 doi:10.20944/preprints202209.0271.v1

Article

Does the Restriction of Human Mobility Significantly Control COVID-19 Transmission in Jakarta, Indonesia? Global Versus Local Regression Models

I Gede Nyoman Mindra Jaya [1*], Anna Chadidjah [2], Gumgum Darmawan [3], Jane Christine Princidy [4], and Farah Kristiani [5]

[1] Statistics Department, Universitas Padjadjaran, Bandung, Indonesia 1; mindra@unpad.ac.id
[2] Statistics Department, Universitas Padjadjaran, Bandung, Indonesia 2; anna.chadidjah@unpad.ac.id
[3] Statistics Department, Universitas Padjadjaran, Bandung, Indonesia 3; gumgum@unpad.ac.id
[4] Statistics Department, Universitas Padjadjaran, Bandung, Indonesia 4; jane18001@mail.unpad.ac.id
[5] Mathematics Department, Parahyangan Catholic University, Bandung, Indonesia 5; farah@unpar.ac.id

* Correspondence: mindra@unpad.ac.id; Tel.:+6281809319977

Abstract: COVID-19 is the most severe health crisis of the 21st century. COVID-19 presents a threat to almost all countries world-wide. The restriction of human mobility is one of the strategies used to control the transmission of COVID-19. However, it has yet to be determined how effective this restriction is in controlling the rise in COVID-19 cases, particularly in major capital cities such as Jakarta, Indonesia. Using Facebook's mobility data, our study explores the impact of restricting human mobility on COVID-19 case control in Jakarta. Our main contribution is showing how the restriction of human mobility data can give important information about how COVID-19 spreads in different places. We proposed modifying a global regression model into a local regression model by accounting for the spatial and temporal interdependence of COVID-19 transmission across space and time. We applied Bayesian hierarchical Poisson spatiotemporal models with spatially varying regression coefficients. We estimated the regression parameters using an Integrated Nested Laplace Approximation. We found that the local regression model with spatially varying regression coefficients outperforms the global regression model based on DIC, WAIC, MPL, and R^2 criteria for model selection. In Jakarta's 44 districts, the impact of human mobility varies significantly. The impacts of human mobility on the log relative risk of COVID-19 range from –4.445 to 2.353. The prevention strategy involving the restriction of human mobility may be beneficial in some districts but ineffective in others. Therefore, a cost-effective strategy had to be adopted.

Keywords: COVID-19; human mobility; spatial autocorrelation; temporal autocorrelation; Facebook mobility data

1. Introduction

COVID-19 is the greatest health threat of the 21st century [1]. According to [2], more than 200 countries were exposed to COVID-19 from December 2019 to August 2022, resulting in more than 500 million cases and approximately 6.4 million deaths. In the absence of a widely available vaccination, restricting human mobility is argued as the most effective method for preventing the spread of COVID-19. According to [3], human mobility is essential for the transmission of infectious diseases. As transportation networks and globalization grow and make people more mobile, COVID-19 can spread quickly and be hard to stop.

Accurate models that anticipate the transmission of COVID-19, are necessary to inform population-level intervention decisions [4]. The accurate COVID-19 risk estimate could be determined by examining human mobility. Numerous investigations have

 doi:10.20944/preprints202209.0271.v1

shown that human mobility is successful in reducing the transmission of COVID-19 [5-9]. However, due to the various characteristics of different regions, human mobility may impact COVID-19 transmission differently between regions [10]. Disease risk mapping can contribute to a better understanding of risk evolution over space and time [11-13]. Ecological regression models with global (fixed) regression coefficients are frequently used to identify relevant risk factors in order to provide precise risk estimates and covariate effects [14-15]. These global regression models often rely on an important assumption regarding the homogeneity of the population. It figures out the fixed regression coefficients for the covariates based on the assumption that the parameters are constant over time. This assumption does not fit the reality because of the socioeconomic conditions and environments that may influence the human behaviors that contribute to the transmission of COVID-19 in different populations [10]. Given the complexity of the relationship between COVID-19 and human mobility variables, it is logical that the impact of human mobility factors be non-stationary. For example, the effect of restricting human mobility on COVID-19 has two opposing sides: on the one hand, mobility restrictions reduce COVID-19 transmission. However, this may result in a decline in economic growth, which may increase stress levels vulnerable to COVID-19. We can fix the issue if we permit location-dependent variation in the mobility-related effects of COVID-19 transmission. According to [14], taking into account the variance instability caused by unobserved risk factors, different populations, and spatial interdependency caused by similar identical risk factor conditions improves the accuracy of disease risk prediction.

Our research offers an alternative to the global regression models for disease modeling and mapping, which presume that regression coefficients are constant across neighborhoods. The spatially varying coefficient (SVC) and geographically weighted regression (GWR) models can accommodate the non-stationarity effects of the covariates. GWR is focused on continuous response variable, which are, however, rarely employed in disease mapping [16]. SVC models, on the other hand, can be used to model discrete outcomes (like binary and count) and can be easily adapted to random effects components in generalized linear models (GLM), which can be estimated using Bayesian techniques (see, for example, [17–18]). The advantage of SVC, according to [19], is that it provides a probability model from which all model parameters can be estimated and relative risk estimates can be derived from the full posterior distribution. SVC models provide a more robust inferential framework when testing hypotheses regarding model parameters or evaluating prediction uncertainty. In addition, Bayesian is a smoothing method that can reduce the noise, resulting in a reliable estimation of the relative risk [20-21]. Thus, SVC models are appropriate for correctly describing regionally variable health issues [22]. The Bayesian smoothing is modeled via a random effects model. Commonly, the Bayesian conditional autoregressive (CAR) random effect model is employed to describe the spatially structured effects in ecological regression for areal data in which the number of COVID-19 cases is typically aggregated by administrative area [23]. To account for temporal autocorrelation, random walk or autoregressive models are commonly employed.

Our empirical application is motivated by the epidemiology of COVID-19 in Jakarta, Indonesia. Jakarta is the capital of Indonesia, where the implementation of Community Activity Restrictions (called PPKM) restricts the movement of people in an effort to prevent the spread of COVID-19. Digital data on human mobility are being utilized to study shifting patterns and understand the effects of prevention and treatment actions. Throughout the COVID-19 pandemic, Facebook mobility data primarily derived from mobile phone usage has been utilized and made accessible [24-25]. Here, we examine aggregated and anonymized Facebook data on the mobility patterns of active users in Jakarta who made use of geolocation services between 3 July 2021 and 6 August 2021. The goal of this study is to explain the methodological and substantive epidemiological implications that can be used to develop an early warning system for future disease outbreaks.

Using a hierarchical Bayesian modeling framework, we evaluate many spatiotemporal Poisson regression models that take into account spatiotemporal dependence and heterogeneity. Simulations of numerically evaluating complicated integrals with Markov

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

Chain Monte Carlo (MCMC) can be challenging and aggravating [26-27]. For approximate Bayesian inference, we utilized the Integrated Nested Laplace Approximation (INLA). The study has two goals. Firstly, by analysing the spatially varying effects of human mobility on COVID-19 risk; and secondly, by mapping the relative risk estimates over space and time. The following section discusses the method, statistical analysis, empirical results, and discussion. Finally, we end with the conclusions.

2. Materials and Methods

2.1 Study area and data

Jakarta is the most populous city in Southeast Asia and the capital of Indonesia which consist of 44 districts. It is located on Java, the most populous island in the world. Its population will reach 10,562,080 in 2020. Jakarta has the smallest land area of any province in Indonesia, but its metropolitan area covers 9,957.08 km2 and is projected to have 35 million residents by 2021, making it the largest urban area in Indonesia. Jakarta is the most developed province in Indonesia. People from all over Indonesia have moved to Jakarta because it is a good place to do business and has a higher standard of living.

