

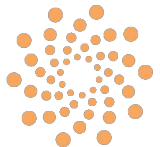


Update on ESGF and Model Benchmarking

10th United States Climate Modeling Summit (2–3 May 2024)

Forrest Hoffman, Oak Ridge National Laboratory

3 May 2024

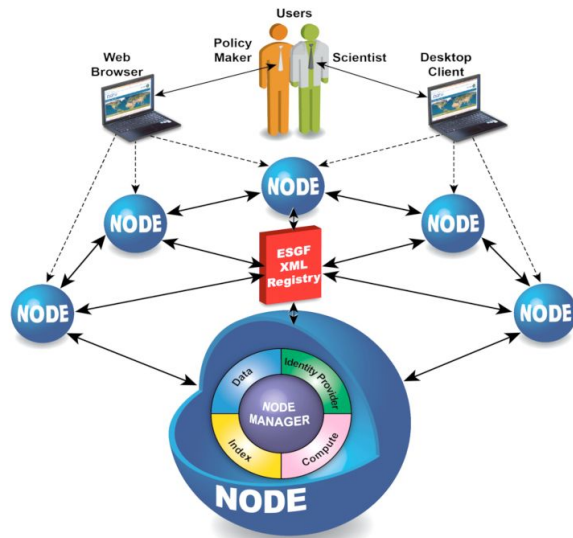




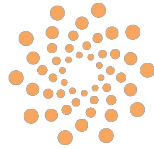
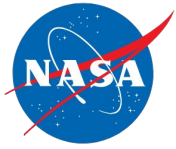
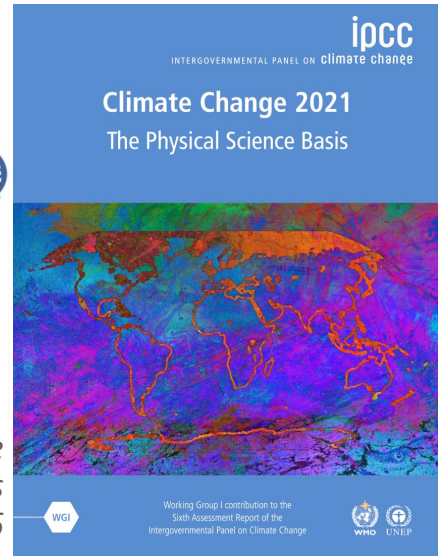
What is the Earth System Grid Federation?

- **Earth System Grid Federation (ESGF)** is an *international consortium* and a *globally distributed peer-to-peer network of data servers* using a common set of protocols & interfaces to archive and distribute climate & Earth system model output and related input, observational, and reanalysis data
- **Open Science data** are used by scientists all over the world to investigate consequences of possible climate change scenarios and Earth system feedbacks

ESGF Conceptual Diagram



Model output data from ESGF are used for research that underpins IPCC Assessment Reports, like AR6



Logos represent primary international contributors: US Department of Energy, NASA, NOAA, NSF, European IS-ENES Project, and Australian NCI



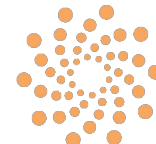
ESGF Organization

ESGF Steering Committee

Justin Hnilo (Chair), DOE
Ben Evans, NCI
Adrienne Simonson, NOAA
Sylvie Joussaume, IPSL (ENES)
Tsengdar Lee, NASA
Forrest Hoffman (Ex-Officio), DOE/ORNL
Philip Kershaw (Ex-Officio), CEDA

ESGF Executive Committee

Forrest Hoffman (Co-Chair), DOE/ORNL
Philip Kershaw (Co-Chair), CEDA
Sasha Ames, DOE/LLNL
Rachana Ananthakrishnan, DOE/ANL
Laura Carriere, NASA
Ben Evans, NCI
Stephan Kindermann, DKRZ
Christian Pagé, CERFACS
Aparna Radhakrishnan, NOAA/GFDL





Tenth ESGF Conference

- Held last week in Rockville, Maryland
- John Dunne joined the meeting to share CMIP priorities and current timeline
- ~50 in-person attendees from 8 countries (Australia, France, Germany, Italy, Japan, Sweden, United Kingdom, USA)
- ~69 virtual registrants from 18 countries
- **Primary objectives** of conference were to
 - **Share all current development activities** across the Federation
 - **Develop a roadmap** for collaborative activities necessary to deploy operational ESGF infrastructure to support CMIP AR7 Fast Track





Conference Agenda

- Accomplished through a minimal set of plenary presentations and a **large amount of time addressing specific topics in breakout groups**
- Message from Jay Hnilo (DOE) and the ESGF Steering Committee
 - US-DOE is committed to continue helping deliver a solution for CMIP
 - The SC is committed to collaborating across institutions for building out innovative solutions
 - We have confidence in the assembled groups to build a modernized framework that is more available and resilient and that enables more community access and innovation
- Agenda is available at <https://bit.ly/3vZlTsj>



Agenda

This agenda is designed to facilitate collaborative interactions and creation of a roadmap for ESGF development activities aimed at modernizing the system for CMIP AR7 Fast Track simulation data. While important categories of development, deployment, and operations are included, the agenda may be changed when necessary to maintain productive use of face-to-face meeting time. Some open meeting times are available in breakout rooms. Plenary Sessions will be broadcast by video conference. Separate meeting rooms will be available for breakout sessions.

