

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# **Application of Wavelet Decomposition and Phase Space Reconstruction in Urban Water Consumption Forecasting: Chaotic Approach (Case Study)**

---

Peyman Yousefi, Gholamreza Naser and Hadi Mohammadi

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.76537>

---

## **Abstract**

The forecasting of future value of water consumption in an urban area is highly complex and nonlinear. It often exhibits a high degree of spatial and temporal variability. It is a crucial factor for long-term sustainable management and improvement of the operation of urban water allocation system. This chapter will study the application of two pre-processing phase space reconstruction (PSR) and wavelet decomposition transform (WDT) methods to investigate the behavior of time series to forecast short-term water demand value of Kelowna City (BC, Canada). The research proposes two pre-process technique to improve the accuracy of the models. Artificial neural networks (ANNs), gene expression programming (GEP) and multilinear regression (MLR) methods are the tools that considered for forecasting the demand values. Evaluation of the tools is based on two steps with and without applying the pre-processing methods. Moreover, autocorrelation function (ACF) is used to calculate the lag time. Correlation dimension is used to study the chaotic behavior of the dataset. The models' relative performance is compared using three different fitness indexes; coefficient of determination (CD), root mean square error (RMSE) and mean absolute error (MAE). The results showed how pre-processing combination of WDT and PSR improved the performance of the models in forecasting short-term demand values.

**Keywords:** artificial neural network, correlation dimension, chaos, gene expression programming, Kelowna, water demand, wavelet

---

## 1. Introduction

Climate change significantly affects the water availability all around the world. This effect plays a crucial role in arid and semiarid regions. On the other hand, urban development, population growth, industrial development and economic expansion also increases water scarcity concerns critically worldwide. Therefore, the governments have to be prepared beforehand for any consequences related to water problems, especially drinking water. The efficient operation and a management plan of urban water supply requires information about the value of consumption in the future. For using different standards to simulate hydraulic constitutions in pipeline systems (to improve the reliability of the system), it is necessary to have an accurate simulation of consumption value in a specific period. In other words, "The purpose of water demand forecast is to demonstrate futuristic information available for public water suppliers as they conduct their business" [1, 2]. Short-term (e.g., less than a week), mid-term (e.g., weekly to monthly) and long-term (e.g., greater than monthly) period forecast demand values are critical for daily operations and future management of the system. Long-term urban demand forecasting (up to 25 years), mid-term (up to 2 years) and short-term values (up to 2 days) depends upon vital factors such as water supply planning, pipeline maintenance, and water distribution system optimization (e.g. optimized pumping, pipeline maintenance, minimize energy cost and water supply cost, improving system reliability and water quality), respectively [3–5]. While studies have advanced the understanding of nonlinear characteristics and high complexity of water consumption factors, further research is still required. The present accepted knowledge for these factors is still limited and depends upon (1) accurate estimation and forecast water consumption and (2) determination of type and degree of nonlinearity among the effective variables [6]. Over the past decades, two groups of deterministic and probabilistic methods have been proposed to forecast urban water demand. The deterministic approach is solely based on the input variables and their initial conditions, whereas a probabilistic model relies on modeling uncertainties and randomness of the input variables.

Given the significant challenges and complexity of probabilistic methods and the fact that pre-processing methods can provide a useful approximation to their probabilistic counterparts, this research focused on the application of pre-processing to forecast short-term consumption.

## 2. Literature review

Midterm water demand forecast helps the water management authorities to develop an integrated plan which balances supply and demand in a given period. Water stress of an area can be reduced by accurate estimation of drinking water supply demand [3, 7–9]. Moreover, management can provide water sustainability based on their experience as well as the accurate and reliable value of future demand [10].

Compared to other hydrological forecast studies (e.g., river discharge, sedimentation, rainfall, etc.) water consumption is not as influenced by the input factors as other studies do. The

most significant input variables are temperature, precipitation, and past demand values that were popular in most of the studies [11–13]. Two different types of variables affecting water demand: climatic (e.g., temperature, relative humidity, rainfall, etc.) and socioeconomic (e.g., population and income) [14]. Climatic variables can affect short-term and mid-term values while socioeconomic variables are useful for long-term forecasting [11, 15, 16]. However, a few studies investigated the impact of climatic variables on demand forecasting [17–19]. Literature enlists various deterministic and probabilistic techniques for forecasting urban drinking water demand. In general, conventional methods were prevalent for a better understanding of determinants of water demand [20–22], which consider linear relationships between effective variables and water demand, which is nonlinear. The mentioned studies are broadly categorized into two-fold: physical based and black box models. Without analyzing the physical processes, the second one applies artificial intelligence techniques (artificial neural networks, genetic programming, etc.), fuzzy-based (fuzzy logic, neuro-fuzzy, etc.), soft computing (support vector machine, etc.), and nonlinear deterministic (nonlinear local approximation, etc.) to identify the relationship between the input and output variables. Conventional regression models [3], autoregressive integrated moving average (ARIMA) [23], autoregressive integrated moving average with explanatory variable (ARIMAX) [24, 25], artificial neural networks (ANN) [9, 26–29], a combination of conventional and ANN [11, 12, 30], feedforward neural networks [12, 31], general regression neural networks [32, 33], support vector machines [14, 9, 34–37], gene expression programming [14, 38], fuzzy regression [39], neuro-fuzzy systems [40, 41], Fourier analysis [4], hybrid models (e.g. combined wavelet-ANN and wavelet-GEP) [13, 38], fuzzy cognitive map learning method [42, 43]. This research applies probabilistic ANN, GEP approach and a conventional method (MLR) to determine the performance of the methods with/without phase space reconstruction and wavelet decomposition in the case.

The chaotic nature has been addressed for various systems [44–49]. Any chaotic system is deterministic in which minor changes in the initial conditions could lead to entire different behaviors in the next periods [44]. Chaos theory was successfully used to understand the nonlinear dynamic of the system. The models that are based on chaos theory and nonlinear dynamics are a better representative of the behavior of dynamic of observed data [50]. In general, chaos theory improves the understanding of nonlinear dynamics [51]. Ng et al. applied chaos theory on noisy time series of discharge in Saugeen River (Canada) [52]. They argued that noisy time series not only increase the complications of the data but also gave high embedding dimension. Sivakumar et al. utilized the concept of nonlinear dynamic behavior to classify rivers from phase-space data reconstruction perspective [53].

Genetic programming (GP) and gene expression programming (GEP) are among the heuristic algorithms based on Darwin's evolution theory [53]. GP was employed to complete missing data in wave records and forecasting [55–57]. Aytek and Kishi used GP model to suspended sediment in the Tongue River (United States) and found GP more accurate than sediment rating curves and multiple linear regressions (MLR) [58]. Ghorbani et al. investigated the chaos theory, artificial neural network (ANN) and GEP in estimating suspended sediment in the Mississippi River (United States) [59]. GEP is superior to GP as it is more convenient to interpret the results by a GEP tree that comes along with output results. GEP also performs

better at extracting a mathematical equation which shows the relation between input and output variables [59–61]. Nasserri et al. developed a hybrid model combining the extended Kalman filter with genetic programming for monthly water demand forecasting in Tehran [62]. Shabani et al. proposed a new rationale and a novel technique in forecasting water demand using lag time to feed the determinants of water demand by the development of GEP and SVM models [14]. Yousefi et al. implemented sophisticated mathematical models to forecast water demand of City of Kelowna in monthly temporal scale. Their study assessed the performance of GEP using wavelet decomposition [38].

Among the variety of examined methods Artificial Neural Networks (ANNs), have been applied to the various period in the wide variety of hydrological issues. The main reason of ANNs frequent usage is its ability to overcome the relationship in determining the complexity of time series, even with the shortage of amount of data available to train the models. Therefore, most of the studies applicable in area of water resources demand applies ANNs to forecast short, mid and long-term demand values [13, 30, 31].

Regarding the literature review reported by Nourani et al. concluded about the dominant application of wavelet-based models [63]. Moreover, Labat notified about the improving ability of wavelet in models' performance [64]. Therefore, the application of wavelet brought researchers attention into the area such as denoising [65]; stream flow and water resources [66]; evaporation and climatic models [67]; groundwater level modeling [68]; water demand forecasting [13, 38], where in most of the mentioned studies combination of Wavelet-ANNs performed accurately over conventional models without hybrid wavelet models (e.g. ARIMA, MLR, ANN and etc.).

The objectives of this study are four-fold: (1) to investigate chaotic behavior of case data and finding the proper lag time; (2) to find the accuracy of the forecasting for one-day ahead lead time with various input combination, and (3) to study if phase space reconstruction (PSR) based on optimum embedding dimension would improve the accuracy of the models, and 4) application of wavelet decomposition by five different transform functions combined with all the mentioned models with and without PSR.

## 3. Methodology

### 3.1. Case study and data information

#### 3.1.1. Understandings

Unlike natural water resources like rainfall, the lower percentage of drinking water which is change to waste water after use, back to the cycle. Water pressure in a pipeline, water quality, supply peak consumption time, pipeline maintenance, maintenance cost, specialist and educated human resources, pipeline failure management, etc. are the variables that all of them should be under control at the same time. Also, to develop an integrated long-term plan, availability of resources is crucial. Therefore, knowing about the value of consumption in a

specific period is the first step for any management plan beyond urban drinking water supply and allocation. This chapter investigates the first step of every long-term plan development in urban drinking water as discussed below. Water utility management needs drinking water long-term forecasted values in several terms. (1) water distribution network design; (2) supply and consumption management; (3) efficient application of distribution network; (4) pipeline pressure management; (5) network development; (6) optimizing the cost of water supply and network maintenance.

