Phase-correction Data Assimilation and Application to Storm-scale Numerical Weather Prediction. Part I: Method Description and Simulation Testing

Keith A. Brewster

Center for Analysis and Prediction of Storms

Norman, OK

Submitted to Monthly Weather Review

May 2001

Revised, February, 2002

Corresponding author address:

Keith Brewster

Center for Analysis and Prediction of Storms

University of Oklahoma

100 E Boyd, RM 1110

Norman, OK 73019

kbrewster@ou.edu

Abstract

An objective method of determining and correcting phase or position errors in numerical weather prediction is described and tested in a radar data observing system simulation experiment (OSSE) addressing the forecasting of ongoing thunderstorms. Such phase or position errors are common in numerical forecasts at grid resolutions of 1-to-20 km (meso-γ scale). It is proposed that the process of correcting a numerical forecast field can be simplified if such errors are addressed directly. An objective method of determining the phase error in the forecast by searching for a field of shift vectors that minimizes a squared-error difference from high-resolution observations is described.

Three methods of applying a phase error correction to a forecast model are detailed. The first applies the entire correction at the initial time, the second in discrete steps during an assimilation window, and the third applies the correction continuously through the model's horizontal advection process.

It is shown that the phase correction method is effective in producing an analysis field that agrees with the data yet preserves the structure developed by the model. The three methods of assimilating the correction in the forecast are successful, and a long-term positive effect on the thunderstorm simulation is achieved in the simulations, even as the modeled storms go through a cycle of decline and regeneration.

1. Introduction

During the past 20 years thunderstorm modeling has advanced as a tool for studying the dynamics and morphology of thunderstorms. At the same time, new operational mesoscale instrument systems such as the wind profiler demonstration network, surface mesonets and the WSR-88D (Weather Surveillance Radar-1988 Doppler) radar network (Klazura and Imy 1993) are increasing our ability to monitor the atmosphere at the meso-β- (20-200 km) and meso-γ-scales (2-20 km). Although the WSR-88D radar network was primarily intended to provide precipitation monitoring and to allow for short-term (0-30 min.) warning of weather hazards, the radial wind and hydrometeor reflectivity data from the radars also represent rich sources of data that can potentially be used to improve weather forecasts.

The improved modeling capability, data availability, and increasing computer processing power lead to the hope of increasing the accuracy of short term (0-12 hour) forecasts of precipitation, thunderstorm initiation, and severe weather hazards (Lilly 1990; Droegemeier 1990). Quantitative precipitation forecasting and prediction of high impact weather events (such as severe thunderstorms, hurricane landfall, flooding rains or aviation hazards) are of particular interest.

A property of Doppler radar data is that they are incomplete, that is, they only provide the radial component of the wind (the radial velocity) and a measure of the total hydrometeor density (the reflectivity) within each sampled volume of the atmosphere.

We seek to exploit the information in radar data and, where possible, to deduce the unmeasured wind components and other atmospheric variables. Some schemes devised to use radar data involve retrieval of model fields not observed by radar by making use of time-series of radar data and building on a horizontally homogeneous background state (e.g., Shapiro et al., 1995, Weygandt et al., 2002a, 2002b).

Other approaches follow analysis and assimilation procedures developed for larger-scale numerical weather analyses by applying local corrections to forecast fields based on surrounding observations (e.g., Bergthórssen and Döös 1955; Bratseth 1986). Data from storm-scale observing systems, including radar, can be added as increments to a background forecast, either directly (e.g., Brewster 1996) or through three- or four-dimensional variational approaches (e.g., Sun and Crook, 2001, and Gao et al., 2001). A difficulty in using storm-scale forecast fields as analysis background fields arises from the fact that small scale features (such as fronts, thunderstorms or outflow boundaries) may have been properly developed in the model but are incorrectly positioned in some way as a result of an error in their propagation velocity or the mispositioning of a mesoscale feature. The objective analysis, retrieval or assimilation system must then remove the incorrectly forecasted disturbances and rebuild them in the proper location. That is a difficult task when the features are temporally or spatially intermittent and an accurate estimate of the background error is lacking at such scales. The scarcity and incompleteness of data at small scales adds to the difficulty. Alternatively, one can try to correct the phase errors in the background forecast. In other words, the radar-observed variables can be used to determine the position errors at a given time, and all the variables of the background field are then adjusted for those errors. This process will, for the most part, preserve the atmospheric structures generated by the forecast model. Furthermore, without relying on any inductive or retrieval processes necessary for time-dependent schemes, an incomplete measure of the wind field, such as the Doppler radial velocity, can have a much larger impact on the other model variables. The phase-corrected fields can then be used as initial conditions for the model, as pre-conditioned backgrounds for objective analysis techniques and retrievals, or as the first guess in iterative three- or four-dimensional variational data analysis or assimilation schemes.

The concept of directly addressing location errors has already received some attention. Thiebaux et al. (1990) noted that forecast position errors of major features were negatively affecting quality control decisions in the National Center for Environmental Prediction (NCEP) operational models. Quality control decisions were improved when the background fields were adjusted for possible position errors before computing the observation-forecast residuals used in those decisions.

