



ÁLLATORVOSTUDOMÁNYI
EGYETEM · BUDAPEST

Department of Animal Hygiene, Herd-health and
Veterinary Ethology
University of Veterinary Medicine Budapest

**Spatial Cluster Analyses of Infectious Animal Diseases
Using VetEpiGIS**

By

Thomas Lillis

Supervisor: Dr Norbert Solymosi, associate professor, PhD

Budapest, 2017

Table of Contents

Introduction	1
Literature Review	3
1. The historical use of spatial analysis with regard infectious diseases	4
2. The application of GIS as an aid in infectious diseases analysis;	12
3. Future application of GIS in Veterinary Epidemiology	15
4. A review of relevant literature pertaining to African Swine Fever case study	16
A) Distribution	16
B) Diagnosis and Control Measures.....	19
Goals/Questions	22
Materials and Methods	23
Results and Discussion of the Analysis of African Swine Fever in Ukraine	33
Conclusions.....	40
Summary/Abstract	43
Bibliography.....	44
Acknowledgments.....	52
Appendices	53
<i>Appendix 1. Attribute Table Utilized in the Analysis of African Swine Fever.....</i>	<i>53</i>
Electronic Licence Agreement and Copyright Declaration.....	Error! Bookmark not defined.

List of Abbreviations

GIS – Geographical Information Systems

QGIS – Quantum Geographical Information Systems

GPS – Geographical Positioning System

OIE – World Organisation for Animal Health

FAO – Food and Agriculture Organization of the United Nations

ASF – African Swine Fever

ASFV – African Swine Fever Virus

Introduction

“The first law of geography: Everything is related to everything else, but near things are more related than distant things” (Tobler, 1970).

The use of statistics in the analysis of infectious diseases has been in use for over 150 years. The fundamental basis of this analysis is somewhat unchanged but the tools by which one analyses this data has changed substantially, especially with the advent of the personal computer and subsequently geographical positioning systems (GPS). This has led to an increased efficiency for researchers of analysis with regard spatial clustering. Spatial clustering can be defined as the collection or aggregation of objects due to the value of certain attributes (Han *et al.*, 2001). In terms of infectious diseases, spatial clustering is interpreted as the uncommonly high density or concentration of disease outbreaks in an area which would have improbably occurred randomly (Knox, 1989). The clustering of emerging infectious diseases is often described as patterns of occurrence that can cause “significant burdens on global economics and public health” (Jones *et al.*, 2008). The application of cluster analysis is often used as a means of recognizing and surveying possible disease outbreaks.

The importance of spatial clustering is described as the “investigation of possible clustering of disease occurrence is a foundation of epidemiology, providing valuable information on possible causes of the disease of interest and methods that may be used for disease control and prevention” (Ward & Carpenter, 2000). Every epidemiological inquiry’s primary goal is to prevent further outbreaks or spread of a disease by implementing control measures (Goodman *et al.*, 1990) and spatial analysis serves as necessary part of this inquiry.

Spatial Clustering statistics are commonly used in the second stage of investigating possible outbreaks by the CDC – “(1) initial contact and response; (2) assessment; (3) major feasibility study; (4) aetiological investigation” (US Department of Human Health and Services, 1990) (Ward & Carpenter, 2000).

The advent of Geographical Information Systems (GIS) a computer based data collection software that displays, manages and analyses data that is geographically referenced has led to a major progression in spatial analysis (Smith *et al.*, 2015). GIS has taken a number of

forms with subsidiaries of the system developed like Quantum Geographical Systems QGIS, ArcGIS, ArcMap, ArcView and Google Earth. Due to the fact, that software like Quantum Geographical Information Systems (QGIS) is free and open sourced, its applications are far-reaching. As a result, plug-in suites like VetEpiGIS have been developed. The remit of this paper includes the application of this plug-in VetEpiGIS within QGIS model in the provided case study.

In the field of veterinary epidemiology, GIS uses include (Mengistu & Hailes, 2017):

- Animal disease surveillance and monitoring.
- Information recording and reporting
- Emergency outbreak control strategies
- Modelling the distribution and spread of diseases

In our era of globalization and the resultant movement of peoples, animals and produce, the ever-increasing risk of pathogen distribution has highlighted a need for the evolution of new methods of disease analysis and control (Pfeiffer & Stevens, 2015). Spatial epidemiology defined as “the description and analysis of geographic variations in disease with respect to demographic, environmental, behavioural, socioeconomic, genetic and infectious risk factor” a subfield of spatial analysis, has become the cornerstone in disease control (Elliot & Wartenberg, 2004). Spatial epidemiology itself can be subdivided into:

- Disease mapping.
- Studies on geographic correlation.
- The use of clustering in surveillance.

Initial studies in disease control used disease mapping as simple way of representing standardized mortality/morbidity ratios of a disease. Historically this was used in public health studies in the early 20th century in tracking outbreaks of cholera, cancer rates (in England) and Typhoid (Snow, 1855). This has progressed to bulk studies using larger and larger data sets with the dawn of the computer age.

Literature Review

As an ever-evolving niche of statistical analysis, whether in software or hardware that crosses varying disciplines, spatial clustering techniques needs regular reviews in order to gauge their effectiveness. The question is always which method is correct or “more correct”? However due to the ranging disciplines using spatial clustering methodology one cannot assume one method fits all or is better than all others. Taking into consideration the recently introduced techniques and methods in the field, provides one with the best direction to undertake this research. The objective of this literature review is to provide an overview of the history of the use of geographical spatial analysis in the field of veterinary epidemiology and its application in the prevention and control of outbreaks.

The sequence of this review was set out as follows:

1. The historical use of spatial analysis with regard infectious diseases.
2. The application of GIS as an aid in infectious diseases analysis and VetEpiGIS.
3. Future application of GIS in Veterinary Epidemiology
4. A review of relevant literature pertaining to African Swine Fever case study

1. The historical use of spatial analysis with regard infectious diseases

Maps have historically been the basis of epidemiological inquiries. They are at their core, easy to understand and eye catching, thus forming the foundation of spatial analysis. In an outbreak of cholera in London in 1854, John Snow demonstrated that the cholera was a water-borne pathogen by mapping out the outbreaks in an area of London and then overlay them with streets and nearest water pumps. In doing so, he was able to predict which pump was the source of the pathogen (Snow, 1855).

It was stated (Ward & Carpenter, 2000) that map-construction forms the basis of the “visualization of disease pattern”. It is useful in:

- Comprehending the epidemiological nature of a disease
- Describing the pathogen quantitatively
- Comparing and interpreting the various maps and plots.

Though this is often difficult and sometimes an invalid method, it shapes the core of most epidemiological investigations. It also provides the researcher with a visual representation of obvious clustering of disease events, which can be conveyed easily to others. Diseases often form clusters and these are the foundation of many analyses. They are often valuable in the examination and control of public and animal health events (Kulldorff & Nagarwalla, 1995).

The techniques most often used in examining spatial distribution of diseases include:

- Spatial Scan Statistic (Bernoulli or Poisson Distribution)
- Global Moran’s I
- Geary’s C
- Local Moran’s I (LISA)
- Nearest Neighbour
- Cuzick and Edwards Test

The distribution of a dataset can be described as random, uniform, or clustered. These can be calculated with some of the following spatial clustering tests. Average Nearest Neighbour Test is described as the geographical distance between two centroids i.e. the centre of mass of a geometric object of uniform (Mitchell, 2005). In the case of veterinary

epidemiology, it could take the form of an outbreak location between two neighbouring farms or affected regions. All the affected farms nearest neighbour distances are calculated and an average is computed and given as index. This index is calculated as ratio of observed distance divided by the expected distance. If the average of these lengths is below the hypothetical random distances (i.e. <1), the analysed farms are deemed clustered (Ward & Carpenter, 2000). If the mean lengths are above the random (i.e. >1), the farms are deemed dispersed (or non-clustered) (Ebdon, 1985). One rule associated with this type of analysis is that the dataset being assessed must be able to situate anywhere within the area of analysis and are independent of each other. The study area must be free of obstacles or barriers that may hinder the movement or spread of a subject (in this case an animal or pathogen) (Mitchell, 2005).

Another test of significance is the Cuzick and Edward Test. Its outcome quantifies the possibility that clustering is occurring in a subpopulation in an already clustered population (Cuzick & Edwards, 1990). It also analyses the clustering of subpopulations. Its applications in veterinary epidemiology have included the analysis of papillomatous digital dermatitis clustering (Rodriguez-Lainz *et al.*, 1996) and *Pasteurella multocida* and *P.haemolytica* resistant strains in US farms (Singer *et al.*, 1998).

Methods used (not exclusively used in this paper) for the examination of clustering events interacting with time-space in veterinary epidemiology include (Ward & Carpenter, 1999):

- Mantel Test
- Barton's Method
- Nearest Neighbour
- Knox Test (and Modified Knox Test) (most used)

Though these are temporal tools are important and useful methods of analysis, often problems arise due to the lack of statistical power. This is as a result of relatively small datasets.

The aim of this review is to describe in detail the four algorithms which form the core of the VetEpiGIS-stat plug-in. These include Global Moran's I, Geary's C, Local Moran's I and the Kulldorff Spatial Scan Statistic.

A. Geary's C

In the examination of distribution, a spatial hypothesis, known as spatial autocorrelation, can be identified in order to determine whether a dataset is randomly distributed or clustering is present. In this form, it tests the hypothesis using a single variable at multiple locations (Hungerford, 1991). A method of spatial autocorrelation of a whole map as a "single qualitative value" has been labelled as Global Spatial Analysis or Global Clustering Analysis (Robertson & Nelson, 2014). The two most common forms of testing to this scale are Geary's C and Global Moran's I (Robertson & Nelson, 2014).

Autocorrelation was defined (by Durr & Gatrell, 2004) as "...a tendency for nearby spatial units to record similar values". This tendency towards similarity has led to the development of a number of statistical techniques in order to quantify it. This includes the Moran's I coefficient. With regard to spatial autocorrelation, pairs of random variables attained values at locations a prescribed distance apart. These values were either positively or negatively correlated giving a more or less similarity to the randomly associated pairs (Legendre, 1993). In terms of infectious diseases, "positive spatial autocorrelation" may infer the infectiousness while the negative correlation may suggest dispersal or inadequate surveillance mechanisms in a particular country (Robertson & Nelson, 2014).

Geary's C or Geary's Ratio is a statistical instrument developed by R. C. Geary, (his published material in 1954 describes it as "contiguity ratio c") and was used to determine whether the dataset within a region is randomly distributed or a pattern is formed (Geary, 1954) (Jeffers, 1973). The algorithm, attached below, can be utilized in both one or multidimensional applications (Geary, 1954) utilizing this autocorrelation hypothesis (Hungerford, 1991). The applications of Geary C in veterinary epidemiology have mainly been to identify large scale clusters (Moore & Carpenter, 1999). Its uses have been described in several veterinary epidemiological papers including a study on the prevalence of anaplasmosis in cattle (Hungerford, 1991) and the clustering of pseudorabies cases (Austin & Weigel, 1992) (Moore & Carpenter, 1999).

However, a number of shortcomings have been identified in the use of Geary's C. "Low samples sizes and non-nominal data" can affect its statistical significance (Robertson & Nelson, 2014). Moreover, data points within clusters observed to be very closely related may not be taken into account and thus the detection of hot spots can be overlooked (Moore & Carpenter, 1999).

$$C = \frac{(n-1)}{2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Figure 1 Geary's C or Geary's Ratio Algorithm (Geary, 1954)

B. Global Moran's I

Moran's, I was developed by Patrick Moran in 1950 as a form of quantifying autocorrelation within a dataset (Moran, 1950). This statistic “describes the relationship between a variable of interest, such as disease prevalence, that is associated with the spatial location of points e.g. cattle herds” (Perez *et al.*, 2002). This process is carried out by measuring spatial autocorrelation using feature values and feature locations at the same time (Li *et al.*, 2007). This in turn deems whether there is random, clustering or dispersed distribution. In addition, the significance of the distribution that has been calculated is assigned a confidence level with a Z-Score (the degree of deviation from the mean when standard normal distribution occurs) and P-Value (the evidence against the null-hypothesis) (Mitchell, 2005) (Ebdon, 1985). A P-Value of <0.05 and a significant Z-Score, one can assume the Global Moran's I is of strong statistical confidence (Mitchell, 2005). Though it's a powerful tool in the recognition of large-scale clusters, one disadvantage of using Global Moran's I is that it may ignore “close, non-adjacent areas” (Moore & Carpenter, 1999).

