Assimilation of Surface Pressure Observations Using an Ensemble Filter in an Idealized Global Atmospheric Prediction System

JEFFREY L. ANDERSON

National Center for Atmospheric Research, Boulder, Colorado

BRUCE WYMAN AND SHAOQING ZHANG

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

TIMOTHY HOAR

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 1 June 2004, in final form 24 January 2005)

ABSTRACT

An ensemble filter data assimilation system is tested in a perfect model setting using a low resolution Held–Suarez configuration of an atmospheric GCM. The assimilation system is able to reconstruct details of the model's state at all levels when only observations of surface pressure (PS) are available. The impacts of varying the spatial density and temporal frequency of PS observations are examined. The error of the ensemble mean assimilation prior estimate appears to saturate at some point as the number of PS observations available once every 24 h is increased. However, increasing the frequency with which PS observations are available from a fixed network of 1800 randomly located stations results in an apparently unbounded decrease in the assimilation's prior error for both PS and all other model state variables. The error reduces smoothly as a function of observation frequency except for a band with observation periods around 4 h. Assimilated states are found to display enhanced amplitude high-frequency gravity wave oscillations when observations are taken once every few hours, and this adversely impacts the assimilation quality. Assimilations of only surface temperature and only surface wind components are also examined.

The results indicate that, in a perfect model context, ensemble filters are able to extract surprising amounts of information from observations of only a small portion of a model's spatial domain. This suggests that most of the remaining challenges for ensemble filter assimilation are confined to problems such as model error, observation representativeness error, and unknown instrument error characteristics that are outside the scope of perfect model experiments. While it is dangerous to extrapolate from these simple experiments to operational atmospheric assimilation, the results also suggest that exploring the frequency with which observations are used for assimilation may lead to significant enhancements to assimilated state estimates.

1. Introduction

The use of ensemble (Kalman) filters for data assimilation applications in the atmospheric and oceanic sciences is growing very rapidly. Ensemble filters are often easier to implement than most traditional data assimilation methods and can produce reasonably good results. This has led to a large number of both idealized applications (Evensen and van Leeuwen 1996; Snyder and Zhang 2003) and attempts to build operational ensemble assimilation/prediction systems in realistic models with real data (Houtekamer et al. 2005; Keppenne and Rienecker 2002; Reichle et al. 2002).

Here, the ability of ensemble filters to extract information from a limited set of observations in idealized situations is highlighted by performing perfect model experiments in a simple atmospheric GCM. In such cases, ensemble filters are capable of extracting information about all state variables from observations of only a localized portion of the state (Daley 1992; Dee 1995). While similar capabilities are expected from four-dimensional variational algorithms (Le Dimet and

Corresponding author address: Jeffrey Anderson, NCAR, P.O. Box 3000, Boulder, CO 80307-3000.
E-mail: jla@ucar.edu

© 2005 American Meteorological Society

Talagrand 1986), such algorithms require the still difficult and time-consuming process of developing adjoints for both the forecast model and the forward observation operators. Three-dimensional variational algorithms require accurate specifications of the background error characteristics; these are extremely difficult to develop for multivariate applications that are essential when only a limited portion of the state variables is well observed (Lorenc 1981; Hamill and Snyder 2000; Hamill et al. 2002; Wu et al. 2002). In addition, three-dimensional variational applications may still require the development of adjoints for forward observation operators.

The idealized experiments here examine the capabilities of ensemble filters with only 20 members in perfect model situations. Unfortunately, real assimilation applications include a number of additional challenges (Mitchell et al. 2002), and ensemble methods that are competitive or superior to existing variational algorithms have proved challenging to develop (Houtekamer et al. 2005). This suggests that future research on ensemble filters should focus on dealing with complications not found in the results presented here including model systematic error, dealing with non-Gaussian observational error characteristics, dealing with model balances, and understanding the interaction of filter sampling errors with these additional error sources.

The experiments described also provide clues about the information content, in the perfect model context, of a variety of surface observing configurations. While extreme caution must be used in extrapolating these results to real applications, there are some hints about the relative value of increasing spatial versus temporal density of observations and about the information available from different types of surface observations (Zhang et al. 2004b, manuscript submitted to *Mon. Wea. Rev.*).

The present work begins by describing the ensemble filter used for assimilation and the simple general circulation model used for perfect model experiments. The experimental design is outlined in section 4. Section 5 presents results for a variety of experiments in which only synthetic surface pressure (PS) observations are assimilated. Sections 6 and 7 examine the assimilation of other surface variables and the impact of having regionally confined PS observations.

2. Ensemble filter

The ensemble filter used here was first described in the literature as an ensemble adjustment Kalman filter (Anderson 2001). Here, it is referred to as an ensemble adjustment filter (EAF) and is one of a class of deterministic square root ensemble filters (Tippett et al. 2003).

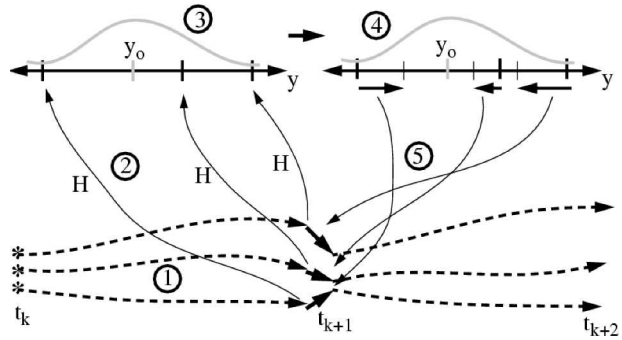

FIG. 1. Schematic description of the least squares framework for sequential ensemble filters for an idealized three-member ensemble case. Please refer to the detailed discussion of this algorithmic representation in section 2 of the text.

As pointed out in Anderson (2003), for observations with uncorrelated error distributions it is possible to describe the operation of a variety of ensemble filters, including the EAF, by describing only the impact of a single scalar observation on a single state variable. Figure 1 shows a schematic of a five-step implementation for a variety of ensemble (Kalman) filter algorithms. In step 1, an ensemble of model state estimates is integrated forward from the time of the previous observation, t_k, to the time of the next available observation(s), t_{k+1}. Prior estimates of the state of an observation, y, are computed by applying the appropriate forward observation operator, H, to each ensemble member in step 2. Step 3 obtains the corresponding value of y from an instrument and an estimate of the instrument's error distribution including representativeness and other errors; these are indicated by the light tick mark and light distribution curve in step 3 of Fig. 1. The information from the prior and observed estimates of y can be combined according to Bayes' theorem by any of a number of algorithms in step 4. The result is an ensemble of estimates of the updated (posterior) state for y and corresponding increments for the prior ensemble estimates, indicated by the dark vectors in step 4. Differences between most variants of ensemble filters that have been described in the literature are confined to the details of the computation of step 4 (Evensen 1994; Burgers et al. 1998; Pham 2001; Whitaker and Hamill 2002). Finally, in step 5, each ensemble member for each state variable is updated by doing a linear regression of the increments for y onto each state variable in turn. This regression is done using the prior ensemble sample of the joint distribution of y and a state variable. This process is then repeated for each additional obser-

vation that is available at time t_{k+1} before the ensemble is advanced to the time of the next observation set.

