Universitat Autònoma de Barcelona
Departament de Sanitat i d'Anatomia Animals

CReSA
Centre de Recerca en Sanitat Animal

# Spatial analysis of Aujeszky's disease eradication in Catalonia, Spain

---

**PhD Thesis**
**Alberto Allepuz**
**2008**

---

Jordi Casal i Fàbrega, catedrático del Departament de Sanitat i d'Anatomia Animals de la Facultat de Veterinària de la Universitat Autònoma de Barcelona y Marc Saez Zafra, catedrático del Departament d'Economia de la Universitat de Girona,

HACEN CONSTAR,

Que el trabajo "Spatial analysis of Aujeszky`s disease eradication in Catalonia, Spain", presentado por Alberto Allepuz para la obtención del grado de Doctor en Medicina y Sanidad Animal, ha sido realizado en el Centre de Recerca en Sanitat Animal (CReSA) bajo su dirección.

Para que conste, firman la presente,

Jordi Casal i Fàbrega

Marc Saez Zafra

Bellaterra, Junio 2008

# AGRADECIMIENTOS

Pienso que esta tesis se la tengo que agradecer en primer lugar (y así de paso hago un poco la pelotilla), a Jordi, por la flexibilidad que tuvo en su momento de cambiar de tema de tesis y lanzarnos (no sin algunas reticencias) al mundillo de la epidemiología espacial. Agradecerle también su constante insistencia y sus continuas frases al entrar al despacho del tipo: "¿cómo va Alberto?, ¿has acabado ya los resultados?, ¡¡te quedan dos semanas!!,….", que aunque me estresaran, sin eso, esta tesis no estaría ahora mismo terminada. Desde luego, también le tengo que agradecer a Marc Saez, de la Universidad de Girona, mi otro director de tesis, que me escuchara y mostrara tanto interés y entusiasmo cuando le propuse que nos ayudara con este trabajo, y que no nos abandonara cuando las cosas no salían como esperábamos que salieran. Sin su ayuda y asesoramiento, tampoco habría sido posible hacer este trabajo. Y evidentemente, tengo que agradecer, y mucho, a todas aquellas personas que han y que están trabajado en los programas de erradicación de Aujeszky. A todos aquellos que van a las granjas, que meten los datos y que nos facilitaron toda la información para que pudiéramos hacer este trabajo: al Departament de Agricultura i Ramaderia (DAR), a l'associació tecnicosanitària del porcí (ASSAPORC), al grup de sanejament de Lleida (GSP Lleida) y al grup de sanejament de Girona (GSP Girona).

También quiero dar las gracias a todas aquellas personas, que de una manera u otra, me han facilitado las cosas y ayudado con el trabajo. A Ruben (el "chico" de los ordenadores) por sus consejos y ayuda a la hora de trabajar con las bases de datos, descubrirme cosas que no sabía ni que se podían hacer y hasta escucharme cuando le contaba que es lo que estaba haciendo. A mis compañeros de epidemio; Anna, Sebas y recientemente Nacho y Meritxell, por ser más amigos que simples compañeros de trabajo, aguantar que usara "de vez en cuando" sus ordenadores y leer los ladrillos de la tesis que les he ido pasando. A Joan Pujols, por escucharme cuando le contaba algún resultado y darle vueltas siempre en como aquello se podría aplicar a la realidad. A toda la gente de infecciosas, por seguir contando conmigo, a pesar de no haber pisado demasiado el departamento estos últimos años y por preocuparse por mí y preguntarme tantas veces que como lo llevaba. A mi hermana, por hacer en un momento diferentes pruebas para la ilustración de la portada y a mi hermano por dejarse robar un "par" de libros (que todavía no le he devuelto) y como no, a mi madre, por ser mi madre, y por

estar tan orgullosa de los logros de sus hijos. A Nuria, por perdonarme el tiempo que no le he dado, apoyarme y aguantarme más que un par de días. Y desde luego, a todas las personas del CReSA, la facultad y también de fuera del CReSA y la facultad, que aunque no estén en esta hoja con su nombre, han hecho que me encuentre tan a gusto aquí donde estoy.

**- INDEX**

# 1. INTRODUCTION

## 1.1. Aujeszky's disease

Aujeszky's disease (AD) also known as pseudorabies, is caused by the pseudorabies virus which belongs to the *Alphaherpesvirinae* subfamily of the *Herpesviridae* (Mettenleiter, 2000). It was first described in 1813 in cattle suffering with extreme pruritus. The term "pseudorabies" was first used in Switzerland in 1849 because the clinical signs in cattle were considered similar to those of rabies. In 1902 Aladar Aujeszky, the Hungarian for whom the disease is named, determined that the etiologic agent was filterable, i.e., not bacterial and also conducted research on the disease in the dog and cat. The agent was first recovered from swine in 1909 by Weiss (Pejsak and Truszcynski, 2006).

Aujeszky's disease virus (ADV) is composed of a nucleoprotein core which contains the genome- a DNA single chain with 145 kb- , an icosahedral capsid of 162 capsomers, a proteinaceous tegument, and a lipid bilayer envelope derived from cellular membranes which contains virally encoded (glyco) proteins (Mettenleiter, 2000).

ADV displays a very broad host range and is able to infect most mammals except higher primates, including humans (Schmidt et al., 2001). ADV is highly virulent for most of the species: in cattle, sheep and goats produces a uniformly fatal infection characterized by clinical pruritus and encephalitis (Pejsak and Truszcynski, 2006). The pig is the only host which is able to survive a productive infection and serve as a virus reservoir (Mettenleiter, 2000).

Following oronasal exposure, primary replication occurs in the epithelia of the upper respiratory tract (Pejsak and Truszcynski, 2006). After a first round of replication in the epithelia, progeny virus is abundantly produced (Mettenleiter, 2000) and thereafter, infection of tonsils and lungs ensues and the virus is disseminated through the body either in free form or via infected leukocytes. In addition, the virus enters the trigeminal and olfactory nerve endings and invades the central nervous system (Pejsak and Truszcynski, 2006). Surviving pigs can remain latently infected and productive viral replication can be reactivated in response to various stimuli (Schmidt et al., 2001). Major sites of ADV latency are the trigeminal ganglion, the olfactory bulb and the tonsil. In these organs, viral DNA can be detected in the absence of infectious virus production (Mettenleiter, 2000).

The severity of the clinical outcome resulting from the infection is determined by the age and immunological status of the animal as well as the virulence and dose of exposure to the virus (Zuckermann, 2000). Infections can result in neurological signs in piglets, respiratory tract illness in older pigs and abortion in pregnant sows (Pejsak and Truszcynski, 2006). Some strains of ADV are associated with reproductive failure and others with respiratory disease, but all of them are able to cause neurological disorders in young pigs. Nervous disorders are very occasionally seen in fattening pigs (Segalés et al., 2003). Young piglets are highly susceptible, with mortality rates reaching 100% during the first 2 weeks and decreasing to 50% during the third and fourth week of age. In grower-finishing pigs, respiratory signs are the most common observations. Morbidity in an infected group often reaches 100%, but in cases uncomplicated by secondary infections, mortality ranges from 1-2%. Complications caused by bacteria, particularly *A. pleuropneumoniae,* increase the losses dramatically. In pregnant females, abortion often occurs and may be the first sign of AD. Sows infected in the first trimester may resorb fetuses and return to estrus. If infection takes place in the second or third trimester of pregnancy, it is usually manifested by abortion or stillborn and/or

weak pigs, particularly if the infection occurs close to term (Pejsak and Truszcynski, 2006).
The diagnosis of AD can be confirmed by isolation of ADV from the oropahryngeal fluid, nasal fluid (swabs) or tonsil biopsies of living pigs. For post-mortem isolation samples of brain and tonsils are the preferred specimens. In addition, the virus can sometimes be found in other tissues including the lung, spleen, liver, kidney, lymph nodes and pharyngeal mucosa. In latently infected animals, the most consistent site for virus isolation is the trigeminal ganglion in domestic pigs (Pejsak and Truszcynski, 2006) and the sacral ganglia in wild boars (Romero et al., 2003).
Isolation can be made by inoculating a tissue homogenate into a susceptible cell line; porcine kidney (PK-15) cells are most often used. Specificity of the cytopathic effect can be verified by immunofluorescence, immunoperoxidase or neutralisation with specific antiserum (OIE, 2004).
The polymerase chain reaction (PCR) can be used to identify ADV genomes in secretions or organ samples, being the primary advantage its speed. Immnunofluorescence (IF) is still of some importance for detection of viral antigen in tissue sections and nasal swabs (Pejsak and Truszcynski, 2006). Aujeszky's disease antibodies are demonstrated by virus neutralisation or enzyme-linked immunosorbent assay (ELISA). Virus neutralisation (VN) has been recognised as the reference method for serology, but for general diagnostic purposes it has been widely replaced by the enzyme-linked immunosorbent assay (ELISA) because of its suitability for large-scale testing and the fact that ELISAs can distinguish vaccinated and infected pigs, if gene-deleted vaccines are used. The tests can be performed on meat juice as well as serum (OIE, 2004).

- **Epidemiology:**

Pigs are the only natural host of ADV. Other animals found on farms such as rats and mice are dead-end hosts (Mettenleiter, 2000). The primary means of transmission of the virus among swine farms is thought to be by direct contact between infected and susceptible pigs because of movement of infected animals (Austin and Weigel 1992). Indirect contacts by vehicles, equipment, personnel, feed or artificial insemination can also spread the virus among farms (Tamba et al., 2002, Solymosi et al., 2004). Moreover, the virus may be transmitted among swine farms by aerosol suspensions of virus (Christensen et al., 1990, 1993; Casal et al., 1997).

Transmission of ADV within farm usually occurs by the oropharyngeal route after direct contact of infected and susceptible animals and the ingestion or aspiration of infected aerosols, secretions and excretions (Romero et al., 2003).
During acute infections, the virus is present in the tonsil epithelium, milk, urine and vaginal and preputial secretions (Anonymous 2007). The infectious period is up to 16 days (Bouma et al., 1995) as antibodies against gE appear approximately 14 days after infection (Van Oirschot et al., 1991).
In the environment, the virus can remain infectious for as long as seven hours in the air, if the relative humidity is at least 55%. It is inactivated by sunlight, drying and high temperatures (Anonymous, 2007a). In slurry, ADV is inactivated in 2-4 minutes at 62ºC. The levels of virus likely to be present in slurry following AD outbreaks are assumed to be low (Turner et al., 2000).

- **Wildboars:**

Within the European Union (EU), all disease programs have been directed towards the control of AD in domestic pigs. However, as this goal is approached and an AD-free status has to be maintained, a wildlife reservoir becomes increasingly important. On the basis of EU regulations, there is no need to monitor wild boar for the occurrence of economically important diseases, with the exception of classical swine fever. Thus, in Europe, serological surveys in wild boar mainly depend on the current disease situation, public interest and funding (Müller et al., 2003). AD has been reported in wild boars in different countries: Germany (Müller et al., 1998), France (Albina et al., 2000), Croacia (Zupancic et al., 2002), Spain (Vicente et al, 2005), Slovenia (Vengust et al, 2006) and Switzerland (Leuenberger et al., 2007).
Nevertheless, in Germany Müller et al., (1998) did not found evidence of disease transmission to domestic pigs despite the presence of an endemic focus in wild boars. Also, in France AD infection is enzootic in wild boars in areas where it has been eradicated in domestic pigs (Albina et al., 2000) and in a study carried out in south-central Spain it was concluded that there was no evidence of interaction between the epidemiology of ADV in domestic pig and the wild boar (Ruiz-Fons, 2006). The infectious AD cycle within the wild boar population appears to be independent of that in domestic pigs (Westergaard, 2000). However, domestic pigs kept outdoors may be one exception to the rule since they may have increased chances of direct contacts with infected wild boars (Albina et al., 2000). Hars and Rossi (2005) suggested that the wild boar was responsible of the appearance of AD outbreaks in open-air domestic pig herds in the French department of Loire. AD in wild boar populations is preferably transmitted by the venereal route (Romero et al., 2001).

- **AD distribution:**

AD can be found in parts of Europe, Southeast Asia and Central and South America (Anonymous 2007). Canada, New Zealand and the United States are free of AD since 2004 (Pejsak and Truszcynski, 2006). In Europe, AD has never been reported in Norway, Finland or Malta. In the European Union, by April 2008, Finland, Austria, Cyprus, Czech Republic, Germany, Denmark, Luxembourg, Slovakia, England, Scotland, Wales, Sweden and France are free of AD and vaccination is prohibited. There is an approved eradication programme in Belgium, Netherlands, Bolzano province in Italy and some provinces of Spain. In Figure 1, the European Union situation, as included in European decision 2008/269/CE is represented (Anonymous, 2008).

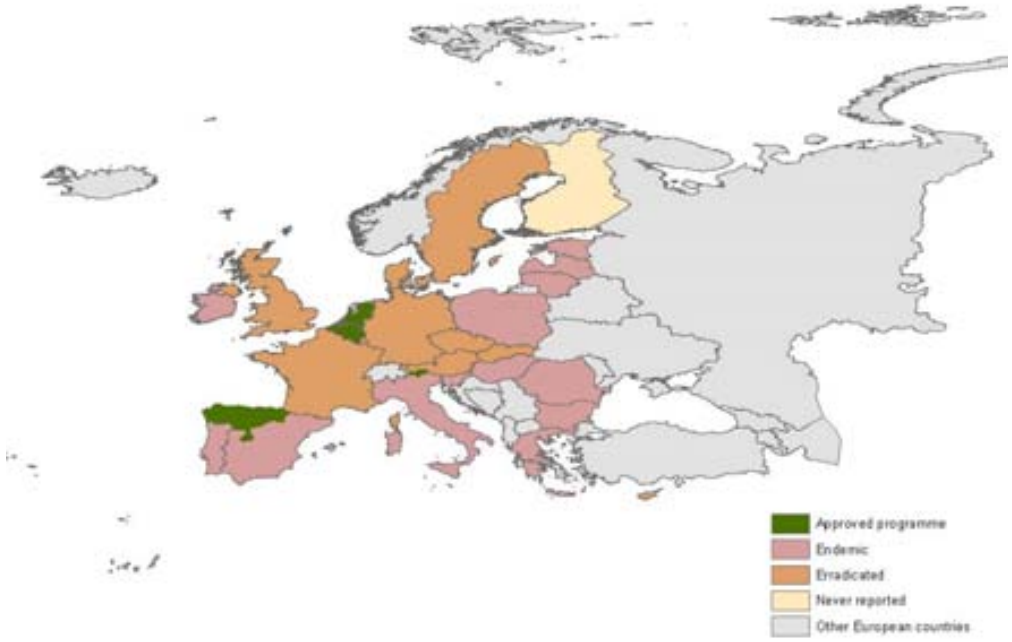


Figure 1. European Union situation as included in 2008/269/CE (source: own production)

In **Spain**, the eradication is in process. Each year, the Spanish government publishes the farm prevalence by county. A farm is defined as positive if at least one positive animal is confirmed by the official gE enzyme-linked immunosorbent assay method. The last published data relates to the end of 2007 serological results. At that point, in 348 counties there were no positive farms, in 103 less than 10% were positive and in 16 counties between 10 and 58% of the farms were positive. The geographical distribution of these counties has been represented in a map by the agricultural ministry (figure 2).

