Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 5, October 2024, pp. 3686~3694
ISSN: 2302-9285, DOI: 10.11591/eei.v13i5.6863

Ensemble neural networks with input optimization for flood forecasting

Nazli Mohd Khairudin, Norwati Mustapha, Teh Noranis Mohd Aris, Maslina Zolkepli
Faculty of Computer Sciences and Information Technology, Universiti Putra Malaysia, Selangor, Malaysia

Article Info

Article history:

Received Jun 2, 2023
Revised Mar 23, 2024
Accepted Mar 28, 2024

Keywords:

Discrete wavelet transform
Ensemble neural networks
Flood forecasting
Generalization
Mutual information

ABSTRACT

Machine learning model has been widely used to provide flood forecasting including the ensemble model. This paper proposed an ensemble of neural networks for long-term flood forecasting that combine the output of backpropagation neural network (BPNN) and extreme learning machine (ELM). The proposed ensemble neural networks model has been applied towards the rainfall data from eight rainfall stations of Kelantan River Basin to forecast the water level of Kuala Krai. The aim is to highlight the improvement on accuracy of the forecast. Prior to the development of such ensemble model, data are optimized in two steps which are decomposed it using discrete wavelet transform (DWT) to reduce variations in the rainfall series and selecting dominant features using entropy called mutual information (MI) for the model. The result of the experiments indicates that ensemble neural networks model based on the data decomposition and entropy feature selection has outperformed individually executed forecast model in term of RMSE, MSE and NSE. This study proved that the proposed method has reduce the data variance and provide better forecasting with minimal error. With minimal forecast error the generalization of the model is improved.

Corresponding Author:

Nazli Mohd Khairudin
Faculty of Computer Sciences and Information Technology, Universiti Putra Malaysia
Serdang, Selangor, Malaysia
Email: nazmkhair@gmail.com

1. INTRODUCTION

Machine learning model has been widely adopted by researchers in developing flood forecast. Recently, many non-linear machine learning models has been developed to forecast flood in long-term manner such as support vector machine (SVM) [1], support vector regression (SVR) [2], [3], artificial neural network (ANN) [4], [5], extreme learning machine (ELM) [6], random forest (RF) [7], and neuro fuzzy [8]. These machine learning model has demonstrated the use of single model in flood forecasting, but it can provide bias and keeping the variance high within the network. This may impact the generalization ability of the model. Another adaptation of machine learning model can be found as an ensemble model. Ensemble model usually done by the combination of various model such as by bagging or stacking technique [9]. Ensemble model has been deemed to provide better performance of forecasting and achieving higher accuracy [10], [11].

Neural network model such as ANN has been widely adopted by researchers. ANN is first developed by [12]. It is a system that simulate how human brains analyze information. ANN works as if they are interconnected neurons in brains that process information and works self-adaptively by producing better results for each of the learning cycle. The complex relationship of the data is defined by its input and output [13] and can produced various patterns. ANN has proven its potential to producing an accurate flood forecast using various predictors in recent studies, but it has some drawbacks in which ANN suffers from local minimum problem and slow convergence [14].

Noticing the drawback of slow convergence in ANN, ELM was introduce by [15] that able to speed up the learning algorithm. ELM is single-layer feedforward neural networks to discover hidden neurons at random and learn the pattern encoded in input variables [16]. Unlike ANN, gradient-based backpropagation isn't required for ELM to work. It sets its weights by using Moore-Penrose generalized inverse. Regardless of these advantages and the faster convergence of the random hidden nodes ELM also tend to produce performance instability [17].

Both ANN and ELM are non-linear method that can flex themselves towards the intricate underlying structure of the data. This high flexibility allows these models to learn in a stochastic way in which each time they are trained, they may find a different set of weights resulting various forecast. With various forecast results produce, they keep the variance of the model high, thus making the final forecast a very challenging task.

When utilising machine learning, specifically neural network models, to forecast long-term floods, it has been demonstrated that incorporating a pre-processing strategy improves the accuracy of the forecasting model [18], [19]. Utilising historical data in a machine learning model for forecasting can have an impact on the model’s behaviour [20], [21]. A recently popular pre-processing method is the discrete wavelet transform (DWT), which involves decomposing data into shifted wavelets. This method offers an approach for analysing time-series variance, providing insights into both the time and frequency domains of the signal [22]. The DWT is more efficient than the continuous wavelet transforms, which involves calculating the wavelet coefficient at each conceivable scale and requires more computational resources [23].

Decomposing the hydrological times series data especially in a single dimension manner has been a great way to provide input for a forecast, but it will only optimize specific input. In the event when datasets are consists of different features, decomposition alone doesn’t reveal the relationship between various data, and it is a challenging task to determine which of these features in the dataset can provide a better forecast. One way to measure the relationship of data is by using entropy. Entropy has proven its ability in hydrological forecasting as it been used in various studies such as to construct flood mapping [24], and forecasting the long-term streamflow [25].

