https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions

Deep, Wide, or Shallow? Artificial Neural Network Topologies for Predicting Intermittent Flows

Farhang Forghanparast, Elma Annette Hernandez, and Venkatesh Uddameri
TTU Water Resources Center, Department of Civil, Environmental and Construction Engineering, Texas Tech University
Correspondence: Elma Annette Hernandez (annette.hernandez@ttu.edu)

Abstract. Intermittent Rivers and Ephemeral Streams (IRES) comprise 60% of all streams in the US and about 50% of the streams worldwide. Furthermore, climate-driven changes are expected to force a shift towards intermittency in currently perennial streams. Most modeling studies have treated intermittent streamflows as a continuum. However, it is better to envision flow data of IRES as a "mixture-type", comprised of both flow and no-flow regimes. It is therefore hypothesized that data-driven models with both classification and regression cells can improve the streamflow forecasting abilities in these streams. Deep and wide Artificial Neural Networks (ANNs) comprising of classification and regression cells were developed here by stacking them in series and parallel configurations. These deep and wide network architectures were compared against the commonly used single hidden layer ANNs (shallow), as a baseline, for modeling IRES flow series under the continuum assumption. New metrics focused on no-flow persistence and transitions between flow and no-flow states were formulated using contingency tables and Markov chain analysis. Nine IRES across the state of Texas, US, were used as a wide range of testbeds with different hydro-climatic characteristics. Model overfitting and the curse-of-dimensionality were reduced using extreme learning machines (ELM), and balancing training data using the synthetic minority oversampling technique (SMOTE), greedy learning and Least Absolute Shrinkage and Selection Operator (LASSO). The addition of classifier cells greatly improved the ability to distinguish between no-flow and flow states, in turn, improving the ability to capture no-flow persistence (dryness) and transitions to and from flow states (dryness initiation and cessation). The wide network topology provided better results when the focus was on capturing low flows and the deep topology did well in capturing extreme flows (zero and >75th percentile).

1 Introduction

Intermittent rivers and ephemeral streams (IRES) are characterized by no-flow conditions for at least some duration over the course of a year (Datry et al., 2017). While these systems have not been studied extensively, there is an increasing recognition that they provide critical environmental and hydrological connectivity within a landscape (Leigh et al., 2016; Boulton et al., 2017; Jaeger et al., 2017; Billi et al., 2018; Hill and Milner, 2018). In addition, they provide valuable ecosystem services such as forage, nesting, and movement corridors for aquatic species and terrestrial wildlife (Richardson, 1990; Steward et al.,

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

2011; Courtwright and May, 2013; Leigh, 2013; Datry et al., 2016; Datry et al., 2018; Karaouzas et al., 2018; Vander Vorste et al., 2020). As IRES make up nearly 60% of all streams in the United States (Levick et al., 2008; Eng et al., 2016) and over 50% of all rivers and streams worldwide (Gutiérrez-Jurado et al., 2019), there is a growing interest in exploiting IRES systems for meeting anthropogenic water supply needs (Lowe et al., 2006; Grey and Sadoff, 2007). As IRES can transport significant amounts of water and materials during flooding, forecasting their flows is of importance in flood control and management applications (Gallart et al., 2012; Datry et al., 2016).

Forecasting flows in intermittent streams is a challenging endeavor. Streamflows in IRES often vary over several orders of magnitude (Sazib et al., 2020). Furthermore, in arid and semi-arid regions, streamflow in IRES exhibits considerable month-to-month variability ranging from very high flows in one month and quickly turning dry the next (Tooth, 2000; Katz et al., 2012). These sharp discontinuities are difficult to capture using water budget methods as they are based on the continuum assumption of flow. Therefore, data-driven modeling is often applied to capture discontinuous streamflow dynamics (Chebaane et al., 1995; Aksoy and Bayazit, 2000).

From a statistical standpoint, intermittent streamflows can be viewed as zero or non-zero inflated datasets based on the extent of intermittency. The presence of a significant number of zeros (or non-zeros) tends to violate the distributional assumptions underlying the data (Tu and Liu, 2016). Therefore, it is often best to separately address zero and non-zero values as they are likely to arise from two different distributions. In the field of hydrology, Van Ogtrop et al. (2011) extended this approach to heterogeneous and nonstationary conditions, wherein the parameters of the distributions (i.e., location, shape, scale) are functions of other covariates. Stochastic approaches predict the probability of flowrates being below a specific threshold (or between two specified values) and are, of great value in risk assessment studies. In many water resources applications, a (point) estimate of the flowrate is often necessary to quantify available water within IRES. To address this need, machine learning methods, particularly artificial neural networks and their variants, have been utilized to model flowrates in intermittent rivers and streams in recent times (Cigizoglu, 2005; Jothiprakash et al., 2009; Kişi, 2009; Kisi et al., 2012; Makwana and Tiwari, 2014; Danandeh Mehr, 2018; Rahmani-Rezaeieh et al., 2020). While these methods have improved the forecasting abilities of IRES systems, modeling the sharp transitions from flow to no-flow regimes continues to be a challenge. As machine learning models presented in the literature often treat both the zero and non-zero flows to be part of the continuum, they tend to generally over-predict low flows and underpredict high flows as these models try to balance the swings between low and high flow regimes while trying to best fit the entire dataset. This phenomenon can also be observed in non-intermittent systems where flowrates exhibit large (several orders of magnitude) variability (Cigizoglu, 2003; Nanda et al., 2019; Wagena et al., 2020).

Given the hydrological and ecological importance of zero flow conditions, the need for accurate prediction of this flow state within an IRES is of considerable importance. Decisions based on over-prediction of zero flows can have deleterious consequences as the prediction would imply the presence of water and flow that are vital to the ecological and hydrological functioning of the stream, which in reality would not be the case (Kaletová et al., 2019). In a similar vein, incorrectly predicting no-flow when there is actually flow (under-prediction) often has economic consequences as it precludes the beneficial use of water that may be available but not accounted for during planning. Improper estimates during the planning phase leads

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

to making alternative provisions for water that would otherwise not have been necessary had the prediction been accurate. Improving the accuracy of intermittent flow predictions and forecasts is therefore critical from both ecological and economic points of view when managing IRES systems.

In this study, intermittent flow data at a gaging station are viewed as a 'mixture-type' data comprising of both categorical (flow, no-flow) and continuous (non-zero flows) data types (Eris et al., 2018; Costa et al., 2020). It is therefore hypothesized that a combined classification-regression machine learning approach is likely to provide better results than models that treat the intermittent flow data as a continuum. To test this hypothesis, two innovative machine learning topologies for combining classification and regression models are proposed in this study. In the first topology, the classification and regression models are stacked sequentially to create a deep learner. In the second topology, the classifier and the regressor are placed in a parallel configuration to create a wide learner. Machine learning models based on these proposed topologies are described in detail and compared to a singular regression-type modeling approach that has been used in previous studies. Model evaluation and comparisons have been carried out at nine intermittent flow stations covering a wide range of hydro-climatic conditions using new tools and metrics to evaluate intermittent flows that are based on contingency tables and Markov Chains. These contributions present state-of-the-art approaches for modeling flows in intermittent streams.

2 Methodology

2.1 Deep topology for modeling intermittent streamflows

The conceptualization of the deep topology is shown in Fig. 1 (A). The algorithm consists of two cells, a classification cell and a regression cell, that are stacked in series to create a deep learner. The inputs are first passed into the classification cell, where a binary classification algorithm is used to predict No-flow (Qo) and non-zero flow (Q+) states. If the prediction of the classification cell is a non-zero flow (Q+) state, then the regression cell is activated. The regression cell comprises a suitable regression-type machine learning algorithm which makes a prediction of the non-zero flow. The concatenation layer receives as inputs the predictions from either the classification and regression cell and provides a final estimate of the flow. In theory, the inputs for the classification cell and regression cells can be completely different. However, as both these cells predict different aspects of the intermittent flow system, they are likely to share at least a subset of inputs. The most parsimonious formulation would result in the inputs being the same for both cells.

2.2 Wide topology for modeling intermittent streamflows

A schematic of the wide-topology model is depicted in Fig. 1 (B). In this approach, the classification and regression cells are placed in parallel. Inputs are simultaneously received by both these cells. Again, the inputs for classification and regression cells can be completely different but will most likely be shared as they are modeling different aspects of intermittent flow and can be the same for maximizing parsimony. The outflow from the classification cell is coded as 0 for no-flow and 1 for flow. The regression cell, in this topology, treats the zero and non-zero flows as a continuum and makes a prediction.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

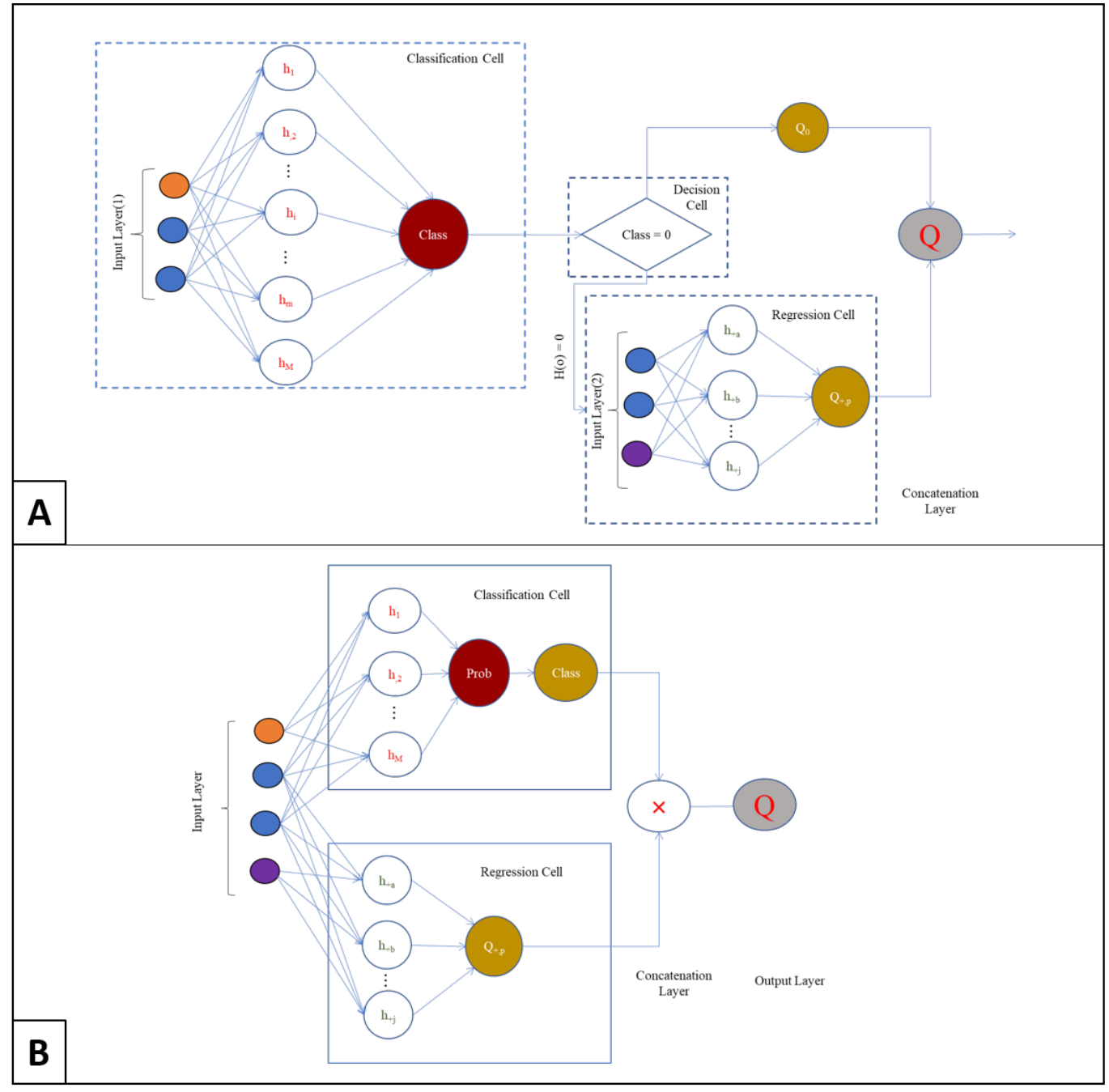


Figure 1. An illustration of the developed architectures for A) the deep model, and B) the wide model for IRES flow prediction.

The concatenation layer comprises a multiplication operator and multiplies the outputs from classification and regression

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

cells, essentially replacing the regression cell prediction to zero when the classification cell predicts no flow and outputs the prediction of the regression cell when the classification cell predicts flow.

The training of the classification cells is carried out using the same dataset in both deep and wide configurations. On the other hand, the algorithms in the regression cell are trained using different datasets. In the deep configuration, only the positive (non-zero) flows within the training dataset are used to train the regression-type model. However, all of the training dataset is used to train the regression-type model in the wide configuration.

Conventionally used models for streamflow forecasting only include the regression cell of the wide network (Cigizoglu, 2005; Kişi, 2009; Makwana and Tiwari, 2014; Rahmani-Rezaeieh et al., 2020). This configuration typically has a single input layer, a hidden layer and an output layer and is referred to as shallow topology (or shallow model) in this study.

2.3 Data imbalance in IRES datasets

Imbalanced datasets (i.e., having an unequal number of zero and non-zero values) are likely to be the norm in IRES datasets rather than an exception. Data imbalance is a major constraint when performing discrete state classification and must be properly dealt with to obtain reliable classifications (Thabtah et al., 2020). Over-sampling of the minority class, under sampling of the majority class and/or random (equal) sampling of both classes are three techniques that have been used to induce balance in classification datasets (Haixiang et al., 2017).

In this study, the Synthetic Minority Oversampling TEchnique (SMOTE) proposed by Chawla et al. (2002) and since extended to higher-dimensional datasets (Blagus and Lusa, 2013) was used to minimize the effects of data imbalance as this approach has been widely used in many applications and known to provide excellent compensation even under high levels of imbalance (Fernández et al., 2018; Raghuwanshi and Shukla, 2020). SMOTE reduces the imbalance by creating synthetic samples for the minority class. These new samples are obtained by combining two similar samples (x and x^r) via weighted addition (see Eq. (1)):

$$s = x + \mu(x - x^r) \tag{1}$$

Where μ is a random weight ($0 \leq \mu \leq 1$) and x^r is randomly chosen from a set of nearest neighbors of x. As SMOTE creates synthetic data, it can only be applied to the training dataset used to obtain the model coefficients and not to the independent testing dataset. Once the SMOTE operation is performed to balance the training dataset, a suitable classification scheme can be adopted.

2.4 Algorithms for classification cell

Artificial Neural Networks (ANN) are known to provide high classification performance over a wide range of datasets and applications (e.g.,Al-Shayea, 2011; Amato et al., 2013; Wang et al., 2017; Bektas et al., 2017).

A schematic of the Multi-Layer Perceptron Artificial Neural Network (MLP-ANN) is shown in the classification cells in Fig. 1 (A) and (B). A conventional architecture includes an input layer, a hidden layer, and an output layer. In the fully

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions
Open Access
EGU

connected feed-forward architecture, all input nodes are connected to all hidden nodes, and all hidden nodes are connected to the output node. Therefore, the outputs of the hidden node serve as the input of the output node. The connection weight, $w_{i,j}$, between any two nodes defines the strength of the relationship between them.

The weighted sum of the inputs are first aggregated and then passed through a (nonlinear) activation function.

$$Z_j = \sum_{i=1}^{M} w_{i,j} I_i + b_j \quad \forall \quad j = 1, ..., H \tag{2}$$

Where Z is the aggregated sum at the jth hidden node, I_i corresponds to the ith input, M is the number of inputs, and H is the total number of hidden nodes within the hidden layer. The bias term (intercept) of the jth node is denoted by, b_j.

The output of the hidden node is obtained by passing the aggregated sum from Eq. (2) through an activation function:

$$h_{j.o} = g(Z_j) \tag{3}$$

In recent years, the non-saturation type activation functions are gaining popularity as they overcome some of the deficiencies associated with saturation activation functions (e.g., tanh), such as the development of vanishing gradients during the parameter estimation process (Ramachandran et al., 2018; Eger et al., 2018). The Rectified Linear Unit (ReLU) and the Leaky Rectified Linear Unit (LRELU) are two widely used non-saturation activation functions. As model results are sensitive to the choice of the activation function, and the ideal activation function for a given application is not known a-priori it must be ascertained empirically.