We conducted a local regression analysis between the mortality rate of COVID-19 and the human mobility index in Jakarta province. This study employs secondary data regarding the cumulative daily number of COVID-19 cases per week and the indicator of community movement (Facebook Mobility Index) in 44 sub-districts of Jakarta Province per week from July 3, 2021 to August 6, 2021, or during the enforcement of restrictions on community activities period. This data was gathered from the Jakarta COVID-19 file history website (https://riwayat-file-covid-19-dki-jakarta-jakartagis.hub.arcgis.com/) and the Data for Good Facebook website (https://dataforgood.facebook.com/).

2.2. Facebook mobility data

The Facebook Mobility Data [25] reveals the locations, movements, and connections of active Facebook users. The data is created by Facebook's location tracking, which integrates geolocation tools and connectivity information (e.g., wi-fi) from smart phones with the Facebook app installed to give users a geographical position at a given time [24].

3. Statistical Analysis

3.1. Model specification

We consider y_{it} representing the spatiotemporal outcomes of COVID-19, and N_{it} denotes the population at risk aggregated by districts $i = 1, \dots, n$ and time $t = 1, \dots, T$. According to y_{it} is count data follows a Poisson distribution, that is, $y_{it} \sim Poisson(E_{it}\theta_{it})$ with the likelihood function:

$$L(\mathbf{y}|E_{it}, \theta_{it}) = \prod_{i=1}^{n}\prod_{t=1}^{T} \frac{\exp(-E_{it}, \theta_{it})\,(E_{it}, \theta_{it})^{y_{it}}}{y_{it}!} \quad (1)$$

where E_{it} and θ_{it} denote the expected number of cases and the relative risk in area i and time t respectively. The relative risk θ_{it} is the health indicator necessary in disease mapping studies to advise policymakers about when and where outbreaks are occurring. The "crude" estimate of the relative risk is called standardized incidence ratio (SIR) is the ratio of observed cases to the expected number of cases, $SIR_{it} = y_{it}/E_{it}$. The E_{it} is given by:

$$E_{it} = N_{it}\frac{\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T} N_{it}} \quad (2)$$

Note that population heterogeneity will influence the estimation of the SIR. As a result, the fixed effect parameter of small areas with a small population size N_{it} will have

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

a high degree of variability. The high variability amongst areas due to population heterogeneity is usually overcome by imposing an independent Gaussian (exchangeable) prior distribution on log-relative risk, $\log\theta_{it} \sim N(\alpha, \sigma_u^2)$ resulting in a log-linear Poisson regression model with random intercept. Thus, $\log\theta_{it} = \alpha + u_i$, implying $\theta_{it} = \exp(\alpha + u_i)$, where $u_i \sim N(0, \sigma_u^2)$ and $p(\boldsymbol{u}|\sigma_u^2) \propto \sigma_u^{-n} \exp(-0.5\sigma_u^{-2} \sum_{i=1}^{n} u_i^2)$. Where α is the global logarithmic level of the relative risk and u_i represents the spatially unstructured random effects. It is likely that the relative risks in a number of adjacent regions will reflect a geographical pattern [14].

Unmeasured confounding variables are also possibly spatially continuous and can display spatial correlation. Typically, such confounding variables are accounted for by introducing a spatially structured random effects component ω_i that describes the probability distribution of ω_i conditional on the set $\boldsymbol{\omega}_{-i} = \boldsymbol{\omega}_{j\neq i} = \{\omega_1, \ldots, \omega_{i-1}, \omega_{i+1}, \ldots, \omega_n\}$. The intrinsic conditional autoregressive (ICAR) prior is a frequently used method for representing irregularly shaped regions, where the conditional distribution of ω_i is given by [14]:

$$\omega_i|\boldsymbol{\omega}_{-i} \sim N\left(\bar{\omega}_i, \frac{\sigma_\omega^2}{w_{i+}}\right) \tag{3}$$

where $\bar{\omega}_i = \sum_j w_{ij}\omega_j / w_{i+}$ and w_{i+} with w_{ij} denotes $n \times n$ binary spatial weight matrix that express the spatial dependency structure with $w_{ij} = 1$ if i and j are neighbors $i \sim j$ (have a shared boundary) and 0 otherwise. The given specification (3) results in the joint distribution for vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)'$ [14]:

$$p(\boldsymbol{\omega}|\sigma_\omega^2) \propto \sigma_\omega^{-n} \exp\left(-0.5\sigma_\omega^{-2} \sum_{i=1} \sum_{i\sim j} w_{ij}(\omega_i - \omega_j)^2\right) \tag{4}$$

This prior is improper prior such that the sum to zero constraint $\sum_{i=1} \omega_i = 0$ is required to ensure identifiability. To avoid the challenge of selecting between spatially structured and unstructured effects, it is possible to combine these two priors as follow [14]:

$$\log\theta_{it} = \alpha + (u_i + \omega_i) \tag{5}$$

The linear predictor $\alpha + u_i + \omega_i$ denotes the random intercepts over areas. Accounting for temporally unstructured and structured effects and spatiotemporal interaction effects, model (5) can be modified becomes [14]:

$$\log\theta_{it} = \alpha + u_i + \omega_i + \varsigma_t + v_t + \delta_{it} \tag{6}$$

where φ_t and v_t denote the temporally unstructured and structured effects respectively. The temporally unstructured effects is usually modeled as exchangeable prior distribution (i.e., $\varphi_t \sim N(0, \sigma_\varphi^2)$. We consider random walk model of order 1 (RW1) to model temporal autocorrelation. The Gaussian vector $\boldsymbol{v} = (v_1, \ldots, v_T)'$ is defined by assuming independent differencing [14]:

$$\Delta v_t = v_t - v_{t-1} \sim N(0, \sigma_v^2) \tag{7}$$

The density function for $\boldsymbol{v}$ is obtained from its $T-1$ increments as follows:

$$p(\boldsymbol{v}|\sigma_v^2) \propto \sigma_v^{-T} \exp\left(-0.5\sigma_v^{-2} \sum_{t=1} (\Delta v_t)^2\right) \tag{8}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

The spatiotemporal interaction effects component δ_{it} can be specified in four different types [28]. Type I denotes the space-time interaction between temporally and spatially unstructured effects, type II describes the space-item interaction between temporally structured and spatially unstructured effects, type III is the interaction between temporally unstructured and spatially structured effects, and finally type IV is the interaction between temporally and spatially structured effects. In addition, the model (6) can be modified to account for the effect of risk factor x_{it}:

$$\log\theta_{it} = \alpha + u_i + \omega_i + \varsigma_t + v_t + \delta_{it} + \beta x_{it} \quad (9)$$

where x_{it} is the risk factors of the human mobility index at location i and time t with coefficient β. Note that the covariate can be continuous or discrete variables. In order to account for spatial heterogeneity in effects of the risk factor x_{it}, the model can be reparametrized by varying the coefficients over areas. A model with spatially varying coefficients is defined:

$$\log\theta_{it} = \alpha + u_i + \omega_i + \varsigma_t + v_t + \delta_{it} + (\beta + \varphi_i)x_{it} \quad (10)$$

where φ_i represents differential spatially varying regression effects which account for spatial heterogeneity effects of the risk factor x_{it}. Consequently, $\beta_i = \beta + \varphi_i$ describes the random slope processes. The common specification for φ_i is either ICAR process $\varphi_i \sim ICAR\left(w, \sigma_\varphi^2\right)$ or exchangeable Gaussian processes $\varphi_i \sim N\left(0, \sigma_\varphi^2\right)$. Some random effects components in model (10) should be eliminated to avoid the overfitting and confounding issues.

