Journal of Water and Wastewater, Vol. 31, No. 6, pp: 60-69

Probabilistic Modeling and Estimation of Flow Rate of Sewage Treatment Plant Using Monte Carlo Hybrid Method

J. Piri[1], B. Pirzadeh[2], B. Keshtegar[3], M. Givehchi[4]

1. PhD Student, Dept. of Civil Engineering, University of Sistan
and Baluchestan, Zahedan, Iran
2. Assist. Prof., Dept. of Civil Engineering, University of Sistan
and Baluchestan, Zahedan, Iran
(Corresponding Author) b_pirzadeh@eng.usb.ac.ir)
3. Assoc. Prof., Dept. of Civil Engineering,
University of Zabol, Zabol, Iran
4. Assist. Prof., Dept. of Civil Engineering, University of Sistan
and Baluchestan, Zahedan, Iran

(Received Jan. 29, 2020 Accepted Sep. 1, 2020)

To cite this article:
Piri, J., Pirzadeh, B., Keshtegar, B., Givehchi, M. 2021. "Probabilistic modeling and estimation of flow rate of sewage treatment plant using monte carlo hybrid method" Journal of Water and Wastewater, 31(6), 60-69.
Doi: 10.22093/wwj.2020.217770.2982.

Abstract

One of the most important results of hydraulic and hydrological modeling of the urban drainage network is hydrograph estimation. Annual journals on hydraulic and hydrological problems, especially in developing countries, are full of missing data, discrete and continuous gaps in most hydrological data, such as inlet flow and other series of flow data. This is due to reasons such as not registering statistics, deleting wrong statistics, or failure of measuring devices. The present paper attempts to investigate the flow rate of wastewater entering the sewage treatment plant (STP) for pumping station planning. The novelty of the research is using the Monte Carlo method, which is one of the simulators of the effect of the uncertainties in the timing and cost of the project, and Fourier series regression model to randomly generate data. These methods were used to simulate a 1-minute time scale of sewage flow data for Zahedan. Considering the construction of the treatment plant in the last decade, this is the first research with this approach in this treatment plant. The results indicate that, this method has been successful in estimating sewage flow data in the hourly and minute intervals. Finally, 270 days of flow data were obtained from a time interval of one minute with two methods of distribution: Lognormal function and a nonlinear Fourier series. Among these two methods, Fourier series' accuracy was higher in terms of statistical indicators. In this simulation, RMSE, d, CI and EF values for Fourier regression model are 0.29, 0.99, 0.99, and 0.99, respectively.

Keywords: Probability Distribution, Monte Carlo, Lognormal, Fourier Series, Wastewater Inflow.

1. Introduction

It is easily possible to determine the characteristics of random variables based on statistics and probabilities. The demand of water for different needs, due to the impossibility of accurate prediction of the factors affecting it in the future, has a random nature and is considered as a continuous random variable with positive values and, on the one hand, the number of failures in supply water demand, such as a pump failure at treatment plant station in a distribution network in a given time period, is a discontinuous random variable that can only take positive integer values. Time series analysis is widely used in many engineering fields. Time series analysis usually has two objectives: first, modeling is based on the random hydraulic and hydrological variables' mechanism, and second, predicting the future values of the series based on its past (Shirvanimoghaddam et al., 2019, Smith, 2001, Te Chow, 2010, Alizadeh, 2013).

There are several methods in which Monte Carlo simulation method has been used for solving hydraulic and hydrological problems. Monte Carlo method is a comprehensive and reliable method. In fact, the Monte Carlo method is versus simulation because the time factor is involved in the simulation, a method to solve a non-random or random problem that over time plays no essential role, random numbers (random numbers uniformly in the range of zero to one) are used. Over the past decades, analysis and modeling of time series in the field of water resource management has attracted considerable attention to simulate and predict hydraulic and hydrological variables (Akan and Houghtalen, 2003, Bhaduri, 2004, Cho and Liu, 2018, Rad and Ardeshir, 2015). One of the most important results of hydraulic and hydrological modeling of urban drainage networks is hydrograph estimation in different parts of the drainage network.

These hydrographs will be used in the design of structures as well as in the risk evaluation of each region (Basser et al., 2015). Rad & Ardeshir (2015) performed infrastructure projects with limited water supplies, to get the most efficiency with the least risk (Rad and Ardeshir, 2015). One of the main problems of researchers for modeling the time series is whether the process is linearly or nonlinearly modeled. Therefore, for proper design of pumping stations and utilization of these stations, their accurate flow estimates are of particular importance. For the accurate estimation of pumped water, it is necessary to select appropriate models. In this study, for simulation of input flow data, Monte Carlo hybrid simulator has been used with Fourier series and distribution methods, which respectively generate artificial data sets according to their statistical and nonlinear regression distributions. In order to develop the model structure, several studies have been conducted on simulation models, the most important of which are Lee (Lee, 2006) & Hatteb (Hattab et al., 2013).

The simulation models, often implemented using the Monte Carlo method, do not have the potential effects of identifying certainty on project objectives. In this present simulation, input data using different models such as stochastic are random samples of the function for each variable. By repeating these samplings, the model is simulated many times and a probability distribution is obtained for the desired objectives (Lu et al., 2016). Khosravi et al., in order to investigate the accuracy of regression and stochastic methods for estimating maximum annual flow in the central Alborz region, used the maximum daily flow rate to reconstruct maximum annual flow (Khosravi et al., 2012). In their study, correlation method was used for reconstruction.

