The Tangent Linear Normal Mode Constraint in GSI: Theory and Early Tests

David F. Parrish[2], Daryl T. Kleist[1], and Catherine Thomas[1,2,3]

[1] University of Maryland-College Park
[2] NOAA/NWS/NCEP/Environmental Modeling Center
[3] IMSG

Some history

The gridpoint statistical interpolation scheme (GSI) was developed as a replacement for the operational spectral statistical interpolation code (SSI) starting in 2001-2002, with Wan-Shu Wu the principle developer. It was based on recursive filters developed by Jim Purser and applied by Dave Parrish starting in 1995 to regional analysis for the NAM eta model.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev*., **120**, 1747-1763.

Wu,W.-S., D. F. Parrish, and R. J. Purser, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.* , **130**, 2905–2916.

Some history

Due to significant code incompatibility between the eta analysis (mainly radiance data processing) and the evolving global GSI analysis, a decision was made in 2003 to make the new GSI code work for both global and regional models. The unified GSI became operational for the regional model on June 13, 2006, with the non-hydrostatic WRF-NMM replacing the earlier eta model.

Some history

While the GSI was now operational for the regional model, the GSI based global parallel system was still consistently inferior to the operational SSI. Daryl Kleist decided to try transferring weak dynamic constraints (turned off in operations) from the SSI code to see if the skill of GSI based forecasts could be improved. But this resulted in a highly ill-conditioned problem.

Some history

In June of 2006, while on vacation, I wondered if it would be possible to create something based on non-linear normal mode initialization (NLNMI), but applied as part of the forward model from the analysis control variable to the observations. Both SSI and GSI already had a regression based linear balance built in to the background error covariance. This would just be adding a 1st order refinement based somehow on NLNMI.

Baer, F., and J. Tribbia, 1977: On complete filtering of gravity waves through nonlinear initialization. M*on. Wea. Rev.*, **105**, 1536–1539.

Machenhauer,B.,1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialization. *Contrib. Atmos. Phys.*, **50**, 253-271.

Some history

The idea came together very quickly. Instead of full variable NLNMI, the new scheme would be based on the tangent linear equations of a dry adiabatic version of the global spectral model. When I got back from vacation, I met with Daryl and we planned a strategy to quickly have a code in place for experiments. Daryl worked on the tangent linear and adjoint of a dry adiabatic version of the global spectral model, while I developed the normal mode part.

Some history

After a lot of debugging and testing, the GSI with the new TLNMI constraint maintained a consistent improvement over the operational SSI based system. The GSI was implemented in the global data assimilation system (GDAS) on May 1, 2007.

During this period, an independent effort was underway in Canada for essentially the same idea, which we first heard about from an ECMWF workshop.

Fillion, L., and Coauthors, 2007: Case dependent implicit normal mode balance operators. Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation, Reading, United Kingdom, ECMWF, 125–142.

Description of the tangent linear normal mode constraint

The basic 3dvar cost function is

$$J(\mathbf{x})= \frac{1}{2} [\ \mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} + (\mathbf{Hx} - \mathbf{y})^T\mathbf{R}^{-1}(\mathbf{Hx} - \mathbf{y})\ ]$$

where

$\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$, $\mathbf{x}_a$ = analysis, $\mathbf{x}_b$ = background,

$\mathbf{y} = \mathbf{y}_o - \boldsymbol{H}(\mathbf{x}_b)$, $\mathbf{y}_o$ = observations,

$\boldsymbol{H}$ = observation forward operator,

$\mathbf{H}$ = tangent linear of $\boldsymbol{H}$,

$\mathbf{B}$ = background error covariance,

$\mathbf{R}$ = observation error covariance.

Ide, K., P. Courtier, M. Ghil, and A. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181-189.

Description of the tangent linear normal mode constraint

With TLNMC, the cost function is now

$$J(\mathbf{x}_c) = \frac{1}{2} \left[\mathbf{x}_c^T \mathbf{B}_c^{-1} \mathbf{x}_c + (\mathbf{H}\mathbf{x}_c - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}_c - \mathbf{y}) \right]$$

where

$\mathbf{x}_c = \mathbf{C}\mathbf{x}$

$\mathbf{B}_c = \mathbf{C}\mathbf{B}\mathbf{C}^T$

and

C is the TLNMC operator.