The details of step 4 for the EAF used here are depicted in Fig. 2 for an idealized three-member ensemble whose prior estimates are indicated by the tick marks on the top axis. The prior sample mean (asterisk on the top axis) and variance (indicated by the solid distribution) are computed and it is assumed that the observing instrument has given a value y_o (indicated by the o on the top axis in the figure) and has a Gaussian error distribution (the dashed distribution on the top axis) with variance Σ_o. An estimate of the updated (posterior) variance, Σ_u, and mean, $\bar{y}_u$, for y is computed by

$$\Sigma_u = (\Sigma_p^{-1} + \Sigma_o^{-1})^{-1}, \tag{1}$$

$$\bar{y}_u = \Sigma_u(\Sigma_p^{-1}\,\bar{y}_p + \Sigma_o^{-1}\,y_o). \tag{2}$$

The updated mean, $\bar{y}_u$, is indicated in the figure by the x on the second and third axes while the updated variance is represented by the dashed distribution on the lowest axis. The prior ensemble is translated (small arrows between top and middle axis tick marks) and linearly compacted (small arrows between middle and bottom axis tick marks) to give a new ensemble with sample mean and variance identical to that computed from (1) and (2). The increment for a given ensemble member y_i is

$$\Delta y_i = (\bar{y}_u - y_i) + A(y_i - \bar{y}_p), \tag{3}$$

$$A = \sqrt{(\Sigma_u/\Sigma_p)}. \tag{4}$$

Applications with small ensembles and large model state vector dimensions have traditionally required several heuristic modifications to the EAF algorithm. First, ensemble filters often produce prior distributions with insufficient ensemble variance that can eventually lead to filter divergence. A heuristic method of adding additional variance such as covariance inflation in which the prior state estimates are linearly expanded about the mean by a constant factor (Anderson and Anderson 1999) is often used. In the results presented here, no covariance inflation is applied in any results.

A second heuristic modification involves limiting the impact of observations to some set of nearby state variables. This localization was developed in terms of a Hadamard product in the ensemble filter based algorithms of Houtekamer and Mitchell (Houtekamer and Mitchell 1998; Mitchell and Houtekamer 2000; Hamill et al. 2001). When ensemble filters are implemented using the sequential framework outlined here, localization is implemented by simply multiplying the regression coefficients computed in step 5 of the filter algorithm by a function of the distance between an observation and a state variable.

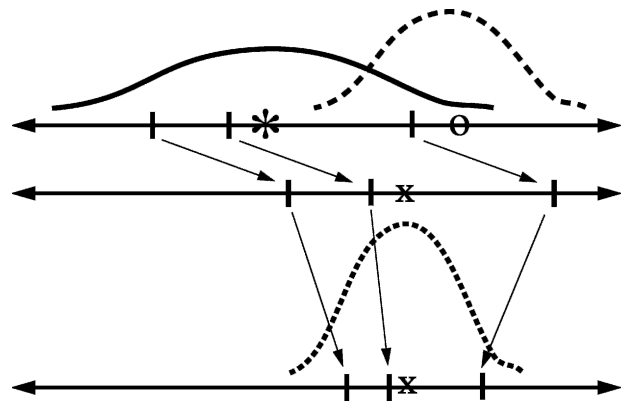

FIG. 2. Schematic description of the observation space ensemble adjustment filter update step used in the ensemble assimilations. Please refer to the detailed discussion of this algorithmic representation in section 2 of the text.

The compactly supported fifth-order polynomial approximation of a Gaussian developed by Gaspari and Cohn (1999) is used for localization here. The distance between observations and state variables is defined as the horizontal distance on the surface of the sphere between the two (there is no notion of a distance in the vertical or a distance between observations and state variables of different types). The Gaspari–Cohn function used is specified by a half-width in radians so that the impact of an observation goes to zero if it is more than twice this many radians from a state variable. The half-width used was developed by a heuristic tuning in a control experiment (section 4) and then was held fixed throughout all additional experiments. The results in all additional experiments would generally be improved if tuning of the half-width were performed independently in each experiment. Localization is required to limit the impacts of sampling error in ensemble filters and more robust ways of defining this have been developed (Anderson 2004, manuscript submitted to *Mon. Wea. Rev.*).

3. Model description

The model is a modified version of the dynamical core used in the Geophysical Fluid Dynamics Laboratory (GFDL) global atmospheric model (Anderson et al. 2004) with forcing and dissipation provided by the GCM benchmark calculation proposed by Held and Suarez (1994) and has evolved from the E-grid dynamical core described in Wyman (1996). The forcing is applied with Newtonian damping toward a zonally symmetric state and simple Rayleigh damping is applied near the surface for dissipation.

The hydrostatic, B-grid dynamical core prognostic variables are the zonal and meridional wind components, temperature, and PS. The model uses a two-level time-differencing scheme. Gravity waves are integrated using the forward–backward scheme (Mesinger 1977) and a split time-differencing scheme (Gadd 1978) is used for a longer advective time step. The gravity wave and advective time step are 400 and 1200 s, respectively, and the forcing is applied every 3600 s. For those runs that require a time step less than 3600 s, the gravity wave and advective time steps are set to one ninth and one third of the time step, respectively.

The B-grid dynamical core uses the vertical discretization and pressure gradient described by Simmons and Burridge (1981). The horizontal and vertical advection of the momentum uses fourth-order, centered spatial differencing, while temperature advection uses second-order differencing. Grid point noise and the 2 delta-x computational mode of the B-grid are controlled with linear fourth order horizontal diffusion. A second-order Shapiro (1970) filter is applied to momentum and temperature at the top level in order to damp eddies and reduce the reflection of waves. Fourier filtering is applied poleward of 60° latitude to damp the shortest resolvable waves so that a longer time step can be taken. The filter is applied to the mass divergence, horizontal omega-alpha term, horizontal advective tendency of temperature and tracers, and the momentum components.

The B-grid model was configured with the minimum horizontal and vertical resolution required in order to generate baroclinic instability and a time mean state that somewhat resembles the observed climate. A grid with 30 latitudes, 60 longitudes and five levels spaced equally in pressure in the vertical was found to be sufficient, resulting in a total of 28 200 state variables. The model's 1-h time step was used in all experiments unless the frequency of observations is higher than 1 h; in such cases the time step is reduced to be equal to the observational frequency.

Figure 3 shows a time mean cross section in height and latitude of the zonal velocity field from a long equilibrated run of the model with the grid superposed. Figure 4 shows a snapshot of the T field at level 3, also with the grid marked. Eastward moving waves in midlatitudes and westward moving waves in the Tropics characterize the climatological behavior of the model. Table 1 includes the climatological standard deviation of the PS, temperature, and u-wind component state variables at the various vertical levels to serve as a baseline for assimilation results.

The highest level (level 1) of the model has some idiosyncratic behavior in the Held–Suarez configuration. The temperature at this level equilibrates very slowly when the model is spun up from a state of rest (it can take years or even decades to equilibrate while all other variables appear to be mostly equilibrated after a year). This same very slow equilibration can impact the results for level 1 in assimilation experiments.

