

# Spatial Statistical Analysis in Cow Disease Monitoring Based on GIS

Lin Li, Yong Yang, Hongbin Wang, Jing Dong, Yujun Zhao, Jianbin He

► **To cite this version:**

Lin Li, Yong Yang, Hongbin Wang, Jing Dong, Yujun Zhao, et al.. Spatial Statistical Analysis in Cow Disease Monitoring Based on GIS. Daoliang Li; Yande Liu; Yingyi Chen. 4th Conference on Computer and Computing Technologies in Agriculture (CCTA), Oct 2010, Nanchang, China. Springer, IFIP Advances in Information and Communication Technology, AICT-345 (Part II), pp.561-566, 2011, Computer and Computing Technologies in Agriculture IV. <10.1007/978-3-642-18336-2\_68>. <hal-01562729>

**HAL Id: hal-01562729**

**<https://hal.inria.fr/hal-01562729>**

Submitted on 17 Jul 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Spatial Statistical Analysis in Cow Disease Monitoring Based on GIS

Lin Li<sup>1</sup>, Yong Yang<sup>2</sup>, Hongbin Wang<sup>3</sup>, Jing Dong<sup>1</sup>, Yujun Zhao<sup>1</sup>, Jianbin He<sup>1</sup>

<sup>1</sup>School of Animal Husbandry and Veterinary Medicine, Shenyang Agricultural University, 110866 Shenyang, P. R. China, E-mail: lilin619619@yahoo.com.cn

<sup>2</sup>Corresponding Author: School of Information and Electrical Engineering, Shenyang Agricultural University, 110866 Shenyang, P. R. China, E-mail: [yangsyau@163.com](mailto:yangsyau@163.com)

<sup>3</sup>School of Animal Medicine, Northeast Agricultural University, 150030 Harbin, P. R. China

**Abstract.** In this paper, we analyze cow endemic fluorosis spatial distribution characteristics in the target district based on combination of spatial statistical analysis and GIS. This disease geographical distribution relates closely to the environmental factors of habitats and the correct analysis of the spatial distribution of the disease for monitoring and prevention plays an important role. However, the amount of monitoring spots is smaller and the monitored data are very limited. How to use the limited data to show the general spatial distribution of the disease is a key issue in building the efficient spatial monitoring method. We use GIS application software ArcGIS9.1 to overcome the lack of data of sampling sites and establish a prediction model in estimation of the disease distribution. Comparing with the true disease distribution, the prediction model has a well coincidence and is feasible and provides a fundamental application for other study on space-related cow diseases.

**Keywords:** Cow disease, GIS, Spatial statistic analysis

## 1 Introduction

The use of geographical information systems (GIS) and spatial statistical analysis in public health is now widespread. GIS is powerful automated system for the storage, retrieval, analysis, and display of spatial data. The system offers new and expanding opportunities for epidemiology because it allows an informed user to choose between options when geographic distributions are part of the problem. Even when used minimally, this system allows a spatial perspective on disease. Used to their optimum level, as tools for analysis and decision making, they are indeed a new information management vehicle with a rich potential for animal epidemiological research. Epidemiologists have traditionally used maps when analyzing associations between location, environment, and the disease (Busgeeth et al., 2004). GIS is particularly well suited for studying these associations because of its spatial analysis and display capabilities. Recently GIS has been used in the surveillance and monitoring of vector-borne diseases (Glass GE et al., 1995; Beck et al., 1994; Richards et al., 1993; Clarke

et al., 1991), water borne diseases (Braddock et al.,1994), in environmental health (Barnes, 1994; Wartenberg et al., 1993; Wartenberg et al.,1992), and the analysis of disease policy and planning (Marilyn et al.,2004).Spatial analysis function of GIS can be widened and strengthened by using spatial statistical analysis, allowing for the deeper exploration, analysis, manipulation and interpretation of spatial pattern and spatial correlation of the animal disease.

Cow endemic fluorosis is a disease of Signs of dental discoloration, difficulty in mastication, bony exostosis and debility. It can cause lameness, dental lesions and illthrift in an extensive cow. The herd was lame and the disease forced the culling of large numbers of cows in northern Australia. Cow endemic fluorosis is biogeochemical disease. Its geographical distribution is related closely to the environmental factors of habitats and has some spatial characteristics, and therefore the correct analysis of the spatial distribution of cow endemic fluorosis for monitoring and the prevention and control of endemic fluorosis has a very important role. However, the application of classic statistical methods in some areas is very difficult because of the pastoral nomadic context. The high mobility of livestock and the lack of enough suitable sampling for the some of the difficulties in monitoring currently make it nearly impossible to apply rigorous random sampling methods. It is thus necessary to develop an alternative sampling method, which could overcome the lack of sampling and meet the requirements for randomness.

