Land modeling challenges and the emergence of CTSM

Martyn P. Clark, University of Saskatchewan at Canmore

4 February, 2019

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA

Main collaborators/funding


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

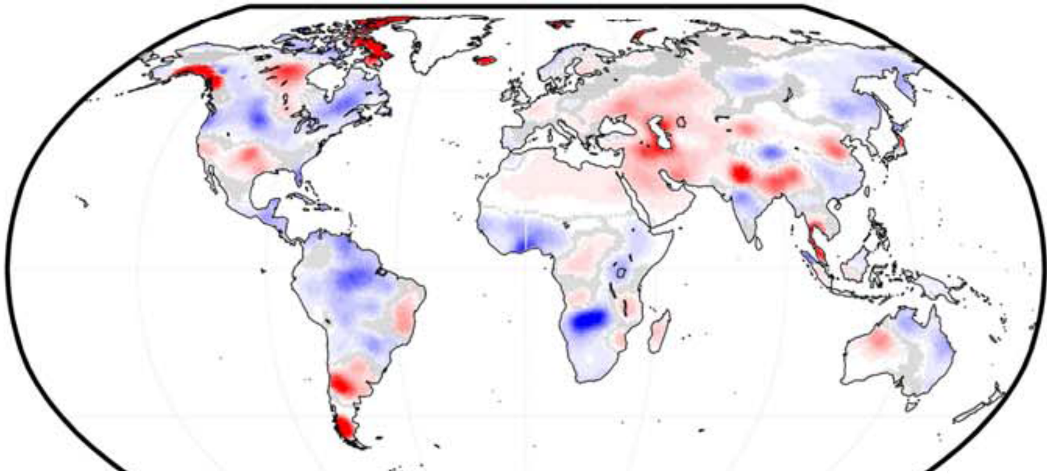
NASA

NSF

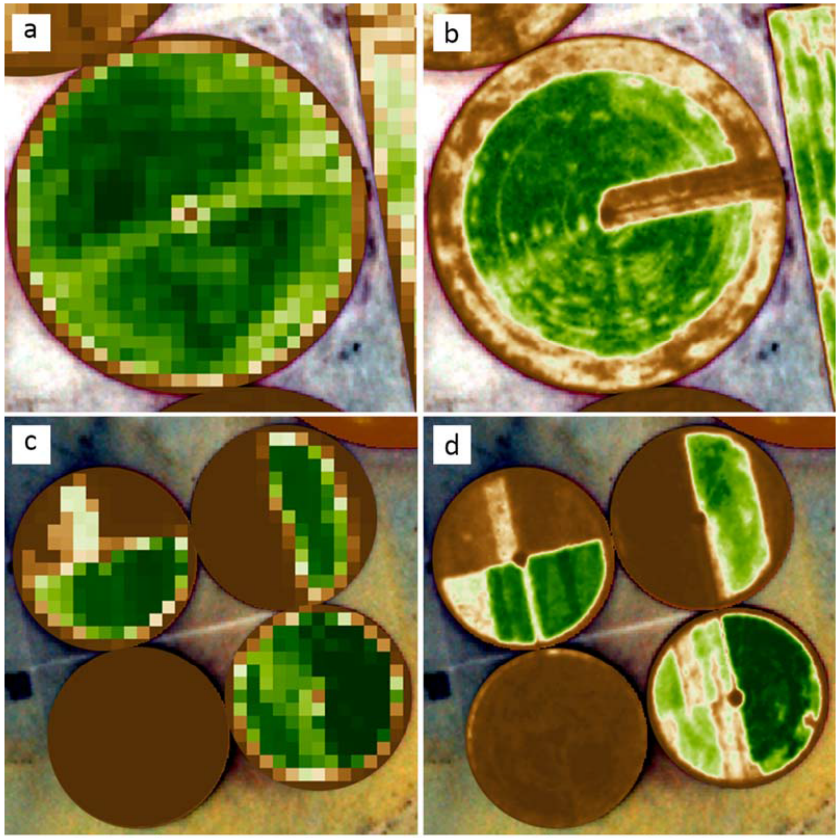
Outline

- Background
  - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The emergence of CTSM
- Summary and research needs

Advances in remote sensing

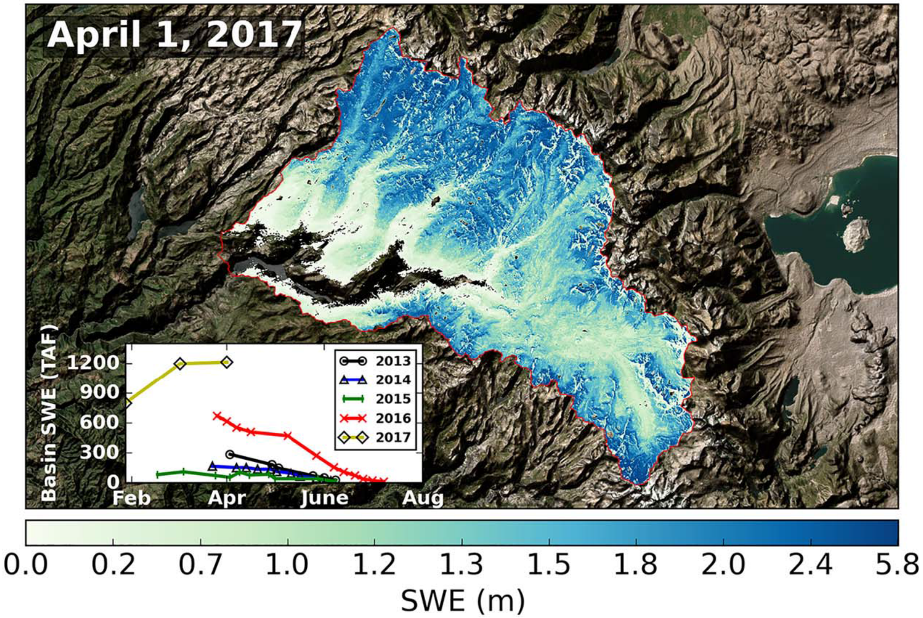


GRACE total water storage estimates
Scanlon et al., WRR 2016



Landsat DOY 339
Planet DOY 340

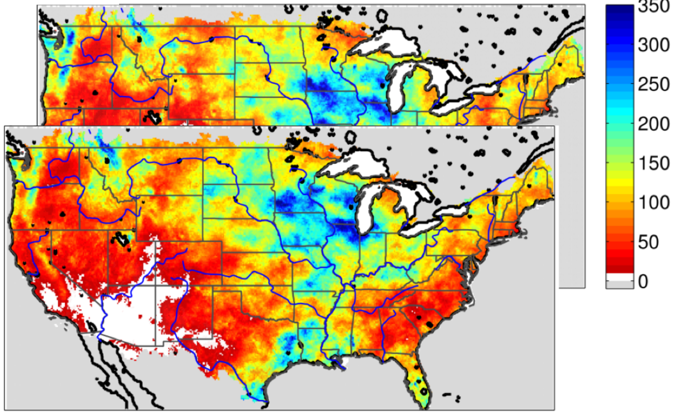
Cubesats in hydrology
McCabe et al., WRR 2017



Airborne LIDAR estimates of snow
Lettenmaier et al., WRR 2017

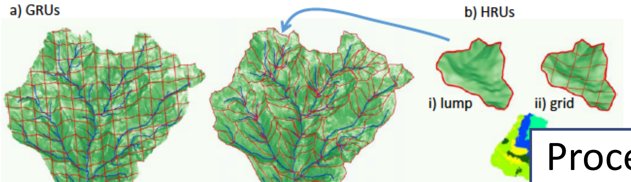
Advances in hydrologic modeling

Ensemble forcing (GMET)

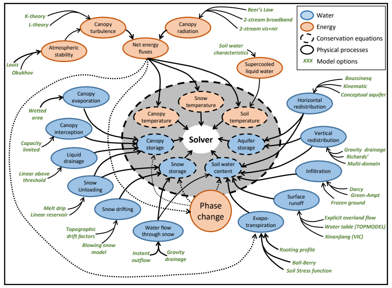


Clark and Slater, 2006; Newman et al., 2015

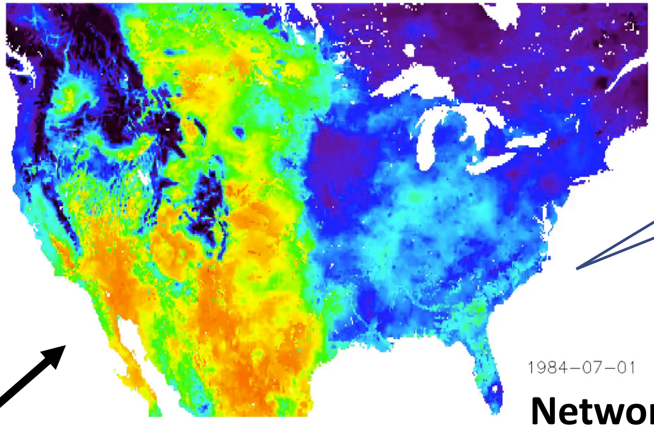
Spatial flexibility (lateral flow)



Process flexibility



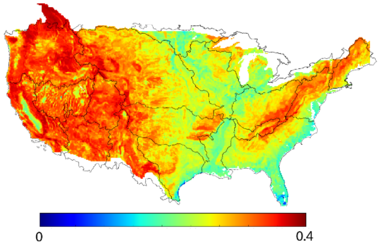
Land Modeling (SUMMA)



1984–07–01

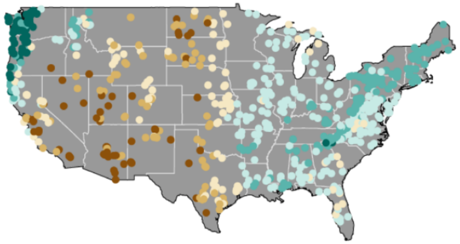
Clark et al., 2015a; 2015b

Large-domain parameter estimation



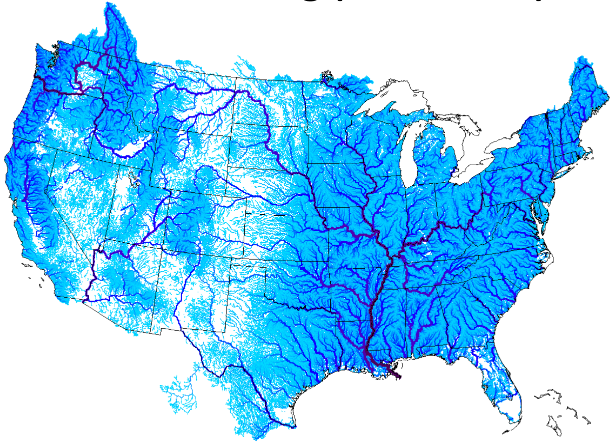
Mizukami et al., 2017

Model benchmarking



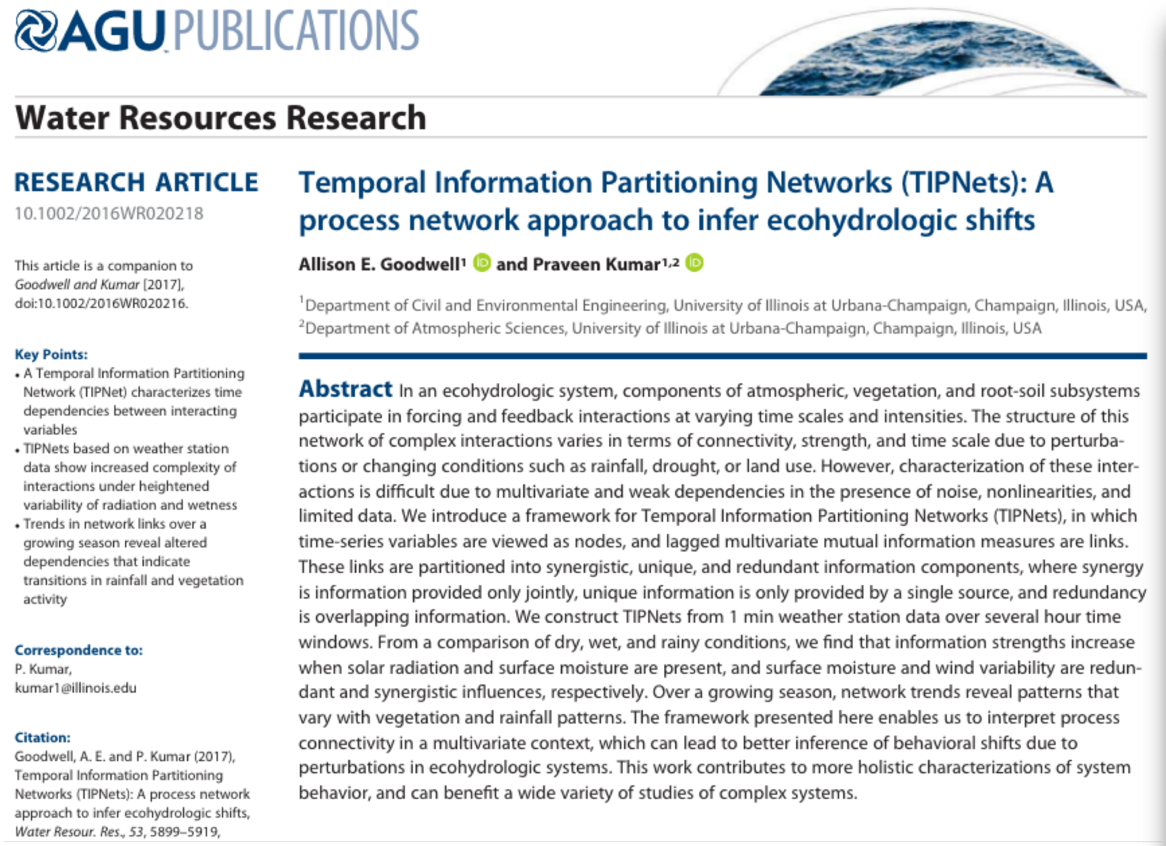
Addor et al., 2017; Newman et al., 2017

Network Routing (mizuRoute)



Clark et al., 2008; Mizukami et al., 2016

Information theory



AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR020218

This article is a companion to *Goodwell and Kumar* [2017], doi:10.1002/2016WR020216.

Key Points:

- A Temporal Information Partitioning Network (TIPNet) characterizes time dependencies between interacting variables
- TIPNets based on weather station data show increased complexity of interactions under heightened variability of radiation and wetness
- Trends in network links over a growing season reveal altered dependencies that indicate transitions in rainfall and vegetation activity

Correspondence to:
P. Kumar,
kumar1@illinois.edu

Citation:
Goodwell, A. E. and P. Kumar (2017), Temporal Information Partitioning Networks (TIPNets): A process network approach to infer ecohydrologic shifts, *Water Resour. Res.*, *53*, 5899–5919,

Temporal Information Partitioning Networks (TIPNets): A process network approach to infer ecohydrologic shifts

Allison E. Goodwell[1] and **Praveen Kumar**[1,2]

[1]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA,
[2]Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

Abstract In an ecohydrologic system, components of atmospheric, vegetation, and root-soil subsystems participate in forcing and feedback interactions at varying time scales and intensities. The structure of this network of complex interactions varies in terms of connectivity, strength, and time scale due to perturbations or changing conditions such as rainfall, drought, or land use. However, characterization of these interactions is difficult due to multivariate and weak dependencies in the presence of noise, nonlinearities, and limited data. We introduce a framework for Temporal Information Partitioning Networks (TIPNets), in which time-series variables are viewed as nodes, and lagged multivariate mutual information measures are links. These links are partitioned into synergistic, unique, and redundant information components, where synergy is information provided only jointly, unique information is only provided by a single source, and redundancy is overlapping information. We construct TIPNets from 1 min weather station data over several hour time windows. From a comparison of dry, wet, and rainy conditions, we find that information strengths increase when solar radiation and surface moisture are present, and surface moisture and wind variability are redundant and synergistic influences, respectively. Over a growing season, network trends reveal patterns that vary with vegetation and rainfall patterns. The framework presented here enables us to interpret process connectivity in a multivariate context, which can lead to better inference of behavioral shifts due to perturbations in ecohydrologic systems. This work contributes to more holistic characterizations of system behavior, and can benefit a wide variety of studies of complex systems.

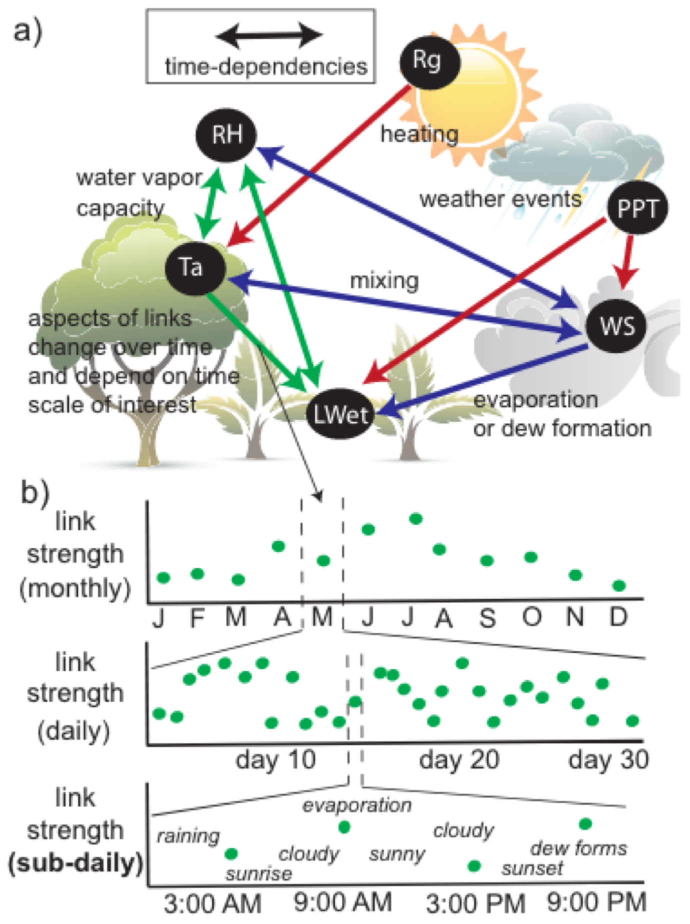
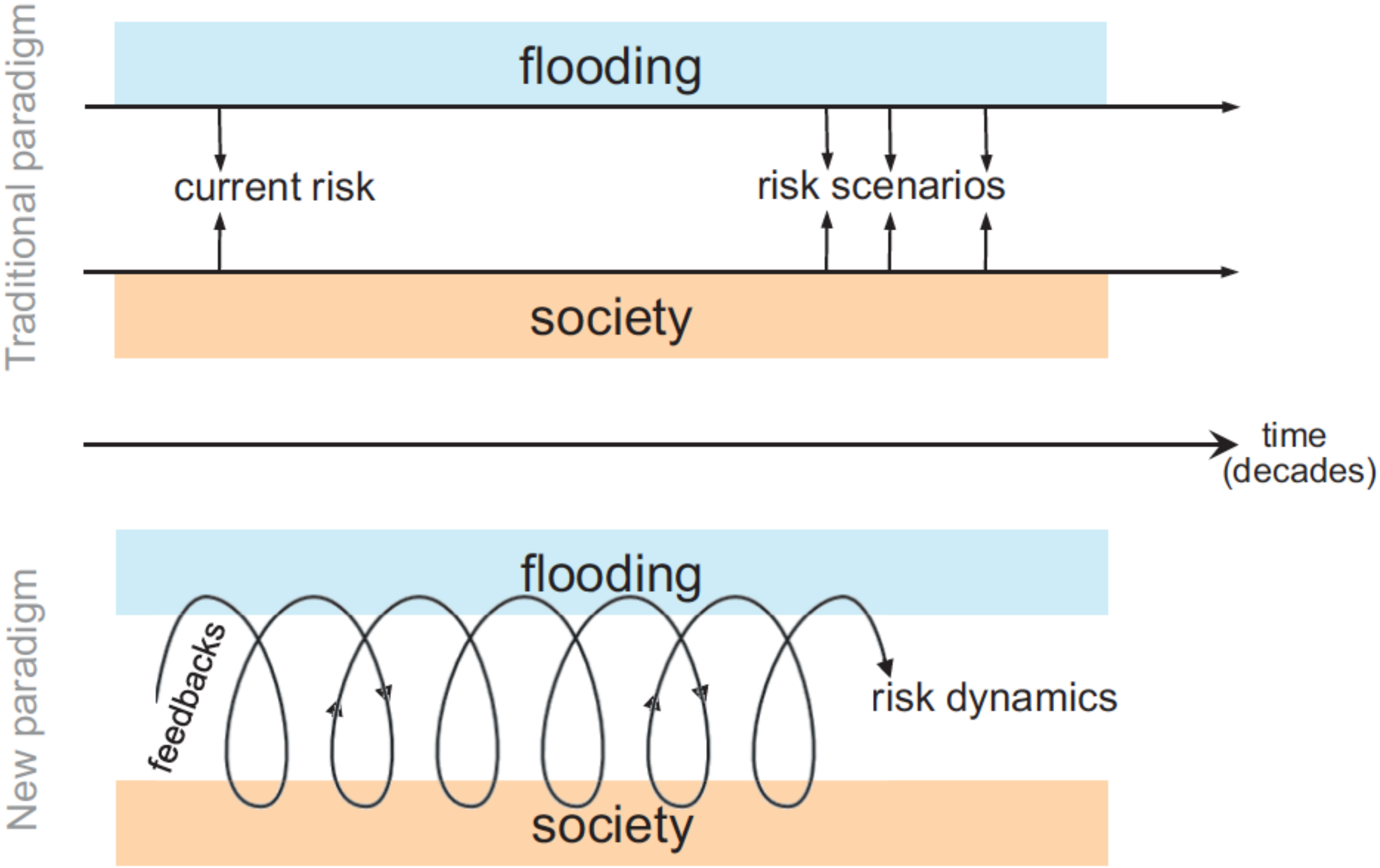


