A dynamic-time dependent spatial autocorrelation detection for East Java's Covid-19 regional percent of cases, March 2020–March 2021 (Indonesia)

Rahma Fitriani
Department of Statistics,
University of Brawijaya,
Malang, Indonesia
E-mail:rahmafitriani@ub.ac.id

Darmanto Darmanto
Department of Statistics,
University of Brawijaya,
Malang, Indonesia
E-mail: darman_stat@ub.ac.id

Zerlita F. Pusdiktasari
Department of Statistics,
University of Brawijaya,
Malang, Indonesia
E-mail: zerlitafahdha@gmail.com

Covid-19 regional percent of cases is one of the regional variables that dynamically interact across space and time. It exhibits a time trend, and at one point in time, it may form clusters of regions with similar values. Since Covid-19 is an infectious disease, the regional percent of cases also exhibits spatial dependence across regions. The time trend indicates the possible time lag of the spatial dependence, and the spatial dependence analysed at one point in time may be undetected. This situation was observed in the 38 regions of East Java. It gives an incorrect impression of the nature of spatial dependence, leading to an improper policy formulation. To capture the spatial interaction more accurately, this study accommodates the time-dependent dynamic nature of the variable into the formulation of the Moran's I index for a set of spatial panel data. A simulation study is conducted to confirm the accuracy of the proposed index, especially when the degree of contemporaneous spatial autocorrelation is high. The proposed index also succeeds in detecting the time-lagged spatial autocorrelation of East Java's Covid-19 regional percent of cases. It provides a better understanding and policy recommendations regarding the spread of this disease in this province.

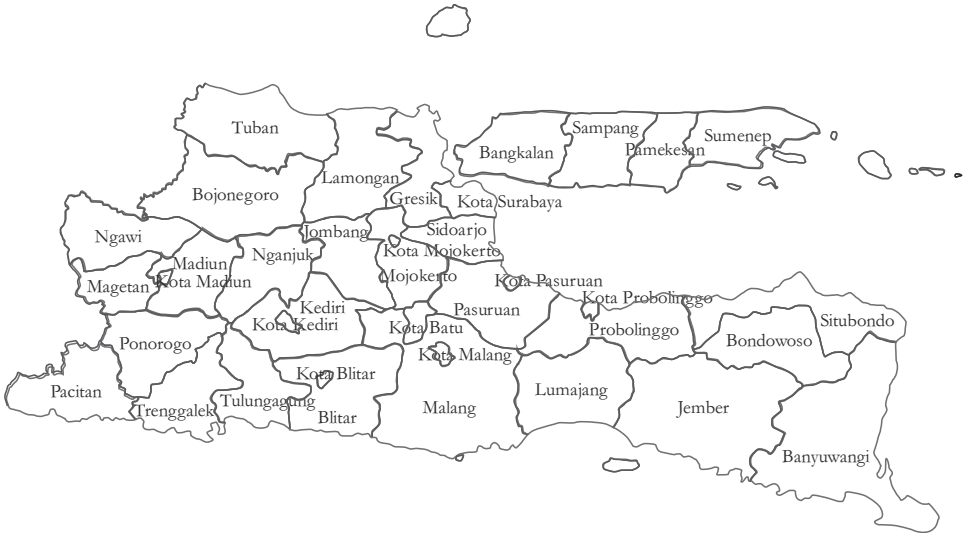
Keywords:
Moran's I,
dynamic spatial autocorrelation,
contemporaneous,
time lagged,
Covid-19 percent of cases,
East Java

Introduction

After being first detected in Wuhan at the end of 2019, Covid-19 spread across the world and has become a pandemic (Kincses–Tóth 2020, Nyikos et al. 2021). While most countries have shown a decreasing number of new cases, as of the middle of 2021, Indonesia is still among the top ten countries with a high number of daily new cases (Worldometer.com 2021). The pandemic has hit the country hard in every aspect of life, such as economic, poverty, and health care systems (Olivia et al. 2020, Suryahadi et al. 2020). During the pandemic, East Java, one of the provinces in Java, Indonesia, consisting of 38 regions (regencies/municipalities) (Figure 1), has always been categorised as a province with a high number of cases and infection rates (Kemenkes 2021). The number of new confirmed cases in a certain period is a common measure of the spread of the disease in this province, as well as in other parts of the world. At the regional level, however, the extent of the spread must be measured relative to the population number. Therefore, the percentage of cases (i.e. the number of new confirmed cases divided by the population number, in percentage value) will be a more accurate measure.

Figure 1

The map of regencies/municipalities of East Java



Note: List of East Java 38 regions (regencies/municipalities) see in the Appendix.

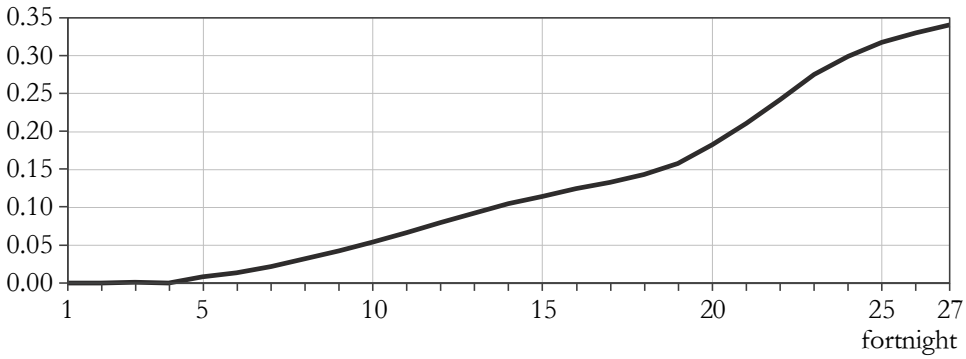
The spread of the disease in this province, as well as in other Indonesian provinces, is faster than the government's response in limiting people's mobility from one region to another (Olivia et al. 2020). Even though people may travel to distant regions, this type of mobility is less frequent than that between shared border regions. Therefore, it is more likely that the increasing percent of cases in one regency/municipality

eventually increases that in its nearby regencies/municipalities. It indicates the presence of first-degree spatial dependence across regions in terms of the percent of cases. This follows the First Law of Geography that „everything is related to everything else, but near things are more related than distant things" (Tobler 1970). The spread of the disease can also be observed through the time series of each regional percent of cases. Given the nature of the disease, the current local percent of cases determines the future local as well as nearby regions' percent of cases. For this specific virus, a duration of two weeks (a fortnight) is the upper bound of the incubation period (Backer et al. 2020, Lauer et al. 2020, Quesada et al. 2021); that is, it takes approximately two weeks for the current cases to create additional new cases. The time series of the aggregated two-weekly percent of cases in East Java, from March 2020 to March 2021, exhibits such a trend (see the time series plot in Figure 2), which confirms the time-dependent property of this variable. A period of two weeks is used, assuming that the strongest time dependence is observed between the two-weekly percent of cases.

The characteristics of this disease suggest that the percent of Covid-19 cases within an administrative region is one of the regional variables (e.g., economic growth, population density, inflation rate) which is distributed and evolved in space over time. The time trend of the two-weekly percent of cases further suggests that, even though the spatial dependence of this variable exists across regions, it has a certain time lag. Due to the time lag, the spatial dependence of the percent of cases across regions at one point in time may be undetected. It gives an incorrect impression regarding the spread of the disease, which consequently leads to improper policy formulation. Thus, a better understanding of the interregional interaction that leads to the spread of this disease is necessary. Based on a set of spatial panel data, this study argues that to capture a more accurate description of the spatial interaction of East Java's regional Covid-19 percent of cases, the spatial dependence detection procedure must accommodate the time-dependent dynamic nature of the variable.

Figure 2

Time series of East Java's Covid-19 percent of cases, March 2020–March 2021



The spatial pattern of the variable under study at one point in time can be assessed by calculating the degree of spatial autocorrelation. Several statistics such as Moran's I index (Moran 1950), Geary's C index (Geary 1954), and Getis-Ord's G index (Getis–Ord 1992) have been developed for this purpose. The clustering spatial pattern leads to a positive spatial autocorrelation, and a negative spatial autocorrelation indicates a dispersing spatial pattern. To observe the degree of spatial autocorrelation as a function of distance, Chen (2021) developed a spatial autocorrelation function based on these indexes. When spatial panel data are available (for N locations and T time units), the degree of spatial autocorrelation can still be measured using one of the indexes. Generally, these indexes use a spatial weight matrix (W_S) which captures the connectivity of every pair of spatial units. To match the time dimension of the spatial panel data, a Kronecker product between I_T, an identity matrix of size T, and a spatial weight matrix (W_S) is used. It is an operation (using an operator $\otimes$) which transforms those two matrices into a larger matrix that contains all the possible products of the entries of the two matrices. The product ($I_T \otimes W_S$) is defined by assuming contemporaneous spatial interaction and no time dependence. Unfortunately, these indexes are incapable of identifying or accurately detecting the spatial interaction with time lag or time dynamics. This study therefore modifies the formulation of the index by accommodating the time-dependent component.

Previous researchers (Gao et al. 2019, Jaya et al. 2019, Lee–Li 2017, Shen et al. 2016, Wang–Lam 2020) developed a similar index using a spatiotemporal weight matrix. However, except for Wang–Lam (2020), these studies did not specifically develop an index for spatial panel data. Shen et al. (2016) used the detrended variable to calculate the index in the presence of non-stationarity. The current study specifically utilises the same spatiotemporal weight matrix proposed by Wang–Lam (2020). Unlike Wang–Lam (2020), who use the weight matrix to modify the Getis-Ord's index, this study uses the weight to modify the Moran's I index. It focuses on Moran's I index because it is defined as the slope between the variable under study and its corresponding spatial lag (Anselin 1999, 2001), which enables the visualisation of the degree of spatial autocorrelation via a scatter plot between those two variables. The index is modified into the space-time Moran's I index to detect the dynamics of the spatial interaction in a set of spatial panel data.

This study proposes two possible specifications in the modified index which may capture the time-dependent dynamic spatial interaction of East Java's regional Covid-19 percent of cases, or any regional variable in general. They are the specifications for contemporaneous space-time interaction and space-time lagged interaction. However, the accuracy of the proposed index needs to be evaluated. For this purpose, before applying the index for East Java's regional Covid-19 percent of cases, a simulation study was conducted. In the simulation, several sets of spatial panel data were generated. Each set has different combinations between the degree of spatial interaction and the degree of time dependence. The proposed index is then applied

to measure the pre-determined dynamic time-dependent spatial interaction within the generated spatial panel data. The result will be useful to define in which condition and specification of the dynamic spatial dependence the index performs well. It assists this study to reach its objective, namely providing a more accurate dynamic time-dependent spatial interaction of East Java's regional Covid-19 percent of cases.

Materials and methods

The proposed index is formulated by modifying the weight matrix of the Moran's I index. Therefore, this section begins with an explanation of the index and proceeds with the formulation of the modified weight.

Moran's I index

It is a commonly used index to measure the degree of spatial autocorrelation in a set of spatial data (Anselin 1988, Egri–Tánczos 2018, Fitriani et al. 2021). The index was developed by Moran (1950) as a two-dimensional version of the test for univariate time-series correlation. The index is defined as:

$$I = \frac{\sum_{j=1}^{N}\sum_{i=1}^{N} w_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}. \quad (1)$$

where Y_i is the observed variable value at spatial unit i, $i = 1,2,\ldots, N$, and $w_{i,j}$, is the ij-th element spatial weight matrix W_S, for $i, j = 1,2,\ldots, N$. The range of the index is from –1 to 1. A positive index indicates a cluster of locations with similar values, with 1 representing the strongest/perfect similarity. A negative index indicates dissimilar values of nearby regions, with –1 capturing the most contrasting values. A 0-value index indicates a random pattern or no spatial autocorrelation. This study focuses on the positive index, which most likely occurs in economic, demographic, or epidemiological cases.

