Plan for Test of Grell-Freitas Convective Parameterization

Global Model Test Bed

POC: Ligia Bernardet (ligia.Bernardet@noaa.gov)
June 2016

Introduction

Convective parameterizations are used in NWP models to represent the effect of subgrid scale convective clouds on the scales of motion that can be resolved by the model. Most modern models have both deep and shallow convection parameterizations, to represent the precipitating and non-precipitating subgrid-scale convective clouds, respectively.

The NCEP operational GFS uses the mass-flux based SAS convective parameterization based on Arakawa and Schubert (1974) and simplified to consider only one cloud top at a

specified time and location and not the spectrum of cloud sizes, which considerably reduced computation time (Grell 1993; Pan and Wu 1995). The current implementation of SAS in the operational GFS has additional modifications made to alter the convective trigger function (Hong and Pan 1998 and Han and Pan 2011), to change the cloud top selection algorithm (http://www.emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf), to reduce the vertical momentum mixing (Han and Pan 2006), and to add a shallow convection component (Han and Pan 2011).

One of the primary assumptions used to derive the SAS convective parameterization is that of scale separation. The fraction of grid cell covered by cumulus clouds is assumed to be much smaller than one, and all the subsidence from the cumulus clouds is assumed to occur in the same grid cell as the ascending motion. While this assumption is valid for coarse scale models with grid spacing of 50 km or larger, it is questionable for the current GFS operational configuration (~13 km) and becomes invalid for higher resolution models.

The limitation of the SAS convective parameterization sparked an interest in exploring an alternate convective parameterization for the UCGS, the future NCEP global model being developed under the auspices of NGGPS. Several efforts are taking place in the community to address what is known as the gray zone problem, namely the representation of cloud and convective transports in models with grid spacings on the order of 1-10 km, in which eddies are partially resolved but still require some parameterization. One of the approaches to address this problem is the superparameterization discussed by Randall et al. (2003), in which a cloud-system resolving model is embedded within each grid cell of a global model. However, given that this approach is too costly for operational implementation and the current and near-term planned computational resources available to NCEP, other avenues are currently being considered.

One option is the G3 cumulus parameterization, a three-dimensional modification of the Grell and Devenyi (2002) scheme that spreads the subsidence due to convection in a grid cell onto neighboring grid cells. Yet another option is to apply an extension of the original SAS parameterization to eliminate the scale-separation assumption (Arakawa et al. 2011). Two implementations of this approach were tested for HWRF, the meso-SAS (Pan et al. 2014) and the scale-aware SAS, with the latter being accepted for the 2016 version of the operational HWRF at NCEP.

Another implementation is the GF parameterization (Grell and Freitas 2014), which is currently used operationally in the NCEP Rapid Refresh model (Benjamin et al. 2016) and was tested for use in MPAS (Fowler et al. 2016). Development and testing of the GF parameterization are the topic of an external award granted by the NWS R2O initiative, and interest in this scheme by EMC and the NGGPS Program Office has led to the selection of this parameterization for testing by GMTB. The experiments will be conducted in close collaboration with the main developer of the GF scheme, G. Grell of NOAA ESRL/GSD, who has agreed to consult with the GMTB team as needed to set up the test.

This exercise will constitute the first test to be conducted by GMTB, a new entity formed within the DTC in 2015 to support R2O for global NWP at NCEP. A primary focus of GMTB is to create information that can be used for an evidence-based decision making process at NCEP. One of the goals for this test is to demonstrate the hierarchical testing capability of GMTB (described here), which ranges from a SCM through an end-to-end automated workflow to run the GSM and post-process, verify, and conduct diagnostics on the forecasts. This test will compare control forecasts using SAS against an experimental configuration using GF, both run over identical cases over a warm and a cold season using current developmental versions of the NEMS and GSM codes, as well as of the EMC parallel suite execution scripts.

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:

- Install systems and tools to use for current and future evaluation of advanced physical parameterizations.
- Demonstrate the GMTB hierarchical global model testbed capability (SCM through global workflow).
- Test ability of GFS operational physics, using SAS and GF parameterizations, to represent a meteorological phenomenon of interest, as represented in a GEWEX case.
- Conduct preliminary evaluation of the GF parameterization as a potential replacement for the SAS parameterization in GFS.

