Received: 26 March 2018 | Accepted: 13 July 2018

DOI: 10.1002/nrm.12189

WILEY Natural Resource Modeling

Neural network modeling of hydrological systems: A review of implementation techniques

Oluwaseun Oyebode | **Derek Stretch**

Department of Civil Engineering, Centre for Research in Environmental, Coastal and Hydrological Engineering (CRECHE), University of KwaZulu-Natal, Durban, South Africa

Correspondence
Oluwaseun Oyebode, Department of Civil Engineering, Centre for Research in Environmental, Coastal and Hydrological Engineering (CRECHE), University of KwaZulu-Natal, Durban, South Africa.
Email: oluwaseun.oyebsode@gmail.com

Abstract

This paper presents a review of procedural steps and implementation techniques used in the development of artificial intelligence models, generally referred to as artificial neural networks (ANNs), within the water resources domain. It focusses on identifying different areas wherein ANNs have found application thereby elucidating its advantages and disadvantages as well as various challenges encountered in its use. Results from this review provide useful insights into how the performance of ANNs can be improved and potential areas of application that are yet to be explored in hydrological modeling.

Recommendations for Resource Managers

- Development of integrated and hybrid artificial intelligent tools is critical to achieving improved forecasts in hydrological modeling studies.
- Further research into comprehending the internal mechanisms of neural networks is required to obtain a practical meaning of each network component deployed to solve real-world problems.
- More robust optimization techniques and tools like differential evolution, particle swarm optimization and deep neural nets, are yet to be fully explored in the water resources analysis, and should be given more attention to enhance neural networks aptitude for modeling complex and nonlinear hydrological processes.

KEYWORDS
artificial neural network (ANN), hydrological modeling, model development, water resources

1 | INTRODUCTION

The inspiration for developing artificial intelligence models (also referred to as artificial neural networks [ANN]) originated from the desire to develop a system that mirrors the working principles of the human brain for making decisions. ANN comprises a collection of processors called neurons, which accept, analyzes, and exchange information over a network of weighted connections. Thus, ANN acquires knowledge via a pragmatic technique that involves identifying a set of weights for the connections and boundary values, known as biases, for the neurons (Sharghi, Nourani, Najafi, & Molajou, 2018). In the same manner as a human brain works, the knowledge gained by ANNs is stored and made available for future use (Elshorbagy, Corzo, Srinivasulu, & Solomatine, 2010).

ANNs can be categorized as single or multilayer based on the number of layers and as feed-forward, recurrent, or self-organizing based on the flow of signal and process configuration (Govindaraju, 2000). The multilayer perceptron (MLP)—a multilayer feed-forward ANN—and radial basis function (RBF) networks are the most widely applied ANNs. The MLP is regarded as the most popular in water resources. So, this paper concentrates on the MLP configuration. Other network configurations that have been applied in water resources include RBF (Kagoda, Ndiritu, Ntuli, & Mwaka, 2010), Bayesian neural network (Zhang, Liang, Srinivasan, & Van Liew, 2009), Kohenen self-organizing feature maps (SOFM; Kalteh, Hjorth, & Berndtsson, 2008), and stacked neural networks (Yonaba, Anctil, & Fortin, 2010).

An MLP ANN comprises an input layer, single or multiple hidden layers, and an output layer. Figure 1 presents a typical structure of a three-layer feed-forward neural network (FFNN). Each neuron is represented by a ring and each connection weight by a line. The configuration of a neuron is also illustrated. Based on a weighted sum of all inputs received and according to a nonlinear function called the transfer or activation function, each neuron returns an output (Kalteh et al., 2008). Thus, in reality, the function of the neuron is to receive input from neighboring or external sources and compute an output signal, which is propagated to other units.

The activation function used in an ANN is dependent on the type, learning algorithm, and scaling approach used. The activation function may be linear, Gaussian, logistic sigmoid, threshold, or hyperbolic tangent sigmoid functions (Bennett, Stewart, & Beal, 2013; Esfe, 2017). The activation functions that are commonly used in hydrological studies are the sigmoid functions such as the hyperbolic tangent and logistic sigmoid functions (Maier & Dandy, 2000), which can be computed using the following mathematical expressions, respectively:

$$\boldsymbol{f(x) = [2/(1 + e^{-2(x)})] - 1} \tag{1}$$

$$\boldsymbol{f(x) = [1/(1 + e^{-(x)})]} \tag{2}$$

where, x is the weighted sum of inputs to the neuron, and $f(x)$ is the output of the neuron.

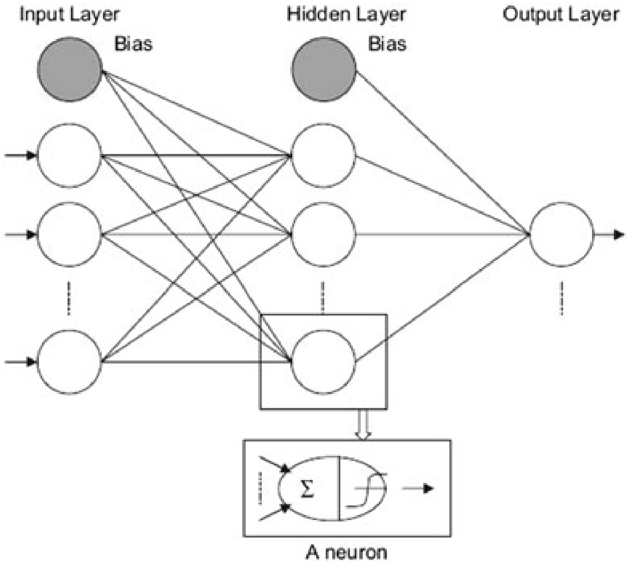

FIGURE 1 Configuration of an FFNN (Kalteh et al., 2008). FFNN: feed-forward neural network

The connection weights are a function of the contribution of each input to the neuron, and the threshold value (bias) must be exceeded for a neuron to be activated. The connection weights and biases are parameters that are typically established by subjecting the network to training using a defined set of data (Arunkumar, Jothiprakash, & Sharma, 2017).

