Volume 9, Special Issue No. 1, February 2018

ISBN: 978-93-5300-455-2

International Journal of Advanced Research in Computer Science
(ISSN: 0976-5697)
RESEARCH PAPER

SHRI SHANKARLAL SUNDARBAI SHASUN JAIN COLLEGE FOR WOMEN
ICRACC 2018
(23rd–24th February 2018)

Available Online at www.ijarcs.info

A SURVEY ON WASTEWATER TREATMENT (WWT) ANALYSIS USING VARIOUS TECHNIQUES

Dr. C. Victoria Priscilla ,
Associate professor & Head,
P.G. Department of computer science,
S.D.N.B Vaishnav College for women,
Chennai, India

A. Anusuya,
Student, P.G. Department of computer science,
S.D.N.B Vaishnav College for women,
Chennai, India
email:anusuyaarumugam1@gmail.com

ABSTRACT

Wastewater is contaminated water that has been affected by human use and also causes environmental pollution. Waste Water Treatment (WWT) is a process of removing contaminants from wastewater and reuses the water in various applications such as Hydroelectric Power Generation, Agriculture, and Radioactivity etc. The developments of various techniques for WWT have been implemented by many researchers. The choice of the suitable treatment techniques is dependent on the wastewater pollutant concentrations such as Biochemical oxygen demand (BOD) and Chemical oxygen demand (COD). This paper explores various techniques like Data Mining, Machine learning, Principal Component Analysis (PCA), Support Vector Machine (SVM), Regression Trees (RT) and provides the estimation of the wastewater quality characteristics.

Keywords: Waste Water Treatment; Data Mining; Machine learning; SVM; PCA; RT;

INTRODUCTION

Wastewater is used water which includes substances such as human waste, solid waste, chemicals, oils, and storm sewer. When wastewater is not properly treated it causes health problems and environmental pollution. [2] In every year 1.8 million people die due to waterborne disease. So the permanent solution is to reduce the threat of water pollution in wastewater treatments. Wastewater treatment is an essential scheme which has to be taken more seriously and also make improvements in our society and future generation. Wastewater management with water and nutrient loop is shown in figure1.Wastewater treatment is a process used to convert wastewater into an effluent which can be returned through the water cycle with minimum contact with the environment or directly reused. It is also called as water reclamation. Place of wastewater treatment is called as wastewater treatment plant (WWTP), repeatedly referred as Water Resource Recovery Facility (WRRF) or Sewage Treatment Plant. Wastewater Treatment Methods are categorized into three subdivisions such as Physical, Biological, and Chemical. Various types of treatment plant Help to reduce pollutants in wastewater. Types of Wastewater Treatments are shown in figure 2.

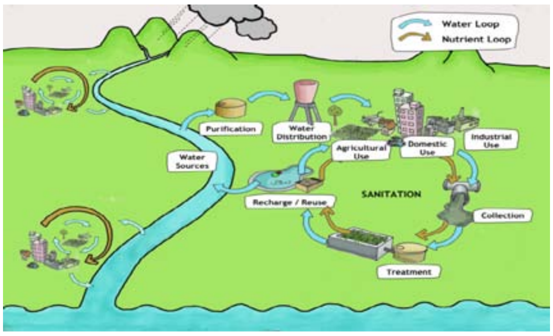

Figure 1: Waste Water Management [1]

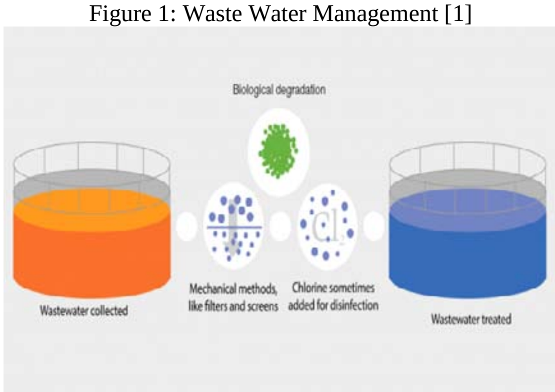

Figure 2: Types of Wastewater Treatments [3]

