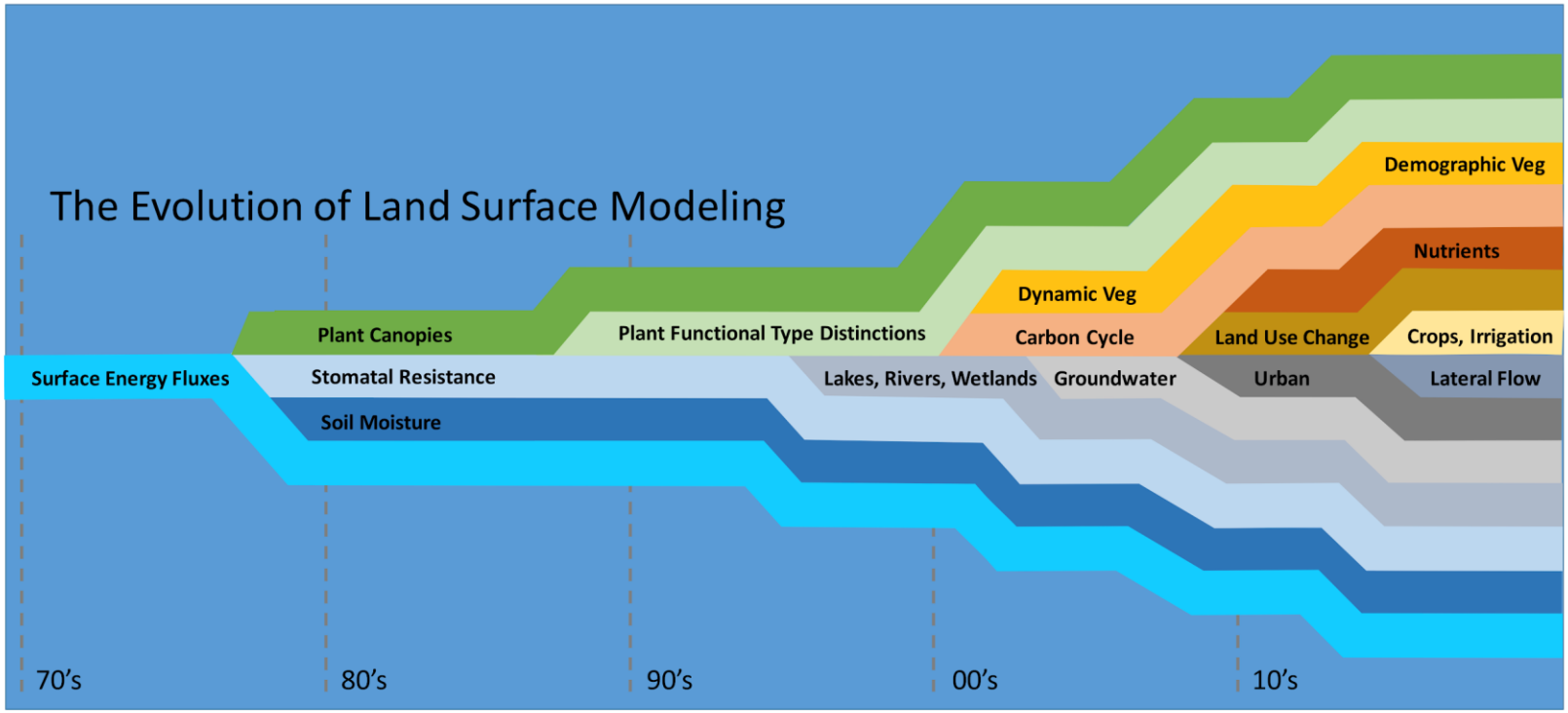
Managing model complexity

Charlie Koven, Lawrence Berkeley National Lab

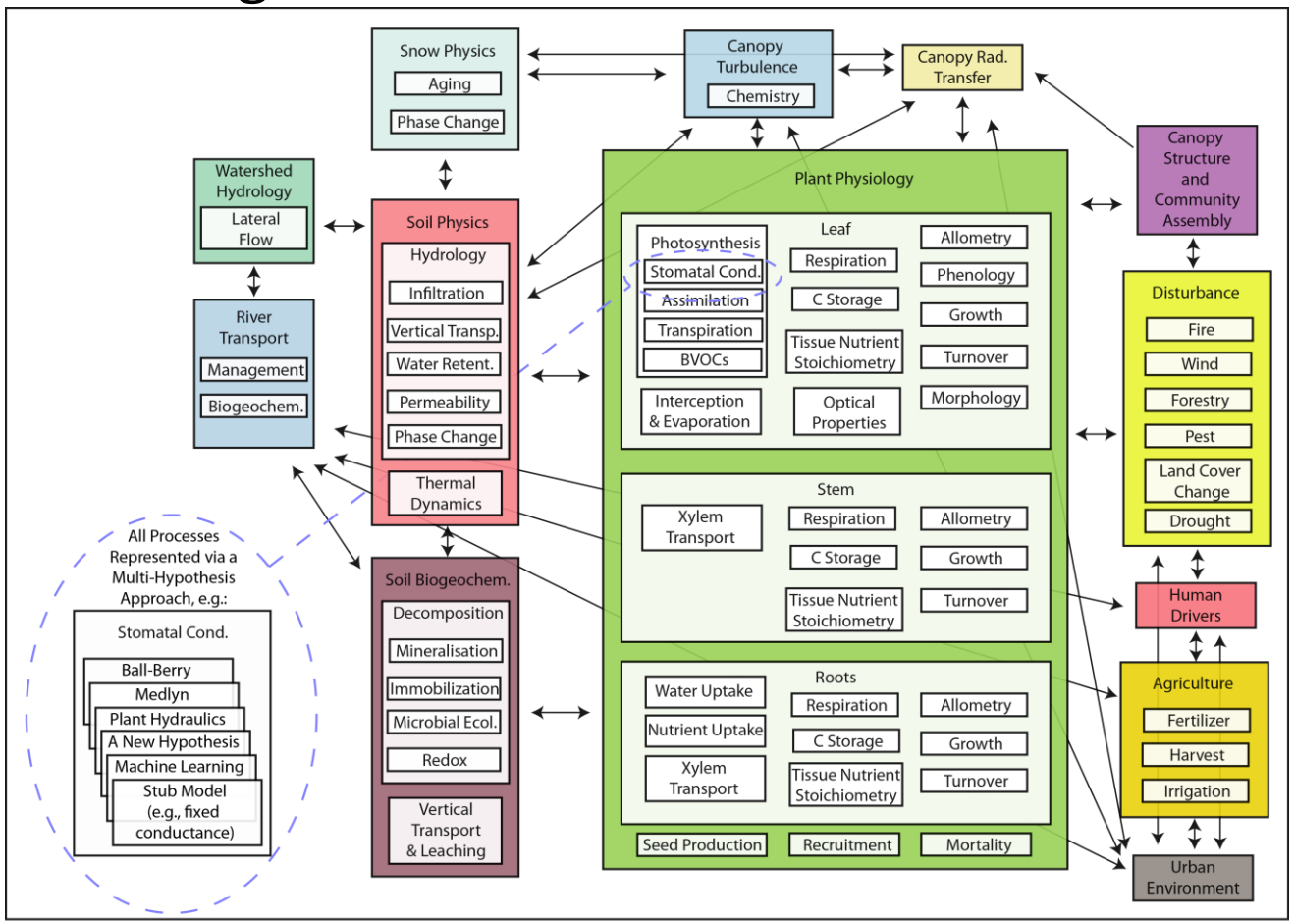
With input from: Rosie Fisher, Jessica Needham, Jennifer Holm, Forrest Hoffman, David Lawrence, Ben Smith, Victor Brovkin, Kei Yoshimura, Andy Pitman, Andy Wiltshire, Vivek Arora, Ben Bond-Lamberty

The thing about models is that they only get more complex over time.



