Community tools for predictive ecosystem assembly

Rosie Fisher
Terrestrial Sciences Section
National Center for Atmospheric Research
Boulder, Colorado

With thanks to
Charlie Koven, Ryan Knox, Lara Kueppers, Chonggang Xu
Dave Lawrence, Gordon Bonan & many others
“What you would do now if you could start over developing an Earth System Model and would have complete freedom in doing so?”
1. I don’t think we should start again...
2. Why would one want to start over again?
1. I don’t think we should start again...
2. Why would one want to start over again?

To make a significantly simpler model?

To make some sort of data-driven machine-learning land surface scheme?

To construct the basic model around some alternative theoretical representation for which existing codebases are not useful?
1. I don’t think we should start again…
2. Why would one want to start over again?

To make a significantly simpler model? (see Abby’s talk)
Experiments and observations happen in the real world…

To make some sort of data-driven machine-learning land surface scheme?

To construct the basic model around some alternative theoretical representation for which existing codebases are not useful?
1. I don’t think we should start again…
2. Why would one want to start over again?

To make a significantly simpler model? (see Abby’s talk)
Experiments and observations happen in the real world…

To make some sort of data-driven machine-learning land surface scheme?
Available “training data” are not relevant to high CO2 etc. world… Out of sample problem

To construct the basic model around some alternative theoretical representation for which existing codebases are not useful?
1. I don’t think we should start again...

2. Why would one want to start over again?

To make a significantly simpler model? (see Abby’s talk)
Experiments and observations happen in the real world...

To make some sort of data-driven machine-learning land surface scheme?
Available “training data” are not relevant to high CO2 etc. world...
Out of sample problem

To construct the basic model around some alternative theoretical representation for which existing codebases are not useful?
Not sure I know of any theory sufficiently revolutionary to justify starting –again-, yet...
1. I don’t think we should start again...
2. Why would one want to start over again?
3. If not start again, then what should we do?
1. I don’t think we should start again…
2. Why would one want to start over again?
3. If not start again, then what should we do?
   What is the problem we are trying to address?

   Models are too complicated?

   Models are too slow?

   Models are not enough like the real world?
1. I don’t think we should start again...
2. Why would one want to start over again?
3. If not start again, then what should we do?
   What is the problem we are trying to address?

Models are too complicated?
   We can build hierarchical complexity into ESMs. (see Abby & Charlie’s talks)
Models are too slow?
   We haven’t even really started proper code optimization, in most cases.
Models are not enough like the real world?
   Improving this is “business as usual”
1. I don’t think we should start again…
2. Why would one want to start over again?
3. If not start again, then what should we do?
4. What business do I have being un-cynical about the status quo?

- Building open community tools can and will accelerate progress on all fronts
- Given appropriate funding, emerging data products will massively improve model constraints, if compare to equivalent processes in models
- Usage of modern data-model integration packages (ILAMB, PECAN, etc.) will accelerate model improvement process.
1. I don’t think we should start again...
2. Why would one want to start over again?
3. If not start again, then what should we do?
4. What business do I have being un-cynical about the status quo?
5. Still, there are some things that would help accelerate progress
Some cultural shifts that might help

1. **Stop it with the “first to add/change a thing” papers**: Editors should mandate sober analysis of new components, discussion of uncertainties in structure & parameters, limitations, etc.

2. **Give less weight to the concept of “the model”**: Single sets of structure and parameters that happen to give OK results are not magic. Consideration of model uncertainty and ensemble simulations should be mandatory.

