5.3 TESTING AND EVALUATION OF THE GSI DATA ASSIMILATION SYSTEM

Kathryn M Newman*, C. Zhou, H. Shao, X.Y. Huang, M. Hu
National Center for Atmospheric Research, Boulder, CO
Developmental Testbed Center

1. INTRODUCTION

The Gridpoint Statistical Interpolation (GSI) is a three dimensional variational (3D-Var) data assimilation system that is being developed at NCEP/EMC, NOAA/GSD, NASA/GMAO, and NCAR/MMM. A community code version of the GSI system is also available and supported through the Developmental Testbed Center (DTC). GSI is run for both regional and global operations including the Global Forecast System (GFS), North American Mesoscale Model (NAM), WRF Rapid Refresh (RR), and Hurricane WRF (HWRF). This report will outline the current testing and evaluation efforts on the end-to-end system of WRF and GSI in regional applications being conducted at the DTC. These tests are conducted in order to determine the capability and robustness of the GSI coupled with WRF ARW in regional applications, as well as to evaluate the impact of a variety of existing and proposed operational data types. Finally, through this testing, the DTC strives to provide a rational basis for operational centers and the research community for advancements in the Numerical Weather Predication (NWP) systems.

2. EXPERIMENTAL DESIGN

The DTC set up an end-to-end system consisting of the WRF Pre-processing System (WPS, v3.2), GSI (Q1FY11), Boundary Update (UPDATE_BC) utility, Advanced Research WRF (ARW, v3.2), WRF Post-Processor (WPP, v3.1), and Model Evaluation Tools (MET, v2.0), for this GSI testbed.

WPS, ARW, and WPP are all community models supported by the DTC and NCAR/MMM. UPDATE_BC is a utility to update upper and lower boundary conditions using the analysis from the data assimilation system. This utility was both developed and is supported through NCAR/MMM.

Extended testing using this testbed were conducted using GSI Q1FY11 coupled with WRF-ARW v3.2. A one-month testing period was used from 15 August 2007 through 15 September 2007. The Air Force Weather Agency (AFWA) T8 domain was used (Fig. 1), which is a near tropical Caribbean domain, featuring the Atlantic Basin.

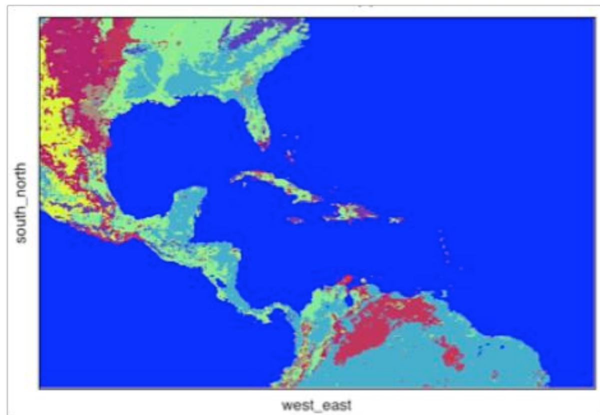


Figure 1: AFWA T8 domain used in all experiments

The model was configured with a 15 km horizontal resolution, 57 full vertical sigma levels, and a 10 mb model top. The domain has 418 west-east and 280 north-south grid points. Configuration studies were performed, resulting in the DTC namelist configuration closely representing that of the NAM, with one exception being the qoption flag. The flag qoption=1 was used for all extended tests, rather than qoption=2 following the NAM configuration. Preliminary tests using qoption=2 were performed; however further testing needs to be conducted before the DTC is confident in implementing this option for all ARW applications. Forecasts were run for 48 hours at 00 and 12 UTC for verification purposes.

For the background error (BE) information used in the tests, two types of BE statistics were provided by NCEP/EMC: The “global” BE statistics were calculated based on the GFS forecast difference and can be interpolated by the GSI code to any

* Corresponding author address: Kathryn M. Newman, National Center for Atmospheric Research, Boulder, CO, 80307-3000; e-mail: knewman@ucar.edu

regional domain; The regional BEs were produced from the NAM model in a CONUS domain, therefore, reflecting the NAM specific background features and might introduce uncertainties to our studies. Two types of BEs were compared through a series of pseudo single observations tests (not shown). Based on the results, global BEs were used for all extended tests.

