Data Stewardship from Science to Society

Aparna Radhakrishnan

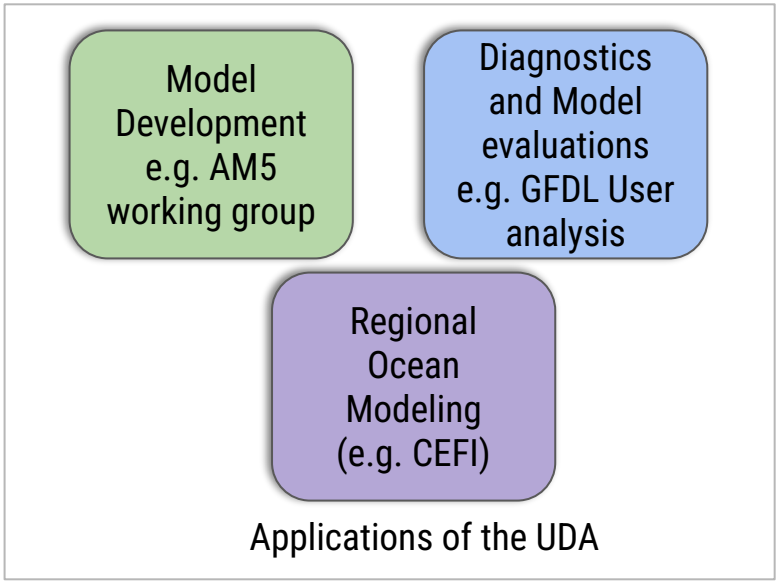
Chandin Wilson, Chris Blanton, Ciheim Brown, Ian Laflotte, Hans Vahlenkamp, Jacob Mims, Jessica Liptak, John Krasting, Kirsten Findell, V. Ramaswamy, Kristen Schepel, Kristopher Rand, Lauren Koellermeier, Mike Skowvron, Utheri Wagura, Wenhao Dong, Whit Anderson, Yi-Hung Kuo, and many others.

Q3: How can GFDL research and modeling be further utilized to meet NOAA stakeholder needs and enhance research partnerships to ensure GFDL's success?

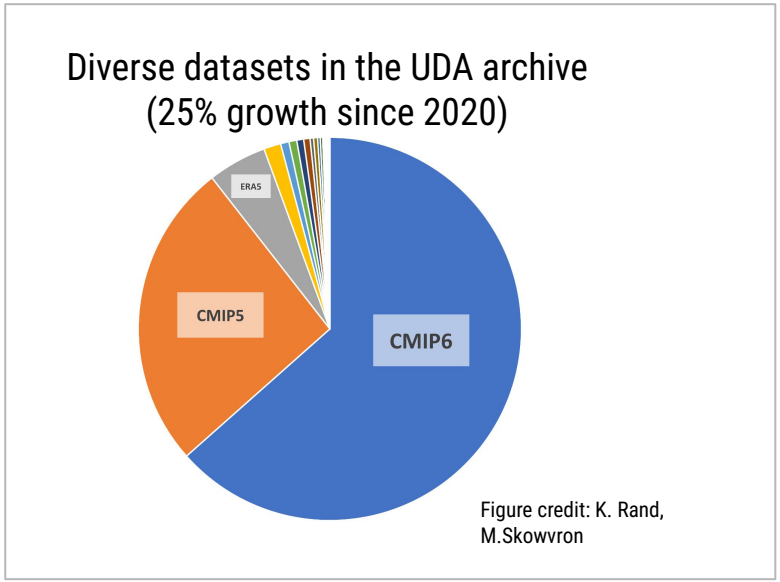
OAR strategic goals and the role of Data Stewardship

EXPLORE THE MARINE ENVIRONMENT	DETECT CHANGES IN THE OCEAN AND ATMOSPHERE	MAKE FORECASTS BETTER	DRIVE INNOVATIVE SCIENCE
Increase knowledge of the oceans, coastal areas, and Great Lakes to support resource management and **public awareness.**	Produce, analyze, and interpret observation records to understand the Earth System and **inform the public.**	Improve accuracy, precision, and efficiency of forecasts and predictions to **save lives** and property and support a vibrant **economy.**	Cultivate and deliver mission-relevant research to **lead** the environmental science **community**

Strengthening GFDL research through the Unified Data Archive

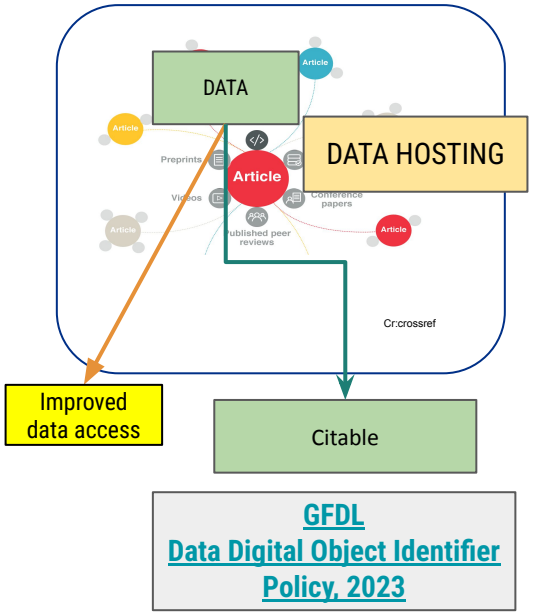


Applications of the UDA



GFDL manages a Unified Data Archive (UDA) with user-requested external data to strengthen GFDL research.

