Regional Ensemble–Variational Data Assimilation Using Global Ensemble Forecasts

WAN-SHU WU, DAVID F. PARRISH, ERIC ROGERS, AND YING LIN

National Centers for Environmental Prediction, College Park, Maryland

(Manuscript received 8 March 2016, in final form 13 October 2016)

ABSTRACT

At the National Centers for Environmental Prediction, the global ensemble forecasts from the ensemble Kalman filter scheme in the Global Forecast System are applied in a regional three-dimensional (3D) and a four dimensional (4D) ensemble–variational (EnVar) data assimilation system. The application is a one-way variational method using hybrid static and ensemble error covariances. To enhance impact, three new features have been added to the existing EnVar system in the Gridpoint Statistical Interpolation (GSI). First, the constant coefficients that assign relative weight between the ensemble and static background error are now allowed to vary in the vertical. Second, a new formulation is introduced for the ensemble contribution to the analysis surface pressure. Finally, in order to make use of the information in the ensemble mean that is disregarded in the existing EnVar in GSI, the trajectory correction, a novel approach, is introduced. Relative to the application of a 3D variational data assimilation algorithm, a clear positive impact on 1–3-day forecasts is realized when applying 3DEnVar analyses in the North American Mesoscale Forecast System (NAM). The 3DEnVar DA system was operationally implemented in the NAM Data Assimilation System in August 2014. Application of a 4DEnVar algorithm is shown to further improve forecast accuracy relative to the 3DEnVar. The approach described in this paper effectively combines contributions from both the regional and the global forecast systems to produce the initial conditions for the regional NAM system.

1. Introduction

Lorenc (2003) provided a thorough comparison between ensemble Kalman filter (EnKF) and four-dimensional variational data assimilation (4DVar) and proposed a hybrid method with extended control variable formulation for regional mesoscale numerical weather prediction (NWP) systems. Kalnay et al. (2007) and Gustafsson (2007) discussed the advantages and the disadvantages of these two approaches and also concluded that a hybrid method would be beneficial to meteorological data assimilation.

The hybrid ensemble–variational [EnVar as recommended nomenclature in Lorenc (2013)] data assimilation (DA) algorithms employ both the static and ensemble-based, flow-dependent error covariances within a variational framework. For variational data assimilation (Var), full-rank static background error covariances are usually derived with assumptions of isotropy and homogeneity in space and time (Parrish and Derber 1992). On the other hand, the error amplitudes and error structures and correlations among variables with the ensemble-based background error covariances are flow dependent and fully nonlinear without the limitation of a certain shape of the structure or using a simplified projection relationship between variables.

Without using the tangent linear and the adjoint of the forecast model as in 4DVar, the 4DEnVar algorithms are used with an ensemble of nonlinear model integrations over the data assimilation window to provide model guidance (Liu et al. 2008, 2009; Liu and Xiao 2013; Buehner et al. 2010a,b; Desroziers et al. 2014; Wang and Lei 2014; Kleist and Ide 2015b). In this application, the 4D advantage is mainly through the model evolution represented in the ensemble perturbations of each ensemble member. Gustafsson and Bojarova (2014) reported that 4DEnVar for the High Resolution Limited Area Model outperformed 4DVar, its computational costs were significantly reduced in comparison with standard 4DVar, and the scalability of the algorithm was improved.

The 3DEnVar and 4DEnVar approaches have been used within both global and regional frameworks. For example, in regional applications there have been studies of various hybrid algorithms (Wang et al. 2007, 2008a,b;...
Liu et al. 2009; Zhang and Zhang 2012; Liu and Xiao 2013; Zhang et al. 2013; Schwartz et al. 2013; Schwartz and Liu 2014; Pan et al. 2014; Schwartz et al. 2015; Caron et al. 2015). It has also been demonstrated that 3DEnVar and 4DEnVar DA methods are suitable for operational NWP applications (Hamill et al. 2011; Buehner et al. 2013; Clayton et al. 2013; Kuhl et al. 2013; Wang et al. 2013; Gustafsson and Bojarova 2014; Buehner et al. 2015; Caron et al. 2015). Following Wang (2010) as first proposed by Lorenc (2003), the 3DEnVar was developed in the Gridpoint Statistical Interpolation (GSI) system using an extended control variable at the National Centers for Environmental Prediction (NCEP). The 3DEnVar was implemented in the operational Global Data Assimilation System (GDAS) in May 2012 and was used in studies with the Global Forecast System (GFS) system (Wang et al. 2013; Kleist and Ide 2015a,b). The same 3DEnVar scheme in the GSI is employed in this study for the North American Mesoscale Forecast System (NAM) Data Assimilation System (NDAS) at NCEP.