This index was further used to test the null hypothesis of spatial randomness. Under the null hypothesis, Moran's I in has the following expected value:

$$E(I) = \frac{-1}{N-1} \quad (2)$$

and variance:

$$var(I) = \frac{NS_4 - S_3S_1(1-2N)}{(N-1)(N-2)(N-3)\left(\sum_{j=1}^{N}\sum_{i=1}^{N} w_{ij}\right)^3} \quad (3)$$

in which:

$$S_1 = \sum_{j=1}^{N}\sum_{i=1}^{N}\left(w_{ij} + w_{ji}\right)^2, S_2 = \sum_{i=1}^{N}\left(\sum_{j=1}^{N} w_{ij} + \sum_{j=1}^{N} w_{ji}\right)^2$$

$$S_3 = \frac{N^{-1}\sum_{i=1}^{N}(X_i - \bar{X})^4}{\left(N^{-1}\sum_{i=1}^{N}(X_i - \bar{X})^2\right)^2},$$

$$S_4 = (N^2 - 3N + 3)S_1 - NS_2 + 3\left(\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\right)^2$$

Using the assumption of normality (independent normal random variates), an inference for Moran's I can be made by employing an approximately standard normal distribution:

$$Z = \frac{I - E(I)}{\sqrt{var(I)}} \sim N(0,1) \tag{4}$$

and an extreme value of this statistic leads to rejection of the null hypothesis. However, this approach may be inappropriate when the underlying assumption is violated (Cliff–Ord 1981). Instead, a computational approach based on permutation was used (Anselin 1995, 2001). Under the null hypothesis, the reference distribution of the statistic is the distribution of Moran's I which is calculated from several random data sets. The random data sets are the permutation of the observed values over the locations. A pseudo p-value was then calculated from the reference distribution using the following:

$$p = \frac{R + 1}{M + 1} \tag{5}$$

where R is the frequency of the computed Moran's I from the permuted data sets which are equal to or greater than the observed Moran's I (from the original data set), and M is the number of permutations. To yield a good approximation of the p-value, M is usually set as 99, 999, etc.

For a set of spatial data, Moran's I basically measures the correlation between the observed variable at a location and the average observed variables from its neighbouring locations (Anselin 1988). These are denoted by Y and $W_S Y$, respectively. The latter is also known as the corresponding spatial lag variable. The formulation of the index depends on the definition of the spatial weight matrix W_S, which is an N x N positive matrix. The positive non-zero elements of the matrix indicate whether the two locations are neighbours. The following definition is commonly used:

$$w_{ij} = \begin{cases} 1, & if\ i\ and\ j\ are\ neighbors \\ 0, & otherwise \end{cases} \tag{6}$$

Other definitions of ij-th matrix element, based on any decreasing function of distance between locations i and j, can also be applied. However, whatever function is used, by convention, the diagonal elements w_{ii} are all set to zero to exclude self-neighbours. The matrix is generally transformed into a row-standardised form.

Spatiotemporal weight matrix

For a spatial panel data setting, the spatial weight matrix used in the Moran's I index in needs to be modified. The weights do not define only the spatial configuration, but also the time dependence. This study refers to the specification of the space-time weight matrix proposed by Wang–Lam (2020). They defined two types of weights based on the following situations:

1. Space-time dependency structure with contemporaneous specification

The variable observed at location i (i = 1,…,N) at time t, t = 1,…,T (Y_{it}) depends on the variable previously observed at the same location ($Y_{i(t-1)}$) and on the neighbouring observed variable at time t (Y_{jt}, j ≠ i,j ∈ Neighbouring of i). By the local time dependence nature, Y_{it} will indirectly affect the future neighbouring variable ($Y_{j(t+1)}$, j ≠ i,j ∈ $N(i)$). For this situation, the weight is defined as:

$$V_{ST}^{C} = I_T \otimes W_s + W_T \otimes I_s \tag{7}$$

2. Space-time dependency structure with lagged specification

The variable observed in location i ($i = 1, \dots, N$) at time t (Y_{it}) depends on the variable previously observed at the same location ($Y_{i(t-1)}$) and at the neighbouring locations ($Y_{j(t-1)}, j \neq i, j \in$ Neighbouring of i). In this case, there is a time lag for the spatial effect. The weight is defined as:

$$V_{ST}^{L} = W_T \otimes W_s + W_T \otimes I_s \tag{8}$$

The matrices in (7) and (8) are $N \times T$ space-time weight matrices. In these matrices:

$$W_T = \begin{bmatrix} t_{11} & \dots & t_{1T} \\ \vdots & \vdots & \vdots \\ t_{T1} & \dots & t_{TT} \end{bmatrix} \tag{9}$$

is a $T \times T$ time weight matrix, in which:

$$t_{ij} = \begin{cases} 1, if\ |t_i - t_j| \leq t_\theta \\ 0, if\ |t_i - t_j| > t_\theta \end{cases} \tag{10}$$

In (10), t_i is the occurrence of the i-th observation, t_j is the occurrence of the j-th observation, $i, j = 1, \dots, T$ and t_θ is the time duration threshold, in which two observations with t_θ time unit apart still affect each other. By assuming that the strongest effect comes from two observations which are one time unit apart, this study uses $t_\theta = 1$. In (7) and (8) W_S is the common spatial weight matrix. It may use the definition in (6) to define the connectivity between locations or any other distance-based weight concepts. I_T is an identity matrix of size T, and I_s is an N x N identity matrix. For comparison purposes, another weight matrix which is also compatible with the panel setting can be defined as:

$$V = I_T \otimes W_s \tag{11}$$

with the same I_T and W_s defined in (7) and (8). In contrast with the previous matrices, the weight matrix in (11) is formulated by assuming that there is no time dependence. Spatial interaction occurs instantly, without being transferred to the next period or affected by observation during the previous period.

Spatiotemporal Moran's I index

The index is a modified version of Moran's I index. It is presented in a matrix/vector notation as follows:

$$I_{ST} = \frac{X' V_{ST} Y}{Y' Y} \tag{12}$$

using V_{ST}, the space-time weight matrices defined in (7) or (8), and

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y_1} - \bar{\boldsymbol{Y}} \\ \vdots \\ \boldsymbol{Y_T} - \bar{\boldsymbol{Y}} \end{bmatrix} = \begin{bmatrix} Y_{11} - \bar{Y} \\ \vdots \\ Y_{N1} - \bar{Y} \\ Y_{12} - \bar{Y} \\ \vdots \\ Y_{N2} - \bar{Y} \\ \vdots \\ Y_{1T} - \bar{Y} \\ \vdots \\ Y_{NT} - \bar{Y} \end{bmatrix}$$

which is a vector of deviation from the mean of the observed variable, in the form of a stacked cross section.

Generated data

Each set of the spatial panel data are generated based on the following model:

$$Y_t = \rho W_S Y_t + \varepsilon_t \quad (13)$$

which Arbia (2006) and Anselin (2001) define as a pure autoregressive model. The model is chosen because it focuses on the interaction between the observed variable in a location and its spatial lag. In other words, the local observed situation is purely the product of the surrounding situation.

Following the notation in Elhorst (2014), in (13), $\boldsymbol{Y_t} = \begin{bmatrix} Y_{1t} \\ \vdots \\ Y_{38,t} \end{bmatrix}$ is a vector of spatial data at time $t, t = 1, \ldots, 24$, and W_s is a 38 x 38 spatial queen contiguity matrix. The matrix is defined based on the spatial configuration of East Java's 38 regencies/municipalities (see the map of the province in Figure 1). ρ is a parameter which determines the degree of spatial autocorrelation at one point in time. $T = 24$ is chosen accordingly, such that the pattern of time dependence is apparent. To create time dependence, the error terms are generated based on the first-order autoregressive process (AR(1)), such that

$$\begin{gathered} \varepsilon_t = \phi \varepsilon_{t-1} + v_t, \\ \varepsilon_0 \sim N_N(0, I_S \sigma^2), and\ v_t \sim N_N(0, I_S \sigma^2) \end{gathered} \quad (14)$$

where $\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{38,t} \end{bmatrix}$ is a vector of the error terms for the spatial data observed at time t, I_S is a 38 x 38 identity matrix, and ϕ is the degree of time dependence in every spatial unit. Based on the values of ρ and ϕ, several scenarios are used to generate the spatial panel data (see Table 1). These are the parameters of the spatial autoregressive and (time) autoregressive processes, respectively.

Table 1

Combination of ρ and ϕ for every scenario

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ϱ	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.5	0.5	0.5	0.9	0.9	0.9	0.9
φ	0.1	0.5	0.9	1.0	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	1.0

When $\rho > 0$ and $\phi = 0$, the data exhibit a certain degree of contemporaneous spatial autocorrelation without any time dependence. On the other hand, when $\rho = 0$ and $0 < \phi \leq 1$, the time series of each spatial unit exhibits time dependence, but there is no contemporaneous spatial dependence in each time unit. The rest of the combinations in Table 1 capture the presence of interaction between the spatial and time dependences, which produce a dynamic time-dependent spatial effect. $\rho = 0$ and $\phi = 0$ represent no contemporaneous spatial dependence and no time dependence, respectively. $\rho = 0.5$ and $\phi = 0.5$ represent medium contemporaneous spatial dependence and medium time dependence, respectively. $\rho \to 1$ and $\phi \to 1$ represent strong contemporaneous spatial dependence and time dependence, respectively. This study uses the autoregressive parameter $\phi = 1$ to create a random walk, which is a non-stationary time series (Cryer–Chan 2008). The simulation does not include $\rho = 1$ because the spatial process is assumed to be stationary. Kelejian–Prucha (1998) derived $-1 < \rho < 1$ as the stationary condition for the spatial process.

For each scenario, 500 sets of spatial data were generated. The index defined in (12) was applied to test the null hypothesis of no space-time autocorrelation of each data set per scenario. The space-time weight matrices in (7) and (8) are used to calculate the index. The accuracy of the test, based on the two types of weight matrices, for each scenario was analysed based on the percentage of rejections among the 500 tests.

East Java's Covid-19 percent of cases data

The data are a set of fortnightly Covid-19 percent of cases from March 2020 to March 2021 for East Java's 38 regencies/municipalities. The fortnightly data were chosen because a two-week period is considered as the upper bound of the incubation period of the virus, a duration of time until the spatial effect comes into effect. For this set of spatial panel data, the space-time Moran's I index in (12) is calculated using two types of the space-time weight matrix, in (7) and (8), followed by testing the null hypothesis of no space-time autocorrelation. The result for each type of matrix is then compared to understand the underlying nature of the space-time autocorrelation of the data on the number of Covid-19 cases in East Java.

Results

Simulation study

This section begins with a graphical presentation of the nature of the time and spatial dependence of the generated spatial panel data. Data are generated following models (13) and (14) such that when $\rho = 0$, several units of time series without contemporaneous spatial dependence are obtained, and when $\phi = 0$, several sets of spatial data without time dependence are obtained. The graphical presentation is made only for the datasets for the two extreme cases. A time series plot of each spatial unit

was made to visualise the time dependence. Conversely, because the spatial autocorrelation is basically a correlation between Y and its spatial lag WY, a scatter plot between these two variables is made to visualise the contemporaneous spatial dependence for each time unit.

A combination of $\rho = 0$ and $\phi = 1$ is used to generate several units of time series without contemporaneous spatial dependence. The time series plots for all 38 spatial units of the generated data are depicted in Figure A1 in the Appendix. For the same generated data, the scatter plots between Y and WY for all 24 time units are presented in Figure A2 in the Appendix. These two figures indicate that the data generation concept with a combination of $\rho = 0$ and $\phi = 1$ produces a set of spatial panel data in which each time series exhibits a trend and no spatial dependence within each time unit.

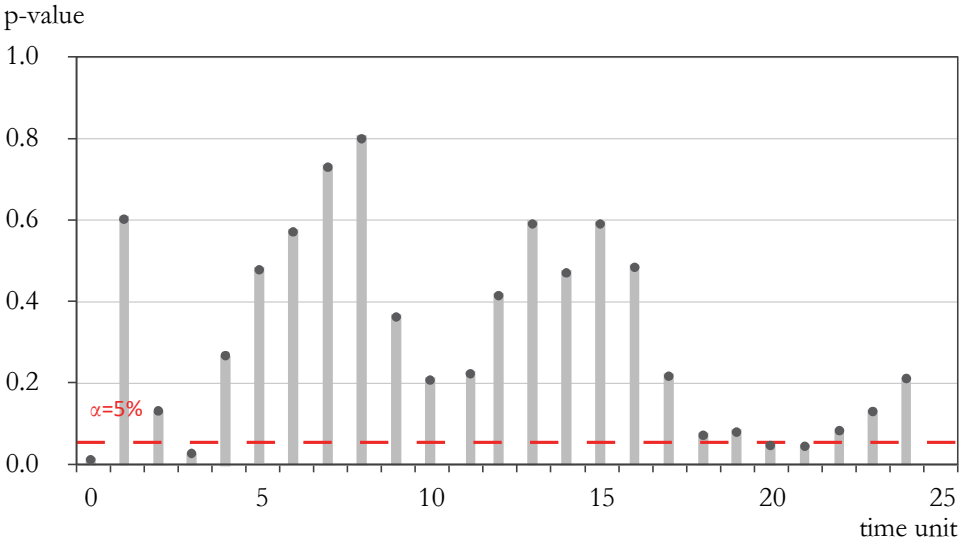
A combination of $\rho = 0.9$ and $\phi = 0$ is used to generate several sets of spatial data without time dependence. The time series plots for all spatial units of the generated data are depicted in Figure A3 and the scatter plots between Y and WY for all 24 time units are presented in Figure A4 in the Appendix. These two figures confirm that the data generation concept with a combination of $\rho = 0.9$ and $\phi = 0$, produces a random time series for every spatial unit and spatial data with positive spatial autocorrelation for every time unit.

