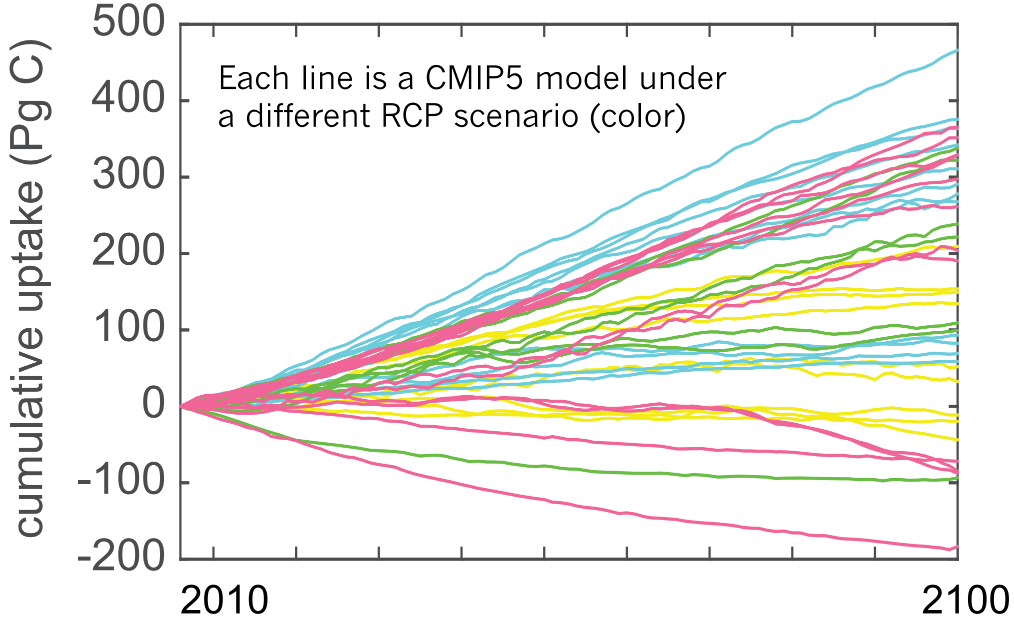
Prescribing parameter–environment relationships improves prediction of NBE

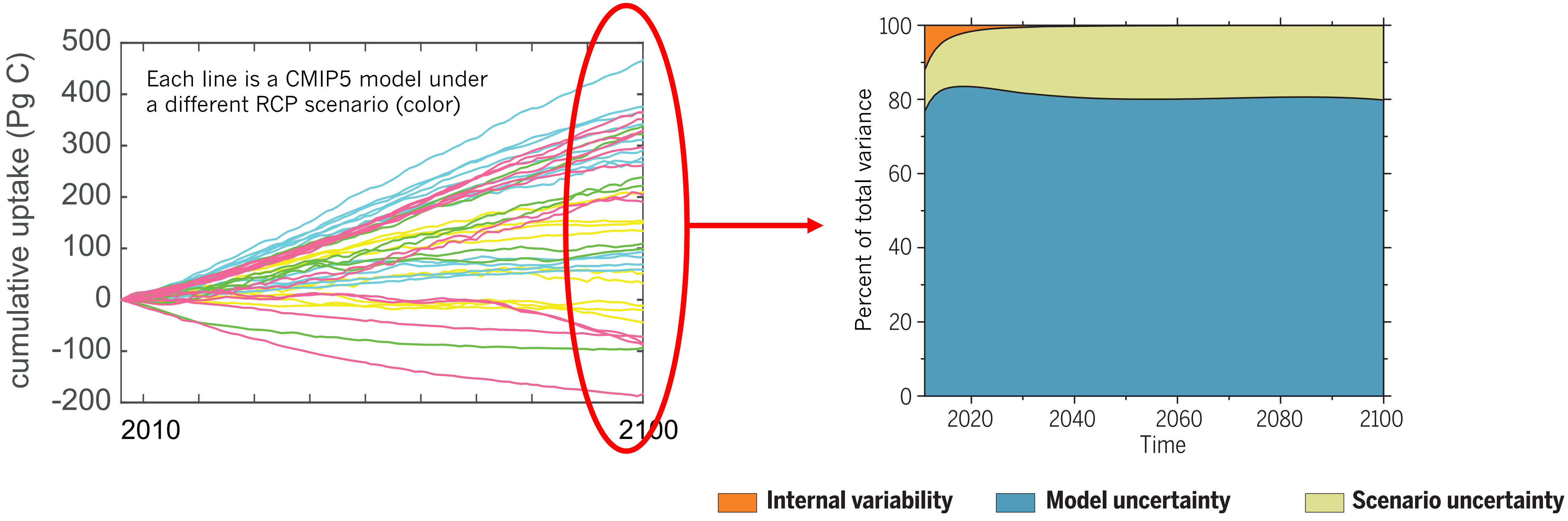
Caroline Famiglietti, Matthew Worden, Greg Quetin,
Luke Smallman, Uma Dayal, Anthony Bloom,
Mathew Williams, & Alexandra Konings

Despite its importance in the Earth system, NBE is challenging to predict



Bonan & Doney, 2018
Lovenduski & Bonan, 2017

Model uncertainty is a key determinant of the spread in future terrestrial carbon cycle predictions



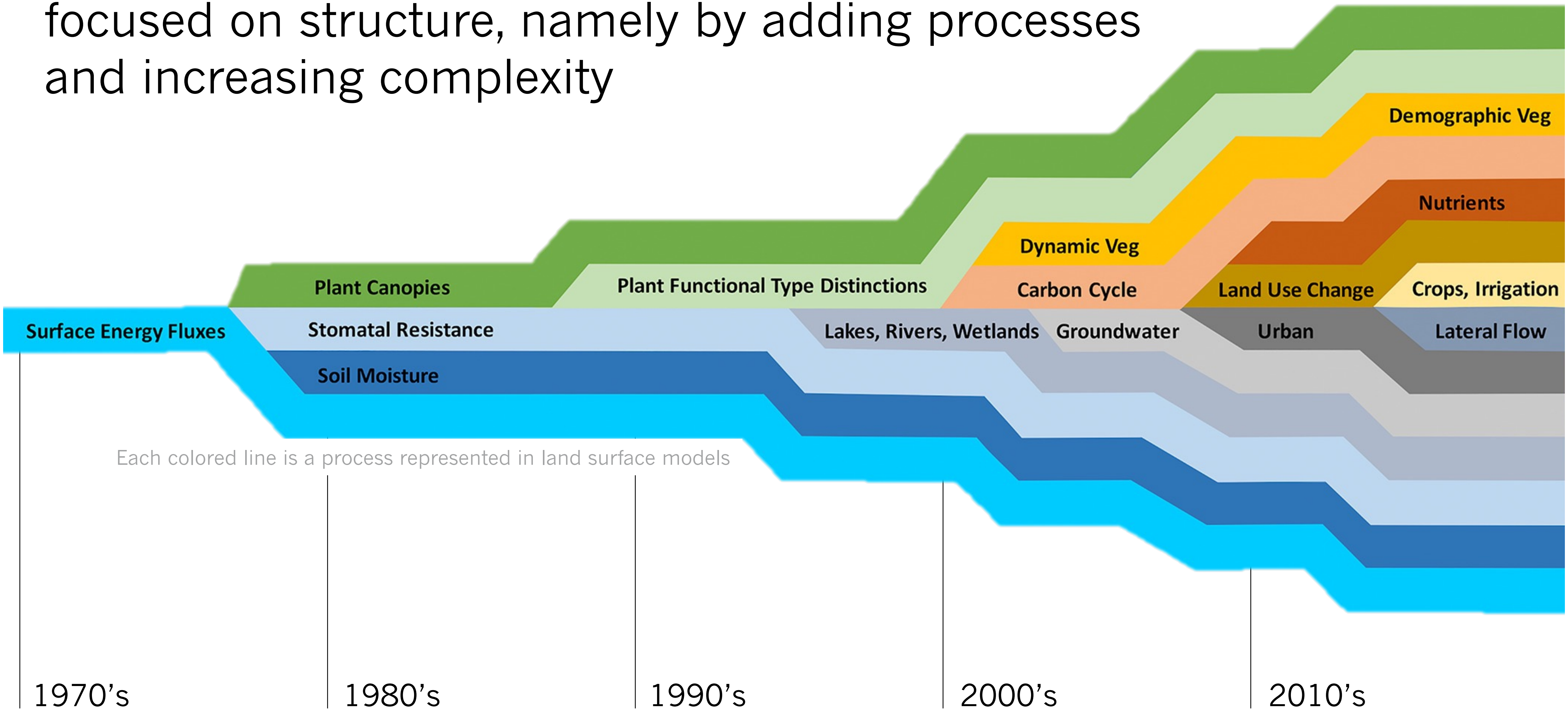
Bonan & Doney, 2018
Lovenduski & Bonan, 2017

Model uncertainty comprises:

Structural uncertainty *(i.e., how realistic is the model's representation of different processes?)*

