



**FACULTY OF ENGINEERING AND TECHNOLOGY**

**DEPARTMENT OF WATER RESOURCES ENGINEERING**

**PREDICTING EFFLUENT CHEMICAL OXYGEN DEMAND (COD) IN  
INDUSTRIAL WASTEWATER USING ARTIFICIAL NEURAL NETWORKS  
(ANNs): A CASE STUDY OF KAKIRA SUGAR EFFLUENT TREATMENT PLANT  
(ETP)**

**BY**

**MWASE BRIAN**

**BU/UG/2021/0310**

**SUPERVISOR: MR. MUYINGO EMMANUEL**

**This Final Year Project Report is submitted to the Department of  
Water Resources Engineering in Partial Fulfillment of the  
Requirement for the Award of the Degree of Bachelor of Science in  
Water Resources Engineering of Busitema University.**

# ABSTRACT

This Study explores the application of Artificial Neural Networks (ANN) to enhance wastewater treatment processes at Kakira Sugar's Effluent Treatment Plant (ETP). Leveraging three years of historical data combined with daily monitoring of key operational parameters—including FlowRate, pH, Total Suspended Solids (TSS), temperature, Total Dissolved Solids (TDS), and Electrical Conductivity (EC)—the study developed an ANN model capable of accurately predicting effluent Chemical Oxygen Demand (COD) levels. The model was constructed using MATLAB's Neural Network Toolbox and trained with a 5-fold cross-validation approach, ensuring robust generalizability and minimizing overfitting through iterative weight adjustments via the Levenberg-Marquardt algorithm. Performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination ( $R^2$ ) were used to evaluate the model's accuracy, revealing strong predictive performance on training data and highlighting areas for improvement in validation and testing phases. Sensitivity analysis using Monte Carlo Simulation (MCS) further identified FlowRate, pH, and temperature as the most influential factors affecting COD predictions. The outcomes of this project underscore the potential of integrating ANN into wastewater treatment operations, offering a promising tool for real-time process optimization, regulatory compliance, and enhanced resource management. Future research is recommended to expand datasets, refine model generalization, and explore hybrid modelling approaches to further elevate the predictive capabilities and operational efficiency of wastewater treatment systems.

**Keywords:** Artificial Neural Networks (ANNs), Chemical Oxygen Demand (COD), Sensitivity Analysis, Monte Carlo Simulation (MCS)

## DECLARATION

I **MWASE BRIAN** hereby declare that, the work in this final year project is my own original work except where citations have been done and has not been presented to any higher institutions of learning for any academic award.

NAME: MWASE BRIAN

Signature.....

Date.....

## APPROVAL

This is to certify that this final year project report is under the supervision of:

SUPERVISOR: MR. MUYINGO EMMANUEL

Signature...

A handwritten signature in blue ink, appearing to read 'E. Muyingo', is shown within a rectangular frame.

Date.....

## ACKNOWLEDGEMENT

I am deeply grateful to the Almighty God for blessing me with life and guiding me to this academic milestone, providing all the necessary resources and strength along the way. My heartfelt thanks go to my cherished family and friends for their unwavering support, encouragement, prayers, and invaluable advice throughout the course of this project. I also extend my sincere appreciation to my supervisors, whose constant guidance, insightful advice, and dedicated efforts in overseeing the progress of this work were instrumental in its completion. A special mention goes to the water resources engineering class of 2025 for their camaraderie, guidance, and assistance in helping me achieve this accomplishment.

## DEDICATION

I dedicate this report to my family members who endured all the hard work put upon them. Am very grateful especially my beloved Aunt **NAMUKASA ESTHER**, my brothers and sisters, may the almighty God bless you all and repay you in abundance and excel in your various activities.

