A vision for continental-domain hydrologic modeling

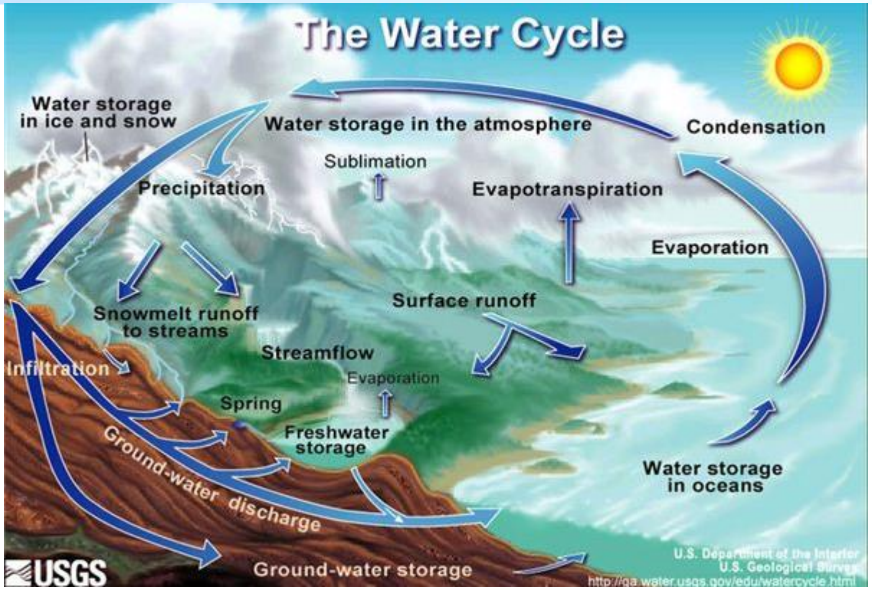
IMPC Meeting, Saskatoon, 19 July 2018

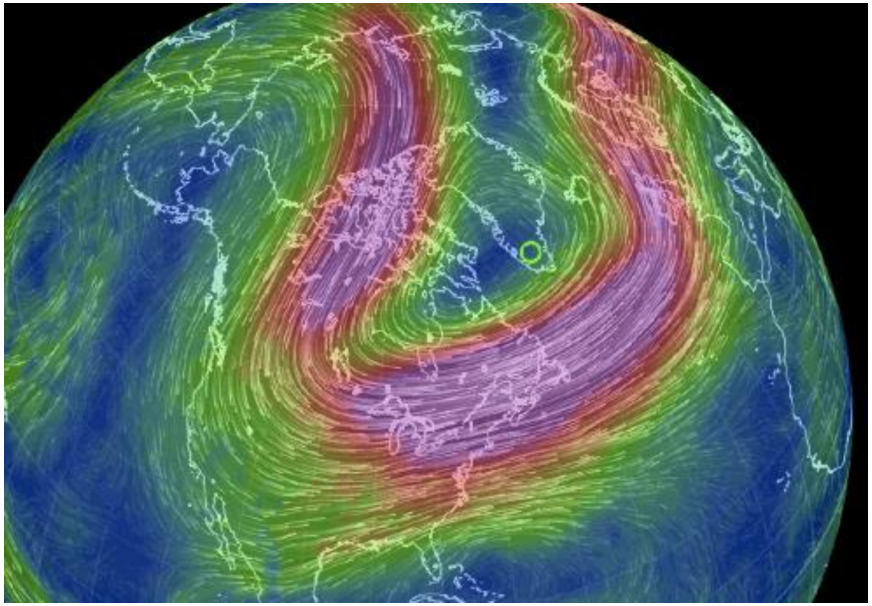
Outline

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- Research applications
  - Developing storylines of the future
  - Advancing streamflow forecasting capabilities
- Summary and research needs

The nature of the hydrologic modeling problem

- 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 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

→

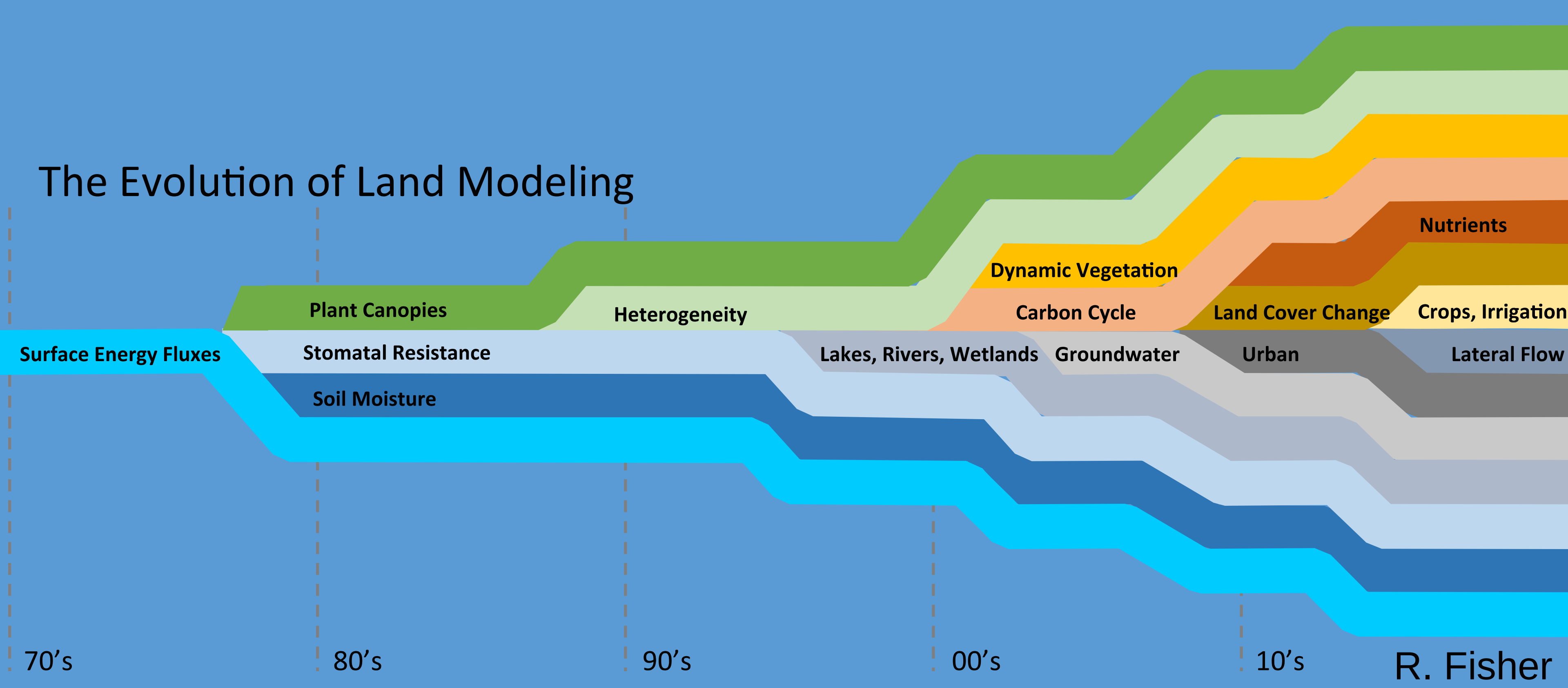
Mechanistic modeling of land processes

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)



The path to model improvement is not obvious…

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

Outline

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- Research applications
  - Developing storylines of the future
  - Advancing streamflow forecasting capabilities
- 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 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