Xu et al. (2008) and Zhao et al. (2015) showed that the time-expanded sampling method could be used to reduce the number of forecast runs needed to produce an ensemble with the desired sample size. Gustafsson et al. (2014) also used a +2-h lagged valid time ensemble to double the size of the ensemble applied in their 4DVar hybrid experiments. It was reported in both papers that the ensemble perturbations seem to have dynamical consistency and can be used with a time shift. In this study, whether the perturbations have the dynamical consistency to be used in another forecast system will be tested. The flow-dependent error covariances derived from the global ensemble will be used to produce the initial conditions for the regional NAM system with the assumption that they can provide useful information to improve the large-scale components of the regional background error within the DA system.

Because of an improved forecast impact over 3DVar in NCEP’s regional NAM forecast system, a 3DEnVar DA using global EnKF ensemble forecasts from NCEP’s GFS was implemented into NCEP’s regional NDAS on 12 August 2014. A similar approach was later used in the Regional Deterministic Prediction System in Environment Canada (Caron et al. 2015). This paper documents the work of implementing the 3DEnVar DA and the extension to a 4DEnVar in the regional NAM system.

The regional model and DA systems, the 4DEnVar setup, the global ensemble, and the proposed trajectory correction and its associated impact are described in section 2. The comparison of the forecast impact from the 3DVar, 3DEnVar, and 4DEnVar with its final configuration in NDAS is included in section 3. The conclusions and discussion are provided in section 4.

2. Model and DA configuration

a. Forecast model NMMB and NDAS

The NAM is one of the primary forecast systems by which NCEP provides mesoscale guidance to public- and private-sector meteorologists. It is run four times daily at 0000, 0600, 1200, and 1800 UTC using the NOAA Environmental Modeling System (NEMS) version of the Nonhydrostatic Multiscale Model Model with Arakawa B-grid staggering (NMMB; Janjic’ et al. 2001). The North American parent domain, as shown in Fig. 1, has a horizontal grid spacing of 12 km and a model top at 2 hPa with 60 vertical levels. The NAM run makes forecasts in its parent domain out to 84 h. There are four fixed nested domains: the continental United States (CONUS), Alaska, Hawaii, and Puerto Rico, with grid spacings of 4, 6, 3, and 3 km, respectively, run to 60 h, which are all one-way nests inside the parent domain. One high-resolution fire weather nest runs to 36 h with grid spacing of 1.33 km if over the CONUS and 1.5 km if over Alaska. The NAM is initialized with a 12-h run of NDAS, which runs a sequence of four GSI (Wu et al. 2002; Kleist et al. 2009a,b) analyses and 3-h NEMS–NMMB forecasts using all available observations to provide a first guess to the NAM “on time” analysis. NDAS 3-h forecasts are initialized with a diabatic digital filter. Figure 2 shows a schematic diagram of NDAS/NAM partial cycling. Each forecast cycle begins with a 12-h analysis–forecast window during which analyses are conducted at 3-h intervals.
FIG. 2. A schematic diagram of NDAS/NAM DA cycling. Each forecast cycle begins with a 12-h analysis–forecast window during which analyses are conducted at 3-h intervals (TM12, TM09, etc.). TM00 refers to the forecast initialization time (e.g., 0000, 0600, 1200, or 1800 UTC). At TM12 the first guess for the atmosphere is a 6-h forecast from the GDAS, and the land states are cycled from the previous NAM/NDAS cycle.

