ANALYSING SPATIAL PATTERNS OF THE COVID-19 OUTBREAK IN TURKEY

Türkiye'de COVID-19 Salgının Mekansal Yapısı

Fatma ZEREN[*]
Veli YILANCI[**]

Geliş: *01.09.2020/* ***Kabul:*** *02.12.2020*
DOI: *10.33399/biibfad.789117*

Abstract

COVID-19 is first detected on 12 March 2020 in Turkey, and since that day more than 100 thousand people are infected. In this study, we aim to determine risky provinces in terms of COVID-19 outbreak and also explore the spatial dynamics of the outbreak in Turkey using province-level data. To analyze spatial patterns of COVID-19, we employ spatial dependence statistics Moran-I. Also, we employ Local Indicator Spatial Association-LISA to detect the hot-spots and cold-spots. Moran-I coefficient found as low and statistically significant that shows spatial interaction is not strong in the context of the whole country. Also using LISA, we found Düzce, Kocaeli, Ordu, Tekirdağ, and Trabzon as hot-spots for data period, which indicates these cities can be classified as risky in terms of COVID-19 outbreak. There are more spatial interaction with their neighbours cities. In terms of the COVID-19 variable, in hot-spot provinces and neighboring provinces of these provinces, measures should be intensified, and control should be increased.

Keywords: COVID-19, spatial autocorrelation, hot-spot, cold spots, Moran I, LISA, Turkey.

Jel Kodları: C11, H12, I18, R10

[*] Doç.Dr, İnönü Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Ekonometri, fatma.zeren@inonu.edu.tr, ORCID: https://orcid.org/0000-0002-1661-3587
[**] Doç.Dr. Sakarya Üniversitesi, SBF, Ekonometri, yilanci@sakarya.edu.tr, ORCID: https://orcid.org/0000-0001-5738-690X

Öz

Türkiye'de COVID-19 ilk olarak 12 Mart 2020'de tespit edildi ve o günden bu yana 100 binden fazla kişiye bulaştı. Bu çalışmada hem COVID-19 salgını açısından riskli illeri belirlemeyi hem de il bazındaki verileri kullanarak Türkiye'deki salgının mekansal dinamiklerini keşfetmeyi hedefliyoruz. COVID-19' un mekansal yapısını ortaya çıkarmak amacıyla mekansal bağımlılık istatistiklerinden Moran-I istatistiği kullanılmıştır. Ayrıca, "hot point" ve "cold point" alanları belirlemek amacıyla mekansal etkileşim istatistiklerinden *Lokal Moran I-LISA* istatistiği kullanılmıştır. İstatistiksel olarak anlamlı bulunan *Moran I* katsayısı tüm ülke bağlamında mekansal etkileşimin güçlü olmadığını göstermektedir. LISA istatistiğine göre ise Düzce, Kocaeli, Ordu, Tekirdağ ve Trabzon illeri veri dönemi için "hot point"alanındaki şehirlerdir. Yani bu şehirlerin COVID-19 salgını açısından en riskli alanlardır ve komşu şehirlerle daha fazla mekansal etkileşimleri bulunmaktadır. COVID-19 değişkeni açısından, "hot point" şehirler ve onların komşularında tedbirler yoğunlaştırılmalı ve kontrol artırılmalıdır.

Anahtar Kelimeler: COVID-19, mekansal otokorelasyon, hot-spot, cold spot, Moran I, LISA, Türkiye

Jel Codes: C11, H12, I18, R10

1. Introduction

Coronaviruses are a diverse, enveloped, positively-sensitive, single-stranded RNA virus that causes various disorders in the respiratory tract, intestine, and liver in humans and a wide variety of animal species. It has four types: alfacoronavirus, betacoronavirus, gammecoronavirus and deltacoronavirus (Zumla et.al, 2016). Until November 2019, only six different coronaviruses were known to have an effect on humans. Four of them (HCoV-NL63, HCoV-229E, HCoV-OC43, and HKU1) have caused mild cold symptoms in immune-healthy people, while the remaining two have caused a pandemic over the past two decades (Rabi et.al, 2016). While 8422 people were infected from the severe acute respiratory syndrome coronavirus (SARS-CoV) that first appeared in Guangdong Province in China in November 2002, 916 of these people died. This disease, which was effective from November 2002 to July 2003, had a mortality rate of 11% (Yeung and Xu, 2003). From the Middle East respiratory failure syndrome coronavirus (MERS-CoV), which first appeared in Jordan and Saudi Arabia in April-June 2002, 2506 people were infected until January 15, 2020, and 862 of them died. In other words, this disease has a fairly high mortality rate of 34% (Groot et.al, 2013). It is determined that Sars-CoV, which is found to be of zoonotic origin, like other coronaviruses, passes from exotic animals, and MERS-CoV passes from camels to people and causes epidemics (Er and Unal, 2020).

