Effect of the leakage location pattern on the speed of recovery in water supply networks

W Nugroho¹,* and N Iriawan²

¹ Student of Technology Management Department, Sepuluh November Institute of Technology, Surabaya 60111, East Java, Indonesia
² Statistical Department, Sepuluh November Institute of Technology, Surabaya 60111, East Java, Indonesia

*widyo.nugroho@gmail.com

Abstract. The International Water Association states that Non-Revenue Water is an efficient indicator of the performance of a water company that has been implemented internationally. The volume of water loss expressed as Non-Revenue Water at the global level shows a high number. Every year more than 32 billion m³ of treated water is lost due to leakage from the distribution network. There are four pillars in the network leakage management strategy include management of pressure, speed of recovery, active leakage control, and asset management. These components influence how leakage is managed and affect the volume and value of economic leakage in a distribution network of water companies. The current annual volume of physical losses tends to increase along with the expansion service area of the distribution network. The rate of increase can be inhibited by the right strategy combining the four components into an effective leakage management strategy. This study analyzes the distribution pattern of leakage points in water distribution pipelines in Malang City, and how much they affect the speed of recovery. The data in this study consist of 5449 pipeline leakage points which include spatial and attributes data. The method used in this study is the Spatial Point Process. The results of the analysis show that the distribution pattern of leakage is an inhomogeneous pattern. It is found that the parameters of the model -with Euclidean distance- the leakage pattern affects the occurrence of leakage and the Geographical Weighted Regression model -with fixed Gaussian Kernel Weight Function with Manhattan distance- the leakage pattern affects the speed of recovery.

1. Introduction

Public infrastructure has an important influence in improving the quality degree of the community life, as well as being a driving factor for economic growth. Infrastructure in the water supply sector, especially the water supply system should provide assurance of public access to the availability of water. But in its implementation, when reviewed both in terms of quality, quantity, and continuity, it has not run continuously. On the other hand, the demand for healthy standard water continues to increase without being balanced with improvements in the quality of services. The water supply system has decreased in quality over time, which usually leads to a variety of problems such as the reduced utility of hydraulic facilities, increased water loss, service disruption, and reduced water quality. In addition, a gradual increase in consumer demand for water creates new problems such as low pressure on customer connections. This increases air pressure in the water distribution system, which in the future can increase the frequency of network leakages. The International Water Association (IWA) states that the Non-
Revenue Water parameter is an efficient performance indicator of a water company that has been applied internationally. Obradovic and Landsdale state that Non-Revenue Water, in general, is the difference between the amounts of water supplied with water consumed [1]. This loss can be technical in nature, for example, water loss in the pipeline itself, while non-technical loss such as water theft in the distribution pipeline. Furthermore, Obradovic and Landsdale stated that in the development of water supply systems, water loss was calculated so that service points could still be fulfilled continuously [1].

Farley and Liemberger state in more detail that water loss, expressed as the ratio of Non-Revenue Water equals the amount of water supplied into the water distribution network (as System Input Volume) minus the amount of water that officially becomes the account of the customer (as legal Consumption) [1]. Non-Revenue Water components cover the entire water company service system starting from the water treatment plant (WTP) outlet meter recorder to the customer meter recorder. The volume of water loss expressed as a ratio of NRW at the global level shows astounding numbers. Every year more than 32 billion m$^3$ of treated water is lost due to leakage from the distribution network. Meanwhile, 16 billion other m$^3$ is distributed to customers without being billed for water theft, inaccurate meter reading, or corruption. A conservative estimate of the total annual cost of water companies worldwide is the US $14 billion. In a number of developing countries, this loss represents 50% to 60% of water services with 35% of global estimation. It is estimated that half of this amount will have implications for increasing water services to 100 million other residents without further investment [2].

In developed countries, management and control of water loss are supported by the existence of adequate infrastructure and operational implementation and is supported by the awareness that the NRW ratio is an important parameter for the performance of the water supply authority. Researchers in South Korea are aware that NRW is an important parameter for setting priorities for improving water distribution systems. For this reason, several researchers in South Korea have identified parameters that affect water loss. In a study conducted in Incheon City, South Korea, Dongwoo Jang and Gyeewon Choi implemented Artificial Neural Network (ANN) to improve the accuracy of calculations and identify parameters that significantly affected NRW ratio [3]. There were two classifications of parameters, including operational parameters that consists of the ratio of energy requirements and the number of network leaks, as well as physical parameters consisting of the average pipe diameter, pipe length per number of customer connections, supply water discharge per number of customer connections, and pipe deterioration ratio [4].