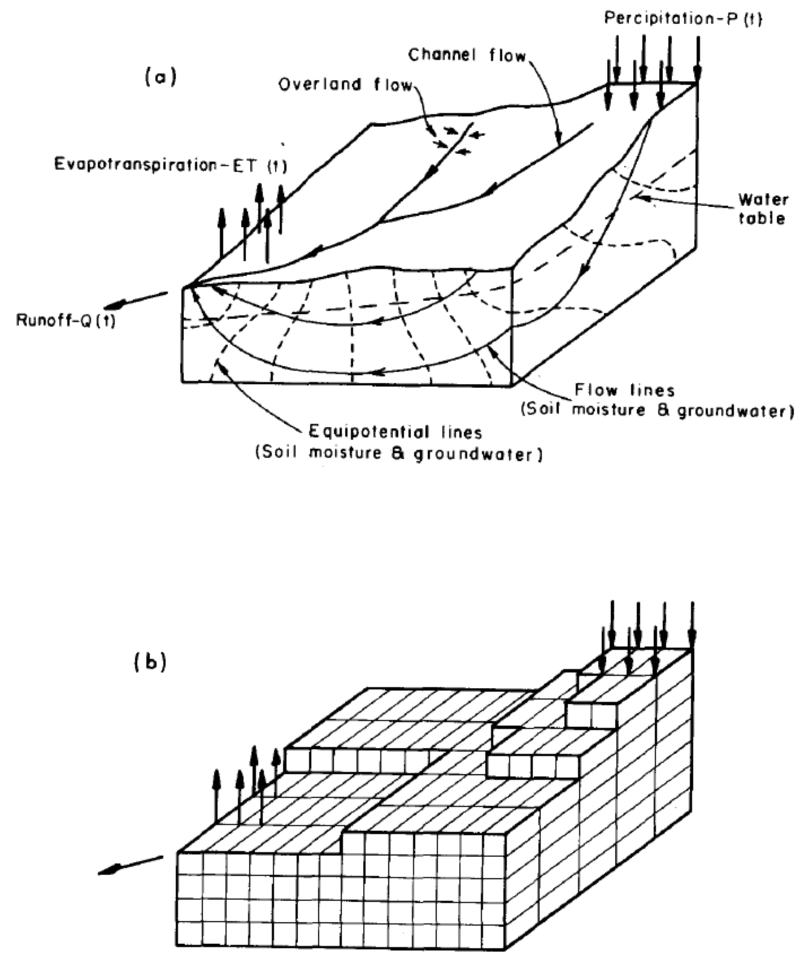


Fig. 3. Schematic diagram of (a) Hydrologic basin and (b) Three dimensional nodal model of hydrologic basin.

Questions posed by Freeze and Harlan

NCAR UCAR

- 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?

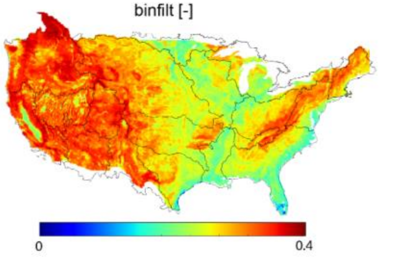
- Have the earlier computer limitations of storage capacity and speed of computation been overcome? Is the application of digital computers to this type of problem economically feasible?

Key challenges

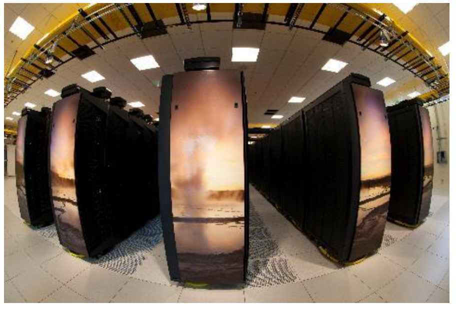
- The choice of modeling approaches (arguably) stems from personal preferences for physics or parsimony
- Need a stronger scientific basis for model development/improvement
    - Treat numerical modeling as a subjective decision-making process – *carefully evaluate all modeling decisions in a controlled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

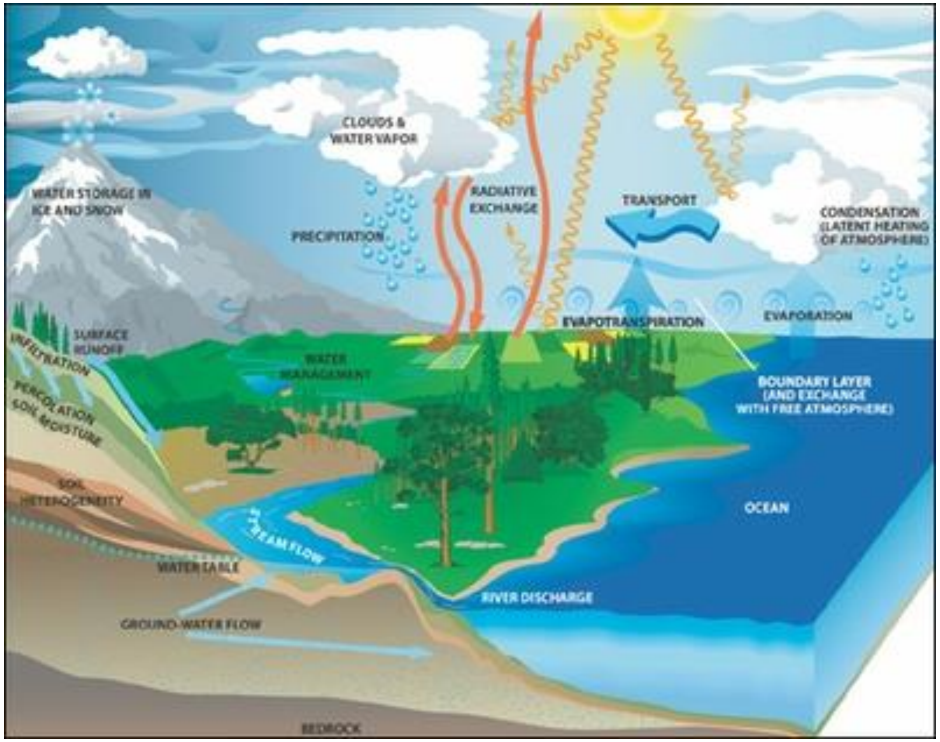


Challenge 1: Modeling processes

- 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
  - …

Modeling approach



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

<u>Conceptual basis</u>:

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

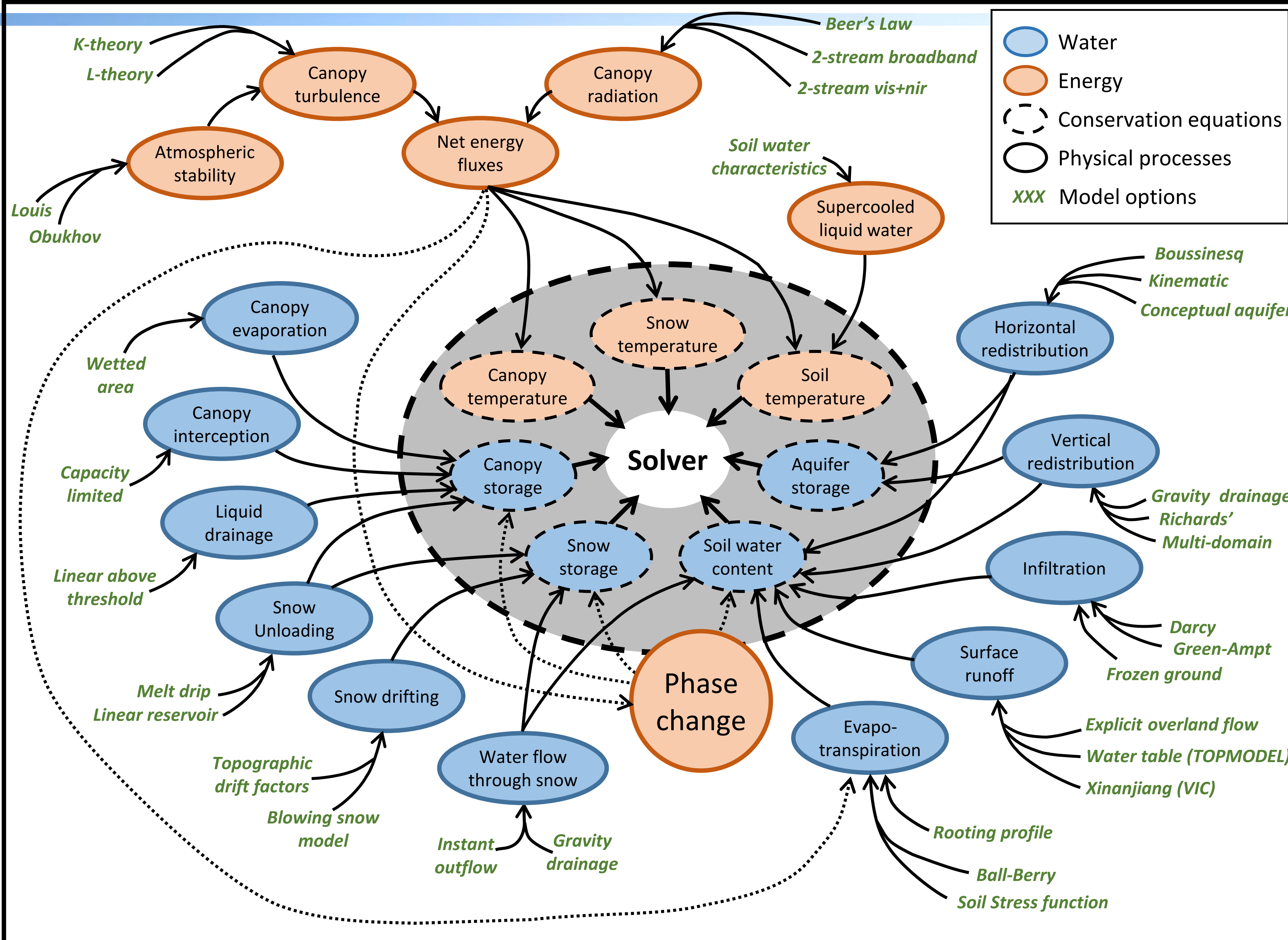
2. Differences among models relate to
   a) the spatial discretization of the model domain;
   b) the approaches used to parameterize individual fluxes (including model parameter values); and
   c) the methods used to solve the governing model equations.

The Structure for Unifying Multiple Modeling Alternatives (SUMMA):

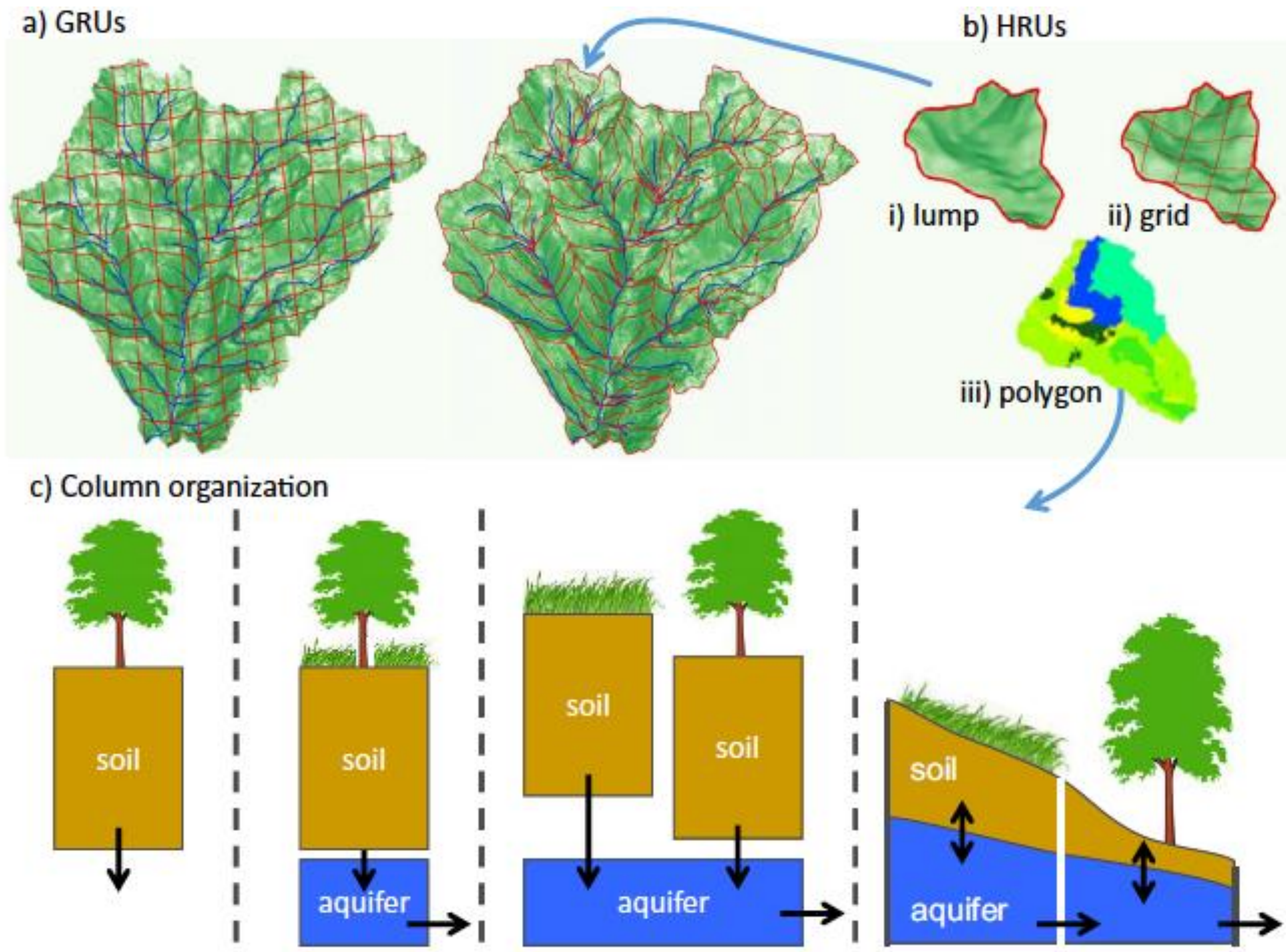
Defines a single set of conservation equations for land biogeophysics, with the capability to use different spatial discretizations, different flux parameterizations and model parameters, & different time stepping schemes

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

Process flexibility



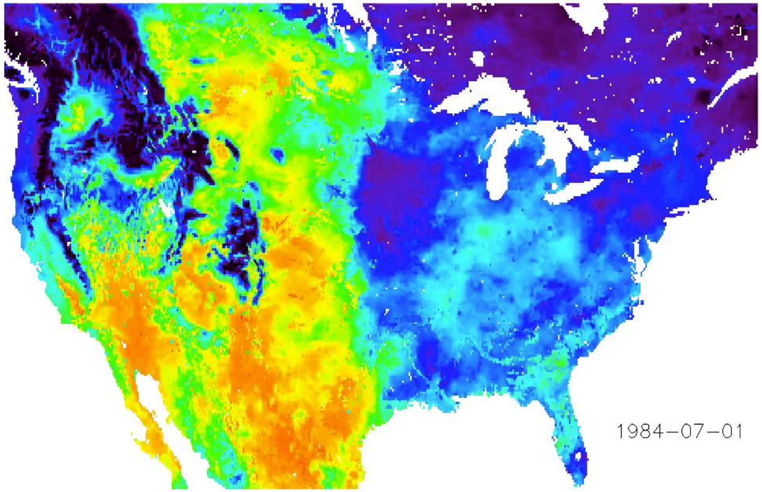
Spatial flexibility

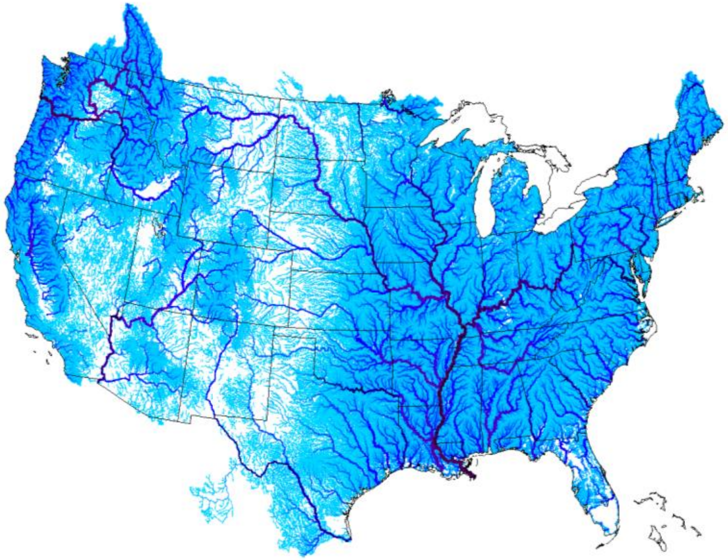


Use cases

- 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

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]

