations is determined by statistics (variances and covariances) provided for both the expected errors in the background field and the expected errors in the observations. The algorithm used is the standard 3DVAR algorithm given in the appendix and derived in PD92 and Morss (1999).

The data assimilation system can easily be modified to include many types of observations at any location. For simplicity, however, all observations in these experiments simulate rawinsondes measuring winds and potential temperature at model grid points and model levels. The observation errors specified in the 3DVAR are assumed to be uncorrelated between different rawinsondes and between wind and temperature observations (Dey and Morone 1985). The variances for wind and temperature observation errors at different levels were adapted from the values in PD92 for rawinsonde observation errors. The vertical correlations between observation errors were obtained from the simple function given in Eq. (3.19) in Bergman (1979). The full matrix of observation error covariances is given in Morss (1999); the sensitivity of the results to the specific values used has been tested and is small.

The background error statistics specified in the 3DVAR control how the data assimilation spreads the observational information to nonobserved locations and variables; they are defined in terms of the QG model variables (interior potential vorticity and boundary potential temperature). As in most other implementations of 3DVAR, these statistics must be simplified to create an algorithm which is computationally feasible. Following PD92, the background error covariances are assumed to be fixed in time, to be diagonal in spectral space, and to have separable vertical and horizontal structures. In addition, only simple vertical correlations are specified. Since streamfunction, and thus wind and

temperature, are determined in the model and the data assimilation algorithm by inverting potential vorticity, even this simple form of background error covariances yields complex vertical covariance structures for the wind and temperature observations.

To allow us to separate the effects of changes in the observation network from the effects of changes in the data assimilation system, in this study we also use the same background error statistics for all observing networks and all data assimilation intervals. The background error statistics can then be optimized for a fixed network and a standard data assimilation interval and used in the 3DVAR for all observing networks. Unlike in operational 3DVAR systems, where the background error is not known and must be estimated, in this idealized system the true background error can be used to develop the background error covariances directly. Thus, to the extent that the assumptions about their structure are valid, the weights in our 3DVAR are correct for a particular network. The appendix provides more detail on how the statistics were generated and shows sample statistics.

The analysis increments produced by this 3DVAR for one observation are qualitatively similar to those shown in PD92 for midlatitudes; an example is shown in the appendix. As in PD92, the background error statistics used here are nearly isotropic, are zonally invariant, and have no knowledge of specific atmospheric dynamical situations. They are, however, similar to those in most currently operational data assimilation systems.

We have tested how this simulated system behaves when several of the assumptions and parameters in the data assimilation system are altered; the results are summarized in section 6 and presented in further detail in Morss (1999). Nonetheless, the data assimilation is a crucial part of how the forecast system responds to any changes in observation networks, and it is not possible to evaluate the importance of all of its aspects. Therefore, potential limitations of the assimilation scheme, and of atmospheric data assimilation in general, are considered when evaluating the results.

4. Experimental design

The observing system simulation experiments (OSSEs) are set up by first defining a ''truth'' solution as an arbitrary state of the QG model described in section 2. An initial ''model'' (also referred to as ''control'') solution is produced by adding random noise to the initial truth state. Then:

1) The forecast (QG) model equations are used to integrate both the model and the truth states forward in time until the next data assimilation period.
2) Observation locations are selected (targeted) according to one of the strategies described in section 5.
3) Simulated rawinsonde observations are constructed by sampling winds and temperature from the truth state.
4) The observational data are assimilated into the model background field to form an analysis.
5) The analysis becomes the new model state in step 1. The analysis can also be integrated forward in time to create forecasts, and ensembles of perturbed forecasts can be generated.

This idealized setup mimics the analysis cycle in real numerical weather prediction, with one major advantage: because the actual true state is known at all times, analyses and forecasts are evaluated using this true state, instead of using limited information from observations or analyses.

In the results shown here, simulated rawinsondes observe wind and temperature at all model levels (including the upper and lower boundaries) at each selected observation location, with random observation errors added. The random observation errors are produced as described in Houtekamer (1993), using a Gaussian distribution and the eigenvectors and eigenvalues of the observation error covariance matrix used in the data assimilation system. The statistics of the observation errors are then the same as those assumed in the data assimilation system. The error structure in the simulated radiosondes is similar to that in Houtekamer (1993); for example, the first eigenvector of the covariance matrix is nearly equivalent barotropic.

To prevent results from being biased by one very ''bad'' observation in a key location, the Gaussian error distribution used to generate the observation errors is truncated at one standard deviation. This simulates a quality control algorithm. The experiments have also been run without errors in the observations, and the results are similar to those shown (Morss 1999).

In this study, the same equations are used to integrate both the truth and model states. This simulates perfect knowledge of the atmospheric dynamic equations. Assuming a perfect model limits forecast errors to result only from errors in the initial conditions and thus simplifies understanding the results. Once the basic interactions between observation networks, data assimilation systems, and forecast models are better understood in the context of a perfect model, adding forecast model error is an important next step.

The standard observation and data assimilation intervals are both 12 h. For each of the standard resolution experiments with a 12-h or shorter interval, the experiment was spun up for 90 days, to equilibrate the model state to the specific observational network. Error statistics were then gathered over a 90-day run. For the experiments with data assimilation intervals of 1 day and longer, the spinup and run times were each 180 days. Each experiment was performed three times, with different initial truth states, different initial model state perturbations, a different set of fixed or random observation locations, and a different set of random errors in

the observations. The spread of results from the three runs suggests the possible spread in the results for each observation density; experiments with error distributions that cannot clearly be separated are considered indistinguishable. Several results are also presented for a single 360-day run following a 30-day spinup period.

5. Observing strategies

In order to compare different types of observation networks, we define three basic strategies for allocating observations: fixed, random, and adaptive. Although there are many ways to adapt observations, in this study only two simplified sample adaptive strategies—one idealized and the other a more realizable approximation to the idealized strategy—are tested. All strategies have approximately the same overall distribution of observation locations.

Unless otherwise stated, observations are allocated globally, all according to the same strategy, and they are allocated at each data assimilation time just prior to taking the observations. In many scenarios, it is unrealistic to have such a short time period between selecting adaptive observation locations and observing at them; this issue, however, is addressed only briefly, in section 8c.

a. Fixed observations

For fixed observations, observation locations are selected randomly prior to the experiment, then left fixed throughout the experiment. This simulates a fixed, inhomogeneously distributed rawinsonde network. Two constraints are placed on the randomly selected fixed observation locations. First, except at observation densities greater than 25% of grid points, observation locations are selected at grid points that are not directly adjacent (i.e., the observation locations are separated by more than 250 km). Second, the overall distribution of the fixed locations is weighted to approximately match the overall distribution of adaptive observation locations. Because the adaptive strategy tends to choose more observation locations near the center of the channel (section 5c), the meridional locations of the fixed observations are sampled from a Gaussian distribution centered at midchannel with a standard deviation of 1500 km.

The observations are not equally spaced, so some regions have fewer observations than others; there are, however, no systematic data voids in these experiments. We have tested a variety of observation spacing and distribution constraints, and except as noted for dense observations in section 6, the specific choices do not affect the results. A sample fixed observation network is shown in Fig. 1a.

b. Random observations

For random observations, observation locations are selected randomly at each targeting time. This strategy is tested in order to differentiate the potential benefits of targeting observations consciously from the benefits of simply moving observation locations. Any effective adaptive strategy should therefore, at minimum, reduce errors more than a random, ''null'' targeting strategy.

The random observation locations are selected according to the same constraints as the fixed observation locations: they are at nonadjacent grid points and are weighted toward the center of the channel. Again, we have tested different constraints for spacing and distributing the observations, and they do not significantly affect the results. A sample random observation network is shown in Fig. 1a.

c. Adaptive observations: General considerations

Adaptive observations are defined as observation locations selected ''intelligently'' according to some strategy at each targeting time. The first strategy tested is idealized, based on the actual error in the initial conditions (which is known in these experiments). Although this strategy cannot be implemented in the real atmosphere, it allows us to compare different configurations of adaptive observations knowing that the strategy is not limited by imperfect knowledge of the true atmospheric state. The second strategy is a semirealistic approximation to the idealized strategy; it estimates errors in the initial conditions from an ensemble of perturbed forecasts.

The only constraint placed on the observation locations selected by the two adaptive strategies is that they be at nonadjacent grid points. We have tested larger minimum observation spacing constraints, and as in the case of fixed observations, except where noted they do not significantly affect the results. Figure 1b shows a sample adaptive observation network at an arbitrary time.

As demonstrated in Fig. 1b, the adaptive strategies tested tend to select observation locations near the center of the channel, in the vicinity of the jet. Thus, as previously described, we constrained the fixed and random strategies to have a similar preference. The adaptively selected locations also tend to cluster near each other to some extent. To assess the importance of the clustering of adaptive observations, we have implemented both the random and adaptive strategies with simulated rawinsondes deployed in clusters of prespecified shapes and sizes, rather than singly; results are discussed in sections 7a and 8.

d. Specific adaptive strategies tested

The adaptive strategies tested here incorporate only information about errors in the initial conditions. Al-

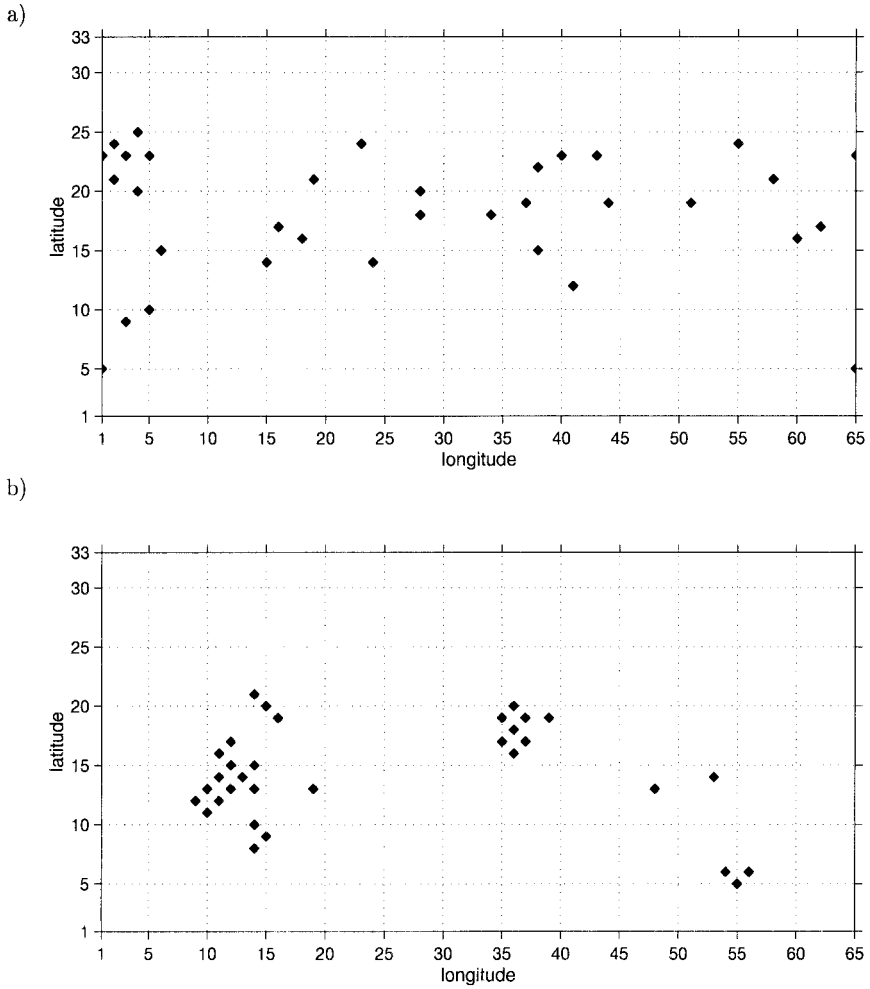

FIG. 1. (a) Set of 32 locations selected by the fixed or random observation strategy at a sample observing time. (b) Set of 32 locations selected by the idealized AER adaptive observation strategy at a sample observing time. Longitude (periodic) is plotted on the *x* axis and latitude on the *y* axis. The grid spacing is 250 km.