## List of Abbreviations

**ANN:** Artificial Neural Network

**COD:** Chemical Oxygen Demand

**BOD** – Biological Oxygen Demand

**ETP:** Effluent Treatment Plant

**MSE:** Mean Squared Error

**RMSE:** Root Mean Squared Error

**R<sup>2</sup>:** Coefficient of Determination (R-squared)

**ReLU:** Rectified Linear Unit

**KSL** – Kakira Sugar Limited

**NEMA** – National Environment Management Authority

**UASB** – Upflow Anaerobic Sludge Blanket

**DAP** – Di-Ammonium Phosphate

**MATLAB**: Matrix Laboratory (software used for numerical computing)

**pH**: Potential of Hydrogen (measure of acidity/alkalinity)

**ML**: Machine Learning

**LM**: Levenberg-Marquardt (algorithm used in training ANNs)

**MLR**: Multiple Linear Regression

**SVMs**: Support Vector Machines

**MCS**: Monte Carlo Simulation

## Table of contents

|                                                                                                         |    |
|---------------------------------------------------------------------------------------------------------|----|
| <b>ABSTRACT</b> .....                                                                                   | 2  |
| <b>DECLARATION</b> .....                                                                                | 3  |
| <b>APPROVAL</b> .....                                                                                   | 4  |
| <b>ACKNOWLEDGEMENT</b> .....                                                                            | 5  |
| <b>DEDICATION</b> .....                                                                                 | 5  |
| <b>List of Abbreviations</b> .....                                                                      | 5  |
| List of figures.....                                                                                    | 10 |
| List of tables.....                                                                                     | 10 |
| <b>CHAPTER ONE: INTRODUCTION</b> .....                                                                  | 11 |
| 1.1 BACKGROUND .....                                                                                    | 11 |
| 1.2 PROBLEM STATEMENT .....                                                                             | 13 |
| 1.3 JUSTIFICATION OF THE STUDY .....                                                                    | 14 |
| 1.4 OBJECTIVES OF THE STUDY .....                                                                       | 15 |
| 1.5 SIGNIFICANCE.....                                                                                   | 16 |
| 1.6 RESEARCH QUESTIONS .....                                                                            | 17 |
| 1.7 SCOPE OF THE STUDY .....                                                                            | 17 |
| <b>CHAPTER TWO: LITERATURE REVIEW</b> .....                                                             | 19 |
| <b>2.1 CHARACTERIZATION OF INFLUENT WASTEWATER SAMPLES</b> .....                                        | 19 |
| 2.1.1 IDENTIFICATION AND SELECTION OF SAMPLING POINTS WITHIN THE<br>EFFLUENT TREATMENT PLANT (ETP)..... | 19 |
| 2.1.2 COORDINATION WITH LABORATORY TECHNICIANS FOR SAMPLE<br>ANALYSIS.....                              | 20 |
| 2.1.3 PREPROCESSING AND CLEANING OF THE DATASET .....                                                   | 20 |
| <b>2.2 DEVELOPMENT AND TRAINING OF AN ANN MODEL TO PREDICT COD<br/>LEVELS</b> .....                     | 22 |
| 2.2.1 DEFINITION AND TRAINING OF THE ANN MODEL ARCHITECTURE<br>USING MATLAB .....                       | 22 |

|                                                                                                      |           |
|------------------------------------------------------------------------------------------------------|-----------|
| 2.2.2 GENERATING COD PREDICTIONS USING THE TESTING DATASET AND CALCULATING PERFORMANCE METRICS ..... | 22        |
| 2.2.3 CONDUCTING SENSITIVITY ANALYSIS .....                                                          | 23        |
| <b>CHAPTER THREE: METHODOLOGY .....</b>                                                              | <b>25</b> |
| STUDY AREA .....                                                                                     | 25        |
| 3.1 CHARACTERIZATION OF INFLUENT WASTEWATER SAMPLES BY ANALYZING COD-IMPACTING PARAMETERS.....       | 26        |
| 3.1.1 IDENTIFICATION AND SELECTION OF SAMPLING POINTS WITHIN THE EFFLUENT TREATMENT PLANT (ETP)..... | 26        |
| 3.1.2 COORDINATION WITH LABORATORY TECHNICIANS FOR SAMPLE ANALYSIS.....                              | 27        |
| 3.1.3 PREPROCESSING AND CLEANING OF DATASET .....                                                    | 28        |
| 3.2 DEVELOPMENT AND TRAINING OF AN ANN MODEL TO PREDICT COD LEVELS.....                              | 29        |
| 3.2.1 DEFINITION AND TRAINING OF ANN MODEL ARCHITECTURE USING MATLAB.....                            | 29        |
| 3.2.2 GENERATING COD PREDICTIONS USING THE TESTING DATASET .....                                     | 30        |
| 3.2.3 CALCULATING PERFORMANCE METRICS .....                                                          | 30        |
| 3.2.4 OPTIMIZATION AND FINE-TUNING OF MODEL PARAMETERS.....                                          | 31        |
| <b>3.3 CONDUCTING SENSITIVITY ANALYSIS.....</b>                                                      | <b>32</b> |
| 3.3.1 OPTIMIZING THE ANN MODEL BASED ON SENSITIVITY ANALYSIS RESULTS .....                           | 32        |
| <b>CHAPTER FOUR: RESULTS AND DISCUSSION .....</b>                                                    | <b>33</b> |
| 4.1 CHARACTERIZATION OF INFLUENT WASTEWATER SAMPLES BY ANALYZING COD-IMPACTING PARAMETERS.....       | 33        |
| 4.2 DEVELOPMENT AND TRAINING OF AN ANN MODEL TO PREDICT COD LEVELS.....                              | 35        |
| 4.2.1 DEFINITION AND IMPLEMENTATION OF ANN MODEL ARCHITECTURE USING MATLAB .....                     | 35        |