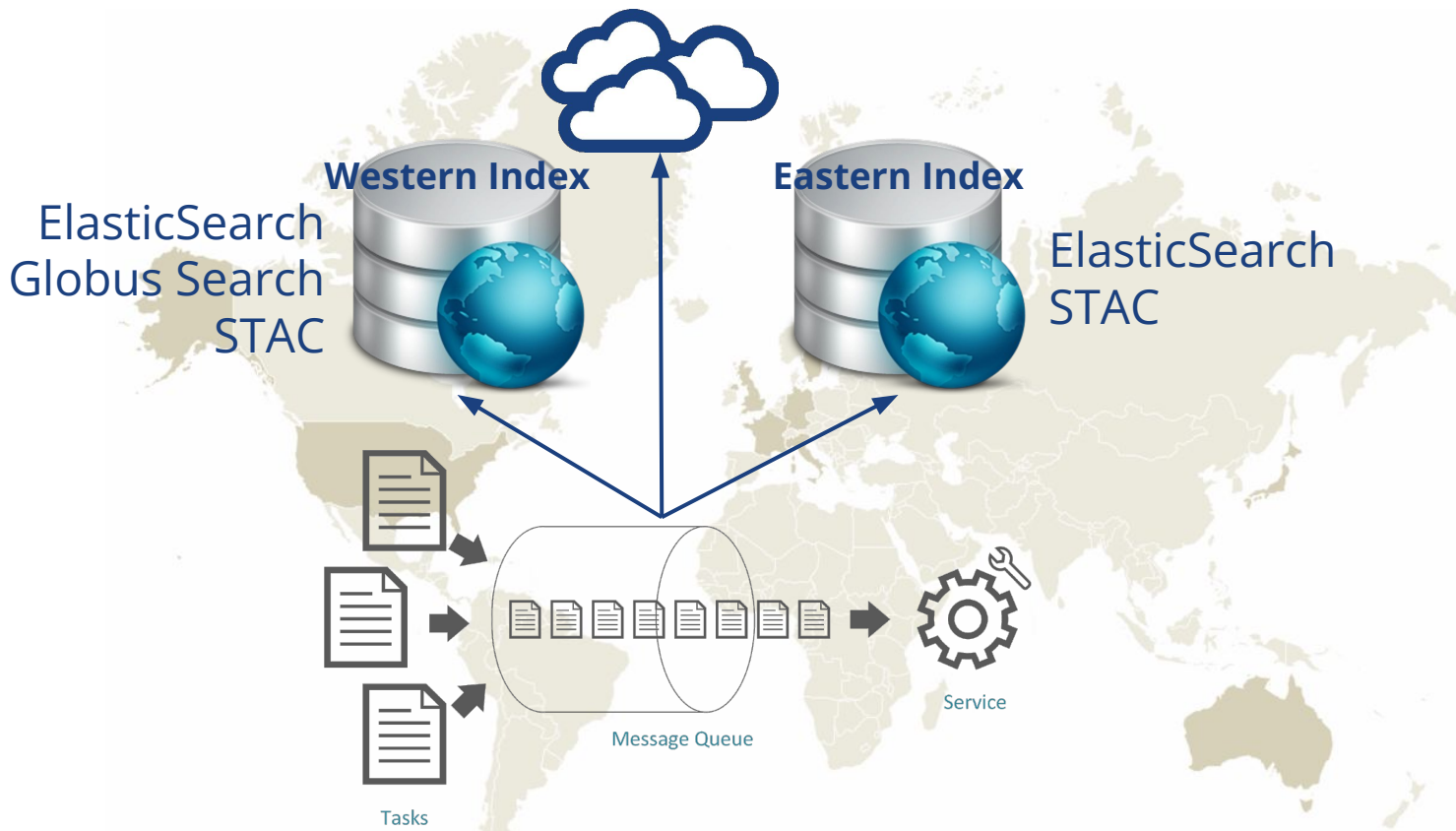
Plenary sessions will be held in (Grand Ballroom) Salon F&G on the Lobby Level of the hotel. Breakout Room 1 is in Salon H, and Breakout Room 2 is in Salon C, both also on the Lobby Level of the hotel. Side meetings can be held in the Seneca Boardroom, one level down from the meeting rooms.

Contacts: Taylor Clower-Mettica <clowertm@ornl.gov> and Forrest M. Hoffman <forrest@climatemodeling.org>

Tuesday, April 23, 2024			
EDT	Description	Session Chair(s)	Location
7:30	ESGF Conference Registration Desk Opens		
8:00	Buffet Breakfast		
8:00	Networking Session		Plenary (F&G)
9:00	Session 1: Welcome and Introductions		Plenary (F&G)
9:00	Welcome to the ESGF Conference and Meeting Logistics	Gary Geernaert, Forrest Hoffman	
9:10	Introduction to the ESGF Steering Committee and Supporting Programs	Justin Hnilo, Tsengdar Lee, Sylvie Joussaume, Ben Evans, Adrienne Simonson	
9:20	Introduction to the ESGF Executive Committee	Forrest Hoffman, Philip Kershaw	
9:25	Introduction to the Coupled Model Intercomparison Project (CMIP), the CMIP International Program Office (CMIP-IPO), CMIP Task Teams & Fresh Eyes on CMIP, and the CMIP7 Timeline	John Dunne	
9:45	Introduction to the WGCM Infrastructure Panel (WIP)	Matthew Mizieliński, Paul Durack	
9:55	Overview of the CMIP Node Operations Team (CDNOT)	Sasha Ames	Plenary (F&G)
10:00	Charge for the meeting: Establish roadmap for CMIP7 readiness • ESGF Data Lifecycle: from Model to Analysis - https://docs.google.com/presentation/d/18kJXCShAl2RBrFeDBpzF2Aiv08bXPbxjv1H6_pFE/edit	Phil Kershaw, Sasha Ames	Plenary (F&G)
10:30	Morning Break		
11:00	Charge for the meeting (Continued) • Present ESGF Roadmap for CMIP7 as it stands (25 min) • Outline session structure (5 min): - Charge - SMART objective (which must tie into the roadmap: what must be done, who will do it, what timeline, what are the dependencies) - Discussion time, work (up to three breakouts) - Report back, decision on next step i.e. how it fits into roadmap (including timings, e.g. task X will take 2 months starting in Oct 2024), what effort/resources needed and who will do it	Phil Kershaw, Forrest Hoffman	Plenary (F&G)
11:30	Session 2: Data Publication	Sasha Ames	Plenary (F&G)
11:40	Breakout Sessions		



