

Plan for Test of Grell-Freitas Convective Parameterization using Cycled DA Initialization

Global Model Test Bed (GMTB)

Contacts: Ligia Bernardet (ligia.Bernardet@noaa.gov) and Josh Hacker (hacker@ucar.edu)

Contributing GMTB staff: Grant Firl, Michelle Harrold, Judy Henderson, Hongli Jiang, and Man Zhang

1st Draft March 2017; Updated April 2017 with feedback from EMC and Physics Team

[Introduction](#)

[Goals](#)

[Experiment design](#)

[Source Codes](#)

[Single Column Model \(Grant\)](#)

[Global workflow and automation](#)

[NEMS and GSM \(including physics\)](#)

[Initial conditions](#)

[Single Column Model](#)

[Global model](#)

[Forecast periods and length](#)

[Single Column Model](#)

[Global model](#)

[Post-processing, graphics, and diagnostics](#)

[Forecast verification](#)

[Data archival](#)

[Computational resources](#)

[Deliverables](#)

[Timeline](#)

[Risks and mitigation](#)

[References](#)

[Appendix A. List of acronyms](#)

Introduction

With the goal of supporting the development of an advanced physics suite for the GFS, in early 2017 the GMTB conducted an assessment of the GFS using the scale-aware Grell-Freitas (GF; Grell and Freitas 2014) deep and shallow convective parameterization (GFS-GF) and compared the results against a control run using the GFS operational convective parameterization, composed of the Simplified Arakawa Schubert (SAS) deep convective scheme and the mass flux shallow convection scheme (collectively referred to as GFS-SAS). These results were described in detail in a [website](#) and [report](#). As a first step in the evaluation of the GF scheme, global forecasts were run at a relatively low resolution (T574), in free-forecast mode (no data assimilation or cycling) and without tuning of the physics suite. Results of RMSE and bias comparisons varied by forecast lead time, level, and region. While upper-level wind speed RMSE was generally similar between the model configurations, temperature, relative humidity, and precipitation forecasts were more significantly impacted by the cumulus parameterization selected. Overall, GFS-SAS displayed superior forecasts in more instances than GFS-GF. The advantage of GFS-SAS over GFS-GF was greater and more frequent earlier in the forecast. As forecast lead time progressed, the gap in performance narrowed and GFS-GF was superior to GFS-SAS for some levels, lead times, and regions. This suggests that the GF scheme may not be in balance with the initial conditions used in that test (operational GFS analyses), and that the GF might perform better in a cycled experiment.

The next test will expand the scope of the previous test by adopting cycled Data Assimilation (DA) for both the control and the experimental cumulus parameterizations. In this test, the two model configurations will have their own independent initial conditions, both generated by running the NCEP GDAS. Running the cycled DA experiment may expose problems with the parameterization, leading to them being later addressed.

To produce relevant results for NCEP as it explores options for its future advanced physics suite, the control forecasts for this test will be updated to use the most recent version of the GFS, which is the code planned for the June 2017 operational implementation at NCEP, commonly referred to as GFS_Q3FY17. In preparation for the upcoming operational implementation, the GFS physics suite was updated by EMC to include the scale-aware SAS (SASAS). Therefore, this test will be comprised of a control with cycled SASAS and the mass flux shallow convection (jointly referred to as GFS-SASAS) and experiment with cycled GF (GFS-GF). Additionally, a baseline of GF runs cold started from the GFS-SASAS analyses (GFS-GFCOLD) will be produced for the assessment of the impact of using cycled DA.

Versions of GF and of the combined mass flux shallow convection and SASAS are already used operationally at NCEP: the former is employed in the RAP (Benjamin et al. 2016), and the latter in HWRF (Biswas et al. 2016). Both schemes implement the Arakawa et al. (2011) extension to the original SAS scheme, which renders it scale-aware. The GF scheme incorporates an ensemble approach to the representation of convection, which can improve the

forecast by using a collection of parameters and algorithms to represent the convective triggers, vertical mass flux, and closures. The ensembles can optionally be perturbed by stochastic fields for deterministic forecasting as well as ensemble data assimilation. Flux-form tracer transport, wet scavenging, and aerosol awareness are also options in this scheme.

The next sections in this document summarize the goals for the test, the experiment design (including details of the source codes, initial conditions, forecast periods, forecast configuration, post-processing, graphics, diagnostics, verification, and archival), computational resources, timeline, and deliverables. A list of references and definition of all acronyms is also included.

Goals

The goals for this test are to:

- Test the ability of the GFS-SASAS and GFS-GF to represent deep convection in the tropics using a SCM forced by data from an ARM field campaign.
- Conduct preliminary evaluation of the GF parameterization as a potential replacement for the SASAS and mass flux shallow convection schemes in GFS.
- Assess the impact of cycled DA when introducing a new parameterization in the GFS suite.