Kleist, D.T., D.F. Parrish, J.C. Derber, R. Treadon, R.M. Errico, and R. Yang, 2009: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137**, 1046-1060.

Description of the tangent linear normal mode constraint

The TLNMC operator **C** is an application of the Machenhauer/Baer-Tribbia style NLNMI iteration, but applied to the tangent linear formulation of a dry adiabatic version of the global spectral model (GFS). It has the option of doing more than 1 iteration, but in practice, 1 iteration is sufficient.

Description of the tangent linear normal mode constraint

$$x_c = C\,x = [\,I + V\,S\,D\,S^{-1}\,V^{-1}\,T\,]^p\,x$$

where

T = tangent linear of dry adiabatic spectral model,

V = n x m matrix of m vertical modes (n=# vert levels)

S = spherical harmonic to lat-lon grid transform

D = Temperton spectral implicit NMI operator

(computes corrections to **x** which reduce the amplitude of propagating gravity waves)

p = number of iterations for Machenhauer scheme

(p = 1 is default and probably best setting)

Temperton, C., 1989: Implicit normal mode initialization for spectral models. *Mon. Wea. Rev.* , **117**, 436–451.

Description of the tangent linear normal mode constraint

$$\mathbf{x}_c = \mathbf{C}\,\mathbf{x} = [\,\mathbf{I} + \mathbf{V}\,\mathbf{S}\,\mathbf{D}\,\mathbf{S}^{-1}\,\mathbf{V}^{-1}\,\mathbf{T}\,]^p\,\mathbf{x}$$

The operator **C** corrects the analysis increment at every iteration of the minimization, ensuring that relatively well balanced model variables are being used to compare against observations. This is more desirable than correcting full fields after the analysis, which moves the analysis away from the data.

Description of the tangent linear normal mode constraint

$$\mathbf{x}_c = \mathbf{C}\,\mathbf{x} = [\,\mathbf{I} + \mathbf{V}\,\mathbf{S}\,\mathbf{D}\,\mathbf{S}^{-1}\,\mathbf{V}^{-1}\,\mathbf{T}\,]^p\,\mathbf{x}$$

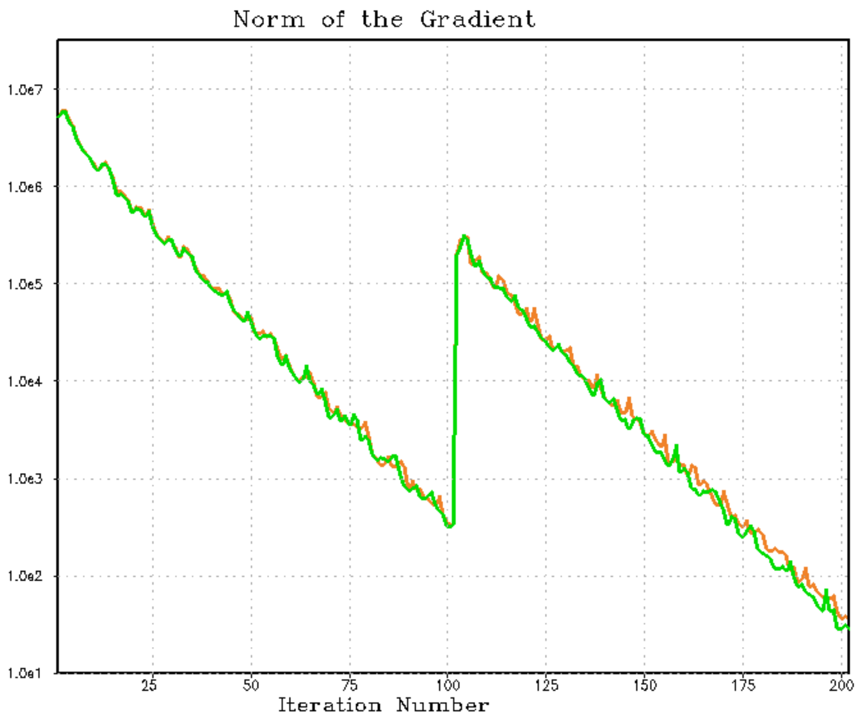
Note that with TLNMC turned on, the background error is now $\mathbf{CBC}^T$, so there are fewer degrees of freedom available to fit observations. This could slow down convergence and/or fit the observations less well. In practice, the convergence rate is about the same, but the fit to observations is not as good. This is offset by improved forecasts.

