Zubaidi, S, Ortega Martorell, S, Al-Bugharbee, H, Olier-Caparroso, I, Hashim, KS, Gharghan, SK, Kot, P and Al Khaddar, RM

Urban Water Demand Prediction for a City that Suffers from Climate Change and Population Growth: Gauteng Province case study

Urban Water Demand Prediction for a City that Suffers from Climate Change and Population Growth: Gauteng Province case study

Salah L. Zubaidi [1,*], Sandra Ortega-Martorell [2], Hussein Al-Bugharbee[3], Ivan Olier [2], Khalid S. Hashim [4,5], Sadik Kamel Gharghan [6], Patryk Kot [4], Rafid Al-Khaddar [4]

[1] Department of Civil Engineering, Wasit University, Wasit, Iraq,

[2] Department of Applied Mathematics, Liverpool John Moores University, Liverpool, UK,

[3] Department of Mechanical Engineering, Wasit University, Wasit, Iraq,

[4] Built Environment and Sustainable Technologies (BEST) Research Institute, Liverpool John Moores University, United Kingdom.

[5] Department of Environment Engineering, Babylon University, Babylon, Iraq

[6] Department of Medical Instrumentation Techniques Engineering, Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq.

[*]Email: salahlafta@uowasit.edu.iq

Abstract

The proper management of municipal water system is essential to sustain cities and support water security of societies. Urban water estimating has always been a challenging task for managers of water utilities and policymakers. This paper applies a novel methodology that includes data pre-processing and Artificial Neural Network (ANN) optimized with Backtracking Search Algorithm (BSA-ANN) to estimate monthly water demand in relation to previous water consumption. Historical data of monthly water consumption in the Gauteng Province, South Africa, for the period 2007–2016, were selected for the creation and evaluation of the methodology. Data pre-processing techniques played a crucial role in the enhancing of the quality of the data before creating the prediction model. The BSA-ANN model yielded the best result with a root mean square error and a coefficient of efficiency of 0.0099 mega liters and 0.979, respectively. Also, it proved more efficient and reliable than the Crow Search Algorithm (CSA-ANN), based on the scale of error. Overall, this paper presents a new application for the hybrid model BSA-ANN that can be successfully used to predict water demand with high accuracy, in a city that heavily suffers from the impact of climate change and population growth.

Keywords: Artificial Neural Network; Backtracking Search Algorithm; Municipal water demand; Climate Change; Population Growth.

1. Introduction

Urban water security is essential to get a resilient environment in smart cities, particularly under the stress of climate change and socio-economic factors [1,2]. Also, cities located close to water resources are driven by all kinds of industries hence, water lack is considered a classic problem for decision-makers [3,4]. Since the last century, gradual changes in freshwater resources have been observed [5]. Recent studies related to climate change have shown that it plays a key role on freshwater resources due to the potential decrease in rainfall amount [6]. Specifically, it has been shown that climate change adversely impacts freshwater resources in the centre of cities, which in turn impacts the sustainable development of water availability and consequently, impacts socio-economic activities [7]. In addition, several studies have shown that freshwater resources are generally adversely affected by pollution [8,9].

Different regions in the world have been facing water scarcity situations, which implies that the gap between water supply and demand is likely to increase in the future. The European Environment Agency in 2010 reported that municipal water consumption is driven by complicated interactions between anthropogenic and natural system factors at multiple spatial and temporal scales [10-12]. In the Gauteng Province, the Republic of South Africa, the municipal water delivered has been less than the demand. This imbalance is due to the impact of climate change, rainfall reduction, as well as others that are human-related, such as economic expansion and population growth. The lack of freshwater resources and the increase in water demand has put pressure on the municipal water supply system. Hence the importance of using the prediction of water demands as an effective approach for optimizing the operation and management of the system, or plan for future expansion or reduction under the variability of climate and socio-economic factors [2,13,14].

House-Peters and Chang [15], Donkor, *et al.* [16], Ghalehkhondabi, *et al.* [17] and de Souza Groppo, *et al.* [18] stated that different methods and models have been applied in previous studies to predict municipal water demand, including traditional, Artificial Intelligence (AI), and hybrid AI models. Traditional models, such as time-series analysis and regression [19,20], were firstly employed in water demand simulation. However, traditional approaches lacked accuracy when forecasting water demand, which can cause significant issues in the operation and management of the water supply system. Additionally, the growth of the impact of climate change and urbanization cause high uncertainty, making the prediction and forecasting more complex, which also motivated researchers to further develop their models [21], including the use of AI techniques.

Data-driven techniques have a far-ranging application such as wastewater [22,23], water demand [24,25], groundwater levels [26]. Some of these techniques include support vector machine (SVM) [27], extreme learning machine (ELM) [24], and random forest (RF)[28]. One of these AI techniques is Artificial Neural Networks (ANN) [29], which is a powerful technique that has been widely used in hydraulic modelling in recent years. It has the capability to deal with complex and nonlinear relationships between inputs and outputs [30,31]. The results obtained when applying ANN have been superior to all types of conventional model in many scenarios, for example Mouatadid and Adamowski [32] and Guo, *et al.* [33]. However, there are cases where conventional methods performed as well as or even better than ANN in terms of accuracy such as Li, et al. [27]. The latter can be due to a number of reasons, for example that the models falling into a local instead of the global minimum, leading to a sub-optimal solution [34], or not using the right network design or hyperparameters for training the neural network [35]. Hence, in order to avoid these drawbacks different approaches have been combined with the ANN model such as heuristic algorithms [36], and different hybrid models have been proposed.

A hybrid model contains two or more techniques; one of them would work as the primary model, while others would act as pre-processing or post-processing approaches [37]. Hybrid models have been used to simulate municipal water demand using different techniques and in different scenarios, and the results have revealed that these models are robust and insightful, e.g. Altunkaynak and Nigussie [38], Seo, et al. [24], Pacchin, *et al.* [39], Ebrahim Banihabib and Mousavi-Mirkalaei [2] and Rasifaghihi, *et al.* [40].

Eggimann, *et al.* [41] reviewed various techniques of data pre-processing that have been used for municipal water management. The reviewed article reveals that data pre-processing techniques have an important potential advantage for optimizing the performance of prediction models. It has applied successfully in different areas of study, e.g., monthly rainfall forecasting [42], irrigation water prediction [43] and urban water demand prediction [24].

Various optimization techniques have been applied to solve problems in engineering applications. The optimization algorithms aim to detect optimal values for the parameters of the system under various conditions [44]. Lately, the crow search algorithm (CSA), a recently proposed metaheuristic algorithm, has been used to tackle a variety of optimization engineering issues [45].

CSA was applied to solve optimization issues in different engineering sectors such as the optimization of energy problems [45], economic environmental dispatch [46], the selection of the optimal size of conductor in radial distribution networks [47], water demand prediction [48] and to solve constrained engineering [49]. In this study, the CSA will be hybridized with the ANN model to select the best hyperparameters of the ANN model.

From the application area viewpoint, another significant consideration is that the selection of best model input that drives the dependent variable [50,51]. Several techniques were applied in different studies such principal component analysis (PCA) [52,53], variance inflation factor (VIF) [21,35] and mutual information (MI) [54,55]. In this study, mutual information technique will use to select the best scenario of model input based on several historical observed water consumption data.

According to the literature review, another significant consideration is that most of the studies focus on short-term water demand estimate, while only a few deals with medium to long-term prediction. Lately, various studies such as [33,56-58] have employed historical data of water consumption as a single input in their short-term prediction models.

However, a challenge still exists for managers of water utilities and policymakers due to the uncertainty to gain knowledge about the capacity of water system under a potential rapid growth in urban water demand as a consequence of socio-economic, demographic and climate factors. Also, as mentioned previously, only a few studies have considered medium-term municipal water demand based on previous water consumption. Therefore, these aforementioned problems motivated us to propose an approach that would refine those existing approached, providing managers with scientific, more accurate insights about the future water demand, reducing the uncertainty.

