IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 5 January 2023, accepted 17 January 2023, date of publication 23 January 2023, date of current version 30 January 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3238717

RESEARCH ARTICLE

Projected Water Levels and Identified Future Floods: A Comparative Analysis for Mahaweli River, Sri Lanka

NAMAL RATHNAYAKE[1], UPAKA RATHNAYAKE[2], IMIYA CHATHURANIKA[3], TUAN LINH DANG[4], AND YUKINOBU HOSHINO[1], (Member, IEEE)

[1]School of Systems Engineering, Kochi University of Technology, Kami, Kochi 782-8502, Japan
[2]Department of Civil Engineering and Construction, Faculty of Engineering and Design, Atlantic Technological University at Sligo, Sligo, F91 YW50 Ireland
[3]Department of Civil Engineering, Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka
[4]School of Information and Communications Technology, Hanoi University of Science and Technology, Hanoi 100000, Vietnam

Corresponding author: Namal Rathnayake (namalhappy@gmail.com)

ABSTRACT The Rainfall-Runoff (R-R) relationship is essential to the hydrological cycle. Sophisticated hydrological models can accurately investigate R-R relationships; however, they require many data. Therefore, machine learning and soft computing techniques have taken the attention in the environment of limited hydrological, meteorological, and geological data. The accuracy of such models depends on the various parameters, including the quality of inputs and outputs and the used algorithms. However, identifying a perfect algorithm is still challenging. This study develops a fuzzy logic-based algorithm called Cascaded-ANFIS to accurately predict runoff based on rainfall. The model was compared against three regression algorithms: Long Short-Term Memory, Grated Recurrent Unit, and Recurrent Neural Networks. These algorithms have been selected due to their outstanding performances in similar studies. The models were tested on the Mahaweli River, the longest in Sri Lanka. The results showcase that the Cascaded-ANFIS-based model outperforms the other algorithms. The correlation coefficient of each algorithm's predictions was 0.9330, 0.9120, 0.9133, 0.8915, 0.6811, 0.6811, and 0.6734 for the Cascaded-ANFIS, LSTM, GRU, RNN, Linear, Ridge, and Lasso regression models respectively. Hence, this study concludes that the proposed algorithm is 21% more accurate than the second-best LSTM algorithm. In addition, Shared Socio-economic Pathways (SSP2-4.5 and SSP5-8.5 scenarios) were used to generate future rainfalls, forecast the near-future and mid-future water levels, and identify potential flood events. The future forecasting results indicate a decrease in flood events and magnitudes in both SSP2-4.5 and SSP5-8.5 scenarios. Furthermore, the SSP5-8.5 scenario shows drought weather from May to August yearly. The results of this study can effectively be used to manage and control water resources and mitigate flood damages.

INDEX TERMS Cascaded-ANFIS, flood forecasting, GRU, LSTM, rainfall-runoff, RNN, SSP-245, SSP-585.

I. INTRODUCTION

Natural disasters often occur due to recent climate changes. Several studies have focused on climate change and its' effect detection where Remote sensing methods are highly used in these methodologies [71].

Floods are frequently observed in natural disasters. However, they are one of the direct outcomes of the rainfall-runoff (R-R) process [77]. Due to their severity and frequent occurrence, flood prediction has taken significant attention in R-R modelling [1]. Even though they are natural disasters, their severity has been impacted by anthropogenic activities. Flow hydrographs are drastically changed to have higher peaks quickly due to ongoing urbanization [2], [3], [4]. Flash floods are often in urbanized areas [5], [6]. Hence, urbanization is one of the most impacting factors in today's floods.

In addition to urbanization, changing climate has adversely impacted today's floods. Some regions receive higher and intensified rainfall events [7], [8], [9], [10] whereas some

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

other areas receive reduced rainfall events [11], [12] due to ongoing climate change. Frequent floods are expected in areas with projected increased rainfall events. Many studies in the literature support this observation [13], [14], [15].

Therefore, accurate modelling of runoff-rainfall relationships to catchments is in high demand. It is important to note that each catchment has to be modelled to find its R-R relationship. Commercial and non-commercial hydrological computer packages are available to simulate the R-R relationships of catchments. However, these computer packages require various data related to digital elevation models, soil data, meteorological data, and discharge data. [17]. The accuracy of the catchment models is highly varied due to the quality of catchment data [39]. Only some catchments are gauged to have meteorological and discharge data and other catchment characteristics on a temporal and spatial basis. Thus, the catchment models always need help achieving the required accuracy to model the runoff and then predict the floods.

In the event of limited data, soft computing [38], [40], and machine learning techniques [19], [20], [21], [22] are helpful to model the R-R processes. R-R processes can be modelled only using the known rainfall and measured discharges and, importantly, without any catchment characteristics. Hence, numerous methodologies under soft computing and machine learning have been developed using various algorithms and study cases. One of these data-driven methods is the artificial neural network (ANN), which has been used in various fields, including hydrology and water resources. It has gained popularity because it can address, model, and forecast stochastic and nonlinear situations in the system [41], [42], [43], [44], [45], [46], [47] The algorithm does not replace conceptual watershed modelling of the impossibility of describing the catchment's internal structure and handling the data disseminated relating to the physical properties. Nevertheless, they have gained acceptance as a practical substitute for conceptual models for forecasting because of numerous benefits, such as the ability to produce simple and accurate models [48] and the computation speed [49]. Additionally, this study has demonstrated its strength and ability to mimic hydrological events. As a result, ANN models are suggested for rainfall-runoff modelling due to their straightforward designs and accuracy, enabling addressing the issues of managing water resources.

To create ANN models, most studies have used feed-forward and backpropagation (FFBP) networks. Although relatively well known for their ability to anticipate floods, neither model's performance in a particular application has been determined [43]. Since several learning methods may be used to improve ANN, there is still a wide range of probability. Gradient descent (GD) is frequently used in neural network training at the backpropagation stage [50]. GD has been used in recent years to increase the potential of the backpropagation algorithm. However, the GD may experience problems with convergence, training technique slowdown, overfitting, and stocking inside local minima. The performance of the training algorithm can lower the performance when the structure of the model is complex, and the parameter set is significant [11], [51], [52].

Moreover, Feed-forward deep neural networks (FF-DNNs) have been used widely in climate change-related studies. A case study in Kastoria Lake in Greece used FF-DNN to predict dissolved oxygen. They have obtained maximum NSE efficiency of 0.89 [70]. Forecasting of dissolved oxygen was studied using three methods such as the Autoregressive integrated moving average (ARIMA) method, Transfer Function (TF) method, and NN method [72]. They concluded that the ARIMA method provides significant results compared to TF and NN. Additionally, A combination of tools such as remote sensing, weather forecasting, and Artificial Intelligence was used to improve irrigation management in Mediterranean Basins. This study suggests that comprehensively using these tools can enhance the irrigation system rapidly [73].

Recently, several novel evaluations of CNN models were implemented: the Extreme Gradient Boosting (XGBoost) and CNN-transformer. These algorithms have been widely tested for uncertain and nonlinear data. Many studies recommended ANFIS as a highly accurate algorithm for predictions [38], [40]. Xuan-Nam et al. [39] have proposed an ML model for blast-induced ground vibration predictions in quarries. They have employed several state-of-the-art algorithms, such as Moth-flame optimization-based ANFIS, XGBoost, ANN, and SVM. The study showcased that the ANIFS-based algorithm outperformed the other model with an accuracy of 98.62%. Moreover, two environmental types of research have been introduced by Hameed et al. [38] and Junliang et al. [40] employing ANFIS and XGBoost algorithms.

On the other hand, Genetic Algorithms (GA) in the hydrological sciences have been the subject of several investigations to train (ANN) rainfall-runoff models that are more accurate than backpropagation technique-based ANN models in anticipating the quotidian flow [59] using natural code GAs. In conjunction with intelligence approaches, the GA has developed into a potent tool for modelling and optimizing complicated processes [56], [57], [58]. It is commonly used in ANN to enhance efficiency by tuning the parameters [54], [55]. Roy and Singh [11] developed a novel hybrid metaheuristic method for simulating the rainfall-runoff process that integrates Biogeography-Based Optimization (BBO), Particle Swarm Optimization (PSO), and grey wolf optimizer (GWO) combining ANN and Adaptive Network-based Fuzzy Inference Systems (ANFIS). Moreover, three optimization algorithms integrated with ANFIS were introduced for rainfall-runoff predictions, namely, Differential Evolution algorithm based ANFIS (ANFIS-DE), Particle Swarm Optimization based ANFIS (ANFIS-PSO), and Genetic Algorithm based ANFIS (ANFIS-GA) [53]. Investigating and contrasting these models in hydrology is strongly advised because the different algorithms have various advantages and

distinct methods for complex modelling phenomena. The investigations in hydrology, particularly rainfall-runoff modelling, are still in the early stages. Hence, the computational analysis has to be comprehensively conducted for a better outcome. Therefore, this research study aims to contribute to scientific society by achieving the following objectives.

1) Designing and developing an accurate, low computational complex machine learning model for rainfall-runoff forecasting. (The Cascaded-ANFIS)
2) Conducting comprehensive experiments to support the proposed algorithm using six regression algorithms (Linear regression, Ridge regression, Lasso regression, Long Short-Term Memory (LSTM), Grated Recurrent Unit (GRU), and Recurrent Neural Networks (RNN)) using an important river basin in Sri Lanka
3) Predict the future water levels for the near-future (2022 - 2030) and mid-future (2031 - 2050) using Shared Socio-economic pathways (SSP245 and SSP585) and then analyze the flood events in the future.

II. PROJECTING ALGORITHM - CASCADED-ANFIS

In 1993, Jang [76] unveiled the Adaptive Network-based Fuzzy Inference System (ANFIS), a versatile and ingenious hybrid system. Fuzzy inference systems (FIS) and neural networks (NN) perform flawlessly together in ANFIS [29]. The ANFIS system benefits from NN and FIS's collaboration by utilizing their advantages. The system's conversion to straightforward if-then rules is another crucial benefit of this network. ANFIS's if-then control structure gives it the capacity to handle non-linear functions. It is shown that ANFIS has been applied in several study fields and yields generally effective outcomes. It is usually known that ANFIS may be used with many algorithms to reduce training phase error. For instance, the least square approach and gradient descent can increase the efficiency of finding the optimal parameters.

This research shows that ANFIS functions similarly to the fuzzy system that Takagi and Sugeno presented in 1985 [30]. The forward section's consequence factors are determined using a least-squares method.

Input, membership function, fuzzification, defuzzification, normalization, and output are the five general layers that makeup ANFIS. An additional explanation is based on Figure 1 and assumes that the ANFIS system has two inputs, x, and y, while the output is denoted as f in Equation (1) and (2).

$$f_1 = p_1 x + q_1 y + r_1, \quad \text{assume} \quad x = A_1, \quad y = B_1 \tag{1}$$
$$f_2 = p_2 x + q_2 y + r_2, \quad \text{assume} \quad x = A_2, \quad y = B_2 \tag{2}$$

Fuzzy sets A_1 and B_1 are used here, and design parameters p_i, q_i, and r_i are used where $i = 1, 2$. The membership makes up the top layer of the ANFIS structure. This layer's nodes are adaptable. For each input, membership ratings are created in this layer. The following equations can be used to explain the functionality:

$$O_{1,i} = \mu_{Ai}(x) \quad i = 1, 2 \tag{3}$$
$$O_{1,j} = \mu_{Bj}(y) \quad j = 1, 2 \tag{4}$$

Nodes' linguistic labels are shown as Ai and Bi when x and y are the inputs. The grades of the membership for a set A ($A1, A2, B1$ and $B2$) are $\mu_{Ai}(x)$, and $\mu_{Bj}(y)$, respectively, which are adaptive. For instance, the following equation is used when the bell-shaped is employed.

$$\mu_{Ai}(x) = \frac{1}{1 + \left\{ \left(\frac{x - c_i}{a_i} \right)^2 \right\}^{b_i}} \tag{5}$$

In this case, the bell-shaped function's corresponding parameters are a_i, b_i, and c_i. Simple multiplication is carried out in the following layer, which comprises fixed nodes. Following is a presentation of the layer's mathematical expression.

$$O_{2,i} = w_i = \mu_{Ai}(x) \times \mu_{Bi}(x) \qquad i = 1, 2 \tag{6}$$

A fixed node normalization layer comes after that. This layer is where the output from the second layer is normalized. The operation is demonstrated by the equation below.

$$O_{3,i} = \overline{w} = \frac{w_i}{w_1 + w_2} \qquad i = 1, 2 \tag{7}$$

Here, w_i displays the firing power of node i.

