Malaysian Journal of Computing, 6 (1): 642-649, 2021

eISSN: 2600-8238 online

DETERMINANTS OF NON-REVENUE WATER

Haslinda Ab Malek[1*], Mohamad Hafizi Zakaria[2], Muhammad Luqman Zulkifli[3] and Nur Farahin Roslan[4]

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Negeri Sembilan, Kampus Seremban, Negeri Sembilan, Malaysia

[1*]haslinda8311@uitm.edu.my, [2]hafizi.zakaria97@gmail.com, [3]muhammadluqmanzulkifli@gmail.com, [4]alinroslan@gmail.com

ABSTRACT

Water is one of the most essential needs in human daily life. Water losses or Non-Revenue Water (NRW) refers to the treated water that has been produced from water plant which did not reach to the customer. This waste of water has caused the company to suffer losses and hence, burdens the people with increasing water tariff. Moreover, it becomes one of the challenges for commercial water system management because the water company must fulfil the demand from the society which keep increasing day by day. In addition, the demand for water is increasing, as the population is growing. Despite having the rainfall throughout the year in Malaysia, many cities are experiencing water shortage and frequent water supply disruptions. Therefore, efficient management of water distribution is required to minimise the water losses and to make sure the sustainability of water reserve for a long period. This study focuses on identifying the significant factors that influence the Non-Revenue Water and modelling the data using Multiple Linear Regression Model. The sample size used in this study were 212 observations and the variables involved were Length of Connection, Number of Connection, Production Quantity, Consumption Quantity and Non-Revenue Water. It is found that the variables of Number of Connection, Consumption Quantity and Production Quantity were significant to Non-Revenue Water whereas the variable of Length of Connection was not significant. It is hoped that the result from this study can be used by the water authority company in improving the water distribution and thus reduce water losses and cost.

Keywords: *Non-Revenue Water, Multiple Linear Regression, Consumption, Water Losses, Customer.*

Received for review: 20-08-2020; Accepted: 23-11-2020; Published: 01-01-2021

1. Introduction

Water is one of the most essential needs in human daily life. Despite having a large percentage of water on Earth, only one percent of water that is safe to drink or use for washing, showering and many other activities for our daily routine. In addition, the increase of the world population every year causes the rise of the demand of the water and therefore leads to water depletion. It is expected in the year 2030, almost 50 percent of the world population will face water crisis. Therefore, efficient management of water distribution is required to minimise the water losses and to make sure the sustainability of water reserve for a long period.

Water losses or known as Non-Revenue Water (NRW) refers to the treated water that has produced from water plant which did not reach to the customer and did not give any revenue. In Malaysia, water losses in 2014 are estimated to achieve 35 percent which means that more than 4.27 billion litres of treated water are leaking out before the water supply reaches consumer's house. Non-Revenue Water is the indicator of the quality of the national water supply services. The high Non-Revenue Water level of the country reflects that huge volume of water lost during the water distribution to the consumer and thus indicates the poorly managed water system by the water supply services.

In addition, with the growing population, the demand of water is increasing. Despite having the rainfall throughout the year in Malaysia, many cities have reported to be experienced water disruption. Water supply disruption may lead to cost increases and decline in productivity. It has become a thriving concern since the loss of water through the water distribution system is achieving to more than 30 percent per year. The losses also affect the consumers because they may result in rising water tariffs to cover the lost revenue.

Therefore, effective water management plays an important role in reducing water losses. Hence, this study is conducted to investigate the determinants that causing the Non-Revenue Water. This study is expected to help to improve the analysis of water distribution systems in Malaysia. It is hoped that this study will assist the management in implementing effective strategies to reduce the Non-Revenue Water level. The developed model is expected to improve the analysis of water distribution systems in Malaysia and optimise them. Hence, this study can help water management authority to predict the Non-Revenue Water level using the best model and minimize the maintenance and recovery cost.

Length of connection or length of network per kilometre is known as a common factor affecting to water loss. The study carried by Alkasseh *et al.* (2013) stated that length of network had a significant positive relationship with Non-Revenue Water. It indicates that the longer the length of the network produced the higher water loss percentage. This statement corroborates with a study done by Van den Berg (2014). The study claimed a positive linear relationship between length of network and the water loss rate because a larger network is more costly to maintain and preserve functionally.

