Growth interdependence in the presence of spatial outliers: Implementation of an average difference algorithm on East Java regional economic growth, 2011–2016

Rahma Fitriani
(corresponding author)
University of Brawijaya, Malang, Indonesia
E-mail: rahmafitriani@ub.ac.id

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

Herman C. Diartho
University of Jember, Jember, Indonesia
E-mail: hermancahyodiartho.feb@unej.ac.id

Keywords:
regional growth,
spatial outliers,
Global Moran's I,
average difference algorithm

Recent analysis regarding economic growth has started to accommodate growth interdependence. Spatial econometrics, which relies on the assumption of spatially autocorrelated economic growth, is the main tool of analysis. The spatial autocorrelation is commonly measured by Global Moran's I, which is not robust to spatial outliers. The visualisation of available data indicates that East Java has experienced economic growth interdependence, but the Global Moran's I for the GDP growth shows otherwise. The presence of spatial outliers is suspected as the cause of the insignificant spatial interaction. Using an average difference algorithm, this study detects the spatially outlying regencies/municipalities in terms of their yearly economic growth during 2011–2016. Based on the data, this study detects some spatially outlying regions, which mask the significance of the spatial interaction of the GDP growth among the regencies/ municipalities.

Introduction

Recent studies on economic growth have given more attention on the significance of growth interdependence among countries, provinces, or any other administrative regions. They argue that in terms of productivity and economic growth, regions cannot be treated as spatially independent observation units. The interdependence has been associated with the existence of externalities that spill over the administrative boundaries. The source of spatial spillover is the spread of technology and information across regions (Vayá et al. 2004, Fingleton–López-Bazo 2006, Ertur–Koch 2007). The intensity of the technological transfer declines with the geographical distance or with the difference of local conditions between regions.

These studies extend the Solow productivity model by assuming that the local technology depends on the technological level of the neighbouring regions, which implies that the local technology is related to the neighbouring capital investment in technology. This extended productivity model is the basis of defining the theoretical growth model that accommodates the interdependence among regions. The model indicates that local growth is a function of local initial income, neighbourhood initial income, neighbourhood growth, and the local control factors. The empirical version of this growth model falls within a spatial econometric framework, such as a Spatial Durbin Model (SDM).

Several empirical works within the spatial econometric framework show the significance of the growth interdependence (Vayá et al. 2004, Fingleton–López-Bazo 2006, Ertur–Koch 2007, Egri–Tánczos 2018). An analysis of the growth based on the spatial model relies on the degree of growth interdependence, which is measured by a global spatial autocorrelation (Anselin 1988). It is formally defined as a correlation between an attribute value of a location and that of its neighbours. The attribute value in this case is the regional economic growth. Among several measures of the global spatial autocorrelation (e.g. Gamma, Joint count, Geary's C, the variogram, Ripley's K function), the Global Moran's I is the most commonly used measure (Getis 2010). The concept of neighbours based on distance is still the focus in this study, even though other studies have started to implement other concepts, such as network linkages among regions (Járosi 2017) or the intensity of social interaction (Conley–Topa 2002).

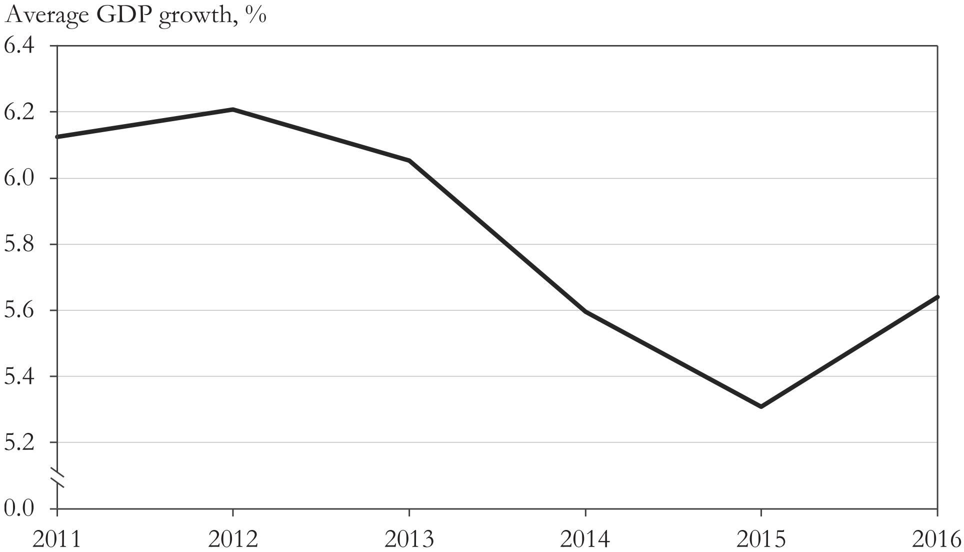
Getis (2010) demonstrates that the Global Moran's I is a Pearson's correlation between the attribute value of a region and the average attribute value of its neighbouring regions. The Pearson's correlation is sensitive to outliers (Kim et al. 2015) and the Global Moran's I suffers the same problem. The outlier in this case is known as a spatial outlier. It is an observation that appears to be inconsistent with the observations in its neighbourhood. Theoretically, the presence of a spatial outlier may lead to the negative Moran's I, even though most of the nearby regions have similar attribute values. Consequently, the significance of the statistic will also be affected.

