Machine Learning-based Parameterizations

Nuno Carvalhais, Bernhard Ahrens, Shanning Bao,
Fabian Gans, Basil Kraft, Rackhun Son, Siyuan Wang,
Lazaro Alonso and Markus Reichstein

[ncarvalhais@bgc-jena.mpg.de]

Max Planck Institute for Biogeochemistry

ellis European Laboratory for Learning and Intelligent Systems

rationale

Forward view:

model + forcing + parameters -> model_data

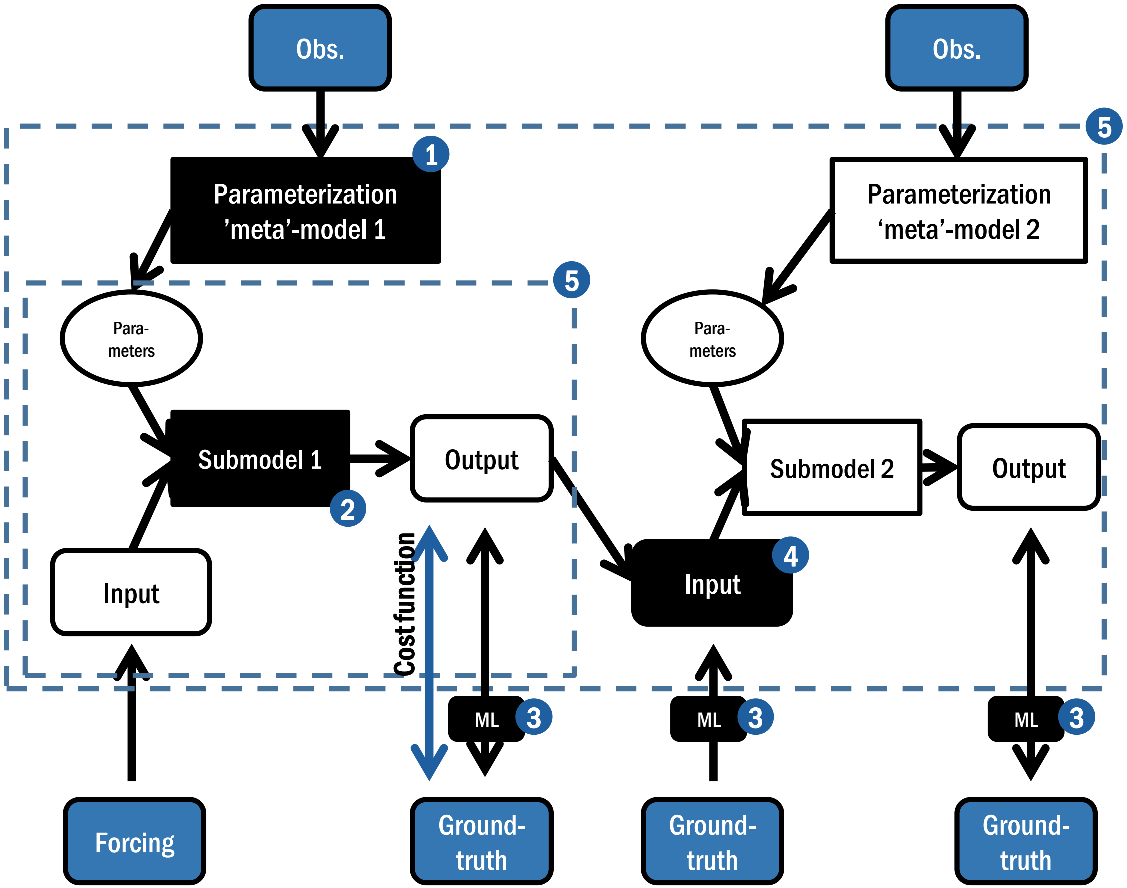
Inverse view:

model + forcing + obs_data -> parameters + model_data

Hybrid view:

forcing + obs_data -> model + parameters + model_data

Model-data-machine-learning integration…



1 - Model parameterization
2 - Hybrid modelling
3 - Pattern-oriented model evaluation and calibration
4 - Driving a model with machine learning output
5 - Model emulation

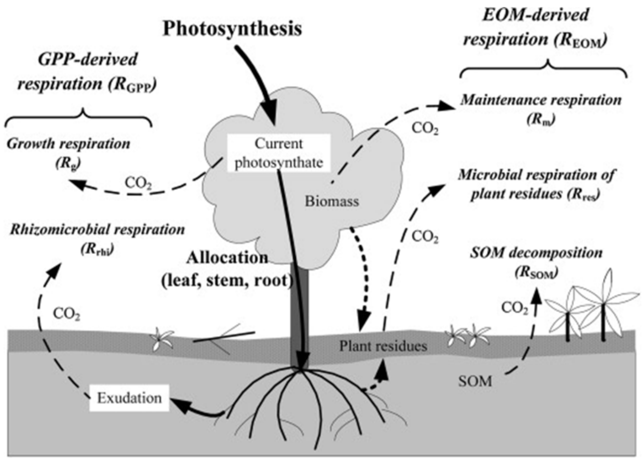
[Reichstein et al., 2019]

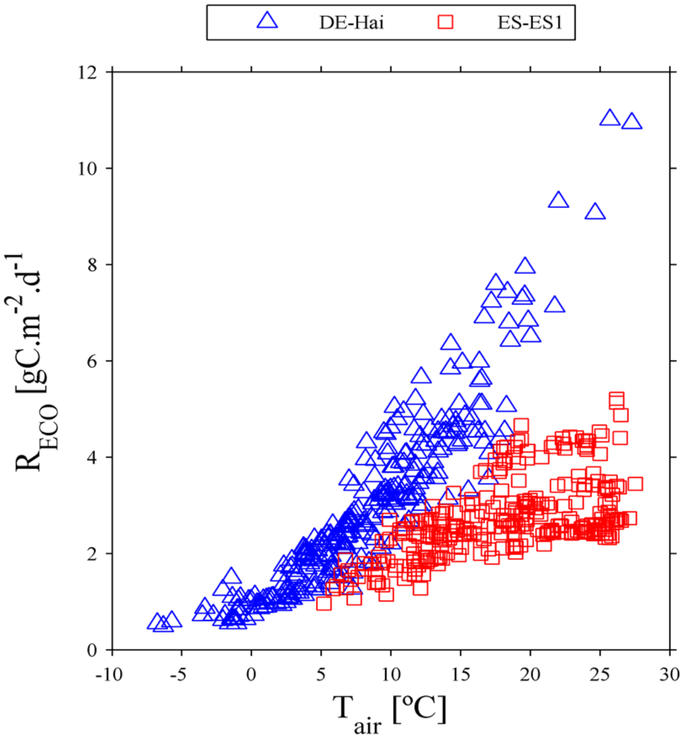
EXPERIMENTAL RESULTS

[Reichstein et al., 2022]

RESPIRATION TEMPERATURE SENSITIVITY

Simple respiration modeling example





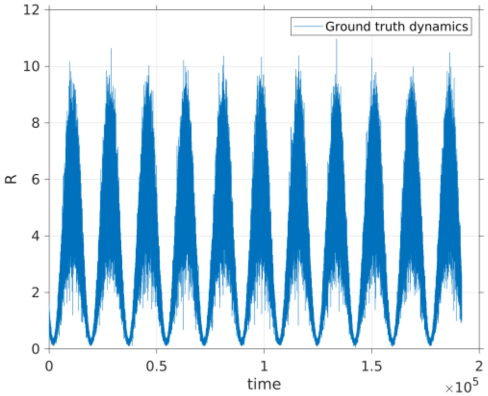
$$Resp(t) = \underbrace{R_b(t)}_{\text{Biological "base" activity = f(all kind of potential factors)}} * \underbrace{Q_{10}^{\frac{T(t)-T_{ref}}{10}}}_{\text{Physico-chemical temperature dependence}}$$

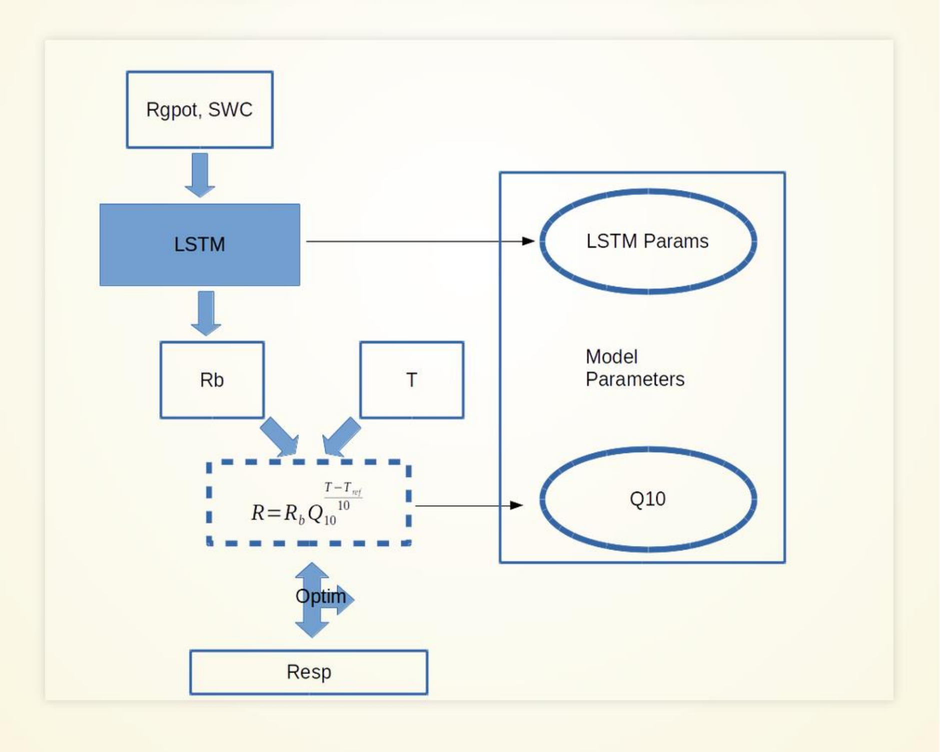
Biological "base" activity = f(all kind of potential factors)