3.2. Bayesian inference

Let $\boldsymbol{\Phi} = \{\alpha, \beta, \boldsymbol{u}, \boldsymbol{\omega}, \boldsymbol{\varsigma}, \boldsymbol{v}, \boldsymbol{\delta}, \boldsymbol{\varphi}\}$ denotes the vector Gaussian latent (unobservable) field and $\boldsymbol{\Psi} = \left\{\sigma_u^2, \sigma_\omega^2, \sigma_\varsigma^2, \sigma_v^2, \sigma_\delta^2, \sigma_\varphi^2\right\}$ be the vector of hyper-parameters. The vector $\boldsymbol{\Phi}$ are conditionally independent multivariate Gaussian distribution with the sparse precision matrix $Q_{ij} = 0$ for $\Phi_i \perp \Phi_j | \Phi_{-ij}$. Bayesian inference is introduced in three stages as follows:

Stage 1: $\mathbf{y}|\boldsymbol{\Phi}, \boldsymbol{\Psi} \sim p(\mathbf{y}|\boldsymbol{\Phi}, \boldsymbol{\Psi})$

Stage 2: $\boldsymbol{\Phi}|\boldsymbol{\Psi} \sim p(\boldsymbol{\Phi}|\boldsymbol{\Psi})$

Stage 3: $\boldsymbol{\Psi} \sim p(\boldsymbol{\Psi})$

The joint posterior distribution of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ conditional on the data likelihood is:

$$p(\boldsymbol{\Phi}, \boldsymbol{\Psi}|\mathbf{y}) = \frac{p(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \mathbf{y})}{p(\mathbf{y})} = \frac{p(\mathbf{y}|\boldsymbol{\Phi}, \boldsymbol{\Psi})p(\boldsymbol{\Phi}|\boldsymbol{\Psi})p(\boldsymbol{\Psi})}{\int_{\boldsymbol{\Phi}} \int_{\boldsymbol{\Psi}} p(\mathbf{y}|\boldsymbol{\Phi}, \boldsymbol{\Psi})p(\boldsymbol{\Phi}|\boldsymbol{\Psi})p(\boldsymbol{\Psi})d\boldsymbol{\Phi}d\boldsymbol{\Psi}} \quad (11)$$

The joint posterior distribution can be expressed as $p(\boldsymbol{\Phi}, \boldsymbol{\Psi}|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\Phi}, \boldsymbol{\Psi})p(\boldsymbol{\Phi}|\boldsymbol{\Psi})p(\boldsymbol{\Psi})$ since the denominator is integrated across the parameters of the latent field. Integrals can be solved via simulation or numerical approaches. We considered making use of Integrated Nested Laplace Approximation (INLA), in which the complex integral can be computed numerically with a faster computation process compared to the Markov Chain Monte Carlo simulation approach (MCMC). The INLA calculating procedure may be summed up as:

(i) Approximate the posterior distribution of the hyper-parameter via nested approach:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

$$p(\boldsymbol{\Psi}|\mathbf{y}) = \frac{p(\boldsymbol{\Phi},\boldsymbol{\Psi}|\mathbf{y})}{p(\boldsymbol{\Phi}|\boldsymbol{\Psi},\mathbf{y})} \approx \left.\frac{p(\mathbf{y}|\boldsymbol{\Phi},\boldsymbol{\Psi})p(\boldsymbol{\Phi}|\boldsymbol{\Psi}),p(\boldsymbol{\Psi})}{p(\boldsymbol{\Phi}|\boldsymbol{\Psi},\mathbf{y})}\right|_{\boldsymbol{\Phi}=\boldsymbol{\Phi}^*(\boldsymbol{\Psi})} = \tilde{p}(\boldsymbol{\Psi}|\mathbf{y})$$

(ii) Utilize simplified Laplace's approximation of posterior marginal distribution using Tylor's series expansion:

$$\tilde{p}(\Phi_i|\boldsymbol{\Psi},\mathbf{y}) = \left.\frac{p(\boldsymbol{\Phi},\boldsymbol{\Psi}|\mathbf{y})}{\tilde{p}(\boldsymbol{\Phi}_{-i}|\Phi_i,\boldsymbol{\Psi},\mathbf{y})}\right|_{\boldsymbol{\Phi}_{-i}=\boldsymbol{\Phi}^*_{-i}(\Phi_i,\boldsymbol{\Psi})}$$

where $\tilde{p}(\boldsymbol{\Phi}_{-i}|\Phi_i,\boldsymbol{\Psi},\mathbf{y})$ is the Laplace-Gaussian approximation to $p(\boldsymbol{\Phi}_{-i}|\Phi_i,\boldsymbol{\Psi},\mathbf{y})$ and $\boldsymbol{\Phi}^*_{-i}(\Phi_i,\boldsymbol{\Psi})$ is its mode. Finally, the marginal posterior distributions are computed as $\tilde{p}(\Phi_i|\mathbf{y}) \approx \int \tilde{p}(\boldsymbol{\Phi}_{-i}|\boldsymbol{\Psi},\mathbf{y})\tilde{p}(\boldsymbol{\Psi}|\mathbf{y})\mathrm{d}\boldsymbol{\Psi}$ [14]. The marginal posterior distribution is utilized for parameter estimation. The parameters of the model are then utilized to estimate the relative risk over space and time. In addition, an exceedance probability is performed to determine whether a are has a notably high risk. More details can be seen in [23].

3.3. Model implementation

The case study utilized COVID-19 outcomes disaggregated by $i = 1,\ldots,44$ districts over $t = 1,\ldots,5$ weeks from 3 July 2021 to 6 August 2021 for model specification. We outfitted three distinct models. For the unstructured spatial effects, we specified $u_i \sim N(0,\sigma_u^2)$ and for the structured spatial effects, $\omega_i \sim ICAR(w,\sigma_\omega^2)$. Due to the limited temporal range of the data, we defined a first-order random walk prior $v_t = v_{t-1} + \Delta v_t, \Delta v_t \sim N(0,\sigma_v^2)$. We picked interaction type IV that is, $\delta_{it} \sim N(0,\sigma_\delta^2)$ to account spatiotemporal interaction because ω_i and v_t capture the spatially and temporally structured variations.

Model 1: Global model fixed effect regression model

$$\log\theta_{it} = \alpha + \beta x_{it}$$

Model 2: Spatially varying coefficient model with exchangeable prior

$$\log\theta_{it} = \alpha + v_t + \delta_{it} + (\beta + \varphi_i)x_{it};\ \varphi_i \sim N\left(0,\sigma_\varphi^2\right)$$

Model 3: Spatially varying coefficient model with ICAR prior

$$\log\theta_{it} = \alpha + v_t + \delta_{it} + (\beta + \varphi_i)x_{it};\ \varphi_i \sim ICAR\left(w,\sigma_\varphi^2\right)$$

Model 3 is supported by evidence of the spatial autocorrelation of human mobility. We examined the significance of spatial autocorrelation of human mobility using Moran's index.

To complete the Bayesian inference, we suggest a Gaussian prior distribution with a zero mean and a huge variance for α and β that is $\{\alpha,\beta\} \sim N(0,10^6)$. In addition, the variance hyperparameters $\boldsymbol{\Psi} = \{\sigma_v^2,\sigma_\delta^2,\sigma_\varphi^2\}$ require priors (hyperpriors). We utilized the proper half-Cauchy distribution with a scale parameter of 25 [23].

4. Results

4.1 Distribution of COVID-19 cases and human mobility

Figure 1 depicts the distribution of COVID-19 counts from Week 1 (3–9 July 2021) to Week 5 (31 July–6 August 2021). Cases ranged from 49 to 3,294 in the first week, 77 to 4,273 in the second, 29 to 1,949 in the third, 30 to 1,375 in the fourth, and 13 to 762 in the fifth. In addition, maps of human mobility indexes are shown in Figure 2. The human mobility

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

index fluctuated between -0.341 and -0.255 during Week 1, -0.420 and -0.264 during Week 2, -0.441 and -0.334 during Week 3, -0.433 and -0.263 during Week 4, and -0.447 and -0.285 during Week 5. A large negative index indicates limited human mobility. Consequently, it appears that the number of cases declines as the human mobility index falls (see Figure 3).

