JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2025JH000749

Baseflow Identification via Explainable AI With Kolmogorov-Arnold Networks

Chuyang Liu[1], Tirthankar Roy[2], Daniel M. Tartakovsky[3], and Dipankar Dwivedi[1]

[1]Geochemistry Department, Lawrence Berkeley National Lab, Berkeley, CA, USA, [2]Department of Civil and Environmental Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA, [3]Department of Energy Science and Engineering, Stanford University, Stanford, CA, USA

Key Points:

- Kolmogorov-Arnold networks (KANs) enhance interpretability of machine-learned hydrological models
- KAN-derived symbolic formulations outperform state-of-the-art semi-empirical aridity indices
- KAN-identified functional form yields an analytical index with fewer fitting parameters and improved performance

Correspondence to:
D. M. Tartakovsky,
tartakovsky@stanford.edu

Citation:
Liu, C., Roy, T., Tartakovsky, D. M., & Dwivedi, D. (2025). Baseflow identification via explainable AI with Kolmogorov-Arnold networks. *Journal of Geophysical Research: Machine Learning and Computation*, *2*, e2025JH000749. https://doi.org/10.1029/2025JH000749

Received 5 JUN 2025
Accepted 16 SEP 2025

Author Contributions:
Conceptualization: Chuyang Liu, Dipankar Dwivedi
Data curation: Chuyang Liu
Formal analysis: Chuyang Liu
Funding acquisition: Daniel M. Tartakovsky, Dipankar Dwivedi
Investigation: Chuyang Liu
Methodology: Chuyang Liu, Tirthankar Roy, Daniel M. Tartakovsky, Dipankar Dwivedi
Project administration: Dipankar Dwivedi
Resources: Dipankar Dwivedi
Supervision: Dipankar Dwivedi
Visualization: Chuyang Liu
Writing – original draft: Chuyang Liu
Writing – review & editing: Tirthankar Roy, Daniel M. Tartakovsky, Dipankar Dwivedi

Abstract Hydrological models often involve constitutive laws that may not be optimal in every application. We propose to replace such laws with the Kolmogorov-Arnold networks (KANs), a class of neural networks designed to identify symbolic expressions. We demonstrate KAN's potential on the problem of baseflow identification, a notoriously challenging task plagued by significant uncertainty. KAN-derived functional dependencies of the baseflow components on the aridity index outperform their original counterparts; they demonstrate that water availability, rather than potential evapotranspiration, drives baseflow by constraining actual evapotranspiration under arid conditions. On a test set, they increase the Nash-Sutcliffe efficiency (NSE) by 65%, decrease the root mean squared error by 29%, and increase the Kling-Gupta efficiency by 34%. This superior performance is achieved while reducing the number of fitting parameters from three to two. Next, we use data from 378 catchments across the continental United States to refine the water-balance equation at the mean-annual scale. The KAN-derived equations based on the refined water balance outperform both the current aridity index model, with up to a 105% increase in NSE, and the KAN-derived equations based on the original water balance. While the performance of our model and tree-based machine learning methods is similar, KANs offer the advantage of simplicity and transparency and require no specific software or computational tools. This case study focuses on the aridity index formulation, but the approach is flexible and transferable to other hydrological processes.

Plain Language Summary Equations used in hydrologic model are often suboptimal, resulting in reduced prediction accuracy and efficiency. We implemented Kolmogorov-Arnold networks (KAN), a machine learning algorithm for deriving symbolic formulations, to estimate groundwater recharge and showed that it outperforms an existing state-of-the-art semi-empirical formulation. In hydrology, Nash-Sutcliffe efficiency (NSE), root mean squared error (RMSE), and Kling-Gupta efficiency (KGE) are commonly used to evaluate model performance. Higher NSE and KGE values indicate better performance, while lower RMSE values are preferable. Our results show that NSE increased by 71%, RMSE decreased by 32%, and KGE improved by 25%. In addition, KAN identifies an optimal functional form and can be used to derive new analytical formulas using the prior knowledge. The KAN-inspired equation outperformed the original formulation and reduced the fitting parameters. Furthermore, we refined the water-balance equation at the mean-annual scale and showed that, based on the new water-balance equation, KAN can derive new formulations that are superior to the original aridity index formulations (up to 105% increase in NSE) and KAN-derived equations based on the original water balance. These findings highlight the significant potential of KAN to advance the scientific understanding of a wide range of hydrologic processes.

1. Introduction

Hydrological models are used to quantify the water cycle and its interactions with human activities and ecosystems. They provide crucial insights into water resource dynamics and support decision-making in various applications. Data-driven construction of such models, or empirical constitutive laws therein, typically involves either parametric (non-symbolic) or symbolic regression. Non-symbolic regression fits data to a predefined model, while symbolic regression discovers the best-fitting equations from the data. Examples of non-symbolic regression include the curve number method (Ponce & Hawkins, 1996), which uses an empirical formula to relate runoff to land use, soil type, and hydrological conditions; the Penman-Monteith equation to estimate evapotranspiration from multiple meteorological and physiological factors (Allen et al., 1998); and various functional forms of the moisture retention curve (Assouline & Tartakovsky, 2001) that relates soil saturation to pore

pressure. Such predefined constitutive relations introduce bias due to their failure to capture complex and dynamic relationships between relevant hydrologic variables. Further improvements are needed to enhance generalizibility across diverse hydrological conditions and ecosystems.

In principle, deep learning and artificial intelligence, for example, (Hong et al., 2023, 2024; Li et al., 2023), can be used to achieve this goal. A prime example from this class is a deep neural network (NN), a universal approximator of any continuous function. It yields black-box models, aka equation-free constitutive relations, which are hard to interpret. NNs-based symbolic regression can ameliorate this limitation by yielding a closed-form formula for various constitutive laws (Kubalík et al., 2023; Martius & Lampert, 2016; Sahoo et al., 2018; Zhou & Pan, 2022). Such symbolic NNs can handle binary (e.g., addition, subtraction) and unary (e.g., exponential, sine, logarithm) operations in different units. The significance of each mathematical operation is represented by the magnitude of its weight that is adjusted, together with the biases of neurons, during the NN training. The combination of these optimized weights and operations gives a final formula. Despite these advantages, symbolic regression has limited transparency in the intermediate process, leading to difficulty in error identification.

The Kolmogorov-Arnold networks (KANs) address these challenges by leveraging the Kolmogorov-Arnold representation theorem. The latter states that any continuous function can be represented by a combination of continuous functions of a single variable and enhances the efficiency of searching the solution spaces (Z. Liu et al., 2024). While KANs use gradient-based optimization like traditional NNs, the selection of activation functions is fundamentally different. In contrast to fixed activation functions, KAN learns the optimal activation functions from data. This endows it with the capability both to visualize learned activation functions and to incorporate prior knowledge into symbolic formulas. These improvements are especially important when the target function is a non-symbolic function where symbolic regression may fail but KAN can handle it through the numerical approximation (Z. Liu et al., 2024). In addition, unlike symbolic regression providing only binary success/failure outcomes, KAN can visualize intermediate processes, facilitating interpretation and error identification (Z. Liu et al., 2024).

