

# Data Visualization of Influent and Effluent Parameters of UASB-based Wastewater Treatment Plant in Uttar Pradesh

Parul Yadav<sup>1</sup>, Aditya Chaudhary<sup>2</sup>, Anand Keshari<sup>3</sup>, Nitish Kumar Chaudhary<sup>4</sup>  
Priyanshu Sharma<sup>5</sup>, Kumar Saurabh<sup>6</sup>, Brijesh Singh Yadav<sup>7</sup>

Department of CSE, Institute of Engineering and Technology, Lucknow, India<sup>1</sup>  
Department of CSE, Institute of Engineering and Technology, Lucknow, India<sup>2,3,4,5</sup>  
Namami Gange STP Project, Voltas Ltd. Patna, India<sup>6</sup>  
Uttar Pradesh Rajya Vidyut Utpadan Nigam Limited, Lucknow<sup>7</sup>

**Abstract**—A rise in the population of a region implies an increase in water consumption and such a continuous increase in the usage of water worsens wastewater generation by the region. This escalation in wastewater (influent) requires the Wastewater Treatment Plants (WWTPs) to operate efficiently in order to process the demand for sewage disposal (effluent). This research paper is based upon visualizing and analyzing the parameters of influent like COD, BOD, TSS, pH, MPN and also, the parameters of effluent like COD, BOD, DO, pH and MPN of Bharwara WWTP situated in Lucknow, India which is the largest UASB-based wastewater treatment plant in Asia. We also design and implement an initial model using the machine learning based techniques to analyze as well as predict the parameters of influent and effluent of the WWTP. Model Performance is measured using Mean Squared Error (MSE) and Correlation Coefficient (R). For analyzing and designing the model, the parameters of influent and effluent have been collected over a period of 26 months on a daily basis covering the variations between seasons and climate. As a result, the model shall provide a better quality of effluent along with consuming the plant resources in an efficient manner.

**Keywords**—Wastewater treatment plant; Bharwara STP; UASB-based plant; influent or effluent prediction; data visualization of influent and effluent; machine learning based for WWTPs

## I. INTRODUCTION

Wastewater Treatment Plants (WWTPs) take part in playing a crucial role in shaping the urban and rural environments as they are used for processing sewage water and removal of various particles and chemicals which are harmful for the water hydrosphere and the organisms which are dependent on it. An increase in the population of a region implies an increase in water consumption and such a continuous increase in the usage of water results in an increase in the wastewater generated by the region [1]. This increase in influent requires the wastewater treatment plants to operate efficiently in order to process the demand for effluent (sewage disposal) [2, 3, 5].

Besides increase in influent, another more challenging issue in a wastewater treatment plant is the fluctuating or uncertain behaviour of various parameters of the influent in

the plant which can be due to varying environmental factors also [15]. To maintain the effluent parameters within the standard range, the wastewater treatment plants need to operate and process on the influent coping up with its varying parameters. On the other side, the wastewater treatment plants require to do optimum utilization of resources during the treatment of influent. Consequently, this uncertain nature of influent parameters demands to find insights and hidden patterns by applying visualization and analytics on the real time historical/ recorded data which in turn shall help to provide/estimate better and efficient (optimized) utilization of resources at wastewater treatment plants. Further knowing the flow and parameters of influent and parameters of effluent in advance shall reduce operational cost of the wastewater treatment plants.

With this objective, we collected and recorded the water parameters for over 26 months (April 2019 to May 2021) from Bharwara wastewater treatment plant situated in Lucknow district which is the largest UASB-based wastewater treatment plant in Asia as it has the capacity to operate and process an average flow rate of 345 Millions of Liter per Day (MLD) with the ability to handle a peak load of 517 MLD of sewage daily. In this paper, we analyze the influent and effluent parameters of Bharwara wastewater treatment plant.

The paper is organized in six sections. Section II presents the research works done in the line of analyzing and predicting influent and effluent parameters across the globe. Section III describes the working of the Bharwara wastewater treatment plant and also its current technological status. In Section IV, we elaborate the methodology of the proposed model. Section V highlights the results obtained and analysis done. Section VI summarizes the conclusions and shows line of the future works.

## II. RELATED WORK

We have carried out the literature survey in the line of research work under two dimensions. The first dimension of study is in the line of exploring technological status of wastewater treatment plants outside India, and the second dimension of study is in the line of exploring technological status of wastewater treatment plants inside India. We shall

discuss both dimensions one by one in the next two subsections.

#### A. International Status

1) *Konya wastewater treatment plant [konya, turkey]:* Tümer Abdullah et al. [2], proposed a model using Artificial Neural Network (ANN) for the prediction of Total Suspended Solids (TSS) based on the input parameters Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), TSS. Model performance was evaluated via Mean Squared Error and Correlation Coefficient (R) for the Konya Wastewater treatment plant. Neural Networks of various hidden layers were used and the correlation coefficient in the training set has reached up to 0.99, that is, a satisfactory result for the proposed model. The model was implemented using MATLAB which increased the complexity of designing ANN models as compared to Python. Training and testing take a lot of effort and scaling the models to other WWTP will require ample resources. In this model, the number of layers and number of neurons per layer were decided/ obtained on the basis of analyzing and comparing error on various trails. In the research paper, the performances of nine different models were compared.

2) *Wastewater treatment plant in South Korea:* Guo Hong et al. proposed ANN and Support Vector Machine (SVM) models to predict the Total Nitrogen (T-N) concentration in the Yong-Yeon (YY) WWTP in Ulsan, South Korea [3]. For evaluation of the model, Coefficient of Determination ( $R^2$ ), Nash-Sutcliffe efficiency and relative efficiency criteria were used. A sensitivity analysis was done using a pattern search algorithm and Latin Hypercube One Factor at a Time (LH-OAT) [4] which showed that the ANN model gave superior results as compared to the SVM model. Resources used in the research are costly and it might not be possible to propose this model for the prediction of other effluent particles like TSS as the quality parameters are dependent on the site. In WWTPs at Korea, the anaerobic digestion process of sewage sludge with Food Waste has been increased. The increase of the Food Waste adversely affects digestion process and results in getting poor quality of effluent water from WWTPs.