Experiment design

The experiment will consist of a control configuration using the SAS parameterization and an experimental configuration using the GF scheme. Both will be run in a SCM and global configuration. The primary resolution for the global test will be T574, chosen to provide sufficient information while fitting in GMTB's limited computational resources. Selected case studies at higher resolutions will be considered if resources allow.

The global configuration will consist of free forecasts, that is, non-cycled forecasts initialized from GFS analyses. The main components of the global workflow are:

- Input task: script to gather GFS input files and verification data from HPSS (this utility will be leveraged from the HWRF Python-based scripts).
- *global_chgres*: utility to convert the T1534 GFS analyses to the necessary resolution for model initialization, in this case T574.
- GSM: NEMS-based, atmosphere-only, forecast application.
- UPP: NCEP Unified Post-Processor.
- Tropical Cyclone tracker: utility for identifying tropical cyclogenesis and tracking TCs.
- Graphics: GMTB Python-based graphics suite.
- Diagnostics: add more detail.
- MET: tool for model evaluation.
- Archival.

In addition to the runs to be performed by GMTB, the GFS operational output will be verified and used as a baseline in order to provide a reference. This comparison should only be interpreted subjectively, due to a variety of differences between the GFS operational and the GMTB runs, including resolution, version of the code, and computational platform.

Source Codes

The provenance of scripts and source codes is described in the subsections below. All revision numbers will be recorded when the test starts.

Single Column Model

The SCM source code can be divided into three "groups" by function and source, each with its own ongoing development.

- First is the SCM infrastructure code necessary for grid setup, I/O, time-stepping, and integration with the IPD. This group of source code is currently in "prototype" form and is housed in NOAA VLab under the "gmtb-scm" project. This code is undergoing active development and will be frozen and tagged prior to the test.

- Second is the IPD, which can be thought of as the code that functionally connects a physics suite with a host atmospheric model, in this case, the SCM infrastructure code. The IPD is under development as a collaborative effort between NOAA EMC and DTC. For the initial test, the IPD code will match that used in the global tests, namely frozen code obtained from the trunk of the NEMS repository as of June 1, 2016.

- The final group is the code for the individual physical parameterizations, including the operational GFS physical parameterizations, the GF code, *gbphys.f*., and the IPD. The operational GFS physical physical parameterization code will be taken from the top of the NEMS trunk as of June 1, 2016, the GF code is as described

in the NEMS/GSM section, and the *gbphys.f* code will match the modified version used in the global runs.

Scripts and automation for global workflow

For this test, two sets of scripts and automation will be used. The first set of scripts, referred to as workflow_v2, provided by Kate Howard of EMC (/scratch4/NCEPDEV/global/save/Kate.Howard/para_gfs/prnems/para_config_NEMS), will be used to run the NEMS-based GFS. This version of the workflow is compatible with the NEMS code after the structuring associated with the NUOPC merge on May 02, 2016 (r75030). These scripts are responsible for a variety of tasks, including setting up environment variables, running the forecast model, post-processing, tracking tropical cyclones, and detecting tropical cyclogenesis. Additionally, these scripts are used for automation purposes, as the various tasks are submitted to the batch system incrementally as dependencies are met. These scripts evoke a number of executables installed by EMC staff on Theia; GMTB will use these executables as-is, except for the NEMS-based GSM, which GMTB will build in order to exercise the change between SAS and GF schemes.

The second set of scripts, contributed by GMTB and automated through the Rocoto Workflow Management System, will be used to stage datasets, preprocess the initial conditions to match the resolution of the runs, create forecast graphics, run forecast verification, archive results, and purge the disk. These scripts are 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 IPD by setting *use_nuopc=true*. A top of trunk checkout on June 07, 2016 of the NEMS and GSM code repositories housed in the EMC SVN servers will be used as the initial code base. Branches named *gf_test* will be created in the NEMSLegacy, NEMS, and GSM EMC SVN code repositories to a) house the GF code provided by the developer (file *module_cu_gf.f90*), and b) keep track of any changes made by GMTB staff for the test. It is anticipated that changes will be needed in the GSM's *gbphys.f* file to control the use of SAS or GF in the runs.

