Indonesian Journal of Electrical Engineering and Computer Science
Vol. 35, No. 3, September 2024, pp. 1618~1626
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v35.i3.pp1618-1626

Urban traffic congestion and its association with gas station density: insights from Google Maps data

Rafif Hasabi[1], Robert Kurniawan[1], Sugiarto[2], Ribut Nurul Tri Wahyuni[2], Erna Nurmawati[1]
[1]Department of Statistical Computing, Polytechnic Statistics STIS, Jakarta, Indonesia
[2]Department of Statistics, Polytechnic Statistics STIS, Jakarta, Indonesia

Article Info

Article history:

Received Feb 23, 2024
Revised Apr 5, 2024
Accepted May 6, 2024

Keywords:

Air pollution
Gas station density
Google Maps
Map digitization
Traffic congestion

ABSTRACT

Analyzing air pollution caused by traffic conditions requires appropriate indicators. Currently, air pollution indicators are approximated by the number of vehicles and gas station density. However, this approach cannot provide information at a smaller level. This study aims to identify traffic congestion distribution from Google Maps data as an alternative air pollution indicator at smaller level using map digitization method. In addition, this study examines its relationship with the existing indicator called gas station density. The results show that the digitization method can map the traffic congestion distribution where most areas in West, North, and Central Jakarta are classified as high traffic. In addition, this study found that there is a strong and significant relationship of 0.58277 between traffic congestion distribution and gas station density. Thus, traffic congestion distribution and gas station density data from Google Maps can be used as an indicator of traffic-related air pollution, especially land transportation. Furthermore, this research is expected to serve as a basis for the government in determining mitigation strategies related to traffic congestion and the resulting emissions.

Corresponding Author:

Robert Kurniawan
Department of Statistical Computing, Polytechnic Statistics STIS
Jatinegara, 13330, Jakarta Timur, DKI Jakarta, Indonesia
Email: robertk@stis.ac.id

1. INTRODUCTION

Air pollution is one of the environmental problems in big cities [1], which has a significant impact on human health [2]. The main contributor to air pollution is motor vehicle emissions, which are exacerbated by traffic conditions [3]. In a 2018 report, the World Health Organization (WHO) estimated that 93 percent of the world's children are exposed to toxic air every day, threatening their health [4]. One of these situations is in Jakarta, which has recently become the capital city with the highest air pollution levels in the world [5].

Air pollution indicators are needed to describe pollution conditions in an area, especially those caused by traffic conditions. Air pollution due to traffic conditions is often associated with congestion [6]. So, the indicator that is often used is the number of motorized vehicles [7]. However, official data from the Indonesian National Police regarding the number of vehicles as a proxy for congestion is also not available at smaller levels such as sub-districts. In fact, data on traffic is needed for urban planning analysis. In areas with urban characteristics such as Jakarta, a more detailed analysis of the sub-district level is required.

Another indicator that can be used as an indicator of air pollution and is still related to congestion is gas station density [8]. This is based on the fact that congestion also increases fuel consumption which contributes to increased air pollution [9]. Thus, the increase in fuel consumption results in a high demand for fuel, especially in areas that are congested or with a large number of vehicles [10]. This has an impact on the number of gas stations in the region. The clustering of gas stations within an area has negative implications for traffic flow [11]. However, the use of these indicators is not directly related to congestion. Thus, its utilization as an indicator of air pollution is less appropriate.

An alternative air pollution indicator that is directly related to traffic conditions is congestion data itself. Currently, congestion levels are measured by indexes such as the TomTom traffic index [12]. This approach only provides general conditions in a region with wide coverage. In addition, all API features of the TomTom traffic index are not freely accessible [13]. On the other hand, data from Google Maps has a lot of information such as various points of interest [14], road maps [15], geographic coordinates [16], and congestion [17]. However, research related to the utilization of Google Maps data for sub-district level air pollution indicators is very rare. In addition, there is no research that uses a map digitization method approach to obtain traffic congestion distribution data. Thus, this research takes the opportunity to use the Google Maps approach as an air pollution indicator that is directly related to traffic conditions in an area using this new approach. Furthermore, the fact that gas station density can be an indicator of traffic-related air pollution is interesting to see how it relates to traffic data obtained from Google Maps. This is important because air pollution is caused by many factors such as traffic [18], industry [19], and households [20].

So, this research aims to identify the distribution of traffic congestion at the sub-district level using the map digitization method in Jakarta as a new indicator of air pollution due to traffic conditions and see its association with existing indicators, namely gas station density. This research is expected to contribute to researchers or the government related to alternative indicators that can describe air pollution due to traffic conditions. With this, the government can develop new policies to limit the use of fossil fuels and switch to environmentally friendly electrical energy. In addition, a good identification of congestion can provide recommendations for more effective urban planning, traffic management, and fuel use in the future, as well as assist relevant stakeholders in decision-making [21]. This research has the potential to support the acceleration of sustainable development (SDGs) goals 7, 9, and 11 [22].

