Estimation of Initial and Forecast Error Variances for the NCEP's operational Short-Range Ensemble Forecast (SREF) system

Hongli Wang

Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO

Visitor's Program Project Report

Developmental Testbed Center (DTC)

Dr. Hongli Wang

E-mail: Hongli.Wang@noaa.gov

Phone (office): 303-497-6032

ABSTRACT

Estimation of initial condition-related numerical weather prediction uncertainty is a prime topic in fields of data assimilation and ensemble forecasting research. An Analysis and Forecast Error Variances Estimation System (AFEVES) had been developed by solving a constraint nonlinear optimization problem. The system uses information from differences between forecast and analysis fields (perceived error, PE) with different lead times with prior knowledge regarding the time evolution of (a) forecast error variance and (b) correlation relation between errors in forecasts and analyses. This report documents the technique details on AFEVES such as cost function, and minimization algorithm. The system was first examined in a framework of Observing System Simulation Experiment (OSSE) and applied to the NCEP GFS history analyses and forecasts. The AFEVES produced almost the same geopotential height error variances estimation reported by Peña and Toth (2014), which indicate the robustness of the system.

1. Background

Estimation of initial condition-related numerical weather prediction uncertainty is a prime topic in fields of data assimilation and ensemble forecasting. A accurate specification of short-term forecast error (background error) statistics is crucial to the quality of the analysis in a data assimilation system, because the background error along with observations error determine to what extent the background fields will be corrected to fit observations (e.g. Courtier et al. 1998; Bannister 2008). However, the estimation of the background error statistics is not straightforward, since true atmospheric state ("truth") is not known. Two methods are mainly used in current data assimilation systems. The so-called NMC (named for the National Meteorological Center, now called the National Centers for Environmental Prediction) method (Parrish and Derber 1992) is one approach that is widely employed to estimate the climatological background error covariances. The background error generated by this method is not forecast error at a lead time (e.g. 6 hours), but a combination of forecast errors of different lengths and their correlations (Bannister 2008; Wang et al. 2014). An alternative method is to use an ensemble of short-term forecasts at a specific time to evaluate the flow-dependent covariances (Houtekamer et al. 1996; Fisher 2003). In this method, the closer the ensemble mean is to "truth", the more accurate the background error estimation. A method that directly deals with the forecast error (difference between forecast and truth) and produces the estimates of forecast error is of interest.

Accurate estimates of error variances are also crucial for generating initial perturbations in ensemble forecast systems. Research community has recognized that the representation in initial condition-related uncertainty is an important aspect of ensemble systems (Tollerud et al. 2013). As a first attempt at estimating spatial variation of rescaling parameters, the ratios of ensemble spread to ensemble mean forecast error at a few vertical levels were investigated for potential application in rescaling initial ensemble perturbations for the operational SREF system (Du et al. 2012). Results suggested that the present method of computing initial perturbations is likely fine as is for the previous SREF implementation from 500 hPa to the top of the atmosphere (Tollerud et al. 2013).

In this report, the method proposed by Peña and Toth (2014) is briefly introduced in section 2. Section 3 describes the nonlinear constraint optimization problem and its new solution scheme. The results from OSSE and a few operational systems will be shown in section 4. The final section gives a summary.

2. Methodology

Peña and Toth (2014, hereafter PT2014) proposed the use of differences between forecast and analysis fields ("perceived forecast errors") to provide the unbiased estimation of analysis and forecast errors. Here we adopt the idea but also use the differences between forecasts valid at the same time but with different lead times to derive the initial and forecast error variance estimations.

2.1 Decomposition of forecast differences

Assuming there are two forecasts F_i and F_j valid at the same time, one with a lead time i, and the other with a *shorter* lead time j, the forecast difference d_{ij} is

$$d_{ij}^2 = (F_i - F_j)^2 = ((F_i - T) - (F_j - T))^2 \equiv (x_i - x_j)^2 \qquad (1)$$

where T is truth, $x_i = F_i - T$ and $x_j = F_j - T$ are forecast errors for lead times i and j respectively.

Following the law of sum of variances to rewrite the right hand side of (1):

$$d_{ij}^2 = x_i^2 + x_j^2 - 2\rho_{ij} x_i x_j \quad (2)$$

where , ρ_{ij} is forecast error correlation between x_i and x_j. Equation (2) builds a relation between variance of forecast difference $(F_i - F_j)^2$ and the unknown true forecast errors x_i and x_j and their correlation ρ_{ij} .