The number of connections refers to the number of pipes connected legally with the main pipeline that carries water from the water supply. Liemberger (2010) stated that to estimate the Non-Revenue Water, apart from the length of the distribution system, the number of service connections also needed to be identified. Moreover, Korevaar *et al.* (2013) also mentioned that the number of connections can lead to the increase of Non-Revenue Water. A regression analysis that was carried out by Murrar (2017) indicated that the more connection in the network, the more increasing the water loss is. This is because when the number of connection increases, the probability of water out through the intersection will also increase.

Moreover, water consumption also influences the non–revenue water ratio. Water consumption can be defined as the total volume of treated water that passes through the consumer's registered meter. This consumption is important because the water supplier can identify if there are water leakage and water loss by looking at the differences the total water treated supplied and the total of water treated passes through the consumer's meter. Aubuchon & Roberson (2012) stated that the increase of water consumption with increasing total profit, may provide the budget to recover any damage or leakage and can reduce the Non-Revenue Water Ratio by investing in maintenance cost. Murrar (2017) found that there is a negative relationship between consumption volume on Non-Revenue Water per connection. This indicates that more water consumption will reduce the Non-Revenue Water.

In addition, production quantity also contributes to the non – revenue water ratio. Production quantity refers to the total volume of treated water that is distributed from the water supply. Liemberger (2010) concluded that increasing water production is a key strategy in reducing Non-Revenue Water. Murrar (2017) mentioned that Non-Revenue Water in the year 2015 decreased from 50% in the year 2003 to 23% due to increase in production quantity from 12 in the year 2003 to 91 litres per person per day in the year 2015. This indicates that there is a negative relationship between production quantity and Non-Revenue Water.

Numerous researchers in determining the factors that contribute to Non-Revenue Water had used multiple linear regression analysis. Van den Berg (2016) conducted a study covering the utilities in 69 countries from 2011 to 2016 to determine the effective factors to reduce the Non-Revenue Water using regression analysis. Moreover, Korevaar *et al.* (2013) also used regression analysis for modelling the Non-Revenue Water performance. In the study, the backward selection method was used for selecting the significant variables that contribute to the Non- Revenue Water performance.

2. Methodology

2.1 Description of Variables

The data used in this study is secondary data consists of 212 observations. The independent variables involved in this study are Length of Connection, Number of Connection, Production Quantity, Consumption Quantity while the dependent variable is Non-Revenue Water. The description of each variable is shown in Table 1.

Table 1. Description of the Variables

Variables	Description
Non-Revenue Water	The measurement of water losses (percentage)
Length of Connection	The length of service connection (kilometres)
Number of Connection	The number of service connection
Production Quantity	Total volume of treated water that are distributed from the water supply (million litres per day)
Consumption Quantity	Total volume of treated water that passes the consumer's registered (million litres per day)

2.2 Method of Analysis

Multiple regression analysis is an analytical technique that estimates causality between variables by statistical methods, and a method to analyze the regression model with a dependent variable and two or more independent variables (Sekaran & Bougie, 2016). The multiple linear regression model with independent variables was expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon_i \quad (1)$$

Where
Y = Non-Revenue Water
β_0 = Regression intercept
$\beta_{1, 2, 3, 4}$ = Regression coefficients
x_1 = Length of Connection
x_2 = Number of Connection
x_3 = Production Quantity
x_4 = Consumption Quantity

ε_i = Error

The model needs to fulfil several criteria to make sure it is valid to make any inference. The assumptions of multiple linear regression include the facts that there is a linear relationship between the dependent variable and the independent variables; the error term must be normally distributed and independent, no multicollinearity exists, and the error variances must be constant.

3. Results

3.1 Multiple Linear Regression

Multiple linear regression was used to determine the significant factors that influence the Non-Revenue Water. This regression modelling carried out by using Statistical Package for the Social Sciences (SPSS). Table 2 shows the result of multiple linear regression analysed by using all the variables that have been identified namely as Length of Connection, Number of Connection, Production Quantity and Consumption Quantity as the independent variables and Non-Revenue Water as the dependent variable. The result of regression analysis indicates that only three variables were significant which are Number of Connection (p-value=0.010), Production Quantity (p-value = 0.000) and Consumption Quantity (p-value = 0.000). Other than that, the variable of Length of Connection was insignificant since the p-value was 0.587 which is more than significant value of 0.05.

