https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

HESS Opinion: Beyond the Machine Learning Paradox in Hydrology: A Framework for Scientifically Useful Modeling

Vidya Samadi [1, 2, 3], Hayley J. Fowler[4,5], Daniel P. Ames[6], Edoardo Borgomeo[7], Jonathan L. Goodall[8,9], Jeffery S. Horsburgh[10], Anthony J. Jakeman[11,12], Venkataraman Lakshmi [13], Venkatesh M. Merwade[14], Vimal Mishra [15,16], Yadu Pokhrel[17], Mostafa Saberian[18], Thomas Wahl[19,20], Conrad Wasko[21], Catherine A. M. E. Wilson[22], Nima Zafarmomen[1]

1. Department of Agricultural Sciences, Clemson University, Clemson, SC, USA.
2. School of Computing, Clemson University, SC, USA.
3. Artificial Intelligence Research Institute for Science and Engineering (AIRISE), School of Computing, Clemson University, Clemson, SC, USA.
4. School of Engineering, Newcastle University, Newcastle upon Tyne, United Kingdom.
5. Tyndall Centre for Climate Change Research, Newcastle University, Newcastle upon Tyne, United Kingdom.
6. Department of Civil and Construction Engineering, Brigham Young University, Provo, UT, USA.
7. Centre for Sustainable Development, Department of Engineering, University of Cambridge, Cambridge, United Kingdom.
8. Department of Civil and Environmental Engineering, University of Virginia, Charlottesville, VA, USA.
9. School of Engineering and Applied Science, University of Virginia, Charlottesville, VA, USA.
10. Department of Civil and Environmental Engineering and Utah Water Research Laboratory, Utah State University, Logan, UT, USA.
11. Fenner School of Environment and Society, Australian National University, Canberra, ACT, Australia.
12. Institute for Water Futures, Australian National University, Canberra, ACT, Australia.
13. Department of Environmental Health and Engineering Whiting School of Engineering Johns Hopkins University, USA
14. Lyles School of Civil and Construction Engineering, Purdue University, West Lafayette, IN, USA.
15. Earth Sciences Department, Indian Institute of Technology Gandhinagar, Gandhinagar, India.
16. Civil Engineering, Indian Institute of Technology Gandhinagar, Gandhinagar, India.
17. Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI, USA.
18. Department of Civil Engineering, Clemson University, Clemson, SC, USA.
19. Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, FL, USA.
20. Center for Integrated Coastal Research, University of Central Florida, Orlando, FL, USA.
21. School of Civil Engineering, University of Sydney, Sydney, NSW, Australia.
22. School of Engineering, Cardiff University, Cardiff, United Kingdom.

Corresponding author: samadi@clemson.edu

Abstract

Machine learning (ML) is rapidly transforming hydrologic modeling, delivering remarkable predictive capabilities across applications. Yet a fundamental question remains: does greater predictive performance necessarily advance hydrologic process understanding? While ML models can outperform process-based approaches, they often provide limited mechanistic insight, violate physical constraints, and struggle beyond their training data. We refer to this tension as the "Machine Learning Paradox in Hydrology": the flexibility that enables ML to capture nonlinear catchment responses can also decouple predictive performance from mechanistic interpretability. This tension raises a broader challenge for hydrologic modeling: how can predictive models be used not only to forecast system behavior, but also to test, refine, and advance understanding of the processes that govern it? In this Perspective, we propose a framework for

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

advancing data-driven hydrologic modeling based on five dimensions: predictive skill, physical plausibility, extrapolative robustness, process-diagnostic value, and theory-building capability. Using examples of ML underperformance across the conterminous United States (CONUS), we illustrate how model failures can reveal unresolved processes, missing information, and limitations in existing physical representations and theory. We then propose four complementary strategies to address the ML paradox: physics-informed ML, explainability, transfer learning, and residual structural learning. Such capabilities would transform hybrid models from predictive tools into process-informed reasoning systems that actively support hypothesis testing, model falsification, mechanistic interpretability, and theory development.

Keywords: Machine learning; Hydrologic modeling; Hybrid machine learning; Explainability; Residual structural learning; Mechanistic interpretability.

Short Summary
Machine learning can predict hydrological variables remarkably well, but accurate predictions do not always explain how catchments respond to rainfall and changing conditions. By examining cases across CONUS where machine learning models underperform, we show how their failures can reveal overlooked processes and limitations in existing models. We propose a framework that combines data-driven methods with physical knowledge to improve predictions and advance hydrologic understanding.

1. When Can Machine Learning Advance Hydrologic Understanding?

Machine learning (ML) is transforming hydrologic science, reshaping how hydrologists model and predict the response of complex environmental systems. Across a growing range of applications including streamflow prediction, flood forecasting, drought monitoring, water quality assessment, and hydrologic model emulation, ML models are routinely achieving predictive performance often exceeding that of traditional conceptual and process-based models (Kratzert et al., 2019; Nearing et al., 2021; Shen et al., 2023; Sadeghi Tabas et al., 2025). The availability of large sample hydrologic datasets, advances in computational power, and increasingly sophisticated ML models have accelerated the adoption of data-driven approaches throughout the hydrologic sciences.

Historically, hydrologic models have been developed for two primary purposes (1) to generate accurate predictions, often to aid decision making, and (2) to improve our understanding of the processes governing the movement, storage, and transformation of water across landscapes. Process-based models therefore serve a dual scientific purpose. They provide practical predictive tools while simultaneously representing hypotheses about catchment behavior. These hypotheses are expressed through conservation laws and representations of process interactions governing the movement of water, energy, and constituents (Abbott et al., 1986; Maxwell and Miller. 2005; Wood et al., 2011; Mendoza et al., 2015; Asher et al., 2015; Elkouk et al., 2024; Staudinger et al., 2025) These models hence enable researchers to test, refine, and potentially falsify hypotheses about hydrologic system behavior through comparison with observations (Wagener et al., 2010; Clark et al., 2011; Beven, 2012). ML models fundamentally challenge this traditional relationship between prediction and understanding. Modern ML models can learn highly complex, nonlinear relationships directly from observations without requiring an explicit representation of the physical processes that generate those observations (Vapnik, 1998; LeCun et al., 2015). As a result, while ML models may achieve remarkable predictive skills, they often provide limited insight into the mechanisms responsible for system behavior (Shalev-Shwartz and Ben-David, 2014), limiting our ability to improve process understanding (Lipton, 2018). This creates a tension between predictive success and broader scientific objectives such as improving our mechanistic understanding of complex hydrological systems.