(TM12, TM09, etc.). TM00 refers to the forecast initialization time (e.g., 0000, 0600, 1200, or 1800 UTC). At TM12 the first guess for the atmosphere is a 6-h forecast from GDAS. The land states are cycled from the previous NAM/NDAS cycle.

b. Hybrid DA

In a hybrid variational–ensemble setting, the ensemble perturbations are incorporated directly into the variational cost function through an augmented control variable (the α method), as suggested by Lorenz (2003). For both 3DEnVar and 4DEnVar, the analysis increments are obtained by minimizing the following hybrid cost function:

\[
J(x_f, \alpha) = \frac{1}{2} (x_f - x_f)^T B^{-1} (x_f - x_f) + \frac{1}{2} \sum_{n=1}^{N} (\alpha^n)^T L^{-1} (\alpha^n)
\]

\[+ \frac{1}{2} \sum_{t=1}^{T} (y_t^f - Hx_f^f)^T R^{-1} (y_t^f - Hx_f^f), \tag{1}\]

where \(x_f\) is the increment associated with the static covariance, \(x_f\) is the total analysis increment at time \(t\) [see Eqs. (2) and (3)], \(T = 1\) for 3DEnVar without a first guess at the appropriate time (FGAT) and is the number of time bins in the assimilation window for 3DEnVar with FGAT and 4DEnVar. \(\alpha^n\) is the control variable for ensemble member \(n\), \(N\) is the total number of ensemble members, \(L\) is the error covariance for \(\alpha\), \(B\) is the static background error covariance, \(H\) is the observation error covariance, \(R\) is the linearized observational operator, and \(y_t^f\) is the observation innovation based on the deterministic first guess in time bin \(t\). The localization included in the error covariance \(L\) is realized through recursive filters (Purser et al. 2003a,b) both in the horizontal and in the vertical for regional applications. The weighting coefficients for the static \(\beta_f\) and the ensemble \(\beta_e\) terms satisfied the equation

\[
\frac{1}{\beta_f} + \frac{1}{\beta_e} = 1.
\]

The total analysis increments at time \(t\) are defined as

\[
x_t' = x_f' + \sum_{n=1}^{N} \alpha^n W_n^e, \tag{2}\]

where \(x_t'\) is the \(n\)th global ensemble perturbation, normalized by \((N - 1)^{1/2}\), and interpolated to the regional analysis grid at time \(t\), and the symbol \(\cdot\) denotes the Schur product. As pointed out by Gustafsson and Bojarova (2014), the analysis increment from the second term is simply a linear combination of the ensemble perturbations. The resolution of \(\alpha^n\) is the same as the analysis grid in this application and no interpolation operator is needed. Note that in both Eqs. (2) and (3) \(t = 1\) for 3DEnVar with or without FGAT, and \(t = 1, \ldots, T\) for 4DEnVar.

In the original GSI, Eq. (2) is applied to the three-dimensional variables (\(u, v, q\)), and \(\alpha^n\) at level 1 (near surface) is applied to the two-dimensional surface pressure. Surface pressure is tied to changes in the column mass and dynamics in the free atmosphere; as such, there is correlation between the surface pressure and the mass and wind variables throughout the vertical column. This correlation is taken into account in this study. A new formulation for the increments of surface pressure is defined as

\[
x_{ps}' = x_{ps}' + \sum_{n=1}^{N} \left[ \sum_{k=1}^{K} (\alpha^n_k W_k^e) \cdot x_{ps}^{en} \right], \tag{3}\]

where \(\alpha^n_k\) is \(\alpha^n\) at level \(k\), \(W_k^e\) is a weighting at the \(k\)th level, and \(x_{ps}^{en}\) is the \(n\)th ensemble perturbation of surface pressure at time bin \(t\). In addition, \(\alpha^n\) is weighted by function \(W\) and integrated vertically over the model layers \(K\) for the contribution to surface pressure. The Schur product operates on the resulting two-dimensional field and the ensemble perturbations of surface pressure. The weighting function \(W\) used in the NDAS is shown in Fig. 3, where the projection matrix from the streamfunction to surface pressure used in the static background error covariance (Wu et al. 2002) is normalized by its vertically integrated value. The dimensions of \(W\) are limited to one (just height dependent) to minimize the sampling error in estimating the statistics and to avoid nonphysical results. The higher weighting near the surface in Fig. 3 is indicative of the fact that the analysis of the surface pressure depends mostly on observations over the lowest levels. Note that the same “alpha” is used for all variables so that the three-dimensional information from other variables will also impact the surface pressure increments. The relevant
observations within this context include conventional observations of wind, temperature, humidity, and surface pressure, along with satellite radiances that have greater sensitivities in the lower troposphere.

