UTILIZATION OF HOT SPOT ANALYSIS IN THE DETECTION OF SPATIAL DETERMINANTS AND CLUSTERS OF THE SPANISH FLU MORTALITY

Suzana Lović Obradović1*, Vladimir Krivošejev2, Anatoliy A. Yamashkin3

1Geographical Institute “Jovan Cvijić” SASA, Belgrade, Serbia; e-mail: s.lovic@gi.sanu.ac.rs
2National Museum “Valjevo”, Valjevo, Serbia; e-mail: vladimir.krivošejev@gmail.com
3National Research Mordovia State University, Geography Faculty, Saransk, Russia; e-mail: yamashkin56@mail.ru

Abstract: The Spanish flu appeared at the end of the First World War and spread around the world in three waves: spring-summer in 1918, which was mild; autumn fatal wave, in the same year; and winter wave in 1919, which also had great consequences. From the United States of America, as the cradle of its origin, the Spanish flu spread to all the inhabited continents, and it did not bypass Serbia either. Research on the Spanish flu, as the deadliest and most widespread pandemic in the human history, was mostly based on statistical researches. The development of the geographic information systems and spatial analyses has enabled the implementation of the information of location in existing researches, allowing the identification of the spatial patterns of infectious diseases. The subject of this paper is the spatial patterns of the share of deaths from the Spanish flu in the total population in Valjevo Srez (in Western Serbia), at the settlement level, and their determination by the geographical characteristics of the studied area—the average altitude and the distance of the settlement from the center of the Srez. This paper adopted hot spot analysis, based on Gi* statistic, and the results indicated pronounced spatial disparities (spatial grouping of values), for all the studied parameters. The conclusions derived from the studying of historical spatial patterns of infectious diseases and mortality can be applied as a platform for defining measures in the case of an epidemic outbreak with similar characteristics.

Keywords: historical demography; mortality spatial patterns; hot spot analysis; Valjevo Srez

Introduction

The deadliest pandemic in the history of mankind was caused by the Spanish flu, which, covering almost the entire world, appeared at the end of the First World War. It spread around the world in three waves: the first was mild, began in the spring and lasted until summer in 1918; the second—autumn fatal wave, in the same year; and the last was the winter wave in 1919. The last wave also had great consequences, but it covered only some territories, including Australia, and had its “tails” in the 1920s. It is assumed that the Spanish flu, in its second mortal wave, arrived in the northern territory of occupied Serbia from Austria–Hungary, from the direction of Bačka, through Zemun to Belgrade, from where it was spread by cars and railway communications throughout Serbia, and, in October 1918, reached Valjevo Srez (Krivošejev, in press).

*Corresponding author, e-mail: s.lovic@gi.sanu.ac.rs
Given that the path of virus spread contains a geographical dimension, it is expected that data on the number of the infected and dead are spatially grouped. For the detection of these patterns, it is possible to apply the principle of spatial dependence, which is based on Tobler’s first law of geography: “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970, p. 236). Accordingly, it is expected to have the clustering of neighboring settlements that have similar values of the number of the infected and the share of deaths from the Spanish flu in the total population, compared to the distant settlements. Mortality, as well as other demographic phenomena or processes are not uniform, but their volume and intensity change regarding the environment (Lović Obradović & Vojković, in press). The emergence of the patterns, taking into consideration the living conditions in the first half of the twentieth century, largely depended on the geographical characteristics of the area—altitude and distance from the center of Valjevo (distance), as the administrative and geographical center of the Srez (according to the then valid administrative division, Serbia was divided into districts, districts into srez or counties, and within the counties, there were municipalities, with one or more settlements) and the district.

In order to determine the spatial patterns of the share of the population who died from the Spanish flu in the total population, as well as the factors that influenced their clustering, the paper adopted hot spot analysis, based on Gi* statistics, for the detection of hotspots and coldspots. The detection of spatial manifestation of the demographic phenomena, such as mortality, enables a better understanding of the spread of the Spanish flu, as well as other infectious diseases throughout history. It can also be applied in modern spatial demography and spatial epidemiology for designing more efficient public health interventions for future outbreaks with similar characteristics (Chowell, Bettencourt, Johnson, Alonso, & Viboud, 2007) and for developing public health countermeasures and implement effective mitigation plans (Reyes et al., 2018).

