Quantifying Drivers of Uncertainty in Land Model Predictions from Site to Global Scales Using Machine Learning

Daniel M. Ricciuto (ORNL) and Khachik Sargsyan (SNL-CA)

Cosmin Safta, Vishagan Ratnaswamy (SNL-CA)
Dan Lu, Peter Thornton (ORNL)
Jennifer Holm (LBNL), Beth Drewniak (ANL)

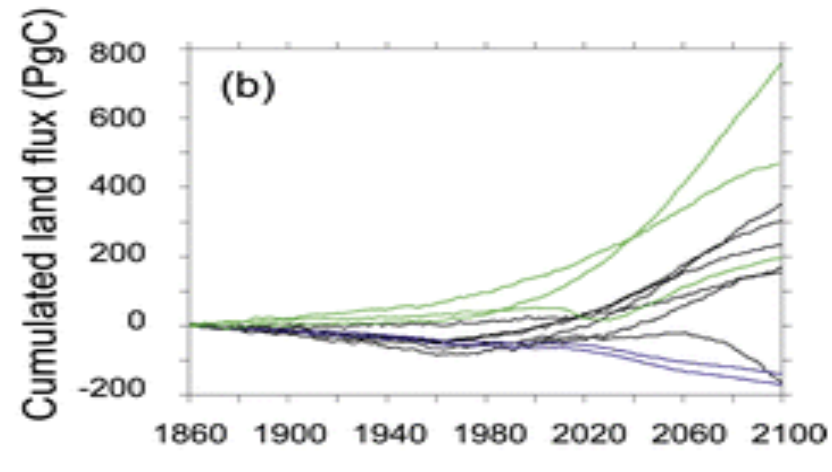
CBGC Webinar
July 28, 2020

Overview and motivation: CBGC models

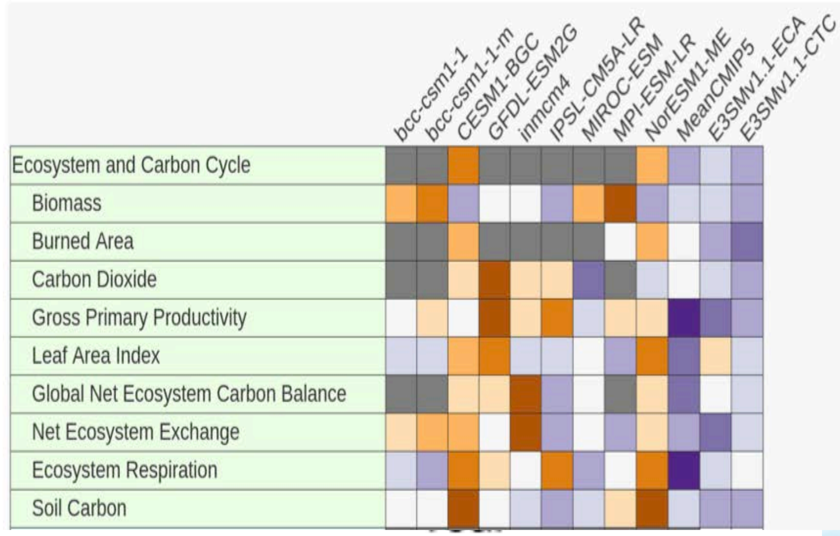
- Uncertainty comes from Multi-model ensembles
- Large spread in outputs
- Many quantities of interest
- Little formal uncertainty quantification (UQ)
  - Expensive model evaluation
  - High dimensionality

UQ challenges in E3SM and SciDAC:

- What processes drive uncertainty?
- What accounts for the key differences among models?
- Can model calibration using observations (e.g. satellite data) reduce uncertainty?



Friedlingstein et al (2014)

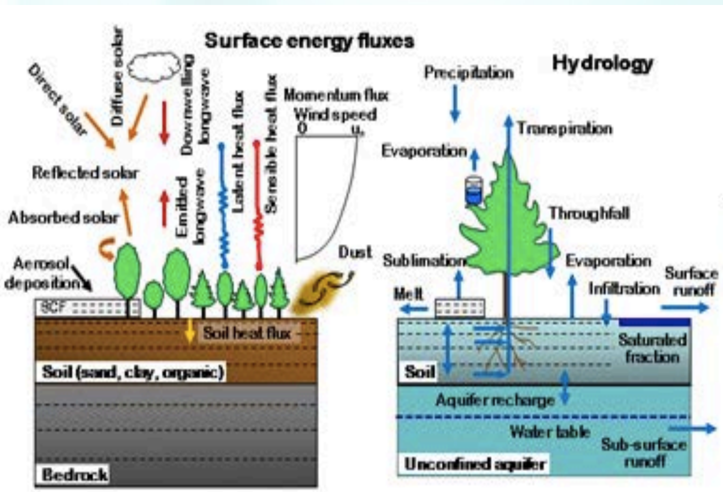


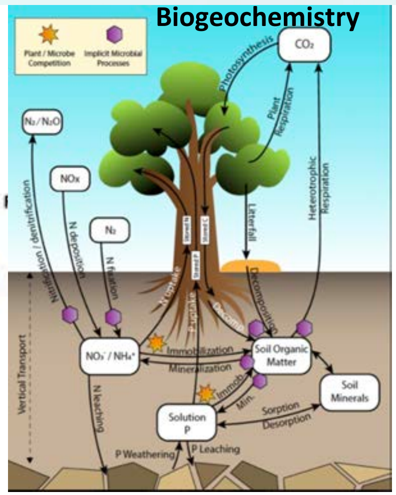
Burrows et al (in review)

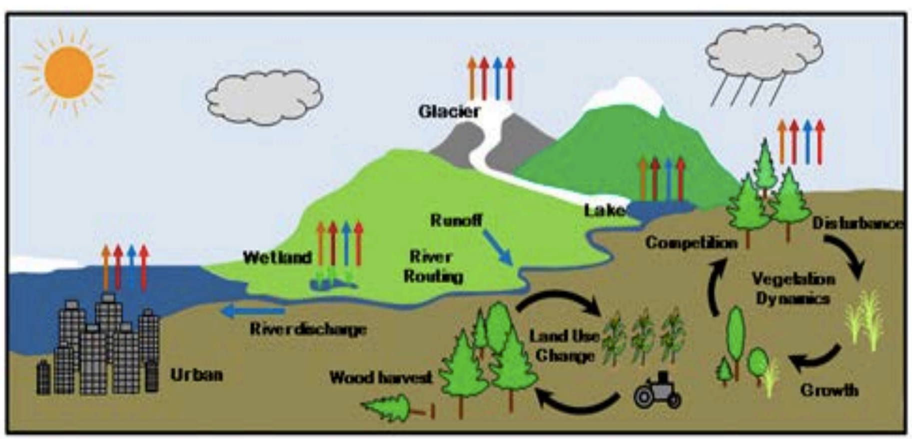
E3SM Energy Exascale Earth System Model

U.S. DEPARTMENT OF ENERGY

Overview and motivation: Land BGC

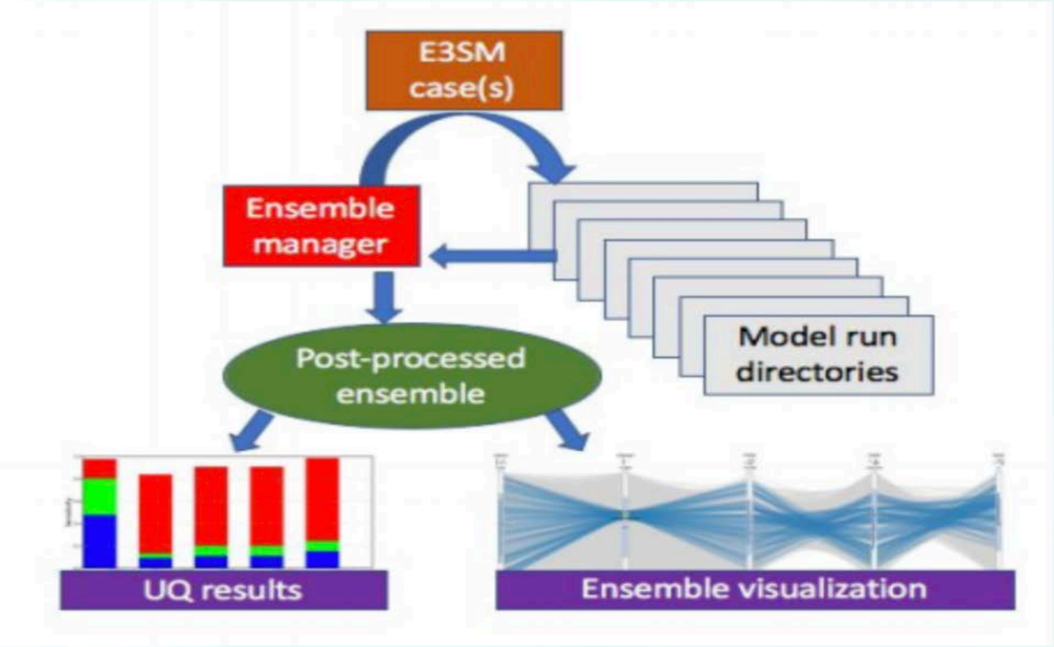






- ELM is an increasingly complex model with many processes
- Large ensembles are needed for UQ
- For BGC simulation, spinup limits ensemble size.
- Surrogate models can increase the efficiency of sensitivity analysis and calibration

Model infrastructure to enable UQ



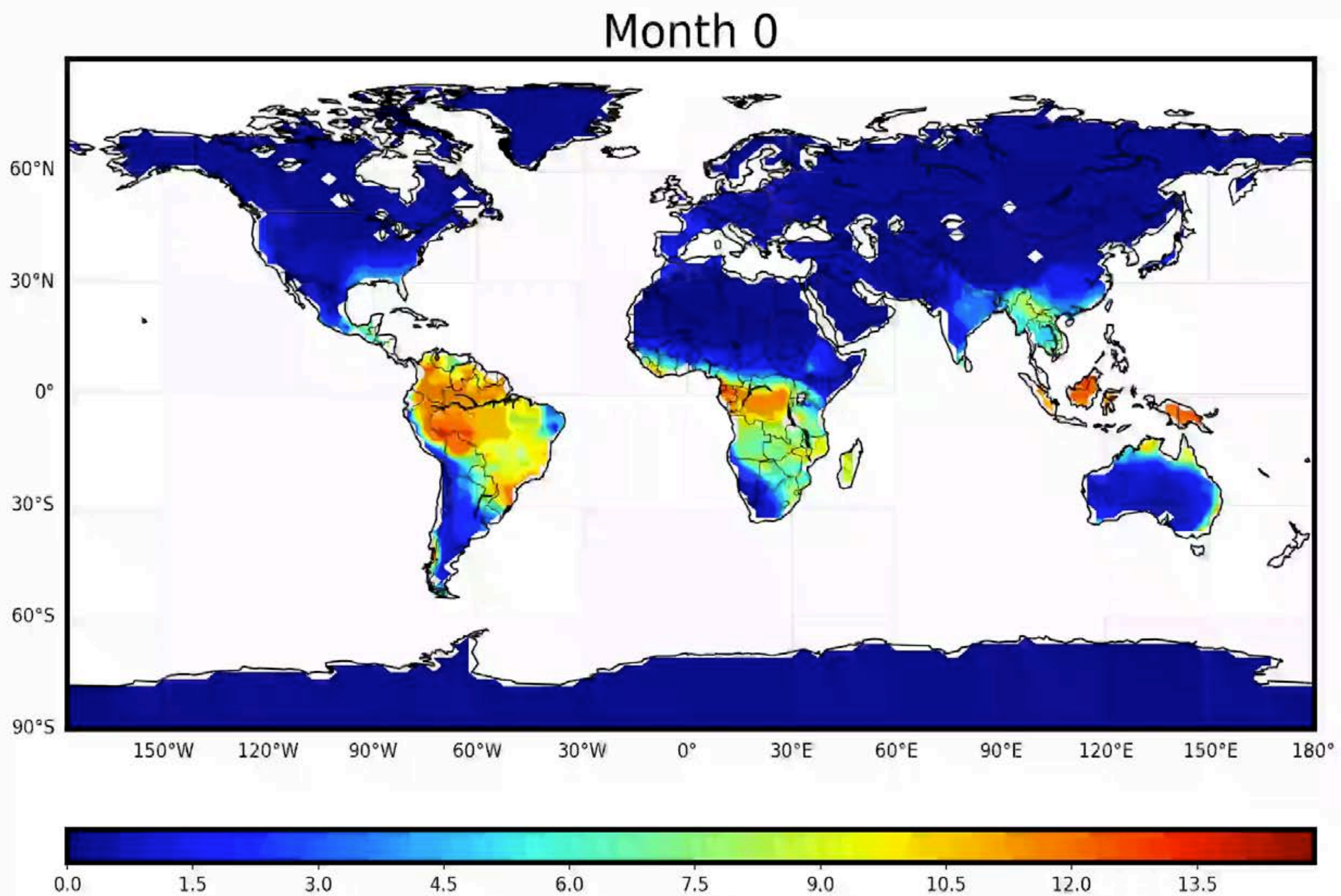
- E3SM single gridcell
  - Default v1 (FLUXNET sites)
  - FATES (Boreal Alaska site)
  - Using mpi4py as part of Offline Land Model Testbed (**OLMT**) to manage ensembles and post-process
  - 100s to 1000s of ensembles