To develop an ANN, there are a series of processes to follow. Model development mainly entails data collection, data splitting, data handling and preprocessing, selection of model inputs, selection of ANN (in terms of architecture, activation function, and learning algorithm), and network training and testing (Abrahart et al., 2012; Maier & Dandy, 2000). However, these procedures are not firmly successive or exclusive, but usually mixed. The objective of this paper is to review the procedural steps and techniques used in the development of ANNs within the water resources domain. It focusses on identifying different areas wherein ANNs have found application thereby elucidating its advantages and disadvantages as well as various challenges encountered from its use, while also providing useful insights into how the performance of ANNs can be improved.

2 | MODEL DEVELOPMENT PROCESSES

2.1 | Data handling and preprocessing

The need for efficient handling and preprocessing of data is of high importance, as variability exists across the different subsets. Thus, data-preprocessing may significantly impact model performance, as it ensures the reduction of noise in data. Maier and Dandy (2000) advocated that all explanatory variables should be standardized to ensure equality in their order of magnitude, and consequently receive equal attention in the training phase. Generally, standardization refers to rescaling of data to fall within the boundaries of the activation function used in the output layer. Minns and Hall (1996) advised that scaling of data must be defined such that the range is proportional to the bounds of the activation function in the output layer; else input variables recorded at different order of magnitudes can influence the training process negatively (Dawson & Wilby, 2001).

Another method used in data-preprocessing is normalization, which entails rescaling the data to a Gaussian function. Maier and Dandy (2000), however, cautioned against the scaling of data to the extreme ranges of the activation function as this reduces the extent of the weight updates, thereby resulting to flat spots during training. Abrahart et al. (2012) advocated for a systematic inquiry into the use of deseasonalized hydrological data, as a series of mixed results show that such inputs may either produce better or worse results than those obtained from the original records with no preprocessing. Other noise-reduction techniques that have been used in water resources applications include wavelet transforms (Kişi, 2009) and single spectrum analysis (Wu & Chau, 2011).

2.2 | Determination of model inputs

One of the initial steps in carrying out any modeling study is to collect data of causal variables that may possibly influence the modeled hydrological system. These sets of data are often large in size, and thus need to be subjected to a screening process. This will ensure the selection of only input variables that will give a good representation of the process (Prasad, Deo, Li, & Maraseni, 2017). Although, data-driven models possess the ability to determine the important model inputs, introducing a large number of inputs to ANNs and trusting the network to decide on the important ones often increases the data size required to effectively compute the connection weights. Furthermore, the addition of unrelated inputs increases the extent of the network, amplifies model complexity, limits knowledge extraction, inhibits the learning process, and thus results in poor generalization (Bowden, Dandy, & Maier, 2005).

Several methods of input selection have been used in hydrological studies. According to Bowden et al. (2005), the main input selection methods can be broadly categorized into five groups namely: (a) methods based on a priori knowledge of the modeled system; (b) correlation analyses; (c) a heuristic approaches; (d) knowledge extraction; and (e) composite methods.

Methods based on a priori knowledge of the modeled process have been widely applied in various water-related modeling studies. The application of a priori methods in determining the most suitable input variables relies on the knowledge base of the modeler concerning the process being modeled. This implies that the effectiveness of the approach is contingent on the expert knowledge, and hence, a subjective- and case-dependent approach (Corominas et al., 2018).

Correlation-based methods are often used for analytical purposes, especially when the system to be modeled is not well-understood. They help provide insight into the relationships that exists between the system processes. Correlation-based methods that have been used for input variable selection include linear correlation analysis and nonlinear correlation analysis. Methods based on linear correlation analysis include cross-correlation (Aqil, Kita, Yano, & Nishiyama, 2007), autocorrelation (Londhe & Charhate, 2010), and partial autocorrelation (Yaseen et al., 2018). Nonlinear correlation analysis is sometimes used due to the inability of linear correlation analysis to capture nonlinear relationships. Nonlinear correlation methods used in hydrological studies include average mutual information (Bhattacharya & Solomatine, 2005), and partial mutual information (Fernando, Maier, & Dandy, 2009).

Heuristic methods involve training several ANN models using a different combination of inputs until the best model is identified (Toth, 2009). Another heuristic approach used in hydrological studies is the stepwise selection approach. The stepwise selection approach serves as an alternative to the input-combination approach, as it eliminates the need for enumerating all input combinations. The stepwise selection approach involves training individual networks for each input variable (Pektas & Cigizoglu, 2017b). Two stepwise approaches usually used include forward and backward selection stepwise approach. The limitation in the use of heuristic approaches is that they are based on trial-and-error, thus there is no assurance of finding the globally best subsets. In addition, the implementation of the stepwise approach is also computationally expensive (Bowden et al., 2005).