A. *Types Of Wastewater Treatments Process:*

a. *Physical Water Treatment Process [4]:*

- Screening used to remove large suspended solids.
- Sand Filtration is used to filter small suspended solids and dissolved solids.
- Membrane Filtration is used to remove total suspended solids and total dissolved solids.

b. *Chemical Water Treatment Process [5]:*

- Chemical Precipitation is used to remove dissolved metals in wastewater.
- Chemical Coagulation is used to destabilizing wastewater particles.
- Chemical Oxidation involves adding or generating oxidants in wastewater.
- Ion Exchange is used to soften the water.
- Chemical Stabilization is the process where sludge is mixed with lime to raise the pH to the high level (>12).

c. *Biological Water Treatment Process [6]:*

- Biological Anaerobic Wastewater Treatment process is used to maintain the organisms in the absence of oxygen.

- Biological Aerobic Wastewater treatment process is used to remove up to 98% of organic contaminants in the presence of oxygen. It is classified as more process as follows:
  - ❖ Activated Sludge Process (ASP) highly concentrates on the degradation of microorganisms and also removes nutrients from wastewater to produce a high-quality outflow.
  - ❖ Trickling Filters are used to remove compounds in the water such as ammonia after primary level treatment.
  - ❖ Aerated Lagoon Process (ALP) introduces oxygen into the oxidation pond in wastewater.
  - ❖ Oxidation Pond contains partially treated wastewater and also reduces the growth of microorganisms such as algae and bacteria.

B. Types Of Wastewater Treatments Plants [7]:

- Effluent Treatment Plants (ETP) are mostly used by various leading companies such as pharmaceutical, chemical industry, Leather industry, steel industry, and tanneries.ETP process used to purify water and removes toxic and non-toxic materials or chemicals from wastewater. It provides environmental protection for various leading companies.

- Sewage Treatment Plants (STP) is used to remove contaminants from wastewater and produce a waste stream and a solid waste. It uses physical, chemical, and biological processes to remove physical, chemical, and biological contaminants in wastewater.
- The Ministry of Environment & Forest, Govt. of India has launched a new scheme called Common Effluent Treatment Plant (CETP) and provides more benefit to industrial technologies.

C. Various Treatment Levels Of ETP [8]:

- Preliminary Level used to do physical separation operations such as screening, sedimentation, and clarification.
  - ❖ Screening: A Screen opens with uniform size used to remove large solids.
  - ❖ Sedimentation used to remove suspended solids from water.
  - ❖ Clarification used for separation of solids from fluids
- Primary Level used to remove suspended and settleable materials. It uses both physical and chemical method treatments.
- Secondary Level uses both Biological and chemical method treatments.
- Tertiary (advanced) Level used for the final cleaning process that improves water quality before recycled, reused or discharged to the environment.

DATA MINING TECHNIQUES:

Francesco et.al [9] has proposed two models, Support Vector Regression (SVR) and Regression Trees (RT) to estimate Wastewater Quality Indicators (WQI). Wastewater Quality indicators (WQI) were biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), and total dissolved solids (TDS) in WWT. Support Vector Regression (SVR) and Regression Trees (RT) provide robustness, reliability and high capability. SVR is same as Support Vector Machine (SVM) which used to analyze non-linear input data. Regression Trees (RT) is used to predict an outcome of real-valued functions. Both machine learning models are compared with statistical measures which are correlation coefficient (R^2) and root-mean-square-error (RMSE) that is shown in table 1.

Table 1: Comparison between SVR and RT [9]

Statistical Measures	RMSE		R^2	
WQI	**SVR**	**RT**	**SVR**	**RT**
TSS	**1049**	3486	**0.97**	0.906
TDS	**1549**	1826	**0.851**	0.828
COD	**1172**	2395	**0.893**	0.784
BOD	**104**	103	**0.87**	0.871

Finally, the result shows that SVR has better performance than RT in forecasting water quality indicators in WWT.

