

# Challenges in continental-domain hydrologic modeling and prediction

**Martyn P. Clark**, University of Saskatchewan at Canmore

**Andrew W. Wood**, National Center for Atmospheric Research

*20 February, 2019*

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA

# Outline



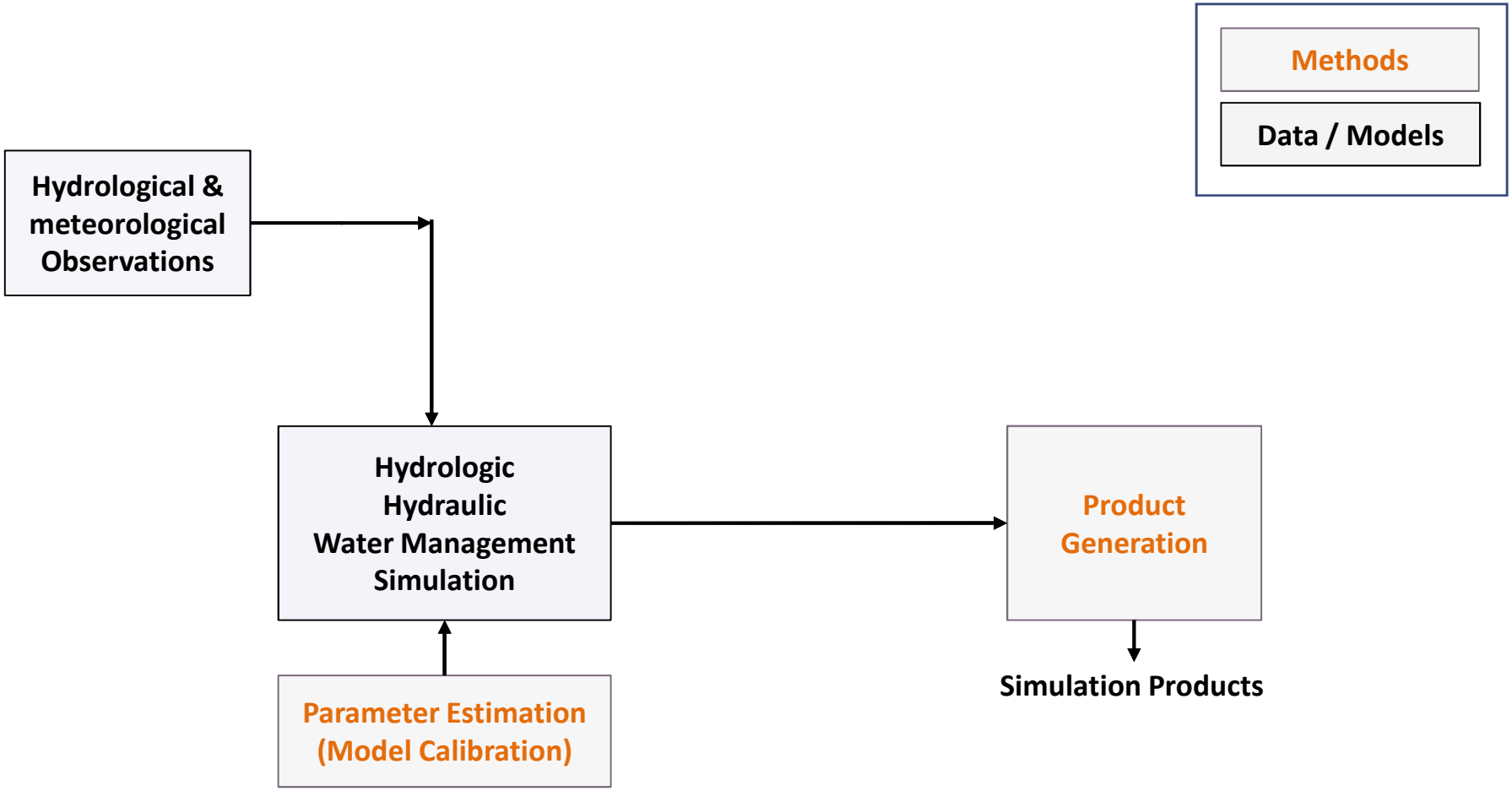
- Motivation
  - The hydrologic prediction challenge
  - System requirements

- Paradigms in hydrologic prediction
  - Traditional (forecaster in the loop)
  - Ensemble intermediate complexity
  - Hyper-resolution land modeling
  - Coupled prediction systems

- Challenges in continental-domain hydrologic modeling/prediction
  - Ensemble forcing
  - “Agile” modeling
  - Parameter estimation
  - Data assimilation

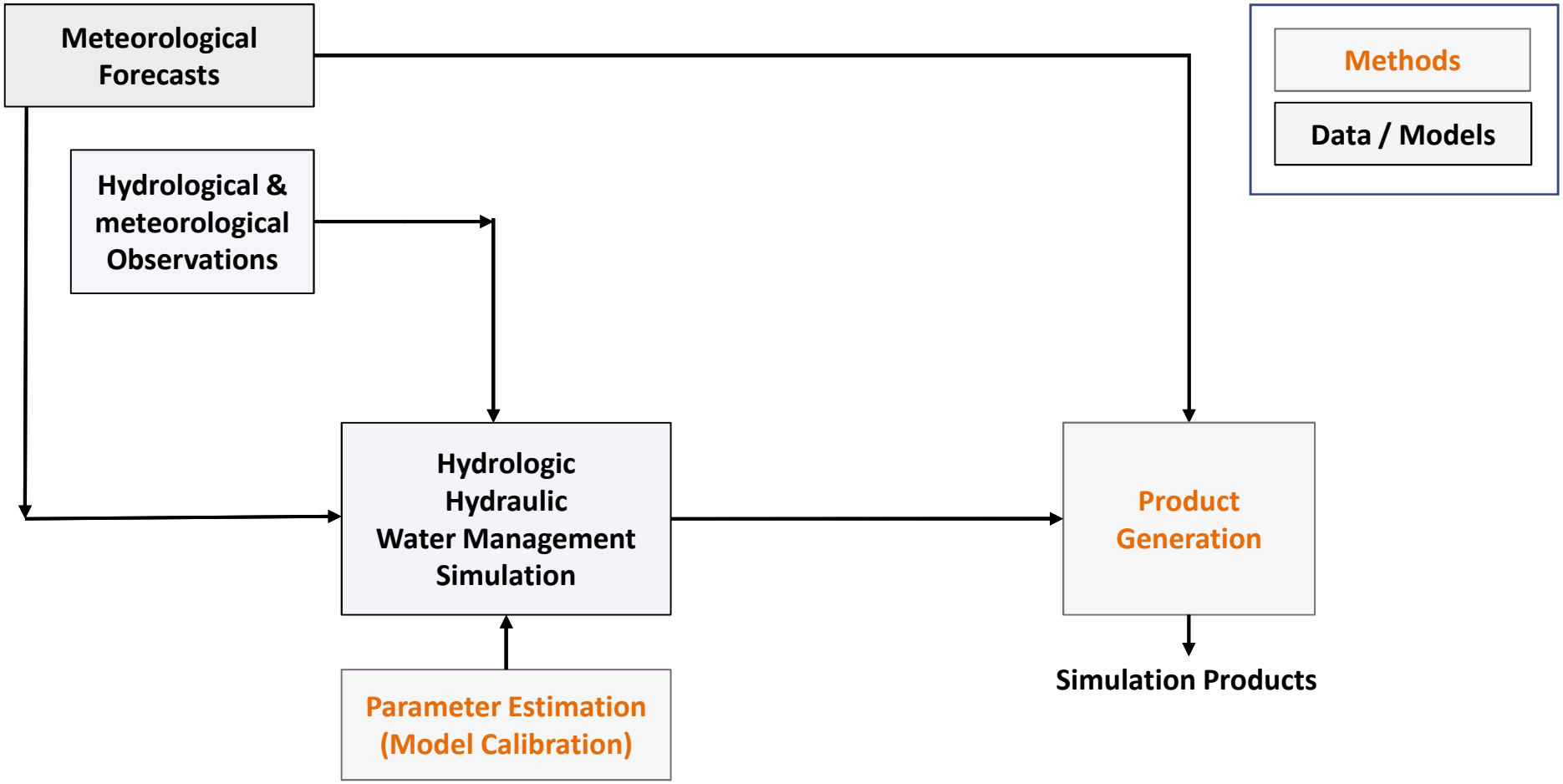
- Summary and research needs

# Broad community experience in modeling



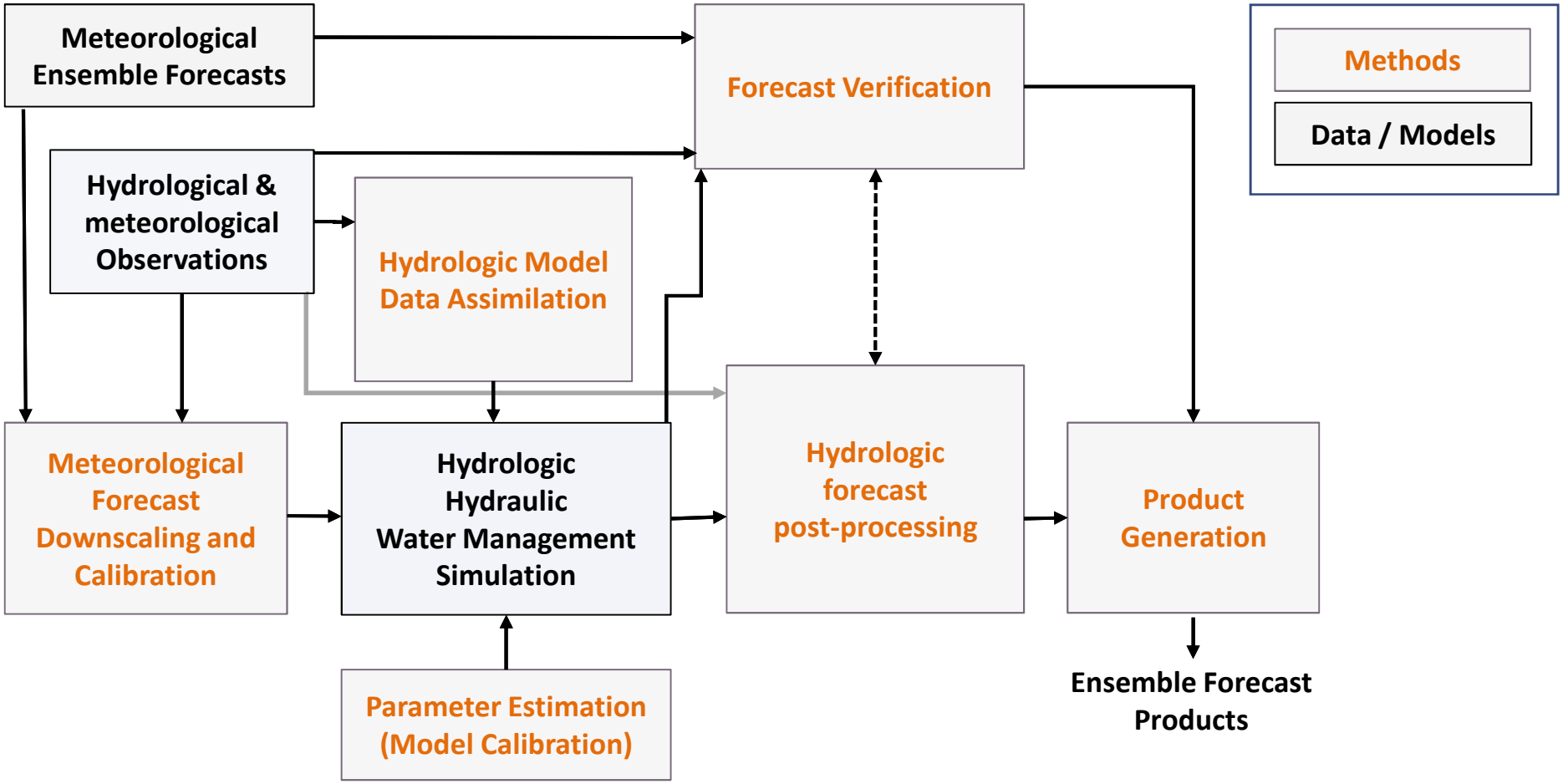

- Uncoupled systems – derive forcing meteorology from external sources to run hydrologic model

# Generating forecasts seems a trivial leap




- Obtain similar forcings from NWP to run through same model, creating forecasts

# Generating skillful forecasts is more complex




- Interdisciplinary
  - geosciences, statistics, computing and software engr., social science
- Not taught as a coherent program in any discipline

# System requirements depend on user needs



## Flash Flood / Short-range (hours to 2 days)

- Watch and warning programs
- Local emergency management activities
- Flood control and energy system management
- Water deliveries (irrigation)



## Medium-range (2-10 days)

- Reservoir management
- Local emergency management preparedness
- Coordinated regional operations
- Water deliveries (irrigation)

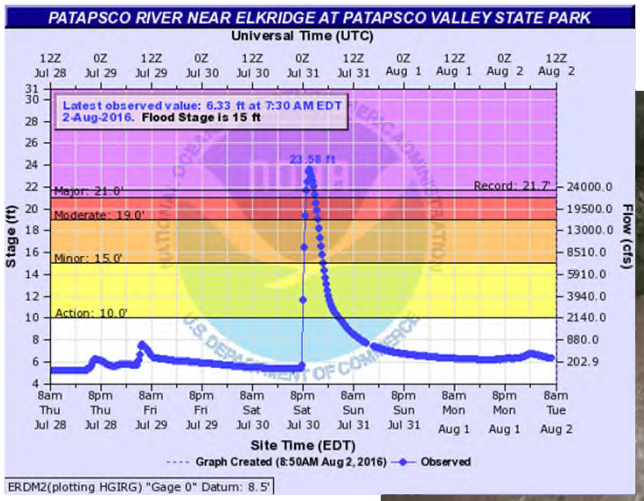


## Long-range (3 weeks to 12 months or longer)

- Water supply planning
- Reservoir system management
- Agriculture, Navigation, Energy, and other management applications

# Linking Uses and Science to Methodological Requirements

| Context | Uses | Potential Scientific Areas for Advancement | Implications for Methods -- Requirements |
|---|---|---|---|
| **Flash Flooding (0-24 hours)** | • Short lead warning at high resolution for communities and infrastructure | • Assimilation of inundated area imagery and reports (hydrologic and channel routing)<br>• Improvement of short range analyses and NWP (including ensembles)<br>• Post-processing based on retro simulations and hindcasting<br>• Improved routing model physics, eg channel representation, including losses | • High to hyper-resolution terrain routing is essential to product value<br>• Fast update cycling is needed<br>• Hydrologic model physics less critical, as primary initial conditions will be saturated<br>• Uncertainty quantification is desirable but not essential<br>• Long retrospectives / hindcasts less essential<br>• Absolute accuracy may be less important than large event detection capability |