Experiment design

The experiment will consist of a control (GFS-SASAS) and an experimental configuration (GFS-GF), both run in a SCM and cycled global configuration. A second experiment configuration will consist of non-cycled GF runs as a baseline to help understand the differences between cycled and non-cycled GF results.

The primary resolution for the global test will be T574, chosen to provide sufficient information while fitting in GMTB's limited computational resources. The global configuration will consist of forecasts initialized from cycled GDAS analyses. The main components of the global workflow (rocoto-based `gfs_workflow.v3.0.0`) are:

- Input task: Utility for gathering input files, observational files, and verification data according to the user defined configuration file.
- GDAS: System used to create a first-guess for the GFS data assimilation. Employs a T254 80-member EnKF and a deterministic 9-h forecast with six hourly cycling (06, 12, 18, 00 UTC).
- GSI: Data assimilation for the GFS.
- GSM: NEMS-based, atmosphere-only, forecast application at T574.
- UPP: NCEP Unified Post-Processor.
- Tropical Cyclone tracker: utility for identifying tropical cyclogenesis and tracking TCs.

- Graphics: GMTB Python-based graphics suite.
- MET: Tool for model evaluation.
- Archival.

Source Codes

The source code for this experiment will be based on the GFS code undergoing final tests for the June 2017 (Q3FY17) operational implementation at NCEP, with a modification to add the GF parameterization. The scripts and automation system will be based on the Rocoto-based `gfs_workflow_v3.0.0` from EMC, with additional verification and diagnostics tasks added by GMTB. The provenance of scripts and source codes is described in detail below. All revision numbers will be recorded when the test starts.

Single Column Model

The code for running the SCM portion of the test resides in NOAA's Virtual Laboratory (VLab) under the "gmtb-scm" project name (further information can be found [here](#)). Within this Git project, the specific code for running this test can be found in the "gf_da_test" branch. A tagged version will be created prior to running the test. This repository contains both the GMTB SCM infrastructure code and the branch of GSM containing the physics code planned for the global portion of the test. The GMTB SCM code interfaces with the GFS physics through the version of `nuopc_physics.f90` found in the specified GSM branch.

Global workflow and automation

Automation of tasks for this test will be done using the Rocoto Workflow Management System, as currently employed in parallel tests at EMC. The xml file used to describe the tasks and their interdependencies will be based on v3.0.0 of the Python-based xml generator developed by EMC. An initial xml file will be created using the following command:

```
/scratch4/BMC/gmtb/svn/gfs/gfs_workflow.v3.0.0/para/exp/rocoto/gfs_workflow_generator.py -c user.conf -w gfs-workflow.xml -t full
```

The xml created by this command contains a variety of tasks, including setting up environment variables, creating the initial conditions by running GDAS, running the forecast model, post-processing, tracking tropical cyclones, and detecting tropical cyclogenesis. The various tasks (each of which evokes a number of executables) are submitted to the batch system incrementally as dependencies are met within the workflow. This preliminary xml will be augmented by GMTB with additional tasks to stage datasets, create forecast graphics, run forecast verification, archive results, and purge the disk. The modified xml will be kept under version control using the Git server in NOAA's VLab.

NEMS and GSM (including physics)

All runs will be performed using the NEMS-based GSM model employing the Physics Driver v3 by setting `use_nuopc=true`. The GFS-SASAS and GFS-GF runs will be conducted using code from <https://svnemc.ncep.noaa.gov/projects/gsm/branches/laurie/gsm.GFmerge>. This branch was created by GMTB on March 15, 2017 from the GSM top of trunk, revision 89613, and contains identical code to the trunk and to the GSM undergoing final testing for use the operational implementation of the GFS planned for June 2017 (https://svnemc.ncep.noaa.gov/projects/gsm/tags/gsm_q3fy2017_kappa), with the exception that the GF code provided by the developer has been added and integrated. Files `module_cu_gf_driver.f`, `module_cu_gf_deep.f`, and `module_cu_gf_sh.f` were added and changes were made to three files in the GSM code: `phys/gbphys.f`, `phys/comprns_physics.f90`, and `phys/gloopr.f90` to accommodate the requirements of the GF scheme, and to create the ability to select between the SASAS and GF in the runs.

In addition to GSM, the NEMSGSM application requires two other components, NEMS and Chem. As described in Fig. 1, for this test, the tags for the Q3FY17 tags will be employed: https://svnemc.ncep.noaa.gov/projects/nems/tags/nems_q3fy2017_kappa and https://svnemc.ncep.noaa.gov/projects/aerosol/chem/tags/chem_q3fy2017_kappa.

