



EPiC Series in Engineering

Volume 3, 2018, Pages 2408–2416

HIC 2018. 13th International
Conference on Hydroinformatics



Estimating High Resolution Temporal Scale of Water Demand Time Series – Disaggregation Approach (Case Study)

Peyman Yousefi^{1*}, Gholamreza Naser^{1†}, Hadi Mohammadi^{1‡}

¹Okanagan School of Engineering, The University of British Columbia, Canada.
pe.yousefi@alumni.ubc.ca, bahman.naser@ubc.ca, hadi.mohammadi@ubc.ca

Abstract

A comprehensive understanding of water demand and its availability is essential for decision-makers to manage their resources and understand related risks effectively. Historical data play a crucial role in developing an integrated plan for management of water distribution system. The key is to provide high-resolution temporal-scale of demand data in urban areas. In the literature, many studies on water demand forecasting are available; most of them were focused on monthly-scales. Since monitoring of time series is a prolonged and costly procedure, the popularity of disaggregation methods is a most recent desirable trend. The objective of this research is to transfer low-resolution into high-resolution temporal scale using random cascade disaggregation and non-linear deterministic methods. This study defines a new technique to apply previously proposed random cascade method to disaggregate continuous data of the city of Peachland. The accuracy of the results is more than 90%. It represents a satisfactory application of the models. The proposed approach helps operators to have access to daily demand without acquiring high-resolution temporal scale values. Although the disaggregated values may not be precisely equal with observed values, it offers a practical solution for the low equipped WDS and leads to lesser number of drinking water-related problems.

Keywords: Disaggregation, Embedding Dimension, Peachland, Continuous Random Cascade, Temporal Scale, Water Demand.

* Ph.D. Student, Department of Civil Engineering

† Assistant Professor, Department of Civil Engineering, Corresponding author.

Tel.: +1-(250)-807-8464; Fax: +1-(250)-807-9850.

E-mail address: bahman.naser@ubc.ca

‡ Assistant Professor, Department of Mechanical Engineering

1 Introduction

To develop an integrated dynamic plan for water distribution systems (WDS), an approximate estimation of future values of water demand is required. This plan is crucial for efficient customer satisfaction and the optimum allocation among industrial, commercial and individual consumption. The estimation depends upon the increase of the rate of population and climatic changes, which are limited by the scarcity of resources. Hence, most of the research studies were done to evaluate water demand for handling future probabilistic water disasters. Majority of well-known methods need to use high-resolution temporal scales of the data for reliable estimation of water consumption. Nevertheless, only well-equipped SCADA systems can record such a high-resolution data (e.g., high-resolution temporal scale, proper time series of pressure values in pipelines, etc.). Therefore, development of a method to transfer low-resolution scales (e.g., annual and monthly values) to high-resolution (e.g., hourly or daily events) scale is necessary. A reliable demand management plan of an urban area requires the understanding of the behavior of series and consumption pattern in different spatial and temporal resolutions (Jorgensen B. 2013, Cominola 2018). Because, temporal high-resolution time series enable better estimation of peak demands that are very important in a short-term plan along with to develop a maintaining plan for pipeline in long-term period to overcome to probabilistic failures in the network (Beal 2016, Shabani 2016, Yousefi 2017, Yousefi P. 2018). These theories also have interpolation and extrapolation methods based on recorded values in SCADA. Those recorded values are time-dependent that makes them unable to provide satisfying high temporal scales. This study aims to present a method that allows transforming the data from low-resolution to high-resolution temporal scale. The concept of transformation from one temporal scale to another is a phenomenon and varies according to the particular case. The first step of recommendation to the researchers in the area is investigation of scaling behavior. Referring to literature to use disaggregation methods, the time series need to have scaling behavior. (B. W. Sivakumar 2004) and (B. Sivakumar 2006) considered the subject from the scaling point of view and studied the possibility of scaling behavior in suspended sediment load of rivers. (D. S. Valencia 1972, D. R. Valencia 1973) proposed disaggregation methods to transfer low-resolution time series to high-resolution temporal scales. They applied their proposed technique to disaggregate annual river discharge into seasonal values. Various downscaling and disaggregating methods were used to estimate high-resolution values. The value of rainfall in short-term temporal scales (Tarboton 1998, Hingray 2005, Gyasi-Agyei 2005, Segond 2006), down-scaling streamflow time series (Rebora 2016), evaluation of weather data (Debele 2017, Balaji 1999, Yates 2003), disaggregation are based on false-nearest neighbor method (Nowak 2010, Prairie 2007). Random cascade method is one of the well-known disaggregation methods that have been used in the area of hydrology. (Olsson 1998) applied random cascade method to disaggregate rainfall values in Sweden. (Molnar 2005) evaluated two various stochastic disaggregation model based on random cascade method to grain observed values into high-resolution scale. The mentioned methods were developed to apply in non-continues data (e.g., rainfall).