Major Accomplishment: New Index Strategy

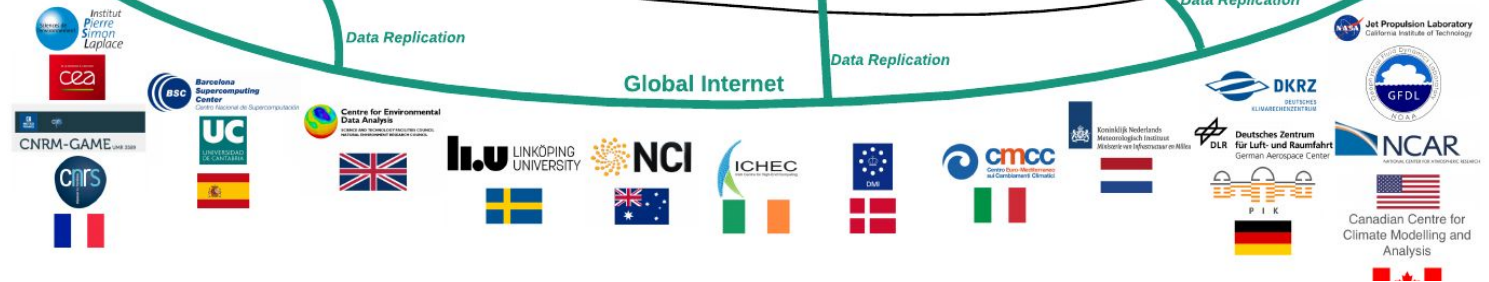
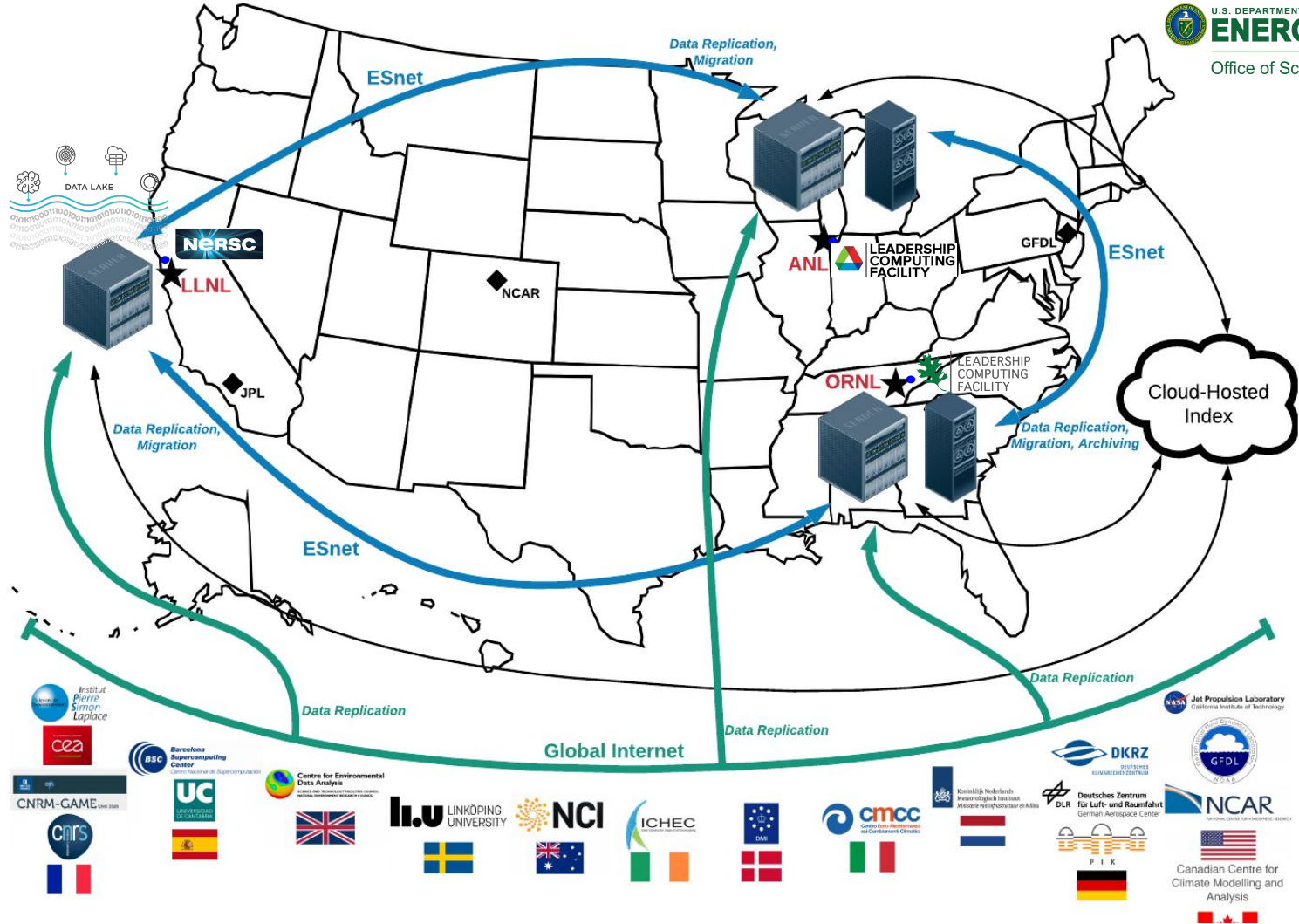


Publication and Replication



DOE's Next Generation Earth System Grid Federation

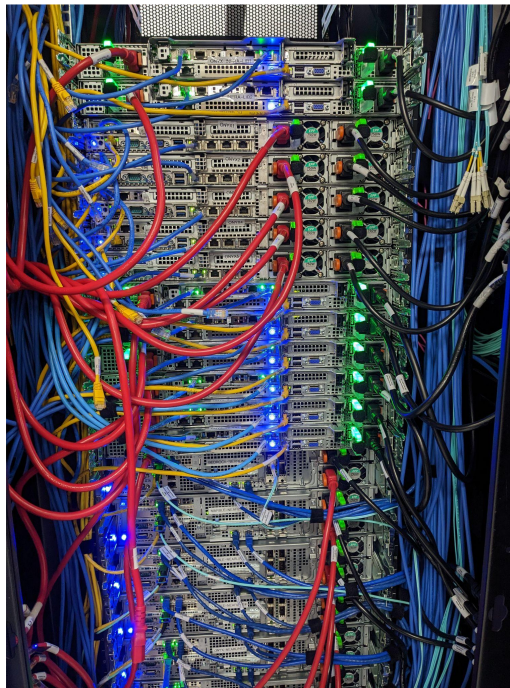
- As many as three nodes co-located at DOE's major computing facilities
- Replicating data from the global Federation
- Providing cloud indexing and tape archiving



Data and Index Nodes Deployed at ORNL

- Containerized server software deployed on the shared Onyx cluster is serving 8 PB of Coupled Model Intercomparison Project (CMIP5 and CMIP6) data at ORNL
- Data are stored on the new Themis hierarchical storage platform, providing on-disk copy for fast access to frequently used data and backup copies on two tapes for all data
- Hardware investment at ORNL has been in storage capacity (fully operational)
 - 15 PB of disk
 - 30 PB of tape (for redundant backup)

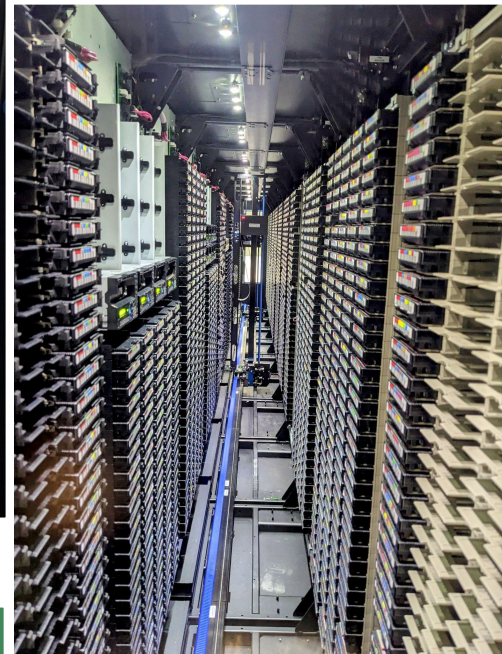
Delivered ahead of schedule and under budget!



