KnE Life Sciences
ICMScE
International Conference On Mathematics And Science Education
Volume 2024

Research Article

COVID-19 Death Risk in Surabaya: Modeling by Spatial Point Process

Dora Isnaini Putri, Vanda Fitriyanah, Achmad Choiruddin*, and Jerry Dwi Trijoyo Purnomo

Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Indonesia

ORCID
Dora Isnaini Putri: https://orcid.org/0009-0000-1497-8200
Vanda Fitriyanah: https://orcid.org/0009-0004-4018-2491
Achmad Choiruddin: https://orcid.org/0000-0003-2568-2274
Jerry Dwi Trijoyo Purnomo: https://orcid.org/0000-0001-5150-5066

Corresponding Author: Achmad Choiruddin; email: choiruddin@its.ac.id

Published: 27 March 2024

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the ICMScE Conference Committee.

Abstract.
The total death rate or Case Fatality Rate (CFR) due to COVID-19 in Surabaya is high, that is almost twice of the global CFR (1.4%). Utilization of high-resolution data has the potential to explore COVID-19 cases, not only recording cases at the district or city level but also at the patient's domicile level so that they can provide more detailed spatial information. Meanwhile, research exploring the risk of death from COVID-19, especially in Surabaya using spatial point process model, has not yet been carried out. In this study, an analysis of the risk of death from COVID-19 in Surabaya will be carried out using the inhomogeneous Poisson point process model with covariates or external factors used including the density of the COVID-19 referral hospital location and the proportion of confirmed COVID-19 population aged > 60 years per districts. Our model shows that referral hospitals (exp( ) = 1.03295) and places of worship (exp( ) = 1.03835) have a significant effect on death risk from COVID-19. So, there is a need for special handling for areas that have a population with a vulnerable age (> 60 years) where at this age the human immune system will decrease.

Keywords: COVID-19, heath risk, spatial point process, surabaya

1. INTRODUCTION

The COVID-19 pandemic is a global pandemic that affects most countries in the world. WHO declared COVID-19 as a pandemic and since March 16th, 2020 this pandemic has begun to spread in East Java Province [1]. Surabaya became the epicenter of the spread of COVID-19 from 38 regencies/cities in East Java with the highest number of positive COVID-19 patients, namely 70729 cases and 2564 deaths [2]. The government has made several efforts to reduce the rate of spread and death of COVID-19, such as health protocols, large-scale social restrictions (PSBB), and vaccination programs. Despite these efforts, there are still many cases of death reported in the province of East Java, especially in Surabaya. The total death rate or Case Fatality Rate (CFR) for Surabaya is quite high, namely 2.57% above the Global CFR (1.4%). Research on the risk of COVID-19 with spatial analysis has been done before [3], especially in Surabaya using high-resolution data has been carried out previously but only focused on the risk of spreading [4–6]. Meanwhile, research exploring the risk of death from COVID-19, especially in Surabaya in a spatial model, has not yet been carried out.

Several factors are suspected to be the cause of the high risk of death in COVID-19 cases, including health facilities and demographics. Health facilities include the density of COVID-19 referral hospital locations in each sub-district. The high number of active cases and deaths or the Case Fatality Rate (CFR) causes various impacts, such as increasing the occupancy of hospital beds or the Bed Occupancy Rate. If this condition is not immediately addressed, it is certain that the number of deaths will continue to increase. Demographics is also one of the factors causing the death of COVID-19 cases in Surabaya. This demographic includes the proportion of the population with confirmed COVID-19 aged >60 years. Elderly or commonly referred to as the elderly are very vulnerable to being exposed to COVID-19 because the immune system has decreased and has a higher risk of severity and death due to complications arising from COVID-19 [7]. In Indonesia, the elderly are aged 60 years and over. This is confirmed in Law Number 13 of 1998 concerning the welfare of the elderly in Chapter 1 Article 1 Paragraph 2 [8].

