The GMAO 4DVAR and its Adjoint Tools

Ricardo Todling, and Yannick Tremolet
Global Modeling and Assimilation Office, NASA/GSFC, Greenbelt, MD 20771

Abstract

The fifth generation of the Goddard Earth Observing System (GEOS-5) Data Assimilation System (DAS) is a 3d-var system that uses the Grid-point Statistical Interpolation (GSI) system developed in collaboration with NCEP, and a general circulation model developed at Goddard, that includes the finite-volume hydrodynamics of GEOS-4 wrapped in the Earth System Modeling Framework and physical packages tuned to provide a reliable hydrological cycle for the integration of the Modern Era Retrospective-analysis for Research and Applications (MERRA). This MERRA system is essentially complete and the next generation GEOS is under intense development. A prototype next generation system is now complete and has been producing preliminary results. This prototype system replaces the GSI-based Incremental Analysis Update procedure with a GSI-based 4d-var which uses the adjoint of the finite-volume hydrodynamics of GEOS-4 together with a vertical diffusing scheme for simplified physics. As part of this development we have kept the GEOS-5 IAU procedure as an option and have added the capability to experiment with a First Guess at the Appropriate Time (FGAT) procedure, thus allowing for at least three modes of running the data assimilation experiments.

The prototype system is a large extension of GEOS-5 as it also includes various adjoint-based tools, namely, a forecast sensitivity tool, a singular vector tool, and an observation impact tool, that combines the model sensitivity tool with a GSI-based adjoint tool. These features bring the global data assimilation effort at Goddard up to date with technologies used in data assimilation systems at major meteorological centers elsewhere.

Four-dimensional Variational Approach

The general cost function of the variational formulation

$$\begin{aligned} J(\mathrm{x}) &= \frac{1}{2}(\mathrm{x}_0-\mathrm{x}^b)^T\mathrm{B}^{-1}(\mathrm{x}_0-\mathrm{x}^b)+J_c \\ &+ \frac{1}{2}\sum_{k=0}^{K}[\mathrm{h}(\mathrm{x}_k)-\mathrm{y}_k]^T\mathrm{R}_k^{-1}[\mathrm{h}(\mathrm{x}_k)-\mathrm{y}_k] \\ &+ \frac{1}{2}\sum_{k=1}^{K}[\mathrm{m}(\mathrm{x}_k)-\mathrm{x}_k]^T\mathrm{Q}_k^{-1}[\mathrm{m}(\mathrm{x}_k)-\mathrm{x}_k] \end{aligned}$$

where

- ▷ $\mathrm{x}\equiv[\mathrm{x}_0,\mathrm{x}_1,\cdots,\mathrm{x}_K]^T$ is a 4d state vector;
- ▷ $\mathbf{h}_k$ and $\mathbf{m}_k$ are the nonlinear observation and dynamical model operators, respectively;
- ▷ $\mathbf{B}$, $\mathbf{Q}_k$, and $\mathbf{R}_k$ are the background, model, and observation error covariances, respectively.
- ▷ Strong constraint formulation: $\mathbf{Q}_k \to \infty$;
- ▷ Weak constraint formulation, $\mathbf{Q} \neq \mathbf{0}$ accounts for imperfections in the model $\mathbf{m}$;
- ▷ J_c represents a balance constraint.

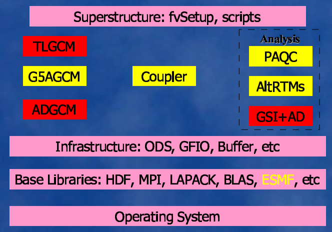
Incremental Variational Formulation

The minimization of J is generally treated in the Gauss-Newton sense where an iterative procedure linearizes the cost function at each, so called, outer loop, turning the problem in to a quadratic minimization problem for the following function

$$\begin{aligned} J_j \equiv J(\delta\mathrm{x}_j) &= \frac{1}{2}(\delta\mathrm{x}_j-\delta\mathrm{x}_j^b)^T\mathrm{B}^{-1}(\delta\mathrm{x}_j-\delta\mathrm{x}_j^b) \\ &+ \frac{1}{2}(\mathrm{H}_j\mathrm{x}_j-\mathrm{d}_j)^T\mathrm{R}^{-1}(\mathrm{H}_j\mathrm{x}_j-\mathrm{d}_j) \end{aligned}$$

where $\mathrm{d}_j \equiv \mathrm{y}-\mathrm{h}(\mathrm{x}_{j-1})$, $\delta\mathrm{x}_j^b \equiv \mathrm{x}^b-\mathrm{x}_{j-1}$, and

