Prediction of Water Loss in Hydraulic Distribution System in Minna, Nigeria Using Artificial Neural Network

Yaba, T.[1,2], **Jimoh, O.D.**[1a], **Adesiji, A. R.**[1b]
[1]Civil Engineering Department, Federal University of Technology, Minna, Nigeria
[2] Niger State Water and Sewage Corporation, Minna, Niger State
[1.2]tanimuyaba@gmail.com; [1a]odjimoh@gmail.com; [1b]ade.richard@futminna.edu.ng
Corresponding Author: tanimuyaba@gmail.com

Abstract

Water supply network are prone to leakages resulting to a loss of large volume of water. Hence it is required to implement a leak detection/prediction technique through water simulation and machine learning. The main objective of this study is to model water loss in the distribution network of Shiroro District Metered Area. This is important because leak is a measure of efficiency of water distribution network. The hydraulic machine, EPANET was used for the hydraulic modelling of the networks. Emitters were used to simulate leakages at thirty-seven nodes in water distribution system. Physical measurement was carried out also at thirty-seven nodes in the network using measuring can, hose, GPS, meter, stop clock. Nash-Sutcliffe simulation efficiency (ENS) indicates how well the plot of observed versus simulated value fits the 1:1 line. The value of efficiency of 1 (when E = 1) means there is a perfect match of modelled discharge relative to the observed data. The observed and model data were loaded into NSE model using coefficient of 0.1, 0.15 0.2 and 0.3. The performance of the model has suggested that using the emitter coefficient of 0.2 can model the study area. Having established this, the values of the model could be used to predict leakages in the DMA using Artificial neural Network, ANN. This study was based on Multi-Layer Perception which was trained and tested using DMA flow data. The objective was to develop an ANN-based model using flow data generated in the selected DMA in Minna, Niger State, Nigeria. The input variables are elevation, base demand, demand and pressure of the network. The data was trained tested and validated in neural network. The study has shown 17.15% of loss from the nodes in the network. The sum of square errors 13.4% and 5.1% respectively for training and testing of the variables in the machine learning. R square is 97%. The model developed can be used in any district metered area of a distribution network to estimate or predict loss. The developed model is expected to help set the direction of improvement of the analysis of water distribution system and optimal operation of water supply in the studied DMA and other DMAs.

Keywords: Water Distribution System, EPANET, Emitter Coefficient, Artificial Neural Network

Introduction

Water distribution systems are primary means of safe drinking water supply to the system. Water produced and delivered to the distribution system is intended for the customers or users. However, a significant amount of water is lost in the system before it gets to its intended users as leak which is termed a physical component of Non-Revenue Water. The occurrence of leaks depends on the factors like materials, composition, age, pressure and joining. Due to complexity of the distribution system, it may be difficult for the utility personnel to identify and fix all the leaks. Hence the need for the development of methodology to identify the leaks using model by integrating observation data

Current statistical surveys indicated that NRW in developing countries is around 45 to 50% that is half of the total system input volume. A high level of apparent losses reduces the principal revenue stream to the utility. Zabidi *et al*. (2020) reported that losses in water distribution system in some urban areas in Nigeria is as high as 50%. High levels of water losses are indicative of poor governance and poor physical condition of the Water Distribution System, WDS, (Mamlook *et al*., 2003). The amount of water loss in water distribution systems varies widely from one system to another, from as low as 3–7 % to as high as 50 % of distribution input volume in the well-maintained systems of developed countries and less maintained system in developing countries respectively (Lambert, 2002).

Regular maintenance of infrastructure also helps to maintain water efficiency levels and is more cost-effective than rehabilitation (Makaya,2014). Many water distribution systems in developing countries are operated under intermittent conditions (WWAP, 2014). As a result, water supply efficiency in these countries is compromised.

Losses from leaks that are discovered and repaired should be measured to determine the rate of loss and the total volume lost during the life of the leak. Three methods are suggested (from Leak Detection Productivity '') by Douglas (AWWA California Nevada section, 1992).

1. Use a container of known volume.
2. Use a hose and a meter.
3. Calculate losses using modified orifice and friction formula.

