WGNE/WGCM Metrics and Diagnostics Panel

Members selected by relevant and diverse experience, and potential for liaison with key WCRP activities

Current members:

Beth Ebert (BMRC) – JWGV/WWRP, WMO forecast metrics

Veronika Eyring (DLR) – WGCM/SPARC/CMIP6, stratosphere, ESMs

Pierre Friedlingstein (U. Exeter) – IGBP, carbon cycle

Peter Gleckler (PCMDI), chair – WDAC, atmosphere and ocean

Simon Marsland (CSIRO) – CLIVAR OMDP, WGCM, ocean

Robert Pincus (NOAA) – GEWEX/GCSS, clouds/radiation

Karl Taylor (PCMDI) – WGCM/CMIP6, atmosphere

Keith Williams (Met Office) – WGNE, Transpose AMIP, clouds

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC. LLNL-PRES-XXXXXX

Broadening scope: WGNE - WGCM - WMAC

- WGNE formed a small panel to identify a limited set of "performance metrics" for climate models (2009)
- Panel expanded, becomes a joint WGNE/WGCM effort (2013)
- Broadened scope recommended by WMAC to include "diagnostics" (2015)

Some context

In recent years, extensive CMIP research related to objective measures of model performance, including:

- New metrics, methods and multi variate scores
- Process-oriented
- MME vs PPE
- Model weighting
- Model dependence
- Emergent constraints
- Tuning

- CMIP DECK/Historical was established to provide continuity, inspiring the ongoing benchmarking of models simulations

- DECK/Historical has brought focus to the panel's remit

Towards routine benchmarking of the CMIP DECK/historical simulations

Motivation

- Routine summaries with less re-inventing
- Facilitate national assessments, the IPCC process, etc.
- Provide more rigorous testing
- More directly contribute to model development via quick feedback

Robust analysis codes are a viable mechanism to accomplish this, thanks to the design target provided by the CMIP data conventions.

An incomplete listing of community-based capabilities that may be relevant for routine evaluation of CMIP DECK simulations

ESMValTool (Eyring et al, GMD, 2016)

PCMDI Metrics Package (Gleckler et al., EOS, 2016)

NCAR's Climate Variability Diagnostics Package (Phillips et al,2014)

CFMIP diagnostics

ILAMB (Luo et al., 2012)

ARM Diagnostics and Metrics package

TECA (Prabhat et al., 2012)

MJO task team diagnostics

NOAA MAPP process-level team

CLIVAR basin panels

….

What can the panel do to help advance routine evaluation?

- Work with expert groups within WCRP who may be developing targeted metrics or diagnostic capabilities; identify key gaps and strive to fill them
- Maintain a catalogue of developing model evaluation capabilities and expert groups defining metrics
- Work with WGCM's WIP to ensure required infrastructure is advancing (Eyring et al., ESD, 2016)
- Establish and promote a set of "best practices" for making results publically available

An incomplete list of community-based capabilities that may be relevant for routine evaluation of CMIP DECK simulations

ESMValTool (Eyring et al, GMD, 2016)

PCMDI Metrics Package (Gleckler et al., EOS, 2016)

NCAR's Climate Variability and Diagnostics Package (Phillips et al,2014)

CFMIP metrics and diagnostics

ILAMB (Luo et al., 2012)

ARM Diagnostics and Metrics package

TECA (Prabhat et al., 2012)

MJO task team diagnostics

NOAA MAPP process-level team

CLIVAR basin panels

….

Closing thoughts

- New terrain for CMIP: “ongoing” experimentation (DECK/Historical) combined with community benchmarking tools
- Having several capabilities with some overlap will help us demonstrate reproducibility and learn more quickly how to improve approach
- Important how the results are documented and disseminated

Terms of Reference: Metrics & Diagnostics Panel

(Updated 2016)

- Foster an environment to advance routine evaluation of climate models
- Coordinate with other WCRP/CLIVAR activities that are actively developing diagnostics and performance metrics
- Identify analysis routines and packages that may be of potential use to modeling groups and researchers, and encourage functionality with the CMIP data conventions
- Ensure that well-established capabilities are applied to the CMIP DECK and Historical experiments, with results made readily accessible
- Encourage and facilitate model evaluation research by identifying key areas needing work and possibly organizing workshops
- Progress and terms to be reviewed annually by and WGCM and WGNE.