In summary, to have an accurate long-term flood forecasting: i) input must be optimized as it will affect the behavior of the model and ii) forecast model should be able to maintain low variance and minimize the generalization error. The aim for this paper is to fulfill these two goals by proposing long term flood forecasting model based on decomposition-entropy-ensemble machine learning framework. This model is intended to: i) decomposing and optimizing the original input to provide cleaner input and dominant feature for the forecasting model and ii) to improve the flood forecasting performance by reducing the generalization error of the model by combining the outputs of the neural networks model.

2. METHOD

2.1. Discrete wavelet transform with entropy

The use of wavelet decomposition, specifically the DWT, involves the breakdown of time series data into a wavelet representation that has been shifted and scaled. The determination of the decomposition level of DWT is very subjective and lacks a straightforward approach [26]. The decomposition level is determined using the empirical (1) as described in reference [27].

$$L = int[\log(N)] \tag{1}$$

where L is the level of decomposition, N is the length of time-series data, and *int []* is the integer-part function. Daubechies with three vanishing moments (db3) is used as the mother wavelet. It provide high vanishing moments for any given support width [28].

While DWT decomposed each of the features in a multi-dimensional dataset, it is become a challenging of which of these features that are dominant as inputs to provide higher accuracy of the forecast model. The selection of features can be done by evaluating the features using entropy called mutual information (MI) and derived the most dominant features from the MI’s rank. Figure 1 present the DWT with MI used in this study. To gain the most superior input that could give better forecast, dataset can be analyze in term of their correlation between the information dispersion with the variable. For this purpose, MI is well-suited for synchronizing non-linear time-series data and handling a non-linear relationship between input and output [29]. Consider X and Y as the two random variables. The amount of information about X that is contained in Y (and vice versa) is measured as MI. Both linear and non-linear interactions between the variables are captured by it. MI is calculated by (2) [29]:

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) \quad (2)$$

where H(X) and H(Y) are the entropy of X and Y, and the joint entropy of H (X, Y) would be as in (3):

$$H(X, Y) = -\sum_{x\in X}\sum_{y\in Y} p_{XY}(x, y) \log p_{XY}(x, y) \quad (3)$$

where *x* and *y* are the specific value of *X* and *Y*, and *p(x,y)* is the joint probability of these values occurring together. MI has ranked each feature in every dataset according to the MI evaluated score. MI has a value between 0 and positive infinity. An increased value denotes a stronger correlation between the variables, whereby understanding the value of one reveals more details about the other. Figure 1 present the DWT with MI used in this study.

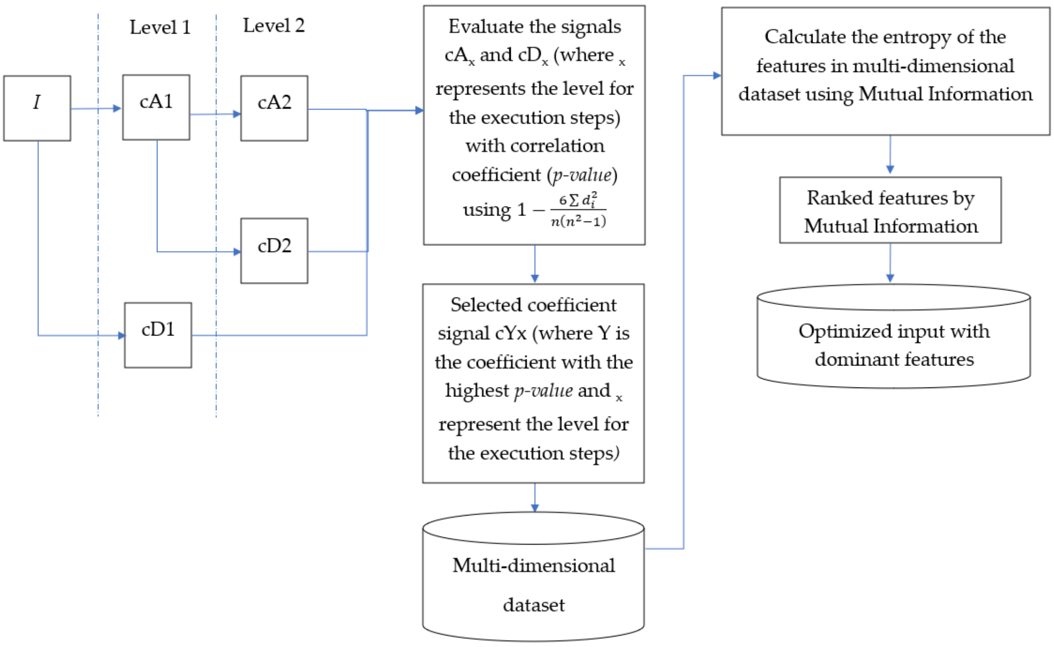

Figure 1. Decomposition of input using DWT and features ranked by MI

2.2. Ensemble neural networks based on DWT decomposition and MI

ANN is a machine learning model that can be utilized to make a forecast through pattern learning. ANN can be a feedforward process that involve the signals been process in one flow direction without any cycle from the input layer to the output layer. Another type is backpropagation neural network (BPNN) in which weight of the networks is adjusted through the error rate gain from previous iteration. The reason behind the weight adjustment is to reduce the error rate to increase generalization. The network relationship of the BPNN in this study can be present as in (4):