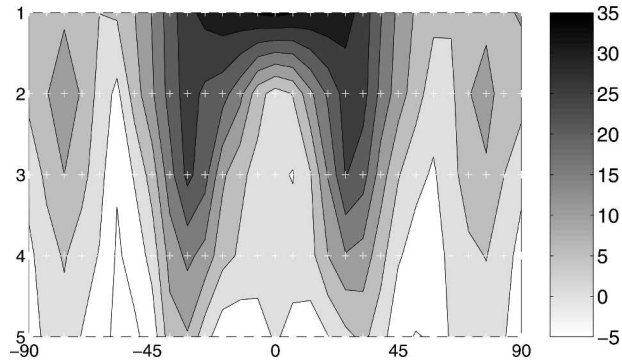

FIG. 3. Zonal mean cross section of time mean zonal velocity from a control integration of the dynamical core; contours every 5 m s^{-1} from 0 to 30 m s^{-1}. Model grid points are marked by the white +'s. The vertical coordinate is model level, with levels (bottom) 5 and (top) 1, and the remaining levels equally spaced in pressure.

4. Experimental design

Perfect model experiments are used throughout this study to evaluate the capabilities of ensemble filters and to explore upper bounds on the information content of surface observations. In perfect model experiments, a very long integration of a numerical model is used to simulate the atmosphere and is referred to as the truth. Synthetic observations of the truth are generated by applying a forward observation operator, H, to the model state and then adding a random sample from a specified observational error distribution to the

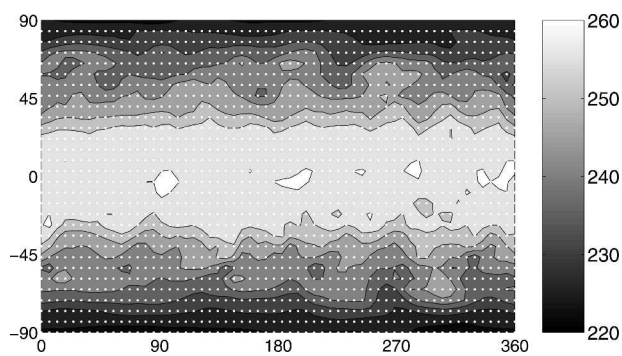

FIG. 4. Instantaneous temperature field at level 3 from a long control integration of the dynamical core; contours every 5 K from 220 to 260 K. Model grid points are marked by white dots.

TABLE 1. Time mean climatological global mean standard deviation and time mean ensemble mean prior rms error from base case (1800 PS observations every 24 h), enhanced spatial density case (28 800 PS observations every 24 h), and high frequency case (1800 PS observations every 5 min) for surface pressure, and T and u at levels 2 through 5.

Experiment	PS (hPa)	T2 (K)	T3 (K)	T4 (K)	T5 (K)	U2 (m s^{-1})	U3 (m s^{-1})	U4 (m s^{-1})	U5 (m s^{-1})
Climotology	7.10	2.42	2.65	3.22	4.22	7.86	8.58	6.42	6.28
Base	0.545	0.272	0.351	0.364	0.456	0.951	1.19	1.02	0.791
28 800	0.364	0.225	0.286	0.287	0.325	0.900	1.03	0.832	0.618
5 min	0.019	0.0066	0.0074	0.0092	0.013	0.0157	0.0131	0.00956	0.0128

result. The observation operator used here is horizontal spatial bilinear interpolation. Observational error distributions are normal with mean 0 and a specified standard deviation. In perfect model experiments, assimilations are performed using the same model that generated the truth, but the only pieces of information available from the truth integration are the values of the synthetic observations. The resulting assimilated model state estimates can be compared to the truth (impossible for the real problem of interest) in order to evaluate the quality of the assimilation algorithm for a given set of observations.

A single perfect model integration is used for the truth in all experiments with assimilation frequency of 1 h or greater. The GCM is first integrated for 100 yr from a state of rest. At the end of the 100-yr integration, 20 perturbed states are generated by adding a random number selected from a normal distribution with mean 0 and standard deviation 10^{-7} to each model state variable. The truth integration and each of these 20 ensemble members are then advanced for an additional 10 yr. The result is a set of 20 ensemble members that can be viewed as random samples selected from the model's climatological distribution. This 20-member ensemble provides the initial condition for all assimilation experiments described here. Starting ensemble assimilations from a climatological distribution of this type is much safer than starting from ensembles that are generated only by adding random noise to the true state. It can be dangerously easy to generate apparently successful assimilation results from poor filters in the latter case.

The truth is integrated for an additional 400 days at the end of the 110 yr and synthetic observations are generated from this time series. Assimilations for these 400 days are generated and the first 200 days of each are discarded to eliminate the impacts of the large errors that occur due to starting with a climatological ensemble distribution. Examining global mean rms error curves by eye suggests that the assimilations have generally asymptoted to stable values after about 100 days of assimilation. All summary results are for the second 200 days of the 400-day assimilation period. All assimilation results shown in the following are for the ensemble mean prior (first guess) estimates generated by the filter. For each of the 30-, 15-, and 5-min assimilation frequency cases, a separate 400-day truth integration is made with the same initial conditions but with a model time step equal to the assimilation frequency.

5. Assimilation of surface pressure observations

This section examines the quality of assimilations that use only PS observations. These experiments are designed to evaluate the capabilities of this rudimentary ensemble filter and to provide a cursory look at the information available from different configurations of PS observations. Whitaker et al. (2004) examine the assimilation of only PS observations in a real model with real observations.

a. Base observing network

The first case assimilates PS observations taken once every 24 h from a network of 1800 PS observing stations. The stations are randomly located on the surface of the sphere, all stations observe simultaneously once a day, and all observations have an error standard deviation of 1 hPa. The forward operator for PS observations is a simple bilinear interpolation from the model grid.

This base configuration is used to tune the value of the half-width of the Gaspari–Cohn localization envelope (section 2). The assimilation is run for localization half-widths of 0.1, 0.15, 0.2, 0.25, 0.3, and 0.35 rad, and the rms error of the assimilations is evaluated. The 0.2 rad case produced the smallest rms errors over the second 200 days of the assimilation for all state variables and all levels and this value is used for all other experiments described here.

The prior rms errors of this assimilation for PS, temperature and u-wind component are found in Table 1, which shows that they are approximately an order of magnitude less than the climatological standard deviations. The rms error for PS is 0.545 hPa, about half the observational error standard deviation. Figure 5 plots

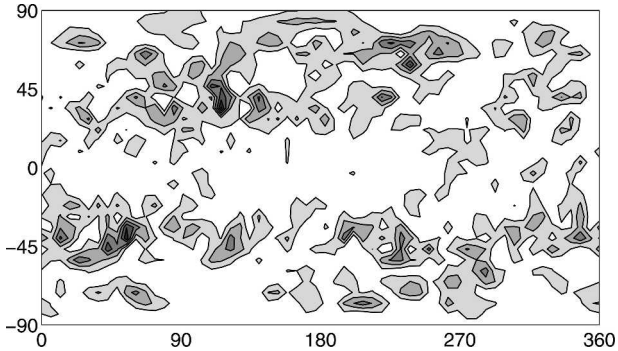

FIG. 5. Absolute value of prior ensemble mean surface pressure error at day 400 of an assimilation using observations from 1800 randomly located surface pressure stations observing once every 24 h with a 1.0-hPa error standard deviation. Contours are every 0.4 hPa from 0.4 to 2.8 hPa.

the absolute value of the prior ensemble mean error of PS at day 400 from this assimilation. The largest errors are confined to the midlatitudes and are similar in scale to the synoptic variability of the model while the smallest errors are found near the equator and appear to have a somewhat larger zonal scale.

