Systematic Land Model Calibration

Quantifying parametric uncertainty and working towards automated calibration

Linnia Hawkins, Daniel Kennedy, Katie Dagon, Pierre Gentine, Dave Lawrence

NCAR

LEAP

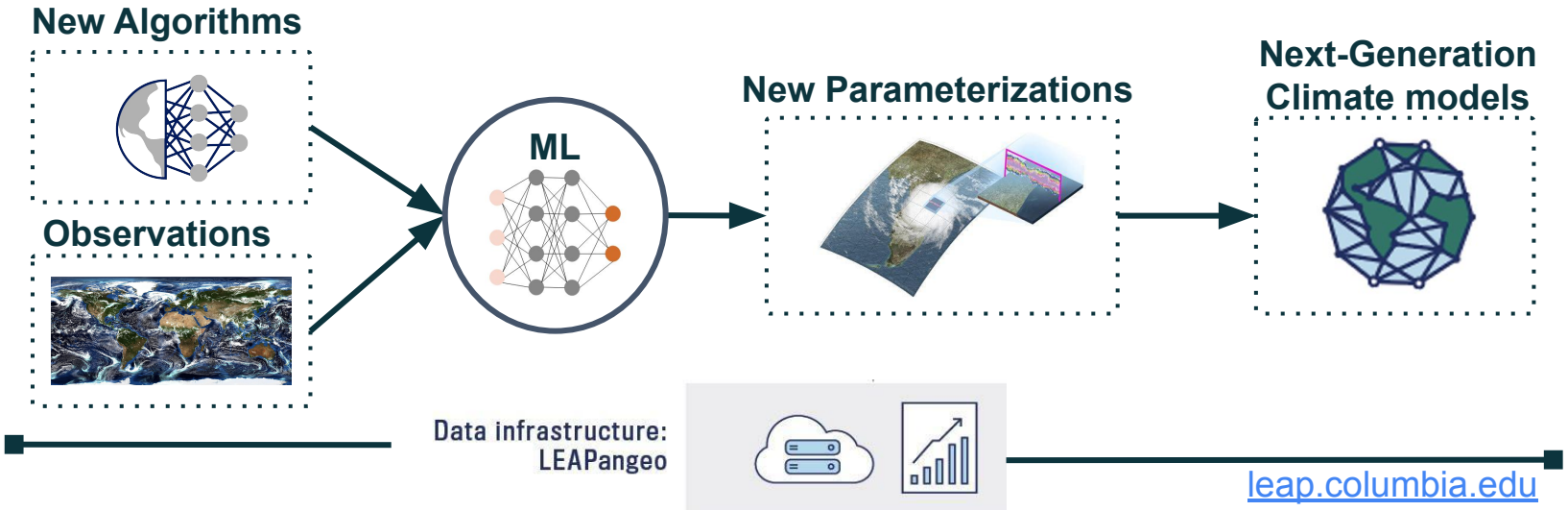
NSF

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

LEAP

LEARNING THE EARTH WITH ARTIFICIAL INTELLIGENCE AND PHYSICS

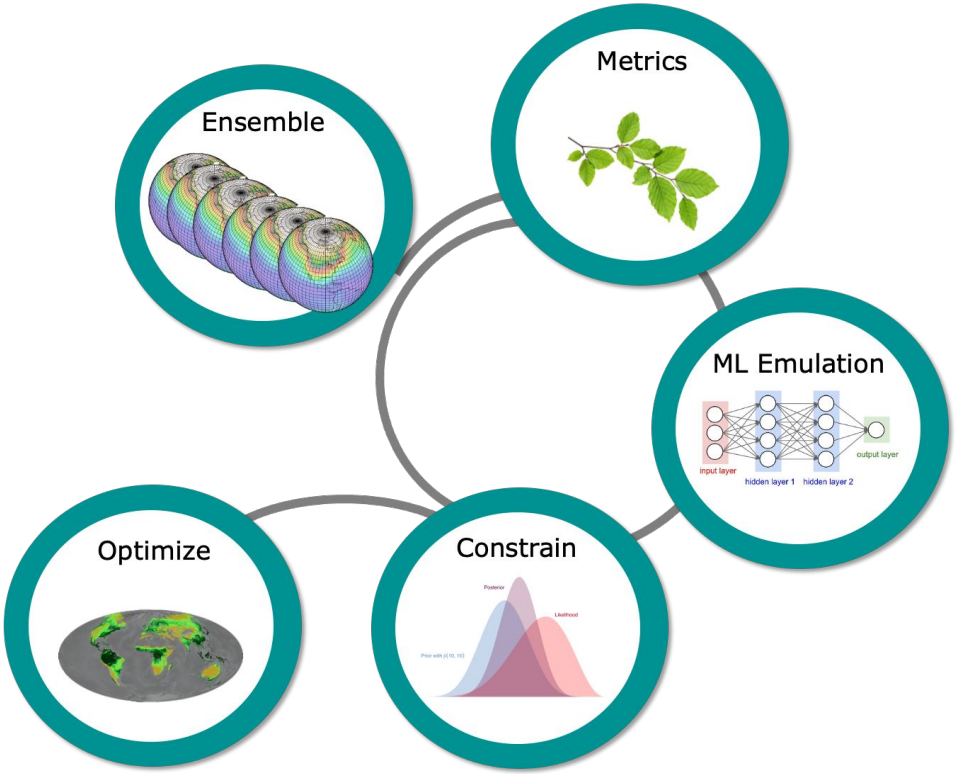
1. Push the frontiers of ML to create algorithms that discover and leverage physical and causal knowledge
2. Accelerate CESM development with novel parameterizations enabled by ML and growing datasets
3. Establish and deploy a modern cloud computing infrastructure for climate data, LEAPangeo.