The spatial outliers presumably affect the significance of the Global Moran's I of the economic growth among East Java's regencies/municipalities. With Surabaya as its capital city, this province is one of the Indonesian provinces with high economic growth. During 2011–2016, it experienced the highest average gross domestic product (GDP) growth of 6.21% (in 2012) and the lowest average GDP growth of 5.13% (in 2015). Fitriani et al. (2020) reveal that the high economic growth in this province creates a certain degree of growth disparity. The province consists of 38 regencies/municipalities, which are in the eastern part of Java Island and the island in its north east, Madura Island. Among the 38 regencies/municipalities, four regencies (Bangkalan, Sampang, Pamekasan and Sumenep) are on Madura Island.

The Suramadu (Surabaya–Madura) bridge, which is cable-stayed bridge, has been operated since 2009 to connect the two islands of Surabaya and Bangkalan across the Madura Strait. The dynamic of GDP growth is depicted in Figure 1. Clusters of regencies/municipalities with similar GDP growth are observed from the distribution map of the 2011 GDP growth (see Figure 2). These clusters indicate the presence of economic growth interdependence. However, when the Global Moran's I test is conducted for yearly growth (2011–2016), the null hypothesis is rejected only for 2011 growth among 2 and 3 nearest regencies/municipalities (see Table 1). The results do not provide enough empirical evidence to support the theory of the growth interdependence in this province.

Figure1

Dynamics of average GDP growth in East Java's regencies/municipalities



Motivated by these results, this study raises several questions. Does the insignificance of the spatial autocorrelation of East Java's economic growth in some of these years indicate the absence of economic interdependence among the regencies/municipalities? Is it possible that the insignificance of the Global Moran's I is due to the presence of spatial outliers? Spatial outliers in this case are regencies/municipalities with distinct economic growth (relatively high or relatively low) from that of their respective neighbouring regencies/municipalities. When the presence of spatial outliers masks the significance of the growth interdependence, the null hypothesis of the Moran test will be mostly not rejected. The result leads to the choice of a non-spatial technique to model the economic growth. This model cannot explain the growth interdependence. The presence of outliers can also indicate variance instability of the underlying spatial process. Any effort to analyse the regional economic growth based on the spatial econometrics models gives inaccurate conclusions due to the violation of spatial variance homogeneity.

This study argues that the insignificant global spatial autocorrelation of the economic growth among East Java's regencies/municipalities is due to the presence of spatial outliers. This study detects the spatially outlying regencies/municipalities in terms of their yearly economic growth based on 2011–2016 data. Kou's (2006) average difference algorithm is used as a spatial outlier detection method.