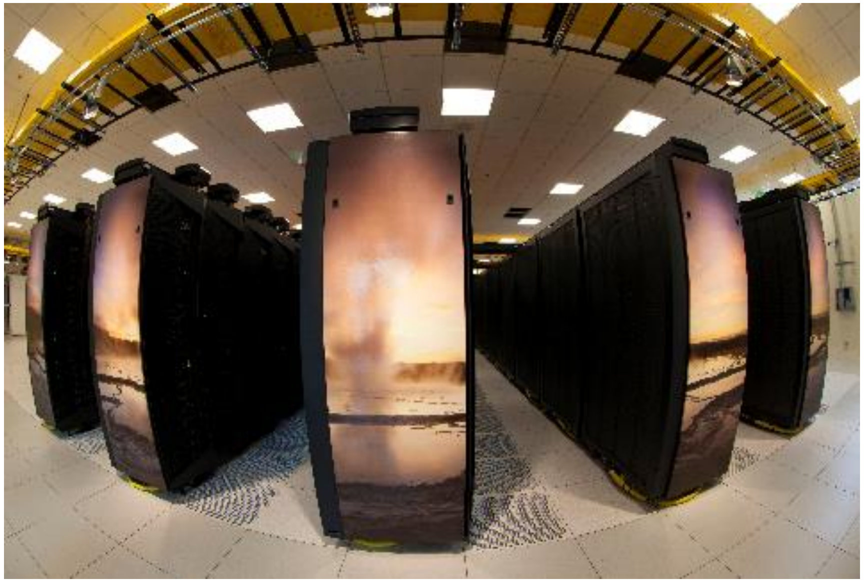
Figure 1. Illustration of complex network behavior in an ecohydrologic system. (a) Time-dependent interactions occur between solar radiation (*Rg*), precipitation (*PPT*), leaf wetness (*LWet*) or moisture condition, wind speed (*WS*), relative humidity (*RH*), and air temperature (*Ta*). We characterize these dependencies as information transfers within a network that are associated with properties of time scale, strength, uniqueness, redundancy, and synergy. (b) Network properties that may be detected on a seasonal time scale (top) result from an accumulation of interactions that vary on much shorter time scales such as daily or subdaily (bottom).

Coupled human-hydrology interactions



Di Baldassarre et al., WRR 2015

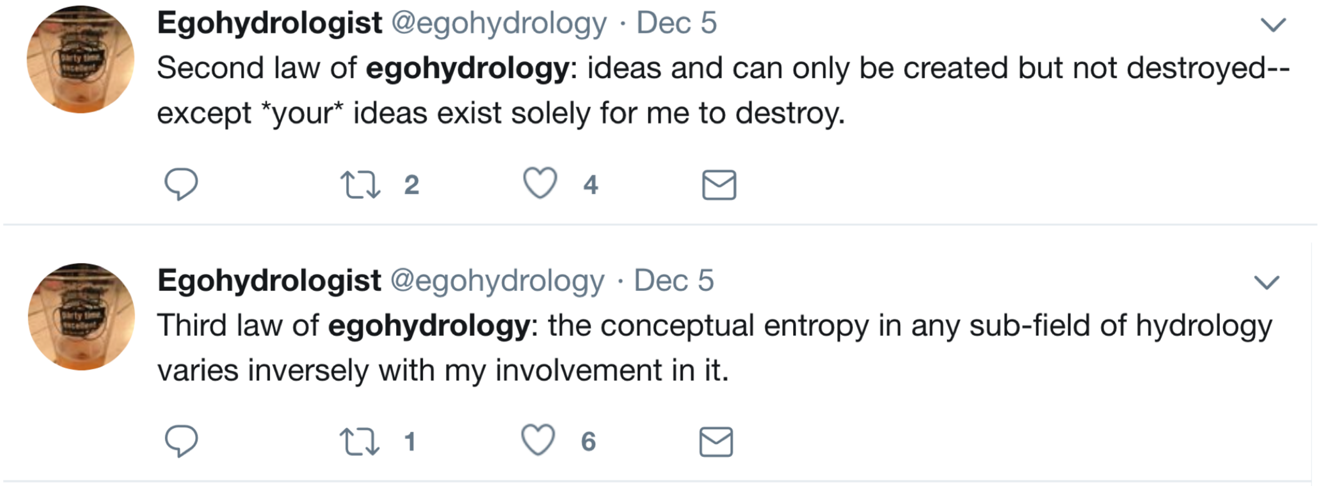
Technological advances



High performance computing

Open data / open models

Findable
Accessible
Make Data FAIR
AGU
Interoperable
Reusable



Egohydrologist @egohydrology · Dec 5
Second law of **egohydrology**: ideas and can only be created but not destroyed--except *your* ideas exist solely for me to destroy.

Egohydrologist @egohydrology · Dec 5
Third law of **egohydrology**: the conceptual entropy in any sub-field of hydrology varies inversely with my involvement in it.

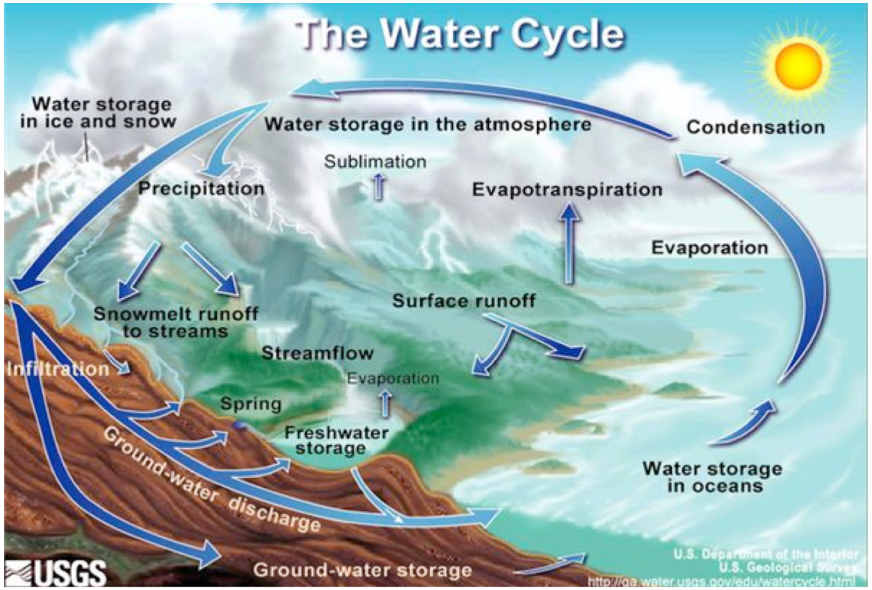
Social media…

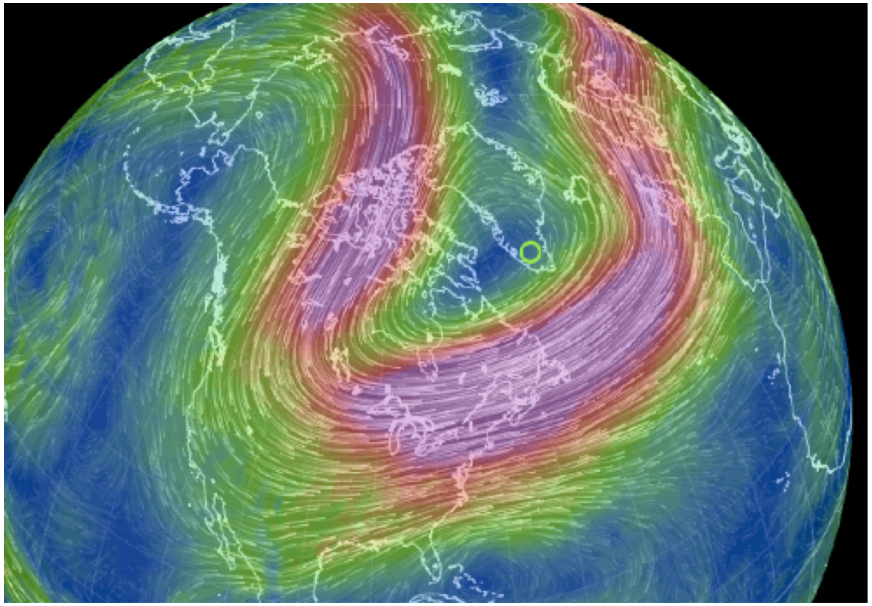
Outline

- Background
  - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The emergence of CTSM
- Summary and research needs

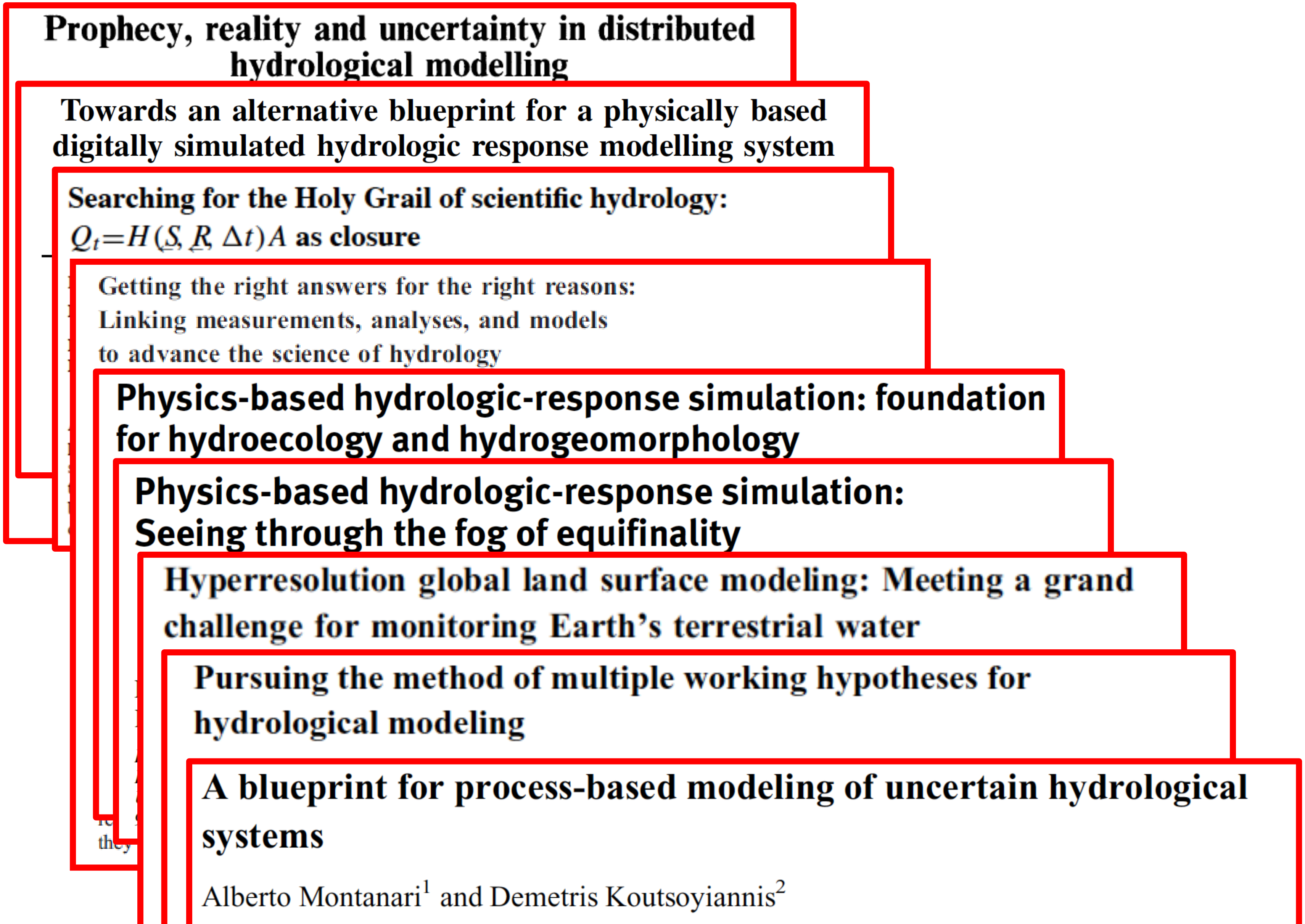
Hydrologic vs. atmospheric modeling

- 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…



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:
$Q_t = H(\underset{\sim}{S}, \underset{\sim}{R}, \Delta t)A$ 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

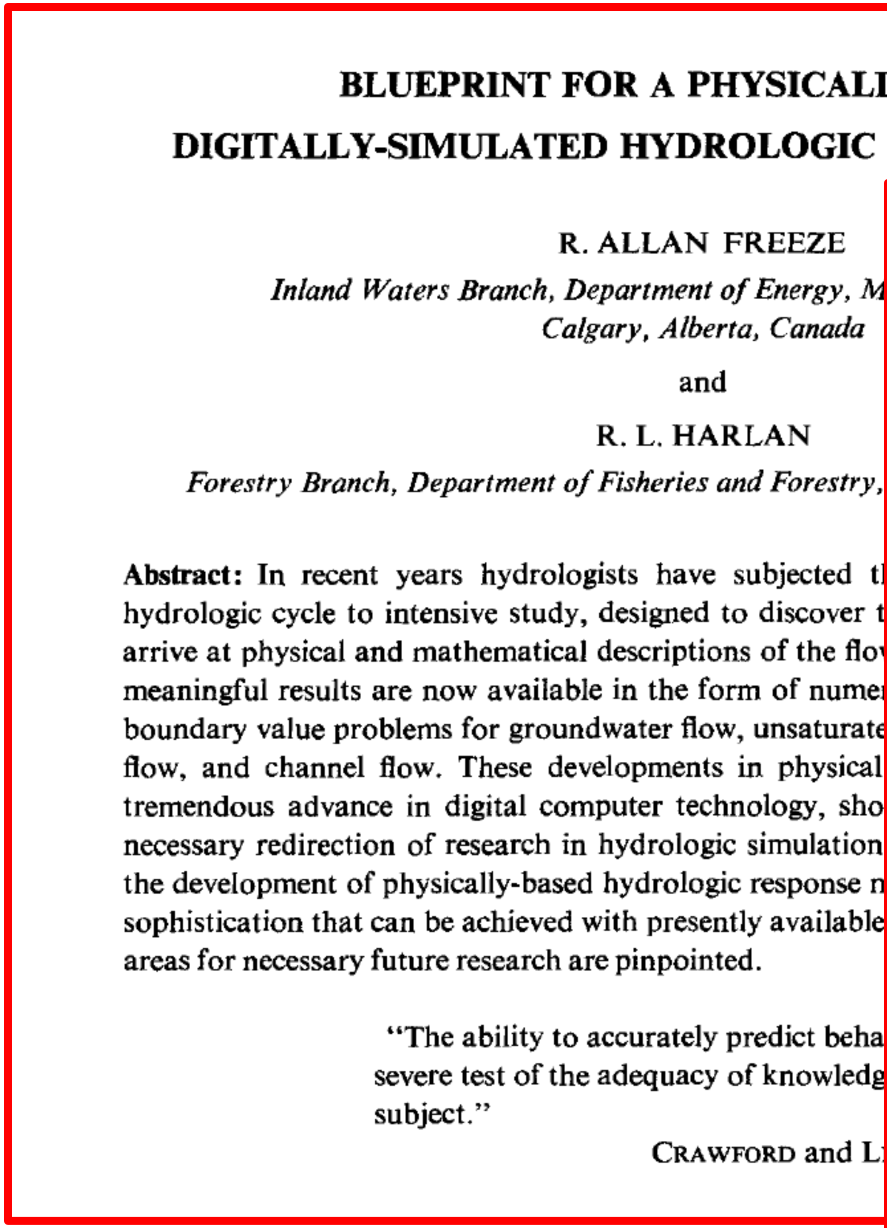
A blueprint for process-based modeling of uncertain hydrological systems

Alberto Montanari[1] and Demetris Koutsoyiannis[2]

Beyond "faith-based modeling"?

