Correcting for Position Errors in Variational Data Assimilation

THOMAS NEHRKORN, BRYAN K. WOODS, AND ROSS N. HOFFMAN

Atmospheric and Environmental Research, Inc., Lexington, Massachusetts

THOMAS AULIGNÉ

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 17 April 2014, in final form 24 October 2014)

ABSTRACT

The Feature Calibration and Alignment technique (FCA) has been developed to characterize errors that a human would ascribe to a change in the position or intensity of a coherent feature, such as a hurricane. Here the feature alignment part of FCA is implemented in the Weather Research and Forecasting Data Assimilation system (WRFDA) to correct position errors in background fields and tested in simulation for the case of Hurricane Katrina (2005). The displacement vectors determined by feature alignment can be used to explain part of the background error and make the residual background errors smaller and more Gaussian. Here a set of 2D displacement vectors to improve the alignment of features in the forecast and observations is determined by solving the usual variational data assimilation problem—simultaneously minimizing the misfit to observations and a constraint on the displacements. This latter constraint is currently implemented by hijacking the usual background term for the midlevel \( u \)- and \( v \)-wind components. The full model fields are then aligned using a procedure that minimizes dynamical imbalances by displacing only conserved or quasi-conserved quantities. Simulation experiments show the effectiveness of these procedures in correcting gross position errors and improving short-term forecasts. Compared to earlier experiments, even this initial implementation of feature alignment produces improved short-term forecasts. Adding the calculation of displacements to WRFDA advances the key contribution of FCA toward mainstream implementation since all observations with a corresponding observation operator may be used and the existing methodology for estimating the background error covariances may be used to refine the displacement error covariances.

1. Introduction

In meteorology and other geophysical fluid contexts it is often useful to characterize the flow in terms of features—a hurricane, the Gulf Stream, and so forth. In discussing differences between two estimates of a feature, it is often useful to describe the differences in terms of the intensity and position of key features in the flow. Differences in the position of a feature especially result in errors with substantial and complex spatial correlations that may also be non-Gaussian (Lawson and Hansen 2005). Such errors are certainly not captured by static error covariance models. For some variables (precipitation, chlorophyll concentration, ground-water variables) anamorphic transformations can reduce the non-Gaussianity of errors (e.g., Simon and Bertino 2009; Schöniger et al. 2012). Such “Gaussian anamorphisms” can also increase the background error correlation radius, which should improve the data assimilation as this tends to increase the benefit from each observation (Brankart et al. 2012).

For data assimilation it is assumed that error statistics are Gaussian and known. Actual errors may be non-Gaussian and, at best, the error statistics are approximately known. Specifying the error statistics is itself an estimation exercise. With limited data available it is desirable to have parsimonious representation of these error statistics.\(^2\) Errors caused by a feature being

\(^{1}\) If one is taken to be the truth, then the differences are errors.

\(^{2}\) For example, errors of a particular variable, say ocean surface wind speed, may vary with location and season in some complex way, while in fact the errors depend principally on wind speed magnitude itself and then on location and season through the variability of wind speed (Hoffman et al. 2013).
incorrectly positioned can be parsimoniously described in terms of displacements. This suggests solving for the displacements, anticipating that the residual errors after aligning the background field will be smaller and more nearly Gaussian. We have developed the Feature Calibration and Alignment technique (FCA) to solve this problem (Hoffman et al. 1995; Hoffman and Grassotti 1996; Grassotti et al. 1999; Nehrkorn et al. 2003, 2014, hereafter N14). In this study the feature alignment technique (FAT) part of FCA is used within the Weather Research and Forecasting Data Assimilation system (WRFDA, which is described by Barker et al. 2012) to correct position errors in background fields. To do this, 2D displacement vectors are determined by solving the usual variational data assimilation problem, but with the displacements as the control vector. This new variant of WRFDA is denoted Displacement WRF Data Assimilation (dWRF) here. To maximize reuse of existing WRFDA software, in the prototype implementation used here, the displacement vectors make use of the storage allocated for the u and v components of the wind increment during the control vector transformation. To apply the displacements to the 3D model state without introducing large imbalances, N14 generalized the method of Hsiao et al. (2010), and this method is used here. Many additional studies (Abijevych et al. 2009, and references therein) have taken a feature-oriented approach for verification and sometimes data assimilation. For example, Beezley and Mandel (2008) make use of “morphing.” Additional examples of feature alignment in data assimilation are given by N14.

