Geospatial Health 2023; volume 18:1161

Does mobility restriction significantly control infectious disease transmission? Accounting for non-stationarity in the impact of COVID-19 based on Bayesian spatially varying coefficient models

I Gede Nyoman Mindra Jaya,[1] Anna Chadidjah,[1] Farah Kristiani,[2] Gumgum Darmawan,[1] Jane Christine Princidy[1]

[1]Statistics Department, Universitas Padjadjaran, Bandung; [2]Mathematics Department, Parahyangan Catholic University, Bandung, Indonesia

Abstract

Coronavirus disease (COVID-19) remains a worldwide threat. Restriction of human mobility is one of the strategies used to control the transmission. But how effective this restriction is, particularly in small areas, has yet to be determined. Using Facebook's mobility data (Facebook Data for Good, 2022), we explored the impact of restricting human mobility on COVID-19 cases in several small districts in Jakarta, Indonesia. This was done by modifying a global regression model into a local regression approach 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 to account for non-stationarity regarding human mobility. We estimated the regression parameters using an integrated nested Laplace approximation and found that the local regression model with spatially varying regression coefficients outperforms global regression based on Deviance Information Criteria (DIC), Watanabe Akaike Information Criteria (WAIC), the Marginal Predictive-Likelihood (MPL) and the coefficient of determination (R^2) criteria for model selection. In Jakarta's 44 districts, the impact of human mobility varies significantly; it was found to range from –4.445 to 2.353 on the log relative risk of COVID-19. We propose a cost-effective strategy as the preventive restriction strategy was found to be beneficial in some districts but ineffective in others. Our main contribution by this research was to show how the restriction of human mobility data can give important information about the transmission of COVID-19 in different small areas.

Correspondence: I Gede Nyoman Mindra Jaya, Department of Statistics, Padjadjaran University, West Java, Indonesia, Kampus Unpad, Jalan Ir.Soekarno Km. 21 Jatinangor, Kabupaten Sumedang, Jawa Barat 45363, Indonesia.
Tel.: +62.227.796002 - Fax: +62.227.796002.
E-mail: mindra@unpad.ac.id

Key words: COVID-19; human mobility; spatial and temporal autocorrelations; Facebook mobility data; Indonesia.

Acknowledgments: this work was partially supported by funding from the Universitas Padjadjaran (DIKTI: 094/E5/PG.02.00.PT/2022 & DRPM: 1318/UN6.3.1/PT.00/2022).

Conflict of interest: the Authors declare no conflict of interest.

Received for publication: 5 October 2022.
Revision received: 29 November 2022.
Accepted for publication: 29 November 2022.

Geospatial Health 2023; 18:1161
doi:10.4081/gh.2023.1161

Introduction

The coronavirus disease (COVID-19) has been deemed the greatest health threat of the 21st century (Karcıoğlu *et al.,* 2020). According to the World Health Organization (WHO), 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 (WHO, 2022). In the absence of widely available vaccination, as the situation were until the beginning of 2021, restricted mobility was arguably the most effective method for preventing viral spread. It goes without question that human mobility is essential for the transmission of infectious diseases and Zhang *et al*., (2022) has reviewed systematically with emphasis on statistical method. Due to modern transportation networks and growing globalization, it took COVID-19 less than four months to become a pandemic (WHO, 2020)

Accurate models that anticipate the transmission of COVID-19, are necessary to support population-level intervention decisions (Sperrin and McMillan, 2020). The accurate COVID-19 risk estimate could be determined by examining human mobility. Numerous investigations have shown that restricted mobility is successful in reducing the transmission of COVID-19 (Hou *et al.,* 2021; Kraemer *et al.,* 2020; Kucharski *et al.,* 2020; Wang *et al.,* 2020; Yabe *et al.,* 2020; Yuan *et al.,* 2020). However, due to the various regional characteristics, human mobility does not impact COVID-19 transmission uniformly everywhere (Firza & Monaco, 2022). Disease risk mapping can contribute to a better understanding of risk evolution over space and time (Bauer *et al.,* 2020; Lome-Hurtado *et al.,* 2021; Vicente *et al.,* 2022). Ecological regression models with global (fixed) regression coefficients are frequently used to identify relevant risk factors to provide precise risk estimates and covariate effects (Wakefield, 2007; Osei & Stein, 2017). These global regression models commonly rely on assumptions regarding homogeneity of the population and fixed covariate regression coefficients. However, these parameters are not likely constant over time and the socioeconomic conditions and environmental variables that may influence local human behaviour that contribute to the transmission of COVID-19 vary strongly between different populations (Firza & Monaco, 2022). Given this potential complexity, it is logical that the impact of human mobility factors be non-stationary. For example, the effect of restricting human mobility can have different effects that can even oppose each other; while mobility restrictions reduce COVID-19 transmission, the resulting reduction in economic growth may increase stress levels and make people more vulnerable to COVID-19. However, location-dependent variation in the mobility-related effects can be permitted and thus achieve a modulate COVID-19 transmission level. On the other hand, considering the variance instability caused by unobserved risk factors in different populations as discussed by Osei and Stein (2017), it is possible that spatial interdependency caused by similar or identical risk factor conditions improves the accuracy of disease risk prediction. Our research offers an alternative to the global regression models for disease modelling and mapping that presume that regression coefficients are constant across neighbourhoods. The Spatially Varying Coefficient (SVC) and Geographically Weighted Regression (GWR) models (Osei & Stein, 2017) can accommodate the non-stationary effects of the covariates. GWR is focused on continuous response variables, which are, however, rarely employed in disease mapping (Jaya & Folmer, 2020). 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, Fiebig *et al.,* 1991 or Congdon, 2018). The advantage of SVC, according to Finley (2011), is that it provides a probability model from which all model parameters can be estimated, and relative risk estimates 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 (Jaya *et al.* 2017; Jaya & Folmer 2020). Thus, SVC models are appropriate for correctly describing regionally variable health issues (Sparks, 2015). Bayesian smoothing is modelled 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, where the number of COVID-19 cases is typically aggregated by administrative area (Jaya & Folmer, 2021a). 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, the capital of Indonesia, where the implementation of Community Activity Restrictions (locally named PPKM) aims to constrain s people's movement 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 (Shepherd *et al.,* 2021). Here, we examined aggregated and anonymized Facebook data on the mobility patterns of active users in Jakarta utilizing geolocation services between 3 July 2021 and 6 August 2021. The current study was undertaken to explain the methodological and substantive epidemiological implications that can be used to develop an Early Warning System (EWS) for future disease outbreaks.

