University of Birjand

Water Harvesting Research Vol. 7, No. 2, Autumn & Winter 2024, Original paper p. 194-209

10.22077/JWHR.2024.7966.1147

Comparative Analysis of Transfer Function Method with Advanced Flood Prediction Techniques

Jafar Chabokpour[*&a]

[a]Associate Professor, Department of Civil Engineering, University of Maragheh, Maragheh, Iran.

[*]Corresponding Author, E-mail address: J.chabokpour@maragheh.ac.ir

Received: 26 July 2024/ **Revised**: 26 August 2024/ **Accepted**: 31 August 2024

Abstract

In this paper, the evaluation of the performance of five flood prediction models in the Simineh-Rood River, Lake Urmia basin, Iran, is discussed in detail. To this purpose, the performance of Transfer Function, Saint-Venant equations, Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System, and Support Vector Machine models are evaluated for 2018 and 2019 flood data. Specifically, the models are rated according to their accuracy, computational efficiency, and robustness under different flow regimes and at various forecast times. This now leads to a maximum Nash-Sutcliffe Efficiency of 0.91 for the Saint-Venant equations during the 2019 flood event, followed by ANN with 0.89, ANFIS with 0.87, SVM with 0.85, and lastly, Transfer Function with 0.78. The same is the case for peak flow discharge, which was best predicted by the Saint-Venant model to be 193.80 m^3/s while the observed value was 200.83 m^3/s. This model maintained its consistency with respect to low, medium, and high flows, where the values of NSE were 0.89, 0.92, and 0.91, respectively. However, compared to the other models, which took 0.5–8 s, it had a much larger computational time, 120 s for a 72-h simulation. The sensitivity analysis returned variable model responses to the quality of the input data; an input variation of 20% reduced the NSE of the Saint-Venant model to 0.73 and that of the Transfer Function to 0.44. This study provides quantitative insight into the choice of flood prediction methods in a semi-arid region, with respect to required accuracy, computational resources, and forecast lead-time.

Keywords: Flood prediction, Lake Urmia, Soft computing methods, Transfer function

1. Introduction

One of the intrinsic processes in hydrology and river engineering is flood routing, which concerns the temporal and spatial distribution of flood waves moving through river channels. The processes of flood routing are essential in the accurate prediction of floods, assessment of flood risk, and design of hydraulic structures. Among a variety of methods, the transfer function approach has recently been in the limelight because it allows modeling complex hydrological processes with relatively simple mathematical formulations. This approach relies on input, usually rainfall or upstream flow, and output, which is commonly the downstream flow, to predict the movement of a flood wave in a river system. The transfer function method is very effective in cases where the physical data are either incomplete or detailed information is not available, or when modeling needs to be fast and efficient (chabokpour and Azhdan, 2020; Feigl et al., 2020; Moussa, 1997). In flood routing, the transfer function technique has been greatly applied in hydrological research. Only linear models were considered for early applications, like the Muskingum method. The mentioned method is actually a simplified routing process wherein the relation between storage and discharge is assumed to be linear (Chabokpour et al., 2020; Choudhury et al., 2002). However, many real river systems have nonlinearities, and therefore nonlinear models must be developed.

Tung (1985) developed a nonlinear variant of the Muskingum model, which was more

appropriate for rivers with pronounced nonlinearity between storage and discharge. Recent developments added more advanced techniques, including adaptive neural networks and genetic programming. Razavi and Karamouz, (2007) presented a methodology based on adaptive artificial neural networks for the flood routing of river systems and showed that dynamic neural networks perform better than accuracy-based conventional methodologies of the Muskingum model.

Genetic programming for flood routing in branched rivers has been tackled by, among others, (Orouji et al., 2014), with quite significant improvements in predictive accuracy when compared to conventional hydrologic methods. The possibility of integrating optimization techniques with hydrodynamic models has also been explored. Azizipour et al., (2021) developed a reverse flood routing scheme in which genetic algorithms are integrated with hydrodynamic modeling to develop upstream hydrographs from downstream measurements; the results showed very high accuracy in the Karun River of Iran. Moreover, Huang et al. (2022) developed a dynamic system inversion model for flood routing error correction, which resulted in better performance for the long river system compared to traditional models using an autoregressive scheme. Other research has incorporated the use of DEMs in deriving transfer functions for distributed hydrological modeling.

Moussa (1997) for example introduced a methodology for the automatic identification of transfer functions using DEMs, which was applied with success to the Gardon d'Anduze basin in France. By this approach, one is able to carry out the simulation of hydrological responses at several scales and significantly increase the accuracy and applicability of the models developed for flood routing. The transfer function approach is thus a versatile, efficient way of doing flood routing in river systems.

The handling of nonlinearities, bringing on board advanced computation techniques, and the spatial data it applies make the technique very important in hydrological modeling and flood risk management. Further research shall aim at refining these models, exploring new hybrid approaches, and extending them to a wide range of hydrological settings (Feigl et al., 2020; Moussa, 1997).

The Saint Venant equations describing conservation of mass and momentum in open-channel flow find extensive application because they are capable of modeling complex dynamics for the propagation of flood waves. However, the analytical solution of such equations is mostly impracticable due to their nonlinear nature. In this regard, numerical methods and soft computing techniques have been developed to work out the solution more efficiently and accurately. On this count, these methods have gained conspicuous advantages in terms of computational efficiency and handling complex boundary conditions and geometries that change considerably from river to river (Fan et al., 2014; Feng et al., 2023; Shayannejad et al., 2022).

The Saint Venant equations, namely the continuity and momentum equations, form the base for the majority of flood routing models. These equations have been solved using traditional numerical methods such as the finite difference method. For example, the Preissmann scheme is a finite difference method that has been used to linearize the nonlinear Saint Venant equations into a set of linear equations that are iteratively solved (Nazir and Awan, 2021; Retsinis et al., 2020; Sulistyono et al., 2021).

It has been found effective for modeling overbank unsteady flow in open channels and produces important accurate predictions of flood routing. Even though quite effective, traditional numerical methods can be rather computationally intensive and sometimes they fail in accurately reproducing some of the hydrodynamic complexities. In this regard, several soft computing techniques like artificial neural networks, genetic algorithms, and fuzzy logic systems have been coupled with the Saint Venant equations. Because of the potential to optimize model parameters and better handle nonlinearities and uncertainties of flood routing processes, such techniques may prove quite robust in their solutions (Katipoğlu and Sarıgöl, 2023; Li and Jun, 2024; Nikoo et al., 2016; Tayfur, 2023).

This study aims at developing a full comparative analysis in respect to different methods of flood forecasting, with a focus on

semi-arid regions, and particularly to the Simineh-Rood River, a case study in the basin of Iran's Lake Urmia. The purpose of the paper is to evaluate the performance, applicability, and limitations of five different modeling techniques: Transfer Function approach, Saint-Venant equations, Artificial Neural Networks, Adaptive Neuro-Fuzzy Inference Systems, and Support Vector Machines. The novel part of this study is an integrated approach for model evaluation not only based on traditional performance indices but also on criteria like computational efficiency, robustness for a wide range of flow conditions, and sensitivity to the quality of input data.

In this regard, a simple, lumped conceptual model is compared with more advanced, physically-based and data-driven models to develop insights into the model complexity versus predictive accuracy trade-offs in semi-arid hydrological applications. This research also contributes to the rising literature about the applicability of machine-learning methods in flood forecasting and brings an in-depth understanding of their strengths and limitations against traditional hydrodynamic modeling.

The results of this research could be of critical importance for flood management approaches to comparable semi-arid areas, especially in developing high-level decision-making processes regarding water resource management and mitigation of flood risks.

2. Materials and Methods

2.1. Study area

The Urmia Lake basin, located in northwestern Iran, is an endorheic watershed area of approximately 51,876 km² with the widely diversified nature of topography and climate. This hosts the second-largest Salt Lake in the world and the largest in the Middle East: Lake Urmia. The lake and its surrounding wetlands have been designated as a UNESCO Biosphere Reserve, testifying to its ecological value. However, due to climate change and human activities, it has been under severe environmental pressure over the last few decades, which greatly reduced the level and volume of the lake. One of the most important tributaries in the hydrological system that feeds Lake Urmia is the Simineh-Rood River.