Global ELM
  - SP version so far
  - Modified OLMT framework
  - Low resolution (1.9x2.5)
  - 10 ELM parameters
  - Photosynthesis/leaf parameters
  - 200 simulations, GPP as QoI

Model parameter ensemble

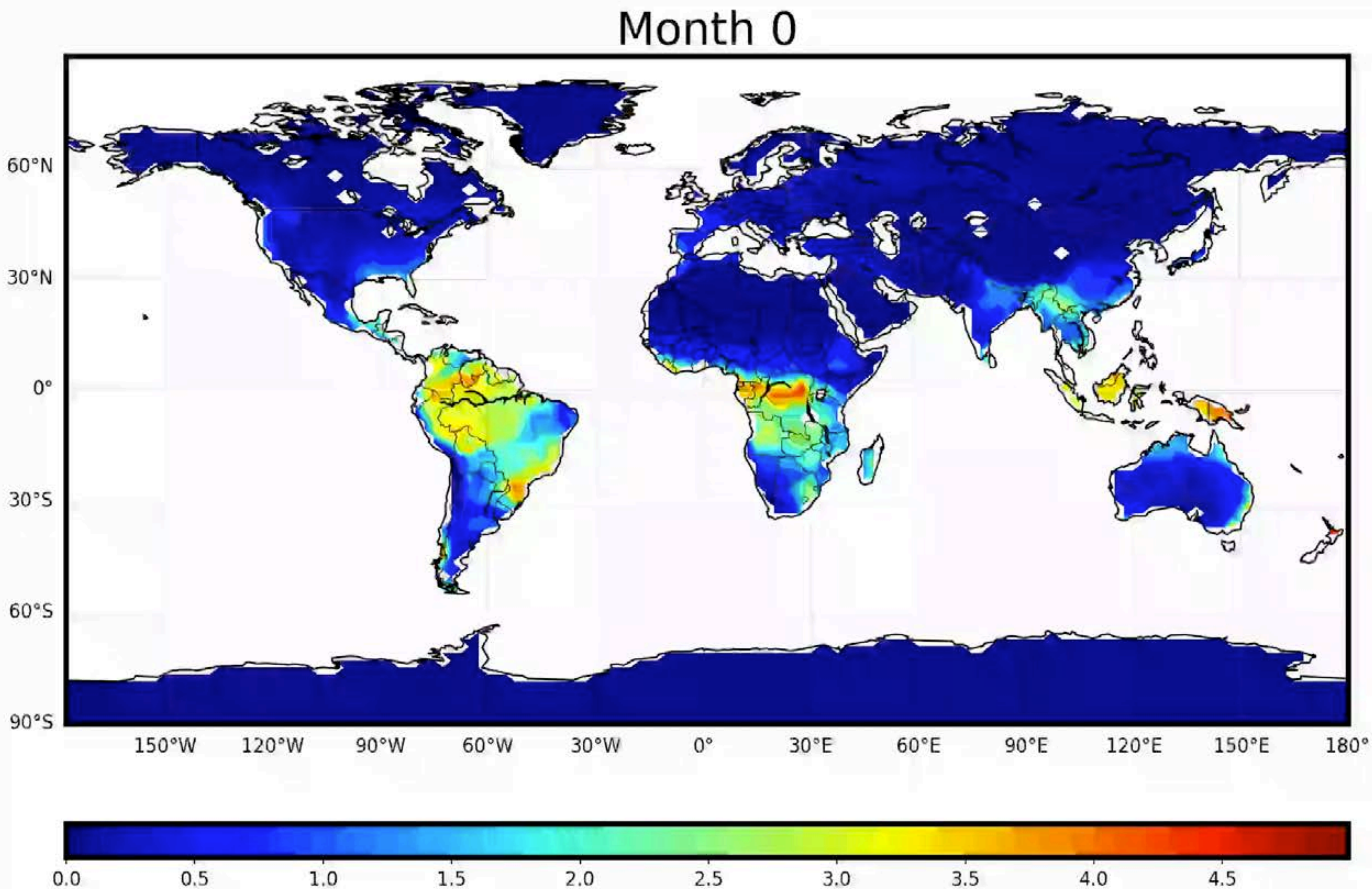
Parameter	description	Min	Max	Default range
flnr	Fraction of leaf N in RuBisCO	0.03	0.25	[0.042,0.176]
mbbopt	Ball-Berry slope parameter	4.0	13.0	[4,9]
bbbopt	Ball-Berry intercept parameter	1000	40000	[10000,40000]
roota_par	Rooting depth distribution parameter	1	10	[3,10]
vcmaxha	Activation energy for Vcmax	50000	90000	72000
vcmaxse	Entropy for Vcmax	640	700	670
dayl_scaling	Day length scaling parameter	1.0	2.5	2.0
dleaf	Characteristic leaf dimension	0.01	0.1	0.04
xl	Leaf/stem orientation index	-0.6	0.8	[-0.5,0.65]

Mean ensemble GPP



GPP (gC m^{-2} day^{-1})

GPP standard deviation

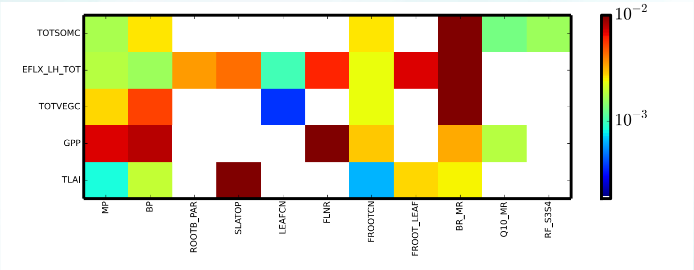


GPP standard deviation (gC m^{-2} day^{-1})

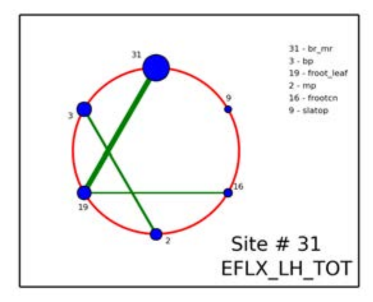
Global Sensitivity Analysis (GSA) enables parameter selection

... otherwise called Sobol indices, variance-based decomposition
Attribute fractions of output variance to input parameters

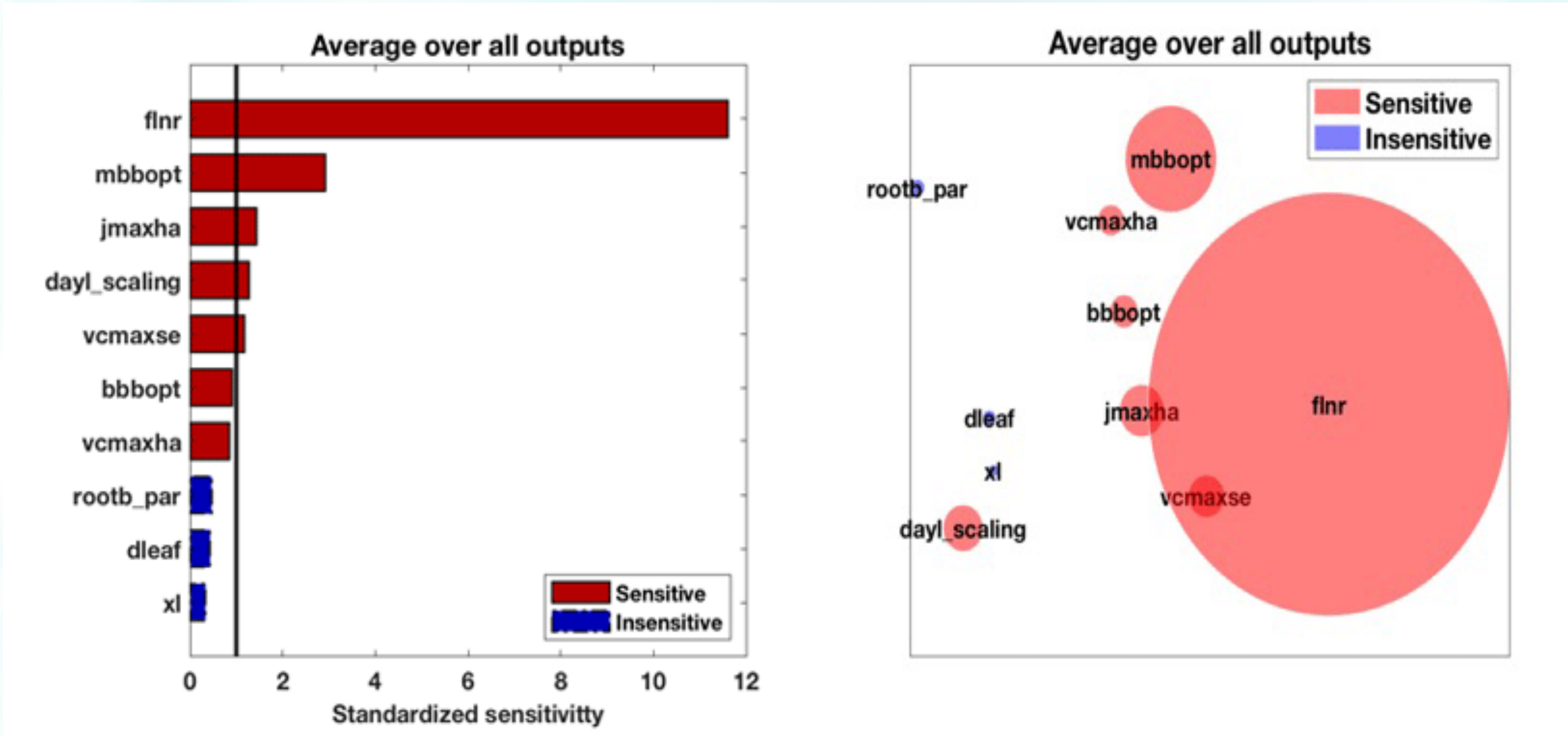
- i.e., how much output variance would reduce if a given parameter is fixed to its nominal value