Figure 2

Distribution of 2011 growth each regency/municipality in East Java, %

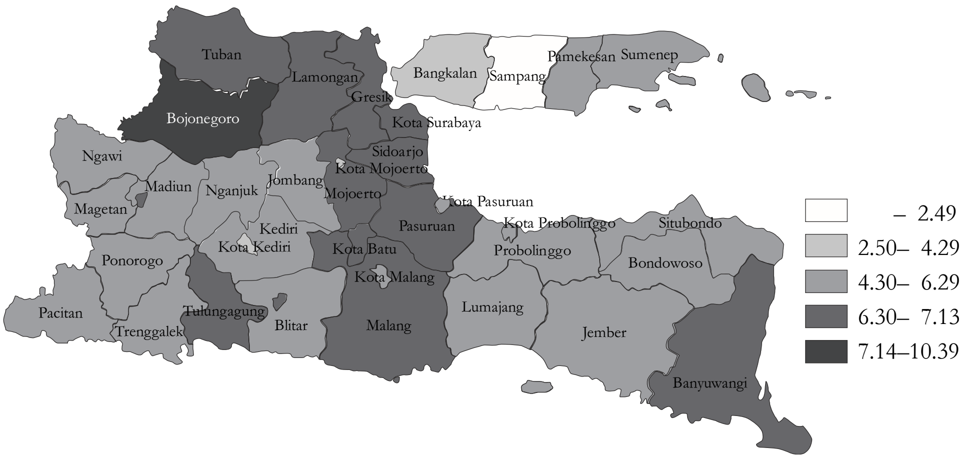


Table 1

Moran's I of yearly growth of GDP for each *k* nearest neighbours (*k*=2, 3, 4, 5) and its significance

Year	Growth of GDP							
	k nearest neighbours							
	2		3		4		5	
	Moran's I	p value	Moran's I	p value	Moran's I	p value	Moran's I	p value
2011	0.221	0.021*	0.099	0.098*	0.079	0.103	0.061	0.126
2012	–0.040	0.397	–0.066	0.277	–0.064	0.26	–0.042	0.36
2013	–0.044	0.368	–0.0334	0.439	–0.022	0.491	–0.017	0.475
2014	–0.028	0.471	–0.0129	0.398	–0.025	0.455	–0.003	0.342
2015	0.068	0.117	0.0143	0.337	0.033	0.195	0.026	0.213
2016	–0.048	0.204	–0.033	0.308	–0.025	0.389	–0.024	0.392

(*) for significant Moran's I.

Spatial outliers and detection methods

A spatial outlier is an object with spatial reference, whose non-spatial attribute value is significantly different from those of other spatially referenced objects in its neighbourhood (Sun–Chawla 2004, Chawla–Sun 2006, Chen et al. 2008, Shekhar et al. 2009). The neighbourhood is defined based on the location of the objects under study, using spatial relationships such as distance or adjacency. Chen et al. (2008) categorises the available spatial outlier detection methods into graphical and quantitative approaches. The graphical approach is based on visualisation of the data, which differentiates the spatial outliers from the rest of the data. Variogram clouds and pocket plots are the techniques in this category. The quantitative approach distinguishes the spatial outliers from the remainder of the data using tests, and provide the visualisation using a Moran scatterplot.

The spatial outlier detection methods generally assign a degree of outlierness to each observation. It is formed by comparing the attribute value of a location with the value of the neighbourhood function (i.e. mean, median). The comparison can be done by via ratio or difference between these two values. The methods are then identify the spatial outlier as an observation with the degree of outlierness that is beyond a defined threshold value (Lu et al. 2003, Sun–Chawla 2004, Chawla–Sun 2006, Kou et al. 2006).

In some cases, a spatially referenced observation is considered significantly different from its neighbouring observations based on not only a single non-spatial attribute value but on multiple non-spatial attribute values. Sun–Chawla (2004), and Chawla–Sun (2006) have accommodated this possibility. They also accommodate the presence of spatial autocorrelation and spatial heteroscedasticity, which are the main characteristics of the spatial data. However, they did not elaborate the formation of the neighbourhood set, when in practice the definition of such set plays an important role in detecting the outliers. The wrong definition of the set may lead to identify the false spatial outliers, while the true spatial outliers are ignored.

In response to the major drawback of the available detection approaches, Lu et al. (2003) use the k nearest neighbours to define the neighbouring set. Median is used instead of average as the neighbourhood function. This is their attempt to reduce the negative impact caused by the presence of neighbouring locations with extremely high/low attribute values. The use of simple average or median is an approach that considers every member of the neighbourhood set has equal contribution. In fact, the closer the distance, or the longer the shared border between two locations, the more intensive the interaction between them. Kou et al. (2006) define this as the impact of the spatial relationship on the neighbourhood comparison. They accommodate this impact in the neighbourhood function, such that each pair of locations has different weight.

The spatial outlier detection method developed by Kou et al. (2006) is known as the averaged difference algorithm. It maintains the spatial variability by calculating the difference between the non-spatial attribute of a location and that of each of its neighbours. The degree of the spatial relationship for each neighbour is defined by weight, which is used further to calculate the weighted average of the difference between a location's non-spatial attribute and the attribute of its neighbours. The standardised weighted average of each location defines the outlierness factor of the location's non-spatial attribute.