2. METHOD

2.1. Study area

This study takes DKI Jakarta Province, Indonesia, as the research location as shown in Figure 1. A total of 42 sub-districts in the administrative city area are used as the unit of analysis in this study [23]. As the center of government and economy, Jakarta faces many challenges such as traffic congestion and air pollution that can affect people's quality of life [24]. Therefore, Jakarta as a research location is suitable for the purpose of this study.

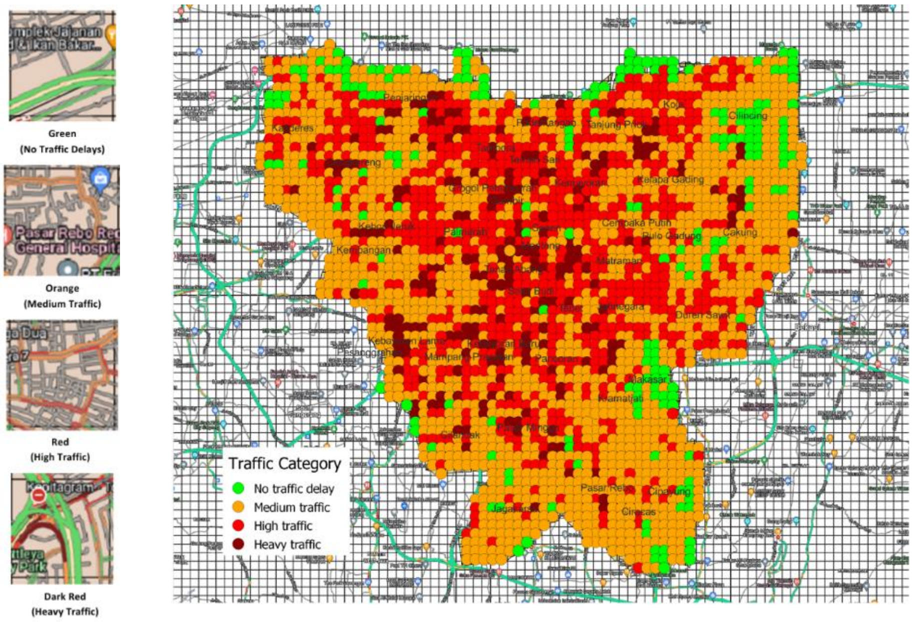
Figure 1. DKI Jakarta Province of Indonesia as study area

2.2. Traffic congestion data

Traffic congestion data was obtained by digitizing Google Maps data through the traffic viewer plugin in the quantum GIS (QGIS) application [25] as shown in Figure 2. Traffic digitization is done to get alternative traffic data in the past period [26]. The colors on the traffic map include green (no traffic delays), orange (medium traffic), red (high traffic), and dark red (heavy traffic) [27]. The digitization stages carried out in this study are as follows:

- Import base map from the Google traffic viewer plugin a specific date. The time taken is 11:00 WIB to avoid data noise that is too dense or too quiet.
- Create a 500m grid so that the entire study area is covered.
- Set the scale on the map to 1:25000 [28].
- Digitizing each grid to determine the traffic category based on the highest traffic as shown in Figure 2(a).
- The process was carried out with a 1-week sample every month in the first week. This is based on previous research that the traffic pattern each week is generally the same [29]–[31]. So that the distribution of traffic congestion for each traffic category is obtained as shown in Figure 2(b).
- The previous four categories can be simplified into two categories, namely no traffic delays (sum of no traffic delays and medium traffic) and high traffic (sum of high traffic and heavy traffic). In addition, the proportion of congestion (CP) can be obtained through (1).

$$CP = \frac{High\ traffic}{Total\ traffic} \quad (1)$$



(a) (b)

Figure 2. Traffic digitization (a) example grid and its traffic categories, and (b) traffic digitization process

2.3. Gas station density data

Data on the number of gas stations was collected using the web scraping point of interest (POI) technique in the form of gas stations in the study area through the Google Maps page. POIs categorized as gas stations include public filling stations (SPBU) and mini SPBU. To obtain the density, the number of gas stations for each sub-district was summed and divided by its area [32]. Furthermore, the density data obtained was adjusted into three categories (low, moderate, high) using quantile cut with QGIS. Quantile cut divides the data into equal-sized data [33]. Gas station density is used as an indicator of air pollution caused by urban traffic conditions [8], [32], [34]. This is also in line with the relationship between gas station distribution and traffic volume [11], as congestion can increase fuel consumption [9], [35].