The output node also performs a weighted aggregation of the input it receives (i.e., the outputs of the hidden node), and in the case of a binary classifier, it passes the summation value through a sigmoid activation function to obtain a value between 0 and 1.

$$Z_o = \sum_{j=1}^{H} w_{j,o} h_{j,o} + b_o \tag{4}$$

The subscript, o, refers to the output node in the above equation. Using Eq. (4), the output (O) can be computed as:

$$O = \frac{1}{1 + e^{-Z_o}} \tag{5}$$

The value of O is a value between 0 and 1. For a dichotomous variable Q_c, which can take values 0 (no flow) or 1 (flow), the output O corresponds to the probability of obtaining a value of 1 (i.e., flow) or P(Q_c = 1). When the value of O is low (typically below 0.5), there is insufficient evidence of a non-zero flow, and as such Q_c is classified as 0 (i.e., No Flow). When the output O has a value greater than 0.5, Q_c is classified as 1 (i.e., flow). Note that when O is interpreted as P(Q_c = 1), Eq. (5) has the same mathematical structure as the logistic regression.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

2.5 Algorithm Selection for the Regression Cell

There is a long history of using artificial neural networks (ANNs) for modeling streamflows, with applications dating back to the early 1990s (e.g., Kang et al., 1993). Based on this large body of literature, the MLP ANN model was again selected as the algorithm for the regression cell. The MLP ANN for a regression problem is very similar to the MLP ANN for classification in that it also has one input, hidden, and output layers with weights defining the strength of the connection between the nodes. The operations at the hidden nodes are also represented by Eq. (2) and Eq. (3). However, as the output of a regression model is continuous and not dichotomous, a linear activation function is used in the output node:

$$Q_p = (\sum_{j=1}^{H_r} w_{j,o} h_{j,o} + b_o) \times 1 \tag{6}$$

Where Q_p is the predicted flowrate within the regression cell, and all other variables have the same meaning as before. Notice that Eq. (6) has the same mathematical form as the ordinary linear regression.

3 Parameter estimation for deep and wide artificial neural network architectures

The presented classification model contains several unknown model coefficients ($w_{i,j}$,b_j, $w_{j,o}$ and b_o for all i = 1,2,...,M; j = 1,...,J) that have to be estimated from a synthetically balanced training dataset. The binary cross-entropy (BCE) loss function is commonly minimized to obtain the unknown model parameters of the classification cell. The minimization of the BCE loss function (Eq. (7)) is similar to the maximization of the log-likelihood function.

$$L_c = -\frac{1}{N}\sum_{r=1}^{N}(1 - Q_{co,r})P(Q_c = 0|I_r) + Q_{co,r}P(Q_c = 1|I_r) \tag{7}$$

Where, $Q_{co,r}$ is the state of the rth record of the observed streamflow (0 or 1) and I_r is the corresponding input record. N is the total number of records in the training dataset.

Using the minimization of squared errors as the basis, the loss function for estimating the parameters of the model within the regression cell can be written as:

$$L_r = \frac{1}{N^+}\sum_{k=1}^{N^+}(Q_{o,j} - Q_{p,j})^2 \tag{8}$$

For the wide network topology, N^+= N, the length of the training dataset, while N^+ denotes the number of records with positive streamflow values in the case of the Deep network topology ($N^+ \leq$ N). The loss function presented in Eq. (8) seeks to minimize the variance and is analogous to the objective function of the least-squares regression.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

The combined loss function (Lcr) can be written by combining Eq. (7) and Eq. (8):

$$L_{cr} = L_c + L_r \tag{9}$$

The parameter estimation problem entails minimization of the combined loss function (Eq. (9)) to obtain a total of (I_c * H_c + 2*H_c + I_r * H_r + 2*H_r + 2) unknowns. Optimization algorithms such as the gradient descent and its variants often have difficulties as the number of parameters to be estimated increases. The estimation of parameters farther from the input node suffer from vanishing (or exploding) gradient problems which lead to weights that are close to zero and leading to models with poor predictive strengths. Therefore, strategies that help mitigate this "curse of dimensionality" become important.

3.1 Strategies for improving the tractability of parameter estimation

3.1.1 Greedy learning

Greedy learning is a widely used strategy in machine learning for training sequential models such as regression trees, random forests, and deep neural networks (Friedman, 2001; Hinton et al., 2006; Bengio et al., 2006; Larochelle et al., 2009; Johnson and Zhang, 2013; Liu et al., 2017; Naghizadeh et al., 2021). In this approach, parameter estimation is not carried out on a global objective function but conducted in a piece-wise manner. This simplification reduces the number of parameters to be estimated and therefore makes the optimization problem mathematically tractable. Despite the lack of a global objective function, greedy learning algorithms are known to produce useful machine learning models that exhibit a high degree of accuracy (Knoblock et al., 2003; Su et al., 2018; Wu et al., 2018; Belilovsky et al., 2019).

Adopting the greedy learning approach here essentially decouples the global objective function (Eq. (9)) into two separate optimization problems whose objective functions are given by Eq. (7) and Eq. (8). In other words, the models in the classification and regression cells are fit separately to estimate the unknowns within each cell. Generally, the increased computation burden of solving two optimization problems is offset by the gains obtained by separating the overall search space of the global objective function. Therefore, the greedy optimization approach was adopted here to solve Eq. (9).

3.1.2 Extreme Learning Machine configuration

An Extreme Learning Machine (ELM) is a special form of MLP wherein the weights for the input-hidden nodes connections and the associated bias terms are randomly assigned, rather than being estimated via optimization. This strategy greatly reduces the complexity of the parameter estimation process as the weights connecting the inputs to hidden nodes and the associated bias terms need not be estimated, and only the weights and bias associated with the output node need to be estimated.

From a conceptual standpoint, as the input-output computations (Eq. (2) and Eq. (3)) are not part of the parameter estimation process, they only need to be performed once. This is tantamount to applying a randomized nonlinear transformation to the original inputs to create a transformed set of variables (i.e., the outputs of the hidden nodes). As the hidden node-output sub-

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

model is a logistic regression formulation in case of a classification problem and linear regression formulation in case of a continuous output, the optimization can be performed with relative ease using analytical approaches.

Despite the random nature of the input-hidden node transformation, ELMs have been shown to have universal approximation capabilities (Huang et al., 2006; Cocco Mariani et al., 2019). From a practical standpoint, they are noted to perform well and provide results that are comparable to other machine learning methods, especially MLPs that have been fitted using nonlinear gradient descent approaches (Zeng et al., 2015; Yaseen et al., 2019; Adnan et al., 2019). ELMs are increasingly being used in hydrology for a wide range of problems (Deo and Şahin, 2015; Atiquzzaman and Kandasamy, 2015; Deo et al., 2016; Mouatadid and Adamowski, 2017; Seo et al., 2018; Afkhamifar and Sarraf, 2020), especially streamflow forecasting (Lima et al., 2016; Rezaie-Balf and Kisi, 2017; Yaseen et al., 2019; Niu et al., 2020).

The use of Greedy learning and ELM configuration greatly reduces the mathematical complexity of the parameter estimation process for the proposed deep and wide topologies for predicting intermittent flow time-series. However, the problem of overfitting (Uddameri, 2007) cannot be ruled out, especially when the hidden layer contains a large number of nodes. Overfitting must be addressed to ensure the proposed deep and wide topologies learn the insights in the training dataset and are able to generalize to other inputs that are presented to the model during the calibration phase.

3.1.3 Regularization for robust estimation for hidden node selection

While the ELM greatly reduces the computational complexity, the randomization of input-hidden node weights implies that the overall model fits are subject to chance. The number of hidden nodes is an important hyper-parameter that critically controls the performance of ANNs, in general, and ELMs, in particular (Huang and Chen, 2007; Rong et al., 2008; Feng et al., 2009; Lan et al., 2010; Zhang et al., 2012; Ding et al., 2014). If the number of hidden nodes is set too low, then the improper specification of hidden node weights due to random selection is hard to correct.

Having a large number of hidden nodes improves the chances of at least some of them having high weights. However, the nodes with the smaller weights tend to learn the noise in the data resulting in poor generalizing capabilities. Reducing overfitting while maintaining a sufficient number of hidden nodes to capture nonlinear input-output relationships using ELM has received a significant amount of attention in recent years (Yu et al., 2014; Shukla et al., ; Feng et al., 2017; Zhou et al., 2018; Duan et al., 2018; Lai et al., 2020).

The second part of the ELM develops a linear least-squares relationship between the output of the hidden nodes and the ultimate output (predictand). When there are a large number of hidden nodes, correlations between them are to be expected. The presence of correlated inputs results in multicollinearity issues when performing ordinary least squares regression (Hamilton, 1992).

Regularization approaches are commonly used to reduce the impacts of correlated inputs and have been used with ELMs to minimize the overfitting problem (Inaba et al., 2018; Zhang et al., 2020). In this approach, an additional term, which is a function of the weights connecting the hidden node and output weights, is added to the loss function (and is referred to as L-norm). The revised objective function (see Eq. (10)) not only minimizes the sum of squares of residuals but also the number of hidden nodes. The L2-norm, also referred to as Ridge norm or Tikhanov regularization, is a function of squares of the

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

weights (see Eq. (11)). This approach typically forces weights with small singular values to be small numbers (as close to zero as possible), which can be ignored during predictions.

The L1-norm, also referred to as LASSO norm (Eq. (12)), is a special case of the sparsity norm, which minimizes the absolute value of the weights and actually sets the insignificant weights to a value of zero. The loss function with L1-norm results in a convex optimization problem that can be solved via linear programming and, therefore, commonly adopted (Zhang and Xu, 2016). Furthermore, the L1-norm is shown to induce a greater degree of sparseness than the L2-norm without sacrificing prediction accuracy (Fakhr et al.,). The L1-norm is also more robust to outliers than the L2-norm (Zhang and Luo, 2015). Outliers are of particular concern when dealing with highly variable intermittent flow. The λ value in Equation 10 is a weighting factor that denotes the relative importance of the regularization term vis-a-vis the error minimization term and can be obtained via cross-validation procedure (Martínez-Martínez et al., 2011).

$$L_{cr} = L_c + L_r + \lambda \sum_{j=1}^{H} g(w_{ho,j}) \tag{10}$$

$$For \quad L2 - Norm : \sum_{j=1}^{H} g(w_{ho,j}) = \sum_{j=1}^{H} w_{ho,j}^2 \tag{11}$$

$$For \quad L1 - Norm : \sum_{j=1}^{H} g(w_{ho,j}) = \sum_{j=1}^{H} |w_{ho,j}| \tag{12}$$

Custom scripts were developed in R (R Core Team, 2019) to implement the deep and wide network models presented above. In particular, the L1-logistic regression for binary flow and no-flow categorization and L1-linear regression necessary for solving the ELM models within classification and regression cells were implemented using procedures outlined by Hastie et al. (2021). The "glmnet" package (Friedman et al., 2010) and the "caret" package (Kuhn, 2008) were used in the model development phase, the "DMwR" package (Torgo, 2010) was used to apply SMOTE, the "FSelectorRcpp "package (Zawadzki and Kosinski, 2020) was used to assess the variable importance via the information gain algorithm in the pre-processing phase, and the "pROC" package (Robin et al., 2011) and the "hydroGOF" package (Zambrano-Bigiarini, 2017) were utilized to assess the discrete and continuous performance of the models, respectively. The developed R scripts are presented in Supplementary Information.

4 Metrics for evaluation of intermittent flow predictions

It is important to keep the purpose of the model in mind when evaluating its performance (Oreskes, 1998). The primary goal of this modeling effort is to improve the prediction capabilities of zero flow as it denotes the flow intermittency in a stream. Current evaluation metrics largely focus on assessment of continuous flows and as such are not readily suitable for appraising intermittent flows. Therefore, a new set of metrics were developed in this study.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

Several questions related to flow intermittency that are of interest in practical applications form the basis for model evaluation scheme presented here. They are based on contingency table presented in Table A1 and a first-order Markov Chain Transition Matrix presented in Table A2.

4.1 How well does the model predict zero-flow states?

Information from contingency tables can be used to estimate how well the model predicts the zero flow rates as:

$$\text{Accuracy of Zero Flow Predictions (AZF): } \frac{N_{o,o}}{N_{o,o} + N_{o,+}} \tag{13}$$

A perfect score for AZF is unity. A value of AZF < 1 indicates the model underpredicts the zero flow occurrences. The AZF metric is not defined for perennial streams as the denominator will contain a zero value in that case.

There are two types of errors associated with flow and no-flow state prediction which can also provide valuable insights related to the model's utility. The over-prediction error (OPE) occurs when the model predicts that there is flow when there is no flow observed. This error is particularly problematic in applications where the existence of a flow is beneficial. For example, when the water in the intermittent stream is needed for sustaining instream aquatic habitats or when the water can be used for other applications (e.g., ranching or agriculture). OPE can be computed from the contingency table as:

$$OPE = \frac{N_{o,+}}{N_{o,o} + N_{o,+}} \tag{14}$$

An under-prediction error (UPE) occurs when the model predicts a no-flow when there is actually an observed flow. From a water availability standpoint, this assumption is conservative. However, when the magnitude of the observed flow is large this under-prediction can actually be problematic as it does not forecast a non-zero flow, and as such limits the ability to react to potential flooding or water logging conditions in riparian areas.

$$UPE = \frac{N_{+,o}}{N_{+,+} + N_{+,o}} \tag{15}$$

The contingency table can also be used to construct other receiver operating characteristic (ROC) based metrics. In particular, the area under the curve (AUC) of the ROC is an important metric that evaluates whether a classifier's predictions are better than random guessing (AUC > 0.5) and can be used for initial screening of models.

4.2 How well does the model predict no-flow persistence?

In most intermittent streams, once zero-flow conditions are set, they often tend to persist for some period of time. While the AZF metric measures paired occurrences of zero flows in observed and predicted data, it does not measure the persistence of the zero flow state. No flow persistence restricts transport of mass and energy across habitats, alters habitat and biodiversity

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

makeup, and affects the trophic structure and carrying capacity of the stream. Therefore, this no-flow persistence is an important ecological indicator (Rolls et al., 2012).

No-flow persistence (NFP) can be defined as the typical duration the stream spends in the no-flow state. However, NFP (i.e., duration of zero flow events) is a random variable. For a confirmatory analysis, the null hypothesis that both observed and predicted no-flow durations arise from the same distribution can be tested against the alternative hypothesis that they are from different distributions using for example, the two sample Kolmogorov-Smirnov (KS) test.

For a more qualitative comparison the expected value of no-flow persistence for observed and predicted data can be separately obtained and compared. The qualitative comparison is useful to comparing the performance of multiple modeling schemes. The expected value can be computed from a theoretical or empirical CDF as:

$$NFP = E(Q_o) = \int_0^{\infty} (1 - F(Q_o))dQ_o - \int_{-\infty}^{0} (F(Q_o))dQ_o \tag{16}$$

As zero-flow durations are of finite length, the second integral of Equation 16 is equal to zero and only presented here for the sake of completeness. Flow persistence (FP) can be calculated similarly for non-zero flow durations.

4.3 How well does the model predict transitions to and from zero-flow states?

Accurate predictions of no flow to flow transitions and vice-versa are critical for evaluating aquatic system connectivity, aquatic habitat alterations and onset and cessation of water mediated biogeochemical reactions (Larned et al., 2010). The "first passage time" denotes the time that the system takes to transition from one state (e.g., no-flow state) to another (e.g., flow state). While the NFP measures the average time of the no-flow duration, the "mean first passage time" measures the average time for a flow to no-flow transition.

The "mean first passage time" is used to define two metrics to evaluate flow transitions. The no-flow to flow transition (NF2FT) is equal to the "mean first passage time" from no-flow to flow state. Similarly, the F2NFT is equal to the mean first passage time from flow to a no-flow state. Comparisons of model predictions of NF2FT and F2NFT with those computed using observed data are useful to evaluate the ability of the model to correctly capture flow transitions.

4.4 How Well Does the Model Predict Non-Zero Flowrates?

While the focus of this study is largely on zero flow states, the developed models also predict non-zero flow values. As such, the overall accuracy of the model hinges on how well the models predict flow values as well. In addition to their ecological relevance, the prediction of non-zero flow values are useful to assess water availability in the stream as well as the likelihood of any catastrophic hazards such as floods. Metrics such as the root mean absolute error (MAE), root mean squared error (RMSE), Pearson's R and Spearman's rho (ρ) correlation coefficients can be made use of on this portion of the dataset (Danandeh Mehr, 2018; Barber et al., 2020). These equations are well known and as such not presented here for brevity. The use of efficiency metrics (e.g., Nash-Sutclife Efficiency) is generally not suitable for low (and no-flow) situations as they are conditioned by

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

high flow values (Pushpalatha et al., 2012) and also assume the predictions are drawn from a single distribution which may not be the case in IRES. The large variability of flows noted in IRES also likely point to the dependence of variance on flow values and states.

5 Input specification for deep and wide ANNs for predicting intermittent streamflows

The selection of appropriate inputs should be guided by the physics of the system to the maximum extent possible, but statistical and information-theoretic methods are useful to select among competing input variables (Ghaseminejad and Uddameri, 2020).

From a physical standpoint, the streamflow is a function of rainfall, antecedent moisture content as well as other losses including infiltration, evapotranspiration, and water withdrawals. Soil and precipitation characteristics are reported to have great influence on streamflow generation in IRES (Kampf et al., 2018; Azarnivand et al., 2020; Gutiérrez-Jurado et al., 2019). Many hydrological processes exhibit persistence and, therefore, lagged values of these hydrological fluxes could also be useful predictors of flows.

The number of lags to include depends upon the temporal resolution of the time-series forecasting problem. Monthly scale forecasts are generally adopted in many water planning and management endeavors, especially those focused on climate change impacts. Therefore, this time-scale was chosen here to illustrate the proposed approach. Climate-driven processes are the most important factors in determining the hydrology (e.g., occurrence, frequency, and persistence) of IRES flows (Goodrich et al., 2018; Borg Galea et al., 2019). Precipitation, Potential evapotranspiration (PET) were chosen to model input and demand fluxes in the watershed and capture a measure of wetness (Azarnivand et al., 2020; Min et al., 2020). The antecedent soil moisture is an important driver of runoff (Zimmer and McGlynn, 2017), but observed values of soil moisture are generally not available, therefore the soil moisture index (SMI), which represents the interaction of temperature and rainfall, was chosen as the surrogate variable following Augustin et al. (2008).