The knowledge extraction method, as the name suggests, entails extracting knowledge from trained ANNs. The most widely used method of extracting knowledge from trained ANNs is the sensitivity method (Dibike & Coulibaly, 2006). In the sensitivity method, graphical representations of the responsiveness to the inputs are critically examined, and significant explanatory inputs are selected based on human assessment (Cao et al., 2017). However, the difficulty with this method is knowing a plausible value by which to perturb the input, thereby choosing a suitable threshold for input significance.

Composite methods refer to the combined use of any of the afore-mentioned approaches. Silverman and Dracup (2000) used a combination of a priori knowledge and trial-and-error method to determine ANN model inputs for long-term precipitation forecasts in California, USA. In addition, the combined use of a priori knowledge and other methods have been endorsed as a viable approach for selecting suitable model inputs and their corresponding lags (Yu, Chen, & Chang, 2006). Other methods used in selecting model inputs but not captured in the above categories include principal component analysis, which is often used when the number of inputs is large (Hu, Wu, & Zhang, 2007), and evolutionary optimization methods, such as genetic algorithm (GA), and multiobjective GA to simultaneously optimize the model inputs and network architecture (Giustolisi & Simeone, 2006).

2.3 | Data splitting

In data-driven modeling, the norm is to split a given set of historical observations into training and validation sets. There is a need to ensure that the training and validation data sets exhibit similar statistical properties, as ANNs are incapable to accurately extrapolate beyond the limits of the training data sets. It has been established that when validation data sets fall outside the range of the training data sets, there is a strong tendency to produce poor predictions (Maier & Dandy, 2000). This usually occurs when a limited amount of data set is used.

According to Maier and Dandy (2000), measures that have been applied to maximize the use of available data sets include the use of a "hold-out method," which involves withholding a little portion of the data for validation and training the network on the remaining data sets. Once a significant degree of convergence is observed in the network using the validation set, a different subset of data is withheld and the process is repeated until the network attains significant generalization for the entire data set (Sun & Trevor, 2018). Another method that is being used to maximize available data is the "cross-validation" technique. This method involves using a distinct test set to evaluate the model performance at different learning phases (Modaresi, Araghinejad, & Ebrahimi, 2018). In the cross-validation method, the norm is to split the available data into three: training, testing, and validation sets (Kisi, Mansouri, & Hu, 2017). The training data set is used to train a number of different ANN configurations. The test set is also used to determine when to halt training, so as to prevent overfitting, and to identify the model with optimal network architecture. The validation set is used to evaluate the preferred model against unseen data. The method, however, requires large amounts of data and is not normally applied when the amount of data is limited.

2.4 | Determination of network architecture

Identification of optimal network architecture is one of the most significant and most challenging tasks in ANN model development (Maier & Dandy, 2000). The development of ANN architecture consists of the number of hidden layers, number of neurons in the input layer, hidden layer(s), and output layer. The most popular ANN architecture used in forecasting hydrological variables is the single hidden-layered FFNN because its networks provide adequate complexity to precisely represent the dynamic and nonlinear behavior of hydrological processes (de Vos & Rientjes, 2005). Therefore, the major task in building FFNN architecture is either specifying model inputs and the number of hidden layer neurons, or only the number of hidden layer neurons, given the fact that the number of model inputs has been predetermined. The optimal network architecture generally finds a middle ground between generalization and network complexity (Maier, Jain, Dandy, & Sudheer, 2010).

Methods for determining the optimal network architecture for ANNs can be categorized into three, namely global, stepwise, or ad hoc methods. The global methods comprise methods that uses the principle of evolutionary search to find the optimal network architecture. Examples of such methods include GA, differential evolution (DE), particle swarm optimization (PSO), and simulated annealing. These methods allow for simultaneous optimization of the network parameters (network weights and biases) and the model structure (number of hidden layer neurons). Maier et al. (2010) opined that if the global methods are used appropriately, they are likely to produce the best model structure and/or parameters. It is, however, important to note that the application of these global methods maybe computationally expensive.

An alternative to the use of global methods is to employ a stepwise trial-and-error approach, which starts with an assumed basic ANN structure and thereafter adjusts after each trial with the aim of obtaining a structure that has optimal complexity. The stepwise methods can be further subdivided into two categories, one based on pruning algorithms and the other on constructive algorithms (Sharma & Chandra, 2010). The pruning algorithm initializes with a suitably complex network structure and sequentially removes network parameters and the associated neurons one after the other until model performance deteriorates significantly. Constructive algorithms, however, start-off with the simplest model structure and sequentially adds hidden neurons in succession until no substantial increase in model performance is noticed. The use of pruning and constructive algorithms, thus require high computational resources, as different model structures undergo training and evaluation before obtaining an optimal structure (Maier et al., 2010).

The selection of an appropriate model structure and an optimal number of hidden layer neurons via pure trial-and-error, without a priori knowledge or other approaches other than the two methods discussed earlier, can be categorized as ad hoc. Maier et al's. (2010) study found that the ad hoc methods were by far the most popular, followed by the constructive and stepwise approaches. Conversely, it was observed that the potential of global optimization methods has not been fully exploited. This implies that global optimization methods need to be further integrated into ANN model development to ensure that more robust ANNs are developed.

