United States Department of Agriculture
Natural Resources Conservation Service

USDA NRCS

An introduction to the new NRCS water supply forecast platform for the US West: a metasystem blending AI, geophysical knowledge, multi-model inference, and practical requirements

Sean W. Fleming, David C. Garen, Angus G. Goodbody, Cara S. McCarthy, Lexi C. Landers

Colorado River Hydrology Research Symposium Webinar Series

8 October 2020

Acknowledgements

- Valuable technical conversations with several people including V. Muggeo (Dept. Economics, Business, and Statistics, University of Palermo), A. Cannon (Climate Research Division, Environment and Climate Change Canada), C. Schwarz (Dept. Statistics and Actuarial Science, Simon Fraser University), W. Hsieh (Dept. Physics and Astronomy, University of British Columbia), and B. Meredig (Citrine Informatics Inc. and Dept. Materials Science and Engineering, Stanford University)
- Presentation draws some background material from prior talks at the Smithsonian Museum, Oregon State University and OMSI Science Pubs, NRCS Snopac meeting, and Los Alamos National Laboratory Center for Nonlinear Studies

Context & motivation

NRCS water supply forecasting: continuing a 5,000 year mission

https://commons.wikimedia.org/wiki/File:N-Mesopotamia_and_Syria_english.svg (Goran tek-en)

Ancient Mesopotamia: the first (or among the first) to invent

- the wheel
- agriculture
- writing
- math
- and most importantly of all, beer

Name derives from the ancient Greek for "between rivers"

Water resource management was foundational to the development of human civilization – and still is

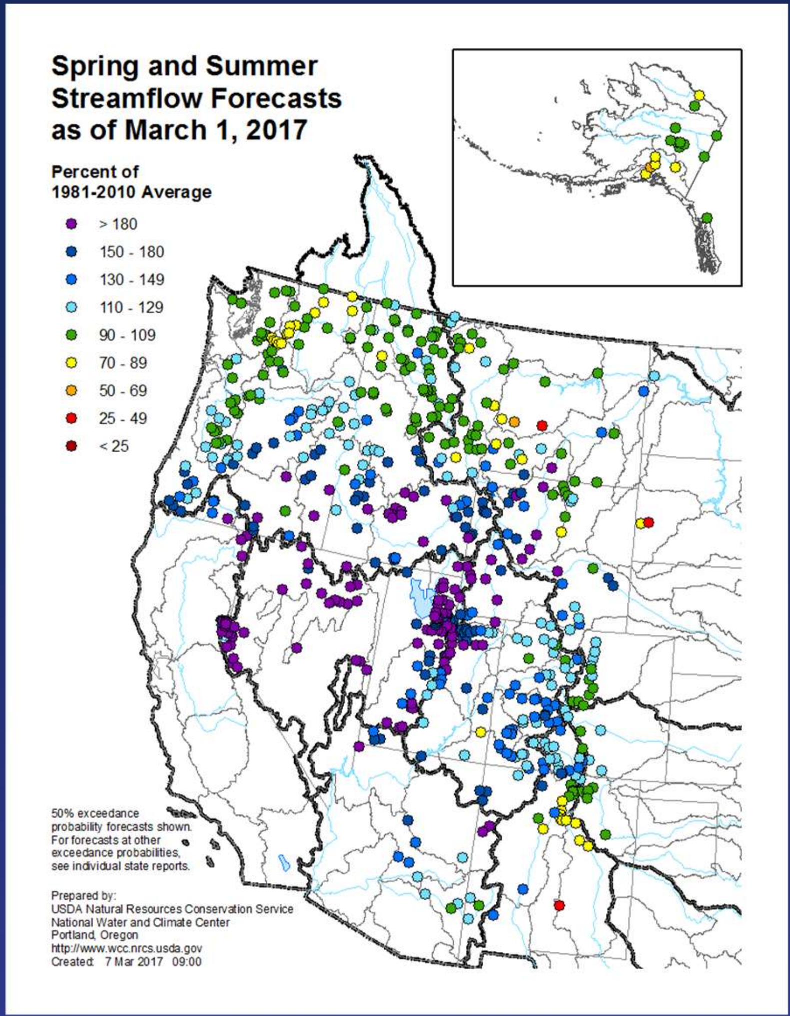
Operational water supply forecasting (WSF) in the US West

Forecasts that are typically issued once per month (officially; also more frequent updates), usually beginning in winter, of spring-summer total flow volume for a given point on a given river, performed by institutions having direct accountabilities to end users around reliable generation of forecast products

- Operational WSFs & the hydrological models and data underlying them provide a practical basis for water resource management throughout the North American West
- Snowpack:
  - US West precipitation is primarily in winter, and mainly as mountain snowpack stored until spring
  - 70-80% of total regional runoff in the 3 largest ranges (Rockies, Sierra Nevada, Cascades): snowmelt
  - Water demand: strongest in spring-summer
  - Snowpack = huge de facto reservoir, and snowpack data = main source of WSF skill (predictors in statistical models, features in machine learning models, data assimilation in process models)
  - USDA SNOTEL network and products derived from it are primary source of snowpack data
- Various other variables also used operationally (precipitation, antecedent flows, climate indices)
- Intensive research on additional predictors (snow remote sensing, seasonal climate models, etc.)

