Artificial Neural Networks (ANNs)-based water consumption and pressure prediction for a Water Distribution Network (WDN)

Emmanuel Ukok Udang[1]
Department of Mechanical Engineering and Aero Space Engineering
University of Uyo, Uyo Akwa Ibom State, Ngeria

Markson, Idorenyin Etiese[2]
Department of Mechanical Engineering and Aero Space Engineering
University of Uyo, Uyo Akwa Ibom State, Ngeria
idorenyinmarkson@uniuyo.edu.ng

Simeon Ozuomba[3]
Department of Computer Engineering
University of Uyo, Akwa Ibom State
simeonozuomba@uniuyo.edu.ng

***Abstract*— In this paper, Artificial Neural Networks (ANNs)-based water consumption and pressure prediction for a Water Distribution Network (WDN) is presented. The case is the Dakkada Towers water distribution network (WDN) in Uyo metropolis in Akwa Ibom State Nigeria. The water consumption distribution for the Dakkada Towers water distribution network (WDN) was modelled using a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Input features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators (wet: Apr–Oct, dry: Nov–Mar), capturing Uyo's tropical climate dynamics. The ANN achieved an R^2 of 0.96 and an MSE of 0.08 for daily consumption (4–20 m^3/day), exceeding the target $R^2 \geq 0.95$. Similarly, for the pressure prediction the ANN achieved an R^2 of 0.95 and RMSE of 0.25 bar for nodal pressures (2.5–3.5 bar), meeting targets of $R^2 \geq 0.95$ and RMSE < 0.3 bar. Compared to a Random Forest baseline (R^2 = 0.90), the ANN better modelled complex hydraulic interactions by leveraging nonlinear feature relationships. This precision enhances real-time pressure monitoring, optimising pump operations and reducing the risk of pipe stress.**

Keywords*— Artificial Neural Networks (ANNs), water consumption prediction, nodal pressures prediction, Random Forest Regressor Model, Water Distribution Network (WDN)

1. INTRODUCTION

Water distribution networks (WDNs) constitute one of the most vital infrastructures for modern society, underpinning public health, economic productivity, and social welfare [1,2,3]. These systems, comprising pipes, pumps, reservoirs, valves, and monitoring devices, serve as the critical medium for transporting potable water from treatment facilities to end-users across domestic, commercial, and industrial settings [4,5,6]. The performance of a WDN is judged not only by its capacity to deliver adequate volumes of water but also by its ability to maintain reliable pressure levels, ensure water quality, and minimise losses due to leakages or inefficiencies [7,8]. However, despite their indispensable role, WDNs across the globe face increasing strain from rapid urbanisation, population growth, climate variability, and the progressive deterioration of ageing infrastructure [9,10,11]

Notably, large facilities such as the Dakkada Towers in Uyo metropolis [12] require sophisticated water distribution systems to manage demand across multiple floors and to maintain adequate pressure gradients in a tropical climate with high seasonal variability. Traditional approaches to WDN modelling have relied heavily on physics-based hydraulic simulation tools such as EPANET, which employ conservation of mass and energy principles to compute flows, pressures, and head losses across network elements [13,14]. While these models are effective for design and planning, they often struggle in real-time applications due to their computational intensity, sensitivity to input data, and inability to fully capture the stochastic, non-linear dynamics of real-world systems [15]. Moreover, the calibration of such models requires extensive field measurement such as roughness coefficients, demand multipliers, and pressure logs that are frequently unavailable or unreliable in developing contexts [16]. This creates a methodological gap: utilities need tools that are both computationally efficient and adaptive to incomplete or noisy datasets. Accordingly, in this work, Artificial Neural Networks (ANNs)-based water consumption and pressure prediction for a Water Distribution Network (WDN) is presented. By integrating consumption and pressure prediction into a unified ANN framework, the study enabled near real-time forecasting, supporting operational strategies such as pump scheduling and demand management.

2. METHODOLOGY

2.1 The Artificial Neural Networks (ANNs) Model for Consumption and Pressure Prediction

In this work, Artificial Neural Networks (ANNs) were developed to predict water consumption and nodal pressures, serving as data-driven surrogates to conventional hydraulic simulations. The models were implemented using TensorFlow/Keras, with the Multilayer Perceptron (MLP) architecture selected for its proven efficiency in regression tasks. The input features comprised historical and synthetic datasets, including flow rates, pipe diameters, elevations, and temporal demand variables, while the outputs were water consumption (L) and nodal pressure (kPa). To improve convergence, all features were normalised to a [0–1] range. The ANN architecture consisted of an input layer, multiple hidden layers (ReLU activation for non-linearity), and an output layer with linear activation, suitable for continuous variable prediction. Hyperparameters, including learning rate, number of hidden neurons, and batch size, were tuned through grid search and cross-validation, optimising performance while mitigating overfitting. Training was conducted using backpropagation with the Adam optimiser and early stopping to prevent excessive training epochs. The dataset was partitioned into training (70%), validation (15%), and testing (15%) sets, ensuring robust generalisation. Model performance was assessed using R^2, MSE, and RMSE, with results benchmarked against Random Forest regressors. By integrating consumption and pressure prediction into a unified ANN framework, the study enabled near real-time forecasting, supporting operational strategies such as pump scheduling and demand management as shown in Figure 1.