Global Moran's I is used as a statistic of inference, therefore one can only interpret the results within the frame of reference of the null hypothesis. The null hypothesis in the case of Global Moran's I is that the dataset being analysed, is distributed randomly or dispersed (Mitchell, 2005). Its applications are far reaching in veterinary epidemiology and is the most common spatial analysis tool for infectious diseases. Bluetongue Virus clustering was analysed using this method in Australian cattle (Ward & Carpenter, 1995). In one study (Perez *et al.*, 2002), researchers used Global Moran's I to determine whether clustering was observed in bovine tuberculosis in regions of Argentina. Another study,

used the statistic to determine whether there was a cluster pattern of the reproductive ratio (R_0) of African Swine Fever in the Russian Federation outbreaks (Iglesias *et al.*, 2016). The same researchers once again implemented the tool to observe clustering of outbreaks in African Swine Fever on the island of Sardinia (Iglesias *et al.*, 2017). This demonstrated its robustness as a spatial analysis statistic.

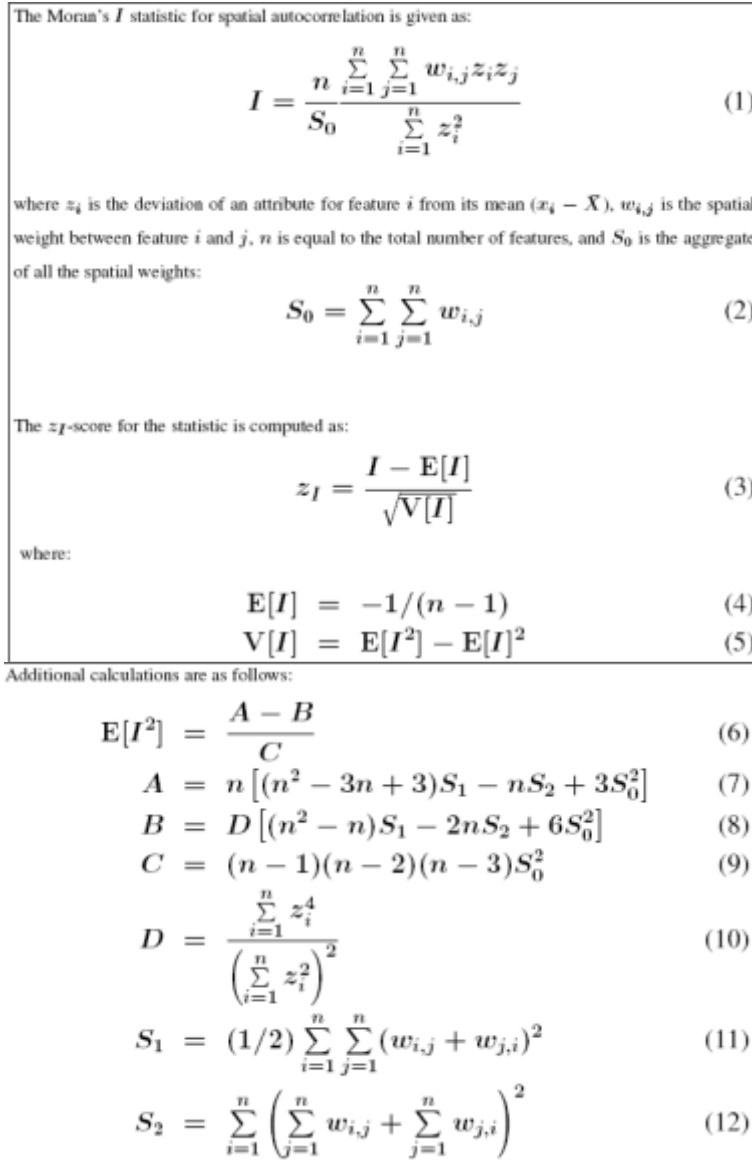


Figure 2 Moran's I and Global Moran's I Algorithm (Mitchell, 2005)

C. Local Moran's I

Global Moran's I can be distinguished from Local Moran's I as the former results in one value to summarize a whole study area while the latter analyses at the smaller local area, resulting in multiple values for each area analysed. If there is no global autocorrelation or clustering in the whole study area, clustering may still occur at a local level (Anselin, 1995). The aggregate values taken from each local study area may yield overall clustering. Like Nearest Neighbour, if the calculated index is more than one (>1), the distribution is considered dispersed and less than one (<1) as clustered. The Z-Score and P-Value are also important in determining Local Moran's I significance. Local Moran's I as such, is used as a tool to analyse map hotspots, identifying clusters and outlier values in diseases (Robertson & Nelson, 2014). Its application in testing spatial autocorrelation and clustering in sub-regions has been utilized in the case of avian influenza (H5N1) in Thailand (Tiensin *et al.*, 2009) and of foot-and-mouth disease in China (Zhang & Zhao, 2015).

The Local Moran's I statistic of spatial association is given as:

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^n w_{i,j} (x_i - \bar{X}) \quad (1)$$

where x_i is an attribute for feature i , \bar{X} is the mean of the corresponding attribute, $w_{i,j}$ is the spatial weight between feature i and j , and:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^n w_{ij}}{n-1} - \bar{X}^2 \quad (2)$$

with n equating to the total number of features.

The z_{I_i} -score for the statistics are computed as:

$$z_{I_i} = \frac{I_i - \mathbf{E}[I_i]}{\sqrt{\mathbf{V}[I_i]}} \quad (3)$$

where:

$$\mathbf{E}[I_i] = -\frac{\sum_{j=1, j \neq i}^n w_{i,j}}{n-1} \quad (4)$$

$$\mathbf{V}[I_i] = \mathbf{E}[I_i^2] - \mathbf{E}[I_i]^2 \quad (5)$$

$$\mathbf{E}[I^2] = A - B \quad (6)$$

$$A = \frac{(n - b_{2_i}) \sum_{j=1, j \neq i}^n w_{i,j}^2}{n-1} \quad (7)$$

$$B = \frac{(2b_{2_i} - n) \sum_{k=1, k \neq i}^n \sum_{h=1, h \neq i}^n w_{i,k} w_{i,h}}{(n-1)(n-2)} \quad (8)$$

$$b_{2_i} = \frac{\sum_{i=1, i \neq j}^n (x_i - \bar{X})^4}{\left(\sum_{i=1, i \neq j}^n (x_i - \bar{X})^2 \right)^2} \quad (9)$$

Figure 3 Local Moran's I Algorithm (Mitchell, 2005)

D. Kulldorff Scan Statistic

Another technique to determine if data points are clustered or random “in a multidimensional point process” is the Spatial Scan Statistic (Kulldorf, 1997). This is achieved by creating a hypothetical “circular window” on a map within which all the data points to be examined are included (Ward & Carpenter, 2000).

The circular window is centred at each point creating a possible cluster with neighbouring points within the window. The technique then uses Poisson and Bernoulli distribution analysis to determine any clustering (Ward & Carpenter, 2000).

Poisson distribution in spatial analysis calculates the probability that an event or number of events may occur in particular space (Haight, 1967). This can be done only if the events can be given a whole number and are independent of each other. Bernoulli distribution describes the density of probability of event occurring by assigning double (Yes/No or True/False) outcome to each data point. (Upensky, 1937). One drawback identified in its application is that small numbers of values increases the likelihood of false positives and there has been shown a limited utility when disease pattern changing rapidly (Moore & Carpenter, 1999). Its applications include the epidemiological analysis of porcine epidemic disease virus (PEDV) in the United States (Alvarez *et al.*, 2016) and in the investigation of the spatial relationship of bovine tuberculosis strains between cattle and badgers (Olea-Popelka *et al.*, 2005).

2. The application of GIS as an aid in infectious diseases analysis;

The geographical location of a disease, outbreak or vector has always been important in the understanding the epidemiological aspects of a pathogen. Since the late 1980s and early 1990s, the availability of personal computers (pcs) and the development of the internet has led to a surge in research in the area of spatial analysis of infectious disease, public health and animal health (Auchincloss *et al.*, 2012).

GIS has substantially increased the number of tools available to analysts in epidemiology (Smith *et al.*, 2015). Though mapping provides a certain amount of visual information, it may not show any apparent spatial relationship between data points. GIS overcomes this problem using methods described below. As a result, GIS is often used a method of proving that an outbreak of a pathogen has occurred.

A previous review on spatial clustering (Auchincloss *et al.*, 2012) states that the most common spatial methods or tools used by researchers in GIS-based spatial analysis includes:

- Spatial Proximity Calculations – calculation of distances
- Aggregation Methods – approximation of summary measures across pre-specified geographic locations
- Cluster Assessments – evaluation of none random spatial patterns
- Spatial smoothing – developing an approximating function that captures important patterns in data.
- Interpolation methods – a method of forming new data points from current data.
- Spatial regression – used to examine the strength or direction of a relationship between variables.

In 1988 one of the first, emergency response-oriented, spatial analysis tools was developed. Its aim was to be able to provide:

- Printable up-to-the minute summary maps.
- Maps of at-risk properties.
- Investigative-insight into airborne transmission of pathogens
- Farm identification within defined regional parameters.
- Scenario based estimations of future epidemics.

This project, known as EpiMan-FMD became an integral part of the Central Disease Control and UK governmental response in the 2001 FMD outbreak (Durr & Gatrell, 2004).

The use of GIS in the event of an epidemic outbreak is varied. It includes:

- Indication of control areas.
- Surveillance zone formation.
- Map production for governmental control agencies and the public.
- Risk analysis of at risk areas/ nearest neighbours.
- Investigation of probable sources of infection.
- Planning surveillance and vaccination schemes.
- Wildlife control measures.
- Cull programs.
- Organization of movement of animals (Durr & Gatrell, 2004).

The use of spatial analytics tools within GIS have been shown to be vital in providing insights into the epidemiological nature of a pathogen. It is most notably pertinent to the area of regional risk factors, economic costs and the effective of current and future control measures.

One important aspect of spatial analysis of an epidemic is the availability of geographical information. The key information required for an effective and accurate analysis include, accurate point data of local outbreaks and accurate locations of farms or other at-risk locations e.g. zoos, petting farms, menageries. New Zealand became the forerunner in the provision of a national database (*AgriBase*) of geocoded farms throughout both islands (Sanson & Pearse, 1997). Due to the fact that this database is necessary for various applications like emergency services, environmental investigations and utility network developments, the database is constantly updated with farm locations and point data on farm-gate locations for physical access to farms. The mapping is then linked to the national land registry and open-sourced making it accurate, free and ideal for disease control investigations (Sanson & Pearse, 1997).

When the outbreak of Foot and Mouth Disease was confirmed in 2001, the UK's Department of Environment, Food and Rural affairs (DEFRA) (formally Ministry of Agriculture, Fisheries and Food) decided to use GIS to analyse the data. However, the farm locational data recorded in the past had been as Ordnance Survey local grid

references, which were then converted to the national map and inputted into GIS. This laborious job involved numerous field investigators locating farms and recording their respective GPS coordinates. The result of which was the formation of a database that would in future aid in the control of future outbreaks in FMD and other pathogens (Durr & Gatrell, 2004).

The availability of disease outbreak reports by the World Organisation for Animal Health (OIE) has provided a vital platform in the investigation of the pathogens on a continental or global basis. This collation of data, in a free database, in chronological order for each disease provides the bulk data necessary for spatial and spatiotemporal analysis as well as possible insights into the basic spread of pathogens over a large geographical area.