NRCS WSF

- National Water and Climate Center of the NRCS operates the largest stand-alone operational WSF system in the US West
- May be the largest statistical operational WSF system in the world
- > 600 forecast locations
- Each with multiple issue dates (forecast dates) and target periods (forecast horizons)
- Several statistical and process simulation systems in use
- Core method forming official forecast is principal component regression (PCR)



US Department of Agriculture NRCS NWCC

What is PCR?

- **Principal component regression (PCR)**
  - Introduced to WSF by David Garen (1992) at NRCS, since adopted for operational WSF at many other organizations elsewhere
  - Principal component analysis (PCA) is applied to input variables
  - Resulting PC scores are candidate predictor variates in forward stepwise multiple linear regression
- **PCR addresses multicollinearity & reduces dimensionality of input variables**
  - E.g.: SWE measurements at multiple SNOTEL sites provide operational redundancy & capture spatial heterogeneity, but form a high-dimensional & multicollinear dataset
- **NRCS implementation follows a specific path**
  - Simple tree-based algorithm used for optimal input variable selection
  - Statistical significance tests used for optimal PC mode retention
  - Probabilistic method: prediction intervals generated using a common heuristic based on the quantiles of a normal distribution having a mean equal to the regression prediction and a standard deviation equal to the regression standard error
  - Official forecasts (end products) are 90%, 70%, 50%, 30%, and 10% exceedance probability flows

Essential context to operational WSF system design

Preparatory steps

Thorough assessment of needs and options before undertaking system design:

- **Clearly document existing NRCS WSF system**
  First implemented about 25 years ago; operational forecasting systems are complex & organic, evolving over time
- **Assess abilities and limitations of current system**
  Included documentation of known issues & completion of advanced statistical diagnostics
- **Comprehensively review data-driven WSF modeling progress**
  Assess for potential relevance to NRCS operations, including topics like longer lead times using climatic forcing data, more advanced statistical & machine learning methods, ensemble modeling, statistical-process simulation modeling hybrids, & other research directions
- **Assess implications of global anthropogenic climate change on WSF**
  Mainly around a need for improved seasonal prediction capabilities given an increasingly unpredictable hydroclimatic system, with lower snowpack and possibly greater variability, while water demand increases
- **Develop initial blueprint & several preliminary scoping models**
  On the basis of foregoing preliminary steps and NRCS system requirements, experiment with some concepts potentially underlying a new approach and assess their suitability for inclusion in the full prototype system