The main objectives of this research study are:

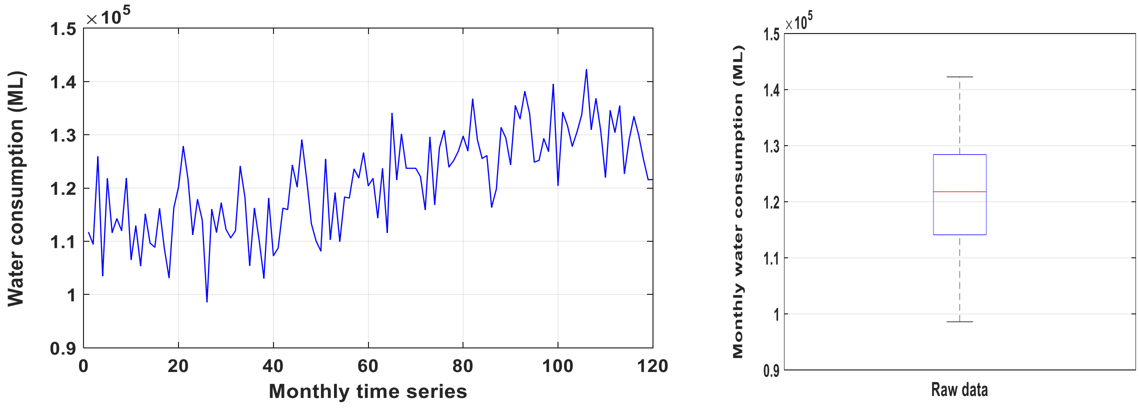
1- To improve the quality of the data and to choose the best model input scenario by applying data pre-processing techniques.
2- To select the optimum values of ANN hyperparameters by using Backtracking Search Algorithm and Artificial Neural Network (BSA-ANN) technique. Also, to evaluate how BSA-ANN performs in comparison with a CSA-ANN algorithm.
3- To assess the performance of the novel methodology to predict medium-term municipal water demand in relation to some lags time of observed water consumption.
4- To reduce the uncertainty for decision-makers by using a novel and refined model, which involves data pre-processing methods (to improve the quality of data and select the model input), and employing a more sophisticated approach for model prediction (using combined techniques to enhance the accuracy of results, and the stand-alone ANN to confirm the results of the hybrid model).

Based on the literature review, the research is thought to be the first study that used this novel combined methodology, which include data pre-processing and automated machine learning to forecast municipal water demand depend on some lags' values of water consumption as model input. As such, it is considering the effect of all climate, demographic and socio-economic factors.

2. Study Area and Data Collection

Gauteng province is the economic powerhouse of the Republic of South Africa, which has eight metropolitan municipalities. This city faced water stress that resulted from climate change, the average annual rainfall was below the world's average of 363mm, and from human relation such as population growth and economic expansion. More than 60% of the population living in the urban regions in South Africa, and Gauteng province receives most migrants in this country. For this city, it is anticipated that the water demand would outstrip the water delivered by 2025. For more than a century, the company Rand Water has delivered municipal water to more than 9 million people and different industries in the Gauteng province, with more than 3000 km of pipeline. The lack of freshwater resources in the Gauteng province has motivated Rand Water to increase storage capacity by constructing new dams and water transfer schemes from several rivers of different regions such as the Vaal, Tugela and Orange rivers [13,59,60].

Historical monthly data of municipal water consumption (in Mega liters, ML) over ten years from 2007 to 2016 were provided by Rand Water and used to build and assess the model. Two pre-tests were applied to these data by SPSS (24) package, one of them being Komarov-Semenove test to assess normality and the other one being a box-whisker test to check for outliers. The results showed that these data are normally distributed, the value of significance is 0.2 > 0.05, and data are clean from outliers, data lies between ±1.5 IQR. These results increase the reliability on the quality of data received from the company. Figure 1 shows the municipal water consumption: a) monthly time series, b) boxplot for Rand Water company.



a b

Figure 1. Municipal water consumption: a) monthly time series, b) boxplot for Rand Water company.

3. Methodology

The proposed methodology can be divided into four parts, including data pre-processing, Artificial Neural Network, Backtracking Search Algorithm and model evaluation.

3.1. Data Pre-processing

Pre-processing the data has a significant effect on the quality of the model produced. At this stage, we perform three steps: normalisation, cleaning and selection of best model inputs. Data normalisation aims to have the same range of values for each of the inputs to the ANN model and to make the time series normally or close to normally distributed, as it would assist the stable convergence of the weights and biases as well as reduce the impact of noise [61]. In this research, natural logarithm was used for normalizing the data because it has the ability to minimize the effects of the multicollinearity between independent variables [37].

The aim of the cleaning approach is to detect and remove the noise from the time series to increase the regression coefficient and decrease the scale of error [21]. All the time series have different components of noise, and the pre-treatment signal is one of the best approaches that denoises the raw time series by decomposing them into different components [62]. This approach can be applied for both linear and nonlinear time series with different sample size -short, medium and long-term. It does not need any assumption of statistical criteria such as normality of error, linearity and stationery of the series [62,63]. More details about the pre-treatment technique can be found in Golyandina and Zhigljavsky [64]. This technique has been applied in several research areas, including predicting stochastic processes [65], hydrology [66] and economics [63].

The selection of the best model input represents one of the most important stages in data pre-processing in general, which is also the case when modelling the forecast of water demand [31]. In this research, the choice of the best explanatory variables is performed by applying Mutual Information (MI) technique. It is used for measuring the statistical correlation between the original time series and the lagged components. This technique enables the selection of the highest correlation components which have the greater mutual information [67].

3.2. Artificial Neural Network (ANN)

ANN is a method inspired by the way the human brain processes data, and emulates its functionality by using similar operations and connectivity as a biological neural system [29,30,68]. Recently, ANN models have been widely utilised in water resources and hydrology applications because of its ability to extract complex nonlinear relationships, which exist within the hydrology data [30,31].

In this study, the multilayer perceptron (MLP) is applied to simulate municipal water demand. MLP has been frequently and successfully used for the forecast of water resources and hydrology applications. Its architecture and hyperparameters (as shown in Table 1) is layered as a feedforward neural network (FFNN) and can be trained using learning algorithms such as the backpropagation of the error (BP) [69] and the Levenberg-Marquardt (LM) [70,71]. It has been reported that the latter is better at limiting the errors of the ANN [30,31]. As in Zubaidi et al., [37,48], the structure of the MLP contains four layers, the first one being the input layer, which has the model inputs representing water consumption lags, followed by two hidden layers and one output layer, which has the water demand. Two types of activation functions have been used: a tan-sigmoidal function in the hidden layers as in Yonaba*, et al.* [72], and a linear activation function in the output layer for covering the positive values of urban water demand as successfully used in Zubaidi, et al. [21]. The ANN model was integrated by using backtracking search optimization algorithm (BSA-ANN) to locate the optimum hidden neurons' number and optimal coefficient of learning rate that maximizes the ability and reliability of the ANN technique [36,73]. The training process of the ANN model is repeated a large number of times over an epoch (i.e., 1000 iterations) until the error between the observed and simulated urban water reaches its minimum. The data were split randomly into three sets 70% for training, 15% for testing and 15% for validation, as previously conducted by Zubaidi, et al. [21] and Zubaidi, et al. [35]. As in Gharghan, et al. [36], cross-validation was used to ensure the generalization capabilities of the model and avoid overfitting, and the stopping criterion for training was done using the root mean square error (RMSE) as an objective function (i.e., error not more than the value of RMSE in the testing stage). This procedure was also used successfully by Zubaidi et al., [37,48].

Table 1. The ANN hyperparameters.

Parameter	Type
Number of inputs	Estimated by Mutual Information (MI) technique
Number of outputs	Our target, which is water demand
Number of hidden layers	Two hidden layers
Number of neurons in hidden layer N1	Estimated by metaheuristic algorithm
Number of neurons in hidden layer N2	Estimated by metaheuristic algorithm
Learning rate coefficient	Estimated by metaheuristic algorithm
Learning algorithm	Levenberg-Marquardt (LM)
Activation function in hidden layer N1	Tansigmoidal activation function
Activation function in hidden layer N2	Linear activation function
Number of epochs	1000 iterations

3.3. Backtracking Search Algorithm (BSA)

The BSA algorithm is an evolutionary algorithm, proposed by Civicioglu to remedy the complex problems of numerical optimization, e.g. highly nonlinear, non-differentiable, constrained design problems and multimodality [73-75]. BSA has been broadly applied to tackle different types of engineering optimization issues, e.g. numerical function optimization [74], constrained engineering optimization problems [75], wireless sensor [36], and home energy management [44]. It can be sorted into five stages: initialization, selection-I, mutation, crossover, and selection-II [75].