Fisher and Koven, 2020

How do we build models that both:
(a) allow comprehensive assessment of the myriad processes and feedbacks in the land system, and
(b) allow controlled experiments, calibration, and understanding?



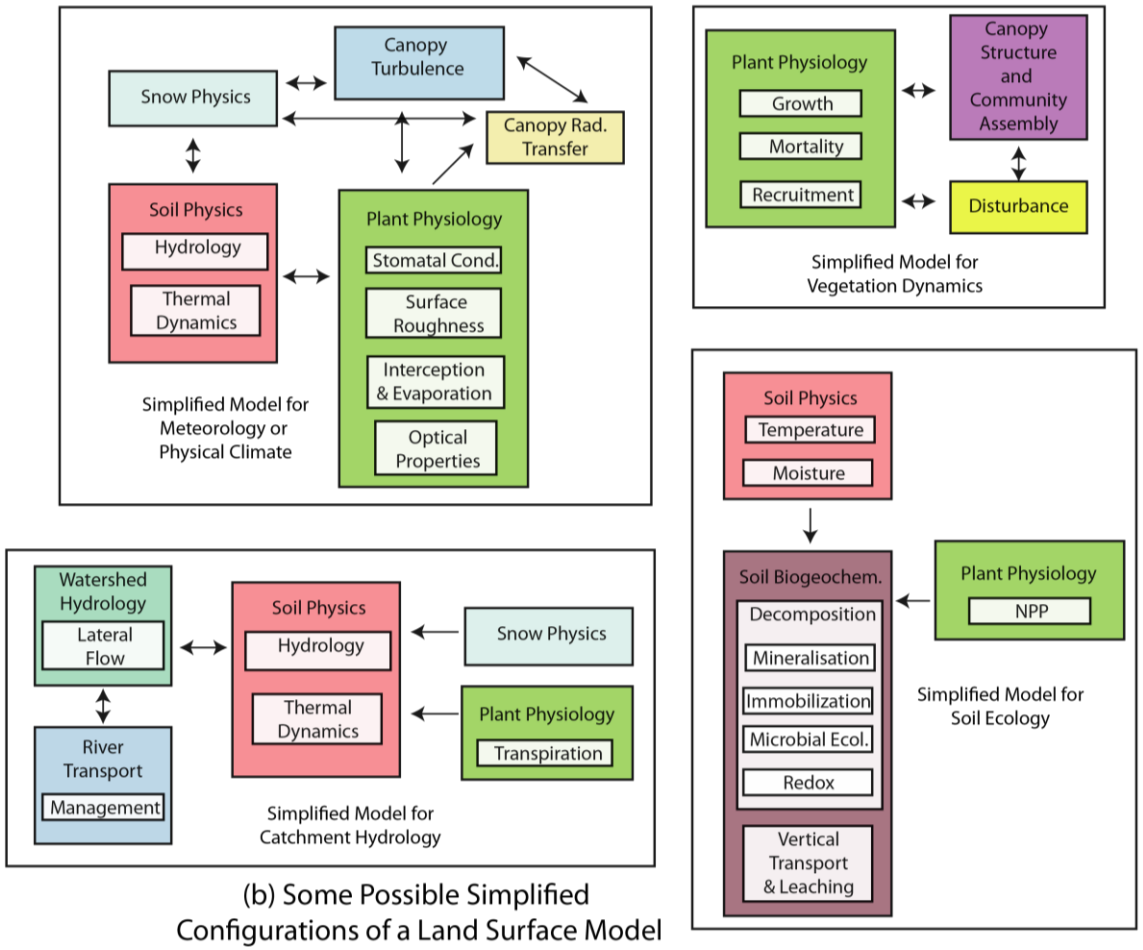
(a) Process Schematic of a Possible Full-Complexity Configuration of a Land Surface Model