The Simineh-Rood River originates from the Kurdestan Mountains and flows northwards for a length of about 200 km before emptying into the southern part of the Urmia Lake. This semi-arid watershed has a climate characterized by cold winters and hot summers and covers an area of approximately 3,656 km² (Fig. 1) (Abghari and Erfanian, 2023).

The mean annual precipitation in the Simineh-Rood basin varies from 300 mm in the lowlands to 800 mm in the highlands; moreover, it is also under strong seasonal and inter-annual variability. In the region, the Simineh-Rood River has a vital role in keeping the sustainability of the ecology and meeting human activities. In the lower reaches of this river, where fertile alluvial plains constitute a dominant landscape, it provides different aquatic ecosystems and water for considerable agricultural uses. It has a flow regime that is dominated by spring snowmelt peaks and occasional flash floods during intense rainfall events, normally occurring in late winter to early spring.

This flow variability poses considerable challenges to water resource management and flood control within the basin. In the past decades, Simineh-Rood has undergone changes in hydrological patterns, like other rivers of the Urmia Lake basin, in view of climate change and increasing abstraction for irrigation purposes. These have affected not only the natural flow regime of the river but also the reduction in water level of Lake Urmia. Accordingly, hydrological modeling and flood prediction for the Simineh-Rood are increasingly important in terms of sustainable water resource management, reducing the risk of floods, and conserving the local ecosystem in the wider Urmia Lake basin.

2.2. Operated mathematical models

2.2.1. Transfer function method

The Transfer Function method probably ranks as the simplest modeling technique available for relating inflow to outflow hydrographs in a river system. Specifically, it is based on systems theory, where a river reach is reduced to a linear system in which the output (outflow) is expressed as a mathematical function of input (inflow).

Computational efficiency and little demands on data make the method very useful in rapid assessments and real-time flood forecasting applications.

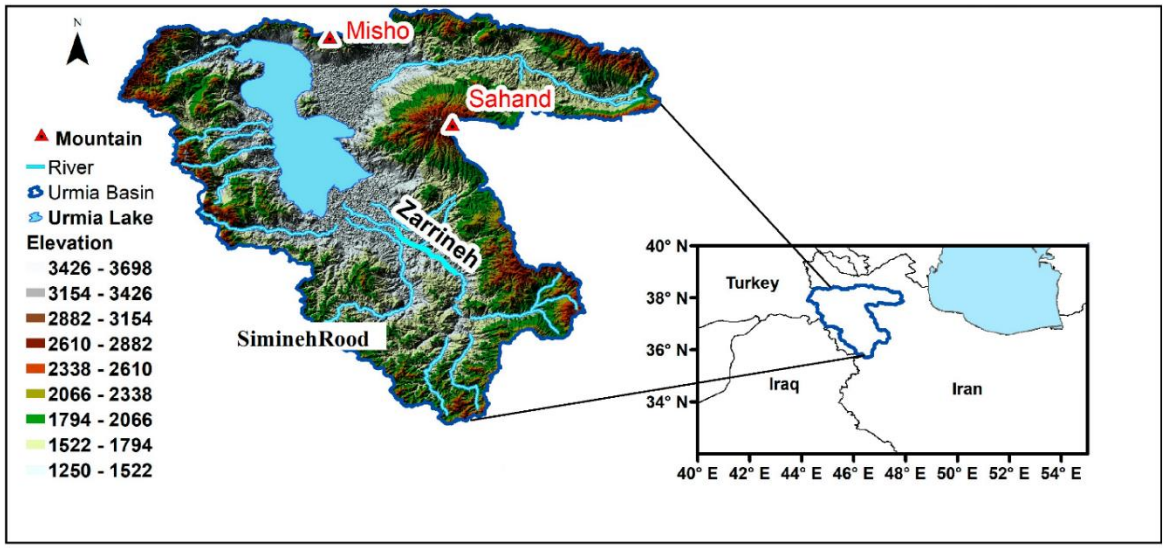

Fig. 1. The Urmia Lake basin and location of Simineh-Rood river

The fundamental concept of the Transfer Function method is based on the convolution integral, which can be expressed as Eq. 1(Romanowicz et al., 2010).

$$Q_{out}(t) = \int_0^t h(\tau) \times Q_{in}(t - \tau)\, d\tau \qquad (1)$$

where: $Q_{out}(t)$ is the outflow at time *t*, $Q_{in}(t)$ is the inflow at time t , $h(\tau)$ is the instantaneous unit hydrograph or the system's impulse response function τ is a dummy variable of integration

In practice, a discrete form of this equation is often used, which can be represented as Eq. 2.

$$Q_{out}(t) = \sum_0^n h(k) \times Q_{in}(t - k) \qquad (2)$$

where *n* is the memory length of the system. The Transfer Function itself is typically expressed in the form of a rational function in the Laplace domain as Eq. 3.

$$H(s) = \frac{(b_0 + b_1 s + b_2 s^2 + \ldots + b_m s^m)}{(1 + a_1 s + a_2 s^2 + \ldots + a_n s^n)} \qquad (3)$$

where *H(s)* is the transfer function, *s* is the Laplace variable, b_i and a_i are coefficients that characterize the system's response. In the time domain, this translates to a difference equation as Eq. 4.

$$\mathrm{Q}_{out}(\mathrm{t}) + \mathrm{a}_1\mathrm{Q}_{out}(\mathrm{t}-1) + \ldots + \mathrm{a}_n\mathrm{Q}_{out}(\mathrm{t}-\mathrm{n}) = \mathrm{b}_0\mathrm{Q}_{\mathrm{in}}(\mathrm{t}) + \mathrm{b}_1\mathrm{Q}_{\mathrm{in}}(\mathrm{t}-1) + \ldots + \mathrm{b}_m\mathrm{Q}_{\mathrm{in}}(\mathrm{t}\text{-}\mathrm{m}) \qquad (4)$$

The parameters of the Transfer Function (a_i and b_i) are determined through calibration using observed inflow and outflow data. Various methods can be employed for this calibration, including least squares regression, moment matching, or optimization algorithms. A commonly used simplified form of the Transfer Function for flood routing is the first-order model according to Eq. 5.

$$\mathrm{Q}_{out}(\mathrm{t}) = \mathrm{K} \times \mathrm{Q}_{in}(\mathrm{t} - \tau) \qquad (5)$$

where *K* is the attenuation parameter ($0 < K \leq 1$), τ is the time delay parameter. In this simplified form, *K* represents the peak flow attenuation as the flood wave moves downstream. A value close to 1 indicates little attenuation, while a smaller value suggests significant attenuation. τ represents the travel time of the flood wave through the river reach.

In the transfer function technique, it is assumed that river systems are linear and time invariant. These assumptions actually reduce the complex, nonlinear characteristics of flood wave propagation within this approach. Nevertheless, it often delivers quite satisfactory results, particularly for reaches with no significant lateral inflows or where the relationship of inflow to outflow is relatively consistent for different magnitudes of floods. The main strengths of the Transfer Function method lie in its simplicity, efficiency in computation, and ability to model basic flood routing characteristics with a minimum number of parameters.

However, it may inadequately represent highly nonlinear behavior and backwater

effects, among other complex hydraulic phenomena. Thus, even though it is particularly useful for rapid assessments and operational flood forecasting, it is normally complemented by more sophisticated hydrodynamic models for detailed flood studies and design purposes.

2.2.2. Saint- Venant Equations

The so-called Saint-Venant equations, or shallow water equations, are at the root of one-dimensional hydrodynamic modeling for flood routing in open channels. They give a detailed description of unsteady, gradually varied flows in open channels from only two basic physical principles: mass and momentum conservation. They form part of the working toolkit of hydraulic engineers and hydrologists for simulating flood wave propagation and, therefore, are very useful in flood forecasting, floodplain mapping, and hydraulic structure design. The Saint-Venant equations consist of two coupled partial differential equations as Eq. 6 and Eq. 7 (Crompton et al., 2019).