[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)

Challenge 2: Model parameters

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

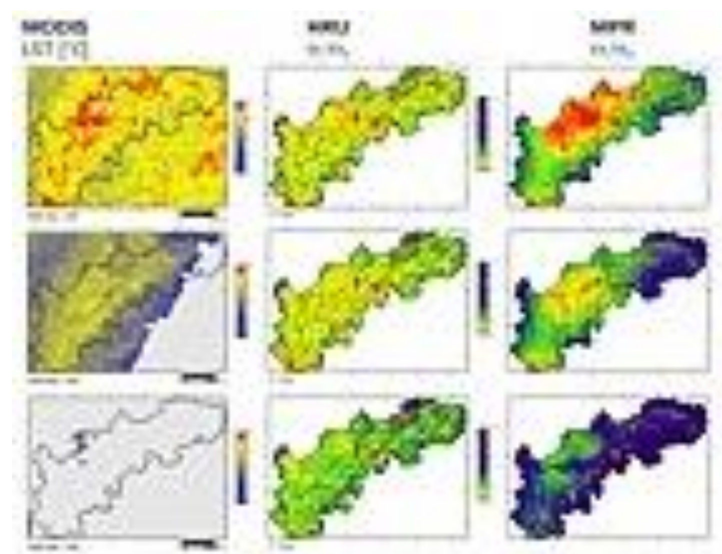
```
! ------------------------ local variables -----------------------------------------
  INTEGER :: IB            !waveband class

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

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

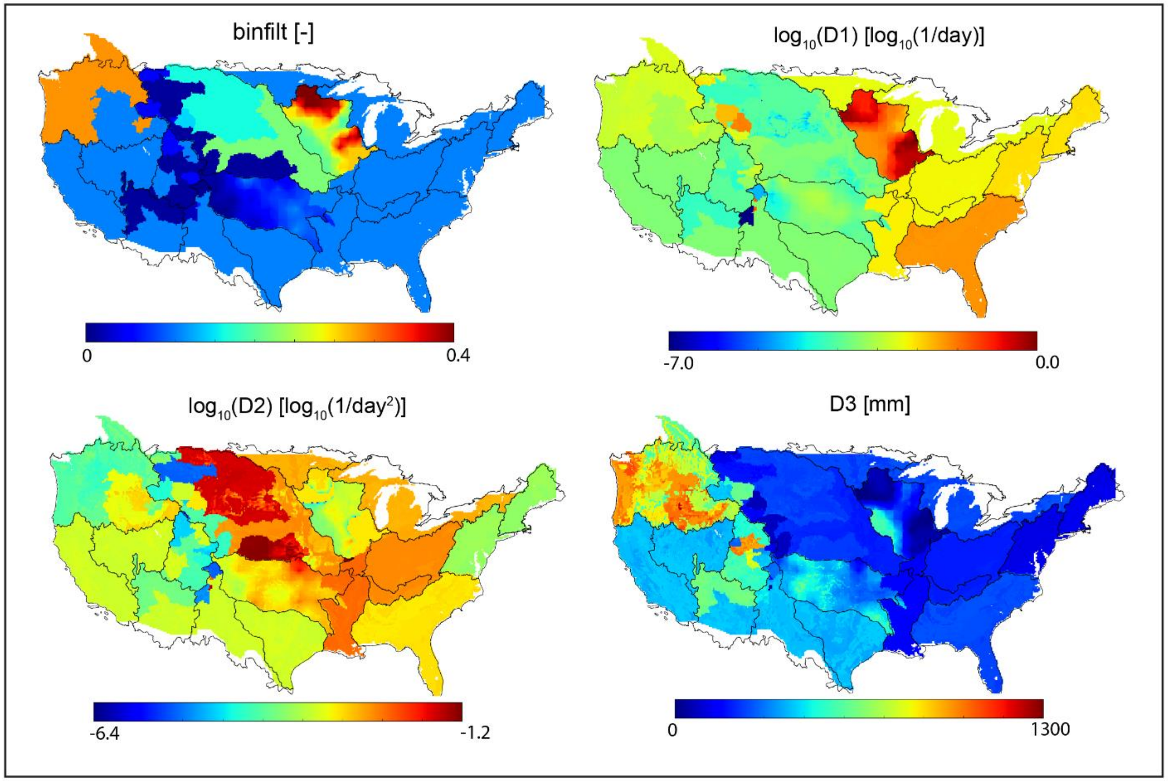
- Another solution: Focus squarely on relating geophysical attributes to model parameters (MPR)



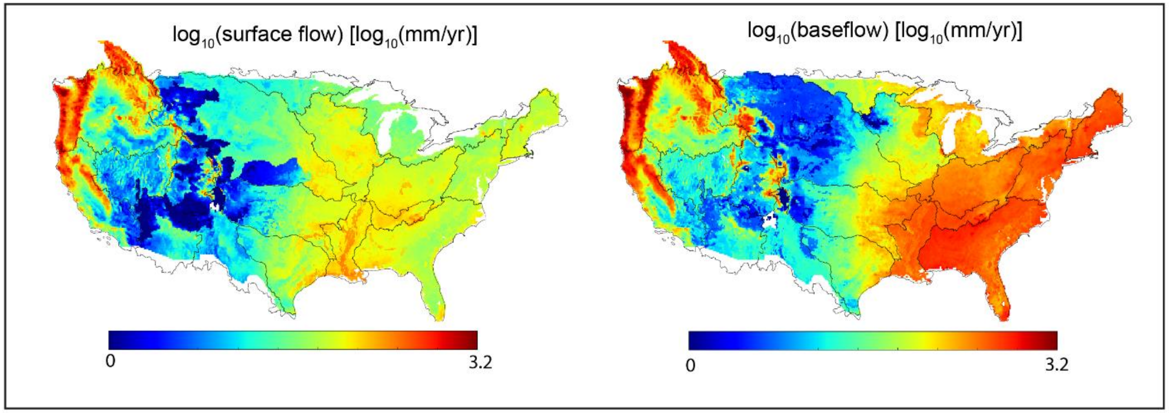
Default params

- 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

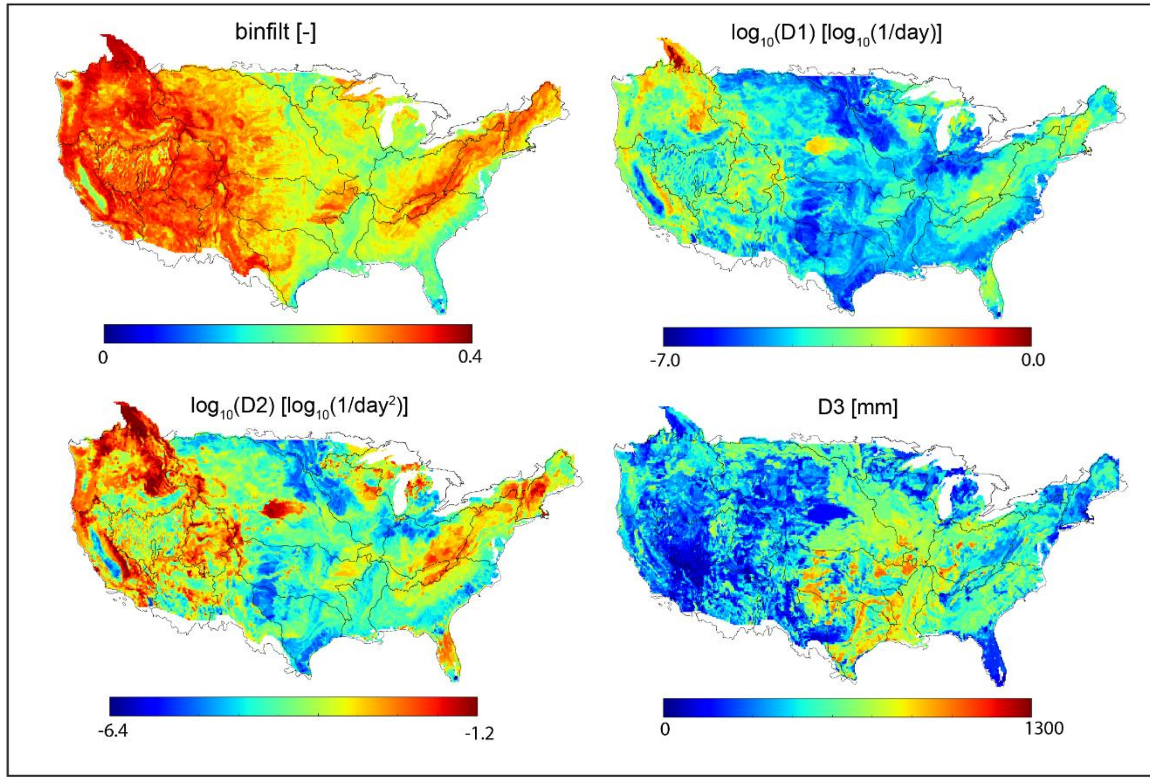
MPR-flex

- 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

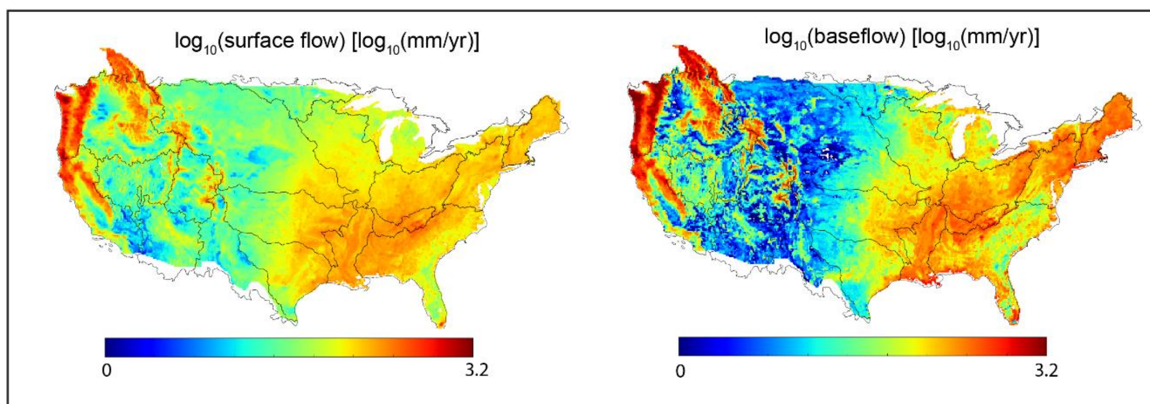
- *Define transfer functions for new models – develop model agnostic MPR (MPR-Flex)*

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

VIC Soil parameters – MPR

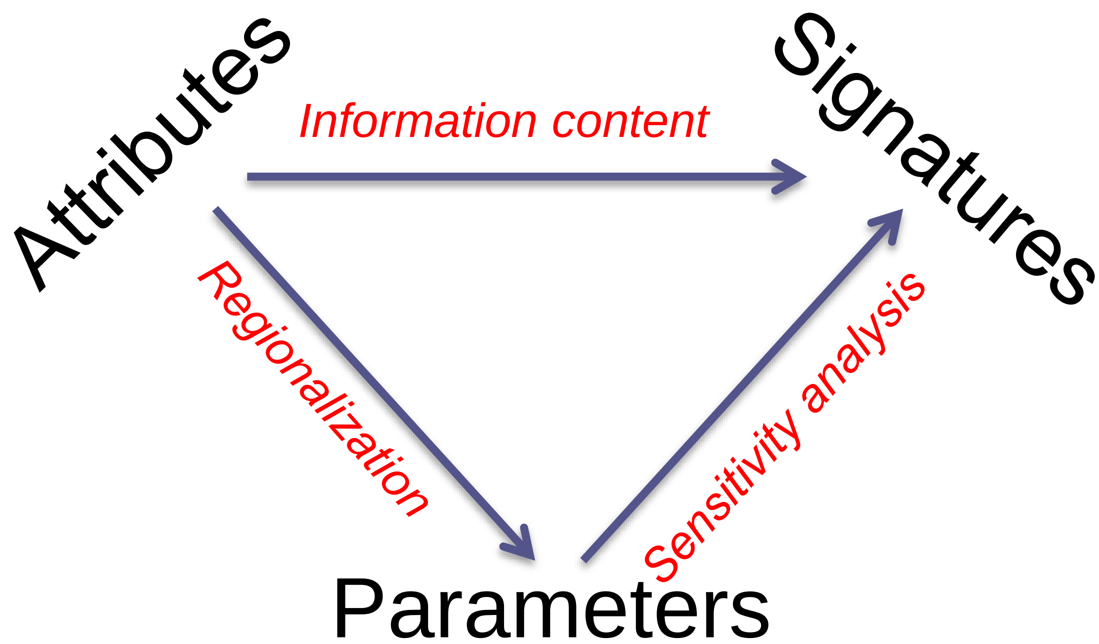


1950-1999 annual mean runoff



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

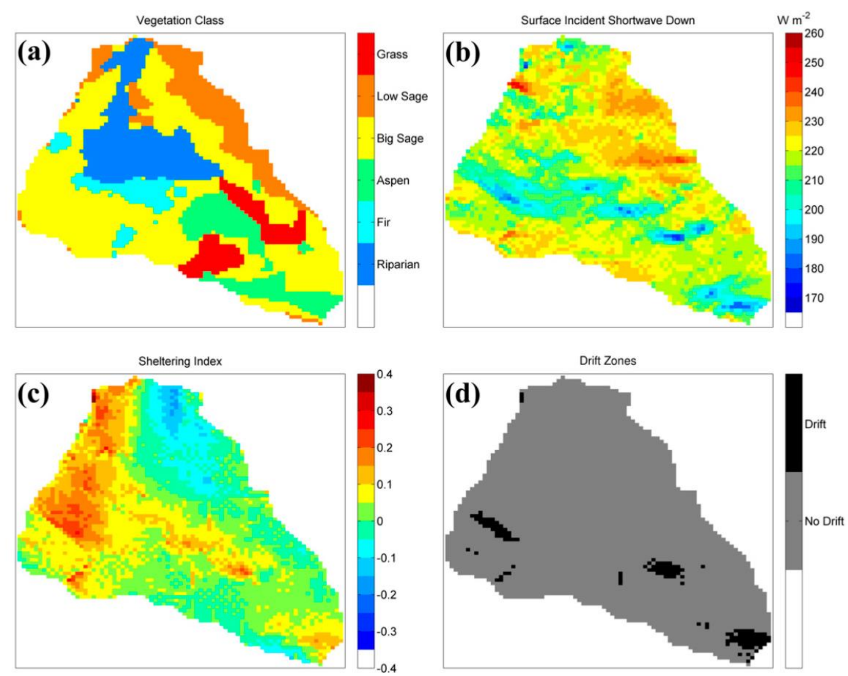