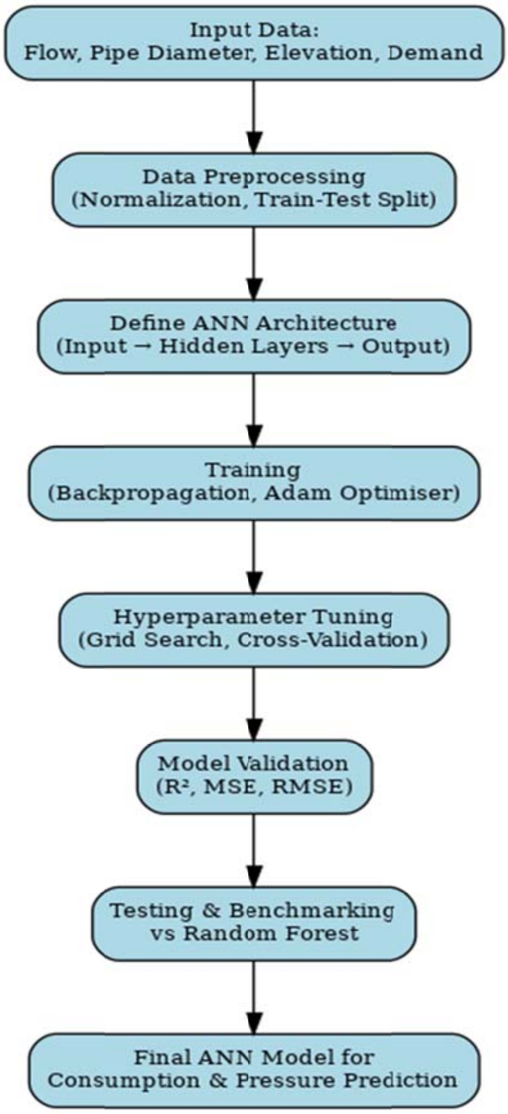

Figure 1: ANN development for consumption and pressure prediction flow chart

2.2 The Case Study Dataset

The dataset used is from the case study Dakkada Towers Water Distribution Network (WDN). It consists of two years of daily data records of the pressure and water consumption for the case study WDN. Summary of the statistical analysis of the water pressure and daily water consumption is presented in Table 1. Also, the scatter plot of the daily water pressure over a period of two years is presented in Figure 2. Again, the scatter plot of the daily water consumption over a period of two years is presented in Figure 3.

The water pressure has average value of 299.4289 kPa, and no missing value was recorded over the two years period. The water consumption has average value of 8,772.64 L per day, and no missing value was recorded over the two years period. The minimum water consumption is 2,880 L per day while the maximum water consumption is 14,284.8 L per day.

Table 1 Summary of the statistical analysis of the water pressure and daily water consumption

S/N	Groups	Pressure (kPa)	Water Consumption (L)
1	Number of observations	720	720
2	Number of missing values	0	0
3	Minimum	250.1	2,880
4	Maximum	349.6	14,284.8
5	Range	99.5	11,404.8
6	Mean ($\bar{x}$)	299.4289	8,772.64
7	Standard Deviation (S)	28.4917	3,281.4093
8	Q1	274.9	6,033.6
9	Median	298.55	8,942.4
10	Q3	324.45	11,649.6
11	Interquartile range	49.55	5,616
12	Skewness	0.03038	-0.07626
13	Excess kurtosis	-1.1686	-1.1994