Six 1-month experiments were conducted for this study:

- GFSWRF: Benchmark run. ARW runs started from GFS analysis every 6 hours

Six-hour full cycling runs using GSI included:

- CYC_CONV_default: GSI and ARW assimilating conventional prepBUFR data
- CYC_CONV_allobs: conventional prepBUFR data were assimilated with increased surface observations
- CYC_CONV_nosfc: No surface prepBUFR data were assimilated
- CYC_GPS: default conventional data and GPS Radio Occultation (RO) data were assimilated
- CYC_AMSUA: default conventional data and AMSU-A radiance data were assimilated
- CYC_AMSUB: default conventional data and AMSU-A and AMSU-B radiance data were assimilated

The GFSWRF run was done from a 'cold-start', meaning it was initialized from a background of an independent system (e.g., GFS analysis in this study). The remaining runs used GSI 6-hour full cycling data assimilation initialized with the 6-hour WRF ARW forecast from the previous cycle to conduct impact studies. For both the GFSWRF and the 6-hour full cycling GSI runs, the lateral boundary conditions (LBC) come from the GFS forecast.

All experiment results were verified using MET v2.0 against GDAS conventional prepBUFR data. Statistics were computed at 17 vertical levels up to 10 mb (plots up to 20mb) for temperature, u-component wind, and v-component wind. Specific humidity was verified up to 150 mb. Statistical significance was tested at the 95% confidence interval.

3. DATA IMPACT

The analysis and forecasts generated from CYC_CONV_default were first compared with those from GFSWRF to ensure the cycling runs were correctly conducted. It should be noted that the CYC_CONV_default run only assimilated prepBUFR conventional observations, where the initial conditions for the GFSWRF run are from the GFS analysis with all operational data types assimilated. The monthly scores are also over long time scales and large-scale dynamics. Figure 2 shows RMSE values for temperature and u-component wind for the analysis time, 24-hour forecast, and 48-hour forecast times verified against sounding data. Statistically significant (SS) differences at the analysis time highlight the differences in the initialization of the GSI runs compared to the benchmark GFS run. Forecast times show neutral to slightly negative differences favoring WRFGFS.

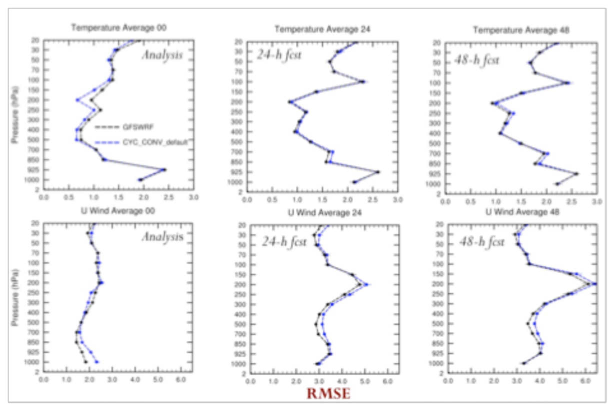


Figure 2: Vertical profiles for the analysis (left), 24-hour forecast (middle), and 48-hour forecast (right) for the temperature (upper) and u-component wind (lower) for GFSWRF (black) and CYC_CONV_default (blue).

