

Multivariable NARX Based Neural Networks Models for Short-Term Water Level Forecasting [†]

Jackson B. Renteria-Mena ¹, Douglas Plaza ² and Eduardo Giraldo ^{3,*}

¹ Facultad de Ingeniería, Universidad Tecnológica Del Chocó, Quibdó 270001, Colombia; jackson.renteria@utp.edu.co

² Facultad de Ingeniería en Electricidad y Computación, Escuela Superior Politécnica del Litoral, Guayaquil 090902, Ecuador; douplaza@espol.edu.ec

³ Research Group in Automatic Control, Electrical Engineering Department, Universidad Tecnológica de Pereira, Pereira 660003, Colombia

* Correspondence: egiraldo@utp.edu.co

[†] Presented at the 9th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 12–14 July 2023.

Abstract: In this work a novel application for multivariable forecasting is presented, applied to hydrological variables and based on a multivariable NARX model. The proposed approach is designed for two hydrological stations located at the Atrato River in Colombia where the variables of water level, water flow and water precipitation are correlated by using the NARX model based on a neural network structure. The structure of the NARX-based neural network is designed in order to consider the complex dynamics of hydrological variables and their corresponding cross-correlations. A short-term water level forecasting is designed based on the NARX model, to be used as an early warning flood system. The validation of the proposed approach is performed by comparing the estimation error with an ARX dynamic model. As a result, it is shown that a NARX model structure is more suitable for water level forecasting than simplified structures.

Keywords: forecast; water level; neural network



Citation: Renteria-Mena, J.B.; Plaza, D.; Giraldo, E. Multivariable NARX Based Neural Networks Models for Short-Term Water Level Forecasting. *Eng. Proc.* **2023**, *39*, 60. <https://doi.org/10.3390/engproc2023039060>

Academic Editors: Ignacio Rojas, Hector Pomares, Luis Javier Herrera, Fernando Rojas and Olga Valenzuela

Published: 5 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The accurate modeling of hydrological variables is crucial for effective flood forecasting and the design and operation of water resource systems, as stated in the reference [1]. To achieve this, two types of techniques are commonly used: white-box algorithms, which rely on mathematical modeling, and black-box algorithms, which employ non-linear neural network techniques based on artificial intelligence. In the context of early warning systems, the latter technique has been used to great effect, as highlighted in [2], which discussed the use of artificial neural networks (ANNs) for prediction and forecasting. Furthermore, the combination of ANNs with the Soil and Water Assessment Tool (SWAT) has been applied for runoff prediction and water management resources, as described in [3].

Hydrological models have been developed to support flood early warning systems through the use of estimation and prediction algorithms. In [4], the authors focused on developing a flood forecasting system (FFS) capable of providing early warning to UDS managers of potential flooding, using a nonlinear autoregressive neural network with exogenous inputs (NARX) to predict the impact of a storm. Meanwhile, in [5], a neural network short-term memory model (LSTM) was proposed for flood forecasting, using daily flow and rainfall as input data, and analyzing features that may affect the model's performance. In [6], the authors presented flood early warning systems that used machine learning (ML) techniques, comparing the performance of five ML classification techniques for short-duration flood forecasting. Lastly, Bande and Shete [7] described an IoT-based flood monitoring and flood prediction system based on artificial neural networks (ANN),

which aimed to monitor humidity, temperature, pressure, rainfall, and water level of rivers and analyze their temporal correlation for flood prediction. The system was designed to improve the scalability and reliability of the flood management system.

Several studies related to flood forecasting and monitoring using various ANN techniques have also been developed. For example, in [8], the authors evaluated the bias correction of real-time precipitation data and the improvement of hydrological models using the ANN bias correction method for real-time flood forecasting. In [9], the authors focused on developing five different ANN models for flood forecasting and compared their performance. In [10], the authors use a multilayer perceptron to design a flood prediction model with flow as input-output variables, and the proposed model's effectiveness was demonstrated through intensive experiments. In [11], the authors designed a flood monitoring system that integrated flow and water level sensors and used a two-class neural network to predict flood status from data stored in the database. Finally, in [12], the authors employed a convolutional neural network (CNN) to predict time series variables such as water level in a flood model, despite CNNs typically being used for two-dimensional image classification with transfer learning.