3) *Wastewater treatment plant in Italy:* Granata Francesco et al. [5] conducted a study on stormwater discharge and proposed a model for the estimation of COD, BOD, TSS, and Total Dissolved Solids (TDS) in the wastewater. Support Vector Regression (SVR) and Regression Tree algorithms were used for modeling, and Coefficient of determination ( $R^2$ ) and Root Mean Squared Error (RMSE) were the performance evaluators. For COD, TSS and TDS, the SVR model performed better than the Regression tree while for BOD, Regression Trees gave better results than SVR. Extending the proposed model to another treatment plant might not be a good choice as the conditions satisfying the development are heavily dependent on rainwater and only in a specific climate and rainy weather.

4) *Wastewater treatment plant in Hong-Kong:* Qin

Xusong et al. [6], showed that wastewater quality can be monitored online. In this paper, UV/ VIS spectrometry and a turbid-meter were used to monitor COD, TSS, and Oil & Grease concentrations. Signals from the two sensors were fused using Sensor fusion technique. Boosting-Partial Least Squares (Boosting-PLS) [6] method was used to make the model and predict the wastewater quality based on the fused information.

5) *Gongxian wastewater treatment plant in yibin, china:* Wang Rui et al. [7] used four machine learning methods (Linear Regression, Ridge, Lasso and ElasticNet) for predicting the influent parameters. For influent parameter predictions, these methods showed high accuracy. The model proposed is used as warning module for assisting in daily operations of WWTP.

#### B. National Status

1) *Wastewater treatment plant in Mangalore:* D S Manu et al. used an artificial intelligence-based model [8] to predict the performance of the treatment plant for the removal of effluent nitrogen particles. Three different techniques/ models, SVM, ANFIS trapezoidal MF model and ANFIS Gbell MF model [8] were used and were implemented in MATLAB. Influent parameters taken were pH, ammonia nitrogen, free ammonia, and Kjeldahl nitrogen. Performance evaluation was done by RMSE, MSE, and Correlation Coefficient (R). The SVM model gave satisfactory results. This research is limited to a biological wastewater treatment plant and the study is conducted considering the biological waste only, thereby extending the proposed model to another site with a different type of wastes might not be possible.

2) *Sewage treatment plant (STP) in Delhi:* Gautam et al. [9] focused on the monitoring of inlet and outlet parameters and measuring the effectiveness of STPs in Delhi, India. The cluster analysis approach was performed to find any relation between the current site and other sites, aiming to find similar sites. Sulfate, Nitrates, Chloride and Phosphate, and Bi-carbonates concentrations were measured and the results showed that STP efficiency was not up to the mark. Samples were collected manually and the scope of automating the analyzing process is very limited. As the study is a decade old, the method used might be good for the estimation but the time consumed in the process can be reduced, if it were to be conducted and monitored online.

3) *345 MLD UASB-based Bharwara STP/ WWTP:* Banerjee et al. [10] focused on the working performance of STP and upgrading Up-flow Anaerobic Sludge Blanket (UASB) reactor technology. The removal efficiency of COD, BOD, and TSS was measured and the relation between pH and influent parameters was determined. Measurement of the parameter was done two times in a month over a period of four months. Hence, this model is susceptible to climate changes, and a small amount of training data might not be a good choice for predicting the quality parameters.

### III. CURRENT TECHNOLOGICAL STATUS

As of now many researchers have contributed in proposing machine learning models [2, 3, 5, 6, 7, 8, 9, 10] for wastewater treatment plants. However, the available monitoring technologies used for analyzing and predicting parameters of wastewater quality have a number of limitations or drawbacks e.g., models are suited for plants of outside India [2, 3, 5, 6, 7]. In India, such smart systems for monitoring, analyzing, predicting quantity and quality parameters of wastewater are at a very preliminary stage. Smart, advanced and/ or automated systems are required not only for efficient utilization of resources but for smooth functioning of the wastewater treatment plants also, which can directly affect the health of humans/ living beings dependent on them. In India, no existing standard smart monitoring tool is currently available/ used [8, 9, 10] that predicts flow of influent and parameters of influent and effluent in advance for effective resource utilization at plant. The automation of system at WWTPs is still at an immature stage and not as developed as other process industries. The authors are researching in the line of implementing one such model/ infrastructure/ framework for wastewater treatment plant where data analysis shall be performed on real time data collected from Bharwara wastewater treatment plant, Lucknow (Fig. 1). In this paper, authors have presented the preliminary findings related to influent and effluent parameters of Bharwara wastewater treatment plant.

#### A. Pre-Treatment

The raw water is brought to the inlet chamber of Bharwara wastewater treatment plant (345 MLD UASB-based Bharwara STP) from the existing rising main of Gwari pumping station. The function of inlet chamber is to break the pressure flow and allow the sewage to flow by gravity to treatment unit. This chamber shall also function as flow distribution chamber to the screen channel.

Bharwara wastewater treatment plant is Asia's Largest UASB-based Wastewater Treatment Plant. The process flow of Bharwara wastewater treatment plant is shown in Fig. 2 and the brief description to its treatment scheme is given above.



Fig. 1. Aerial Photograph of Bharwara WWTP.