Since KAN is domain-independent and does not rely on assumptions specific to hydrology, it can be applied across various fields for scientific discoveries. We demonstrate the value of KANs by identifying optimal functional dependencies between the runoff components and the aridity index. The latter is defined as the ratio of mean annual potential evapotranspiration (PET) to mean annual precipitation. It represents regional dryness and its impact on water balance. This well-established indicator has significant implications for hydrological modeling. For example, semi-empirical models were used to estimate long-term (mean-annual) direct runoff and baseflow from the aridity index across 378 catchments in the contiguous United States (Meira Neto et al., 2020). KANs are suitable for baseflow identification because of their capability to capture nonlinear and complex relationships between hydrological variables effectively. The modular structure of KANs can be visualized and derived as mathematical operations to provide insights into how hydrological variables affect the baseflow processes and potentially enhance the physical understanding of underlying hydrological mechanisms. While this study primarily focuses on the impacts of the aridity index on baseflow, the modular framework of KAN can be easily extended to consider additional hydrological variables to improve the baseflow estimation in the future.

By providing more precise formulations of the aridity index, we improve the accuracy and interpretability of hydrological models by addressing two key questions. First, can a KAN identify novel functional dependencies of both baseflow and direct runoff on the aridity index? Second, how do the KAN-derived symbolic relations involving the aridity index compare with their current counterparts in terms of accuracy, simplicity, and predictive power? By exploring these questions, we demonstrate that KANs yield superior models, highlighting their potential to improve hydrological modeling in various applications.

2. Materials and Methods

2.1. Data Sets

The Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) data set (Addor et al., 2017) includes daily streamflow, meteorology, and other attributes for 671 catchments across the contiguous United States. Using this data set, Meira Neto et al. (2020) calculated the mean-annual precipitation, P; the mean-annual PET; the mean-annual direct runoff, Q_D; the mean-annual baseflow, Q_B; and the aridity index $\phi = \mathrm{PET}/P$. Briefly, after excluding catchments that did not meet the criteria for data completeness, catchment size, snow

fraction, and positive annual evaporation, 378 out of 671 CAMELS catchments were selected for analysis. Daily streamflow data were partitioned into direct runoff and baseflow using the low-pass filter method (Lyne & Hollick, 1979). Daily PET was calculated using the Reference-crop Penman-Monteith formulation, based on incoming solar radiation and wind speed at two m from the gridMET data set (Abatzoglou, 2013), along with other variables from the CAMELS data set. Finally, daily values were aggregated to mean-annual variables for each catchment. The processed data set is available at the repository (Meira, 2020). Our analysis relies on this data set, including the definition of direct runoff. This definition represents the catchment's fast response to the precipitation and comprises infiltration-excess and saturation-excess runoff, as well as subsurface storm flow. Additional details of data processing are available in Meira Neto et al. (2020).

2.2. Aridity Index Formulations

Aridity index formulations of Meira Neto et al. (2020) are derived using the Budyko and Budyko (1974) and the L'vovich (1979) frameworks. In the former, P is decomposed into mean annual streamflow, Q, and evaporation, E, such that

$$\frac{E}{P} = 1 - \frac{Q}{P} \equiv f_E(\phi), \tag{1}$$

where the function f_E describes a relationship between E/P and ϕ (aridity index).

In the L'vovich (1979) framework, annual precipitation p is partitioned into annual direct runoff q_D, annual baseflow q_B, and annual evapotranspiration e. At the mean-annual time scale, this partition is written as

$$\frac{Q}{P} = \frac{Q_D}{P} + \frac{Q_B}{P}. \tag{2}$$

From Equation 1, Q/P is a function of ϕ. Assuming that both terms on the right-hand-side of Equation 2 are also functions of ϕ,

$$\frac{Q}{P} = f_D(\phi) + f_B(\phi), \tag{3a}$$

where

$$f_D(\phi) = \frac{Q_D}{P} \quad \text{and} \quad f_B(\phi) = \frac{Q_B}{P}. \tag{3b}$$

Based on the observed patterns of $f_D(\phi)$ and $f_B(\phi)$ in the aforementioned data sets, Meira Neto et al. (2020) assigned to these functions an exponential form,

$$f_i(\phi) = \exp(-\phi^{a_i} + b_i)^{c_i}, \qquad i = D, B. \tag{4}$$

where a_i, b_i, and c_i are the fitting parameters in the ith function. After randomly splitting the data sets into halves to fit parameters and evaluate model performance 100 times, Meira Neto et al. (2020) obtained

$$\frac{Q_B}{P} \equiv f_B(\phi) = \exp\left(-\phi^{1.71} - 0.873\right)^{1.05} \tag{5}$$

and

$$\frac{Q_D}{P} \equiv f_D(\phi) = \exp\left(-\phi^{0.77} - 0.864\right)^{1.06}. \tag{6}$$

These relations capture the significant variability in Q_D/P and Q_D/P across 378 catchments in the contiguous US (Meira Neto et al., 2020).

2.3. Development of KAN Models

We train KAN models on the same data set, which was randomly split into a training set and a test set with an 8 to 2 ratio. To avoid overfitting, we used a 10-fold cross-validated grid-search method (Pedregosa et al., 2011). In this method, the training data set was split into a pre-training set (any nine folds out of the 10 folds) and a validation set (the remaining one fold out of the 10 folds). The pre-training set was used to develop a KAN with a given set of hyperparameters, and the validation set was used to evaluate the model performance by the coefficient of determination, R^2. The Broyden Fletcher Goldfarb Shanno algorithm (Head & Zerner, 1985) was used to find model parameters that minimize the loss, defined as the root mean squared error|root mean square error (RMSE) (Chai & Draxler, 2014). The grid search in this method iterated through all sets of hyperparameters and each set of hyperparameters has an average R^2. The optimal set of hyperparameters was chosen based on the highest average R^2. The model was trained with the optimal hyperparameters and the training set was selected as the final KAN model.

In the 10-fold cross-validation with a single set of hyperparameters, 10 KAN models were developed. Each KAN model was developed through five steps. In step 1, a fully connected KAN is initially trained to be sparse by regularization. To retain significant connections, pruning is used in step 2 to remove insignificant ones. Based on learned activation functions in the remaining connections, symbolic functions are set in step 3 through the optimal curve-fitting with the highest R^2. In step 4, affine parameters are further refined by additional training. In the final step, the combination of trained parameters and symbolic functions is expressed as a symbolic formula. After training each KAN model, the ten R^2 values were averaged to obtain the average R^2 for a single set of hyperparameters.