Although only PS is observed, Table 1 shows that the prior ensemble mean errors for all other variables in the free atmosphere are also greatly reduced from their climatological standard deviation. Figure 6 displays the error of temperature at level 3 at day 400. For temperature in the middle of the atmosphere, the largest errors are found in the Tropics and appear to have a somewhat smaller spatial scale than was found for the PS errors. Outside of the Tropics, the largest T errors are considerably smaller and appear to have a larger spatial scale, exactly the opposite of the PS errors.

b. Varying observation spatial density

Next, the impact of changing the number of PS observing stations is examined. Assimilation cases with 150, 300, 450, 900, 1800 (the case just discussed), 3600, 7200, 14 400, and 28 800 randomly located PS observations with observational error standard deviation of 1 hPa taken once every 24 h have been assessed.

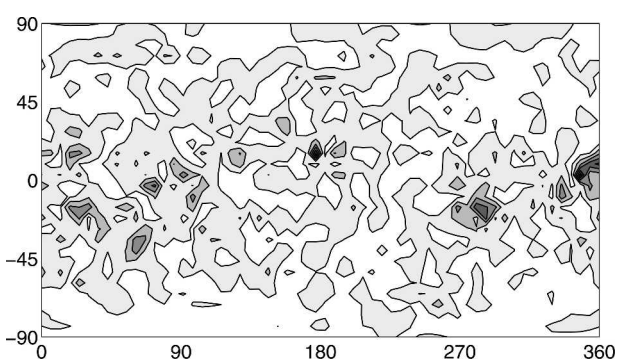

FIG. 6. As in Fig. 5 except for temperature field at level 3. Contours are every 0.4 K from 0.2 to 2.2 K.

Figure 7 plots the global rms error of the prior ensemble mean of PS as a function of the number of PS observations. The case with only 150 observations is able to reduce the prior error variance below the climatological variance but the error is much larger than the observational error level. As the number of observations is increased, the error levels fall rapidly up to about 1800 observations (the control case). As the number of observations is further increased, there is little additional reduction in the prior error up to 28 800 PS observations.

Figure 8 displays the impact of increasing the number of PS observing stations reporting every 24 h on the temperature prior error at each model level. Although temperature is not being observed directly, the curves in Fig. 8 are qualitatively similar to the PS error in Fig. 7. The error falls rapidly as the number of observing stations is increased to 1800 and then decreases slowly as additional stations are added. The T errors are reduced by an additional 10% to 20% when 28 800 observations are used. Table 1 contains the global mean time mean rms error values for PS, temperature, and u-wind component for the 28 800 observation case.

Morss et al. (2001) examined the impact of increasing observational density in a quasigeostrophic channel model using a 3D-variational assimilation method. They observed all model state variables in a column and located their columns of observations at model grid points. They produced curves that had slopes close to a curve with the form $\alpha n^{-1/2}$, the asymptotic behavior expected for large numbers of observations, n. The thick straight lines in Figs. 7 and 8 have this slope. For relatively sparse observations, the slope of the assimilated results is somewhat steeper than expected, again consistent with Morss et al. (2001). However, for very dense observations, it appears that the slope from the assimilations is too shallow. A possible explanation is that as the number of observations becomes very large, other sources of error because of approximations in the filtering algorithm may begin to dominate the assimilation error.

c. Varying frequency of observations

Next, the impact of varying the frequency with which observations are available from a fixed set of 1800 PS observing stations is examined. Cases with observations available every 24 (the base case), 12, 6, 4, 3, 2, 1 h, and every 30, 15, and 5 min are available. Figure 9 displays the prior global mean rms error for PS as a function of

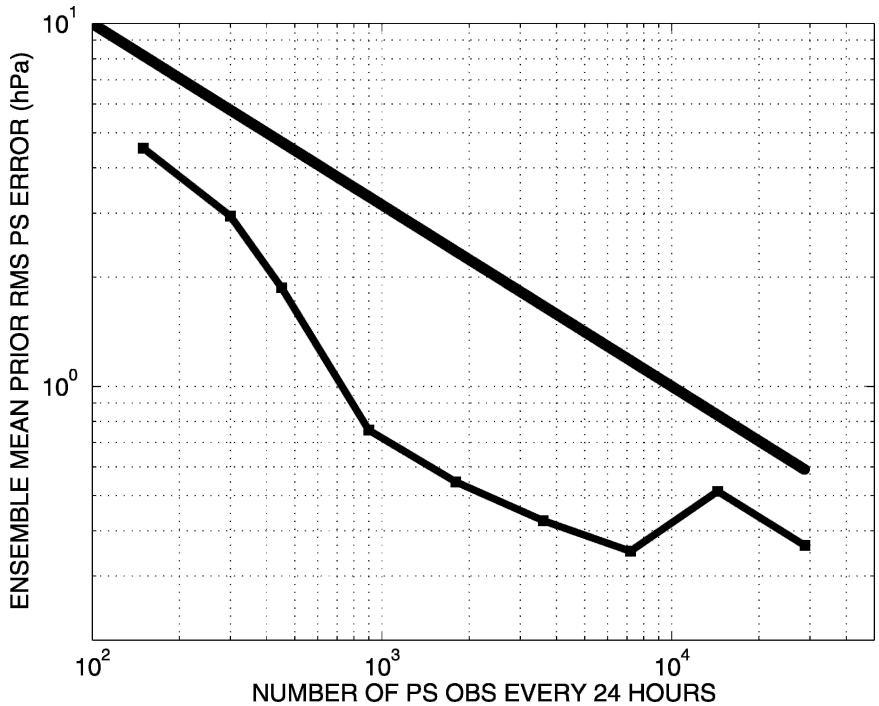

FIG. 7. Log–log plot of global mean rms error of the prior ensemble mean surface pressure as a function of the number of randomly located surface pressure stations observing once every 24 h with a 1.0-hPa error standard deviation. Cases plotted are for 150, 300, 450, 900, 1800, 3600, 7200, 14 400, and 28 800 stations. The thick straight line is a line of the form $\alpha n^{-1/2}$ where n is the number of PS observations and α is a constant.

observation frequency. The error is reduced as the observation frequency is increased and the rate of the decrease in error increases as the frequency increases. The curve is generally fairly smooth except for frequencies between about 2 and 6 h. For very high frequency observations, the prior error becomes very small with a global mean prior rms of only 0.019 hPa for observations every 5 min (Table 1).

Figure 10 displays the global mean prior temperature errors for all five levels as a function of observation frequency. Even though temperature is not observed directly, the error behavior is very similar to that for PS. Error reduces uniformly as observation frequency is reduced except in the vicinity of the 4-h observation period. The prior errors for temperature in the free atmosphere are reduced to approximately 0.01 K for observations assimilated every 5 min. There is no evidence that this error reduction is saturating, suggesting that prior errors for all state variables can be made arbitrarily small by increasing the frequency with which PS is observed. Results are qualitatively similar for the other state variables, the u- and v-wind components. Table 1 includes the rms errors for PS, temperature, and u-wind component for the 5-min case.