The development of dWRF presented here represents a major step toward operationalizing the FAT. Since this is the first description of dWRF in the literature, we do the following:

- Define the cost function and control vector (section 2a).
- Describe how the control vector is used in the FAT prototype (section 2b).
- Define the forecast error covariance (B) for the displacement vectors (section 2c) and justify approximating B, at least initially, by the nominal forecast error covariance for midlevel wind components.
- Describe the method of incremental solution that iteratively solves a linearized minimization problem within each nonlinear outer loop step (section 2d).
- Describe how to parallelize the algorithm (section 2e).

This work extends similar experiments using displacements reported by N14, but in the earlier work FCA was implemented as an offline, stand-alone prototype, which presents an impediment for operational implementation. Results of N14 using FCA serve as a benchmark here to validate dWRF. Therefore, some figure panels from N14 are redrawn in this article. Now, with the FAT integrated into the dWRF (see section 2), (pseudo)GPS precipitable water observations were used to determine the alignment. In an idealized experiment for Hurricane Katrina (2005) (section 3) we demonstrate how dWRF corrects gross position errors, that these corrections improve short-term forecasts, and that using integrated water vapor (IWV) observations dWRF results in greater improvement compared to using WRFDA (section 4). The development of dWRF advances the key contribution of FCA toward mainstream implementation since all observations with a corresponding WRFDA observation operator may be used to determine the optimum displacements, and the WRFDA methodology for estimating the background error covariances may be used to refine the displacement error covariances (section 5), without the need for additional code development. Furthermore, the development of dWRF here provides a roadmap for the application of FAT beyond 3DVar to 4DVar, ensemble, and hybrid data assimilations systems (section 5).

2. Formulation of feature alignment in dWRF

The integration of the FAT into WRFDA is described here. For ease of discussion, we first introduce the classic variational data assimilation system (DAS) formulation, following the notation used in Barker et al. (2004) and Huang et al. (2009), and then describe the general case of combined additive and displacement analysis increments, before describing details of the current prototype implementation, in which only displacements increments are used.

a. Control vector and cost function

We start with the standard 3DVar objective function:

\[ J = J_b + J_o, \]

with \( J_b = (x - x^b)^T B^{-1} (x - x^b) \) and \( J_o = [y - H(x)]^T R^{-1} [y - H(x)] \).

\[ (1) \]

Here \( x \) is the current estimate of the atmospheric state vector, and \( x^b \) is the corresponding a priori, or background estimate. The observations are represented by \( y \). Terms \( B \) and \( R \) are the background and observation error covariance matrices, respectively. The term \( H(x) \)
descriptions of
This choice reduces the degrees of freedom in
in the vertical), but these assumptions are not necessary.
The displacements to be horizontal and 2D (i.e., constant
model guess state
vector
as the initial first guess
initial incremental
and the observation cost function is approximated via
increment:
and the observation cost function is approximated via
where Note that the innovation vector is computed from the
date model state vector
Also in analogy to the additive increments, we define a background error covariance
matrix of displacements (B), a corresponding control variable transform
Assuming that displacement and additive errors are uncorrelated, the total cost function then be-
comes J = Jb + Jbs + Joa. A schematic overview of the
inner and outer loop computations is shown in Fig. 1.
In its current prototype implementation, dWRF only uses displacements. Then in the transformation to s we select a single midlevel from the wind storage arrays that hold the optimal displacement vector components, and scale from winds to displacements (section 2c).