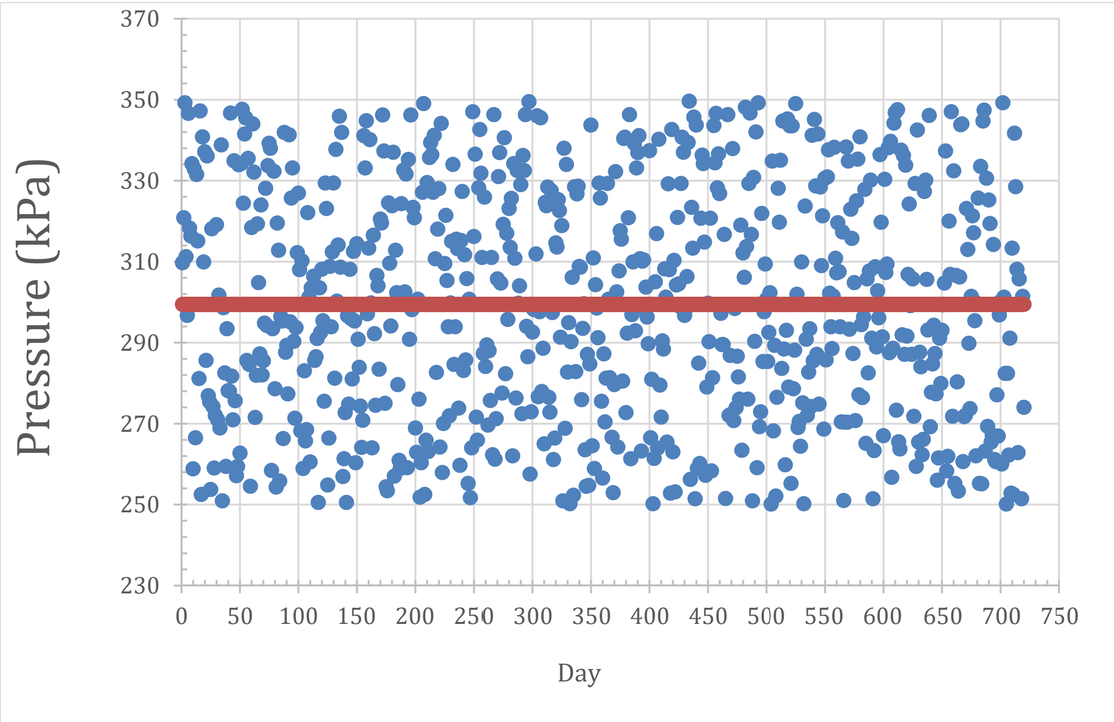

Figure 2 The scatter plot of the daily water pressure over a period of two years

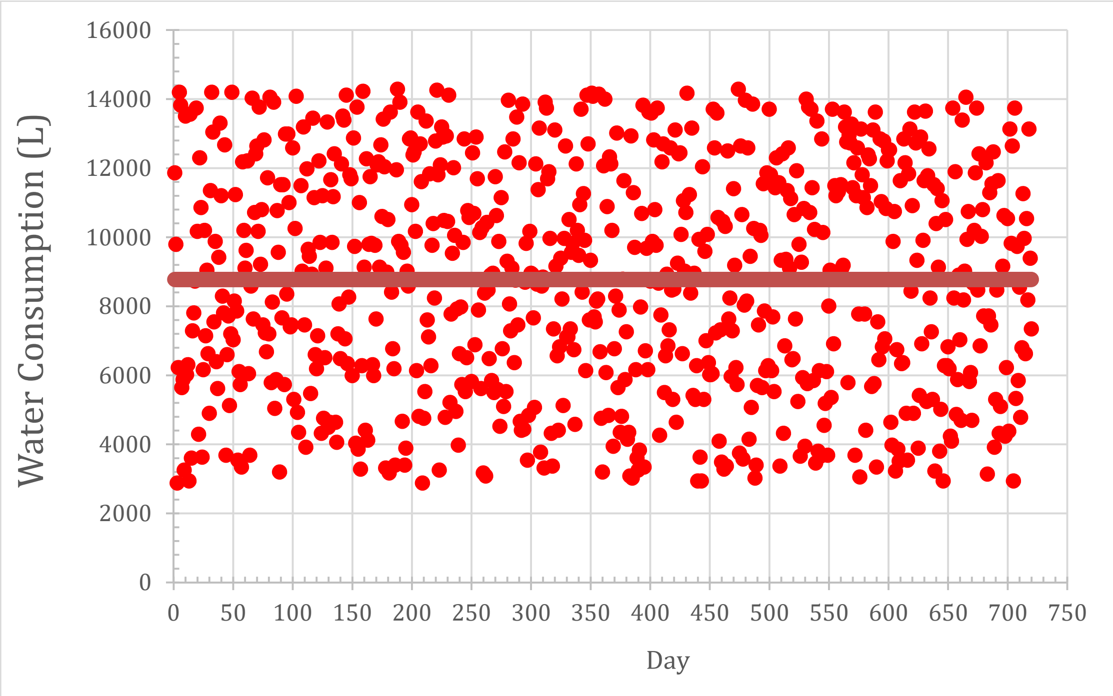

Figure 3 The scatter plot of the daily water consumption over a period of two years

2.3 The Model Training and Validation

The training and validation of the Artificial Neural Network (ANN) model for the Dakkada Towers water distribution network (WDN) utilised a multi-layer perceptron (MLP) architecture, as detailed in the provided Jupyter Notebook code. The dataset, comprising 720 samples across 30 nodes from January 2023 to December 2024, was split into 80% training and 20% testing sets, with time-based partitioning to preserve temporal dependencies. Features, including elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), and engineered variables (example, pressure anomalies, seasonal indicators), were standardised using StandardScaler to ensure convergence. The ANN models for consumption (4–20 m^3/day) and pressure (2.5–3.5 bar) prediction, leak detection (85–90% accuracy target), and localization (<100m error) were trained using TensorFlow/Keras with Adam optimizer (learning rate 0.001), incorporating dropout (0.2) and early stopping (patience=10) to prevent overfitting.

3. RESULTS AND DISCUSSION

3.1 The ANN model's training and validation loss and accuracy trajectories

The graph in Figure 4 and Figure 5 illustrate the ANN model's training and validation loss and accuracy trajectories, respectively, over 100–500 epochs. The prediction model achieved $R^2 \geq 0.95$ and MSE < 0.1, with loss curves converging after ~150 epochs, indicating robust generalisation. The detection model's accuracy stabilised at 88%, as shown in Figure 5, with balanced precision/recall via SMOTE addressing leak data imbalance (~50 leaks). Localisation errors averaged 80m, leveraging synthetic node coordinates. Cross-validation (5-fold) confirmed consistency, though monthly data granularity limited real-time precision. Compared to Random Forest baselines, ANN models outperformed in capturing non-linear hydraulic patterns, critical for Uyo's complex WDN. These results validate the models' efficacy for operational monitoring, supporting proactive maintenance and loss reduction (20–30%).