In this study, hyperparameters, including the number of layers (2–3), the number of neurons in each layer (1–5), the seed number to initialize the model (1–10), and the number of grid intervals for each spline (2–4), are evaluated. Specifically, the number of layers was predefined first (either 2 or 3), and the remaining hyperparameters were tuned subsequently. For each predefined number of layers, the optimal combinations of remaining hyperparameters were selected based on the highest average R^2. The maximum of highest average R^2 was further used to select the optional number of layers along with associated optional hyperparameters.

2.3.1. Kolmogorov-Arnold Networks

The Kolmogorov-Arnold representation theorem states that a continuous multivariate function on a bounded domain can be expressed as a summation of finite univariate functions (Braun & Griebel, 2009; Kolmogorov, 1956). Based on this theorem, Z. Liu et al. (2024) proposed a NN architecture called KAN as an alternative to multi-layer perceptrons (MLPs) with arbitrary widths and depths. Unlike MLPs, which have fixed activation functions, KAN utilizes learnable univariate functions parameterized by B-splines. Each activation function in a KAN can be set symbolically, forming the final symbolic formula.

Given an input vector $\mathbf{x}$, the output of a KAN with L layers is the composition of each layer:

$$\mathrm{KAN}(\mathbf{x}) = (\Psi_{L-1} \circ \Psi_{L-2} \circ \cdots \circ \Psi_1 \circ \Psi_0)\,\mathbf{x}, \tag{7}$$

where Ψ_l is the activation function matrix for the lth KAN layer. In a KAN, neuron i at layer l is denoted by (l,i), neuron j at layer $l+1$ is denoted by $(l+1,j)$, and the activation function connecting (l,i) and $(l+1,j)$ is denoted by $\psi_{l,j,i}$. The input to neuron $(l+1,j)$ is the summation of all post-activation values from connected neurons at layer l:

$$x_{l+1,j} = \sum_{i=1}^{n_l} \psi_{l,j,i}(x_{l,i}), \quad j = 1, \ldots, n_{l+1} \tag{8}$$

Additional information about KANs is available in Appendix A. The learned activation function can be set to an evaluated symbolic function with the highest R^2. Currently, KAN supports $x, x^2, x^3, x^4, 1/x, 1/x^2, 1/x^3, 1/x^4, \sqrt{x}, 1/\sqrt{x}, \exp(x), \log(x), |x|, \sin(x), \tan(x), \tanh(x), \mathrm{sigmoid}(x), \mathrm{sign}(x), \arcsin(x), \arctan(x), \mathrm{arctanh}(x), 0$, Gaussian function, $\cosh(x)$, and other user-defined functions.

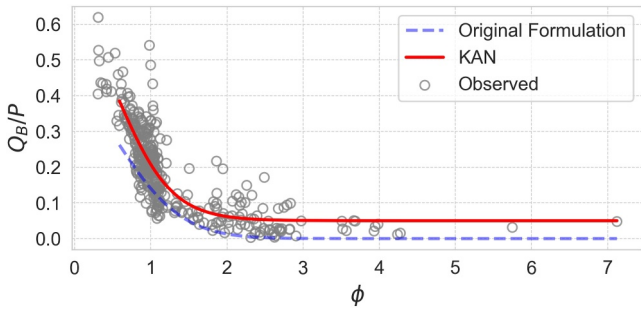

Figure 1. Observed and simulated dependence of normalized baseflow, Q_B/P, on aridity index, ϕ, at 378 catchment Attributes and Meteorology for Large-sample Studies catchments. The field observations are represented by the circles. The predictions obtained via the KAN-derived relation (Equation 9) and the original aridity index formulation (Equation 5) are shown by the dashed and solid lines, respectively. Graphical representation begins at $\phi \approx 0.05$ instead of 0 due to limited data availability.

2.4. Model Performance Evaluation

The test data set contains data unseen by the KAN during its training. We evaluate the model's performance on this data set in terms of the following metrics: the Nash-Sutcliffe efficiency (NSE), Kling-Gupta efficiency (KGE), RMSE, and R^2. Their precise definitions are provided in Appendix B. Higher values of NSE, KGE, and R^2 signify a better model performance; lower values of RMSE signify the same. While not typically used for daily or seasonal forecasting, these statistics help us evaluate how well our models capture the overall variability in mean annual runoff components across diverse catchments.

3. Results and Discussion

3.1. KAN-Derived Function f_B

To ensure the robustness of KAN, we use the 10-fold cross-validated grid-search method to train a KAN on the input-output pairs $\{\phi, Q_B/P\}$ extracted from the training data set (Methods). After iterating through all hyper-parameters in the grid search, the final KAN model structure is (1,1), that is, a single learned activation function explains the relationship between ϕ and Q_B/P. In the final KAN model, the learned activation function is a symbolic function, and the corresponding KAN-derived f_B equation is

$$\frac{Q_B}{P} = f_B(\phi) = 0.39 - 0.34\tanh(1.42\phi - 0.82). \tag{9}$$

Figure 1 depicts the curves of the aridity index $f_B(\phi)$ corresponding to the KAN-derived relation (Equation 9) and the original model (Equation 5). These are compared to the observed data $\{\phi, Q_B/P\}$ from the test data set across 378 CAMELS catchments. The observed data shows that baseflow normalized by precipitation has a monotonically inverse relationship with the aridity index. In relatively humid regions ($\phi < 1$; PET $< P$), baseflow declines rapidly with the increasing aridity index. This pattern aligns with expectations, as PET is a significant sink in the water balance. Given the same precipitation amount, higher PET leads to higher water loss and consequently results in lower baseflow. Both models capture this pattern but the KAN-derived relation has a steeper slope to better agree with the observations. In relatively arid regions ($\phi > 1$; PET $> P$), baseflow decreases more gradually with increasing aridity index. This reflects physical processes, where water availability—rather than PET—primarily limits actual evapotranspiration.

Table 1 confirms the visual comparison. The performance of both models on the training data set is consistent with that on the test data set, implying that both models neither overfit nor underfit. Under both training data set and the test data set, all model performance metrics of the KAN-derived f_B equation outperform the original aridity index formulation. The NSE increases by 71.3%, the RMSE decreases by 32.4%, and the KGE increases by 25% under the test data set.

Table 1
Performance Metrics of the Two Models of $Q_B/P = f_B(\phi)$, Where f_B Represents the Functional Relationship Between the Aridity Index and Baseflow Normalized by Precipitation

	Training data set		Test data set		
Metric	Original f_B	KAN-derived f_B	Original f_B	KAN-derived f_B	Improvement (%)
NSE	0.387	0.747	0.432	0.741	71.3
KGE	0.622	0.804	0.636	0.795	25.0
RMSE	0.0936	0.0601	0.0866	0.0585	32.4
R^2	0.740	0.747	0.733	0.744	1.50

Note. The KAN-derived f_B equation improves the model performance.