Background and literature review

The origin of the Spanish flu pandemic was on the territory of the USA and with its entry into World War I and the arrival of American recruits at the European battlefield, the virus flooded Europe and other continents with tremendous speed. The war censorship of that time, hiding the appearance of the disease as a military secret, suppressed the information about its appearance. Spain enjoyed neutrality in the war and that is why its media wrote more openly about the scale of the epidemic, especially when King Alfonso XIII, the Prime Minister, and several members of the government became ill. Thus, a false first impression that the origin of the infection was in Spain was created, which gave the name to the disease (Radusin, 2012a). Although the pandemic had great consequences, they were neglected for a long time, since the disease was seen as a continuation of war suffering (Anušić, 2015). It was estimated that about 500 million people became infected, and about 50 million died as a result of the Spanish flu (Hutinec, 2006). Based on these claims, it can be concluded that during the pandemic, a third of the world’s population at the time became ill, and died much more from its consequences than from all known pandemics together (Radusin, 2012a, 2012b).

In scientific circles, the Spanish flu pandemic began to receive more attention during the second half of the 20th century, with the appearance of other related pandemics, first Asian (1957/58), then Hong Kong flu (1968/1969), and additional interest was noticed at the beginning of this century, after the outbreak of new pandemics: SARS, avian flu, swine flu, and MERS. However, the “great fear” caused by these diseases was not reflected in Serbia, so there were several studies dealing with the Spanish flu. The increased interest emerged in the early 20th century (Radusin, 2006, 2012a, 2012b),
although before these studies, many authors had also paid attention to the appearance of the Spanish flu (Dragić, 1980; Đenić, 1985; Gavrilović, 1995; Milenković, 2007). Dragić (1980) described the occurrence of the Spanish flu in Belgrade but pointed out that there were no data on how many people died, while Đenić (1985) and Milenković (2007), partially quantified the victims in the territories of Trstenik Srez and Zlatibor Srez, and Milenković (2007) pointed out the number of deaths in Šajkaš villages, without additional quantifications. Their research, with additional analyses, indicated that more than 1,000 people, i.e., about 5% of the population, died in the distinctly hilly and mountainous Zlatibor Srez (Đenić, 1985; Krivošejev, 2020), while in the partly valley and partly hilly Trstenik Srez 1,267 people died, i.e. 3.74% of the population (Krivošejev, 2020; Milenković, 2007).

Most traditional epidemiological studies are based on the statistical analysis rather than on spatial information (Jiang, Zhang, Jin, Zhang, & Wang, 2011). The development of geographic information systems and spatial analyses has enabled the implementation of location information in contemporary demographic and epidemiological research. This further enabled the detection of the spatial patterns of spreading infectious diseases, such as the Spanish flu, as well as the mortality. The research of available data has shown that the spatial patterns of the mortality of the Spanish flu depended on the local specifics of the area, both geographical and socio-economic. A group of researchers led by Chowell et al. (2007), considered that spatial variations in disease and mortality patterns of the 1918–1919 influenza pandemic remain poorly studied, and that the substantial geographical variations in pandemic mortality impact can occur within and between countries, perhaps due to the differences in prior immunity, economy, background mortality levels, and population density (Chowell, Viboud, Simonsen, Miller, & Acuna-Soto, 2010). The results of the study about the disparities in influenza mortality and transmission related to socio-demographic factors within Chicago showed that the spatial variation in mortality in 1918 had strong patterns of spatial clustering (Grantz et al., 2016). Geographic variation in pandemic influenza-related death patterns in Chile were conditioned by a combination of local factors: host-specific susceptibility, population density, baseline death rate, and climate (Chowell, Simonsen, Flores, Miller, & Viboud, 2014). Cilek (2019) considers that demographic and spatial determinants conditioned specific patterns of the Spanish flu mortality in Madrid. The research of the spatial patterns of the Spanish flu mortality based on the available data for Valjevo Srez, as well as the geographical characteristics of the area affected by the epidemic, which determine the spatial heterogeneity of these patterns, is a pioneering endeavor in the field of historical spatial demography and spatial epidemiology in Serbia.

**Methodology**

This study covers the territory of Valjevo Srez, which was one of the five srez of Valjevo district. It was composed of 14 municipalities with 53 settlements. According to the 1910 census, there were 36,867 inhabitants in the Srez, 27,746 in 1916 (Popović, 2000), and 33,600 inhabitants were registered in 1921 (Opšta državna statistika, 1932). Ten church registers of the deceased, in which data for the deceased from 51 settlements of Valjevo district had been registered, were analyzed. No data were found for the Slovac settlement (1% of the population of Valjevo district). The total number of deaths from the Spanish flu on the territory of Valjevo Srez was 635, or 2.29% of the population registered in the 1916 census. The number of inhabitants at the settlement level was taken from the official data from the 1916 census, as the year of the closest census of the studied period of the epidemic. The average altitude of the settlement was generated from European Digital Elevation Model, while the distance was calculated using Planplus.rs application.
Earth application was used for geocoding (obtaining longitude and latitude coordinates), to determine the centers of the settlements from a contemporary network of settlement boundaries.

