



Peishi Jiang

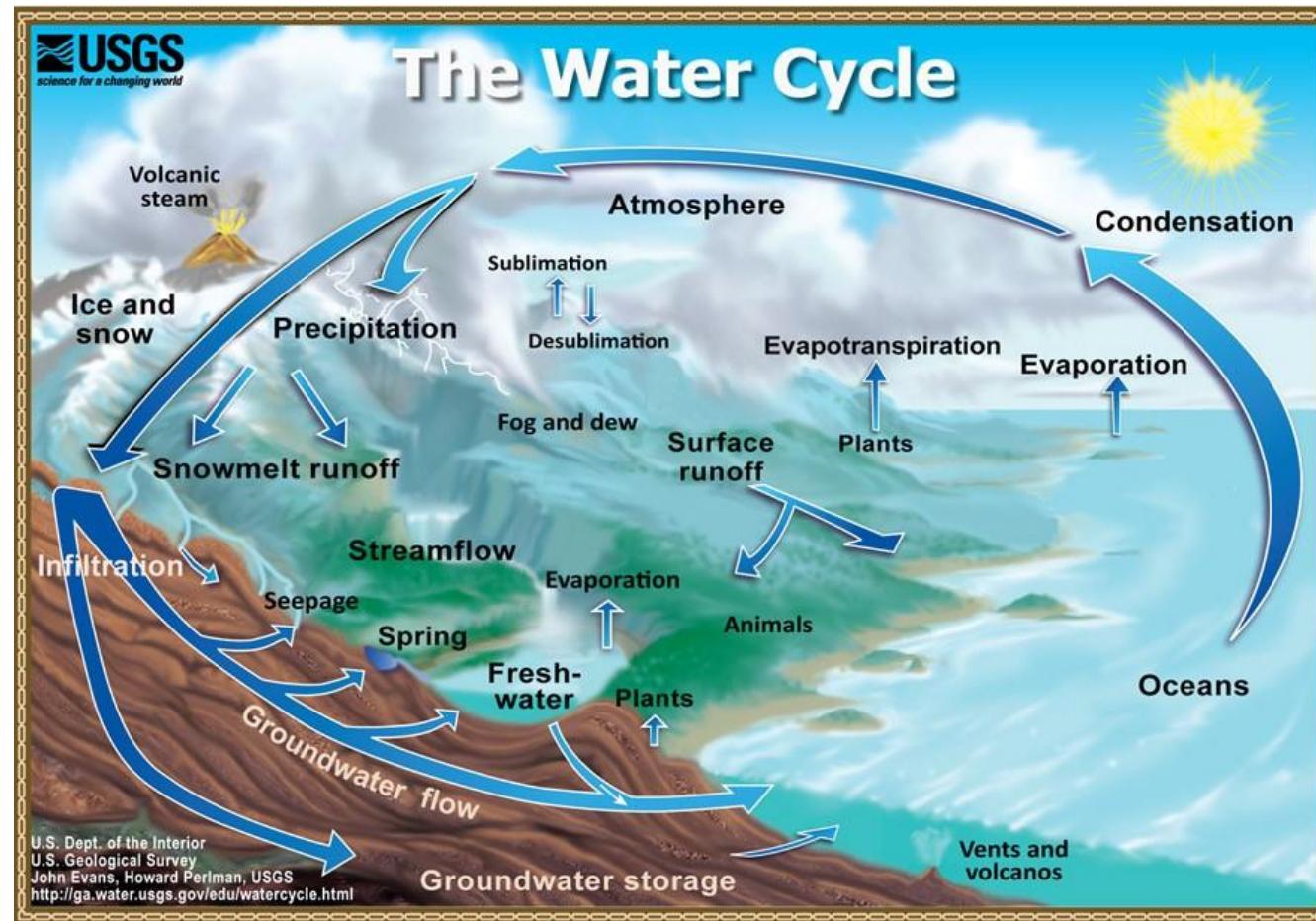
Pacific Northwest National Laboratory

Machine Learning (ML)
Applications in Hydrology

Hydrology: the study of water

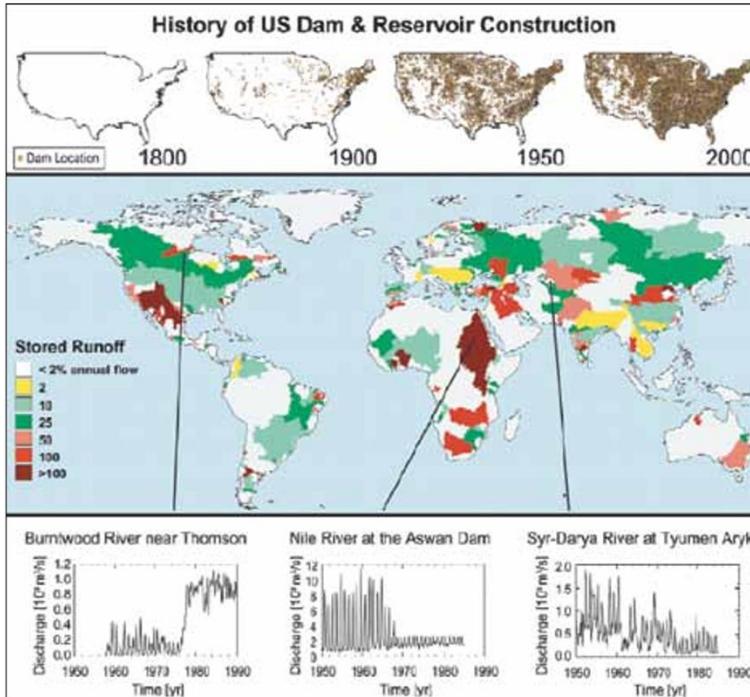
“Hydrology is the science that encompasses the occurrence, distribution, movement and properties of the waters of the earth and their relationship with the environment within each phase of the hydrologic cycle.”

by United States Geological Survey

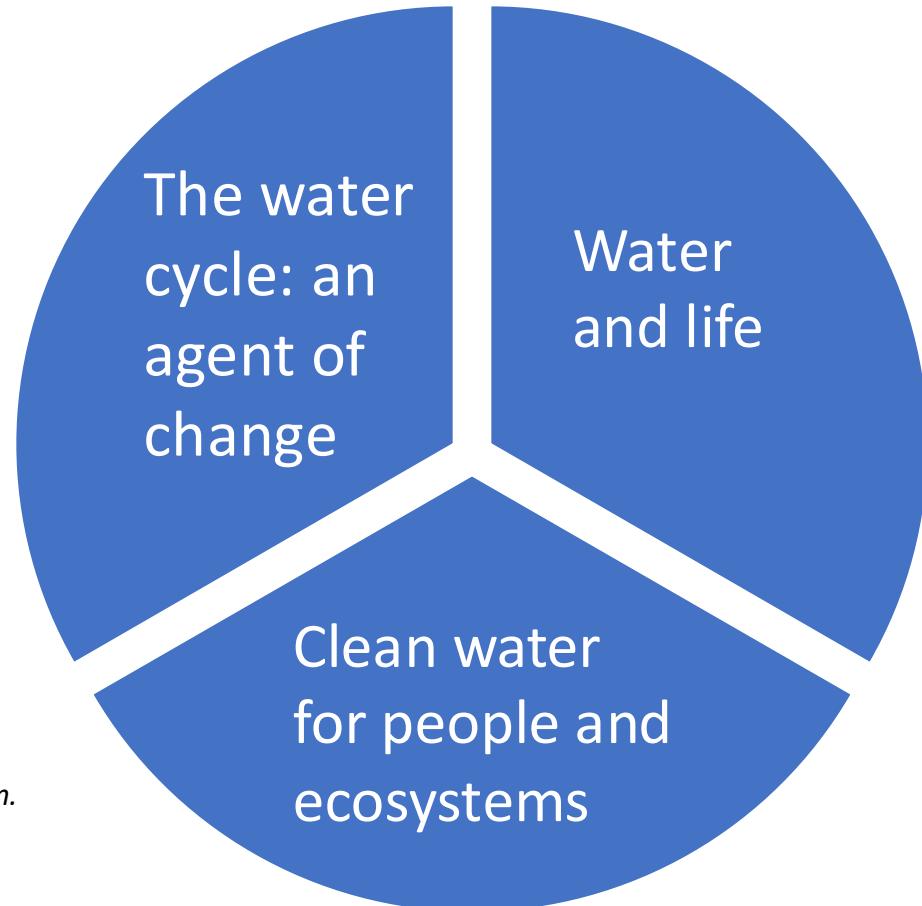


What are the grand challenges in Hydrology?

History of US Dam & Reservoir Construction



SOURCE: Vörösmarty et al. (2004) American Geophysical Union.



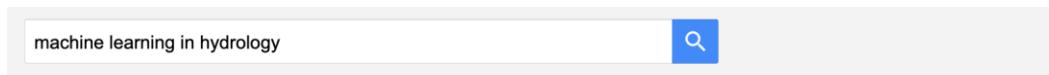
National Research Council, et al.
"Challenges and opportunities in the hydrologic sciences." (2012).

Flowing meltwater on the surface of the Greenland ice sheet



SOURCE: Roger Braithwaite, University of Manchester and Specialiststock.com.

How have Machine Learning (ML) been applied in Hydrology?



About 171,000 results from Google search

Machine learning for hydrologic sciences: An introductory overview

T Xu, F Liang - Wiley Interdisciplinary Reviews: Water, 2021 - Wiley Online Library

... machine learning algorithms and deep learning ... machine learning in hydrologic sciences.

Finally, we conclude with challenges associated with applying machine learning for hydrologic ...

☆ Save 99 Cite Cited by 116 Related articles All 11 versions ☺

Machine learning applications in hydrology

H Lange, S Sippl - Forest-water interactions, 2020 - Springer

... In this section, we provide a short overview on some important machine learning techniques used in hydrology. Given the pace of development in this field, the overview is necessarily ...

☆ Save 99 Cite Cited by 125 Related articles All 5 versions ☺

Ensemble machine learning paradigms in hydrology: A review

M Zounemat-Kermani, O Batelaan, M Fadaee... - ... of Hydrology, 2021 - Elsevier

... generalized stacked, in different application fields of hydrology. The main hydrological topics in this review study cover subjects such as surface hydrology, river water quality, rainfall-...

☆ Save 99 Cite Cited by 383 Related articles All 7 versions

Application of machine learning algorithms in hydrology

H Mosaffa, M Sadeghi, I Mallakpour... - Computers in earth and ..., 2022 - Elsevier

... in hydrology. The Science Direct and Springer databases were used to investigate the number of studies in the subfield of hydrology with machine learning and hydrology keywords for ...

☆ Save 99 Cite Cited by 49 Related articles ☺

Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review

S Ardabili, A Mosavi, M Dehghanj... - ... for Sustainable Future ..., 2020 - Springer

... of machine learning and deep learning methods ... machine learning and deep learning are presented through a novel classification of methods. The paper concludes that deep learning ...

☆ Save 99 Cite Cited by 170 Related articles All 8 versions ☺

Broadening the use of machine learning in hydrology

C Shen, X Chen, E Laloy - Frontiers in Water, 2021 - frontiersin.org

... machine learning in various subdomains of hydrology. In this Research Topic, we sought to broaden the use of machine learning (ML) in hydrology ... applications of machine learning in ...

☆ Save 99 Cite Cited by 70 Related articles All 9 versions ☺

What role does hydrological science play in the age of machine learning?

GS Nearing, F Kratzert, AK Sampson... - Water Resources ..., 2021 - Wiley Online Library

... learning to rainfall-runoff simulation indicate that there is significantly more information in large-scale hydrological ... interest in machine learning in the hydrological sciences community, ...

[PDF] wiley.com

Full View

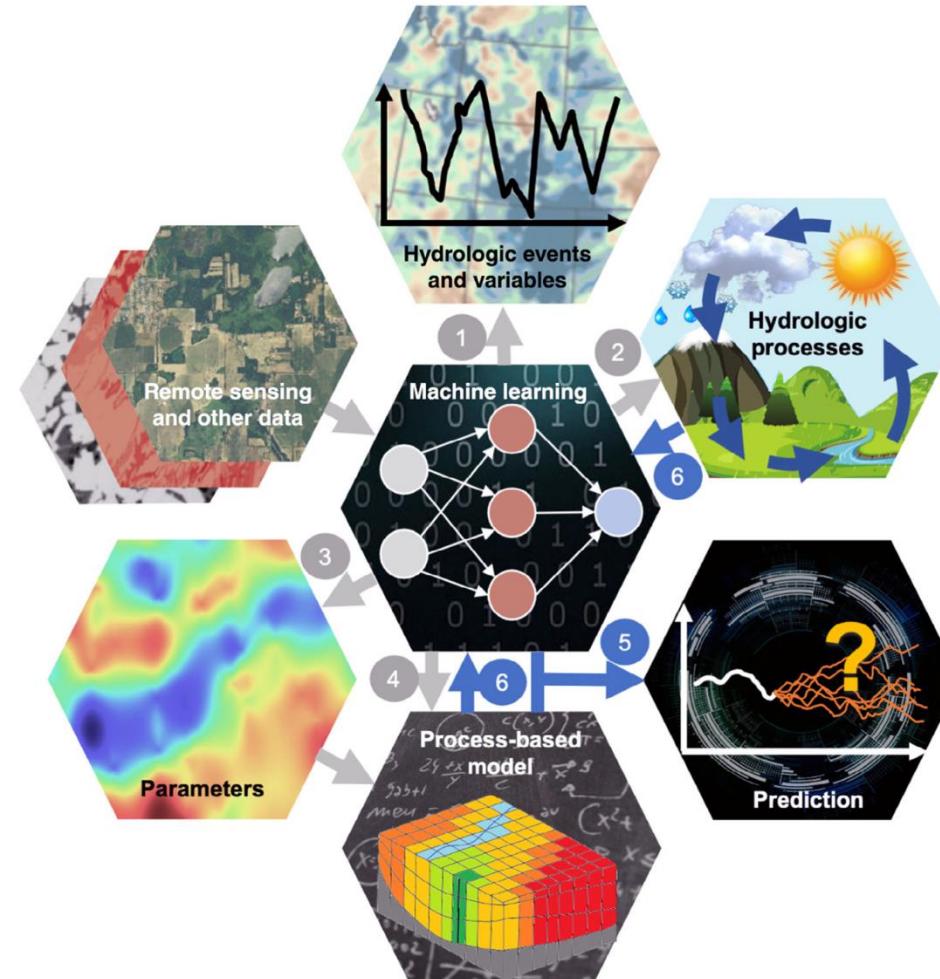
[PDF] researchgate.net

[HTML] sciencedirect.com
ViewIt@pnnl.gov

[PDF] preprints.org

[PDF] frontiersin.org
ViewIt@pnnl.gov

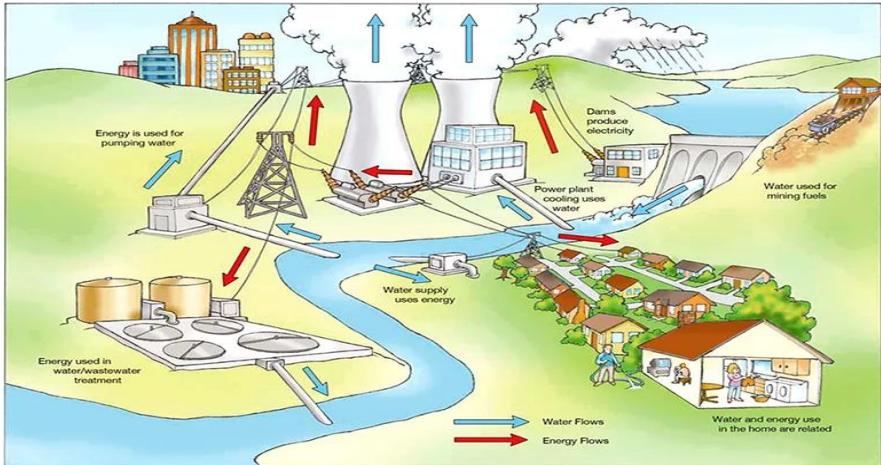
[PDF] wiley.com
Full View



Xu, Tianfang, and Feng Liang. "Machine learning for hydrologic sciences: An introductory overview." *Wiley Interdisciplinary Reviews: Water* 8.5 (2021): e1533.