The study results indicated that these methods are very accurate for maximum annual flow reconstruction. NJ simulated appropriate values of different inputs for pollution in the treatment plant using Monte Carlo method, and argued that this method worked well in the proper design of the treatment plant (NJ, 2011). Lu et al., using Monte Carlo method simulated the flow rate in four wells, and using this simulation a significant amount of wells' flow measurement costs were reduced (Lu et al., 2018). In a study by Cho and Lu, using Monte Carlo combined with Markov chain, simulated pollution random data on observation wells (Cho and Liu, 2018). However, they stated that the combination in simulation is not always easily achieved and certainly is not a guarantee. In a study, Norouzi Khatiri et al., to investigate the uncertainty of the input parameters of the Madflu model, used the Monte Carlo method and the optimal amount of salinity and head in the aquifer were obtained (Norouzi Khatiri et al., 2020).

When the Monte Carlo method has been used to recognize the significance of model uncertainty on the optimization results, the simulation model demands to be requested many times, which will cause a considerable amount of calculation (Fan et al., 2020). An incorrect estimate of the input current in the treatment plant will show an increase in energy consumption, so, in the present study, for the first time, Fourier nonlinear series and continuous random variable methods are used to estimate the sewage flow hydrograph. Then, sewage flow data are obtained from the Monte Carlo method for one minute interval. All computations are performed using two methods of distribution and Fourier nonlinear series with Monte Carlo modeling in MATLAB version 2014.

2. Materials and Methods

2.1. Study area

Zahedan Station Sewage Treatment Plant[1] is located in southeastern Zahedan. Since 2012 the Treatment Plant

---

[1] Sewage Treatment Plant (STP)

Table 1. Models related to probability distribution functions used in this research

Probability distribution function	Parameters of model	Kind of distribution
Lognormal	Parameter scale σ Parameter position μ	$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp^{-(\ln x-\mu)^2/(2\sigma^2)}$ $-\infty<\mu<+\infty, \quad \sigma>0$
Logistic	Parameter scale σ Parameter position μ $Z=\frac{x-\mu}{\sigma}$	$f(x)=\frac{\exp(-Z)}{\sigma(1+\exp(-Z))^2}$ $-\infty<\mu<+\infty, \quad \sigma>0$
Gamma	Parameter from α Parameter scale β Gamma function Γ	$f(x)=\frac{(x^{\alpha-1})}{\beta^{\alpha}T(\alpha)}\exp(-\frac{x}{\beta})$ $\alpha>0, \quad \beta>0$

has been used as a new source of fresh water in the area for various purposes including artificial nutrition of underground aquifers, green spaces and forestry (Fig. 1). Currently, 14,000 sewage users are connected to the sewage collection network and the sewage is transferred to the STP from these users. In this treatment plant, the total predicted population for the plan is more than 1 million persons by 2032. The sewage data is measured 24 hours a day at an hourly rate at this station. Fig. 2 shows the mean variation in sewage flow and standard deviation overnight in the time interval of 9 months per year. In this paper, we have studied and analyzed the data for 9-month flow on a daily basis from April 2018 to December 2018 (Zahedan, 2018).

Distributions like Lognormal, Logistic, and Gama have commonly been used in hydraulic and hydrologic data analysis (Table 1). In this study, the same distributions have been used, and the summary of the relations related to these distributions is listed in Table 1.

2.2. Lognormal distribution

The normal distribution is symmetric around its average, and lognormal distribution is usually suggested for modeling physical phenomena such as water flow (Chen and Rao, 2003). The lognormal distribution is a continuous widely used distribution in the state where random variables have non-negative values. The random variable X has a lognormal distribution, if logarithmic conversion Y = LnX has a normal distribution with mean μ_{LnX} and variance σ^2_{LnX}. PDF[1] lognormal random variable is shown in Table 1. In each equation, X indicates the variable sewage flow and Greek letters represent the parameters of the probability density function. In order to study the goodness of fit, Kolmogrov-Smirnov, Anderson-Darling and Chi-square tests have been used (Kosugi, 1994).

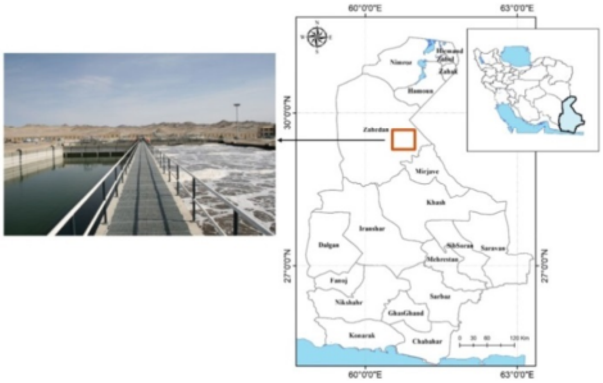


Fig. 1. Geographical location of Zahedan sewage treatment plant

2.3. Fourier series

Fourier conversion and Fourier analysis are used in various hydraulic and hydrological problems. If the function f(x) is a periodic function with the period of P = 2L, thenwe can use the above function as a sum of

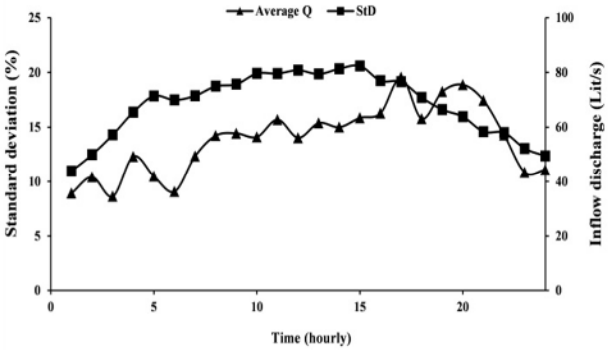


Fig. 2. Changes in mean sewage flow rate and standard deviation hourly for 9 months in 2018