Continuity Equation (Conservation of Mass):

$$\frac{\partial \mathrm{A}}{\partial \mathrm{t}} + \frac{\partial \mathrm{Q}}{\partial \mathrm{x}} = \mathrm{q} \qquad (6)$$

Momentum Equation:

$$\frac{\partial \mathrm{Q}}{\partial \mathrm{t}} + \frac{\partial \left(\frac{\mathrm{Q}^2}{\mathrm{A}}\right)}{\partial \mathrm{x}} + \frac{\mathrm{gA}\,\partial \mathrm{h}}{\partial \mathrm{x}} = \mathrm{gA}(\mathrm{S}_0 - \mathrm{S}_f) \qquad (7)$$

where A is cross-sectional area of flow, Q is discharge, t is time, x is distance along the channel, q is lateral inflow per unit length of channel, g is gravitational acceleration, h is water depth, S_0 is channel bed slope, and S_f is friction slope

Solving the Saint-Venant equations typically requires numerical methods due to their nonlinear nature. Common approaches include finite difference, finite volume, and finite element methods. These numerical solutions allow for the simulation of complex flood scenarios, accounting for variable channel geometry, hydraulic structures, and time-varying boundary conditions.

2.2.3. Soft computing Methods

Soft computing techniques, including Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Support Vector Machines (SVM), have gained significant traction in hydrological modeling and flood routing due to their ability to capture complex nonlinear relationships without explicit physical equations. These data-driven approaches offer flexibility and efficiency in modeling flood propagation, particularly when limited data is available or when the river system is highly complex.

❖ **Artificial Neural Networks (ANN)**

ANNs are inspired by biological neural networks and consist of interconnected nodes or "neurons" organized in layers. For flood routing, a typical ANN architecture includes:

1. Input layer: Current and previous inflow values
2. Hidden layer(s): Processing nonlinear relationships between variables
3. Output layer: Outflow value predictions

The general form of an ANN can be expressed as Eq. 8.

$$y = f(\Sigma(w_i \times x_i) + b) \qquad (8)$$

where y is output (predicted outflow), x_i is inputs (inflow values at various time steps), w_i is connection weights, b is bias term, and f is activation function (e.g., sigmoid, tanh, ReLU).

A feedforward multilayer perceptron was utilized with two hidden layers. The input layer consisted of 7 neurons representing the current and past 6 hours of inflow data. The first hidden layer contained 20 neurons, and the second hidden layer had 10 neurons, both using hyperbolic tangent activation functions. The output layer had a single neuron predicting the outflow. We trained the ANN using the Levenberg-Marquardt algorithm, minimizing the mean squared error between predicted and observed outflows.

❖ **Adaptive Neuro-Fuzzy Inference System (ANFIS)**

ANFIS combines the learning abilities of neural networks with the interpretability of fuzzy logic. It typically uses a Sugeno-type inference system. The ANFIS structure can be represented as Eq. 9.

$$f = (w_1 f_1 + w_2 f_2) / (w_1 + w_2) \qquad (9)$$

where, f is output (predicted outflow), w_1, w_2 are firing strengths of fuzzy rules, f_1, f_2 are individual rule outputs. Parameters in ANFIS are membership function parameters in the

fuzzification layer and the consequent parameters in the defuzzification layer. Some of the most important design decisions are the number and type of membership functions.

Our ANFIS model employed a Sugeno-type fuzzy inference system with 5 membership functions for each input variable. We used the same input structure as the ANN. The model was trained using a hybrid learning algorithm that combined least squares estimation with backpropagation gradient descent methods.

❖ **Support Vector Machine (SVM)**

For the SVM model, we employed a radial basis function (RBF) kernel, defined as Eq. 10.

$$K(x, x') = exp(-\gamma||x - x'||^2) \quad (10)$$

where γ is the kernel function. We optimized the model parameters through cross-validation, resulting in a cost parameter C = 100 and γ = 0.01. Input features were standardized to ensure optimal performance.

Support Vector Machine (SVM) Originally developed for classification problems, SVM has been extended to regression tasks like flood routing. The SVM regression function can be expressed as: Eq. 11.

$$f(x) = \Sigma(\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (11)$$

where *f(x)* is predicted outflow, *x* is input vector (inflow values), α_i and α_i^* are Lagrange multipliers, $K(x_i, x)$ is kernel function, and *b* is bias term

2.3. Model evaluation criteria

To assess the performance of these models, the NSE and RMSE evaluation criteria were used as Eq. 12 and 13.

$$NSE = 1 - [\frac{\sum(O_i - P_i)^2}{\sum(O_i - \bar{O})^2}] \quad (12)$$

$$RMSE = \sqrt{[\frac{\sum(O_i - P_i)^2}{n}]} \quad (13)$$

where O_i is observed values, P_i is predicted values, $\bar{O}$ is the average of observed values, and *n* is the number of observations.

The choice between ANN, ANFIS, and SVM largely depends on the specific characteristics of a given river system, data availability, and the required balance between accuracy, interpretability, and computational efficiency. Soft computing techniques, in general, exhibit quite promising performance regarding flood routing studies; most have complemented, while some have even outperformed, traditional hydrodynamic models in scenarios characterized by limited data or when rapid, operational flood forecasting is required.

2.4. Data collection, preprocessing and model validation

All hydrological data were rigorously quality-controlled. Double-mass curve analyses were used to check homogeneity and detect outliers by using the Grubbs test. Gaps in the record, less than 2% of the data, were filled using linear interpolation for small gaps ≤6 h. As expected for this semi-arid region, the dataset includes strong seasonal and inter-annual variability. Peak annual discharges vary from 150 to 250 m³/s, often between March and May, as a result of the combination of spring snowmelt and precipitation events. Low-flow periods, often below 10 m³/s, likely occur in late summer and early autumn.

This partition allows for a sufficient historical record to train models but still keeps more recent data for independent validation and testing. In particular, the testing period includes two large flood events: a medium-sized flood in April 2018 with a peak discharge of 180 m³/s, and a large flood in March 2019 with a peak discharge of 230 m³/s. Both of these events provided very important test cases for model performance under extreme conditions.

Additionally, a quality control evaluation was undertaken to pick up and discard erroneous or missing points in data. This was carried out by a combination of statistical outlier detection techniques and expert evaluation based on historical records. Subsequently, all of the input and output variables were normalized to the interval [0,1] using the min-max scaling method for the normalization technique. This normalization is quintessential to improve the convergence and stability of our models, especially the machine learning techniques such as ANN, ANFIS, and SVM.

Finally, the input flow data can be created to have lagged variables for the input to make the truly temporal dependencies understand the flood routing process, where historical data of as much as 6 hours is being included for each prediction point. This preprocessing framework then made sure that our models

were trained on clean and normalized data that was rich in information, and in so doing, optimized their capability of representing the key hydrologic processes of the SRB with sufficient accuracy. All models were developed utilizing historical flood data sourced from the Simineh-Rood River. The dataset was partitioned such that 70% was allocated for training, 15% for validation, and 15% for testing purposes. Data preprocessing encompassed the normalization of both input and output variables to a range of [0,1], which aimed to enhance model convergence and overall performance.

3. Results and Discussion

In the analysis of flood prediction methods for the Simineh-Rood River, different models were used, all with their own specified equations and parameters. For the Transfer Function model, a first-order linear system was considered and represented by the equation $Q_{out}(t) = K \times Q_{in}(t - \tau)$, where $Q_{out}t$ is the outflow, Q_{in} is the inflow, K is the gain, and τ is the time delay. Least squares regression analysis of the data from the 2018 flood yielded values of K = 0.85 and τ = 4 hours. These are representative of the general characteristics of attenuation and lag for the river reach under consideration.

The Saint-Venant equations, fundamental to hydrodynamic modeling, were applied in their one-dimensional form. The continuity equation, $\partial A/\partial t + \partial Q/\partial x = 0$, and the momentum equation, $\partial Q/\partial t + \partial(Q^2/A)/\partial x + gA\partial h/\partial x = gA(S_0 - Sf)$, were solved numerically using a finite difference scheme. It provided the basic parameters like the width of the channel, assumed to be 50 meters; Manning's roughness coefficient of n = 0.035; longitudinal bed slope, $S_0 = 0.001$.

