Showcasing the hierarchical testing framework established by the Global Model Test Bed (GMTB)

Michelle Harrold*[1,4], Grant Firl[1,4], Man Zhang[2,3,4], Jamie Wolff[1,4], Judy Henderson[2,4], Ligia Bernardet[2,3,4], Joshua Hacker[%]

1 National Center for Atmospheric Research/Research Applications Laboratory • 2 NOAA/Earth System Research Laboratory/Global Systems Division

3 Cooperative Institute for Research in the Environmental Sciences • 4 Developmental Testbed Center

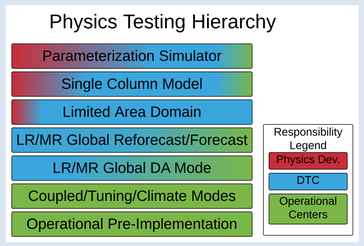
*harrold@ucar.edu

What is the GMTB?

The Global Model Test Bed (GMTB) team was established within the Developmental Testbed Center (DTC) to facilitate community involvement in the development of the Next Generation Global Prediction System (NGGPS) by supporting a hierarchical testing framework.

The GMTB is actively developing a uniform 'test harness' to enable in-depth investigation of various physical parameterizations and advanced physics suites. The goals of the hierarchical testing framework are to provide:

- A common infrastructure for testing physics developments that works across all temporal and spatial scales and facilitates an efficient R2O pipeline
- Simple-to-complex testing
- A framework for evidenced-based decision making
- Streamlined testing to accelerate transfer of worthy improvements into operations



Hierarchical Testing

- The hierarchical testing capability within the GMTB was used to support a testing and evaluation effort to compare the GFS's operational convective parameterization [Scale-Aware Simplified Arakawa Schubert (SASAS)] against an experimental configuration using a more advanced, scale-aware parameterization, the Grell–Freitas (GF) scheme
- Test plan was created jointly with EMC, NGGPS, Program Office, and the physics developer (G. Grell)
- Invokes concept of hierarchical testbed
  - SCM
  - Global retrospective runs
    - ✓ 20160601 – 20160615; 00 UTC initializations
    - ✓ Global diagnostics
    - ✓ Global verification
- All experiments were performed at a resolution of T574; cycled DA was employed for global runs

Global Workflow

Initialization Data → Pre-processing → Forecast → Post-processing → Graphics → Diagnostics → Verification

Workflow from EMC
- ✓ GMTB keeping pace with EMC
- ✓ GMTB & EMC collaboration

GMTB Workflow
- ✓ Highly flexible and configurable
- ✓ Python for graphics
- ✓ DTC's Model Evaluation Tools for verification

- **For full DTC testing of GF scheme:**
- ✓ *No tuning was performed*
- ✓ Deep and shallow convection turned on; mid-level scheme was not used
- ✓ Diurnal cycle adjustment was employed
- ✓ For closure option, average of all possible closures was used
- **Testing was first performed with non-cycled runs → progression to next tier of advanced testing included cycled DA**

Hierarchical Testing Results

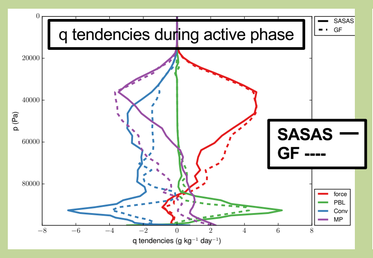
Single Column Model

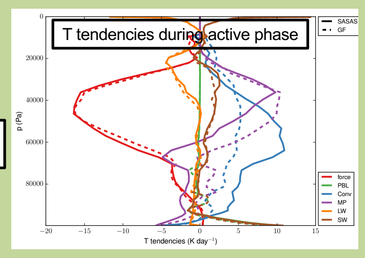
SCM developed by GMTB:
- Uses Interoperable Physics Driver (IPD) to connect to GFS operational physics
- Portable code has minimal dependencies
- Two cases currently available:
  - ✓ Maritime deep convective GCSS case based on TWP-ICE field campaign
  - ✓ Continental deep convective GCSS case based on ARM SGP Summer 1997 IOP
- Available to friendly users via NOAA's Vlab
- Features complete documentation and User's Guide:

  http://www.dtcenter.org/GMTB/gmtb_scm_doc/

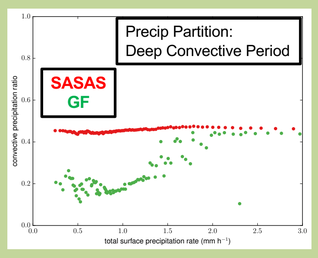
Case Setup
- GCSS case from TWP-ICE field campaign
- Deep and suppressed convection near Darwin, Australia during 1/20-2/12/2006
- Forced by fixed SSTs, prescribed horizontal advective tendencies of T, q, prescribed vertical velocity, and nudged u,v
- 100-member forcing ensemble that varies forcing based on uncertainty in precip measurement