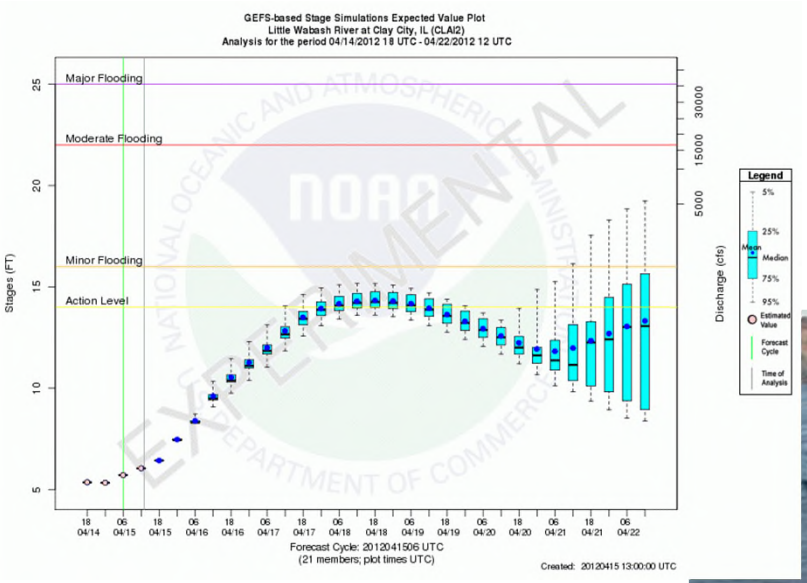


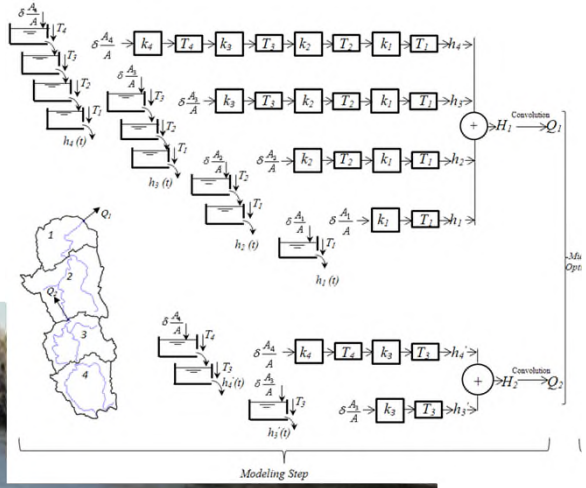





# Linking Uses and Science to Methodological Requirements

| Context | Uses | Potential Scientific Areas for Advancement | Implications for Methods -- Requirements |
|---|---|---|---|
| **River Prediction (0-15 days)** | • Operate water infrastructure for myriad uses -- flood protection, power, agricultural and municipal needs, nagivation etc.<br>• Support flood related emergency management | • Hydrologic model data assimilation (streamflow, snow, ponded water, etc.)<br>• Improved analyses and NWP (especially ensembles) and uncertainty quantification<br>• Improved model parameter estimation<br>• Post-processing to achieve unbiased, reliable forecasts<br>• Improved model physics to capture hydrologic variability | • Agility to support ensemble methods:<br>  • data assimilation<br>  • model calibration (1000s of trials)<br>  • ensemble prediction at longer leads<br>  • hindcasting (supports verification and post-processing)<br>• Increased complexity (resolution & physics) without sacrificing agility |

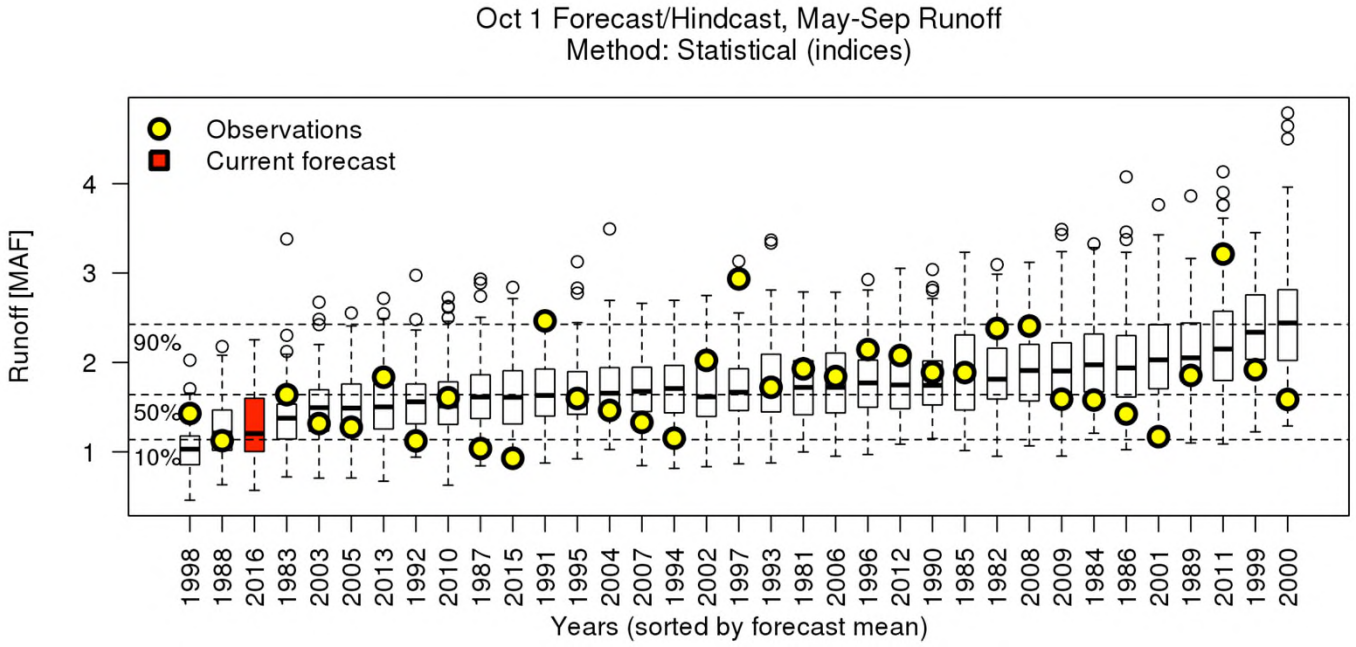





# Linking Uses and Science to Methodological Requirements

| Context | Uses | Potential Scientific Areas for Advancement | Implications for Methods -- Requirements |
|---|---|---|---|
| **Seasonal runoff prediction (0-12 months)** | • risk-based management of water infrastructure at specific control locations for myriad applications<br>• monthly to seasonal inflows to reservoirs and allocations to stakeholders | • Hydrologic model data assimilation (streamflow, snow)<br>• Improved analyses and climate forecasts (especially ensembles) and uncertainty quantification<br>• Improved model parameter estimation<br>• Post-processing and multi-model combination<br>• Improved model physics | • Forecast skill at control locations is overarching objective<br>• Risk quantification through reliable uncertainty estimates is required (i.e., ensembles)<br>• Forecast combination, verification and post-processing emphasis requires hindcasts<br>• Model calibration is critical because biases become exposed |




Hungry Horse Reservoir, MT



# Outline



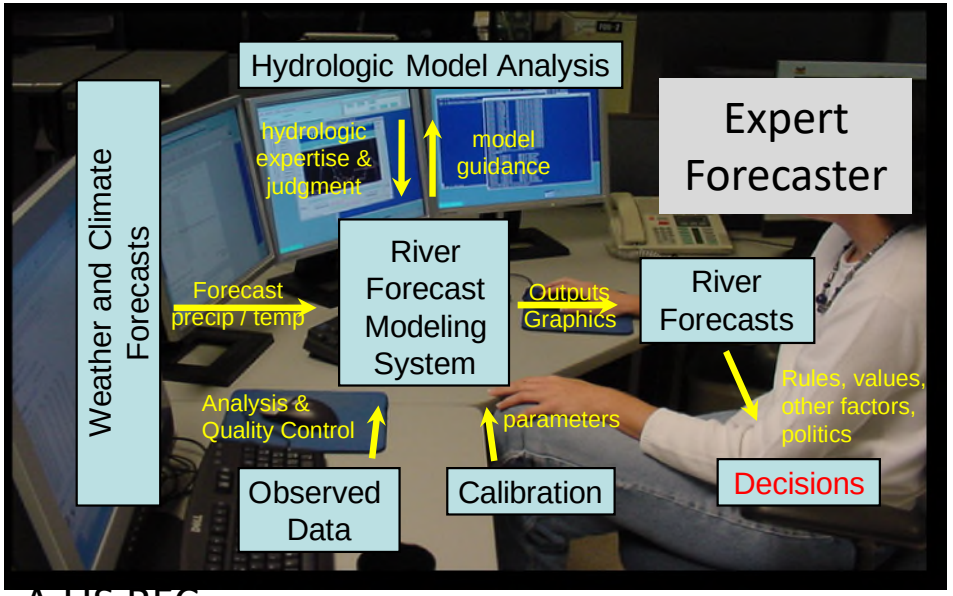
- Motivation
  - The hydrologic prediction challenge
  - System requirements
- Paradigms in hydrologic prediction
  - Traditional (forecaster in the loop)
  - Ensemble intermediate complexity
  - Hyper-resolution land modeling
  - Coupled prediction systems
- Challenges in continental-domain hydrologic modeling/prediction
  - Ensemble forcing
  - “Agile” modeling
  - Parameter estimation
  - Data assimilation
- Summary and research needs

# Traditional interactive (in-the-loop) paradigm

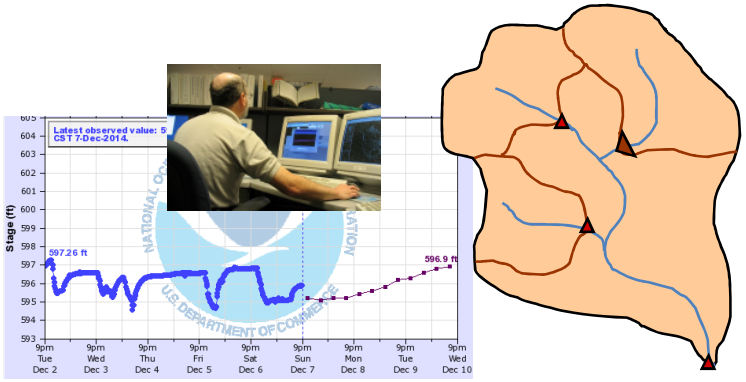


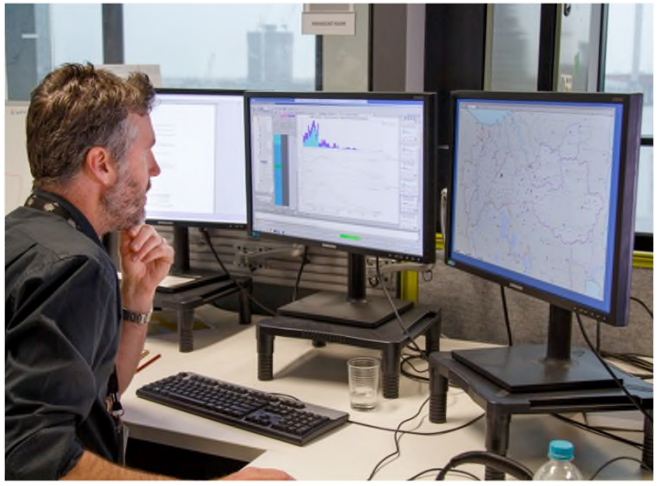
**Operational, 1980s => present**

- Provide tailored, mostly deterministic forecasts that are inputs to water management
- Conceptual models that can be adjusted manually, in real-time
- Heavy use of manual model calibration
- Reliance on human expertise at the model/data system level
- Non-reproducible, non-scalable forecast process
- Run on workstations




A US RFC

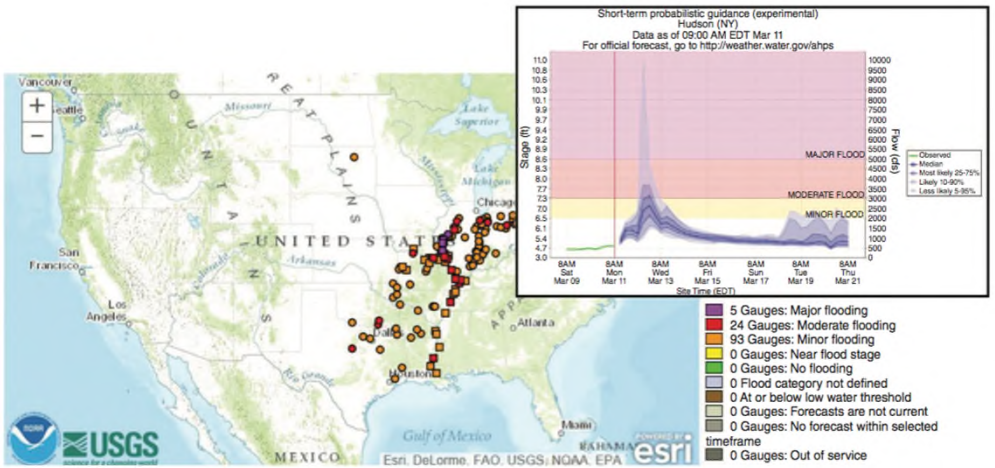


An Australian RFC

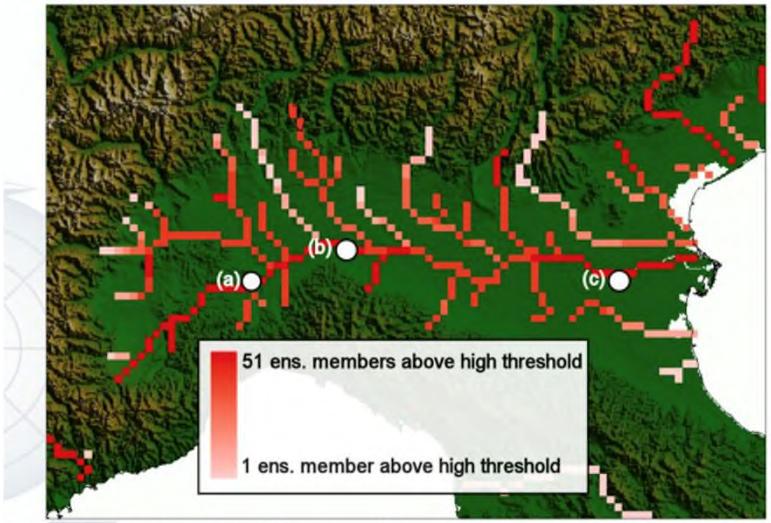
**Hurdles: scalability, reproducibility**