In addition to the graphical representations, to confirm the nature of the spatial interaction produced by the data-generating setting, a hypothesis test for no spatial autocorrelation was conducted independently for each set of spatial data. Because the simulation uses 24 time units, there are 24 sets of spatial data. The p-value of each test for the combination of $\rho = 0$ and $\phi = 1$ is illustrated in Figure 3 a). The same results of the test for the combination of $\rho = 0.9$ and $\phi = 0$ are presented in Figure 3 b). Using a significance level of 5 %, the result in Figure 3 a) reveals that the null hypothesis of no spatial dependency is rejected only for the spatial data set of time unit 13. This confirms that this scenario ($\rho = 0$ and $\phi = 1$) produces individual (time unit) generated spatial data which generally have no significant spatial autocorrelation. The opposite holds for the combination of $\rho = 0.9$ and $\phi = 0$. The result in Figure 3 b) demonstrates that for every time unit, the null hypothesis of no spatial dependency is rejected. The test results are in line with the scatter plots depicted in Figure A2 and A4 in the Appendix for each scenario.

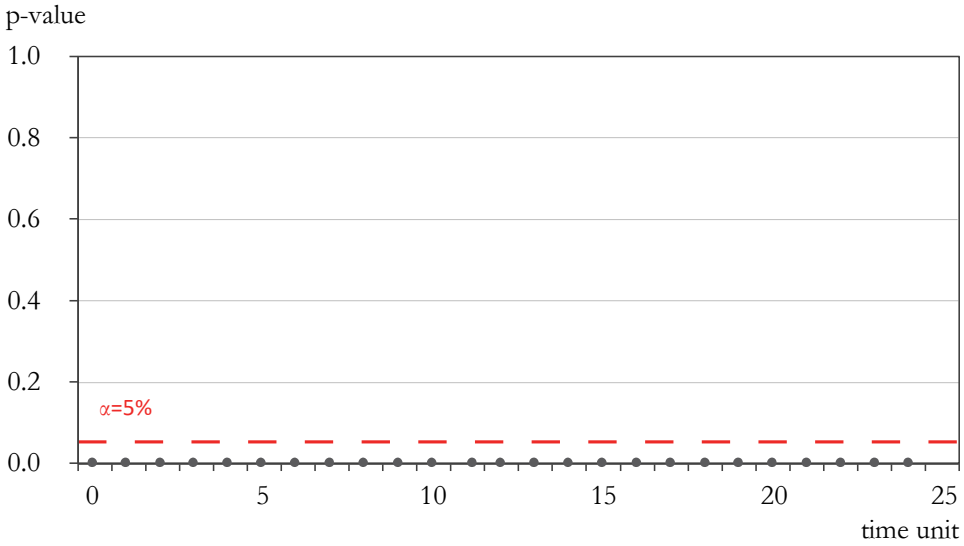
Figure 3

p-value of the Moran's I, spatial autocorrelation test for each time unit

a) Generated data $\rho = 0$ and $\phi = 1$



b) Generated data $\rho = 0.9$ and $\phi = 0$



After confirming the intended nature of the spatial interaction of the generated data, the settings in (13) and (14) are used to generate the spatial panel data for several combinations of ρ and ϕ (scenarios). The scenarios are presented in Table 1. Each scenario is used to generate 500 sets of spatial panel data, and the null hypothesis of no space-time autocorrelation for every data set is tested. The test is conducted using the space-time Moran's I index defined in (12). A total of two types of weight were used to define the index, the weight with contemporaneous space-time dependency setting defined in (7) (V_{ST}^{C}) and that with space-time lagged dependency setting defined in (8) (V_{ST}^{L}). The simulation output of each scenario is the percentage of the null hypothesis' rejection among the 500 tests for each type of weight. Here, the percentage of rejection was used as the measure of the accuracy of the index.

The plots in Figure A5 in the Appendix present the percentage of rejection of the null hypothesis of no space-time autocorrelation for every $\rho, \rho = 0, 0.1, 0.5, 0.9$. They were made to understand the effect of the time dependence parameter pattern ϕ on the accuracy of the test. Each plot has a solid line and a dashed line connecting the percentage of rejection of the test for every chosen ϕ, based on the space-time contemporaneous (V_{ST}^{C}) and space-time lagged (V_{ST}^{L}) dependency specification indexes, respectively.

In Figure A5 in the Appendix a), $\rho = 0$ defines the scenarios of no contemporaneous spatial interaction in each independent time unit. Therefore, when conducted separately for each time unit, the null hypothesis of no spatial autocorrelation will not be rejected. However, when the time series of each spatial unit has time dependence, even at the slightest degree, the spatial dependence is identified as the space-time lagged dependence. In this case, the V_{ST}^{L}-based index performs better than the V_{ST}^{C}-based index. The test using the first specification index has a 100% rejection percentage for most of the scenarios, except for that of weak time dependence ($\phi = 0.1$). However, the test using the latter specification index has a low rejection percentage for all scenarios (less than 30.4%).

The plots in Figure A5 in the Appendix b), c), and d) reveal that when the time trend is combined with an increasing degree of contemporaneous spatial dependence, the spatial dependence can be detected as space-time contemporaneous as well as space-time lagged. In this case, both the V_{ST}^{C}- and the V_{ST}^{L}-based indexes perform well. For all simulated values of ρ, the rejection percentage failed to reach 100% only when the degree of time dependence was low. Therefore, Figure A5 in the Appendix generally confirms the accuracy of the V_{ST}^{C}-as well as the V_{ST}^{L}-based indexes when both the degree of time dependence and the contemporaneous spatial dependence are strong.

The plots in Figure A6 in the Appendix depict the same simulation results for every scenario as the plots in Figure A5 in the Appendix. However, they focus more on the effect of the spatial dependence parameter pattern ρ on the test's accuracy for every $\phi, \phi = 0.1,0.5,0.9,1$. Each plot in Figure A6 in the Appendix has a solid line and

a dashed line to connect the percentage of rejection of the test for every chosen ρ, using the space-time contemporaneous dependency (V_{ST}^{C}) and the space-time lagged spatial dependency (V_{ST}^{L})-based indexes, respectively.

The patterns observed in Figure A6 are in line with the findings in Figure A5 in the Appendix. Every plot in Figure A6 indicates that in the presence of time dependence, the nature of spatial autocorrelation, in every degree, is mostly detected as the space-time lagged autocorrelation by the V_{ST}^{L}-based index. The contemporaneous space-time autocorrelation is detected perfectly by the V_{ST}^{C}-based index only when the contemporaneous spatial dependency is strong ($\rho = 0.9$).

The simulation results confirm the accuracy of the space-time lagged specification Moran's I index to detect the dynamic spatial dependency in a set of spatial panel data. However, the accuracy of the contemporaneous space-time specification Moran's I index depends on the nature of the contemporaneous spatial interaction of the observed spatial panel data.

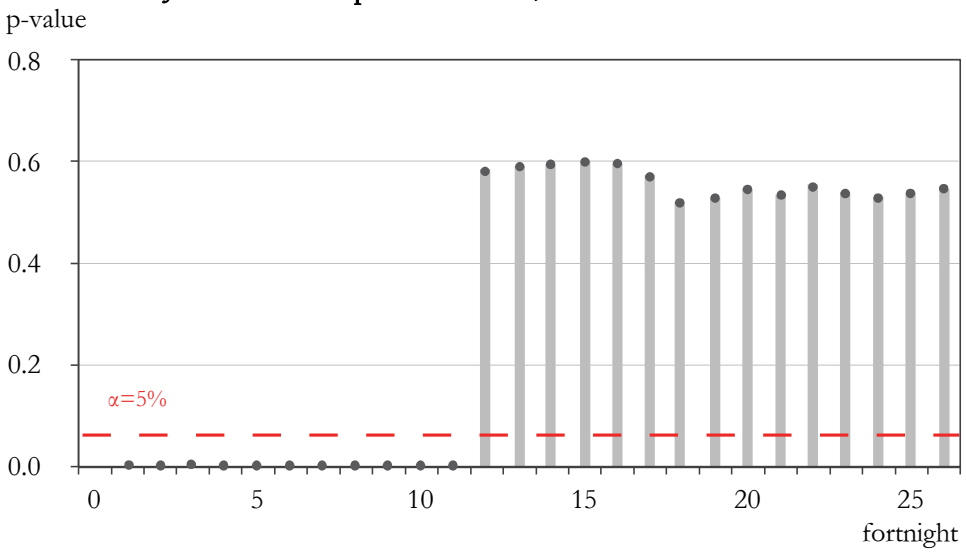
East Java's Covid-19 percent of cases data

It is a set of spatial panel data, the Covid-19 percent of cases of East Java's 38 regencies/municipalities, which are observed fortnightly, from March 2020–March 2021 (27 fortnights). Time series plots for every regency/municipality are presented in Figure A7 in the Appendix. They indicate a certain trend over time for the local percent of cases. Furthermore, the plots between fortnightly local percent of cases and their corresponding neighbours' percent of cases are presented in Figure A8 in the Appendix. The plot for each fortnight exhibits no clear pattern between the local vs the neighbours' percent of cases. In most of the depicted plots, a few points have an outlying pattern compared to the rest of the points. These points belong to the province's capital city (Surabaya) and its surrounding regencies/municipalities, which have always had a relatively higher percent of Covid-19 cases than the rest of the regions since the beginning of the pandemic.

By assuming that there is no time dependence, the null hypothesis of no spatial autocorrelation is tested independently for each (fortnight) spatial data. The test is based on the original Moran's I index, as defined in (1). The p-values of the test for each (fortnight) spatial data are presented in Figure 4. The spatial autocorrelation is significant only at the initial fortnights, that is, in the beginning of the pandemic, when clusters of regencies/municipalities with no cases were found. The test fails to reject the null hypothesis of no spatial autocorrelation for the rest of the fortnights.

Figure 4

p-value of the Moran's I, spatial autocorrelation test for each fortnight unit, East Java's Covid-19 percent of cases, March 2020–March 2021



For East Java's regional Covid-19 percent of cases data, two types of weight matrices in (7) and (8) are used to calculate the space-time Moran's I index in (12), each for different space-time dependency structure. The results of the tests are reported in Table 2. The p-value of the test indicates that the test using the contemporaneous space-time dependency (V_{ST}^{C}) as well as the space-time lagged dependency (V_{ST}^{L})-based indexes leads to the same conclusion. This confirms the dynamic, time-dependent nature of the spatial autocorrelation in the East Java's 38 regencies/municipalities fortnightly Covid-19 percent of cases.

Table 2

Results of hypothesis test of no space-time dependence based on the space-time Moran's I with two different specifications (weights)

Weight	Moran's I Statistic	E(I)	var(I)	p-value
V_{ST}^{C}	0.47384	–0.001	0.0003	2.2 x 10^{-16}
V_{ST}^{L}	0.36388	–0.001	0.0002	2.2 x 10^{-16}

Discussion

The modification of the Moran's I index for a set of spatial panel data is motivated by the presence of the dynamic-time dependent nature of the spatial autocorrelation of several regional variables, particularly Covid-19 percent of cases. However, the accuracy of the proposed index in capturing the dynamic time-dependent spatial autocorrelation still needs to be analysed. For this purpose, a simulation study was conducted as an initial step.

In general, the simulation study reveals that the modified Moran's I for both specifications perform well. The accuracy of the index for each specification depends on the nature of the contemporaneous spatial interaction of the observed spatial panel data. When the (contemporaneous) spatial dependence dominates the time dependence, the contemporaneous space-time specification Moran's I index performs well in detecting the dynamic spatial dependence. On the other hand, when time dependence dominates (contemporaneous) spatial dependence, the space-time lagged specification Moran's I index is more accurate.

From this result, prior knowledge regarding the nature of time dependence and spatial interaction will be useful for determining the specification which best captures the dynamic time-dependent spatial autocorrelation. Visualising the nature of time dependence and spatial interaction is necessary to obtain this knowledge. Time series plots for each spatial unit can be used to visualise the time dependence, whereas the scatter plot of the variable under study (Y) vs its corresponding spatial lag (WY) for each time unit is useful for visualising the spatial interaction. The visible time trend of each spatial unit but no clear pattern between Y and WY of each time unit are an indication of the domination of time dependence over the (contemporaneous) spatial dependence. On the other hand, the obvious pattern between Y and WY of each time unit and no visible time trend of each spatial unit suggests that the (contemporaneous) spatial dependence dominates the time dependence.