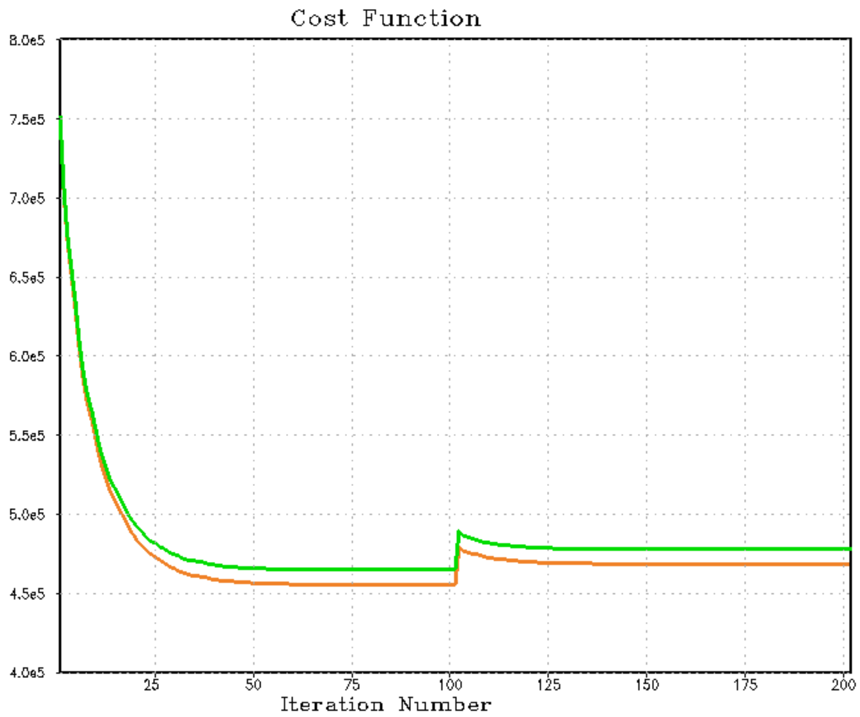
Because $\mathbf{C}^T$ is required, the adjoints of $\mathbf{V}$, $\mathbf{S}$, $\mathbf{D}$, $\mathbf{S}^{-1}$,$\mathbf{V}^{-1}$ and $\mathbf{T}$ are needed.

Some results generated by Daryl Kleist

The following 4 slides illustrate the positive impact of the TLNMC on global data assimilation. These results were an important step towards implementation of GSI in the NCEP Global Data Assimilation System (GDAS)