$$Y = f_a\left[\sum_j W_{mk} . f_s\left(\sum_i W_{ki} x_i + b_k\right) + b_o\right] \quad (4)$$

where *x* is the vector of input, W_{ki} is the weight of the connection between *i*th node in input layer and *k*th node in hidden layer, b_k is the bias of *k*th hidden neuron, W_{mk} is the weight of the connection between *k*th node in hidden layer and *m*th node in the output layer, b_o is the bias of *k*th output neuron, and activation functions is represented as f_a and f_s. During the backpropagation, the signal is sent in reverse manner. BPNN works by tuning the neural network weight and biases to minimize error.

ELM is based on the feedforward neural networks that uses Moore-Penrose generalized inverse to adjust the networks weights [16]. ELM is run without iterative tuning and the learning parameters are assign independently and automatically. In hydrological forecasting, ELM has been found to provide flood forecasting using various parameters such as watershed [30], groundwater [5], and stream flow [31]. ELM can be formulated in mathematical (5) as [15]:

$$Y_j = \sum_{i=1}^{L} w_i g\left(W_{in(i)}, b_i, x_j\right) = \sum_{i=1}^{L} w_i g\left(W_{in(i)} . x_j + b_i\right) \quad j=1 \ldots\ldots\ldots N \quad (5)$$

where x_j is the vector of input, $W_{in(i)}$ is the weight vector for input, $W_{in(i)}.x_j$ is the inner product of $W_{in(i)}$ and x_j, b_i is the bias of the ith hidden node, $g(\cdot)$ is the approximation function (sigmoid), w_i is the weight matrix of the output and Y_j is the output of the ELM. In initial stage of ELM, the weight and bias are randomly assigned.

BPNN and ELM are powerful tool and widely used by researcher in flood forecasting, but due to their high flexibility towards the training data, it may produce high variance in the forecast and reduce the generalization. The generalization of these neural networks can be elevated by proper ensemble model. Forecasted results produce by individual model may varied according to their datapoints, thus combining the results of those individuals model can indemnify the error [32], [33]. The ensemble neural networks model and independent models are executed with the optimized datasets gain from the pre-processing phased using DWT decomposition and MI. Feature scaling using standardization is used to make the features in common scales. In ensemble model, the output of each independently executed neural networks then been averaged to gain better forecast. Combining the results of different neural network model tends to add bias that balance out the variance of the independent model. It is less dependent towards the input data and the selected training scheme. The proposed architecture of the ensemble neural networks model based on the input decomposition of DWT and features selection by MI is present in Figure 2.