How does landscape evolution define the storage and transmission properties of the landscape?

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

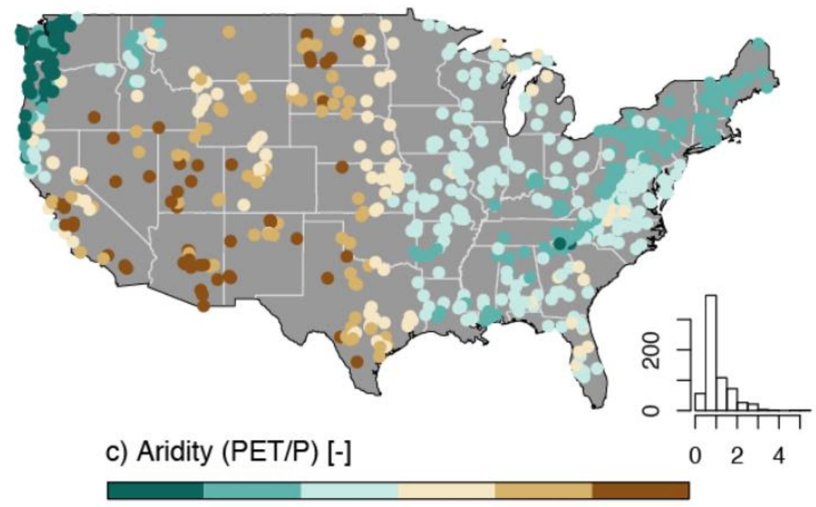
- A continuum of complexity
  - ***Process complexity***: Which processes are represented explicitly?
  - ***Spatial complexity***: To what extent do we explicitly represent details of the landscape, and spatial connections (flow of water) across model elements?

- 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*

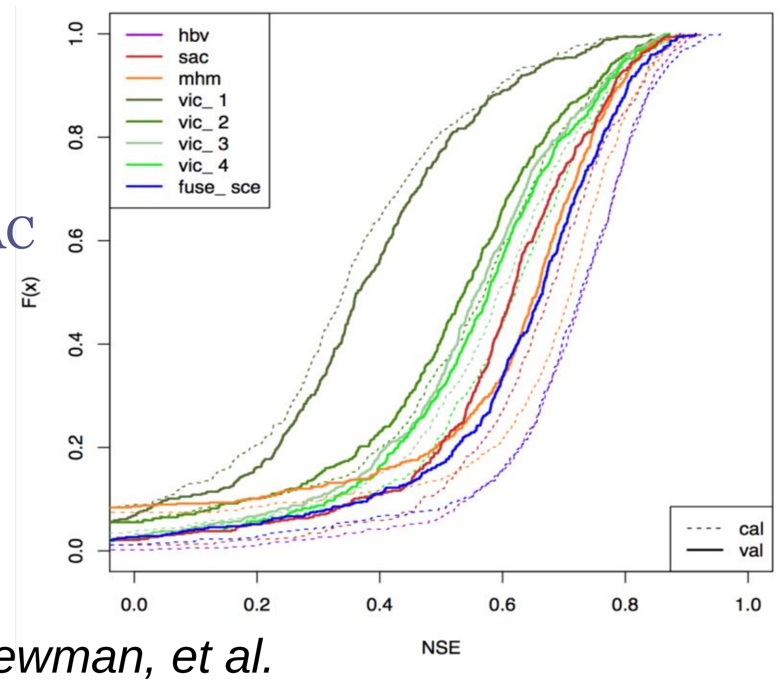
- Process-based hydrologic models
  - Explicitly represent dominant hydrologic and biophysical processes; explicitly represent details of the landscape
  - Reliance on geophysical data to estimate model parameters and widespread use of spatially constant parameters obtained from limited experimental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



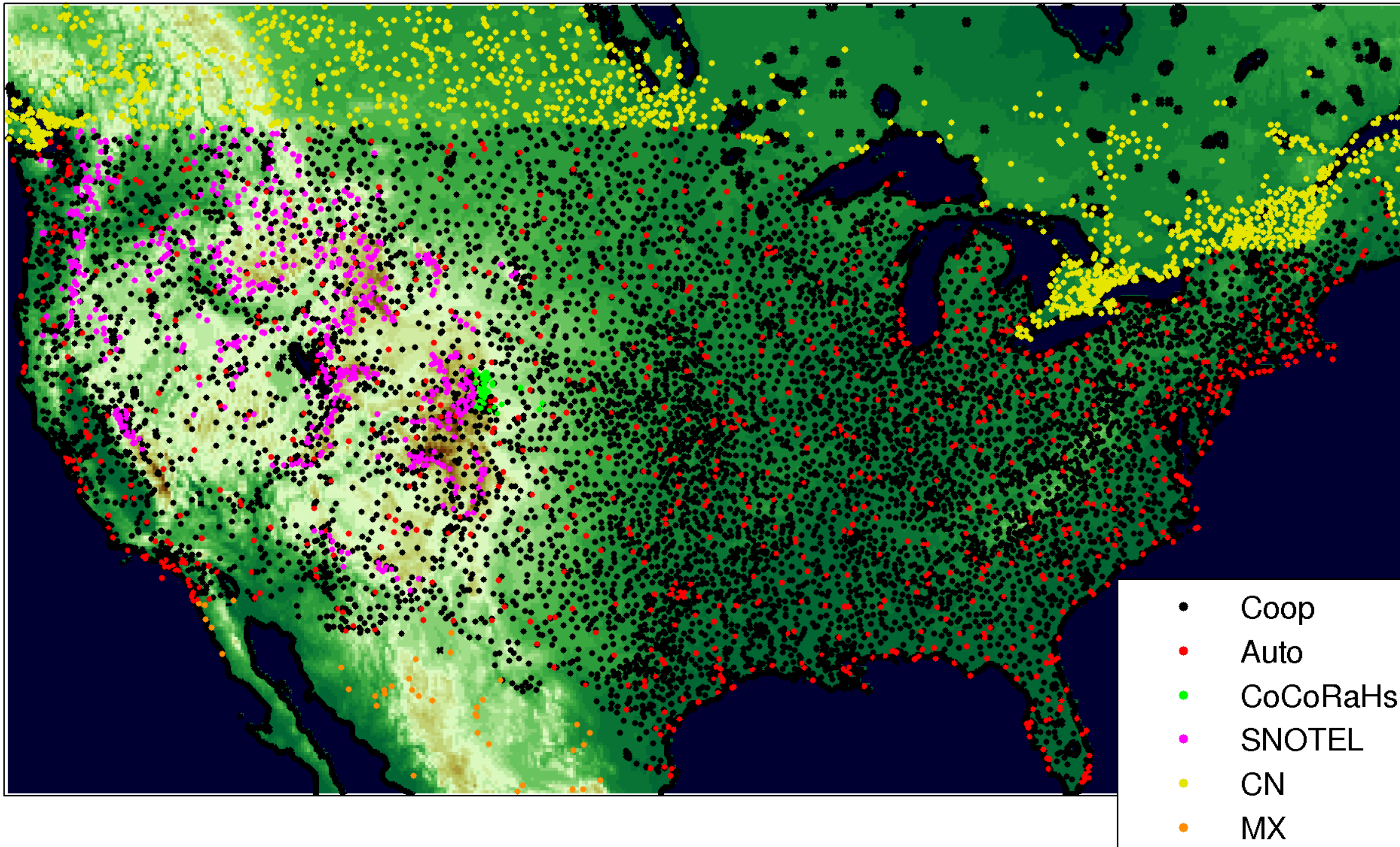
- Large model sample
  - Existing models
    - VIC, CLM, Noah-MP, PRMS, HBV, MHM, SAC
  - Multiple hypothesis frameworks
    - FUSE and SUMMA
    - Clark et al., 2008; 2011; 2015a,b