Here, we characterize this tension as the “Machine Learning Paradox in Hydrology”. The paradox arises because the same flexibility that allows ML models to capture complex hydrologic behavior can also obscure the mechanisms responsible for that behavior. Unlike process-based models, whose structures are

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

at least partially linked to physical hypotheses and governing principles, ML models can achieve highly accurate forecasts while providing limited insight into the governing hydrologic processes (Reichstein et al., 2019). This tension manifests through several interconnected challenges that affect ML models, including the balance between prediction and explanation, the need for physical consistency, robustness of models under extrapolation, and the difficulty of extracting mechanistic interpretability from learned representations (Karpatne et al., 2017; Reichstein et al., 2019; Nearing et al., 2021; Beucler et al., 2021; Frame et al., 2022; Sadeghi Tabas et al., 2025). Framing the discussion as a paradox serves to surface a more important scientific question. The question should not be whether ML or process-based models are inherently better, rather, the more productive scientific question is: under what conditions does ML contribute to process understanding?

This question echoes Beven's (2012) vision of "models of everywhere" calling for moving beyond prediction toward models that actively advance process understanding and hydrologic theory. In this Perspective, we posit that the research community should move beyond the age-old debate between process-based and data-driven modelling and instead ask how ML can contribute to hydrologic theory development.

2. When Machine Learning Models Fail

While ML models often demonstrate exceptional predictive skill, their performance varies across landscapes, hydroclimatic settings, and hydrologic regimes. These systematic failures can serve as diagnostic signals, revealing unresolved process interactions, missing information, and limitations in current hydrologic understanding. For example, we identified catchments across the conterminous United States (CONUS) where ML models such as Long Short-Term Memory (LSTM; Hochreiter and Schmidhuber, 1997) and Temporal Fusion Transformer (TFT; Lim et al, 2021) consistently underperformed conceptual and process-based models for daily streamflow prediction (Sadeghi-Tabas et al., 2025; Figure 1), occur systematically in catchments with strong physiographic heterogeneity and complex storage dynamics. This suggests that there are key controls on streamflow variability that remain poorly represented in available training data (Table 1).

Table 1. Typical failure modes of ML models in hydrology and their implications for hydrologic understanding.

Failure mode	Hydrologic characteristics	Why ML struggles	Scientific insight revealed	Representative examples	Criteria primarily stressed	References
Type A: Localized process dominance (flashy headwater systems)	Small-to-medium catchments with rapid runoff generation, strong spatial heterogeneity, hillslope–channel connectivity, urban modification	Important controls occur at spatial scales not represented by forcing data or catchment descriptors	Highlights limitations in observations of fine-scale connectivity and anthropogenic influences	Western Run, Jacob Fork, South Fork Roanoke	Predictive skill; process-diagnostic value	Sadeghi Tabas et al. (2025); Wang & Karimi (2022); Song et al. (2025); Vischer et al. (2025); Jeong et al. (2026)
Type B: Storage-dominated systems	Strong groundwater contributions, subsurface connectivity, delayed release, long catchment memory	Internal storage states are poorly observed and difficult to infer from meteorological inputs alone	Reveals importance of hidden states, catchment memory, and subsurface processes	Cossatot River, Cypress Creek	Process-diagnostic value; theory-building capability	Sadeghi Tabas et al. (2025); Lees et al. (2021, 2022); De la Fuente et al. (2024)

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Type C: Intermittent and Threshold-Dominated Systems	Ephemeral streams, transmission losses, runoff thresholds, long zero-flow periods interrupted by floods	Rare events are underrepresented in training data and difficult to learn statistically	Reveals importance of threshold behavior and nonlinear process transitions	Nueces River	Predictive skill; theory-building capability	Sadeghi Tabas et al. (2025); Gutiérrez-Jurado et al. (2019); Costigan et al. (2015); Kaplan et al. (2022)

2.1. Type A: Localized Process Dominance

Flashy headwater catchments represent one of the most persistent challenges for streamflow prediction using ML models. These rapidly responding catchments are characterized by short flow paths and localized runoff generation, that occur over spatial and temporal scales often finer than those represented within available forcing datasets (Figure 1). Examples include Western Run, Maryland (USGS 01583500; 154.9 km²), Jacob Fork, North Carolina (USGS 02143040; 66.6 km²), and the South Fork Roanoke River near Shawsville, Virginia (USGS 02053800; 282.3 km²), where ML models consistently underperform relative to conceptual and process-based models (Tabas and Samadi., 2022; Tabas et al., 2023; Sadeghi Tabas et al., 2025). Across the three catchments, hydrological models including Sacramento Soil Moisture Accounting model (SAC-SMA; Burnash, 1995), Variable Infiltration Capacity model (VIC; Liang et al., 1994), and Hydrologiska Byråns Vattenbalansavdelning (HBV; Bergström, 1976) achieved moderate but consistent performance (NSE ~ 0.54–0.66), whereas ML models achieved substantially lower skill (NSE ~ 0.13-0.39).

Western Run is located within the Mid-Atlantic Piedmont near the Baltimore metropolitan region, where mixed rural–suburban–urban land use has substantially altered natural drainage connectivity through channel modification, road–bridge crossings, and localized stormwater inputs (Thomas et al., 2000). Consequently, streamflow dynamics are influenced not only by rainfall and antecedent moisture conditions but also by anthropogenic modifications to surface and subsurface flow pathways. In contrast, Jacob Fork is a steep, forested headwater catchment in the North Carolina Piedmont and Foothills region, where short flow paths and steep channel gradients produce rapid hydrologic response (Wilder and Simmons, 1982). Streamflow dynamics are strongly influenced by localized rainfall patterns, shallow subsurface storage, and hillslope–channel connectivity. Consequently, small differences in rainfall location, antecedent conditions, or connectivity can produce markedly different runoff responses.

In the South Fork Roanoke River near Shawsville, VA, streamflow variability reflects the combined influence of complex topography, groundwater contributions, and spatially heterogenous runoff generation processes (Ward et al., 2011; Sanford et al., 2012). In addition, the catchment is moderately influenced by infrastructure, including culverts and bridges, and localized flow regulation structures such as small weirs (Jastram et al., 2015), which further modify flow pathways and channel dynamics. Poor TFT and LSTM performance in this catchment suggests that important hydrologic controls governing streamflow variability are either insufficiently observed or inadequately represented within the predictor variables used for model training.

These catchments highlight a key challenge for hydrologic modelling: many of the dominant controls on runoff generation function at scales that are poorly observed and therefore remain outside the information content of typical training datasets. Rather than representing a weakness of ML alone, such failures expose limitations in understanding fine-scale hydrologic connectivity.

2.2. Type B: Storage-Dominated Systems

A second failure mode emerges in catchments characterized by strong groundwater contributions, delayed subsurface drainage, and long hydrologic memory. Examples include the Cossatot River near Vandervoort, Arkansas (USGS 07340300; 89.6 Km^2), where streamflow reflects both rapid storm response and delayed

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

subsurface drainage (Green, 1994), and Cypress Creek, Mississippi (USGS 02479155), where streamflow depends strongly on internal storage states, subsurface connectivity, and antecedent conditions. In both catchments, hydrologic models achieve NSE values of approximately 0.53–0.65, compared with 0.16–0.34 for ML models.