Baseflow is largely negligible in intermittent streams and the observed runoff is largely a function of quickflow component of the streamflow (Jothiprakash et al., 2009). In the drier arid and semi-arid regions, most of the precipitation that falls within a watershed converts into runoff immediately (Sazib et al., 2020). Coarse-scale runoff estimates generated from regional Variable Infiltration Capacity (VIC) models using 21 different GCM model forcings (see Table S1 in Supplementary Information for a list of models), were obtained (Cherkauer et al., 2003; Brekke et al., 2013) and used as an input to condition model predictions (i.e., inform the model of the best initial guess of the likely streamflow). From a Bayesian perspective, the runoff from the regional-scale VIC model serves as the prior information of the runoff which is then refined (downscaled) by point-scale hydrological fluxes (P, PET) and antecedent soil moisture conditions (SMI). Given the monthly time-step, the first lag of all the variables were also considered to account for conditions that may have initiated in the latter parts of the month and persisted into a significant portion of the next month.

Lagged values of observed streamflows have also been used as inputs to explicitly capture streamflow persistence in many streamflow forecasting studies (Badrzadeh et al., 2017; Hadi and Tombul, 2018; Danandeh Mehr, 2018; Rahmani-Rezaeieh et al., 2020). However, the use of lagged streamflows (which is an output of the model) is not suitable when multi-step ahead

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

forecasts are needed as forecasts often deteriorate at large times (Uddameri et al., 2019; Li et al., 2020). Furthermore, as the intermittent outflows contain both zero and non-zero flows, persistence of flow regimes tends to be complex in intermittent rivers and ephemeral streams. Therefore, lagged values of streamflows were not used in this study. Limiting the model inputs to other hydroclimatic variables allows the models developed here to be used in long-term climate studies, as projections of various hydrological fluxes (inputs) are available from downscaling of global climate model inputs.

6 Model evaluation testbeds

The performance of the proposed models was evaluated at 9 intermittent streamflow sites (gaging stations) in Texas. Table 1 presents salient characteristics of the selected stations and the location of these sites are shown in Fig. 2. The selected sites exhibited a wide range of flow intermittency and are also scattered across different land use (Fig. 2(a)), soil types (Fig. 2(b)), and climatic conditions (Fig. 2(c) and Fig. 2(d)) providing a comprehensive evaluation of the proposed approach over a wide range of hydro-climatic conditions.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

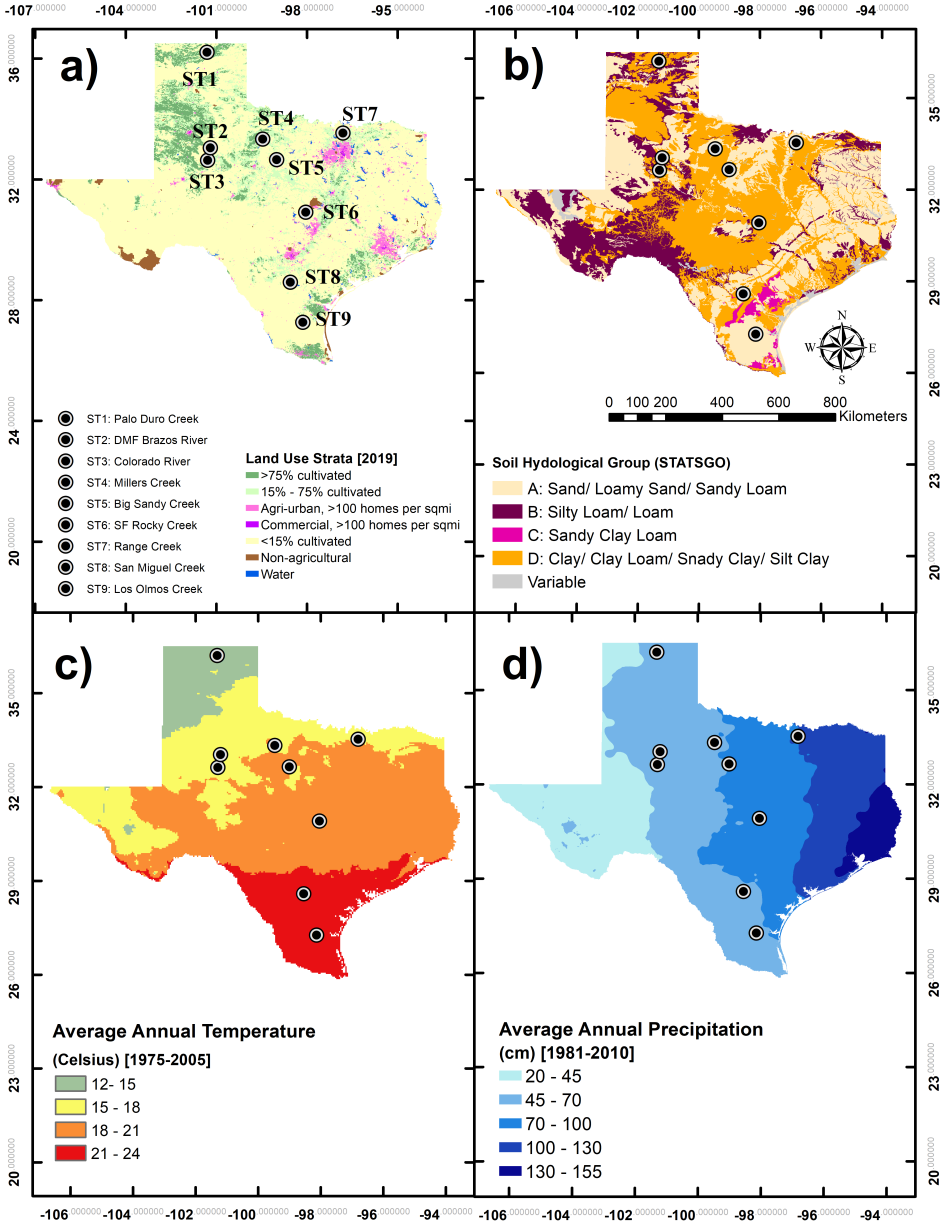


Figure 2. The location of the nine IRES of study in Texas over maps of A) Land Use Strata, B) Soil Hydrological Group, C) Average Annual Temperature, and D) Average Annual Precipitation.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions

Table 1. Summary of information on the nine streamflow monitoring stations.

Station ID	USGS ID	Stream	Location	Intermittency Ratio	Interquartile Range (m^3/s)	Maximum Recorded Flowrate (m^3/s)	Range of Records
ST1	7233500	Palo Duro Creek	Near Spearman, Hansford County, TX	65%	0.03	5.584	1999/7 – 2020/6
ST2	8079600	Double Mountain Fork Brazos River	At Justiceberg, Garza County, TX	12%	0.627	24.797	1961/12 – 2020/6
ST3	8117995	Colorado River	Near Gail, Borden County, TX	33%	0.215	20.102	1988/3 – 2020/5
ST4	8082700	Millers Creek	Near Munday, TX	41%	0.032	12.283	1963/8 – 2020/6
ST5	8086290	Big Sandy Creek	Above Breckenridge, TX	17%	0.414	35.42	1962/2 – 2020/11
ST6	8103900	South Fork Rocky	Near Briggs, Burnet County, TX	24%	0.297	5.352	1963/4 – 2020/9
ST7	8050840	Range Creek	Near Collinsville, TX	32%	0.471	9.299	1992/10 – 2020/9
ST8	8206700	San Miguel Creek	Near Tilden, TX	32%	0.343	51.76	1992/10 – 2020/11
ST9	8212400	Los Olmos Creek	Near Falfurrias, Brooks County, TX	78%	0	3.879	1999/4 – 2020/11

7 Results and discussion

7.1 Model calibrations and testing

The time-series predictions of models over training and testing periods are shown in Figs. 3- 5. In addition to dichotomous flow states, IRES also exhibit very high flow variability often ranging over two orders of magnitude or more. Capturing such large variability is a common problem while forecasting streamflows. The log transformation with a unit shift parameter, i.e., log(Q+1), is commonly used to shrink the range of variability while preserving zero flows (Curran-Everett, 2018; Gao and Martos, 2019). This strategy was also adopted here to reduce the range of streamflows prior to their calibration. Generally speaking, the testing period exhibited greater variability comprising of both sequences of no flows but also higher than normal rainfalls, especially in some parts of the state (see stations ST2, ST3, and ST9). As can be seen from Figs. 3- 5, all models were able to capture the general trends in intermittent flowrates. However, the continuous model did not predict absolute zero values and also yielded physically unrealistic negative flows at 4 stations (see Fig. A1).

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

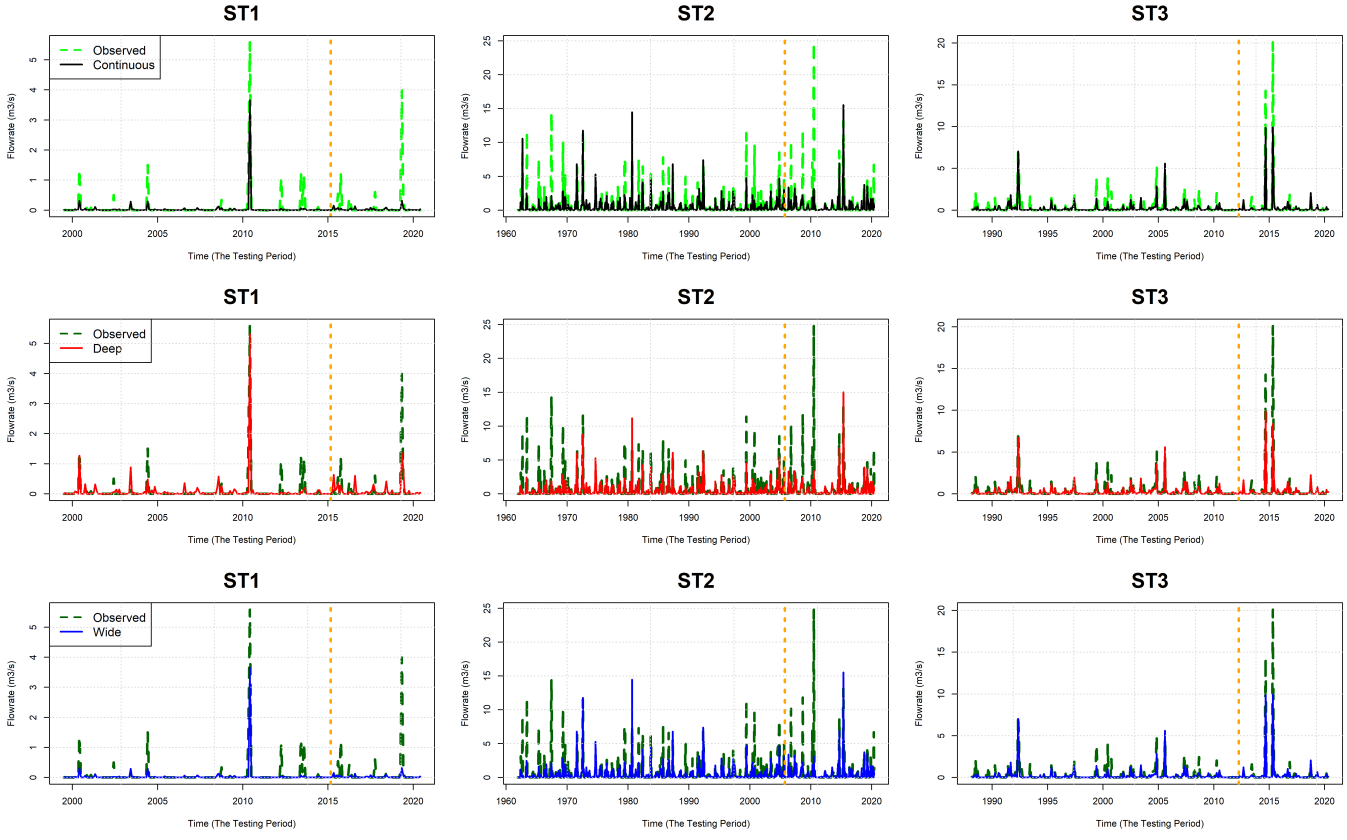


Figure 3. The observed vs. the predicted streamflow time-series of the shallow, deep, and wide models for ST1, ST2, and ST3.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

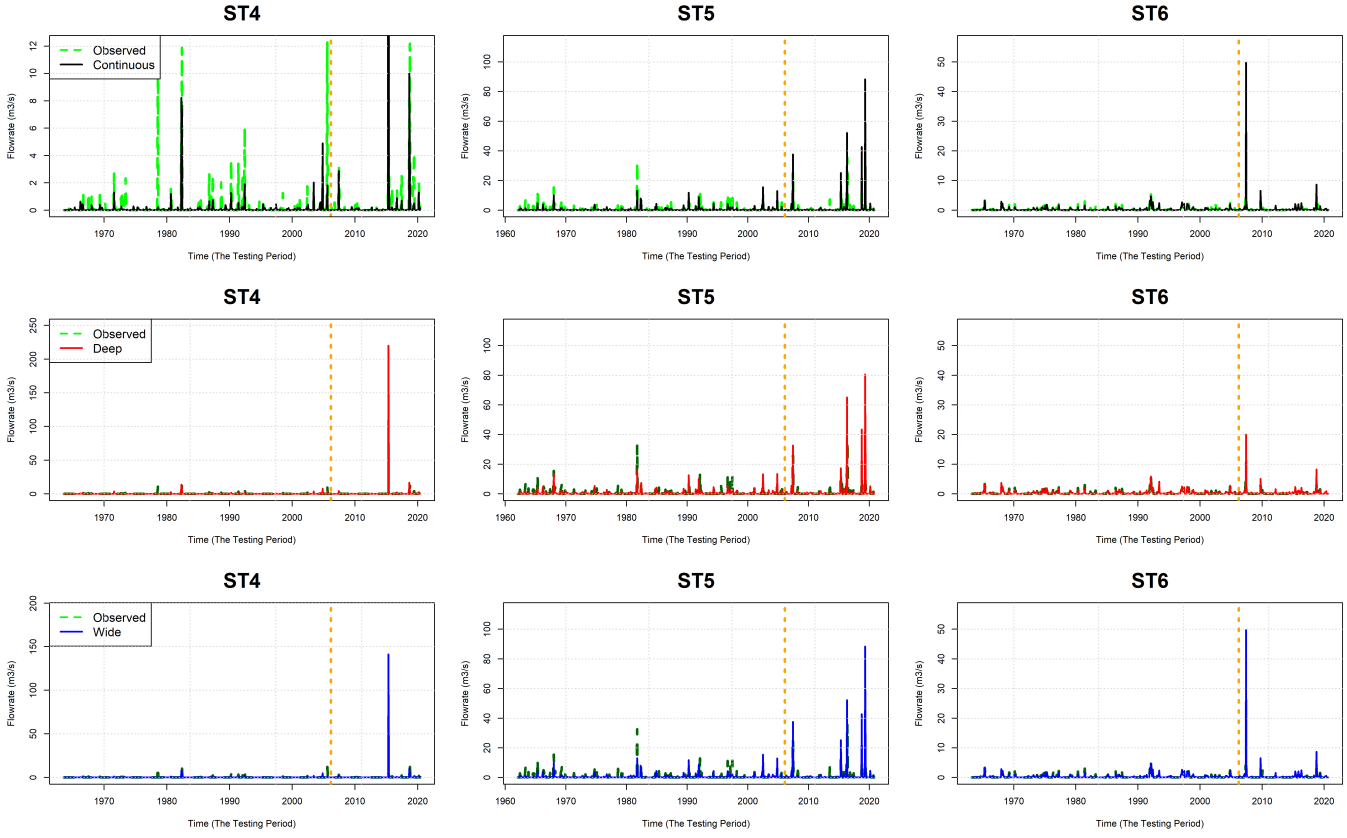


Figure 4. The observed vs. the predicted streamflow time-series of the shallow, deep, and wide models for ST4, ST5, and ST6.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