though other criteria, such as future error growth, are important when adapting observations, we limited the initial investigation to strategies based on reducing analysis error for several reasons. First, the main purpose of this study is to examine how and when forecast errors can be reduced when an intelligent adaptive strategy is used, not to compare different adaptive strategies. Error in the initial conditions is one of the likely important criteria for targeting observations, and basing a strategy on it is both very simple and computationally quick relative to calculating rapidly growing perturbations (Lorenz and Emanuel 1998). In addition, more complex strategies are likely to place observations at locations with smaller existing errors, where the data assimilation system may not perform as well (since the signal-to-noise ratio in the observations is smaller). Testing strategies based on errors in the initial conditions thus reduces the confusion between the observing network results and the limitations of the data assimilation system. Finally, by estimating the theoretical upper bound of error reduction possible using strategies based on initial condition errors, we provide results with which to compare more complex and more realistic strategies.

1) IDEALIZED ANALYSIS ERROR REDUCTION ADAPTIVE STRATEGY

The idealized analysis error reduction (idealized AER) strategy selects observation locations, prior to each observation time, where the background and truth states differ the most. In the absence of observations, the analysis error is largest where the background error is largest; thus, this strategy assumes that observing at these locations will minimize the analysis error. Reducing the analysis error will ideally then, on average, reduce the forecast error. Note that this strategy does

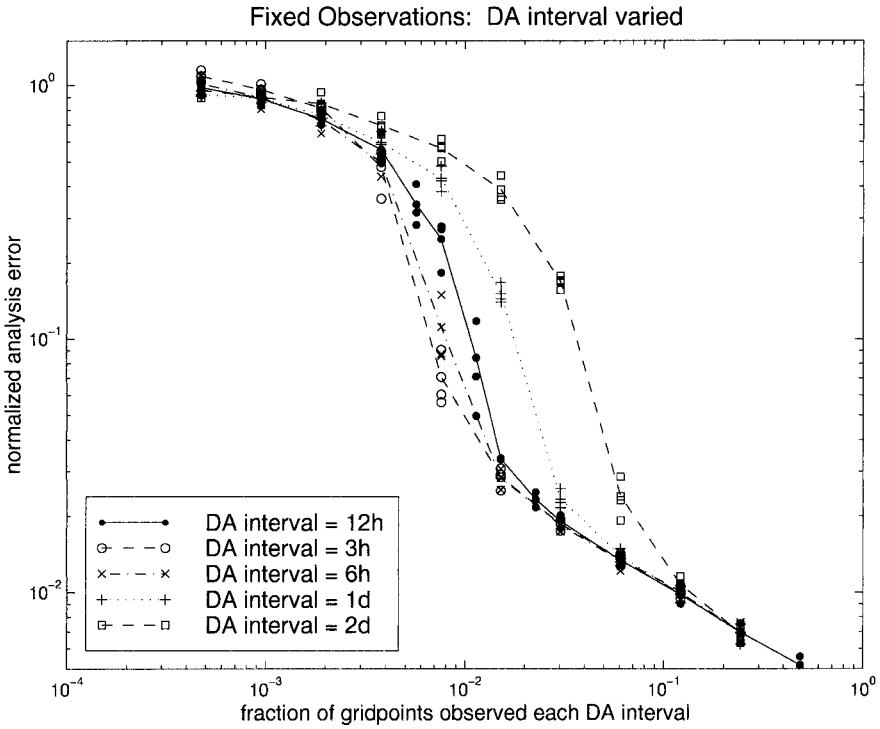

FIG. 2. Domain- and time- root-mean-square-averaged analysis error (streamfunction norm) as a function of the density of fixed observations, for a 3-h, 6-h, 12-h, 1-day, and 2-day data assimilation interval. Three of the symbols plotted for each observation density and data assimilation interval are the root-mean-square-average analysis errors for each of the three separate runs described in section 4; the spread of the symbols suggests the range of errors for each observation density. The fourth symbol and the connecting line are the root-mean-square-average of the three runs. For each of the runs, the *y* axis is normalized by the average error in the absence of observations, that is, the saturation error. The *x* axis is normalized by the maximum number of observation locations (the number of gridpoints in the *x*–*y* plane, 2112) and represents the spatial observation density at each data assimilation time. Here, 1% of grid points observed corresponds to an average of approximately 1 observation every (2000 km)^2 in the dynamically active portion of the domain at every data assimulation time; the Rossby radius of deformation (R_d) $\approx$ 1000 km.

not assess how important any specific part of the initial conditions is for a future forecast, nor does it know how the data assimilation will incorporate the observations. Therefore, it is not guaranteed to select the locations where observations will reduce analysis or forecast error the most.

2) ESTIMATED ANALYSIS ERROR REDUCTION ADAPTIVE STRATEGY

The estimated AER strategy selects observation locations where the background error is estimated, based on the spread of an ensemble of QG model forecasts, to be large. The ensemble in this study is produced by assimilating a different set of perturbed observations, at the same locations, separately into each ensemble member at each analysis time. The observation perturbations are generated by sampling the same observation error distribution described in section 4; the perturbations are added to the errors that are already present in the control observations. This ensemble simulates errors propagating through an analysis and forecast cycle. It is similar to the multiple replication ensemble tested in Lorenz and Emanuel (1998), to the perturbed observation ensemble in Hamill et al. (2000), to the OSSE-MC procedure described in Houtekamer and Derome (1995), and to the ensemble currently operational at the Atmospheric Environment Service in Canada (with an unperturbed forecast model).

6. Average errors for different fixed observation densities

Before testing adaptive observations, it is important to understand how the simulated observing system adjusts to changes in a typical observation network. Figure 2 shows, for a fixed observation network and several data assimilation intervals, how the time- and domain-averaged analysis error decreases as the number of observations increases; note that both axes are logarithmic.

TABLE 1. Means, standard deviations, and maxima of domain-averaged analysis errors and 3-day forecast errors for different fixed observation densities during a 360-day time series (following a 30-day spinup period). The error norm is streamfunction in nondimensional units × 100. Because we have only tested one observation network for each observation density, small changes between results for different observation networks are not statistically significant.

	Analysis error			3-day forecast error		
Observation network	Mean	Std dev	Max	Mean	Std dev	Max
None	40.16	6.74	57.33	40.12	6.74	57.26
4 fixed	30.00	8.24	58.61	34.34	7.71	56.62
8 fixed	23.03	6.19	43.58	30.75	7.80	57.93
16 fixed	7.64	2.47	15.95	15.96	5.20	32.14
32 fixed	1.17	0.27	2.60	2.79	0.97	7.21
64 fixed	0.75	0.13	1.35	1.88	0.58	4.60
128 fixed	0.54	0.07	0.81	1.38	0.43	3.75
256 fixed	0.39	0.05	0.58	1.05	0.30	2.44

In Fig. 2 and subsequent figures, the observation density is normalized by the maximum number of observation locations (2112) and the analysis error is normalized by the error in the absence of observations, that is, the saturation error. As discussed in section 6a, all comparisons shown are similar for forecast errors and for other error norms. For each observation density and data assimilation interval, the symbols plotted represent the results for the three different runs described in section 4. The spread of the symbols suggests the range of results for each observation density; it also indicates how dependent the results can be on minor parameters, even with 90 or more days of cases. In this section, we discuss primarily the results for the standard, 12-h data assimilation interval; the results for different data assimilation intervals are compared in section 6b.

This simulated system has three general regimes in observation density. On the left of Fig. 2, for few observations, the data assimilation system does not have enough information to resolve even the large-scale errors in the initial conditions. It can make only small local improvements that are quickly swamped by the error growth; the errors in the model state remain essentially saturated, and only a small benefit is accumulated from adding a few observations.

As more fixed observations are added, the data assimilation system begins to resolve more synoptic features at more times. In the middle observation density regime, therefore, both the mean error and the time variability in the error decrease dramatically as the observation density increases. This is evident from the symbols and their spread in Fig. 2 and from the results for a 360-day time series in Table 1.

On the right of Fig. 2, for many observations, adding observations again produces only a small additional benefit. Given the errors in the observations, the assumptions in the data assimilation system, and the relatively small error growth at small scales in the forecast model, in the dense observation regime the analysis errors are already small. Not only are the average errors small, but so is the error variability; as the maximum errors in Table 1 show, dense observations also leave no cases with large errors. This dense observation limit was also evident in observing system simulation experiments performed several decades ago (Bengtsson and Gustavsson 1971; Bengtsson and Gustavsson 1972; Morel et al. 1971).

The apparent power-law behavior for dense observations is suggestive of the $n^{-1/2}$ asymptotic behavior expected in the limit that n, the number of observations, is much larger than the number of degrees of freedom in the analysis.[1] Further examination, however, indicates a more subtle relationship between the curves shown in Fig. 2 and this asymptotic result. First, the number of observations is not large compared to the number of degrees of freedom, and fits to the curves yield slopes that are significantly different from ½. In addition, the slopes themselves depend on a number of aspects of the experimental design and the implementation of the data assimilation scheme, such the convergence tolerance for the 3DVAR solver (Morss 1999).

For a 12-h data assimilation interval, the middle observation density regime occurs in the range of an average of 1 fixed observation every $(2000\ \text{km})^2$ in the most dynamically active areas of the domain. With on the order of 1 observation per Rossby radius of deformation (1000 km), then, the data assimilation system receives information that is sufficiently dense in space and time to resolve the dominant features effectively at all times. The error reduction is dominated by observations at synoptic scales not just for streamfunction errors but for all error norms tested (Morss 1999).