Materials and Methods

Using a hierarchical Bayesian framework, we evaluated many spatiotemporal Poisson regression models that consider spatiotemporal dependence and heterogeneity. Simulations of numerically evaluating complicated integrals with Markov Chain Monte Carlo (MCMC) can be challenging and aggravating (Rue *et al.,* 2017, 2009). For approximate Bayesian inference, we utilized the Integrated Nested Laplace Approximation (INLA). The study had two goals: i) analyzing the spatially varying effects of human mobility on COVID-19 risk, and ii) mapping the relative risk estimates over space and time.

Study area and data

Jakarta is the most populous city in Southeast Asia and the capital of Indonesia comprising 44 districts. It is located on Java and had an estimated population of 10,562,080 in 2020 (Figure 1). Jakarta's metropolitan area covers 9,957.08 km^2 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 and a strong inflow of people because it provides business opportunities 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 employed 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 3July 2021 to 6 August 2021, or during the enforcement of restrictions on community activities period. These data were gathered from the Jakarta COVID-19 history file website (https://riwayat-file-covid-19-dki-jakarta-jakartagis.hub.arcgis.com/) and the Data-for-Good Facebook website (https://dataforgood.facebook.com/).Facebook mobility data reveal the locations, movements, and connections of active Facebook users. The data are 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 (Shepherd *et al.,* 2021).

Statistical analysis

Model specification

We considered y_{it} as representing the spatiotemporal outcomes of COVID-19 infections and the population at risk aggregated by districts i=1, ..., n and time t=1, …, T. The term y_{it} expresses count data following a Poisson distribution, that is, $y_{it} \sim Poisson\ (E_{it}\, q_{it})$ with the likelihood function according to Jaya and Folmer (2020, 2021b) that is calculated by equation 1:

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

where E_{it} and q_{it} denote the expected number of cases and the relative risk in area and at time 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, the standardized incidence ratio (SIR) is the ratio of observed cases to the expected number of cases at each place and time: $SIR_{it} = y_{it}/E_{it}$, with the last term as described by Jaya and Folmer (2020, 2021a, 2021b) given by equation 2:

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

Note that population heterogeneity influences the estimation of the SIR value. As a result, the fixed effect parameter of small areas with a small population size N_{it} generally have a high degree of variability. High areal variability due to population heterogeneity is usually overcome by imposing an independent Gaussian (exchangeable) prior distribution on the log-relative risk, *i.e.*, log $\theta_{it} \sim N(\alpha,\ \sigma_u^2)$ which results in a log-linear Poisson regression model with a random intercept. Thus log $\theta_{it} = \alpha + u_i$, where α is the global logarithmic level of the relative risk and u_i the spatially unstructured random effects, which implies that $\theta_{it} = \exp(\alpha + u_i)$, where $u_i \sim N(0, \sigma_u^2)$ and $p(u|\sigma_u^2) \propto \sigma_u^{-n} \exp(-0.5\sigma_u^{-2}\Sigma_{i=1}^{n} u_i^2)$. It is likely that the relative risks in several adjacent regions would reflect a geographical pattern (Osei and Stein, 2017).

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 its probability distribution conditional on the set $\omega_{-i} = \{\omega_1, ..., \omega_{i\neq 1}, \omega_{i+1}, \omega_n\}$. The Intrinsic Conditional Autoregressive (ICAR) prior is a frequently used method for representing irregularly shaped regions where, as discussed by Osei and Stein (2017) and Jaya and Folmer (2020), the conditional distribution of is given by equation 3:

$$\omega_i|\boldsymbol{\omega}_{-i} \sim N\left(\overline{\omega}_i, \frac{\sigma_\omega^2}{w_{i+}}\right) \quad \text{(Eq. 3)}$$

where $\varpi_i = \Sigma_j\, \omega_{ij}\omega_j /\ \omega_{i+}$ and ω_{i+} with ω_{ij} denoting $n \times n$ binary spatial weight matrices that express the spatial dependency structure with $\omega_{ij} = 1$ if i and j are neighbours (i~j) and otherwise 0. The given specification (Eq. 3) results in the joint distribution for vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$' as discussed by Osei and Stein (2017):

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

This prior is unusual as it requires that the sum to zero constraint $\Sigma_{i=1}\ \omega_i = 0$ to ensure identifiability. To avoid the challenge of selecting between spatially structured and unstructured effects, it is possible to combine these two priors as done by Osei and Stein (2017):

$$\log\theta_{it} = \alpha + (u_i + \omega_i) \quad \text{(Eq. 5)}$$

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, Eq. 5 can be modified to become:

$$\log\theta_{it} = \alpha + u_i + \omega_i + \varsigma_t + v_t + \delta_{it} \quad \text{(Eq. 6)}$$

Figure 1. Jakarta is located on Java Island of Indonesia.

where φ_t and n_t denote the temporally unstructured and structured effects, respectively. The temporally unstructured effects are usually modeled as exchangeable prior distribution (*i.e.*, $\varphi_t \sim N(0, \sigma_\varphi^2)$). We considered a random walk model of order 1 (RW1) to model temporal autocorrelation. The Gaussian vector $\nu = (\nu_1, \ldots, \nu_T)'$ is defined by assuming independent differencing (Osei and Stein, 2017):

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

The density function for ν is obtained from its $T - 1$ increments as follows:

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

The spatiotemporal interaction effects component δ_{it} can be specified in four different types (Knorr-Held, 2020). 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 type IV is the interaction between temporally and spatially structured effects. In addition, the Eq. 6 can be modified to account for the effect of risk factor x_{it} as follows:

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

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 variables can be continuous or discrete. To account for spatial heterogeneity with respect to 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 as:

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

where φ_t represents differential spatially varying regression effects accounting 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 the ICAR process $\varphi_i \sim ICAR(w, \sigma_\varphi^2)$ or exchangeable Gaussian processes $\varphi_i \sim N(0, \sigma_\varphi^2)$. Some random-effect components in Eq. 10 should be eliminated to avoid the overfitting and confounding issues.

