1M.1

A COMPARISON OF SENSITIVITY ANALYSES FROM THREE METHODS: AN ADJOINT, A VERY LARGE ENSEMBLE, AND A NEW METHOD OF RANDOM PERTURBATIONS

William J. Martin[1,*] and Ming Xue[1,2]
[1]Center for Analysis and Prediction of Storms and [2]School of Meteorology
University of Oklahoma
Norman, Oklahoma 73019

1. INTRODUCTION

The sensitivity of a model forecast to the model state at an earlier time is a basic calculation used in data assimilation by the adjoint (or 4DVAR) method. The adjoint of a model specifically is the derivative of a forecast model scalar to initial model fields. If the scalar field is the error between the model forecast and some set of observations, then this derivative can be used in a minimization scheme to find the initial condition which minimizes the forecast error.

Sensitivity fields are also of interest by themselves as a way to discover physical connections between forecast items of interest and possible causative mechanisms. Such sensitivity fields can also be used in targeting observations to improve a forecast and for potential weather modification.

Adjoint methods currently compete with ensemble Kalman filter methods (enKf). In the latter method, an ensemble of typically 20 to 100 members of the forward model is run with the initial condition perturbed differently in each ensemble member. Error covariances between forecast variables are calculated statistically and these covariances can then be used to assimilate data in a maximum likelihood sense. An ensemble is needed for atmospheric models because it is not practical to directly calculate forecast covariances for such complex models.

It has long been known that for linear models the adjoint and enKf methods produce mathematically the same result (e.g., Li and Navon, 2001). This is because the maximum likelihood result found from the enKf method is the same answer found by minimizing the error using the 4DVAR method. Differences between the two methods arise from non-linearity. Adjoint derivatives are, in fact, based on a linearization of the model. However, the enKf uses the full non-linear model, so that differences can arise. If the perturbations used for the enKf method are small and the time integrations also small, then the two methods can be expected to closely agree as these two conditions will tend to minimize the effects of nonlinearity. It is not necessarily clear which method gives better results in the presence of nonlinearity as the maximum likelihood weighting invoked by an enKf assumes the model is linear. Nonlinearity could lead to poor results from either method.

This paper explores the use of an ensemble for calculating forecast sensitivities in a manner analogous to the calculation of error covariances in the enKf method. In this case, the ensemble is used to calculate sensitivities of a forecast scalar to initial fields. These sensitivity fields are directly comparable to those calculated by an adjoint with differences being due to nonlinearity and random noise error in the statistical estimation. A direction comparison should shed some light on the accuracy of both methods in light of ensemble size and inherent nonlinearities. The use of random perturbations in a model ensemble is also a potentially useful new technique for calculating sensitivities by itself. This work is partly an extension of Martin and Xue (2005) in which sensitivity fields were calculated by using a very large ensemble (VLE) of model runs using defined (rather than random) perturbations. The use of random perturbations has some significant advantages over the VLE method as will be discussed below.

* *Corresponding author address:* William J. Martin, Center for Analysis and prediction of Storms, 100 East Boyd, SEC Room 1110, Norman, OK, 73019; wjmartin@ou.edu

2. THE ARPS MODEL AND TEST CASE

The initial condition analysis of the case to be analyzed is at 1800 UTC on 24 May 2002 over a region of the southern Plains centered on Oklahoma. The initial field of surface water vapor over the entire domain is shown in Fig. 1 and the initial field of 10 m wind vectors is shown in Fig. 2. This analysis was obtained using the ADAS analysis package (Xue et al., 2003) utilizing the standard National Weather Service rawinsondes, surface observations, with the NAM (ETA) 1800 UTC analysis as a background field. Additionally, special observations taken during the IHOP field program were used including aircraft dropsondes, and Kansas, Texas, and Oklahoma mesonet surface data. Many of these special observations were taken near the southwest corner of Oklahoma, north of where initiation actually occurred.