Streamflow prediction

Water resources



Source: medium

Drought



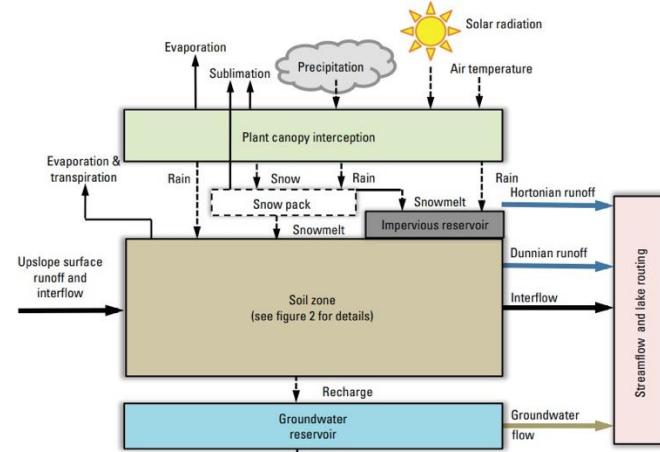
Source: Phys.org

Flooding

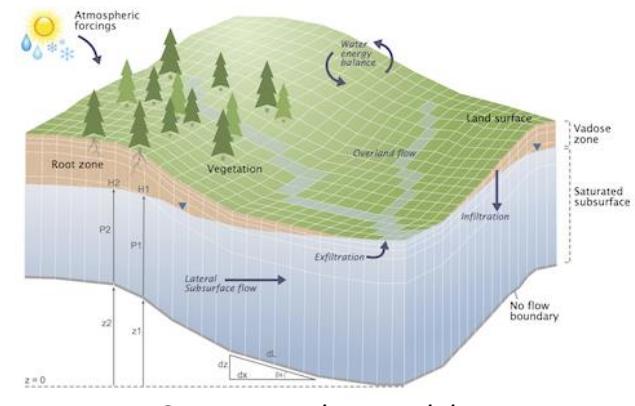


Source: CNN

Process-based hydrological models



Source: PRMS model



Source: ParFlow model

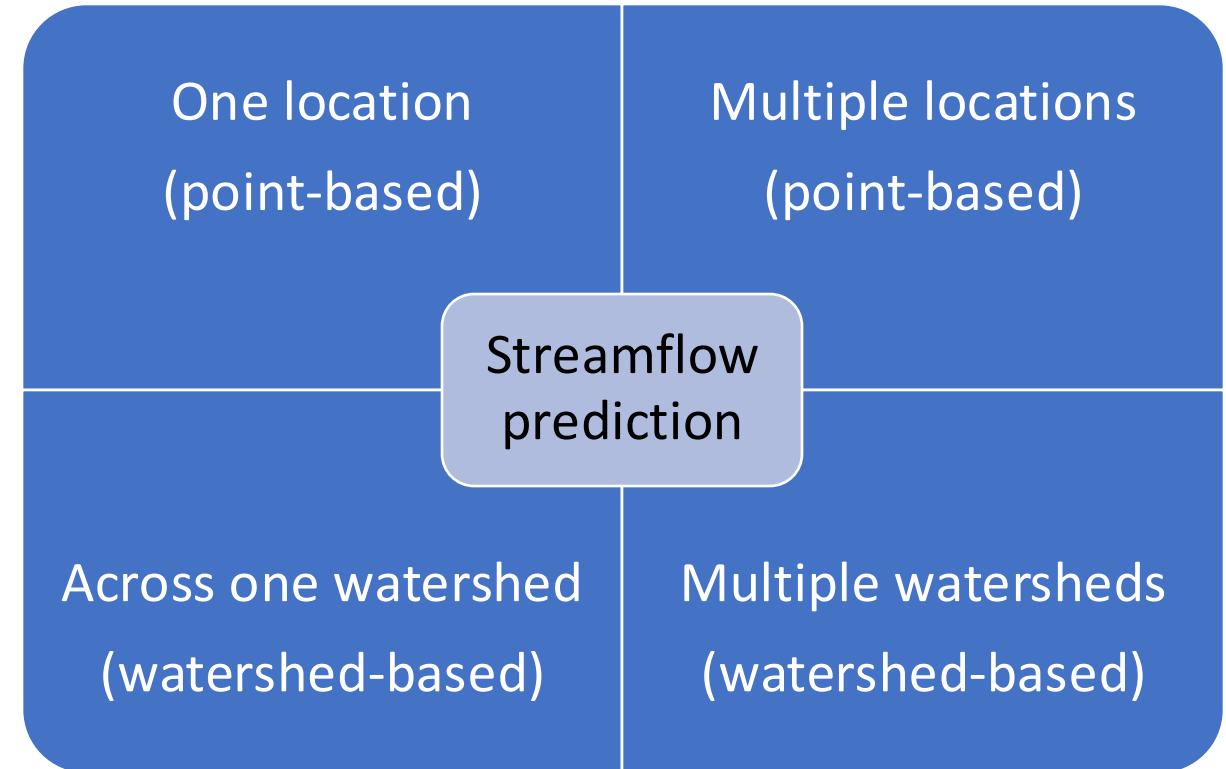
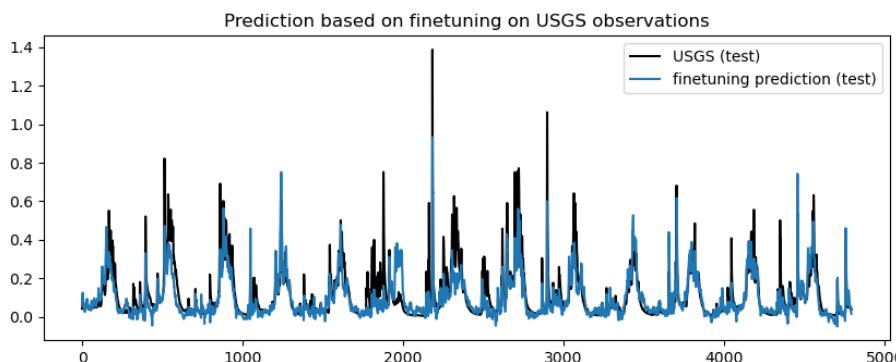
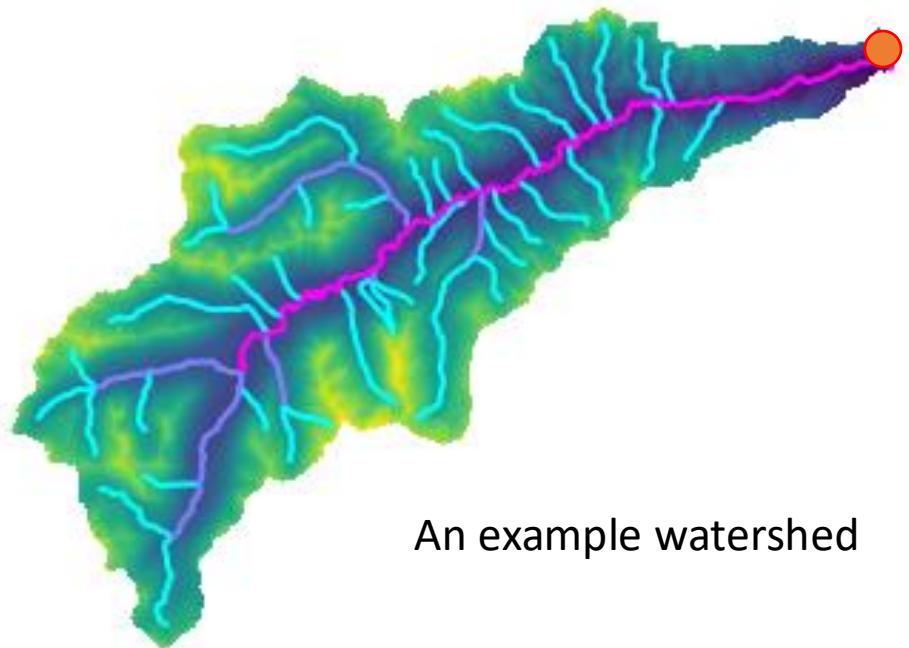
Bucket model:

- Simplified process
- Lack of accuracy

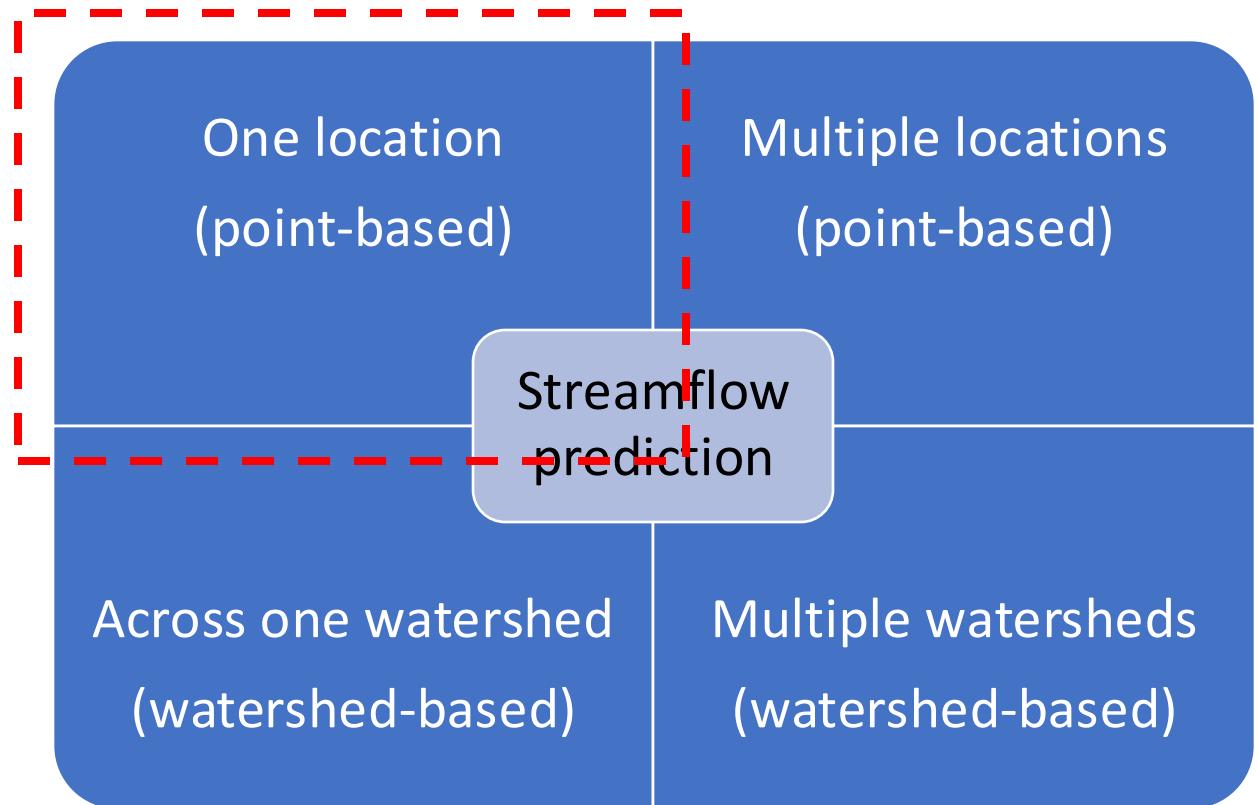
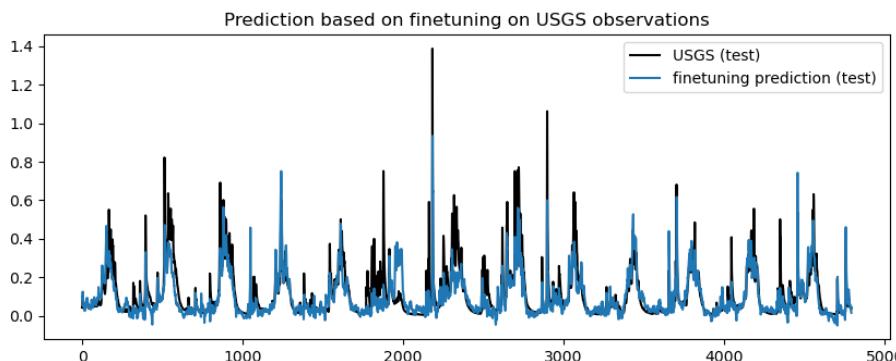
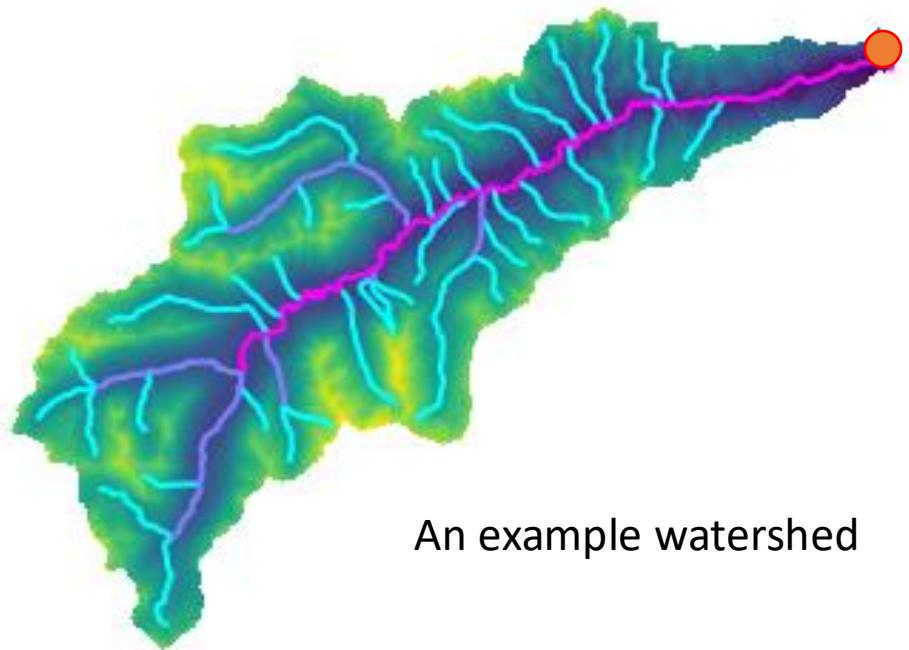
Fully integrated model:

- Physically accurate
- Computationally expensive

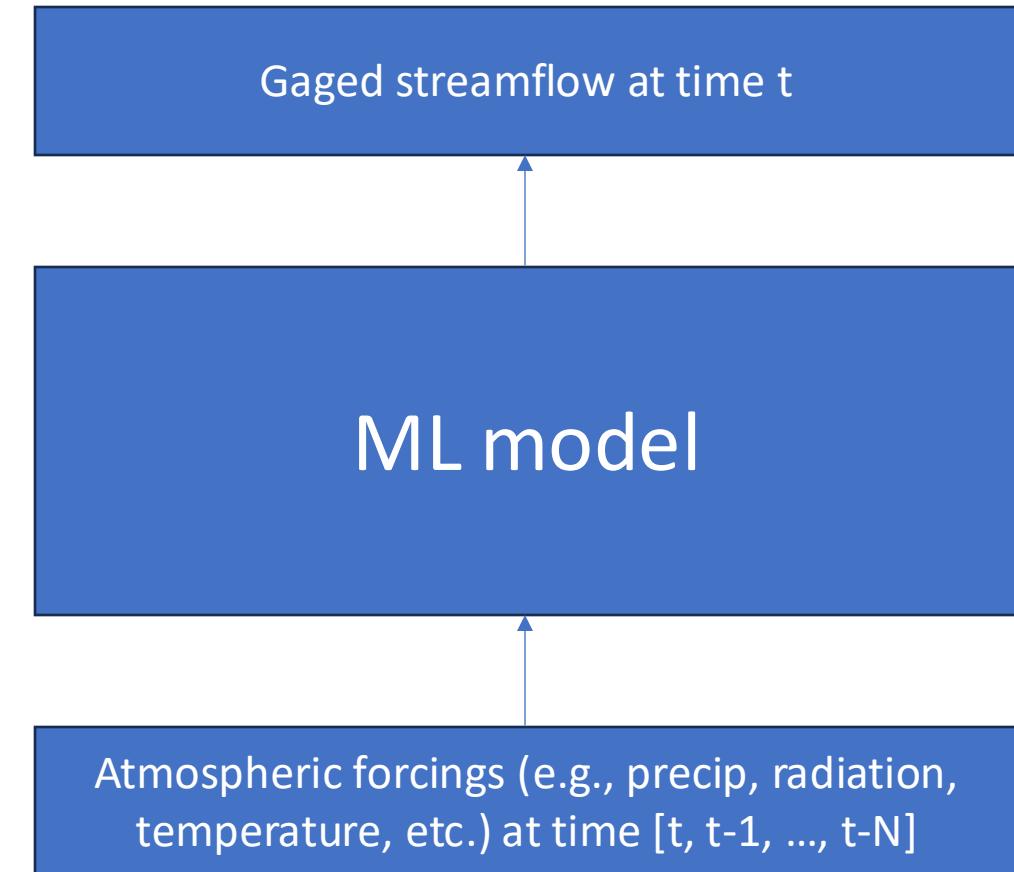
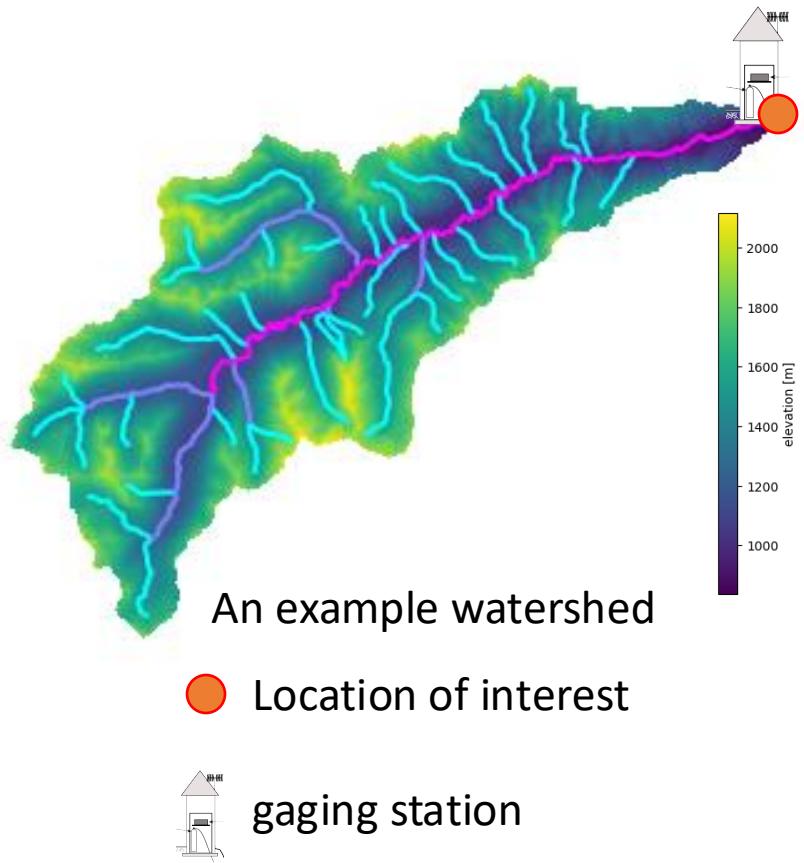
ML application in streamflow prediction



ML application in streamflow prediction



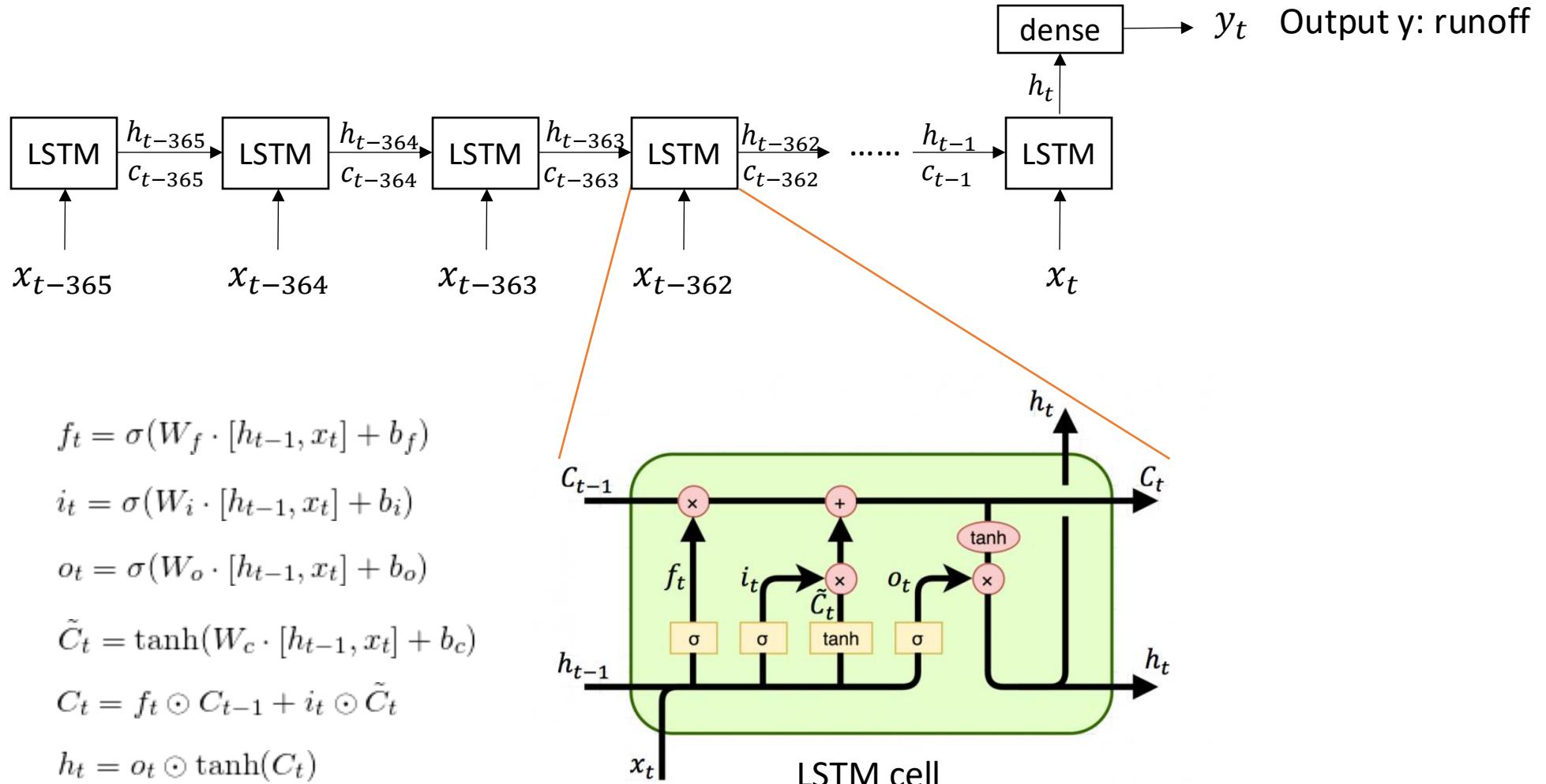
Streamflow prediction at one gauged location:



Streamflow prediction at one gauged location: Long Short-Term Memory model (LSTM)

Input x:

- precipitation
- temperature
- solar radiation
- vapor pressure

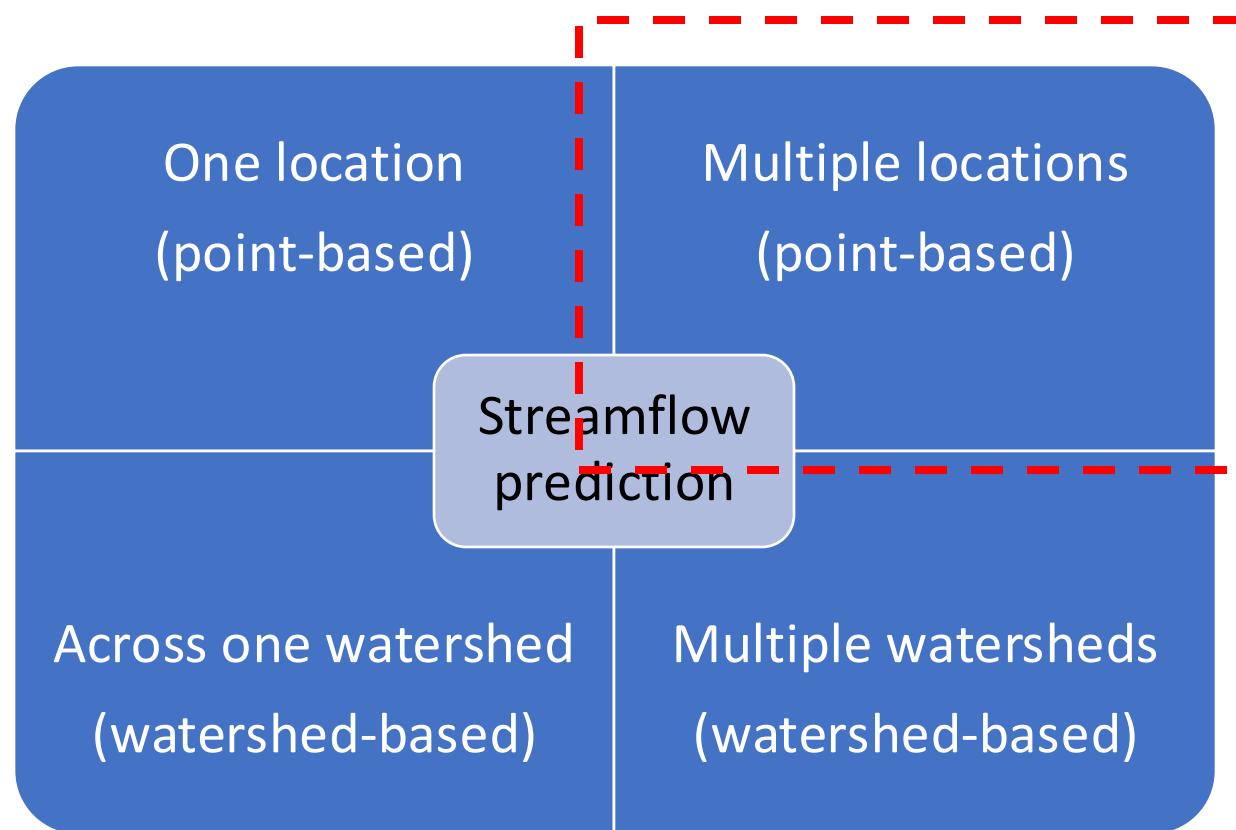
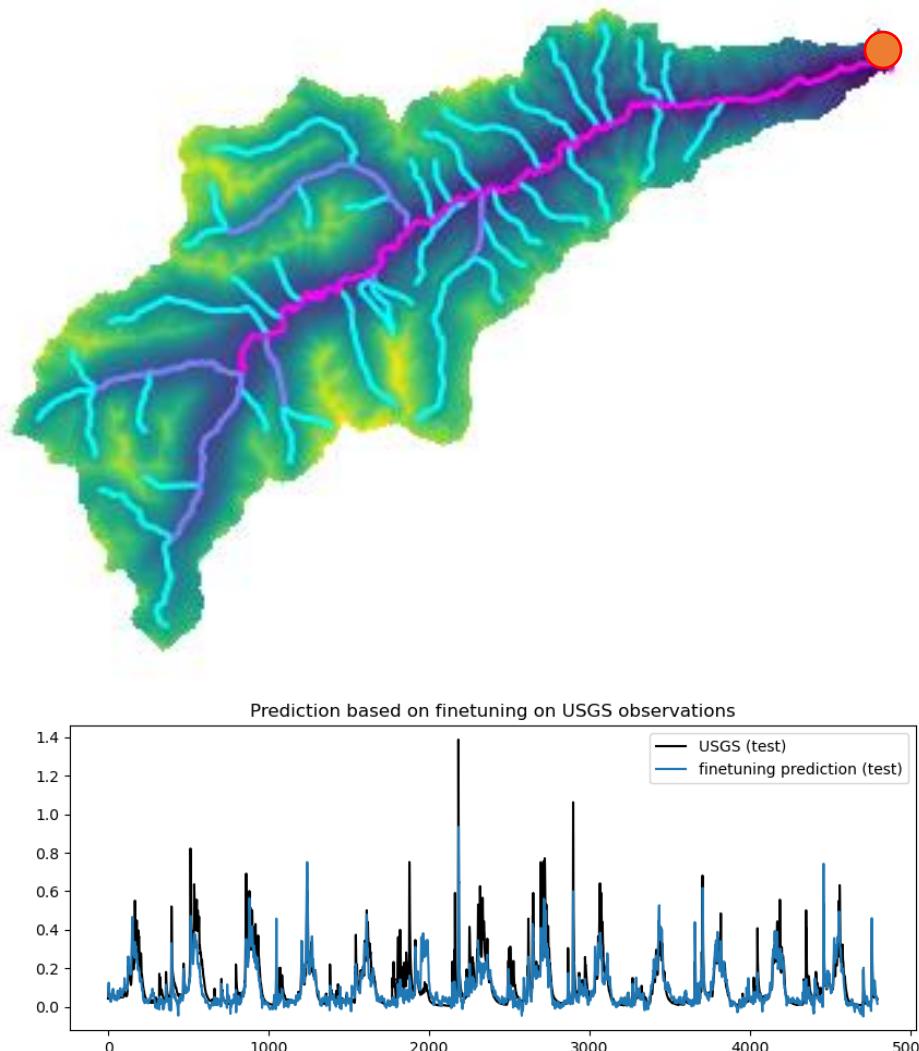


Streamflow prediction at one gauged location: Long Short-Term Memory model (LSTM)

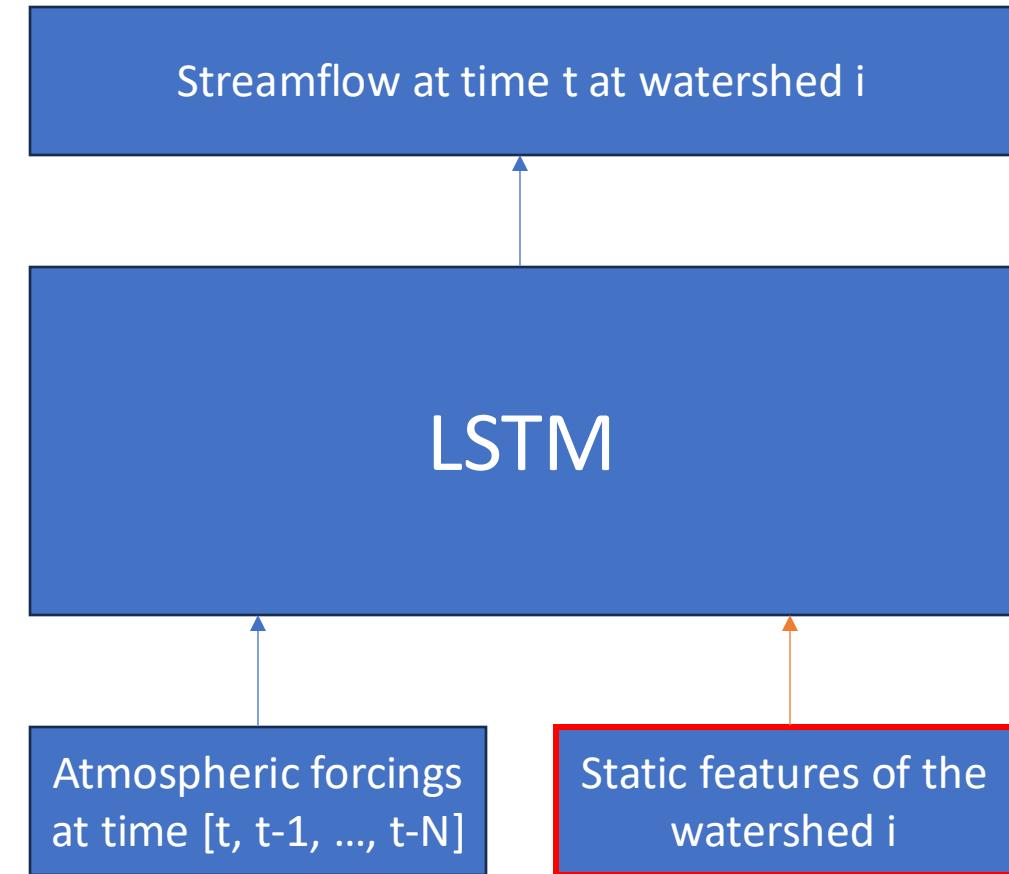
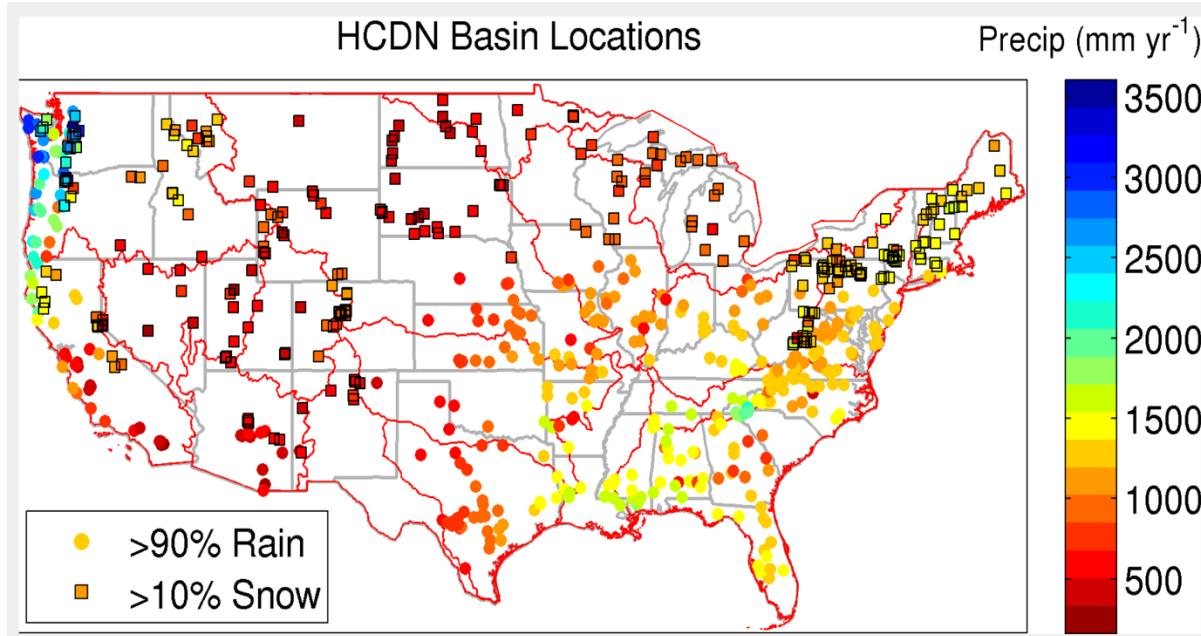