---

[1] Probability Distribution Function (PDF)

sinusoidal and cosine sentences with different arguments that are called Fourier series written as follows (Asmar, 2016)

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right) \qquad (1)$$

Where

a_n, b_n and a_0 are called Fourier series coefficients and are calculated using the following equations

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi}{L}xdx \quad , \qquad n = 1,2,3,\ldots\ldots \qquad (2)$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi}{L}xdx \quad , \qquad n = 1,2,3,\ldots\ldots \qquad (3)$$

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(x) \qquad (4)$$

2.4. Monte Carlo Simulator

Monte Carlo simulation is the most commonly used technique for the release of uncertainties in various aspects of a system to the predicted efficiency (Lu et al., 2018, Lu et al., 2016).

Each computational algorithm and Monte Carlo simulation in a simple or complex form contain several main components. These components can be repeated more than once, independently or intermittently, and in different parts of the algorithm. These components consist of random numbers, probability distribution function, sampling rules, error estimation and dispersion standard deviation reduction, and metropolis algorithms. The cornerstone of any random simulation method or Monte Carlo method is based on the continuous use of random numbers. In Monte Carlo simulation methods, the use of probabilistic distribution is very important, and its selection is very effective on the process economics and accelerating the response. A probability distribution function for a random number (x) in a given range, continuous or discrete, yields the probability of its occurrence as f(x, t), where t denotes the existence of effective parameters in the distribution function. Moreover, the integral of the distribution function as a cumulative distribution function is also used to study and evaluate the performance of Monte Carlo algorithm (Fig. 3). For the distribution function f(x), the value of cumulative distribution function is defined as the integral mode on the entire range as follows

$$F(x) = \int_{a}^{x} f(\tau)d\tau \qquad (5)$$

In general, f(x) denotes the chance or probability of an event for a random amount, and F(x) represents the chance or probability of occurrence of all events for a random range $0 < x < 1$. The normal distribution functions are commonly used in Monte Carlo algorithms. This function is as follows

$$f(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma}} \qquad (6)$$

Where

μ is the mean value of the data and σ is the amount of dispersion or standard deviation.

The results of the preliminary statistical analysis of sewage flow data showed that the standard deviation from sewage flow data q(t) has an insignificant relationship with time (t) (Table 2). Therefore, the lognormal distribution function was chosen for subsequent calculations (Figs. 4 and 5), and then Monte Carlo method was selected to obtain 270 days of simulation data for a 1 minute interval (i.e. the conversion of 1 hour to 1 minute). Also, using Fourier nonlinear regression method, first the best series was obtained for one-hour sewage flow data, then Monte-Carlo method simulated 270 days of data in 1 minute intervals. Monte Carlo program flowchart for both Monte Carlo method and steps is described in the following 1 minute intervals.

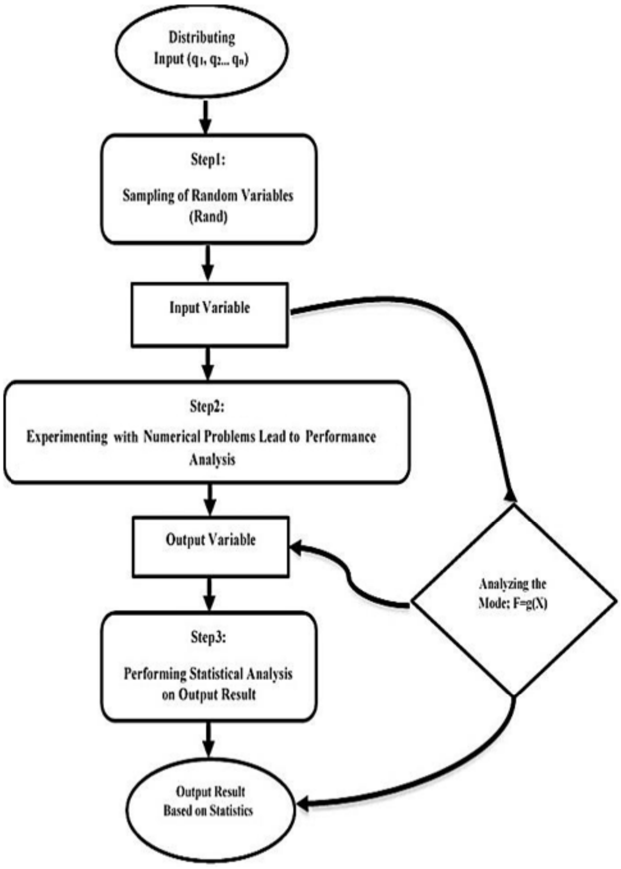


Fig. 3. Monte Carlo program flowchart for simulating sewage flow data

Table 2. The goodness of fit values test statistics

Distribution	Kolmogorov-Smirnov statistic	Anderson-Darling statistic	Qi-square statistic
Lognormal	0.092	68.29	1859.7
Logistic	0.103	74.36	2502.8
Gamma	0.111	72.18	2735.5