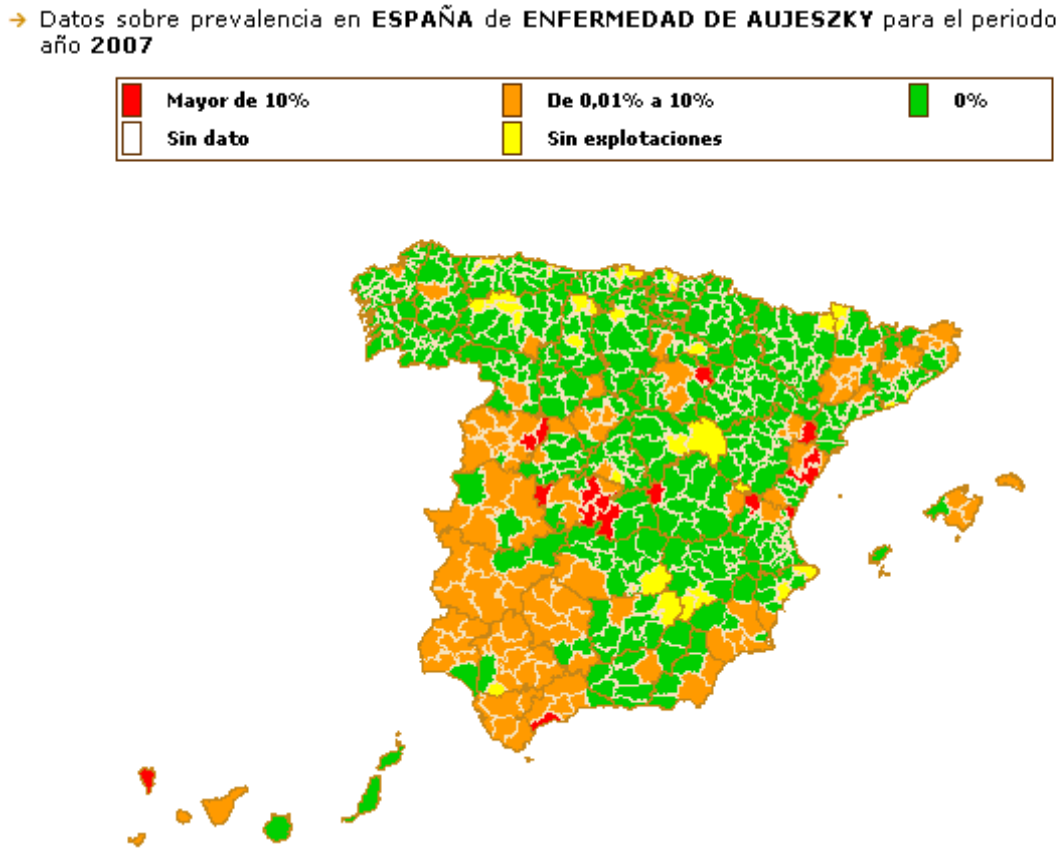


Figure 2. Serological prevalence at the end of 2007 (from Ministery of Agriculture, fisheries and food ; MAPA).

- **Control measures and legislation:**

The framework of animal health problems affecting intra-Community trade in bovine and swine animals is defined in the Council Directive 64/432/EEC of 26 June 1964 (Anonymous, 1964). It defined health requirements to be met if live pigs were to be moved between member states and established the legal basis for AD control in the EU (Müller et al., 2003).
In 1990, this basic trade Directive was amended in such a way that measures relevant to the control and eradication of AD could be made by Commission Decisions. In 1993, the European Commission adopted two major Decisions concerning protective measures in relation to AD and trade in live pigs. Commission Decision 93/24/EEC (Anonymous, 1993a) set out the health conditions which apply to pigs moving into areas where the disease is absent, while Commission Decision 93/244/EEC (Anonymous, 1993b) defined health requirements applying to pigs destined to member states where an approved eradication programme was in place. The two Decisions divided the territory of the EU into areas where the disease was absent; areas where approved eradication programs were being implemented and areas where AD was either not under official control or where control/eradication programs did not yet meet the criteria for approval by the Commission (Westergaard, 2000). In July 2001, both Commission Decisions were amended and merged into the Commission decision

2001/618/EC (Anonymous, 2001) which provided with additional guarantees in intra-community trade of pigs relating to Aujeszky's disease.
In the last decades, big efforts have been made to eradicate AD. To reach this goal, the development and approval of a glycoprotein-E (gE) deleted vaccine was the decisive break-through, as this vaccine made it possible for the first time to distinguish infected from vaccinated pigs (Van Oirschot et al., 1986).This led to large application of gE-deleted vaccines and to new control strategies.
In Spain, a decree (RD) published in 1995 (Anonymous, 1995) started the AD eradication programme. This was based on the forbidding of the use of gE positive vaccines, the establishment of a compulsory vaccination program and animal movements restrictions. The vaccination program consisted in vaccination of breeding sows three times per year, fatteners at least once between 10 and 12 weeks of age and two vaccinations of gilts before entering the reproductive cycle. Movement of animals was allowed between farms of similar AD status and a serological control of farms was implemented. In 2003, due to the additional guarantees in intra-community trade of pigs relating to Aujeszky's disease included in Commision Decision 2001/618/CE, the AD eradication programme was adapted and reinforced by the 427/2003 decree (RD) (Anonymous, 2003a). In this new programme, vaccination of breeding sows was compulsory also three times per year, but it had to be done at regular intervals. Fatteners had to be vaccinated at least two times; the first application between 10 and 12 weeks of age and the second one three or four weeks after. Gilts had to be vaccinated three times before entering the reproductive cycle. Vaccination was not compulsory in farms classified as officially free or in process to obtain this status. Annual serological testing of farrow to weaning and farrow to finish farms was established, the sample size was calculated based on a detection of a 5% of prevalence with a 95% of confidence. Also, all gilts of auto-replacement were to be tested in order to ensure that only gE-negative gilts were used as reproductors. On the other hand, movements of animals were allowed only between farms or areas with similar AD status.
In 2005 (206/2005 decree (RD)), a new modification in the AD eradication programme was implemented. Farms were classified according to the following classification:

- A0: Farms with unknown Aujeszky's diseases status.
- A1: Farms were the vaccination protocol was known and with positive results in the last serological control.
- A2: Farms were the vaccination protocol was known and with negative results in the last serological control.
- A3: Farms classified as free.
- A4: Farms classified as officially free.

It was established the figure of veterinary responsible of the implementation of AD eradication programme in each farm. Compulsory annual serological testing of farrow to weaning and farrow to finish farms was continued and counties began to be classified based on the proportion of positive sow farms in the county:

- County with high prevalence: proportion of positive sow farms higher than 10%.
- County with low prevalence: proportion of positive sow farms equal or lower than 10%.
- County free of AD: no positive sow farms

A compulsory serological control of fattening farms was also established; the number of samples to be tested was calculated based on a detection of a 20% of prevalence with a 95% of confidence. Movement of animals was allowed between counties with the same AD classification (Anonymous, 2005a). Finally, in 2006, the programme was reinforced again (636/2006 decree (RD)); the main changes were related to the process to be classified as AD free: necessary two consecutive negative serological results and no clinical signs in the last 12 months, serological control of fattening farms was based on risk assessment results (Anonymous, 2006a).

In Catalonia region, the adaptation of these programmes were done and completed by the design of specific eradication programmes for the region (Anonymous, 2003b, Anonymous, 2006b). The eradication of AD was design to be achieved through different phases of 1 year of duration each one:

***Phase 1*** (April 2003 – May 2004): National measures were complemented by serological testing of fatteners in purely fattening units and in fatteners of farrow to finish farms and the application of movement restrictions on farms with prevalence higher than 20%.
***Phase 2.1*** (June 2004 – May 2005): During this period movement restrictions were applied on those farms with prevalence higher than 10%. More efforts were applied to ensure those points included in the previous phase and it was initiated the process to qualify farms or areas as AD free.
***Phase 2.2*** (June 2005 – May 2006): In this period, efforts were made in order to qualify farms and areas as AD free by a voluntary programme. Movement restrictions were applied on those farms with prevalence higher than 0%. Also movement restrictions were applied based not only on a farm basis, also based on area AD status.
***Phase 3*** (June 2006 – May 2007): Qualification of farms and areas by a compulsory programme and movement restrictions on those farms with prevalence higher than 0%. Also, in this period, a compulsory programme of slaughter of positive animals was established.

## 1.2. Spatial epidemiology

The term spatial epidemiology is used to define studies where the geographical location of the events is a fundamental component (Saez and Saurina, 2007). It is consider a subdiscipline of epidemiology whose primary purpose is to describe and explain the spatial pattern of disease (Durr and Gatrell, 2004). Knowledge of where and when a disease occurs enables the generation of disease causation hypothesis for diseases with unknown or poorly characterised aetiology, identification of disease risk factors and the design of efficient disease surveillance and control programmes in animal health (Ward, 2007).

Up to the 1980s, it is difficult to find examples in the veterinary literature; the exceptions are works undertaken by parasitologists, interested in the interaction between climate and disease via its effect on vectors and intermediate hosts. One of the first works was conducted by Robson et al., (1961) who showed that the East Coast fever in Tanzania was confined to areas where tsetse flies were absent and cattle was present. Also in Tanzania (Lake Victoria) Yeoman (1966) by carefully mapping disease outbreaks in relation to the cattle population, draw a line separating enzootic and epizootic areas and map the spatial development of the disease. In human medicine, there are studies dating from the beginning of the 1800s in which maps were employed to demonstrate the distribution of disease (Lawson and Williams, 2001). Possibly the most famous use of mapping in epidemiology in this period were the studies by John Snow of the cholera epidemics in London in 1854 (Figure 3). Through observation of the adresses of the people who die. Snow was among the first to show clearly that cholera could be spread through a contaminated water supply (Lawson and Williams, 2001).


Figure 3. John Snow map of cholera deaths and water supply in London.

Nevertheless, it was not until the 1980s, thanks to the technical breakthroughs in computing, that spatial epidemiology was really developed. Advances in geographic information systems (GIS), statistical epidemiology and availability of high-resolution, geographically referenced health and environment data, created new opportunities to investigate geographic variations in disease occurrence (Elliot and Wartenberg, 2004).
GIS, spatial analysis and remote sensing are the main tools employed in spatial epidemiology (Durr and Gatrell, 2004). GIS are powerful tools for displaying and querying spatial information and in the last years they have been build more user-friendly and powerful software packages. The fundamental ingredient of these packages is the use of map layers which contain different information about the mapped area. Each layer can be manipulated interactively (edited) to provide a composite map (Lawson and Williams, 2001). Spatial analysis techniques have been developed in parallel, but largely independently. As a result, modern GIS software still has fairly limited spatial analysis functionality (Pfeiffer, 2004). Spatial analysis deals with the exploration, description and analysis of data taking into account their geographical distribution (Saez and Saurina, 2007). Spatial data are defined as geographical features and the attributes of these features, each feature will often have multiple attributes (Pfeiffer and Jones, 2002). There are different types of spatial data:

- ***Lattice data:*** observations are aggregated at specific areas. For example number of cases of a specific disease by parish.
- ***Point processes:*** the data set may consist of locations only, or it may be a *marked* point process, with data values associated with each location (marks).
- ***Geostatistical data:*** In this case measures are done at specific locations but the observation of interest is continuous. For example, measures of rainfall or temperature.

To manage spatial data, a GIS requires both spatial and non-spatial database management functionality. The geographical features are managed by the spatial data functions, which also maintain links to the attribute data. The latter are often stored in a standard database management system (Pfeiffer and Jones, 2002). The spatial analysis techniques to be applied will be dependent on the type of spatial data that we have and the objective of the study (Saez and Saurina, 2007). In general, the analysis of spatial data consists of 3 components: ***exploratory analysis*** (finding interesting patterns), ***visualization*** (showing interesting patterns) and ***spatial modelling*** (explaining interesting patterns).

## 1.2.1. Exploratory Analysis

Exploratory analysis involves statistical examination of the data for the presence of any patterns. Such analysis allows determining geographical trends related with the disease, but can not be used to test causal hypothesis (Pfeiffer and Jones, 2002).

***1- Disease clustering:***

**a) Focused**-cluster tests are used when there is a pre-specified point source for a presumed epidemic, and the question of interest is whether there is an elevated risk of the disease around that specific source. These tests do not appear to have been extensively used in veterinary epidemiology (Ward and Carpenter, 2000).

**b)** Most proposed tests for spatial clustering are tests for **global clustering** (Kulldorf, 2006). These methods test for clustering throughout the study region without the ability to locate the sites of specific clusters (Ward and Carpenter, 2000). The data are typically measured on a **categorical** or **continuous** scale. If the data can be treated as **continuous**, the presence of spatial autocorrelation can be assessed using *Moran´s I* statistic or a variogram. Spatial autocorrelation is the tendency for nearby spatial units to record similar values (Durr and Gatrell, 2004). The ***Moran´s I*** coefficient is essentially a modification of the ordinary (Pearson) correlation coefficient but with an added term which measures spatial proximity. However, what is mean by proximity has to be defined (Durr and Gatrell, 2004) and also weights to these values must be given (Saez and Saurina, 2007). With **lattice** spatial data, one common definition for spatial proximity is that the area units must have a common boundary. Alternatively, if the distance between the centroids is measured, proximity can be defined in terms of a threshold distance (Durr and Gatrell, 2004). The weights for the neighbourhood values can be specified as **binary** (1/0) depending if the neighbour is adjacent or inside the threshold distance or as a function of the distance between pairs of points (Saez and Saurina, 2007). The ***Moran´s I*** coefficient, provides a one-number overall measure of spatial autocorrelation with a numerator that measures the extent to which adjacent points (as described by the elements of 'w') have similar deviations about the mean of the data, and a denominator that standardizes that quantity to reflect the scale or variability of the variable being examined. This statistic is given by:

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

Where 'n' *is* is the number of pair of points to be compared, 'x' is the variable of interest, '$\overline{x}$' is the mean of 'x' and 'w' is the neighbourhood weigth matrix describing the adjacency or distance between the i-th and j-th point. The statistical significance is calculated by a z-statistic, using the standard deviation of Moran (Ward and Carpenter, 2000).