One might ask if the distinctive shape of Fig. 2 is a consequence of the experimental design. We have tested modifying many aspects of the experiments and have found that the specific appearance of Fig. 2 is sensitive to changes in the 3DVAR and experimental setup, but the general shape is not. The results for different data

[1] In this limit, there are many observations of each variable (u, v, θ) at each grid point and level. These sets of redundant observations can then, in essence, be replaced by single observations equal to the mean of each set and having expected errors that decrease as $n^{-1/2}$.

assimilation intervals in Fig. 2 are typical: changes can alter the slope of the curve in the three regimes and the range of observation densities occupied by each of the regimes, but both the data sparse and data dense limits remain [further examples and discussion can be found in Morss (1999)]. The only aspects of the experiments that have not been modified in this study and could affect the curve shape are the quasigeostrophic dynamics, the perfect model assumption, and several of the assumptions in the data assimilation system. We expect these to be more important for dense than for sparse observations.

In the remainder of the study, we focus on the nondense (sparse and middle) observation regime for several reasons. First, as shown in Morss (1999), the detailed results for dense observations are sensitive to small changes in the data assimilation system and experimental setup. Also, the results for dense observations are likely to be most affected by the simplified model dynamics, the perfect model assumption, and the 3DVAR data assimilation scheme. Finally, it is outside the dense regime that additional observations have the greatest benefit, and thus that adaptive observations have the greatest potential. An effective adaptive observation strategy will move the middle regime to a lower observation density (to the left) compared to a fixed or random strategy, producing the same average error reduction with fewer observations.

Note that this regime of highest influence was not selected a priori but rather was defined by the system, and that it occurs at a fairly low observation density. In order to better allocate observational resources for NWP, further study is needed to understand when and where different observation density regimes are relevant for operational atmospheric forecast models, data assimilation systems, and observation networks.

a. Sensitivity to error norm

The error depicted in Fig. 2 and subsequent figures is root-mean-square analysis error averaged in time, throughout the domain, and at all levels, with a streamfunction norm. The general curve shape and the comparisons shown are similar for all other analysis error norms tested (including energy, potential vorticity, winds, and temperature), and for the analysis error at individual levels. Results are also qualitatively similar for various time- and domain-averaged forecast error norms. Only the specific slopes and values change, with differences primarily in the data-dense regime (Morss 1999).

When adapting observations, we would like to reduce not only domain-averaged errors but also errors in significant forecast events. We have tested using the forecast error associated with important individual atmospheric systems as an error norm, and on a statistical basis in these experiments, results are similar to those for domain-averaged error norms. This occurs because at any given time, the domain tends to be dominated by a few atmospheric systems. The large analysis and forecast errors tend to be localized near these systems, and reducing errors in these regions is what reduces domain-averaged analysis and forecast errors. Thus, all results are presented for domain-averaged analysis error with a streamfunction norm, but they are similar to results for other error norms.

b. Sensitivity to data assimilation interval

Figure 2 shows the average error for different spatial densities of fixed observations when data are taken and assimilated at different regular intervals. Changing the data assimilation interval explores the effects of incorporating data from fixed locations less or more frequently. It also tests the sensitivity of the results to the ratio between the data input interval and the timescales for advection and error growth. Consistent with the assumption of fixed data assimilation for all observation densities, the 3DVAR data assimilation remains fixed, optimized for a single observation density and a 12-h interval.

The x axis in Fig. 2 is the spatial density of observations at each data assimilation time. For each density along the x axis, as the data assimilation interval is halved (say from 12 to 6 h), the same number of observations is taken twice as often. Intuitively, one might think that twice as much data would produce a better analysis, but Fig. 2 demonstrates that this is not always true. For both sparse and dense observations, for example, observing at fixed locations more frequently improves the analysis only slightly.[2]

In fact, the 3- and 6-h observation interval results are similar not only for sparse and dense observations, but also in the middle observation density regime. Given that the advective timescale (on the order of 12 h) and the error growth timescale (several days) are longer than the interval between observations, this redundancy of frequent fixed observations is expected, since the analysis error at low and middle observational densities is controlled by the incompleteness of the observations rather than their accuracy. With the QG model run at approximately double resolution, so that the advective timescale remains the same but the error growth is faster, fixed observations are redundant on approximately the same timescale (Morss 1999). This suggests that the advective timescale dominates in determining the redundancy of fixed observations. Only in the middle re-

[2] Because of the assumptions used to develop the 3DVAR input statistics, the data assimilation system uses additional data suboptimally, especially in the data-dense regime. Thus, with a better data assimilation system, more frequent observations might be more beneficial. The dense observation results are also likely to be affected by the lack of model error and the unrealistic subsynoptic-scale dynamics.

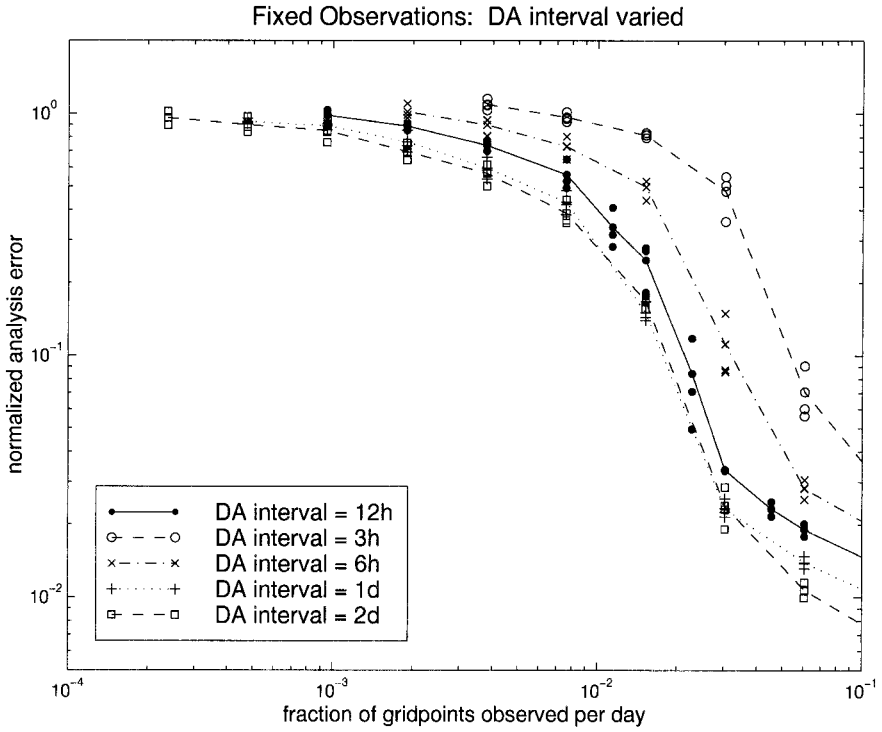

FIG. 3. As in Fig. 2, but with the *x* axis normalized both by the number of grid points and by the number of data assimilation intervals per day. The *x* axis represents the observation density in space and time, that is, allocation of a fixed amount of observational resources.

gime in data density, and only for longer data assimilation intervals, does changing the rate at which data are taken significantly affect the results.

Figure 3 presents the same results as in Fig. 2, but with the *x* axis of each curve normalized by the number of data assimilation times per day (and focused on the regime of interest, nondense observations). The *x* axis is now observation density in space *and* time, and might be thought of as the rate at which observation resources are expended. Each observation density represents a constant amount of data per day but input at different frequencies. Given a certain number of observations each day, we can now ask how frequently we would like to observe, at the expense of observing at fewer locations, to maximize error reduction. In section 7b, the results in Fig. 3 are compared with similar results for random and adaptive observations.

If we are restricted to fixed observation locations, Fig. 3 shows that, for a constant rate of expenditure of observations, the analysis error is reduced most when we sample at more locations less frequently (instead of at fewer locations more frequently). This is as expected, since as described earlier, data gathered frequently at fixed locations is on average redundant; filling in the gaps between the observations is more beneficial than adding observations at the same locations more frequently than the advective timescale.[3] For data assimilation less frequent than every 12–24 h, the fixed observations are no longer redundant, and there is no clear preference for allocating the observations in space and time.

Fundamentally, the results from the fixed observation experiments demonstrate how, for a given forecasting system, the effect of adding observations in space and time can depend strongly on the observation density and the error regime. Understanding that this occurs even for simple observation networks is important for interpreting the results for adaptive observations presented next.

7. Comparison of fixed, random, and idealized adaptive observing strategies

Figure 4 compares the average analysis error as a function of observation density for three types of global

[3] This should not be interpreted to indicate that rawinsondes more frequent than 12 h are not useful in real atmospheric prediction. Further study is needed with a more realistic forecasting system (including a more complex and imperfect forecast model). In addition, more frequent data from fixed platforms may be beneficial for applications not addressed here.

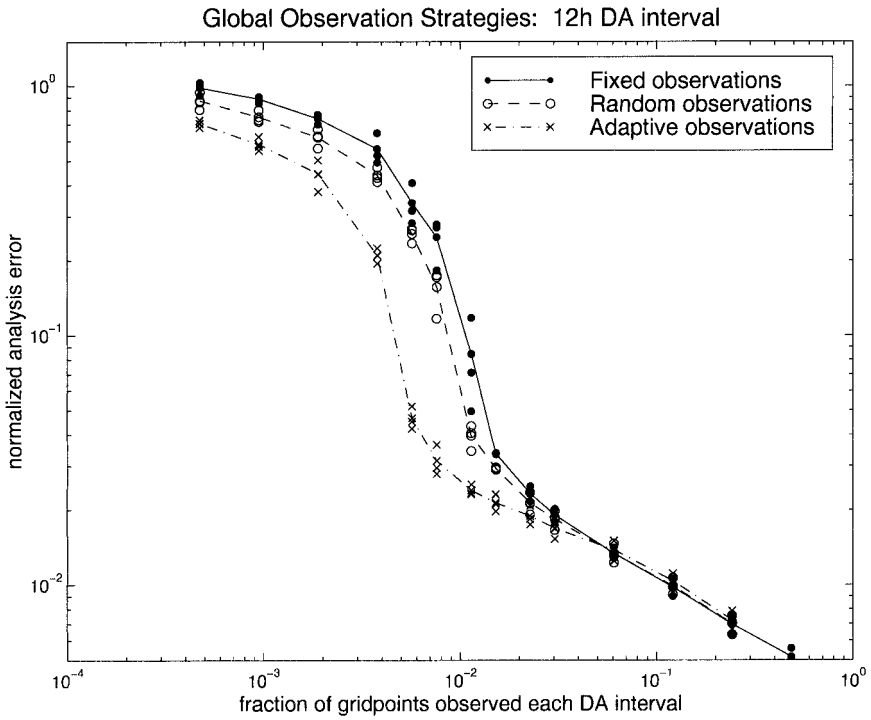

FIG. 4. As in Fig. 2, but for global fixed, random, and idealized AER adaptive observations, all for a 12-h data assimilation interval. The fixed observation results are the same as those for the 12-h data assimilation interval in Fig. 2. The observation strategies are defined in section 5.

observation networks: fixed, random, and idealized AER adaptive (described in section 5). The data assimilation and targeting intervals are 12 h. As in the fixed observation experiments, the data assimilation system is the same for all observation networks. The spread of the symbols again suggests the range of results for each type of observation network. The influence of the observations is measured in terms of analysis error averaged throughout the domain and over many model states. As discussed in section 6a and demonstrated in Morss (1999), comparisons for other time-averaged analysis and forecast error norms are qualitatively similar to those shown.