Methodology

This study uses regency/municipality–level yearly GDP growth for all 38 regencies/municipalities in the province over 2011–2016, from BPS *(Badan Pusat Statistik)* (2017a). All data are expressed in 2016 constant price. The outlying regencies/municipalities in terms of their yearly GDP growth are identified using the average difference algorithm. It is a spatial outlier detection method based on a weighted average of the difference between the non-spatial attribute of location i and that of each of its neighbours. Kou et al. (2006) define the possibility of combining several spatial relationships, such as the length of shared border and inverse distance, to define the weight, which is larger the more intense is the relationship between a pair of regions. This study adopts the weight's definition, but it only uses one measure of spatial relationship, namely the inverse distance. It also requires the definition of k nearest neighbours. The two nearest neighbours are chosen since the most significant spatial autocorrelation is observed among the 2011 GDP growth of two nearest regencies/municipalities (see the significance of Global Moran's I in Table 1).

The algorithm is implemented as follows:

1. Define the yearly GDP growth of each regency/municipality as the observed non-spatial attribute:
$$y_i,\ i = 1,2,\ldots,38.$$
2. Calculate the difference of the yearly GDP growth between regency/municipality i and its nearest neighbours:
$$diffji = | y_i - y_j |$$
$$i = 1,2,\ldots,38$$
$$j \in J_k(\mathrm{i})$$
in which J_k(i) is the set of k nearest neighbours of regency/municipality i.
3. Calculate the weight of every regency/municipality j which is the member of Jk(i), based on the following formula:
$$w_{ji} = \frac{invDist_{ji}}{\sum_{l \in J_k(i)} invDist_{li}}, \text{ for } j \in J_k(\mathrm{i}),\ i = 1,2,\ldots 38,$$

The value of this weight is between 0 and 1, and by definition $\sum_{j \in J_k(i)} w_{ji} = 1$.

4. Calculate the weighted average difference of every regency/municipality:

$$M_w(i) = \frac{\sum_{j\epsilon J_k(i)} w_{ji}\ diff_{ji}}{\sum_{j\epsilon J_k(i)} w_{ji}} = \sum_{j\epsilon J_k(i)} w_{ji}\ diff_{ji}$$

$$i = 1,2,\ldots,38$$

A regency/municipality with large average difference is indicated as a spatial outlier.

5. Identify the outlying status of each regency/municipality. Regency/municipality i is identified as a spatial outlier if:

$$|M_w(i) - \hat{\mu}_w| > Z_{\alpha/2}\sigma_w \text{ for } i = 1,2,\ldots 38,$$

This criterion is derived by assuming that $M_w(i) \sim N(\mu_w, \sigma_w^2)$, in which:

$$\hat{\mu}_w = \frac{\sum_{i=1}^{n} M_w(i)}{n}$$

$$\hat{\sigma}_w = \frac{\sum_{i=1}^{n} (M_w(i) - \hat{\mu}_w)^2}{n-1}$$

Results

The implemented algorithm detects some outlying regencies/municipalities in terms of their GDP growth. The results in Table 2 show that, based on 2011–2016 data, the algorithm identifies one or some regions/municipalities with relatively different yearly GDP growth from their neighbouring regencies/municipalities.