The spatial dependence over various distances can be represented by a ***variogram*** (Gatrell, 2004). These are presented as graphs with distance (spatial lag) on the x-axis and variation (semivariance) on the y-axis. In spatial statistics the semivariance is described by:

$$\gamma(h) = \sum_{i=1}^{n(h)} \frac{(z(x_i + h) - z(x_i))^2}{n(h)}$$

Where '*z*' is a datum at a particular location, '*h*' is the distance between data, and '*n(h)*' is the number of paired data at a distance of '*h*'.
Semivariance is the converse of autocorrelation, in that it is low in the presence of local spatial effects and increases to a maximum where there is no longer spatial dependence (Durr and Gatrell, 2004). A variogram curve with a flat shape suggests the absence of

spatial dependence. A curve with an exponential shape, expressing increasing variability between pairs of locations with distance, reflects the presence of spatial dependence (Pfeiffer, 2004).
For data measured as categorical, **nearest-neighbour test** (Clark and Evans, 1954), **Cuzick-Edwards´s test (Cuzick and Edwards, 1990)** and **K-function (Diggle et al., 1995)** among others, have been used to detect case clustering.
Results of the **nearest-neighbour test** are subject to confounding because it does not take into account the underlying population at risk. If the population at risk is clustered, then disease cases arising from that population are also expected to be clustered (Ward, 2007). The index of the nearest neighbour test is the ratio of mean Euclidean distance between nearest neighbour points in a given area ($D_{observed}$) to mean distance expected from a randomly distributed series of points in that area ($D_{random}$), where:

$$D_{random} = 0.5 \sqrt{(A/N)} \text{ and } D_{observed} = \sum d_{ij} / N .$$

'N' is the number of points, 'A' is the study area size and '$d_{ij}$' is the distance between nearest-neighbour points '*i*' and '*j*'. The statistical significance is determined by calculating a z-statistic (Ward and Carpenter, 2000).
To take into account the underlying population, and avoid thus confounding, the **Cuzick-Edwards´s** test and the **K-function** tests adjust for the presence of heterogeneously distributed population at risk. This is accomplished by selecting appropriate controls for the cases. Controls are drawn from the same underlying population as cases, thereby accounting for clustering that may occur in the population regardless of the clustering of cases (Ward, 2007). Under the null hypothesis, a similar cluster would be expected for the cases and the controls, if no disease-association was present (Carpenter, 2001).In **Cuzick-Edwards´s test**, the statistic is derived by first counting, for a given case, the number of nearest neighbours that are also cases. This is then summed over all cases to yield the global statistic. The statistical significance is assessed by comparing its value with its expected value under the null hypothesis of random labelling of case and control locations (Rogerson, 2006).
The **K-function**, is calculated as the number of events within a distance 's'. For clustered patterns, each event is likely to be surrounded by further members of the same cluster, therefore, for smaller 's' values the K-function will be relatively large. Conversely, if events are regularly spaced, each event is likely to be surrounded empty space and, for small values of 's', the K-funtion will be small. To take into account the heterogeneous spatial distribution of the underlying population at risk, separate K-functions for cases ($K_{case}$) and controls ($K_{controls}$) are computed. The observed difference function:

$$D = K_{case} - K_{controls}$$

is interpreted as a measure of the extra aggregation of case events over and above that observed for controls (Stevenson, 2003).

**c) Local**-cluster tests: The spatial **scan statistic** (Kulldorf, 1997) is a cluster detection test able to locate the site and to test the significance of specific clusters. It is a recommended test for the detection of local clusters (Song and Kulldorf, 2003). The test can be used for spatially aggregated data as well as when exact geographic coordinates are known for each individual. Therefore it can be used for lattice or point spatial data.

It can also be used with categorical data (Bernoulli model) as well as for continuous data (Poisson model).
With the Bernoulli model there are cases and non-cases represented by a 0/1 variable and with the Poisson model, the number of cases in each location is assumed to be Poisson-distributed. Under the null hypothesis, the expected number of cases in each area is proportional to its population size (Kulldorf, 2006). The SaTScan software uses this test, it searches for clusters by using a variable circular window size to detect spatial clusters in large areas, while controlling for the underlying population (Kulldorf, 1997). The circle is centred on each of the points. For each point, the size of the circle varies continuously from zero to some upper limit specified by the user. Each circle is a possible candidate cluster. For each circle a likelihood ratio statistic is computed, based on the number of observed and expected cases within and outside the circle and compared with the likelihood under the null hypothesis (Aamodt et al., 2006). The likelihood function under the Poisson assumption is proportional to:

$$\left(\frac{c}{E[c]}\right)^{c}\left(\frac{C-c}{C-E[c]}\right)^{C-c} I0$$

where C is the total number of cases, c is the observed number of cases within the window and E[c] is the expected number of cases within the window under the null-hypothesis. I() is an indicator function. When SaTScan is set to scan only for clusters with high rates, I() is equal to 1 when the window has more cases than expected under the null-hypothesis, and 0 otherwise (Kulldorf, 2006).
For the Bernoulli model the likelihood function is:

$$\left(\frac{c}{n}\right)^{c}\left(\frac{n-c}{n}\right)^{n-c}\left(\frac{C-c}{N-n}\right)^{C-c}\left(\frac{(N-n)-(C-c)}{N-n}\right)^{(N-n)-(C-c)} I0$$

where c and C are defined as above, n is the total number of cases and controls within the window, while N is the combined total number of cases and controls in the data set.

The likelihood function is maximized over all window locations and sizes, and the one with the maximum likelihood constitutes the most likely cluster. This is the cluster that is least likely to have occurred by chance. The likelihood ratio for this window constitutes the maximum likelihood ratio test statistic (Kulldorf, 2006). Its distribution under the null-hypothesis is obtained by repeating the same analytic exercise on a large number of randomly selected replications of the data set generated under the null-hypothesis (Ward and Carpenter, 2000). The p-value is obtained through Monte Carlo hypothesis testing by comparing the rank of the maximum likelihood from the real data set with the maximum likelihood from the random data sets (Kulldorf, 2006).

***2 - Disease mapping:***

Disease mapping studies aim to summarize spatial variation in disease risk, in order to assess and quantify the amount of true spatial heterogeneity and the associated patterns, to highlight areas of elevated or lowered risk and to obtain clues as to the disease aetiology (Best et al., 2005).

The detection of disease clusters has typically been approached as a hypothesis testing problem; whether the geographical distribution of disease is random or not, adjusting for the geographical distribution of the population (Ugarte et al., 2005). Disease mapping methods are most useful for capturing gradual regional changes in disease rates, and are less useful in detecting abrupt localized changes indicative of clustering (Gangnon and Clayton, 2000). As the objectives of these map presentations are to identify locations with unusually high or low disease levels, a common parameter represented is the ratio between the observed and expected cases (Elliot and Wartenberg, 2004). Traditionally, the ratio of observed to expected counts is called a Standardized Incidence Ratio (SIR) and this ratio is an estimate of relative risk within each area (the ratio describes the odds of being in the disease group rather that the background group) (Lawson et al., 2003).

Due to the counts of cases observed in a particular location ($O_i$) is a counter variable, it can be modelled by a Poisson distribution (likelihood function). The local variability of the counts is thus modelled as follows:

$O_i \sim \text{Poisson}(E_i\theta_i)$

The parameter of interest is $\theta_i$ , the relative risk (estimation of SIR) that quantifies whether the area 'i' has a higher ($\theta_i > 1$) or lower ($\theta_i < 1$) occurrence of cases than that expected from the reference rates (Richardson et al., 2004).

In order to avoid confounding, the SIR must be standardised by those known variables that are linked with the disease of study. The standardisation of this ratio can be done by the direct or the indirect method. In the direct standardisation the expected cases are calculated using the rate of a reference population while the indirect standardisation uses the rate calculated after aggregating the data from all regions included in the analysis. In human epidemiology, the most common is to standardise by age, in order to take into account the age structure between areas to be compared (Saez and Saurina 2007). In veterinary medicine the standardisation by breed has been used to compare bovine spongiform disease risk between areas (Abrial et al., 2005a; Stevenson et al., 2005a; Allepuz et al., 2007).

## 1.2.2. Data visualization

The results of the statistical procedures are represented visually in mapped form. Hence, some consideration must be given to the purely cartographic issues that affect the representation of geographical information (Lawson et al., 2003). The type of map presentation depends on the type of data available, either the actual event locations (such as the x-y coordinates) or aggregate data (Pfeiffer and Jones, 2002).

To visualise ***point data,*** the oldest and most frequently-used method is to plot the locations of the study subjects using their Cartesian coordinates. Whereas plots of point events provide a general impression of the spatial characteristics of the process under investigation, they present problems where there are multiple events at the same location since no indication of event density can be appreciated. Because of this, point maps are best suited for displaying location information for small numbers of events (Stevenson, 2003). If a continuous surface is to be mapped based on a discrete set of

observation points, then interpolation techniques, based on **geostatistical methods**, must be used (Lawson and Williams, 2001). Interpolation techniques enable the construction of *isopleth* maps. These maps show the distribution of spatially continuous phenomena by a logical sequence of tones colour that symbolizes equal values. Isolines are often overlaid on top of an isopleth map to indicate threshold value.

Choropleth maps are a common method for visualising ***lattice data.*** Areas within a region of interest are shaded according to a discrete scale based on the recorded value of the attribute of interest for that area. The hatching pattern or colour is based on a class interval or continuous scale derived from a descriptive statistic of the aggregated data, such as the prevalence of disease for example. When inspecting choropleth maps the following factors should be taken into consideration (Pfeiffer and Jones, 2002):

a) The boundaries are usually chosen for political or other reasons irrelevant to disease spread, although such boundaries may have a direct impact on the reporting of disease.

b) Large areas may visually dominate the map, even though their importance may be limited with respect to the corresponding population size (Berke, 2005). In the geography literature, this is called the modifiable area unit problem (Elliot and Wartenberg, 2004).

A range of methods have been developed to deal with these problems. The modifiable area unit problem can be solved by analysing the data on the basis of the smallest area units for which information is available. Moreover, the use of different polygonal boundary definitions is advised, if practical (Stevenson, 2003). Also, to overcome the *visual bias* and for a more natural representation of the continuous spatial variation in risk, *isopleth* maps can be used to visualize the underlying risk of disease using **geostatistical methods** (Berke, 2005).

There are a number of techniques that can be used to **interpolate** data by use of deterministic methods or to predict the values statistically at the grid-coordinates (Berke, 2004). Among the possible choices is inverse distance weighting (IDW), kernel smoothing and kriging.

**Inverse distance weighting** (IDW) is a method of interpolation that estimates values by averaging the values of data points in the neighborhood of each point weighted only by distance (d). The weights (w) are a decreasing function of distance. The closer a point is to the data point being estimated, the more influence, or weight; it has in the averaging process.

$$w(d) = \frac{1}{d^p}$$

The value of p is specified by the user, the most common choice is p= 2. The size of the neighborhood can be expressed as a radius or as a number of points. In the first case, the radius is constant independently of the density of farms, while in the second one, the size of the radius increases until achive a given number of farms. All the measured points that fall within the radius will be used in the calculation of each interpolated data point. The search radius will increase until it can encompass the minimum number of points.

***Kernel density estimation:*** This technique involves the placement of a regular grid over a region of interest and the construction of an area (known as the kernel) around each observed point ($x_1, x_2, \ldots, x_n$) to define a distance-decaying density estimate. The kernel density estimate for each grid cell equals the sum of the density estimates that fall within the cell. The intensity of a spatial point pattern $x$, $f(x_h)$, is given by:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - x(i)}{h}\right)$$

Where the term 'K' defines the shape of the kernel structure to be used (Gaussian, triangular, quartic,…) being the Gaussian the most-commonly used and the term 'h' defines the radius of the kernel (known as the bandwidth or smoothing parameter) (Stevenson, 2003). The larger the bandwidth, the smoother is the surface map; whereas when the selected bandwidth is too small the resulting density will resemble the original point pattern. Diverse methods have been proposed to adjust the window size appropriately (Berke, 2005) but, sensitivity analysis, by repeating the same analytical exercise with different bandwidths, is a practical way to choose the radius that better represents the data. With marked point processes where attribute variables are recorded at each location, the density of the attribute value is estimated by constructing a non-parametric kernel regression surface where the Cartesian coordinates provide the covariates of the regression and the attribute values provide the vector of responses. Where no attribute values are provided (for example, where interest lies in determining the density of farm holdings in an area) a kernel density surface may be constructed using the Cartesian coordinates of each farm holding (Stevenson, 2003).
Recently, Benschop et al., (2007) used a novel kernel estimator with a spatially adaptive smoothing bandwidth to visualise the spatial pattern of Salmonella risk in the Danish swine Salmonella control programme. This approach takes into account the population density among the region resulting in a more detailed analysis of the data.

***Kriging:*** is a method for spatial interpolation or prediction based on a set of observations at point locations by the use of a linear combination of surrounding sampled values. Unlike kernel density estimation, kriging is based on a parametric spatial model (Berke, 2005). The kriging predictor is a weighted average that is calculated from the entire sample with weights depending on the semi-variogram. The weights are constructed to give regional risk estimates more influence on the prediction the closer they are to the prediction sites and to downplay a cluster of points that contains largely redundant information (Berke, 2005). Within a probabilistic framework, kriging attempts to: (a) minimize the error variance; and (b) systematically set the mean of the prediction errors to zero, so that there are no over- or under-estimates (Anonymus, 2005b).
The problems with the kriging method for generating isopleth maps of disease occurrence are (1) heterogenous variances in the regional estimates and (2) the potential of negative interpolations. The first problem can be ameloriated by the use of smoothing techniques to the data prior to kriging. Appropiate geostatistical modelling can solve the second problem (Berke, 2004).

## 1.2.3. Modelling

Once SIR has been standardised by confounding variables, if any, the model can be extended to include specific covariates (Stevenson, 2003):

$$\log(\mu_i)=\log(E_i) + \alpha + \beta_1 x_{1i} + ... + \beta_m x_{mi}$$

Where $\mu_i = E_i\theta_i$

The use of the SIR for disease mapping, assumes that once all known and observable confounding variables are included, the resulting map will be clean of all artefacts and hence depicts the true excess risk surface. However, unknown or unmeasured risk factors may lead to extra variation in the observed counts (greater variation than that expected from a Poisson distribution) and also they will vary in space, which in turn induces spatial correlation between the observed disease counts in nearby units (Stevenson, 2003). These two sources of extra variation are called the *uncorrelated heterogeneity* and *correlated heterogeneity* respectively.

The correlated and uncorrelated heterogeneity of $\theta_i$ should be included in the analysis in order to clean the SIR and obtain a true risk surface:

$$\log(\mu_i) = \log(E_i) + \alpha + \beta_1 x_{1i} + ... + \beta_m x_{mi} + s_i + a_i$$

where '$s_i$' represents the structured (*correlated heterogeneity*) and '$a_i$' the unstructured (*uncorrelated heterogeneity)*. They can be included as random effects in a generalized linear mixed model (GLMM); but Bayesian modelling approaches allow the exact analysis of these random effects (Lawson and Zhou, 2005). Also, in the frequentist framework is not possible to include neighbourhood matrixes to take into account the *correlated heterogeneity*. Alternatively, Bayesian methods enable the use of big and complicate neighbour matrixes and therefore, they are a necessary tool when dealing with such analyses.