Initialization: this stage initializes primary population P and history population oldP with Equations. (1) and (2), consecutive:

$$P_{i,j} \sim \mathrm{U}(low_j, up_j) \qquad (1)$$

$$oldP_{i,j} \sim \mathrm{U}(low_j, up_j) \qquad (2)$$

Where,

i= 1, 2, 3, …., N; N is the population size; U is the uniform distribution.

j= 1, 2, 3… D; D is the problem dimension.

BSA's Selection-I: in this stage, BSA algorithm re-chooses a new oldP to calculate the search direction through the 'if-then' rule in Equation (3) and the permuting's function in Equation (4) is utilized to change randomly individuals order in oldP. This stage confirms that the BSA algorithm has memory.

$$oldP \coloneqq P \;/\; a, b \sim \mathrm{U}(0,1) \qquad (3)$$

$$oldP \coloneqq permuting(oldP) \qquad (4)$$

Mutation: in this stage, BSA algorithm generates the initial trail population form M based on Equation (5)

$$M = P + F.(oldP - P) \qquad (5)$$

Where, F is the responsible for controlling the amplitude of the search direction matrix. It can be obtained by applying Equation (6), where randn is a standard normal random number.

$$\mathrm{F} = 3 \cdot \mathrm{randn} \qquad (6)$$

In this study, we used F=3 as was used before in Gharghan, et al. [36].

Crossover: the last formula of trial population T is generated at this stage. The value of T is limited within the acceptable boundary limitations. The unique crossover phase of BSA algorithm contains two primary phases. The first stage is to adjust a binary integer-valued matrix (map) with size N * D via utilizing map(1: N,1:D) = 1. Then, two various crossover strategies are randomly conducted to set the map, as presented in Equation (7). The second stage is used for updating T based on the defined map utilizing Equation (8).

$$map_{i,u} = 0 \begin{cases} u = \lceil mixrate \cdot rand \cdot D \rceil, & if\ c < d/c, d \sim \mathrm{U}(0,1), \\ u = randi(D), & else, \end{cases} \qquad (7)$$

$$T_{i,j} = \begin{cases} M_{i,j}, & if\ map_{i,j} = 0, \\ P_{i,j}, & else, \end{cases} \qquad (8)$$

Where,

mixrate: is the mix rate parameter, which controls the elements' number that will be altered.

A boundary control mechanism is conducted via applying Equation (9), for avoiding the individuals in T exceeding the search space limits.

$$T_{i,j} = rand \cdot \left(up_j - low_j\right) + low_j, if \quad (T_{i,j} < low_j)\ or \left(T_{i,j} > up_j\right). \qquad (9)$$

Selection-II: this is the final stage of the BSA algorithm, which evaluates the fitness values of the trial population T and population P, and updates the individuals of P according to a greedy selection, as presented in Equation (10).

$$P_i = \begin{cases} T_i, & \text{if } fitness(T_i) < fitness(P_i), \\ P_{i,}, & else. \end{cases} \qquad (10)$$

More details about the BSA algorithm can be found in Civicioglu [73]. In our research study, we have hybridized BSA with ANN to choose the best hyperparameters of the ANN model, as opposed to using trial and error as it may not be reliable. As briefly mentioned earlier, these ANN hyperparameters include the neurons' number in both hidden layers and the coefficient of the learning rate.

3.4. Evaluation Model

Several standard statistical measures can be employed to appraise the performance of the methodology in the validation stage for the selection of the best model that has a minimum mean error to decrease deviations in future forecasts [16]. In this research five criteria were utilised to examine the accuracy of the forecast model: root mean square error (RMSE), mean absolute error (MAE), mean absolute relative error (MARE), coefficient of efficiency (CE) and coefficient of determination (R^2). Also, four tests were applied to assess residual data include Kolmogorov-Smirnov, Shapiro-Wilk, Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test.

4. Results and Discussion

4.1. Development Model Input

After normalizing the data by applying the natural logarithm, the pre-treatment signal technique was employed to obtain the time series data of urban water consumption without noise (this was performed by decomposing the original time series into three signals). Figure 2 shows the original time series (top row), the new time series (second row) and two noise signals (third and fourth rows). Data pre-processing enhance the correlation coefficients between dependent and independents variables for different lags of monthly water consumption e.g. the correlation coefficient of raw data of Lag_1 increased significantly from 0.63 to 0.96. The correlation coefficients for the first four lags are 0.96, 0.91, 0.84 and 0.78, respectively.

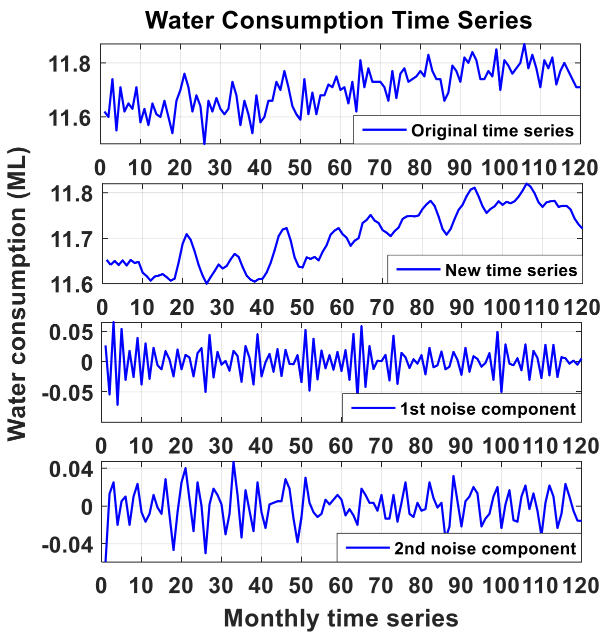


Figure 2. Original time series (top row) and three components of water consumption obtained by the pre-treatment signal technique (2^{nd} to 4^{th} rows). 2^{nd} row represents the new time series, while 3rd and 4th represent noise.

Two boxplots' shapes for normalised and denoised data are shown in Figure 3. It can be seen that there are no outlier's data for both shapes. Additionally, both shapes almost have the same median, the upper and lower quartiles, while the upper and lower extremes of the denoised data are less than those for normalised data because of noise elimination. Moreover, the shape of denoised data is near to normal distribution pattern, better than the normalised data shape.

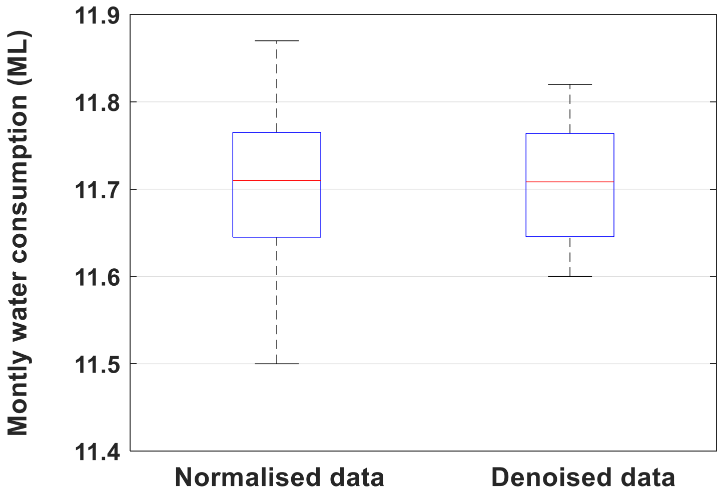


Figure 3. The box plot distribution for normalised, and denoised data.

Further to this, the MI technique was applied to select the best scenario of model input for the prediction model as shown in Figure 4. According to the literature, the first minimum of average mutual information (AMI) is selected as the time lag [76,77]. Based on the figure of AMI, four lags (Lag_1 to Lag_4) of monthly historical water consumption were used to simulate future water demand.

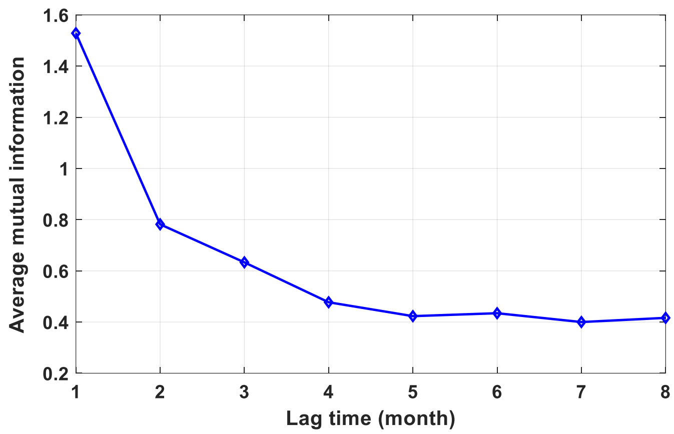


Figure 4. AMI function of the water consumption time series.