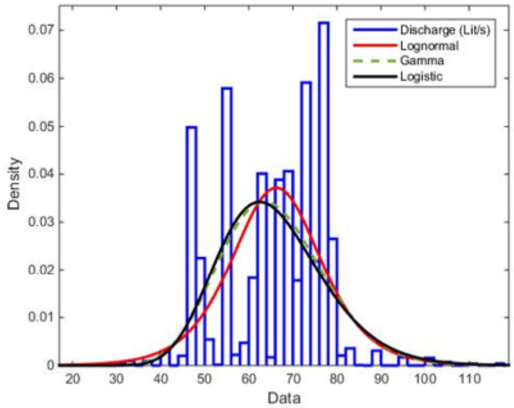


Fig. 4. PDF for four lognormal, gamma, and logistic distributions

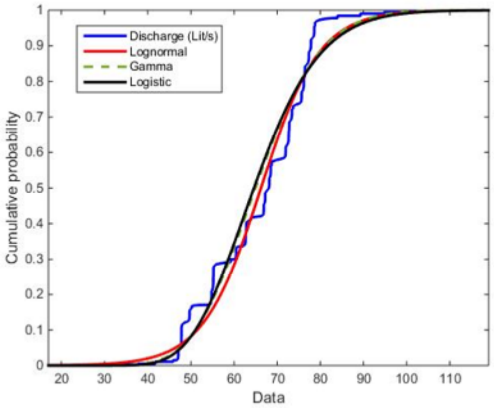


Fig. 5. CDF[1] for four distributions of lognormal, gamma, and logistic distributions

A set of variables the relationship between X and F in Monte Carlo is calculated for sewage flow data as follows:

Step 1: Give data for the time interval of 1-hour for the actual data (Q_i)

Step 2: Calculate the mean of data ($\bar{Q}_{\Delta t,i}$)

Step 3: Determine $Q_{N,\, i}$ for any time interval: $Q_{N,i}=\frac{Q_i}{Q_{\Delta t,i}}$

Step 4: For each minute, a Random Number[2] is generated between zero and one

Step 5: Determine the equation

$$\frac{1}{2}+\frac{1}{2}\mathrm{erf}\left(\frac{\ln Q_{N,i}-\mu}{\sqrt{2\sigma}}\right)-RN=0 \text{ to obtain } \tilde{Q}_{N,i} \quad (7)$$

where

μ and σ and are respectively the mean and standard deviation of the variable (where Ln is the logarithmic operator, and the erf is error function):

Finally, by using Steps 1-5, the new flow at a given time interval (one minute) is obtained

$$Q_i = \tilde{Q}_{N,i}\times\bar{Q}_{\Delta t,i} \quad (8)$$

Therefore, according to the steps of Monte Carlo combined method, Flowchart will be as shown in Fig. 6.

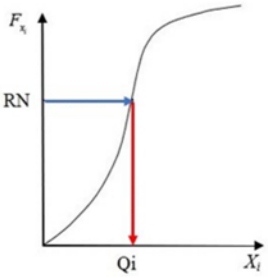


Fig. 6. Estimation of flow data based on Monte Carlo method

2.5. Model evaluation

In the present study, for evaluating and comparing models we used mean square error (RMSE) (Piri and Kisi, 2015), adaptive index (d)(Keshtegar et al., 2016), model efficiency (NSE) (Keshtegar et al., 2019) and confidence index (CI)((Keshtegar et al., 2019)

$$RMSE=\sqrt{\frac{\sum_{i=1}^{NE}(O(X_i)-P_i)^2}{NE}} \quad (9)$$

$$d=1-\frac{\sum_{i=1}^{NE}(O(X_i)-P_i)^2}{\sum_{i=1}^{NE}(\left|O(X_i)-\bar{O}\right|+\left|Pi-\bar{P}\right|)^2} \quad 0\le d\le 1 \quad (10)$$

$$NSE=1-\frac{\sum_{i=1}^{NE}(O(X_i)-P_i)}{\sum_{i=1}^{NE}(P_i-\bar{O})} \quad 0\le NSE\le 1 \quad (11)$$

$$CI = NSE \times d \quad (12)$$

---

[1] Cumulative Distribution Function (CDF)

[2] Random Number (RN)

O is measured flow values, P is simulated flow values, $\overline{O}$ is mean flow measured and $\overline{P}$ is simulated mean values.

When comparing models the lower the RMSE index value, the closer NSE, CI, and D indicators' values are to 1, the higher the accuracy of the model (Keshtgar et al., 2016).

3. Results and discussion

After preparing and completing sewage flow data of the treatment plant station, they were fitted with the current distribution of hydrology, such as 2-parameter lognormal, logistic, and gamma. In Qi-square statistical method, the amount of Qi-square theory distribution was compared with a certain level of acceptance and degree of freedom. The Kolmogorov-Smirnov test is another test used to measure the distribution of a particular distribution sample.

In Anderson-Darling method, this test checks whether the data follows the specified distribution. Due to this issue, these tests can be considered as a non-parametric statistical method. Using graphics methods, the Probability Distribution Function (PDF) and Cumulative Distribution Function (CDF) are selected as the best statistical data distribution functions for analyzing station flow data.

Kolmogorov-Smirnov, Anderson-Darling and Qi-square for the proper probability distribution function for sewage flow have shown that the lognormal function was more suitable for continuation of the calculations (Table 2, Figs. 4 and 5). It can also be concluded that the sum of standard random variables has a normal distribution; the result of the multiplication of random variables of the lognormal will also have a lognormal distribution.