Bayesian inference

Let $\Phi = \{\alpha, \beta, \mu, \omega, \varsigma, \nu, \delta, \varphi\}$ denotes the vector Gaussian latent (unobservable) field and $\Psi = \{\sigma_\upsilon^2, \sigma_\omega^2, \sigma_\varsigma^2, \sigma_\nu^2, \sigma_\delta^2, \sigma_\varphi^2\}$ 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 $\boldsymbol{\Phi}_i \perp \boldsymbol{\Phi}_j | \boldsymbol{\Phi}_{-ij}$. The Bayesian inference is introduced in three stages as follows:

$y\|\Phi, \Psi \sim p(y\|\Phi, \Psi)$	Stage 1
$\Phi\|\Psi \sim p(\Psi\|\Phi)$	Stage 2
$\Psi \sim p(\Psi)$	Stage 3

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

$$p(\Phi, \Psi|y) = \frac{p(\Phi, \Psi, y)}{p(y)} = \frac{p(y|\Phi, \Psi) p(\Phi|\Psi) p(\Psi)}{\int_\Phi \int_\Psi p(y|\Phi, \Psi) p(\Phi|\Psi) p(\Psi) d\Phi d\Psi} \quad \text{(Eq. 11)}$$

The joint posterior distribution can be expressed as $p(\Phi, \Psi|y) \propto p(y|\Phi, \Psi)\, p(\Phi|\Psi)\, p(\Psi)$ 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 the INLA approach, in which the complex integral can be numerically computed through a faster computation process compared to MCMC simulation. The INLA calculating procedure may be summed up as suggested by Osei and Stein (2017) and folled up by Jaya and Folmer (2020):

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

$$p(\Psi|y) = \frac{p(\Phi, \Psi|y)}{p(\Phi|\Psi, y)} \approx \left.\frac{p(y|\Phi, \Psi) p(\Phi|\Psi), p(\Psi)}{p(\Phi|\Psi, y)}\right|_{\Phi = \Phi^*(\Psi)} = \tilde{p}(\Psi|y) \quad \text{(Eq. 12)}$$

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

$$\tilde{p}(\Phi_i|\Psi, y) = \left.\frac{p(\Phi, \Psi|y)}{\tilde{p}(\Phi_{-i}|\Phi_i, \Psi, y)}\right|_{\Phi_{-i} = \Phi^*_{-i}(\Phi_i, \Psi)} \quad \text{(Eq. 13)}$$

where $\tilde{p}(\Phi_{-i}|\Phi_i, \Psi, y)$ is the Laplace-Gaussian approximation to $p(\Phi_{-i}|\Phi_i, \Psi, y)$ and $\Phi^*{-i}(\Phi_i, \Psi)$ is its mode. Finally, the marginal posterior distributions are computed as $\tilde{p}(\Phi_i|y) \approx \int \tilde{p}(\Phi_{-i}|\Phi_i, \Psi, y)\, \tilde{p}(\Psi|y) d\Psi$ according to Osei and Stein (2017) with the marginal posterior distribution 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 an area has a notably high risk. More details can be seen in the papers by Jaya and Folmer (2020, 2021b).

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 $\omega_i \sim ICAR(w, \sigma_u^2)$. Due to the limited temporal range of the data, we defined a RW1 prior $\nu_t = \nu_{t-1} + \Delta\nu_t, \Delta\nu_t \sim N(0, \sigma_u^2)$. We picked interaction type IV, *i.e.*, $\delta_{it} \sim N(0, \sigma_u^2)$, to account for spatiotemporal interaction because ω_i and ν_t capture the spatially and temporally structured variations.

Global model fixed effect regression — Model 1

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

SWC with exchangeable prior — Model 2

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

SVC with ICAR prior — Model 3

$$\log\theta_{it} = \alpha + \upsilon_t + \delta_{it} + (\beta + \varphi_{it})\, x_{it};\ \varphi_i \sim ICAR(w, \sigma\psi_\psi^2)$$

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 *I*. To complete the Bayesian inference, we turned to a Gaussian prior distribution with a zero mean and a huge variance for α and β: $\{\alpha,\beta\}\sim N(0, 10^6)$. In addition, the variance hyperparameters $\Psi = \{\sigma_v^2, \sigma_\delta^2, \sigma_\varphi^2\}$ require priors (hyperpriors). We utilized the proper half-Cauchy distribution with a scale parameter of 25 (Jaya & Folmer 2021b).