Physico-chemical temperature dependence

experiment

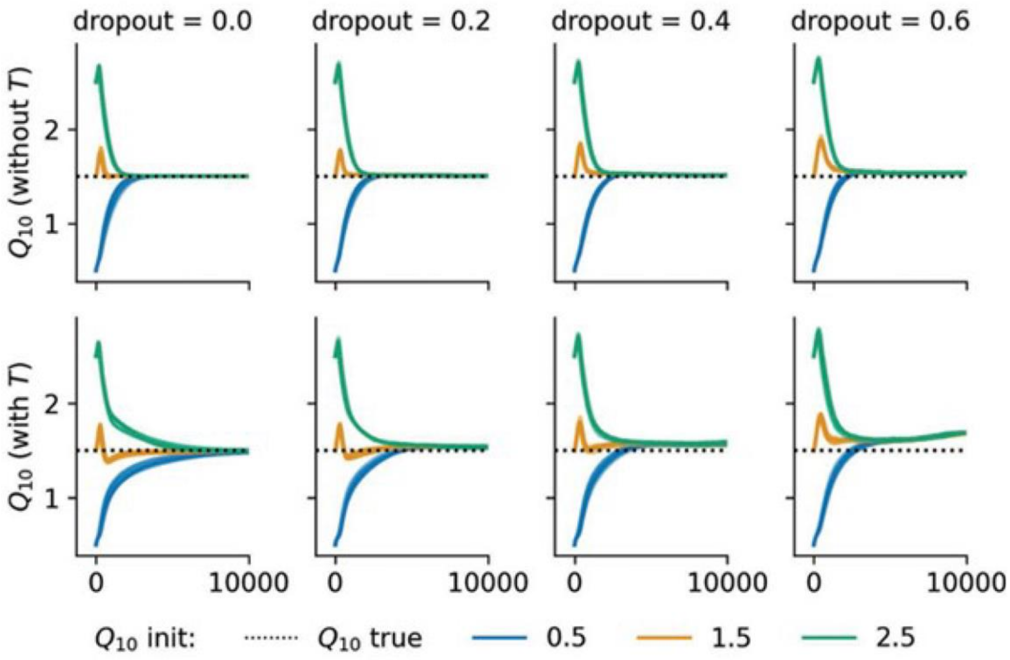
- Resp = $Q10^{((t-15)/10)}$ x f(Rgpot,SM)
- Resp = $Q10^{((t-15)/10)}$ x NN(X)

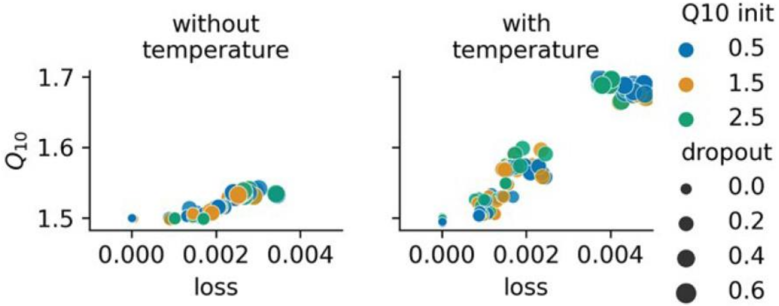




[Gans et al., in prep.]

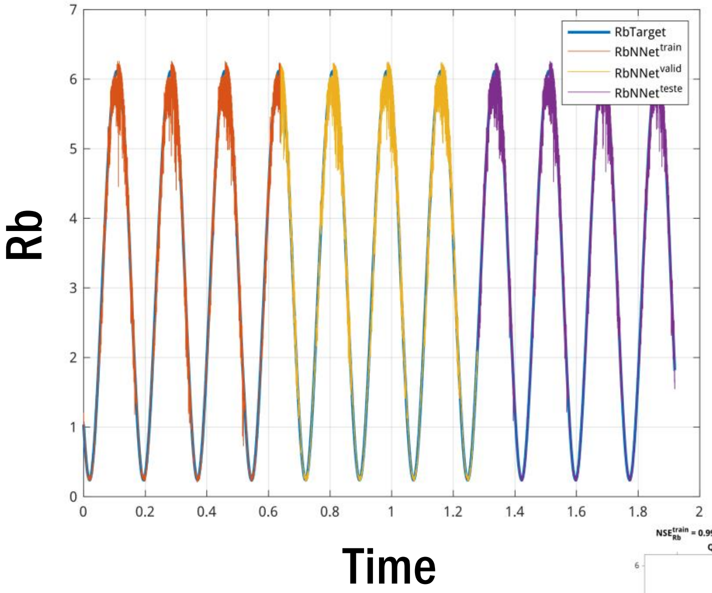
Q10 results

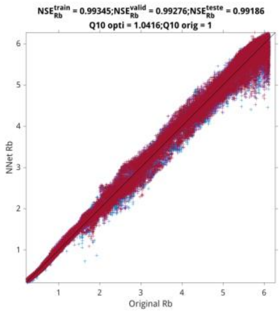


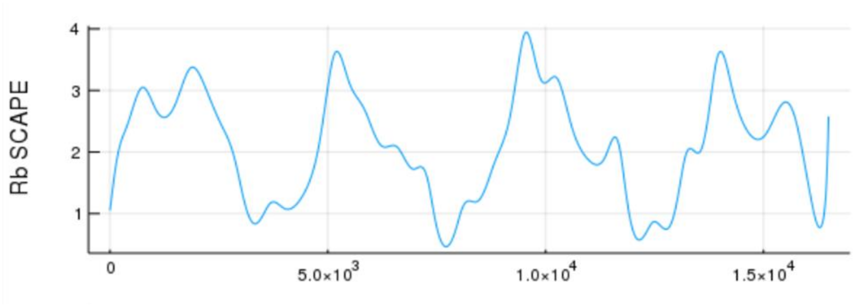


- Tight constraints on Q10
- Error relates to bias
- Low (no) impact of initial value

Emerging Rb dynamics







[Mahecha et al., 2010]

Generating patterns on daily/subdaily Rb dynamics [Gans et al., in prep]

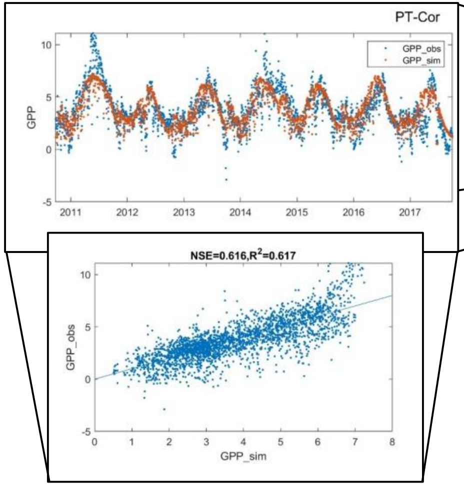
Contributing hypothesis for internal/external constrols on dynamics of substrate availability [Ahrens et al., in prep]

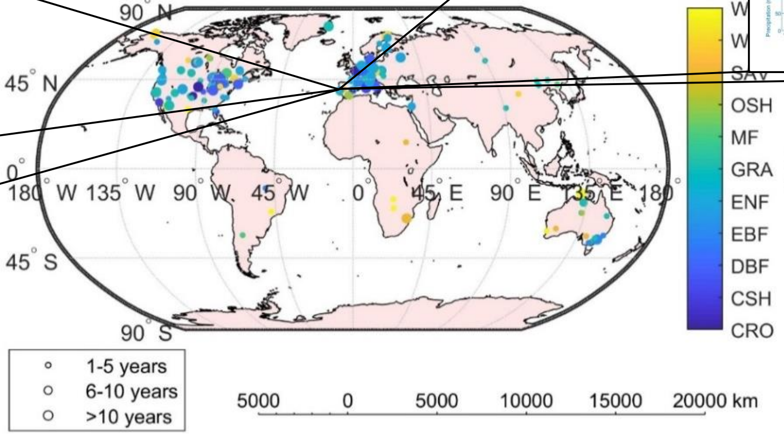
[Bao et al., in prep.]

