IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 December 2023, accepted 3 January 2024, date of publication 9 January 2024, date of current version 24 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3351754

RESEARCH ARTICLE

Comparative Analysis of Artificial Intelligence Methods for Streamflow Forecasting

YAXING WEI[1], HUZAIFA BIN HASHIM[1], SAI HIN LAI[1,2], KAI LUN CHONG[3], YUK FENG HUANG[4], ALI NAJAH AHMED[5], MOHSEN SHERIF[6,7], AND AHMED EL-SHAFIE[1]

[1]Department of Civil Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia
[2]Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak 94300, Malaysia
[3]Faculty of Engineering and Quantity Surveying, INTI International University (INTI-IU), Nilai, Negeri Sembilan 71800, Malaysia
[4]Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Kajang, Selangor 43000, Malaysia
[5]School of Engineering and Technology, Sunway University, Petaling Jaya 47500, Malaysia
[6]Civil and Environmental Engineering Department, College of Engineering, United Arab Emirates University, Al Ain, United Arab Emirates
[7]National Water and Energy Center, United Arab Emirates University, Al Ain, United Arab Emirates

Corresponding author: Huzaifa Bin Hashim (huzaifahashim@um.edu.my)

This work was supported by the National Water and Energy Center, United Arab Emirates University, United Arab Emirates, for APC.

ABSTRACT Deep learning excels at managing spatial and temporal time series with variable patterns for streamflow forecasting, but traditional machine learning algorithms may struggle with complicated data, including non-linear and multidimensional complexity. Empirical heterogeneity within watersheds and limitations inherent to each estimation methodology pose challenges in effectively measuring and appraising hydrological statistical frameworks of spatial and temporal variables. This study emphasizes streamflow forecasting in the region of Johor, a coastal state in Peninsular Malaysia, utilizing a 28-year streamflow-pattern dataset from Malaysia's Department of Irrigation and Drainage for the Johor River and its tropical rainforest environment. For this dataset, wavelet transformation significantly improves the resolution of lag noise when historical streamflow data are used as lagged input variables, producing a 6% reduction in the root-mean-square error. A comparative analysis of convolutional neural networks and artificial neural networks reveals these models' distinct behavioral patterns. Convolutional neural networks exhibit lower stochasticity than artificial neural networks when dealing with complex time series data and with data transformed into a format suitable for modeling. However, convolutional neural networks may suffer from overfitting, particularly in cases in which the structure of the time series is overly simplified. Using Bayesian neural networks, we modeled network weights and biases as probability distributions to assess aleatoric and epistemic variability, employing Markov chain Monte Carlo and bootstrap resampling techniques. This modeling allowed us to quantify uncertainty, providing confidence intervals and metrics for a robust quantitative assessment of model prediction variability.

INDEX TERMS Artificial neural network, deep learning convolutional neural network, Bayesian statistic, streamflow, time series, uncertainty analysis.

ABBREVIATIONS AND ACRONYMS

A	Approximation Component.
AdaBoost	Adaptive Boosting.
AI	Artificial Intelligence.
ANFIS	Adaptive Neuro-Fuzzy Inference Systems.
ANN	Artificial Neural Network.
ARIMA	Autoregressive Integrated Moving Average.
CNN	Convolutional Neural Network.
COD	Chemical Oxygen Demand.
DW	Detail Component.
ELM	Elman Neural Network.
GP	Gaussian Processes.
GRU	Gated Recurrent Units.
LSTM	Long Short Term Memory.

The associate editor coordinating the review of this manuscript and approving it for publication was Massimo Cafaro.

MLP	Multi-Layer Perceptron.
MSE	Mean Square Error.
NSE	Nash-Sutcliffe Efficiency.
NH_4^+	Ammonium Concentration.
PCA	Principal Component Analysis.
RF	Random Forest.
RFBN	Radial Basis Function Neural Networks.
RMSE	Root Mean Square Error.
RNN	Recurrent Neural Network.
SVM	Support Vector Machine.
SVR	Support Vector Regression.
SWAT	Soil & Water Assessment Tool.
SWE	Snow Water Equivalent.
TSL	Total Sediment Load.
TSS	Total Suspended Solids.
WT	Wavelet Transform.
WWTP	Wastewater Treatment Plant.
XGB	eXtreme Gradient Boosting.

I. INTRODUCTION

Streamflow monitoring is critical for estimating the availability and distribution of water resources for human water demands, which is essential for agricultural irrigation, industrial operations, and municipal water supply planning. Water resource pressures rise and environmental concerns grow, prompting an urgent need to address a critical question: Why is streamflow important, and how can a nuanced understanding of its patterns and dynamics be leveraged for sustainable water management, ecological health, and infrastructure planning in the face of evolving environmental and societal demands?

More broadly, the significance of streamflow extends beyond the context of human water demands. Terrestrial aquatic ecosystems are intricately linked to streamflow dynamics, and their health depends on consistent adequate flow [1]. Streamflow can impact the prevalence and transmission of illnesses in aquatic ecosystems. While streamflow does not directly cause diseases, it does play a crucial role in generating habitat conditions that can impact pathogen dynamics and their consequences for aquatic animals. Streamflow facilitates the transport of pathogens, including bacteria, viruses, and parasites, through aquatic environments [2]. Increased flow can disperse pathogens over greater distances, potentially affecting a broader range of species and ecosystems. Streamflow influences water quality parameters such as temperature, oxygen level, and nutrient concentrations [3]. Changes in these factors due to variations in streamflow can stress aquatic organisms, making them more susceptible to diseases.

Understanding streamflow patterns not only is essential for ecological health but also forms the basis for sedimentary budget forecasting in fluvial flows [4]. Streamflow modeling becomes a key tool in this regard, as it simulates hydrodynamic processes within river systems, including flow velocity, discharge, and channel morphology. These models make it possible to identify areas prone to erosion or sediment deposition [5]. Inadequate streamflow management can lead to increased erosion along riverbanks and within the river channel. Without sufficient flow to transport sediment downstream, sediments may accumulate due to elevated net sedimentation rates. This can adversely affect water quality, aquatic habitats, and infrastructure [6].

Streamflow modeling is a vital technique for managing water resources, especially in the early detection of flood dangers [7], [8], [9]. Several types of advanced models can operate across the range of climate zones. Especially during flood events, fast efficient replication of streamflow is crucial to the forecasting process and is accomplished by hydrodynamic models [10], [11]. For instance, Mahdian et al. [12] employed the Soil and Water Assessment Tool model to simulate and analyze streamflow dynamics under various climate and land use scenarios. Their findings help clarify how changes in climate and human activities, such as deforestation and urbanization, can impact streamflow and sediment inputs to ecosystems, in their case the Anzali wetland ecosystem. However, these complex models require precise river geometry data, which is not always accessible. In contrast, artificial intelligence (AI) tools, such as artificial neural networks and deep learning, have abolished the necessity for detailed knowledge of river geometry [13]. Furthermore, their modeling can be nonlinear, piecewise, or discontinuous, among other types of relationships [14].

II. LITERATURE REVIEW

Existing literature reveals successful demonstrations of AI-based models in hydrological forecasting, leading to improved accuracy and predictive capabilities in hydrological and water resource management. Models range from widely used artificial neural networks (ANNs) to state-of-the-art algorithms, such as deep learning, which include convolutional neural networks (CNNs), long short-term memory (LSTM), and generative adversarial networks.

Recurrent neural networks (RNNs), which include LSTM, gated recurrent units (GRU), and standard RNN, are widely employed for time series forecasting due to their proficiency in handling sequential data. Sahoo et al. [15] illustrated their effectiveness, finding that RNN outperforms radial basis function neural networks (RFBN) for streamflow forecasting. However, standard RNNs can encounter issues related to vanishing or exploding gradients, potentially affecting their performance. Samantaray et al. [16] demonstrated this challenge by showing that support vector machines (SVMs) and adaptive neuro-fuzzy inference systems (ANFISs) outperformed RNN in rainfall modeling.

In response to the vanishing/exploding gradient problem, LSTM has gained prominence. LSTM models exhibit a lower susceptibility to these pitfalls and offer enhanced performance. For instance, Bala et al. [17] achieved success using LSTM for rainfall prediction, with the LSTM approach surpassing the Elman neural network (ELM) and autore-

gressive integrated moving average (ARIMA) approaches. Researchers can also incorporate dynamic sliding windows when working with time series data exhibiting a range of periodicities. For instance, Dong et al. [18] applied LSTM with dynamic sliding windows to predict streamflow at monthly intervals, resulting in an impressive 8.63% RMSE improvement. LSTM models have also demonstrated their suitability for real-time streamflow prediction. In a comparative study by Cheng et al. [19], their research employed a recursive forecasting procedure for long lead-time forecasting, where the last one-step-ahead forecast was used as a new input for the next-step-ahead forecast in both ANN and LSTM systems. The LSTM model outperformed the ANN model in long lead-time daily forecasting, showcasing superior performance. However, the LSTM model's performance was less satisfactory in multi-monthly forecasting, attributed to the limited availability of a large monthly training dataset.

Deep learning algorithms used in hydrology produce not only one-dimensional (1D) time-series forecasts but also significant supplemental information. Recent research demonstrates the effectiveness of CNNs in various groundwater mapping tasks. Panahi introduced a CNN model for creating groundwater potential maps, which offer insights into the spatial distribution of groundwater resources essential for sustainable freshwater management. Furthermore, Hakim et al. [20] utilized a CNN–LSTM approach to unveil patterns among 14 parameters and responses related to groundwater potential, further enhancing groundwater mapping endeavors. Other research, delving into three-dimensional (3D) mapping, has predicted flow rates through rock formations in 3D. Santos et al. [21] successfully employed a 3D CNN to forecast flow rates, presenting an alternative model for the Navier–Stokes equation, a pivotal component of hydrodynamic simulations. These advancements in deep learning hold significant promise for enhancing hydrological modeling and resource management.

Aside from deep learning algorithms, recent research in hydrological problems concerning streamflow has proven the Gaussian processes (GPs) method as a viable route in forecasting streamflow Donnelly et al. [22]. These authors demonstrated that GPs consistently outperform other models, including CNNs, excelling in streamflow forecasting due to their adaptability to complex hydrological systems, precise uncertainty estimation for flood risk management, and computational efficiency. GP models exhibit minimal inaccuracies in predicting streamflow, even in complex metropolitan settings. These attributes make GPs a practical choice for dependable efficient streamflow prediction.

A further investigation of streamflow hydrological processes is that of Khosravi et al. [23] concerning soil erosion by flowing water. That study identified elevation as the most influential factor for soil water erosion (SWE) susceptibility. They achieved high predictive performance using various deep learning models, such as CNN, LSTM, and RNN, with RNN exhibiting a slight advantage during testing. These models provided valuable tools for SWE susceptibility mapping in data-poor regions but inherently lacked explanatory capability. Predicting sediment loads has also been a research focus. For instance, Noori et al. [24] intriguing study used principal component analysis (PCA) to address multicollinearity among ten input factors impacting total sediment load in rivers. They created PCA-based models using multiple linear regression or support vector regression (SVR). During verification, the PCA-based RBF–SVR model attained an exceptional Nash–Sutcliffe efficiency (NSE) of 0.86, outperforming previous empirical models and confirming its resilience in forecasting extreme sediment concentrations.

Wastewater treatment facilities represent an additional hydrological concern. Borzooei et al. [25] conducted extensive data analysis, which involved identifying outliers and addressing missing values, emphasizing significant outlier events in chemical oxygen demand (COD) and total suspended solids (TSS) data. They established peaking factors for flow rate and pollutant characteristics. Their regression analysis demonstrated a moderate positive association of precipitation intensity (PI) with influent flow (Q_{in}) and a low negative correlation with pollutant concentration. Wet weather, as determined by PI criteria, resulted in higher flow rates and lower pollutant concentrations owing to dilution effects. COD and ammonium concentration (NH_4^+) were more susceptible to meteorological conditions than TSS. Another study by Borzooei et al. [26] aimed to estimate the impact of wet-weather events on the Q_{in} and wastewater characteristics at a prominent Italian wastewater treatment plant (WWTP) in Castiglione Torinese. They analyzed eight years of collected data on influent flow and daily precipitation to investigate the relationships between PI, Q_{in}, COD, N-NH4, and TSS. The study employed time-series data mining to segment data into wet and dry weather events, with a wet-weather definition proposed based on PI criteria. Their results suggest that this methodology provides a practical alternative to costly time-consuming data collection efforts for emergency response and climate preparedness.

However, the complexity of time series forecasting procedures and the reliance on high-quality data in data-driven approaches further enhance the desirability of integrating AI models with data-preprocessing methods [27]. Data preprocessing methods data mining methods for transforming raw data into data that can be efficiently used [28], [29]. The wavelet transform (WT) is one of the most used time series preprocessing techniques for temporal extraction of crucial properties, including short- and long-term fluctuations. Among the methods for accomplishing such extraction is decomposition of the time series into its several significant subcomponents [30]. Including WT in AI models has become an integral aspect of time series forecasting applications, particularly in univariate streamflow, where the modeling processes considers only historical datasets. Several researchers have used the wavelet analysis principle to explore the association and coherency between streamflow

and climate given constantly changing climatic conditions. Ghaderpour et al. [31] conducted a study to understand the interaction between climatic factors and streamflow. Their research focused on uncovering trends and patterns that might not be evident in conventional temporal or frequency domain analysis. This approach is valuable for identifying the cyclic or periodic patterns that influence streamflow, especially over extended periods. Moreover, the coherency established through this approach aids in comprehending when and to what extent climate variables and streamflow are correlated. Notably, in this research, the analysis of time lags can reveal changes in precipitation precede or follow changes in streamflow. Such information is crucial for understanding cause-and-effect relationships and improving forecasting accuracy.

