Challenges in process-based hydrologic modeling

GEWEX Open Science Conference, Canmore, 9 May 2018

Main collaborators/funding

NCAR UCAR


Bart Nijssen


Jessica Lundquist


Andy Wood


Gab Abramowitz


Dmitri Kavetski


Julie Vano

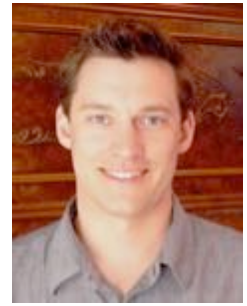

Luis Samaniego


Ying Fan


Ethan Gutmann


Naoki Mizukami


Sean Swenson


Keith Musselman


Ken Nowak


Dave Lawrence


Andy Newman


Jeff Arnold


Tom Giambelluca


Nans Addor


Roy Rasmussen

Funding: US Army Corps of Engineers, U.S. Department of the Interior Bureau of Reclamation, NOAA, NASA, NSF

The interdisciplinary evolution of land models

Land as a lower boundary to the atmosphere → **Land as an integral component of the Earth System**

Focus on land-atmosphere energy fluxes

Limited representation of land processes & feedbacks

→

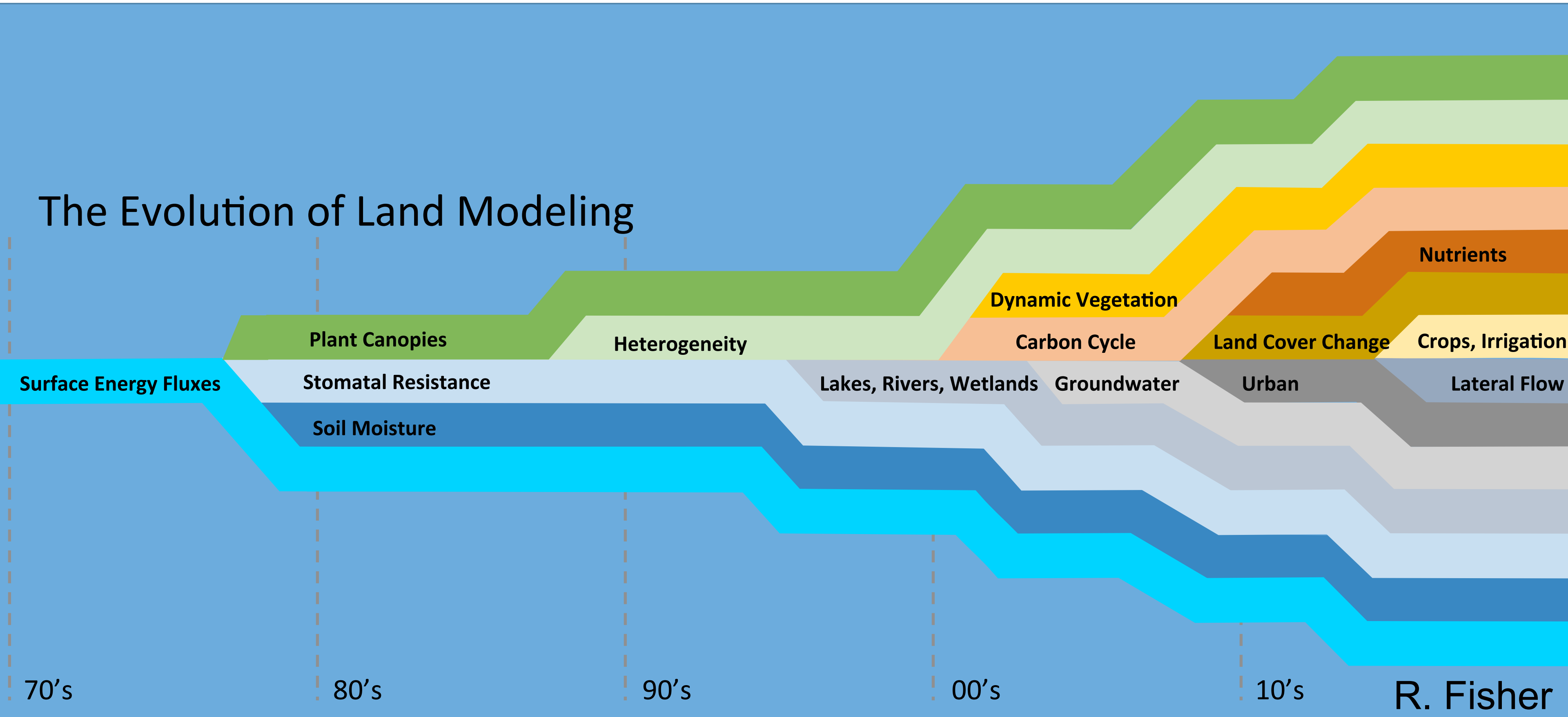
Mechanistic modeling of land processes

Properties define processes (focus on short-term fluxes)

→

Simulate the dynamics of change (e.g., dynamic vegetation)

Processes define properties (feedbacks and interactions across time scales)



Context (Clark et al., HESS 2017)

NCAR UCAR

The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism

Martyn P. Clark[1], Marc F. P. Bierkens[2], Luis Samaniego[3], Ross A. Woods[4], Remko Uijlenhoet[5], Katrina E. Bennett[6], Valentijn R. N. Pauwels[7], Xitian Cai[8], Andrew W. Wood[1], and Christa D. Peters-Lidard[9]

[1]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO 80301, USA
[2]Faculty of Geosciences, Utrecht University, 3508 TC Utrecht, The Netherlands
[3]UFZ-Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany
[4]Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, UK
[5]Hydrology and Quantitative Water Management Group, Wageningen University, 6700 AA Wageningen, the Netherlands
[6]Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[7]Department of Civil Engineering, Monash University, Victoria 3800, Australia
[8]Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA
[9]Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

Correspondence to: Martyn P. Clark (mclark@ucar.edu)

Received: 28 December 2016 – Discussion started: 16 January 2017
Accepted: 31 May 2017 – Published: 11 July 2017

- Paper in celebration of Eric Wood's research contributions over the last 40 years
- Recognize that all modeling groups are trying to solve the same problems, regardless of model type and complexity

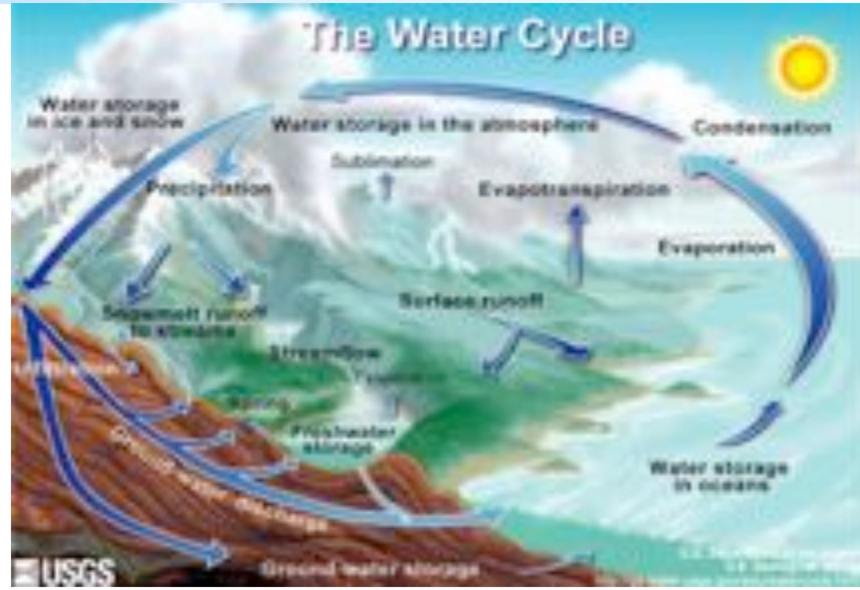
Outline

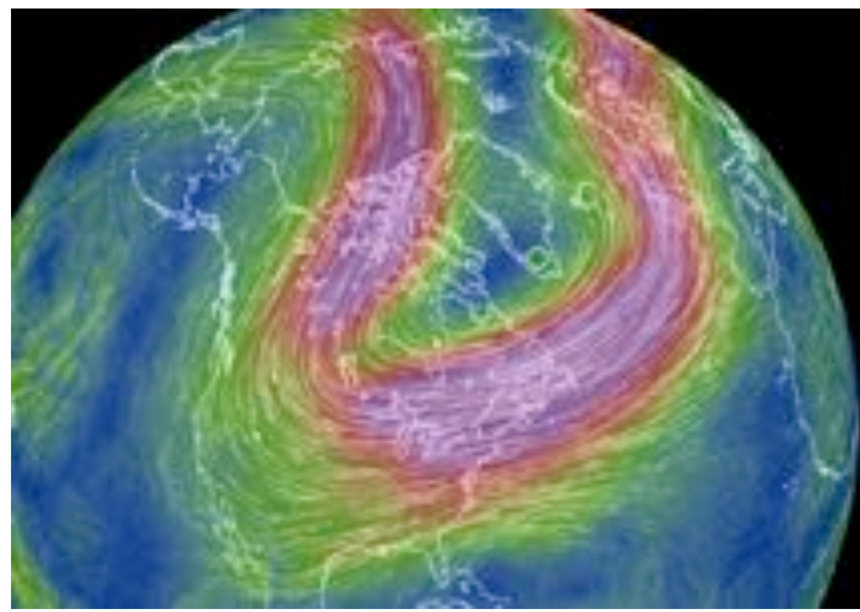
NCAR UCAR

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The Community Terrestrial System Model (CTSM)
  - Development process
  - Status
- Summary and research needs

Hydrologic vs. atmospheric modeling

- Modeling the terrestrial water cycle depends on the (unknown) details of the landscape
- Increases in horizontal resolution often do not lead to increases in hydrologic model performance (especially at larger scales)
- Need creativity in spatial discretization of the model domain and the way that we parameterize fluxes
- Hydrologists have developed a glut of models that differ in almost every aspect of their conceptualization and implementation





The path to model improvement is not obvious…

NCAR UCAR

Physically Based Hydrologic Modeling
2. Is the Concept Realistic?

Prophecy, reality and uncertainty in distributed hydrological modelling

Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system

Searching for the Holy Grail of scientific hydrology: $Q_t = H(S, R, \Delta t)A$ as closure

Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology

Physics-based hydrologic-response simulation: foundation for hydroecology and hydrogeomorphology

Physics-based hydrologic-response simulation: Seeing through the fog of equifinality

Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water

Pursuing the method of multiple working hypotheses for hydrological modeling

A blueprint for process-based modeling of uncertain hydrological systems

Alberto Montanari[1] and Demetris Koutsoyiannis[2]

Beyond "faith-based modeling"?