In addition, in [13], a flood forecasting model that predicted future flood occurrence was designed and evaluated by constructing a hybrid deep learning algorithm called ConvLSTM, which integrated the predictive merits of CNN and a long-term memory network (LSTM). In [14], a fuzzy neural network that used fuzzy numbers to account for uncertainty in the results and model parameters was proposed to predict the peak flow in an urban river. In [15], the potential of the AI computational paradigm for modeling streamflow was explored by developing nine different flood prediction models using all available training algorithms of ANN, fuzzy logic, and adaptive neuro-fuzzy inference systems (ANFIS) algorithms. Lastly, in [16], a deep neural network was used to predict floods as a function of temperature and rainfall intensity, and its accuracy and error were compared with other machine learning models, such as the support vector machine (SVM), K-nearest neighbor (KNN), and Naïve Bayes.

The use of real-time methods based on ANN was also proposed based on neural networks. For example, in [17], a system for predicting flood levels was developed based on real-time sensor data. The system used a multi-layer artificial neural network model created with MATLAB to predict flood levels in advance using data collected from sensors in a real-time monitoring system. In [18], a hybrid river flood forecasting model was presented using time series analysis and artificial neural networks to explain and forecast daily water discharge of the Mohawk River in New York. Multiple linear regression (MLR) models and an ANN model were used to describe each component for predicting the water discharge time series. In [19], five alternative machine learning techniques were used to improve the hydrological model, including linear regression, neural network regression, Bayesian linear regression, and reinforced decision tree regression, with the MIKE-11 hydrologic forecasting model used as a test system. In [20], a machine learning method was presented that uses historical typhoon paths to predict flood hydrographs of a Taiwanese watershed. Finally, in [21], a general framework for probabilistic flood forecasting was introduced, which uses an unaccented Kalman filter (UKF) postprocessing technique to model point forecasts made with a recurrent neural network and their corresponding observations. The methodology was tested using a 6 h long-term time series of the Three Gorges reservoir in China.

In this work, the application of a multivariable NARX model based on neural network for short-term water level forecasting along the Atrato river in Colombia is presented and evaluated. To this end, the present work uses data from two hydrological stations located in the Atrato river, which are monitored by the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM). The data includes measurements of flow, precipitation, and water levels sampled every 12 h over a period of 789 days. The multivariable NARX model is trained to predict the water levels of each station based on the inputs of water level, water flow and water precipitation by considering the inherent dynamic and correlation of

the process. The performance of the models is evaluated based on the mean square error of the estimated outputs compared to the actual data. The performance of the proposed approach is compared to a multivariable ARX system. The main contribution of this paper is a general design of a multivariable NARX model structure based on neural networks for short-term water level forecasting. This work is organized as follows; in Section 2 the theoretical framework where the hydrological variables and the NARX model for their corresponding dynamic approximation is proposed. In Section 3 the experimental setup and the estimation results are shown, and finally in Section 4 the conclusions and final remarks are presented.

2. Theoretical Framework

2.1. Hydrological Variables

In order to perform water level forecasting based on the dynamic of a river, two hydrological stations are located on a river in two different positions. In order to consider the correlation among all the variables of the system and their corresponding nonlinearities, a nonlinear dynamical model is proposed. In (1) the inputs and outputs of the proposed model are shown.

$$y[k] = \begin{bmatrix} y_{L_1}[k] \\ y_{L_2}[k] \end{bmatrix}, u[k] = \begin{bmatrix} u_{F_1}[k] \\ u_{PT_1}[k] \\ u_{F_2}[k] \\ u_{PT_2}[k] \end{bmatrix} \quad (1)$$

where $y_{L_1}[k]$, $y_{L_2}[k]$ correspond to the two outputs of the level variable of the two stations, $u_{F_1}[k]$, $u_{F_2}[k]$, $u_{PT_1}[k]$, $u_{PT_2}[k]$ are the four inputs of the neural network system corresponding to the two stations of the multivariable system, i.e., the $u_{F_j}[k]$ represents the j -th two inputs of the flow variable of the two stations and $u_{PT_j}[k]$ represents j -th; two more inputs of the rainfall variable of the two stations already mentioned, thus obtaining a multivariable system with four inputs and two outputs.