The analysis shows a generally north-south oriented dryline in the eastern Texas panhandle and an east-northeast to west-southwest oriented cold front across southern Kansas. This initial condition is integrated forward for 6 hours using the ARPS model (Xue et al., 2003) and lateral boundary forcing from the 1800 UTC NAM forecast fields. The numerical domain consists of 135 by 135 horizontal grid cells with uniform 9 km grid spacing. There are 53 vertical grid levels on a stretched vertical grid with a minimum vertical spacing of 20 m at the surface. Despite the 9 km horizontal resolution, the model is integrated using an explicit representation of convection rather than a convective parameterization. Full ice, surface, and radiation microphysical packages are employed.

Six hour forecast fields of surface water vapor, 10 m wind barbs, and total accumulated precipitation are shown in Figs. 3, 4, and 5 respectively. After 6 hours of integration, the water vapor field shows a considerable amount of strengthening of the moisture gradient, and also shows the effects of convection along the cold front in northwest Oklahoma and near the dryline-cold front triple point in the southeast Texas panhandle.

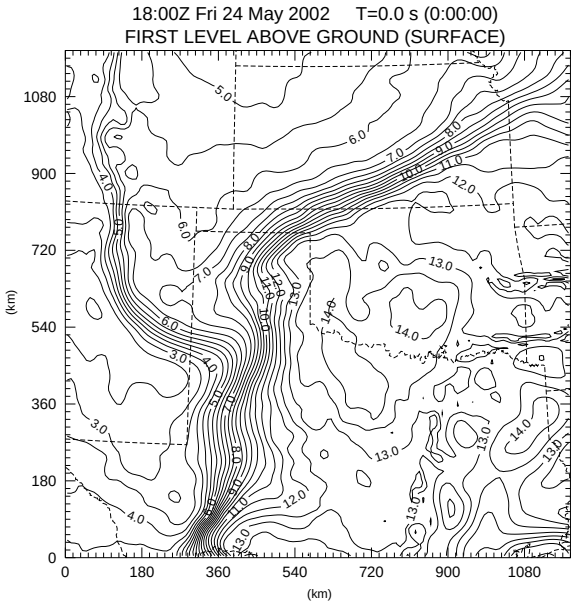

Fig. 1. Initial surface water vapor field in g kg^{-1} for 24 May 2002 at 1800 UTC. Level is at 10 m above the ground.

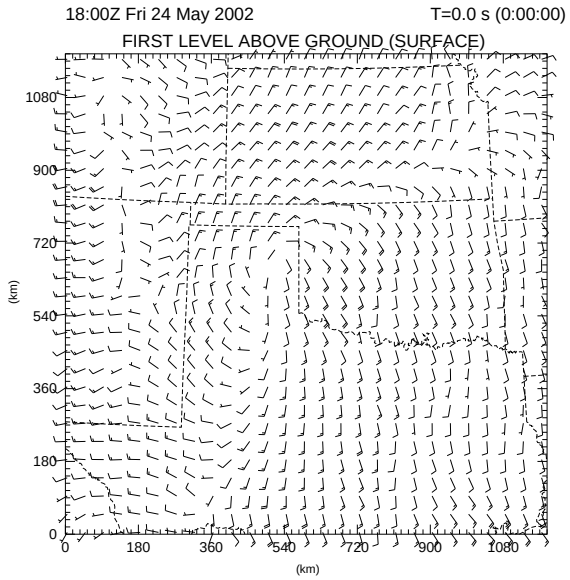

Fig. 2. Initial field of wind barbs for 24 May 2002 at 1800 UTC, 10 m above the surface. One full wind barb is 5 m/s. A half wind barb is 2.5 m/s.