For Covid-19, since it is an infectious disease, the percentage of cases for every region/municipality exhibits a trend over time (see time series plots in Figure A7 in the Appendix). Due to people's mobility, the disease is easily spread outside an administrative boundary, such that the increasing percent of cases in one regency/municipality will eventually increase that of cases in its nearby regencies/municipalities. However, the potential spatial dependence cannot be observed instantly. This characteristic is illustrated in Figure A8, in which there is no clear pattern between the local vs neighbourhood percent of cases for each fortnight. The patterns in these figures indicate the dominance of time trends over spatial dependence. The possible time lag of the spatial effect was suspected to be the main reason. In this situation, conducting a hypothesis test of no spatial autocorrelation for each time unit separately potentially leads to incorrect decisions. Based solely on the results in Figure 4, policymakers would be falsely informed that there is no statistical evidence regarding the spread of this disease spatially. The disease may be interpreted as a locally contained one that does not spread into neighbouring regencies/ municipalities. Therefore, accommodating the dynamic time-dependence of spatial autocorrelation is necessary.

The dynamic time-dependent spatial autocorrelation of East Java's regional Covid-19 percent of cases is analysed based on both specifications (based on V_{ST}^{C} and V_{ST}^{L}). This step is performed because each specification has a different mechanism in explaining the space-time effect of East Java's 38 regencies/municipalities Covid-19

percent of cases. The contemporaneous space-time dependency (using V_{ST}^{C}) indicates that the current percent of cases in a regency/municipality depends on the previous fortnight local and the current neighbourhood percent of cases. In this way, the current local percentage of cases directly affects the current neighbouring and the next fortnight local percent of cases. By time dependence, it also indirectly affects the next fortnight neighbouring percent of cases. On the other hand, the space-time lagged dependency (using V_{ST}^{L}) indicates that the current percent of cases in a regency/municipality depends on the percent of previous fortnight local and neighbourhood observed cases. In this way, the current observed local percentage of cases directly affects the next fortnight local percent of cases as well as the neighbourhood percent of cases.

The results in Table 2 confirm the dynamic time-dependent nature of East Java's regional Covid-19 percent of cases. This supports the argument about the importance of accommodating the time-dependent dynamic nature in the spatial dependence detection procedure for East Java's regional Covid 19 percent of cases, or other regional variables with similar nature. The results in Table 2 also indicate that the magnitude of V_{ST}^{C} based Moran's I statistic is larger than that of the V_{ST}^{L} based statistic. This indicates that even though the space-time dependency of the regional Covid-19 percent of cases can be seen as contemporaneous as well as time lagged in nature, the first specification is more probable than the second. In this case, it is expected that the increase in the current local percent of cases not only increases the current neighbouring regencies/municipalities percent of cases, but within the next two weeks, it also increases that of neighbouring regencies/municipalities.

Conclusion

Spatial interaction of the regional variable is the main interest when a set of spatial data is available. However, when the variable has time dependence, analysis based on one time-period spatial data might fail to detect spatial autocorrelation. However, this situation does not imply the absence of spatial interaction. The problem is in the detection method, which does not accommodate time dependence. This situation applies to East Java's regional Covid-19 percent of cases, in which the spatial autocorrelation of the regional Covid-19 percent of cases cannot be detected from each time unit spatial data using existing methods. In contrast to previous methods which are based on one set of spatial data, the index developed in this study works for a set of spatial data such that the dynamic time-dependent nature of the spatial interaction can be captured more accurately. The proposed index confirms the dynamic time-dependent spatial interaction of East Java's regional Covid-19 percent of cases.

Based on the results of the test, this study highlights several implications which are useful for formulating policies to control the spread of this disease. The increase

in the local percent of cases will increase the local as well as the neighbourhood percent of cases, but not instantly. This study finds that two weeks is a reasonable duration for the effect to be fully observed. In this way, the regional percent of cases dynamically changes over time. Consequently, the effect of the implemented policy cannot be observed instantly. It takes some time for the effect to be fully realised. Therefore, the control policy (social restriction, travel ban, etc.) must be implemented continuously, reviewed periodically, and integrated over the local and interregional levels.

Even though the nature of Covid-19 regional percent of cases has been the main motivation behind the development of the index, it will also be useful as a detection method of the dynamic-time dependence spatial interaction of any other regional variables with a similar nature. Further research can be directed to accommodate the time dependence the arises from two observations which are more than one time unit apart.

Acknowledgements

This study was supported by the 2021 internal research fund (*DPP/SPP*) of the Faculty of Mathematics and Natural Sciences, University of Brawijaya, Malang, Indonesia.

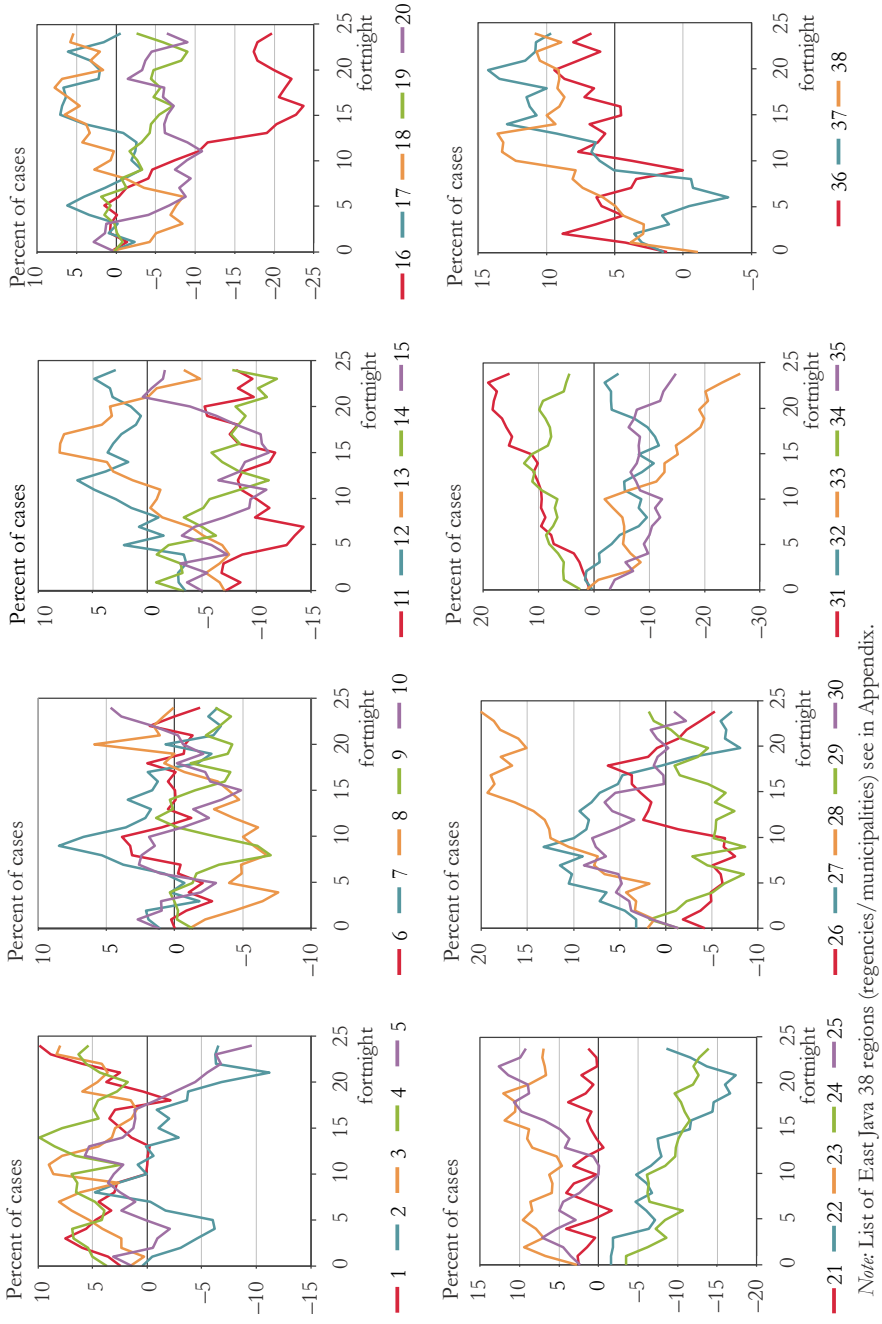
Appendix

List of East Java 38 regions (regencies/municipalities)

Banyuwangi Regency	Mojokerto Kota
Batu Kota	Mojokerto Regency
Bengkalan Regency	Nganjuk Regency
Blitar Regency	Ngawi Regency
Bojonegoro Regency	Pacitan Regency
Bondowoso Regency	Pamekesan Regency
Gresik Regency	Pasuruan Kota
Jember Regency	Pasuruan Regency
Jombang Regency	Ponorogo Regency
Kediri Kota	Probolinggo Kota
Kediri Regency	Probolinggo Regency
Kota Blitar	Sampang Regency
Lamongan Regency	Siduarjo Regency
Lumajang Regency	Situbondo Regency
Madiun Kota	Sumenep Regency
Madiun Regency	Surabaya Kota
Magetan Regency	Trenggalek Regency
Malang Kota	Tuban Regency
Malang Regency	Tulungagung Regency

Figure A1

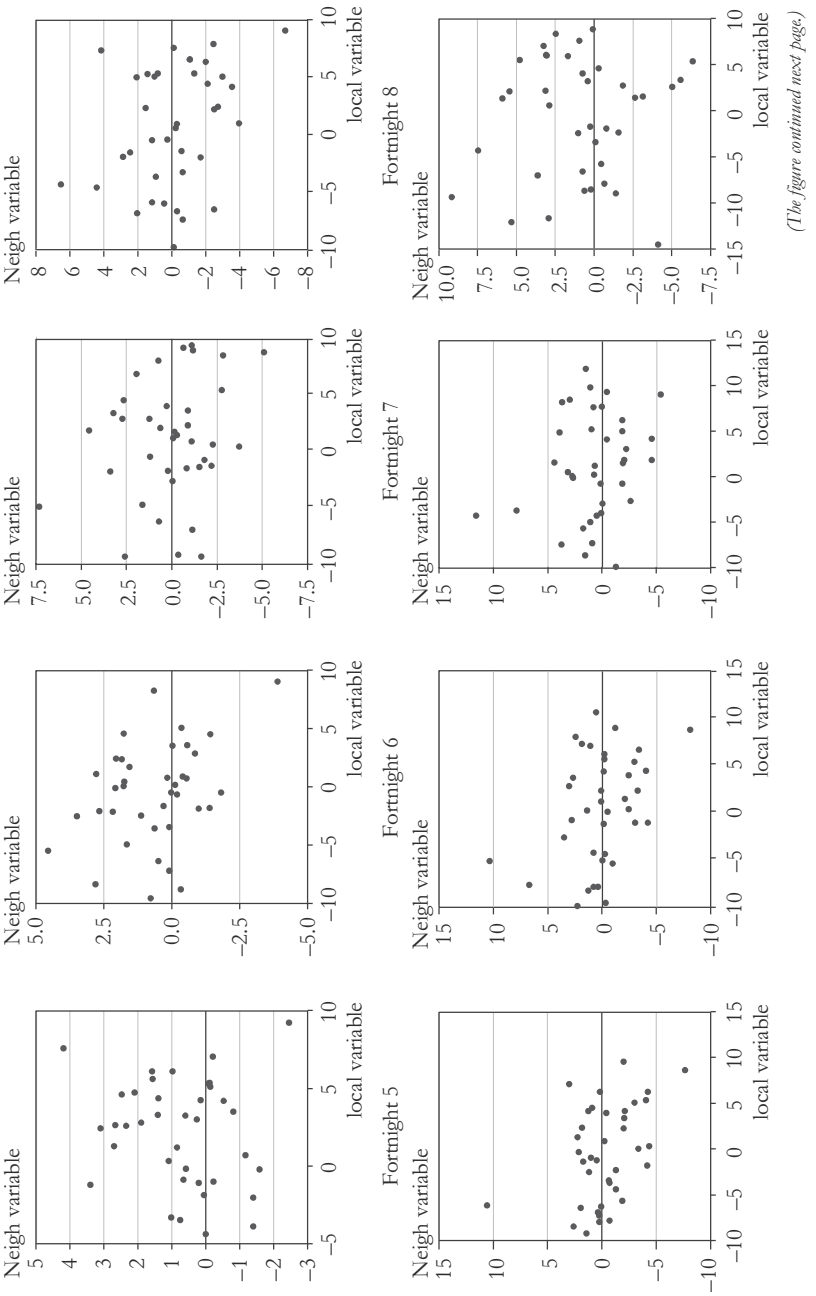
Time series plots for all spatial units, generated data with ρ=0 and ϕ=1