Jupyter Notebook

ML application in streamflow prediction

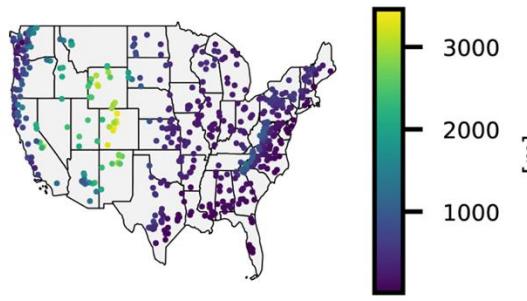


Streamflow prediction at multiple gauged locations: LSTM

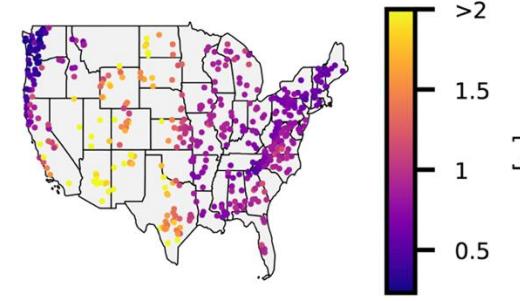


Streamflow prediction at multiple gauged locations: LSTM

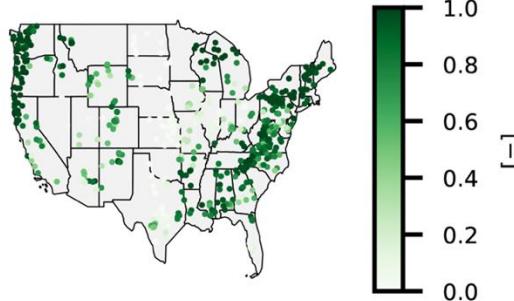
(a) Mean elevation



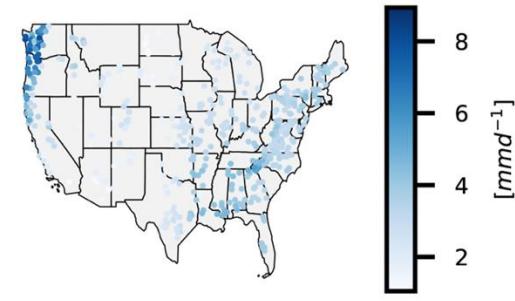
(b) Aridity



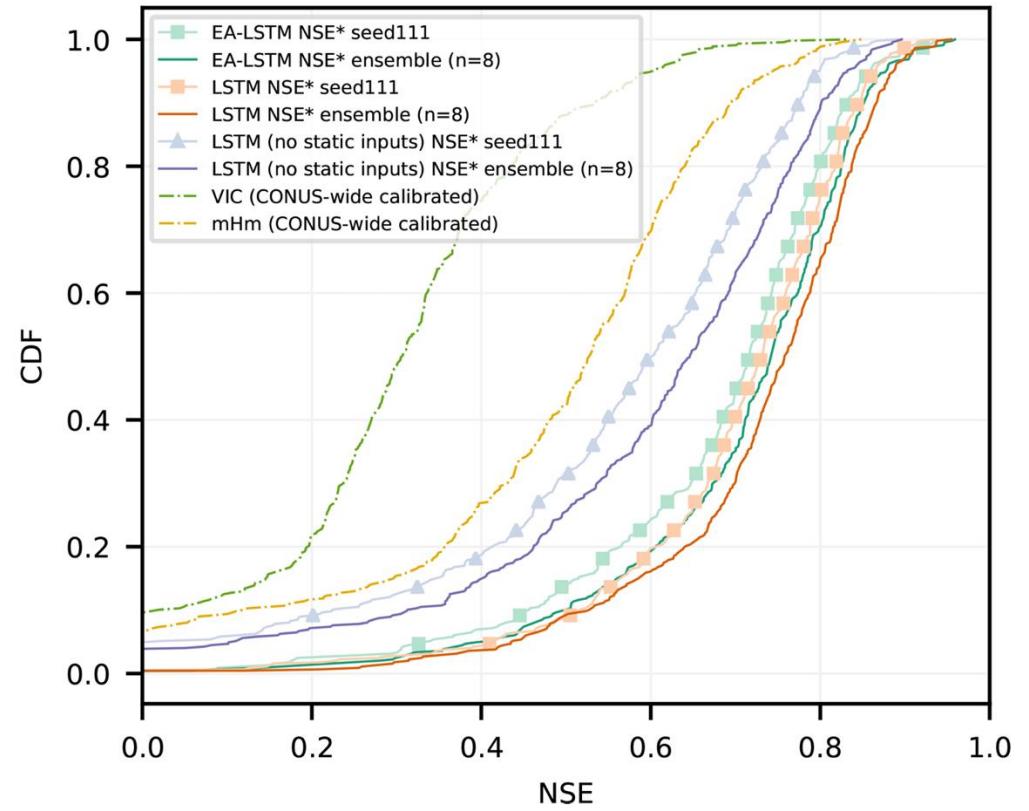
(c) Forest fraction



(d) Mean daily precipitation



Benchmarking vs CONUS-wide calibrated models

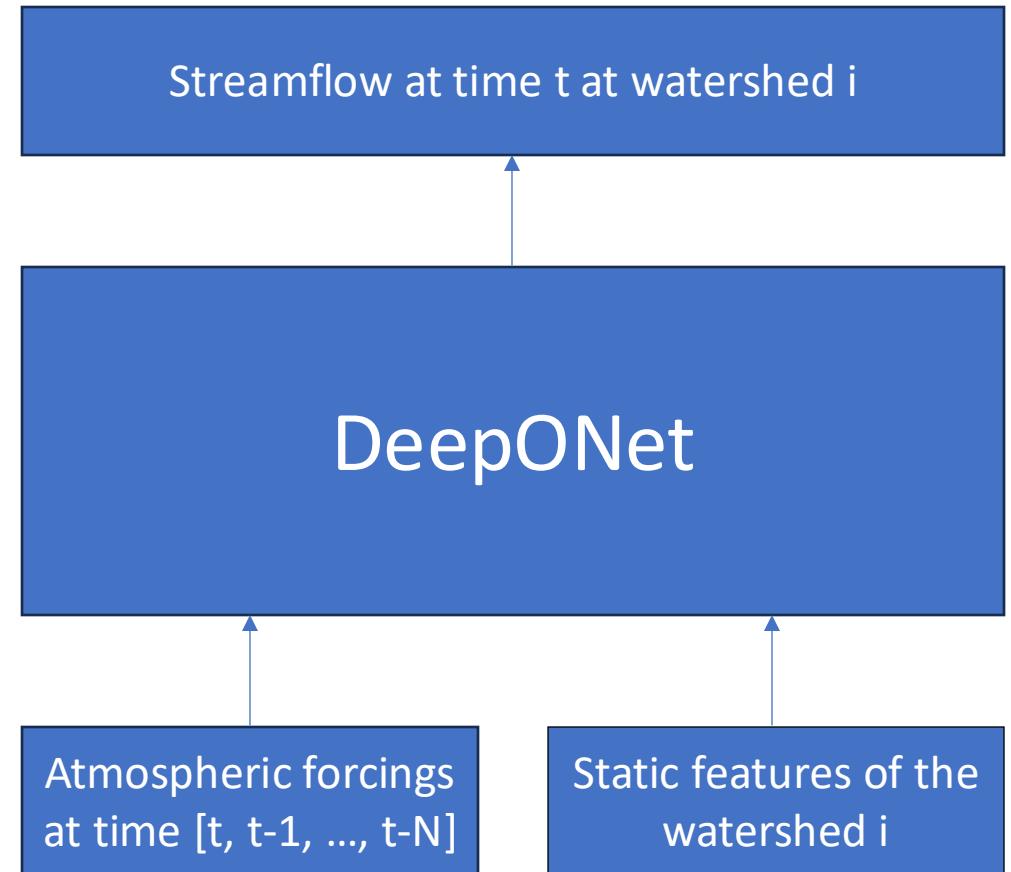
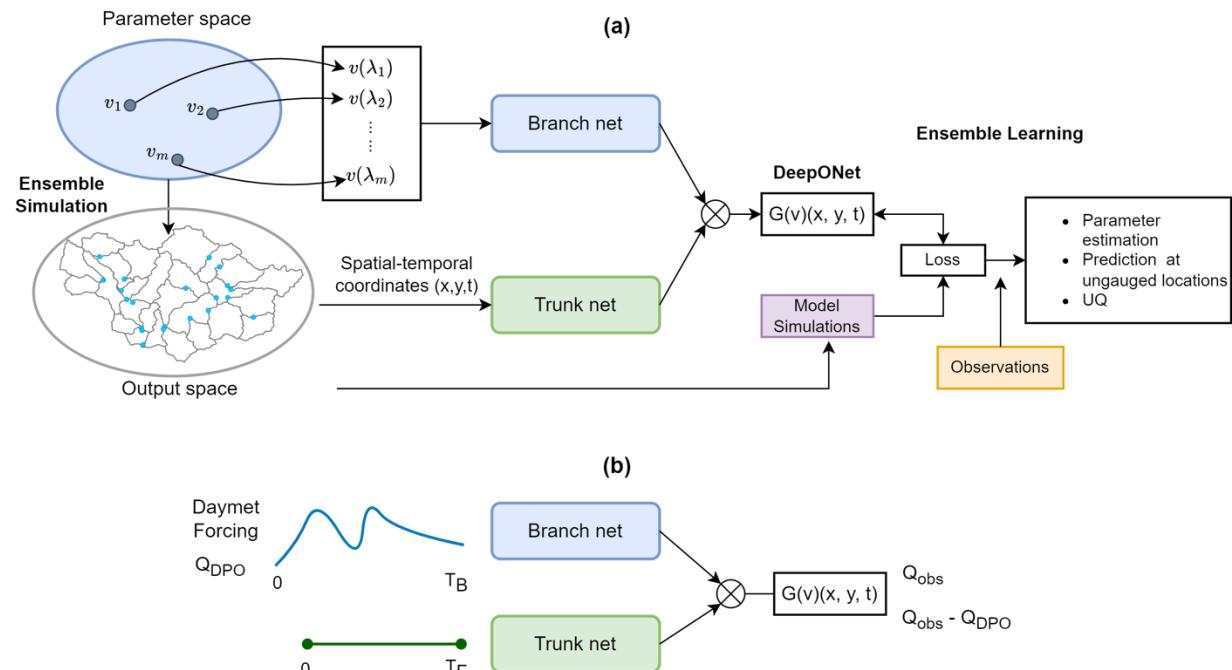


VIC, mHm: two hydrological models

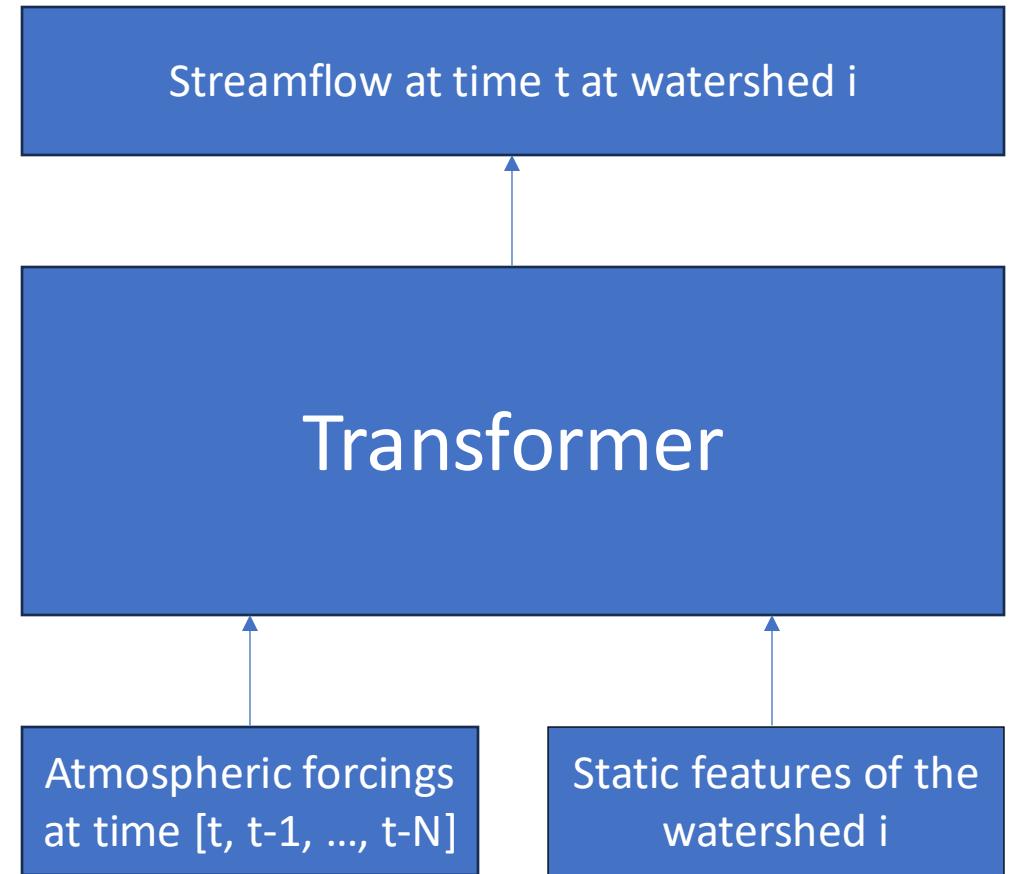
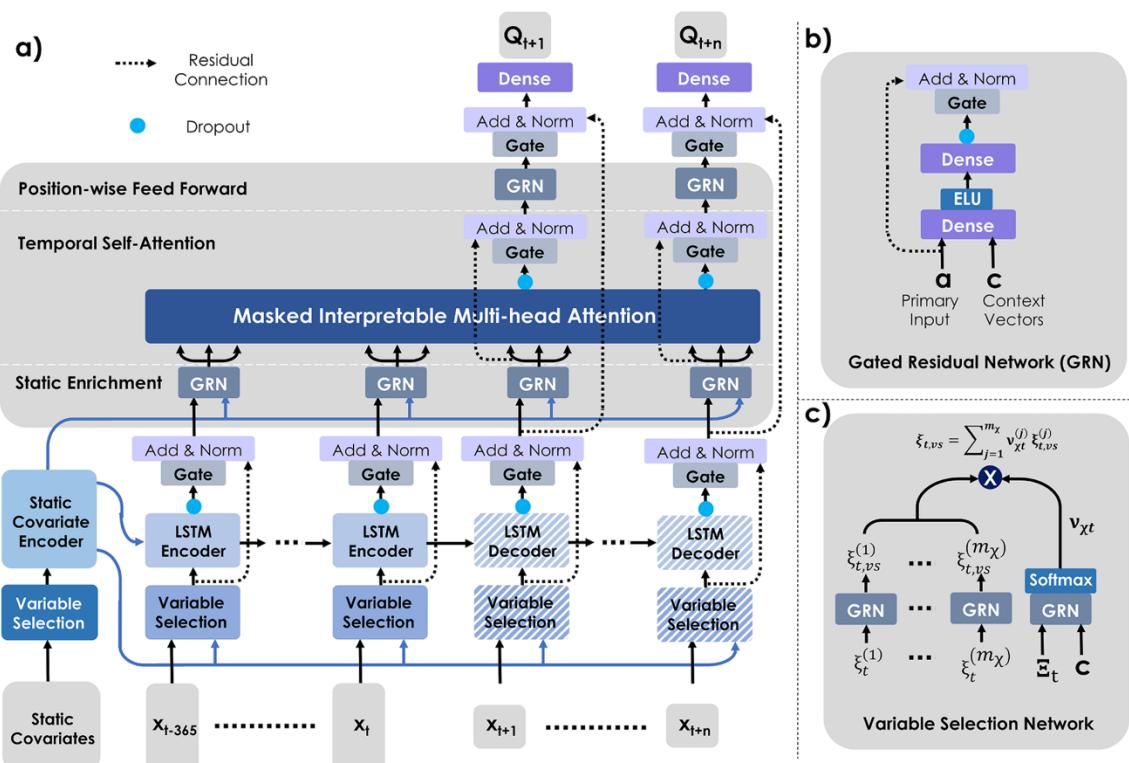
NSE: Nash–Sutcliffe model efficiency coefficient

CDF: Cumulative distribution function

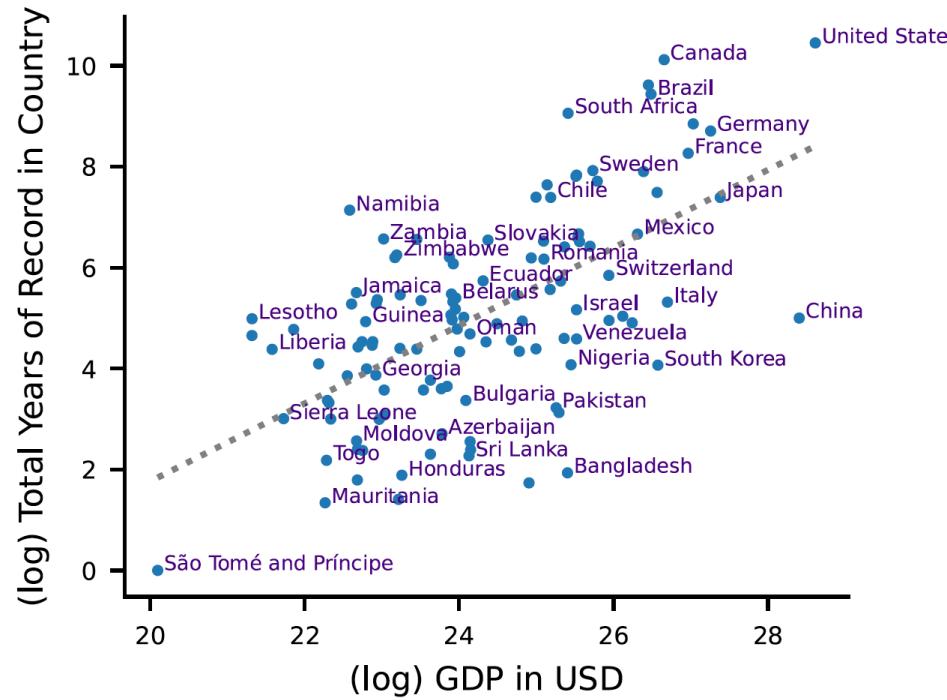
Streamflow prediction at multiple gauged locations: Other ML models



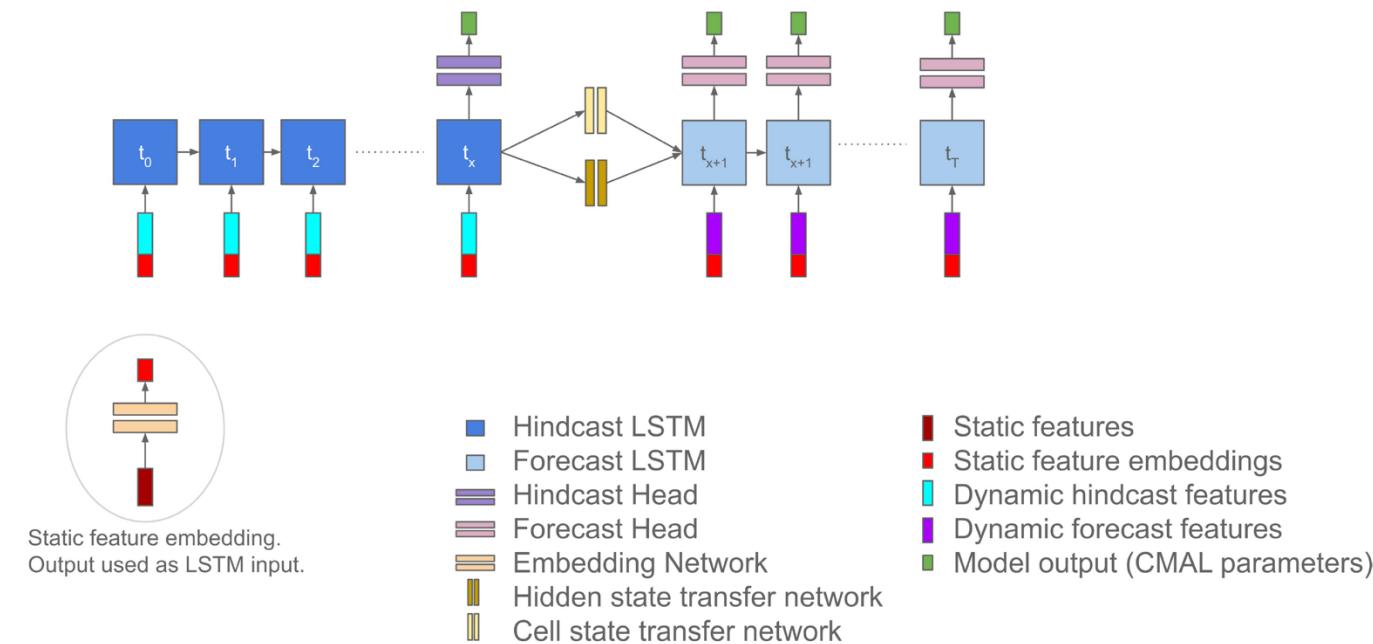
Streamflow prediction at multiple gauged locations: Other ML models



Streamflow prediction at ungauged locations: LSTM & transfer learning



Extended Data Fig. 1 | Streamflow data availability correlates with national GDP. There is a log-log correlation ($r=0.611$; $N=117$) between national Gross Domestic Product (GDP) and the total number of years worth of daily streamflow data available in a country from the Global Runoff Data Center. GDP data are sourced from The World Bank⁴⁶.