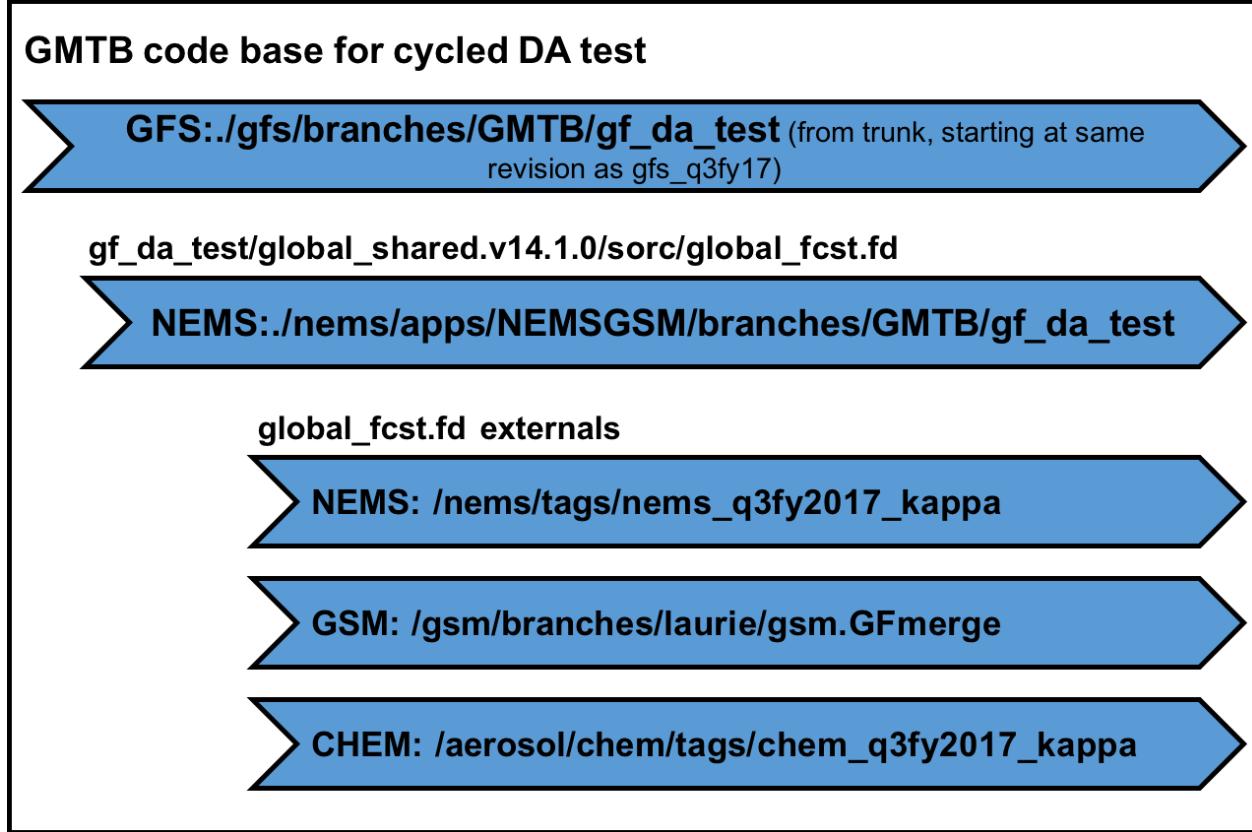


Figure 1. Schematic of GMTB code base for cycled DA test.

Initial conditions

Single Column Model

The SCM will be configured to run the GCSS Working Group 4's sixth intercomparison case based on ARM's TWP-ICE field campaign as described in Davies et al. (2013). The case is based on a suite of observations obtained near Darwin, Australia in January and February of 2006. Meteorological conditions observed included deep convection associated with an active phase of the monsoon and suppressed convection and clear sky associated with the inactive phase. The initial profiles of temperature, moisture, and horizontal winds reflect average conditions over the study area (centered on 12.425°S, 130.891°E) at 3 UTC on January 19, 2006. The surface is oceanic with a fixed SST, implying interactive surface fluxes calculated by a surface layer scheme. An observed ozone profile is included for use with interactive radiation, and large-scale horizontal advective tendencies for temperature and moisture as well as mean vertical motion are included from variational analysis performed on the observational data. Horizontal wind profiles are relaxed to observed profiles on a timescale of two hours.

Global model

Initial conditions for the global model will be generated by running GDAS, which employs a 80-member T254 EnKF on a six-hourly cycle. GDAS will be run separately for each model configuration (GFS-SASAS and GFS-GF), so that initial conditions are consistent with the physics being tested. GFS forecasts will be launched at 00 UTC only.

Forecast periods and length

Single Column Model

Forcing for the SCM is supplied for the entire length of the TWP-ICE field campaign from 03 UTC on January 17, 2006 to 21 UTC on February 12, 2006. The case coordinators have supplied a “best estimate” forcing dataset that is derived using the constrained variational analysis technique and a 100-member forcing ensemble dataset derived using the same method that varies the forcing based on the uncertainty in the surface precipitation measurements. For this test, the SCM will be run for all ensemble forcing datasets along with the “best estimate” forcing dataset in order to evaluate how the applied forcing affects the results.