Creating the normalized output from the third layer can be more straightforward than the fourth. The outcome of this adaptive layer may be shown using the equation below.

$$O_{4,i} = \overline{w} f_i = \overline{w_i}(p_i + q_i + r_i) \qquad i = 1, 2 \tag{8}$$

Only one fixed node makes up the final layer. This node adds up all the inputs that are received. Ultimately, the complete result can be extracted by applying the equation below.

$$O_{5,i} = \sum_{i=1}^{2} \overline{w_i} f_i = \frac{w_i f_1 + w_2 f_2}{w_1 + w_2} \tag{9}$$

Since back-propagation and least squares techniques improve the method's accuracy and speed up convergence, ANFIS has a more substantial capacity for learning. As previously stated, this system uses six tunable parameters (while a bell shape is used). The primary goal of this ANFIS system is to tune these settings to get the lowest cost. The first layer's parameters will be adjusted through back-propagation, and the fourth layer's parameters will be adjusted by the least squares technique [34].

With two main inputs and one main output, the Cascaded-ANFIS algorithm is a repeated ANFIS algorithm. The critical difference between the Cascaded-ANFIS algorithm and the classic ANFIS algorithm is that the output of the conventional ANFIS technique is used as the input for future applications of the traditional ANFIS method. Figure 2 is a valuable tool for presenting the building of this method.

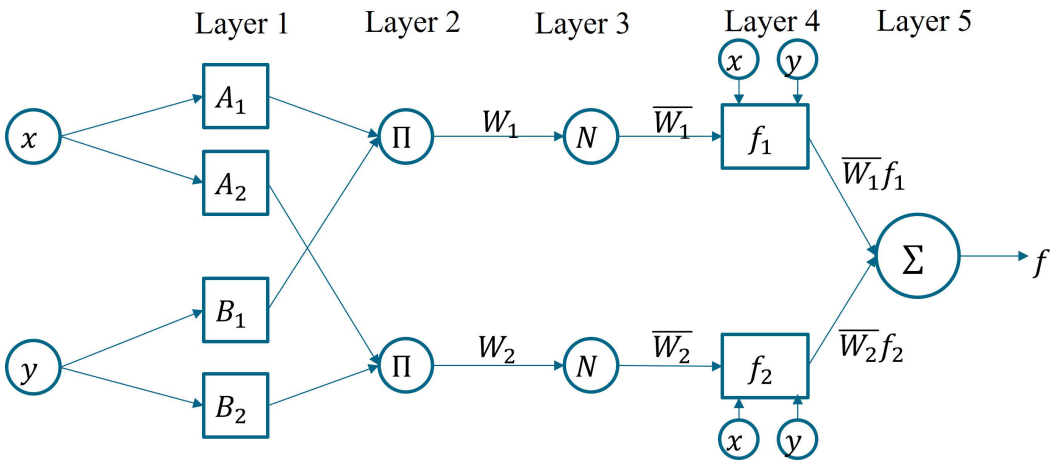


FIGURE 1. General adaptive network-based fuzzy inference system.

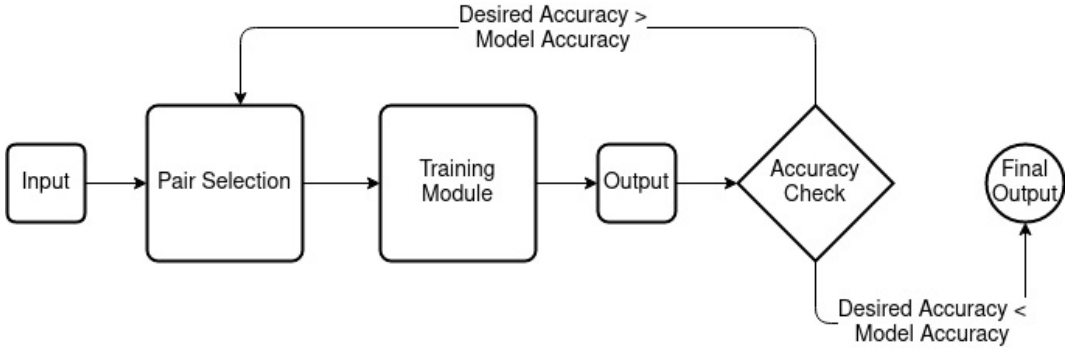


FIGURE 2. The general structure of the Cascaded-ANFIS algorithm.

The Cascaded-ANFIS algorithm comprises two major parts: the pair selection method and the training method. The Pair Selection module solves the first considerable issue with ANFIS. The Pair Selection module solves the first significant issue with ANFIS. However, the inner layers of the ANFIS model use fuzzy, merely like the standard ANFIS technique. Membership functions convert numerical data into fuzzy members and are used to achieve fuzzification.

However, the original method uses each characteristic to build a robust model, equally valid for noisy data sets. The novel Cascaded-ANFIS method manages computational complexity through its Training.

The pair Selection takes advantage of sequential feature selection (SFS). This technique employs a 2-input, 1-output ANFIS model to find the best match for each input. The training module also makes use of the 2-input ANFIS model. The ANFIS module may receive the input variables directly since they are connected to the preceding module's best match, which results in current outputs and Root Mean Squared Error (RMSE) for each data pair. The expected error is then contrasted with the RMSE. There is now an error with a pre-determined aim as well. The procedure can be finished if the goal error is attained. If not, the algorithm advances to the next iteration.

This article for implementation provides a thorough introduction to the Cascaded-ANFIS algorithm with pseudo-code [31], [32], [69].

III. METHODOLOGY

A. PROBLEM FORMULATIONS

The following relationship shown in Equation 10 was modelled using the Cascaded-ANFIS algorithm. The relationship was trained using the ground-measured rainfall at i^{th} station and water level. Subscript t in Equation 10 denotes the time domain of the R-R relationship.

$$WaterLevel_t = f(RainFall_{i,t}) \tag{10}$$

However, it is well noted that time domains can be shifted from rainfall to runoff from that rainfall due to the catchment characteristics like river length, catchment area, land use patterns, and soil type. The travel time of a particular rainfall event has to be clearly understood.

Figure 3 develops the flowchart for the developed Cascaded-ANFIS model. As shown in the Figure, the rainfall data is used as the primary input of the system. Then the input data are re-arranged with a delay of one day and two days. The inputs were then removed based on the computation of the correlation between each input and the output of the flow

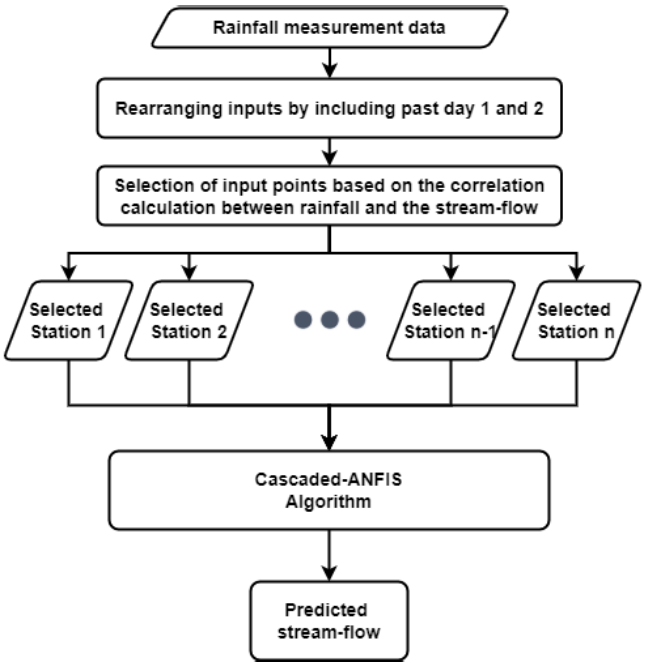


FIGURE 3. The overall structure of the Cascaded-ANFIS implementation using the selected inputs.

level. A minimal correlation of 0.40 between an input and an output was used in this case. The selection methodology of inputs is discussed in later sections.

B. COMPARATIVE ANALYSIS TO IDENTIFY THE BEST ALGORITHM

Six regression algorithms (Linear Regression, Ridge Regression, Lasso Regression, Long Short-Term Memory (LSTM), Grated Recurrent Unit (GRU), and Recurrent Neural Network (RNN)) together with the Cascaded-ANFIS algorithm were used to formulate the R-R relationship. These ML algorithms were considered in this study due to a few specific reasons, such as algorithms being similar and easy implementation. Moreover, they are low in weight and can be processed in a general computer without GPU support. Table 1 shows the parameter values considered for tuning hyper parameters of Ridge, Lasso, LSTM, GRU, and RNN tuning. These parameters were selected based on trial and error methods. Each parameter is tested with the datasets used in this study and employs the optimum value.

The Cascaded-ANFIS used three Gaussian membership functions for each input in the system. The whole cascades were ten to achieve satisfactory accuracy and error value.

C. THE MAHAWELI RIVER SUB-CATCHMENT ANALYSIS

Localized floods can be observed in sub-catchments in Figure 5 without showcasing major floods downstream due to the catchment characteristics. Therefore, the downstream river gauge may not observe any flood situation. However, upstream sub-catchments might have experienced localized floods. Therefore, it is essential to cluster larger catchments into sub-catchments and then analyze them separately. This scenario was analyzed in this research work and formulated Equation 10 for sub-catchments.

TABLE 1. Parameter settings of the algorithms used in this study.

Algorithm	Tuned Parameters	
	Parameter	Value
LSTM	Batch Size	16
	Epoch	10
	Optimizer	adam
GRU	Batch Size	16
	Epoch	10
	Optimizer	adam
RNN	Batch Size	16
	Epoch	10
	Optimizer	adam
Ridge	Alpha	10
	Iterations	10000
	Solver	sparse_cg
Lasso R	Alpha	10
	Iterations	10000
	Selection	Random
	Random State	Random
Casaceded-ANFIS	Iterations	100
	Membership Function	Gaussian
	Number of MF	3
	Step Size	0.1
	Decrease rate	0.9
	Increase rate	1.1

D. FLOOD IDENTIFICATION

According to the desinventar dataset of natural disasters [78], there has been significant damage due to flooding in Sri Lanka. In most cases, the damage has increased due to unexpected heavy rainfall and poor irrigation management. The database reveals that in the past events from 2005 to 2018, there was at least one death due to flooding. The highest number of deaths, injured and missing personals were recorded in 2017, with 67, 73, and 63, respectively.

Historical water levels were analyzed to define threshold water levels to identify floods in the basin. Here, water levels were considered because the authorities recorded the data as water levels instead of the water flows. If the water levels or stream flows exceed the threshold, that flow may be a flood. However, this can be confirmed with the ground-measured discharge data and by comparing flood data to the catchment. Nevertheless, many countries do not have these flood databases, so there can be some issues with the accuracy [79].

E. SHARED SOCIO-ECONOMIC PATHWAYS (SSP) CLIMATE DATA EXTRACTION

IPCC's sixth report [60] presented a new set of scenarios based on greenhouse gas emissions to project the future climates until 2100. Practitioners who engage with future climate data may investigate climate changes across a range

of quite diverse futures thanks to the availability of climate forecasts for numerous Shared Socio-Economic Pathways (SSPs). These SSPs are titled SSP1, SSP2, SSP3, SSP4, and SSP5 under several Socioeconomic Pathways. SSPs describe potential future growth pathways for human cultures. A set of models combine assumptions on the ambitions for reducing the impact of climate change with predictions about how population, education, energy usage, technology, and other factors may evolve over the next century. Various conceivable future climates, from a pessimistic high-carbon scenario to a low-carbon one that satisfies the goals of the 2015 Paris Agreement, are described in the climate change forecasts from these scenarios [25], [26].