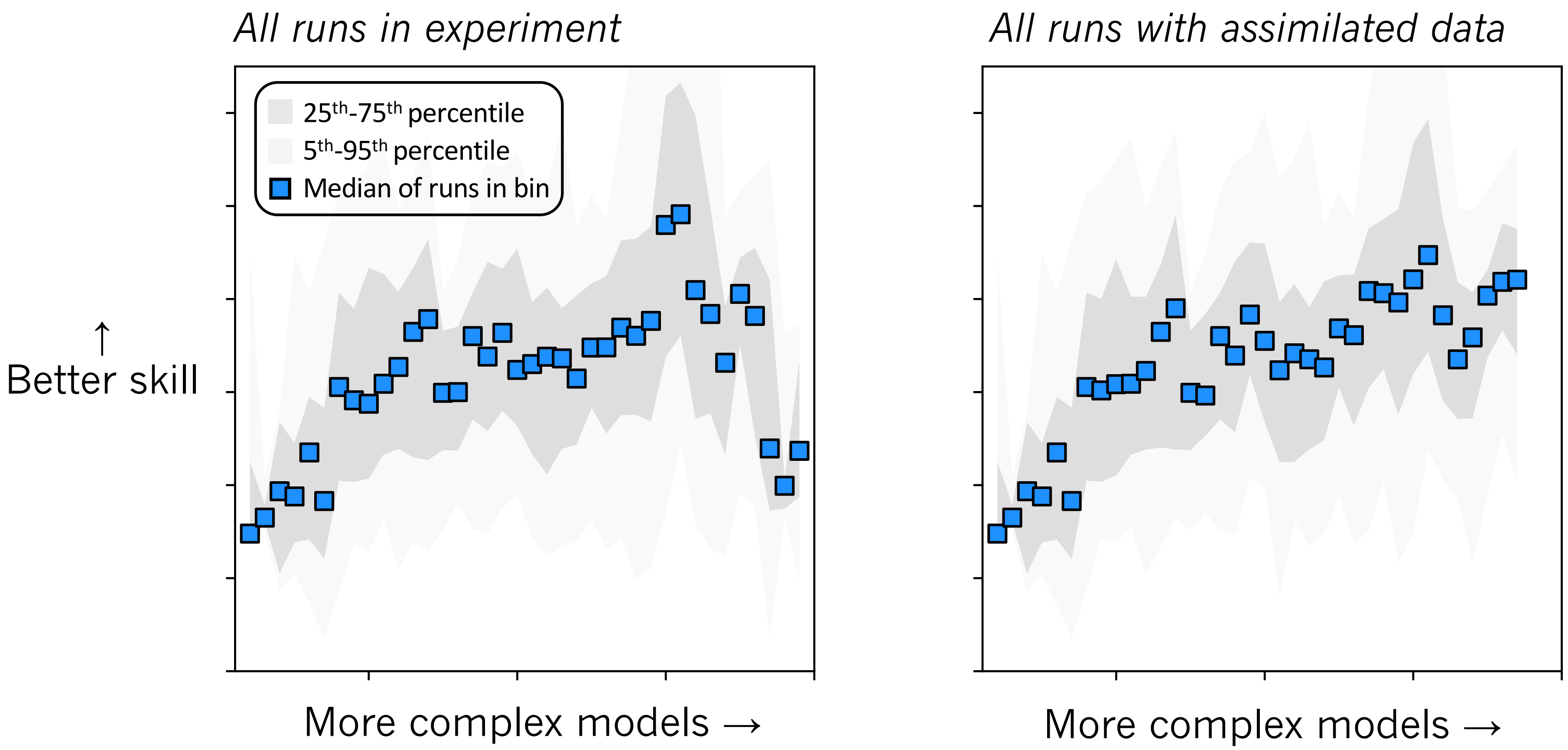
Parametric uncertainty *(i.e., how accurate are the model's parameter values?)*

Efforts to reduce model uncertainty have mainly focused on structure, namely by adding processes and increasing complexity



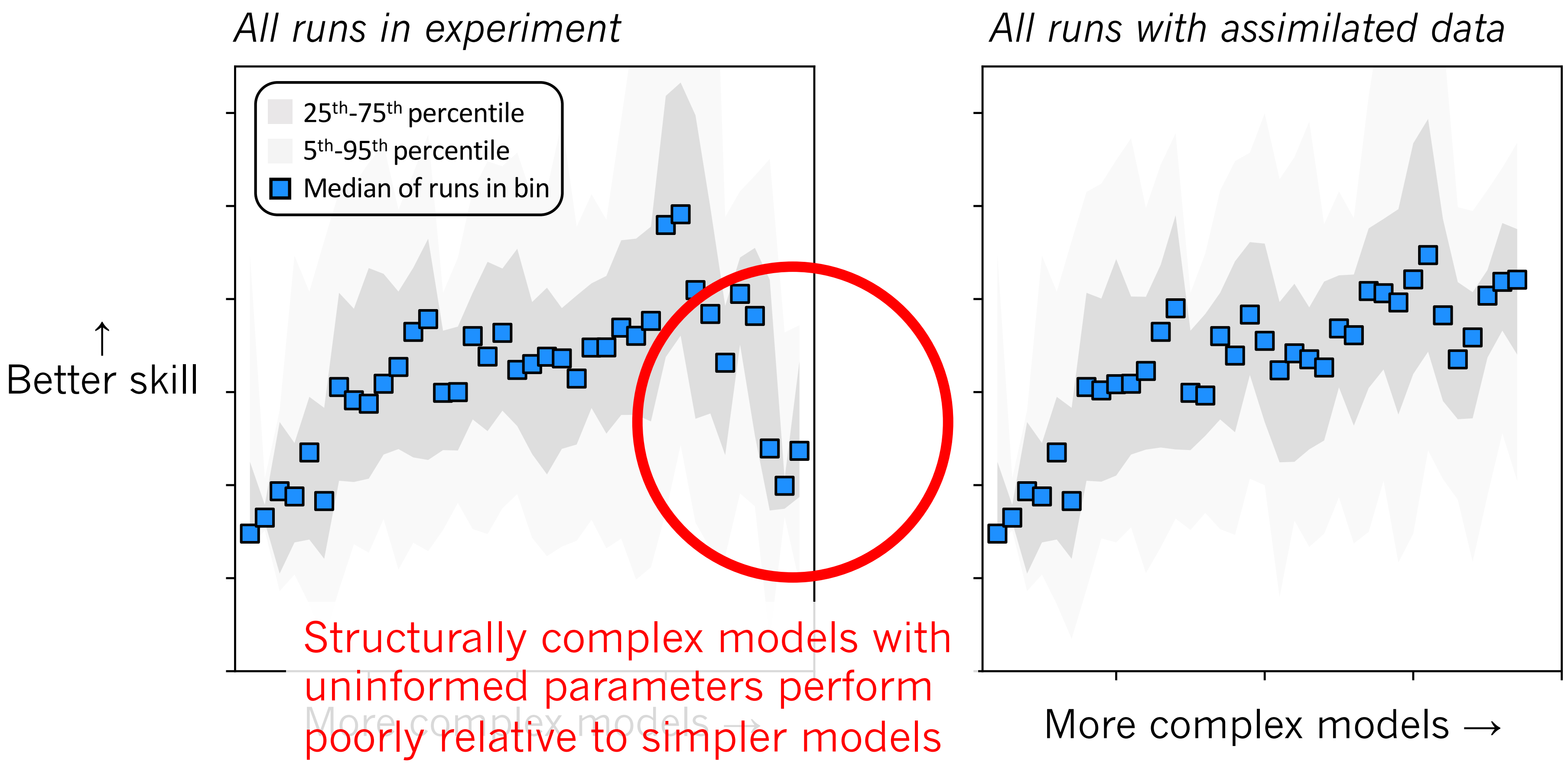
Fisher & Koven, 2020

Without accurate parameters, increasing complexity can degrade skill



Famiglietti et al., 2021

Without accurate parameters, increasing complexity can degrade skill



Famiglietti et al., 2021

Parameterizing a global model is challenging

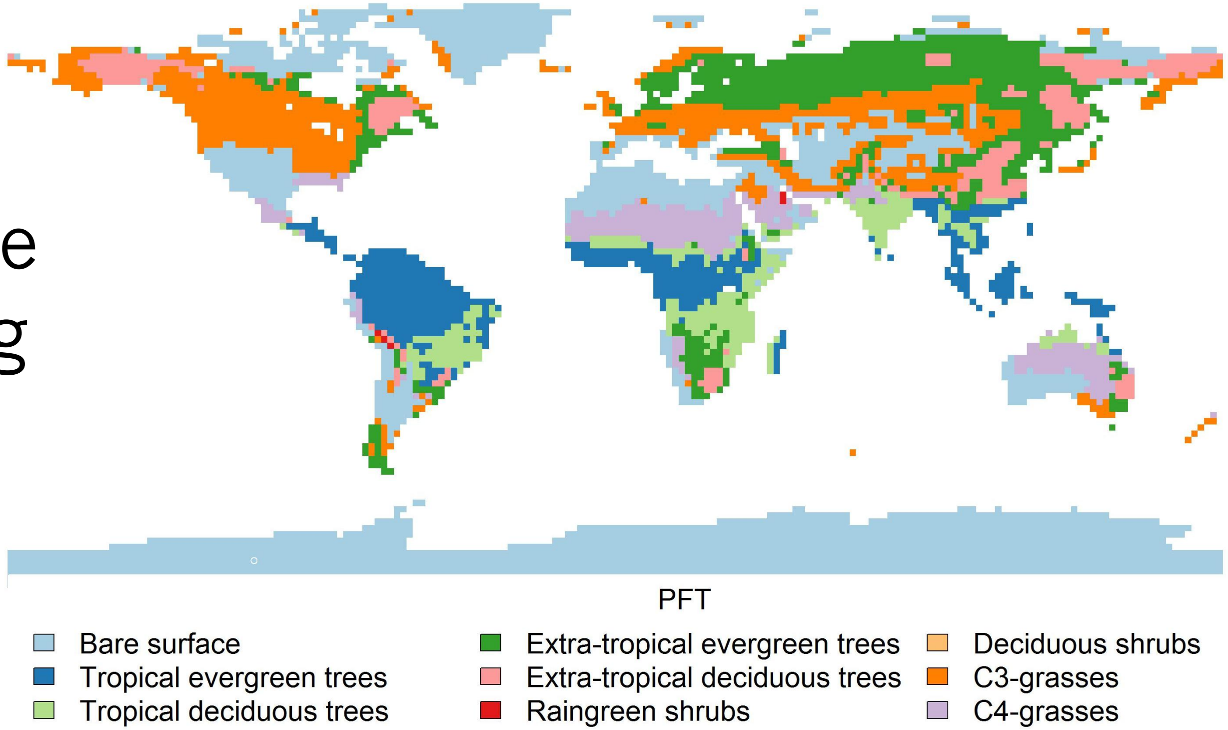
e.g., How to assign a leaf lifespan parameter globally?

→ Simplifying assumptions are necessary

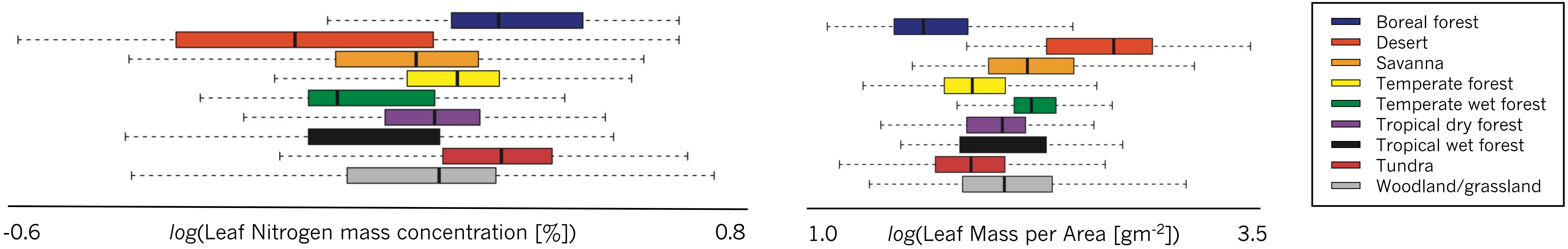
Here we ask:

How does the **choice of parameterization assumption** affect NBE prediction error?