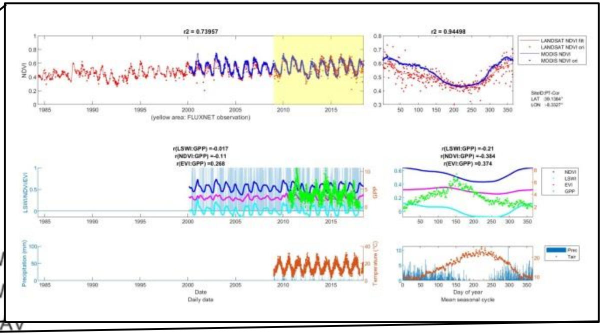
SPATIAL EXTRAPOLATION OF PHOTOSYNTHESIS PARAMETERS

$$GPP = \varepsilon_{max} \cdot PAR \cdot FAPAR \cdot fT \cdot fVPD \cdot fW \cdot fL \cdot fCI$$

- Semi-empirical descriptions of “f”
  - Sensitivity of ecosystem GPP to different forcing



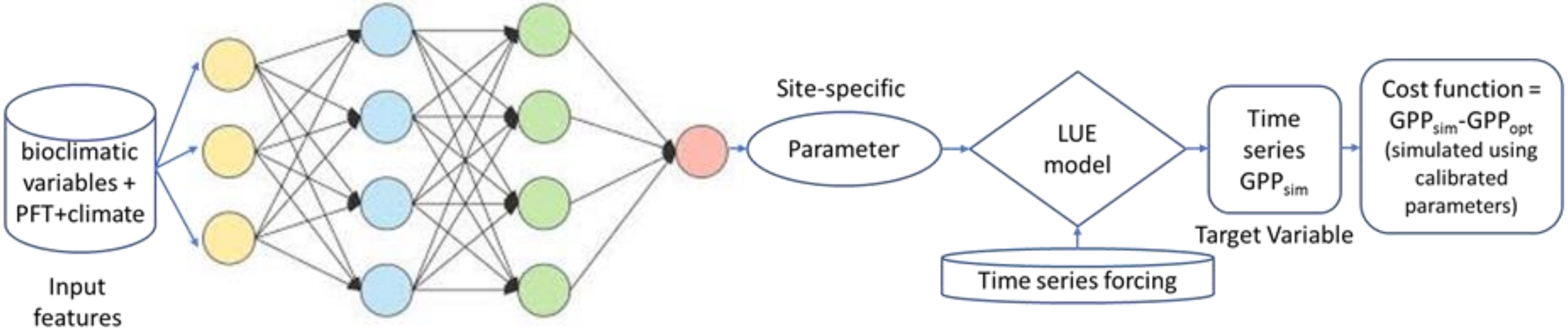




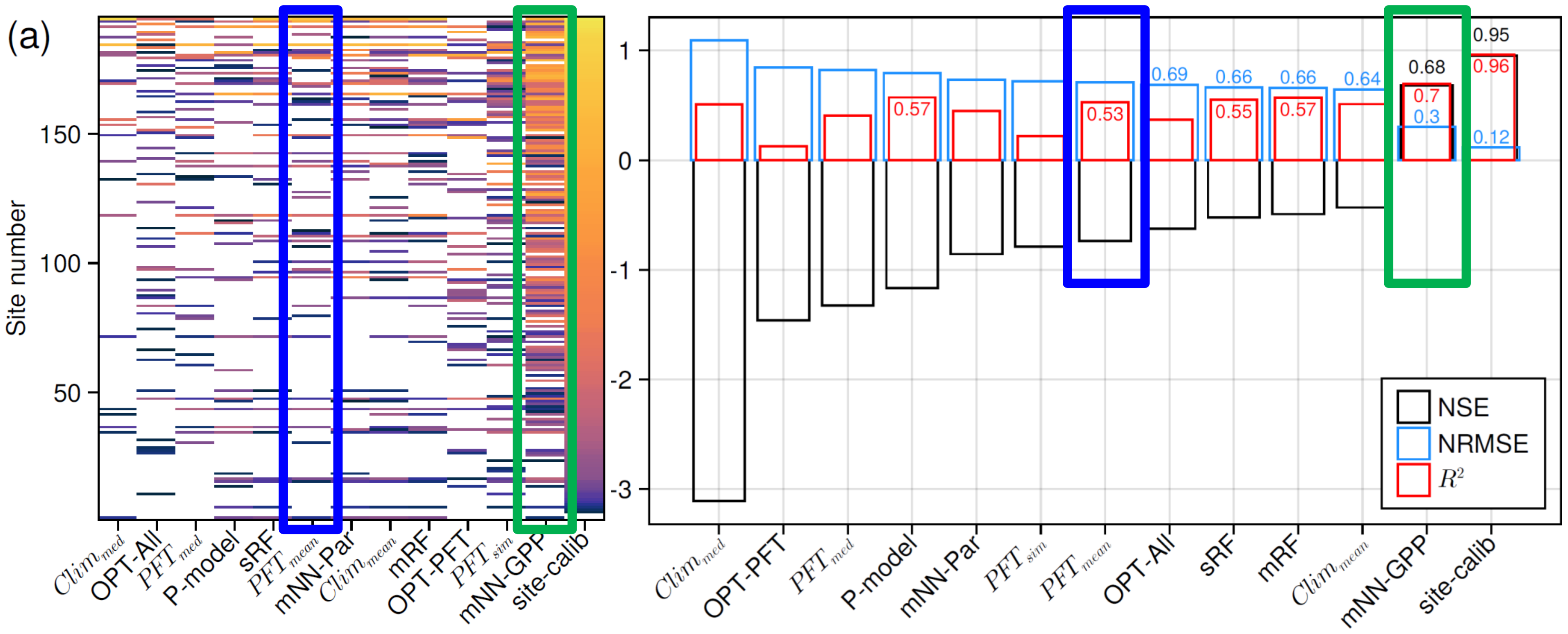
[Bao et al., 2021]

what controls parameter variability?

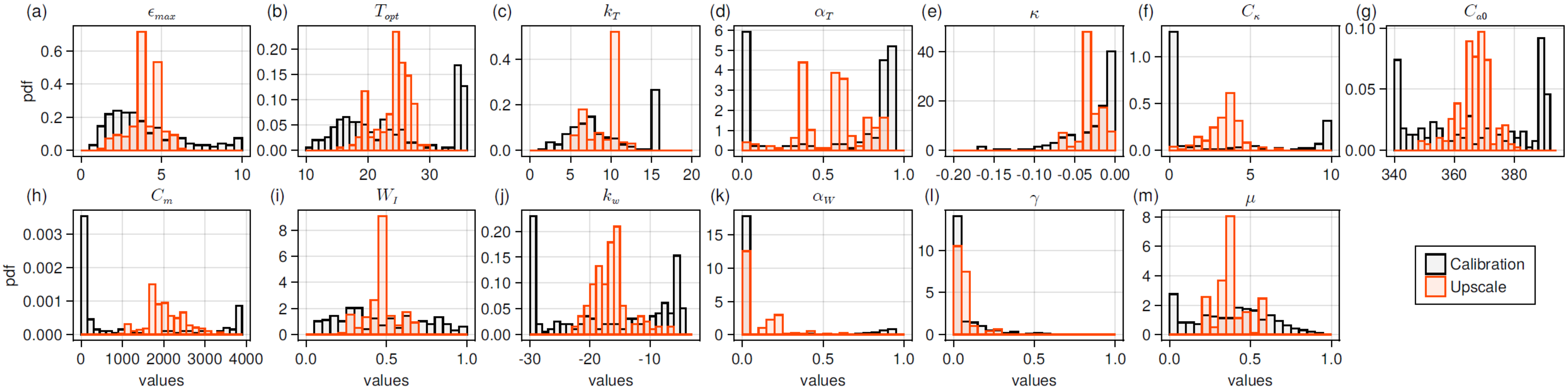
PFT? Climate? Soil properties? All? More?