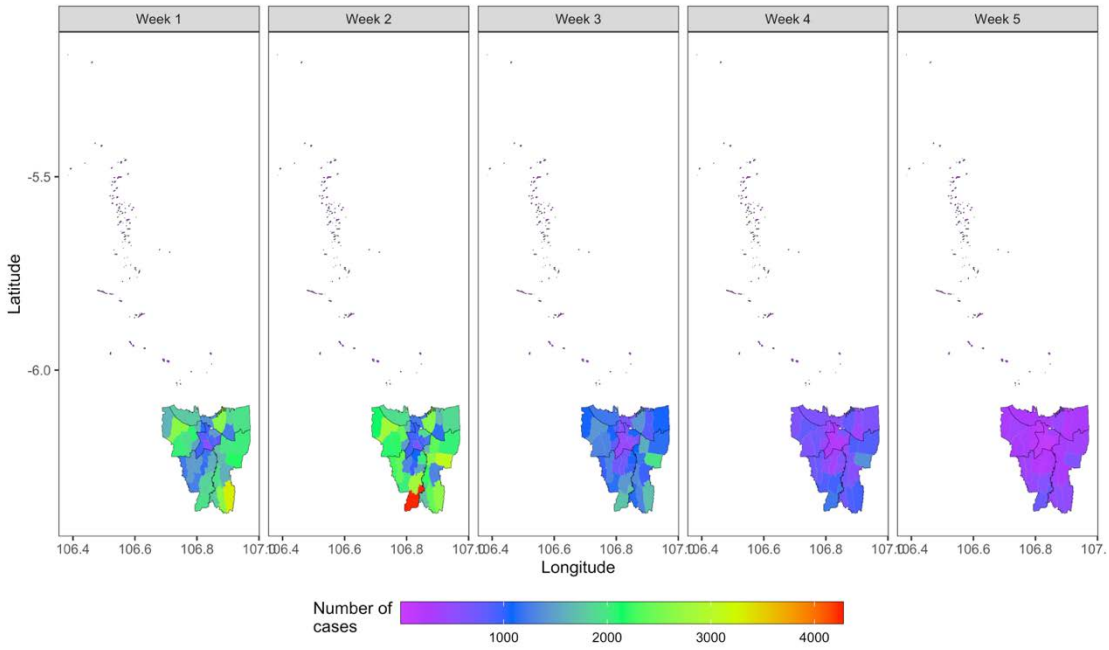


Figure 1. Mapped cases of COVID-19 from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021). These maps were created using R software

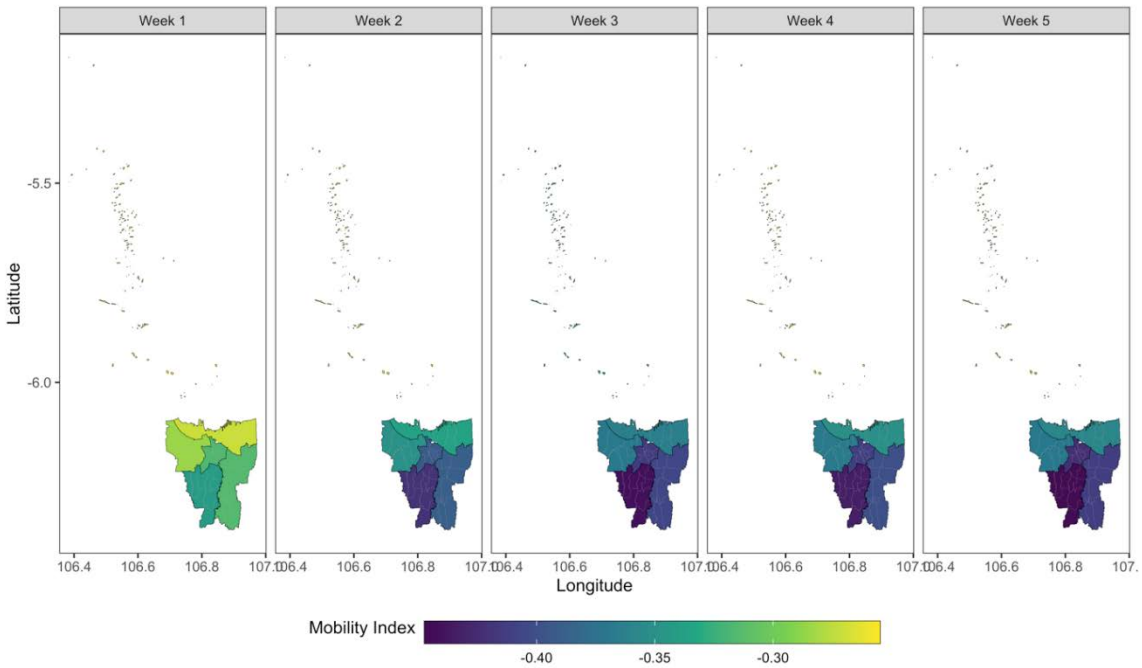


Figure 2. Mapped human mobility index from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021). These maps were created using R software

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

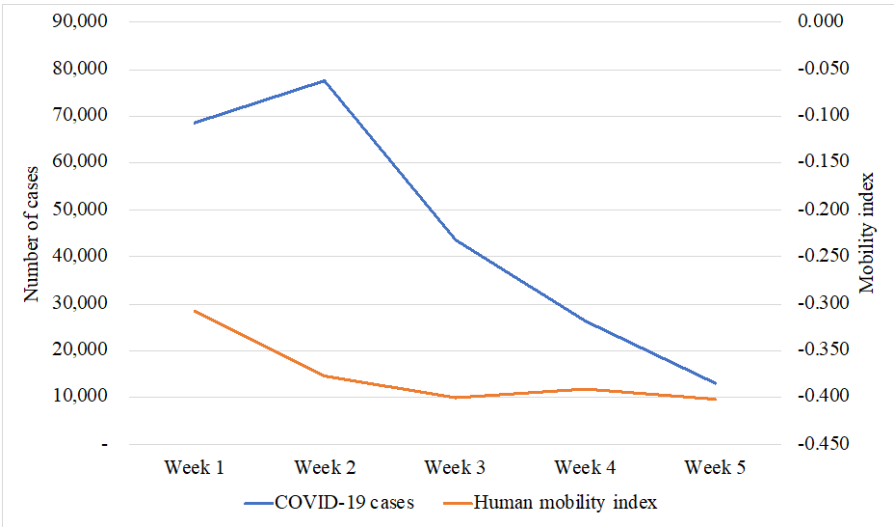


Figure 3. The temporal trend of number of cases COVID-19 and human mobility index

4.2. Model selection.

For model selection, we utilized the deviance information criteria (DIC), Watanabe akaike information criteria (WAIC), marginal predictive-likelihood (MPL), and determination coefficients (R2). Refer to [21] for a comprehensive analysis of these criteria. Three models were fitted, with Model 1 being the simplest and Model 3 being the most complex. The first model is a global regression with fixed regression coefficients. Model 2 emphasizes the varying coefficients of human mobility via exchangeable priors. Model 3 emphasizes the spatially varying regression effects of human mobility, which indicates epidemiological advantages. By transitioning from Model 1 to Model 3, the fit changed as the DIC values decreased from 98,274.33 to 2,283.98, WAIC values from 14,975.05 to 2,241.47 and MPL values increased from -49,165.41 to -1,543.64 and R^2 values increased from 0.0541 to 0.999 (See Table 1).

Table 1. Results of the three models comparison

Model	Specification	p_{DIC}	DIC	WAIC	MPL	R^2
Model 1	$\log\theta_{it} = \beta_0 + \beta_1 x_{it}$	36.22	98274.33	14975.05	-49165.41	0.0541
Model 2	$\log\theta_{it} = \alpha + v_t + \delta_{it} + (\beta + \varphi_i)x_{it}$; $\varphi_i \sim N(0, \sigma_\varphi^2)$	202.48	2283.50	2240.59	-1518.91	0.999
Model 3	$\log\theta_{it} = \alpha + v_t + \delta_{it} + (\beta + \varphi_i)x_{it}$; $\varphi_i \sim ICAR(w, \sigma_\varphi^2)$	202.56	2283.98	2241.47	-1543.64	0.999

Table 1 demonstrates that models 2 and 3 have comparable prediction performance, which is much superior to model 1. To compare the performance of models 2 and 3, we analyze the spatial autocorrelation of the index of human mobility over weeks. The results of the spatial autocorrelation test for human mobility indexes are displayed in Table 2.