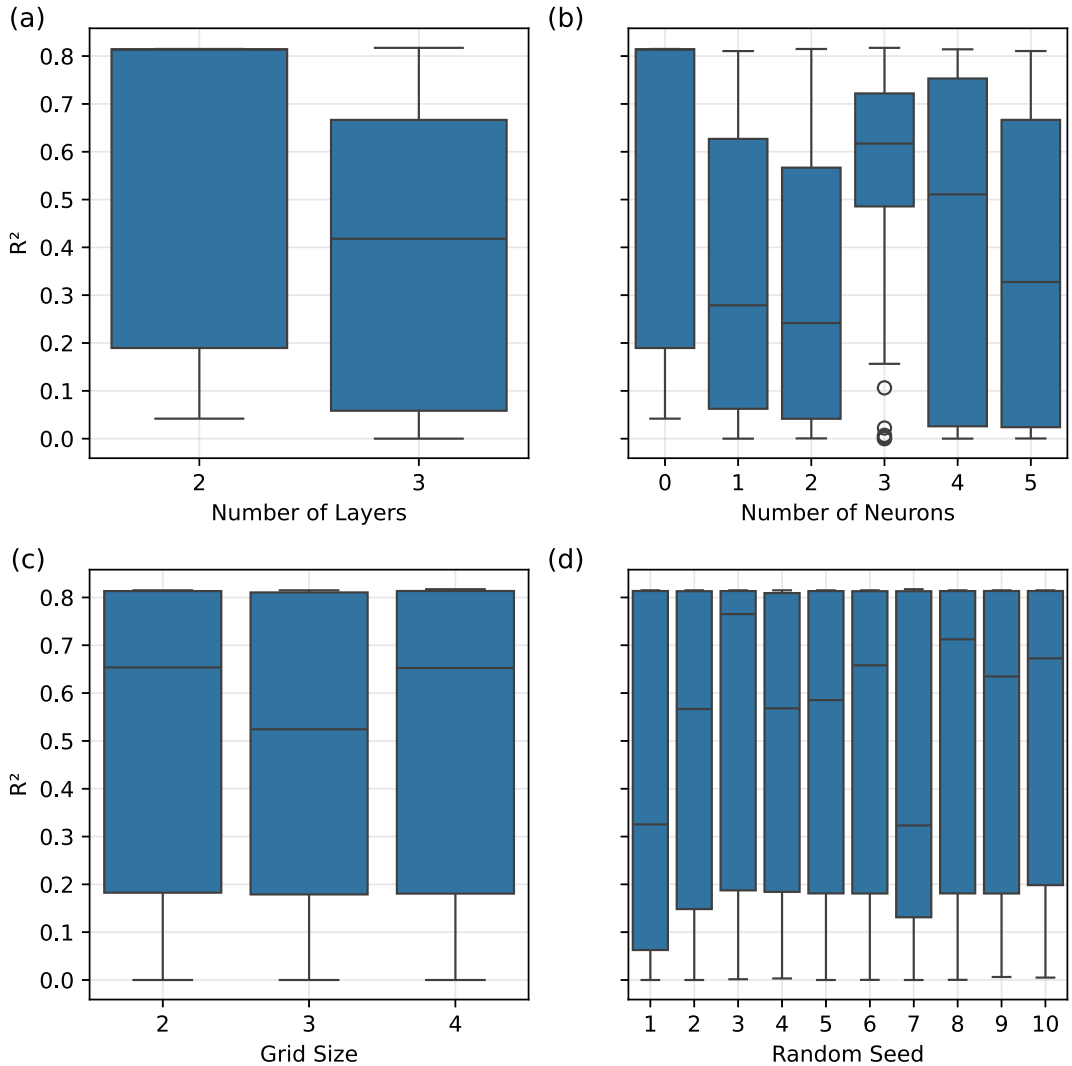

Figure 2. Sensitivity of KANs to hyperparameters (a) number of layers, (b) number of neurons in the hidden layer, (c) grid size, and (d) random seed. The box-plots represent the distribution of R^2 values on the test data set across different hyperparameter values.

While the KAN-derived relation (Equation 9) outperforms the original model (Equation 5) across all evaluated metrics, we further assess its robustness under varying climatic conditions. Following Shukla et al. (2019), we classify the climate based on aridity index values into humid ($\phi < 1.53$), semi-humid ($1.53 \leq \phi < 2.0$), and arid ($\phi \geq 2.0$). The KAN-derived equation has the best performance under humid climate conditions with KGE of 0.652, RMSE of 0.061, and R^2 of 0.601. This is expected because the majority of the sample belongs to this category ($n = 287$). Under semi-humid and arid climates, RMSE remains comparable (0.036–0.049), but KGE and R^2 decline substantially (KGE from −0.44 to −0.17, and R^2 from 0.04 to 0.12). This performance decline is attributable to limited data size. In this study, semi-humid climate and arid climate only have 26 or 65 data points, respectively. Future studies are suggested to focus on data collection in semi-humid and arid regions to enhance model performance under these climatic conditions.

To address the role of hyperparameters in KAN model performance, we conduct the sensitivity analysis of four evaluated hyperparameters (i.e., number of layers, number of neurons in the hidden layer, grid size, and random seed; Figure 2). Among all evaluated hyperparameters, KANs are most sensitive to the number of layers, since two-layer models show higher median R^2 on the test data set and less variability in model performance compared to three-layer models (Figure 2a). For three-layer models (i.e., models with hidden neurons), model performance

Table 2
Performance Metrics of the Two Models of $Q_B/P = f_B(\phi)$, Where f_B Represents the Functional Relationship Between the Aridity Index and Baseflow Normalized by Precipitation

	Training data set		Test data set		
Metric	Original f_B	KAN-inspired f_B	Original f_B	KAN-inspired f_B	Improvement (%)
NSE	0.387	0.730	0.432	0.712	65.1
KGE	0.622	0.855	0.636	0.849	33.6
RMSE	0.0936	0.062	0.0866	0.062	28.9
R^2	0.740	0.741	0.733	0.726	−0.98

Note. The KAN-inspired f_B equation improves the model performance.

varies significantly with the number of neurons (Figure 2b). However, increasing the number of neurons does not necessarily improve the model's performance. For instance, models with five hidden neurons show lower median model performance than those with four hidden neurons. In contrast, grid size and random seed have a relatively minor impact, as both parameters show similar variability ranges in model performance. This suggests that KANs are least sensitive to grid size and random seed, with the number of layers being the most sensitive parameter, followed by the number of neurons.

Training and hyperparameters tuning of KAN models for Function f_B took 5 hr and 49 min on a single core (Apple M1 Max). This computational cost is primarily attributable to unoptimized algorithms for efficiency and exhaustive hyperparameter search across 1,700 combinations. This computational efficiency may limit its application to scientific discoveries with big data. However, the computational efficiency can be enhanced by starting with a simple model structure and gradually increasing the complexity.