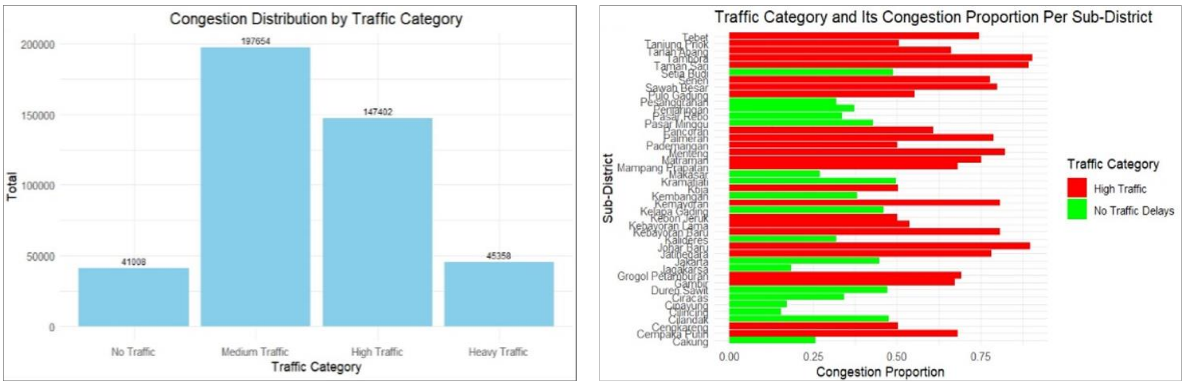
2.4. Chi-square association

The Chi-square association test was conducted to see the relationship between traffic congestion distribution and gas station density. Chi-square association test is one of the comparative tests to see the relationship between two variables, where the variable scale used is nominal [36]. The size of the level of association or relationship can be seen through the contingency coefficient (C_c) [37]. Cohen suggested the following guidelines for interpreting C_c; weak > 0.1, moderate > 0.3, and strong > 0.5 [38].

3. RESULTS AND DISCUSSION
3.1. Traffic congestion distribution

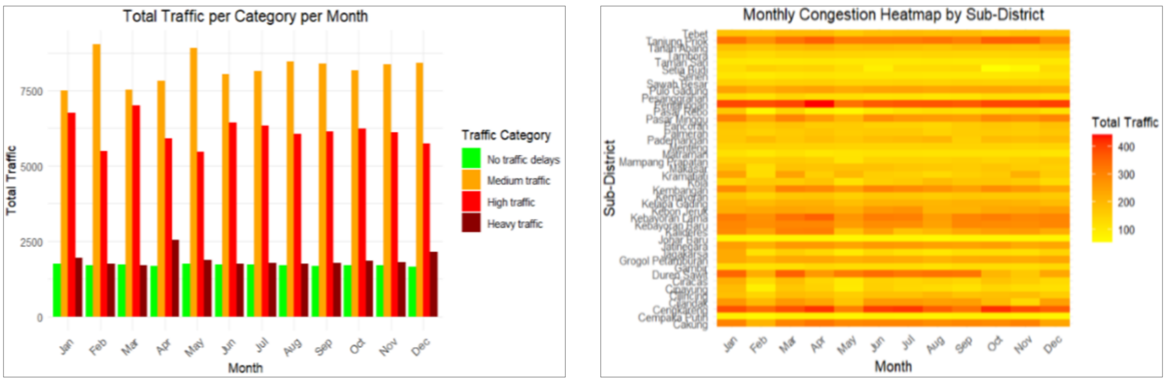
The results of digitizing traffic congestion distribution every day are 2,568 points. The distribution of traffic congestion data is then aggregated on an annual (Figure 3) and monthly (Figure 4) basis using summation. In addition, the traffic congestion category is simplified to simplify the analysis. The medium traffic category on an annual basis has the largest number of points, at 197,654 points as shown in Figure 3(a). This shows the general condition of traffic in Jakarta in 2023 which is not congested. The traffic category with the next highest number of points is high traffic with 147,402 points. Then, the number of points for the no traffic delays and heavy traffic categories are 41,008 and 45,358 points. Furthermore, the distribution of the proportion of congestion (CP) in Figure 3(b) is highest in Tambora District, West Jakarta at 0.90476, while the lowest CP value is in Cilincing District, North Jakarta at 0.15282. There are 25 sub-districts in the administrative city of Jakarta that are classified as no traffic delays where the CP value is less than 0.5. Then, 17 other sub-districts are classified as high traffic where the CP value is more than 0.5.