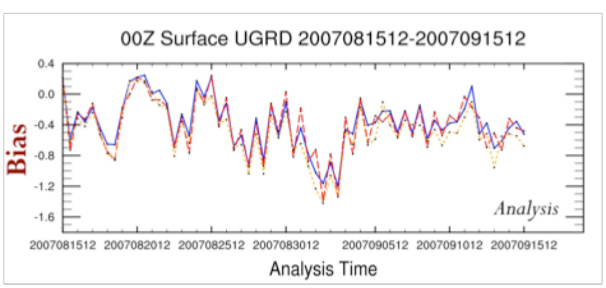
3.1 SURFACE DATA IMPACT

In order to evaluate the impact of surface data on the GSI data assimilation system, a series of surface impact runs were performed. First, the CYC_CONV_default run was conducted, assimilating observations with the default GFS quality marks for the GDAS data. Because the QC marks resulted in most of the surface data being rejected during GSI minimization, a second run was conducted. This run, CYC_CONV_allobs, loosened the QC mark threshold values for the surface data assimilation to match the data types allowed in the NAM. This process may introduce some bad observations to the analysis because the NCEP NAM QC procedure was essentially skipped; however, this run was for the purpose of evaluating the impact of surface observations and any potential additional data. Finally, a third run was conducted which removed all prepBUFR surface observations from the assimilation. Table 1 shows the QC mark value for each corresponding prepBUFR report type. All runs followed the NAM configurations, while changing the 'qcmark' threshold altered the amount of surface data assimilated. The CYC_CONV_default qcmark was set to 2.00, therefore only allowing surface observations marked in green in Table 1. The qcmark threshold was altered to 9.00 for CYC_CONV_allobs, allowing both green and red values listed in Table 1.

	ADPSFC 181		ADPSFC/ SFCSHP 183	ADPSFC 187	
Ps Q T	2.00 9.00 8.00		9.00 9.00 8.00	6.00 9.00 8.00	
	ADPS FC 281	SPSS MI 283	ADPSFC/ SFCSHP 284	QKSWN D 285	ADPSFC 287
U V	9.00	2.00	9.00	2.00	9.00

Table 1: QC mark value for corresponding prepbufr report type. Green values used in CYC_CONV_default run, red and green values used in CYC_CONV_allobs, and all observations in table were not used in CYC_CONV_nosfc.

Figure 3a shows u-component wind bias verification against prepBUFR surface observations for each surface impact run. Assimilation of the surface data reduced the wind bias through the 48 hour forecast. This positive impact has also propagated from the surface into the upper levels shown in the Figure 3b at the 24-hour and 48-hour forecast times. However, compared with CYC_CONV_default, impact of additional surface observations (CYC_CONV_allobs) are neutral.



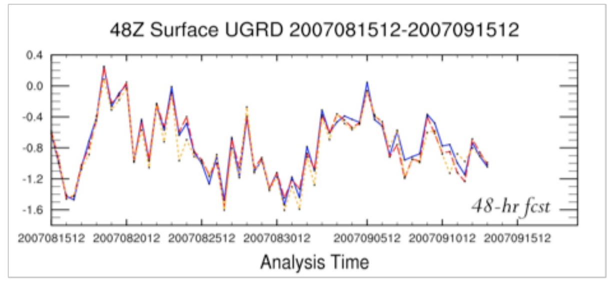


Figure 3a: Time series of surface u-component wind bias for analysis (top) and 48-hour forecast (bottom) for CYC_CONV_default (blue), CYC_CONV_allobs (red), and CYC_CONV_nosfc (yellow).

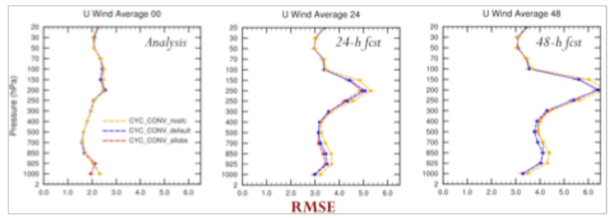


Figure 3b: Vertical profiles for the analysis (left), 24-hour forecast (middle), and 48-hour forecast (right) for the u-component of the wind field for CYC_CONV_default (blue), CYC_CONV_allobs (red) and CYC_CONV_nosfc (yellow).