The PCMDI Metrics Package (PMP)

https://github.com/PCMDI/pcmdi_metrics

Emphasizes a diverse suite of relatively robust high level summary statistics objectively comparing models and observations across space and time scales

- End-to-end provenance to ensure reproducibility

- Contributions from:
  - PCMDI research
  - Collaborations with expert teams
    (e.g., CLIVAR Pacific metrics group, Monsoon Panel, MJO task force, Ocean Model development Panel)

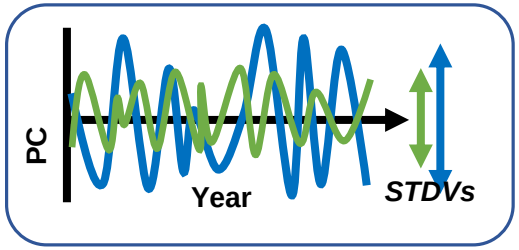
- 6 modeling groups using the PMP with 3 more getting started

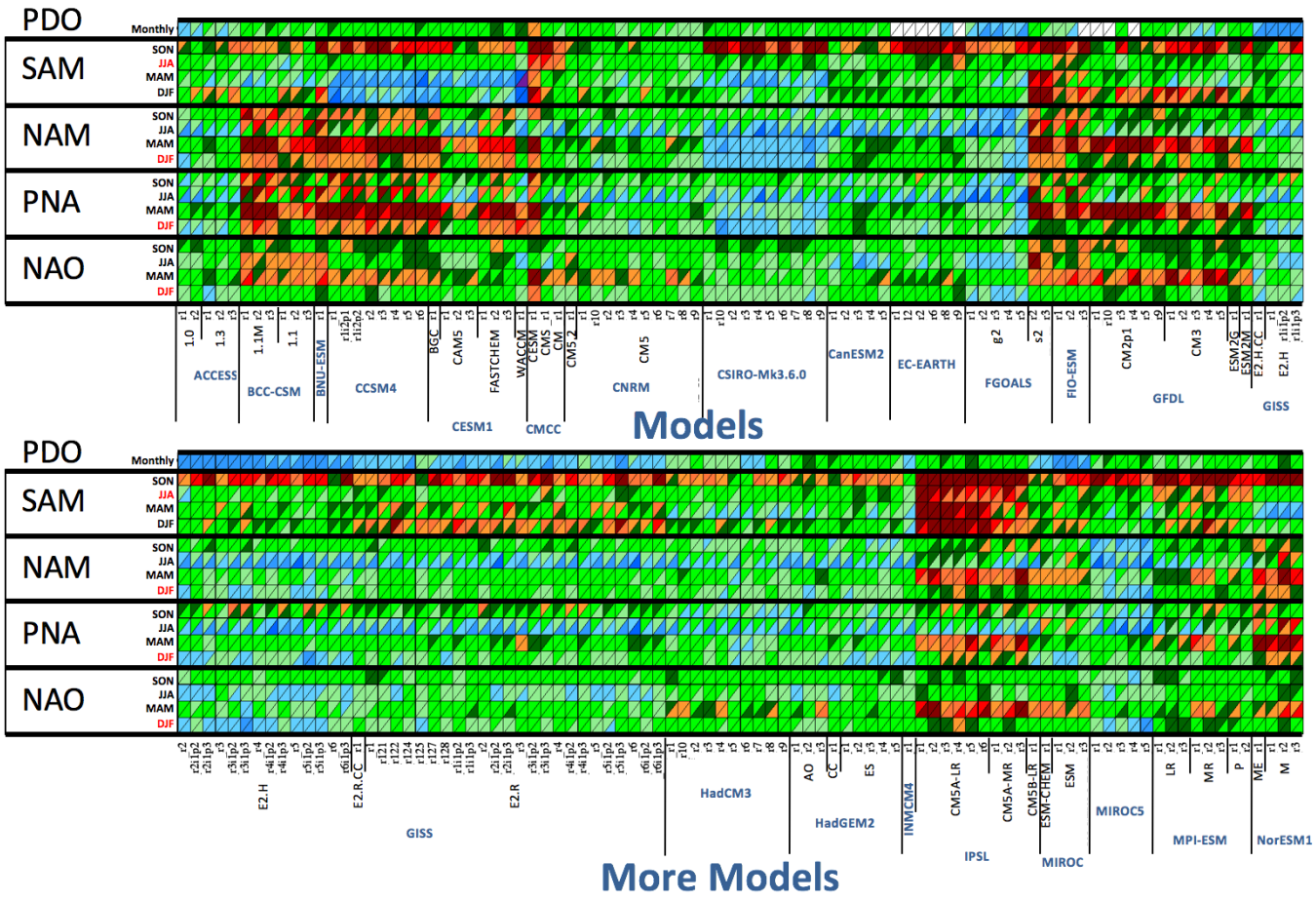
Gleckler et al., EOS, 2016

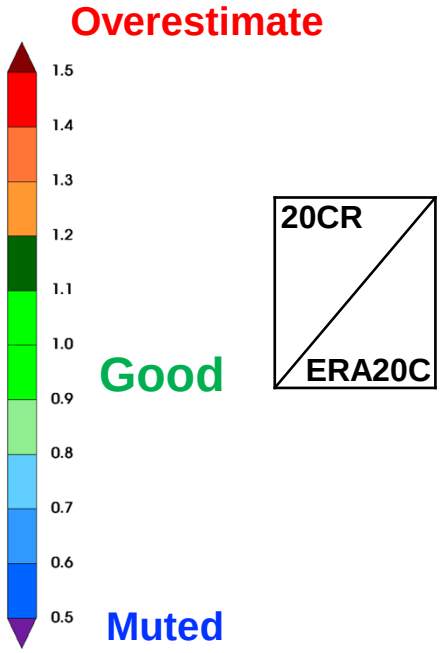
The PCMDI Metrics Package (PMP) v1.2

- Orthogonally decomposed climatological error statistics
- Monsoon precipitation indices
- Sea-ice “sector scale” metrics
- Ocean Argo based T & S
- Cloud properties (via S. Klein group)
- Precipitation diurnal cycle and intermittency (incl sub daily)
- ENSO metrics in collaboration with CLIVAR Pacific Basin Panel
- Dominant extra-tropical modes of interannual variability

Simulated/reference amplitude ratios
CMIP5 historical simulations (1900-2005)







PMP feedback to modeling groups

- Simulation summaries will be provided to modeling groups soon after their DECK and Historical simulations are made available via ESGF
- We provide support to modeling groups interested in using the PMP

A possible aid to modelers

- Help identify unexpected degradation against backdrop of general improvement
- Determine if “red flags” are significant (in MME context), to help decide if they should influence development priorities

Coordinated Model Evaluation Capabilities (CMEC)

- developing high level API and provenance syntax that can be shared across independent packages