The objective of the VetEpiGIS programme was to develop a suite of tools for the management spatial information in the investigation of infectious diseases (Mazzucato *et al.*, 2017). In doing so, these proposed, easy-to-use tools, which could allow researchers or veterinary public health officials to implement complex methods of spatial analysis without the years of GIS training usually required. In doing so, quickly organize a framework and execute disease control measures. This basic idea of breaking down the barrier of technological inexperience necessary for this field could help makes epidemiology more accessible to researchers and the public alike. Furthermore, the simplification of GIS specific tasks can remove or stem the amount of errors related to analysis that is particularly prevalent in users who are at an “Enthusiasts” level (URISA, 2013).

The project developed a number of plug-ins to be used in conjunction with the functionalities provided by QGIS. This allowed users to:

1. “Capture, store and manage geospatial data on farms, outbreaks, and disease response measures (e.g. the creation of restriction areas)” (VetEpiGIS-Tool);
2. “Share data among the VetEpiGIS users and between a VetEpiGIS user and a centralized data collector” (VetEpiGIS-Group);
3. “Analyse spatial data, through a set of predefined statistical methods” (VetEpiGIS-Stat).

(Mazzucato *et al.*, 2017)

3. Future application of GIS in Veterinary Epidemiology

Future applications of geographical information systems will involve the use of “data mining” i.e. the examination of large database to generate information (Moore & Carpenter, 1999). This has to some extent already occurred with the implementation of spatial analysis of large scale data in the context of epidemiology in what has been described as the “Big Data Era” (Mooney *et al.*, 2015). The cornerstones of “Big Data” have been described as the “3V’s: high variety, high volume, and/or high velocity information assets” (Mooney *et al.*, 2015). This type of analysis of this size of data could allow epidemiologists to better predict future outbreaks from occurring and thus preventing them.

(Mooney *et al.*, 2015) go on to predict that all future epidemiologists will need basic programming skills in order to apply the information and knowledge gathered into future hardware and software developed.

It was proposed to depart from the tradition of two dimensional maps to the development of new mapping techniques, which may serve to understand and visualize epidemiological data in new ways (Moore & Carpenter, 1999). This has come to a certain realisation with advent of virtual reality software and could open directions in understanding georeferenced data that (Vitek *et al.*, 1996) previously stated would cause a “renaissance” in the use of maps. The link between GIS and virtual reality has been developed with the use of “virtual reality modelling language” (VRML) (Huang *et al.*, 2001) and a “virtual reality GIS analysis platform” has been proposed (Wang *et al.*, 2017).

Though the software and hardware is ever changing, like cloud-based computer and data-capturing, the use of GIS in the study of diseases will become a more essential skill required by veterinary epidemiologists in the future.

4. A review of relevant literature pertaining to African Swine Fever case study

African Swine Fever Virus (ASFV) is an arthropod-borne DNA virus that causes African Swine Fever (ASF). This arbovirus is part of the Asfarviridae family with a single serotype though differences in virulence are observed. ASFV, a non-zoonotic virus, affects domestic pigs and *Sus scrofa* (wild boar) of all ages (Arias *et al.*, 2002) (OIE 1 African Swine Fever Disease Card, n.d). The arthropod vector mainly responsible for transmission of the virus is the soft tick species including *Ornithodoros erraticus* (located in Iberian region) and the portions of the *Ornithodoros moubata* complex. In addition, the virus can be transmitted *per os*, through fomites and bodily excretions. Normally the virus causes 100% morbidity in swine and mortality rates rely on the virulence level. It's normally associated with high mortality (>80-90%) but lower virulence strains can cause chronic forms of the disease. With high environmental resistance levels, this may account for the transmission pathogen over a large geographic distance (OIE 1 African Swine Fever Disease Card, n.d)

A) Distribution

Up to the middle of the 20th century, ASF was confined to the continent of Africa, and endemic to most Sub-Saharan countries. The first recorded outbreak was recorded and then described in Kenya in 1907 and 1921 respectively (Arzt *et al.*, 2010). In 1957, outbreaks were recorded in Portugal, with the Iberian Peninsula becoming endemic. The disease was contained and eradicated in this region in the mid 1990's by using strict slaughtering legislation (Costard *et al.*, 2013).

Up until 2012, it was thought that the virus was endemic only to Africa, Italy and Sardinia with sporadic outbreaks in Europe and America (Costard *et al.*, 2012). However, the introduction of the (highly virulent) genotype II ASFV to Georgia in 2007 saw the progressive spread of the disease across The Caucasus or Caucasia with outbreaks reported in Armenia, Azerbaijan as well as Iran and Russia (Chapman *et al.*, 2011) (Gogin *et al.*, 2013).

According to a Food and Agriculture Organization of the United Nations (FAO) report, Russia had become an endemic country that could place Europe at risk to African Swine

Fever (Khomenko *et al.*, 2013). The spread of ASF within the Russian Federation starting in 2007, according to a FAO 2013 report, attributes the initial spread of the virus through wild boar (*Sus scrofa*) populations (Sylvatic Cycle). After the establishment in the wild boar populations, there was a spill-over into domestic pigs. The majority of which occurred on low biosecurity farms or backyard farms where contact between wild boar and domestic breeds occurred frequently (Gogin *et al.*, 2013). This connection with low security farms was confirmed because the timing of outbreaks reported also reflected the production-seasonality of small-holding farms, with 75% of outbreaks occurring between June and November. The FAO predicted the spread of ASFV into European Union countries due to the high density of wild boar distributed across central and western Europe (Putmen *et al.*, 2011). In addition, the long distance spread of the ASFV across the Russian Federation can be attributed to the practice of swill feeding to domestic pigs i.e. the feeding of food scraps and waste to pigs (Oganesyanyan *et al.*, 2013). This practice has been banned in the EU and most other countries worldwide by 2002. A cycle was formed known as the “Catering Cycle” by the FAO. This describes the process of pig products being produced and sold by farms infected with ASF to consumers in another part of Russia, generally areas with less pig production. It is then consumed and the leftovers/waste is then sold/given to other pig farms in this new vicinity, causing outbreaks in this area thus creating a “Catering Cycle”. These small backyard farms formed a reservoir for ASF in Russia as a result (Khomenko *et al.*, 2013). This prediction of EU outbreaks was correct, with the first outbreaks in European Union countries reported in 2014 in Lithuania, Poland, Latvia and subsequently followed by Estonia in 2015. In June 2017, Czech Republic had its first outbreak of ASF.

The 2013 FAO report cited porous borders between Russia and several Eastern European countries like Ukraine as a source of transfer of infected products. In addition, countries like Latvia and Ukraine have high proportions of low security pig farms that could form similar reservoirs found in Russia (Khomenko *et al.*, 2013). This is confirmed by the mapping analysis carried out by Eurostat. As the figure below illustrates, a high proportion of farms in Eastern Europe can be classified as low income with low numbers livestock (Khomenko *et al.*, 2013). However other articles suggest wild boar movements from Russia to Belarus and subsequently into the European Union as the cause of the spread (Gallardo *et al.*, 2014). The use of swill feeding despite being banned is still prevalent in

many backyard holdings, and illegal trade of pork products may account for the trans-frontier transmission and the establishment of the disease (Sánchez-Vizcaíno *et al.*, 2013).

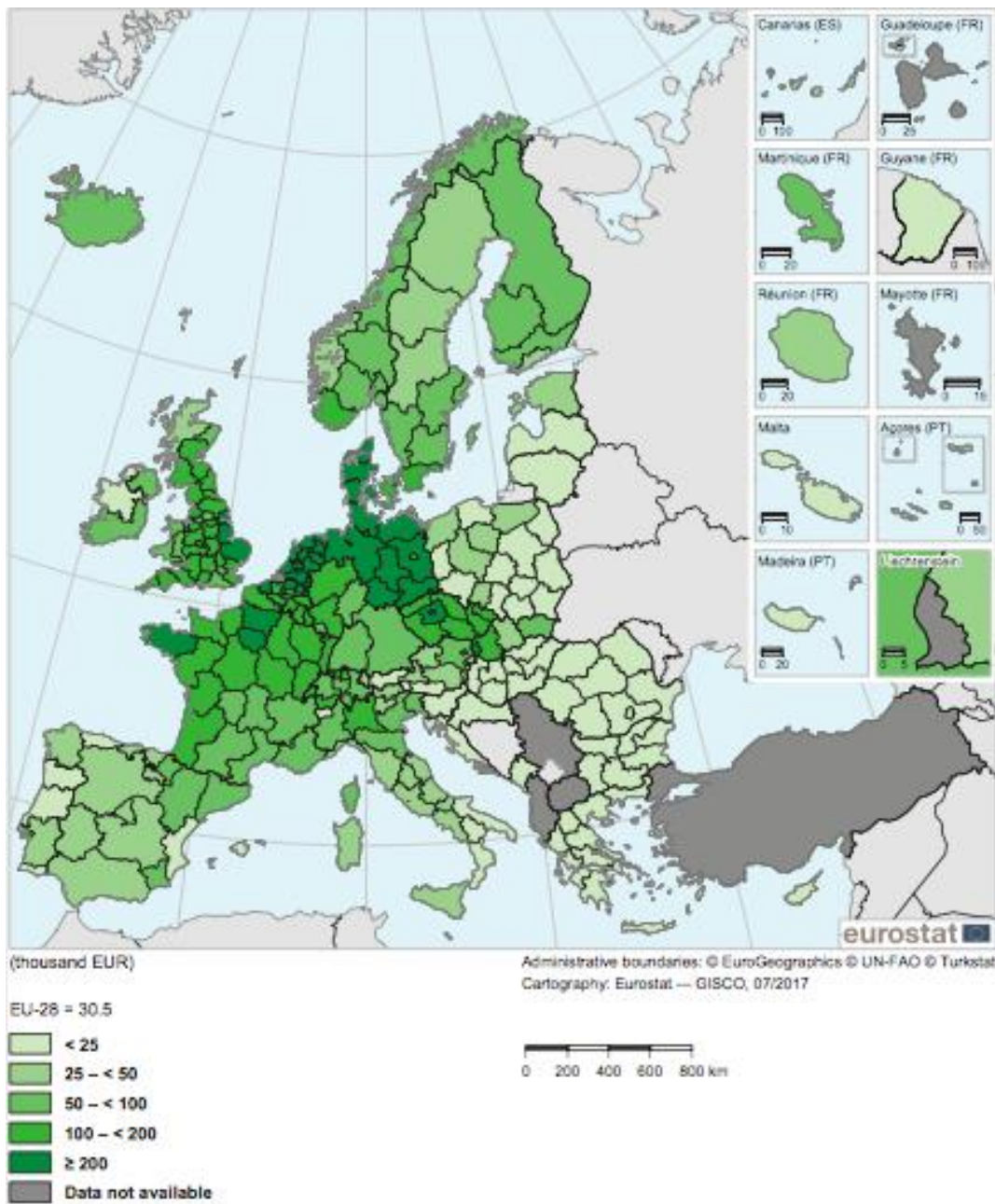


Figure 4 Average Economic Size of Farm Holdings, by NUTS 2 REGIONS, 2013 – (thousand Euro) (Agricultural Statistics at the Regional Level at 2013, EUROSTAT, Europe).

Past outbreaks analysed in Zambia have found that catering waste was the probable cause of many outbreaks. During the detailed analysis, no evidence of sylvatic hosts (warthog and soft ticks) were found leading the investigators to the conclusion that kitchen waste

fed to the pigs was the probable source of infection (Simulundu *et al.*, 2017). This would lead one to think that outbreaks in Eastern Europe could, at least in part, be attributed to the process of swill feeding to pigs.

B) Diagnosis and Control Measures

African Swine Fever cannot be differentiated solely on clinical signs from Classical Swine Fever (Arias *et al.*, 2002). As a result, different diagnostic tools are employed.

Identification tests include:

- Isolation through cell culture inoculation
- Haemadsorption Test (HAD) – in endemic regions
- Fluorescent Antibody Test (FAT)
- Polymerase Chain Reactions (PCR)

Serological Tests include:

- Enzyme-linked Immunosorbent Assay (ELISA)
- Indirect Fluorescent Antibody Test (IFA)
- Immunoblotting Test (IB)

(OIE 1 African Swine Fever Disease Card, n.d)

In recent years, with the geographical spread of ASFV, a number of diagnostic tools were improved (Mur *et al.*, 2016). Researchers in Spain have developed an in-field technique to detect ASF antibodies in less than 10 minutes using immuno-chromatographic devices with very high accuracy (99% specificity and 100% sensitivity) (Perez *et al.*, 2011). Filter paper used on the samples in the field has provided another technique that relies on fewer samples to be taken and transported. It detects ASFV antibodies and antigens in the filter paper (Uttenthal *et al.*, 2013).

