International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 12, ISSN No. 2455-2143, Pages 142-146
Published Online April 2023 in IJEAST (http://www.ijeast.com)

ANALYSIS OF EFFICIENCY OF SEWAGE TREATMENT PLANT USING DATA SCIENCE

Sourabh Sarpanchal, Tejashri Agale, Pratika Jadhav
Dept. of IT/ Computer Engg.
Government College of Engineering and Research Avasari Kh, Pune
Maharashtra, India

Abstract—**As a study location, The Sewage Treatment Plant (STP) of New Naidu was chosen. The facility is close to Naidu Hospital. It can currently hold 115 MLD. The recorded 304 days of daily inflow data from the New Naidu Sewage Treatment Plant(From 1June 2021 to 31 March 2022) that was read by the flow meter on the daily basis was collected in order to undertake time series analysis and forecasting the inflow of wastewater. The average daily inflow was recorder& the data that was obtained was then utilised to build the Autoregressive Integrated Moving Average Model (ARIMA) Model. The STP provided the lab report samples. Value of the various parameters, including pH, Total Suspended Solids (TSS), Chemical Oxygen Demand (COD), and Biological Oxygen Demand (BOD), which have a standard, were obtained to determine whether the water was reusable by making use of different classification techniques.**

Keywords—**STP, ARIMA Model, COD, BOD, TSS.**

I. INTRODUCTION

The Mula-Mutha river, situated in India, is facing severe pollution due to various reasons such as inadequate sewerage systems, construction, and material, serial and sewerage discharge in polluted banks. As a result, the quality of the river water has significantly deteriorated over time, leading to a number of waterborne illnesses such as cholera, diarrhoea, typhoid, hepatitis, and other skin conditions. The contamination of the river with pollutants also increases the temperature of the water, which encourages the growth of harmful algae and results in a decrease in dissolved oxygen, ultimately leading to the death of aquatic life.

To combat this problem, the Pune Municipal Corporation (PMC) has installed Sewage Treatment Plants (STPs) at various locations across the city to prevent further pollution of the river. However, it is crucial to check the effectiveness of these STPs regularly to ensure regulatory compliance, minimize environmental impacts, protect public health, and maintain operational performance. This is important because if the STPs become overburdened with wastewater due to population growth, PMC may have to build additional STPs or expand the capacity of existing treatment facilities.

At present, the New Naidu Sewage Treatment Plant, which is the study area, has a capacity of 115 MLD (million litres per day). The project aims to evaluate the effectiveness of the STPs and predict future wastewater inputs to ensure that the STPs can continue to function efficiently, and the Mula-Mutha river can be restored to its original state.

II. RELATED WORK

Some of the previously carried out papers which proved to be very insightful for our project are enlisted here. N. Bolong, M. Q. N. Sabli, I. Saad and A. N. A. Ali [1], This study evaluated the efficiency of a sewage treatment plant (STP) at University Malaysia Sabah (UMS) based on Malaysian guidelines. Results showed that the STP design complied with guidelines, and peak usage resulted in higher influent pollutants. Effluent quality was generally within limits, except for occasional excess of oil and grease. The STP achieved consistent 80% removal efficiency for BOD. Effluent reuse for agriculture and landscape irrigation is proposed due to potential environmental benefits. Jorgen Thorn, Erika Kerekes [2], This paper reviews the health risks for employees working in sewage treatment plants. Exposure to microorganisms and chemicals in this environment can lead to infections, with hepatitis A being a major concern. Gastrointestinal, respiratory, and other symptoms have been reported among workers, potentially caused by inflammation from endotoxins. While some studies suggest an increased risk of certain cancers, further investigation is needed to determine the causal agents and work-related effects.M. Bhargavi, E. Ananta Rao, T. Pravallika,Y. Sri Teja [3], This study aims to design a sewage treatment plant for Vizianagaram municipality due to a steady increase in population resulting in excess sewage. The plant includes various components such as screens, grit chamber, sedimentation tank, and sludge drying beds, following standards and permissible limits. The treated water will be used for irrigation, and sludge will be used as manure to increase soil fertility and reduce groundwater usage.Rob Jamieson, Wendy Krkosek, Leah Boutilier and Graham Gagnon [4],The Canadian Municipal Wastewater Effluent Strategy was developed to provide a