- also generalizes to joint sensitivities: joint parameter impact to a given QoI



Sensitivities of global average GPP



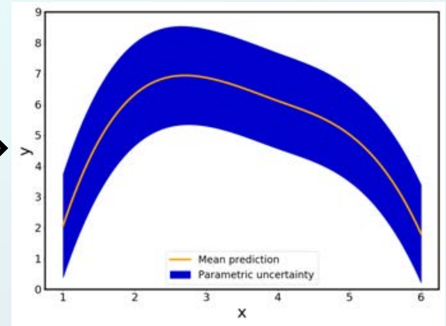
Uncertainty Propagation
… enabled by Surrogate Models

λ

Input parameters

Comp. model

$f(x;\lambda)$



Output prediction

- *Forward predictions:*
  - surrogate models,
  - sensitivity analysis,
  - parametric uncertainty

Work with the model as a black-box (**non-intrusive**):
create an ensemble of simulations with varying/perturbing λ

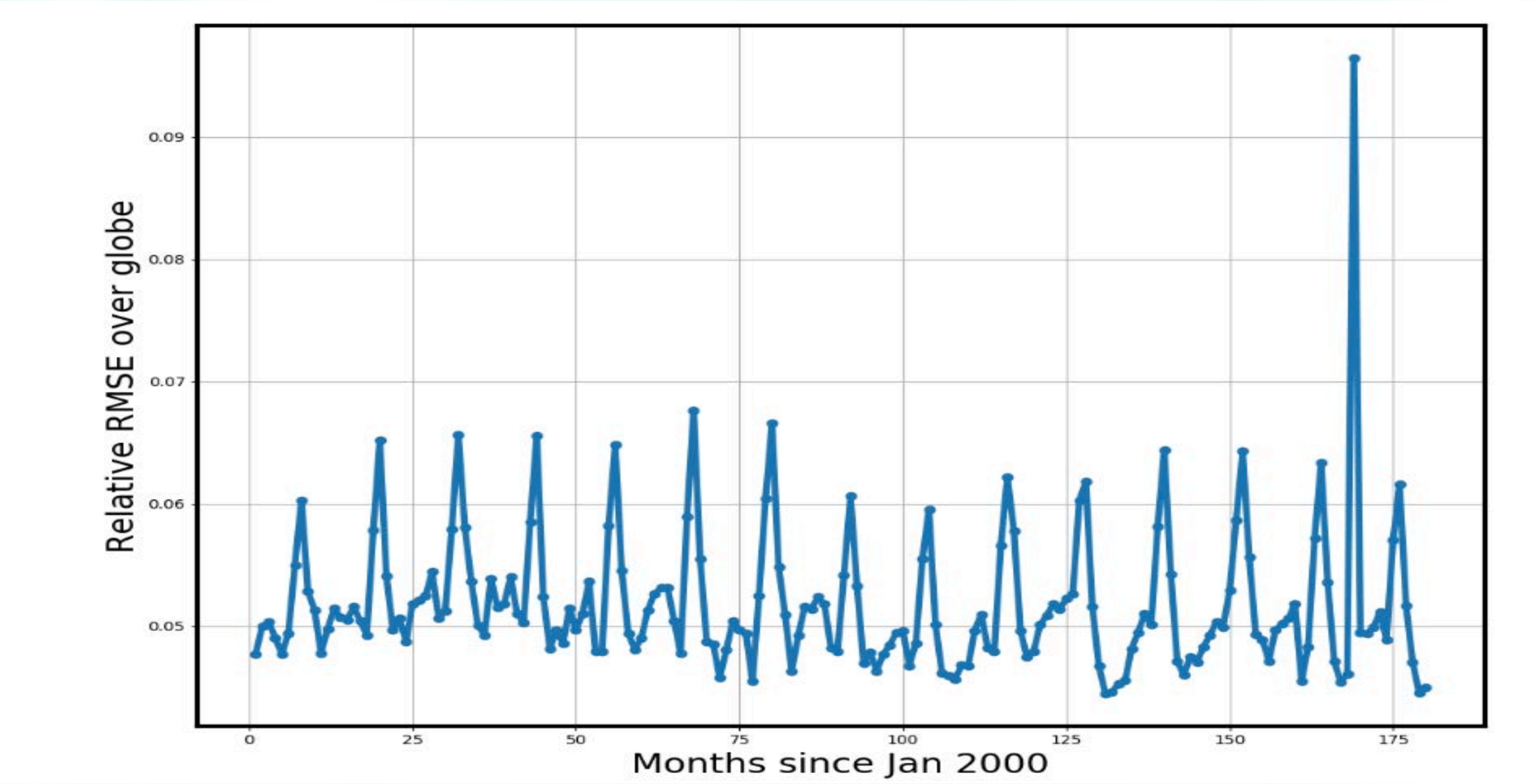
Never analyze the ensemble directly:
build a **surrogate** first
… otherwise called proxy, metamodel, emulator, response surface, supervised ML

$$f(x;\lambda) \approx f_s(x;\lambda)$$

Spatio-temporal surrogate model via Karhunen-Loève expansions

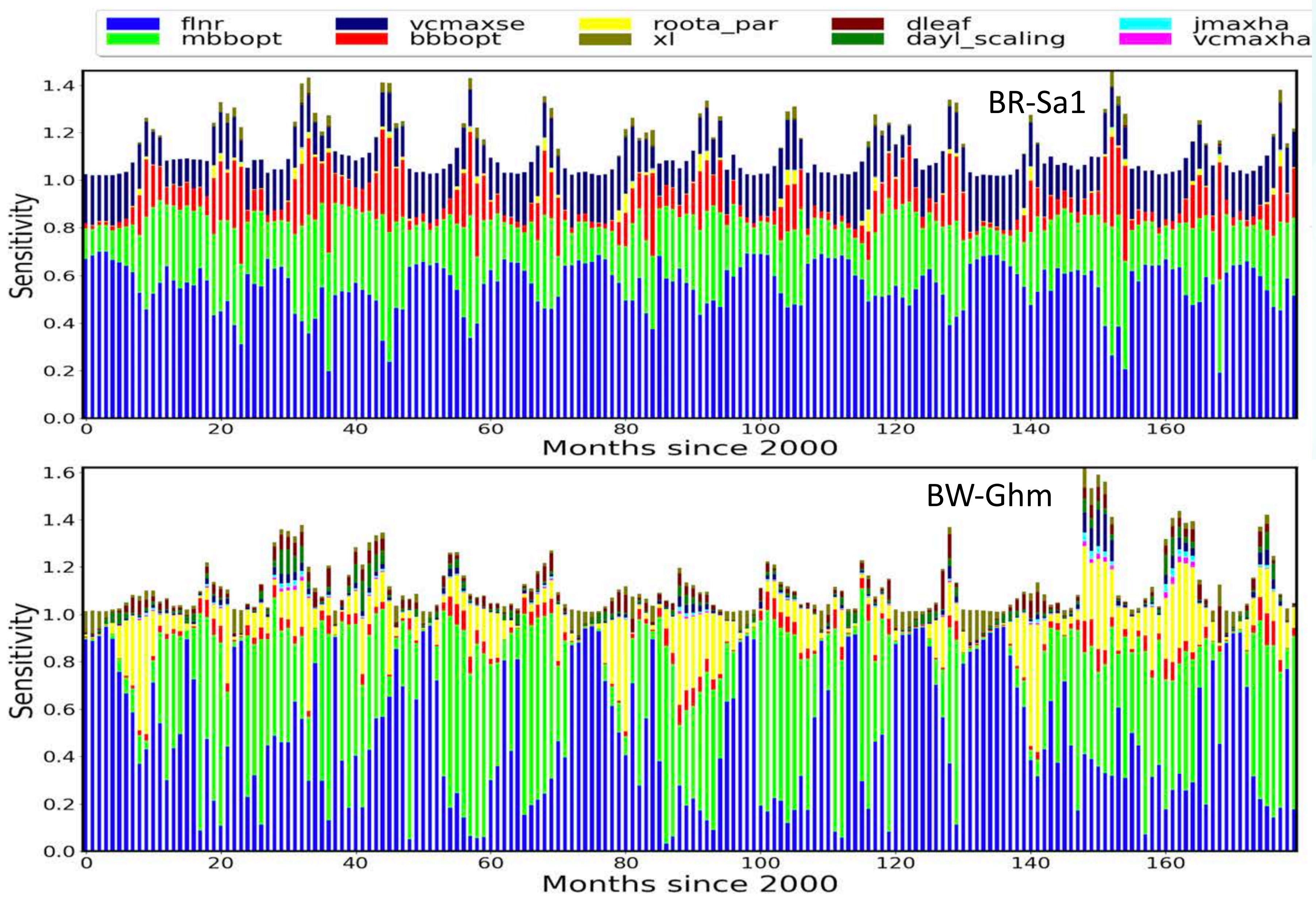
- 3183 active land cells over 180 months is > 500,000 outputs
- Karhunen-Loève expansions help reduce dimensionality due to strong spatio-temporal correlations (think of Principal Component Analysis in stochastic space)
- Instead of 500K surrogates, we build about 2K surrogates, one for each eigen-component
- End result: a single surrogate, resolved in space and time, with about 5% relative error compared to true ELM
- Surrogate ELM is extremely cheap to evaluate and is being used online to calibrate the parameters

Surrogate model accuracy

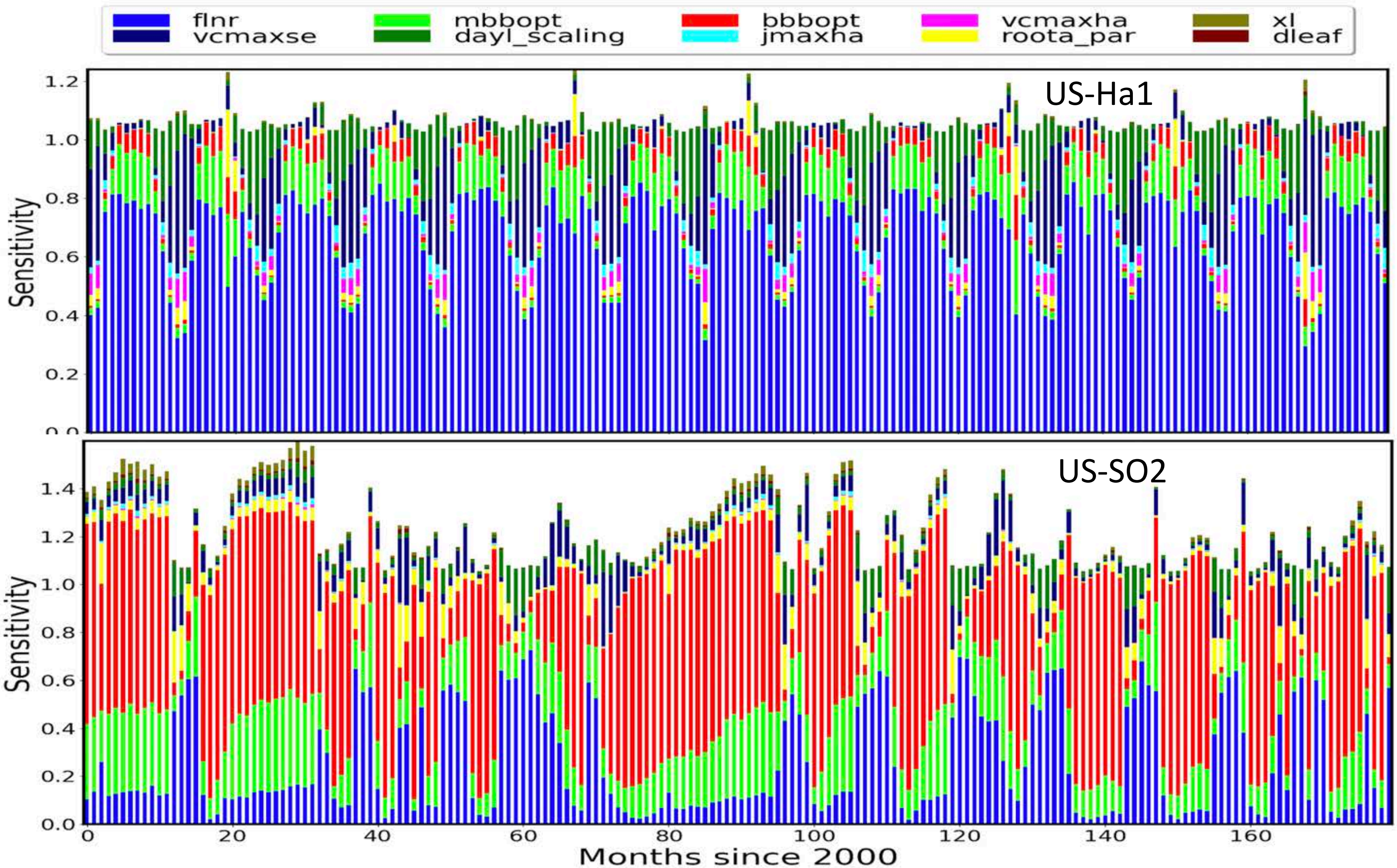


Relative error of the surrogate model averaged over all land cells
Highest errors in August (most complex behavior)