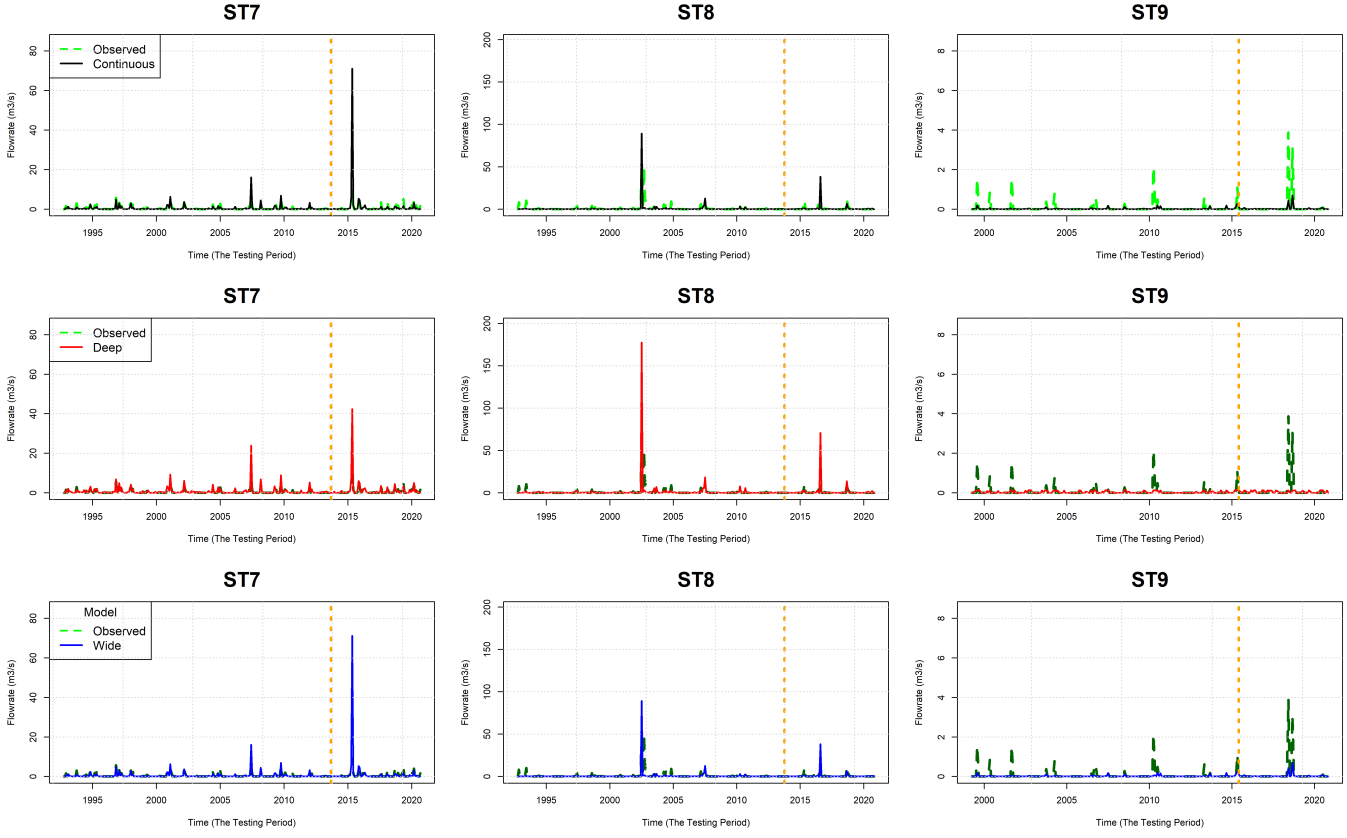


Figure 5. The observed vs. the predicted streamflow time-series of the shallow, deep, and wide models for ST7, ST8, and ST9.

The predictive performance of the shallow, deep, and wide models was compared across the nine IRES of study over the entire range of flowrates (Table 2). In general, the shallow and deep model had lower Mean Absolute Errors (MAE), while the deep model has smaller root mean square error (RMSE). As Errors in estimating higher flow rates lead to greater RMSE errors, the deep model did better in capturing these extreme high end flows. This result is to be expected because, the regression cell of the deep model is trained only with non-zero flow data while the shallow and wide models are trained using both zero and non-zero flows, which also explains their relative ability to capture lower flows well. Therefore, misclassification of the zero flow state in the classification cell leads to over-estimation of flows by the deep model. A separate comparison of misclassification flows indicated that the over-estimation was no more than 0.15 m3/s (typically < 0.05 m3/s) in comparison to predictions by shallow and wide models.

The correlation coefficients (R and ρ) were generally similar for all models at most stations, a few exceptions were noted when the testing dataset was influenced by flow abnormalities (e.g., ST6, ST7, ST9). The Spearman rank correlation coefficients were generally higher at stations were the parametric Pearson product moment correlation coefficient, indicating the models did a better job capturing monotonic trends at these stations. Comparing goodness-of-fit metrics in Table 2, indicated that the

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 2. Summary of continuous performance evaluation metrics (MAE, RMSE, Pearson's r, and Spearman's Rho) for the shallow, deep, and wide models during testing with and without applying log transform for the entire flow range. Abbreviations: MAE, mean absolute error; RMSE, root mean squared error.

Station ID	Model	Log(Q+1) Transform				No Transform			
		MAE	RMSE	Pearson's r	Spearman's Rho	MAE	RMSE	Pearson's r	Spearman's Rho
ST1	Shallow	0.121	0.500	0.780	0.660	0.266	0.556	0.320	0.400
ST1	Deep	0.132	0.384	0.760	0.700	0.298	0.619	0.560	0.510
ST1	Wide	0.120	0.500	0.780	0.720	0.225	0.550	0.330	0.510
ST2	Shallow	0.549	2.068	0.640	0.860	0.766	2.005	0.670	0.720
ST2	Deep	0.556	2.034	0.650	0.820	0.720	2.044	0.680	0.800
ST2	Wide	0.556	2.073	0.640	0.830	0.717	1.998	0.670	0.760
ST3	Shallow	0.290	1.160	0.970	0.730	0.496	1.920	0.780	0.620
ST3	Deep	0.340	1.317	0.940	0.690	0.546	1.918	0.740	0.660
ST3	Wide	0.285	1.160	0.970	0.700	0.477	1.919	0.790	0.640
ST4	Shallow	1.012	10.454	0.350	0.720	0.381	0.926	0.740	0.370
ST4	Deep	1.503	16.478	0.350	0.740	0.302	0.807	0.780	0.740
ST4	Wide	1.011	10.454	0.350	0.750	0.309	0.913	0.740	0.420
ST5	Shallow	1.536	7.514	0.550	0.710	0.894	3.041	0.650	0.650
ST5	Deep	1.499	7.420	0.550	0.670	0.938	3.061	0.660	0.670
ST5	Wide	1.537	7.516	0.550	0.670	0.868	3.046	0.650	0.640
ST6	Shallow	0.526	3.528	0.590	0.600	0.261	0.420	0.790	0.670
ST6	Deep	0.357	1.336	0.700	0.570	0.281	0.433	0.780	0.610
ST6	Wide	0.522	3.528	0.590	0.570	0.248	0.418	0.790	0.620
ST7	Shallow	1.144	6.795	0.680	0.890	0.325	0.562	0.950	0.890
ST7	Deep	0.750	3.673	0.790	0.840	0.317	0.515	0.950	0.840
ST7	Wide	1.145	6.795	0.680	0.840	0.313	0.563	0.950	0.840
ST8	Shallow	0.767	3.304	0.620	0.670	1.398	3.429	0.790	0.250
ST8	Deep	1.182	6.625	0.630	0.630	1.201	2.910	0.790	0.490
ST8	Wide	0.764	3.305	0.620	0.620	1.258	3.416	0.800	0.240
ST9	Shallow	0.100	0.520	0.930	0.520	0.146	0.559	0.790	0.500
ST9	Deep	0.144	0.595	0.260	0.530	0.210	0.588	0.230	0.550
ST9	Wide	0.098	0.520	0.930	0.570	0.121	0.558	0.760	0.570

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

log(Q+1) transform was useful to better capture observed flows at stations dominated by low flows. However, as most stations exhibited high flow variability, the sensitivity the transformation to high flows could not be fully ruled out (Krause et al., 2005). The use of raw (non-transformed) flowrates helped better capture these high flows as even small prediction errors in the log-transformed data at high flowrates gets magnified when the predictions are transformed back to the raw (original) scale.

Overall, the performance of all three models were comparable to the reported streamflow forecasting models in the literature (Adnan et al., 2017; Danandeh Mehr, 2018; Adnan et al., 2019; Yaseen et al., 2019; Cheng et al., 2020; Kisi et al., 2021) when viewed over the entire flow regime. As the focus of the study is on improving zero flow predictions, additional tests were carried out to evaluate this aspect further.

7.2 Deep and wide topologies for predicting zero flow states

If the same set of inputs are used in wide and deep formulation, then the classification cell essentially has the same architecture in both these topologies. Therefore, the comparison here is limited to a single classification cell. As seen from Fig. 6, the Area under the curve (AUC) of the receiver operator characteristics (ROC) curve was greater than 0.5 for all models indicating the classifiers yielded better results than random guessing. The application of SMOTE to balance the training dataset generally improved the prediction accuracy of testing data (which was not subject to SMOTE). At sites where intermittency ratios (#zeros/total # of records) were less than 50%, SMOTE was instrumental in reducing OPE and thus improving zero flow predictions by over 50% on average overall sites. The reduction in OPE was generally accompanied by a small increase in UPE ($\sim 15\%$ overall). When the intermittency ratio is high, the non-zero flows represent the minority class that SMOTE tries to balance. Here, the application of SMOTE caused a reduction in the accuracy of predicting zero flows but improved the accuracy of predicting non-zero flows. The overall accuracy of predictions as well as the area under the curve (AUC) the Receiver Operating Characteristics (ROC) curve generally improved with the application of SMOTE except under very high intermittency conditions (e.g., ST9 which has an intermittency of 78%). Thus, SMOTE is a useful data balancing method to improve the predictions of zero flow states in streams with low intermittency and to enhance the accuracy of non-zero flows in streams with high intermittency.

7.2.1 Comparison of zero flow predictions with shallow model

The shallow model, treating the flow time-series of IRES as a continuum, failed to predict any absolute zero flowrates. While a pattern of over-prediction of low flows were observed in the results of the shallow model, in 45% of cases, physically unrealistic, negative flowrates were also predicted by shallow networks (Fig. A1). The addition of a classification cell effectively addressed this issue as the deep and wide models were capable of predicting absolute zero flowrates with no negative predictions.

7.3 Comparison of no-flow state persistence and transitions

As the continuous model did not provide absolute zero flow predictions, a nominal flow of 0.0003 m3/s (0.01 cfs - the instrument detection limit) was used as the cut-off to delineate no-flow and flow states. All the predicted flowrates below the

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

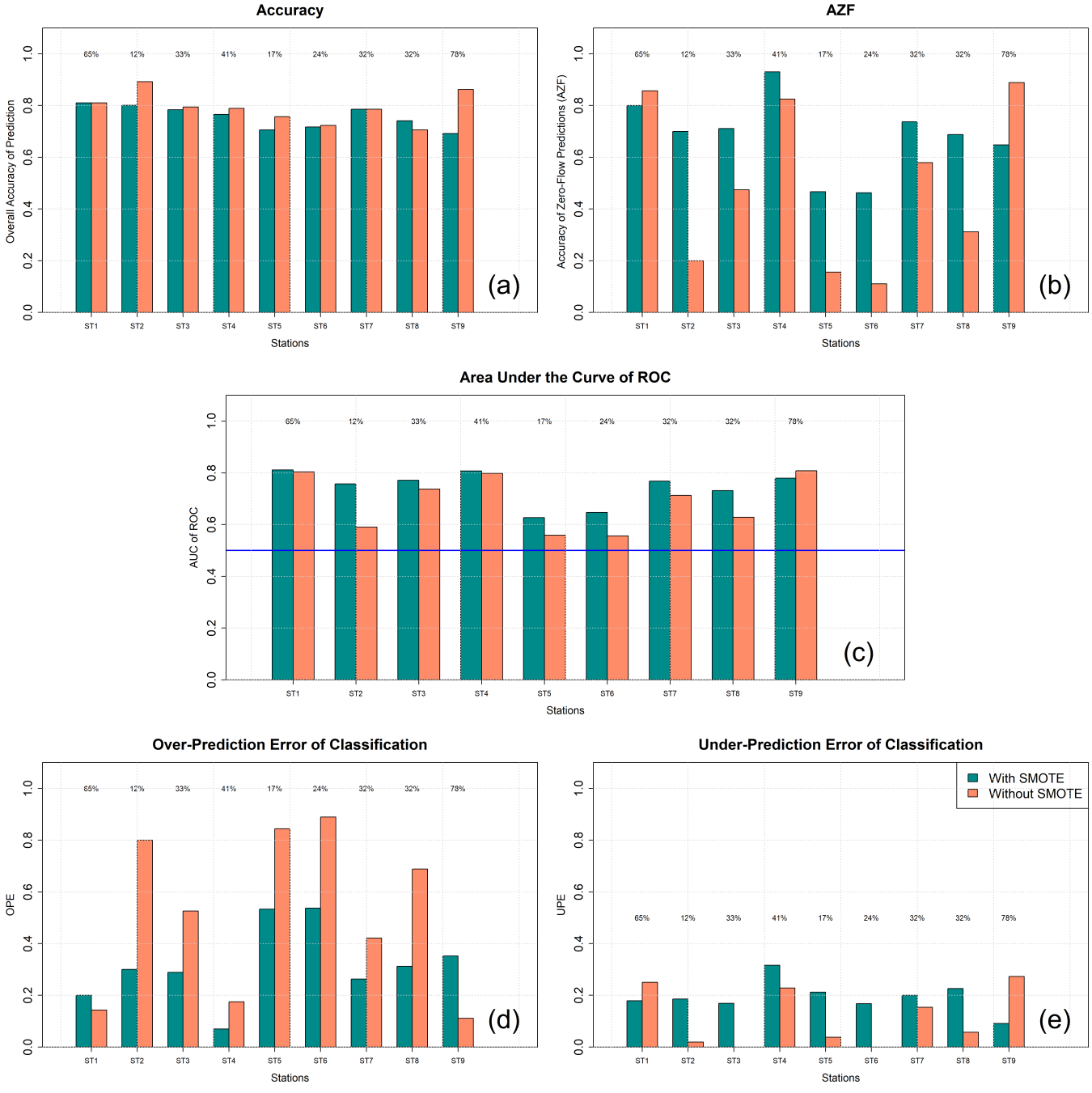


Figure 6. Predictive performance of the classification model during Testing with respect to ACC, AZF, OPE, UPE, and Area Under the Curve of ROC with and without applying SMOTE.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions

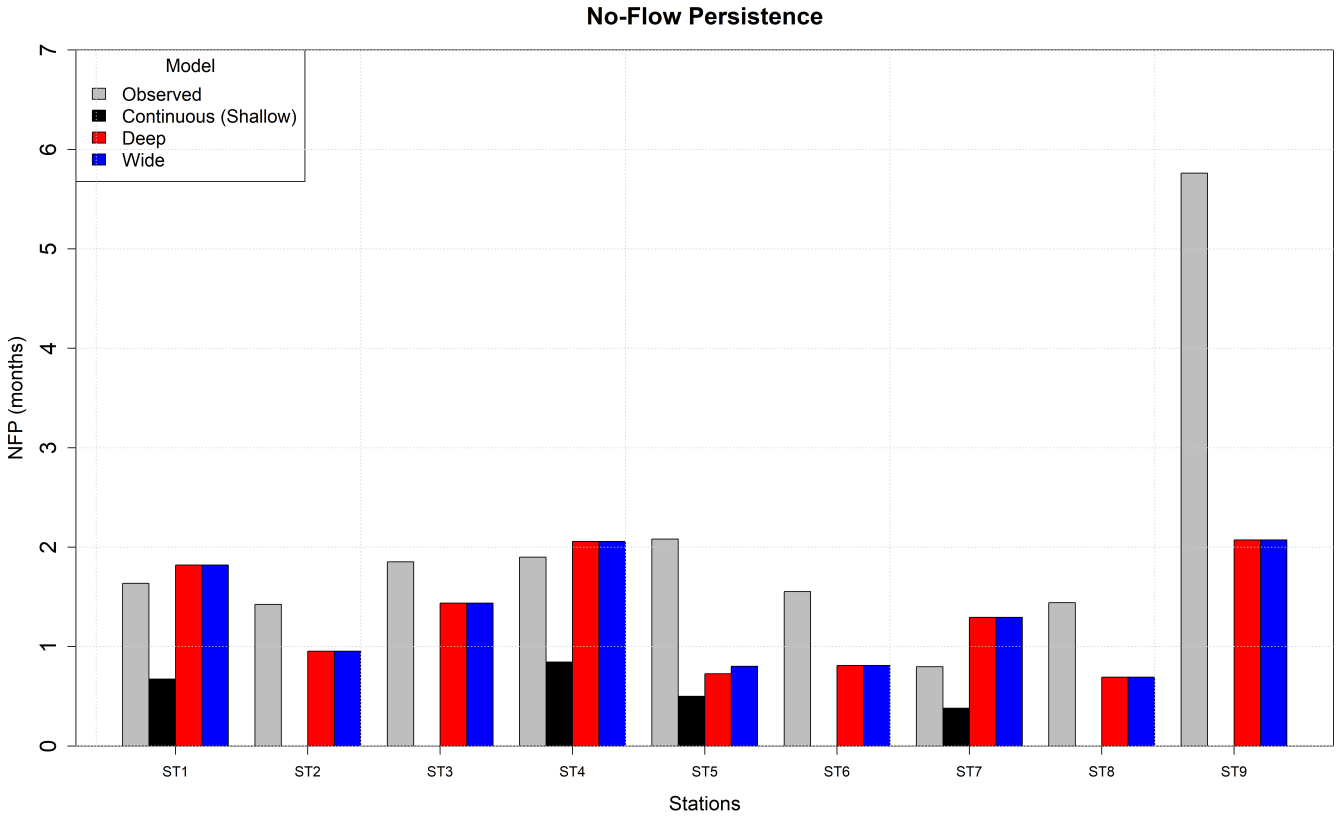


Figure 7. Comparison of observed No-Flow Persistence (NFP) and the predictions of the shallow, deep and wide models for the nine IRES of study.

threshold were assumed to be no-flow events predicted by each model. Then, NFP, FP, NF2FT and F2NFT metrics were computed to compare the performance of the shallow, deep, and wide models in capturing the persistence of no-flow state and transitions between the wet and dry states (Fig. 7-10).