NCAR UCAR

- The choice of modeling approaches (arguably) stems from personal preferences for physics or parsimony

  - Bucket-style rainfall-runoff models
    - ***Assume that we know nothing***
  - Process-based hydrologic models
    - ***Assume that we know everything***

  - **Need a stronger scientific basis for model development/improvement**
    - Treat numerical modeling as a subjective decision-making process – *carefully evaluate all modeling decisions in a controlled and systematic way*

Outline

NCAR UCAR

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The Community Terrestrial System Model (CTSM)
  - Development process
  - Status
- Summary and research needs

The Freeze and Harlan blueprint (1969)

NCAR UCAR

BLUEPRINT FOR A PHYSICALLY-BASED, DIGITALLY-SIMULATED HYDROLOGIC RESPONSE MODEL

R. ALLAN FREEZE

Inland Waters Branch, Department of Energy, M

Calgary, Alberta, Canada

and

R. L. HARLAN

Forestry Branch, Department of Fisheries and Forestry,

Abstract: In recent years hydrologists have subjected t
hydrologic cycle to intensive study, designed to discover t
arrive at physical and mathematical descriptions of the flo
meaningful results are now available in the form of nume
boundary value problems for groundwater flow, unsaturate
flow, and channel flow. These developments in physical
tremendous advance in digital computer technology, sho
necessary redirection of research in hydrologic simulation
the development of physically-based hydrologic response r
sophistication that can be achieved with presently available
areas for necessary future research are pinpointed.

"The ability to accurately predict beha
severe test of the adequacy of knowledg
subject."

CRAWFORD and L

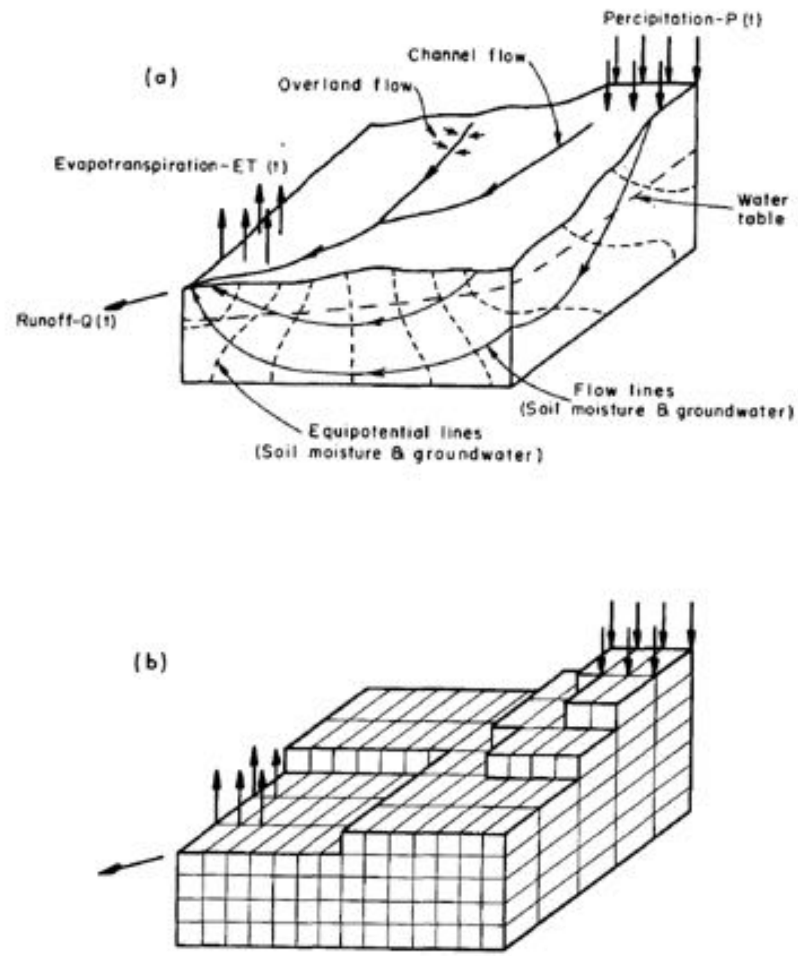


Fig. 3. Schematic diagram of (a) Hydrologic basin and (b) Three dimensional nodal model of hydrologic basin.

Questions posed by Freeze and Harlan

- Are physically based mathematical descriptions of hydrologic processes available? Are the interrelationships between the component phenomena well enough understood? Are the developments adaptable to a simulation of the entire hydrologic cycle?

- Is it possible to measure or estimate accurately the controlling hydrologic parameters? Are the amounts of necessary input data prohibitive?

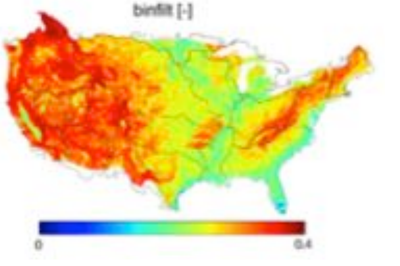
- Have the earlier computer limitations of storage capacity and speed of computation been overcome? Is the application of digital computers to this type of problem economically feasible?

Key challenges

- The choice of modeling approaches (arguably) stems from personal preferences for physics or parsimony
- Need a stronger scientific basis for model development/improvement
    - Treat numerical modeling as a subjective decision-making process – *carefully evaluate all modeling decisions in a controlled and systematic way*

- Processes
    - Many models do not adequately represent dominant processes
    - The spatial gradients that drive flow occur at very small spatial scales and are not resolved by even the finest terrain grid used in large-domain hyper-resolution models

- Parameters
    - Models as mathematical marionettes
    - Vegetation and soils datasets have limited resolution and information content



- Computing
    - The rapid advances in computing are revolutionizing capabilities for simulations with large domain size, more detailed process representation, fine horizontal resolution, and large ensembles
    - The expense of complex models can sacrifice opportunities for model analysis, model improvement, and uncertainty characterization

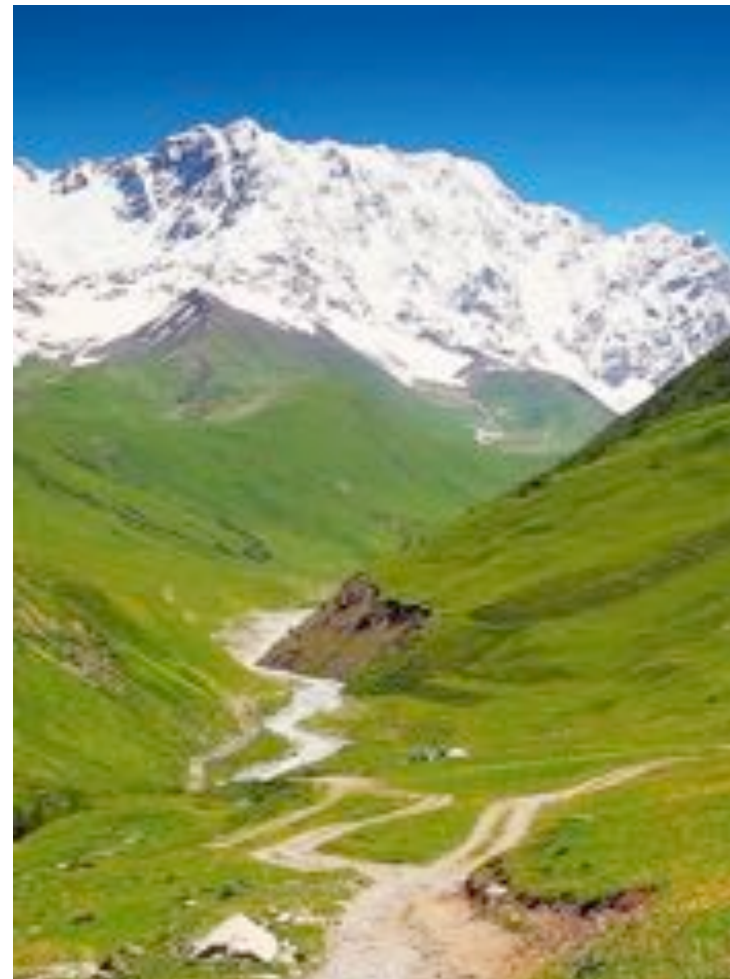
Challenge 1: Modeling processes

NCAR UCAR

- The spatial gradients that drive flow occur at very small spatial scales and are not resolved by even the finest terrain grid used in large-domain hyper-resolution models

- Hot spots and hot moments
  - Small areas of the landscape and short periods of time have a disproportionate impact on large-scale fluxes

- Examples
  - Variable source areas
  - Intermittent turbulence
  - Localized rainfall/snowmelt
  - Riparian transpiration
  - Macropore flow
  - Fill-and-spill
  - …



Two issues: Model proliferation and the shantytown syndrome

NCAR UCAR

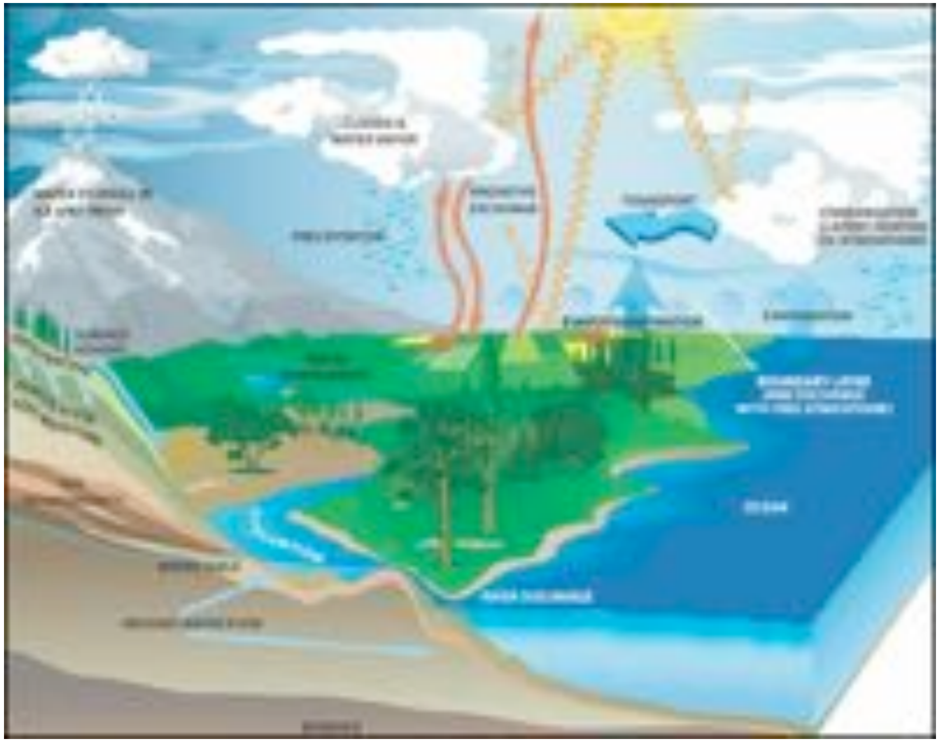
- **Model proliferation**: Every hydrologist has their own model, making different decisions at different points in the model development process