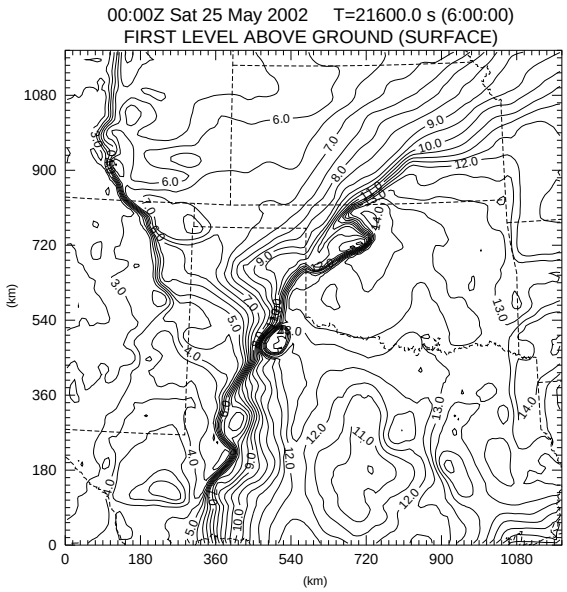

Fig. 3. 6 hour forecast water vapor field in g/kg, 10 m above the surface.

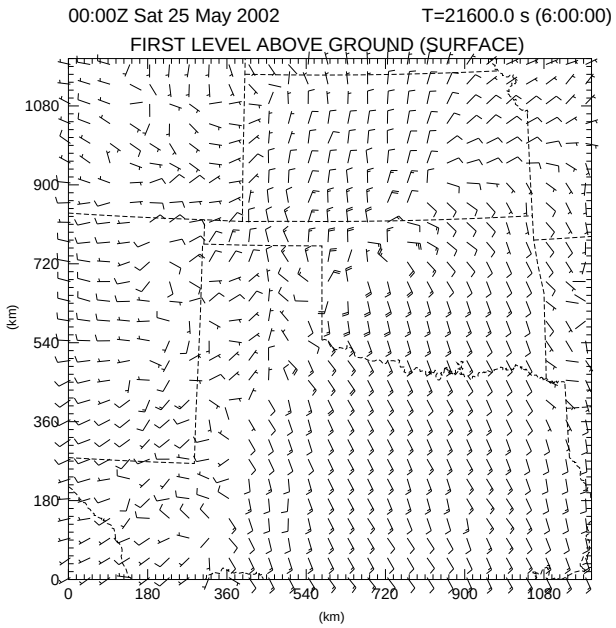

Fig. 4. 6 hour forecast field of 10 m wind barbs. Full wind barb corresponds to 5 m/s.

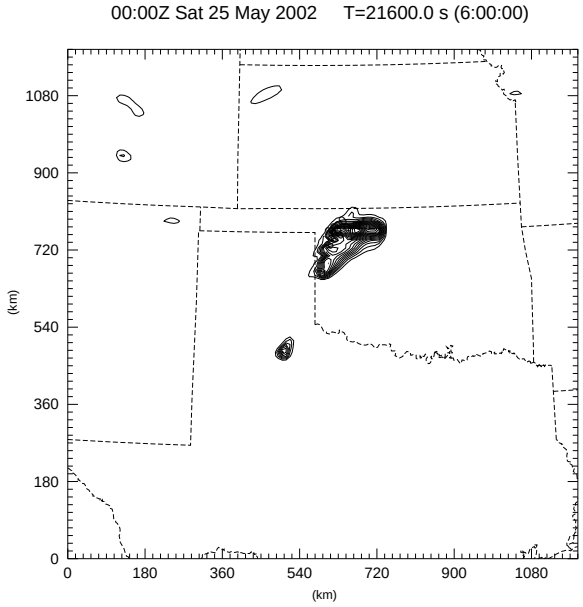

Fig. 5. 6 hour forecast of accumulated precipitation. Contour increments are 10 mm with a local maximum of 205 mm.

3. THE VLE TECHNIQUE

In the VLE technique (Martin and Xue, 2005), the initial condition is systematically perturbed with a small perturbation at different spatial locations and the model run for each perturbation. For this work, we use a boundary layer perturbation patch that is 27 km by 27 km by 1 km deep in the lowest levels of the model. This patch occupies 3 by 3 by 9 grid cells. To tile the entire 135 by 135 grid cell horizontal domain requires 2025 different patches, consequently, the forward run is executed 2026 times (once with no perturbation). This method is efficient because of the existence of large parallel computing machines in which each member of the ensemble can be run simultaneously. This work specifically used the cluster of 3000 workstations known as the "Lemieux" supercomputer at the Pittsburgh Supercomputer Center.