The Onyx cluster hosts the ESGF containerized data & index nodes

Data and services reside in the Open Network Enclave of NCCS to provide fast and open access to data

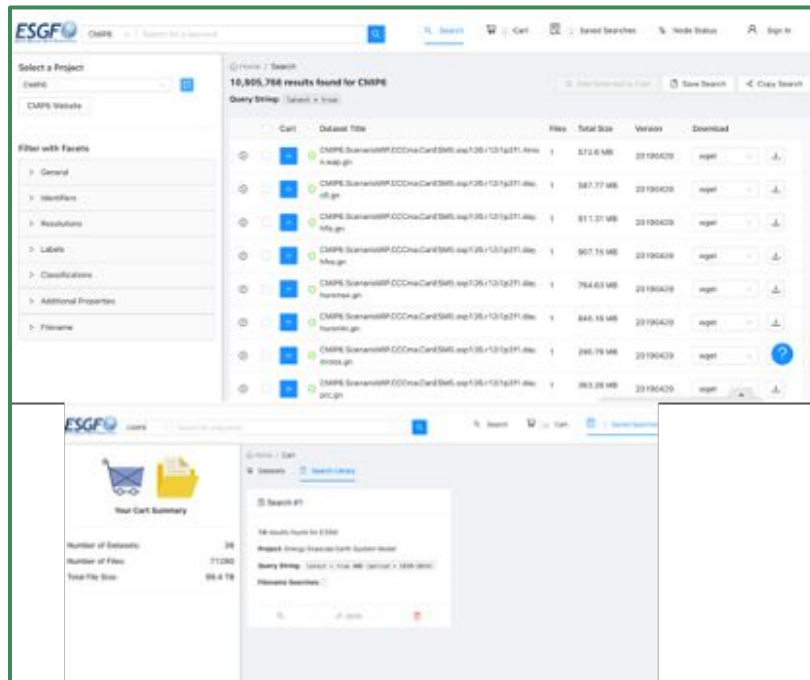
In partnership with the ORNL National Center for Computational Sciences (NCCS)



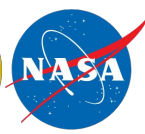
Expandable tape subsystem of the Themis storage system

Metagrid Enhances ESGF Search

- New **Metagrid faceted search user interface**, developed at LLNL on popular React Javascript framework, deployed at ORNL, LLNL and ANL
- Offers new features, including a **shopping cart**, ability to **save and share searches**, integration with **Globus authentication & transfer** and a search page **tour & support dialog**
- User experience enhancements make it **faster and easier** to discover published data
- **Globus integration** offers faster and more reliable data access
- Will be deployed internationally across the Federation by mid-2024



The Metagrid Web Interface for ESGF search is a completely redesigned interface from CoG. It features a familiar faceted search and a new capability to save searches.

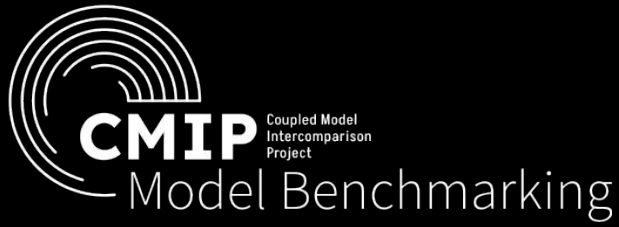


ESGF² Outreach Activities

- Organize **Webinars, Tutorials, and Bootcamps**
 - Data management lessons learned, ingest best practices
 - Data discovery and access, analysis frameworks and tools
- **ESGF Webinar series playlist at <https://www.youtube.com/@esgf2432>**
- **Hackathons and Workshops**
 - Data standards, data node deployment and user compute resources
 - Hold at large relevant conferences, e.g., AGU, EGU, AMS
- **Open ESGF Workshop at AGU 2022 (Chicago)**
- **Open ESGF Workshop & Tutorial at AGU 2023 (San Francisco)**
- Organize / host annual **ESGF Developer and User Conferences**
- **Ninth ESGF Developer and User Dual-Hybrid Conference was held January 18–20, 2023 at ORNL and Toulouse**
- **Tenth ESGF Developer and User Conference scheduled for Rockville, MD, on April 23–26, 2024**

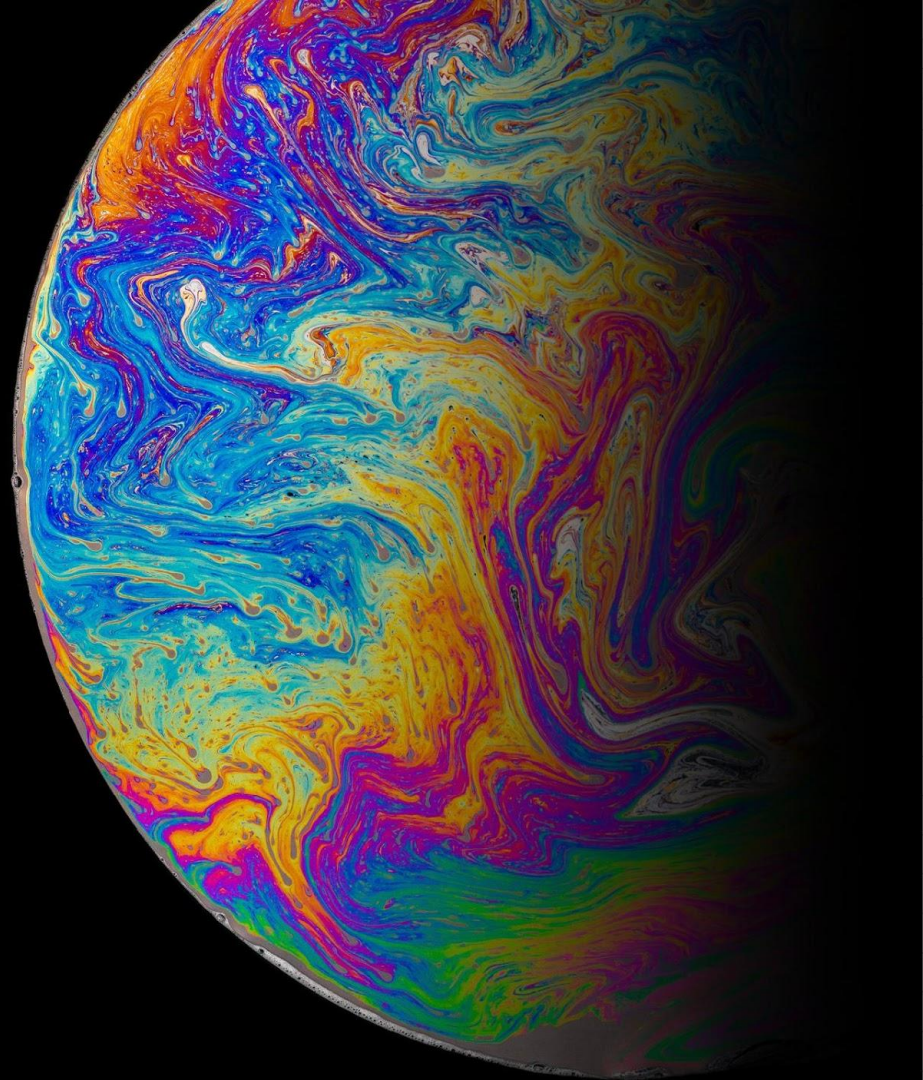


Ninth ESGF Developer and User Conference, held jointly between Oak Ridge National Laboratory (USA) and Toulouse (France), January 18–20, 2023