# Ensemble intermediate-scale paradigm

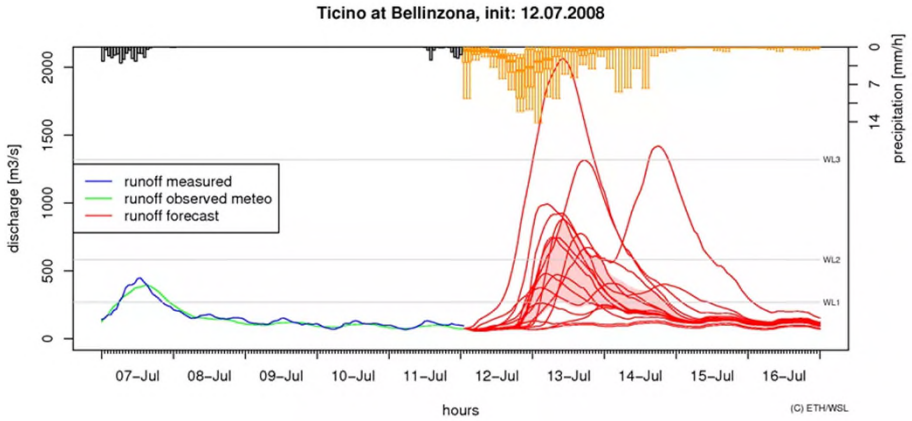



*(from Emerton et al, 2017 – on HEFS)*

- Uncertainty quantification to support risk-based decisions
  - semi-agile intermediate scale modeling
  - forecaster interaction at warning decision
  - allow for calibration, DA, hindcasting, ensembles
- **Hurdle**: Resolution and complexity limited by need for agility

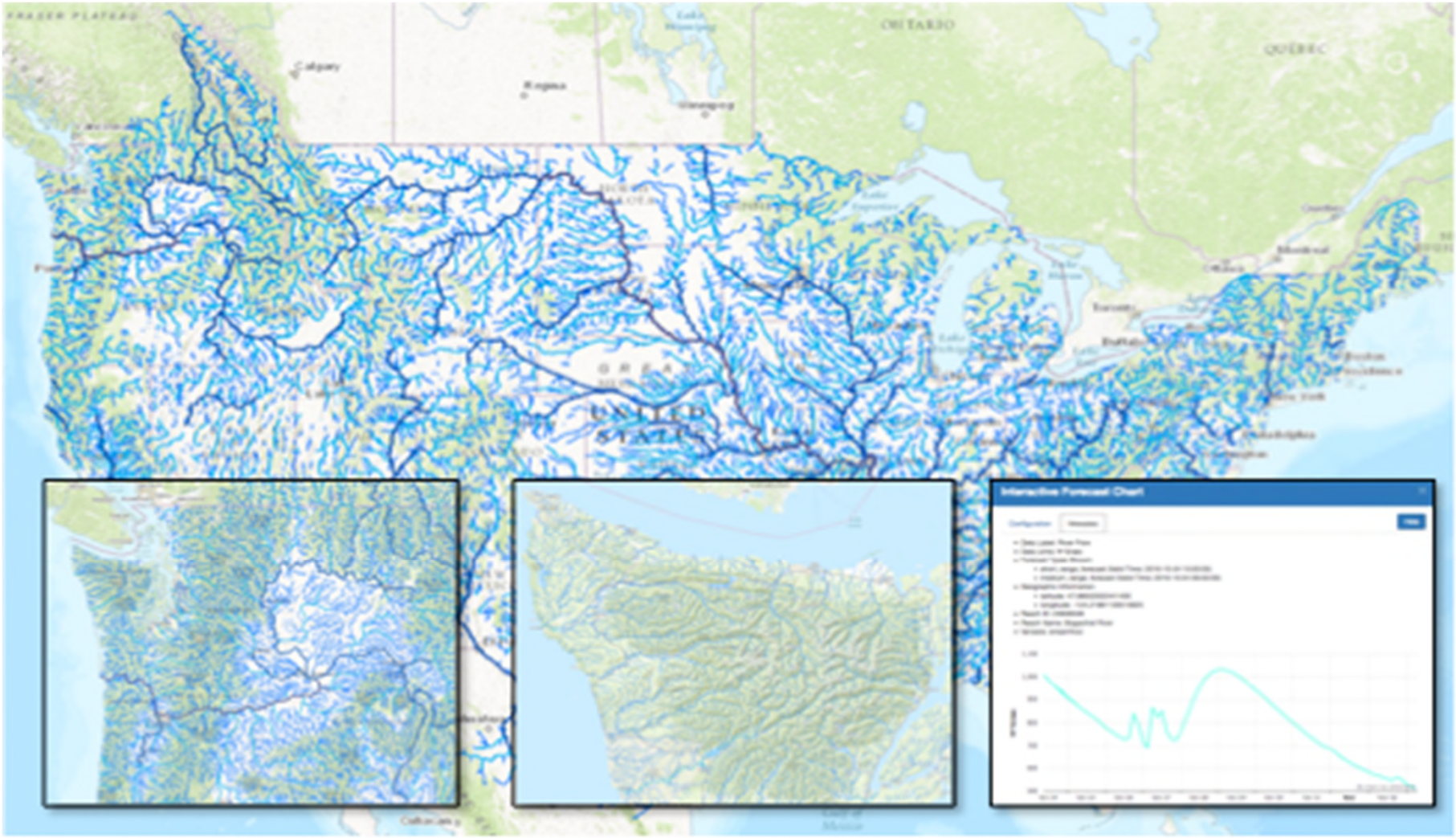



***Number of ensemble members based on the ECMWF EPS forecast from the 24th of April 00 UTC simulating discharges which exceed the EFAS high alert level for the Po river basin. The circles denote the locations of reporting points.***



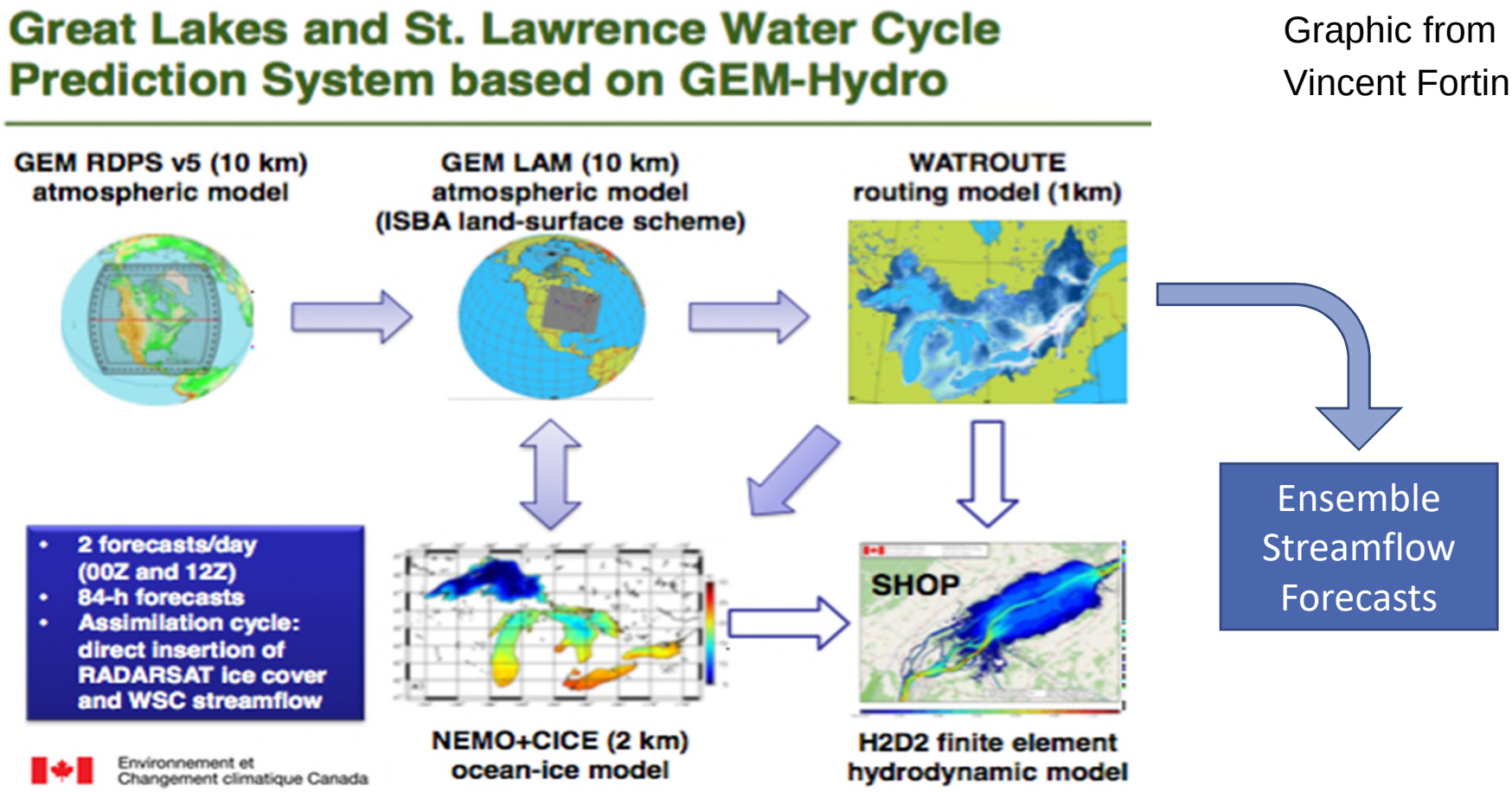

*(Massimiliano Zappa and Kaethi Liechti, from hepex.org)*

# Hyper-resolution Monitoring & Prediction – National Water Model




- 250m resol.: output for 2.7 million stream reaches, hourly timestep, deterministic
- Philosophy that higher resolution and physics reduce the need for parameter estimation
- **Hurdles**: calibration, ensembles, data assimilation, verification, longer-range forecasts

# A *Coupled* NWP-based Hydrologic Forecast System -- GEM-Hydro

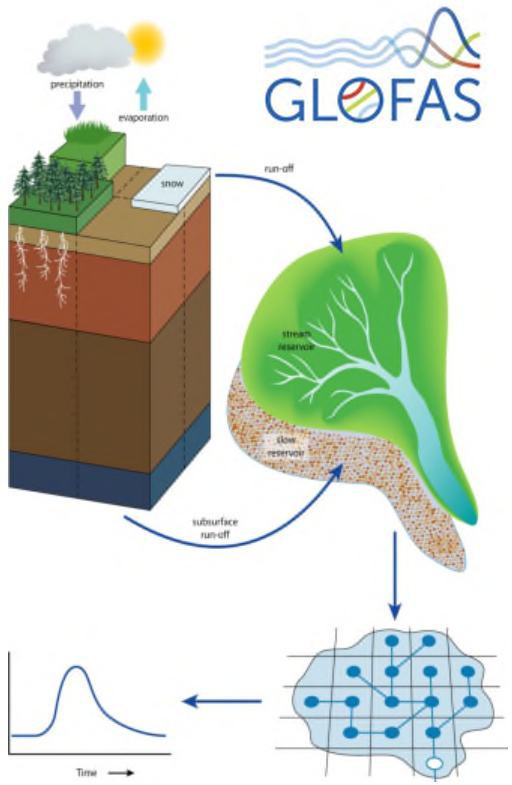



Graphic from Vincent Fortin

- No offline hydrologic forecasting component
- Intermediate scale LSM (10 km) with high resolution channel routing (1 km)
- Consistent model states and fluxes (precipitation, evaporation, runoff)
- **Hurdles**: more complicated to develop/improve (must fit NWP system R&D priorities)
- see also: Snow et al, 2016, JAWRA (GLOFAS runoff + RAPID routing)

# The Global Flood Awareness System (ECMWF)






Output from global NWP land-surface scheme ensemble forecast: HTESSEL (ECMWF)

-Surface heat & evaporation
-Soil water budget
Output: surface flux & subsurface flux

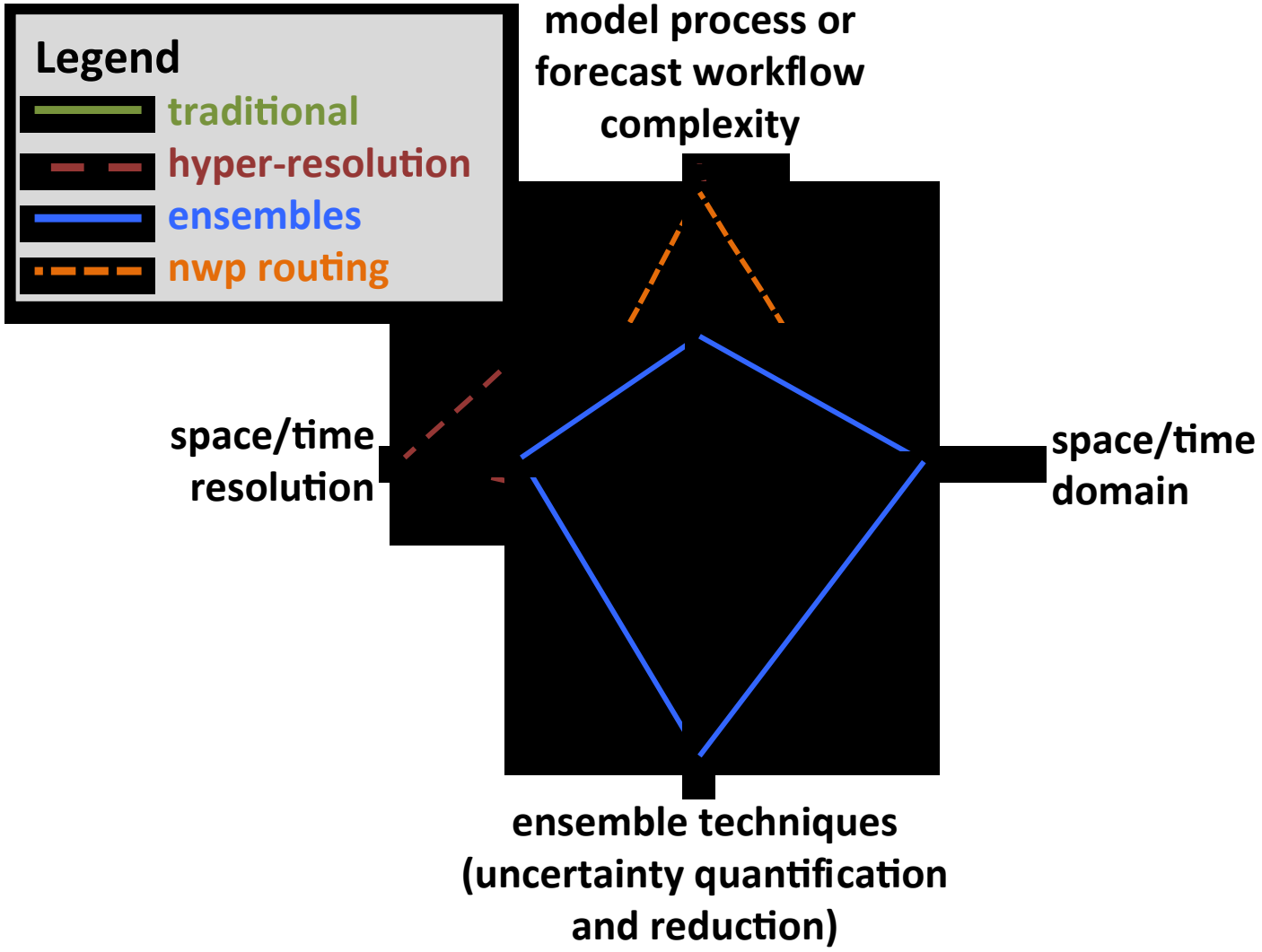
Routing models
Simplified LISFLOOD (JRC) / CAMAFLOOD

