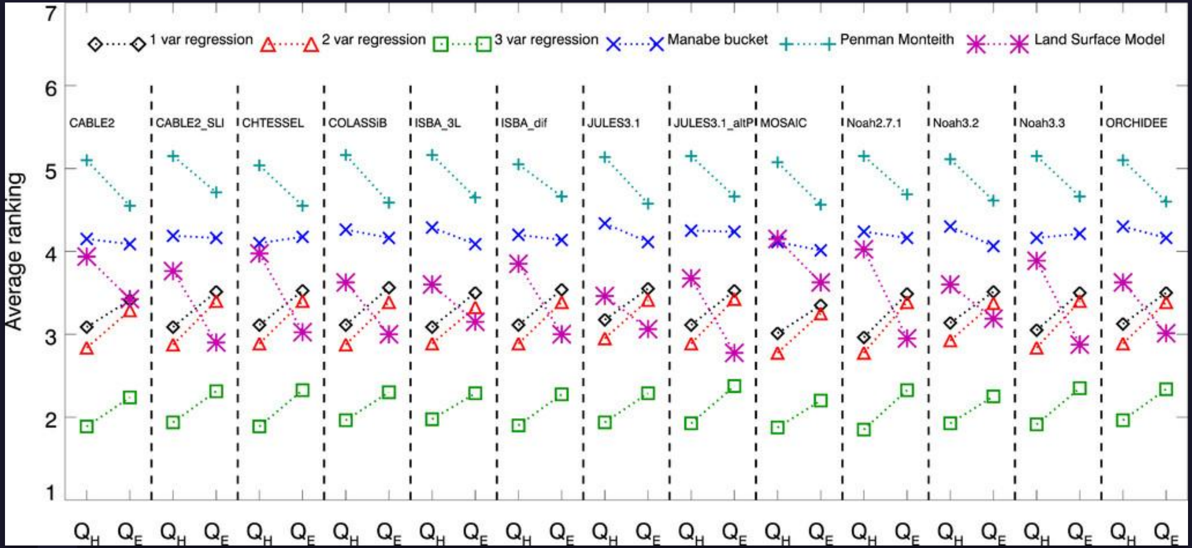
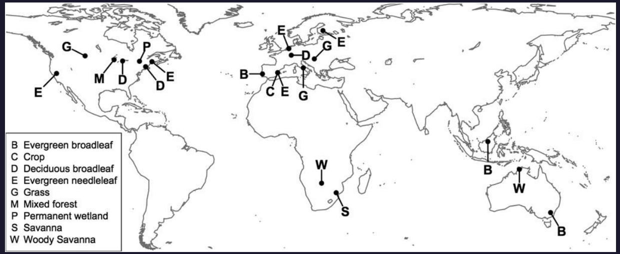
Models, Hypothesis Testing, & Machine Learning

A tale of two experiments:
(1) The Plumber Experiments

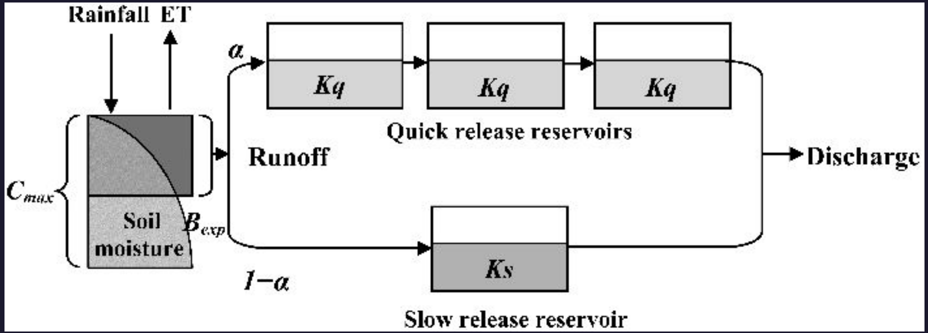




Best, Martin J., et al. "The plumbing of land surface models: benchmarking model performance." *Journal of Hydrometeorology* 16.3 (2015): 1425-1442.

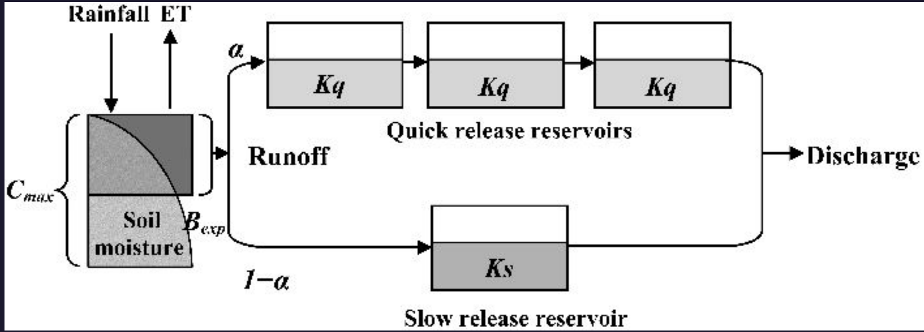
A tale of two experiments:

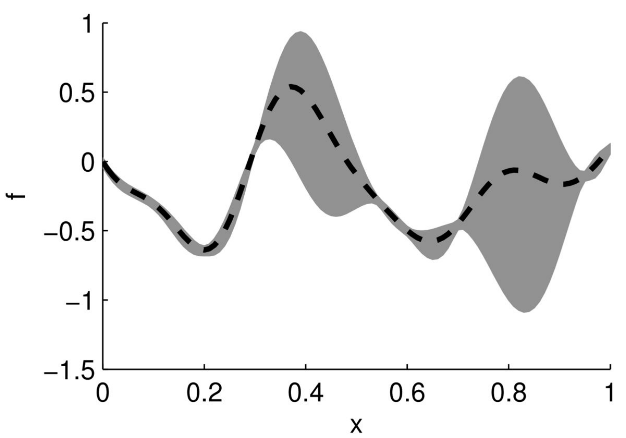
(2) State-Space Learning



$$\frac{dX}{dt} = f(X, u, \theta)$$

A tale of two experiments:
(2) State-Space Learning



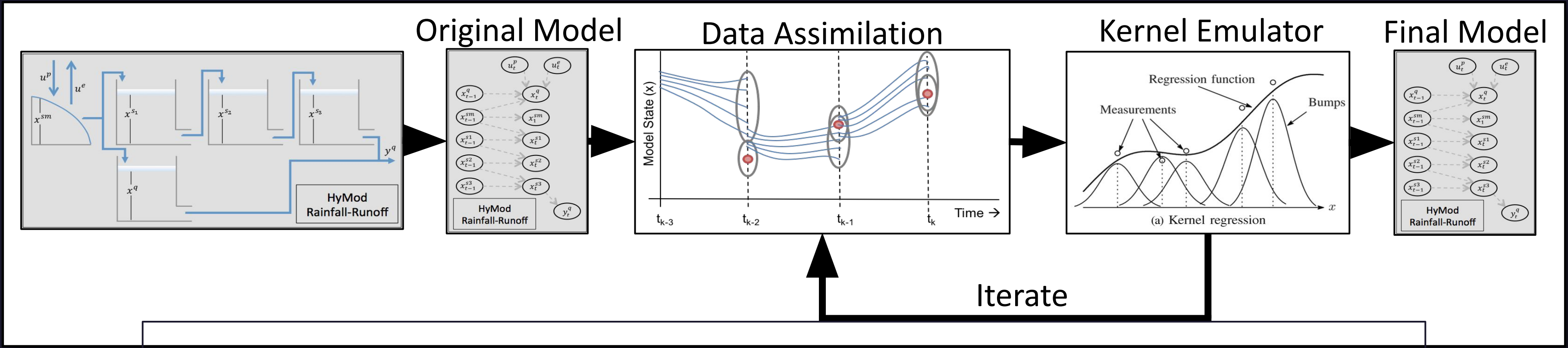


$$\frac{dX}{dt} = f(X, u, \theta)$$

$$+\, g(X, u, \theta)$$

A tale of two experiments:

