23.3

CONSTRUCTION OF AN ENSEMBLE OF FORECASTS USING ADJOINT-DERIVED SENSITIVITIES

Michael C. Morgan and Daryl T. Kleist[2]*

University of Wisconsin – Madison
Madison, Wisconsin

[2]Science Applications International Corporation
National Centers for Environmental Prediction
Camp Springs, MD

1. INTRODUCTION

Uncertainty in a particular numerical weather prediction (NWP) forecast can arise from imperfect specification of the initial conditions as well as from errors associated with finite differencing (or spectral truncation) or physical parameterizations. Despite the observed success of day-to-day short range forecasts, it is becoming increasingly recognized that no NWP forecast can be considered complete without a concomitant forecast of the flow-dependent predictability (Palmer 2000). In principle, this flow-dependent predictability could be assessed by predicting the evolution of the probability density function (PDF) of the atmospheric state (or more accurately the PDF of the model's representation of that state). In practice however, uncertainties in the specification of the initial PDF and the large dimensionality of the phase space of the atmosphere (NWP model) make this PDF prediction difficult. In order to reduce the dimensionality of the phase space that needs be explored in predictability assessments, carefully chosen samples of estimates of the atmospheric state are chosen as initial conditions for NWP models. The NWP model is then integrated forward in time from each initial estimate of the atmospheric state to create an ensemble of forecasts. The statistics of the ensemble output are subsequently used to evaluate the predictability of the flow.

Ensemble members are formulated to account for uncertainties in the initial conditions or uncertainties in the model physics (Palmer 2000). Typically, this involves constructing a suite of initial conditions and integrating a model forward several times, to generate an ensemble of forecasts. Ensemble member spread can also be achieved through the use of different physics packages, such as a suite of cumulus parameterizations schemes, or by adding stochastic perturbations to the model physics at intermediate times. Lastly, ensemble spread can also be achieved through the addition of random noise to a forecast, either to the initial conditions or model forecast state.

Most work in ensemble forecasting has been focused on the generation of initial condition perturbations through a variety of methods. The generation of initial perturbations for an ensemble is governed by two principles: 1) the initial perturbations should be constructed using some knowledge of the statistics of the analysis error and 2) the initial ensemble perturbations should share the structure of those perturbations which amplify rapidly over the forecast interval of interest. Two methods are used operationally to generate initial perturbations for ensemble forecasting applications. One of these methods, currently in use at the European Centre for Medium Range Forecasts (ECMWF), involves the use of singular vectors (SVs, also known as optimal perturbations), which are those perturbations which amplify linearly most rapidly for a given norm, given basic state, over a prescribed time interval (Molteni et al. 1996). The second method, in use at the National Centers for Environmental Prediction (NCEP), utilizes a bred mode technique, which uses previous forecasts to ensure that the perturbations will amplify over the next forecast interval (Toth and Kalnay 1997).

SVs are particularly useful for ensemble forecasting, as they provide large spread (though not necessarily sufficiently large spread) over a prescribed forecast interval. Additionally, the choice of the analysis error covariance metric (or some appropriate surrogate for the analysis error covariance metric) as a measure of initial amplitude ensures that the SVs are constructed using knowledge of the characteristics of analysis uncertainty. Although SVs have shown utility in their use for the generation of initial condition ensemble members, there are serious limitations to their efficacy including the computation cost of calculating SVs, the validity of the linear assumption, and the number of members needed to construct a reasonably sized ensemble. Calculating SVs is computationally expensive, as multiple integrations of both the linearized version of an NWP model and its adjoint are required in the iterative schemes used to solve the eigenvalue problem that defines the SVs. For any ensemble prediction system, there is no *a priori* means of identifying the number of ensemble members necessary needed in an ensemble for any particular forecast.