An effective leakage management strategy should take into account the pressure dynamics of a water distribution network. This is because pressure plays a pivotal role in enhancing the magnitude of water leakage. This is because there is a physical relationship between leakage flow rate and pressure. Thus, the pressure exerted by either gravity or by water pumps results in a corresponding change in leakage rate. The frequency of new pipe bursts is also a function of pressure such that the higher or lower the pressure, the higher or lower the leakage. Pressure level and pressure cycling strongly influence burst frequency. Some of the most important ways of managing pressure is by either using pressure reducing valves (manual or automatic) or by using variable speed pump controllers. Under normal circumstances a pressure reducing valve is used to maintain a fixed downstream pressure regardless of the upstream pressure dynamics. The leakage from water distribution systems has been shown to be directly proportional to the square root of the distribution system pressure as indicated by the relationship (Wallingford, 2003).

Evidence shows that the rate of increase of bursts is more than linearly proportional to pressure. Indeed, it has even been suggested that there could be a cubic relationship, i.e. burst frequency proportional to pressure cubed (Farley and Trow, 2003).

Most software such as EPANET, is a widely used water distribution network simulator developed by the Environmental Protection Agency (EPA), requires that sub-components for distribution storage and piping be inputted with the necessary information.

Nash-Sutcliffe simulation efficiency (ENS) indicates how well the plot of observed versus simulated value fits the 1:1 line. The Nash–Sutcliffe model efficiency coefficient is used in assessing the predictive power of hydrological models. Nash–Sutcliffe efficiency ranges from infinity to 1. The value of efficiency of 1 (when E = 1) means there is a perfect match of modeled discharge relative to the observed data. The value of efficiency equal to (when E = 0) shows that the predictions of model are as accurate as the mean of the observed data, whereas an efficiency below zero (E < 0) occurs when the observed mean is a better predictor than the model or, in other words, when the residual variance, is larger than the data variance (the denominator). Therefore, the closer the model efficiency is to 1, the more accurate the model is (Karthikeyan et al. 2013). And according to Dongquan et al. (2009), an E_{NS} greater than 0.5 indicates acceptable model performance for model simulation.

Artificial Neural Networks (ANN) comprise of a network of neurons and take the cue from their biological counterparts. ANNs have found wide application in modelling water resources management problems including leakage detection, water distribution network optimisation, water pipeline replacement and rehabilitation, water demand forecasting, and pressure monitoring. Hamideh *et al*. (2021) proposed a new method to locate a leakage in WDNs using feedforward artificial neural networks (ANNs).

Methodology

Water Distribution Network Simulation

The hydraulic machine, EPANET was used for the hydraulic modelling of the networks

Other software machines employed are for data collection to accomplish this assignment include: ArcGIS, AutoCAD and Google Earth Pro.

Shapefiles from digitized map of transmission and distribution mains, reservoirs, tanks and valves were loaded to AutoCAD all geo referenced. These shape files loaded into AutoCAD were converted to metafile and used as backdrop in EPANET. The simulated backdrops were saved as NET File or INP file in EPANET interface

The shapefiles were as well converted to KML and superimposed in google earth to obtain nodal elevation values. The shapefiles were equally loaded in AutoCAD and then converted to DXF file for terrain extractor to assign the nodal elevation values as check for nodal values. TCX converter utilized as well to verify correctness of key nodal point values which were viewed in excel sheet..

Comprehensive data analyses were carried out, Geo referenced network maps successfully loaded on to EPANET interface for modelling.

Model Calibration

Nash-Sutcliffe simulation efficiency (ENS) indicates how well the plot of observed versus simulated value fits the 1:1 line. The Nash–Sutcliffe model efficiency coefficient is used in assessing the predictive power of hydrological models, and it is defined as

$$E = 1 - \frac{\sum_{t=1}^{T}(Q_o^t - Q_m^t)^2}{\sum_{t=1}^{T}(Q_o^t - \bar{Q}_o)^2} \quad (1)$$

Where;

Q_o = mean of observed discharges, and

Q_m = modeled discharge and

Qo^t = observed discharge at time t.