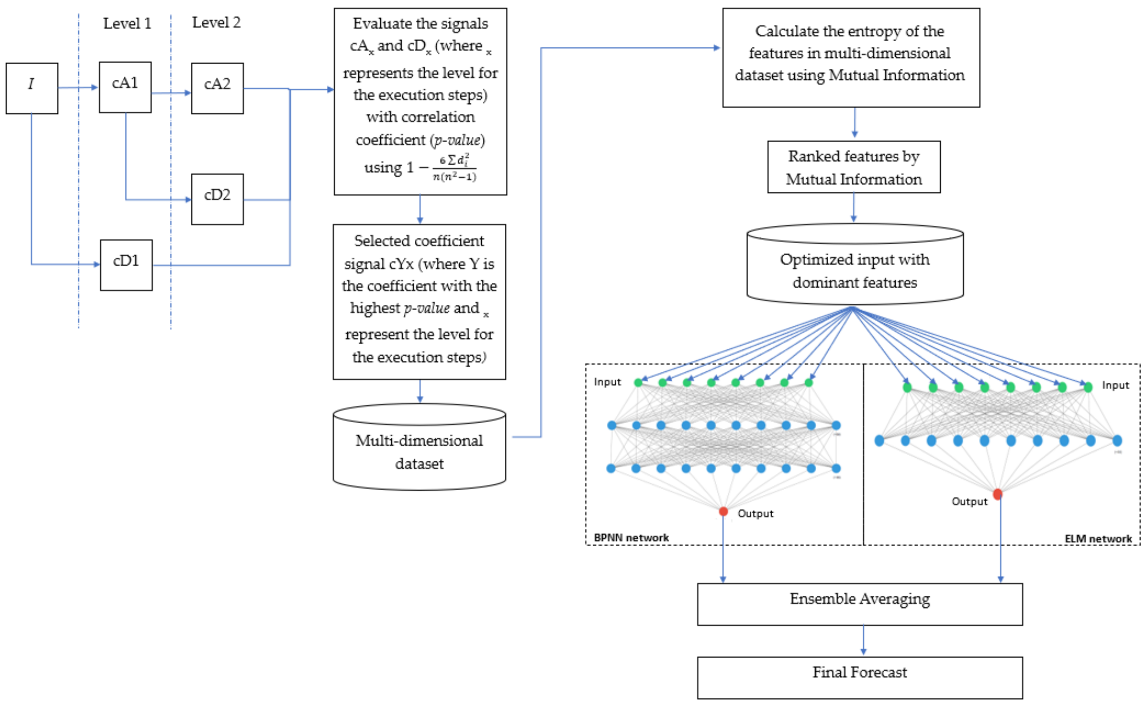

Figure 2. Ensemble neural networks architecture

By combining the predictions from various distinct classifiers tends to produce better prediction and reduce the generalization error. The ensemble averaging is taking the idea of arithmetic mean. Let say we have an ensemble neural network with N sub-neural networks, the ensemble averaging can be defined as (6):

$$\bar{y}=\frac{1}{N}\sum_{n=1}^{N} y_i^n \quad i=1,2,3\ldots\ldots X \tag{6}$$

where y_i^n is the output of the n-th sub-neural networks and X denotes the data set length.

To comprehensively measuring the performance of the ensemble model, three statistical methods has been used to evaluate the ensemble forecast result against the observed value. These methods are root mean squared error (RMSE), mean absolute error (MAE), and nash-sutcliffe efficiency (NSE). The smaller the number of RMSE and MAE represent better forecast while NSE represents powerful forecast with higher value produce.

2.3. Model development

In developing forecast model for water level, data from eight rainfall stations along Galas River and Lebir River in Kelantan River Basin are utilized which include Gunung Gagau, Kuala Koh, Kampung Aring, Kampung Lalok, Kampung Tualang, Kuala Krai, Dabong and Limau Kasturi. These rainfall stations are all located in the upstream. These 2 rivers are the primary tributaries of Kelantan River. The rainfall data from these stations are the predictors to forecast the Kuala Krai water level that located at the downstream.

These data range from April 2011 to November 2019 and in monthly pattern, 75% of this data are used in the training phase of the model while the rest are used to test the model. As Kuala Krai is prone to flooding, water level forecasting is considered vital. To accurately forecast the monthly water level of Kuala Krai, an ensemble neural networks based on DWT decomposition and MI are developed. To verify the ensemble model, BPNN and ELM model also been developed as independent model for comparison and their forecast result will be assessed. To accommodate validation of the results, every independently developed model and the ensemble model are executed three times and averaged the recorded performance to assess the accuracy. The forecast result of the ensemble model then be compared to the forecast result of the independent models.

3. RESULTS AND DISCUSSION

Kuala Krai is prone to flood every year due to the changes of Northeast Monsoon especially during December-January. In nature, the relation between rainfall and water level is highly nonlinear. An accurate water level forecast is crucial to minimize the flood risk. Four category of water level has been outlined by Department of Irrigation and Drainage Malaysia which are normal level (17.00 m), alert level (20.00 m), warning level (22.5 m) and danger level (25.00 m).

The original time-series data contains various variability which can affect the model performance. The pre-processing of original data in this study tends to provide the model with cleaner input that could elevate the model performance. The data first been decomposed using DWT. The coefficient of the sub-signal produces then been evaluated using the Spearman correlation coefficient (p-value). Selected sub-signal with higher coefficient with original data then been selected to become a feature in the multidimensional dataset.

After all the original timeseries data from all stations are decomposed, and sub-signal has been selected, the multi-dimensional dataset is constructed by treating each of these selected signals as the features of the dataset. Hence, altogether there are eight features (selected sub-signal of each of the rainfall station) and one output (water level of Kuala Krai) defined in the multidimensional dataset. The result of selected sub-signal is as in Table 1. The selected feature is labelled with features identification (ID) for ease of references. The data structure for the multidimensional dataset is as in Figure 3. Further optimization is done by evaluating all the features in the multidimensional dataset using MI. From the process, features are rank, and the evaluation result is presented in Table 2. All features in the rank are executed repeatedly as combination with elimination of the lowest rank in each repetition until only the highest rank is executed.

Table 1. Selected sub-signal for each rainfall stations after DWT decomposition

Rainfall station	Total sub-signal	Selected sub-signal	p-value	Features ID for selected sub-signal
Gunung Gagau	cA2, cD2, cD1	cA2	0.753	GG
Kuala Koh	cA2, cD2, cD1	cA2	0.664	KK
Kampung Aring	cA2, cD2, cD1	cA2	0.552	KA
Tualang	cA2, cD2, cD1	cA2	0.642	TU
Kampung Lalok	cA2, cD2, cD1	cD1	0.581	KL
Kuala Krai	cA2, cD2, cD1	cA2	0.612	KR
Limau Kasturi	cA2, cD2, cD1	cA2	0.711	LK
Dabong	cA2, cD2, cD1	cA2	0.626	DA