Global model

As an initial assessment, the test will cover 10 forecasts within summer (June, July, and August 2016). While cycled DA will run every 6h, forecasts will be launched once daily at 00 UTC and run out to ten days with output every six hours.

Post-processing, graphics, and diagnostics

The *unipost* program within NCEP's UPP will be used to output the necessary variables at specified levels, derive additional meteorological fields, and vertically interpolate fields to isobaric levels. The post-processed forecast files will include two- and three-dimensional fields, which are necessary for both the plotting routines and verification tools. The necessary parameter files for *unipost* will be based on what is currently being utilized at NCEP for parallel testing; however, minor modifications may be made to remove legacy variables as a means to reduce file sizes. Output from *unipost* will be in GRIB2 format, and the *wgrib2* utility will be used to interpolate the post-processed files to a 0.25° global grid (G193).

Graphics will include a suite of figures created by ingesting the 0.25° GRIB2 files and creating plots over two projections: 1) the global grid from the 0.25° GRIB2 files, and 2) G218, a 12-km Lambert Conformal grid.

The following variables will be plotted for each grid:

- 250-hPa wind speed
- 250-hPa temperature
- 500-hPa temperature
- 500-hPa vorticity
- 700-hPa temperature
- 700-hPa vertical Velocity
- 850-hPa height
- 850-hPa temperature
- 850-hPa relative Humidity
- 2-m temperature
- 2-m dewpoint temperature
- 6-h accumulated convective precipitation
- 6-h accumulated total precipitation

In addition to plotting model output, the GMTB will also include diagnostics such as area-averaged precipitation accumulation over specified regions (i.e., Amazon).

Forecast verification

Objective model verification statistics will be generated using the MET package. MET is capable of pairing forecast and verification datasets in multiple ways, such as:

- Grid-to-point: utilized to compare gridded surface and upper-air model data to point observations.
- Grid-to-grid: utilized to compare gridded surface and upper-air model data to gridded observations (e.g., QPE and radar reflectivity) or gridded model analyses.

For point-based verification, post-processed model output for select surface (Table 1) and upper-air (Table 2) variables will be compared to observations (METARs and RAOBs) using the MET point-stat tool. The 0.25° model output will be regridded to G218, a 12-km Lambert Conformal grid covering the CONUS (Figure 2) and evaluated using the NAM NDAS PrepBUFR files as the observational dataset for the surface verification. For upper-air verification, the 0.25° model output will be regridded to both the G218 and G3 (a global 1.0° latitude-longitude domain shown in Figure 3) and evaluated using the GDAS PrepBUFR files as the observational dataset. Bias (or Mean Error - ME), RMSE, and BCRMSE will be computed separately for each variable at the surface and upper-air levels. Verification statistics will be stratified by forecast lead time, vertical level, regional area, and season. For the surface variables, statistics will be aggregated over the CONUS domain along with 14 sub-regions (Figure 4). Upper-air statistics will be aggregated over the CONUS domain (for forecasts regridded to G218), along with global, Northern Hemisphere (NH; $20^{\circ} - 80^{\circ}$ N), Southern Hemisphere (SH; $20^{\circ} - 80^{\circ}$ S), and Tropics (20° S – 20° N) domains for G3.

Precipitation verification will be performed over CONUS and over the entire globe. For the CONUS domain, a grid-to-grid comparison will be made using the QPE from the CCPA dataset, which has a resolution of ~ 4.8 km. Both the CCPA QPE analyses and the 0.25° post-processed model output will be interpolated to G218 and compared over the CONUS domain and 14 sub-regions. For the global evaluation, CMORPH precipitation analyses (60° N- 60° S) will be used due to their high spatial (8 km at the equator, $\sim 0.07^{\circ}$) and temporal resolution. Both the CMORPH analyses and the 0.25° post-processed model output will be interpolated to G3 and compared over the NH ($20^{\circ} - 60^{\circ}$ N), SH ($20^{\circ} - 60^{\circ}$ S), and Tropics (20° S – 20° N). In addition, a verification region will be placed over the Amazon to better evaluate precipitation, including the diurnal signal, over a convectively active area. Precipitation verification will be conducted for both 6-h and 24-h accumulation period (valid from 12 UTC to 12 UTC) using the MET grid-stat tool. Traditional verification metrics computed for both CONUS and global regions will include the frequency bias (FBias) and the Equitable Threat Score (ETS).

Table 1. Description of the surface verification to be performed using the listed observation dataset for the specified variables, levels, metrics, and grids. Z2 and Z10 refer to 2- and 10-m AGL.