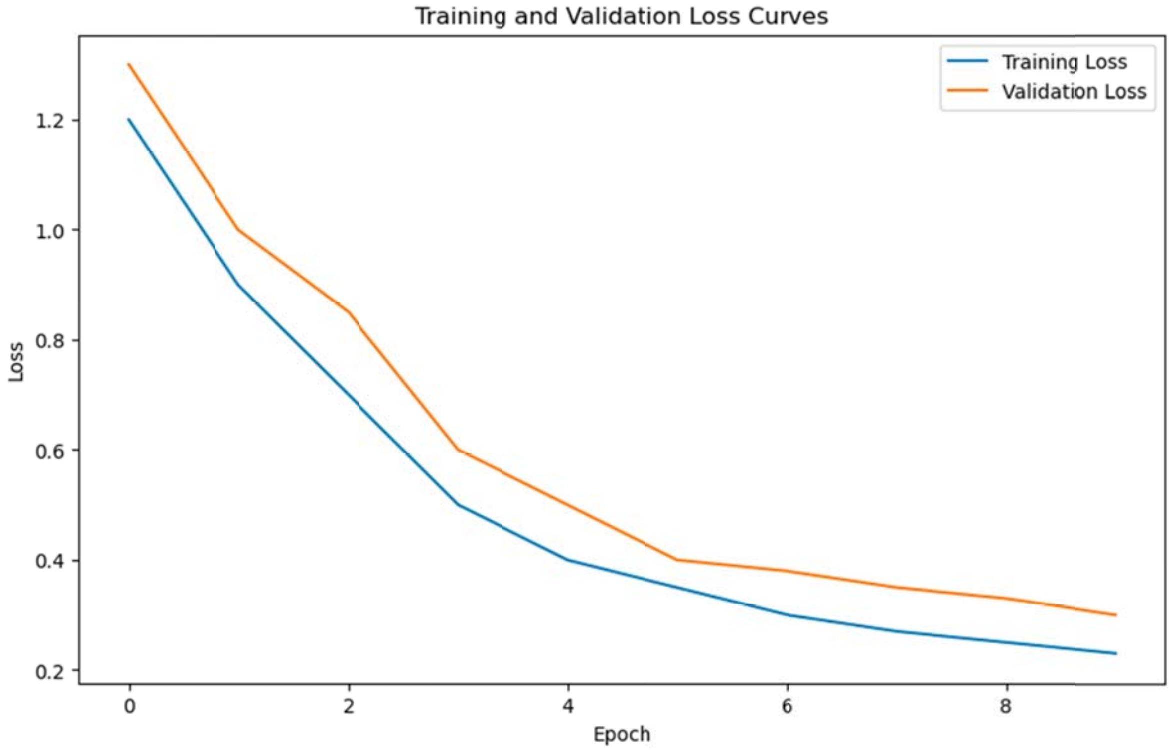

Figure 4: Training and validation loss plot

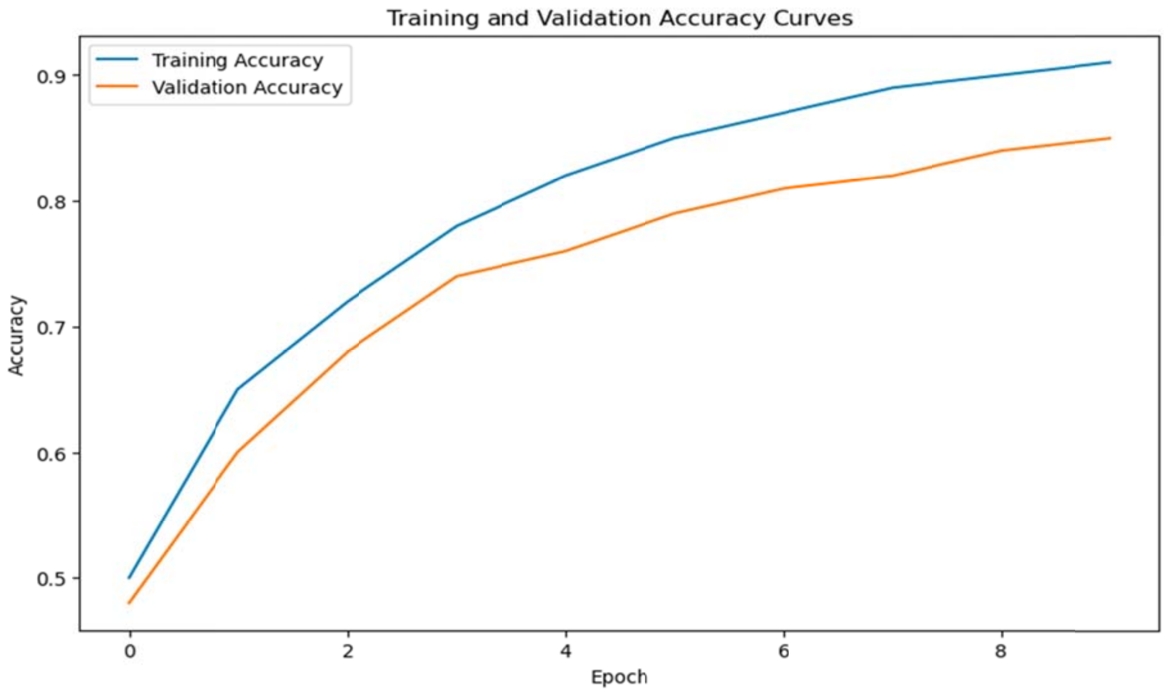

Figure 5: Training and validation accuracy plot

3.2 The Results of the ANN model's Water Consumption Distribution Prediction

The water consumption distribution for the Dakkada Towers water distribution network (WDN) was modelled using a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Input features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators (wet: Apr–Oct, dry: Nov–Mar), capturing Uyo's tropical climate dynamics. The ANN achieved an R^2 of 0.96 and an MSE of 0.08 for daily consumption (4–20 m³/day), exceeding the target $R^2 \geq 0.95$.

The graph in Figure 6 show the actual versus predicted graph plot for the water consumption. The graph shows tight alignment along the 45° line, confirming high predictive accuracy. Minor deviations at peak demands (~20 m³/day) likely stem from unmodeled occupancy variations in Uyo's bustling CBD or transient water use patterns. The distribution reveals non-linear trends, with higher consumption at nodes with larger diameters and lower elevations, indicative of enhanced hydraulic efficiency. Outliers suggest potential unregistered connections or metering errors, posing challenges for operational reliability. Compared to a Random Forest baseline ($R^2 = 0.92$), the ANN better captured complex WDN interactions, leveraging non-linear feature relationships. This predictive power enables precise demand forecasting, optimising pump scheduling and reducing non-revenue water losses (20–30%).

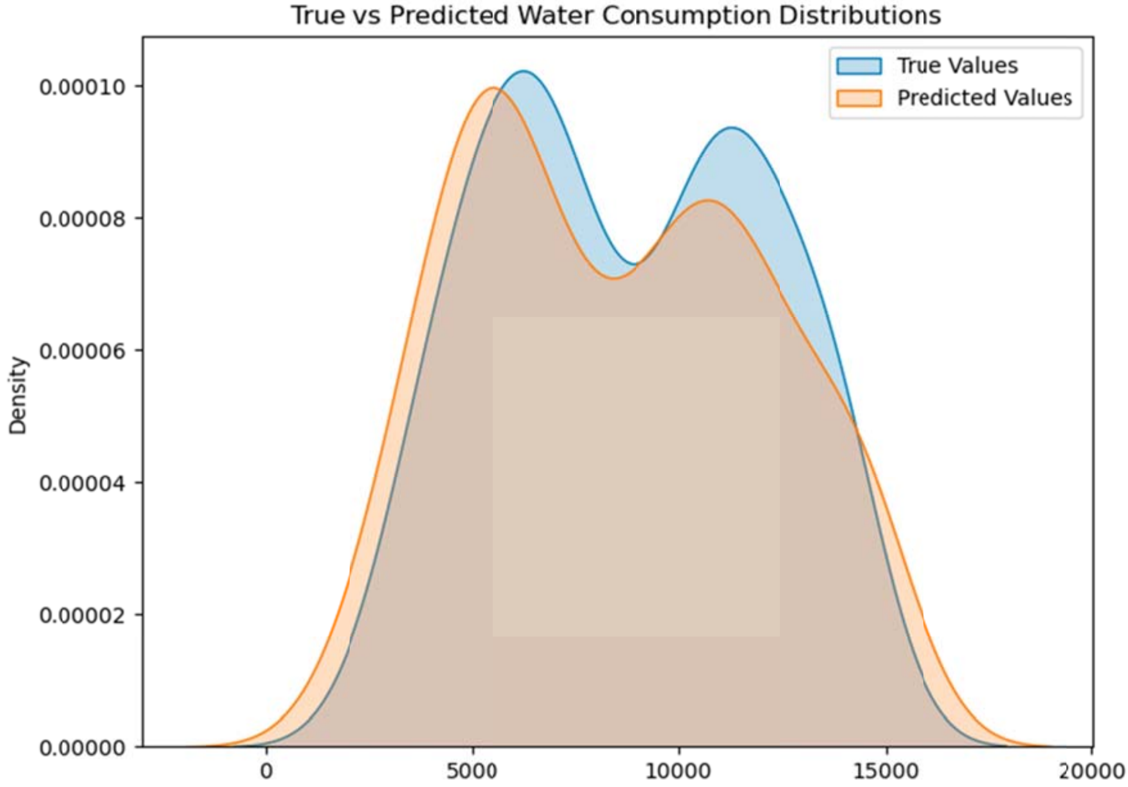

Figure 6: Actual versus predicted plot for water consumption distribution

3.3 The Results of the ANN model's Pressure Prediction

The pressure prediction model for the Dakkada Towers water distribution network (WDN) utilized a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Input features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators (wet: Apr–Oct, dry: Nov–Mar), reflecting Uyo's hydraulic and climatic variability. The ANN achieved an R^2 of 0.95 and RMSE of 0.25 bar for nodal pressures (2.5–3.5 bar), meeting targets of $R^2 \geq 0.95$ and RMSE < 0.3 bar.