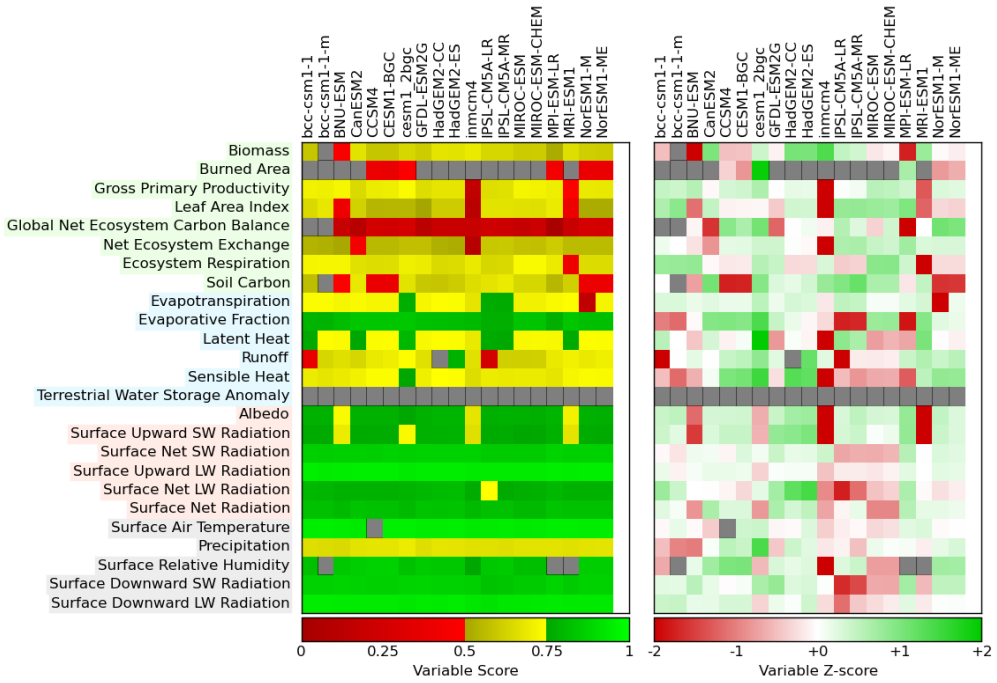
- expected outcome: much easier for modeling groups to use multiple packages

Currently includes:

  - PCMDI Metrics Package (PMP)
  - ARM diagnostics for GCMs (ARMDiag)
  - Parallel Toolkit for Extreme Climate Analysis (TECA)
  - International Land Model Benchmarking Project (ILAMB)
  - . . .

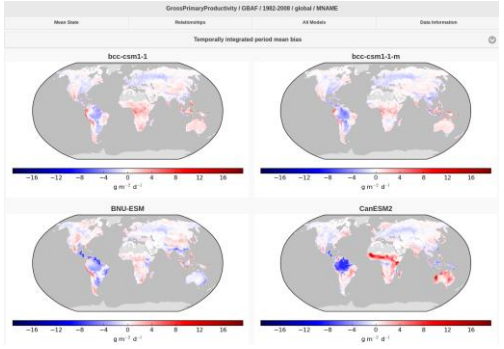
- Will provide DECK summaries as a collective

International Land Model Benchmarking (ILAMB) Package



- Provides systematic assessment of land model results compared with observations
- Scores model performance across a wide range of independent benchmark data
- Includes comparison of functional (variable-to-variable) relationships
- Results from an international community coordination effort for designing metrics
- Built on python and open standards

ILAMB is open source model benchmarking package that runs in parallel.
ILAMBv2 can be downloaded from https://www.bgc-feedbacks.org/software/



BGC Feedbacks | Argonne NATIONAL LABORATORY | Berkeley Lab | Los Alamos NATIONAL LABORATORY | NCAR | Oak Ridge National Laboratory | UCI | MICHIGAN

An update on obs4MIPs

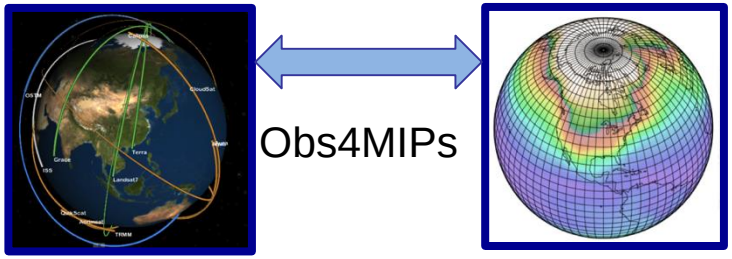
WDAC Observations for Model Evaluation Task Team