For the 3DEnVar in NDAS, Eqs. (2) and (3) are used at the center of the analysis window. With the 4DEnVar for NAM, Eqs. (2) and (3) are applied at each time bin in the assimilation window. While the analysis is available at each time bin with 4DEnVar, the four-dimensional incremental analysis update (4DIAU) with time-varying increments proposed by Buehner et al. (2015) is not used in this application. The analysis at the center of the window is chosen to initialize the NAM free forecast.

The background error covariances currently used in NDAS are derived with the National Meteorological Center (NMC, now known as NCEP) method (Parrish and Derber 1992). Forecast differences are used to estimate the balance projection, the variances, and the error structure statistics for the analysis variables. With the forecasts of the NAM parent domain in Fig. 1, spectral calculations are used in finding the streamfunction and velocity potential from the wind fields. The structure functions are estimated with an autocorrelation method that generates more stable and robust statistics than those with other methods. For the cycle at TM12 in Fig. 2, when the first guess for the atmosphere is a 6-h forecast from the GDAS, global background error statistics derived from global forecasts are employed. The statistics are interpolated to the latitude and height of the regional grid, and the differences in the resolutions are taken into account in the application of the error structure. For the other cycles, from TM09 to TM00 when regional forecasts are used as the first guess, regional background error statistics are used.

The 80-member T574 global EnKF ensemble is used in the NDAS 3DEnVar. The EnVar method is one-way coupled. The analysis increment grid is 3 times coarser (36 km) than the model grid. The analysis increments are interpolated to the original 12-km grid and added back to the deterministic background. The global T574 ensemble is of 35-km grid spacing. Since the global ensemble perturbations used in this study are of similar resolution to the regional analysis, single-resolution EnVar is chosen. Unlike Schwartz et al. (2015), no interpolation operator is used in Eqs. (2) or (3). The ensemble perturbation is interpolated to the regional analysis grid before being used in the minimization.

A slight degradation over 3DVar near the upper and lower boundaries was observed when a 3DEnVar analysis using global ensemble was first applied in NDAS. There are important differences between GFS and the regional NAM related to model resolution, physics, and numerical approximations at the top and bottom of the model. To mitigate the problem, $\beta_f$ and $\beta_e$ are made height dependent in the regional application, as shown in Fig. 4, where $\beta_e=0.5$ in the middle of the atmosphere and decreases below (above) 850 (150) hPa to 0.25 at the boundaries. The degradation near the boundaries is eliminated with the vertically varying weighting in 3DEnVar. However, the vertical variation of $\beta_f$ and $\beta_e$ is turned off in the regional 4DEnVar since no obvious degradation is observed without the variation. The reason for this discrepancy may be the stronger constraint to fit the model trajectory within the 4D application.

c. 3DEnVar and 4DEnVar setup

For the 3DEnVar application, $T=1$ in Eq. (1) and $t=1$ in Eqs. (2) and (3) are employed. FGAT is not used in the current operational NDAS with 3DEnVar because it produces no impact on forecast quality. For each of the five analysis cycles in Fig. 2, a set of 80-member T574 global EnKF ensemble forecasts and a regional first guess valid at the analysis time are needed. The tangent-linear normal-mode constraint as described in Kleist et al. (2009a) is not utilized in the regional mode.