- **The shantytown syndrome**: Ad-hoc approach to model development

- Model proliferation & the shantytown syndrome make it difficult to test underlying hypotheses and identify a clear path to model improvement
- With current model structures, it is easy to incorporate new equations for a given process, ***but very difficult to incorporate new approaches that cut across multiple model components (multi-layer canopy example)***

Modeling approach

NCAR UCAR



General schematic of the terrestrial water cycle, showing dominant fluxes of water and energy

Conceptual basis:

1. Most modelers share a common understanding of how the dominant fluxes of water and energy affect the time evolution of model states

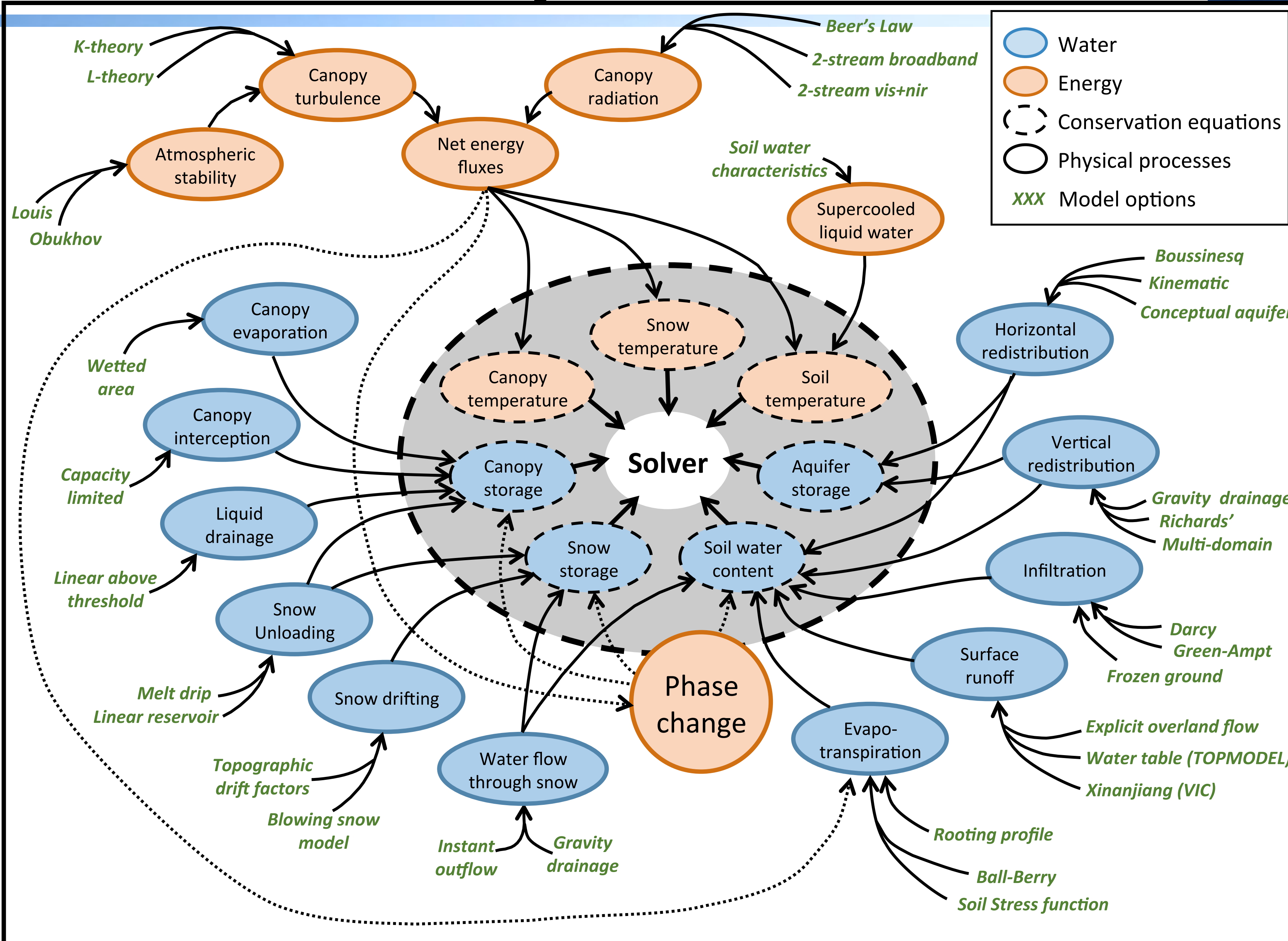
2. Differences among models relate to
   a) the spatial discretization of the model domain;
   b) the approaches used to parameterize individual fluxes (including model parameter values); and
   c) the methods used to solve the governing model equations.

The Structure for Unifying Multiple Modeling Alternatives (SUMMA):

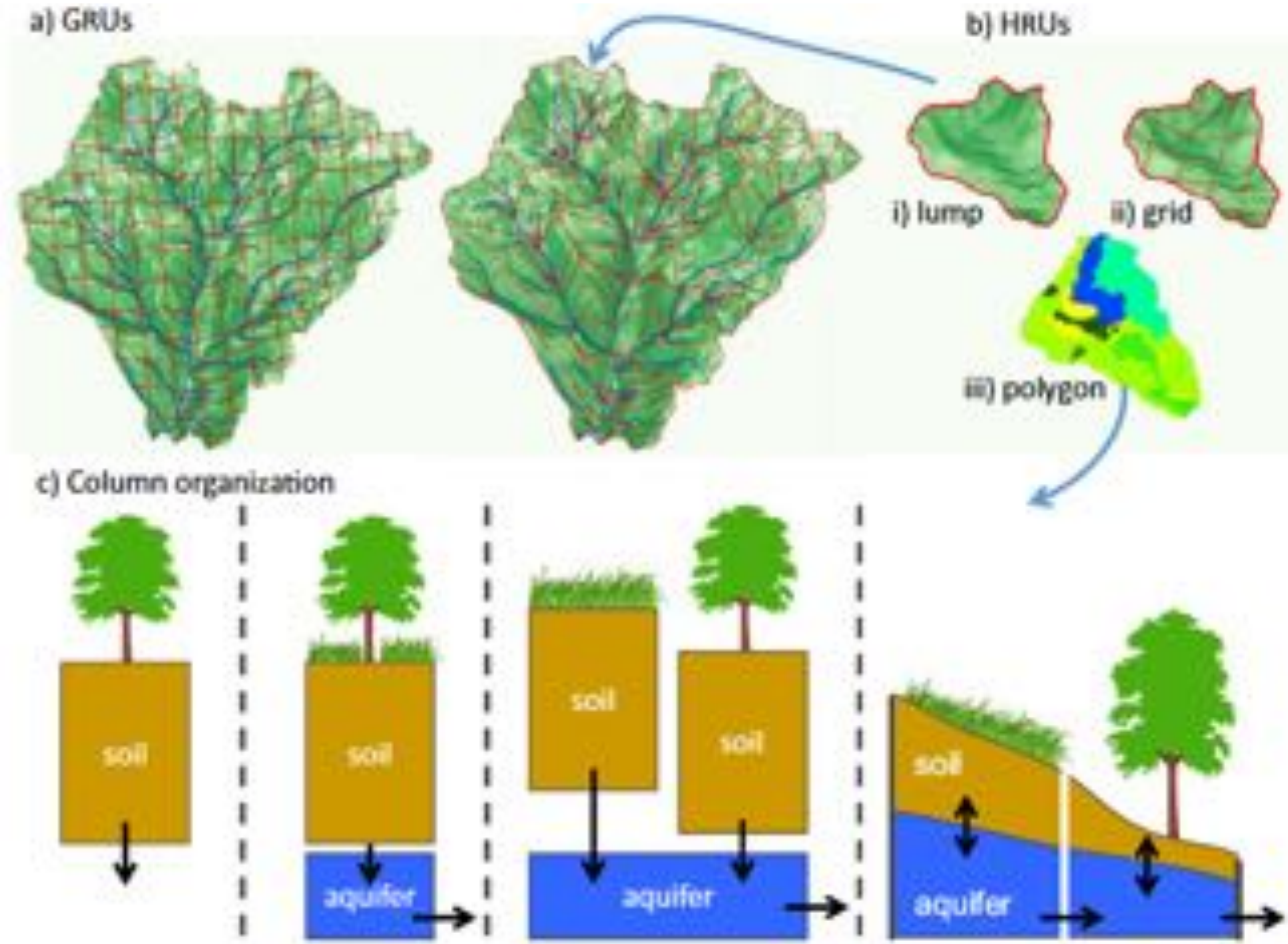
Defines a single set of conservation equations for land biogeophysics, with the capability to use different spatial discretizations, different flux parameterizations and model parameters, & different time stepping schemes

Clark et al. (WRR 2011); Clark et al. (WRR 2015a; 2015b)

Process flexibility



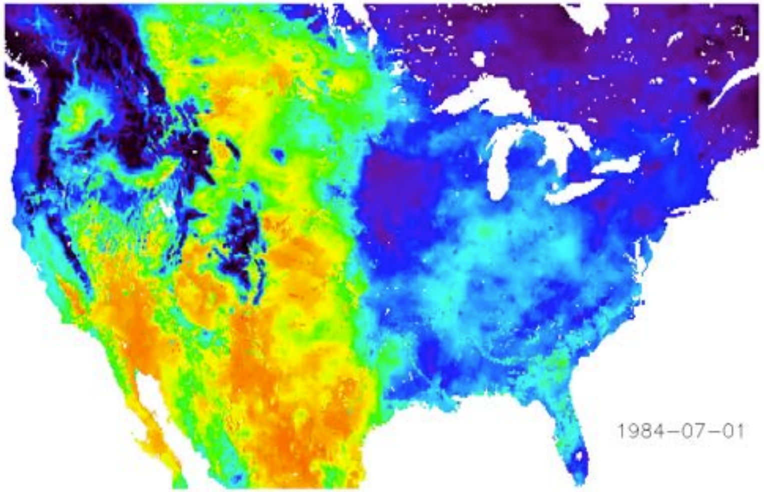
Spatial flexibility



Use cases

- Large-domain extensions
  - Continental-domain simulations now feasible
  - Coupled to mizuRoute, enabling routing on multiple networks
- Model usability
  - A growing set of synthetic test cases and model use cases
  - Extensive stress testing
  - SUMMA in hydroShare

