Leonidas Doukissas, Statistician, Phd Candidate

Department of Geography, Harokopio University

doukissas@yahoo.com

Stamatis Kalogirou, Assistant Professor

Department of Geography, Harokopio University

skalo@hua.gr

POINT PATTERN ANALYSIS OF CLOSED SCHOOLS DURING THE PANDEMIC OF THE VIRUS A(H1N1)pdm09 IN 2009. THE CASE OF THESSALONIKI

Summary

The pandemic A(H1N1)pdm09, which was first detected in Mexico and spread rapidly throughout the whole world, is considered the fastest flu in history (WHO, 2009). In Greece, concerning the pandemic, according to the last epidemiological report of Hellenic Center for Disease and Prevention (HCDCP) on 26 May 2010, during the period 2009 - 2010, 18.230 laboratory confirmed cases were reported (average age 21 years old) (HCDCP, 2010). The Pandemic was at its peak during the 48th- 49th week of 2009. In 2009 Pandemic, young children experienced the highest rates of disease, and country reports reveal that the highest rates of infection were in school - aged children (ECDC, 2010). As a result, for the 2009 pandemic, schools (Health Protection in West Midlands, 2009) and households (Nishiura et al, 2009) were especially important. This particular assignment conducts a spatial analysis on closed schools due to A(H1N1) pandemic flu during the onset of the pandemic in a Greek Region, municipality of Thessaloniki, from October to December 2009. The documentation of the specific data was done by the Institution of Computer and Technology center in Patra (CTI) in cooperation with HCDCP, the National school of Public Health and the Ministry of Education and Lifelong Learning (ICT, 2009). The week when the most laboratory confirmed cases were reported coincides with the week when the number of closed schools in primary school were at their highest. First-order and second-order properties characterize the comportment of a general spatial distribution process. The spatially varying intensity of a point pattern can be described with *first-order properties*. Intensity is defined as the mean value of the distribution at locations throughout the region of interest (Diggle, 1983). The covariance structure of the point pattern described with second-order properties. These properties identify and analyze the distribution of distances between those sample points (King 1962; Wiegand et al, 2004; Ripley 1977). There are several functions for modeling spatial patterns. A suitable tool to describe second-order properties of a point pattern is Ripley's K function. If the points display a clustered pattern, an excess of sample points at short distances can be shown. It's worth noting that in the years 2009 - 2011 Central Macedonia had an average of 44 deaths / per year from H1N1 with a mortality ratio higher by 33% than the entire country (Doukissas et al, 2016a). Visualizing and analyzing closed schools in a future pandemic wave is crucial for its early mitigation aiming at the non-propagation of the virus.

Key words: Point pattern Analysis, Pandemic H1N1, spatial analysis, primary education

Introduction

According to Trifonov et al (2010) the A(H1N1)pdm09 pandemic which was first detected in Mexico in April 2009, “where a previous triple reassortment of bird, swine and human flu viruses further combined with a Eurasian pig flu virus appeared to be a new strain of H1N1”.

In Greece, the first known case of the virus H1N1 was reported by a 19-year-old man on May 18th in 2009, who had returned two days earlier from New York. On the 26th and 27th May the second and third incident was diagnosed in two university students who had returned from the UK. These two cases were the first cases of influenza A(H1N1)pdm09 imported from another country within the European Union, as mentioned in the article of Panagiotopoulos et al. (2009). In national studies, around 25 to 30% of deaths attributed to the pandemic were in entirely healthy young adults and outside the traditional risk groups (ECDC, 2010). In Greece, concerning the pandemic, according to the last epidemiological report of HCDCP on 26 May 2010, during the period 2009 - 2010, 18.230 laboratory-confirmed cases were reported (average age 21 years old) (HCDCP, 2010). Glezen (1996) compares the pandemics H1N1 during the years 1892, 1918, 1936, 1957 per age and by reviewing the special infection rate per age concludes that school children hold the highest infection rate during the pandemics (and inter-pandemics). Moreover, children play an important role in the spread of the virus in the community. Peak hospitalization for pneumonia and deaths usually occur with two weeks delay after the peak incidence of a disease in the community.

As it was detected by the European Centre for Child Disease Prevention and Control, children were particularly vulnerable to the virus A(H1N1). As a result, reduction strategies of the phenomenon for the protection of the younger population and in general the whole community are necessary. For example, closing the schools constitutes a non-pharmaceutical prevention, which has often been proposed for controlling the pandemic of the virus among children and the population of adults. The basic cause, as it was discovered, is that children rapidly transmit diseases, as they are more infectious and susceptible to different strains of viruses than adults. Consequently, the high percentage of contact in schools will increase transmissions. This is a strong argument since the 60% of people who were infected by H1N1 are 18 years old or younger.