Table 2. Results of the Moran's Index of human mobility for 5 weeks.

Parameter	Week 1	Week 2	Week 3	Week 4	Week 5
Moran's I	0.529	0.387	0.499	0.335	0.400

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

p-value	0.000	0.000	0.000	0.000	0.000

The spatial autocorrelation of the human mobility indexes for all weeks was significant, supporting model 3, which includes spatially structured coefficients with varying ICAR values. Models 2 and 3 improved model 1's fit as a result of its increased complexity, as the DIC and WAIC values were decreased and the MPL and R^2 values were increased. Models 2 and 3 exhibit comparable prediction performance based on all comparison criteria. However, model 3 highlights a fundamental gain in terms of the disease transmission control implications of comprehending the spatially varying regression effects of human mobility. Since the present analysis is concerned with the spatial interdependence that is used to account for the non-stationarity issue, we will focus on Model 3 in our subsequent analysis and discussions. Figure 4 demonstrates that the local regression Pearson's residuals of model 3 are less than the global regression residuals of model 1. This provides additional support for choosing the local regression models with spatially varying coefficients over the global regression models.

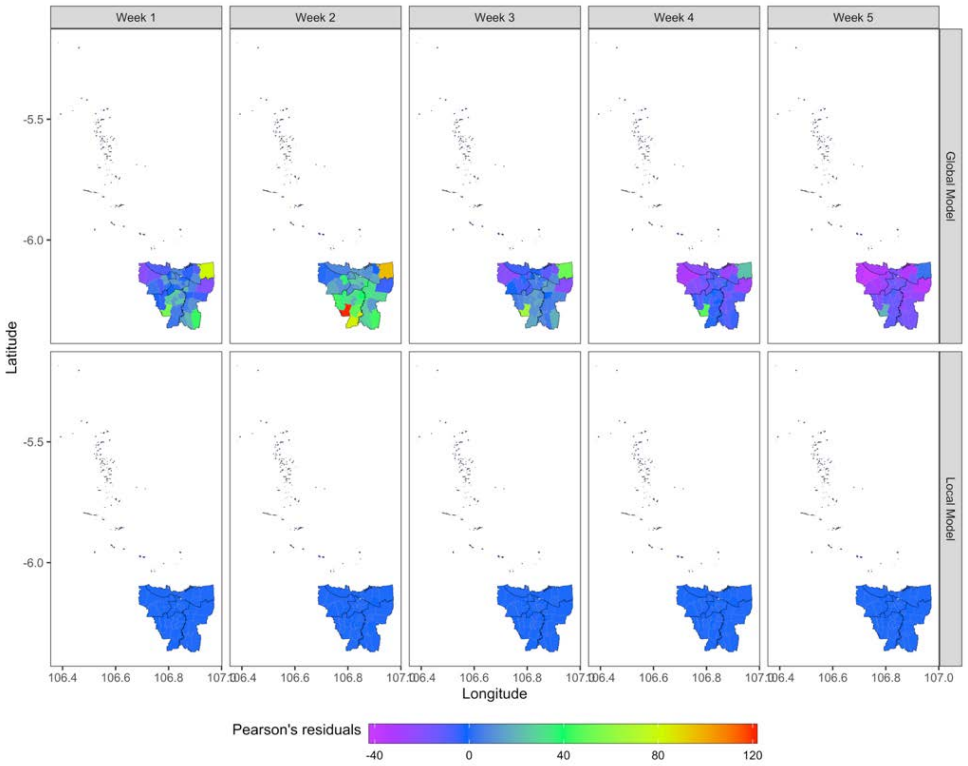


Figure 4. Pearson's residual global model (M1) versus local model (M3)

4.3. Estimation result

The estimation result from 3 different models are presented in Table 2.

Table 2. Results of the various models fitted

Parameters	Mean	Sd	$q(0.025)$	$q(0.975)$	Fraction Variance (%)
Model 1: Global model					
Fixed Effect					
Intercept (α)	1.591	0.016	1.559	1.622	
Slope (β)	4.281	0.044	4.195	4.367	

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

Parameters	Mean	Sd	$q(0.025)$	$q(0.975)$	Fraction Variance (%)
Model 2: Exchangeability varying coefficient Model					
Fixed Effect					
Intercept (α)	-0.991	0.604	-2.185	0.193	
Slope (β)	-2.247	1.634	-5.477	0.954	
Random effect					
Spatially varying coefficient ($\sigma^2_{\beta_i}$)	1.166	0.129	0.939	1.446	47.920
Temporally structured ($\sigma^2_{v_i}$)	1.066	0.569	0.394	2.565	43.820
Spatiotemporal interaction ($\sigma^2_{\delta_{it}}$)	0.201	0.013	0.178	0.227	8.260
Model 3: Spatially varying coefficient model					
Fixed Effect					
Intercept (α)	-0.050	0.985	-2.055	1.817	
Slope (β)	0.241	2.609	-5.072	5.189	
Random effect					
Spatially varying coefficient ($\sigma^2_{\beta_i}$)	5.596	0.072	3.609	8.717	86.128
Temporally structured ($\sigma^2_{v_i}$)	0.861	0.270	0.107	5.299	13.249
Spatiotemporal interaction ($\sigma^2_{\delta_{it}}$)	0.041	0.000	0.032	0.052	0.624

According to the three model specifications presented in Table 2, there are distinct fixed effects (slope) of human mobility on the COVID-19 risk. The regression slope for the global model (model 1) is 4.281, while it is -2.247 for the exchangeability varying coefficient model (model 2) and 0.241 for the spatially varying coefficient model (model 3). Models 1 and 3 demonstrate the positive impact of human mobility on the risk of COVID-19. It indicates that the number of COVID-19 will increase as human movement increases. Nonetheless, model 2 yields a peculiar conclusion. Mobility has a detrimental effect on the chance of contracting COVID-19. The outcome is inconsistent with reality. It could be because of the spatially ambiguous issue [29]. According to the result of the models' comparison, the discussion that follows will concentrate on model 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

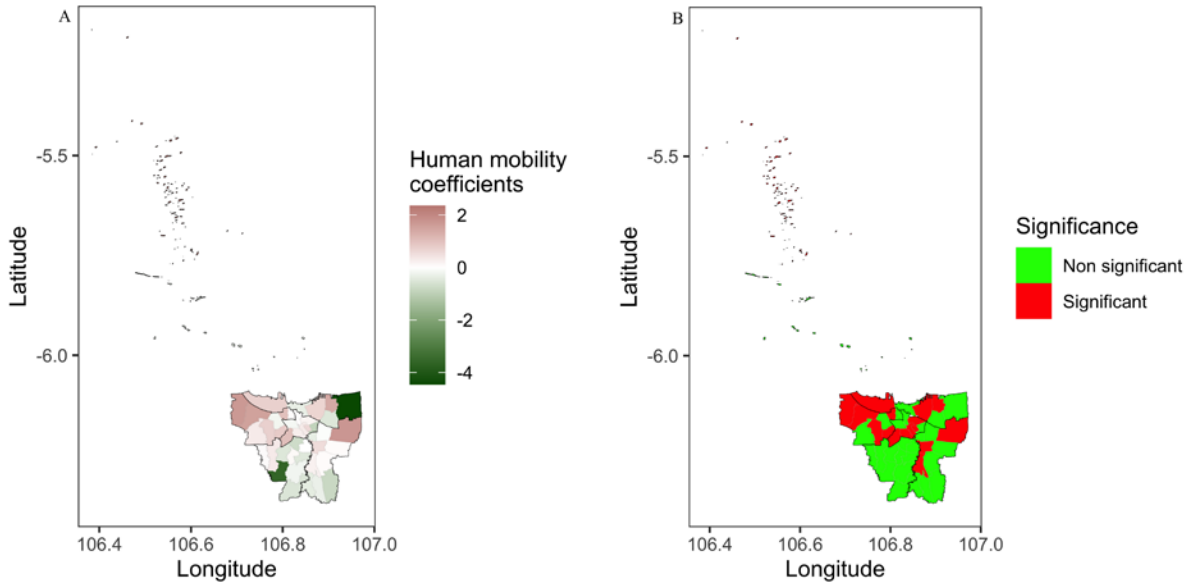


Figure 5. (A) Spatially varying coefficients and (B) Significance of the varying coefficient of human mobility

According to model 3, the spatial distribution of the regression coefficient of human mobility on COVID-19 risk differs. The impacts range between -4.455 and 2.353 (Figure 5A). West Jakarta and a few districts in North, Central, South, and East Jakarta had the most districts with a positive effect. Using hypothesis testing (H_0: $\beta_i = 0$ versus H_1: $\beta_i > 0$) we determined that the effects of human mobility were significant in some districts but not in others (Figure 5B). The spatially structured random effect component accounts for the majority of the COVID-19 risk's unexplained variation. The fractional variance is 86.128%.