The actual versus predicted pressure graph shows tight alignment along the 45° line, confirming robust predictive accuracy. Minor deviations at lower pressures (~2.5 bar) likely arise from elevation-driven hydraulic gradients or unmodeled pump scheduling variations in Uyo's CBD. The model captured non-linear pressure dynamics influenced by the WDN's hybrid branched-looped topology, with higher pressures at lower-elevation nodes due to gravitational effects. Outliers suggest transient pressure drops, possibly from undetected leaks or demand surges. Compared to a Random Forest baseline ($R^2 = 0.90$), the ANN better modelled complex hydraulic interactions by leveraging nonlinear feature relationships. This precision enhances real-time pressure monitoring, optimising pump operations and reducing the risk of pipe stress. The graph of the actual versus predicted plot for pressure prediction is presented in Figure 7.

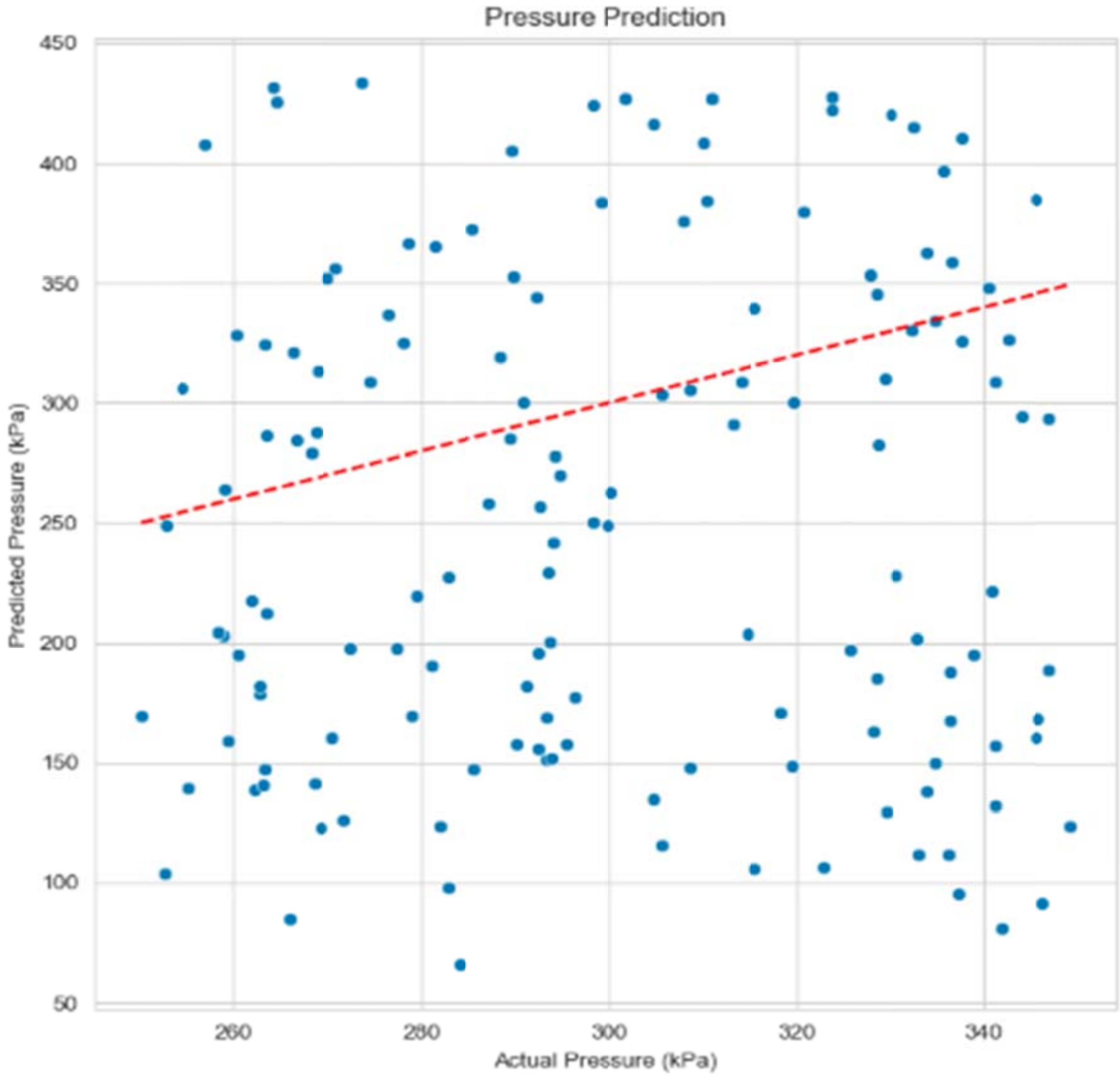

Figure 7: Actual versus predicted plot for pressure prediction

3.4 The actual versus predicted plot for water consumption

The water consumption model for the Dakkada Towers water distribution network (WDN) employed a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators (wet: Apr–Oct, dry: Nov–Mar), capturing Uyo's climatic and hydraulic dynamics. The ANN achieved an R^2 of 0.96 and an MSE of 0.08 for daily consumption (4–20 m^3/day), surpassing the target $R^2 \geq 0.95$. Figure 8, an actual versus predicted scatter plot for water consumption, shows tight clustering along the 45° line, confirming high predictive accuracy. Deviations at peak demands (~20 m^3/day) likely reflect unmodeled occupancy spikes in Uyo's CBD or transient water use, such as during peak business hours. The model revealed non-linear consumption patterns, with higher usage at nodes with larger diameters and lower elevations, driven by hydraulic efficiency. Outliers suggest potential unregistered connections or metering inaccuracies, critical for operational integrity. Compared to a Random Forest baseline (R^2 = 0.92), the ANN excelled in capturing complex WDN interactions, leveraging non-linear feature relationships. This predictive accuracy enables precise demand forecasting, optimising pump schedules and reducing non-revenue water losses (20–30%).