The Cossatot River in the Ouachita Mountains is characterized by steep hillslopes, narrow valleys, and heterogeneous forested terrain. Streamflow in this system reflects a combination of rapid flow contributions from connected hillslopes and delayed releases from subsurface storage and spring systems (e.g., Ludwig, 1992). Similarly, the Cypress Creek catchment in Mississippi (USGS 02479155) situated in the low-gradient Gulf Coastal Plain, exhibits ponding conditions due to wetlands, shallow groundwater tables, floodplain storage, and slow drainage processes (Myers, 2022). These two systems represent contrasting extremes of hydrological complexity; one dominated by rapid flow connectivity in steep terrain, and the other one by distributed storage and delayed release in low-relief landscapes. Although these catchments differ substantially in physiography, they share an important characteristic: similar meteorological inputs can produce markedly different streamflow responses depending on antecedent groundwater conditions, subsurface storage states, and channel connectivity.

Although modern ML models such as TFTs can learn long-term temporal dependencies, their latent representations may capture statistical memory without resolving the physical storage processes that generate streamflow. These failures therefore highlight a fundamental limitation of ML models: when key hydrologic states are poorly represented in observed data, predictive learning alone may be insufficient to discover the mechanisms governing catchment response.

2.3. Type C: Intermittent and Threshold-Dominated Systems

Semi-arid and intermittently flowing catchment systems, such as the Nueces River at Laguna, Texas (USGS 08190000; 1,908.8 km²), are characterized by highly variable runoff generation, prolonged periods of low- or zero-flow, interrupted by infrequent high-magnitude flood events (Thomas et al., 2019). In such environments, streamflow generation is controlled by threshold processes, nonlinear runoff initiation mechanisms, and episodic connectivity between hillslopes and channels.

Both LSTM and TFT models struggle to reproduce the timing and magnitude of flow variability (NSE = 0.08), whereas hydrological models which explicitly incorporate thresholds for runoff generation perform comparatively better (NSE ≈ 0.44–0.58). ML model failure occurs because the observational record is dominated by prolonged low- or zero-flow periods, while the most important hydrologic behavior is concentrated within a small number of extreme episodes. As a result, the statistical information available for ML training is highly imbalanced. ML models can struggle to reproduce both the timing and magnitude of flow initiation and recession, particularly when flow events occupy only a small fraction of the training record.

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

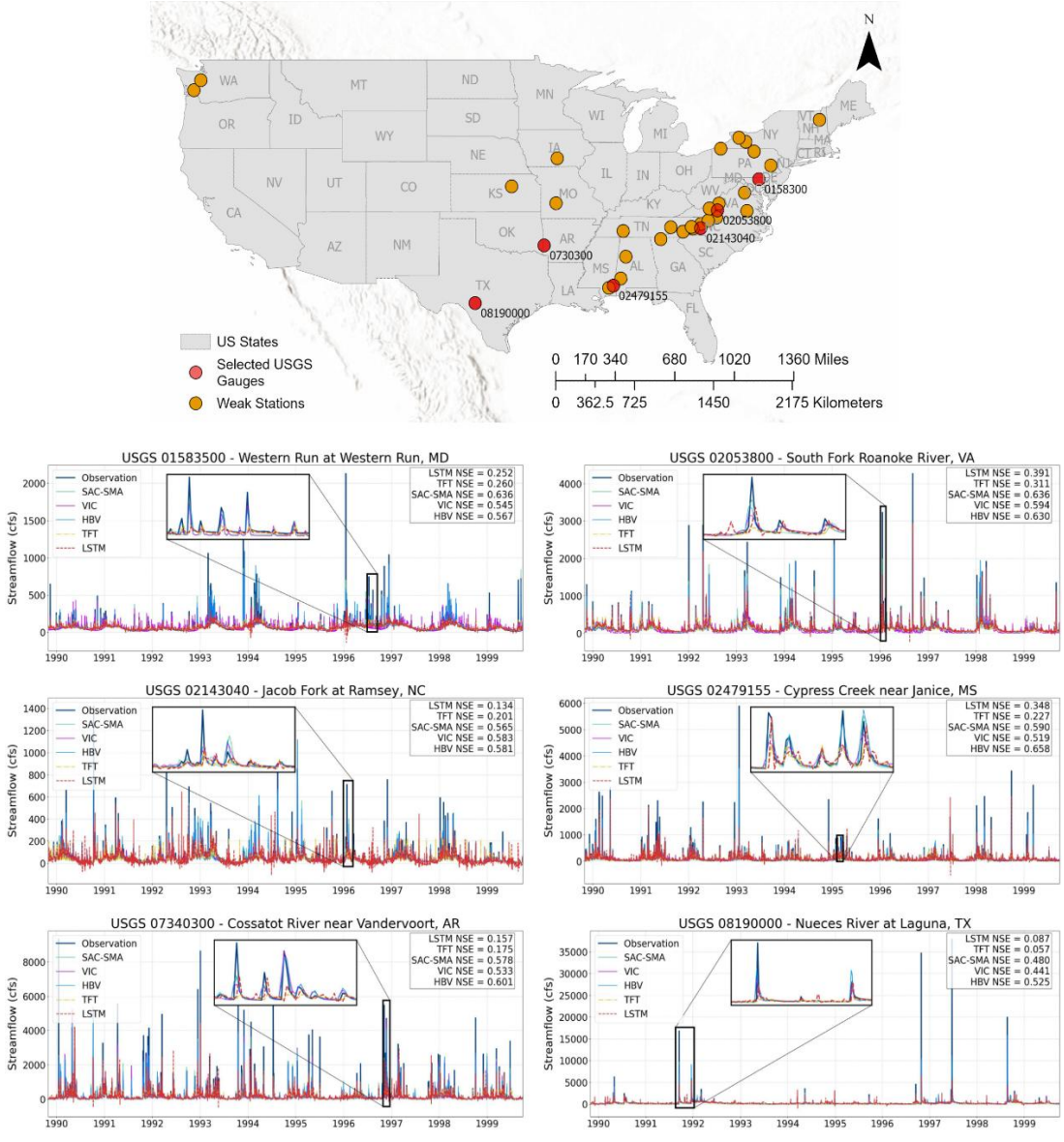


Figure 1. Hydrologically informative ML failure modes across six hydroclimatically diverse basins. The six basins (red) were selected from catchments where ML models (LSTM and TFT) consistently underperformed relative to conceptual and process-based models (SAC-SMA, VIC and HBV) (orange). The selected basins (red) illustrate three recurring classes of failure identified in this study: flashy headwater systems (Western Run, Jacob Fork, South Fork Roanoke), storage-dominated systems (Cossatot River, Cypress Creek), and intermittent systems (Nueces River). Insets show selected high-flow periods, and NSE values summarize model performance. The figure demonstrates that ML performance degrades systematically in environments characterized by localized runoff generation, hidden storage dynamics, threshold behavior, and poorly observed controls on streamflow variability, providing insight into the broader failure modes discussed in Section 2.