The results presented in Fig. 7 highlight the superiority of the deep and wide topologies in capturing no-flow persistence especially when the intermittency ratio was less than 65%. As the deep and shallow models utilize the same classifier, they provide similar results. The extent of over-prediction of non-zero flows (Q >0.0003 m3/s) was high for the shallow models that in the 55% of cases no zero-flow states were predicted despite applying the cut-off.

NFP, observed to be less than 2 months for 88% of cases, were closely approximated by the deep and wide models (within than half a month) for most streams. The shallow models either failed to capture any dry states or exhibited under-prediction of NFP and hence, were outperformed by the deep and wide models. The deep and wide architectures were also superior in capturing the flow persistence, as the shallow models were prone to over-prediction (up to 5 months). Therefore, the inclusion of a classification cell was essential to capture no-flow persistence.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions
Open Access
EGU

The two-sample Kolmogorov-Smirnov test was used to test the null hypothesis that the observed and modeled zero-flow persistence came from the same distribution against the alternative hypothesis that they came from different distributions. The null hypothesis could not be rejected for any of the models indicating that empirical CDF of the predicted NFP was similar to observed NFP. However, the K-S statistic which is a measure of the distance (D) between two CDFs was shorter (D~ 0.18) in case of deep and wide topologies as compared to shallow topologies (D~ 0.3) indicating a better fit with the observations.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

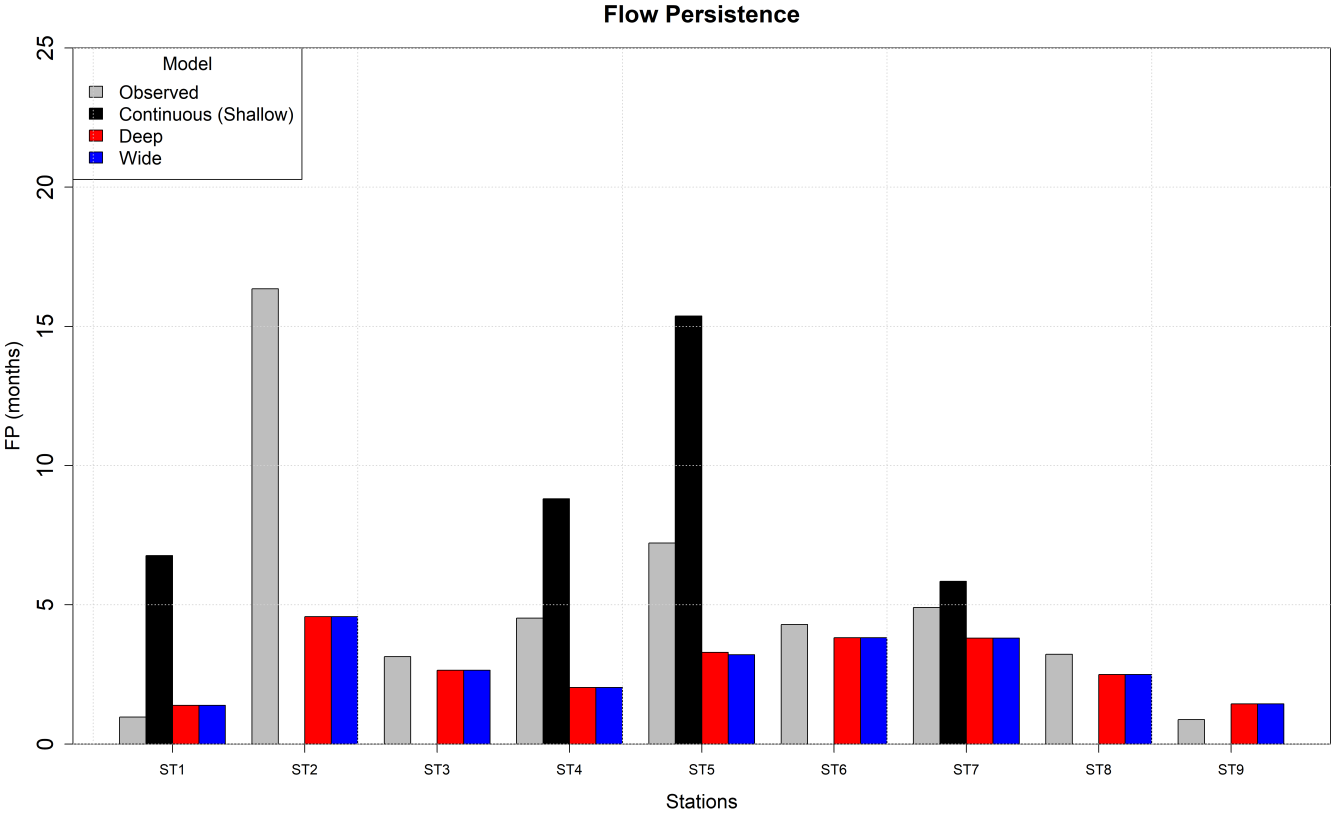


Figure 8. Comparison of observed Flow Persistence and the predictions of the shallow, deep and wide models for the nine IRES of study over the Testing Period.

The shallow architecture over-predicted flow persistence (FP) in all cases which is to be expected because of its inability to correctly predict no-flow states (Fig. 8). The deep and wide models performed much better in predicting flow persistence except when the intermittency ratio was too low (ST2 and ST5 with 12% and 17% intermittency, respectively). However, there was no general trend between the extent of intermittency and flow persistence otherwise. The testing period was generally wetter than normal and also more erratic at these locations compared to the training periods. These microclimatic fluctuations plausibly explain the noted discrepancies at these stations as ANNs are universal interpolators with minimal extrapolation abilities, if any.

The shallow model as before could not predict flow persistence in five of the nine stations. In the four stations that the shallow model was able to discern zero flow and flow states the model was able to not reject the null hypothesis of both observed data and modeled results were from the same distribution in three of them (p-value > 0.05 at ST4, ST5 and ST7) and rejected the null hypothesis in one (ST1, p-value ≤ 0.002). On the other hand, the shallow and deep models did not reject the null-hypothesis of the two-sample Kolmogorov-Smirnov (K-S) test at any of the stations (p-value > 0.05) indicating they better

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

captured the variability in the observed data. Furthermore, the distance metric of the K-S test was smaller for deep and wide models compared to the shallow model at the four sites, indicating the deep and wide models better approximated the CDF of flow persistence. This result once again highlights the advantage of including a separate classifier as part of the intermittent flow prediction algorithm.

Unlike NFP, the flow persistence varies much more widely across the stations (1 month - 16 months) and the deep and wide topologies were able to capture these transitions reasonably well and certainly better than the continuous (Shallow) model. Wetter than normal conditions during the testing phase did affect some results locally but overall the results were reasonable for the most part indicating the utility of treating intermittent flow data as a mixture and thus warranting a separate treatment of flow and no-flow states both for prediction of no-flow and flow persistence.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

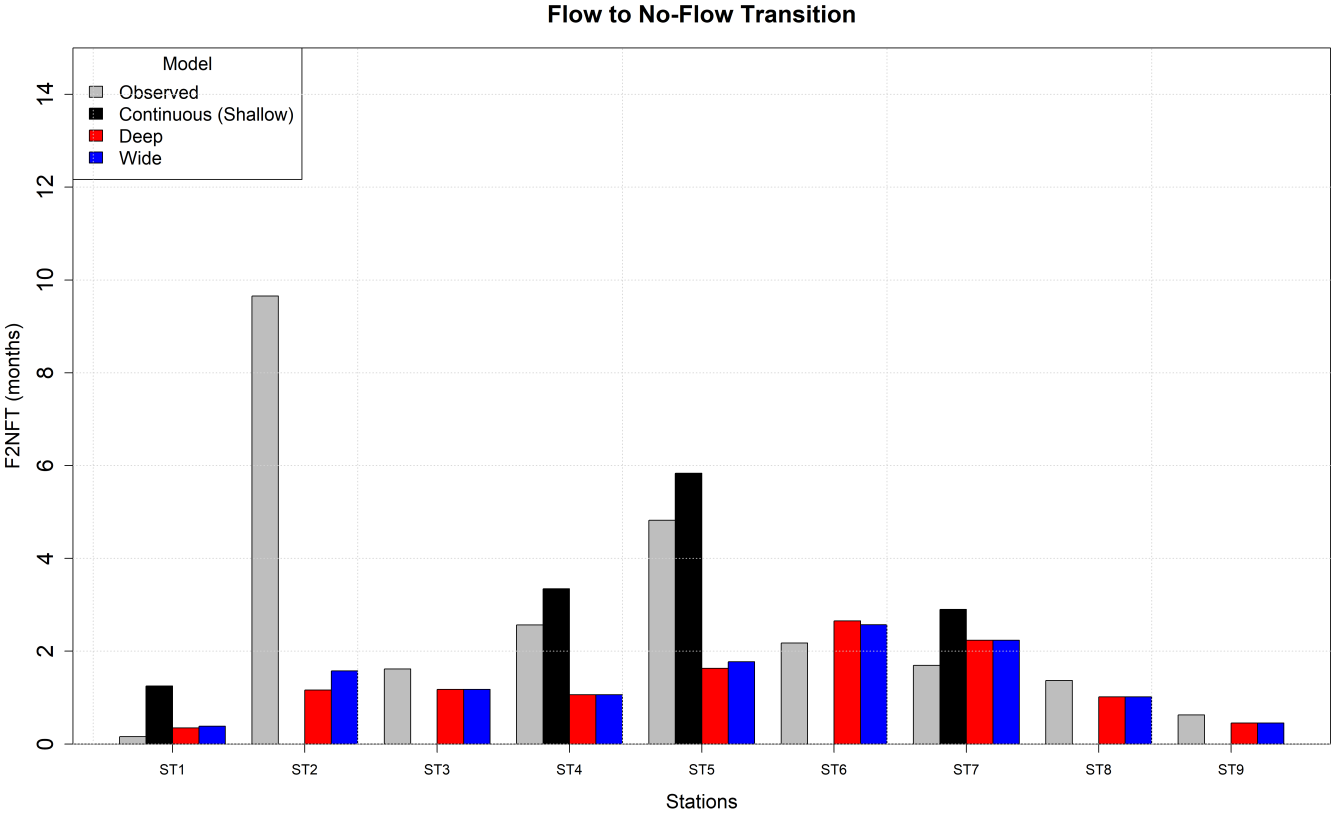


Figure 9. Comparison of observed Flow to No-Flow Transition (F2NFT) and the predictions of the shallow, deep and wide models at nine stations.

455 The F2NFT metric indicates the expected time it takes to move to a no-flow state when the system enters a flow state and is presented in Fig. 9. For highly intermittent streams this value tends to be small (e.g., ST1 and ST9) while for less intermittent streams longer transition times are to be expected (e.g, ST2 and ST5). The shallow model either over-predicts the F2NFT or does not provide any information due to lack of transitions. The over-prediction is mainly caused by over-estimation of no-flow values.

460 The predictions of deep and wide topologies tend to be fairly close in almost all cases and they definitely capture the F2NFT better than the shallow model. The predictions of deep and wide topologies are also close to observed values in most cases except for ST2 and ST5 which, as explained previously, had significantly different climatic conditions over the testing period as compared to their training. The F2NFT show similar trends as FP (see Fig. 8) but tend to be of shorter duration. This is to be expected because,for an example, a no-flow event is followed by a 3 month flow event and a no-flow event. Then, the FP is
465 3 months but F2NFT equals the average of 1, 2 and 3 months, i.e., 2 months.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions

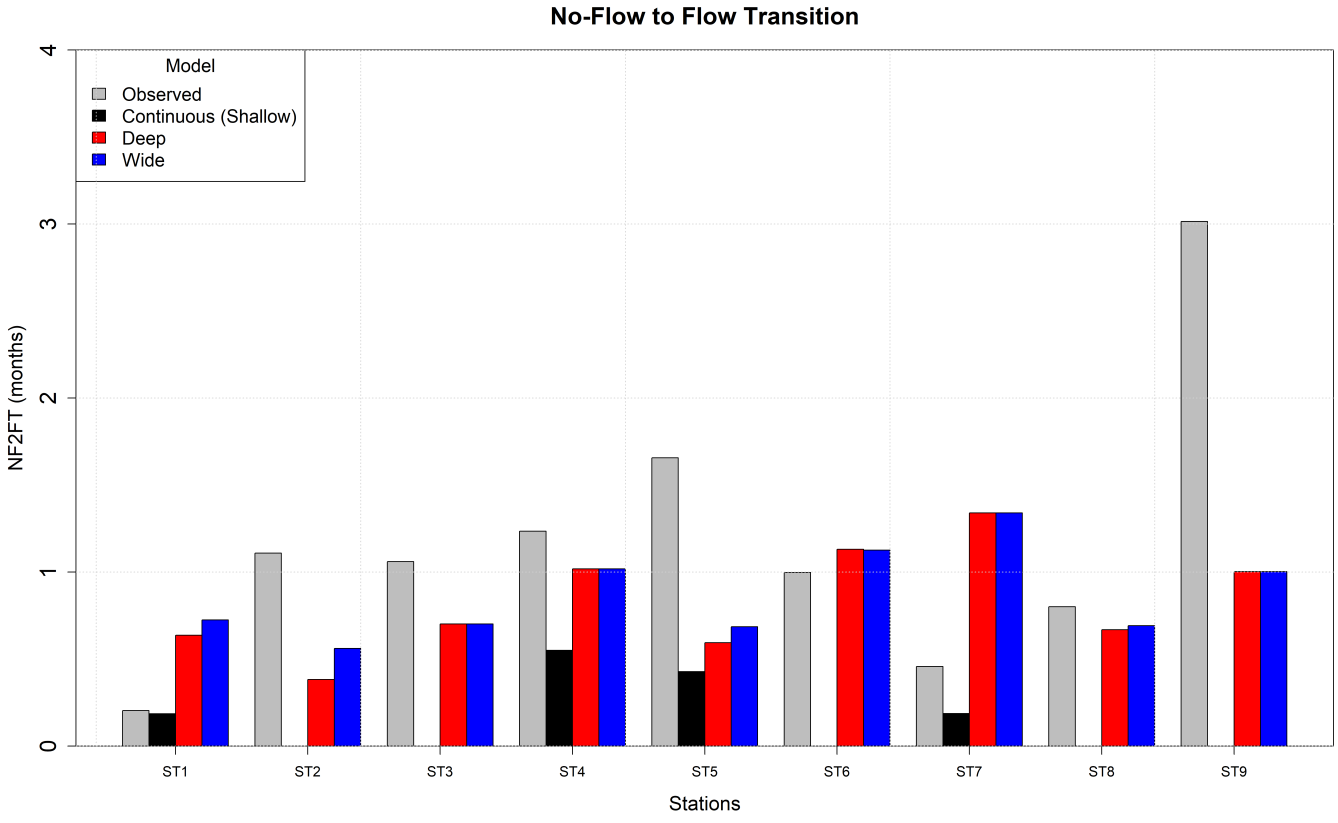


Figure 10. Comparison of observed No-Flow to Flow (NF2FT) and the predictions of the shallow, deep and wide models for the nine IRES of study.

Unlike the F2NFT discussed above, the NF2FT represents the expected time necessary for transitioning from a no-flow to a flow state. The NF2FT were much shorter than F2NFT (compare the Y-axis of Fig. 9 and 10). While the deep and wide models are unable to correctly predict the NF2FT transitions, the errors tend to be within a month (the time-scale of prediction) at 8 of the 9 stations. ST9 with highest level of intermittency has an expected NF2FT of 3 months while the deep and wide models underpredict this value by two months. This result largely arises because the models tend to have higher OPE (over-prediction error) at this station (see Fig. 6d). The performance of the shallow model is not too far off from the deep and wide topologies at the 4 stations where the model is able to discern between flow and no-flow states after application of the cut-off. However, the deep and wide topologies consistently provide estimates of NF2FT at all stations while, the shallow model is unable to do so at 55% of the sites.