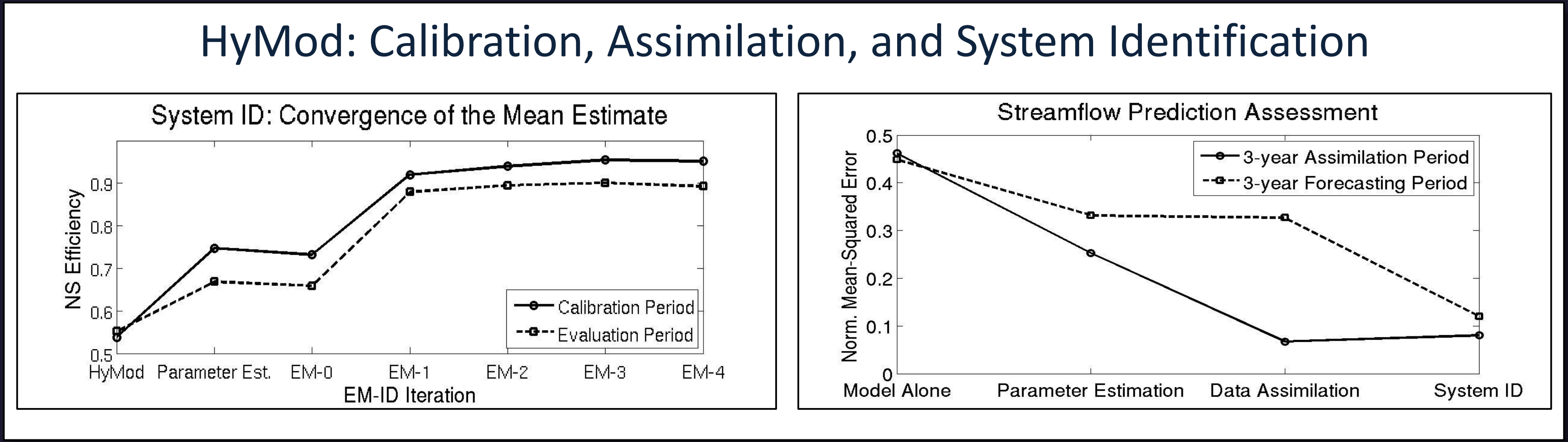
(2) State-Space Learning



Expectation-Maximization Data Assimilation

A tale of two experiments:

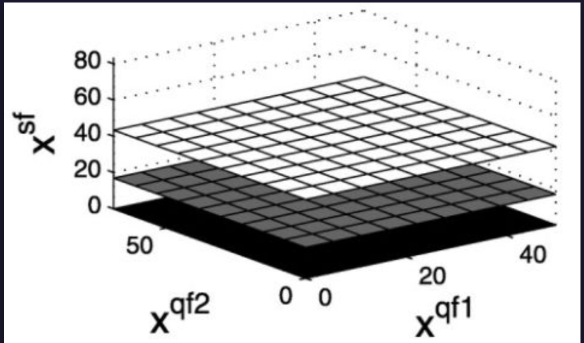
(2) State-Space Learning



Nearing, Grey S., and Hoshin V. Gupta. "The quantity and quality of information in hydrologic models." *Water Resources Research* 51.1 (2015): 524-538.

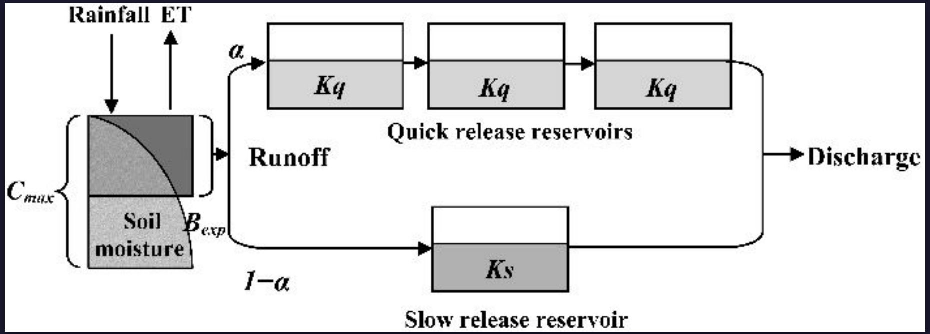
A tale of two experiments:

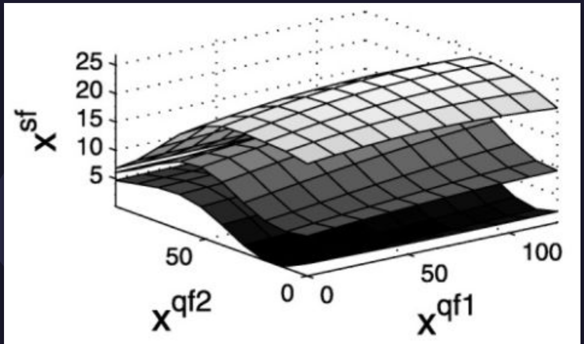
(2) State-Space Learning



before ML

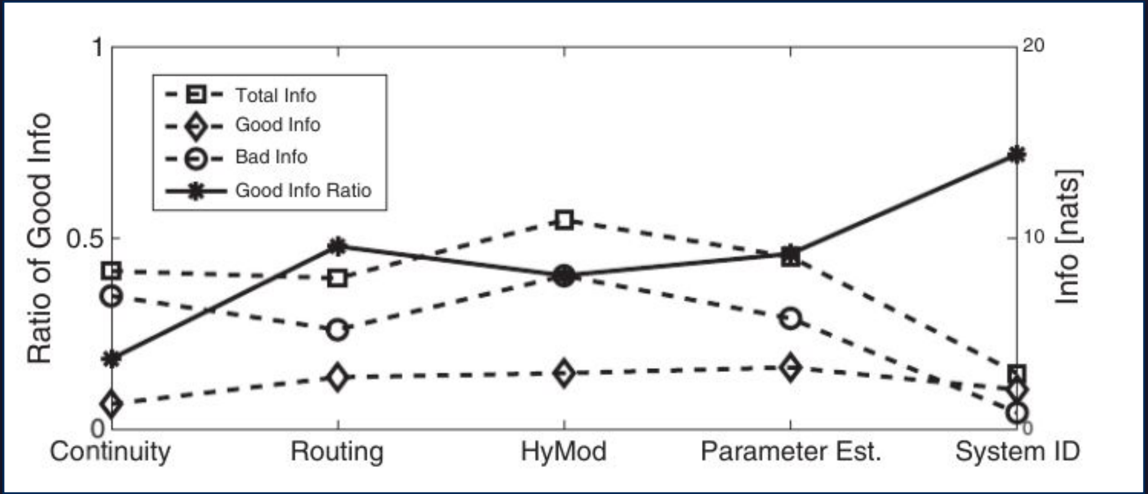
original model





after ML

A tale of two experiments:
(2) State-Space Learning



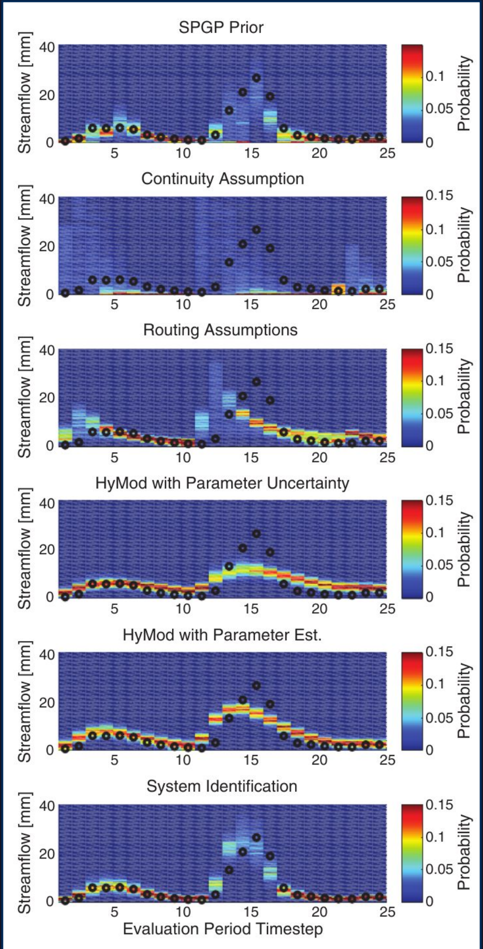
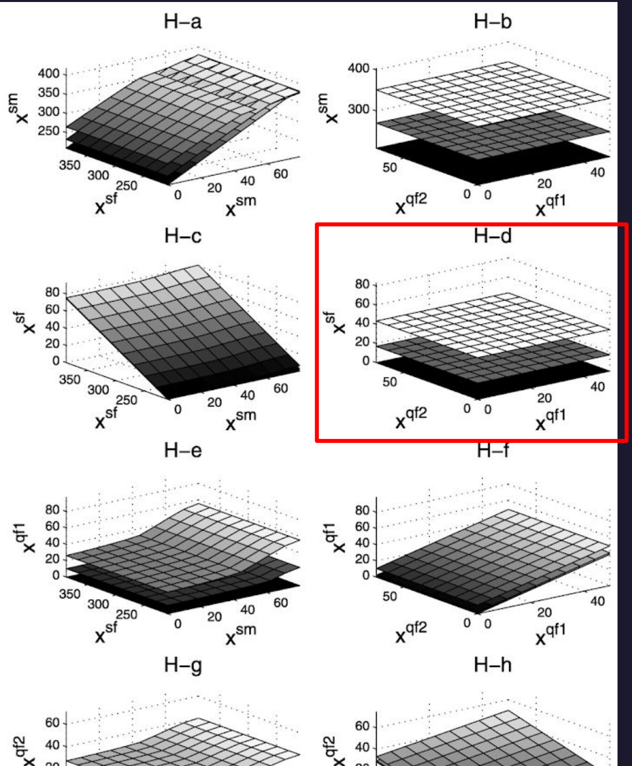


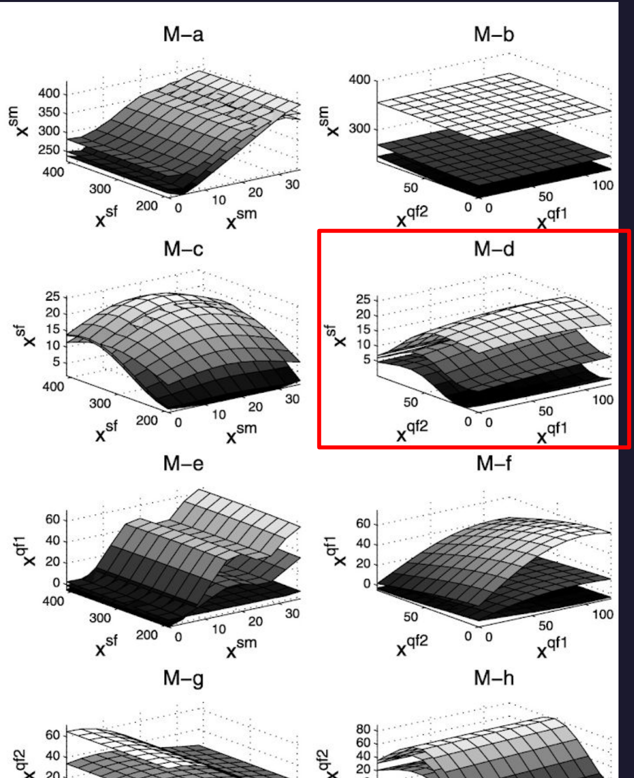
Figure 3. Evaluation period predictions of daily streamflow made by the empirical SGPR prior and each of the five inductive models. Black circles mark daily streamflow observations.

Nearing, Grey S., and Hoshin V. Gupta. "The quantity and quality of information in hydrologic models." *Water Resources Research* 51.1 (2015): 524-538.

A tale of two experiments:

(2) State-Space Learning



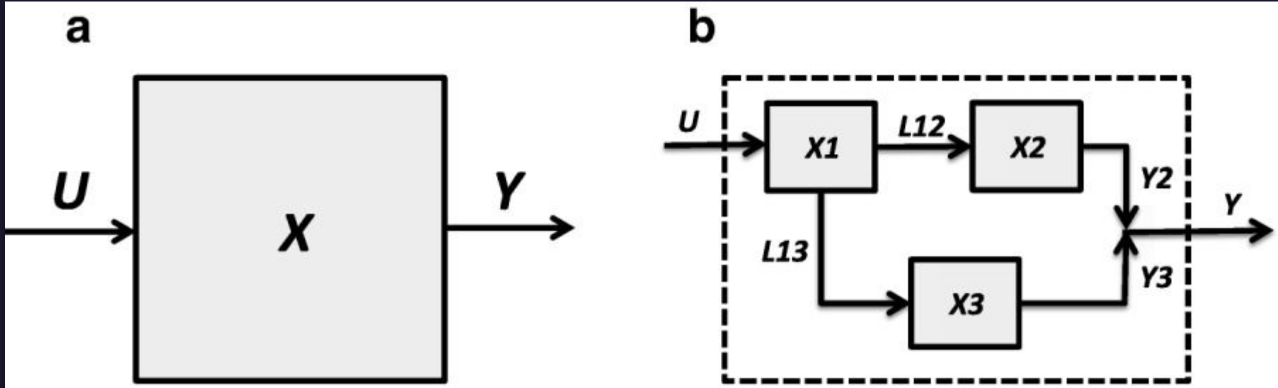


Which experiment is more interesting?