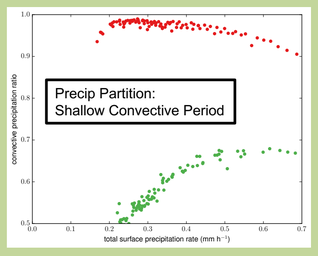
Physics Scheme	Control / GF
Surface	Noah (ocean surface)
Radiation	RRTMG
PBL	Hybrid EDMF
Microphyics	Zhao-Carr
Deep & Shallow Con.	**SASAS / GF**

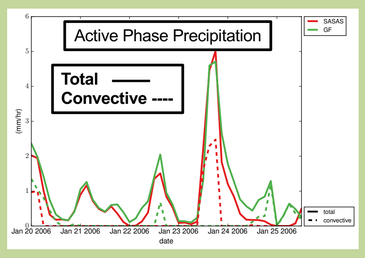




- Key SCM results:
  - GF produces weaker tendencies, leaving the microphysics to do more "work" to balance the forcing
  - The GF suite, for both the deep and shallow convective periods, produced much lower convective precipitation ratio
  - SASAS produces more convective precipitation



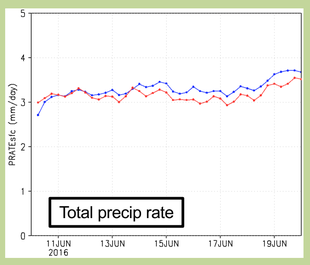




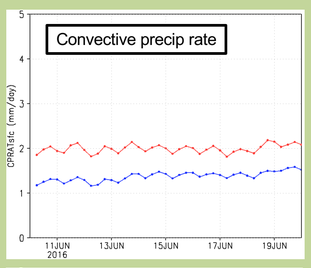
Global Runs – Diagnostics

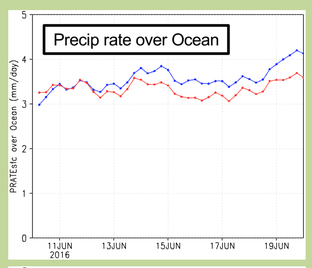
SASAS GF

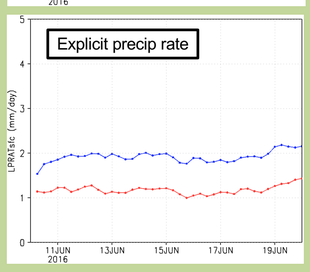
Diagnostics provided by GMTB:
- For a single case (20160610), several diagnostics were investigated to better understand the behavior of the two configurations:
  - ✓ Precipitation
    - Precipitation rate
    - Convective/explicit partition
  - ✓ Components of water budget
  - ✓ Cloud and radiation variables
    - Low, mid, and high cloud fraction
    - LW and SW and surface and TOA

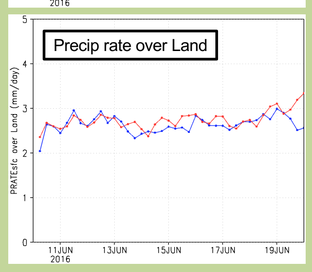


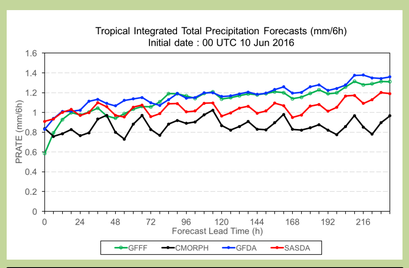
- Total precipitation partitions:
  - GF has more explicit and less convective precip than SASAS
  - GF produces more precip over ocean than SASAS; both produce similar amounts over land





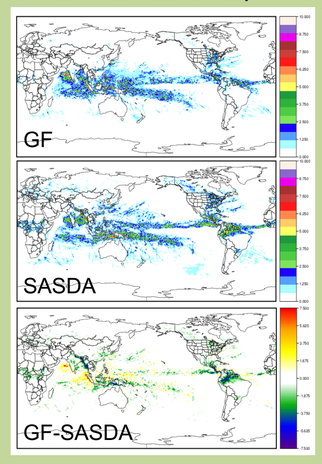






- Regional investigation over the Tropics:
  - Tropical region: 20° S – 20° N
  - Uses gridded CMORPH analyses (satellite microwave observations) as truth
  - Compared to CMORPH, both SASAS and GF overestimate precipitation, with GF overestimating at a higher level
  - Both SASAS and GF display an increasing trend in precipitation as the forecast progresses in time

6-h Accum. Convective Precip. – f240



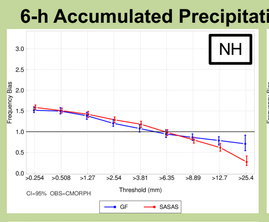
- Given the results in both the SCM and case study, the average convective precipitation over the retrospective period was calculated to visualize the differences between SASAS and GF
- Both GF and SASAS are appropriately capturing the ITCZ, including the SPCZ
- Large differences over northern part of South America, where GF has minimal convective precipitation
- With exception to areas off the west coasts of India and Indonesia, SASAS typically has more organized areas of stronger convective precipitation