3.2. KAN-Inspired f_B Function

While the KAN-derived f_B Equation 9 significantly outperforms the original model, it also increases the number of fitting parameters from three to four. Evaluation of the available symbolic functions used by the KAN to approximate the learned activation function revealed $\exp(x)$, $\cosh(x)$, $\arctan(x)$, x^4, $\tanh(x)$ to be the top five symbolic functions, arranged in the ascending order with the corresponding R^2 increasing from 0.9986 to 0.9993. The exponent fits the activation function with $R^2 = 0.9986$, indicating that the original functional form of the aridity index (Equation 4) is a reasonable choice. However, the function $\tanh(x)$ fits the activation function with $R^2 = 0.9993$, suggesting that $\tanh(x)$ might be a slightly better choice.

Although $\exp(x)$ and $\tanh(x)$ give comparable performance in function estimation, they have different hydrological implications. The exponent is more suitable for processes that occur monotonically, such as radioactive decay. In contrast, being bounded by -1 and 1 for all x, $\tanh(x)$ is suitable for processes with limits, such as ET constrained by hydrological processes. That explains why $\tanh(x)$ is a better choice for this case. It indicates underlying processes that stabilize water balance partition, exhibiting inherent limits or balancing mechanisms in the hydrological response. An additional physical constraint is a requirement for the normalized baseflow to approach zero when the aridity index approaches infinity. These considerations suggest the aridity index of the form $f_B(\phi) = a - a\tanh(\phi + b)$. Using the training data set to infer the fitting parameters a and b as

$$f_B(\phi) = 0.536 - 0.536\tanh(\phi - 0.283). \tag{10}$$

Table 2 summarizes the model performance metrics for the original aridity index formulation f_B in Equation 5 and the KAN-inspired model (Equation 10). The two models neither overfit nor underfit. The new aridity-index model shows a superior performance on both the training data set and the test data set, except the 0.98% decrease in R^2 for the test data set. The KAN-inspired model is better than the original equation in terms of correlation, bias, and variability. In addition, the KAN-derived equation uses three fitting parameters, while the KAN-inspired model uses only two fitting parameters based on domain knowledge. This demonstrates that the KAN approach can effectively reduce model complexity without sacrificing performance. Both models utilize a functional form identified by KAN, highlighting KAN's ability to discover optimal functional relationships in hydrological

Table 3
Performance Metrics of the Two Models of $Q_D/P = f_D(\phi)$, Where f_D Represents the Functional Relationship Between the Aridity Index and Direct Runoff Normalized by Precipitation

	Training data set		Test data set		
Metric	Original f_B	KAN-derived f_B	Original f_B	KAN-derived f_B	Improvement (%)
NSE	0.530	0.541	0.532	0.539	1.22
KGE	0.604	0.631	0.601	0.625	4.07
RMSE	0.044	0.044	0.044	0.043	0.696
R^2	0.543	0.542	0.536	0.539	0.594

Note. The KAN-derived f_D equation improves the model performance.

processes. For $\phi = 0$, these models predict $Q_B/P \approx 0.62$ and 0.684. Hence, both formulations have a consistent maximum baseflow generation under extremely humid conditions. Thus, KANs identify relevant functional forms to improve model accuracy, leading to a physically meaningful and efficient model.

3.3. KAN-Derived and KAN-Inspired f_D Functions

Based on the definition, $f_B(\phi) + f_D(\phi) = 1$ for $\phi = 0$. The standard method of KAN training cannot incorporate such physical constraints into a learned model. We achieve that by adding artificial points near the boundaries $\phi = 0$ and $\phi \to \infty$ to the training data set: we set $Q_D/P = 1 - 0.684 = 0.316$ for $0 \le \phi \le 0.2$, and $f_B(\phi = 0) = 0.684$ in the KAN-inspired f_B; we set $Q_D/P = 0$ for $\phi \ge 7$. This KAN training strategy returns

$$f_D(\phi) = 0.31 \exp(-0.83\phi). \tag{11}$$

Table 3 summarizes the model performance metrics for the original aridity index formulation f_D in Equation 6 and the KAN-derived model (Equation 11). The KAN-derived model slightly improves the model performance across all evaluated metrics, yet it predicts $Q_D/P = 0.31$ at $\phi = 0$, which is close but not strictly equal to 0.316 (=1 − 0.684). To strictly enforce 0.316 at $\phi = 0$, we also consider the aridity index of the form $f_D(\phi) = 0.316 \exp(a\phi)$, with the only fitting parameter a. Based on training data, the KAN-inspired f_D is

$$f_D(\phi) = 0.316 \exp(-0.875\phi). \tag{12}$$

Table 4 summarizes the model performance metrics for the original aridity index formulation f_D in Equation 6 and the KAN-inspired model (Equation 12). In conjunction with Equation 10, the KAN-inspired models not only satisfy $f_B(\phi) + f_D(\phi) = 1$ at $\phi = 0$, but also further slightly enhance the model performance in terms of KGE, RMSE, and R^2. Figure 3 exhibits the sum, $Q/P = Q_B/P + Q_D/P$, across the full range of observed ϕ values. The error metrics $R^2 = 0.84$ and RMSE = 0.066 demonstrate the model's consistency for all ϕ values.

Table 4
Performance Metrics of the Two Models of $Q_D/P = f_D(\phi)$, Where f_D Represents the Functional Relationship Between the Aridity Index and Direct Runoff Normalized by Precipitation

	Training data set		Test data set		
Metric	Original f_B	KAN-inspired f_B	Original f_B	KAN-inspired f_B	Improvement (%)
NSE	0.530	0.543	0.532	0.535	0.538
KGE	0.604	0.650	0.601	0.643	7.079
RMSE	0.044	0.044	0.044	0.044	0.307
R^2	0.543	0.544	0.536	0.540	0.670

Note. The KAN-inspired f_D equation improves the model performance.

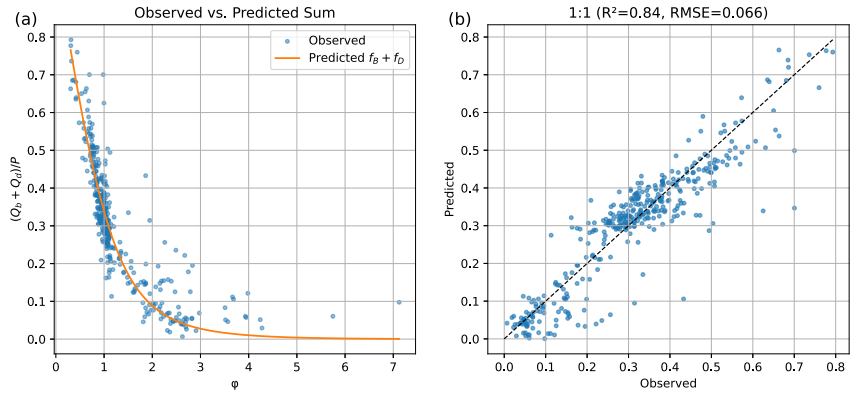

Figure 3. Simulated and observed sum of normalized baseflow and direct runoff, $Q/P = Q_B/P + Q_D/P$. (a) Dependence of Q/P on the aridity index ϕ, at 378 catchment Attributes and Meteorology for Large-sample Studies catchments. The field observations are represented by the blue dots. The predictions obtained via the sum of KAN-inspired relations (Equations 10 and 12) are shown by the orange solid line. (b) One-to-one scatter of the predicted and observed values of Q/P. The black dashed line represents the 1:1 reference line. Graphical representation begins at $\phi \approx 0.05$, instead of 0, due to limited data availability.