Andrew et.al [10] proposed Data-Mining models such as Support Vector Machine (SVM), Adaptive Neuro-Fuzzy Inference System (ANFIS), neural network (NN), k-nearest neighbor (K-nn) and Random Forest Tree for methane production in WWT. Biological anaerobic wastewater treatment process is used here. The Architecture of ANFIS is shown in figure 3. Both models are compared with statistical measures like percentage error (PE), fractional bias (FB), root mean square error (RMSE), normalized mean square error (NMSE), and index of agreement (IA) which is shown in table 2.

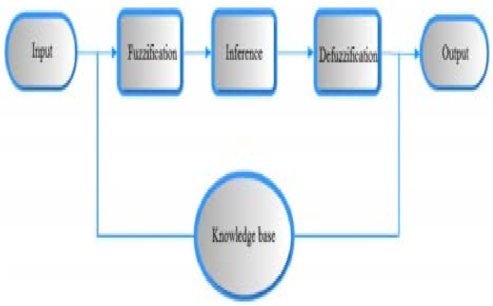

Figure 3: Architecture of ANFIS [10]

Finally, the result shows that Adaptive Neuro-Fuzzy Inference System (ANFIS) model provides better performance for methane production.

Table 2: Statistical metrics for all algorithms [10]

Models	PE	FB	RMSE	NMSE	IA
ANFIS	**0.08**	**0.00**	**74144**	**0.01**	**0.99**
NN	0.16	0.07	144787	0.04	0.99
SVM	0.10	0.03	84089	0.01	0.99
Random Forest Tree	0.10	0.06	92629	0.02	0.99
K-nn	0.13	0.02	123363	0.03	0.99

Hamed et.al [11] has proposed two models, ANN and ANN with PCA and two methods, multilayer perceptron (MLP) feed-forward neural network and stop training method used to predict industrial wastewater. Principal Component Analysis (PCA) is used to modify and improve the performance of the neural network and also used for reduction of dimensionality. This model used to estimate pH, Chemical Oxygen Demand (COD), Total Dissolved Solids (TDS), total nitrogen (Total N) and total phosphorous (Total P). Both models are compared with statistical methods such as Mean Absolute Error (MAE), determination coefficients (R2) and (RMSE) root mean square error which shown in figure 4.

Davut et.al [12] proposed Intelligent model which is based on wavelet packet decomposition, entropy and the neural network is used to predict the total suspended solids (TSS) in wastewater treatment. Wavelet packet decomposition is used to reduce the input vectors dimensions of the intelligent model. Structure of Intelligent model is shown in figure 5. Wavelet expansion function, wavelet basis function, mother wavelet function, scaling function and J-level wavelet decomposition were used in wavelet transforms. Wavelet packet and NN structure used to compose two layers such as wavelet packet layer and multilayer perceptron layer. Wavelet packet layer has two stages that are wavelet packet decomposition and wavelet entropy. Multilayer perceptron layer used to classify the features of wavelet packet layer. Wavelet packet and NN structure combine as WPNN.