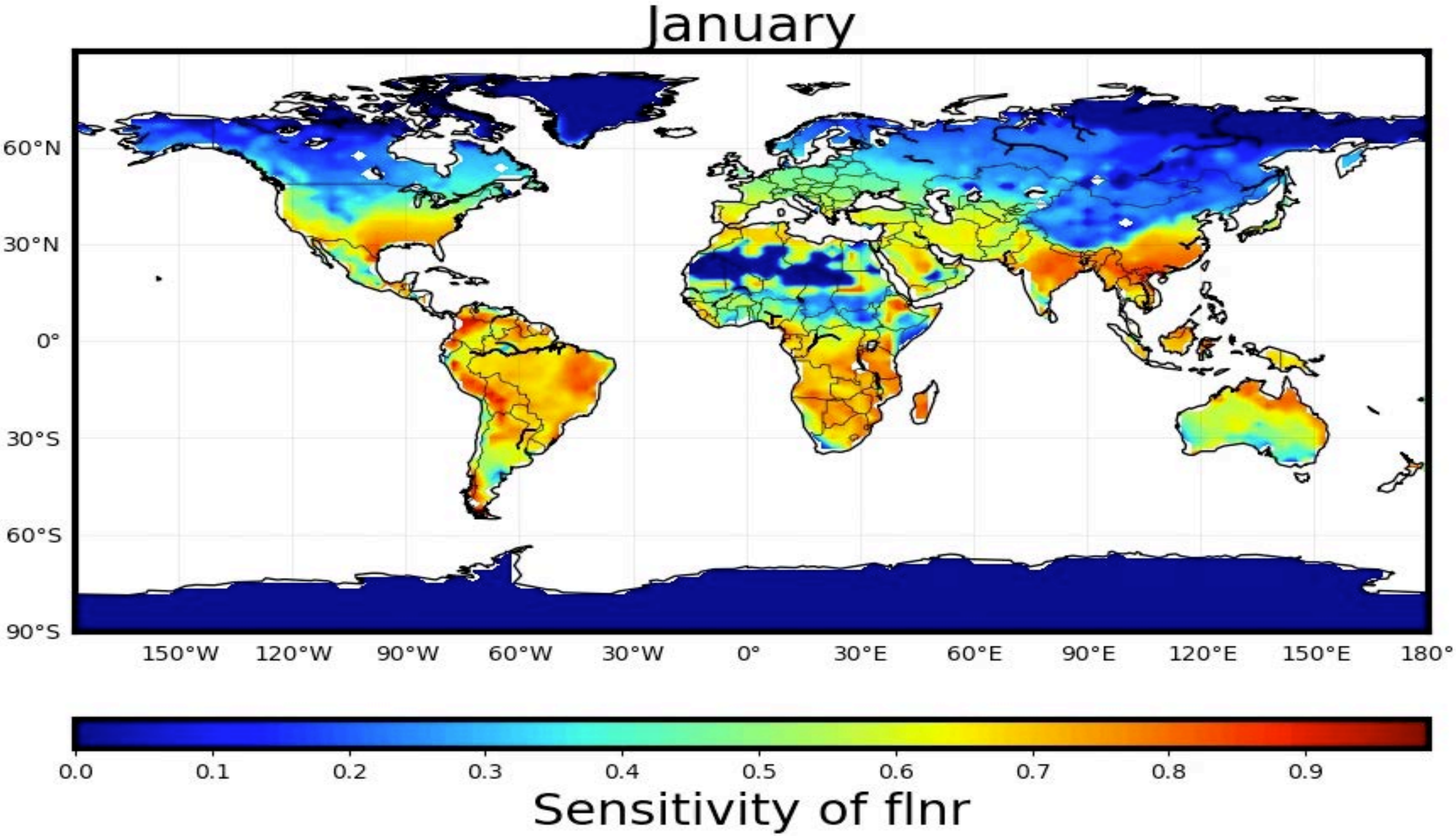
Gridcell-level sensitivities (tropical)



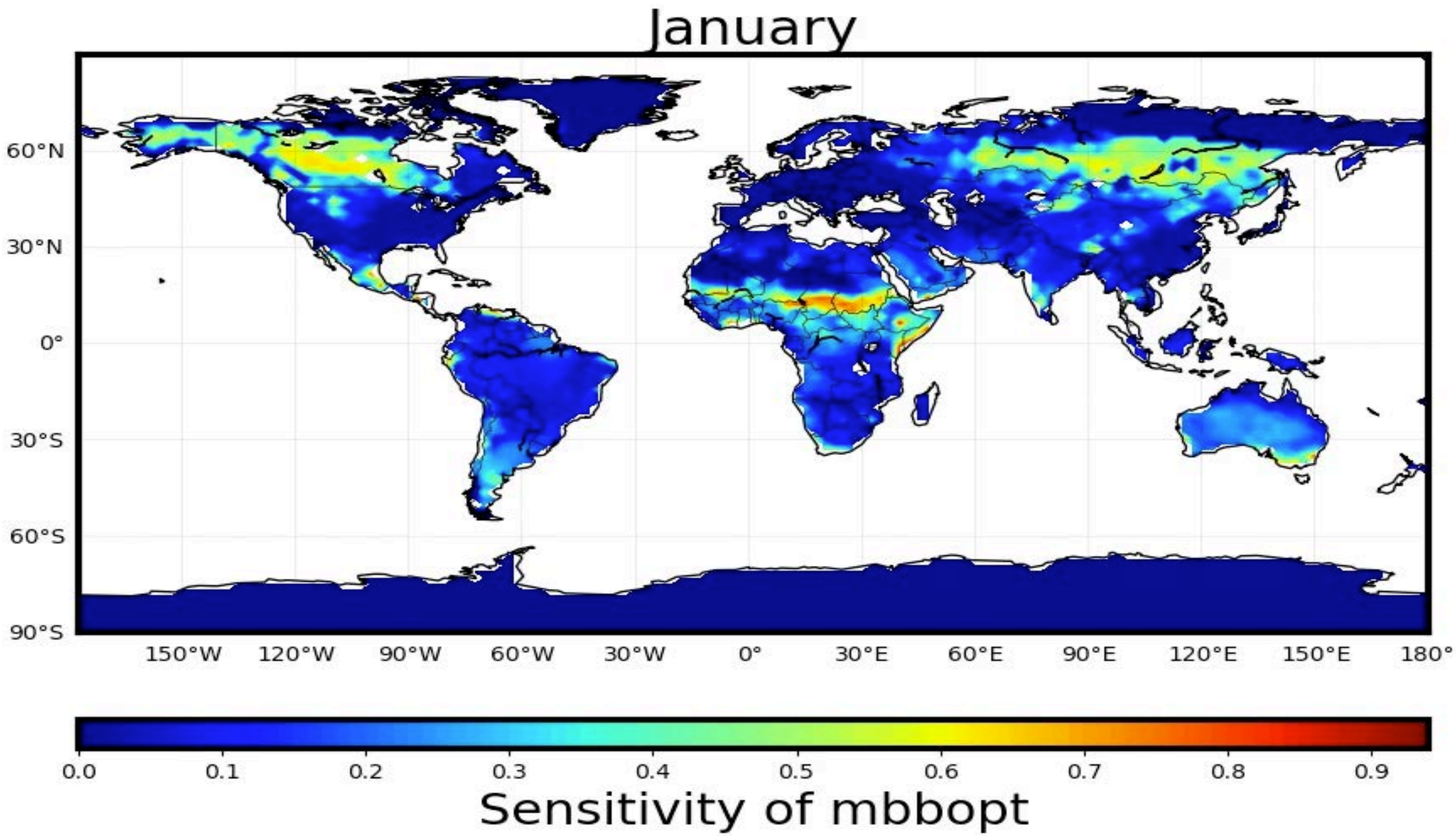
Gridcell-level sensitivities (temperate)