Model characteristics & selection matrix

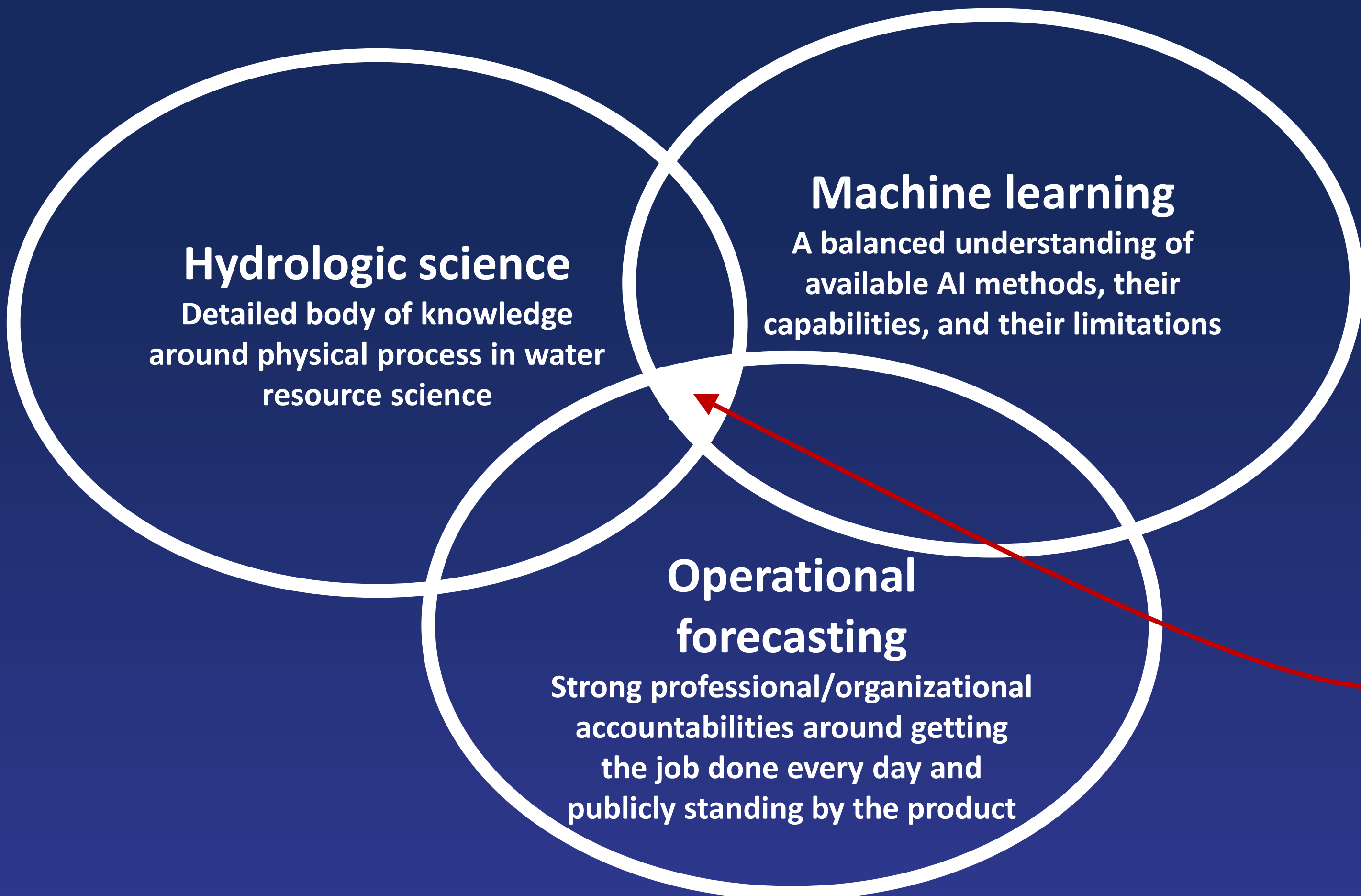
Characteristic	Process simulation	Statistical regression model	Machine learning
Cost/ease/requirements	terrible to moderate	very good to excellent	moderate to very good
Code reliability	very poor to excellent	moderate to excellent	good to excellent
Code transparency	terrible to excellent	very good to excellent	moderate to excellent
Physical interpretability	moderate to excellent	poor to very good	very poor to very good
Deterministic accuracy	terrible to very good	moderate to very good	very good to excellent
Probabilistic accuracy	terrible to moderate	moderate to excellent	moderate to excellent
Innovation/opportunities	terrible to very good	terrible to good	excellent

- **Which class of models is "best" depends on what you're aiming to do**
- **Main punch lines around machine learning as relevant to NRCS WSF:**
  - Overwhelmingly good bang for the buck as a predictive tool
  - Leverage the data science all around us in our everyday lives to produce better, cheaper WSFs
  - But can't just throw machine learning at WSF and expect it to work right: must approach this in a very application-specific way, carefully choosing/developing the right tools for the job

Machine learning and hydrologic prediction

- **Long and somewhat mixed history**
  - Research papers using AI for river forecasting date back to 1995
  - Yet true operationalization is mostly very recent (as far as most of us know…)
- **Several specific stumbling blocks**
  - Lack of probabilistic forecasts
  - Absence of a good storyline around physical hydroclimatic process for most AI models
  - Skepticism in high-stakes operational settings where existing techniques are well-established and new methods in the research literature are unproven and sometimes oversold
  - Misunderstandings and lack of technical training and professional familiarity among many physical scientists with AI, its strengths, and its limitations
- **But things have changed**
  - Just about everyone is getting used to the idea of AI
  - Tools are far more accessible, and methods are far more diversified
  - Explainability, overtraining, and prediction uncertainty estimation questions can be addressed with careful AI design and implementation choices

Convergence of knowledge, tools, and requirements



Intersection:

Need to build AI prediction systems that align with domain-expert expectations around consistency with both physical hydrologic science and organizational requirements of operational prediction organizations