Peter Gleckler, co-chair, PCMDI and Duane Waliser, co-chair, JPL/NASA
Mike Bosilovich, GSFC/NASA
Helene Chepfer, IPSL
Carol Anne Clayson, WHOI
Veronika Erying, DLR
Robert Ferraro, JPL/NASA
Pierre-Phillipe Mathieu, ESA
Jerry Potter GSFC
Roger Saunders, UKMO
Jörg Schulz, EUMETSAT
Karl Taylor, PCMDI
Jean-Noël Thépaut, ECMWF

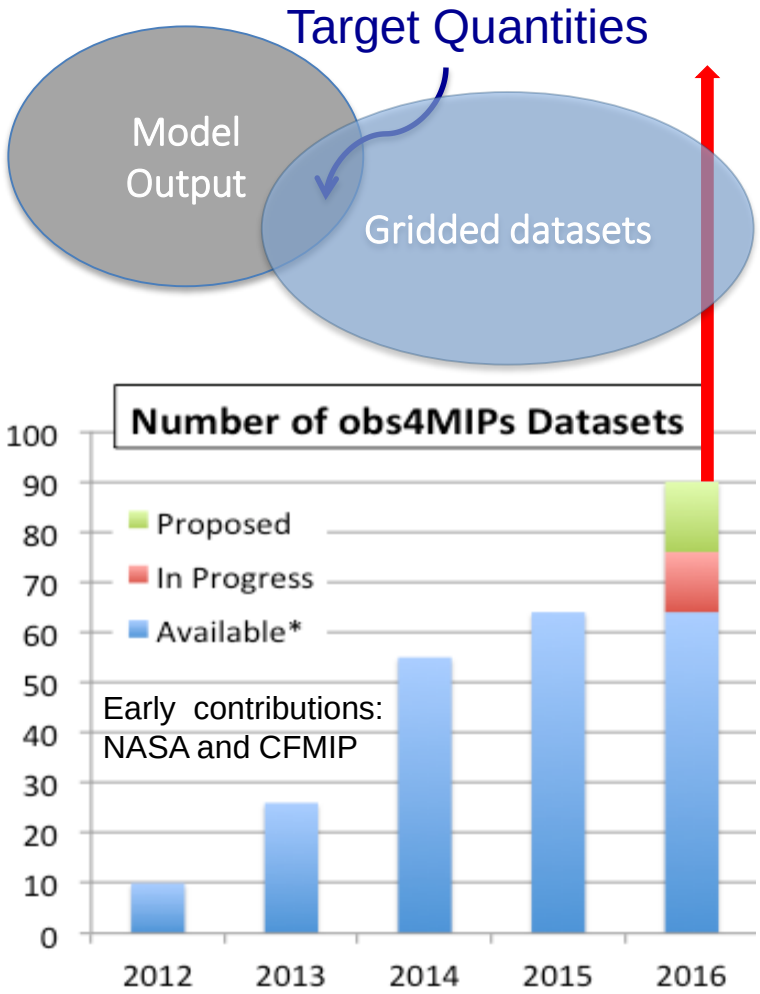
Additional regular contributors: Otis Brown, Michel Rixen, Sophie Cloché (IPSL)
Tsengdar Lee (NASA) and Renu Joseph (DOE)
Luca Cinquini (JPL) – CoG technical support
Denis Nadeu (PCMDI) – CMOR development
Paul Durack (PCMDI) – Data specifications
Jim Biard (NCEI) and Matthias Tuma (WCRP) – beta testers
… and many others