EXTENSIONS
The method described above could be used as a pre-
processor. That is, the standard WRFDA could use the output from dWRF as the background for a second anal-
ysis, which would calculate additive increments. This
two-step approach uses the data twice. While this may be
justified when our knowledge of the background error
covariance is limited, the full implementation of dWRF
as described above would calculate displacements and additive analysis increments simultaneously. While the
prototype implementation of dWRF calculates only dis-
placements, the prototype solves all the problems necessary to implement the full dWRF. Note that to enable
simultaneous optimization of displacement and additive
analysis increments, the current prototype implementation
would need to be modified to allocate separate storage for the
displacement vectors, and to concatenate these with the
normal WRFDA control vector.

b. Displacement computation details
As described in section 2a above, displacement vec-
tors are assumed to be 2D, and constant with height. Even with these simplifying assumptions, changes made

\[ J_{oa} = (H[Uv^n + Ss] - d)^T R^{-1} (H[Uv^n + Ss] - d) \]  
while the nonlinear version \( S \) is used to compute the up-
dated model state vector \( x^n \). Also in analogy to the additive
increments, we define a background error covariance
matrix of displacements (B), a corresponding control variable transform \( s = B^{-1/2} v \), resulting in a background
term \( J_{bs} = v^T v \). Assuming that displacement and additive
errors are uncorrelated, the total cost function then be-
comes \( J = J_b + J_{bs} + J_{oa} \). A schematic overview of the
inner and outer loop computations is shown in Fig. 1.

In its current prototype implementation, dWRF only uses displacements. Then in the transformation to s we select a single midlevel from the wind storage arrays that hold the optimal displacement vector components, and scale from winds to displacements (section 2c).

EXTENSIONS
The method described above could be used as a pre-
processor. That is, the standard WRFDA could use the output from dWRF as the background for a second anal-
ysis, which would calculate additive increments. This
two-step approach uses the data twice. While this may be
justified when our knowledge of the background error
covariance is limited, the full implementation of dWRF
as described above would calculate displacements and additive analysis increments simultaneously. While the
prototype implementation of dWRF calculates only dis-
placements, the prototype solves all the problems necessary to implement the full dWRF. Note that to enable
simultaneous optimization of displacement and additive
analysis increments, the current prototype implementation
would need to be modified to allocate separate storage for the
displacement vectors, and to concatenate these with the
normal WRFDA control vector.

b. Displacement computation details
As described in section 2a above, displacement vec-
tors are assumed to be 2D, and constant with height. Even with these simplifying assumptions, changes made

\[ J_{oa} = (H[Uv^n + Ss] - d)^T R^{-1} (H[Uv^n + Ss] - d) \]  
while the nonlinear version \( S \) is used to compute the up-
dated model state vector \( x^n \). Also in analogy to the additive
increments, we define a background error covariance
matrix of displacements (B), a corresponding control variable transform \( s = B^{-1/2} v \), resulting in a background
term \( J_{bs} = v^T v \). Assuming that displacement and additive
errors are uncorrelated, the total cost function then be-
comes \( J = J_b + J_{bs} + J_{oa} \). A schematic overview of the
inner and outer loop computations is shown in Fig. 1.

In its current prototype implementation, dWRF only uses displacements. Then in the transformation to s we select a single midlevel from the wind storage arrays that hold the optimal displacement vector components, and scale from winds to displacements (section 2c).

EXTENSIONS
The method described above could be used as a pre-
processor. That is, the standard WRFDA could use the output from dWRF as the background for a second anal-
ysis, which would calculate additive increments. This
two-step approach uses the data twice. While this may be
justified when our knowledge of the background error
covariance is limited, the full implementation of dWRF
as described above would calculate displacements and additive analysis increments simultaneously. While the
prototype implementation of dWRF calculates only dis-
placements, the prototype solves all the problems necessary to implement the full dWRF. Note that to enable
simultaneous optimization of displacement and additive
analysis increments, the current prototype implementation
would need to be modified to allocate separate storage for the
displacement vectors, and to concatenate these with the
normal WRFDA control vector.

b. Displacement computation details
As described in section 2a above, displacement vec-
tors are assumed to be 2D, and constant with height. Even with these simplifying assumptions, changes made

\[ J_{oa} = (H[Uv^n + Ss] - d)^T R^{-1} (H[Uv^n + Ss] - d) \]  
while the nonlinear version \( S \) is used to compute the up-
dated model state vector \( x^n \). Also in analogy to the additive
increments, we define a background error covariance
matrix of displacements (B), a corresponding control variable transform \( s = B^{-1/2} v \), resulting in a background
term \( J_{bs} = v^T v \). Assuming that displacement and additive
errors are uncorrelated, the total cost function then be-
comes \( J = J_b + J_{bs} + J_{oa} \). A schematic overview of the
inner and outer loop computations is shown in Fig. 1.