On a monthly basis in Figure 4(a), there are fluctuations in the number of traffic category points, especially in the medium traffic and high traffic categories. The most medium traffic category occurred in February with 9,038 points, while the highest traffic category occurred in March with 7,001 points. From these results, the number tends to be constant for the no traffic delays and heavy traffic categories. Figure 4(b) shows the total monthly high and heavy traffic categories for each sub-district. The redder the color of the heatmap, the more congested the sub-district is, Figure 5 shows the data distribution of urban traffic congestion and gas station density. The distribution of urban traffic congestion based on Figure 5(a) shows that sub-districts that fall under no traffic delays are mostly in South Jakarta, while those that fall under high traffic are mostly in West, North, and Central Jakarta.



(a) (b)

Figure 3. Traffic congestion data (annual) (a) congestion distribution by traffic category and (b) traffic category and its congestion proportion per sub-district

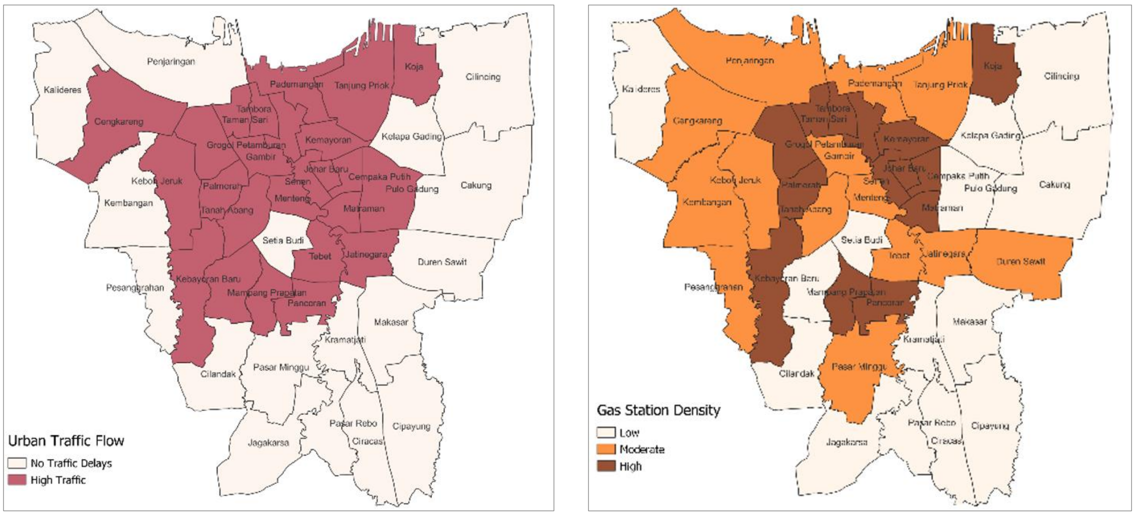


(a) (b)

Figure 4. Traffic congestion data (monthly) (a) trend of total traffic per category and (b) traffic congestion heatmap per sub-district

3.2. Gas station density

A total of 1,044 gas station data in Jakarta was collected using Google Maps website scraping technique. The results show that South Jakarta has the greatest number of gas stations with 269 gas stations. One of the sub-districts in South Jakarta, Kebayoran Lama, also has the highest number of gas stations compared to other sub-districts. Gas station density data from the calculation and quantile cut results are mapped in Figure 5(b). In general, the density of moderate and high gas stations is located in West, North, and Central Jakarta. The low-density category is mostly located in South and East Jakarta.



(a) (b)

Figure 5. Distribution of (a) urban traffic congestion and (b) gas station density

3.3. Association between traffic congestion distribution and gas station density

The relationship between traffic congestion distribution and gas station density was tested using Chi-square association. Table 1 shows that districts with high gas station density tend to have high traffic congestion distribution. This is in line with the results of the Chi-square association test in Table 2 which shows a significant relationship (p-value = 0.00) at the 0.05 level of significance. The strength of the relationship can be seen through the value of $C_c = 0.58227$ which shows a strong relationship between traffic congestion distribution and gas station density.