|                                                                  |           |
|------------------------------------------------------------------|-----------|
| 4.2.2 TRAINING THE ANN MODEL.....                                | 35        |
| 4.2.3 GENERATING COD PREDICTIONS USING THE TESTING DATASET ..... | 36        |
| 4.2.4 CALCULATING PERFORMANCE METRICS .....                      | 36        |
| OPTIMIZATION AND FINE-TUNING OF MODEL PARAMETERS.....            | 38        |
| 4.3 CONDUCTING SENSITIVITY ANALYSIS .....                        | 40        |
| <b>CHAPTER FIVE: CONCLUSION AND RECOMMENDATION .....</b>         | <b>41</b> |
| <b>5.1 CONCLUSION .....</b>                                      | <b>41</b> |
| <b>5.2 RECOMMENDATION .....</b>                                  | <b>41</b> |

## List of figures

|                                                                                                          |           |
|----------------------------------------------------------------------------------------------------------|-----------|
| <i>Figure 1 conceptual framework showing all the three specific objectives .....</i>                     | <i>18</i> |
| <b>Figure 3 shows Kakira sampling points blue colour inlet point and red indicates outlet point.....</b> | <b>26</b> |
| <b>Figure 4 laboratory technician and I carrying out the daily necessary COD-related tests .....</b>     | <b>27</b> |
| <b>Figure 5 optimal ANN structure for COD prediction model.....</b>                                      | <b>35</b> |
| <b>Figure 6 shows performance of ANN model.....</b>                                                      | <b>38</b> |
| <b>Figure 7 shows parity plot of COD prediction model.....</b>                                           | <b>36</b> |
| <b>Figure 8 shows Regression plot between predicted and measured data for the ANN model .....</b>        | <b>37</b> |
| <b>Figure 9 showing Sensitivity Analysis results (Importance of input parameters) .....</b>              | <b>40</b> |

## List of tables

|                                                                                    |                                     |
|------------------------------------------------------------------------------------|-------------------------------------|
| <b>Table 1 showing daily parameters dataset .....</b>                              | <b>Error! Bookmark not defined.</b> |
| <b>Table 2 Basic statistics and data cleaning of Kakira ETP water quality.....</b> | <b>34</b>                           |
| <b>Table 3 sensitivity-guided optimization results .....</b>                       | <b>Error! Bookmark not defined.</b> |

- **Real-time Implementation and Sensor Integration:** Explore the integration of the developed ANN model with online sensors for key input parameters (FlowRate, pH, temperature) to enable real-time COD<sub>eff</sub> prediction and automated alerts at Kakira Sugar ETP.
- **Expansion of Dataset:** Continuously update the dataset with new operational data to retrain and refine the model, improving its accuracy and adaptability to changing plant conditions or influent characteristics over time.
- **Exploration of Hybrid Models:** Investigate hybrid modeling approaches (e.g., ANN combined with fuzzy logic or genetic algorithms) or other machine learning techniques (e.g., Support Vector Machines, Random Forests) to potentially further enhance predictive accuracy or interpretability.
- **Multi-Parameter Prediction:** Extend the modeling approach to predict other critical effluent parameters, such as Biological Oxygen Demand (BOD) or Total Suspended Solids (TSS), for a more comprehensive ETP performance monitoring system.
- **Transferability Assessment:** Conduct studies to assess the transferability of the modeling methodology to other similar industrial ETPs, adapting the model with site-specific data.