Variable	Level	Metrics	Observation dataset	Grid to verify	Aggregated verification region
TMP	Z2	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
RH	Z2	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
SPFH	Z2	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
HGT	Z0	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
UGRD	Z10	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
VGRD	Z10	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
WIND	Z10	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions

PRMSL	Z0	ME, RMSE, BCRMSE	NDAS (NAM if NDAS not available)	G218	CONUS and 14 sub-regions
--------------	----	------------------------	--	------	-----------------------------

Table 2. Description of the upper-air verification to be performed using the listed observation dataset for the specified variables, levels, metrics, and grids.

Variable	Level (hPa)	Metrics	Observation dataset	Grid to verify	Aggregated verification region
TMP	10, 20, 50, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics
RH	300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics
SPFH	300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics
HGT	10, 20, 50, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics
UGRD	10, 20, 50, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics

VGRD	10, 20, 50, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics
Wind	10, 20, 50, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925, 1000	ME, RMSE, BCRMSE	GDAS (GFS if GDAS not available)	G218	CONUS
				G3	Global, NH, SH, Tropics

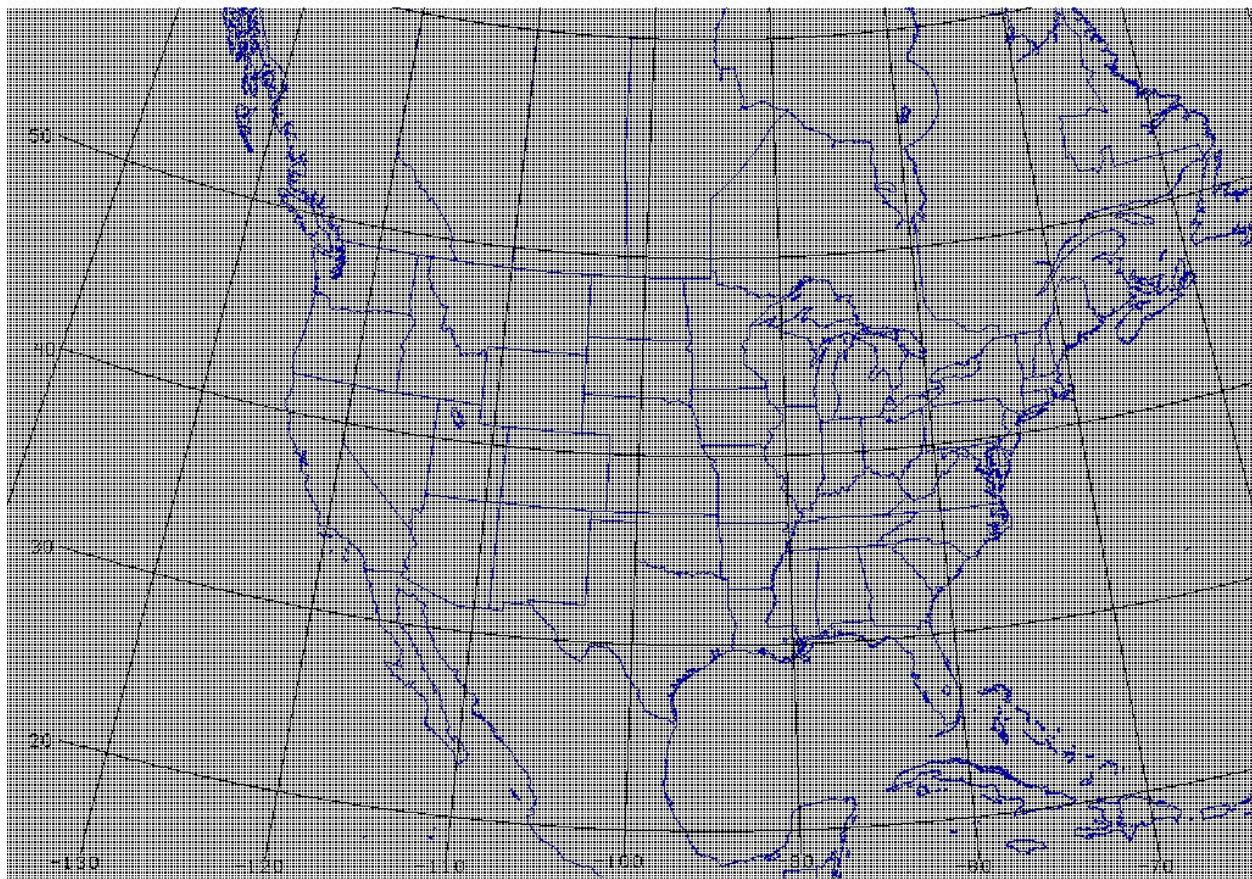


Figure 2. Map showing the NCEP ~12-km Lambert Conformal CONUS domain (G218).