Results

Distribution of COVID-19 cases and human mobility

Figure 2A 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

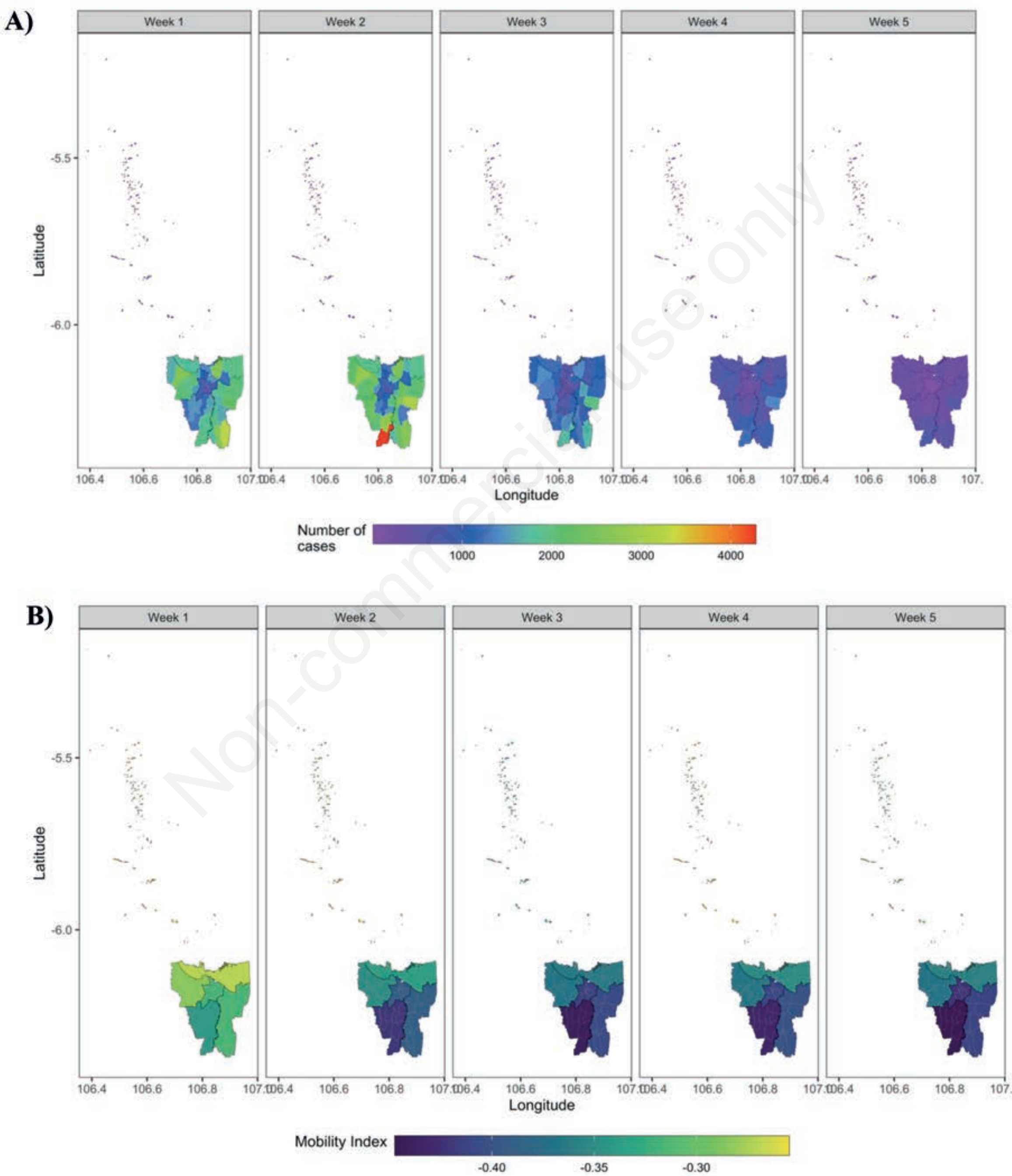

Figure 2. A) The number of COVID-19 cases in 2021 from first week in July (week 1) to first week in August (week 5). B) The human mobility index in 2021 from first week in July (week 1) to first week in August (week 5). Maps created using R software.

in the fifth. The maps of human mobility indexes are shown in Figure 2B. The human mobility index fluctuated between -0.341 and -0.255 during week 1, between -0.420 and -0.264 during week 2, between -0.441 and -0.334 during week 4, between -0.433 and -0.263 during week 4, and between -0.447 and -0.285 during week 5. As a large negative index indicates limited human mobility, it is obvious that the number of cases declines as the human mobility index falls, which is clear from Figure 3.

Model selection

For model selection, we utilized the deviance information criteria (DIC), Watanabe Akaike information criteria (WAIC), the marginal predictive likelihood (MPL) and the coefficient of determination (R^2). We refer to Jaya and Folmer (2020) for a comprehensive discussion of these criteria. Three models were fitted, with Model 1 being the simplest and Model 3 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 and 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 and the WAIC ones from 14,975.05 to 2,241.47, while the MPL increased from -49,165.41 to -1,543.64 and R^2 from 0.0541 to 0.999 (Table 1). According to the model selection result, we considered the Model 3 to be the best model.

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 analyzed the spatial autocorrelation of Moran's *I* of human mobility over the five weeks. The results are displayed in Table 2.

Estimation results

The spatial autocorrelation of the human mobility indices for all weeks was significant, supporting Model 3, which includeed spatially structured coefficients with varying ICAR values. As the

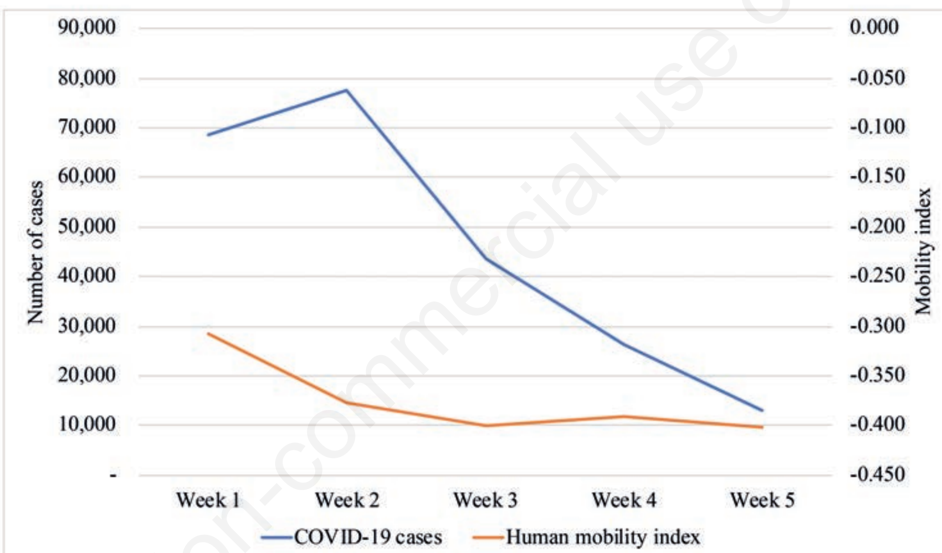

Figure 3. The temporal trend of the number of cases COVID-19 along the human mobility index.

Table 1. Summary of model 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	$\log \theta_{it} = \alpha + \nu_t + \delta_{it};\ \varphi_i \sim N(0, \sigma\varphi^2)$	202.48	2283.50	2240.59	-1518.91	0.999
Model 3	$\log \theta_{it} = \alpha + \nu_t + \delta_{it} + (\beta + \varphi_i)\, x_{it};\ \varphi_i \sim \text{ICAR}\ (\omega, \sigma_\psi^2)$	202.56	2283.98	2241.47	-1543.64	0.999

DIC, deviance information criteria; WAIC, Watanabe Akaike information criteria; MPL, marginal predictive-likelihood; R^2, coefficient of determination.

Table 2. Changing of Moran's I of human mobility over 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
p-value	0.000	0.000	0.000	0.000	0.000

DIC and WAIC values decreased and the MPL and R^2 values increased, Models 2 and 3 clearly improved Model 1's fit, but the cost of increased complexity. 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 was concerned with spatial interdependence used to account for the non-stationarity issue, we focused 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.