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 12, ISSN No. 2455-2143, Pages 142-146
Published Online April 2023 in IJEAST (http://www.ijeast.com)

framework for managing wastewater, with the Far North given a 5-year window to research and develop feasible standards. Wastewater treatment in Nunavut is regulated through different forms of legislation, with waste stabilization pond and lake treatment systems being the most common due to capacity challenges for mechanical plants. Treatment performance in cold climates is highly variable and site-specific, with very little peer-reviewed literature documenting passive system treatment performance in arctic climates. Design guidelines and approaches for modeling the treatment performance of passive systems operating in arctic regions have yet to be produced. Niraj S. Topare, S. J. Attar and Mosleh M. Manfe [5], It is crucial for
users to ensure that their Sewage/Wastewater treatment process adheres to regulatory guidelines. The objective of the treatment process is to eliminate various constituents of the polluting load, including solids, organic carbon, nutrients, inorganic salts, metals, and pathogens. Proper wastewater collection and treatment are vital for both environmental protection and public health. Wastewater management involves different methods to reduce water and organic content, with the primary goal being the protection of the environment while considering public health and socio-economic concerns. This article provides a brief discussion of Sewage/Wastewater treatment techniques, the factors that influence system selection and design, and their importance.Puspalatha P., Kalpana P [6], the population growth in Srikakulam has led to an increase in domestic sewage production, and there is currently no sewage treatment plant in place. It is therefore necessary to construct a plant with sufficient capacity to manage the generated sewage, although there are challenges associated with excess sludge treatment and disposal. The main goal of sewage treatment is to reduce the water and organic content of the waste, while also protecting the environment and public health. This paper discusses the sewage generation in Srikakulam and proposes the design of a sewage treatment plant to handle an estimated 22.2 MLD of sewage per day, based on projected population growth over the next 30 years. The plant will include various components such as screening, grit chamber, primary sedimentation tank, biological reactor, secondary clarifier, activated sludge tank, and drying beds, all designed to meet applicable standards and permissible limits for treated sewage water. The treated water will be used for crop irrigation, while the sludge generated from the treatment process can be used as manure to enhance soil fertility and reduce groundwater usage.

III. METHODOLOGY

For the process of forecasting the inflow of waste water in a sewage treatment plant, the wastewater engineers face numerous challenges in form of non-linear precipitation-runoff relationships, unpredictability due to ageing infrastructure and inconsistency in data monitoring and data quality.

Moreover, considering the laboratory tests performed on the treated water, the sewage treatment plant determines the reusability of waste water based on various parameters which have standard and pre-defined values.

Generic Methodology:

1. To address challenges related to inflow of water, time series analysis models such as ARIMA (Auto regressive integrated moving average) and ARTIFICIAL NEURAL NETWORK MODELS (the multilayer perceptron neural network) are generally used. The performances of the respective models are evaluated using ROOT MEAN SQUARE ERROR, MEAN SQUARED Errored
2. To estimate whether the treated water is reusable or not, different classification techniques are employed. Principal Component Analysis, Decision Trees, Logistic Regression models, Naive Bayes classifier are some of them. Some methodologies also include determining WATER QUALITY INDEX using weighted arithmetic index method

Proposed Methodology:

1. The proposed system makes use ARIMA model in order to forecast the inflow of waste water based on historical values of data.

There are 4 components of Time Series Analysis – trend, seasonal variations, cyclic variations and irregular variations.

a. Augmented Dickey- Fuller Test:

For forecasting, time series must be stationary. Here, we make use of Augmented Dickey- Fuller Test in order to check the stationarity of data.

The data that is obtained does not follow any trend. Moreover, it has been observed that the data is stationary and does not include any seasonal variations.

b. Auto Correlation function (ACF)

This plot summarises the correlation of an observation with its lag values. The x- axis shows the lag values and the y-axis shows the correlation coefficient.

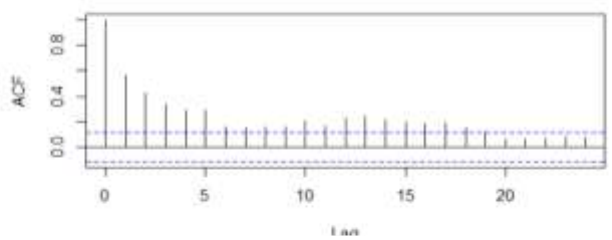

Fig.1. Series INFLOW(ACF)

c. Partial Auto correlation Function (PACF):

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 12, ISSN No. 2455-2143, Pages 142-146
Published Online April 2023 in IJEAST (http://www.ijeast.com)

This plot summarises the correlation of an observation with its lag values that is not accounted for by prior lagged observations.

