Determining physical and operational factors influencing pipeline leakage location pattern in water distribution networks using spatial poisson point process

W Nugroho1,*, N Iriawan2 and C Utomo3

1Doctoral Student of Civil Engineering Department, Sepuluh November Institute of Technology, Surabaya 60111, East Java, Indonesia.
2Professor of Statistical Department, Sepuluh November Institute of Technology, Surabaya 60111, East Java, Indonesia.
3Associate Professor of Civil Engineering Department, Sepuluh November Institute of Technology, Surabaya 60111, East Java, Indonesia.

*widyo.nugroho@gmail.com

Abstract. There are four pillars in leakage management strategy for water distribution networks, consist of pressure management, speed of repairs, active leakage control, and asset management. Leakage can be managed by applying the right combination of those factors in the right leakage management strategy. This study conducted in Malang City, East Java, Indonesia. The factors that influence physical water loss in the water distribution system are analyzed by focusing on the physical and operational parameters. The data in this study consist of spatial data and quantitative data. Poisson distribution in the spatial point process is used as the leakage point distribution approximation. From the analysis, it is found that the distribution pattern of pipeline leakage location is categorized as inhomogeneous poisson process, with two types of intensity area. Based on the results of parameter estimation, it is found that the high-intensity area is affected by time of repairs, pipe diameter, pipe length ratio, water supply ratio, and pipe deterioration ratio but not affected by the hydraulic energy demand. While the low-intensity area is influenced by pipe diameter, pipe length ratio, water supply ratio, pipe deterioration ratio, and the hydraulic energy demand but not affected by the time of repairs.

1. Introduction
Water supply authority implements a leakage control strategy based on a scheme recommended by Malcolm Farley and Roland Liemberger which focuses on Active Leakage Control (ALC), network asset management, hydraulic pressure management and the speed of repairs [1]. However, it was never discovered which factors has a significant influence related to the characteristics of each service zone or district in the water distribution networks. The leakage management strategy in each service zone or district will be various according to its unique characteristics of the distribution network system [2].

Based on four factors recommended by Malcolm Farley and Roland Liemberger, some researchers propose parameters that affect water loss in water distribution networks. South Korean researcher H.S. Park proposes several parameters including reservoir capacity, number of customers, speed of repairs, water meter replacement ratio, total daily water supply per demand junction, pipe deterioration and pipel
length, water supply scale, facility scale, and number of damaged facilities. J.K. Jang proposes parameters for the number of leakage and customer connections, pipe age, hydraulic pressure, and minimum night flow ratio. J.M Lee proposes the use of parameters for pipe replacement ratio in service zones, water meter replacement ratio and speed of leakage repairs. Chung et.al propose parameters the number of demand junction per pipe length in the service zone, the age of the residential building, and the level of water meter replacement [3].

D. Jang and G. Choi propose pipe diameter, pipe length per demand junction, total water supply per demand junction, pipe deterioration ratio, number of leakage events, hydraulic energy demand ratio. Based on studies conducted by H.S. Park, J.K. Jang, JM Lee, Chung et.al and D. Jang and G. Choi, Dongwoo Jang conducted a qualitative study to develop a parameter classification system that affects water loss, especially physical loss [3].

Based on a qualitative review of previous research, Dongwoo Jang states that real losses are influenced by physical parameters and operational parameters. Parameters that affect physical loss include the type of pipe material, pipe diameter, pipe length, the quantity of water supply and the pipe deterioration ratio. The parameters that affect water loss in the operational process are the number of leakage events, the speed of repairs and the ratio of hydraulic energy demand [4]. Further, Nugroho and Iriawan examined that there was a significant relationship between speed of repairs and pipeline leakage pattern [5].

This study conducted in Malang, East Java, Indonesia. Malang city water distribution network is managed by Malang City Public Water Company. In the network operational and maintenance practices, the water company divides operational and maintenance divisions into several sub-divisions, which is so far the organizational structure only follows the water supply service zone boundaries [6]. While on the other hand, not all zones require the same priority in handling leakage. This will affect the effectiveness of leakage management strategy in the region. Thus, there is a necessity of conducting study that involving spatial parameters in leakage management strategy [7]. Based on this background, this study will analyse the pipeline leakage location pattern using the spatial poisson point process [8]. The use of poisson distribution in the spatial point process is proposed as an approximation to the pipeline leakage pattern to answer problem statements below:

- Is the distribution pattern of pipeline leakage locations categorized as a homogeneous or inhomogeneous poisson process?
- What parameters play a significant role in the distribution of pipeline leakage points in each service zone of water distribution networks?