Tabachnick and Fidell [61] indicated that the relationship between the size's sample (N) and the independent variables' number should comply with Equation (14).

$$N \geq 50+8m \tag{14}$$

m = number of predictors variables.

In this research, the cases' number is N = 116, which is more than the 82 needed, which indicates compliance with the proposition from Tabachnick and Fidell [61].

4.2. Application Hybrid Heuristic Algorithms-ANN Techniques

After performing data pre-processing methods, data were split into three datasets include training, testing and validation as presented in Table 2. The table tabulates four statistical standards for all data sets include maximum consumption (C_{max}), minimum consumption (C_{min}), mean consumption (C_{mean}), standard deviation (C_{std}) and total sample size for each data set (T). The outcomes show that all sets mostly have the same style.

Table 2. The statistical parameters for training, testing, and validation sets.

Water Consumption (ML)	C_{max}	C_{min}	C_{mean}	C_{Std}	T
Training set	11.81	11.60	11.70	0.062	82
Testing set	11.82	11.61	11.71	0.070	17
Validation set	11.79	11.61	11.72	0.057	17

Five sizes of the population (10, 20, 30, 40 and 50) were used to simulate the hybrid BSA-ANN algorithm in MATLAB toolbox, to locate the optimal population size that offers best learning rate coefficient and number of neurons in both hidden layers of the ANN technique. Figure 5a shows that the population size of 40 offers the optimal answer with less fitness function equal to (0.00608×10^{-3}) after 149 iterations. A CSA-ANN algorithm is applied as well to attain the same objective for the same populations' size and to then to be compared with the outcomes from the hybrid BSA-ANN algorithm, as revealed in Figure 3b. Figure 5b reveals that the population size of 40 gives the optimal solution with less fitness function equal to (0.006497×10^{-3}) after 181 iterations. The result gained from the BSA-ANN algorithm was associated with these from the CSA-ANN algorithm to compare with the new technique. The hybrid BSA-ANN model has a lower RMSE (with less iteration) in comparison to the CAS-ANN. The results of BSA algorithm have been employed to enhance the ANN capabilities in the modeling of municipal water demand. Accordingly, the hyperparameters of the ANN obtained from the best population size were: learning rate coefficient: 0.3954, number of neurons:5 and 2 for hidden layer one and two, respectively.

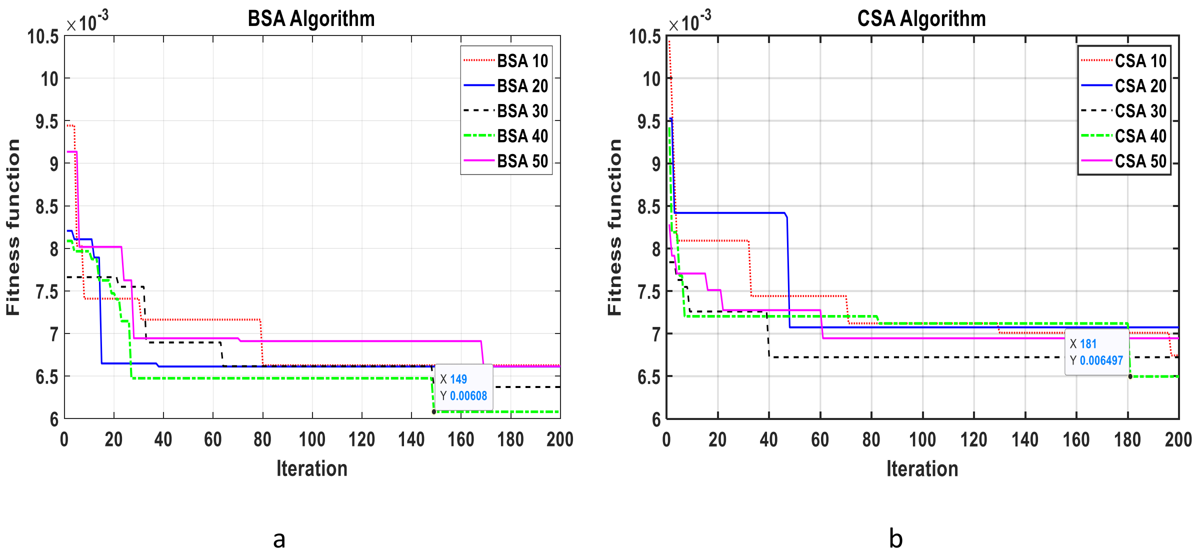


Figure 5. Metaheuristic algorithms simulation for five population size.

The ANN technique was design to estimate the effect of using BSA algorithm in conjunction with the ANN, and to validate the results of the combined model. Consequently, extensively trial and error technique's scenarios were implemented to determine the ANN model's factors (LR, N1, and N2) that offer the optimal precise of prediction. Accordingly, the outcomes present that the values of LR, N1, and N2 are 0.3, 7, and 10, respectively.

To explore the capability and accuracy of the combined model for generalization, the coefficient of determination (R^2) was estimated between the observed and simulated water demand for training, testing and validation sets, as presented in Figure 6. The measured municipal water consumption is indicated in the x-axis and plotted against the simulated water demand in the y-axis. Also, the dataset of the testing stage was employed to plot a regression calibration curve between the observed versus simulated water consumption time series, with a 95% confidence interval (CI). The figure shows that there are neither any irregular data nor a particular pattern trend, and high levels of consistency between the observed and simulated data. Also, the hybrid model was significant R^2= 0.97, 0.97, and 0.98 for training, testing, and validation datasets, respectively. These results support the capabilities of the BSA-ANN model to accurately generalise unseen data (i.e. dataset that was not considered before in training and testing stages).