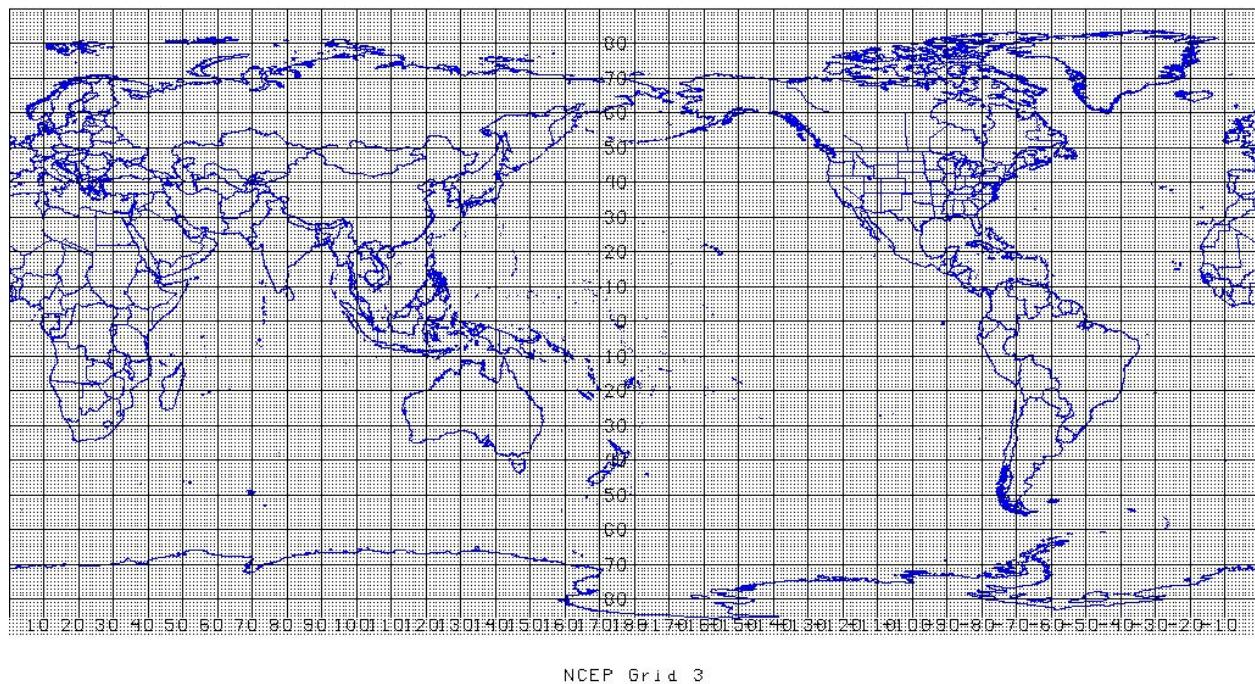


Figure 3. Map showing the NCEP 1.0° global latitude-longitude domain (G3).

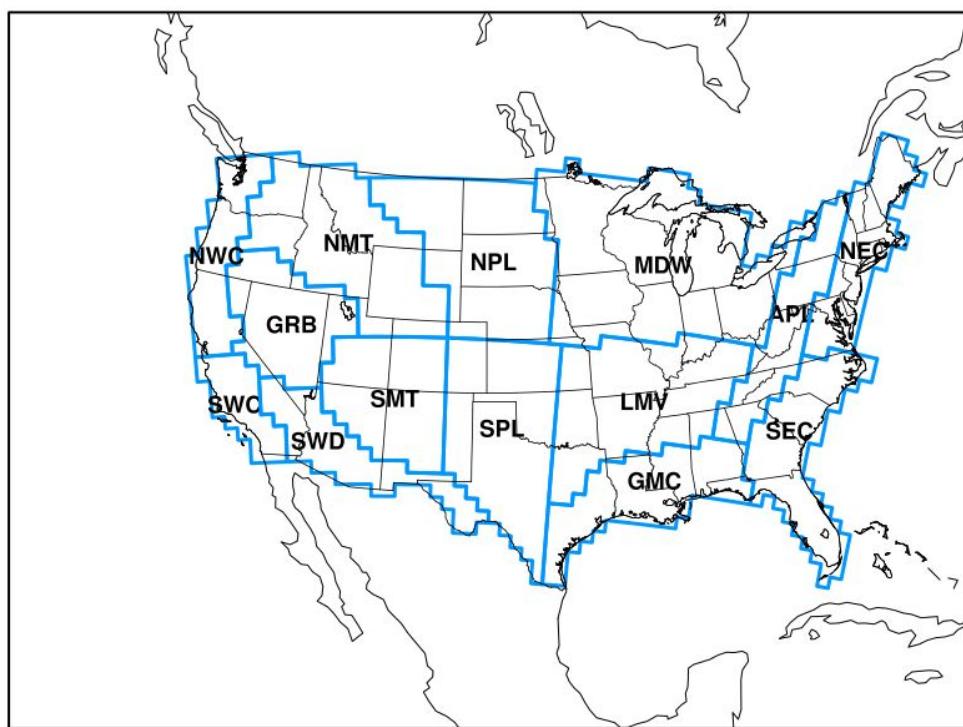


Figure 4. Map showing the CONUS (outer boundary of blue line) and 14 NCEP subregion verification domains.