In streamflow forecasting, WT machine learning applications have demonstrated significant success. Chong et al. [32] presented an application that involved the aggregation of decomposed wavelets to reconstruct the streamflow series, resulting in performance superior to that of conventional machine learning. Wu and Wang [33] used an LSTM model rather than a CNN in research following these same principles. In this context, deep learning may sometimes be viewed as superior to machine learning due to its ability to execute complex tasks. However, such an assumption should not be made by default. Indeed, one well-known problem with deep learning is its propensity to overfit the training data, especially for a relatively simple dataset, since deep learning automatically learns a massive number of parameters when constructing the model [34]. However, machine learning, such as by an ANN, tends to underfit a complicated dataset according to a set architectural form, thus suffering from a flaw complementary to that from which deep learning can suffer [35]. However, several studies have shown that deep learning can outperform other methods for handling time series problems when the spatial and temporal variations are critical [36], [37]. This question of methodological superiority has remained overlooked in prior WT research since deep learning is typically seen as inherently superior to machine learning. However, this question could be effectively and meaningfully tackled by raising it in the context of a set of streamflow impacts that includes simple one and spatially and temporally variable ones.

Recognizing that spatially and temporally variable streamflows add complexity to the modeling process, previous research has demonstrated that the most extensively used deep learning algorithms, such as LSTM and CNN, can account for spatial and temporal streamflow variability. For instance, Hussain et al. [38] investigated the deep learning technique, specifically a 1D-CNN, for one-step-ahead streamflow forecasting, where the streamflow time series is convoluted with a 1D- CNN kernel. Van et al. [39] adopted the multiple time series approach, consisting of several rainfall stations' data when utilizing CNN in hydrological forecasting to exploit the correlation structure among the time series. Furthermore, hybridizing 1D-CNN with another deep learning algorithm, such as LSTM and GRU, may provide a better solution. In the hybrid approach outlined above, CNN processes and sends the information to LSTM. However, under a CNN, the convoluted region of input data frequently overlaps at grid sizes smaller than the original grid size as the CNN traverses the whole domain of inputs to find a better feature extraction procedure. This point is crucial in signal processing and classification applications. Overlapping across the entire input domain, however, might cause a significant obstacle in time series forecasting: data leakage. Previous research has thus concentrated on the spatial and temporal effects of streamflow forecasting rather than the practical relevance of time series forecasting procedures.

Approaches previously employed in real-time applications have typically produced only a single prediction due to the lack of time to perform numerous model runs. However, such approaches are coming to be seen as less satisfactory, since numerous model runs are necessary to account for the variability of the forecasting outcomes, given the uncertainty surrounding the future in real-time applications. Technical issues frequently undermine the credibility of such approaches, and the absence of adequate calibration makes it impossible to specify effective mitigation solutions [40], [41]. Overestimating the technical limitation may result in over-engineered design features, while underestimating it may increase the probability of model failure [42]. Measuring the output behavior of a mathematical model is essentially focused on computing statistical values such as mean, variance, and distribution quantiles [43], [44]. These statistical qualities are often quantified along with a set of confidence levels, which give a more robust uncertainty assessment [45], [46]. Modeling the aleatoric and epistemic variation also generally helps in assessing the variation among model outcomes [47].

One computationally relatively simple approach, known as Monte Carlo analysis, conveys prediction variations in the form of a probabilistic result [48]. An alternative approach to Monte Carlo analysis is the ensemble prediction approach. Ensemble predictions are collections of a limited number of diverse models, offering greater variety than single-model prediction. For example, Sharma et al. [8] incorporated two separate models: a weather forecasting model and a simulation hydrology model. This integration resulted in an ensemble forecasting typically performing better with medium-range time-frames than with shorter lead periods. The ensemble technique is not limited to stacking only models that serve different functions. For instance, Lu et al. [49] adopted a stacking ensemble strategy for predicting daily streamflow in the Qiantang River Basin and demonstrated that the final ensembled forecasting model outperformed each of its constituent models, including random forest (RF), adaptive boosting (AdaBoost), and extreme gradient boosting (XGB), each trained individually on the same set of data. Without questioning the probabilistic approach in general, these authors could argue that single-model predictions were

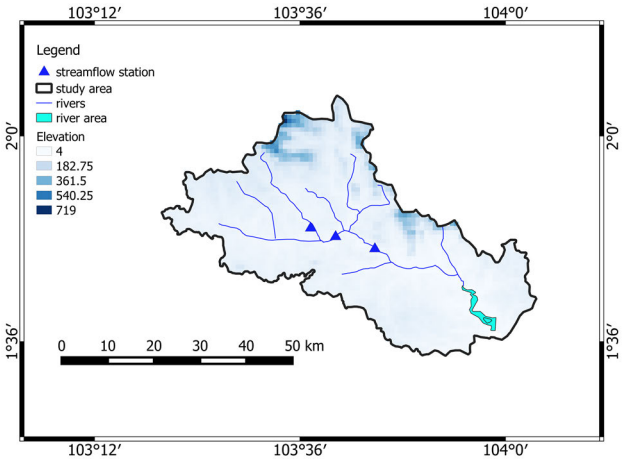

FIGURE 1. Location of Johor river basin.

either relatively accurate or relatively inaccurate. Their means or some other means of elucidating the variability among predictions is crucial.

III. CONTRIBUTIONS OF THE PRESENT STUDY

The paper emphasizes the following research questions, given the progress made in previous studies: 1) How different is a deep learning algorithm from a conventional machine learning algorithm in handling a set of sub-series that result from WT decomposition? 2) What forecasting advantages do developing machine learning regression processes have, especially when considering aleatoric and epistemic variables?

The primary objective of the present research is to address significant gaps in streamflow forecasting through the application of artificial neural networks and deep learning convolutional neural networks. The research is aimed at addressing the following main issues.

- To investigate the robustness and adaptability of AI modeling streamflow forecasting under varying time series complexity, providing insight into the models' applicability and limitations.
- To address the practical water resources management challenges associated with streamflow forecasting, namely the temporal constraints imposed by deep learning CNN.
- To address the aleatoric and epistemic variables within AI modeling streamflow forecasting to improve forecasting reliability using a Bayesian neural network, in which quantification is achieved through confidence interval estimation using the bootstrap method.

IV. METHODOLOGY

A. CASE STUDY

As illustrated in Fig. 1, Johor is a coastal state on the east coast of Peninsular Malaysia, near the border with Singapore. It boasts a 400-kilometer coastline that includes both the east and west coasts and is recognized for its tropical rainforest climate, with the South China Sea monsoon season lasting from November to February each year. The average annual rainfall is 1,788 mm, and the mean annual relative humidity is 84%. Temperatures typically range between 21°C and 32°C, with a mean annual temperature of 26.7 °C. The Johor River, the principal river of Johor, has an approximate length of 130 km and a catchment area of 2,600 km^2. Table 1 shows the average, minimum, and maximum annual Johor River streamflows recorded over the 1977–2005 period. Its principal tributaries include the Sayong, Lebam, Linggui, and Tiram Rivers. Malaysia's Department of Irrigation and Drainage provided the streamflow data for this 28-year study.

TABLE 1. Monthly Streamflow raw data: Maximum, mean, and minimum (m^3/s).

Month	Maximum	Year of maximum	Minimum	Year of minimum	Mean
January	1,694.3	1997	130.7	1984	712.8
February	1964.8	1984	68.4	1989	459.1
March	2,349.9	2004	8.6	2005	527.7
April	1,032.0	1979	12.9	2005	531.5
May	986.5	1981	28.5	2005	533.3
June	391.7	2000	101.4	2005	404.0
July	888.7	1996	430.9	1994	405.9
August	945.9	1988	96.8	2001	417.1
September	1,498.9	1988	148.2	1997	446.4
October	1,395.5	2003	242.1	1992	636.4
November	1,623.4	1987	222.6	2005	774.3
December	2,305.9	1983	199.3	2001	815.9

B. MACHINE LEARNING APPROACHES

In our study of streamflow forecasting, we used an MLP neural network, a traditional ANN architecture. Traditionally, the choice of MLP is foundational for benchmarking against more advanced models. Additionally, we explored the use of a CNN, which is a state-of-the-art ANN variant tailored for image recognition tasks. Here, we specifically designed it for the analysis of spatial data. This traditional CNN model was used with the aim of capturing spatial patterns in streamflow data.

1) DATA COMPONENT

The main goal was to build a streamflow forecasting model that accurately reflects the complex link between a collection of input variables and the related streamflow values. These input properties $x_p \in R^I$ include a variety of potentially significant factors such as rainfall, temperature, and humidity, that collectively impact streamflow dynamics. Streamflow values $x_p \in R^k$ indicate the outcomes of interest, representing the volume of water flow in a river or stream per time step. Through the mapping of input characteristics and streamflow values, the model attempts to uncover patterns and correlations in past data comprising both input attributes and streamflow values, allowing projections of future or alternative streamflow outcomes. The process of neural network approximation has the goal of approximating the underlying mapping $F : x_p \rightarrow t_p$ using a neural network.

The activation function ϕ and weight parameter w are interconnected throughout the layers of the neural network architecture. The predicted streamflow $\mathrm{Y}(x_p, w)$, which emerges through a sequence of weighted transformations and activation, is determined by the output layer of the network. Mathematically, we can symbolize the predicted streamflow $\mathrm{Y}(x_p, w)$ as

$$\mathrm{Y}(x_p, w) = \phi_{\mathrm{L}}[w_L \times \phi_{\mathrm{L}-1}(w_{L-1} \times \ldots \phi_1(w_1 \times z_p))], \quad (1)$$

where L signifies the number of layers in the neural network. The input feature vector z_p undergoes a series of operations (additive and multiplicative) involving weight matrices w_L and activation functions ϕ to yield the final streamflow value prediction. The activation function ϕ of each layer transforms the operation's outcome from the previous layer, progressively mapping the initial input features to the ultimate output through a series of weight parameters $\{w_1, w_2, \ldots w_L\}$.

Due to the complexity of natural systems such as streamflow, attaining a flawless alignment between $\mathrm{F}: x_p \rightarrow t_p$ and $\mathrm{Y}(x_p, w)$ is often unattainable. We considered a Euclidean norm as a way to quantify and address the discrepancies between our predictions and the actual values. This norm serves as a measure of the distance between two vectors: the predicted streamflow value $\mathrm{Y}(x_p,w)$ and the actual target streamflow value t_p for a given data component. The difference between these vectors is given as $\mathrm{Y}(x_p, w) - t_p$, which represents the prediction error associated with that single data component.

A loss function was introduced to quantify the discrepancy between expected outputs and actual targets. This function directs the network to modify its settings to reduce this discrepancy, increasing predictive accuracy. The loss is frequently determined using the sum of squared errors (SSE), obtained by adding all the patterns. However, comparing SSE values across simulations of various sizes might lead to inaccurate conclusions. Among the array of widely favored loss functions, notable candidates include the mean squared error (MSE) loss, cross-entropy loss [50], hinge loss [51], Kullback–Leibler loss [52], and Huber loss [53]. Selection among these loss functions hinges on whether the neural network serves a classification role or takes on a regression task. A comprehensive review of the loss functions in machine learning can be found in Wang et al. [54].

The MSE cost function was adopted to evaluate streamflow forecasting performance. The model offers an in-depth representation of the overall prediction error by adding the squared values of these errors across all data components in the historical streamflow datasets D. This summation, $\sum_{\mathrm{d}=1}^{\mathrm{D}} [\mathrm{Y}(x_p, w) - t_p]^2$, summarizes the collective disparity between predictions and actual values across the entire set of datasets. By iteratively updating the weight parameters of network w through the optimization process, the model converges toward minimization of this prediction error across all data components D. This iterative adjustment process enables the neural network to progressively enhance its predictive capabilities and model the complex relationship between input features and streamflow values with increasing accuracy.

These criteria are relevant in that the lower the values of these criteria, the more accurate the anticipated outcome. The minimal value for these criteria is 0, indicating flawless prediction (no difference between predicted and actual values). However, the result (always a positive value) depends on the scale of the numbers utilized. For example, an RMSE of 0.7 is modest for data that range from 0 to 1,000. However, when the data range is reduced to 0 to 1, the same RMSE of 0.7 is no longer considered small; it assumes a more substantial proportion relative to the data range.

2) DATA COLLECTION AND PREPROCESSING

Data preparation is considered a pivotal phase within the ML process. Often, raw data sourced from diverse channels harbors anomalies, outliers, missing values, or noise that can potentially impact the efficacy of the predictive models [55]. The core of data preparation consists of a series of procedures intended to clean up, transform, and organize the data into a format that complies with model training specifications, thereby alleviating these concerns. These procedures include addressing missing values through approaches such as interpolation or imputation, identifying and managing outliers to avert their distortion of outcomes, ensuring consistent impact by scaling input features, and transforming categorical variables into numerical representations utilizing methods such as one-hot encoding. Additionally, the aggregation of data into specific time intervals holds key importance for discerning temporal relationships.

3) SCALING AND TRANSFORMATION AND PERFORMANCE CRITERION

As a necessary part of data preparation for machine learning, scaling emerges as a fundamental facet, requiring the intricate task of aligning input values with the dynamic range of the chosen activation function, that is, the range over which the activation function exhibits heightened sensitivity to variation in computing the information, particularly in the hidden layer. Employing this approach prevents scenarios in which the activation function operates within dormant domains, thereby potentially impeding meaningful adjustments to minor weight factors.

Within the extensive toolkit of scaling and transformation techniques, a suite of strategies exists to normalize those values within datasets that might introduce difficult complexities. These complexities could stem from a lack of meaningful minimum and maximum values or the amalgamation of disparate measurement units. Notably, when dealing with bounded activation functions, there is an imperative to scale target values for each data point. This adaptive step enables the seamless integration of target values into an active range. A commonly used linear scaling function is

$$x_s = \frac{x_u - x_{u,\min}}{x_{u,\max} - x_{u,\min}} \times \left(x_{s,\max} - x_{s,\min}\right) + x_{s,\min}, \quad (2)$$

where x_u represents the unscaled value, x_s signifies the scaled value, and $x_{u,\max}$ and $x_{u,\min}$ correspond to the maximum and minimum values of the unscaled range, respectively. $x_{s,\min}$ and $x_{s,\max}$ denote the minimum and maximum values of the scaled range.