Adjoint derived forecast sensitivities can be used to estimate the change in a particular response function given any arbitrary, but small, perturbation (e.g., Errico 1997). In particular, the sensitivity gradient with respect to the model initial conditions can be used to estimate

*Corresponding author address: Daryl T. Kleist, W/NP2, NOAA, WWB#207, 5200 Auth Rd., Camp Springs, Maryland, 20746-4304
e-mail: daryl.kleist@noaa.gov

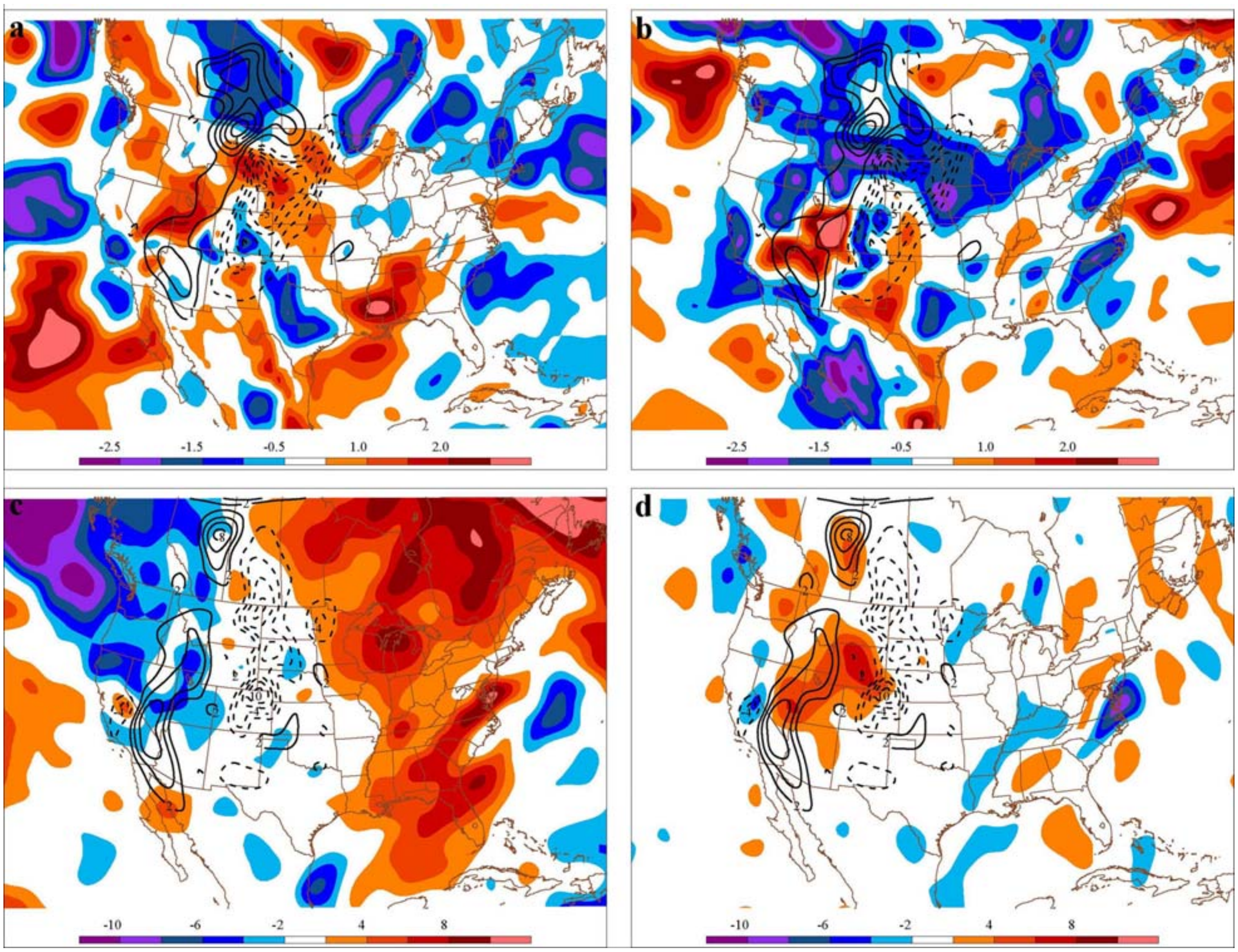

Figure 1. Temperature difference at $\sigma = 0.85$ (color filled, interval 0.5 K) between (a) Eta and AVN model analyses and (b) UKMET and NOGAPS model analyses, as well as meridional wind difference at $\sigma = 0.75$ (color filled, interval 2 m s^{-1}) between (c) Eta and AVN model analyses and (d) UKMET and NOGAPS model analyses valid at 0000 UTC 20 January 2002. Also plotted is the sensitivity of the forecasted average temperature with respect to initial distribution of temperature (a) and (b) at $\sigma = 0.85$ (contours, interval 1 x 10^{-3} K K^{-1}), as well as with respect to meridional wind (c) and (d) at $\sigma = 0.75$ (contours, interval 2 x 10^{-4} K s m^{-1}) at 0000 UTC 20 January 2002; where negative values are dashed and the zero contour has been omitted.

the change in response function given changes to the model initial state. Given a measure of analysis error and a sensitivity gradient for a particular response function derived from a single integration of the adjoint model, one can estimate the likely ranges of values of a response function without integrating the NWP model more than once. Although this method does not suffer the burden of the computational cost of calculating SVs, the linearity assumption is still present, as well as concerns about the ensemble size.

In this presentation, adjoint-derived forecast sensitivity gradients are used in conjunction with differences between operational analyses to construct an ensemble of forecasts for a particular response function. The procedure for creating an ensemble of forecasts using this methodology is outlined in section 2. In section 3, examples from the ensemble procedure are presented and discussed. A summary and outline for future work can be found in section 4.