2.2 Parameters estimation

The estimations of unknowns x_j, α, and ρ_{ij} are achieved by the minimization of the following cost function

$$J(x_i, x_i, \rho_{ij}) = \max(\left|d_{ij}^2 - \hat{d}_{ij}^2\right| \cdot w_{ij}^{-1}) \quad (3)$$

where w_{ij} are specified weights and $\hat{d}_{ij}$ is measured variance of forecast difference calculated from real model forecasts. It is seen that the method is independent of any assumption or tuning parameter used in data assimilation schemes. To facilitate the estimation of the three unknowns, one can introduce more equations like (2) valid for various other lead times. By doing so, however, additional unknown variables (x_i and ρ_{ij}) are also introduced. In PT2014, a nonlinear relation between $x_i = f(a, x_j)$ is introduced to reduce the optimization variables to the unknowns x_j, α, and ρ_{ij}, where α is the error growth rate. It is noted that x_j can be regarded as initial condition error (analysis error if F_j is analysis) with lead time j=0.

3. Analysis and Forecast Error Variance Estimation System

The minimization of the cost function defined by Eq. 3 is a nonlinear constraint optimization problem. In PT2014, the Nelder-Mead Simplex method (Lagarias et al., 1998) in the Matlab software was used to minimize eq. (3). However, this method is sensitive to first guess parameters; it is thus important to ensure that the starting point of the minimization of the cost function is located close to the absolute minimum.

Moreover, the gradient of the cost function defined by Eq. (3) does not exist, which makes it not straightforward to use minimization algorithms that require the gradient of a cost function.

To overcome the limitation in the Nelder-Mead Simplex method, and make the codes easily portable on different computational platforms, the minimization problem (Eq. 3) is redefined, and a limited-memory BFGS (L-BFGS or LM-BFGS) algorithm was introduced to solve the constraint nonlinear optimization problem. The new optimization system is referred to as the Analysis and Forecast Error Variances Estimation System (AFEVES).

The newly proposed cost function with various constraints can be written as,

$$\begin{cases} J(x_0,\alpha,\rho_1) = \sum_i [(d_i^2 - \hat{d}_i^2)w_i^{-1}]^2 \\ d_i^2 = x_i^2 + x_0^2 - 2\rho_i x_i x_0 \\ x_i^2 = x_0^2 e^{\alpha t_i} \\ \rho_i = \rho_1^i \\ 1 \ge \rho_i \ge 0 \\ \alpha \ge 0 \end{cases} \tag{4}$$

x_0 and x_i is analysis error and forecast error with leading time *i,* respectively. ρ_1 is an optimized variable denotes the correlation between forecast error with lead time j=1 and analysis error. ρ_i is correlation between x_0 and x_i. d_i^2 is measured variance of difference between forecast with lead time *i* and analysis. This is called perceived error in PT2014. A limited-memory BFGS (L-BFGS or LM-BFGS) algorithm (Byrd et al. 1995) is used to obtain the minimization of the constraint cost function. LM-BFGS is a kind of quasi-Newton method that approximates the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm using a limited amount of computer memory. The minimization code and its interfaces for the above defined cost function and its gradient were written in the Fortran language and thus it is portable to different computational platforms. The system includes an input and output (IO) component and a key component of a LM-BFGS minimization algorithm with specified cost and its gradient.

NCL scripts were also written to provide input files for the AFEVES.

4. Results

4.1 OSSE results

The AFEVES is first tested in the framework of a Observing System Simulation Experiment (OSSE). The OSSE configuration is as follows:

- Period: Aug 2- 31 2005
- Nature run: ECMWF T511
- Data assimilation and forecasting system
    - NCEP's GFS system
    - T382for both GSI and GFS model
    - GSI with static background error covariance
    - Conventional and remote observations
    - 3-day forecasts initiated at 00Z and 12Z everyday

The perceived error variances were calculated using the 3-day forecasts during 2-31Aug 2005 via the GSI analysis. The true error variances were also obtained by comparing the analysis and forecasts with the ECMWF nature run. The optimization period is 2.5 day with 12-hour interval. In the report only the error estimation for geopotential height (GPH) is investigated.