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

Ensemble spatial met. fields

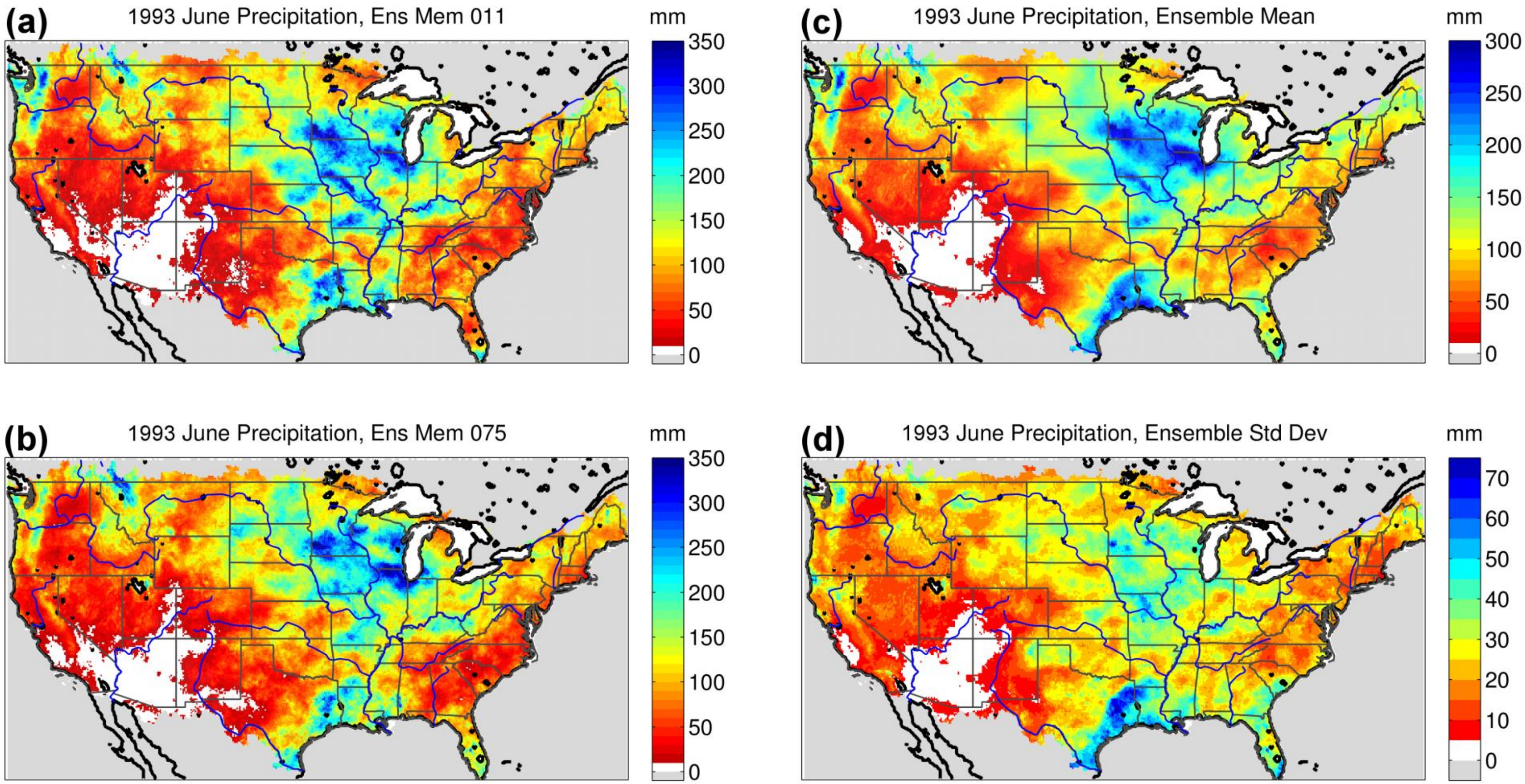
Spatial extrapolation from 12,000+ stations across the CONUS



CONUS product

- Dataset constructed from 1980-2012
- Daily spatial fields of precipitation and temperature
- Dataset freely available

Example output for June 1993



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

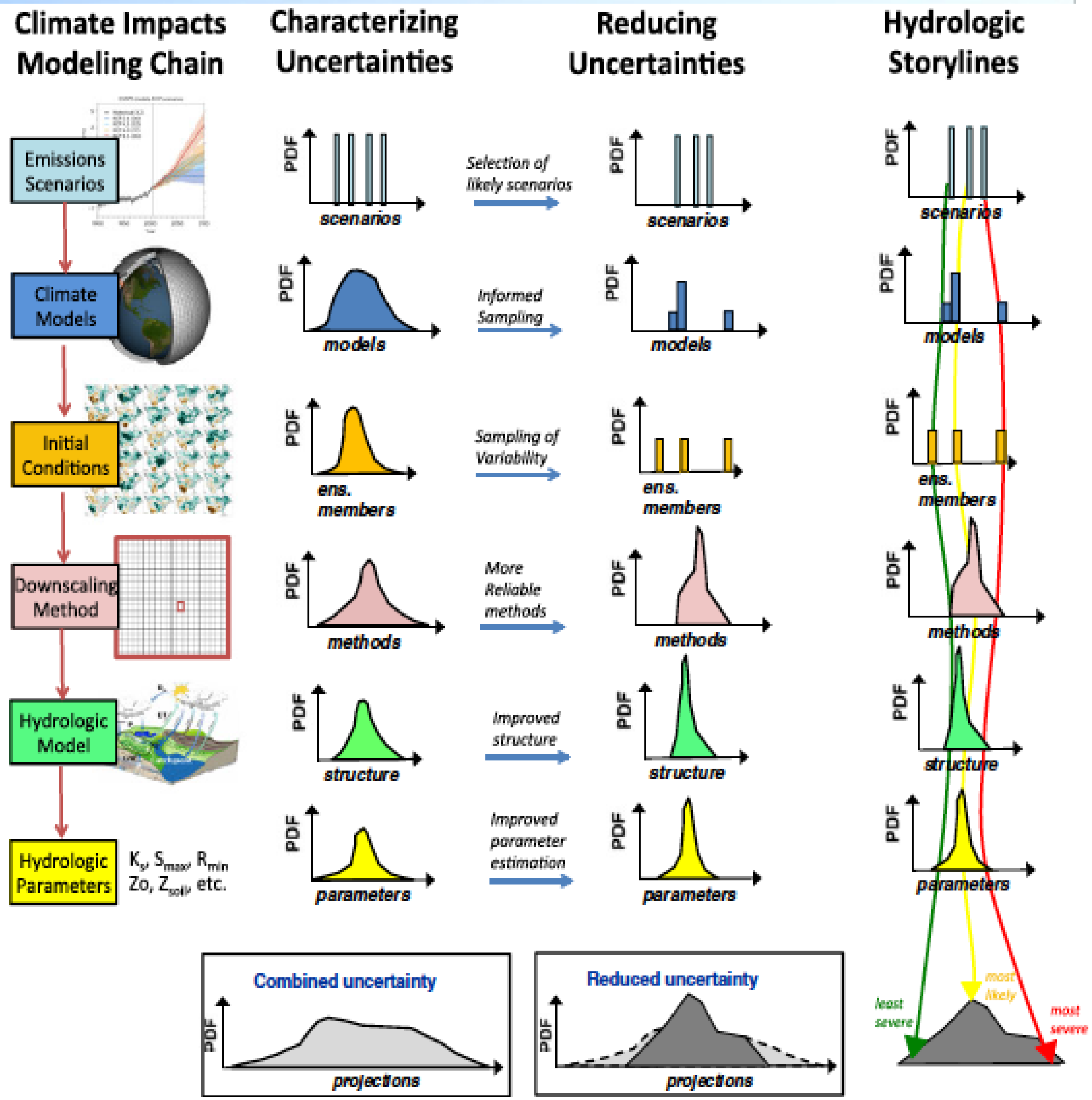
Outline

NCAR UCAR

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- Research applications
  - Developing storylines of the future
  - Advancing streamflow forecasting capabilities
- Summary and research needs

Developing storylines of the future

- Characterize uncertainty: "full" coverage of model hypothesis space
- Reduce uncertainty: cull bad models and methods



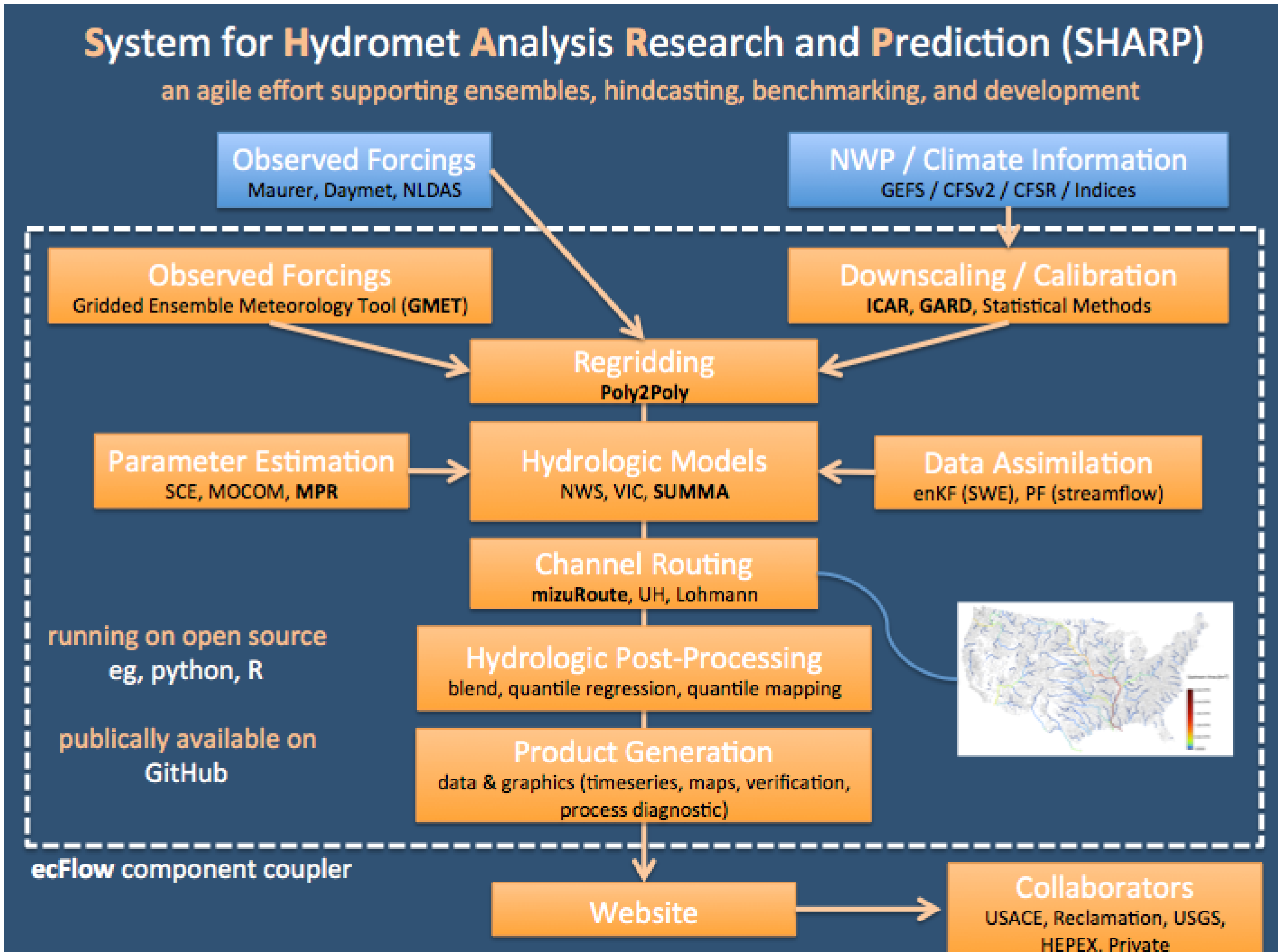
- Not adding uncertainty, but revealing uncertainty

Clark et al., WRR 2015; Clark et al., Current Climate Change Reports 2016

Advancing streamflow forecasting capabilities

NCAR UCAR

The SHARP system is now running at NCAR to generate real time short and seasonal range forecasts for a number of pilot case study basins