Most commonly, parameters are assigned in global models using plant functional types (PFTs)



… but the true variability in parameters within a given PFT can often exceed that between them

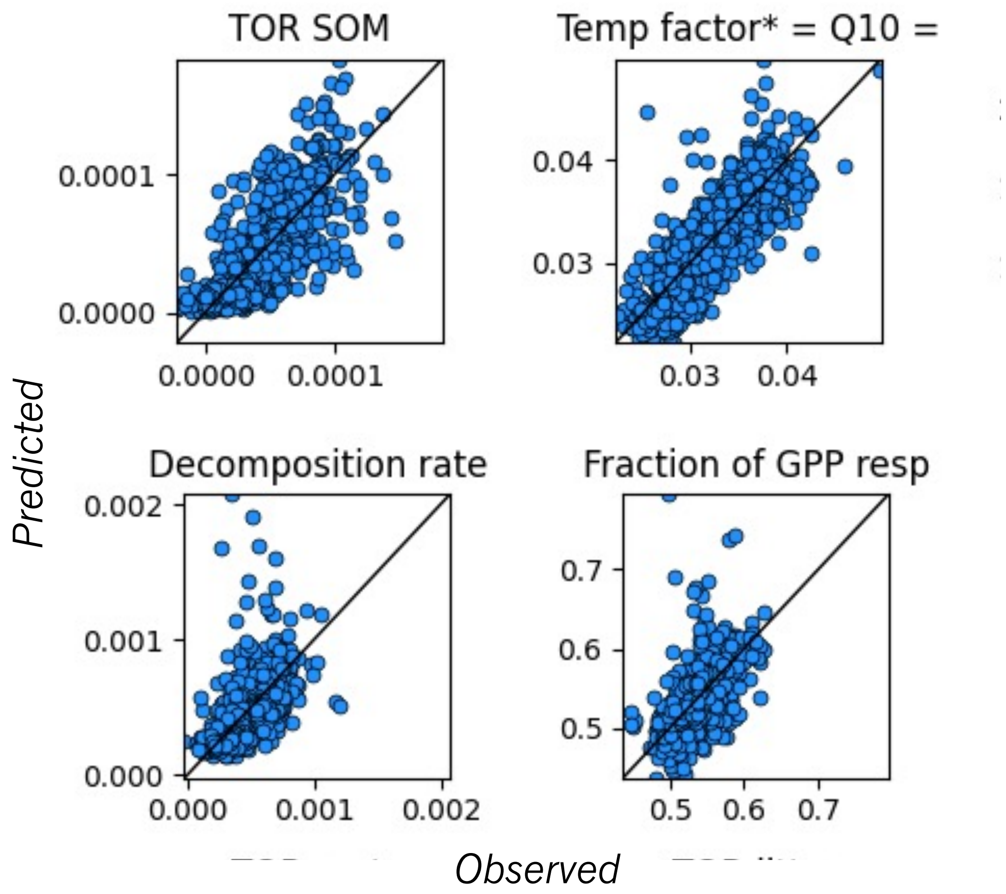


Adapted from *van Bodegom et al.*, 2012

Alternative approaches—like the environmental filtering hypothesis—may represent a way forward

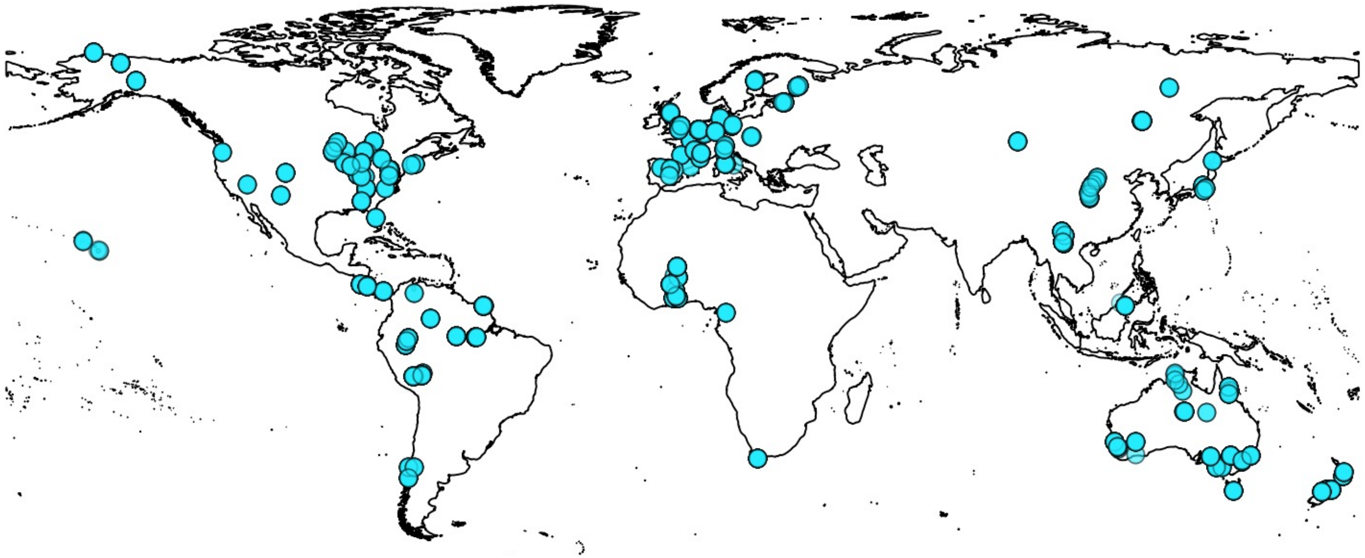
Alternative approaches—like the **environmental filtering hypothesis**—may represent a way forward

A pixel's parameters can be predicted as a function of local climate, soil, and canopy characteristics



... but the degree to which EF-based assumptions can improve C cycle predictions is unclear

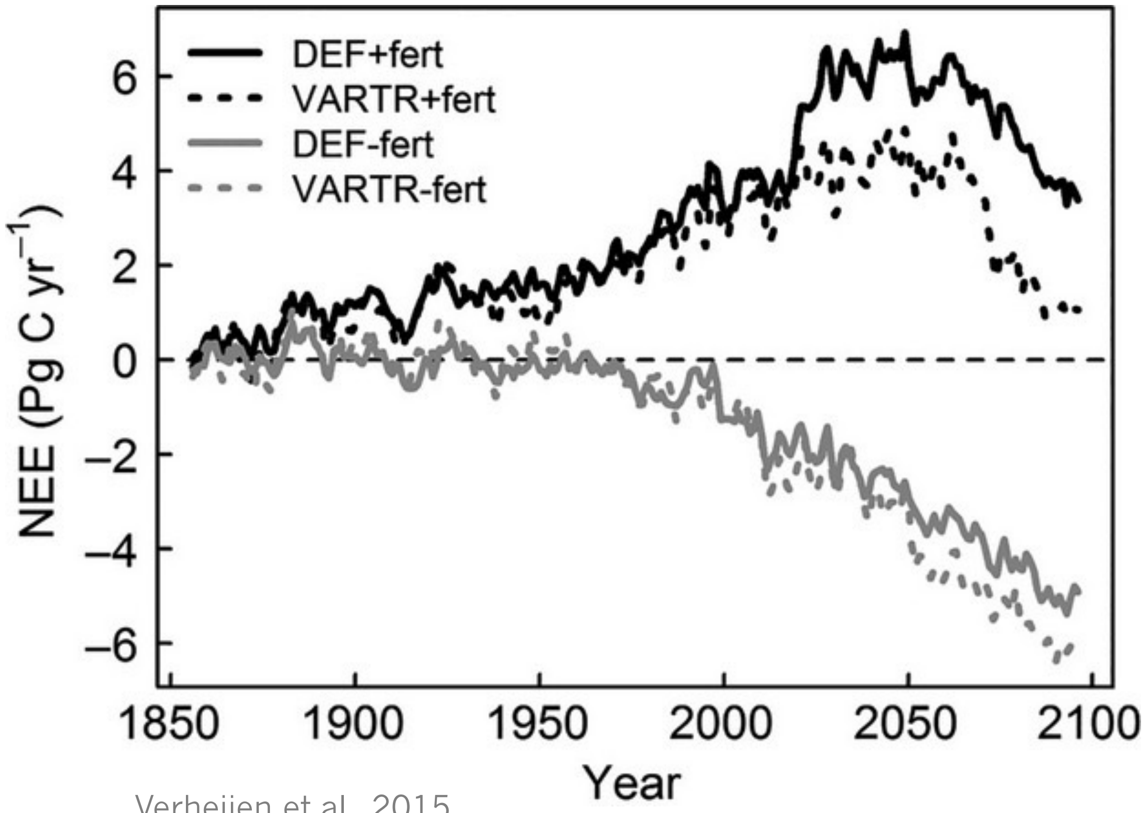
1. Previously EF relationships were developed using trait observations from the TRY database, which has significant **spatial- and species-related biases**



Example of locations of V_{cmax} measurements from the TRY database

... but the degree to which EF-based assumptions can improve C cycle predictions is unclear