In PT2014, it is assumed that error follows an exponential growth function ($x_i^2 = x_0^2 e^{\alpha t_i}$) for short-range forecasts, which is also used a strong constraint in the cost function. This means the method can estimate the growth components in the analysis error variances. However, if the analysis error variances have a decaying component for short-range forecasts, one needs to be cautious when interpreting the results.

It is found that there are decaying components in the GPH analysis error variances (Fig. 1). This is mainly because the analyses were produced by the GSI 3DVAR with climatological background error covariance that cannot describe multi-variate covariance. It indicates the method may underestimate analysis error variances since it will always try to extract the growth component using modeled perceived error information.

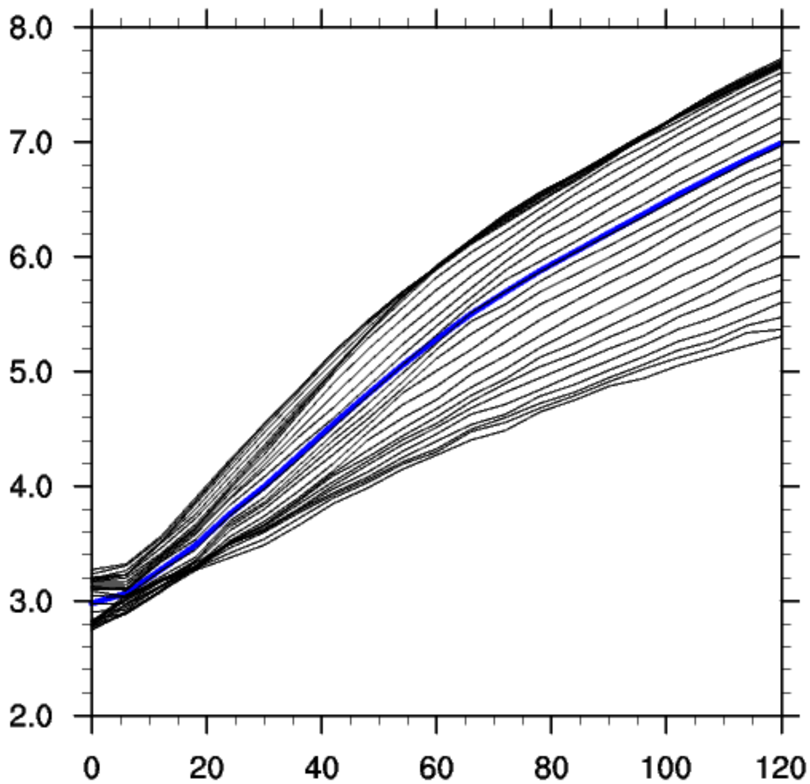

Fig. 1. The GPH error variance evolution with forecast time (unit: hour). The $\ln(x(t)^2/x(0)^2)$ is plotted. The black curves present the snapshot of error variance evolution between 30°N-60°N. The blue curve describes the mean error variance evolution between 30°N-60°N.

Before investigating the GPH error variance estimated by the AFEVES, the fitting error is examined first. It is expected that the absolute value of the difference between the observed perceived error variance and modeled perceived error variance (ABSE) is smaller than the standard error of the mean (SEM) if the AFEVES works well. This standard is also used as an index to justify the minimization algorithm in PT2014. Figure 2 shows SEM and the ABSE in geopotential height at 200hPa, 500hPa and 700hPa levels.

It is seen that the AFEVES produces very good fitting to the perceived errors and the ABSEs are smaller than the SEMs. This confirms the good performance of the AFEVES.