- 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*

The Freeze and Harlan blueprint



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 m
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

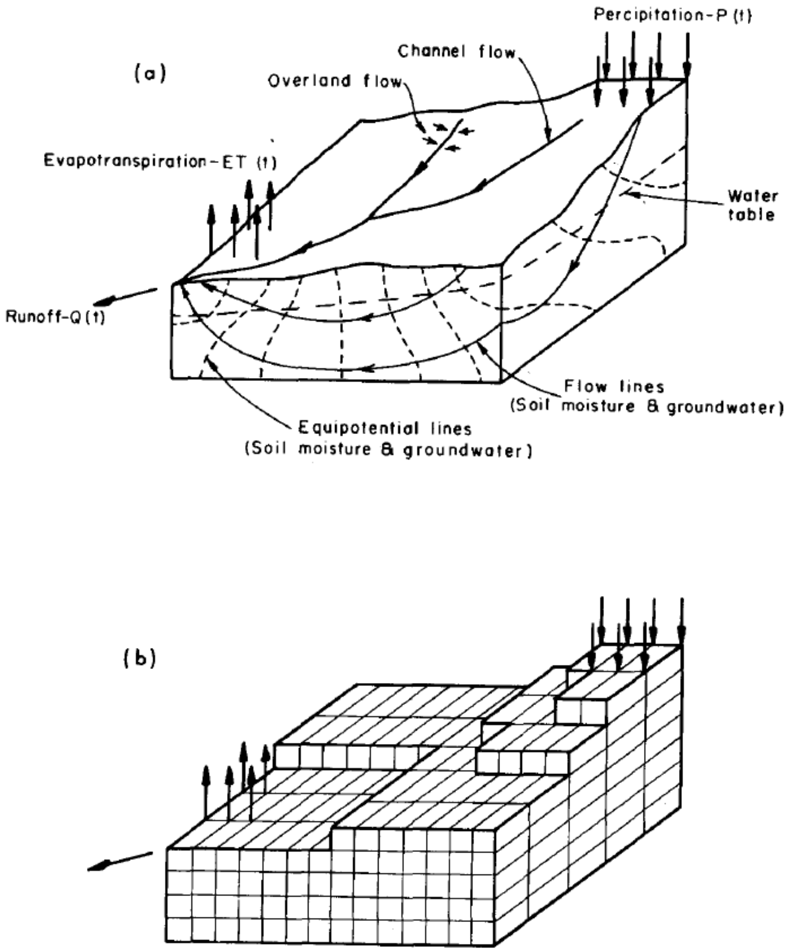


Fig. 3. Schematic diagram of (a) Hydrologic basin and (b) Three dimensional nodal model of hydrologic basin.

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?

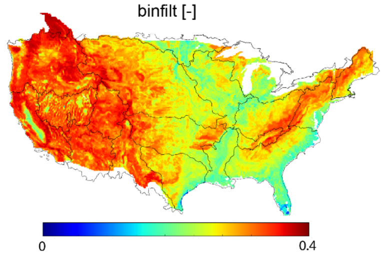
- Have the earlier computer limitations of storage capacity and speed of computation been overcome? Is the application of digital computers to this type of problem economically feasible?

Key challenges

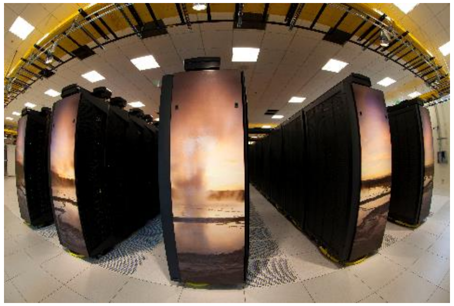
- The choice of modeling approaches (arguably) stems from personal preferences for physics or parsimony
- Need a stronger scientific basis for model development/improvement
    - Treat numerical modeling as a subjective decision-making process – *carefully evaluate all modeling decisions in a controlled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



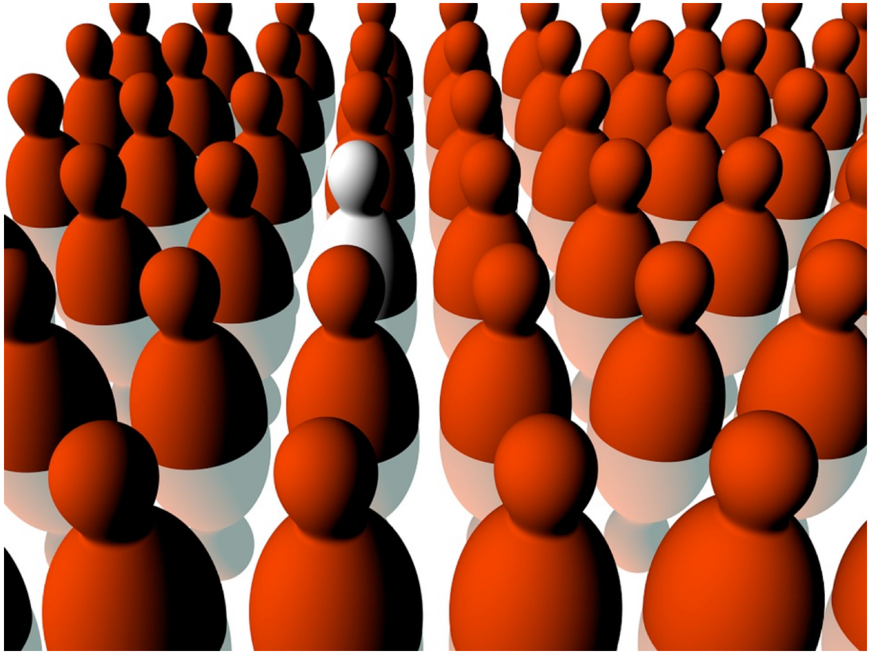
Outline

- Background
  - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
  → Processes
  - Parameters
  - Computing
- The emergence of CTSM
- Summary and research needs

Two issues:

Model proliferation and the shantytown syndrome

- **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)***

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

→

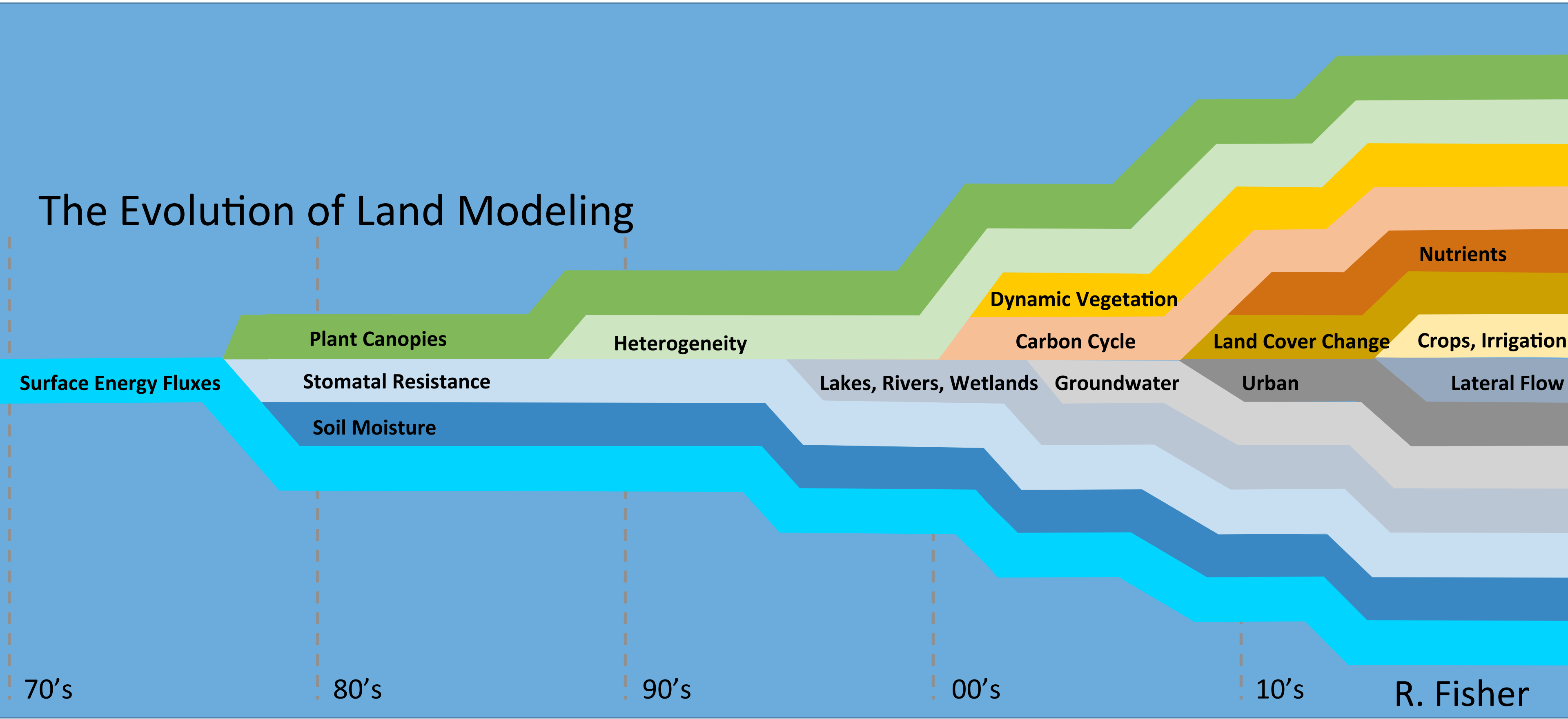
Mechanistic modeling of land processes

Properties define processes (focus on short-term fluxes)

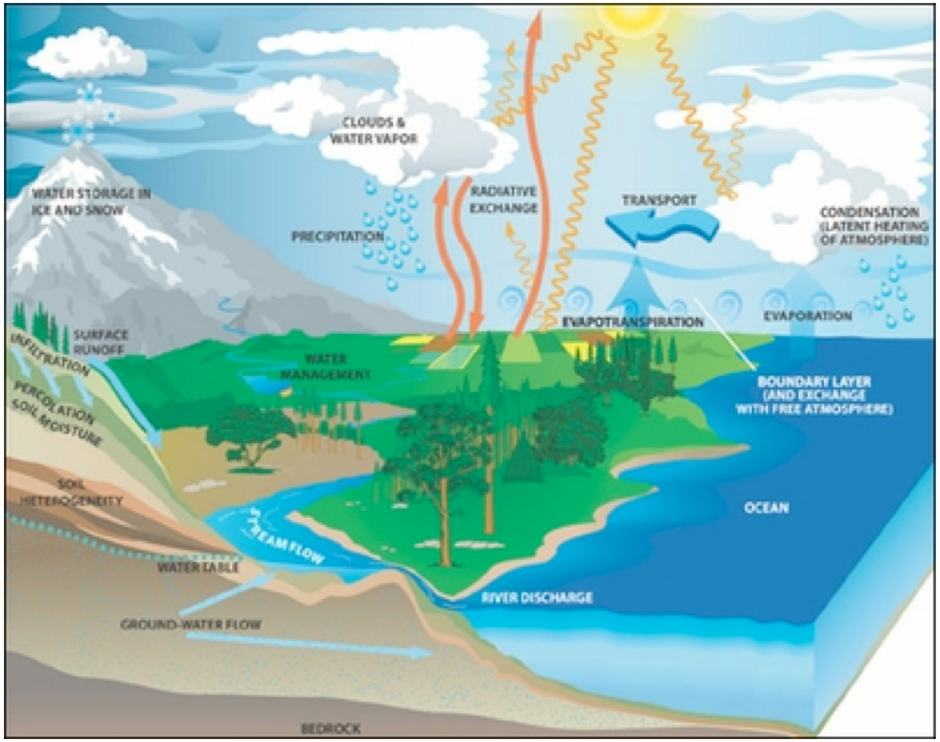
→

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

Processes define properties (feedbacks and interactions across time scales)



Unifying models



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

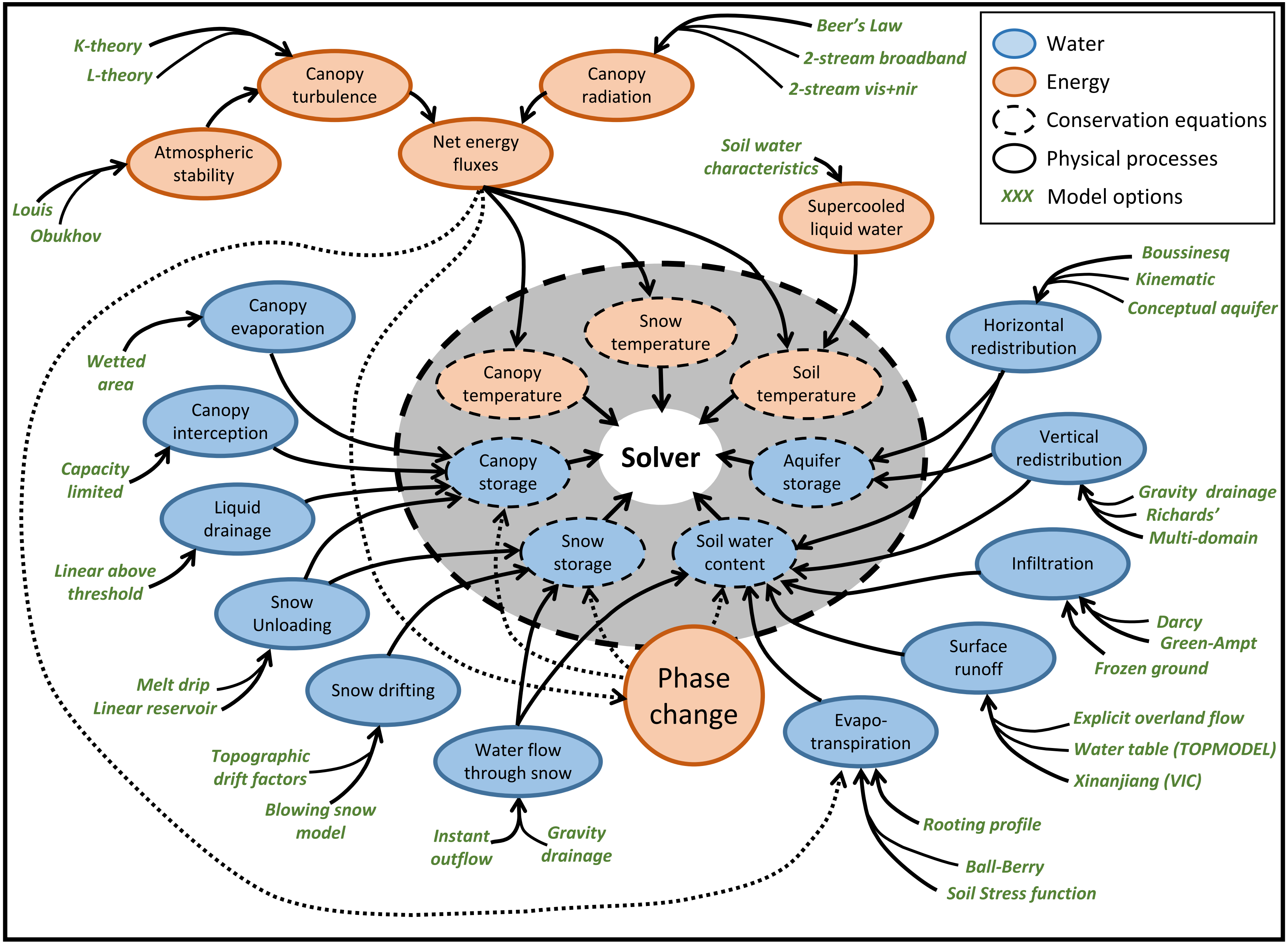
2. Differences among models relate to
   a) the spatial discretization of the model domain;
   b) the approaches used to parameterize individual fluxes (including model parameter values); and
   c) the methods used to solve the governing model equations.

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)

Unifying process representations



- ○ Water
- ○ Energy
- ◌ Conservation equations
- ○ Physical processes
- *XXX* Model options

Solver

Canopy temperature
Snow temperature
Soil temperature
Canopy storage
Aquifer storage
Snow storage
Soil water content

Canopy turbulence
K-theory
L-theory

Atmospheric stability
Louis
Obukhov

Canopy radiation
Beer's Law
2-stream broadband
2-stream vis+nir

Net energy fluxes

Supercooled liquid water
Soil water characteristics

Canopy evaporation
Wetted area

Canopy interception
Capacity limited

Liquid drainage
Linear above threshold

Snow Unloading
Melt drip
Linear reservoir

Snow drifting
Topographic drift factors
Blowing snow model

Water flow through snow
Instant outflow
Gravity drainage

Phase change

Evapo-transpiration
Rooting profile
Ball-Berry
Soil Stress function

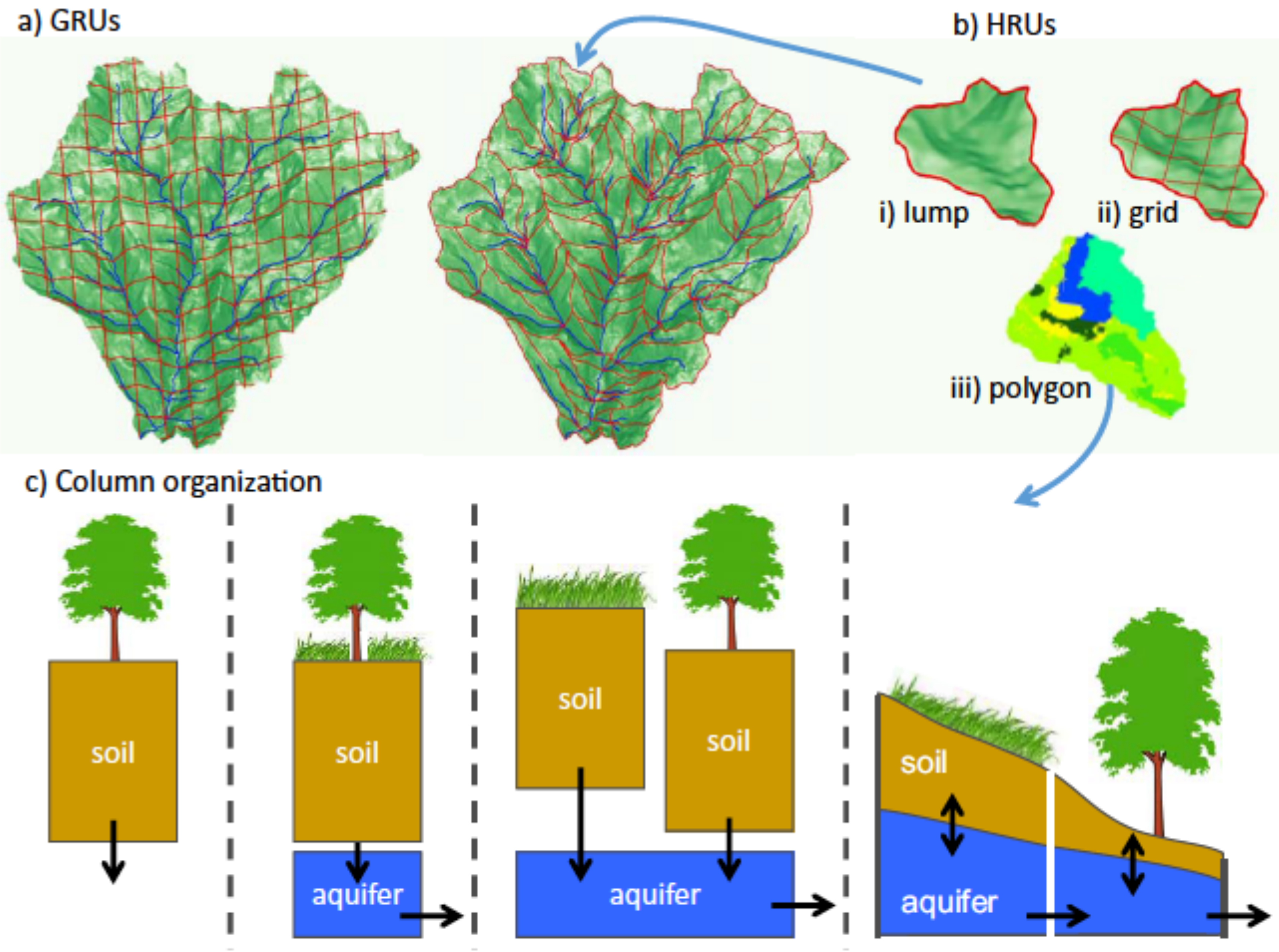
Surface runoff
Explicit overland flow
Water table (TOPMODEL)
Xinanjiang (VIC)