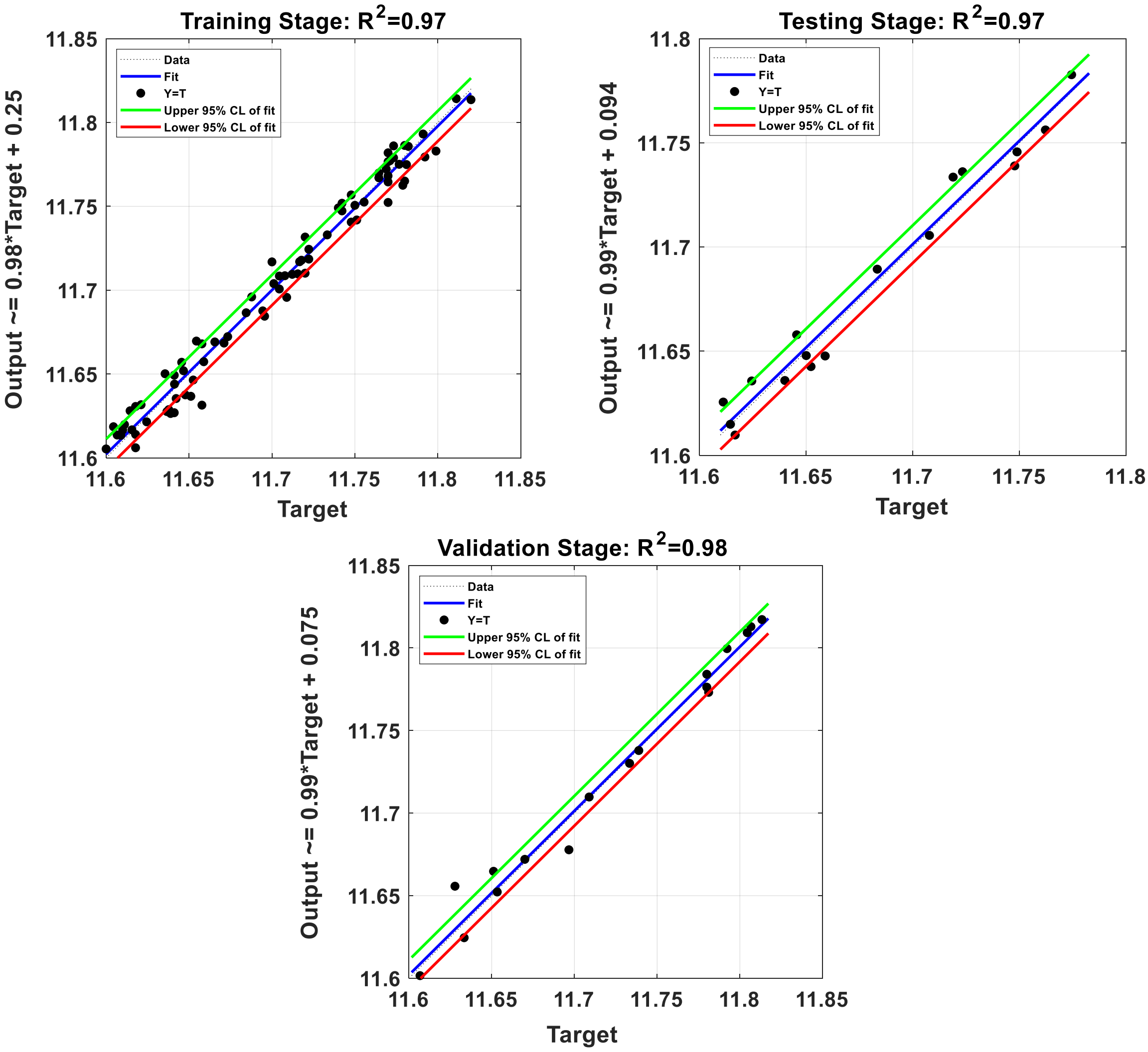


Figure 6. The performance of combined model in training, testing and validation stages.

The coefficient of determination (R^2) criterion was utilised again to evaluate the accuracy of ANN model (stand-alone) and its capability for generalizing data in the validation stage, as presented in Figure 7. The figure shows that R^2= 0.98, 0.96 and 0.95 for training, testing and validation datasets. Although the values of coefficient of determinations for training and testing stages are slightly bigger than the value of the same criteria for validation stage, this is not considered a problem, as was also discussed in Dawson, *et al.* [78], . Hence, we can confidently say that this statistical criterion supports the increased generalization capabilities of the BSA-ANN model compared with the ANN model (stand-alone).

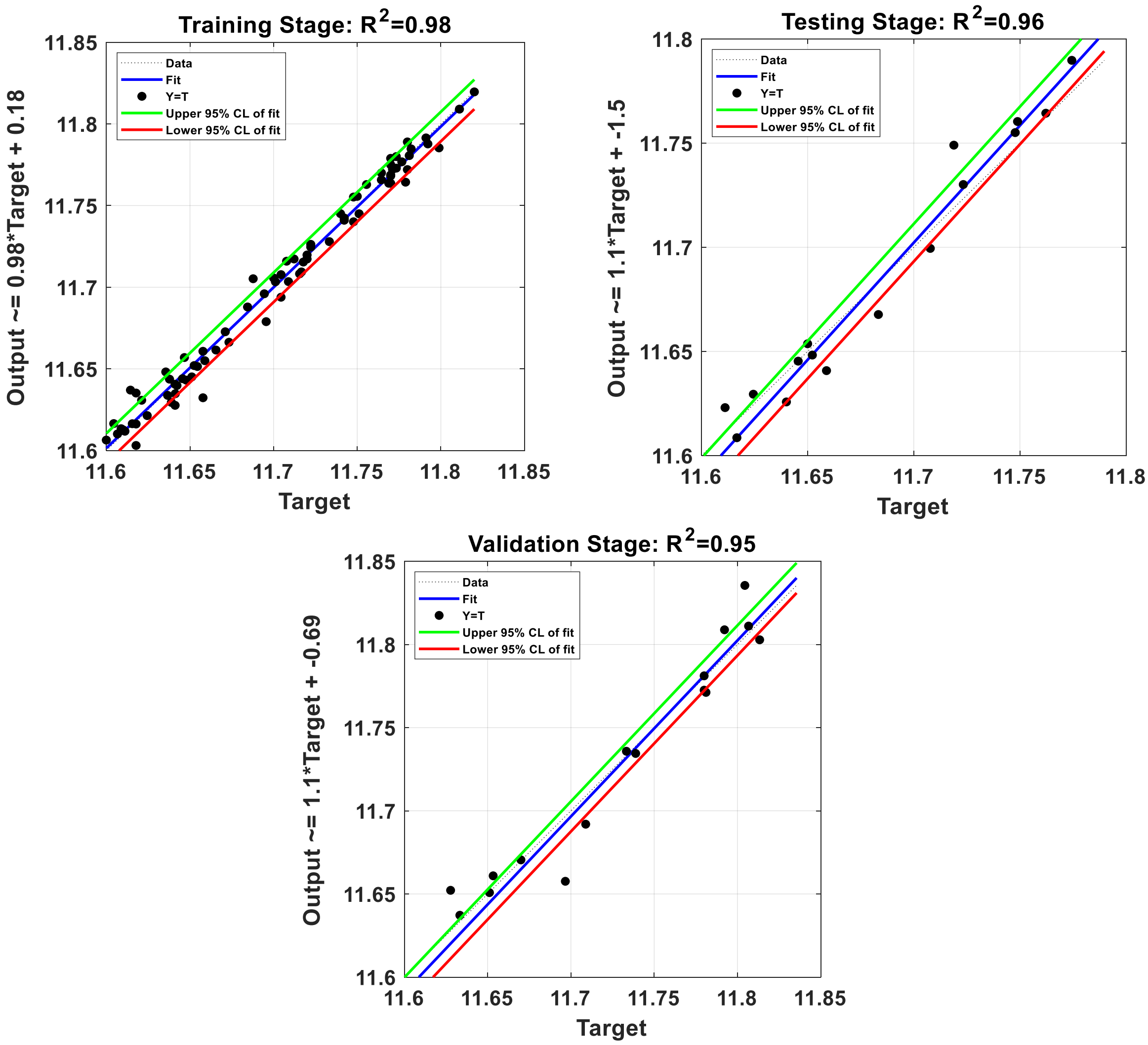


Figure 7. The performance of ANN (stand-alone) model in training, testing and validation stages.

Also, the performance of the BSA-ANN and ANN model (stand-alone) was further examined by using four different statistical indicators RMSE, MAE, MARE and CE for training, testing and validation stages. These indicators are a valuable criterion for examining the nonlinear time series as municipal water time series as presented in Table 3. According to Dawson, et al. [78], the results of these four statistical criteria indicate the ability of the models, BSA-ANN and ANN (stand-alone), to simulate accurately municipal water demand. But, the capability of BSA-ANN model for generalizing data in the validation stage is still better than ANN (stand-alone) model (e.g. the value of CE=0.979 for BSA-ANN is better than CE=0.931 for ANN (stand-alone) model.

Table 3 Performance evaluation for validation data stage.

Model	Data Stage	RMSE	MAE	MARE	CE
	Training	0.0091	0.0075	0.00064	0.999
BSA-ANN	Testing	0.0090	0.0079	0.00044	0.972
	Validation	0.0099	0.0071	0.00040	0.979
ANN	Training	0.0078	0.0058	0.00049	1.0

(stand-alone)	Testing	0.0138	0.0112	0.00063	0.935
	Validation	0.0181	0.0129	0.00072	0.931

Furthermore, a graphical test was utilised to examine the capability of the combined model to generalize water data time series in the validation stage. Figure 8 presents the observed water data in blue colour and predicted water data by BSA-ANN and ANN (stand-alone) in red and black colour, respectively. It can be noticed that the predicted data by BSA-ANN follow the trend and periodicity of the observed data, and it is very close to the observed data based on the scale of error better than data that predicted by ANN (stand-alone). Therefore, these results support the generalization capability of the combined model to forecast the municipal water time series compare with the ANN (stand-alone) model.

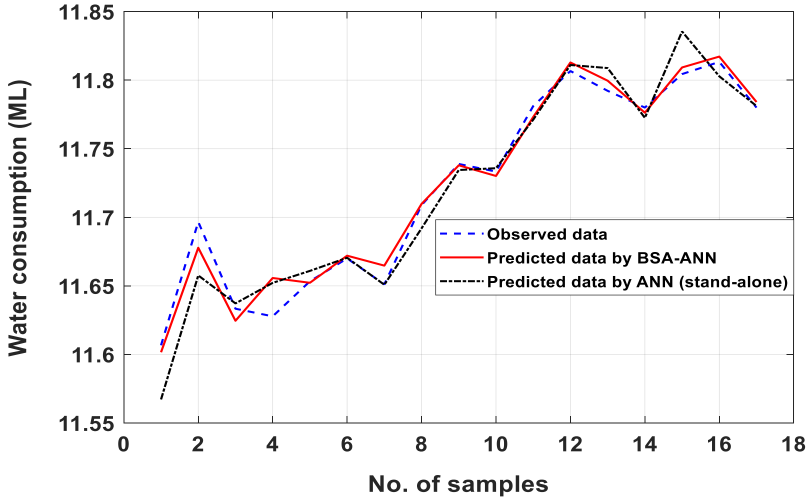


Figure 8. Presents the comparison between observed and predicted data for BSA-ANN and ANN (stand-alone) for the validation stage.