Previous research on cases of death due to COVID-19 using the Geographically Weighted Random Forest method to estimate the nonlinear relationship between COVID-19 mortality and 47 risk factors including environmental and socioeconomic factors [9]. Then, previous research using spatio-temporal analysis on the risk of spreading COVID-19 [4, 10, 11]. Based on this background, this study will conduct an analysis of the risk of death from COVID-19 in Surabaya using the inhomogeneous Poisson point process model associated with several covariates or external factors such as health facilities including the density of the COVID-19 referral hospital location and demographics including the proportion population with confirmed COVID-19 age > 60 years.

2. RESEARCH METHOD

2.1. Dataset

This study uses several data, namely the addresses of patients confirmed to have died from COVID-19, the proportion of patients with confirmed COVID-19 age >60 years, and the number of referral hospitals for treating COVID-19 patients in the city of Surabaya.

This data is secondary data obtained from the Surabaya City COVID-19 Task Force and the Surabaya City Health Office, from the beginning of the COVID-19 case in May 2020 to April 4, 2021. The experimental unit is the location of the address point for the death of COVID-19 patients in Surabaya.

2.2. Inhomogeneous Poisson Point Process

The spatial point process is a random mechanism that generates a point pattern, which is a collection of data that shows the spatial location of an object or event being observed [12]. The realization form of the spatial point process is a spatial point pattern **x**=$\{x_1,...,x_n\}$ from n≥ 0 points that are in an observation window **W** [13]. x_i is the location of the research object represented by Cartesian coordinates, namely **u**= (longitude, latitude), where **u** $\in W, W \subseteq R^2$.

Inhomogeneous Poisson process is a Poisson process **X** in $\mathbb{R}^2$ with intensity function λ(**u**) which is not constant or varies according to changes in location and observation window $W \subset R^2$ with $\mu(W) > 0$ if it meets the following conditions [12].

1. The number of random points in *W* namely *N(W)* has a Poisson distribution with an average of *μ(W)*.
2. The points *N(W)* is independent with the intensity function (**u**) where $u \in W$.
3. The expectation value of the point in the *W* region is *E(N(W))*=$\int_W \lambda(u)\, du$.
4. If $W_1, W_2, ..., W_n$ is a region that does not intersect from a space, then $N(W_1), N(W_2), ..., N(W_n)$ are independent random variables.
5. If *N(W)=n*, then the n points have identical and independent distributions, with a probability density function (pdf) as in equation (1).

f(**u**) = $\frac{\lambda(u)}{I}$, where $I = \int_W \lambda(u)\, du$ (1)

The intensity function of the Inhomogeneous Poisson point process forms a loglinear model that depends on the parameters θ and is affected by covariates [12][14], written as follows

$$\lambda_\beta(u) = expexp\left(\beta^T Z(u)\right) \quad (2)$$

where $Z(u) = \left(Z_1(u), \ldots, Z_p(u)\right)^T$ is covariate vector and $\beta = \left(\beta_0, \ldots, \beta_p\right)$ is the regression parameter for each predictor variable in the model [14].

2.3. Homogeneity and Independent Testing

To detect a spatial inhomogeneity of the death cases distribution, we use the quadrat.test function of the spatstat R package to apply quadrat counting test with a null hypothesis of stationary Poisson process against an inhomogeneous Poisson process closeness of the relationship between variables is correlation. In spatial point process analysis, a very popular technique used to analyze correlations in point patterns is the K-Function [15]. The concept of the K-Function is to calculate the distance d_{ij} = ||$\mathbf{u}_i$- $\mathbf{u}_j$|| between all pairs of different points $\mathbf{u}_i$ and $\mathbf{u}_j$ in a point pattern **x**. With this distance, the spatial pattern formed will be known. If the pattern forms groups/clusters, almost every distance between pairs of points is of small value, whereas if the pattern is regular then there are only a few distances between pairs of points of small value.