In this study, according to fit tests including Kolmogorov-Smirnov, Anderson-Darling and Qi-square, and drawings of CDF and PDF, sewage flow data, as shown in Table 1 and Figs. 3 and 6, and the lognormal distribution function were selected. The results of the simulation of the lognormal distribution function and Fourier non-linear series have been compared with Monte Carlo. First, daily flow data were fitted to the lognormal probability distribution in an hourly interval. Then, it was simulated using a combined method of Monte Carlo and distribution function in a time interval of one minute. Similarly, using Fourier non-linear series, daily data was fitted in hourly intervals.

It was then simulated using a non-linear Fourier series in a 1-minute interval. Similarly, using Fourier nonlinear series, daily data was fitted in hourly interval. It was then simulated using a non-linear Fourier series in a 1-minute interval. Table 3 shows statistical comparison of distribution methods using Fourier series in hourly interval. The two methods in Table 3 and Figs. 7a, 7b, 8a and 8b were first compared statistically with RMSE, d, CI and EF values per day, hourly.

The results showed higher accuracy of the normal distribution function than Fourier method. Zahedan is a tourist city, in addition to the fixed population, the free community (tourists) will increase and decrease the discharge of wastewater per unit time (Fig. 7a)

According to Fourier method, within a time interval of one minute, the optimal matching in the fifth series is achieved. The nonlinear regression function for estimating data in a time interval of one minute is as follows. Fourier series 5-order equation (Eq. 13)

$$\text{Fitted Fourier }(x) = a_0+a_1\times\cos(X\times W)+b_1\times\sin(X\times W)+a_2\times\cos(2\times X\times W)+b_2\times\sin(2\times X\times W)+a_3\times\cos(3\times X\times W)+b_3\times\sin(3\times X\times W)+a_4\times\cos(4\times X\times W)+b4\times\sin(4\times X\times W)+a_5\times\cos(5\times X\times W)+b_5\times\sin(5\times X\times W) \qquad (13)$$

W and X are angular frequency and time. Coefficients of Eq. 13 are listed in Table 4.

Figs. 9a and 10a compare the discharge estimate of 270 days in a one-minute time for nine months of the year, with observed data from the two methods of distributing the lognormal and the Fourier series, respectively. Table 5 and Figs. 9a, 9b, 9c, 10a, 10b, and 10c illustrate the

Table 3. Statistical comparison of distribution methods using Fourier series in hourly interval

	EF (%)	CI (%)	d	RMSE (lit/s)
Lognormal	0.9992	0.9999	0.9998	0.1013
Fourier	0.8648	0.8322	0.9623	4.2712

Table 4. Coefficient of Eq. 13

Parameters	Value	with 95% confidence bounds
a_0	66.79	(64.24, 69.35)
a_1	-16.49	(-18.76, -14.23)
b_1	1.679	(-3.719, 7.076)
a_2	-3.315	(-4.399, -2.231)
b_2	1.058	(-0.4801, 2.596)
a_3	-1.835	(-3.319, -0.3505)
b_3	-1.368	(-2.567, -0.1691)
a_4	-1.386	(-2.609, -0.1625)
b_4	-1.329	(-3.33, 0.6728)
a_5	0.3846	(-0.5497, 1.319)
b_5	-0.6204	(-1.472, 0.2312)
W	0.004143	(0.003703, 0.004582)

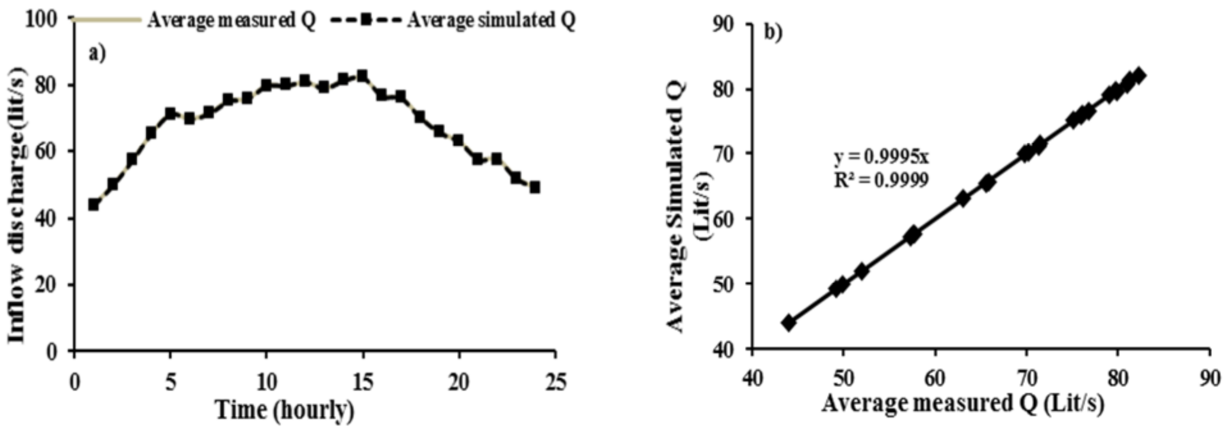


Fig. 7. Comparison of the mean measured and simulated flow data based on the lognormal distribution function in hourly interval, a) measured and simulated hourly data b) scatter plots of measured and simulated data