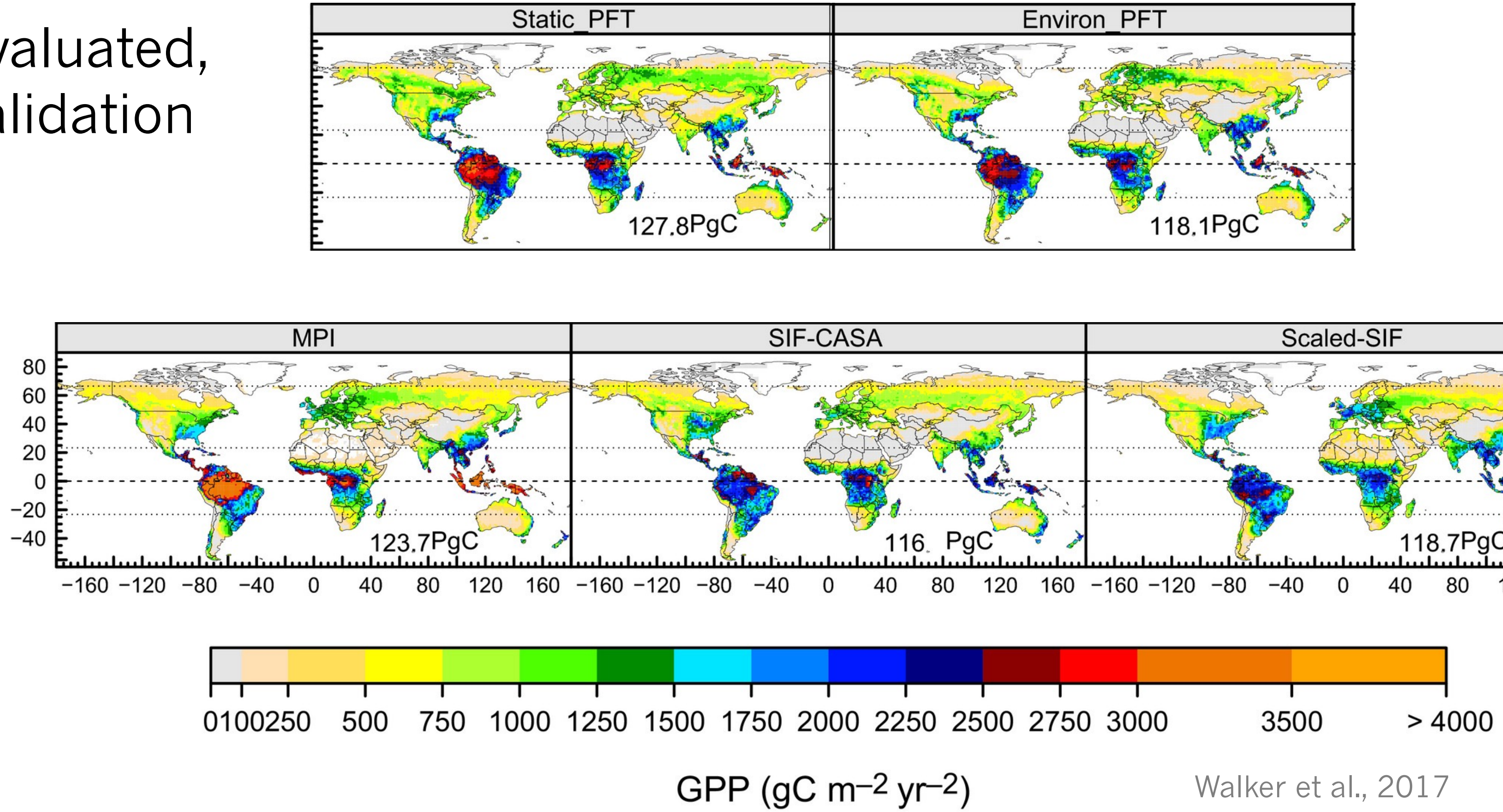
2. Many studies have only assessed differences in predictions, not errors



Verheijen et al., 2015

… but the degree to which EF-based assumptions can improve C cycle predictions is unclear

3. When errors are evaluated, proxies used for validation can themselves be uncertain



Walker et al., 2017

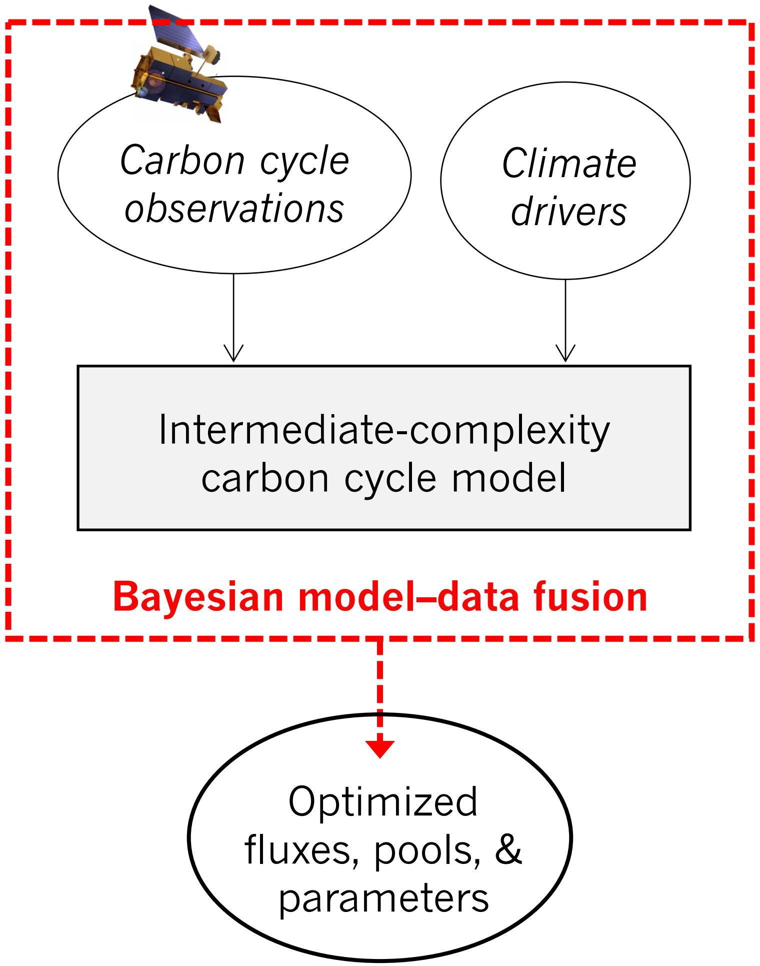
Can CARDAMOM help us avoid these issues?

We performed a simulation experiment to directly compare each parameterization assumption

1. CARDAMOM's flexible structure allows us to substitute either EF- or PFT-based assumptions into DALEC

2. CARDAMOM's optimal retrievals and resulting NBE predictions can be <u>used as a benchmark</u>

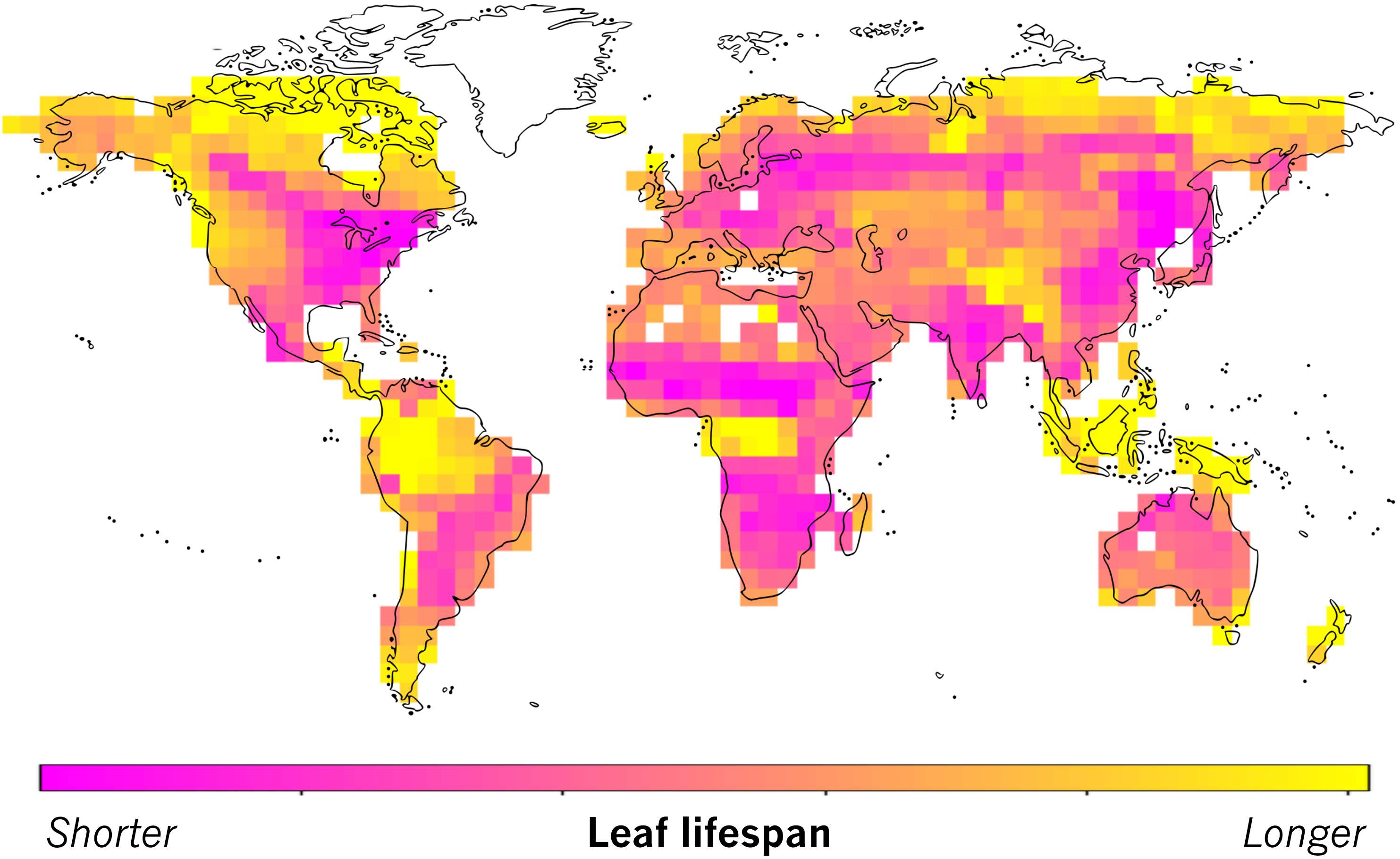
→ Because we compare to optimal predictions rather than to observations, mismatches across simulations are wholly attributable to parametric uncertainty



Methods summary

1. Retrieve optimal model parameters globally using CARDAMOM's standard inversion approach.

For example:



1. Retrieve optimal model parameters globally using CARDAMOM's standard inversion approach.

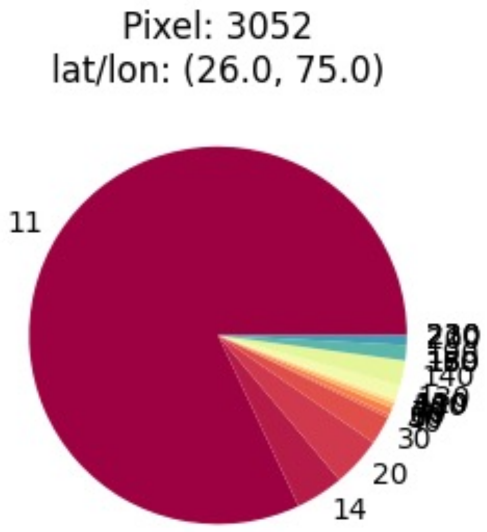
Adapted from Quetin et al., in revision

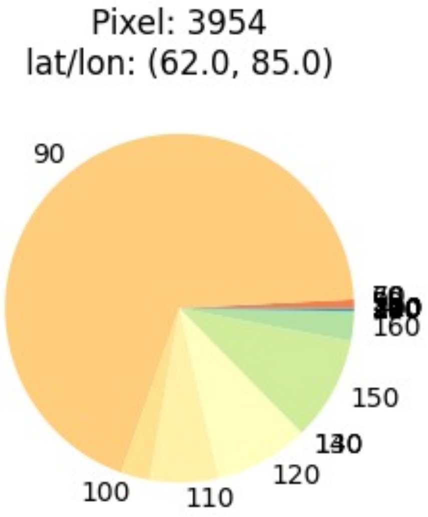
Observation	Source	Years	Uncertainty	Reference
Net biome exchange (NBE)	CMS-Flux	2010–2015	Optimized	Liu et al., 2017, 2021
Leaf area index (LAI)	MODIS	2010–2015	±log(1.2)	Myneni et al., 2002
Solar-induced fluorescence (SIF)	GOSAT	2010–2015	±log(2)	Frankenberg et al., 2011
Above- and below-ground biomass (ABGB)	Multiple	2000	≥±log(1.5)	Saatchi et al., 2011
Soil organic matter (SOM)	SoilGrids	2000	±log(1.5)	Poggio et al., 2021
Fire C emissions	MOPITT	2010–2015	±20%	Bowman et al., 2017; Worden et al., 2017

1. Retrieve optimal model parameters globally using CARDAMOM's standard MCMC approach.

2. Simulate PFTs by sampling optimal parameters from a few "sites" for each vegetation type

   a) For each coarse-scale CARDAMOM pixel, determine its PFT composition using an underlying land cover map (GlobCover).



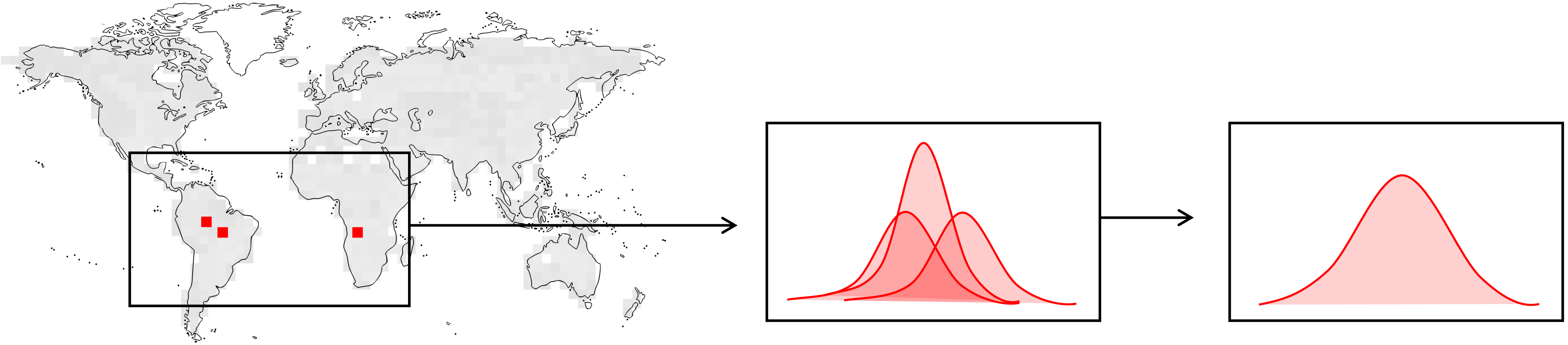


1. Retrieve optimal model parameters globally using CARDAMOM's standard MCMC approach.

2. Simulate PFTs by sampling optimal parameters from a few "sites" for each vegetation type

   b) Identify representative pixels for each PFT. Create aggregated parameter set.

   c) Assume the aggregated parameter sets are representative of all pixels sharing that vegetation type. **Exception for initial conditions*



Representative pixels for PFT X

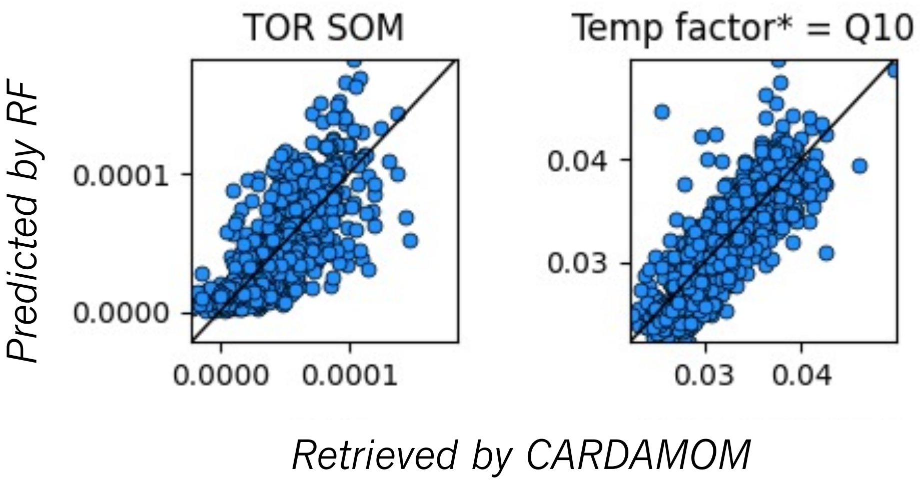
Representative pixels' optimal retrievals for parameter i

Aggregated distribution for parameter i (PFT X)

1. Retrieve optimal model parameters globally using CARDAMOM's standard MCMC approach.

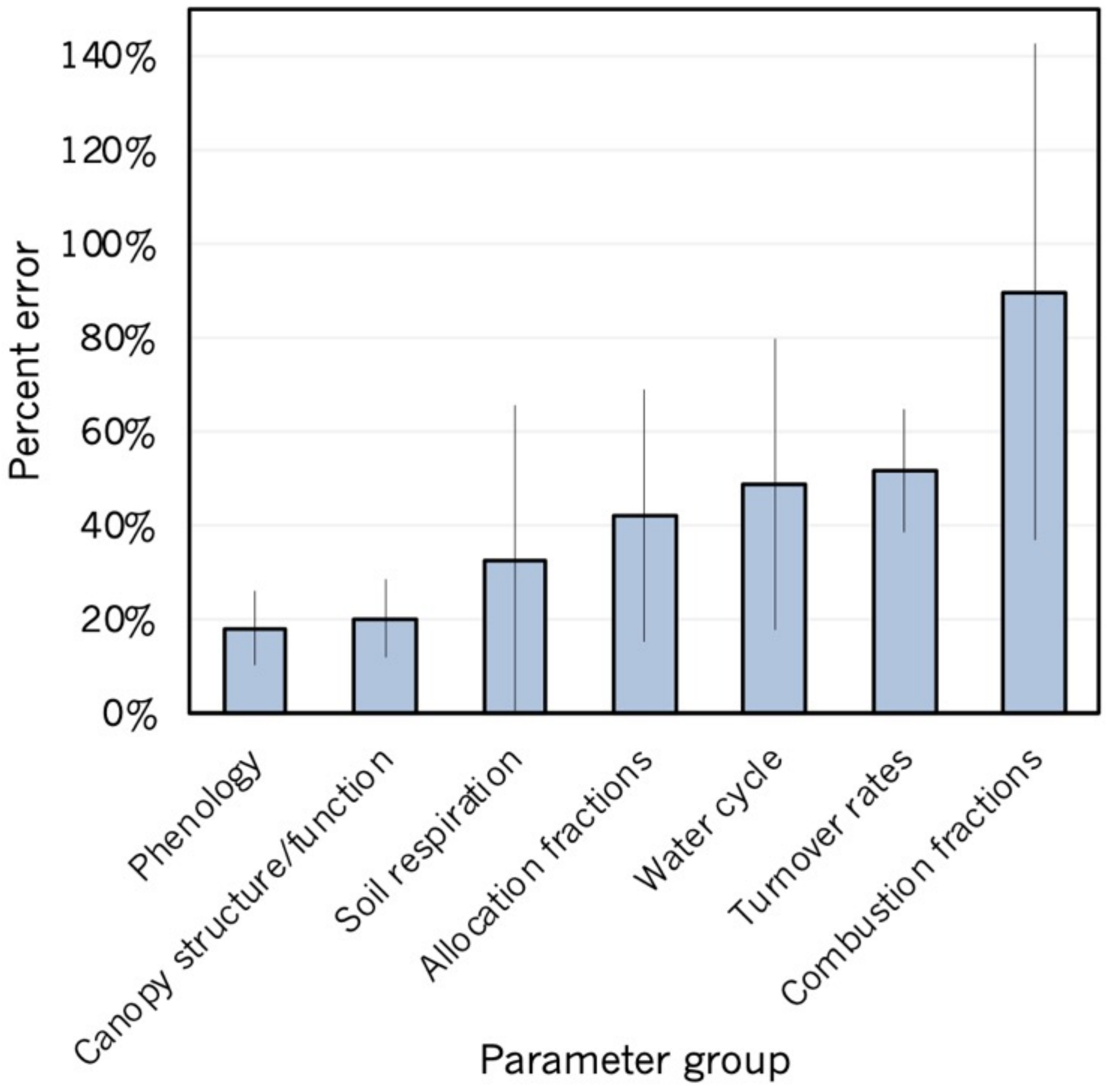
2. Simulate PFTs by sampling optimal parameters from a few "sites" for each vegetation type
   a) Assume the aggregated parameter sets are representative of all pixels sharing that vegetation type.

3. Develop **"top-down"** environmental filtering relationships to predict parameters using pixels' climate, soil, and canopy properties
   a) Train random forest models for each parameter using optimal retrievals. **Exception for initial conditions*
   b) Predictors include climate *(e.g.,* temperature, VPD*)*, canopy *(e.g.,* leaf area index), and soil *(e.g.,* soil pH, clay fraction) variables.