Post-processing for end users

Made possible by the <u>archiving of hydrological variables</u> from forecasts and reanalysis at ECMWF

# System Design Tradeoffs

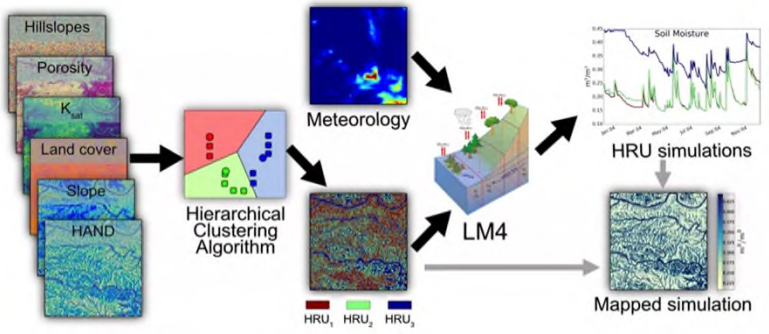
Given Fixed Resources…



# Innovative strategies for managing tradeoffs

## Application of principles of hydrologic similarity

- avoid explicit simulation of all high resolution model pixels
- identify major modes of watershed variability, simulate those as point processes and map back to their watershed locations
  - Chaney et al, Hydrol. Proc. (2016)
  - Newman et al, JHM (2014)

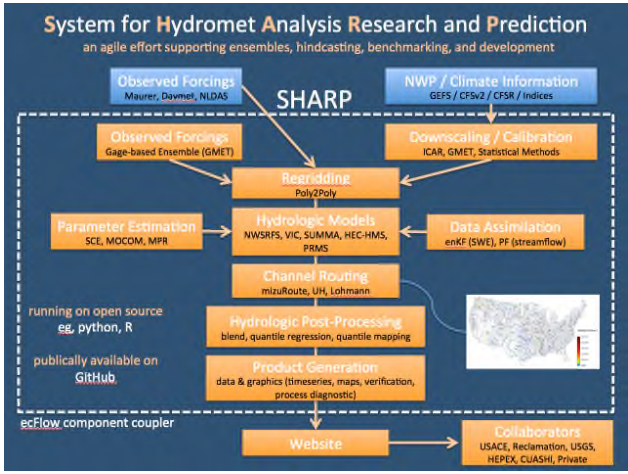



## Leveraging multi-resolution systems

- ex. 1 – ensemble members at both high resolution and lower resolution to see both high resolution patterns and uncertainty represented in larger/low resolution ensemble
- ex. 2 – switch from fine resolution for short lead times to lower resolution for longer leads
  - similar to many NWP systems

## Creation of real-time streamflow forecast testbeds

- value of demonstration in operational context
- data/latency/reliability constraints alter feasibility of methods
- facilitate comparison with existing approaches



# Outline

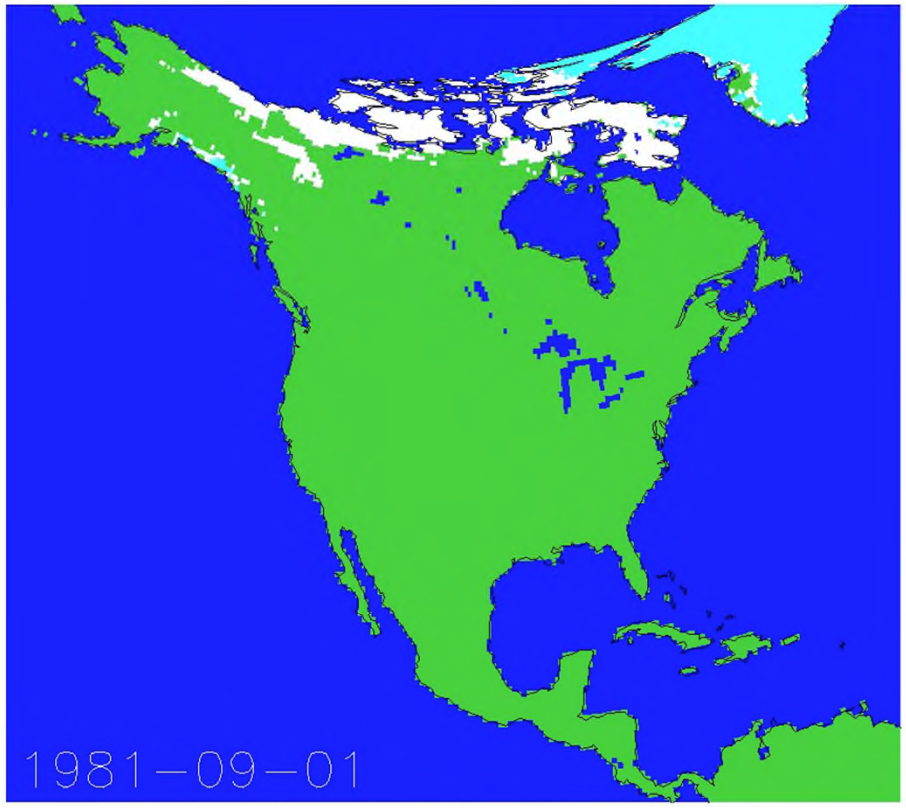


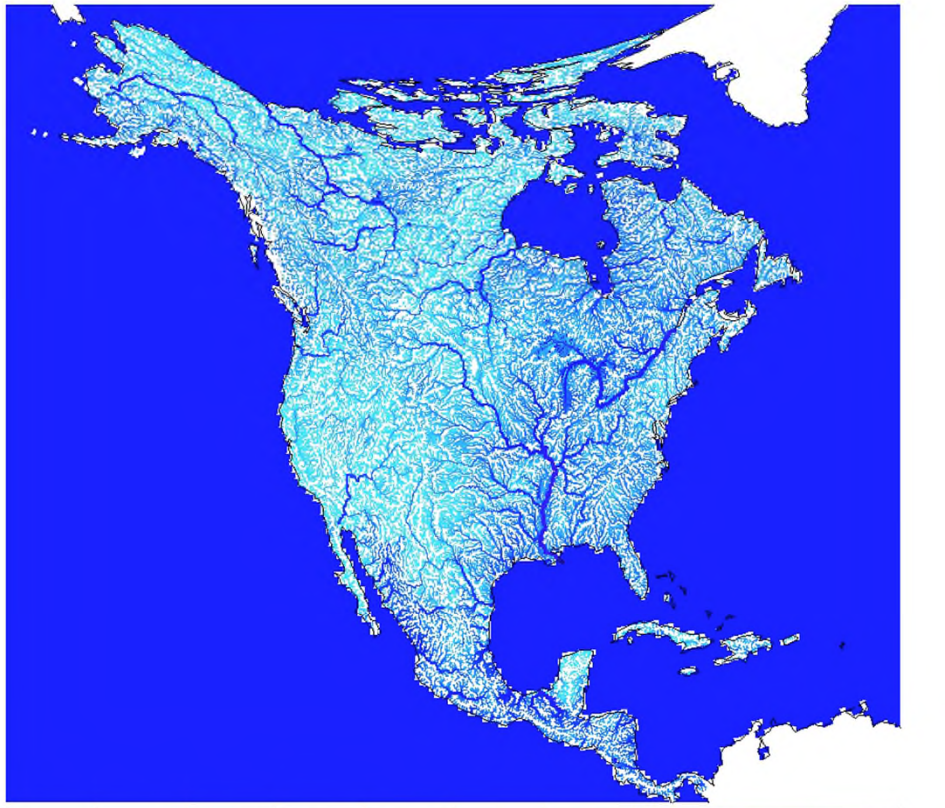
- Motivation
  - The hydrologic prediction challenge
  - System requirements
- Paradigms in hydrologic prediction
  - Traditional (forecaster in the loop)
  - Ensemble intermediate complexity
  - Hyper-resolution land modeling
  - Coupled prediction systems
- Challenges in continental-domain hydrologic modeling/prediction
  - Ensemble forcing
  - “Agile” modeling
  - Parameter estimation
  - Data assimilation
- Summary and research needs

# Continental-domain modeling

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA

- Advance simulations of hydrologic processes across North America

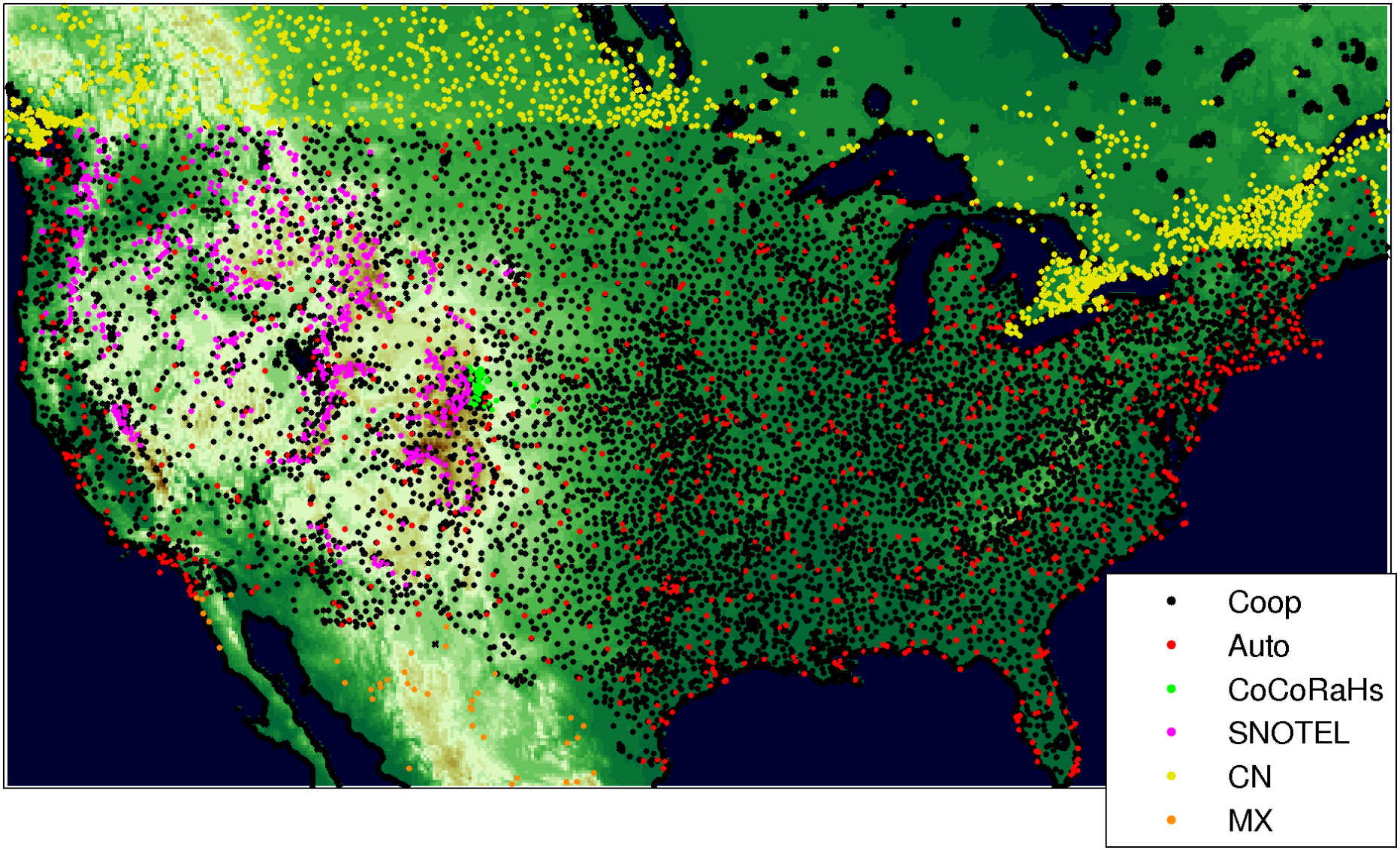


# Ensemble spatial met. fields

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA

## Spatial extrapolation from 12,000+ stations across the CONUS



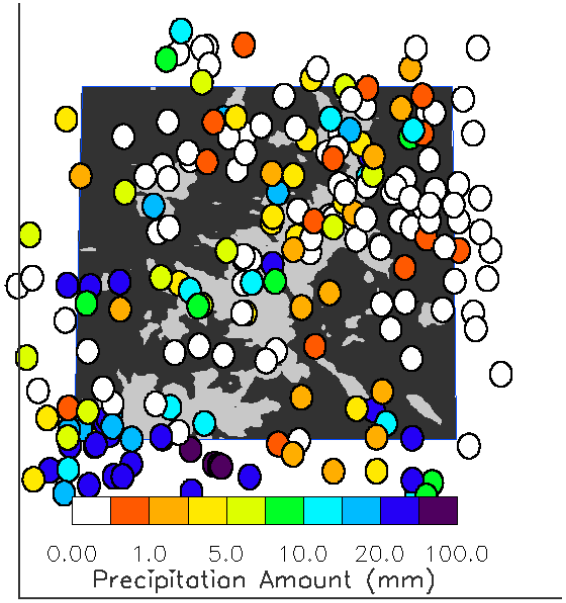
# Ensemble Generation: Generating grids

Step 1:

Locally weighted regression at each grid cell:

- Probability of Precipitation via logistic regression
- Amount and uncertainty (least squares mean & residuals)

observations

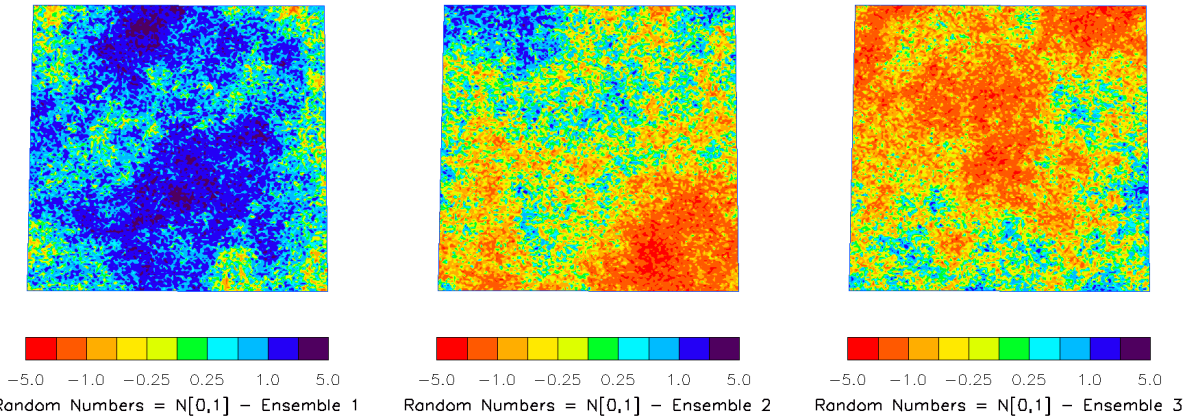



*Example over the Colorado Headwaters*

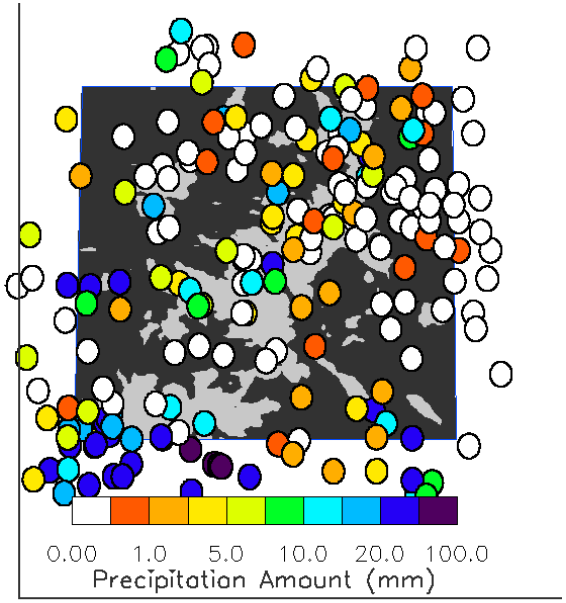
# Ensemble Generation: Generating grids

Step 2:
Synthesize ensembles using spatially correlated random fields

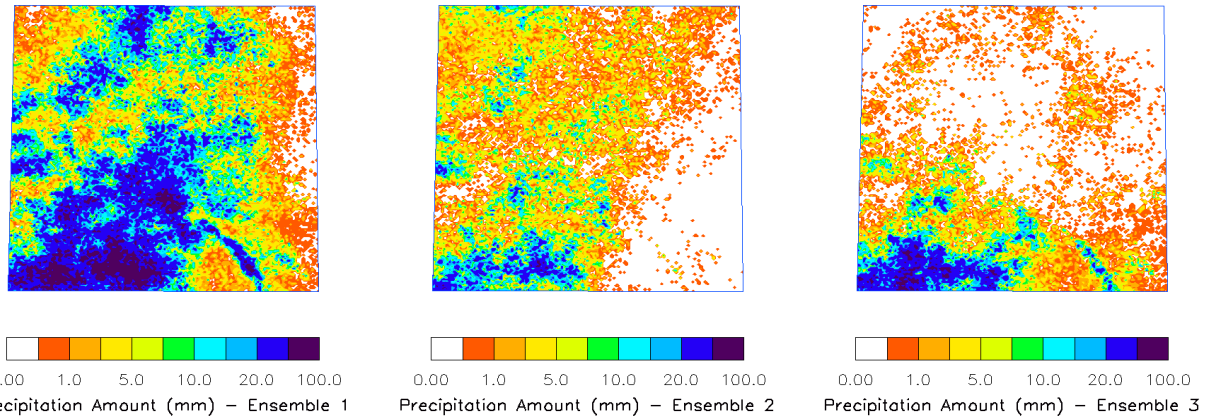
Spatially Correlated Random Fields: examples




observations




Ensemble member: examples



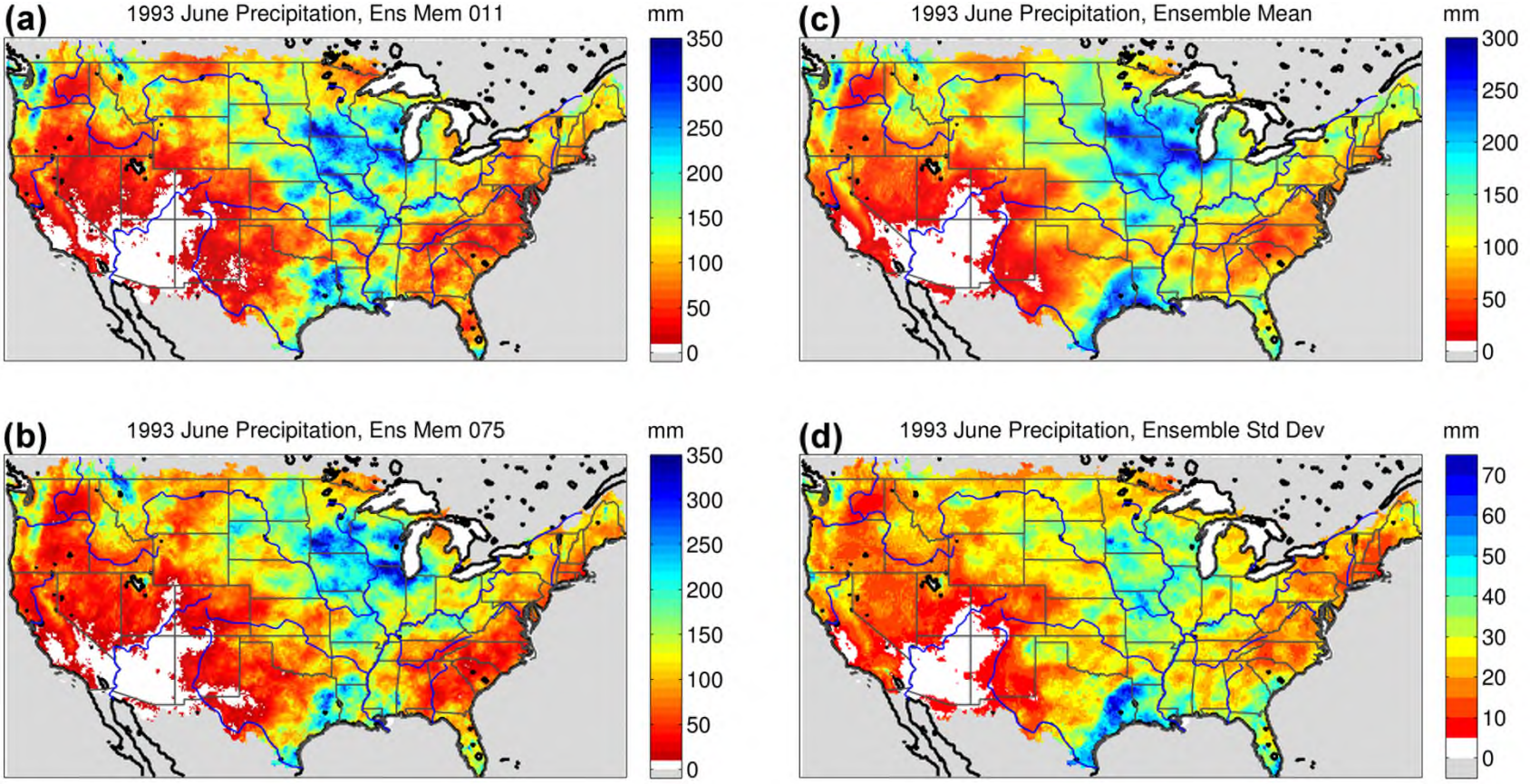

*Example over the Colorado Headwaters*

Clark & Slater (2006), Newman et al. (2015, JHM)

# CONUS product

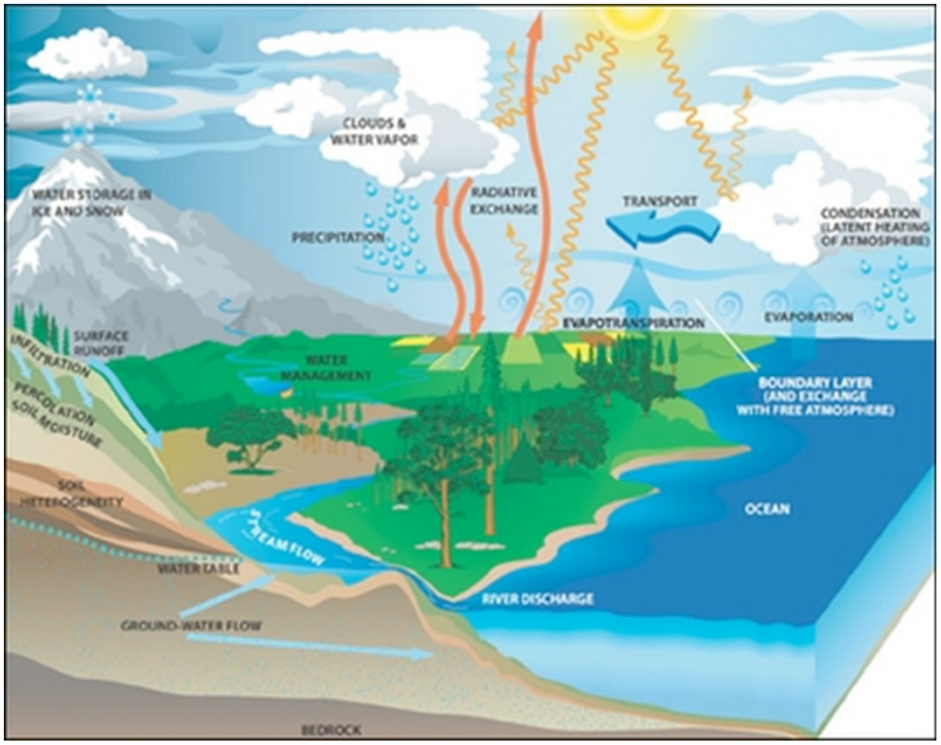
- Dataset constructed from 1980-2012
- Daily spatial fields of precipitation and temperature
- Dataset freely available

*Example output for June 1993*




Newman et al., JHM 2015

# Agile modeling




*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

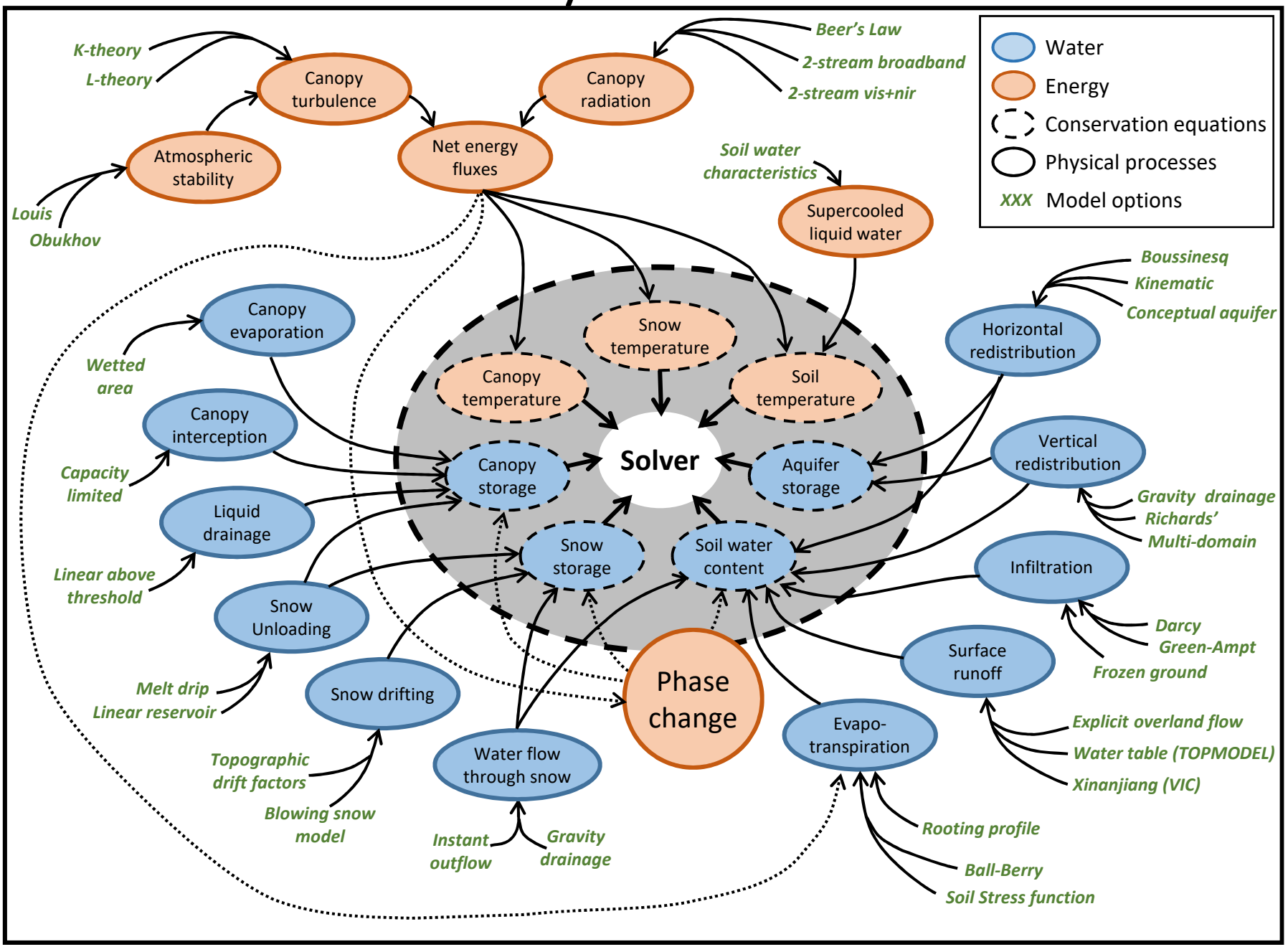
2. Differences among models relate to
   a) the spatial discretization of the model domain;
   b) the approaches used to parameterize individual fluxes (including model parameter values); and
   c) the methods used to solve the governing model equations.