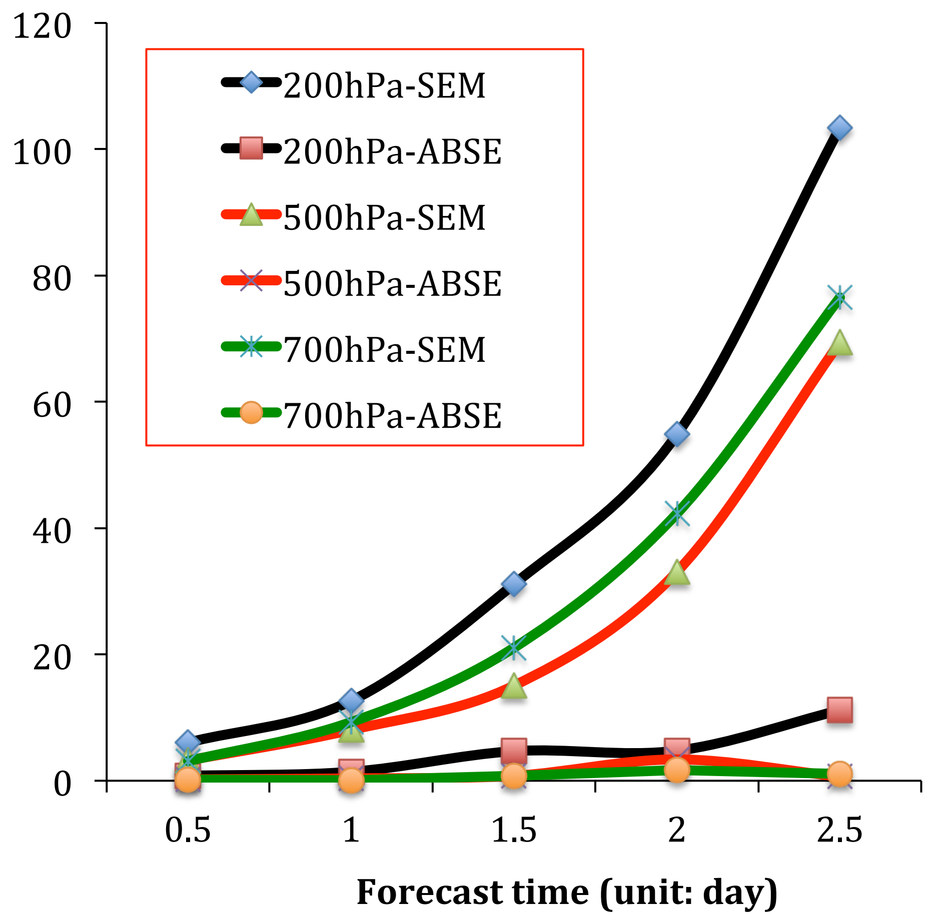

Fig. 2. Absolute value of the difference between the observed perceived error variance and modeled perceived error variance (ABSE), and the standard error of the mean (SEM) in goepotential height at 200hPa, 500hPa and 700hPa levels.

Figure 3 displays the vertical profiles of "true" GPH analysis error variance and the estimated error variance. Except for levels 200 and 300 hPa, the estimated analyses error variance is smaller than the "truth" one. One of the reasons is that there are decaying components in the analysis error variance, whereas the current method only extracts the growing components of it.

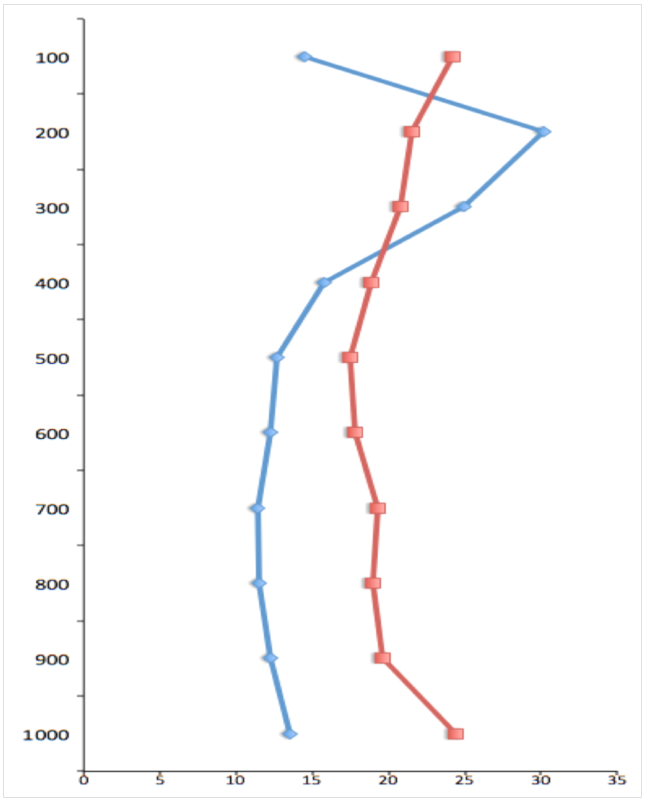

Fig. 3 Vertical profiles of GPH analysis error variance. Estimated analysis error variance is the blue curve and "truth" analysis error variance is the red curve.