The temporally structured random effect component accounts for 13.25% of the total variation of random effects, and from the first week PPKM was deployed, the relative risk of COVID-19 has steadily decreased (Fig 6).

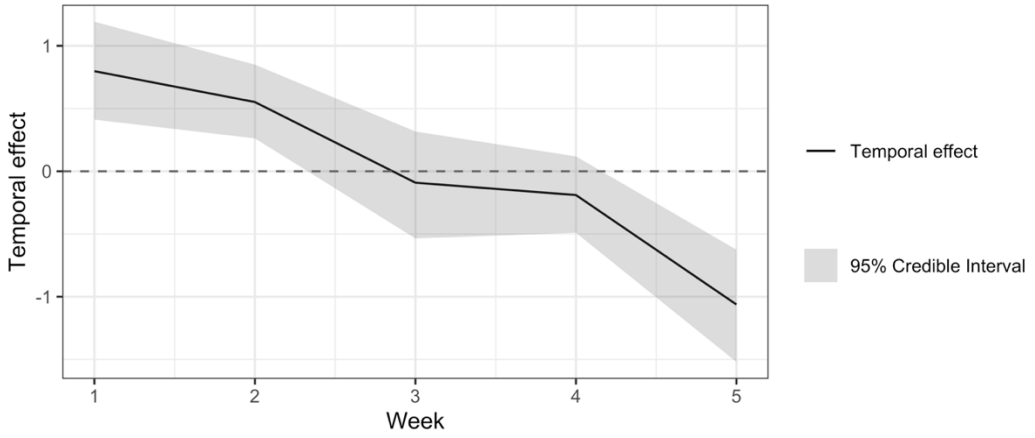


Figure 6. Structured temporal effect from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

The proportion of variance explained by spatiotemporal interaction relative to the total random effects is 0.64%, indicating that space and time interact less. Figure 7 depicts the plots of the spatiotemporal interaction components. Figure 7 demonstrates that nearly all districts share a similar temporal pattern, indicating the significant impact of the temporal trend.

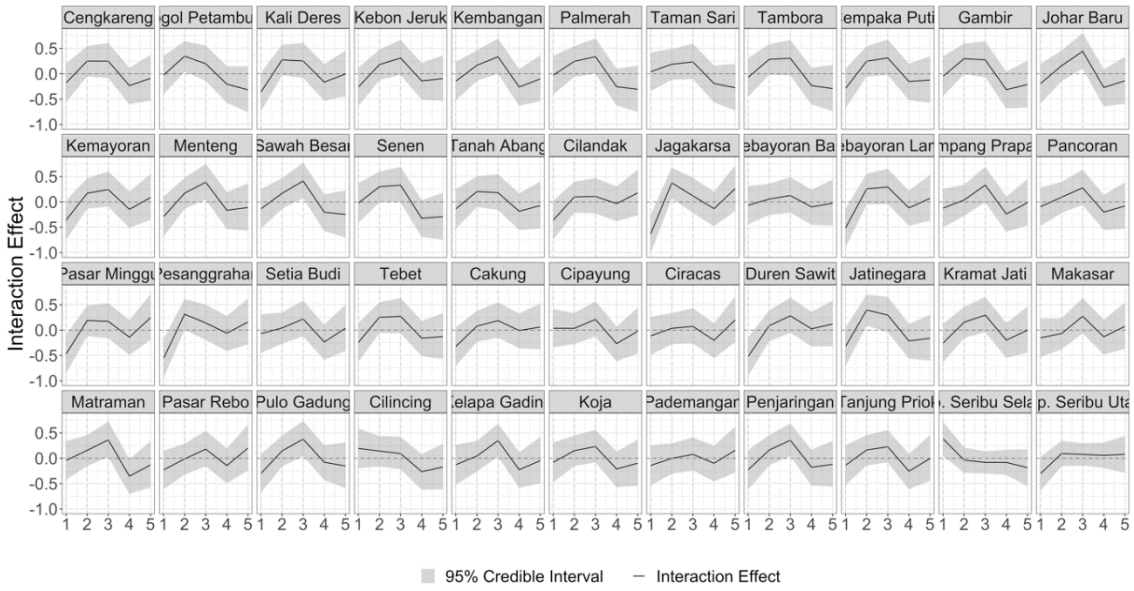
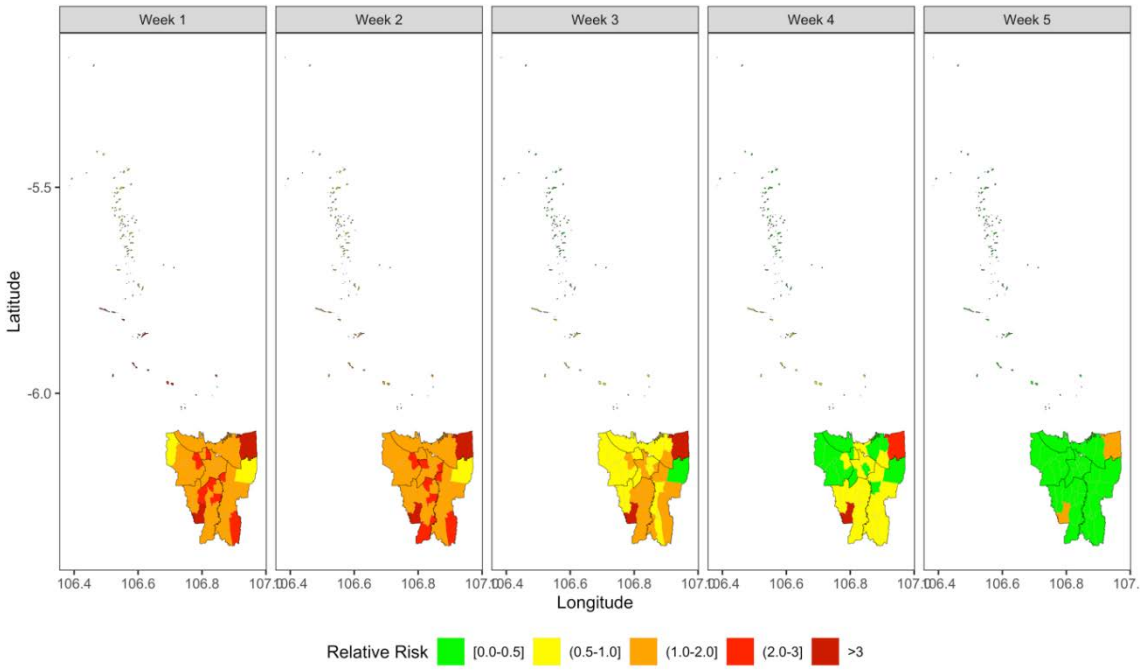


Figure 7. Spatiotemporal interaction effect for 44 districts from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021)

4.4. Relative risk estimate

We calculate the relative risk of COVID-19 for 44 districts from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021) using a spatially varying coefficients regression model (Figure 8) and calculate for the exceedance probability to identify the districts with substantial high risk over weeks (Fig 9)



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 September 2022 doi:10.20944/preprints202209.0271.v1

Figure 8. Mapped spatiotemporal distribution of the posterior means of the relative risk of COVID-19 from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021).