Infiltration
Darcy
Green-Ampt
Frozen ground

Vertical redistribution
Gravity drainage
Richards'
Multi-domain

Horizontal redistribution
Boussinesq
Kinematic
Conceptual aquifer

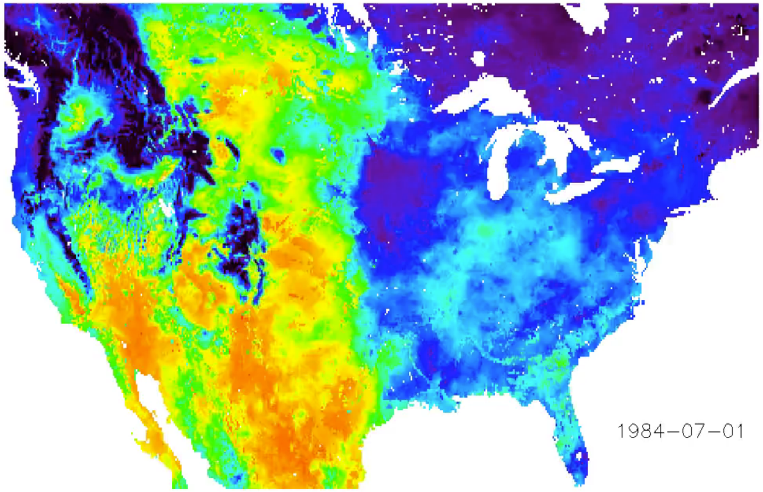
Unifying spatial configurations

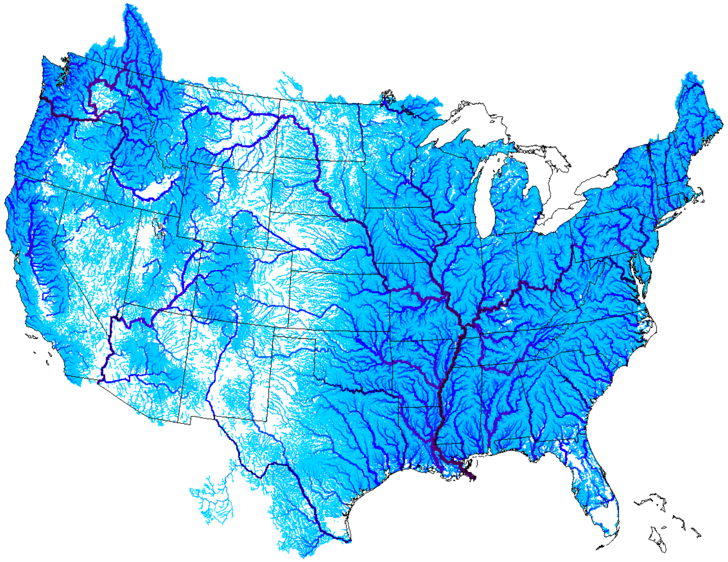


Use cases

- 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: spatial scaling

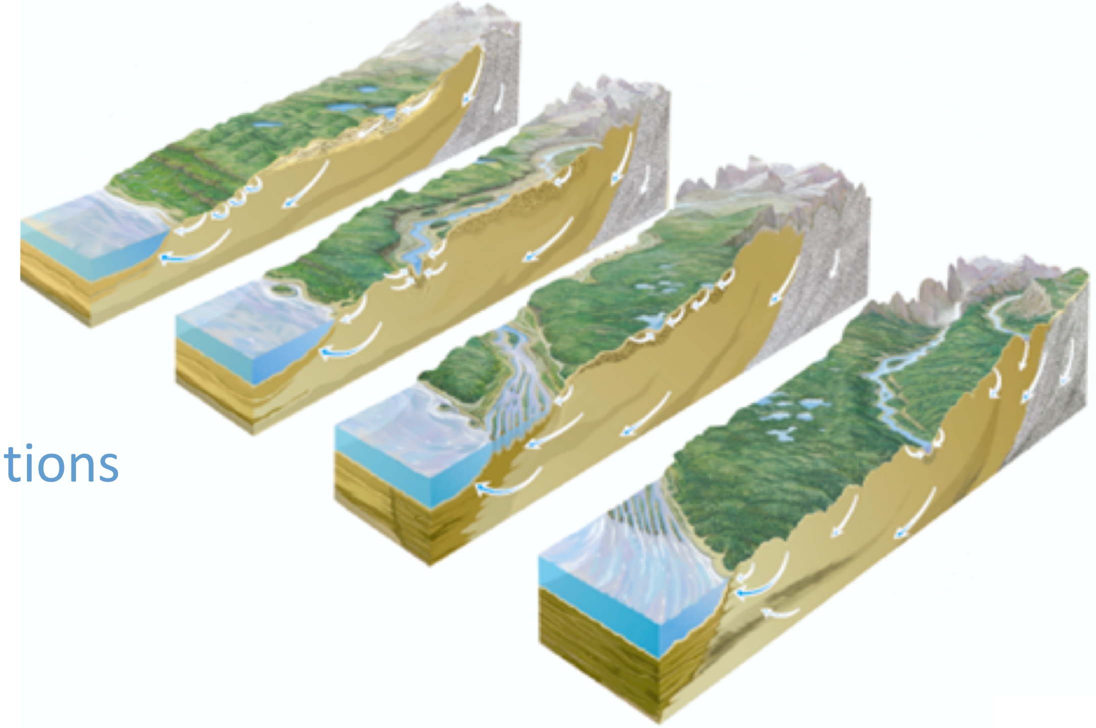
- The spatial gradients that drive flow occur at very small spatial scales and are not resolved by even the finest terrain grid used in large-domain hyper-resolution models

- 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
  - …

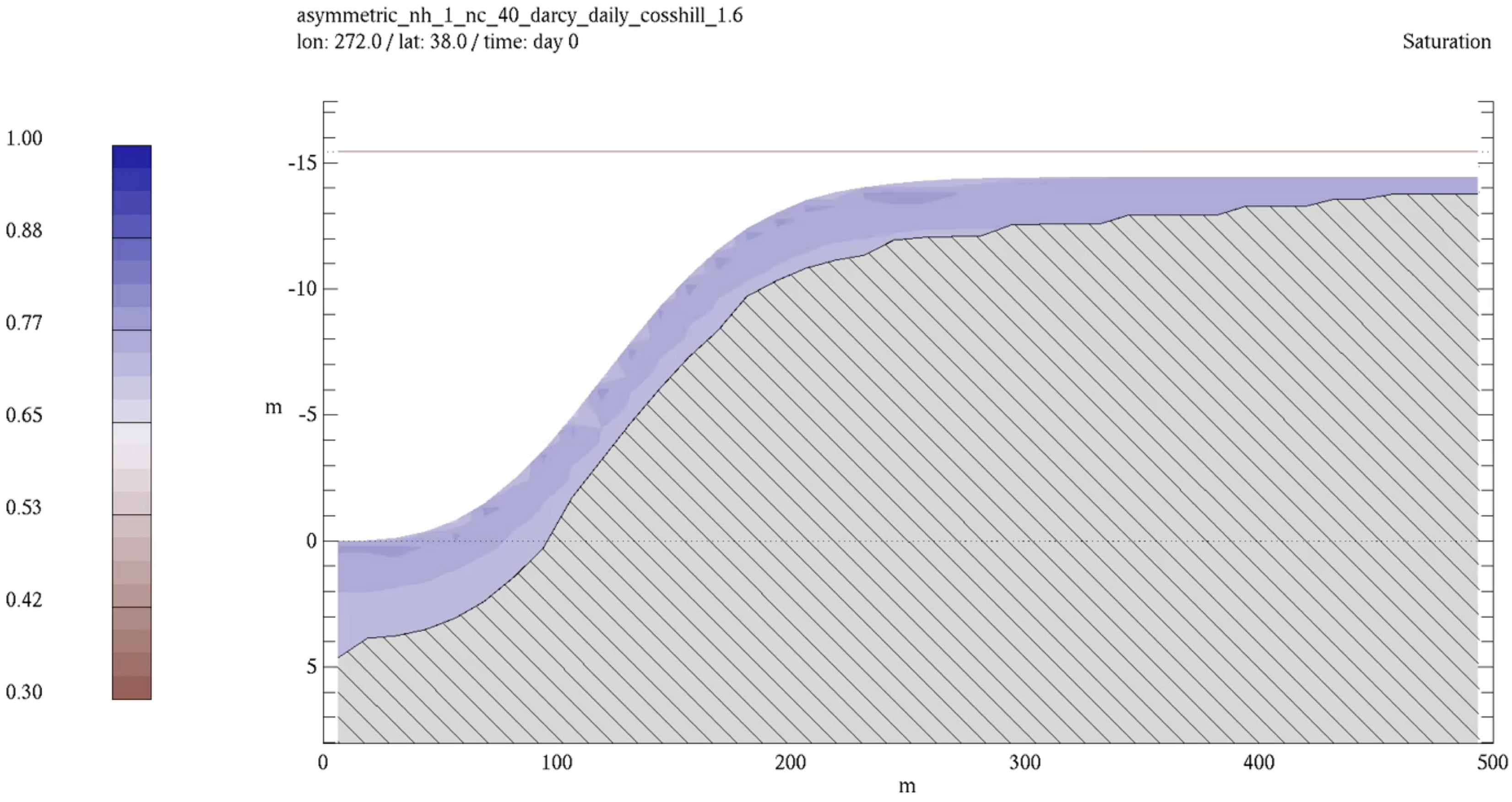
Challenge: spatial scaling

- Modeling challenge: Develop flux parameterizations that represent the aggregate impact of sub-grid-scale heterogeneities.



- Grid-average fluxes
  - Upscaled parameter values
  - New flux parameterizations
  - Sub-grid probability distributions
  - More…
- Spatial discretization
  - Hydrologic similarity
  - Representative hillslopes
  - Separate computations for process subsets
  - More…

Example: Representative hillslopes



Courtesy Sean Swenson (NCAR)

Outline

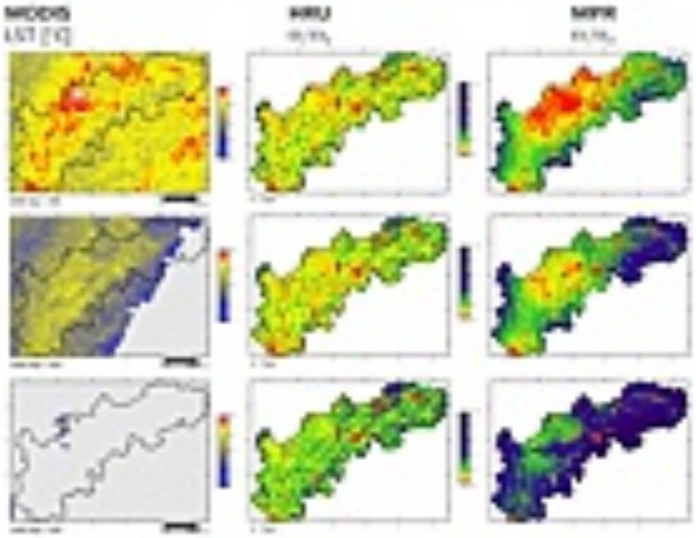
- Background
  - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
  - Processes
  - ➡ Parameters
  - Computing
- The emergence of CTSM
- Summary and research needs

Challenge 2: Model parameters

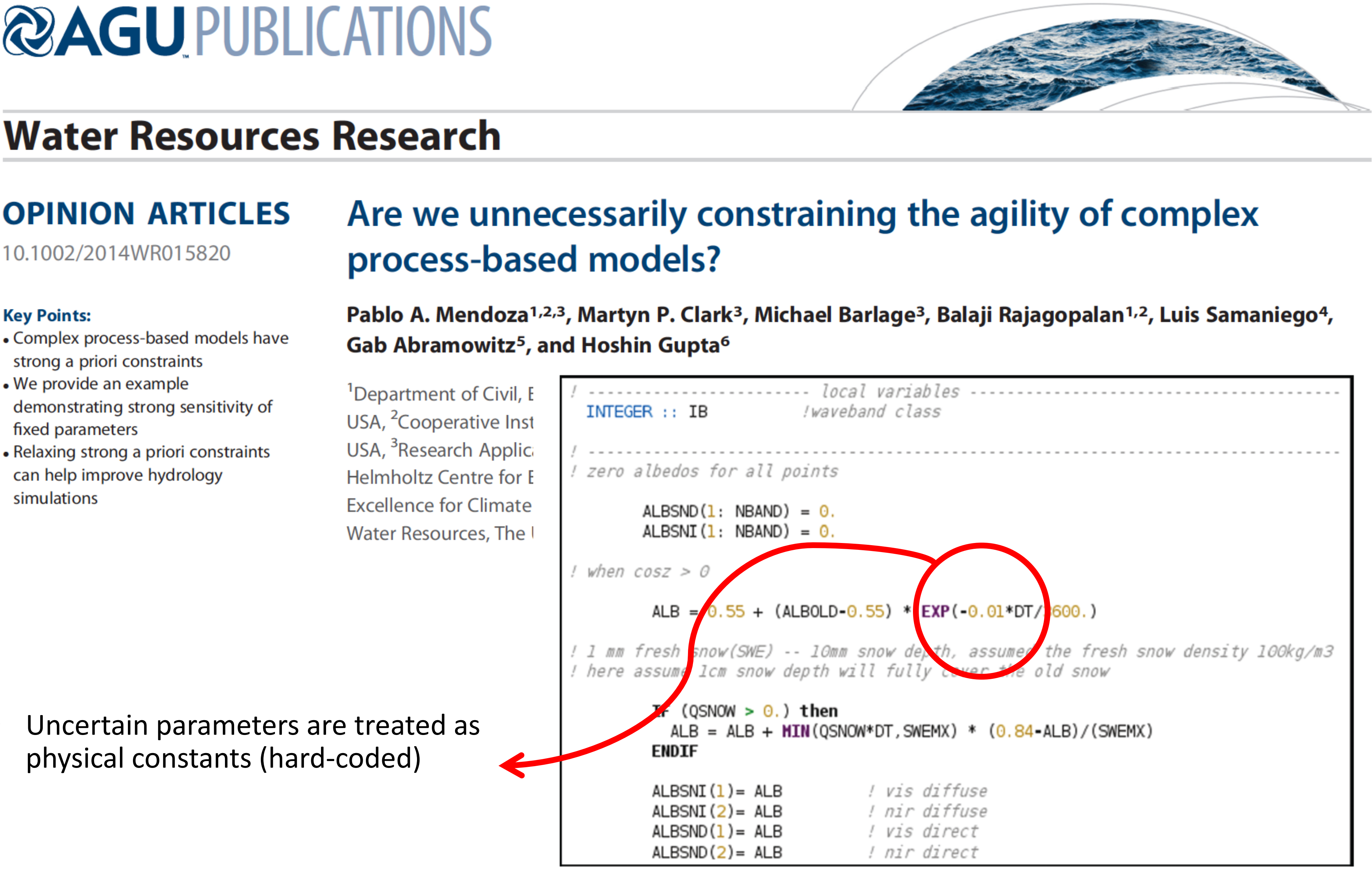
Model parameters

- 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)



Model parameters



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]

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

Key Points:

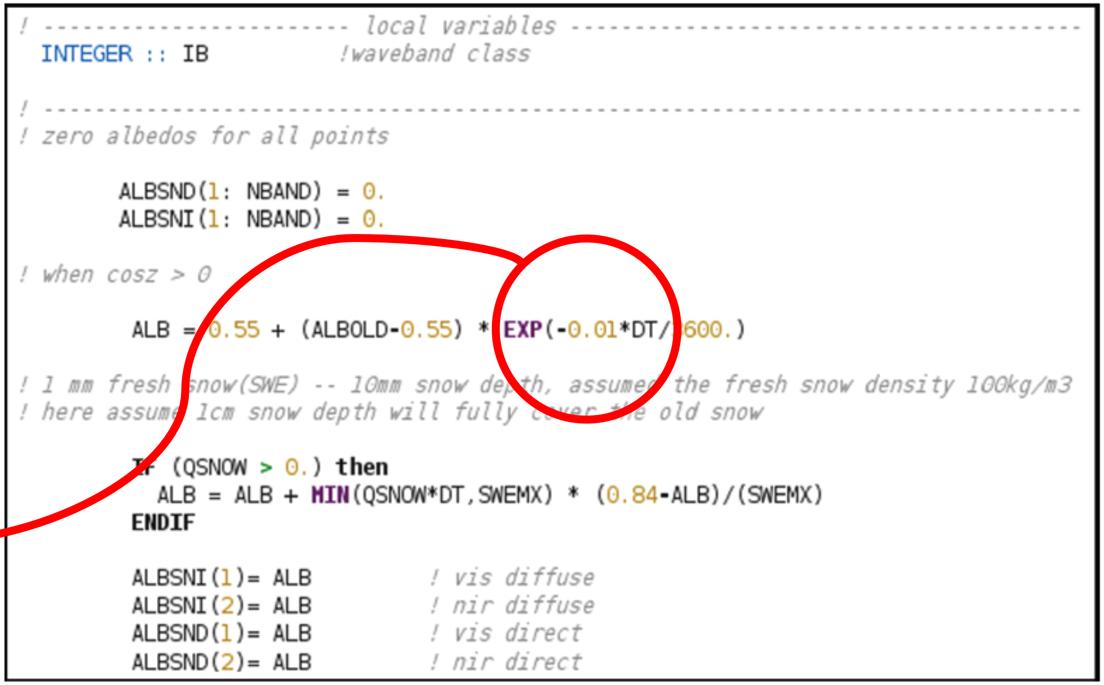
- Complex process-based models have strong a priori constraints
- We provide an example demonstrating strong sensitivity of fixed parameters
- Relaxing strong a priori constraints can help improve hydrology simulations

```
! ------------------------ 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)

Model parameters

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



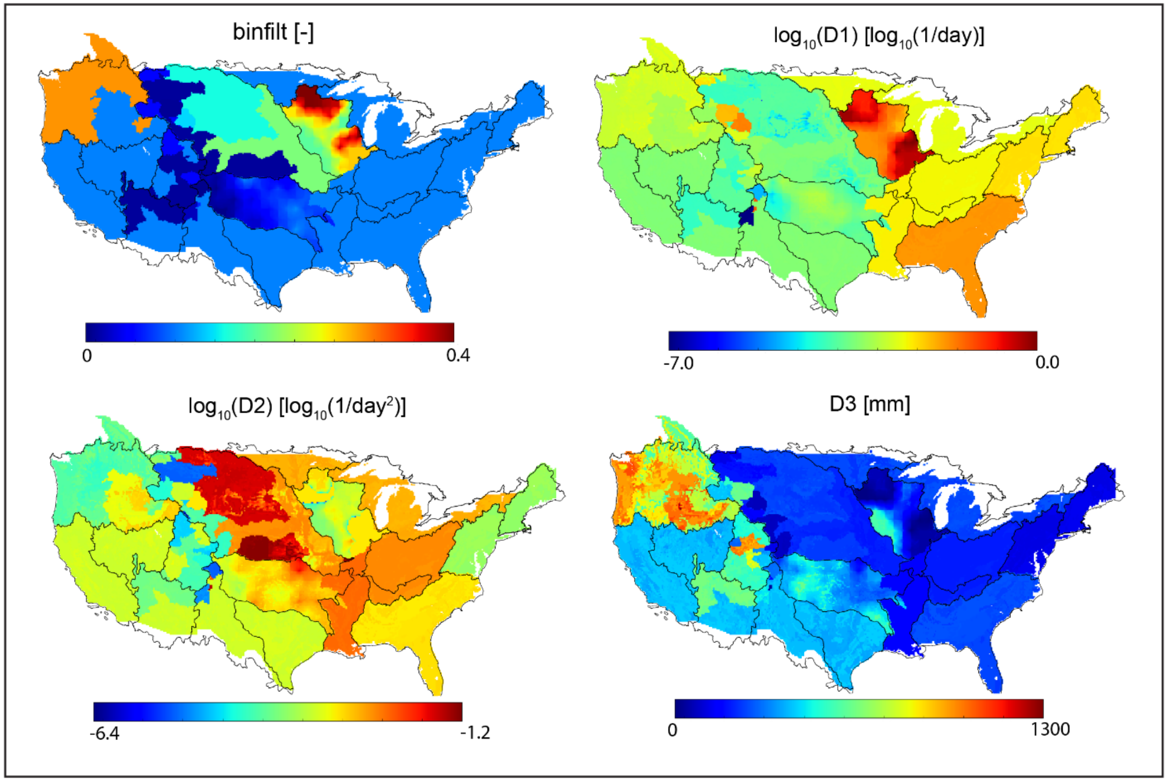
```
! ------------------------ local variables ------------------------------------
 INTEGER :: IB              !waveband class