The dynamic of the hydrological variables is defined by considering a Nonlinear function with an Auto-Regressive and exogenous inputs (NARX) as follows:

$$y[k] = f(y[k-1], \dots, y[k-n], u[k-1], \dots, u[k-n]) + \eta[k] \quad (2)$$

being n the order of the NARX model and $f(\cdot)$ the nonlinear function, and $\eta[k]$ the additive noise at time instant k .

2.2. Narx Based Neural Network Structure

In order to consider the NARX model of (2), the inputs are selected as $u[k-j]$ and $y[k-j]$, with $j = 1, \dots, n$, which correspond to a n -th order model. In this work a 4-th order model ($n = 4$) is considered according to [22] where an analysis of the order selection is performed and the lowest estimations error is obtained for the 3-rd order model or higher. Therefore, by considering the variables described in (1), the proposed NARX model consists of 24 inputs and 2 outputs.

In order to approximate the nonlinear function of (2), the nonlinear function $f(\cdot)$ is approximated by using a neural network structure $f^*(\cdot)$, as depicted in Figure 1.

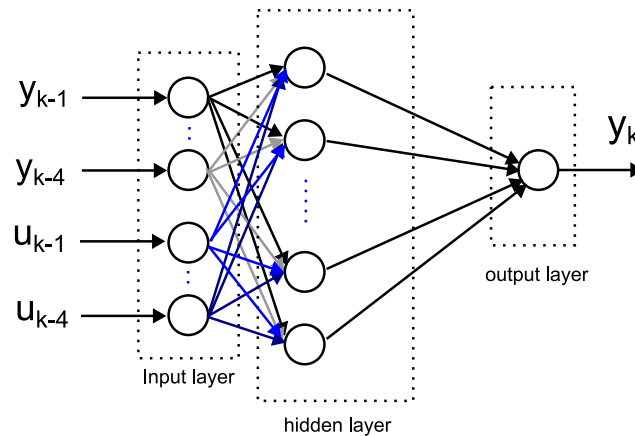


Figure 1. NARX based Neural Networks Structure.

where the NARX model can be defined as follows:

$$y[k] = f^*(y[k-1], \dots, y[k-n], u[k-1], \dots, u[k-n]) + \eta[k] \quad (3)$$

To this end, 24 input activation functions with one hidden layer and 2 outputs are considered. A feed-forward network is selected as a candidate for the NARX model in order to speed-up the training process. The training of the NARX model is performed offline by considering the data sample.

A linear ARX structure can be obtained by neglecting the hidden layer as depicted in Figure 2.

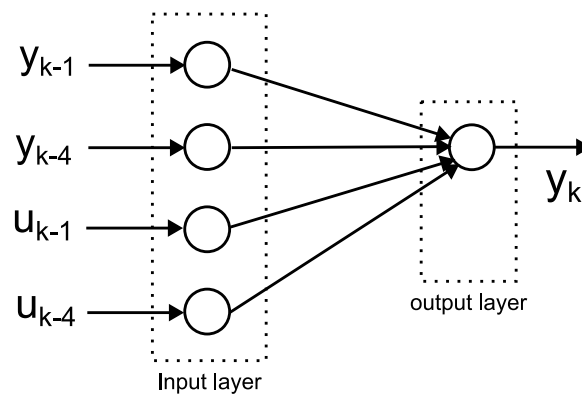


Figure 2. ARX Structure.

where the ARX model can be defined as follows:

$$y[k] = - \sum_{j=1}^4 a_j y[k-j] + \sum_{j=1}^4 b_j u[k-j] \quad (4)$$

where $A_j \in \mathbb{R}^{m \times m}$ and $B_j \in \mathbb{R}^{m \times m}$ are the parameters of the model matrix, where y are the outputs and u are the inputs; with $j = 1$, being p the order of the system, $e[k]$ the noise with m , the number of outputs and inputs of the system, $y[k] \in \mathbb{R}^{m \times 1}$ and $u[k] \in \mathbb{R}^{m \times 1}$.