There are numerous crucial reasons why RMSE is among the primary measures used in statistical analysis. It is commonly used in streamflow forecasting, ensuring the comparability and understanding of results within the framework of hydrology. RMSE returns findings in the original data units; this behavior boosts interpretability, which is crucial in streamflow forecasting, since flow magnitude truly matters. RMSE's ability to measure prediction errors is critical for practical applications such as water resource management and flood control. RMSE provides for objective model comparison, assisting in selecting the best-suited model. Its application adheres to accepted hydrological principles, allowing for straightforward communication and result comparison with other hydrological specialists. RMSE, mathematically represented as the square root of the average of squared differences between observed and predicted values, is calculated as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}, \tag{3}$$

where n is the total number of data points, Y_i represents the observed (actual) values, and $\hat{Y}_i$ represents the predicted values.

4) FORMULATION OF INPUT DATA USING WAVELET TRANSFORM

The wavelet transform is a time-frequency–localized analysis that extracts information from the hydrological parameters in the form of time series through stretching and shifting processes. The wavelet transform works on a simple principle: it objectively decomposes the hydrological parameters into a new hierarchy that is easier to define and forecast. We employed the discrete wavelet transform in this study, represented as

$$\Psi_{m,n}\left(\frac{t-r}{s}\right) = s_0^{-\frac{m}{2}}\Psi\left(\frac{t - nr_0s_0^m}{s_0^m}\right) \tag{4}$$

where m represents the dilated parameter, n represents the translation parameter, s_0 represents the dilation step (> 1), and r_0 represents the position parameter. The dilation term s_0^m in the preceding equation governs the dilation process in the complex parameter $nr_0s_0^m$.

5) MODEL ARCHITECTURE

a: ARTIFICIAL NEURAL NETWORK (ANN)

An ANN is a parametric model with a collection of parameters, such as weights and biases, that are trainable. It has several hyperparameters that need tuning, such as learning rate and hidden layer size. At its last layer, it contains only a fully connected layer, in which each neuron is connected to all other neurons, as shown in Fig. 2.

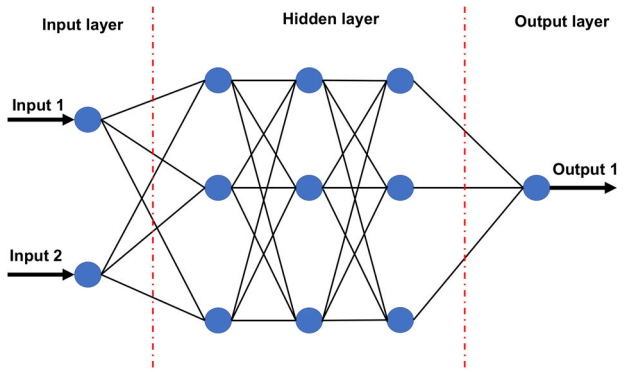

FIGURE 2. Schematic diagram of standard artificial neural network.

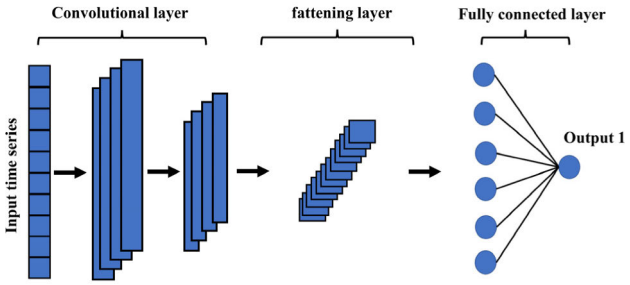

FIGURE 3. Schematic diagram of convolutional neural network.

b: CONVOLUTIONAL NEURAL NETWORK FORECASTING MODEL

While a CNN is a neural network model developed for working with a two-dimensional image dataset, it can work with one-dimensional or three-dimensional data. The fundamental building components of a CNN are its convolutional layers, which serve as the core building blocks of the networks. In the convolutional layer, the hydrological parameters convoluted with the kernel filter, resulting in a feature map. Fig. 3 depicts the model architecture and structure of the forecasting method. The CNN network layer and the fully connected layer comprise the core of the architecture (dense layer).

c: BAYESIAN PROBABILISTIC MODEL

An uncertainty estimation method is necessary to address the questions raised in this study. In this study, the uncertainty of each parameter was assessed using the Bayesian approach. Bayes' rule describes the probability of an event (hypothesis or degree of belief) based on prior knowledge (evidence). Generally, the theorem can be expressed as

$$Posterior\ P(H \mid E) = \frac{Prior\ P(H) \times likelihood\ P(E|H)}{Evidence\ P(E)}, \tag{5}$$

where H is the hypothesis whose probability may be affected by evidence E (data), $P(H|E)$ is the posterior predictive distribution, $P(E)$ is the marginal likelihood, and $P(H)$ is the prior probability. In Bayesian inference, the likelihood function plays a pivotal role. This function contains all

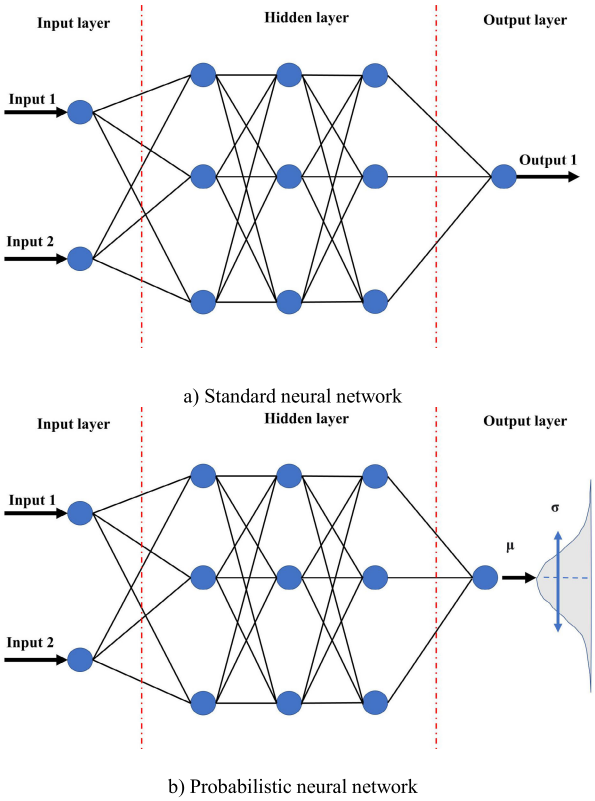

FIGURE 4. Framework for a) standard neural network and b) probabilistic neural network.

of the information needed to evaluate statistical evidence. In hydrology, the likelihood function is frequently built-in based on the assumption that any residual errors are independent and follow a Gaussian distribution. Fig. 4 presents the probabilistic framework used to predict system performance, along with relevant uncertainties.

Probabilistic models $p(t|x, w)$ can be used to address supervised learning problems and model the relationship between the input and output, as represented by $Z(t, x)$, making predictions based on probabilistic reasoning. Here the terms x and t are the input series and the desired output value, respectively, in supervised learning networks. With reference to Bayes' theorem above, the problem can be formulated as

$$Posterior\ P\,(w\,|\,Z) = \frac{PriorP\,(w) \times likelihood\ P(Z|w)}{Evidence\ P\,(Z)}. \tag{6}$$

This probabilistic model can provide a distribution for future data that are likely to occur, known as a predictive posterior distribution (PPD). The PPD incorporates the uncertainty associated with the parameter w. In essence, the PPD of future observations is calculated based on current data by adding the distribution of y over w to the posterior distribution of w over x.

$$\mathrm{PPD}(t, x) = \int \mathrm{P}\,(t\,|\,w)\,\mathrm{P}\,(w\,|\,x)\,\mathrm{d}w \tag{7}$$

Another analytical approximation of posterior distribution is through variational inference, where a variation distribution, denoted as $Q(w|\theta)$, is established to make statistical inferences about the true posterior distribution. A smaller Kullback–Leibler divergence indicates a higher likelihood that the two distributions are similar. The discrepancy between the approximation and actual value can thus be reduced by measuring and then iteratively minimizing the Kullback–Leibler divergence, as follows:

$$KL\,(Q\,|\,P) = Q\,(x)\,log\frac{Q}{P} \tag{8}$$

A confidence interval (CI) characterizes the degree of certainty in the forecasted value. There are many strategies for constructing the CI, offering various choices. Several studies have reviewed CI techniques [56]. The bootstrap methods, such as the approximate bootstrap, percentile-t method, bias-corrected percentile method, and accelerated bias-corrected percentile method, can be employed in many applications and fields.

The percentile-t method or bootstrap-t method can be adopted to estimate the hydrological parameters quantile as suggested by [57] and [58]. An intriguing study by Ghiasi et al. [59] involved the resampling technique; they adopted a modified bootstrap-t method developed from [60] for uncertainty quantification. They employed the modified bootstrap method for resampling distinct training patterns to ensure that the selected training patterns accurately reflected the statistical characteristics of the dataset used in the study. This approach targets mitigation of potential limitations associated with conventional bootstrap techniques when dealing with rare instances. However, to align with the objective of this research, we utilized resampling for CI estimation to thoroughly examine the range of possible results and the level of confidence associated with these outcomes. This procedure necessitated accounting for both aleatoric and epistemic variability in our streamflow forecasting models.

To refer to the percentiles, the data collected were used to estimate the sampling distribution, which in turn yields the CI [61]. Based on the forecast distribution of the series, the confidence interval was formed using the quantiles of the forecast distribution. In this case, we computed the CI at a statistical significance level of 5%.

$$\mathrm{CI} = \mathrm{forecast} + (\mathrm{t_{critical}} - \mathrm{value}) \times (\sigma_{\mathrm{error}}) \tag{9}$$

V. RESULTS AND DISCUSSION

A. MODEL VALIDATION

Statistical and machine learning techniques assume that the individual observations in a dataset are independent and identically distributed (i.i.d.). However, this assumption does not hold in the case of streamflow data, as individual observations are not statistically independent; instead, they depend on previous observations [62]. For example, the flow of a river today is influenced by the flow in the preceding days or weeks due to factors such as rainfall, snowmelt, and watershed characteristics. Streamflow data exhibit significant

autocorrelation, requiring correlation of observed values in a streamflow time series with their corresponding past values. This autocorrelation arises from the continuity and flow dynamics of water in rivers.

To apply cross-validation to time series forecasting, certain modifications are made to better adapt the method to sequential data, one of which is the blocked cross-validation (Bl-CV) strategy [63]. Bl-CV shares similarities with the standard k-fold strategy but differs in that it does not involve a random shuffling of the dataset. Instead, it partitions the data into k subsets of a specific length, creating k consecutive blocks of data observations. Bl-CV maintains the temporal integrity of data within each block of observations, which is critical for streamflow data, since it preserves the temporal order of observations within each block.

However, as mentioned in the study by Cerqueira et al. [64], Bl-CV is most suitable for a time series that is short and stationary. Streamflow forecasting often encounters short-time series data, particularly in regions with limited historical streamflow records. The stationarity of streamflow may vary depending on the interval of analysis. However, over extended periods (e.g., monthly or yearly), extreme events, such as floods and climate change, disrupt the typical streamflow pattern, resulting in nonstationary data.

Therefore, the hold-out method was adopted, acknowledging the temporal nature of time series data. This method involves dividing the data into two sets: a training set, which contains historical observations up to a specific point in time, and a validation set, consisting of data points from beyond the training period.

B. ACCOUNTING FOR TEMPORAL VARIABILITY OF STREAMFLOW

Despite the effectiveness of CNNs in the context of two-dimensional images, a modified CNN was developed and utilized to address the real-world optimization problem adopted in this research study. As a CNN requires working with two-dimensional datasets, the boundary to which CNN extracts the information would be an (N×M) dimension, where N is the width and M the length of each image. The overlays of (N×M) dimension are essential in processing an image by reducing the original size of the image to a form that is easier to process toward a good prediction without losing features. However, this extraction process may disrupt the order in which the model receives the input. Such a disruption is problematic, as the chronological sequence of the time series data is crucial. The proposed CNN was designed to retain temporality while not unraveling the sequence in which the input is fed into the model: the prediction $p(x_{k+1}{:}x_k, x_{k-1}, \ldots x_{k-n})$ omits any future time steps, such as $x_{k+1}, x_{k+2}, \ldots x_{k+T}$. The proposed framework thus does not reduce the number of features of (N×M) dimension, so the reception field would still provide the same coverage (Fig. 5).

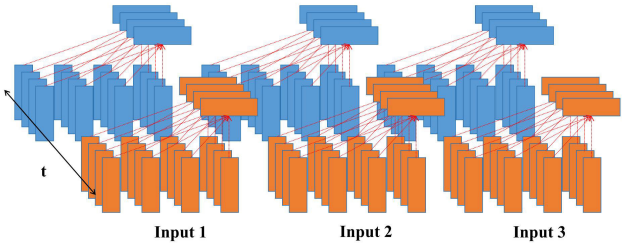

FIGURE 5. Proposed framework of receiving the inputs per time step.

C. SIGNIFICANCE OF WAVELET TRANSFORM IN AI MODEL UTILITY THROUGHOUT THE FORECASTING PHASE

To verify the proposed time series framework, the developed CNN was verified, along with ANN, on a univariate time series dataset, that is, a sequence of observations in temporal order. This developed CNN needed this dataset as a model to learn from in order to anticipate future values. Accurate prediction involves accurate monitoring of the timing and magnitude of streamflow in sustaining the water supply system and reducing the risk of flood. Timing errors of the models are a common problem in neural network rainfall-runoff models. For comparison, naive prediction is a prediction methods involving random and seasonal random walks wherein all forecasts are based solely on the value of the last observation. The lagging effect issue may occur because naive methods do not incorporate sophisticated models or account for dynamic changes in the underlying data. This issue are not limited to the present research, as indicated in Fig. 6, but have also been observed in previous published research. As seen in Fig. 6, the predicted results from CNN and ANN methodologies appear to lag behind the actual data; such lag in the results is a common issue when using lagged variables (historical data) as input features.