Considering the application of random cascade methods, this study aims to estimate high-resolution temporal scales of urban drinking water demand. A newly developed cascade method that is based on weights and probability statistics of the recorded data would be used to transfer continuous low-resolution, i.e. monthly values into high-resolution daily value in Peachland's urban drinking water demand (BC, Canada). The city of Peachland is selected to evaluate the performance of the model based on their drinking water demand (Table 1). Along with the consideration of scaling behavior of the used demand data, embedded dimensions with phase space reconstruction, lag time and disaggregation calculations have been done in five different temporal scales. Then the results of disaggregated values are compared with recorded data of the city. Six years monthly data are considered for the test case.

Property	Monthly	16-Days	8-Days	4-Days	2-Days	Daily
Data	64	128	256	512	1024	2048
Max. value*	1998263	1018734	532808	273002	138667	70336
Min. value*	83028	41189	20444	9995	4936	2406
Average*	566144	284548	141536	70768	35384	17692
St. deviation*	521088.9	264745.3	132795.7	66621.2	33428.8	16736.4
Kurtosis	0.641	0.527	0.587	0.591	0.597	0.616

Table 1: Statistics values of water consumption of Peachland (* m³).

2 Methodology

In the present study, the power spectrum method is used for investigation of scaling characteristics of water demand data. Use of the power spectrum indicates the existence of a specific temporal lobe distinctively which provides the fractal behavior (B. W. Sivakumar 2004, B. Sivakumar 2002).

2.1 Random Cascade Method

To disaggregate the time series, the cascade model divides the value of the demand (V) of each temporal scale into two scales (V_1, V_2). Each of the two scaled values resulted from multiplying the weight value ($W_1, W_2; W_1+W_2=1$) to (V) (Gaume 2007). Figure 1 shows the brief scheme of the process in random cascade disaggregating.

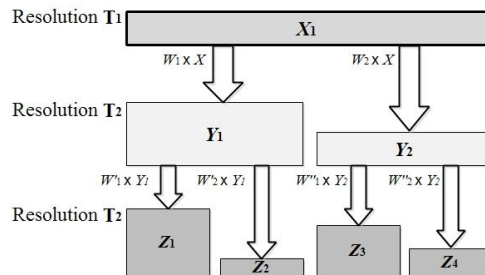


Figure 1: Schematic representation of cascade weight distribution.

The idea of subdivision and boxes in cascade method has been taken from the proposed technique of (Olsson 1998). This method disaggregates non-continuous rainfall data into short-term values (see the details for distribution of the weight in (Olsson 1998)). Since the demand series are continuous values, this study contributes to upgrade the technique to be applicable for continues data (e.g., discharge, sedimentation, water consumption, etc.). Each box is divided into four subs; 1) have the value of demand greater than previous and successive time; 2) lesser than previous and greater than successive time; 3) greater than previous and lesser than successive; 4) lesser than both previous and successive ones. The mean and median of the demand values are considered as another variable to divide different boxes. By combining all the information about the boxes, the probabilities of each box $P(x|x)$ were estimated, and a proper weight (W_i) were given to each value to be divided into two different values.

2.2 Non-Linear Deterministic Method

To determine the degree of correlation of time series and the lag time, average mutual information (AMI) has been utilized. This function is a periodic phenomenon, which indicates stability of a relationship between repetitive values. It is assumed that $x_i, i = 1, 2, 3, \dots, n$ is a series of water demand with a resolution of T_1 and aim to find the values of $(z_i)_k, k = 1, 2, 3, \dots, p$ in a high resolution, and also $p = T_1 / T_2$ and values of x_i according to $(z_i)_k, k = (W_i)_k = (W_i)_k \times x_i$ to $(z_i)_k, (W_i)_k$ is the distribution of transfers weights of X_i which is written to $(z_i)_k$ (for more information about the details see (B. W. Sivakumar 2004)).