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

3. Resolving the ML Paradox

The three failure modes above demonstrate that ML model failures can expose unresolved process interactions, hidden hydrologic states, threshold behavior, and deficiencies in physical representation. Viewed through this perspective, model failure becomes a source of systematic insight rather than merely modeling deficiency. In many applications ML models outperform traditional hydrologic models (Kratzert et al., 2019; Nearing et al., 2021; Shen et al., 2023; Saberian et al., 2026). The ML paradox is not a competition between data-driven and process-based modelling, it reflects a broader challenge: determining when predictive models contribute to scientific understanding, pointing to the fact that predictive accuracy alone is not a measure of scientific value. A model may achieve excellent predictive performance while violating physical constraints, failing under changing environmental conditions, or providing little insight into the mechanisms responsible for its predictions (Karpatne et al., 2017; Reichstein et al., 2019; Beucler et al., 2021). Conversely, a physically interpretable model may advance understanding despite sacrificing predictive performance. Resolving the ML paradox therefore requires moving beyond viewing ML and process-based modelling as competing paradigms. This brings us to the tenet of this perspective: how can we ensure that ML techniques deliver outcomes which are not just numerically superior but also advance our scientific understanding?

Hybrid modeling architectures move beyond treating ML as a stand-alone "black box" and instead position it as a complementary component within a unified, physically grounded modelling framework. In such systems, ML can serve multiple scientific roles: acting as a computationally efficient surrogate for expensive process-based model components (McSpadden et al., 2024; Roy et al., 2025; Mangukiya and Sharma, 2025; ; Acuña Espinoza et al., 2025), learning residual structural errors that remain unexplained by existing theories, and identifying hidden process relationships or emergent system behaviors that are difficult to represent using conventional parameterizations (Reichstein et al., 2019; Slater et al., 2023). Conversely, process-based models provide the physical scaffold within which learning occurs, supplying the conservation laws, process constraints and mechanistic structure necessary to ensure that the ML optimization remains within the physically plausible bounds. Here we identify four complementary integration strategies to build hybrid modeling architectures (Figure 2) that take ML models towards what we term “scientifically useful” hydrologic models that progress our hydrologic process understanding:

(i) **Explainability**, which seeks to reverse-engineer feature contributions to extract new physical insight from ML models whose internal representations lack direct physical interpretation;
(ii) **Physics-Informed ML**, which embeds physical constraints, such as mass balance, directly within ML algorithms to ensure consistency;
(iii) **Transfer Learning**, which transfers knowledge learned from data-rich source domains to poorly gauged or ungauged target regions, enabling hydrologic information to be leveraged across hydroclimatic and observational settings; and
(iv) **Residual Structural Learning**, which uses ML to identify deficiencies in existing theories and process representations after leveraging process-based models to capture known physical laws.

Collectively, these approaches establish a framework in which physical laws constrain the search space of data-driven learning, while ML identifies structural gaps and structural uncertainty in existing physical theories, hidden relationships, and sources of uncertainty within existing hydrologic theories, thereby driving future model refinement in both fields.

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

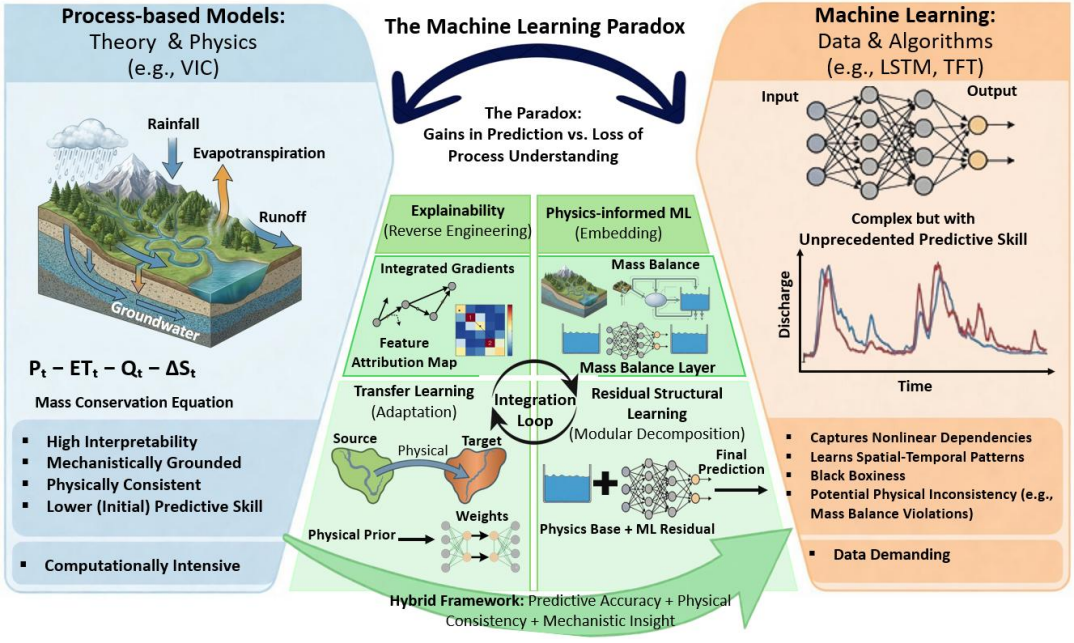


Figure 2. Conceptual framework for addressing the "Machine Learning Paradox in Hydrology" through hybrid modeling that integrates process-based reasoning with data-driven learning. Process-based models provide mechanistic representations of hydrologic processes and encode physical principles such as mass, energy, and momentum conservation. Their strengths lie in physical interpretation and process-based extrapolation, although simplified structures and parameterizations may limit predictive performance. ML models learn complex, nonlinear patterns from data and can achieve high predictive skill, but their internal representations may lack physical plausibility and mechanistic interpretability, particularly under distributional change. The proposed framework connects these paradigms through four complementary strategies: explainability, physics-informed ML, transfer learning, and residual structural learning. The framework is bidirectional: physical knowledge constrains data-driven learning, while ML-generated evidence supports process diagnosis, hypothesis development, and refinement of hydrologic theory.

3.1. Explainability and Process-Diagnostic Value

Although modern ML architecture frequently achieves remarkable predictive skill, their internal representations are rarely expressed in terms of physical processes. Consequently, an important scientific question is whether information contained within these learned representations can be used to improve hydrologic understanding rather than simply explain model behavior. Explainability seeks to address this challenge by quantifying how individual variables, internal states, and learned representations contribute to model predictions. Rather than treating neural networks as irreducible "black boxes", explainability methods attempt to identify which meteorological forcings, catchment attributes, or hidden features exert the greatest influence on predicted streamflow.