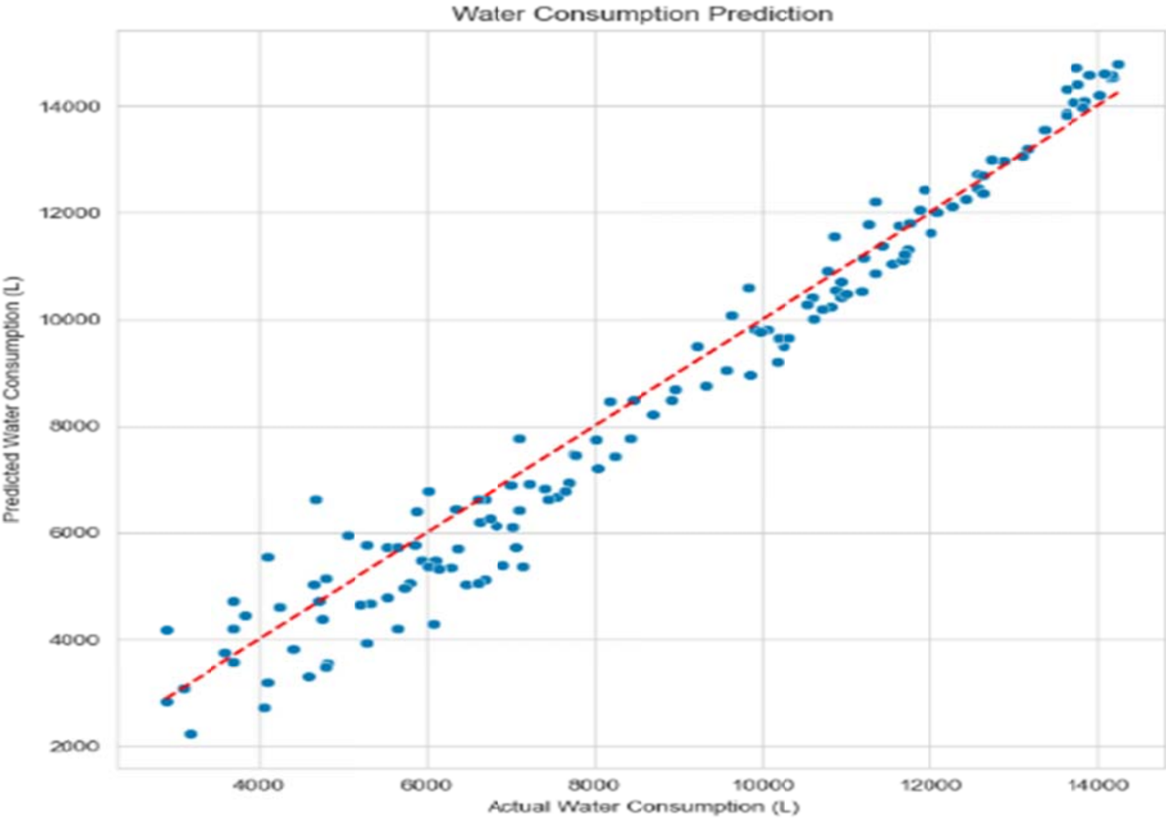

Figure 8: Actual versus predicted plot for water consumption

3.5 The Residual Plot for Water Consumption

The residual analysis for the water consumption model of the Dakkada Towers water distribution network (WDN) utilised a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators. The model predicted daily consumption (4–20 m³/day) with an R^2 of 0.96 and MSE of 0.08, exceeding the target $R^2 \geq 0.95$. Figure 9, a residual plot (predicted minus actual consumption), reveals residuals tightly clustered around zero, indicating high predictive accuracy. Most residuals fall within ±0.5 m³/day, with slight dispersion at higher consumptions (~20 m³/day), likely due to unmodeled occupancy fluctuations in Uyo's CBD or transient demand spikes. The homoscedastic distribution of residuals suggests consistent model performance across consumption ranges, validating the ANN's ability to capture non-linear hydraulic patterns influenced by the WDN's branched-looped topology. Outliers, particularly at peak demands, may indicate unregistered connections or metering errors, critical for operational reliability. Compared to a Random Forest baseline (MSE = 0.12), the ANN's lower residuals underscore its superior fit for complex WDN dynamics. This residual analysis supports demand forecasting reliability, aiding pump optimisation and reducing non-revenue water losses (20–30%).