Table 2. Results of Multiple Regression Analysis Using All Parameters

Model 1	**Unstandardized Coefficients**		**Standardized Coefficients**	**t**	**Sig.**
	B	**Std. Error**	**Beta**		
(Constant)	33.858	0.990		34.194	**0.000**
Length of Connection	-0.060	0.111	-0.21	-0.545	0.587
Number of Connection	-0.003	0.001	-0.117	-2.592	**0.010**
Production Quantity	0.032	0.001	1.831	31.303	**0.000**
Consumption Quantity	-0.047	0.002	-1.671	-29.165	**0.000**

Table 3 shows the results of multiple regression with the model regressed with significant independent variables only. It also revealed that variables of Number of Connection (p-value=0.010), Production Quantity (p-value=0.000) and Consumption Quantity (p-value=0.000) were significant since the p-value for each variable is less than significant value of 0.05. It means that for every increase in the number of connections, the non-revenue water is decreased by 0.004%. For the variable of Production Quantity, it indicates that for every 1 million litres per day increase in production quantity, the non-revenue water also will increase by 0.032%. Moreover, for the Consumption Quantity, it can be concluded that there is decreased of non-revenue water for every 1 million litres per day increase in Consumption Quantity.

Table 3. Results of Multiple Regression Analysis using Significant Independent Variables

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	35.732	1.158		35.084	0.000
Number of Connection	-0.004	0.001	-0.128	-3.186	0.002
Production Quantity	0.032	0.001	1.823	32.260	0.000
Consumption Quantity	-0.047	0.002	-1.667	-29.357	0.000

Table 4 shows the result of Analysis of Variance (ANOVA). The p-value of the F test has a lower value (0.000) than significance value, α (0.05) indicates that the model is valid. Therefore, the model can be used for modelling the data and predict the Non-Revenue Water.

Table 4. Analysis of Variance (ANOVA)

Source of Variation	Sum of Squares	df	Mean Square	F	Sig.
Regression	49544.533	3	16514.851	371.053	0.000
Residual	9257.690	208	44.508		
Total	58802.243	211			

The value of R Square from Table 5 is 0.843. It shows that the independent variables had 84.3% explained on the non-revenue water while other factors explain 15.7%. The value of standard error of estimate represents how far the data fall from the regression line or regression predictions on the scale of the measurements result. Table 5 shows that the average distance of the data falls from the regression lines is about 6.7% on the dependent variable (Non-Revenue Water). In addition, it shows that the value of Durbin-Watson is 2.162. According to Field (2009), this value is relatively normal and acceptable since it lies in the range of 1.5 to 2.5, which indicates that the residuals are independent.

Table 5. Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.918	0.843	0.840	0.067	2.162

3.2 Model Adequacy Checking

The model adequacy for the regression model found needs to be checked to satisfy the assumption of multiple linear regression model. The scatter matrix was plotted in Figure 1 to check the linearity of Non-Revenue Water against the independent variables (Number of Connection, Production Quantity and Consumption Quantity). Figure 1 depicts a negative linear relationship between Number of Connection and Non-Revenue Water and Consumption Quantity and Non-Revenue Water. It proved that the value of Non-Revenue Water is decreased as the Number of Connection and Consumption Quantity increases. Moreover, Figure 1 also shows a positive linear relationship between Production Quantity and Non-Revenue Water. It indicated that the Non-Revenue Water increases when the Production Quantity increases.