If the intensity is inhomogeneous, then the weight for $\mathbf{u}_i$ is added by b_i=1/(*u*$_i$). b_i is the intensity proportional to the points $\mathbf{u}_i$. While the pair $\mathbf{u}_i$ and $\mathbf{u}_j$ will add a weight of b_{ij}= b_i b_j = 1/(($\mathbf{u}_i$) ($\mathbf{u}_j$)), so that the empirical K-Function formula with inhomogeneous intensity can be written as in equation (3).

$$\hat{K}_{inhom}(r) = \frac{1}{D^p W}\sum_{i=1}^{n}\sum_{j=1\ j\neq 1}^{n} I\left\{\frac{d_{ij}}{\hat{\lambda}(u_i)\hat{\lambda}(u_j)}\right\} h_{ij}(u_i;u_j;r)(3)$$

with $D = \frac{1}{W}\sum_{i=1}^{n}\frac{1}{\hat{\lambda}(u_i)}$ where $h_{ij}(u_i;u_j;r)$ is the edge correction weight and $i - u_j$ calculating the distance from point $u_i$to point u_j.

2.4. Parameter Estimation

In the inhomogeneous Poisson process model with intensity $\lambda\left(u;\beta\right)$ there is a parameter denoted by β. The likelihood function used to estimate β can be written as in equation (4).

$$L\left(\beta\right) = L\left(\beta;u\right) = \lambda(\beta;u_1)\ldots\ \lambda(\beta;u_n)\exp\left(\int_W \left(1-\lambda\left(u;\beta\right)\right)du\right)\ (4)$$

Based on equation (4), a log-likelihood function can be formed as in equation (5).

$$\text{Log } L(\ )=\sum_{i=1}^{n} loglog\ \lambda(\beta;u_{\ i})\text{-} \int_W \lambda\left(u;\beta\right)du\ (5)$$

Based on equation (5), it can be seen that the log-likelihood function of the inhomogeneous Poisson process involves the integral of the observation window **W**. The numerical quadrature method was developed for the inhomogeneous Poisson Point Process so that the likelihood function is close to the likelihood function of the Generalized Linear Poisson Model. Using the numerical quadrature approach, then $\int_W \lambda\left(u;\beta\right)du$

can be approximated by $\sum_{i=1}^{n+q} \lambda(\beta; u_i) w_i$, where w_i is the quadrature weight and q is the number of dummy points so that equation (5) can be written as equation (6).

Log $L(\beta)=\sum_{i=1}^{n} loglog\ \lambda(\beta; u_i)- \sum_{i=1}^{n+q} \lambda(\beta; u_i) w_i$ (6)

The simple form of equation (6) can be written as in equation (7).

Log $L(\beta)=\sum_{i=1}^{n+q} (I_i\ loglog\ \lambda(\beta; u_i)- \lambda(\mathbf{u}_i;\beta) w_i$ (7)

I_i value 1 if u_i is a data point, while u_i he dummy point I_i is 0. Equation (7) can also be written as in equation [16]

Log $L(\beta)=\sum_{i=1}^{n+q} (y_i\ loglog\ \lambda(\beta; u_i)- \lambda(\mathbf{u}_i;\beta) w_i$ (8)

where $y_i = I_i/w_i$, equation (8) is equivalent to the Weighted Poisson likelihood function which is weighted w_i.

2.5. Model Assessment

The goodness of the resulting model can be checked using the K-Function envelope plot and the BIC value. Envelope is the critical limit of the statistical test of the K-Function function which validates the suitability of the point pattern data with the point process model [12]. The model will be said to be good in modeling certain data if the K-Function plot of the original point data is in the K-Function envelope interval data. In addition to using the K-Function envelope plot, in this study the best model was selected based on the smallest BIC value. The use of the BIC value takes into account the large heterogeneity of the data and is preferred if the aim is to select the model. Where the BIC criteria can be defined as in equation (9) [14].

$$BIC = -2L_n(\hat{\beta}) + plog(N(W))(9)$$

where L_n is a function of the log likelihood of the model and *N(W)* is the number of points in an observation window.

3. RESULTS AND DISCUSSION

3.1. Data Characteristics

The cumulative number of deaths of COVID-19 patients in the observation area of Surabaya City from May 1, 2020 to April 30, 2021 after pre-processing was recorded as many as 254 cases of death, where the distribution is shown in Fig. 1.