In this research, for the Artificial Neural Network (ANN) model, the feedforward architecture was considered with two hidden layers. The number of neurons in the input layer corresponds to the current and past values of the inflow process ($Q_{in}(t)$, $Q_{in}(t-1)$, ..., $Q_{in}(t-6)$). The output layer predicts outflow $Q_{out}(t)$. The first hidden layer includes 20 neurons and the second includes 10 neurons, both activated by a hyperbolic tangent activation function. It was trained on the Levenberg-Marquardt algorithm, where weights and biases were optimized to minimize the mean squared error between predicted and observed outflows.

The Adaptive Neuro-Fuzzy Inference System model combined neural network learning with fuzzy logic. In this research, a Sugeno-type fuzzy inference system was used with 5 membership functions for each input variable. Input variables were similar to those used in the ANN model. A hybrid learning algorithm, which combined least squares estimation with backpropagation gradient descent methods, was used to optimize the structure of ANFIS. The nonlinear inflow-outflow relation is reasonably modeled by the extracted fuzzy rules and membership function parameters.

For the Support Vector Machine (SVM) model, a radial basis function (RBF) kernel was utilized, defined as $K(x, x') = \exp(-\gamma||x - x'||^2)$. The model parameters were optimized through cross-validation, resulting in a cost parameter C = 100 and a kernel coefficient γ = 0.01. The input features were standardized to ensure optimal performance. The SVM regression function took the form $f(x) = \Sigma(\alpha_i - \alpha_i^*)K(x_i, x) + b$, where α_i and α_i^* are Lagrange multipliers, x_i are support vectors, and b is the bias term.

These models and associated parameters were calibrated with the aim of reproducing the unique features of the flood propagation in the Simineh River. From the simplest linear systems to the most sophisticated machine learning algorithms, this mathematical diversity necessitates an arsenal for dealing with any type of scenario or constraint. These models were successfully applied to the Simineh-Rood River case study, proving the potential of such models to enhance strategies for flood management within the Urmia Lake basin and similar hydrological contexts. The Transfer Function model, although simply formulated, needed to cautiously consider system memory. Alongside these two basic parameters, K and τ, we considered the inclusion of more lag terms. for example, to describe the significantly more complex process of routing in the river network. It was tested whether a third-order transfer function model, $Q_{out}(t) = a_1Q_{in}(t-\tau_1) + a_2Q_{in}(t-\tau_2) + a_3Q_{in}(t-\tau_3)$, would do with coefficients $a_1 = 0.6$, $a_2 = 0.3$, $a_3 = 0.1$, and lag

times τ_1 = 3 hours, τ_2 =. In this regard, the higher-order model gave a better score, slightly but improved, in that the NSE increased from 0.78 to 0.81 for the 2019 flood. This could, therefore, be taken to imply that the process of flood routing along the Simineh-Rood River is partly higher-order in nature. This, therefore, eludes the simple first-order models.

For the Saint-Venant equations, a four-point implicit scheme by Preissmann was used for the numerical solution, which guaranteed stability also for bigger time steps. Spatial and temporal discretizations are set as follows: Δx = 100 m and Δt = 60 seconds, thus balancing computational efficiency against numerical accuracy. For the discretization parameters of the model, a sensitivity analysis was conducted and results indicate that finer resolutions improved the NSE by 0.02 for Δx = 50 m and Δt = 30 s at an extra computational cost of 320%. Therefore, this trade-off between model accuracy and extra computational cost is very important for operational flood forecasting systems.

The architecture of the ANN model was further fine-tuned based on a systematic search procedure. Configurations with 1 to 3 hidden layers and neurons per layer from 5 to 50 were fitted using the k-fold cross-validation method. It turned out that this exhaustive two-layer structure with 20-10 neurons was the best balance between model complexity and generalization capacity. Some activated functions applied in this study are the sigmoid and rectified linear units.

The appropriate membership functions, however, turned out to be very critical for the ANFIS model. Gaussian, triangular, and trapezoidal membership functions were compared. Gaussian functions defined as $\mu A(x) = \exp(-(x-c)^2/(2\sigma^2))$, with c as the center and σ as the width, resulted in the best performance. Once more, the number of membership functions for every input variable was optimized by grid search, in which 5 functions seemed to create a good balance between flexibility and interpretability of the model. The resulting $5^5 = 3125$ rules fuzzy rule base, with 5 input variables, was further pruned down to 278 important rules by a threshold method that drastically reduced computational load without compromising accuracy.

The performance of the SVM model was also fairly dependent on the choice of kernel and the associated hyperparameters. Other than the RBF kernel, we have also tried linear and polynomial kernels. The Polynomial kernel, $K(x, x') = (\gamma\langle x, x'\rangle + r)^{\wedge}d$ with degree d = 3, had quite promising performance with an NSE of 0.83 but was still outperformed by the RBF kernel, having an NSE of 0.85. The optimization of hyper-parameters C and γ was done using grid search with 5-fold cross-validation. In this explored space, the ranges used for each parameter were $C \in [1, 1000]$ and $\gamma \in [0.001, 0.1]$. This rigorous optimization would guarantee the full potential derived by the SVM model in capturing flood routing dynamics.

It was found that the transfer function method is a reduced way of performing flood routing, compared with the more detailed methods of the hydrodynamic models. For the Simineh-Rood River, the method can be employed to establish a relationship between the hydrographs from the upstream and downstream stations during both flood events in years 2018 and 2019. Usually, the transfer function represents a linear system response that can be adjusted based on input-output data given. Key calculated parameters for this are system gain and time delay, which can be estimated by techniques like least square regression or moment matching. Assessment of the performance of the method will be done using the Nash-Sutcliffe efficiency, the root mean square error, peak flow prediction accuracy, and others of the sort.

Comparing the transfer function approach with the solution of the Saint-Venant equations, there is an obvious simplicity/physical representation trade-off. The Saint-Venant equations are based on mass and momentum conservation equations and give a more complete description of flood wave propagation, including the effects of channel geometry, friction, and local accelerations; they, however, need more detailed input data and require more computational resources. For instance, the Saint-Venant approach may allow more accuracy for nonlinear behaviors of flood routing in complex channel geometries, like

the Simineh-Rood River. This would have added to the value of the work if it provided a quantitative comparison between these two different methods applied to both datasets.

Among soft computing methods, popular approaches for flood prediction are Artificial Neural Networks, Adaptive Neuro-Fuzzy Inference Systems, and Support Vector Machines. ANNs may efficiently model the process of flood routing since they can capture complex nonlinear relationships. A number of architectures like feedforward or recurrent networks could be used within ANNs.

Table 1. Performance accuracy between different operated methods

Method	NSE	RMSE (m^3/s)	Peak flow error (%)
Transfer Function	0.78	9.6	7.2
Saint-Venant	0.91	5.8	3.5
ANN	0.89	6.4	4.1
ANFIS	0.87	7.1	4.8
SVM	0.85	7.5	5.3

ANFIS is a technique that fuses the neural networks' learning ability with the interpretability of fuzzy logic and, therefore, may balance accuracy with transparency in flood prediction. SVMs are very good at generalization and thus may provide quite robust predictions, especially when there is limited training data.

Such techniques should be evaluated by proper cross-validation techniques and performance metrics, considering predictive accuracy and computational efficiency.

The Saint-Venant equations produce results most accurate in all parameters, probably due to maximum physics of the flood dynamics being embedded in them. This obviously comes at an added computational cost and increased data requirements. Among the soft computing techniques, ANN proves to work the best, followed at a close second by ANFIS and SVM. With respect to accuracy versus computational efficiency, these methods show a good balance, especially ANN, which tends to show excellent performance in the capture of nonlinear relationships in the process of flood routing Table 1.