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 the T1534 analyses created as part of the retrospective GFS runs using the GFS implemented operationally in May 2016. Those are located in the NOAA HPSS, and It should be noted that the files in those tarballs are in the parallel naming convention.

For the summer 2015 cases, the relevant dataset is the *prnems8s* located at: /5year/NCEPDEV/emc-global/emc.glopara/WCOSS/prnems8s.

For the winter 2015-2016 cases, the relevant dataset is *prnems8w* located at: /5year/NCEPDEV/emc-global/emc.glopara/WCOSS/prnems8w. Forecast periods and length

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. In addition to a "best estimate" forcing dataset for the time period, a 100-member ensemble dataset is also supplied. For the GF test, the SCM will be run for the 5th, 25th, 75th and 95th percentile forcing datasets along with the "best estimate" forcing dataset in order to evaluate how the applied forcing affects the results.

Global model

To allow the evaluation of statistical significance, the test will cover two three-months periods, one in summer (June, July, and August 2015) and one in winter (December 2015, and January and February 2016). Forecasts will be launched once a day at 00 UTC and run out to fifteen 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 either plotting the gridded data directly, or regridding it to various verification grids used by NCEP. Those grids currently include G104, a north polar stereographic view of the CONUS at approximately 90-km resolution, and G218, a Lambert Conformal grid used by the 12-km NAM model.

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

Community input and collaboration is an essential piece of the GMTB's hierarchical testing. For this test, several diagnostics used within Dr. Zhou Wang's research group will be included. Dr. Wang is a NGGPS-funded principal investigator working on physics-oriented diagnostic tools for model evaluation and improvement. Dr. Wang and her research group graciously shared code for creating several diagnostics described in Li et al. (2014), including zonally averaged diabatic heating rate and zonal and meridional winds. In addition, the Li et al. (2014) calculations for investigating average daily precipitation rate over several regions (e.g., tropics and eastern and western Pacific) will be included as well.

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 1) 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 2) 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 3). Upper-air statistics will be aggregated over the CONUS domain (for forecasts regridded to G218), along with global, Northern Hemisphere (NH; 20° – 80° N), Southern Hemisphere (SH; 20° – 80° 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, ~0.07°) 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° – 60° N), SH (20° – 60° S), and Tropics (20° S – 20° N) . Precipitation verification will be conducted for a 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 GSS, also known as the ETS.
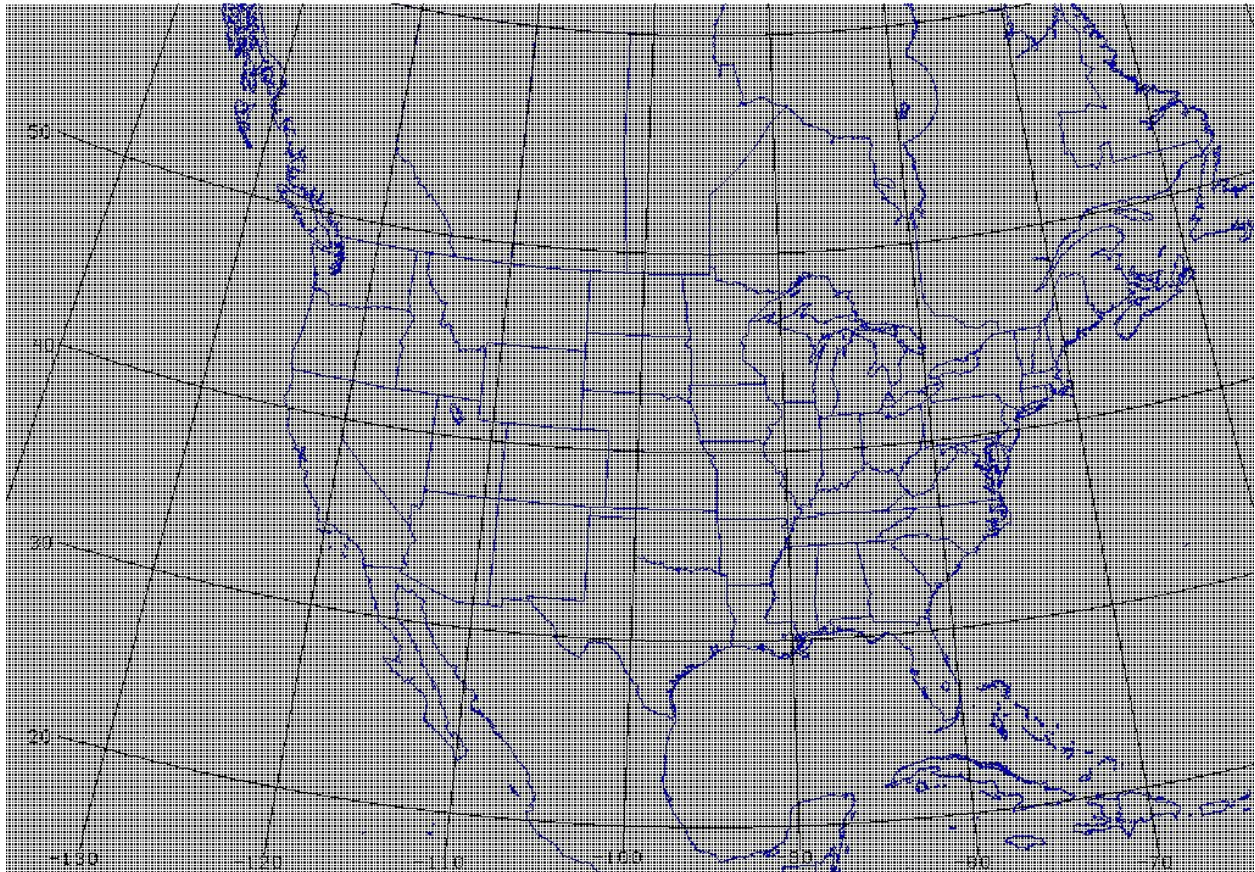
Table 1. Description of the surface verification to be performed using the listed observation dataset for the specified variables, levels, metrics, and grids. Z2 and Z10 refer to 2- and 10-m AGL.