Applications in hydrology have already demonstrated that explainable ML can recover physically meaningful relationships associated with snow accumulation and melt, catchment storage dynamics, and climatic controls on runoff generation (Kratzert et al., 2019; Nearing et al., 2021; Shen et al., 2023; De la Fuente et al., 2024). More importantly, these methods provide a mechanism for generating new hypotheses regarding hydrologic system behavior, linking predictive modelling with scientific discovery (Reichstein et al., 2019). One widely adopted approach is Integrated Gradients (IG), which quantifies the contribution

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

of each input variable to an individual prediction (Sundararajan et al., 2017), providing a mathematically rigorous decomposition of the prediction into contributions from individual variables.

From a hydrologic perspective, this formulation enables interpretation. For example, physically realistic behavior would be expected if large, predicted flood peaks were strongly attributable to antecedent soil moisture, intense rainfall, snowmelt, or other hydrologically plausible controls. Conversely, attribution to variables lacking a clear physical connection may indicate that the model has learned spurious statistical relationships rather than true hydrologic processes. Explainability therefore bridges statistical learning and mechanistic interpretation allowing learned representations to be evaluated against established hydrologic theory (Reichstein et al., 2019; Molnar, 2022).

3.2. Physics-informed ML and Physical Plausibility

ML architecture is not automatically required to satisfy conservation laws. An ML model may therefore reproduce observed streamflow with high predictive skill while generating internally inconsistent estimates of storage, evapotranspiration, energy fluxes, or flow propagation. This distinction is central to the ML paradox: predictive accuracy does not necessarily imply that the learned representation is physically plausible. Physics-informed ML addresses this limitation by integrating physical knowledge into the learning process without eliminating the flexibility that makes ML effective (e.g., Raissi et al., 2019; Karniadakis et al., 2021). Several studies have demonstrated the potential of physics-informed ML approaches in hydrologic applications (e.g., Beucler et al., 2021; Shen et al., 2023; Sadeghi Tabas et al., 2025). The broader scientific challenge, however, is to determine how to incorporate hydrological constraints so that models remain accurate, physically consistent, interpretable, and transferable across diverse hydrologic regimes.

Physical constrains can be introduced into the ML framework through several complementary mechanisms. These include: (i) adding physical constraints to the objective (loss) function used in model calibration, (ii) representing physically meaningful states, fluxes, and parameters explicitly within the network architecture, (iii) coupling ML components with governing equations or process-based model structures, and (iv) enforcing conservation principles during training or prediction. Among these strategies, physics-informed loss functions provide one of the most direct and widely used means for guiding model optimization towards physically plausible solutions.

Minimizing the loss function encourages the model to reproduce observations, often yielding highly accurate predictions, but it does not guarantee compliance with fundamental physical principles such as mass conservation, energy balance, or flow continuity. Different internal combinations of states and fluxes may yield similar prediction errors (equifinality as stated in Beven, 2006) even when some combinations violate mass conservation, energy balance, or flow continuity. Physics-informed ML addresses this limitation by augmenting the conventional total loss function with one or more terms that quantify violations of physical relationships (Karniadakis et al., 2021; Wang et al., 2024). During training, the ML model is optimized to simultaneously reduce prediction error and physical inconsistency, and the ML model is penalized not only for inaccurate predictions but also for generating physically implausible system behavior. This restricts the set of solutions explored during training by penalizing violations of fundamental hydrological principles.

Physics-informed loss functions do not automatically guarantee physically correct models. Their effectiveness depends on the suitability of the selected constraints, the accuracy of the observations used to evaluate them, and the choice of penalty weights. Moreover, apparent violations may reflect measurement uncertainty, unresolved exchanges across system boundaries, or scale mismatches rather than deficient model behavior (Wang et al., 2023). Physical constraints should therefore be treated as testable scientific assumptions, not as unquestionably correct representations of the system. The scientific value of physics-

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

informed ML consequently extends beyond producing more plausible predictions. By examining where data-fitting and physical constraints are compatible, and where they remain in tension, one can identify missing processes, inadequate boundary conditions, uncertain observations, and limitations in existing hydrologic theory (e.g., Gupta et al., 2012). Physics-informed learning thus addresses the physical plausibility dimension of scientifically useful ML while also providing a diagnostic framework for refining the physical knowledge embedded within hydrologic models.

3.3. Transfer Learning and Generalization

Although ML models can learn complex hydrologic relationships from large datasets, their performance is often strongly dependent on the quantity, quality, and representativeness of the available observations. This creates a fundamental difficulty for hydrologic applications because many catchments are poorly gauged, have short observational records, or lack measurements of important fluxes. Models trained locally under these conditions may overfit limited observations and fail to represent the broader hydrologic behavior of the target basin.

Transfer learning seeks to address this limitation by reusing knowledge acquired from a data-rich source domain to improve learning in a data-sparse target domain. The source domain may be a well-gauged catchment, a collection of hydrologically similar basins, a large-sample regional dataset, or a process-based simulation archive. The target domain may be a poorly gauged or ungauged catchment for which only limited observations are available. Rather than estimating a target model entirely from sparse local data, transfer learning constrains its parameters or learned representations using information acquired during source-model training.

Several transfer mechanisms in ML are possible. These include transferring complete model parameters, fine-tuning selected network layers, reusing learned latent representations, transferring model states, or pre-training on data generated by process-based models. The suitability of each approach depends on the similarity between the source and target domains and on which hydrologic information is expected to be transferable. These approaches allow the ML model to leverage physical representations learned in data-rich domains to constrain behavior in ungauged catchments, effectively acting as a "physical prior" for the target catchment and mitigate the risk of overfitting in data-poor environments. Constraining the target model parameters reduces the effective degrees of freedom available during local calibration and limiting overfitting in data-poor environments. The scientific value of transfer learning depends on whether the knowledge transferred between domains reflects hydrologically relevant similarities. In this way, transfer learning addresses the observational dependency gap while also contributing to process diagnosis. Its value lies not merely in borrowing information from data-rich catchments, but in determining what hydrologic knowledge is transferable, under which conditions, and with what uncertainty.

3.4. Residual Structural Learning

Neither process-based nor data-driven modelling approaches can fully represent all controls on hydrologic behavior. Process-based models encode established physical concepts but may omit poorly understood processes, simplified interactions, human influences, and non-stationary system responses. ML models can approximate these unresolved relationships, but their unconstrained use may obscure rather than clarify the source of model improvement. Residual structural learning combines the two approaches by assigning each a distinct role: the process-based model represents the hydrologic behavior explained by current theory, while the ML component learns the systematic component that remains unexplained.

The total prediction can be decomposed into a physical component from a process-based model and a learned residual from an ML model. The objective function then optimizes the ML model to learn only the systematic errors of the physical model. Such a structure ensures that known physics provides the baseline, while the ML model is tasked exclusively with learning structural uncertainty (or “structural residuals”)

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

capturing processes not resolved explicitly in the process-based model. While more sophisticated architectures may update both parameter sets jointly the aim remains the same – to represent the processes not resolved in the physically based model. Systematic residuals may arise from simplified process representations, inadequate parameterizations, scale mismatch, missing storage mechanisms, imperfect routing, or interactions that are difficult to formulate explicitly (Smith et al., 2008; Gupta et al., 2012). They may also reflect anthropogenic influences, including reservoir operations, groundwater abstraction, irrigation return flows, drainage modification, and land-use change, that are often incompletely represented in conventional process-based models (Fowler and Wilby, 2010; Samadi and Meadows, 2017; Fowler et al., 2021; Zafarmomen et al., 2024). In this way the ML model training reveals a new process-based understanding that was previously undisclosed.