***The Structure for Unifying Multiple Modeling Alternatives (SUMMA):***

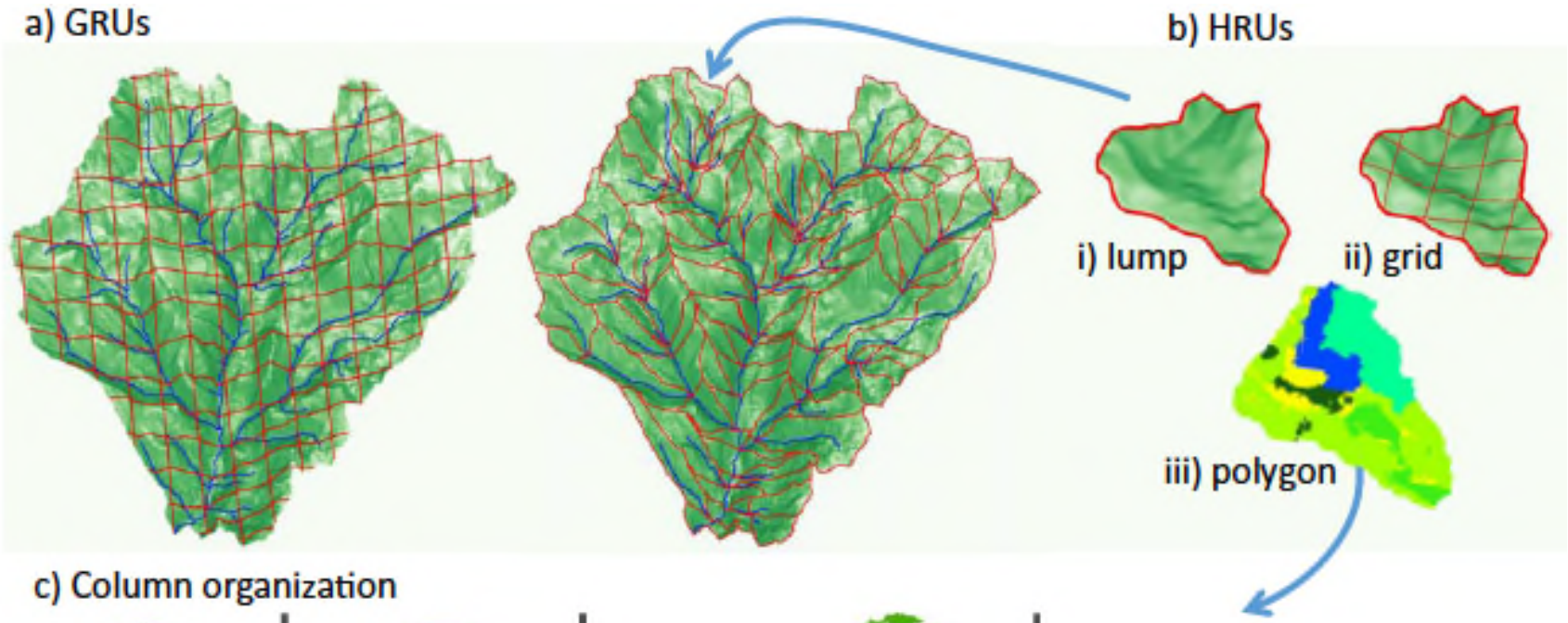
Defines a single set of conservation equations for land biogeophysics, with the capability to use different spatial discretizations, different flux parameterizations and model parameters, & different time stepping schemes

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

# Process flexibility



# Spatial flexibility

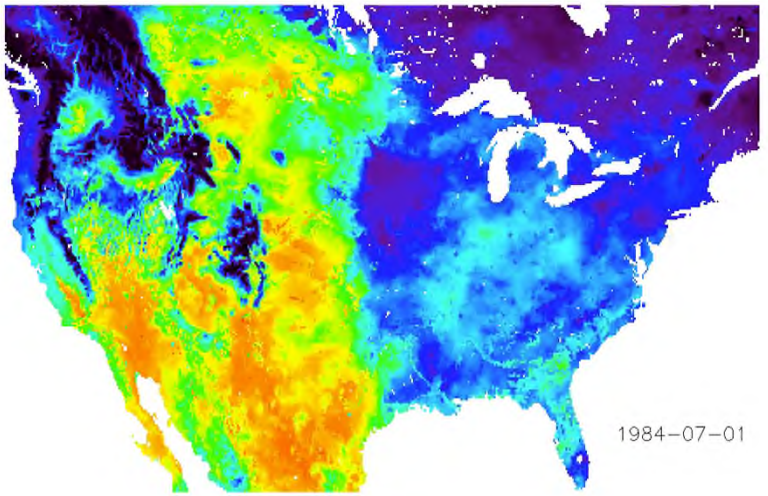


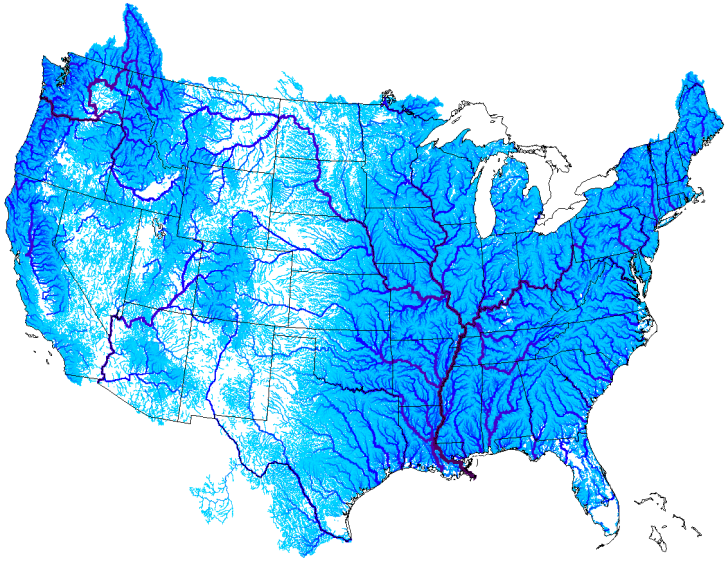
# Use cases



- 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)**



# Understanding complexity

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA

- 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*

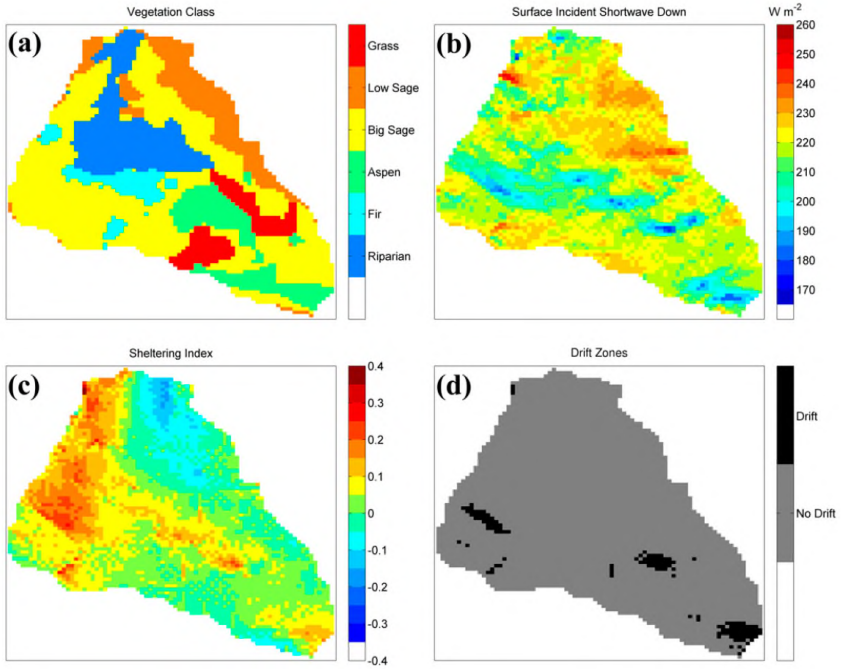
- Process-based hydrologic models
  - Explicitly represent dominant hydrologic and biophysical processes; explicitly represent details of the landscape
  - Reliance on geophysical data to estimate model parameters and widespread use of spatially constant parameters obtained from limited experimental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

# Understanding complexity

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

- Possible 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

# Model parameters

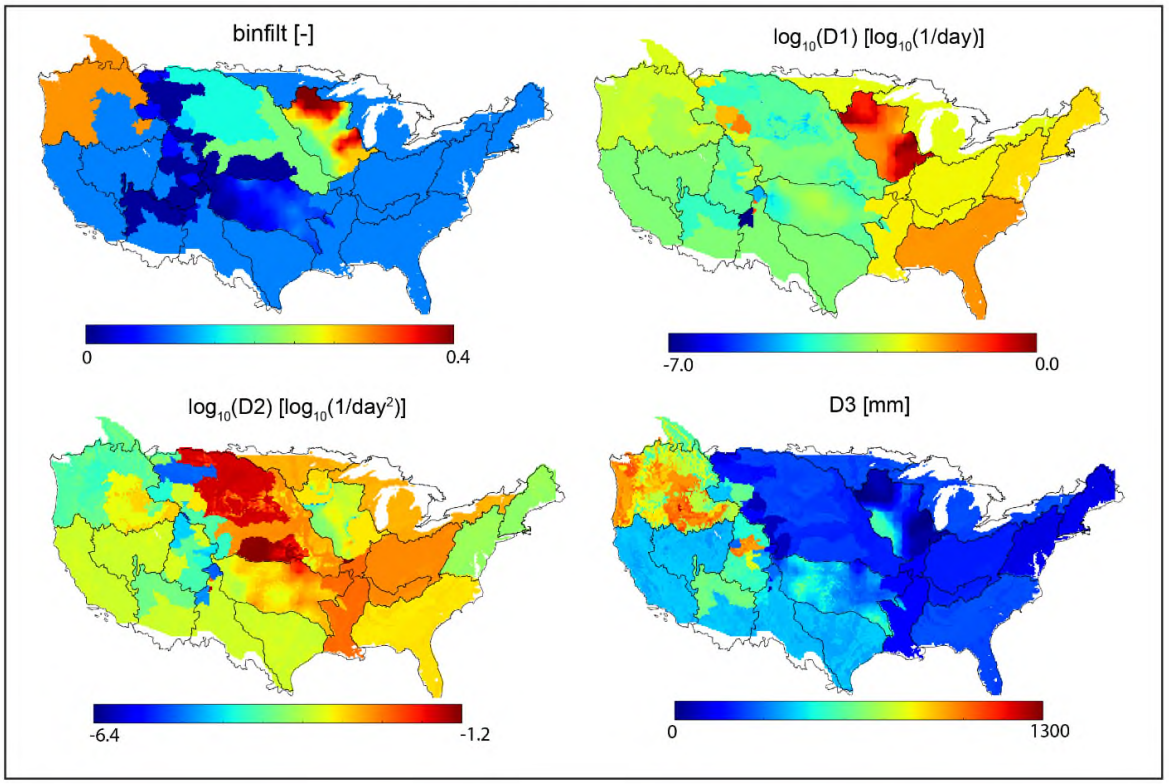
- Lack of knowledge of model parameters
  - Vegetation and soils datasets do not have sufficient resolution and information content
    - *Same soil type across large areas (assume no heterogeneity)*
    - *Often limited information on hydraulic properties necessary to simulate heterogeneous hydrologic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

# Default params

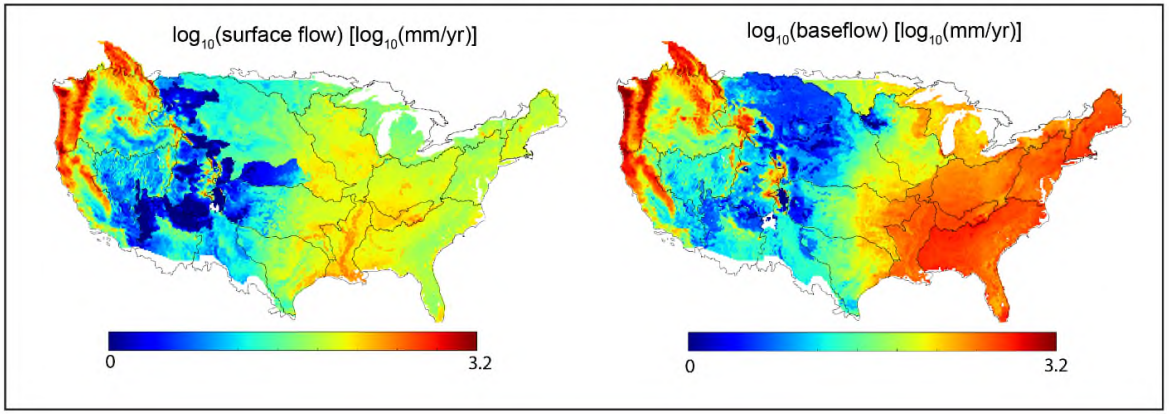
- Spatial discontinuities in model parameters