When Eq. (1) is applied to 4DEnVar, the static part of the solution $x_f$ and $\alpha$ are constant in the assimilation time
window, while the ensemble perturbations $x_e$ evolve with time. The time-evolving ensemble perturbations are the only four-dimensional components of the algorithm. For NDAS 4DEnVar, a 6-h assimilation window was chosen. A set of the ensemble and the first guess valid at each time bin is needed, and the length of the time bin needs to be determined. Time bin refers to the discretization in time for the 4D analysis, which represents the frequency with which ensemble perturbations and first-guess fields are used throughout the DA window. It was found in Wang and Lei (2014) that the forecast accuracy and the balances of analysis of 4DEnVar was degraded when less frequent ensemble perturbations, 2-h instead of 1-h bins, were used in a global system. One would reason that since the flow solutions could be fit to the observations more often with shorter bins, the discretization error in time can be decreased with the shorter ones. However, for this regional 4DEnVar system mixed results are found when using hourly bins as opposed to 3-hourly bins. The time evolution of the domain-averaged root-mean-square errors (RMSEs) against radiosondes from experiments using 4DEnVar with hourly bins and with 3-hourly bins are shown in Fig. 5. The RMSEs for wind, temperature, relative humidity, and surface pressure are of 24-h forecasts initialized at 0000 UTC from 8 April to 3 May 2015. The time means listed in the insets in Fig. 5 indicate a neutral forecast impact from the length of the time bins. A possible explanation for these neutral results is the change in data density because of the temporal discretization of the 4DEnVar algorithm. Smaller discretization error with an hourly bin is not the only factor that impacts the forecast quality. With hourly binning the observations are divided into more bins so that there are fewer data points in each bin. Higher sampling error (i.e., not enough data in each bin to solve for the optimal $\alpha$ in the cost function) might be the reason when the experiment with hourly bins performs worse than that with 3-hourly bins. Since the impact from the size of the bins is neutral, in the remainder of the study 3-hourly bins were used because of the availability of a global ensemble from the operational system at NCEP.

Figure 6 shows a schematic diagram of the final 4DEnVar setup in NDAS. Each forecast cycle begins with a 9-h NMMB forecast from the previous cycle. The 3-, 6-, and 9-h forecasts and corresponding global EnKF ensemble are used at $T_2$, $T_3$, and $T_4$ of the assimilation window. The 4DEnVar analysis valid at the center of the assimilation window is used as the initial conditions for the 84-h NAM forecast. To run 4DEnVar, it is necessary to apply FGAT, which allows the observation innovation to be evaluated in each time bin of the assimilation window. For the regional 4DEnVar experiments, a better forecast impact was found in the experiments using the 3-, 6-, and 9-h forecasts (see Fig. 6) from the previous cycle as the first guess than those using three 3-h forecasts produced during the 3-hourly cycles (see Fig. 2) of NDAS assimilation.

The forecast accuracies for the 3DEnVar and 4DEnVar approaches are sensitive to the scale of the horizontal localization. For 4DEnVar, horizontal localization with a Gaussian scale of 800 km is found to have the optimal impact on the NAM forecasts while 300 km is used for 3DEnVar in the operational NDAS. For vertical localization a recursive filter with a scale of five model grid spaces is used in both 3DEnVar and 4DEnVar. A different partitioning between the static and the ensemble background error covariances is used for 4DEnVar ($\beta_e = 0.75$) as compared with 3DEnVar ($\beta_e = 0.5$). The choices are based on the impact to the quality of the NAM forecasts.

All the observations used in the operational NDAS are used in the experiments reported upon in this study. The types of the conventional data are available online in a table showing the PREPBUFR report type for regional GSI (EMC 2015). Another table showing the current use of data not passing through PREPBUFR
processing (EMC 2016) includes satellite radiances, satellite-retrieved products (sea surface temperature, sea surface wind, and atmospheric motion winds), radar, and GPS radio occultation.