Architecture of the LSTM-based forecast model developed for this project

Streamflow prediction at ungauged locations: LSTM & transfer learning

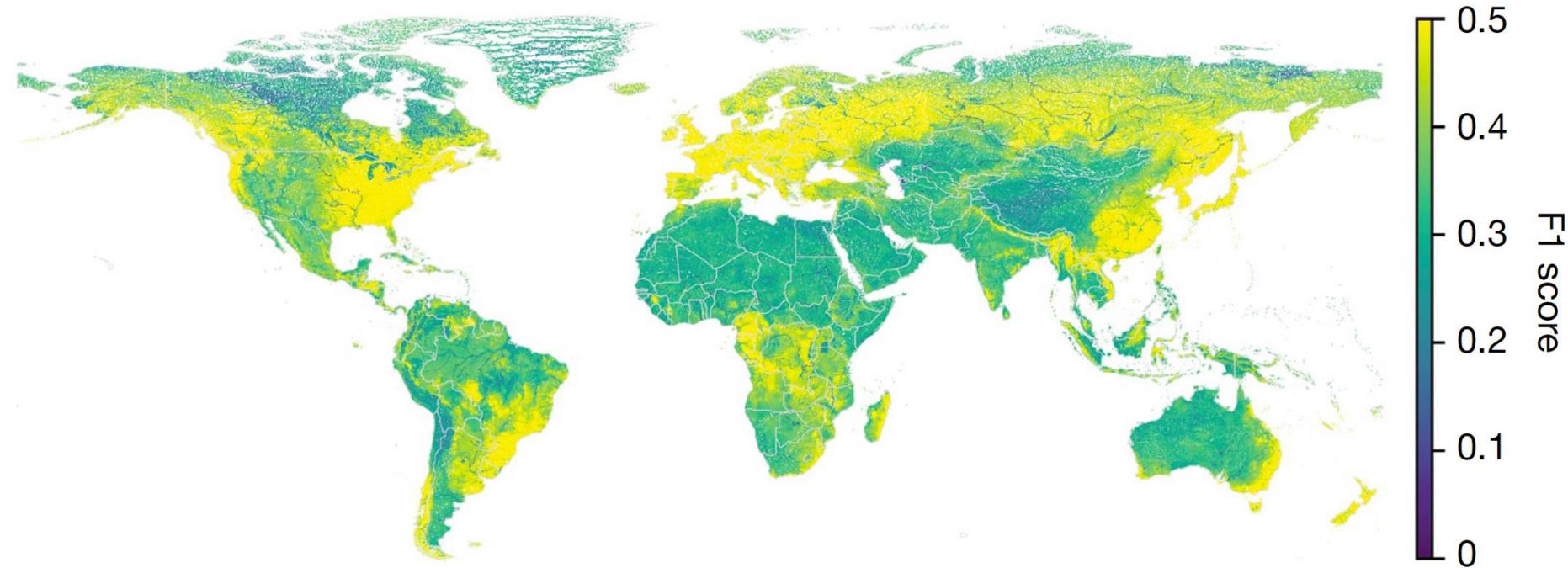
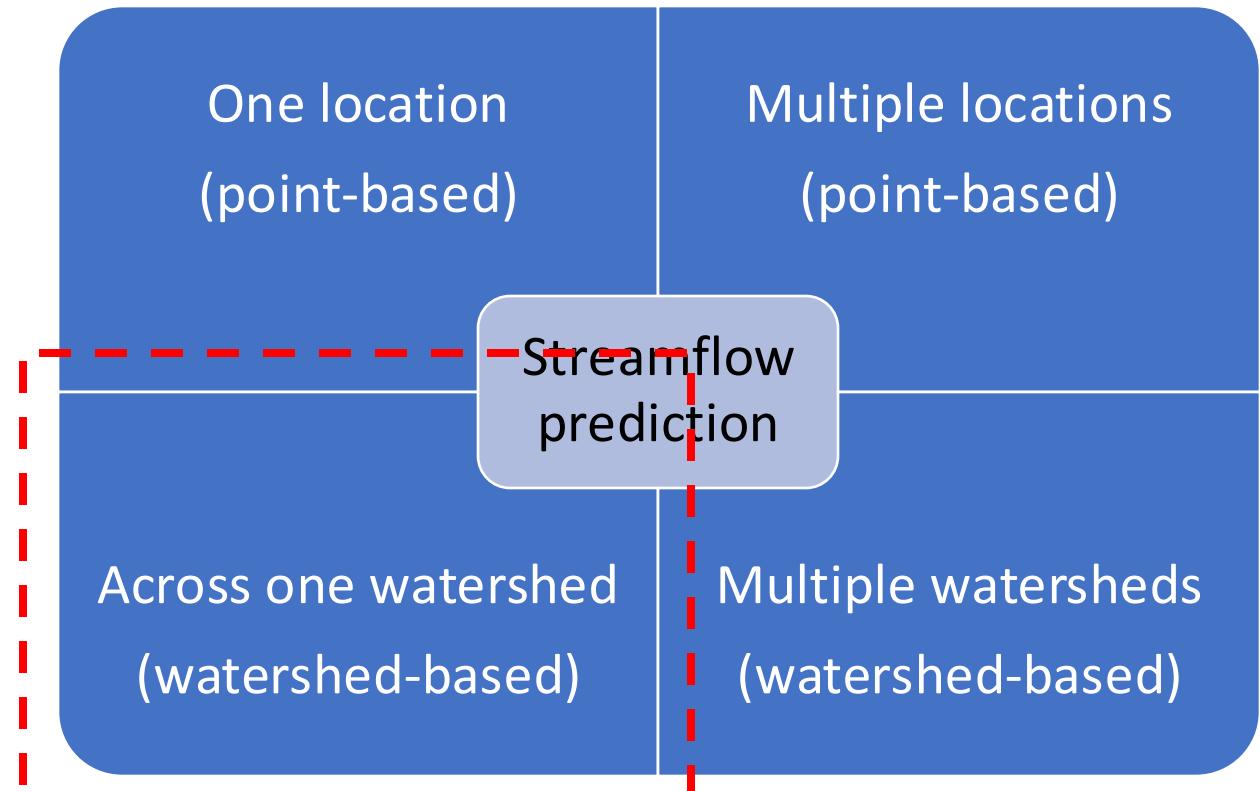
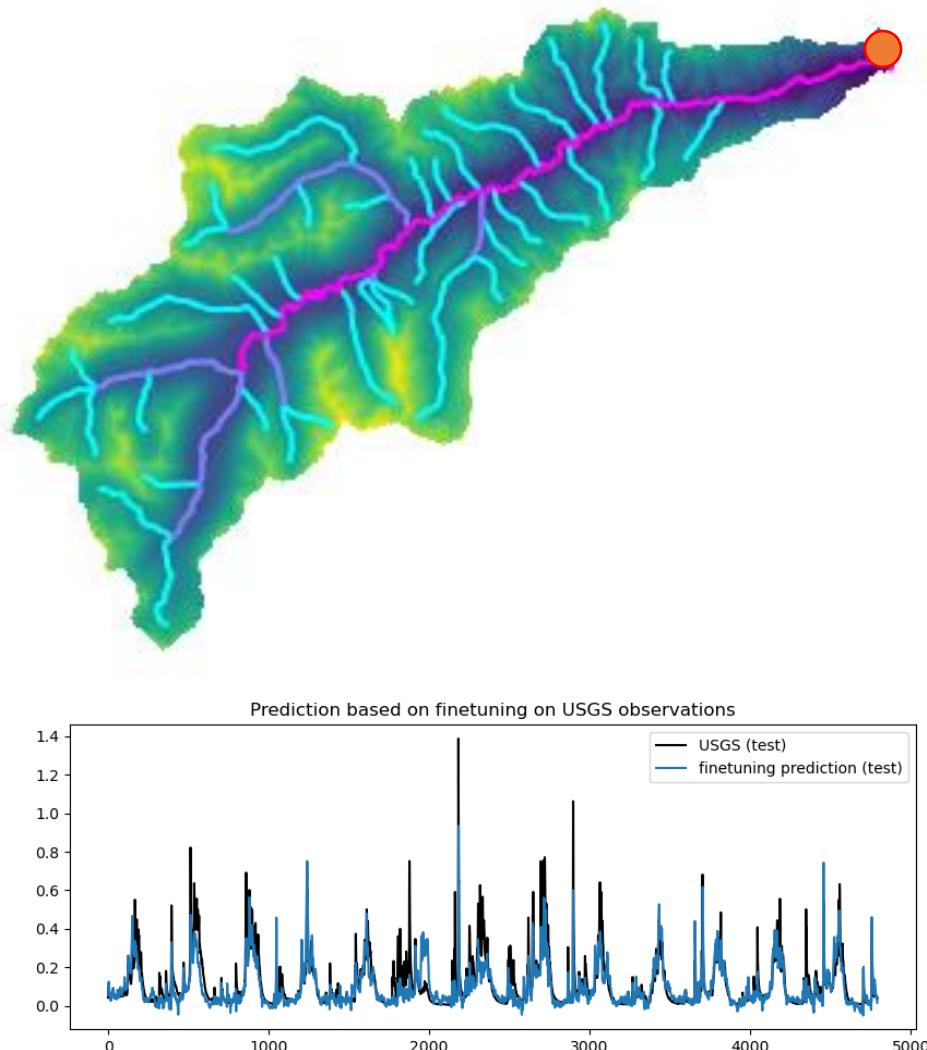


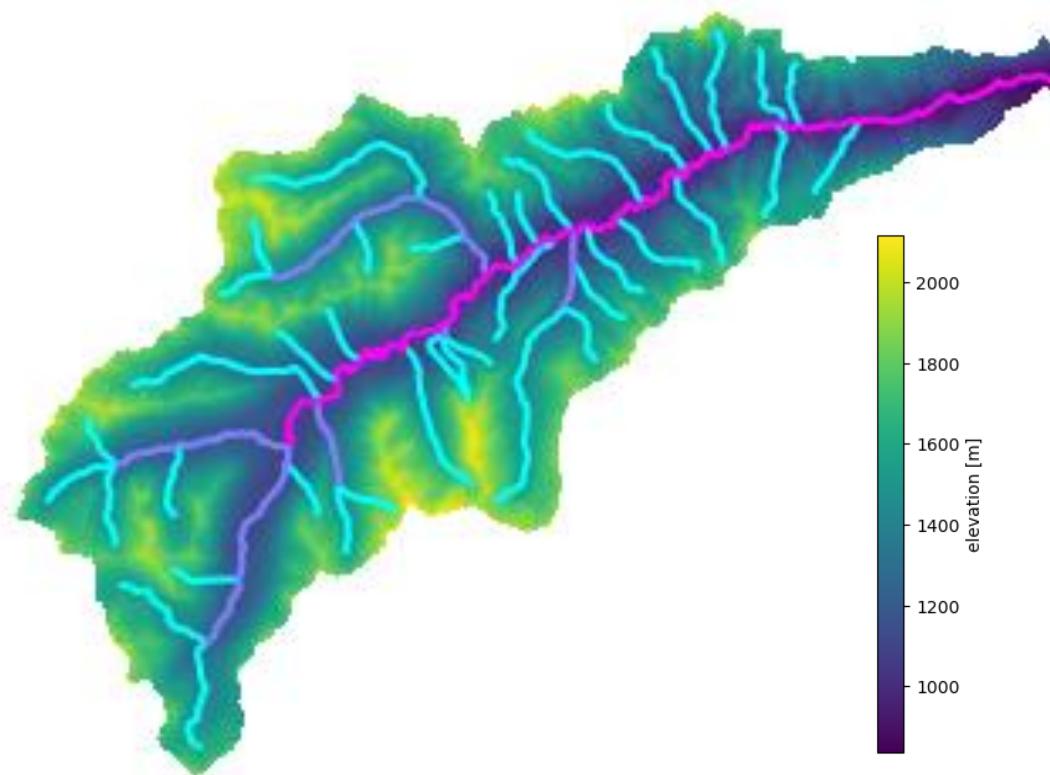
Fig. 6 | Global predicted skill. This map shows predictions of 2-year return period F1 scores over 1.03 million HydroBASINS level-12 watersheds for the AI forecast model. Basemap from GeoPandas³⁴.

F1 score: the harmonic mean of precision and recall

ML application in streamflow prediction

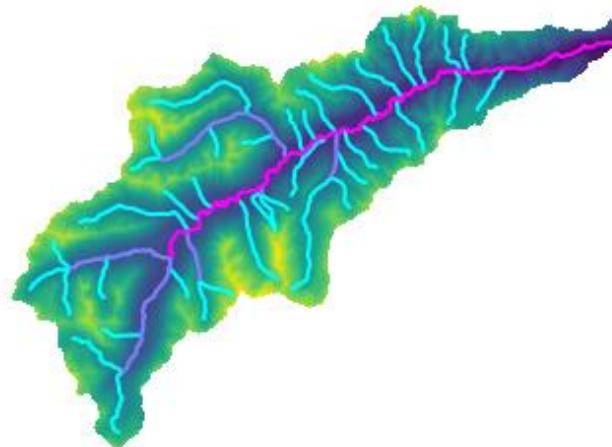
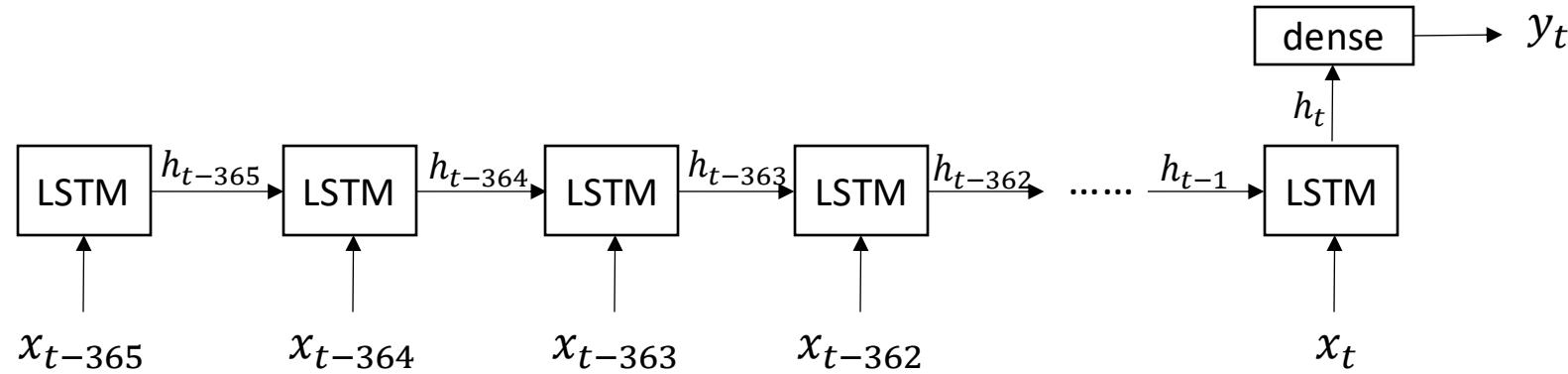


Streamflow prediction across a watershed



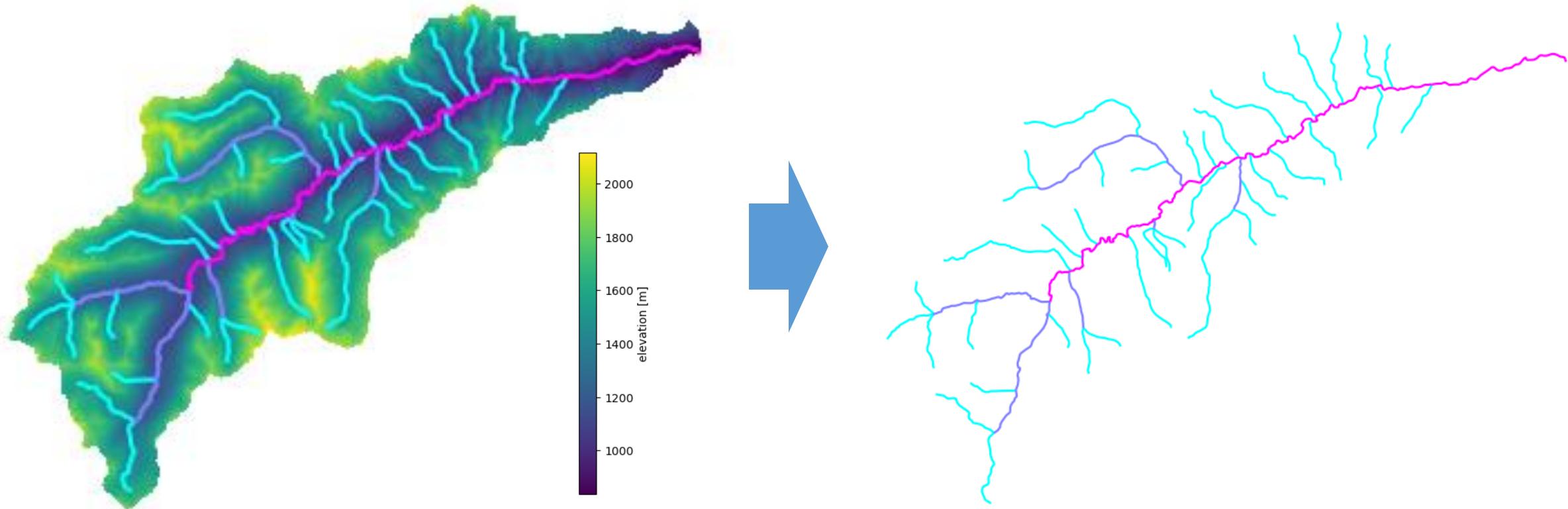
An example watershed

Streamflow prediction across a watershed: Issues with LSTM



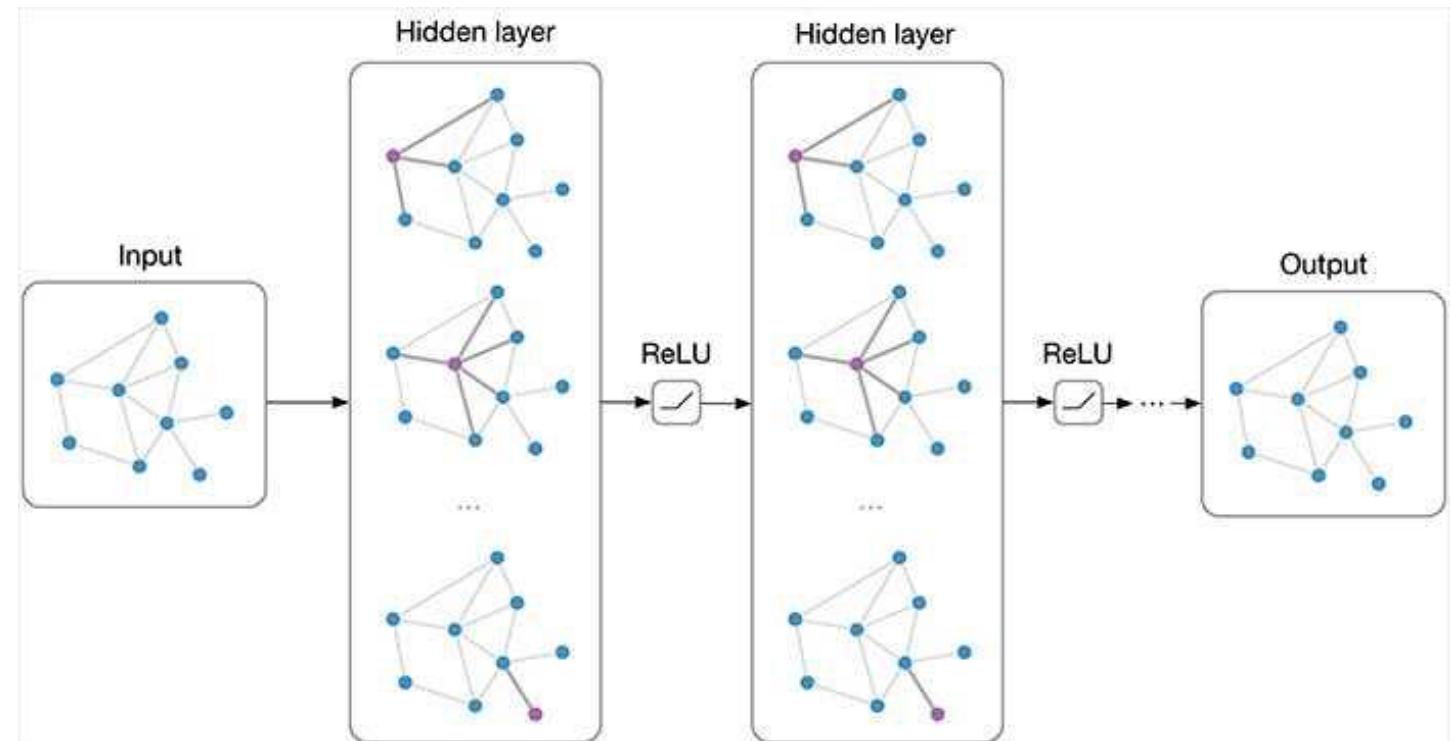
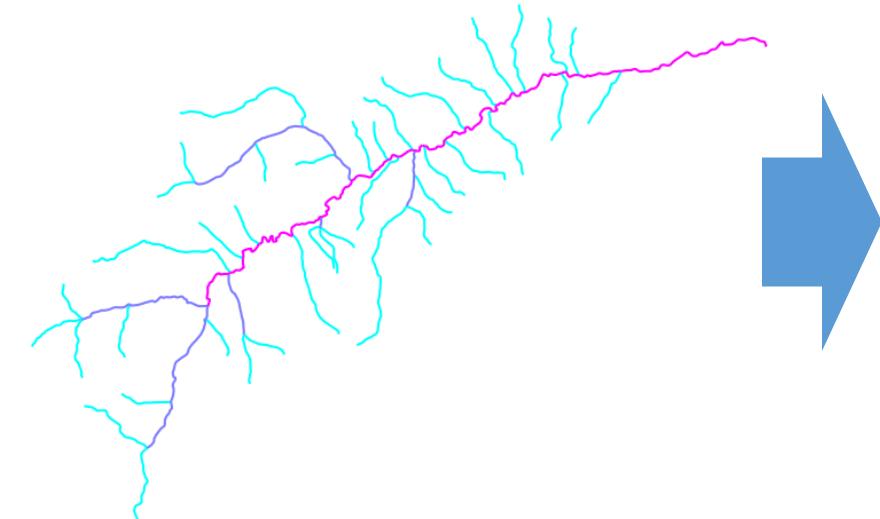
LSTM does not explicitly capture the physical connections between river reaches