Table 1. Cross tabulation of traffic category and gas station density category

Traffic category	Gas station density category		
	Low	Moderate	High
No traffic delays	12	5	0
High traffic	2	9	14

Table 2. Chi-square test result

Chi-square test	Value
Pearson chi-square	21.54353
p-value	0.00000
Contingency coefficient	0.58227

3.4. Discussion

This research provides a new approach to collecting sub-district level traffic data through Google Maps data using the map digitization method as an alternative indicator of traffic-related air pollution. This research addresses the lack of alternative indicators for traffic-related air pollution that have been explored in previous studies such as number of motorized vehicles [7], gas station density [8], [34], and traffic index e.g., TomTom traffic index [12]. This method approach will be suitable to be implemented in areas with urban characteristics such as Jakarta. This study found that the congested areas are mostly in West, North, and Central Jakarta. The results obtained are quite similar to the data on congestion hotspots that have been released by the transportation department [39] and information on congestion hotspots on online news sites. Furthermore, the congested areas are also in line with population growth and density in Jakarta. Central Jakarta has the highest population density compared to other administrative cities. This relationship has been explored in previous research. Where rapid population growth and density in urban areas go hand in hand with an increase in the number of vehicles [40]. This shows that the method used is sufficient to describe the general condition of traffic congestion in Jakarta even though the sample used is only one week each month. The use of a one-week sample complements the findings of previous studies which state that the traffic pattern each week is generally the same [29]–[31], [41]. However, accuracy and consistency depend on the data provided by Google Maps [42]. In addition, the digitization process is also inseparable from human error [43].

This study compares the proposed air pollution indicator with the existing gas station density indicator [8], [32], [34]. Through the Chi-square association test and C_c value, this research found that there is a strong and significant relationship between traffic congestion distribution and gas station density. This result indicates the alignment between traffic data and gas station density. Thus, the proposed traffic data has the potential to be an alternative indicator of air pollution caused by traffic conditions. Traffic data is more directly related to traffic conditions because it describes areas with high traffic or no traffic delays. When compared to gas station density data, it is not directly related to traffic conditions. However, a previous study found that clustering gas stations within an area has negative implications for traffic flow [11]. The existence of a strong significant relationship between traffic data and density in this study provides insight that both variables can be indicators of traffic-related air pollution because they are related to congestion. Traffic congestion that occurs in almost all areas of Jakarta is a focus in air pollution monitoring because it contributes to increased pollutants [18] and energy consumption [35].

This research explores alternative indicators of traffic-related air pollution using a map digitization method that still has shortcomings. Futher in-depth studies may be needed to confirm the distribution of congestion in Jakarta. By using indicators from traffic data and gas station density data, air pollution monitoring can be done earlier in an area. However, this indicator only looks at air pollution in terms of traffic conditions, especially land transportation. Where it is approached with traffic congestion data and gas station density data. In fact, air pollution does not only come from transportation. Further exploration of indicators that can cover air pollution more thoroughly has the potential to be done. In addition, as it relates to air pollution, information on the relationship of traffic congestion and gas station density data with several pollutants produced by motor vehicles can be used to confirm the accuracy of the proposed indicators related to air pollution.

From the results obtained, this study recommends the use of traffic congestion distribution and gas station density data as indicators of air pollution, especially those caused by land transportation traffic conditions. Moreover, when official statistics data such as the number of vehicles is difficult to obtain. So that air pollution monitoring can be done earlier both through congestion data and gas station density data in an area. This can make it easier for the government to develop strategies related to urban planning, traffic management, and the use of environmentally friendly fuels. By knowing the distribution of traffic congestion and gas stations, mitigation efforts will be more effective and targeted. An effort to reduce air pollution from motor vehicle emissions is traffic management itself. With good congestion management efforts, it can contribute to the reduction of combustion air pollution and reduce energy consumption [44]. This can affect the development of refueling station infrastructure that is more evenly distributed and not clustered in an area.

4. CONCLUSION

This research provides an initial basis for utilizing Google Maps data as an alternative indicator of air pollution caused by traffic conditions. Through the map digitization method, traffic congestion distribution data can be obtained at a smaller presentation level, namely sub-districts. The traffic-related air pollution indicator proposed in this research is approached with information from Google Maps that shows congestion. The more congested an area is, the worse the air quality in that area will be. This research found that there is a strong and significant relationship between traffic congestion distribution and gas station density. The alignment shows that traffic congestion data can be used as an indicator of air pollution in an area in addition to using gas station density data or the number of vehicles. The results of this study are expected to support the government in formulating more targeted strategies and policies to reduce traffic congestion and the resulting emissions. Further research can use more effective methods in obtaining traffic congestion data from Google Maps and use other data that can be an indicator of air pollution more thoroughly.

ACKNOWLEDGEMENTS

The authors would like to thank Politeknik Statistika STIS for providing the essential support and resources that greatly assisted this research. Their contribution has been a cornerstone in the culmination of this study. Thank you to reviewers for giving very helpful feedback that made it possible for this paper to be published.