Collectively, the four strategies discussed here address different but interconnected dimensions of the ML paradox. Explainability interrogates learned representations; physics-informed learning constrains predictions using established principles; transfer learning tests which knowledge can be generalized across systems; and residual structural learning reveals where current process representations remain incomplete. Their integration creates a bidirectional relationship in which hydrologic theory guides ML, while ML diagnoses, challenges, and potentially extends hydrologic theory. Process-based models provide the mechanistic and theoretical foundation, while ML models capture complex, nonlinear relationships and system behaviors from observations. By treating these as complementary and bidirectional approaches, ML models can be simultaneously predictive, physically consistent, mechanistically interpretable, and hence "scientifically useful".

4. Conclusions and Future Directions

ML has transformed hydrologic modeling by demonstrating that highly accurate predictions can often be achieved without explicitly representing the physical processes that explain catchment behavior. This capability challenges long-standing assumptions regarding the relationship between prediction and understanding, giving rise to the "Machine Learning Paradox in Hydrology": the same flexibility that enables ML to learn complex, nonlinear hydrologic behavior can also limit its ability to provide physically interpretable explanations of that behavior. The paradox is therefore not simply a technical issue arising from black-box algorithms but reflects a long-standing question of whether predictive success translates into meaningful advances in understanding, and whether prediction can be transformed into explanation and testable hypotheses.

Here, we argue that resolving this paradox does not require choosing between ML and process-based modelling. Rather, it requires redefining their respective roles within the scientific process. Consistent with the philosophy articulated by Beven (2012), hydrologic models should not be viewed as static, definitive representations of reality, but as dynamically evolving and falsifiable hypotheses of catchment system behavior that are continually tested, refined, and challenged through new observations and theories. Resolving the tension between predictive skill and mechanistic interpretability therefore requires a transition toward a hybrid modeling paradigm structured around the four foundational integration strategies.

The four integration strategies of explainability, physics-informed ML, transfer learning, and residual structural learning, represent complementary pathways towards the goal of ML being a tool for improved process understanding. These approaches must become increasingly capable of distinguishing errors arising from observations, forcing, parameterization, and model structure, rather than treating all prediction errors as equivalent. They must also use these diagnostic strategies to determine when existing process representations are inadequate and when learned relationships indicate previously unresolved hydrologic behavior. Most importantly, future research should focus not simply on improving predictive accuracy, but on determining why models succeed or fail, which physical assumptions challenge those outcomes, and whether the resulting evidence supports refinement of existing theory or development of new hypotheses.

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Finally, the next frontier lies in hybrid modeling frameworks capable of scientific self-diagnosis. Rather than simply correcting prediction errors, future models should self-identify when those errors arise from missing physical processes, inadequate observations, or limitations in existing theory. Equally important, they should distinguish situations in which statistical learning has merely compensated for structural weaknesses from those in which it has revealed new and generalizable hydrologic behavior. Such capabilities would transform hybrid models from predictive tools into mechanistic reasoning systems that actively support hypothesis testing, model falsification, mechanistic interpretability, and theory development.

5. Acknowledgements

VS is supported by the US National Science Foundation Directorate of Engineering (Grant # CMMI 2125283 and CBET 2429082). DA was supported by the Cooperative Institute for Research to Operations in Hydrology (CIROH) with funding under award NA22NWS4320003 from the NOAA Cooperative Institute Program. VM is supported by the US National Science Foundation (Grant # 2118329). CW is supported by a Sydney Horizon Fellowship. JG is supported by the US National Science Foundation (#2209139). T.W. acknowledges support by the U.S. Department of Defense's Strategic Environmental Research and Development Program under Project RC23-3707. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies. The authors also thank Dr. Louise Slater (University of Oxford) for her valuable comments on an earlier draft of this manuscript.

Author Contributions

VS and HJF conceptualized the manuscript and wrote the original draft. All authors (VS, HJF, DPA, EB, JLG, JSH, AJJ, VL, VMM, VM, YP, MS, TW, CW, CAMW, and NZ) contributed to reviewing, editing, and revising the manuscript and approved the final version.

Competing Interests

The authors declare that they have no competing interests.

6. References

1. Abbott, M., J. Bathurst, J. Cunge, P. O'Connell, and J. Rasmussen.1986. An introduction to the European Hydrological System—Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system, J. Hydrol., 87(1), 61–77.
2. Acuña Espinoza, E., Loritz, R., Kratzert, F., Klotz, D., Gauch, M., Álvarez Chaves, M. and Ehret, U., 2025. Analyzing the generalization capabilities of a hybrid hydrological model for extrapolation to extreme events. *Hydrology and Earth System Sciences*, *29*(5), pp.1277-1294. 10.5194/hess-29-1277-2025
3. Asher, M., Croke, B.F.W., Jakeman, A.J. and Peeters, L. 2015. A review of surrogate models and their application to groundwater modeling. Water Resources Research, 51(8): 5957-5973. 10.1002/2015WR016967
4. Bergström, S. 1976. Development and Application of a Conceptual Runoff Model for Scandinavian Catchments. SMHI Report RHO 7.
5. Beucler, T., Pritchard, M., Rasp, S., Ott, J., Baldi, P. and Gentine, P., 2021. Enforcing analytic constraints in neural networks emulating physical systems. Physical review letters, 126(9), p.098302. 10.1103/PhysRevLett.126.098302
6. Beven, K.J., 2006. A manifesto for the equifinality thesis. Journal of Hydrology, 320(1–2), pp.18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007.
7. Beven, K.J., 2012. Rainfall-runoff modelling: the primer. John Wiley & Sons.