SUMMA simulation of soil water (mm)

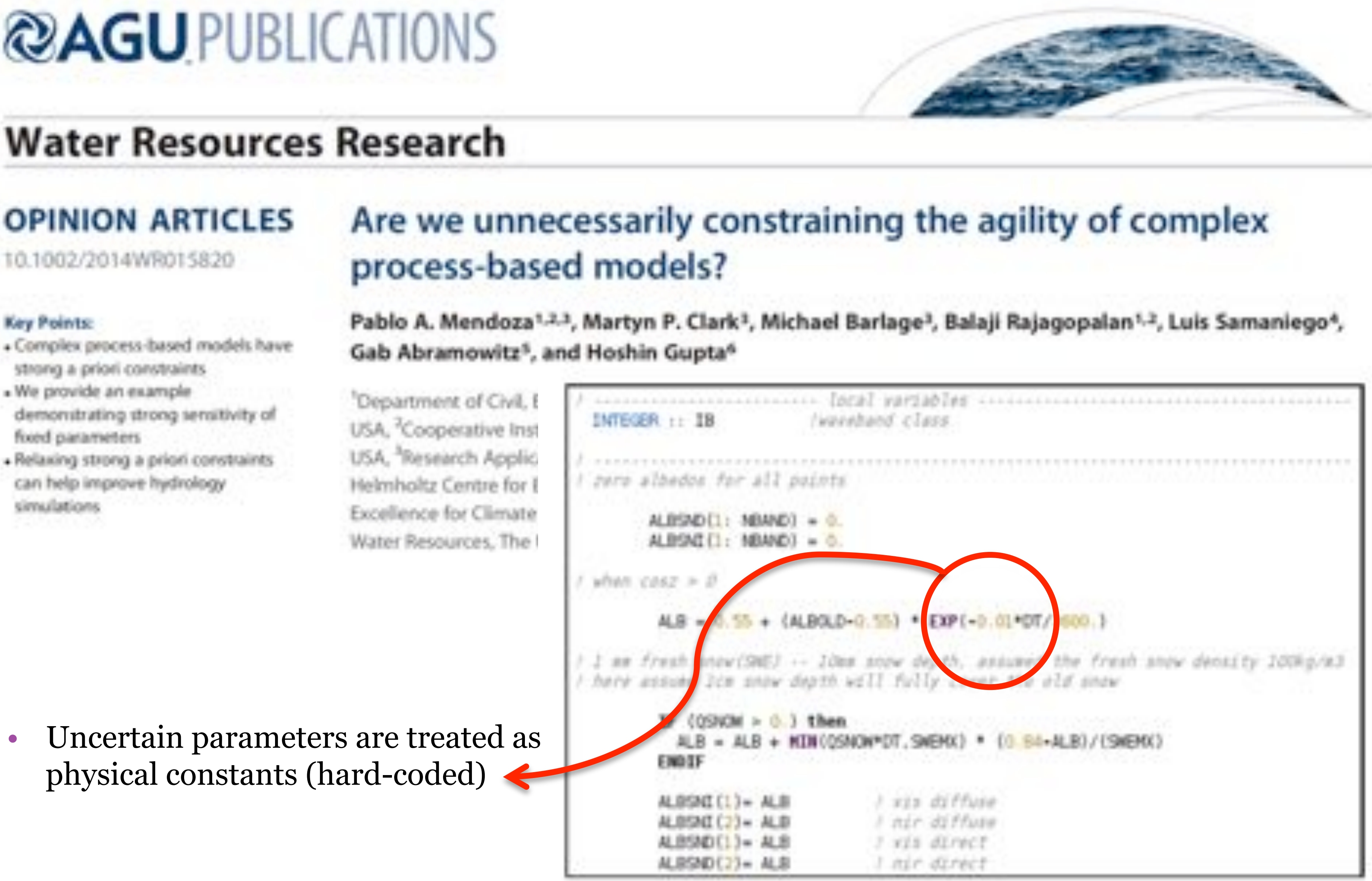


Challenge 2: Model parameters

NCAR UCAR

Challenge 2: Model parameters

NCAR UCAR



AGU PUBLICATIONS

Water Resources Research

OPINION ARTICLES

10.1002/2014WR015820

Are we unnecessarily constraining the agility of complex process-based models?

Pablo A. Mendoza[1,2,3], Martyn P. Clark[3], Michael Barlage[3], Balaji Rajagopalan[1,2], Luis Samaniego[4], Gab Abramowitz[5], and Hoshin Gupta[6]

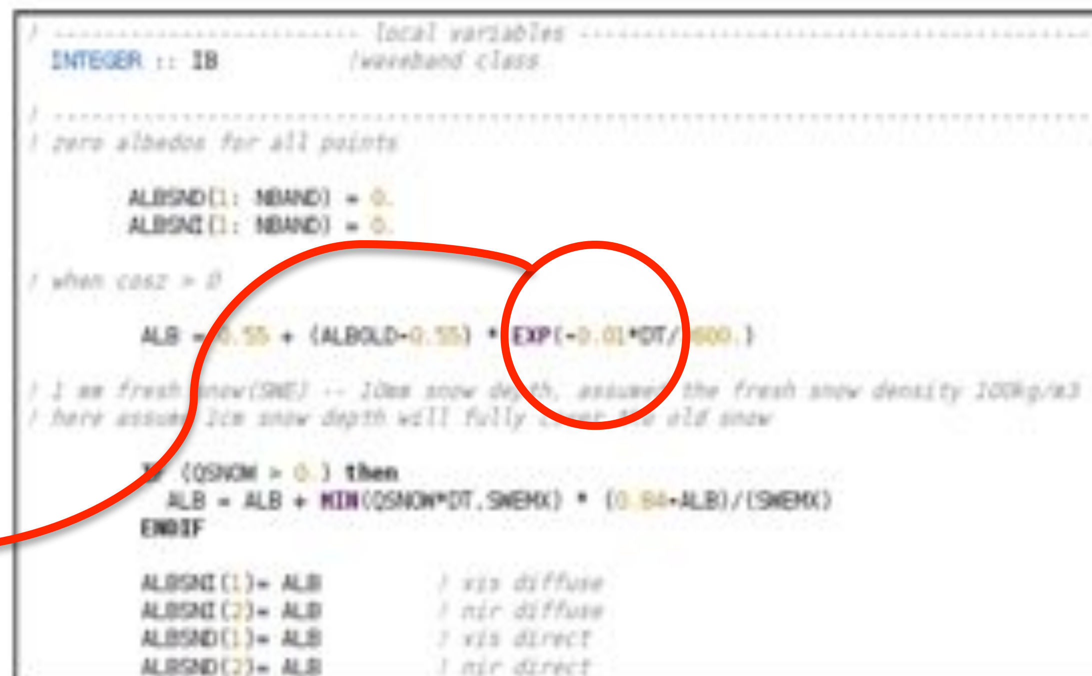
Key Points:
- Complex process-based models have strong a priori constraints
- We provide an example demonstrating strong sensitivity of fixed parameters
- Relaxing strong a priori constraints can help improve hydrology simulations

[1]Department of Civil, E
USA, [2]Cooperative Inst
USA, [3]Research Applic
Helmholtz Centre for E
Excellence for Climate
Water Resources, The U

```
! ------------------------- local variables ---------------------------
  INTEGER :: IB          !waveband class

! ---------------------------------------------------------------------
! zero albedos for all points

       ALBSND(1: NBAND) = 0.
       ALBSNI(1: NBAND) = 0.

! when cosz > 0

       ALB = 0.55 + (ALBOLD-0.55) * EXP(-0.01*DT/3600.)

! 1 mm fresh snow(SWE) -- 10mm snow depth, assumed the fresh snow density 100kg/m3
! here assume 1cm snow depth will fully cover the old snow

       IF (QSNOW > 0.) then
         ALB = ALB + MIN(QSNOW*DT,SWEMX) * (0.84-ALB)/(SWEMX)
       ENDIF

       ALBSNI(1)= ALB          ! vis diffuse
       ALBSNI(2)= ALB          ! nir diffuse
       ALBSND(1)= ALB          ! vis direct
       ALBSND(2)= ALB          ! nir direct
```

- Uncertain parameters are treated as physical constants (hard-coded)

Challenge 2: Model parameters



```
! ------------------------- local variables ---------------------------
  INTEGER :: IB          !waveband class

! -------------------------------------------------------------------
! zero albedos for all points

        ALBSND(1: NBAND) = 0.
        ALBSNI(1: NBAND) = 0.

! when cosz > 0

         ALB = 0.55 + (ALBOLD-0.55) * EXP(-0.01*DT/3600.)

! 1 mm fresh snow(SWE) -- 10mm snow depth, assumed the fresh snow density 100kg/m3
! here assume 1cm snow depth will fully cover the old snow

         IF (QSNOW > 0.) then
           ALB = ALB + MIN(QSNOW*DT,SWEMX) * (0.84-ALB)/(SWEMX)
         ENDIF

         ALBSNI(1)= ALB         ! vis diffuse
         ALBSNI(2)= ALB         ! nir diffuse
         ALBSND(1)= ALB         ! vis direct
         ALBSND(2)= ALB         ! nir direct
```

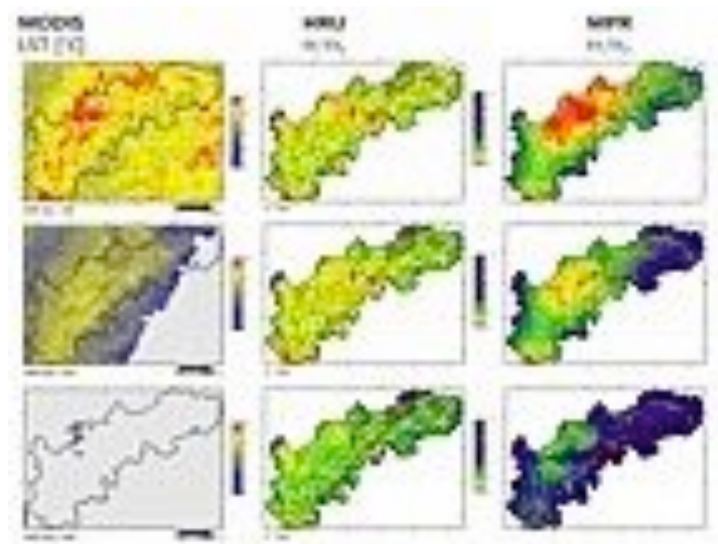
- Uncertain parameters are treated as physical constants (hard-coded)