3. **Land surface models aren’t just accessories to the atmosphere, they are the only way we can understand impacts on people and ecosystems**: We need to make that clearer, to funding agencies in particular.
Understanding model components

Initial parameter sensitivity analysis in CLM5

http://www.cesm.ucar.edu/models/cesm2.0/land/
The CLM is now available publically at https://github.com/ESCOMP/ctsm

+ More eyes on code is always better

+ Better coordination of development

+ Forum for collaboration: questions can be directed to whole community

- This requires
Solid funding for maintenance of system
Development of community ethical guidelines (no scooping, etc.)
CLM5 has:

1. A new Nitrogen Cycle, with carbon marketplace for N uptake (FUN), optimal allocation of N within leaves based on optimal/co-limitation theory (LUNA) and flexible vegetation C:N ratios (FLEXCN)
2. Plant Hydrodynamics Scheme (based on Sperry & Love 2015)
4. Revised interception, snow, depth-to-bedrock map, soil evaporation and aquifers.
5. Dynamic crops, harvest, fertilizer, irrigation etc.

Productivity in CLM5 responds less to Nitrogen and more to CO2 than its forbears.

What controls this response?
### Parameter Perturbations: Focus on Carbon and Nitrogen cycling parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Name</th>
<th>Range Determined By</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Leaf Area</td>
<td>SLATOP</td>
<td>TRY database</td>
</tr>
<tr>
<td>Leaf C:N ratio</td>
<td>LEAFCN</td>
<td>TRY database</td>
</tr>
<tr>
<td>Root:leaf ratio</td>
<td>FROOT_LEAF</td>
<td>Litton et al. (2011)</td>
</tr>
<tr>
<td>Stem:leaf ratio</td>
<td>STEM_LEAF</td>
<td>Litton et al. (2011)</td>
</tr>
<tr>
<td>Fraction N fixers</td>
<td>FRACFIXERS</td>
<td>Logical Range (0-1)</td>
</tr>
<tr>
<td>Growth Respiration</td>
<td>GRPERC</td>
<td>Atkin et al. 2018</td>
</tr>
<tr>
<td>Stomatal Slope</td>
<td>MEDLYN_SLOPE</td>
<td>Medlyn et al. 2011</td>
</tr>
<tr>
<td>Respiration BaseRate</td>
<td>LMR_INTERCEPT</td>
<td>Atkin et al. 2015</td>
</tr>
<tr>
<td>Fraction Ectomyccorrhizl fungi</td>
<td>PERECM</td>
<td>Logical Range (0-1)</td>
</tr>
<tr>
<td>Flexible CN ratio ‘a’</td>
<td>FUN_FLEX_CN_A</td>
<td>Logical Range (0-1)</td>
</tr>
<tr>
<td>Flexible CN ratio ‘b’</td>
<td>FUN_FLEX_CN_B</td>
<td>Sensitive Range (1-400)</td>
</tr>
<tr>
<td>Flexible CN ratio ‘c’</td>
<td>FUN_FLEX_CN_C</td>
<td>Sensitive Range (1-32)</td>
</tr>
<tr>
<td>N Costs (x6 parameters)</td>
<td>N_COSTS</td>
<td>Sensitive range (4 ord.magnitude)</td>
</tr>
</tbody>
</table>
No-one should be very surprised if changing parameters affects state, etc.
Free Air Carbon Enrichment (FACE) @ Oak Ridge.
CO2 response: Higher leaf allocation

Allocation parameters don’t affect CO2 response
(these act linearly on C cycle)
N uptake parameters DO affect response

Fraction of fixers significantly alters result

As does cost of N uptake

So, the Nitrogen cycle is constraining the CO2 response...

..and in CLM5, plants can buy their way out of that problem!
Photosynthesis parameters are less fundamental to impact.

CO2 response: Higher leaf allocation
C:N flexibility parameters are not of 1st order importance

This is mildly surprising, since ‘diluting’ Nitrogen is one way around limitation

(but diluting Nitrogen also reduces growth)

These parameters are highly unconstrained...

CO2 response: Higher leaf allocation
CLM5 Parameter Sensitivity: Nitrogen Deposition Response

Both CO2 and N responses are dominated by parameters related to Nitrogen cycling.