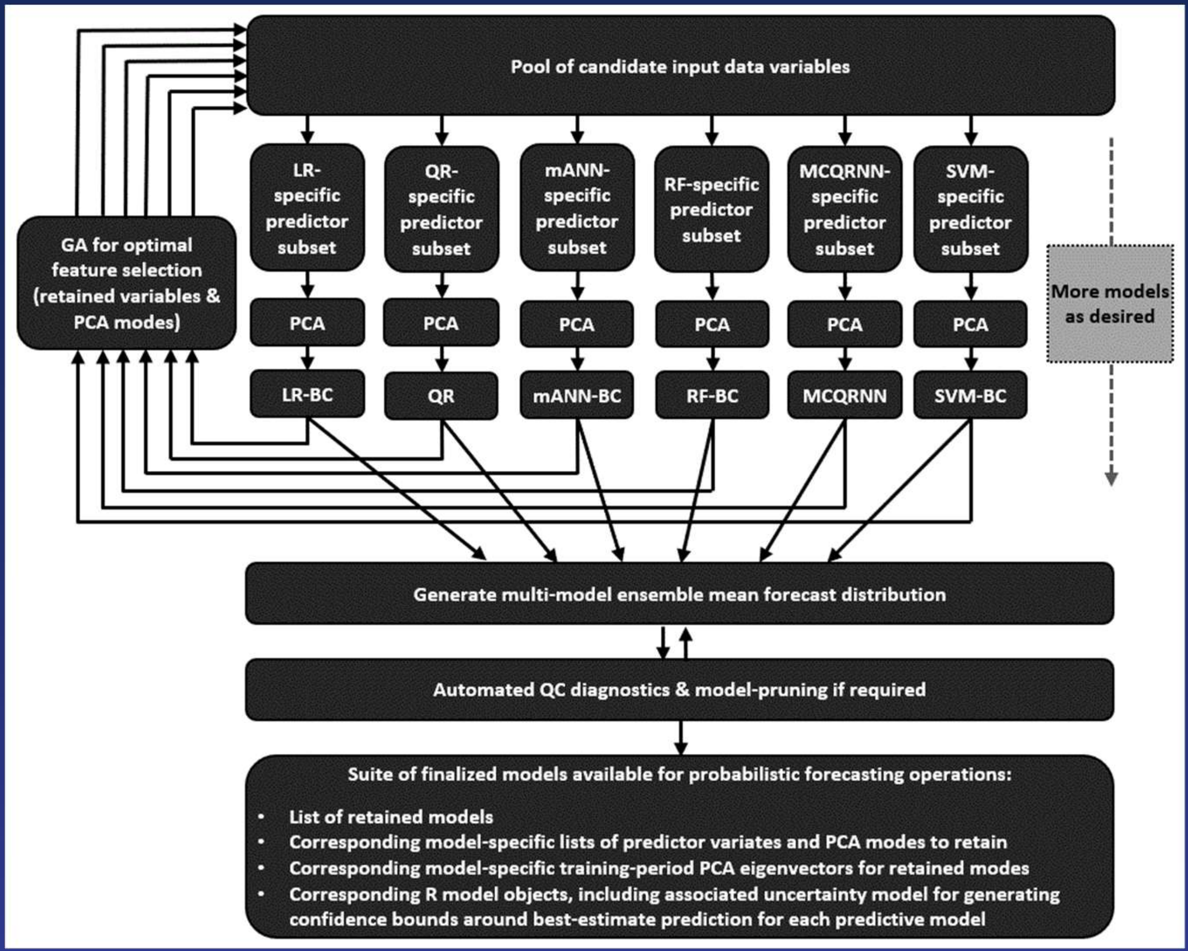
Some key design criteria

1. Improved forecast accuracy
2. Improved potential for automation
3. Relatively low cost and good ease of development, implementation, and operation
4. Seamlessly address three known issues: (a) nonlinear functional forms, and (b) heteroscedastic and (c) non-Gaussian residuals
5. Modular and expandable
6. Geophysically interpretable forecasts
7. Achieve balance between visibly demonstrating innovation & performance gains vs. construction from established building blocks using proven tools
8. Multi-method ensemble approach: required to address equifinality and model selection uncertainty over diverse geophysical environments across the US West
9. Accommodate high-dimensional multicollinear predictor datasets and potential for multiple independent input signals: will only grow more important in the future with more spatially distributed inputs

Multi-method machine learning metasystem (M^4)

Solution: a hybrid prediction analytics engine

- Several very carefully selected modeling threads: handle heteroscedastic & non-Gaussian residuals + nonlinearity, allow physicality constraints like monotonicity & non-negativity
- Combine with dimensionality reduction using statistical pattern recognition, a genetic algorithm for optimizing feature creation and selection, some parallelization across processor cores, a flexible architecture, a degree of AutoML, and physicality checks and adjustments
- Forms an integrated, modular, ensemble forecasting metasystem