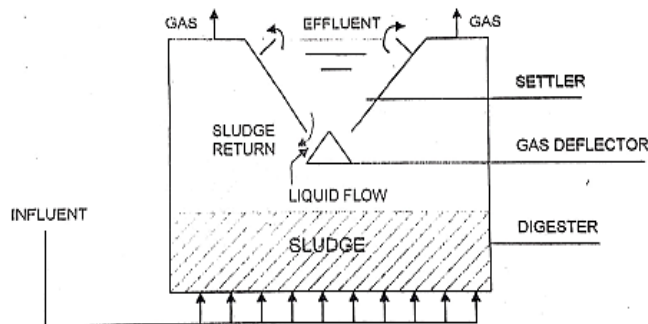


Fig. 2. Schematic of a UASB Reactor.

#### B. Primary Treatment

Primary treatment consists of Fine Screening and De-Gritting. Objective of Screening is to remove floating matters and other large size objects from the sewage stream. Objective of De-gritting is to remove grit particles from the sewage stream by gravity separation process. Screening and De-gritting are physical processes and they are accomplished in the units provided for the same with assistance of equipment provided. In the plant, nine automatic fine screening with 6 mm bar spacing are used. The cleaning of the screen is automated by mechanical means. Mechanical Fine Screen Channel always remains connected to the system, except during maintenance period. In the plant, three manual screening with a bar spacing of 12 mm (Standby for the automatic) are used. Six De-gritting Units with Grit Removal Mechanism and grit washing system are used for separation of gritty matter from sewage stream. Then, it is washed to make it free from organic matter and to transfer organic matter back to sewage stream. All the Grit Chambers are provided with Grit Collector, reciprocating type grit washing mechanism and Organic Return Pumps. These Grit Chambers remain in operation all the time. Three of them are manual De-gritting units for handling 50% of raw sewage flow and three of them are Parshall Flume with flow measurement followed by thirty Up-flow Anaerobic Sludge Blanket (UASB) Reactors as shown in Fig. 3. Inside the Reactors four reactions (Hydrolysis, Acidogenesis, Acetogenesis and Methanogenesis) account for the whole process. The system achieves a removal efficiency of 70%-80%, even when receiving organic loads greater than 15kg COD/m<sup>3</sup> of reactor per day at 8 hours HRT. The biogas is made up of 75%-85% methane. The sludge at the bottom has a concentration of about 40-70g Volatile Suspended Solids (VSS)/ l. In the plant, three Primary Sludge sump and pump house ultrasonic flow meters are provided in the Parshall flume for flow measurement with flow indicator, totalizer and recorder.

#### C. Secondary Treatment

In Secondary treatment, the refined water coming from primary treatment passes through three Pre-Aeration Tanks with six surface Aerators, further it has three Polishing ponds with two compartments having 18 Floating Aerators for polishing pond compartment 1 and eleven fountain pumps for polishing pond compartment 2.

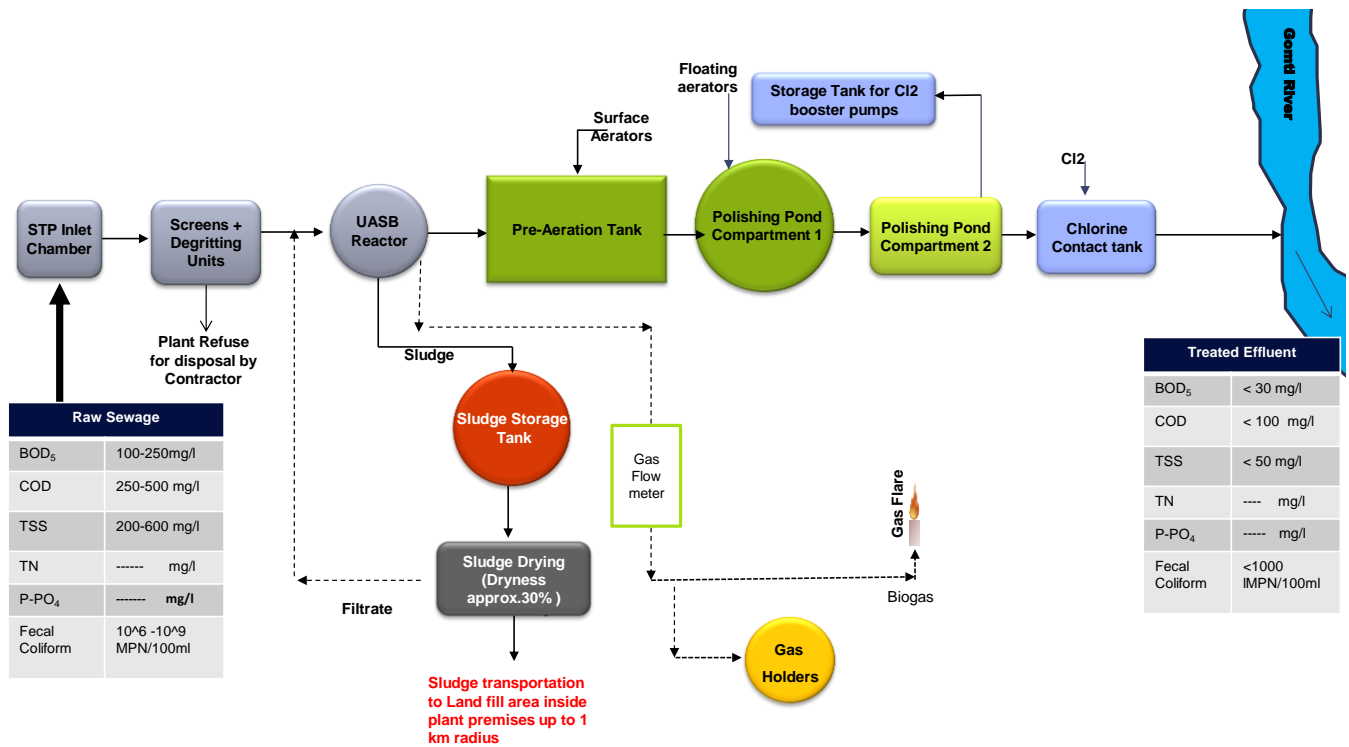


Fig. 3. Process Flow for 345 MLD UASB-based Bharwara STP.