3. Results

3.1. Experimental Setup

In order to validate the proposed approach, real measurements from two hydrological stations located at the Atrato river are considered. The measured variables are: water level, water flow, and water precipitation. The Atrato river is located in Colombia (South America) and has a total length of 750 km with a variable width in a range of 150 m to

500 m. In addition, the depth of the river has a variability in the range of 31m to 38 m. A total amount of 789 days of data with a sample time of 12 h is considered, with initial date 1 January 2021.

In Table 1 the geographical positions of the hydrological stations are shown.

Table 1. Location of the hydrological stations.

	Station 1 (E1)	Station 2 (E2)
Longitude	76°40'10.75" W	76°39'44.13" W
Latitude	5°45'53.38" N	5°41'52.77" N
Altitude	20.579 MASL	20.83 MASL
City	Belén de Bajirá	Quibdó

In addition, it is worth noting that the distance in kilometers between the two stations E1 and E2 is 447.1 km.

In order to validate the proposed approach, a comparison analysis of the proposed NARX approach based on neural networks (3) is performed with a multivariable ARX model (4). A visual comparison of the real and estimated signals is presented for the ARX and NARX methods and also a quantitative evaluation based on the mean squared error is performed. Both systems (ARX and NARX) are trained offline by considering the measurement data. An evaluation in terms of the training error is also presented. The feed-forward network structure for the NARX approach considers one hidden layer with 256 units. The Rectified Linear Unit (ReLU) Activation Function is selected for the proposed approach, where the ReLU is a piecewise linear function that will output the input directly if it is positive, or zero otherwise.

The implementation of the proposed NARX model based on neural networks and also the ARX model is performed in Python by using Tensorflow, which is an open-source machine learning library developed by Google.

3.2. Estimation Results

In this subsection the forecasting results for the two considered methods are presented: The proposed multivariable NARX approach, and the multivariable ARX.

In Figure 3 the estimation results for the ARX method for each of the two water level outputs are presented. In Figure 3a the short-term estimation of the first water level output and also the real measurements are shown. In Figure 3b the short-term water level forecasting of the second output as well as the real corresponding measurements are depicted.

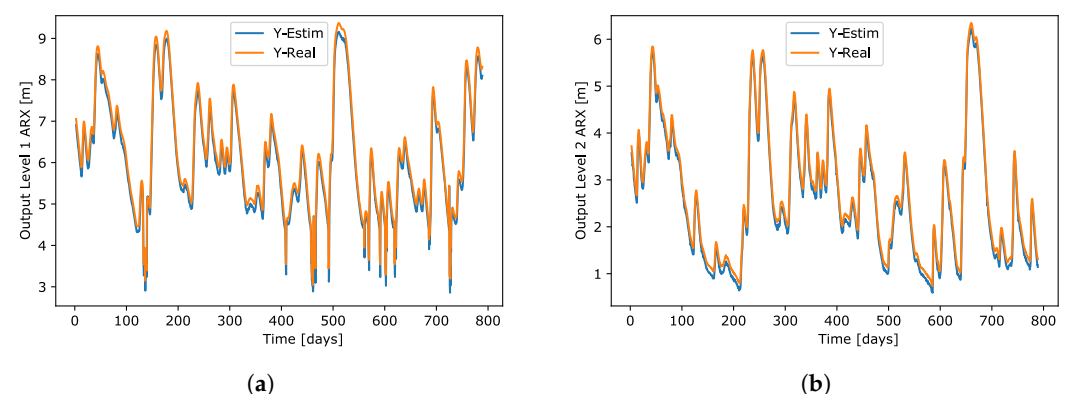


Figure 3. Multivariable ARX short-term forecasting results for two water level outputs. (a) First water level output and short-term estimation based on a multivariable ARX model. (b) Second water level output and short-term estimation based on a multivariable ARX model.

In Figure 3a,b shows that the real measurements are adequately estimated by using the multivariable ARX model.

In Table 2 an analysis for the nonlinear neural network NARX in terms of the number of nodes in the hidden layer and their corresponding mean squared estimation error is shown.

From Table 2 it can be seen that the total estimation error is reduced by increasing the number of nodes in the hidden layer. It is noticeable that there is no significant reduction in the total estimation error between 256 and 512 nodes. Therefore, in this work, 256 nodes in the hidden layer are used for evaluation of the NARX models.