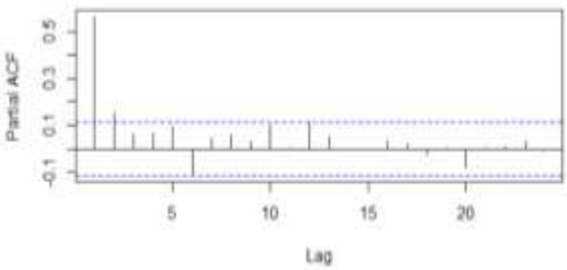

Fig.2. Series INFLOW (PACF)

The ACF and PACF plots should be considered together to define the process. The above plots show geometric decreasing pattern.

d. Model-fitting:

```
Fitting models using approximations to speed things up...

ARIMA(2,0,2) with non-zero mean : Inf
ARIMA(0,0,0) with non-zero mean : 1965.568
ARIMA(1,0,0) with non-zero mean : 1892.571
ARIMA(0,0,1) with non-zero mean : 1915.624
ARIMA(0,0,0) with zero mean     : 3028.958
ARIMA(2,0,0) with non-zero mean : 1890.713
ARIMA(3,0,0) with non-zero mean : 1893.208
ARIMA(2,0,1) with non-zero mean : 1890.507
ARIMA(1,0,1) with non-zero mean : 1889.078
ARIMA(1,0,2) with non-zero mean : 1890.839
ARIMA(0,0,2) with non-zero mean : 1898.61
ARIMA(1,0,1) with zero mean     : Inf

Now re-fitting the best model(s) without approximations...

ARIMA(1,0,1) with non-zero mean : 1889.108

Best model: ARIMA(1,0,1) with non-zero mean
```

Hence, considering the type and constraints of the data used as input, ARIMA (1, 0, 1) with non-zero mean model has been deployed.

e. Holt – Winters Forecasting:
It is used for exponential smoothening for level, trend, and seasonal components. As trend and seasonal components are absent in the data used in the proposed system, hence simple exponential smoothing is used for removing irregularities.

2. To estimate whether the treated water is reusable or not, firstly, exploratory data analysis was performed on data.
a. Parameter Testing :

Shapiro- Wilk Normality Test :

Significance value = 0.05
H_0 : pH is normally distributed
H_1 : pH is not normally distributed

It was found that the p- value for each test was less than 0.05 hence, the null hypothesis is rejected.
The output of the Shapiro Test shows that the data does not follow Normal Distribution. This is an indication to opt for non-parametric tests.

Wilcoxon' Signed Rank Test:
The results of Wilcoxon's test show that all parameters are in given range. Hence, the given data is non-parametric.

b.Heatmap:
Heatmap is used as a correlation matrix in order to determine the strength of relationships between the parameters. It is also used in detecting outliers and linear and non-linear relationships.
To check the multicollinearity between regressors, heatmap is employed.

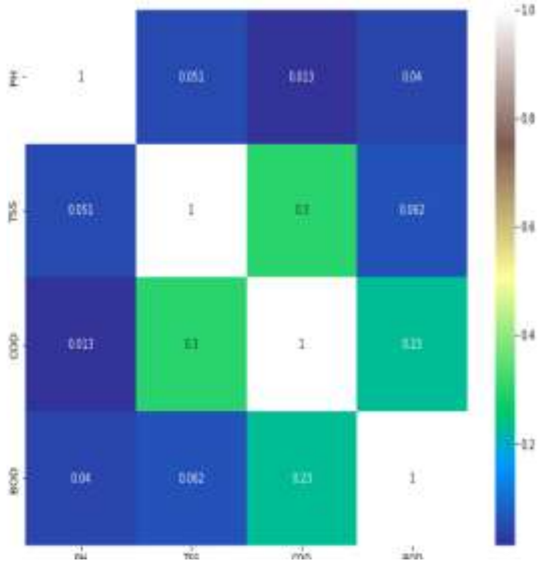

Fig.3. Heatmap

3. Further analysis has shown that the data is not normalized. Therefore, there were certain limitations with respect to the classification models that could be employed. The proposed system employs the following models for classification:

a. Logistic Regression Model:
Examines the association between one or more independent variables and a binary dependent variable.