Fisher and Koven, 2020

Several distinct problems introduced by complexity

- Barriers to entry: how can one use a model if one doesn't fully understand all of the processes in it?
- Calibration: how to calibrate models with so many internal feedbacks from many different processes?
- Instabilities: how to prevent one bad prediction in a model from taking down the whole thing?
- Experimental design: how to design simulations to allow one to focus a model on only the desired processes?

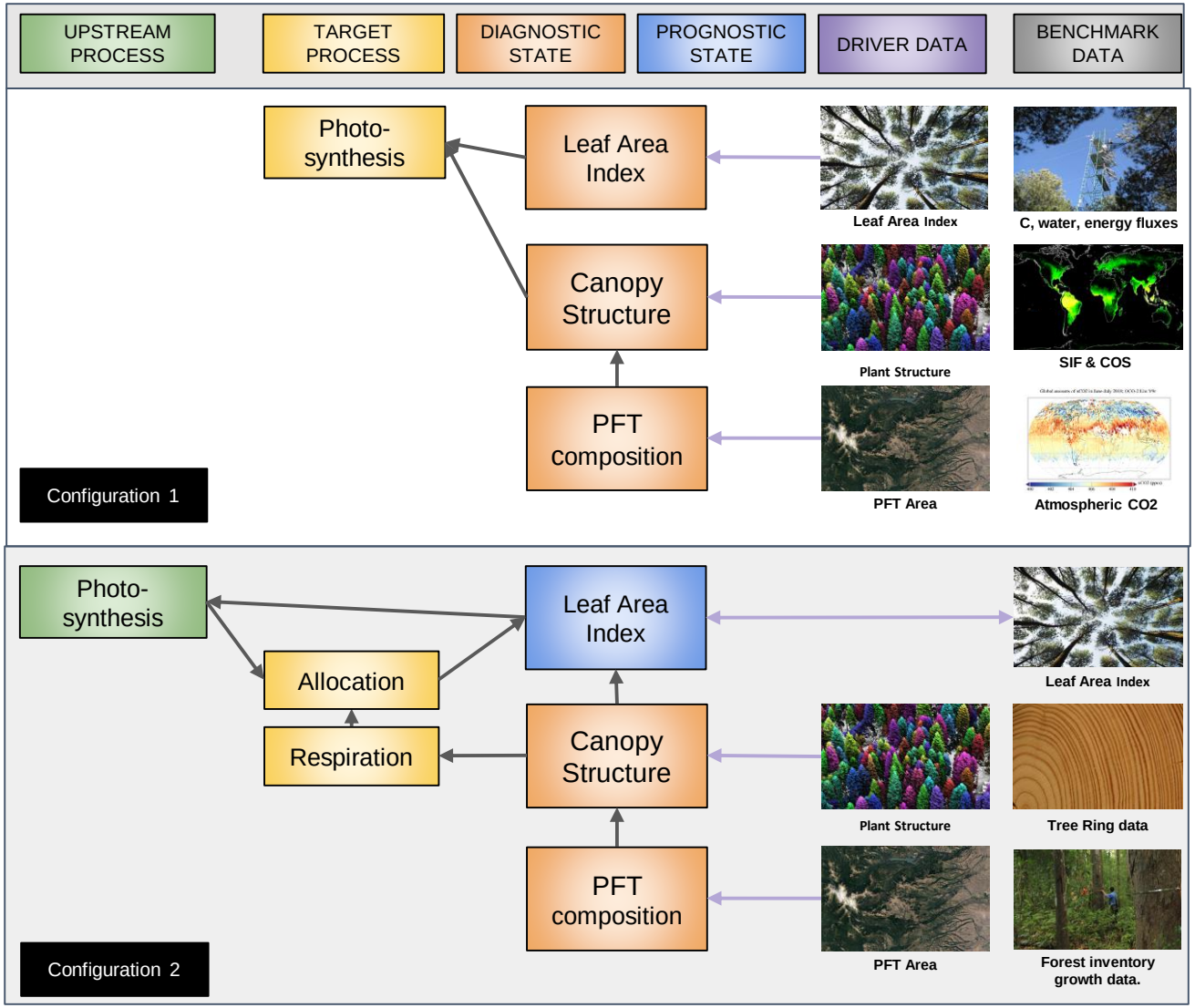
One strategy, “modular complexity”, might be to build models that can be configurable to either complex or (multiple) simple representations.