Challenge 2: Model parameters

- Lack of knowledge of model parameters
  - Vegetation and soils datasets do not have sufficient resolution and information content
    - *Same soil type across large areas (assume no heterogeneity)*
    - *Often limited information on hydraulic properties necessary to simulate heterogeneous hydrologic processes*
  - The rigid structure of complex models (e.g., treating uncertain parameters as physical constants) constrains capabilities to represent spatial variations in hydrologic processes

- One solution: Stochastic hyper-resolution simulation

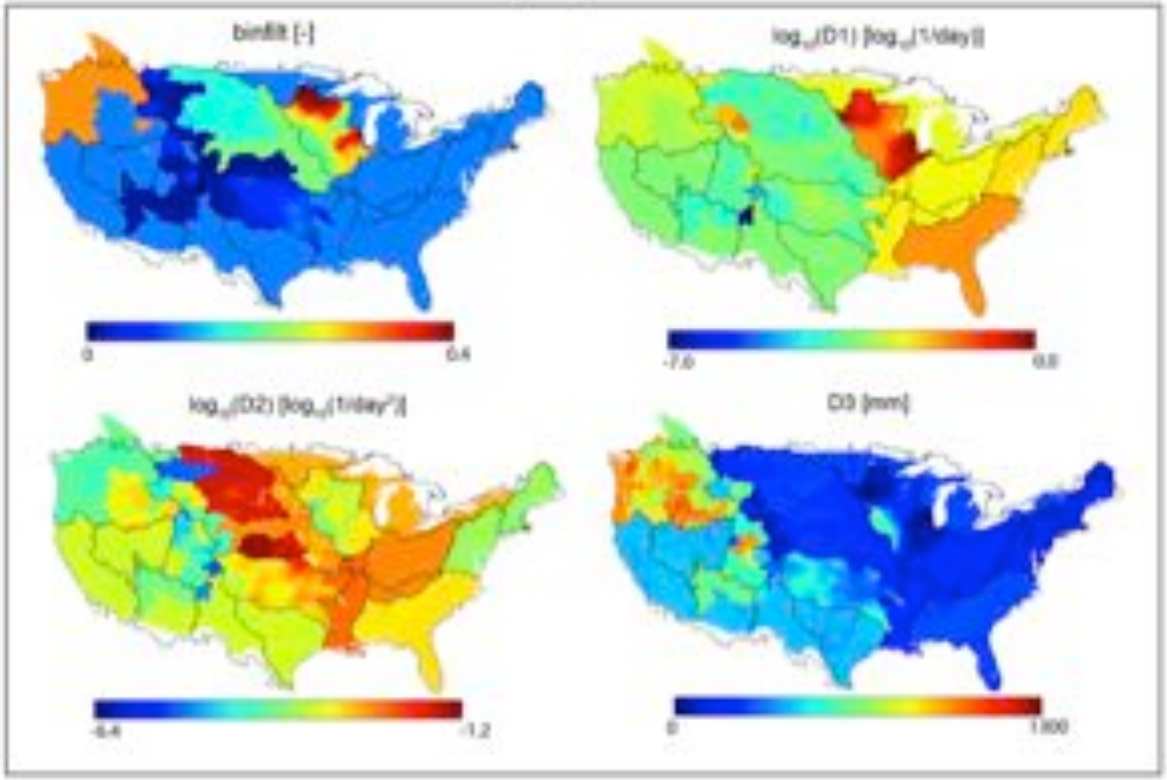
- Another solution: Focus squarely on relating geophysical attributes to model parameters (MPR)



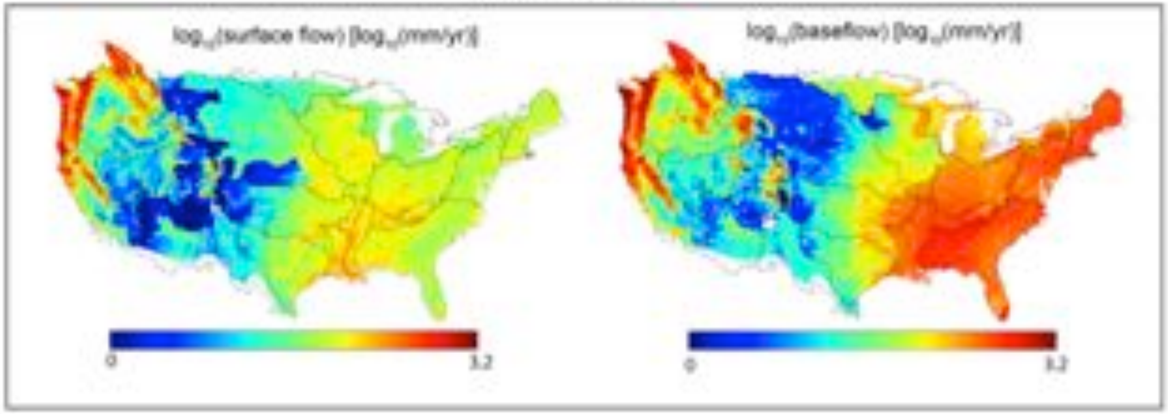
Default params

- Spatial discontinuities in model parameters
- Spatial discontinuities in model simulations

VIC Soil parameters – CMIP5 default



1950-1999 annual mean runoff



Mizukami et al., WRR 2017

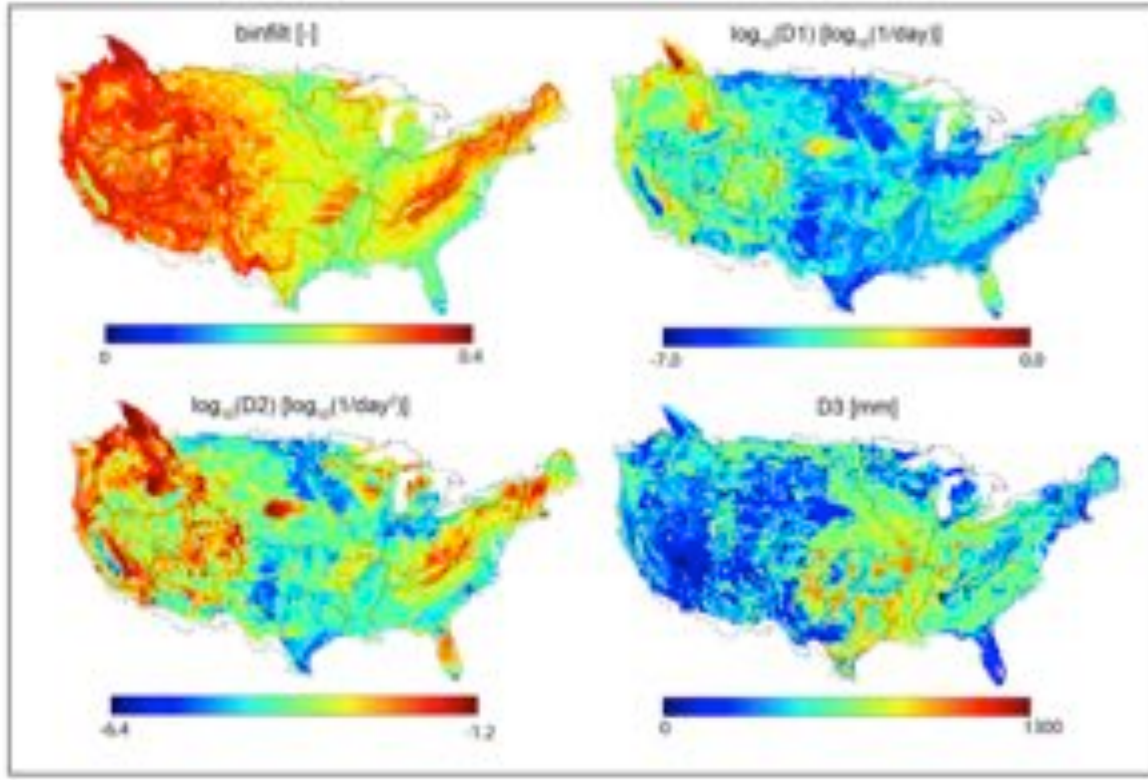
MPR-flex

- Modify coefficients in transfer functions that relate physical attributes (soil, veg, topography) to model parameters
- Use parameter-specific upscaling operators to represent multi-scale behavior

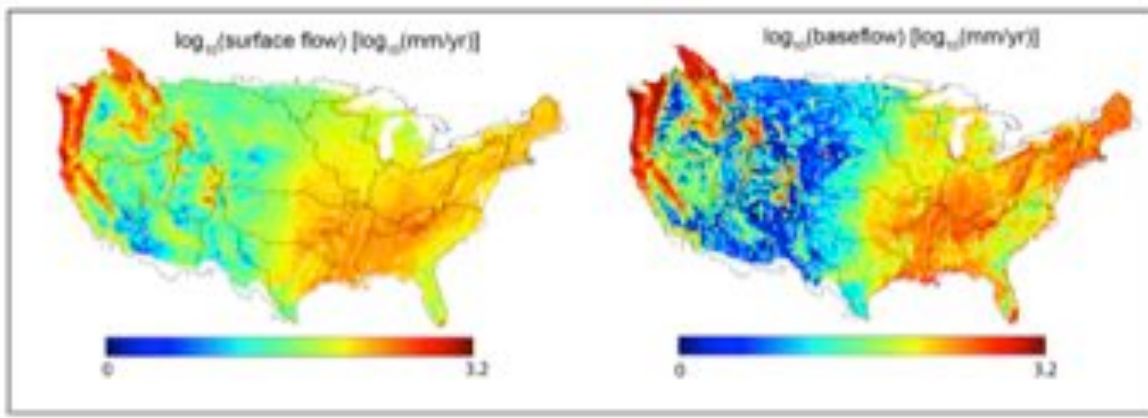
- *Define transfer functions for new models – develop model agnostic MPR (MPR-Flex)*

- No flux discontinuities
- Parameters more closely related to geophysical attributes