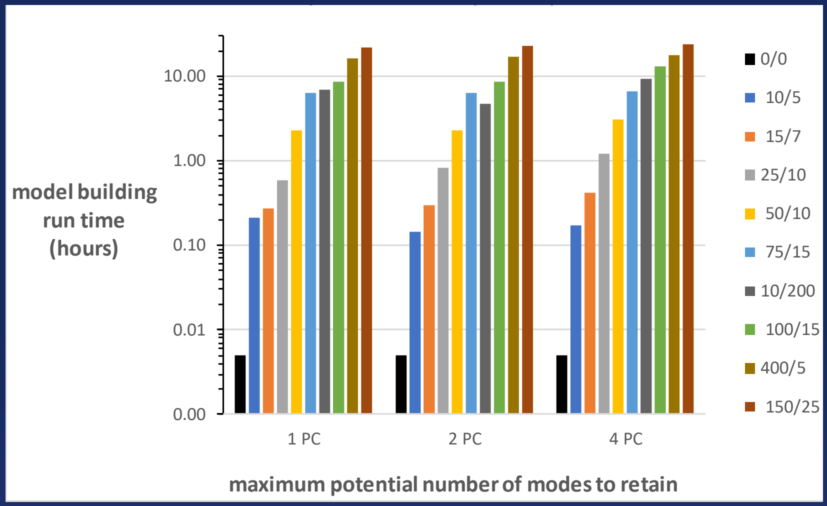


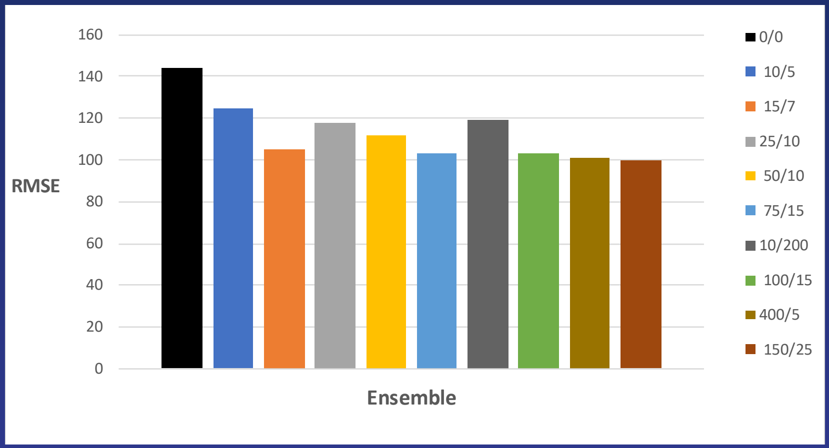
Toward achieving AI intersections with process physics

- Operational hydrologist-directed features engineering
  - Reflects end user knowledge around representativeness, reliability, quirks and capabilities, of potential input variables & measurement sites & geophysical interpretations of PCA modes
  - Key location in AI development process for domain experts to insert physical hydrologic knowledge
- Reframe as a low-dimensional problem with a parsimonious solution
  - PCA data pre-processing & compact ML architectures enable visualizing input-output relationships
- Monotonicity constraints
  - Some feature-target relationships are known to be significantly nonlinear but monotonic
  - Certain AI methods allow a monotonicity constraint to be enforced; also encourages regularization
- Non-negativity constraints
  - Negative-valued flow volume predictions are non-physical yet can happen in some prediction systems
  - Certain AI methods allow a non-negativity constraint to be enforced
  - M^4 also includes algorithmic logic to test final forecast distribution for non-negativity and sequentially prune individual ensemble members if constraint is violated