Thus, the responsible authorities claim that precautionary the closure of schools will lead to a lower transmission of the disease. (Fraser et al., 2009; Team, 2009). During a public health emergency there are three different causalities for closing schools which emerged in the 2009 H1N1 outbreak (Klaiman et al 2011). First school closures “limit the spread of the virus in the community, secondly protect vulnerable children and thirdly react to staff shortages or children kept at home because of infection or parents’ fears of infection”. However, this strategy may have some benefits as far as public health, in the bibliography there is a rich debate concerning the dangers and the advantages of closing the schools. Recent studies show the lack of satisfactory evidence for the efficiency of such public measures as is the closure of schools. Even if there were benefits, these must be weighted up against the potential high risk of economic and social cost caused by closing schools taking into account the consequences that households and the workforce take on (Cauchemez et al., 2008).

Data

The data refer to the closed schools of primary education on December 7 of 2009 (48th - 49th week) which coincide with the peak of confirmed cases. The specific data was gathered by the National School of Public Health in cooperation with the Institute of Technology, Computer and Education and supervised by the Ministry of Education and Lifelong Learning (ICT, 2009).

Methodology

Point pattern, is the estimation of the pattern, or distribution, of a set of points on a study area. It can refer to the actual spatial or temporal exact location of these points (events) in the given study area. The points could be as in our analysis the closed schools due to the virus A(H1N1)pdm09. The term intensity is the mean number of points or in other words the expected number of points per spatial unit. So, we will thoroughly analyze the intensity which can be defined either as constant (homogeneous) or it can differ from area to area (inhomogeneous). In the current study we will deal with the uniform intensity. Furthermore, we will mention a rather important concept, that of *stochastic dependence*. Stochastic dependence is the interaction among the points in our point pattern. Usually we expect a stronger dependence among the points closer to one to another. Most times the observed point pattern which is a vector **x** with coordinates $\mathbf{x}_\mathrm{i} = (\mathrm{x}_\mathrm{i}, \mathrm{y}_\mathrm{i})$ is dealt as realization of a random process **X** in the two dimensional area. Extending the meaning of stochastic process we set the stochastic point process as a random set of points with random locations. Our primary goal is to estimate the distribution of **X**.

First-order and second-order properties characterize the comportment of a general spatial distribution process. The spatially varying intensity of a point pattern can be described with *first-order properties*. Intensity is defined as the mean value of the distribution at locations throughout the region of interest (Diggle, 1983). The methods of second order states that the marginal distributions of points have a constant frequency but the marginal densities of all points is such that marginal distributions of points is not independent (Brunsdon, 2016). This procedure can describe infectious diseases where a case can follow others in a nearby area. Usually the function of distance is used (Ripley, 1981) and is set as:

$$K(d) = \frac{E(N_d)}{\lambda}$$

Where N_d are the number of events x_i in distance d from a randomly chosen event out of all events $\{x_1, \dots., x_n\}$, and λ the intensity of the process measured in events per unit area. If we assume that the distribution of $\mathbf{x_i}$ are independent and the marginal densities are uniform it is often called *Poisson Process* or *Complete Spatial randomness (CSR)*. Then, we would expect a large number of events, at a distance d of a randomly set of events to be equal to the frequency λ multiplied by the area of a circle of radius d that is:

$$K_{CSR}(d) = \pi d^2,$$

This constitutes a benchmark to assess the clustering of the processes. To conclude, for a process having a K – function $K(d)$, if $K(d) > K_{CSR}(d)$ this suggests that there is an excess of nearby points so there exists clustering at the spatial scale associated with the distance d.

Spatial dispersion at this scale means that $K(d) < K_{CSR}(d)$ (other points are less likely to appear nearby one point).

When working with sample data then K – function is not known and therefore must be estimated. If d_{ij} is the distance between $\mathbf{x_i}$ and $\mathbf{x_j}$ then an estimate of the K Function (Ripley, 1976) is:

$$\widehat{K}(d) = \widehat{\lambda^{-1}} \sum_{\iota} \sum_{j \neq i} \frac{2I(d_{ij} < d)}{n(n-1)w_{ij}} \text{ (1),}$$

Where w_{ij} is the area of intersection between a circle centered at $\mathbf{x_i}$ passing through $\mathbf{x_j}$ and the study area A.

