GSI Data Assimilation System: From Operations to the Community

Hui Shao, Ming Hu, Louisa Nance, Hans Huang, and John Derber
NCAR, Boulder, CO

Abstract
The Gridpoint Statistical Interpolation (GSI) 3D-Var system is currently part of NOAA/NCEP's operational Global Forecast System (GFS) and North American Mesoscale Model (NAM) and will be part of the WRF Rapid Refresh, which is slated to replace the Rapid Update Cycle run at NCEP in early 2010. It is also planned to become operational at the Air Force Weather Agency (AFWA) by the end of 2010 or the beginning of 2011. In terms of the WRF community, the WRF Developmental Testbed Center (DTC) plans to add the GSI DA system to the NWP code it supports and provide an official GSI code release in the coming months. Currently, the DTC is working closely with the NCEP/EMC, NCAR/RAL/DATC, NOAA/ESRL/GSD and NCAR/ESSL/MMM to conduct a review of the GSI system and prepare for its release. This paper will introduce the current efforts on reviewing the GSI system and present some preliminary testing results.

1. Overview

Data assimilation (DA) is a core component of any earth system model and has provided many of the major advances in Numerical Weather Prediction (NWP) in recent years. The advent of advanced, internationally-coordinated observing systems (e.g., the National Polar-Orbiting Operational Environmental Satellite System (NPOESS)) and new observation types (e.g., hyperspectral sounders, GPS radio occultation, dual polarization Doppler radar) means this trend is likely to continue for the foreseeable future.

The development of the Gridpoint Statistical Interpolation (GSI) 3D-Var system has provided the WRF research and operational communities with state-of-the-art, efficient DA capabilities. In terms of operations, the GSI is currently part of NOAA/NCEP's operational Global Forecast System (GFS) and North American Mesoscale Model (NAM) and will be part of the WRF Rapid Refresh (RR), which is slated to replace the Rapid Update Cycle (RUC) run at NCEP in early 2010. In 2008, the Air Force Weather Agency (AFWA) decided to transition from its current operational DA system to the GSI for its high-resolution, multi-theater and world-wide NWP by the end of 2010 or the beginning of 2011. Hence, applications of the GSI DA system are planned to span from global to regional scales at high resolution by 2011. The number of groups involved in GSI development has recently expanded from the central development group located in Maryland (NCEP and NASA) to include two development groups located in Boulder, Colorado (NOAA/ESRL/GSD and NCAR/ESSL/MMM).

Starting from 2009, the DTC plans to add the GSI DA system to the NWP code it supports for the WRF community. The DTC currently focuses on testing and evaluation of WRF for limited area NWP applications, but transitioning to include global forecast applications is included in the DTC long range planning, so it is in the DTC's interest to support a GSI system that includes the various capabilities developed by all the GSI developers. Hence, the DTC is ideally situated to play a role in coordinating code sharing among the distributed development groups.

As the DTC adds the GSI to the suite of codes it supports for the WRF community, it will also be adding this capability to its testing and evaluation efforts, making the DTC ideally situated to play a key role in the testing activities directed at transitioning the AFWA operational forecast system to include GSI while not degrading the forecast performance. During this transition period, an end-to-end WRF-ARW and GSI system will be set up at NCAR parallel to the AFWA operational system. This end-to-end system will

undergo an extensive series of experiments to test and tune the GSI system to prepare it for operational implementation. These experiments will also help to establish and maintain WRF Reference Configurations.

The DTC is a component of the Joint Numerical Testbed (JNT) within the NCAR's Research Applications Laboratory (RAL). The DTC will work closely with the Data Assimilation Testbed Center (DATC), which also resides in the JNT. The cost of coordinating the GSI development groups and supporting the GSI to the community will be shared with other WRF DTC partners.

This paper presents the on-going efforts at the DTC to bring the GSI to the community. Section 2 introduced the GSI system briefly. Section 3 introduced the current work at the DTC on the GSI public release and community support. Finally, Section 4 gives a brief summary.

2. GSI System

The GSI was initially developed as a next generation global analysis system based on the Spectral Statistical Interpolation (SSI) analysis system used operationally at NOAA/NCEP. Instead of being constructed in spectral space like the SSI, the GSI is constructed in physical space by replacing spectral definition for background errors with grid point definition based on recursive filters. After initial development, the GSI analysis system is evolving and has become a unified data analysis system for both regional and global applications. It became operational in June 2006 for regional and May 2007 for global analysis. It is also being used to generate the Real-Time Mesoscale Analysis (RTMA) products.

The cost function to be minimized in the GSI is

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

where x is a vector of the analysis increments, B is the background error covariance matrix, y is a vector of the observational residuals, $y=y_{obs}-Hx_b$, R is the observational and representativeness error covariance matrix, H is a transformation operator from the analysis variable to the form of the observation vector, and Jc is a constraint term. The analysis variables are streamfunction, unbalanced part of velocity potential, unbalanced part of temperature, unbalanced part of surface pressure, pseudo-relative humidity (or normalized relative humidity), ozone, and cloud liquid water. The GSI can assimilate various observational data types, including conventional data, GPS radio occultation refractivity, SSMI-rain, TMI-rain, SBUV, GOES, AMSU-A and B, HIRS2, 3 and 4, MHS, MSU, and AIRS data.

The most significant differences between the global and regional systems include the basic coordinate systems of the two and the current use of the strong constraint in the global system. Other differences are primarily options, e.g., which data are used, specification of background errors, etc.