obs4MIPs

https://www.earthsystemcog.org/projects/obs4mips/



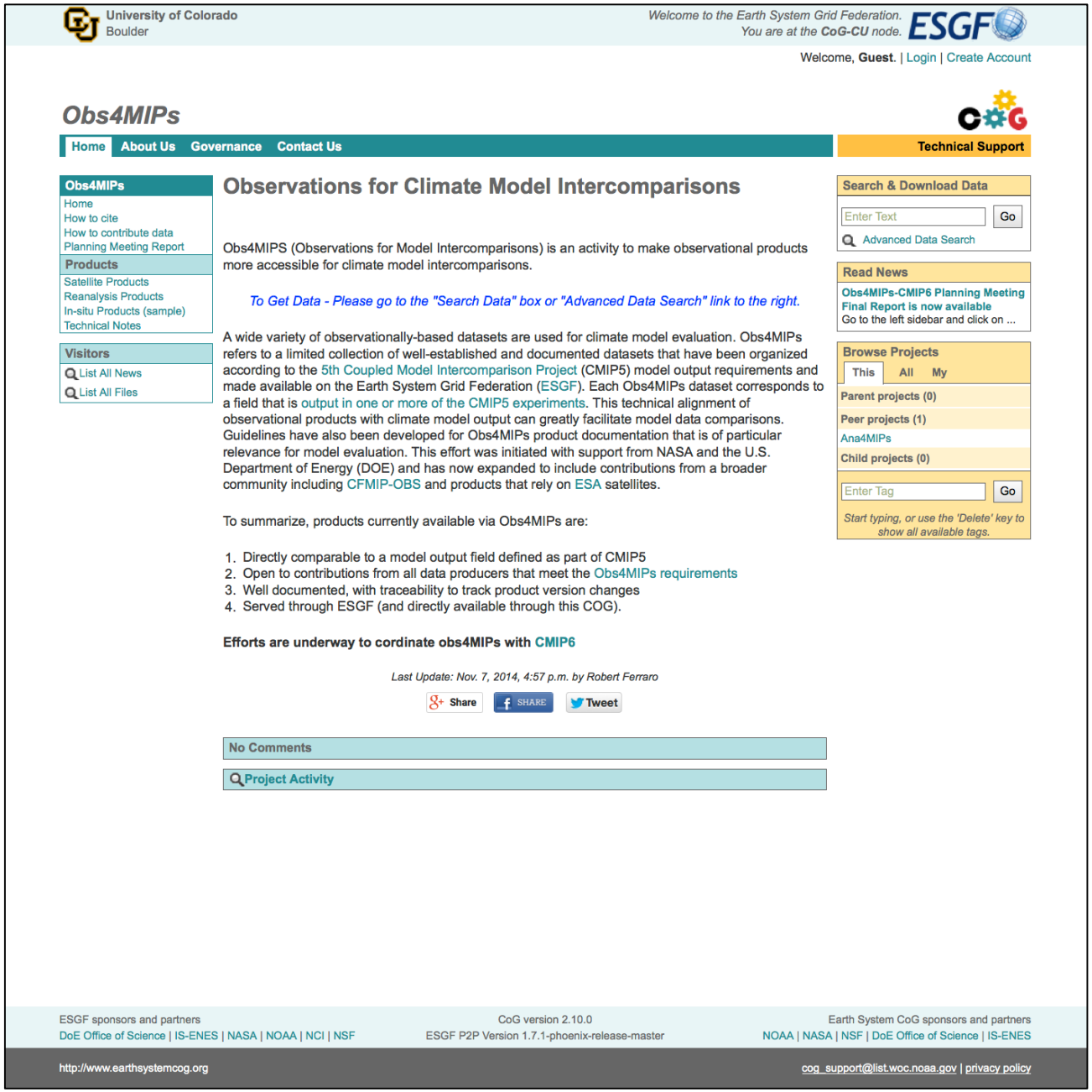
- A project for documenting and disseminating observations for climate model evaluation in WCRP MIPs, notably CMIP.
- Data accessible on ESGF with CMIP model output, adhering to coordinated conventions
- A well defined tech note protocol
- Guided by the WCRP Data Advisory Council obs4MIPS Task Team



Complete (~125*)
In Progress (~25)
New datasets in queue (~100)

…. and growing!