In its current prototype implementation, dWRF only uses displacements. Then in the transformation to s we select a single midlevel from the wind storage arrays that hold the optimal displacement vector components, and scale from winds to displacements (section 2c).
to model variables must be formulated so that they minimize thermodynamic and dynamic imbalances. This question was addressed for model fields in the context of vortex relocation by Hsiao et al. (2010), who devised a procedure in which some of the model variables are adjusted directly, whereas others are diagnosed using the equation of state and hydrostatic relationships. We adopted their procedure, and modified it to allow displacements over varying terrain (a complication that was avoided in their study of oceanic tropical cyclones). The modified procedure implemented here consists of horizontal displacements of some model variables (wind components, potential temperature) and derived quantities (relative humidity, mean sea level pressure). These three-dimensional model fields, as well as hydrometeor mixing ratios, are displaced along the terrain-following sigma surfaces, except for potential temperature, for which displacements are computed along constant-height surfaces to avoid unrealistic changes in vertical stability near steep gradients of model terrain. A detailed description of the procedure is given in N14.

In all experiments presented here, displacement vectors are not allowed to extend beyond the domain boundaries, and the horizontal interpolation is bilinear. These restrictions could be relaxed by using enclosing global model fields beyond the WRF domain, and by choosing any interpolation that depends linearly on the gridpoint values. Some improvements in convergence might result from a method, such as a spline method, that has continuous derivatives at the gridbox boundaries.

c. Modified background functional and error covariance scaling

The experiments described here make use of the default background error statistics included in the WRFDA distribution (Wu et al. 2002, control vector option 3). The control variables for this choice consist of streamfunction, unbalanced velocity potential,
unbalanced temperature, unbalanced surface pressure, and pseudorelative humidity. Vertical and horizontal correlation length scales are modeled using recursive filters. The filter characteristics can be modified using namelist variables that allow rescaling of the error magnitudes and correlation length scales (details are available in Wang et al. 2012).

We make the implicit assumption that the patterns in the displacement covariance statistics should be similar to that for horizontal wind vectors since it is the wind itself that is advecting the coherent features in the model fields. Since dWRF uses the additive midlevel $u$- and $v$-wind component increments (in m s$^{-1}$) to store the displacement vectors (in units of grid lengths), wind background error covariances need to be rescaled accordingly for their application as displacement vector background error covariances. The effective parameters can be determined by running a simple single-observation test of the (standard) DAS. The effective background error standard deviation was determined to be approximately 2.7 m s$^{-1}$ for the test case with default namelist settings for background error variance scaling (0.25). If used without rescaling, this would correspond to a displacement length scale of 2.7 grid lengths (30 km in our test case), or approximately 81 km. The standalone prototype used a displacement length scale of 150 km, so we multiply the standard variance rescaling by a factor of $(150/81)^2$. A similar procedure is used for the length scale tuning. The single observation test revealed an effective $e$-folding length of the analysis increment of about 400–600 km, which is in rough agreement with the smallest scale allowed by the truncated spectral representation (total wavenumber 5) used by Grassotti et al. (1999).

The sensitivity of the dWRF solution to the size of the background error variance was examined by using both the rescaled (150 km) and unscaled (81 km) displacement length scale. The solution changed very little as a result. This is consistent with the findings of N14 for this case. Their results were also not very sensitive to the tunable FCA parameters or whether IWV alone or IWV and surface pressure observations were used.