Because the perturbations are small, each forecast of the ensemble is nearly identical; however, measurable differences exist due to linear and nonlinear sensitivities to small perturbations. A sensitivity field is obtained by taking the difference between some defined response function (many of which can be defined) of the unperturbed forecast and each member of the ensemble. The value of the sensitivity field at some location is then simply the difference between the response function of

the control run and the run with the perturbation at that location. Figure 6 shows the sensitivity field derived from a response function, J, defined as the area average surface potential temperature:

$$J = \frac{\sum_i (\theta_i)\Delta x \Delta y}{\sum_i \Delta x \Delta y} \quad (1)$$

where the summation is taken over all the surface grid points shown in the box in north central Texas shown in Fig. 6. The sensitivity field, S(x,y), is the difference between J for the control run and J for each perturbed run divided by the initial perturbation at the (x,y) location:

$$S(x, y) = (J(x, y) - J_{control}) / \delta\theta_0 \quad (2)$$

The sensitivity is nondimensionalized by dividing the numerator in (2) by the value of the response function from the unperturbed run, J_0, and dividing the denominator in (2) by the value of the unperturbed potential temperature field. This gives:

$$s = \frac{\delta J / J_0}{\delta\theta / \theta 0} \quad (3)$$

Figure 6 shows a well-defined, though modest, sensitivity of the boundary layer temperature forecast to perturbations in the temperature field 6 hours earlier. Potential temperature perturbations of 1 K were used. The location and shape of the sensitivity contours indicates the combined effects of advection and diffusion in an expected manner. The region had southerly flow, so that the water vapor forecast depends on perturbations upstream. Also, diffusion, both numerical and physical, spreads the water vapor perturbation as might have been anticipated.

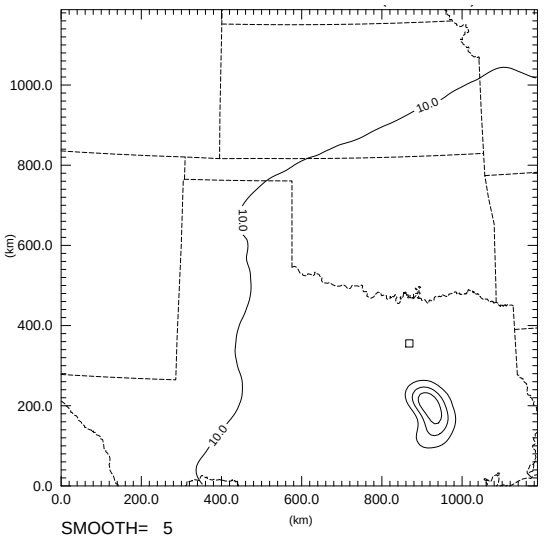

Fig. 6. Nondimensional sensitivity field for the dependence on the area average potential temperature in the box indicated to a +1 K perturbation in the initial condition. Contours are every .0005 with a maximum of .0020. The 10 g kg^{-1} isoline of moisture is drawn as a proxy for the dryline.

Figure 7 shows a more interesting sensitivity field calculated from the response function defined as the total accumulated rainfall that fell in the box indicated throughout the six hours of integration. The box surrounds the area in which rainfall initiated along the dryline. In this case, the physical mechanisms are less clear as the sensitivity is about the same on both the dry and moist sides of the dryline.

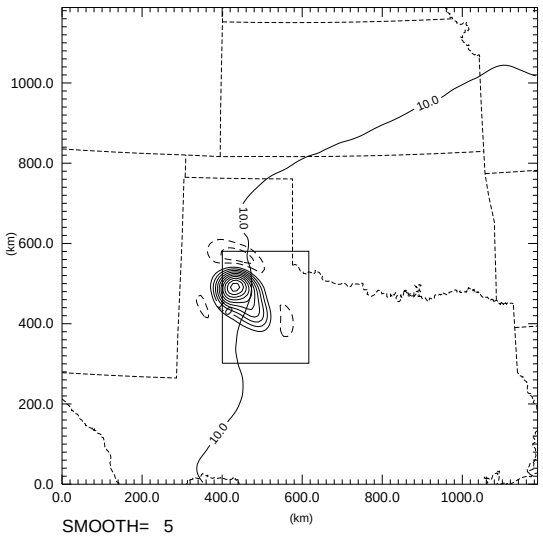

Fig. 7. Nondimensional sensitivity field for the dependence of total accumulated precipitation in the indicate box on initial +1 K boundary layer temperature perturbations. Contours are drawn every .5, with a maximum of 5.2. The 10 g kg^{-1} isoline of moisture is also drawn as a proxy for the dryline.