Mariano (1990) has developed an analysis technique based on the consideration of position errors in the Gulf Stream current and its ring vortices. Called "contour analysis", Mariano's technique involves the contouring of two or more fields, the adjustment of the contours to their optimal location and the inversion of contour fields to obtain gridded data. It is intended for melding of the location information from different sources (e.g., a forecast and satellite map of gulf stream meanders) and requires a complete specification of the field from these sources.

The idea of phase-correcting assimilation for storm scale models was presented and tested in some simple one-dimensional models by Brewster (1991). It was found that knowledge of phase errors determined from one variable in the dynamic system could effectively be used to update the system, correct errors in all the fields and improve the forward forecast of the system state. The present work extends this idea and tests it in three-dimensional simulations.

Hoffman et al. (1995) developed a method for expressing the error in a forecast in terms of a "distortion error" consisting of a displacement of each analyzed field (compared to the forecast) and an amplification of the anomaly. Correlation maximization and RMS error reduction were compared as a metric for finding the displacement error as in Brewster (1991). They also show how this could produce new analyses from forecast fields using data to provide the correction. The method was demonstrated using single fields of data such as SSM/I precipitable water (or integrated water vapor, IWV), ERS-1 backscatter wind data, and a 500-hPa geopotential height field. No forecast experiments were performed.

Hoffman and Grassotti (1996) extended the work of Hoffman et al. (1995) by performing distortion analyses on the precipitable water fields of European Center for Medium Range Weather Forecasting (ECMWF) analyses (on a 2.5° X 2.5° archived grid) made without those data. They used the position and amplitude corrected fields to make new forecasts. That study focused on fronts and marine tropical cyclones in the otherwise data-sparse South Pacific Ocean, on those large-scale grids. They minimized a functional formed by

$$J = J_r + J_d + J_a \tag{1}$$

where J_r is a specific measure of the fit of the distorted field to the data, J_d is a measure of the smoothness of the displacement vectors, and J_a is a measure of the magnitude of the displacement vectors. The functional is minimized after transforming the data to spectral space. The spectral transform and the smoothness constraints are designed to focus the displacements on the large-scale displacement errors. Hoffman and Grassoti showed that the distortion analysis improved the ECMWF analyses by decreasing the differences between the analyses and satellite IWV observations. Further, subjective improvements in the shape and scale of frontal features were noted.

Alexander et al. (1998) used a technique termed “digital warping” to transform water vapor fields using SSM/I IWV observations and the Pennsylvania State University-National Center for Atmospheric Research mesoscale model (MM5). The scheme is based the on warping or “morphing” technique sometimes employed in animated motion picture production to transform images from one character or shape to another. It requires manually determined pairs of “tie points” which connect areas of similar features between the observation and background field. The meteorological fields are then transformed from the original to a warped field by applying a cubic transformation of the grid locations and minimizing the difference in the cubic translation with respect to the translations dictated by the tie points. For each test case the authors used 130 tie points to describe the warping of IWV fields in a mid-Atlantic frontal system, a North Atlantic cyclone, and the 13 March 1993 “superstorm” cyclone in the Southeast United States. The transformation was applied to all of the model’s moisture fields, and the resultant fields were then used as part of the initial conditions in the MM5 model. A forecast made with warping was judged superior to a forecast made with no adjustments and to a forecast employing nudging assimilation. Improvements were noted in fields such as the surface pressure even though the transformation was only applied to moisture variables.

Huo and Strum (1999) have developed a graphical user interface (GUI) to manually identify phase corrections to apply to the Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). Their scheme can also remove spurious storms that are identified by an analyst. Early results show that short-term improvements can be made to forecasts by applying a uniform-in-height phase correction and removal of spuriously generated storms. Testing was done on a model with 24-km grid spacing that employed a convective parameterization scheme. In their case, the advantage due to the phase correction lasted about 5 hours before being overcome by other forecast errors. However, as in the digital warping scheme, manual intervention is required by this technique to identify the forecast phase error.

In this work, an automated system of identifying and correcting forecast displacement errors using high-resolution data is developed. The method is suitable for gridded data without requiring spectral transforms, and is flexible for direct and indirect observations. The method is described in Section 2 and tested by means of an observing system simulation experiment (OSSE) described in Section 3. Results are presented in Section 4, and additional discussion is given in Section 5. In the Part II companion paper (Brewster 2002) the method is applied to a thunderstorm outbreak case using observed mesoscale surface network and radar data.

2. Phase Correcting Technique

2.1 Introduction

Details of a method for correcting phase or position errors in high resolution, non-hydrostatic, numerical weather forecast models are presented in this section; they include the process for finding the three dimensional field of horizontal phase correction vectors and three methods for applying the corrections to the forecast fields.

2.2 Estimation of Phase Error

The phase correction method seeks a local translation, described by a field of translation vectors, $\delta\vec{x}$, that is to be applied to the forecast field in order to shift and distort it to match the observed data. Several methods might be appropriate for finding the best shift vector; the one used here is a minimization of the mean square difference from the observations. In previous work (Brewster 1999) it was found that this produces better results in simple numerical systems than a method of maximizing correlation. This is consistent with our expectations, because, for example, a correlation-maximizing scheme would not be able to pick-up a position error in a field with a linear trend.