For example, Snieder et al. [65] provided a demonstration that lagging effect often occur when dealing with streamflow data characterized by high seasonality. The lagging effects are evident in cases in which predictions resemble naive forecasts, as observed for the Bow River (high seasonality) but not for the Don River (no significant seasonality). Rivers in Malaysia also exhibit high seasonality, primarily due to the influence of the monsoon, making such timing errors a likely outcome for Malaysian data. Similarly, Mehr and Kahya [66] observed that univariate streamflow forecasting using machine learning often leads to low quantitative efficiency, primarily because of delayed predictions. Univariate models may also have restricted predictive power when compared to multivariate models. They might not be suitable for long forecast horizons or scenarios where a comprehensive understanding of the system, considering multiple variables simultaneously, is essential [67]. In contrast, Shu et al. [68] reported that the machine learning algorithms utilized, including CNNs and ANNs, were less affected, by timing errors despite their lack of data preprocessing, which contrasts with findings reported by other studies. This

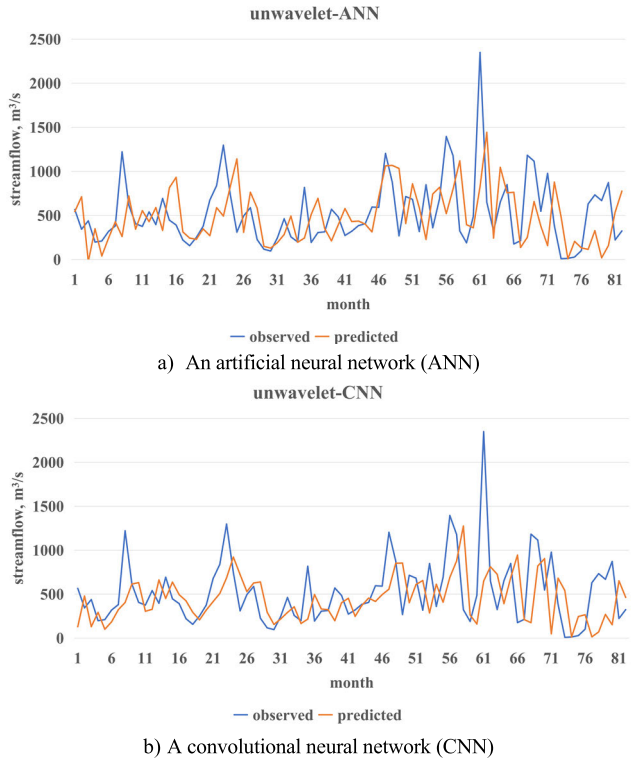

FIGURE 6. Forecasting results using a) an artificial neural network (ANN) and b) a convolutional neural network (CNN).

divergence in findings highlights differing perspectives on the effectiveness of data preprocessing methods.

Further discussion is necessary regarding the shortcomings illustrated in Fig. 6, including overestimating low flows and underestimating high flows. These discrepancies, in part, can be attributed to timing errors, which encompass disparities between projected and actual flow timings. Timing inaccuracy significantly impacts forecast accuracy by altering the timing of flow events. Streamflow forecasting necessarily involves temporal dependence, in which past states influence the current state of streamflow. Markov chains are one method for representing this temporal dependency [69]. Markov chains predict the possibility of shifting from one flow state to another in the future based on previous observations. This modeling method gives additional weight to recent lagged variables for estimating future streamflow, highlighting the impact of current states on the prediction. It is essential to note that while Markov chains can be a valuable tool in specific streamflow forecasting scenarios, they represent just one of several modeling methodologies. The choice of forecasting approach depends on the unique characteristics of the data and the research objectives, with more advanced models and machine learning techniques also being employed.

In this research, based on the ML approaches, the RMSE value for a hybrid wavelet–ANN (WA) model (119.25) was significantly smaller than for a corresponding ANN model without WT (126.88). This result highlights the substantial potential of wavelet-aided hydrological prediction.

The lagging problem, as demonstrated in Fig. 6, shows the critical importance of data preparation in developing effective data-driven models with accurate representative data collection. The quality of data significantly influences the robustness of neural network models. In this context, WA models play a pivotal role by employing data sub-series acquired through multi-resolution analysis as input data. This approach, as illustrated in Figs. 7 and 8, demonstrates how incorporating wavelet components resolves the lagging problem associated with lagged variables used as input features. Notably, the WT used in this research divides the original signal into multiple resolution levels (scales), resulting in low-frequency approximation components (A) and high-frequency detail components (DW). These characteristics empower the models to effectively capture less complex data series, even when dealing with considerably nonlinear datasets. Ultimately, the forecasted streamflow is derived by summing all the decomposed wavelets, providing a comprehensive robust solution for addressing the lagging problem.

It is also essential here to consider the distinction between stationary and nonstationary time series data. Stationarity is a critical factor in data analysis. Stationary time series maintain consistent statistical properties over time, while nonstationary time series can exhibit changing means, variances, or other statistical characteristics, posing more challenges for accurate prediction and modeling. Achieving data stationarity is one of the crucial ways for obtaining reliable insights and consistent forecasting results. This WT transformation allows for the computation of meaningful sample statistics of data in improving the predictiveness of the model

While WT certainly helps, a perspective drawn from the study of Ghaderpour et al. [70], who also demonstrated the utility of machine learning in monthly hydrological (precipitation) forecasting, highlights the challenge of detecting abrupt change. This observation, prompts a consideration of the limitation of the current WT implementation. Our study employed different machine learning techniques but has similar findings, where the abrupt changes of streamflow were not captured completely, the main divergence being in the performance criteria used. They emphasized the importance of exploring alternative data sources, such as for temperature, soil moisture, and wind, which may have direct or indirect causal relationships with streamflow. Historical streamflow data may not accurately capture extreme weather events, such as heat waves, heavy rainfall, or prolonged droughts, which can swiftly and substantially impact streamflow. Temperature and wind data can provide early warnings of such occurrences. Utilizing a diverse set of input variables can also enhance the performance of machine learning and data-driven algorithms. The inclusion of temperature and wind data as predictors alongside historical streamflow data can enhance the predictive capabilities of these models.

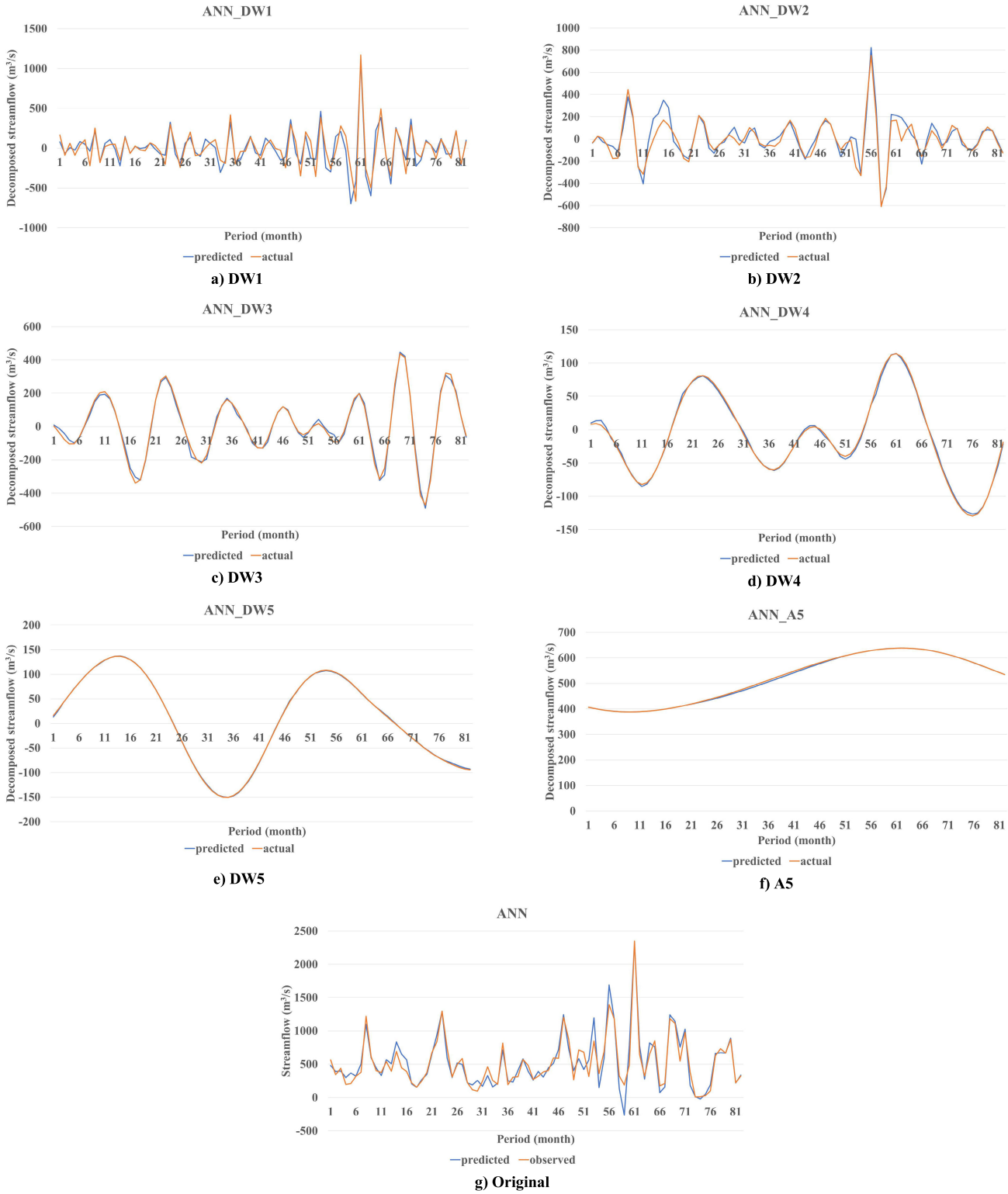

FIGURE 7. Wavelet decomposition series for ANN: individual components (a) DW1, (b) DW2, (c) DW3, (d) DW4, (e) DW5, (f) A5, and (g) original resulting from wavelet analysis.

D. CNN ARCHITECTURAL INFLUENCE

The filter in CNN has a role similar to that of the neurons in ANN, as an N×M filter contains N×M neurons, which perform convolution on the input to form a feature map (output). The full factorial design used in the current study allows the effects of one component to be evaluated at several

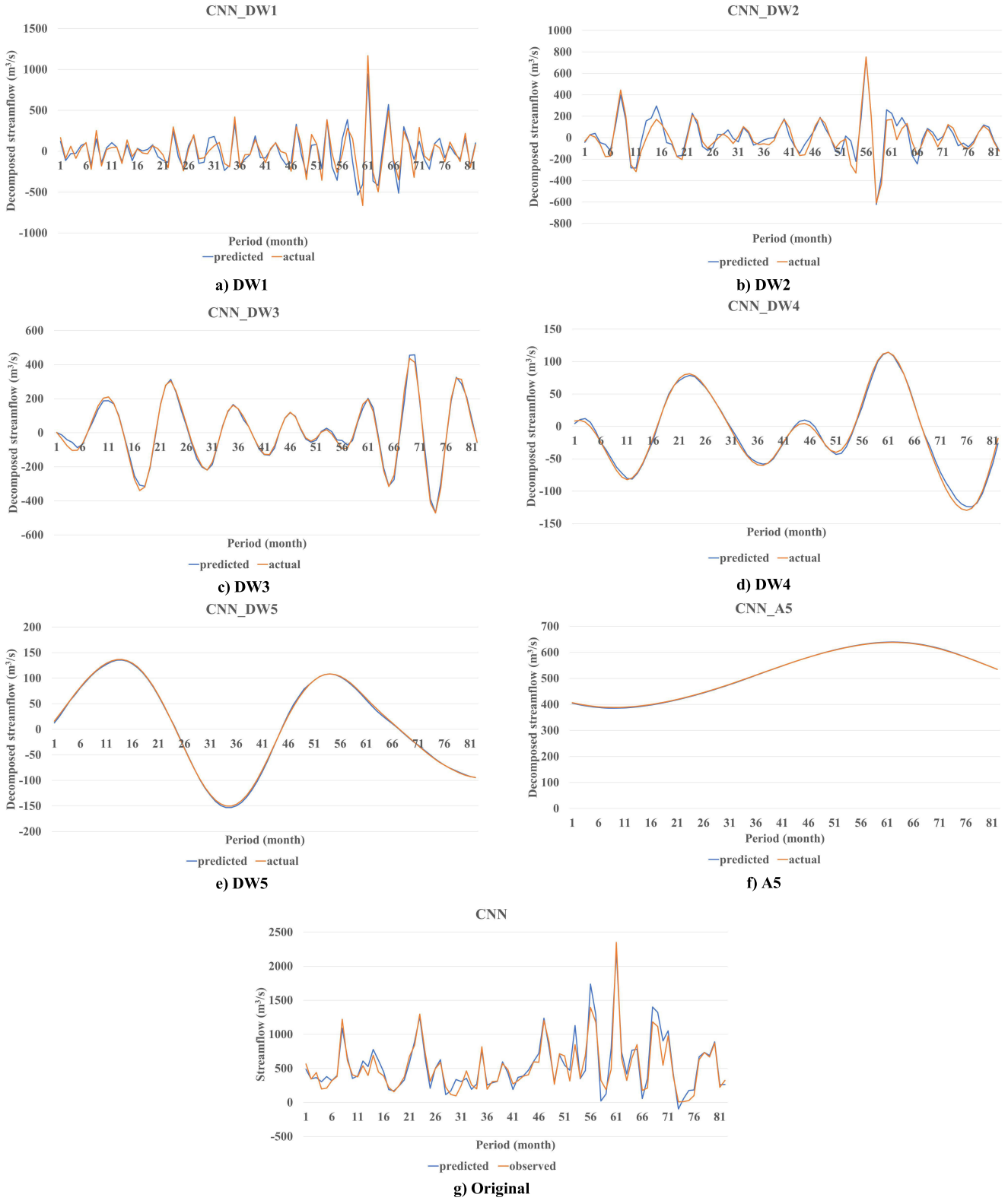

FIGURE 8. Wavelet decomposition series for CNN: individual components (a) DW1, (b) DW2, (c) DW3, (d) DW4, (e) DW5, (f) A5, and (g) original resulting from wavelet analysis.

levels of the other factors, resulting in a wide variety of experimental settings. An alternative method, the Taguchi technique [71], is based on orthogonal arrays and defines how to carry out the smallest number of tests necessary to