Table 3. Description of the accumulated precipitation verification to be performed using the listed observation dataset for the specified temporal intervals, metrics, and grids.

Variable	Accumulation interval (h)	Metrics	Observation dataset	Grid to verify	Aggregated verification region
APCP_06	6	FBias, ETS	CCPA	G218	CONUS and 14 subregions
APCP_24	24	FBias, ETS	CCPA	G218	CONUS and 14 subregions
			CMORPH	G3	NH, SH, Tropics

In order to gain further insight into model biases, the MET package will also be used to compare observations and forecasts contained in the diagnostic files output by GSI. These binary files contain the innovations associated with each assimilated observation, that is, the Observation - Background (O-B) value. The above processing will be done to the diagnostic files produced during the GSI run within the GDAS system, as they have a later cut off time for observations compared to the diagnostic file produced by the GFS system, and will therefore be more valuable.

Anomaly correlation is a measure of the ability of an NWP model to forecast synoptic-scale weather patterns (e.g., high pressure ridges and low pressure troughs), as well as the location of frontal and storm systems. Since it is a well-accepted verification metric used among operational centers and the research community, it will be included in the evaluation. To compute the AC, the mean climatology will be removed from the forecast and observations so that the strength of the linear association between the forecast and observed anomalies can be evaluated. The climatology files that will be used for this test are the same 1.0° GRIB1 files that are currently being used by NCEP. In order to pair the gridded forecast and analyses files with the climatology, the 0.25° post-processed global forecasts will be read into MET's grid-stat tool and then re-gridded to a 1.0° grid before performing the AC calculation.

Another component of the evaluation will be TC attributes (position, intensity, and structure) and genesis. Forecasts of existing storms, obtained with a vortex tracker, will be compared against the Best Track dataset using MET-TC, a module within the MET tools. Conversely, forecasts of TC genesis will be compared against the Best Track dataset using code adapted from Halperin et al. (2016). Since every forecast will be run for both configurations of the model, the presentation of the results will take advantage of the pairwise nature of the

test. With this methodology, differences between the verification statistics will be computed for the GSM T574 runs with SAS versus the runs with GF.

For surface and upper-air, both the individual and pairwise verification statistics will be accompanied by CIs computed from standard error estimates using a correction for autocorrelation. The CIs will be computed on the median values of the aggregated results for the surface and upper-air statistics using parametric tests. For the precipitation statistics, a bootstrapping method (using 1500 replicates) will be used. The CIs on the pairwise differences between statistics for two configurations will assist in determining whether the differences are statistically significant.

With a large amount of verification results being produced for this test, a “scorecard” is a straight-forward way to identify patterns in the difference of performance between two configurations, including level of significance, for specified metrics, variables, levels, regions, and times. Using a development version of the Developmental Testbed Center (DTC) METViewer application, this test will include scorecards for surface and upper-air verification over regions and forecast hours of interest.

Data archival

Input and output data files from multiple stages of the global workflow system will be archived to the NOAA HPSS. Archives contain files typically archived using EMC parallel procedures, as well as additional verification and graphics added by GMTB.

Computational resources

The T574 runs will be computed on the NOAA R&D platform Theia using project *gmtb*, which has an allocation of 101,917 core-hours/month and 8 TB of disk. Selected files will be archived in the NOAA HPSS, and results will be displayed in the DTC website (dtcenter.org).

The SCM runs do not require any HPC resources and can be executed on local machines, although the capability exists to replicate these runs on Theia.

Deliverables

The following deliverables will be produced in this test:

- Archives of forecasts in NOAA HPSS, accessible by NGGPS collaborators for further analysis.
- Website with test results.
- Verification statistics loaded in database and accessible through MET Viewer.
- Final report of evaluation results.

Timeline

The timeline and dependencies for the various tasks involved in this test are outlined in Table 4. It should be noted that there is high uncertainty in the time it will take to conduct the actual runs because the estimation of required computational resources is underway.

Table 4. Timeline and dependencies for this test. Numbers on the left column indicate number of weeks needed to complete each activity. Horizontal staggering of activities indicates dependencies among them.

Weeks	Test Plan	GF code	SCM	Workflow v3	Website and Report
Done	Draft	Place code in GSM branch	Obtain code from GSM branch	Obtain from EMC	
1	Collect feedback	Get updates from developer	Update SCM and test	Test	
2-3	Finalize	Finalize preliminary tests with updates	Finalize and run	Test, add MET, graphics and finalize	
4-7			Analyze	Run (estimate)	
8-9				Analyze	
10-12					Finalize

Risks and mitigation

Table 5 lists risks associated with this test, along with a strategy to mitigate them. It should be noted that the need to implement these mitigation strategies could lead to a longer time being needed to conduct the test.

