OAK RIDGE National Laboratory

Katherine Evans

Director, Computational Sciences and Engineering

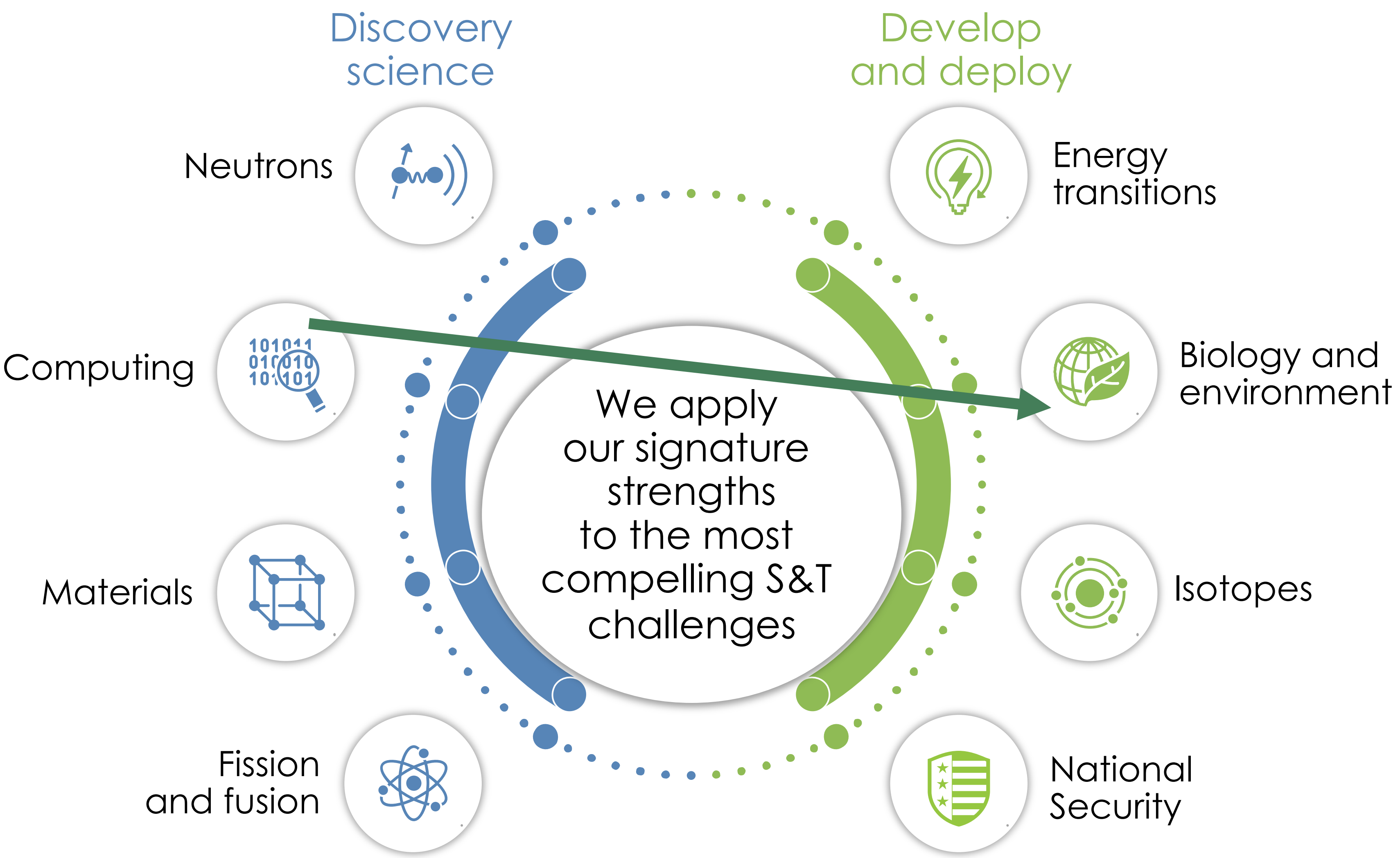
Mission

Transdisciplinary computational science and analytics at scale to enable scientific discovery

- HPC Modeling/simulation
- Data analytics at scale
- Quantum Information Sciences

Dan Lu, Prasanna Balaprakash, Forrest Hoffman, and Matt Norman at ORNL contributed content for this talk