To minimize the weighted-squared differences between the observations and the background, an error sum is formed over each test volume in the domain containing observations:

$$J[\delta\vec{x}] = \frac{s\left(|\delta\vec{x}|l^{-1}\right)}{N_\alpha} \sum_{j=1}^{n_v} \alpha_j \sum_{i=1}^{n_o} \frac{\left\{H\left[\overline{F}(\vec{x}_i + \delta\vec{x})\right] - o_{i,j}(\vec{x}_i)\right\}^2}{\sigma_{i,j}^2} \tag{2}$$

where $o_{i,j}$ is an observation, $\vec{x}_i$ is the observation location, $\delta\vec{x}$ is the horizontal displacement vector, $\sigma_{i,j}^2$ is the expected observation variance, which, in general, is a function of variable, data source and height. $\overline{F}$ is the forecast field smoothed by a 9-point filter in two horizontal dimensions (the smoothing is done to avoid fitting small-scale noise to the observations). *H* represents a transformation, if necessary, from the forecast variables to the observed quantity. For example, the transformation could be a projection of wind vectors to the radial directions of a radar, the conversion of liquid water and ice to reflectivity, or an integration of moisture to obtain precipitable water. Each of the n_v variables is weighted by α, which may account for the total number of observations of each type in the region being considered (n_o) and/or the anticipated usefulness of a particular variable in determining the displacement error.

The multiplier on the right hand side, *s*, is a distance-dependent function that serves as a penalty for straying too far from the original position. The function used here follows from Thiebaux et al. (1990), the inverse of the SOAR function:

$$s\left(|\delta\vec{x}|l^{-1}\right) = \frac{\exp\left(|\delta\vec{x}|l^{-1}\right)}{\left(1 + |\delta\vec{x}|l^{-1}\right)} \tag{4}$$

where *l* is a length scale parameter. For the work presented here *l* is set as:

$$l = 0.5\sqrt{L_x^2 + L_y^2}$$

where L_x and L_y are the lengths of the sides of the summation test volumes (discussed below) in the x and y directions, respectively. Finally, the functional includes a normalization factor, N_α:

$$N_\alpha = \sum_{j=1}^{n_v} \sum_{i=1}^{n_o} \alpha_j \quad (5)$$

The normalization factor is to account for the fact that observations may drop out of the calculation of J in the special case where the test volume is near the domain boundary and the test shift takes the observation outside the forecast domain. This could otherwise artificially decrease J just because the total number of observations in the sum is decreasing, and could potentially lead to a false minimum in the functional.

For the purpose of matching with radar data, the model hydrometeors are converted to reflectivity factor, Z, based on the following formula, combining formulae found in Kessler (1969) and Rogers and Yau (1989):

$$Z = 1.73 \times 10^4 \left(10^3 \overline{\rho} q_r\right)^{7/4} + 3.8 \times 10^4 \left[10^3 \overline{\rho}\left(q_s + q_h\right)\right]^{2.2} \quad (6)$$

where q_r, q_s, and q_h are the forecast rainwater, snow, and hail concentrations (kg/kg), respectively, and $\overline{\rho}$ refers to the horizontal mean atmospheric air density in kg/m^3. The Z is converted to reflectivity in dBZ, the variable and units of the radar data, using

$$Z_{dBZ} = 10 \log_{10}(Z)$$

The model Cartesian winds (*u, v*) are projected to the radial directions for comparison to the observed radial velocity data from radar, using the azimuth of the radial direction and the local slope of the radar beam, following a curving ray path:

$$v_r = \frac{dh}{dr} w + \frac{ds}{dr}(u \sin\phi + v \cos\phi) \quad (7)$$

where v_r is the projected radial velocity, r is the slant range (ray path distance), h is the height above the curving earth's surface, s is the distance along the earth's surface and ϕ is the radar azimuth angle. Differentiating the equation for the height of the radar beam from the four-thirds-earth ray path equations (from e.g., Doviak and Zrnić 1984) one can derive the local slope of the ray path:

$$\frac{dh}{dr} = \left(r + \tfrac{4}{3} a \sin\theta_e\right)\left[r^2 + \left(\tfrac{4}{3} a\right)^2 + \tfrac{8}{3} ar \sin\theta_e\right]^{-1/2} \tag{8}$$

where a is the earth's radius and θ_e is the radar elevation angle. From the grid point's location with respect to the radar and the four-thirds-earth model for the ray path one can obtain the needed elevation angle and slant range:

$$\theta_e = \tan^{-1}\left[\frac{H \cos S - k_e a}{H \sin S}\right] \tag{9a}$$

$$r = \frac{H \sin S}{\cos\theta_e} \tag{9b}$$

where

$$H = \tfrac{4}{3} a + \left(z - z_{radar}\right) \quad \text{and} \qquad S = 3s/4a$$

where z_{radar} is the height of the radar.

The algorithm to find the shift field that will minimize the error sum proceeds by dividing the domain into overlapping test volumes. The size of the volumes is flexible, and is determined by considering the data density, the scale of displacement error one is seeking to correct, and the size of any discrete precipitation systems resolvable by the model. One needs to have sufficient data in the volumes to make a stable calculation of the error sum without being dominated by any single observation. In contrast to that consideration, it is desirable to consider volumes that are as small as possible to accurately position the significant storm-scale features, for instance, down to the scale of individual thunderstorm cells. Because of this conflict, an iterative approach using a cascade of test volume sizes has been developed, and is illustrated in Fig. 1. First, the large-scale phase errors are found using large-scale data and large test volume sizes, and then the displacements are refined to correct small-scale errors using smaller test volume sizes and high-resolution data. In the context of correcting a mesoscale forecast including thunderstorms, one can think of the process as correcting the position of a squall line first, then adjusting the relative location of individual cells.