! ----------------------------------------------------------------------------
! 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
```

Default parameters

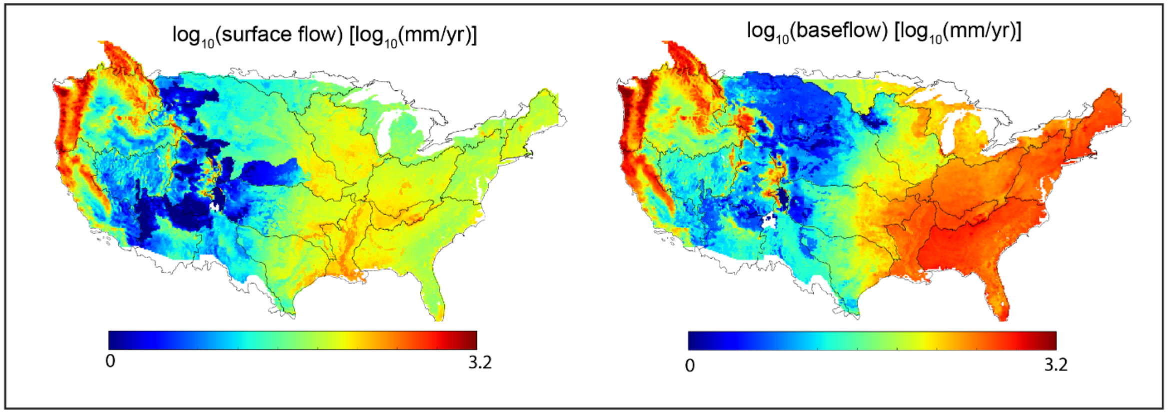
- Spatial discontinuities in model parameters

- Spatial discontinuities in model simulations

VIC Soil parameters – CMIP5 default



1950-1999 annual mean runoff



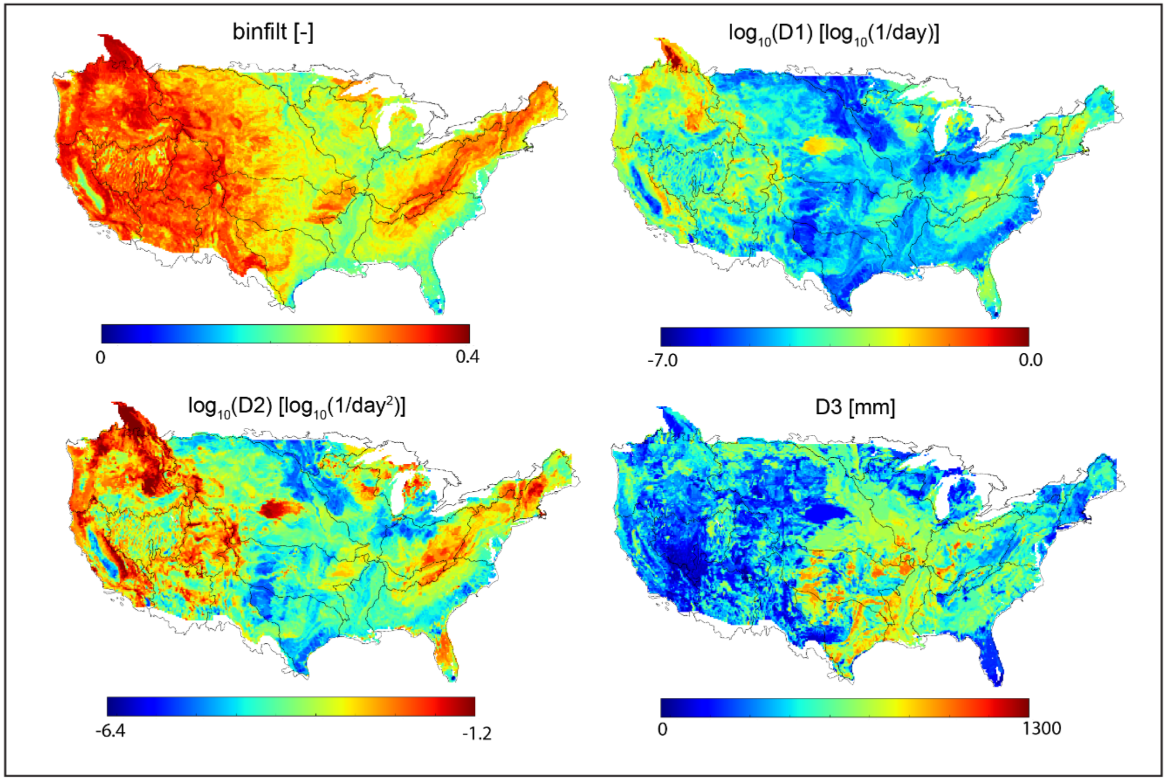
Mizukami et al., WRR 2017

MPR-flex

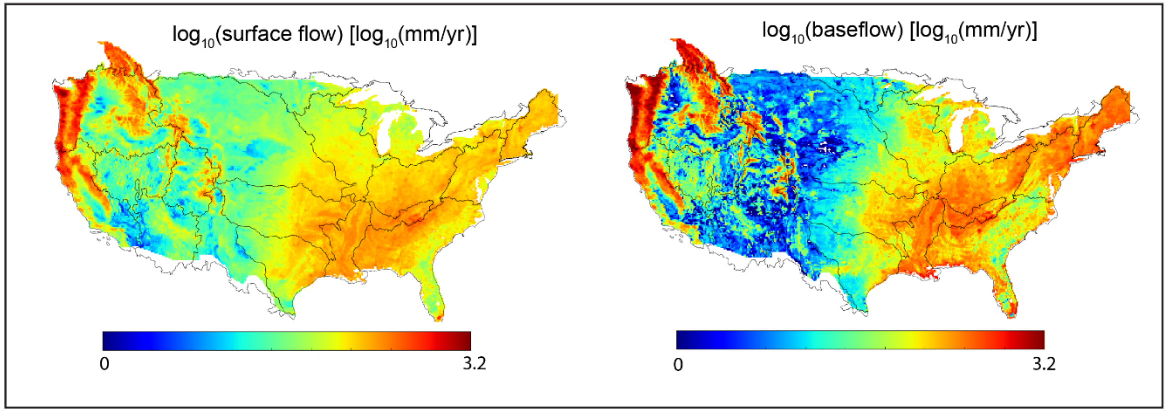
- Modify coefficients in transfer functions that relate physical attributes (soil, veg, topography) to model parameters
- Use parameter-specific upscaling operators to represent multi-scale behavior
- *Define transfer functions for new models – develop model agnostic MPR (MPR-Flex)*

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

VIC Soil parameters – MPR



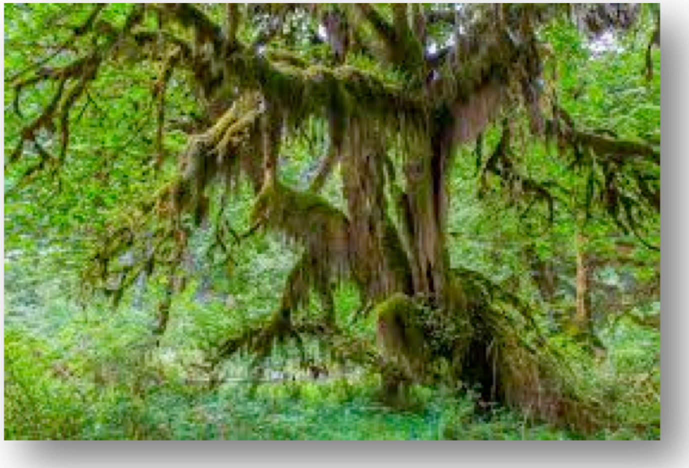
1950-1999 annual mean runoff



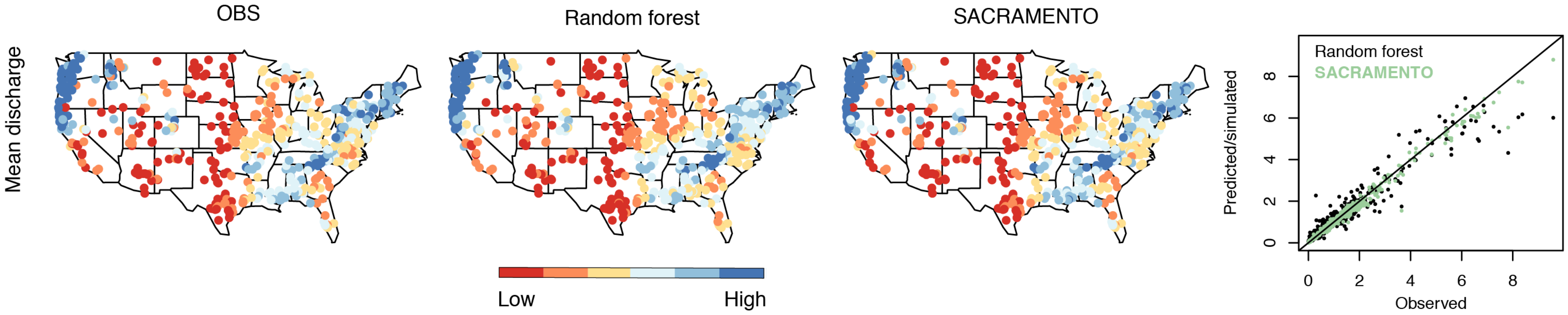
Mizukami et al., WRR 2017

Transferring parameters across space
What are the key challenges?

- Representing landscape (vegetation, soil, climate, topography) in the models
- Which attributes have the most influence on catchment behavior (i.e. on the dominant hydrological processes)?

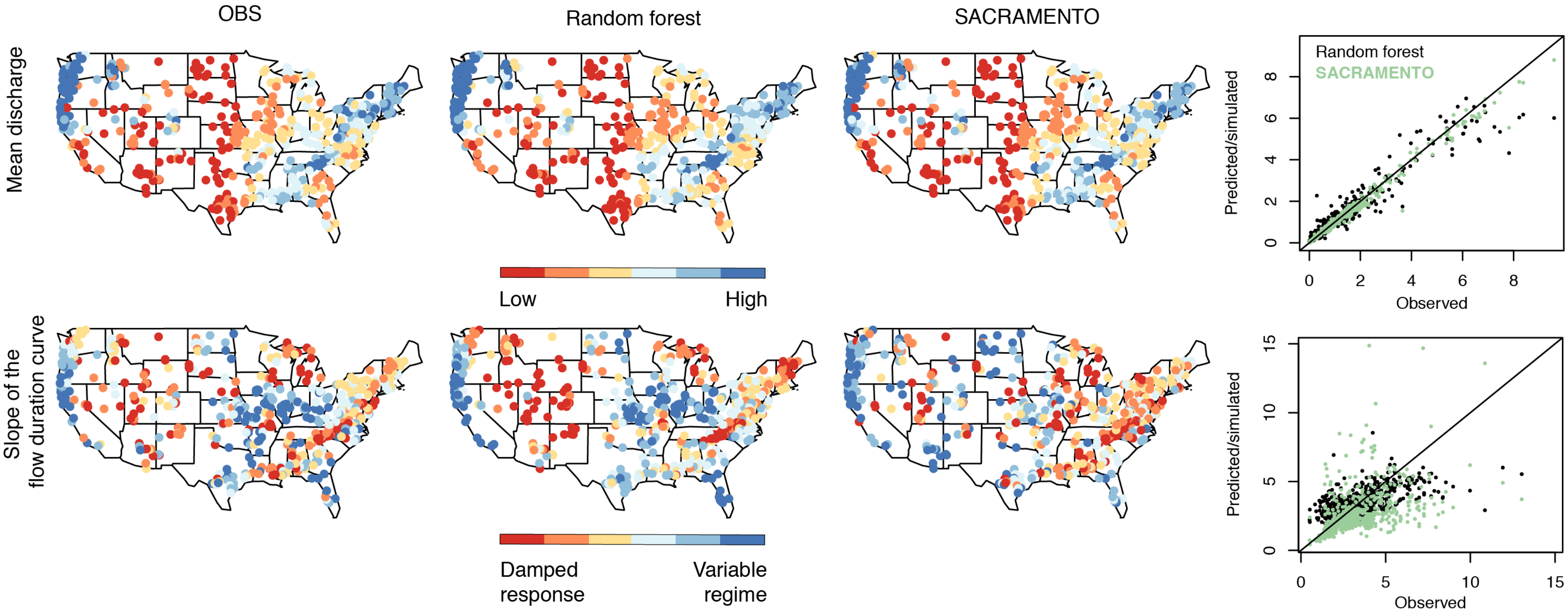


How well can we capture those hydrological signatures?



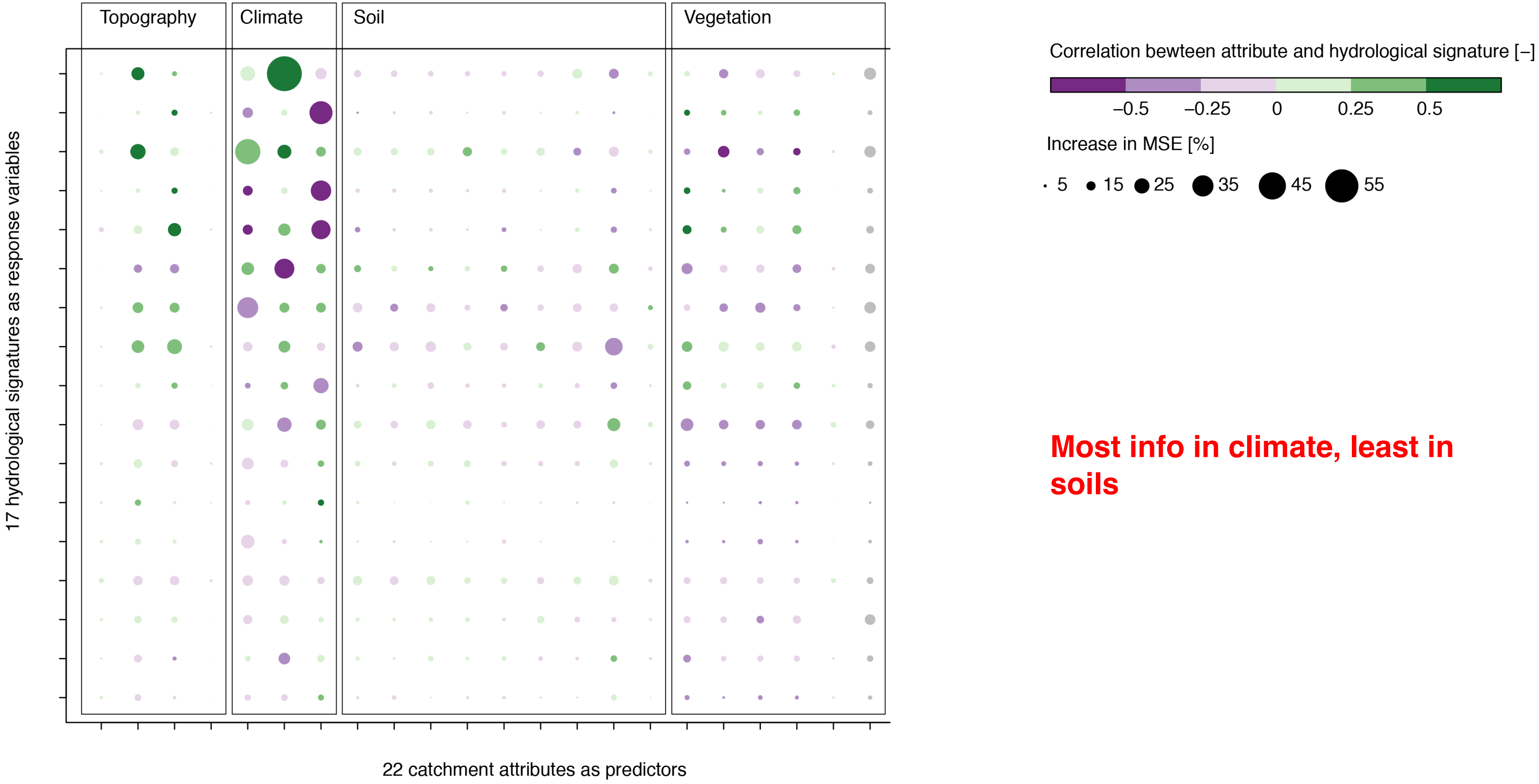
Addor et al., WRR, in press

How well can we capture those hydrological signatures?



Addor et al., WRR, in press

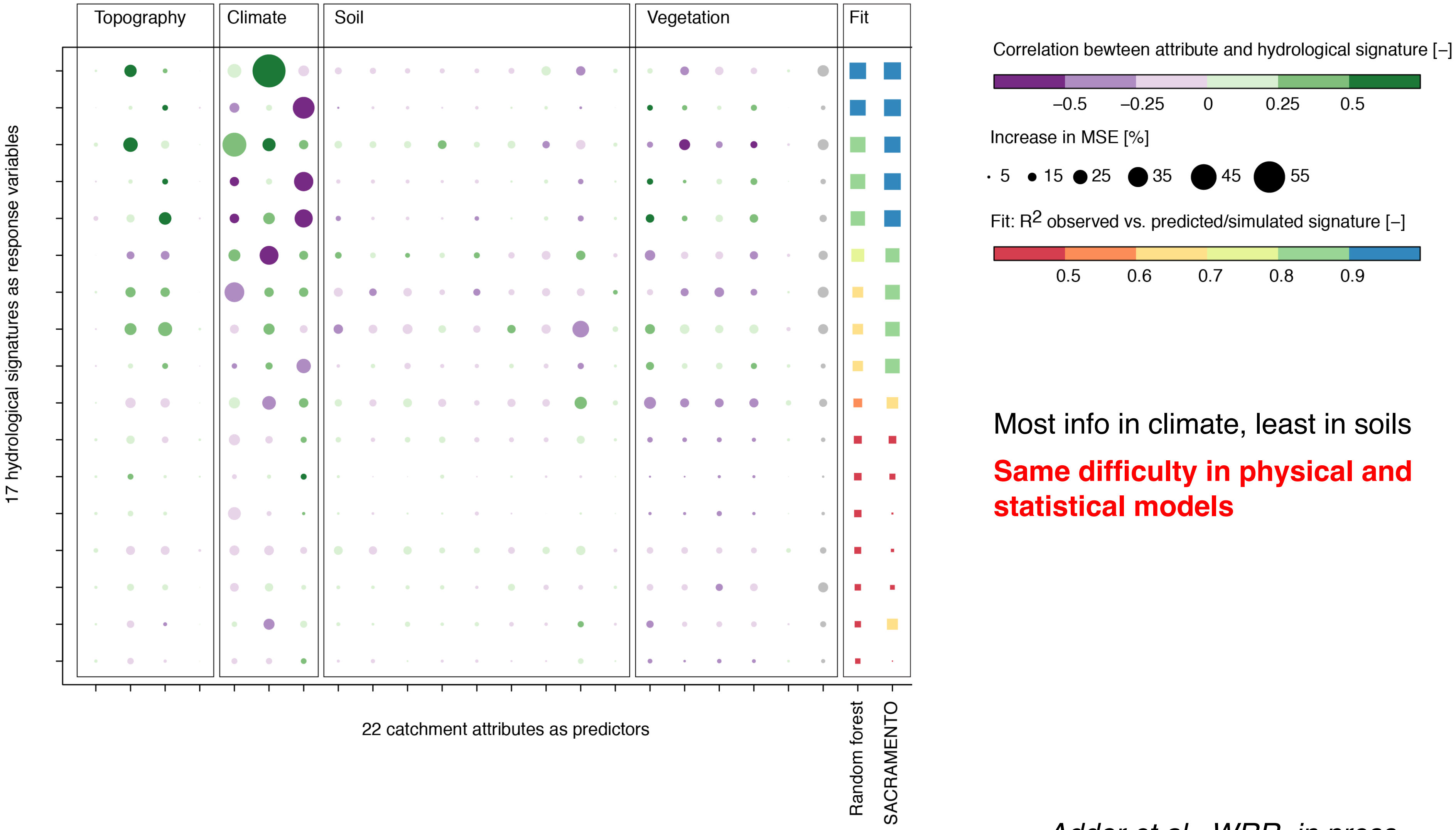
Which catchment attributes are most important?



Most info in climate, least in soils

Addor et al., WRR, in press

How do physical models compare to statistical ones?

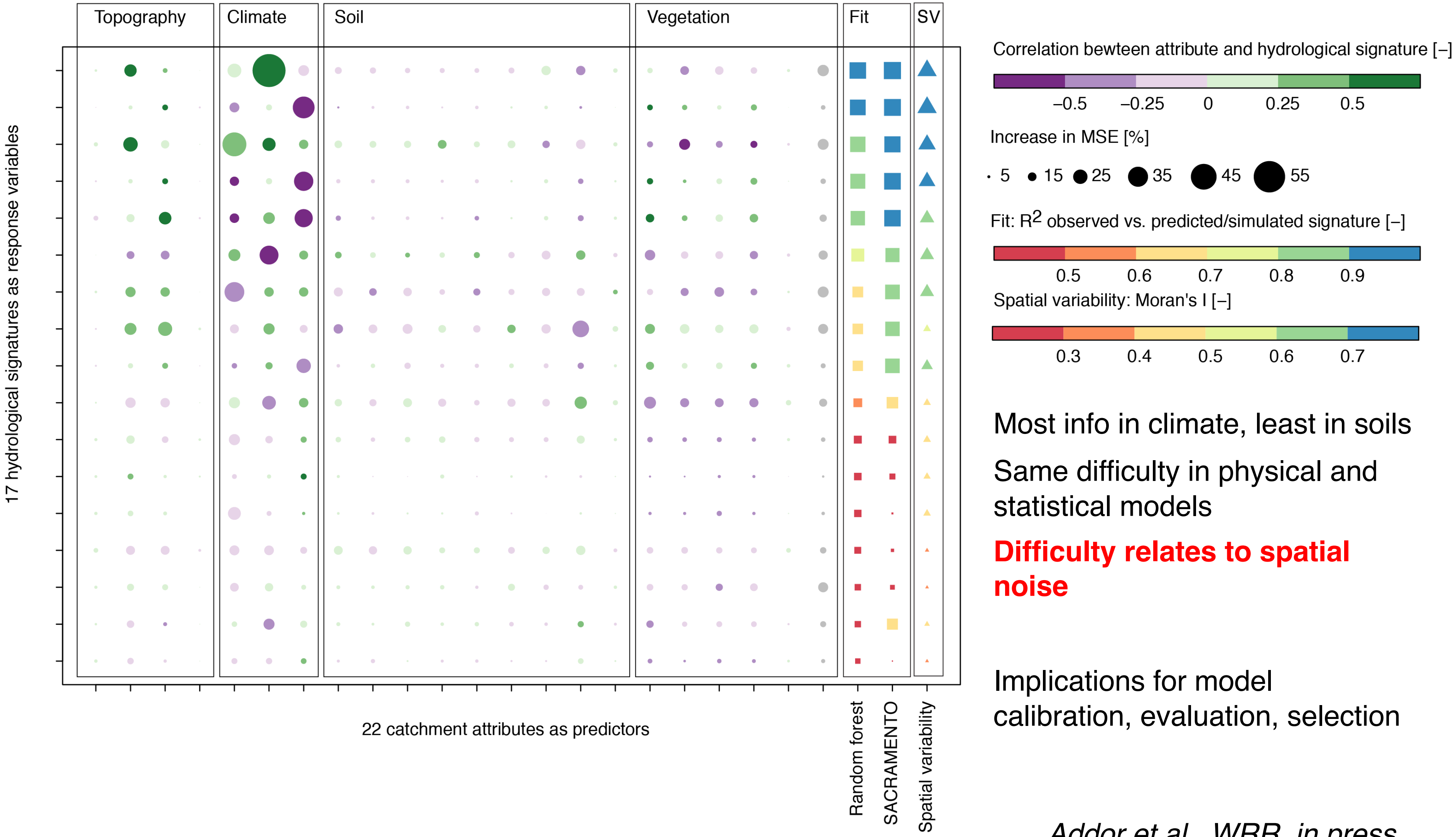


Most info in climate, least in soils

Same difficulty in physical and statistical models

Addor et al., WRR, in press

Why are simulations good/poor?



Most info in climate, least in soils

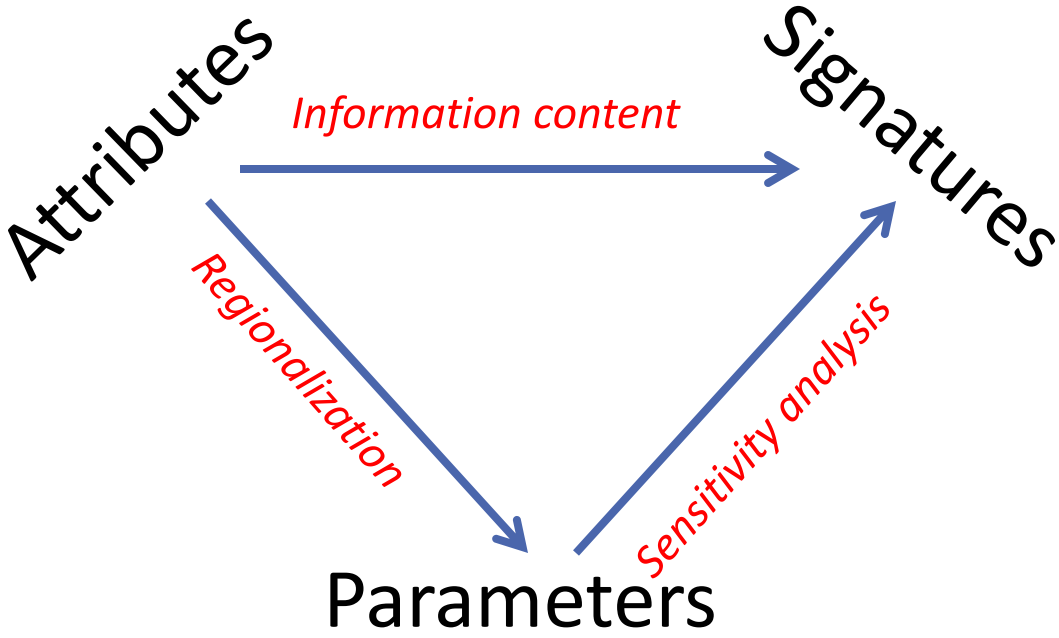
Same difficulty in physical and statistical models

Difficulty relates to spatial noise

Implications for model calibration, evaluation, selection

Addor et al., WRR, in press

Process-based parameter estimation?



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?

Outline

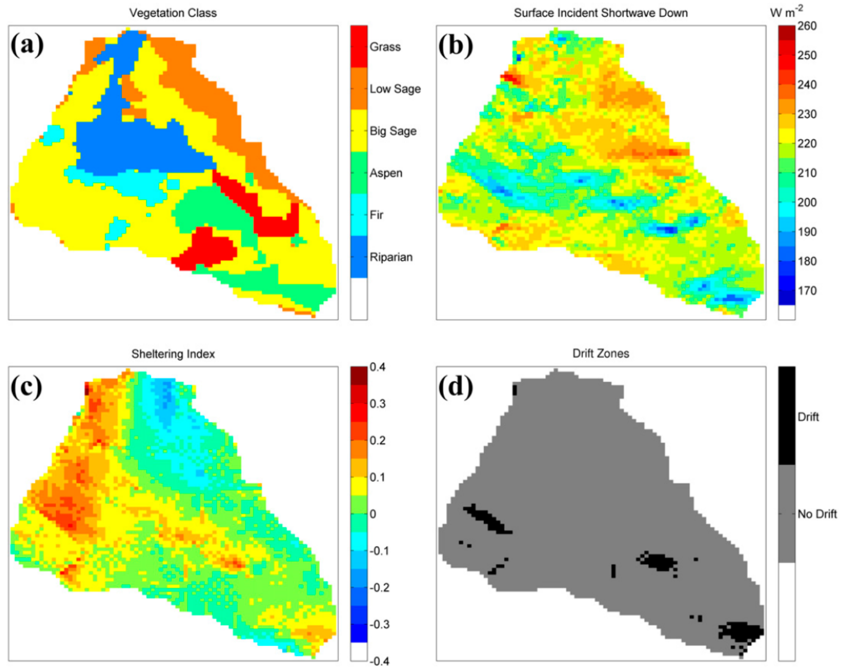
- Background
  - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
  - Processes
  - Parameters
  - ➡ Computing
- The emergence of CTSM
- Summary and research needs

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

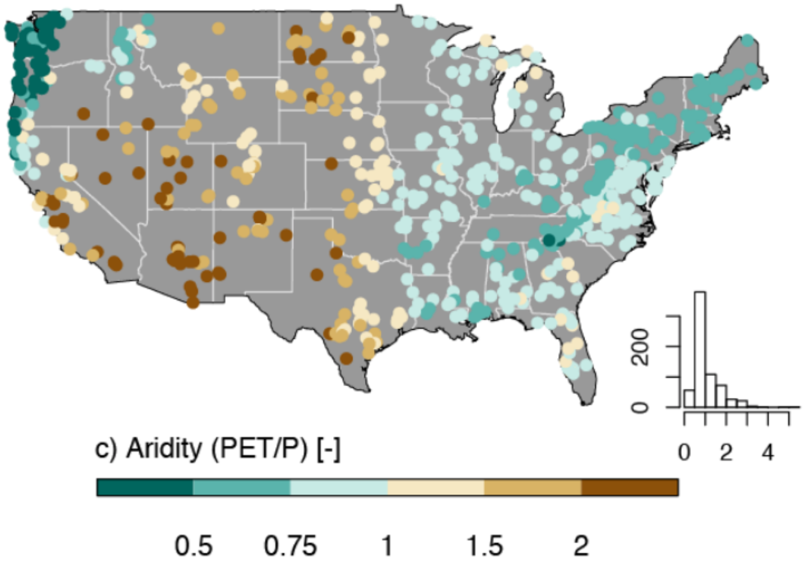
- A continuum of complexity
  - ***Process complexity***: Which processes are represented explicitly?
  - ***Spatial complexity***: To what extent do we explicitly represent details of the landscape, and spatial connections (flow of water) across model elements?

- 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*

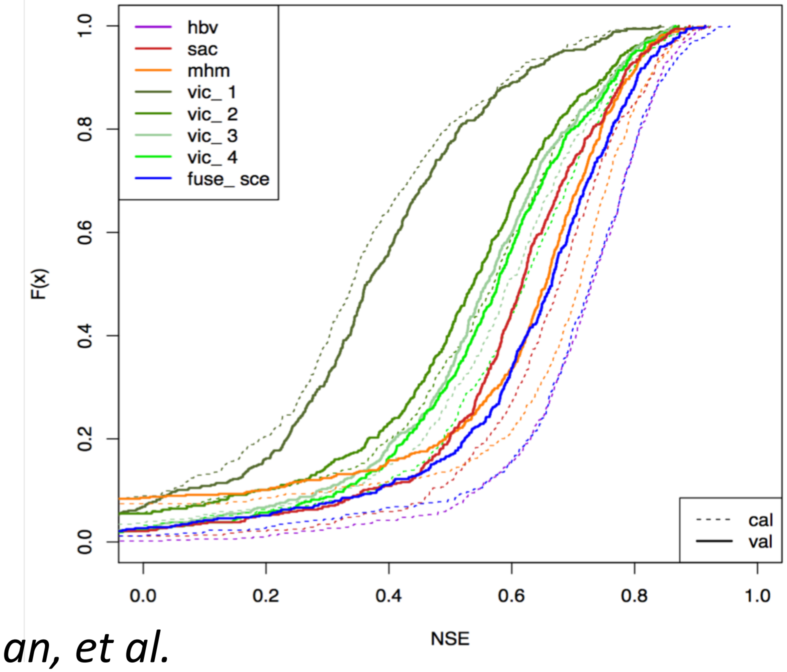
- Process-based hydrologic models
  - Explicitly represent dominant hydrologic and biophysical processes; explicitly represent details of the landscape