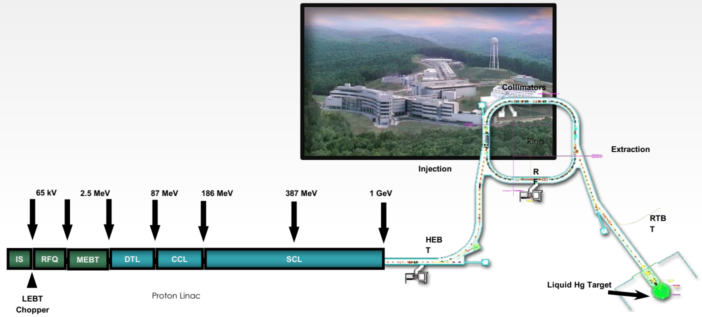
ORNL delivers translational research for national priorities

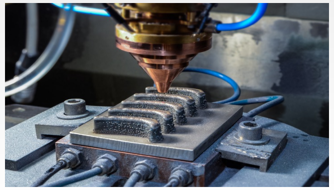


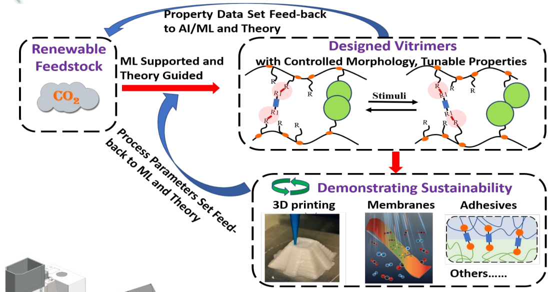
Overview of ORNL’s AI initiative

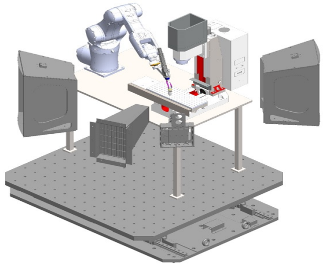
ORNL invests in

- The science and engineering application of AI
- Fundamental AI/ML in support of the science and engineering mission









Strategic focus for foundational research

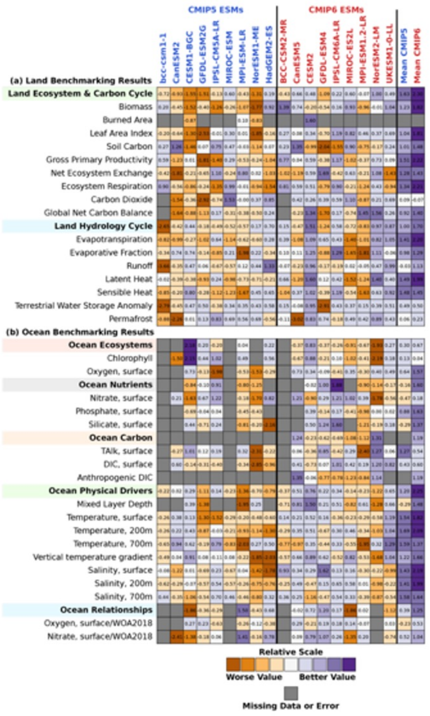
- AI for Robust Engineering and Science (AIRES)
  - Digital twin, digital thread
  - Complex networked systems
  - Control
  - Prognostics
- AI for Scientific Discovery and Design (AISD2)
  - ML surrogate models for multiscale systems and processes
  - AI-based optimization and system design
  - Causal analysis and design of experiment
- Assurance
  - Uncertainty quantification
  - Verification and validation
  - Explainability and interpretability
- Scalable AI

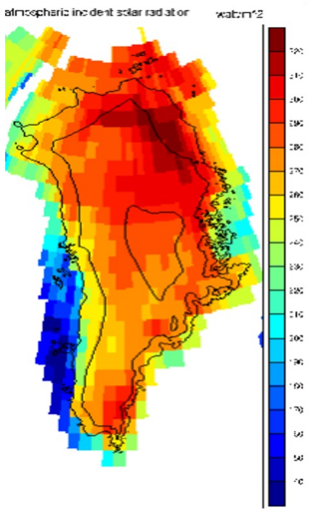
Computational Earth Sciences @ ORNL

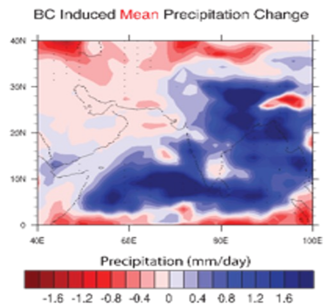
- Improves process understanding of the global Earth system by developing and applying models,
- Quantifies interactions, feedbacks and uncertainty within and between the Earth system and its cycles
- Develops and applies methods and tools, including AI and machine learning, for quantitative assessment and benchmarking of coupled, multiscale Earth system models
- Provides metrics that connect to integrated and vulnerability assessment and adaptation projects