The Representative Concentration Pathways, or RCPs, or earlier projections of greenhouse gas concentration, are improved upon by SSP-based scenarios. To investigate the consequences of various emission trajectories or emissions concentrations, RCPs were explicitly created for the community of climate modellers. It is challenging to relate social trends such as population growth, educational attainment, and government policies to climate objectives like limiting global warming to below 2 $^{\circ}C$ since the socioeconomic factors used to establish RCPs need to be standardized. To address this, SSPs outline how societal decisions might alter Radioactive Forcing towards the end of the century. As a result, SSPs were built on RCPs to enable a uniform comparison of societal decisions and the degrees of climate change they cause. These SSP data are used in various recent research studies such as flood forecasting [35], land use optimization [36], and prediction of air pollution for the future [37]. Climate change research [37]. According to these studies, the reliability of SSP data is much higher than the RCP data. Therefore, this study employed SSP projections for daily rainfall data acquisition [27], [28]. Here, two SSP scenarios have been used for the data acquisition, such as SSP2-4.5 and SSP5-8.5. SSP2-4.5 redevelops the low carbon impact globally, while SSP5-8.5 is the high carbon scenario.

F. BIAS CORRECTION

The extracted rainfall data under SSP2-4.5 and SSP5-8.5 were corrected using linear bias correction factors. Usually, the data extracted from climate models may have some systematic errors [61]. Therefore, the model's extracted climate data are corrected for bias using the ground-measured climate data. Various bias correction techniques are available [62]; however, the linear bias correction method was selected in this research work. Equation 11 gives the simple mathematical formulation for linear bias correction. More details on this can be found in Chaturanika et al. [63].

$$RF^{*}_{sim}(d) = RF_{sim} \times \frac{\mu_m(RF_{obs}d))}{\mu_m(RF_{his}(d))} \tag{11}$$

where RF, d, μ_m, his, obs, and sim are rainfall, daily, long-term monthly mean, raw SSP data, observed/measured data, and raw RCM forecast. The symbol $*$ denotes the bias-corrected datasets.

G. PROJECTED WATER LEVELS AND FLOODS

Bias-corrected SSP rainfall data were fed to the developed R-R relationship in Equation 10. Based on these future rainfalls under two SSP scenarios, the stream flows in the means of water levels were predicted for future years. These predicted water levels for the whole catchment were tested for the extreme values in the time series and then identified localised and downstream floods. These predicted floods are given for the near future (from 2022-2030) and mid-future (2031-2050).

IV. CASE STUDY

Sri Lanka is a country blessed with water resources. However, heavy monsoon rainfall drives many rivers into floods, and annual floods are quite often [64]. Sri Lanka has many rivers, tanks and lakes, and these watersheds are flooded during the monsoon periods. Several deaths and excessive structural damage are annually reported due to extreme weather conditions. Sri Lanka has 103 rivers, and the total length of the rivers is around 4500 km. The longest river in Sri Lanka is the Mahaweli River. It is 335 km long and covers a 10488 km^2 river basin which covers almost one-fifth of the total area of the island [65], [66]. The river has several branches along the way to the sea. 40% of the total electricity demand of Sri Lanka is provided by the hydropower generated by the Mahaweli River. Nevertheless, the Mahaweli River is known to provide a vast water supply for the cultivation of crops such as rice and vegetables [67]. Several Mahaweli River developments have been for hydroelectric generation and irrigation purposes. Many dams were constructed along the river to enhance energy generation, which led to flood risk changes. Kothmale dam was one of those constructed to generate electricity; however, indirectly, it has mitigated the floods downstream [68]. The Mahaweli River was selected for this research study due to its importance in many utilities and its frequent floods in the northeastern monsoon period (from December to February).

A. STUDY AREA AND SUB-CATCHMENTS

The Mahaweli River starts from the central hills of Sri Lanka with several small creeks. Agra Oya from Horton Plains is one of the starting creeks of the Mahaweli River. The river reaches the Bay of Bengal on the southwestern side of Trincomalee Bay. The bay includes the first of several submarine canyons, making Trincomalee one of the finest deep-sea harbours in the world. As part of the Mahaweli Development program, the river and its tributaries are dammed at several locations to allow irrigation in the dry zone, with almost 1,000 km^2 (386 sq mi) of land irrigated. Figure 5 develops the primary catchment and sub-catchments, whereas Figure 4 shows the catchment of the Mahaweli River basin.

Two sub-catchments were identified along two tributaries of the Mahaweli River. The catchment above Peradeniya (for Kothmala Oya and other parts upstream creeks of Mahaweli River) is given in sub-figure (a). In contrast, the catchment

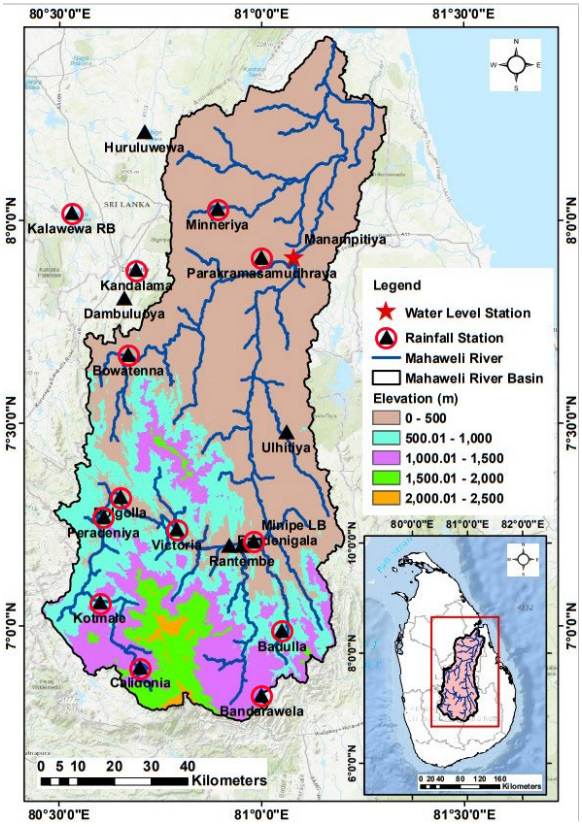


FIGURE 4. Study area of the Mahaweli river catchment map.

above Thaldena for Badulu Oya is given in sub-figure (b) in Figure 5. The sub-catchment at Peradeniya is in the wet zone of the country; thus, heavy rainfall can be experienced. However, the sub-catchment at Thaldena is in the wet and intermediate zone. Thus, the rainfall in that sub-catchment is not as high as that at Peradeniya. However, these two sub-catchments are essential in terrain, land use, and urbanization. In addition, two flow gauges can also be found in these two sub-catchments.

B. DATA

Figure 4 shows rain gauges for the Mahaweli River basin. Due to the unavailability of complete data in most of the years, the daily rainfall data from 2000 to 2017 were purchased from the Department of Meteorology, Sri Lanka. The missing data percentage for the selected years was less than 1%. The rain gauges were selected to represent the whole catchment covering as much as its area. In addition, the stream flow gauge at Manampitiya was selected to model the R-R relationship. This is the most downstream stream flow gauge available. The water levels at the station were purchased from the Department of Irrigation, Sri Lanka. Furthermore, two water level measuring stations were identified for the selected two sub-catchments: Pereadeniya and Thaldena (refer to Figure 5). The water levels for these two stations were also purchased for the same period from the Department of Irrigation, Sri Lanka.

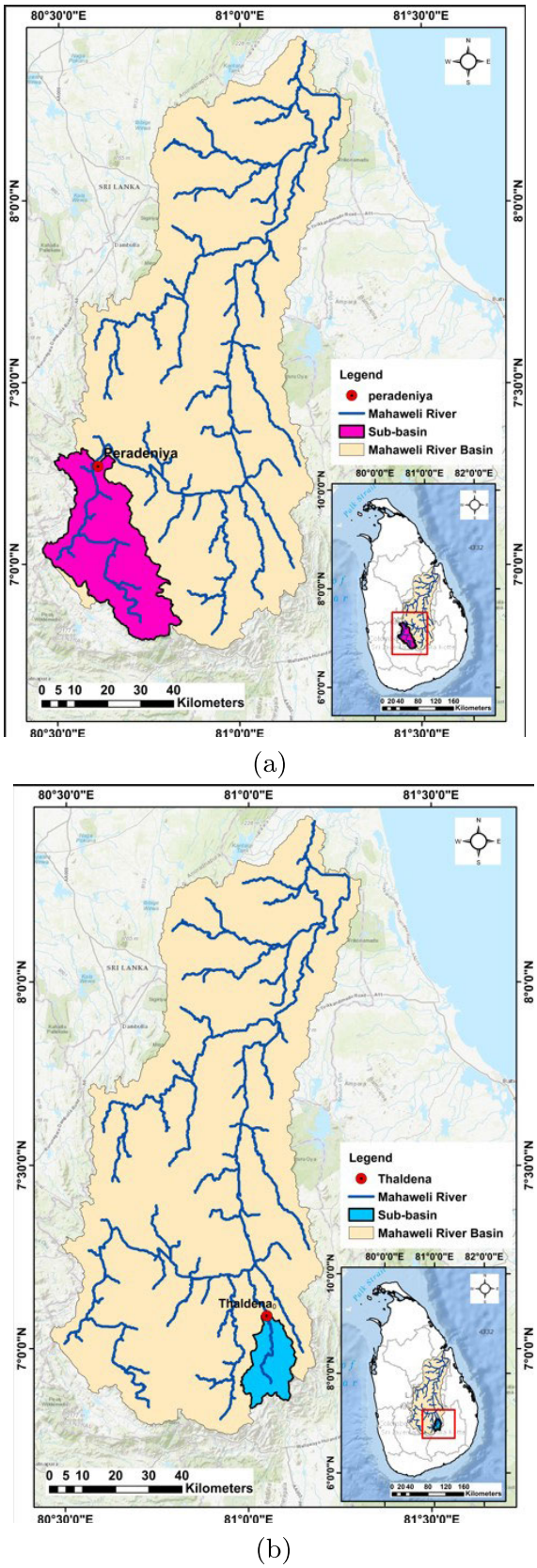


FIGURE 5. Sub catchment study areas; (a) Catchment map at the Peradeniya sub-catchment; (b) Catchment map at the Thaldena sub catchment.

A descriptive analysis of the dataset used in this analysis are shown in Table 2. There were 6207 data samples in the dataset. The selected dataset was divided at a ratio of 7:3 for the training and testing. These sub-dataset samples were used to train and test the algorithms used in this study. The water levels are presented in centimetres, whereas the rainfalls are millimetres. Moreover, several homogeneity tests were conducted, such as the Standard normal homogeneity test (SNHT), Buishand range (BR) test, Pettitt test, and von Neumann ratio (VNR) test to evaluate the dataset before employing it in training models.

Due to the missing data in a significant time frame, few rainfall stations were omitted in the evaluation of the case study. The missing data were presented in Huruluwewa, Dambuluoya, Ulhitiya, Minipe LB, and Rantembe. Therefore, as shown in Table 2, 13 rainfall station data were considered as the inputs.

The correlation calculation in subsection III-C is given in Table3. The selected outputs are highlighted with a minimum of 0.4 correlation. Twelve inputs were selected using the correlation method to train the R-R model. The trial and error method made the selection based on the correlation. At a correlation value of 0.40, the maximum accuracy was obtained. Then the general structure of the Cascaded-ANFIS was used to generate the final outputs of predicted water levels. Additionally, according to the literature, it is considered negligible if a correlation is 0.30 or below. Therefore, 0.40 and above values were considered safe marginal inputs in the system [80].

C. RECENT FLOODS FOR THE RIVER BASIN

Figure 6 shows the annual water level measurements at each of the observation points, such as the primary catchment of Mahaweli River (Mannampitiya) and sub-catchments of Mahaweli River (Peradeniya and Thaldena). It can be seen that Mannampitiya water outlets record a higher level of water when compared with the sub-catchments. As indicated by the figure (refer to the rectangular section), the water levels on some of the days of 2011 (21.8 m on 10/01/2011), 2012 (25.6 m on 18/12/2012 and 21.7 m on 27/12/2012), and 2014 (26.9 m on 27/12/2014) were higher than 20m. These can be identified as flood thresholds to the Manampitiya river gauge.