In the study conducted in Incheon City, Dongwoo Jang and Gyewoon Choi estimated the NRW ratio using Artificial Neural Network (ANN) and Multivariate Regression Analysis (MRA) by including specific parameters that affect the frequency of leaks in the water distribution system in Incheon City [5]. This study was carried out in the initial stages using MRA and ANN to estimate the NRW ratio, then comparing the results of the two methods. Accuracy assessment shows that the ANN model has a higher prediction accuracy than the MRA. The results of this study subsequently become a model for estimating the NRW ratio using ANN based on certain parameters that affect leakage in the water distribution system in South Korea. Accuracy assessment and plot distribution analysis was used to select the optimal ANN model cases, which in this case were special parameters that affected leaks in the Incheon water distribution system. In the study, Dongwoo Jang and Gyewoon Choi found that the quantity of water supply and pipe deterioration ratio were the main factors that had the most significant influence on NRW management and the ANN method had a higher accuracy compared to the MRA method [6]. On the other hand, many developing countries that instead manage NRW effectively, even to ensure that customers get a safe supply of water in sufficient quantities still requires hard effort. Coupled with the problem of distribution in which processing water is channeled through inadequate pipelines with weak management systems and low levels of technical and technological capabilities. The tariff setting and revenue collection policies often do not reflect the true value of supplied water, limiting the company's income and encouraging customers to give low value to the services provided. This is because not all cities or regions, especially in developing countries, have the infrastructure and operational procedures that meet the requirements for handling NRW [7].
Water losses can occur in all parts of the distribution network, and can even occur on newly installed networks. Physical water loss is generally the largest volume loss component in the water balance analysis. Physical water loss is often referred to as actual water loss or in terms of practitioners known as leakage. This loss includes the total volume of water loss but does not include losses due to business and commercial processes. In the analysis of the water balance, it is shown that some of the physical loss rates are the result of estimation and thus the results of calculating the leakage volume are approaches that can be potentially wrong or inaccurate. The management of the water supply authority must verify the results of their analysis by conducting a component analysis that involves a top-down approach or a deeper assessment of the physical loss component [8].

There are four pillars in the physical loss management strategy which include pressure management, speed of recovery, active leak control, and asset management. These factors influence how the leakage can be managed and thus the volume and value of economic leakage, in the water company distribution network can be determined. The current annual volume of physical loss tends to increase along with the distribution network service area expansion. The rate of increase can be inhibited by applying the right combination of the four components into an effective leak management strategy [2].

In a water distribution network operations and maintenance practices, it is common for water supply authority to divide operational and maintenance divisions into several sub-divisions, which have only driven by the service zone boundaries [4]. While on the other hand, not all service zones require the same priority in handling recovery. Of course, this will determine the amount allocated to the region. An analysis that is involving spatial parameters in an effective leak management strategy is the needs of the current water supply authority. Based on the background above, in this study the distribution pattern analysis of the leakage location will be carried out to answer problem statements below:

- Is the distribution pattern of pipeline leakage locations categorized as a homogeneous pattern?
- Does the distribution pattern of pipeline leakage locations have a significant effect on the speed of recovery?

2. Methods and materials

In this study, the factors that influence water loss in the water distribution system are analyzed by focusing on the physical and operational aspects of the network. With the expansion of service areas, location distribution has an influence on the non-revenue water ratio. Finally, the expected result of this study is to find out the significance of the relationship between the pattern of pipeline leakage location with operational and physical network parameters. Speed of recovery is one of the water distribution operational parameter [1]. The data in this study consist of 5449 pipeline leakage points from year of 2018 operation period which include spatial and attributes data. The Poisson distribution in the Spatial Point Process was chosen in the approximation to the pattern of leakage points [9]. The homogeneity test in the Spatial Point Process aims to find out the intensity of the point pattern categorized into the homogeneous point pattern or the inhomogeneous point pattern.

The research variables used in this study consist of response and covariate. The covariate variable is the number of leakage events that have certain geographical coordinates, responding by the speed of recovery. The service zone boundary is used as the grid of model density, where the covariate variable is transformed into a pixel image and then extracted into 13 grids. Geographical Weighted Regression model -with a fixed Gaussian kernel weighting function- used as an approach to obtain the effect of the leakage location pattern on the speed of recovery.

The Bayesian algorithm used to estimate parameters as follows [10]:

- Forming the likelihood function of the Geographical Weighted Regression.
- Determining the prior distribution that depends on the frequentist result.
- Forming a posterior distribution based on the prior distribution.
- Forming full conditional posterior distribution for estimated parameters.
- Performing an iterative process of estimating parameters using Gibbs sampling.
- Determining the best model parameters.
3. Result and discussion

The homogeneity test in the Spatial Point Process aims to find the intensity of the leakage pattern categorized into the homogeneous point pattern or the inhomogeneous point pattern. The homogeneity test results in Table 1 show that the leakage pattern is categorized as an inhomogeneous point process.

| Stat Value | df | p-value |
|------------|----|---------|
| 20,846     | 11 | 0.035   |

\( H_0 \) : the intensity of the leakage pattern is homogenous

\( H_1 \) : the intensity of the leakage pattern is inhomogeneous

![Image of Planar point pattern of the leakage locations.](image)

Parameter estimation of the Geographical Weighted Regression model with a fixed Gaussian kernel weighting function is carried out using the weighted Maximum Likelihood Estimation method. Descriptive parameter estimation results are presented in Table 2 show that the speed of recovery carried out is between 1.973 and 2.023 days.

| Variable | Minimum Value | Maximum Value | Range |
|----------|---------------|---------------|-------|
| Intercept| 1.9740        | 2.0230        | 0.049 |
| \( X_1 \) | 0.3614        | 0.3734        | 0.012 |