cross-validated model performance

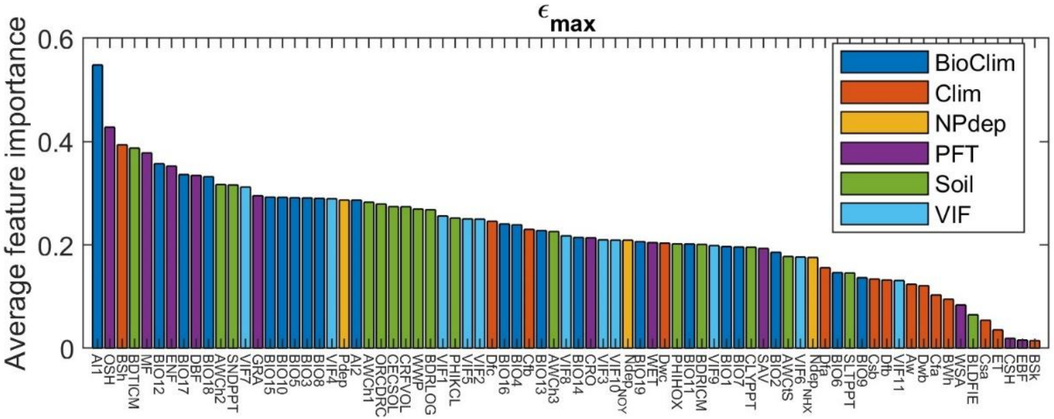


resulting constraints in parameters



large reduction in spatial variance of parameters

drivers of parameter variability



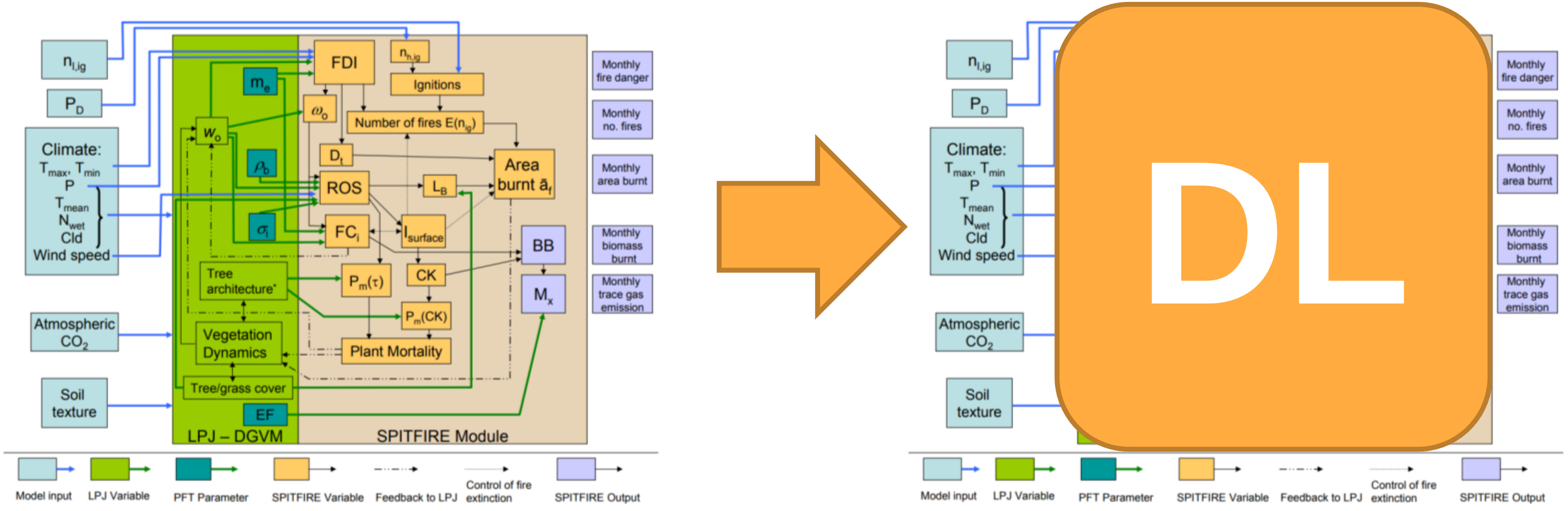
- Several classes of factors used to predict parameter variability
- Link between feature dependencies and long term responses to local conditions?

[Son et al., in prep.]

HYBRID FIRE MODELING

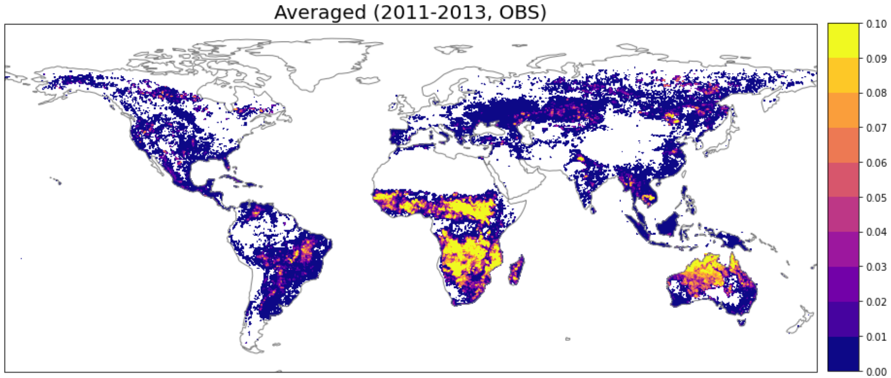
motivation

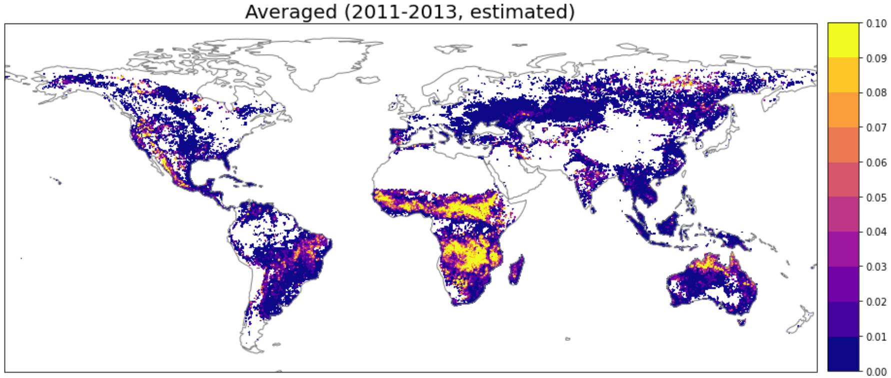
- Challenges in ESM fire modeling

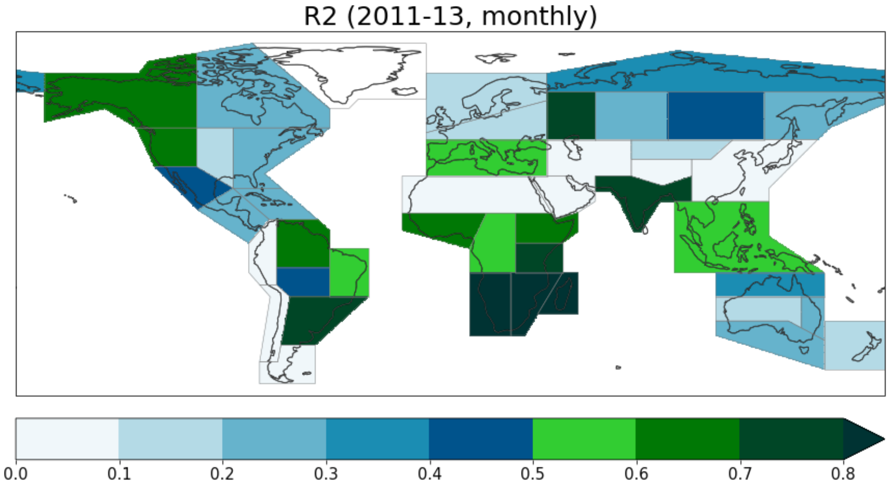


K. Thonicke et al.,(2010) The influence of vegetation, fire spread and fire behavior on biomass burning and trace gas emissions: results from a process-based model.

Results: evaluation period





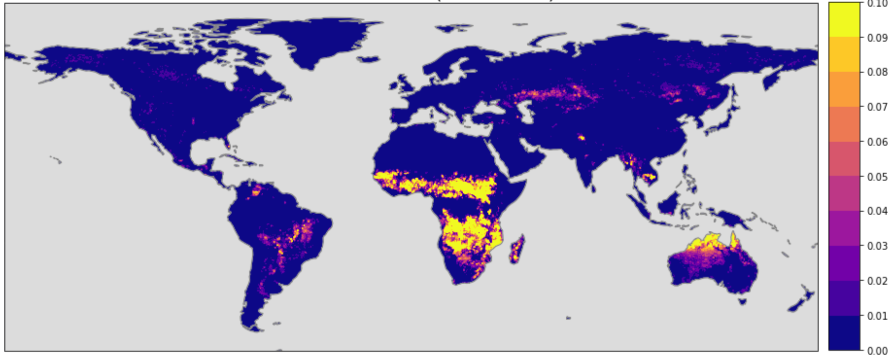


- Global spatial patterns
- Regionally variable temporal perfomance

JSB4-DL simulation (2001-13)

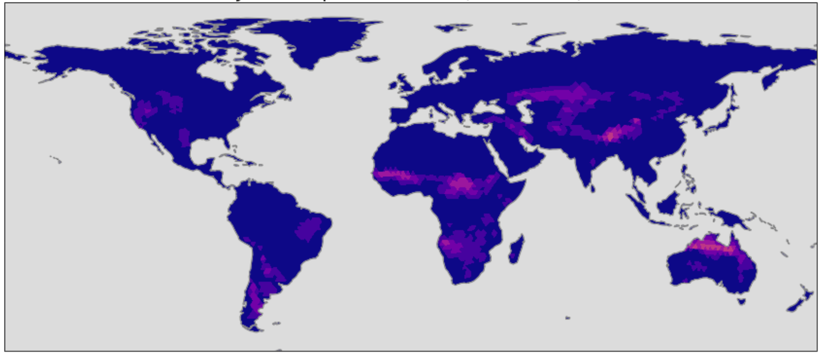
OBS

FBA from GFED4 (2001-2013)



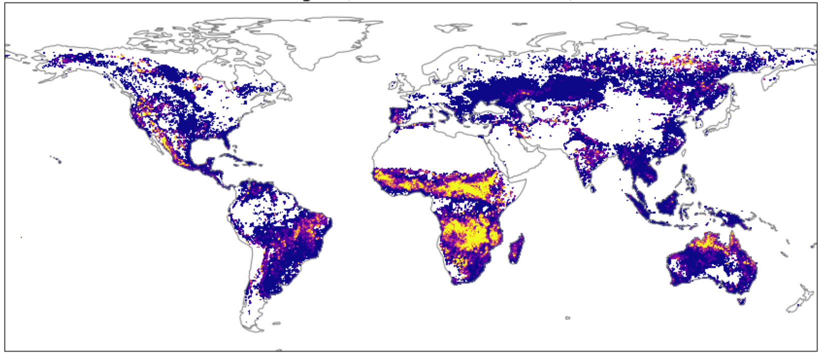
JSB

JSB4 simple fire model (2001-2013)



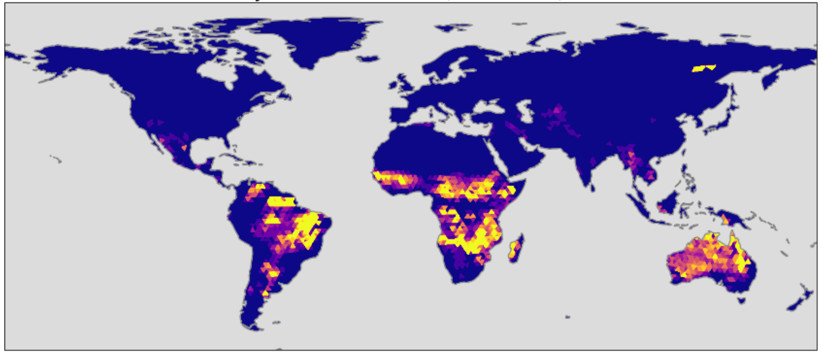
OBS-DL

Averaged (2011-2013, estimated)