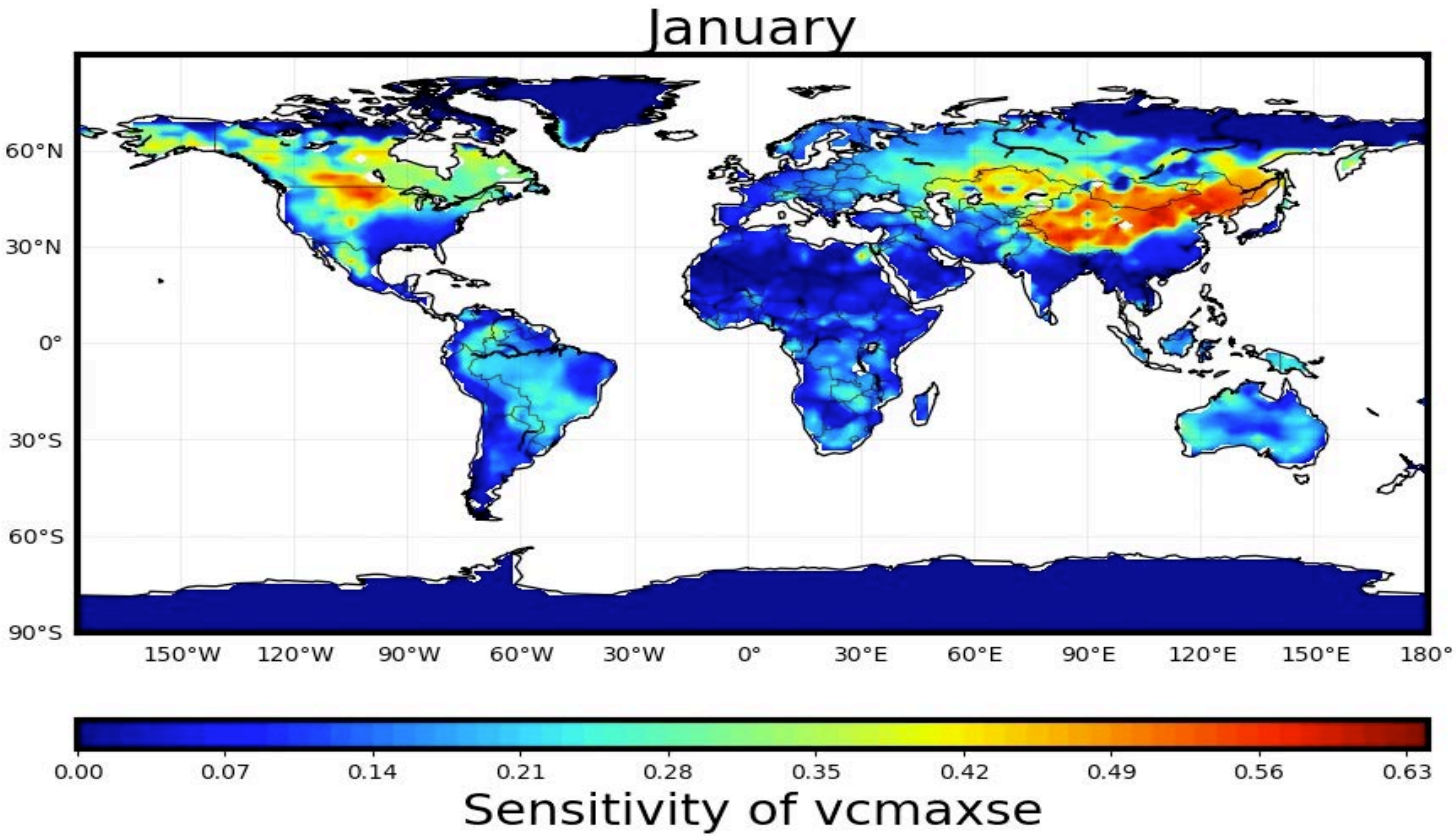
Globally mapped sensitivities



Globally mapped sensitivities



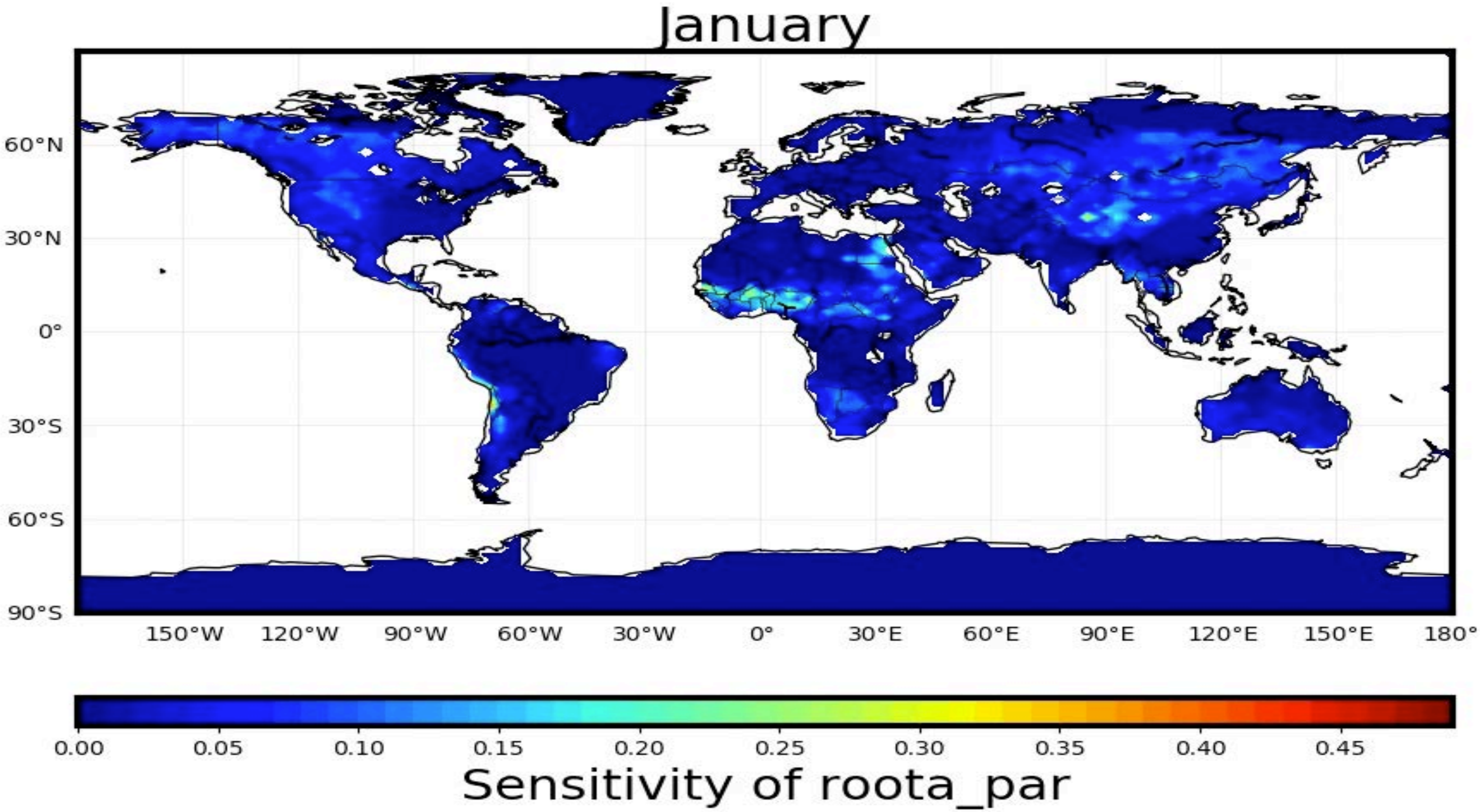
Globally mapped sensitivities



E3SM Energy Exascale Earth System Model

U.S. DEPARTMENT OF ENERGY

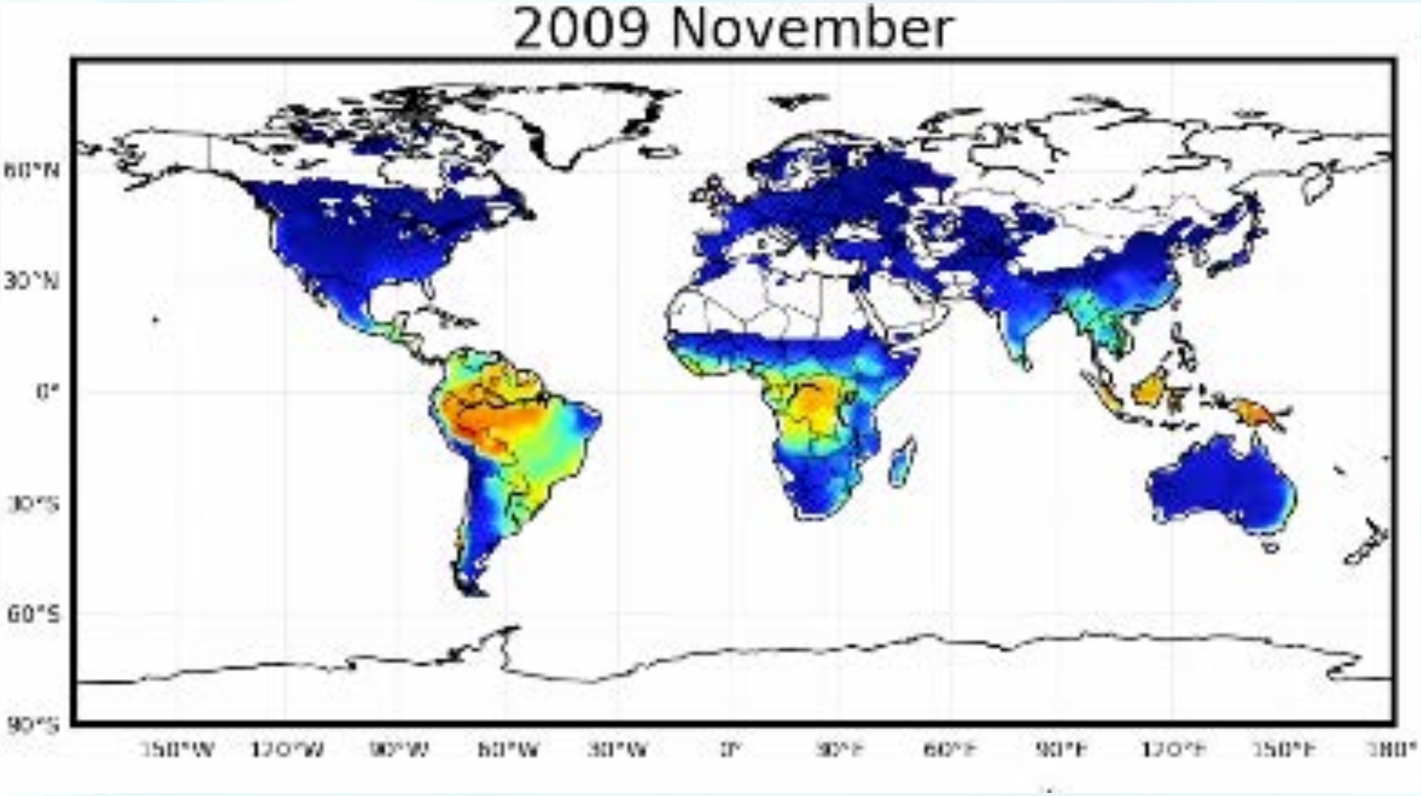
Globally mapped sensitivities



E3SM Energy Exascale Earth System Model

U.S. DEPARTMENT OF ENERGY

We have gridded observations and benchmarks

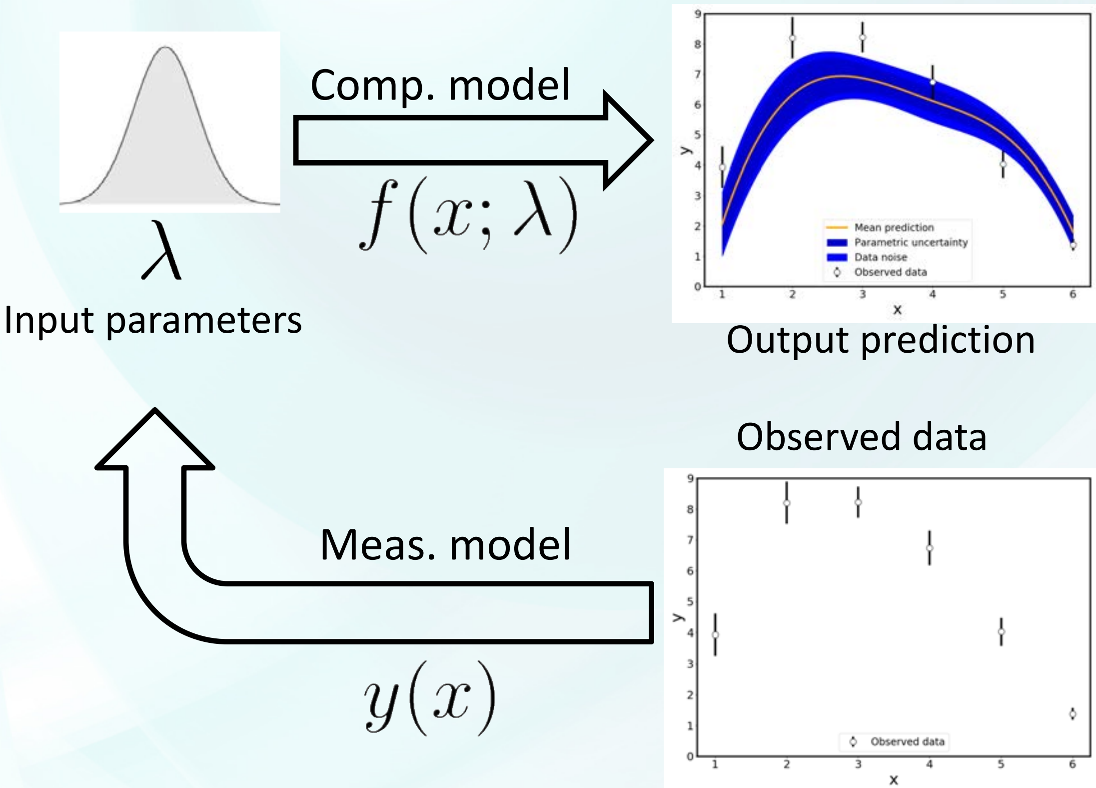


ILAMB GPP

Benchmark	[-]	118.
bcc-csm1-1	[-]	123.
bcc-csm1-1-m	[-]	112.
CESM1-BGC	[-]	129.
E3SMv1-CTC	[-]	124.
E3SMv1-ECA	[-]	106.
GFDL-ESM2G	[-]	167.
inmcm4	[-]	136.
IPSL-CM5A-LR	[-]	165.
MeanCMIP5	[-]	138.
MIROC-ESM	[-]	129.
MPI-ESM-LR	[-]	170.
NorESM1-ME	[-]	129.