Medical prophylactic measures are limited in African Swine Fever due to the fact that no vaccine is available and no treatment has been shown to be effective.

Prevention through sanitary prophylactic measures published by the OIE have formed the guidelines employed by most countries to control the spread of ASF.

The measures vary depending on the epidemiological nature of the country. Non-infected countries are recommended to adhere to strict import legislation with regard to animal transport and to dispose of imported waste (from transport vessels) appropriately.

Endemic or infected countries should prevent the interaction between arthropod vectors like *Ornithodoros* species and swine.

State veterinary measures in the event of an outbreak include:

- Slaughter of all pigs on the holding regardless of infectivity.
- Disposal of waste and remains in the correct manner.
- Supervised cleaning and disinfection of the holding.
- Epidemiological enquiry formation to investigate any sources of infection
- Formation of protection and surveillance zones and control of movement of susceptible species.
- The holding is totally destocked of animals for a period before restocking (OIE 1 African Swine Fever Disease Card, n.d)

According to a 2013 FAO report the lack of treatment or vaccinations for ASF has resulted in many countries attempting to control ASF by introducing movement bans of pigs and pork products. However, these efforts are often limited due to the fact that state veterinary services lack the relevant funding and expertise. Furthermore, the report goes on to infer that economic losses due to culling results in small scale farmers being less likely to report an outbreak due to the lack of state compensation (Khomenko *et al.*, 2013).

The FAO report goes on to make a number recommendations for areas that have not yet been infected. These include:

- Domestic pig and *Sus scrofa* population quantification and distribution analysis
- Put policies in place for cooperation between intergovernmental agencies
- Identification of low biosecurity pig farms in the region vulnerable to ASF
- Rapid containment protocols in place
- Appropriate compensation from the state or European Union to promote compliance from low income farmers.
- Preventive or emergency population control of *Sus scrofa* (Khomenko *et al.*, 2013).

Sardinia has long been affected by African Swine Fever with ASFV being endemic there since 1978. A review of control measures has found a number of factors that are impeding

the total eradication of ASF (Mur *et al.*, 2016). These factors may also be associated with maintaining outbreaks in Eastern Europe. They include:

- High proportion of backyard/small farmers with lacking appropriate knowledge, may not adhere to best practices (e.g. feeding and breeding) and possess deficient holdings in the area biosecurity.
- The practice of free range farming with communal breeding areas within a district of town.
- Black-market trading of swine products.
- Large areas void of humans where reservoirs for disease may be harboured without surveillance.
- Lack of disposal of hunting by-products like offal may account for the spread (Mur *et al.*, 2016).

In summary, the literature with regard to African Swine Fever outbreaks in Europe is vast, and one can infer that the large and varied numbers of scientific studies in this field is a reflection of potentially detrimental impacts it can have on pig production in Europe and the Eurasian region.

Goals/Questions

The goal of this paper in its basic form was to convey the importance of spatial analysis in the field of veterinary epidemiology for those with limited experience of Geographical Information Systems (GIS). This was undertaken by using case study example utilizing the software “VetEpiGIS” developed by Dr Solymosi. This QGIS suite of plugins is a free independent piece of software with an intuitive interface, which allows the user to undertake spatial cluster analysis of datasets within the QGIS programme. The hope of this paper was to illustrate the plugin’s practical abilities and to demonstrate its necessity not only in research based compositions but in the wider practical field like in state veterinary institutions by implementing the four statistical tools of VetEpiGIS for the analysis of geographical pattern in animal health event distribution. The methodology has been described in minute detail in an effort to make the results reproducible and to furthermore be applied to other similar studies for anyone who lacks QGIS or mapping experience. In doing so, this paper may allow persons at a district veterinarian level run cluster analysis “on the fly” to identify clustering in their region and implement or adapt the relevant control measures. At an undergraduate veterinary level, this paper hoped to illustrate the benefits of this field of study as a worthwhile and practical outlet that could be implemented in future veterinarians’ careers.

Materials and Methods¹

In this African Swine Fever case study, the outbreak data points were attained by manually downloading every available outbreak report² for African Swine Fever between 2007 and 2017, which numbered approximately 5020 outbreak reports. These reports are available from the WAHIS Interface on the World Organisation for Animal Health website (OIE 2 Disease Information Report Archive, n.d.)

Pre-2006 reports had limited availability online and probably have not been digitized accordingly. The fact that African Swine Fever has spread into Eurasia and consequently, Russia and Eastern Europe in the years 2007 to 2017, the acquired downloaded outbreak reports encompass these events. In the process of the downloading the animals developed a linux code in order to download the OIE reports rapidly and in bulk. This sped up the process that previously required individually downloading all ~5000 reports by individually downloading each report by clicking on year and finding each outbreak along the list of various disease reports. The linux code can be adapted to any other disease reports on the WAHIS Interface with the possibility of adapting the code to other disease reporting databases. This code alone could prove a very important development in processing bulk data related to infectious data.

Dr Solymosi parcelled the longitude and latitude coordinates of each outbreak report using an R-Script and saving them to a .CSV file with an ID number and X (latitude), Y (longitude) coordinates and Datum (dates) for each report.

1. The .CSV file was added to QGIS Version 2. 18 as a text delimited file with a Semicolon Custom Delimiter, Geometry Definition of Point Coordinates. In the Coordinate Reference System Selector (CRS) WGS 84 / Pseudo Mercator (Authority ID EPSG:3857).

¹ The methodology of this section will be described in detail in order for it to be replicated for those researchers with little to no experience in QGIS mapping and using the VetEpiGIS plugin.

² Please find attached example of an OIE ASF outbreak report at the end of the document.

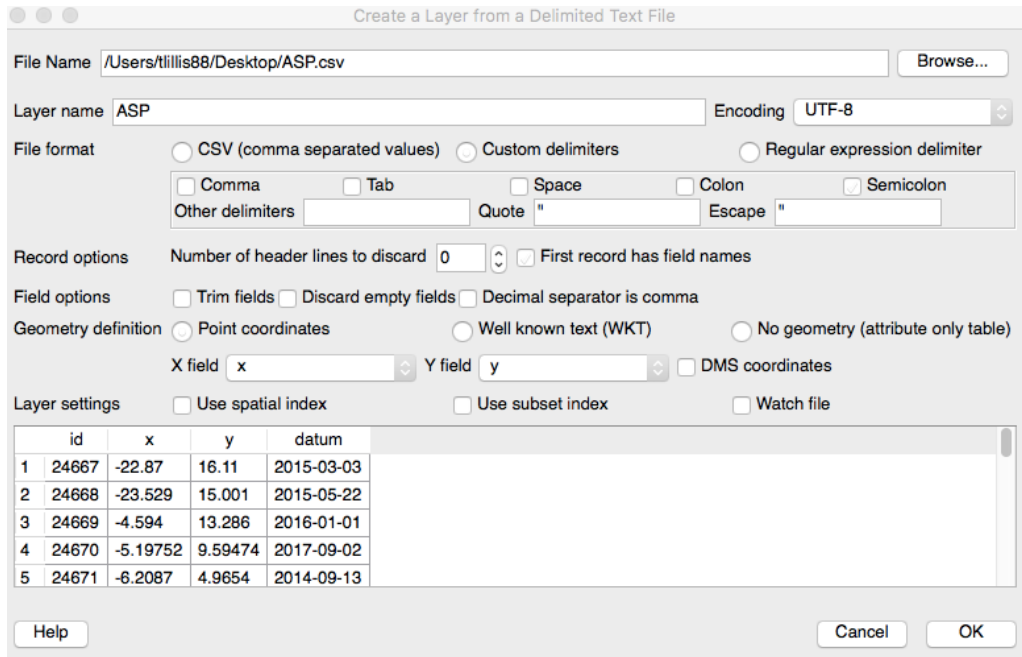


Figure 5 Create a layer from a delimited text file. Note custom delimiters using semicolon and point coordinates geometry definition



Figure 6 Output map formed from text delimited file

The Shapefile “gadm28_levels.shp.zip” was downloaded from the above website. This whole world base map comes in six dissolved layers with one for each level of subdivision in shapefile. The downloaded files were added by dragging and dropping the downloaded files onto the QGIS mapping window. This allows any researcher to add a mapping layer with need for knowing the type being added. QGIS will filter out the files that are not recognized and identify those files as invalid data sources. The valid sources were uploaded to the system with multiple layers of varying detailing shown in the layers section. At first, nothing was visible on the map, the layers were edited on the Layers Order Panel. Furthermore, the Coordinate Reference System (CRS) was set to WGS 84 / Pseudo Mercator (Authority ID EPSG:3857) to coincide with the .CSV data points. A map like that one found below should be generated and various layers selected.

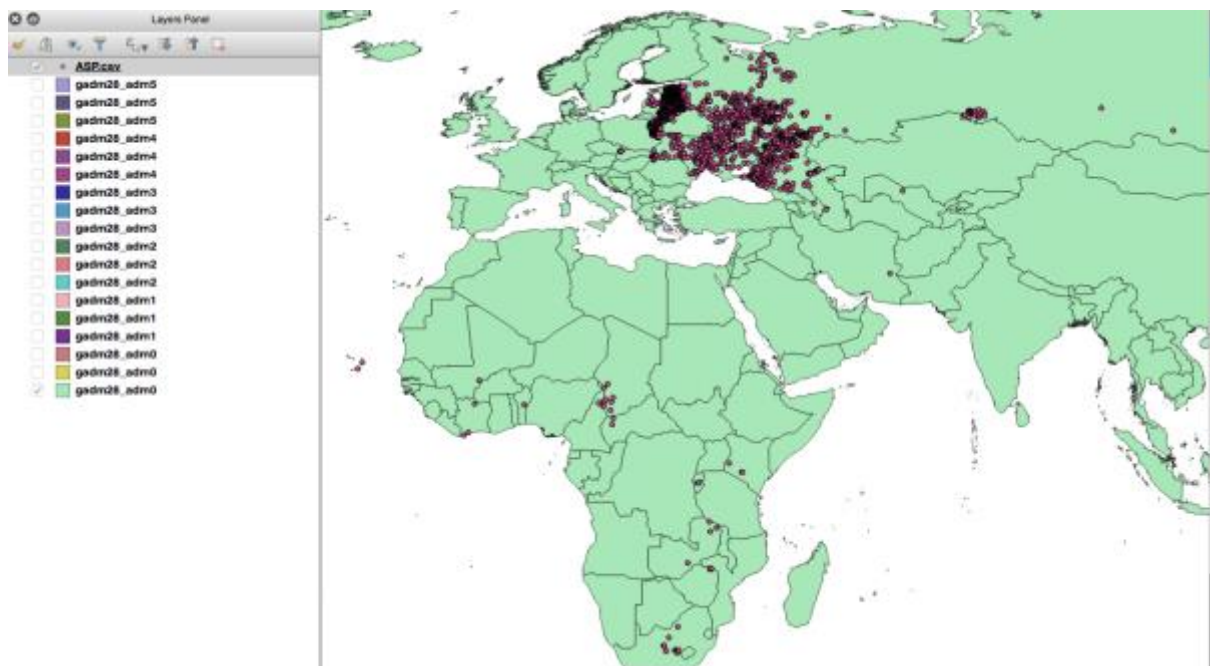


Figure 7 African Swine Fever Dataset Outbreaks and GADM mapping data base layer

3. To analyse of the data points generated by .CSV file, the QGIS plugin was downloaded by opening Plugins dropdown menu and the accessing the Plugin Repository/Database through the “Manage and Install Plugins” Button. VetEpiGIS plugin can be located using the search box and installed. Three toolbars as illustrated below should appear. From left to right the following symbols represent:



Figure 8 VetEpiGIS Suite Toolbar (Left to Right) Stat, Group and Tool

VetEpiGIS-Stat

- A. Global Moran's I/Geary's C
- B. Local Moran's I (LISA)
- C. Scan Statistic

VetEpiGIS-Group

- D. Create Outbreak Layer with a dropdown menu to Create Point, Create Polygon Drawing and Add Element to Outbreak Layer.
- E. Create Point of Interest
- F. Create Buffers with a dropdown menu to Create Points from POI layer and Create Zones.
- G. Database Layers dropdown menu

VetEpiGIS-Tool

- H. Set Up Working Database
- I. Merging Databases

4. For the purposes of this dissertation, the focus was put on VetEpiGIS-Stat, which allows the user to avail of Global Moran's I, Geary's C, Local Moran's I and a Scan Statistic tools.