Climate Model Benchmarking for CMIP7

Forrest M. Hoffman and Birgit Hassler
CMIP Climate Model Benchmarking Task Team Co-Leads



The Model Benchmarking TT



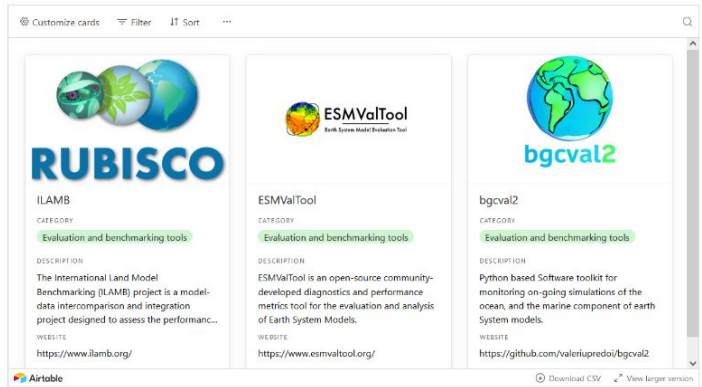
- Rebecca Beadling, *USA*
- Ed Blockley, *UK*
- Jiwoo Lee, *USA*
- Valerio Lembo, *Italy*
- Jared Lewis, *Australia*
- Jianhua Lu, *China*
- Luke Madaus, *USA*
- Elizaveta Malinina, *Canada*
- Brian Medeiros, *USA*
- Wilfried Pokam Mba, *Cameroon*
- Enrico Scoccimarro, *Italy*
- Ranjini Swaminathan, *UK*

- **Diversity** in expertise (realms and methods), user group representation, gender, location, career stage
- **Overarching goals:**
 - Systematic and rapid performance assessment of the expected models participating in CMIP7 (including the model output and documentation)
 - Enhancing existing community evaluation tools that facilitate performance assessment of models
 - Integration of evaluation tools into CMIP publication workflows and fostering publication of their diagnostic outputs alongside the model output on the ESGF
- Collaboration with two **Fresh Eyes on CMIP** Subgroups
 - Model Evaluation
 - Data Analysis

Model Benchmarking Tools – Info “Cards” & Videos

- **Main characteristics** of (open source) benchmarking and evaluation tools available for analyses of CMIP-style data summarized in an overview “card” or an information video
- Collected information **presented centrally** on the CMIP website for easy access
- Cards can be filled out for **all available open source benchmarking and evaluation tools** if they can be used for CMIP data analysis; pre-defined questionnaire available on the CMIP website

<https://wcrp-cmip.org/tools/model-benchmarking-and-evaluation-tools/>



The screenshot shows a web interface with three tool cards. Each card includes a logo, a title, a category, a description, and a website URL.

Tool Name	Category	Description	Website
RUBISCO	Evaluation and benchmarking tools	The International Land Model Benchmarking (ILAMB) project is a model-data intercomparison and integration project designed to assess the performance...	https://www.ilamb.org/
ESMValTool	Evaluation and benchmarking tools	ESMValTool is an open-source community-developed diagnostics and performance metrics tool for the evaluation and analysis of Earth System Models.	https://www.esmvaltool.org/
bgcval2	Evaluation and benchmarking tools	Python based Software toolkit for monitoring on-going simulations of the ocean, and the marine component of earth System models.	https://github.com/valeriupredoii/bgcval2

Status: first cards available

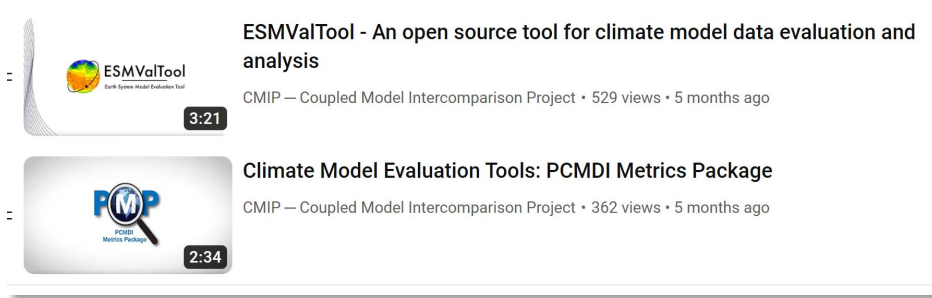
Started: October 2023



#CMIP

Model Benchmarking Tools – Information Videos

- **Videos** with descriptions of different benchmarking and evaluation tools
- Contain also the **main characteristics** of the different tools, just presented in a different way than the “cards”
- Videos can also be of **different style**
- All videos are presented in **one central location** linked to CMIP



- ▣ More **videos** of tools welcome!
- ▣ More **info cards** about evaluation/benchmarking tools welcome!

<https://wcrp-cmip.org/tools/model-benchmarking-and-evaluation-tools/>

#CMIP

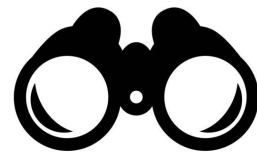
Retrospective paper



- **Definitions** of “evaluation”, “validation”, and “benchmarking”
- Retrospective look at **evolution of evaluation & benchmarking metrics**
- What tools were **available** for CMIP6 (methods, philosophies, tools)?
- What approaches were **used** for CMIP6?
- Which of them **worked well** for CMIP6 and what **did not work** for CMIP6?
- **Extensive information** about different benchmarking and evaluation tools

Status: Currently being finalized

Planned submission: May 2024



What is the way forward?

- Based on the findings of the extensive information collected about different tools, and the retrospective paper – What do we think should be the **benchmarking/evaluation focus for CMIP7?**
- What **framework** would ideally be available for instantaneous benchmarking and evaluation at the time of data submission? Is such a framework even possible?
- How to **avoid the bottlenecks** encountered in CMIP6 benchmarking/evaluation?
- Comprehensive community evaluation in **near-real time** possible (i.e., on ESGF)?