Little Impact on Minimization

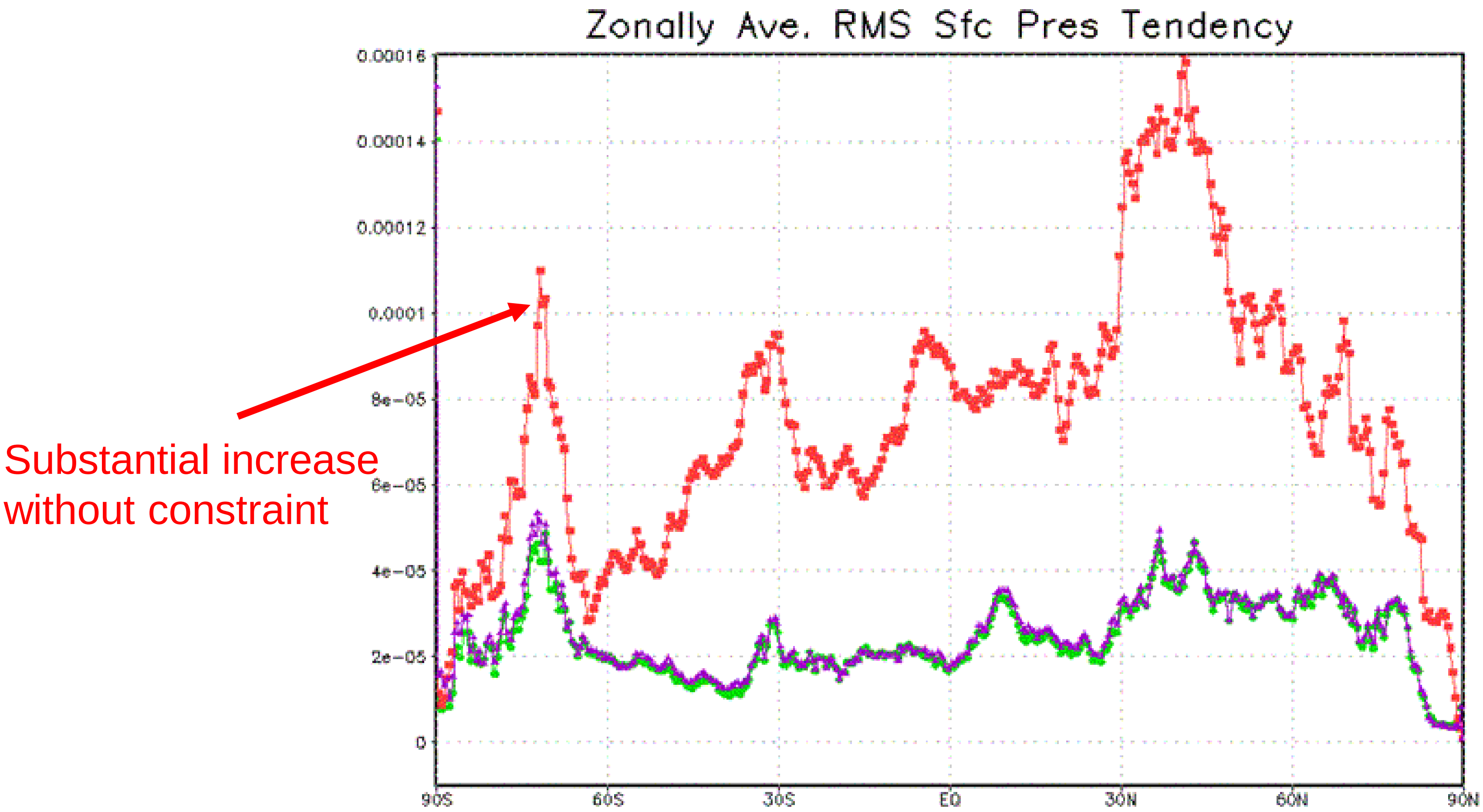




Norm of gradient (left) and total penalty (right) for each iteration for analysis at 12Z 09 October 2007 [Jump at iteration 100 from outer loop update]

