Coordinated Model Evaluation Capabilities

Paul Ullrich

Peter Gleckler (PCMDI)

Forrest Hoffman (RUBISCO)

William Collins (CASCADE)

Travis O'Brien (CASCADE)

U.S. DEPARTMENT OF ENERGY

Office of Science

Motivating Metric Coordination

- The number of metrics and diagnostics available to global research groups has been growing rapidly.
- Multiple metrics are available from groups around the world (polished to various degrees).
- Little coordination between separate projects, often requiring extensive developer-intervention when others want to use a particular code.
- Clear value in the use of a large suite of metrics:
    - Are features and their characteristics correct? (Feature-Based Metrics)
    - Are the associated upstream processes correct? (Process-Oriented Metrics)
    - Are quantities of value to downstream users / stakeholders correct? (Use-Inspired Metrics)

DOE RGMA Metric Development (not comprehensive)

Metrics	Leads
Drought	Paul Ullrich
Monsoons	Bill Boos, Colin Zarzycki, Paul Ullrich
Coastal Storms	Colin Zarzycki, Kevin Reed, Paul Ullrich
Forecasting	Weiran Liu, Colin Zarzycki, Paul Ullrich
Precipitation (base)	Peter Gleckler, Min-Seop Ahn, Angie Pendergrass
Precipitation Scales	Nick Klingaman
ENSO	Jiwoo Lee, Peter Gleckler
ETCCDI	Michael Wehner, Peter Gleckler
MJO	Daehyun Kim, Charlotte Demott, Jiwoo Lee and Min-Seop Ahn
Sea Ice	Julie McClean, Detelina Ivanova, Sarah Gille, Peter Gleckler
AMOC	Gokkan Danabasaglu and others
Arctic (POLeSTAr)	Andrew Roberts
Oceans	Wilbert Weijer, Wei Cheng

Flagship Packages

PMP
ILAMB / IOMB
TECA
TempestExtremes

E3SM / SciDAC efforts

E3SMdiags
ARMdiags
LIVVkit
MPAS-Analysis

CESM Diagnostics

PCWG sea-ice diags
AMWG diags
Climate Variability diags

Pathways for Metric Development

1 Contribution to existing flagship packages

PCMDI Metrics Package

ILAMB

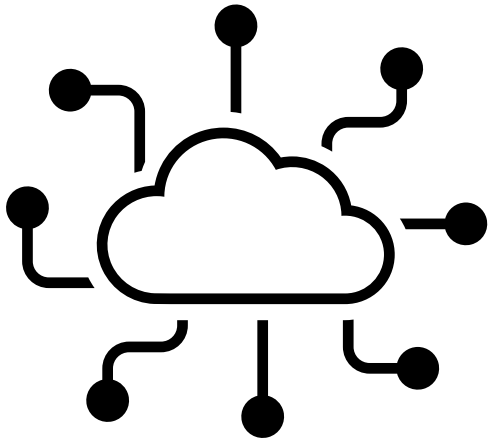
International Land Model Benchmarking Project

TEMPEST EXTREMES

CASCADE TECA

Toolkit for Extreme Climate Analysis

2 Development of Standalone Package



CMEC Project Goals

1. Develop robust **standards** for the development of metrics and diagnostics packages.

2. Develop accompanying tools for **coordinated execution** of metric packages and **visualization of / interaction with** metrics and diagnostics package output.

3. Build **connections across projects and agencies** related to model evaluation activities.

Notably: CMEC is not intended to curate or maintain metrics packages.