#### D. Final Disinfection

In this process, three Chlorinator / one Chlorinator House is used. The chlorinator flow is manually adjusted. All the chlorinators are of the vacuum type. Hence, since as the chlorine gas is fed to the injector (located close to the chlorinators) at a pressure lower than atmospheric, no leak will occur. It has one water reservoir for chlorinating system and three Chlorine Contact Tank, One Final effluent chamber and One Final Effluent pipe.

The dewatering is carried out by means of Sludge drying beds. The 106 numbers of drying beds each of size 729 m<sup>2</sup> are feed by pumping from the SDB (Sludge Drying Bed) Feed sludge pumps. The consistency of dewatered sludge is around 30 – 35%.

The 345 MLD UASB-based Bharwara STP has following process objectives: reusable treated effluent, generating biogas according to raw sewage effluents and constantly delivering required quality of treated effluent.

In this section, we have briefly discussed major components of process flow for the plant.

In the next section, we present the methodology used to analyze and design a model to predict the parameters of influent/ effluent at wastewater treatment plant. This study shall provide foundation to improve reactor work performance in wastewater treatment plant.

#### IV. METHODOLOGY

We propose a machine learning based model to predict parameters of influent and effluent which shall provide efficient utilization of chemical resources during treatment

process ensuring the desired level of quality indicators in effluent. We collected real time dataset of 345 MLD UASB-based Bharwara STP using manual process for data analysis. The methodology for the proposed model is briefed using the following four steps:

##### A. Identification of Locations and Water Parameters to be captured at Plant

We along with supporting staff at 345 MLD UASB-based Bharwara STP identified five locations where the water parameters are to be captured. The placing of various locations in the plant are shown in Fig. 2. At each location, we identified and listed the water parameters like BOD, COD, DO, SS, temperature, pH, Residual Chlorine etc. be measured. The basis of identifying water parameters at a particular location in the plant is the process/ treatment/ chemical reactions taking place at these locations. These identified locations and respective parameters to be measured at these locations are listed in Table I.

TABLE I. LOCATIONS AND MEASURING PARAMETERS

Location	Parameters
Inlet Chamber	BOD, pH, Suspended Solids, Temperature, COD, oil, flow, Phosphorous, DO
Outlet of UASB Reactor	BOD, Suspended Solids, pH, COD
Polishing pond	Dissolved Oxygen, pH
Outlet of Chlorine contact Tank	BOD, Suspended solids, pH, COD, Fecal Coliform, Residual Chlorine, Dissolved Oxygen.
Primary sludge	pH, Total Solids, Volatile solids.

### B. Data Collection

We collected a real-time data set of the 26 months (April 2019 to May 2021) from the plant. In the data set, selected parameters of influent and effluent are collected/ captured and recorded using manual process adopted at the plant.

### C. Data Preprocessing

We do pre-processing on the recorded data set. For preprocessing, we treat missing values and outliers using standard procedures and kNN, and further normalized the data set. The outlier treatment is performed using statistical techniques i.e., calculating interquartile range and neglecting the values above lower limit and upper limit [11]. The normalization of the data set is performed using the following formula (1):

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

where  $x'$  is the normalized value,  $x$  is the original value, and  $\min(x)$  and  $\max(x)$  respectively are the minimum and maximum values. The data is normalized in the range between 0 and 1.

### D. Discovering Unknown Patterns

We discover various patterns or relations within the collected data sets. We visualize the patterns in the data set. We design and implement a machine learning-based model to analyze and predict the parameters of influent/ effluent in the wastewater treatment Plant. We used Linear Regression to design the preliminary prediction model. Linear regression [12] is a statistical tool for the prediction of a dependent variable from an independent variable. It establishes a linear relationship between the independent (input) and dependent (output) variables. Linear Regression is a modeling technique where a dependent variable is predicted based on the independent variables. Linear Regression is the most widely used technique among all statistical techniques. The linear regression model is designed on Google Colab using python 3.7.12 for performing analysis.

Let us discuss the dependent variable, independent variable, line of regression, data preprocessing, model properties for the linear regression model.

**Dependent variable:** It is a variable that depends on other factors (independent variables) that are measured.

**Independent variable:** It is the variable [13] that is stable and unaffected by another variable which we are trying to measure independent variables (predictors) are used to predict the value of the dependent variable (target variable).

**Line of regression model:** It is the relationship between independent and dependent variables.

**Model Properties:** We implemented the initial model using Linear Regression in Python Implementation environment for the model is given in Table II.

TABLE II. IMPLEMENTATION ENVIRONMENT

<b>Language</b>	Python (version 3.7.12)
<b>Tool</b>	Google Colaboratory
<b>Libraries</b>	NumPy, Pandas, Matplotlib, Scikit Learn, SciPy and Seaborn

The model Properties are as follows:

- Model inputs: Inlet COD, BOD, PH, TSS, MLD and MPN
- Model outputs: Outlet COD, BOD, PH, TSS, DO, MPN
- Training - Test split: 80/20
- Estimator Function: Mean Square Error

In order to measure the performance of the model, Mean Square Error (MSE) is used. Formula for MSE [14] is given as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

We improved and tested the performance of the model by minimizing MSE and maximizing the correlation coefficient (R).

## V. RESULTS AND ANALYSIS

We collected the raw data from 345 MLD UASB-based Bharwara STP during April 2019 to May 2021. The summary of the collected data, analysed using Python is shown in Table III. Based upon the initial analysis, it is found that the data has some missing facts/ details/ values and the outliers under few variables. Therefore, we applied Mean method and KNN to treat missing values in the given dataset. Fig. 4 shows the results after treating missing values on OUT\_MPN. However, the similar results are obtained for the other variables (columns) with missing values in the dataset.