VIC Soil parameters – MPR

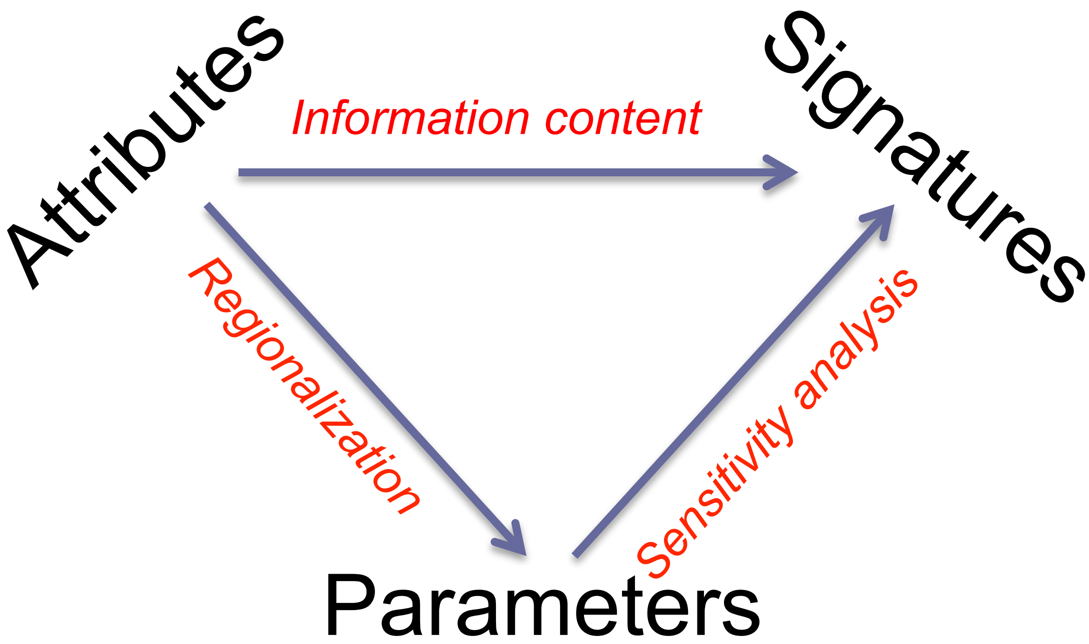


1950-1999 annual mean runoff



Mizukami et al., WRR 2017

Current approaches are unsatisfying



Need to study process interactions across time scales

Instead of the traditional paradigm of properties define processes, study how processes define properties

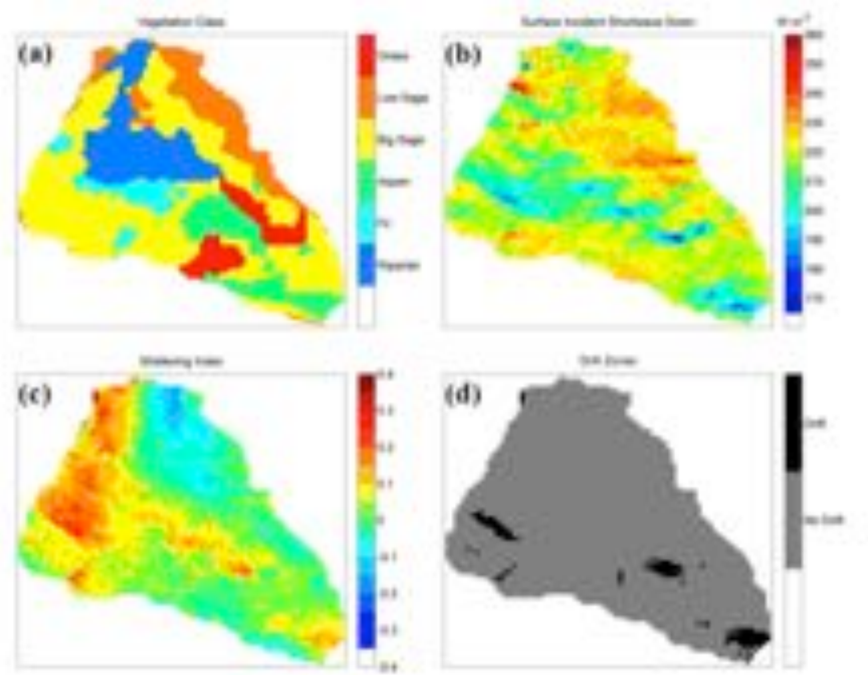
How does landscape evolution define the storage and transmission properties of the landscape?

Challenge 3: Computing

- The computational expense of complex models can sacrifice opportunities for model analysis, model improvement, and uncertainty characterization

- Solutions
  - Hydrologic similarity
  - Representative hillslopes
  - Separate computations for process subsets
  - …

- Recent studies show that similarity methods have the same information content as hyper-resolution models, and orders of magnitude faster



Newman et al., JHM 2014

Computing = understanding complexity

NCAR UCAR

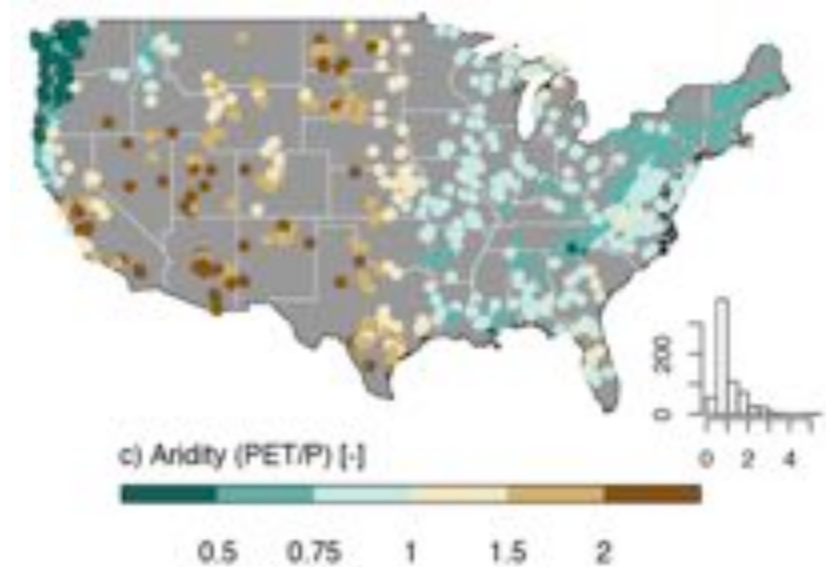
- A continuum of complexity
  - ***Process complexity***: Which processes are represented explicitly?
  - ***Spatial complexity***: To what extent do we explicitly represent details of the landscape, and spatial connections (flow of water) across model elements?

- Bucket-style rainfall-runoff models
  - Lumping of processes, and lumping of the landscape
  - Reliance on inverse methods (calibration) to estimate model parameters
    - *Models as mathematical marionettes, giving the "right" answers for the wrong reasons*
    - *Theoretically unsatisfying*
  - Computationally frugal
    - *Enables use of ensemble methods*
    - *Enables extensive experimentation with different model parameters*

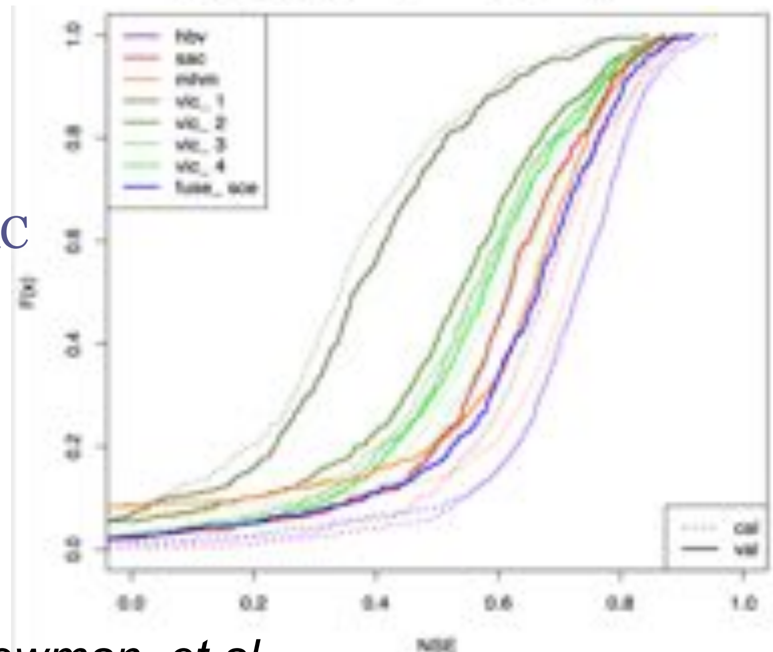
- Process-based hydrologic models
  - Explicitly represent dominant hydrologic and biophysical processes; explicitly represent details of the landscape
  - Reliance on geophysical data to estimate model parameters and widespread use of spatially constant parameters obtained from limited experimental data
    - *Huge challenge in relating geophysical data to model parameters*
    - *Common approach of treating uncertain model parameters as (hard-coded) physical constants*
  - Computationally expensive
    - *Often restricted to a single deterministic simulation*
    - *Limited model analysis (and "tuning") since model is too expensive to calibrate*

Results from many catchments/models

- Large catchment sample
  - Include catchments of varying topography, climate, vegetation and soils
  - Newman et al. (2015), Addor et al. (2017)



- Large model sample
  - Existing models
    - VIC, CLM, Noah-MP, PRMS, HBV, MHM, SAC
  - Multiple hypothesis frameworks
    - FUSE and SUMMA
    - Clark et al., 2008; 2011; 2015a,b