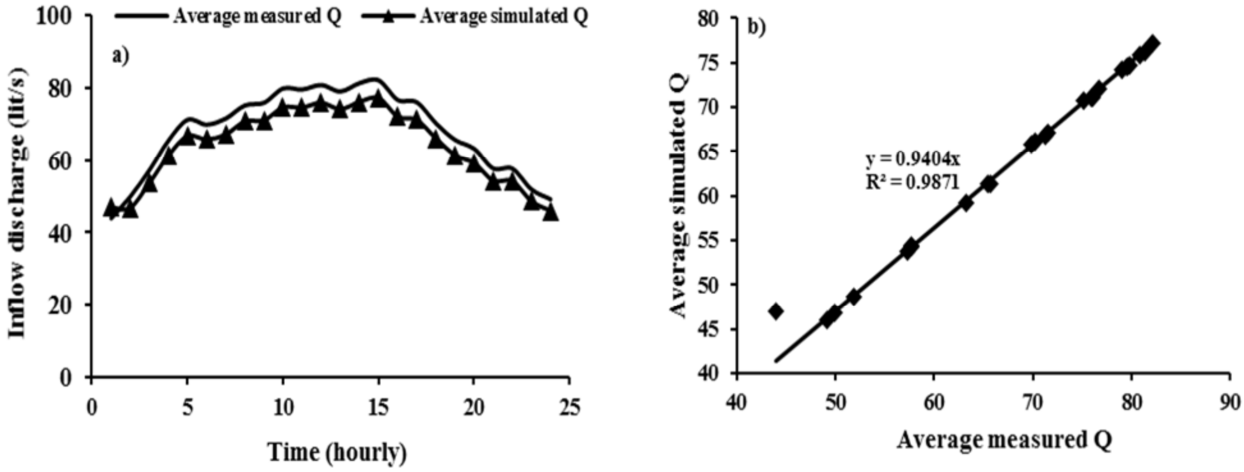


Fig. 8. Comparison of the mean measured and simulated flow data based on Fourier series in hourly interval,a) measured and simulated hourly data b) scatter plots of measured and simulated data

simulation of sewage flow data in a time interval of one minute. The statistical analysis indicates that the Fourier series is more accurate than the distribution method in a 270-day simulation over one minute. Another advantage of Fourier nonlinear method has been the output equation compared to the distribution method, i.e., in the Fourier method, it has a 5-order comparison, which is very accurate in the simulation of sewage data and can easily be used at any time interval of flow data simulation for Zahedan.

Table 5. Shows statistical comparisons of Fourier series distribution method in a 1-minute interval

	EF (%)	CI (%)	d	RMSE (lit/s)
Lognormal	0.9948	0.9935	0.9987	0.8221
Fourier	0.9994	0.9992	0.9998	0.2931

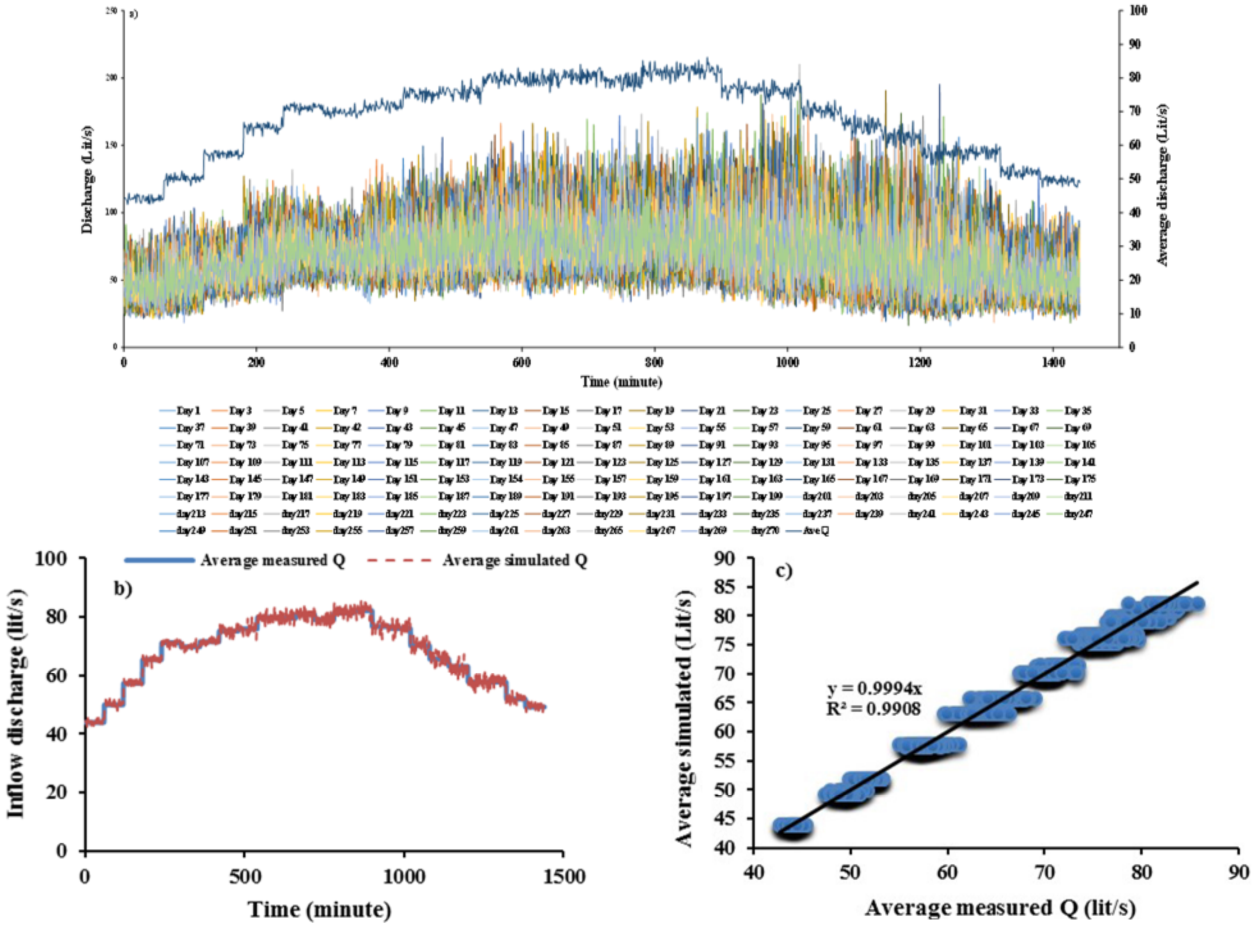
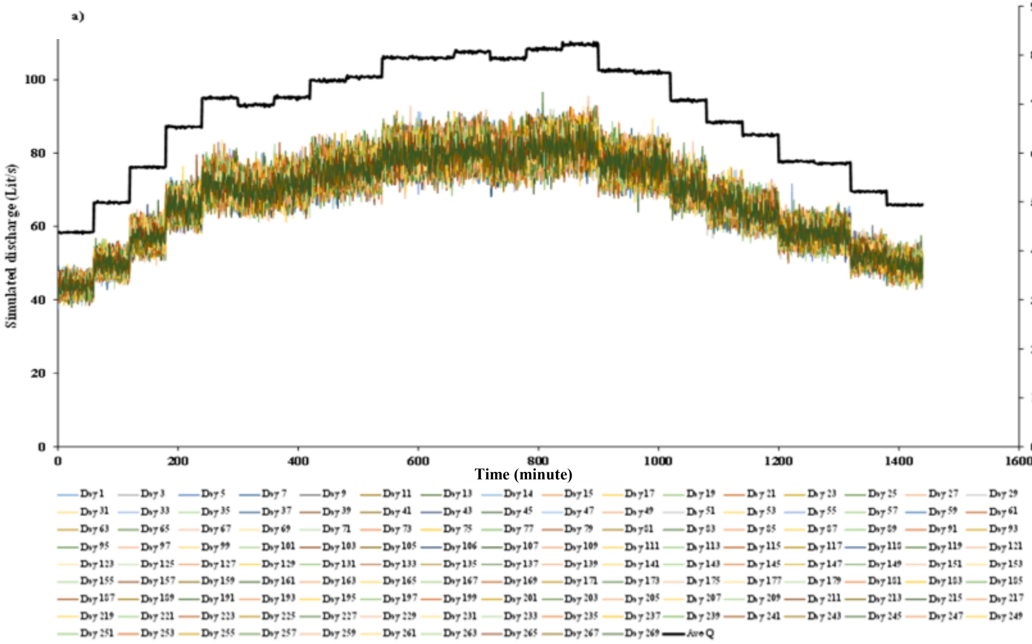