Moreover, Kolmogorov-Smirnov and Shapiro-Wilk tests agree that the residual data are normally distributed base on the significant values. In addition, the residual data are stationary based on ADF and KPSS tests. Accordingly, the values of residual data and its pattern distribution confirm the capabilities of the combined model.

Based on the above outcomes of statistical criteria, data analysis and graphical test, it can be concluded that: 1) data pre-processing techniques have been applied successfully for enhancing the quality of the data and to choose the best model input scenario. 2) BSA-ANN algorithm is more efficient and accurate than CSA-ANN algorithm, based on the fitness function value (RMSE), to locate the optimum hyperparameters of the ANN model. 3) The hybrid model BSA-ANN can accurately generalise data in validation stage compared with the ANN (stand-alone) model based on several statistical criteria. 4) The combined technique, data pre-processing and BSA-ANN algorithm, has proven to be robust for the prediction of water demand with less error, in relation to previous water consumption. 5) Using metaheuristic algorithms to detect best hyperparameters of the ANN method and comparing the outcomes of the hybrid technique with the results of the ANN (stand-alone) model leads to increasing the validation of the proposed methodology and reduce the uncertainty.

Finally, this study highlights the importance and suitability of data pre-processing and hybrid model in predicting medium-term urban water demand for the city that suffers from variability in climate and socio-economic factors such the Gauteng province. Rand Water can take benefit from the outcomes of this research to evolve effective plans for optimised system operation and ensure balancing between water delivered and need under good quality and sufficient pressure. Also, this combined technique was considered all the factors that affect water demand include socio-economic, strategic, demographic and climatic. So, it is recommended to be applied in different cities that suffer from the impact of the same factors.

5. Conclusion

In this manuscript, the performance of novel combined models that include pre-treatment signal, mutual information and BSA-ANN technique were assessed to estimate monthly municipal water needed based on previous water consumption. Historical data of monthly water consumption over ten years from the Gauteng province, South Africa, was utilised to build and evaluate the predictive model developed. The outcomes show that data pre-processing is a crucial step to enhance the quality of the data before feeding it into the model by denoising time series and selecting the best scenario of model input. Also, the hybrid BSA-ANN algorithm can be successfully applied to select optimum ANN hyperparameters, and it outperforms CSA-ANN algorithm based on fitness function (RMSE). In addition, the ANN model (stand-alone) was used to decrease the uncertainty by validating the outcomes of the hybrid model (BSA-ANN). Moreover, the results confirm the appropriateness of the combined model to forecast water demand depend on historical water consumption of a city under variability in climate and socio-economic factors such the Gauteng province. The advantages of the proposed methodology are: easy to implemented, high accuracy with less uncertainty, time-saving, and applicable when the climate and socio-economic factors are missing (i.e. lost the information of factors that drive water demand). Hence, these results can accurately inform Rand Water (i.e. its decision-makers and managers), helping this water utility company to better manage the existing municipal water system and to better plan for extensions in response to the increasing consumption, which would lead to better service and better management of resources in the Gauteng province. Therefore, taking into consideration all the benefits mentioned before, we recommend that additional studies are conducted in other regions with similar or different climatic and socio-economic factors, or regions that lack climatic and socio-economic factors but have reliable water consumption data. Also, based on the outputs of the current study, we recommend exploring the use of different techniques of data pre-processing and several hybrid models in the simulation of municipal water demand depend on historical water consumption for other cities in the world due to there is no global method that surpasses all the models for prediction water demand.

Acknowledgements: Authors are grateful the Rand Water company for providing the historical municipal water data for this study.

Author Contributions: Each of the authors contributed to the design, analysis, and writing of the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farhan, S.L.; Abdelmonem, M.G.; Nasar, Z.A. The Urban Transformation of Traditional City Centres: Holy Karbala as a Case Study. *International Journal of Architectural Research: ArchNet-IJAR* **2018**, *12*, 53-67, doi:10.26687/archnet-ijar.v12i3.1625.
2. Ebrahim Banihabib, M.; Mousavi-Mirkalaei, P. Extended linear and non-linear auto-regressive models for forecasting the urban water consumption of a fast-growing city in an arid region. *Sustainable Cities and Society* **2019**, *48*, doi:10.1016/j.scs.2019.101585.
3. Farhan, S.L.; Hashim, I.A.J.; Naji, A.A. The Sustainable House: Comparative Analysis of Houses in Al Kut Neighborhoods-Iraq. In Proceedings of 2019 12th International Conference on Developments in eSystems Engineering (DeSE); pp. 1031-1036.
4. Farhan, S.L.; Jasim, I.A.; Al-Mamoori, S.K. The Transformation of The City of Najaf, Iraq: Analysis, Reality and Future Prospects. *Journal of Urban Regeneration and Renewal* **2019**, *13*, 1-12.

5. Zubaidi, S.L.; Kot, P.; Hashim, K.; Alkhaddar, R.; Abdellatif, M.; Muhsin, Y.R. Using LARS –WG model for prediction of temperature in Columbia City, USA. *IOP Conference Series: Materials Science and Engineering* **2019**, *584*, 1-9, doi:10.1088/1757-899x/584/1/012026.
6. Osman, Y.Z.; Abdellatif, M.; Al-Ansari, N.; Knutsson, S.; Jawad, S. Climate Change and Future Precipitation in An Arid Environment of The Middle East: Case Study of Iraq. *Journal of Environmental Hydrology* **2017**, *25*, 1-18.
7. Zubaidi, S.L.; Al-Bugharbee, H.; Muhsen, Y.R.; Hashim, K.; Alkhaddar, R.M.; Hmeesh, W.H. The Prediction of Municipal Water Demand in Iraq: A Case Study of Baghdad Governorate. In Proceedings of 2019 12th International Conference on Developments in eSystems Engineering (DeSE); pp. 274-277.
8. Hashim, K.S.; Al-Saati, N.H.; Alquzweeni, S.S.; Zubaidi, S.L.; Kot, P.; Kraidi, L.; Hussein, A.H.; Alkhaddar, R.; Shaw, A.; Alwash, R. Decolourization of Dye Solutions by Electrocoagulation: An Investigation of The Effect of Operational Parameters. *IOP Conference Series: Materials Science and Engineering* **2019**, *584*, 1-8, doi:10.1088/1757-899x/584/1/012024.
9. Hashim, K.S.; Hussein, A.H.; Zubaidi, S.L.; Kot, P.; Kraidi, L.; Alkhaddar, R.; Shaw, A.; Alwash, R. Effect of Initial Ph Value on The Removal of Reactive Black Dye from Water by Electrocoagulation (EC) Method. *Journal of Physics: Conference Series* **2019**, *1294*, 1-6, doi:10.1088/1742-6596/1294/7/072017.
10. Ashoori, N.; Dzombak, D.A.; Small, M.J. Identifying water price and population criteria for meeting future urban water demand targets. *Journal of Hydrology* **2017**, *555*, 547-556, doi:10.1016/j.jhydrol.2017.10.047.
11. Toth, E.; Bragalli, C.; Neri, M. Assessing the significance of tourism and climate on residential water demand: Panel-data analysis and non-linear modelling of monthly water consumptions. *Environmental Modelling & Software* **2018**, *103*, 52-61, doi:10.1016/j.envsoft.2018.01.011.
12. Hashim, K.S.; Kot, P.; Zubaidi, S.L.; Alwash, R.; Al Khaddar, R.; Shaw, A.; Al-Jumeily, D.; Aljefery, M.H. Energy Efficient Electrocoagulation Using Baffle-Plates Electrodes for Efficient Escherichia Coli Removal from Wastewater. *Journal of Water Process Engineering* **2020**, *33*, 1-7, doi:10.1016/j.jwpe.2019.101079.
13. Msiza, I.S.; Nelwamondo, F.V.; Marwala, T. Water demand forecasting using multi-layer perceptron and radial basis functions. In Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007.
14. Kusangaya, S.; Warburton Toucher, M.L.; van Garderen, E.A. Evaluation of uncertainty in capturing the spatial variability and magnitudes of extreme hydrological events for the uMngeni catchment, South Africa. *Journal of Hydrology* **2018**, *557*, 931-946, doi:10.1016/j.jhydrol.2018.01.017.
15. House-Peters, L.A.; Chang, H. Urban water demand modeling: Review of concepts, methods, and organizing principles. *Water Resources Research* **2011**, *47*, doi:10.1029/2010wr009624.
16. Donkor, E.A.; Mazzuchi, T.H.; Soyer, R.; Roberson, J.A. Urban water demand forecasting: review of methods and models. *Journal of Water Resources Planing and Management* **2014**, *140*, 146-159, doi:10.1061/(ASCE)WR.1943-5452.
17. Ghalehkhondabi, I.; Ardjmand, E.; Young, W.A., 2nd; Weckman, G.R. Water demand forecasting: review of soft computing methods. *Environ Monit Assess* **2017**, *189*, 313, doi:10.1007/s10661-017-6030-3.
18. de Souza Groppo, G.; Costa, M.A.; Libânio, M. Predicting water demand: a review of the methods employed and future possibilities. *Water Supply* **2019**, 10.2166/ws.2019.122, doi:10.2166/ws.2019.122.
19. Gato, S.; Jayasuriya, N.; Roberts, P. Forecasting residential water demand: case study. *JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT* **2007**, *133*, 309-319, doi:10.1061/(ASCE)0733-9496(2007)133:4(309