  - Reliance on geophysical data to estimate model parameters and widespread use of spatially constant parameters obtained from limited experimental 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 mode is too expensive to calibrate*

Results from many catchments and 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

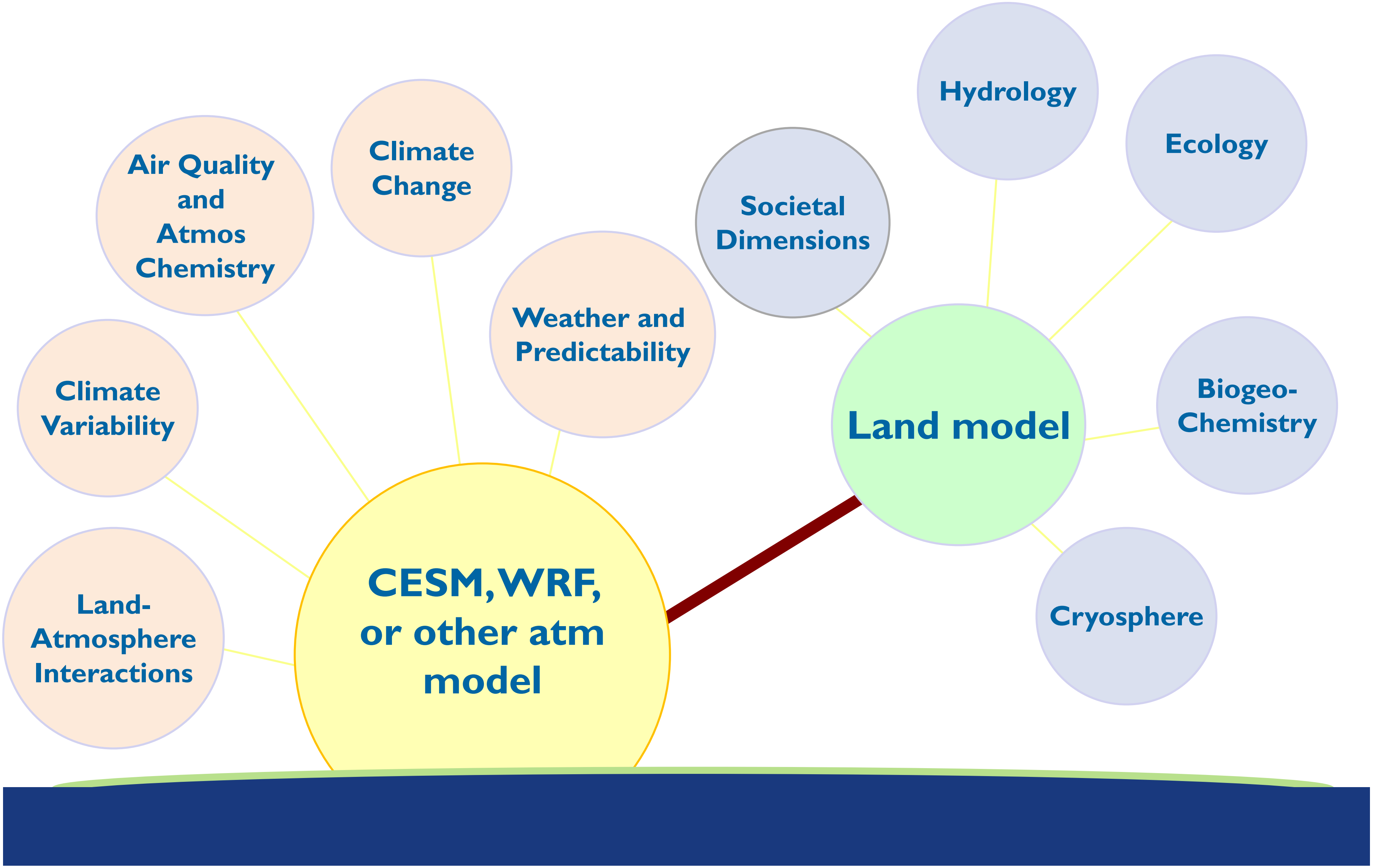


Efforts from Nans Addor, Naoki Mizukami, Andy Newman, et al.

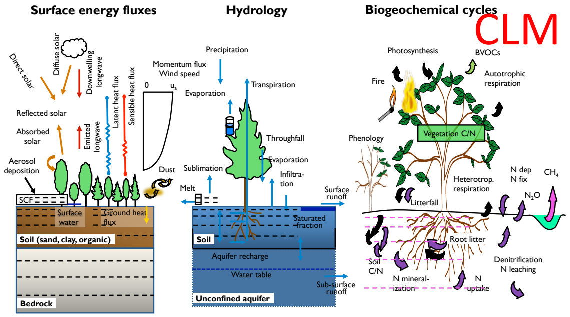
Outline

- Background
  - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- The emergence of CTSM
- Summary and research needs

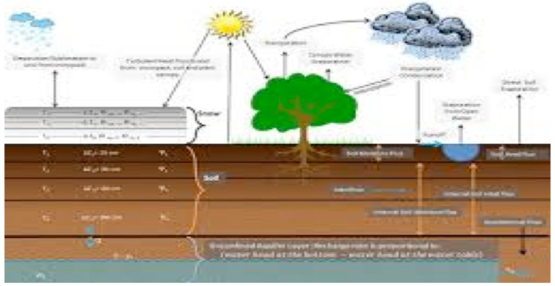
The interdisciplinary challenge of land modeling



CTSM



CLM

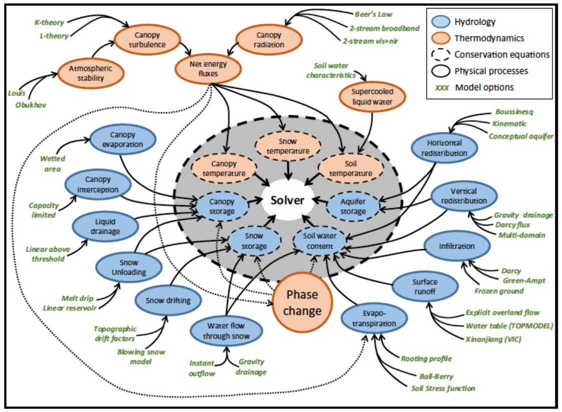


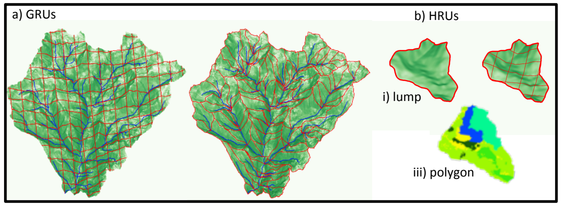
Noah-MP

Conceptual basis

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

SUMMA

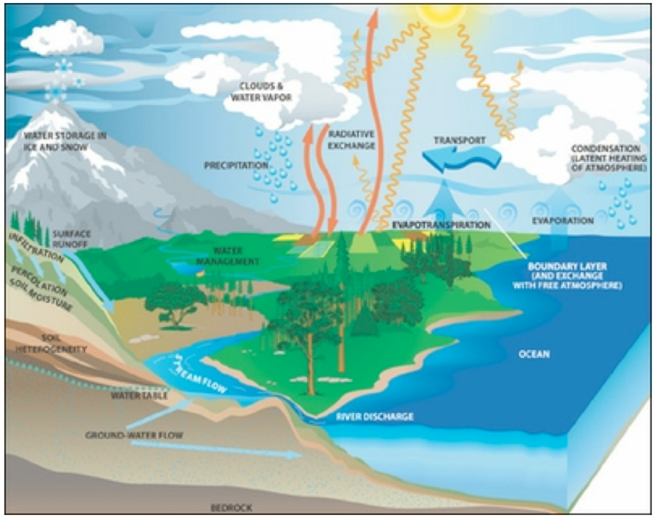




Formulates master model template which multiple models can be derived

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

The Community Terrestrial Systems Model (CTSM)

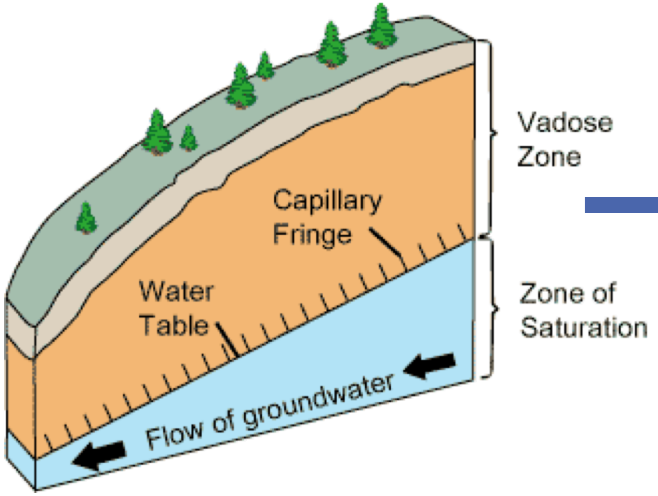


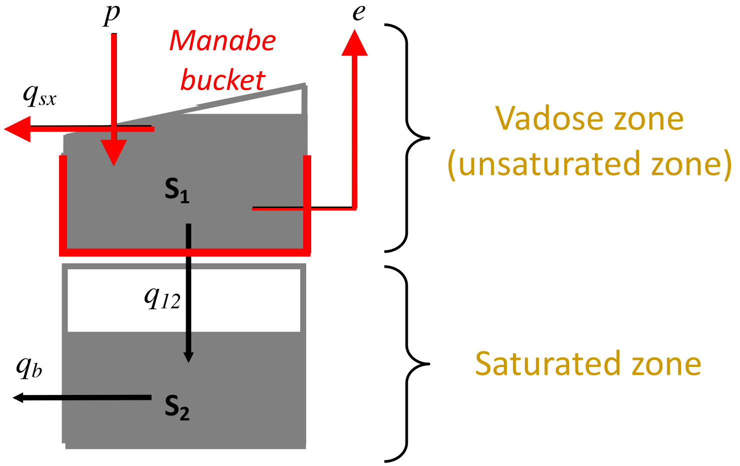
Unifies land models across climate, weather, water, and ecology

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

Model construction…

- Consider a very simple land model…





- Conservation equations

Multiple flux options (more complexity)

$$\frac{dS_1}{dt} = p - q_{sx} - e - \boxed{q_{12}}$$

$$\frac{dS_2}{dt} = q_{12} - q_b$$

Model parameters

- Flux parameterizations

$$q_{sx} = pA_s; \qquad A_s = 1 - \left(1 - \frac{S_1}{S_{1,\max}}\right)^b$$

Surface runoff

$$q_{12} = k_u \left(\frac{S_1}{S_{1,\max}}\right)^c$$

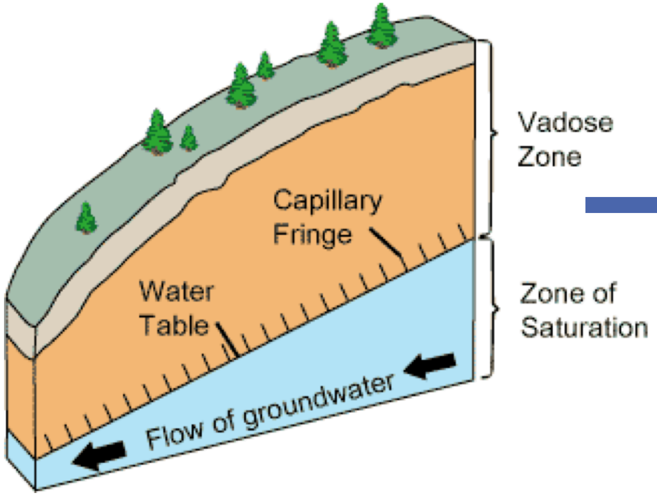
Vertical percolation

$$q_b = k_s \left(\frac{S_2}{S_{2,\max}}\right)$$

Baseflow

On top of spaghetti???

- Our very simple land model…



- Conservation equations

$$\frac{dS_1}{dt} = p - q_{sx} - e - q_{12}$$

$$\frac{dS_2}{dt} = q_{12} - q_b$$

- Common numerical implementation

$$\left.\begin{aligned} S_1^{n+1,*} &= S_1^n + p\Delta t \\ S_1^{n+1,**} &= S_1^{n+1,*} + q_{sx}\Delta t \\ S_1^{n+1,***} &= S_1^{n+1,**} + e\Delta t \\ \\ S_1^{n+1} &= S_1^{n+1,***} + q_{12}\Delta t \end{aligned}\right\}$$

Non-standard!
Physics are intertwined with numerics

Can't capitalize on decades of progress in applied math

More standard implementations

- The model state equations can be written as

$$\frac{\mathrm{d}\mathbf{S}}{\mathrm{d}t} = \mathbf{g}\left(\mathbf{S}, t\right)$$

- The exact solution of the average flux over the interval t^n (start of the time step) to t^{n+1} (end of the time step) is

$$\overline{\mathbf{g}}^{n \rightarrow n+1} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \left(\mathbf{g}(\mathbf{S}, \varsigma), \varsigma \right) \mathbf{d}\varsigma$$

- Given an estimate of the average flux, the model state variables can be updated as

$$\mathbf{S}\left(t^{n+1}\right) = \mathbf{S}\left(t^{n}\right) + \Delta t \overline{\mathbf{g}}^{n \rightarrow n+1}$$

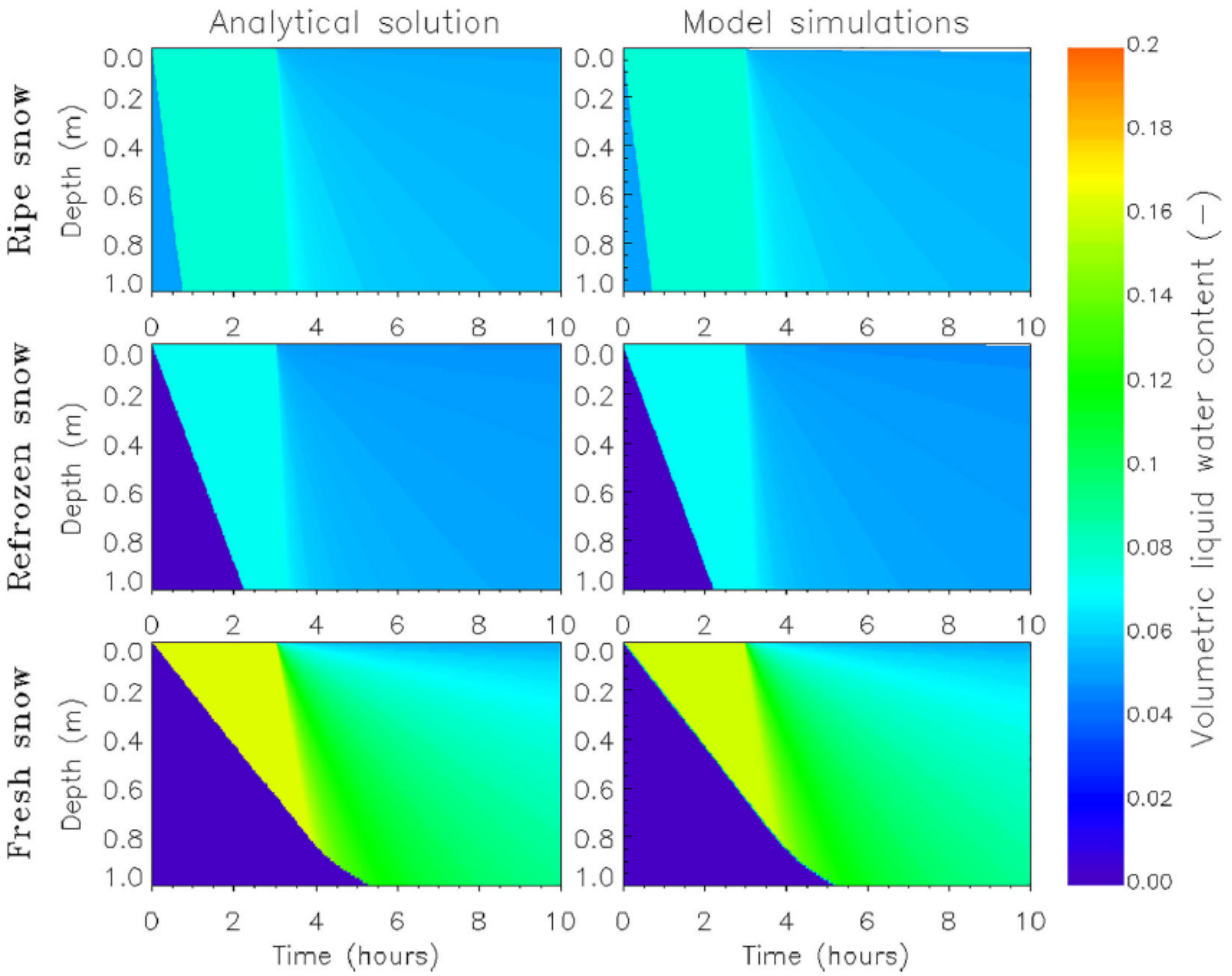
- The exact solution is computationally expensive, so approximations to the exact solution are used
