6A.2

GSI DATA ASSIMILATION SYSTEM: COMMUNITY SUPPORT AND PRELIMINARY TESTING RESULTS

Hui Shao *, Ming Hu, Kyungjeen Park, Kathryn Crosby
Hans Huang, Louisa Nance, Bill Kuo
NCAR, Boulder, CO
and
John Derber, Michael Lueken
NCEP/EMC, Camp Springs, CO

1. Introduction

The Gridpoint Statistical Interpolation (GSI) 3D-Var system is currently part of the operational Global Forecast System (GFS) and North American Mesoscale (NAM) Model supported by National Centers for Environmental Prediction (NCEP) and will be part of the WRF Rapid Refresh system, which is slated to replace the Rapid Update Cycle run at NCEP in 2010. It is also planned to become operational at the Air Force Weather Agency (AFWA). In terms of the Weather Research and Forecasting (WRF) community, the Developmental Testbed Center (DTC) has added the GSI data assimilation (DA) system to the numerical weather predication code it supports and provided an official GSI code release starting from 2009. Currently, the DTC is working closely with various groups to conduct testing and evaluation of the evolving GSI system. This paper will introduce the current efforts on the GSI community release and support and present some preliminary testing results on the end-to-end system of WRF and GSI in regional applications.

2. GSI code management and support

The cornerstone of the GSI community work is to set up a community GSI repository to accommodate distributed developments. The DTC worked with the various developer groups to put together a plan for maintaining this community GSI repository. This plan includes the following components:

- A description of how the code is 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.

The code management plan was approved by the NCEP/Environmental Modeling Center (EMC) and other GSI developers and has been implemented.

* *Corresponding author address*: Hui Shao, NOAA/NWS/NCEP/EMC, 5200 Auto Rd, Camp Springs, MD 20746; e-mail: huishao@ucar.edu

The DTC has started to use this repository for community release and support as well as community tool developments. As a starting point, the 2009 operational GSI code and libraries in NCEP global applications (Q1FY09) were ported, tested, and built into this repository. This version of GSI has been tested on IBM and Linux with various compilers. The DTC has built a configure/compile scripting tool for the community GSI code to make it easy to install on different platforms. The DTC has also developed a regression test suite including test cases in global and regional domains working with both binary and NETCDF formatted initial conditions.

Using this Community GSI Repository, Beta version 1.0 of the GSI code and related libraries were released in June 2009. The formal release of the GSI system (v1.0) was in September, 2009. The DTC also started preparing for the second GSI community release (v2.0). which will be based on the Q1FY10 code implemented in the GFS system by NCEP Central Operations (NCO) in December, 2009. The DTC has been testing this code on various platforms and is ready for a beta release at the beginning of February, 2010. The formal release of v2.0 is scheduled in April, 2010 after finishing more tests and updating the GSI User's Manual.

The DTC has been working on various components to provide GSI community support. The DTC has set up a GSI helpdesk (gsi_help@ucar.edu) and appointed dedicated staff working as a frontline to receive help requests and answer questions or distribute questions to developers/experts. The GSI Users' Guide is updated and reviewed by the Boulder GSI groups and NCEP/EMC before release, and is consistent with the latest release. The DTC has built a GSI website (http://www.dtcenter.org/com-GSI/users/) accessible by the general user community, including GSI related announcements, system downloading, GSI documents, on-line tutorial, etc. Community users are required to register through this GSI website to download the GSI release code and documentation. The DTC has also set up a GSI wiki page resident at UCAR with authorized access for GSI developers to share GSI information. With help from the NCEP GSI group, the DTC staff gave a GSI instructional session at the 10th Annual WRF Users' Workshop in June, 2009 and a GSI lecture at the

WRF Data Assimilation Tutorial in July, 2009. The DTC has been working with NCEP/EMC to prepare the first Community GSI Tutorial scheduled for June 28-30, 2010.