Above a certain observation density, the three strategies shown in Fig. 4 produce similar results. Thus, for spatially dense observations, this data assimilation and forecast model system has on average little preference among the observation strategies tested. The remainder of the discussion focuses in the spatially nondense data regime, where adding and adapting observations have the greatest potential to reduce errors.

Recall from section 6b that fixed observations are redundant at shorter data assimilation intervals. Thus, for nondense observations taken at a 12-h data assimilation interval, a randomly moving observation network can perform on average slightly better than a fixed network. Adaptive strategies must improve upon both the fixed and random strategies to be considered effective.

As Fig. 4 clearly shows, adaptive observations can, on average, reduce errors significantly more than the same number of random or fixed observations. The idealized AER strategy in Fig. 4 uses energy averaged at all levels as the norm for targeting. Most other targeting norms tested—including background error measured in terms of root-mean-square-averaged potential vorticity (averaged enstrophy), winds, and temperature at all levels, and each of these norms at individual levels—produce similar average improvements. Further tests have revealed that a streamfunction norm is less effective for AER targeting, presumably because it smooths smaller-scale structures.

Although the specific values depend on the details of the experiment and the adaptive strategy tested, Fig. 4 illustrates the range of error reduction we can expect from the idealized AER adaptive strategy for nondense observations. In the example shown, to achieve the rapid drop-off of error, we only need to take about one half as many adaptive observations as we would fixed or random observations. We thus consider the idealized AER adaptive strategy effective for nondense observations. At least in the context of a QG model, then, with good information about the background errors, it is possible to adapt observation networks objectively to

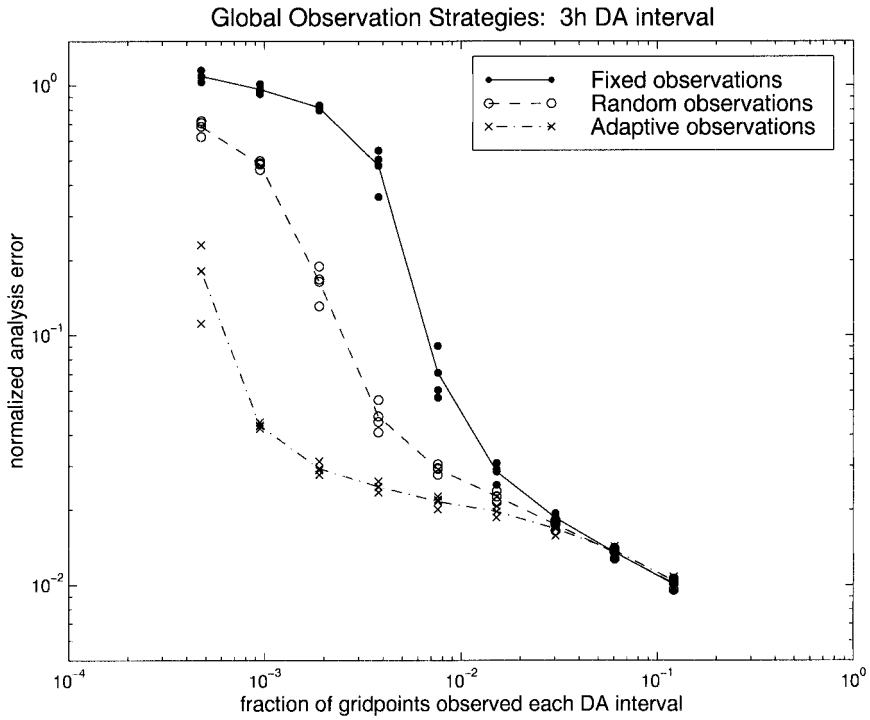

FIG. 5. As in Fig. 4, but for a 3-h data assimilation interval.

reduce average analysis and forecast errors—as long as one is in the appropriate data density and frequency regime.

a. Clustered observations

As discussed in section 6a, at any given time the largest errors tend to focus in a few regions of the domain. Thus, the adaptive strategy tested tends to cluster the observations in one or several groups (e.g., Fig. 1b). To ensure that the different results for the random and idealized adaptive strategies are not caused by their different observation spacing tendencies, we have also tested observing the targeted locations with ''clusters'' of rawinsondes, rather than with single rawinsondes as in previous experiments. For all strategies, in the clustered observation experiments we observe simultaneously at several locations in a prespecified pattern around each targeted location, spacing the targeted locations so that the observation clusters do not overlap. This simulates, for example, one or more aircraft sent with infinite speed to drop several rawinsondes within a specific region.

In general, the adaptive and random strategies compare similarly for single observations and for various-sized and various-shaped observation clusters. The difference between the strategies is therefore not due solely, or even primarily, to how they space the observations.

For less effective AER adaptive strategies or at longer data assimilation intervals, however, the results can be sensitive to whether and how the observations are clustered. For the estimated AER adaptive strategy, for example, observations in prespecified patterns can reduce errors more than the same number of single observations (section 8). Some comparisons among observation clusters of different shapes and sizes are discussed in Morss (1999). Note, however, that the relative effectiveness of different observation clusters is highly dependent on the specific situation and the data assimilation system.

b. Sensitivity to data assimilation interval

Given a certain number of observations, we would like to know not only how to distribute them in space, among strategies and targeted locations, but also how to distribute them in time. For example, if there are eight observation platforms available, each of which can be sent anywhere to observe but only once per day, should they be used all at once or at different times? To address this issue, in this section we compare how the same three observation strategies (fixed, random, and idealized AER adaptive) behave as the time interval between observations is changed. As in section 6b, this also tests how sensitive the results are to the ratio between the data input interval and the model timescales.

The same set of experiments as in Fig. 4 have been performed for data assimilation intervals from 3 h to 2 days; results for a 3-h interval are shown in Fig. 5. As discussed in section 6b, observations taken frequently at fixed locations are on average redundant. Thus, as

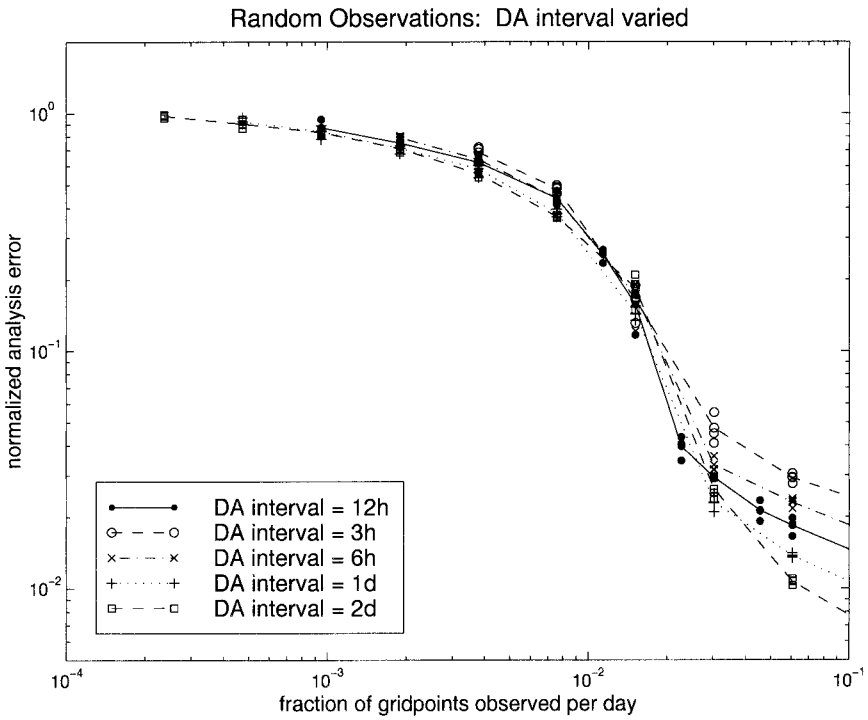

FIG. 6. As in Fig. 3, for random observations. The *x* axis represents observation density in space and time, that is, allocation of a fixed amount of observational resources.

data are taken more frequently, moving observation locations randomly becomes more beneficial compared to keeping them fixed.

Moving observations adaptively also becomes more beneficial as data are taken more frequently, compared to both fixed and random observations. For example, recall from Fig. 4 that to reduce the average analysis error by a factor of about 30 compared to saturation, for a 12-h data assimilation interval we require about twice as many fixed or random observations as we do adaptive observations. For a 3-h interval (Fig. 5), to reduce the errors by the same factor of 30, we require about 4 times as many random observations or about 8 times as many fixed observations as we do adaptive observations.

With less frequent data input (not shown), the random and fixed strategies converge, confirming that observations at fixed locations are not redundant at periods longer than 12 h. The adaptive strategy is also less beneficial when data is taken less frequently, and its influence depends more on the background error norm used to target and the prespecified clustering and spacing of the observations.

For the model at approximately double resolution, where the advective timescale is the same but the error growth is faster, more frequent targeting and observing is generally needed to see the same benefit from adapting observations (Morss 1999). Thus, the error growth timescale is likely more important than the advective timescale when determining the effectiveness of AER adaptive observations.

To compare the strategies at different rates of data input in another way, Figs. 6 and 7 allocate the same total number of random or idealized AER adaptive observations at different data assimilation intervals (as in Fig. 3 for fixed observations). For dense observations, the results are similar to those shown in Fig. 2 and discussed in section 6b for fixed observations; again, we focus on the results for nondense observations.

If observation locations are randomly selected at each time, Fig. 6 indicates that this forecast model and data assimilation system has no preference for how the observations are allocated in time. If observation locations are fixed, there is either no preference for allocation in time or, at shorter data assimilation intervals, errors are reduced the most when observations are taken at more locations less often (Fig. 3). Figure 7, on the other hand, demonstrates that the adaptive strategy performs best when observations are taken at fewer locations more often.