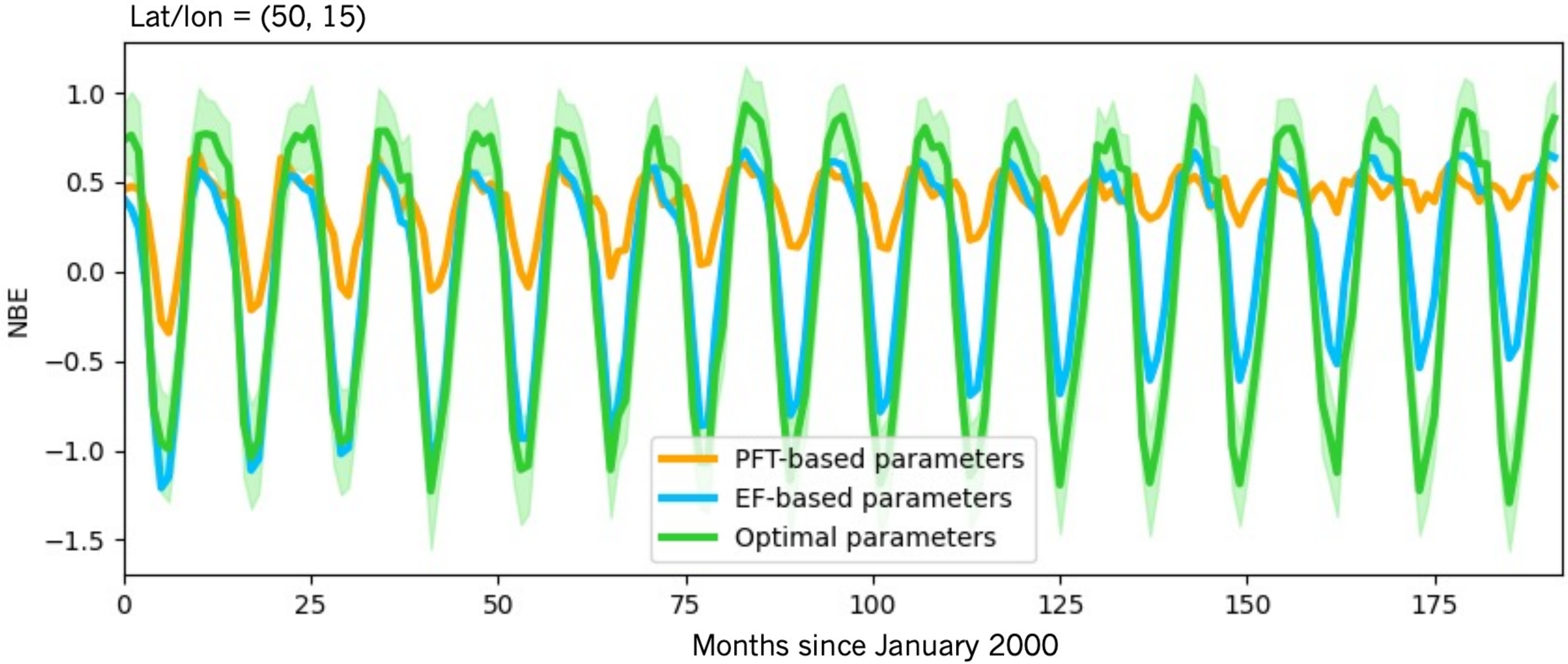


- Fire- and combustion-related parameters are poorly predicted by the EF approach

- Phenology & canopy structure/ function parameters are well predicted

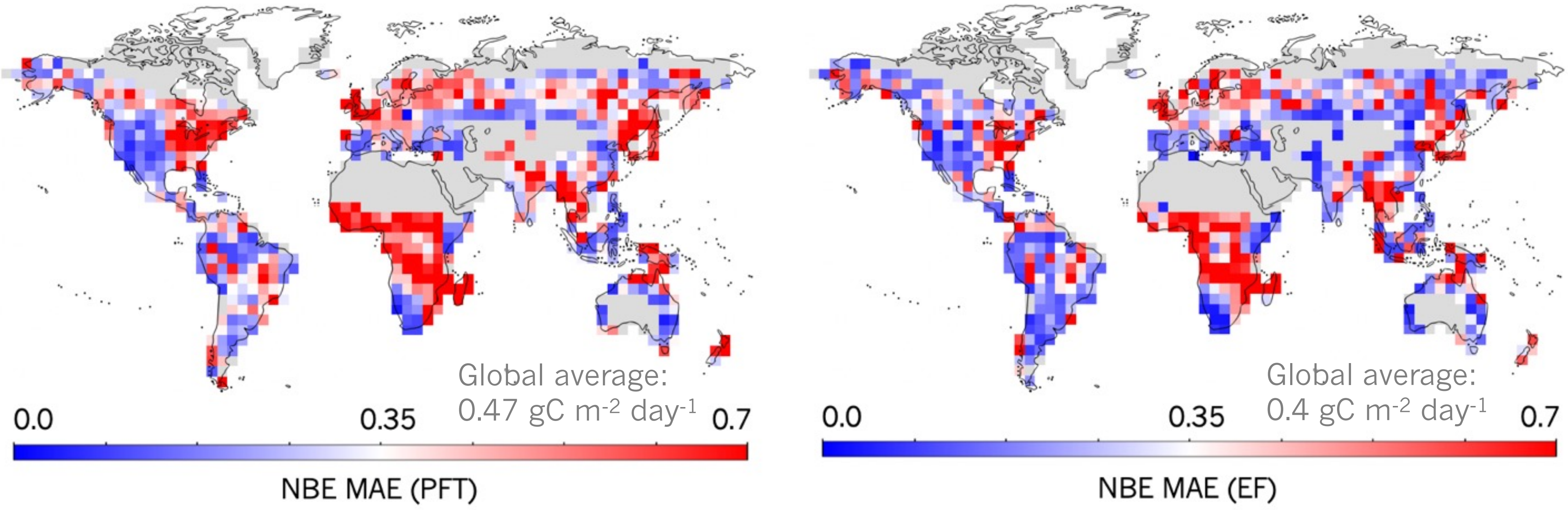


4. At each vegetated pixel globally, run CARDAMOM forward with each parameter set to produce NBE (as well as component flux) time series.