Model	Parameter	Target		PCA-ANN			ANN		
Model 1	COD, pH		Data	R^2	[1]RMSE	[2]MAE	R^2	RMSE	MAE
	TDS	pH	All	0.76			0.63		
	N		Train	0.8	0.1569	0.1214	0.74	0.1850	0.1348
	P		Test	0.71	0.2606	0.2107	0.59	0.2240	0.1773
Model 2	COD, pH		Data	R^2	RMSE	MAE	R^2	RMSE	MAE
	TDS	TDS	All	0.72			0.68		
	N		Train	0.82	0.99e3	0.76e3	0.80	1.07e3	0.83e3
	P		Test	0.40	1.86e3	1.63e3	0.14	1.85e3	1.49e3
Model 3	COD, pH		Data	R^2	RMSE	MAE	R^2	RMSE	MAE
	TDS	N	All	0.84			0.76		
	N		Train	0.90	2.7207	3.4785	0.80	4.0307	2.9149
	P		Test	0.62	4.0246	4.9533	0.78	5.4452	4.3869
Model 4	COD, pH		Data	R^2	RMSE	MAE	R^2	RMSE	MAE
	TDS	p	All	0.7			0.66		
	N		Train	0.80	0.2479	0.1925	0.74	0.2676	0.2074
	P		Test	0.59	0.2828	0.2189	0.38	0.2796	0.2192
Model 5	COD, pH		Data	R^2	RMSE	MAE	R^2	RMSE	MAE
	TDS	COD	All	0.59			0.53		
	N		Train	0.71	43.673	32.706	0.64	43.765	31.8084
	P		Test	0.44	48.164	36.788	0.50	50.026	37.7631

Figure 4: Statistical parameters of various models [11]

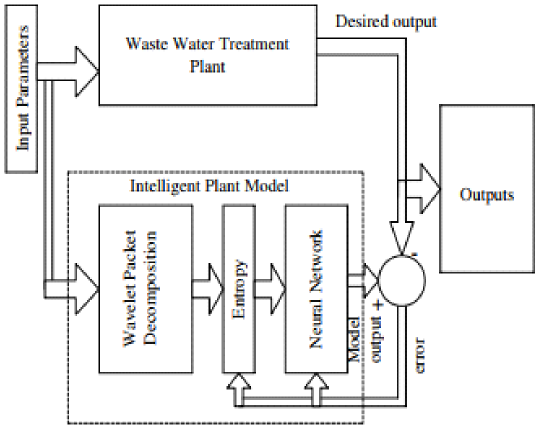

Figure 5: Structure of Intelligent model [12]

WPNN model used 50 different input data for training and 12 input data for testing. Finally, the result shows that WPNN model provides effectiveness and reliability in extracting features of input data.

Djeddou et.al [13] has proposed ANN model for prediction of sludge Volume Index (SVI) of activated sludge process (ASP) to avoid high non-linearity and non-uniformity task issues in Municipal Waste Water.ANN used a standard feed-forward back propagation neural network with three layers are input layer, one hidden layer, and an output layer with Log-sigmoid transfer function. Fifteen variables were used as input parameters. These variables were TSS, COD, BOD, pH, conductivity, NH^+_4, NO_3, P,TSS efficiency, COD efficiency, BOD efficiency, N-NH^+_4 efficiency, N-NO_3 efficiency, P-PO_4efficiency.Kolmogrov theorem used to determine no of hidden layer and no of hidden nodes (NHN).ANN model is evaluated through statistical methods like correlation coefficient (R), mean absolute error (MAE), root mean square error (RMSE) and mean absolute error (MAPE) that is shown in table 3.ANN model 15-13-1 has one input layer with fifteen input variables, one hidden layer with thirteen nodes and one output layer with one output variable. Finally, the result shows that ANN model 15-13-1 is a useful tool for activated sludge process (ASP) and also increases the performance of Waste Water Treatment Plant.

'Table 3: Performance of ANN prediction models [13]

ANN model	$R_{TRAINING}$	R_{ALL}	MAE	RMSE	MAPE (%)
ANN 15-5-1	0.9490	0.8091	0.364	0.548	18.76
ANN 15-6-1	0.9705	0.9126	1.141	1.346	14.79
ANN 15-7-1	0.8164	0.7565	0.394	0.484	28.43
ANN 15-9-1	0.9922	0.8246	0.292	0.479	21.35
ANN 15-10-1	0.977	0.8476	0.268	0.446	18.67
ANN 15-11-1	0.9191	0.8217	0.347	0.454	22.82
ANN 15-13-1	**0.9993**	**0.8432**	**0.186**	**0.443**	**10.98**
ANN 15-15-1	0.983	0.8784	0.254	0.393	17.12

Samaneh et.al [14] proposed hybrid ANN-COA model which offers a high degree of accuracy for predicting and controlling WWTP. The model was designed through Artificial Neural Network (ANN) and Cuckoo Optimization Algorithm (COA).