A disease epidemic, known as pneumonia of unknown cause, first discovered in Wuhan, China, in late December 2019. A few days later, this causative virus was identified as a new coronavirus by many independent laboratories and was temporarily named SARS-CoV-2, and finally named as Coronavirus disease 2019 (COVID-19) by the World Health Organization (He et. al, 2020). Although the origin of COVID-19, which is species of a betacoranavirus and shares 79.5% of the genetic sequence of SARS-CoV, is unknown, the first cases are associated with the Huanan Seafood Market in Wuhan (Lu et.al, 2020; Zhou et.al, 2020; Mahase, 2020).

The average number of basic reproduction(R_0) of COVID-19, which was declared as an "international public health emergency" by the World Health Organization on 30 January 2020, and was announced as a pandemic on 11 March 2020, was estimated to be 3.28 (Liu et.al, 2020), and the mortality rate is estimated to be between 3-4 %[1].

There are studies in the literature examining spatial dynamics of various diseases in different countries and regions. For example, Chaikaew et al., (2009) examined the spatial patterns of diarrhea cases in Thailand and found that Chiang Mai province was a hot-spot between 2001-2006 (Chaikaew,2009). Al-Ahmedi and Al-Zahrani (2013) examined the spatial distribution of cancer cases in Saudi Arabia. While the results of the study show that there is no significant spatial autocorrelation, it is also stated that there is a clustered pattern in lung cancer in men and breast cancer in women. Bhunia et al., (2013) examined the spatial spread of kala-azar disease in Vaishali region of

[1] Baud et al. (2020) stated that this rate was 5.7%, but Kim and Goel (2020) and Lipsitch (2020) emphasized that Baud et al. (2020) might be wrong in their studies (Baud et.al, 2020; Kim and Goel, 2020; Lipsitch, 2020)

India and concluded that there are spatially clustered patterns, although there are significant differences in the villages.

Kang et al., (2020) studied the spatial dynamics of COVID-19 in China. On 22 January 2020, they concluded that the COVID-19 pandemic had a significant spatial spread. Guliyev (2020), on the other hand, examined the relationship between the number of COVID-19 cases and the number of deaths and the number of recoveries using spatial panel techniques in China and found that these two variables were effective on the number of cases .

This study examines the distribution of COVID-19 among provinces in Turkey, where more than a hundred thousand people are currently infected. The second part of the study addresses COVID-19 case and death numbers in Turkey and in the World, and in the third part, data, the econometric theory, and empirical results will be presented and the study will be concluded with the conclusion and assessment part.

2. The Progress of COVID-19 in Turkey and the World

The epidemic center of the COVID-19 outbreak, which was first detected in China, became Europe on 13 March 2020. By this date, while 80932 cases were determined in China, 3172 people died. After this date, the number of cases and deaths started to increase rapidly in Europe, especially in France, Spain, and Italy. As of today (April, 23 2020), 119151 cases were detected in France, 208389 in Spain, and 187327 in Italy, the total number of deaths was 21340 in France, 21717 in Spain and 25085 in Italy. While the number of people infected in Europe today is 1.13 million, 83876 people in China have been infected; While the total number of deaths was 110692 in Europe, it was 4636 in China (Bhunia et.al, 2013).

COVID-19 has been seen in 210 countries and regions and two ships (Diamond Princess and MS Zaandam) so far, affecting about 2.5 million people and killed about 180 thousand people. The figure 1 shows the total number of cases and deaths detected worldwide over time:

As it can be seen from the chart above, the number of 500 thousand cases was exceeded on March 27, the threshold value of 1 million was exceeded on April 3, and the threshold value of 2 million was exceeded on April 15. It is noteworthy that the number of cases followed an exponential course.

Although the first cases in Turkey detected on 12 March 2020, Turkey has begun to take measures about two months ago. On 10 January 2020, the Coronavirus Scientific Board was established within the Ministry of Health, and flights were stopped with China on February 5 and Iran with February 23.