- The approximation controls the stability, accuracy, smoothness, and efficiency of the solution

A controlled approach to model development

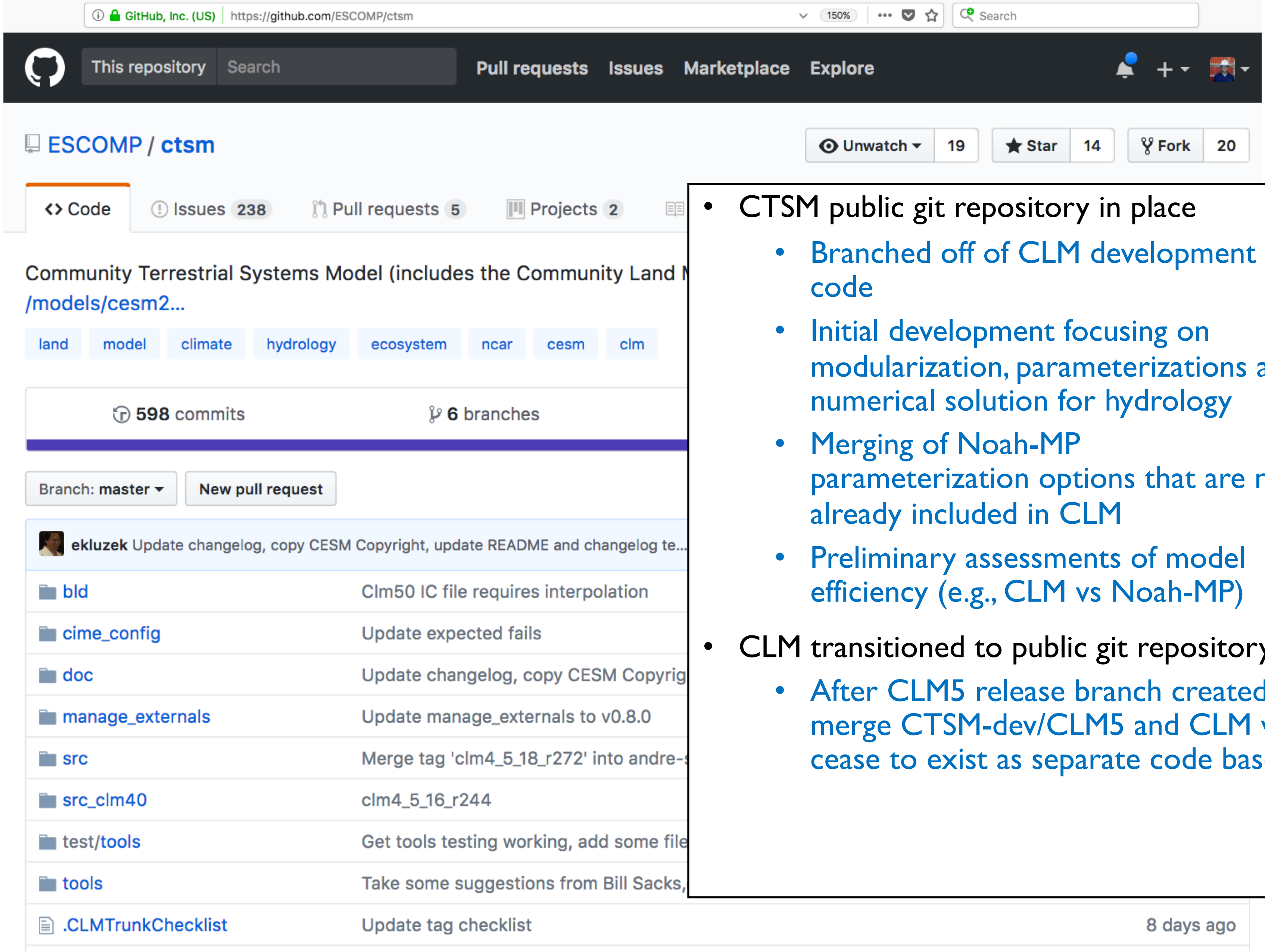
Laugh tests for land models

Constant precip for three hours at top of a 1-m snowpack

Analytical solution



CTSM is public



GitHub, Inc. (US) https://github.com/ESCOMP/ctsm 150%

ESCOMP / ctsm

Community Terrestrial Systems Model (includes the Community Land M
/models/cesm2...

598 commits
6 branches

ekluzek Update changelog, copy CESM Copyright, update README and changelog te...

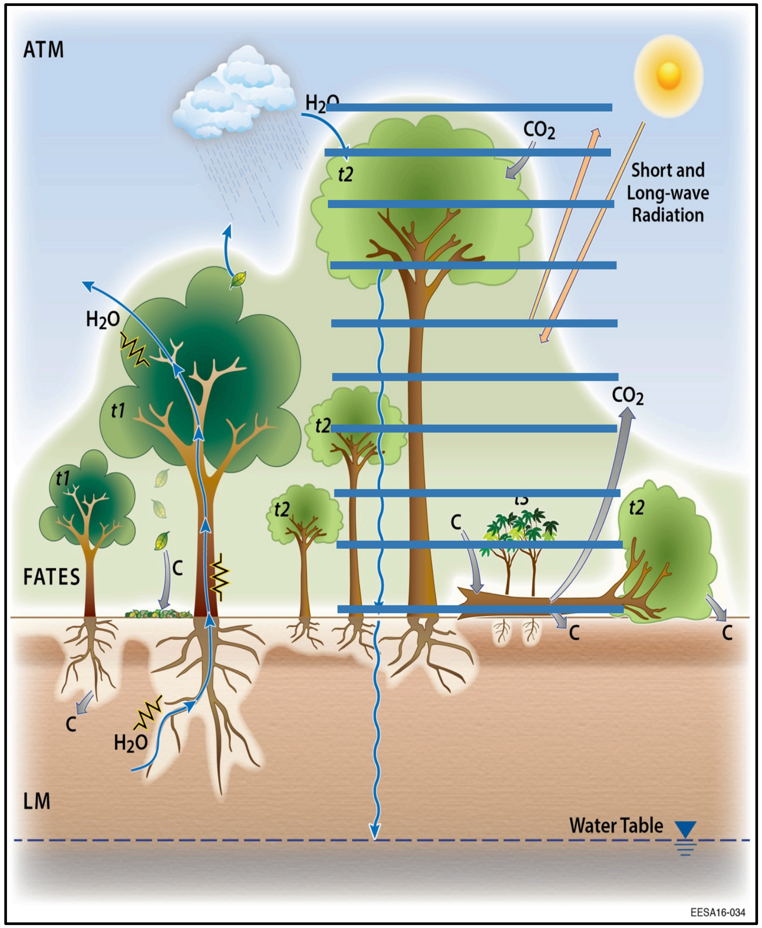
bld	Clm50 IC file requires interpolation	
cime_config	Update expected fails	
doc	Update changelog, copy CESM Copyrig	
manage_externals	Update manage_externals to v0.8.0	
src	Merge tag 'clm4_5_18_r272' into andre-	
src_clm40	clm4_5_16_r244	
test/tools	Get tools testing working, add some file	
tools	Take some suggestions from Bill Sacks,	
.CLMTrunkChecklist	Update tag checklist	8 days ago

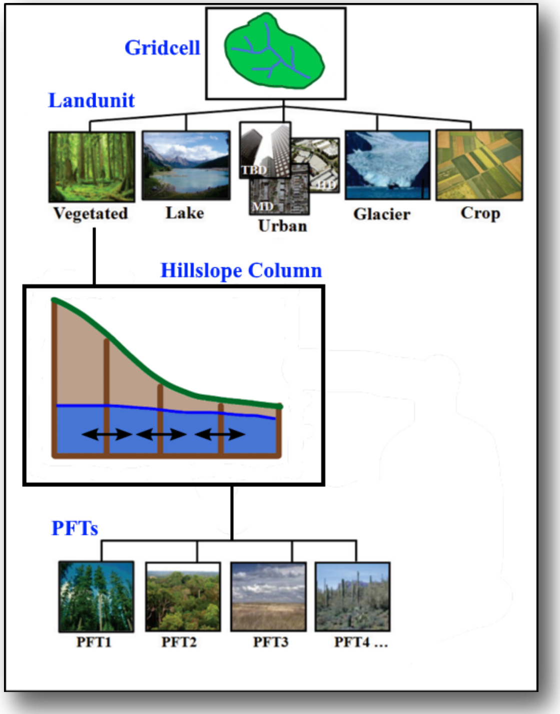
- CTSM public git repository in place
  - Branched off of CLM development code
  - Initial development focusing on modularization, parameterizations and numerical solution for hydrology
  - Merging of Noah-MP parameterization options that are not already included in CLM
  - Preliminary assessments of 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



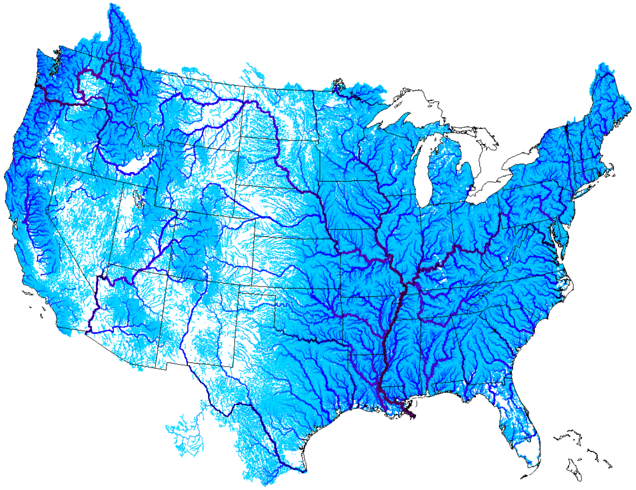


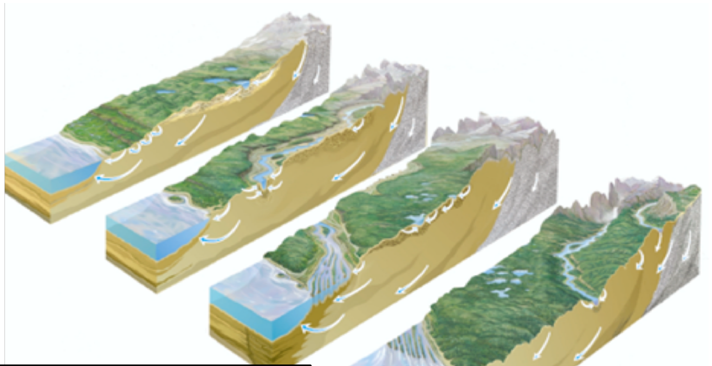
Lateral fluxes of water

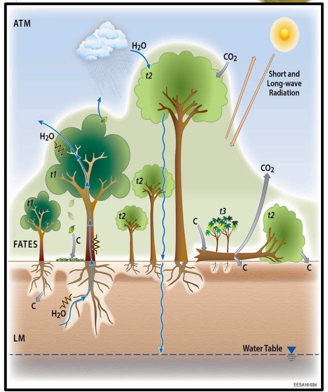
Water and land management

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

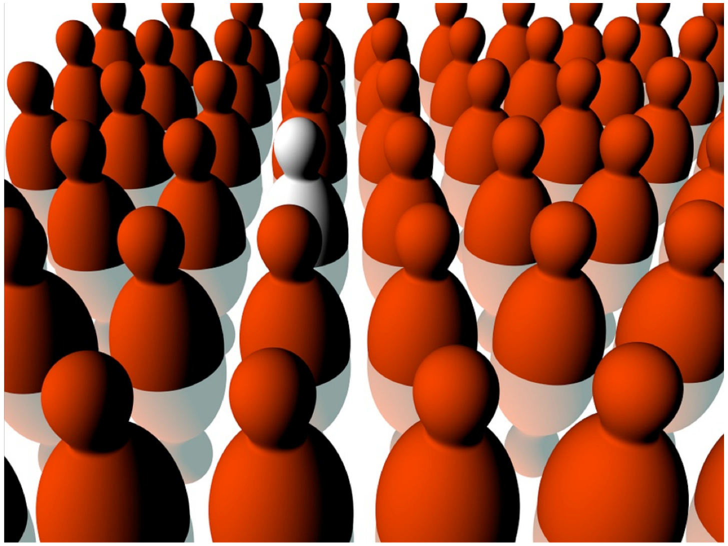






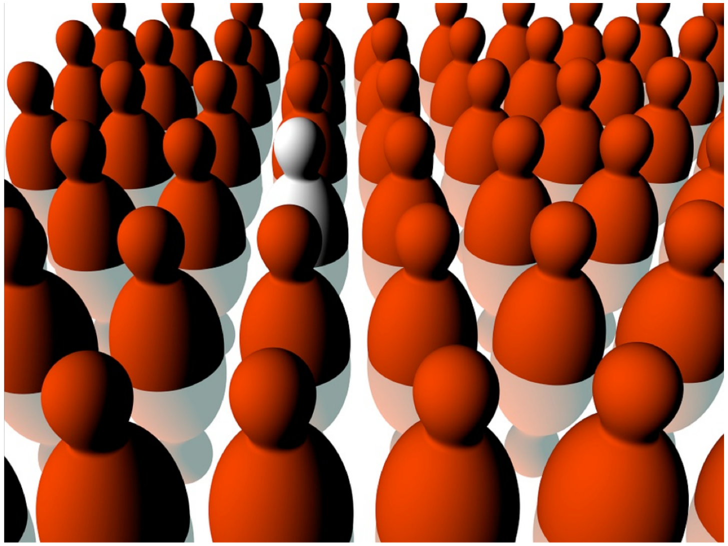
Benefits of a unified land model

- Improve understanding of differences among models (debate about processes)
  - Model inter-comparison experiments flawed because too many differences among participating models
- Improve understanding of model limitations
  - Most models not constructed to enable a controlled and systematic approach to model development and improvement
- Improve characterization of model uncertainty
  - Explicitly characterize uncertainty in individual modeling decisions
  - Enables shift from small-ensemble to large-ensemble framework
- Unite disparate (disciplinary) modeling efforts
  - Without a unified modeling framework the community cannot effectively work together, learn from each other, and accelerate model development
- Reduce duplication of effort



Benefits of the proposed model structure

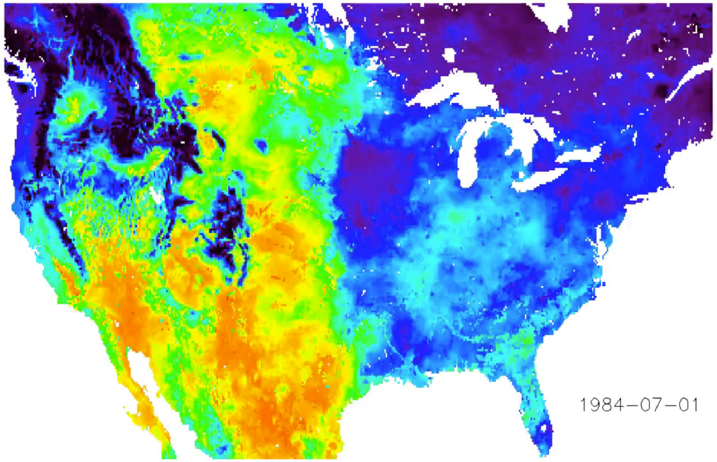
- Simplifies sharing of code and concepts across different model development groups
  - Separating physics from numerics (the “structural core”) and modularity at the flux level accelerates the process of adding/testing new capabilities
- Enables users to include/exclude specific processes
  - Model can be tailored to suit multiple applications
  - Model simplification opens up new possibilities for teaching and research
- Simplifies data assimilation efforts