2.5 | Network training

The objective of ANN training is to obtain a set of optimal model parameters (connection weights) that allows a model to give the best representation of the input–output relationship of a particular system. This usually involves error minimization between the network and target outputs to achieve optimal generalization of the model (Pektas & Cigizoglu, 2017a). However, finding the model parameters or combination of model parameters that will produce the smallest error estimates is not a simple task. The degree of difficulty in searching for the optimal set of model parameters is dependent on the "ruggedness" or "smoothness" of the error surface (Maier et al., 2010).

Owing to the challenges related to ANN training, optimization algorithms are typically used for training ANNs. Training algorithms can be classified into either local or global optimization approaches. Local approaches are generally established on the gradient (first-order and second-order) principle. The most widely applied local-based training algorithm is the back-propagation (BP) algorithm (Rumelhart, Hinton, & Williams, 1985), instituted on the first-order gradient of the slope of the objective function. Second-order gradient techniques include the Levenberg–Marquardt (LM) and conjugate gradient (CG) algorithms. The local approaches are often computationally proficient algorithms and have been implemented in several studies (Elshorbagy et al., 2010; Toth, 2009; Wang, Chau, Cheng, & Qiu, 2009). However, there is a tendency for local approaches to get confined in local optima and a high likelihood of producing negative figures, especially if the error surface is not smooth (Maier et al., 2010).

Global optimization approaches like PSO, DE, GA, and evolutionary programming are now being tested in ANN modeling of water resources. These global optimization approaches are being applied to achieve optimal model parameterization (Adeyemo, Oyebode, & Stretch, 2018; Oyebode & Adeyemo, 2014) or coupled with the gradient based training techniques like BP, LM, CG algorithms (Kişi, 2010; Subudhi & Jena, 2011) to optimize the ANN learning process. Findings from the literature show that, unlike local optimization approaches that are susceptible to being trapped in local optima, global optimization approaches exhibit greater aptitude for overcoming local optima, thereby obtaining more stable solutions (Ding, Li, Su, Yu, & Jin, 2013). They, however, use greater computational resources.

Stochastic-based approaches such as Bayesian methods have also been used to achieve the distribution of model parameters, rather than searching for single parameter values (Rasouli, Hsieh, & Cannon, 2012). The major benefit of these methods is that prediction intervals can be automatically obtained for the model predictions.

3 | APPLICATIONS OF ANNs IN HYDROLOGICAL MODELING

Over the past two decades, the use of ANNs for the forecasting of water resources variables has been firmly established. ANNs have been applied in diverse fields of water resources, such as rainfall estimation, rainfall-runoff modeling, streamflow modeling, groundwater modeling, sediment transport modeling, water quality, hydrological time series, precipitation forecasting, climate change impact modeling, reservoir operations, ecomodeling, and uncertainty estimation (Abrahart et al., 2012).

With specific reference to streamflow prediction, numerous researchers have adopted the cause-effect form of modeling to predict streamflow. Historical observations of causative variables, such as rainfall, temperature, soil moisture content, evaporation, and runoff, have been used to predict streamflow either individually or in combinatory form (Baratti et al., 2003). However, in some studies, only previous values of streamflow were used for streamflow prediction, as in univariate time series modeling (Londhe & Charhate, 2010).

Some studies have focused on performance comparison between ANN and traditional statistical methods such as autoregressive and autoregressive-moving-average methods (Jain & Kumar, 2007). Generally, the majority of studies have proven ANNs to have better predictive performance than the traditional statistical methods. In addition, the dominance of ANN over linear regression can be ascribed to the intrinsic nonlinearity in hydrological processes, which are inadequately represented by the regression methods.

Some researchers have used modular or hybrid procedures for streamflow prediction. Abrahart and See (2002) proposed a hybrid model based on the clustering of historical observations with an SOFM neural network. Chang and Chen (2001) developed a hybrid neural network by the fusion of a neural network and fuzzy arithmetic, which was successful in streamflow prediction. Sivakumar, Jayawardena, and Fernando (2002) combined an ANN with a phase-space reconstruction method for river flow prediction. Wang, Van Gelder, Vrijling, and Ma (2006) also developed a hybrid model to forecast daily streamflow in the Yellow River, China. Jain and Kumar (2007) used a combination of ANNs and traditional time series techniques for monthly streamflow prediction.

ANNs have also been coupled with other data-driven models (DDMs), such as fuzzy rule-based systems (FRBS), genetic programming, Bayesian methods, etc. (Adeyemo et al., 2018; Khan & Coulibaly, 2006). Although the use of modular ANNs ensures higher forecast accuracy by splitting the various hydrological processes among individual models (Solomatine & Ostfeld, 2008), the synergistic use of ANNs and other DDMs improves forecast accuracy by making-up for the limitations that might be inherent in any of the techniques.

Finally, ANNs have also been deployed as complementary tools in water-related studies, such as for error updates in process-based and other DDMs (Toth, 2009), for infilling of missing hydrological records (Coulibaly & Evora, 2007) and for estimating prediction intervals of model outputs (Shrestha, Kayastha, & Solomatine, 2009).