- Spatial discontinuities in model simulations

**VIC Soil parameters – CMIP5 default**




**1950-1999 annual mean runoff**



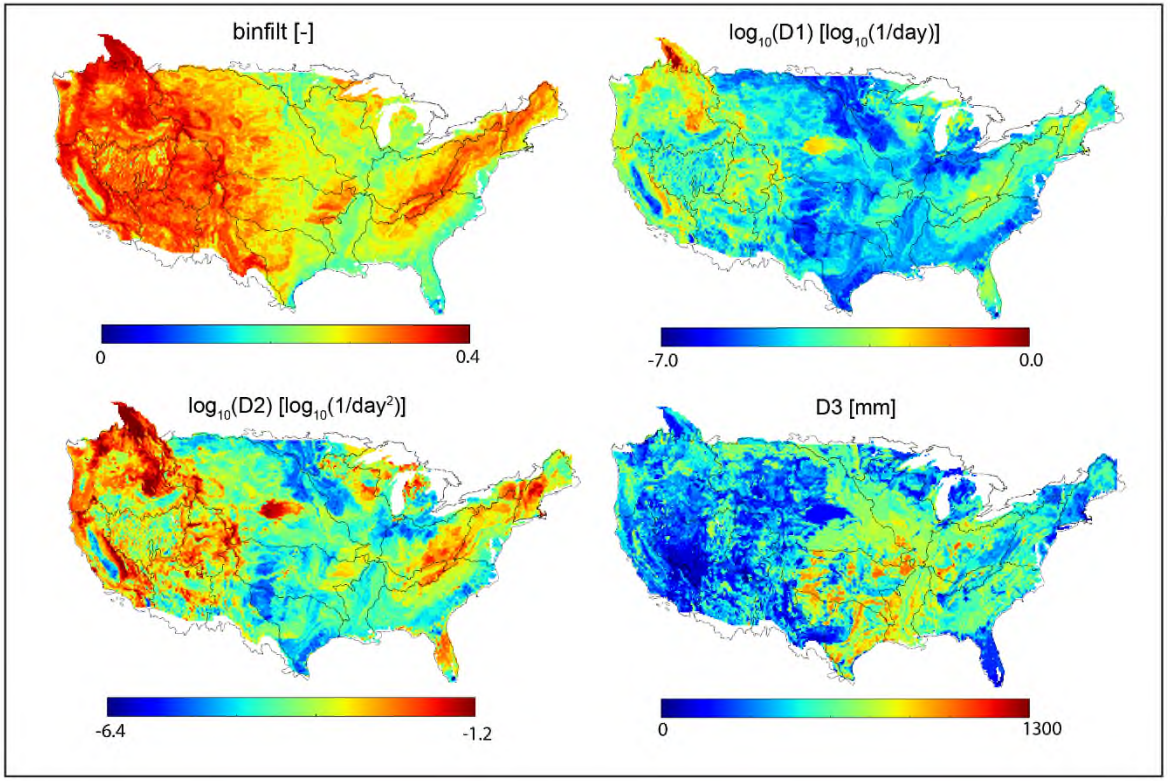

Mizukami et al., WRR 2017

# MPR-flex

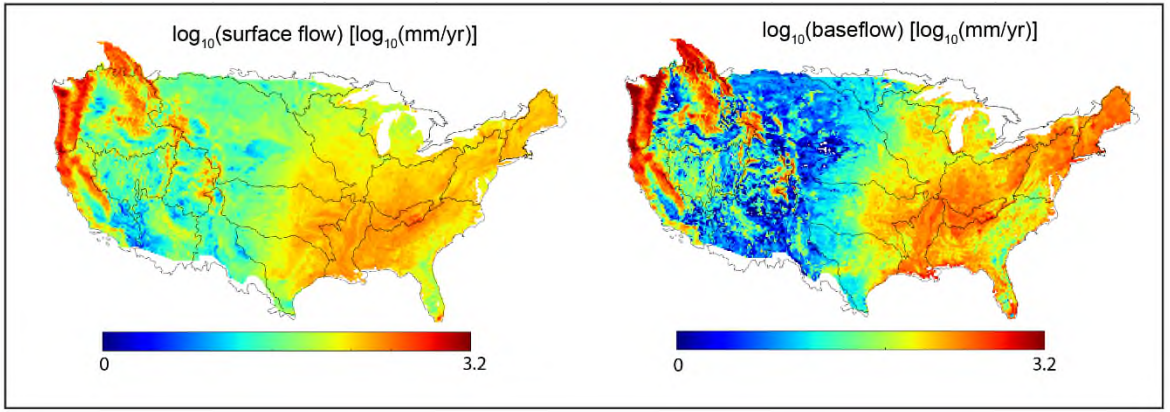
- Modify coefficients in transfer functions that relate physical attributes (soil, veg, topography) to model parameters
- Use parameter-specific upscaling operators to represent multi-scale behavior
- *Define transfer functions for new models – develop model agnostic MPR (MPR-Flex)*

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

**VIC Soil parameters – MPR**




**1950-1999 annual mean runoff**

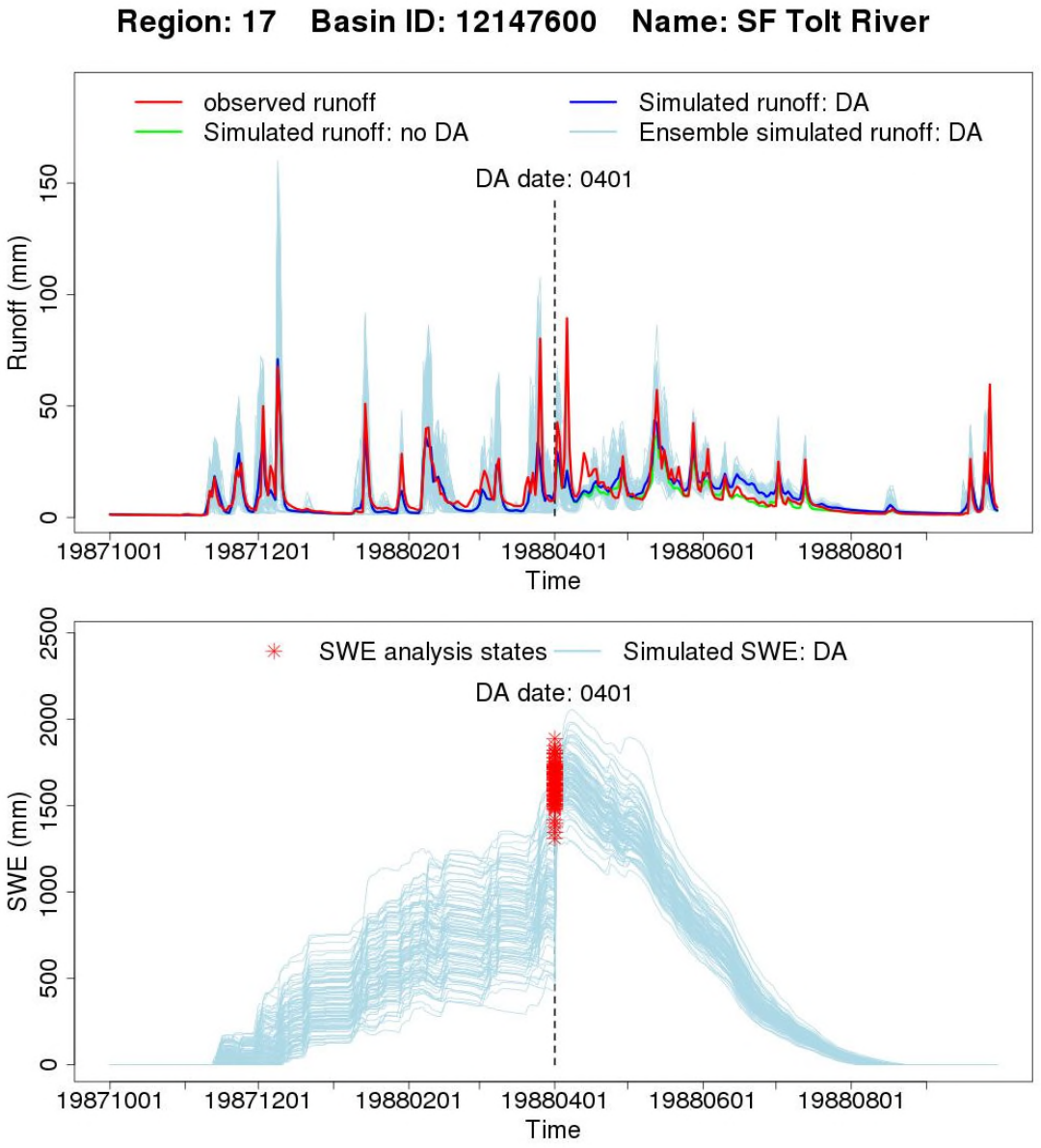



Mizukami et al., WRR 2017

# Hydrologic DA for seasonal prediction

Example

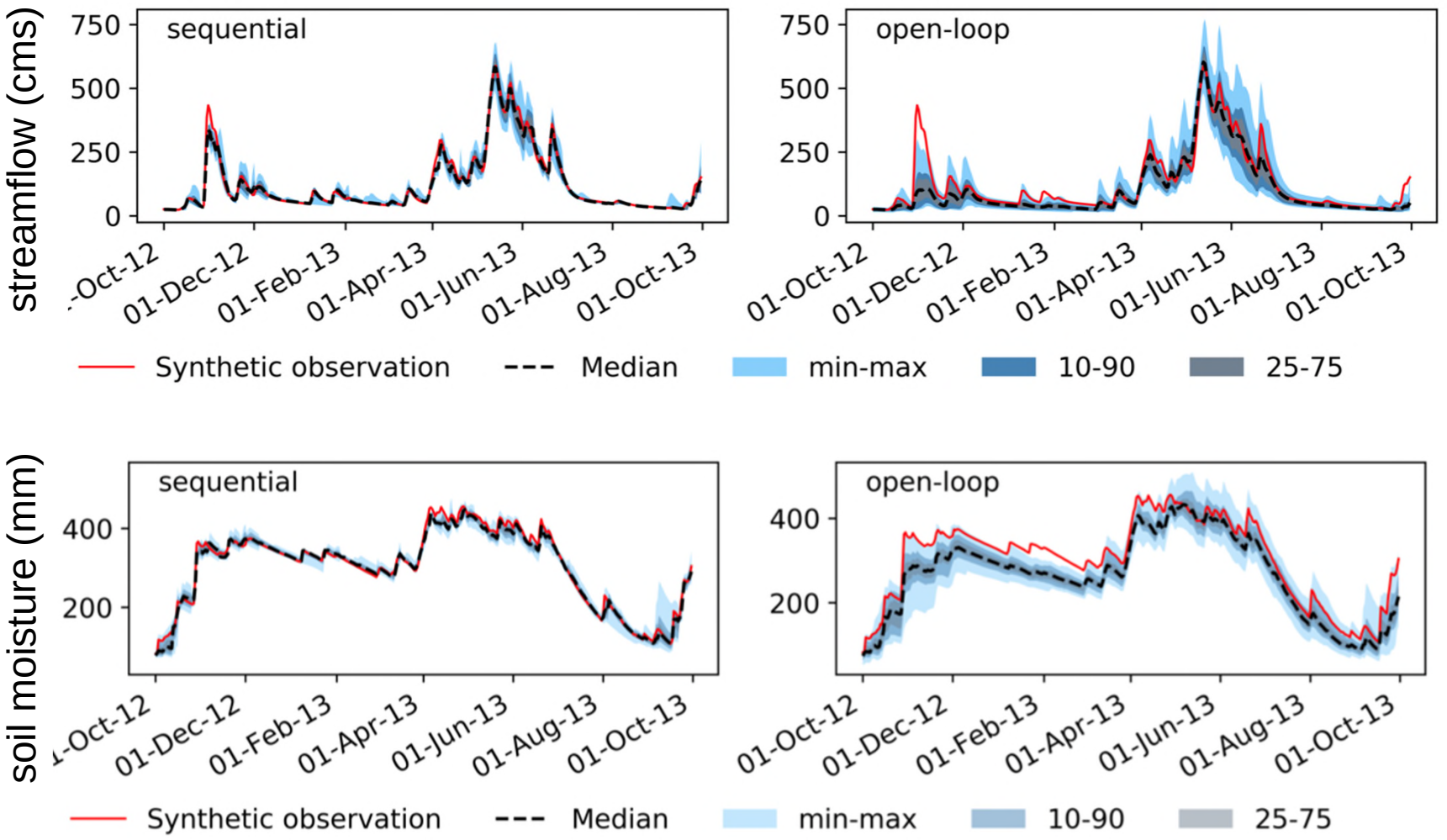
- Use an ensemble method to estimate initial conditions
- Update those conditions with SWE observations
- Make ensemble predictions from mean model states
- Assess forecast skill after assimilation



# Hydrologic DA for seasonal prediction

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA

## Particle Filter DA (Sequential Importance Resampling) improves both flow and SM

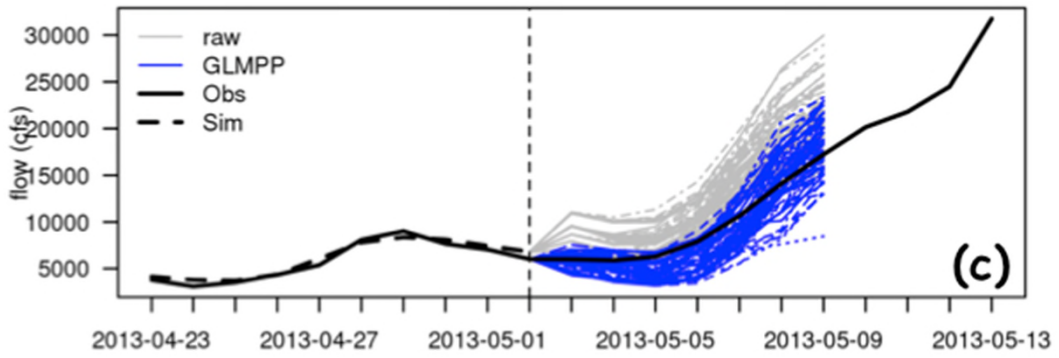



*Elizabeth Clark, A. Wood, B. Nijssen, M. Clark, and others*

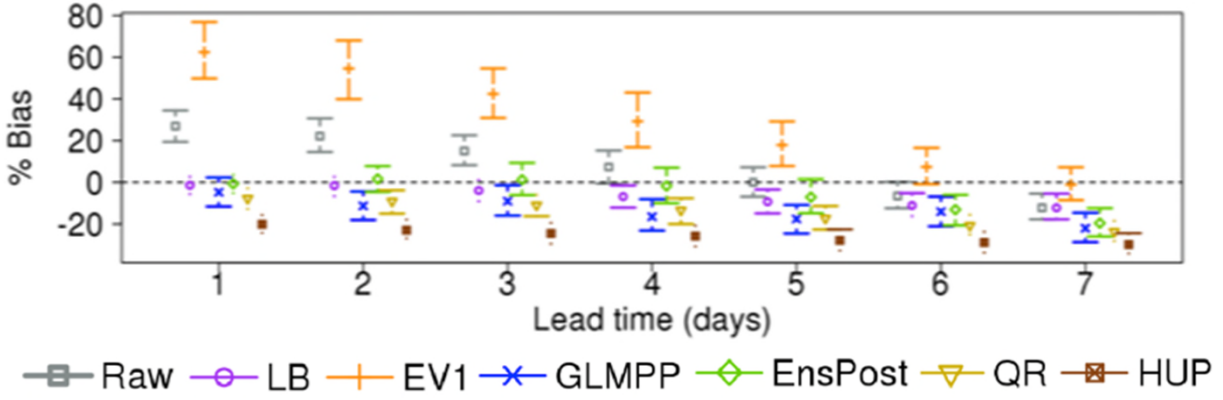
# Post-processing is also critical

- reduces residual uncertainty after other parts of the process

| Method | Name |
|---|---|
| **LB** | Linear blending |
| **EV1** | Error in variable Model Output Statistics with one variable |
| **GLMPP** | Generalized Linear Model Post-Processor |
| **EnsPost** | Ensemble Post-Processor |
| **QR** | Quantile Regression |
| **HUP** | Hydrologic Uncertainty Processor |