4.2 Error Estimation for Operational Forecasting System

In this section, the AFEVES was applied to estimate analysis error variance for a few operational forecasting systems and compared results in PT2014. The 500 hPa geopotential height error variances over the Northern Hemisphere (hereafter NH; 30°N to 90°N) from the four models are analyzed. The dataset and region are same to PT2014 for direct comparison of results. Hence the SREF products were not used. The model versions were in operation during the Fall of 2008 at the National Centers for Environmental Prediction (NCEP), the Canadian Meteorological Center (CMC), the European Center for Medium range Weather Forecast (ECMWF) and the Fleet Numerical Meteorology and Oceanography Center (FNMOC). The forecasts and the analysis verification data are on regular grids of 1×1 degrees of resolution in latitude and longitude. The forecast perceived errors are computed as area-average error variances over the domains available. Forecast data consist of output every 12 h out to 2.5 d initialized at 0000 UTC daily from 1 Sep 2008 through 30 Nov 2008. Forecasts are verified against their own analyses. Table 1 shows analysis error variance and error growth rate estimated by AFEVES for the NCEP, CMC, ECMWF and FNMOC forecast systems. The results reported by PT2014 are shown as well for direct comparison.

It is seen that though different optimization systems are used, the AFEVES produces almost the same analysis error variance estimation as PT2014 and exactly the same error growth rate estimation as PT2014. The results solidify the performance of the newly developed AFEVES system.

Table. 1. Analysis error variance and error growth rate estimated by AFEVES and comparison with PT2014 for the NCEP, CMC, ECMWF and FNMOC forecast systems.

Model	Analysis Error Variance		Error Growth Rate	
	PT2014	AFEVES	PT2014	AFEVES
NCEP	38.0	37.78	0.25	0.25
CMC	29.5	30.14	0.27	0.27
ECMWF	11.5	11.40	0.30	0.30
FNMOC	49.2	52.29	0.26	0.26

Figure 4 displays the estimated true error variances at analysis time and their evolution for the four models. It is seen that the ECMWF model has the smallest analysis and forecast errors followed by the NCEP and CMC models, and the FNMOC model shows the largest error variance. Though different minimization algorithms and cost function specifications, these results are quantitatively consistent with the results reported in PT2014 (See their Fig. 6).