4 | AREAS OF CONCERN

The major areas of concern in the use of ANNs relates to the methodologies used in building the model network (Figure 2). Most of these issues have been discussed one way or the other while enumerating the model development processes in previous sections. However, the synopsis of the major areas of concern in the application of ANNs according to Abrahart et al. (2012) include (a) the fact that ANNs are “black-box” models raises questions on their ability to reveal the physics of underlying processes, as the final product and the development process are not fully transparent. The need for knowledge-based networks and rule-extraction techniques aimed at opening the box were, therefore, suggested; (b) identification of the optimal training set is challenging as no defined procedure exist for selecting or splitting data sets for model development purposes; (c) in terms of ANN model structures, the use of simple models may produce diminished performance. Conversely, complex models may also produce insufficient generalization and its parameters could be more difficult to calibrate or interpret. Thus, striking a balance between costs and benefits still remains a source of concern; (d) the extrapolation ability of ANNs is also a subject of concern. Although, ANNs are seen as good interpolators,

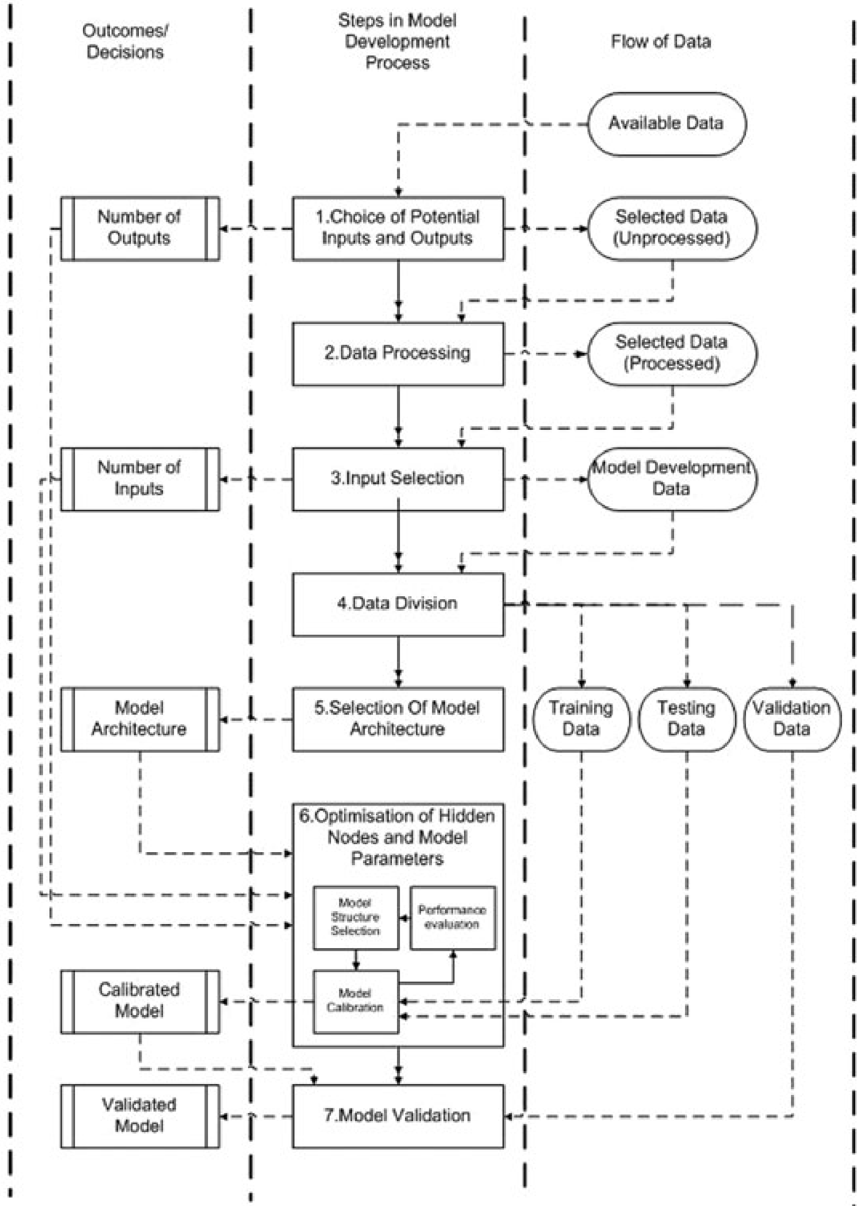

FIGURE 2 Steps in ANN development process (Maier et al., 2010). ANN: artificial neural network

however, predictions made beyond the range of the training set (i.e., extrapolation) tend to generate poor and unreliable predictions. To give room for extrapolation, methods such as scaling of activation functions have been used, but at the expense of introducing an extra, initially arbitrary parameter.

Other areas of concern in the application of ANNs reported in the literature include: (a) inability to accurately predict extreme events unless the training data sets comprise similar extreme events (Sayagavi & Charhate, 2017); (b) failure to provide explicit information regarding prediction uncertainty (Kasiviswanathan & Sudheer, 2017); and (c) disability in vague data processing (Adib & Mahmoodi, 2017).