Obs4MIPs data (and ana4MIPs) are available through the CoG/ESGF



7 ESGF notes currently supporting obs4MIPs

Key challenges being addressed

- Making the process for contributing data more efficient

- Broadening scope (recommendations from 2014 workshop)
  - Higher frequency data
  - Data uncertainties
  - Obs “ensembles” and other supporting data
  - Allow forward models or simulators
  - Other classes of data (notably in-situ)

- How to accommodate the diverse spectrum in data quality and “apples to apples” model/data comparisons ?

- Keeping products up-to-date and version control

Recent progress

- The obs4MIPs data specifications (ODS 2.1) are closely aligned with CMIP6 (will enable enhanced searching methods)
- CMOR3 now accommodates observational gridded data
- Metadata alignment provides a hook to es-doc to support Tech Notes
- DKRZ-provided citation service coordinated with CMIP6
- Pending additions: NOAA, ESA CCI, DWD and others now actively preparing new data products
- Solutions for several longstanding difficulties (discussed next)

Recent strategic advancements

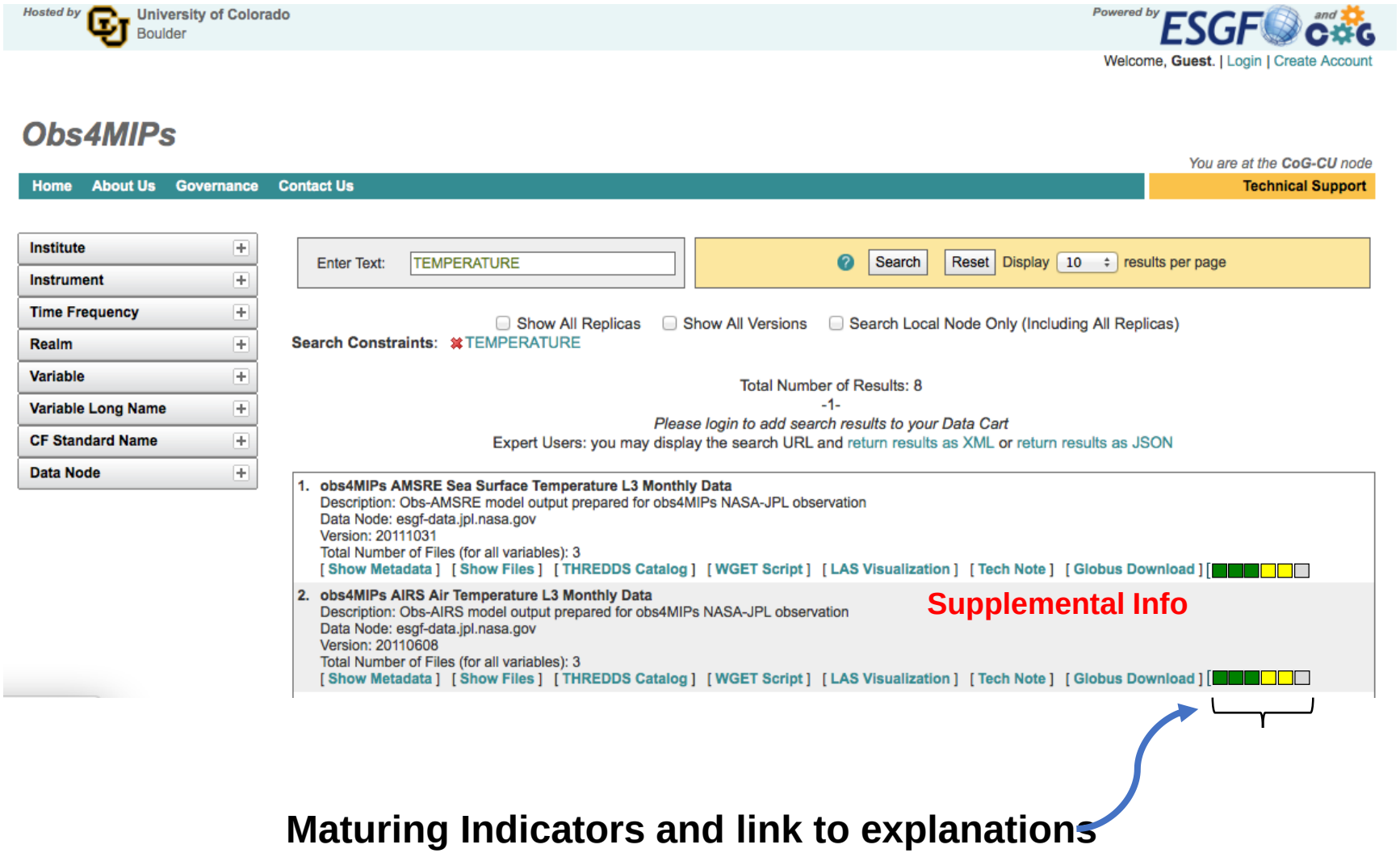
- A concise set of Dataset Suitability and & Maturity Indicators
- Accommodating a wide range of supplemental data and metadata beyond the “best estimate”