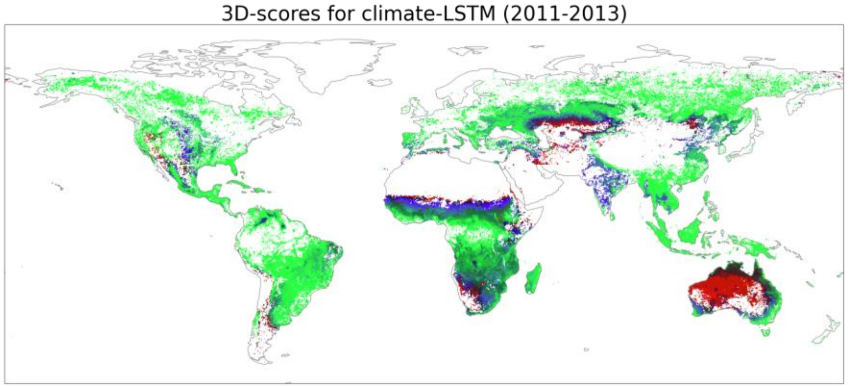


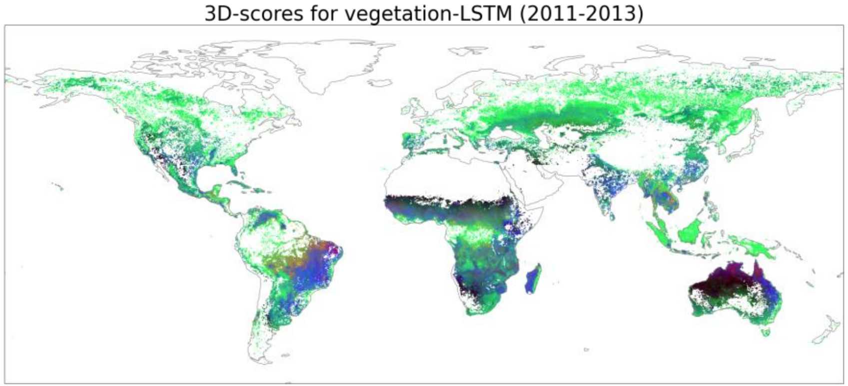
JSB-DL

JSB4 DL fire model (2001-2013)



emerging responses





Here, scores as a measure of complexity

Explore fuel and atmospheric construtions to burned area

Contrast, e.g., South America with Africa, East-West patterns in Asia

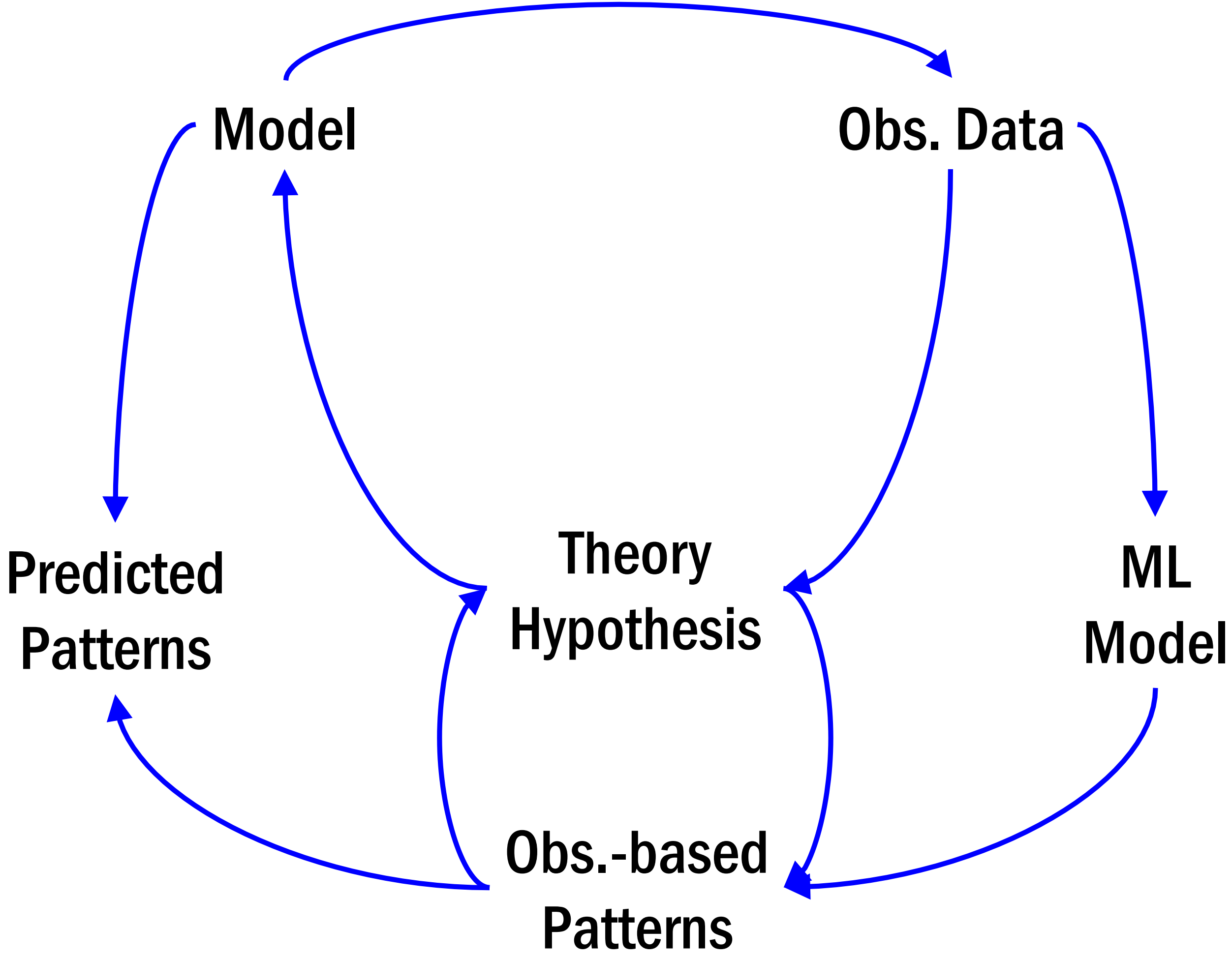
OVERALL

current results

- Support parameterizations, abstract structural sub-model limitations, correct biases

- Hybrid modeling as an approach for expanding information content on physically-bounded models

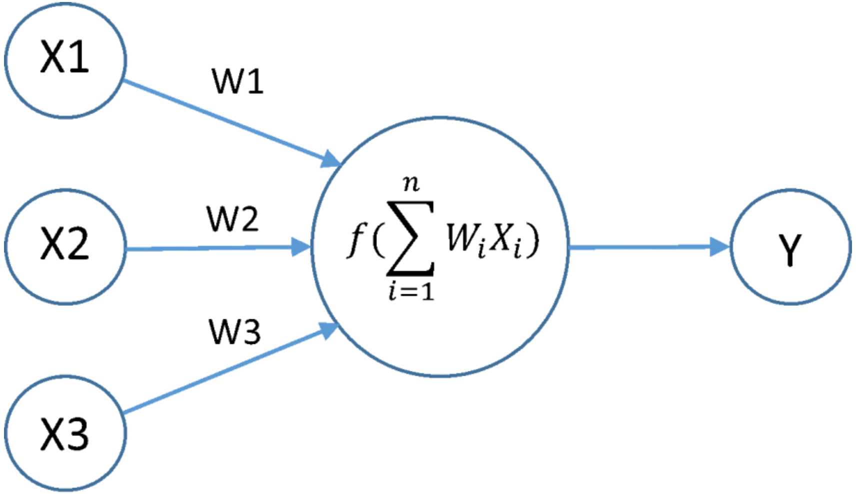
HYPOTHESIS-DRIVEN / DATA-DRIVEN SCIENCE

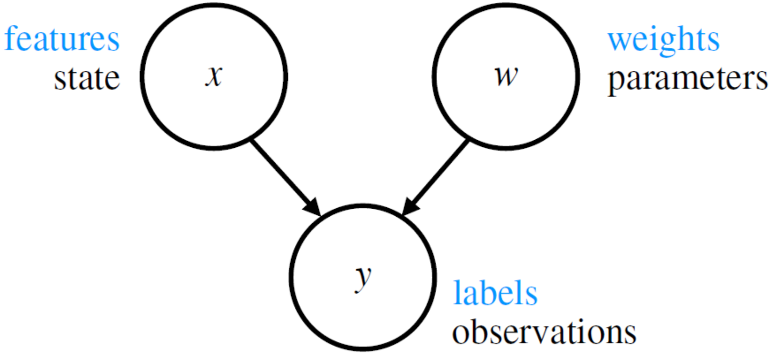


[adapted from Reichstein et al., 2019]

THANK YOU!