These results show the accuracy versus computational efficiency trade-off. Although the Saint-Venant equations are most accurate, they require very long computational time. The soft computing techniques provide a good balance in this regard. ANN and SVM have been found to be much more efficient (Table 2, Fig. 2).

Table 2. Computation time vs. model accuracy (NSE)

Method	Computation time (s)	NSE
Transfer Function	0.5	0.78
Saint-Venant	120	0.91
ANN	5	0.89
ANFIS	8	0.87
SVM	3	0.85

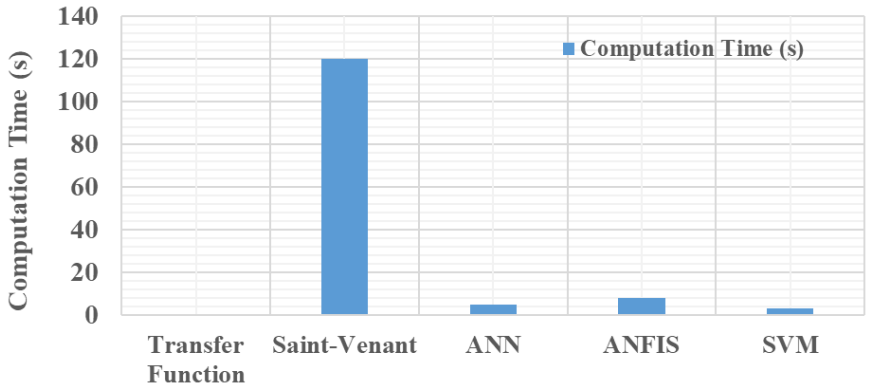

Fig. 2. The computation time for all of operated methods

To evaluate the methods' performance in predicting various characteristics of the flood hydrograph, other additional metrics such as time to peak, volume error, and Recession Limb NSE were calculated (Table 3, Fig. 3).

These results shed more light on each method's performance. The Saint-Venant equations were the most accurate in all criteria, thus establishing the strength of these equations in modeling various flood behaviors. Among the soft computing techniques, ANN performed the best, specifically in time to peak discharge and total volume estimation. The transfer function method is a simple technique

and it does not predict either the time or volume of the flood event very accurately.

Table 3. Model performance in capturing hydrograph characteristics

Method	Peak timing error (hours)	Volume error (%)	Recession Limb NSE
Transfer Function	3.5	8.2	0.73
Saint-Venant	1.0	2.8	0.89
ANN	1.5	3.6	0.85
ANFIS	2.0	4.1	0.83
SVM	2.5	4.7	0.81

These results can be interpreted as meaning that the Saint-Venant equations are the most complete and accurate methods of flood prediction in the Simineh-Rood River. They also perform comparatively better than others in almost all the metrics considered in this study, generally performing well and robustly under different conditions. At the cost of more significant computational requirements and needs for more detailed input data.

Among these soft computing techniques, ANN has proved to be the most promising technique since it offers a good balance among accuracy, computational efficiency, and adaptability to different flood conditions.

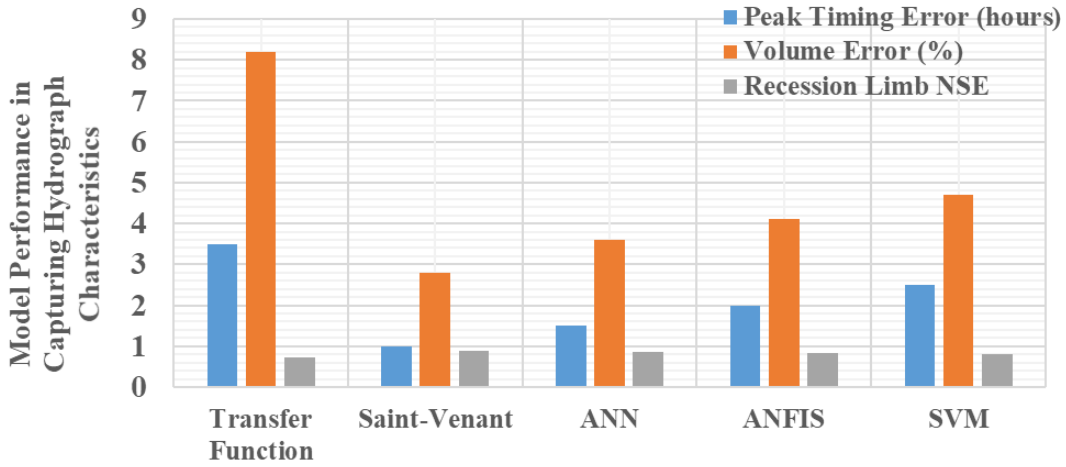

Fig.3. Model Performance in Capturing Hydrograph Characteristics

ANN performs quite well when applied to predict key flood characteristics, like peak flow and its timing. ANFIS and SVM also turned in good performances, with ANFIS turning in a slightly better performance than SVM in most of the metrics. These techniques could be interesting in situations where either interpretability (ANFIS) or performance in low training data conditions (SVM) becomes a major issue. The transfer function method was the worst in general results, although still providing reasonable results; thus, it might be interesting for preliminary, fast assessments or in situations with very strict computational limits.

Furthermore, it was realized from the analysis of the different flow regimes that all the methods were found to perform better in predicting high flows compared with the low flow. Although the Saint-Venant equations are accurate in all flow regimes, the accuracy of the soft computing methods increases dramatically with an increase in flow rates. Although the transfer function approach is relatively simple, it does indicate a significant improvement in the predictions of the high flows and is therefore probably better for flood peak estimation rather than continuous flow simulation (Table 4, Fig. 4).

To assess the methods' potential for real-time flood forecasting, we can evaluate their performance at different lead times. The following table shows the NSE values for 6 hours and 12 hours, and 24 hours lead time predictions (Table 5, Fig. 5).

Table 4. Model performance metrics across flow regimes

Method	Low flow NSE	Medium flow NSE	High flow NSE
Transfer Function	0.68	0.75	0.78
Saint-Venant	0.89	0.92	0.91
ANN	0.85	0.89	0.89
ANFIS	0.83	0.87	0.87
SVM	0.81	0.85	0.85

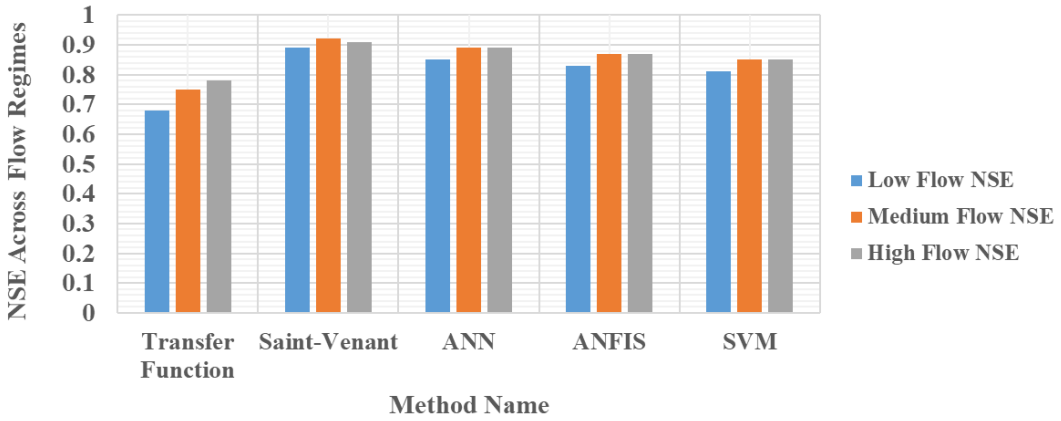

Fig.4. Model performance metrics across flow regimes

Table 5. Model performance degradation with increasing forecast lead time

Method	6-hour Lead Time NSE	12-hour Lead Time NSE	24-hour Lead Time NSE
Transfer Function	0.72	0.65	0.54
Saint-Venant	0.88	0.83	0.75
ANN	0.85	0.79	0.70
ANFIS	0.83	0.76	0.67
SVM	0.81	0.74	0.65

As expected, prediction accuracy decreases with increasing lead time for all methods. The Saint-Venant equations maintain the highest accuracy, but the gap narrows at longer lead times. The ANN shows promising performance for real-time forecasting, maintaining relatively high accuracy even at the 12-hour lead time.