3.4. Analysis of Observed Mean-Annual Variables

To further improve the accuracy of estimating runoff components, we evaluate the relationship among mean-annual variables. Figure 4 exhibits the observed mean-annual variables Q_B, Q_D, P, and PET, along with two normalized streamflow components Q_B/P and Q_D/P, all plotted against ϕ for the 378 CAMELS catchments. In accordance with Meira Neto et al. (2020), Q_B/P and Q_D/P have a similar dependence on ϕ (Figures 4a and 4c). Our analysis demonstrates that all plotted variables, with the exception of PET, decrease with ϕ. Across the 378 CAMELS catchments, Q_B and Q_D show the smallest spread around the mean behavior, indicating that ϕ is a more reliable predictor of the variability of Q_B and Q_D than of their normalized counterparts Q_B/P and Q_D/P.

We explore the relationship between this variability and ϕ to uncover potential explanations. The data show that variability exists across the entire range of ϕ, albeit with differing magnitudes. The most significant variation occurs when $\phi \in [0.5, 3]$. This range also corresponds to significant variations in P and PET (Figures 4e and 4f). However, these variations are moderated by ϕ. For instance, in catchments where a ϕ value approaches 1.0, P varies significantly from 800 to 1,500 mm/year, and PET ranges from 500 to 1,500 mm/year. Such substantial variations in both P and PET are encapsulated by a similar ϕ value, contributing to a small spread of both Q_B and Q_D over ϕ. P and PET significantly influence the water balance as source and sink terms, respectively. Since $\phi = \mathrm{PET}/P$, it indicates the degree to which streamflow is partitioned into direct runoff and baseflow. When these streamflow components are further normalized by P, additional variability from P is introduced and cannot be mitigated since PET is not included in the normalization.

Based on these findings, we refine the water-balance Equations 3a and 3b to isolate the streamflow components,

$$Q = F_D(\phi) + F_B(\phi), \qquad F_D(\phi) \equiv Q_D, \qquad F_B(\phi) \equiv Q_B. \tag{13}$$

While f_B and f_D represent streamflow components normalized by P, F_B and F_D represent absolute streamflow components without normalization. The KAN-derived symbolic equations for F_B and F_D are

$$F_B(\phi) = 47.13 + 1932.52 \exp\left(-1.42(-\phi - 0.29)^2\right) \tag{14}$$

and

$$F_D(\phi) = 616.82 - 418.39 \arctan(2.84\phi - 0.87). \tag{15}$$

Tables 5 and 6 collate the performance metrics for the original aridity-index expressions for Q_B/P and Q_D/P and KAN-derived expressions for Q_B and Q_D. (We do not report RMSE because of the differing magnitudes of the

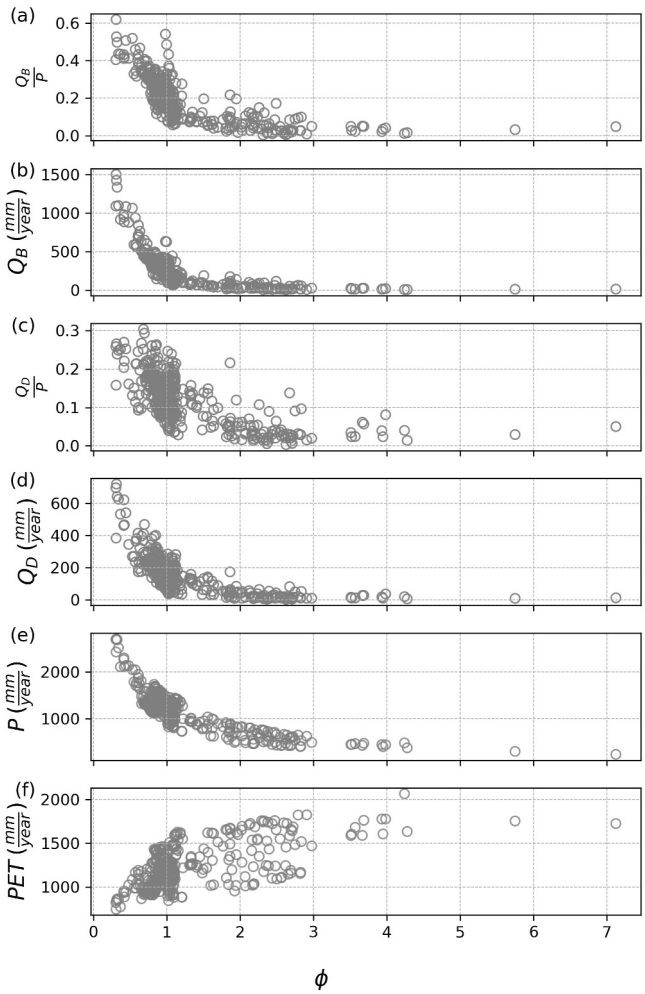

Figure 4. Observed mean-annual (a) baseflow normalized by precipitation (Q_B/P), (b) baseflow (Q_B), (c) direct runoff normalized by precipitation (Q_D/P), (d) direct runoff (Q_D), (e) precipitation (P), and (f) potential evapotranspiration (PET) over the aridity index ϕ = PET/P, for 378 catchment Attributes and Meteorology for Large-sample Studies catchments. All variables except PET show monotonic inverse relationships with the aridity index. Streamflow components exhibit less variation than their normalized terms with the aridity index, implying the potential to enhance the model prediction accuracy.

target variables caused by normalization.) The KAN-derived equations significantly outperform their original counterparts in terms of all metrics. While, the KAN-derived expression for Q_B/P achieves $R^2 = 0.744$ (Table 1), the KAN-derived formula for Q_B has $R^2 = 0.886$. This 16.4% increase in R^2 is attributable to observed Q_B with less uncertainty around its mean behavior across the aridity index, compared to its normalized counterpart (Q_B/P). Since PET and P are important factors in water balance to control the baseflow, the aridity index (PET/P) alone with a small variation can explain the partition of water flow across catchments with a magnitude difference of PET and P. The aridity index already contains the information on precipitation and PET. Normalizing baseflow by P again can introduce additional uncertainty from P, leading to reduced modeling accuracy. These findings underscore the significance of refining the original water-balance Equations 3a and 3b.