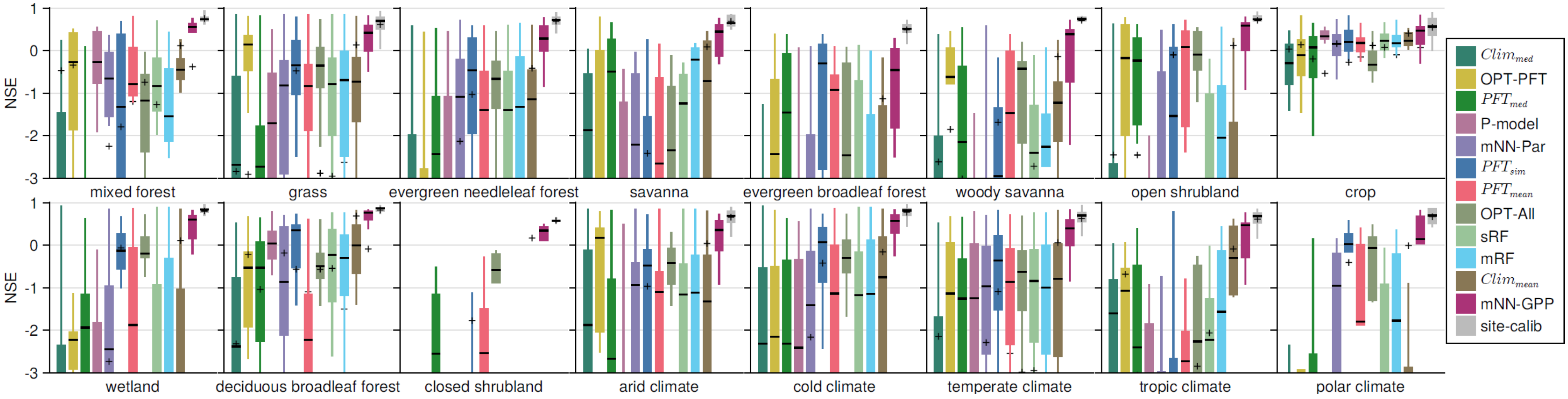
Basic building block: a single neuron





[Geer, 2020]

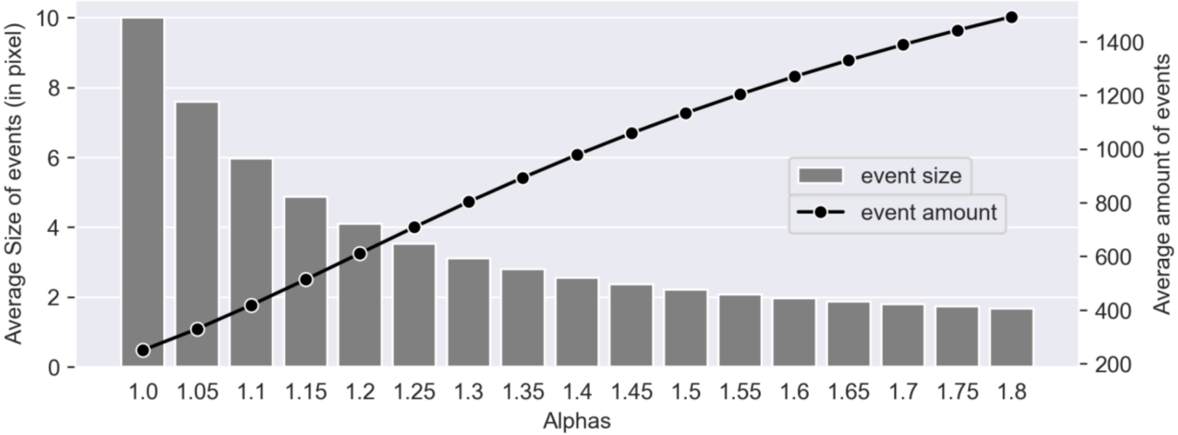
cross-validated model performance across climate/PFT classes



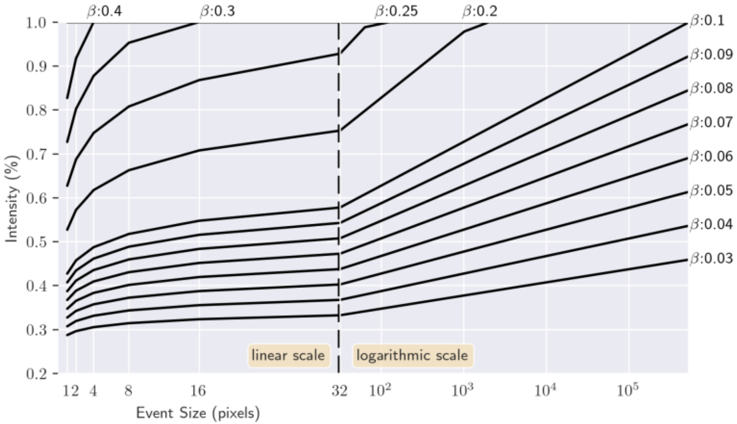
Wang et al., 2022 (pre-print)

DISTURBANCE REGIMES

observational synthesis on disturbance regimes

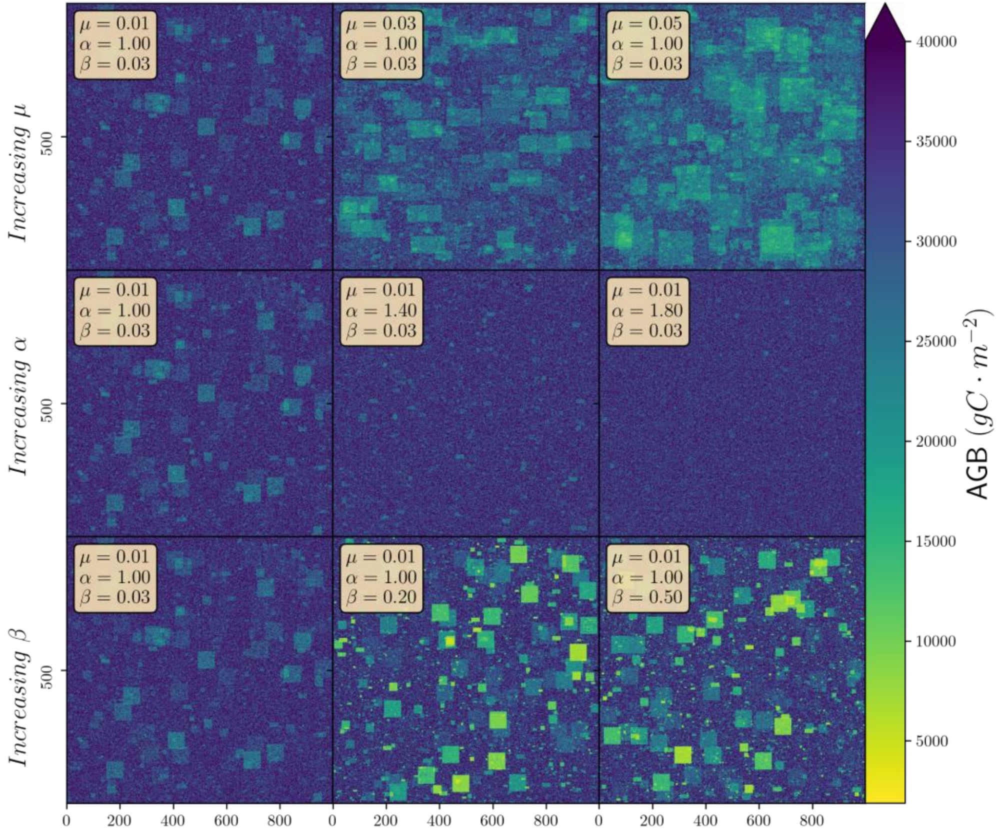


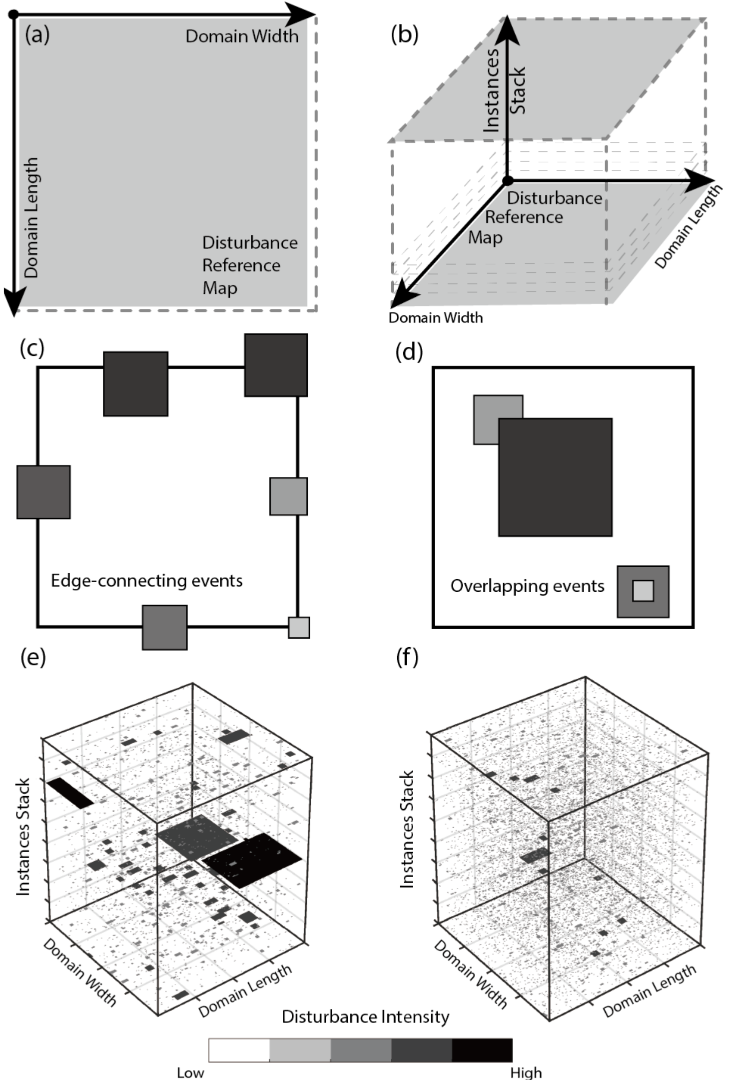
[Fisher et al., 2008]