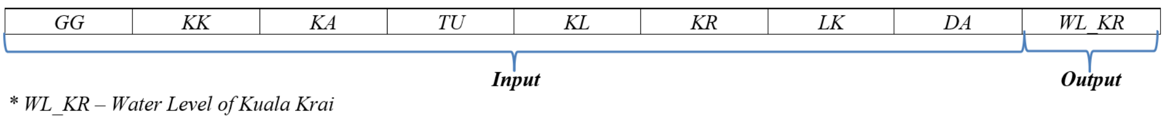

Figure 3. Data structure for the multi-dimensional dataset

Table 2. MI rank

More to less dominant							
GG	KR	DA	TU	KK	KA	KL	LK

By using the optimized dataset gained in the pre-processing phase, the ensemble of neural networks is executed three times for each combination order and the performance measurement of all execution is averaged to validate the result. The ensemble model performance is statistically measured using RMSE, MAE and NSE. The performance measurement comparison is extended to the model using original data without any optimization. Table 3 present the forecast result for BPNN, ELM and ensemble neural networks model with original data without any optimization, while Table 4 present the performance measurement of RMSE, MAE and NSE for the individual and ensemble neural networks model using optimized datasets based on DWT and MI.

Table 3. BPNN, ELM, and ensemble neural networks forecast result using unoptimized data

	RMSE	MAE	NSE
BPNN	1.4805	1.0608	-6.4814
ELM	1.4266	0.8918	-0.7956
Ensemble neural networks	1.3065	0.9054	-2.1685

Table 4. BPNN, ELM, and ensemble neural networks forecast result using data optimized by DWT and MI

		Number of features							
		8	7	6	5	4	3	2	1
BPNN	RMSE	1.3748	1.1576	1.3194	1.3521	1.1846	1.1876	1.3518	1.3652
	MAE	1.0502	0.9343	0.9924	1.0385	0.9389	0.9762	1.0376	1.0611
	NSE	-6.4678	-2.1375	-4.6931	-4.4639	-3.1113	-2.8168	-4.4582	-4.9719
ELM	RMSE	1.0535	1.1208	1.0492	0.9650	1.1437	1.2443	1.1428	1.3465
	MAE	0.7335	0.7808	0.7156	0.7217	0.7586	0.8102	0.8365	0.9400
	NSE	0.1689	-0.1626	0.2487	0.2380	0.2320	0.1872	-0.6184	-0.9372
Ensemble neural networks	RMSE	1.0682	1.0817	1.0465	1.0156	0.8855	1.0463	1.1687	1.3565
	MAE	0.7830	0.7498	0.7722	0.7444	0.6559	0.8062	0.8679	0.9437
	NSE	-1.2421	-1.6688	-0.6290	-0.3462	0.4054	-0.7565	-1.6825	-2.3024

The performance measurement indicates that based on the pattern, the use of ensemble neural networks model in forecasting water level of Kuala Krai has improved the overall forecast performance and minimizing the error compared to the model that use unoptimized data and independently executed model. The lowest score for RMSE and MAE, and the highest score for NSE are achieved by the ensemble model with the value of 0.8855, 0.6559 and 0.4054 respectively. This has been achieved by executing the ensemble model using four features which are decomposed rainfall data from Gunung Gagau, Kuala Krai, Dabong and Tualang stations. The optimization process using DWT and MI has reduced the dimensionality of the dataset. This means we can use less data to achieve better forecast. It is also found that the proposed ensemble model has outperformed the single independently executed model which are BPN and ELM.

Figure 4 present the forecasted water level by the ensemble neural networks model compared to the observed water level. Here, we analyze the best forecasted water level for the ensemble model with DWT and MI optimization using four features. The model is executed three times and the performance measurement are averaged, thus Figure 4 shown the results for every execution. Based on Figure 4, the ensemble model pattern works well following the observed data trend, but when spike is occurred in the observed data, it seems to behave differently.

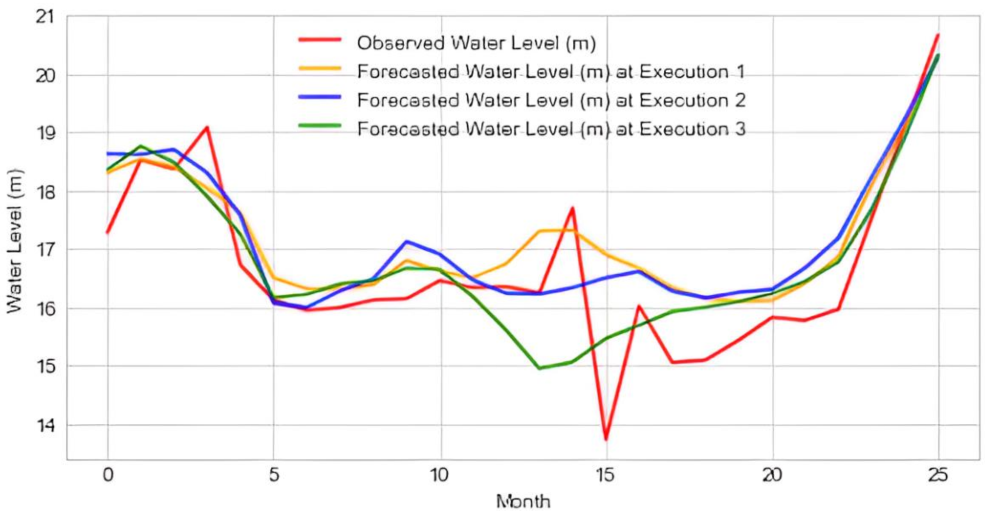

Figure 4. Observed vs forecasted water level for each execution of the ensemble model