Figure 1 shows that the central and northern areas of Surabaya City are the areas with the highest COVID-19 death cases with a case percentage of 38.5% then spread to the

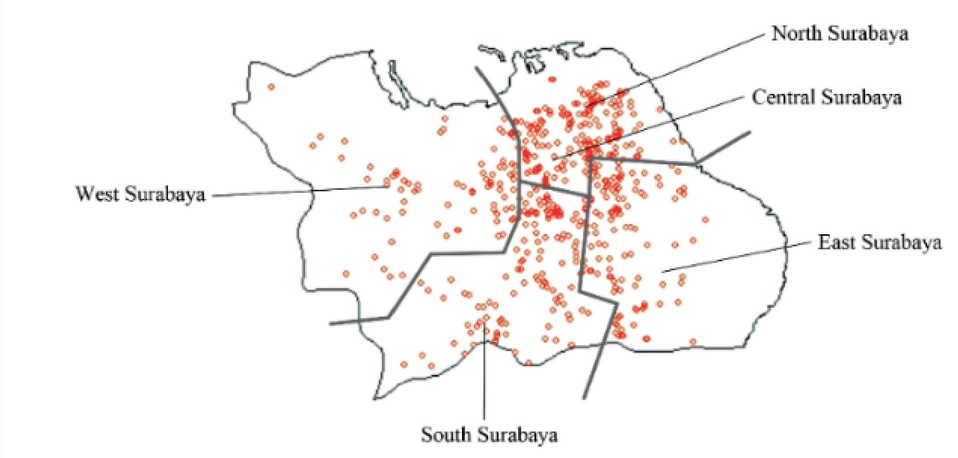
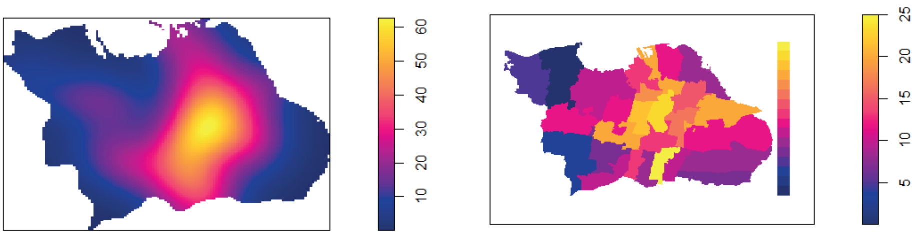


Figure 1: Map of the distribution of death cases of Covid-19 patients in Surabaya from 1 May 2020 to 30 April 2021.

eastern region with a case percentage of 29.4% and the south with a case percentage of 21.6 %. Figure 2 shows the density plot for the density of the location of the COVID-19 referral hospital and the proportion of the population with confirmed COVID-19 age >60 years. Based on Fig. 2 (a), it can be seen that the density of referral hospital locations with high intensity is dominated in the central area of Surabaya City which spreads to the north and south areas. Meanwhile, the density of the proportion of people with confirmed COVID-19 aged >60 years as shown in Fig. 2 (b) has the highest intensity in the southern and central Surabaya areas.



(a) (b)

Figure 2: Density plot of covariate density in percent a) COVID-19 referral hospital b) Proportion of population confirmed positive with age >60 years per sub-district. The brighter the density plot, the higher the intensity or density.

3.2. Spatial Trend and Clustering Detection

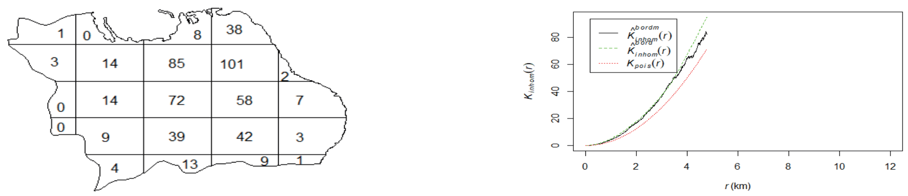
To test data homogeneity of death cases due to COVID-19 in Surabaya using the Chi-Squared test in equation (10) [12].

$$X^2_{hit} = \sum_{j=1}^{m} \frac{(n_j e_j)^2}{e_j} = \sum_{j=1}^{m} \frac{(n_j - \frac{n}{m})^2}{n/m} (10)$$

TABLE 1: Chi-square test of Surabaya City COVID-19 death data.