Allowing for sampling variation via simulation which referred to envelope analysis we can create the highest and lowest values of $\widehat{K}(d)$. More specifically, we have the potentiality to conduct hypothesis testing using the graphical display of the functions $\widehat{K}(d)$ and $K_{CSR}(d)$. The null hypothesis is the complete spatial randomness (CSR), in other words the homogeneous procedure Poisson and the alternative includes all other procedures. Besag et al (1977) and (Ripley, (1977); Ripley, (1981)) proposed hypothesis testing using the principle of Monte Carlo (Hope, 1968; Besag et al, 1989). One such test is the maximum absolute deviation (MAD) which is the absolute value of the largest discrepancy between the two functions:

$$MAD = max_d \left| \widehat{K}(d) - K_{CSR}(d) \right| \text{ (2)}$$

An alternative test is advocated by Loosmore and Ford (2006) with test statistic:

$$u_i = \sum_{d_k = d_{min}}^{d_{max}} \left[\widehat{K}_\iota(d_k) - \overline{K}_\iota(d_k) \right]^2 \delta_k \text{ (3)}$$

Where $\overline{K}_\iota(t_k)$ is the average value of $\widehat{K}(d)$ over the simulations, the d_k are a sequence of sample distances ranging from d_{min} to d_{max} and $\delta_k = d_{k+1} - d_k$.The specific statistic measures the sum of the squared distance between the functions.

Analysis and Results

The next figure presents the closed primary schools due to the A(H1N1)pdm09 per county. More specifically the regions which had the highest number of closed schools was the region of Attiki (203 schools), the region of Ioannina (108 schools), the region of Kozani (84 schools), the region of Imathia (60 Schools) and the region of Thessaloniki with 34 schools. We can notice that the phenomenon of the epidemic in the student body in primary education was much more intense in the northern part of the country and especially in the regions of central Macedonia and that of Epirus as well. Interpreting the phenomenon of closed schools per area in total 705 primary schools that remained closed we observe the following. In the region of central Macedonia 129 closed schools were reported due to the virus A(H1N1)pdm09. In the region of western Macedonia 92 schools whereas in the region of eastern Macedonia 13 primary schools remained closed. Later in the region of Epirus 129 primary schools closed and in the region of Thessaly 33 schools. By adding the above five regional units we notice that the percentage of closed schools in the northern but also central Greece (taking into account only the region of Thessaly) reaches that of 56,2% closed schools . The magnitude of the intensity of the phenomenon in the above regions is therefore very intense. By adding the region of Attiki which is the 28,8% of closed schools the percentage comes to 85%. Conclusively, the above 6 regions constitute the 85% of the total closed primary schools in Greece.

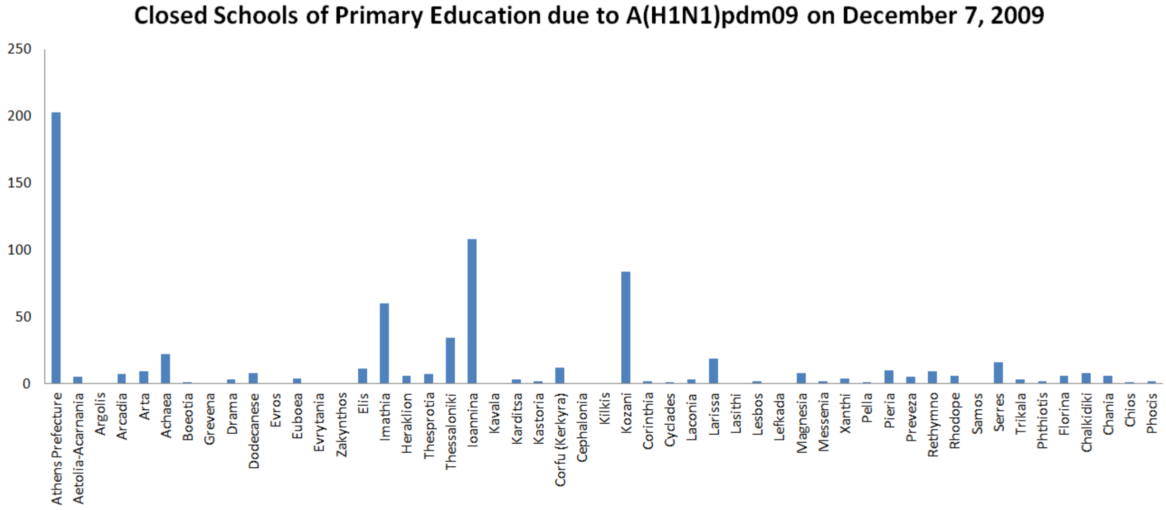


Figure 1. Closed Schools of Primary Education due to A(H1N1)pdm09 on December 7, 2009.