2. Materials and methods

The research variables that proposed in this study consist of response and covariate variable. Response variable proposed in this study is the number of pipeline leakage event with a certain latitude and longitude geographical coordinate. The data will be displayed visually as a planar point pattern then transformed into a pixel image display [9]. The covariate variable proposed in this study is the operational and physical network variables at each pipeline leakage point, displayed as pixel images or marks point patterns [9]. The water supply 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, based on the number of water supply service zone [8].

| Variable | Scale | Description                  |
|----------|-------|------------------------------|
| $Y$      | Ratio | Leakage event numbers       |
| $X_1$    | Ratio | Leakage time of repairs      |
| $X_2$    | Ratio | Leakage Pipeline diameter    |
| $X_3$    | Ratio | Pipeline length per demand junction |
| $X_4$    | Ratio | Water supply per demand junction |
| $X_5$    | Ratio | Pipeline deterioration ratio  |
| $X_6$    | Ratio | Hydraulic demand ratio       |
### Table 2. Data structure.

| Grid | $Y_1$ | $X_{1.1}$ | $X_{1.2}$ | $X_{1.3}$ | $X_{1.4}$ | $X_{1.5}$ | $X_{1.6}$ |
|------|-------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1    | $Y_2$ | $X_{2.1}$ | $X_{2.2}$ | $X_{2.3}$ | $X_{2.4}$ | $X_{2.5}$ | $X_{2.6}$ |
| 2    |       |           |           |           |           |           |           |
| ...  |       |           |           |           |           |           |           |
| 13   | $Y_{13}$ | $X_{13.1}$ | $X_{13.2}$ | $X_{13.3}$ | $X_{13.4}$ | $X_{13.5}$ | $X_{13.6}$ |

The data analysis process proposed in this study as follows:

- Analyse the characteristics of leakage distribution patterns.
- Conduct homogeneity testing of intensity.
- Identify mixture models.
- Estimate model parameters using the Bayesian method.
- Make conclusions and recommendations from the results of the analysis.

To estimate parameters, the Bayesian algorithm used as follows [10]:

- Form the probability function.
- Determine the prior distribution that depends on the frequentist result.
- Form a posterior distribution based on the prior distribution.
- Form full conditional posterior distribution for estimated parameters.
- Perform an iterative process of estimating parameters using Gibbs sampling, meet irreducible, repetitive, and aperiodic characteristics.
- Determine the best model parameters.

### 3. Results and discussion

Figure 1 (a) shows that visually the distribution from the location of leak points is inhomogeneous. It appears in the southern service area there is a higher density point compared to other service areas. From Figure 1 (b) shows that there is a peak point in the intensity of the distribution of leakage pipeline locations, in the southern service area, precisely in the Wendit Zone and the Buring Zone.

![Fig 1a](a) ![Fig 1b](b)

**Figure 1.** Leakage point distribution in planar point pattern (a) and pixel image (b).

By the aspect of the number of customers, both zones have the largest number of customers, where the Wendit Zone serves 28.16% of the total number of customers. Meanwhile the Buring Zone serves 21.48% of all customer connections. From the network asset aspects, the two zones have the largest pipeline length, where the Wendit Zone has 30.71% of the total length of the distribution network pipeline. While the Buring Zone has 18.87% of the total length of the distribution network pipeline. Figure 1 (b) shows gradation of leakage intensity is uneven distributed by visual observation on pixel images of leakage intensity. This indicates that the distribution of leakage locations is not homogeneous in water service area, which there are two intensity clusters, divided into high leakage intensity and low leakage intensity. Data on leakage locations that have been explored visually are then identified as
inhomogeneous poisson processes through statistical distribution goodness of fit test and homogeneity test of intensity. If the leakage distribution has more than one pattern or bimodal, then the mixture model must be identified. The estimated parameters of the spatial poisson point process model using the Bayesian method will be carried out after both identification completed [8].

Homogeneity test of intensity is conducted to find out whether the pattern of leakage is distributed evenly or unevenly. The results that will be obtained from the homogeneity test of intensity is one of the characteristics of the type of Poisson process, homogeneous poisson processes or inhomogeneous poisson processes [11].