Selected input variables were T, pH, DO, BOD, COD, TSS, TDS, NO_3 and PO_4. Neural network with three layers is input layer, one hidden layer and an output layer with two functions such as hyperbolic tangent sigmoid and linear transfer function. An Aggregated measure (AM) has three parameters are model bias (MB), NS and r_{mod}. Here ANN and ANN-COA models are evaluated through statistical methods were Regression coefficient (R), mean square error (MSE), root mean square error (RMSE) and aggregated measure (AM) results shown in table 4. Finally, the result shows that ANN-COA model provides better performance than ANN.

Table 4: Performances of ANN and COA-ANN models [14]

	Models	R	MSE	RMSE	MB	NS	r_{mod}	AM
ANN	ANNBOD	0.72	36.41	6.03	0.07	0.7	0.62	0.75
	ANNCOD	0.8	29.7	5.44	0.06	0.82	0.79	0.8
	ANNTSS	0.7	48.29	6.94	0.08	0.6	0.54	0.68
COA-ANN	**COA-ANNBOD**	0.58	82.1	9.06	0.05	0.85	0.8	0.86
	COA-ANNCOD	0.68	66.75	8.17	0.032	0.9	0.85	0.9
	COA-ANNTSS	0.36	73.11	8.55	0.06	0.85	0.74	0.84

Hossien et.al [15] proposed Neural Network model used for prediction of Perchlorate wastewater treatment. Perchlorate (C104) is a strong electrochemical oxidant. The variables were speed of particle, pH particle, temperature of particle, density, wastewater, sewage of TSS, flue and waste rate of the tank. These variables used to predict Perchlorate density of wastewater. The entire dataset was normalized with two functions such as logsin and tansig. Different trail network models are evaluated through statistical methods were correlation coefficient (R), mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE).the statistical measurement is shown in table 5. Finally, the result shows that Trail 2 model provides better performance.

Table 5: Statistical metrics of different models [15]

Frequency Network	**No of neurons**	**MAE**	**R**	**MAPE (%)**	**RMSE**
Trail 1	16	7.453	0.8507	4.595	11.964
Trail 2	**17**	**7.373**	**0.8846**	**4.831**	**10.827**
Trail 3	16	7.457	0.8642	4.745	11.55
Trail 4	15	7.517	0.8723	4.625	11.15
Trail 5	16	7.561	0.863	4.515	11.709

Abdullah et.al [16] proposed Artificial Neural Network model to predict the performance of WWTP. Data was taken from Konya WWTP. Neural network with three layers is input layer, one hidden layer, an output layer and many neurons in each layer. The input values are pH, temperature, BOD, COD, TSS, and TDS. The entire dataset was normalized with three functions such as logsin, Purelin, and tansig. ANN model was evaluated through statistical methods are correlation coefficient (R) and mean square error (MSE).ANN 3-3-1 model shows maximum correlation coefficient and minimum mean square error.ANN 3-3-1 model is used as a valuable performance assessment tool for Wastewater plant operations. Summary of ANN 3-3-1 model is shown in table 6.