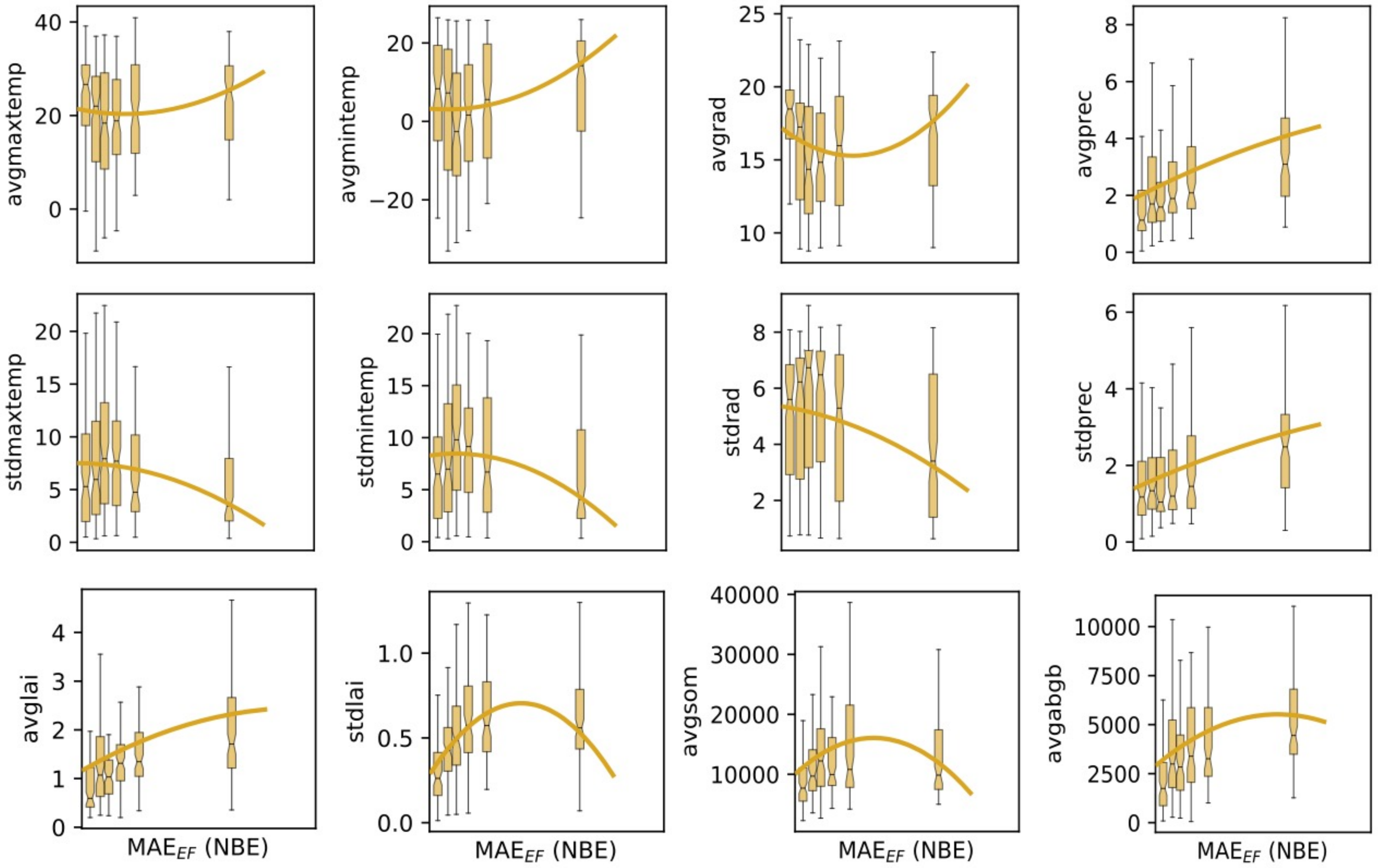
Results

NBE MAE hotspots largely overlap between PFT and EF-based approaches

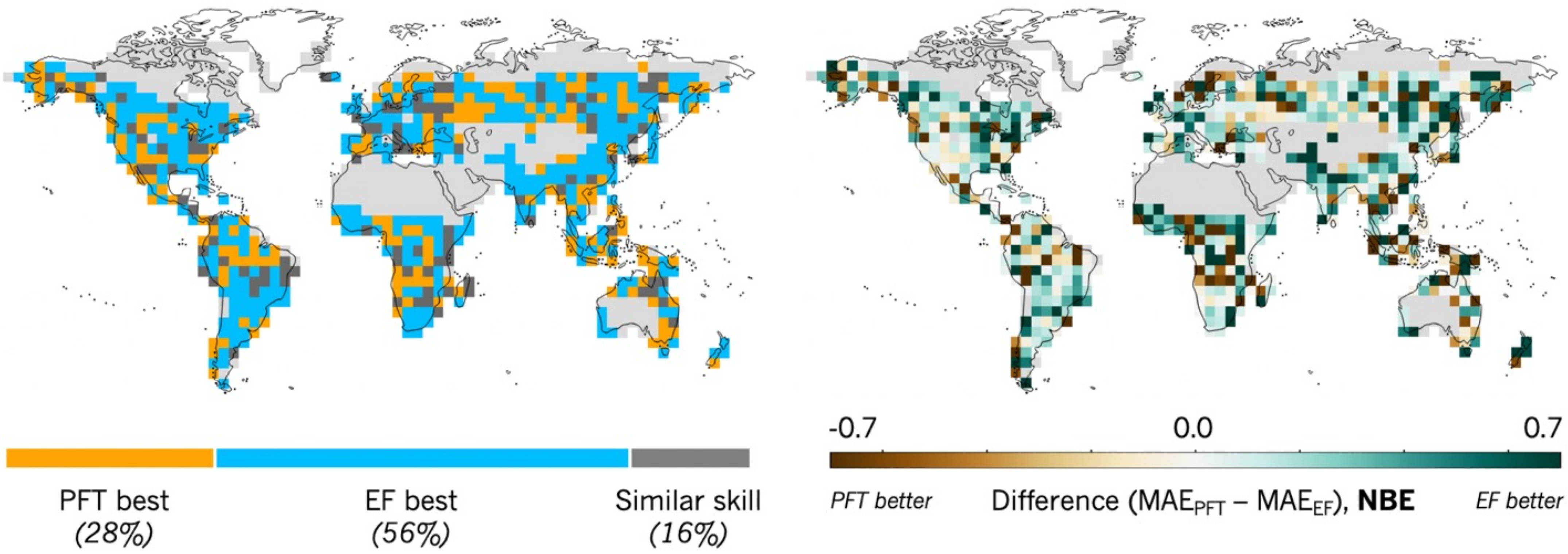


Skill determined using mean absolute error in NBE over the period 2000-2015, relative to the optimally parameterized NBE estimates

Error patterns broadly follow gradients of climate and vegetation

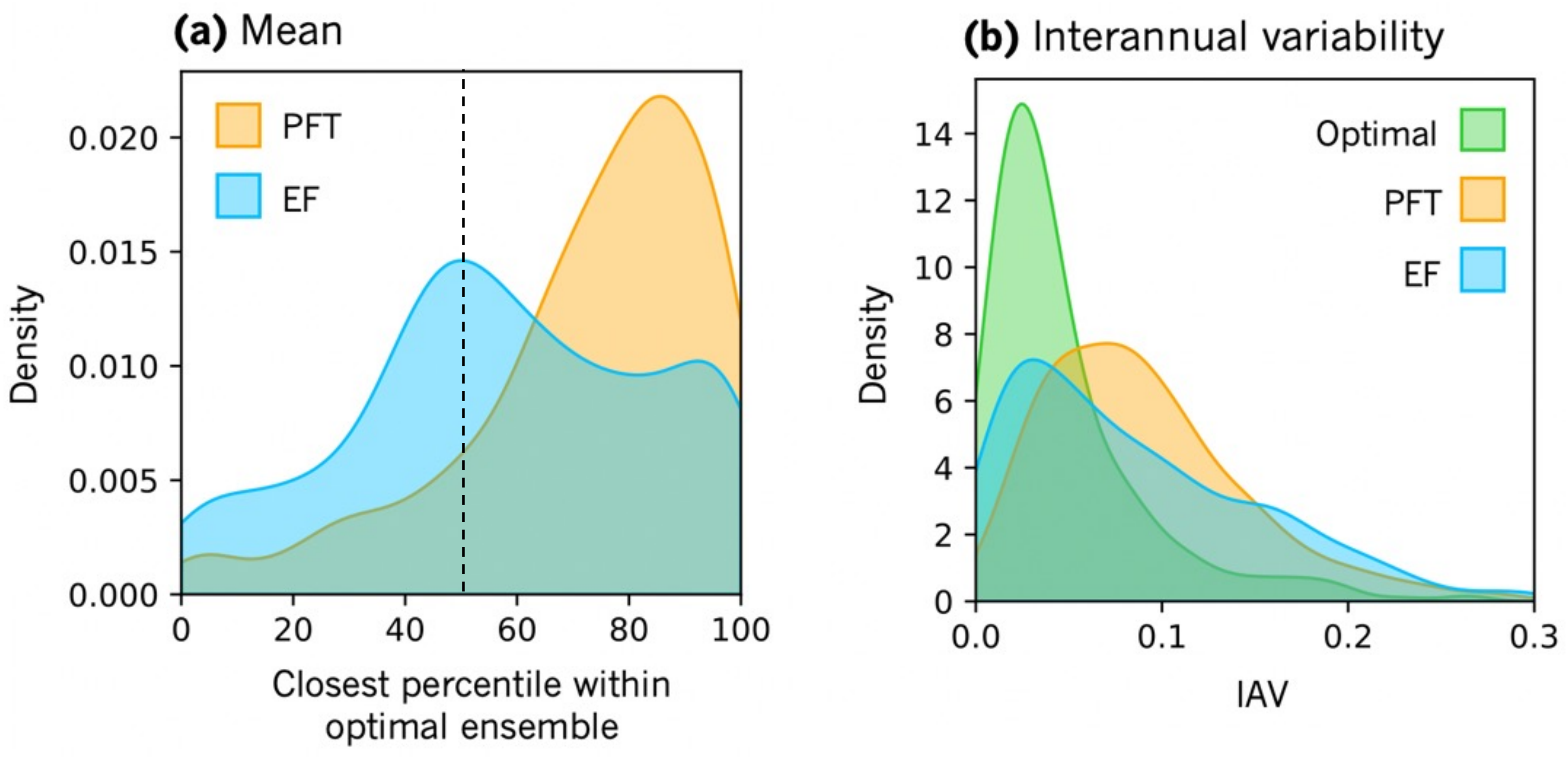


EF outperforms PFT-based approach at 2x as many pixels as the converse



Skill determined using mean absolute error in NBE over the period 2000-2015, relative to the optimally parameterized NBE estimates

EF captures NBE mean and IAV more accurately than PFT-based approach



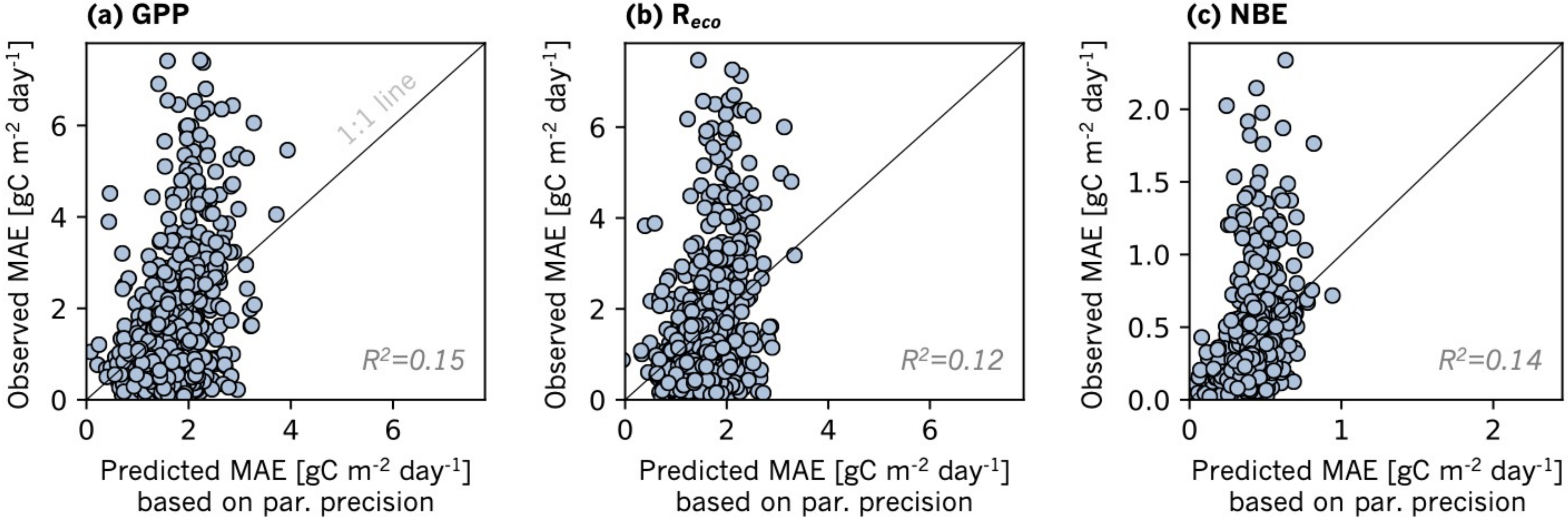
What controls variations in the EF-based model's relative NBE performance across space?

H1. Accuracy of EF-based parameter predictions

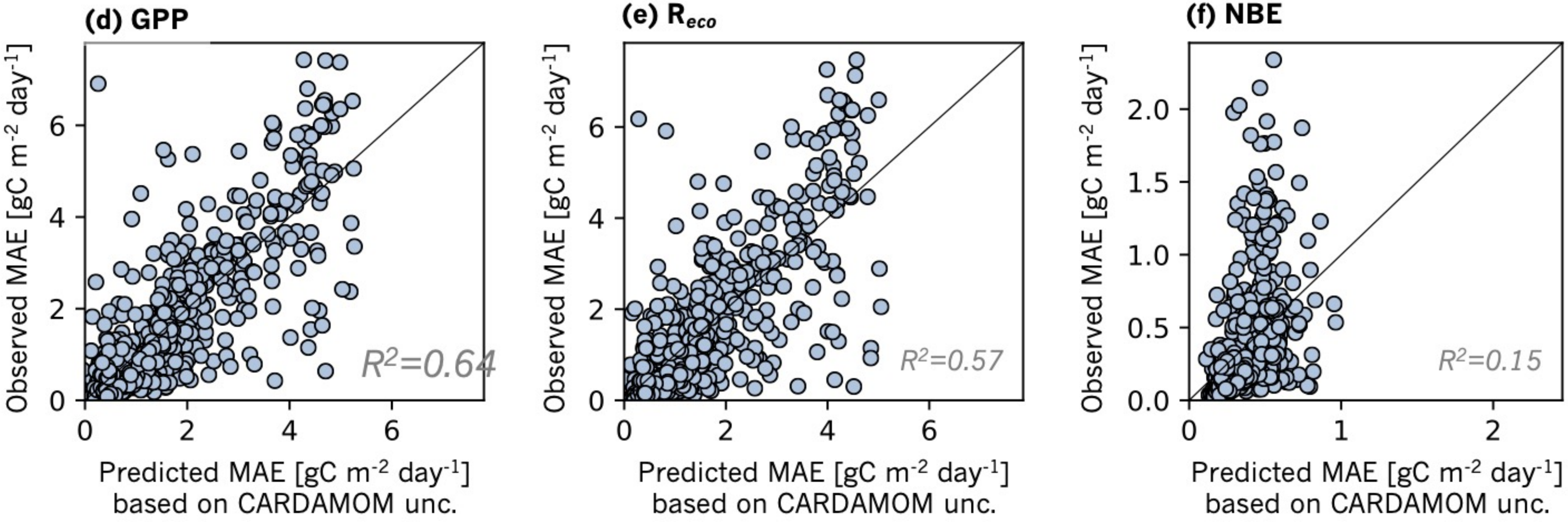
H2. Uncertainty of CARDAMOM's retrievals

Here we expand our lens to also consider NBE's component fluxes.