Figure 1. Map of the settlements of Valjevo Srez with the number and percentage share of deaths from Spanish flu in the total population of the settlement (Author of the map: Zoran Mujbegović)

We conducted hot spot analysis to identify statistically significant hot and coldspots using Getis-Ord $G^*_i$ statistic. Getis and Ord (1996) developed $G^*_i$ statistics to detect local spatial dependence, which remains hidden when applying previously developed $G_i$ global statistics. It is defined based on the formula:

$$G^*_i = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{\sqrt{\left[ n \sum_{j=1}^{n} w_{i,j}^2 - (\sum_{j=1}^{n} w_{i,j})^2 \right] / n - 1}}$$

(1)

where $x_j$ is the attribute value for feature $j$, $w_{i,j}$ is the spatial weight between features $i$ and $j$, $n$ is the total feature number.
With the analysis it is possible to determine if there was a spatial pattern of the Spanish flu mortality and what the chances are that those patterns are a result of a random spatial process. The output of the hot spot analysis tool is GiZScore and GiPValue for each feature, and these values represent the statistical significance of the spatial clustering of values, given the conceptualization of spatial relationships and the scale of analysis (Prasannakuma, Vijith, Charutha, & Geetha, 2011). P-value represents probabilities and z-score standard deviation. Very low p-value and high values of z-score indicate hotspots (red shades), while very high values of p-value and low values of z-score indicate coldspots (blue shades). Hot/cold spot does not always indicate the highest/lowest value of the feature in the study area. Instead, the feature is observed within the context of predefined neighborhood (a group of features around it, including the feature itself). If the neighborhood is significantly different from the study area, the feature is marked as a hotspot (the value is significantly higher than the study area) or coldspot (the value is significantly lower than the study area). There are three different confidence levels – 90%, 95%, and 99% for cold- and hotspots. In case that spatial cluster is not statistically significant, the patterns are random (grey dots) (Esri, 2020). The analysis was conducted in Version 2.5 of ArcGIS Pro (Esri, 2020).

Results and discussion

The results of the research are presented within the contemporary boundaries of the municipality of Valjevo (78 settlements), whose area is slightly larger than the area of the Srez (52 settlements). Distinct spatial disparities, i.e., spatial grouping of values, were identified for all the studied parameters. The separation of two different clusters of hot- and coldspots in the area is evident. Settlements marked as non-significant do not belong to hot- or coldspots. These settlements have very different values compared to the values in settlements in their immediate environment (neighborhood), indicating certain atypical events in the space. They can refer to the differential factors (average altitude and the distance) that caused different levels of demographic phenomena, in this case, the share of deaths (Lović Obradović, 2019).

Figure 2 presents the mapped results of the hot spot analysis of the share of deaths from the Spanish flu in the total population. The cluster of hotspots is located in the southwestern part of Valjevo Srez and includes 17 settlements (55% of the total population of the Srez) marked as hotspots. These are: Balinović, Bačeveci, Bogatić, Brezovice, Vujinovača, Gornje and Donje Leskovce, Kunice, Lelić, Pakljæ, Rebelj, Rovni, Sandalj, Sovač, Stubo, Sušice, and Tubravić. This means that in the mentioned settlements, the share of deaths from the Spanish flu is higher regarding the average value of Valjevo Srez at the aggregate level and that they are surrounded by the settlements with the same characteristics, pointing to the location of the most affected settlements. The other cluster

Figure 2. The hot spot analysis of the share of deaths from the Spanish flu in the total population of the settlement.
is located in the northeastern part of Valjevo Srez, with 22 identified coldspots: Babina Luka, Beloševac, Blizorje, Brankovina, Bujačić, Valjevo, Veselinovac, Gornja Grabovica, Divci, Dupljaj, Žabari, Donja Zabrdica, Jasenica, Joševa, Klanic Kozličić, Loznica, Lukavac, Petnica, Popučke, and Rađevo (27.9% of the total population of the Srez).