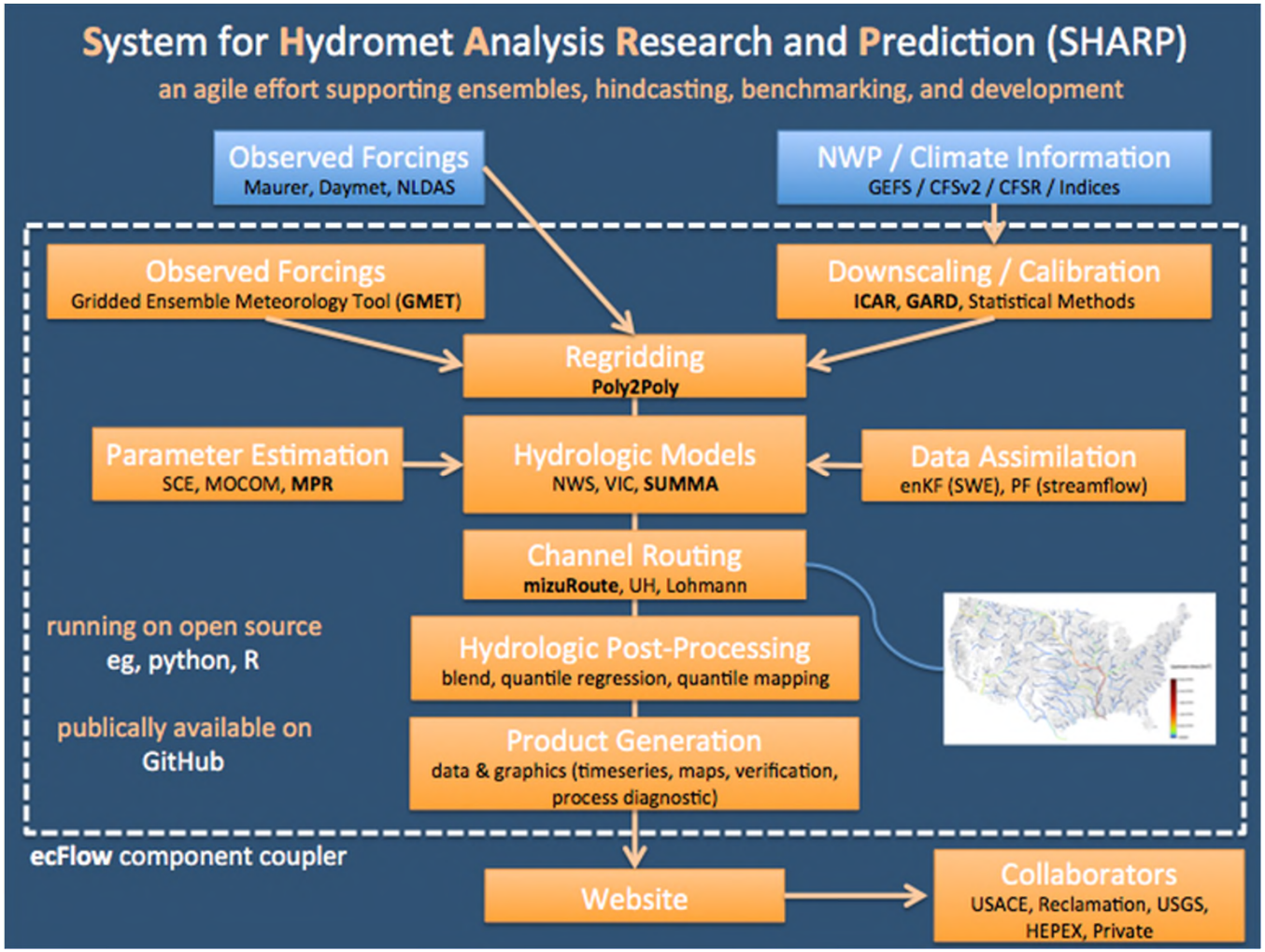

hindcast based skill evaluation




contribution by Pablo Mendoza

# SHARP Testbed forecasting system

We have developed an operational system to assess methods for real-time short and seasonal range forecasts




**workflow web monitor**

**SHARP System Status Report**

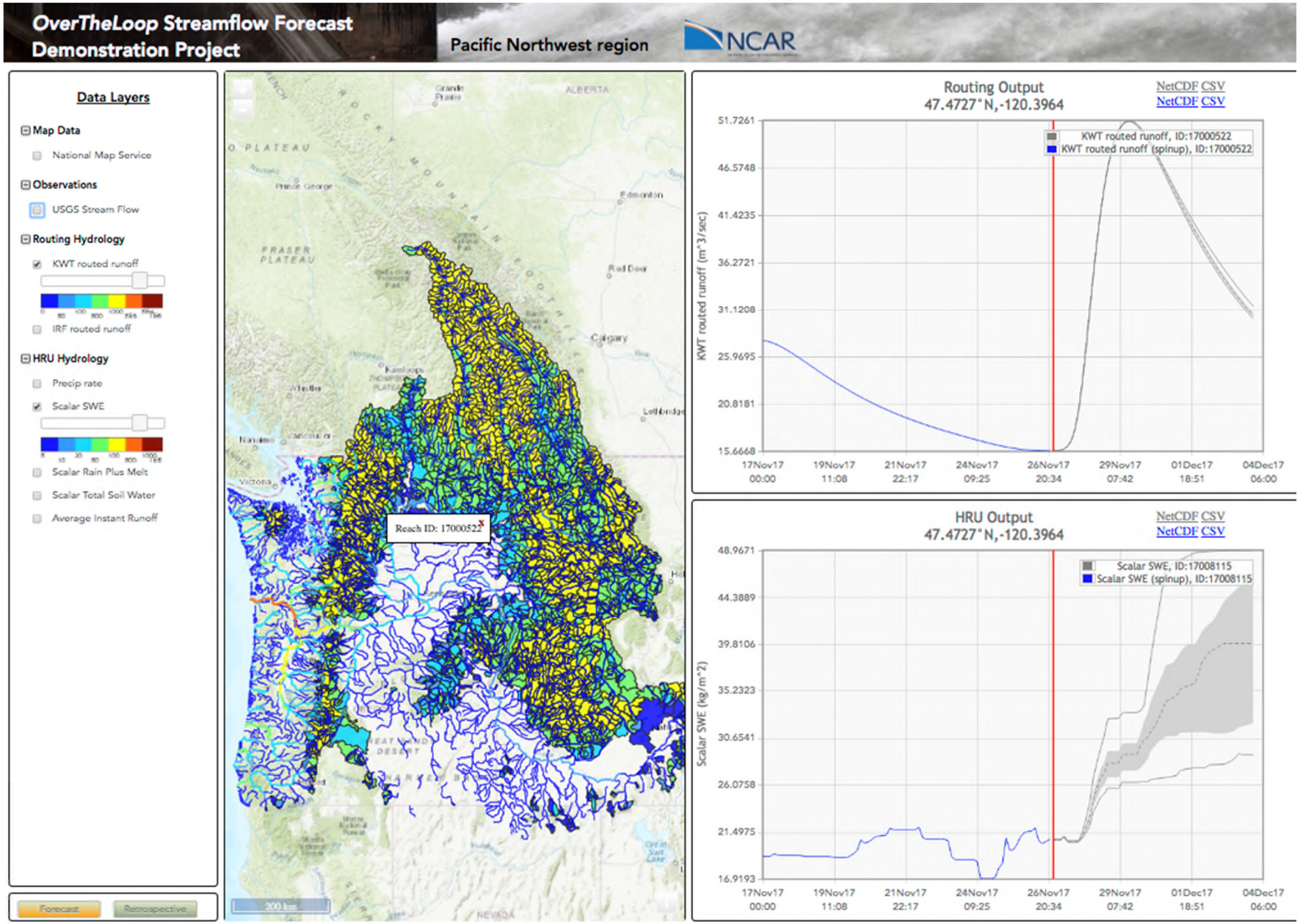
**Updated:** Tue Dec 13 15:13:57 UTC 2016

| 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/

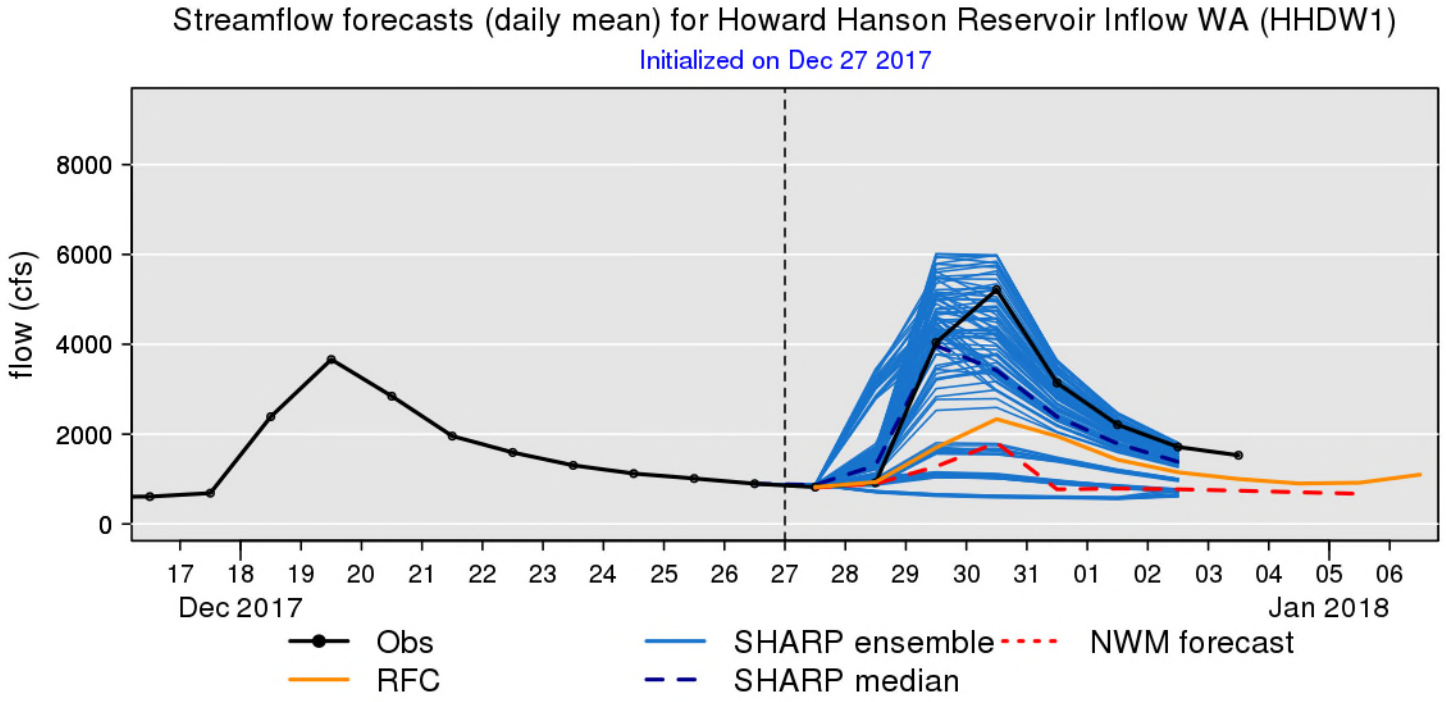
# Intermediate scale modeling enables ensembles

UNIVERSITY OF SASKATCHEWAN
Global Water Futures
GWF.USASK.CA



# Experimental short-range predictions

- Operational demonstration shows hydrologic DA is workable in the pilot watersheds
- Hydrologic DA can replace some of the manual adjustment (MODs) now made by human forecasters





- Clark, E.A., A.W. Wood, and B. Nijssen, 2017: Assessing ensemble particle filters for the estimation of model states for streamflow forecasting, Water Resources Research (in review).
- Clark, E. A., A. W. Wood, B. Nijssen, and M. P. Clark. Implications of streamflow data assimilation via particle filter on streamflow forecasts in basins with seasonal snow. *Hydrology and Earth System Sciences*, in preparation.

# Outline



- Motivation
  - The hydrologic prediction challenge
  - System requirements
- Paradigms in hydrologic prediction
  - Traditional (forecaster in the loop)
  - Ensemble intermediate complexity
  - Hyper-resolution land modeling
  - Coupled prediction systems
- Challenges in continental-domain hydrologic modeling/prediction
  - Ensemble forcing
  - “Agile” modeling
  - Parameter estimation
  - Data assimilation
- Summary and research needs

# Modeling strategy



- 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.

# Looking Forward

- For tailored watershed applications, ensemble flow forecasts are operational
    - over a decade of learning and experimentation have been successful
    - yet key components such as automated DA require more work for widespread use

- For large domain applications (national, global), ensemble hydrologic forecasts are also now operational
    - these support medium-range qualitative early warning
    - yet substantial science and data challenges undermine their quality
        - lack of constraining observations
        - weak parameter estimation / calibration
        - lack of DA and post-processing

- Continued effort is required to make most national and global-scale ensemble systems usable locally for decision-making
    - need convergence between ESM and operational watershed communities

- There is still a need for communication about ensemble forecast value and usability for decision-making
    - early adopters have been successful, but broader understanding and acceptance is still a challenge


QUESTIONS??