Sub-catchments Pereadeniya and Thaldena showcased some higher water levels comparable to the higher water levels at Manampitiya; however, some differences can also be observed (refer to Table 4). Thaldena has not showed a significantly higher water level in 2012. Still, higher water levels were observed at Manampitiya during the same time (t_1, t_2, and t_3 in sub-figure (a) in Figure 6). Similar trends can be observed in Peradeniya too. Therefore, the analysis of sub-catchments for floods is highly justified. Comparable observations have led the authors to define flood thresholds for Peradeniya and Thaldena. The threshold for Peradeniya was considered 6 m, while 3 m was considered for Thaldena. The flood events were identified in Peradeniya and presented as t_1, t_2, and t_3 in sub-figure (b) in Figure 6 (6.7 m on 03/06/2013, 6.9 m on 14/09/2013, and 6.7 m on 26/12/2014). In comparison, two incidents were identified for Thaldena and presented as t_1 and t_2 in sub-figure (c) Figure 6 (3.1 m on 02/02/2011 and 3.5 m on 26/12/2014).

V. EXPERIMENTAL RESULTS

A. EVALUATION PARAMETERS

The algorithm performances were then tested by several metrics, including root mean square error (RMSE), bias, Nash-Sutcliffe efficiency (NSE), Kling- Gupta Efficiency (KGE), and correlation coefficient (R). These performance evaluating metrics are given in Equations (12), (13), (14), and (15).

$$RMSE = \sqrt{\frac{1}{q}\sum_{t=1}^{q}(u(t) - \bar{u}(t))^2} \tag{12}$$

$$bias = \frac{\sum_{j=1}^{k} u(t) - \bar{u}(t)}{\sum_{j=1}^{k} u(t)} \tag{13}$$

$$NSE = 1 - \frac{\sum_{j=1}^{k}(u(t) - \bar{u}(t))^2}{\sum_{j=1}^{k}(u(t) - \bar{v}(t))^2} \tag{14}$$

$$R = \frac{\sum(v(t) - \bar{v}(t))(u(t) - \bar{u}(t))}{\sqrt{\sum(v(t) - \bar{v}(t))^2 \sum(u(t) - \bar{u}(t))^2}} \tag{15}$$

where $u(t)$ is the predicted parameter, $\bar{u}(t)$ is the mean of predicted parameter$v(t)$ is the measured parameter, k is the population size, and $\bar{v}(t)$ is the mean of the measured parameter. The correlation coefficient (R) redevelops the goodness of fit. It varies from -1 to 1; the best is when it becomes 1. Bias tells the differences between predicted to measured values. The ideal bias value is 0, and 1 becomes the worst. NSE calculates the perfectness of the match between actual and prediction. The results of the NSE can vary between minus infinity being the worst and 1 being the ideal [75]. KGE is a combined calculation of three primary parameters: NSE, bias, and coefficient of variation. Recently it has been used rapidly in hydrological model performance calculations [74].

B. PERFORMANCE EVALUATION

The river in this case study is the longest in Sri Lanka. According to the geographical experts in Sri Lanka, it is considered that the maximum time duration of travelling water from the start to the end of the river is less than three days. However, there are several reservoirs and dams along the river. Hence, we have considered 1-day, 2-day, and 3-day lags to include all corresponding scenarios in the calculation.

1) CORRELATION OF COEFFICIENTS CALCULATION FOR THE PRIMARY CATCHMENT

The primary catchment of the Mahaweli River consists of 13 rain gauges, all of which were used to predict the water level at Manampitiya. As mentioned in the previous sections, the experiment was designed to identify the best R-R prediction algorithm. Figure 7 shows the coefficient of correlation

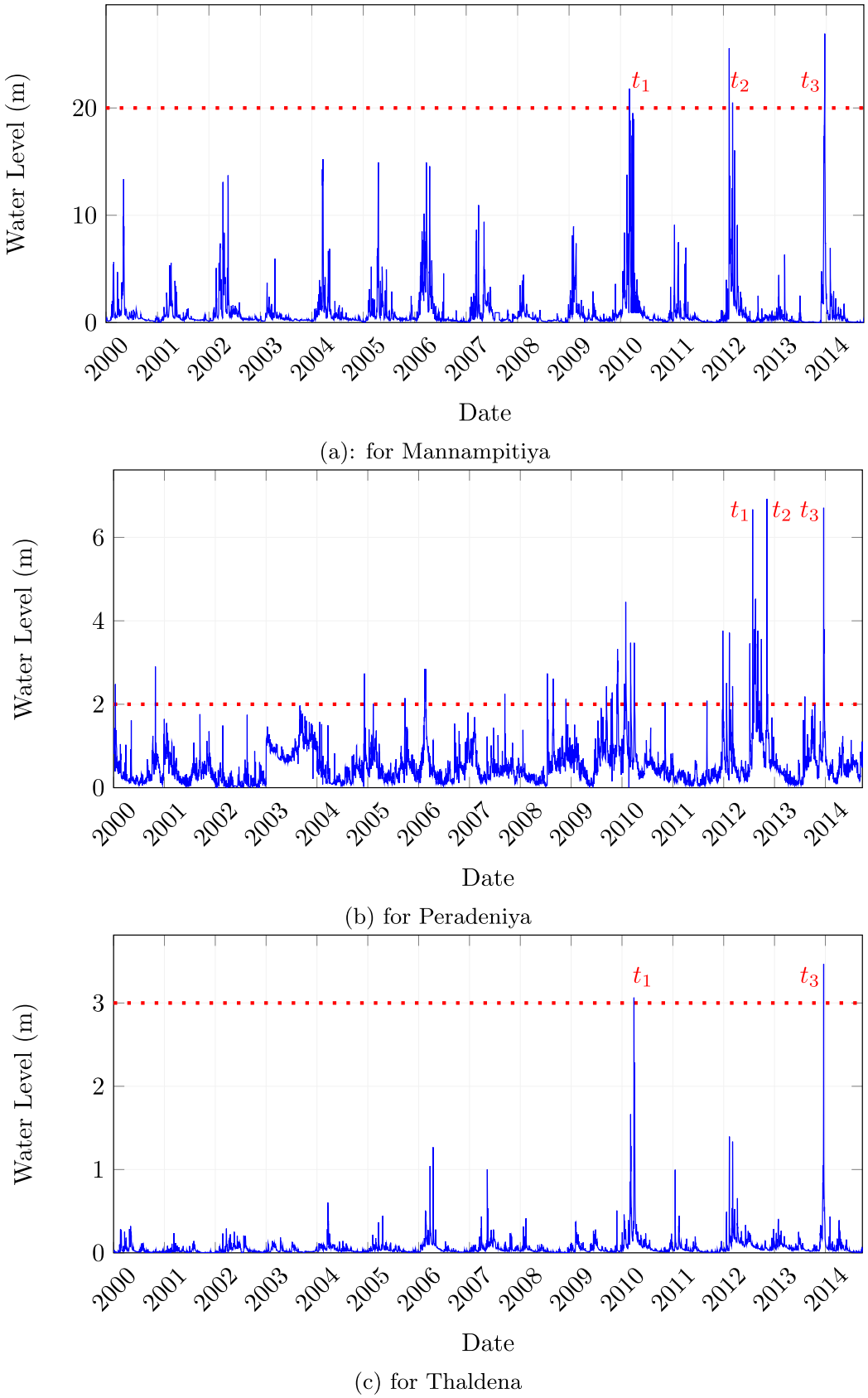


FIGURE 6. Historical water level measurements from the year 2000 to 2015: (a) at the Manampitiya water level measurement station, (b) at the Peradeniya water level measurement station, and (c) at the Thaldena water level measurement station. Here, the flood incidents according to the historical data are presented as t_1, t_2, and t_3.

TABLE 2. Descriptive analysis of the data for the Mahaweli River basin.

Variable	Sample Data	mean	std	min	25%	50%	75%	max
Peradeniya	6207	5.27	13.56	0	0	0	3.9	194.3
Minneriya	6207	4.28	14.09	0	0	0	0	210
Calidonia	6207	5.84	11.87	0	0	0	7	144.6
Parakrama samudhraya	6207	5.13	15.69	0	0	0	0	222
Kandalama	6207	4.55	13.78	0	0	0	0	198.6
Kalawewa RB	6207	3.52	11.51	0	0	0	0	166
Bowatenna	6207	5.09	15.29	0	0	0	1.1	242.5
Kotmale	6207	6.50	13.95	0	0	0	7	191
Polgolla	6207	4.61	11.98	0	0	0	3	170
Randenigala	6207	4.61	13.70	0	0	0	1.3	270.9
Victoria	6207	3.98	11.89	0	0	0	1.3	295
Badulla	6207	5.05	11.94	0	0	0	3.7	195.9
Bandarawela	6207	4.48	10.36	0	0	0	3.4	134.9
Mannampitiya Water Level	6207	0.92	1.94	0.02	0.17	0.34	0.80	26.93

TABLE 3. Correlations between inputs and the flow levels.

RF Stations	Correlation		
	t	t-1	t-2
Peradeniya	0.14	0.25	0.22
Minneriya	0.30	0.41	0.38
Calidonia	0.08	0.16	0.14
Parakramasamudhraya	0.34	0.49	0.43
Kandalama	0.28	0.40	0.36
Kalawewa	0.22	0.31	0.27
Bowatenna	0.32	0.47	0.40
Kotmale	0.08	0.15	0.12
Polgolla	0.19	0.33	0.30
Randenigala	0.29	0.48	0.48
Victoria	0.29	0.49	0.45
Badulla	0.25	0.42	0.42
Bandarawela	0.17	0.29	0.28

TABLE 4. Water levels at river gauges.

Date	Water Levels (m)		
	At Manampitiya	At Peradeniya	At Thaldena
10/01/2011	**21.8**	2.5	0.98
02/02/2011	19.2	1.6	**3.1**
18/12/2012	**25.6**	3.7	1.4
03/06/2013	0.1	**6.7**	0.1
14/09/2013	0.4	**6.9**	0.05
26/12/2014	**21.7**	**6.7**	**3.5**
27/12/2014	**26.9**	3.3	1.2

Bold values are identified as floods

of the predicted water to the ground-measured water level at the Manampitiya river gauge.

Figure 8 develops the prediction accuracy under combined scenarios initially identified as per Table 3 for the predicted water levels at Manampitiya.

2) CORRELATION OF COEFFICIENTS CALCULATION FOR SUB-CATCHMENTS

Figure 9 and 10 shows the prediction accuracy of water levels for each algorithm for the sub-catchments Peradeniya and Thaldena.

Additionally, other parameters were used to evaluate the results, such as Bias, NSE, RMSE, and KGE. The evaluation results are presented in Table 5.

C. PROJECTED WATER LEVELS AT MANAMPITIYA

Figures 11 illustrate the projected future water levels at Manampitiya under the two SSP scenarios for the near future (2022-2030) and mid-future (2031-2050). These results project some exciting interpretations. None of the scenarios develops extreme flood situations for any year from 2022 to 2050. This is very surprising. This can be due to several reasons, including the future data quality and bias correction technique. However, these strange results imply that the researchers conducted some extensive projected flood analysis based on the ground-measured flow situations. In addition, the R-R model can be implemented for Representative Concentration Pathways (RCPs) and then analyze the differences.

VI. DISCUSSION

A. MODEL EVALUATIONS

According to the sub-figure (i) in Figure 7, it can be seen herein that the GRU algorithm with an R of 0.9301 performed the best prediction. In addition, the LSTM algorithm with 2-day back rainfall data (t-2 scenario) performed as the second best with 0.9265 (refer to the sub-figure (l) in Figure 7). Interestingly, as per sub-figure (b) in Figure 7, the Cascaded-ANFIS algorithm showcased its highest R-value at 0.9140 for 1-day back rainfall data (t-1 scenario). However, it can be clearly understood that three scenarios separately cannot be used to model the R-R relationship. In other words, the rainfall which occurs two days back for the most upstream location can reach Manampitiya on the current day. Similarly, rainfall received one day back in another location can reach Manampitiya on the current day. Therefore, a combination of these three scenarios has to be considered.