sample real-time workflow web monitor

SHARP System Status Report

Updated: Tue Dec 13 15:13:57 UTC 2016

Job	Submitted	Completed	Failed
get_ghcnd	14:00:00	14:16:35	
get_nwcc	14:00:01	14:05:17	
get_gefs	pending	pending	
get_cfsr	14:00:01	14:02:13	
get_flow	14:00:01	14:01:03	
reformat_ghcnd	14:16:36	14:33:58	
reformat_nwcc	14:33:59	14:34:12	
QC_stn_data	14:34:12	14:38:30	
fill_stn_data_pass1	14:38:31	15:13:56	
fill_stn_data_pass2	15:13:56	pending	
fill_stn_data_pass3	pending	pending	
fill_stn_data_pass4	pending	pending	
gen_ens	pending	pending	
grid2poly	pending	pending	
make_nws_forc	pending	pending	
run_nws_spinup	pending	pending	
downscale_gefs_fcst	pending	pending	
downscale_gefs_fcst_regr	pending	pending	
reformat_gard_output	pending	pending	
reformat_gard_output_regr	pending	pending	
met_forecast_grid2poly	pending	pending	
make_nws_met_forecast	pending	pending	
run_nws_gefs_fcst	pending	pending	
plot_stn_data_map	pending	pending	
plot_mr_fcst	pending	pending	

ecFlow -- https://software.ecmwf.int/wiki/display/ECFLOW/

Outline

NCAR UCAR

- Motivation
  - The nature of the hydrologic modeling problem