4. THE RANDOM PERTURBATION TECHNIQUE

To calculate a sensitivity field using random perturbations, the initial condition is perturbed randomly by the addition of fields of random numbers to one or more variables. For the test case presented here, the initial field of boundary layer potential temperature through a depth of 1 km was randomly perturbed. Forward runs from 400 different randomly perturbed initial conditions were made. As for the VLE technique, a defined response function is calculated from each of these runs. These 400 values are then linearly regressed against the perturbation magnitude at a particular spatial location. The slope of this regression gives an estimate of the sensitivity at that location, and sensitivity fields are constructed by calculating this sensitivity at each spatial location.

This invokes the Monte Carlo concept. Each of the randomly perturbed forecasts will have a lot of low-level random noise obscuring the impact of any particular perturbation. Essentially, the effects of all the perturbations are present simultaneously. By regressing the response function against the initial perturbations at one point, the effects of all the other perturbations (which are random) will be reduced by averaging. The larger the ensemble size, the less noisy the calculated sensitivity fields become.

This work uses binary perturbations. The perturbation field at each horizontal location is either +1 K or -1 K, depending on a random number coin flip. The random number generator is specifically the RANDOM_NUMBER function native to FORTRAN90. A sample regression is shown in Fig. 8 which plots the value of the response function of total accumulated rainfall along the dryline, J, versus the random perturbations at a location which we knew to have a strong impact.

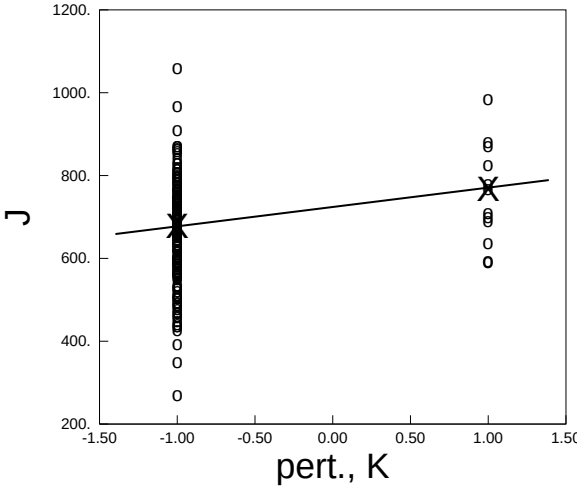

Figure 8. Scatter plot of value of the response function J (with units of total mm of rain) versus size of the boundary layer potential temperature perturbation in the initial condition at each of the 400 members of the ensemble, at a location near the maximum of Fig. 10. The large X's mark the average of J for the positive and negative perturbations. The line drawn through the X's is the best fit line, the slope of which is the sensitivity of J to potential temperature at that one point.

We note two things about Fig. 8. First, there is a great deal of noise. The slope of the regression (indicated by the line drawn in Fig. 8) is small relative to the variance in the values of the response function. Second, there a great many more negative perturbations than positive perturbations. This was unintentional and is caused by imperfections in the random number generator. Better random number generators are available, though good results have been

obtained so far using the generator supplied with FORTRAN90.

A calculation to obtain the slope as done for Fig. 8 is done at every horizontal location and a two-dimensional sensitivity field is thus constructed. Figures 9 and 10 show the results for the same response functions as Figs. 6 and 7, respectively, obtained using the VLE technique.