As in the selected rainfall gauge analysis, it was clear that the results were more consistent and accurate. The Cascaded-ANFIS algorithm-based prediction model had

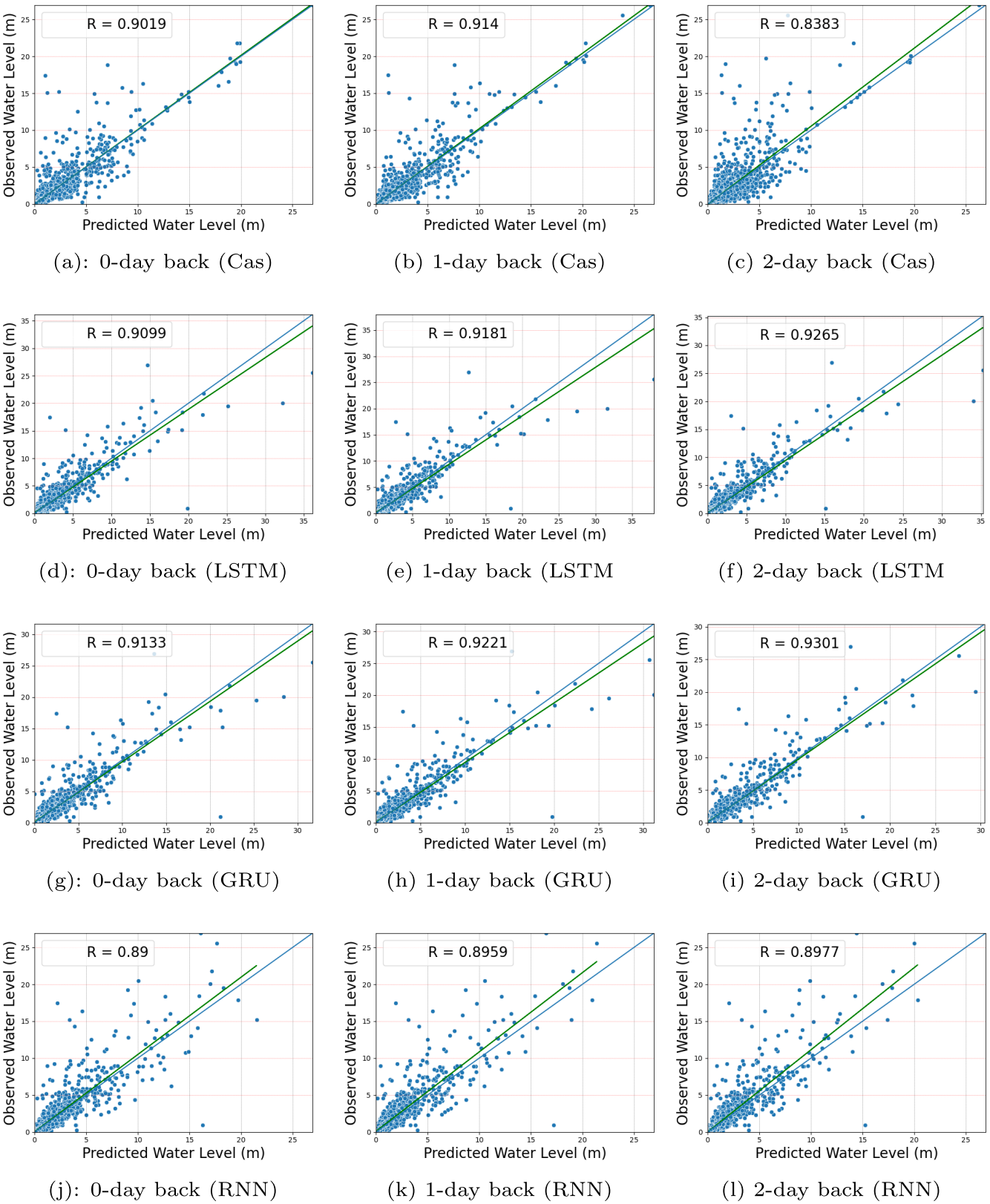
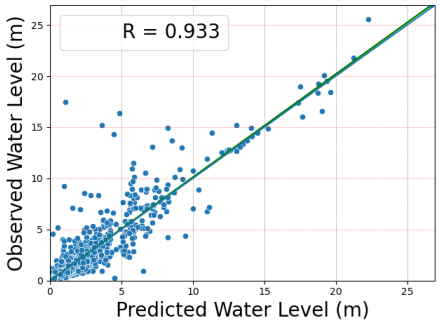
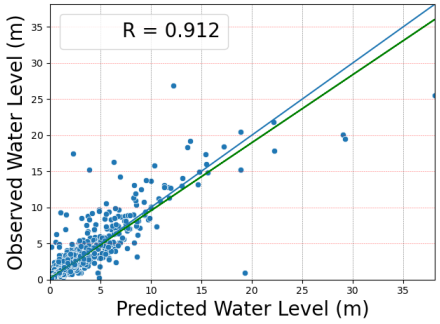


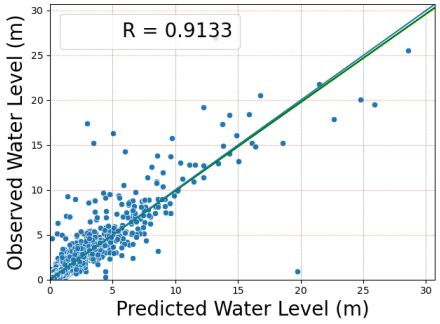
FIGURE 7. **Mahaweli catchment water level prediction and observed values with calculated correlation coefficient (R): 0-days (current day inputs), 1-day (the current day and past 1-day inputs), and 2-day (the current day and past one and 2-day inputs).**



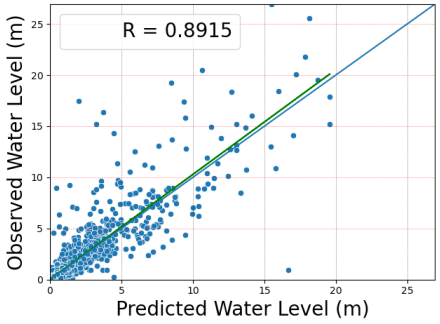
(a): Cascaded-ANFIS



(b) LSTM

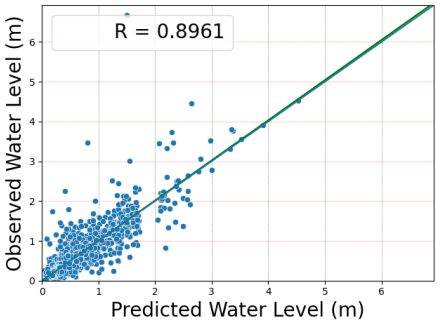


(c) GRU

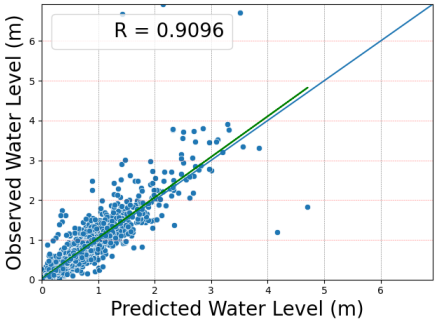


(d): RNN

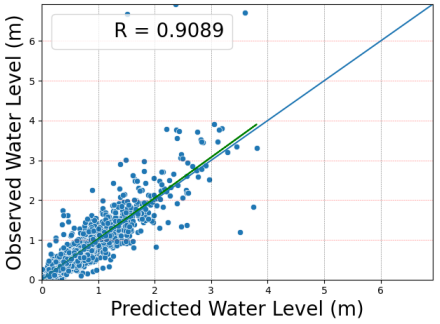
FIGURE 8. Prediction accuracy for water levels combined scenarios at Manampitiya: (a) Cascaded-ANFIS test predictions for selected inputs; (b) LSTM test predictions for selected inputs; (c) GRU test predictions for selected inputs; (d) RNN test predictions for selected inputs.



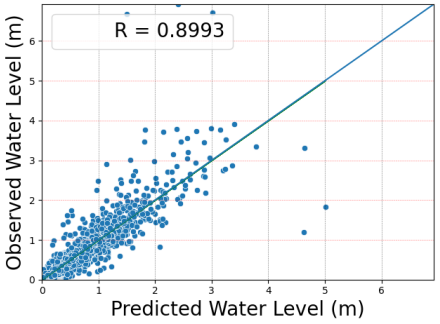
(a): Cascaded-ANFIS



(b) LSTM



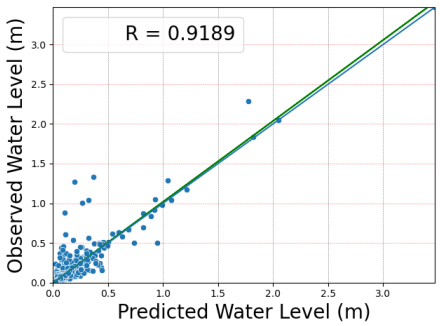
(c) GRU



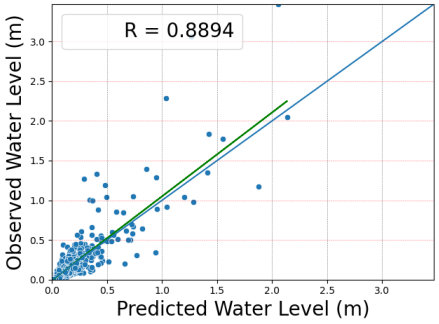
(d): RNN

FIGURE 9. Prediction accuracy for water levels at Peradeniya; (a) Cascaded-ANFIS test predictions; (b) LSTM test predictions; (c) GRU test predictions; (d) RNN test predictions.

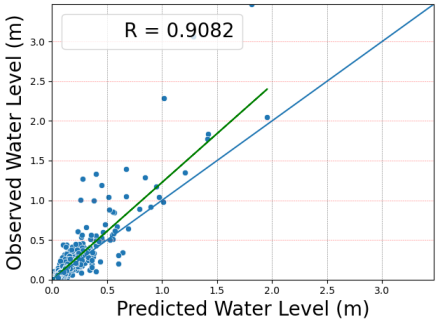
an R of 0.933 for selected inputs (refer to the sub-figure (a) in Figure 8). GRU, LSTM, and RNN showed R values of 0.9133, 0.9120, and 0.8915, respectively, and were outperformed by Cascaded-ANFIS. Therefore, the Cascaded-ANFIS algorithm can be used effectively to predictions of water levels.



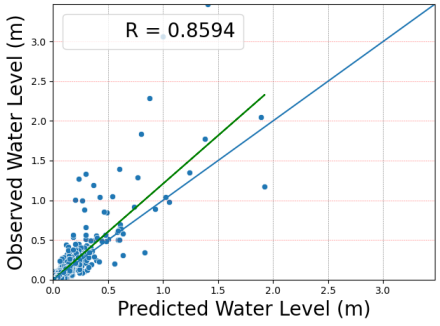
(a): Cascaded-ANFIS



(b) LSTM



(c) GRU



(d): RNN

FIGURE 10. Prediction accuracy for water levels at Thaldena; (a) Cascaded-ANFIS test predictions; (b) LSTM test predictions; (c) GRU test predictions; (d) RNN test predictions.

The sub-catchment correlation coefficient analysis in Figures 9 and 10 shows that the Cascaded-ANFIS algorithm has outperformed the other three algorithms in predicting water levels at the sub-catchment level. In Figure 10, the correlation coefficients were found to be 0.9188 for Cascaded-ANFIS, 0.8894 for LSTM, 0.9082 for GRU, and 0.8594 for RNN. Therefore, the water level prediction for the Thaldena sub-catchment also succeeded by the prediction model developed based on the Cascaded-ANFIS algorithm.

