ESS Cyberinfrastructure Working Groups

Common tools and approaches for constraining models, quantifying uncertainty, and performing big (and little!) data analytics supporting MoDEx:

Working Group on Model-Data Integration Co-Leads

Xingyuan Chen (PNNL)

Forrest Hoffman (ORNL)

- Modeling workflows, model–data comparison, model evaluation and benchmarking
- Uncertainty quantification and data assimilation
- Management and fusion of model output and observational data (Data Management)
- Geospatial and remote sensing data analysis
- Data analytics methods and techniques, e.g., data mining, machine learning, genetic algorithms, visual analytics
- Hybrid ML/process-based models & data-driven models (Software Engineering & Integration)

EESSD's Model-Data-Experiment Enterprise

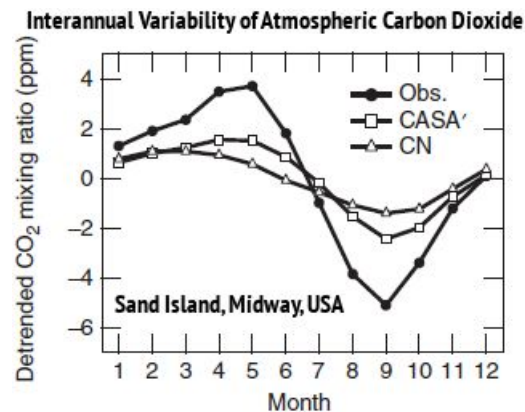U.S. DEPARTMENT OF ENERGY | Office of Science
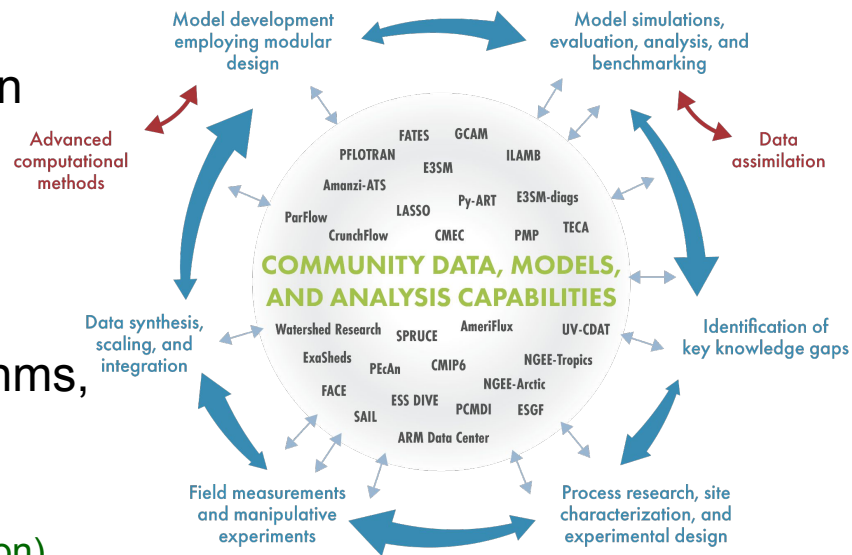
Virtual Hackathon for E3SM Land Model (ELM) Watershed Simulation

- GitHub Link:
  https://github.com/doe-ess-community/elm_hackathon_20230413
- Video Recording: http://www.youtube.com/watch?v=WgkX2QWKBF4

What is a Benchmark?

Interannual Variability of Atmospheric Carbon Dioxide

Sand Island, Midway, USA

Models often fail to capture the amplitude of the seasonal cycle of atmospheric CO_2

- A **benchmark** is a quantitative test of model function achieved through comparison of model results with observational data
- Acceptable performance on a benchmark **is a necessary but not sufficient condition** for a fully functioning model
- **Functional relationship benchmarks** offer tests of model responses to forcings and yield insights into ecosystem processes
- Effective benchmarks must draw upon **a broad set of independent observations** to evaluate model performance at multiple scales

(Randerson et al., 2009)

Models may reproduce correct responses over only a limited range of forcing variables

Argonne NATIONAL LABORATORY   BROOKHAVEN NATIONAL LABORATORY   BERKELEY LAB   Los Alamos NATIONAL LABORATORY   NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH   OAK RIDGE National Laboratory   UCI   MICHIGAN

What is ILAMB?