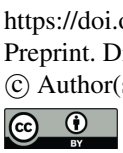
https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

7.4 Deep and wide topologies for IRES non-zero flowrate prediction

The prediction of the non-zero flow states were compared to observed values by constructing empirical CDFs using Gringorten plotting positions (Gringorten, 1963) and generating P-P plots (Gan et al., 1991) and are depicted in Fig. 11 for log(Q+1) transformation and in Fig. 12 for non-transformed data.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021



Hydrology and Earth System Sciences
Discussions
Open Access
EGU

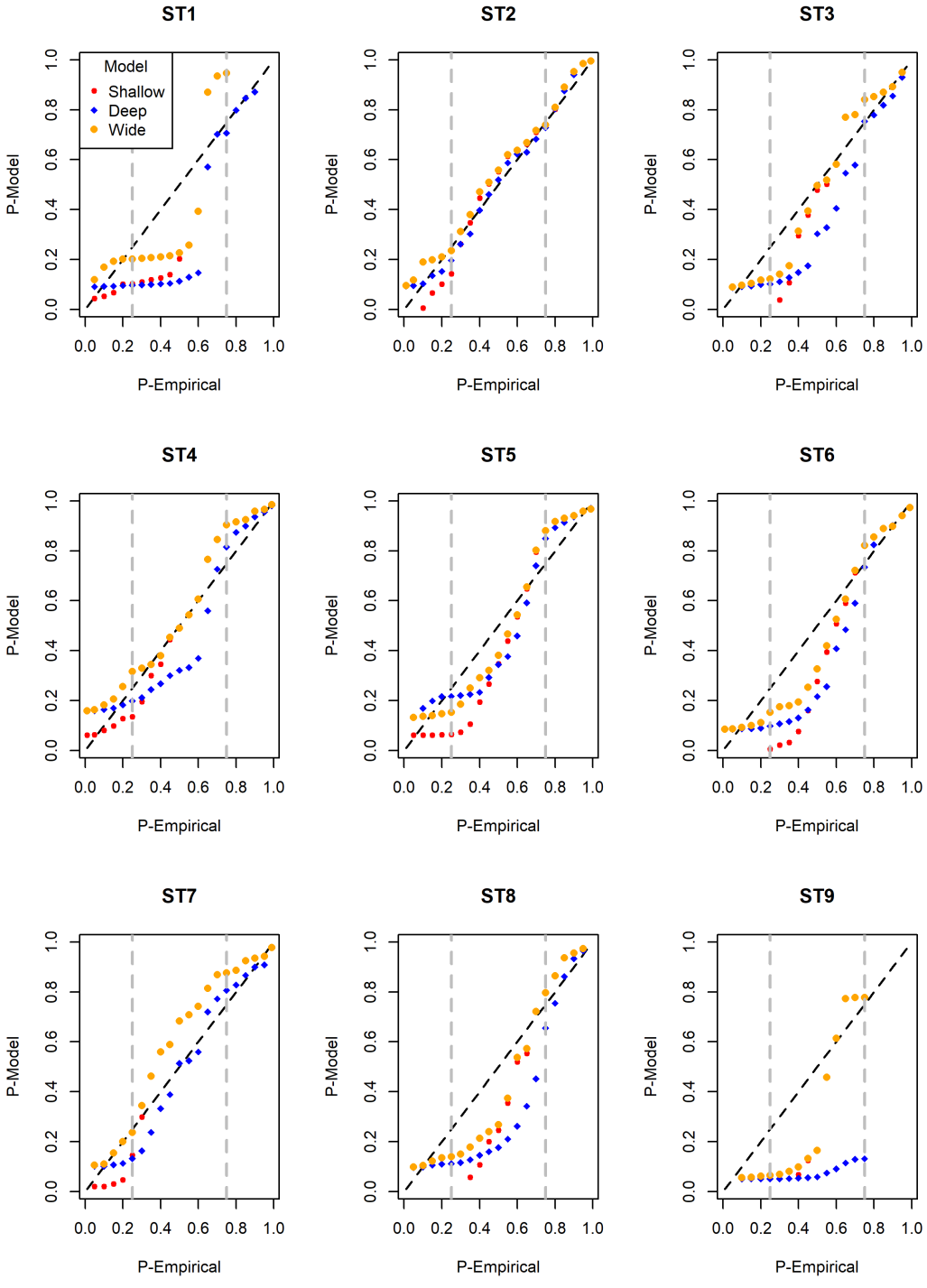


Figure 11. The probability- probability plot for the shallow, deep and wide models across the nine IRES of study over log(Q+1) transformed flow data.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

As seen in Fig. 11 and 12 the application of the log(Q+1) transformation is recommended for low flow analysis of IRES. The models trained on the transformed datasets provided closer predictions with lower errors. Also, for 55% of the studied IRES, the models trained on transformed flowrates provided better predictions for the entire flow range with notably lower errors (see Table 2). For the other four streams (ST4, ST5, ST6, and ST7), using the transformed data resulted in large errors (by orders of magnitude) in the prediction of extreme high flows. These latter stations were the ones which had one or more recording-breaking flood events during the testing period, several times greater than the maximum peak flow of the training period. Extrapolation beyond the training set, as discussed before, limits the performance of the neural network models and the risk of predicting with large errors is higher when using log(Q+1) transformation.

Three distinct flow regimes were observed in the observed and predicted intermittent flow time-series: the zero and extreme low flows, the mid-flow pulses, and the extreme high flows. The wide models dominate the prediction of the no- and low-flow events for all IRES of study. For predicting the lower mid-flow pulses, the wide topologies were still reasonable and the deep models provided closer approximations for higher flow rates. Following the same predictive behavior, for the extreme high flows that were several orders of magnitude greater than mid-flow pulses and very limited (less than 5 events during the testing period), the deep models had the least errors and the superior performance among the three models. Capturing the extreme large peak flows was a challenging task, even for the deep models and there were multiple cases of under-predictions and few cases of over-prediction. Overall, the deep network topology is recommended as a more appropriate tool when zero and extremely high flows are of simultaneous interest, such as maintaining dry and flooding cycles to ensure sustainable spawning of both small and large fauna over the course of a year (Postel and Richter, 2012). .

8 Summary and conclusions

In this paper, three Artificial Neural Network-type architectures were investigated for streamflow forecasting in IRES: The shallow model which treated IRES flow time-series as a continuum, and the deep and wide topologies that treated the IRES flow data as a mixture-type consisting of flow and no-flow records using a separate classification cell that is stacked either in series or parallel with a regression cell for the deep and wide models, respectively. The performance of the models was evaluated with a suite of continuous and discrete goodness-of-fit metrics and was explored across nine IRES with varying hydro-climatic features.

The addition of a separate classifier is recommended here, as without one, the shallow model was unable to predict absolute zero-flow events and sometimes even predicted physically unrealistic negative flowrates. Effective capturing of no-flow states by the deep and wide topologies provided useful insights of the IRES flow time-series (e.g., persistence and transitions of states) that otherwise would not have been possible with a shallow network.

For low flow estimation, the wide models were found superior as in the case of a misclassification of no-flow states (OPE) the errors in the estimates of the deep model were higher. However, for high flow analysis, the deep model outperformed the other counterparts and provided closer predictions.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

The application of two pre-processing techniques, namely SMOTE and Log(Q+1) transformation was also investigated. As IRES flowrates are imbalanced datasets, SMOTE was useful in reducing the extent of imbalance, providing a better trade-off of under-prediction error (UPE) and over-prediction error (OPE), and thereby achieving better classification performance at most stations, except under very high level of flow intermittency. Applying a log transformation with a one-unit shift was found beneficial, specifically for low flow analysis. However, for several stations, the models using the transformed data showed large errors in predicting extreme high flows, limiting the utility of this transformation technique for high flow analysis.

It is recommended that intermittent flow time-series not be treated as a continuum but rather be modeled as a mixture containing zero (or non-zero) inflated data. A stacked topology neural network comprised of an initial classifier cell provides better zero flow predictions and helps maintain its Markovian properties (no-flow and flow persistence and no-flow to flow and flow to no-flow transitions). A parallel coupling (Wide model) of the classification cell with a regression cell is beneficial when the focus is on estimating low flows and having lower errors in case of zero-flow misclassification. On the other hand, a series coupling (deep model) with the regression cell is beneficial when modeling extreme (i.e., zero and very high) flows. Extreme learning machines (ELMs) calibrated using greedy algorithms and LASSO regularization enabled computationally efficient calibration while avoiding issues related to overfitting and curse-of-dimensionality and can be adopted to train stacked ANNs for improved forecasting of intermittent flows.

Code and data availability. The data set and custom scripts (in R) developed for this study are available at https://doi.org/10.17605/OSF.IO/UKAGQ (Forghanparast et al., 2021)

Author contributions. VU and EAH conceptualized the study. FF carried out data curation and investigation and visualized the project. FF and EAH validated the data. FF and VU conducted the formal analysis and developed the model software. FF, EAH, and VU developed the methodology and wrote, reviewed, and edited the manuscript. EAH acquired the funding, provided resources, and administered the project. VU supervised the project.

Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgements. The authors would like to acknowledge financial support from J.T. and Margaret Talkington Fellowship, TTU Water Resources Center, and the Department of Civil, Environmental, and Construction Engineering at Texas Tech University.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table S1 of this paper) for producing and making available their model output. For CMIP the

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

References

Adnan, R. M., Yuan, X., Kisi, O., and Yuan, Y.: Streamflow forecasting using artificial neural network and support vector machine models, American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 29, 286–294, https://www.asrjetsjournal.org/index.php/American_Scientific_Journal/article/view/2814/1086, 2017.

Adnan, R. M., Liang, Z., Trajkovic, S., Zounemat-Kermani, M., Li, B., and Kisi, O.: Daily streamflow prediction using optimally pruned extreme learning machine, Journal of Hydrology, 577, 123 981, https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.123981, 2019.

Afkhamifar, S. and Sarraf, A.: Prediction of groundwater level in Urmia Plain aquifer using hybrid model of wavelet Transform-Extreme Learning Machine based on quantum particle swarm optimization, Watershed Engineering and Management, 12, 351–364, https://doi.org/10.22092/ijwmse.2019.126515.1669, 2020.

Aksoy, H. and Bayazit, M.: A model for daily flows of intermittent streams, Hydrological Processes, 14, 1725–1744, https://doi.org/https://doi.org/10.1002/1099-1085(200007)14:10<1725::AID-HYP108>3.0.CO;2-L, 2000.

Al-Shayea, Q.: Artificial Neural Networks in Medical Diagnosis, International Journal of Computer Science Issues, 8, 150–154, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.402.9531&rep=rep1&type=pdf#page=174, 2011.

Amato, F., López, A., Peña-Méndez, E. M., Vaňhara, P., Hampl, A., and Havel, J.: Artificial neural networks in medical diagnosis, Journal of Applied Biomedicine, 11, 47–58, https://doi.org/https://doi.org/10.2478/v10136-012-0031-x, 2013.

Atiquzzaman, M. and Kandasamy, J.: Prediction of hydrological time-series using extreme learning machine, Journal of Hydroinformatics, 18, 345–353, https://doi.org/10.2166/hydro.2015.020, 2015.

Augustin, N. H., Beevers, L., and Sloan, W. T.: Predicting river flows for future climates using an autoregressive multinomial logit model, Water Resources Research, 44, https://doi.org/https://doi.org/10.1029/2006WR005127, 2008.

Azarnivand, A., Camporese, M., Alaghmand, S., and Daly, E.: Simulated response of an intermittent stream to rainfall frequency patterns, Hydrological Processes, 34, 615–632, https://doi.org/https://doi.org/10.1002/hyp.13610, 2020.

Badrzadeh, H., Sarukkalige, R., and Jayawardena, A. W.: Intermittent stream flow forecasting and modelling with hybrid wavelet neuro-fuzzy model, Hydrology Research, 49, 27–40, https://doi.org/10.2166/nh.2017.163, 2017.

Barber, C., Lamontagne, J. R., and Vogel, R. M.: Improved estimators of correlation and R2 for skewed hydrologic data, Hydrological Sciences Journal, 65, 87–101, https://doi.org/10.1080/02626667.2019.1686639, 2020.

Bektas, J., Ibrikci, T., and Ozcan, I. T.: Classification of Real Imbalanced Cardiovascular Data Using Feature Selection and Sampling Methods: A Case Study with Neural Networks and Logistic Regression, International Journal on Artificial Intelligence Tools, 26, 1750 019, https://doi.org/10.1142/S0218213017500191, 2017.

Belilovsky, E., Eickenberg, M., and Oyallon, E.: Greedy Layerwise Learning Can Scale To ImageNet, in: Proceedings of the 36th International Conference on Machine Learning, edited by Chaudhuri, K. and Salakhutdinov, R., vol. 97 of *Proceedings of Machine Learning Research*, pp. 583–593, PMLR, http://proceedings.mlr.press/v97/belilovsky19a.html, 2019.

Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., et al.: Greedy layer-wise training of deep networks, Advances in neural information processing systems, 19, 153, 2006.

Billi, P., Demissie, B., Nyssen, J., Moges, G., and Fazzini, M.: Meander hydromorphology of ephemeral streams: Similarities and differences with perennial rivers, Geomorphology, 319, 35–46, https://doi.org/https://doi.org/10.1016/j.geomorph.2018.07.003, 2018.

Blagus, R. and Lusa, L.: Evaluation of SMOTE for High-Dimensional Class-Imbalanced Microarray Data, in: 2012 11th International Conference on Machine Learning and Applications, vol. 2, pp. 89–94, https://doi.org/10.1109/ICMLA.2012.183, 2013.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Borg Galea, A., Sadler, J. P., Hannah, D. M., Datry, T., and Dugdale, S. J.: Mediterranean intermittent rivers and ephemeral streams: Challenges in monitoring complexity, Ecohydrology, 12, e2149, https://doi.org/https://doi.org/10.1002/eco.2149, 2019.

Boulton, A. J., Rolls, R. J., Jaeger, K. L., and Datry, T.: Chapter 2.3 - Hydrological Connectivity in Intermittent Rivers and Ephemeral Streams, pp. 79–108, Academic Press, https://doi.org/https://doi.org/10.1016/B978-0-12-803835-2.00004-8, 2017.

Brekke, L., Thrasher, B. L., Maurer, E. P., and Pruitt, T.: Downscaled CMIP3 and CMIP5 climate projections: release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs, US Department of the Interior, Bureau of Reclamation, Technical Service Center, Denver, Colorado, USA, 2013.

Chawla, N. V., Bowyer, K. W., O. Hall, L., and Kegelmeyer, W. P.: SMOTE: synthetic minority over-sampling technique, J. Artif. Int. Res., 16, 321–357, https://doi.org/https://doi.org/10.1613/jair.953, 2002.

Chebaane, M., Salas, J. D., and Boes, D. C.: Product Periodic Autoregressive Processes for Modeling Intermittent Monthly Streamflows, Water Resources Research, 31, 1513–1518, https://doi.org/https://doi.org/10.1029/95WR00144, 1995.

Cheng, M., Fang, F., Kinouchi, T., Navon, I. M., and Pain, C. C.: Long lead-time daily and monthly streamflow forecasting using machine learning methods, Journal of Hydrology, 590, 125 376, https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125376, 2020.

Cherkauer, K. A., Bowling, L. C., and Lettenmaier, D. P.: Variable infiltration capacity cold land process model updates, Global and Planetary Change, 38, 151–159, https://doi.org/https://doi.org/10.1016/S0921-8181(03)00025-0, 2003.

Cigizoglu, H. K.: Estimation, forecasting and extrapolation of river flows by artificial neural networks, Hydrological Sciences Journal, 48, 349–361, https://doi.org/10.1623/hysj.48.3.349.45288, 2003.

Cigizoglu, H. K.: Application of Generalized Regression Neural Networks to Intermittent Flow Forecasting and Estimation, Journal of Hydrologic Engineering, 10, 336–341, https://doi.org/doi:10.1061/(ASCE)1084-0699(2005)10:4(336), 2005.

Cocco Mariani, V., Hennings Och, S., dos Santos Coelho, L., and Domingues, E.: Pressure prediction of a spark ignition single cylinder engine using optimized extreme learning machine models, Applied Energy, 249, 204–221, https://doi.org/https://doi.org/10.1016/j.apenergy.2019.04.126, 2019.

Costa, V., Fernandes, W., and Starick, n.: Identifying Regional Models for Flow Duration Curves with Evolutionary Polynomial Regression: Application for Intermittent Streams, Journal of Hydrologic Engineering, 25, 04019 059, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001873, 2020.

Courtwright, J. and May, C. L.: Importance of terrestrial subsidies for native brook trout in Appalachian intermittent streams, Freshwater Biology, 58, 2423–2438, https://doi.org/https://doi.org/10.1111/fwb.12221, 2013.

Curran-Everett, D.: Explorations in statistics: the log transformation, Advances in Physiology Education, 42, 343–347, https://doi.org/10.1152/advan.00018.2018, 2018.

Danandeh Mehr, A.: An improved gene expression programming model for streamflow forecasting in intermittent streams, Journal of Hydrology, 563, 669–678, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.06.049, 2018.

Datry, T., Pella, H., Leigh, C., Bonada, N., and Hugueny, B.: A landscape approach to advance intermittent river ecology, Freshwater Biology, 61, 1200–1213, https://doi.org/https://doi.org/10.1111/fwb.12645, 2016.

Datry, T., Singer, G., Sauquet, E., Jorda Capdevilla, D., Von Schiller, D., Subbington, R., Magrand, C., Paril, P., Milisa, M., Acuna, V., Alves, M., Augeard, B., Brunke, M., Cid, N., Csabai, Z., England, J., Froebrich, J., Koundouri, P., Lamouroux, N., Marti, E., Morais, M., Munne, A., Mutz, M., Pesic, V., Previsic, A., Reynaud, A., Robinson, C., Sadler, J., Skoulikidis, N., Terrier, B., Tockner, K., Vesely, D., and Zoppini, A.: Science and management of intermittent rivers and ephemeral streams (SMIRES), Research Ideas and Outcomes, 3, 23 p., https://doi.org/10.3897/rio.3.e21774, 2017.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions

Datry, T., Boulton, A. J., Bonada, N., Fritz, K., Leigh, C., Sauquet, E., Tockner, K., Hugueny, B., and Dahm, C. N.: Flow intermittence and ecosystem services in rivers of the Anthropocene, Journal of Applied Ecology, 55, 353–364, https://doi.org/https://doi.org/10.1111/1365-2664.12941, 2018.