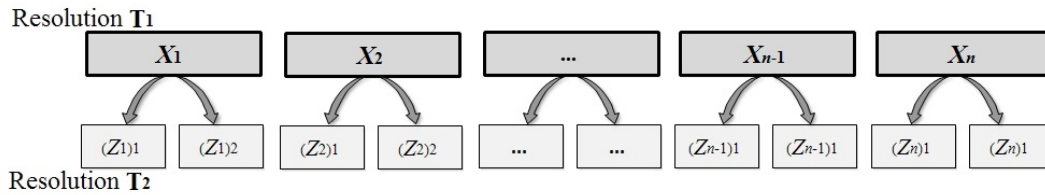


Figure 2: Schematic representation of distributions of weights between resolutions.

According to nonlinear behavior of demand data, the evolution of sub-mechanisms must be considered by using phase space reconstruction (see (Takens 1981)). This reconstruction makes it possible to draw a relationship between the first condition and next one by a functional equation.

$$Y_{j+T} = f_T(Y_j) \tag{1}$$

By disaggregating the values of water consumption, the resolution increases from T_1 to T_2 and in this way the distributed weights of $(W_{n+1})_k$ can be determined.

3 Results and Discussion

Figure 3 shows the result of power spectrum for daily scale of water demand. The value of β can be calculated by the gradient of line, approximately 1.775, which is more than 1, and represents scaling behaviour of the time series.

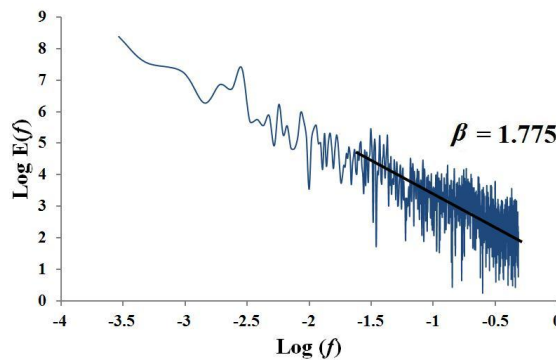


Figure 3: Scaling analysis of time series.

The results of continuous cascade method show a satisfactory application of the technique. Referring to Table 2, lowering the temporal scale, the accuracy decreases. It is because the number of available data that causes a better calibration for the models. Regarding the cascade method, which is classified into four different boxes (subs), each box has a probabilistic value. In other words, cascade

method is similar to other non-linear forecasting and estimation methods because of being dependent on the amount of data set. Figure 4 shows the comparison of actual recorded and disaggregated values of water demand in Peachland City (BC, Canada). However, the accuracy of the results is not perfectly aligned in simulating peak demand values; it provides precise values with 100% accuracy in calculation of cumulative values demand in respect to define a valid urban water plan. However, the used model in the present study considered one variable (consumption values), applying the impact of related variables in water consumption such as weather information, holidays, peak hours, etc., seems to be more helpful for calibration of the model to estimate subs’ weight more accurately. It has been not reported yet to consider the mentioned variables in the models if they need to have scaling behavior, for which the authors are working on it to upgrade the continuous disaggregation techniques to be applicable in different conditions.

Temporal Scale	Correlation Coefficient	RMSE (m^3)	MAE
2-Days to Daily	0.9142	7386	4595
4-Days to 2-Days	0.9149	14662	9095
8-Days to 4-Days	0.8996	31423	19138
16-Days to 8-Days	0.8938	65399	40561
Monthly to 16-Days	0.8997	123404	72765

Table 2: Accuracy of the disaggregated values in different temporal scales by cascade mdoel.

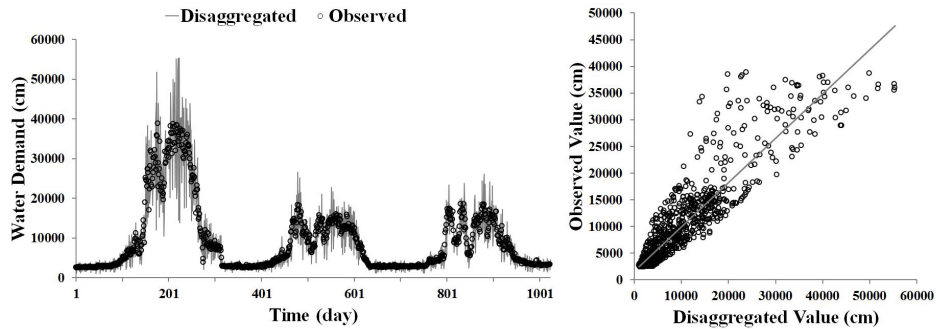


Figure 4: Disaggregated values of water demand in comparison with observed values (Daily).