Senior staff: Forrest Hoffman, Moet Ashfaq, Salil Mahajan, and Dan Lu. Funding from DOE ASCR and BER, NOAA and US Air Force as part of an HPC collaboration

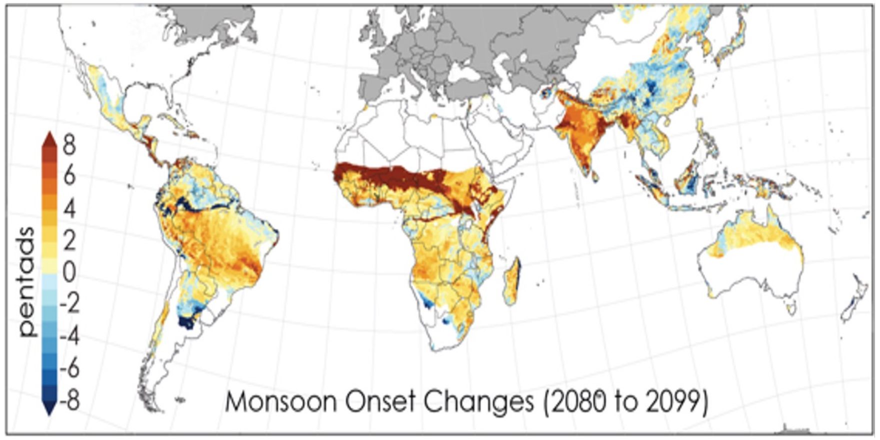
(PCC AR6 WG1, 2021, Figure 5.22







(Batibeniz et al., *Earth's Future*, 2020)

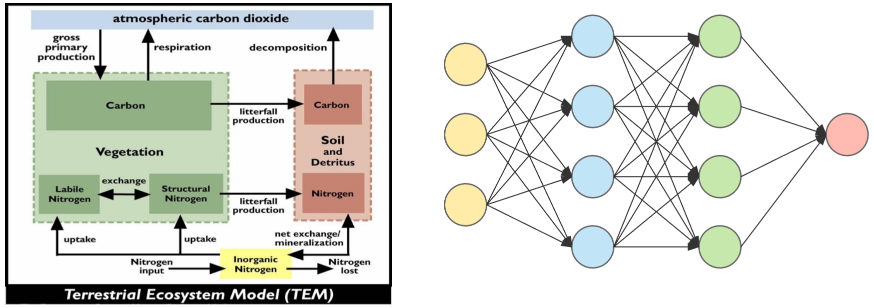


(Ashfaq et al., *Clim. Dyn.*, 2020)

Interpretable and reliable ML methods to advance Earth system predictability with uncertainty quantification

Surrogate modeling

Use ML to build a fast surrogate of the expensive physics-based model

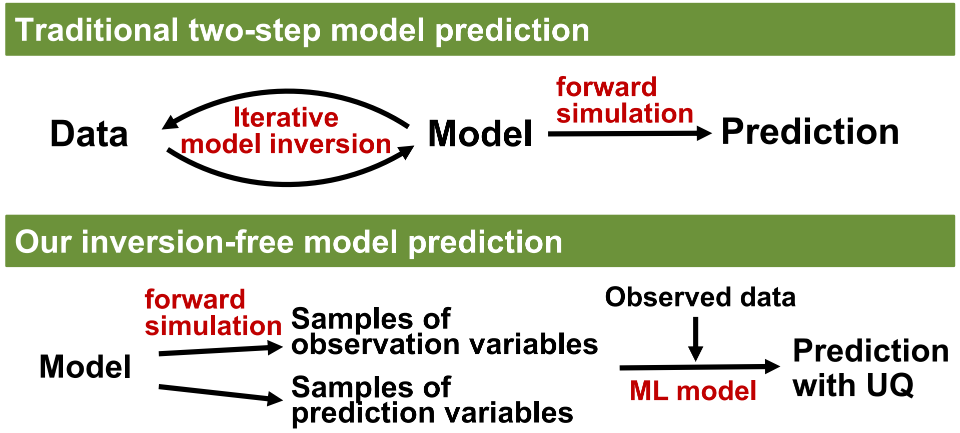


Neural network ML model as a surrogate of the terrestrial ecosystem model

- Surrogate modeling builds a fast surrogate of the physics-based model and replace the expensive model in prediction and UQ.
- It reduces time of a single model run and thus the total computational cost of the ensemble simulations.