Fig. 9. Estimation of 270 days of simulated flow data based on Monte Carlo method with lognormal distribution function, based on an interval of one minute for 270 days, a) Estimation of 270 days b) comparison of the mean measured data and simulated distribution dispersion data c) scatter plots of measured and simulated data



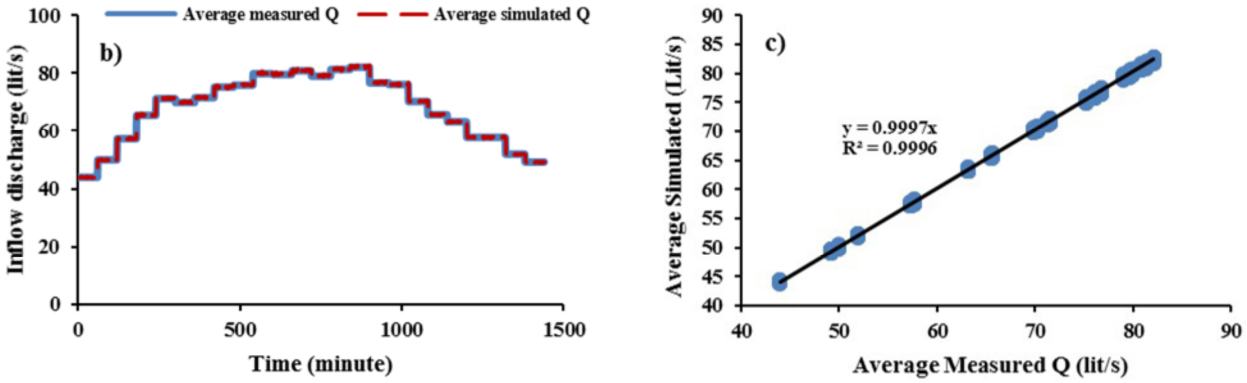


Fig. 10. Estimation of 270-days of simulated flow data based on Fourier series method, based on an interval of one minute for 270 days, a) estimation of 270 days b) comparison of the mean measured data and simulated Fourier series data c) scatter plots of measured and simulated data

4. Conclusion

Accurate estimation of the input flow can help managers in designing the pumping station and reducing energy consumption. In this paper, a data mining approach has been developed to model the sewage flow data of the sewage treatment plant (STP). This data mining approach is based on randomized data based on the Monte Carlo simulation method. For simulating the data, known probabilistic functions have been used, such as 2-parameter lognormal, Logistic, and Gamma. Using a graphic technique, the best statistical distribution of the data, lognormal, was selected for analysis of the station's flow data. The data were randomly generated based on the Fourier series regression model. This is an innovation in estimating the flow rate inversely for optimization methods. In Fourier linear and non-linear distribution models, values of hourly mean and standard deviation were used to produce data on a 1-minute time scale. The results of the probabilistic method based on probability distribution and non-linear Fourier series fit method using the Monte Carlo simulation method have been compared with Zahedan sewage flow data. The results indicate that 270 days of data were simulated in one minute, and both ways had an acceptable estimate. In this simulation, RMSE, d, CI, and EF values for the Fourier regression model were obtained about 0.29, 0.99, 0.99, and 0.99, respectively. This method provides a superior prediction given to form the best statistical results. The nonlinear Fourier series as a probabilistic data-driven approach was presented the high-performances for accuracy and tendency compared to the distribution method for simulating the nonlinear response of sewage flow. Also, the Fourier nonlinear method is superior to the distribution method given from the lognormal function thus, it can simulate the flow values at the desired time interval.