Table 2

The identified outlying regencies/municipalities in terms of GDP growth

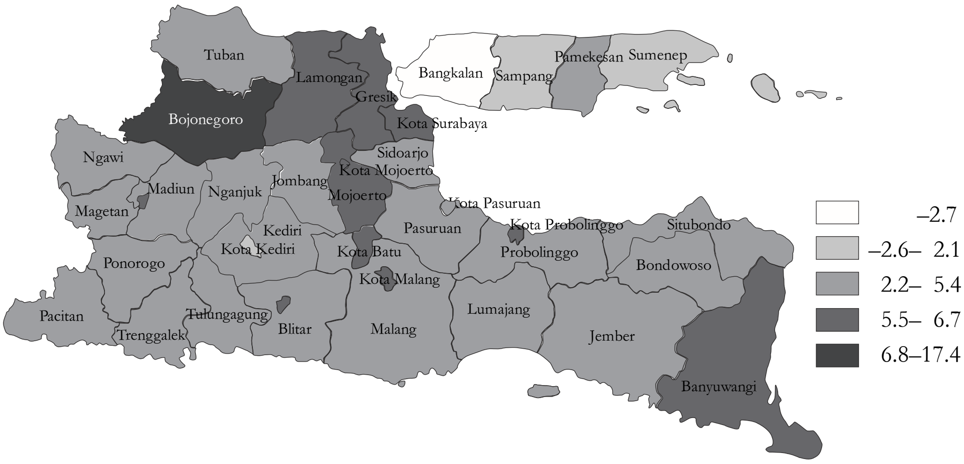
Year	Outlying regencies/municipalities
2011	Bojonegoro
2012	Bangkalan, dan Sumenep
2013	Bangkalan, dan Sumenep
2014	Bangkalan, dan Sampang
2015	Bojonegoro, Tuban, dan Bangkalan
2016	Bojonegoro

To visually understand the nature of the spatial interaction of the growth among regencies/municipalities, we create a distribution map of yearly GDP growth. The map is made based on classifying the yearly GDP of each regency/municipality, in which a different colour is assigned for each class. There are several classification methods for statistical mapping. Jenks' (1967) natural breaks method is known to best represent the spatial characteristic of values. Using this method, the classification of GDP growth is made such that the variance between classes is maximised and that within classes is minimised. In this way the clusters of regencies/municipalities according to their GDP growth are made naturally. The distribution map is compared to a map that represents the outlier status of every regency/municipality. Since the yearly pattern of the growth is similar, only the distribution map for 2015 is presented here (see Figure 3). This year is chosen because among the available years, the algorithm detects the most outlying regencies in terms of their 2015 growth.

Figure 3

GDP growth, 2015

(a) Distribution, %



(b) Outlier status of every regency/municipality

Table 3

Moran's I of the yearly GDP growth and the corresponding p value before and after deleting outlier(s) (using the 2 nearest neighbours concept)

Year	Before deleting outlier(s)		After deleting outlier(s)	
	Moran's I	P value	Moran's I	P value
2011	0.22	0.021	0.26	0.006
2012	–0.04	0.397	0.189	0.001
2013	–0.044	0.368	0.165	0.001
2014	–0.028	0.471	0.215	0.001
2015	0.068	0.117	0.065	0.006
2016	–0.049	0.204	–0.185	0.001

To shed light on the nature of the yearly economic growth interaction without the detected spatially outlying regencies/municipalities, the outliers are excluded from the analysis. The Global Moran's I is calculated for each year of economic growth based on the remaining regencies/municipalities. This approach is possible without losing the spatial reference, since the spatial weights are adjusted using the set of the remaining regencies/municipalities, which is a subset of the complete regions. Anselin (1996) refers to these statistics as a regionalised Moran's I. The comparison of the Moran's I (the value and the significance) before and after deleting the outliers in terms of yearly economic growth (2011–2016) of East Java's regencies/municipalities is presented in Table 3.

Discussion

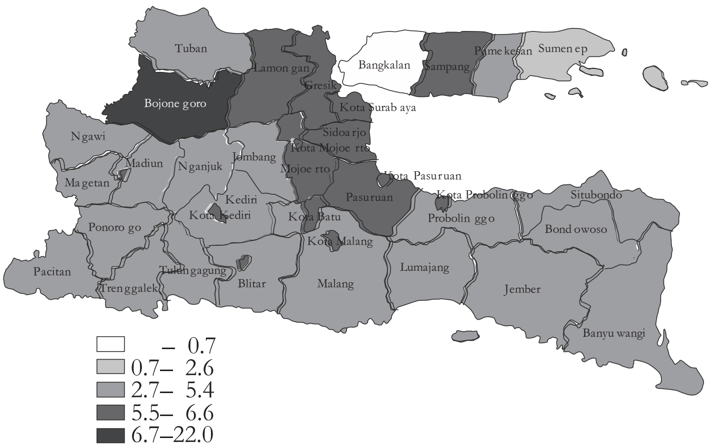
Several years' data of GDP growth (2011–2016) of East Java's regencies/ municipalities were used to explore the dynamic of the spatial dependency among regions in terms of their yearly economic growth. Among the observed years, only 2013 data shows a distinct pattern of GDP growth interaction among the regencies/municipalities (see Figures 1 and 4). The patterns of the economic interaction for the rest of the observed years are not much different. A dynamic change in GDP growth has been observed in Surabaya, the province's capital city, and its nearby regencies/municipalities. These regions always fall into the medium and high category of economic growth. It indicates the significant role of Surabaya as the main centre of growth in the province.