The raw data contained the outliers which were impacting/ decreasing the efficiency of the model(s) created. So, the treatment for removing the outliers from the data is done during pre-processing. The Fig. 5 shows the graph for inlet BOD before the outlier treatment and Fig. 6 shows the graph for inlet BOD after outliers' treatment. Similarly, we did the outlier treatment for other variables (columns) in the data set.

After doing the missing values treatment and outliers' treatment, the data set is used for further data visualization. We obtained the summary of the pre-processed dataset as shown in Table IV.

Further, we did data visualization and selected the influent parameters and obtained the results. The obtained results are elaborated in this section.

TABLE III. SUMMARY OF RAW DATA BEFORE PRE-PROCESSING

	DATA QUALITY PARAMETERS											
	Month Count	MLD	Influent PH	Effluent PH	Influent TSS	Effluent TSS	Influent COD	Effluent COD	Influent BOD	Effluent BOD	Influent MPN	Effluent DO
COUNT	781	780	780	780	780	780	780	780	774	774	572	780
MEAN	13.31	324.8	7.342	7.58	233.38	41.97	279.18	74.05	146.12	26.13	533836.14	4.58
STD	7.40	40.84	0.13	0.11	37.81	3.96	52.03	8.66	21.22	1.73	776153.06	1.41
MIN	1	157.11	6.93	7.23	145	4	125	28	75	19	1.4	4
25%	7	304.14	7.25	7.52	212	39	240	68	135	25	11	4.3
50%	13	344	7.34	7.6	230	42	264	72	148	26	780000	4.5
75%	20	349.79	7.43	7.66	254	45	324	80	160	28	920000	4.8
MAX	26	437.93	7.75	7.87	556	72	528	96	220	29	14000000	43

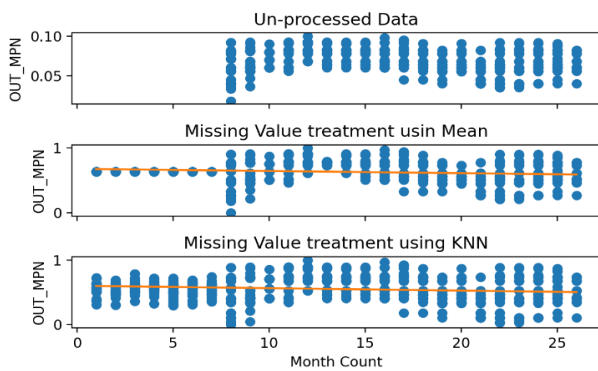


Fig. 4. Preprocessing of OUT\_MPN.

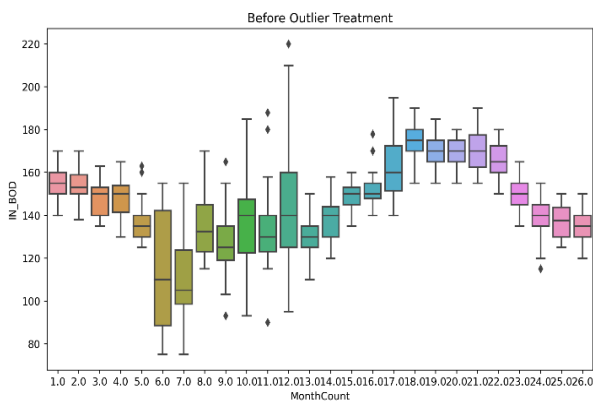


Fig. 5. Before Outlier Treatment.

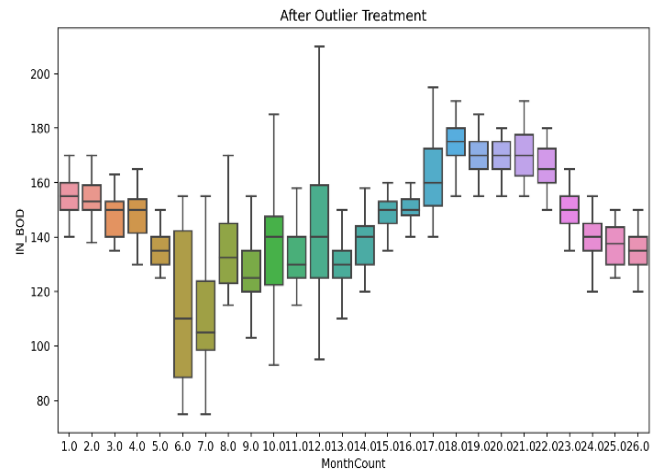


Fig. 6. After Outlier Treatment.

Fig. 7 shows the flow rate (in MLD) of influent with respect to month. In Fig. 8, we can clearly observe the inconsistency in the flow. Fig. 9 shows day wise flow of influent in the month of January of 2020 and 2021. For the month of February 2020, the flow is around the 7000 million litres but it rises too nearly 12000 million litres in the month of July. Also, the same months of different years have shown the major differences in the data which can be clearly observed in Fig. 9. We have further created the linear regression model for prediction of flow in 345 MLD UASB-based Bharwara STP/ WWTP.