To understand in more detail why the adaptive strategy performs best when taking fewer observations more frequently, we analyzed results from several other experiments. First, we tested moving the adaptive observation locations only at alternate data assimilation times. If data is taken every 12 h, for example, but the adaptive observation locations are only moved every 24 h, the errors are between those for the standard 12-h and the

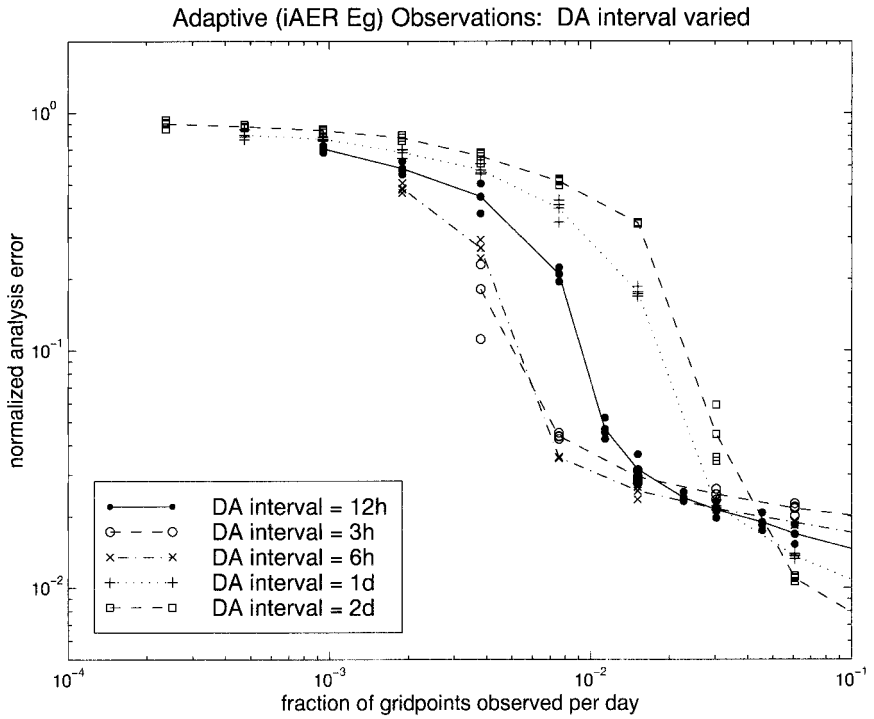

FIG. 7. As in Fig. 3, for idealized AER adaptive observations. The *x* axis represents observation density in space and time, that is, allocation of a fixed amount of observational resources.

standard 24-h adaptive observation and data assimilation experiments (Morss 1999). We also tracked the sequence of adaptive locations selected in many different cases, as synoptic systems and errors evolved. An example is shown in Fig. 8; note how, at some times, the adaptive strategy chooses to observe at locations that are close together or in regions that have recently been observed.

These results suggest that, in this simulated system, the AER adaptive strategy benefits on average both from repeating observation locations and from having the opportunity to choose observation locations more frequently. This is, in large part, due to deficiencies in the data assimilation system; if errors remain large despite previous observations, observing frequently gives the data assimilation more chances to reduce them. Even if

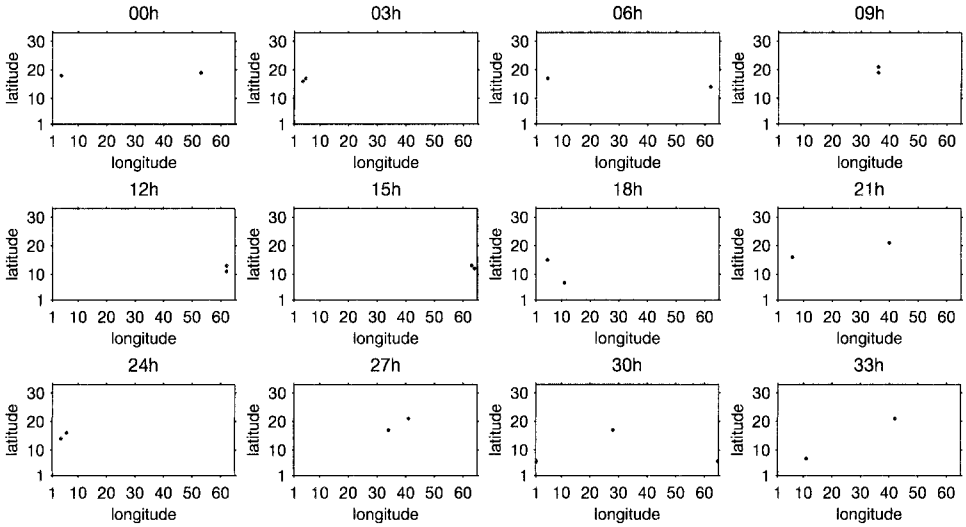

FIG. 8. Sample sequence of adaptive observation locations selected at 3-h intervals over the course of 1.5 days, from an experiment with 2 single idealized AER adaptive observations assimilated every 3 h.

the 3DVAR is able to use the adaptive observations to reduce errors, selecting observation locations frequently gives the adaptive strategy the opportunity to target those errors that continue to grow rapidly. Observing more frequently may also help reduce errors as they begin to grow, before they become very large. Frequent targeting, then, allows the adaptive strategy both to compensate for deficiencies in the assimilation scheme and to control the magnitude of rapidly growing errors.

With a strategy that has more information about future error growth or about how the data assimilation system performs, adaptive observations may not always so clearly prefer to be allocated at more times, rather than at more locations. Nevertheless, observing frequently with any strategy can help compensate for the imperfect data assimilation systems and the imperfect information available in the real world.

8. Adaptive sampling using ensemble spread to estimate background error

The idealized AER adaptive strategy is a useful tool for exploring a large number of adaptive observation configurations. Unfortunately, it is impossible to implement in the real world. Therefore, we next test a more realizable adaptive strategy, one that uses ensemble spread to estimate the locations with the largest background errors. The estimated AER adaptive strategy and the ensemble generation technique are described in section 5d. The ensemble-based strategy tested here is relatively simple, incorporating limited information from the ensemble; almost certainly neither the strategy nor the ensemble formulation is ideal for adapting observations. Nevertheless, it can be used to begin evaluating the potential for estimating background error for adaptive observations.

Observations are assimilated every 12 h, again using the standard 3DVAR data assimilation system. The results shown in this section are for a 13-member ensemble (12 perturbed trajectories in addition to the control); the effects of ensemble size are discussed in section 8b. The results shown are also for observation locations selected using an ensemble of 12-h forecasts generated at the data assimilation time prior to the observation time, that is, with observation locations selected only 12 h before the data are taken. The potential for targeting observation locations at longer lead times is discussed briefly in section 8c.

a. Comparison with fixed, random, and idealized adaptive observing strategies

If a single simulated rawinsonde is deployed at each selected location, the estimated AER adaptive strategy is not an improvement over the fixed or random strategy. However, if observations are taken from a prespecified cluster of rawinsondes at each location (as described in section 7a), the estimated AER adaptive strategy is beneficial. Figure 9 compares the average error for random, estimated AER adaptive, and idealized AER adaptive clustered observations; in this example, the cluster is a triangle of three observations around each location selected by each strategy. The differences among results for the strategies vary with the cluster pattern, but the estimated strategy generally performs on average better than the random strategy but worse than the idealized strategy.

In Fig. 9, the ensemble spread is defined as the average difference between the ensemble members and the mean of the ensemble, with a vertically averaged energy norm (the same norm as for the background error in the idealized AER strategy results shown). As with the different background error norms tested for the idealized strategy, using a root-mean-square potential vorticity norm for estimated AER adaptive observations produces results similar to an energy norm, while a root-mean-square streamfunction norm is less effective. We have tested several other ensemble spread norms, including an extremum energy norm (based on results in Buizza and Palmer 1998), and all produce similar results. We have also found that the estimated adaptive strategy tends to perform better when there are a few fixed observations in the domain than when all observations are taken adaptively (Morss 1999).

To understand why the estimated AER adaptive strategy is more effective for clustered than for single observations, we have compared the ensemble spread to the background error on a case-by-case basis (Morss 1999). As discussed in sections 6.1, at any time there are generally a few regions with large forecast errors, associated with dynamically active synoptic systems. In these same dynamically active regions, the ensemble members diverge, producing large spread. Thus, the ensemble spread and the background error often identify the same general regions. Because the ensemble does not know which specific errors are present in the control forecast initial conditions, however, it estimates background error imperfectly in several ways. First, the actual errors are often large in only one part of a region identified by the ensemble spread as likely to have errors. The ensemble spread thus tends to have less small-scale structure than the actual background error, and it is less focused on specific subregions. In addition, in some cases a region of large ensemble spread is somewhat offset from a region with large background errors. Finally, even when the ensemble spread and the background error identify similar regions, they sometimes prioritize among those regions differently.

Such discrepancies arise whenever the statistics of a random variable (such as the background error) are compared to a specific realization of that variable. Such discrepancies are thus inherent in any practical estimate of background error. Consequently, with imperfect information about background errors, adapting observations in clusters rather than singly may be beneficial

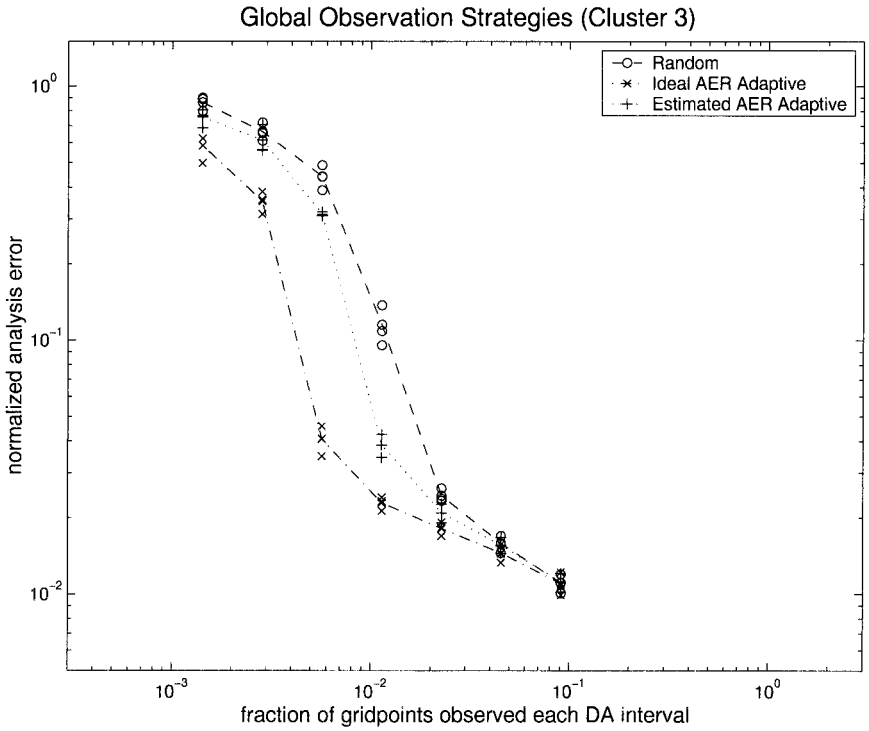

FIG. 9. Average reduction in the domain-averaged error in the control analysis as a function of observation density for random, idealized AER adaptive, and estimated AER adaptive clustered observations. A triangle of three observations is taken around each targeted location for all of the strategies. The observation and data assimilation intervals are 12 h. For the estimated AER strategy, the ensemble has 13 members (12 perturbed forecasts in addition to the control). The *x* axis is normalized both by the number of grid points (2112) and by the number of observations taken for each targeted location (3); the *x* axis is also shifted so that the maximum observation density is 100% × the number of observations for each targeted location.

because it minimizes the effects of the mismatch between background error and its estimate.