For each of the test volumes, the algorithm gathers the data that lie within the volume. It then seeks the displacement in location of the grid that will minimize the error sum (Eq. 2) between the observations and the shifted grid. The algorithm considers horizontal displacements of the forecast grid in intervals of three grid lengths in both x and y directions and for each of the test volumes. Searching is refined to single grid length shifts within a square five grid lengths wide centered on zero offset and in any other area where the value of the test function has not exceeded twice the minimum value found within five grid lengths of zero offset. Where there is overlap among test volumes, the components of the shift vectors found in each of the overlapping volumes are averaged. The amount of overlap (in the horizontal and vertical) is allowed as a user-specified quantity, but overlaps of 50% in the horizontal and 20% in the vertical seem to provide sufficient continuity. Where there are no data in an individual test volume $\delta\vec{x}$ is set to zero for that particular volume (before averaging with overlapping regions and subsequent smoothing).

This brute-force search method is chosen over an iterative minimization procedure utilizing local *J* gradients with the anticipation that the presence of individual thunderstorm cells in the radar data and in the model would not, in general, yield a smoothly varying field of the error sum, and may have multiple local minima and maxima. To illustrate this, sample square error topologies are presented in Fig. 2 for a real-data case using radar and surface mesonet data. In the figure, each plotted number represents the square error statistic divided by 10 for each offset in i,j space. Topologies for two sample test volumes are displayed; each contains some radar data, but the topologies are quite different. For volume "A", the field is well behaved with a well-defined trough containing a single local minimum (at I offset=5,4). For another volume, volume "B", the topology is more chaotic. There are a number of small values in a region also displaying a number of large values as "noisier" thunderstorm data find their match within the forecast grid. The minimum is found at i,j offset of -5,-9.

The cause of the irregular pattern is shown schematically in Fig. 3. The square error field is relatively flat and largely governed by the distance penalty where the data in the sample volume are being compared to regions of the forecast without thunderstorms (Fig. 3a). As the offset testing approaches a thunderstorm in the forecast grid (Fig. 3b) there are places where the forecast thunderstorm perturbation winds may oppose those observed creating high values of *J*. A further shift causes the forecasted reflectivity and wind to line-up with the observed fields and produces the sought-after minimum (Fig 3c).

Despite the lack of sophistication in the search scheme, the searching is computationally very efficient. The program, when used in a real case including radar data with four search iterations uses less computer time than the objective analysis code.

The result of this process is a three-dimensional field of horizontal shift vectors. These are interpolated to the model grid from the test volumes and smoothed to provide a smooth transition across the entire domain. Six passes of a 27-point 3-dimensional filter are used to ensure smoothness.

2.3 Applying the Phase Correction

Once a smoothly varying field of $\delta\vec{x}$ is determined, it is necessary to correct the model forecast fields for this error. Three different phase correction schemes have been devised to accomplish this: a single-step shift, a multiple-step phase correction assimilation (shifting in small steps over a period of time), and the introduction of a phase-correcting pseudo-wind in the horizontal advection terms of the model's prognostic equations.

The transformation can be applied to one or more model variables. When the transformation is applied to multiple variables, the structure and local balances among the variables generated or maintained by the model will be nearly preserved, the exception being those involving local gradients, which may have been modified by the horizontally varying shift vectors. The translation is applied equally to all prognostic atmospheric variables. To avoid upsetting the hydrostatic balance in the base state pressure and temperature, the translation is applied to the perturbation pressure and temperature; otherwise the translation is applied to the total fields.

With the single-step adjustment method, the field of shift vectors is used to distort the fields in one step, before execution of the forward forecast. Due to the averaging done in the overlap regions and the smoothing done in the last step, the final field of shift vectors will not be of unit grid lengths. The shift therefore involves an interpolation process to find the value of variables at locations not on model grid points. In the work presented here, a second-order interpolation scheme is employed for this task.

To ease the transition of the model to the phase-corrected solution, the phase correction can be applied in multiple small steps during the model integration rather than all at once at the initial time. The total shift vector is divided into *N* equal parts, where *N* is the number of shift intervals during the assimilation period. For the thunderstorm cases presented here, the adjustment is done over a short time frame, 10 minutes, with a shift applied every minute, for a total of 10 equal steps. In general the phase assimilation time window should be no more than the life cycle of the individual features being shifted, so for larger-scale applications, the window could be significantly longer.