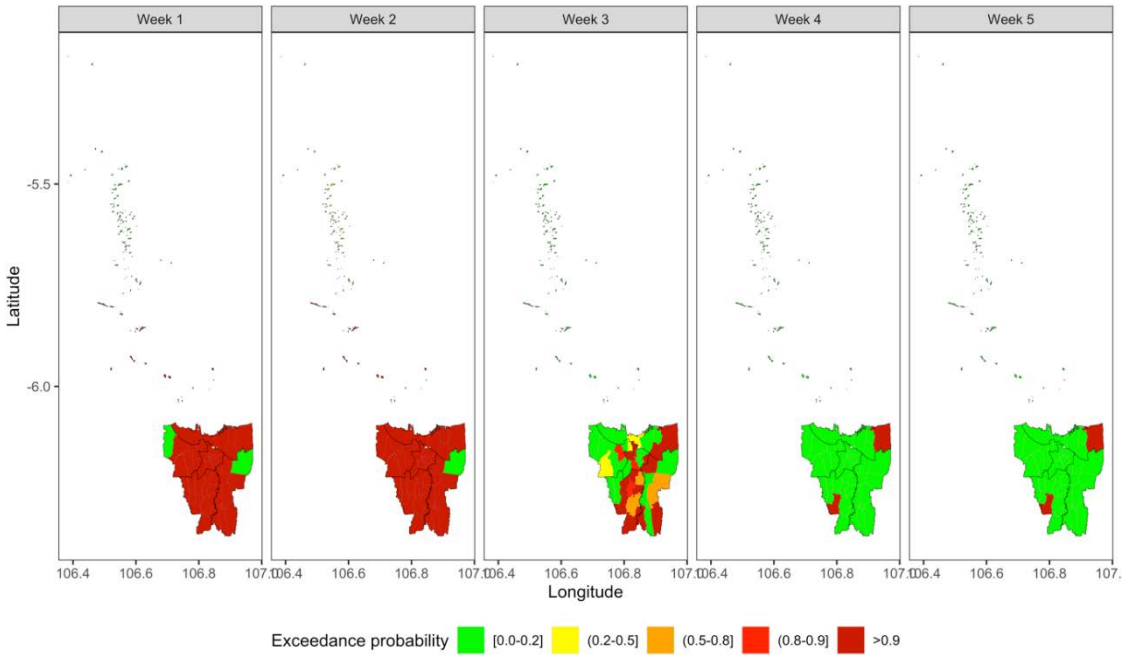


Figure 9. Mapped the spatiotemporal exceedance probabilities of $\Pr(\theta_{it} > 1|\mathbf{y})$, from Week 1 (3 – 9 July 2021) to Week 5 (31 July – 6 August 2021).

Figure 7 depicts the spatiotemporal distribution the relative risks θ_{it} from week 1 to week 5 when spatially varying coefficients, temporal effects, and space-time interactions are accounted for. We interpret these as relative risks based on a model. The relative risks were significantly clustered and steadily decreased beginning in the second week. Those areas with $\theta_{it} > 1$ have number of COVID-19 cases higher than expected cases, while those areas with $\theta_{it} < 1$ have lower than expected cases. A few particularly high- and low-risk districts appeared to form and gradually fade over time. Figure 8 shows the corresponding exceedance probabilities, $\Pr(\theta_{it} > 1|\mathbf{y})$. Dark red and red indicate regions with high probabilities, while yellow and green indicate regions with low probabilities. The patterns of districts with relative risks of $\theta_{it} >$ appear spatially continuous, as evidenced by the maps of probability of exceedance. For the first two weeks, exceeding threshold 1 indicated that more than 98% of the areas were identified as high-risk areas. After the second week, the relative risk decreased further. This is consistent with the temporal pattern that revealed a sharp decrease in risk from week 3 to week 5 (Figure 5).

5. Discussion

This paper illustrated the Bayesian Poisson spatiotemporal local regression model to evaluate the spatial heterogeneous effects of human mobility on COVID-19 risk transmission in 44 districts in Jakarta, Indonesia, as well as developed spatiotemporal choropleth maps of COVID-19 risk. Using spatially varying coefficients with the intrinsic conditional autoregressive (ICAR) prior, we were able to account for local variations in the effects of neighborhood human mobility. The most important implication of our findings is that human mobility is spatially continuous, so a global regression model is insufficient for

 doi:10.20944/preprints202209.0271.v1

quantifying human mobility effects. Usually, neighborhood disease morbidities are collected over distinct administrative areas such as districts, cities, or countries, which is incongruous with the transmission dynamics of infectious diseases such as COVID-19. Inaccurate reporting of diseases that cross neighborhood boundaries can result in spatial spillovers. Demographic variations also contribute to the non-stationary effects of the risk factors. Adopting spatially structured random effects components on the effects of the covariate using ICAR can account for the confounding variables. Models that attempt to account for spatially varying effects of covariates on the discrete outcome are extremely poorly understood, particularly in epidemiological research where discrete outcomes are common.

GWR estimates varying coefficients of risk factors using the weighted least square (WLS) estimation method to fit regression models. The Spatially varying coefficient (SVC) model considers regression coefficients as random components and employs a random effects approach to account for all relevant spatially confounding variables. This approach can simply be expanded to include spatially and temporally structured and unstructured random effects, as well as their interaction. Our empirical research on COVID-19 in Jakarta, Indonesia, demonstrates that local regression models by means of SVC are superior to global regression models in terms of fit and epidemiological significance. The results suggest that human mobility has spatially varying effects on COVID-19 risk. There was high variation in the local regression coefficients. Consequently, models with spatially varying coefficients can be beneficial for understanding the ecological importance of the various consequences of human mobility. Estimates of the local regression may have significant effects on the organization and evaluation of treatments.

Using the SVC model with ICAR prior, the district-specific relative risk is calculated. Since the third week, these choropleth maps (Figure 8) demonstrate a significant discrepancy between regions and a gradual fall. This phenomena could be explained by confounding variables accounted for by the spatially varying effects of human mobility.

District-specific treatments must take the relative importance of various targeted transmission paths into account. The model-based risk maps emphasize the significance of human mobility in specific locations for reducing the risk of COVID-19 transmission. We found spatially variable relationships between COVID-19 risk and human mobility.

Using exceedance probability, we discovered statistical evidence that the relative risk in the north east and south west is consistently high during PPKM. In these areas we found that the restricting human mobility does not have a significant effect on reducing disease transmission.

6. Conclusions

Our research contributes to the field of spatiotemporal epidemiology by illustrating the technical and empirical advantages of local regression model with SVC for assessing COVID-19 risk. In contrast to the global regression model, the SVC model had the extra benefit of highlighting the varying effects of the human mobility across areas. It has the practical implication of establishing a scientific foundation that allows precise intervention targeting of district-specific risk. Our research revealed that the association between COVID-19 risks and human movement is local. Our findings suggest that a reduction in human mobility could dramatically reduce transmission of COVID-19 in a few places, but would have no noticeable impact in other regions. The relative risk and exceedance probability maps provide a factual foundation for local medical planning and resource deployment. Moreover, the study provided a method for practitioners to quantify and map the relative COVID-19 risk across space and time. In addition, we offered a detailed methodology for modeling the effect of risk factors on disease risk in a heterogeneous population. It would be beneficial to do additional research to determine the geographically variable effects of climatic conditions on COVID-19. Sensitivity of spatially varying regression coefficients to complicated network structural dependencies and hyperprior distribution is a deserving area for further study.

 doi:10.20944/preprints202209.0271.v1

7. Research in Context

Evidence Before This Study. The majority of human mobility-related COVID-19 transmission reports are based on a global regression model. Few research have examined geographical heterogeneity, incorporating local regression models in particular.