Nash–Sutcliffe efficiency ranges from infinity to 1. The value of efficiency of 1 (when E = 1) means there is a perfect match of modeled discharge relative to the observed data. The value of efficiency equal to (when E = 0) shows that the predictions of model are as accurate as the mean of the observed data, whereas an efficiency below zero (E < 0) occurs when the observed mean is a better predictor than the model or, in other words, when the residual variance (numerator in equation (1), is larger than the data variance (the denominator). Therefore, the closer the model efficiency is to 1, the more accurate the model is (Karthikeyan et al. 2013). And according to Dongquan et al. (2009), an E_{NS} greater than 0.5 indicates acceptable model performance for model simulation.

Neural network construction predicts the independent variable giving the available information of independent variables, Neural networks are made up of a series of layers with each layer comprising at least one neuron. While intermediate layers (hidden layers) perform the data processing functions of the network, the first and last layers input and output variables respectively. Within the hidden layers, weights to the neurons are adjusted by training the network in accordance with the stipulated learning rule (Zealand et al., 1999).

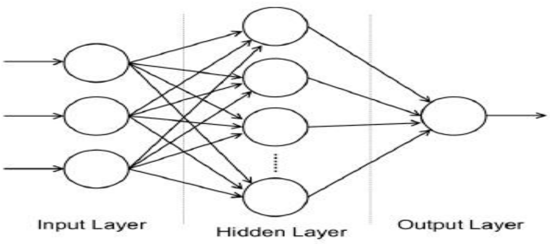

Figure1: Neural Network Diagram

3Model Calibration using NASH Sutcliffe Efficiency Coefficient and Artificial neural network

Emitters were used to simulate leakages at nodes. This is given by the equation

$$Q = a * P^b \qquad (2)$$

Where Q = leakage (Q_{leak}), a and b are discharge coefficient and emitter exponent respectively and P is the pressure at the node.

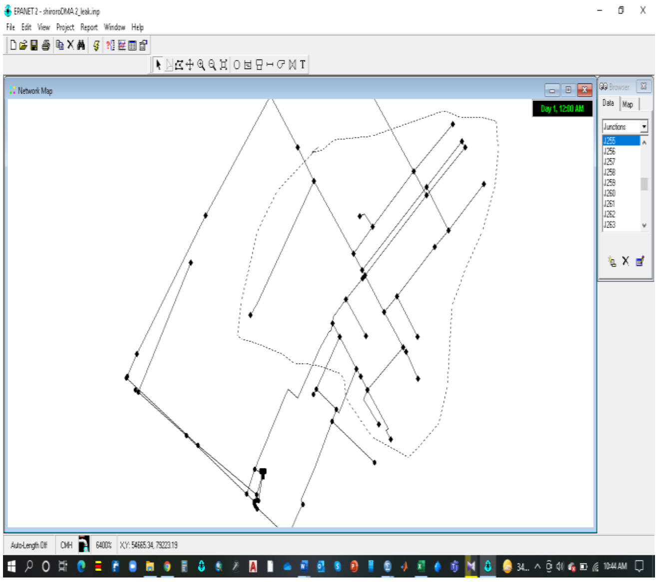

Figure 2: EPANET interface of Shiroro DMA showing the selected nodes for analysis

Result

Modelled and Observed Data Test in NS

Analyses at 8 to 11th hours

Using the leak coefficients of 0.1, 0.15, 0.2 and 0.3 in the emitter equation

$Q_{leak} = a * P^N$ at 8 and 9 hours, observed and modelled data loaded in the NASH provided the NASH Sutcliffe Efficiency Coefficients of -4.552, 0.092, 0.73, 0.187 and -3.777, 0.143, 0.68 and -0.07. NSE at 10 and 11 hours are -2.573, 0.286, 0.582, -0.288 and -0.689, 0.256, 0.516 and -0.826 These values deviated from the required standards of perfect or nearly perfect match except at 0.2 which gives a nearly perfect match

The performance of the model has suggested that using the emitter coefficient of 0.2 can model the study area. Table 1 shows the model performance in NSE

Table 1: Summary of the Model performance in NSE

	a			
Hour	0.1	0.15	0.2	0.3
8	-4.552	0.092	0.73	0.187
9	-3.777	0.143	0.68	-0.07
10	-2.573	0.286	0.582	-0.288
11	-0.689	0.256	0.516	-0.826