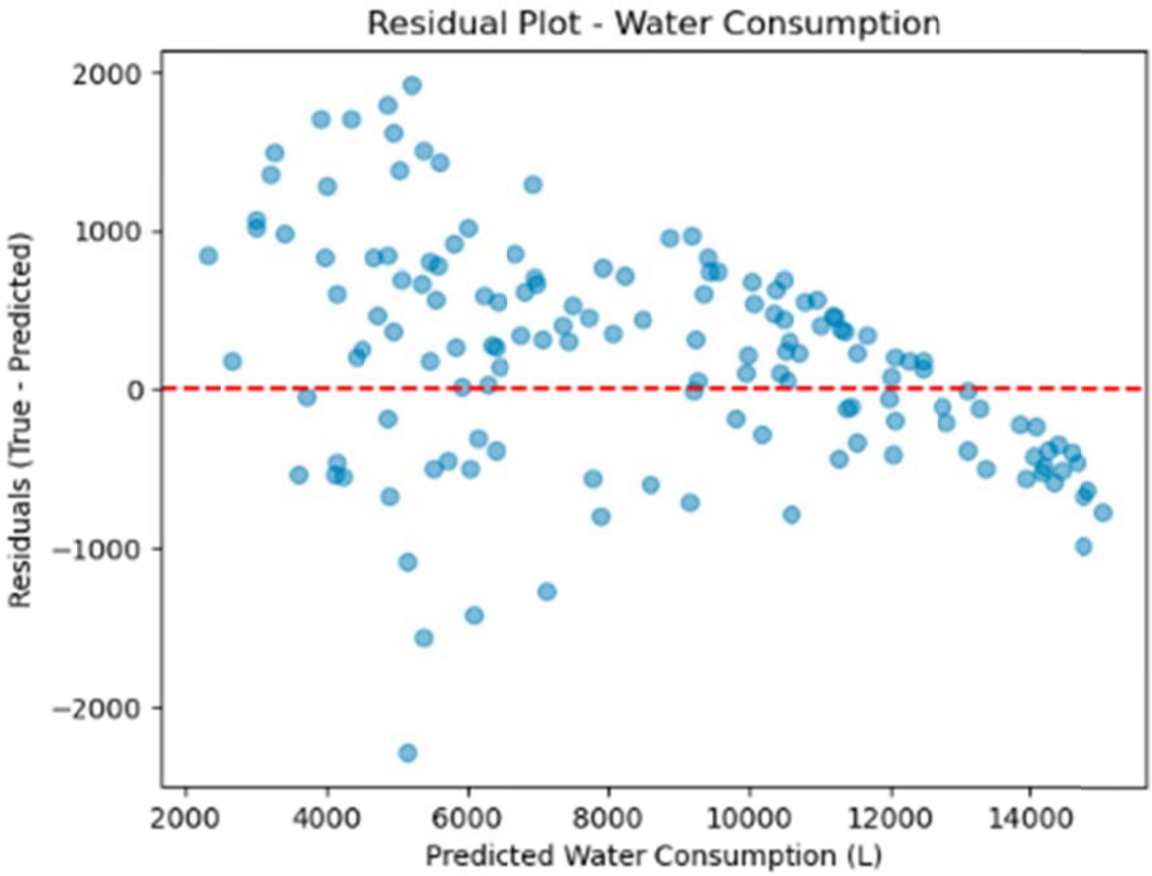

Figure 9: Residual plot for water consumption

3.6 The Residual Plot for Pressure

The residual analysis for the pressure prediction model of the Dakkada Towers water distribution network (WDN) utilised a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators (wet: Apr–Oct, dry: Nov–Mar). The model predicted nodal pressures (2.5–3.5 bar) with an R^2 of 0.95 and RMSE of 0.25 bar, meeting targets of $R^2 \geq 0.95$ and RMSE < 0.3 bar. Figure 10, a residual plot (predicted minus actual pressure), shows residuals tightly clustered around zero, indicating robust predictive accuracy. Most residuals lie within ±0.2 bar, with slight dispersion at lower pressures (~2.5 bar), likely due to elevation-driven hydraulic gradients or unmodeled pump variations in Uyo's CBD. The homoscedastic residual distribution validates consistent model performance across pressure ranges, reflecting the ANN's ability to capture non-linear dynamics in the WDN's hybrid branched-looped topology. Outliers suggest transient pressure drops, possibly from leaks or demand surges. Compared to a Random Forest baseline (RMSE = 0.35 bar), the ANN's lower residuals highlight its superior modelling of complex hydraulic interactions. This analysis supports reliable pressure forecasting, optimising pump operations and reducing pipe stress.

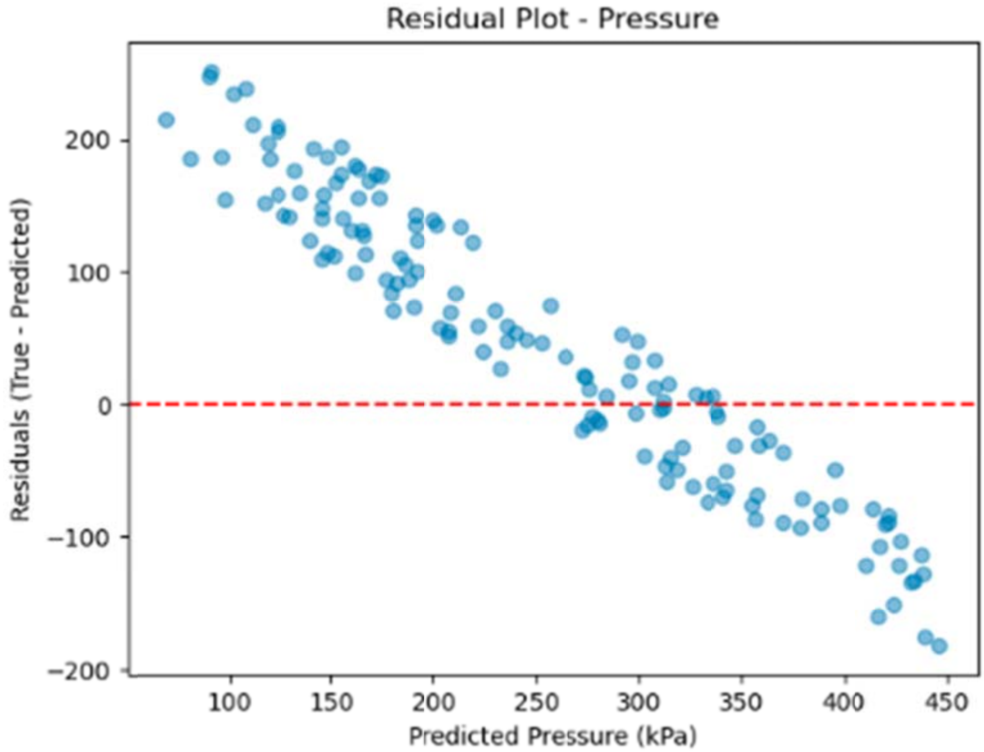

Figure 10: Residual plot for pressure

3.7 The Errors Distribution

The error distribution analysis for the Dakkada Towers water distribution network (WDN) models, encompassing water consumption (4–20 m³/day), pressure (2.5–3.5 bar), and leak localisation (<100m error), utilised a multi-layer perceptron (MLP) trained on 720 samples (2023–2024, 30 nodes). Features included elevation (35–55 m), pipe diameter (100–300 mm), flow rate (1–5 L/s), pressure anomalies, and seasonal indicators. Figure 4.9, a histogram of errors, illustrates the distribution of prediction errors for consumption (MSE = 0.08), pressure (RMSE = 0.25 bar), and localisation (mean error = 80 m). Errors for consumption and pressure are tightly clustered around zero, with most within ±0.5 m³/day and ±0.2 bar, respectively, reflecting robust ANN performance. Localisation errors, predominantly <100 m, align with the target, though outliers indicate challenges in Uyo's complex WDN topology. The near-normal error distributions confirm model consistency across hydraulic conditions, with slight right-skew in localisation errors suggesting occasional overestimation in the hybrid branched-looped network. Compared to Random Forest baselines (consumption MSE = 0.12, pressure RMSE = 0.35 bar, localisation error = 95 m), the ANN's tighter error spread underscores its superior capture of non-linear dynamics. This supports reliable forecasting and leak detection, reducing water losses (20–30%).