TABLE 5. The evaluation results of the study; Percent bias value (Bias), Root Mean Squared Error (RMSE), Nash-Sutcliffe efficiency (NSE), Kling-Gupta efficiency (KGE), and Correlation Coefficient (R).

Algorithm	Bias	RMSE	NSE	KGE	R
Cascade-ANFIS	1.5156	0.6591	0.8704	0.8979	0.9330
GRU	4.0945	0.7895	0.8335	0.8791	0.9133
Lasso R	-0.6269	1.4332	0.4480	0.5869	0.6734
LSTM	8.8764	0.8057	0.8266	0.8711	0.9119
Linear R	-0.6673	1.4171	0.4638	0.5558	0.6811
Ridge R	-0.6673	1.4171	0.4638	0.5558	0.6811
RNN	6.6573	0.8800	0.7932	0.8180	0.8914

The proposed algorithm shows the least RMSE with 0.66. The proposed algorithm also scored the highest NSE and KGE values, with 0.87 and 0.90. The second-best performances were shown by the GRU algorithm having RMSE, NSE, and KGE as 0.79, 0.83, and 00.88. When considering the bias factor of the predicted outputs, the Cascaded-ANFIS model shows a significantly low value of 1.52. This low score for the bias provides a certification that the model can predict the water levels with higher accuracy and lower bias. The overall results are shown in Table 5. It is also clear that the Linear, Ridge, and Lasso algorithms' performances are significantly miniature compared to the other LSTM, GRU, RNN, and Cascaded-ANFIS algorithms.

B. FORECASTING OF THE RIVER WATER LEVEL

Let the predictions be accurate (assumed). Then, there is a severe issue in the water levels, thus the river flow at Manampitiya. The average water levels for Manampitiya are around 10 m (from its historical data). However, the projected water levels are around 6 m (60% of the average). Therefore, drought conditions can be projected. The predicted outcomes of the trained model can be a result of the dataset. The dataset provides a short range of rainfall data. Therefore, more than the sample size may be needed to train a perfect R-R model. However, this cannot be considered a conclusion of this study. Even though the prediction accuracy is good in the Cascaded-ANFIS model, future data quality is critical in a solid prediction. Therefore, Figures 11 cannot be treated as a conclusion of this study.

However, these water levels were presented in Figure 12 shows the forecasting of water levels at Manampitiya for the projected rainfalls. From the year 2031 to 2050, forecasting is shown in sub-figures (c) and (d) in Figure 12 respectively for SSP2-4.5 and SSP5-8.5. The X-axis contains 365 ticks representing days of the year, and the scale bar on the right side of Figure 12 showcases the intensity of the water level. During the northeaster monsoon (December to February), the water levels can be observed at higher levels, as predicted at

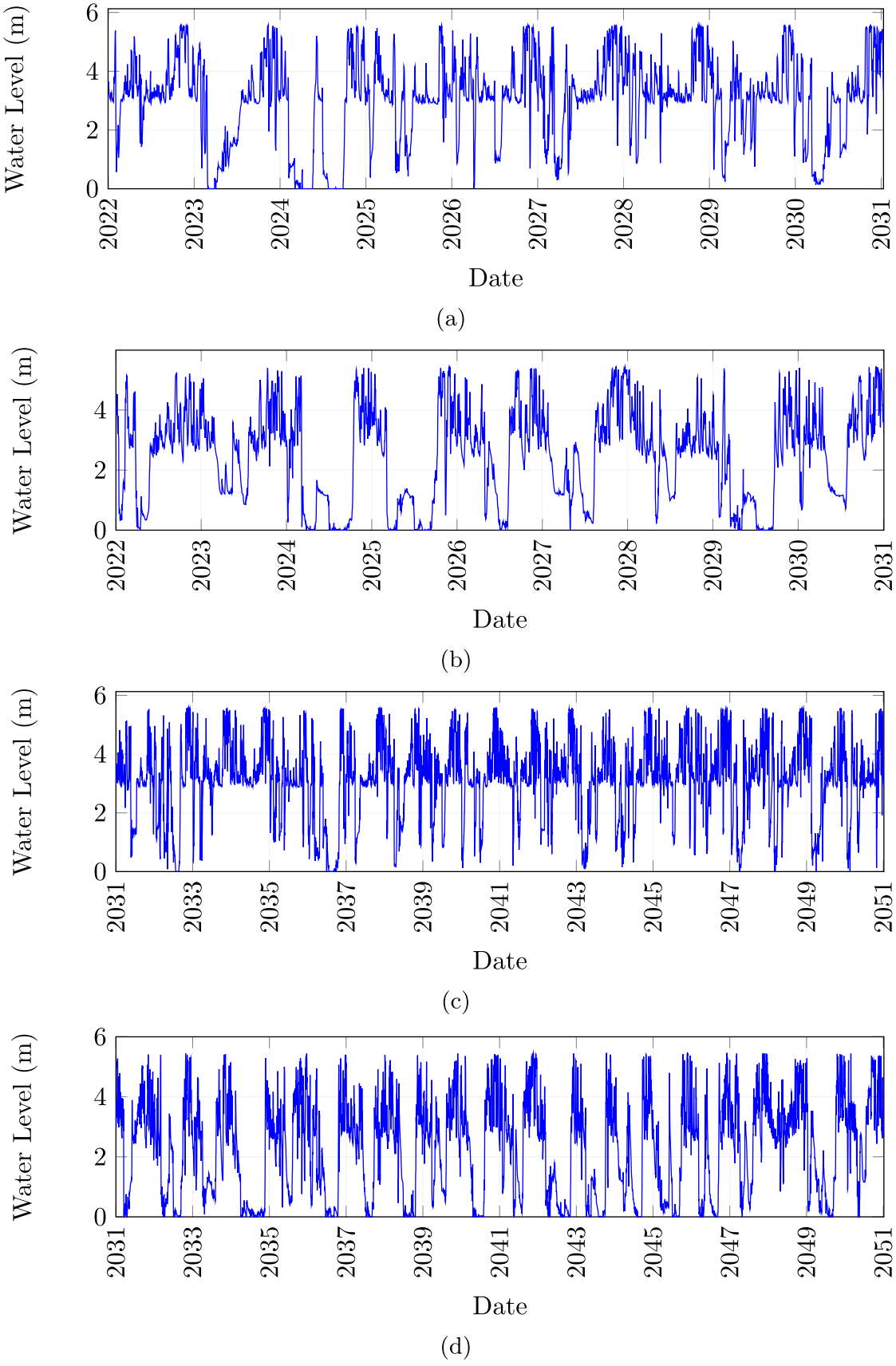


FIGURE 11. Projected future water levels: (a) for near-future at SSP2-4.5; (b) for near-future at SSP5-8.5; (c) for mid-future at SSP2-4.5; (b) for mid-future at SSP5-8.5.

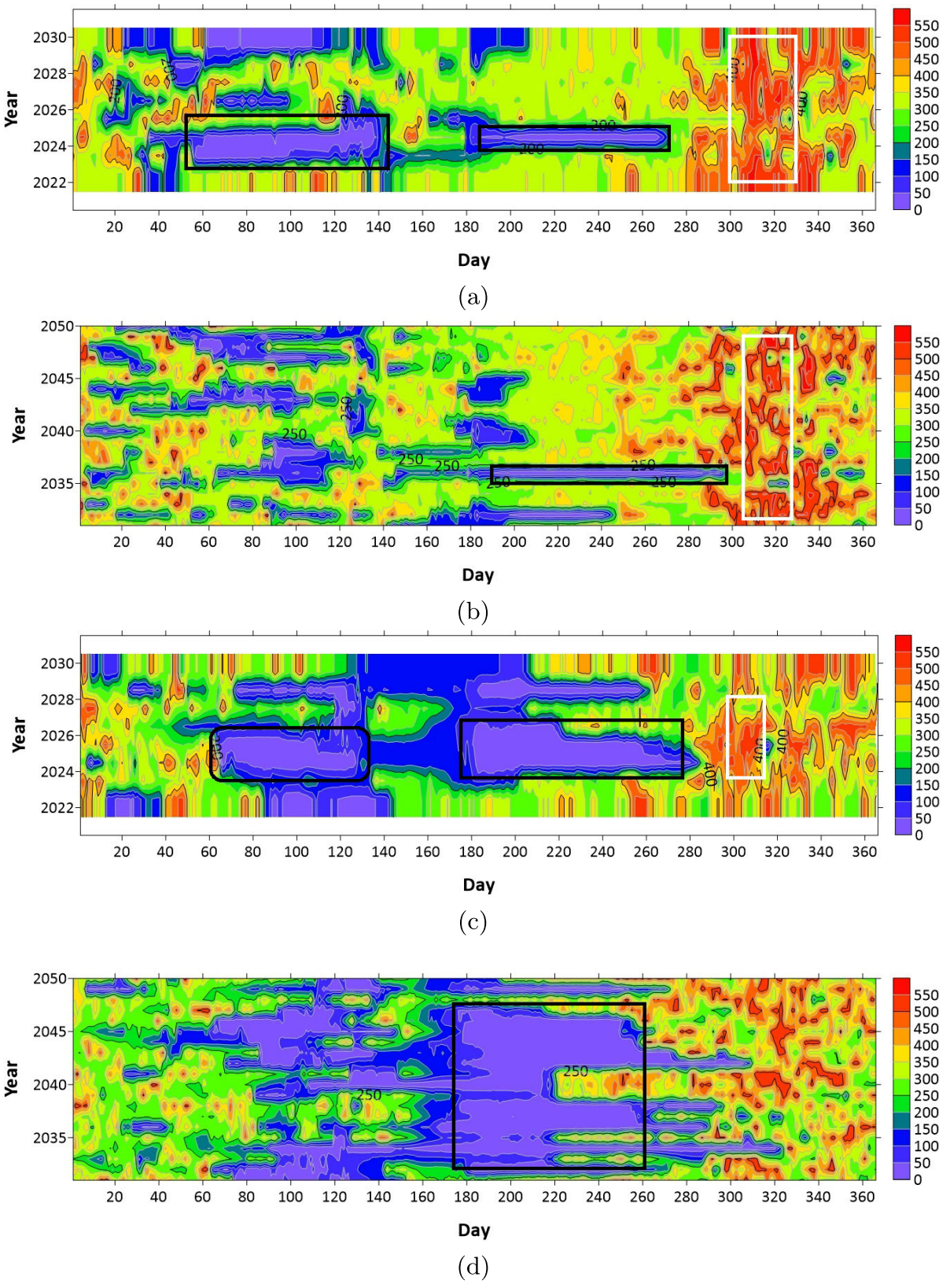


FIGURE 12. Water level predictions in a graphical way: (a) SSP2-4.5 for the year 2022 to 2030; (b) SSP5-8.5 for the year 2022 to 2030; (c) SSP2-4.5 for the year 2031 to 2050;(d) SSP5-8.5 for the year 2031 to 2050. Here, the color map notation indicate as: Red = Highest, Purple = Lowest. The black and white boxes indicate the regions of interest of lowest and highest water levels, consequently.

Manampitiya. However, the SSP5-8.5 scenario has projected lower water levels for mid-year, reaching less than 1 m. These can be droughts. However, the SSP5-8.5 is a higher scenario for fossil-fueled development. This observation cannot be seen in the SSP2-4.5 scenario. The key observations are indicated using black and white squares where black being lower water level periods and white being higher water level periods. Nevertheless, as discussed, more research is needed for a solid conclusion on future water levels.

VII. CONCLUSION

An R-R prediction model was developed using the Cascaded-ANFIS algorithm for the Mahaweli River, the longest river in Sri Lanka. The R-R model was developed for the sub-catchment levels as well. The dataset used in the case study was well evaluated using four different methods of homogeneity tests Standard normal homogeneity test (SNHT), Buishand range (BR) test, Pettitt test, and von Neumann ratio (VNR) test. The algorithm was tested against six regression algorithms used in most past studies: Linear regression, Ridge regression, Lasso regression GRU, LSTM, and RNN. The results were comparatively studied using correlation coefficient, bias, RMSE, NSE, and KGE. The highest correlation coefficient was recorded by the Cascaded-ANFIS when utilizing the selected rainfall gauges to train the models having 0.933 where Linear, Ridge, Lasso, GRU, LSTM, and RNN showed the R values of 0.6811, 0.6811, 0.6734, 0.9133, 0.9120, and 0.8915, respectively.