## References

- Aani, S. Al, Bonny, T., Hasan, S.W. & Hilal, N., 2019, 'Can machine language and artificial intelligence revolutionize process automation for water treatment and desalination?', *Desalination*, 458, 84–96.
- Al, R., Behera, C.R., Zubov, A., Gernaey, K. V. & Sin, G., 2019, 'Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants – An application to the BSM2 model', *Computers and Chemical Engineering*, 127, 233–246.
- Almomani, F., 2020, 'Prediction the performance of multistage moving bed biological process using artificial neural network (ANN)', *Science of The Total Environment*, 744, 140854.
- Alsulaili, A. & Refaie, A., 2021, 'Artificial neural network modeling approach for the prediction of five-day biological oxygen demand and wastewater treatment plant performance', *Water Supply*, 21(5), 1861–1877.
- Ay, M. & Kisi, O., 2014, 'Modelling of chemical oxygen demand by using ANNs, ANFIS and k-means clustering techniques', *Journal of Hydrology*, 511, 279–289.

- Bachir, S., Samir, B., Hicham, C. & Azzedine, H., 2021, 'Prediction of Effluent Chemical Oxygen Demand and Suspended Solids From a Domestic Wastewater Treatment Plant Using SVM and ANN', *Soft Computing Techniques in Solid Waste and Wastewater Management*, pp. 275–288, Elsevier.
- Bagheri, M., Akbari, A. & Mirbagheri, S.A., 2019, 'Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: A critical review', *Process Safety and Environmental Protection*, 123, 229–252.
- Bagheri, M., Farshforoush, N., Bagheri, K. & Shemirani, A.I., 2023, 'Applications of artificial intelligence technologies in water environments: From basic techniques to novel tiny machine learning systems', *Process Safety and Environmental Protection*, 180, 10–22.
- Bahramian, M., Dereli, R.K., Zhao, W., Giberti, M. & Casey, E., 2023, 'Data to intelligence: The role of data-driven models in wastewater treatment', *Expert Systems with Applications*, 217, 119453.
- Balogun, S. & Ogwueleka, T.C., 2023, 'Performance prediction for wastewater treatment plant effluent cod using artificial neural network', *International Journal of Environmental Science and Technology*, 20(11), 12659–12668.
- Behera, S.K., Karthika, S., Mahanty, B., Meher, S.K., Zafar, M., Baskaran, D., Rajamanickam, R., Das, R., Pakshirajan, K., Bilyaminu, A.M. & Rene, E.R., 2024, 'Application of artificial intelligence tools in wastewater and waste gas treatment systems: Recent advances and prospects', *Journal of Environmental Management*, 370, 122386.
- Bekkari, N. & Zeddouri, A., 2019, 'Using artificial neural network for predicting and controlling the effluent chemical oxygen demand in wastewater treatment plant', *Management of Environmental Quality: An International Journal*, 30(3), 593–608.
- Cairone, S., Hasan, S.W., Choo, K.-H., Li, C.-W., Zarra, T., Belgiorno, V. & Naddeo, V., 2024, 'Integrating artificial intelligence modeling and membrane technologies for advanced wastewater treatment: Research progress and future perspectives', *Science of The Total Environment*, 944, 173999.
- Corominas, Ll., Garrido-Baserba, M., Villez, K., Olsson, G., Cortés, U. & Poch, M., 2018, 'Transforming data into knowledge for improved wastewater treatment operation: A critical review of techniques', *Environmental Modelling & Software*, 106, 89–103.
- El-Din, A.G. & Smith, D.W., 2002, 'A neural network model to predict the wastewater inflow incorporating rainfall events', *Water Research*, 36(5), 1115–1126.