Finally, the translation can be achieved by adding an artificial wind to the advection terms in the model equations. The "pseudo-wind" then continuously advects the structures into their proper position during the assimilation time window. For example, in the Advanced Regional Prediction System (ARPS, Xue et. al, 2001, Xue et al., 2000, Xue et al., 1995), the forecast model used for this study, the horizontal advection of a scalar perturbation, Q', is written as:

$$\frac{\partial}{\partial t}(\rho^* Q')_{hadv} = -\overline{u^* \delta_\xi Q'}^{\xi} - \overline{v^* \delta_\eta Q'}^{\eta}$$

where

$$\rho^* = |J_3|\overline{\rho}, \qquad u^* = \rho^* u, \qquad \text{and } v^* = \rho^* v$$

δ_ξ is a finite difference operator in the east-west direction, δ_η is the finite difference operator in the north-south direction, $\overline{\rho}$ is the horizontally averaged air density for the base-state and J_3 is the Jacobian representing the stretching of the vertical coordinate in the vertical direction. The overbar denotes averaging in the indicated direction to account for grid staggering. In the application of the pseudo-wind scheme, the *u* and *v* wind components are replaced by

$$\hat{u} = u + \frac{\delta x}{\delta T_n}, \qquad \text{and } \hat{v} = v + \frac{\delta y}{\delta T_n} \tag{10}$$

where δx and δy are the *x* and *y* displacement components, respectively, and δT_n is the length of the assimilation time window. It is desirable for the assimilation time window to be large enough that the pseudo-wind does not dominate the physical wind in the advection process. For short-lived features, this may not be possible however.

Employing this technique in a model with an explicit finite-difference representation of advection carries the risk of violating the Courant-Friedrichs-Levy (CFL) condition, that is, the stability limitation on the time-step size must now include the effect of the pseudo-wind, which could increase the effective maximum wind (or phase speed). The maximum pseudo-wind or the duration of the assimilation window could automatically be adjusted, if necessary to prevent such a problem. However, in the experiments performed for this work, no such adjustments were required, as the time-step limit for the model was largely determined by the small grid spacing in the vertical combined with vertical mixing and advection processes which are not affected by the addition of the pseudo-wind.

Although the addition of the pseudo-wind in the advection step might be viewed as adding momentum to the system, this energy is not available to be converted to other forms and is removed after the assimilation window, so it should not affect the energy balance of the system.

It is anticipated that the phase shifting technique could serve as a preconditioning of a background field before the application of other analysis schemes. Because changes in amplitude are not applied in this scheme, other methods will be necessary to complete the correction of the background forecast to obtain an accurate initial condition. Figure 4 is a flow chart showing how the phase correction is applied using the ARPS with the ARPS Data Assimilation System (ADAS, Brewster, 1996).

3. Observing System Simulation Experiment

In order to test the phase-error detection scheme and determine which of the proposed correction methods might work best in a numerical forecasting system, a demonstration involving a thunderstorm simulation is devised. The ARPS forecast model is run using a horizontally homogeneous environment, specified by a sounding launched at Ft. Sill, Oklahoma on the afternoon of May 20, 1977. This case has been studied from both an observational (e.g., Ray et al. 1981) and numerical modeling standpoint (e.g., Klemp et al. 1981; Klemp and Rotunno 1983). The ARPS model is known to produce a supercell thunderstorm in this environment when initialized with a single warm temperature perturbation in the shape of an ellipsoid, i.e., a warm bubble (Xue et al. 1995). In reality, a supercell thunderstorm developed on this day, produced a damaging tornado in central Oklahoma, in Del City. The case is colloquially known as, “The Del City Storm.”

The model is run using a horizontal grid spacing of 2-km and a uniform vertical grid spacing of 500 m. A domain of 98 x 98-km is used, in a reference frame moving with the mean wind. This allows the simulations to be run in a reasonably small domain as the storms won't move too far from their initial position in the model grid. For this work convection is initiated by two distinct ellipsoids of horizontal radius 8 km and vertical dimension 1.5 km, and a maximum positive temperature perturbation of 4° C. Two 2-hour simulations are performed. The two simulations differ in the location of the initial warm bubbles. In the first, "Control", the bubbles are placed at (56, 20) km and (56, 36) km and in the second, "Verification", the bubbles are located at (70, 14) km and (63,37) km. The goal of the test is to use simulated radar data to correct the control simulation for its position error with respect to the verification run.

After the model is allowed half an hour to develop thunderstorms, radar observations remapped to the model grid are simulated by "observing" the control run with a radar located 10 km south of the model domain's southern boundary (66,-10). The radar observations are simulated assuming the WSR-88D storm-mode volume scanning pattern (14 distinct elevation angles) and observations are created only where the modeled reflectivity (determined from the model hydrometeor fields) exceeds 10 dBZ. To simulate the superobbing and remapping preprocessing of real radar observations in ADAS, the data are taken on a regular grid corresponding to the model grid. Gaussian random errors with standard deviations of the 1 ms^{-1} and 2 dBZ are added to the gridded velocity and reflectivity pseudo-observations, respectively. Although the storms are well sampled, there are still gaps in the coverage due to the finite number of elevation angles in the scan pattern.

In the application of the phase correction scheme for this demonstration, the Doppler radial velocities and reflectivities are used as data. Because this test case involves features of only one scale and a single high-resolution data source, the calculation of the shift vectors was done in a single pass with the domain divided into volumes measuring 16-by-16 km in the horizontal and 3.5 km in the vertical. The phase correction volumes overlapped by 8 km on each side in the horizontal and 0.5 km in the vertical.