Chi-Square Test	
X^2	931.68
df	79
p-value	<2.2 x 10^{-16}

Since the p-value less than α = 0,05 we conclude that the distribution of death cases of COVID-19 in Surabaya is not distributed homogeneity. This is clarified by Fig. 3(a), there are areas with a large number of deaths, but in other areas there are small deaths, namely in the eastern and western areas of Surabaya. The results of the Chi-Squared test show that the pattern of address data for COVID-19 deaths in Surabaya is not stationary (inhomogeneous), so a spatial correlation analysis can be carried out to detect whether the pattern of address points for COVID-19 deaths in Surabaya tends to be regular (away from each other), independent. (random), or cluster (clustered) using the inhomogeneous K-function. Visualization results of inhomogeneous K-function using border edge correction (green curve) and modified border edge correction (black curve) are shown in Fig. 3(b).



(a) (b)

Figure 3: a) Quadrat count plot of the distribution of COVID-19 death cases in Surabaya b) Inhomogeneous K-function plot data on COVID-19 death cases in Surabaya.

Based on Fig. 3(b), it is obtained that from testing with the inhomogeneous K-Function plot, it shows that the short distance is around 0-4 km, empirical curve $\hat{K}$(r) (green, black curve) above the theoretical curve $\hat{K}$(r) (red curve) but not too far away. So, in this study using the Poison model according to the fact that there is no interaction between one death and another in this case. We further test this using the envelope detected in Fig. 5.

3.3. Modelling Results

Table 2 shows the results of the parameter estimation and the parameter significance test of the inhomogeneous Poisson point process model from the address data for COVID-19 deaths in Surabaya.

Table 2: Inhomogeneous poisson point process model parameter estimation.

Variable	Coefisien	Exp (coef)	lZ-valuel	Significant
Intercept	-0.754332	0.4703247	-5.7354	√
Spatial covariate:				
COVID-19 referral hospital	0.032418	1.03295	10.2460	√
Proportion of positive confirmed population aged >60 years per sub-district	0.037634	1.03835	4.19290	√

The critical value used is $Z_{\alpha/2}$ =1.96 and all covariates have a test statistic value of more than the critical value, so the null hypothesis is rejected, which means that at the 95% confidence level, it is statistically proven that all covariates, namely the COVID-19 referral hospital and the proportion of positive confirmed population aged >60 years, have a significant effect on the intensity or risk of COVID-19 death in Surabaya.

Evaluation of the goodness of the inhomogeneous Poisson point process model can be done using the BIC value and visualization of the inhomogeneous K-function envelope. The comparison of BIC values is used to determine the best inhomogeneous Poisson point process model in predicting the risk of COVID-19 death in Surabaya, which is shown in Table 3.

Table 3: BIC value comparison.

Model	BIC
Covariate model of the COVID-19 referral hospital and the proportion of confirmed COVID-19 population aged >60 years per sub-district	158.8779
Covariate model of the COVID-19 referral hospital	343.8003
Covariate model of the proportion of confirmed COVID-19 population aged >60 years per sub-district	254.4867

A good model is determined from the lowest BIC value. Based on Table 3, the inhomogeneous Poisson point process model with the covariate model of the COVID-19 referral hospital and the proportion of confirmed COVID-19 population aged >60 years per sub-district has a lower BIC value than the other model. so this model is the best model to explain the influence of spatial trends on the risk of death from COVID-19 in Surabaya. Visualization of the inhomogeneous K-function envelope for the inhomogeneous Poisson point process model with edge correction using all covariates is shown in Fig. 5.

In Fig. 4 it can be seen that at close distances (under 5 km), $\hat{K}$(r) the empirical curve tends to be relatively close to the Poisson process curve $\hat{K}$(r) and around the confidence interval area so that the overall model is quite good. The intensity function shows the