d. Global ensemble

An 80-member global EnKF ensemble with T574 resolution is used in the operational NDAS with 3DEnVar. Table 1 shows the experiments designed to test the impact of various versions of the global ensemble used with regional 4DEnVar. There are the operationally generated 80-member T574 global EnKF ensemble with a 3DEnVar recentering and an experimental global EnKF ensemble with 4DEnVar recentering. The recentering refers to the global EnKF ensemble mean being replaced every cycle by the global EnVar analysis. A recentering procedure ensures that the EnKF analysis ensemble is always centered on the hybrid EnVar analysis to maintain synergy between the EnKF and hybrid EnVar analyses (Wang et al. 2013; Kleist and Ide 2015a). The control was initiated from the regional 3DEnVar. The regional 4DEnVar with a global ensemble using EnKF from the operational GDAS with 3DEnVar recentering was used in the first experiment. The second and third experiments were with T254 and T574 using EnKF in an experimental GDAS with 4DEnVar recentering. The domain-mean RMSEs compared with conventional data for the 24-h forecasts are included in Table 2. Using the EnKF ensemble with the 4DEnVar recentered from GDAS in experiments T254_4d and T574_4d produced no forecast advantage over using...
the ensemble with the 3DEnVar recentered in experiment T574_op. It also shows little impact from using the full-resolution ensemble in experiment T574_4d compared with experiment T254_4d. This result seems to suggest that the impact from the global ensemble on the regional 4DEnVar DA system is mostly at the larger scales of the background error covariances, which are resolved in both the truncated T254 and full-resolution T574 ensembles. The same finding was reached by Schwartz et al. (2015), that EnVar with a lower-resolution ensemble can initialize forecasts at a quality similar to EnVar with a full-resolution ensemble. The setup as in experiment T574_op is used in the following experiments.

e. Trajectory correction

In the original GSI EnVar, only the ensemble perturbations are used, and the ensemble mean is disregarded. The trajectory of the ensemble mean can be different from that of the first guess, and the information in the ensemble mean can potentially be beneficial to the EnVar DA system. For a regional system using a global ensemble, the domain sizes, model resolutions, physics parameterizations and numerical methods employed in each system can all contribute to the differences between the regional forecast used as a first guess and the global ensemble mean. Even for a “two way” coupled system, the trajectory of the low-resolution ensemble mean could also be deviated from the high-resolution deterministic forecast. For example, the recentering procedure is needed in the global EnKF/EnVar system in NCEP. To explicitly make use of the information in the ensemble mean, trajectory correction (TC) was added to GSI as an option in the regional EnVar system. When the TC option is turned on, an extra member is added to the ensemble. The perturbation for this extra member is the difference between the global ensemble mean and the regional first guess. In 3DEnVar, TC enables the solution to be drawn, locally, close to the global ensemble mean. In 4DEnVar, TC enables the solution in the assimilation window to be locally close to the trajectory of the global ensemble mean.

The $a$ for this member is given the same weight (the same error covariance) as the $a$ for any of the other 80 ensemble members. The influence of the extra member (the magnitude of $a$) on the analysis is determined by the solution of Eq. (1). Since localization is applied within the EnVar approach, TC is also local. Within the localization distance, if the perturbations associated with this extra member fit well with the innovations throughout the assimilation window, $a$ for the member could be close to one. By adding the increments of this member back to the first guess, the resulting analysis would be close to the global ensemble mean. Note that this process does not forcefully correct the regional analysis toward the global ensemble mean. The correction happens automatically when Eq. (1) is solved. The addition of the extra member with TC is made just to improve the coverage of the ensemble in phase space (to enrich the ensemble).

Figure 7 shows the time evolution of the domain-averaged 24-h forecast RMSE against the radiosonde observations from experiments using 3DEnVar with TC (3DTC) and 3DEnVar without TC (3DNTC). Although the TC produced a limited impact in the regional 3DEnVar, the benefits to the wind forecasts are

| Table 1. Description of experiments with NDAS EnVar at 0000 UTC 4 Feb 2015. |
|---|---|
| Expt | Description |
| Cntl | Regional 3DEnVar using the global T574 EnKF ensemble with 3DEnVar recentering |
| T574_op | Regional 4DEnVar using the global T574 EnKF ensemble with 3DEnVar recentering |
| T254_4d | Regional 4DEnVar using the global T254 EnKF ensemble with 4DEnVar recentering |
| T574_4d | Regional 4DEnVar using the global T574 EnKF ensemble with 4DEnVar recentering |
Fig. 7. Time evolution of the domain-averaged 24-h forecast RMSEs from the 3DTC (blue) and 3DNTC (black) experiments against radiosonde observations for forecasts initialized at 0000 UTC during 3–22 Dec 2015 of (a) vector wind ($\text{m s}^{-1}$), (b) temperature (K), (c) relative humidity (%), and (d) surface pressure (hPa). The time means are shown in insets.