Morss et al. (2001) also examined the impact of observation frequency for a temporally fixed observation set. They found that there was relatively little benefit from increasing the observation frequency below 12 h, which is inconsistent with the results here. This suggests that deficiencies in their 3D-variational assimilation implementation may have resulted in an inability to exploit information from more frequent observations. In particular, their choice of using the same background error statistics in all cases may have negatively impacted the performance of their assimilation algorithm. The filter results here show that the prior correlation structure between observations and state variables changes significantly as the frequency of observations is increased.

d. Discussion

It is important to understand the internal error growth characteristics of the global model to be certain that the assimilation results are meaningful. Figure 11 plots the rms error for T at all five levels as a function of forecast lead time for an ensemble forecast initiated at the end of the 400-day assimilation for 5-min observation frequency. The error growth is nearly exponential for about 50 days before asymptoting at the level of the climatological variability. This is consistent with an

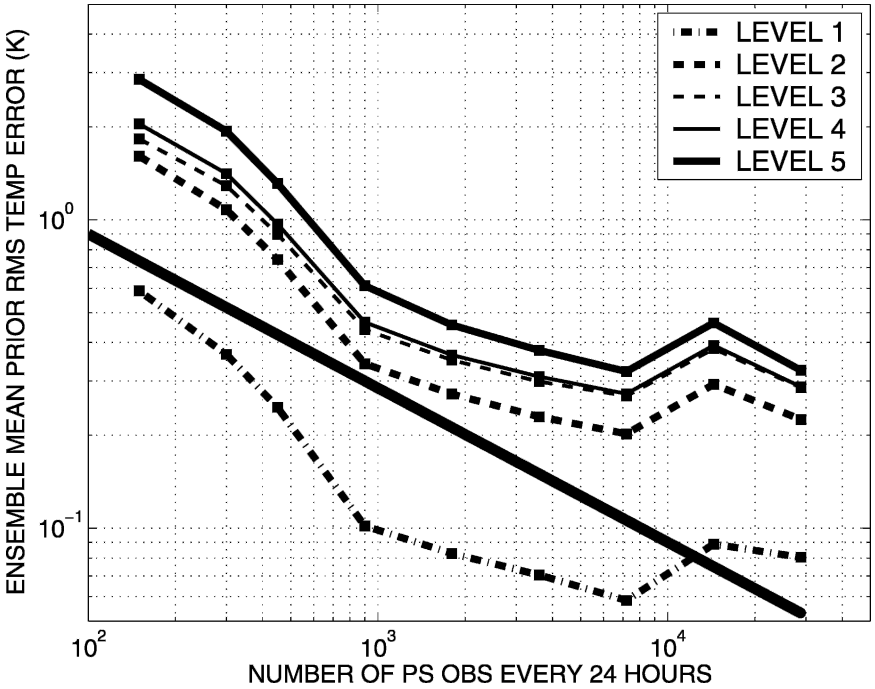

FIG. 8. As in Fig. 7 except for temperature field at each of the five model levels.

error doubling time of approximately 5.5 days, which is considerably slower than that found in real forecast GCMs or believed to hold for the real atmosphere (Simmons et al. 1995; Simmons and Hollingsworth 2002; Zhang et al. 2004a).

It is interesting to examine the relative benefits of increasing observational spatial density versus temporal frequency. First, it is important to note that results have been shown only for prior errors (i.e., forecast errors from forecasts with leads equal to the observation frequency). An alternative would have been to display error results from forecasts of the same lead for all observational frequency cases. However, this leads to only minor quantitative differences in the results as can be gleaned from an examination of the forecast error growth curve in Fig. 11.

There are several possible explanations for why in-

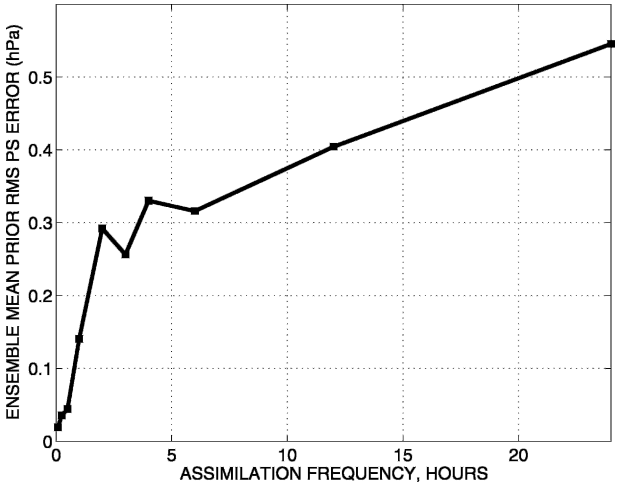

FIG. 9. Global mean rms error of the prior ensemble mean surface pressure as a function of the frequency with which 1800 randomly located surface pressure stations take observations with a 1.0-hPa error standard deviation. Cases plotted are for observations taken every 24, 12, 6, 4, 3, 2, 1 h and 30, 15, and 5 min.

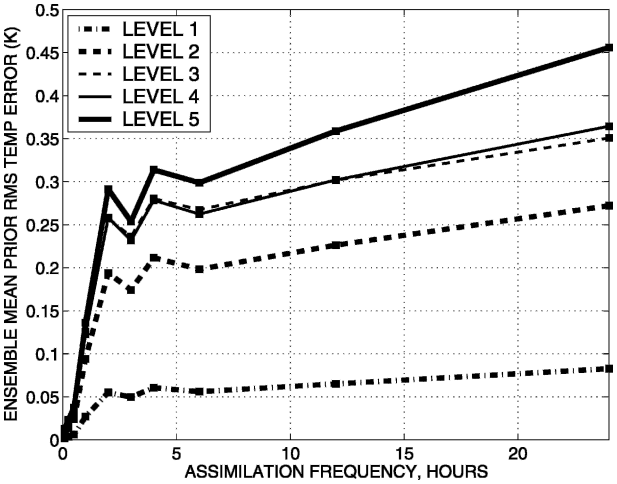

FIG. 10. As in Fig. 9 except for temperature field at each of the five model levels.