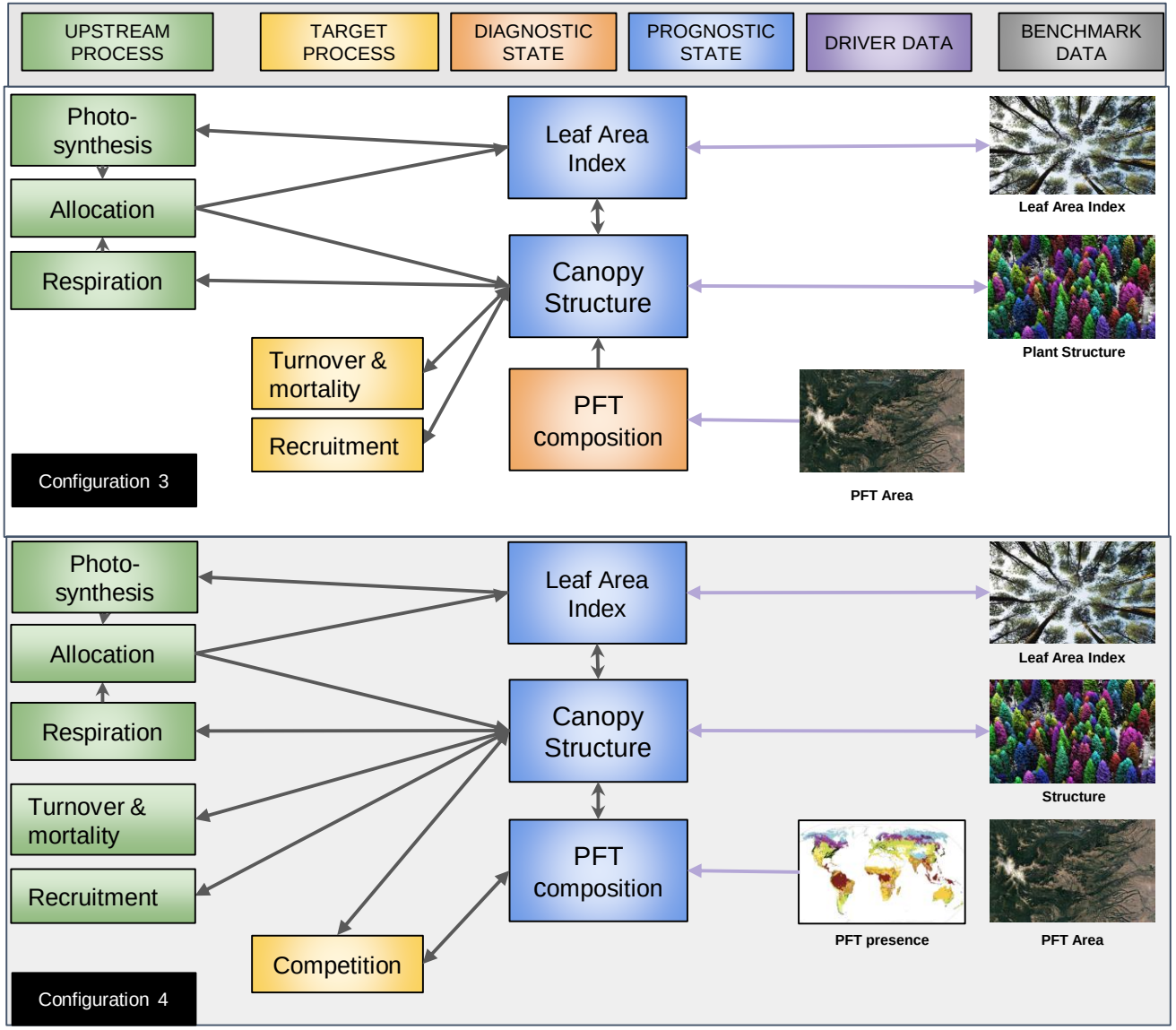
(b) Some Possible Simplified Configurations of a Land Surface Model