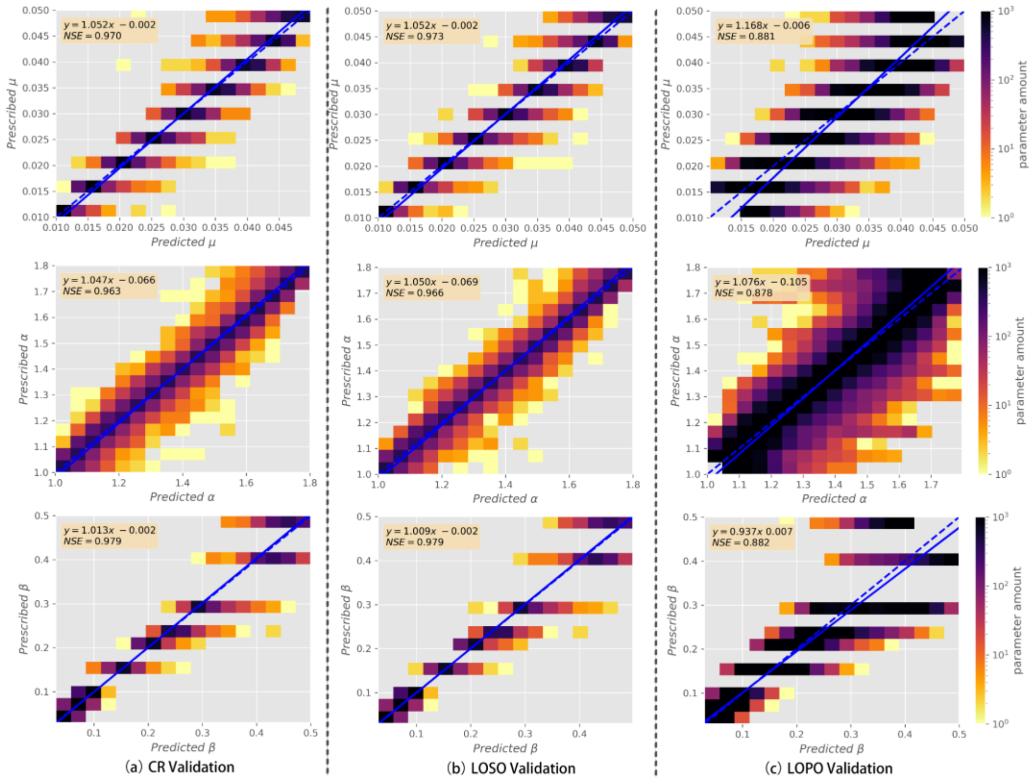
[Chambers et al., 2013]

Regime -> sp. pattern

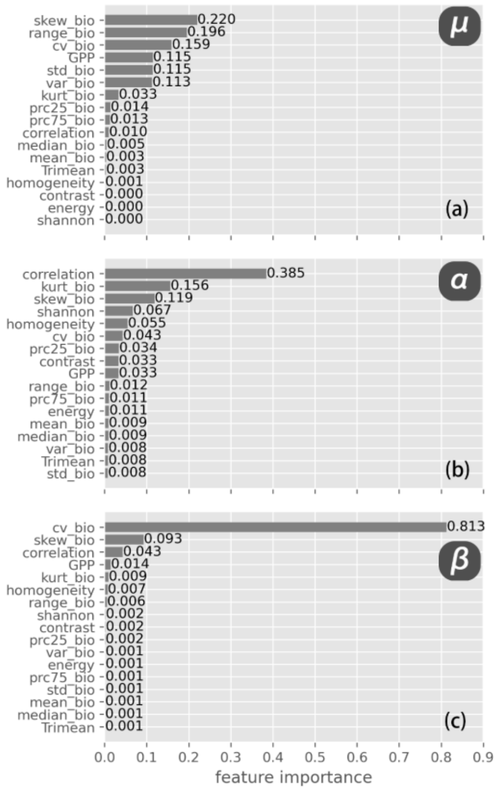




[μ, α, β] = ML(AGB)



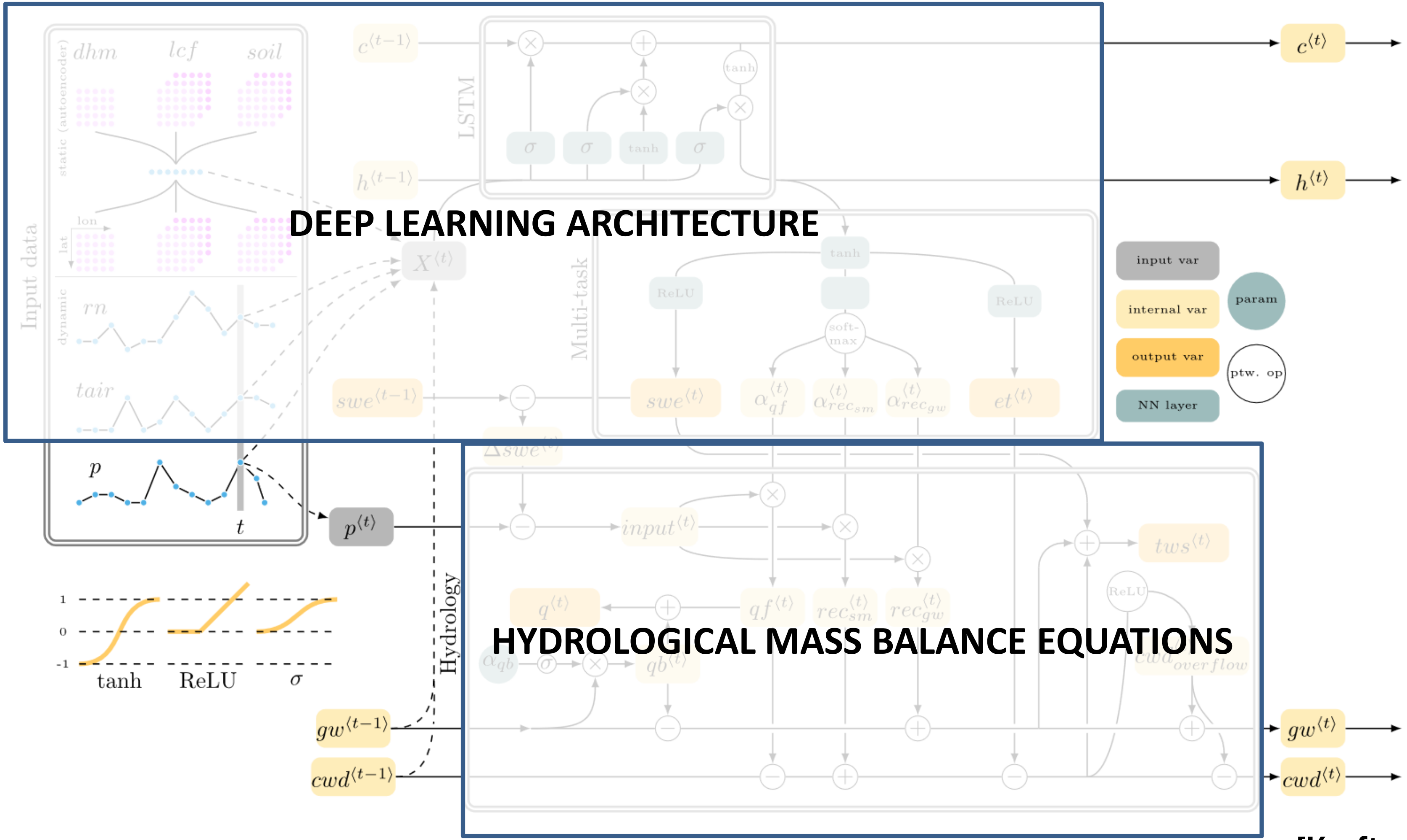
y = 1.052x − 0.002
NSE = 0.970
y = 1.052x − 0.002
NSE = 0.973
y = 1.168x − 0.006
NSE = 0.881
y = 1.047x − 0.066
NSE = 0.963
y = 1.050x − 0.069
NSE = 0.966
y = 1.076x − 0.105
NSE = 0.878
y = 1.013x − 0.002
NSE = 0.979
y = 1.009x − 0.002
NSE = 0.979
y = 0.937x 0.007
NSE = 0.882
Prescribed μ
Predicted μ
Prescribed α
Predicted α
Prescribed β
Predicted β
parameter amount
(a) CR Validation
(b) LOSO Validation
(c) LOPO Validation