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

8. Burnash, R. J. C. 1995. The NWS River Forecast System—Catchment Modeling. In: Singh, V. P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, Colorado, USA.
9. Clark, M. P., D. Kavetski, and F. Fenicia. 2011. Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resour. Res., 47, W09301, doi:10.1029/2010WR009827.
10. Costigan, K.H., Daniels, M.D., and Dodds, W.K., 2015. Fundamental spatial and temporal disconnections in the hydrology of an intermittent prairie headwater network. Journal of Hydrology, 522, pp.305–316. https://doi.org/10.1016/j.jhydrol.2014.12.031.
11. De la Fuente, L. A., Ehsani, M. R., Gupta, H. V., and Condon, L. E.: Toward interpretable LSTM-based modeling of hydrological systems, Hydrol. Earth Syst. Sci., 28, 945–971, https://doi.org/10.5194/hess-28-945-2024, 2024.
12. Elkouk, A., Pokhrel, Y., Livneh, B., Payton, E., Luo, L., Cheng, Y., Dagon, K., Swenson, S., Wood, A.W., Lawrence, D.M., 2024. Toward understanding parametric controls on runoff sensitivity to climate in the community land model: A case study over the Colorado River headwaters. Water Resources Research, 60(12): e2024WR037718. 10.1029/2024WR037718.
13. Fowler, H.J. and Wilby, R.L., 2010. Detecting changes in seasonal precipitation extremes using regional climate model projections: Implications for managing fluvial flood risk. Water Resources Research, 46(3). 10.1029/2008WR007636.
14. Fowler, H.J., Lenderink, G., Prein, A.F., Westra, S., Allan, R.P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Do, H.X. and Guerreiro, S., 2021. Anthropogenic intensification of short-duration rainfall extremes. Nature Reviews Earth & Environment, 2(2), pp.107-122. https://doi.org/10.1038/s43017-020-00128-6
15. Frame, J.M., Kratzert, F., Klotz, D., Gauch, M., Shalev, G., Gilon, O., Qualls, L.M., Gupta, H.V., and Nearing, G.S., 2022. Deep learning rainfall–runoff predictions of extreme events. Hydrology and Earth System Sciences, 26, pp.3377–3392. https://doi.org/10.5194/hess-26-3377-2022
16. Green, W.R., 1994. Water quality assessment of Maumelle and Winona reservoir systems, central Arkansas, May 1989-October 1992 (Vol. 93, No. 4218). US Geological Survey.
17. Gupta, H.V., Clark, M.P., Vrugt, J.A., Abramowitz, G. and Ye, M., 2012. Towards a comprehensive assessment of model structural adequacy. Water Resources Research, 48(8). 10.1029/2011WR011044.
18. Gutiérrez-Jurado, K.Y., Partington, D., Batelaan, O., Cook, P., and Shanafield, M., 2019. What triggers streamflow for intermittent rivers and ephemeral streams in low-gradient catchments in Mediterranean climates. Water Resources Research, 55, pp.9926–9946. https://doi.org/10.1029/2019WR025041.
19. Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, *9*(8), pp.1735-1780. 10.1162/neco.1997.9.8.1735.
20. Jastram, J.D., Krstolic, J.L., Moyer, D. and Hyer, K., 2015. Fluvial geomorphology and suspended-sediment transport during construction of the Roanoke River Flood Reduction Project in Roanoke, Virginia, 2005–2012 (No. 2015-5111). US Geological Survey.
21. Jeong, J., Goodall, J. L., and Quinn, J. D. 2026. Comparing strategies for training LSTM models for street-scale urban flood prediction in Norfolk, Virginia. Journal of Hydrology: Regional Studies, 65, 103405. https://doi.org/10.1016/j.ejrh.2026.103405.
22. Kaplan, N.H., Blume, T., and Weiler, M., 2022. Event controls on intermittent streamflow in a temperate climate. Hydrology and Earth System Sciences, 26, pp.2671–2696. https://doi.org/10.5194/hess-26-2671-2022.
23. Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S. and Yang, L., 2021. Physics-informed machine learning. Nature Reviews Physics, 3(6), pp.422-440.10.1038/s42254-021-00314-5
24. Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N. and Kumar, V., 2017. Theory-guided data science: A new paradigm for scientific

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

discovery from data. IEEE Transactions on knowledge and data engineering, 29(10), pp.2318-2331. 10.1109/TKDE.2017.2720168

25. Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A.K., Hochreiter, S. and Nearing, G.S., 2019. Toward improved predictions in ungauged basins: Exploiting the power of machine learning. Water Resources Research, 55(12), pp.11344-11354. 10.1029/2019WR026065
26. LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444. 10.1038/nature14539
27. Lees, T., Buechel, M., Anderson, B., Slater, L., Reece, S., Coxon, G., and Dadson, S.J., 2021. Benchmarking data-driven rainfall–runoff models in Great Britain: A comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. Hydrology and Earth System Sciences, 25, pp.5517–5534. https://doi.org/10.5194/hess-25-5517-2021.
28. Lees, T., Reece, S., Kratzert, F., Klotz, D., Gauch, M., De Bruijn, J., Kumar Sahu, R., Greve, P., Slater, L., and Dadson, S.J., 2022. Hydrological concept formation inside long short-term memory (LSTM) networks. Hydrology and Earth System Sciences, 26, pp.3079–3101. https://doi.org/10.5194/hess-26-3079-2022.
29. Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99(D7), pp.14415-14428. 10.1029/94JD00483.
30. Lim, B., Arık, S.Ö., Loeff, N. and Pfister, T., 2021. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International journal of forecasting*, *37*(4), pp.1748-1764.
31. Lipton, Z.C., 2018. The mythos of model interpretability. *Communications of the ACM*, *61*(10), pp.36-43. 10.1016/j.ijforecast.2021.03.012.
32. Ludwig, A.H., 1992. Flow duration and low-flow characteristics of selected Arkansas streams (Vol. 92, No. 4026). US Department of the Interior, US Geological Survey.
33. Mangukiya, N.K. and Sharma, A., 2025. Integrating reservoir dynamics into differentiable process-based hydrological model for enhanced streamflow estimation. Water Resources Research, 61(7), p.e2025WR040268.
34. Maxwell, R., and N. Miller. 2005. Development of a coupled land surface and groundwater model., J. Hydrometeorol., 6, 233–247. 10.1175/JHM422.1.
35. McSpadden, D., Goldenberg, S., Roy, B., Schram, M., Goodall, J.L. and Richter, H., 2024. A comparison of machine learning surrogate models of street-scale flooding in Norfolk, Virginia. Machine Learning with Applications, 15, p.100518. 10.1016/j.mlwa.2023.100518.
36. Mendoza, P.A., Clark, M.P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G. and Gupta, H., 2015. Are we unnecessarily constraining the agility of complex process-based models?. Water Resources Research, 51(1), pp.716-728. 10.1002/2014WR015820.
37. Molnar, C., 2022. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. 2nd ed. Self-published. Available at: https://christophm.github.io/interpretable-ml-book/
38. Myers, R.D., 2022. Floodplain Sedimentation in the Lowermost Mississippi River. Tulane University.
39. Nearing, G.S., Kratzert, F., Sampson, A.K., Pelissier, C.S., Klotz, D., Frame, J.M., Prieto, C. and Gupta, H.V., 2021. What role does hydrological science play in the age of machine learning?. Water Resources Research, 57(3), p.e2020WR028091.
40. Raissi, M., Perdikaris, P. and Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 378, pp.686-707. 10.1016/j.jcp.2018.10.045.
41. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. and Prabhat, F., 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), pp.195-204. 10.1038/s41586-019-0912-1