Model parameters of greatest impact are:
- frac_fixers
- lmr_intercept
- leafcn
- medlyn_slope
This is the simplest example. Across the CLM/ELM/FATES community we are making further efforts in:

1. Inverse Calibration using emulators (Dagon, Sanderson, Lawrence, Fisher et al.)

2. Global sensitivity analysis using FAST (Massoud, Xu, et al.)

3. Robust Assessment of structural uncertainty of individual components (Walker, Rogers, et al.)

4. Variance Decomposition in PECAN (Serbin, Dietze et al.)

5. Data Assimilation (Fox, Hoar et al.)

6. DREAM inverse calibration at site scale (Post, Fox et al.)

But global tuning (geographically and parameterically) remains slightly out of reach (see CLM5 parameter tuning exercise)

We shouldn’t let the perfect be the enemy of the good!
Community Tools for Predictive Ecosystem Assembly
Alternative vegetation representations in ESMs

- Big Leaf Models (‘DGVMs’)
- Cohort models (ED, FATES, etc.)
- Stochastic Individual Models (LPJ-GUESS, SEIB, etc.)
Vegetation Demographics representation is becoming increasingly commonplace in ESMs

<table>
<thead>
<tr>
<th>Model acronym</th>
<th>Name</th>
<th>Vegetation representation</th>
<th>Coupled to ESM?</th>
<th>Stochastic?</th>
<th>Canopy structure</th>
<th>Disturbance history patches?</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEIB</td>
<td>Spatially Explicit Individual-Based model</td>
<td>Individual</td>
<td>MIROC-ESM</td>
<td>Yes</td>
<td>Individuals</td>
<td>No</td>
</tr>
<tr>
<td>LPJ-GUESS</td>
<td>Lund-Potsdam-Jena General Ecosystem Simulator</td>
<td>Individual or Cohort</td>
<td>EC-Earth, RCA-GUESS</td>
<td>Yes (optional for some processes)</td>
<td>Flat-top</td>
<td>Yes</td>
</tr>
<tr>
<td>LM3-PPA</td>
<td>Perfect Plasticity Approximation</td>
<td>Cohort</td>
<td>GFDL-ESM</td>
<td>No</td>
<td>PPA</td>
<td>No</td>
</tr>
<tr>
<td>ED</td>
<td>Ecosystem Demography model</td>
<td>Cohort</td>
<td>RAMS</td>
<td>No</td>
<td>Flat-top</td>
<td>Yes</td>
</tr>
<tr>
<td>ED2</td>
<td>Ecosystem Demography model v2</td>
<td>Cohort</td>
<td>RAMS</td>
<td>No</td>
<td>Flat-top</td>
<td>Yes</td>
</tr>
<tr>
<td>CLM(ED)</td>
<td>Community Land Model with Ecosystem Demography</td>
<td>Cohort</td>
<td>CESM</td>
<td>No</td>
<td>PPA</td>
<td>Yes</td>
</tr>
</tbody>
</table>
A note on FATES

• FATES is the “Functionally Assembled Terrestrial Ecosystem Simulator”
• FATES is a module, designed to run within land surface models, that simulates plant physiology, competition processes, ecosystem assembly and vegetation distribution
• FATES is based on the Ecosystem Demography Model.
• FATES is managed by the NGEE-tropics team at LBNL
• It lives at: https://github.com/NGEET/fates
• ...soon to be open source
Vegetation structure in FATES

Each time-since-disturbance tile contains cohorts of plants, defined by PFT and size.
Benefits of FATES
What do we get for these two axes of extra complexity?

Vertical Competition for light resources, gap formation, succession

Representation of co-existence and functional diversity

Functional diversity mediates responses to climate

Recovery from disturbances

Arbitrary (non-fixed) PFT definition

Closer links between model world and real-world observations
“Trait Filtering” example

PFT distribution depends only on assigned plant physiology...
Two PFT example

How much carbon do plants target in their savings account?

Growing, rather than storing, is a good idea wherever the canopy is closed...
Biome boundary positions for nine different trait trade-offs

- Selection is typically not along only temperature or precip gradients.