μ
skew_bio 0.220
range_bio 0.196
cv_bio 0.159
GPP 0.115
std_bio 0.115
var_bio 0.113
kurt_bio 0.033
prc25_bio 0.014
prc75_bio 0.013
correlation 0.010
median_bio 0.005
mean_bio 0.003
Trimean 0.003
homogeneity 0.001
contrast 0.000
energy 0.000
shannon 0.000
(a)
α
correlation 0.385
kurt_bio 0.156
skew_bio 0.119
shannon 0.067
homogeneity 0.055
cv_bio 0.043
prc25_bio 0.034
contrast 0.033
GPP 0.033
range_bio 0.012
prc75_bio 0.011
energy 0.011
mean_bio 0.009
median_bio 0.009
var_bio 0.008
Trimean 0.008
std_bio 0.008
(b)
β
cv_bio 0.813
skew_bio 0.093
correlation 0.043
GPP 0.014
kurt_bio 0.009
homogeneity 0.007
range_bio 0.006
shannon 0.002
contrast 0.002
prc25_bio 0.002
var_bio 0.001
energy 0.001
prc75_bio 0.001
std_bio 0.001
mean_bio 0.001
median_bio 0.001
Trimean 0.001
(c)
feature importance

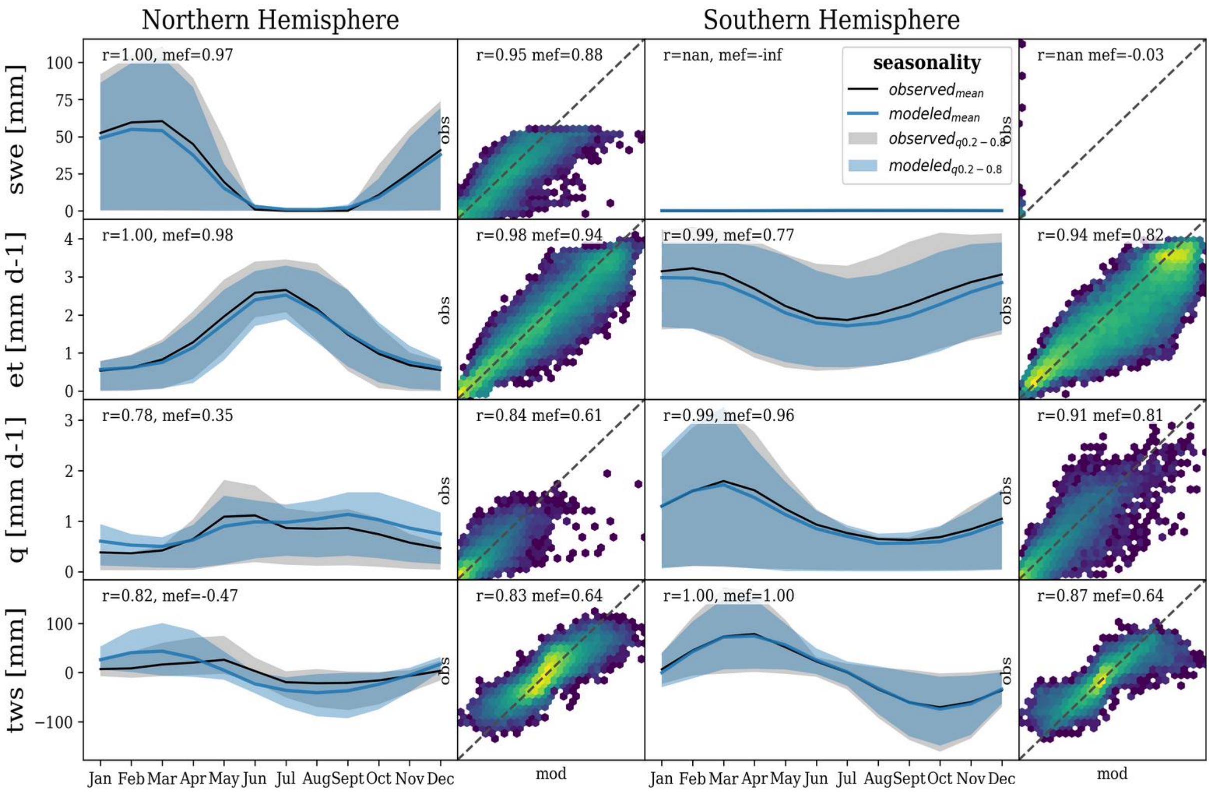
[Kraft et al., 2020-2022]

GLOBAL HYDROLOGICAL MODEL



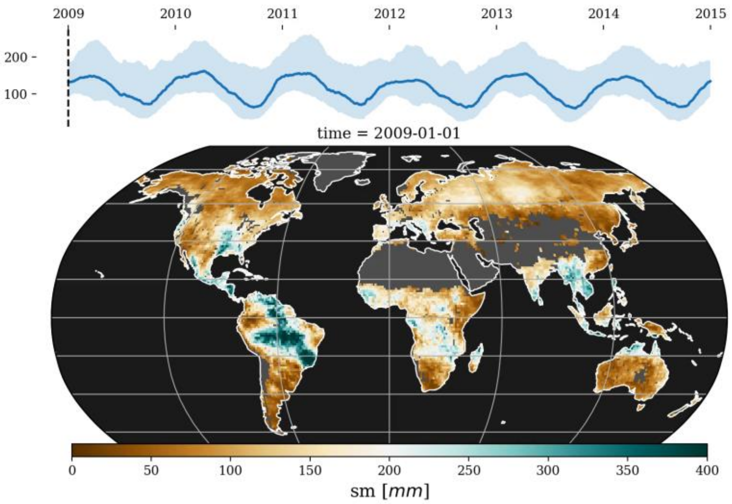
[Kraft et al. 2020]

aggregated performance

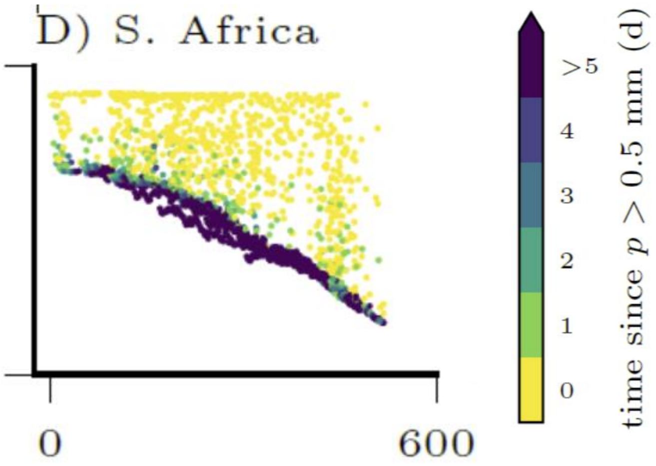


[cf. Kraft et al. 2020-22]

diagnostic of water cycle components



Root zone soil moisture



Evaporative fraction

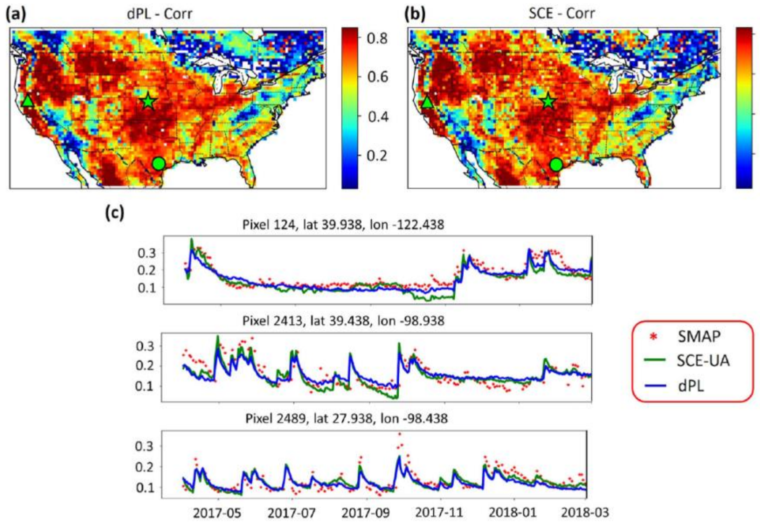
Soil water deficit [mm]

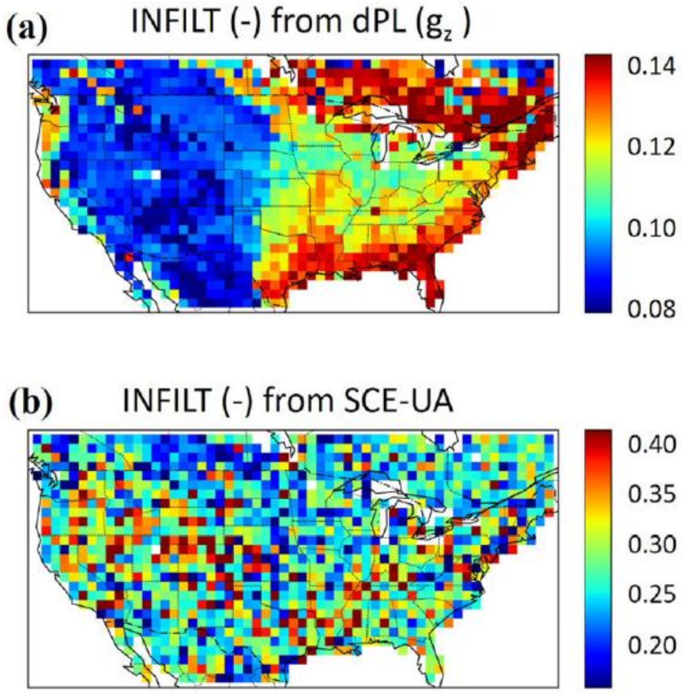
[Kraft et al. 2021, HESSD]

[Tsai et al., 2021]

SPATIAL EXTRAPOLATION OF HYDROLOGICAL PARAMETERS

Hydrological parameter upscaling...





[Tsai et al., 2021]