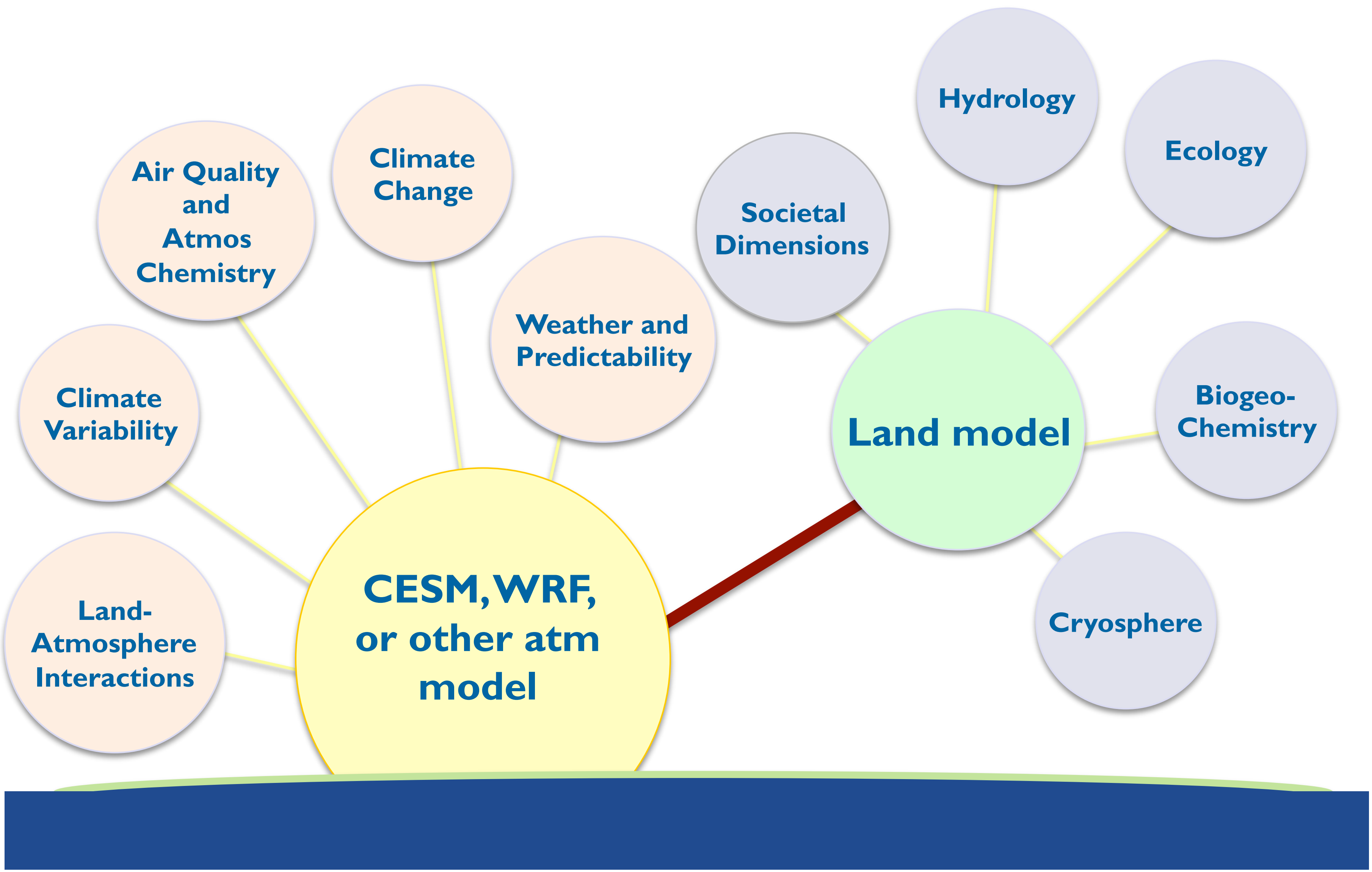
Efforts from Nans Addor, Naoki Mizukami, Andy Newman, et al.

Outline

NCAR UCAR

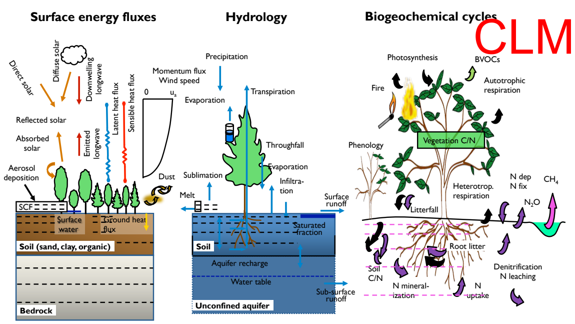
- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The Community Terrestrial System Model (CTSM)
  - Development process
  - Status
- Summary and research needs

The interdisciplinary challenge of land modeling



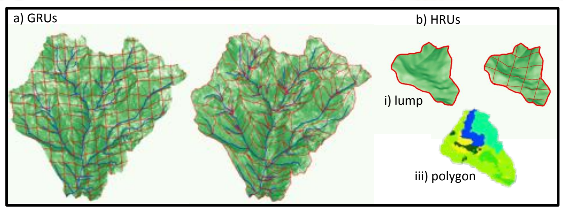
The Community Terrestrial Systems Model

CLM

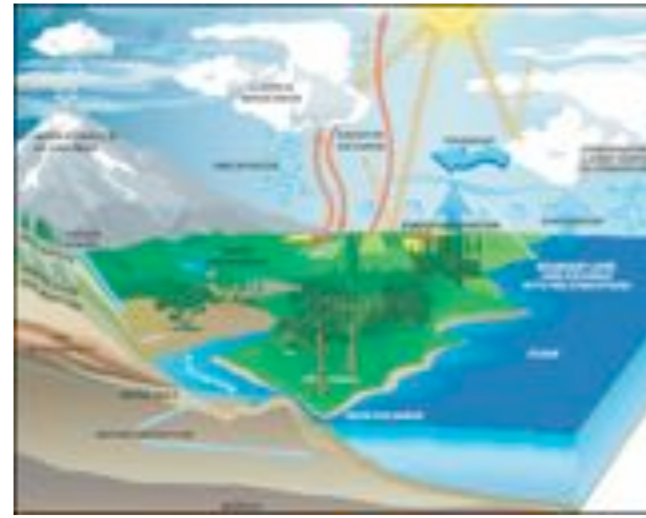


Noah-MP

SUMMA



The Community Terrestrial Systems Model (CTSM)



Conceptual basis

- Modelers agree on many aspects of terrestrial system science
- Differences among models relate to
  - *Flux parameterizations*
  - *Spatial discretization*
  - *Numerical solution*

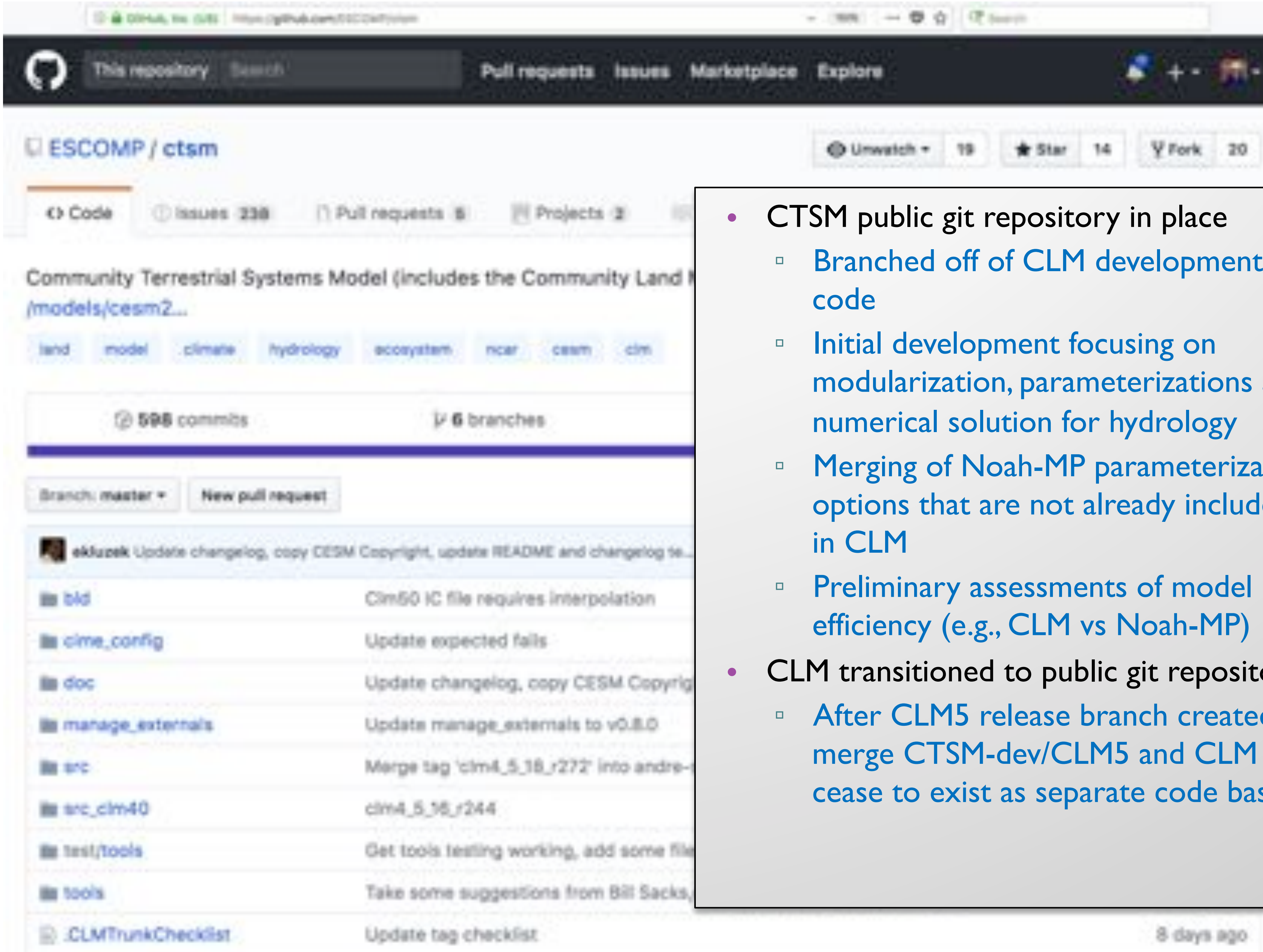
Formulates master model template which multiple models can be derived