- ***Bayesian methods:***

In the past, statistical analyses based on **Bayes** theorem were often daunting because of the numerical integrations needed. Recently, computer intensive sampling methods of estimation have revolutionised the application of Bayesian methods in fields as diverse as biostatistics, econometrics and genetic mapping (Congdon, 2001).
In the classical statistics the probability is based on the relative frequency achieve when an experiment is repeated a lot of times in similar conditions, however, in the Bayesian perspective, the uncertainty about parameters or hypothesis is represented as a probability (Saez and Saurina, 2007). In the frequentist framework, parameters are fixed non-random quantities and the probability statements concern the data. In the Bayesian framework, probability statements are made about model parameters (Congdon, 2001) and it is assumed that the parameters have distributions. These distributions control the form of the parameters and are specified by the investigator, usually based on their prior belief concerning their behaviour.
Prior distributions and likelihood provide two sources of information about any problem. The likelihood informs about the parameter via the data, while the prior distributions inform via prior beliefs or assumptions. The product of the likelihood and

the prior distributions is called the posterior distribution. This distribution describes the behaviour of the parameters after the data are observed and prior assumptions are made. In Bayesian modelling all the inference about parameters is made from the posterior distribution. Posterior distributions are sampled to give a range of values (posterior sample) this contains a large amount of information about the parameter of interest (Lawson et al., 2003). The starting point of Bayesian statistics began in XVIII century with the Bayes theorem (Bayes 1702 – 1761) (Saez and Saurina, 2007):

$$p(\theta / X) \propto p(X / \theta) p(\theta)$$

Where:

$p(\theta)$ is the ***prior distribution***.
$p(X / \theta)$ is the probability distribution of the data, also named ***likelihood function***.
$p(\theta / X)$ is the ***posterior distribution***. In this probability distribution, the information contained in the prior knowledge about the parameter and the observed data has been combined to obtain the posterior distribution of the parameter of interest ($\theta$).
The election of the prior distributions is a subjective election and is base on expert knowledge and on the type of data that we have. When the prior and posterior distributions come from the same family of distributions the prior is said to be conjugate to the likelihood. The most common prior conjugate distributions for the usual likelihood functions are (Saez and Saurina, 2007):

| **Likelihood** | **Prior distribution** | **Posterior distribution** |
|---|---|---|
| Normal | Normal | Normal |
| Binomial | Beta | Beta |
| Poisson | Gamma | Gamma |

When a Bayesian hierarchical model is employed it is no longer possible to provide a simple point estimate for any parameter of interest. The parameters are no longer assumed to be fixed, they are assumed to arise from a distribution of possible values. Given the observed data, the parameter or parameters of interest will be described by the posterior distribution, and hence this distribution must be found and examined. It is possible to examine the mean or mode of the posterior distribution to give a point estimate for a parameter (Lawson et al., 2003).
In general, a posterior distribution can be so complex that simulation methods must be used to obtain samples. There are two basic methods for this:

- Gibbs sampling
- Metropolis-Hastings sampling

Both methods use simulation to generate sample values over large numbers of iterations. Samples are drawn from distributions to provide starting values that should converge to a stationary distribution (the posterior distribution) (Pfeiffer, 2004).
The 'Gibbs sampling' algorithm is one of the most simply and used for Monte Carlo Markov Chain (MCMC) calculations (Saez and Saurina, 2007). Moreover, WinBUGS (Bayesian Inference Using Gibbs Sampling) was developed for Gibbs sampling. In figure 4, Bayesian methodology is represented.

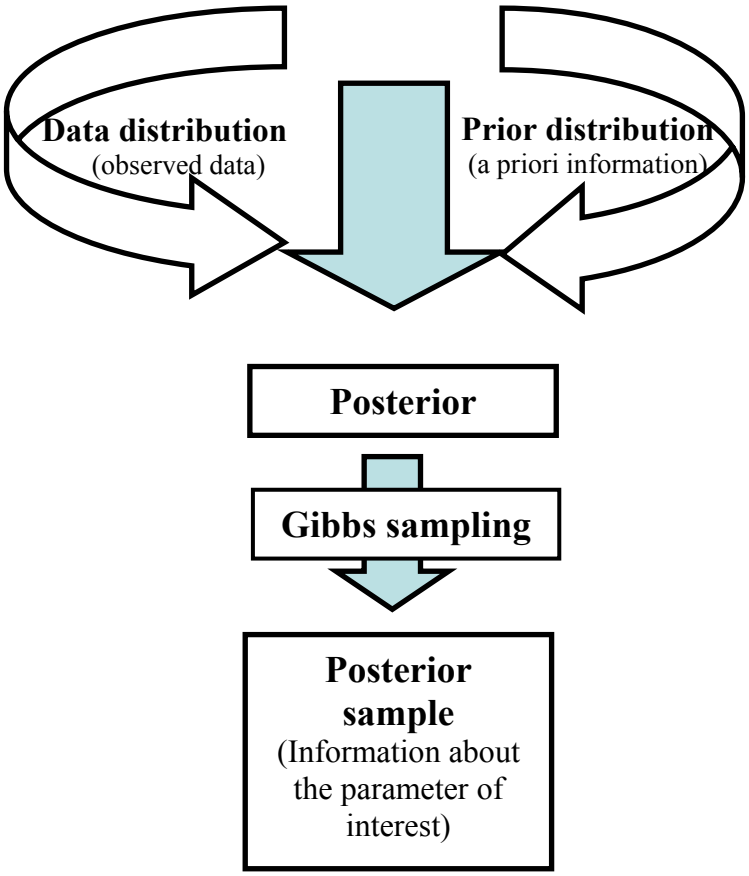


Figure 4: Schematic representation of Bayesian analysis methodology (source: own production)

Sampled chains require an initial burn-in period until they can be assumed to converge to the posterior distribution of interest (Lawson et al., 2003). Unlike most numerical methods used in statistical inference, MCMC does not give a clear indication of whether it has converged (Plummer et al., 2006). Judging convergence has been the subject of much debate and can still be regarded as art rather than science. There are a wide variety of methods now available to assess convergence of chains within MCMC (Lawson et al., 2003). Usual practice is to run multiple chains for the same model, starting for different initial values, and visually inspect 'trace' plots for convergence. Once convergence reached, samples should look like a random scatter about a stable mean value (Figure 5) (Best and Abellan, 2006). It is also possible to check the convergence of the algorithm using the Gelman-Rubin convergence diagnostic implemented in WinBUGS. This test is based on running parallel chains from different starting values; at convergence their ratio (R) is close to 1 (Lawson et al., 2003).

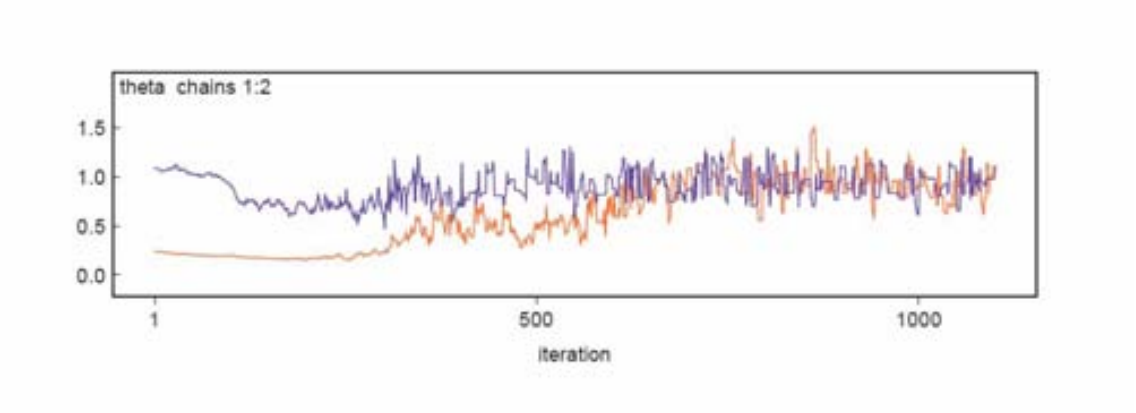


Figure 5. Trace plot to visually inspect for convergence. Convergence has been reached; as samples look like a random scatter about a stable mean value (source: Best and Abellan, 2006).

Accuracy of the posterior estimates can be assessed by using the Monte Carlo standard error for each parameter. To reach efficiency, the Monte Carlo error should be small in relation to the posterior standard deviation. As a rule of thumb, the simulation should be run until the Monte Carlo error for each parameter of interest is less than about 5% of the sample standard deviation (Lawson et al., 2003). A powerful feature of the Bayesian approach is that all the inference about parameters is made from the posterior distribution, so it is simple to calculate the probability e.g. $Pr(\theta > 1)$ (Best and Abellan, 2006).

For model goodness-of-fit measure, the deviance information criterion (DIC) proposed by Spiegelhalter (2002) measures the goodness-of-fit of any given model (Lawson and Williams, 2001). DIC is a measure of model fit penalised by the complexity of the model and its value is calculated by adding the effective number of parameters to the posterior mean deviance of the model. The ''best fit'' model is the one with the smallest DIC value (Lawson et al., 2003).

As mentioned before, the **correlated and uncorrelated heterogeneity** of the SIR can be included in the analysis by the use of random effects. The random effect for the **correlated heterogeneity ($b_i$)** can be linearly added to the model:

$O_i \sim \text{Poisson}(E_i\theta_i)$

$Ln(\mu_i) = \ln(E_i) + (\alpha + \beta_1 x_{1i} + ... + \beta_m x_{mi}) + s_i + a_i$

For the ($s_i$) term, the correlated heterogeneity, various choices exist, but the most common ones are the **Multivariate Normal model** (MVN) or the **Conditional Autoregressive Model** (CAR).

With the **MVN** the ($s_i$) term follows a multivariate normal distribution with spatially structures covariance matrix (Best and Abellan, 2006):

$s_i \sim MVN\ (0_i, \nu)$

where $\nu$ denotes the correlation between $s_i$ and $s_j$. A common practice is to model the correlation as function of distance:

$\nu_{ij} = f(d_{ij}\ ;-\ \varphi)$ ; where $\varphi$ describes the strength of the correlation and $d_{ij}$ denotes the distance between '*i'* and '*j'*.

Instead of directly modelling the correlation between log relative risks as a function of distance between areas, an alternative approach is to model conditional dependence between risks in different areas. Besag et al. (1991) proposed a conditional autoregressive (**CAR**) prior for the log relative risks in a disease mapping context in which the relative risk in one area is influenced by the average log relative risk of its neighbours. CAR models are the most commonly used models for smoothing risks (Ugarte et al., 2005). Besag et al., (1991) recommended modelling the $Ln(\theta_i)$ as the sum of two random components; a CAR model ($s_i$) and a random effect without spatial structure ($a_i$); the **uncorrelated heterogeneity**. This is the Besag, York and Mollie model:

$\log(\theta_i) = a_i + s_i$

Where;

$a_i \sim Normal\ (0,\nu)$

$s_i \sim CAR\ (W, \tau\ )$; W=0 if is not a neighbour and W=1 if they are neighbours.

Several times, certain areas or locations, presents a number of cases equal to cero. The excess of zeros in the data count relative to the Poisson distribution, leads to different approximations than the Besag, York and Mollie model, based on the standard Poisson distribution. In this context, data may follow a mixture distribution rather than a Poisson distribution. Lertxundi (2006) suggested that the threshold to consider that there is an excess of zeros in the data is when this proportion is over 40-50%. The use of mixture models is a common approximation for this problem. These models assume an underlying partition of the population into *k* homogeneous components, where each component has a different risk level, depending on possibly different covariates. Usually, there are two homogeneous components, one for zeros (non observed cases) and the other for those that are not zero (positive values). The most common approximations for mixture models are (Lertxundi 2006):

***Zero Inflated Poisson*** (ZIP) and ***Zero Inflated Negative Binomial*** (ZINB) distributions:
In ZIP models the response variable is modelled as a mixture of the Bernoulli (zero-counts) and Poisson (non-zero counts) distributions.
In practice, count data are often overdispersed so that alternative distributions such as the zero inflated negative binomial (ZINB) may be more appropiate than the ZIP (Xiang et al., 2007).

The ZINB distribution is a mixture distribution assigning a mass of *p* to 'extra' zeroes and a mass of (1 – p) to a negative binomial distribution, where $0 \leq p \leq 1$. A negative binomial distribution approaches a Poisson distribution when there is not overdispersion. The ZINB model therefore accounts for the 'excess of zeroes' and also for the extra heterogeneity in the positive outcome (Mwalili et al., 2004).

***Hurdle models:***

In this type of models zero-counts and non zero-counts are treat completely separately (Lertxundi, 2006). Locations with zero-counts could not have a non-zero count. In the case of veterinary medicine, this is not logical, as far as there are animals it must exist a risk level. Therefore, these types of models are not applicable in veterinary epidemiology.

Spatial statistical **disease modelling** can be aimed at investigating possible causal effects which are considered to be associated with the disease occurrence. In this framework, the model does not use any information on any known mechanisms of transmission or incubation for the disease. These models use a retrospective approach (Lawson and Zhou, 2005), usually require a large amount of data, and are generated from existing georeferenced data sets about disease occurrence as well as potential risk factors (Pfeiffer, 2004). They provide summaries and parameter estimates relevant to disease management (French and White, 2004). These studies are used in conjunction with simple descriptive studies of geographical variation in an attempt to determine how much of the variation in disease rates is associated with variation in exposure (Lawson et al., 2003). When working with aggregated data, the so-called ecological fallacy can arise. The ecological fallacy arises when an attempt is made to ascribe to individuals the characteristics which have been derived from the properties of groups of population (Lawson and Williams, 2001). For this reason, observations at the ecologic scale will usually need validation and replication at the individual level, for example, through cohort or case-control studies. The problem of ecological bias may be lessened when the analysis is carried out at a more local or small-area scale as the analysis is closer to the level of the individual (Elliot and Wartenberg, 2004). Despite this problem, ecological studies have some advantages over other studies because they include the ability to study a large population at a low cost (Pfeiffer and Jones, 2002).
On the other hand, mechanistic models include some form of transmission dynamic within their structure (Lawson and Zhou, 2005). Model parameters are used within a spatial and temporal framework to generate data in the form of predicted patterns of disease (French and White, 2004).

## 1.3. Application of Spatial Epidemiology tools in veterinary epidemiology

- **Application for research on Aujeszky's disease:**

In the **USA**, GIS and spatial statistics were used in the early 90`s to investigate about the spatial distribution of AD in different states and those factors implicated in the geographic distribution of the virus (Marsh et al., 1991; Austin and Weigel 1992; Norman et al., 1996):

In 1991, Marsh et al., (1991) developed a computerized database linked to a geographic information system developed by the Land Management Information Center of the **Minessota** State Planning Agency, to study the spread of ADV among swine herds in the state of Minesotta. They conducted a pilot study in one Minessota County with 280 swine herds, of which AD status was known in 115. GIS technology was used to obtain data sets describing topographical features: distance to nearest road, rivers and lakes; distance to nearest known infected herd and to count the number of herds within a given area. They used proportional hazard models (Cox regression analysis) to identify factors associated with AD infection. They did not found association between AD status of the herd and distance to nearest road or nearest infected herd. However, farrow to finish production system, a complete confinement management and density of swine herds within a 5 km radius, were considered risk factors. The presence of a river or lake within a 1km distance had a protective effect.
Austin and Weigel (1992) used second-order spatial analysis (spatial autocorrelation) methods to analyze the distribution of county AD prevalence rates (proportion of infected herds) in **Illinois**. The degree of spatial dependence among variables was calculated as a function of distance between points. For such analysis they used the centroids of the counties. This analysis revealed a clustering of counties with higher prevalence (tended to be close to one another), more so than the clustering observed for counties with larger numbers of swine herds. They conducted a multiple linear regression analysis to test for the association of county AD prevalence with county herd density, average herd size in the county and regional density of AD infection rates. These variables accounted for the 69% of the variance in AD county prevalence rates. The most important factor was the regional density of AD. County herd density and average herd size were also related but at a lower level.
Norman et al., (1996) conducted a spatial analysis of AD in **Pennsylvania** State. They developed a case-control study involving 123 infected herds and 162 non infected herds. In buffers of 1.61 km, 3.22 km and 6.44 km radius around each farm they calculated the density of infected, non infected, vaccinated, farrow to finish and farrow to weaning herds. They build multivariate models and used unconditional logistic regression to analyze the data. They did not have enough data to evaluate variables included in the 1.61 km buffer, but a decreased on the risk of AD was associated with high density of non infected herds and low density of infected herds within 3.22 km. Increased risk of AD was associated with low density of vaccinated herds and being a farrow to finish herd increased the probability for becoming infected, compared with farrow to weaning farms. They obtained similar results in the 6.44 km buffer.
In all of these studies conducted in different states of the USA, a particular route of transmission of the virus between farms was not ruled out, but all three highlighted the

importance of area spread in the transmission of AD infection between herds, and recommended to take into account for eradication efforts.