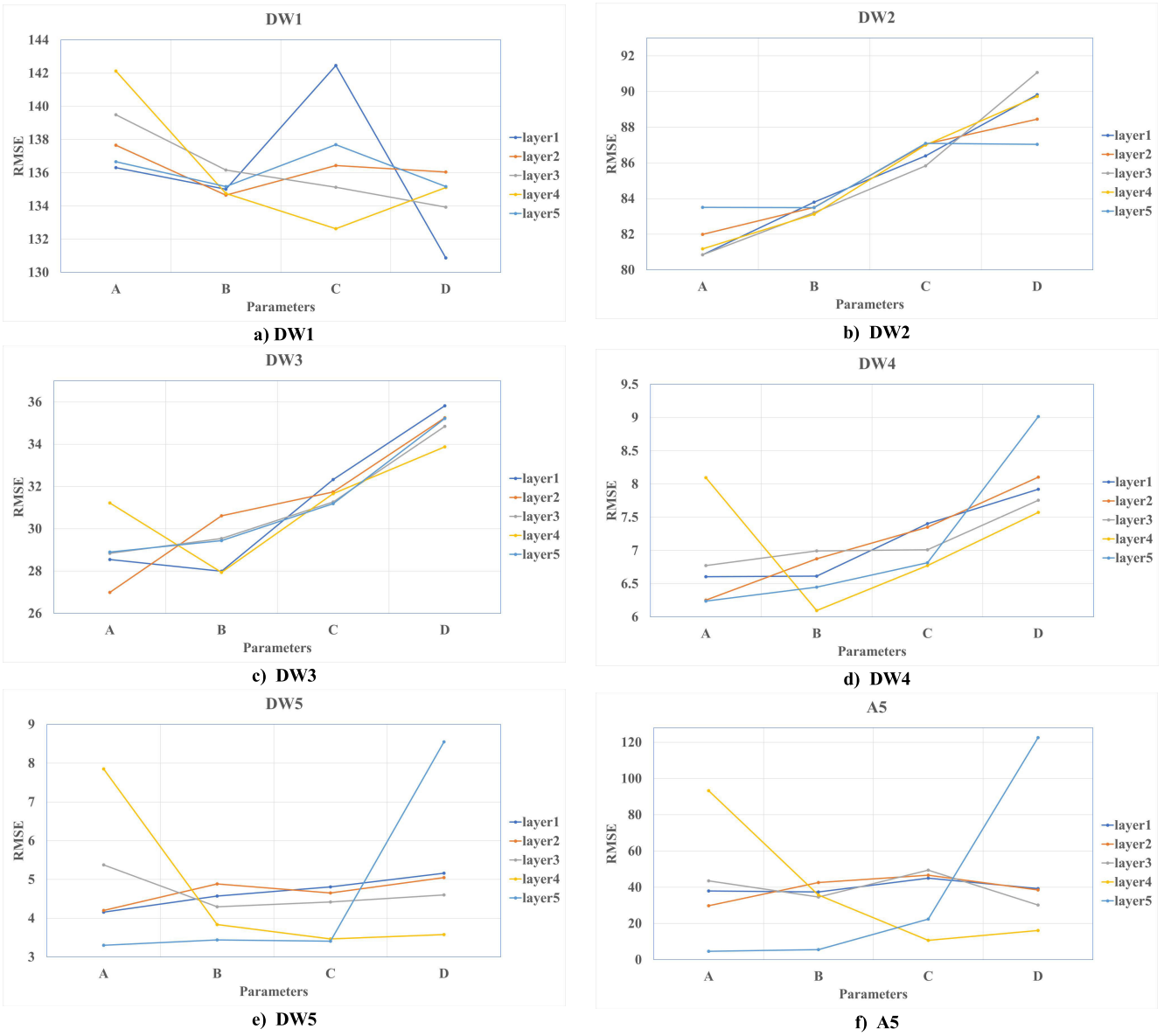

FIGURE 9. Main-effect plots of the decomposed wavelets for the convolutional neural network (CNN) model: (a) DW1, (b) DW2, (c) DW3, (d) DW4, (e) DW5, and (f) A5.

offer comprehensive information for all factors impacting the output parameter. The reason for using a full factorial design rather than the Taguchi method is the higher accuracy of the full factorial array, although the cost of the experiment (including the time to run the experiment) is admittedly higher. On the other hand, the Taguchi method was adopted for ANN, since the computational cost of running each parameter ten times was unsustainably high. The letters A–E are used to denote the components for simplicity. The set of weight values for the four hidden layers in ANN was (20,15,10,5), whereas the set of weight values for the first four hidden layers in CNN was (128,64,32,16), and the last layer of CNN is (20,15,10,5). Main effects plots were used to examine variations around means for each factor level. The main effect graphs for the decomposed wavelets are shown in Fig. 9.

1) PERFORMANCE IMPACT OF PARAMETERS

Analysis of variance (ANOVA) was used to examine the influence of the number of CNN filters on the performance of streamflow model predictions. ANOVA is a reliable approach for extracting useful information from data and is among the most powerful statistical tests [72]. Table 2 presents the ANOVA results for each parameter of the neural network. Each p-value below 0.05 indicates a statistically significant impact of that parameter on the performance metric.

A graphical examination of Pareto charts may also be used to visually assess the amount and relevance of the impact of

TABLE 2. Analysis of variance (ANOVA) of convolutional neural network parameters.

Dw1					
Source	DF	Adj SS	Adj MS	F-value	*p*-value
layer1	3	17,289	5,763.03	29.87	0
layer2	3	1,130	376.75	1.95	0.12
layer3	3	4,371	1,457.12	7.55	0
layer4	3	12,954	4,318.01	22.38	0
layer5	3	1,259	419.71	2.18	0.089
layer1*layer2	9	14,782	1,642.47	8.51	0
layer1*layer3	9	3,674	408.25	2.12	0.026
layer1*layer4	9	864	95.99	0.5	0.877
layer1*layer5	9	1,199	133.24	0.69	0.718
layer2*layer3	9	987	109.63	0.57	0.824
layer2*layer4	9	4,499	499.86	2.59	0.006
layer2*layer5	9	2,532	281.32	1.46	0.159
layer3*layer4	9	2,836	315.07	1.63	0.101
layer3*layer5	9	3,317	368.57	1.91	0.047
layer4*layer5	9	4,969	552.14	2.86	0.002
Dw2					
Source	DF	Adj SS	Adj MS	F-value	*p*-value
layer1	3	11,045	3,681.58	61.06	0
layer2	3	6,966	2,322.15	38.51	0
layer3	3	14,880	4,959.88	82.26	0
layer4	3	11,570	3,856.69	63.96	0
layer5	3	3,301	1,100.4	18.25	0
layer1*layer2	9	1,773	196.98	3.27	0.001
layer1*layer3	9	2,483	275.87	4.58	0
layer1*layer4	9	500	55.52	0.92	0.506
layer1*layer5	9	269	29.89	0.5	0.878
layer2*layer3	9	2,896	321.77	5.34	0
layer2*layer4	9	751	83.42	1.38	0.191
layer2*layer5	9	388	43.08	0.71	0.696
layer3*layer4	9	3,657	406.35	6.74	0
layer3*layer5	9	376	41.73	0.69	0.717
layer4*layer5	9	4,442	493.57	8.19	0
Dw3					
Source	DF	Adj SS	Adj MS	F-value	*p*-value
layer1	3	10,363	3,454.32	29.65	0
layer2	3	8,949	2,982.92	25.6	0
layer3	3	5,740	1913.17	16.42	0
layer4	3	4,824	1,607.97	13.8	0
layer5	3	6,372	2,124.15	18.23	0
layer1*layer2	9	1968	218.66	1.88	0.052
layer1*layer3	9	2,081	231.25	1.98	0.038
layer1*layer4	9	3,690	409.99	3.52	0
layer1*layer5	9	1,741	193.49	1.66	0.094
layer2*layer3	9	1,614	179.32	1.54	0.13
layer2*layer4	9	1,806	200.71	1.72	0.08
layer2*layer5	9	1,429	158.74	1.36	0.201
layer3*layer4	9	3,782	420.19	3.61	0
layer3*layer5	9	713	79.2	0.68	0.728
layer4*layer5	9	11,038	1,226.49	10.53	0
Dw4					
Source	DF	Adj SS	Adj MS	F-value	*p*-value
layer1	3	322.3	107.441	7.46	0
layer2	3	476.7	158.889	11.03	0
layer3	3	144.3	48.101	3.34	0.019
layer4	3	609.6	203.184	14.1	0
layer5	3	1,272.4	424.124	29.44	0
layer1*layer2	9	266.4	29.602	2.05	0.031
layer1*layer3	9	230.5	25.613	1.78	0.068
layer1*layer4	9	294.1	32.683	2.27	0.016
layer1*layer5	9	195	21.669	1.5	0.142
layer2*layer3	9	77.6	8.618	0.6	0.799
layer2*layer4	9	82.3	9.147	0.63	0.768
layer2*layer5	9	298.6	33.174	2.3	0.015
layer3*layer4	9	514.7	57.192	3.97	0
layer3*layer5	9	145.6	16.175	1.12	0.343
layer4*layer5	9	2,733.6	303.738	21.08	0

each parameter (Fig. 10). The bars representing components D, E, DE, and BC, for example, cross the red reference line in

TABLE 2. *(Continued.)* Analysis of variance (ANOVA) of convolutional neural network parameters.

Dw5					
Source	DF	Adj SS	Adj MS	F-value	*p*-value
layer1	3	127.3	42.42	0.73	0.535
layer2	3	103.2	34.4	0.59	0.621
layer3	3	169.3	56.43	0.97	0.407
layer4	3	3,428.4	1,142.79	19.62	0
layer5	3	5,265.4	1,755.14	30.13	0
layer1*layer2	9	724.2	80.47	1.38	0.192
layer1*layer3	9	301.4	33.49	0.57	0.818
layer1*layer4	9	382.5	42.5	0.73	0.682
layer1*layer5	9	203.5	22.61	0.39	0.941
layer2*layer3	9	577	64.11	1.1	0.359
layer2*layer4	9	65.4	7.27	0.12	0.999
layer2*layer5	9	120	13.34	0.23	0.99
layer3*layer4	9	1,541	171.22	2.94	0.002
layer3*layer5	9	818.9	90.99	1.56	0.122
layer4*layer5	9	11,382.4	1,264.71	21.71	0
A5					
Source	DF	Adj SS	Adj MS	F-value	*p*-value
layer1	3	6,492	2,164	0.24	0.868
layer2	3	41,472	13,824	1.54	0.204
layer3	3	64,448	21,483	2.39	0.068
layer4	3	1,066,567	355,522	39.49	0
layer5	3	2,431,690	810,563	90.02	0
layer1*layer2	9	74,149	8,239	0.92	0.511
layer1*layer3	9	69,311	7,701	0.86	0.565
layer1*layer4	9	66,545	7,394	0.82	0.597
layer1*layer5	9	7,139	793	0.09	1
layer2*layer3	9	71,248	7,916	0.88	0.543
layer2*layer4	9	96,892	10,766	1.2	0.294
layer2*layer5	9	222,553	24,728	2.75	0.004
layer3*layer4	9	100,375	11,153	1.24	0.267
layer3*layer5	9	102,223	11,358	1.26	0.254
layer4*layer5	9	3,940,413	437,824	48.63	0

the Pareto graph for decomposed wavelet A5, indicating that these components are statistically significant. These factors exhibit statistical significance at a $p < 0.05$ level.

In addition, some of the interactions among parameters are significant, as indicated in Table 2. For example, when forecasting the dw1 component, the magnitudes of the effect estimates show that A is by far the most dominant factor. D plays the next most critical role, followed by C. Fig. 8 also shows the effects of interactions between the parameters on model performance. Although parameters such as B and E do not play important roles individually, their interactions with the other parameters can obscure the main effects. The weights of these intermediate layers are interconnected, meaning that altering one initial layer affects subsequent layers in the series. Hence, selecting the best hyperparameters setting based solely on the main-effect plot for streamflow forecasting is not advisable. In this study, the best set of factors was determined from the designs that provided the lowest average RMSE (the full factorial design) and the best setting, as shown in Table 3.

2) STATISTICAL EVALUATION: CNN VS. ANN

A comparison between ANN and CNN was made to test the effectiveness of deep learning in streamflow prediction. Table 4 complements the information provided in

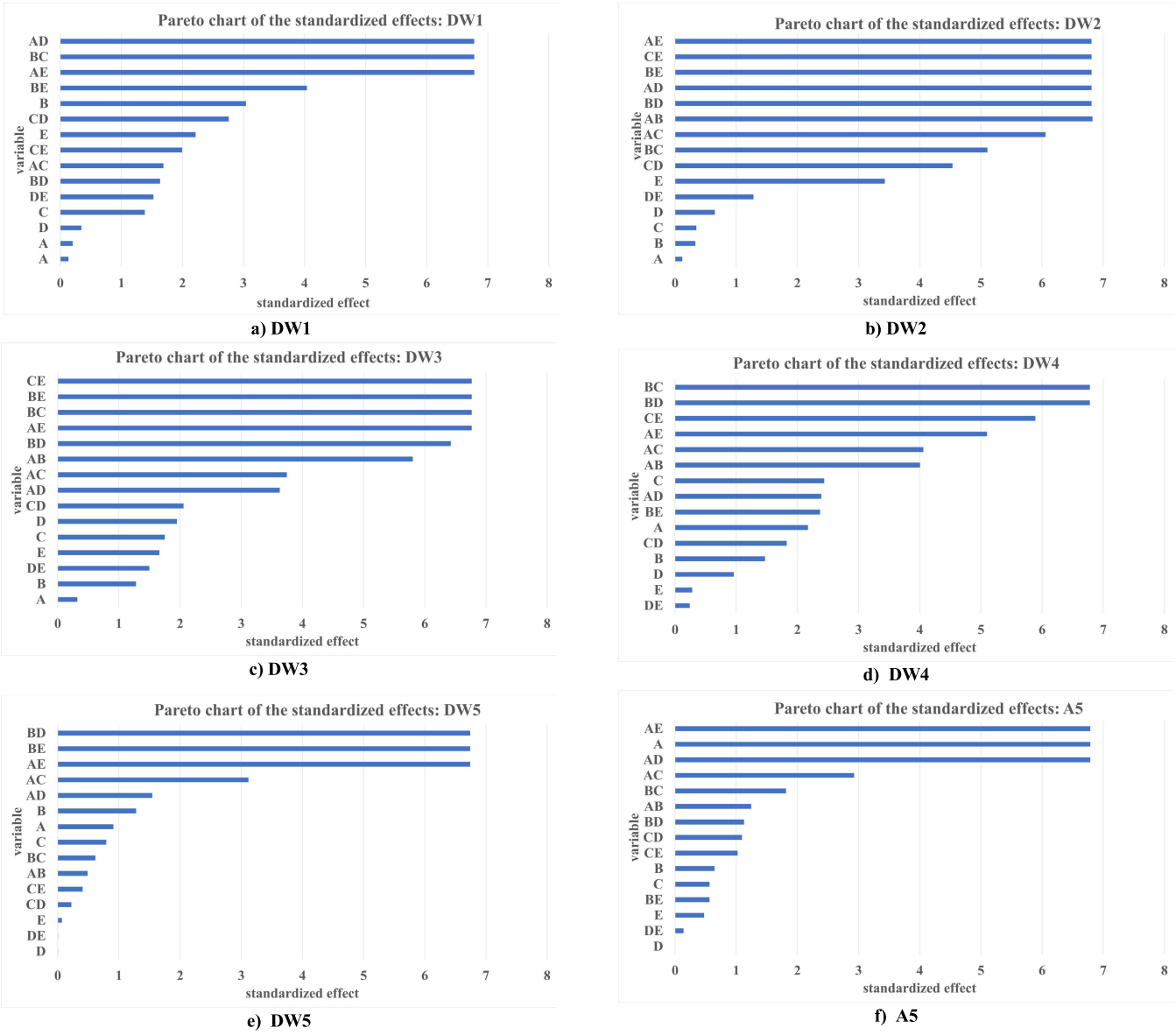

FIGURE 10. Pareto charts for the effects of each of the decomposed wavelets: (a) dw1, (b) dw2, (c) dw3, (d) dw4, (e) dw5, and (f) A5.