3. GSI Status at the DTC

Supporting and managing a version of the GSI DA system applicable to the community model system requires a plan for coordinating the sharing of code between the various developers. During this funding cycle, the DTC is working with the various developer groups to put together a plan for maintaining a community repository. This plan includes the following components:

- A plan for how the code will be stored and supported in a repository under version control.
- A policy on who can check-in new code to the community-contributed repository, as well as testing requirements and approval process.

- A schedule for syncing repositories if multiple repositories exist.
- Release schedules and testing requirements.

3.1 Repository

The primary GSI repository will be at the NCEP/EMC (consistent with operational forecasting resident at NCEP). The NCEP GSI repository will be closely linked with a second complementary GSI repository in Boulder to be used by Boulder GSI partners as a focal point for Boulder GSI related research and as the DTC community repository. The code contained in the repository is developed by members of the GSI community. Newly developed methods or enhancements to existing methods offered as contributions to the GSI system must address a potential need or make improvements over current techniques. Before code can be committed into the GSI repository, it must be written in the GSI software framework, meet GSI coding conventions and standards, and pass designated testing.

Proposed changes offered by a developer to the GSI repository are reviewed at regular (e.g., bi-weekly or weekly) meetings of the GSI developers. Once changes are approved, and before they are checked into the GSI code repository, each bundle of proposed changes must pass a basic regression test.

3.2 Code release

The release schedule is drafted as show in Table 1.

Tasks	Time
Beta release (Q1FY09)	Jun 26, Friday, 2009
WRF Workshop: Instructional session	Jun 26, Friday, 2009
WRF/WRF-Var tutorial: GSI introduction	July 20-22, 2009
First release (Q1FY09)	Sep, 2009
Tutorial	Jan, 2010

3.3 User's support

A draft of the GSI User's Guide is under construction by the DTC. This User's Guide provides instructions on the GSI code installation and compilation and descriptions of GSI implementation and diagnostics. A GSI webpage has been set up within the NCAR's RAL and will serve as a platform for the code release and user's support. Related information on the GSI will be posted over there. A regular GSI tutorial will be held (annual or bi-annual). The first full tutorial is scheduled in January, 2010.

3.4 Testbed

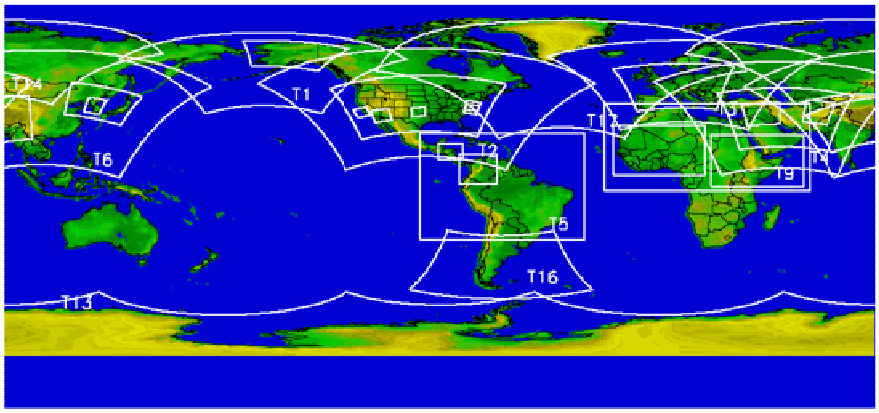

Fig.1 The AFWA theater.

Testing and evaluation of the GSI DA system in regional applications will be carried out by staff affiliated with the DATC in order to take advantage of the DA expertise of this group. Fig.1 shows the possible testing domains in the AFWA theater. The DATC will set up an end-to-end system (Fig.2) including the latest available version of GSI from the GSI repository. This end-to-end system will include the WPS and WRF-ARW version selected in collaboration with MMM and the AFWA DA project. This end-to-end system will use the DTC's Model Evaluation Tool (MET) to compute objective verification scores. One extensive run of the WRF-ARW (standard "NODA" run) will be generated to serve as a benchmark. The second step in this testing process will be to carry out extensive runs of the WRF-ARW and GSI system. The capability and robustness of the WRF-ARW and GSI system in regional applications will be evaluated by comparing the results with the NODA results. This test will be extended to consider the performance of different cycling schemes (e.g., full cycling and limited cycling) of the WRF-ARW and GSI system.

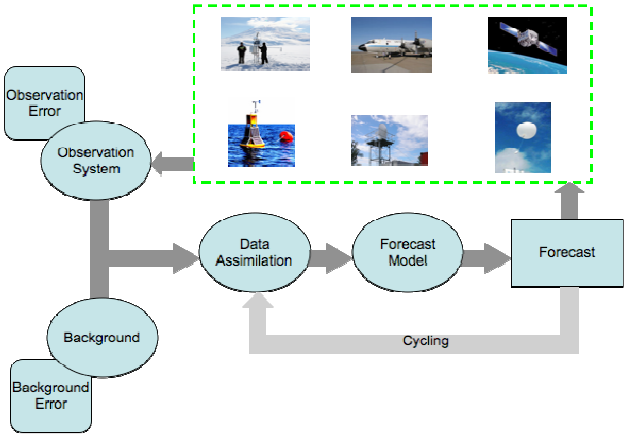

Fig.2 DATC end-to-end system.

4. Summary

The WRF Developmental Testbed Center plans to add the GSI DA system to the NWP code it supports and provide an official GSI code release in the coming months. Currently, the DTC is working closely with the NCEP/EMC, NCAR/RAL/DATC, NOAA/ESRL/GSD and NCAR/ESSL/MMM to conduct a review of the GSI system and prepare for its release. This paper introduces the on-going efforts in the DTC for code release and support of the GSI.