According to the three model specifications presented in Table 3, there are distinct fixed effects (slope) of human mobility on the COVID-19 risk. The regression slope for the global model (Model 1) was 4.281, while it was –2.247 for the exchangeability varying coefficient model (Model 2) and 0.241 for the SVC model (Model 3). Models 1 and 3 demonstrate the impact indicating that the number of COVID-19 increases with human movement. Nonetheless, Model 2 yields a peculiar conclusion: As increased mobility has a detrimental effect on the chance of contracting COVID-19, the outcome is inconsistent with reality. This could be explained by the spatially ambiguous issue (Adin *et al*., 2022). 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 each of the North, Central, South, and East parts of the city had showed a positive effect (increased risk). 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 accounted for most of the COVID-19 risk's unexplained variation. The fractional variance was 86.13%. The temporally structured random effect component accounts for 13.25% of the total variation of random effects, and from the first week that the PPKM programme was deployed, the relative risk of COVID-19 steadily

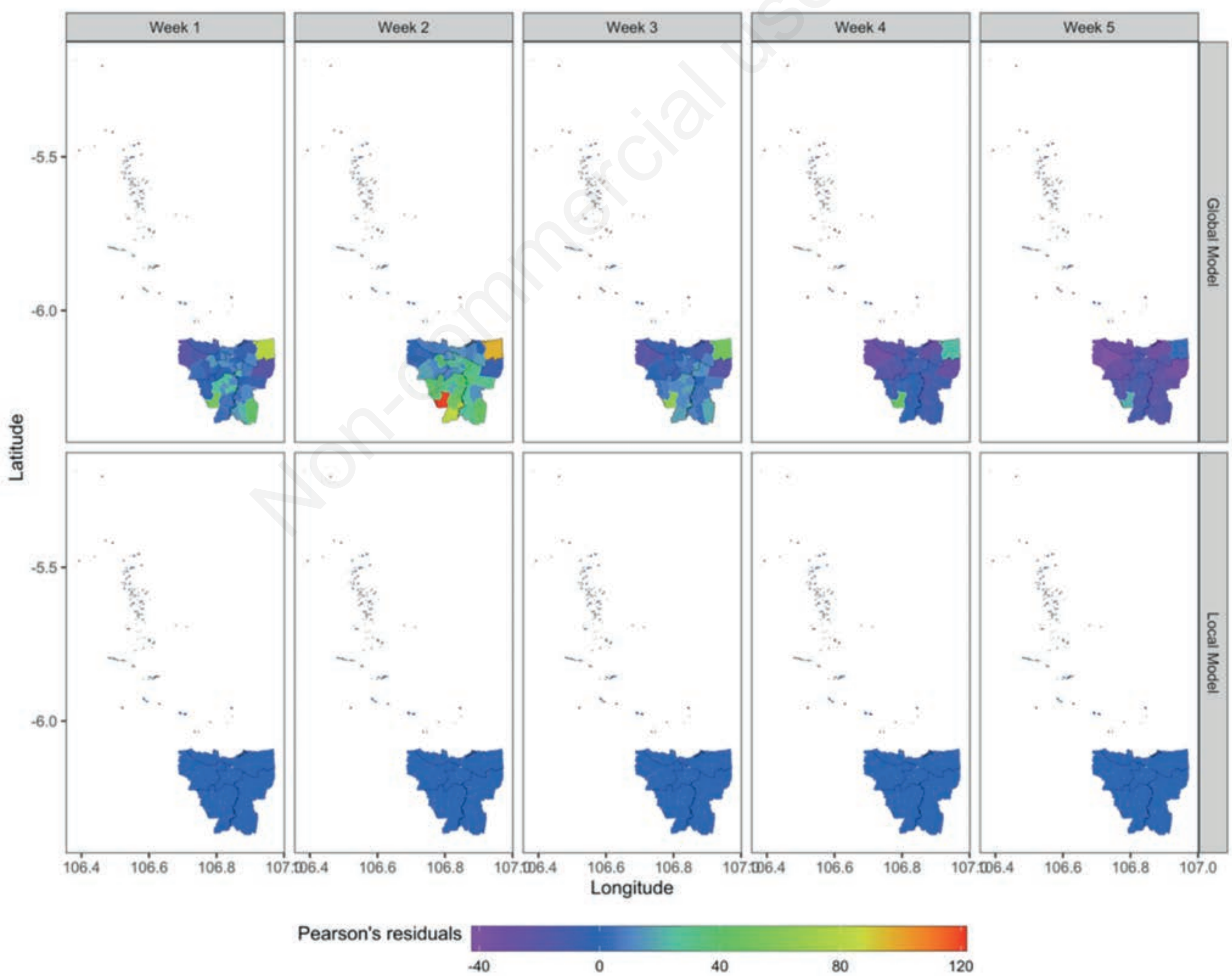

Figure 4. Pearson's residual global model *versus* local model (Model 1 *versus* Model 3).

Table 3. Model fitting.

Parameter	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	
Model 2: Exchangeability varying coefficient					
Fixed Effect					
Intercept (α)	-0.991	0.604	-2.185	0.193	
Slope (β)	-2.247	1.634	-5.477	0.954	
Random effect					
Exchangeability varying coefficient ($\sigma_{\beta} i^2$)	1.166	0.129	0.939	1.446	47.920
Temporally structured ($\sigma_{v} i^2$)	1.066	0.569	0.394	2.565	43.820
Spatiotemporal interaction ($\sigma_{\delta it}^2$)	0.201	0.013	0.178	0.227	8.260
Model 3: Spatially varying coefficient					
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 ($\sigma_{\beta} i^2$)	5.596	0.072	3.609	8.717	86.128
Temporally structured ($\sigma_{v} i^2$)	0.861	0.270	0.107	5.299	13.249
Spatiotemporal interaction ($\sigma_{\delta it}^2$)	0.041	0.000	0.032	0.052	0.624