### Table 3. Homogeneity test of intensity.

| Statistics | alpha | p-value |
|------------|-------|---------|
| 46537      | 0.05  | 2.2 x 10^{-16} |

From the homogeneity test results in table 3, found that the chi-square statistical value = 46537 and p-value = 2.2 x 10^{-16} using alpha number 0.05 or 5%. The decision obtained based on the results of the test is intensity of the distribution of leakage locations is inhomogeneous poisson process [9].

Goodness of fit test is conducted to find out that the leakage pattern in the water service area categorized as unimodal or bimodal poisson distribution [11].

### Table 4. Goodness of fit test.

| Statistics | Critical Value | Alpha |
|------------|----------------|-------|
| 141,34     | 2.5018         | 0.05  |

It is found that the leakage point pattern does not have a unimodal poisson distribution so it must go through a mixture poisson model analysis process. The mixture identification results show that in this study a mixture poisson regression model with two components will be used. Mixing Poisson regression modelling is done using Win BUGS software. The structure parameters of the mixture poisson regression model can be formed with the Directed Acyclic Graph (DAG) as shown in Figure 2. The parameter structure of the mixture poisson regression model in Figure 2 which consists of the variable y [i] as the node distributed poisson obtained from log lambda [i]. The form of the lambda equation [i] is as follows:

$$\lambda_{[i]} = (\beta_0 \tau[i] + \beta_1 \tau[i] + \beta_2 \tau[i] + \beta_3 \tau[i] + \beta_4 \tau[i] + \beta_5 \tau[i] + \beta_6 \tau[i] + X_d[i] + X_a[i] + X_b[i] + X_c[i] + X_d[i])$$

(1)

The value of T [i] is used to identify lambda [i], where the value of T [i] changes according to the distribution of P [1: 2] charters which have two categories according to the number of mixture components in the model.
The prior distribution used in this study is the pseudo prior, where prior depends on the frequentist results shown by nodes outside the box in Figure 2. Priors are raised normally with the mean and standard deviation obtained from the estimated parameters and from the frequentist GLMs function. The posterior distribution is used to obtain the estimation of model parameters, where the Markov Chain Monte Carlo (MCMC) method with Gibbs Sampling is used. The results obtained have to meet the nature of irreducible, aperiodic, and recursive which can be seen from the results of history plots, autocorrelation, and kernel density. Based on Table 5 it is found that almost all parameters are significant except for the Beta 1 [2] and Beta 6 [1] parameters. This is because credible intervals do not pass zero, so all parameters except Beta1 [2] and Beta 6 [1] are significant for use in the model. Based on the model, it is known that the highest proportion of leakage points per grid is found in component 1, this

The 5th Annual Applied Science and Engineering Conference (AASEC 2020)  
IOP Conf. Series: Materials Science and Engineering 1098 (2021) 022051 doi:10.1088/1757-899X/1098/2/022051

Figure 2. Directed Acyclic Graph (DAG) mixture poisson regression.

Table 5. Parameter estimation result.

| Parameter | Average | Standard Deviation | 2.5% | Median | 97.5% |
|-----------|---------|--------------------|------|--------|-------|
| P[1]      | 0.7654  | 0.00567            | 0.5894 | 0.7645 | 0.8956 |
| P[2]      | 0.2346  | 0.07843            | 0.1563 | 0.2355 | 0.3457 |
| Beta 0 [1]| 6.333   | 0.09194            | 6.152  | 6.333  | 6.513  |
| Beta 0 [2]| 4.695   | 0.03919            | 4.623  | 4.696  | 4.766  |
| Beta 1 [1]| -1.002  | 0.175              | -1.346 | -1.002 | -0.6593|
| Beta 1 [2]| -0.02485| 0.0407             | -0.105 | -0.02472| 0.0544 |
| Beta 2 [1]| 0.2917  | 0.1456             | 0.00478| 0.2924 | 0.5764 |
| Beta 2 [2]| -0.2863 | 0.04637            | -0.3778| -0.286 | -0.1967|
| Beta 3 [1]| 0.9196  | 0.1372             | 0.6505 | 0.9195 | 1.188  |
| Beta 3 [2]| 0.5916  | 0.1231             | 0.3518 | 0.5915 | 0.8318 |
| Beta 4 [1]| -1.302  | 0.1875             | -1.671 | -1.302 | -0.9366|
| Beta 4 [2]| -0.6442 | 0.1158             | -0.8715| -0.6441| -0.4176|
| Beta 5 [1]| 0.6465  | 0.1473             | 0.3576 | 0.6468 | 0.9349 |
| Beta 5 [2]| 0.3789  | 0.05417            | 0.2739 | 0.3785 | 0.4861 |
| Beta 6 [1]| 0.01481 | 0.03794            | -0.05982| 0.01496| 0.08878|
| Beta 6 [2]| 0.7729  | 0.0648             | 0.6468 | 0.7727 | 0.9008 |
can be seen from the proportion value in component 1 which is 0.7654 or 76.54% of leakage events is in the component 1 model area (high leakage intensity).