In Kelantan River, rainfall is influenced by monsoon movement that occurred mostly during the end and early of the year. During this season, rainfall might reach to a certain category level that can lead to flood. To validate the ensemble neural networks model towards the observed data, the forecast value resulted from the ensemble model with input optimization is distributed within the boxplot in Figure 5. Based on the four category of water level outlined by the Department of Irrigation and Drainage Malaysia, it is indicated that all three executions (labelled as exc1 for first execution, exc2 for second execution and exc3 for third execution) of the ensemble model have produced maximum forecast point of over twenty in which reaching the alert level. This has foreseen the alert level of Kuala Krai's water level for the same period. The ensemble model has proven to provide better generalization of the water level forecast using rainfall data for Kelantan River by averaging the outputs of independent neural networks.

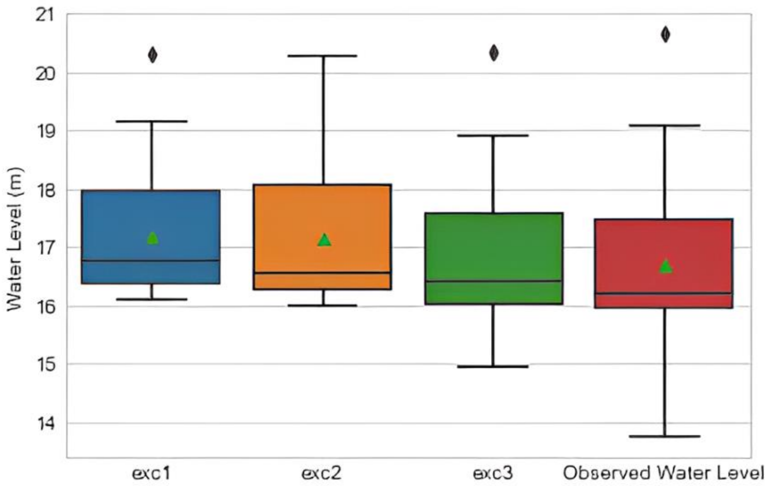

Figure 5. Observed vs forecasted water level for ensemble neural networks model

4. CONCLUSION

This study proposed an ensemble neural networks model with input optimization based on DWT and MI for long-term flood forecasting. The ensemble neural networks model is used to forecast the monthly water level of Kuala Krai. Neural network has been deemed to be a useful tool in flood forecasting, but they imposed high flexibility that led to high variance of the model and poor generalization. The ensemble of neural networks based on decomposition by DWT and ranked by MI has reduced the generalization error compared to the independently executed neural network model by selecting the dominant features for the forecast. It also reduced dimensionality of dataset in which an accurate forecast can be done with minimal inputs. The ability of ensemble model to provide an accurate forecast with reduce RMSE, MAE and higher NSE flood forecasting can help relevant agencies and citizen to gain early knowledge and information to manage the risk of flooding.

ACKNOWLEDGEMENTS

The authors of this study express their gratitude to the Department of Irrigation and Drainage Malaysia, under the Ministry of Energy Transition and Water Transformation, for their provision of the necessary data.

REFERENCES

[1] H. Huang, Z. Liang, B. Li, D. Wang, Y. Hu, and Y. Li, "Combination of multiple data-driven models for long-term monthly runoff predictions based on bayesian model averaging," *Water Resources Management*, vol. 33, no. 9. pp. 3321–3338, 2019, doi: 10.1007/s11269-019-02305-9.

[2] B. Yaghoubi, S. A. Hosseini, and S. Nazif, "Monthly prediction of streamflow using data-driven models," *Journal of Earth System Science*, vol. 128, no. 6, pp. 1–15, 2019, doi: 10.1007/s12040-019-1170-1.

[3] B. B. Sahoo, R. Jha, A. Singh, and D. Kumar, "Application of support vector regression for modeling low flow time series," *KSCE Journal of Civil Engineering*, vol. 23, no. 2, pp. 923–934, 2019, doi: 10.1007/s12205-018-0128-1.