Fisher and Koven, 2020

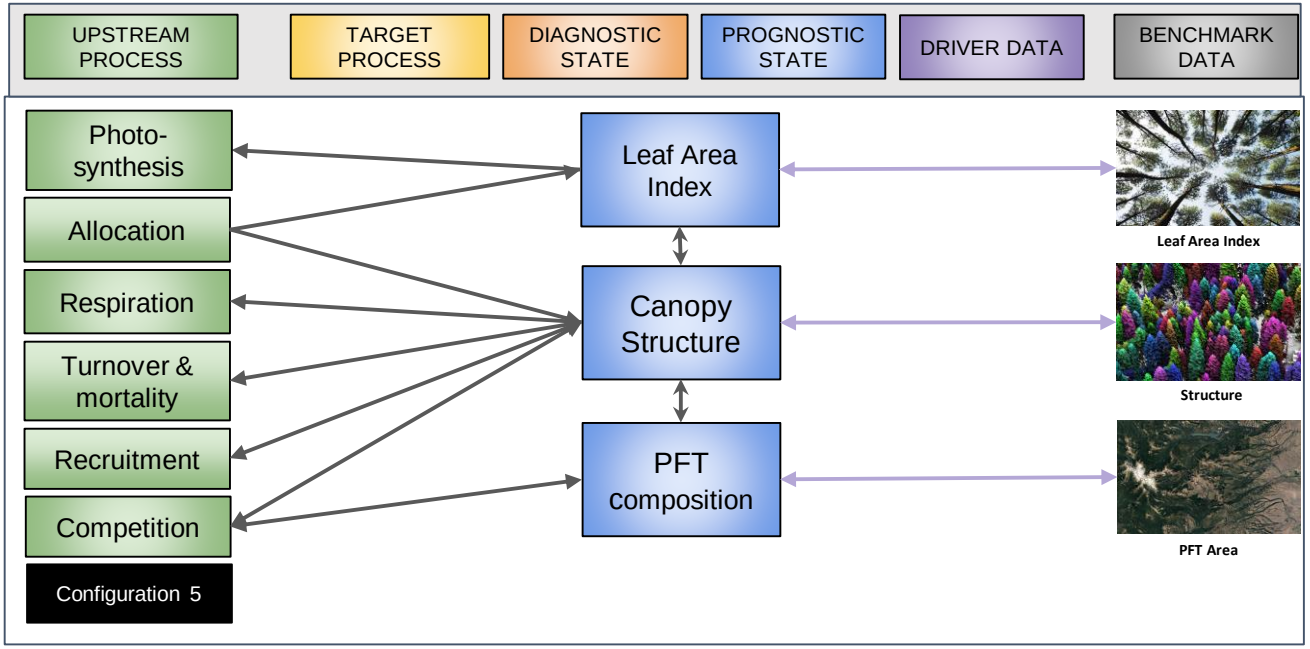
An example of the modular complexity approach: FATES “calibration cascade”



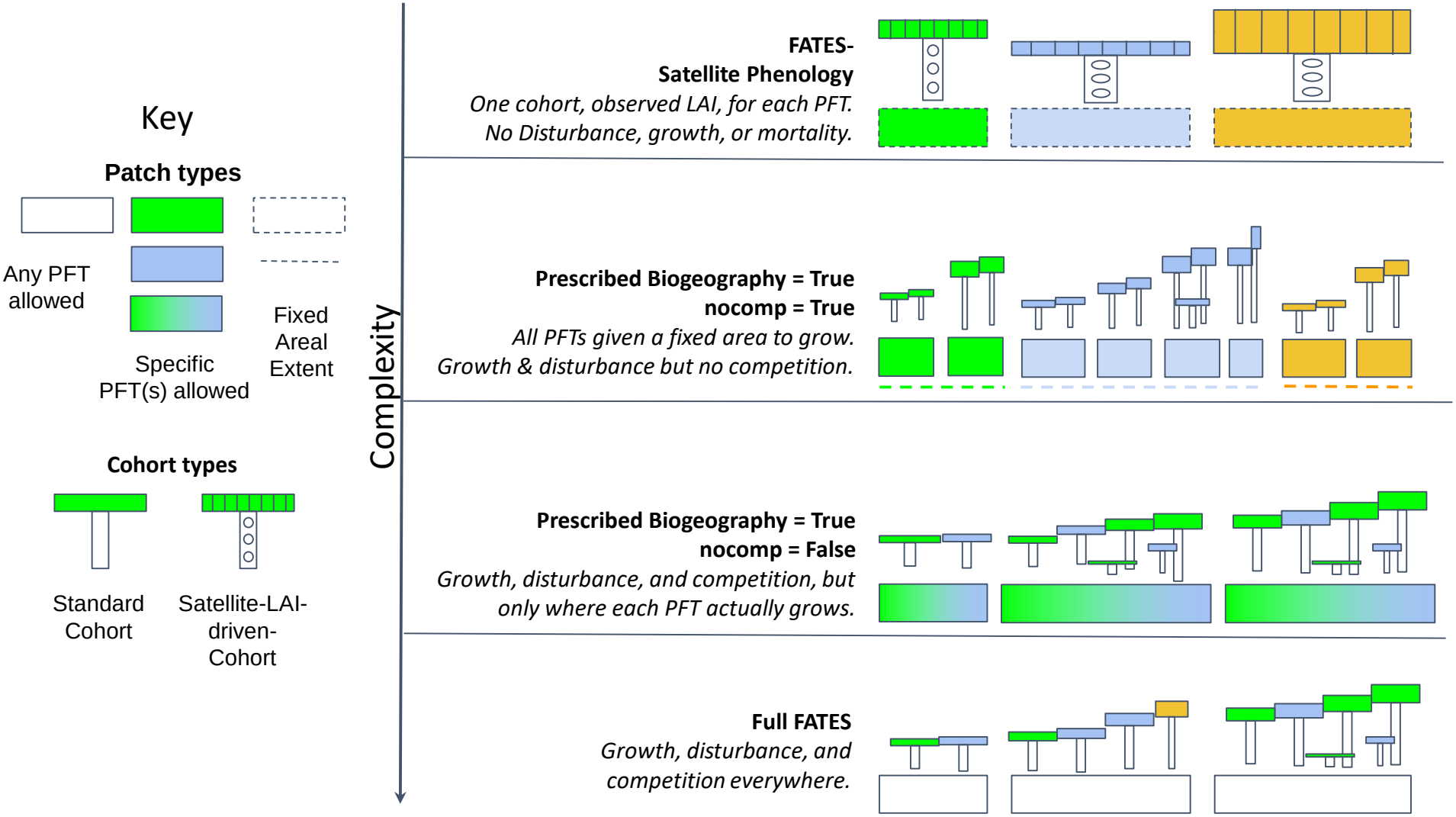
An example of the modular complexity approach: FATES “calibration cascade”