Summary of the of the modelled and measured leak is shown in Table 2

Table 2: Simulated and Observed Leaks at the site

	Base Demand (m^3/h)	Demand (m^3/h)	Pressure (m)	Simu_Q_{leak} (m^3/h)	Obs_Q_{leak} (m^3/h)
249	0.87	1.67	16.08	0.8	1
242	0.87	0.87	24.16	1	0.7
252	3.86	4.61	13.88	0.7	0.8
253	0.87	1.59	12.83	0.7	0.6
250	0.87	1.65	15.32	0.8	0.1
252	3.86	4.59	13.28	0.7	0.9
252	0.87	1.6	13.27	0.7	1
243	0.87	1.81	22.06	0.9	1
254	3.86	4.51	10.53	0.6	0.5
253	0.87	1.55	11.53	0.7	0.6
252	0	0.72	12.79	0.7	0.6
252	0	0.71	12.72	0.7	0.8
254	0.87	1.52	10.53	0.6	1.1
254	3.86	4.51	10.55	0.6	0.8
251	3.86	4.59	13.33	0.7	0.6
251	3.86	4.6	13.67	0.7	0.6
0	3.86	7.11	264.68	3.3	3
248	0.87	1.69	16.72	0.8	0.8
246	0.87	1.74	18.71	0.9	1

Having established this, the values of the model can now be used to predict leakages in the DMA using Artificial neural Network, ANN. The study has shown 17.1% of loss in the network.

Table 3: Summary of flow logging data

Elevation (m)	Base Demand (m^3/h	Demand (m^3/h)	Pressure (m)	Simu_Q_{leak} (m^3/h)
248	0.87	1.69	16.72	0.8
246	0.87	1.74	18.71	0.9
246	0.87	1.74	18.71	0.9
252	0.87	1.58	12.71	0.7
251	0.87	1.61	13.62	0.7
248	0.87	1.69	16.61	0.8
250	0.87	1.63	14.55	0.8
249	0.87	1.66	15.52	0.8
255	3.86	4.47	9.19	0.6
247	3.86	4.69	17.14	0.8
254	3.86	4.5	10.29	0.6
248	3.86	4.67	16.29	0.8

Table 4: Model Validation Result in ANN

Elevation (m)	Base demand (m^3/h)	Actual Demand (m^3/h)	Pressure (m)	Leak (m^3/h)	MLP_PredictedValue (m^3/h)
249	0.87	1.67	16.08	0.8	0.79
242	0.87	0.87	24.16	1	0.98
252	3.86	4.61	13.88	0.7	0.7
253	0.87	1.59	12.83	0.7	0.7
250	0.87	1.65	15.32	0.8	0.77
248	0.87	1.67	15.89		0.8
246	0.87	1.72	17.89		0.85
246	0.87	1.72	17.89		0.85
252	0.87	1.56	11.89		0.69
251	0.87	1.59	12.8		0.72

In this model calibration, sum of square errors for training and testing are 13.45% and 5.1% respectively.

The sum of square errors for samples trained and tested is depicted in Table 5 Table 6 indicate the model summary in percentages of the valid samples.

Table 5: Model Summary

Training sum of square error	Testing sum of square error
	0.051

Table 6: Case Processing Summary

Training samples	Testing samples	Validity	% trained	% tested	% valid	Samples excluded
83	28	111	74.8	25.2	100	37