### 3.1.2. Study area

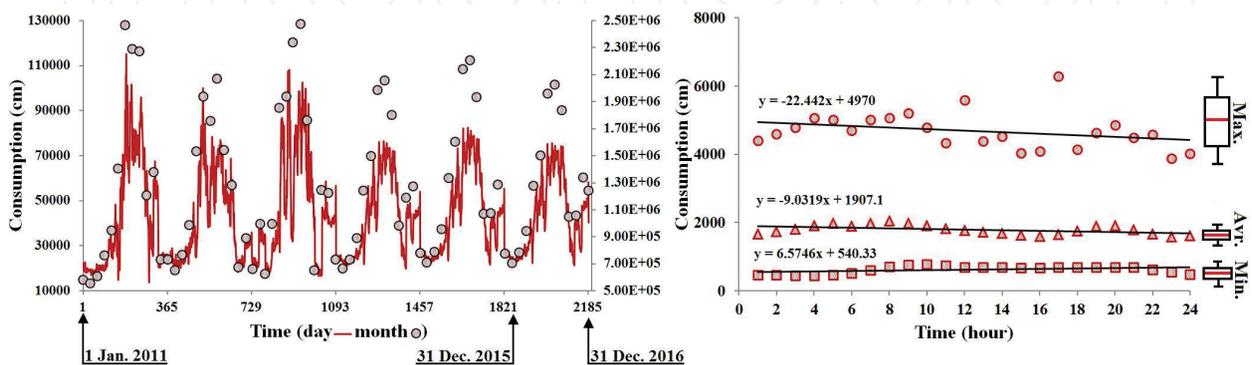
The present research selected Water consumption of the City of Kelowna (BC, Canada) as the test case. The city of Kelowna water utility provides services for approximately 65,000 residents. Poplar Point, Eldorado, Cedar Creek and Swick Road pump stations cover services for 99% of the population of the area [69]. However, few areas in the boundary are named as “Future City” where does not contain any population yet, land development plan shows water servicing is considered in the area. Monitoring of water quality, the operation of the pumps, water level in reservoirs, and pipeline pressure are conducted by the use of Supervisory Control and Data Acquisition Software (SCADA).

### 3.1.3. Review of data records

Hourly water demand for the above-mentioned stations has been made available by the city utility of Kelowna. The data used 6 years (approximately 52,464 hourly consumption) starting from January 1st, 2011 to 30th December 2016. **Figure 1** shows the variation of daily and monthly water demand and the consumption pattern. Concerning the 6 years water demand samples of daily scale (2186 points), the first 5 years (1882 points) are used for calibrating the models and the last year (365 points – 2016) is considered as the test period. **Table 1** shows the characteristics of the dataset in the test case.

## 3.2. Phase space reconstruction (PSR)

Given a set of physical variables and their interactions, the dynamics of a system (e.g., water consumption) can be defined by a single point moving on a trajectory, where each of its points



**Figure 1.** Time series plot of (a) daily water demand; (b) average of the consumption pattern in 24 h within 6 years.

Property	Number of Data	Max. Value*	Min. Value*	Average*	Standard deviation*	Coefficient of variation	Skew	Kurtosis
Data	2186	114597.2	14,124	43046.4	20074.5	0.46	0.73	-0.38

**Table 1.** Statistics of water consumption of Kelowna City in different temporal resolutions (\*m<sup>3</sup>).

represents a state of the system. The lag-embedding method reconstructs phase-space from a univariate or multivariate time series generated by a deterministic dynamic system [70]. The underlying dynamics can be studied by building an  $m$ -dimensional space  $X_t$  defined by [49]:

$$X_t = \{x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(m-1)\tau}\}, t = 1, 2, \dots, N \quad (1)$$

where  $X_t$  is a vector of the observed data of  $\{x_t\}$   $t = 1, \dots, N$ ,  $N$  is the total number of observed data,  $\tau$  is the lag time, and  $m$  is embedding dimension. The embedding dimension ( $m$ ) is typically in the range of 1–10 [53, 54]. The lag-embedding method is sensitive to both embedding parameters of  $\tau$  and  $m$ . Average mutual information (AMI) and autocorrelation function (ACF) are the two well-known methods for estimating the lag time [71, 72]. More details about ACF and different functions are available at [73].

### 3.3. Correlation dimension (Chaos investigation)

Correlation dimension is a nonlinear measure of the correlation between pairs lying on the attractor. The dimension of a system reveals the number of effective variables in the system. Kermani (2016) classified different dimensions in a system as topological, Hausdorff, box counting, point-wise, and correlation dimension. These dimensions are nearly equal in chaotic systems [52, 74]. This research employed correlation dimension, as it is a lower bound measure of the fractal dimension [59, 74]. For time series whose underlying dynamics is chaotic, the correlation dimension gets a finite fractional value, whereas it is infinite for stochastic systems. The later does not saturate to a specific amount of correlation exponent [75]. For an  $m$ -dimensional phase-space, the correlation function,  $C_m(r)$ , is defined as the fraction of states closer than  $r$  [76].

$$C_m(r) = \lim_{N_p \rightarrow \infty} \frac{2}{(N_p - w)(N_p - w - 1)} \sum_{i=1}^{N_p} \sum_{j=i+1+w}^{N_p} H(r - |X_i - X_j|) \quad (2)$$

where  $H$  is the Heaviside step function,  $X_i$  is the  $i$ <sub>th</sub> state vector,  $N_p$  is the number of points on the reconstructed attractor,  $r$  is the radius of a sphere with the content of  $X_i$  or  $X_j$ . The Theiler window ( $w$ ) is the correction needed to avoid spurious results due to temporal correlations instead of dynamical ones.  $C_m(r)$  is proportional to  $r$  for stochastic time series, whereas for chaotic time series it scales with  $r$  as:

$$C_m(r) \propto r^{c_e} \quad (3)$$

where  $c_e$  is correlation exponent defined by:

$$c_e = \lim_{r \rightarrow 0} \frac{\ln C_m(r)}{\ln r}. \quad (4)$$

The parameters  $m$  and  $C_e$  can be determined as the slopes of the lines when plotted  $C_m(r)$  against  $r$  in logarithmic scale. In a deterministic system,  $C_e$  increases by increasing  $m$  until eventually remaining unchanged. The correlation dimension of time series is defined as the specific value of  $m$  after which  $C_e$  remains unchanged [54, 59].

### 3.4. Artificial neural networks

When ANN is based roughly on the neural layout of the human brain and is capable of non-linear modeling processes that can classify the patterns and recognize the capabilities [77]. Regarding the ability of multilayer perceptron (MLP)-ANN outperformance as a conventional ANN approaches [77, 78, 79], this research employed three-layer MLP-ANN (input, hidden and output layers) and the different number of neurons. In hidden layer, neurons are calculated by the summation of demand values ( $d_i$ ) with the given weight for each value ( $w_{ij}$ ) to determine the output signal as ( $u_j$ ).

$$u_j = \sum_{i=1}^t w_{ij}d_i \quad (5)$$

$$O_j = \phi(u_j - \theta_j) \quad (6)$$

where  $\phi$  is the transfer function and  $\theta$  is a threshold limit [80, 81, 82]. Among various transfer functions (e.g., sigmoid shape, piecewise, step, linear and non-linear functions), the logistic sigmoid and Purelin (linear) transfer functions. Regarding the large number of input variables in the present study, no transfer function is applied to reduce the computationally demanding. While, the logistic sigmoid and Purelin transfer functions that are commonly used in literature [79, 81, 82] are provided at the output and hidden layers, respectively (further details about the bias and transfer functions are available at [79, 81]). Feed-forward multi linear perceptron is employed in this study containing input, hidden and output layers. The number of neurons in the input layer varies from 1 to 10 (without decomposition) and from 4 to 24 (with decomposition). Moreover, the neurons of the layers are connected with the neurons in the next layer by weights. Also, to consider all optimal solutions with the highest probable accuracy, this study investigated the number of HLN from 1 to 20 in 1 to 200 epochs.

### 3.5. Gene expression programming

Evolutionary computation has received significant attention among researchers for studying complex engineering systems. Genetic algorithm (GA), genetic programming (GP), and gene expression programming (GEP) were inspired by Darwin's theory of evolution [60, 61]. GEP defines an algorithm and equation which shows the relation between input and output variables. GA and GP rely on a string of numbers with defined length called "chromosomes", while GEP employs a set of nonlinear entities with different shapes and sizes, "expression/

parse trees". The expression tree accommodates the ease of a GA solution as well as the capability of accepting the nonlinear/complex behavior in a typical GP solution. The chromosome can have one or more genes of equal length. A gene represents a set of symbols containing two parts; a head which has functions and terminals and a tail which only has terminals. Initiating with the random generation of chromosomes, GEP is followed by different applications of genetic operators like replication, recombination, mutation, etc. The terminating condition for developing GEP depends upon the selection of maximum fitness. This research applied 30 chromosomes, eight head sizes, three genes, and arithmetic operators of  $\{+, -, \times, \div, x, x^2, \sqrt{x}\}$ .