2. MOTIVATION AND METHODOLOGY

a) Motivation

At any given analysis time, a comparison of analyses within and between operational centers reveals that there may be considerable discrepancies between the initial conditions used for the various operational models. These differences between analyses can be attributed to differing data assimilation algorithms, as well as the usage of different types and subsets of data. To directly compare various operational analyses, all analyses are first interpolated onto a common grid. Assuming the interpolation does not introduce new error, it is clear that operational analyses valid at the same time can exhibit considerable differences (Fig. 1). A comparison of temperature at $\sigma = 0.85$ between NCEP's Eta and Aviation model analyses, as well as between the Navy NOGAPS and United Kingdom Meteorological Office's (UKMET) model analyses reveals that lower tropospheric temperature differences can be significantly large, and in fact can exceed 2 K in certain locations (Figs. 1a and b). Likewise, a comparison of the analyzed meridional wind at σ = 0.75 shows the differences between the operational analyses exceed 10 m s^{-1} in certain locations. The comparison also reveals that the size of the analysis differences depends on which models are being compared, as well as which variable and level are chosen for the comparison (Fig. 1).

Although the analysis differences are not necessarily representative of the actual analysis error, they do provide a means for representing analysis uncertainty. Because of this, the analysis differences can be used as initial time perturbations to create an ensemble of forecasts. Each analysis could be used to initialize an NWP model, and an ensemble of forecasts could be created, similar to performing a model impact study. However, the effect of the analysis perturbations on a particular response function can be estimated by using an adjoint-derived sensitivity gradient (see below), without running the nonlinear model several times.