In **France**, Auvigne and Hery (1997) described the spatial distribution of AD in Brittany (Northwest France) and analysed the relationship between seroprevalence and pig density. They used a spatial moving average to reduce the influence of geographical units with few observations. This method calculates a corrected value for each variable which depends on its observed value and on the observed value in a radius. The corrected values were included in a linear regression model: dependent variable seroprevalence and pig density as explanatory variable. They run separately 4 regressions, one for each departament of Brittany. As previous studies conducted in the USA, pig density was found to be related to AD seroprevalence, but the intensity of the relation (slope and R-squares) was different from one department to another. The authors attribute these differences to the fact that pig density might be insufficient to describe the structure of the pig industry in one area (they did not take into account the ratio sows/fatteners). Also, other variables like mean size of herds or vaccination schemes were not included in the analysis.
More recently, Solymosi et al., (2004),conducted a spatial risk assessment of herd sero-status in a county in **Hungary** with all known pig herds (except pure fattening units). They used a stepwise logistic regression to test the association between sero-status of pig herds and presence of topographical features (lake, forest and highway) and uninfected herds in circular radius 1-10 km (in 1 km increments; one logistic regression for each radius) around each herd. They found that lake and highway were positively associated with AD seropositivity whereas the presence of a forest and uninfected herds had a protective effect. The authors speculate that the effect of the lake could be due to the existence of more abundant vegetation around the lake, which increases the risk through wildlife or wildlife vectors (e.g. rodents, foxes, birds) or to an increase in the possibility of airborne virus transmission because of the fog and higher humidity around a lake. On the other hand, the forest could decrease the risk of infection by decreasing air-borne transmission.
Berke and Beilage (2003) and Berke (2005) used data about AD infections at farm level (only from herds with breeding pigs) in a region of high animal density in **Germany** at the beginning of the national eradication project to propose an exploratory relative risk mapping approach. This approach enables the investigation of unknown spatially varying risk factors and facilitates hypothesis generation about putative risk factors (Berke, 2005). On the basis of serological findings 186 farms were classified as positive out of a total of 482 investigated farms (Berke and Beilage, 2003). This approach proceeds in four steps. First, they search for disease clusters, if any, in the study area using the spatial scan statistic and the population is classified as exposed (inside the cluster) and unexposed (outside the cluster). Second, they calculate the background risk (prevalence in the unexposed population) from the population at risk outside the high risk areas. Third, they create the risk surface map. To produce the risk surface map Kernel density estimates were used in order to give weighted means for each location in the study area. The risk surface is calculated as the ratio between the kernel density of cases and the kernel density of the population at risk. In the fourth last step, the risk surface was scaled using the background risk to obtain the exploratory relative risk map (Berke, 2005). The use of this approach to investigate AD distribution on Germany gave a complete presentation of the geographical variation in risk in the study area. Also, two clusters were identified using spatial scan statistics implemented in SaTScan software (www.satscan.com).

- **Application for research on other diseases different than Aujeszky's disease**:

Until recently, GIS, remote sensing and spatial statistics were underexploited in the field of veterinary epidemiology, due to the prohibitive cost of hardware and the complexity of GIS software that required a high level of expertise . The developments in computer performance of the last decade have not only reduced the costs of equipment but have made available easy-to-use Web-based software which in turn have meant that GIS are more widely accessible by veterinary services at all levels (Calistri et al., 2007a).
Also, in the last years, due to computational advances, spatial analysis has become more easily accessible for epidemiologists. However, while developments in spatial statistics within the human health field have advanced considerably, there has been less evidence of a corresponding interest in such methodology within veterinary medicine (Lawson and Zhou, 2005).

- **Application of GIS in veterinary activities:**

The spatial tools available in modern GIS software such as theme overlays, buffering and spatial filtering, have been used to generate single or combinations of variables to be used within a modelling framework (Durr et al., 2000). Also, the increased awareness of the possibilities offered by GIS has created new opportunities for decision-makers to enhance their planning, analysis and monitoring capabilities. A picture is worth a thousands words and decision-makers increasingly require that information is mapped (Wint, 2007). GIS technology has been used for the monitoring of animal diseases and zoonoses, for surveillance activities and as a web tool for data and knowledge sharing.

Among others, it has been used by Perruci et al., (2007) to **monitor** parasitic infections in an organic grazing cattle herd in the Natural Park of Migliarino-San Rossore-Massaciuccoli (Pisa). They collected samples from cattle between 2002 and 2006 every two to three months and also environmental parameters and data on biodiversity.
GIS was used to determine the land-use heterogeneity based on land-use sequences between the pixels in six different transects made in the area, to create vector maps of the different grazing areas and to plot on a map all collected data and evaluate the presence, prevalence and intensity of endoparasitic diseases.

GIS technology has been used for the **surveillance** of different diseases. In Bluetongue, it has been used for serological and vector (*Culicodes* biting midges) surveillance activities in different countries (Capela et al., 2003, Torina et al., 2004, Calvete et al., 2006, Purse et al., 2006). GIS has been useful to produce detailed maps of the spatial distribution of the main vector in the Mediterranean Basin (*Culicoides imicola*) and that of potential novel vectors, and also for visual comparison of the spatial distribution of Culicoides *complexes* and disease outbreaks.
Also in the **surveillance** framework, it has been used as a system applied to the international surveillance and control of transboundary animal diseases by the development of the global animal health information system (EMPRES-i). This is a programme developed by the Food and Agriculture Organization (FAO) that supports the Emergency Prevention System for Transboundary Animal and Plant Pests and

Diseases (EMPRES). EMPRES-i monitors data on major transboundary animal diseases: Highly Pathogenic Avian Influenza (HPAI), Foot and Mouth Disease (FMD), Rift Valley fever (RVF), Contagious Bovine Pleuropneumonia (CBPP), African swine fever (ASF) and rinderpest. Through the use of GIS, disease observation data entered into EMPRES-i can be visualised as basic maps and combined with additional geographical layers such as land use, poultry density, remote sensed imagery or other variables hypothesised as relevant to the epidemiology of the disease, providing with timely and accurate situation updates (Martin et al., 2007).

The use of GIS as a **web tool** for **data and knowledge sharing** is exploding in the last years mainly due to the spread of globalisation with its consequent impact on trade, information exchange networks and emerging diseases. The veterinary and research communities that utilise and adapt GIS analyses are expanding and the demand from government and media for disease-related information and early warning systems is increasing. All of this depends on networked data exchange and standardised data processing (Wint, 2007). The United States Department of Agriculture (USDA) Veterinary Services, to date, have the following mapping applications: Equine Infectious Anemia Testing, National Tick survey distribution maps, the emergency Management Response System-Mapping Module for disease investigations and emergency outbreaks, and the Scrapie mapping module to assist with the control and eradication of this disease. Thanks to this Web-based interactive mapping service, users can use spatial data through a familiar internet browser, interact with data and create a customised output, a map or a spreadsheet. The functionality of the applications can be used without having to purchase GIS software or learn GIS software (Maroney et al., 2007). Also, since the Bluetongue spread across the Mediterranean Basin and into the Balkans, an information network based on internet links and on geographic information system website technologies have been established between Balkan and Eastern Mediterranean countries (East-BTNet project). The GIS application developed it is not only used for the presentation and analysis of spatial data, but it represents an exclusive interface for accessing data entry and retrieval forms (Calistri et al., 2007b).

- **Application of Remote Sensing images in veterinary activities:**

Applied to epidemiological research, remote-sensing techniques can be used to determine and monitor factors involved in disease transmission. Historically, the first studies were carried out on schistosomosis, malaria, tick-diseases and trypanosomiasis and then applied on many other vector-borne diseases (De La Rocque et al., 2004).
Tatem et al., (2003) processed remotely sensed imagery to obtain variables with environmental significance like the normalized difference vegetation index (NDVI), middle infra-red reflectance (MIR), land surface temperature (LST) and air temperature (TAIR). By discriminant analysis, they identified the combination of remotely sensed variables that most effectively allocated the presence-absence and abundance of *Culiciodes imicola* captured during summer 2000 and 2001 in 87 sites across Portugal. The models where then used to predict the distribution and abundance of *Culiciodes imicola* for the rest of Europe and North Africa. *Culiciodes imicola* was predicted to be present and in high abundance in Baleares, Sardinia, Corsica, Sicily, areas of mainland Italy, large areas of Greece, western Turkey, northern Algeria and Tunisia.
Guis et al., (2007) used environmental data related to **bluetongue (BT)** occurrence in Corsica (France). They assumed that the active flight range of *C. imicola* should range between a few hundred meters and a few kilometres, and therefore they tested three

sizes of buffer zones: 0.5, 1 and 2 km in radius around each farm. In order to identify potential landscape risk factors, they described the environmental surrounding using a high spatial resolution SPOT image and a digital elevation model (DEM). From the DEM they obtained the altitude, slope, sunshine index and aspect and from the SPOT image the NDVI, a land-cover map and landscape metrics. The results reveal the role of landscape structure, particularly those characterizing prairies and woodlands, as well as farm type, latitude and sunshine to explain the presence of bluetongue. This approach was adapted to a different French region where *C. imicola* was present but not the virus, to help focus the trapping sites of the entomological surveillance system in the areas where BT risk was the greatest.
Leblond et al., (2007) also used remote sensing technologies to identify landscape features of Camargue area (France) associated with risk of **West Nile virus** transmission as defined by the presence of confirmed horse cases. They classified the landscape categories (open water, rice fields, salty ponds,…) from a SPOT-4 satellite image. By using generalized linear models they tested the relationship between the presence of cases and proportions of landscape categories in the wet area. They found that rice fields, wet "sansouire" and open water were the major components of the landscape that were associated with the presence of West Nile virus cases.

- **Application of spatial statistics in veterinary activities:**

As mentioned before, the analysis of spatial data consists of 3 components: ***exploratory analysis*** (finding interesting patterns), ***visualization*** (showing interesting patterns) and ***spatial modelling*** (explaining interesting patterns).

In the framework of **exploratory analysis**, different techniques have been used: the Cuzick-Edwards´s test (global cluster test), has been used to identify clusters of antibodies against bluetongue virus in cattle in Australia (Ward and Carpenter, 1995), *Corynebacterium pseudotuberculosis* cases in horses in California (Doherr et al., 1999), cases of bovine spongiform encephalopathy in animals born after the feed ban in Switzerland (Doherr et al., 2002) and of bovine tuberculosis in Argentina (Perez et al., 2002). Also, the K-function, as a global cluster test has been used by different authors; it has been used by O'Brien et al., (2000) to study the spatial and temporal distribution of cancers in Michigan and to compare this distribution with cancers in human population. French et al., (1999) used the method to investigate the space-time pattern of sheep scab in Great Britain between 1973 and 1992. Also it was used by Picado et al., (2007) to describe the spatio-temporal pattern of the risk of infection during the 2001 Foot and Mouth Disease outbreak in United Kingdom.
Perhaps the most widely exploratory technique used has been the local cluster test, spatial scan statistic implemented in SaTScan software (www.satscan.org). In the last years, SaTScan has been widely used to search for clusters on domestic animals as well as on wildlife populations. Among others, the spatial scan statistic has been used to identify clusters of bovine spongiform encephalopathy in Great Britain (Stevenson et al., 2000) Switzerland (Doherr et al., 2002) Ireland (Sheridan et al., 2005) France (Abrial et al., 2003) and Spain (Allepuz et al., 2007). It has also been used to identify clusters of bovine tuberculosis in Argentina (Perez et al., 2002); bovine tuberculosis in badgers populations in Ireland (Olea-Popelka et al., 2003) and to assess, also in Ireland, the spatial relationship between Myocbacterium bovis strains in cattle and badgers (Olea-Popelka et al., 2005). It has also been used to search for clusters of Aujeszky's disease in an animal-dense region in Germany at the beginning of the national

eradication programme (Berke and Beilage, 2003). In the United States has been used to propose an early warning system for West Nile Virus activity based on the detection of dead bird clusters (Mostashari et al., 2003). In Denmark Vigre et al., (2005) used it to investigate if the incidence of herds diagnosed with Postweaning Multisystemic Wasting Syndrome (PMWS) were particularly high in some geographical areas. Mainar et al., (2005) used spatial scan statistics to identify and locate significant spatial clusters of small ruminant brucellosis in a northern province of Spain between 1997 and 1999, and Alba et al. (2008) used it to study the spatial distribution of Maedi and Pestivirus infections in sheep in Catalonia. In Ontario (Canada) it was used to test for spatial clustering of swine influenza (Poljak et al., 2007).
As mentioned before, while exploration and visualisation are used as descriptive tools for generating causal hypothesis, **epidemiological modelling** of spatial data aims to explain or predict the occurrence of disease (Pfeiffer, 2004).

In 2001 **Foot and Mouth Disease** (FMD) Outbreak in Great Britain **mechanistic spatial models** were constructed, predicting the effectiveness of intervention efforts and thus helping to inform policy-making (Boender et al., 2007). Among others Keeling et al., (2001) and Morris et al., (2001) developed mechanistic models for FMD. In the model of Keeling et al., (2001) the likelihood of transmission between farms was determined by their spatial separation and a distance kernel. The kernel was estimated from contact tracing performed during the epidemic and described the relationship between spatial separation and the likelihood of transmission (by any route). Every day, the probability of infection was calculated for each farm, and this was used to determine, by Monte Carlo simulation, whether the event happened. The model was used to produce a map of predicted cases. Morris et al., (2001) like Keeling et al., (2001) developed a model that was location-specific and predicted both the scale and spatial pattern of the disease. The probability that each infected farm would transmit to any other farm on a given day was determined by a set of probability distributions. Four mechanisms of transmission were simulated: local spread to nearby farms via fomites or personnel, spread by the movement of animals to farms or markets, long-distance wind-borne spread, and dairy tanker movements. These models were a significant development and they contributed to the outbreak and its control (French and White, 2004).
Boender et al., (2007), developed a spatial mechanistic model for the spread of Highly Pathogenic Avian influenza poultry in the Netherlands. On each day, farms were classified as susceptible (S), being infected on that day (B), infected but not infectious (E), infectious (I) or removed (R). Based on data collected during an outbreak of a highly pathogenic H7N7 avian influenza virus in The Netherlands in 2003, they estimated model parameters. The central concept of the model was the transmission kernel, which determines the probability of pathogen transmission from infected to uninfected farms as a function of interfarm distance. Their model generated risk maps for the spread of avian influenza in poultry and showed that there are two poultry-dense areas in The Netherlands where epidemic spread is possible, and in which local control measures are unlikely to stop the disease. In these areas an epidemic can only be brought to an end by the depletion of susceptible farms by infection or massive culling.

Bovine Spongiform Encephalopathy (BSE), Influenza and Tuberculosis (TB) among others, have been studied by different authors in the framework of spatial statistical disease modelling aimed at **describe spatio-temporal spread** and **investigate hypothesised causal effects**.

In relation to **BSE**, Ducrot et al., (2005) by means of hierarchical Bayesian models based on a Poisson distribution with spatial smoothing, described the spatial evolution of BSE risk in different periods in France. They built four different analyses to examine the spatial distribution of the BSE risk for different successive birth periods and they found that there was a significant spatial heterogeneity of the BSE risk for each of them. Also, the areas with the highest BSE risk were different for each birth cohort. Also, in Great Britain Stevenson et al., (2005a) used Bayesian hierarchical models to evaluate area-level BSE risk for cattle born before and after the July 1988 ban on feeding ruminant-derived meat and bone meal to ruminants. In Spain, these models were used by Allepuz et al., (2007) to evaluate the BSE-risk distribution before and after the implementation on changes in MBM processing. In France, Great Britain and Spain covariates were included in the Bayesian hierarchical models in order to test hypothesis that could explain the BSE-risk spatial distribution. This BSE risk seemed to be linked with low levels of cross-contamination of cattle feed by pig feed (Abrial et al., 2005b; Stevenson et al., 2005a; Allepuz et al., 2007).