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
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 |

NCEP Grid 218

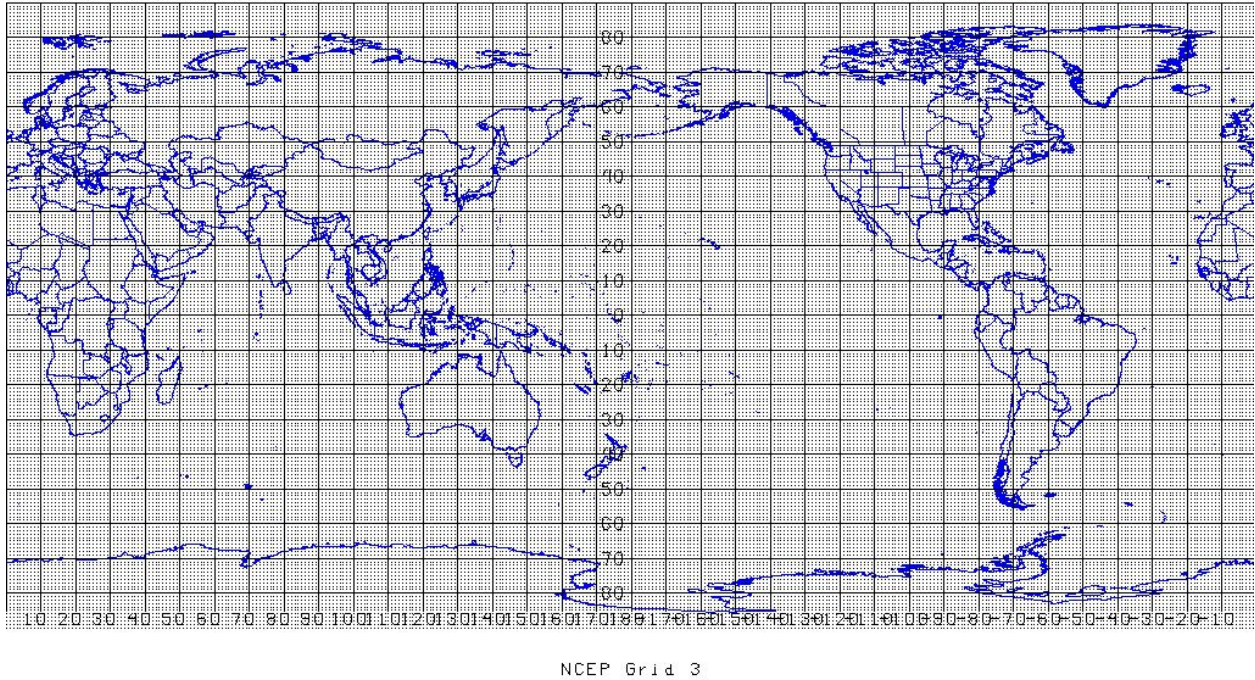Figure 1. Map showing the NCEP ~12-km Lambert Conformal CONUS domain (G218).