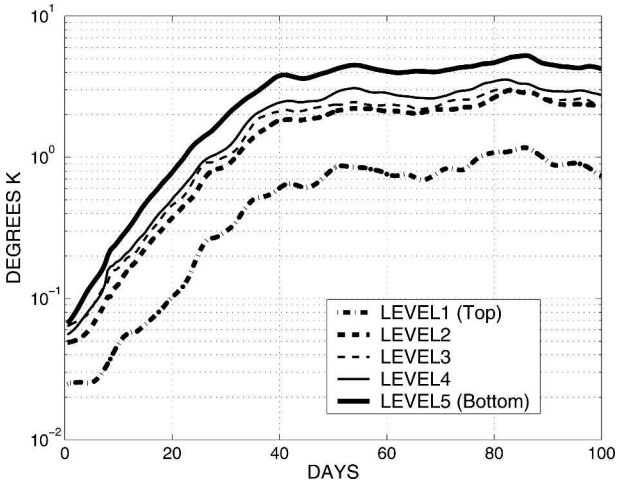

FIG. 11. Global mean rms error of ensemble mean temperature as a function of level and lead time for forecast initiated from day 400 of an assimilation using 1800 randomly located surface pressure stations observing once every 5 min with a 1.0-hPa error standard deviation.

creasing observation frequency is much more effective than increasing spatial density at a fixed frequency. First, the increased frequency observations are associated with an increased number of sample estimates of the covariance between observations and state variables. While these different samples are clearly not independent, they may still provide additional information about the relation between observation and state variables that can reduce error levels. Second, more frequent observations imply that wavelike structures are impacted by observations at many more phases of the wave. At wave crests or troughs, for instance, the information in observations has relatively little impact on adjacent state variables so observing only sporadically can lead to portions of a wave that are not tightly constrained by observations. It should be possible to explore further these possibilities by working with large ensembles that would reduce the prior covariance sampling error.

Another interesting aspect of the results is the behavior of the prior errors as a function of observation frequency for periods of a few hours. In these cases, Figs. 9 and 10 indicate some discontinuities in what appears to be otherwise a fairly regular curve. This behavior can also be seen clearly in a log–log plot of the PS rms as a function of observation frequency, Fig. 12. This has nearly straight and parallel segments at the high and low frequencies but has error above the continuations of these lines for intermediate frequencies.

The primitive equation B-grid dynamical core can produce gravity wave oscillations. The damping in the model is designed to quickly reduce the amplitude of gravity waves, however, the introduction of noise to a model state can generate high amplitude transient gravity waves. Sampling error and simplifying assumptions

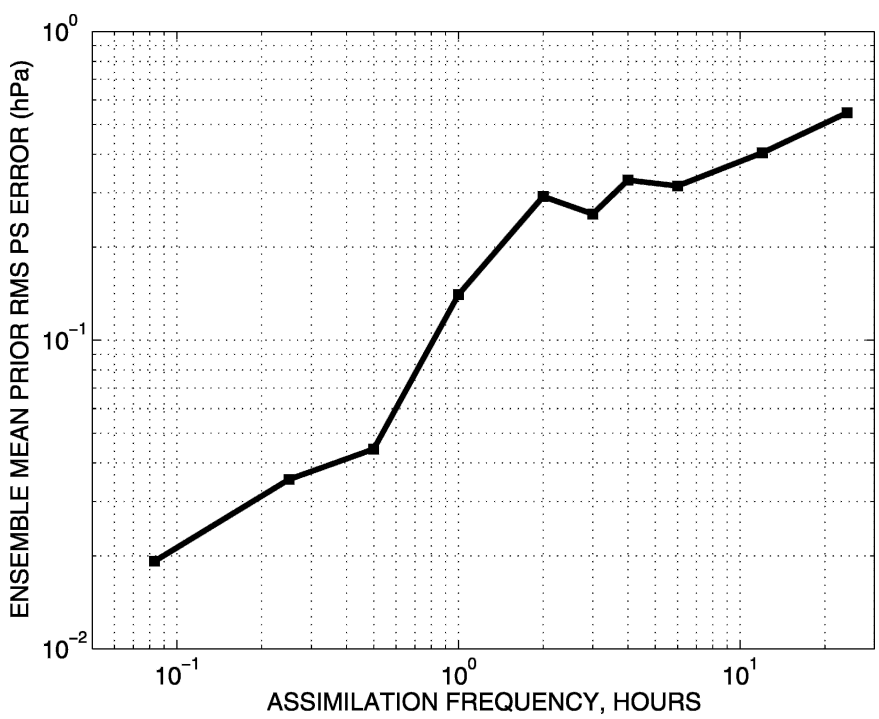

FIG. 12. Log–log plot of same quantity plotted on conventional axes in Fig. 9.

made in the ensemble filter can be viewed as leading to the introduction of ensemble increments with incorrect statistics into the assimilated field; this can be viewed as introducing noise to the model states. The gravity wave structure and period vary as a function of level and latitude with average periods of about 4 h. Figure 13 plots a time series of PS at a grid point at 30°N for 10 randomly selected members of a 20-member ensemble forecast. The forecast is initiated from the end of the assimilation with PS observations every 4 h and is integrated for 24 h. The time series of the truth is also shown in Fig. 13 as a darker line. The true trajectory is very smooth in midlatitudes over 24 h while the forecast initiated from the assimilation has high amplitude oscillations with a period of about 4 h that appear to gradually damp out during the forecast integration. The waves from all 10 ensemble members are nearly in phase.

Similar ensemble forecasts initiated from assimilations with higher or lower observational frequency show a greatly reduced amplitude of oscillations like those in Fig. 13. Apparently, intermediate frequency observations lead to assimilations that include relatively high amplitude gravity waves that are in phase for all ensemble members. In cases with high frequency observations, there are many observations per gravity wave period and the gravity wave amplitude is removed from the assimilations by the observations. In cases with low frequency observations, the time between observations is sufficient to allow the model to significantly damp the gravity wave amplitude between observations. In either case, the result is much reduced gravity wave amplitude in the resulting forecasts.

In the intermediate frequency cases with large gravity wave amplitude, the efficiency of the assimilation scheme is reduced because the prior estimates can be heavily biased since all ensemble members produce gravity waves that are in phase. For instance, if the next assimilation were performed using the 4-h forecast in Fig. 13 as a prior, the resulting posterior would continue to be heavily biased to a value greater than the truth. This in turn would tend to project on gravity waves that would continue to the next forecast step.

If ensemble filters are applied with primitive equation models and real data, this problem of generating gravity waves in the assimilated state is only expected to be worse. Some method for reducing this problem will probably be important in real applications (Mitchell et al. 2002). Results here suggest that one solution might be to increase the frequency of use of observations if this were possible. Other solutions might attempt to increase the model damping of gravity waves or to remove gravity waves from the assimilated ensembles by applying some sort of initialization procedure.

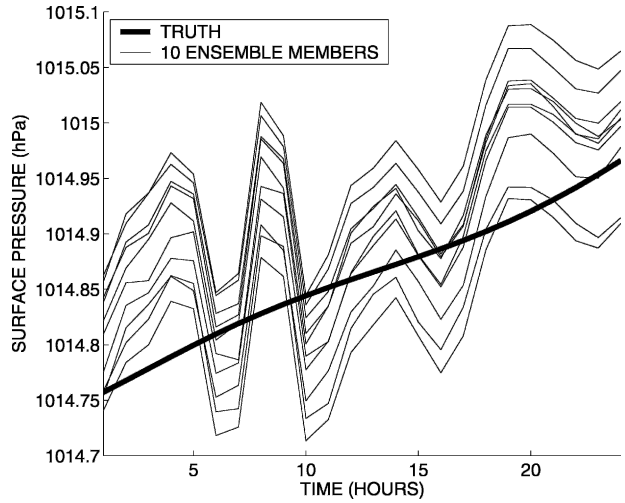

FIG. 13. Time series of surface pressure at a grid point in the midlatitudes (40°N) for a forecast initiated from day 400 of an assimilation using 1800 randomly located surface pressure stations observing once every 4 h with a 1.0-hPa error standard deviation. The thick curve is the truth from a long control integration while the 10 thin curves are the first 10 ensemble members from the 20-member ensemble forecast.