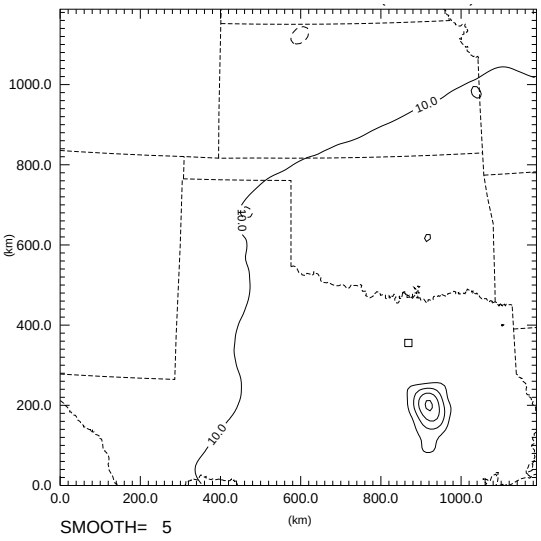

Figure 9. Nondimensional sensitivity field for the dependence on the area average potential temperature in the box indicated to a +1 K perturbation in the initial condition. The contour increment is .0005 and the maximum is .0021.

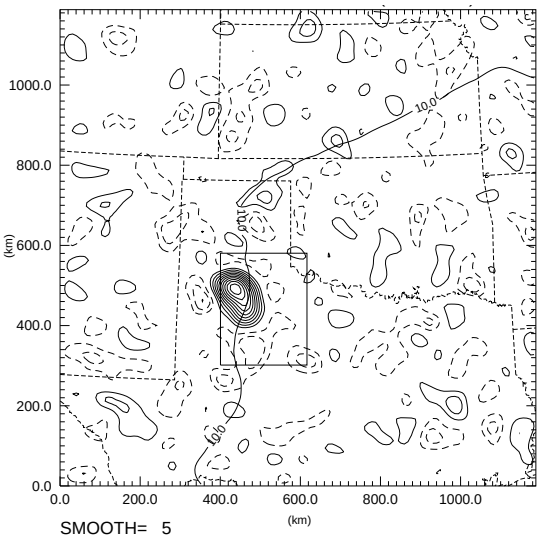

Figure 10. Nondimensional sensitivity field for the dependence of total area accumulated precipitation in the indicate box on initial +1 K boundary layer temperature perturbations. The contour increment is 0.5 and the maximum is 5.4.

Figure 9 is directly comparable with Fig. 6, with both being plotted with the same contour increment and with the same 9-point smoother having been applied 5 times. While the shape of the central contour is a little different, the magnitude of the sensitivity compares quite will with .0020 versus .0021. There is also a little noise apparent in Fig. 9, though this has been much reduced by smoothing. Figure 10 is directly comparable with Fig. 7, again with both having the same contour increment and smoothing level. Again, while the shapes of the central maxima are slightly different, the magnitudes of the maxima agree quite well with 5.2 verus 5.4. Figure 10 shows some significant noise, even after the 9-point smoother was applied 5 times.

To reduce the noise in the random perturbation technique, a larger ensemble size will need to be attempted. Theoretically increasing the size of the ensemble by a factor of N will reduce the size of the random error by the square root of N. Consequently, reduction of random error is expensive. Figure 11 plots the standard deviation of the random noise from an unsmoothed sensitivity field similar to Fig. 9 versus ensemble size (up to 180 members). The reduction in error with increasing ensemble size agrees well with the theoretical values plotted as open circles in Fig. 11.

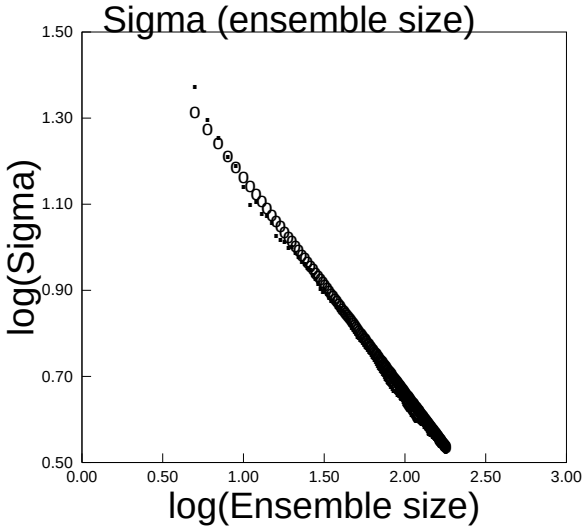

Figure 11. Log-log plot of the RMS noise of a sensitivity field calculated using random perturbations versus size of the ensemble. Dots are the calculated noise values and open circles are the theoretical values.