Table 6: Summary of ANN 3-3-1 model [16]

No of neurons	Transfer functions	Training R	Testing R	MSE
7	Logsig	0.98	0.93	0.0078904
	Purelin	0.89	0.58	0.0052331
	Tansig	0.93	0.86	0.0044014
8	Logsig	0.99	0.87	0.0075730
	Purelin	0.88	0.58	0.0058282
	Tansig	0.96	0.57	0.0044315
9	**Logsig**	**0.99**	**0.95**	**0.0023310**
	Purelin	0.61	0.38	0.0084474
	Tansig	0.99	0.79	0.0058679
10	Logsig	0.99	0.86	0.0071530
	Purelin	0.79	0.34	0.0049574
	Tansig	0.99	0.83	0.0067905
11	Logsig	0.99	0.79	0.0067391
	Purelin	0.61	0.50	0.0043681
	Tansig	0.81	0.58	0.0443300

Xiupeng et.al [17] proposed Neural Network model for short-term prediction of influent flow rate in WWTP. Multilayer Perceptron Neural Network Algorithm (MLP) used to build the prediction model by different time horizons with various data such as influent flow rate, rainfall rate, and radar reflectivity. Data was taken from WRF Iowa. Different time horizons were current time (t), t+15, t+30, t+60, t+90, t+120, t+150, t+180. Multilayer Perceptron with three layers is an input layer with 34 inputs, an output layer, and one or more hidden layers. NN model is built by various data mining algorithms such as MLP, random forest, boosted tree and SVM .all data mining algorithm is evaluated through statistical methods are correlation coefficient (R), mean absolute error (MAE) and mean square error (MSE).The entire dataset was normalized with five functions such as logistic, identity, hyperbolic, exponential and sine function. Comparison of all algorithm with statistical metrics proves that MLP model has MSE and MAE values are smallest one and highest correlation coefficient (R=0.988). Comparison of all algorithms with statistical metrics is shown in table 7. Based on comparison MLP model was constructed with statistical metrics and different time horizons as t to t+180.Prediction of influent flow rate with two models were graphically designed and shown in figures (7, 8). Finally, the result shows that prediction model predicted influent flow rate well until t+150.

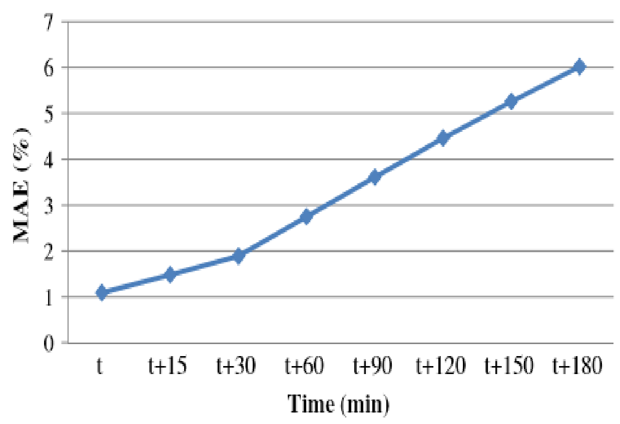

Figure 5: MAE of model for influent rate [17]

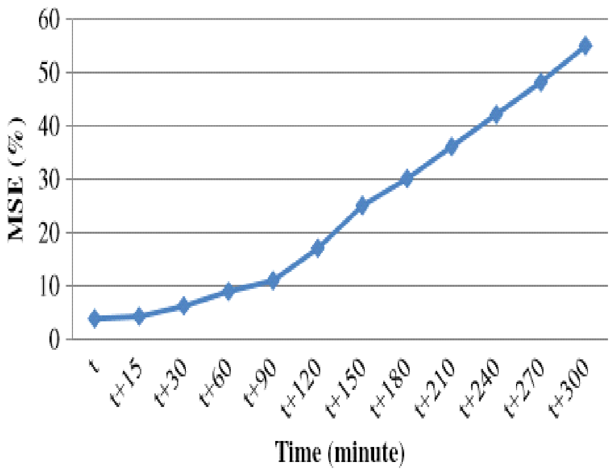

Figure 6: MSE of model for influent rate [17]

Table 7: Prediction Accuracy of various algorithms [17]

Algorithm	MAE	MSE	R
MLP	**1.09**	**4.21**	**0.988**
Random forest	3.04	20.69	0.945
Boosted tree	1.77	11.16	0.97
SVM	1.47	5.46	0.985

CONCLUSION AND FUTURE WORK:

Water plays big role in world economy. Water management and Wastewater treatment practice is most important for future generation to avoid environmental pollution and health problems. In this paper, an analysis is presented for Wastewater Treatment Plant (WWTP) using various techniques at different locations. Based on this survey several techniques are analyzed for the implementation of the waste water treatment. this study shows when the waste water is recycle the quantity of the waste water is reduced in urban area. The aim of the future plan is to analyses the data set from UCI Machine Learning Repository and implements the water quality indicator to evaluate various data mining algorithms which examine the performance of WWTP.