In **Influenza**, Pfeiffer et al., (2007), used a random effects logistic regression model to analyse the association between various risk factors and the risk of outbreak occurrence at commune level during a particular epidemic wave period of the HPAI occurrence in Vietnam. The risk of outbreak occurrence increased with a greater percentage of rice paddy fields, increasing domestic water bird and chicken density. The authors proposed to consider the resulting risk maps when defining areas of high and low risk as part of risk-based surveillance. On a cross-sectional study of swine influenza in Ontario (Canada) in finisher herds in 2004 and 2005, Poljak et al., (2008), observed a different transmission pattern of the epidemic H3N2 subtype compared to the H1N1 subtype. Only the H3N2 positive herds in 2005 clustered in space. They used logistic regression modelling and their results suggest that both, infection between neighbouring farms or a common geographical risk factor, might have played a role in the H3N2 transmission pattern. A negative association between the probability of being an H3N2 positive herd and the reported distance to the nearest pig herd was suggestive of local spread either by airborne, trough biological vectors or mechanical vectors and fomites. Type of gilt source might be a common geographical risk factor if groups of premises located in close proximity used the same source of H3N2-infected breeding animals.

In south-east of New Zealand Porphyre et al., (2007), described the temporal and geographical distribution of confirmed cases of **bovine tuberculosis** (TB). They defined six-year periods, between 1980 and 2003, in order to coincide with changes against the possum (TB reservoir) depopulation strategies. Prior to map the TB incidence rate by period they applied empirical Bayesian methods to smooth it. These methods estimate the posterior distribution on the basis of applying maximum likelihood procedures to existing data, whereas fully Bayesian methods generate the posterior using a sampling process (Pfeiffer, 2004). They also used SaTScan to identify clusters of disease in each period and results were compared with disease mapping. The Bayes adjusted incidence maps, and the scan statistic analyses showed that changes in the spatial and temporal pattern of TB were associated with spatial and temporal changes in possum control strategies applied throughout the study period.

Miller et al., (2007) used spatial scan statistics to identify TB clusters in white-tailed deer in north-eastern Michigan. This analysis enabled them to divide the deer population into cases (TB cluster areas) and controls (areas outside TB clusters). They used logistic-regression models with spatial effects, by using a neighbourhood

approach, to identify associations between possible risk factors and cluster membership. They included as covariates in the model: land-use, land-type, numbers of large feeding sites collected by aerial surveillance, deer related data (average age and percent group that were males) and interaction terms. They found that factors that promoted congregation of deer for extended periods of time (natural cover, access to water, and less human contact) appeared to be associated with increased odds of TB positivity.

# Conclusion:

In the last years, due to computational advances, spatial epidemiological analysis has become more easily accessible for epidemiologists (Lawson and Zhou, 2005). Increased data quality and availability through the development of modern animal disease and production monitoring and surveillance systems, the ability to sort and recombine data using GIS, and the increasing availability of software packages over the past three decades, have created an ideal environment for epidemiologists to apply spatial and temporal analytical techniques to disease problems (Ward, 2007).
At the same time, the increased awareness of the possibilities offered by GIS, spatial analysis and remote sensing has created new opportunities for decision-makers to enhance their planning, analysis and monitoring capabilities (Calistri et al., 2007a).
Nowadays spatial epidemiological analysis methods are becoming a standard component of the epidemiologist's tool chest of analysis techniques.

# 2. OBJECTIVES

**Main objective**: Conduct a spatial analysis of the Aujeszky's disease (AD) eradication programme in Catalonia, Spain, from 2003 to 2007. This objective can be detailed in the following points:

**1. -** Explore for *high risk areas* (clusters) in order to test whether the spatial distribution of AD in the region during the consecutive eradication periods was homogeneously distributed over the territory or clustered in space.

**2.-** Identify areas where reinfection and elimination of AD in sow farms (farrow to weaning and farrow to finish) were more likely to occur.

**3. -** Evaluate the effect of geographical factors on the success of the eradication of AD during the different phases of the eradication programme in Catalonia.

# 3. MATERIAL AND METHODS

## 3.1. Study area

The study was conducted in Catalonia in northeast Spain. This is the largest pig producing region in the country with a census of 6.3 million pigs and 5,700 farms, which represents 25% and 4% of the national and European pig population respectively (Anonymous, 2007b).

## 3.2. Data

Number of positive and negative animals tested by the gE official enzyme-linked immunosorbent assay method (ELISA) (INGEZIM® ADV gE, Ingenasa; CIVTEST® suis ADV gE, Hipra) for each farm were obtained through the Department of Agriculture of the Autonomous Government of Catalonia and through pig health associations. They also provided information on the type of farm (farrow to weaning, farrow to finish and fattening), date of sampling, geographical coordinates and census. The herd was considered the epidemiological unit of study. A herd was considered positive if at least one animal presented a positive result to the gE in the official competitive enzyme-linked immunosorbent assay method. The geographical coordinates of pig slaughterhouses were also obtained from the Department of Agriculture. Cartographic database of roads of Catalonia were obtained from the geographical information system services of Girona province (SIGTE). Only conventional roads were included in the analyses. Motorways and pathways were excluded from the cartographic database of roads.

## 3.3. Study periods

The study has been conducted through the different phases considered in the official eradication program. Each phase elapsed for about 1 year and stablished more strict measures to control the infection:
*Phase 1*: April 2003 – May 2004: Serological testing of fatteners. Movement restrictions were applied on farms with prevalence higher than 20%.
*Phase 2.1*: June 2004 – May 2005. Movement restrictions were applied on those farms with prevalence higher than 10%
*Phase 2.2:* June 2005 – May 2006. Movement restrictions were applied on those farms with at least one seropositive animal.
*Phase 3:* June 2006 – May 2007. Qualification of farms and areas was applied by a compulsory programme.
Measures adopted during each of the phases are explained with more detail in the introduction chapter.

## 3.4. Descriptive epidemiology

For the descriptive epidemiology SPSS (v. 14) (SPSS Inc. Chicago, USA) and Epi Info (v 3.3.2) (http://www.cdc.gov/epiinfo) softwares were used. ArcMap® v9.1 of ESRI, Redlands, CA was used to represent pig farm density. This density was represented with a Kernel density surface calculated with a bandwidth of 5km.

## 3.5. Exploratory analysis

The objective of this part of the study was aimed at exploring whether the spatial distribution of AD in Catalonia, at the end of each eradication period from 2003 to 2007, was homogeneously distributed over the territory or clustered in space. We also aimed at determining areas where reinfection and elimination of AD in sow farms (farrow to weaning and farrow to finish) were more likely to occur.
For that purpose, we used SaTScan® v6.1 (http://www.satscan.org). As explained in the introduction chapter, the spatial scan statistic searches for clusters by using a variable circular window size to detect spatial clusters in large areas while controlling for the underlying population (Kulldorff, 1997). The circle is centred on each of the points. For each point, the size of the circle varies from zero to an upper limit specified by the user. As recommended by Kulldorf (1997), a 50% of the population at risk was selected as the upper limit for the circles.
To explore the spatial distribution of Aujeszky's disease, we ran different purely spatial analyses based on the Bernoulli model in each eradication period. We only scanned for clusters with high rates (more cases than expected). The number of Monte Carlo simulations was set at 999, and clusters with p-value<0.05 were considered as statistically significant. Only statistically significant non-overlapping clusters were reported (Kulldorff, 2006).
We performed four different purely spatial analyses in each period with the following case/control definitions represented by a 1/0 variable:

a) Positive sow farms: Sow farms with at least one positive sample (cases) and sow farms with no positive samples (controls) at the end of each eradication period.
b) Positive fattening farms: purely fattening farms with positive samples in any of the analyses carried out during each period (cases) and negative fattening farms during the whole period (controls).
c) Elimination analysis (based on sow farms which became negative): Sow farms with positive samples at the beginning of each period and negative at the end of the period (cases) and positive sow farms during the whole period (controls).
d) Reinfection analysis (based on sow farms which became positive): Sow farms with negative samples at the beginning of each period and positive at the end (cases) and negative sow farms during the entire period (controls).

The resulting clusters were coded using numbers and letters. The first number corresponds to the period of the AD-eradication programme in Catalonia: period 0 (at the start of the study: April 2003), period 1 (from April 2003 to May 2004), period 2.1 (from June 2004 to May 2005), period 2.2 (from June 2005 to May 2006) and period 3 (from June 2006 to May 2007). The capital letter indicates the type of cluster analysis:

'S': Positive sow farms. 'F': Positive fattening farms. 'E': Elimination analysis 'R': Reinfection analysis. The lower case letter indicates the cluster detected by the spatial scan statistic, where the letter 'a' represents the most likely cluster and 'b', 'c', etc; the secondary clusters.
In order to explore the relation between the results of the analyses, we exported the cluster information file to ArcMap® v9.1. The different clusters were represented on a map of Catalonia and over a Kernel density surface of pig farms (farrow to weaning, farrow to finish and fattening) calculated with a bandwidth of 5km.

The density of pig farms was calculated for elimination clusters and reinfection clusters. In order to do so, the area of each cluster was calculated with the Spatial Statistical Tools of ArcToolbox from ArcMap®. Then, the number of pig farms inside each cluster was calculated by merging the cluster and pig farm themes. Finally, pig farm density for each cluster was obtained by dividing the number of farms per cluster by the cluster's area.
The proportion of positive fattening pig farms within eradication and reinfection clusters in each period was obtained from the database by merging the cluster and pig farm themes. The Shapiro Wilk test was used to test for the normality of the farm density variable; the p-value was below 0.05, and therefore the variable was not considered to be normally distributed. Due to this, a non-parametric alternative to a t-test (Wilcoxon rank-sum test) was used to test for farm density differences between clusters of reinfection and elimination. A Pearson's chi-square test was used to test for differences in the proportion of positive fattening farms between reinfection and elimination clusters. Odds ratios and their 95% confidence intervals were also calculated. Differences were considered statistically significant when $p<0.05$. SAS software (version 9.1 SAS Institute Inc., Cary, NC) was used for these analyses.

## 3.6. Modelling

The aim of this part was to evaluate the effect of geographical factors on the probability of being AD positive in Catalonia during the different phases of the eradication programme.

- **Explanatory variables included in the analysis:**

Four geographic variables were included in the analysis: a) Distance to the nearest slaughterhouse, b) distance to the nearest conventional road, c) mean number of AD serological positive sows in the neighbourhood (750 meters) of each sow farm, and d) mean number of AD serological positive fattening pigs in the neighbourhood (750 meters) of each sow farm. A non geographic variable, type of farm (farrow to finish *vs* farrow to weaning) was also included in the analysis.
ArcGIS 9.1 was used to map and manage spatial information. A joint, based on spatial location, was used to calculate the shortest Euclidean distance between sow farms (farrow to weaning and farrow to finish) and slaughterhouses and conventional roads.
For each period, the mean number of positive animals in the neighbourhood of each sow farm was calculated: a buffer radius of 750 meters was created around each sow farm. This buffer was merged with pig farm themes, and a list with the serological results of all fattening and sow farms on the neighbourhood of each sow farm was obtained. Density of positive sow and fattening animals was calculated as:

$\sum(p_i*n_i)/N_i$

Where i=1,…,n indexes neighbour farm (sow or fattening), 'p' is the proportion of positive animals to the gE official ELISA, 'n' is the number of sows or fattening animals in the farm and 'N' is the total number of fattening or sow farms in the 750m buffer radius.
In order to simplify the numerical fitting of the models, distance to nearest slaughterhouse and nearest conventional road variables were standardised resting them by their mean and dividing them by their standard deviation. Mean number of AD serological positive sows and AD serological positive fattening pigs in the neighbourhood of each sow farm were divided by 1,000.

- **Disease mapping:**

The aggregation level used in the analysis was the farm (i = 1, . . . , n). Being 'n' the number of sow farms (farrow to weaning and farrow to finish) in each eradication period. A positive sow farm was defined as a sow farm (farrow to weaning or farrow to finish) where there is at least one positive animal to the gE official ELISA.
In disease mapping, it is often to find two basic forms of extra variation. First, as in the no spatial scale, a form of independent and spatially uncorrelated extra variation can be assumed. This is called *uncorrelated heterogeneity* (Lawson et al., 2003). Secondly, geographically close areas tend to have similar disease rates (Clayton and Bernardinelli, 1992) so independence of risk between close areas cannot be assumed. This is called *correlated heterogeneity*. This form of extra variation implies that there exists spatial autocorrelation between spatial units. This autocorrelation could arise because the disease of concern could be naturally clustered in its spatial distribution at the scale of observation. Many infectious diseases display such spatial clustering, and a number of apparently non-infectious diseases also cluster (Lawson et al., 2003).
To model the extra variation of the parameter of interest and take into account the risk dependence between close areas, spatial hierarchical Bayesian models were used.
These models have the advantage of the possibility to control the variability of the parameter of interest by the use of prior distributions, overcoming therefore the classical problem of overdispersion in disease mapping (Abrial et al., 2005).

- **Specification of the model**

Because a sow farm can be positive or negative we assumed that the number of positive sow farms ($O_i$) follows a binomial distribution:

$O_i \sim$ binomial ($p_i$ ,$n_i$)

Where i=1,...,N indexes the farm, 'n' is the number of sow farms in each eradication period and 'p' is the probability of being a positive sow farm in each period.

The probability of being a positive sow farm is given by:

logit ($p_i$)= $\beta_0 + a_i + s_i$

Where '$a_i$' and '$s_i$' indicate the uncorrelated heterogeneity and spatially-correlated heterogeneity random effects respectively.

- **Uncorrelated and correlated heterogeneity**

For the prior distribution of $a_i$ we chose a normal distribution centred on zero:

$a_i \sim N(0,v)$

where '$v$' is the precision (inverse of variance) of the prior distribution.

To take into account the spatial autocorrelation, we used the conditional autoregressive (CAR) model with a spatial component ($s_i$) proposed by Besag et al., (1991).
This spatial component takes into account the spatial dependence between neighbour farms (500 meters radius). The prior distribution of this parameter follows a normal distribution:

$s_i \sim N(b_i, \sigma_{ni})$

where '$b_i$' is the mean of the spatial component in the set of farms neighbours to farm '$i$' and '$\sigma$' is the precision, weighted by '$n_i$', the number of neighbours of farm '$i$'.

- **Incorporating covariates in the models**

The number of slaughtered animals is not homogeneous over the different slaughterhouses of the region. Also, the number of transported pigs is not homogenous in the different roads. Because of that, we assumed that the different slaughterhouses and roads influence '$p_i$' differently so the regression parameters were individualised. The prior distribution for these parameters was assumed to follow a normal distribution:

$\beta_1 \sim N(0,\tau_1)$
$\beta_2 \sim N(0,\tau_2)$

where '$\tau_1$' and '$\tau_2$' are the precision of the prior distribution of '$\beta_1$' and '$\beta_2$' respectively.

Finally, the four random effects (*uncorrelated and correlated heterogeneity, distance to nearest slaughterhouse and conventional road*) and the three covariates (*type of farm, mean number of positive sows and positive fattening animals)* were linearly added to the prior distribution of '$p_i$':

$$(1)\ \text{logit}(p_i) = \beta_0 + \beta_1\, type_i + \beta_2\, sows_i + \beta_3 fattening_i + \beta_{4i}\, slaughterhouse_i + \beta_{5i} road_i + a_i + s_i$$

The residuals of the model (resid) were computed as:

$resid_i = O_i - p_i$

The spatial locations of the residuals were useful to evidence the geographical areas where the probability of being a positive sow farm were less explained by the explanatory variables included in the model.