From March 16, 2020, theater, cinema, etc. the activities of such gatherings were temporarily suspended. Besides, primary, secondary and high schools, and universities were also closed down. As of March 21, 2020, a curfew has been imposed on people aged 65 and over and people with chronic diseases. As of 3 April 2020, 30 metropolitan areas and Zonguldak, where lung diseases are common, are prohibited from entering and exiting vehicles. And at the same date people under the age of 20 are prohibited from going out on the streets. Also, it is obligatory to wear a mask in crowded areas. On April 5, the Ministry of Health informed that free masks will be provided for people who do not have a curfew. There are curfews in 31 provinces on April 11, 12, 18, and 19 and an additional curfew was also declared on April 20 for the April 23-26.

The first death occurred on March 17, 2020 in Turkey, where there are currently 98674 cases, 2276 deaths (April 23, 2020). These values of COVID-19 are summarized in Figure 2 (Roser et.al, 2020).

Although the rate of change in the number of cases in Turkey seems very similar to most other countries, the number of deaths has remained relatively slower than in other countries because in Turkey, measures were taken by the earliest date.

3. Material and Methods

In Turkey, the number of COVID-19 cases per province only announced twice, on April 1, 2020 and April 4, 2020. In this study, we employed the latest announced date gathered from Dokuz8Haber, by computing COVID-19 cases per 100 thousand people. We take the logarithm of the data. In this study, we will consider the spatial patterns of COVID-19. Tobler's first "law" of geography is that: "everything is related to everything else, but near things are more related than distant things". Depending on this assumption, which states close observations are associated with each other, we want to measure whether there is a spatial association of COVID-19 among provinces in Turkey.

There are many tests in the literature that investigate whether spatial autocorrelation exists according to a parametric approach. The most commonly used test from these tests is the Moran I statistic. This statistic is the autocorrelation coefficient between the observations of a variable (COVID-19) at a particular location and the mean (spatial lag variable) of the same variable in its neighbors. Are the provinces with similar characteristics clustered together for the examined variable? Is there any interaction between the channel and neighboring provinces (communication or access)? This autocorrelation coefficient gives the answer to the questions. We test the null hypothesis of no spatial dependence against the alternative of spatial dependence, that is, observations are spatially clustered.

Moran I statistics are as follows (Moran, 1948; Moran, 1950; Anselin, 1988).

$$I = \frac{N}{S_0} \frac{y^{\star 1} W y^{\star}}{y^{\star 1} y^{\star}} \tag{1}$$

Here the variable y is the total number of COVID-19 cases per 100 thousand (hereafter only COVID-19) for all 81 provinces and y_i shows the same variable of the province. $y_i^{\star} = y_i - \bar{y}$ and W, is a (NxN) dimensional weight matrix that measures spatial interaction or similarity. N, is the total number of observations in the study area and is 81 for this study. In this study, the quenn weight matrix will be used. According to this weight matrix, the provinces sharing common boundaries and corners are assumed to be neighbors. w_{ij}, shows each element of the weight matrix W. S_0, shows the sum of these elements $S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}$. w_{ij}, takes a value of 0 or 1 and is defined as follows.

$$w_{ij} = \begin{cases} 1\ , & while \quad j \in N(i) \\ 0, & others \end{cases} \qquad (2)$$

$N(i)$ is a cluster of neighbors of the province and if j. is the element of this set according to weight matrix, the corresponding observation value of the weight matrix takes the value 1, and if not it takes the value zero.

Moran I coefficient takes value in the range [-1, 1] like other correlation coefficients. When it gets zero, it means that there is no correlation and that all provinces have a random distribution, and there is no cluster in the relevant space (all sample-provinces). -1 is the perfect negative correlation situation where non-similar provinces are clustered together. When it takes the value 1, it is a perfect positive correlation situation where similar provinces cluster together or interact with their neighbors. In the case of positive autocorrelation, high-value provinces or low-value provinces cluster together. In the case of negative autocorrelation, high and low-value provinces (ie, non-similar provinces) cluster together.

4. Results

4.1. Global Indicator Spatial Association- Moran I

Moran scatter plot of COVID-19 variables is shown in Figure 3. As seen in this figure, there are four different sections. The upper right section shows a positive autocorrelation region where the spatial association is high, and the lower-left section is the positive correlation region with low-value observations. In the lower-right and upper-left regions, there is the negative autocorrelation region with spatial deviations, that is, clusters of dissimilar values. In terms of spatial analysis, the clustering of these last two regions has no meaning. These areas are areas of dispersion.