TABLE 2. The 24-h forecast domain-averaged RMSEs for the experiments in Table 1 compared with conventional observations of surface pressure, relative humidity, temperature, and vector wind.

<table>
<thead>
<tr>
<th>Expt</th>
<th>Surface pressure (psfc; hPa)</th>
<th>RH (%)</th>
<th>Temp (K)</th>
<th>Wind ($\text{m s}^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cntl</td>
<td>1.0866</td>
<td>17.17</td>
<td>1.50</td>
<td>5.74</td>
</tr>
<tr>
<td>T574_op</td>
<td>1.0787</td>
<td>16.90</td>
<td>1.47</td>
<td>5.81</td>
</tr>
<tr>
<td>T254_4d</td>
<td>1.1427</td>
<td>17.10</td>
<td>1.48</td>
<td>6.06</td>
</tr>
<tr>
<td>T574_4d</td>
<td>1.1467</td>
<td>17.05</td>
<td>1.48</td>
<td>6.09</td>
</tr>
</tbody>
</table>

consistently positive in Fig. 7a. The time evolution of domain-averaged 24-h forecast RMSEs for forecasts initialized from 4DEnVar with TC (4DTC), and from 4DEnVar without TC (4DNTC), is depicted in Fig. 8. The data assimilation and forecast experiment was carried out for the period 3 January–8 February 2016. The positive impact on all the variables from TC is more accentuated for the 4DEnVar than for the 3DEnVar. Note that the RMSEs in Figs. 7 and 8 represent 3D (2D for surface pressure) domain averages, and the RMSEs have significant spatial variations. To further illustrate this, we present a vertical profile of horizontally domain-averaged 24-h forecast RMSEs of wind to radiosondes for one randomly selected case: 0000 UTC 14 January 2016 (time 14 in Fig. 8a) in Fig. 9. The result of 3DTC is also included as reference. The impact from applying 4DEnVar (4DTC) as opposed to 3DEnVar (3DTC) is positive at most pressure levels except at 950 hPa, with the maximum benefit seen between 500 and 150 hPa. The impact from using TC (4DTC versus 4DNTC) is positive at all pressure levels and is most...
substantial in the mid- to upper troposphere. The horizontally averaged fit to radiosonde winds is improved by 0.9 m s$^{-1}$ at 275 hPa from using 4DEnVar instead of 3DEnVar. Note that the impact of adding the extra member onto the existing 80 members with TC is nearly half of that magnitude (0.4 m s$^{-1}$, comparing 4DNTC and 4DTC) at that level. With the encouraging results, TC is used in NDAS parallel tests for EnVar reported in the next section.

It was shown in Gustafsson et al. (2014) that the richness in spatial structures of the ensemble is more important than the scaling of the amplitude of the ensemble perturbations in experiments with HIRLAM 4DVar Hybrid assimilation. The proposed TC in this study provides a way to enrich the ensemble. Our results with NAM 3D and 4DEnVar assimilation indicate that the forecast quality is improved with TC, which confirms the importance of the richness in the spatial structure of the ensemble perturbations to the hybrid ensemble assimilations.

3. Forecast impact with 3DEnVar and 4DEnVar

The final configurations of the 3DEnVar analysis systems were tested in regional NDAS parallels. The geopotential height RMSE of the preimplementation parallel tests of 3DEnVar, averaged from 25 April to 6 June 2012, are shown in Fig. 10. The 3DEnVar has a consistent and
substantial positive impact at all levels over the 3DVar in the control with regard to the 24-, 48-, and 72-h forecast quality. A similar impact was observed in all variables (not shown). Consequently, the 3DEnVar analysis system was implemented in the operational NDAS on 12 August 2014.