Deo, R. C. and Şahin, M.: Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia, Atmospheric Research, 153, 512–525, https://doi.org/https://doi.org/10.1016/j.atmosres.2014.10.016, 2015.

Deo, R. C., Samui, P., and Kim, D.: Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models, Stochastic Environmental Research and Risk Assessment, 30, 1769–1784, https://doi.org/10.1007/s00477-015-1153-y, 2016.

Ding, S., Xu, X., and Nie, R.: Extreme learning machine and its applications, Neural Computing and Applications, 25, 549–556, https://doi.org/10.1007/s00521-013-1522-8, 2014.

Duan, M., Li, K., Yang, C., and Li, K.: A hybrid deep learning CNN–ELM for age and gender classification, Neurocomputing, 275, 448–461, https://doi.org/https://doi.org/10.1016/j.neucom.2017.08.062, 2018.

Eger, S., Youssef, P., and Gurevych, I.: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, https://doi.org/10.18653/v1/D18-1472, 2018.

Eng, K., Wolock, D. M., and Dettinger, M. D.: Sensitivity of Intermittent Streams to Climate Variations in the USA, River Research and Applications, 32, 885–895, https://doi.org/https://doi.org/10.1002/rra.2939, 2016.

Eris, E., Aksoy, H., Onoz, B., Cetin, M., Yuce, M. I., Selek, B., Aksu, H., Burgan, H. I., Esit, M., Yildirim, I., and Karakus, E. U.: Frequency analysis of low flows in intermittent and non-intermittent rivers from hydrological basins in Turkey, Water Supply, 19, 30–39, https://doi.org/10.2166/ws.2018.051, 2018.

Fakhr, M. W., Youssef, E. S., and El-Mahallawy, M. S.: L1-regularized least squares sparse extreme learning machine for classification, in: 2015 International Conference on Information and Communication Technology Research (ICTRC), pp. 222–225, https://doi.org/10.1109/ICTRC.2015.7156462.

Feng, C., Huang, G., Lin, Q., and Gay, R.: Error Minimized Extreme Learning Machine With Growth of Hidden Nodes and Incremental Learning, IEEE Transactions on Neural Networks, 20, 1352–1357, https://doi.org/10.1109/TNN.2009.2024147, 2009.

Feng, X., Liang, Y., Shi, X., Xu, D., Wang, X., and Guan, R.: Overfitting Reduction of Text Classification Based on AdaBELM, Entropy, 19, 330, https://doi.org/https://doi.org/10.3390/e19070330, 2017.

Fernández, A., Garcia, S., Herrera, F., and Chawla, N. V.: SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary, Journal of artificial intelligence research, 61, 863–905, https://doi.org/https://doi.org/10.1613/jair.1.11192, 2018.

Forghanparast, F., Hernandez, E. A., and Uddameri, V.: Data and Codes for - Deep, Wide or Shallow - Artificial Neural Network Topologies for Predicting Intermittent Flows, Open Science Framework (OSF) data repository, https://doi.org/10.17605/OSF.IO/UKAGQ, 2021.

Friedman, J., Hastie, T., and Tibshirani, R.: Regularization Paths for Generalized Linear Models via Coordinate Descent, Journal of statistical software, 33, 1–22, https://pubmed.ncbi.nlm.nih.gov/20808728https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929880/, 2010.

Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, 29, 1189–1232, http://www.jstor.org/stable/2699986, 2001.

Gallart, F., Prat, N., García-Roger, E. M., Latron, J., Rieradevall, M., Llorens, P., Barberá, G., Brito, D., De Girolamo, A., and Porto, A. L.: A novel approach to analysing the regimes of temporary streams in relation to their controls on the composition and structure of aquatic biota, Hydrology and Earth System Sciences, 16, 3165–3182, 2012.

Gan, F. F., Koehler, K. J., and Thompson, J. C.: Probability Plots and Distribution Curves for Assessing the Fit of Probability Models, The American Statistician, 45, 14–21, https://doi.org/10.1080/00031305.1991.10475759, 1991.

Gao, A. and Martos, P.: Log Transformation and the Effect on Estimation, Implication, and Interpretation of Mean and Measurement Uncertainty in Microbial Enumeration, Journal of AOAC INTERNATIONAL, 102, 233–238, https://doi.org/10.5740/jaoacint.18-0161, 2019.

Ghaseminejad, A. and Uddameri, V.: Physics-inspired integrated space–time artificial neural networks for regional groundwater flow modeling, Hydrol. Earth Syst. Sci., 24, 5759–5779, https://doi.org/10.5194/hess-24-5759-2020, 2020.

Goodrich, D. C., Kepner, W. G., Levick, L. R., and Wigington Jr, P. J.: Southwestern Intermittent and Ephemeral Stream Connectivity, JAWRA Journal of the American Water Resources Association, 54, 400–422, https://doi.org/https://doi.org/10.1111/1752-1688.12636, 2018.

Grey, D. and Sadoff, C. W.: Sink or Swim? Water security for growth and development, Water Policy, 9, 545–571, https://doi.org/10.2166/wp.2007.021, 2007.

Gringorten, I. I.: A plotting rule for extreme probability paper, Journal of Geophysical Research (1896-1977), 68, 813–814, https://doi.org/https://doi.org/10.1029/JZ068i003p00813, 1963.

Gutiérrez-Jurado, K. Y., Partington, D., Batelaan, O., Cook, P., and Shanafield, M.: What Triggers Streamflow for Intermittent Rivers and Ephemeral Streams in Low-Gradient Catchments in Mediterranean Climates, Water Resources Research, 55, 9926–9946, https://doi.org/https://doi.org/10.1029/2019WR025041, 2019.

Hadi, S. J. and Tombul, M.: Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination, Journal of Hydrology, 561, 674–687, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.04.036, 2018.

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., and Bing, G.: Learning from class-imbalanced data: Review of methods and applications, Expert Systems with Applications, 73, 220–239, https://doi.org/https://doi.org/10.1016/j.eswa.2016.12.035, 2017.

Hamilton, L. C.: Regression with graphics: A second course in applied statistics, 1992.

Hastie, T., Narasimhan, B., and Tibshirani, R.: The Relaxed Lasso, 2021.

Hill, M. J. and Milner, V. S.: Ponding in intermittent streams: A refuge for lotic taxa and a habitat for newly colonising taxa?, Science of The Total Environment, 628-629, 1308–1316, https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.02.162, 2018.

Hinton, G. E., Osindero, S., and Teh, Y.-W.: A Fast Learning Algorithm for Deep Belief Nets, Neural Computation, 18, 1527–1554, https://doi.org/10.1162/neco.2006.18.7.1527, 2006.

Huang, G.-B. and Chen, L.: Convex incremental extreme learning machine, Neurocomputing, 70, 3056–3062, https://doi.org/https://doi.org/10.1016/j.neucom.2007.02.009, 2007.

Huang, G.-B., Zhu, Q.-Y., and Siew, C.-K.: Extreme learning machine: Theory and applications, Neurocomputing, 70, 489–501, https://doi.org/https://doi.org/10.1016/j.neucom.2005.12.126, 2006.

Inaba, F. K., Teatini Salles, E. O., Perron, S., and Caporossi, G.: DGR-ELM–Distributed Generalized Regularized ELM for classification, Neurocomputing, 275, 1522–1530, https://doi.org/https://doi.org/10.1016/j.neucom.2017.09.090, 2018.

Jaeger, K. L., Sutfin, N. A., Tooth, S., Michaelides, K., and Singer, M.: Chapter 2.1 - Geomorphology and Sediment Regimes of Intermittent Rivers and Ephemeral Streams, pp. 21–49, Academic Press, https://doi.org/https://doi.org/10.1016/B978-0-12-803835-2.00002-4, 2017.

Johnson, R. and Zhang, T.: Accelerating Stochastic Gradient Descent using Predictive Variance Reduction, in: Advances in Neural Information Processing Systems, edited by Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., vol. 26, pp. 315–323, Curran Associates, Inc., https://proceedings.neurips.cc/paper/2013/file/ac1dd209cbcc5e5d1c6e28598e8cbbe8-Paper.pdf, 2013.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021


Hydrology and Earth System Sciences
Discussions

Jothiprakash, V., Magar, R., and Kalkutki, S.: Rainfall–runoff models using adaptive neuro–fuzzy inference system (ANFIS) for an intermit-
690 tent River, Int J Artif Intell, 3, 1–23, 2009.

Kaletová, T., Loures, L., Castanho, R. A., Aydin, E., Gama, J. T., Loures, A., and Truchy, A.: Relevance of Intermittent Rivers and Streams in Agricultural Landscape and Their Impact on Provided Ecosystem Services—A Mediterranean Case Study, International Journal of Environmental Research and Public Health, 16, https://doi.org/10.3390/ijerph16152693, 2019.

Kampf, S. K., Faulconer, J., Shaw, J. R., Lefsky, M., Wagenbrenner, J. W., and Cooper, D. J.: Rainfall Thresholds for Flow Generation in
695 Desert Ephemeral Streams, Water Resources Research, 54, 9935–9950, https://doi.org/https://doi.org/10.1029/2018WR023714, 2018.

Kang, K.-W., Park, C.-Y., and Kim, J.-H.: Neural network and its application to rainfall-runoff forecasting, Korean journal of hydrosciences, 4, 1–9, https://www.koreascience.or.kr/article/JAKO199311920101523.page, 1993.

Karaouzas, I., Theodoropoulos, C., Vardakas, L., Kalogianni, E., and Th. Skoulikidis, N.: A review of the effects of pollution and water scarcity on the stream biota of an intermittent Mediterranean basin, River Research and Applications, 34, 291–299,
700 https://doi.org/https://doi.org/10.1002/rra.3254, 2018.

Katz, G. L., Denslow, M. W., and Stromberg, J. C.: The Goldilocks effect: intermittent streams sustain more plant species than those with perennial or ephemeral flow, Freshwater Biology, 57, 467–480, https://doi.org/https://doi.org/10.1111/j.1365-2427.2011.02714.x, 2012.

Kişi, z.: Neural Networks and Wavelet Conjunction Model for Intermittent Streamflow Forecasting, Journal of Hydrologic Engineering, 14, 773–782, https://doi.org/doi:10.1061/(ASCE)HE.1943-5584.0000053, 2009.

705 Kisi, O., Nia, A. M., Gosheh, M. G., Tajabadi, M. R. J., and Ahmadi, A.: Intermittent Streamflow Forecasting by Using Several Data Driven Techniques, Water Resources Management, 26, 457–474, https://doi.org/10.1007/s11269-011-9926-7, 2012.

Kisi, O., Alizamir, M., and Shiri, J.: Conjunction Model Design for Intermittent Streamflow Forecasts: Extreme Learning Machine with Discrete Wavelet Transform, pp. 171–181, Springer, https://doi.org/https://doi.org/10.1007/978-981-15-5772-9_9, 2021.

Knoblock, C. A., Lerman, K., Minton, S., and Muslea, I.: Accurately and reliably extracting data from the web: A machine learning approach,
710 vol. 111, pp. 275–287, Springer, https://doi.org/https://doi.org/10.1007/978-3-7908-1772-0_17, 2003.

Krause, P., Boyle, D. P., and Bäse, F.: Comparison of different efficiency criteria for hydrological model assessment, Advances in Geosciences, 5, 89–97, https://doi.org/10.5194/adgeo-5-89-2005, 2005.

Kuhn, M.: Building Predictive Models in R Using the caret Package, Journal of Statistical Software, 028, https://EconPapers.repec.org/RePEc:jss:jstsof:v:028:i05, 2008.

715 Lai, J., Wang, X., Li, R., Song, Y., and Lei, L.: BD-ELM: A Regularized Extreme Learning Machine Using Biased DropConnect and Biased Dropout, Mathematical Problems in Engineering, 2020, 3604 579, https://doi.org/10.1155/2020/3604579, 2020.

Lan, Y., Soh, Y. C., and Huang, G.-B.: Constructive hidden nodes selection of extreme learning machine for regression, Neurocomputing, 73, 3191–3199, https://doi.org/https://doi.org/10.1016/j.neucom.2010.05.022, 2010.

Larned, S. T., Datry, T., Arscott, D. B., and Tockner, K.: Emerging concepts in temporary-river ecology, Freshwater Biology, 55, 717–738,
720 https://doi.org/https://doi.org/10.1111/j.1365-2427.2009.02322.x, 2010.

Larochelle, H., Bengio, Y., Louradour, J., and Lamblin, P.: Exploring Strategies for Training Deep Neural Networks, J. Mach. Learn. Res., 10, 1–40, https://www.jmlr.org/papers/volume10/larochelle09a/larochelle09a.pdf, 2009.

Leigh, C.: Dry-season changes in macroinvertebrate assemblages of highly seasonal rivers: responses to low flow, no flow and antecedent hydrology, Hydrobiologia, 703, 95–112, https://doi.org/10.1007/s10750-012-1347-y, 2013.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Leigh, C., Boulton, A. J., Courtwright, J. L., Fritz, K., May, C. L., Walker, R. H., and Datry, T.: Ecological research and management of intermittent rivers: an historical review and future directions, Freshwater Biology, 61, 1181–1199, https://doi.org/https://doi.org/10.1111/fwb.12646, 2016.

Levick, L. R., Goodrich, D. C., Hernandez, M., Fonseca, J., Semmens, D. J., Stromberg, J. C., Tluczek, M., Leidy, R. A., Scianni, M., and Guertin, D. P.: The ecological and hydrological significance of ephemeral and intermittent streams in the arid and semi-arid American Southwest, US Environmental Protection Agency, Office of Research and Development, 2008.

Li, M., Robertson, D. E., Wang, Q. J., Bennett, J. C., and Perraud, J.-M.: Reliable hourly streamflow forecasting with emphasis on ephemeral rivers, Journal of Hydrology, p. 125739, https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125739, 2020.

Lima, A. R., Cannon, A. J., and Hsieh, W. W.: Forecasting daily streamflow using online sequential extreme learning machines, Journal of Hydrology, 537, 431–443, https://doi.org/https://doi.org/10.1016/j.jhydrol.2016.03.017, 2016.

Liu, X., Yang, T., Hsu, K., Liu, C., and Sorooshian, S.: Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau, Hydrol. Earth Syst. Sci., 21, 169–181, https://doi.org/10.5194/hess-21-169-2017, 2017.

Lowe, W. H., Likens, G. E., and Power, M. E.: Linking Scales in Stream Ecology, BioScience, 56, 591–597, https://doi.org/10.1641/0006-3568(2006)56[591:Lsise]2.0.Co;2, 2006.

Makwana, J. J. and Tiwari, M. K.: Intermittent Streamflow Forecasting and Extreme Event Modelling using Wavelet based Artificial Neural Networks, Water Resources Management, 28, 4857–4873, https://doi.org/10.1007/s11269-014-0781-1, 2014.

Martínez-Martínez, J. M., Escandell-Montero, P., Soria-Olivas, E., Martín-Guerrero, J. D., Magdalena-Benedito, R., and Gómez-Sanchis, J.: Regularized extreme learning machine for regression problems, Neurocomputing, 74, 3716–3721, https://doi.org/https://doi.org/10.1016/j.neucom.2011.06.013, 2011.

Min, L., Vasilevskiy, P. Y., Wang, P., Pozdniakov, S. P., and Yu, J.: Numerical Approaches for Estimating Daily River Leakage from Arid Ephemeral Streams, Water, 12, 499, https://doi.org/10.3390/w12020499, 2020.

Mouatadid, S. and Adamowski, J.: Using extreme learning machines for short-term urban water demand forecasting, Urban Water Journal, 14, 630–638, https://doi.org/10.1080/1573062X.2016.1236133, 2017.

Naghizadeh, A., Metaxas, D. N., and Liu, D.: Greedy auto-augmentation for n-shot learning using deep neural networks, Neural Networks, 135, 68–77, https://doi.org/https://doi.org/10.1016/j.neunet.2020.11.015, 2021.

Nanda, T., Sahoo, B., and Chatterjee, C.: Enhancing real-time streamflow forecasts with wavelet-neural network based error-updating schemes and ECMWF meteorological predictions in Variable Infiltration Capacity model, Journal of Hydrology, 575, 890–910, https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.05.051, 2019.

Niu, W.-j., Feng, Z.-k., Chen, Y.-b., Zhang, H.-r., and Cheng, C.-t.: Annual Streamflow Time Series Prediction Using Extreme Learning Machine Based on Gravitational Search Algorithm and Variational Mode Decomposition, Journal of Hydrologic Engineering, 25, 04020 008, https://doi.org/doi:10.1061/(ASCE)HE.1943-5584.0001902, 2020.

Oreskes, N.: Evaluation (not validation) of quantitative models, Environmental Health Perspectives, 106, 1453–1460, https://doi.org/10.1289/ehp.98106s61453, 1998.

Postel, S. and Richter, B.: Rivers for life: managing water for people and nature, Island Press, 2012.