4. **Establish systematic ML-based methodology for calibration of Earth System Models**



leap.columbia.edu

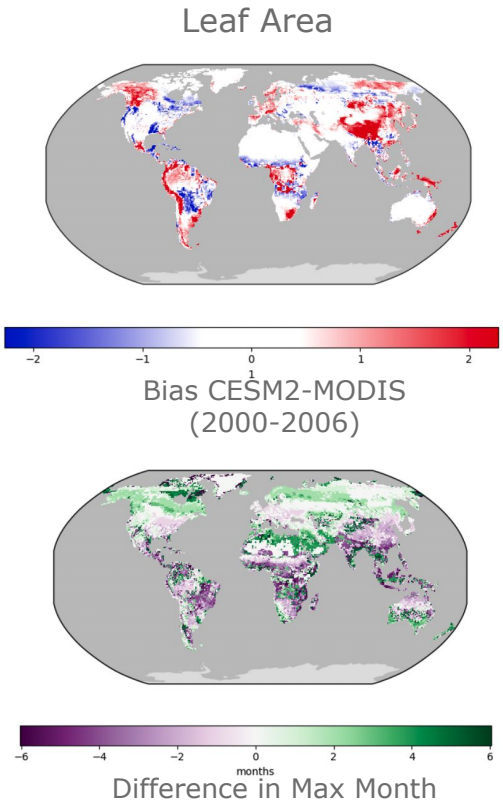
Approach

Establish systematic ML-based methodology for calibration of Earth System Model parameters

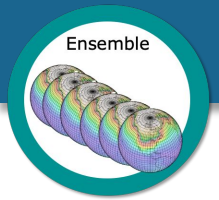


Leaf Area Index (m^2 leaf m^{-2} ground)

- Observations of leaf area are global and robust
- Vegetation structure is foundational
- CLM5.0 biases:
  - too high in tropics and arctic
  - timing of seasonal peak
    - early in tropics & late in arctic
  - trend is too strong
  - interannual variability is underestimated