Added Value of this Study. The analysis in this study is based on Facebook mobility data. The Facebook Data For Good Program makes aggregate mobility data accessible by utilizing the mobile apps' location service records of the user's GPS positions. Facebook's mobile application has a higher spatial resolution than Google's, and consequently gives more accurate results.

Implications of the Available Evidence. This exhaustive examination of the morbidity of the COVID-19 pandemic in various regions provides a clearer knowledge of the effect of human migration in places that are beneficial for combating the pandemic. In some regions, restricting human mobility may not be an appropriate method for fighting the epidemic.

Author Contributions

I.G.N.M.J. and A.C. collected and analyzed data and drafted the manuscript; G.D., F.K., and J.C. contributed to the data collection and confirmation; I.G.N.M.J., A.C., G.D., F.K., and J.C. contributed to the data analysis and manuscript draft; I.G.N.M.J., A.C., G.D., F.K All authors have read and approved the version of the paper that has been published.

Funding

This work was partially supported by funding from the Universitas Padjadjaran (DIKTI: 094/E5/PG.02.00.PT/2022 & DRPM1318/UN6.3.1/PT.00/2022)

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

All COVID-19 cases available online at https://riwayat-file-covid-19-dki-jakarta-jakartagis.hub.arcgis.com/ and Facebook mobility data available at https://dataforgood.facebook.com/

Acknowledgments

The views expressed in this article are the sole responsibility of the writers and do not necessarily reflect the decisions, policies, or opinions of any institution or organization represented by the authors.

Conflicts of Interest

We declare no competing interest.

References

1. Karcıoğlu, O.; Yüksel, A.; Baha, A.; Er B. A.; Esendağlı, D.; Gülhan, P. Y.; Karaoğlanoğlu, S.; Erçelik, M.; Şerifoğlu, İ. Covid-19: The biggest threat of the 21st century: In respectful memory of the warriors all over the world. *Turk. Thorac. J.* ***2020****, 21,* 409-418.

2. WHO Coronavirus (COVID-19) dashboard. Available online: https://covid19.who.int/. (accessed on 25 August 2022).
3. Zhang, M.; Wang, S.; Hu, T.; Fu, X.; Wang, X.; Halloran, B.; Li, Z.; Cui, Y.; Liu, H.; Liu Z.; Bao, S. Human mobility and COVID-19 transmission: a systematic review and future directions. *Annals of GIS* **2022**, 1-14.
4. Sperrin, M.; McMillan, B. Prediction models for covid-19 outcomes. *BMJ* **2020**, *371*, m3777.
5. Kraemer, M.U.G.; Yang, C.H.; Gutierrez, B.; Wu, C.H.; Klein, B.; Pigott, D.M. Open COVID-19 Data Working Group; Plessi, L.D. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **2020**, *368*, 493-497.
6. Kucharski, A.J.; Russell, T.W.; Diamo, C.; Liu, Y.; Edmunds, J.; Funk, S.; Eggo, R.M. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **2020**, *20*, 553–558.
7. Wang, S.; Liu, Y.; Hu, T. Examining the change of human mobility adherent to social restriction policies and its effect on COVID-19 cases in Australia. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7930.
8. Yuan, Z.; Xiao, Y.; Dai, Z.; Huang, J.; Zhang, Z.; Chen, Y. Modelling the effects of Wuhan's lockdown during COVID-19, China. *Bull. World Health Organ*. **2020**, *98*, 84–494.
9. Yabe, T.; Tsubouchi, K.; Fujiwara, N.; Wada, T.; Sekimoto, Y.; Ukkusur, S.V. Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. *Sci. Rep.* **2020**, *10*, 18053.
10. Hou, X.; Gao, S.; Li, Q.; Kang, Y.; Chen, N.; Chen, K.; Rao, J.; Ellenberg, J.S.; Patz, J. A. Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race. *PNAS* **2021**, *118*, e2020524118.
11. Vicente, G.; Goicoa, T.; Ugarte, M. Bayesian inference in multivariate spatio-temporal areal models using INLA: analysis of gender-based violence in small areas. *Stoch. Environ. Res. Risk Assess*. **2020**, *34*, 1421–1440.
12. Lome-Hurtado, A.; Lartigue-Mendoza, J.; Trujillo, J.C. Modelling local patterns of child mortality risk: a Bayesian Spatio-temporal analysis. *BMC Public Health* **2021**, *21*, 1-12.
13. Bauer, C.; Wakefield, J.; Rue, H.; Self, S.; Feng, Z.; Wang, Y. Bayesian penalized spline models for the analysis of spatiotemporal count data. Stat. Med. **2016**, 35, 1848–1865.
14. Osei, F.; Stein, A. Diarrhea morbidities in small areas: Accounting for non-stationarity in sociodemographic impacts using Bayesian spatially varying coefficient modelling. *Sci. Rep.* **2017**, *7*, 1-15.
15. Wakefield, J. Disease mapping and spatial regression with count data. *Biostatistics* **2007**, *8*, 158-183.
16. Jaya, I.G.N.M. Folmer, H. Identifying spatiotemporal clusters by means of agglomerative hierarchical clustering and Bayesian regression analysis with spatiotemporally varying coefficients: methodology and application to dengue disease in Bandung, Indonesia. *Geogr. Anal.* **2020**, 1-57.
17. Fiebig, D.G.; Bartels, R.; Aigner, D. J. A random coefficient approach to the estimation of residential end-use load profiles. *J. Econom.* **1991**, *50*, 297-327.
18. Congdon, P. Spatial heterogeneity in Bayesian disease mapping. *GeoJournal* **2018**, *1*, 1-14.
19. Finley, A.O. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. *Methods Ecol. Evol.* **2011**, 2, 143–154.
20. Jaya, I.G.N.M. Folmer, H. Ruchjana, B.N.; Kristiani, F.; Yudhie, A. Modeling of infectious diseases: A core research topic for the next hundred years. In Regional Research Frontiers: Methodological Advances, Regional Systems Modeling and Open Sciences; Jackson, R., Schaeffer, P., Eds.; Springer International Publishing: USA, 2017; 2 pp. 239-254.
21. Jaya, I.G.N.M. Folmer, H. Bayesian spatiotemporal mapping of relative dengue disease risk in Bandung, Indonesia. *J. Geogr. Syst.* **2020**, *22*, 105-142.

 doi:10.20944/preprints202209.0271.v1

22. Sparks, C. An examination of disparities in cancer incidence in Texas using Bayesian random coefficient models. *PeerJ* ***2015***, *3*, e1283.
23. Jaya, I.G.N.M. Folmer, H. Bayesian spatiotemporal forecasting and mapping of COVID-19 risk with application to West Java Province, Indonesia. *J. Reg. Sci.* ***2021***, *61*, 849-881.
24. Shepherd, H.E.; Atherden, F.S.; Chan, H.M.T.; Loveridge, A.; Tatem, A.J. Domestic and international mobility trends in the United Kingdom during the COVID-19 pandemic: an analysis of facebook data. *Int. J. Health Geogr.* ***2021***, 20, 1-13.
25. Facebook Data for Good: Disease Prevention Maps. Available on https://dataforgood.facebook.com/ (accessed 21 August 2022)
26. Rue, H.; Riebler, A.; Sørbye, S.H.; Illian, J.B.; Simpson, D.P.; Lindgren, F.K. Bayesian computing with INLA: A review. *Annu. Rev. Stat. Appl.* ***2017***, *4*, 395–421.
27. Rue, H.; Martino, S.; Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc.* ***2009***, *7*, 319–392.
28. Knorr-Held, L. Bayesian modelling of inseparable space-time variation in disease risk. *Stat Med* ***2000***, *19*, 15-30.
29. Adin, A.; Goicoa, T.; Hodges, J.S.; Schnell, P.M.; Ugarte, M.D. Alleviating confounding in spatio-temporal areal models with an application on crimes against women in India. *Stat. Modelling* ***2022***, 1-22