Toward community-based evaluation of climate models: Obs4MIPs, the metrics panel and evaluation tools for CMIP6

Peter J. Gleckler

WGCM18, Grainau/Garmisch-Partenkirchen, October 8-10, 2014
Opportunities . . .

• As an ongoing exercise, the DECK provides a design target

• Scientists and software developers can further exploit the data conventions

• Many modeling groups now maintain a CMIP compliant data stream

• This environment commonality presents possibilities to greatly advance how we evaluate models by:
  
  • Applying same data conventions, more effectively organize observations used for model evaluation (obs4MIPs)
  
  • Retain CMIP evaluation capabilities for repeated use
obs4MIPs update
https://www.earthsystemcog.org/projects/obs4mips/

- A centralizing activity of the WCRP Data Advisory Council (WDAC)
- Now hosted on CoG with direct data search on ESGF
- All data is technically aligned with CMIP data structure
- Over 50 products with technical notes, and growing

WDAC observations for model evaluation task team:
Direct access to data on ESGF
Several consensus recommendations:

- Expand the dataset holdings (no consensus on priorities)
  - Higher frequency data and model output for more process-oriented evaluation, e.g., a “golden era” of sampling?
  - Relax the strictness of the model equivalent requirement
  - Reliable and defendable error characterization/estimation of observations

BAMS meeting summary (Ferraro et al., in review) with detailed meeting report to be made available on obs4MIPs CoG
Advancing obs4MIPs

Challenges:

• Task team refining protocol for including and documenting products (e.g., establishing a maturity matrix)

• Lots of technical details (preparing data, publishing on ESGF, etc.) .... WIP is made aware of infrastructural challenges

Opportunities

• More emphasis on higher frequency satellite data

• Setting a precedent for other observational communities

• Further alignment with CMIP (e.g., server side analysis via ESGF)
The WGNE/WGCM Climate Model Metrics Panel
Members: B. Ebert, V. Eyring, P. Friedlingstein, P. Gleckler (chair), H. Hewitt, Simon Marsland, R Pincus, K Taylor

**Status report**

Successfully promoted the development of model performance metrics

Identified some useful benchmarks (well-established in literature)

Monitored a rapidly evolving research topic (e.g., skill scores, process oriented, model independence, emergent constraints, ...)

Have *not* selected a "limited set of most important" limited set

Strategy for CMIP6 now formulated
Making analysis capabilities more accessible for repeated use

• The metrics panel working to develop a catalogue of CMIP related analysis tools and monitoring characteristics such as:
  
  • What software is required (e.g., CDO, R, NCL, CDAT, python, Fortran, etc)?
  • Are codes designed to work, as is, with the CMIP data structure?
  • What is the installation procedure/effort required?
  • ........

• Metrics panel to motivate adherence to CMIP standards, and encourage CMIP research community (MIPs) to contribute “shareable” codes

• Significant technical challenge, but potential for great benefit
Recent examples of developing capabilities for model evaluation

- PCMDI’s metrics package / ESMValTool
- NCAR Climate variability and diagnostics package
- CFMIP diagnostics and metrics codes
- ILAMB benchmarking
- GEWEX GASS
- Toolkit for Extreme Climate Analysis (TECA)

....
Summary

Significant challenges (examples)

- Many hurdles to organize a diversity of observations
- Scientists use different analysis tools (will continue to do so), complicating collaboration/sharing

Possibilities for CMIP6 and beyond

- At a minimum – some useful tools made available to modeling groups, standard diagnostics performed on all simulations with better access to obs
- Target – a transformation in how we do collaborative research