4. Results

Figure 5 shows a horizontal cross-section of the wind vectors and vertical velocity at 30 min at a height of 4.75 km through a sub-domain of the control and verification simulations. The bold squares in Fig. 5b, and subsequent figures indicate the approximate center of the maximum updraft in the verification run at the corresponding time. Due to the differences in the locations of the initial temperature perturbations, we see that the storms in the control simulation are incorrectly positioned relative to the verification, with the southern storm showing a larger error. There are also some small differences in storm structure in the northern storm, due to differences in storm interactions.

Phase correction vectors identified by the position error detection algorithm at the 4.75 km level are depicted in Fig. 6. Generally, the algorithm correctly diagnosed the position error, though there are some compromises in the region between the storms. This is due to the close proximity of two storms requiring different corrections and to the application of the smoother to the correction vectors which is required to make the corrections vary smoothly in space.

The effect of the single-step phase-correction procedure can be seen in Fig. 7. The locations of the updraft centers are handled very well, as are the vertical velocity extrema. The northern cell is smoothed and spread out slightly compared to its initial shape. This is especially evident in the downdraft on the south edge of the cell. The multiple-step phase-correction assimilation was applied in 10 steps over a 10-minute window from t=20 min. to t=30 min. As a result of that process (Fig. 8), the positions of the storms were adjusted fairly well, but the northern cell is not as quite as well positioned as in the single-step case (with the updraft maximum being about 3 km southeast of the maximum in the verification run). There is a slight amplification in the updraft maximum in the southern cell, and there is some loss in magnitude of the downdrafts on the flanks of the northern storm.

Figure 9 shows the storms after the third type of correction, the advection adjustment, has been applied in a window from 20-30 min. The storm locations are fairly accurate, with an error in position in the northern cell similar to that of the multiple-step assimilation experiment, there is slight amplification to the updraft in the southern cell and amplification in the downdraft on the northwestern flank of the southern cell. Overall, it appears that all three methods have done a good job of correcting the storm position during the period of assimilation.

In addition to examining the effect of the corrections at the time of observations, it is important to examine how the corrections affect the forecasts beyond that time. Figures 10a and 10b show the prediction at 2 hours for the control and verification runs, respectively. This is 90 min. after the data were provided to the experimental runs. During that period, the initial storms decayed and a single storm developed from their outflow boundaries. In the control forecast the resultant cell is dissipating at this time (at location 40,78 km) and a new cell was produced on its outflow that is exiting the domain along the eastern boundary. Both of those cells are considerable further north than the cell in the verification run.

The forecasts valid at 2 hours for the single-step correction, multi-step correction and the advection correction are shown in Figs. 11, 12 and 13, respectively. There is little difference among the experimental forecasts. They are certainly closer to the verification than the control, in position and in storm morphology. However the experiments show the storm at a more advanced stage of splitting than the verification. There is a small position error in the phase-corrected runs compared to the control – about 8-km to east. Considering the non-linearity of the processes that occurred in the interim, the difference between the phase-corrected runs and the control is remarkably small at this particular time.

Examining the time series of maximum vertical velocity, shown in Fig. 14, reveals there is a difference among the forecasts in that the phase-corrected forecasts generate the second cell about 13-15 minutes before the verification run. So, although the location error has been corrected, a timing error is introduced. Given the sensitivity of storm initiation to the strength of the low-level forcing versus the convective inhibition in the capping inversion, this result is not too surprising.

The forecasts are quantitatively compared by calculating an error score, a sum of the root-mean-square differences of eight forecast variables weighted to balance their relative magnitudes, namely:

$$S = R_u + R_v + R_w + 10^{-2} R_p + R_\theta + 10^4 \left(R_{qv} + R_{qc} + R_{qr} \right) \quad (11)$$

where R represents the root-mean-square difference of the forecast from the verification run for the subscripted variable (u, v and w wind components (ms^{-1}), pressure (Pa), potential temperature (K), water vapor (kg kg^{-1}), cloud water and rain water (kg kg^{-1}). The differences are calculated over the entire domain except the first point within the boundaries. Figure 15 shows the total error scores for the forecast experiments. The phase correction schemes are all able to reduce the forecast error significantly by the time of the inserted data. There was a rise in the error at the time of the most rapid growth of the secondary storm in the assimilation runs. This error leveled off as the verification storm strengthened. The assimilation runs maintain their advantage to t=2 hours. The forecasts have similar errors until they diverge somewhat after 1.25 h when the advection correction run seems to have a slight advantage over the other two methods.

It is of interest to consider the noise generated by the various phase correction strategies, as it may play a role in differentiating among the schemes. Figure 16 shows the mean absolute pressure change (Pa s^{-1}) for the verification run and the three phase correction methods. As expected, the advection correction, which applies the correction continuously, generates less pressure noise than the other two methods. Pressure tendency noise in all methods is less than 0.1 Pa s^{-1} within 8 minutes of the last correction application. Although the phase correction runs initiate the second generation cell earlier than the verification, the role of the adjustment noise was likely minor as early re-generation was produced in the advection-correction forecast, even though it was virtually noise-free.

5. Discussion

The radar OSSE shows that position errors in forecasts of thunderstorms can be successfully identified using Doppler radar data, even with a simple scheme for searching for relative minima. It was demonstrated that it is possible to correct such position errors with the phase-correction data assimilation scheme, and such corrections can improve the ensuing forecasts, even in a case involving storm interactions and secondary storm development. An error in timing of secondary storm development was introduced, but the experiments with phase correction did correct the errors in the location and morphology of the secondary storm.