### 3.6. Multilinear regression

When MLR corresponds to a linear combination of the components of multiple signals  $x$  (e.g. recorded discharge, lag time discharge, or combination of both) to a single output signal  $y$  (Demand) by:

$$y = b + \sum_{i=0}^N a_i x_i \quad (7)$$

where  $x_i$  is the defined input (demand) and  $a_i$  is regression coefficient determined by the least square method with the residual  $r$  defined by:

$$r = y - a_1 x_1 - a_2 x_2 - \dots - b. \quad (8)$$

### 3.7. Wavelet decomposition

Commonly wavelet transforms are used for decomposition, de-noising, and compression of the time series [83]. Time series have a combination of low and high frequency which represent improved features (e.g., cyclical trends) and chaotic element, respectively [84]. Considering these frequencies, separation of low and high frequency is helpful in studying the original pattern and behavior of the time series. One of the mentioned methods is discrete wavelet transform (DWT) to separate per level of frequencies in time series. One of the common discretion ways proposed by Mallat that this study used the mentioned DTW method to separate the frequencies of the applied data [85]. The level of the decomposition shows the subseries. For example, for level 1 decomposition, the number of subseries is two. Therefore, the number of levels indicates the number of subseries plus one. Level 3 is considered as suitable decomposition level in the present study regarding the number of data (2186 day) and following Nourani et al. (2009) that offered [83]:

$$L_n = \text{int}[\log(N)]. \quad (9)$$

where  $L_n$  is the number, the level of decomposition and  $N$  is the number of used data. Thus, the proper level in this study is considered as 3. However, increasing the level number does not necessarily improve the accuracy of the models. Therefore, the original data are discretized in a high-frequency subset ( $a_3$ ) and three high frequencies as ( $d_1$ ), ( $d_2$ ) and ( $d_3$ ), where the

summation of all is equal with the value of original data. This research employed Haar, the second and fourth order Daubechies (db2, db4), and the second and fourth order Symlets (Sym2, sym4) wavelets to decompose daily water demand time series into sub-series. The software MATLAB 2015 (<https://www.mathworks.com>) was employed for the analysis.

### 3.8. Evaluation of models' performance

This research measured the models' accuracy by coefficient of determination (CD), root mean squared error (RMSE) and mean absolute error (MAE) defined as:

$$CD = \left[ \frac{\sum_{i=1}^{N_t} (O_i - \bar{O})(F_i - \bar{F})}{\left[ \sum_{i=1}^{N_t} (O_i - \bar{O})^2 \right]^{\frac{1}{2}} \left[ \sum_{i=1}^{N_t} (F_i - \bar{F})^2 \right]^{\frac{1}{2}}} \right]^2 \quad (10)$$

$$MSE = \sqrt{\frac{\sum_{i=1}^{N_t} (O_i - F_i)^2}{N_t}} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{j=1}^n |O_i - F_i| \quad (12)$$

where  $N_t$  is the number of values,  $O$  and  $F$  are the observed and forecasted values of demand, respectively.  $\bar{O}$  and  $\bar{F}$  are the mean of the observed and forecasted demand values, respectively. Note that the range of CD is between 0 and 1 with higher positive values indicate better agreement. A lower value of RMSE and MAE indicates better agreement between the observed and forecasted values.

## 4. Preliminary results

### 4.1. Phase space reconstruction and investigation of chaotic behavior

Existence of chaotic behavior in the time series is shown in **Figure 2**. However, the results are not entirely based on the proof of having chaotic behavior, as the figure only shows possible low-dimensional chaotic behavior. Theoretically, several methods are well known for investigating the chaotic behavior such as lag time calculation method (e.g., average mutual information (AMI), Autocorrelation function (ACF)), correlation dimension, largest Lyapunov exponent, etc.). This study investigates the chaotic behavior by applying ACF and correlation dimension. Having chaotic behavior allows using ACF to calculate the lag time of the time series. The value of lag time is considered as the first approach of ACF to 0 (**Figure 2**).

The results show 83-days as the lag time of the time series. Therefore, 83-day is used to design combination of inputs as phase space for the time series. In this study, the difference between 1st day and 83rd day is used as delay period for phase space reconstruction varying embedding dimensions from 1 to 10 ( $m_1: D_t; m_2: D_t, D_{t-\tau}; m_{10}: D_t, \dots, D_{t-10\tau}$ ). It should be noticed that

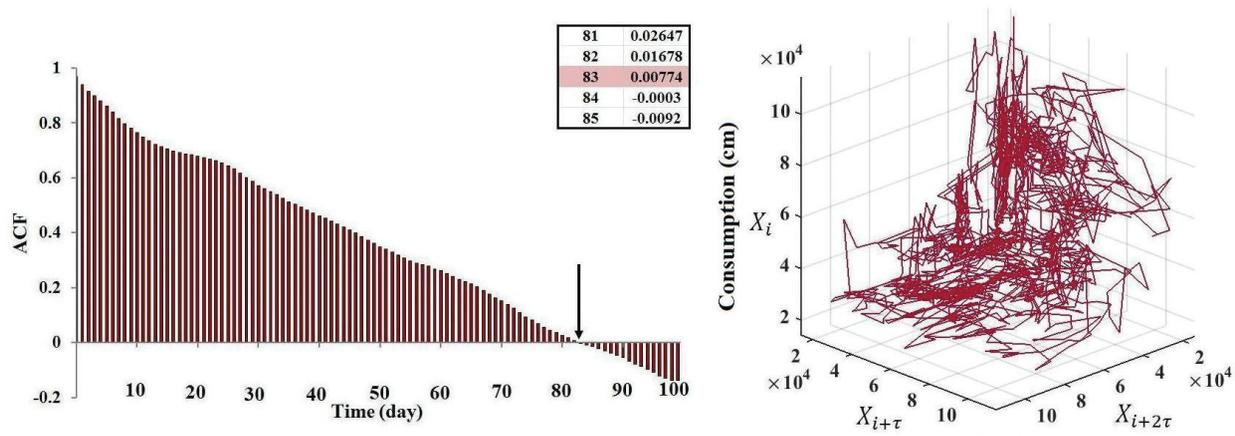


Figure 2. (a) Autocorrelation function ( $\tau$ ); (b) reconstructed phase space by ( $\tau$  and  $2\tau$  -day lag time).

several methods were introduced in literature to calculate the value of optimum embedding dimension which may be more than 10 for the used time series in this study. This study aims at showing the performance of embedding dimension and reconstructed phase space, where  $m$  is only considered 1 to 10. **Figure 2** shows the value of ACF for the demand series and reconstructed phase space ( $\tau = 83$ ). **Figure 3a** shows the relation between  $C(r)$  and  $r$  and (3b) correlation exponent by varying  $m$ . **Figure 3b** shows that the value of correlation exponent increases by  $m$  and as  $m = 17$ , the correlation exponent reaches a specific value ( $Ce = 3.41$ ). This constant value of  $Ce$  at  $m = 17$  indicates the existence of the deterministic behavior of the time series.

### 4.2. Multilinear regression

Excel 2010 was used to implement MLR model. The train period was used to derive regression coefficient from getting the value of variables in the linear equation. The availability of trained equation, helped in testifying the last year data as the test period. In the first fold, the 1-day delay was considered for  $m$  1 to 10, and second fold applied 83-day delay. **Table 2** shows the results of both MLR and PSR-MLR in the test period.

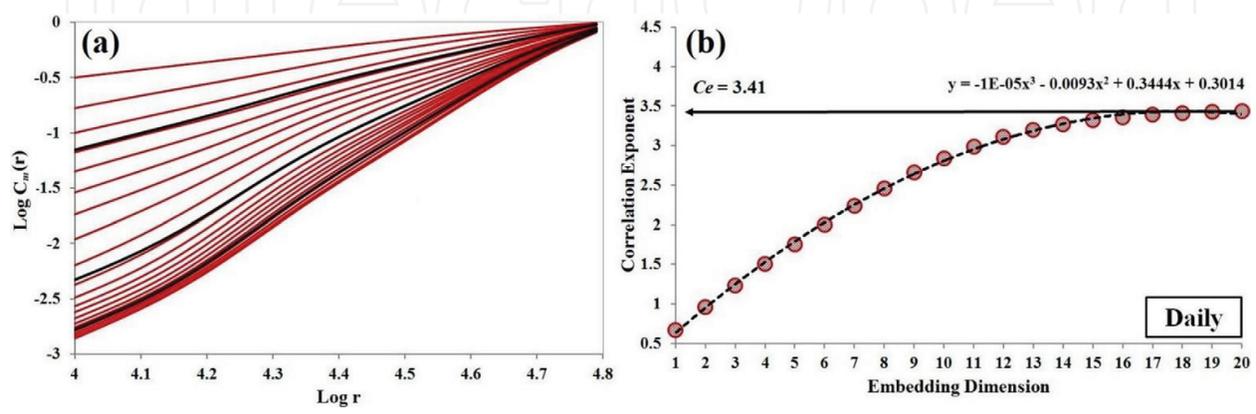


Figure 3. (a) The relation between correlation function  $C(r)$  and  $r$  by various  $m$ ; (b) Saturation of correlation dimension  $Ce$  ( $m$ ) with embedding dimensions.