- ▷ $\delta\mathrm{x}_j \equiv \mathrm{x}_j-\mathrm{x}_{j-1}$ is the control variable;
- ▷ $\mathbf{R}$ is a 4d matrix combining the matrices $\mathbf{R}_k$ and $\mathbf{Q}_k$;
- ▷ The inner loop minimization of J_j can be solved by
  - Conjugate gradient
  - Quasi-Newton (such as L-BFGS)
  - Lanczos
- ▷ Conditioning of the J_j minimization is determined by the Hessian $\nabla^2 J_j = \mathrm{B}^{-1}+\mathrm{H}_j^T\mathrm{R}^{-1}\mathrm{H}_j$, which spectrum is such that a good preconditioning is essential, particularly in 4dvar.



Superstructure: fvSetup, scripts

TLGCM

G5AGCM

ADGCM

Coupler

Analysis

PAQC

AltRTMs

GSI+AD

Infrastructure: ODS, GFIO, Buffer, etc

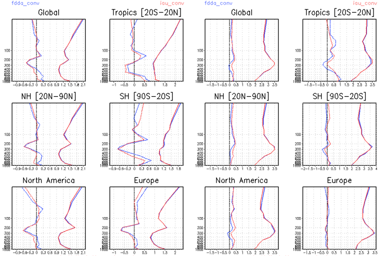
Base Libraries: HDF, MPI, LAPACK, BLAS, ESMF, etc

Operating System

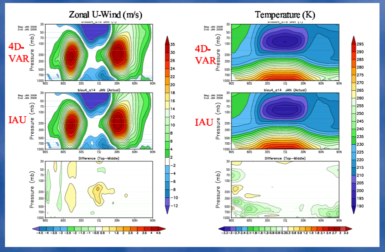
Experimental Setup
IAU and 4DVAR in GEOS-5

- 2x2.5x72 resolution
- Period: January 2006
- Comparison based on Monthly Means and Residual Statistics
- Both 3dvar and 4dvar use Lanczos-based CG; 50/30iterationss

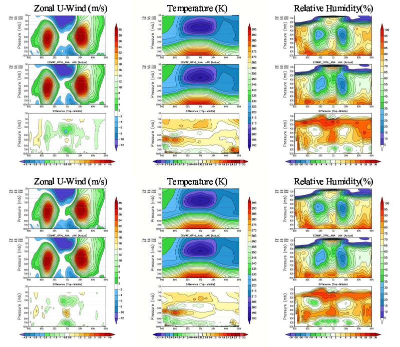
The first test of the newly implemented 4DVAR is to compare its results against the present GMAO IAU-based 3dvar assimilation system. Using the set up above, we ran an experiment for January 2006 and illustrate the comparison with IAU in the following few figures. The statistics of various observation-minus-background residuals is a crucial quantity to evaluate. The figure below shows zonal wind (6 panels on the left) and temperature residuals (6 panels on the right) for radiosondes. Introduction of 4DVAR affects mainly the biases in these residuals and indeed, in the case of the winds mid-tropospheric wind and temperature biases seem to improve slightly with 4DVAR, with a small degradation of the low-tropospheric temperatures.



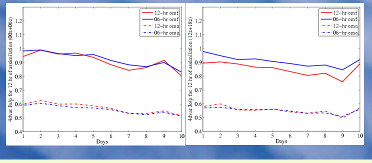
In the zonally-averaged monthly mean sense IAU and 4DVAR behave fairly similar at a first glance. Closer examination of the difference shows more explicitly the differences in biases between IAU and 4DVAR. This is illustrated in the figure below for zonal wind (left) and temperature.



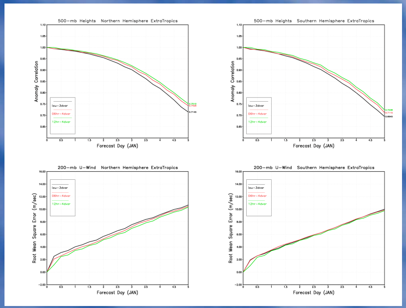
Though the observation-minus-background residual statistics is already very telling for when and where 4DVAR is an improvement over IAU, or is still deficient, comparison against other center fields can help further illustrate some points. Below, we compare the zonally-averaged monthly means of zonal wind, temperature and relative humidity with ECMWF operational fields for IAU (top) and 4DVAR (bottom).,



First we note that at this low 2-degree resolution the GEOS-5 IAU-3dvar system has a strong bias in the tropical zonal wind (not necessarily typical of the GEOS-5 full resolution system); a troposheric mid-latitude wind bias is also present to a lesser extent. Similarly, IAU-3dvar shows some mid-troposphere and low-stratospheric biases in the temperature field. The comparison with ECMWF seems to corroborate the residual statistics results showing Improved tropical tropospheric winds; indeed even the mid-latitude winds seems to improve considerably with 4DVAR. The temperature fields also improve in the low-stratosphere, mid-troposheric tropics and NH, with some deterioration over the mid-latitudes SH mid-to-low troposhere. The relative humidity fields are consistent with the temperature changes.