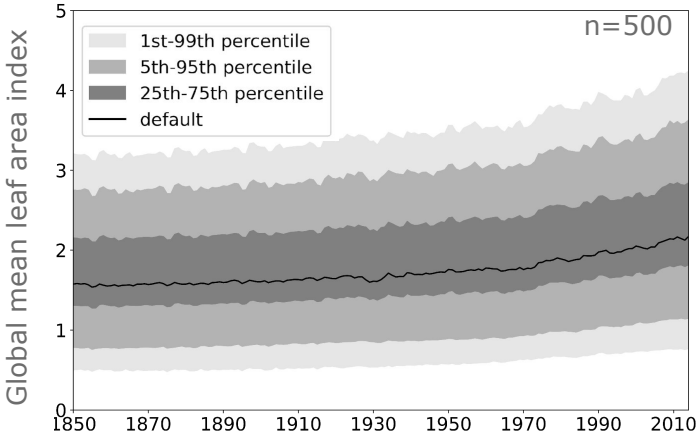
Figures from ILAMB.org



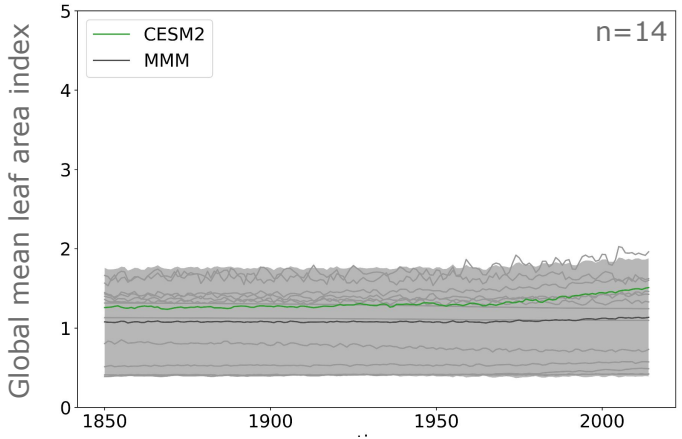
Perturbed Parameter Ensemble

- Latin Hypercube Ensemble
  - 500 ensemble members (1850-2014)
  - 32 parameters varying simultaneously

CLM5-PPE



CMIP6



Surrogate models

Training ML-based emulators for an ESM

- Fast & computationally cheap
- Explore parameter space
- Assess parameter sensitivity
- Detect interactions and non-linearities

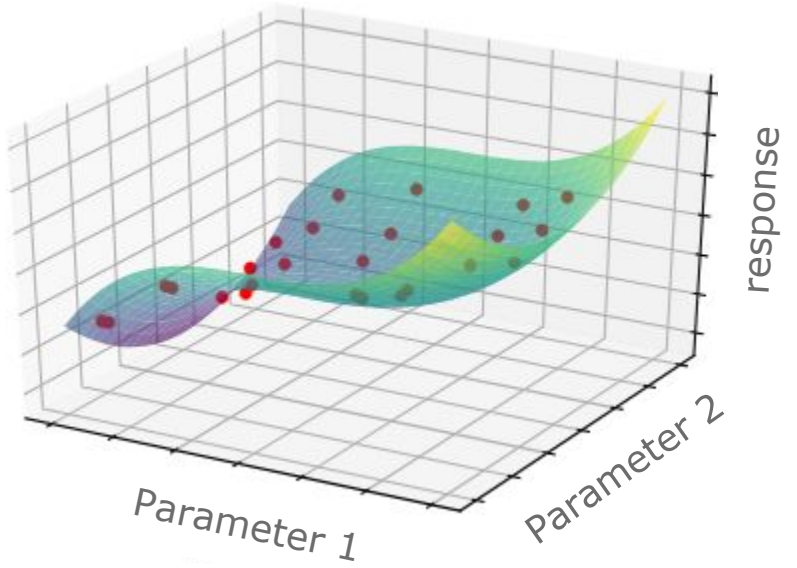
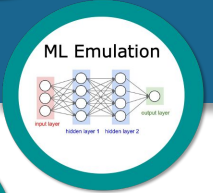
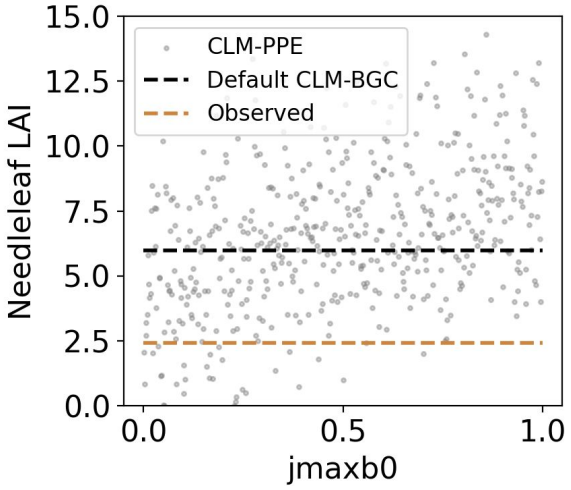