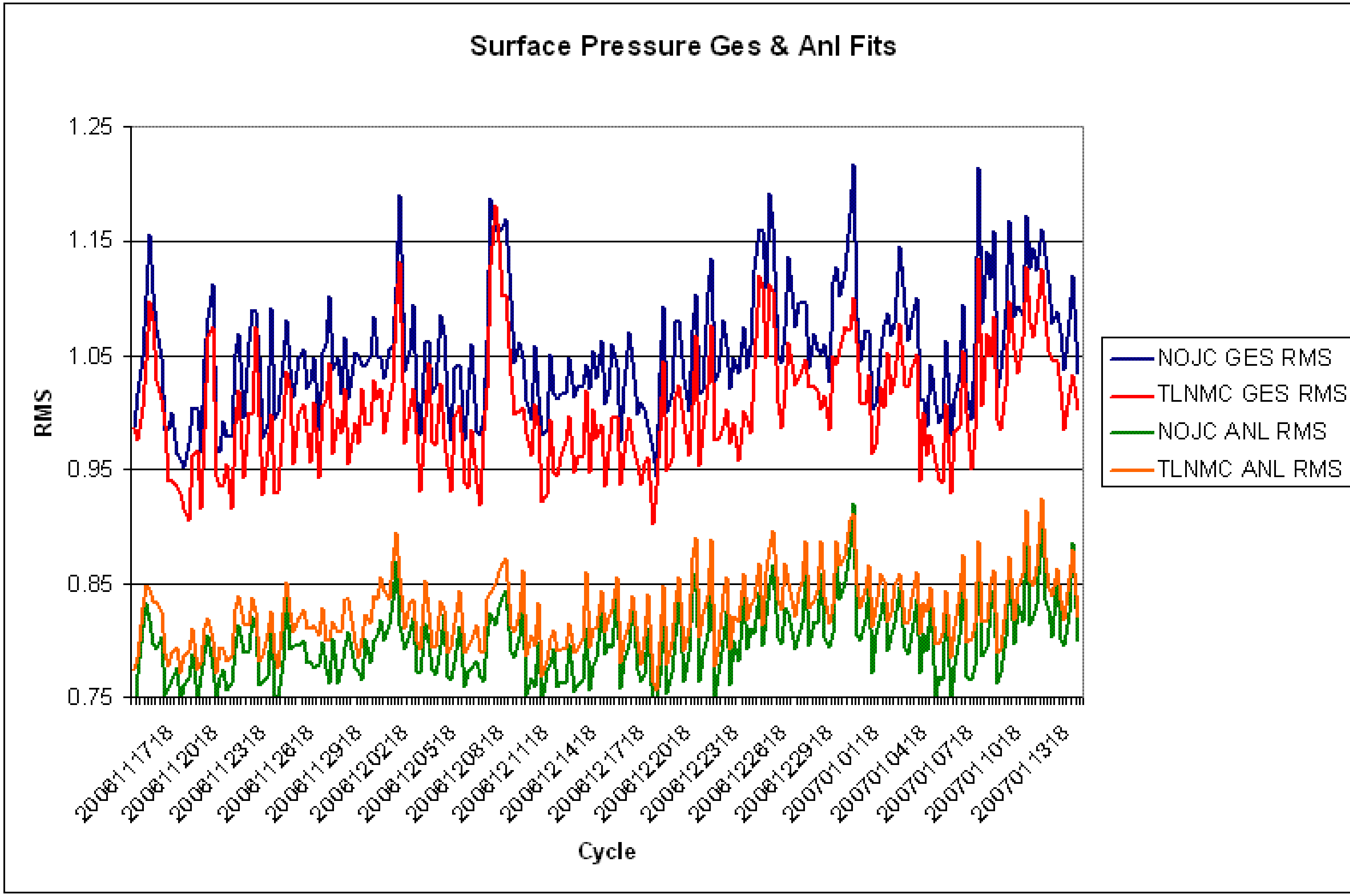
No Constraint (orange) versus TLNMC (green)