- Existing models (*CLM, Noah-MP, WRF-Hydro, etc.*) as a special case

Unifies land models across climate, weather, water, and ecology

- Multiple configurations
- Easy to modify/use
- Centralized support

CTSM is now public

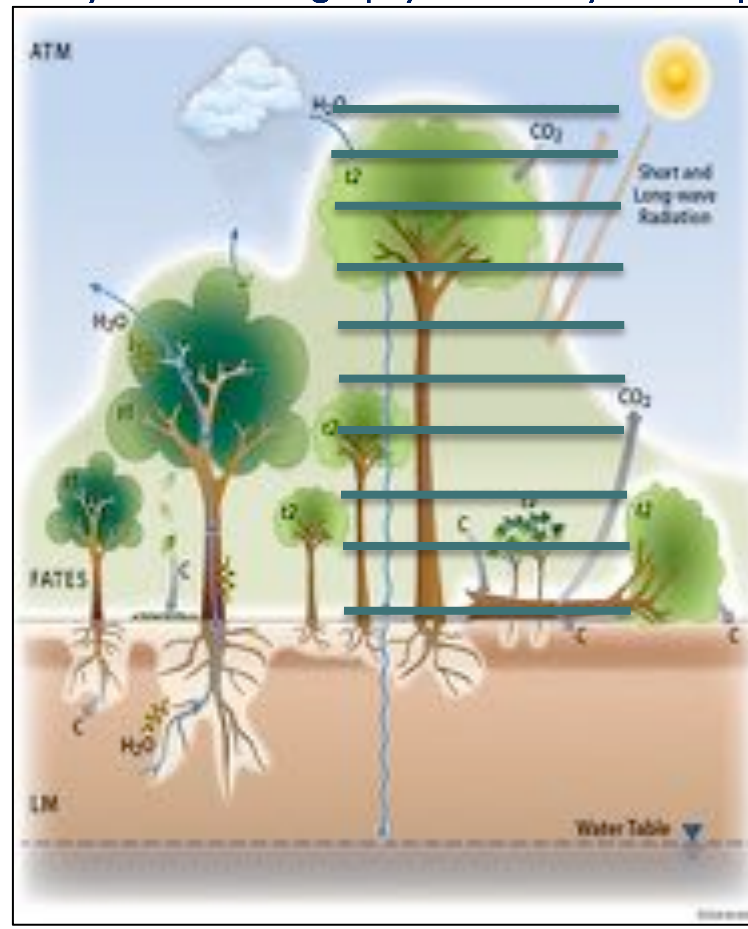


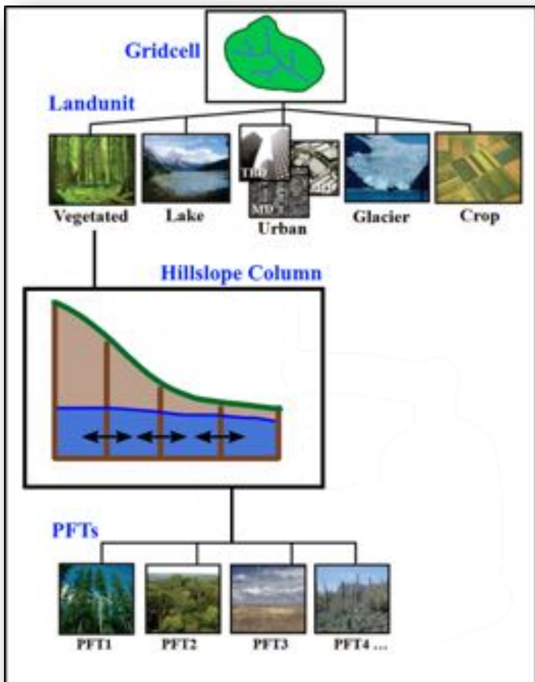
- CTSM public git repository in place
  - Branched off of CLM development code
  - Initial development focusing on modularization, parameterizations and numerical solution for hydrology
  - Merging of Noah-MP parameterization options that are not already included in CLM
  - Preliminary assessments of model efficiency (e.g., CLM vs Noah-MP)
- CLM transitioned to public git repository
  - After CLM5 release branch created, merge CTSM-dev/CLM5 and CLM will cease to exist as separate code base

Plans for the next-generation land model

- Ecosystem vulnerability and impacts on carbon cycle and ecosystem services
- Sources of predictability from land processes
- Impacts of land use and land-use change on climate, carbon, water, and extremes
- Water and food security in context of climate change, climate variability, and extreme weather

Ecosystem Demography / Multi-layer canopy



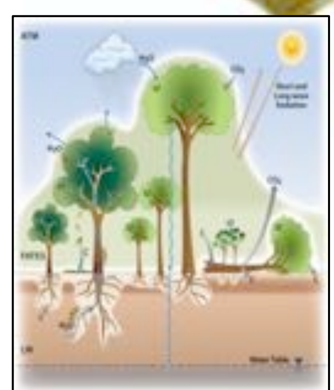


Lateral fluxes of water

Water and land management

Key opportunities

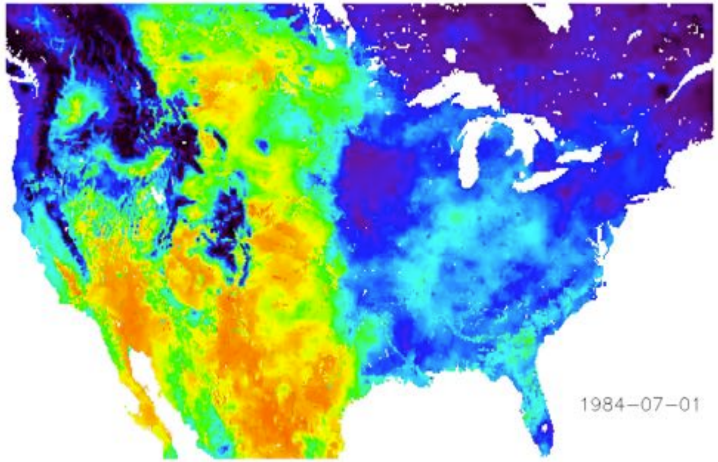
- Land modeling applications in climate, weather, water, and ecology
  - Hydrologic prediction across scales / hydrologic ensemble methods
  - Interdisciplinary advances (e.g., the union of hillslope hydrology and FATES)
  - ESM concepts for short-term prediction problems (e.g., impact of vegetation phenology on meteorological prediction, estimating fuel loads for fire)
- Integrate land modeling expertise
  - Land-atmosphere interactions, hydrologic prediction, water and land management, data assimilation, model analysis
  - Monthly NCAR-wide science discussions
- Simplify incorporating new capabilities in land models
  - Modular structure and separating physics from numerics reduces the in-person cost of modifying CLM, a cost borne by NCAR scientists and software engineers and university collaborators

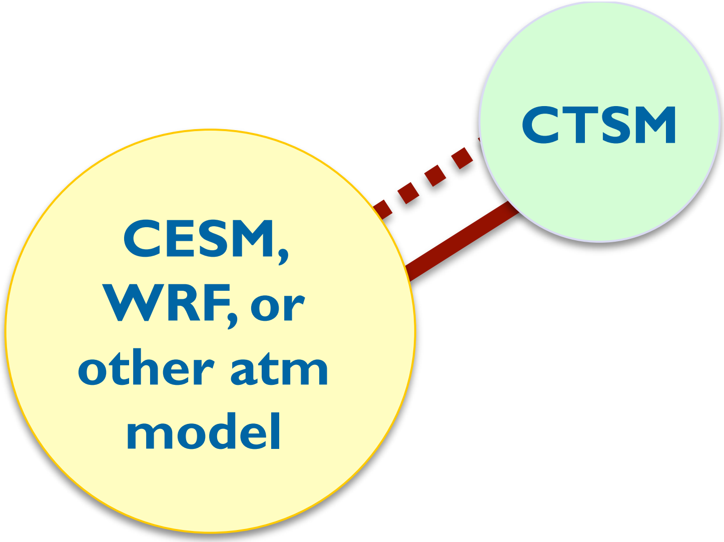


CTSM challenges

- Parallel development
  - Existing models currently used across multiple projects
  - Initially the effort is diffuse (e.g., individuals developing code for both Noah-MP and CTSM)
  - Need to accelerate early applications for different model use cases
  - Rapid prototyping in SUMMA
- Modularity/coupling
  - Support contributions at multiple levels of granularity (e.g., FATES)
  - Community standards for model construction, to simplify sharing code/concepts across model development groups
  - Simplify coupling/ease of use across multiple communities
- Funding
  - Support the interdisciplinary challenge of land modeling

SUMMA simulation of soil water (mm)





LILAC
Lightweight Infrastructure for Land-Atmosphere Coupling
Funded NSF Infrastructure project

Outline

NCAR UCAR

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The Community Terrestrial System Model (CTSM)
  - Development process
  - Status
- Summary and research needs

Summary

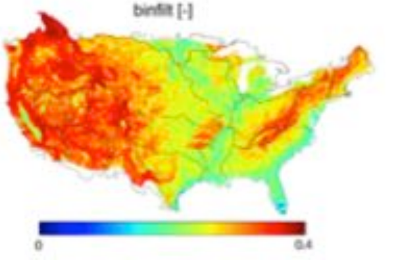
NCAR UCAR

- We need better frameworks to evaluate the myriad of decisions made during model development (multiple hypothesis frameworks + information theory + ...)
- We need to treat parameter estimation as a model development problem

---

- Processes
  - We really need to focus on the scaling problem – how do inter-model differences in large-scale flux parameterizations affect the flow of information through models

- Parameters
  - We really need to incorporate stronger hydrologic theory when evaluating model parameters
  - Process parameterizations and model parameters are highly inter-related and should be considered together



- Computing
  - We should not let the allure of computing advances constrain our capabilities for model analysis (let's not get ahead of our skis)
  - Always make room for model analysis

Modeling strategy

NCAR UCAR

- A three-pronged strategy to improve the physical realism of process-based hydrologic models
  - **Processes**: Isolate and evaluate competing modeling approaches.
  - **Parameters**: Improve the agility of process-based models, and focus squarely on relating geophysical attributes to model parameters
  - **Computing**: Take advantage of hydrologic similarity methods to reduce redundancies in hydrologic models and enable extensive analysis. Explore accuracy-efficiency tradeoffs in numerical solutions.

- Modeling strategy explicitly characterizes model uncertainty, as well as uncertainty in model input/response data
  - Probabilistic QPE
  - Ensembles of alternative model configurations
  - Seek to characterize and reduce uncertainties

- Overall goal: Improve the physical realism of models at any scale through better informed choices about the physics.

Specific research needs

NCAR UCAR

1. Improve the theoretical underpinnings of our hydrologic models
   - A more productive dialog between the experimentalist and modeler

2. Expand our prominence in community hydrologic modeling
   - Provide accessible and extensible modeling tools
   - Provide key research datasets and model test cases
   - Increase the effectiveness and efficiency in sharing data and model source code (simplify the sharing of data and source code developed by different groups)

3. Systematically explore the benefits of competing modeling approaches
   - Scrutinize models using data from research watersheds
   - Evaluate information gains/losses using models of varying complexity
   - Use applications of information theory to quantify how effectively models use the available information

4. Develop new modeling approaches that simulate the temporal dynamics of environmental change
   - How natural selection favors plants
   - How energy gradients dictate landscape evolution
   - How the dynamic interactions between humans and the environment shapes the storage and transmission of water across the landscape

Specific research needs (continued)

NCAR UCAR

5. Advance research on process-oriented approaches to estimate spatial fields of model parameters
   - Estimate spatial variations in storage/transmission properties of the landscape
   - New data sources on geophysical attributes, new approaches to link geophysical attributes to model parameters, and new diagnostics to infer model parameters

6. Obtain better data on hydrologic processes.
   - Motivate and design new field experiments to advance understanding of the terrestrial component of the water cycle across scales and locations.
   - Ensure model development is not unduly constrained by the limited experimental field data that we have at present.

7. Advance methods for model analysis, especially for complex models.
   - Currently very little insight into process/parameter dominance and process/parameter interactions in very complex models
   - Information is desperately needed to inform parameter estimation strategies

8. Improve the construction of hydrologic models.
   - Move beyond the "shantytown" syndrome
   - Enable greater model extensibility and code reuse


QUESTIONS??