Any response function can be used in this calculation, but for this work, R_1 (defined as the 36-h forecast temperature averaged over the Midwest on the $\sigma = 0.85$ surface) is chosen for its simple interpretation and potential operational interest for the upper Midwest. If the analysis differences have any projection on the initial time sensitivity gradients for R_1, estimates of values for R_1 can be calculated. The estimate of the change in the response function is simply an inner product, and as such, will be largest if both the analysis differences and sensitivity gradients are large, or small if either the sensitivity gradients or analysis differences are small. The actual calculation for the estimated change in the response function is performed by summing over the entire domain point by point the product of the analysis difference with respect to each variable and the initial condition sensitivity with respect to that same variable. There are certain locations where the analysis differences have a large projection onto the sensitivity gradient, such as in temperature at $\sigma = 0.85$ over southern Canada and the north central United States (Figs. 1a and b) for the analysis valid at 0000 UTC 20 January 2002. However, there are also regions in which the analysis differences are very large, but the sensitivity gradients are extremely small or near zero, such as over eastern North America and off of the coast of British Columbia (Fig. 1c).

b) Methodology

The adjoint model used for this study is a component of the MM5 Adjoint Modeling System (Zou et al. 1997). This modeling system, based upon version one of the Pennsylvania State University/National Center for Atmospheric Research fifth generation mesoscale model (MM5), includes the nonlinear MM5 model, its TLM, and corresponding adjoint. The MM5 model is a nonhydrostatic, limited area, primitive equation model which uses as its vertical coordinate a terrain following sigma coordinate. For all of the sensitivity calculations performed, the nonlinear version of MM5 is used to create a basic state about which the TLM and adjoint models are linearized. For the TLM and adjoint integrations, the basic state is updated every time step.

The domain for the nonlinear, TLM, and adjoint integrations is a 90 km, 70 x 48 horizontal grid, with 10 evenly spaced sigma levels in the vertical (top pressure level in the model is 100 hPa). The nonlinear model is initialized from the National Centers for Environmental Prediction (NCEP) Eta model analysis (on the AWIPS 104 grid) interpolated to the MM5 grid, and lateral boundaries are updated using the NCEP Eta model forecast. The nonlinear integrations use the following physical parameterizations: the Grell convective scheme, a bulk aerodynamic formulation of the planetary boundary layer, horizontal and vertical diffusion, dry convective adjustment, and explicit treatment of cloud water, rain, snow and ice. The TLM and adjoint integrations use the same parameterizations (or their adjoints), but the effects of moisture are neglected. This means that the TLM and adjoint models are integrated using only dry dynamics about the moist basic state created from the nonlinear model run.

The first step in the ensemble procedure is to first define a control 36-h MM5 forecast trajectory initialized using the NCEP Eta model analysis and forecast for initial and boundary conditions respectively. The adjoint model is then run about this basic state trajectory, using as its input, the gradient of the response function R_1 (defined above), to calculate sensitivity with respect to the initial conditions. Analysis differences $\left(\delta \mathbf{x}_0^i\right)$ are then determined from the differences between the NCEP Eta, NCEP Aviation, NCEP NGM, UKMET, and Navy NOGAPS model analyses interpolated to the MM5 grid. From these five model analyses, we may construct 20 initial perturbations (10 positive, and 10 negative). Estimates for the change in the response function are then calculated using these analysis perturbations:

$$\delta R_1^i = \left\langle \frac{\partial R_1}{\partial \mathbf{x}_0}, \delta \mathbf{x}_0^i \right\rangle .$$

Because the calculation is a linear estimate, only 10 independent (positive) perturbations are necessary, as the change for the negative perturbations is determined by multiplying the result by negative one. From these linear estimates, bounds on the value of the response function may be determined by the largest δR_1^i calculated.

As a check on the validity of our linear estimate, the change in R_1, ΔR_1, is evaluated from differences between the nonlinear model runs, which are initialized from perturbed analyses (by adding and subtracting the analysis perturbations to come up with two new analyses). As an example, one of the sets of perturbed nonlinear model integrations involves re-running the model after adding and subtracting perturbations derived from differences between the NCEP Eta and NCEP Aviation model analyses:

$$\Delta R_1^{\pm} = R_1\left(\mathbf{x}_0^{eta} \pm \delta \mathbf{x}_0^{eta-avn}\right) - R_1\left(\mathbf{x}_0^{eta}\right),$$

where $\mathbf{x}_0^{eta}$ is the analysis used to create the control forecast, and $\mathbf{x}_0^{eta} \pm \delta \mathbf{x}_0^{eta-avn}$ are the two perturbed analysis. If the dynamics are linear, then the nonlinearly calculated change in the response function $\left(\Delta R_1^{\pm(eta-avn)}\right)$ will be the same as the linear estimate $\left(\delta R_1^{\pm(eta-avn)}\right)$ for the change in the response function.