3.2 GPS RO IMPACT

The CYC_GPS run was performed in order to evaluate the impact of assimilating GPS data compared to conventional data assimilation alone. Original testing for FY2009 on GSI v1.0 resulted in nearly all the data being rejected. Updates made to GSI v2.0 for the GPS RO assimilation reduced observation errors in the tropics, therefore increasing the data usage (Cucurull, 2010). These changes were evident in extended runs, as over 40% more data were used in the assimilation compared to GSI v1.0 (Table 2). The GSI v1.0 w/ '2.0' QC refers to code changes in FY2009 based on documented v2.0 changes in order to increase GPS RO data usage. GSI v2.0 reflects the changes after running GSI Q1FY11 for FY2010 testing. The 40% increase in data usage can be seen after v2.0 changes were implemented. Using GSI v2.0, the background fit to observations was also slightly improved with lower RMS values (0.97).

	Ob Rejection	**RMS**
GSI v1.0	~100%	-
GSI v1.0 w/ '2.0' QC	61.2%	1.16
GSI v2.0	59.1%	0.97

Table 2: COSMIC RO data rejection and diagnostic statistics for 2007081512

To ensure the GSI system was properly using the GPS data, the analyses produced were verified against conventional observations. Figure 4 shows three time series of the number of GPS RO observations, mean, and standard deviation of the analysis increment (OMA) and background innovation (OMB) with and without quality control (QC). The significant improvement in the analysis from the QC, as well as the smaller RMS value in the analysis increment over the background innovation showed the system was using the GPS RO data, and the QC is working.

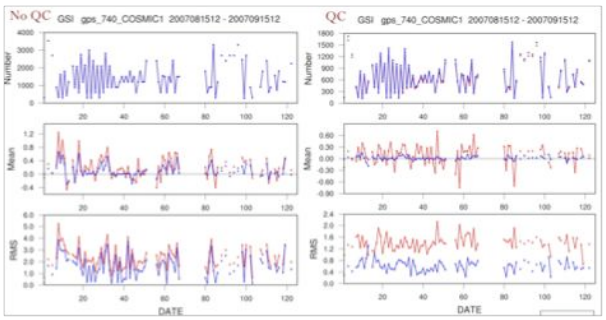


Figure 4: Time series of number of GPS RO observations, mean and standard deviation of the analysis increment (red) and background innovation (blue) from experiment CYC_GPS.

Figure 5 shows the forecast verification against prepBUFR observations for CYC_GPS and CYC_CONV_default. A slight positive SS impact can be seen in the GPS RO over the conventional assimilation alone. These impacts are especially noticeable in the upper troposphere and lower stratosphere in the wind field.

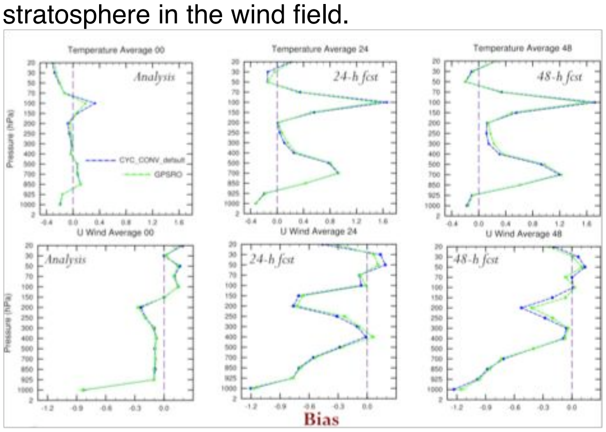


Figure 5: Vertical profiles for the analysis (left), 24-hour forecast (middle), and 48-hour forecast (right) of temperature (upper) and u-component wind (lower) from CYC_GPS (green) and CYC_CONV_default (blue).

3.3 AMSU-A AND AMSU-B RADIANCE IMPACT

For the CYC_AMSUA and CYC_AMSUB runs, a variational bias correction (BC) within GSI was used for these runs. Figure 6 has three time series showing the number of brightness temperature observations, mean, and standard deviation of the brightness temperature background with and without BC. Significant reduction in bias and standard deviation after the GSI bias correction can be seen at this channel. Even with BC, the time series of Figure 6 present diurnal oscillations, which is due to the diurnal changes in the BC coefficient not being considered. Such effects were detected and confirmed in the experiments discussed later.