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



42. Roy, B., Goodall, J.L., McSpadden, D., Goldenberg, S. and Schram, M., 2025. Forecasting multi-step-ahead street-scale nuisance flooding using a Seq2seq LSTM surrogate model for real-time application in a coastal-Urban City. Journal of Hydrology, 656, p.132697. 10.1016/j.jhydrol.2025.132697.
43. Saberian, M., Samadi, V., & Popescu, I. 2026. Probabilistic hierarchical interpolation and interpretable neural network configurations for flood prediction. Hydrology and Earth System Sciences, 30(2), 371-399.
44. Tabas, S. and Samadi, S., 2022. Variational Bayesian dropout with a Gaussian prior for recurrent neural networks application in rainfall–runoff modeling. Environmental Research Letters, 17(6), p.065012. 10.1088/1748-9326/ac7247.
45. Sadeghi Tabas, S., Samadi, V., Wilson, C. and Bhattacharya, B. 2025. Probabilistic physics-guided deep neural networks with recurrence and attention mechanisms for interpretable daily streamflow simulation. Water Resources Research, 61(9), p.e2025WR040173.
46. Samadi, S. and Meadows, M.E., 2017. The transferability of terrestrial water balance components under uncertainty and nonstationarity: A case study of the Coastal Plain watershed in the Southeastern USA. *River Research and Applications*, *33*(5), pp.796-808.
47. Sanford, W. E., Nelms, D. L., Pope, J. P., & Selnick, D. L. 2012. Quantifying components of the hydrologic cycle in Virginia using chemical hydrograph separation and multiple regression analysis (No. 2011-5198). US Geological Survey.
48. Smith, P.J., Beven, K.J. and Tawn, J.A., 2008. Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. *Water resources research*, *44*(1). 10.1029/2006WR005205.
49. Shalev-Shwartz, S. and Ben-David, S. 2014. Understanding machine learning: From theory to algorithms. Cambridge university press.
50. Shen, C., Appling, A.P., Gentine, P., Bandai, T., Gupta, H., Tartakovsky, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li, L. and Liu, X., 2023. Differentiable modelling to unify machine learning and physical models for geosciences. Nature Reviews Earth & Environment, 4(8), pp.552-567. 10.1038/s43017-023-00450-9
51. Slater, L.J., Arnal, L., Boucher, M.A., Chang, A.Y.Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L. and Villarini, G., 2023. Hybrid forecasting: blending climate predictions with AI models. Hydrology and earth system sciences, 27(9), pp.1865-1889. 10.5194/hess-27-1865-2023.
52. Song, Y., Bindas, T., Shen, C., Ji, H., Knoben, W.J.M., Lonzarich, L., Clark, M.P., Liu, J., van Werkhoven, K., Lamont, S., Denno, M., Pan, M., Yang, Y., Rapp, J., Kumar, M., Rahmani, F., Thébault, C., Adkins, R., Halgren, J., Patel, T., Patel, A., Sawadekar, K.A., and Lawson, K., 2025. High-resolution national-scale water modeling is enhanced by multiscale differentiable physics-informed machine learning. Water Resources Research, 61, e2024WR038928. https://doi.org/10.1029/2024WR038928
53. Staudinger, M., Herzog, A., Loritz, R., Houska, T., Pool, S., Spieler, D., Wagner, P.D., Mai, J., Kiesel, J., Thober, S. and Guse, B., 2025. How well do process-based and data-driven hydrological models learn from limited discharge data?. Hydrology and Earth System Sciences, 29(19), pp.5005-5029. 10.5194/hess-29-5005-2025.
54. Sundararajan, M., Taly, A. and Yan, Q., 2017, July. Axiomatic attribution for deep networks. In International conference on machine learning (pp. 3319-3328). PMLR.
55. Thomas Jr, W.O., Monde, M.C. and Davis, S.R., 2000. Estimation of time of concentration for Maryland streams. Transportation research record, 1720(1), pp.95-99. 10.3390/hydrology6010020.
56. Thomas, E.D., Venkataraman, K., Chraibi, V. and Kannan, N., 2019. Hydrologic trends in the upper Nueces River basin of Texas—implications for water resource management and ecological health. Hydrology, 6(1), p.20.
57. Vapnik, V.N., 1998. Statistical learning theory wiley-interscience. New York.

https://doi.org/10.5194/egusphere-2026-6531
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



58. Vischer, M.A., Otero, N., and Ma, J., 2025. Spatially resolved rainfall streamflow modeling in central Europe. Hydrology and Earth System Sciences, 29, pp.5233–5250. 10.5194/hess-29-5233-2025.
59. Wagener, T., Sivapalan, M., Troch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Kumar, P., Rao, P.S.C., Basu, N.B. and Wilson, J.S., 2010. The future of hydrology: An evolving science for a changing world. Water Resources Research, 46(5), W05301. https://doi.org/10.1029/2009WR008906.
60. Wang, Q., Guillaume, J.H.A., Jakeman, J.D., Bennett, F.R., Croke, B,F.W., Yang, T. and Jakeman, A.J. 2023. A decision-relevant factor-fixing framework: application to uncertainty analysis uncertainty analysis of a high-dimensional water quality model. Water Resources Research, 59(8), https://doi.org/10.1029/2022WR032194.
61. Wang, C., Jiang, S., Zheng, Y., Han, F., Kumar, R., Rakovec, O. and Li, S., 2024. Distributed hydrological modeling with physics-encoded deep learning: A general framework and its application in the Amazon. *Water Resources Research*, *60*(4), p.e2023WR036170.
62. Wang, Y., and Karimi, H.A., 2022. Impact of spatial distribution information of rainfall in runoff simulation using deep learning method. Hydrology and Earth System Sciences, 26, pp.2387–2403. https://doi.org/10.5194/hess-26-2387-2022.
63. Ward E. Sanford, David L. Nelms, Jason P. Pope, and David L. Selnick. 2011. Quantifying Components of the Hydrologic Cycle in Virginia using Chemical Hydrograph Separation and Multiple Regression Analysis. Accessed on June 17, 2026. available at https://www.researchgate.net/profile/Ward-Sanford-2/publication/260638698_Quantifying_Components_of_the_Hydrologic_Cycle_in_Virginia_using_Chemical_Hydrograph_Separation_and_Multiple_Regression_Analysis/links/0c960531e2792dac7f000000/Quantifying-Components-of-the-Hydrologic-Cycle-in-Virginia-using-Chemical-Hydrograph-Separation-and-Multiple-Regression-Analysis.pdf
64. Wilder, H.B. and Simmons, C.E., 1982. Program for evaluating stream quality in North Carolina (No. 2185-A). USGPO.
65. Wood, E.F., Roundy, J.K., Troy, T.J., Van Beek, L.P.H., Bierkens, M.F., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J. and Gochis, D., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. Water Resources Research, 47(5). 10.1029/2010WR010090
66. Zafarmomen, N., Alizadeh, H., Bayat, M., Ehtiat, M., & Moradkhani, H. 2024. Assimilation of sentinel-based leaf area index for modeling surface-ground water interactions in irrigation districts. Water Resources Research, 60(10), e2023WR036080. https://doi.org/10.1029/2023WR036080