MLR, $\tau = 1$				PSR-MLR, $\tau = 83$			
$m$	CD	RMSE( $m^3/day$ )	MAE	$m$	CD	RMSE( $m^3/day$ )	MAE
1	<b>0.9565</b>	<b>3642.89</b>	<b>50.42</b>	1	0.9565	3642.89	50.42
2	0.9565	3804.14	52.14	2	0.9565	3804.14	52.14
3	0.9468	14106.70	112.82	3	0.9570	5319.51	66.90
4	0.9473	13174.97	108.82	4	<b>0.9572</b>	<b>3636.34</b>	<b>51.04</b>
5	0.9505	3724.99	49.81	5	0.9568	4167.55	56.45
6	0.9503	3746.33	50.09	6	0.9569	5907.90	71.65
7	0.9503	3747.49	50.10	7	0.9565	4370.03	58.86
8	0.9493	6058.34	70.88	8	0.9566	4581.10	60.89
9	0.9505	3736.33	50.02	9	0.9566	5023.16	64.71
10	0.9506	3738.35	50.07	10	0.9566	4327.34	58.48

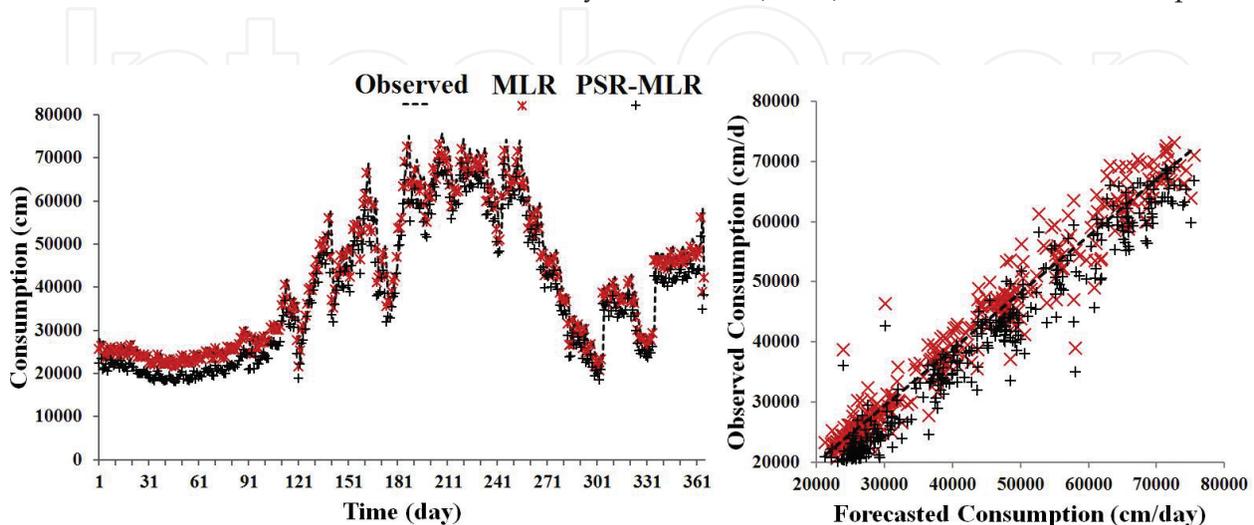
**Table 2.** Fitness values for MLR and PSR-MLR methods in different embedding dimensions (bolded lines are the most accurate values).

Statistical indices for the fitness values showed  $m = 1$  for 1-day delay and  $m = 4$  for the reconstructed phase space with the value of (CD = 0.9565, RMSE = 3642.89 and MAE = 50.42) and (CD = 0.9572, RMSE = 3636.34 and MAE = 51.04), respectively. However, the difference between the two models is not considerable, in the large value of demand in long-term this difference can come into account. **Figure 4** shows the comparison of observed and demand values. Moreover, the suggested equation for the best result by MLR is given by:

$$D_{t+1} = -0.00854D_t - 0.0366D_{t+\tau} - 0.0128D_{t+2\tau} + 0.9427D_{t+3\tau}. \quad (13)$$

### 4.3. Performance of artificial neural network

ANN is another approach to model the demand values which represented in Section 3.4. ANN's structures have different hidden layer neurons (HLN) from 1 to 20 with 200 epochs



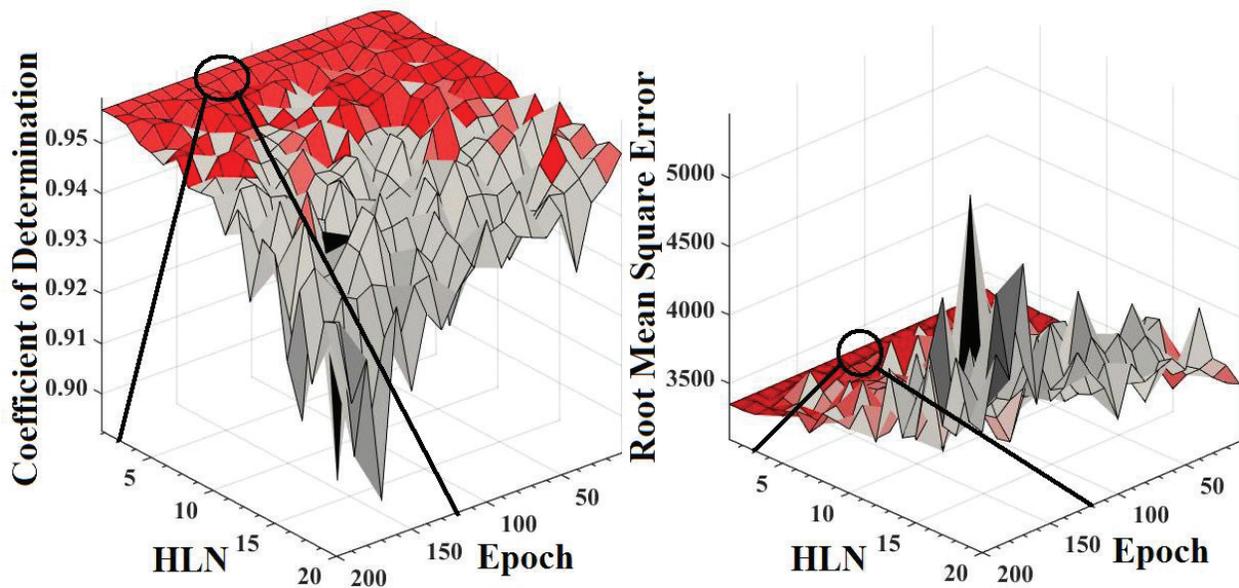
**Figure 4.** The performance of MLR and PSR-MLR in comparison with observed values.

for each model. **Table 3** represents the result of ANN for both 1-day delay and PSR values. The results in the table for each  $m$ , are extracted from the result of various HLN and epochs. **Figure 5** shows the example for selecting  $m = 3$  among ( $20 \times 200 = 4000$ ). This calculation has been done for all  $m$  from 1 to 10 for both 1-day delay and PSR. ( $4000 \times 10 \times 2 = 80,000$ ) number of calculations where the best 10 values have been selected (**Table 3**).

Selection of ANN structures are represented in **Table 3** for the test period. Statistical indices for the fitness values showed  $m = 6$  for 1-day delay and  $m = 3$  for PSR, with the values of (CD = 0.9520, RMSE = 3535.66 and MAE = 47.58) and (CD = 0.9578, RMSE = 3330.53 and

ANN, $\tau = 1$						PSR-ANN, $\tau = 83$					
$m$	Structure	Epoch	CD	RMSE*	MAE	$m$	Structure	Epoch	CD	RMSE*	MAE
1	1-5-1	110	0.9505	3611.56	48.34	1	1-6-1	150	0.9573	3369.55	47.63
2	1-3-1	140	0.9509	3602.25	48.25	2	1-4-1	20	0.9568	3369.50	47.83
3	1-16-1	20	0.9514	3554.13	48.06	3	<b>1-2-1</b>	<b>120</b>	<b>0.9578</b>	<b>3330.53</b>	<b>47.13</b>
4	1-16-1	170	0.9516	3550.99	47.90	4	1-2-1	70	0.9575	3333.67	47.25
5	1-3-1	160	0.9513	3561.33	47.98	5	1-3-1	110	0.9578	3340.36	47.15
6	<b>1-9-1</b>	<b>50</b>	<b>0.9520</b>	<b>3535.66</b>	<b>47.58</b>	6	1-3-1	40	0.9572	3340.16	47.46
7	1-3-1	100	0.9511	3563.14	48.08	7	1-3-1	100	0.9570	3348.68	47.80
8	1-8-1	20	0.9510	3570.70	47.84	8	1-2-1	150	0.9573	3333.88	47.24
9	1-4-1	200	0.9511	3566.00	47.69	9	1-2-1	140	0.9571	3338.53	47.89
10	1-3-1	100	0.9515	3546.72	47.94	10	1-4-1	10	0.9539	3518.86	49.28

**Table 3.** Fitness values for ANN and PSR-ANN in different embedding dimensions\*  $m^3$ /day). (bolded lines are the most accurate values).



**Figure 5.** The results of ANN for  $\tau = 83$  PSR by various HLN and epochs.

MAE = 47.13), respectively. Regarding the results, PSR-ANN mostly dominates in all embedding dimensions for the fitness accuracy indices. **Figure 6** shows the comparison of observed and demand values in the test period for both ANN and PSR-ANN in  $m = 6$  and 3, respectively. The results showed  $(D_t, D_{t+\tau}, D_{t+2\tau})$  as the best input combination for the models.