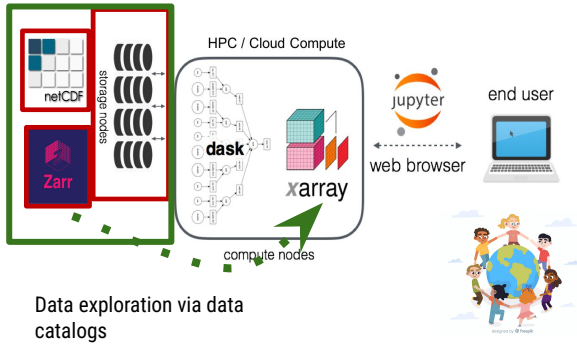
NOAA GOAL#1.4: Bolster Authoritative data and information stewardship



- Collaborations with Princeton University Library, NOAA and OAR Data Management Working Group, LEGEND Task Team.
- Ground-breaking progress to make GFDL data citable and accessible through GFDL data DOI policy, Globus, leveraging collaborative work.
- Process and planning in place to adhere to NOAA's Public Access to Research Results (PARR)

Increase ability to access and use Earth system data

- Data access: Publication efforts for CMIP (400+ data citations), TOAR-II, QBOi, etc.
- Data usability: Cloud-optimized data publication and data analysis through collaborations with Princeton, AWS, Pangeo and ESGF, NOAA NODD.
- Discoverability and interoperability: Community-driven Data exploration utilities for seamless analysis



Exploring cloud infrastructure to maximize accessibility, transparency, usability (NOAA GOAL #1.4)

Related slides in [Question 3] V.Naik, National and International Assessments and Multimodel Intercomparison Projects and
J.Durachta et al Modeling Systems Division [Question 1]

NOAA's Model Diagnostics Task Force, funded by NOAA MAPP

Connect model development efforts with expertise in the academic and private sectors to evaluate and improve model performance and process representation

NOAA Model Diagnostics Task Force

Meet the MDTF team

led by JD Neelin, UCLA

What We Have Done

Established community of >90 diagnosticians

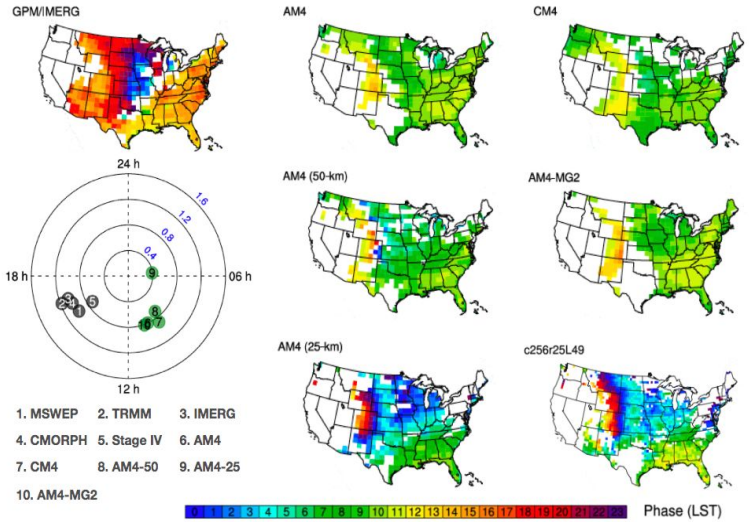
- Built a Python-based framework
- Collaboration with DOE on Earth System Metrics and Diagnostics Standards
- Prepared extensive documentation & tutorials
- Moved to an open-development environment
- Data processor redesign for an agile framework
- Ability to do multi-model comparisons
- Robust Testing

Where We are Going

- Expand to more Earth system components
- Explore more weather-scale phenomena
- Process-oriented diagnostics for AI emulators
- Model workflow integration
- Support for Jupyter notebook based diagnostics
- Align with GFDL/NOAA objectives while collaborating with other community partners

MAPP Modeling, Analysis, Predictions, and Projections | UCLA | GFDL | UCAR | SAIC | Princeton University | NCAR | Florida State University | University of Washington | UCDAVIS University of California | U.S. Department of Energy