Streamflow prediction across a watershed: river network as a graph



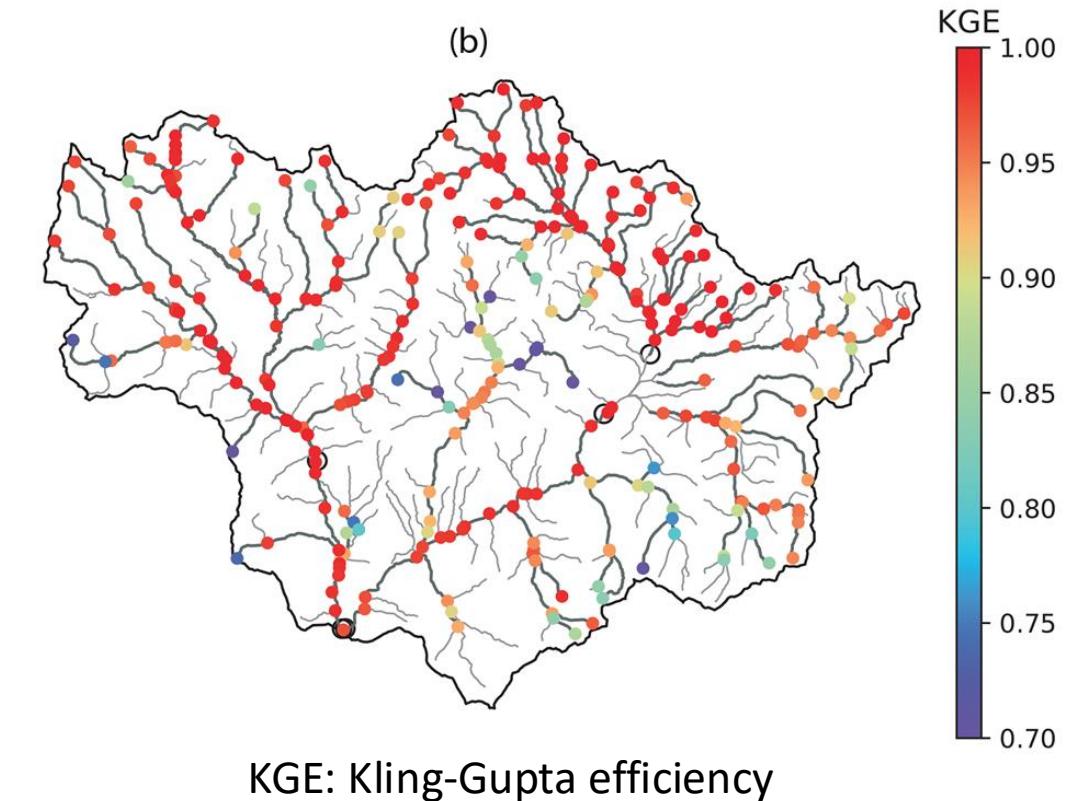
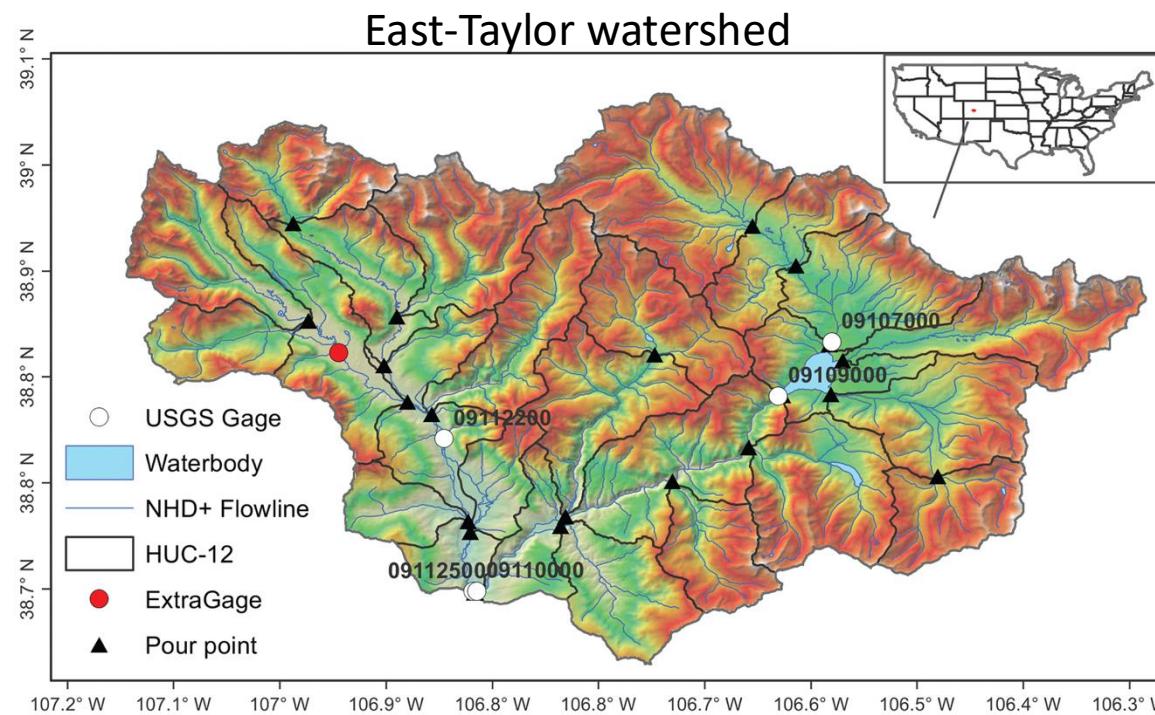
An example watershed

Streamflow prediction across a watershed: Graph Neural Network



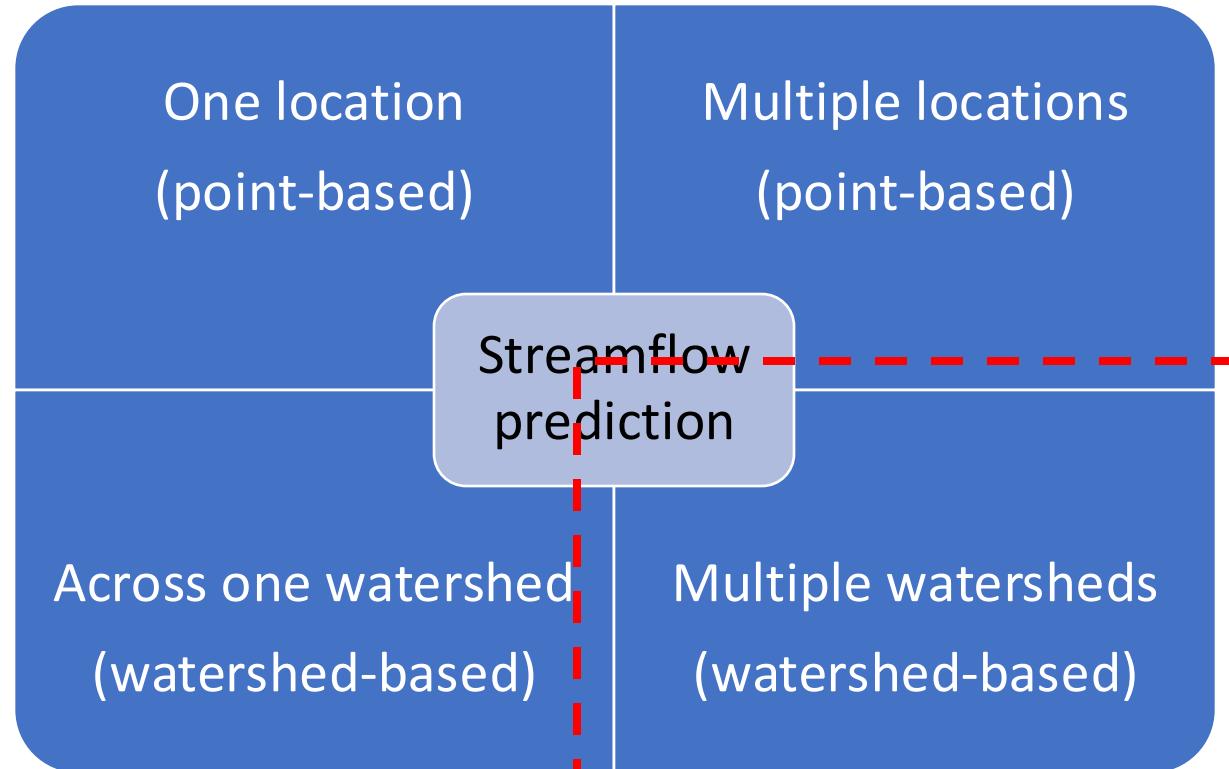
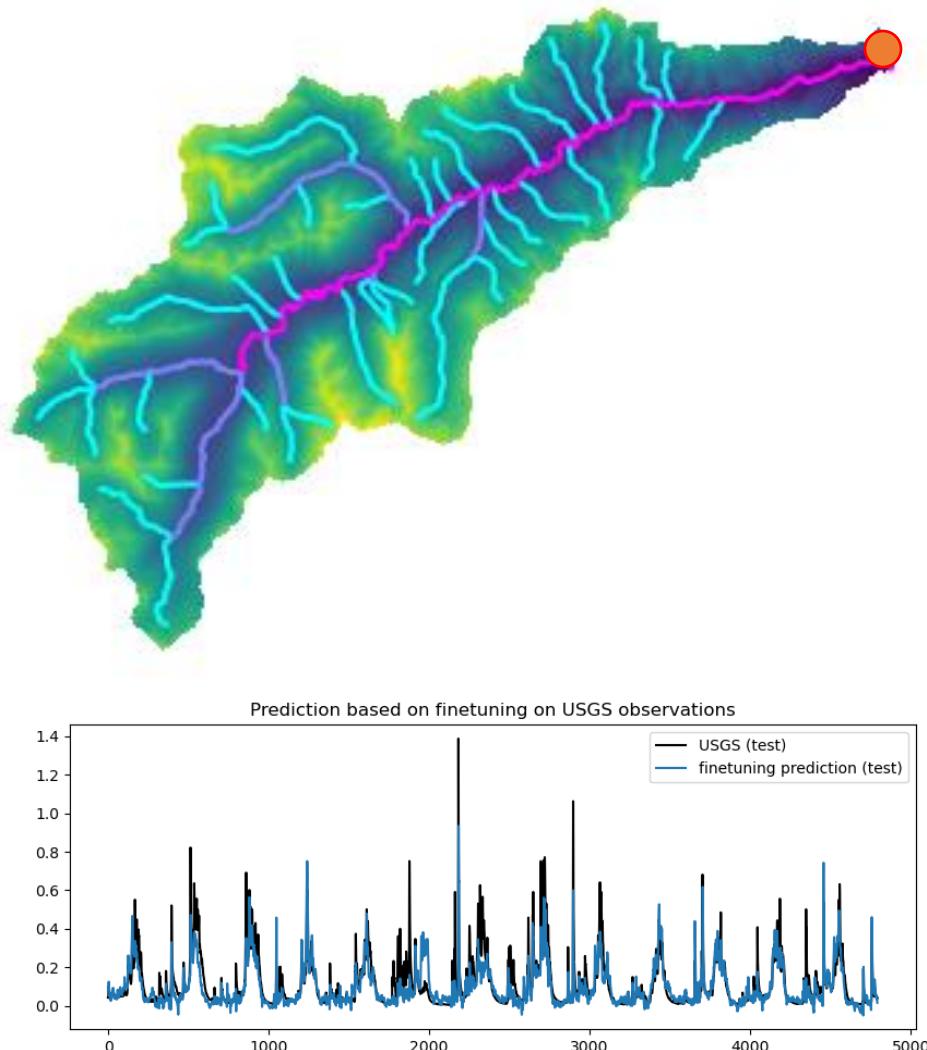
Kipf, Thomas N., and Max Welling. "Semi-supervised classification with graph convolutional networks." *arXiv preprint arXiv:1609.02907* (2016).

Streamflow prediction across a watershed: Graph Neural Network



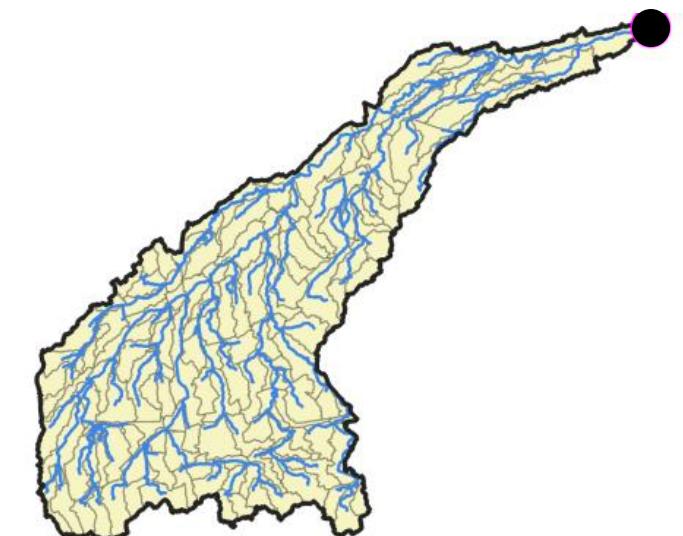
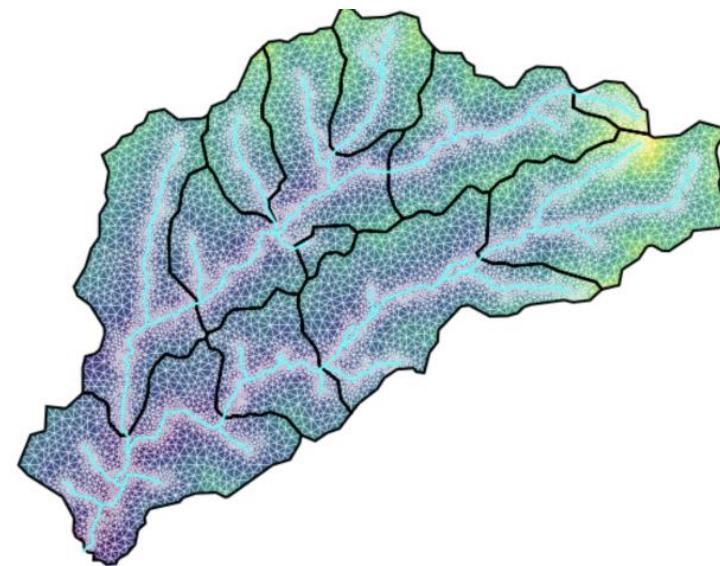
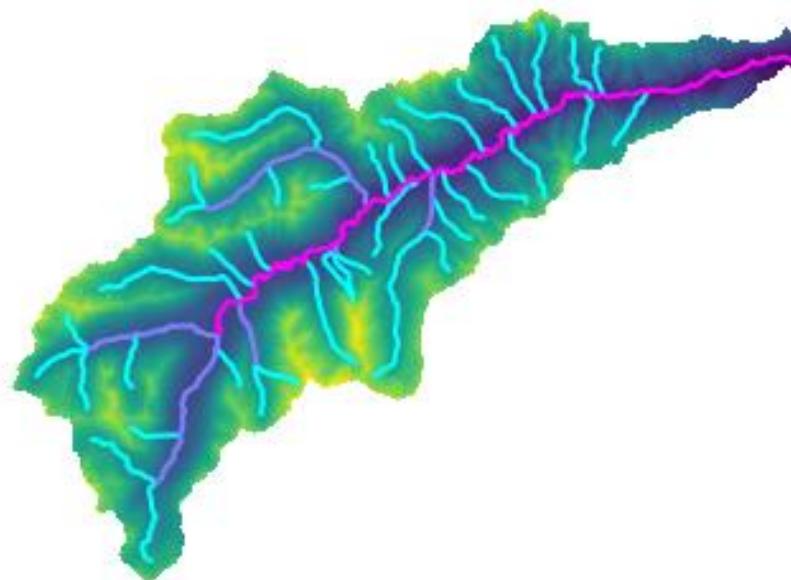
Sun, Alexander Y., Peishi Jiang, Zong-Liang Yang, Yangxinyu Xie, and Xingyuan Chen. "A graph neural network (GNN) approach to basin-scale river network learning: the role of physics-based connectivity and data fusion." *Hydrology and Earth System Sciences* 26, no. 19 (2022): 5163-5184.

ML application in streamflow prediction



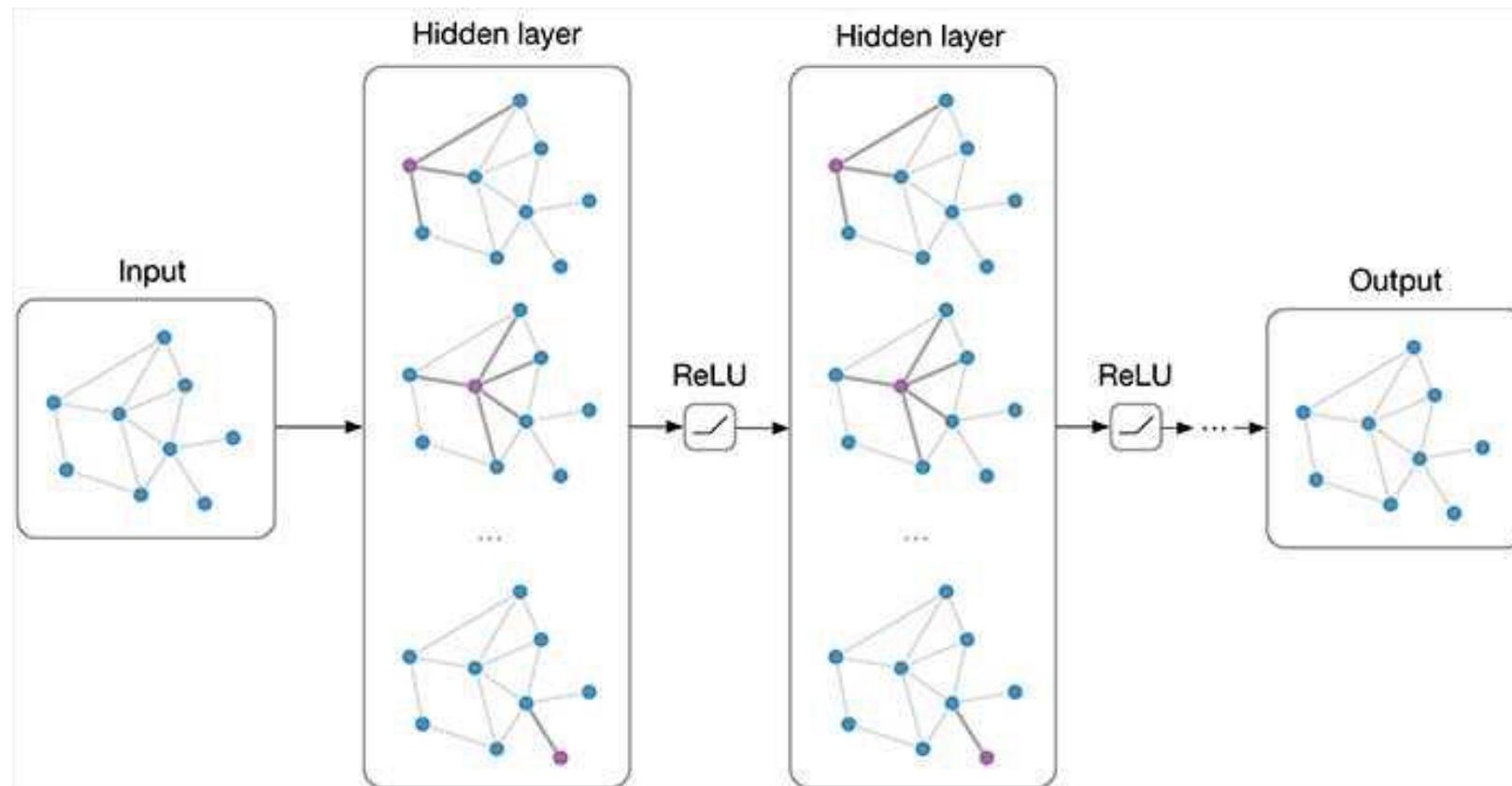
Streamflow prediction across ungauged watersheds

What if the goal is to predict streamflows at reach scale across watersheds?