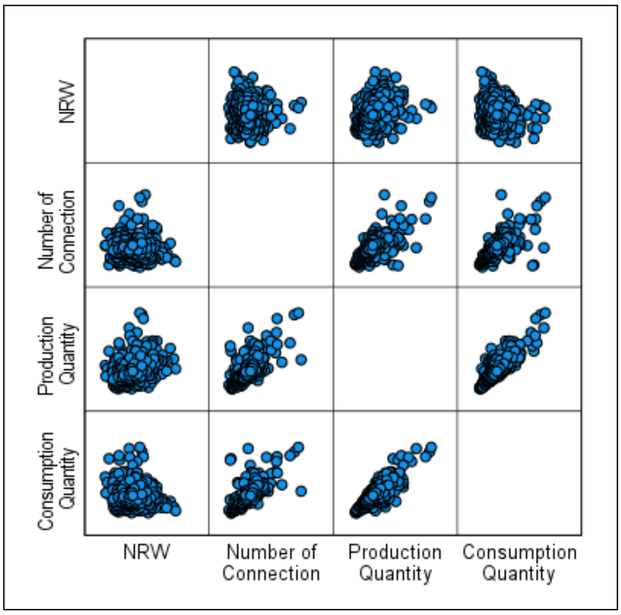

Figure 1. The Partial Regression Plot

The normal probability plot of residuals was produced in Figure 2. Since the plots are close on the 45-degree line, it can be concluded that the error terms are approximately normally distributed. In addition, Table 6 shows the values of skewness and kurtosis to check the normality of the error term. The descriptive statistic of skewness and kurtosis value were -0.197 and 0.055 respectively. Since both values are in the range of ±2, this implies that error terms are normally distributed. Moreover, the assumption of the normality of error term also has been confirmed by conducting the test of normality as shown in Table 7. The p-value (0.077) of Shapiro-Wilk was more than the level of significance(α=0.05) indicating that the error term is normally distributed.

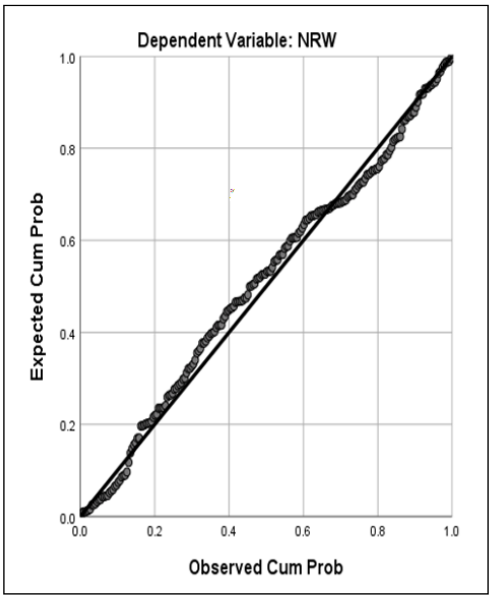

Figure 2. Normal P-P Plot of Regression Standardized Residual

Table 6. Descriptive Statistic of Skewness and Kurtosis

		Statistic	Std. Error
Unstandardized Residual	Skewness	-0.197	0.167
	Kurtosis	0.055	0.333

Table 7. Test of Normality

	Shapiro-Wilk		
Unstandardized Residual	Statistic	df	P-Value
	0.988	212	0.077

Table 8 shows the value of Tolerance and Variance Inflation Factor (VIF) to diagnose the multi-collinearity problem. The value of tolerance statistics for the variables exceeds the cut-off point which is 0.1 and the VIF values less than 10. These indicate that there was no multi-collinearity exist.

Table 8. Collinearity Diagnostic

Model	Collinearity Statistics	
	Tolerance	VIF
Number of Connection	0.470	2.126
Production Quantity	0.237	4.218
Consumption Quantity	0.235	4.262