Ensembles of perturbed forecasts contain information besides an estimate of background error, and the background error is not the only criterion important for adapting observations. Thus, we could almost certainly use ensembles to develop a more sophisticated, more effective adaptive observation strategy than the one implemented here. As a first attempt at ensemble-based targeting in this idealized system, therefore, the results in Fig. 9 are encouraging.

b. Ensemble size

The results in Fig. 9 are for a 13-member ensemble. We have also performed the same set of experiments with different-sized ensembles; for all of the ensemble and observation configurations tested, having more than 6 perturbed forecasts benefits the estimated AER adaptive strategy little if at all (Morss 1999). Comparisons between background errors and spreads of different-sized ensembles in individual cases suggest that large ensembles do not benefit the estimated AER adaptive strategy because with only a few perturbed members, an ensemble is on average able to identify the few general regions with large background errors (Morss 1999). The specific number of ensemble members required, however, is likely to depend on the simplified dynamics and the simplified geometry of the QG forecast model. In addition, we have only tested simple uses of the ensemble; larger ensembles may be useful for more sophisticated targeting techniques.

c. Targeting lead time

In the real world, many adaptive observing platforms (particularly moving in situ platforms) require that observation locations be selected well in advance of the observation time. In recent field experiments, for example, aircraft constraints required that adaptive observation locations be selected 36–48 h in advance, with preliminary flight planning beginning even earlier. Longer targeting lead times complicate adapting observations for two reasons. First, adaptive observations are likely to be most important in uncertain forecast situations, when analysis and forecast errors make it most difficult to select observation locations in advance. Second, currently most proposed adaptive strategies incor-

porate no information about future, already planned observations. If observations are taken in important regions between the targeting time and the observation time, the adaptive strategy may have overestimated the need for data in those regions. These issues have raised concerns that adaptive strategies, particularly those based on estimating probable errors, are not practical when advance planning is required.

The results in Fig. 9 are for observation locations selected with a 12-h lead time. We have also tested targeting at longer lead times (Morss 1999); the idealized AER strategy then uses errors in a longer lead time control forecast, and the estimated AER strategy uses the spread in a longer lead time ensemble forecast. The results suggest that, as expected, both the idealized and estimated AER adaptive strategies become much less effective as the targeting lead time increases. On average, however, the adaptive strategies are still an improvement over random observations for lead times of up to 1–3 days (depending on the observation configuration).

We have compared actual background errors to longer lead time ensemble spreads and control forecast errors in individual cases. The comparisons suggest that, at longer lead times, the AER adaptive strategies are still able to identify the general regions with large errors but not the specific subregions (Morss 1999). How useful estimates of errors in the initial conditions will be at longer lead times, then, depends on how much detail we require in the estimate and how rapidly our forecast skill decays with time. Estimates of errors at longer lead times may also be more useful if we are able to observe in a larger region.

The longer lead time targeting results are preliminary and are likely to depend on several limitations of these experiments, particularly the perfect model assumption. Thus, further study is needed to determine the extent to which lead time requirements limit the potential of adaptive sampling. Nevertheless, the results from this study suggest that, with a relatively small ensemble, estimates of background error may be useful for targeting observations at lead times of 12 h or longer.

9. Observations added to a preexisting observation network

It is unlikely that all real atmospheric observations will be taken adaptively in the near future. Initially, however, we allocated all observations according to only one strategy at a time so that we could avoid having to consider how fixed and moving observations might interact. Figure 10 compares the global observation strategy results from Fig. 4 with results for idealized adaptive observations in a more realistic scenario: observations added to a preexisting network of fixed observations. The targeted observations are added every 12 h to 16 fixed observations, and they are constrained to be more than 250 km (1 grid spacing) from the preexisting observations. The average error is plotted as a function of the density of all (fixed and moving) observations.

As adaptive observations are added to the network, the results from the mixed observation network rapidly asymptote to the results from the all adaptive network. As Table 2 shows, regularly adding AER adaptive observations to a reasonably sparse fixed observation network not only significantly reduces the mean error more than adding fixed or random observations—it also significantly reduces the day-to-day variability in the error. The regularly added adaptive observations reduce the mean error and the error variability primarily by reducing the error in the situations when the error would otherwise be very large (Fig. 11). Results for the estimated AER adaptive strategy lie between those for the random and idealized AER strategies; again, the estimated strategy is more effective for clustered than for single observations.

Figure 11 also demonstrates how the results for a given strategy can vary significantly over time. Over the full 360-day run and for single rather than clustered observations, the results vary even more dramatically than those shown, and they can do so over time periods of several months or more. Thus, even in this idealized system, many cases and much data are required to definitively distinguish between strategies. This is particularly true for strategies for which the benefit is less apparent.

For adaptive observations added to other preexisting fixed observation densities, the time-averaged results asymptote to the results for an all adaptive network as they do in Fig. 10. For dense observations, the errors for a global adaptive network are only slightly smaller than those for a global nonadaptive network, if at all, so there is much less potential for reducing errors by adding adaptive observations. As we would expect, therefore, for preexisting networks with more than approximately 64 fixed observations (3.0% of grid points), the influence of added fixed, random, and adaptive observations is nearly indistinguishable for both average errors and error variability (shown in Table 3 and in Morss 1999). As discussed in section 6, for dense observation networks there are no cases with large errors, so the added adaptive observations also reduce the maximum error only slightly. Regular adaptive observations may still be beneficial for dense observing networks, but they are likely to require a more specific error norm, a more sophisticated strategy, and a better data assimilation system.

The results demonstrate that in this idealized system, we can reduce the largest domain-averaged analysis and forecast errors simply by observing at every data assimilation time at the few locations with the largest background errors, without explicitly taking future error growth into account. Given the limitations of the data assimilation system, the QG model, and the adaptive strategy, only one or a few observations taken regularly at well-selected locations are needed to improve the

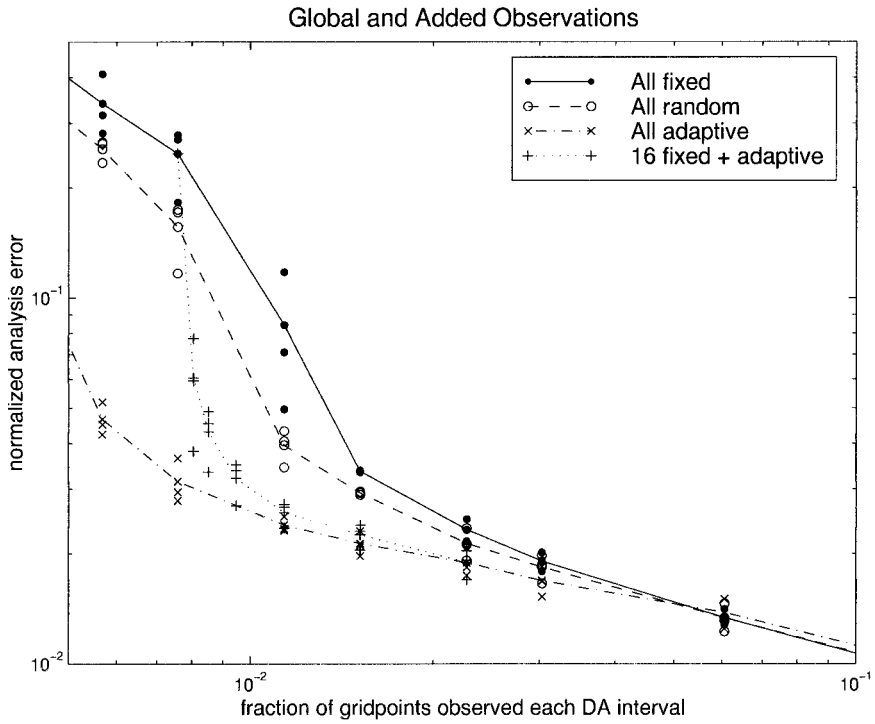

FIG. 10. Analysis error reduction as a function of the total number of observations for single global fixed, random, and idealized AER adaptive observations, and for single adaptive observations added to a preexisting network of 16 fixed observations. All observations are allocated, taken, and assimilated every 12 h. The global observation strategy results are the same as those in Fig. 4 but are shown only for a subset of the observation densities. The added adaptive observation results are plotted as a function of the total observation density, that is, the density of the fixed and adaptive observations together. The axes are magnified to show the regime of interest.

analysis nearly as much as is possible. This suggests that at least for nondense observation networks, regular adaptive observations have significant potential to improve analysis and forecast errors both on average and in cases with large errors.

10. Summary

The results presented show, first of all, that for observation networks that are sufficiently sparse in space, regular adaptive observations can on average improve analyses and forecasts. This has not previously been demonstrated on a statistically significant basis with a three-dimensional forecast model and a realistic data assimilation system. Adaptive observations are beneficial in two ways. First, for a fixed amount of observational resources, adaptive observations can reduce errors more than fixed or random observations. The errors are reduced both on average and in the situations in which the errors would otherwise be largest. Second, if we wish to reduce errors to a certain level, adaptive observations can save observational resources.

Unfortunately, even the most effective adaptive strategy tested in this study is not always beneficial. If there are enough observations to resolve synoptic-scale features reasonably well, adding or redistributing observations according to *any* strategy, adaptive or nonadaptive, improves analyses and forecasts only a small amount. The benefit from modifying the observation network thus depends strongly on the total observation density, and on the preexisting errors. This concept, while not new, is reemphasized by the results shown, and its importance when evaluating observation networks is confirmed.

The type of adaptive strategy tested in this study performs best when selecting a few observation locations frequently, rather than more locations less frequently. When a few adaptive observations are added to a preexisting observation network at every data assimilation time, the results rapidly asymptote to those for a global adaptive network. Thus, adding adaptive observations to a fixed network is on average beneficial (or not beneficial) in the same scenarios as adapting all observations is. This result indicates that regularly adapting only part of an observation network can improve analyses and forecasts.

The adaptive strategies tested are based only on error in the initial conditions, with no information about future error growth. With only a small number of ensemble members, the spread of a multiple replication ensemble

TABLE 2. As in Table 1, for a 360-day run with 1 single observation or 1 cluster of observations added to 16 fixed observations at every data assimilation time according to one of three strategies: random, estimated AER adaptive, or idealized AER adaptive. The observation cluster is 3 observations in a triangle around the selected location, the same cluster as for the results in Fig. 9. The ensemble for the estimated strategy has 13 members. A time series of the errors during part of the run with 1 added cluster of observations is shown in Fig. 11.