Figure courtesy of Duncan Watson-Parris
duncanwp.github.io

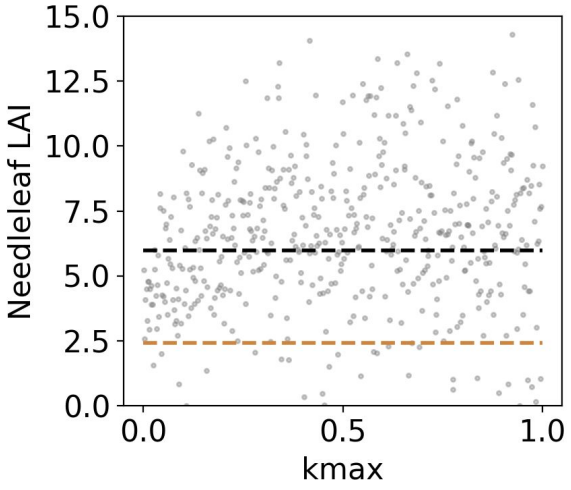


PPE and emulators

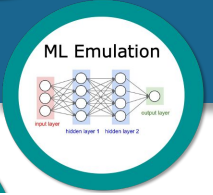
Use the PPE and ML based emulators to identify non-linearities and interactions



Photosynthetic capacity

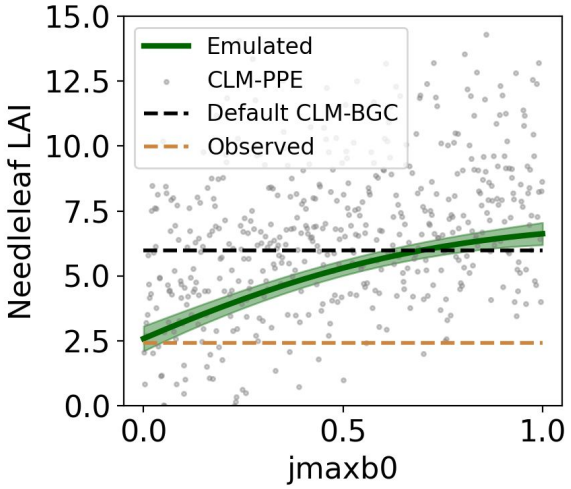


Max Conductance

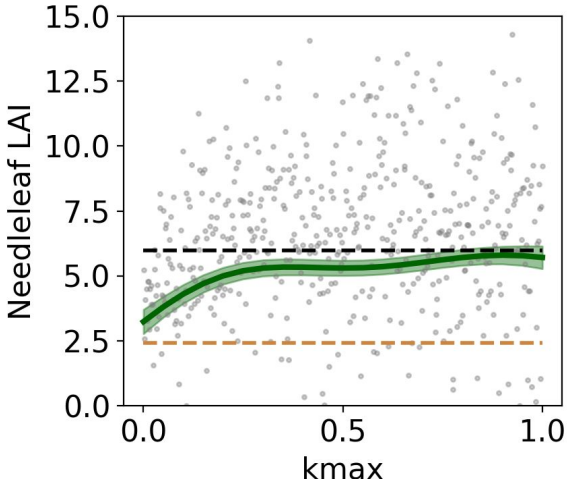


PPE and emulators

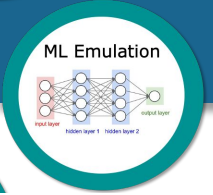
Use the PPE and ML based emulators to identify non-linearities and interactions