FIG. 2. IWV in kg m$^{-2}$ for the NR at (a) 0600, (b) 1200, and (c) 1800 UTC. (d) The black box within the base map shows the location of (a)–(c) and all similar subsequent figures. The domain for all FAT calculations and for all WRF forecasts is a 30-km grid covering the entire area of (d). IWV values are indicated in the color scales to the right. (The color scales used for IWV and IWV errors are the same in all figures.) State boundaries and latitude–longitude lines are plotted in gray. Latitude north and longitude east are labeled within (d), axes in (a) and (c) are labeled with gridpoint index. The IWV signature of Katrina is tracking steadily to the NW and the three “×” marks in (a)–(c) (and on all similar panels of subsequent figures) indicate the centers of the IWV feature at the three synoptic times shown.
d. Linearization and incremental solution strategy

During the forward (nonlinear and tangent linear) computations, displacement vectors are extracted and applied to the 3D model fields using the method described by N14. During the backward (adjoint) computation, the cost function gradient is computed with respect to the displacement vectors and stored in the $u$ and $v$ components of the gradient data structure (Fig. 1). dWRF makes use of a fully nonlinear outer loop and a linearized inner loop (section 2a). During the inner loop minimization, linearized versions of the displacements and observation operators are used to determine the incremental displacement. The full nonlinear displacement algorithm is then applied in the outer loop to update the current estimate (i.e., the analysis) of the 3D model fields, which serves as the new linearization point for the next outer loop step (OLS). In this way the displacements are calculated as a series of incremental displacements, which are effectively applied cumulatively at each OLS of the DAS. Since this linearization makes the inner loop cost function quadratic, the conjugate gradient minimization algorithm that is used is efficient and has good convergence properties.

e. Parallel processing considerations

For applications to large domains that cannot be stored in memory available to a single compute node, dWRF is implemented using the message passing interface (MPI) and halo exchanges, allowing the use of multiple cores within a given node or across nodes.6 The native WRFDA domain decomposition is used to divide the domain into several patches to spread the load across multiple MPI processes running on separate cores (in hybrid shared-distributed memory implementations, multiple cores from a single node are assigned to a single MPI process, and the patch is further subdivided into tiles operated on by multiple threads within each process, but this does not change the following discussion). Complications arise when a displacement for the current gridpoint references a location outside the local patch.

When this occurs, separate MPI interprocess communication is used. Since the same set of displacement vectors is applied to multiple model fields (and levels), the needed communication was organized into two steps such that information related to displacement vector origin and end points is exchanged only once. During the first step of this procedure, each process compiles a list of the displacement vectors with end points in its own patch of the model domain for which model values are needed at origin points outside its patch. The processes then exchange their list of needed points to create a global list of needed origin grid points. Each process searches the global list of needed points, and extracts two sets of information: 1) a set of displacement vectors with origin points inside its own patch, and 2) a mapping

---

6 Each core is capable of executing one instance of a process or thread.
of the previously determined needed origin points (outside its patch) to the global list of needed origin points. This information is then used in multiple applications of the second step, in which interpolated values at vector origin points are exchanged between processes. Finally, halo updates are called following the computation of displacement analysis increments, so that updated halo values are available during the remainder of the DAS processing (e.g., for the computation of the innovation vector).

3. Simulation experiments

Identical-twin observing system simulation experiments (OSSEs) for Hurricane Katrina (2005) are used here and by N14 to test the utility of displacements determined by dWRF and FCA to align 3D model fields. The default WRF tutorial Katrina case, run for 24 h from 0000 UTC 28 August 2005, provides the nature run (NR). Figures 2a–c show the evolution of IWV in the NR in the subdomain containing Katrina. The full model domain is shown in Fig. 2d. For this NR, Global Forecast System (GFS) analyses provide the initial and boundary conditions, grid spacing is 30 km, there are 28 levels, and the time step is 180 s. Over this period, Katrina moves steadily toward the northwest (NW). The NR at 0600 UTC is then taken to be the background or first guess for the 1200 UTC analysis. That is, the NR at 1200 UTC is the truth. Simulated IWV observations are constructed by spatially subsampling the NR at 1200 UTC every third grid point in each direction, excluding a five-gridpoint border, to create a set of simulated observations. This provides a set of observations that are synoptically consistent with the truth (i.e., the NR at 1200 UTC) and a background showing the hurricane displaced to the southeast. The dWRF- and FCA-determined displacements are used to align the 3D background fields and then short-term forecasts are run from 1200 to 1800 UTC.