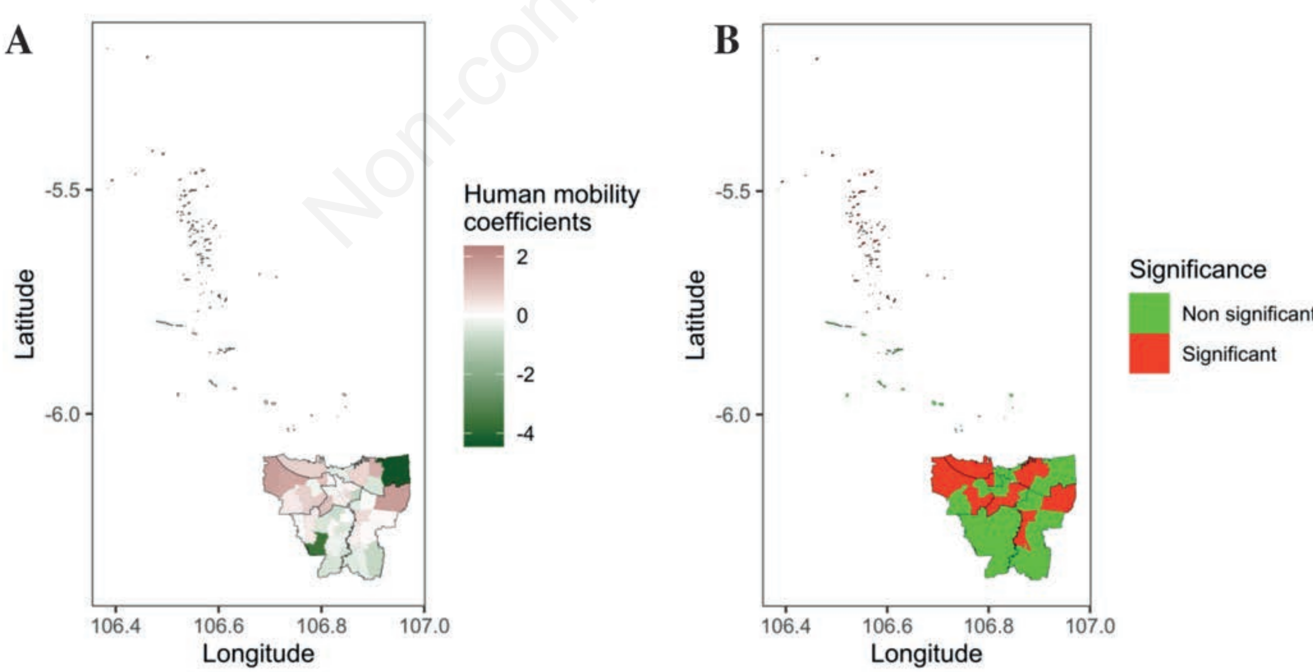

Figure 5. Coefficients of human mobility. A) Spatially varying coefficients; B) Significance of spatially varying coefficient.

decreased (Figure 6A). The proportion of variance explained by spatiotemporal interaction relative to the total random effects was 0.64% indicating that space and time interact less. Figure 6B depicts the plots of the spatiotemporal interaction components demonstrating that nearly all districts shared a similar temporal pattern and indicating the significant impact of the temporal trend.

Relative risk estimate

We calculated the relative risk of COVID-19 for the 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 7A), which helped to identify districts with a substantial high risk based on the exceedance probability (Figure 7B). Figure 6B depicts the spatiotemporal distribution of the relative risk θ_{it} from over the weeks when spatially varying coefficients, temporal effects and space-time interactions had been accounted for. The relative risks were significantly clustered and began to steadily decrease from the beginning of the second week. Those areas with $\theta_{it} > 1$ showed number of COVID-19 cases to be higher than what was expected, while the situation was the opposite for the areas with $\theta_{it} < 1$ have lower than expected cases. A few districts with particularly high and low risk appeared to form and gradually fade