[4] Z. Zahmatkesh and E. Goharian, “Comparing machine learning and decision making approaches to forecast long lead monthly rainfall: The city of Vancouver, Canada,” *Hydrology*, vol. 5, no. 1, pp. 1–22, 2018, doi: 10.3390/hydrology5010010.
[5] M. Alizamir, O. Kisi, and M. Zounemat-Kermani, “Modelling long-term groundwater fluctuations by extreme learning machine using hydro-climatic data,” *Hydrological Sciences Journal*, vol. 63, no. 1, pp. 63–73, 2018, doi: 10.1080/02626667.2017.1410891.
[6] Z. M. Yaseen, R. C. Deo, I. Ebtehaj, and H. Bonakdari, “Hybrid data intelligent models and applications for water level prediction,” in *Handbook of research on predictive modeling and optimization methods in science and engineering*, no. 1, 2018, pp. 121–139, doi: 10.4018/978-1-5225-4766-2.ch006.
[7] A. M. Al-Juboori, “Generating Monthly Stream Flow Using Nearest River Data: Assessing Different Trees Models,” *Water Resources Management*, vol. 33, no. 9, pp. 3257–3270, 2019, doi: 10.1007/s11269-019-02299-4.
[8] A. Mosavi and M. Edalatifar, "A hybrid neuro-fuzzy algorithm for prediction of reference Evapotranspiration," *Confetences Lecture Notes in Networks and Systems*, 2019, vol. 53, pp, 235–243, doi: 10.1007/978-3-319-99834-3_31.
[9] N. M. Khairudin, N. Mustapha, T. N. M. Aris, and M. Zolkepli, “in-Depth Review on machine learning models for long-term flood forecasting,” *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 10, pp. 3360–3378, 2022.
[10] A. Mosavi, P. Ozturk, and K. W. Chau, “Flood prediction using machine learning models: Literature review,” *Water (Switzerland)*, vol. 10, no. 11, pp. 1–40, 2018, doi: 10.3390/w10111536.
[11] F. Sajedi-Hosseini *et al.*, “A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination,” *Science of the Total Environment*, vol. 644, pp. 954–962, 2018, doi: 10.1016/j.scitotenv.2018.07.054.
[12] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The Bulletin of mathematical Biophysics*, vol. 5, pp. 115–133, 1943.
[13] S. V. Lakshminarayana, “Rainfall forecasting using artificial neural networks (ANNs): A comprehensive literature review,” *Indian Journal of Pure and Applied Biosciences*, vol. 8, no. 4, pp. 589–599, 2020, doi: 10.18782/2582-2845.8250.
[14] M. A. A. Albadr and S. Tiun, “Extreme learning machine: A review,” *International Journal of Applied Engineering Research*, vol. 12, no. 14, pp. 4610–4623, 2017.
[15] G. Bin Huang, Q. Y. Zhu, and C. K. Siew, “Extreme learning machine: Theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006, doi: 10.1016/j.neucom.2005.12.126.
[16] D. Sotelo, D. Velásquez, C. Cobos, M. Mendoza, and L. Gómez, “Optimization of neural network training with ELM based on the iterative hybridization of differential evolution with local search and restarts,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 38–50, 2019, doi: 10.1007/978-3-030-13709-0_4.
[17] C. M. Saporetti, G. R. Duarte, T. L. Fonseca, L. G. Da Fonseca, and E. Pereira, “Extreme learning machine combined with a differential evolution algorithm for lithology identification,” *Revista de Informática Teórica e Aplicada*, vol. 25, no. 4, 2018, doi: 10.22456/2175-2745.80702.
[18] E. S. K. Tiu, Y. F. Huang, J. L. Ng, N. AlDahoul, A. N. Ahmed, and A. Elshafie, *An evaluation of various data pre-processing techniques with machine learning models for water level prediction*, vol. 110, no. 1. 2022, doi: 10.1007/s11069-021-04939-8.
[19] Y. Yu, H. Zhang, and V. P. Singh, “Forward prediction of Runoffdata in data-scarce basins with an improved ensemble empirical mode decomposition (EEMD) model,” *Water (Switzerland)*, vol. 10, no. 4, pp. 1–15, 2018, doi: 10.3390/w10040388.
[20] N. Khairuddin, A. Z. Aris, A. Elshafie, T. Sheikhy Narany, M. Y. Ishak, and N. M. Isa, “Efficient forecasting model technique for river stream flow in tropical environment,” *Urban Water Journal*, vol. 16, no. 3, pp. 183–192, 2019, doi: 10.1080/1573062X.2019.1637906.
[21] Y. Zhou, S. Guo, and F. J. Chang, “Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts,” *Journal of Hydrology*, vol. 570, pp. 343–355, 2019, doi: 10.1016/j.jhydrol.2018.12.040.
[22] M. Tayyab, J. Zhou, X. Dong, I. Ahmad, and N. Sun, “Rainfall-runoff modeling at Jinsha River basin by integrated neural network with discrete wavelet transform,” *Meteorology and Atmospheric Physics*, vol. 131, no. 1, pp. 115–125, 2019, doi: 10.1007/s00703-017-0546-5.
[23] G. Zuo, J. Luo, N. Wang, Y. Lian, and X. He, “Decomposition ensemble model based on variational mode decomposition and long short-term memory for streamflow forecasting,” *Journal of Hydrology*, vol. 585, 2020, doi: 10.1016/j.jhydrol.2020.124776.
[24] H. Shafizadeh-Moghadam, R. Valavi, H. Shahabi, K. Chapi, and A. Shirzadi, “Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping,” *Journal of Environmental Management*, vol. 217, pp. 1–11, 2018, doi: 10.1016/j.jenvman.2018.03.089.
[25] X. M. Kong, G. H. Huang, Y. R. Fan, and Y. P. Li, “Maximum entropy-Gumbel-Hougaard copula method for simulation of monthly streamflow in Xiangxi river, China,” *Stochastic Environmental Research and Risk Assessment*, vol. 29, no. 3, pp. 833–846, 2015, doi: 10.1007/s00477-014-0978-0.
[26] E. Neachell, “Book review-environmental flows: saving rivers in the thrid millennium,” *River Research and Applications*, vol. 30, no. January, pp. 132–133, 2014, doi: 10.1002/rra.
[27] A. Azizpour, M. A. Izadbakhsh, S. Shabanlou, F. Yosefvand, and A. Rajabi, “Estimation of water level fluctuations in groundwater through a hybrid learning machine,” *Groundwater for Sustainable Development*, vol. 15, Nov. 2021, doi: 10.1016/j.gsd.2021.100687.
[28] M. Hammad, M. Shoaib, H. Salahudin, M. A. I. Baig, M. M. Khan, and M. K. Ullah, “Rainfall forecasting in upper Indus basin using various artificial intelligence techniques,” *Stochastic Environmental Research and Risk Assessment*, vol. 35, no. 11, pp. 2213–2235, 2021, doi: 10.1007/s00477-021-02013-0.
[29] N. Lv *et al.*, “A long short-term memory cyclic model with mutual information for hydrology forecasting: A Case study in the xixian basin,” *Advances in Water Resources*, vol. 141, 2020, doi: 10.1016/j.advwatres.2020.103622.
[30] J. Ren, B. Ren, Q. Zhang, and X. Zheng, “A novel hybrid extreme learning machine approach improved by K nearest neighbor method and fireworks algorithm for flood forecasting in medium and small watershed of Loess region,” *Water (Switzerland)*, vol. 11, no. 9, 2019, doi: 10.3390/w11091848.
[31] Z. M. Yaseen, M. F. Allawi, A. A. Yousif, O. Jaafar, F. M. Hamzah, and A. El-Shafie, “Non-tuned machine learning approach for hydrological time series forecasting,” *Neural Computing and Applications*, vol. 30, no. 5, pp. 1479–1491, 2018, doi: 10.1007/s00521-016-2763-0.
[32] B. S. Prasad, S. Shreya, T. Sinha, and S. Priya, “Ensemble model-based prediction for river management: A case study on River Kaveri and Coastal Karnataka,” 2021, doi: 10.1109/SMARTGENCON51891.2021.9645804.
[33] R. Jovanović, A. A. Sretenović, and B. D. Živković, “Ensemble of various neural networks for prediction of heating energy consumption,” *Energy and Buildings*, vol. 94, pp. 189–199, 2015, doi: 10.1016/j.enbuild.2015.02.052.