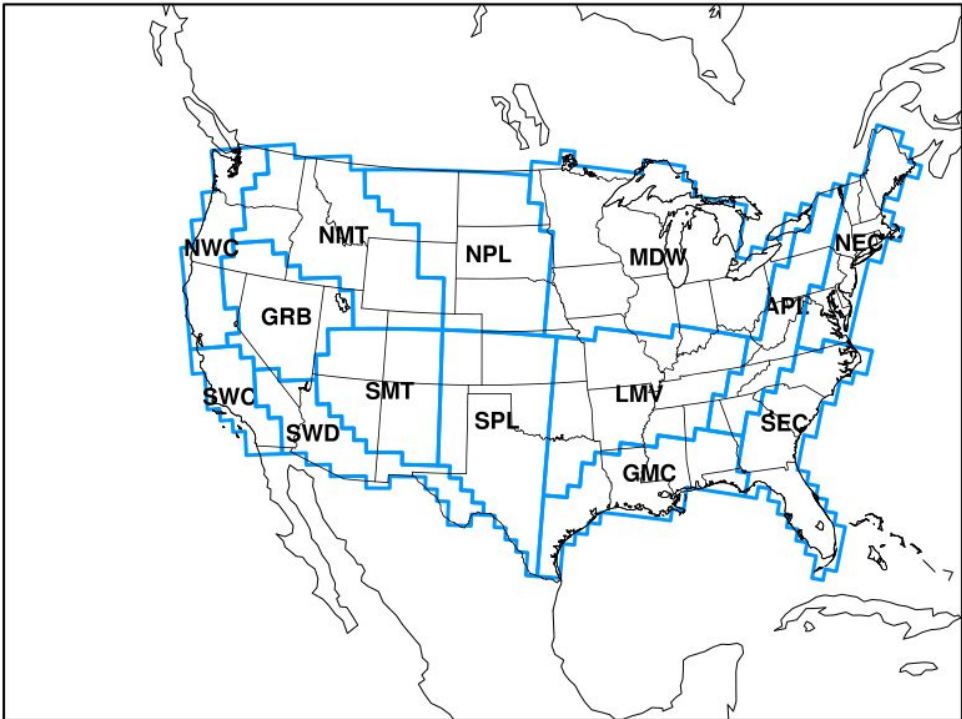Figure 2. Map showing the NCEP 1.0° global latitude-longitude domain (G3).

Figure 3. 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, GSS	CCPA	G218	CONUS and 14 subregions
APCP_24	24	FBias, GSS	CCPA	G218	CONUS and 14 subregions
			CMORPH	G3	NH, SH, Tropics

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 position, intensity, and structure verification. Forecasts obtained with a vortex and genesis tracker will be compared against the Best Track dataset using MET-TC, a module within the MET tools.

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.

Data archival

Input and output data files from multiple stages of the global workflow system will be archived to the NOAA HPSS. Archives will include:

- Input files (T574) that have been run through *global_chgres*.
- Configuration files and namelists specific to each forecast cycle.
- 0.25^o GRIB2 forecast files from *unipost* (analysis and forecasts at 6-hour increments).
- Graphics from Python plotting suite and diagnostic routines.
- Output from MET and MET-TC.

Computational resources

The SCM and T574 runs will be computed on the NOAA R&D platform Theia using project *gmtb*, which has an allocation of 100,000 core-hours/month and 3 TB of disk.