REFERENCES

[1] A. Piracha and M. T. Chaudhary, "Urban air pollution, urban heat island and human health: a review of the literature," *Sustainability (Switzerland)*, vol. 14, no. 15, 2022, doi: 10.3390/su14159234.
[2] I. Manisalidis, E. Stavropoulou, A. Stavropoulos, and E. Bezirtzoglou, "Environmental and health impacts of air pollution: a review," *Frontiers in Public Health*, vol. 8, Feb. 2020, doi: 10.3389/fpubh.2020.00014.
[3] J. C. Adiatma, "A transition towards low carbon transportation in Indonesia: a technological perspective," *Study Report IESR*, vol. Oct 2020, p. 71, 2020, [Online]. Available: https://iesr.or.id/en/pustaka/a-transition-towards-low-carbon-transport-in-indonesia-a-technological-perspective.
[4] WHO, "Air pollution and child health," 2018. [Online]. Available: http://pediatrics.aappublications.org/content/113/Supplement_3/1037.full.html.
[5] Y. Sunitiyoso, P. F. Belgiawan, M. Rizki, and V. Hasyimi, "Public acceptance and the environmental impact of electric bus services," *Transp. Res. Part D Transp. Environ.*, vol. 109, no. June, p. 103358, 2022, doi: 10.1016/j.trd.2022.103358.
[6] C. K. Gately, L. R. Hutyra, S. Peterson, and I. S. Wing, "Urban emissions hotspots: quantifying vehicle congestion and air pollution using mobile phone GPS data," *Environmental Pollution*, vol. 229, pp. 496–504, Oct. 2017, doi: 10.1016/j.envpol.2017.05.091.
[7] S. Chen, S. Bekhor, Yuval, and D. M. Broday, "Aggregated GPS tracking of vehicles and its use as a proxy of traffic-related air pollution emissions," *Atmospheric Environment*, vol. 142, pp. 351–359, Oct. 2016, doi: 10.1016/j.atmosenv.2016.08.015.
[8] H. F. Chiu *et al.*, "Traffic air pollution and risk of death from gastric cancer in Taiwan: petrol station density as an indicator of air pollutant exposure," *Journal of Toxicology and Environmental Health - Part A: Current Issues*, vol. 74, no. 18, pp. 1215–1224, 2011, doi: 10.1080/15287394.2011.590100.
[9] R. Bista and S. Paneru, "Does road traffic congestion increase fuel consumption of households in Kathmandu City?," *Journal of Economic Impact*, vol. 3, no. 2, pp. 67–79, 2021, doi: 10.52223/jei3022102.
[10] A. Konieczna, K. Roman, and W. Rzodkiewicz, "Fuel consumption, emissions of air pollutants and opportunities for reducing CO_2 emissions from linear sources in the model rural municipality," *Energies*, vol. 16, no. 14, p. 5553, Jul. 2023, doi: 10.3390/en16145553.
[11] E. O. Akinsulire and S. O. Fadare, "An assessment on the locational pattern of petrol filling stations along lasu-isheri road corridor," *American International Journal of Multidisciplinary Scientific Research*, vol. 6, no. 2, pp. 6–29, 2020, doi: 10.46281/aijmsr.v6i2.705.
[12] T. Moyo, S. Mbatha, O. O. Aderibigbe, T. Gumbo, and I. Musonda, "Assessing spatial variations of traffic congestion using traffic index data in a developing city: lessons from Johannesburg, South Africa," *Sustainability (Switzerland)*, vol. 14, no. 14, 2022, doi: 10.3390/su14148809.
[13] T. P. Bauer, J. Edinger, and C. Becker, "A qualitative and quantitative analysis of real time traffic information providers," *2019 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2019*, pp. 113–118, Mar. 2019, doi: 10.1109/PERCOMW.2019.8730795.
[14] B. Firmansyah and W. S. Wahyuni, "Analysis of using Google Maps data to measure the presence or accessibility of urban facilities for BPS - statistics Indonesia classification of urban and rural villages," *Journal of Information Systems and Informatics*, vol. 4, no. 4, pp. 823–832, 2022, doi: 10.51519/journalisi.v4i4.337.
[15] B. G. Grepon, J. Margallo, J. Maserin, and R. A.-D. Dompol, "RUI: a web-based road updates information system using Google Maps API," *International Journal of Computing Sciences Research*, vol. 7, pp. 2253–2271, 2023, doi: 10.25147/ijcsr.2017.001.1.158.
[16] N. I. Akanda, M. A. Hossain, M. M. I. Fahad, M. N. Rahman, and Khairunnaher, "Cost-effective and user-friendly vehicle tracking system using GPS and GSM technology based on IoT," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 28, no. 3, pp. 1826–1833, 2022, doi: 10.11591/ijeecs.v28.i3.pp1826-1833.
[17] A. Muñoz-Villamizar, E. L. Solano-Charris, M. AzadDisfany, and L. Reyes-Rubiano, "Study of urban-traffic congestion based on google maps API: The case of Boston," *IFAC-PapersOnLine*, vol. 54, no. 1, pp. 211–216, 2021, doi: 10.1016/j.ifacol.2021.08.079.
[18] G. A. Nanditho and L. Yola, "Urban development and traffic congestion: Jakarta study during the pandemic," *Lecture Notes in Civil Engineering*, vol. 243, pp. 135–142, 2022, doi: 10.1007/978-3-030-99979-7_16.
[19] R. Kurniawan *et al.*, "Impacts of industrial production and air quality by remote sensing on nitrogen dioxide concentration and related effects: An econometric approach," *Environmental Pollution*, vol. 334, p. 122212, Oct. 2023, doi: 10.1016/j.envpol.2023.122212.
[20] F. Ahmed *et al.*, "Impact of household air pollution on human health: source identification and systematic management approach," *SN Applied Sciences*, vol. 1, no. 5, pp. 1–19, May 2019, doi: 10.1007/s42452-019-0405-8.
[21] H. Shabanpour, A. Fathi, S. Yousefi, and R. F. Saen, "Ranking sustainable suppliers using congestion approach of data envelopment analysis," *Journal of Cleaner Production*, vol. 240, 2019, doi: 10.1016/j.jclepro.2019.118190.
[22] A. Sharifi, Z. Allam, S. E. Bibri, and A. R. Khavarian-Garmsir, "Smart cities and sustainable development goals (SDGs): a systematic literature review of co-benefits and trade-offs," *Cities*, vol. 146, no. October 2023, p. 104659, 2024, doi: 10.1016/j.cities.2023.104659.
[23] W. Alrasyid, D. Lestari, F. Novkaniza, A. Haqqi, and S. Devila, "Modeling the number of acute hepatitis sufferers in DKI Jakarta using negative binomial regression," vol. 03, pp. 4–5, 2015.
[24] S. I. Syalianda and R. D. Kusumastuti, "Implementation of smart city concept: a case of Jakarta Smart City, Indonesia," *IOP Conference Series: Earth and Environmental Science*, vol. 716, no. 1, 2021, doi: 10.1088/1755-1315/716/1/012128.
[25] A. Haddid and A. Larbi, "Traffic Viewer — QGIS Python Plugins Repository," 2021. https://plugins.qgis.org/plugins/Traffic-Viewer/#plugin-details.