Inversion-free prediction

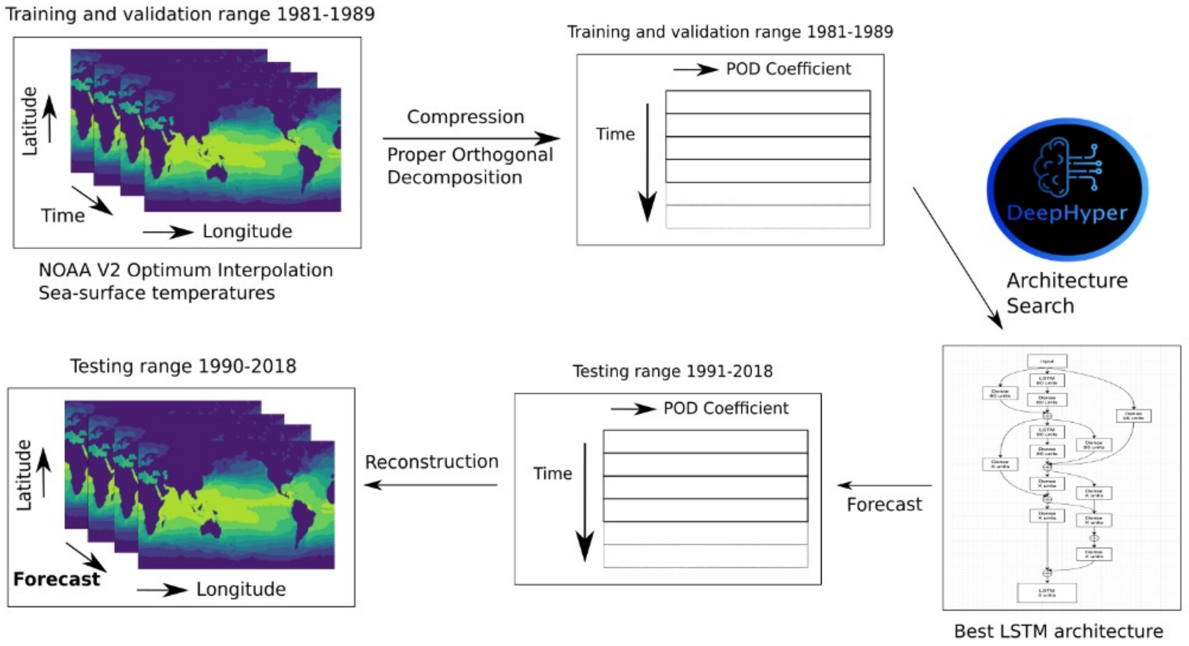
Use ML to learn observation-prediction relationship and make predictions directly



- This method avoids computationally expensive, iterative inverse modeling.
- It infers prediction directly by learning observation-prediction relationship.
- We use ML model to learn the obs-pred relationship from their simulation samples.

Scalable and optimized neural networks reduce computational and development costs for simulation of sea surface temperatures

- Use distributed neural architecture search to construct a purely data-driven surrogate model for SST forecasting, trained using publicly available NOAA SST data
- Process-based models for SST are costly. Using data-driven emulators reduces cost and can be automatically discovered by scalable search methods on leadership class systems.



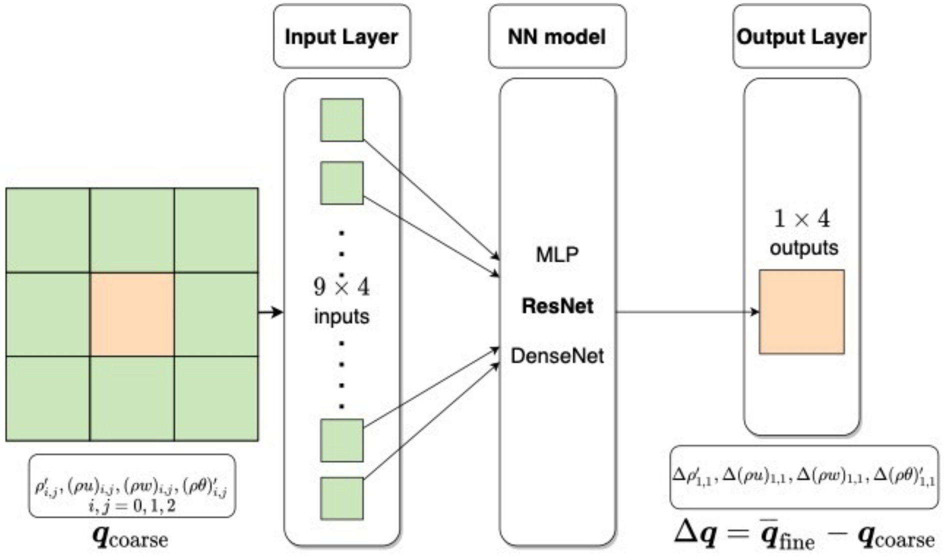
A schematic of the surrogate model search with DeepHyper