Pushpalatha, R., Perrin, C., Moine, N. L., and Andréassian, V.: A review of efficiency criteria suitable for evaluating low-flow simulations, Journal of Hydrology, 420-421, 171–182, https://doi.org/https://doi.org/10.1016/j.jhydrol.2011.11.055, 2012.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2019.

Raghuwanshi, B. S. and Shukla, S.: SMOTE based class-specific extreme learning machine for imbalanced learning, Knowledge-Based Systems, 187, 104 814, https://doi.org/https://doi.org/10.1016/j.knosys.2019.06.022, 2020.

Rahmani-Rezaeieh, A., Mohammadi, M., and Danandeh Mehr, A.: Ensemble gene expression programming: a new approach for evolution of parsimonious streamflow forecasting model, Theoretical and Applied Climatology, 139, 549–564, https://doi.org/10.1007/s00704-019-02982-x, 2020.

Ramachandran, P., Zoph, B., and Le, Q. V.: Searching for Activation Functions, ArXiv, abs/1710.05941, https://doi.org/arXiv:1710.05941, 2018.

Rezaie-Balf, M. and Kisi, O.: New formulation for forecasting streamflow: evolutionary polynomial regression vs. extreme learning machine, Hydrology Research, 49, 939–953, https://doi.org/10.2166/nh.2017.283, 2017.

Richardson, W. B.: A Comparison of Detritus Processing Between Permanent and Intermittent Headwater Streams, Journal of Freshwater Ecology, 5, 341–357, https://doi.org/10.1080/02705060.1990.9665247, 1990.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., and Müller, M.: pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinformatics, 12, 77, https://doi.org/10.1186/1471-2105-12-77, 2011.

Rolls, R. J., Leigh, C., and Sheldon, F.: Mechanistic effects of low-flow hydrology on riverine ecosystems: ecological principles and consequences of alteration, Freshwater Science, 31, 1163–1186, https://doi.org/10.1899/12-002.1, 2012.

Rong, H.-J., Ong, Y.-S., Tan, A.-H., and Zhu, Z.: A fast pruned-extreme learning machine for classification problem, Neurocomputing, 72, 359–366, https://doi.org/https://doi.org/10.1016/j.neucom.2008.01.005, 2008.

Sazib, N., Bolten, J., and Mladenova, I.: Exploring Spatiotemporal Relations between Soil Moisture, Precipitation, and Streamflow for a Large Set of Watersheds Using Google Earth Engine, Water, 12, https://doi.org/10.3390/w12051371, 2020.

Seo, Y., Kwon, S., and Choi, Y.: Short-Term Water Demand Forecasting Model Combining Variational Mode Decomposition and Extreme Learning Machine, Hydrology, 5, https://doi.org/10.3390/hydrology5040054, 2018.

Shukla, S., Yadav, R. N., and Naktode, L.: Correlation Based Extreme Learning Machine, in: 2016 9th International Conference on Developments in eSystems Engineering (DeSE), pp. 268–272, https://doi.org/10.1109/DeSE.2016.41.

Steward, A. L., Marshall, J. C., Sheldon, F., Harch, B., Choy, S., Bunn, S. E., and Tockner, K.: Terrestrial invertebrates of dry river beds are not simply subsets of riparian assemblages, Aquatic Sciences, 73, 551, https://doi.org/10.1007/s00027-011-0217-4, 2011.

Su, S., Zhang, C., Han, K., and Tian, Y.: Greedy Hash Towards Fast Optimization for Accurate Hash Coding in CNN, pp. 806–815, http://papers.nips.cc/paper/7360-greedy-hash-towards-fast-optimization-for-accurate-hash-coding-in-cnn, 2018.

Thabtah, F., Hammoud, S., Kamalov, F., and Gonsalves, A.: Data imbalance in classification: Experimental evaluation, Information Sciences, 513, 429–441, https://doi.org/https://doi.org/10.1016/j.ins.2019.11.004, 2020.

Tooth, S.: Process, form and change in dryland rivers: a review of recent research, Earth-Science Reviews, 51, 67–107, https://doi.org/https://doi.org/10.1016/S0012-8252(00)00014-3, 2000.

Torgo, L.: Data Mining with R, learning with case studies, Chapman and Hall/CRC, http://www.dcc.fc.up.pt/~ltorgo/DataMiningWithR, 2010.

Tu, W. and Liu, H.: Zero-Inflated Data, pp. 1–7, American Cancer Society, https://doi.org/https://doi.org/10.1002/9781118445112.stat07451.pub2, 2016.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Uddameri, V.: Using statistical and artificial neural network models to forecast potentiometric levels at a deep well in South Texas, Environmental Geology, 51, 885–895, https://doi.org/10.1007/s00254-006-0452-5, 2007.

Uddameri, V., Singaraju, S., and Hernandez, E. A.: Is Standardized Precipitation Index (SPI) a Useful Indicator to Forecast Groundwater Droughts? — Insights from a Karst Aquifer, JAWRA Journal of the American Water Resources Association, 55, 70–88, https://doi.org/https://doi.org/10.1111/1752-1688.12698, 2019.

Van Ogtrop, F., Vervoort, R., Heller, G., Stasinopoulos, D., and Rigby, R.: Long-range forecasting of intermittent streamflow, Hydrology and Earth System Sciences, 15, 3343–3354, 2011.

Vander Vorste, R., Obedzinski, M., Nossaman Pierce, S., Carlson, S. M., and Grantham, T. E.: Refuges and ecological traps: Extreme drought threatens persistence of an endangered fish in intermittent streams, Global Change Biology, 26, 3834–3845, https://doi.org/https://doi.org/10.1111/gcb.15116, 2020.

Wagena, M. B., Goering, D., Collick, A. S., Bock, E., Fuka, D. R., Buda, A., and Easton, Z. M.: Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models, Environmental Modelling and Software, 126, 104 669, https://doi.org/https://doi.org/10.1016/j.envsoft.2020.104669, 2020.

Wang, L., Yang, B., Chen, Y., Zhang, X., and Orchard, J.: Improving Neural-Network Classifiers Using Nearest Neighbor Partitioning, IEEE Transactions on Neural Networks and Learning Systems, 28, 2255–2267, https://doi.org/10.1109/TNNLS.2016.2580570, 2017.

Wu, C., Luo, C., Xiong, N., Zhang, W., and Kim, T.: A Greedy Deep Learning Method for Medical Disease Analysis, IEEE Access, 6, 20 021–20 030, https://doi.org/10.1109/ACCESS.2018.2823979, 2018.

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., and Chau, K.-W.: An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction, Journal of Hydrology, 569, 387–408, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.11.069, 2019.

Yu, Q., van Heeswijk, M., Miche, Y., Nian, R., He, B., Séverin, E., and Lendasse, A.: Ensemble delta test-extreme learning machine (DT-ELM) for regression, Neurocomputing, 129, 153–158, https://doi.org/https://doi.org/10.1016/j.neucom.2013.08.041, 2014.

Zambrano-Bigiarini, M.: hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series, https://doi.org/10.5281/zenodo.840087, r package version 0.3-10, 2017.

Zawadzki, Z. and Kosinski, M.: FSelectorRcpp: 'Rcpp' Implementation of 'FSelector' Entropy-Based Feature Selection Algorithms with a Sparse Matrix Support, 1, https://CRAN.R-project.org/package=FSelectorRcpp, 2020.

Zeng, Y., Xu, X., Fang, Y., and Zhao, K.: Traffic sign recognition using deep convolutional networks and extreme learning machine, vol. 9242, pp. 272–280, Springer, https://doi.org/https://doi.org/10.1007/978-3-319-23989-7_28, 2015.

Zhang, C. D. and Xu, Y. L.: Comparative studies on damage identification with Tikhonov regularization and sparse regularization, Structural Control and Health Monitoring, 23, 560–579, https://doi.org/https://doi.org/10.1002/stc.1785, 2016.

Zhang, G., Li, Y., Cui, D., Mao, S., and Huang, G.-B.: R-ELMNet: Regularized extreme learning machine network, Neural Networks, 130, 49–59, https://doi.org/https://doi.org/10.1016/j.neunet.2020.06.009, 2020.

Zhang, K. and Luo, M.: Outlier-robust extreme learning machine for regression problems, Neurocomputing, 151, 1519–1527, https://doi.org/https://doi.org/10.1016/j.neucom.2014.09.022, 2015.

Zhang, R., Lan, Y., Huang, G., and Xu, Z.: Universal Approximation of Extreme Learning Machine With Adaptive Growth of Hidden Nodes, IEEE Transactions on Neural Networks and Learning Systems, 23, 365–371, https://doi.org/10.1109/TNNLS.2011.2178124, 2012.

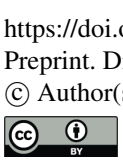
Zhou, Z., Chen, J., and Zhu, Z.: Regularization incremental extreme learning machine with random reduced kernel for regression, Neurocomputing, 321, 72–81, https://doi.org/https://doi.org/10.1016/j.neucom.2018.08.082, 2018.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Zimmer, M. A. and McGlynn, B. L.: Ephemeral and intermittent runoff generation processes in a low relief, highly weathered catchment, Water Resources Research, 53, 7055–7077, https://doi.org/https://doi.org/10.1002/2016WR019742, 2017.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021



Hydrology and Earth System Sciences
Discussions
Open Access
EGU

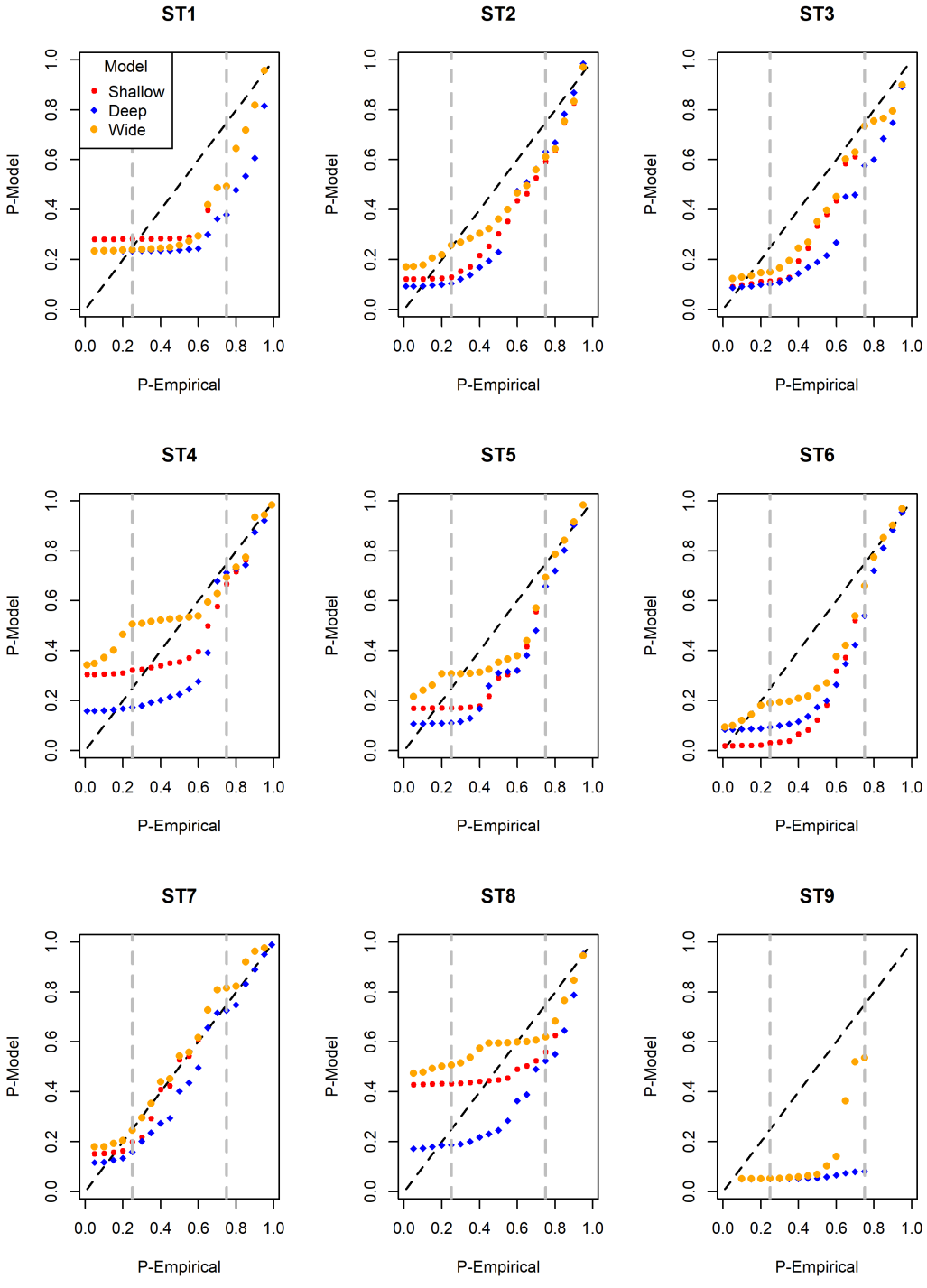


Figure 12. The probability- probability plot for the shallow, deep and wide models across the nine IRES of study over raw (not transformed) flow data.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions

Appendix A

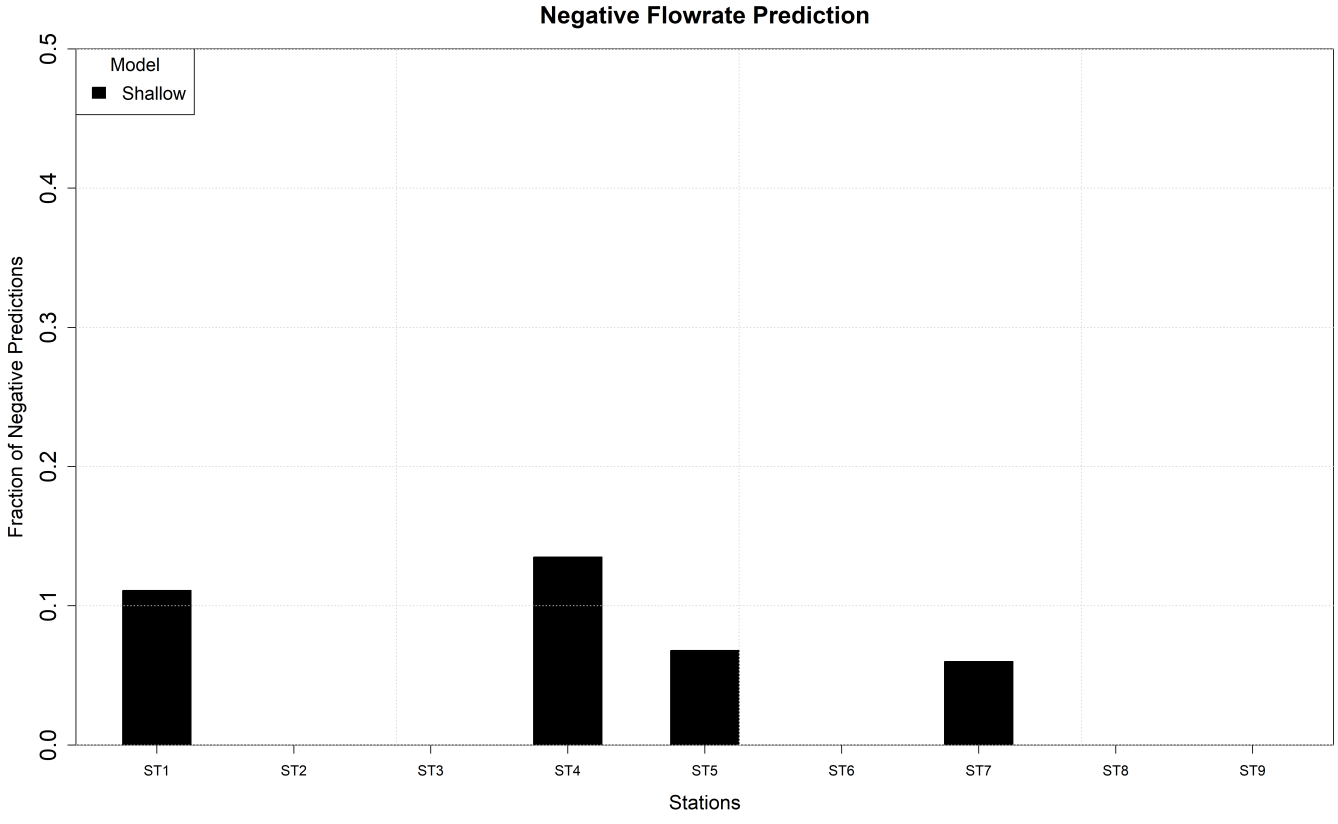


Figure A1. Fraction of negative flow predictions by the shallow, deep and wide models for the nine IRES of study.

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table A1. Contingency Table for Evaluation of Classifier Cell Predictions

Contingency Table		Predicted	
		0 (No-flow)	+ (Flow)
Observed	0 (No-Flow)	N_{00}	N_{0+}
	+ (Flow)	N_{+0}	N_{++}

https://doi.org/10.5194/hess-2021-176
Preprint. Discussion started: 14 April 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table A2. Transition Matrix for Flow and No-Flow States

Transition Matrix		Next State	
		0 (No Flow)	1 (Flow)
Current State	0 (No Flow)	p_{00}	p_{01}
	1 (Flow)	p_{10}	p_{11}