The predicted and the real values of leaks are depicted in Figure 3

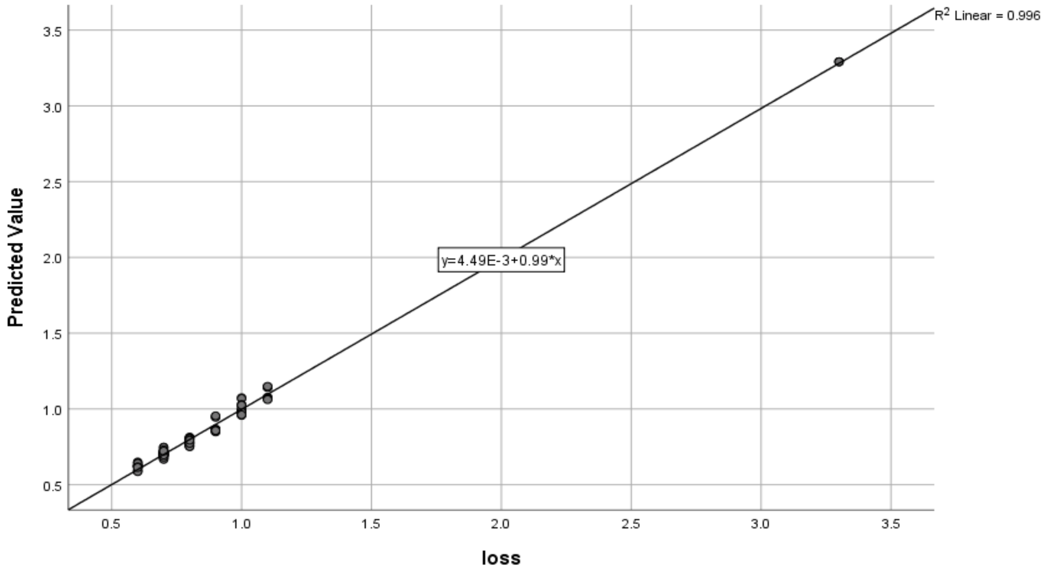

Figure 3: Real loss and predicted values of loss

The result showed that the model built can estimate the amount of leak, given elevation, base demand, demand, pressure and head as variables. This can be useful for water utilities in pipe inspection and maintenance. The value of R^2 indicates the model is doing well in terms of prediction

Conclusions

The main objective of this study is to model water loss in the distribution network of Shiroro District Metered Area. The model developed can be used in any district metered area of a distribution network to estimate the loss. R^2 linear .97% The errors are 13.4% and 5.1% respectively for training and testing of the variables in the machine learning. The input variables are elevation, base demand, demand and pressure of the network. The developed model is expected to help set the direction of improvement of the analysis of water distribution system and optimal operation of water supply in the studied DMA and other DMAs. This study has shown 17.1% of physical or real loss as NRW. This study has shown 17.1% of physical or real loss as NRW.

References

Dong-quan, Z. O., Hao-zheng, W. G., Ji-ning, C. N. & Hao-chang, W. G. (2009). Parameters uncertainty analysis of urban rainfall-runoff simulation, 20(1), 45-51.

Farley, M. (2003). Non-revenue water - International Best Practice for Assessment, Monitoring and Control Annual CWWA Water, Wastewater and Solid Waste Conference. Atlantis, Paradise Island, Bahamas.

Hamideh, F., Muhammadreza, J., Babak, A. & Jafar, Y. (2021). Ghazizadeh Leakage Detection in Water Distribution Networks Using Hybrid Feedforward Artificial Neural Networks

Karthikeyan, L., Kumar, D. N., Graillot, D. & Gaur, S. (2013). Prediction of ground water levels in the uplands of a tropical coastal riparian wetland using artificial neural networks. *Water resources management,* 27(3), 871-883.

Lambert A. O. (2002). A Water losses management and techniques. Water Science and Technology: Water Supply, 2(4), pp1–20,

Makaya, E. (2014). Water distribution systems losses. A guide for operation and maintenance. Cibew Publishing Company, Botswana, ISBN 978-99968-0-269-0.

Mamlook, R. and H. Al-Jayyous. (2003) Fuzzy sets analysis for leak detection in infrastructure systems: a proposed methodology. Cleaning Techniques and Environmental Policy, 6(1), pp 26-33.

Rossman, L. A. (2000). EPANET 2 User Manual. Water Supply and Water Resources Division National Risk Management Research Laboratory, Cincinnati, OH, EPA/600/R-00/057.

Segmasin Nig. Ltd. (1998) Assessment of unaccounted water in Bida and New Bussa, Nigeria. A report prepared in collaboration with Parkman Nig Ltd and submitted to Niger State

WWAP, (2014). United Nations World Water Assessment Programme, United Nations, New York.

Zabidi, H. A., Goh, H. W., Chang, C. K., Chan, N. W., & Zakaria, N. A. (2020). A review of roof and pond rainwater harvesting systems for water security: The design, performance and way forward. *Water*, *12*(11), 3163.

Zealand, C.M., Burn, D.H. and Simonovic, S.P. (1999). Short Term Streamflow Forecasting using Artificial Neural Networks. Journal of Hydrology, 214, 32-48.