Observation network	Analysis error			3-day forecast error		
	Mean	Std dev	Max	Mean	Std dev	Max
16 fixed	7.64	2.47	15.95	15.96	5.20	32.14
17 fixed	5.72	2.61	15.48	12.99	5.56	33.31
16 fixed + 1 random	7.34	2.41	16.05	15.90	4.92	36.77
16 fixed + 1 estimated adaptive	4.27	2.00	12.21	9.95	4.58	31.91
16 fixed + 1 idealized adaptive	1.94	0.57	5.67	4.63	1.74	13.12
16 fixed + 1 cluster random	4.28	2.04	11.97	9.43	4.49	27.17
16 fixed + 1 cluster estimated adaptive	1.96	0.55	5.44	4.58	1.66	13.52
16 fixed + 1 cluster idealized adaptive	1.38	0.31	2.67	3.19	1.09	9.99

provides a useful (but by no means ideal) estimate of initial condition errors for adapting observations. The results suggest, therefore, that it may be possible to adapt observations effectively with the imperfect information available in the real world. Although the experiments in this study are idealized, the success of a simple strategy with a relatively simple data assimilation system makes us optimistic that it may be possible, given the appropriate considerations, to use adaptive sampling to improve atmospheric analyses and weather forecasts.

11. Discussion and future work

In this simulated system, the adaptive strategy tested performs best when observing regularly and frequently. Our experience suggests that this occurs because the adaptive strategy does not know specifically how the data assimilation system will incorporate the data into the model nor how the forecast model will integrate the resulting analysis increment forward in time. Without this knowledge, the adaptive strategy tested in this study, and perhaps any feasible realistic strategy, cannot predict whether analysis or forecast errors will actually be reduced once an observation is taken. More frequent targeting, then, gives the AER adaptive strategy the opportunity to try to reduce large errors as they move and evolve, whether they occur because an atmospheric feature has developed or is developing rapidly or because past observations did not reduce the error sufficiently.

The influence of modifying the observation network depends on how well the data assimilation system and forecast model are able to use the new data. Therefore, as observing platforms, data assimilation systems, and forecast models evolve, the most important considerations for adapting observations and the best observing strategies are likely to change. For example, improving the data assimilation system will make observations more likely to improve analyses and forecasts; frequent targeting may then not be as important. An improved data assimilation is also likely to improve the benefit of adding and adapting observations for denser observation networks. Similarly, targeting less frequently or for denser observations might be more effective using a more sophisticated adaptive strategy, one which incorporates information about future error growth and/or about the data assimilation system. However, both data assimilation and the criteria for selecting observations locations are statistical procedures, so one can never guarantee that a particular observation, even if near per-

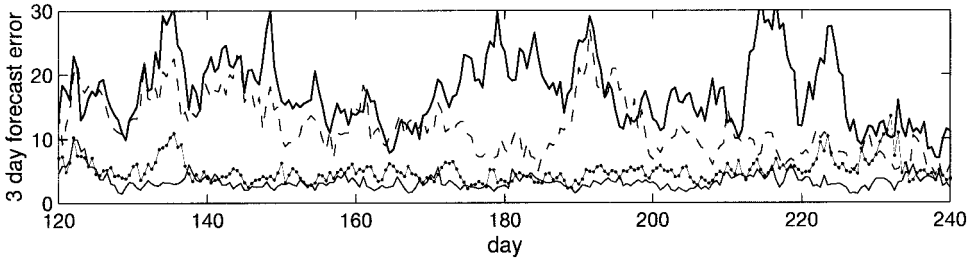

FIG. 11. Domain-averaged 3-day forecast error plotted every 12 h during one 120-day period for 16 fixed observations (thick solid line), 16 fixed observations + 1 random observation cluster (dashed), 16 fixed observations + 1 estimated AER adaptive observation cluster (thin solid with circles), and 16 fixed observations + 1 idealized AER adaptive observation cluster (medium solid). The observation cluster is a triangle of three observations around each selected location. The ensemble for the estimated strategy has 13 members. All observation strategies are implemented at each data assimilation time, every 12 h. The model state was spun up prior to $t = 0$ for each observation network. The error scale is streamfunction in nondimensional units $\times 100$.

TABLE 3. As in Tables 1 and 2, for 2 single observations added to 64 fixed observations at every data assimilation time. Again, small changes are not statistically significant.

Observation network	Analysis error			3-day forecast error		
	Mean	Std dev	Max	Mean	Std dev	Max
64 fixed	0.75	0.13	1.35	1.88	0.58	4.60
66 fixed	0.74	0.12	1.28	1.79	0.56	4.31
64 fixed + 2 random	0.75	0.12	1.31	1.85	0.58	5.10
64 fixed + 2 idealized adaptive	0.69	0.11	1.18	1.64	0.47	3.84

fect, will improve the forecast of interest. Thus, adaptive observing networks are likely, in general, to be most beneficial when they are designed with the strengths and weaknesses of data assimilation and forecasting systems in mind.

How to best estimate initial condition errors for adaptive observations depends on how well the data assimilation scheme can use observational data in different situations. It is almost certain that an ensemble can be used to estimate initial condition errors more effectively than our first attempt does. For example, the data assimilation system used in this study has more difficulty correcting small background errors than it does medium or large errors; we might therefore want to minimize the risk of observing at locations with small errors by not significantly overestimating background error. Before we can develop more useful estimates of initial condition errors, however, further work is needed to understand which aspects of initial condition errors are the most important to estimate well (and not to estimate poorly) for adapting observations.

Several aspects of our simulated system must be generalized before the results can be applied to real NWP. First, we have assumed a perfect forecast model; the model dynamical equations are the same as the true dynamical equations, and the truth contains no subgrid-scale processes that are not resolved by the model. Second, the forecast model used for the OSSEs contains only quasigeostrophic dynamics. These two assumptions are likely to be most important when modeling subsynoptic-scale structures, that is, in our data-dense regime. They are also likely to modify the governing spatial scales and timescales, including the influence of observations at specific observation densities and the results for forecast errors. In addition, we have studied the influence of observation networks only on a control forecast, rather than on an ensemble of forecasts, and we have tested only simplified strategies and idealized observing platforms. Therefore, further study is needed in a more complex and more realistic system to investigate how the results from this study extend to real observing networks.

Acknowledgments. This work was supported by the National Science Foundation under Grant ATM-9634239.

APPENDIX

3DVAR Data Assimilation System: Further Detail

Given M observations and N analysis variables at any time, we can define

$\mathbf{L}$ = an $(M \times N)$ operator that transforms from the analysis variables and locations to the observation variables and locations.

From the observations ($\mathbf{y}_o$) and a background (first guess) field ($\mathbf{x}_b$), we can then calculate

$\mathbf{y}$ = an M-component vector of observation residuals ($=\mathbf{y}_o - \mathbf{L}\mathbf{x}_b$, the difference between observations and the background field at observation locations).

To assimilate the observations, we solve for

$\mathbf{x}$ = an N-component vector of analysis increments (at analysis locations)

by inverting

$$[\mathbf{I} + \mathbf{B}\mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}\mathbf{L}]\mathbf{x} = \mathbf{B}\mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}\mathbf{y}, \quad \text{(A1)}$$

where

$\mathbf{B}$ = the $N \times N$ matrix of covariances between the background errors (called the background error covariance matrix) and

$\mathbf{O} + \mathbf{F}$ = the $M \times M$ observation error and representativeness covariance matrix.

Equation (A1) is derived in Morss (1999) and PD92. At each assimilation time, the 3DVAR solves Eq. (A1) for the analysis increments $\mathbf{x}$ using an iterative conjugate residual solver (Morss 1999; the CR2 scheme in Smolarkiewicz and Margolin 1994). The analysis is generated by adding the analysis increment to the background field ($\mathbf{x}_a = \mathbf{x}_b + \mathbf{x}$); it then becomes the initial conditions for the next model run.

To generate the background error statistics for this study, we used the experimental setup described in section 4 to accumulate 12-h forecast error statistics over a large number of runs with a 12-h data assimilation interval and different distributions of 32 fixed simulated rawinsondes (observations at approximately 1.5% of grid points). We then inserted the output statistics back into the 3DVAR and accumulated new statistics, iterating sev-

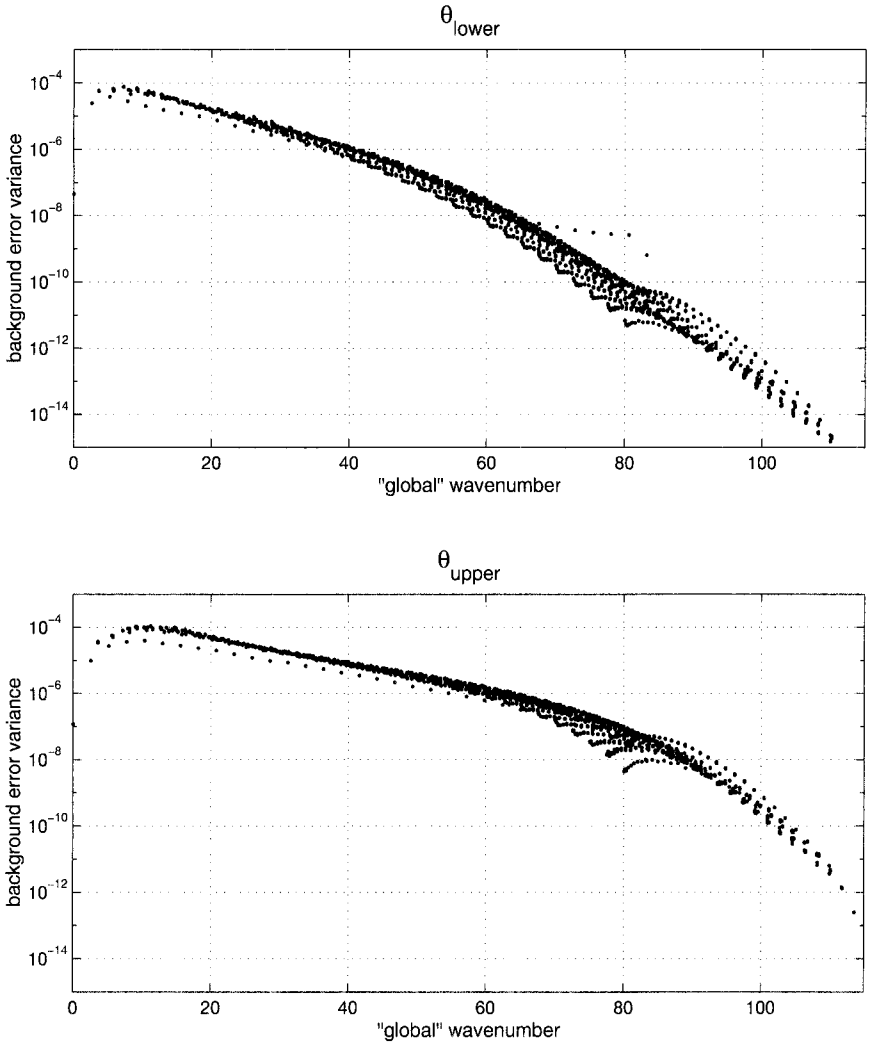

FIG. A1. Horizontal background error variances (diagonal elements of **B**) in the standard-resolution 3DVAR as a function of approximate global wavenumber, for nondimensionalized (a) lower-boundary potential temperature and (b) upper-boundary potential temperature. The statistics were calculated as described in section 3 and the appendix. Global wavenumber is not well defined in the channel model, and it is shown here only to suggest how the error variances compare at different spatial scales. For zonal wavenumber k and meridional half-wavenumber l as represented in the quasigeostrophic model, we approximate global wavenumber as $[(2.5 \times k)^2 + (5.2 \times 0.5 \times l)^2]^{1/2}$. The factors of 2.5 and 5.2 are included to adjust for the zonal and meridional extent of the channel model compared to the ''real'' globe.

eral times. Sample background error covariance statistics are shown in Fig. A1. Morss (1999) evaluates the sensitivity of the results to some of the assumptions made to develop the background error statistics.