Photosynthetic capacity

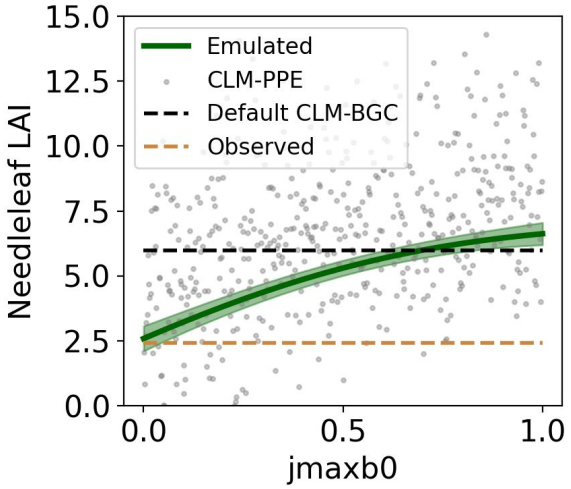


Max Conductance

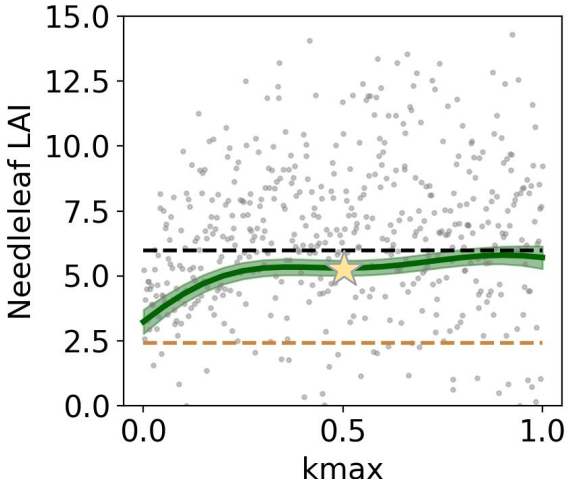


PPE and emulators

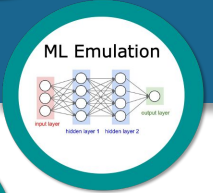
Use the PPE and ML based emulators to identify non-linearities and interactions



Photosynthetic capacity

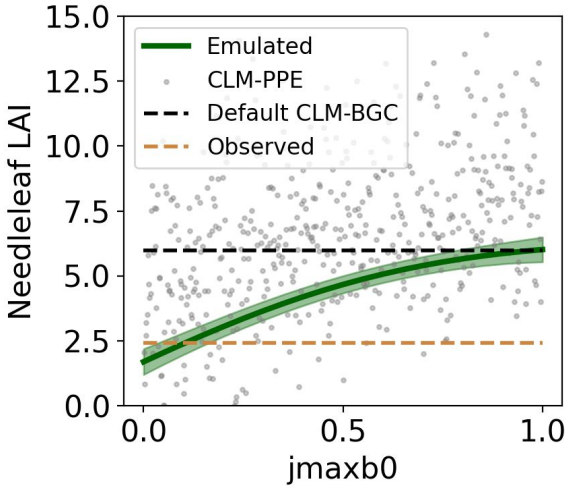


Max Conductance

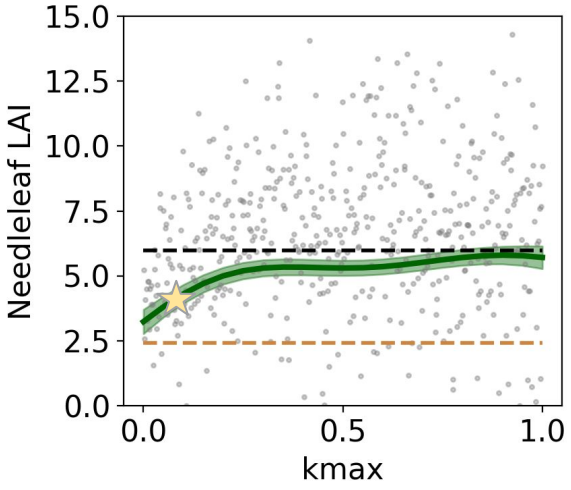


PPE and emulators

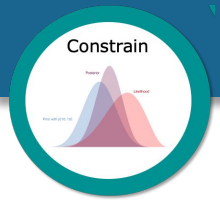
Use the PPE and ML based emulators to identify non-linearities and interactions



Photosynthetic capacity

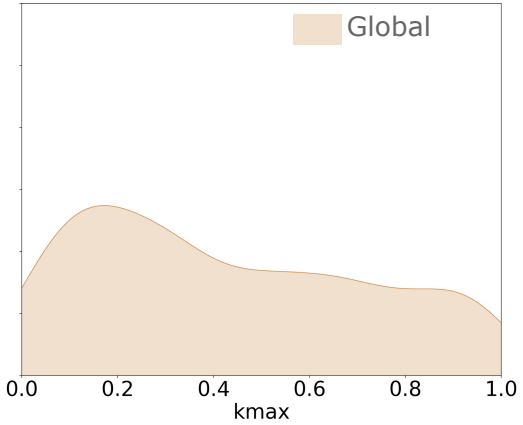


Max Conductance

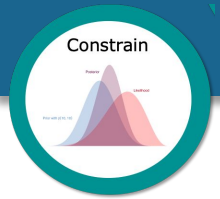


Constrain parameter space

Global mean doesn’t considerably constrain parameter space
- trade-offs & compensation between parameters
- equifinality between parameter sets