  - Formalizes the input-state-output relationships, meaning land model construction matches data assimilation methods
- Reduces development costs
  - 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

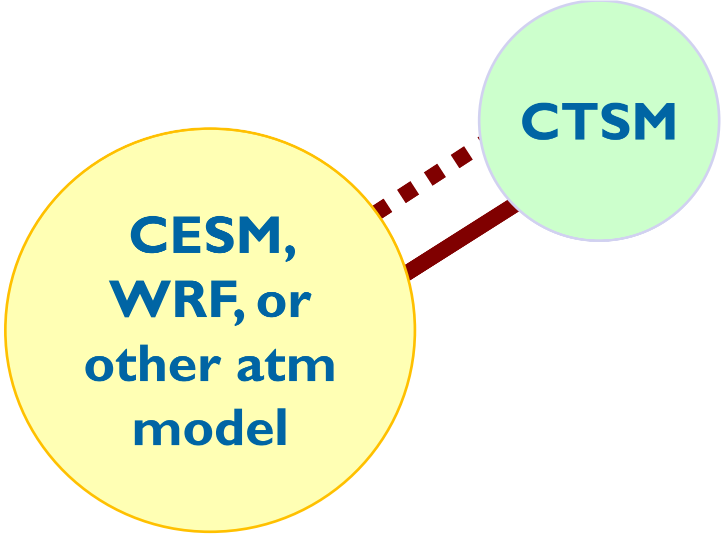


CTSM challenges

- 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

- Background
    - Remarkable scientific and technical advances in many areas supporting hydrologic modeling and prediction
- Modeling challenges
    - Processes
    - Parameters
    - Computing
- The emergence of CTSM
- Summary and research needs

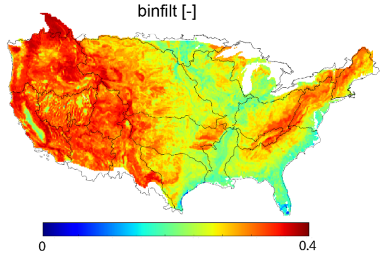
Modeling opinions

- 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

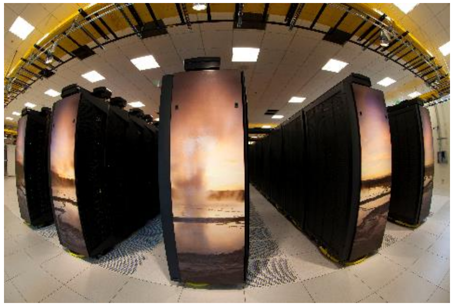
---

- Processes
  - We really need to focus on the scaling problem – use a mix of explicit discretization and implicit parameterizations to improve simulations of large-scale fluxes

- Parameters
  - We really need to incorporate stronger hydrologic theory when evaluating model parameters – it's a physics problem
  - 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

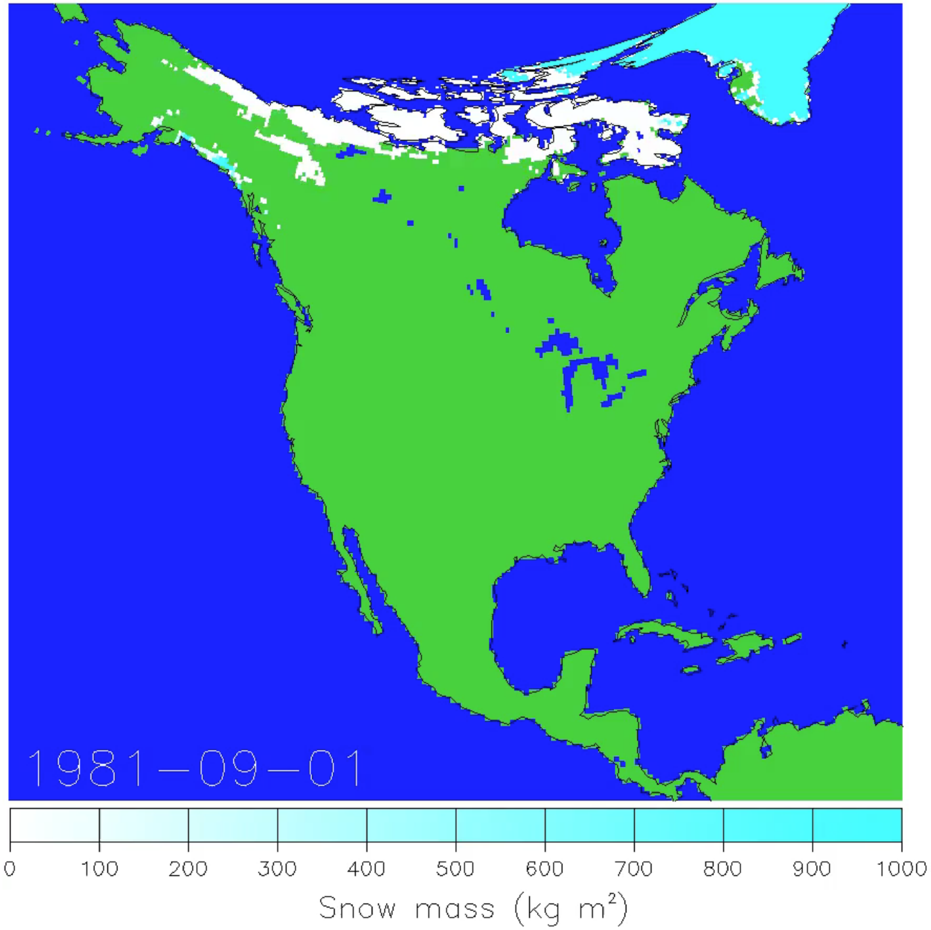


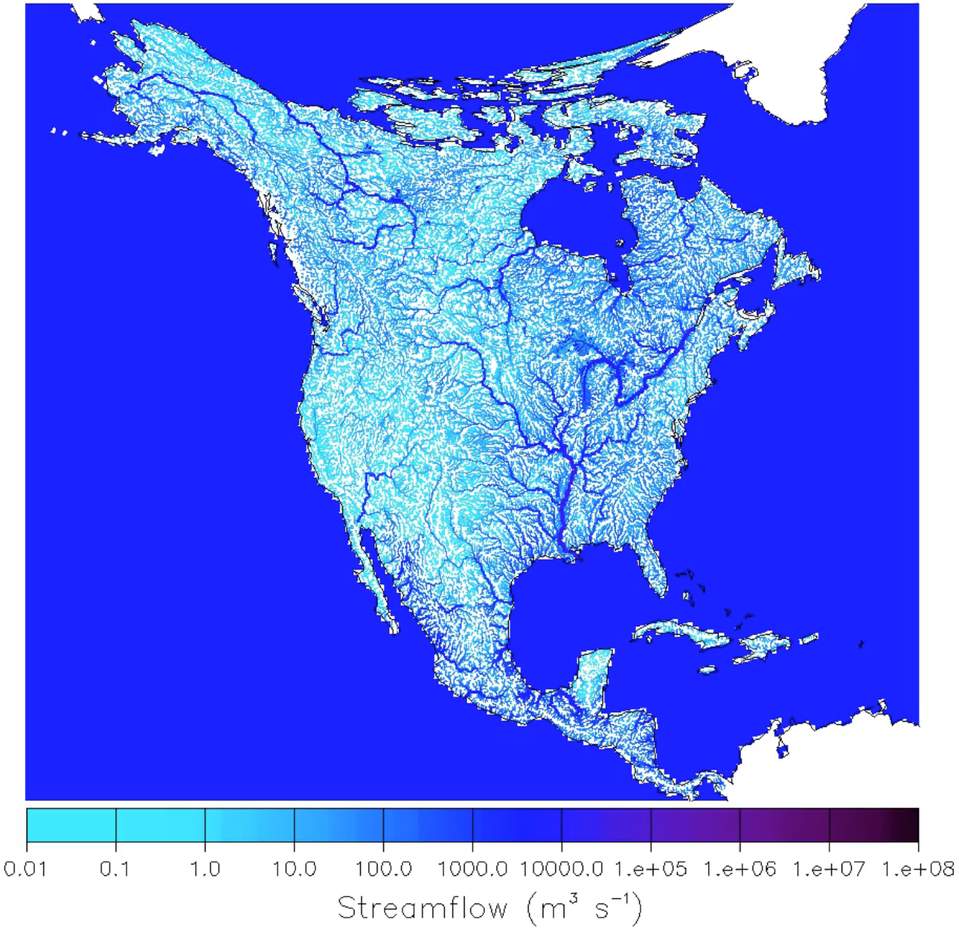
Modeling strategy

- A three-pronged modeling strategy
  - **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.

Continental-domain modeling

- Advance large-domain simulations





Specific research needs

1. Unify process-based land modeling
   - Inter-component coupling (make use of legacy models)
   - Intra-component coupling (advance model construction)
2. Advance community hydrologic modeling (rather than a single model)
   - 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. Include/improve missing/poorly represented processes in land models
   - Glaciers, permafrost dynamics, water quality, stream temperature, river ice, etc.
   - Groundwater, humans as an endogenous component of the Earth System
4. Systematically explore the benefits of competing modeling approaches
   - Scrutinize models using data from research watersheds
   - Evaluate information gains/losses using models of varying complexity
5. Construct variable-complexity models
   - Capabilities to simplify process complexity and spatial complexity
   - Advance applications that require “agile” models
   - Evaluate accuracy-efficiency tradeoffs

Specific research needs (cont.)

6. Develop better continental-domain forcing data
   - Probabilistic approach to combine NWP models, radar, and station data
   - Meaningful multi-scale structure and inter-variable relationships

7. Advance research on process-oriented approaches to estimate spatial fields of model parameters – *parameter estimation as a physics problem*
   - 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

8. 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

9. Advance methods to characterize and quantify uncertainty
   - Epistemic and aleatory uncertainty
   - Ensure conclusions are not contaminated by over-confidence

10. 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.
   - A more productive dialog between experimentalists and modelers

QUESTIONS?

Martyn P. Clark
University of Saskatchewan at Canmore
martyn.clark@usask.ca