For a sufficiently large ensemble size, and for a realistic estimate of the initial condition uncertainty, periods for which the chosen response function has enhanced or decreased forecast uncertainty can be determined. If the estimated change in the response function is relatively large, the analysis differences or the forecast sensitivity gradients (or both) must be large, consistent with that particular forecast exhibiting relatively high forecast uncertainty. Similarly, if the estimated change in the response function is relatively small, higher confidence can be placed on the control forecast for the response function, as a small linear estimate would be consistent with a particular forecast exhibiting rather low forecast uncertainty.

3. EXAMPLES

As with any application which utilizes an adjoint method, the validity of the linearity assumption must be checked. For the ensemble calculations, this involves comparing the estimated change in the response function (δR) with the actual change in the response function (ΔR) calculated as a difference in the response function evaluated from two nonlinear forecasts. For the forecast period spanning 0000 UTC 18 January 2002 through 1200 UTC 24 January 2002 (hereafter referred to as the winter week), the linearity assumption holds quite well, when comparing a time series of the estimated and actual changes in the 36-h forecasted response function (Fig. 2). For the winter week ensemble estimates, correlation coefficients between δR and the ΔR are about 0.85, and as high as 0.95, with the exception being for the Eta-NGM analysis perturbation (Table 2). This suggests that the linear estimates for the change in the response function can serve as a computationally inexpensive proxy for the actual change.

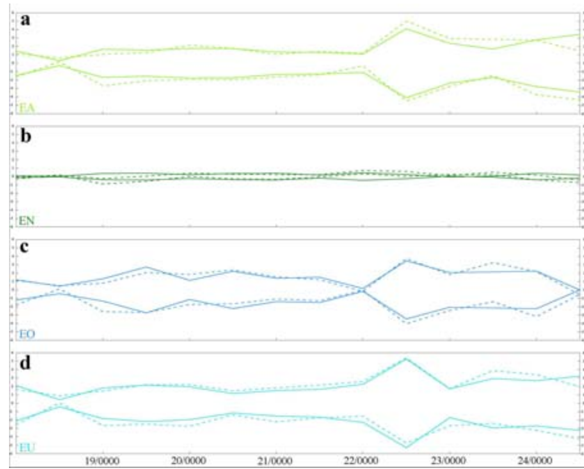

Figure 2. Time series of the actual change in response function (ΔR_1, dashed lines) and estimates for the change in response function (δR_1, solid lines) relative to the control forecast (zero value) valid from 0000 UTC 18 January 2002 through 1200 UTC 24 January 2002 for perturbations derived from analysis differences between (a) EA, (b) EN, (c) EO, and (D) EU. Abbreviations for model analysis differences are identified in Table 1.

The problem with the linearity test for the Eta-NGM perturbation can be attributed to the size of the analysis differences between the two model analyses. Since both of the models are initialized using the same data assimilation algorithm, just mapped to a different grid, their analyses are extremely similar, and as a result, the analysis differences are typically an order of magnitude or more smaller than the differences calculated between the other operational analyses. The problem arises in the nonlinear model because of the inclusion of diabatic processes, which are excluded in the linear estimate for the change in the response function. This is of little consequence to these calculations, as the spread for this particular ensemble member is in general much smaller than for the other members (Fig. 2), and as a result, contains little useful information.