In the listed settlements, the values of the share of deaths are lower regarding the average of the Srez, and those settlements are surrounded by the settlements with also low values of the share of deaths.

Figure 3 shows statistically significant hotspots and coldspots of the average altitude of the settlements. The cluster, consisting of 15 settlements (39.5% of the total population of the Srez), marked as hotspots, is located in the eastern part of the Srez (Belić, Beloševac, Bujačić, Degurić, Dračić, Gornja Grabovica, Jasenica, Klinci, Petnica, Popučke, Rađevo, Sedlari, Valjevo, Žabari, and Žarube). These settlements have higher values of average altitudes regarding the average value of the Srez and are surrounded by the settlements with also higher values of average altitudes. The cluster of nine coldspots (13.5% of the total population of the Srez) is located in the southwestern part of the Srez and consists of the settlements: Gornje Leskovice, Kunice, Paklje, Rebelj, Sovač, Stubo, Sušice, Tubravić, and Vujinovača and they form a neighborhood with lower values than the average. The presented segment of the research indicates a significantly higher mortality rate in villages that are mainly located on the southern slopes of the Valjevo Mountains.

Figure 4 presents the clusters of hotspots and coldspots of the distance. Twenty-three statically significant coldspots (52.3% of the total population of the Srez): Belić, Beloševac, Brganović, Bujačić, Valjevo, Gornja Grabovica, Degurić, Dračić, Žabari, Zabrdica, Zarube, Zlatarić, Jasenica, Klinci, Kovačice, Lelić, Petnica, Popučke, Prijezdić, Ravanje, Rađevo Selo, Sedlari, and Strmna Gora, form a cluster in the central and south-eastern part of the Srez. All these settlements are at smaller distances from the center of Valjevo. The cluster of six hotspots (8.9% of the total population of the Srez): Brezovica, Vujinovac, Kunica, Sovač, Sušice, and Tubravić, is located in the south-western part of the Srez and these settlements are at significantly greater distances from the center of Valjevo.
In the conditions of limited movement, such as the one in the Valjevo Srez in the first half of the 20th century, the distance had a significant role in the forming of behavioral patterns of the population, as well as their isolation. Everyday life in an isolated rural environment meant a much lower level of health culture. On the one hand, this refers to the specifics of family cooperatives, with a large number of household members and a kind of semi-collective accommodation, food from the same dishes, as well as ignorance and non-application of the basic measures of prevention and self-treatment. On the other hand, the population of isolated areas was provided with medical care to a much lesser extent, which was more accessible to the inhabitants of towns and nearby villages. Besides, it should be added that due to difficult communications, the inhabitants of isolated rural settlements had fewer contacts, and were not able to "get through" and thus create a stronger collective immunity.

**Conclusion**

This study adopted hot spot analysis based on Getis-Ord Gi* statistic to identify spatial clusters of hot- and coldspots of the share of deaths from the Spanish flu in the total population of settlements in Valjevo Srez, as well as geographical characteristics of the area—the average altitude of the settlement and the distance of the center of the settlement from the center of the Srez. Our results suggest that the cluster of the coldspots of deaths is located in the area where the cluster of the hotspots of average altitudes are and the cluster of the coldspots of the average distance are located. The hotspots cluster of deaths is concentrated in the area of the cluster of the coldspots of average altitudes and the cluster of the hotspot of the average distance. The presented results indicate that in Valjevo (an urbanized center of the Srez with higher population density), rural settlements that surround it, and villages that are relatively more distant, but with good road communication with the center of the Srez and other urban centers from the wider environment, mortality from the Spanish flu was significantly smaller than in villages with poorer road communication. This should not be directly related to the altitude of the settlements, but to their isolation from urban centers. Based on the obtained results, we may conclude that the share of deaths from the Spanish flu in the total number of inhabitants of the settlement was determined by the geographical characteristics of the studied territory.

Based on the collected data from the church registers of other parishes, it is possible to determine the spatial patterns of mortality from the Spanish flu and make a comparative analysis between parishes, srez or districts and thus contribute to the development of historical geography, historical demography, and historical spatial epidemiology in Serbia. A better understanding of these historical issues can help contemporary epidemiologists and policymakers to better prepare for future outbreaks (Cilek, 2019). Determining the spatial patterns of infectious diseases and their spatial extent, can also represent a framework for determining the spatial patterns of the number/share of the infected or dead from the ongoing Covid-19 pandemic.

**Acknowledgements**

The authors of this paper would like to thank Zoran Mujbegović for his support in making Figure 1.

**References**