The forecasts from the three correction methods were quite similar. The advection correction method had the least noise as measured by the magnitude of the pressure tendencies. However, the noise in the other methods was short-lived and it is not apparent that the noise had any significant effect on the subsequent forecast.

The approach of correcting the phase error by means of multiple small shifts might be improved by using higher order interpolation methods. This would allow the use of more adjustment steps (each of which carries a price of smoothing), and allow the scheme to more closely approach the result of the advection scheme without the advection scheme's risk of violating the CFL stability criterion during model integration. Similarly, a semi-Lagrangian advection method applied to the advection terms is desirable, as it could also eliminate concerns about the pseudo-wind advection scheme violating the CFL criterion.

Given that the simulations presented here were well observed (close-in simulated radar) and had one-to-one correspondence between the control and verification storm cells, it is recognized that the robustness of the method has not been thoroughly tested. The penalty for more distant straying from the background forecast locations seems to be adequate for preventing aliasing in the presence of two-storms, and the overlapping test volumes seem capable of preventing large gradients between neighboring test volumes, but more complex situations may present themselves where this is not the case. A more complex method of combining the results from the individual test volumes may be devised to better insure the resultant shift vector field is well behaved.

The simulations presented here represent a relatively easy task for many aspects of the scheme. In Part II of this work (Brewster 2002), the phase correction technique is applied to a case using a complete real data set and further applications of the technique are discussed.

6. Acknowledgments

The author is grateful for the support and guidance of Dr. Frederick Carr during the development and testing of this concept. Discussions with Drs. Alan Shapiro and Ming Xue contributed to this work.

Support for this work was from a portion of NSF grants to the Center for Analysis and Prediction of Storms, ATM91-200009 and ATM9909007, and a grant from the Federal Aviation Administration, DOT-FAA NA17RJ1227-01.

7. References

Alexander, G.D., J.A. Weinman, J.L. Schols, 1998: The use of digital warping of microwave integrated water vapor imagery to improve forecasts of marine extratropical cyclones. *Mon. Wea. Rev.*, **126**, 1469-1495.

Bergthórssen, P. and B. Döös, 1955: Numerical weather map analysis. *Tellus*, **7**, 329-340.

Bratseth, A.M., 1986: Statistical interpolation by means of successive corrections. *Tellus*, **38A**, 439-447.

Brewster, K.A., 1991: Phase correcting data assimilation. Preprints*, Ninth Conference on Num. Wea. Prediction*, Denver, CO., AMS, Boston, 610-613.

Brewster, K., 1996: Implementation of a Bratseth analysis scheme including Doppler radar. *Preprints*, *15th Conf. on Weather Analysis and Forecasting*, AMS, Boston, 92-95.

Brewster, K.A., 1999: Phase-correcting data assimilation and application to storm-scale weather prediction, Ph.D. Dissertation, University of Oklahoma, School of Meteorology, Norman, OK, 73019. Available from UMI Dissertation Archive, Ann Arbor, MI.

Brewster, K.A., 2002: Phase-correcting data assimilation and application to storm scale numerical weather prediction. Part II: Application to a Severe Storm Outbreak. *Mon Wea. Rev*., **130**, submitted for publication.

Doviak, R.J. and D.S. Zrnić, 1984: *Doppler Radar and Weather Observations*, Academic Press, Orlando, FL, 458 pp.

Droegemeier, K.K., 1990: Toward a science of storm-scale prediction. *Preprints, 16th Conf. on Severe Local Storms,* Kananaskis Park, Alta., Canada, AMS, Boston, 256-262.

Droegemeier, K.K., 1997: The numerical prediction of thunderstorms: Challenges potential benefits, and results from real-time operational tests. *WMO Bulletin*.

Gao, J., M. Xue, K. Brewster, F. Carr, and K.K. Droegemeier, 2001: A three-dimensional variational data assimilation scheme for a storm-scale model. Preprints, 14th Conf. On Numerical Weather Prediction, Orlando, FL, Amer. Meteor. Soc.,

Hoffman, R.N. and C. Grassotti, 1996: A technique for assimilating SSM/I observations of marine atmospheric storms with ECMWF analyses. *J. Appl. Metr.* , **35**, 1177-1188.

Hoffman, R.N., Z. Liu, J.-F. Louis and C. Grassotti, 1995: Distortion representation of forecast errors. *Mon. Wea. Rev.*, **123**, 2758-2770.

Hou, Z. and S. Strum, 1999: Brutal but effective data assimilation for convective weather. Preprints, *Third Conf. on Coastal Atmospheric and Oceanic Prediction and Processes,* New Orleans, LA, AMS, Boston, 279-284.

Kessler, E., 1969: On the distribution and continuity of water substance in atmospheric circulation. Meteor. Monogr., **10**, no. 32, AMS, Boston.

Klazura, G.E. and D. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. *Bull. Amer. Meteor. Soc.*, **74**, 1293-1311.

Klemp, J.B., R.B. Wilhelmson and P.S. Ray, 1981: Observed and numerically simulated structure of a mature supercell thunderstorm. *J. Atmos. Sci*., **38**, 1558-1580.