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 12, ISSN No. 2455-2143, Pages 142-146
Published Online April 2023 in IJEAST (http://www.ijeast.com)

b. Decision Trees:
For making decisions by mapping possible consequences of choices and predicting outcomes.

c. K - nearest neighbour model:
A classification method that assigns a class label to a new data point by considering the class labels of the k-nearest neighbours in a training set.

d. Naive Bayes Classifier:
An algorithm that calculates the probability of a given class using Bayes' theorem and assumes that the features are conditionally independent.

IV. RESULT ANALYSIS

Based on the assessment of STP effectiveness and prediction of wastewater input, it was determined that additional STPs or expansion of existing facilities may be required in the future due to population growth. The treated water was tested for various factors including pH, TSS, COD, and BOD, and it was found that the values of these parameters were within the specified range.

Table.1.1. Criteria for Water Quality Assessment

Parameter	pH	COD	BOD	TSS
Standard	6.5-8.5	50 mg/L	10 mg/L	20 mg/L

[Source: http://pmc.gov.in/]

The flow rate and concentration of water can vary throughout the day and across seasons, with certain patterns emerging. For instance, in the morning, both flow rate and BOD values tend to be higher. During the rainy season, surface run-off can increase the amount of solid contents in the water.
After performing model fitting to determine the best fitting model for these values, it was concluded that the treated water is reusable. This is a positive outcome as it means that the treated water can be used for various purposes such as irrigation, which can reduce the demand for freshwater resources.
Overall, the results and analysis indicate that the STP is effective in treating wastewater and the treated water is of good quality, making it suitable for reuse. However, it is important to continue monitoring and assessing the performance of the STP in order to ensure that it continues to operate effectively and efficiently.

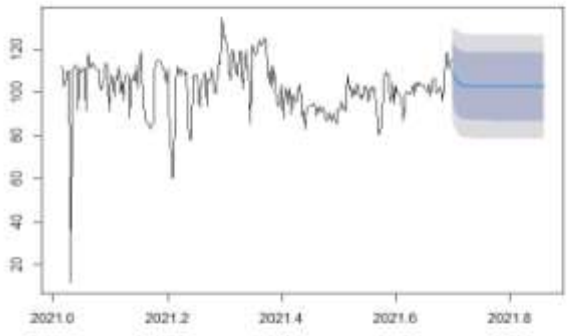

Fig.4. Forecasts from ARIMA (1,0,1) with non-zero mean

The graph shown above predicts the average inflow of wastewater for next 84 days. We can determine that there is no trend being followed in the inflow of water in next 84 days. The inflow is more or less the same.

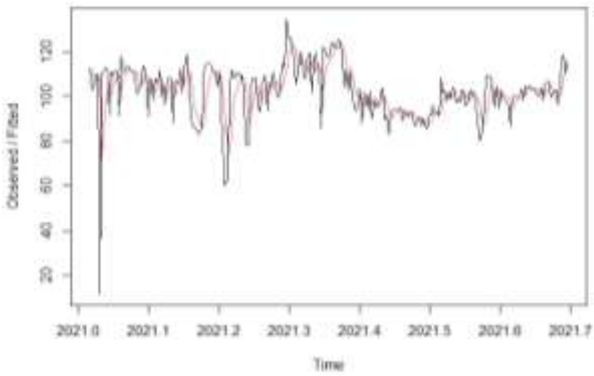

Fig.5. Forecasts from Holt-Winters Forecasting Model

Holt Winters forecasting model deployed for the purpose of forecasting uses exponential smoothening parameters to reduce irregularities in data. However, since our data does not follow
any trend or regularity, we have used simple exponential smoothening.

Table.1.2. Models and their predicted accuracy

Sr. No.	Model	Accuracy
1.	Logistic Regression	97.61%
2.	Decision Tree	97.13%
3.	K-Nearest Neighbour	93.68%
4.	Naïve Bayes	92.74%

V. CONCLUSION

To summarize, this STP project aimed to understand the sewage inflow variations of New Naidu STP using time series and ARIMA models, as well as the use of the Holt-Winters model for forecasting. The results from these models will aid in monitoring sewage load and planning for future STPs. Additionally, the project utilized hypothesis testing to determine the efficiency of the STPs in treating wastewater for reuse. Machine learning techniques were also employed to classify water quality based on pH, TSS, BOD, and COD parameters, using supervised classification models such as decision trees, KNN, and Logistic Regression. Overall, the findings from this project can help inform better water management practices and ensure water quality meets standard ranges set by standard regulatory authorities and laboratory reports.