In this case study, Ukraine was identified as a viable example to analyse the possibility of clustering patterns of African Swine Fever from the outbreak reports published. The following methodology was undertaken to investigate clustering in the relevant districts. the above methods, the following tasks were undertaken.

Using the *Select Features by area or single click* tool was selected and Ukraine was highlighted, its features and districts copied and pasted as new vector layer and assigned as a "Study Area". The gadm28_adm2 map with the detailed districts of Ukraine was selected

using a Query Builder. In doing so, this illustrated the regions of Ukraine only, thus removing the rest of world map from the viewer. The ASF data set was then filtered to show only data points or outbreaks that had occurred within Ukraine using the *Select by Location* function on the *Attributes Toolbar*. This gave the researchers the ability to then analyse the relevant filtered data set within the districts of Ukraine alone. The result of which was illustrated below in cartological form below.

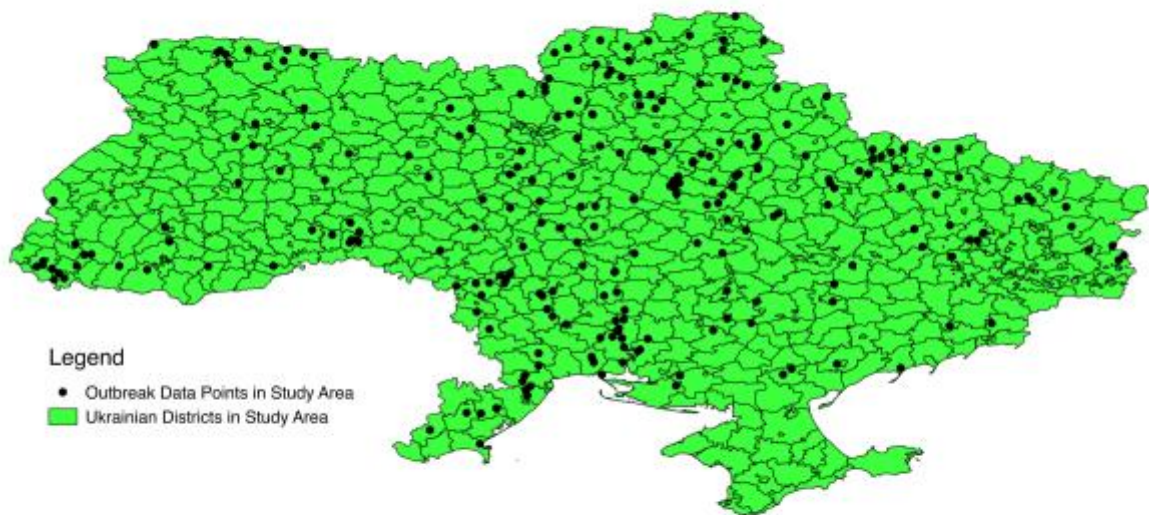


Figure 9 Map of Ukraine including the state's districts and the distribution of African Swine Fever Outbreaks

Within the QGIS *Analysis Toolbar*, the *Count Points in Polygon* tool was selected and the data points within the study area (Ukrainian districts) were counted and labelled “Study_dat”.

The attribute table was then edited with the area of each district in Ukraine calculated with the addition of a *New Field* to the table labelled “Uarea’ The number data points per district area were then calculated and added to the table using the *New Field* tool and the

Dialog Expression page with the following expression “NUMPOINTS”/”Uarea”. This expression in the attribute table was labelled “Casparea”.

To summarize a number of extra fields were added to the original attribute:

- A. NUMPOINTS – the number of data points (outbreaks) within each district.
- B. Uarea – the area of each district.
- C. Casparea – the number of points per area within each district.

Following this, the 4 tools within the VetEpiGIS-Stat plugin were applied as follows:

Global Moran’s I and Geary’s C

The dataset was selected and the Global Moran’s I/Geary’s C tool was selected (see previously for the relevant icon) and Casparea was selected from the Data Field giving an output as follows.

Moran's I: 0.023924266769

Expectation: -0.0015923566879

Variance: 0.000194209111579

Moran's I standard deviate: 1.83100025213

p-value: 0.0335502527224

Geary's c: 0.266032823049

Expectation: 1.0

Variance: 0.139526315877

Geary's c standard deviate: 1.96493674939

p-value: 0.0247107793317

Local Moran's I

The dataset was selected and the Local Moran's I tool was selected (see previously for the relevant icon) and Casperea was selected from the Data Field. The output of the test was then illustrated on the map viewer by editing the properties of the data set. This was achieved by using a Categorized Style and the P-values of less than 0.05. Two classifications were labelled: Significant and Not Significant. The Figure below demonstrates the output.

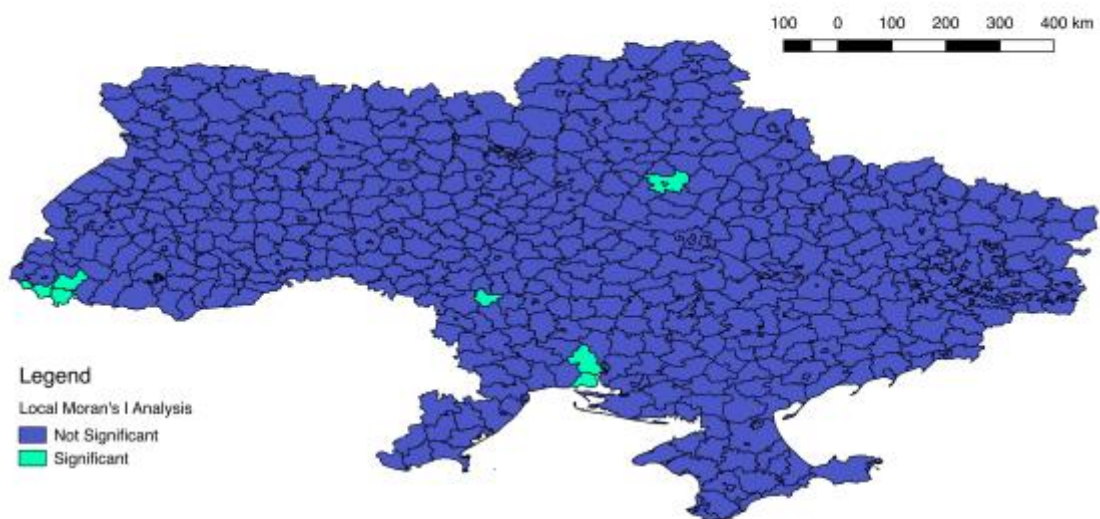


Figure 10 Map of Ukraine illustrating the statistically significant regions (p -value > 0.05) of clustering in local districts using Local Moran's I analysis

Kulldorff Scan Statistic

Finally, the Kulldorff Scan Statistic was employed. However, in order to utilize this tool, the Expected Case Number had to be calculated for each district. This was achieved by adding a new field to the attribute table. This employed the expression dialog and employed this calculation: “Uarea” * (sum(“NUMPOINTS”)/sum(“Uarea”))
i.e. the area of the districts multiplied by the sum of the number of data points divided by the sum of the area of the districts of Ukraine. This new field was labelled “Expnum” and saved.

The Kulldorff Scan Statistic tool was selected and the following input was entered:

- Likelihood Type: Poisson
- Case Field: NUMPOINTS
- Population Field: Uarea
- Expected Case Number: Expnum

The following output was generated after approximately 15 minutes with a default setting of an upper bound of 0.50.

Most likely cluster details:

Feature IDs: 475, 476, 486, 490, 480, 491, 489, 474, 473, 478, 477, 38, 31, 483, 493, 492, 487, 488, 23, 482, 44, 481, 46, 42, 479, 30, 425, 25, 45, 494, 436, 29, 485, 484, 34, 47, 450, 43, 422, 495, 440, 39, 40, 36, 35, 438, 448, 432, 445, 437, 193, 33, 424, 181, 28, 449, 429, 427, 37, 209, 442, 292, 446, 444, 443, 423, 26, 273, 204, 198, 24, 447, 4, 185, 27, 426, 41, 439, 441, 428, 190, 263, 191, 2, 433, 32, 199, 282, 21, 281, 192, 435, 270, 434, 431, 430, 207, 0, 208, 189, 268, 269, 183, 206

Case number: 89

Population: 173619866263

Expected case number: 47.38

SMR: 1.88

Log-likelihood ratio: 19.09

Monte Carlo rank: 1

P-value: 0.001

Secondary cluster details:

Feature IDs: 402, 415, 406, 396, 399, 418, 421, 412, 411, 413, 373, 388, 401, 376, 397, 391, 398, 393, 389, 410, 381, 390, 384, 416, 409, 379, 408, 385, 407, 386, 383, 371, 380, 417, 419, 395, 375, 377, 392

Case number: 43

Population: 61087260580

Expected case number: 16.67

SMR: 2.58

Log-likelihood ratio: 15.96

Monte Carlo rank: 1

P-value: 0.001

The output was then illustrated in map form by editing the style properties of the data set. This was achieved by using a Categorized Style and selecting the Primary Cluster data. Two classifications were labelled: True and False, representing the clustering in the Ukrainian districts. This was repeated for the secondary cluster in the output. *Figures 11 & 12* demonstrate the output.



Figure 11 Kulldorff Scan Statistic of Ukrainian ASF Outbreaks – Primary Cluster



Figure 12 Kulldorff Scan Statistic of Ukrainian ASF Outbreaks - Secondary Cluster

Results and Discussion of the Analysis of African Swine Fever in Ukraine

Global Moran's I

Global Moran's I is also known as a correlation coefficient that computes the overall spatial autocorrelation of a data set. The null hypothesis (H_0) used for this statistical tool states that the data points being analysed are randomly dispersed. The index value calculated using VetEpiGIS was 0.02392. Due to the fact that the Global Moran's index value of -1 indicates ideal clustering of varying values, a value of 0 (or more correctly $-1/(n-1)$) equates to no autocorrelation and a value of +1 demonstrates ideal clustering of similar values, one can interpret that these results have attained a certain degree of clustering tending towards the aggregation of similar values (Mitchell, 2005). The statistical significance of this result can be deemed robust due to the fact the standard deviate (1.8310) is positive and the p-value is less than 0.05 (0.0336). In summary, it can be deduced that the Global Moran's I results show a slight positive autocorrelation of African Swine Fever outbreaks in the country of Ukraine to a confidence level of 95%.

Table 1 Results of Global Moran's I Analysis of African Swine Fever in Ukraine

<i>Moran's I</i>	<i>0.023924266769</i>
<i>Expectation</i>	<i>-0.0015923566879</i>
<i>Variance</i>	<i>0.000194209111579</i>
<i>Moran's I Standard Deviate</i>	<i>1.83100025213</i>
<i>p-value</i>	<i>0.0335502527224</i>

Geary's C

The Geary's C ratio was calculated as 0.2660 using the VetEpiGIS-Stat plugin.

In contrast to the Global Moran's I statistic, Geary's c ratio ranges from 0 to 2, whereby a value tending towards 0 indicates positive spatial correlation, a value of 1 indicates no spatial autocorrelation (i.e. values are spatially independent or randomly distributed) and values tending towards 2 are negatively spatially correlated (i.e. dissimilar values aggregate or colloquially as a checkerboard pattern). The null hypothesis used in this statistical tool states that the data points being analysed are randomly dispersed.