In the next map of Thessaloniki we illustrate the closed schools due to the virus A(H1N1)pdm09 pointing them out with dots. We notice that in the centre of Thessaloniki there is an increased number of closed schools. In order to come to the conclusion to whether there is clustering or not we will conduct the testing according the K- function estimator as well as the hypothesis testing of equations (2) and (3).

Map 1. Closed Primary Schools due to A(H1N1)pdm09 in Thessaloniki.

The result compares the estimated function ($\widehat{K}_{bord}$) to the theoretical function under CSR ($\widehat{K}_{CSR}$). It may been seen that the data appear to be clustered because the empirical K function is greater than that for CSR, suggesting that more points occur close together than would be expected under CSR. Performing with the simulation technique the envelope analysis it can been seen that the estimated K-function for the sample takes on a higher value than the envelope of simulated K-functions for CSR suggesting strong evidence that the locations of closed primary schools do indeed exhibit clustering.

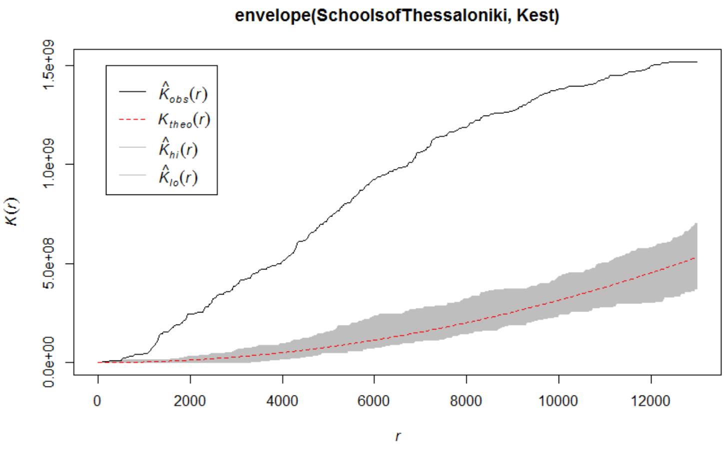
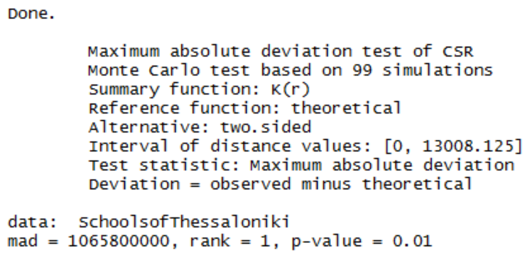


Figure 2. Ripley's K Function plot for Closed School in Thessaloniki.

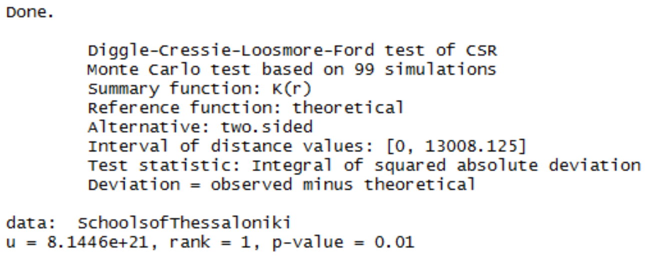
Next Figure shows MAD test in order to evaluate the null hypothesis of CSR. It can be seen that the null hypothesis of CSR can be rejected at the 1% level. At the same conclusion ends up rejecting the null hypothesis with the test statistic of equation (3). Concluding, there exists statistically significant difference from the hypothesis of CSR suggesting intense clustering.



```
Done.

          Maximum absolute deviation test of CSR
          Monte Carlo test based on 99 simulations
          Summary function: K(r)
          Reference function: theoretical
          Alternative: two.sided
          Interval of distance values: [0, 13008.125]
          Test statistic: Maximum absolute deviation
          Deviation = observed minus theoretical

data:  SchoolsofThessaloniki
mad = 1065800000, rank = 1, p-value = 0.01
```



```
Done.

          Diggle-Cressie-Loosmore-Ford test of CSR
          Monte Carlo test based on 99 simulations
          Summary function: K(r)
          Reference function: theoretical
          Alternative: two.sided
          Interval of distance values: [0, 13008.125]
          Test statistic: Integral of squared absolute deviation
          Deviation = observed minus theoretical

data:  SchoolsofThessaloniki
u = 8.1446e+21, rank = 1, p-value = 0.01
```

Figure 3. MAD Test for clustering

Figure 4. Loosmore Test for clustering

Conclusions

In order to plan for future epidemics, nonpharmaceutical interventions should be considered for inclusion as companion measures aiming at earning precious time for the preparation of the vaccines and medications for public health (Markel et al, 2007).