While extending the time-window of 4DVAR from 6- to 12-hr we looked at the normalized observation cost function, indicative of the goodness of fit, before the minimization starts and at the end of the minimization. A 10-day time series of Jo/p is shown above for both the 6- and 12-hr 4DVAR. To make the proper comparison we separate the results in two consecutive 6-hr batches for the 6-hr 4DVAR. The results are an indication that in the average the assimilation convergences similar in both cases. Indeed, the 12-hr 4DVAR seems to give slightly better initial fits for the latter part of the solar day.



Finally, we show results of forecast skill scores for IAU and both 4DVAR setups using either a 6- or a 12-hr time window. We remind the reader that these scores are for a low resolution GEOS-5 DAS and do not reflect the actual scores when the full resolution DAS is used. The figures above confirm what the residuals and the comparison with ECMWF operational analysis indicate, that is, that the 6-hr 4DVAR is indeed an improvement over IAU-3dvar particularly in the NH extra-tropics. Furthermore, the figures also show that the 12-hr 4DVAR provides either neutral or improved scores over the 6-hr 4DVAR results. Though preliminary, these results are very encouraging.

Observation Sensitivity and Impact

Take $\mathcal{F}$ to be the measure of an aspect of the forecast one wishes to examine. Since the forecast if a function of the background field x^b and the observations y, this measure is a convolution of operations:

$$\mathcal{F}(\mathrm{x}^b,\mathrm{y}) = F \bullet M \bullet G(\mathrm{x}^b,\mathrm{y})$$

where M and G represent the forecast model and the data assimilation system, respectively.

The sensitivity of the forecast to observations is

$$\frac{\partial\mathcal{F}}{\partial\mathrm{y}} = \mathrm{G}^T\mathrm{M}^T\frac{\partial F}{\partial\mathrm{x}^f} = \mathrm{G}^T\frac{\partial F}{\partial\mathrm{x}^a}$$

which in general requires second order adjoint information (Le Dimet et al 2002).

The observation impact, defined as the change in the forecast aspect $\mathcal{F}$ due to a set of observations can be approximated to first order by

$$I_1 \equiv <\frac{\partial F}{\partial\mathrm{x}},\delta\mathrm{x}>$$

where $\delta\mathrm{x}$ represents an analysis increment.

For a linear DAS, $\delta\mathrm{x}=\mathbf{Kd}$, here,

$$I_1 = <\frac{\partial F}{\partial\mathrm{x}},\mathbf{Kd}> = <\mathbf{K}^T\frac{\partial F}{\partial\mathrm{x}},\mathbf{d}> = <\frac{\partial F}{\partial\mathrm{y}},\mathbf{d}> .$$

For a nonlinear DAS, the increment is a successive correction of the linear-type increment, that is,

$$\delta\mathrm{x}_j = \mathbf{K}_j\mathbf{d}_j + (\mathbf{I}-\mathbf{K}_j\mathbf{H}_j)(\mathrm{x}^b-\mathrm{x}_{j-1}) .$$

The final (total) increment is

$$\delta\mathrm{x} = \sum_{j=1}^{m}\mathbf{K}_m\mathbf{H}_m\cdots\mathbf{K}_{j+1}\mathbf{H}_{j+1}\mathbf{K}_j\mathbf{d}_j .$$

Therefore, for a nonlinear DAS, the first order impact is approximately

$$I_1 = \sum_{j=1}^{m} <\mathbf{K}_m\mathbf{H}_m\cdots\mathbf{K}_{j+1}\mathbf{H}_{j+1}\mathbf{K}_j\frac{\partial F}{\partial\mathrm{x}},\mathbf{d}_j> .$$

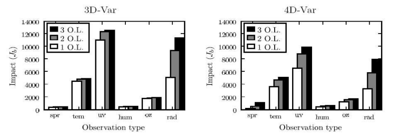
Higher order terms for linear DAS have been derived and discussed in Errico (2007) and Gelaro et. (2007); Tremolet (2008) gives a comprehensive discussion and derivation for nonlinear DAS.

The Adjoint of a Variational Analysis System

Observation sensitivity and impact studies require the adjoint of the underlying data assimilation system. The model adjoint provides the model sensitivity as the input to the analysis adjoint for calculation of the observation sensitivities.