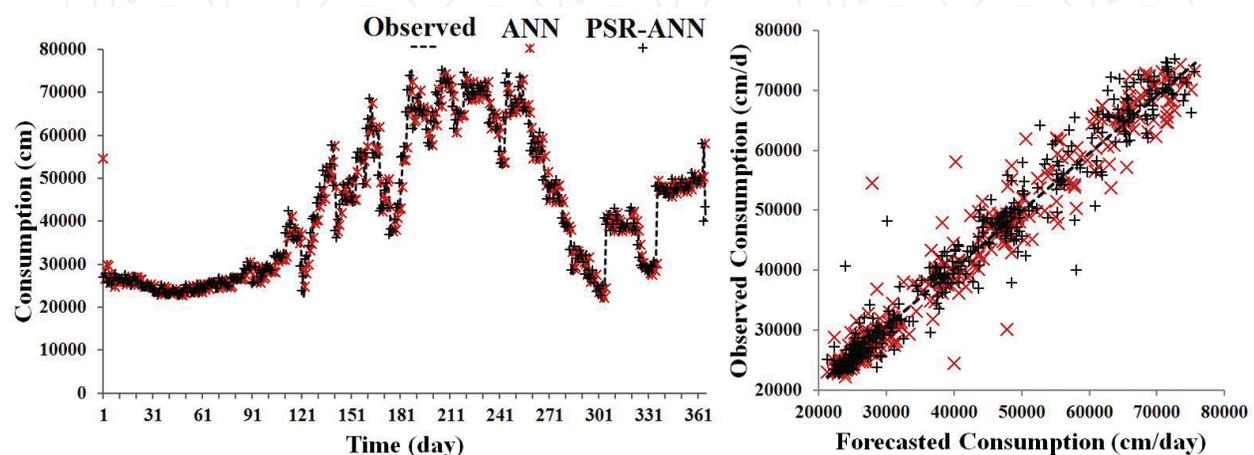
#### 4.4. Performance of gene expression programming

GEP preliminarily investigates the relationship between input and output as discussed in Section 3.5. Unlike the other models in this study, 1-day ahead is output, and various combinations of input in terms of  $m$  are considered as input variables. The arithmetic operations used in this study are  $\{+, -, \times, x, x_2, \sqrt{x}\}$ , and GEP applies them to fit the best accuracy between input and output variables. Further details of GEP initial term values are in following of [14, 38, 59] to extract the GEP model for both 1-day delay and PSR. The results are shown in the **Table 4** for the test period.

According to the **Table 4**, there is not much difference among the different  $m$ . But the difference in PSR-GEP results can be considered as a proof of sensitivity to the initial values of specific time lags where the variations of the results for different  $m$  are more than 1-day delay. There is not a significant difference in the results in this study comparing to other alternative models, especially PSR-ANN is not an advantage of GEP. However, extracting the mathematical equation through GEP is one of advantage of GEP comparing to other artificial models. As a result of given model, equation for  $m = 3$  (PSR-GEP) can calculate the demand value for 1-day ahead by:

$$D_{t+1} = 0.0529\sqrt{D_{t+\tau} + D_{t+2\tau}} + D_t - 7.0838 \tag{14}$$

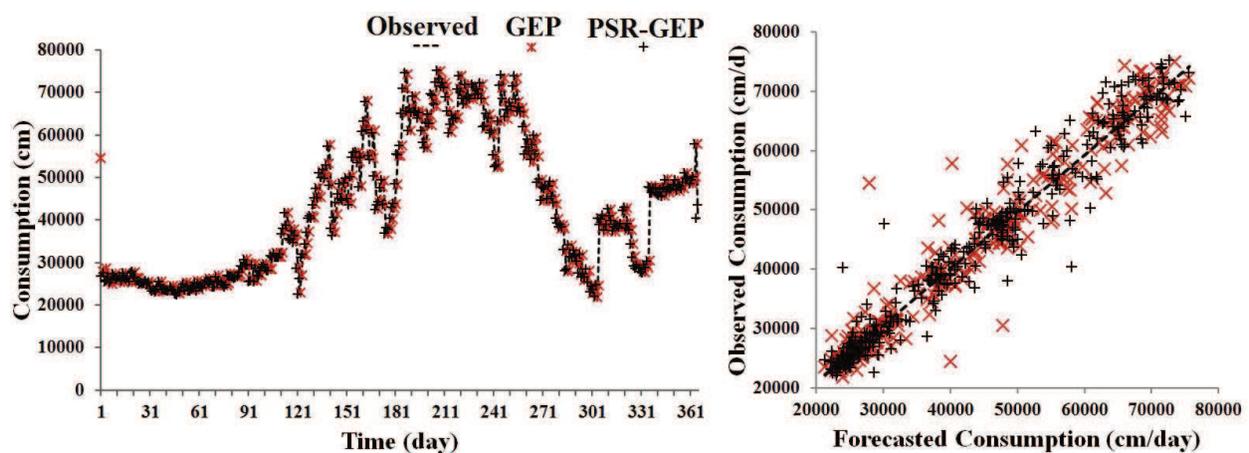
Although, variety of other arithmetic operations may have been applied here but focusing on the aim of study, only simple known operations were applied to extract the GEP equation. The results of PSR-GEP and alternative ones prove the advantage of PSR to improve the accuracy of the models. Statistical indices for the fitness values showed  $m = 2$  for 1-day delay and  $m = 3$  for the reconstructed phase space with the value of (CD = 0.9497, RMSE = 3609.82, and



**Figure 6.** The performance of ANN and PSR-ANN in comparison with observed values.

GEP, $\tau = 1$				PSR-GEP, $\tau = 83$			
$m$	CD	RMSE( $m^3/day$ )	MAE	$m$	CD	RMSE( $m^3/day$ )	MAE
1	0.9494	3621.87	48.59	1	0.9565	3363.46	48.03
2	<b>0.9497</b>	<b>3609.82</b>	<b>48.37</b>	2	0.9565	3357.00	47.82
3	0.9494	3633.87	48.42	3	<b>0.9569</b>	<b>3343.36</b>	<b>47.50</b>
4	0.9494	3637.74	48.42	4	0.9566	3359.53	47.95
5	0.9494	3639.05	48.43	5	0.9562	3372.70	48.04
6	0.9494	3619.77	48.60	6	0.9566	3359.64	48.08
7	0.9495	3630.44	48.38	7	0.9564	3365.04	47.95
8	0.9494	3634.41	48.42	8	0.9567	3353.24	47.62
9	0.9494	3628.46	48.42	9	0.9562	3370.08	48.05
10	0.9494	3631.12	48.40	10	0.9565	3356.68	47.84

**Table 4.** Fitness values for GEP and PSR-GEP in different embedding dimensions (bolded lines are the most accurate values).



**Figure 7.** The performance of GEP and PSR-GEP in comparison with observed values.

MAE = 48.37) and (CD = 0.9569, RMSE = 3343.36, and MAE = 47.50), respectively. **Figure 7** shows the comparison of observed and demand values in the test period for both GEP and PSR-GEP in  $m = 2$  and 3, respectively.

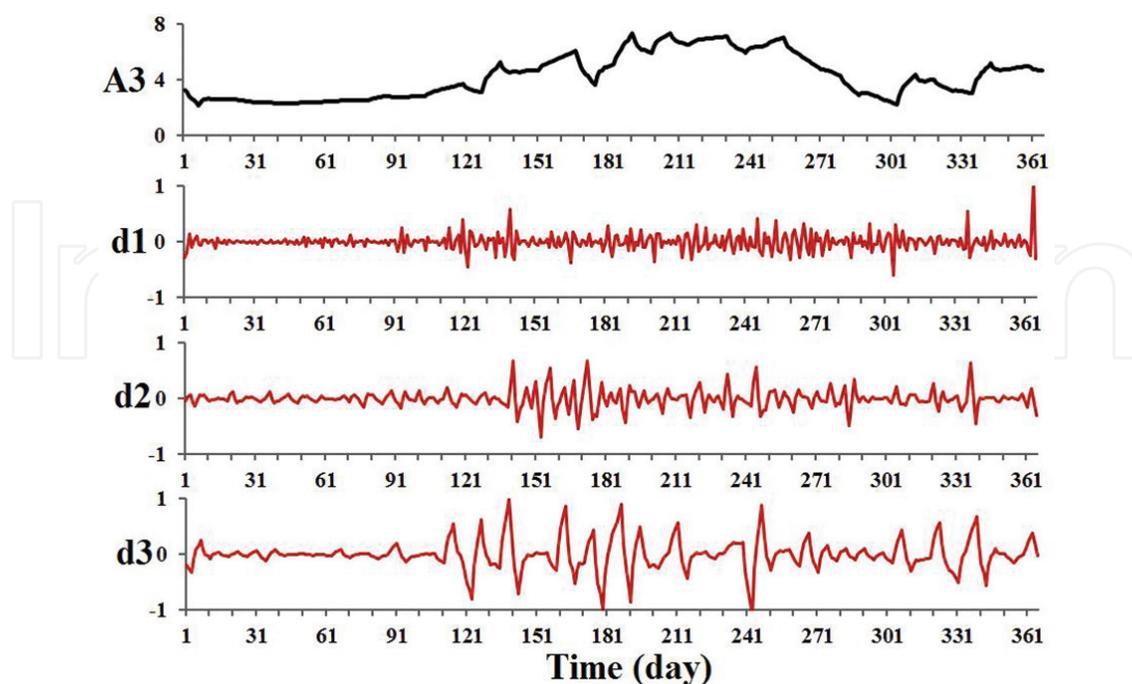
## 5. Wavelet decomposition and models' performance

The combination of models with wavelet decomposition is derived by adding the output of each wavelet to the input of the models. **Figure 8** shows the example of the decomposed values for water demand time series by db2 transform function. To discrete the demand values, five

wavelet transforms were applied (Section 3.7.). As suggested by Nourani et al. [83], 3rd level decomposition is recommended for 2186 point data.