One advantage of the random perturbation method is that it can be executed with very little disk storage. Since random number generators are deterministic, the initial perturbations used for regression do not need to be stored as they can be recalculated any time they are needed (in practice, however, we have found that it is quicker to read the random perturbations in from a file). Also, if one knows what response function is needed, then only the value of that response function need be stored for each member of the ensemble. To generate Fig. 10, we only needed to store the 400 values of the response function. The same 400 values are regressed against the (recalculated) random perturbations at each spatial location. For research, however, we store complete output fields from each model so that any response function can be calculated later. It is useful to reflect that the information content of Fig. 10 is limited by the 400 real numbers used to generate it. The random noise reflects this limitation in information content.

5. THE ADJOINT TECHNIQUE

Comparable results using the adjoint of the ARPS (Xiao et al., 2005) will be obtained as this work is developed further.

6. SUMMARY

This study investigates the feasibility of calculating sensitivity fields using random perturbations. Nominally the same fields can be calculated by an Adjoint and by the VLE technique. The results from random perturbations agreed remarkably well with those from VLEs with the VLE requiring 2026 model runs and the random method requiring 400. Comparisons with the results of the adjoint are yet to be completed. The various techniques have different advantages and disadvantages.

The main advantage of the adjoint is that it can require the least amount of computation. Only one forward model integration and one integration of the adjoint are required (both requiring about the same amount of time). Also, the results of the adjoint are exact for assumed infinitesimal perturbations. The main disadvantages of the adjoint method are the difficulty in implementing a derivative of the model code, and the large amount of computer memory required (the adjoint integration requires the storage of the complete model fields at every time step of the forward model).

The principal advantages of the VLE technique are that it is easy to implement and that it provides the actual sensitivity to finite perturbations, including the effects of potential nonlinearities. The principal disadvantage of the VLE method is the large number of forward model runs required. If three-dimensional sensitivity fields are desired, then the number of forward model runs is currently prohibitive. The VLE method also requires the storage of a large number of forecast model fields.

The principal advantage of the random perturbation technique is that it does not require as many forward model integrations as the VLE method, if the level of noise is acceptable. The random perturbation method is exponentially more efficient than the VLE method because the number of forward model runs required by the VLE method is equal to the number of degrees of freedom of the model, while that required by the random perturbation method is fixed (though potentially large). Also, the random perturbation technique can, potentially at least, require very little storage as all that is needed in principle are the values of the scalar response function from each forward model run and knowledge of the random number generator.

7. ACKNOWLEDGEMENTS

This research was supported by NSF Grant ATM0129892. The second author was also supported by NSF grant ATM9909007 and a DOT-FAA grant. The computations were performed on the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputing Center.

8. REFERENCES

Li, Z. and I. M. Navon, 2001: Optimality of 4D-var and its relationship with Kalman filter and Kalman smoother. *Quart. J. Roy. Meteor. Soc.*, 127, 661-84.

Martin, W. J. and M. Xue, 2005: Sensitivity analysis of convection of the 24 May 2002 IHOP case using very large ensembles. *Mon. Wea. Rev.* Accepted.

Xiao, Y. M., M. Xue, W. J. Martin, and J. D. Gao, 2005: Development of an adjoint of a complex atmospheric model, the ARPS, using TAF. In *Automatic Differentiation: Applications, Theory and Tools*, Martin Bruecker, George Corliss, Paul Hovland, Uwe Naumann, and Boyanna Norris, eds. Lecture Notes in Computational Science and Engineering.

Xue, M., D. Wang, J. Gao, K. Brewster and K. Droegemeir, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteorol. Atmos. Phys.*, **82**: 139-170.