[26] E. Piškinaitė and D. Veteikis, "The results of digitizing historical maps: comparison of lithuanian land-use structure in the 19th and 21st Centuries," *Land*, vol. 12, no. 5, 2023, doi: 10.3390/land12050946.
[27] Google Maps Help, "Use layers to find places, traffic, terrain, biking & transit," 2023. https://support.google.com/maps/answer/3092439?hl=en&co=GENIE.Platform%3DDesktop#zippy=%2Ctraffic.
[28] S. Hamad, "Updating topographic Maps at scale 1:250000 for libyan territory using quantum GIS (QGIS) and open geospatial data: libya topo-project," *Journal of Geographical Studies*, vol. 4, no. 1, pp. 22–34, 2020, doi: 10.21523/gcj5.20040103.
[29] F. G. Habtemichael and M. Cetin, "Methodology for quantifying incident-induced delays," no. June, pp. 1–18, 2015.
[30] G. Lonati, M. Giugliano, and S. Cernuschi, "The role of traffic emissions from weekends' and weekdays' fine PM data in Milan," *Atmospheric Environment*, vol. 40, no. 31, pp. 5998–6011, 2006, doi: 10.1016/j.atmosenv.2005.12.033.
[31] S. Yang, J. Wu, G. Qi, and K. Tian, "Analysis of traffic state variation patterns for urban road network based on spectral clustering," *Advances in Mechanical Engineering*, vol. 9, no. 9, pp. 1–11, 2017, doi: 10.1177/1687814017723790.
[32] I. M. M. Terrés, M. D. Miñarro, E. G. Ferradas, A. B. Caracena, and J. B. Rico, "Assessing the impact of petrol stations on their immediate surroundings," *Journal of Environmental Management*, vol. 91, no. 12, pp. 2754–2762, 2010, doi: 10.1016/j.jenvman.2010.08.009.
[33] B. Swiderski, S. Osowski, M. Kruk, and W. Barhoumi, "Aggregation of classifiers ensemble using local discriminatory power and quantiles," *Expert Systems with Applications*, vol. 46, pp. 316–323, 2016, doi: 10.1016/j.eswa.2015.10.038.
[34] H. H. Weng, S. S. Tsai, H. F. Chiu, T. N. Wu, and C. Y. Yang, "Childhood leukemia and traffic air pollution in Taiwan: petrol station density as an indicator," *Journal of Toxicology and Environmental Health - Part A: Current Issues*, vol. 72, no. 2, pp. 83–87, 2009, doi: 10.1080/15287390802477338.
[35] A. Al Tamseh, E. Al Karabsheh, and S. Al Kasassbeh, "Traffic demands policies effect on congestion, delay and fuel consumption," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 5, pp. 521–527, 2020, doi: 10.35940/ijitee.e2487.039520.
[36] W. Miller, "Non-parametric statistics," *Statistics and Measurement Concepts with OpenStat*, pp. 159–166, 2013, doi: 10.1007/978-1-4614-5743-5_6.
[37] N. Salkind, *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, 2012.
[38] J. Cohen, "Statistical Power Analysis," *Current Directions in Psychological Science*, vol. 1, no. 3, pp. 98–101, 1992, doi: 10.1111/1467-8721.ep10768783.
[39] Satlantas DKI Jakarta, "Data on Congestion Hotspots in DKI Jakarta (In Indonesia)," 2023. https://satudata.jakarta.go.id/open-data/detail/data-titik-rawan-kemacetan-di-dki-jakarta (accessed Feb. 18, 2024).
[40] R. Ramanathan, "Link between population and number of vehicles. Evidence from Indian cities," *Cities*, vol. 17, no. 4, pp. 263–269, 2000, doi: 10.1016/S0264-2751(00)00022-6.
[41] T. D. Wemegah and S. Zhu, "Big data challenges in transportation: a case study of traffic volume count from massive radio frequency identification(RFID) data," *Conference Proceedings - 2017 International Conference on the Frontiers and Advances in Data Science, FADS 2017*, vol. 2018-January, pp. 58–63, 2017, doi: 10.1109/FADS.2017.8253194.
[42] M. Strong, "Google Maps," *Geography Today: An Encyclopedia of Concepts, Issues, and Technology*, vol. 178, no. 8, pp. 281–282, 2019, doi: 10.5120/ijca2019918791.
[43] Z. Xie, Q. Jiang, H. Wu, and B. Fu, "Uncertainty analysis of map scanning digitization and data quality control," *2010 18th International Conference on Geoinformatics, Geoinformatics 2010*, 2010, doi: 10.1109/GEOINFORMATICS.2010.5567530.
[44] F. Fachinger, F. Drewnick, and S. Borrmann, "How villages contribute to their local air quality – the influence of traffic- and biomass combustion-related emissions assessed by mobile mappings of PM and its components," *Atmospheric Environment*, vol. 263, no. July, p. 118648, 2021, doi: 10.1016/j.atmosenv.2021.118648.