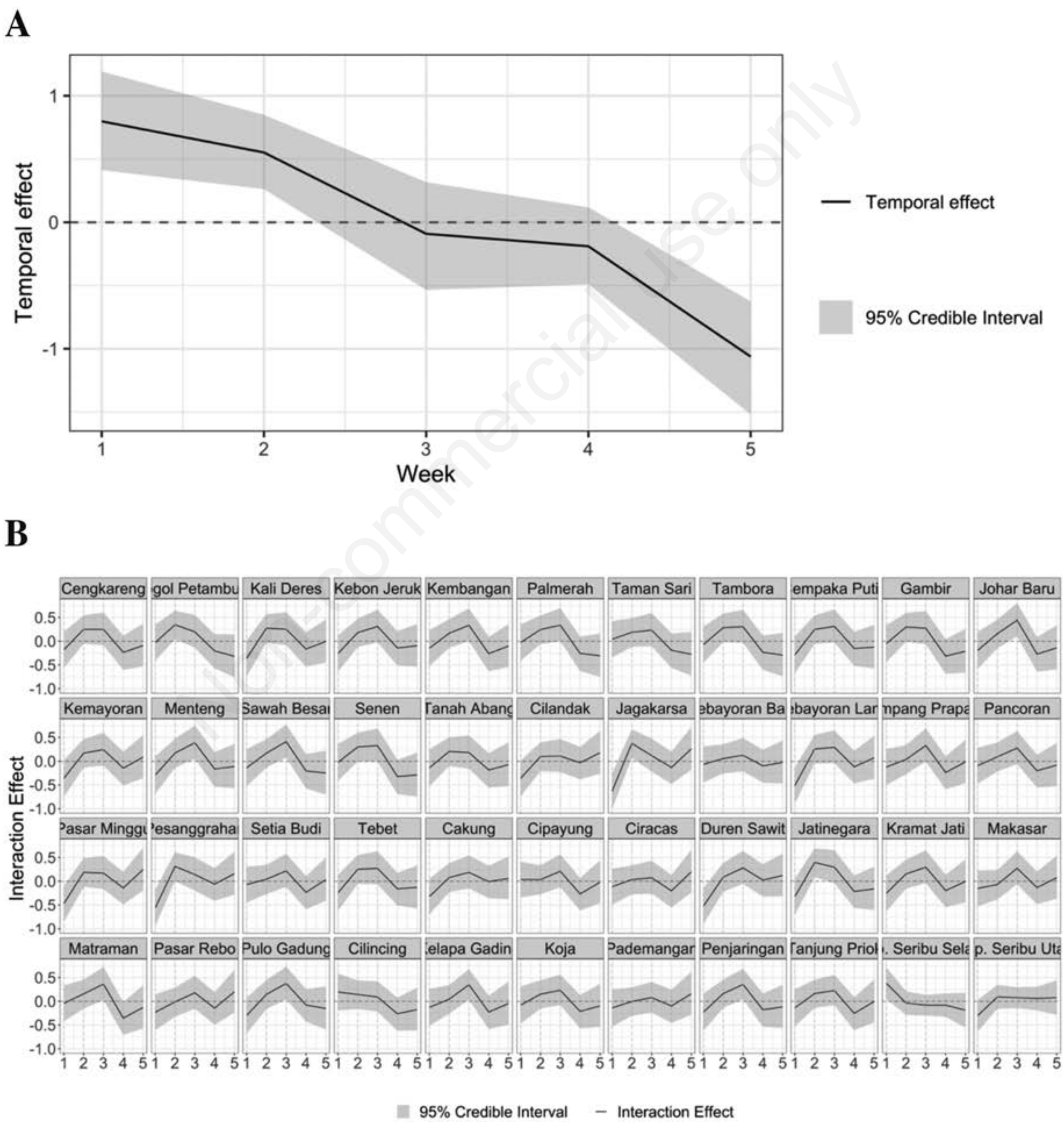

Figure 6. A) Structured temporal effect in 2021 from first week in July (week 1) to first week in August (week 5); B) Spatiotemporal interaction effect in 2021 for 44 districts from first week in July (week 1) to first week in August (week 5).

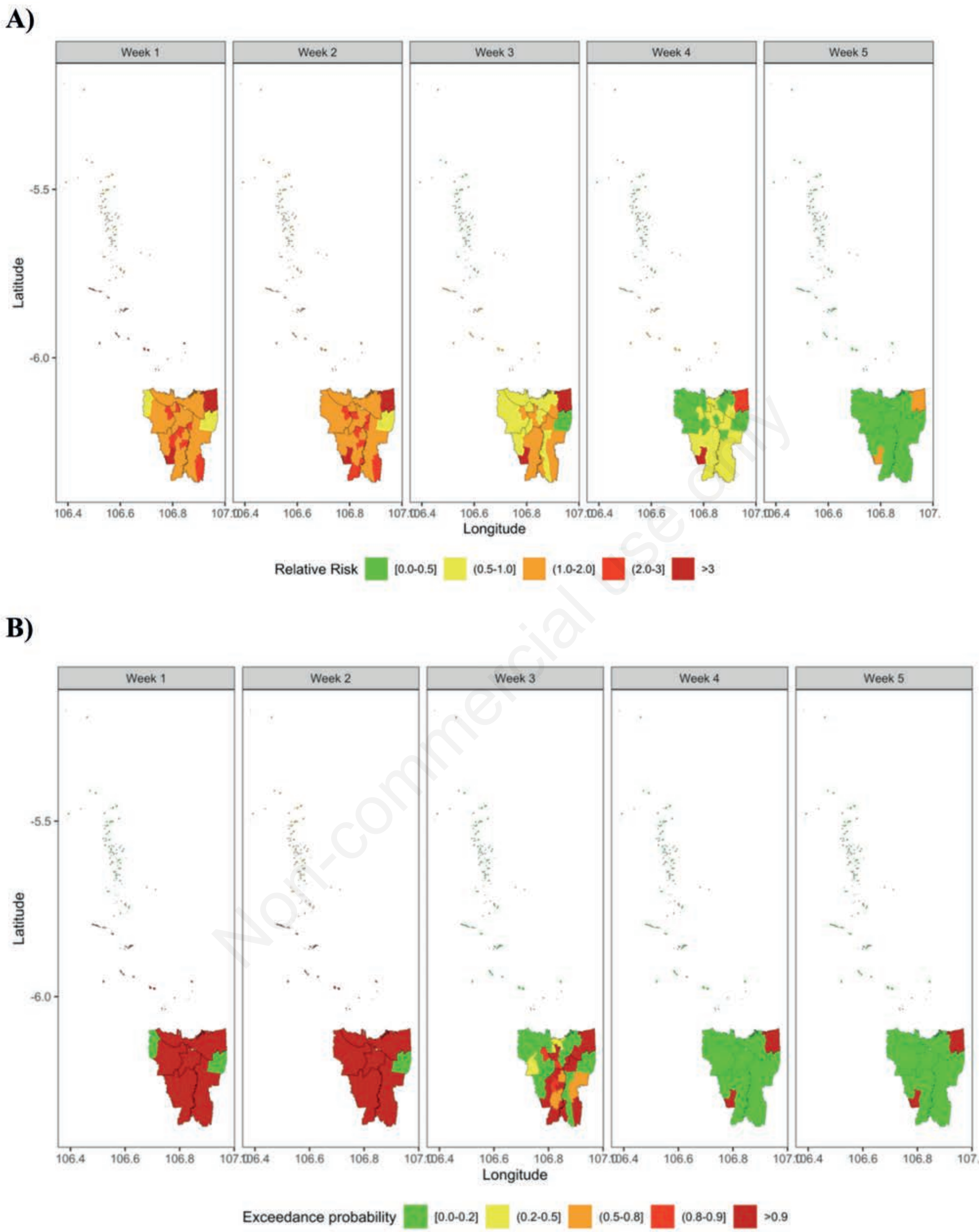

Figure 7. A) Mapped spatiotemporal distribution of the posterior means of the relative risk of COVID-19 in 2021 from first week in July (week 1) to first week in August (week 5). Red indicates regions with high probabilities, while yellow and green indicate regions with low probabilities; B) The spatiotemporal exceedance probabilities of Pr (θ_{it} > 1|y), in 2021from first week in July (week 1) to first week in August (week 5). Maps created using R software.

over time. Figure 7A shows the corresponding exceedance probabilities, *i.e*., those characterized as Pr (θ_{it}> 1|y). The pattern of districts with relative risks of θ_{it}> 1 appeared to be spatially continuous, as evidenced by the maps of probability of exceedance. For the first two weeks, more than 98% of the areas were identified as high-risk ones. After the second week, the relative risk decreased further, which is consistent with the temporal pattern that revealed a sharp decrease in risk from week 3 to week 5 (Figure 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. Using spatially varying coefficients with the ICAR prior, we were able to account for local variations in the effects of neighbourhood human mobility. The most important implication of our findings is that human mobility is spatially continuous, so a global regression model is insufficient for quantifying human mobility effects. Usually, neighbourhood 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 boundaries can result in spatial spill-over. Demographic variations also contribute to the non-stationary effects of the risk factors. Adopting spatially structured random-effect 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 estimation method to fit regression models. The 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(s). Our empirical research on COVID-19 in Jakarta 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 the 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. Local regression estimates may have significant effects on the organization and evaluation of treatments.