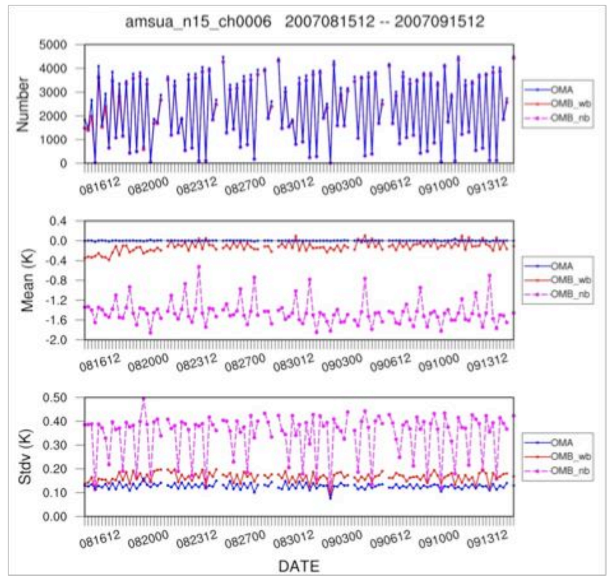


Figure 6: Time series of the number of brightness temperature observations, mean, and standard deviation of brightness temperature background with BC (red), without (pink) BC, and analysis increment (blue) for the CYC_AMSUA experiment

Forecast verification was performed against prepBUFR sounding observations for temperature, wind, and specific humidity, shown in Figure 7. A slightly SS positive impact can be seen in the humidity and wind forecasts. Adding AMSU-B radiance data in addition to AMSU-A shows no SS improvement or degradation over conventional and AMSU-A radiance data alone in this case. A slight negative SS impact can be seen in Figure 7, particularly in the u-component wind from 50-10mb. These differences may show a slightly unfair negative impact due to inclusion of the radiance data from channels with noticeable weighting function values beyond 10 mb model top in the assimilation. The vertical weighting function describes the relative contribution that microwave radiation emitted by a layer in the atmosphere makes to the total intensity measured above the atmosphere by the satellite.

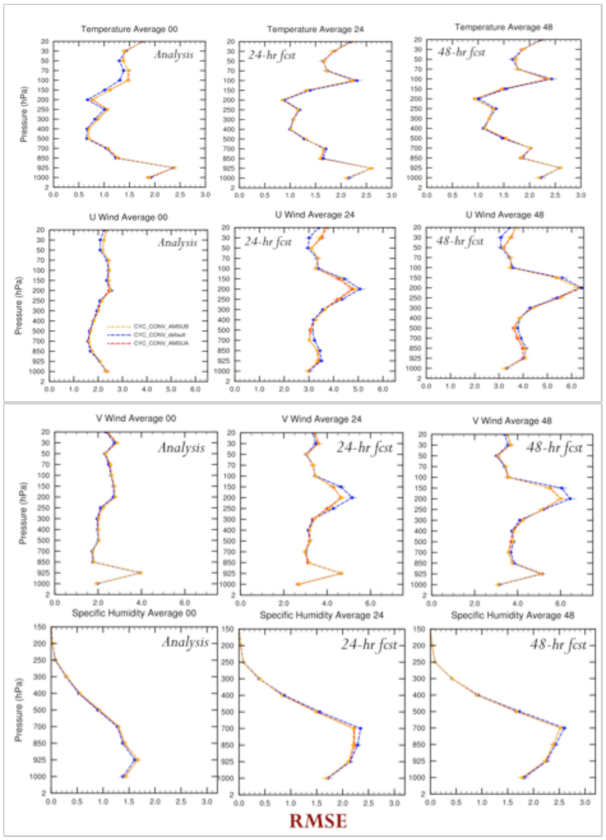


Figure 7: Vertical profiles for the analysis (left), 24-hour forecast (middle), and 48-hour forecast (right) of temperature (upper), u-component wind (upper-middle), v-component wind (lower-middle), and specific humidity (lower) from CYC_CONV_default (blue), CYC_ AMSUA (red) and CYC_AMSUB (orange). Verification is against prepBUFR sounding data and 95% confidence intervals.