The workflow_v1 provided by EMC and tested by GMTB requires approximately 460 core-hours (for GSM) and 2500 core-hours (for UPP). GMTB is working with EMC to strategize how to reduce the footprint of the model runs, and with the Theia allocation officer to discuss options for additional allocation or windfall resources.

Selected files will be archived in the NOAA HPSS, and results will be displayed in the DTC website (dtcenter.org).

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 forecast images of 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.

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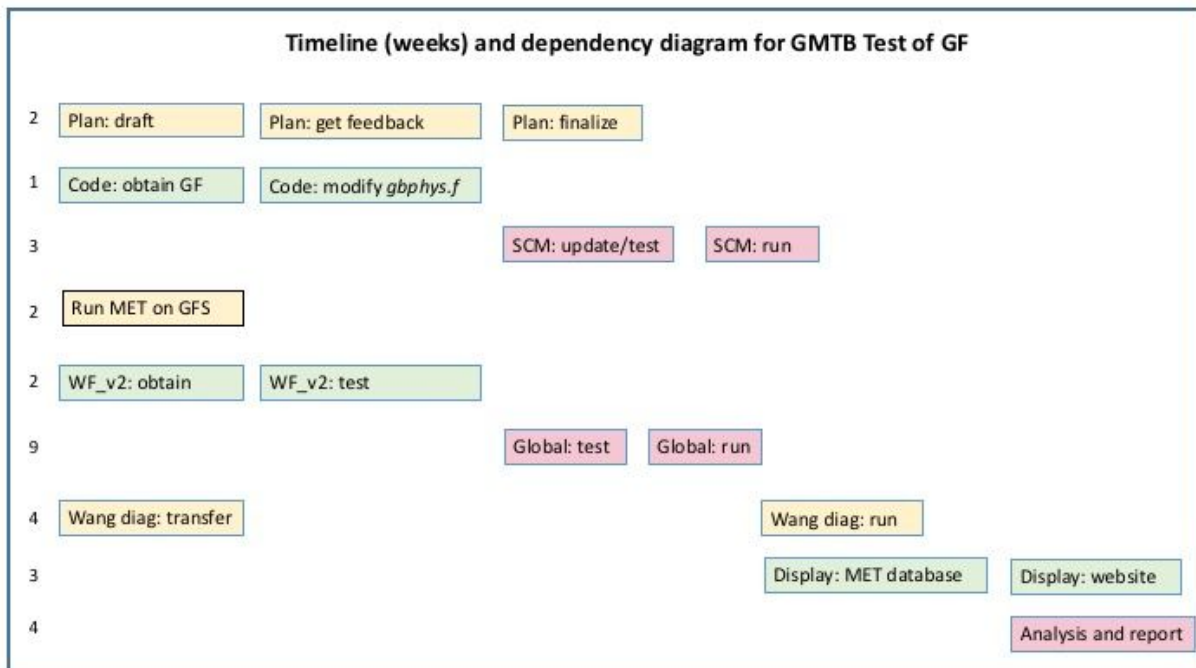
Risks and mitigation

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 workflow_v2	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	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.

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment. *J. Atmos. Sci.*, **31**, 674–701. DOI: *http://dx.doi.org/10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2*.

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*

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.

Fowler, L., W. C. Skamarock, G. A. Grell, S. R. Freitas, and M. G. Duda, 2016: Analyzing the Grell-Freitas convection scheme from hydrostatic to nonhydrostatic scales within a global model. *Mon. Wea., Rev*. early release. DOI: http://dx.doi.org/10.1175/MWR-D-15-0311.1

Grell, G.A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.*, **121**, 764-787. DOI: http://dx.doi.org/10.1175/1520-0493(1993)121<0764:PEOAUB>2.0.CO;2.

Grell, G. A. and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geoph. Res. Let.,* **29**, 38.1-38.4, http://dx.doi.org/doi:10.1029/2002GL015311.