Moreover, the bias value of the proposed algorithm is significantly low (1.52) compared with the other algorithms. The Cascaded-ANFIS model scored 0.66, 0.87, and 0.90 for RMSE, NSE, and KGE, respectively. These results outperformed the other algorithms used in this study.

According to the overall results, it can be concluded herein that the Cascaded-ANFIS algorithm-based prediction model has outperformed the other six algorithms. The second-best algorithm that performed well in prediction was the GRU algorithm. However, the Cascaded-ANFIS algorithm has advantages compared to the black-box regression models, such as lightweight, lower computational cost, easy real-time implementation, and efficiency. Therefore, the Cascaded-ANFIS algorithm can predict the water levels of various catchments under the requirement of measured rainfalls and water levels. More importantly, the model can be developed under mixed rainfall input along the timeline due to the upstream waterś travel time to the riverś downstream.

Overall the prediction model based on the Cascaded-ANFIS algorithm predicts accurate results using the ground-measured rainfalls. The future water levels were projected under two SSP scenarios for the Manampitya station. However, promising results were only found under the near future and mid-future SSP rainfalls. None of the years was projected to have unacceptable floods (by looking at the records). Therefore, this research does not provide any conclusions about the future projected water levels. More research is needed for a solid outcome for future water levels.

REFERENCES

[1] F.-J. Chang, J.-M. Liang, and Y.-C. Chen, "Flood forecasting using radial basis function neural networks," *IEEE Trans. Syst., Man Cybern., C, Appl. Rev.*, vol. 31, no. 4, pp. 530–535, Nov. 2001.

[2] D. B. De Lima, M. D. C. E. Lima, and R. M. Saigado, "An empirical analysis of MLP neural networks applied to streamflow forecasting," *IEEE Latin Amer. Trans.*, vol. 9, no. 3, pp. 295–301, Jun. 2011.

[3] M. Romlay, M. Rashid, and S. A. T. Ibrahim, "Rainfall-runoff model based on ANN with LM, BR and PSO as learning algorithms," *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 971–979, Sep. 2019.

[4] M. Rajurkar, U. U. Kothyari, and U. C. Chaube, "Artificial neural networks for daily rainfall-runoff modelling," *Hydrol. Sci. J.*, vol. 47, pp. 865–877, Dec. 2002.

[5] R. S. Vilanova, S. S. Zanetti, and R. A. Cecílio, "Assessing combinations of artificial neural networks input/output parameters to better simulate daily streamflow: Case of Brazilian Atlantic rainforest watersheds," *Comput. Electron. Agricult.*, vol. 167, Dec. 2019, Art. no. 105080.

[6] W.-C. Wang, D.-M. Xu, K.-W. Chau, and S. Chen, "Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD," *J. Hydroinformatics*, vol. 15, no. 4, pp. 1377–1390, Oct. 2013.

[7] Z. M. Yaseen, M. Fu, C. Wang, W. H. M. W. Mohtar, R. C. Deo, and A. El-shafie, "Application of the hybrid artificial neural network coupled with rolling mechanism and grey model algorithms for streamflow forecasting over multiple time horizons," *Water Resour. Manage.*, vol. 32, no. 5, pp. 1883–1899, Mar. 2018.

[8] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, Art. no. e00938.

[9] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[10] G. Corani and G. Guariso, "Coupling fuzzy modeling and neural networks for river flood prediction," *IEEE Trans. Syst., Man Cybern., C, Appl. Rev.*, vol. 35, no. 3, pp. 382–390, Aug. 2005.

[11] B. Roy and M. P. Singh, "An empirical-based rainfall-runoff modelling using optimization technique," *Int. J. River Basin Manage.*, vol. 18, no. 1, pp. 49–67, Jan. 2020.

[12] V. K. Vidyarthi and S. Chourasiya, "Particle swarm optimization for training artificial neural network-based rainfall-runoff model, case study: Jardine River Basin," in *Micro-Electronics and Telecommunication Engineering*. Singapore: Springer, 2020, pp. 641–647.

[13] M. Asadnia, L. H. C. Chua, X. S. Qin, and A. Talei, "Improved particle swarm Optimization–Based artificial neural network for rainfall-runoff modeling," *J. Hydrologic Eng.*, vol. 19, no. 7, pp. 1320–1329, Jul. 2014.

[14] R. J. Bensingh, R. Machavaram, S. R. Boopathy, and C. Jebaraj, "Injection molding process optimization of a bi-aspheric lens using hybrid artificial neural networks (ANNs) and particle swarm optimization (PSO)," *Measurement*, vol. 134, pp. 359–374, Feb. 2019.

[15] Q. Fang, H. Nguyen, X.-N. Bui, T. Nguyen-Thoi, and J. Zhou, "Modeling of rock fragmentation by firefly optimization algorithm and boosted generalized additive model," *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3503–3519, Apr. 2021.

[16] H. Tian, J. Shu, and L. Han, "The effect of ICA and PSO on ANN results in approximating elasticity modulus of rock material," *Eng. Comput.*, vol. 35, no. 1, pp. 305–314, Jan. 2019.

[17] B. Jamali, M. Rasekh, F. Jamadi, R. Gandomkar, and F. Makiabadi, "Using PSO-GA algorithm for training artificial neural network to forecast solar space heating system parameters," *Appl. Thermal Eng.*, vol. 147, pp. 647–660, Jan. 2019.

[18] A. Jain and S. Srinivasulu, "Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques," *Water Resour. Res.*, vol. 40, no. 4, 2004. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2003WR002355, doi: 10.1029/2003WR002355.

[19] I. Istadi and N. A. S. Amin, "Modelling and optimization of catalytic–dielectric barrier discharge plasma reactor for methane and carbon dioxide conversion using hybrid artificial neural network—Genetic algorithm technique," *Chem. Eng. Sci.*, vol. 62, no. 23, pp. 6568–6581, Dec. 2007.

[20] S. Nandi, Y. Badhe, J. Lonari, U. Sridevi, B. S. Rao, S. S. Tambe, and B. D. Kulkarni, "Hybrid process modeling and optimization strategies integrating neural networks/support vector regression and genetic algorithms: Study of benzene isopropylation on hbeta catalyst," *Chem. Eng. J.*, vol. 97, nos. 2–3, pp. 115–129, Feb. 2004.
[21] A. L. Ahmad, I. A. Azid, A. R. Yusof, and K. N. Seetharamu, "Emission control in palm oil mills using artificial neural network and genetic algorithm," *Comput. Chem. Eng.*, vol. 28, no. 12, pp. 2709–2715, Nov. 2004.
[22] A. Adib and A. Mahmoodi, "Prediction of suspended sediment load using ANN GA conjunction model with Markov chain approach at flood conditions," *KSCE J. Civil Eng.*, vol. 21, no. 1, pp. 447–457, Jan. 2017.
[23] A. Mukerji, C. Chatterjee, and N. S. Raghuwanshi, "Flood forecasting using ANN, neuro-fuzzy, and neuro-GA models," *J. Hydrologic Eng.*, vol. 14, no. 6, pp. 647–652, Jun. 2009.
[24] Z. M. Yaseen, W. H. M. W. Mohtar, A. M. S. Ameen, I. Ebtehaj, S. F. M. Razali, H. Bonakdari, S. Q. Salih, N. Al-Ansari, and S. Shahid, "Implementation of univariate paradigm for streamflow simulation using hybrid data-driven model: Case study in tropical region," *IEEE Access*, vol. 7, pp. 74471–74481, 2019.
[25] K. Riahi, D. P. Van Vuuren, E. Kriegler, J. Edmonds, B. C. O'neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, O. Fricko, and W. Lutz, "The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview," *Global Environ. Change*, vol. 42, pp. 153–168, Jan. 2017.
[26] J. Rogelj, A. Popp, K. V. Calvin, G. Luderer, J. Emmerling, D. Gernaat, S. Fujimori, J. Strefler, T. Hasegawa, G. Marangoni, and V. Krey, "Scenarios towards limiting global mean temperature increase below 1.5C," *Nature Climate Change*, vol. 8, pp. 325–332, Apr. 2018.
[27] M. Gidden, K. Riahi, S. J. Smith, S. Fujimori, G. Luderer, E. Kriegler, D. P. Van Vuuren, M. Van Den Berg, L. Feng, D. Klein, and K. Calvin, "Global emissions pathways under different socioeconomic scenarios for use in CMIP6: A dataset of harmonized emissions trajectories through the end of the century," *Geosci. Model Develop.*, vol. 12, pp. 1443–1475, Apr. 2019. [Online]. Available: https://www.geosci-model-dev.net/12/1443/2019/
[28] M. Meinshausen, Z. Nicholls, J. Lewis, M. Gidden, E. Vogel, M. Freund, U. Beyerle, C. Gessner, A. Nauels, and N. Bauer, "The shared socioeconomic pathway (SSP) greenhouse gas concentrations and their extensions to 2500," *Geosci. Model Develop.*, vol. 13, pp. 3571–3605, Aug. 2020.
[29] A. Baghban, "Application of the ANFIS strategy to estimate vaporization enthalpies of petroleum fractions and pure hydrocarbons," *Petroleum Sci. Technol.*, vol. 34, no. 15, pp. 1359–1366, Aug. 2016.
[30] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 1, pp. 116–132, Jan. 1985.
[31] N. Rathnayake, T. L. Dang, and Y. Hoshino "A novel optimization algorithm: Cascaded adaptive neuro-fuzzy inference system," *Int. J. Fuzzy Syst.*, vol. 23, no. 7, pp. 1–17, 2021.
[32] N. Rathnayake, U. Rathnayake, T. L. Dang, and Y. Hoshino, "A cascaded adaptive network-based fuzzy inference system for hydropower forecasting," *Sensors*, vol. 22, no. 8, p. 2905, Apr. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/8/2905
[33] M. Shrestha, "Linear scaling bias correction (V.1.0)," Microsoft Excel File, 2015, doi: 10.13140/RG.2.1.3365.8967.
[34] S. Shamshirband, M. Hadipoor, A. Baghban, A. Mosavi, J. Bukor, and A. Várkonyi-Kóczy, "Developing an ANFIS-PSO model to predict mercury emissions in combustion flue gases," *Mathematics*, vol. 7, no. 10, p. 965, Oct. 2019.
[35] R. Samu and B. Akıntuğ, "Pre-disaster planning and preparedness: Drought and flood forecasting and analysis in Zimbabwe," *Water SA*, vol. 46, no. 3 July, pp. 448–457, Jul. 2020.
[36] A. Popp, K. Calvin, S. Fujimori, P. Havlik, F. Humpenöder, E. Stehfest, B. Bodirsky, J. Dietrich, J. Doelmann, and M. Gusti, "Land-use futures in the shared socio-economic pathways," *Global Environ. Change*, vol. 42, pp. 331–345, Jan. 2017.
[37] H. Schandl, Y. Lu, N. Che, D. Newth, J. West, S. Frank, M. Obersteiner, A. Rendall, and S. Hatfield-Dodds, "Shared socio-economic pathways and their implications for global materials use," *Resour., Conservation Recycling*, vol. 160, Sep. 2020, Art. no. 104866.
[38] M. M. Hameed, M. K. AlOmar, S. F. M. Razali, M. A. K. Khalaf, W. J. Baniya, A. Sharafati, and M. A. AlSaadi, "Application of artificial intelligence models for evapotranspiration prediction along the southern coast of Turkey," *Complexity*, vol. 2021, p. 20, 2021, Art. no. 8850243, doi: 10.1155/2021/8850243.
[39] X.-N. Bui, H. Nguyen, Q.-H. Tran, D.-A. Nguyen, and H.-B. Bui, "Predicting blast-induced ground vibration in quarries using adaptive fuzzy inference neural network and Moth–Flame optimization," *Natural Resour. Res.*, vol. 30, no. 6, pp. 4719–4734, Dec. 2021.
[40] J. Fan, L. Wu, F. Zhang, H. Cai, W. Zeng, X. Wang, and H. Zou, "Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China," *Renew. Sustain. Energy Rev.*, vol. 100, pp. 186–212, 2019.
[41] A. Mosavi, P. Ozturk, and K.-W. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, p. 1536, Oct. 2018.
[42] M. Shao, G. Zhao, S.-C. Kao, L. Cuo, C. Rankin, and H. Gao, "Quantifying the effects of urbanization on floods in a changing environment to promote water security—A case study of two adjacent basins in Texas," *J. Hydrol.*, vol. 589, Oct. 2020, Art. no. 125154.
[43] K. Luo and X. Zhang, "Increasing urban flood risk in China over recent 40 years induced by LUCC," *Landscape Urban Planning*, vol. 219, Mar. 2022, Art. no. 104317.
[44] G. E. Hollis, "The effect of urbanization on floods of different recurrence interval," *Water Resour. Res.*, vol. 11, no. 3, pp. 431–435, Jun. 1975.
[45] Y. Bayazıt, C. Koç, and R. Bakış, "Urbanization impacts on flash urban floods in bodrum province, Turkey," *Hydrological Sci. J.*, vol. 66, no. 1, pp. 118–133, Jan. 2021.
[46] B. Feng, Y. Zhang, and R. Bourke, "Urbanization impacts on flood risks based on urban growth data and coupled flood models," *Natural Hazards*, vol. 106, no. 1, pp. 613–627, Mar. 2021.
[47] M. Noor, T. Ismail, E.-S. Chung, S. Shahid, and J. Sung, "Uncertainty in rainfall intensity duration frequency curves of peninsular Malaysia under changing climate scenarios," *Water*, vol. 10, no. 12, p. 1750, Nov. 2018.
[48] R. Bhatla, S. Verma, R. Pandey, and A. Tripathi, "Evolution of extreme rainfall events over indo-gangetic plain in changing climate during 1901–2010," *J. Earth Syst. Sci.*, vol. 128, no. 5, pp. 1–14, Jul. 2019.
[49] J. Z. Khan and M. Zaheer, "Hydrological response to changing climate: A case study of large watershed, xianning, hubei, China," *Geol., Ecol., Landscapes*, vol. 6, no. 2, pp. 141–147, Apr. 2022.
[50] A. F. Prein, R. M. Rasmussen, K. Ikeda, C. Liu, M. P. Clark, and G. J. Holland, "The future intensification of hourly precipitation extremes," *Nature Climate Change*, vol. 7, no. 1, pp. 48–52, Jan. 2017.
[51] S. Wongchuig, J. C. Espinoza, T. Condom, H. Segura, J. Ronchail, P. A. Arias, C. Junquas, A. Rabatel, and T. Lebel, "A regional view of the linkages between hydro-climatic changes and deforestation in the southern Amazon," *Int. J. Climatol.*, vol. 42, no. 7, pp. 3757–3775, 2022. [Online]. Available: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.7443, doi: 10.1002/joc.7443.
[52] K. Abass, G. Dumedah, F. Frempong, A. S. Muntaka, D. O. Appiah, E. K. Garsonu, and R. M. Gyasi, "Rising incidence and risks of floods in urban ghana: Is climate change to blame?" *Cities*, vol. 121, Feb. 2022, Art. no. 103495.
[53] B. Mohammadi, "A review on the applications of machine learning for runoff modeling," *Sustain. Water Resour. Manage.*, vol. 7, no. 6, pp. 1–11, Dec. 2021.
[54] S. Chen, M. Ren, and W. Sun, "Combining two-stage decomposition based machine learning methods for annual runoff forecasting," *J. Hydrol.*, vol. 603, Dec. 2021, Art. no. 126945.
[55] H. M. V. V. Herath, J. Chadalawada, and V. Babovic, "Hydrologically informed machine learning for rainfall-runoff modelling: Towards distributed modelling," *Hydrol. Earth Syst. Sci.*, vol. 25, pp. 4373–4401, Aug. 2021.
[56] R. M. Adnan, A. Petroselli, S. Heddam, C. A. G. Santos, and O. Kisi, "Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model," *Stochastic Environ. Res. Risk Assessment*, vol. 35, no. 3, pp. 597–616, Mar. 2021.
[57] F. Fahimi, Z. M. Yaseen, and A. El-shafie, "Application of soft computing based hybrid models in hydrological variables modeling: A comprehensive review," *Theor. Appl. Climatol.*, vol. 128, nos. 3–4, pp. 875–903, May 2017.
[58] M. Rezaie-Balf, Z. Zahmatkesh, and S. Kim, "Soft computing techniques for rainfall-runoff simulation: Local Non–Parametric paradigm vs. Model classification methods," *Water Resour. Manage.*, vol. 31, no. 12, pp. 3843–3865, Sep. 2017.
[59] O. Kisi, J. Shiri, and M. Tombul, "Modeling rainfall-runoff process using soft computing techniques," *Comput. Geosci.*, vol. 51, pp. 108–117, Feb. 2013.