- Most trait filtering is related to light competition intensity

- Are we missing processes/traits that allow filtering by temperature?

- Topt, freeze tolerance, recruitment traits?

Fraction of PFT#1
Some concrete goals that remain for the VDM community

- **Process understanding of the location of biome boundaries**
  - "Default" global trait input set (tbd) is a superset of these problem
  - Requires extensive experimentation & detailed analysis of trait databases

- Achieve an operational multi-scale **model-data-fusion toolkit for VDMs**
  - Link site data and experiments into existing benchmarking software

- Closely and specifically **link emerging Earth Observation data streams** (hyperspectral, canopy structure, biomass, ecostress, etc.) with VDM structures for initialization & testing.
IF WE DON’T START AGAIN?
HOW MIGHT WE ORGANIZE IMPROVEMENT OF LSMS?
(with an emphasis on disturbance)

**Inputs**

Input needs (specific to dynamics & disturbance)

- Drivers
  - Exogenous disturbances
  - Fire ignitions

- Initial Conditions
  - Canopy Structure
  - Observable trait distributions?

- Parameters:
  - Detailed analysis of trait databases
  - Add new traits critical to VDMs
  - Define trade-off surface & environmental responses

**Model Development**

Process gaps (specific to dynamics & disturbance)

- Partial Disturbance
  - (defoliation, burning, freezing, drought)

- Resprouting
  - (fire, hurricanes)

- Seed Dispersal
  - (also recruitment triggers)

- Land Use Change Interpretation
  - Interaction with fire, crops

- Insect Outbreaks

**Testbeds**

DVMs predict vastly more things than normal LSMs

- **Intensive testbeds**: Internal mechanisms of resource competition, allocation, physiology etc.

- **Extensive testbeds**: Predictions of trait distributions in space/time.

**Software architectures** (ILAMB, PECAN) for standardized model-data synthesis

---

**calibration**

**process updates**
The End
FATES tutorial (February 2018)

Tutorial lecture.
https://docs.google.com/presentation/d/1x5BDSgFfETdoQd_lxF49tQsDJZzDTaqxhrlYQfCy73Y

Practical session
Methodological Issues


1. Modularity: implementation of alternative physics in same codebase

1. Collaboration. We can and must prevent conflicts and code divergence with frequent communication. There are always more questions to answer than you think.

1. Robust and commonplace interrogation of uncertainties. 'The Model' is one somewhat arbitrary instance of a massive structural and parametric space.

1. Encourage use of common scientific 'testbeds'. (ILAMB, FLUXNET, NGEET)

1. Software Engineers are extremely important. Put them in your budget. Buy them chocolate. Don’t be mean to them ever.
Where FATES lives inside a system of earth system models

- CESM
- E3SM
- CLM
- ELM
- FATES

Earth System Models
Host Land Model
Demographic Vegetation Model
Abrupt changes in ecosystem type might produce large climate impacts!

This depends a lot on which DGVM you use…

1st gen. DGVMs are not fit for purpose

We need to integrate VDMs into ESMs

How close are we to a predictive science of the biosphere?
The Evolution of Land Surface Modeling

- Surface Energy Fluxes
- Stomatal Resistance
- Soil Moisture
- Plant Canopies
- Heterogeneity
- Lakes, Rivers, Wetlands
- Groundwater
- Dynamic Vegetation
- Carbon Cycle
- Land Cover Change
- Crops, Irrigation
- Nutrients
- Urban
- Lateral Flow

- 70's
- 80's
- 90's
- 00's
- 10's
The Evolution of Land Surface Modeling

- Surface Energy Fluxes
- Stomatal Resistance
- Soil Moisture
- Plant Canopies
- Heterogeneity
- Lakes, Rivers, Wetlands
- Groundwater
- Dynamic Vegetation
- Carbon Cycle
- Land Cover Change
- Urban
- Nutrients
- Crops, Irrigation
- Lateral Flow
- Demographics

70's  80's  90's  00's  10's
To DO

1. Benefits of fates slide