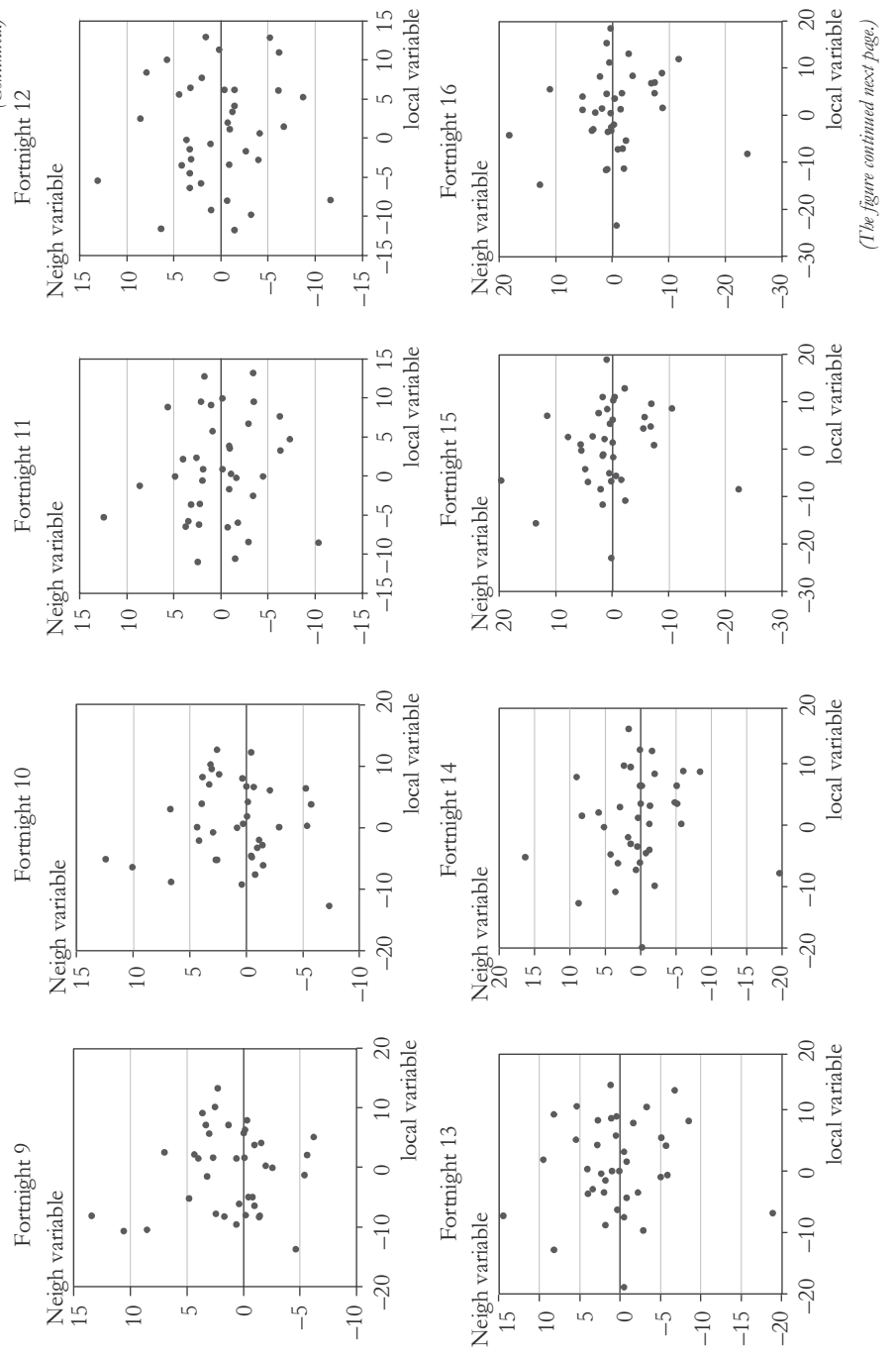
Note: List of East Java 38 regions (regencies/municipalities) see in Appendix.

Figure A2

Plots of Y and WY, visualisation of the spatial dependency for all time unit, generated data with ρ=0 and ϕ=1



(The figure continued next page.)


(Continued.)
Fortnight 9
Fortnight 10
Fortnight 11
Fortnight 12
Fortnight 13
Fortnight 14
Fortnight 15
Fortnight 16
Neigh variable
local variable
(The figure continued next page.)

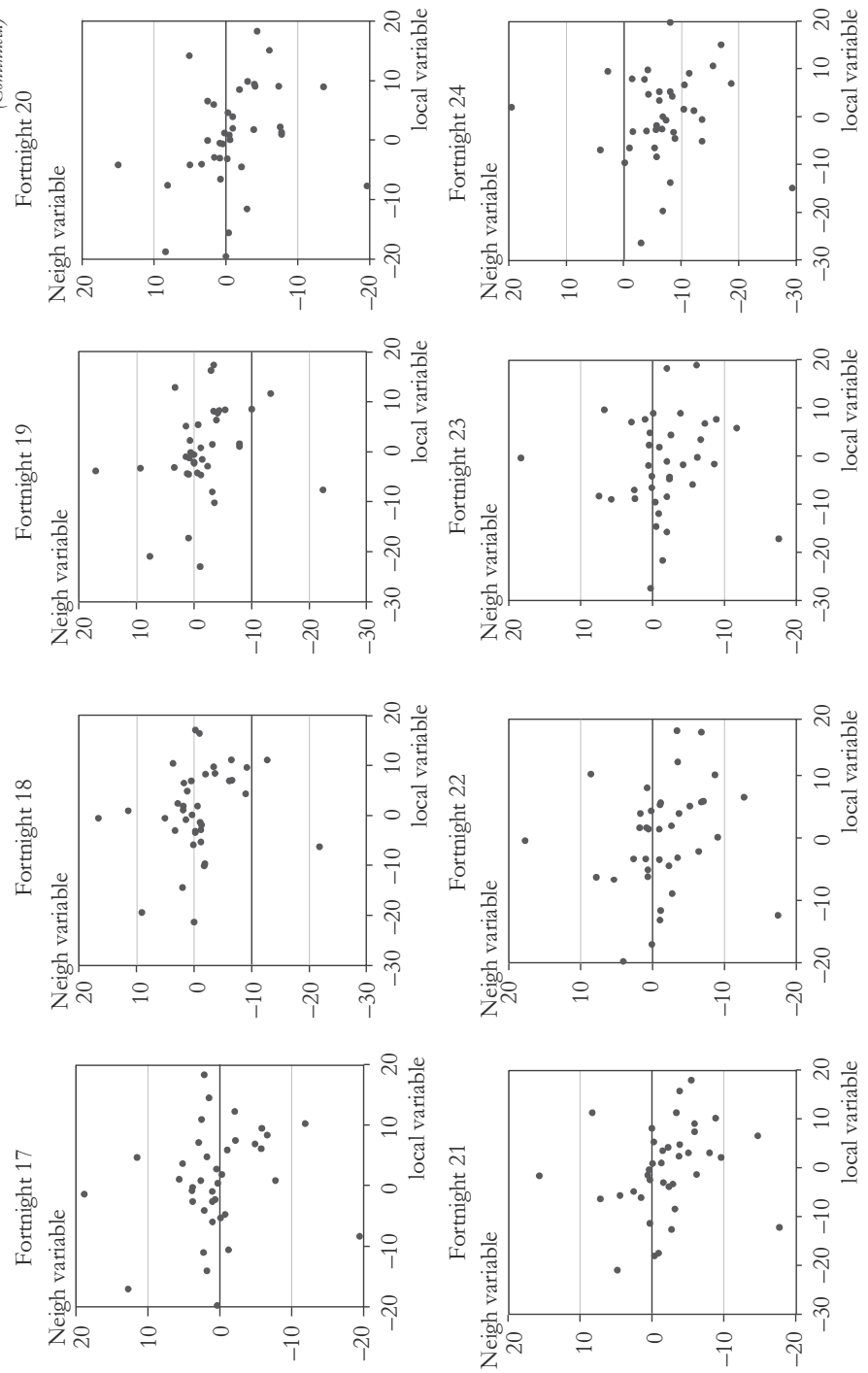
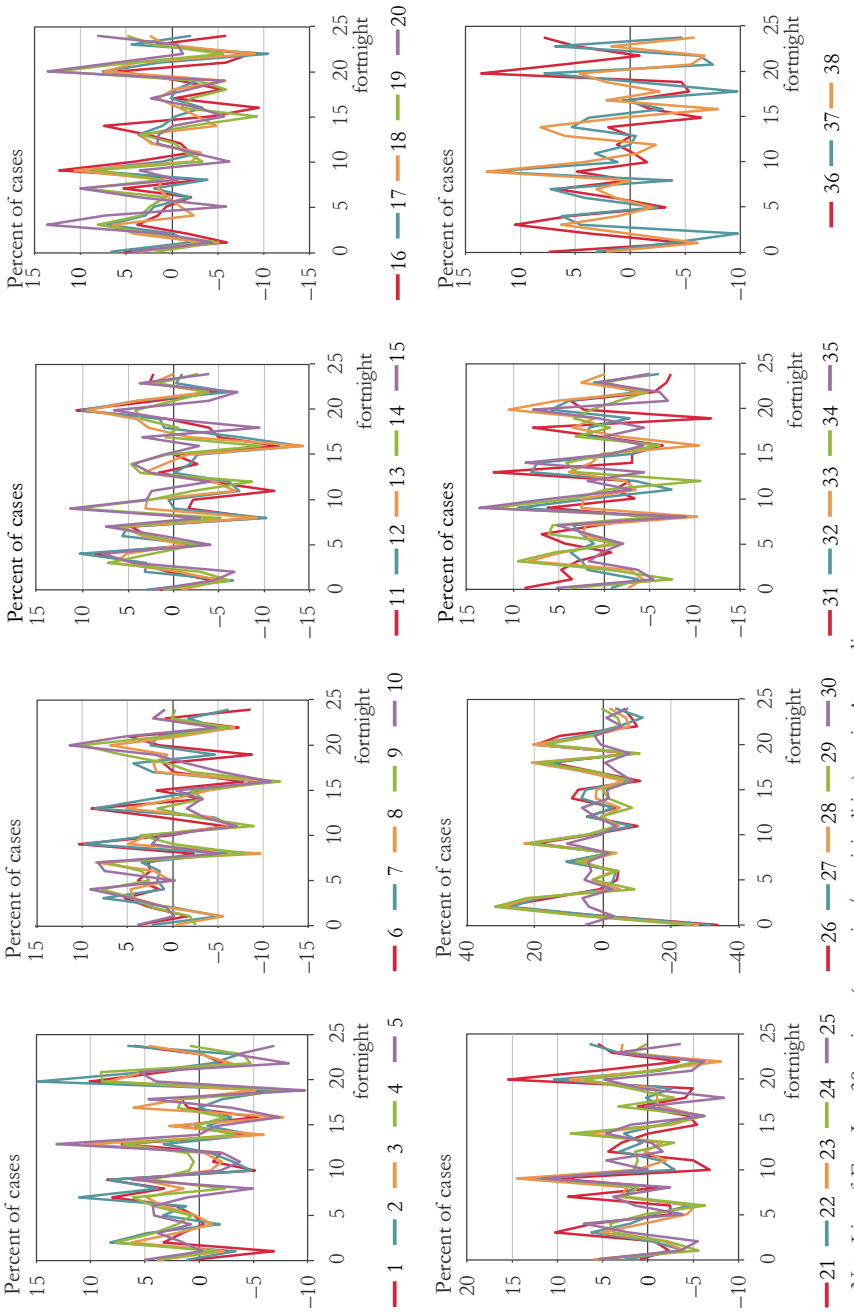

(Continued.)
Fortnight 17
Fortnight 18
Fortnight 19
Fortnight 20
Fortnight 21
Fortnight 22
Fortnight 23
Fortnight 24
Neigh variable
local variable