Klemp, J.B. and R. Rotunno, 1983: A study of the tornadic region within a supercell thunderstorm. *J. Atmos. Sci*., **40**, 359-377.

Lilly, D.K., 1990: Numerical predictions of thunderstorms -- has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779-798.

Mariano, A.J., 1990: Contour analysis, a new approach for melding geophysical fields. *J. Oceanic and Atmos. Technol.*, **7**, 285-295.

Ray, P.S., B.C. Johnson, K.W. Johnson, J.S. Bradberry, J.J. Stephens, K.K. Wagner, R.B. Wilhelmson and J.B. Klemp, 1981: The morphology of severe tornadic storms on 20 May 1977. *J. Atmos. Sci.*, **38**, 1643-1663.

Rogers, R.R. and M.K. Yau, 1989: *A Short Course in Cloud Physics*, Pergammon Press, 293 pp.

Shapiro, A., T. Gal-Chen, J. Zhang, L. Zhao, Q. Xu, H. Gu, C-J Qiu, I. Zawadski, S. Laroche, and J. Tuttle, 1995: Highlights from a single-Doppler velocity retrieval intercomparison project ("bake-off"). Preprints, *Sixth Conf. on Aviation Weather Systems*, Dallas, TX, AMS, Boston.

Huo, Z. and S. Strum, 1999: Brutal but effective data assimilation for convective weather. Preprints, *Third Conf. on Coastal Atmospheric and Oceanic Prediction and Processes,* New Orleans, LA, AMS, Boston, 279-284.

Sun, J. and N.A. Crook, 2001: Real-time Low-level Wind and Temperature Analysis Using Single WSR-88D Data. *Wea. and Forecasting,* **16**, 117-132.

Thiebaux, H. J., P. R. Julian and G. J. DiMego, 1990: Areal versus collocation data quality control. *Preprints, International Symposium on Assimilation of Observations in Meteorology and Oceanography*, Clermont-Ferrand, France, WMO, Geneva, 255-260.

Weygandt, S.S., A. Shapiro, and K.K. Droegemeier, 2002a: Retrieval of initial fields from single-Doppler observations of a supercell thunderstorm. Part I: Single-Doppler velocity retrieval. *Mon. Wea. Rev.*, **130**, 433-453.

Weygandt, S.S., A. Shapiro, and K.K. Droegemeier, 2002b: Retrieval of initial fields from single-Doppler observations of a supercell thunderstorm. Part II: Thermodynamic retrievals and numerical prediction. *Mon. Wea. Rev.*, **130**, 454-476.

Xue, M., K.K. Droegemeier, V. Wong, A. Shapiro and K. Brewster, 1995: *ARPS Version 4.0 User's Guide*. 381 pp. Available from Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, OK, 73019. http://www.caps.ou.edu

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Physics*, **75**, 161-193.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Physics*, **76**, 143-165.

8. Figure Captions

FIG. 1 Flow chart for determining phase error field.

FIG. 2.   Sample square error topology for RMS-based shift correction search. June 8, 1995 2010 UTC, including radar data. Plotted is the sum of square errors divided by 10 for each i,j offset. Offsets marked with a dot were not tested by the algorithm due to lack of positive results in that direction. A and B are for two separate shift data volumes.

FIG. 3 Schematic of the change in the error statistic for matching an observed and forecast thunderstorm. Low-level wind vectors and arbitrary reflectivity contours. a) forecast and observed storm widely displaced, b) searching scheme approaching correct offset, c) forecast and observation matched.

FIG. 4   Forecast update cycle combining phase error correction and analysis steps.

FIG. 5. Thunderstorm simulation at 30 min. Horizontal winds and vertical velocities (0.5 ms-1) intervals) at 4.75-km AGL. Distance scale is in km. a) Verification b) Control. Bold squares in figures indicate the location of the maximum updraft in the verification run.

FIG. 6. Phase correction vectors at a height of 4.75-km necessary to correct the control run to match the verification run. Vector units are km.

FIG. 7. Thunderstorm simulation at 30 min. (data time) for single-step phase correction. Vectors and contours as in Fig. 5.

FIG. 8. Thunderstorm simulation at 30 min. (data time) for multiple-step phase correction. Vectors and contours as in Fig. 5

FIG. 9. Thunderstorm simulation at 30 min. (data time) for pseudo-wind advection phase correction. Vectors and contours as in Fig. 5

FIG. 10. Thunderstorm simulation at 2 hours. (90 min. forecast), a) verification, b) control. Vectors and contours as in Fig. 5

FIG. 11. Thunderstorm simulation at 2 hours for single-step phase correction. Vectors and contours as in Fig. 5

FIG. 12. Thunderstorm simulation at 2 hours for multiple-step phase correction. Vectors and contours as in Fig. 5

FIG. 13. Thunderstorm simulation at 2 hours for the pseudo-wind advection phase correction. Vectors and contours as in Fig. 5

FIG. 14. Time-series of maximum vertical velocity in the domain. Bold line is verification, bold dashed line is the control run, dotted line is the single-step method, thin solid line is the multi-step method, and thin dashed line is the advection correction method.

FIG. 15. Forecast RMS error score. See text, Eq. 11, for definition. Line textures as in Fig. 14.

FIG 16. Time-series of domain-mean absolute pressure tendency for the verification run and the three forecast correction methods. Line textures as in Fig 14.