TABLE IV. SUMMARY OF DATASET AFTER PRE-PROCESSING

	DATA QUALITY PARAMETERS												
	Month Count	MLD	Influent PH	Effluent PH	Influent TSS	Effluent TSS	Influent COD	Effluent COD	Influent BOD	Effluent BOD	Influent MPN	Effluent MPN	Effluent DO
COUNT	773	773	773	773	773	773	773	773	773	773	773	773	773
MEAN	13.23	326.2	7.34	7.58	231.51	42.01	276.81	74.01	145.77	26.14	10.31	0.06	4.53
STD	7.34	39.66	0.13	0.11	34.21	3.57	50.79	8.24	20.72	1.72	1.72	0.01	0.31
MIN	1	157.1	6.98	7.23	145	32	125	56	75	19	5.61	0.033	4
25%	7	306.8	7.25	7.52	212	39	240	68	135	25	9.2	0.061	4.3
50%	13	345.2	7.34	7.61	227	42	264	72	145	26	10	0.06	4.5
75%	20	350	7.43	7.66	251	45	320	80	160	28	11.21	0.08	4.8
MAX	26	403	7.75	7.87	443	49	464	96	210	29	15.51	0.1	5.3

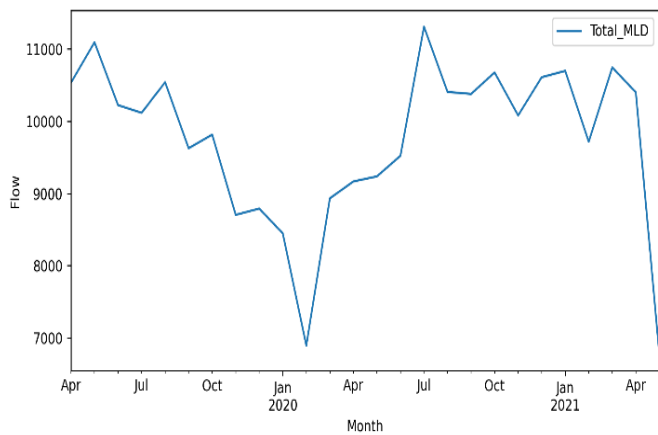


Fig. 7. Flow Rate of Influent by Month.

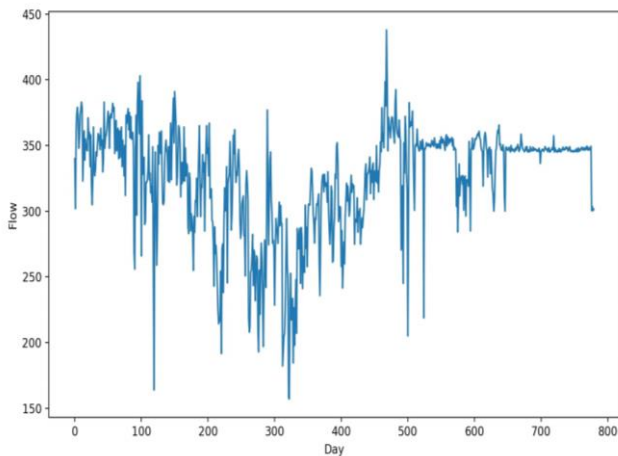


Fig. 8. Flow Rate of Influent by Day.

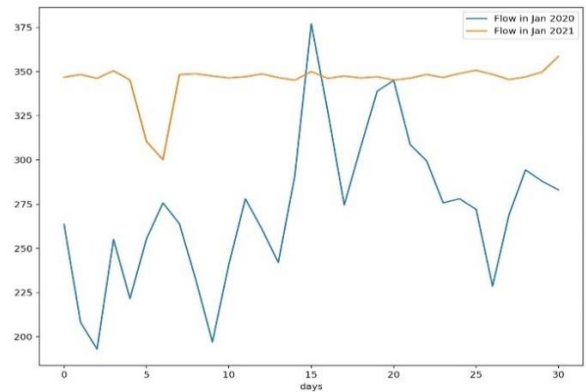


Fig. 9. Comparison of Flow Rate between Jan 2020 and Jan 2021 by Day.

Fig. 10 signifies the effectiveness of WWTP by showing the relationship between influent (untreated water) and Effluent (Treated and processed) water. The parameters include Total Suspended Solids (TSS), Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), Myeloproliferative Neoplasms (MPN), Dissolved Oxygen (DO) and pH.

These graphs in Fig. 10, show that there is a great fluctuation/ variation in the influent parameters which are the main factors affecting the efficiency of the plant. Therefore, a prediction by a Machine Learning model shall greatly help in managing and enhancing the quality and effectiveness of waste water treatment processes used in the plant.

The relationship between parameters can be analysed by the correlation coefficient. It can be used to obtain the effectiveness of the relationship among the parameters and can be used for further analysis and modelling. The positive correlation signifies that if one value increases another also increases, higher value shows the stronger correlation.

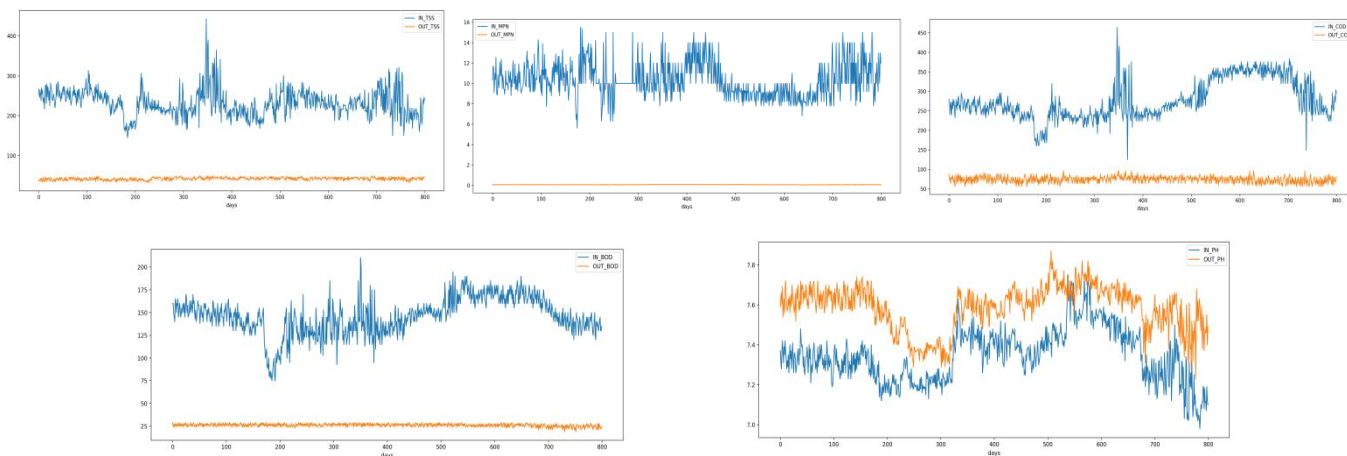


Fig. 10. Relationship between Influent and Effluent Parameters.