- Directly comparable with model GPP output
- Can be used for single gridcell to global-scale calibration
- Methods extensible to other ILAMB products

Inverse modeling: tuning model parameters with observational data

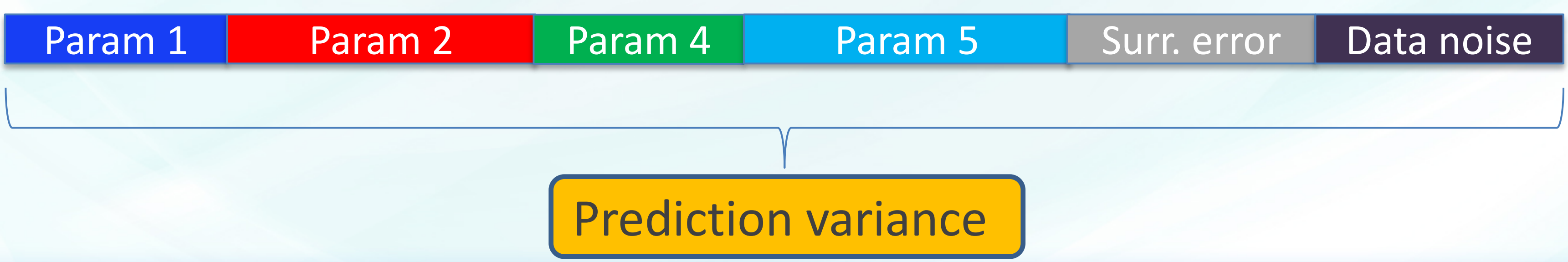


- *Forward predictions:*
  - surrogate models
  - sensitivity analysis,
  - parametric uncertainty

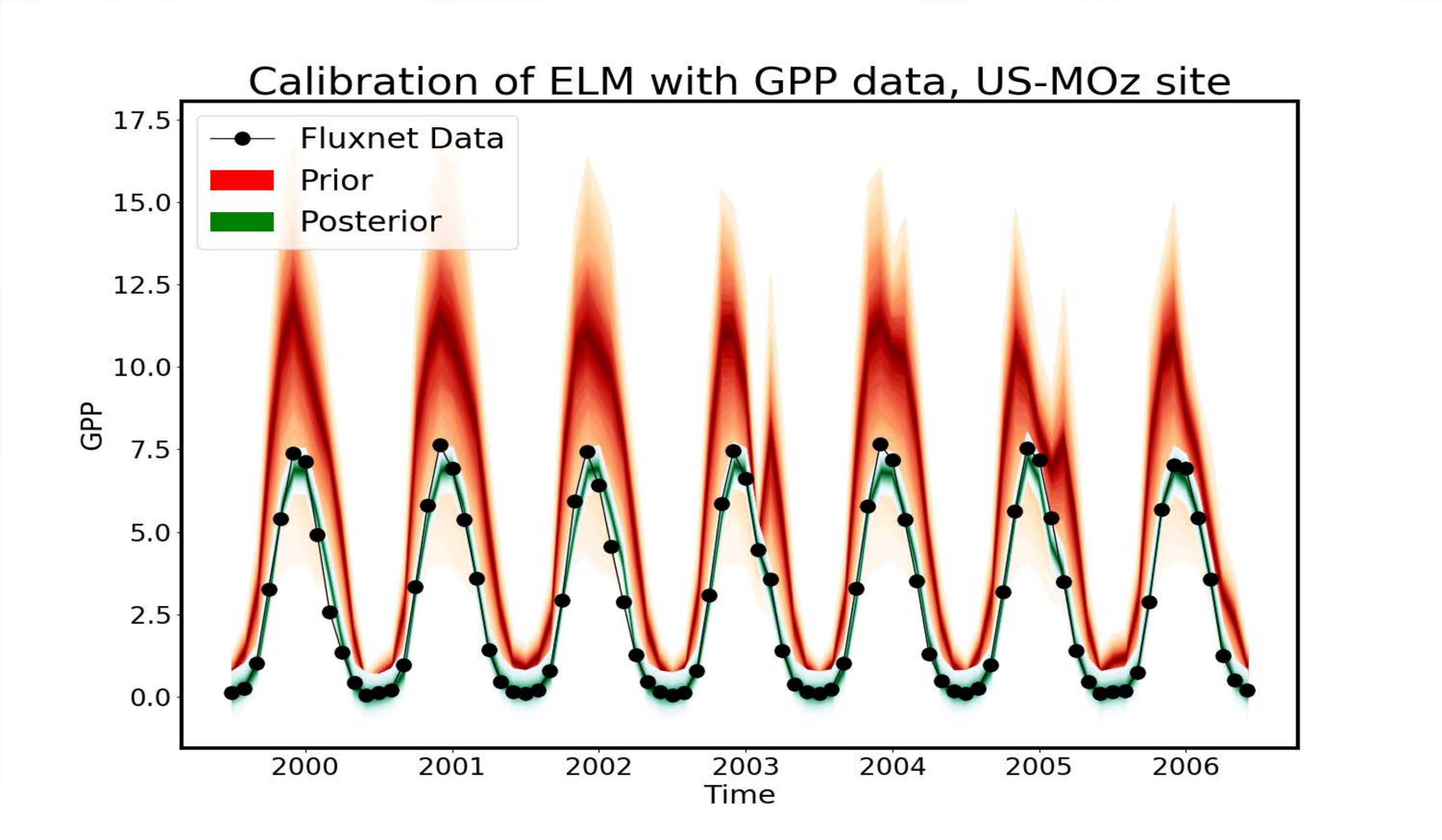
- *Inverse modeling:*
  - parameter tuning
  - calibration
  - data noise incorp.

Bayesian approach is main tool for parameter calibration

- Bayesian inference allows incorporation of various sources of uncertainty
- Markov chain Monte Carlo (MCMC) for building posterior PDFs
  - Non-linear models → Ugly high-dimensional parameter PDFs, but advanced MCMC methods are available
- Requires many online, mostly serial evaluations of the model
  - This is why surrogate models are handy!
- Predictive uncertainty decomposition augmented with surrogate error and observational noise



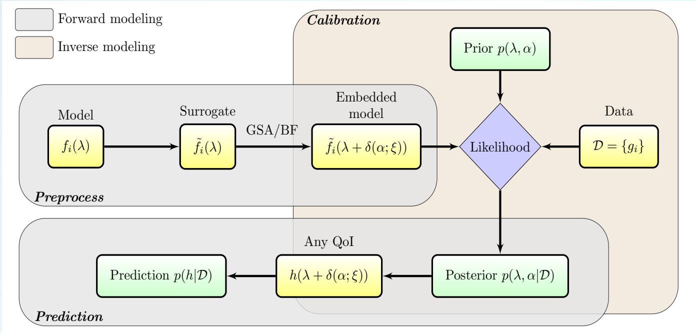
Initial calibration at a single gridcell



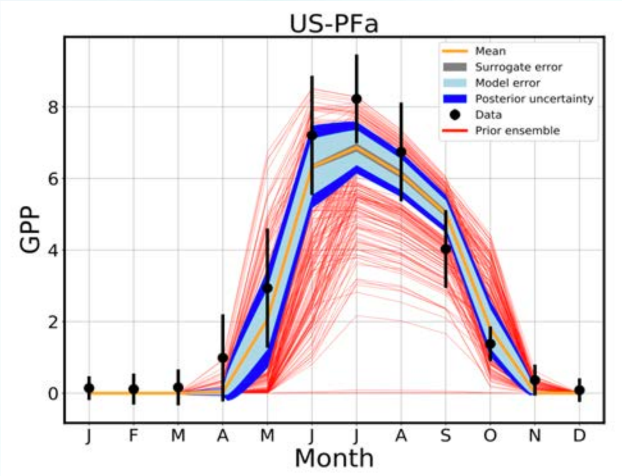
- Calibrated model is able to reproduce the observed with high accuracy
- Posterior uncertainties are greatly reduced
- Next test: By PFT and by gridcell calibrations.

Calibration with *Embedded* Model Structural Error

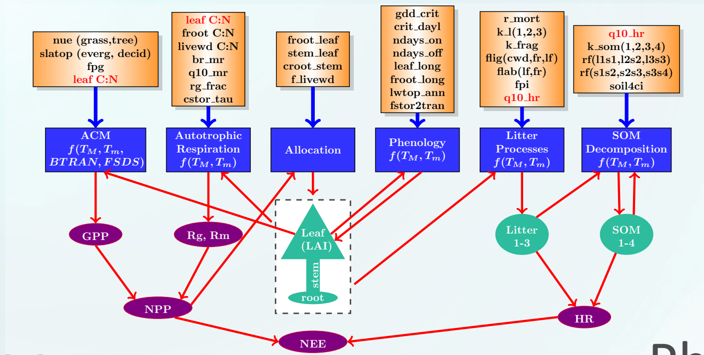
- Model structural error embedding approach *[Sargsyan et. al., 2015, 2018]*
  - Embedded, but not intrusive, i.e. black-box
  - Physics-driven model correction
  - Meaningful extrapolation to full set of QoI predictions
  - Disambiguation between model error and data noise
  - Core FASTMath capability



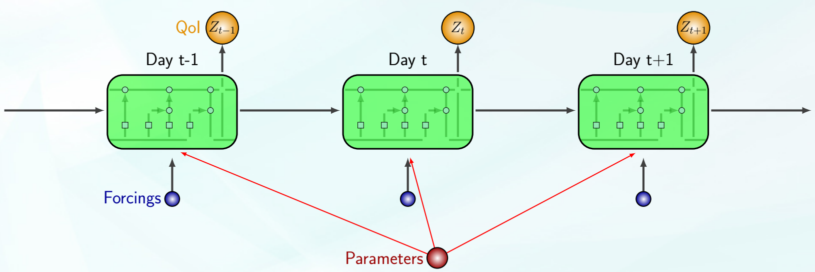
- Calibration of sELM with FLUXNET sites data
- Model error is the dominant uncertainty component
- Removes parameter biases and overfitting
- Points to submodels/parameters that are the culprit



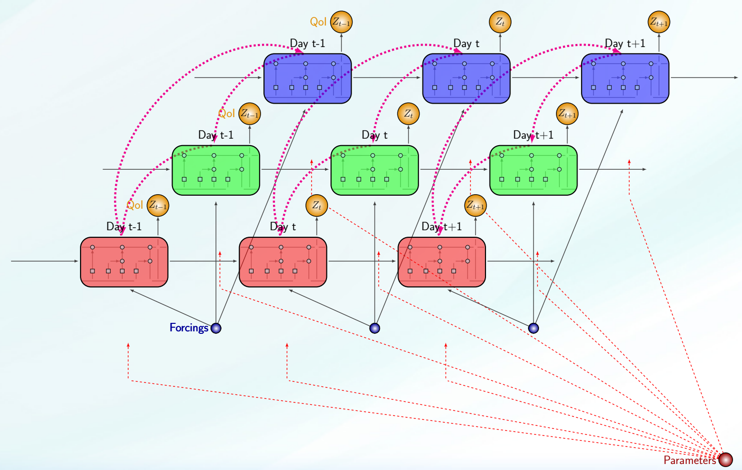
Improving surrogate models further: LSTM architecture complies with known physical connections