(1) ML finds extra information in data

(2) ML finds specific errors in a model

What is a model?



$$\frac{dX_t}{dt} = U_t - Y_t$$

$$\{U_t, X_t, Y_t\} \geq 0$$

$$X_t \geq Y_t$$

$$\frac{dX1_t}{dt} = U_t - L12_t - L13_t$$

$$\frac{dX2_t}{dt} = L12_t - Y2_t$$

$$\frac{dX3_t}{dt} = L13_t - Y3_t$$

$$\{X1_t, X2_t, X3_t, L12_t, L13_t, Y2_t, Y3_t\} \geq 0$$

$$X1_t \geq L12_t \qquad X2_t \geq Y2_t \qquad X3_t \geq Y3_t$$

Gupta, Hoshin V., and Grey S. Nearing. "Debates—The future of hydrological sciences: A (common) path forward? Using models and data to learn: A systems theoretic perspective on the future of hydrological science." (2014).

What is a model?

Nearing, Grey S., and Hoshin V. Gupta. "Ensembles vs. information theory: supporting science under uncertainty." *Frontiers of Earth Science* (2018).

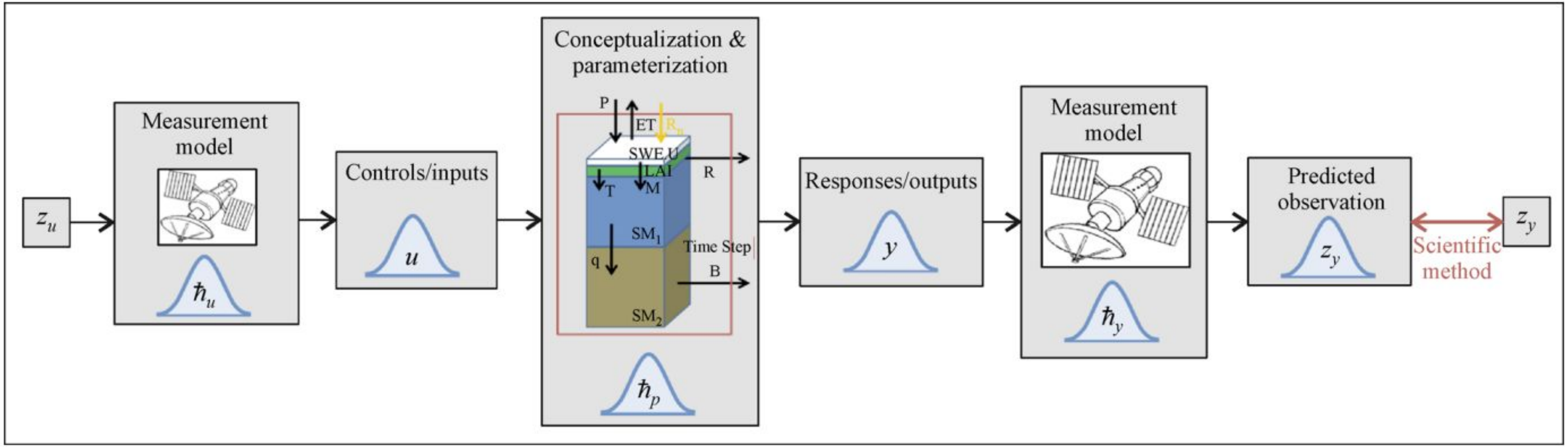
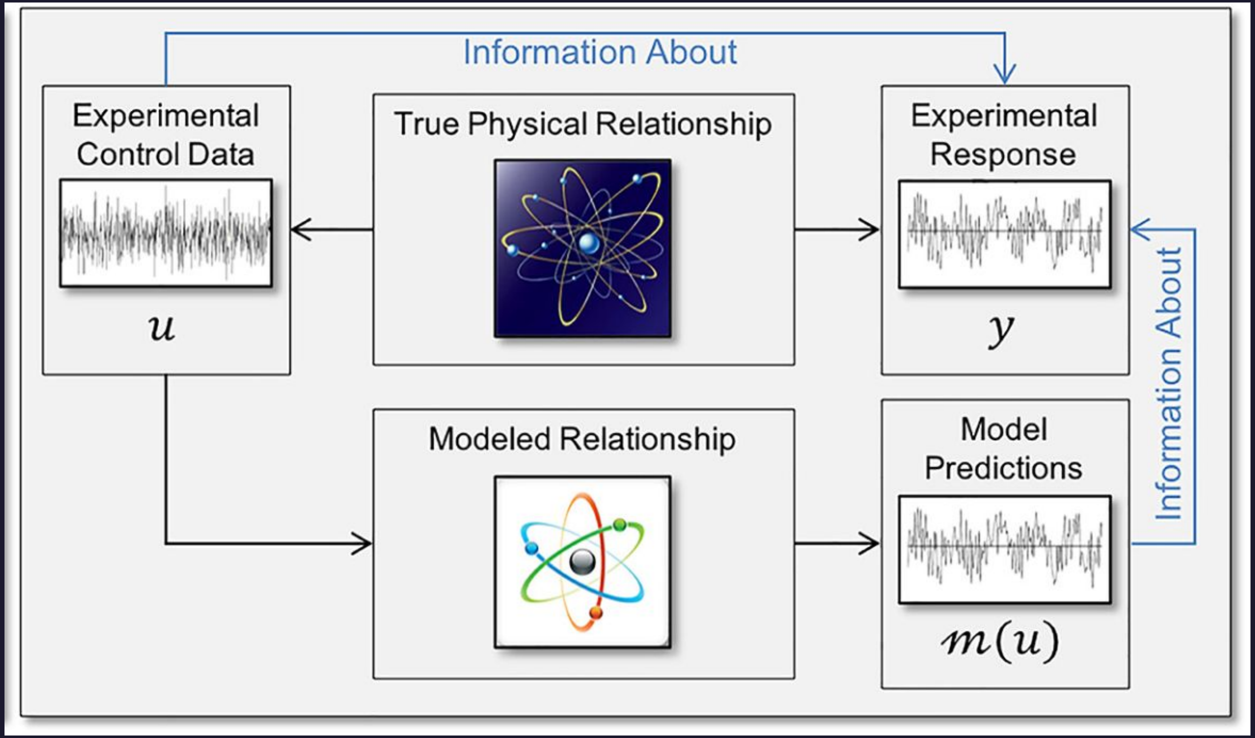