Increase in Ps Tendency found in GSI analyses

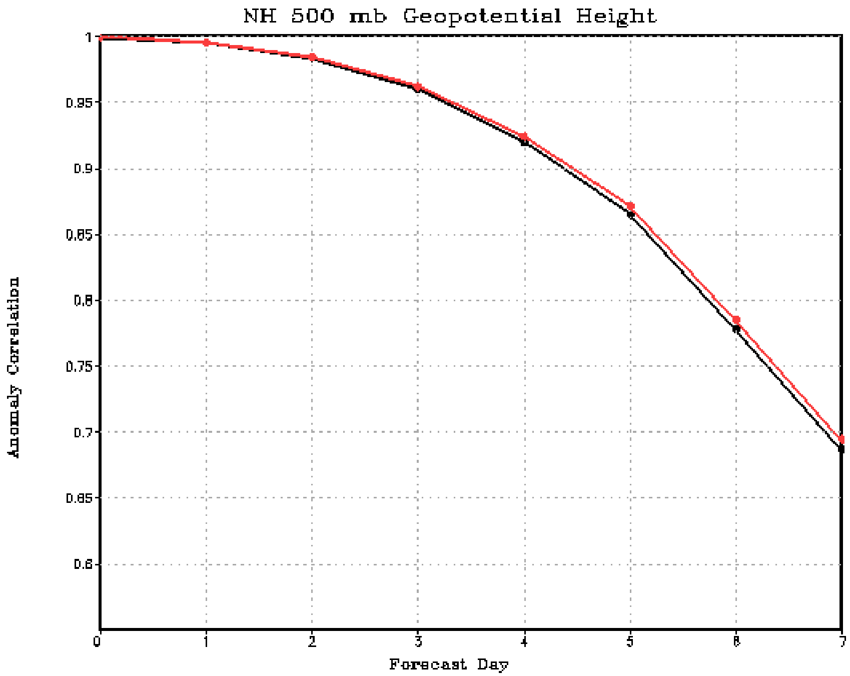


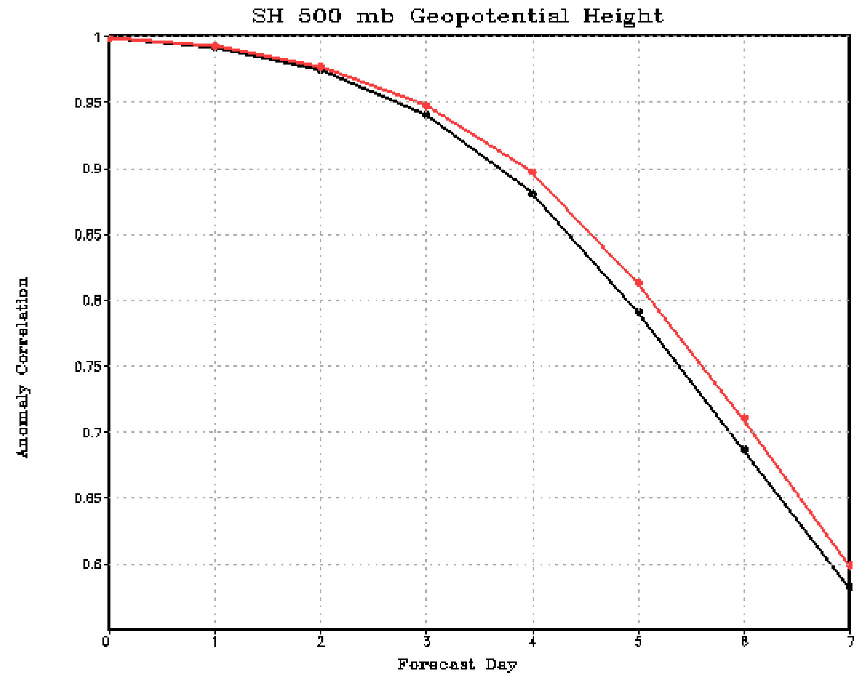
Zonal-average surface pressure tendency for background (green), unconstrained GSI analysis (red), and GSI analysis with TLNMC (purple).

Fits of Surface Pressure Data in Parallel Tests



Impact of TLNMC on 500 hPa AC Scores





No Constraint (control, black) versus TLNMC (red)

500 hPa Geo. Height AC Scores for period 01 Dec. 2006 to 14 Jan. 2007

Application of TLNMC to regional model

A regional version of TLNMC was created after the global version was completed. For the normal mode part, the method of Briere was used. This consisted of using a double sine series expansion in the horizontal over the regional domain. The linear model was based on a double sine series expansion in the horizontal and the assumption of constant coefficients (constant Coriolis parameter and map factor).

Briere, S., 1982: “Nonlinear Normal Mode Initialization of a Limited Area Model”, *Mon. Wea. Rev.*, **110**, 1166-1186.

Application of TLNMC to regional model

Results with the North American Data Assimilation System (NDAS) were initially encouraging, with significant reduction of gravity wave noise and small but consistent improvement in forecasts. However, these tests were run in near real time instead of running the same time period repeatedly, so it was difficult to make meaningful adjustments without this ability.

A parallel system ran for 6 months, with results over the first month significantly better than the operational NDAS, but over time the test system gradually deteriorated until it was a little worse compared to operations.

Application of TLNMC to regional model

Because of the large additional expense of running TLNMC for no apparent benefit, it has been left idle for the regional application. In the meantime, efforts at EMC were redirected towards adding hybrid ensemble var as an option in GSI, for both global and regional application.

In the next talk, Daryl Kleist will continue with more recent developments with GSI and how TLNMC might be used.