For the week spanning 0000 UTC 20 June 2002 through 1200 UTC 27 June 2002 (hereafter referred to as the summer week), the correlation coefficients between the linear estimate and actual nonlinear differences is much smaller than those for the winter week (Table 2). The largest correlation for the summer week is only 0.69, with the smallest being 0.224 (excluding the 0.083 for the Eta-NGM perturbation). This can also be seen in the time series comparing the linear estimates with the nonlinear differences for the summer week (Fig. 3).

In general, the linearity assumption appears to be more valid for this calculation in the cold season, in comparison with the warm season. This statement is

Figure 3. As in Fig. 2, but valid from 0000 UTC 20 June 2002 through 1200 UTC 27 June 2002 for perturbations derived from analysis differences between (a) EA, (b) EN, and (c) EO. Abbreviations for model analysis differences are identified in Table 1.

consistent with calculations that have been performed operationally for the past two years, and not specific to the two weeks chosen for this discussion.

A time series showing all of the linear estimates for the winter week shows that the spread amongst ensemble members varies from forecast to forecast, with the smallest spread being $\pm 1\mathrm{K}$ and the largest being $\pm 5\mathrm{K}$ for the winter week (Fig. 4a). It is evident that our linear estimate is providing about the same spread as the ensemble of forecasts calculated by integrating the nonlinear model several times (Fig. 4). Fig. 4a shows that there appear to be regimes in which the forecast uncertainty is relatively low (low spread, such as between 18 January 2002 and 21 January 2002), and likewise periods of relatively high forecast uncertainty (high spread, such as between 21 January 2002 and 24 January 2002). The verification for the winter week appears to have little correlation with the ensemble spread, which is somewhat different from the findings for other weeks. Although the verification tends to lie in the vicinity of the ensemble spread, there are often times in which the forecast spread is not large enough to account for the forecast error. However, in general, the largest forecast busts are associated with a large spread in the ensemble members, such as for 1200 UTC 22 January 2002 (Fig. 4a), which is particularly useful then for assessing the uncertainty of a particular forecast.

For the summer week, the spread is in general much smaller than what is typical for cold season calculations (Figs. 4 and 5). This can be attributed to smaller analysis differences or smaller forecast sensitivity for the forecasted temperature over the upper Midwest. The fact that the linearity correlation coefficients are so much smaller for the summer week is consistent with the analysis differences being much smaller, and therefore making the linearity assumption

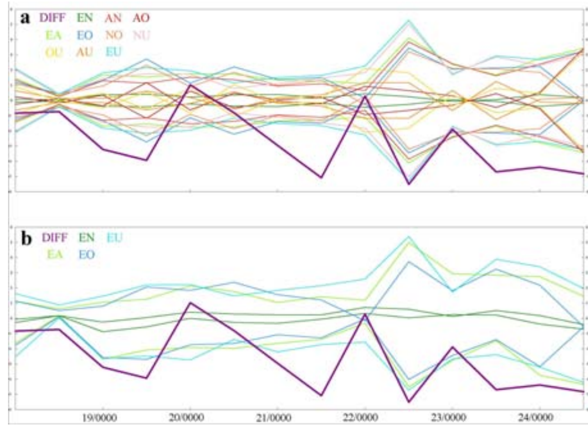

Figure 4. Time series of (a) estimates for the change in response function (δR_1) and (b) actual change in response function (ΔR_1) valid from 0000 UTC 18 January 2002 through 1200 UTC 24 January 2002. The perturbations derived from the model analysis differences are color coded, and the abbreviations are identified in Table 4.1. DIFF is the verified forecast error, and is calculated by taking the difference between the verification of and the control forecast of R_1.