AutoML

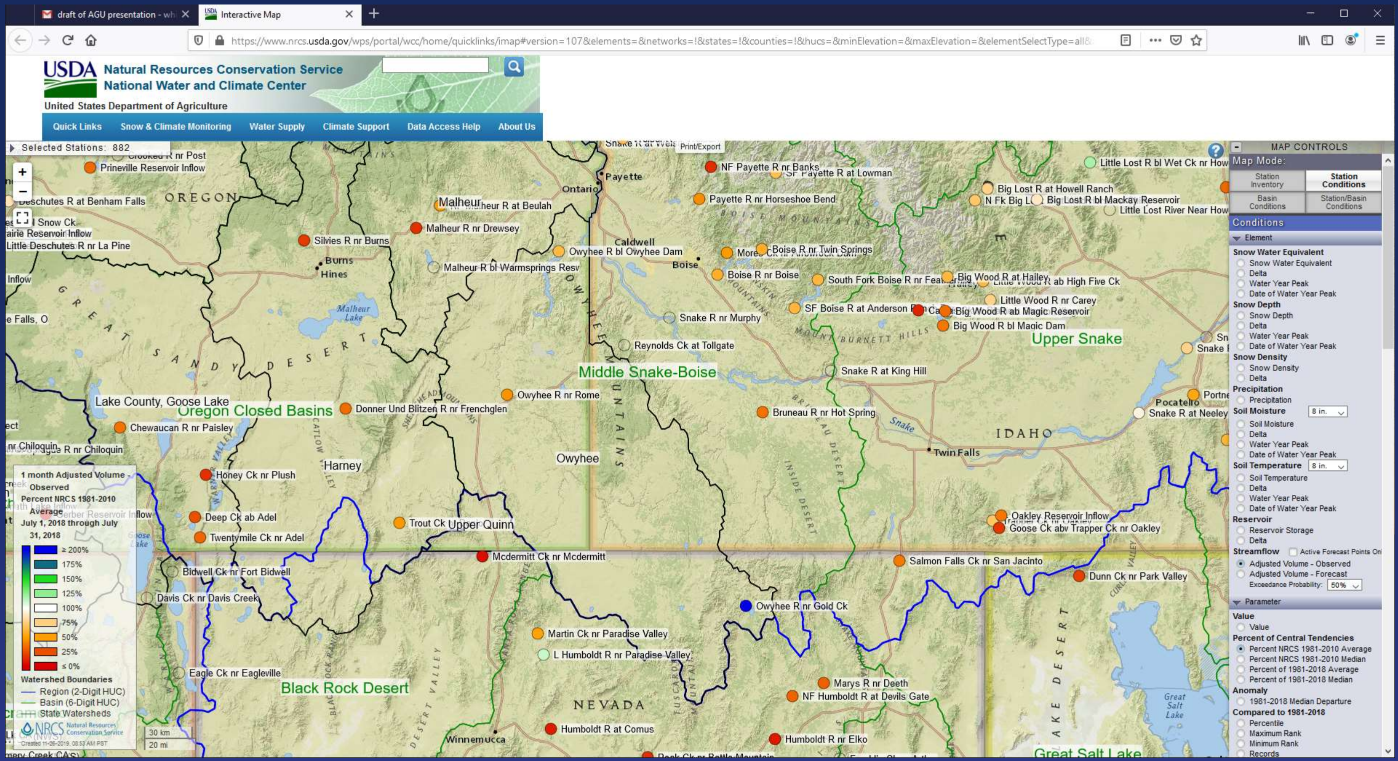
- Completely hands-off operation not desired, but streamlining/automation of many tasks is: leads to judicious use of AutoML
- Most optimization & decision points in the overall prediction algorithm (including values of some machine learning hyperparameters, e.g., around regularization) are automated with options for manual overrides
- Others have been pre-calibrated to reasonable values for this application on the basis of experimentation using various NRCS WSF test cases and problem setups
- Example: locate reasonable default values for pop size & # generations in genetic algorithm



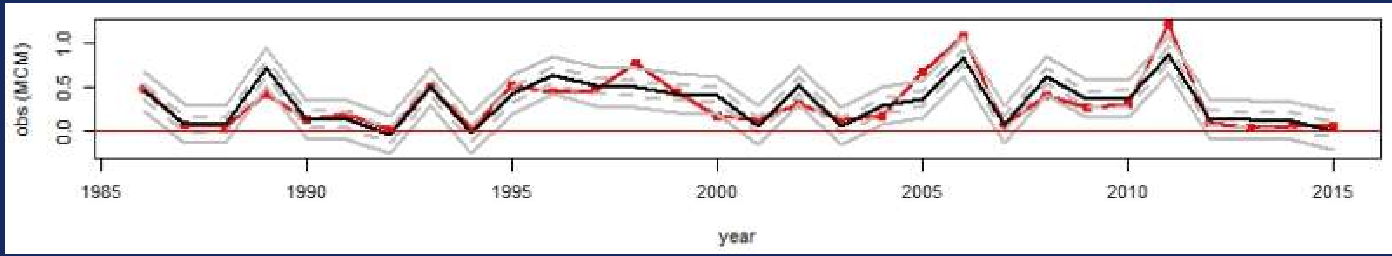


Application example

- Owyhee River near Rome, April 1 forecast of April-July volume
- Particularly troublesome forecast point
- Here, use up to four PCA modes derived from candidate input variable pool of up to 18 SNOTEL SWE & precipitation datasets

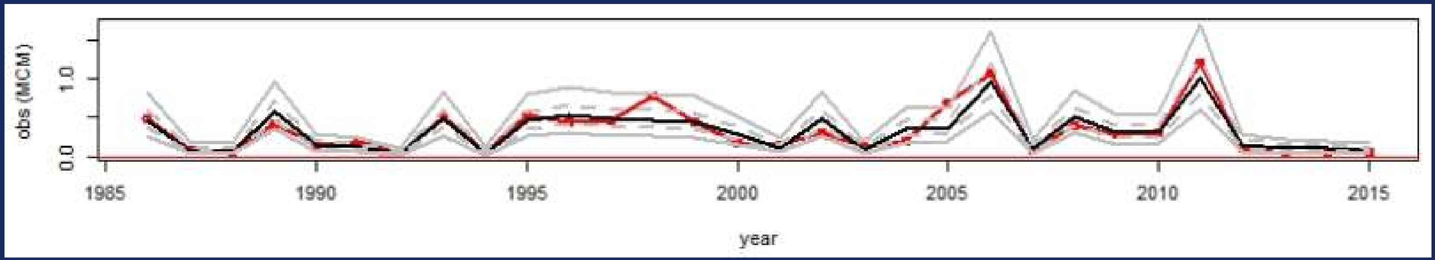


Linear PCR



- **Linear PCR solution produces best-estimate predictions that are physically impossible: runoff volume can't be negative**
- **Homoscedastic Gaussian assumptions do not generate required time-varying prediction bound widths: too wide in low-flow years, too narrow in high-flow years**
- **This is fixable in PCR using predictand transforms – but it's a manual process, and…**
  - **Time-consuming & labor-intensive**
  - **Dependent on expert opinion & not objectively reproducible: reliability & defensibility issues**
  - **Cannot separate distributional issues from functional form issues**
  - **Basically unsatisfying: should pick models that suit the data, not change the data to suit the model**
  - **Smart choice if linear PCR is all that is readily available, but recently, options have grown massively**