As such, it can be deduced that there was a strong, positive autocorrelation of the outbreaks in the Ukrainian study area. With a positive Geary's C standard deviation (of 1.9649) and a confidence level of less than 0.05 (0.0247), one can deduce with a confidence of over 97% that the outbreaks in the study area of Ukraine have a strong, positive, spatial correlation.

Table 2 Results of Geary's C Analysis of African Swine Fever in Ukraine

<i>Geary's C</i>	<i>0.266032823049</i>
<i>Expectation</i>	<i>1.0</i>
<i>Variance</i>	<i>0.139526315877</i>
<i>Geary's c Standard Deviate</i>	<i>1.96493674939</i>
<i>p-value</i>	<i>0.0247107793317</i>

Local Moran's I Analysis

The Local Moran I is a statistic used to identify clusters and outliers at a local level. The results as graphed below show a significant clusters in the following districts (see *Table 4*).

Table 3 Results of Local Moran's I analysis sorted in terms of decreasing, statistically significant districts (p-value >0.05), generated using the VetEpiGIS-Stat plugin.

District Name	Li	E_Li	Var_Li	Z_Li	p_value	Neighbours	Influence
Poltava Lubens'ka	8.804516595	-0.00159235	0.31008184	15.81415125	0	High-high	TRUE
Poltava Lubens'kyi	9.519433638	-0.01273885	2.476254409	6.057511701	6.91216E-10	High-high	TRUE
Transcarpathia Berehivs'kyi	3.264782897	-0.00636942	1.23847194	2.939393181	0.001644278	High-high	FALSE
Transcarpathia Vynohradivs'kyi	2.816126638	-0.00477707	0.929013644	2.926693314	0.001712933	High-high	FALSE
Mykolayiv Ochakivs'kyi	1.80431722	-0.00318471	0.619550278	2.296362745	0.010827575	High-high	FALSE
Mykolayiv Irshavs'kyi	2.747650208	-0.00955414	1.857373317	2.023109333	0.021530938	High-high	TRUE
Transcarpathia Irshavs'kyi	2.379290454	-0.00955414	1.857373317	1.752823942	0.039816116	High-high	TRUE
Odessa	2.157228009	-0.00796178	1.547925164	1.74028746	0.040904277	High-high	FALSE

Derived from the statistically significant data as shown in *Table 4*, one can affirm that clustering has occurred in the above districts. The Local Moran's index (*Li*) for these districts has been calculated as above *I* with a high Z-score (*Z_Li*) over 1.74.

Based on significant levels (p-values) ranging from not significant, >0.05, 0.01 to a confidence level of less than 0.001 (99.9%) the following map was developed.

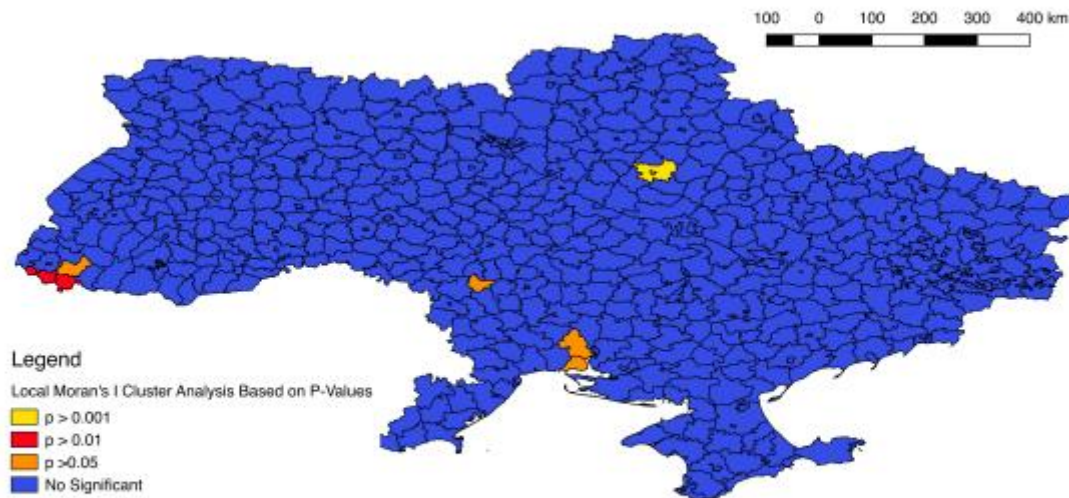


Figure 13 Local Moran's I cluster analysis of African Swine Fever outbreaks in Ukraine

As illustrated there is significant results in 8 districts in Ukraine with the districts of Lubens'ka and Lubens'kyi, within the Poltava region, showing very highly significant clustering with a confidence level of more than 99.9% (p-value > 0.001). The Transcarpathian districts of Berehivs'kyi and Vynohradivs'kyi showed high significant clustering of ASF to a confidence level of 99% (p-value > 0.01). Finally, the districts of Ochakivs'kyi and Mykolasvs'kyi (within the region of Mykolayiv), Irshavs'ky (within the Transcarpathian region) and the Odessa district of Savrans'kyi all indicate significant clustering that can be assessed with a confidence level of 95% (p-value > 0.05).

The Local Moran's I results were further investigated by displaying how the attribute "Neighbours" clustered (as shown in Figure 14), a new case generated during the Local Moran's I analysis. The red districts represent areas where high rates of clustering were observed with a high number of outbreaks recorded. The blue areas represent areas with low clustering and low numbers of outbreaks reported. The pink districts show high areas of clustering with low numbers of outbreaks while purple areas representing districts of low clustering with high number of outbreaks.

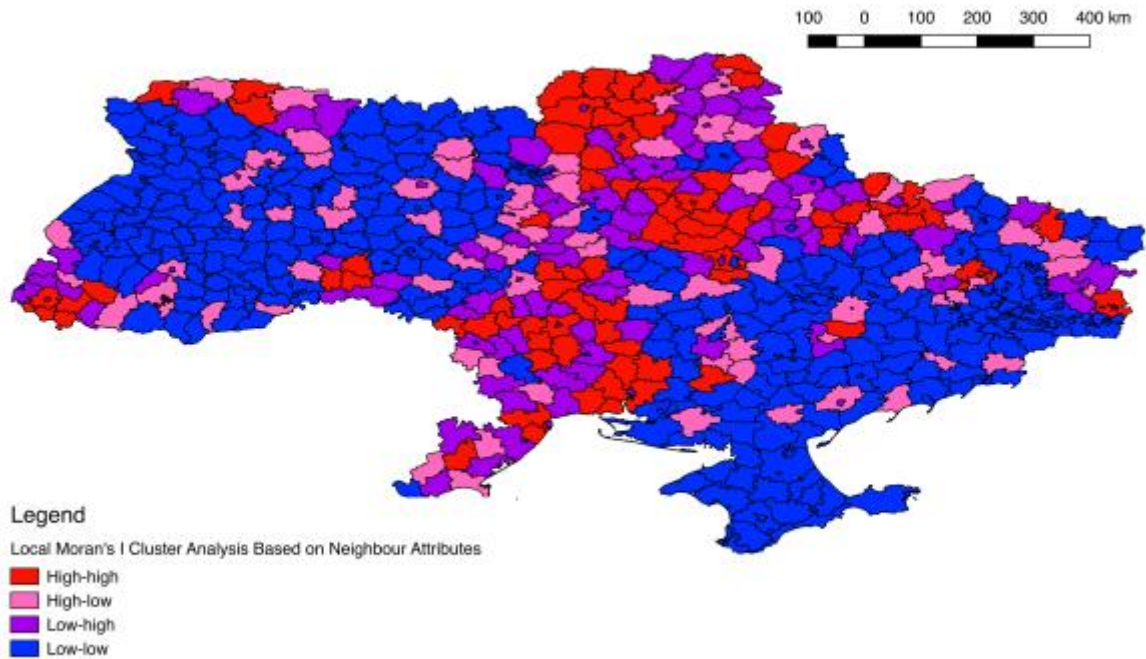


Figure 14 Local Moran's I cluster analysis based on the "neighbour" case field depicting High-High, High-Low, Low-High and Low-Low clustering relationships

As one can see from the *Figure 14*, there are vast regions of high clustering of outbreaks in high numbers in the north, central and south west regions of the country. In terms of the oblasts of Ukraine (the secondary administrative levels or divisions of the state), these high-high clusters were observed in Chernihiv, Sumy, Poltava, Kirovohrad, Mykolaiv and Odesa. Pockets of high-high clustering was also observed in the far east and far west of the country in areas bordering the Russian Federation and Hungary.

Kulldorff Scan Statistic

The Kulldorff Scan Statistic is used in the detection of clusters on a local level. The results have been recorded in *Table 5* below and illustrated in *Figure 15*.

Table 4 Results of Kulldorff Scan Statistic using Poisson likelihood type of African Swine Fever outbreaks in Ukraine generated using the VetEpiGIS plugin on QGIS.

Cluster	Case No.	Population	Expected Case No.	SMR	Log-likelihood Ratio	Monte Carlo Rank	P-value
Primary	89	173619866263	47.38	1.88	19.09	1	0.001
Secondary	43	61087260580	16.67	2.58	15.96	1	0.001

The tabulated data shows the number of cases in each detected cluster, the population field, the expected number of cases, the standardized mortality rate (SMR), log-likelihood ratio, Monte Carlo rank and p-values. At first observation the statistical significance of these clusters are strong with an assigned confidence level of 99.9% (p-value = 0.001).

The null hypothesis of the Kulldorff Scan Statistic is that the outbreaks are randomly dispersed. The data above as a result, rejected that null hypothesis and demonstrated two statistically significant cluster regions as illustrated geographically in *Figure 15*. The log-likelihood ratio was calculated for each cluster of disease events and the highest two values became the primary and secondary clusters. In this case, log-likelihood ratio 19.09 constituted the highest ratio and thus became the most-likely cluster or primary cluster. In other words, this cluster was least likely to occur randomly or by chance. In addition, the second highest log-likelihood ratio of 15.96 constituted the secondary cluster.

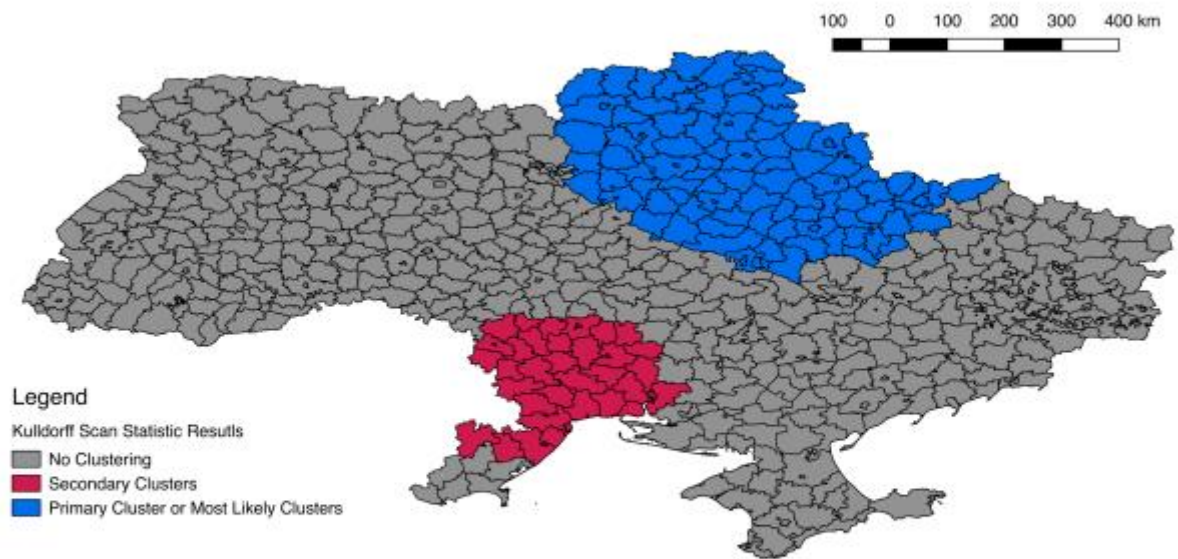


Figure 15 Kulldorff Scan Statistic Analysis of African Swine Fever in Ukraine with primary and secondary Clusters represented

Figure 15 illustrates the extent of the cluster areas. The primary cluster represented in blue is located in the north east of the country and the secondary cluster is observed in the south west of Ukraine.