NOAA dataset https://www.ncdc.noaa.gov/oisst/optimum-interpolation-sea-surface-temperature-oisst-v20

Courtesy: Prasanna Balaprakash

Use ML to reduce computational costs in Earth system model prediction. Example: cost effective turbulence closure

- Goal: capture high resolution dynamics within lower resolution models.
- Fine grained neural networks emulate physics that are difficult to capture with traditional PDEs
- Solution: train NN on differenced hi-res and low res model output, then apply
- Very low latency in-kernel deployment on GPUs at scale with a portable C++ library
- Challenge: Ensuring ML outputs retain physicality with inputs far from the training data



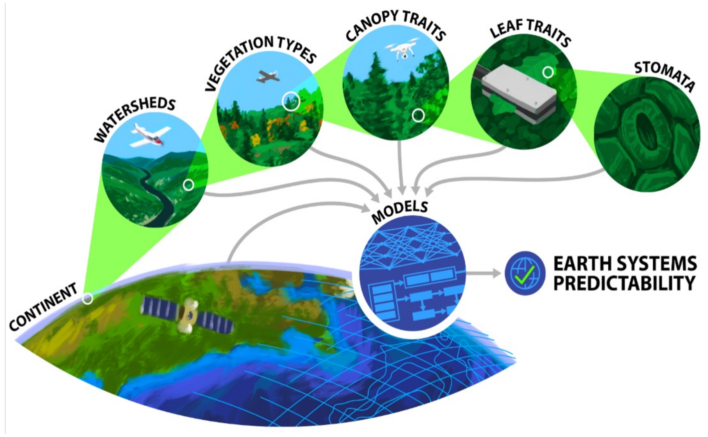
Using a stencil of fluid data to predict sub grid scale turbulent dynamics directly from high resolution data

Note: Physics-informed is not enough. Rigid constraints on model outputs are needed.

Courtesy: Matt Norman, Oak Ridge Leadership Computing Facility

Augment science impact using AI for Earth sciences: Workflows and science efficiency

- **Autonomous discovery**; identify areas of interest or concern
- **Connected instruments and data**: ML used to optimize networks and data transfer, processing, detecting issues
- **Assisted software generation**: off the shelf products as a first draft
- **Literature search, hypothesis generation:** natural language processing tools, making new connections



ML potential is high for improving predictability when (1) *sufficient data are available for process representations* and (2) *process representations are computationally expensive*
Many Earth systems involve unresolved, subgrid-scale processes that strongly influence results at the largest scales

https://ai4esp.org/ https://ai4esp.slack.com/

AI4ESP | Artificial Intelligence for Earth System Predictability

A multi-lab initiative working with the Earth and Environmental Systems Science Division (EESSD) of the Office of Biological and Environmental Research (BER) to develop a new paradigm for Earth system predictability focused on enabling artificial intelligence across field, lab, modeling, and analysis activities.

White papers solicited for development and application of AI methods in areas with an emphasis on quantifying and improving Earth system predictability, particularly related to the integrative water cycle and extreme events.

How can DOE directly leverage artificial intelligence (AI) to engineer a substantial (paradigm-changing) improvement in Earth System Predictability?

156 white papers received and read for the **AI4ESP Workshop on Oct 25–Dec 3, 2021**

Earth System Predictability Sessions

- Atmospheric Modeling
- Land Modeling
- Human Systems & Dynamics
- Hydrology
- Watershed Science
- Ecohydrology
- Aerosols & Clouds
- Climate Variability & Extremes
- Coastal Dynamics, Oceans & Ice

Cross-Cut Sessions

- Data Acquisition
- Neural Networks
- Surrogate models and emulators
- Knowledge-Informed Machine Learning
- Hybrid Modeling
- Explainable/Interpretable/Trustworthy AI
- Knowledge Discovery & Statistical Learning
- AI Architectures and Co-design

Workshop Report

- Posted on ai4esp.org
- Executive Summary
- Long summary
- Earth science chapters
- Computational science chapters

AMS Special Collection

- Open submissions for new AI for the Earth Systems journal