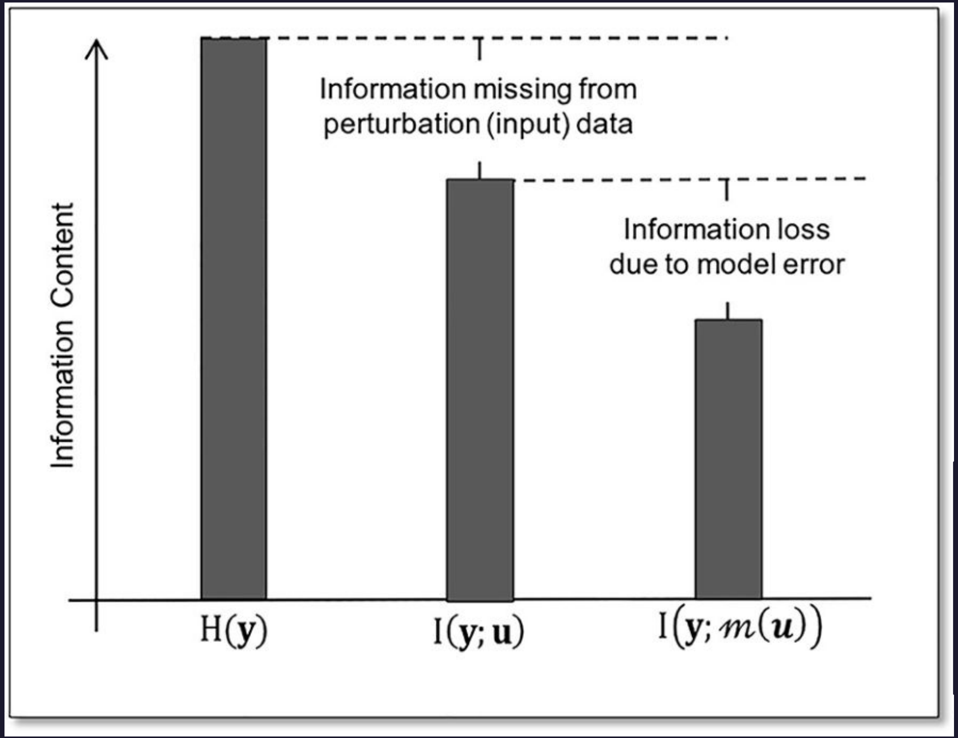
Fig. 1 Diagram of a simple experiment with a single input and single output. Even in the simplest case, we require at least one process hypothesis $\hbar_p$, and at least two measurement models $\hbar_u$ and $\hbar_y$. Aleatory uncertainty is defined as the (unknown) distribution $\mathcal{S}(z_y|z_u)$, which would only be knowable if we had access to both a perfect process model and also perfect measurement models. A full model is the conjunction $\hbar = \{\hbar_p, \hbar_u, \hbar_y\}$.

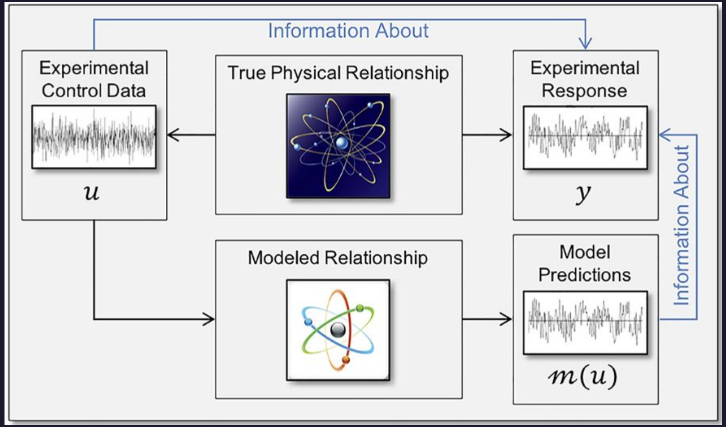
What is hypothesis testing?



Nearing, Grey S., et al. "Does information theory provide a new paradigm for earth science? Hypothesis testing." *Water Resources Research* (2020).

What is hypothesis testing?





Gong, Wei, et al. "Estimating epistemic and aleatory uncertainties during hydrologic modeling: An information theoretic approach." *Water resources research* (2013).

What is a model?

Nearing, Grey S., and Hoshin V. Gupta. "Ensembles vs. information theory: supporting science under uncertainty." *Frontiers of Earth Science* (2018).