To compare the results of the continuous cascade method with another well-known method, the non-linear deterministic method is used to disaggregate the value of water demand. Regarding the literature, the application of this method is well-known among researchers and was applied to various hydrology areas (discharge, sediment, climate, etc.). In the process of AML, the first local minimum is considered as the lag time. Figure 5 shows the results of time determining and the values of 24, 20, 18, 11 and 7 are for the daily, 2-, 4-, 8- and 16-days temporal scale consecutively.

The value of false nearest neighbors (FNN) for daily scale shows 9, which is considered as the number of optimum embedding dimension. The time series were divided into two groups, the first group is used for reconstruction of phase space and the second group, and half of it shows the performed fitness evaluation. Embedding dimension of $m=1$ to 10 shows the dynamic of reconstructed transportation; also, zero degree approximation is utilized to find neighbors in disaggregation process. The presented results (accuracy of disaggregation regarding with correlation coefficient (R) and root mean square error (RMSE) for different embedding dimensions from $m=1$ to 10, go following the results that are obtained from disaggregation of water demand values (See Table 3). Since the understanding of consumers’ behaviour depends on the knowledge of consumption anomalies, peak-hours, usage frequencies, used devices, metrological and social information (Yousefi P. 2018), considering the mentioned factors into account to be considered as a variable to develop

updated disaggregation method would be a great idea to “design targeted personalized demand management strategy, including economic incentives to upgrade inefficient appliances (e.g. (Mayer 2004))” (Cominola 2018).

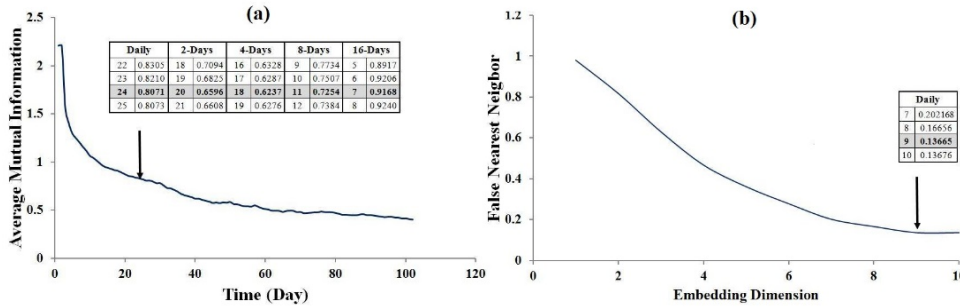


Figure 5: (a) Lag time and; (b) value of false nearest neighbor for different temporal.

Embedding dimension	2-Days to Daily		4-Days to 2-Days		8-Days to 4-Days		16-Days to 8-Days		Monthly to 16-Days	
	R	RMS E	R	RMS E	R	RMS E	R	RMS E	R	RMS E
1	0.9819	1645	0.980	3399	0.981	6663	0.961	1888	0.9819	1645
2	0.9780	1819	0.976	3711	0.977	7274	0.965	1780	0.9780	1819
3	0.9819	1643	0.978	3560	0.977	7197	0.968	1693	0.9819	1643
4	0.9826	1609	0.976	3748	0.976	7379	0.973	1572	0.9826	1609
5	0.9822	1630	0.979	3495	0.979	6872	0.955	2041	0.9822	1630
6	0.9814	1668	0.978	3614	0.982	6473	0.962	1852	0.9814	1668
7	0.9808	1695	0.980	3395	0.979	6954	0.957	1980	0.9808	1695
8	0.9800	1733	0.980	3361	0.977	7207	0.962	1854	0.9800	1733
9	0.9801	1727	0.976	3744	0.980	6826	0.963	1841	0.9801	1727
10	0.9822	1630	0.976	3720	0.979	6907	0.972	1581	0.9822	1630

Table 3: Accuracy of the disaggregated values in different temporal scales.