**Table 5** indicates the results of wavelet decomposition for the selected models in the previous section. As the table highlights, db4 and db2 are the transforms which resulted in the highest accuracy in W-MLR and W-PSR-MLR, with the value of (CD = 0.9697, RMSE = 2804.44 and MAE = 42.11) and (CD = 0.9745, RMSE = 2699.83 and MAE = 43.61), respectively. After implying the decomposed inputs for MLR and PSR-MLR for result comparison improved the results in both models. Also, sym4 and db2 are the transforms which resulted in the highest accuracy in W-ANN and W-PSR-ANN, with the value of (CD = 0.9915, RMSE = 1486.21 and MAE = 30.06) and (CD = 0.9756, RMSE = 2517.24, and MAE = 41.68), respectively. Also, calculations for W-ANN and W-PSR-ANN are done with HLN 1 to 20 and epochs 1 to 200, and the mentioned results in the table are selective of the highest among them. Unlike the results of MLR, W-ANN forecasted accurately than W-PSR-ANN which is the inversion of the results of ANN and PSR-ANN. However, wavelet decomposition improved the results of W-ANN and W-PSR-ANN comparing to the alternative without decomposition (**Table 3**). Moreover, db4 and db2 are the transforms which resulted in the highest accuracy in W-GEP and W-PSR- GEP, with the value of (CD = 0.9845, RMSE = 2027.28 and MAE = 36.62) and (CD = 0.9753, RMSE = 2532.21, and MAE = 41.69), respectively. Following the results of ANN method, W-GEP forecasted accurately than W-PSR-GEP. However, wavelet decomposition improved the results of W-GEP and W-PSR-GEP comparing to the alternative without decomposition (**Table 4**).

All PSR models resulted in the highest values which used the decomposed inputs by db2 transform. It is noticeable that PSR affects the inherent of the time series which the results of performance of all models are in common about improving the accuracy. Considering this fact,



**Figure 8.** Three level DWT of daily water demand time series of Kelowna City in 2016.

Models	Fitness	Transform functions				
		haar	db2	db4	sym2	sym4
W-MLR	CD	0.9612	0.9477	<b>0.9697</b>	0.9677	0.9694
	RMSE( $m^3/day$ )	3168.62	3681.17	<b>2804.44</b>	2893.60	2816.06
	MAE	44.48	49.04	<b>42.11</b>	43.54	42.24
W-PSR-MLR	CD	0.9670	<b>0.9745</b>	0.9719	0.9745	0.9712
	RMSE( $m^3/day$ )	3008.34	<b>2699.83</b>	2811.58	2699.83	2845.69
	MAE	45.39	<b>43.61</b>	43.95	43.61	44.24
W-ANN	CD	0.9868	0.9816	0.9861	0.9856	<b>0.9915</b>
	RMSE( $m^3/day$ )	1853.11	2189.15	2136.25	1948.28	<b>1486.21</b>
	MAE	33.78	36.91	39.50	33.86	<b>30.06</b>
W-PSR-ANN	CD	0.9685	<b>0.9756</b>	0.9723	0.9752	0.9715
	RMSE( $m^3/day$ )	2867.87	<b>2517.24</b>	2677.44	2547.89	2724.56
	MAE	43.16	<b>41.68</b>	42.09	42.19	42.61
W-GEP	CD	0.9721	0.9766	<b>0.9845</b>	0.9297	0.9255
	RMSE( $m^3/day$ )	2698.16	2492.46	<b>2027.28</b>	4311.60	4429.89
	MAE	41.21	39.05	<b>36.62</b>	54.23	55.25
W-PSR-GEP	CD	0.9667	<b>0.9753</b>	0.9721	0.9748	0.9704
	RMSE( $m^3/day$ )	2937.76	<b>2532.21</b>	2689.20	2555.82	2770.80
	MAE	43.66	<b>41.69</b>	42.13	41.90	42.51

**Table 5.** Fitness values for decomposition of selection of models for the test period (bolded lines are the most accurate values).

PSR can be introduced as a pre-processing method like wavelet decomposition; however, complexity and accuracy of PSR cannot be compared with the higher result of wavelet decomposition. **Figure 9** shows the comparison of all selected models with highest accuracy (W-PSR-MLR, W-ANN, and W-GEP) in forecast of short-term water demand values.

The figure shows that the performance of W-ANN and W-GEP is better than W-PSR-MLR, while W-ANN's calculated values are more accurate than W-GEP in simulating peak points. This study eventually would suggest that these peak points are indication of critical issues related to water distribution system (pressure management, peak time demand, etc.) taking in account the performance of the models and simulations of highest and lowest values of demands. Therefore, it is recommended to evaluate models' performance in two separate parts as maximum values and minimum values along with evaluating criteria such as CD, RMSE, and MAE for the test period. The difference is not visible in **Figure 9**. Therefore focusing on **Figure 10**, it shows the performance of models by residual values in the test period.

In **Figure 10** the residual values show the remarkable difference of performance of models. W-ANN values distributed in the area of  $(-15\%, +15\%)$ , unlike other two models. W-GEP dominates over W-PSR-MLR; however, the fitness criteria values for both are very close to each other (**Tables 2 and 4**).

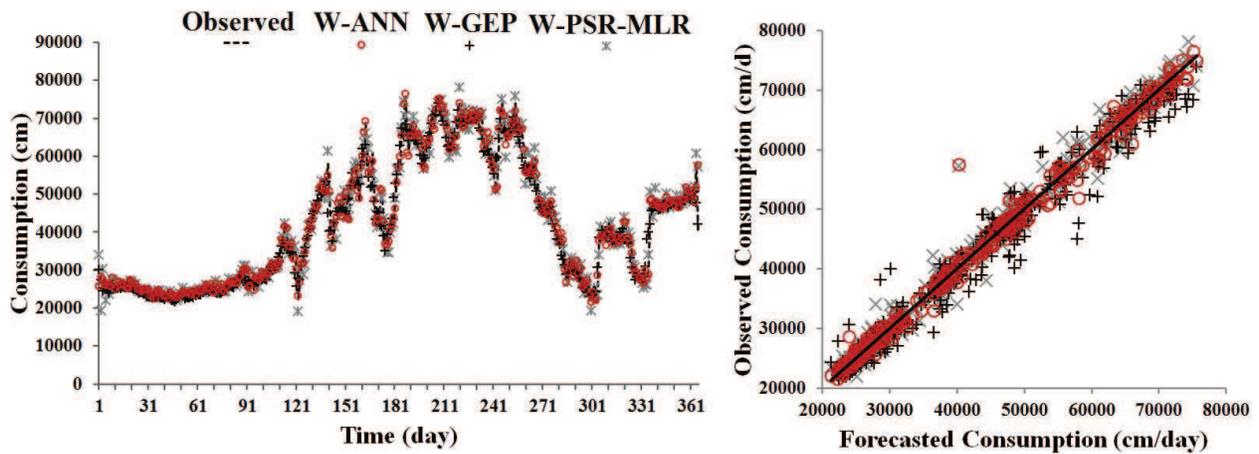


Figure 9. The performance of the W-models in comparison with observed values.

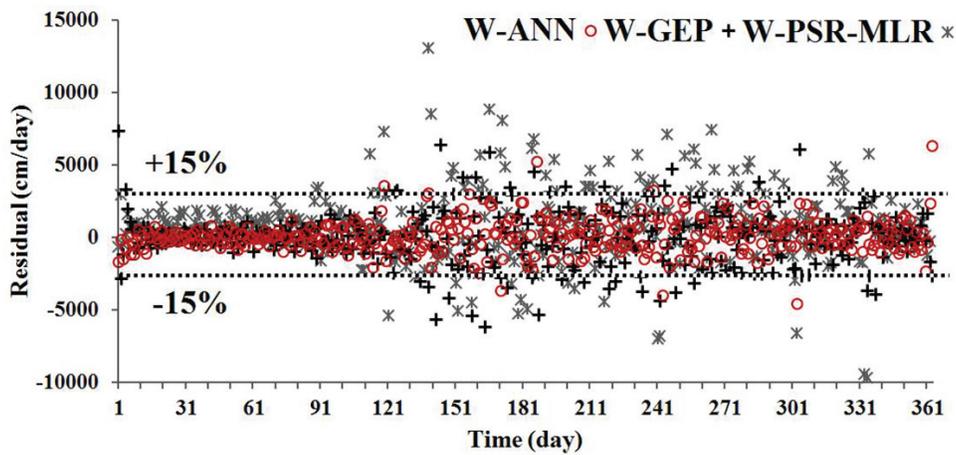


Figure 10. Residual values of the selected W-models.