While simulating an absolute value (Q_B) may indeed be less challenging than simulating a ratio (Q_B/P), this doesn't necessarily invalidate the comparison. Q_B and Q_B/P have different practical applications in hydrological modeling, and comparing them provides valuable insights into model capabilities across different metrics. Our study demonstrates that KAN models for Q_B can achieve significantly better prediction accuracy than models for Q_B/P. This finding is important as it highlights areas where our modeling approaches excel and where they may need improvement. The superior performance in Q_B prediction suggests that the KAN approach is particularly effective at capturing the absolute quantity of baseflow. We acknowledge the challenges in comparing these different metrics and agree that further work is needed. Specifically, we propose to improve P prediction or incorporate additional physical constraint parameters to enhance Q_B/P predictability. This approach will allow us to leverage the strengths of KAN in absolute value prediction while addressing the complexities of ratio prediction.

Finally, we compare the performance of KANs to that of two tree-based machine learning approaches: XGBoost (Chen & Guestrin, 2016) and random forest (Breiman, 2001). On the test data set, the random-forest model achieves NSE = 0.911, KGE = 0.941, and $R^2 = 0.911$. The corresponding statistics for the XGBoost model is NSE = 0.894, KGE = 0.854, and $R^2 = 0.899$. While the metrics of the tree-based models are similar to those of the KAN-derived relations, these tree-based machine learning approaches exhibit several limitations in hydrological modeling. These algorithms are inherently discrete making them inadequate for capturing continuous relationships among hydrological and climatic variables that are common in hydrology. In contrast, KANs leverage data to learn smooth activation

Table 5
Performance Metrics of the Two Models of $Q_D = F_D(\phi)$

	Training data set		Test data set		
Metric	Original f_B	KAN-derived f_B	Original f_B	KAN-derived f_B	Improvement (%)
NSE	0.530	0.769	0.532	0.768	44.2
KGE	0.604	0.816	0.601	0.845	40.6
R^2	0.543	0.769	0.536	0.769	43.3

Note. Here, F_D represents the functional relationship between the aridity index and direct runoff, and f_D represents the functional relationship between the aridity index and direct runoff normalized by precipitation. The KAN-derived F_D formula improves the model performance.

Table 6
Performance Metrics of the Two Models of $Q_B = F_B(\phi)$

	Training data set		Test data set		
Metric	Original f_B	KAN-derived f_B	Original f_B	KAN-derived f_B	Improvement (%)
NSE	0.387	0.885	0.432	0.886	105
KGE	0.622	0.899	0.636	0.920	44.8
R^2	0.740	0.886	0.733	0.886	20.9

Note. Here, F_B represents the functional relationship between the aridity index and baseflow, and f_B represents the functional relationship between the aridity index and baseflow normalized by precipitation. The KAN-derived F_B formula improves the model performance.

functions, which can be visualized and interpreted to gain insights into process understanding. Additionally, tree-based algorithms are struggling with extrapolation beyond the training data set. For instance, when ϕ approaches zero, Q_B predictions from the tree-based methods remain close to 1,500 mm/year, as this exceeds the bounds of the training data(Figure 4b). This limitation poses a challenge for modeling extreme climate conditions, which are common in hydrology and are expected to become more frequent in the future. In contrast, KAN-derived formulations are well-suited for extrapolation due to their continuous nature. For instance, KAN-derived formulation can predict Q_B over 1,500 mm/year. Furthermore, KANs generate closed-form symbolic models that are simple and can be directly used by other users without requiring any specialized software or computational tools.

However, the tree-based algorithms are highly efficient and optimized for parallel computation. Using six cores, both tree-based models can be trained, and finish their hyperparameter tuning within 2 minutes. In contrast, KANs, currently limited to single-core processing, require approximately 6 hours for training and hyperparameter tuning through exhaustive grid search. However, KANs may be preferred in applications where a more in-depth interpretability is crucial, as in the present study. In addition, the simple and accurate expressions derived from KANs can be directly integrated into ecohydrological or process-based models. In contrast, tree-based models are less amenable to such integration due to their reliance on specific programming environments and external libraries. Future advancements in KAN algorithms compatible with parallel computing can make them more suitable for large-scale applications.

4. Conclusion

We used KAN to derive a symbolic formula for estimating the long-term baseflow across 378 CAMELS catchments. The new equation significantly outperforms the original aridity index formula in terms of all evaluated metrics. On the test data, the NSE increases by 71.3%, the RMSE decreases by 32.4%, and the KGE increases by 25%.

KAN has identified a new "optimal" relation between the long-term baseflow and the aridity index. This new model with two fitting parameters outperforms the original model, which has three fitting parameters.

Observed mean-annual variables across 378 catchments in the United States indicate that the streamflow components Q_B and Q_D have more robust dependencies on the aridity index ϕ than their normalized counterparts Q_B/P and Q_D/P do.

Consequently, we refine the water-balance equations to relate Q_B and Q_D to ϕ. The KAN-derived symbolic relations significantly outperform the original formulations and are comparable to the tree-based machine learning methods, highlighting the advantages of KAN in generating interpretable and accurate models for predicting streamflow components. The simple and accurate formations provide a solid foundation for partitioning direct runoff and baseflow, especially under the changing climate conditions.

Additionally, the symbolic relations identified in this study are transferable to other KAN models, allowing for the exploration of new variables or evaluation of new data by introducing more intervals to approximate complex relationships, or by investigating new activation functions while locking the activation functions identified here. These findings highlight that KAN is a promising approach to enhance hydrological modeling. Future studies are suggested to explore underlying physical reasons for KAN models and KAN's broader generalizability.

While the aridity index adequately explains mean annual streamflow components, model performance declines in semi-humid and arid regions, likely due to data sparsity. To improve generalizability, future work could explore strategies such as data augmentation and transfer learning. In addition, further improvements are possible when considering other variables (e.g., soil depth, soil properties, and topography). For example, since topography is a crucial factor in determining groundwater recharge, its spatial distribution significantly affects the amount of baseflow. Including this parameter may further improve prediction ability. The KAN framework can be extended to include these additional variables by adding new input nodes for topography as well as new edges between these nodes and existing aridity index and baseflow nodes. The relationship (i.e., learned activation function) between ϕ and baseflow identified in this study can be kept. After training the model with new variables, the new relationship among topography, aridity index, and baseflow can be visualized through newly learned activations and used to derive new mathematical expressions. If the considered variable is not well-studied, new mechanisms between these variables may be identified. When the target function is not symbolic (e.g., Bessel function), KAN can still approximate it numerically whereas conventional symbolic regression methods may fail (Z. Liu et al., 2024). KANs are domain-independent and do not rely on hydrology-specific assumptions to demonstrate optimal functional dependencies between runoff components and the aridity index (i.e., the focus of this manuscript). This versatility makes them applicable across various fields, including ecology, climate science, and other areas of scientific discovery.

Multiple studies report improved model performance and computational efficiency by enhancing the training method and architecture of KANs (Ji et al., 2025). One such modified KAN architecture (Polo-Molina et al., 2024) replaces B-splines with cubic Hermite splines to ensure monotonicity, resulting in a more explainable model. Despite these advances, current formulations do not explicitly incorporate physical constraints, such as limiting behavior or boundary conditions. Future work will extend the KAN architecture to embed such physical constraints or boundary conditions directly within the training process, thereby improving physical realism and consistency. Additionally, the computational expense of training in this study is relatively high due to the use of a single core. Algorithmic optimization and parallelization will be essential for enabling large-scale applications.