BIOGRAPHIES OF AUTHORS

Rafif Hasabi was born in Blora, Indonesia, in 2002. He is currently pursuing his final year of the statistical computing study program at STIS Polytechnic of Statistics in Jakarta. His research interests include data analysis, big data, statistics, and official statistics. He can be contacted at email: rafifhasabi@gmail.com.

Robert Kurniawan is a researcher at the Statistics Polytechnic STIS Jakarta who focuses on social science, disaster management and the environment, and big data. He is currently pursuing a doctorate in population and environmental education at the State University of Jakarta. He is actively conducting research and writing reference books with ISBNs such as easy understanding of nonparametric statistics in the health sector and regression analysis using R. He hopes that big data can provide new insights into waste management. He can be contacted at email: robertk@stis.ac.id.

Sugiarto is a Lecturer at Politeknik Statistika STIS Jakarta, and focusing on social sciences, economics, official statistics and Management. Currently actively doing research and writing books. He hopes that this paper can provide new insights regarding public health. He can be contacted at email: soegie@stis.ac.id.

Ribut Nurul Tri Wahyuni . is an assistant professor at the Polytechnic of Statistics STIS. She received the S.ST. degree in statistics from the Polytechnic of Statistics STIS, as well as the M.S.E. and doctoral degrees in economics from Universitas Indonesia, respectively. His research interests include labour, urban, public, and regional economics. She can be contacted at email: rnurult@stis.ac.id.

Erna Nurmawati obtained an undergraduate degree, S.ST., at Institute of Statistics Indonesia majoring in computational statistics. She continued the study at Institute of Technology Bandung Indonesia and received a postgraduate/Master degree, M.T, which focused on Information Technology. Her research interest is in statistics, database management, machine learning and Information technology. Currently, she works as lecture at Institute of Statistics Jakarta, Indonesia. She can be reached at email: erna.nurmawati@stis.ac.id.