Fig. 11 shows the result of the linear regression model designed and implemented for prediction of effluent parameters (BOD, TSS, MPN, DO, and COD), here red dots show actual results and blue cross show predicted values of the linear regression model. The Model is trained on considering each influent parameter (BOD, TSS, MPN, DO, and COD) as input variables and the selected effluent

parameter as output variable. The initial model is showing adequate results.

In Fig. 12, heatmap shows this relation with the intensity of the colour used; darker colour shows the stronger relationship. The colour turning to blue shows the negative relationship means the increase in one value will lead to the decrease in another.

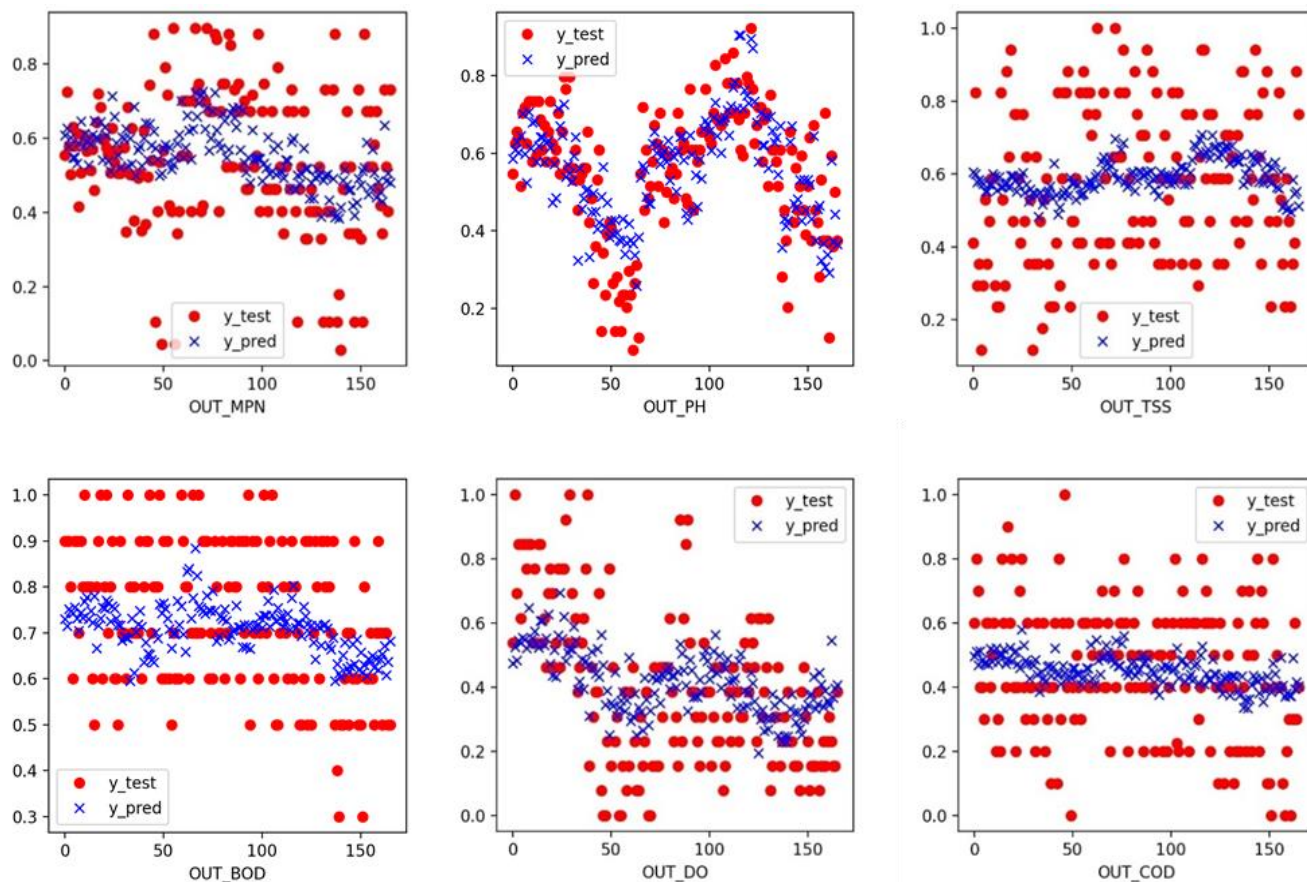


Fig. 11. Linear Regression Output.



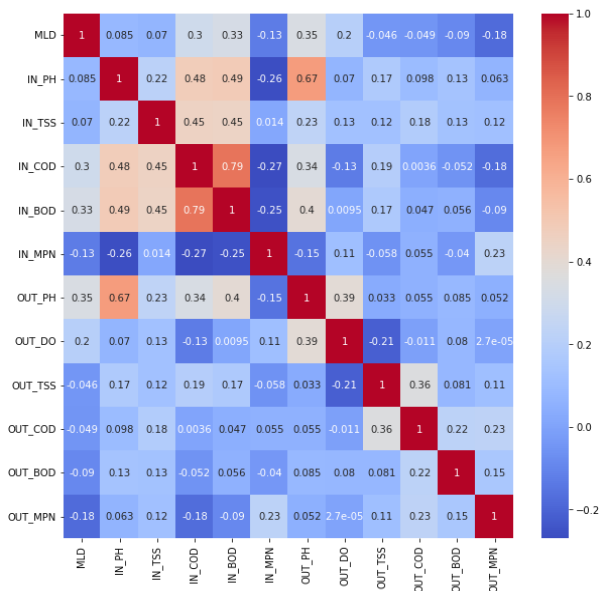


Fig. 12. Heatmap for Correlation.