This chapter presents the performance of two pre-processes methods in improving the accuracy of three models to forecast short-term urban water demand value in Kelowna City, BC, Canada. The first pre-process approach of PSR which is calculated by ACF method has improved the results of all models in this study. However, PSR does not improve the accuracy of models for entire dataset. Based on the behavior of time series, ACF or AMI (two lag time calculation methods) may have improved a non-deterministic dataset, but it seems in a chaotic dataset, PSR improves the performance of models in increasing accuracy with a proper number of embedding dimensions. Wavelet decomposition, the second pre-process method in the present study has also improved the accuracy of the models but, decomposition did not work on PSR based methods except MLR. It can be concluded that PSR and wavelet are in common with their outfits as two applicable pre-process methods. Also, PSR pre-processing is simpler than wavelet. Therefore, it is recommended to use PSR for the models. As per the results of this study it seems PSR works on a chaotic dataset which seems to be considered as disadvantage of PSR. Comparing the mentioned two pre-process methods, wavelet decomposition is significant to use, though, it is time-consuming and complex than PSR. Also, each transform functions have specific application where each of them can be used independently (e.g., seasonal, de-noising, peak points, etc.).

## 6. Conclusion

Over the past decades, hydrologists have paid attention to data-driven modeling techniques. City governments and WDS operators are always looking for an accurate estimation of water demand values not only for future but also focusing on probable failures like peak consumption and pressure values to manage the WDS pipelines. Therefore, the wide variety of modeling techniques such as artificial and evolutionary simulation methods are proposed by researchers. This chapter investigated the performance of three techniques (ANN, GEP, and MLR) in forecasting short-term water demand of Kelowna City (BC, Canada). About 6 years daily dataset was employed for training and testing the models. First 5 years were considered to train the model and the last year as the test period. All three techniques performed considerably accurate, while the focus of this chapter was on improving the accuracy of the models for the same dataset. Firstly, the model was calibrated by different input combination with 1-day lag time. Then, models were calibrated by the lag time of the data set (83-day) which was calculated by ACF method. WDT was combined with the models to capture multi-scale features of the signals by decomposing observed demand values into sub-series. Five WDT functions (haar, Db2, db4, Sym2, and sym4) were employed to decompose the dataset. The results were then compared with the MLR, ANN, and GEP when no pre-processing (PSR, WDT) was applied. The research results were accurate than PSR. WDT have also improved the accuracy of models with PSR and without PSR. However, the impact of wavelet on the models with PSR was not as considerable as without PSR. The lowest error was reported by W-ANN among all alternative models in this chapter. Regarding the improvement of all models combining WDT and PSR, it is recommended to use the method in modeling and forecasting issues, especially about the dataset that the peak points are very critical in the case. The inherent behavior of dataset (deterministic or stochastic) can affect the performance of the pre-processing methods. Therefore, behavior of datasets should be investigated before deciding to combine any pre-process methods.

## Acknowledgements

The authors would like to thank the financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada. The Okanagan Basin Water Board and the City of Kelowna are thanked for providing water consumption data.

## Author details

Peyman Yousefi, Gholamreza Naser\* and Hadi Mohammadi

\*Address all correspondence to: bahman.naser@ubc.ca

Okanagan School of Engineering, The University of British Columbia

## References

- [1] Billings RB, Jones CV. *Forecasting Urban Water Demand*. USA: American Water Works Association; 2011. ISBN: 1-58321-537-9
- [2] Ghalekhondabi I, Ardjmand E, Young WA, Weckman GR. Water demand forecasting: Review of soft computing methods. *Environmental Monitoring and Assessment*. 2017; **189**(7):313
- [3] Ghiassi M, Zimbra DK, Saidane H. Urban water demand forecasting with a dynamic artificial neural network model. *Journal of Water Resources Planning and Management*. 2008; **134**(2):138-146
- [4] Odan FK, Reis LF. Hybrid water demand forecasting model associating artificial neural network with Fourier series. *Journal of Water Resources Planning and Management*. 2012; **138**(3):245-256
- [5] IwAnek M, Kowalska B, Hawryluk E, Kondraciuk K. Distance and time of water effluence on soil surface after failure of buried water pipe. Laboratory investigations and statistical analysis. *Eksploracja i niezawodność – Maintenance and Reliability*. 2016; **18**(2):278-284
- [6] Wang W, Vrijling JK, Van Gelder PH, Ma J. Testing for nonlinearity of streamflow processes at different timescales. *Journal of Hydrology*. 2006; **322**(1-4):247-268
- [7] Jain A, Ormsbee LE. A decision support system for drought characterization and management. *Civil Engineering Systems*. 2001; **18**(2):105-140
- [8] Kame'enuei AE. *Water Demand Forecasting in the Puget Sound Region: Short and Long-Term Models* (Doctoral dissertation, University of Washington)
- [9] Herrera M, Torgo L, Izquierdo J, Pérez-García R. Predictive models for forecasting hourly urban water demand. *Journal of Hydrology*. 2010; **387**(1-2):141-150
- [10] Zhou SL, McMahon TA, Walton A, Lewis J. Forecasting operational demand for an urban water supply zone. *Journal of Hydrology*. 2002; **259**(1-4):189-202
- [11] Jain A, Varshney AK, Joshi UC. Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks. *Water Resources Management*. 2001; **15**(5):299-321
- [12] Bougadis J, Adamowski K, Diduch R. Short-term municipal water demand forecasting. *Hydrological Processes*. 2005; **19**(1):137-148
- [13] Adamowski J, Fung Chan H, Prasher SO, Ozga-Zielinski B, Sliusarieva A. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*. 2012; **48**(1)
- [14] Shabani S, Yousefi P, Adamowski J, Naser G. Intelligent soft computing models in water demand forecasting. In: *Water Stress in Plants*. 2016. Rijeka, Croatia: InTech; ISBN: 978-953-51-2621-8

- [15] Miaou SP. A class of time series urban water demand models with nonlinear climatic effects. *Water Resources Research*. 1990;**26**(2):169-178
- [16] Gato-Trinidad S, Jayasuriya N, Roberts P. Understanding urban residential end uses of water. *Water Science and Technology*. 2011;**64**(1):36-42
- [17] Zhou SL, McMahon TA, Walton A, Lewis J. Forecasting daily urban water demand: A case study of Melbourne. *Journal of Hydrology*. 2000;**236**(3-4):153-164
- [18] Mukhopadhyay A, Akber A, Al-Awadi E. Analysis of freshwater consumption patterns in the private residences of Kuwait. *Urban Water*. 2001;**3**(1-2):53-62
- [19] Dos Santos CC, Pereira Filho AJ. Water demand forecasting model for the metropolitan area of São Paulo, Brazil. *Water Resources Management*. 2014;**28**(13):4401-4414
- [20] Brekke L, Larsen MD, Ausburn M, Takaichi L. Suburban water demand modeling using stepwise regression. *American Water Works Association. Journal*. 2002 Oct 1;**94**(10):65
- [21] Polebitski AS, Palmer RN, Waddell P. Evaluating water demands under climate change and transitions in the urban environment. *Journal of Water Resources Planning and Management*. 2010;**137**(3):249-257
- [22] Lee SJ, Wentz EA, Gober P. Space-time forecasting using soft geostatistics: A case study in forecasting municipal water demand for Phoenix, Arizona. *Stochastic Environmental Research and Risk Assessment*. 2010;**24**(2):283-295
- [23] Adebisi AA, Adewumi AO, Ayo CK. Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*. 2014;**2014**(1):1-7
- [24] Young CC, Liu WC. Prediction and modelling of rainfall-runoff during typhoon events using a physically-based and artificial neural network hybrid model. *Hydrological Sciences Journal*. 2015;**60**(12):2102-2116
- [25] Young CC, Liu WC, Hsieh WL. Predicting the water level fluctuation in an alpine lake using physically based, artificial neural network, and time series forecasting models. *Mathematical Problems in Engineering*. 2015:1-11
- [26] Adamowski J, Karapataki C. Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: Evaluation of different ANN learning algorithms. *Journal of Hydrologic Engineering*. 2010;**15**(10):729-743
- [27] Cutore P, Campisano A, Kapelan Z, Modica C, Savic D. Probabilistic prediction of urban water consumption using the SCEM-UA algorithm. *Urban Water Journal*. 2008;**5**(2):125-132
- [28] Firat M, Yurdusev MA, Turan ME. Evaluation of artificial neural network techniques for municipal water consumption modeling. *Water Resources Management*. 2009;**23**(4):617-632
- [29] Jentgen L, Kidder H, Hill R, Conrad S. Energy management strategies use short-term water consumption forecasting to minimize cost of pumping operations. *American Water Works Association. Journal*. 2007;**99**(6):86-94