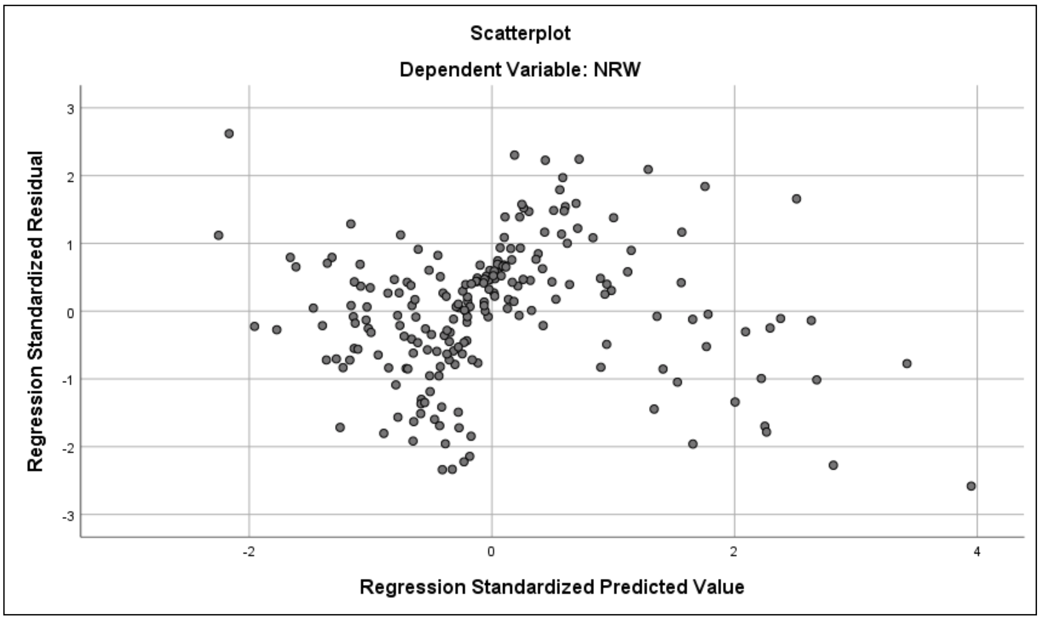

Figure 3. Scatterplot of Residual Against Predicted

The plot of residual against the predicted value is shown in Figure 3 to check the homogeneity of error variances. The scatterplot shows that there is no obvious pattern and indicates that the error variance is constant. This result also confirmed by heteroscedasticity test with p-value of 0.0927. Since this p-value (0.0927) is more than the level of significance(α=0.05), it revealed that the error variances are homogeneous. Since all the assumptions of multiple linear regression were fulfilled, it can be concluded that the regression model found is valid and can be used to make a prediction.

4. Results

Based on the regression analysis that has been conducted, it can be concluded that the variables of Number of Connection, Production Quantity and Consumption Quantity were statistically significant to the Non-Revenue Water. Variables of Number of Connection and Consumption Quantity has a negative linear relationship with the Non-Revenue Water. This result was consistent with the study done by Adams & LutzLey (2012), Aubuchon & Roberson (2012) and Murrar (2017). Meanwhile, the variable of Production Quantity has a positive linear relationship with the Non-Revenue Water. This result also conforms by Van den Berg (2014) in his study, which is the higher the production quantity, the higher the non-revenue water is. Moreover, the variable of Consumption Quantity is the most significant variable since the value of beta (B) coefficients is higher than the variable of Number of Connection and Production Quantity. The value of R square of the model is 0.843 indicates that the model is a good fit and can be used to predict the Non-Revenue Water.

References

Adams, A. S., & LutzLey, A. N. (2012). An empirical study of factors affecting water loss in Mexican cities. Journal of Water Sustainability, 2(3), 167-178.

Alkasseh, J. M., Adlan, M. N., Abustan, I., Aziz, H. A., & Hanif, A. B. M. (2013). Applying minimum night flow to estimate water loss using statistical modeling: a case study in Kinta Valley, Malaysia. Water resources management, 27(5), 1439-1455.

Aubuchon, C. P., & Roberson, J. A. (2012). Price perception and nonprice controls under conservation rate structures. Journal-American Water Works Association, 104(8), E446-E456.

Field, A. (2009). Discovering statistics using SPSS. London: SAGE Publications Ltd.

Korevaar, C. B., Schaefer, W. F., Glumac, B., van Weenen, I. B., & Riemersma, I. M. (2013). Modeling the relationships towards the performance of non-revenue water.

Liemberger, R. (2010). Recommendations for Initial Non-Revenue Water Assessment. IWA Water Loss, 1, 1-13.

Murrar, A. (2017). The Determinants of Non-Revenue Water in Balkan Countries. American Journal of Water Science and Engineering, 3(2), 18-27.

Sekaran, U., & Bougie, R. (2016). Research methods for business: A skill building approach. John Wiley & Sons.

Van den Berg, C. (2014). The Drivers Of Non-Revenue Water: How Effective Are Non-Revenue Water Reduction Programs?. The World Bank.