## **2 Materials and methods**

In this paper, we analyzed the cow endemic fluorosis spatial distribution characteristics in the target district A (due to the secret of epidemic data we call it district A) based on the established GIS of the cow endemic fluorosis in this district in combination of spatial statistical analysis and GIS. We established a prediction model to estimate the endemic fluorosis distribution based on the spatial characteristic of the density of cow endemic fluorosis. A two-stage random sampling was used. It consisted in generating sampling sites at random and then drawing the required number of cow from the nearest herd to each point, and this, for practical reasons, within a specified radius. In this sampling method, the primary unit was defined as a settlement, watering point or grazing area where animals were expected to be found. Due to a lack of an exhaustive list of these locations in some areas and because of the high mobility of pastoral herds, the unpredictability of their movement, the classical random selection of sites was not feasible. Therefore, there was the need to develop an alternative method that allows random selection of the first sampling units. GIS computer application software (ArcGIS) was used.

## **3 Cow endemic fluorosis data sampling based on GIS**

GIS computer application software (ArcGIS) was used. ArcGIS Spatial analyst is an extension to ArcGIS Desktop that provides powerful tools for comprehensive, raster-based spatial modeling and analysis. With its extension to generate random

coordinates, ArcGIS is able to generate at random the required number of sites within the area where sampling needs to take place, be it at zone, country, region or even district level, simultaneously allowing the application of weight factors to different sub-units of the surveyed area. Each site is identified by its longitude and latitude coordinates. In reality, a generated site is materialized only as a point on the map and therefore a fixed radius is defined around each point in order to determine a geographical area within which the sampling of cow endemic fluorosis is carried out. The determination of length of the radius takes into consideration the cow endemic fluorosis density. A radius below 5 km in a dry area may result in too many sampling sites with no animals, whereas in a densely populated area around permanent watering points and grazing areas, a relatively shorter radius suffices. A surplus of the total number of sites is generated at random to replace target sites that could not be accessed due to natural obstacles or sites with no animals within the defined circle.

#### **4 Spatial cluster analysis of cow endemic fluorosis density**

Spatial cluster analysis plays an important role in quantifying geographic variation patterns. It is commonly used in disease surveillance, spatial epidemiology, population genetics, landscape ecology, crime analysis and many other fields, but the underlying principles are the same. Spatial clustering analysis is the main research field of spatial data mining. We analysis cow endemic fluorosis data spatial cluster using High/Low Clustering tool in ArcGIS which can measures the degree of clustering for either high values or low values. The tool calculates the value of Z score for a given input feature class. A z-score is a measure of the divergence of an individual experimental result from the most probable result, the mean. Z is expressed in terms of the number of standard deviations from the mean value.

$$z = \frac{x - \mu}{\sigma}$$

where: x is ExperimentalValue,  $\mu$  is Mean, and  $\sigma$  is StandardDeviation.

This z-value or z score expresses the divergence of the experimental result x from the most probable result  $\mu$  as a number of standard deviations  $\sigma$ . The larger the value of z, the less probable the experimental result is due to chance. For this tool, the null hypothesis states that the values associated with features are randomly distributed. The higher (or lower) the Z score, the stronger the intensity of the clustering. A Z score near zero indicates no apparent clustering within the study area. A positive Z score indicates clustering of high values. A negative Z score indicates clustering of low values.

## 5 Spatial Autocorrelation analysis of cow endemic fluorosis density

Spatial Autocorrelation is correlation of a variable with itself through space. If there is any systematic pattern in the spatial distribution of a variable, it is said to be spatially autocorrelation. If nearby or neighboring areas are more alike, this is positive spatial autocorrelation, Negative autocorrelation describes patterns in which neighboring areas are unlike. Random patterns exhibit no spatial autocorrelation. Spatial autocorrelation is important, because most statistics are based on the assumption that the values of observations in each sample are independent of one another. Positive spatial autocorrelation may violate this, if the samples were taken from nearby areas. Goals of spatial autocorrelation Measure is the strength of spatial autocorrelation in a map and test the assumption of independence or randomness. Spatial autocorrelation is, conceptually as well as empirically, the two-dimensional equivalent of redundancy. It measures the extent to which the occurrence of an event in an area unit constrains, or makes more probable, the occurrence of an event in a neighboring area unit. We analysis cow endemic fluorosis data Spatial Autocorrelation using Global Moran's I tool in GEODA. Moran's I is a measure of spatial autocorrelation developed by Patrick A.P. Moran. Moran's I is defined as:

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

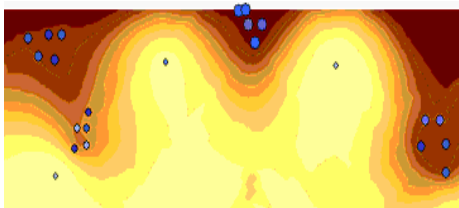
where: N is the number of spatial units indexed by i and j, x is the variable of interest,  $\bar{x}$  is the mean of x, and  $w_{ij}$  is a matrix of spatial weights.