The decision for school closure during a flu epidemic as a mitigation technique is not as easy. Ferguson et al (2006) demonstrated through mathematical simulation models for influenza H5N1 that school closure is effective to reduce peak attack levels but will have little effect on overall attack rates. During the 2009 A(H1N1) pandemic, Cauchemez et al (2009) carried out a multidisciplinary perspective survey for the closure of schools. More specifically, they noted that it is important to plan to mitigate the negative features of closure even if the schools are expected to close as a deliberate policy choice, or simply because of the high levels of staff absenteeism. Possibly, school closure should be taken in a group of parallel measures aimed at better monitoring of the phenomenon. For example as indicated by Seid et al (2007) before arrive an emergency event, a health department improves preparedness by building a combination of skills such as developing and exercising policies and plans, by performing ongoing surveillance and detection (monitoring the health of the community, investigating outbreaks). In the debate on school closures as a measure of prevention there are varying opinions especially after the knowledge learned from the influenza pandemic A(H1N1)pdm09. An interesting point of view is that of Lee et al (2010) who stated that in order to have a significant impact on the overall rate of serological attack the closure of schools may need to be maintained throughout most of the epidemic which is at least 8 weeks. Returning susceptible students back into schools after relatively short school closures (2 weeks or less) in the middle of the epidemic may actually slightly increase the overall attack rate.

In Greece, Sypsa et al (2009) used a stochastic model to assess the impact of various intervention strategies on the spread of the new influenza A(H1N1)pdm09. The article concluded that the combination of antiviral treatment with school closure and social distancing at the assumed thresholds was found to control the spread of influenza A(H1N1)pdm09. Moreover, in Greek territory during the pandemic A(H1N1)pdm09 the number of absences in primary school was at its highest (weeks 46th – 47th) just two weeks before the period that most laboratory-confirmed cases of A(H1N1)pdm09 were reported (weeks 48th – 49th) (Doukissas et al, 2017). This result is quite interesting since the absences can be used as an early warning system of a future influenza outbreak. In particular, Glass et al (2007) concluded that if schools close as soon as 2% of children are infected, it can be a great benefit while if intervention is delayed until infected 20% of children, there is little benefit. However, in Greece the peak of laboratory confirmed cases due to the virus A(H1N1)pdm09 coincides with the peak of closed primary schools. This finding is particularly interesting for the evaluation of a future pandemic.

The technique of point pattern analysis which was developed in this paper may help us as a supplementary sentinel method in a future epidemic in three ways. Firstly, as a quick tool of visualization of closed schools in the study area. Secondly, we can easily locate if there exists concentration (clustering) of the virus at a local level during its first stages of development. Thirdly, using statistical inference we can conduct hypothesis testing of its validity. It is worth noting that point pattern analysis provides us not only with the information of clustering but also with the exact distance (d) where clustering occurs. Therefore, policy-makers can intervene having available a variety of tools such as absences in the municipal level, number of closed schools from the virus, confirmed cases, deaths at local level and region vaccination coverage. It is a holistic approach with more auxiliary “weapons in the quiver” that will provide clearer and optionally stolid picture of the extent of the phenomenon in the course of its development.

Acknowledgments

The authors feel the need to thank professor of the National Public Health School Takis Panagiotopoulos as well as scientific associate Dimitri Papamichail for their scientific contribution, support and guidance in the management of the data that arose in this assignment. Their valuable help was a way to complete the current publication.

References

Besag J.E. and Clifford P (1989). Generalized Monte Carlo significance tests. Biometrika, 76:633- 642.

Brunsdon, C. Comber, L. (2015). An introduction to R for Spatial Analysis and Mapping. Sage publications Ltd.

Cauchemez S, Ferguson NM, Wachtel C, Tegnell A, Saour G, Duncan B, Nicoll A (2009). Closure of schools during an influenza pandemic. The Lancet Infectious Diseases 2009, 9:473-81.