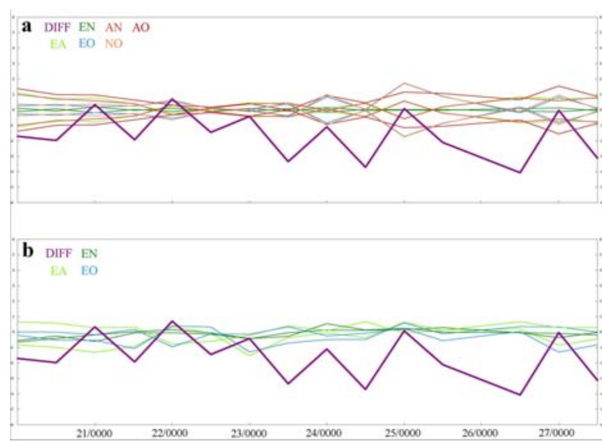

Figure 5. As in Fig. 4, but from 0000 UTC 20 June 2002 through 1200 UTC 27 June 2002.

less valid. However, it is also likely that during the warm season, the forecasted temperature may be more dependent on local effects such as radiational heating, boundary layer mixing, or local circulations, rather than the larger scale dynamics, making the forecast uncertainty much more dependent on the parameterization of such processes. In general, the warm season ensemble spread is much smaller, despite the fact that larger forecast busts occur on a regular basis.

4. FUTURE WORK

The use of adjoint-derived sensitivity gradients in conjunction with analysis differences appears to be an effective, and efficient means of estimating the behavior of a full ensemble, without the need to run a nonlinear model several times. Compared with the cost of SV generated initial ensemble members, the sensitivity based method is computationally inexpensive, as it requires a single iteration of the nonlinear and adjoint models. The choice of response function allows one to tailor this methodology to the needs of the forecaster. The ensemble can be formulated to be specific to a region, or for different forecast parameters, so long as the response function is a differentiable functions of the model forecast variables.

In order for this methodology to be practical, the following questions must be addressed: 1) What is the minimum number of initial perturbations necessary to generate a useful ensemble of forecasts? 2) Are there means of generating more ensemble members using only the currently available operational analyses? 3) What is the maximum forecast length that the linearity assumption can be made to estimate a change in response function? 4) What are the reasons for the notable differences between the warm and cold season ensemble forecasts for our particular response function?

In order to truly assess the practicality of this methodology, a more rigorous evaluation of forecast performance of the ensemble must be done. This is more difficult than evaluating the performance of a single deterministic forecast, as probabilistic information is contained in the ensemble of forecasts. A statistical evaluation of the two years of ensemble calculations needs to be performed, to fully assess the applicability of this methodology. Lastly, this methodology should be applied to response functions other than the one chosen for this work. An ensemble utilizing adjoint-sensitivities could be useful in a variety of forecasting platforms, such as in severe weather forecasting by providing a forecast of ensembles for parameters such as vertical wind shear.

5. REFERENCES

Errico, R., 1997: What is an adjoint model?. *Bull. Amer. Meteor. Soc.*, **78**, 2577-2591.

Molteni, F., R. Buizza, T. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,***119**, 269-298.

Palmer, T. N., 2000: Predicting uncertainty in forecasts of weather and climate, *Rep. Prog. Phys.*, **63**, 71-116.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297-3319.

Abbreviation	Analysis Difference
EA	± (Eta – Aviation)
EN	± (Eta – NGM)
EO	± (Eta – NOGAPS)
EU	± (Eta – UKMET)
AN	± (Aviation – NGM)
AO	± (Aviation – NOGAPS)
AU	± (Aviation – UKMET)
NO	± (NGM – NOGAPS)
NU	± (NGM – UKMET)
OU	± (NOGAPS – UKMET)

Table 1. Abbreviations used for differences between model analyses.

	EO	EU	EN
Winter (+) (-)	0.855 (0.917)	0.864 (0.887)	0.452 (0.112)
Summer (+) (-)	0.224 (0.468)		0.505 (0.083)

Table 2. Correlation coefficients between the linear estimate for the change in response function (δR_1) with the actual change (ΔR_1).

Acknowledgements

This study represents part of the corresponding author's Master's thesis completed at the University of Wisconsin – Madison. This work was supported by the National Science Foundation Grant ATM-0121186.