5 | PERFORMANCE IMPROVEMENT METHODS

A number of methods may be adopted for improving the performance of ANNs. These performance improvements range from investigation of different ANN algorithms and architecture to the development of more robust hybrid models. Abrahart, Kneale, and See (2004) suggested five research directions regarding the improvement of ANN to make meaningful progress in their applications to water resources. The five directions include (a) improving existing ANN models by investigating existing ANN hydrological problems; (b) focusing on comparative analysis of ANNs with operational system-based models; (c) developing more robust modeling assessment standards; (d) further research into comprehending ANN internal mechanisms and obtaining a practical meaning for each network component; and (e) development of integrated and hybrid ANN tools for water resources studies.

Research has shown that, over the last decade, efforts have been made toward implementing the suggested research directions. Generally, these efforts have focused primarily on the development and application of new (single or hybridized) optimization techniques to address overparameterization, overfitting, model complexity, and knowledge extraction problems (Babel & Shinde, 2011; Ding et al., 2013; Londhe & Narkhede, 2017; Londhe & Shah, 2017). However, a provisional search of scholarly papers and databases show that some techniques or tools are yet to be fully explored in ANN modeling of water resources systems. These techniques include evolutionary computation techniques and other metaheuristics like DE, PSO, and deep learning. These techniques have, however, been successfully applied to solving problems in other fields of science and engineering (Abdul-Kader, 2009; Mason, Duggan, & Howley, 2018; Qian, Wang, Hu, Huang, & Wang, 2009). There is, therefore, a need for these techniques to be given more attention in the development of intelligent models for water resources analysis to address the areas of concerns highlighted in this paper. This will assist in fostering the realization of the full potential of modeling complex hydrological processes.

6 | ADVANTAGES AND DISADVANTAGES

Various advantages can be derived from the use of ANNs in water resources. These advantages include (a) aptitude for modeling complex and nonlinear hydrological processes without any assumption for the relationships between input and output variables (Oyebode & Adeyemo, 2014); (b) self-adjusting ability to a given set of data (Sharghi et al., 2018); (c) requires limited knowledge of a given problem to achieve positive results (Bennett et al., 2013); (d) showcase compactness and flexibility in model structure, and thus allow for easy integration of ANNs into other DDMs (Oyebode & Adeyemo, 2014); and (e) relatively low demand for computation resources when compared with some other modeling techniques (Abrahart et al., 2012).

However, a key restraining factor to the use of ANNs is their failure to produce explicit models, as their core operations are unclear and not interpretable. Another major drawback is that the optimal network configuration for each modeling circumstance can differ, as the challenge of "no theoretical support" is evident in the model development stages, with experimentation forming the basis of most of the fundamental aspects (Abrahart et al., 2004).

Furthermore, no standard or fixed rules exist for governing appropriate model design and development, thus making it impossible to establish a suitable model a priori (Abrahart et al., 2012). The inability to incorporate knowledge acquired from existing physical laws into ANNs also serves as a drawback to their use. Finally and most importantly, ANNs are susceptible to overparameterization and overfitting problems, especially when techniques based on optimal input selection and early stopping are not implemented (Adeyemo et al., 2018; Sayagavi & Charhate, 2017).

7 | CONCLUSIONS

This paper has reviewed the application of ANNs to water resources modeling with a specific focus on the different techniques used in model development. Although the application of ANN is an established practice in water resources analysis, some areas of concern have been highlighted in this review. These areas include challenges in determining optimal input parameters and training data sets, model complexity issues, and generalization problems. This review has also highlighted strategies and methods that can be used for improving the performance of ANNs. This includes the development and application of more robust optimization tools like DE, PSO, and deep neural nets, which are yet to be fully explored in the water resources domain. These techniques could be used for parameter selection, model training, architecture selection, and prediction improvement purposes, thereby ensuring that the potential of ANN in water resources is fully used.

ORCID

Oluwaseun Oyebode http://orcid.org/0000-0001-8876-9743

REFERENCES

Abdul-Kader, H. (2009). Neural networks training based on differential evolution al-gorithm compared with other architectures for weather forecasting34. *International Journal of Computer Science and Network Security*, *9*(3), 92–99.

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., ... Wilby, R. L. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography*, *36*(4), 480–513.

Abrahart, R. J., Kneale, P. E., & See, L. M. (2004). Towards a hydrological research agenda. *Neural Networks for Hydrological Modeling*, *289*.

Abrahart, R. J., & See, L. (2002). Multi-model data fusion for river flow forecasting: An evaluation of six alternative methods based on two contrasting catchments. *Hydrology and Earth System Sciences Discussions*, *6*(4), 655–670.

Adeyemo, J., Oyebode, O., & Stretch, D. (2018). River flow forecasting using an improved artificial neural network, *EVOLVE-A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation VI* (pp. 179–193). Cham, Switzerland: Springer.

Adib, A., & Mahmoodi, A. (2017). Prediction of suspended sediment load using ANN GA conjunction model with Markov chain approach at flood conditions. *KSCE Journal of Civil Engineering*, *21*(1), 447–457.

Aqil, M., Kita, I., Yano, A., & Nishiyama, S. (2007). A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *Journal of Hydrology*, *337*(1), 22–34.

Arunkumar, R., Jothiprakash, V., & Sharma, K. (2017). Artificial intelligence techniques for predicting and mapping daily pan evaporation. *Journal of the Institution of Engineers (India): Series A*, *98*(3), 219–231.

Babel, M. S., & Shinde, V. R. (2011). Identifying prominent explanatory variables for water demand prediction using artificial neural networks: A case study of Bangkok. *Water Resources Management*, *25*(6), 1653–1676.