3. Testing and evaluation of GSI DA system in regional applications

3.1 Experiment design

An end-to-end system, including the WRF Pre-processing System (WPS, v3.1), GSI (v1.0), Boundary Update (UPDATE_BC) utility, Advanced Research WRF (ARW, v3.1), WRF Postprocessor (WPP, v3.1), and Model Evaluation Tools (MET, v2.0) system, was set up at the DTC for this GSI Testbed. WPS, ARW and WPP are among the community models supported by DTC and NCAR/MMM. UPDATE_BC is a utility to update lower and lateral boundary conditions using the analysis from a data assimilation system. It has been developed and supported by NCAR/MMM. Figure 1 shows the flowchart of the system.

Using this testbed, several extended runs were conducted with two goals: (1) To test and evaluate the capability and robustness of the GSI+ARW in regional applications; (2) To test and evaluate the performance of alternative cycling schemes. The geographic domain for this task (shown in Fig. 2) corresponds to a Caribbean domain. The forecasts and post-processing are configured with a 15km horizontal resolution, 57 vertical levels and the model top at 10mb. The testing period is August 15 – September 15, 2007.

Five month-long experiments were conducted for the testing period:

- GFSWRF: ARW runs. Started from GFS analysis every 6 hours (“cold-start” mode).
- CYC_CONV: GSI and ARW runs in a full cycling mode. PrepBUFR (NCEP featured BUFR format) data were assimilated every 6 hours.
- CYC_CONV+RAD: Same as CYC_CONV except AMSU-A radiance data were assimilated as an addition.
- CYC_CONV+RAD+GPS: Same as CYC_CONV+RAD except GPS RO refractivity were assimilated as an addition.
- LCYC_CON+RAD+GPS: Same as CYC_CONV+RAD+GPS except the runs were in a partial cycling mode.

The “cold-start” mode here means the GSI runs were initialized by background from other independent systems/analyses/forecasts. The “full cycling” mode means that the GSI runs were initialized with the 6 hour ARW forecasts from the previous cycle. The “partial cycling” is a combination of the “cold-start” and “full cycling” modes. For this specific testbed, it means that an analysis was generated by cycling the GSI and ARW every 6 hours starting from 12 hours before the analysis time. The background for the GSI at -12 hour was from the GFS 6 hour forecasts. For all the above-mentioned experiments, 48 hour forecasts were performed at 00Z and 12Z for verification purposes. All data assimilation runs used the global background error file provided by NCEP/EMC, following the suggestion from the NCEP/EMC. The BE data were examined in the testing domain through standard diagnostics and a series of pseudo single-observation tests.

For verification, the analyses and forecasts were interpolated to the location of the observations. The statistics were computed using MET at 17 vertical levels up to 10hpa except for relative humidity up to 200hpa.

3.2 Analysis Verification

To evaluate the performance of the GSI system, the analyses produced from the experiments were verified against conventional observations and brightness temperature. As an example, Fig. 3 shows three time series of the number of brightness temperature observations and mean and standard deviation of brightness temperature background with and without bias correction and analyses (OMA) from the experiment CYC_CONV+RAD. They show a significant reduction of the data bias and root-mean-square (RMS) errors after the bias correction in GSI for this specific channel.

The first try to assimilate COSMIC data failed since most of COSMIC data were rejected due to the strict QC procedure in the GSI (v1.0) near the tropical area. Based on the discussion with NCEP/EMC, similar QC procedures to the latest GSI code were adopted to the current GSI code for testing purposes. An example of the coverage of the assimilated data within a 6 hour data assimilation window is shown in Fig. 2. Half of the observations were accepted and assimilated by the GSI and the analysis fits to the observations were checked which shows the code works reasonably well.