Figs. 7 and 8. According to Table 4, the acquired data indicate that ANN was more stochastic than CNN, the outcomes fluctuating from run to run. Although both strategies perform well in forecasting streamflow time series, each offers distinct advantages. Because of its unique feature extraction capabilities, CNN may offer good processing efficiency and solutions that are closer to globally optimal. While CNN offers efficient global solutions, ANN prioritizes precision over stability. For the more sophisticated tasks, CNN may outperform ANN. For example, CNN may give better results than ANN for predicting the highest-frequency component (dw1); meanwhile, ANN readily predicts the lowest-frequency component (A5), outperforming CNN. CNN exhibits large numbers of extraction features, causing overfitting of low-frequency components such as A5.

It is worth noting that for real-time forecasting, CNN can be a better for determining reservoir operating rules than ANN, given the stochasticity of the ANN process. When employing a predictive neural network model in real-time, updating is required owing to fluctuations in the predictions, which may be influenced by dynamic real-world changes rather than being merely a statistical technicality. Because the underlying state of the problem may have changed, the impact of retraining a model with new data may cause predictiveness to decrease. As a result, when deploying the model in real-time, it becomes essential to determine whether new input patterns resemble the historical data used to calibrate the model. Given the lengthy time necessary to train the new model, the CNN model is to be preferred.

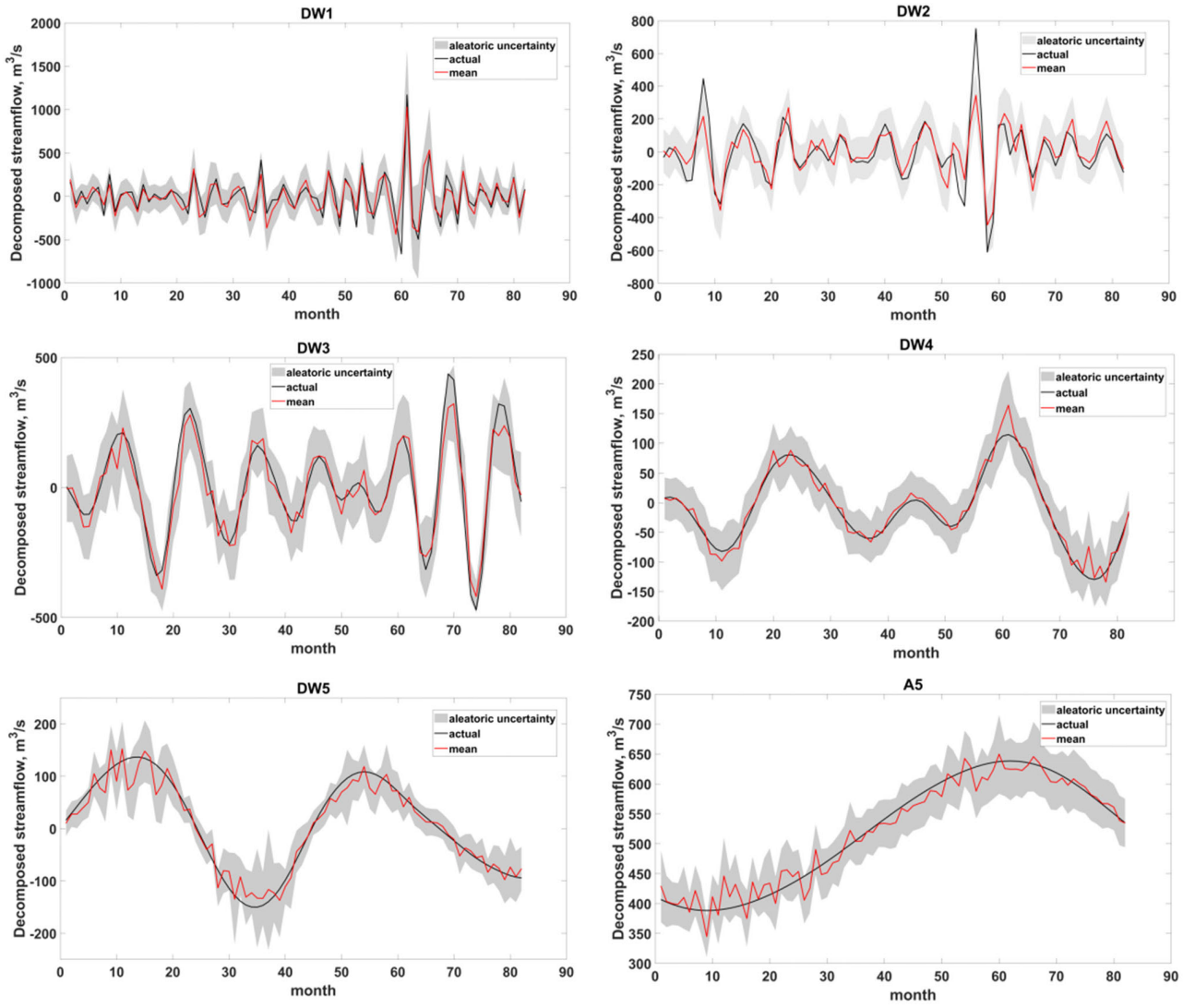

FIGURE 11. Aleatoric uncertainty analysis of the time series.

E. ADDRESSING VARIABILITY IN THE PREDICTION PROCESS

To cope with aleatory and epistemic variability, probabilistic approaches are crucial for producing useful predictions. In this investigation, the Bayes' theorem approach was used to predict variability. The negative log-likelihood term, a cost function that measures how well or poorly the model performs, with lower values indicating higher performance, was used to calculate the loss for each machine learning model. By minimizing the negative log-likelihood, the cost function achieves maximization, which deviates from the conventional minimization. This unique optimization is enabled by representing the data as a probability distribution, thus facilitating the estimation of the probability of selecting the correct label. Using a popular alternative approach, a CI may also be used to assess the variability in the forecast. A wide confidence interval implies poor forecast accuracy and vice versa. In this way, the prediction intervals may be used to assess the amount of variability in the projections. Unlike conventional prediction, which anticipates solely point-by-point values, probabilistic prediction forecasting contains variability information.

1) PREDICTIVE VARIABILITY OWING TO ALEATORIC AND EPISTEMIC VARIABLES

The forecasting procedure grows more sophisticated with increasing variability of the time series. Aleatoric variability is a type of variability that indicates the underlying heterogeneity of the parameters. The mean and standard deviation of the series are required to produce the aleatoric prediction. According to Fig. 11, the red and black lines show the actual and anticipated mean of the streamflow values, respectively, while the shaded region reflects the output variability. The existence of the mean and standard deviation of the series in this form might convey information about the overall trend and variability of the series.

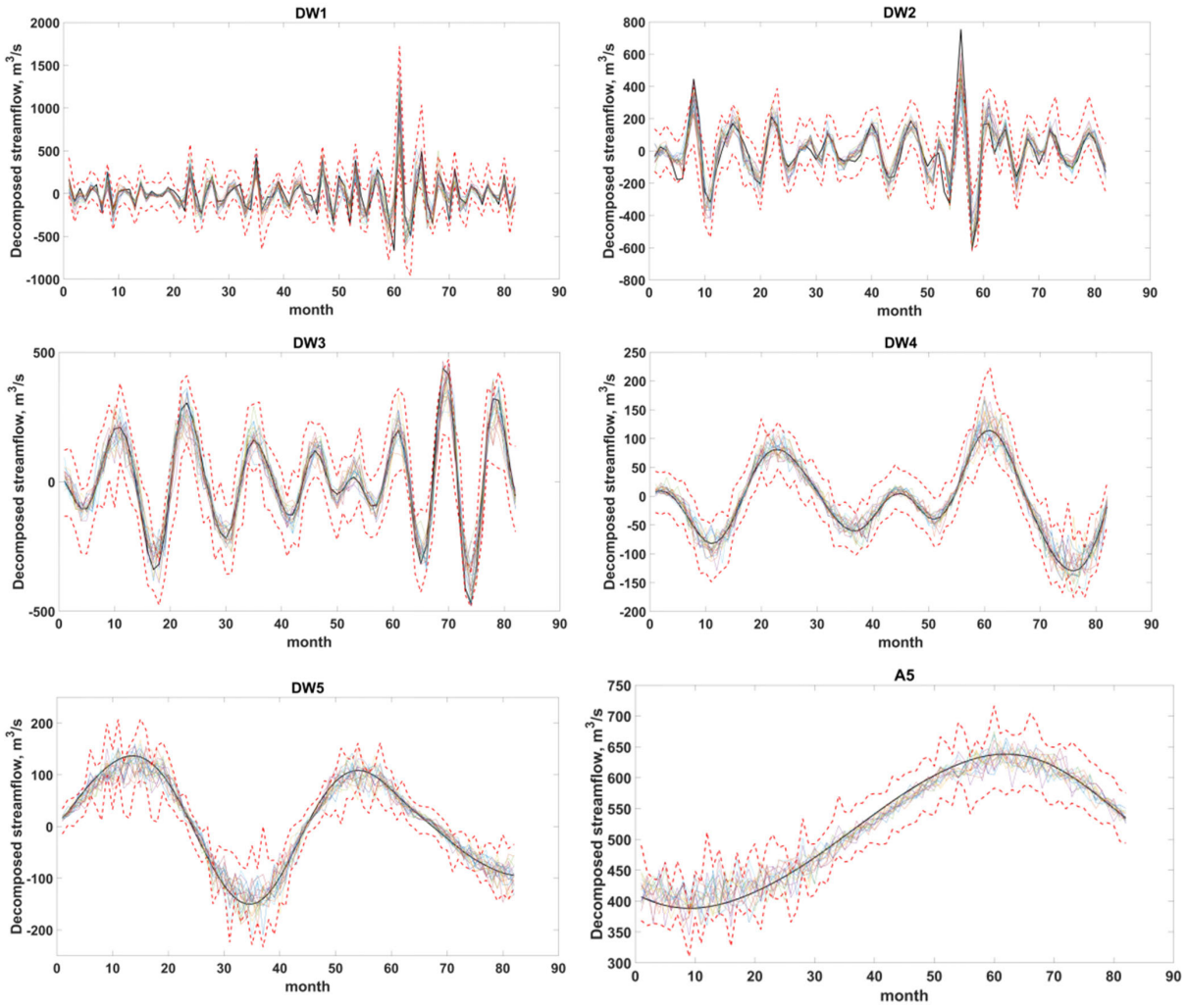

FIGURE 12. Epistemic uncertainty analysis of the time series.

TABLE 3. Numbers of filters in convolutional neural network (CNN) and numbers of filters in artificial neural network (ANN).

CNN	Layer1	Layer2	Layer3	Layer4	Layer5 (for CNN only)
Dw1	128	16	128	32	15
Dw2	128	64	64	32	15
Dw3	128	128	128	128	5
Dw4	128	128	64	32	15
Dw5	16	128	128	128	15
A5	128	128	16	32	20
ANN	Layer 1	Layer 2	Layer 3	Layer 4	
Dw1	10	5	10	20	
Dw2	20	20	15	10	
Dw3	10	15	20	10	
Dw4	20	5	20	15	
Dw5	5	20	10	15	
A5	5	15	15	20	

TABLE 4. Root-mean-square error (RMSE) for convolutional neural network (CNN) and artificial neural network (ANN) in Streamflow forecasting.

Type		RMSE					
		Dw1	Dw2	Dw3	Dw4	Dw5	A5
CNN		97.32	60.96	17.34	4.14	2.01	2.95
ANN	Maximum	143.95	70.44	17.21	2.81	1.28	2.96
	Minimum	90.74	57.37	16.46	2.64	1.00	2.82
	Mean	116.65	65.81	16.907	2.75	1.12	2.92

Correspondingly, when modeling streamflow forecasting variability, we encounter epistemic variability, resulting in a broader range of outcomes, as shown in Fig. 12, that reflects time series volatility.The epistemic variability for continuous random variables is formed using multiple probability density functions and employing the Flipout layers, such as Convolution1DFlipout. This procedure samples from the entire space of potential ensembles rather than separately sampling each model in an ensemble.