Moran I coefficients in Figure 3 are statistically significant according to the quenn neighborhood definition. Inference for the Moran I coefficient is carried out with several different approaches. In this study, we used a computational approach based on *permutation*. This calculates an empirical reference distribution of Moran's I under the null hypothesis of spatial randomness (GeoD Center). The psuedo-p values calculated for 999 permutations for the COVID-19 variables is 0.004.

As a result of the findings, the null hypothesis was rejected. The estimated value of Moran I is 0.193. It is closer to 0 and farther to 1, so they are not very high. Thus, it can be said that spatial interaction is not strong in the context of the whole country or all provinces. This finding is in line with our theoretical expectations. Because Republic of Turkey did not anounce the number of cases of COVID-19 about provinces when the pandemic first appeared. Thus, a spatial spread is not very dense.

In Figure 3, the observations are located both at the right and the left of the average. Besides, observations appear to be distributed both in the positive autocorrelation and in the negative autocorrelation region. Only positive autocorrelation is important in terms of spatial dependence, and it can be said that the distribution in these autocorrelated regions is not intensive, and thus spatial autocorrelation is not very strong.

4.2. Local Indicator Spatial Association - *LISA Statistic*

Moran I coefficient is a statistic that provides only one value for all provinces. Anselin (1995) developed the local indicator spatial association (LISA) statistic that measures spatial autocorrelation for each location i . This statistic is also called Local Moran I statistics and can be calculated as follows: (Anselin, 1995; Getis and Ord, 2010).

$$I_i = (y_i - \bar{y}) \sum_{j=1}^{N} w_{ij} \left(y_j - \bar{y}\right) \quad for\ i \neq j \tag{3}$$

For each province, this statistic is calculated based on the observations of the neighbor provinces or provinces that it shares a border. Thus, it is stated that this calculated statistic measures the similarity of the related location with its neighbors. The significance of LISA statistics is also determined by the psuedo-p-value obtained by the permutation approach.

Many GIS (geographic information system) programs provide a significant spatial association map and spatial cluster map based on the information they obtain with this statistic. Significance maps for Local Moran I statistics are shown in Figure 4 for COVID-19.

According to Figure 4, there are local association in 13 provinces (4 at the 1% significance level and 9 for 5% level). These provinces are as follows: Adana, Bolu, Denizli, Düzce, Mardin, Ordu, Muğla, Trabzon, and Van at the 5% significance level and Hakkari, Kocaeli, Siirt, and Tekirdağ according to the 1% level.

Hot-spot, cold-spold, and spatial outliers locations can be determined on the Local Moran I cluster map. Hot spot regions are provinces in which there are a high number of COVID-19 cases and provinces surrounded by provinces with high COVID-19 cases. Cold spot regions are the areas where low-value provinces are clustered. Spatial outliers are divided into two regions. In high-value outlier, high-case provinces are surrounded by low-case provinces. The low-value outlier is surrounded by low case provinces and high case provinces. The cluster map obtained by Local Moran I statistics is given in Figure 5.

The number of hot spot provinces is five and the number of cold spot provinces is 4 for COVID-19. While the hot spots are Düzce, Kocaeli, Ordu, Tekirdağ, Trabzon, the cold spots are Van, Muğla, Mardin, Hakkari. It is found in only five provinces with spatial outliers, and the spatial pattern is random in the remaining provinces.

5. Conclusion

In this study, which examines the spatial dynamics of Covid-19 cases that are confirmed in Turkey for the first time, the proportion of COVID-19 cases pertaining to 81 provinces, announced on April 4, per 100000 population is addressed.

The fact that the estimation of Moran's I was significant and low indicates that the cluster of COVID-19 occurred only in small regions throughout the country. This information indicates that there is a cluster in the context of COVID-19 and it does not have a random or uniform distribution in Turkey

According to the local spatial association (LISA) statistics, hot spot, and cold spot areas have been determined for the provinces. The hot spot covers the provinces with positive autocorrelation, where the spatial association is most significant. The cold spot includes the provinces where the spatial association is the lowest but significant.

The LISA results demonstrate that Düzce, Kocaeli, Ordu, Tekirdağ, and Trabzon are the cities with a higher risk of COVID-19. The reason for this is that Istanbul, which is the province with the highest number of Covid-19 cases, is neighbor with both cities. According to the LISA statistics, these provinces are hot spot provinces. We think that there may be spatial spreads because some of the people whose residence addresses are in Duzce, Kocaeli, and Tekirdağ may have workplaces in Istanbul.