Vanilla LSTM:
each QoI is dealt with separately

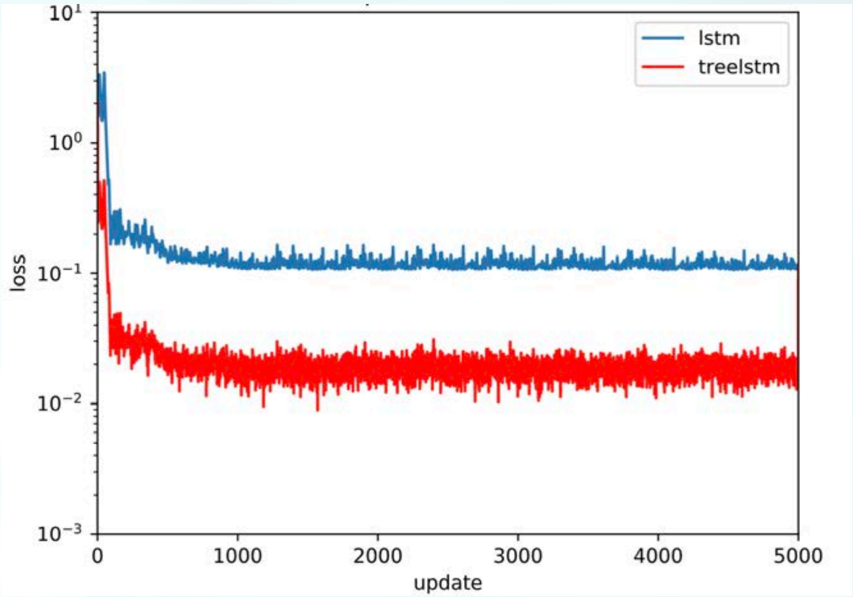


Physics-informed LSTM:
accounts for QoI connections

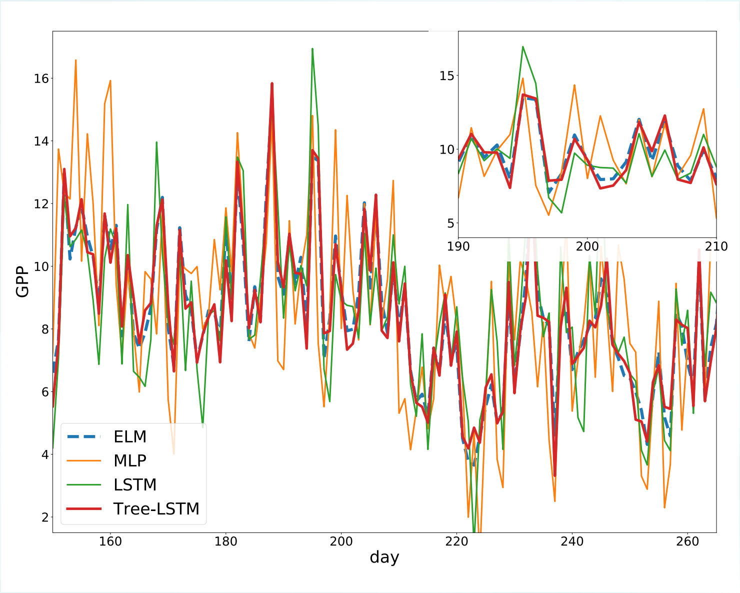


Physics-informed LSTM neural network accurately resolves time evolution

Loss function of vanilla LSTM and physics-informed LSTM

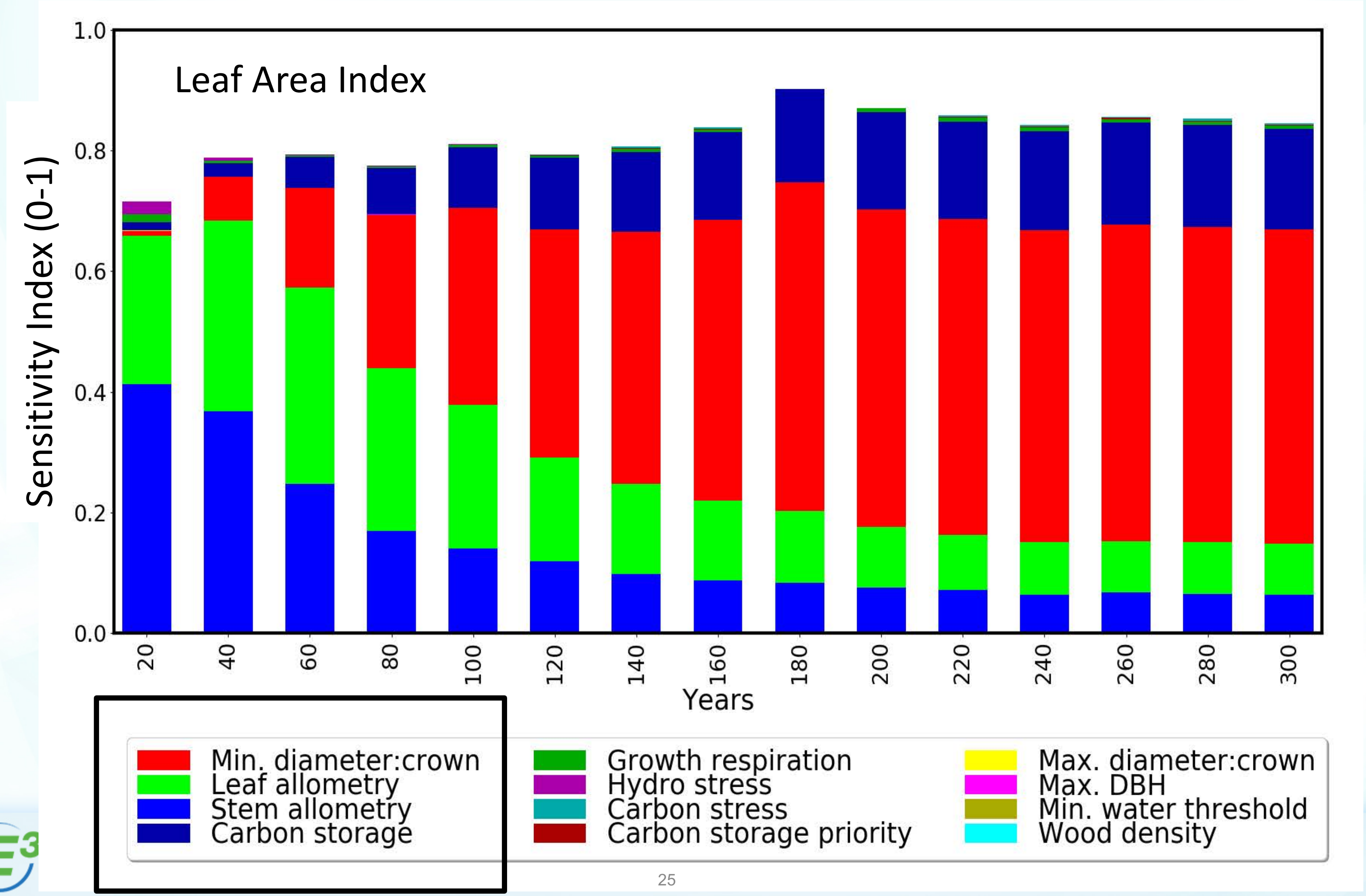


Comparison of sELM and NNs

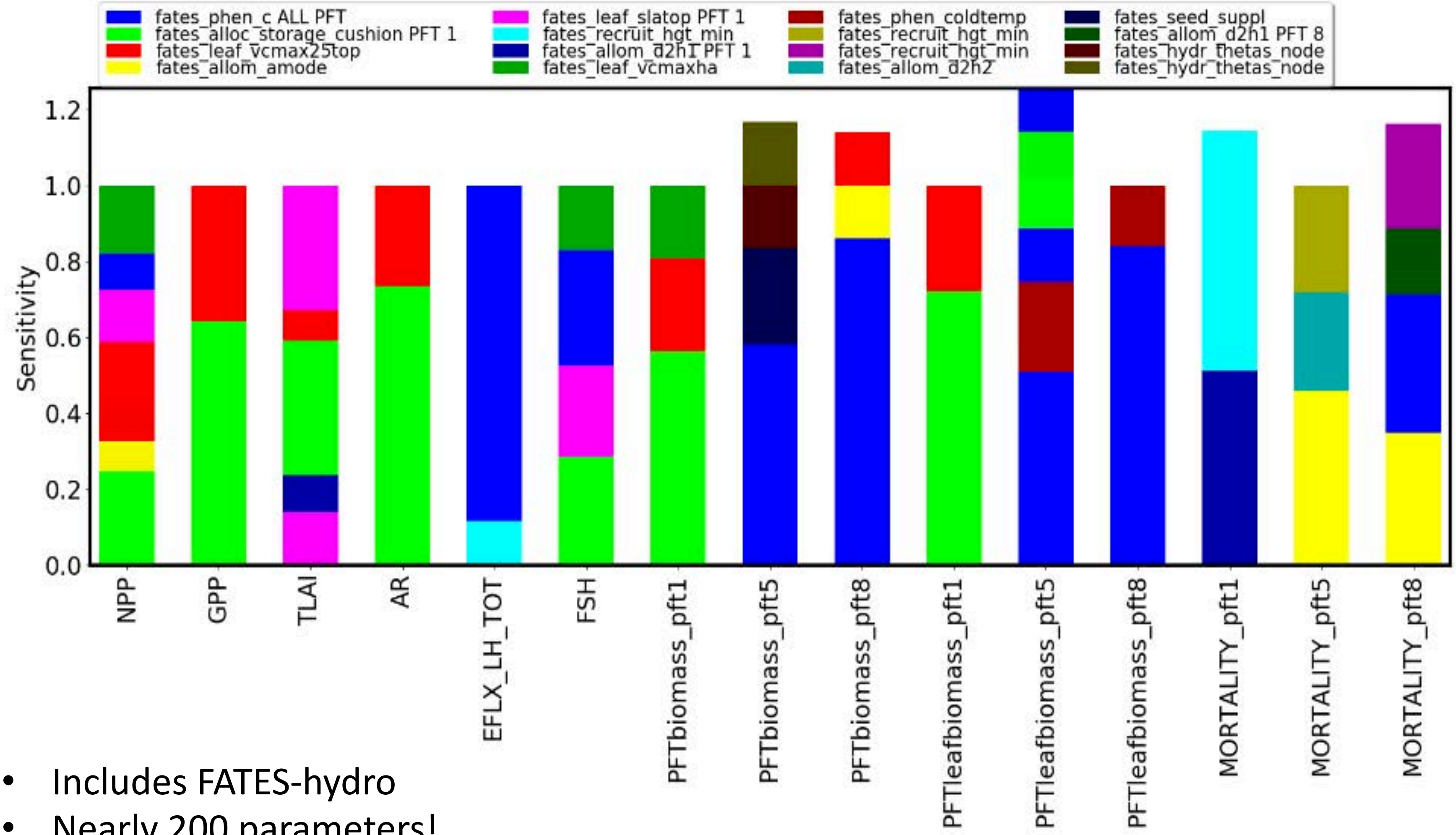


LSTM NNs approximates the sELM behaviour
with respect to perturbations in 47 parameters, with a fraction of the cost