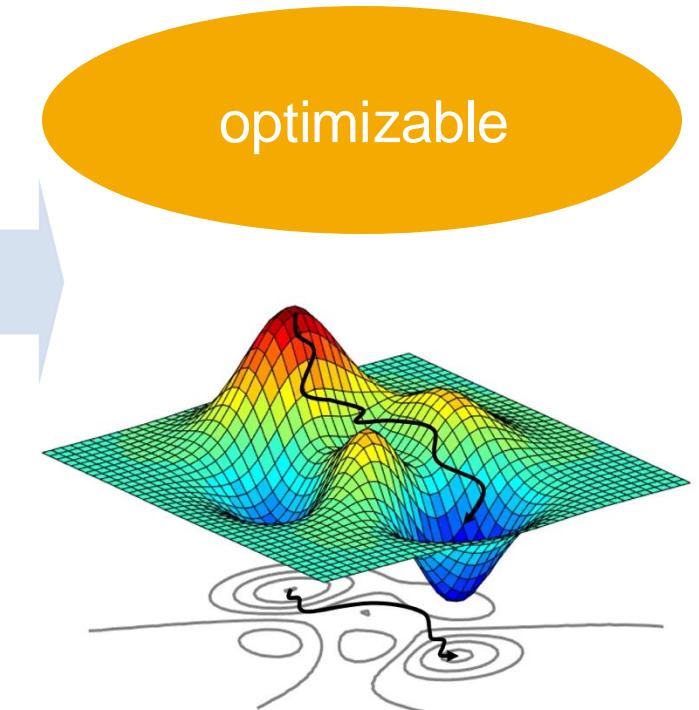
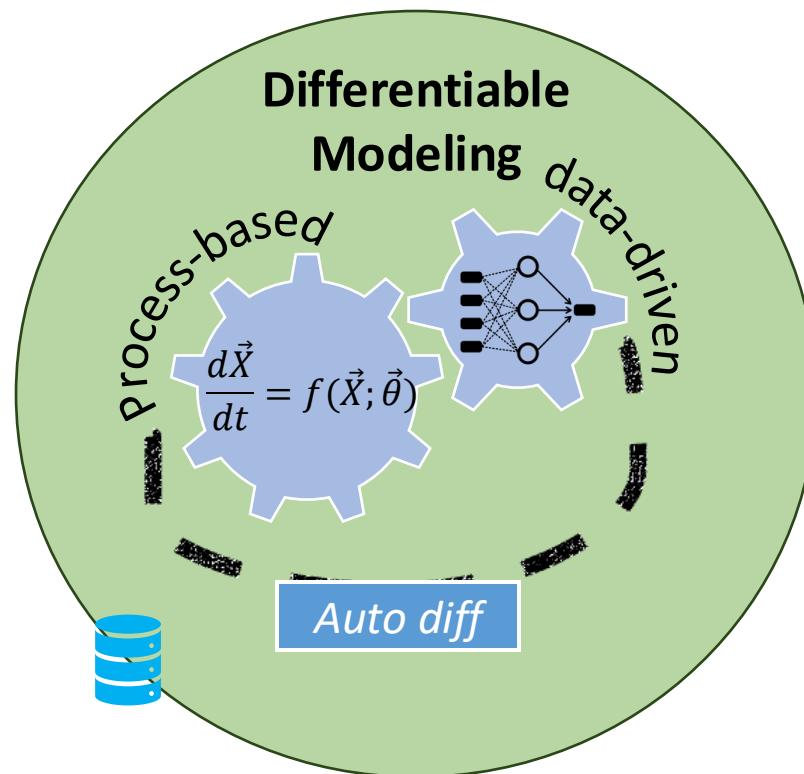
Streamflow prediction across ungauged watersheds

The vanilla Graph Neural Network won't work any more



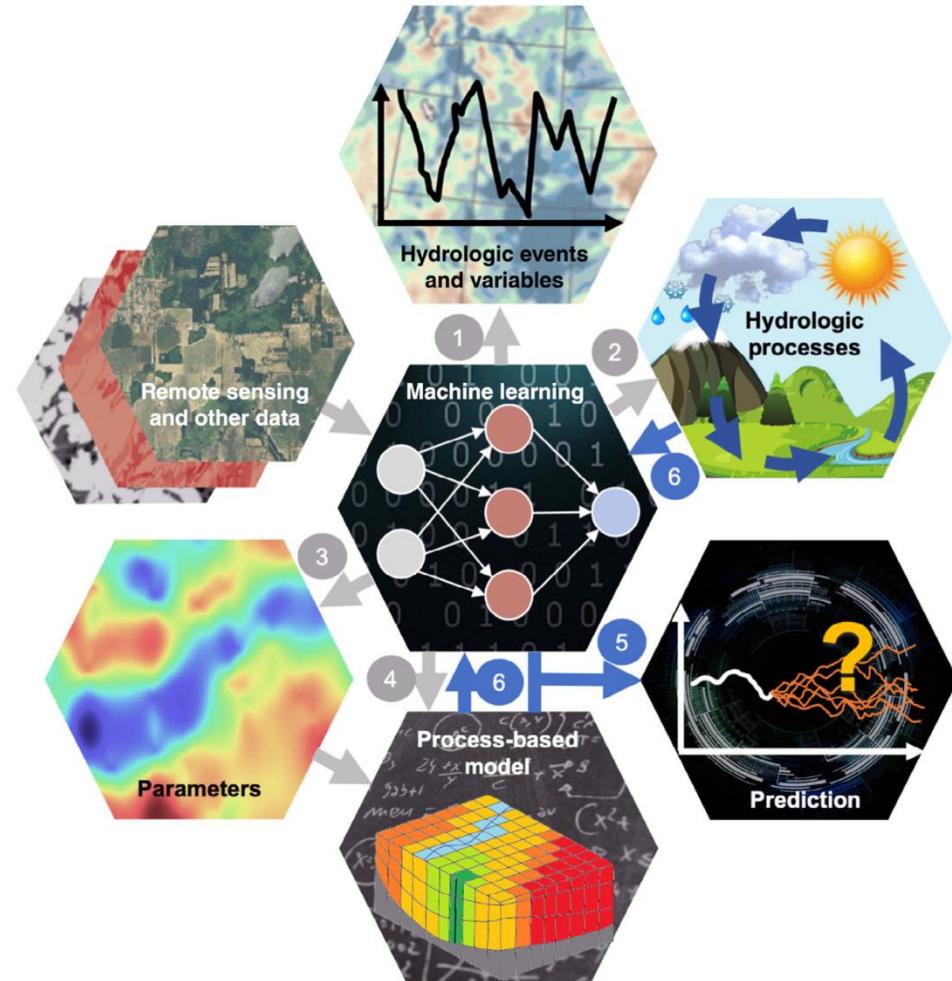
Streamflow prediction across ungauged watersheds: Hybrid Modeling

Ongoing work but there is a promise in using differentiable modeling



Other ML applications in Hydrology

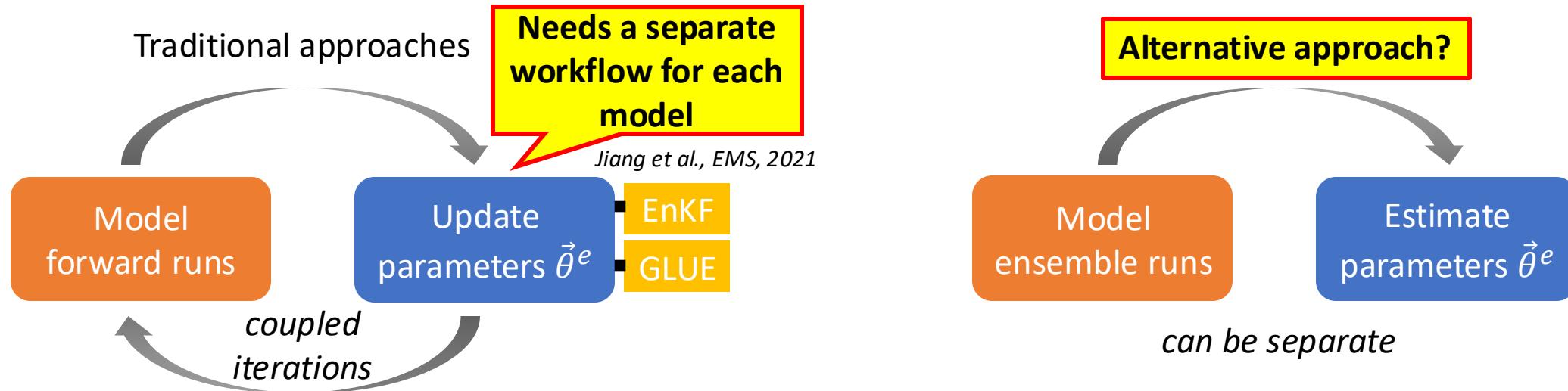
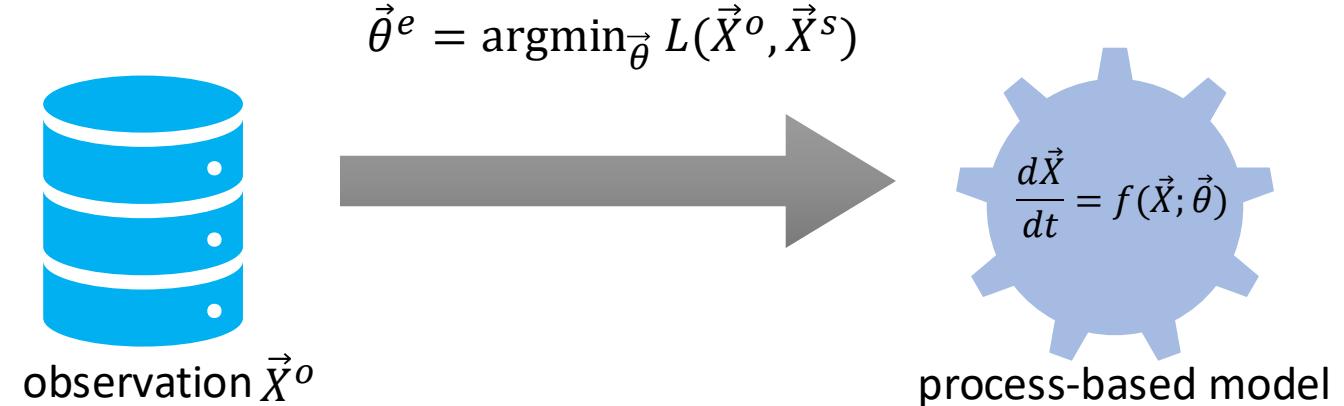
- Surrogate modeling
- Inverse modeling
- Hybrid modeling



Xu, Tianfang, and Feng Liang. "Machine learning for hydrologic sciences: An introductory overview." *Wiley Interdisciplinary Reviews: Water* 8.5 (2021): e1533.

Inverse modeling: How to estimate the parameters of process-based models from observations?

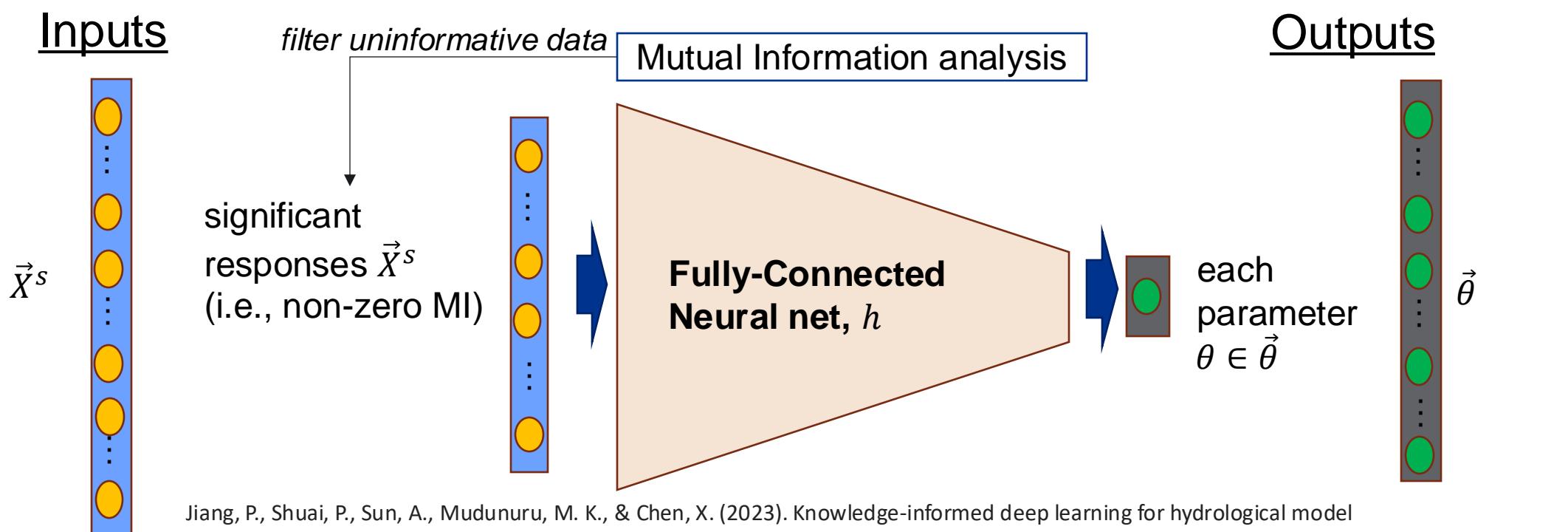
$\vec{\theta}^e$: the optimal parameters
 \vec{X}^o : the observations
 $\vec{X}^s \in \vec{X}$: the observed model states
 L : the loss function



Knowledge-Informed deep learning approach hydrological model calibration

$$\text{Inverse mapping } \vec{\theta}^e = h(\vec{X}^o; \vec{w}^e) \text{ with} \\ \vec{w}^e = \operatorname{argmin}_{\vec{w}} L(\vec{\theta}^e; h(\vec{X}^s; \vec{w}))$$

h : a deep learning model ;
 \vec{w} : the trainable parameters of h ;
 L : the loss function.