In Figure 4 the estimation results for the proposed multivariable NARX method for each of the two water level outputs are presented, by using 256 nodes at the hidden layer according to the results presented in Table 2. In Figure 4a the short-term estimation of the first water level output and also the real measurements are presented. In Figure 4b the short-term water level forecasting of the second output as well as the real corresponding measurements are presented.

Table 2. Mean squared Estimation error for several nodes configurations.

NARX Hidden Layer Nodes	Level 1	Level 2	Total
2	1.6867	1.6859	3.3726
4	0.0266	0.0275	0.0541
8	0.0254	0.0272	0.0526
16	0.0227	0.0257	0.0484
32	0.0232	0.0201	0.0433
64	0.0235	0.0180	0.0415
128	0.0176	0.0133	0.0309
256	0.0085	0.0108	0.0193
512	0.0078	0.0101	0.0179

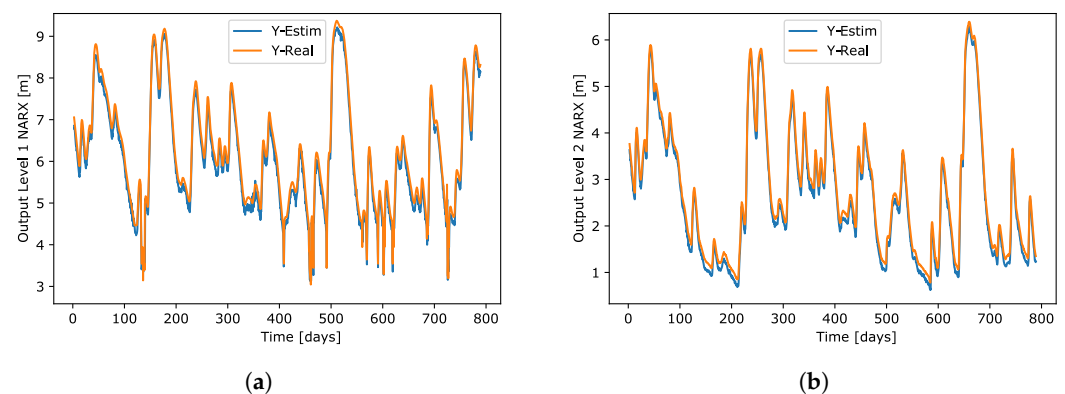


Figure 4. Multivariable NARX short-term forecasting results for two water level outputs. (a) First water level output and short-term estimation based on a multivariable NARX model. (b) Second water level output and short-term estimation based on a multivariable NARX model.

It is worth noting that in Figure 4a,b it is shown that also for the multivariable NARX model, the real measurements are adequately estimated.

By considering the forecasting results presented in Figures 3 and 4, the proposed multivariable NARX and ARX approaches show an adequate performance by visual inspection. In order to determine which approach tracks the dynamics of the hydrology variables more adequately, a quantitative evaluation is performed. To this end, the mean squared error is computed in order to compare the real measurements and their corresponding forecasting for each of the considered methods. As a result, in Table 3 the mean squared error for

the proposed multivariable NARX approach and also the multivariable ARX method are presented. It can be seen that the estimation error for the proposed NARX model is lower than the ARX model. It is worth noting that the reduction of estimation error for the NARX approach in comparison to the ARX approach is over the 50% for all variables.

In Figure 5 the estimation errors during training for each of the considered methods are shown. It can be seen that the multivariable ARX approach converge faster than the proposed multivariable NARX approach. However, the training error was lower for the NARX approach in comparison to the ARX approach. This behaviour validated the fact that there are nonlinear dynamics inherent to the measured hydrological variables and therefore the proposed NARX model forecast the data behaviour more adequately.

Table 3. Mean squared Estimation error.