Results of linear regression model, obtained in the form of MSE and Correlation Matrix (R), are shown in Table V. We are able to improve the efficiency of the initial model up to some extent using the linear regression model.

TABLE V. MSE TABLE AND CORRELATION

MET HOD	DATA QUALITY PARAMETERS					
	EFFLU-ENT PH	EFFLU-ENT DO	EFFLU-ENT TSS	EFFLU-ENT COD	EFFLU-ENT BOD	EFFLU-ENT MPN
MSE	0.015	0.048	0.046	0.038	0.028	0.033
R	0.736	0.413	0.211	0.249	0.239	0.436

## VI. CONCLUSION

Authors have analyzed the flow and quality parameters like COD, BOD, TSS, DO, pH, Temperature, Ammonia, Phosphorous and oil content, etc. in influent, and also parameters like COD, BOD, DO, pH, etc. of effluent in the WWTP. The proposed model provides support to centrally monitor processes and operations of a wastewater treatment plant. This paper depicts and visualizes the fluctuating and varying nature of influent parameters in 345 MLD UASB-based Bharwara STP.

The analysis presented in the paper, provides basis for improving operational efficiency and provides a cost-effective utilization of various resources at wastewater treatment plants by knowing about the pattern of the influent and effluent parameters may be in advance also. We have also designed and implemented an initial model using the Linear Regression algorithm to analyze as well as predict the parameters of influent and effluent of 345 MLD UASB-based Bharwara STP. However, the implemented model shall be applicable for

any UASB based wastewater treatment plant or any wastewater treatment plant after a specific training part.

In future work, the authors shall incorporate SVM (Support Vector Machine) and ANN (Artificial Neural Network) techniques in the model to predict the influent and effluent parameters in WWTPs. It is expected that the model incorporating SVM and ANN shall show more robust relationship among the parameters and give a better estimate than the current model, in future.

## REFERENCES

- [1] Gautam Rajneesh, Verma Saumya, More Nandkishor. "Sewage Generation and Treatment Status for the Capital City of Uttar Pradesh, India", Avicenna Journal of Environmental Health Engineering. vol. 5, no. 1, pp. 8-14, 2018.
- [2] Tümer Abdullah, Edebalı Serpil. (2015). An Artificial Neural Network Model for the Wastewater Treatment Plant of Konya. International Journal of Intelligent Systems and Applications in Engineering. vol. 3, no. 4, pp. 131-135, 2015.
- [3] Guo Hong, Jeong Kwanho, Lim Jiyeon, Jo Jeongwon, Kim Young, Park Jong-pyo, Kim Joon Ha, Cho Kyung, "Prediction of effluent concentration in a wastewater treatment plant using machine learning models", Journal of Environmental Sciences, vol. 32, pp. 90-101, 2015.
- [4] Matala Anna, "Sample size requirement for Monte Carlo simulations using Latin hypercube sampling.", Helsinki University of Technology, Department of Engineering Physics and Mathematics, Systems Analysis Laboratory, 2008.
- [5] Granata Francesco, Papirio Stefano, Esposito Giovanni, Gargano Rudy, DE MARINIS, Giovanni., "Machine Learning Algorithms for the Forecasting of Wastewater Quality Indicators", Water, vol. 9, no. 2, pp. 105, 2017.
- [6] Qin Xusong, Gao Furong, Chen Guohua, "Wastewater quality monitoring system using sensor fusion and machine learning techniques", Water research, vol. 46, no. 4, pp. 1133-1144.
- [7] Wang Rui, Pan Zhicheng, Chen Yangwu, Tan Zhouling, Zhang J. "Influent Quality and Quantity Prediction in Wastewater Treatment Plant: Model Construction and Evaluation", Polish Journal of Environmental Studies, vol. 30, no. 5, 20.
- [8] D S, Manu, Thalla Arun, "Artificial Intelligence Models for Predicting the Performance of Biological Wastewater Treatment Plant in the removal of Kjeldahl Nitrogen from Wastewater", Applied Water Science, vol. 7, no. 7, pp. 3783-3791, 2017.
- [9] Gautam, Sandeep Kumar, Divya Sharma, Jayant Kumar Tripathi, Saroj Ahirwar, and Sudhir Kumar Singh. "A study of the effectiveness of sewage treatment plants in Delhi region." Applied Water Science, vol. 3, no. 1, pp. 57-65, 2013.
- [10] Banerjee, Arif Siddiquiel Rajiv. "Performance Evaluation & Upgradation of UASB Technology used for the Treatment of Sewage Generated from Lucknow City.", 2016.
- [11] Alnaa, Samuel, Ahiakpor, Ferdinand, "ARIMA (autoregressive integrated moving average) approach to predicting inflation in Ghana", Journal of Economics and International Finance, vol.3, pp. 328-336, 2011.
- [12] Kumari, Khushbu, and Suniti Yadav. "Linear regression analysis study." Journal of the practice of Cardiovascular Sciences, vol. 4, no. 1, pp. 33, 2018.
- [13] Cramer, Duncan, and Dennis Laurence Howitt. The Sage dictionary of statistics: a practical resource for students in the social sciences, Sage, 2004.
- [14] Available at: <https://www.datavedas.com/model-evaluation-regression-models/>.
- [15] Sin, Gürkan, Krist V. Gernaey, Marc B. Neumann, Mark CM van Loosdrecht, and Willi Gujer. "Uncertainty analysis in WWTP model applications: a critical discussion using an example from design.", Water Research, vol.43, no. 11, pp. 2894-2906, 2009.