10.1061//ASCE/0733-9496/2007/133:4/309.

20. Gato, S.; Jayasuriya, N.; Roberts, P. Temperature and rainfall thresholds for base use urban water demand modelling. *Journal of Hydrology* **2007**, *337*, 364-376, doi:10.1016/j.jhydrol.2007.02.014.
21. Zubaidi, S.L.; Dooley, J.; Alkhaddar, R.M.; Abdellatif, M.; Al-Bugharbee, H.; Ortega-Martorell, S. A Novel approach for predicting monthly water demand by combining singular spectrum analysis with neural networks. *Journal of Hydrology* **2018**, *561*, 136-145, doi:10.1016/j.jhydrol.2018.03.047.
22. Boyd, G.; Na, D.; Li, Z.; Snowling, S.; Zhang, Q.; Zhou, P. Influent Forecasting for Wastewater Treatment Plants in North America. *Sustainability* **2019**, *11*, doi:10.3390/su11061764.
23. Zhang, Q.; Li, Z.; Snowling, S.; Siam, A.; El-Dakhakhni, W. Predictive models for wastewater flow forecasting based on time series analysis and artificial neural network. *Water Sci Technol* **2019**, *80*, 243-253, doi:10.2166/wst.2019.263.
24. Seo, Y.; Kwon, S.; Choi, Y. Short-Term Water Demand Forecasting Model Combining Variational Mode Decomposition and Extreme Learning Machine. *Hydrology* **2018**, *5*, doi:10.3390/hydrology5040054.
25. Shabani, S.; Candelieri, A.; Archetti, F.; Naser, G. Gene Expression Programming Coupled with Unsupervised Learning: A Two-Stage Learning Process in Multi-Scale, Short-Term Water Demand Forecasts. *Water* **2018**, *10*, doi:10.3390/w10020142.
26. Polomˇcic´, D.; Gligoric´, Z.; Bajic´, D.; Cvijovic´, C.e. A Hybrid Model for Forecasting Groundwater Levels Based on Fuzzy C-Mean Clustering and Singular Spectrum Analysis. *Water* **2017**, *9*, doi:10.3390/w9070541.
27. Li, X.; Li, Z.; Huang, W.; Zhou, P. Performance of statistical and machine learning ensembles for daily temperature downscaling. *Theoretical and Applied Climatology* **2020**, *140*, 571-588, doi:10.1007/s00704-020-03098-3.
28. Seo, Y.; Kim, S.; Singh, V.P. Comparison of different heuristic and decomposition techniques for river stage modeling. *Environ Monit Assess* **2018**, *190*, 392, doi:10.1007/s10661-018-6768-2.
29. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2554–2558, doi:doi:10.1073/pnas.79.8.2554.
30. Hameed, M.; Sharqi, S.S.; Yaseen, Z.M.; Afan, H.A.; Hussain, A.; Elshafie, A. Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications* **2016**, *28*, 893-905, doi:10.1007/s00521-016-2404-7.
31. Bayatvarkeshi, M.; Mohammadi, K.; Kisi, O.; Fasihi, R. A new wavelet conjunction approach for estimation of relative humidity: wavelet principal component analysis combined with ANN. *Neural Computing and Applications* **2018**, 10.1007/s00521-018-3916-0, doi:10.1007/s00521-018-3916-0.
32. Mouatadid, S.; Adamowski, J. Using extreme learning machines for short-term urban water demand forecasting. *Urban Water Journal* **2016**, *14*, 630-638, doi:10.1080/1573062x.2016.1236133.
33. Guo, G.; Liu, S.; Wu, Y.; Li, J.; Zhou, R.; Zhu, X. Short-Term Water Demand Forecast Based on Deep Learning Method. *Journal of Water Resources Planning and Management* **2018**, *144*, doi:10.1061/(asce)wr.1943-5452.0000992.
34. Gharghan, S.K.; Nordin, R.; Ismail, M.; Ali, J.A. Accurate wireless sensor localization technique based on hybrid pso-ann algorithm for indoor and outdoor track cycling. *Institute of Electrical and Electronics Engineers Sensors Journal* **2016**, *16*, 529-541, doi:10.1109/jsen.2015.2483745.
35. Zubaidi, S.L.; Gharghan, S.K.; Dooley, J.; Alkhaddar, R.M.; Abdellatif, M. Short-Term Urban Water Demand Prediction Considering Weather Factors. *Water Resources Management* **2018**, *32*, 4527-4542, doi:10.1007/s11269-018-2061-y.
36. Gharghan, S.K.; Nordin, R.; Ismail, M. A Wireless Sensor Network with Soft Computing Localization Techniques for Track Cycling Applications. *Sensors (Basel)* **2016**, *16*, 1-23, doi:10.3390/s16081043.