Figure A3

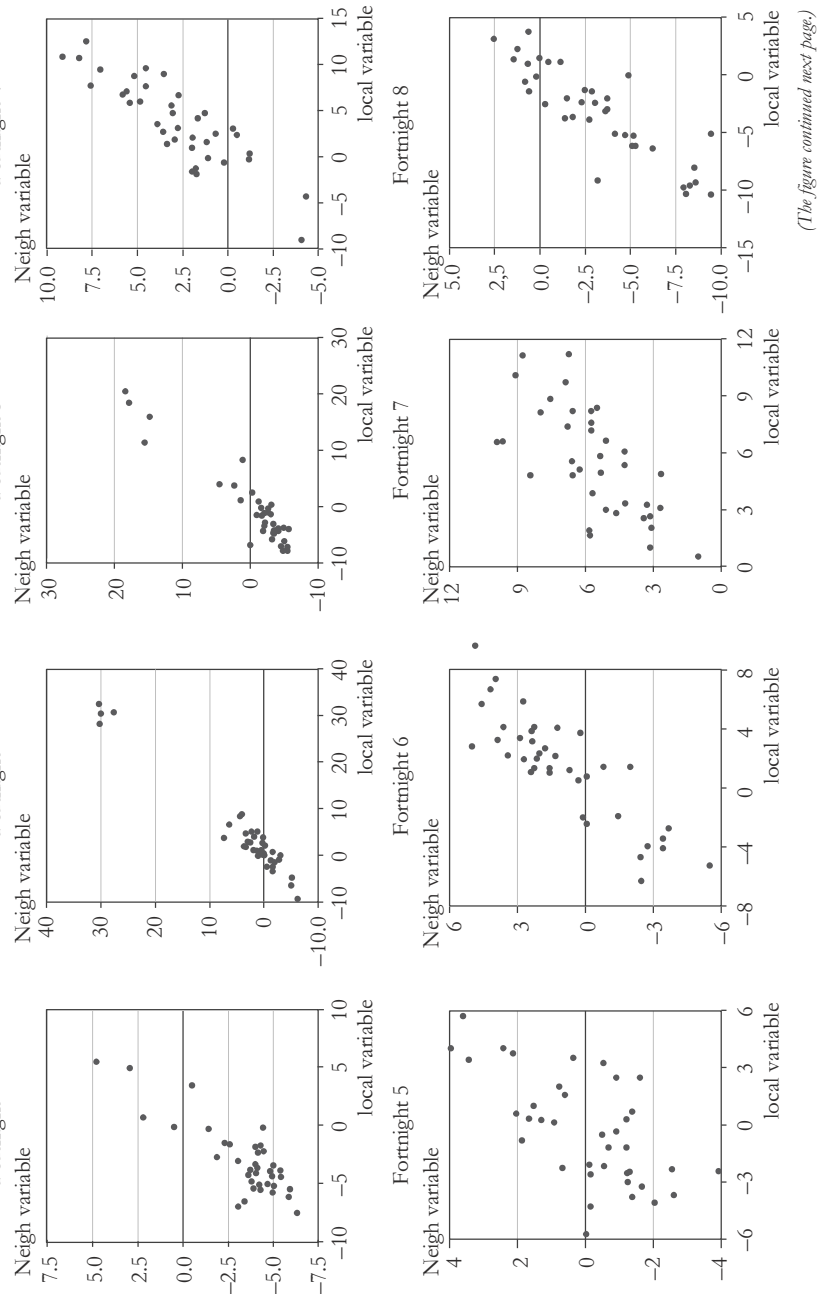
Time series plots for all spatial units, generated data with ρ=0.9 and ϕ=0



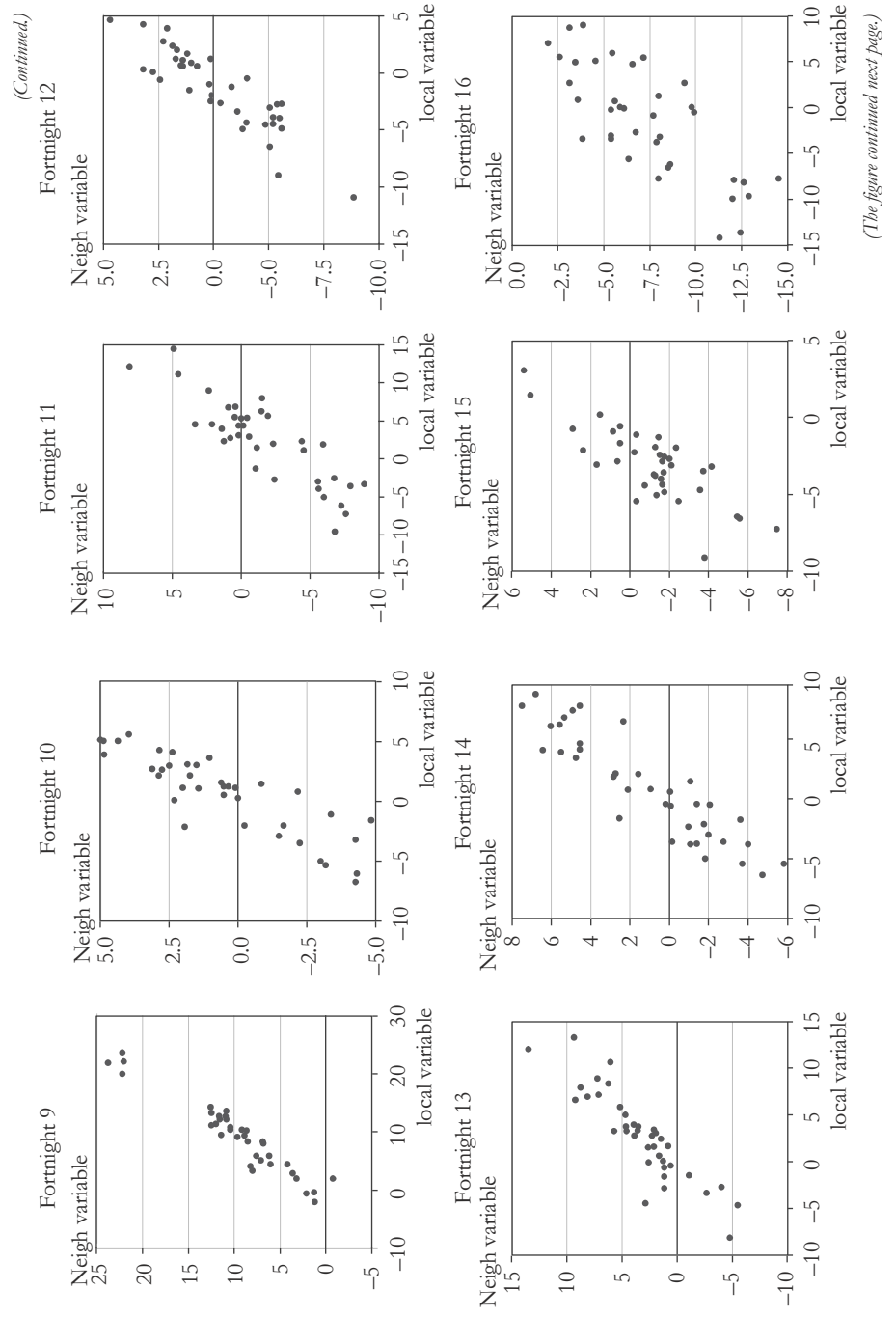
Note: List of East Java 38 regions (regencies/municipalities) see in Appendix.

Figure A4

Plots of Y and WY, visualisation of the spatial dependency for all time unit, generated data with ρ=0.9 and ϕ=0



(The figure continued next page.)


(Continued.)
Fortnight 9
Fortnight 10
Fortnight 11
Fortnight 12
Fortnight 13
Fortnight 14
Fortnight 15
Fortnight 16
Neigh variable
local variable
(The figure continued next page.)

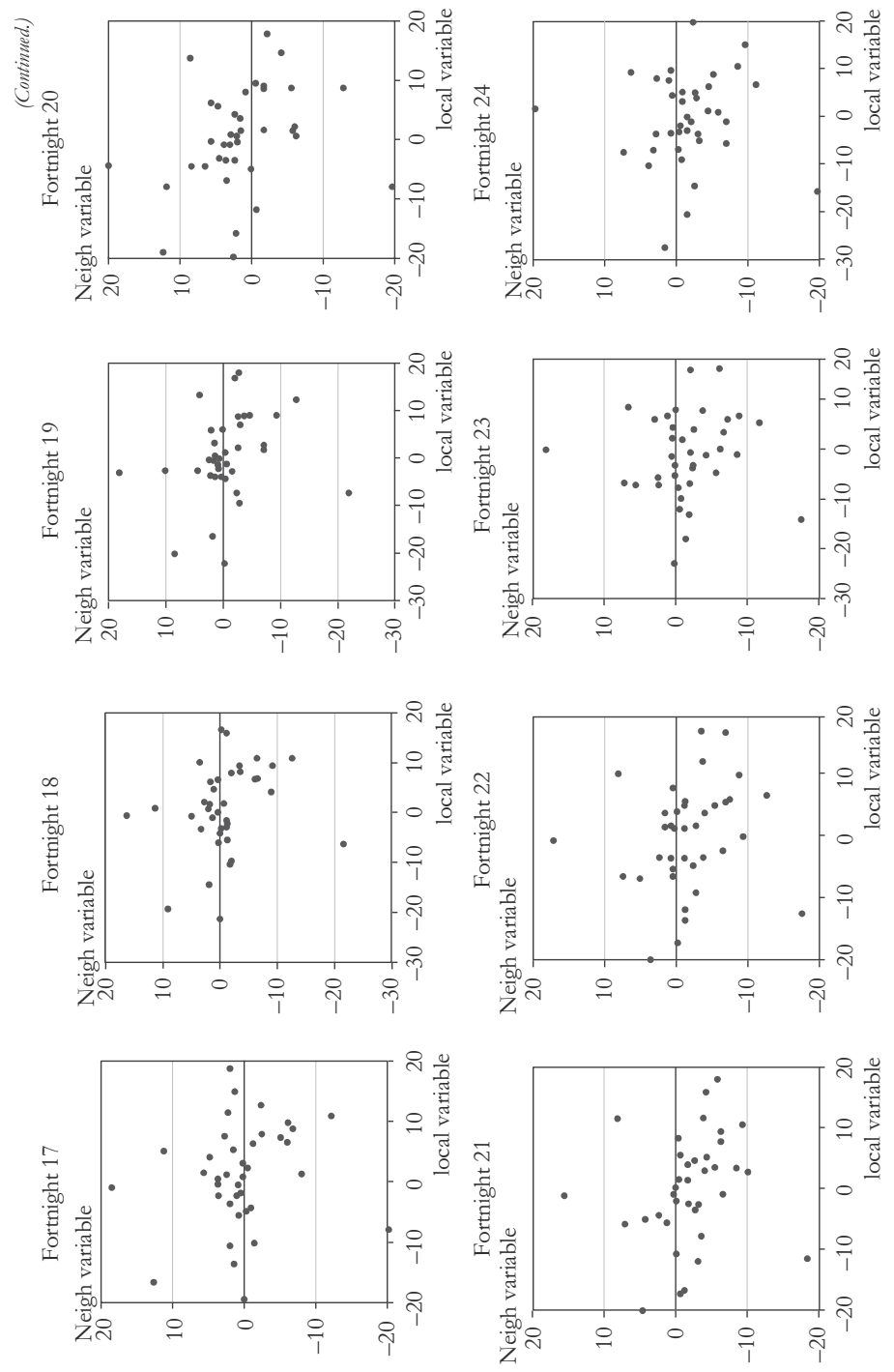

(Continued.)
Fortnight 17
Fortnight 18
Fortnight 19
Fortnight 20
Fortnight 21
Fortnight 22
Fortnight 23
Fortnight 24
Neigh variable
local variable

Figure A5

Percentage of rejection of null hypothesis of
no space-time autocorrelation as a function of ϕ