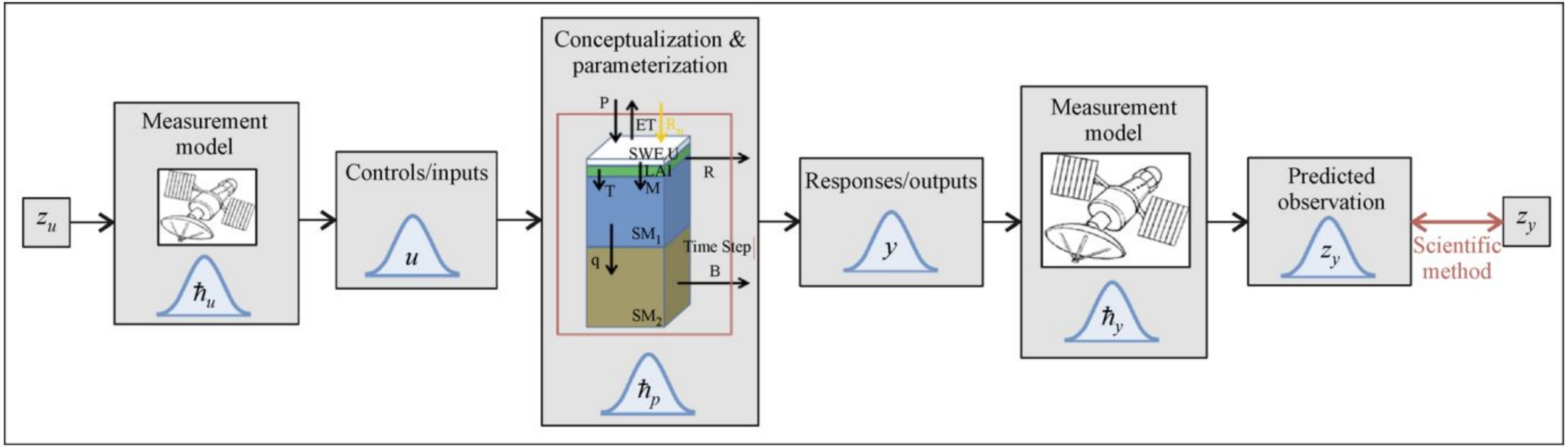
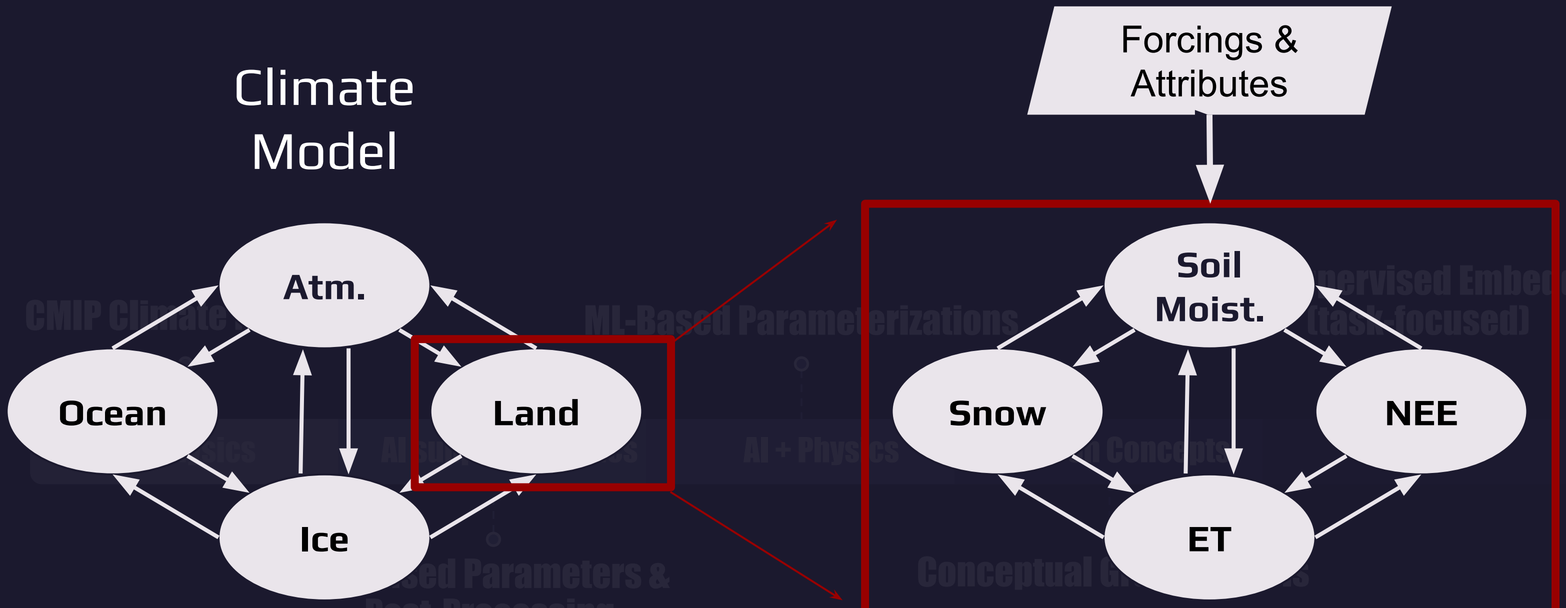
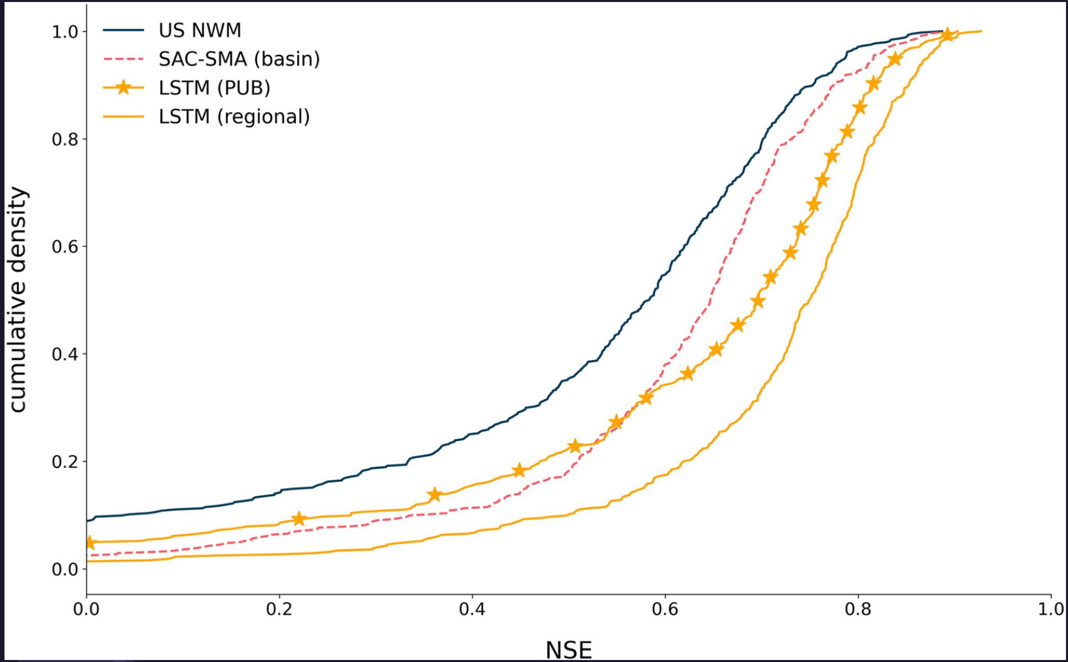


Fig. 1 Diagram of a simple experiment with a single input and single output. Even in the simplest case, we require at least one process hypothesis $\hbar_p$, and at least two measurement models $\hbar_u$ and $\hbar_y$. Aleatory uncertainty is defined as the (unknown) distribution $\mathcal{S}(z_y|z_u)$, which would only be knowable if we had access to both a perfect process model and also perfect measurement models. A full model is the conjunction $\hbar = \{\hbar_p, \hbar_u, \hbar_y\}$.

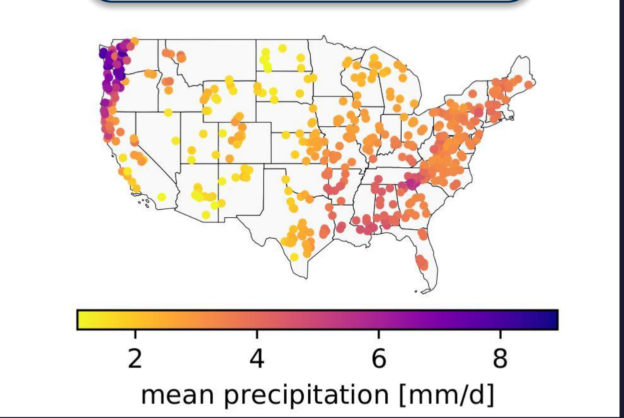
Hierarchical Graph Models



Finding new information *is* interesting



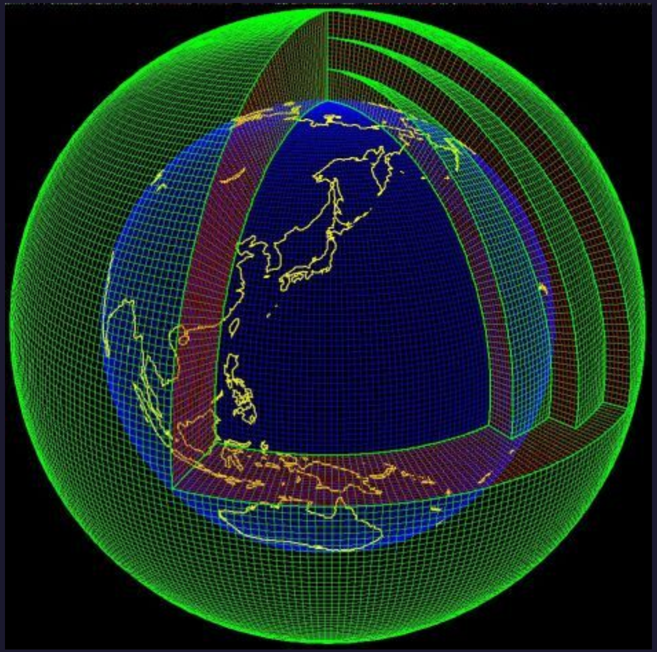
Space/time split validation: tested in “ungagued” basins



Kratzert, Frederik, et al. "Toward improved predictions in ungauged basins: Exploiting the power of machine learning." *Water Resources Research* 55.12 (2019).