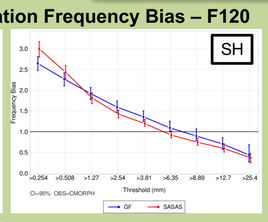
Global Runs – Verification

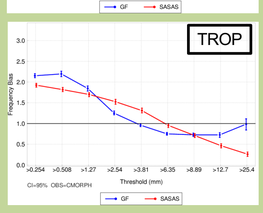
GF SASAS GF-SAS

Verification performed by GMTB:
- **Grid-to-grid**
  - ✓ Precipitation (6-h and daily accum.)
    - Frequency bias
    - Equitable threat score
  - ✓ 500 hPa height
    - Anomaly correlation
- **Grid-to-point**
  - ✓ Upper-air (T, RH, wind, height)
    - Bias, RMSE
  - ✓ Near-surface (T, RH, wind, PRMSL)
    - Bias, RMSE
- **Performed over:**
  - ✓ Global sub-domain (Grid 3, 1°×1°)
  - ✓ CONUS sub-domain (Grid 218, 12km)

6-h Accumulated Precipitation Frequency Bias – F120

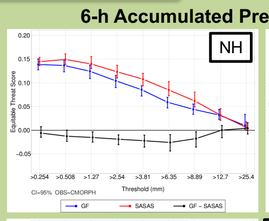


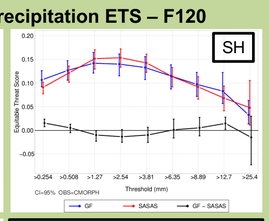


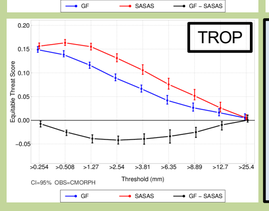


- In all regions, overall transition from high bias at lower thresholds to low bias at higher thresholds
- NH and SH: Minimal statistically significant (SS) differences
- Tropics: GF often SS better

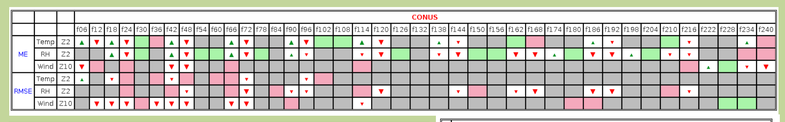
6-h Accumulated Precipitation ETS – F120



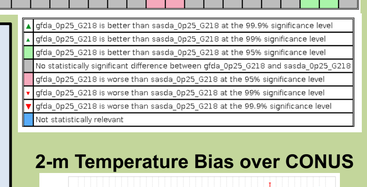




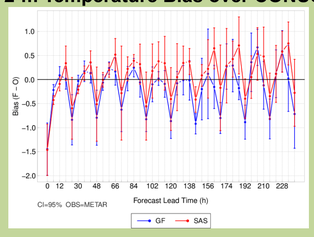
- In NH and Tropics, skill decreases as threshold increases
- In NH and Tropics, SS differences show SASAS as having higher skill
- In the SH, performance is similar – minimal differences



- "Scorecards" are an excellent tool to identify patterns in the difference of performance between two configurations, including level of significance, for metrics, variables, levels, regions, and times
- Scorecard of near-surface temperature, relative humidity (RH), and wind:
  - Clear diurnal signal in statistical results, with differences often at times valid at 18 and 00 UTC
  - SASAS typically better for RH and wind RMSE
  - GF frequently outperforms SASAS when considering temperature and RH bias
- Plot on right shows 2-m temperature bias over the CONUS; corresponds to the first line of scorecard
  - Note the diurnal signal and relationship to performance



2-m Temperature Bias over CONUS



Conclusion & Next Steps

- GMTB successfully developed an SCM that has been a valuable tool for physics developers within the hierarchical testing framework
- GMTB successfully implemented a framework for performing coarse resolution global forecasts – including pre-processing, forecasts, post-processing, graphics, and verification
- Using the SCM and global workflow, the GMTB undertook a testing and evaluation effort to compare 2 schemes: scale-aware simplified Arakawa Schubert and the *untuned* Grell-Freitas
- GMTB will continue to expand aspects of the hierarchical testing framework
- **Input and collaboration from the community is essential to the success of the testbed!**

DTC Visitor Program

- The DTC Visitor Program supports visitors to work with the DTC to test new forecasting and verification techniques, models and model components: ***http://www.dtcenter.org/visitors/***
- **Currently accepting applications!**
- Potential areas related to GMTB:
  - ✓ Additional cases in SCM
  - ✓ Additional model diagnostics
  - ✓ Physics developers modifying schemes for inclusion into Common Community Physics Package (CCPP)

Acknowledgments: Collaboration with Georg Grell was instrumental to the success of implementing the GF code in NEMS. Vijay Tallapragada and the EMC global team shared their workflow capabilities and knowledge of their verification procedures that were heavily leveraged in this effort. The DTC is funded by the National Oceanic and Atmospheric Administration, the Air Force, and the National Center for Atmospheric Research (NCAR). NCAR is sponsored by the National Science Foundation (NSF).

% Currently affiliated w/ Jupiter.