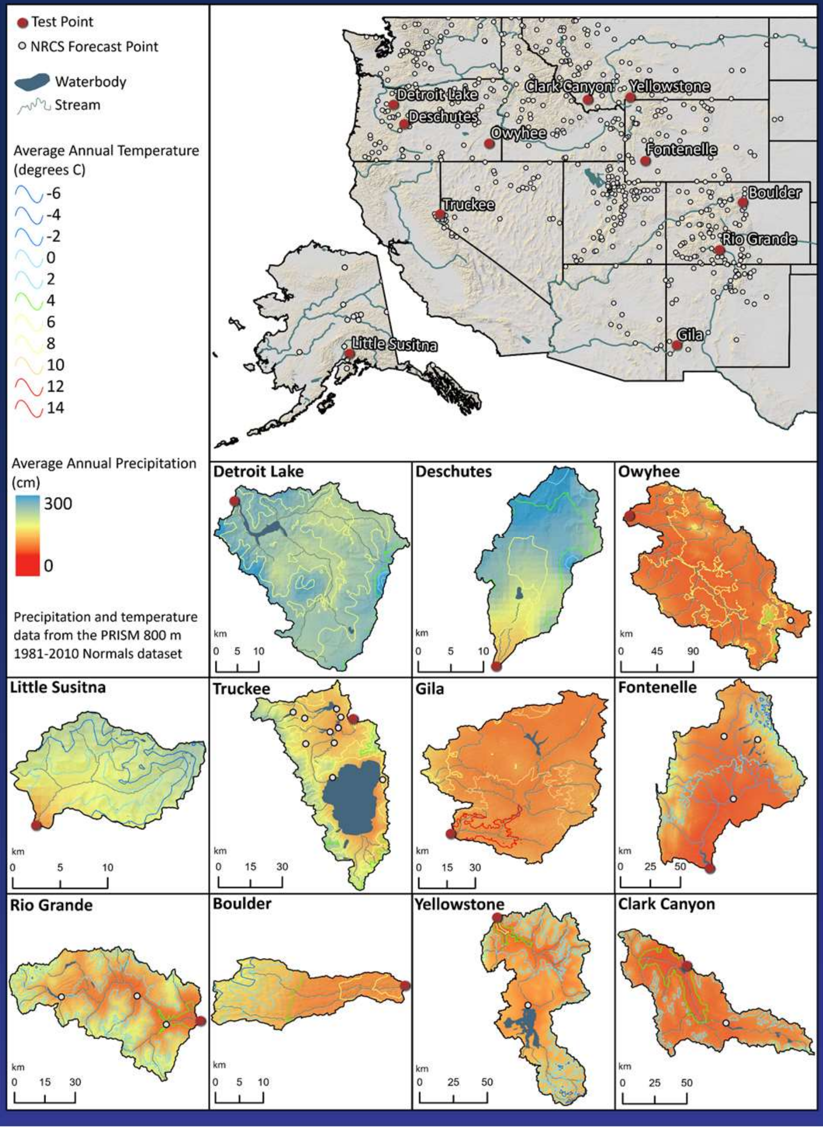
M⁴



- **Multi-model prediction engine generates forecasts that…**
  - **are physically reasonable, in particular non-negative, from the 90% to the 10% exceedance flows**
  - **have prediction bounds that can vary in width from year to year, if needed**
  - **have prediction bounds that can be asymmetric about the best estimate, if needed**
  - **are relatively robust to outliers due to use of median as best estimate in some constituent models**
  - **are automated, don't require extensive user intervention/subjective judgement, and are reproducible: fast, reliable, objective**
  - **use up-to-date advanced statistical and machine learning techniques and philosophies specifically suited to accurate prediction of complex, open, nonlinear systems**
  - **show forecast skill better than linear PCR as measured by several cross-validated performance metrics**
- **… yet is still relatively simple and cheap to build and operate.**