A number of weekly runs were performed with channels above 10 mb removed, including diurnal cycles in the variational bias correction (VarBC) (AMSUA_satbias_DC) and adding angle-dependent VarBC in addition to air-mass VarBC (AMSUA_angle) in the previous runs. Figure 8 shows MET verification against prepBUFR data suggests that removing channels above the current model top and including diurnal cycles in VarBC gives smaller bias and RMSE for the temperature and specific humidity at analysis time, whereas adding angle-dependent VarBC presents no further improvement. This positive impact on temperature and humidity field is not present in the wind field and does not carry over to the forecast times. Extended runs with tuned satellite bias coefficients might be useful in further evaluating the impacts.

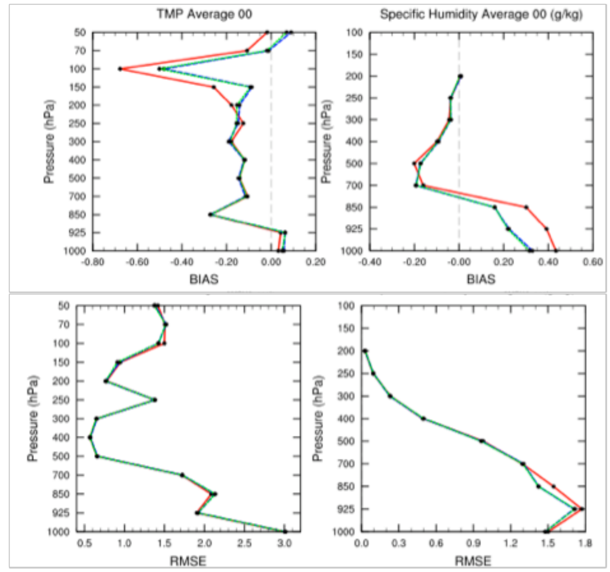


Figure 8: Vertical profiles at the analysis time of temperature (right) and specific humidity (left) for AMSUA_angle (green), AMSUA_satbias_DC (blue), and CYC_AMSUA (red). Verification is against prepBUFR observations. Results from 1-week (2007081512-2007082212) testing period.

4. SUMMARY AND CONCLUSIONS

A series of monthly experiments were run using GSI Q1FY11 coupled with WRF-ARW to investigate the capability and performance of the system. This testing includes observation impacts of surface, GPS RO, and AMSU-A and AMSU-B radiance data.

The surface impact study showed positive impacts from surface data assimilation on the near surface analysis and upper air at the forecast times. Simply adding more surface observations will not increase forecast skill, however, removing all surface observations will provide a SS negative impact compared to using the default number of surface observations following the GFS QC marks.

Assimilating GPS RO data showed that GSI v2.0 changes implemented did increase the data usage in the tropics. The forecast verification of the GPS RO run against prepBUFR observations showed slight SS positive impact over assimilating conventional observations alone. This is particularly true in the upper troposphere and lower stratosphere.

Radiance assimilation showed successful BC using the GSI variational BC, however the forecast verification showed overall neutral impact over conventional data assimilation alone. A slight SS positive impact from radiance assimilation can be seen in the specific humidity RMSE at lower-levels and in the v-component winds RMSE at the upper-levels. A slight negative SS impact for the radiance assimilation could be seen at 50-10mb levels, due to inclusion of channels with noticeable weighting function values above 10 mb in the data assimilation. Removing the diurnal cycle of the observation bias also helped to improve the analysis with radiance data assimilated. Adding AMSU-B radiance data did not result in any positive or negative SS differences over assimilating conventional data and AMSU-A radiance data alone.

ACKNOWLEDGEMENTS

The Air Force Weather Agency (AFWA) funds this research through the DTC. The National Center for Atmospheric Research (NCAR) is funded by the National Science Foundation (NSF).

REFERENCES

Cucurull, L., (2010): Improvement in the Use of an Operational Constellation of GPS Radio Occultation Receivers in Weather Forecasting, *Weather and Forecasting*. **25**,pp. 749-767.