A community coordination activity created to:

- **Develop internationally accepted benchmarks** for land model performance by drawing upon collaborative expertise
- **Promote the use of these benchmarks** for model intercomparison
- **Strengthen linkages between experimental, remote sensing, and Earth system modeling communities** in the design of new model tests and new measurement programs
- **Support the design and development of open source benchmarking tools**
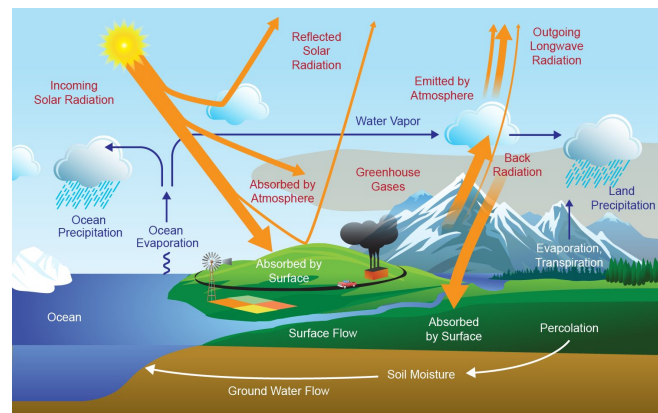
Energy and Water Cycles
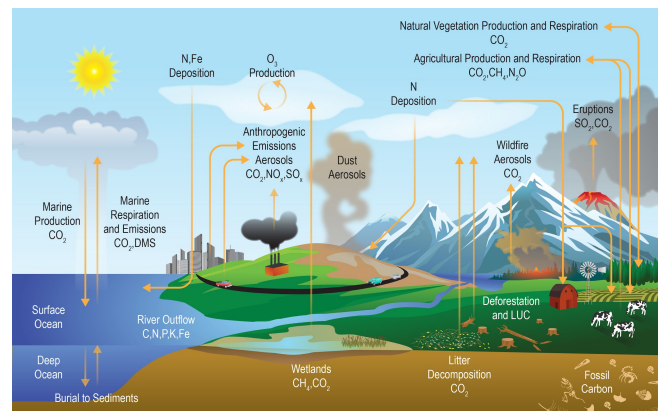
Carbon and Biogeochemical Cycles

ILAMB Produces Diagnostics and Scores Models

- ILAMB generates a top-level **portrait plot** of models scores
- For every variable and dataset, ILAMB can automatically produce
  - **Tables** containing individual metrics and metric scores (when relevant to the data), including
    - Benchmark and model **period mean**
    - **Bias** and **bias score** (S_{bias})
    - **Root-mean-square error (RMSE)** and **RMSE score** (S_{rmse})
    - **Phase shift** and **seasonal cycle score** (S_{phase})
    - **Interannual coefficient of variation** and **IAV score** (S_{iav})
    - **Spatial distribution score** (S_{dist})
    - **Overall score** ($S_{overall}$) $\longrightarrow$ $$S_{overall} = \frac{S_{bias} + 2S_{rmse} + S_{phase} + S_{iav} + S_{dist}}{1 + 2 + 1 + 1 + 1}$$
  - **Graphical diagnostics**
    - Spatial contour maps
    - Time series line plots
    - Spatial Taylor diagrams (Taylor, 2001)
- Similar **tables** and **graphical diagnostics** for functional relationships

ILAMBv2.6 Package Current Variables

- **Biogeochemistry:** Biomass (Contiguous US, Pan Tropical Forest), Burned area (GFED3), CO_2 (NOAA GMD, Mauna Loa), Gross primary production (Fluxnet, GBAF), Leaf area index (AVHRR, MODIS), Global net ecosystem carbon balance (GCP, Khatiwala/Hoffman), Net ecosystem exchange (Fluxnet, GBAF), Ecosystem Respiration (Fluxnet, GBAF), Soil C (HWSD, NCSCDv22, Koven)

- **Hydrology:** Evapotranspiration (GLEAM, MODIS), Evaporative fraction (GBAF), Latent heat (Fluxnet, GBAF, DOLCE), Runoff (Dai, LORA), Sensible heat (Fluxnet, GBAF), Terrestrial water storage anomaly (GRACE), Permafrost (NSIDC)

- **Energy:** Albedo (CERES, GEWEX.SRB), Surface upward and net SW/LW radiation (CERES, GEWEX.SRB, WRMC.BSRN), Surface net radiation (CERES, Fluxnet, GEWEX.SRB, WRMC.BSRN)

- **Forcing:** Surface air temperature (CRU, Fluxnet), Diurnal max/min/range temperature (CRU), Precipitation (CMAP, Fluxnet, GPCC, GPCP2), Surface relative humidity (ERA), Surface down SW/LW radiation (CERES, Fluxnet, GEWEX.SRB, WRMC.BSRN)

Multi-Model Validation Example

RUBISCO

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