37. Zubaidi, S.L.; Ortega-Martorell, S.; Kot, P.; Alkhaddar, R.M.; Abdellatif, M.; Gharghan, S.K.; Ahmed, M.S.; Hashim, K. A Method for Predicting Long-Term Municipal Water Demands Under Climate Change. *Water Resources Management* **2020**, *34*, 1265-1279, doi:10.1007/s11269-020-02500-z.
38. Altunkaynak, A.; Nigussie, T.A. Monthly water demand prediction using wavelet transform, first-order differencing and linear detrending techniques based on multilayer perceptron models. *Urban Water Journal* **2018**, *15*, 177-181, doi:10.1080/1573062x.2018.1424219.
39. Pacchin, E.; Gagliardi, F.; Alvisi, S.; Franchini, M. A Comparison of Short-Term Water Demand Forecasting Models. *Water Resources Management* **2019**, *33*, 1481-1497, doi:10.1007/s11269-019-02213-y.
40. Rasifaghihi, N.; Li, S.S.; Haghighat, F. Forecast of urban water consumption under the impact of climate change. *Sustainable Cities and Society* **2020**, *52*, doi:10.1016/j.scs.2019.101848.
41. Eggimann, S.; Mutzner, L.; Wani, O.; Schneider, M.Y.; Spuhler, D.; Moy de Vitry, M.; Beutler, P.; Maurer, M. The Potential of Knowing More: A Review of Data-Driven Urban Water Management. *Environ Sci Technol* **2017**, *51*, 2538-2553, doi:10.1021/acs.est.6b04267.
42. Ouyang, Q.; Lu, W. Monthly Rainfall Forecasting Using Echo State Networks Coupled with Data Preprocessing Methods. *Water Resources Management* **2017**, *32*, 659-674, doi:10.1007/s11269-017-1832-1.
43. Zhang, J.; Li, H.; Shi, X.; Hong, Y. Wavelet-Nonlinear Cointegration Prediction of Irrigation Water in the Irrigation District. *Water Resources Management* **2019**, *33*, 2941-2954, doi:10.1007/s11269-019-02270-3.
44. Ahmed, M.S.; Mohamed, A.; Khatib, T.; Shareef, H.; Homod, R.Z.; Ali, J.A. Real time optimal schedule controller for home energy management system using new binary backtracking search algorithm. *Energy and Buildings* **2017**, *138*, 215-227, doi:10.1016/j.enbuild.2016.12.052.
45. Díaz, P.; Pérez-Cisneros, M.; Cuevas, E.; Avalos, O.; Gálvez, J.; Hinojosa, S.; Zaldivar, D. An Improved Crow Search Algorithm Applied to Energy Problems. *Energies* **2018**, *11*, doi:10.3390/en11030571.
46. Abou El Ela, A.A.; El-Sehiemy, R.A.; Shaheen, A.M.; Shalaby, A.S. Application of the Crow Search Algorithm for Economic Environmental Dispatch. In Proceedings of Nineteenth International Middle East Power Systems Conference (MEPCON), Menoufia University, Egypt, 19-21 December; pp. 78-83.
47. Abdelaziz, A.Y.; Fathy, A. A novel approach based on crow search algorithm for optimal selection of conductor size in radial distribution networks. *Engineering Science and Technology, an International Journal* **2017**, *20*, 391–402, doi:10.1016/j.jestch.2017.02.004.
48. Zubaidi, S.L.; Al-Bugharbee, H.; Ortega-Martorell, S.; Gharghan, S.K.; Olier, I.; Hashim, K.S.; Al-Bdairi, N.S.S.; Kot, P. A Novel Methodology for Prediction Urban Water Demand by Wavelet Denoising and Adaptive Neuro Fuzzy Inference System Approach. *Water* **2020**, *12*, 1-17, doi:10.3390/w12061628.
49. Askarzadeh, A. A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm. *Computers & Structures* **2016**, *169*, 1-12, doi:10.1016/j.compstruc.2016.03.001.
50. Zhou, P.; Li, Z.; Snowling, S.; Baetz, B.W.; Na, D.; Boyd, G. A random forest model for inflow prediction at wastewater treatment plants. *Stochastic Environmental Research and Risk Assessment* **2019**, *33*, 1781-1792, doi:10.1007/s00477-019-01732-9.
51. Ghaith, M.; Siam, A.; Li, Z.; El-Dakhakhni, W. Hybrid Hydrological Data-Driven Approach for Daily Streamflow Forecasting. *Journal of Hydrologic Engineering* **2020**, *25*, doi:10.1061/(asce)he.1943-5584.0001866.
52. Gedefaw, M.; Hao, W.; Denghua, Y.; Girma, A.; Khamis, M.I. Variable selection methods for water demand forecasting in Ethiopia: Case study Gondar town. *Cogent Environmental Science* **2018**, *4*, doi:10.1080/23311843.2018.1537067.
53. Haque, M.M.; Rahman, A.; Hagare, D.; Chowdhury, R.K. A Comparative Assessment of Variable Selection Methods in Urban Water Demand Forecasting. *Water* **2018**, *10*, 1-15, doi:10.3390/w10040419.

54. Zhang, X.; Qiu, J.; Leng, G.; Yang, Y.; Gao, Q.; Fan, Y.; Luo, J. The Potential Utility of Satellite Soil Moisture Retrievals for Detecting Irrigation Patterns in China. *Water* **2018**, *10*, doi:10.3390/w10111505.
55. Kim, K.; Joo, H.; Han, D.; Kim, S.; Lee, T.; Kim, H.S. On Complex Network Construction of Rain Gauge Stations Considering Nonlinearity of Observed Daily Rainfall Data. *Water* **2019**, *11*, doi:10.3390/w11081578.
56. Gagliardi, F.; Alvisi, S.; Kapelan, Z.; Franchini, M. A Probabilistic Short-Term Water Demand Forecasting Model Based on the Markov Chain. *Water* **2017**, *9*, doi:10.3390/w9070507.
57. Pacchin, E.; Alvisi, S.; Franchini, M. A Short-Term Water Demand Forecasting Model Using a Moving Window on Previously Observed Data. *Water* **2017**, *9*, doi:10.3390/w9030172.
58. Bata, M.t.H.; Carriveau, R.; Ting, D.S.K. Short-Term Water Demand Forecasting Using Nonlinear Autoregressive Artificial Neural Networks. *Journal of Water Resources Planning and Management* **2020**, *146*, doi:10.1061/(asce)wr.1943-5452.0001165.
59. RW. Rand Water’s Integrated Annual Report. RW: South Africa, 2013; p 236.
60. Muringathuparambil, R.J.; Musango, J.K.; Brent, A.C.; Currie, P. Developing building typologies to examine energy efficiency in representative low cost buildings in Cape Town townships. *Sustainable Cities and Society* **2017**, *33*, 1-17, doi:10.1016/j.scs.2017.05.011.
61. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics*, Sixth ed.; Pearson Education, Inc: United States of America, 2013.
62. Al-Bugharbee, H.; Trendafilova, I. A Fault Diagnosis Methodology for Rolling Element Bearings Based on Advanced Signal Pretreatment And Autoregressive Modelling. *Journal of Sound and Vibration* **2016**, *369*, 246-265, doi:10.1016/j.jsv.2015.12.052.
63. Hassani, H.; Webster, A.; Silva, E.S.; Heravi, S. Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis. *Tourism Management* **2015**, *46*, 322-335, doi:10.1016/j.tourman.2014.07.004.
64. Golyandina, N.; Zhigljavsky, A. *Singular Spectrum Analysis for Time Series*; Springer: USA, 2013; DOI 10.1007/978-3-642-34913-3.
65. Khan, M.A.R.; Poskitt, D.S. Forecasting stochastic processes using singular spectrum analysis: Aspects of the theory and application. *International Journal of Forecasting* **2017**, *33*, 199-213, doi:10.1016/j.ijforecast.2016.01.003.
66. Zubaidi, S.L.; Kot, P.; Alkhaddar, R.M.; Abdellatif, M.; Al-Bugharbee, H. Short-Term Water Demand Prediction in Residential Complexes: Case Study in Columbia City, USA. In Proceedings of 2018 11th International Conference on Developments in eSystems Engineering (DeSE), Cambridge, United Kingdom, 2-5 Sept. 2018; pp. 31-35.
67. Kinney, J.B.; Atwal, G.S. Equitability, mutual information, and the maximal information coefficient. *Proceedings of the National Academy of Sciences* **2014**, *111*, 3354-3359.
68. Nguyen-ky, T.; Mushtaq, S.; Loch, A.; Reardon-Smith, K.; An-Vo, D.-A.; Ngo-Cong, D.; Tran-Cong, T. Predicting water allocation trade prices using a hybrid Artificial Neural Network-Bayesian modelling approach. *Journal of Hydrology* **2018**, *567*, 781-791, doi:10.1016/j.jhydrol.2017.11.049.
69. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536.
70. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* **1944**, *2*, 164-168.
71. Marquardt, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics* **1963**, *11*, 431-441.
72. Yonaba, H.; Anctil, F.; Fortin, V. Comparing Sigmoid Transfer Functions for Neural Network Multistep Ahead Streamflow Forecasting. *JOURNAL OF HYDROLOGIC ENGINEERING* **2010**, *15*, 275-283, doi:10.1061/共ASCE兲HE.1943-5584.0000188.
73. Civicioglu, P. Backtracking Search Optimization Algorithm for numerical optimization problems. *Applied Mathematics and Computation* **2013**, *219*, 8121-8144, doi:10.1016/j.amc.2013.02.017.

74. Civicioglu, P.; Besdok, E.; Gunen, M.A.; Atasever, U.H. Weighted differential evolution algorithm for numerical function optimization: a comparative study with cuckoo search, artificial bee colony, adaptive differential evolution, and backtracking search optimization algorithms. *Neural Computing and Applications* **2018**, 10.1007/s00521-018-3822-5, doi:10.1007/s00521-018-3822-5.
75. Wang, H.; Hu, Z.; Sun, Y.; Su, Q.; Xia, X. A novel modified BSA inspired by species evolution rule and simulated annealing principle for constrained engineering optimization problems. *Neural Computing and Applications* **2018**, 10.1007/s00521-017-3329-5, doi:10.1007/s00521-017-3329-5.
76. Stergiou, N. *Nonlinear analysis for human movement variability*; CRC Press: 2016.
77. Aldrich, C.; Auret, L. *Unsupervised process monitoring and fault diagnosis with machine learning methods*; Springer: 2013.
78. Dawson, C.W.; Abrahart, R.J.; See, L.M. HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software* **2007**, *22*, 1034-1052, doi:10.1016/j.envsoft.2006.06.008.