Table 5. Risks and mitigation strategies.

Risk	Mitigation
Problems running gfs_workflow_v3.0.0	Consult with EMC colleagues to get it functioning.
Problems with developer code (software or scientific)	Send code back to developer to address issue. Rerun experiment.
Lack of computational resources on theia	Procure more resources or reduce scope of test through one or more of the following methods: shorter forecast length, less frequent output, fewer cases, lower resolution, less variables in post output

References

Arakawa, A., J. H. Jung, and C. Wu, 2011: Toward unification of the multiscale modeling of the atmosphere. *Atmos. Chem. Phys.*, **11**, 3731–3742., DOI: <http://dx.doi.org/10.5194/acp-11-3731-2011>.

Benjamin, S., S. S. Weygandt, J. M. Brown, M. Hu, C. R. Alexander, T. G. Smirnova, J. B. Olson, E. P. James, D. C. Dowell, G. A. Grell, H. Lin, S. E. Peckham, T. L. Smith, W. R. Moninger, and J. S. Kenyon, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669-1694. DOI: <http://dx.doi.org/10.1175/MWR-D-15-0242.1>

Biswas and co-authors. 2016: [HWRF 2016 Scientific Documentation](#)

Davies, L., C. Jakob, K. Cheung, A. D. Genio, A. Hill, T. Hume, R. J. Keane, T. Komori, V. E. Larson, Y. Lin, X. Liu, B. J., Nielsen, J. Petch, R. S. Plant, M. S. Singh, X. Shi, X. Song, W. Wang, M. A. Whitall, A. Wolf, S. Xie, and G. Zhang, 2013: A single-column model ensemble approach applied to the TWP-ICE experiment. *J. Geophys. Res.: Atmos.* **118**, 12, 6544–6563. DOI: <http://dx.doi.org/10.1002/jgrd.50450>.

Grell, G.A, and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **14**, 5233-

5250. DOI: <http://www.atmos-chem-phys.net/14/5233/2014/>
doi:10.5194/acp-14-5233-2014

Halperin, D. J., H. E. Fuelberg, R. E. Hart, and J. H. Cossuth, 2016: Verification of Tropical Cyclone genesis forecasts from global numerical models: Comparisons between the North Atlantic and Eastern North Pacific Basins. *Wea. Forecasting*, **31**, 947-955. DOI: <http://dx.doi.org/10.1175/WAF-D-15-0157.1>

Appendix A. List of acronyms

ARM: Atmospheric Radiation Measurement
BUFR: Binary Universal Form for Representation of Meteorological Data
CI: Confidence Interval
CCPA: Climatology-Calibrated Precipitation Analysis
CONUS: Contiguous United States
CMORPH: Climate Prediction Center Morphing Technique
DA: Data Assimilation
DTC: Developmental Testbed Center
EMC: Environmental Modeling Center
ETS: Equitable Threat Score
FBias: Frequency Bias
GEWEX: Global Energy and Water cycle EXperiment
GCSS: GEWEX Cloud System Study
GDAS: Global Data Assimilation System
GF: Grell-Freitas cumulus parameterization
GFS: Global Forecast System
GFS-GF: Global Forecast System run with GF
GFS-SASAS: Global Forecast System run with SASAS and mass flux shallow convection
GMTB: Global Model Test Bed
GRIB2: GRIdded Binary file format version2
GSM: Global Spectral Model
GSS: Gilbert Skill Score
HPSS: High Performance Storage System
HWRF: Hurricane Weather Research and Forecasting System
IPD: Interoperable Physics Driver
MET: Model Evaluation Tools
METAR: international standard code format for hourly surface observations
NAM: North American Mesoscale model
NCEP: National Centers for Environmental Prediction
NDAS: NAM Data Assimilation System
NEMS: NOAA Environmental Modeling System
NH: Northern Hemisphere (defined here as 20° – 80° N for upper air verification and 20° – 60° N for precipitation verification)
NGGPS: Next-Generation Global Prediction System
NHC: National Hurricane Center

PrepBUFR: Quality-controlled BUFR
RAP: Rapid Refresh Forecast System
RMSE: Root-Mean-Square Error
SAS: Simplified Arakawa-Schubert cumulus parameterization
SASAS: Scale-Aware SAS cumulus parameterization
SCM: Single Column Model
SH: Southern Hemisphere (defined here as 20° – 80° S for upper air verification and 20° – 60° S for precipitation verification)
SVN: Apache Subversion
TC: Tropical Cyclone
TROP: Tropics (defined here as 20° S – 20° N)
TWP-ICE: Tropical Warm Pool - International Cloud Experiment
UPP: Unified Post Processor
UTC: Coordinated Universal Time
VLab: Virtual Laboratory