Negative (positive) values indicate negative (positive) spatial autocorrelation. Values range from -1 (indicating perfect dispersion) to +1 (perfect correlation). A zero values indicate a random spatial pattern. For statistical hypothesis testing, Moran's I values can be transformed to Z-scores in which values greater than 1.96 or smaller than -1.96 indicate spatial autocorrelation that is significant at the 5% level.

## 6 Results

The results of cow endemic fluorosis density spatial cluster analysis showed that the density of dairy cattle endemic fluorosis distribution in space with a high concentration of the Z Score = 5.5,  $P < 0.01$ , which means that there is a spatial clustering model. The results of Cow endemic fluorosis density spatial Autocorrelation analysis showed Moran's I = 0.4216, which means there is an intensely spatial autocorrelation. Based on the spatial characteristic of the density of cow endemic fluorosis we established a prediction model to estimate the endemic

fluorosis distribution by Ordinary Kriging tool in ArcGIS. Superimposing map of predict map and Distribution of cow endemic fluorosis epidemic focus in district A (Fig1). Comparing with the true distribution, the prediction model has a well coincidence and is feasible to the application.



**Fig. 1.** Superimposing map of predict and the true distribution of cow endemic fluorosis

## 7 Conclusion

Applied Spatial Statistics used in conjunction with geographic information systems (GIS) provide an efficient tool for the surveillance of diseases. In the veterinary epidemiology, the advantage of mapping the locations of farms and other facilities with animals is obvious. In an outbreak of a disease it could make the management of the situation easier, and it could also provide a tool to evaluate different strategies to prevent the spread of infectious diseases. GIS have tremendously enhanced ecological epizootiology, the study of diseases in relation to their ecosystems. They have found increasing application for surveillance and monitoring studies, identification and location of environmental risk factors as well as disease prediction, disease policy planning, prevention and control. In this article we discuss the application of GIS to veterinary and medical research and the monitoring method of cow endemic fluorosis based on GIS and spatial statistical analysis. The method using a GIS tool facilitates is implemented significantly in the cow endemic fluorosis monitoring and investigation, and the space statistics - related prediction model provides a fundamental use for other study on space-related animal diseases.

## Acknowledgement

Funding for this research was in part provided by china postdoctoral science foundation (NO.20090461189), the postdoctoral fund of Shenyang Agricultural University, Research fund for young teachers of Shenyang Agricultural University, Dr. Start Fund of Liaoning Province, P. R. China. The authors are grateful to the Shenyang Agricultural University for providing conditions with finishing this research.

## References

1. Wartenberg D.: Screening for lead exposure using a geographic information system. *J. Environ Res Dec.* 59, 310--317 (1992)
2. Glass GE, Schwartz BS, Morgan JM III, Johnson DT, Noy PM, Israel E. :Environmental risk factors for Lyme disease identified with geographic information systems. *Am J. Public Health.* 85, 944--948 (1995)
3. Beck LR, Rodrigues MH, Dister SW, Rodrigues AD, Rejmankova E, Ulloa A, et al. :Remote sensing as a landscape epidemiologic tool to identify villages at high risk for malaria transmission. *Am J. Trop Med Hyg.* 51, 271--280 (1994)
4. Jubb TF, Annand TE, Main DC, Murphy GM.: Phosphorus supplements and fluorosis in cattle--a northern Australian experience. *J. Aust Vet.* 70, 379--83 (1993)
5. Richards FO, Jr.: Use of geographic information systems in control programs for onchocerciasis in Guatemala. *J. Bull Pan Am Health Organ.* 27, 52--55 (1993)
6. Clarke KC, Osleeb JR, Sherry JM, Meert JP, Larsson RW. : The use of remote sensing and geographic information systems in UNICEF's dracunculiasis (Guinea worm) eradication effort. *J. Prev Vet Med.*11, 229--235 (1991)
7. Braddock M, Lapidus G, Cromley E, Cromley R, Burke G, Branco L. : Using a geographic information system to understand child pedestrian injury. *Am J. Public Health.* 84, 1158--1161 (1994)
8. Queiroz, JW. Dias, GH, Nobre, ML, et al.: Geographic Information Systems and Applied Spatial Statistics Are Efficient Tools to Study Hansen's Disease (Leprosy) and to Determine Areas of Greater Risk of Disease. *J. American journal of tropical medicine and hygiene.* 82, 306-314 (2010)
9. Barnes S, Peck A. :Mapping the future of health care: GIS applications in Health care analysis. *J. Geographic Information systems.* 4, 31--33(1994)
10. Wartenberg D, Greenberg M, Lathrop R. : Identification and characterization of populations living near high-voltage transmission lines: a pilot study. *J. Environ Health Perspect.* 101, 626--632 (1993)
11. Busgeeth, K.: The use of a spatial information system in the management of HIV/AIDS in South Africa, *J. International Journal of Health Geographics.* 3, 13--15 (2004)