Status: Under development

Planned submission: Summer 2024

#CMIP

Other Planned TT Activities

- CMIP Model Benchmarking TT **face-to-face meeting in May** at the German Aerospace Center (DLR) to finish off the retrospective paper and establish vision for the perspectives paper
- In collaboration with **Fresh Eyes on CMIP** groups
 - Develop scope for better quality assurance / quality control (QA/QC) for CMIP model output
 - Develop a white paper on data needs for model benchmarking, including uncertainties
 - Develop a final report from the Model Benchmarking TT that proposes an architecture for automated benchmarking capabilities at the time of CMIP data publication



#CMIP

PCMDI Metrics Package



Near-Surface Air Temperature: MAM

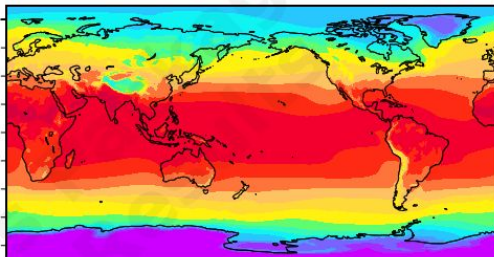
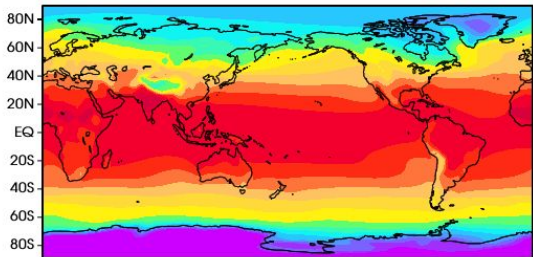
(CMIP6 HISTORICAL: 1981-2005, global)

CESM2

REF (ERA-5)

Mean 14.53 Max 33.4713 Min -57.6086

Mean 14.46 Max 34.7196 Min -59.5538

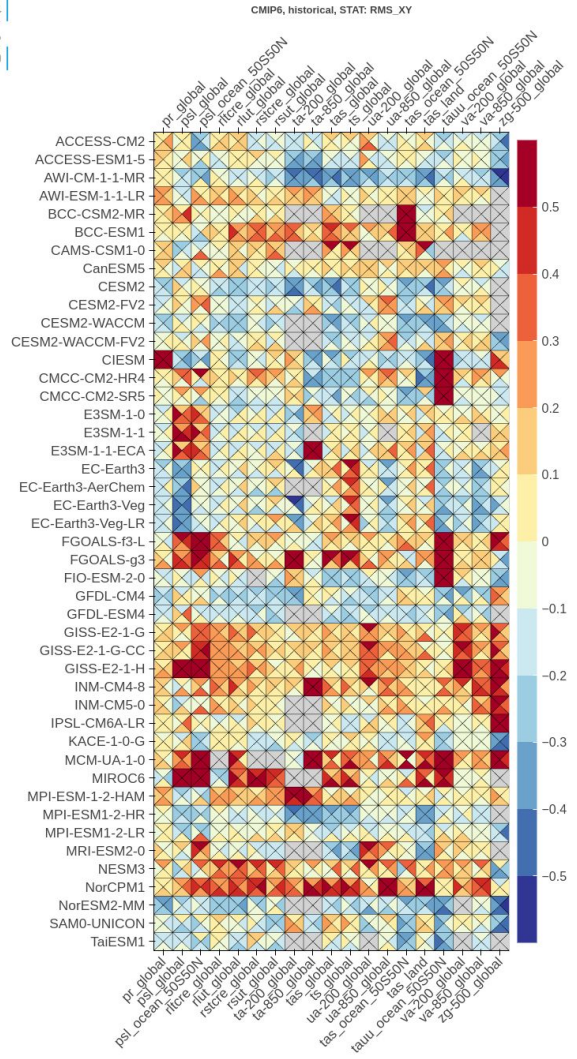
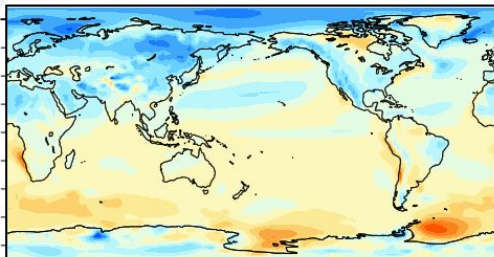
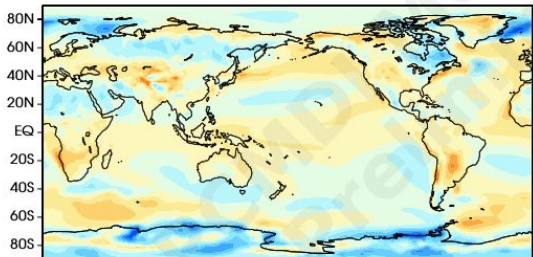


CESM2 - REF (ERA-5)

Mean 0.05646 Max 5.72967 Min -7.72577

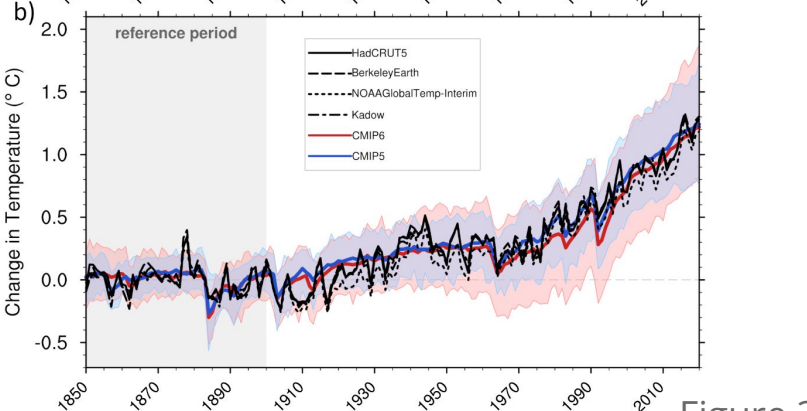
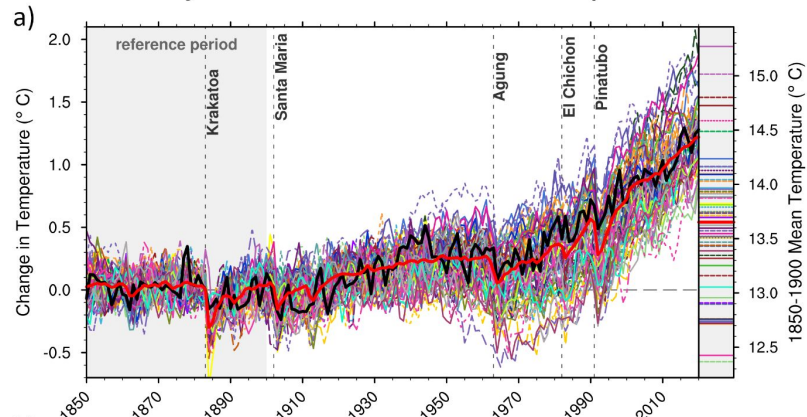
Multi Model Mean - REF (ERA-5)