- [30] Jain A, Ormsbee LE. Short-term water demand forecast modeling techniques—Conventional methods versus AI. *Journal (American Water Works Association)*. 2002;**94**(7):64-72
- [31] Adamowski JF. Peak daily water demand forecast modeling using artificial neural networks. *Journal of Water Resources Planning and Management*. 2008;**134**(2):119-128
- [32] Zhou J, Yang K. General regression neural network forecasting model based on PSO algorithm in water demand. In *Knowledge Acquisition and Modeling (KAM), 2010 3rd International Symposium on 2010 Oct* (pp. 51-54). IEEE
- [33] Firat M, Turan ME, Yurdusev MA. Comparative analysis of neural network techniques for predicting water consumption time series. *Journal of Hydrology*. 2010;**384**(1-2):46-51
- [34] Msiza IS, Nelwamondo FV, Marwala T. Water demand prediction using artificial neural networks and support vector regression. *Journal of Computers*. 2008;**3**(11):1-8
- [35] Brentan BM, Luvizotto E Jr, Herrera M, Izquierdo J, Pérez-García R. Hybrid regression model for near real-time urban water demand forecasting. *Journal of Computational and Applied Mathematics*. 2017;**309**:532-541
- [36] Msiza IS, Nelwamondo FV, Marwala T. Artificial neural networks and support vector machines for water demand time series forecasting. In: *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on 2007 Oct* (pp. 638-643). IEEE
- [37] Trzęsiok M. Symulacyjna ocena jakości zagregowanych modeli zbudowanych metodą wektorów nośnych. *Studia Ekonomiczne*. 2013;**132**:115-126
- [38] Yousefi P, Shabani S, Mohammadi H, Naser G. Gene expression programming in long term water demand forecasts using wavelet decomposition. *Procedia Engineering*. 2017;**186**: 544-550
- [39] Azadeh A, Neshat N, Hamidipour H. Hybrid fuzzy regression—artificial neural network for improvement of short-term water consumption estimation and forecasting in uncertain and complex environments: Case of a large metropolitan city. *Journal of Water Resources Planning and Management*. 2011;**138**(1):71-75
- [40] Atsalakis G, Minoudaki C, Markatos N, Stamou A, Beltrao J, Panagopoulos T. Daily irrigation water demand prediction using adaptive neuro-fuzzy inferences systems (anfis). In: *Proceedings of international conference on energy, environment, Ecosystems and Sustainable Development; 2007 Jul*
- [41] Tabesh M, Dini M. Fuzzy and neuro-fuzzy models for short-term water demand forecasting in Tehran. *Iranian Journal of Science and Technology*. 2009;**33**(B1):61-77
- [42] Papageorgiou EI, Poczęta K, Laspidou C. Application of fuzzy cognitive maps to water demand prediction. In: *Fuzzy systems (FUZZ-IEEE), 2015 IEEE international conference on 2015 Aug* (pp. 1-8). IEEE
- [43] Ahmadi S, Alizadeh S, Forouzideh N, Yeh CH, Martin R, Papageorgiou E. ICLA imperialist competitive learning algorithm for fuzzy cognitive map: Application to water demand

- forecasting. In: Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on 2014 Jul, China: IEEE; (p. 1041-1048)
- [44] Lorenz EN. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*. 1963; **20**(2):130-141
- [45] Lorenz EN. Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*. 1969; **26**(4):636-646
- [46] Jayawardena AW, Lai F. Analysis and prediction of chaos in rainfall and stream flow time series. *Journal of Hydrology*. 1994; **153**(1-4):23-52
- [47] Porporato A, Ridolfi L. Nonlinear analysis of river flow time sequences. *Water Resources Research*. 1997; **33**(6):1353-1367
- [48] Krasovskaia I, Gottschalk L, Kundzewicz ZW. Dimensionality of Scandinavian river flow regimes. *Hydrological Sciences Journal*. 1999; **44**(5):705-723
- [49] Sivakumar B, Berndtsson R, Persson M. Monthly runoff prediction using phase space reconstruction. *Hydrological Sciences Journal*. 2001; **46**(3):377-387
- [50] Gan CB, Lei H. A new procedure for exploring chaotic attractors in nonlinear dynamical systems under random excitations. *Acta Mechanica Sinica*. 2011; **27**(4):593-601
- [51] Casdagli M. Nonlinear forecasting, chaos and statistics. In: *Modeling Complex Phenomena*. New York, NY: Springer; 1992. pp. 131-152
- [52] Ng WW, Panu US, Lennox WC. Chaos based analytical techniques for daily extreme hydrological observations. *Journal of Hydrology*. 2007; **342**(1-2):17-41
- [53] Sivakumar B, Jayawardena AW, Li WK. Hydrologic complexity and classification: A simple data reconstruction approach. *Hydrological Processes*. 2007; **21**(20):2713-2728
- [54] Khatibi R, Ghorbani MA, Aalami MT, Kocak K, Makarynskyy O, Makarynska D, Aalinezhad M. Dynamics of hourly sea level at Hillarys boat harbour, Western Australia: A chaos theory perspective. *Ocean Dynamics*. 2011; **61**(11):1797-1807
- [55] Kalra R, Deo MC. Genetic programming for retrieving missing information in wave records along the west coast of India. *Applied Ocean Research*. 2007; **29**(3):99-111
- [56] Ustoorikar K, Deo MC. Filling up gaps in wave data with genetic programming. *Marine Structures*. 2008; **21**(2-3):177-195
- [57] Gaur S, Deo MC. Real-time wave forecasting using genetic programming. *Ocean Engineering*. 2008; **35**(11-12):1166-1172
- [58] Aytok A, Kişi Ö. A genetic programming approach to suspended sediment modelling. *Journal of Hydrology*. 2008; **351**(3-4):288-298
- [59] Ghorbani MA, Khatibi R, Asadi H, Yousefi P. Inter-comparison of an evolutionary programming model of suspended sediment time-series with other local models. In: *Genetic*

Programming-New Approaches and Successful Applications. Rijeka, Croatia: InTech; 2012. 978-953-51-0809-2

- [60] Ferreira C. Gene expression programming in problem solving. In: *Soft Computing and Industry*. London: Springer; 2002. pp. 635-653
- [61] Ferreira C. Function finding and the creation of numerical constants in gene expression programming. In: *Advances in Soft Computing*. London: Springer; 2003. pp. 257-265
- [62] Nasser M, Moeini A, Tabesh M. Forecasting monthly urban water demand using extended Kalman filter and genetic programming. *Expert Systems with Applications*. 2011;**38**(6):7387-7395
- [63] Nourani V, Baghanam AH, Adamowski J, Kisi O. Applications of hybrid wavelet-artificial intelligence models in hydrology: A review. *Journal of Hydrology*. 2014;**514**:358-377
- [64] Labat D. Recent advances in wavelet analyses: Part 1. A review of concepts. *Journal of Hydrology*. 2005 Nov 25;**314**(1-4):275-288
- [65] Chou CM. Application of set pair analysis-based similarity forecast model and wavelet denoising for runoff forecasting. *Water*. 2014;**6**(4):912-928
- [66] Labat D. Wavelet analysis of the annual discharge records of the world's largest rivers. *Advances in Water Resources*. 2008;**31**(1):109-117
- [67] Partal T, Cigizoglu HK. Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *Journal of Hydrology*. 2008;**358**(3-4):317-331
- [68] Adamowski J, Chan HF. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*. 2011;**407**(1-4):28-40
- [69] Water Quality and Customer Care Supervisor City of Kelowna. City of Kelowna 2016 Annual Water and Filtration Exclusion Report, Report Submitted: July 31, 2017
- [70] Takens F. Detecting strange attractors in turbulence. In: *Dynamical systems and turbulence*. Warwick, Berlin, Heidelberg: Springer; 1980, 1981. pp. 366-381
- [71] Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Physical Review A*. 1986;**33**(2):1134
- [72] Holzfuss J, Mayer-Kress G. An approach to error-estimation in the application of dimension algorithms. In: *Dimensions and Entropies in Chaotic Systems*. Berlin, Heidelberg: Springer; 1986. 114-122
- [73] Dunn PF, Davis MP. *Measurement and Data Analysis for Engineering and Science*. CRC Press; 2017
- [74] Zounemat-Kermani M. Investigating chaos and nonlinear forecasting in short term and mid-term river discharge. *Water Resources Management*. 2016;**30**(5):1851-1865
- [75] Dhanya CT, Kumar DN. Nonlinear ensemble prediction of chaotic daily rainfall. *Advances in Water Resources*. 2010;**33**(3):327-347

- [76] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. In: *The Theory of Chaotic Attractors*. New York, NY: Springer; 2004. pp. 170-189
- [77] Lek S, Guégan JF. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*. 1999 Aug 17;120(2-3):65-73
- [78] Nourani V, Entezari E, Yousefi P. ANN-RBF hybrid model for spatiotemporal estimation of monthly precipitation case study: Ardabil plain. *International Journal of Applied Metaheuristic Computing (IJAMC)*. 2013;4(2):1-6
- [79] Najah A, El-Shafie A, Karim OA, El-Shafie AH. Application of artificial neural networks for water quality prediction. *Neural Computing and Applications*. 2013;22(1):187-201
- [80] Yousefi P, Naser G, Mohammadi H. Surface Water Quality Model: Impacts of Influential Variables. *Journal of Water Resources Planning and Management*. 2018 Feb 22;144(5):04018015
- [81] Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd ed. Upper Saddle River: Prentice Hall; 1999
- [82] Melesse AM, Hanley RS. Artificial neural network application for multi-ecosystem carbon flux simulation. *Ecological Modelling*. 2005;189(3-4):305-314
- [83] Nourani V, Alami MT, Aminfar MH. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence*. 2009;22(3):466-472
- [84] Zhou T, Wang F, Yang Z. Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction. *Water*. 2017 Oct 12;9(10):781
- [85] Mallat SG. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989;11(7):674-693