Hierarchical Testing and GMTB Physics Testbed

Grant Firl* and Michelle Harrold* + GMTB Team

*NCAR - RAL/JNT

Purpose of Physics Testbed

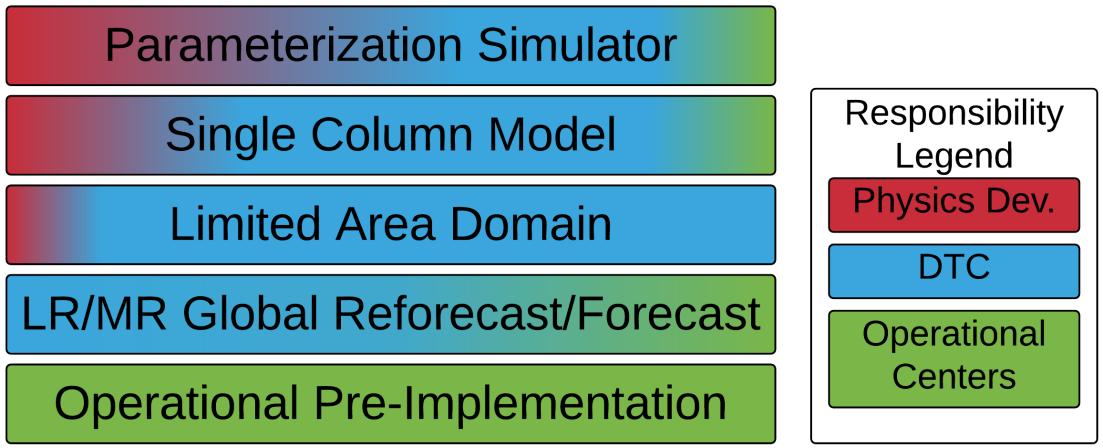
- To provide STANDARDIZED methods and infrastructure for
  1. testing
  2. evaluation
  3. diagnostics

  of CCPP-related physics schemes
- Augmentation of existing testing
- Guiding principles:
  - Objectivity/transparency
  - Replicability
  - Like-to-like comparisons

Potential User Groups

Hierarchical Testing

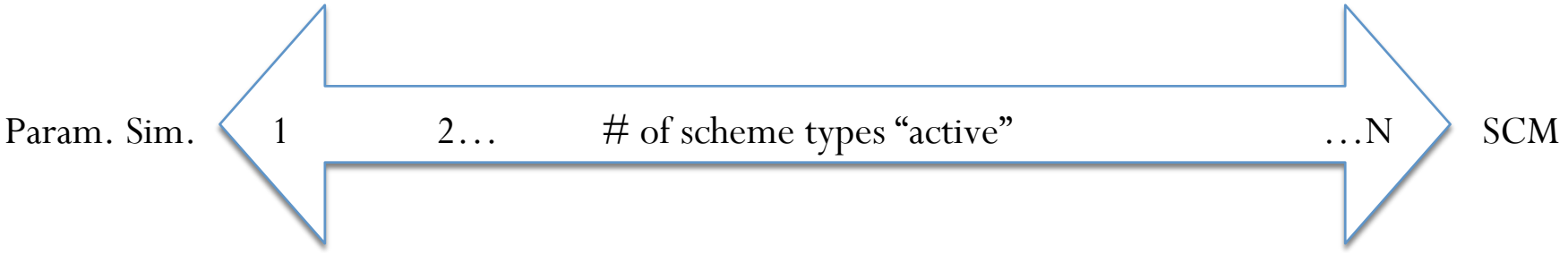
Physics Testing Hierarchy



- Simple-to-complex
- All-inclusive
  - Init./forcing data
  - Evaluative data
  - Opportunity for BYO-case
- Highly collaborative

Tier Focus: Parameterization Simulator

- Catch-all term that encompasses testing done outside of a physics "suite"
- Eliminates as much non-pertinent feedback as possible
  - analogous to unit testing
- Related to SCM
  - continuum between this tier and SCM based on number of scheme types that are active



Tier Focus: SCM

- Physics "suite" level testing
  - includes feedback among physics (either directly or through column state variables)
- Inexpensive, quick, approachable (yet limited)
  - De-facto standard for some types of schemes
  - GCSS/GASS/GABLS intercomparisons
- Can meet mechanistic goals and address science questions
- Basic suite-level tuning