4. Results

The OSSE setup of section 3 is the same as in N14, except that they used simulated IWV and/or surface
pressure observations, and here we use simulated GPS IWV observations. The dWRF prototype was run for 10 OLSs to allow us to examine the convergence behavior.

a. Convergence behavior

We found that the dWRF solution converged after five or six OLSs (Fig. 3) with relatively small increments found after two OLSs (Fig. 4). After a small number of OLSs, little reduction in error was achieved. As is to be expected, the background cost function increases while the observation cost function decreases. Note that here $J_b$ includes displacements from all previous OLSs. In this case, the background cost function is orders of magnitude smaller than the observation cost function (and indistinguishable from the total cost function on this plot) during the first two OLS. As the displacements are applied in the outer loop, each OLS produces its own incremental displacements (Fig. 4).

The dWRF cumulative displacement vectors approximately correspond to the total displacement vectors (Fig. 5a) applied to the background field. In that sense they are comparable to the vectors derived by N14 using the Grassotti et al. (1999) methodology (Fig. 5b); dWRF displaces the vortex in the same way as found by N14 during the first two OLSs (Figs. 4a,b), but shows some differences outside of the vortex that are largely introduced in subsequent OLSs (Figs. 4c,d). This level of agreement is remarkable given that the two methods use different constraints and that FCA uses both IWV and pressure observations.

b. FAT analysis impacts

For comparison purposes we define a control analysis. For the control analysis the displacements are zero, and the analysis equals the background. Figure 6a (which repeats Fig. 2a) is the IWV background field. Because of the movement of Katrina there is a significant dipole error structure (Fig. 6b)—the background IWV is low in the area of the true storm location and high in the area of the storm location in the background. The RMS error in the plotted subdomain is $6.75 \text{ kg m}^{-2}$, with extreme errors of order $20 \text{ kg m}^{-2}$. Basic statistics for the error in the subdomain, as well as for all other difference fields plotted, are collected in Table 1.

Figure 7 shows the resulting IWV field after aligning the model state using the dWRF displacements, and the error of this IWV field. Clearly the displacements are the correct direction and magnitude. The RMS error is now under $3 \text{ kg m}^{-2}$, but extreme errors are still of order $10 \text{ kg m}^{-2}$ (Table 1). The reconstructed IWV field (Fig. 7a) has lost its symmetry—remember we are aligning the fields of potential temperature and relative humidity—but is nearly centered on the correct location. This asymmetry is seen both in the IWV itself (Fig. 7a) and in the IWV error (Fig. 7b).

It is instructive to compare these results to the results of N14. While there is agreement in the broadscale displacement fields, there are some interesting differences. Figure 5b shows the displacements determined by the FCA using the IWV data. Note the generally similarity of the displacements in Figs. 5a and 5b. Figure 8 shows the directly aligned IWV field (Fig. 8a), and its error (Fig. 8b, $A_{\text{IWV}} - O$, where $O$ is the observations). Since the FAT was applied to this field directly, the IWV errors are very small, with an RMS error of $2.64 \text{ kg m}^{-2}$ and extreme error values with magnitudes smaller than $8 \text{ kg m}^{-2}$ (Table 1). Figure 9 is similar to Fig. 8—the displacements are identical—but here IWV is determined

---

7 Here $O$ is used, not $T$. But $O$ is subsampled $T$. 

---
from the model fields that have been aligned as described in section 2. Errors are slightly larger than those of the dWRF analysis (Table 1, Fig. 7b). As was the case in Fig. 7, there is some loss of symmetry (but less here than in the dWRF solution). The center of the IWV signature of the storm is displaced to the south and there is again a definite dipole structure in the errors.