The expanded range of outcomes produced by the probabilistic neural network model had significant implications for the accuracy and reliability of streamflow predictions. This finding indicates that the intrinsic variability of the time series could not be adequately accounted for without a thorough grasp of all impacting variables. Projections of individual variables provided precision, but the integration of epistemic variability introduced a layer of uncertainty, resulting in a broader range of anticipated streamflow values. The results were achieved by applying the Monte Carlo approximation to forecast the overall mean across a finite number of trials. The Monte Carlo approach enabled accurate portrayal of the epistemic limits by directly accounting for uncertainty in the modeling process.

Although the precise insights from the present research are limited, future studies should prioritize reducing epistemic variability. Such a reduction can be achieved through measures such as enhancing data collection and implementing comprehensive environmental monitoring. While epistemic variability broadens the range of expected outcomes, it also opens up new avenues for study and innovation in the field of streamflow forecasting. It is crucial to note that this variability is not an artificial construct but rather an accurate reflection of the inherent imprecision in the real-world conditions, thereby necessitating adaptive approaches in forecasting methodologies.

In theory, the addition of uncertainty knowledge makes a Bayesian approach preferable to a deterministic one. It provides a trustworthy approach for reaching a conclusion when some prior knowledge is absent or foggy, eliminating the need to estimate values for unknown features. Noise cannot be included in any approach that is deterministic, since noise is inherently non-deterministic. Accordingly, stochastic approaches such as Bayesian methods should be employed to address noise-contaminated signals rather than deterministic methods.

VI. CONCLUSION AND RECOMMENDATIONS

This paper presents and analyzes the performance of a strategy for streamflow time-series forecasting based on CNN coupled with WT. A preliminary evaluation demonstrated that forecasting streamflow series utilizing only lagged variables resulted in the trailing of the expected outcome behind the actual data. This finding validated the need for WT. A comparative analysis was also performed to assess the complicated interplay between spatially and temporally varying streamflow impacts and various complexity components (decomposed wavelets). A deep learning CNN algorithm was adopted and tested against ANN to investigate the potential of CNN in hydrological forecasting. According to the findings, CNN performed less stochastically than ANN due to its ability to extract features, rendering CNN less susceptible to changes in input, in particular spatial and temporal variability. However, the simple structure exhibited by some time series produced an overfitting issue for CNN. For example, the forecasting of the highest-frequency component (DW1) produced a better result using CNN than using ANN. On the other hand, ANN easily predicted the lowest-frequency component (A5) and outperformed CNN because of overfitting by the large number of extraction features under CNN.

A probabilistic neural network was then employed to illustrate the difference in prediction produced by building aleatoric and epistemic factors into artificial intelligence models. This network made predictions the same way as the previous non-Bayesian network, but it additionally incorporated a parameter that measured the overall degree of uncertainty. Since the uncertainty was caused by parameter variability or by noise in the data, revealing the mean and variance of the anticipated values increased the reliability of the model, lowering the uncertainty inherent in the forecasting process. The Bayesian probabilistic neural network offers various advantages over typical deterministic ANNs, including the ability to forecast hydrological uncertainty. Reasonable estimates of the forecast uncertainty of hydrological simulations are required for decision-making in real-time water resource management scenarios.

Recognizing changes in streamflow regimes is equally crucial because these shifts can significantly impact the accuracy and relevance of forecasting results. Regime modifications might increase uncertainty when evaluating water-associated risks; unexpected floods or droughts are more likely when streamflow patterns change quickly than forecasting models can fully account for. These unanticipated occurrences may have detrimental effects on infrastructure development and catastrophe preparedness, providing a substantial challenge to overcome in hydrological forecasting. At present, streamflow predictions from our Bayesian approach are available only one day in advance. Subsequent research on the proposed approach might go deeper into developing confidence intervals for forecasting made more than a day in advance by using different time intervals.

REFERENCES

[1] P. S. Fabian, H.-H. Kwon, M. Vithanage, and J.-H. Lee, "Modeling, challenges, and strategies for understanding impacts of climate extremes (droughts and floods) on water quality in Asia: A review," *Environ. Res.*, vol. 225, May 2023, Art. no. 115617.

[2] N. T. Anh, L. D. Can, N. T. Nhan, B. Schmalz, and T. L. Luu, "Influences of key factors on river water quality in urban and rural areas: A review," *Case Stud. Chem. Environ. Eng.*, vol. 8, Dec. 2023, Art. no. 100424.

[3] R. Patil, Y. Wei, D. Pullar, and J. Shulmeister, "Effects of change in streamflow patterns on water quality," *J. Environ. Manage.*, vol. 302, Jan. 2022, Art. no. 113991.

[4] I. D. Goodwin, M. Ribó, and T. Mortlock, "Coastal sediment compartments, wave climate and centennial-scale sediment budget," in *Sandy Beach Morphodynamics*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 615–640.

[5] T. de Haas, W. Nijland, S. M. de Jong, and B. W. McArdell, "How memory effects, check dams, and channel geometry control erosion and deposition by debris flows," *Sci. Rep.*, vol. 10, no. 1, p. 14024, Aug. 2020.

[6] R. C. Grabowski, K. Vercruysse, I. Holman, A. Azhoni, B. Bala, V. Shankar, J. Beale, S. Mukate, A. Poddar, J. Peng, and J. Meersmans, "The land–river interface: A conceptual framework of environmental process interactions to support sustainable development," *Sustainability Sci.*, vol. 17, no. 4, pp. 1677–1693, Jul. 2022.

[7] W. Liu, S. Park, R. T. Bailey, E. Molina-Navarro, H. E. Andersen, H. Thodsen, A. Nielsen, E. Jeppesen, J. S. Jensen, J. B. Jensen, and D. Trolle, "Quantifying the streamflow response to groundwater abstractions for irrigation or drinking water at catchment scale using SWAT and SWAT-MODFLOW," *Environ. Sci. Eur.*, vol. 32, no. 1, pp. 1–25, Dec. 2020.

[8] S. Sharma, G. Raj Ghimire, and R. Siddique, "Machine learning for postprocessing ensemble streamflow forecasts," *J. Hydroinformatics*, vol. 25, no. 1, pp. 126–139, Jan. 2023.

[9] M. H. Mughal, Z. A. Shaikh, K. Ali, S. Ali, and S. Hassan, "IPFS and blockchain based reliability and availability improvement for integrated rivers' streamflow data," *IEEE Access*, vol. 10, pp. 61101–61123, 2022.

[10] D. Goodarzi, K. S. Lari, E. Khavasi, and S. Abolfathi, "Large eddy simulation of turbidity currents in a narrow channel with different obstacle configurations," *Sci. Rep.*, vol. 10, no. 1, p. 12814, 2020.

[11] D. Goodarzi, A. Mohammadian, J. Pearson, and S. Abolfathi, "Numerical modelling of hydraulic efficiency and pollution transport in waste stabilization ponds," *Ecol. Eng.*, vol. 182, Sep. 2022, Art. no. 106702.

[12] M. Mahdian, M. Hosseinzadeh, S. M. Siadatmousavi, Z. Chalipa, M. Delavar, M. Guo, S. Abolfathi, and R. Noori, "Modelling impacts of climate change and anthropogenic activities on inflows and sediment loads of wetlands: Case study of the anzali wetland," *Sci. Rep.*, vol. 13, no. 1, p. 5399, Apr. 2023.

[13] M. A. Ghorbani, R. C. Deo, S. Kim, M. H. Kashani, V. Karimi, and M. Izadkhah, "Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia," *Soft Comput.*, vol. 24, no. 16, pp. 12079–12090, Aug. 2020.

[14] Y. Zhang, L. Jin, Z. Zhang, X. Li, Q. Liu, and H. Wang, "SF-ANN: Leveraging structural features with an attention neural network for candidate fact ranking," *Int. J. Speech Technol.*, vol. 52, no. 5, pp. 5841–5856, Mar. 2022.

[15] A. Sahoo, S. Samantaray, and D. K. Ghose, "Stream flow forecasting in mahanadi river basin using artificial neural networks," *Proc. Comput. Sci.*, vol. 157, no. 9, pp. 168–174, 2019.

[16] S. Samantaray, O. Tripathy, A. Sahoo, and D. K. Ghose, "Rainfall forecasting through ANN and SVM in Bolangir Watershed, India," in *Proc. 3rd Int. Conf. Smart Comput. Inform.* Singapore: Springer, 2020, vol. 1, vol. 1, pp. 767–774.

[17] R. Bala, R. Prasad, and V. P. Yadav, "Quantification of urban heat intensity with land use/land cover changes using Landsat satellite data over urban landscapes," *Theor. Appl. Climatol.*, vol. 145, nos. 1–2, pp. 1–12, Jul. 2021.

[18] L. Dong, D. Fang, X. Wang, W. Wei, R. Damaševičius, R. Scherer, and M. Woźniak, "Prediction of streamflow based on dynamic sliding window LSTM," *Water*, vol. 12, no. 11, p. 3032, Oct. 2020.

[19] M. Cheng, F. Fang, T. Kinouchi, I. M. Navon, and C. C. Pain, "Long lead-time daily and monthly streamflow forecasting using machine learning methods," *J. Hydrol.*, vol. 590, Nov. 2020, Art. no. 125376.

[20] W. L. Hakim, A. S. Nur, F. Rezaie, M. Panahi, C.-W. Lee, and S. Lee, "Convolutional neural network and long short-term memory algorithms for groundwater potential mapping in Anseong, South Korea," *J. Hydrol., Regional Stud.*, vol. 39, Feb. 2022, Art. no. 100990.

[21] J. E. Santos, D. Xu, H. Jo, C. J. Landry, M. Prodanović, and M. J. Pyrcz, "PoreFlow-Net: A 3D convolutional neural network to predict fluid flow through porous media," *Adv. Water Resour.*, vol. 138, Apr. 2020, Art. no. 103539.

[22] J. Donnelly, S. Abolfathi, J. Pearson, O. Chatrabgoun, and A. Daneshkhah, "Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model," *Water Res.*, vol. 225, Oct. 2022, Art. no. 119100.

[23] K. Khosravi, F. Rezaie, J. R. Cooper, Z. Kalantari, S. Abolfathi, and J. Hatamiafkoueieh, "Soil water erosion susceptibility assessment using deep learning algorithms," *J. Hydrol.*, vol. 618, Mar. 2023, Art. no. 129229.

[24] R. Noori, B. Ghiasi, S. Salehi, M. E. Bidhendi, A. Raeisi, S. Partani, R. Meysami, M. Mahdian, M. Hosseinzadeh, and S. Abolfathi, "An efficient data driven-based model for prediction of the total sediment load in rivers," *Hydrology*, vol. 9, no. 2, p. 36, Feb. 2022.

[25] S. Borzooei, R. Teegavarapu, S. Abolfathi, Y. Amerlinck, I. Nopens, and M. C. Zanetti, "Impact evaluation of wet-weather events on influent flow and loadings of a water resource recovery facility," in *New Trends in Urban Drainage Modelling*. Cham, Switzerland: Springer, 2019, pp. 706–711.

[26] S. Borzooei, R. Teegavarapu, S. Abolfathi, Y. Amerlinck, I. Nopens, and M. C. Zanetti, "Data mining application in assessment of weather-based influent scenarios for a WWTP: Getting the most out of plant historical data," *Water, Air, Soil Pollut.*, vol. 230, no. 1, pp. 1–12, Jan. 2019.

[27] Z.-K. Feng, W.-J. Niu, T.-H. Zhang, W.-C. Wang, and T. Yang, "Deriving hydropower reservoir operation policy using data-driven artificial intelligence model based on pattern recognition and metaheuristic optimizer," *J. Hydrol.*, vol. 624, Sep. 2023, Art. no. 129916.

[28] J. Lee, H. Koh, and H. J. Choe, "Learning to trade in financial time series using high-frequency through wavelet transformation and deep reinforcement learning," *Int. J. Speech Technol.*, vol. 51, no. 8, pp. 6202–6223, Aug. 2021.

[29] S. J. Hadi, M. Tombul, S. Q. Salih, N. Al-Ansari, and Z. M. Yaseen, "The capacity of the hybridizing wavelet transformation approach with data-driven models for modeling monthly-scale streamflow," *IEEE Access*, vol. 8, pp. 101993–102006, 2020.

[30] D. Liu, W. Jiang, L. Mu, and S. Wang, "Streamflow prediction using deep learning neural network: Case study of Yangtze river," *IEEE Access*, vol. 8, pp. 90069–90086, 2020.

[31] E. Ghaderpour, M. S. Zaghloul, H. Dastour, A. Gupta, G. Achari, and Q. K. Hassan, "Least-squares triple cross-wavelet and multivariate regression analyses of climate and river flow in the Athabasca river basin," *J. Hydrometeorology*, vol. 24, no. 10, pp. 1883–1900, Oct. 2023.

[32] K. L. Chong, S. H. Lai, Y. Yao, A. N. Ahmed, W. Z. W. Jaafar, and A. El-Shafie, "Performance enhancement model for rainfall forecasting utilizing integrated wavelet-convolutional neural network," *Water Resour. Manage.*, vol. 34, no. 8, pp. 2371–2387, Jun. 2020.

[33] J. Wu and Z. Wang, "A hybrid model for water quality prediction based on an artificial neural network, wavelet transform, and long short-term memory," *Water*, vol. 14, no. 4, p. 610, Feb. 2022.

[34] G. Algan and I. Ulusoy, "Image classification with deep learning in the presence of noisy labels: A survey," *Knowl.-Based Syst.*, vol. 215, Mar. 2021, Art. no. 106771.

[35] H. Gupta, H. Varshney, T. K. Sharma, N. Pachauri, and O. P. Verma, "Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3073–3087, Aug. 2022.

[36] R. Sen, H.-F. Yu, and I. S. Dhillon, "Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, no. 12, 2019.

[37] A. Zeroual, F. Harrou, A. Dairi, and Y. Sun, "Deep learning methods for forecasting COVID-19 time-series data: A comparative study," *Chaos, Solitons Fractals*, vol. 140, Nov. 2020, Art. no. 110121.

[38] D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi, and A. Jamil, "A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin," *Earth Sci. Informat.*, vol. 13, no. 3, pp. 915–927, Sep. 2020.

[39] S. P. Van, H. M. Le, D. V. Thanh, T. D. Dang, H. H. Loc, and D. T. Anh, "Deep learning convolutional neural network in rainfall–runoff modelling," *J. Hydroinformatics*, vol. 22, no. 3, pp. 541–561, May 2020.