Mean -0.2551 Max 5.72786 Min -7.23238



ESMValTool

Anomaly of Near-Surface Air Temperature



- HadCRUT5
- MultiModelMean
- ACCESS-CM2
- ACCESS-ESM1-5
- AWI-CM-1-1-MR
- AWI-ESM-1-1-LR
- BCC-ESM2-MR
- BCC-ESM1-0
- CAMS-CSM1-0
- CanESM5
- CanESM5-CanOE
- CESM2
- CESM2-FV2
- CESM2-WACCM
- CESM2-WACCM-FV2
- CIEM
- CMCC-CM2-HR4
- CMCC-CM2-SR5
- CMCC-ESM2
- CNRM-CM6-1
- CNRM-CM6-1-HR
- CNRM-ESM2-1
- CSIR-Mk3-0
- CSIR-Mk3-1
- CSIR-Mk3-1-ECA
- EC-Earth3
- EC-Earth3-AerChem
- EC-Earth3-CC
- EC-Earth3-Veg
- EC-Earth3-Veg-LR
- FGOALS-g3-L
- FGOALS-g3
- FIO-ESM-2-0
- GFDL-CM4
- GFDL-ESM4
- GISS-E2-1-G
- GISS-E2-1-G-CC
- GISS-E2-1-H
- HadGEM3-GC31-LL
- HadGEM3-GC31-MM
- ITM-ESM
- INM-CM4-8
- INM-CM5-0
- IPSL-CM5A2-INCA
- IPSL-CM6A-LR
- KACE-1-0-G
- KIOST-ESM
- MCM-UA-1-0
- MIROC-ES2L
- MIROC6
- MPI-ESM1-2-HAM
- MPI-ESM1-2-HR
- MPI-ESM1-2-LR
- MPI-ESM2-0
- NESM3
- NorCPM1
- NorESM2-LM
- NorESM2-MM
- SAM0-UNICON
- TaiESM1
- UKESM1-0-LL

Lauer et al., 2002

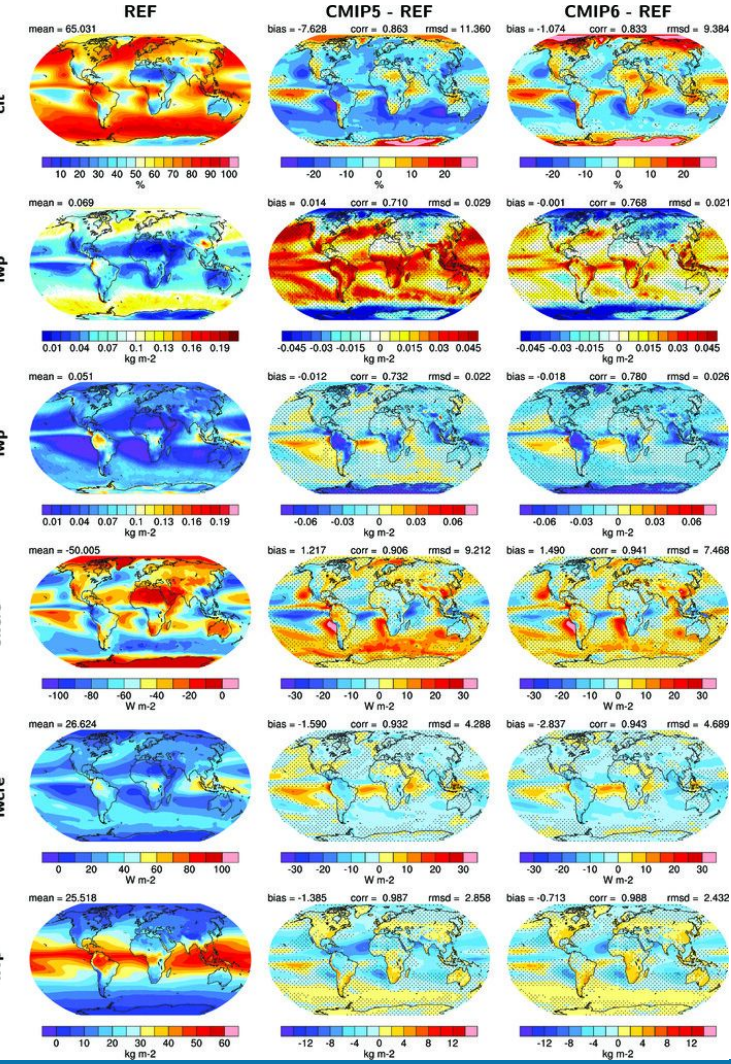


Figure 3.4, IPCC AR6 WG1

Evaluation of CMIP5 vs CMIP6 with ILAMB and IOMB

- (a) ILAMB and (b) IOMB have been used to evaluate how land and ocean model performance has changed from CMIP5 to CMIP6
- Model fidelity is assessed through comparison of historical simulations with a wide variety of contemporary observational datasets
- The UN's Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) from Working Group 1 (WG1) Chapter 5 contains the full ILAMB/IOMB evaluation as Figure 5.22