  - Beyond faith-based modeling?
- Modeling challenges
  - Processes
  - Parameters
  - Computing
- Research applications
  - Developing storylines of the future
  - Advancing streamflow forecasting capabilities
- Summary and research needs

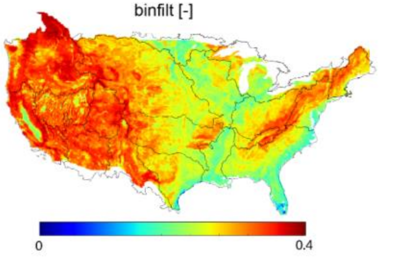
Summary

- 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

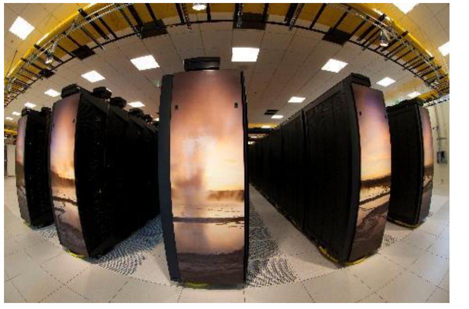
---

- 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

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.

Possible research directions for GWF

NCAR UCAR

1. Unify process-based land modeling across Canada (and beyond!)
   - Inter-component coupling (make use of legacy models)
   - Intra-component coupling (advance model construction)
2. Provide leadership 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. 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

Possible research directions for GWF

NCAR UCAR

6. Develop better continental-domain forcing data
   - Probabilistic approach to combine GEM/WRF, 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 is 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??