Tier Focus: Limited-area Domain

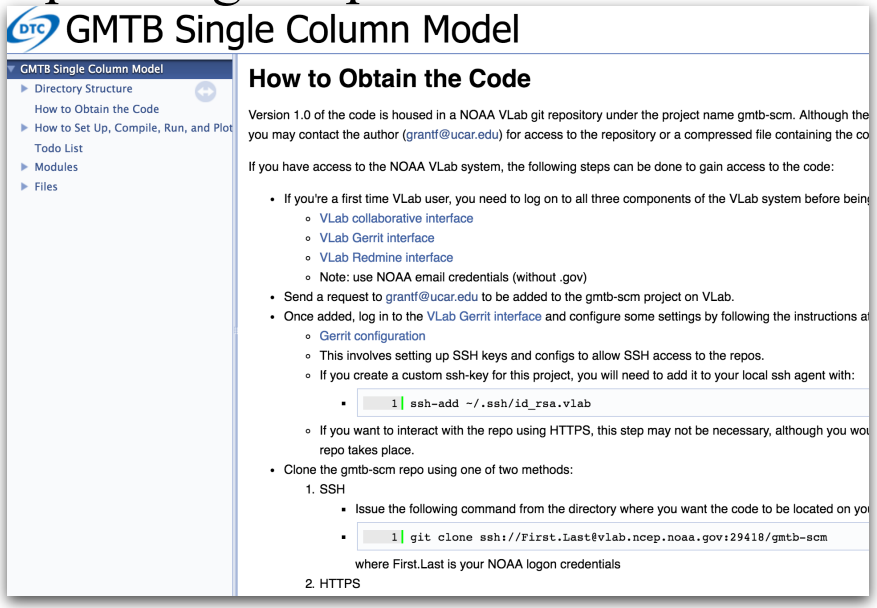
- Pros
  - fills a large spatial and computational gap
  - a large limited-area domain community exists
  - adds 3D interactivity w/o full cost of global run
  - LES configuration can complement SCM testing
- Cons
  - how would this be implemented with two dycores that are not configurable as limited-area models?

Tier Focus: Low-Med Res. Global

- Mechanical connection check
- Basic diagnostics & verification
- Testing and improving portability
- Facilitates easy replication of experiments
- Less expensive than full forecast mode
- "Stepping stone" for physics community who may not be well-versed in global modeling
- Once workflow exists, simpler to add coupling and DA

Testbed Status: SCM Tier

- Ver 1.1 (repo on NOAA VLab)
  - Uses first version of IPD to interface with GFS physics
  - Driven by GCSS/GASS cases
  - Portable (multi-compiler, minimized dependencies)
  - Basic Python-based analysis/plotting script included
  - User + Technical Docs



Testbed Status: SCM Tier

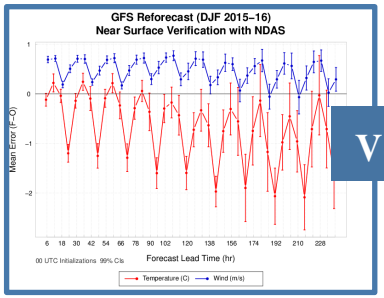
- Looking ahead…
  - Enhanced capabilities
  - User-contributed cases
  - Platform for testing IPD
  - Online interactive notebook format (i.e. Jupyter)
- Issues
  - Staying up-to-date with GFS physics
  - Additional suites
  - Evaluation metrics needed (incl. comp. efficiency)

Testbed Status: Global Tier

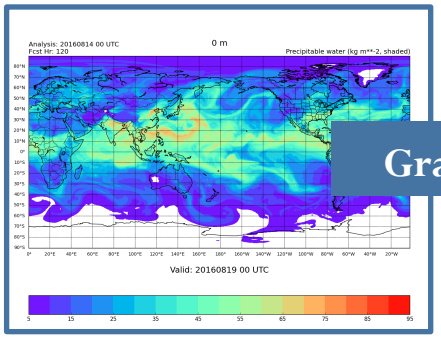
Workflow supplied by NOAA EMC

Init Datasets → Pre-proc → Forecast → Post-proc

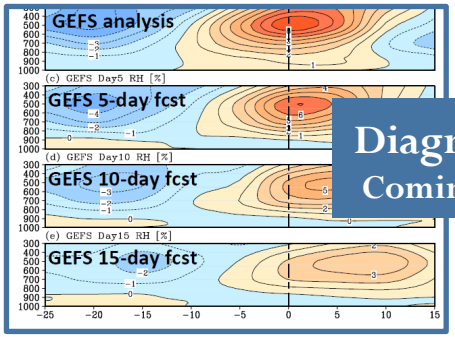
Workflow developed by GMTB



Verification



Graphics



Diagnostics
Coming soon!

Complementary workflows

EMC-based workflow

- GMTB keeping pace with EMC procedures
- GMTB/EMC collaborate to resolve issues on both sides

GMTB workflow

- Highly flexible and configurable
- EMC verification methods within DTC's **Model Evaluation Tools**
- Graphics and diagnostic suite being actively developed

Testbed Status: Global Tier

- **Where are we headed?**
  - **Near future:**
    - Run Grell-Freitas Test (focus: compare control forecasts using SAS against an experimental configuration using GF)
    - Expand diagnostic capabilities in testbed
  - **Looking ahead:**
    - Essential to engage with community in hierarchical testing and respond to feedback
    - Transition to global unified workflow system being actively developed by EMC and built around NITE principles

Discussion Points

- What objective measures of physics performance should the physics testbed provide to reviewers?
- What is the appropriate path forward for inclusion of a limited-area domain model in the testbed?
- Should more complex global runs be facilitated in the testbed? (DA? coupling?)
- What is the best way to engage the community?
- What is the best way to facilitate 2-way feedback?
- Are there resources available to run global tests? Process for obtaining?
- How should testing procedure differ for suites vs. schemes?

Extra

Testbed Status: SCM Tier

- Example of SCM Workflow for "Pre-made" case

1. Find relevant obs-based case with initialization and forcing
2. Use scripts to prepare data for SCM
3. Run SCM using baseline GFS physics configuration
4. Run SCM using modified physics
5. Evaluate. Rinse & repeat