Even though clusters of regencies and municipalities with similar economic growth are observed during 2011–2016 (see Figures 1 and 4), the spatial autocorrelation is significant only for 2011 GDP growth (see Table 3). The results in Table 3 indicate that after deleting the detected outlying regencies/municipalities in terms of their economic growth, the spatial autocorrelation of the GDP growth among the remaining regencies/municipalities becomes significant. This confirms the main argument of this study, in which the insignificant global spatial autocorrelation of the GDP growth among East Java's regencies/municipalities is due to the presence of the spatial outliers. Apart from the detected spatial outliers, East Java has experienced significant yearly economic growth interaction among its regencies/municipalities. This phenomenon supports the argument of economic interdependence among regions or economic spatial interaction.

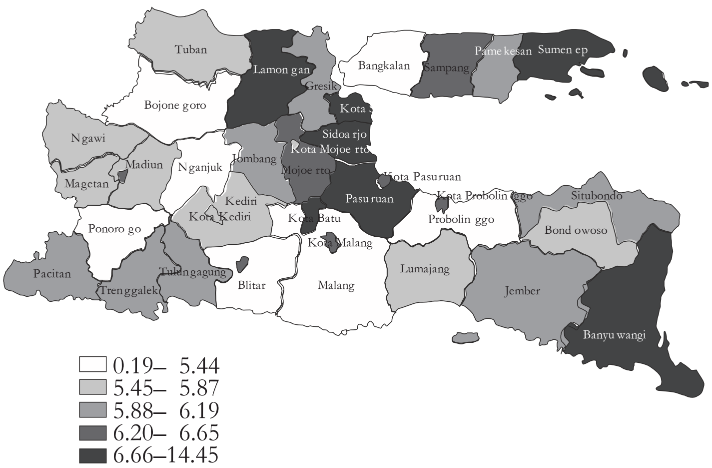
Figure 4

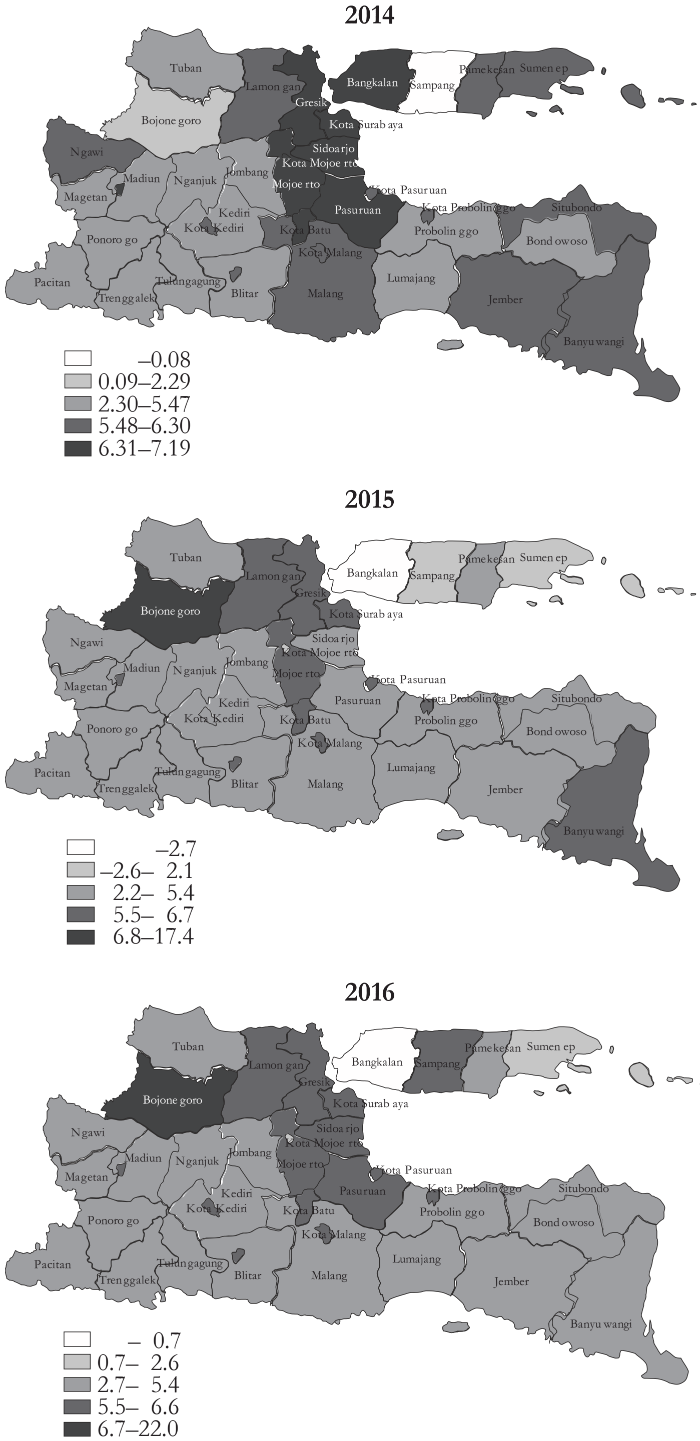
Distribution map of GDP growth for East Java's regencies/municipalities, %