Also shown in Table 1 are analysis errors when the same set of pseudo-observations is used in the regular WRFDA ($A_w$) rather than dWRF ($A$). For this application of WRFDA, the default global background error statistics were rescaled to use a correlation length scale that is 5 times smaller than the default 500 km, and background error standard deviations 5 times larger than the default. As is to be expected, the regular WRFDA is able to fit the observations more closely than dWRF.

c. Short-term forecast impacts

Figure 10a shows the control forecast valid at 1800 UTC. The WRF forecast model, even at the coarse 30-km resolution used, is able to correctly capture the movement of the IWV signature of Katrina. Therefore, the control forecast beginning with IC that lags the NR by 6 h, continues to lag the NR by 6 h (the forecast position of Katrina in the forecast valid at 1800 UTC is close to the true location at 1200 UTC). The RMS forecast error (5.64 kg m$^{-2}$) is similar in size to the RMS background error (6.75 kg m$^{-2}$, Table 1). As a result, the forecast error is similar to the difference of truth at 1200 and 1800 UTC. Similarly the control forecast initial conditions are equal to the truth at 0600 UTC and the initial conditions error is the difference of truth at 0600 and
1200 UTC. Finally, since the core of Katrina in the NR is basically tracking steadily toward the NW, these difference fields are very similar in structure, but shifted along the storm track (cf. Fig. 10b to Fig. 6b).

In the short-term dWRF forecast, by 1800 UTC, the IWV feature has regained its symmetry (Fig. 11a) and is closer to the correct location. Notably, the forecast IWV errors ($F_a - T$, Fig. 11b) are smaller, with an RMS of 3.91 kg m$^{-2}$ (Table 1) and less organized than the aligned initial conditions IWV errors ($A - T$, Fig. 7b). Since the model fields, not the IWV field, are aligned, the resulting IWV field loses its symmetry. However, errors are reduced (cf. Fig. 6b). In a similar fashion, in the short-term FCA forecast, IWV regains a symmetrical structure by 1800 UTC (Fig. 12), and shows comparable overall quality (Table 1): RMS errors for the FCA forecast are slightly smaller (3.57 kg m$^{-2}$), but extreme values are larger (of order 15 kg m$^{-2}$). Interestingly, the forecast initiated from the WRFDA initial conditions ($F_w$) has much larger errors at 1800 UTC (larger even than the control forecast), and shows virtually no change from the control forecast vortex position (Fig. 13), even though it had the smallest errors in IWV at the initial time. This is also evident in the time series of lowest level perturbation pressure shown in Fig. 14 for two points along the vortex track, which shows almost no difference between control and the WRFDA forecast, whereas the dWRF forecast is much closer to the nature run. The alignment produced by dWRF does generate some imbalances that result in wavelike motions during the initial phase of the forecast, as is evident from the noisier appearance of the pressure trace for the dWRF forecast. However, the dWRF forecast is able to maintain the shift in vortex position, while the WRFDA adjustments to IWV do not result in a corresponding correction to vortex position.

d. Sensitivity experiments

The results shown so far used a highly idealized simulation experiment, chosen to demonstrate the potential of the method and compare its performance to the stand-alone FCA implementation. To examine how much these results depend on the particulars of the experimental setup, we have repeated the WRFDA and dWRF experiments using IWV observations with less coverage: using pseudo-observations every third (as before), fifth, or tenth grid point, and excluding observations from a boundary zone of 5 (as before) or 10 grid
The same set of experiments was also repeated using pseudoradiosonde observations (containing surface pressure and profiles of temperature, relative humidity, and horizontal wind). The results are summarized in Table 2, which shows the error statistics of the analysis over the same subdomain as shown in Table 1 and the previous figures. In cases with dense observation spacing, the WRFDA analysis produces smaller analysis errors of IWV than dWRF as was the case for the baseline case shown in Table 1. For the pseudoradiosonde experiments with lower observation density, the displacements introduced by dWRF during later OLS lead to increasing errors when compared to truth, and best results are obtained at OLS $< 10$, as shown in Table 2. For these dWRF experiments, the fit to the observations also decreases in later OLSs. This problem is seen in other settings for strongly nonlinear observation operators, such as for cloudy radiances. In fact there is no guarantee that trying to solve a nonlinear problem through a sequence of linearized problems will converge. If the problem is strongly nonlinear and we perform too many inner loop iterations for each OLS, then we may “overfit” the linearized problem and diverge from the full nonlinear solution. Examination of displacement vectors for those cases (not shown) also revealed nontrivial displacements in the data-void boundary regions, which also points to possible overfitting of the observations, and the need to improve on the simple rescaling of global background error statistics to constrain the dWRF solutions for the more general situation of incomplete data coverage.