Doukissas L, Kalogirou S., Tsibos C., (2016a). Spatial Patterns of Mortality due to the virus H1N1 during the pandemic in Greece in 2009. European Regional Science Association, Greek Section, Panteion.

Doukissas L., Papamichail D., Kalogirou S., Panagiotopoulos T., (2017). Spatial and Temporal Study of the Absenteeism in Primary Education during the Pandemic of Influenza A (H1N1) in 2009. The case of Central Macedonia. Cities and regions in a changing Europe: challenges and prospects 5-7 July 2017, Panteion University, Athens, Greece.

European Center for Disease Control and Prevention (ECDC), (2010). The 2009 A(H1N1) Pandemic in Europe, Special Report.

Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS (2006). Strategies for mitigating and influenza pandemic. Nature, 442:448-52.

Glass K, Barnes B (2007). How much would closing schools reduce transmission during an influenza pandemic? Epidemiology, 18:623-38

Glezen WP., (1996). Emerging infections: Pandemic Influenza. Epidemiol Rev.; 18:64-76.

Health Protection Agency West Midlands H1N1v Investigation Team, (2009). Preliminary descriptive epidemiology of a large school outbreak of influenza A(H1N1)v in the West Midlands, United Kingdom, May 2009. Euro Surveill. Jul 9;14(27):pii=19264.

Hellenic Center for Disease Control and Prevention (HCDC), (2010). Weekly Reports and Epidemiological Surveillance of the Flu During the Period 2009-2010, Department of Epidemiological Surveillance and Intervention. Available at: http://www2.keelpno.gr/images/stories/keelpno/nosimata/H1N1/epaggelmaties/gripi_ebdo_20100526.pdf

Hope A.C.A (1968). A simplified Monte Carlo significance test procedure. Journal of the Royal Statistical Society, series B, 30:582-598.

J. Besag and P.J. Diggle (1977). Simple Monte Carlo tests for spatial pattern. Applied Statistics, 26:327-333.

King LJ (1962). A quantitative expression of the pattern of urban settlements in selected areas of the United States. Journal of Economic and Social Geography 53: 1-7.

Klaiman et al.: Variability in school closure decisions in response to 2009 H1N1 (2011). a qualitative systems improvement analysis. BMC Public Health, 11:73.

Lee BY, Brown ST, Cooley P, Potter MA, Wheaton WD, Voorhees RE, Stebbins S, Greffenstette JJ, Zimmer SM, Simmerman RK, Assi TM, Bailey RR, Wagener DK (2010). Simulating school closure strategies to mitigate an influenza epidemic. J Public Health Manag Pract, 16(3):252-61

Loosmore, N.B. and Ford, E.D. (2006). Statistical inference using the G or K point pattern spatial statistics. Ecology, 87(8):1925-1931.

Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM (2007). Nonpharmaceutical interventions implemented by U.S. cities during the 1918-1919 influenza pandemic. JAMA, 298(6):644-54.

Nishiura H, Wilson N, Baker MG., (2009). Estimating the reproduction number of the novel influenza A virus (H1N1) in a Southern Hemisphere setting: preliminary estimate in New Zealand. The New Zealand medical journal, 122(1299):73-7.

Ripley B.D. (1981). Spatial Statistics. John Wiley and Sons, New York.

Ripley, B. D. (1977). Modelling spatial patterns (with discussion). Journal of the Royal Statistical Society, B 39, 172-212.

Seid M, Lotstein D, Williams VL, Nelson C, Leuschner KJ, Diamant A, Stern S, Wasserman J, Lurie N (2007). Quality improvement in public health emergency preparedness. Annual Review of Public Health 2007, 28:19-31.

Sypsa V, Hatzakis A (2009). School closure is currently the main strategy to mitigate influenza A(H1N1): a modeling study. Euro Surveill, 14(24):19240.

The Computer Technology Institute and Press "Diophantus", ICT, (2009). Database of Daily Absence Recording.

Trifonov, Vladimir; Khiabanian, Hossein; Rabadan, Raul (2009). "Geographic Dependence, Surveillance, and Origins of the 2009 Influenza A (H1N1) Virus". New England Journal of Medicine. 361 (2): 115 119.

WHO, (2009). "Influenza - like illness in the United States and Mexico". Disease Outbreak News. World Health Organization. 2009-04-24. Retrieved from http://www.who.int/csr/don/2009_04_24/en/

Wiegand T, Moloney KA (2004). Rings, circles, and null-models for point pattern analysis in ecology. Oikos 104(2): 209-229