BIOGRAPHIES OF AUTHORS

Nazli Mohd Khairudin is a Postgraduate student at the Faculty of Computer Sciences and Information Technology, Universiti Putra Malaysia. He received B.Sc. (Hons) (Computer) from Universiti Teknologi Malaysia in 2001 and MSc degree in Software Engineering from Universiti Putra Malaysia in 2013. His major interest in research are data mining, artificial intelligence, and intelligence system. He can be contacted at email: nazmkhair@gmail.com.

Norwati Mustapha is an Associate Professor at the Department of Computer Science. She received the B.Sc. in Computer Science from University Putra Malaysia in 1991 and M.Sc. degree in Information Systems from University of Leeds UK in 1995. She obtained her Ph.D. in artificial intelligence from University Putra Malaysia in 2005. She is an active researcher in data mining, web mining, social networks, and intelligent computing. She can be contacted at email: norwati@upm.edu.my.

Teh Noranis Mohd Aris received her bachelor's degree in Information Technology (Hons.) from University Utara Malaysia (UUM), in 1994. She received her Master of Science in Artificial Intelligence from University Putra Malaysia (UPM) in 2002, and Ph.D. Degree from University Kebangsaan Malaysia (UKM) in 2008. Currently, she is an Associate Professor at the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia. Her research interests include artificial intelligence, fuzzy systems, image processing and software agent. She can be contacted at email: nuranis@upm.edu.my.

Maslina Zolkepli currently is a senior lecturer at the Department of Computer Science, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia. She received her Bachelor of Computer Science in 2007, Master of Computer Science in 2010, both from Universiti Putra Malaysia. In 2015, she received her Doctor of Engineering Degree in the field of Computational Intelligence and Systems Science from Tokyo Institute of Technology, Japan. Her research interests include big data, fuzzy systems, and computational intelligence. She can be contacted at email: masz@upm.edu.my.