Baratti, R., Cannas, B., Fanni, A., Pintus, M., Sechi, G. M., & Toreno, N. (2003). River flow forecast for reservoir management through neural networks. *Neurocomputing*, *55*(3), 421–437.

Bennett, C., Stewart, R. A., & Beal, C. D. (2013). ANN-based residential water end-use demand forecasting model. *Expert Systems with Applications*, *40*(4), 1014–1023.

Bhattacharya, B., & Solomatine, D. P. (2005). Neural networks and M5 model trees in modelling water level–discharge relationship. *Neurocomputing*, *63*, 381–396.

Bowden, G. J., Dandy, G. C., & Maier, H. R. (2005). Input determination for neural network models in water resources applications. Part 1—background and methodology. *Journal of Hydrology*, *301*(1), 75–92.

Cao, M. S., Pan, L. X., Gao, Y. F., Novák, D., Ding, Z. C., Lehký, D., & Li, X. L. (2017). Neural network ensemble-based parameter sensitivity analysis in civil engineering systems. *Neural Computing and Applications*, *28*(7), 1583–1590.

Chang, F.-J., & Chen, Y.-C. (2001). A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. *Journal of Hydrology*, *245*(1), 153–164.

Corominas, L., Garrido-Baserba, M., Villez, K., Olsson, G., Cortés, U., & Poch, M. (2018). Transforming data into knowledge for improved wastewater treatment operation: A critical review of techniques. *Environmental Modelling & Software*, *106*(8), 89–103.

Coulibaly, P., & Evora, N. D. (2007). Comparison of neural network methods for infilling missing daily weather records. *Journal of Hydrology*, *341*(1), 27–41.

Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, *25*(1), 80–108.

Dibike, Y. B., & Coulibaly, P. (2006). Temporal neural networks for downscaling climate variability and extremes. *Neural Networks*, *19*(2), 135–144.

Ding, S., Li, H., Su, C., Yu, J., & Jin, F. (2013). Evolutionary artificial neural networks: A review. *Artificial Intelligence Review*, *39*(3), 251–260.

Elshorbagy, A., Corzo, G., Srinivasulu, S., & Solomatine, D. P. (2010). Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology—part 1: Concepts and methodology. *Hydrology and Earth System Sciences*, *14*(10), 1931–1941.

Esfe, M. H. (2017). Designing a neural network for predicting the heat transfer and pressure drop characteristics of Ag/water nanofluids in a heat exchanger. *Applied Thermal Engineering*, *126*, 559–565.

Fernando, T. M. K. G., Maier, H. R., & Dandy, G. C. (2009). Selection of input variables for data driven models: An average shifted histogram partial mutual information estimator approach. *Journal of Hydrology*, *367*(3), 165–176.

Giustolisi, O., & Simeone, V. (2006). Optimal design of artificial neural networks by a multi-objective strategy: Groundwater level predictions. *Hydrological Sciences Journal*, *51*(3), 502–523.

Govindaraju, R. S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of hydrologic Engineering*, *5*(2), 115–123.

Hu, T., Wu, F., & Zhang, X. (2007). Rainfall-runoff modeling using principal component analysis and neural network. *Nordic Hydrology*, *38*(3), 235–248.

Jain, A., & Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, *7*(2), 585–592.

Kagoda, P. A., Ndiritu, J., Ntuli, C., & Mwaka, B. (2010). Application of radial basis function neural networks to short-term streamflow forecasting. *Physics and Chemistry of the Earth, Parts A/B/C*, *35*(13), 571–581.

Kalteh, A., Hjorth, P., & Berndtsson, R. (2008). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling & Software*, *23*(7), 835–845.

Kasiviswanathan, K. S., & Sudheer, K. P. (2017). Methods used for quantifying the prediction uncertainty of artificial neural network based hydrologic models. *Stochastic Environmental Research and Risk Assessment*, *31*(7), 1659–1670.

Khan, M. S., & Coulibaly, P. (2006). Bayesian neural network for rainfall-runoff modeling. *Water Resources Research*, *42*(7).

Kisi, O., Mansouri, I., & Hu, J. W. (2017). A new method for evaporation modeling: Dynamic evolving neural-fuzzy inference system. *Advances in Meteorology*, *2017*, 1–9.

Kişi, Ö. (2009). Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *Journal of Hydrologic Engineering*, *14*(8), 773–782.

Kişi, Ö. (2010). River suspended sediment concentration modeling using a neural differential evolution approach. *Journal of Hydrology*, *389*(1-2), 227–235.

Londhe, S., & Charhate, S. (2010). Comparison of data-driven modelling techniques for river flow forecasting. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, *55*(7), 1163–1174.

Londhe, S. N., & Narkhede, S. (2017). Forecasting stream flow using hybrid neuro-wavelet technique. *ISH Journal of Hydraulic Engineering*, *24*, 1–10.

Londhe, S. N., & Shah, S. (2017). A novel approach for knowledge extraction from artificial neural networks. *ISH Journal of Hydraulic Engineering*, 1–13. https://doi.org/10.1080/90715010.2017.1409667

Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environmental Modelling & Software*, *15*(1), 101–124.

Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software*, *25*(8), 891–909.

Mason, K., Duggan, J., & Howley, E. (2018). A multi-objective neural network trained with differential evolution for dynamic economic emission dispatch. *International Journal of Electrical Power & Energy Systems*, *100*, 201–221.