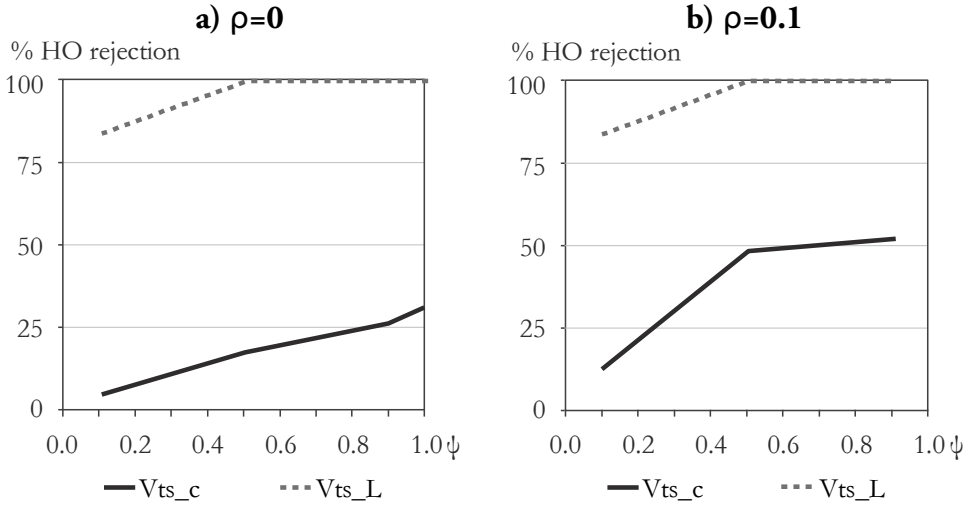


Figure A6

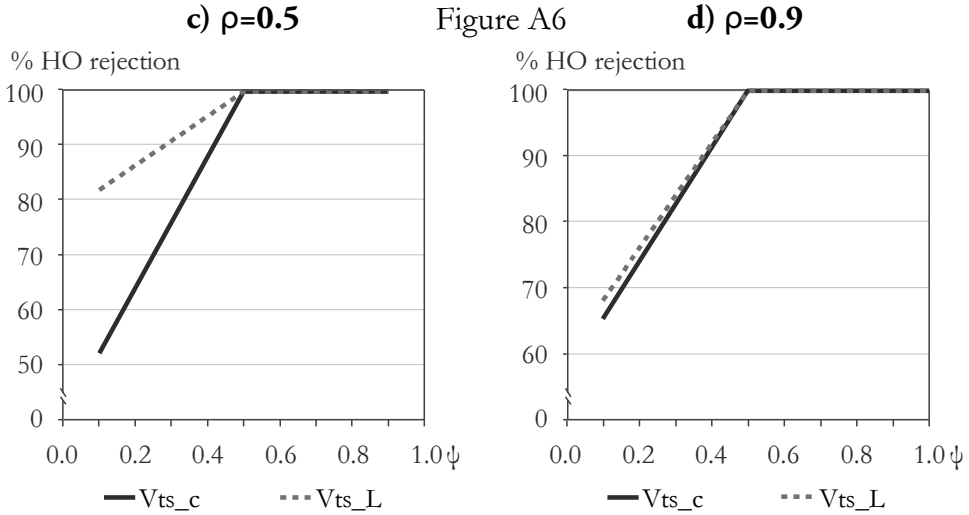


Figure A6

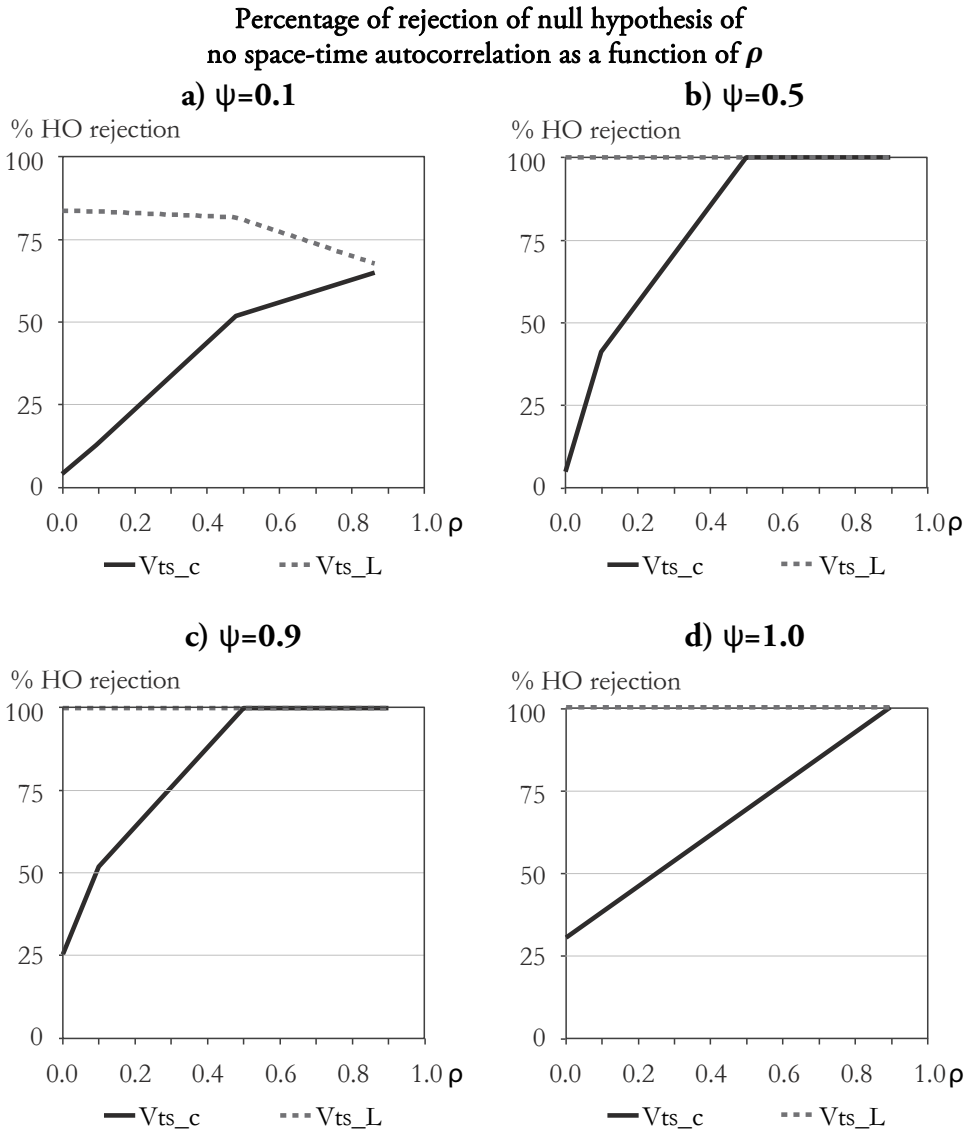
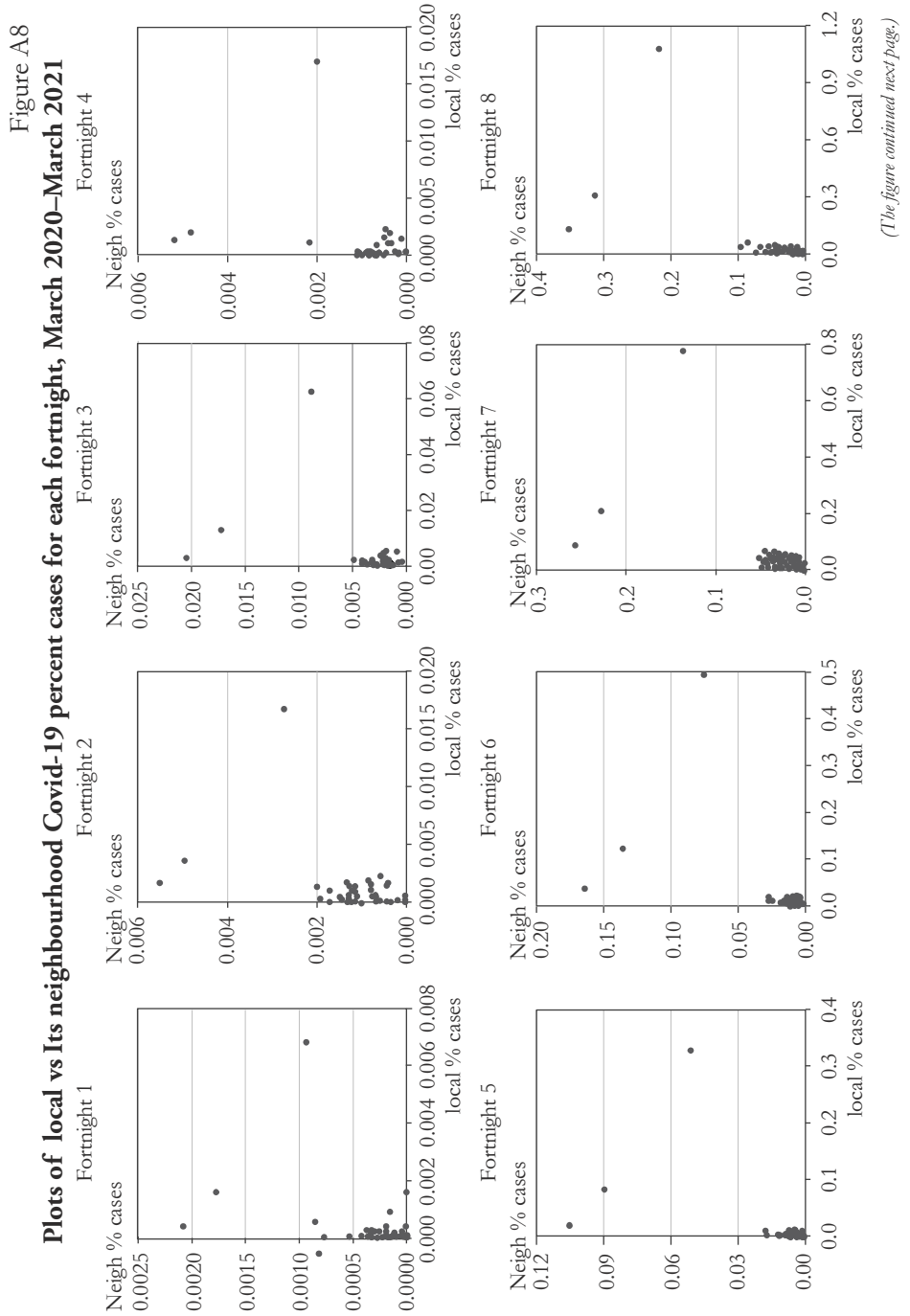

Percentage of rejection of null hypothesis of
no space-time autocorrelation as a function of ρ
a) ψ=0.1
b) ψ=0.5
c) ψ=0.9
d) ψ=1.0
% HO rejection
100
75
50
25
0
0.0
0.2
0.4
0.6
0.8
1.0 ρ
Vts_c
Vts_L

Figure A7

Time series of Covid-19 percent of cases for every regency/municipality, March 2020–March 2021

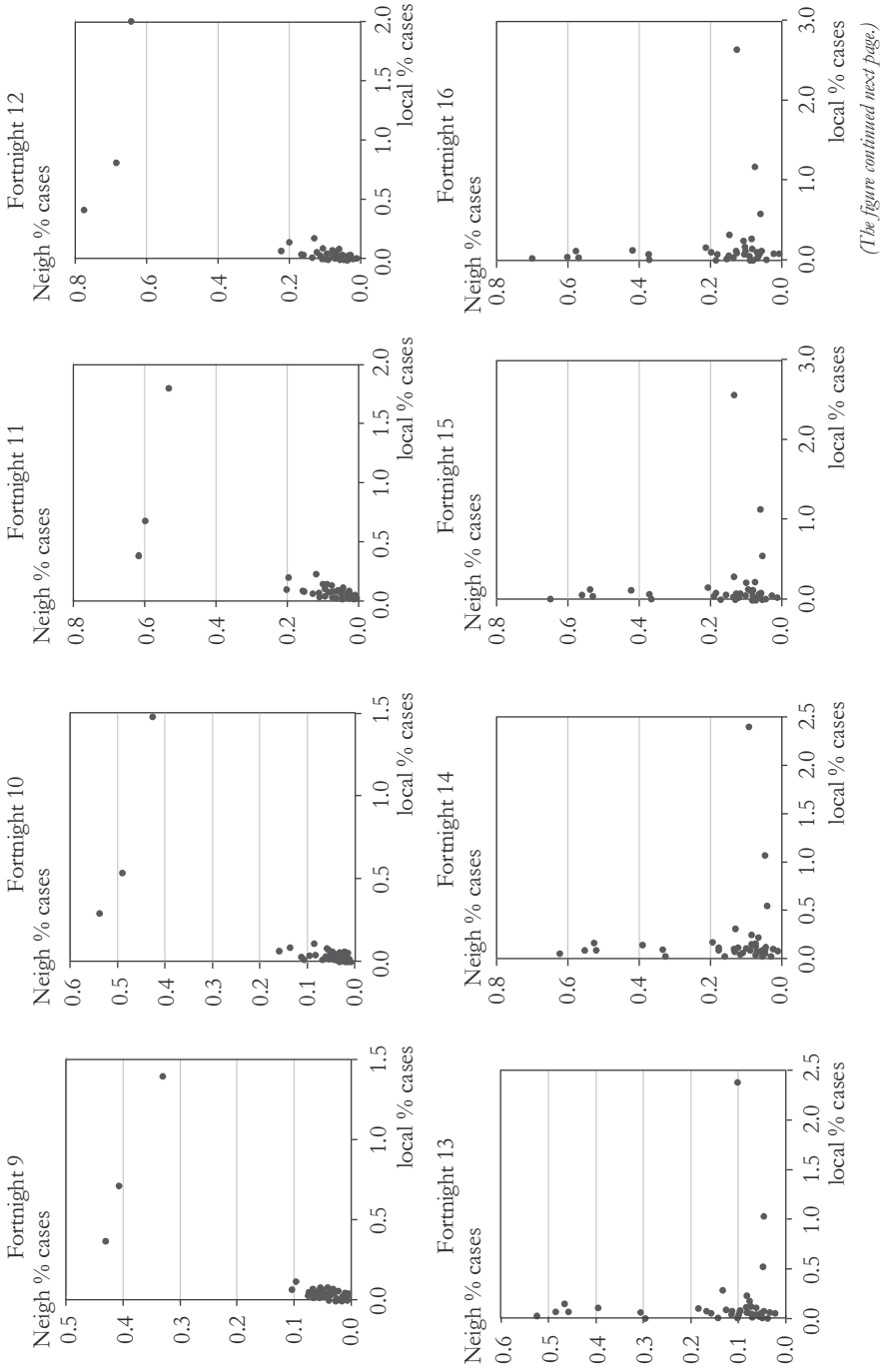
Note: List of East Java 38 regions (regencies/municipalities) see in Appendix.