MDTF Process-Oriented Diagnostics (POD) Examples

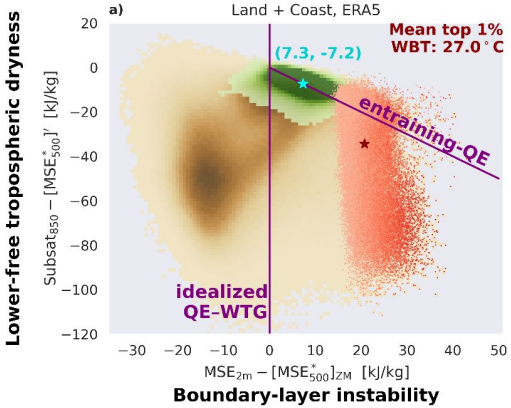


Dong, Guo, Krasting 2023 J Climate
c256r25L49 from Lin et al. 2024 JAMES

- The application of Diurnal cycle of precipitation POD to a series of GFDL AM4 models provides valuable insights to guide the development of the AM5 model.
- Diurnal cycle of precipitation *are crucial for* detecting convective storms, driving the need to represent them better to understand and mitigate societal and environmental impacts.

MDTF POD Examples (contd..)

Observational reanalysis

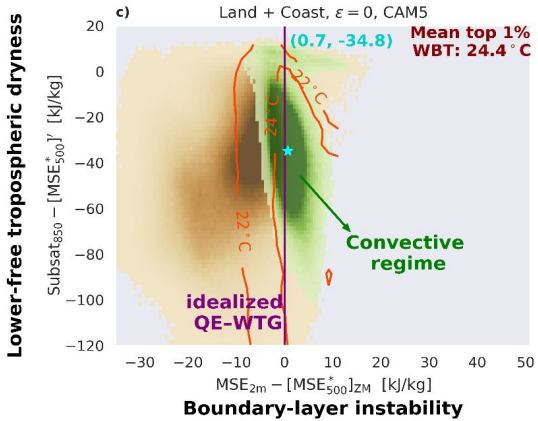


- **Moist heatwaves** place significant stress on human physiology and pose substantial risks to health and society.
- Improved representation of moist heat waves in our models can foster stronger scientific and societal connections.

Entrainment of unsaturated air in the lower troposphere increases the limit for moist-heatwave magnitude posed by deep convection and permits a larger number of extreme moist heat cases to occur.

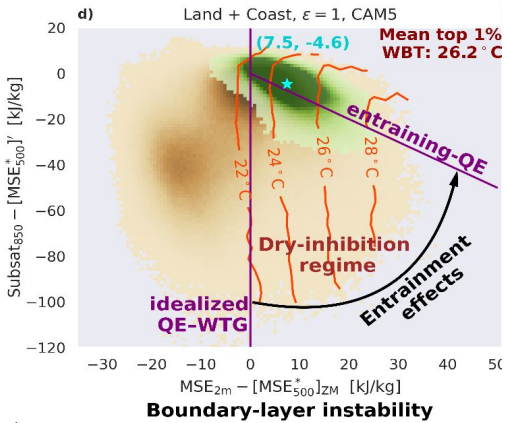
Duan, Ahmed, and Neelin, 2024, *Nature Geosci.*

Model simulation without entrainment



⇨ **cannot produce intense moist heatwaves**

Model simulation with entrainment



⇨ **reproduces intense moist heatwaves**

Opportunities: Why we do what we do?

GFDL data science group aims to facilitate the application of Earth system model data for NOAA's service and stewardship, by strategically identifying pathways and building information-centric processes involving public engagement.

Science

Service

Continuous user engagement

Sustainable tools and services to help plan for future

NOAA Technical Readiness Level 7

Delivering Knowledge for a prosperous society

GFDL SPEAR

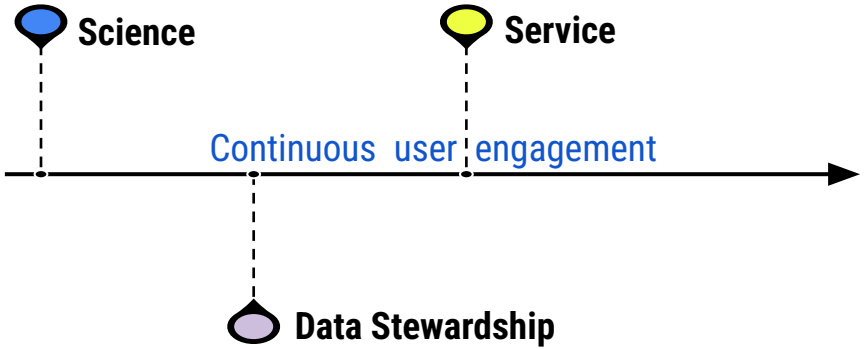
Data Stewardship

FAIR and AI-ready

Seasonal-to-Decadal Earth system projection datasets utilizing GFDL SPEAR

Collaborators: NCEI (lead), GFDL, PSL, CIs, NOAA Climate Adaptation Partnerships / Regional Integrated Science Assessments

Challenges: Expanding Horizons with continuity



Sustainable tools and services to help plan for future