The regional 4DEnVar was applied in an NDAS parallel while a regional parallel with 3DEnVar served as the control. Although the overall 4DEnVar impact was consistently positive over 3DEnVar during the 5-month (15 June–15 November 2015) test, the magnitude of the improvement was more significant during the cold-season months. This might be attributed to the fact that the background error covariance from the global ensemble provides more meaningful error statistics during the winter season because of the additional baroclinicity in the extratropics. The impact from using 4DEnVar to initiate the forecasts compared with using a 3DEnVar analysis in NDAS is depicted in Fig. 11. The forecast RMSEs against radiosonde observations of geopotential height, wind, temperature, and specific humidity averaged for 2 months (from 15 September 2015 to 15 November 2015) are shown. The fit to the height in Fig. 11a is neutral at 24 h and better for 4DEnVar parallel at 48 and 72 h everywhere except near the top of the model domain when compared with the 3DEnVar. The fit to the wind observations in Fig. 11b for the forecasts initiated during the 4DEnVar analysis is better at all layers and all lead times than that from the 3DEnVar approach. The most notable benefit is realized for 48-h forecasts between 200 and 300 hPa. A positive impact on the temperature in Fig. 11c is found at all levels, and the largest improvement over 3DEnVar is realized in the middle of the troposphere. The impact on the specific humidity in Fig. 11d is also positive at all layers. Equitable threat scores and bias scores for day 3 precipitation over the continental United States for the same two months are shown in Fig. 12. In general, the experiment with
4DEnVar has better equitable threat scores and lower biases than that with 3DEnVar.

4. Conclusions and discussion

In this paper we describe the work undertaken to use the global EnKF ensemble in the regional NDAS with the EnVar algorithm at NCEP. Three new features are added to the existing EnVar system in the GSI, namely, the vertically varying weighting coefficients ($\beta_f$ and $\beta_e$) for the static and the ensemble partitioning, the vertically integrated contribution of the ensemble coefficient $\alpha$ to the surface pressure analysis, and TC. The vertical varying $\beta_f$ and $\beta_e$ are necessary because we use the ensemble perturbations from the global system, which has very different vertical model grid spacing than that in the NAM system. However, the feature is only used in the 3DEnVar application but not in 4DEnVar. The second feature is from the physical consideration that changes in surface pressure are tied to changes in the column mass and dynamics in the free atmosphere. With TC, the third feature, both the perturbations, and the ensemble mean from the global EnKF system contribute to the regional analysis through the EnVar algorithm. The difference between the first guess and the ensemble mean is added as the perturbation of an extra member to the ensemble. TC enriches the ensemble by improving the coverage in phase space and allows the analysis with EnVar to be drawn locally close to the ensemble mean. The proposed TC is found to be beneficial, especially with 4DEnVar.

The results of parallel tests in NDAS over extended periods demonstrated that the ensemble from the global GFS/EnKF system could be used to improve the regional NAM analysis with an EnVar system. The regional EnVar with TC effectively combined contributions from both the global and the regional systems to produce initial conditions for the NAM. With a substantially positive impact on the forecasts over those initiated with 3DVar, 3DEnVar with a global EnKF ensemble was implemented in August 2014, and the parallel test results indicated a 4D extension of the EnVar system provided further improvement to the forecasts.

A possible extension of this work is to apply EnVar with TC to combine a superensemble: an ensemble consisting

![Fig. 11. Cumulative RMSEs 24-h (black), 48-h (red), and 72-h (blue) forecast errors vs radiosondes for (a) geopotential height (m), (b) vector wind (m s$^{-1}$), (c) temperature (K), and (d) specific humidity fit (g kg$^{-1}$) over the CONUS for 15 Sep–15 Nov 2015. Solid lines are the control with 3DEnVar and dashed lines are experiments with 4DEnVar.](image-url)
of forecasts from more than one forecast system to produce initial conditions. In the future, in addition to the global EnKF ensemble, both the perturbations and the ensemble means from the regional ensemble (i.e., NCEP’s Short-Range Ensemble Forecast) will be used in the regional EnVar to help define the smaller-scale part of the background error covariances.

Acknowledgments. The work was completed as part of the authors’ job responsibilities at the National Centers for Environmental Prediction. We thank the support of the National Oceanic and Atmospheric Administration/National Weather Service. The authors thank Daryl Kleist, Rahul Mahajan, Russ Treadon, and Jeff Whitaker, whose contributions made this work possible. We express our appreciation to David Groff, Rahul Mahajan, and Mary Hart, whose comments helped to improve the paper. WSW thanks Eugenia Kalnay for the idea of using ensemble perturbations in data assimilation. Qin Xu and two anonymous reviewers helped to significantly improve the manuscript.

REFERENCES