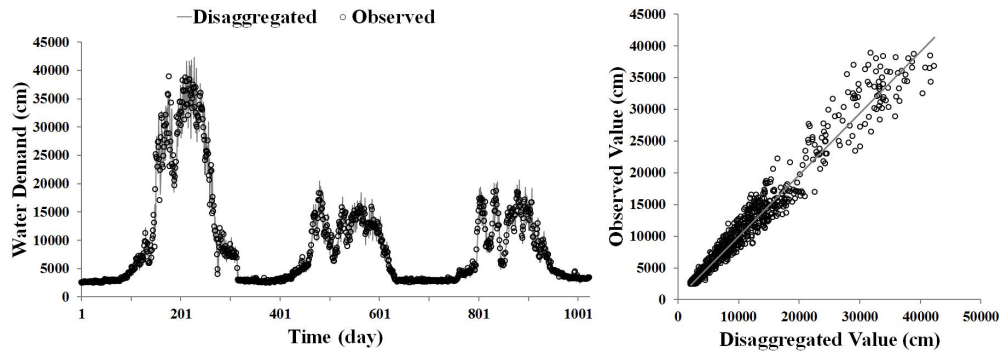


Figure 6: Disaggregated values of water demand in comparison with observed values (Daily).

4 Conclusion

In this study, continuous random cascade and nonlinear deterministic methods were considered in order to attain high resolution temporal data for a given water demand of Peachland City. Power spectrum was used to investigate the scaling behavior of the time series. The value of power spectrum was reported 1.775 which is more than 1 and represents the presence of scaling behavior of the time series. Down scaling was applied to convert monthly values into daily values for both methods. The domain of the given weight for random cascade method was in the range of 0.3 to 0.7 for each defined sub boxes which resulted from investigation of the relation for different temporal scales.

It was suggested to use other effective variables (e.g. temperature, humidity, average income, population growth and etc.) in order to calculate more accurate values of weight for the subs. The well-defined weight will improve the accuracy of continuous random cascade model in disaggregating. The value of R and RMSE for down scaling 2-days to daily scale was more accurate than the other temporal scales. Also, to disaggregate demand values by non-linear deterministic method, various embedding dimensions (EDs) varied from 1 to 10 were considered for different temporal scales. The lag time for different temporal scales was calculated by average mutual information (AMI). The optimum ED that resulted from false nearest neighbour (FNN) was 9, however, the results of model's evaluation for ED of 4, were reported more accurate than the other EDs in this study. By comparing the applications of the two techniques, it was suggested that non-linear deterministic approach is more effective than continuous random cascade method in simulating the peak demand in this stage of the research (without applying the role of related variables). The focus of this research was to realize the application of disaggregation methods, which are mostly used for hydrology issues, into urban management area i.e. drinking water consumption. Having the mentioned methods (or similar techniques) in the area will help the governors of the cities with low SCADA equipment to solve considerable problems. However, the value of fitness functions of the used methods was not so high; the total demand in large time scale is exactly equal to the recorded values. The total values of consumption are considered in WDS to design the network and management planning. The disaggregation methods can be considered as an effective replacement for any other nonlinear evolutionary forecasting methods because the cumulative values of their results are less accurate than disaggregation methods.