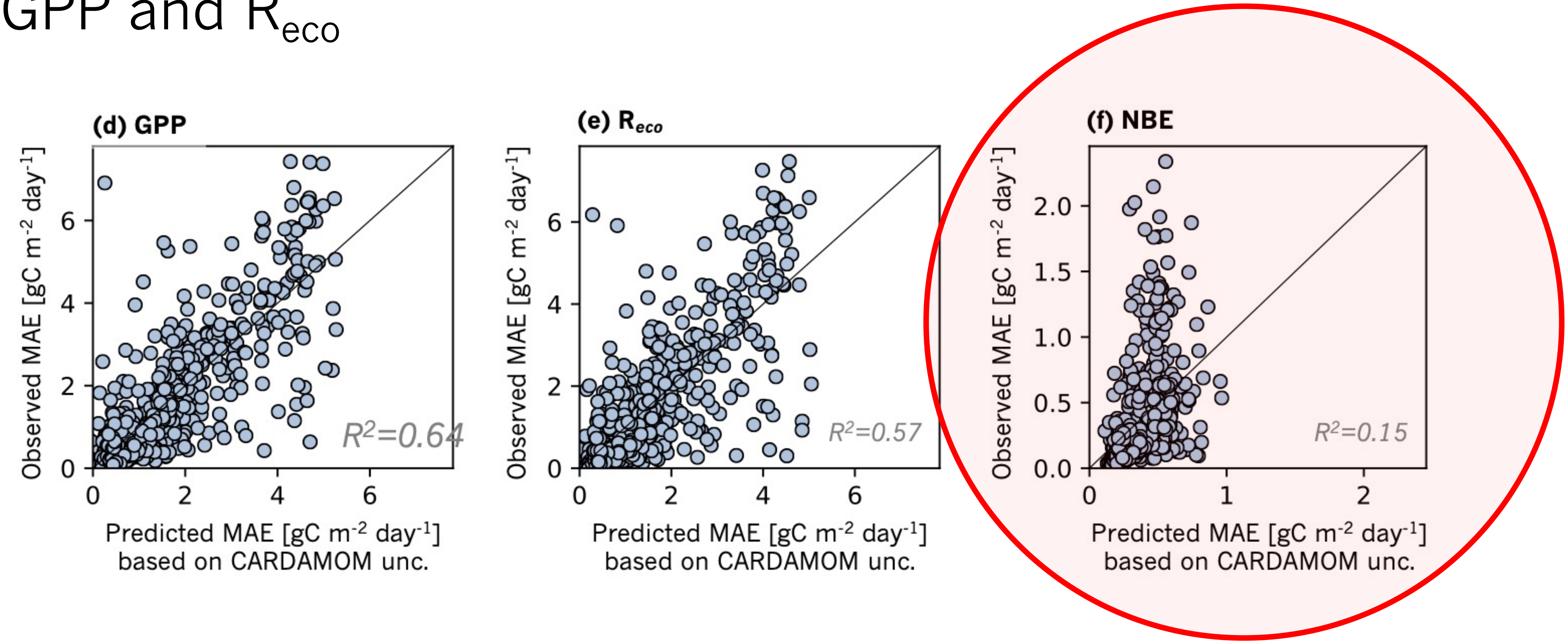
H1. Parameter accuracy is a poor predictor of flux MAE



H2. But CARDAMOM's uncertainty is a strong predictor, at least for GPP and R_{eco}

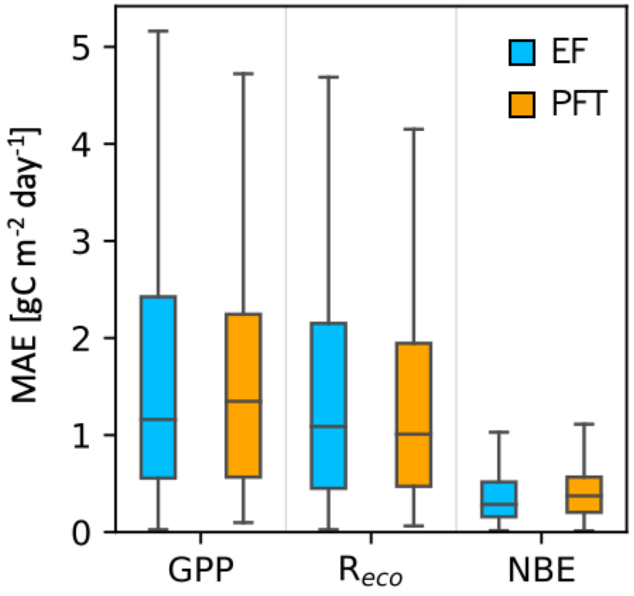


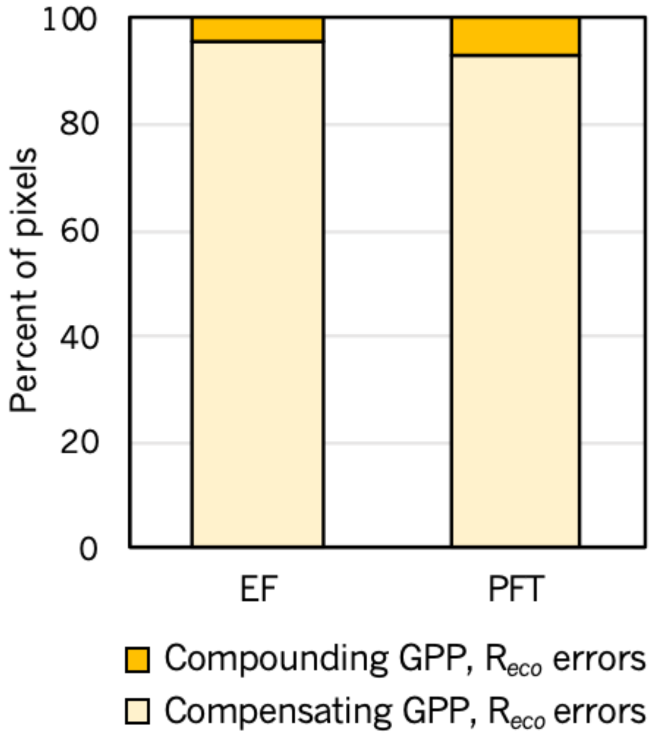
H2. But CARDAMOM's uncertainty is a strong predictor, at least for GPP and R_{eco}



Why not NBE?

NBE errors are the result of strong compensation between component flux predictions

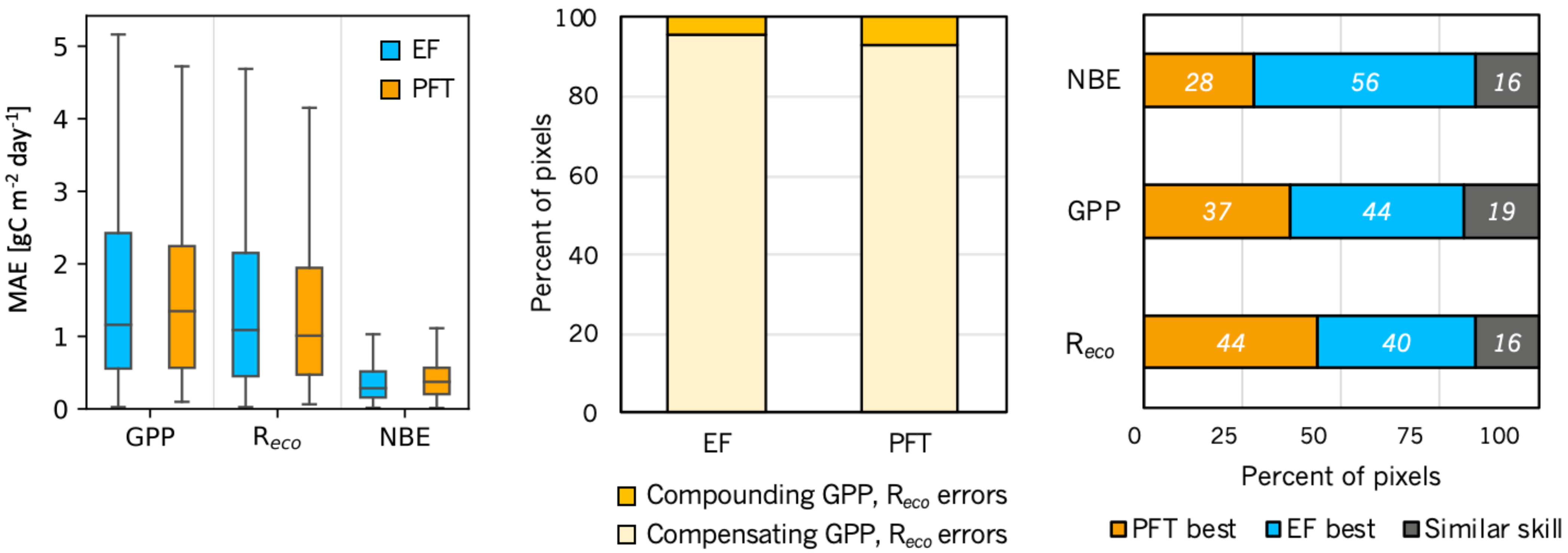




Compounding errors = **NBE error larger** than both GPP and Reco errors

Compensating errors = **NBE error smaller** than either GPP or Reco error, or both

NBE errors are the result of strong compensation between component flux predictions



Compensation occurs in both models, but impacts differ

Key takeaways:

- NBE predictability **strongly controlled** by choice of parameterization assumption
- EF-based model shows lower NBE MAE than PFT-based model at **2x as many pixels** as the converse
- EF-based model matches mean and IAV of NBE more closely than PFT-based model across pixels
- Still, both models show significant **compensation** between component flux (GPP and Reco) errors

Thank you!