Minns, A. W., & Hall, M. J. (1996). Artificial neural networks as rainfall-runoff models. *Hydrological Sciences Journal*, *41*(3), 399–417.

Modaresi, F., Araghinejad, S., & Ebrahimi, K. (2018). A comparative assessment of artificial neural network, generalized regression neural network, least-square support vector regression, and K-nearest neighbor regression for monthly streamflow forecasting in linear and nonlinear conditions. *Water Resources Management*, *32*(1), 243–258.

Oyebode, O., & Adeyemo, J. (2014). Reservoir inflow forecasting using differential evolution trained neural networks, *EVOLVE-A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V* (pp. 307–319). Cham, Switzerland: Springer.

Pektas, A. O., & Cigizoglu, H. K. (2017a). Investigating the extrapolation performance of neural network models in suspended sediment data. *Hydrological Sciences Journal*, *62*(10), 1694–1703.

Pektas, A. O., & Cigizoglu, H. K. (2017b). Long-range forecasting of suspended sediment. *Hydrological Sciences Journal*, *62*(14), 2415–2425.

Prasad, R., Deo, R. C., Li, Y., & Maraseni, T. (2017). Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. *Atmospheric Research*, *197*, 42–63.

Qian, B., Wang, L., Hu, R., Huang, D., & Wang, X. (2009). A DE-based approach to no-wait flow-shop scheduling. *Computers & Industrial Engineering*, *57*(3), 787–805.

Rasouli, K., Hsieh, W. W., & Cannon, A. J. (2012). Daily streamflow forecasting by machine learning methods with weather and climate inputs. *Journal of Hydrology*, *414-415*, 284–293. https://doi.org/10.1016/j.jhydrol.2011.10.039.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation*. San Diego, CA: Institute for Cognitive Science, University of California.

Sayagavi, V., & Charhate, S. (2017). Applications of soft tools to solve hydrological problems for an integrated Indian catchment. *International Journal of Water Resources and Environmental Engineering*, *9*(7), 150–161.

Sharghi, E., Nourani, V., Najafi, H., & Molajou, A. (2018). Emotional ANN (EANN) and wavelet-ANN (WANN) approaches for markovian and seasonal based modeling of rainfall-runoff process. *Water Resources Management*, *32*, 1–16.

Sharma, S. K., & Chandra, P. (2010). Constructive neural networks: A review. *International Journal of Engineering Science and Technology*, *2*(12), 7847–7855.

Shrestha, D. L., Kayastha, N., & Solomatine, D. P. (2009). A novel approach to parameter uncertainty analysis of hydrological models using neural networks. *Hydrology and Earth System Sciences*, *13*(7), 1235–1248.

Silverman, D., & Dracup, J. A. (2000). Artificial neural networks and long-range precipitation prediction in California. *Journal of Applied Meteorology*, *39*(1), 57–66.

Sivakumar, B., Jayawardena, A. W., & Fernando, T. M. K. G. (2002). River flow forecasting: Use of phase-space reconstruction and artificial neural networks approaches. *Journal of Hydrology*, *265*(1), 225–245.

Solomatine, D. P., & Ostfeld, A. (2008). Data-driven modelling: Some past experiences and new approaches. *Journal of Hydroinformatics*, *10*(1), 3–22.

Subudhi, B., & Jena, D. (2011). Nonlinear system identification using memetic differential evolution trained neural networks. *Neurocomputing*, *74*(10), 1696–1709.

Sun, W., & Trevor, B. (2018). Multiple model combination methods for annual maximum water level prediction during river ice breakup. *Hydrological Processes*, *32*(3), 421–435.

Toth, E. (2009). Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting. *Hydrology and Earth System Sciences*, *13*(9), 1555–1566.

de Vos, N. J., & Rientjes, T. H. M. (2005). Constraints of artificial neural networks for rainfall-runoff modelling: Trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences Discussions*, *2*(1), 365–415.

Wang, W., Gelder, P. H. A. J. M. V., Vrijling, J. K., & Ma, J. (2006). Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology*, *324*(1), 383–399.

Wang, W.-C., Chau, K.-W., Cheng, C.-T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, *374*(3), 294–306.

Wu, C. L., & Chau, K. W. (2011). Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology*, *399*(3-4), 394–409.

Yaseen, Z. M., Fu, M., Wang, C., Mohtar, W. H. M. W., Deo, R. C., & El-Shafie, A. (2018). Application of the hybrid artificial neural network coupled with rolling mechanism and grey model algorithms for streamflow forecasting over multiple time horizons. *Water Resources Management*, 1–17.

Yonaba, H., Anctil, F., & Fortin, V. (2010). Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. *Journal of Hydrologic Engineering*, *15*(4), 275–283.

Yu, P.-S., Chen, S.-T., & Chang, I.-F. (2006). Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, *328*(3), 704–716.

Zhang, X., Liang, F., Srinivasan, R., & Van Liew, M. (2009). Estimating uncertainty of streamflow simulation using Bayesian neural networks. *Water Resources Research*, *45*(2).

How to cite this article: Oyebode O, Stretch D. Neural network modeling of hydrological systems: A review of implementation techniques. *Natural Resource Modeling*. 2019;32:e12189. https://doi.org/10.1002/nrm.12189