Data + Domain Science Anecdotes

“Show that precipitation is useful for predicting floods.”



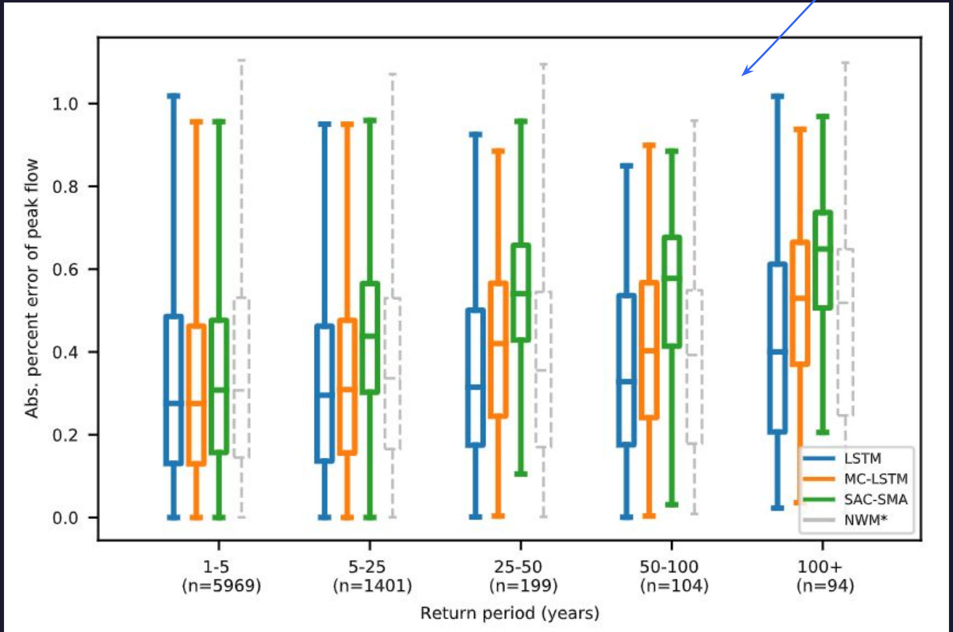
Data + Domain Science Anecdotes

Despite the impressive performance of ML models [for streamflow prediction] … they have not been widely adopted by government agencies and practitioners …

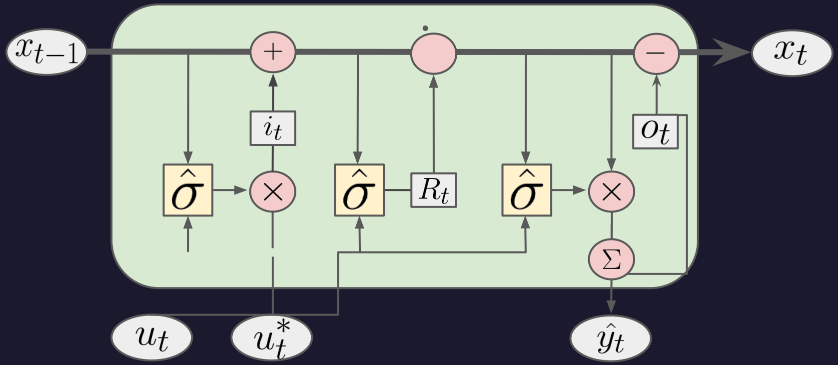
While this is due in part to the AI community's failure to engage with the hydrologic community, [collaborative efforts are emerging] ***leading to the integration of process knowledge with AI technologies****.*

– Unnamed Hydrology Research Group

Mass Balance

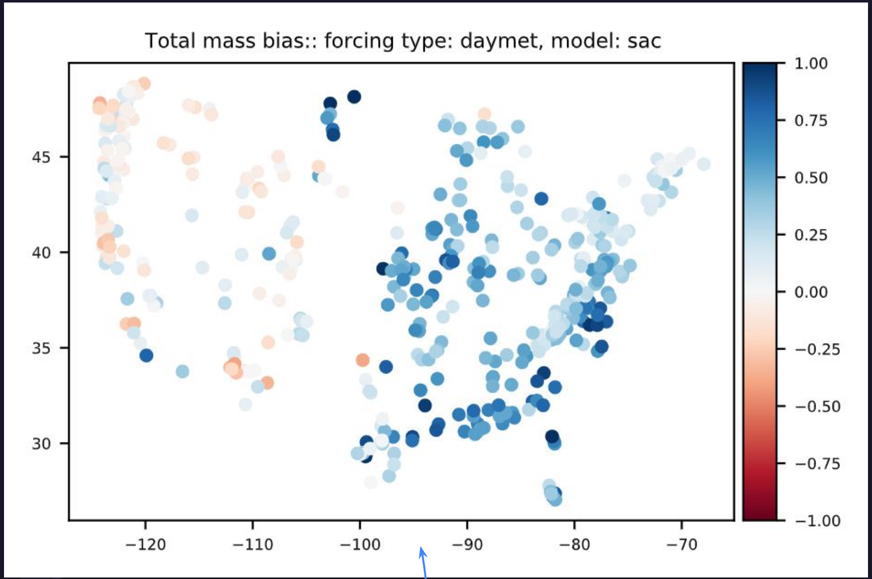


Adding a mass balance constraint does not improve even out-of-sample predictions.