Figure A8

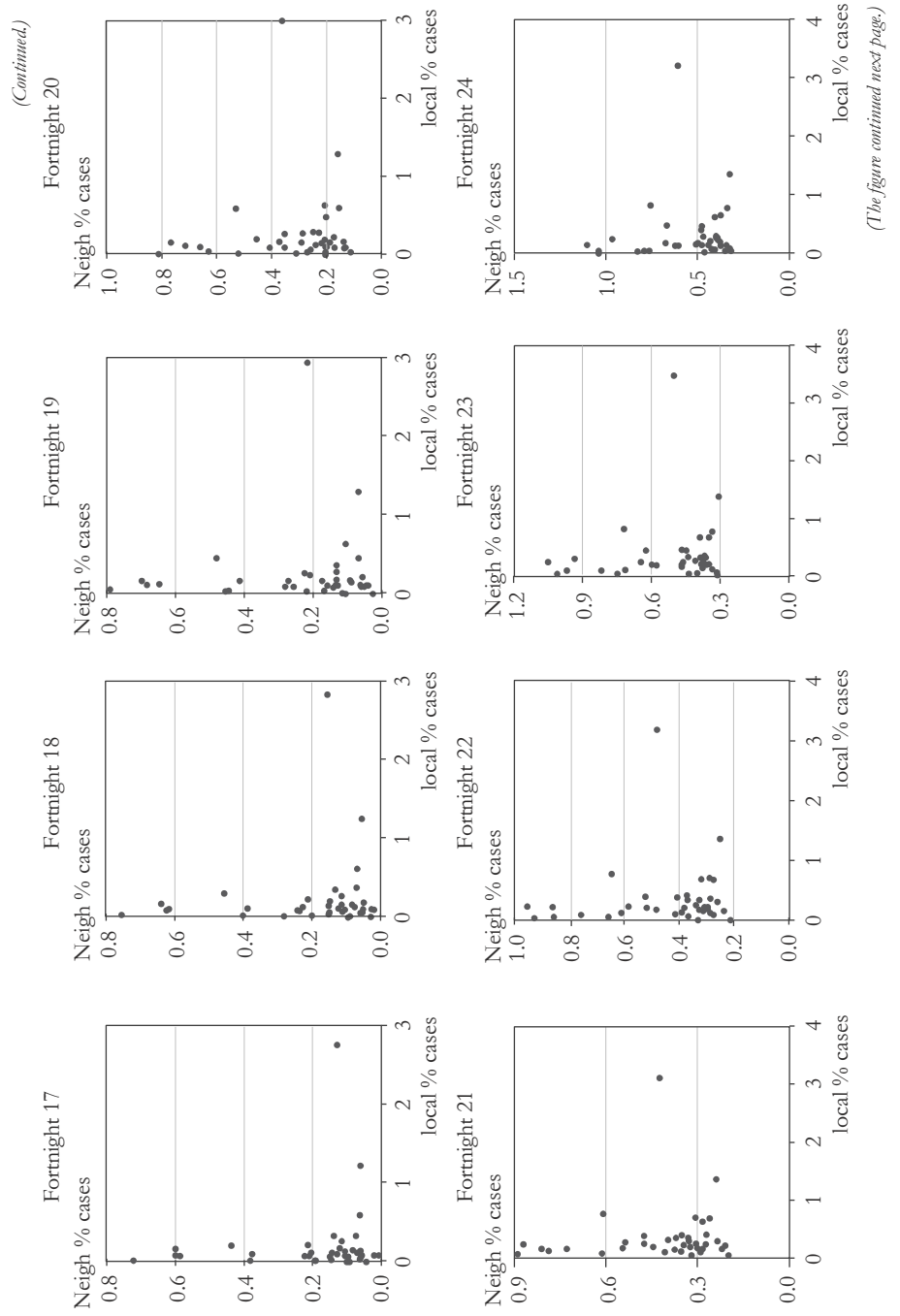


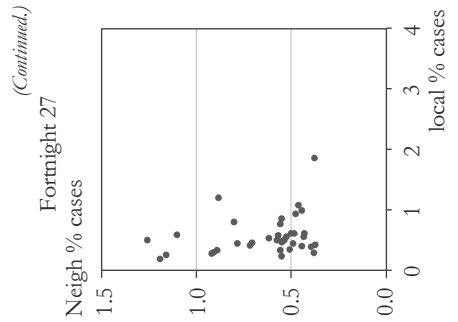
(The figure continued next page.)

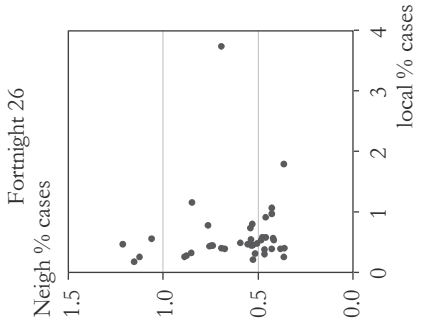
(Continued.)

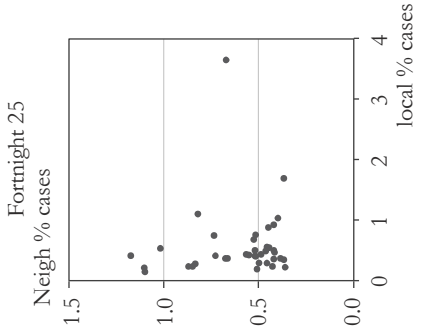


(The figure continued next page.)


(Continued.)
Fortnight 17
Fortnight 18
Fortnight 19
Fortnight 20
Fortnight 21
Fortnight 22
Fortnight 23
Fortnight 24
Neigh % cases
local % cases
(The figure continued next page.)


(Continued.)
Fortnight 27
Neigh % cases
1.5
1.0
0.5
0.0
0
1
2
3
4
local % cases


Fortnight 26
Neigh % cases
1.5
1.0
0.5
0.0
0
1
2
3
4
local % cases


Fortnight 25
Neigh % cases
1.5
1.0
0.5
0.0
0
1
2
3
4
local % cases

REFERENCES

ANSELIN, L. (1988): *Spatial econometrics: methods and models* Kluwer Academic Publishers, Boston, MA.

ANSELIN, L. (1995): Local indicators of spatial association – LISA *Geographical Analysis* 27 (2): 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

ANSELIN, L. (2001): Spatial econometrics. In: BALTAGI, B. H. (ed.): *A companion to theoretical econometrics* pp. 310–330., Blackwell Publishing Ltd., Hoboken, New Jersey.

ARBIA, G. (2006): *Spatial econometrics: statistical foundations and applications to regional convergence* Springer Science & Business Media., Berlin, Heidelberg. https://doi.org/10.1007/3-540-32305-8

BACKER, J. A.–KLINKENBERG, D.–WALLINGA, J. (2020): Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020 *Eurosurveillance* 25 (5): 2000062. https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062

CHEN, Y. (2021): An analytical process of spatial autocorrelation functions based on Moran's index *PLOS ONE* 16 (4): e0249589. https://doi.org/10.1371/journal.pone.0249589

CLIFF, A. D.–ORD, J. K. (1981): *Spatial processes: Models and applications (Pion limited)* Taylor & Francis, London.

CRYER, J. D.–CHAN, K.-S. (2008): Time series regression models. In: *Time series analysis: With applications in R* pp. 249–276., New York: Springer. http://dx.doi.org/10.1007/978-0-387-75959-3

EGRI, Z.–TÁNCZOS, T. (2018): The spatial peculiarities of economic and social convergence in Central and Eastern Europe *Regional Statistics* 8 (1): 49–77. https://doi.org/10.15196/RS080108

ELHORST, J. P. (2014): *Spatial econometrics: From cross-sectional data to spatial panels* Springer Briefs in Regional Science, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40340-8

FITRIANI, R.– PUSDIKTASARI, Z. F.–DIARTHO, H. C. (2021): Growth interdependence in the presence of spatial outliers: Implementation of an average difference algorithm on East Java regional economic growth, 2011–2016 *Regional Statistics* 11 (3): 119–132. https://doi.org/10.15196/RS110306

GAO, Y.–CHENG, J.–MENG, H.–LIU, Y. (2019): Measuring spatio-temporal autocorrelation in time series data of collective human mobility *Geo-Spatial Information Science* 22 (3): 166–173. https://doi.org/10.1080/10095020.2019.1643609

GEARY, R. C. (1954): The contiguity ratio and statistical mapping *The Incorporated Statistician* 5 (3): 115–146.

GETIS, A.–ORD, J. K. (1992): The analysis of spatial association by use of distance statistics *Geographical Analysis* 24 (3): 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

JAYA, I.–ANDRIYANA, Y.–TANTULAR, B.–RUCHJANA, B. N. (2019): Spatiotemporal dengue disease clustering by means local spatiotemporal Moran's index *IOP Conference Series: Materials Science and Engineering* 621 (1): 12017. https://doi.org/10.1088/1757-899X/621/1/012017

KELEJIAN, H. H.–PRUCHA, I. R. (1998): A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances *The Journal of Real Estate Finance and Economics* 17 (1): 99–121. https://doi.org/10.1023/A:1007707430416

KINCSES, Á.–TÓTH, G. (2020): How coronavirus spread in Europe over time: national probabilities based on migration networks *Regional Statistics* 10 (2): 228–231. https://doi.org/10.15196/RS100210

LAUER, S. A.–GRANTZ, K. H.–BI, Q.–JONES, F. K.–ZHENG, Q.–MEREDITH, H. R.–AZMAN, A. S.–REICH, N. G.–LESSLER, J. (2020): The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application *Annals of Internal Medicine* 172 (9): 577–582. https://dx.doi.org/10.7326/M20-0504

LEE, J.–LI, S. (2017): Extending Moran's index for measuring spatiotemporal clustering of geographic events *Geographical Analysis* 49 (1): 36–57. https://doi.org/10.1111/gean.12106

MORAN, P. A. P. (1950): Notes on continuous stochastic phenomena *Biometrika* 37 (1/2): 17–23.

NYIKOS, GY.–SOHA, B.–BÉRES, A. (2021): Entrepreneurial resilience and firm performance during the COVID-19 crisis – Evidence from Hungary *Regional Statistics* 11 (3): 29–59. https://doi.org/10.15196/RS110307

OLIVIA, S.–GIBSON, J.–NASRUDIN, R. (2020): Indonesia in the time of Covid-19 *Bulletin of Indonesian Economic Studies* 56 (2): 143–174. https://doi.org/10.1080/00074918.2020.1798581

QUESADA, J. A.–LÓPEZ-PINEDA, A.–GIL-GUILLÉN, V. F.–ARRIERO-MARÍN, J. M.–GUTIÉRREZ, F.–CARRATALA-MUNUERA, C. (2021): Incubation period of COVID-19: A systematic review and meta-analysis *Revista Clínica Española (English Edition)* 221 (2): 109–117. https://doi.org/https://doi.org/10.1016/j.rceng.2020.08.002

SHEN, C.–LI, C.–SI, Y. (2016): Spatio-temporal autocorrelation measures for nonstationary series: A new temporally detrended spatio-temporal Moran's index *Physics Letters A* 380 (1–2): 106–116. https://doi.org/10.1016/j.physleta.2015.09.039

SURYAHADI, A.–IZZATI, R. A.–SURYADARMA, D. (2020): Estimating the impact of Covid-19 on poverty in Indonesia *Bulletin of Indonesian Economic Studies* 56 (2): 175–192. https://doi.org/10.1080/00074918.2020.1779390

TOBLER, W. R. (1970): A computer movie simulating urban growth in the Detroit region *Economic Geography* 46 (sup1): 234–240. https://doi.org/10.2307/143141

WANG, Z.–LAM, N. S. N. (2020): Extending Getis–Ord statistics to account for local space–time autocorrelation in spatial panel data *The Professional Geographer* 72 (3): 411–420. https://doi.org/10.1080/00330124.2019.1709215

INTERNET SOURCES

ANSELIN, L. (1999): Interactive techniques and exploratory spatial data analysis. In: LONGLEY, P. A.–GOODCHILD, M. F.–MAGUIRE, D. J.–RHIND, D. W. *Geographical information systems: Principles, techniques, management and applications* Vol. 1. pp. 251–264. https://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch17.pdf (downloaded: September 2021).

KEMENKES, R. I. (2021): *Peta Sebaran COVID-19 | Covid19.go.id*. Kementerian Kesehatan RI. https://covid19.go.id/peta-sebaran-covid19 (downloaded: September 2021).

WORLDOMETER.COM (2021). *COVID live update* https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1? (downloaded: September 2021).