2012



2013




2014
Tuban
Lamon gan
Bangkalan
Sampang
Pamekasan
Sumen ep
Gresik
Bojone goro
Kota Surab aya
Ngawi
Sidoa rjo
Kota Mojoe rto
Jombang
Madiun
Nganjuk
Mojoe rto
Magetan
Kota Pasuruan
Kediri
Pasuruan
Kota Probolin ggo
Situbondo
Kota Kediri
Kota Batu
Probolin ggo
Ponoro go
Bond owoso
Kota Malang
Pacitan
Tulungagung
Lumajang
Trenggalek
Blitar
Malang
Jember
Banyu wangi
–0.08
0.09–2.29
2.30–5.47
5.48–6.30
6.31–7.19
2015
–2.7
–2.6– 2.1
2.2– 5.4
5.5– 6.7
6.8–17.4
2016
– 0.7
0.7– 2.6
2.7– 5.4
5.5– 6.6
6.7–22.0

During these observed years, Bojonegoro and the regencies in Madura Island are often detected as spatial outliers in terms of their respective yearly GDP growth. Some following facts justify the result:

– All areas in Bojonegoro have become the focus of exploration and exploitation of oil and gas (Brata et al. 2017). In fact, it is estimated that 20% of Indonesia's oil is in Bojonegoro Regency (Ardhyanti–Hanif 2014). The oil and gas sector contributes almost 50% of the region's economic growth (Qurbania et al. 2020). As of 2010–2011, Bojonegoro receives Profit-Sharing Fund from its oil and gas production. The fund significantly increases its 2011 GDP (Ardhyanti–Hanif 2014). After 2011, even though the Profit-Sharing Fund from oil and gas production decreases, Bojonegoro's GDP is always considered higher than its neighbouring regions.
– The regencies of Madura Island have different cultural and socio-economic conditions than the rest of East Java. The presence of the Suramadu Bridge, which connects Surabaya and Bangkalan, and the airport in Sumenep have not succeed in reducing the gap. The economic performance of the regions of Madura Island is most affected by their relatively low Human Development Index. During the observed years, of all East Java's regencies/municipalities, the regencies of Madura Island always belong to the bottom three in terms of their Human Development Index (BPS 2017b).

Summary and conclusions

The results of this study answer the research questions regarding the nature of spatial interaction of the economic growth among East Java's regencies/ municipalities. The insignificance of the spatial autocorrelation among the regencies during the period is due to the presence of spatial outliers.

In practice, information regarding the spatial outliers, such as Bojonegoro, Bangkalan, Sampang, and Sumenep for East Java's economic growth will be useful for policy formulation. Any regional policy to promote regional development in these regions should be treated according to their unique characteristics. Furthermore, the policy can utilise the nature of spatial interaction for the remaining regencies. It should consider the contagious nature of the effect of any implemented economic policy between neighbouring regencies/municipalities.

In terms of econometrics analysis, without proper identification of the spatial outliers, since the test provides no proof for the spatial autocorrelation, the analysis may proceed without accommodating the spatial interaction. The analysis will be based on a non-spatial technique when the underlying relationship involves the spatial interaction. This of course leads to misleading results. Therefore, when there is a need to properly address the underlying spatial interaction, the spatial outliers should be excluded from the analysis, or treated differently. Alternatively, the presence of outliers may indicate variance instability of the underlying spatial process. In this case, a spatial econometrics technique that accommodates spatial heterogeneity of variance assumption must be implemented.