REFERENCES:

1. http://archive.epi.yale.edu/case-study/primary-vs-secondary -types-wastewater-treatment
2. https://en.m.wikipedia.org/wiki/wastewater
3. https://www.sswm.info/category/implementation-tools/implementation-tools-introduction
4. http://www.waterprofessionals.com/learning-center/articles/physical-water-treatment/
5. https://www.thomasnet.com/articles/chemicals/wastewater-chemical-treatment/
6. http://www.neoakruthi.com/blog/biological- treatment-of- wastewater.html
7. http://www.yourarticlelibrary.com/water/types-of-wastewater-treatment-process-etp-stp-and-cetp/27418
8. Rakesh Singh Aiswal, Dr.Santosh Kumar ,Shweta Singh and Megha Sahu, “Wastewater Treatment by Effluent Treatment Plants”, SSRG International Journal of Civil Engineering,Vol-3, ISSN 2348-8352,32-38,2016
9. Francesco Granata Stefano Papirio, Giovanni Esposito, Rudy Gargano and Giovanni de Marinis, “Machine Learning Algorithms for the Forecasting of Wastewater Quality Indicators”,www.mdpi.com/journal/water, Vol-9, DOI 10.3390/w9020105, 1-12, 2017.
10. Andrew Kusiak, Xiupeng Wei, “Prediction of methane production in wastewater treatment facility: a data-mining approach”, Annals of Operations Research, Vol-216, ISSN 0254-5330, 71-81, 2014.
11. Hamed Hasanlou, Naser Mehrdadi, Mohammad Taghi Jafarzadeh, Hamidreza Hasanlou, “Performance Simulation of H-TDS Unit of Fajr Industrial Wastewater Treatment Plant Using a Combination of NN and PCA ”, Journal of Water Resource and Protection,Vol-4,DOI 10.4236/jwarp.2012.45034 ,311-317 ,2012
12. Davut Hanbay a, Ibrahim Turkoglu a, Yakup Demir b, “Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks”, Expert Systems with Applications, DOI 10.1016/j.eswa.2006.10.030,1038–1043,2008.
13. Djeddou M, Achour B, “The Use of A Neural Network Technique for the Prediction of Sludge Volume Index in Municipal Wastewater Treatment Plant”, Larhyss Journal, ISSN 1112-3680, 351-370, 2015.
14. Samaneh Khademikia a, Ali Haghizadeh c, Hatam Godini d, Ghodratollah Shams Khorramabadi b, “Artificial Neural Network-Cuckoo Optimization Algorithm for Optimal Control of Khorramabad Wastewater Treatment Plant, Iran”, Civil Engineering -- Journal, Vol-2, 2016.
15. Hossien Fakhraee, Seyed Shayan Akbarpour, “Application of Artificial Neural Network for Prediction of Perchlorate Wastewater Treatment”, Computational Research Process in Applied Science and Engineering, Vol 01(03), ISSN 2423-4591, 103-111, 2015.
16. Abdullah Erdal TUMER, Serpil EDEBALI, “An Artificial Neural Network Model for Wastewater Treatment Plant of Konya”, IJISAE, ISSN: 2147-6799, 2015.
17. Xiupeng Wei, Andrew Kusiak and Hosseini Rahil Sadat, “Prediction of Influent Flow Rate: Data Mining Approach”, Journal of Energy Engineering, Vol-139, ISSN 0733-9402, 118-123, 2013