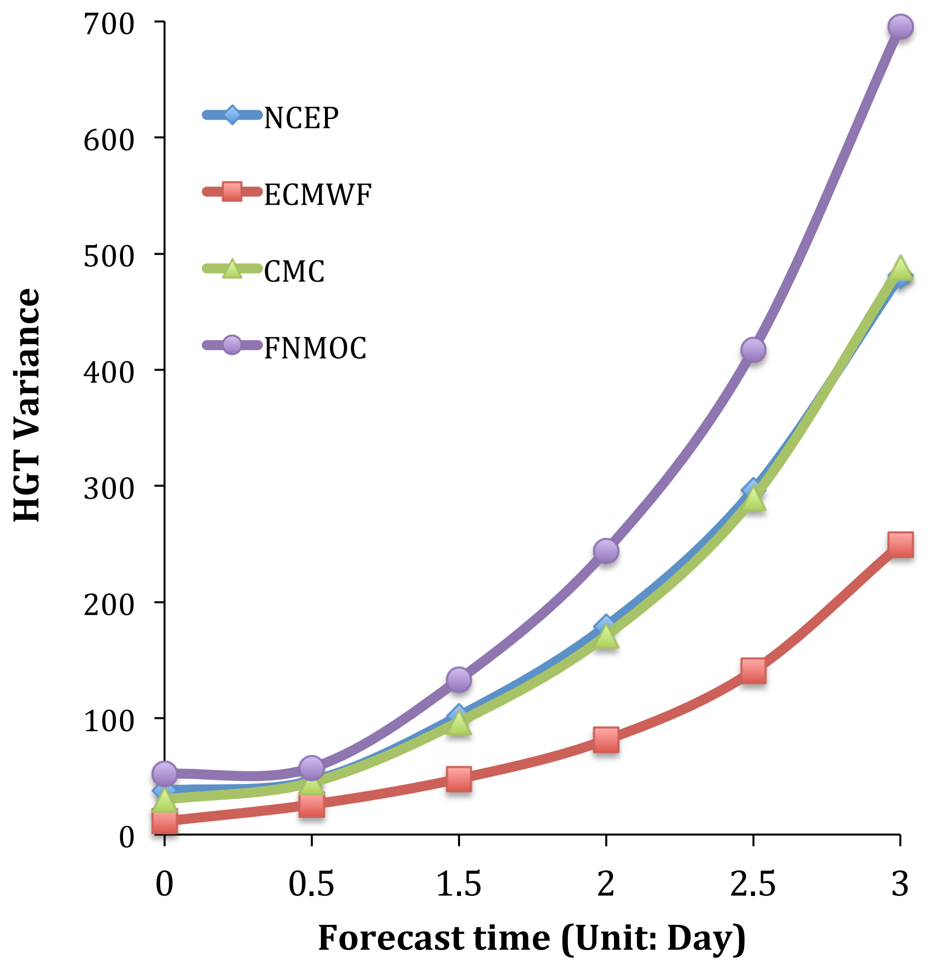

Fig. 4. Estimated true forecast error variances as a function of lead time for four global models: NCEP, CMC, ECMWF and FNMOC for models operational in 2008.

4. Summary

The Analysis and Forecast Error Variances Estimation System (AFEVES) has been developed to estimate the growing components of analysis error variance and following short-range forecast error variance. This work advances PT2014 in the following aspects: 1) a new cost function with L2 norm whose gradient exists were used; 2) the LM-BFGS algorithm was incorporated to solve the constraint nonlinear optimization problem; and 3) the optimization code was written in the Fortran language. The AFEVES has merits of less sensitivity to initial guesses of the best solution and it is easy to port to other computational platforms.

The system was first examined in the framework of Observing System Simulation Experiment (OSSE) and then applied to the NCEP operational GFS history analyses and forecasts. Results from the OSSE show that there are decaying components in analysis error variance. However, the system is designed to extract the fast growing components of the errors, which indicates that the method may underestimate the analysis error varianc. This was evidenced that in most pressure levels (except for 200 and 300 hPa) the error variance of GPH was underestimated. A reason leading to the decaying components in the GFS system is because the analyses were produced by the GSI 3DVAR with climatological background error covariance that cannot accurately describe multi-variate covariance.

Acknowledgement

I thank Dr. Toth and Dr. Pena for extensive discussions and collaboration on this research. Discussions with Dr. Yuanfu Xie, Dr. Yuejian Zhu, and Deqiang Liu are gratefully acknowledged. I also thank Dr. Tara Jensen and Dr. Louisa Nance for insightful discussions and kind help keeping the project on track.

References

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Quart. J. Roy. Meteor. Soc*., 134, 1951–1970.

Byrd, Richard H.; Lu, Peihuang; Nocedal, Jorge; Zhu, Ciyou (1995). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific and Statistical Computing* **16** (5): 1190–1208. doi:10.1137/0916069.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part 1: formulation. *Quart. J. Roy. Meteor. Soc.,* 124, 1783-1807.

Du J. et al. 2012: New 16km NCEP Short-Range Ensemble Forecast (SREF) system: what we have and what we need? Ensemble workshop. 10 - 12 September 2012. NCAR: Boulder, CO.

Fisher M. 2003: Background error covariance modelling. ECMWF Seminar on recent developments in data assimilation for atmosphere and ocean, 8 – 12 September 2003. ECMWF: Reading UK.

Houtekamer PL, L. Lefaivre, J. Derome, H. Ritchie, and H.L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev*.. 124: 1225–1242.

Lagarias, J. L., Reeds, J. A., Wrights, M. H. and Wright, P. E. 1998. Convergence properties of the Nelder-Mead Simplex Method in Low dimensions. *SIAM J. Optim.* **9**, 112–147.

Parrish, D., and J. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev*., 120, 1747–1763.

Peña M. and Z. Toth, 2014: Estimation of Analysis and Forecast Error Variances. Tellus A 2014, **66**, 21767, http://dx.doi.org/10.3402/tellusa.v66.21767

Tollerud, Edward I., and Coauthors, 2013: The DTC Ensembles Task: A New Testing and Evaluation Facility for Mesoscale Ensembles. *Bull. Amer. Meteor. Soc.*, 94, 321–327.

Wang H**.,** X.-Y. Huang, J. Sun, D. Xu, S. Fan, J. Zhong and M. Zhang, 2014: Background Error Modeling for WRF-Var using the NMC method. *J. Appl. Meteor. Climatol.* **53**, 2287–2309.