After obtaining estimators of the model parameters, parameter testing is carried out to determine the influential parameters, presented in Table 3. The parameters testing of the model with the fixed Gaussian Kernel Weighting Function indicates that the test statistic value so that \( H_0 \) is rejected and found that the parameters of the model with Euclidean distance affect the occurrence of leakage. While the parameter testing of the model with Manhattan distance shows the test statistic value that \( H_0 \) is rejected and it is found that the parameters of the Geographical Weighted Regression model with the fixed gaussian kernel weight function with Manhattan distance affect the speed of recovery.
Table 3. Simultaneous parameter testing GWR model with fixed Gaussian weighting function and two distances.

| Fixed Gaussian Kernel       | F Score | F Table | Decision |
|-----------------------------|---------|---------|----------|
| Euclidean Distance          | 14.453  | 2.631   | Reject H0 |
| Manhattan Distance          | 14.452  | 2.631   | Reject H0 |

\[ H_0 : \beta_1(u_j, v_j) = \beta_2(u_j, v_j) = \ldots = \beta_p(u_j, v_j) = 0 \]
\[ H_1 : \text{at least one } \beta_k(u_j, v_j) \neq 0 ; k = 1, 2, \ldots, p \]

In Table 4, partially parameter testing of the model used to form the model and to determine the effect of each predictor variable.

Table 4. Parameters testing of the GWR model with the fixed Gaussian kernel weighting function with Euclidean distance.

| Parameter | Estimator | SD   | Z Score |
|-----------|-----------|------|---------|
| \( \beta_0 \) | 1.989     | 2.061| 0.965   |
| \( \beta_1 \) | 0.371     | 0.095| 3.897   |

\[ H_0 : \beta_k(u_j, v_j) = 0 \]
\[ H_1 : \beta_k(u_j, v_j) \neq 0 ; k = 1, 2, \ldots, p \]

In addition to parameter testing, it is necessary to conduct the goodness of fit test on the model to determine the difference between Geographical Weighted Regression models and Multiple Linear Regression Models. Based on Table 5, the F score is 1.196 with a p-value at the 5% significant level. It is found there is a significant difference between the multiple linear regression models and the weighted fixed Gaussian Kernel with multiple linear regression models, which Geographical Weighted Regression model with weighted fixed Gaussian Kernel is fitter than Multiple Linear Regression Models.

Table 5. The goodness of fit test of GWR model with fixed Gaussian kernel weighting function.

| Source            | JK     | Db      | KT      | F   | p-value |
|-------------------|--------|---------|---------|-----|---------|
| Residual Global   | 42.186 | 34.000  |         |     |         |
| GWR Improvement   | 40.121 | 32.027  | 1.253   | 1.196| 0.632   |
| GWR Residual      | 2.066  | 1.973   | 1.047   |     |         |

4. Conclusion

The results of the analysis show that the distribution pattern of leakage is inhomogeneous point pattern. Based on the model obtained by the computation result it is found that the parameters of the Geographical Weighted Regression model -with the fixed gaussian kernel weight function with Manhattan distance- affect the speed of recovery and it is found that the effectiveness of the recovery needs to consider covariate location variables such as number of demand junctions and boundary of the service zone.

Acknowledgements

We hereby thank you to Technology Management Department, Sepuluh November Institute of Technology, for supporting the publication of this research.
References

[1] Farley M and Liemberger R 2004 Developing a Non-Revenue Water Strategy, Part 1: Investigating and Assessing Water Losses *Journal of Water Science and Technology* **324**

[2] Farley M and Liemberger R 2005 Developing a Non-Revenue Water Strategy, Part 2: Planning and Implementing the Strategy *Journal of Water Science and Technology* **51** 41–50

[3] Jang D and Choi G 2017 Estimation of Non-Revenue Water Ratio for Sustainable Management Using Artificial Neural Network and Z-Score in Incheon, Republic of Korea *Journal of Water Science and Technology: Sustainability* **9**

[4] Jang D 2018 A Parameter Classification System for Non-Revenue Water Management in Water Distribution Networks Hindawi 2018

[5] Jang D and Choi G 2017 Estimation of Non-Revenue Water using MRA and ANN in Water Distribution Networks *Journal of Water, MDPI* **10**

[6] Jang 2018 Estimation of leakage ratio using principal component analysis and artificial neural network in water distribution systems *Journal of Water Science and Technology: Sustainability* **10** 3 1–13

[7] Ellen J L and Kellogg J S 2005 Deficiencies in Drinking Water Distribution Systems in Developing Countries *Journal of Water Health* **3** 109–27

[8] Kanakoudis V K and Tolikas D K 2016 The role of leaks and breaks in water networks-technical and economical solutions *Journal of Water Science and Technology: Sustainability* **10** 3 1–13

[9] Baddeley A, Rubak E and Turner R 2016 *Spatial Point Patterns Methodology and Applications with R* (CRC Press Taylor & Francis Group)

[10] Box G and Tiao G 1973 *Bayesian Inference in Statistical Analysis* (John Wiley and Sons Inc.)