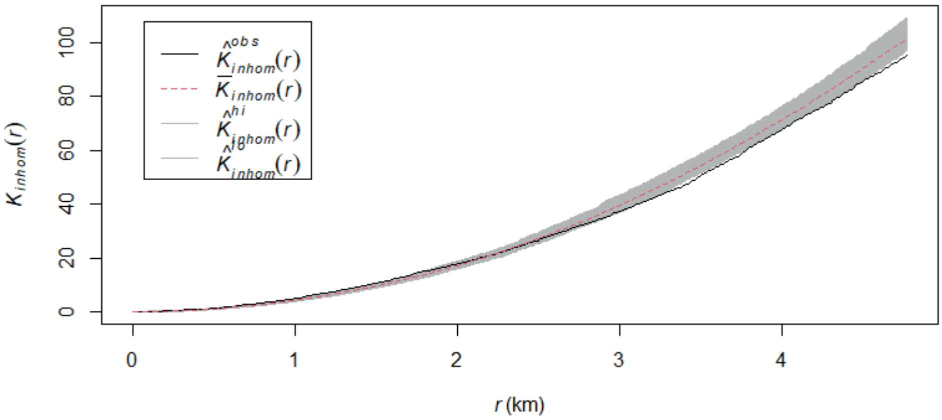


Figure 4: Envelope inhomogeneous K-function with edge correction on covariate model of COVID-19 referral hospital and proportion of population confirmed with COVID-19 Age >60 years per district.

probability of observing a point in the address of a COVID-19 death in a very small area, which is mathematically written by equation (11).

$$\lambda(u) = expexp\ (\ -1.754332 + 0.32418 Z_1 + 0.37634 Z_2)(11)$$

The model equation 11 obtained shows that the risk of COVID-19 death in a location is influenced by the density of the COVID-19 referral hospital, which means that every additional 1 COVID-19 referral hospital in 1 km^2 will increase the risk of COVID-19 1.03295 times compared to without increasing the number of referral hospitals. Meanwhile, if in 1 km^2 there is an addition of 1% of confirmed COVID-19 population with age > 60 years, it will increase the risk of COVID-19 death in Surabaya by 1.03835 times. Between the two variables that have a large effect on the risk of death is the proportion of population confirmed COVID-19 aged >60 years. So, there is a need for special handling for areas that have a high proportion of population confirmed COVID-19 aged >60 years such as vaccination for the elderly, tracing, and testing. Interestingly, health facilities have a positive effect on the risk of death. This means that the higher the number of COVID-19 referral hospitals, the higher the risk of death from COVID-19. This is due to the similarity of the pixel image between the point pattern of death and the covariates of the referral hospital. There are also other things that affect it such as the number of COVID-19 cases. it is possible that areas with small COVID-19 death cases do have a small number of positive COVID-19 cases[4]. So, it is necessary to examine this covariate more deeply for this covariate.

3.4. Prediction

The prediction of the risk of death from COVID-19 in Surabaya using the inhomogeneous Poisson Point Process model is shown in Fig. 5.

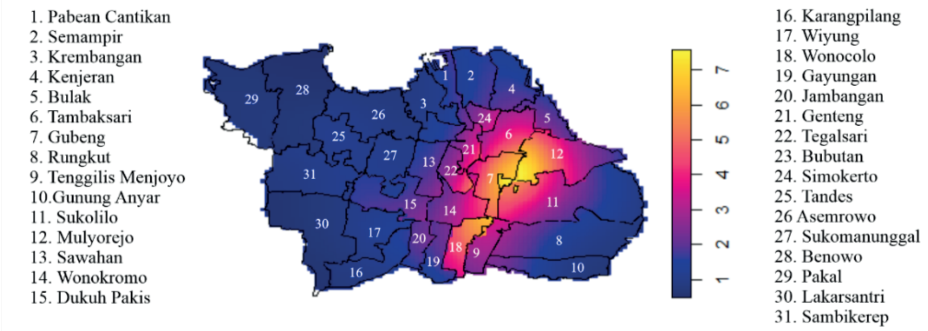


Figure 5: Predicted map of COVID-19 death risk in Surabaya.

The prediction of the risk of COVID-19 death in Surabaya is depicted in Fig. 5 which is a prediction in the form of the intensity of the death of COVID-19 in Surabaya based on the variable density of the COVID-19 referral hospital and the density of the proportion of the population confirmed positive with age >60 years. The prediction of COVID-19 death cases in Surabaya as shown in Fig. 5 shows that the epicenter of the COVID-19 death cases is around Surabaya. It can be seen that the results of the prediction of the risk of death from COVID-19 in the eastern Surabaya area are higher when compared to other areas. Where the prediction of the risk of positive confirmed cases of COVID-19 is the highest, reaching around 7 cases per km^2.