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

Han, J. and H.-L. Pan, 2006: Sensitivity of hurricane intensity forecasts to convective momentum transport parameterization. *Mon. Wea. Rev.,* **134**, 664-674. *DOI: http://dx.doi.org/10.1175/MWR-D-15-0255.1*

Han, J. and H.-L. Pan, 2011: Revision of Convection and Vertical Diffusion Schemes in the NCEP Global Forecast System. *Wea. Forec.*, **26**, 520-533. *DOI:* http://dx.doi.org/10.1175/WAF-D-10-05038.1.

Hong, S.-Y. and H.-L. Pan, 1998: Convective trigger function for a mass flux cumulus parameterization scheme. *Mon. Wea. Rev.*, **126**, 2621-2639. DOI: http://dx.doi.org/10.1175/1520-0493(1998)126<2599:CTFFAM>2.0.CO;2.

Li, W., Z. Wang, M. Peng, and J. Ridout, 2014: Evaluation of Tropical Intraseasonal Variability and Moist Processes in the NOGAPS Analysis and Short-Term Forecasts. *Wea. Forec.*, **29**, 975-995. DOI: http://dx.doi.org/10.1175/WAF-D-14-00010.1.

Pan, H.-L., and W. Wu, 1995: Implementing a mass flux convective parameterization package for the nmc medium-range forecast model. *NMC Office Note*, **409**, available at http://www2.mmm.ucar.edu/wrf/users/phys_refs/CU_PHYS/Old_SAS.pdf.

Pan, H.-L., Q. Liu, J. Han, and R. Sun, 2014: Extending the simplified Arakawa-Schubert scheme for meso-scale model applications. *NCEP Office Note*, **479**, 10 pp, available at http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on479.pdf.

Randall, D. A., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterization deadlock. *Bull. Amer. Meteor. Soc.*, **84 (11)**, 1547-1564. DOI: http://dx.doi.org/10.1175/BAMS-84-11-1547.

Appendix A. List of acronyms

AC: Anomaly Correlation
AGL: Above Ground Level
ARM: Atmospheric Radiation Measurement
BCRMSE: Bias-Corrected Root Mean Square Error
CI: Confidence Interval
CCPA: Climatology-Calibrated Precipitation Analysis
CMORPH: Climate Prediction Center MORPHing technique
CONUS: Contiguous United States
DTC: Developmental Testbed Center
EMC: Environmental Modeling Center
ESRL: Earth System Research Laboratory
ETS: Equitable Threat Score
FBias: Frequency Bias
GCSS: GEWEX Cloud System Study
GDAS: Global Data Assimilation System
GEWEX: Global Energy and Water cycle EXchanges
GF: Grell-Freitas
GFS: Global Forecast System
GMTB: Global Model Test Bed
GRIB1: GRIdded Binary file format version1
GRIB2: GRIdded Binary file format version2
GSD: Global Systems Division
GSM: Global Spectral Model
GSS: Gilbert Skill Score
HPSS: High Performance Storage System
HWRF: Hurricane Weather Research and Forecasting System
IPD: Interoperable Physics Driver
ME: Mean Error
MET: Model Evaluation Tools
METAR: international standard code format for hourly surface observations
MPAS: Model for Prediction Across Scales
NAM: North American Mesoscale
NCEP: National Centers for Environmental Prediction
NDAS: NAM Data Assimilation System
NEMS: NOAA Environmental Modeling System
NGGPS: Next-Generation Global Prediction System
NH: Northern Hemisphere
NOAA: National Oceanic and Atmospheric Administration
NUOPC: National Unified Operational Prediction Capability
NWP: Numerical Weather Prediction
NWS:- National Weather Service
POC: Point Of Contact
QPE: Quantitative Precipitation Estimate
RAOB: RAwinsonde OBservation
R2O: Research-to-Operations

RMSE: Root Mean Square Error
SAS: Simplified Arakawa-Schubert cumulus parameterization
SCM: Single Column Model
SH: Southern Hemisphere
SVN: Apache Subversion
TB: Terabytes
TC: Tropical Cyclone
TWP-ICE: Tropical Warm Pool - International Cloud Experiment
UCGS: Unified Coupled Global System
UPP: Unified Post Processor
UTC - Coordinated Universal Time
VLab: Virtual Laboratory
WF_v2: Version 2 of the EMC workflow used to compute GFS free forecasts