Sensitivity analysis: ELM-FATES

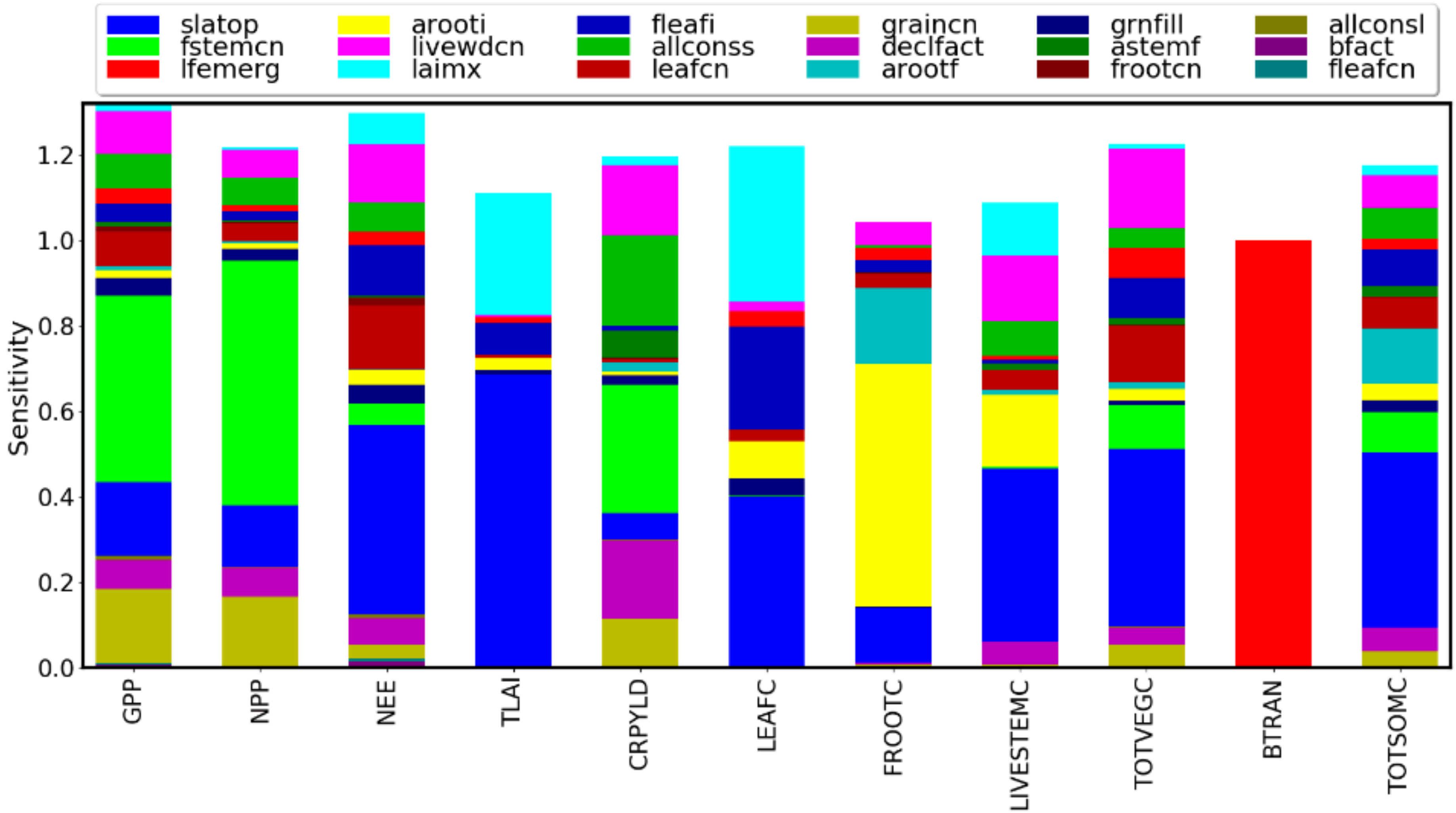


GSA across many QoIs, PFTs

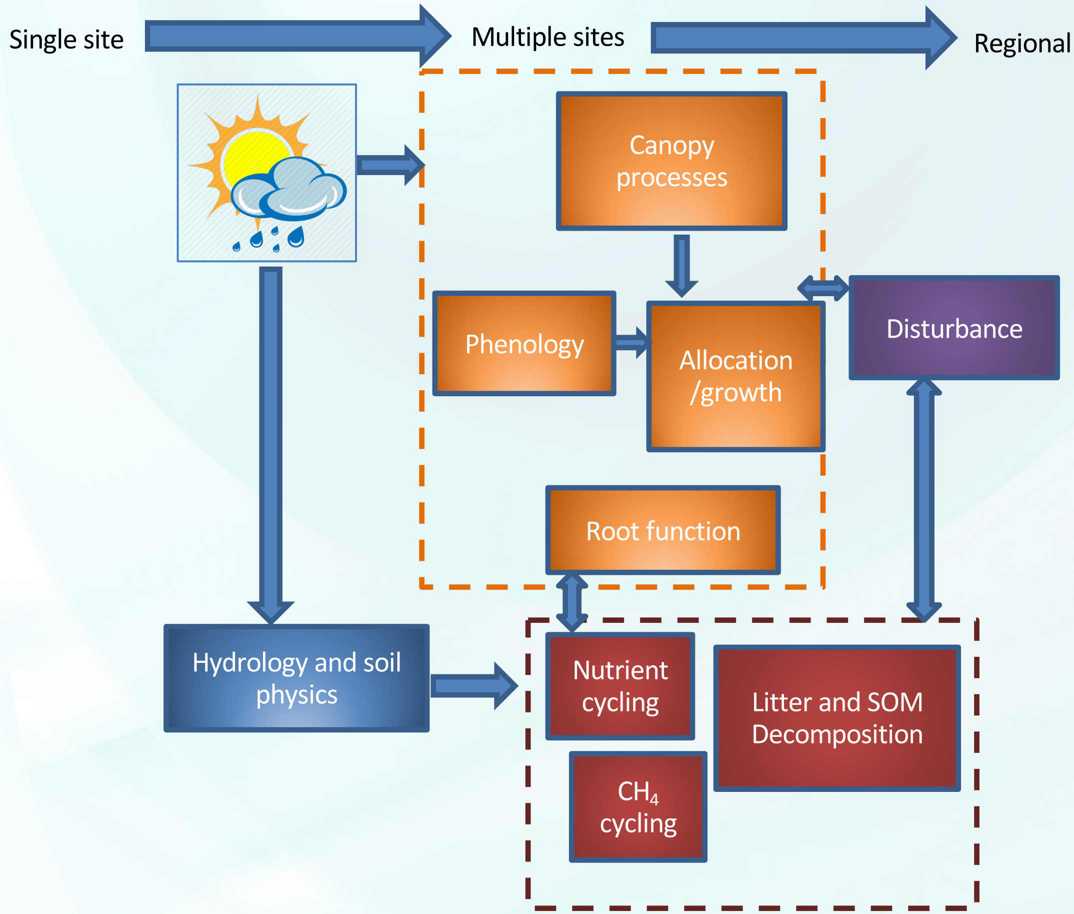


- Includes FATES-hydro
- Nearly 200 parameters!

Sensitivity analysis, ELM-CROP



An ELM UQ testbed

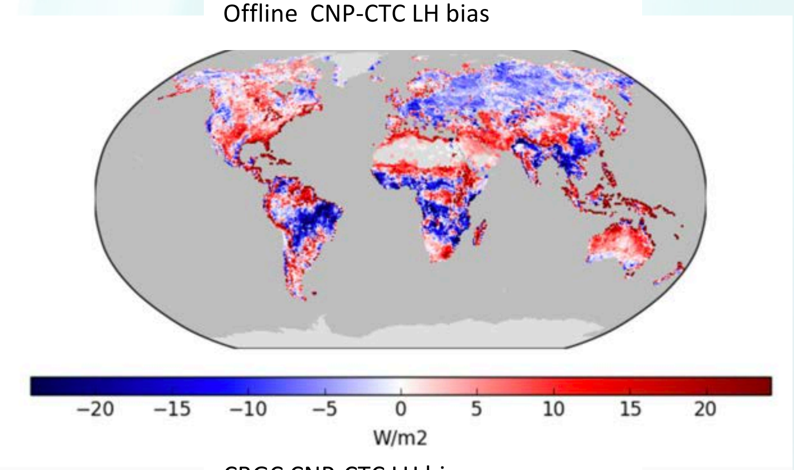


Joint work with the ORNL TES SFA

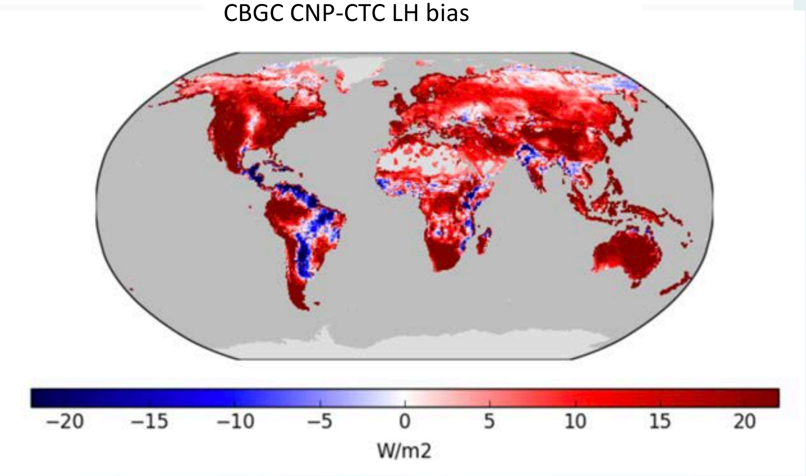
- Breaking down model UQ to manageable problems?
- Develop ELM “functional units” for process submodels
- Develop rapid evaluation capability using surrogates for:
  - Key individual model outputs
  - Each process submodel
- Hierarchical calibration
  - ELM complexity is high
  - Calibrate submodels using process-specific observations
  - Calibrate ELM using integrative observations (e.g. NEE).
- Enable ecological forecasting

Toward Coupled UQ

Offline CNP-CTC LH bias

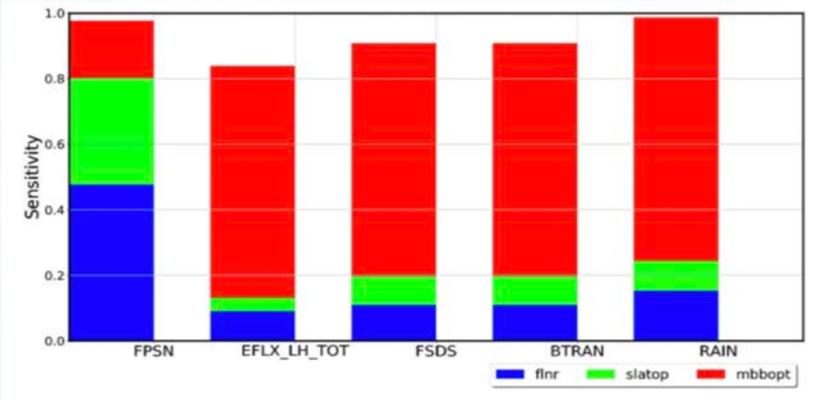


CBGC CNP-CTC LH bias



- Well-tuned/calibrated offline component models may perform poorly in coupled system
- Biases related to other components or coupling between them

- Large computational demand for individual experiments (including spinup)
- Additional outputs/dimensions – even larger ensembles needed.
- Machine learning --> meaningful UQ?
- First step in OSCM : SCM land-atmosphere



Sensitivity analysis of SCM at GOAmazon site

Summary

- **Forward UQ (uncertainty propagation):**
  - Surrogate modeling is the key
    - For point/site simulations, we have well-developed workflows and willing to work with core, NGD and ecosystem projects.
    - With a combination of approaches, we can achieve high surrogate accuracy ***with global relevance*** using a relatively small number of simulations.
  - Global sensitivity analysis or variance decomposition for parameter dimension reduction
  - Extending global work to BGC: Possible but computational resources needed

- **Inverse UQ (parameter calibration):**
  - Bayesian model calibration with Markov chain Monte Carlo (MCMC) sampling
  - Expense is alleviated by using an accurate surrogate: makes ELM calibration feasible
  - Key advances: Spatially explicit surrogates, incorporating model structural error

- **"MODEX" loop enabler**: use attributable model prediction uncertainties to optimally locate new observation locations.