4. CONCLUSION

Apart from the fact that COVID-19 cases have decreased, very important to have a better understanding of the pandemic and effective programs to control COVID-19. Areas with a high incidence of infection in the population aged >60 years have a greater risk of death. In addition, in future studies it is necessary to consider positive cases of COVID-19 because it could be a high case of death in an area because the number of positive cases in that area is also high. Although the results of the analysis on the covariate of referral hospital density have a significant positive effect, further analysis needs to be carried out by considering other covariates such as bed occupancy rate, individual health status including vaccination status, comorbid disease status, and other covariates so that more information is obtained about the risk of death due to COVID-19 in Surabaya.

Acknowledgments

This study is supported by the Direktorat Riset, Teknologi, dan Pengabdian Kepada Masyarakat, Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

References

[1] W.H.O. “Novel corona virus,” https://www.who.int/indonesia/news/novel-coronavirus.

[2] S.C. -19 Jatim, “Jatim tanggap Covid-19,” https://infocovid19.jatimprov.go.id/

[3] Park J, Chang W, Choi B. An interaction Neyman-Scott point process model for coronavirus disease-19. Spat Stat. 2022 Mar;47(47):100561.

[4] A. Choiruddin, F. Hannanu, J. Mateu, and V. Fitriyanah, “COVID-19 transmission risk in Surabaya and Sidoarjo: An inhomogeneous marked Poisson point process approach,” (2022). https://doi.org/10.21203/rs.3.rs-2258695/v1.

[5] Virania TA, Choiruddin A, Ratnasari V. Analisis risiko penyebaran kasus Covid-19 di Surabaya Raya menggunakan model thomas cluster process. Inferensi. 2021;4(1):57.

[6] Izati PP, Choiruddin A. Pemodelan risiko penyebaran COVID-19 di Surabaya Raya menggunakan model cauchy cluster process. Inferensi. 2022;5(1):1.

[7] Nurhayatun S, Fitriyanti N, Sonia D. Analisis Average Length of Stay (AVLOS) kasus covid-19 di rumah sakit x Bandung. Jurnal Ilmiah Manusia Dan Kesehatan. 2021;4(3):412–8.

[8] S.&. N. Tamher, *Kesehatan usia lanjut dengan pendekatan asuhan keperawatan. Salemba Medika*, Jakarta, 2009.

[9] Luo Y, Yan J, McClure S. Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: a spatial nonlinear analysis. Environ Sci Pollut Res Int. 2021 Feb;28(6):6587–99.

[10] Elson R, Davies TM, Lake IR, Vivancos R, Blomquist PB, Charlett A, et al. The spatio-temporal distribution of COVID-19 infection in England between January and June 2020. Epidemiol Infect. 2021 Mar;149:e73.

[11] P. Purwanto, S. Utaya, B. Handoyo, et al., “Spatiotemporal analysis of COVID-19 spread with emerging hotspot analysis and space-time cube models in East Java, Indonesia.,” *ISPRS International Journal of Geo-Information*. vol. 10, no. 3, p. 2021. https://doi.org/10.3390/ijgi10030133.

[12] Baddeley A, Rubak E, Turner R. Spatial point patterns methodology and applications with R. Boca Raton: CRC Press; 2015. https://doi.org/10.1201/b19708.

[13] J. Møller and R.P. Waagepetersen, “Modern statistics for spatial point processes.,” *Scandinavian Journal of Statistics*. vol. 4, no. 34, pp. 070927154002001-???, 2007. https://doi.org/10.1111/j.1467-9469.2007.00569.x.

[14] Choiruddin A, Coeurjolly J, Waagepetersen R. Information criteria for inhomogeneous spatial point processes. Aust N Z J Stat. 2021;63(1):119–43.

[15] Ripley B. Modelling spatial patterns. J R Stat Soc B. 1977;39(2):172–92.

[16] Berman M, Turner TR. Approximating point process likelihoods with GLIM. Appl Stat. 1992;41(1):31.