Prior distributions were established for the parameters. Non-informative prior knowledge was considered with a flat distribution for the intercept ($\beta_0$) and normal distributions for covariate parameters and random effects. These normal distributions have a 0 mean and their precision follow a gamma distribution (hyperprios) of mean 0.1 and variance 1,000 (Lawson et al., 2003).

To implement the model we used WinBUGS version 1.4.3 (Medical Research Council, Biostatistics Unit of Cambridge, London, http://www.mrc-su.cam.ac.uk/bugs/welcome.shtml), a free software for Bayesian Inference using Gibb Sampling. The posterior distribution of the parameters was obtained from 45,000 Markov chain Monte Carlo simulations (MCMC) after a burn-in of 5,000 simulations. Conformity analysis was made from the percentile posterior distribution given by the quantiles of the usable chain, making appropriate interpretations in the context of disease mapping (Richardson et al., 2004). Convergence of the chains was checked using the sample trace plot and Gelman-Rubin convergence diagnostic of the deviance (as is a global parameter of the model) (Lawson et al., 2003). In appendix 1 the sintaxis of the model is displayed.

- **Construction of the neighbouring matrix**

In order to apply the CAR model to our data, we defined a neighbour farm as a farm located in a distance of 500m of each sow farm. In order to do so, a 500m radius buffer around each sow farm was created in each period. The neighbouring matrix (W) was created based on the intersections of the farm buffers. If the buffers of two farms (*i* and *j*) intersect each other, then these farms are neighbouring and the element $w_{ij}$ of W neighbouring matrix is 1 otherwise 0. The neighbourhood matrix calculation was performed using maps2WinBUGS version 1.5.1 (Solymosi et al, 2008).

- **Displaying results into a map**

Because the interpretation problems that present the plots of point events, we decided to use interpolation techniques in order to represent the results on each eradication period. We used the inverse distance weighting (IDW) calculated with a fixed bandwidth of 4km, in order to represent the predicted probability of infection by the model, the observed infected sow farms and the residuals of the model.

# 4. RESULTS

## 4.1. Descriptive analysis

### 4.1.1. Study population

In April 2003, (period 0) just at the beginning of the study, a total of 5,504 farms (with 3,546,656 animals) were included in the analysis; and at the end of the study, in June 2007 (period 3) the number of farms was 5,854 with 4,301,011 animals. The number of farms and animals during the different periods of the study are shown in table 1.

| | | | Period 0 | Period 1 | Period 2.1 | Period 2.2 | Period 3 |
|---|---|---|---|---|---|---|---|
| **Number of farms** | Farrow to finish | | 1,152 | 1,158 | 820 | 874 | 1,053 |
| | Farrow to weaning | | 857 | 869 | 665 | 738 | 800 |
| | Fattening farms | | 3,495 | 3,495 | 1,770 | 3,734 | 4,001 |
| | **Total** | | **5,504** | **5,522** | **3,255** | **5,346** | **5,854** |
| **Number of animals** | Farrow to finish | Sows | 202,775 | 204,890 | 154,344 | 178,164 | 222,019 |
| | | Fatteners | 1,429,094 | 1,843,543 | 588,037 | 648,209 | 813,954 |
| | Farrow to weaning | | 217,510 | 232,348 | 225,854 | 284,354 | 314,392 |
| | Fattening farms | | 3,126,371 | 3,126,371 | 1,424,130 | 3,351,185 | 3,764,600 |
| | **Total** | | **4,975,750** | **5,407,152** | **2,392,365** | **4,461,912** | **5,114,965** |

Table 1. Number of farms and animals during the different periods of study; period 0 (April 2003), period 1 (May 2004), period 2.1 (May 2005), period 2.2 (May 2006) and period 3 (May 2007).

Farms are not homogeneously distributed over the study region. The geographic distribution of fattening and sow farms is represented in figure 6:

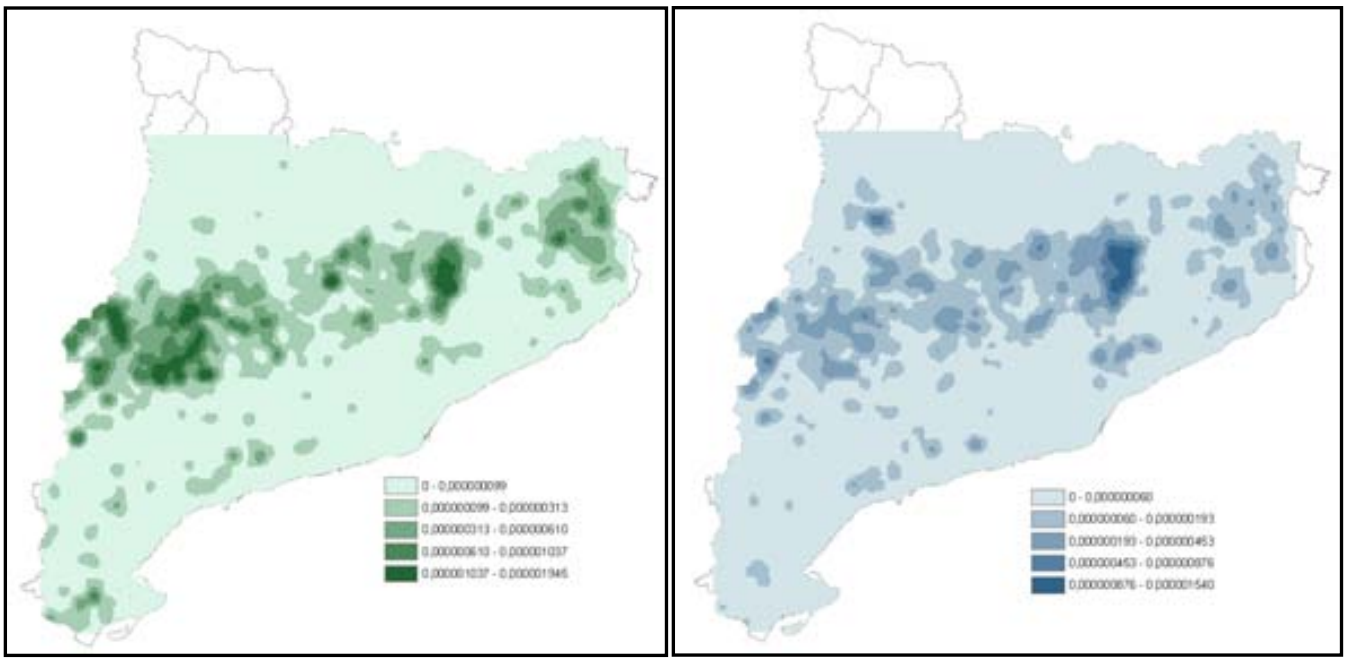

Figure 6. Kernell density surface, calculated with a 5km bandwidth, of fattening farms (left) and sow farms (right) in the study region.

## 4.1.2. Serological results

At the beginning of the study (April 2003) 1,212 out of 2,009 **sow farms** in Catalonia had at least one sow with positive serological results to AD (60.3%). At the end of period 1 (May 2004), the proportion was similar, 1,194 out of 2,027 (58.9%). In May 2005 (end of period 2.1), the proportion decreased to 54.1% (804 with positive serological results out of 1,485 sow farms). At the end of period 2.2 (May 2006) the proportion was 23.4% (377 out of 1,612), and at the end of period 3 (May 2007) it fell to 9.7% (180 out of 1,853). The proportion of positive sow farms (farrow to weaning and farrow to finish) at the end of each eradication period and the within prevalence in positive sow farms are represented in tables 2 and 3.

| | **Proportion of positive farrow to weaning farms** | | | **Prevalence in positive farrow to weaning farms (%)** | | | |
|---|---|---|---|---|---|---|---|
| **Period** | **Pos** | **Neg** | **Perc(%)** | **Mean (sd)** | **25%** | **50%** | **75%** |
| **April 2003** | 493 | 364 | 57.5 | 32.2 (29.8) | 8.3 | 20.3 | 50 |
| **Period 1 (May 2004)** | 477 | 392 | 54.9 | 26.4 (28.2) | 5.4 | 13.9 | 38.6 |
| **Period 2.1 (May 2005)** | 369 | 296 | 55.5 | 19.8 (23.3) | 3.4 | 9.8 | 27.4 |
| **Period 2.2 (May 2006)** | 156 | 582 | 21.1 | 25.2 (27.9) | 2.5 | 12.1 | 41.1 |
| **Period 3 (May 2007)** | 67 | 733 | 8.3 | 30.1 (29.9) | 5.7 | 20 | 44.9 |

Table 2. Positive (Pos), negative (Neg) and Proportion (perc) of Aujeszky's disease positive farrow to weaning farms and within prevalence (mean, standard deviation (sd) and quartile distribution) in positive farrow to weaning farms at the end of each eradication period.

| | **Proportion of positive farrow to finish farms** | | | **Prevalence in positive farrow to finish farms (%)** | | | |
|---|---|---|---|---|---|---|---|
| **Period** | **Pos** | **Neg** | **Perc(%)** | **Mean(sd)** | **25%** | **50%** | **75%** |
| **April 2003** | 719 | 433 | 62.4 | 44.4(33.3) | 13 | 36.9 | 73.6 |
| **Period 1 (May 2004)** | 717 | 441 | 61.9 | 34.4(31.1) | 7.1 | 24.6 | 56.3 |
| **Period 2.1 (May 2005)** | 435 | 385 | 53.1 | 24.8(26.2) | 4.2 | 13.3 | 37.5 |
| **Period 2.2 (May 2006)** | 221 | 653 | 25.3 | 26.1(29.6) | 3.1 | 11.4 | 45.1 |
| **Period 3 (May 2007)** | 113 | 940 | 10.7 | 37.4(31.7) | 10 | 26.7 | 61.5 |

Table 3. Positive (Pos), negative (Neg) and Proportion (perc) of Aujeszky's disease positive farrow to finish farms and within prevalence (mean, standard deviation (sd) and quartile distribution), in positive farrow to finish farms at the end of each eradication period.

In relation to **purely fattening** units, during the first period of study, 598 out of 3,495 (17.11%) had serological positive animals. This proportion was of 7.1% (126 out of 1,770) during period 2.1; 9.4% (351 out of 3,734) during period 2.2 and 8.5% (341 out of 4,001) during period 3.

### 4.1.3. Geographical factors included in the analysis

In figure 7 is shown the distribution of pig slaughterhouses and conventional roads included in the analysis.

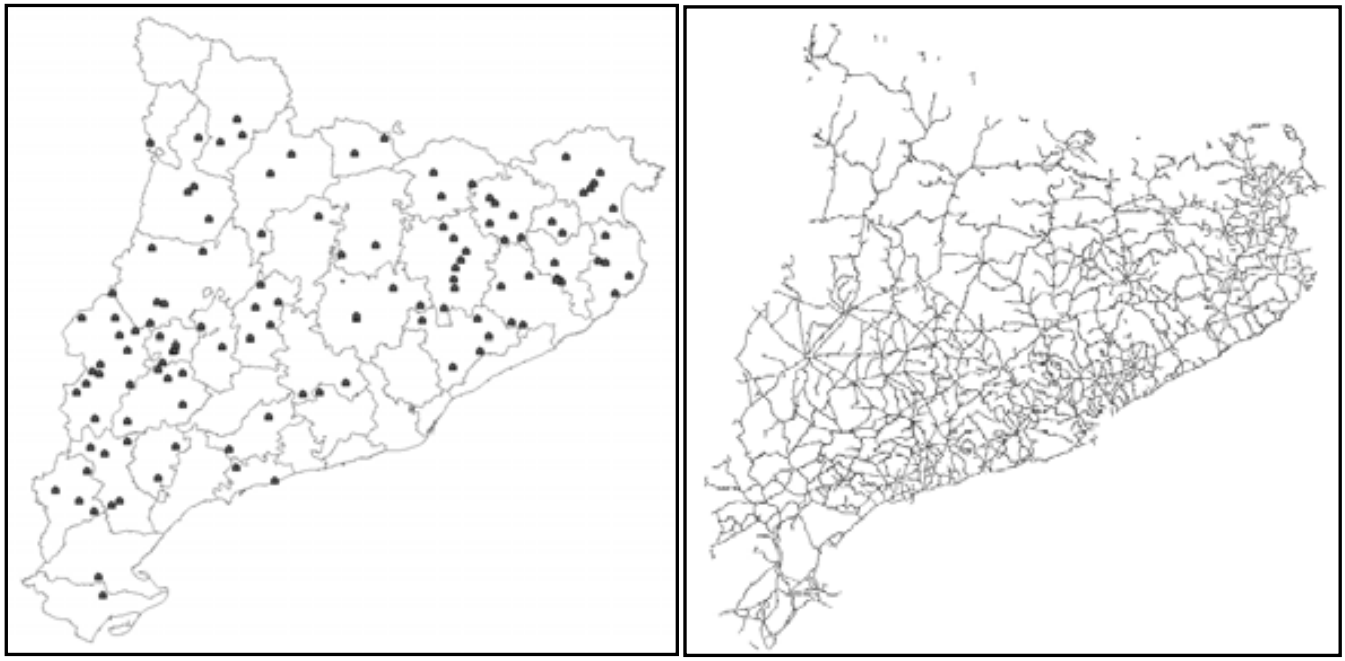

Figure 7. Distribution of pig slaughterhouses (left) and conventional roads (right) present in Catalonia and included in the analysis.

Most of the sow farms (75%) in the area of study had a **pig slaughterhouse** and/or a **conventional road** within a distance of 9km and 1km respectively (table 4).

| | Mean(sd) | Min | 25% | 50% | 75% | Max |
|---|---|---|---|---|---|---|
| **Distance to slaughterhouse** | 6,294 (4,268) | 73 | 3,070 | 5,341 | 8,632 | 36,434 |
| **Distance to conventional road** | 721 (817) | 1 | 175 | 457 | 955 | 6,086 |

Table 4. Mean, standard deviation (sd), minimum (Min), maximum (Max) and quartile distribution of the distance (in meters) from sow farms to the nearest slaughterhouse and nearest conventional road.