Conclusions

Significant Clustering of African Swine Fever outbreaks have been identified and proven statistically significant using the spatial analysis tools employed by the VetEpiGIS plug-in. However, one must consider the application of these results in the control and eradication of the disease. According to the OIE (World Organization for Animal Health), Ukraine utilizes the following disease control measures:

- Disease Control Measures
- Border Control Precautions
- Monitoring
- Screening
- General Surveillance
- Movement Control Inside the Country
- Stamping Out
- Zoning
- Control of Wildlife Reservoirs

(OIE 3 WAHIS Interface Disease Control Measures, n.d.)

These measures are implemented on domestic pig but little or no application of these measures is given to wild boar in Ukraine. Unlike ASF-affected European Union countries who have enlisted the above techniques in both domestic and wild swine. Suggestions laid out in a previous article in the control and eradication of ASF in Europe, include the following measures (Sánchez-Vizcaíno *et al.*, 2013):

- Centralized eradication control programme.
- Control of the illegal trade and movement of animals stemming mainly from backyard holdings.
- Education about the risks involved in the use of swill feeding of pigs in backyard farming, prohibit its use and enforce the prohibition.
- Increase compliance of pig farmers with regard (Sánchez-Vizcaíno *et al.*, 2013).

Further evidence has shown that the implementation of a national control plan can be successful with near eradication of ASF in Spain in approximately two years (Costard *et al.*, 2009)

One could infer from the data presented in this paper that hot-spots of clustering could be targeted for increased control measures. For incidence, taking *Figure 13* into account, one could suggest that these eight statistically significant clusters could be targeted. Protection zones and surveillance zones could be set up around these districts. Targeted surveillance within these districts followed by identifying risk factors for the disease spread could prove successful. A strict ban on the movement of pigs within these districts could also be beneficial. This suggestion seems to be the most feasible as the clusters identified in the subsequent figures suggested larger geographic areas of control that may be not feasible for the State Veterinary Medicine Department of Ukraine. Though to fully infer this conclusion, a feasibility report would need to be undertaken.

If the above plan was successful, subsequent targeted control measures implemented in the clusters identified in *Figure 14* (Local Moran's I – Neighbour clustering) and *Figure 15* (Kulldorff Scan Statistic). This may form a basis for the eradication ASF. Clearly this is a simplified spatial analysis of African Swine Fever clustering in Ukraine but could serve as an indicator of how eradication may occur. Further research on the spatial distribution, as well as the spatiotemporal distribution of ASF in Ukraine and the distribution of wild boar in Ukraine would be vital before any control measures are undertaken.

According to an FAO report in 2010, the western half of Ukraine has very high density of backyard pig holdings with these densities ranging from 8.5 to 20 heads/km² (Watch, 2010). To reiterate, backyard holdings are often associated with low biosecurity measures and are at the highest risk of African Swine Fever outbreaks. The presumption that these areas could have large clusters and numbers of outbreaks of ASF is feasible. However, when the clustering maps were observed (*Figure 13 & 14 & 15*), large regions of western Ukraine were devoid of any reported outbreak. The fact that Local Moran's I and Kulldorff Scan Statistic analysis reported no clustering in this area of Ukraine would lead analysts to two possible conclusions. Firstly, the premise that Ukrainian backyard holdings have excellent biosecurity measures on par with the commercial pig industry throughout Europe. Secondly, and the more likely conclusion, is that there is a gross under-reporting of outbreaks in this area. It could be inferred that this is due to a lack of animal-owner knowledge of the disease, the lack of state veterinary regulation (from the State Veterinary Medicine Department of Ukraine) or the lack of state compensation that could deter farmers from reporting the disease in the first instance. The robustness of spatial analysis of infectious diseases depends upon the compliance with outbreak reporting. As such, this

analysis calls into question the statistical strength of the outbreak reports in Ukraine published by the OIE. Suggestions to rectify this problem would include improving the pig owners knowledge, improving the state veterinary institutions and guaranteeing infected holdings are given appropriate state compensation. This would, in turn, enhance the probability of reporting African Swine Fever, thus enabling the state veterinary institutions to comprehend the real impact of ASF and act accordingly. Accurate outbreak recording could allow the re-analysis of this case study material and therefore, give a better understanding of the clustering distribution of African Swine Fever in Ukraine.

Overall, the application of the VetEpiGIS-Stat plugin was effective and efficient. The four statistical tools Global Moran's I, Geary's C, Local Moran's I and Kulldorff Scan Statistic effectively analysed the outbreak data in the Ukrainian study area. One of the premises of this study was that the use of this plugin would allow researchers, with limited GIS experience, an easy to use tool in the spatial analysis of infectious diseases. In its utilization of this plugin, it was experienced that this plugin does in deed, require a certain level of expertise with regard editing data because a number of obstacles being identified during the case study analysis. Due to the fact that no user-manual was available at the time of writing this dissertation, it impacted the progress of the study. However, a user-manual is currently being developed by the creators of the plugin, which will be essential for all those hoping to undertake similar studies with comparable experience using geographical information systems. A goal that was propositioned in this paper was the hope that this software would future veterinarians an easy to use application for the spatial analysis of disease data. The development of an in-depth user manual will be a step forward to attaining this goal.

To conclude, the benefits of VetEpiGIS-Stat far outweigh its drawbacks. The four statistical tools within one plugin with a focus on infectious diseases, has enabled the user to quickly and effectively calculate the relevant data for disease events. Similar software in the GIS universe focus on one tool or don't focus on its application in veterinary epidemiology. The fact that VetEpiGIS-Stat is part of a suite of plugins, allowing for a variety of disease analyses, has identified a much sought-after niche in geographical information systems.

Summary/Abstract

Spatial clustering of disease events have long been analysed and multiple software programmes run various spatial analysis tools in order to determine this clustering. Analysis of the geographical pattern of the animal health event distributions may help the control and eradication of infectious diseases. VetEpiGIS is a free, platform independent QGIS plugin family. The application of the statistical piece of this tool (VetEpiGIS-Stat) will be presented by analysing the distribution of African Swine Fever Virus (ASFV) outbreaks in Ukraine as reported by the World Organization of Animal Health (OIE). Every ASFV outbreak report published by the OIE was downloaded between 2007 and 2017 and Ukraine was selected for the analysis using VetEpiGIS-Stat. The data was parcelled and filtered to represent only outbreaks occurring within the Ukrainian state. The four statistical tools of VetEpiGIS-Stat include Global Moran's I, Geary's C, Local Moran's I and Kulldorff Scan Statistic. The null hypothesis (H_0) that the distribution of African Swine Fever outbreaks in Ukraine was distributed randomly was employed for this analysis. The data was analysed using these tools. Global Moran's I and Geary's C both showed statistically significant ($p\text{-value} > 0.05$) mild positive autocorrelation (*Global Moran's Index* = 0.0239)(*Geary's C ratio* = 0.2660). Local Moran's I analysis identified 8 districts in Ukraine where clustering of ASF was statistically significant ($p\text{-value} > 0.05$). These districts include Lubens'ka, Lubens'kyi, Berehivs'kyi, Vynohradivs'kyi, Ochakivs'kyi, Mykolavs'kyi, Irshavs'kyi, and Sarans'kyi. The Kulldorff Scan Statistic was utilized, a primary and secondary cluster of ASFV were identified ($p\text{-value} > 0.001$) using Poisson distribution. The results calculated from these analytic tools state that the null hypothesis can be rejected and statistically significant spatial clustering occurred African Swine Fever in Ukraine.

Taking into account the limitations in the collated data, the study suggests these spatial analysis results could be applied to areas that require further control measures for ASF. The 8 districts with significant clustering, identified in Local Moran's I analysis could provide a feasible framework on which areas to implement control measures. The results were also compared to data compiled by an FAO report in 2010 that focused on demonstrating the risks ASF to Europe. This comparison infers that gross under-reporting of ASF outbreaks may be occurring in areas of Western Ukraine with high densities of small/backyard pig holdings.

Bibliography

1. Agricultural Statistics at the Regional Level, Eurostat, Europa, Europe Union
URL:http://ec.europa.eu/eurostat/statisticsexplained/index.php/Agriculture_statistics_at_regional_level. Last updated March 2017. Accessed 20 November 2017.
2. Alvarez, J., Goede, D., Morrison, R. and Perez, A., 2016. Spatial and temporal epidemiology of porcine epidemic diarrhea (PED) in the Midwest and Southeast regions of the United States. *Preventive veterinary medicine*, 123, pp.155-160.
3. Anselin, L., 1995. Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), pp.93-115.
4. Arias, M., Sánchez-Vizcaíno, J.M., Morilla, A., Yoon, K.J. and Zimmerman, J.J., 2002. African swine fever. *Trends in emerging viral infections of swine*, pp.119-124.
5. Arzt, J., White, W.R., Thomsen, B.V. and Brown, C.C., 2010. Agricultural diseases on the move early in the third millennium. *Veterinary Pathology*, 47(1), pp.15-27
6. Chapman, D.A., Darby, A.C., Da Silva, M., Upton, C., Radford, A.D. and Dixon, L.K., 2011. Genomic analysis of highly virulent Georgia 2007/1 isolate of African swine fever virus. *Emerging infectious diseases*, 17(4), p.599.
7. Costard, S., Wieland, B., De Glanville, W., Jori, F., Rowlands, R., Vosloo, W., Roger, F., Pfeiffer, D.U. and Dixon, L.K., 2009. African swine fever: how can global spread be prevented?. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1530), pp.2683-2696.
8. Costard, S., Mur, L., Lubroth, J., Sanchez-Vizcaino, J.M. and Pfeiffer, D.U., 2013. Epidemiology of African swine fever virus. *Virus research*, 173(1), pp.191-197.

9. Cuzick, J. and Edwards, R., 1990. Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp.73-104.
10. Durr, P.A. and Gatrell, A.C. eds., 2004. *GIS and spatial analysis in veterinary science*. Cabi.
11. Ebdon, D. *Statistics in Geography* (second edition). Oxford: Blackwell; 1985
12. Elliott, P. and Wartenberg, D., 2004. Spatial epidemiology: current approaches and future challenges. *Environmental health perspectives*, 112(9), p.998.
13. WATCH, F.E., 2010. FAO takes a close look at the pig Sector in Eastern Europe to better understand the threats of African Swine fever.
14. Gallardo, C., J. Fernandez-Pinero, V. Pelayo, I. Gazaev, I. Markowska-Daniel, G. Pridotkas, R. Nieto, P. Fernandez-Pacheco, S. Bokhan, O. Nevolko, Z. Drozhzhe, C. Perez, A. Soler, D. Kolvaso, and M. Arias, 2014: Genetic variation among African Swine Fever genotype II viruses, Eastern and Central Europe. *Emerg. Infect. Dis.* 20, 1544–1547. G
15. Geary, R.C., 1954. The contiguity ratio and statistical mapping. *The incorporated statistician*, 5(3), pp.115-146.
16. Gogin, A., Gerasimov, V., Malogolovkin, A. and Kolbasov, D., 2013. African swine fever in the North Caucasus region and the Russian Federation in years 2007–2012. *Virus research*, 173(1), pp.198-203.
17. Goodman, R.A., Buehler, J.W. and Koplan, J.P., 1990. The epidemiologic field investigation: science and judgment in public health practice. *American journal of epidemiology*, 132(1), pp.9-16.
18. Haight, F.A., 1967. *Handbook of the Poisson distribution*.