4. Conclusion
Based on the analysis that has been done, the following conclusions are obtained:

- From the intensity homogeneity analysis, the results show that the distribution pattern of leakage pipelines in Malang City is categorized as inhomogeneous poisson process.
- Based on the estimation results of the parameters obtained, the results show that in the area of high leakage intensity, the number of leakage events is influenced by repair time, pipe diameter, ratio of pipe length per demand junction, ratio of supply discharge per demand junction, and pipe deterioration ratio. The greatest significance parameter is the ratio of pipe length per demand junction.
- For areas with low leakage intensity, the number of leakage events is influenced by pipe diameter, ratio of pipe length per demand junction, ratio of supply discharge per demand junction, pipeline deterioration ratio, and hydraulic demand ratio. The greatest significance parameter is the hydraulic demand ratio.
- In the area of high leakage intensity, the maintenance of distribution networks should prioritize Active Leakage Control (ALC) and network asset management. With a high ratio of pipe length per demand junction, there will be a tendency for high hydraulic pressure, which will certainly increase the risk of pipeline leakage. Active leakage monitoring in areas that have a high deterioration ratio can reduce the risk of leakage, as well as pipeline replacement that is nearing or exceeding its theoretical age. In addition, network asset management, especially distribution pipelines and customer connection pipelines, is also important to remember, considering that the pipe deterioration ratio is the second parameter that has a significant influence.
- In the area of low leakage intensity, maintenance of the distribution network should prioritize pressure management and increase the quantity of water supply to the service area. Pressure management can be proposed by placement of pressure relief instruments such as PRV (Pressure Reducing Valve) at several points in the service area.

Acknowledgements
We hereby thank you to Civil Engineering Department, Sepuluh November Institute of Technology and Malang City Public Water Company, for supporting the publication of this research.

References
[1] Liemberger R and Farley M 2004 Developing a Non-Revenue Water Reduction Strategy Part 1: Investigating and Assessing Water Losses Proc IWA 4th World Water Congr. Exhib. 1924 Sept. 2004 Marrakech Morocco
[2] Farley M, Wyeth G, Ghazali Z B M, Istandar A and Sigh S 2008 The Manager’s Non-Revenue Water Handbook A Guid. to Underst. Water Losses, Ranhill Util. Bernhad USAID, Malaysia
[3] Jang D, Park H and Choi G 2018 Estimation of leakage ratio using principal component analysis and artificial neural network in water distribution systems Sustain.
[4] 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 Sustain.
[5] Nugroho W and Iriawan N 2019 Effect of the leakage location pattern on the speed of recovery in water supply networks J. Phys. Conf. Ser. 1402 022023
[6] Alegre H, Baptista J M, Cabrera E, Cubillo F, Duarte P, Himel W, Merkel W and Parena R 2016 Performance Indicators for Water Supply Services: Third Edition Water Intell. Online
[7] Farley M and Liemberger R 2005 Developing a non-revenue water reduction strategy: Planning and implementing the strategy Water Science and Technology: Water Supply
[8] Baddeley A and Turner R 2005 spatstat: An R package for analyzing spatial point patterns J. Stat.
[9] Evangelista P F and Beskow D 2019 Geospatial point density R J.
[10] Stern H S 2015 Bayesian Statistics International Encyclopedia of the Social & Behavioral Sciences: Second Edition
[11] Blei D M, Kucukelbir A and McAuliffe J D 2017 Variational Inference: A Review for Statisticians J. Am. Stat. Assoc.