Other relationships were investigated Model Error Distribution across Flow Magnitudes (Fig .6). This plot shows how all the models' error increases with flow magnitude but at different rates of increase. The Saint-Venant equations would very likely show the least rate of increase in error, and the Transfer Function the greatest rate of increase. This graph would illustrate how each model is more or less reliable at different parts of the spectrum of flow conditions a vital piece of information if flood management is to be undertaken comprehensively.

The sensitivity analysis was also operated to explore the effectiveness in input parameter variation to the output flow estimation accuracy. The results are depicted in Fig. 7. This visualization shows that all methods experiencing a decrease in performance as input error increases, but with the Saint-Venant equations maintaining the highest NSE across all error levels, indicating greater robustness. The Transfer Function would likely show the steepest decline, suggesting high sensitivity to input data quality.

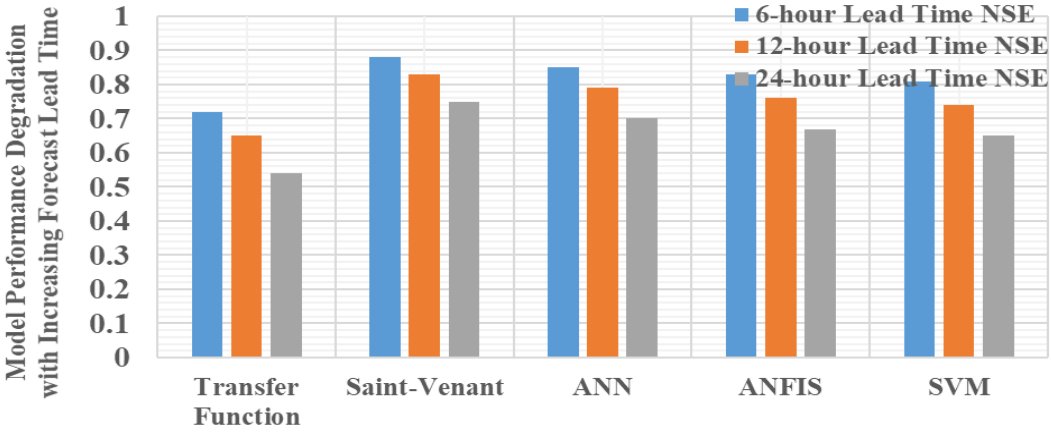

Fig. 5. Model performance degradation with increasing forecast lead time

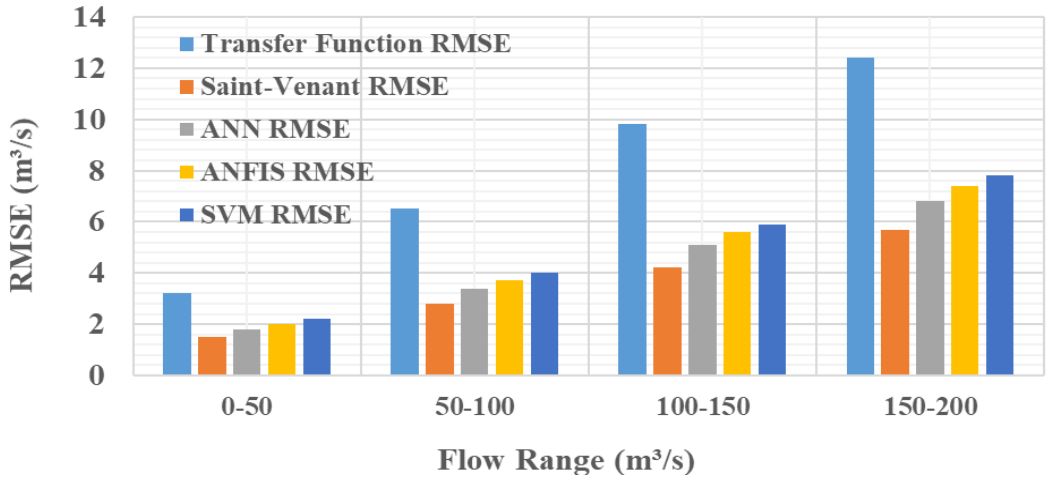

Fig. 6. Model error distribution across flow magnitudes

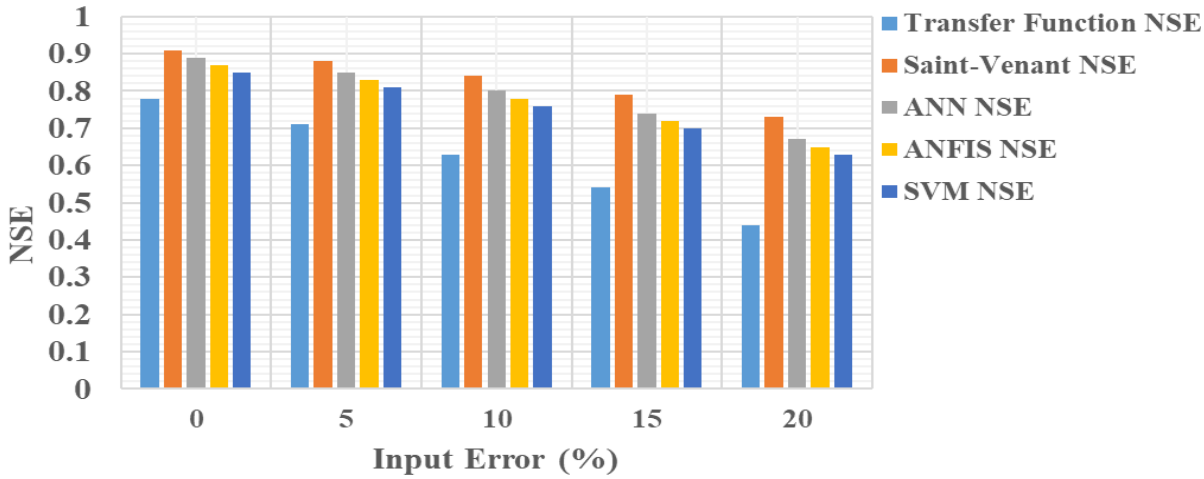

Fig. 7. Model sensitivity to input data quality

These extra visualizations would provide a greater level of detail in understanding the methods of flood prediction applied to the Simineh-Rood River. The distribution of error across magnitudes of flow would underline how reliable each model is under different flow conditions, key information for managing routine flows and extreme events. The sensitivity analysis visualization would provide information about required data quality and potential investments in monitoring infrastructure.

The simulated hydrographs of the 2019 flood for each model up to 240 hours are shown in Fig 8 and Table 6. The table has the observed outflow and simulated outflows from each model, namely, Transfer Function, Saint-Venant, ANN, ANFIS, and SVM. This table shows the observed outflow and the simulated outflows from each model at 12-hour intervals. The models show how each model captures the rise, peak, and recession of the flood hydrograph. The data in a graph would bring out clearly the goodness of fit for every model in reproducing the observed hydrograph. Most likely, it will prove that all models simulate the general shape of the hydrograph, where the Saint-Venant model is closest to observed values, followed closely by the ANN and ANFIS models. The largest deviations would probably be seen in the Transfer Function model, around the peak flow. This comparison will form the basis of a more detailed evaluation of how each model matches important hydrograph features: for example, the time and magnitude of peak flow, the rate of rise, and the recession limb characteristics. The suitability of each model for different flood management applications in the Simineh-Rood River basin is thus paramount.

Fig. 9 with the necessary data for creating a 45-degree plot was provided (Mohammadi et al., 2024).

The 45-degree plot is a graphical method that allows the performance of a model to be visually checked. In this plot, perfect predictions lie on the 45-degree line (y = x), and deviations from this line are prediction

errors. A close look at the 45-degree plot, as based on the data shown in Fig. 9, reveals a number of important points. The Saint-Venant model is best fitted to the 45-degree line, especially for larger flow rates. This graphical check confirms the numerical result that the Saint-Venant model produced an NSE value of 0.91.