A Cross-Agency Solution

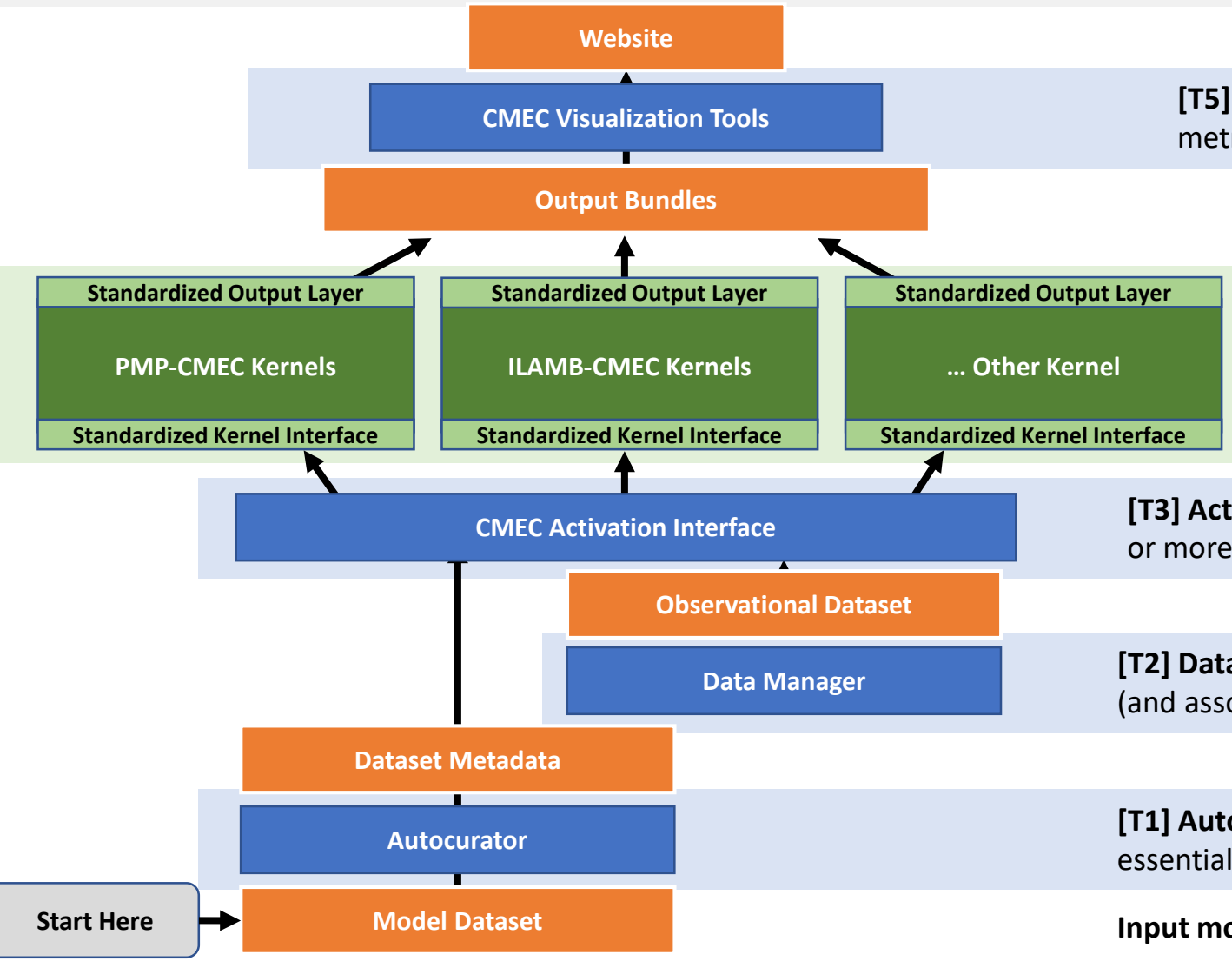
There are substantial overlaps among agency interests when it comes to model evaluation (and within DOE across programs).

CMEC is building towards a US cross-agency collaborative for metric and process-oriented diagnostics (PODs) development activities.

CMEC standards have been developed in collaboration with the NOAA Model Diagnostics Task Force (MDTF).

Similar to RGMA, NOAA MAPP program directly supports the development of diagnostics packages.

CMEC Tools



[T5] Visualization Tools: Utilities that allow a user to interact with the metric and diagnostics data output from one or more metrics packages.

[T4] Standardized Kernels and Interfaces that perform the relevant calculations on the model and observational data. Lightweight interfaces allow for data to maintain CMEC standards.

[T3] Activation Interface: A utility to enable applciation of one or more metric kernels to a particular model dataset.

[T2] Data Manager: A utility for managing observational data (and associated metadata) needed by the metric kernels.

[T1] Autocurator: A utility for generating dataset metadata, essentially a table of contents to the dataset.

Input model dataset to be analyzed.