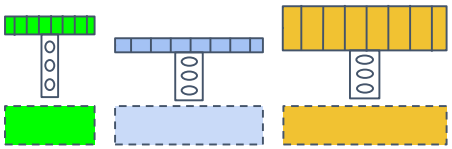
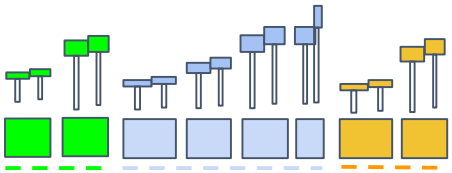
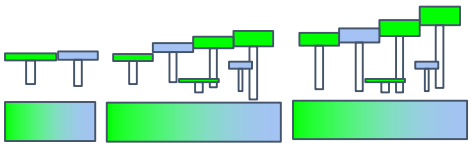
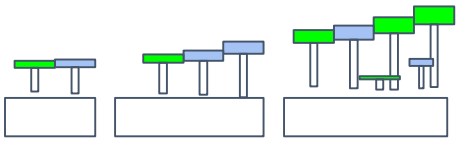
An example of the modular complexity approach: FATES “calibration cascade”



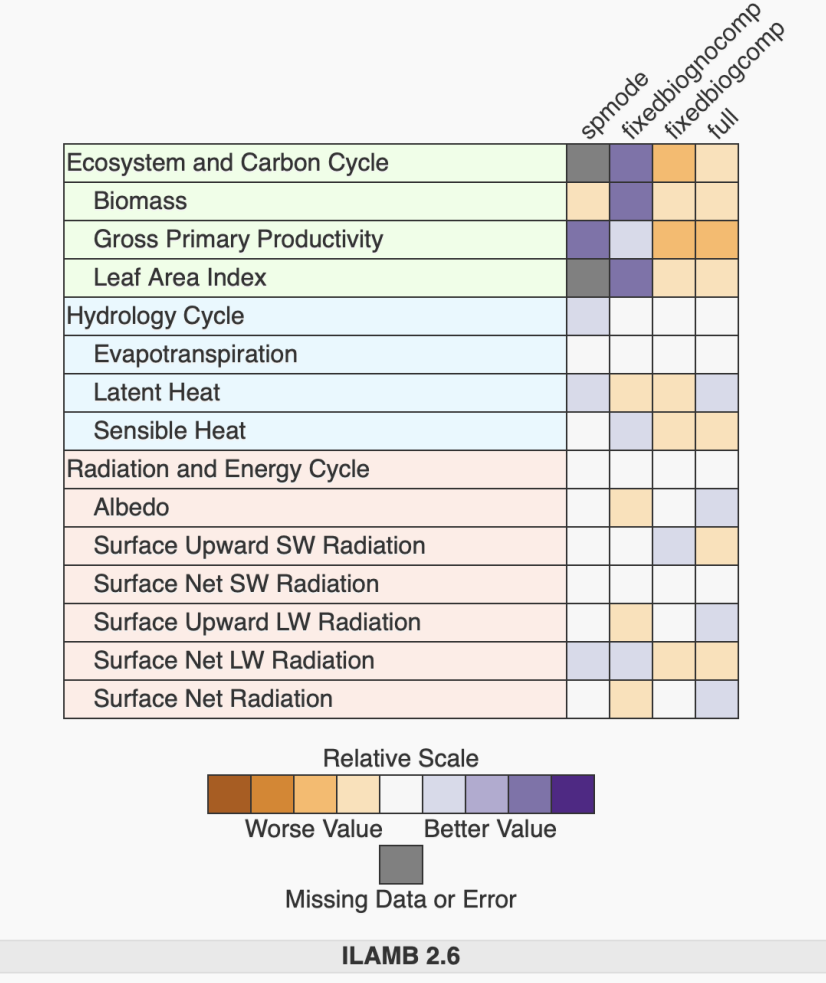
FATES reduced complexity configurations enable calibration cascade



Role of each configuration in calibration cascade and science applications:

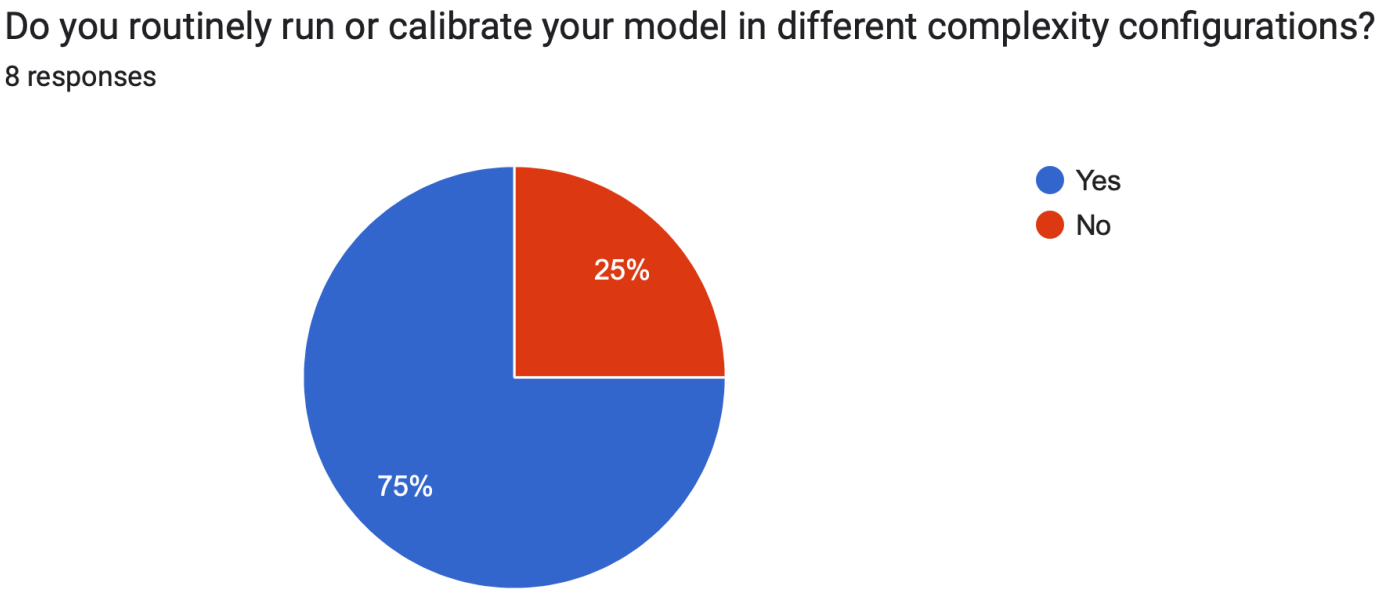
		**Overall role in science and calibrations**	**What variables to calibrate?**
FATES- **Satellite Phenology** *One cohort, observed LAI, for each PFT.* *No Disturbance, growth, or mortality.*		Biophysics and land-atmosphere exchange. Fast spinup, few feedbacks.	Leaf traits, soil parameters, hydraulic conductivities
Prescribed Biogeography = True **nocomp = True** *All PFTs given a fixed area to grow.* *Growth and disturbance but no competition.*		Carbon cycling and demography in absence of competition between PFTs for light	Allometry, allocation, phenology, growth, respiration, mortality parameters
Prescribed Biogeography = True **nocomp = False** *Growth, disturbance, and competition, but only where each PFT actually grows.*		Competition of plants, with some controls over what PFTs can compete	Environmentally-sensitive growth and mortality parameters
Full FATES *Growth, disturbance, and competition everywhere.*		Full dynamics of model	Test of final outcome: does the model capture observed patterns?