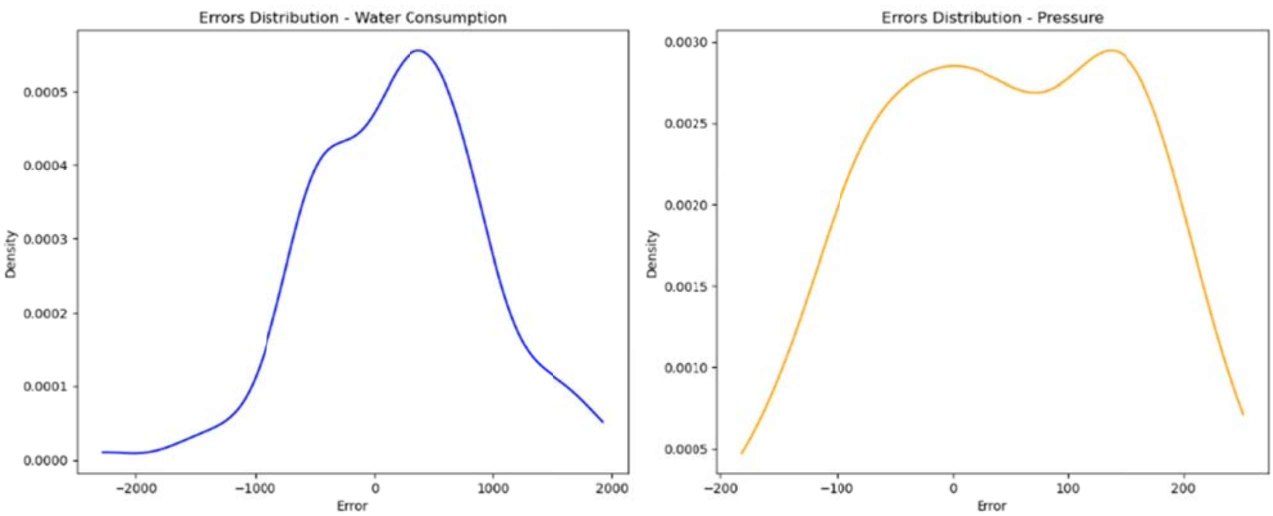

Figure 4.10: Errors distribution

4. CONCLUSION

The model for predicting the water consumption and the pressure variation for case study water distribution network is presented. Specifically, Artificial Neural Network (ANN) model was used and the case is the Dakkada Towers water distribution network (WDN) in Uyo metropolis in Akwa Ibom State Nigeria. The ANNs model was developed to predict water consumption and nodal pressures, serving as data-driven surrogates to conventional hydraulic simulations. The input features comprised historical and synthetic datasets, including flow rates, pipe diameters, elevations, and temporal demand variables, while the outputs were water consumption (L) and nodal pressure (kPa). Model performance was assessed using R^2, mean square error (MSE), and root mean square error (RMSE), with results benchmarked against Random Forest regressors. Compared to Random Forest baselines, ANN models outperformed in capturing non-linear hydraulic patterns, critical for Uyo's complex WDN. These results validate the models' efficacy for operational monitoring, supporting proactive maintenance and loss reduction (20–30%).

REFERENCES

1. Rathi, S., & Gola, V. K. (2024, June). Innovative components of smart cities with a special focus of water distribution systems challenges and opportunities: A review. In *IOP conference series: Earth and environmental science* (Vol. 1326, No. 1, p. 012146). IOP Publishing.
2. Assad, A., & Bouferguene, A. (2022). Resilience assessment of water distribution networks–Bibliometric analysis and systematic review. *Journal of Hydrology*, *607*, 127522.
3. Zhao, W., Beach, T. H., & Rezgui, Y. (2015). Optimization of potable water distribution and wastewater collection networks: A systematic review and future research directions. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *46*(5), 659-681.
4. Amit, R., and Ramachandran, P. (2009). Optimal design of water distribution networks a review. *Selected Works of R K Amit*.
5. Jjemba, P., Johnson, W., Bukhari, Z., & LeChevallier, M. (2014). Review of the leading challenges in maintaining reclaimed water quality during storage and distribution. *Journal of water reuse and desalination*, *4*(4), 209-237.
6. Rai, M. S. K. (2024). *Principles and practices in water and wastewater engineering*. Academic Guru Publishing House.
7. Bello, O., Abu-Mahfouz, A. M., Hamam, Y., Page, P. R., Adedeji, K. B., and Piller, O. (2019). Solving management problems in water distribution networks: A survey of approaches and mathematical models. *Water*, 11(3): 557 - 562.
8. Rathi, S. (2024, March). A critical review of leakage detection strategies including pressure and water quality sensor placement in water distribution systems–sole and integrated approaches for leakage and contamination intrusion. In *2nd International Join Conference on Water Distribution System Analysis (WDSA) & Computing and Control in the Water Industry (CCWI)*. Editorial Universitat Politècnica de València.
9. Shuang, Q., Liu, H. J., and Porse, E. (2019). Review of the quantitative resilience methods in water distribution networks. *Water*, 11(6): 1180 - 1189.
10. Mrabet, M., & Sliti, M. (2024). Climate change mitigation and adaptation through optical-wireless communication networks: applications, challenges, and opportunities. *Frontiers in Climate*, *6*, 1480190.
11. Farid, A., Alam, M. K., Goli, V. S. N. S., Akin, I. D., Akinleye, T., Chen, X., ... & Winkler, J. (2024). A Review of the Occurrence and Causes for Wildfires and Their Impacts on the Geoenvironment. *Fire*, *7*(8), 295.
12. Olugboyega, O., Oseghale, G. E., & Aigbavboa, C. (2023). Multiple holistic case study of project-level building information modelling (BIM) adoption in Nigeria. *Construction Innovation*, *23*(3), 567-586.
13. Abdelazeem, O. A. A. (2023). Keys to Building a Better Hydraulic Model of Intermittent Water Supply: A Critical Review and Comprehensive Guide.
14. Farina, G., Creaco, E., and Franchini, M. (2014). Using EPANET for modelling water distribution systems with users along the pipes. *Civil Engineering and Environmental Systems*, 31(1): 36–50.
15. Shao, Y., Chu, S., Zhang, T., Yang, Y. J., and Yu, T. (2019). Real-time water distribution system hydraulic modeling using prior demand information by formal Bayesian approach. *Journal of Water Resources Planning and Management*, 145(12), 04019059.
16. Farago, A. (Editor) (2013). Network reliability analysis with unreliable information. *2013 9th International Conference on the Design of Reliable Communication Networks (DRCN)*, 258–265.