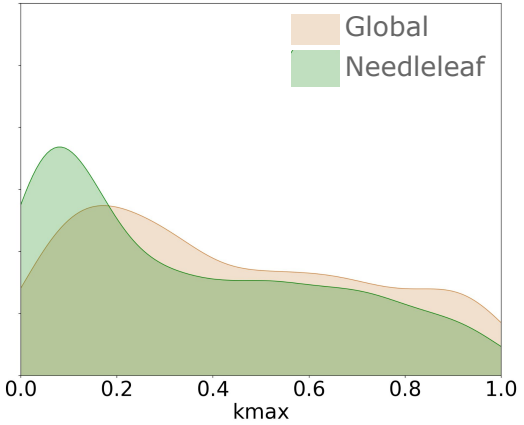
Max Conductance



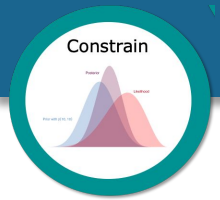
Constrain parameter space

Global mean doesn’t considerably constrain parameter space
- trade-offs & compensation between parameters
- equifinality between parameter sets

Targeting individual PFT’s better constrains posteriors



Max Conductance

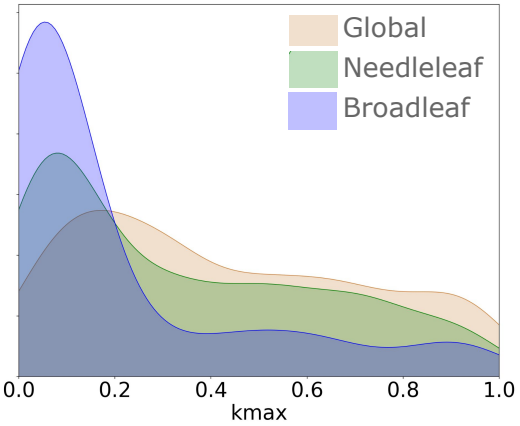


Constrain parameter space

Global mean doesn’t considerably constrain parameter space
- trade-offs & compensation between parameters
- equifinality between parameter sets

Targeting individual PFT’s better constrains posteriors

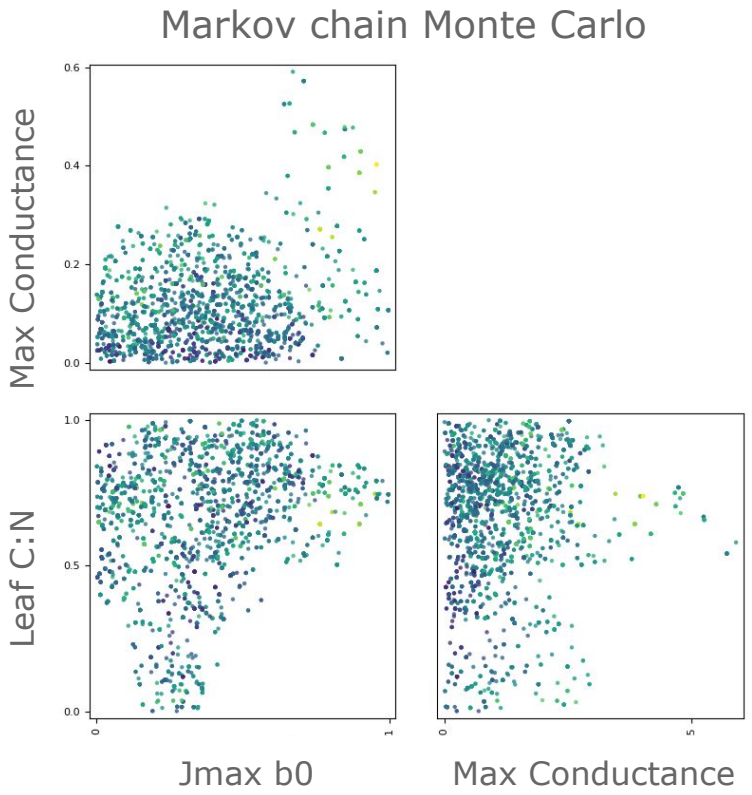
Strategically add metrics to further constrain plausible parameter ranges.