obs4MIPs Dataset Suitability & Maturity Indicators

Technical Requirements		***Dataset Suitability and Maturity***			***Comparison Complexity***
Meets obs4MIPs data technical requirements	Includes obs4MIPs technical note information	Closeness or robustness of measurement to observed reference quantity	Maturity with respect to climate model evaluation	Provision for robust uncertainty information	Complexity of Model Observation Comparison
Data suitably processed with CMOR and/or consistent with obs4MIPs standards	Complete technical note information provided	Firmly established and/or validated methodology	Multiple peer-reviewed examples of application to CMIP climate model evaluation	Uncertainty information provided per retrieval/grid point	Comparison can be made directly with CMIP model output variable
Largely complete with minor metadata inconsistencies	Technical note information incomplete and/or could be improved	Indirect means of calculation or observations only providing partial constraint (e.g. ocean surface latent heat flux)	One peer-reviewed example of application to CMIP climate and/or examples of other sorts of model evaluation.	General uncertainty information given relative to the methodology and dataset as a whole - backed by actual field/in-situ validation exercises	Comparison requires some simple post processing of CMIP output variable(s) (e.g. vertical integral or ratio of two variables)
Non-compliant. Should be removed from database!	Technical note not provided	Largely model-derived quantity (e.g. LAI, root zone soil moisture, NPP)	As of DATE-TBS, no significant application to climate model evlauation	No uncertainty information provided	Comparison requires complex processing of CMIP output (e.g. "simulator", budget calculation)

Prototyping Dataset Suitability And Maturity Indicators

Based on typical obs4MIPs dataset search



Maturing Indicators and link to explanations

Summary and Perspective

- obs4MIPs is closely coordinated with CMIP6 via data specs, etc

- Efforts are underway to expand the scope of the project by adding dataset maturity indicators and enabling supplemental information

- In addition to serving the CMIP analysis community:
  - obs4MIPs benefits ESMValTool, PMP and other evaluation tools
  - value in this effort is appreciated beyond model evaluation

- WDAC is encouraging further expansion, particularly to include insitu data. This is being taken seriously but more work is needed

Links to related material

- ESGF Dec 2016 conference report (pdf)
- CMIP6 data specifications (google docs)
- Climate Model/obs Output Rewritter, CMOR (website)
- obs4MIPs draft data specification (google docs)
- Draft user guide for preparing obs4MIPs (google docs)
- obs4MIPs tables controlled vocabulary (github)