An initial benchmarking of ELM-FATES across the complexity cascade

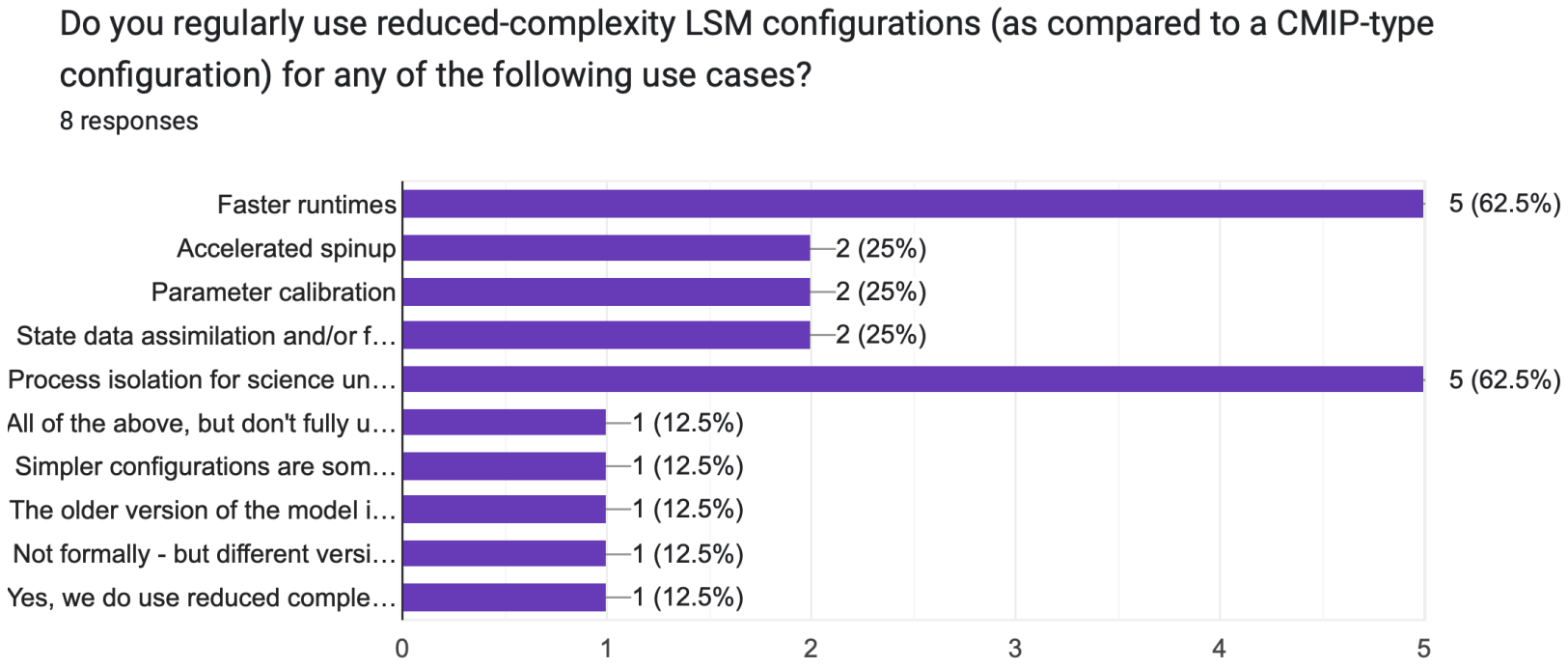


Slide courtesy Jessie Needham

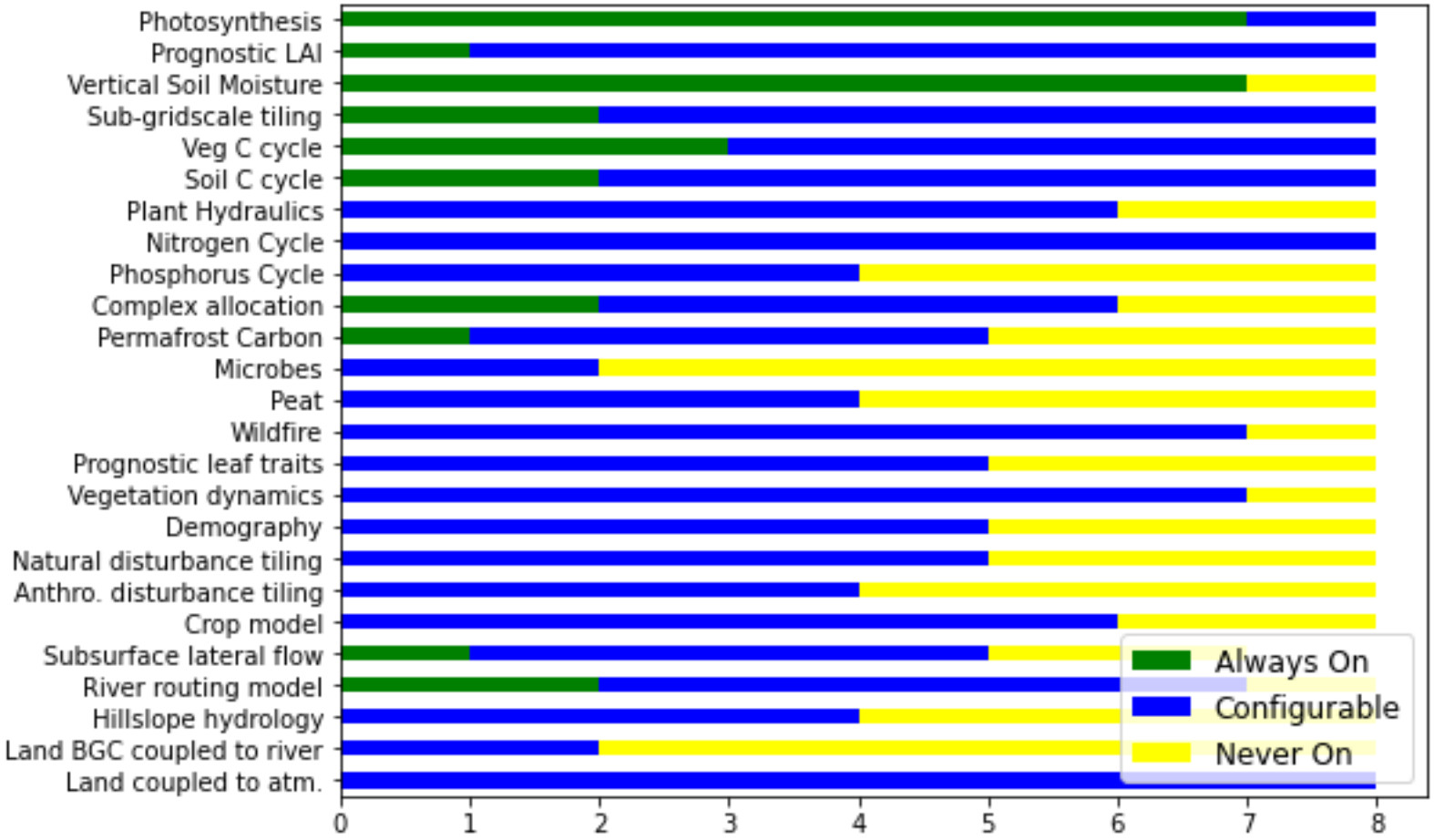
And now some survey results…



And now some survey results…



Which aspects of models are configurable versus always on or off?



Conclusions and possible breakout discussion seeds

- Approaches are needed to manage complexity
- Simplified model configurations can be useful for a huge variety of things: calibration, experiments, …
- Most or all modeling centers are already doing this
- Should we strive to define (more) common reduced complexity configurations and do MIPs, benchmarking or similar around them?
- Are there opportunities for (more) sharing of workflows, reduced-complexity configurations or similar between models?

Thanks!

U.S. DEPARTMENT OF ENERGY | Office of Science