An application at Coal Creek Watershed in Colorado



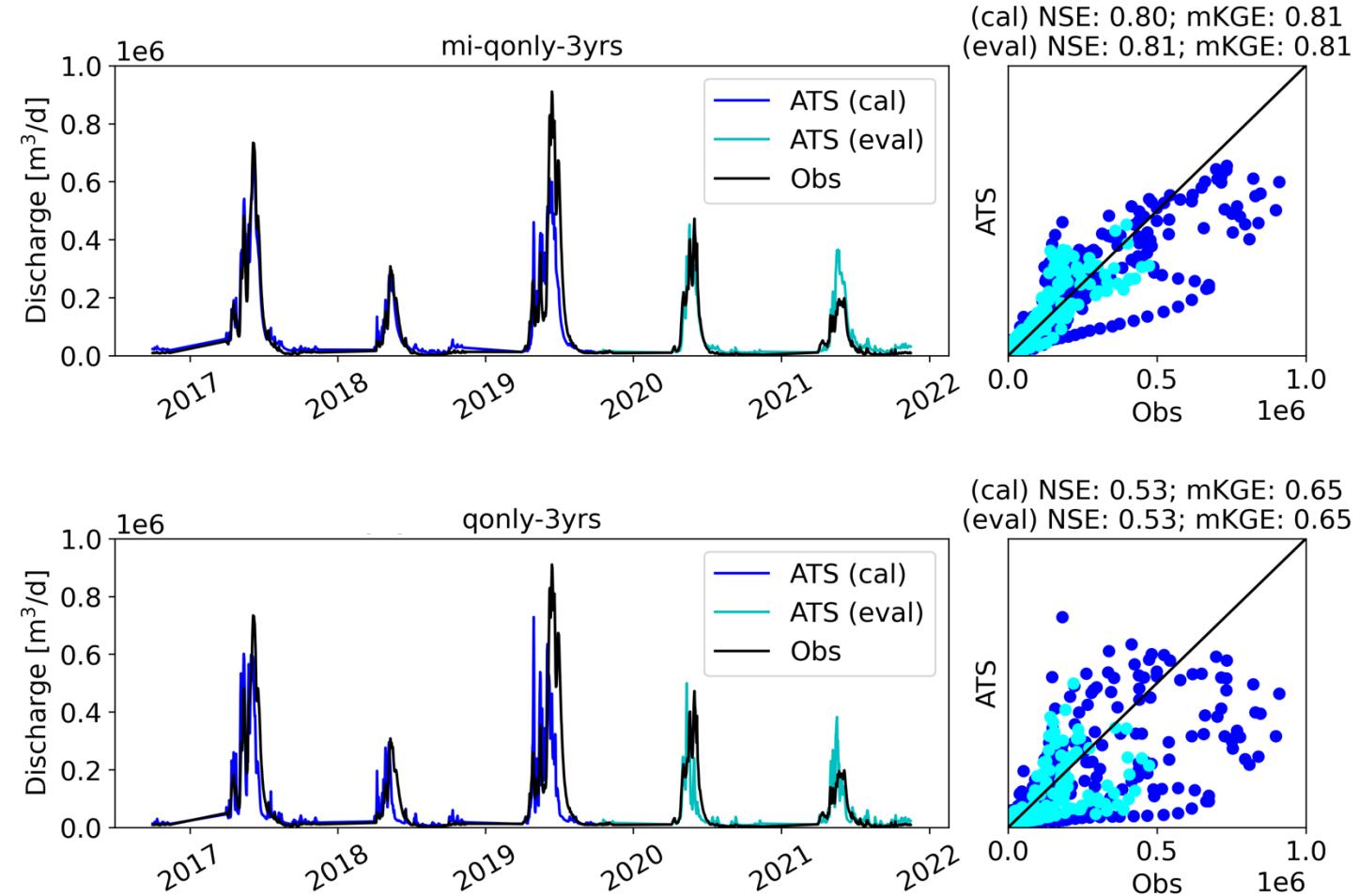
**knowledge-informed
inverse mapping**

NSE: the Nash–Sutcliffe model efficiency coefficient

mKGE: the modified the Kling-Gupta efficiency metric

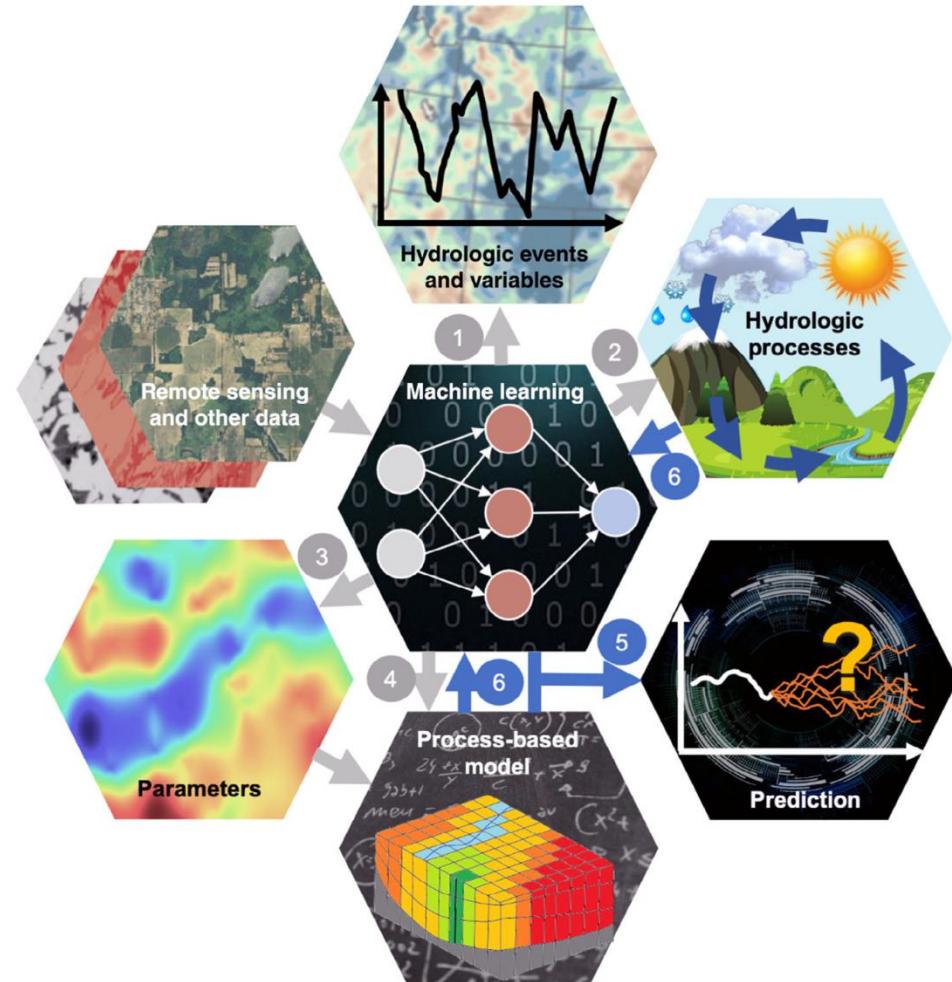


**original inverse
mapping (w/o MI
analysis)**



Other ML applications in Hydrology

- Surrogate modeling
- Inverse modeling
- Hybrid modeling



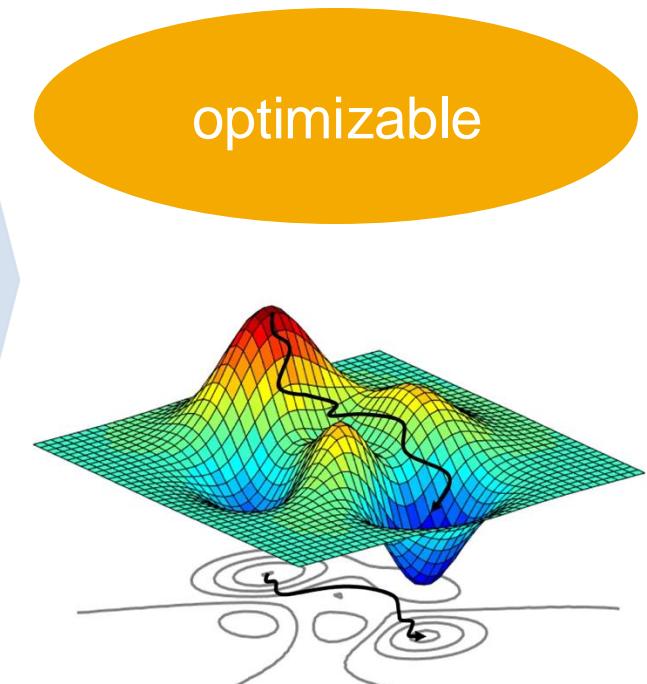
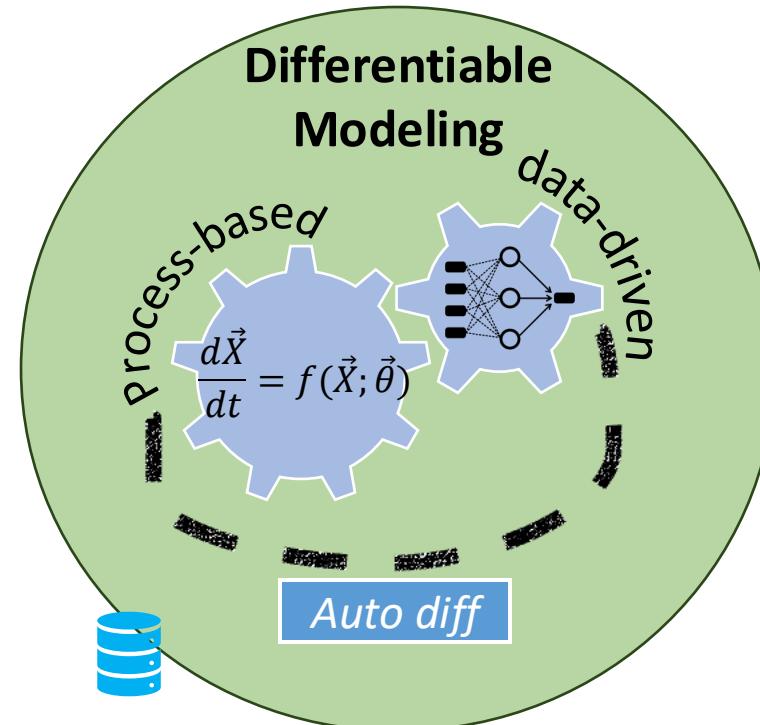
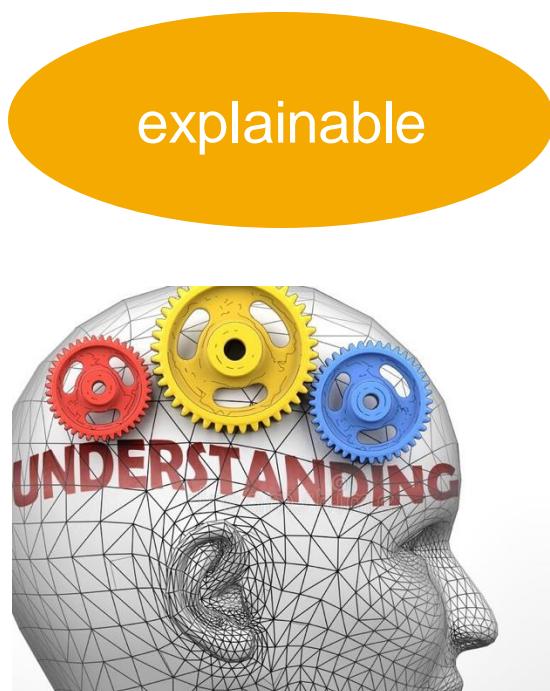
Xu, Tianfang, and Feng Liang. "Machine learning for hydrologic sciences: An introductory overview." *Wiley Interdisciplinary Reviews: Water* 8.5 (2021): e1533.

Differentiable hybrid models for ecohydrological processes

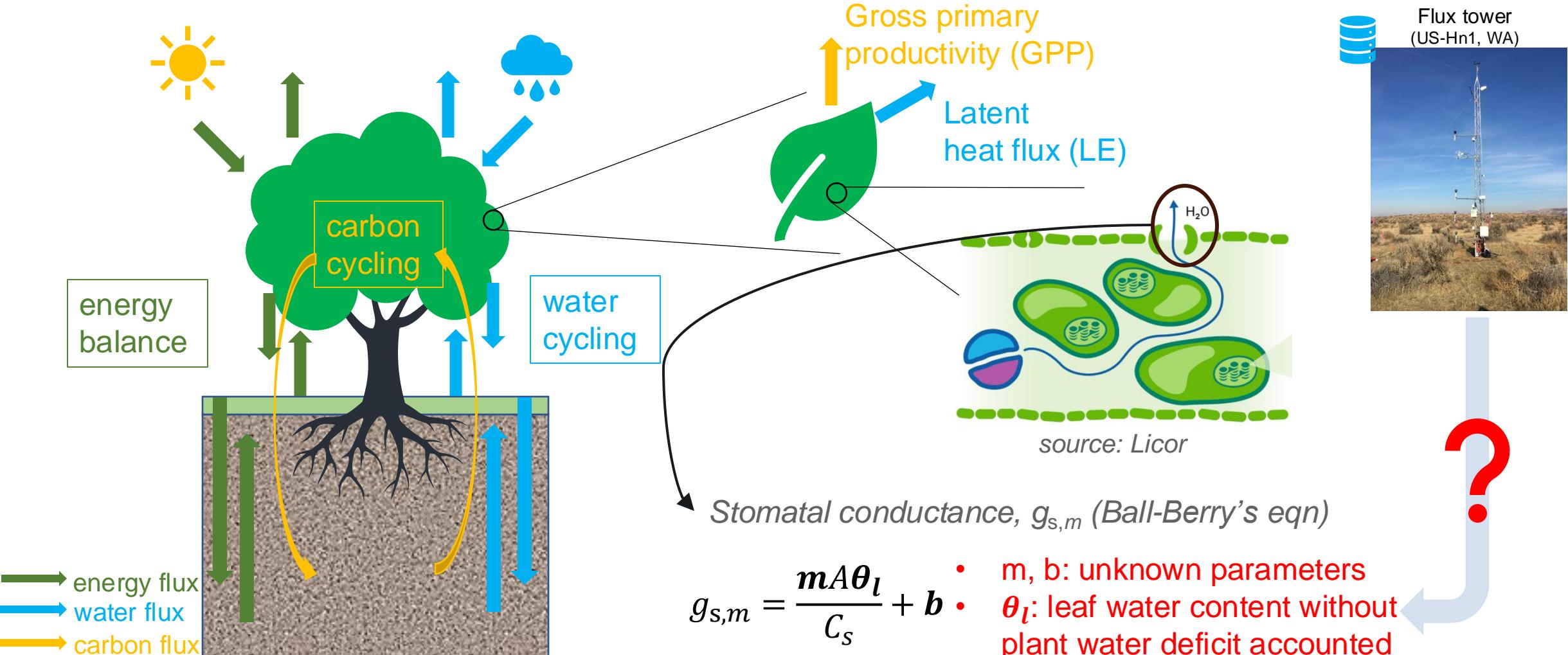
Differentiable hybrid modeling $\frac{d\vec{X}}{dt} = f(\vec{X}; \vec{\theta}^e = h(\vec{X}; \vec{w}^e))$ with $\vec{w}^e = \operatorname{argmin}_{\vec{w}} L(\vec{X}^o, \vec{X}^s)$

h : a deep learning model 

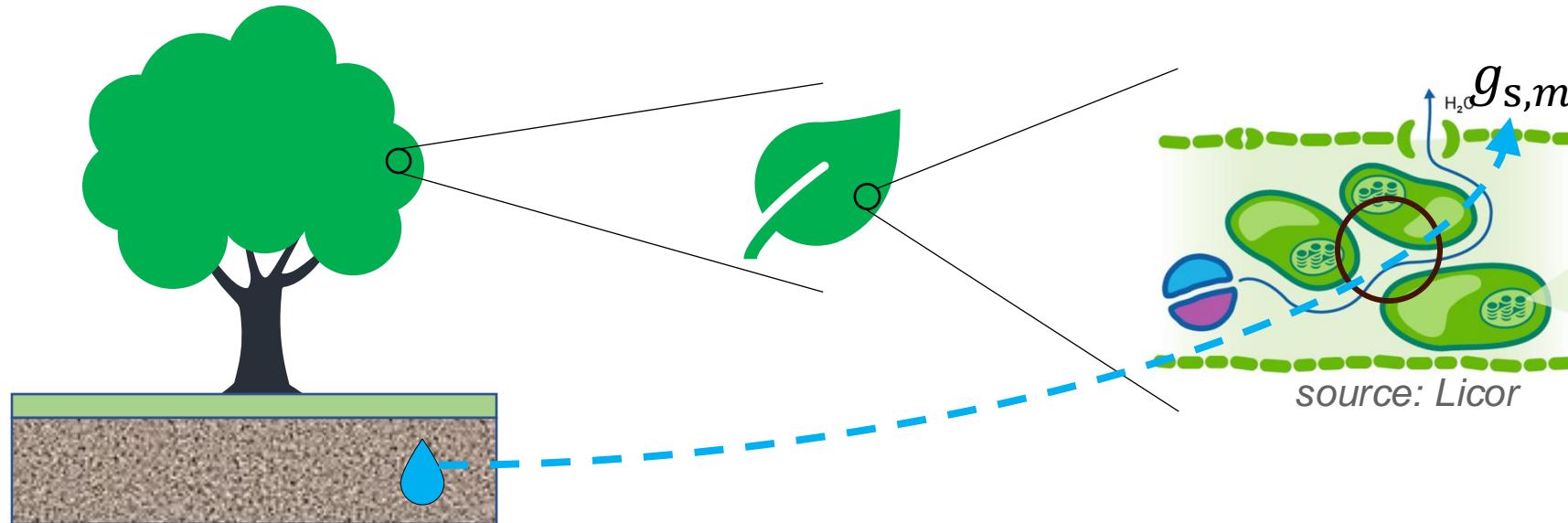
Inverse mapping for parameter estimation $\vec{\theta}^e = h(\vec{X}^o; \vec{w}^e)$ with $\vec{w}^e = \operatorname{argmin}_{\vec{w}} L(\vec{\theta}^{prior}; h(\vec{X}^s; \vec{w}))$



Current land surface models are facing uncertainties in model parameters and mechanisms



We proposed a hybrid version of Ball-Berry eqn calculating stomatal conductance $g_{s,m}$ to account for plant water stress impact



The original Ball-Berry equation

$$g_{s,m} = \frac{mA\theta_l}{C_s} + b$$

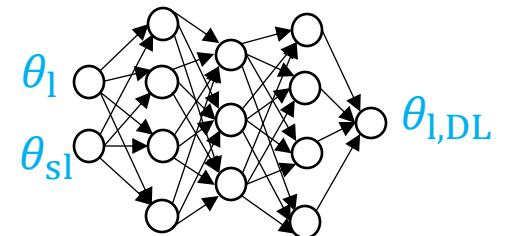
$$\theta_l = \frac{e_a}{e_*(T_l)}$$

- Not accounting for plant water stress

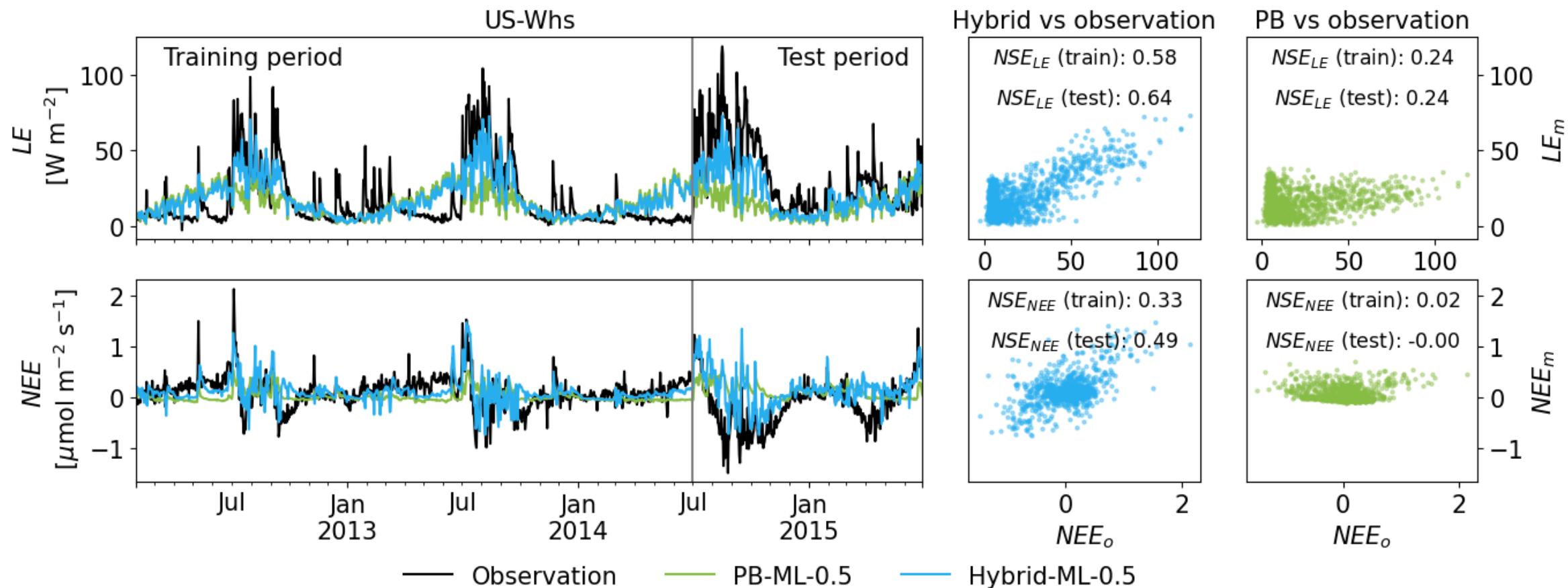
The hybrid Ball-Berry equation approximates soil water stress impact on $g_{s,m}$ using soil water content

$$g_{s,m,DL} = \frac{mA\theta_{l,DL}}{C_s} + g_0$$

$$\theta_{l,DL} = h(\theta_l, \theta_{sL})$$



Hybrid modeling improves both the latent heat fluxes (LE) and net ecosystem exchange (NEE) simulations at a semi-arid ecosystem



Conclusion



Overview of ML application in Hydrology



ML application in streamflow prediction



ML application in parameter estimations of hydrological models



Differentiable hybrid modeling in ecohydrology