References

- Balaji, R., Lall, U. "A k-nearest-neighbor simulator for daily precipitation and other weather variables." *Water Resources Research banner* 35 (1999): 3089-3101.
- Beal, C. D., Gurung, T.R., Stewart, R.A. "Demand-side management for supply-side efficiency: Modeling tailored strategies for reducing peak residential water demand." *Sustainable Production and Consumption* 6 (2016): 1-11.
- Cominola, A., M. Giuliani, A. Castelletti, D. E. Rosenberg, Abdallah, A. M. "Implications of data sampling resolution on water use simulation, end-use disaggregation, and demand management." *Environmental Modelling and Software* 102 (2018): 199-212.
- Debele, B., Srinivasan, R., Parlange, J.Y. "Accuracy evaluation of weather data generation and disaggregation methods at finer timescales." *Advances in Water Resources* 30 (2017): 1286-1300.
- Gaume, E., Mouhous, N, Andrieuc, H. "Rainfall stochastic disaggregation models: Calibration and validation of a multiplicative cascade model." *Advances in Water Resources* 30, no. 5 (2007): 1301-1319.
- Gyasi-Agyei, Y. "Stochastic disaggregation of daily rainfall into one-hour time scale." *Journal of Hydrology* 309 (2005): 178-190.
- Hingray, B., Ben Haha, B. "Statistical performances of various deterministic and stochastic models for rainfall series disaggregation." *Atmospheric Research* 77 (2005): 152-175.
- Jorgensen B., Martin J., Pearce M., Willis E. "Some difficulties and inconsistencies when using habit strength and reasoned action variables in models of metered household water conservation." *Journal of environmental management* 115 (2013): 124-135.
- Mayer, P., DeOreo, W., Towler, E., Martien, L., Lewis, D. "Tampa water department residential water conservation study: the impacts of high efficiency plumbing fixture retrofits in single-family homes." Tampa Water Department and the United States Environmental Protection Agency, 2004.
- Molnar, P., Burlando, P. "Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model." *Atmospheric Research* 77, no. 1-4 (2005): 137-151.
- Nowak, K., Prairie, J., Rajagopalan, B., Lall, U. "A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow ." *Water Resources Research* 46, no. 8 (2010).
- Olsson, J. "Evaluation of a scaling cascade model for temporal rain-fall disaggregation." *Hydrology and Earth System Sciences Discussions* 2, no. 1 (1998): 19-30.
- Prairie, J., Rajagopalan, B., Lall, U., Fulp, T. "A stochastic nonparametric technique for space-time disaggregation of streamflows." *Water Resources Research* 43, no. 3 (2007).
- Rebora, N., Silvestro, F., Rudari, R., Herold, C., Ferraris, L. "Downscaling stream flow time series from monthly to daily scales using an auto-regressive stochastic algorithm: StreamFARM." *Journal of Hydrology* 537 (2016): 297-310.
- Segond, M.L., Onof, C., Wheeler, H. "Spatial-temporal disaggregation of daily rainfall from a generalized linear model." *Journal of Hydrology* 331 (2006): 674-689.
- Shabani, S., Yousefi, P., Adamowski, J., Naser, Gh. "Intelligent Soft Computing Models in Water Demand." In *Plants, Water Stress in*. INTECH, 2016.
- Sivakumar, B. "A phase-space reconstruction approach to prediction of suspended sediment concentration in rivers." *Journal of Hydrology* 258, no. 1-4 (2002): 149-162.
- Sivakumar, B. "Suspended sediment load estimation and the problem of inadequate data sampling: a fractal view." *Earth Surface Processes and Landforms* 31 (2006): 414-427.
- Sivakumar, B., Jayawardena, A.W., Fernando, T.M. "River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches." *Journal of Hydrology* 265, no. 1-4 (2002): 225-245.

- Sivakumar, B., Wallender, W. "Deriving high-resolution sediment load data using a nonlinear deterministic approach." *Water Resources Research banner* 40 (2004).
- Takens, F. "Detecting strange attractors in turbulence." *Dynamical systems and turbulence*. Berlin, Heidelberg: Springer, 1981. 366-381.
- Tarboton, D. G., Sharma, A., Lall, U. "Disaggregation procedures for stochastic hydrology based on nonparametric density estimation." *Water Resources Research banner* 34 (1998): 107-119.
- Valencia, D. R., Schaake J.C. "Disaggregation processes in stochastic hydrology." *Water Resources Research banner* 9 (1973): 580-585.
- Valencia, D.R., Schaake, J.C. *A disaggregation model for time series analysis and synthesis*. U.S.A: School of Engineering, Massachusetts Institute of Technology, 1972.
- Yates, D., Gangopadhyay, S., Rajagopalan, B., Strzepek, K. "A technique for generating regional climate scenarios using a nearest-neighbor algorithm." *Water Resources* 39, no. 7 (2003).
- Yousefi P., Naser, Gh., Mohammadi, H. "Application of wavelet decomposition and phase space reconstruction in urban water consumption forecasting – Chaotic approach (Case Study)." In *Wavelet Theory and Its Applications*. INTECH, 2018.
- Yousefi, P., Shabani, S., Mohammadi, H., Naser, Gh. "Gene Expression Programing in Long Term Water Demand Forecasts Using Wavelet Decomposition." *Procedia Engineering* 186 (2017): 544-550.