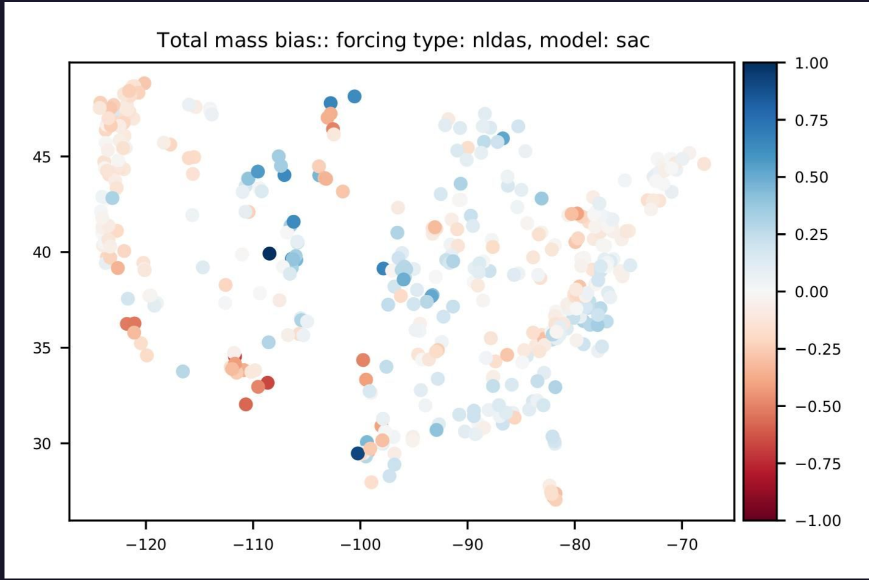


Mass Balance

DayMet Precipitation



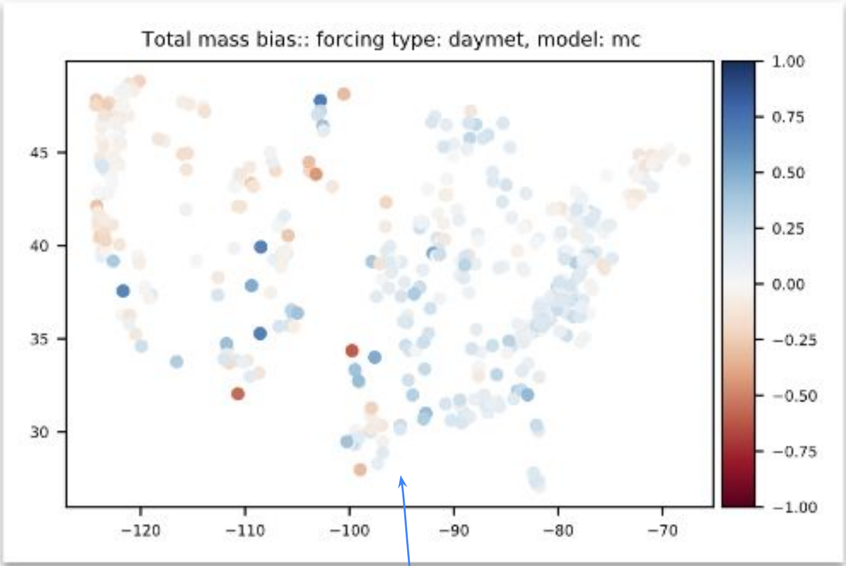
NLDAS Precipitation



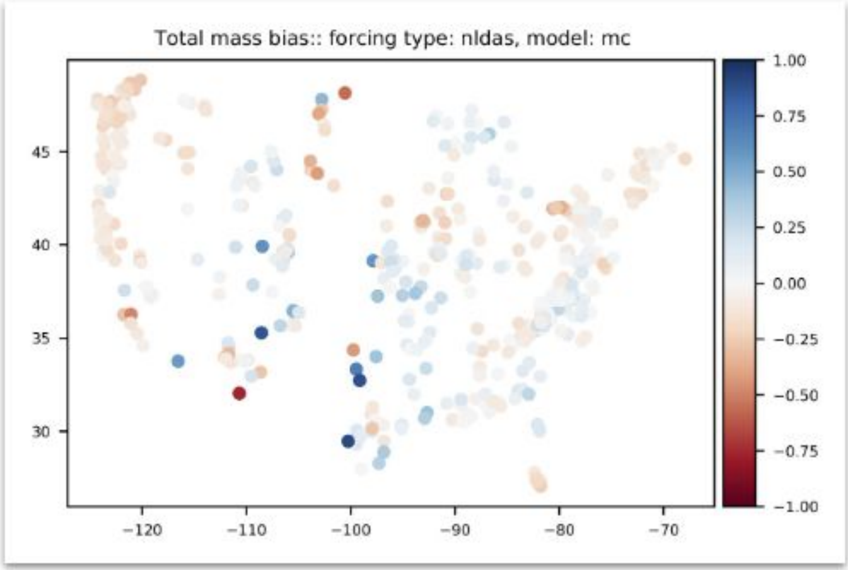
long-term bias in traditional hydrology model

Mass Balance

DayMet Precipitation



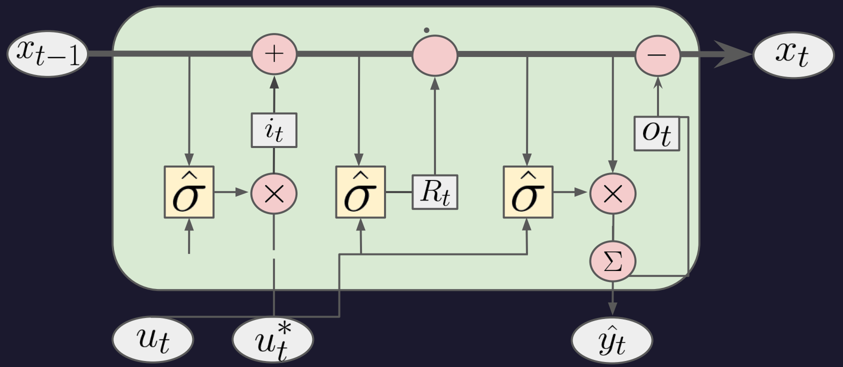
NLDAS Precipitation



Long-term bias is not present in
ML predictions.

Adding physics into ML streamflow models ... does not help (so far)

- Use physical models as inputs to ML models
- Add physical symmetries to ML model structures
- Regularize loss function
- Augmented / synthetic targets



“Isomorphisms” between physics and ML

“Building Blocks” of Physical Intuition:

- Spatial
- Locality
- Non-locality
- Monotonicity
- Causality
- Symmetry (conservation)
- Generalizability

Types of Inductive Bias

- Convolution (spatial & temporal)
- Recurrence (temporal)
- Attention
- Signed Activations
- Networks
- Regularization & Scaling
- Transfer Learning, Few Shot Learning, etc.

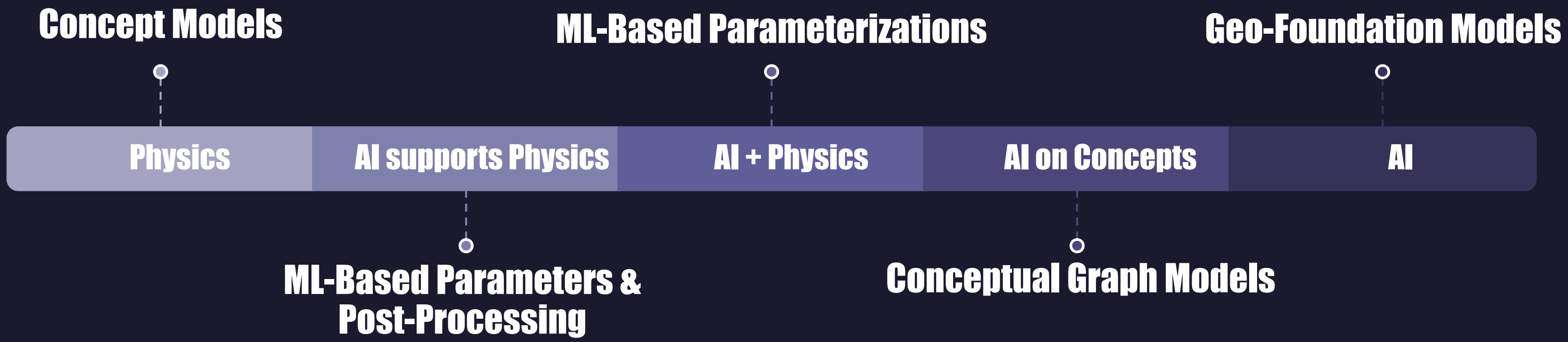
What are the core challenges?

What are the core challenges?

(1) Can physics constrain imperfect optimization?

(2) Can physics regularize outside of the training envelope?

Where is data-based Geoscience modeling headed?



What is a model?

Nearing, Grey S., and Hoshin V. Gupta. "Ensembles vs. information theory: supporting science under uncertainty." *Frontiers of Earth Science* (2018).