The RMS error statistics of the short-term forecasts (Table 2) show the dWRF forecasts to be more skillful than those from WRFDA, except for the case of dense radiosonde coverage. (For ease of comparison across cases with different optimal OLS, all forecasts were initialized from the result of OLS 3.) This shows the potential for dWRF to make maximum use of incomplete observations to correct position errors.

5. Discussion and conclusions

The feature alignment technique (FAT) is used to correct position errors in background fields in dWRF by solving the usual WRFDA variational data assimilation problem, but with 2D displacement vectors as the control vector. The full model fields are then aligned following the generalization of the method of Hsiao et al. (2010) by N14. This procedure minimizes dynamical imbalances by displacing conserved or quasi-conserved quantities and then recalculating needed derived...
quantities; dWRF is thus a refinement of WRFDA that instead of finding optimal additive increments to the background, determines the optimal field of displacements that aligns the background to the data. Katrina OSSEs show the effectiveness of these procedures in correcting gross position errors and improving short-term forecasts.

The dWRF represents a major advance over the stand-alone implementation of FCA in N14. First, dWRF is not restricted to comparing 2D fields. Second, it can simultaneously make use of all observation types supported by WRFDA—no new observation operators are required. Third, it is conceptually straightforward to extend the control vector to include both additive increments and displacements and, thereby, to solve for both simultaneously instead of first applying the FAT as a stand-alone preprocessor. Fourth, the WRFDA estimation of error covariances can be used to refine the displacement error covariances. [In the current work, a simple rescaling of the wind background error covariances was used to characterize the displacement error covariances (section 2c).] Fifth, the parallel implementation of the FAT within dWRF is computationally efficient and scalable. By reducing demands on per-processor memory and overall wall-clock time it is feasible to apply the algorithm to large model domains and/or large numbers of cases.

The approach to implementing FAT in dWRF could be applied to other variational and/or ensemble data assimilation systems in a straightforward manner. Extension of dWRF to 4DVar and ensemble or hybrid approaches is straightforward. The current code, although only tested in the 3DVar context, should work as well with 4DVar, hybrid 3DEnVar, and 4DEnVar. Indeed, the inclusion of FAT happens at the control variable transform stage, regardless of whether the model TL/adjoint is used (4DVar) or if a hybrid ensemble/stationary background error covariance is used. A similar redefinition of the analysis state vector (displacement instead of additive increments) could be developed for the EnKF. This would simply require the use of correlations between the observations and the displacement field, instead of correlations between the observations and the model state. The generation of an accurate flow-dependent ensemble of displacement fields will be the subject of future work.

Comparison with the Katrina case identical-twin OSSE of N14 shows that dWRF performs properly and generally reproduces the major features of the FCA solutions. By design the stand-alone FCA displacements

![Fig. 12. The FCA-aligned (a) forecast and (b) forecast error of IWV at 1800 UTC.](image1)

![Fig. 13. The WRFDA (a) forecast and (b) forecast error of IWV at 1800 UTC.](image2)
greatly reduce the error in the aligned IWV field. In comparison, the alignment of the 3D model fields results in larger residual IWV errors, but these errors are still greatly reduced compared to the control case. When the aligned model fields are used as initial conditions in short-term forecasts, the forecasts of IWV are much improved relative to control, with the greatest improvement seen in the dWRF experiment. Forecasts from regular WRFDA solutions using the same pseudorecords show much less improvement, except for cases with dense coverage by radiosondes.

Future work will investigate various aspects of displacement increments: changes to the background error specification (beyond the basic rescaling exercised so far), the optimal combination of displacement and additive increments, the selective use of different observation types for one or the other, the use of the principle of time continuity for the displacements, and the use of vertically varying displacement vectors.

Acknowledgments. Partial funding for this work was provided by the U.S. Air Force Weather Agency. The final version of this manuscript was greatly improved by the constructive criticism of three anonymous reviewers.

REFERENCES