Using the SVC model with ICAR prior allowed the relative risk at the district level to be calculated. The choropleth maps (Figure 7A) demonstrate a significant discrepancy between regions after the third study week followed by a gradual fall, phenomena that 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. Using exceedance probability, we discovered statistical evidence that the relative risk in the North- east and the South-west remained consistently high during PPKM requirements. Thus, restricting human mobility in this area would not have a significant effect on disease transmission. Importantly, additional research to determine the geographically variable effects of climatic conditions on COVID-19 would be beneficial. Sensitivity of spatially varying regression coefficients to complicated network structural dependencies and hyperprior distribution is a deserving area for further study.

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, and they suggest that a reduction in human mobility could dramatically reduce transmission of COVID-19 in a few places, while not having any noticeable impact elsewhere. 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. Our work also offers a detailed methodology for modelling the effect of risk factors on disease risk in a heterogeneous population.

References

Adin A, Goicoa T, Hodges JS, Schnell PM, Ugarte MD, 2022. Alleviating confounding in spatio-temporal areal models with an application on crimes against women in India. Stat Modelling: 1-22.

Bauer C, Wakefield J, Rue H, Self S, Feng Z, Wang Y, 2016. Bayesian penalized spline models for the analysis of spatiotemporal count data. Stat Med 35:1848-65.

Congdon P, 2018. Spatial heterogeneity in Bayesian disease mapping. Geo Journal 1:1-14.

Facebook Data for Good, 2022. Disease Prevention Maps. Accessed 21 August 2022. Available from: https://dataforgood-facebookcom/

Fiebig DG, Bartels R, Aigner D J, 1991. A random coefficient approach to the estimation of residential end-use load profiles. J Econom 50:297-327.

Finley AO, 2011 Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. Methods Ecol Evol 2:143–54.

Firza N, Monaco A, 2022. forecasting model based on lifestyle risk and health factors to predict COVID-19 severity. Int J Environ Res Public Health 19:12538.

Hou X, Gao S, Li Q, Kang Y, Chen N, Chen K, Rao J, Ellenberg JS, Patz JA, 2021. Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic age and race. PNAS 118:e2020524118.

Jaya IGNM, Folmer H, 2021a. Bayesian spatiotemporal forecast-

ing and mapping of COVID-19 risk with application to West Java Province Indonesia. J Reg Sci 61:849-81.
Jaya IGNM, Folmer H, 2021b. 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 53:1-57.
Jaya IGNM, Folmer H, 2020. Bayesian spatiotemporal mapping of relative dengue disease risk in Bandung Indonesia. J Geogr Syst 22:105-142.
Jaya IGNM, Folmer H, Ruchjana BN, Kristiani F, Yudhie, 2017. 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 2 pp 239-254.
Karcıoğlu O, Yüksel A, Baha A, Er BA, Esendağlı D, Gülhan P Y, Karaoğlanoğlu S, Erçelik M, Şerifoğlu İ, 2020. Covid-19: The biggest threat of the 21st century: In respectful memory of the warriors all over the world. Turk Thorax J 21:409-18.
Knorr-Held L, 2000. Bayesian modelling of inseparable space-time variation in disease risk Stat Med 19: 15-30.
Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM Open COVID-19 Data Working Group; Plessi LD The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368:493-497.
Kucharski AJ, Russell TW, Diamo C, Liu Y, Edmunds J, Funk S, Eggo RM, 2020. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis 20:553-58.
Lome-Hurtado A, Lartigue-Mendoza J, Trujillo JC, 2021. Modelling local patterns of child mortality risk: a Bayesian Spatio-temporal analysis. BMC Public Health 21:1-12.
Osei F, Stein A, 2017. Diarrhea morbidities in small areas: Accounting for non-stationarity in sociodemographic impacts using Bayesian spatially varying coefficient modelling Sci Rep 7:1-15.
Rue H, Martino S, Chopin N, 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc 7:319-92.
Rue H, Riebler A, Sørbye SH, Illian JB, Simpson DP, Lindgren FK, 2017. Bayesian computing with INLA: A review. Annu Rev Stat Appl 4:395-421.
Shepherd HE, Atherden FS, Chan HMT, Loveridge A, Tatem AJ, 2021. Domestic and international mobility trends in the United Kingdom during the COVID-19 pandemic: an analysis of facebook data. Int J Health Geogr 20:1-13.
Sparks C, 2015. An examination of disparities in cancer incidence in Texas using Bayesian random coefficient models. Peer J 3:e1283.
Sperrin M, McMillan B, 2020. Prediction models for covid-19 outcomes. BMJ 371:m3777.
Vicente G, Goicoa T, Ugarte M, 2020. Bayesian inference in multivariate spatio-temporal areal models using INLA: analysis of gender-based violence in small areas. Stoch Environ Res Risk Assess 34:1421–40.
Wakefield J, 2007. Disease mapping and spatial regression with count data. Biostatistics 8:158-83.
Wang S, Liu Y, Hu T, 2020. 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 17:7930.
WHO, 2020. Briefing 11 March Accessed on 18 Novemeber, 2022. Available from: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-march-2020
WHO, 2022. Coronavirus (COVID-19) dashboard. Accessed 25 August 2022. Available from: https://covid19whoint/
Yabe T, Tsubouchi K, Fujiwara N, Wada T, Sekimoto Y, Ukkusur SV, 2020. Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. Sci Rep 10:18053.
Yuan Z, Xiao Y, Dai Z, Huang J, Zhang Z, Chen Y, 2020. Modelling the effects of Wuhan's lockdown during COVID-19 China. Bull World Health Organ 98:84–494.
Zhang M, Wang S, Hu T, Fu X, Wang X, Halloran B, Li Z, Cui Y, Liu H, Liu Z, Bao S, 2022. Human mobility and COVID-19 transmission: a systematic review and future directions. Ann GIS 1-14.