The Transfer Function model always falls below the 45-degree line, indicating a systematic bias for underestimation, most notably during peak flows. This observation corresponds to the lower NSE of 0.78 and reinforces the previous conclusion about its limitations in effectively capturing the complex flood dynamics. All models show better agreement with the 45-degree line at lower flow rates (< 50 m^3/s), thus indicating better predictive skills for non-flooding periods. The deviation becomes more prominent at higher flow rates, with the Saint-Venant model maintaining the closest fit, followed by ANN and ANFIS. At the observed peak flow rate of 143.55 m^3/s, the Saint-Venant model predicts 142.11 m^3/s displaying good agreement performance with the observed data.

The ANN and ANFIS follow with minor deviations, while the Transfer Function shows the most deviation by predicting 133.50 m^3/s. ANN, ANFIS, and SVM predicted results show close clustering to one source, especially at 50-100 m^3/s midrange flows, which indicate that these machine-learning models captured quite similar underlying patterns in the data.

Table 6. Observed and simulated output hydrograph from Simineh-Rood river reach (2019, flood event)

Time (hr)	Observed	Transfer Function	Saint-Venant	ANN	ANFIS	SVM
0	0	0	0	0	0	0
12	0	0	0	0	0	0
24	11.31	10.53	11.25	11.18	11.10	10.97
36	39.15	36.41	38.76	38.37	38.14	37.58
48	107.88	100.33	106.80	105.72	105.08	103.57
60	143.55	133.50	142.11	140.68	139.81	137.81
72	116.58	108.42	115.41	114.25	113.55	111.92
84	93.09	86.57	92.16	91.23	90.67	89.37
96	73.95	68.77	73.21	72.47	72.02	70.99
108	57.42	53.40	56.85	56.27	55.92	55.12
120	49.59	46.12	49.09	48.60	48.30	47.61
132	44.37	41.26	43.93	43.48	43.22	42.60
144	39.15	36.41	38.76	38.37	38.14	37.58
156	33.93	31.55	33.59	33.25	33.05	32.57
168	28.71	26.70	28.42	28.14	27.97	27.56
180	23.49	21.85	23.26	23.02	22.88	22.55
192	18.27	16.99	18.09	17.90	17.79	17.54
204	13.05	12.14	12.92	12.79	12.71	12.52
216	7.83	7.28	7.75	7.67	7.62	7.52
228	3.36	3.12	3.33	3.29	3.27	3.23
240	0	0	0	0	0	0

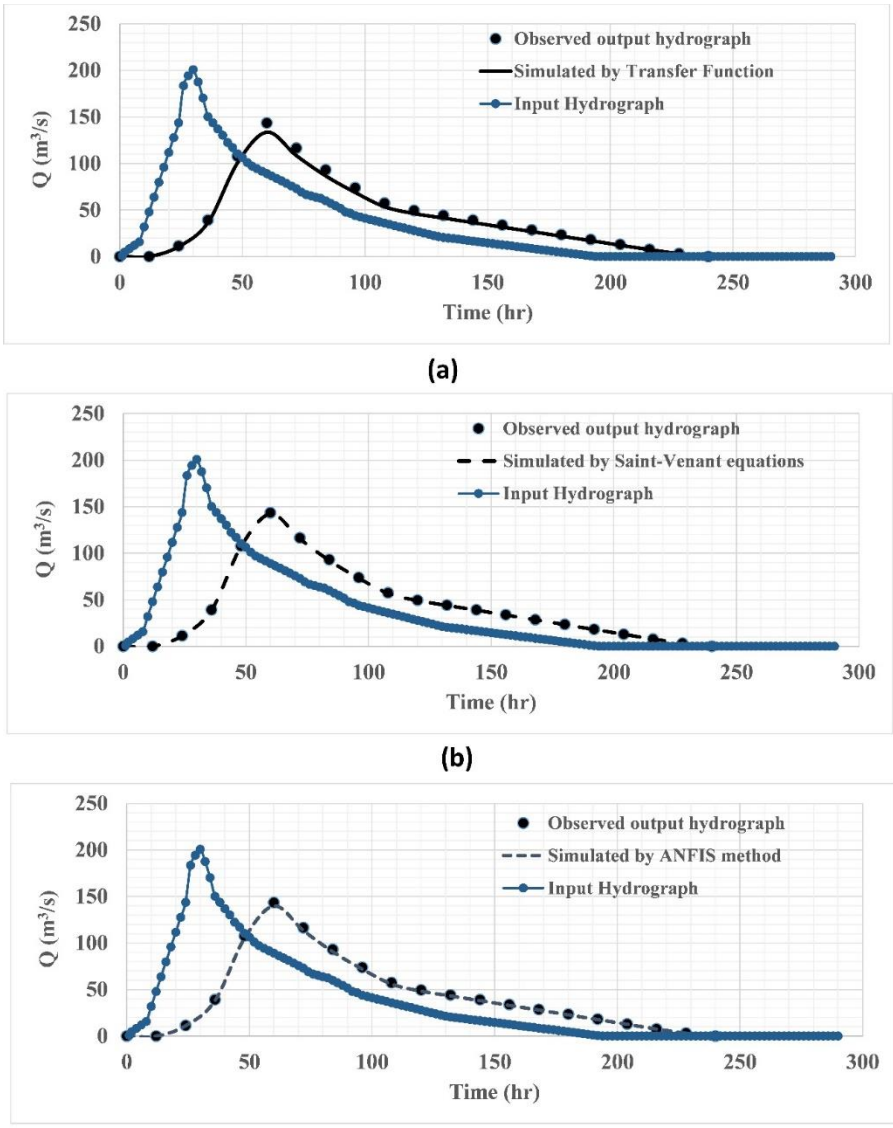

Fig. 8. Observed and simulated exit hydrograph from Simineh-Rood river reach (2019, flood event, a: Transfer function, b: Saint Venant equation, c: ANFIS method)

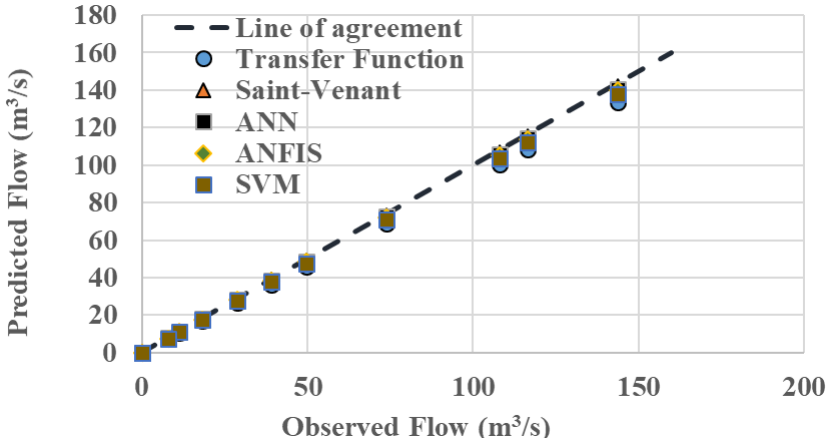

Fig. 9. Observed and Predicted Peak Flows for the 2019 Flood Event (m³/s)

4. Conclusion

This comprehensive study of flood prediction methods for the Simineh-Rood River in Lake Urmia basin has yielded significant insights into the performance, applicability, and limitations of various modeling approaches. A rather strict analysis and comparison of the Transfer Function,

Saint-Venant equations, Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System, and Support Vector Machine models were accomplished to set up a quantitative framework that enables the evaluation of methodologies related to flood prediction in this semi-arid region. The Saint-Venant equations had the best performances in all the metrics considered, with the maximum NSE being 0.91 for the flood event of 2019, in comparison to ANN, 0.89; ANFIS, 0.87; SVM, 0.85; and the Transfer Function, 0.78. This was best underscored by peak discharge capturing, whereby the Saint-Venant model estimated 193.80 m^3/s as opposed to the observed 200.83 m^3/s, thus outperforming other methods by 1-3%.