According to LISA statistics, the majority of these provinces, which are determined to be cold-spot, are located in the southeastern Anatolian region of the country. The reason for the lower number of cases in these regions is that they do not have direct links with international transportation. As it is known in cold spot areas, there is a significant spatial association even if it is at a low level. In this context, they should be more careful at the point of contact with neighboring provinces.

In terms of the COVID-19 variable, especially the Turkish Ministry of Health should pay attention to hot-spot provinces, in the measures and follow-up policies regarding COVID-19. In these provinces and neighboring provinces of these provinces, measures should be intensified, and control should be increased. Cold spot areas are also provinces with positive spatial correlation. There is also partial spatial spread or spatial similarity in these provinces and their neighboring provinces. Similarly, it is necessary to develop a policy for precautionary measures in these regions.

References

Al-Ahmadi, K., & Al-Zahrani, A. (2013). Spatial autocorrelation of cancer incidence in Saudi Arabia. *International Journal of Environmental Research and Public Health*, 10(12), 7207-7228.

Anselin L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht

Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical analysis*, 27(2), 93-115.

Baud, D., Qi, X., Nielsen-Saines, K., Musso, D., Pomar, L., & Favre, G. (2020). Real estimates of mortality following COVID-19 infection. *The Lancet Infectious Diseases*, 20(7), https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30195-X/fulltext?fbclid=IwAR2V4jsykeM-R5lsAjcksGvmw4zYCutcVqLd9btqQkIsldVRoNedkBKDMNs

Bhunia, G. S., Kesari, S., Chatterjee, N., Kumar, V., & Das, P. (2013). Spatial and temporal variation and hotspot detection of kala-azar disease in Vaishali district (Bihar), India. *BMC Infectious Diseases,* 13(1), 1-12.

Chaikaew, N., Tripathi, N. K., & Souris, M. (2009). Exploring spatial patterns and hotspots of diarrhea in Chiang Mai, Thailand. *International Journal of Health Geographics*, 8(1), 36.

Chan-Yeung, M., & Xu, R. H. (2003). SARS: epidemiology. *Respirology*, 8, 9-14

de Groot, R. J., Baker, S. C., Baric, R. S., Brown, C. S., Drosten, C., Enjuanes, L., ... & Perlman, S. (2013), "Commentary: Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. *Journal of Virology*, 87(14), 7790-7792

Dokuz8haber, COVID19 - TÜrkiye Raporu, https://datastudio.google.com/u/0/reporting/1KH9kCoJoh1VgwdFbFPIbX3sONzvrOJ2k/page/fpDLB?s=i2fOJW2TkuU [Erişim Tarihi : 20.04.2020]

Er, A. G. and Ünal, S. (2020). 2019 koronavirüs salgını–anlık durum ve ilk izlenimler. *FLORA*, 25, 1-5.

GeoDa center, https://geodacenter.github.io/workbook/5a_global_auto/lab5a.html

Getis, A., & Ord, J. K. (2010). The analysis of spatial association by use of distance statistics. In *Perspectives on spatial data analysis* (127-145). Springer, Berlin, Heidelberg

Guliyev, H. (2020). Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Statistics*, 100443. https://doi.org/10.1016/j.spasta.2020.100443

He, F., Deng, Y., & Li, W. (2020). Coronavirus disease 2019 (COVID-19): What we know?. *Journal of Medical Virology*, https://doi.org/10.1002/jmv.25766

Kang, D., Choi, H., Kim, J. H., & Choi, J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China. *International Journal of Infectious Diseases*, 94, 96-102

Kim, D. D., & Goel, A. (2020). Estimating case fatality rates of COVID-19. *The Lancet Infectious Diseases.*

Lipsitch, M. (2020). Estimating case fatality rates of COVID-19. *The Lancet Infectious Diseases.*

Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*. 27(2), 1-4.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., & Bi, Y. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224), 565-574.

Mahase, E. (2020). China coronavirus: what do we know so far?. *BMJ 2020*, 368 doi: https://doi.org/10.1136/bmj.m308

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, 10(2), 243-251.

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

Rabi, F. A., Al Zoubi, M. S., Kasasbeh, G. A., Salameh, D. M., & Al-Nasser, A. D. (2020). SARS-CoV-2 and Coronavirus Disease 2019: What We Know So Far. *Pathogens*, 9(3), 1-14.

Roser, M, Ritchie, H, Ortiz-Ospina, E., and Hasell, J. (2020). Coronavirus disease (COVID-19). Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/coronavirus' [Çevrimiçi Kaynak],

Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., & Chen, H. D. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270-273.