REFERENCES

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

ANSELIN, L. (1996): The Moran scatterplot as an ESDA tool to assess local instability in spatial *Spatial Analytical* 4: 111–126.

ARDHYANTI, E.–HANIF, H. (2014): Innovative policy of regional development in decentralized Indonesia.

BPS, J.T. (2017a): *Produk domestik regional bruto provinsi Jawa Timur Menurut Lapangan Usaha 2012–2016*.

BPS, J.T. (2017b): Jawa Timur dalam Angka 2017.

BRATA, N.T.–SAIRIN, S.–SETIADI, S. (2017): The economic activities analysis of the marginal community living around the artisanal oil mining site in Hargomulyo, Bojonegoro Regency *Journal of Degraded and Mining Lands Management* 4 (2): 733–738. https://doi.org/10.15243/jdmlm.2017.042.733

CHAWLA, S.–SUN, P. (2006): SLOM: a new measure for local spatial outliers *Knowledge and Information Systems* 9 (4): 412–429. https://doi.org/10.1007/s10115-005-0200-2

CHEN, D.–LU, C.-T.–KOU, Y.–CHEN, F. (2008): On detecting spatial outliers *Geoinformatica* 12 (4): 455–475.

CONLEY, T. G.–TOPA, G. (2002): Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics* 17 (4): 303–327.

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

ERTUR, C.–KOCH, W. (2007): Growth, technological interdependence and spatial externalities: theory and vidence *Journal of Applied Econometrics* 22 (6): 1033–1062. https://doi.org/10.1002/jae.963

FINGLETON, B.–LÓPEZ-BAZO, E. (2006): Empirical growth models with spatial effects. *Papers in Regional Science* 85 (2): 177–198. https://doi.org/10.1111/j.1435-5957.2006.00074.x

FITRIANI, R.–PUSDIKTASARI, Z. F.–DIARTHO, H. C. (2020): The dynamic of 2011 – 2016 East Java's regional spatial growth, An exploratory spatial data analysis *Journal of Economics and Business* 3 (2): 947–964. https://doi.org/10.31014/aior.1992.03.02.252

GETIS, A. (2010): Spatial autocorrelation. In: *Handbook of applied spatial analysis* pp. 255–278., Springer.

JÁROSI, P. (2017): Modelling network interdependencies of regional economies using spatial econometric techniques. *Regional Statistics* 7 (1): 3–16. https://doi.org/10.15196/RS07101

JENKS, G. F. (1967): The data model concept in statistical mapping *International yearbook of cartography* 7: 186–190.

KIM, Y.–KIM, T.-H.–ERGÜN, T. (2015): The instability of the Pearson correlation coefficient in the presence of coincidental outliers *Finance Research Letters* 13: 243–257. https://doi.org/10.1016/j.frl.2014.12.005

KOU, Y.–LU, C.-T.–CHEN, D. (2006): Spatial weighted outlier detection. In: *Proceedings of the 2006 SIAM international conference on data mining* pp. 614–618., SIAM.

LU, C.-T.–CHEN, D.–KOU, Y. (2003): Algorithms for spatial outlier detection. In: *Third IEEE International Conference on Data Mining* pp. 597–600., IEEE.

QURBANIA, I. D.–SUMARNOB, T.–CASSYC, R. W. N. (2020): Regional economic development strategy: Increasing local generated domestic revenue of the regional-owned enterprises in the oil and gas sector *Brawijaya Law Journal* 7 (1): 70–81. http://dx.doi.org/10.21776/ub.blj.2020.007.01.05

SHEKHAR, S.–ZHANG, P.–HUANG, Y. (2009): Spatial data mining. In: *Data mining and knowledge discovery handbook* pp. 837–854., Springer.

SUN, P.–CHAWLA, S. (2004): On local spatial outliers. In: *Fourth IEEE International Conference on Data Mining (ICDM'04)* pp. 209–216., IEEE.

VAYÁ, E.–LÓPEZ-BAZO, E.–MORENO, R.–SURIÑACH, J. (2004): Growth and externalities across economies: an empirical analysis using spatial econometrics. In: *Advances in spatial econometrics* pp. 433–455., Springer.