References

Akan, A. O. & Houghtalen, R. J. 2003. *Urban hydrology, hydraulics, and stormwater quality: engineering applications and computer modeling*, John Wiley & Sons.

Alizadeh, A. 2013. *The Principles of Applied Hydrology,* Imam Reza (AS) University, Mashhad. (In Persian)

Asmar, N. H. 2016. *Partial differential equations with Fourier series and boundary value problems*, Courier Dover Publications.

Basser, H., Cheraghi, R., Karami, H., Ardeshir, A., Amirmojahedi, M., Akib, S. et al. 2015. Modeling sediment transport around a rectangular bridge abutment. *Environmental Fluid Mechanics,* 15**,** 1105-1114.

Bhaduri, B. 2004. Urban hydrology, hydraulics, and stormwater quality: engineering applications and computer modeling. *Journal of the American Water Resources Association,* 40**,** 1667.

Chow, V. T. 2010. *Applied Hydrology*, Tata McGraw-Hill Education.

Chen, H. L. & Rao, A. R. 2003. Linearity analysis on stationary segments of hydrologic time series. *Journal of Hydrology,* 277**,** 89-99.

Cho, W. K. T. & Liu, Y. Y. 2018. Sampling from complicated and unknown distributions: Monte Carlo and Markov Chain Monte Carlo methods for redistricting. *Physica A: Statistical Mechanics and its Applications,* 506, 170-178.

Fan, Y., Lu, W., Miao, T., An, Y., Li, J. & Luo, J. 2020. Optimal design of groundwater pollution monitoring network based on the SVR surrogate model under uncertainty. *Environmental Science and Pollution Research*, 1-13.

Hattab, N., Hambli, R., Motelica-Heino, M. & Mench, M. 2013. Neural network and Monte Carlo simulation approach to investigate variability of copper concentration in phytoremediated contaminated soils. *Journal of Environmental Management,* 129, 134-142.

Keshtegar, B., Heddam, S. & Hosseinabadi, H. 2019. The employment of polynomial chaos expansion approach for modeling dissolved oxygen concentration in river. *Environmental Earth Sciences,* 78, 34.

Keshtegar, B., Piri, J. & Kisi, O. 2016. A nonlinear mathematical modeling of daily pan evaporation based on conjugate gradient method. *Computers and Electronics in Agriculture,* 127, 120-130.

Khosravi, M., Salajegheh, A. & Mahdavi, M. 2012. Comparing chow regression and stochastic method for annual peak flood (case study: central Alborz region). *Journal of Iran- Water Resources Research*, 8(1), 43-50. (In Persian).

Kosugi, K. I. 1994. Three-parameter lognormal distribution model for soil water retention. *Water Resources Research,* 30, 891-901.

Lee, S. 2006. *Dynamic planning and control methodology: understanding and managing iterative error and change cycles in large-scale concurrent design and construction projects*. Massachusetts Institute of Technology Pub.

Lu, D., Ricciuto, D. & Evans, K. 2018. An efficient Bayesian data-worth analysis using a multilevel Monte Carlo method. *Advances in Water Resources,* 113, 223-235.

Lu, D., Zhang, G., Webster, C. & Barbier, C. 2016. An improved multilevel Monte Carlo method for estimating probability distribution functions in stochastic oil reservoir simulations. *Water Resources Research,* 52, 9642-9660.

Nj, A. 2011. Prevalence and risk factors of helicobacter pylori and intestinal protozoal infections among sewage workers in mansoura, Egypt. *Egyptian Journal of Occupational Medicine,* 35, 67-81.

Norouzi Khatiri, K., Niksokhan, M. H. & Sarang, A. 2020. Choosing various likelihood functions on uncertainty assessment in groundwater simulation-optimization model. *Water Supply,* 20, 737-750.

Piri, J. & Kisi, O. 2015. Modelling solar radiation reached to the Earth using ANFIS, NN-ARX, and empirical models (case studies: Zahedan and Bojnurd stations). *Journal of Atmospheric and Solar-Terrestrial Physics,* 123, 39-47.

Rad, I. Z. & Ardeshir, A. 2015. A multi-objective approach towards urban drainage development prioritization and project definition structure. *Journal of Hydroinformatics,* 17, 959-976.

Shirvani Moghaddam, M., Shirvani Moghaddam, K., Abolhasani, M. M., Farhangi, M., Barsari, V. Z., Liu, H. et al. 2019. Towards a green and self-powered internet of things using piezoelectric energy harvesting. *IEEE Access,* 7, 94533-94556.

Smith, E. 2001. Pollutant concentrations of stormwater and captured sediment in flood control sumps draining an urban watershed. *Water Research,* 35, 3117-3126.

Zahedan, W. T. P. 2018. Sewage collection network, water and wastewater research and design consultant engineers. Zahedan wastewater treatment plant. (In Persian)