Using these background error covariances, the 3DVAR interpolates a single observation to produce analysis increments similar to those depicted in Fig. A2, which are for one zonal wind observation at the middle model level. Because the analysis variables are interior potential vorticity and boundary potential temperature, the 3DVAR infers (from **B** and **L**) that the zonal wind measurement must have resulted from a dipole in potential vorticity (Fig. A2d), spread in the vertical (not shown). This potential vorticity dipole has wind and temperature structures associated with it (Figs. A2a–c), which are recovered using the operator **L**. Since the 3DVAR has no information about the specific atmospheric structure that the observation came from, the observation has been interpolated nearly isotropically.

These analysis increments depict the background error correlations in physical space; one can infer from Fig. A2 the approximate correlation length scale in the 3DVAR with the current statistics. For more than one observation (i.e., observations of more than one variable, at more locations, and/or at more levels), the analysis increments become more complex, and it is more difficult

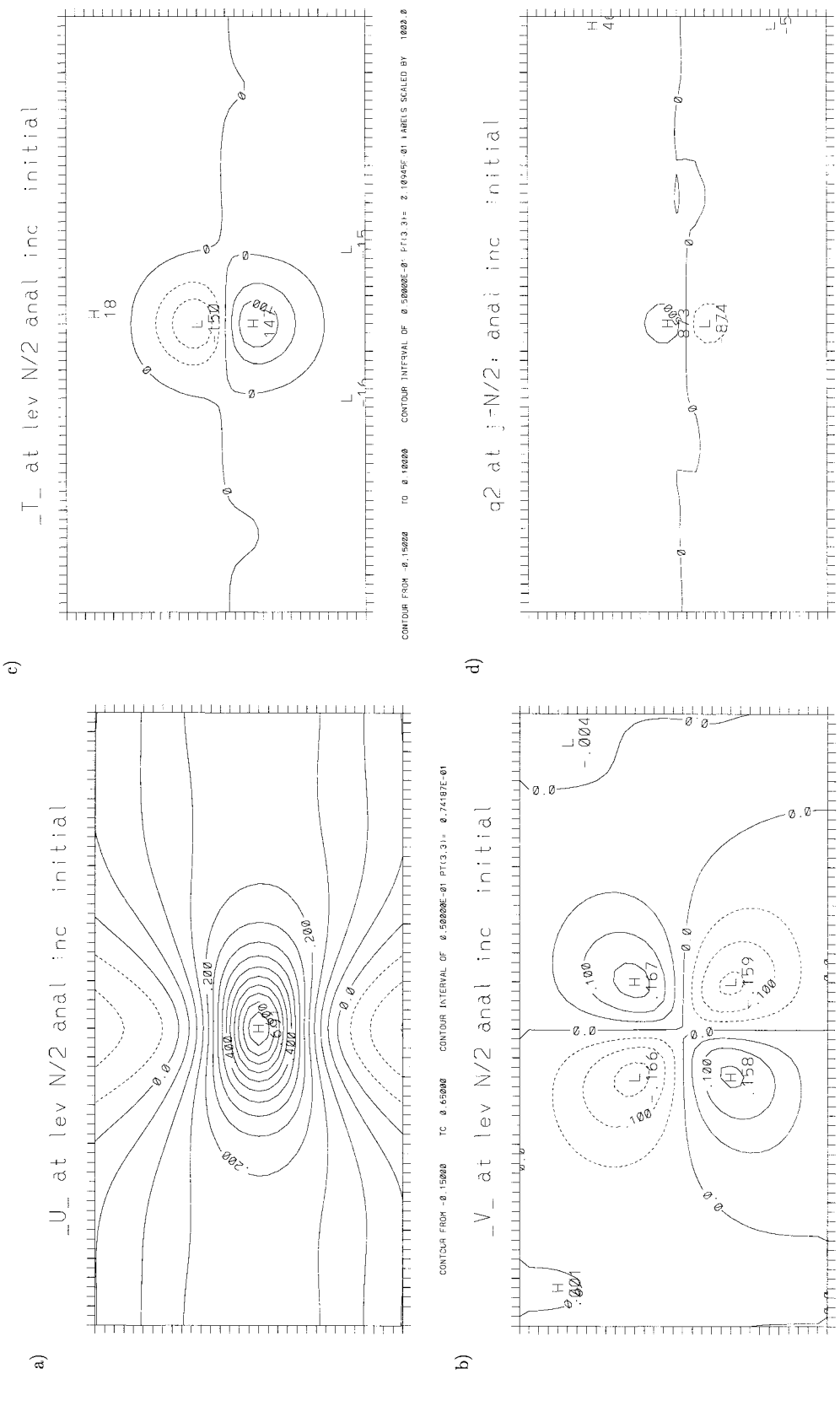

FIG. A2. 3DVAR analysis increments for a single zonal wind observation $u = 1$ at level 3 (the middle level) in the QG model at standard resolution. The variables plotted are nondimensionalized: (a) zonal wind, (b) meridional wind, (c) temperature, and (d) potential vorticity, all at level 3. Longitude is plotted on the x axis and latitude on the y axis, each with 250-km grid spacing. Positive contours are solid lines and negative contours are dashed lines; all contour intervals are 0.05.

to interpret how the data assimilation uses each individual observation. For further detail on the data assimilation system and its behavior, see Morss (1999).

REFERENCES

Bengtsson, L., and N. Gustavsson, 1971: Assimilation of data in dynamical analyses. *Tellus,* **23,** 328–336.

——, and ——, 1972: Assimilation of non-synoptic observations. *Tellus,* **24,** 383–399.

Bergman, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.,* **107,** 1423–1444.

Bergot, T., G. Hello, A. Joly, and S. Marlardel, 1999: Adaptive observations: A feasibility study. *Mon. Wea. Rev.,* **127,** 743–765.

Berliner, L. M., Z.-Q. Lu, and C. Snyder, 1999: Statistical design for adaptive weather observations. *J. Atmos. Sci.,* **56,** 2536–2552.

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

Buizza, R., and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.,* **126,** 2503–2518.

Dey, C. H., and L. L. Morone, 1985: Evolution of the National Meteorological Center global data assimilation system: January 1982–December 1983. *Mon. Wea. Rev.,* **113,** 304–318.

Emanuel, K. A., and R. Langland, 1998: FASTEX adaptive observations workshop. *Bull. Amer. Meteor. Soc.,* **79,** 1915–1919.

Gelaro, R., R. Langland, G. D. Rohaly, and T. E. Rosmond, 1999: An assessment of the singular vector approach to targeted observing using the FASTEX dataset. *Quart. J. Roy. Meteor. Soc.,* **125,** 3299–3328.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter—3D-variational analysis scheme. *Mon. Wea. Rev.,* **128,** 2905–2919.

——, ——, and R. E. Morss, 2000: Error characteristics of bred, singular vector, and perturbed observation ensemble forecasts. *Mon. Wea. Rev.,* **128,** 1835–1851.

Hansen, J. A., and L. A. Smith, 2000: The role of operational constraints in selecting supplementary observations. *J. Atmos. Sci.,* **57,** 2859–2871.

Hoskins, B. J., and N. V. West, 1979: Baroclinic waves and frontogenesis. Part II: Uniform potential vorticity jet flows—Cold and warm fronts. *J. Atmos. Sci.,* **36,** 1663–1680.

Houtekamer, P. L., 1993: Global and local skill forecasts. *Mon. Wea. Rev.,* **121,** 1834–1846.

——, and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.,* **123,** 2181–2196.

Jastrow, R., and M. Halem, 1970: Simulation studies related to GARP. *Bull. Amer. Meteor. Soc.,* **51,** 490–513.

Langland, R. H., 1999: Workshop on targeted observations for extratropical and tropical forecasting. *Bull. Amer. Meteor. Soc.,* **80,** 2331–2338.

——, R. Gelaro, G. D. Rohaly, and M. A. Shapiro, 1999a: Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP 17 and IOP 18. *Quart. J. Roy. Meteor. Soc.,* **125,** 3241–3270.

——, and Coauthors, 1999b: The North Pacific Experiment (NORPEX-98): Targeted observations for improved North American weather forecasts. *Bull. Amer. Meteor. Soc.,* **80,** 1363–1384.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.,* **55,** 399–414.

Morel, P., G. Lefevre, and G. Rabreau, 1971: On initialization and non-synoptic data assimilation. *Tellus,* **23,** 197–206.

Morss, R. E., 1999: Adaptive observations: Idealized sampling strategies for improving numerical weather prediction. Ph.D. thesis, Massachusetts Institute of Technology, 225 pp. [Available from UMI Dissertation Services, 300 N. Zeeb Rd., Ann Arbor, MI 48106.]

——, and K. A. Emanuel, 2000: Influence of added observations on analysis and forecast errors. *Quart. J. Roy. Meteor. Soc.,* in press.

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

Parrish, D. F., and J. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Rotunno, R., and J.-W. Bao, 1996: A case study of cyclogenesis using a model hierarchy. *Mon. Wea. Rev.,* **124,** 1051–1066.

Smolarkiewicz, P. K., and L. G. Margolin, 1994: Variational solver for elliptic problems in atmospheric flows. *Appl. Math. Comput. Sci.,* **4,** 527–551.

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

Szunyogh, I., Z. Toth, K. A. Emanuel, C. Bishop, J. Woolen, T. Marchok, R. Morss, and C. Snyder, 1999: Ensemble-based targeting experiments during FASTEX: The impact of dropsonde data from the Lear jet. *Quart. J. Roy. Meteor. Soc.,* **125,** 3189–3218.

——, ——, R. E. Morss, S. Majumdar, B. Etherton, and C. H. Bishop, 2000: The effect of targeted dropsonde observations during the 1999 Winter Storm Reconnaissance Program. *Mon. Wea. Rev.,* **128,** 3520–3537.

Toth, Z., I. Szunyogh, S. Majumdar, R. Morss, C. Bishop, and S. Lord, 1999: Ensemble-based targeted observations during NORPEX. Preprints, *Third Symp. on Integrated Observing Systems,* Dallas, TX, Amer. Meteor. Soc., 74–81.

——, and Coauthors, 2000: Targeted observations at NCEP: Toward an operational implementation. Preprints, *Fourth Symp. on Integrated Observing Systems,* Long Beach, CA, Amer. Meteor. Soc., 186–193.