[60] P. Arias, N. Bellouin, E. Coppola, R. Jones, G. Krinner, J. Marotzke, V. Naik, M. Palmer, G. Plattner, and J. Rogelj, "Others climate change 2021: The physical science basis. contribution of working group14 I to the sixth assessment report of the intergovernmental panel on climate change; technical summary," Weltklimarat, Oberpfaffenhofen, Tech. Rep., 2021.
[61] S. Dangol, R. Talchabhadel, and V. P. Pandey, "Performance evaluation and bias correction of gridded precipitation products over arun river basin in Nepal for hydrological applications," *Theor. Appl. Climatol.*, vol. 148, nos. 3–4, pp. 1353–1372, May 2022.
[62] T. Lafon, S. Dadson, G. Buys, and C. Prudhomme, "Bias correction of daily precipitation simulated by a regional climate model: A comparison of methods," *Int. J. Climatol.*, vol. 33, no. 6, pp. 1367–1381, May 2013.
[63] I. M. Chathuranika, M. B. Gunathilake, H. M. Azamathulla, and U. Rathnayake, "Evaluation of future streamflow in the upper part of the nilwala river basin (Sri Lanka) under climate change," *Hydrology*, vol. 9, no. 3, p. 48, Mar. 2022.
[64] M. M. G. T. De Silva and A. Kawasaki, "Modeling the association between socioeconomic features and risk of flood damage: A local-scale case study in Sri Lanka," *Risk Anal.*, Feb. 2022.
[65] T. A. J. G. Sirisena, S. Maskey, J. Bamunawala, E. Coppola, and R. Ranasinghe, "Projected streamflow and sediment supply under changing climate to the coast of the kalu river basin in tropical sri Lanka over the 21st century," *Water*, vol. 13, no. 21, p. 3031, Oct. 2021.
[66] S. Shelton and Z. Lin, "Streamflow variability in Mahaweli river basin of Sri Lanka during 1990–2014 and its possible mechanisms. *Water*, vol. 11, p. 2485, Nov. 2019.
[67] S. Withanachchi, S. Köpke, C. Withanachchi, R. Pathiranage, and A. Ploeger, "Water resource management in dry zonal paddy cultivation in Mahaweli River Basin, Sri lanka: An analysis of spatial and temporal climate change impacts and traditional knowledge," *Climate*, vol. 2, no. 4, pp. 329–354, Nov. 2014.
[68] S. Chandrasekara, V. Prasanna, and H.-H. Kwon, "Monitoring water resources over the kotmale reservoir in sri Lanka using ENSO phases," *Adv. Meteorol.*, vol. 2017, pp. 1–9, Jan. 2017.
[69] N. Rathnayake, U. Rathnayake, T. L. Dang, and Y. Hoshino, "An efficient automatic fruit-360 image identification and recognition using a novel modified cascaded-ANFIS algorithm," *Sensors*, vol. 22, no. 12, p. 4401, Jun. 2022.
[70] L. Karamoutsou and A. Psilovikos, "Deep learning in water resources management: The case study of Kastoria lake in Greece," *Water*, vol. 13, no. 23, p. 3364, Nov. 2021.
[71] M. Amadori, V. Zamparelli, G. De Carolis, G. Fornaro, M. Toffolon, M. Bresciani, C. Giardino, and F. De Santi, "Monitoring lakes surface water velocity with SAR: A feasibility study on lake Garda, Italy," *Remote Sens.*, vol. 13, no. 12, p. 2293, Jun. 2021.
[72] A. Sentas, A. Psilovikos, T. Psilovikos, and N. Matzafleri, "Comparison of the performance of stochastic models in forecasting daily dissolved oxygen data in dam-lake thesaurus," *Desalination Water Treatment*, vol. 57, no. 25, pp. 11660–11674, May 2016.
[73] D. I. A. Velasquez, M. Pulido-Velazquez, and M. S. Hector, "Improvement of water management for irrigation in Mediterranean basins combining remote sensing, weather forecasting, and artificial intelligence," 2022.
[74] H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling," *J. Hydrol.*, vol. 377, nos. 1–2, pp. 80–91, Oct. 2009.
[75] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models Part I—A discussion of principles," *J. Hydrol.*, vol. 10, no. 3, pp. 282–290, Apr. 1970.
[76] J.-S.-R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 3, pp. 665–685, May 1993.
[77] D. Parker, *Floods*. Evanston, IL, USA: Routledge, 2014.
[78] V. Panwar and S. Sen, "Disaster damage records of EM-DAT and DesInventar: A systematic comparison," *Econ. Disasters Climate Change*, vol. 4, no. 2, pp. 295–317, Jul. 2020.
[79] R. Taylor, "Interpretation of the correlation coefficient: A basic review," *J. Diagnostic Medical Sonography*, vol. 6, no. 1, pp. 35–39, 1990, doi: 10.1177/875647939000600106.
[80] R. Taylor, "Interpretation of the correlation coefficient: A basic review," *J. Diagnostic Med. Sonography*, vol. 6, pp. 35–39, Jun. 1990.

NAMAL RATHNAYAKE received the B.Sc. degree in electronics from the University of Wolverhampton, in 2016, and the master's degree in robotics and automation from the University of Salford Manchester, in 2019. He is currently pursuing the Ph.D. degree with the Department of Computer and System Engineering, Kochi University of Technology (KUT), Japan.

His research interests include fuzzy logic control, artificial intelligence, and UAV robotics. He was a recipient of the Special Scholarship Program from KUT for the Ph.D. program, in 2020.

UPAKA RATHNAYAKE received the Ph.D. degree from the University of Strathclyde, Glasgow, U.K. (Royal Charter in 1964 as the U.K.'s First Technological University). He is currently a faculty member and Researcher with the Department of Civil Engineering and Construction in Faculty of Engineering and Design, Atlantic Technological University, Sligo, Ireland. He is working on various research topics related to the sustainability of water resources management.

IMIYA CHATHURANIKA received the master's degree in water engineering and management from the Asian Institute of Technology (AIT), Thailand. She is currently a Research Assistant with the Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka. Her current research interests include hydrology, hydrological modeling, climate change and variability, hydropower production, reservoir operation, water-saving agricultural practices, and water accessibility.

TUAN LINH DANG received the Ph.D. degree in computer science from the Kochi University of Technology (KUT), in 2017.

He is currently a Lecturer with the School of Information and Communication Technology, Hanoi University of Science and Technology. His current research interests include machine learning, computer vision, hardware/software co-design, and FPGA. He was a recipient of the Special Scholarship Program from KUT for a Ph.D. program, in 2014, and the Japanese Government Scholarship for a master's program, in 2012.

YUKINOBU HOSHINO (Member, IEEE) received the B.Sc. degree from Westmar University, in 1995, and the master's and Ph.D. degrees from Ritsumeikan University, in 1998 and 2002, respectively.

He is currently an Associate Professor with the Department of System Engineering, Kochi University of Technology, Japan. His current research interests include intelligent systems and machine learning based on SOFT computing for SoC. He is a member of many reputed societies, including the Japan Society for Fuzzy Theory and intelligent informatics (SOFT), the Institute of Systems, Control and Information Engineers (ISCIE), the Society of Instrument and Control Engineers (SICE), and the Japan Society of Kansei Engineering (JSKE).