VI. ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to everyone who has supported us in the successful completion of our research paper "Analysis of Efficiency of Sewage Treatment Plants Using Data Science".

Firstly, we extend our special thanks to our project guide, Dr. S. U. Ghumbre, for his able guidance and support throughout the project. His knowledge and expertise in the field of data science wasinstrumental in shaping our research and bringing it to fruition.

We would also like to thank our faculty for their guidance and constant supervision. Their valuable inputs and feedback helped us to refine our research methodology and ensure the accuracy and validity of our results.

We would also like to offer our sincere gratitude to our principal, Dr. D. R. Pangavhane, for providing us with all the necessary facilities and resources needed for our research. We are grateful for their continued support and encouragement throughout the project.

Furthermore, we would like to thank our parents for their unwavering support and encouragement throughout our academic journey. Their motivation and belief in us have been our driving force and have inspired us to work harder and achieve our goals.

We would also like to express our gratitude to our friends who helped us in finalizing this project within a limited time frame.

Lastly, we are thankful to our classmates whose instructions and suggestions were helpful in shaping our research andimproving its quality. We appreciate their collaboration and support throughout our academic journey.

Once again, we express our sincere gratitude to everyone who has supported us in the successful completion of our research paper. Regards.

VII. REFERENCES

[1]. Bolong N., SabliM. Q. N., Saad I. , and Ali A. N. A.“The efficiency of sewage treatment plant: A case study at the main campus of University Malaysia Sabah (UMS).” In 2022IOP Conf. Series: Materials Science and Engineering, Volume: 1229, Issue: 3.2022.

[2]. Thorn Jorgen, KerekesErika. “Health effects among employees in sewage treatment plants: A literature survey.” In American journal of industrial medicine, Volume: 40, Issue: 8.2001.

[3]. Bhargavi M., Rao Ananta E., Pravallika T., Teja Sri.Y.“Analysis and Design of Sewage Treatment Plant: A Case Study on Vizianagaram Municipality.”InSSRG International Journal of Civil Engineering (SSRG - IJCE), Volume: 5, Issue: 4.2018.

[4]. Jamieson Rob, KrkosekWendy,BoutilierLeah, Gagnon Graham, “Literature Review of Wastewater Treatment Design and Performance in the Far North.”In Centre of Water Resources Studies, Dalhousie University, Halifax, Nova Scotia, Issue:11.2016.

[5]. TopareNiraj S.,Attar S. J. and ManfeMosleh M., “Sewage/Wastewater Treatment Technologies: A Review.” In Department of Chemical Engineering, Bharti Vidyapeeth University College of Engineering, Volume:1, Issue: 10.2011.

[6]. P. Puspalatha, P.Kalpana “Design Approach for Sewage Treatment Plant: A Case Study of Srikakulam Greater Municipality, India.” In Environmental Engineering, GMR Institute of Technology, Rajam, AP, India, Vol:12 Iss: 9.2016.

[7]. Salvi Sahil, Patil Premchandra, “A Case Study on Sewage Treatment Plant” In International Journal of Creative Research Thoughts, India, Vol:9, Iss:4.2021.

[8]. Gupta Shreya, Singh SK, Gandhi Vishal, “A study on Sewage Treatment and Disposal in Delhi” In International Journal of Advance Research and Innovation, India,Vol:6, Iss:2018.

[9]. Khan Maaz Allah, Akhtar Mohd Daud, “Sewage Treatment Plant (Stp)”In International Journal of Advance Research and Management, Lucknow, UP, Vol:2,Iss:4.2017.

[10]. Dixit Dipanshu, Verma Manaswee, Pathak Diksha, Agrawal Antima, “Comparative Case Study of Sewage Treatment Plant” In International Journal of Creative Research Thoughts, India, Vol:8, Iss:6.2020.

[11]. Elangovan G.,Rajanandhini V.M., “Analysis and Design of Sewage Treatment Plant: A Case Study Atnagore” In International Research Journal of Multidisciplinary Technovation, India, Vol:6 ,Iss:11.2019.

[12]. AswathyM. , Hemapriya, “Analysis and design of Sewage Treatment Plant of Apartments in Chennai” In India, Vol: 13,Iss: 2017.

[13]. harathiBattu, Prof. PolasaMurthy, “Design of Sewage Treatment Plant for a Gated community” In Hyderabad, India, Vol: 3, Iss: 2014.