While it is dangerous to extrapolate results from simple idealized experiments to real assimilation prediction systems, the relative benefits of increased frequency versus increased density of observations may be worth further study. For historical and algorithmic reasons, operational atmospheric prediction systems tend to do assimilations at relatively low frequencies (usually around 6 h at present). Observations from a window surrounding the assimilation time are often weighted by a function of their time offset from the assimilation time. Many modern observations are available quite frequently and could be assimilated at times close to the actual observation time. It would be interesting to assess the impact of increasing the assimilation frequency to once an hour, for instance, in an operational filter system. Surface observations and a variety of remote sensing observations are available at hourly or higher frequencies and modern communications makes it possible to use these data.

Caution must be used when increasing the frequency of assimilation in models with multiple time-level time-differencing schemes. A model with leapfrog differencing and a 20-min time stepping would only have three time steps between each assimilation time. If the model differencing were restarted with a forward step after each assimilation, the time-differencing scheme could

easily become numerically unstable (assuming the basic forward scheme by itself is unstable).

6. Varying the observation type and error

Five additional assimilation cases were explored using the 1800 randomly located observing stations with observations once every 24 h. The first observed PS but with an observational error standard deviation of 2 instead of 1 hPa. The next two cases observed the lowest level (level 5) temperature with an observational error standard deviation of 1.0 and 0.5 K, respectively. The final pair observed both the u and v components of the lowest level wind with an observational error of 2.0 and 1.0 m s^{-1} independently on each component at each observing station.

Figure 14 shows the global average prior rms error for the six experiments (including the base PS case) while Fig. 15 shows the corresponding global rms errors for T at each of the five levels. Ordering the experiments from lowest to highest error for PS gives UV 1.0, PS 1.0, UV 2.0, T 0.5, PS 2.0, and T 1.0, while the same ordering for T errors gives UV 1.0, UV 2.0, PS 1.0, T 0.5, T 1.0, and PS 2.0. This implies that the relative impact of different observation types is different for different state variables. Also interesting is the relative improvement for halving the rms error of the observations. The largest decrease in relative error is for the PS observations where halving the observational error rms leads to nearly a halving of the prior error for the PS fields. The error reduction for halving the T observational error is significantly less than that for PS or the velocity components. For both PS and T prior errors, the net improvement factor from halving the observational error is greatest for PS observations and least for T observations with the velocity component observations in between.

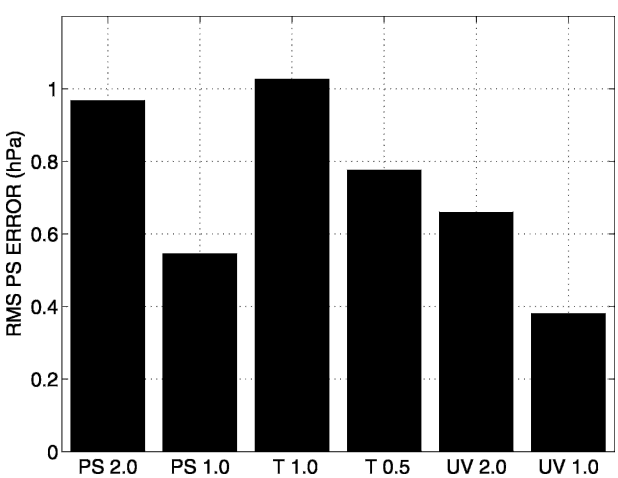

FIG. 14. Time mean global mean surface pressure rms error of ensemble mean prior from assimilations using observations from 1800 randomly located stations that observe once every 24 h. The six cases used observations of surface pressure with a 2.0-hPa error standard deviation, surface pressure with 1.0-hPa error, temperature at the lowest level with a 1-K error, temperature at the lowest level with a 0.5-K error, zonal and meridional wind components at the lowest level observed with independent 2.0 m s^{-1} error, and wind components at the lowest level with 1.0 m s^{-1} error.

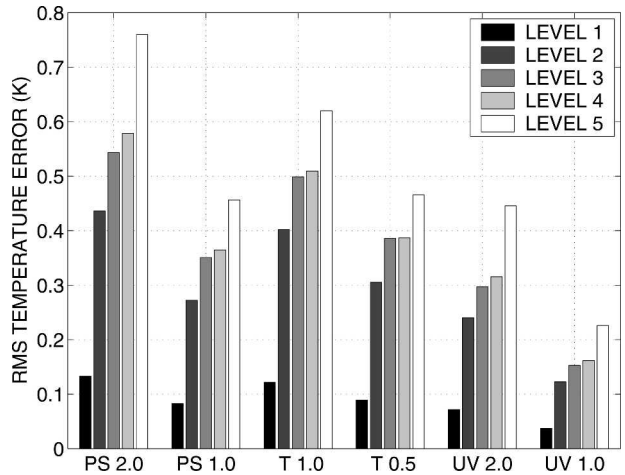

FIG. 15. As in Fig. 14 but for the temperature rms error at each level.

Again, it is dangerous to extrapolate too much from these idealized results to the real atmosphere. However, it does imply that the relative improvement from reducing error in different observing system components can be quite different and this effect is only expected to be enhanced given the complexity of the real assimilation/prediction problem. Similar differences in the idealized system (results not shown) arose for varying the spatial density of the different observation types. Improvement was greater for increasing the number of PS observations and least for increasing the number of lowest level temperature observations. Again, while it is dangerous to extrapolate the details, this does suggest that the value of increasing the density of different types of real observations may vary significantly.

One might expect halving the observation error to lead to a halving of rms errors in the assimilated fields in a perfect model experiment. The fact that this does not happen suggests that the linear approximation used in the regression step of the ensemble filter is one significant source of error. As the spread of the prior ensemble for a particular observation increases, errors in the regression can grow faster than linearly if the underlying relation between the observation and state

variables is nonlinear. Different degrees of nonlinearity in the prior relations between different observation types and model state variables are a likely explanation for the differing degree of improvement found for assimilation of different observation types.

7. Spatially localized observations

The spatial density of many real atmospheric and oceanic observations continues to be quite heterogeneous. Surface and radiosonde data are much denser over the midlatitude continents of the Northern Hemisphere. To explore how ensemble filters deal with horizontally spatially localized observations, a set of assimilation experiments was performed in which 450 surface observing stations were randomly located inside the region north of the equator and bounded by longitudes 90°E and 90°W. With ¼ the number of observations of the control run located over ¼ of the globe, this leaves the observational density the same in the observed region and 0 outside.