Neural Network Model	Level 1	Level 2	Total
ARX	0.0280	0.0263	0.0543
NARX	0.0085	0.0108	0.0193

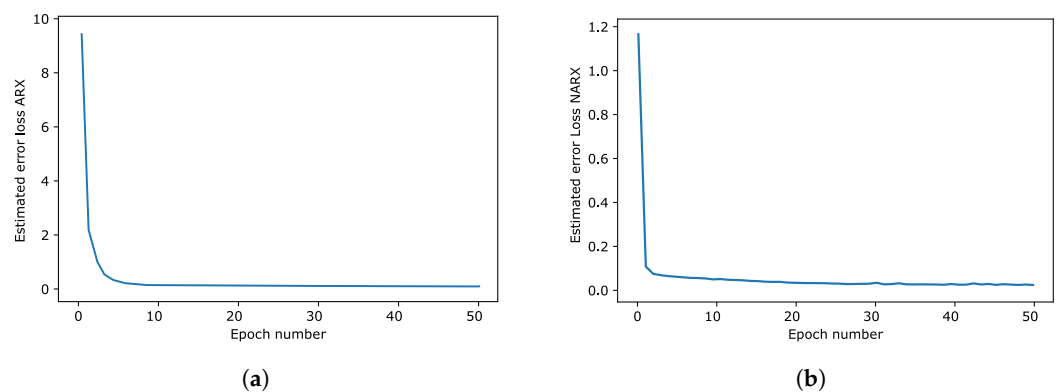


Figure 5. Training error. (a) Training error with the ARX multivariable model. (b) Training error with the NARX model.

4. Discussions and Conclusions

In this work a novel application for multivariable forecasting method for hydrological variables based on multivariable NARX model is presented. To this end, the nonlinear function of the NARX model has been approximated by using neural networks. The proposed multivariable NARX approach is compared to a multivariable ARX approach, where the proposed NARX approach shows a lower estimation error (a reduction over a 50% error as shown in Table 3). By considering the training errors, it can be seen that the training error is lower for the NARX approach in comparison to the ARX approach due to the nonlinear dynamics inherent to the measured hydrological variables. In summary, the proposed multivariable NARX model based on neural networks is an effective tool for water level forecasting by considering the correlation among several hydrological variables and several stations. It is worth noting that the main contribution of the proposed approach is the design of a general structure for modeling that can be extended to several hydrological systems with more stations and variables. In future work, online training, white-box or more complex nonlinear structures can be used to describe the nonlinear behavior of the system.

Author Contributions: All authors contributed to the conception and design of the analysis. Data acquisition was performed by J.B.R.-M. The manuscript was written by J.B.R.-M., and all authors reviewed, commented, and edited the draft version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by project no. 6-22-8 entitled “Identificación y control de sistemas multivariables interconectados a gran escala” by Universidad Tecnológica de Pereira, Pereira, Colombia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All relevant data has been made available in the paper.

Conflicts of Interest: The authors declare that there are no conflict of interest regarding the publication of this paper.