[40] M. Loga, M. Piniewski, and P. Marcinkowski, "Bayesian decision tables for estimation of risk of water management decisions based on uncertain surface water status: A case study of a Polish catchment," *Environ. Sci. Eur.*, vol. 34, no. 1, pp. 1–16, Dec. 2022.

[41] N. Sun, S. Zhang, T. Peng, J. Zhou, and X. Sun, "A composite uncertainty forecasting model for unstable time series: Application of wind speed and streamflow forecasting," *IEEE Access*, vol. 8, pp. 209251–209266, 2020.

[42] N. Seddon, A. Chausson, P. Berry, C. A. J. Girardin, A. Smith, and B. Turner, "Understanding the value and limits of nature-based solutions to climate change and other global challenges," *Phil. Trans. Roy. Soc. B, Biol. Sci.*, vol. 375, no. 1794, Mar. 2020, Art. no. 20190120.

[43] C. Geffray, A. Gerschenfeld, P. Kudinov, I. Mickus, M. Jeltsov, K. Kööp, D. Grishchenko, and D. Pointer, "Verification and validation and uncertainty quantification," in *Thermal Hydraulics Aspects of Liquid Metal Cooled Nuclear Reactors*. Amsterdam, The Netherlands: Elsevier, 2019, pp. 383–405.

[44] M. M. Rajabi, "Review and comparison of two meta-model-based uncertainty propagation analysis methods in groundwater applications: Polynomial chaos expansion and Gaussian process emulation," *Stochastic Environ. Res. Risk Assessment*, vol. 33, no. 2, pp. 607–631, Feb. 2019.

[45] R. M. Cooke, D. Marti, and T. Mazzuchi, "Expert forecasting with and without uncertainty quantification and weighting: What do the data say?" *Int. J. Forecasting*, vol. 37, no. 1, pp. 378–387, Jan. 2021.

[46] A. Abbaszadeh Shahri, C. Shan, and S. Larsson, "A novel approach to uncertainty quantification in groundwater table modeling by automated predictive deep learning," *Natural Resour. Res.*, vol. 31, no. 3, pp. 1351–1373, Jun. 2022.
[47] E. Hüllermeier and W. Waegeman, "Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods," *Mach. Learn.*, vol. 110, no. 3, pp. 457–506, Mar. 2021.
[48] M. Kowsari, B. Halldorsson, B. Hrafnkelsson, J. P. Snæbjörnsson, and S. Jónsson, "Calibration of ground motion models to Icelandic peak ground acceleration data using Bayesian Markov chain Monte Carlo simulation," *Bull. Earthq. Eng.*, vol. 17, no. 6, pp. 2841–2870, Jun. 2019.
[49] M. Lu, Q. Hou, S. Qin, L. Zhou, D. Hua, X. Wang, and L. Cheng, "A stacking ensemble model of various machine learning models for daily runoff forecasting," *Water*, vol. 15, no. 7, p. 1265, Mar. 2023.
[50] K. L. Chong, Y. F. Huang, C. H. Koo, M. Sherif, A. N. Ahmed, and A. El-Shafie, "Investigation of cross-entropy-based streamflow forecasting through an efficient interpretable automated search process," *Appl. Water Sci.*, vol. 13, no. 1, p. 6, Jan. 2023.
[51] S. A. R. Al-Obaidi, D. Zabihzadeh, and H. Hajiabadi, "Robust metric learning based on the rescaled Hinge loss," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 11, pp. 2515–2528, Nov. 2020.
[52] T. Kim, J. Oh, N. Kim, S. Cho, and S.-Y. Yun, "Comparing Kullback–Leibler divergence and mean squared error loss in knowledge distillation," 2021, *arXiv:2105.08919*.
[53] S. Huang and Q. Wu, "Robust pairwise learning with Huber loss," *J. Complex.*, vol. 66, no. 10, 2021, Art. no. 101570.
[54] Q. Wang, Y. Ma, K. Zhao, and Y. Tian, "A comprehensive survey of loss functions in machine learning," *Ann. Data Sci.*, vol. 9, no. 2, pp. 187–212, Apr. 2022.
[55] Y. Liu, T. Dillon, W. Yu, W. Rahayu, and F. Mostafa, "Missing value imputation for industrial IoT sensor data with large gaps," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6855–6867, Aug. 2020.
[56] T. J. Diciccio and J. P. Romano, "A review of bootstrap confidence intervals," *J. Roy. Stat. Soc. B, Stat. Methodol.*, vol. 50, no. 3, pp. 338–354, 1988.
[57] P. K. Dunn, "Bootstrap confidence intervals for predicted rainfall quantiles," *Int. J. Climatol., J. Roy. Meteorol. Soc.*, vol. 21, no. 1, pp. 89–94, 2001.
[58] P. Naveau, A. Hannart, and A. Ribes, "Statistical methods for extreme event attribution in climate science," *Annu. Rev. Statist. Appl.*, vol. 7, no. 1, pp. 89–110, Mar. 2020.
[59] B. Ghiasi, R. Noori, H. Sheikhian, A. Zeynolabedin, Y. Sun, C. Jun, M. Hamouda, S. M. Bateni, and S. Abolfathi, "Uncertainty quantification of granular computing-neural network model for prediction of pollutant longitudinal dispersion coefficient in aquatic streams," *Sci. Rep.*, vol. 12, no. 1, p. 4610, Mar. 2022.
[60] R. Noori, A. Mirchi, F. Hooshyaripor, R. Bhattarai, A. Torabi Haghighi, and B. Kløve, "Reliability of functional forms for calculation of longitudinal dispersion coefficient in rivers," *Sci. Total Environ.*, vol. 791, Oct. 2021, Art. no. 148394.
[61] G. Rousselet, C. R. Pernet, and R. R. Wilcox, "An introduction to the bootstrap: A versatile method to make inferences by using data-driven simulations," *Meta-Psychol.*, vol. 7, 2023.
[62] F. B. Hamzah, F. Mohd Hamzah, S. F. Mohd Razali, O. Jaafar, and N. Abdul Jamil, "Imputation methods for recovering streamflow observation: A methodological review," *Cogent Environ. Sci.*, vol. 6, no. 1, Jan. 2020, Art. no. 1745133.
[63] V. M. A. Cerqueira, "Ensembles for time series forecasting," Universidade do Porto, Porto, Portugal, Tech. Rep., 2020.
[64] V. Cerqueira, L. Torgo, and I. Mozetič, "Evaluating time series forecasting models: An empirical study on performance estimation methods," *Mach. Learn.*, vol. 109, no. 11, pp. 1997–2028, Nov. 2020.
[65] E. Snieder, R. Shakir, and U. T. Khan, "A comprehensive comparison of four input variable selection methods for artificial neural network flow forecasting models," *J. Hydrol.*, vol. 583, no. 4, Apr. 2020, Art. no. 124299.
[66] A. D. Mehr and E. Kahya, "A Pareto-optimal moving average multigene genetic programming model for daily streamflow prediction," *J. Hydrol.*, vol. 549, no. 6, pp. 603–615, Jun. 2017.
[67] M. Kavya, A. Mathew, P. R. Shekar, and P. Sarwesh, "Short term water demand forecast modelling using artificial intelligence for smart water management," *Sustain. Cities Soc.*, vol. 95, Aug. 2023, Art. no. 104610.
[68] X. Shu, W. Ding, Y. Peng, Z. Wang, J. Wu, and M. Li, "Monthly streamflow forecasting using convolutional neural network," *Water Resour. Manage.*, vol. 35, no. 15, pp. 5089–5104, Dec. 2021.
[69] C. Gao, M. J. Booij, and Y.-P. Xu, "Development and hydrometeorological evaluation of a new stochastic daily rainfall model: Coupling Markov chain with rainfall event model," *J. Hydrol.*, vol. 589, Oct. 2020, Art. no. 125337.
[70] E. Ghaderpour, H. Dadkhah, H. Dabiri, F. Bozzano, G. S. Mugnozza, and P. Mazzanti, "Precipitation time series analysis and forecasting for Italian regions," *Eng. Proc.*, vol. 39, no. 1, p. 23, 2023.
[71] W.-R. Jong, Y.-M. Huang, Y.-Z. Lin, S.-C. Chen, and Y.-W. Chen, "Integrating Taguchi method and artificial neural network to explore machine learning of computer aided engineering," *J. Chin. Inst. Engineers*, vol. 43, no. 4, pp. 346–356, May 2020.
[72] P. M. Pujar, H. H. Kenchannavar, R. M. Kulkarni, and U. P. Kulkarni, "Real-time water quality monitoring through Internet of Things and ANOVA-based analysis: A case study on river krishna," *Appl. Water Sci.*, vol. 10, no. 1, pp. 1–16, Jan. 2020.

YAXING WEI received the degree in civil engineering from the Wuhan Institute of Technology, China. She is currently pursuing the Ph.D. degree with the Department of Civil Engineering, University of Malaya. Her doctoral research investigates the use of artificial intelligence techniques in the field of water resources and hydrological process.

She takes a multidisciplinary approach that encompasses the fields of hydrology, statistical analysis, optimization, and machine learning.

HUZAIFA BIN HASHIM was born in 1985. He received the B.E. (civil) degree from Universiti Teknologi Malaysia, in 2008, and the dual degrees in structural engineering and geotechnical engineering from the University of Malaya, in 2010 and 2018, respectively. He is a Senior Lecturer with the Faculty of Engineering. He has made his path unceasingly in civil engineering field, gaining in knowledge over the years both in academic and industrial line. He was a consulting engineering of many civil engineering related projects under the University of Malaya Consulting Unit and UPUM Sdn. Bhd.

SAI HIN LAI is a Professor with the Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak. His research is focused on flood, drought, and water resources management in the context of climate change, which involves computational simulation, development of decision support systems, artificial intelligence models, and latest optimization techniques. He is a registered Professional Engineer (P.Eng., Malaysia), a Chartered Engineer (C.Eng., U.K.), and a fellow of the Asean Academy of Engineering and Technology (FAAET) and the Institution of Engineering and Technology (FIET).

KAI LUN CHONG received the Civil Engineering degree from the University of Malaya, Malaysia. He is a motivated Doctorate with the Faculty of Engineering and Quantity Surveying, INTI International University (INTI-IU). His Ph.D. study focuses on applying artificial intelligence approaches to water resources and hydrological processes.

His work encompasses hydrology, statistical analysis, optimization, and machine learning, using an interdisciplinary approach. During an acdemic time, his study focused on the use of wavelet analysis for trend analysis in hydrological systems. Furthermore, throughout his tenure with the University of Malaya, he actively participated in several academic contests, exhibiting his dedication to academic success.

YUK FENG HUANG received the B.E. degree (Hons.) in biology and agricultural, the M.Sc. degree in civil engineering, and the Ph.D. degree in water resources engineering from Universiti Putra Malaysia.

He is a highly accomplished Professor with the Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science. He is well-versed in his field. His expertise in drought forecasting, flood modeling, hydraulics, hydrology/water resources, and meteorology/climate change, coupled with his professional qualification as a P.Eng., further solidify his distinguished reputation.

ALI NAJAH AHMED received the B.Sc. degree in civil engineering from the University of Technology, Baghdad, Iraq, in 2003 and the M.Eng. and Ph.D. degrees in civil engineering from the National University of Malaysia(UKM), in 2009 and 2012, respectively.

During the period between 2012 and 2015, he worked as a Faculty Member as Senior Lecturer with the School of Ocean Engineering, University Malaysia Terengganu (UMT), Malaysia. He was working as a Senior Lecturer with the Department of Civil Engineering, College of Engineering, National Energy University (Universiti Tenaga Nasional - UNITEN), Malaysia, from 2015 to 2023. Currently he is holding position as an Associate Professor with the School of Engineering and Technology, Sunway University. He is the author of more than 300 scientific articles published in prestigious journals.

Dr. Ahmed serves as a member of the editorial board and guest editor for special issues for various journals. He has been involved as a leader to several research projects, national and international from various agencies with a budget equivalent to 160.000 USD. He is a Chartered Engineers (CEng) by Engineering Council in U.K.

MOHSEN SHERIF is currently appointed as a Professor of water resources and the Director of the National Water Center and the Research and Sponsored Projects with UAE University. Prior to his current assignment, he occupied the positions of an Associate Provost for Academic Personnel, the Acting Dean of the College of Engineering, and the Department Chair–Civil and Environmental Engineering with UAE University; and an Associate Researcher with the Hydrology Department, KISR, Kuwait. He has 35 years of teaching and research experience in groundwater flow and contamination, artificial recharge and management of groundwater systems, hydrology, hydraulics, fluid mechanics, numerical simulation, climate change, and water resources management. He has published more than 120 journal and conference papers and book chapters and completed several large scale projects.

He was a recipient of several international and national awards, including two Fulbright Scholarships, in 1993 and 1997; and the HE Sheikh Khalifa (President of UAE) Award for Education, Higher Education/Research, First Round, in March 2008. He was also awarded the designation of Diplomate and a Water Resources Engineer with the American Academy of Water Resources Engineers. He is an Associate Editor of the *Journal of Hydrology* and *Journal of Hydrologic Engineering*.

AHMED EL-SHAFIE received the B.Sc. and M.Sc. degrees from the Department of Civil Engineering, Cairo University, Giza, Egypt, in 1993 and 1998, respectively, and the Ph.D. degree in water resources management and planning from the Department of Civil Engineering, Cairo University, in 2003, under a collaborative academic channel program with the Civil Engineering Department, University of Calgary, Calgary, AB, Canada.

From 2004 to 2007, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, Royal Military College of Canada, and Queen's University, Kingston, ON, Canada. From 2007 to 2015, he was an Associate Professor of smart engineering system with the Department of Civil and Structural Engineering, University Kebangsaan Malaysia, Malaysia. Currently, he is a Professor with the Department of Civil Engineering, University of Malaya. He has more than 200 research manuscripts that have been published in highly prestigious scientific journals and published five books. His research interest includes artificial intelligence techniques with their applications to several engineering applications, giving an emphasis to hydrological process, environmental and water resources, dam and reservoir operation, and multi-sensor system integration.