CMEC Standards for Metrics and Diagnostics

- **Compatibility with CMEC effectively requires:**
  - A JSON file that provides metadata on the package being executed
  - A bash script that allows for “lowest common denominator” execution of the metrics module
  - Metrics output compatible with the package

- **“Lowest common denominator” (LCD) execution** – modules should run over:
  - Path to the base directory of the metric module
  - Path to the observational data and model data
  - Path where output should be written
  - List of metric modules to be executed (and their configuration name)

Example Module Descriptor File (JSON)

```
{
    "info": "Settings used to prototype application of CMEC standards to the PMP",
    "obslist": {
        "obs_name": "GPCP-2-3",
        "version": "2-3"
    },
    "settings": {
        "async": "DASK",
        "description": "Calculates mean climate metrics",
        "driver": "cmec_pmp_meanclimate_demo1.bash",
        "name": "meanclimate",
        "long_name": "PMP mean climate driver",
        "runtime": {
            "PMP": "1.2",
            "CDAT": "8.2"
        }
    },
    "varlist": {
        "pr": {
            "frequency": "mon",
            "units": "kg m-2 s-1"
        }
    }
}
```

Example Metrics Descriptor File (JSON)

```
{
   "DIMENSIONS":{
      "json_structure":[
         "product",
         "metric",
         "statistic"
       ],
       "dimensions":{
          "metric":{
             "TC Count":{},
             "TC Days":{},
             "Accumulated Cyclone Energy":{},
             "Latitude of Maximum Intensity":{}
          },
          "statistic":{
             "indices":[
                 "Bias",
                 "Seasonal Correlation"
             ]
          },
          "product":{
             "ERA-I":{},
             "20CRv3":{},
             "MERRA":{},
...
```

```
...
    "RESULTS":{
       "ERA-I":{
          "TC Count":{
             "Bias": -7.6,
             "Seasonal Correlation": 0.89
          },
          "TC Days":{
             "Bias": -38.2,
             "Seasonal Correlation": 0.92
          },
          "Accumulated Cyclone Energy":{
             "Bias": -100.3,
             "Seasonal Correlation": 0.95
          },
          "Latitude of Maximum Intensity":{
             "Bias": 5.9,
             "Seasonal Correlation": 0.65
          }
          },
       "20CRv3":{
          "TC Count":{
             "Bias": -6.7,
             "Seasonal Correlation": 0.91
          },
...
```

CMEC Live Demo

CMEC Activation Interface

Metric Visualization with the LMT Unified Dashboard

How to Be Involved?

Email me: **Paul Ullrich <paullrich@ucdavis.edu>**

Contact Metric Flagship Product Managers

Product	Focus	RGMA Project	Contact
PMP	Climate System	PCMDI	**Peter Gleckler <gleckler1@llnl.gov>**
ILAMB / IOMB	Land Surface / Ocean	RUBISCO	**Forrest Hoffman <forrest@climatemodeling.org>**
TECA	Extremes / ML	CASCADE	**Bill Collins <wdcollins@lbl.gov>**
TempestExtremes	Feature-Tracking	HyperFACETS	**Paul Ullrich <paullrich@ucdavis.edu>**

Coordinated Model Evaluation Capabilities

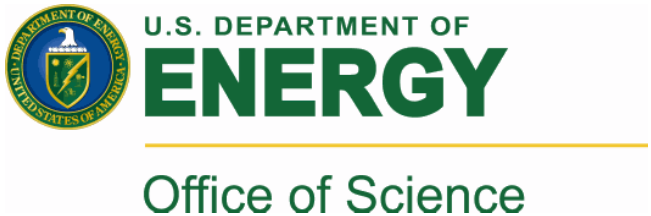
Paul Ullrich

Peter Gleckler (PCMDI)

Forrest Hoffman (RUBISCO)

William Collins (CASCADE)

Travis O'Brien (CASCADE)



Metrics Survey: Biggest Gaps

- How to translate the knowledge gained from the model diagnostics to model physical parameterization development.
- More focus on the full distributions of variables (not just means).
- Need more tools from information theory / machine learning.
- A unified framework for distribution of publicly available observational data for use in metrics packages, not requiring exhaustive downloads.
- Weather diagnostics such as feature tracking have high computational cost.
- High-latitude ocean observations especially in the Southern Ocean and over the Antarctic continental shelf.
- Metrics that look at and connect phenomena at different space/time scales.
- More comparison of different metrics of the same phenomena (e.g. MCSs).
- Systematic comparison using different observational datasets.
- More online evaluation capabilities, including feature tracking.
- Difficult to design metrics equally applicable to observations and models (lack of model data).
- Diagnostics are not formulated as closures on water / energy / carbon cycles.