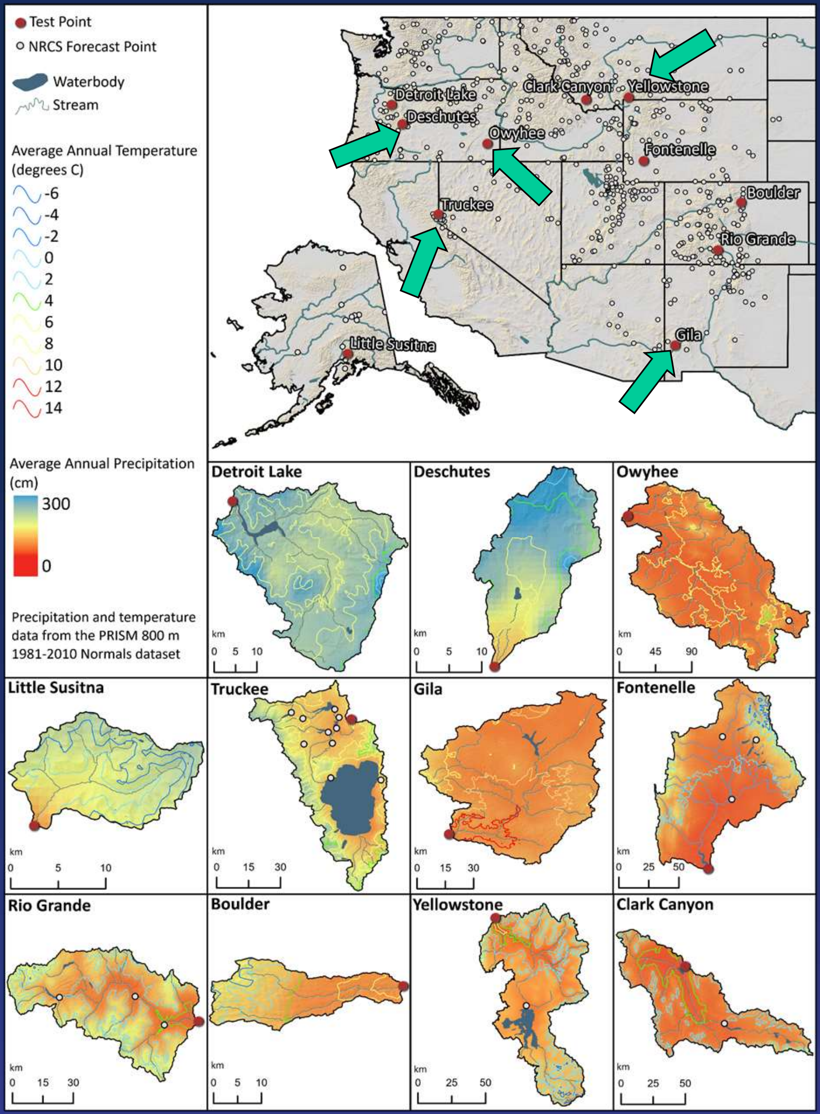
Hindcast testing

- **20 test cases spanning diverse geophysical environments & statistical characteristics across the western US and Alaska**
- **Use PCR as a challenging benchmark: well-established method about on par with ESP**
- **Ensemble mean M^4 prediction always meets or beats linear PCR in terms of deterministic and probabilistic performance metrics and physicality**
- **Always more consistent than, and often outperforms, any of its constituent models**
- **After candidate predictor selection, completely hands-off automated use**



Operational testing

- **Live operational testing at subset of 5 forecast locations during 2020 forecast season alongside existing operational system**
- **Confirmed feasibility of workflows in time-and data-constrained genuine operational setting**
- **Easily rebuilt some models on-the-fly due to COVID-19 related snow course data absences**
- **Confirmed physical interpretability of predictions, forming straightforward 'storyline' for the issued forecast in terms of current hydroclimatic conditions and watershed characteristics**



Conclusions

Current & next steps

- Further software development
  - Streamline code and work with software development team to build UI (GUI, IT/database linkages, interactive capabilities around graphics, mapping, and data pre-processing & forecast distribution post-processing)
- Other R&D directions
  - Continue R&D on topics not covered here: supplementing PCA with NMFk, using snow remote sensing or assimilation data as WSF inputs, integrating other AIs into the modular metasystem, experimenting with new sources of seasonal climate prediction information, etc.
- Explore using M^4 as a more generalized forecast integration platform
  - Expand scope by ingesting probabilistic forecasts from external process-simulation models (e.g., NWS RFS ESPs, NRCS PRMS ESPs, NOAA National Water Model, etc.) into the multi-technique metasystem's final ensemble – opportunity for re-establishing interagency forecast coordination?

For further information

- For technical details of M^4 see peer-reviewed AI literature:

  Fleming SW, Goodbody AG. 2019. A machine learning metasystem for robust probabilistic nonlinear regression-based forecasting of seasonal water availability in the US West. *IEEE Access*, 7, 119943-119964, doi:10.1109/ACCESS.2019.2936989

- Additional papers in progress address application from a water resource perspective & document other AI directions we are pursuing

- Feel free to reach out to us: sean.fleming@usda.gov