References

1. Bras, R.L.; Rodriguez-Iturbe, I. *Random Functions and Hydrology*; Courier Corporation: Chelmsford, MA, USA, 1993.
2. Palchevsky, E.; Antonov, V.; Enikeev, R.; Breikin, T. A system based on an artificial neural network of the second generation for decision support in especially significant situations. *J. Hydrol.* **2023**, *616*, 128844. [\[CrossRef\]](#)
3. Lv, Z.; Zuo, J.; Rodriguez, D. Predicting of Runoff Using an Optimized SWAT-ANN: A Case Study. *J. Hydrol. Reg. Stud.* **2020**, *29*, 100688. [\[CrossRef\]](#)
4. Abou Rjeily, Y.; Abbas, O.; Sadek, M.; Shahrour, I.; Hage Chehade, F. Flood forecasting within urban drainage systems using NARX neural network. *Water Sci. Technol.* **2017**, *76*, 2401–2412. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Le, X.H.; Ho, H.V.; Lee, G.; Jung, S. Application of long short-term memory (LSTM) neural network for flood forecasting. *Water* **2019**, *11*, 1387. [\[CrossRef\]](#)
6. Muñoz, P.; Orellana-Alvear, J.; Bendix, J.; Feyen, J.; Céleri, R. Flood Early Warning Systems Using Machine Learning Techniques: The Case of the Tomebamba Catchment at the Southern Andes of Ecuador. *Hydrology* **2021**, *8*, 183. [\[CrossRef\]](#)
7. Bande, S.; Shete, V.V. Smart flood disaster prediction system using IoT & neural networks. In Proceedings of the 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bengaluru, India, 17–19 August 2017; pp. 189–194.
8. Jabbari, A.; Bae, D.H. Application of Artificial Neural Networks for accuracy enhancements of real-time flood forecasting in the Imjin basin. *Water* **2018**, *10*, 1626. [\[CrossRef\]](#)
9. Tabbussum, R.; Dar, A.Q. Comparative analysis of neural network training algorithms for the flood forecast modelling of an alluvial Himalayan river. *J. Flood Risk Manag.* **2020**, *13*, e12656. [\[CrossRef\]](#)
10. Dtissibe, F.Y.; Ari, A.A.A.; Titouna, C.; Thiare, O.; Gueroui, A.M. Flood forecasting based on an artificial neural network scheme. *Nat. Hazards* **2020**, *104*, 1211–1237. [\[CrossRef\]](#)
11. Abdullahi, S.I.; Habaebi, M.H.; Malik, N.A. Flood disaster warning system on the go. In Proceedings of the 2018 7th International Conference on Computer and Communication Engineering (ICCC), Kuala Lumpur, Malaysia, 19–20 September 2018; pp. 258–263.
12. Kimura, N.; Yoshinaga, I.; Sekijima, K.; Azechi, I.; Baba, D. Convolutional neural network coupled with a transfer-learning approach for time-series flood predictions. *Water* **2019**, *12*, 96. [\[CrossRef\]](#)
13. Moishin, M.; Deo, R.C.; Prasad, R.; Raj, N.; Abdulla, S. Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm. *IEEE Access* **2021**, *9*, 50982–50993. [\[CrossRef\]](#)
14. Khan, U.T.; He, J.; Valeo, C. River flood prediction using fuzzy neural networks: An investigation on automated network architecture. *Water Sci. Technol.* **2018**, *2017*, 238–247. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Tabbussum, R.; Dar, A.Q. Performance evaluation of artificial intelligence paradigms—Artificial neural networks, fuzzy logic, and adaptive neuro-fuzzy inference system for flood prediction. *Environ. Sci. Pollut. Res.* **2021**, *28*, 25265–25282. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Sankaranarayanan, S.; Prabhakar, M.; Satish, S.; Jain, P.; Ramprasad, A.; Krishnan, A. Flood prediction based on weather parameters using deep learning. *J. Water Clim. Chang.* **2020**, *11*, 1766–1783. [\[CrossRef\]](#)
17. Cruz, F.R.G.; Binag, M.G.; Ga, M.R.G.; Uy, F.A.A. Flood prediction using multi-layer artificial neural network in monitoring system with rain gauge, water level, soil moisture sensors. In Proceedings of the TENCON 2018–2018 IEEE Region 10 Conference, Jeju, Republic of Korea, 28–31 October 2018; pp. 2499–2503.
18. Tsakiri, K.; Marsellos, A.; Kapetanakis, S. Artificial neural network and multiple linear regression for flood prediction in Mohawk River, New York. *Water* **2018**, *10*, 1158. [\[CrossRef\]](#)
19. Noymanee, J.; Theeramunkong, T. Flood forecasting with machine learning technique on hydrological modeling. *Procedia Comput. Sci.* **2019**, *156*, 377–386. [\[CrossRef\]](#)
20. Chang, L.C.; Chang, F.J.; Yang, S.N.; Tsai, F.H.; Chang, T.H.; Herricks, E.E. Self-organizing maps of typhoon tracks allow for flood forecasts up to two days in advance. *Nat. Commun.* **2020**, *11*, 1983. [\[CrossRef\]](#) [\[PubMed\]](#)

21. Zhou, Y.; Guo, S.; Xu, C.Y.; Chang, F.J.; Yin, J. Improving the reliability of probabilistic multi-step-ahead flood forecasting by fusing unscented Kalman filter with recurrent neural network. *Water* **2020**, *12*, 578. [[CrossRef](#)]
22. Renteria-Mena, J.B.; Giraldo, E. Multivariable AR Data Assimilation for Water Level, Flow, and Precipitation Data. *IAENG Int. J. Comput. Sci.* **2023**, *50*, 263–273.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.