(a) Land Benchmarking Results

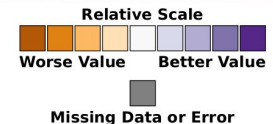
Land Ecosystem & Carbon Cycle

	bcc-csm1-1	CanESM2	CanESM1-BGC	GFDL-ESM2G	IPSL-CM5A-LR	MIROC-ESM	MPI-ESM-LR	NorESM1-ME	HadGEM2-ES	BCC-CSM2-MR	CanESM5	CESM2	GFDL-ESM4	IPSL-CM6A-LR	MIROC-ES2L	MPI-ESM1-2-LR	NorESM2-LM	UKESM1-0-LL	Mean CMIP5	Mean CMIP6
Biomass	-0.72	-0.93	-1.95	-1.51	-0.13	0.60	-0.43	-1.31	0.19	-0.43	0.66	0.48	-1.09	0.22	0.60	-0.07	1.00	0.49	1.63	-2.30
Burned Area	0.20	-0.45	-1.52	-0.40	-1.26	-0.26	-1.07	-1.77	0.92	1.39	0.74	-0.20	-0.54	0.16	0.93	-0.96	-0.01	1.04	1.23	1.82
Leaf Area Index	-0.20	-0.64	-1.30	-2.53	-0.01	0.30	0.01	-1.85	-0.16	0.27	0.08	0.34	-0.70	1.19	0.82	0.46	0.37	0.69	1.04	1.61
Soil Carbon	0.27	1.26	1.46	0.07	0.75	0.47	-0.03	-1.14	0.07	0.23	1.35	-0.99	-2.04	-1.55	0.90	-0.75	-0.17	0.24	1.01	1.48
Gross Primary Productivity	0.59	-1.23	0.01	-1.81	-1.40	0.29	-0.53	-0.24	-1.04	0.77	0.04	0.59	-0.38	1.17	-1.02	-0.37	0.73	0.09	1.51	2.22
Net Ecosystem Exchange	-0.42	1.81	-0.21	-0.65	1.10	-0.24	0.80	0.02	-1.03	-1.02	-1.19	0.59	1.69	-0.42	0.63	-0.21	1.08	-1.43	1.28	1.43
Ecosystem Respiration	0.90	-0.56	-0.86	-0.24	-1.35	0.99	-0.01	-0.94	-1.54	0.81	0.59	0.51	-0.79	0.90	-0.21	-1.24	0.43	-0.94	1.34	2.21
Carbon Dioxide	-1.54	-0.36	-2.92	-0.74	-1.53	-0.00	0.37	0.85		0.42	-0.26	0.39	0.59	1.10	-0.87	0.21	0.69	0.09	-0.07	
Global Net Carbon Balance	-1.64	-0.88	-1.13	0.17	-0.31	-0.38	-0.50	0.24		-0.23	1.34	-1.70	0.17	-0.74	1.45	1.56	0.26	0.92	1.40	
Land Hydrology Cycle	-2.65	-0.42	0.44	-0.18	-0.49	-0.52	-0.57	0.17	0.70	0.15	-0.47	1.51	-1.24	0.58	-0.72	-0.83	0.97	0.87	1.00	1.70
Evapotranspiration	-0.82	-0.99	-0.27	-1.02	0.64	-1.14	-0.62	-0.60	0.28	0.39	-1.08	1.09	0.65	0.43	-1.40	-1.01	0.82	1.05	1.41	2.20
Evaporative Fraction	-0.34	0.74	0.74	-0.14	-0.85	0.21	-1.98	0.22	-0.34	0.10	0.11	1.25	-0.88	1.29	-1.65	-1.81	1.11	-0.06	0.98	1.29
...																				
Terrestrial Water Storage Anomaly	-2.79	-0.45	0.47	0.50	-0.38	0.34	0.35	0.43	0.58	0.15	-0.08	0.95	-2.91	0.43	0.37	0.15	0.39	0.51	0.49	0.50
Permafrost	-0.88	-2.26	0.01	0.13	0.83	0.69	0.56	0.69	-0.56	-0.11	-3.02	0.83	0.74	-0.18	0.49	0.42	0.89	0.43	0.06	0.23

(b) Ocean Benchmarking Results

Ocean Ecosystems

Chlorophyll			2.18	0.20	-0.20		0.04		0.22		-0.37	0.83	-0.37	-0.26	-0.91	-0.67	1.93	0.27	0.30	0.67
Oxygen, surface	-3.50	2.13	0.44	1.02		0.49		0.56		-0.67	0.88	-0.21	0.10	-1.02	-0.41	-2.19	0.18	0.13	0.34	
Ocean Nutrients			0.73	-0.13	-1.98		-0.53	-1.53	-0.29		0.73	0.34	-0.09	-0.41	0.35	-0.30	0.40	0.49	0.64	1.57
Nitrate, surface			-0.84	-0.10	0.91		-0.80	-1.25			-0.02	1.00	1.68		-0.90	-1.14	-0.17	-0.16	1.60	
Phosphate, surface	0.21	-1.63	0.67	1.22		-0.18	-1.70	0.82		1.21	-0.90	0.29	1.21	1.02	0.39	-1.78	-0.56	-0.47	0.18	
Silicate, surface			-0.69	-0.04	0.04		-0.45	-0.43		0.39	-0.14	0.17	-0.41	-0.98	0.00	0.02	0.88	1.63		
Ocean Carbon	0.44	-0.71	0.24			-0.81	-0.20	-2.16		0.50	1.24	1.60		-1.21	-0.19	0.18	0.29	1.37		
TAlk, surface										1.24	-0.23	-0.62	-0.69	-1.08	-1.12	1.31				1.19
Salinity, 700m	-0.27	1.01	0.12	0.19		0.32	-2.31	-0.22		0.06	-0.36	0.85	-0.42	0.29	2.48	1.27	0.06	1.27	0.54	
Ocean Relationships	0.44	-0.35	-1.06	-0.54	0.70	0.46	-0.46	-0.80	0.32	0.36	0.25	-1.16	-0.47	0.54	0.33	-0.39	-0.87	-0.54	1.58	1.64
Oxygen, surface/WOA2018			-1.86	-0.36	-0.29		1.50	-0.43	0.68		-0.02	0.72	1.20	0.17	1.86	0.02		-1.12	0.39	1.25
Nitrate, surface/WOA2018	0.27	0.23	-0.63			-0.26	-0.12	-0.38		0.29	-0.21	0.19	0.18	0.14	-0.07			0.03	-0.23	0.53



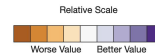
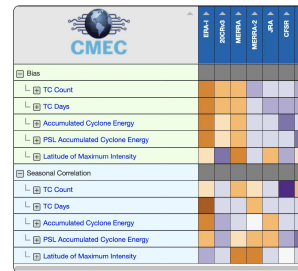
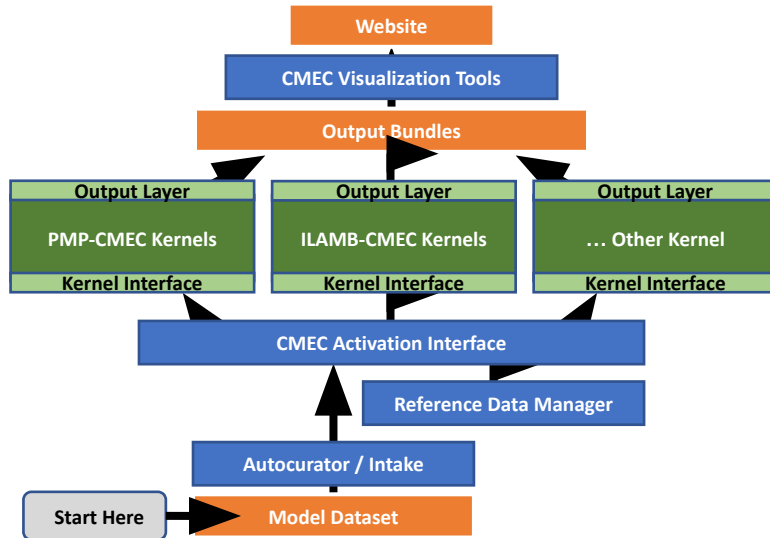


Coordinated Model Evaluation Capabilities

The **CMEC Project** is developing uniform standards and software tools that allow multiple metrics and diagnostics packages to be executed from a unified interface, and results explored in a unified manner.

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 **interactive analysis** of metrics and diagnostics package output.
3. Build **connections and standards across projects and agencies** related to model evaluation (e.g., MDTF).



CMEC supports a comprehensive software suite for analysis of model evaluation output, including both climate data metrics and diagnostics.



Contact Paul Ullrich
ullrich4@lnl.gov