Zumla, A., Chan, J. F., Azhar, E. I., Hui, D. S., & Yuen, K. Y. (2016). Coronaviruses-drug discovery and therapeutic options. *Nature Reviews Drug Discovery*, 15(5), 327-347

FIGURES

Figure 1: Cumulative Number of Cases and Deaths for COVID-19 at Worldwide level

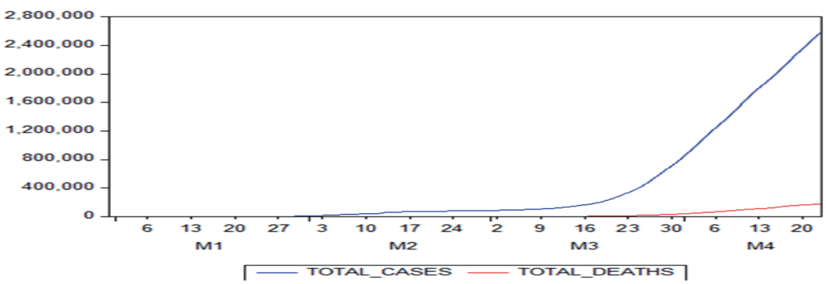

Figure 2: Cumulative Number of Cases and Deaths for COVID-19 at Turkey

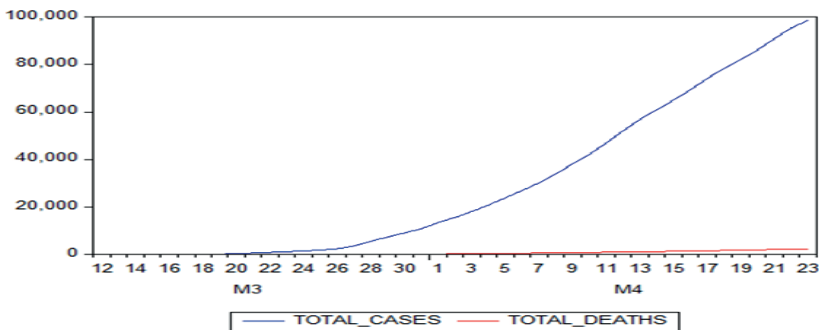

Figure 3: Moran I Scatter Plot for COVID-19

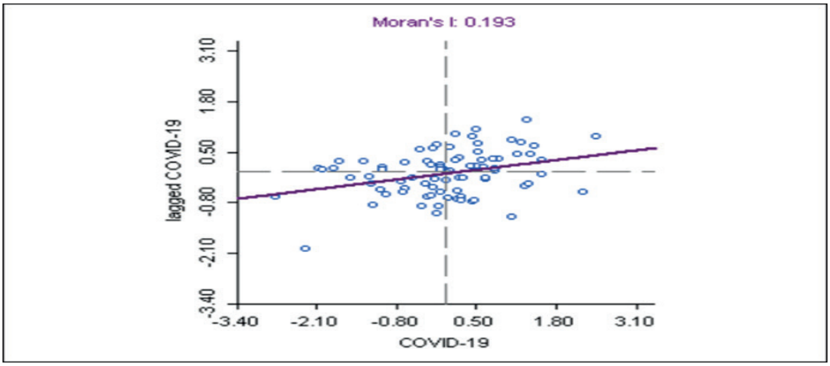

Figure 4: LISA significance Map for COVID-19

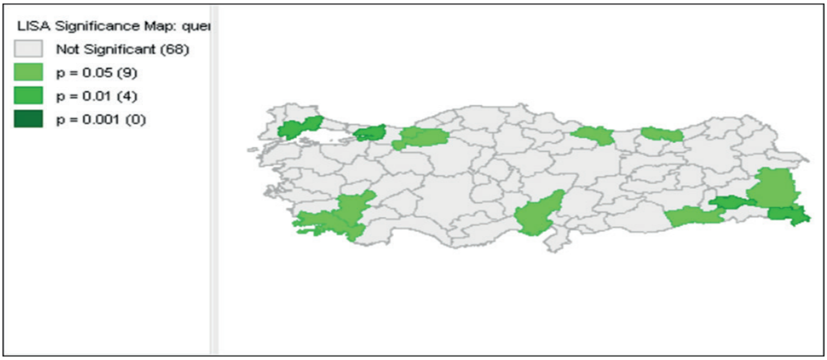

Figure 5: LISA Cluster Map for COVID-19