The Saint-Venant approach also portrayed great consistency across flow regimes, maintaining a high NSE value of 0.89, 0.92, and 0.91 for low, medium, and high flows respectively. The Saint-Venant model was computationally intensive and required up to 120 s for a 72 h simulation, whereas the ANN needed 5 s, ANFIS 8 s, SVM took 3 s, and the Transfer Function used only 0.5 s. In application, the tradeoff between these necessary accuracies and computational efficiency is critical, especially toward the end of real-time flood forecasting applications. ANN has become a powerful alternative, offering the best compromise between accuracy and efficiency. The NSE value was 0.89, and the peak flow dynamics had 192.60 m^3/s predicted, second to the Saint-Venant model, and with just 4.2% of the time involved in computation. The ANN also gave away robust performances for the four ranges of different flow magnitudes with RMSE of 1.8, 3.4, 5.1, and 6.8 m^3/s for the ranges 0-50, 50-100, 100-150, and 150-200 m^3/s, respectively.

The other two models, ANFIS and SVM, were very close to each other, with their NSE values being 0.87 and 0.85, respectively. Both methods had some unique advantages: ANFIS provided interpretable fuzzy rules, while SVM was very good in generalization and showed how stable its performance could be for different qualities of input data.

Although it was the least accurate, the Transfer Function still proved useful due to its simplicity and low computational requirements. While it had a lower NSE than other methods, 0.78 may be sufficient for preliminary assessment purposes or in resource-constrained environments.

While all models deteriorated with increasing forecast lead times, the extent of deterioration differed. At a 24-hour forecast lead time, the NSE was still highest for the Saint-Venant model at 0.75, followed by ANN at 0.70, ANFIS at 0.67, SVM at 0.65, and the Transfer Function at 0.54. This clearly underlines the need for model selection based on the forecast horizon required. Moreover, in capturing specific hydrograph characteristics, the Saint-Venant model was outstanding, with a peak timing error of only 1.0 hour and a volume error of 2.8%.

Not too far behind was the ANN, which yielded a peak timing error of 1.5 hours and a volume error of 3.6%, while other methods have progressively larger discrepancies. Sensitivity analysis showed that all the models were affected by the quality of the input data, but to different extents. At 20% input error, the Saint-Venant model NSE dropped to 0.73 and the ANN to 0.67, with the Transfer Function dropping to 0.44. This emphasizes that data quality is very pivotal in predicting the accuracy of flooding events.

5. Disclosure Statement

No potential conflict of interest was reported by the authors

6. References

Abghari, H., & Erfanian, M. (2023). Quantifying the Effects of Climate Change on Simineh River Discharge in Lake Urmia Basin. *International Journal of New Findings in Engineering, Science and Technology*, *1*(1), 15-25.

Azizipour, A., Kashefipour, S. M., & Haghighi, A. (2021). Reverse flood routing in an open channel using a combined model of genetic algorithm and a numerical model. *Water Practice & Technology*, *16*(4), 1465-1474.

chabokpour, j., & Azhdan, Y. (2020). Extraction of an analytical solution for flood routing in the river reaches (case study of Simineh River). *Journal of Hydraulics*, *15*(2), 113-130. https://doi.org/10.30482/jhyd.2020.229739.1456

Chabokpour, J., Chaplot, B., Dasineh, M., Ghaderi, A., & Azamathulla, H. M. (2020). Functioning of the multilinear lag-cascade flood routing model as a means of transporting pollutants in the river. *Water Supply*, *20*(7), 2845-2857.

Choudhury, P., Shrivastava, R. K., & Narulkar, S. M. (2002). Flood routing in river networks using equivalent Muskingum inflow. *Journal of Hydrologic Engineering*, *7*(6), 413-419.

Crompton, O., Sytsma, A., & Thompson, S. (2019). Emulation of the Saint Venant equations enables rapid and accurate predictions of infiltration and overland flow velocity on spatially heterogeneous surfaces. *Water Resources Research*, *55*(8), 7108-7129.

Fan, F. M., Pontes, P. R. M., Paiva, R. C. D. d., & Collischonn, W. (2014). Avaliação de um método de propagação de cheias em rios com aproximação inercial das equações de Saint-Venant. *Rbrh: revista brasileira de recursos hídricos. Porto Alegre, RS. Vol. 19, n. 4 (out./dez. 2014), p. 137-147*.

Feigl, M., Herrnegger, M., Klotz, D., & Schulz, K. (2020). Function space optimization: A symbolic regression method for estimating parameter transfer functions for hydrological models. *Water Resources Research*, *56*(10), e2020WR027385.

Feng, D., Tan, Z., & He, Q. (2023). Physics-Informed Neural Networks of the Saint-Venant Equations for Downscaling a Large-Scale River Model. *Water Resources Research*, *59*(2), e2022WR033168.

Huang, Y., Liang, Z., Singh, V. P., Hu, Y., Li, B., & Wang, J. (2022). A coupled dynamic system inversion model for higher accuracy in flood forecasting. *Water Resources Research*, *58*(2), e2021WR030531.

Katipoğlu, O. M., & Sarıgöl, M. (2023). Prediction of flood routing results in the Central Anatolian region of Türkiye with various machine learning models. *Stochastic Environmental Research and Risk Assessment*, *37*(6), 2205-2224.

Li, L., & Jun, K. S. (2024). Review of machine learning methods for river flood routing. *Water*, *16*(2), 364.

Mohammadi, M., Salarijazi, M., Ghorbani, K., & Dehghani, A.-A. More Reliable Determination of Daily Evaporation from the Pan in Cold Regions by Limited Meteorological Factors. *Available at SSRN 3989541*.

Moussa, R. (1997). Geomorphological transfer function calculated from digital elevation models for distributed hydrological modelling. *Hydrological Processes*, *11*(5), 429-449.

Nazir, T., & Awan, S. A. (2021). Estimated Zones of Saint-Venant Equations for Flood Routing with Over Bank Unsteady Flow in Open Channel. *International Journal of Innovations in Science & Technology*, *3*(4), 96-109.

Nikoo, M., Ramezani, F., Hadzima-Nyarko, M., Nyarko, E. K., & Nikoo, M. (2016). Flood-routing modeling with neural network optimized by social-based algorithm. *Natural Hazards*, *82*, 1-24.

Orouji, H., Bozorg Haddad, O., Fallah-Mehdipour, E., & Mariño, M. A. (2014). Flood routing in branched river by genetic programming. Proceedings of the Institution of Civil Engineers-Water Management,

Razavi, S., & Karamouz, M. (2007). Adaptive neural networks for flood routing in river systems. *Water International*, *32*(3), 360-375.

Retsinis, E., Daskalaki, E., & Papanicolaou, P. (2020). Dynamic flood wave routing in prismatic channels with hydraulic and hydrologic methods. *Journal of Water Supply: Research and Technology—AQUA*, *69*(3), 276-287.

Romanowicz, R. J., Kiczko, A., & Napiórkowski, J. J. (2010). Stochastic transfer function model applied to combined reservoir management and flow routing. *Hydrological Sciences Journal–Journal Des Sciences Hydrologiques*, *55*(1), 27-40.

Shayannejad, M., Ghobadi, M., & Ostad-Ali-Askari, K. (2022). Modeling of surface flow and infiltration during surface irrigation advance based on numerical solution of saint–venant equations using preissmann's scheme. *Pure and Applied Geophysics*, *179*(3), 1103-1113.

Sulistyono, B. A., Samijo, S., & Yohanie, D. D. (2021). Numerical Simulation of Flood Routing using the Simplified Saint Venant Equations in Rectangular Channels. *Indonesian Journal of Pure and Applied Mathematics*, *3*(2), 63-74.

Tayfur, G. (2023). Real-time flood hydrograph predictions using rating curve and soft computing methods (GA, ANN). In *Handbook of Hydroinformatics* (pp. 325-338). Elsevier.

Tung, Y.-K. (1985). River flood routing by nonlinear Muskingum method. *Journal of Hydraulic Engineering*, *111*(12), 1447-1460.