Concentrating on the analysis adjoint, there are at least three ways to obtain the adjoint of a variational analysis system:

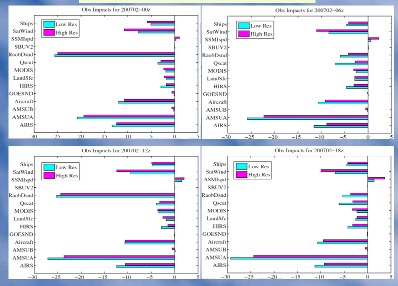
- ▷ Direct, line-by-line, adjoint (Zhu & Gelaro 2007)
- ▷ Operator manipulation:
  - Observation space (Baker & Daley 2000): $\mathrm{K}^T\partial F/\partial\mathrm{x} = \delta\mathrm{z}$
    $$(\mathrm{HBH}^T+\mathrm{R})\delta\mathrm{z} = \mathrm{HB}\frac{\partial F}{\partial\mathrm{x}}$$
  - In physical/spectral space (Tremolet 2008): $\mathrm{K}^T\partial F/\partial\mathrm{x} = \mathrm{R}^{-1}\mathrm{H}\delta\mathrm{x}$
    $$(\mathrm{B}^{-1}+\mathrm{H}^T\mathrm{R}^{-1}\mathrm{H})\delta\mathrm{x} = \frac{\partial F}{\partial\mathrm{x}}$$
- ▷ Approximating the Hessian (Cardinali et al. 2004): $\mathrm{K}^T\partial F/\partial\mathrm{x} = \mathrm{R}^{-1}\mathrm{H}\tilde{\mathrm{A}}\partial F/\partial\mathrm{x}$
  $$\tilde{\mathrm{A}} \sim \sqrt{\mathrm{B}^{-1}+\mathrm{H}^T\mathrm{R}^{-1}\mathrm{H}}$$



The figure above shows the observations impact on Jb for 100 inner-loop iterations and three outer loop iterations. Results for 3dvar (left) and 4dvar (right) are shown for each of the control variables: 4dvar outer loop have larger impact than 3dvar outer loops; the overall impact of 4dvar is smaller than that of 3dvar – 4dvar makes more consistent use of observations through the time window – Jo is smaller in 4dvar even though a larger number of observations end up being used (not shown).

Experimental Setup
Observation Impacts on 24-hr forecast for resolutions of GEOS-5 DAS

- Two experiments one at 2x2.5x72;
- another at 0.5x0.625x72 resolution
- Period: February 2007
- Each experiment verify against its own analysis



The figures above show observation impacts on the 24-hr forecasts from GEOS-5 DAS (3dvar-IAU) over the period of February 2007. Results for a low-resolution (blue bars) and the full high resolution DAS are displayed. The impacts are separated into each of the four synoptic analysis times and grouped as some of the the main observing system types currently used in GEOS-5. Typical of the kinds of results one can extra from these plots is that the observing systems of most significance to the 24-hr forecast are the radiosondes and the AMSUA radiances; the impact of AIRS, aircrafts, and satellite winds are also very significant. Moreover, one see that AMSUA compensates at times when the radiosonde observing network is less active (at 06z and 18z). Also, we see that the relative impacts of the various observing systems does not change significantly when resolution changes making it feasible to use low-resolution experiments to quickly assess overall impact of various (new) data types. *Note: positive impacts on the 24-hr forecast are negative numbers; negative impacts are positive numbers.*

Summary and Conclusions

- NASA GEOS DAS has entered the 4D-world (again!)
- Various adjoint tools are now available in GEOS-DAS, capable of performing studies in forecast sensitivities, singular vectors, analysis sensitivity and observations impact
- First exercise including some of these tools will be the Observations Impact Inter-comparison Study (ECMWF, Env. Canada, Meteo France, NASA, and NRL)
- Hooks for weak constraint are in place in GSI and soon will be in place in the GEOS-5 GCM Work is on way to update the GCM TLM/ADM with cube-sphere core
- Soon we will be able to compare 4DVAR with NCEPs approximate 4D-scheme; First Order Time-interpolation to Observations (FOTO)

The implementations done thus far benefited greatly from the incredible infrastructure of GSI.

Background photo from Cloud Appreciation Society by: Katrin Pfeifer, Afternoon sky with the moon, taken in Castagneto Carducci, Tuskany, Italy.

International TOVS Study Conference, 16th, ITSC-16, Angra dos Reis, Brazil, 7-13 May 2008. Madison, WI, University of Wisconsin-Madison, Space Science and Engineering Center, Cooperative Institute for Meteorological Satellite Studies, 2008.