An assimilation experiment with PS observed once every 24 h at the 450 stations with an error standard deviation of 1 hPa was performed. Figure 16 displays the ensemble mean PS field and Fig. 17 the rms error of the ensemble mean after 400 days of assimilation. Over the region with data, the error is relatively small with the spatial mean of the error in the box being about twice the global mean error for PS from the control experiment with global observations. Inside the box, most of the structure of the midlatitude waves found in the truth is found in the ensemble mean. In the Northern Hemisphere, outside of the observed region, the error increases and the amount of zonally varying structure in the ensemble mean decreases. The increase in error and decrease in structure is slower downstream

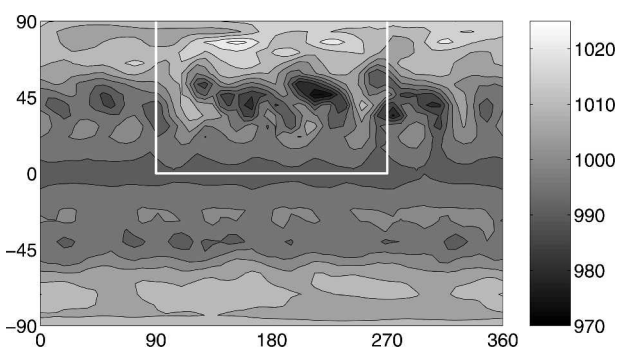

FIG. 16. Instantaneous ensemble mean prior surface pressure field (hPa) from day 400 of assimilation using 450 surface pressure observing stations located randomly in the white box and observing once every 24 h with a 1.0-hPa error standard deviation. Contour interval is 5 hPa from 970 to 1020 hPa.

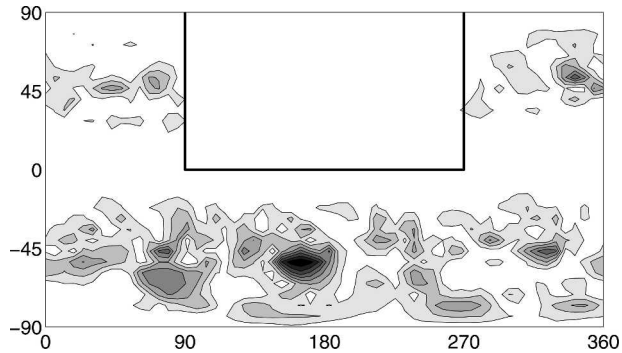

FIG. 17. Absolute value of ensemble mean prior surface pressure error (hPa) at day 400 of assimilation using 450 surface pressure observing stations located randomly in the Northern Hemisphere box and observing once every 24 h with a 1.0-hPa error standard deviation.

and more rapid upstream in the midlatitudes confirming that the information obtained from observations is primarily propagated downstream. In the Southern Hemisphere, error is nearly as large as the climatological error standard deviation and the ensemble mean assimilation is nearly zonally uniform. The error is slightly smaller and the amount of structure slightly greater in the Southern Hemisphere midlatitudes east of the eastern end of the observed region. This behavior appears to be representative of the time mean behavior throughout the assimilation. Similar plots for both temperature and wind component errors show qualitatively similar behavior. Apparently, large horizontal inhomogeneities in data density are not a significant problem for ensemble filters in idealized situations.

Rms error results for the T and wind component fields are qualitatively similar. For all variables and all levels, the rms error within the horizontal box is less than twice the corresponding error from the control experiment with global observations; rms errors are reduced to about one fifth of their climatological errors within the box. Outside the box, the errors are much larger with nearly climatological values in the Southern Hemisphere.

8. Conclusions

The perfect model results described here clearly illustrate the abilities of ensemble filters to extract information about the state of a three-dimensional multivariate model from observations of only one type of variable on the boundary of the domain. Other assimilation methods, especially 4D-variational methods, should also be able to do this. However, implementing

an ensemble filter is particularly simple requiring no a priori specification of background error statistics.

The extrapolation of these results to realistic applications must be undertaken with care. However, it is important to note that ensemble filter assimilations of only PS observations have already been successfully undertaken in real forecast models with real observations (Whitaker et al. 2004). Such experiments have constrained the error in the midtroposphere to considerably less than the levels of climatological variance. Assimilating much higher frequency PS data, perhaps hourly data, would be an interesting experiment in a real model. The results here suggest that error levels in the midtroposphere would be considerably decreased, perhaps approaching levels from using the array of upper air data available. All the problems present in real assimilations but not in this perfect model study certainly preclude such radical improvements, but just how well high frequency surface data can constrain assimilations of the real atmosphere remains uncertain. Other results found here, for instance, the relative value of temperature, wind, and PS data would also be interesting to explore in real assimilation systems. In addition to the many complicating factors already mentioned, the extreme representativeness errors found in the boundary layer for temperature and wind would also be a challenge.

Acknowledgments. Chris Snyder, Rebecca Morss, Alain Caya, Anthony Rosati, Matt Harrison, and an anonymous review all provided a number of excellent comments on earlier versions of this manuscript, leading to a much-improved final product.

REFERENCES

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

——, 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.,* **131,** 634–642.

——, and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.,* **127,** 2741–2758.

——, and Coauthors, 2004: The new GFDL global atmosphere and land model AM2–LM2: Evaluation with prescribed SST simulations. *J. Climate,* **17,** 4641–4673.

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

Daley, R., 1992: Estimating model-error covariances for application to atmospheric data assimilation. *Mon. Wea. Rev.,* **120,** 1735–1746.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.,* **123,** 1128–1145.

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

——, and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model. *Mon. Wea. Rev.,* **124,** 85–96.

Gadd, A. J., 1978: A split explicit integration scheme for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **104,** 569–582.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

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

——, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background-error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129,** 2776–2790.

——, C. Snyder, and R. E. Morss, 2002: Analysis-error statistics of a quasigeostrophic model using three-dimensional variational assimilation. *Mon. Wea. Rev.,* **130,** 2777–2790.

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

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

——, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.,* **133,** 604–620.

Keppenne, C. L., and M. Rienecker, 2002: Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.,* **130,** 2951–2965.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 97–110.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.,* **109,** 701–721.

Mesinger, F., 1977: Forward-backward scheme, and its use in a limited area model. *Contrib. Atmos. Phys.,* **50,** 186–199.

Mitchell, H. L., and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.,* **128,** 416–433.

——, ——, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.,* **130,** 2791–2808.

Morss, R. E., K. A. Emanuel, and C. Snyder, 2001: Idealized adaptive observation strategies for improving numerical weather prediction. *J. Atmos. Sci.,* **58,** 210–232.

Pham, D. T., 2001: Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.,* **129,** 1194–1207.

Reichle, R. H., J. P. Walker, R. D. Koster, and P. R. Houser, 2002: Extended versus ensemble Kalman filtering for land data assimilation. *J. Hydrometeor.,* **3,** 728–740.

Shapiro, R., 1970: Smoothing, filtering and boundary effects. *Rev. Geophys. Space Phys.,* **8,** 359–387.

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

——, and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **128,** 647–677.

——, R. Mureau, and T. Petroliagas, 1995: Error growth and estimates of predictability from the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.,* **121,** 1739–1771.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.,* **131,** 1663–1677.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.,* **131,** 1485–1490.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130,** 1913–1924.

——, G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.,* **132**, 1190–1200.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.,* **130,** 2905–2916.

Wyman, B. L., 1996: A step-mountain coordinate general circulation model: Description and validation of medium range forecasts. *Mon. Wea. Rev.,* **124,** 102–121.

Zhang, S., J. L. Anderson, A. Rosati, M. J. Harrison, S. P. Khare, and A. Wittenberg, 2004a: Multiple time level adjustment for data assimilation. *Tellus,* **56A,** 2–15.