19. Khomenko, S., Beltrán-Alcrudo, D., Rozstalnyy, A., Gogin, A., Kolbasov, D., Pinto, J., Lubroth, J. and Martin, V., 2013. African swine fever in the Russian Federation: Risk Factors.
20. Han, J., Kamber, M. and Tung, K.H., 2001. Spatial Clustering Methods in Data Mining: A Survey. Harvey J. Miller and Jiawei Han (eds.), *Geographic Data Mining and Knowledge Discovery*.
21. Huang, B., Jiang, B. and Li, H., 2001. An integration of GIS, virtual reality and the Internet for visualization, analysis and exploration of spatial data. *International Journal of Geographical Information Science*, 15(5), pp.439-456.
22. Iglesias, I., Muñoz, M.J., Montes, F., Perez, A., Gogin, A., Kolbasov, D. and Torre, A., 2016. Reproductive ratio for the local spread of African swine fever in wild boars in the Russian Federation. *Transboundary and emerging diseases*, 63(6).
23. Iglesias, I., Rodriguez, A., Feliziani, F., Rolesu, S. and Torre, A., 2017. Spatio-temporal Analysis of African Swine Fever in Sardinia (2012–2014): Trends in Domestic Pigs and Wild Boar. *Transboundary and emerging diseases*, 64(2), pp.656-662.
24. Jeffers, J.N.R., 1973. A basic subroutine for Geary's contiguity ratio. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 22(4), pp.299-302.
25. Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L. and Daszak, P., 2008. Global trends in emerging infectious diseases. *Nature*, 451(7181), pp.990-993.
26. Kulldorff, M., 1997. A spatial scan statistic. *Communications in Statistics-Theory and methods*, 26(6), pp.1481-1496.
27. Kulldorff, M. and Nagarwalla, N., 1995. Spatial disease clusters: detection and inference. *Statistics in medicine*, 14(8), pp.799-810.

28. Knox, E.G., 1989. Detection of clusters. *Methodology of enquiries into disease clustering*. London: Small Area Health Statistics Unit, 17, p.20.
29. Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm?. *Ecology*, 74(6), pp.1659-1673.
30. Li, H., Calder, C.A. and Cressie, N., 2007. Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, 39(4), pp.357-375.
31. **Pending Publication** Mazzucato, M., Mulatti, P., Solymosi, N., Lorenzetto, M., Ferrè, N., Development of a suite of tools for the management of spatial information related to animal infectious diseases. VetEPidGIS: an open source plug-in suite for QGIS
32. Mengistu, T.S. and Haile, A.W., 2017. Review on the Application of Geographical Information Systems (GIS) in Veterinary Medicine. *Int J Vet Health Sci Res*, 5(4), pp.176-182.
33. Mitchel, A., 2005. The ESRI Guide to GIS analysis, Volume 2: Spatial measurements and statistics. *ESRI Guide to GIS analysis*.
34. Moran, P.A., 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), pp.17-23.
35. Mooney, S.J., Westreich, D.J. and El-Sayed, A.M., 2015. Epidemiology in the era of big data. *Epidemiology (Cambridge, Mass.)*, 26(3), p.390.
36. Moore, D.A. and Carpenter, T.E., 1999. Spatial analytical methods and geographic information systems: use in health research and epidemiology. *Epidemiologic reviews*, 21(2), pp.143-161.

37. Mur, L., Atzeni, M., Martínez-López, B., Feliziani, F., Rolesu, S. and Sanchez-Vizcaino, J.M., 2016. Thirty-Five-Year Presence of African Swine Fever in Sardinia: History, Evolution and Risk Factors for Disease Maintenance. *Transboundary and emerging diseases*, 63(2).
38. Olea-Popelka, F.J., Flynn, O., Costello, E., McGrath, G., Collins, J.D., O'keeffe, J., Kelton, D.F., Berke, O. and Martin, S.W., 2005. Spatial relationship between *Mycobacterium bovis* strains in cattle and badgers in four areas in Ireland. *Preventive veterinary medicine*, 71(1), pp.57-70.
39. Oganesyanyan, A. S., O. N. Petrova, F. I. Korennoy, N. S. Bardina, and A. Gogin, 2013: African swine fever in the Russian Federation: spatio-temporal analysis and epidemiological overview. *Virus Res.* 173, 204–211.
40. OIE 1 African Swine Fever Disease Card
URL:https://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/Disease_cards/AFRICAN_SWINE_FEVER.pdf. Accessed the 19 November 2017
41. OIE 2 Disease Information Report Archive,
URL:https://www.oie.int/wahis_2/public/wahid.php/Diseasecontrol/measures. Accessed 17 November 2017.
42. (OIE 3, WAHIS Interface Disease Control Measures)
URL: https://www.oie.int/wahis_2/public/wahid.php/Diseasecontrol/measures. Accessed 22 November 2017.
43. Perez, A.M., Ward, M.P., Torres, P. and Ritacco, V., 2002. Use of spatial statistics and monitoring data to identify clustering of bovine tuberculosis in Argentina. *Preventive veterinary medicine*, 56(1), pp.63-74.
44. Perez, T., Gallardo, C., Nieto, R., De Mia, G., Vela, C., Bishop, R., Couacy, E., Martin, H., Arias, M. and Sanz, A., 2011, June. Development and preliminary

- validation of a pen-side test based on the use of vp72 protein for ASFV antibody detection. In *6th International Symposium on Emerging and Re-emerging Pig diseases. Barcelona, Spain*
45. Pfeiffer, D.U. and Stevens, K.B., 2015. Spatial and temporal epidemiological analysis in the Big Data era. *Preventive veterinary medicine*, 122(1), pp.213-220.
 46. Putman, R., Apollonio, M. and Andersen, R. eds., 2011. *Ungulate management in Europe: problems and practices*. Cambridge University Press, p. 410
 47. Robertson, C. and Nelson, T.A., 2014. An overview of spatial analysis of emerging infectious diseases. *The Professional Geographer*, 66(4), pp.579-588.
 48. Rodríguez-Lainz, A., Hird, D.W., Carpenter, T.E. and Read, D.H., 1996. Case-control study of papillomatous digital dermatitis in southern California dairy farms. *Preventive veterinary medicine*, 28(2), pp.117-131.
 49. Sánchez-Vizcaíno, J.M., Mur, L. and Martínez-López, B., 2013. African swine fever (ASF): five years around Europe. *Veterinary microbiology*, 165(1), pp.45-50.
 50. Sanson, R., Pearson, A., 1997. Agribase – A National Spatial Farm Database. *Epidemiology of Animal Health*, pp. 31-32
 51. Simulundu, E., Lubaba, C.H., van Heerden, J., Kajihara, M., Mataa, L., Chambaro, H.M., Sinkala, Y., Munjita, S.M., Munang'andu, H.M., Nalubamba, K.S. and Samui, K., 2017. The Epidemiology of African Swine Fever in “Nonendemic” Regions of Zambia (1989–2015): Implications for Disease Prevention and Control. *Viruses*, 9(9), p.236.
 52. Singer, R.S., Case, J.T., Carpenter, T.E., Walker, R.L. and Hirsh, D.C., 1998. Assessment of spatial and temporal clustering of ampicillin- and tetracycline-resistant strains of *Pasteurella multocida* and *P. haemolytica* isolated from cattle in California. *Journal of the American Veterinary Medical Association*, 212(7), pp.1001-1005.

53. Snow, J., 1855. *On the mode of communication of cholera*. John Churchill.
54. Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1), pp.234-240.
55. URISA Model, G.C.M., 2013. GIS Management Institute® GIS Capability Maturity Model. URL <http://www.urisa.org/clientuploads/directory/GMI/GISCMM-Final201309>
56. Uttenthal, Å., Braae, U.C., Ngowi, H.A., Rasmussen, T.B., Nielsen, J. and Johansen, M.V., 2013. ASFV in Tanzania: Asymptomatic pigs harbor virus of molecular similarity to Georgia 2007. *Veterinary microbiology*, 165(1), pp.173-176.
57. US Department of Health and Human Services, 1990. Guidelines for investigating clusters of health events. *Mort Morb Wkly Rep*, 39, pp.1-23.
58. Wang, W., Lv, Z., Li, X., Xu, W., Zhang, B., Zhu, Y. and Yan, Y., 2017. Spatial query based virtual reality GIS analysis platform. *Neurocomputing*.
59. Ward, M.P., Flanagan, M., Carpenter, T.E., Hird, D.W., Thurmond, M.C., Johnson, S.J. and Dashorst, M.E., 1995. Infection of cattle with bluetongue viruses in Queensland, Australia: results of a sentinel herd study, 1990–1992. *Veterinary microbiology*, 45(1), pp.35-44.
60. Ward, M.P. and Carpenter, T.E., 2000. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Preventive Veterinary Medicine*, 45(3), pp.257-284.
61. WATCH, F.E., 2010. FAO takes a close look at the pig Sector in Eastern Europe to better understand the threats of African Swine fever.

62. Vitek, J.D., Giardino, J.R. and Fitzgerald, J.W., 1996. Mapping geomorphology: A journey from paper maps, through computer mapping to GIS and Virtual Reality. *Geomorphology*, 16(3), pp.233-249.
63. Zhang, S. and Zhao, J., 2015. Spatio-temporal epidemiology of hand, foot and mouth disease in Liaocheng City, North China. *Experimental and therapeutic medicine*, 9(3), pp.811-816.

Acknowledgments

I'd like to acknowledge the hard work and patience of Dr Norbert Solymosi as the supervisor of this thesis. His teaching sparked an interest in GIS and veterinary epidemiology for me while attending a lecture of his and I'm grateful for that.

I'd also like to acknowledge the help of my parents Marie and Michael for their help proofing this paper but more importantly instilling the ambition to attend the University of Veterinary Medicine Budapest and providing the means to do so.

HuVetA

ELECTRONIC LICENSE AGREEMENT AND COPYRIGHT DECLARATION*

Name:..... Thomas Lillis.....
Contact information (e-mail):..... tllillis.88@hotmail.com.....
Title of document (to be uploaded):..... Spatial Cluster Analyses of Infectious Diseases Using VetEpiGIS.....
.....
.....
Publication data of document:.....
.....
Number of files submitted:

By accepting the present agreement the author or copyright owner grants non-exclusive license to HuVetA over the above mentioned document (including its abstract) to be converted to copy protected PDF format without changing its content, in order to archive, reproduce, and make accessible under the conditions specified below.

The author agrees that HuVetA may store more than one copy (accessible only to HuVetA administrators) of the licensed document exclusively for purposes of secure storage and backup, if necessary.

You state that the submission is your original work, and that you have the right to grant the rights contained in this license. You also state that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the document has parts which you are not the copyright owner of, you have to indicate that you have obtained unrestricted permission from the copyright owner to grant the rights required by this Agreement, and that any such third-party owned material is clearly identified and acknowledged within the text of the licensed document.

The copyright owner defines the scope of access to the document stored in HuVetA as follows (mark the appropriate box with an X):

- I grant unlimited online access,
- I grant access only through the intranet (IP range) of the University of Veterinary Medicine,
- I grant access only on one dedicated computer at the Ferenc Hutýra Library,
- I grant unlimited online access only to the bibliographic data and abstract of the document.

Please, define the **in-house accessibility of the document** by marking the below box with an X:

I grant in-house access (namely, reading the hard copy version of the document) at the Library.

If the preparation of the document to be uploaded was supported or sponsored by a firm or an organization, you also declare that you are entitled to sign the present Agreement concerning the document.

The operators of HuVetA do not assume any legal liability or responsibility towards the author/copyright holder/organizations in case somebody uses the material legally uploaded to HuVetA in a way that is unlawful.

Date: Budapest, 24 day 11 month 2017 year



Author/copyright owner
signature

HuVetA Magyar Állatorvos-tudományi Archivum – Hungarian Veterinary Archive is an online veterinary repository operated by the Ferenc Hutýra Library, Archives and Museum. It is an electronic knowledge base which aims to collect, organize, store documents regarding Hungarian veterinary science and history, and make them searchable and accessible in line with current legal requirements and regulations.

HuVetA relies on the latest technology in order to provide easy searchability (by search engines, as well) and access to the full text document, whenever possible.

Based on the above, HuVetA aims to:

- *increase awareness of Hungarian veterinary science not only in Hungary, but also internationally;*
- *increase citation numbers of publications authored by Hungarian veterinarians, thus improve the impact factor of Hungarian veterinary journals;*
- *present the knowledge base of the University of Veterinary Medicine Budapest and its partners in a focussed way in order to improve the prestige of the Hungarian veterinary profession, and the competitiveness of the organizations in question;*
- *facilitate professional relations and collaboration;*
- *support open access.*