Max Conductance

Optimization

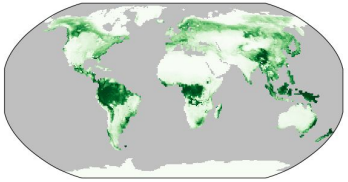
- ✔ Defined target metrics
- ✔ Constrained posterior distributions
- ✔ Robust emulator(s)
  - Optimizations techniques (e.g., MCMC)
  - Identify improved parameter sets
  - Re-run CLM



Systematic calibration requires strategic design

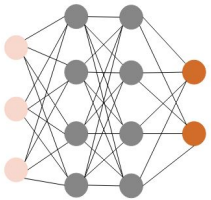
Computational Constraints

Limited CLM runs



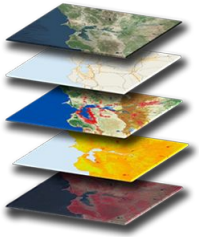
High dimensionality

Many parameters



Multiple target metrics

Vary in space and time



Outcomes

- **Systematic calibration workflow:** transferable across model configurations
- **Accessibility:** Easier for community to apply the model for specific applications
  - Point / Region
  - Tune to your data
- **Insight:** improve our understanding of CLM parameters and processes

Community input!

LEAP leap.columbia.edu

github.com/djk2120/CLM5PPE

LH3194@columbia.edu

Extra Slides

Identify relative influence of parameters

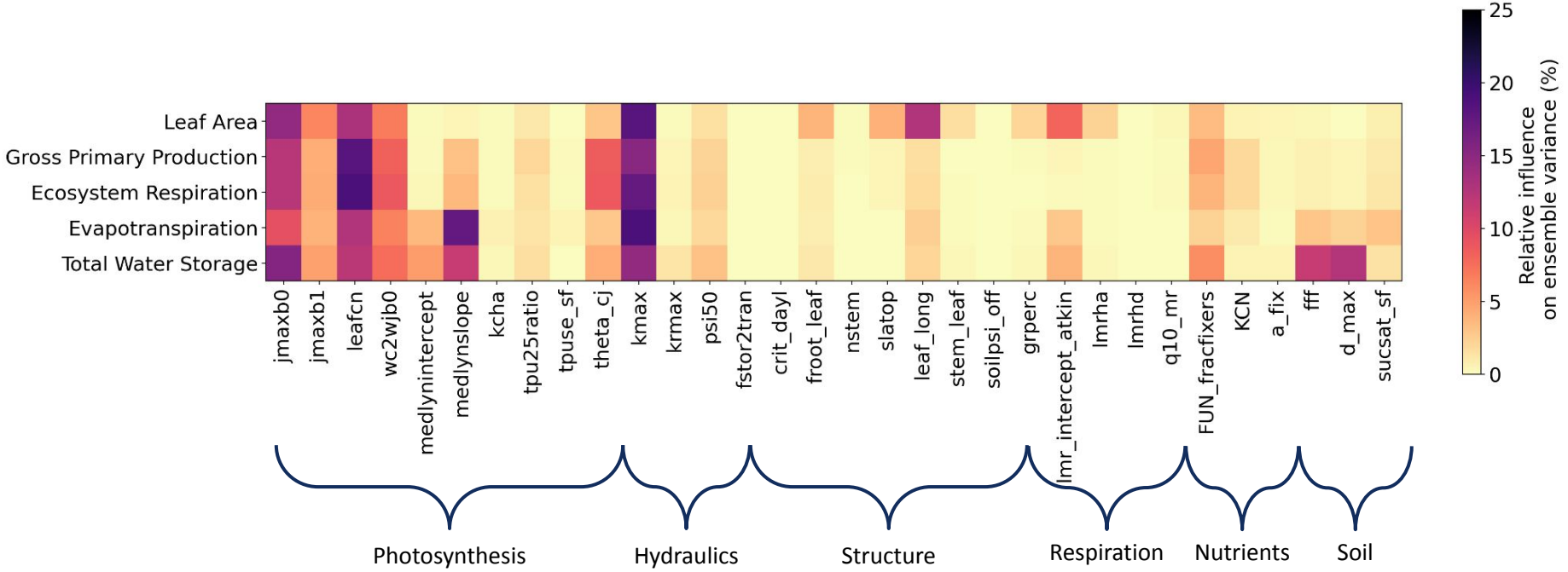


Figure from Linnia Hawkins