Appendix A: Kolmogorov-Arnold Networks

An output of a KAN is written as

$$y = \mathrm{KAN}(\mathbf{x}) = \sum_{i_L=1}^{n_L-1} \psi_{L-1,i_L,i_{L-1}} \left(\sum_{i_{L-2}=1}^{n_{L-2}} \cdots \left(\sum_{i_2=1}^{n_2} \psi_{2,i_3,i_2} \left(\sum_{i_1=1}^{n_1} \psi_{1,i_2,i_1} \left(\sum_{i_0=1}^{n_0} \psi_{0,i,i_0}(x_{i_0}) \right) \right) \right) \cdots \right). \tag{A1}$$

The activation function $\psi(x)$ is defined as

$$\psi(x) = w_b \frac{x}{1+\mathrm{e}^{-x}} + w_c \sum_i c_i B_i(x), \tag{A2}$$

where w_b and w_c are weights, $x/(1+\mathrm{e}^{-x})$ is the SiLU basis function, and $\sum_i c_i B_i(x)$ is a linear combination of B-splines with trainable coefficients c_i.

Appendix B: Model Performance Evaluation Metrics

We evaluate the performace of the proposed algirithms in terms of the following metrics. The NSE,

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(S_i - O_i)^2}{\sum_{i=1}^{n}\left(O_i - \bar{O}\right)^2}; \tag{B1}$$

the standard R^2,

$$R^2 = \frac{\left(\sum_{i=1}^{n}(O_i - \bar{O})(S_i - \bar{S})\right)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(S_i - \bar{S})^2}; \tag{B2}$$

the KGE,

$$\mathrm{KGE} = 1 - \sqrt{(R-1)^2 + \left(\frac{\mu_s}{\mu_o} - 1\right)^2 + \left(\frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} - 1\right)^2}; \tag{B3}$$

and the RMSE,

$$\mathrm{RMSE} = \left(\frac{1}{n}\sum_{i=0}^{n}(S_i - O_i)^2\right)^{\frac{1}{2}}. \tag{B4}$$

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

All the data used in this study are either available in the open literature (Meira Neto et al., 2020) or generated with our research code. The Python code used to generate synthetic data, perform data analysis, and produce the figures is available in the Zenodo repository (Liu et al., 2025).

Acknowledgments
This research was funded in part by the Strategic Environmental Research and Development Program (SERDP) of the Department of Defense under award RC22-3278, by the Air Force Office of Scientific Research under Grant FA9550-21-1-0381, by the National Science Foundation under award 2100927, and by the Office of Advanced Scientific Computing Research (ASCR) within the Department of Energy Office of Science under award number DE-SC0023163.

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131. https://doi.org/10.1002/joc.3413

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-guidelines for computing crop water requirements — FAO irrigation and drainage paper 56. *Fao*, *300*, D05109.

Assouline, S., & Tartakovsky, D. M. (2001). Unsaturated hydraulic conductivity function based on a fragmentation process. *Water Resources Research*, *37*(5), 1309–1312. https://doi.org/10.1029/2000wr900332

Braun, J., & Griebel, M. (2009). On a constructive proof of Kolmogorov's superposition theorem. *Constructive Approximation*, *30*(3), 653–675. https://doi.org/10.1007/s00365-009-9054-2

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Budyko, M. I., & Budyko, M. (1974). Climate and life.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM. https://doi.org/10.1145/2939672.2939785

Head, J. D., & Zerner, M. C. (1985). A Broyden-Fletcher-Goldfarb-Shanno optimization procedure for molecular geometries. *Chemical Physics Letters*, *122*(3), 264–270. https://doi.org/10.1016/0009-2614(85)80574-1

Hong, D., Zhang, B., Li, H., Li, Y., Yao, J., Li, C., et al. (2023). Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks. *Remote Sensing of Environment*, *299*, 113856. https://doi.org/10.1016/j.rse.2023.113856

Hong, D., Zhang, B., Li, X., Li, Y., Li, C., Yao, J., et al. (2024). Spectralgpt: Spectral remote sensing foundation model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *46*(8), 5227–5244. https://doi.org/10.1109/tpami.2024.3362475

Ji, T., Hou, Y., & Zhang, D. (2025). A comprehensive survey on Kolmogorov Arnold networks (KAN). Retrieved from https://arxiv.org/abs/2407.11075

Kolmogorov, A. (1956). On the representation of continuous functions of several variables as superpositions of continuous functions of a smaller number of variables. *Doklady Akademii Nauk*, *108*.

Kubalík, J., Derner, E., & Babuška, R. (2023). Toward physically plausible data-driven models: A novel neural network approach to symbolic regression. *IEEE Access*, *11*, 61481–61501. https://doi.org/10.1109/access.2023.3287397

Li, C., Zhang, B., Hong, D., Yao, J., & Chanussot, J. (2023). LRR-net: An interpretable deep unfolding network for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–12. https://doi.org/10.1109/TGRS.2023.3279834

Liu, C., Roy, T., Tartakovsky, D. M., & Dwivedi, D. (2025). Baseflow identification via explainable ai with Kolmogorov-Arnold networks. *Zenodo*. https://doi.org/10.5281/zenodo.17180307

Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., et al. (2024). KAN: Kolmogorov-Arnold networks.

L'vovich, M. I. (1979). World water resources and their future. *American Geophysical Union*.

Lyne, V., & Hollick, M. (1979). Hydrology and water resources symposium. *National Committee on Hydrology and Water Resources of the Institute of Engineering*.

Martius, G., & Lampert, C. H. (2016). Extrapolation and learning equations.

Meira, A. (2020). Meira Neto et al 2020. https://doi.org/10.6084/m9.figshare.11592015.v3

Meira Neto, A. A., Roy, T., de Oliveira, P. T. S., & Troch, P. A. (2020). An aridity index-based formulation of streamflow components. *Water Resources Research*, *56*(9), e2020WR027123. https://doi.org/10.1029/2020WR027123

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Polo-Molina, A., Alfaya, D., & Portela, J. (2024). Monokan: Certified monotonic Kolmogorov-Arnold network. Retrieved from https://arxiv.org/abs/2409.11078

Ponce, V. M., & Hawkins, R. H. (1996). Runoff curve number: Has it reached maturity? *Journal of Hydrologic Engineering*, *1*(1), 11–19. https://doi.org/10.1061/(asce)1084-0699(1996)1:1(11)

Sahoo, S. S., Lampert, C. H., & Martius, G. (2018). Learning equations for extrapolation and control.

Shukla, P. R., Skeg, J., Buendia, E. C., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D., et al. (2019). Climate change and land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.

Zhou, H., & Pan, W. (2022). Bayesian learning to discover mathematical operations in governing equations of dynamic systems.