3.3 Forecast Verification

To evaluate the performance of the system as well as the impacts of the radiance and GPS observations, monthly averaged statistics were examined. Fig. 4 shows the vertical profiles for the analyses and forecasts of specific humidity and temperature from CYC_CONV, CYC_CONV+RAD and CYC_CONV+RAD+GPS. Overall, the values of the biases and RMS errors of the moisture and temperature analyses and forecasts are reasonable indicating the system is working appropriately. The assimilation of radiance data reduces the biases of the moisture analysis and forecast fields in the lower troposphere and however increases the bias above. For the temperature, the impacts of the radiance data are more significant on the analyses by reducing the temperature biases at high levels. The assimilation of extra GPS data presents minor impacts. This might be due to the limited quantity of the GPS data assimilated in the domain throughout the testing period. The temperature field also shows a large bias at the model top.

To examine the impact of the cycling scheme, the results from LCYC_CONV+RAD+GPS and CYC_CONV+RAD+GPS were compared. The monthly statistical results are mixed and do not show preference toward any scheme. As a case study, the DTC also examined a hurricane event, Hurricane Dean, during the testing period. The tracks of all experiments are quite similar. However, full cycling schemes produced the best density analysis compared with partial cycling runs (Fig. 5).

4. Summary

The DTC has created a GSI community repository to accommodate distributed development. Based on this repository, the DTC provides community support and conduct annual code release, and will hold the first GSI Community Tutorial in June, 2010.

The DTC has set up a script suite to test the GSI v1.0 in an end-to-end system, including WPS, ARW, UPDATE_BC, ARW, WPP and MET. Various month-long experiments were conducted in a Caribbean domain for the period of August 15 through September 15, 2007 for the following purposes: (1) To test and evaluate the capability and robustness of the GSI+ARW in regional applications; (2) To test and evaluate the performance of alternative cycling schemes.

Various aspects and components of the system have been investigated, including the background errors, observation sensitivity studies (conventional, radiance and GPS Radio occultation observations) and running schemes. The experiments show

- The GSI system coupled with WRF-ARW using global background errors is working well.
- AMSU-A radiance data assimilation has positive impacts on analyses and forecasts.
- The QC procedure for GPS data in GSI v1.0 rejected too many data in the testing domain (near-tropical domain). NCEP has found this problem and updated the QC procedure in GSI v2.0. Currently GSI v2.0 is being tested and will be released in April, 2010.
- The full cycling and limited cycling schemes show mixed results for the impact study however the full cycling run produced a better density forecast for the Hurricane Dean.

Figures

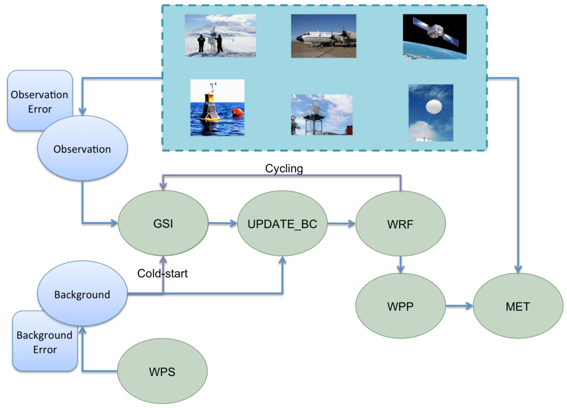

Fig. 1 Flowchart for the GSI Testbed suite.

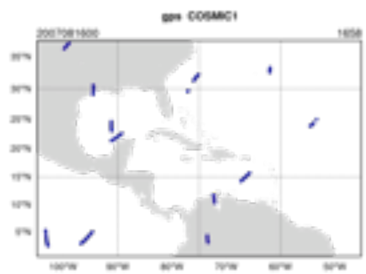

Fig. 2 Geographic domain for the GSI testing and evaluation. The blue dots represents the locations for the COSMIC RO observations within a 6 hour data assimilation window centered at 00Z August 16, 2007.

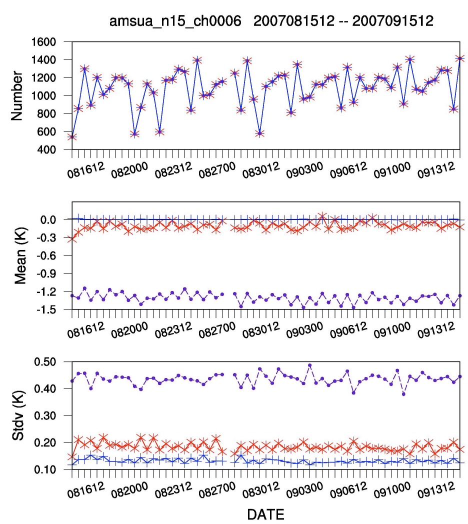

Fig. 3 Time series of the number of brightness temperature observations and mean and standard deviation of brightness temperature background with (red) and without bias correction (purple) and analyses (blue) from the experiment CYC_CONV+RAD.

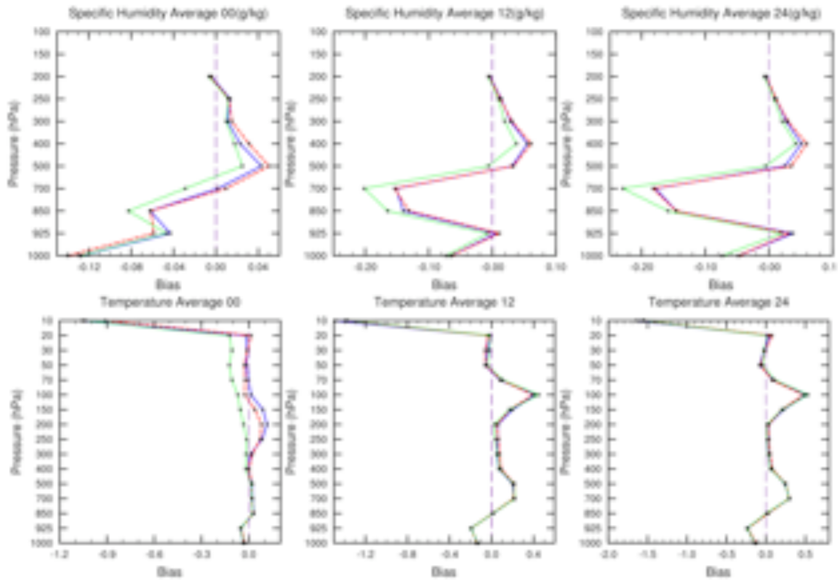

Fig. 4 The vertical profiles for the analyses (left panels) and 12 hour (middle panels) and 24 hour (right panels) forecasts of specific humidity (upper) and temperature (lower) from CYC_CONV (green curves), CYC_CONV+RAD(blue curves) and CYC_CONV+RAD+GPS (red curves).

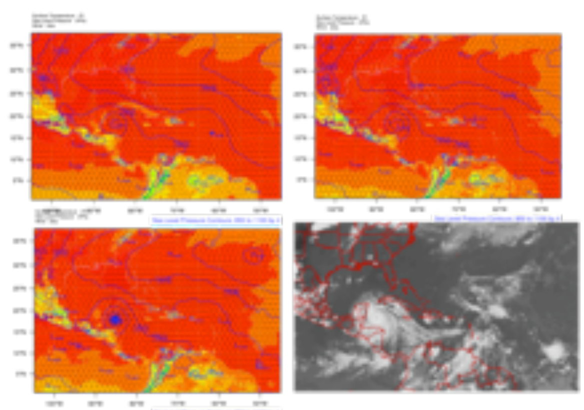

Fig. 5 The analyses of surface pressure (contours), temperature (shade) and winds from WRFGFS (upper left), CYC_CONV+RAD+GPS (lower left) and LCYC_CONV+RAD+GPS (upper right) and satellite observation (lower right).