At the end of the first period; 912 out of the 1,194 positive sow farms (76.4%) had positive sows in their neighbourhood (750 meters radius buffer), 720 (60.3%) had positive fatteners and 629 (52.7%) had both sows and fatteners. This proportion was of 27.2% (219 out of 804) at the end of period 2.1; of 36.9% (139 out of 377) at the end of period 2.2 and decreases to 20.5% (37 out of 180) at the end of period 3. The number of positive sow and fattening animals in the neighbourhood of **positive** sow farms at the end of each eradication period is represented in tables 5 and 6.

| | Positive sow farms | Positive sow farms with positive sows in the neighbourhood | Perc (%) | Mean(sd) | 25% | 50% | 75% |
|---|---|---|---|---|---|---|---|
| **April 2003 May 2004** | 1,194 | 912 | 76.4 | 56.9(93.3) | 13 | 37 | 66 |
| **June2004 May 2005** | 804 | 476 | 59.2 | 46.5(78.8) | 11 | 26 | 49 |
| **June2004 May 2006** | 377 | 225 | 59.7 | 53.2(95.4) | 7 | 19 | 52 |
| **June2006 May 2007** | 180 | 88 | 48.9 | 42.2(82.9) | 3 | 20.5 | 54.5 |

Table 5. Percentage (Perc) of positive sow farms with positive sows in the neighbourhood and mean, standard deviation (sd) and quartile distribution of Aujeszky's disease serological positive sows in the neighbourhood of positive sow farms during each eradication period.

| | Positive sow farms | Positive sow farms with positive fatteners in the neighbourhood | Perc (%) | Mean(sd) | 25% | 50% | 75% |
|---|---|---|---|---|---|---|---|
| **April 2003 May 2004** | 1,194 | 720 | 60.3 | 438.4(454.9) | 112.5 | 378.5 | 598.5 |
| **June2004 May 2005** | 804 | 257 | 31.9 | 382.6(545.2) | 138 | 240 | 486 |
| **June2004 May 2006** | 377 | 200 | 53.1 | 615(517.7) | 203 | 538.5 | 846 |
| **June2006 May 2007** | 180 | 52 | 28.8 | 150.1(124.2) | 11.5 | 148 | 261 |

Table 6. Percentage (Perc) of positive sow farms with positive fatteners in the neighbourhood and mean, standard deviation (sd) and quartile distribution of Aujeszky's disease serological positive fattening animals in the neighbourhood of positive sow farms during each eradication period.

## 4.2. Exploratory analysis

### 4.2.1. Cluster analysis

***a) Positive sow farms:***

Using spatial scan statistics, we identified clusters of Aujeszky's disease positive sow farms (farrow to weaning and farrow to finish) throughout the study period. The prevalence ratio estimates and areas (square kilometres) of the different clusters are presented in Table 7. The prevalence ratio values for the clusters were between 1.56 and 1.73 at the beginning of the study (April 2003) and increased as the study progressed to reach values between 5.15 and 5.48 for the last period (June 2007).

| Period | Clusters | Population | Observed | P-value | Area ($Km^2$) | Prevalence ratio |
|---|---|---|---|---|---|---|
| **April 2003** | 0Sa | 405 | 341 | 0.001 | 39,806 | 1.61 |
| | 0Sb | 95 | 84 | 0.001 | 676 | 1.56 |
| | 0Sc | 53 | 49 | 0.001 | 34 | 1.62 |
| | 0Sd | 24 | 24 | 0.008 | 12 | 1.73 |
| **May 2004** | 1Sa | 242 | 202 | 0.001 | 223 | 1.58 |
| | 1Sb | 476 | 347 | 0.001 | 31,774 | 1.41 |
| | 1Sc | 21 | 21 | 0.018 | 7.8 | 1.79 |
| **May 2005** | 2.1Sa | 267 | 194 | 0.001 | 329 | 1.76 |
| | 2.1Sb | 171 | 133 | 0.001 | 1,974 | 1.82 |
| **June 2006** | 2.2Sa | 171 | 87 | 0.001 | 1,354 | 3.10 |
| | 2.2Sb | 127 | 70 | 0.001 | 1,454 | 3.24 |
| | 2.2Sc | 33 | 24 | 0.001 | 8.8 | 3.88 |
| **June 2007** | 3Sa | 116 | 45 | 0.001 | 745 | 5.15 |
| | 3Sb | 78 | 34 | 0.001 | 903 | 5.48 |

Table 7. Results of the spatial scan statistic analyses on sow farms (farrow to weaning and farrow to finish) with the Bernoulli model and a maximum scanning window of 50% of the population at risk. Population and observed columns indicate the total number of sow farms and number of Aujeszky's disease positive sow farms located inside the cluster respectively. The first number of the code of the cluster corresponds to the period, the capital letter indicates the type of cluster analysis and the lower case letter indicates the cluster detected by the spatial scan statistic, where letter 'a' represents the most likely cluster and 'b', 'c' and 'd' the secondary ones.

***b) Positive fattening farms:***

Comparing the farms that had at least one positive animal with those without positive animals, two significant clusters were identified in the first period, one in the second, two in the third and two in the fourth one. As with sow farms, the ranges of prevalence ratios increased throughout the study period from 1.73 to 2.45 in the first period to 2.77 to 5.52 in the last period (Table 8).

| Period | Cluster | Population | Observed | P-value | Prevalence ratio |
|---|---|---|---|---|---|
| **April 2003 May 2004** | 1Fa | 1,260 | 531 | 0.001 | 1.73 |
| | 1Fb | 46 | 34 | 0.001 | 2.45 |
| **June 2004 May 2005** | 2.1F | 874 | 198 | 0.001 | 3.44 |
| **June 2005 May 2006** | 2.2Fa | 591 | 237 | 0.001 | 3.64 |
| | 2.2Fb | 138 | 66 | 0.001 | 3.28 |
| **June 2006 May 2007** | 3Fa | 1,161 | 181 | 0.001 | 2.77 |
| | 3Fb | 24 | 11 | 0.039 | 5.52 |

Table 8. Results of the spatial scan statistic analyses on fattening farms with the Bernoulli model and a maximum scanning window of 50% of the population at risk. Population and observed columns indicate the total number of fattening farms and Aujeszky's disease positive ones respectively inside the cluster. The first number of the code of the cluster corresponds to the period, the capital letter indicates the type of cluster analysis and the lower case letter indicates the cluster detected by the spatial scan statistic, where letter 'a' represents the most likely cluster and 'b' the secondary cluster.

***c) Elimination analysis* (based on sow farms that became negative):**

Significant clusters were identified in the four study periods, which indicate that the proportion of positive farms that eliminated the infection was higher in some areas. The relative risk values during the different periods were similar, ranging from 1.54 to 2.83 (Table 9a).

***d) Reinfection analysis* (based on sow farms that became positive):**

The spatial scan statistic identified areas where the proportion of farms that became positive was higher in the four study periods. The relative risk values of these clusters were higher than the values of other cluster types, ranging from 3.14 to 18.55. There was an increase in the relative risk values in period 4 compared to the previous ones (Table 9b).

| | Period | Cluster | Population | Observed | P-value | Relative Risk | Area ($Km^2$) |
|---|---|---|---|---|---|---|---|
| a) | **April 2003 May 2004** | 1Ea | 69 | 29 | 0.014 | 2.43 | 392 |
| | | 1Eb | 40 | 20 | 0.025 | 2.83 | 69 |
| | **June 2004 May 2005** | 2.1Ea | 178 | 95 | 0.001 | 1.89 | 1,228 |
| | | 2.1Eb | 30 | 23 | 0.005 | 2.42 | 549 |
| | **June 2005 May 2006** | 2.2E | 269 | 209 | 0.001 | 1.54 | 5,188 |
| | **June 2006 May 2007** | 3E | 144 | 121 | 0.001 | 1.71 | 5,241 |
| b) | **April 2003 May 2004** | 1R | 212 | 126 | 0.001 | 5.41 | 225 |
| | **June 2004 May 2005** | 2.1Ra | 9 | 9 | 0.001 | 5.98 | 6 |
| | | 2.1Rb | 50 | 24 | 0.001 | 3.14 | 963 |
| | **June 2005 May 2006** | 2.2R | 173 | 28 | 0.001 | 5.29 | 2,215 |
| | **June 2006 May 2007** | 3R | 18 | 8 | 0.001 | 18.55 | 239 |

Table 9. Results of the spatial scan statistic analyses to search for clusters of elimination (a) and clusters of re-infection (b), with the Bernoulli model and a maximum scanning window of 50% of the population at risk. The first number of the code of the cluster corresponds to the period, the capital letter indicates the type of cluster analysis and the lower case letter indicates the cluster detected by the spatial scan statistic, where letter 'a' represents the most likely cluster and 'b' the secondary cluster.

## 4.2.2. Geographical location of clusters

The location of clusters from the different analyses and periods are represented in Figures 8, 9, 10 and 11. Three different sized clusters were identified in the study of positive sow farms in the first period in the central (1Sa) and western (1Sb, 1Sc) parts of the region (Fig.8). Clusters 1Sa and 1Sb overlapped with clusters of sow, fattening and reinfection in the same or in the following periods (Fig.8, 9, 10 and 11), in contrast to 1Sc where clusters of sows, fattening and reinfection were not identified in the same or the following periods (Fig.8, 9, 10 and 11).

For the second period, clusters of fattening farms (2.1F) were identified in the north eastern part of the region. During the following periods, clusters of sow farms (2.2Sb, 3Sb), fattening farms (2.2Fb, 3Fb) and reinfection in sow farms (2.2R, 3R) were identified in this area (Fig. 9, 10 and 11).

We identified areas where the proportion of sow farms that became negative (elimination) was higher compared to the whole region throughout the four study periods. During the first period these areas were located in the western (1Ea) and central (1Eb) parts of the region (Fig. 8). In the second period they were located in the western (2.1Eb) and central parts (2.1Ea) (Fig.9), in the third period they moved to the central and north eastern parts (2.2E) (Fig.10), and in the fourth period to the central and western parts of the region (3E) (Fig.11).

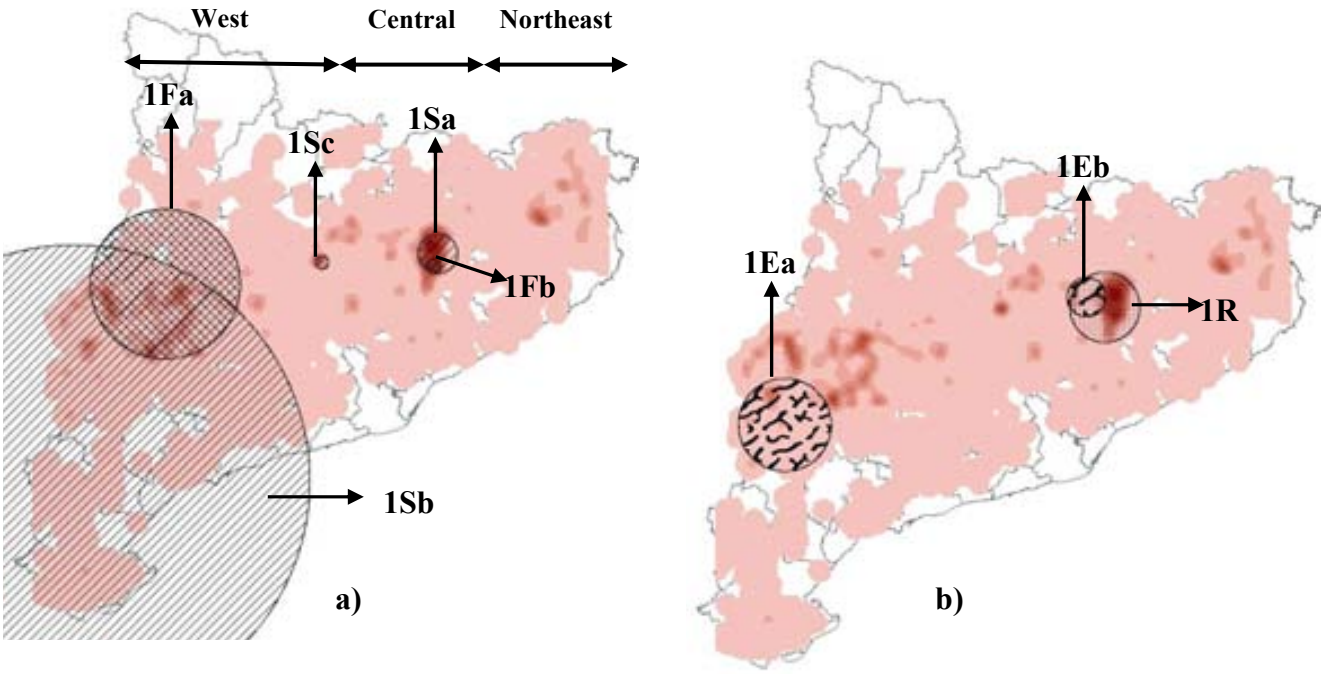


Figure 8. Clusters identified in the first period (April 2003 – May 2004) with the spatial scan statistic (Bernoulli model). a) clusters of positive sow farms (S) and fattening farms (F), b) clusters of elimination (E) and re-infection (R). Clusters are represented over a Kernell density surface of pig farms (farrow to weaning, farrow to finish and fattening). Areas with higher pig farms density are represented with higher colour intensity.

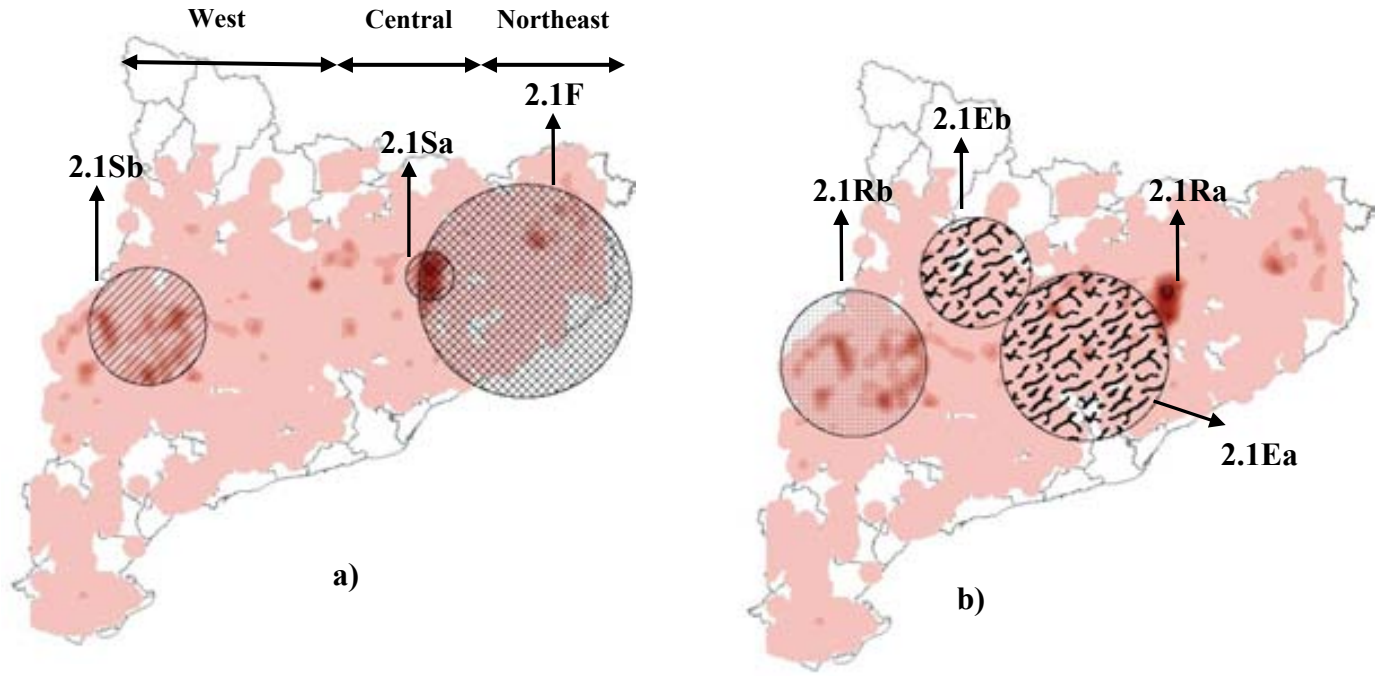


Figure 9. Clusters identified in the second period (June 2004 – May 2005) with the spatial scan statistic (Bernoulli model). a) clusters of positive sow farms (S) and fattening farms (F), b) clusters of elimination (E) and re-infection (R). Clusters are represented over a Kernell density surface of pig farms.

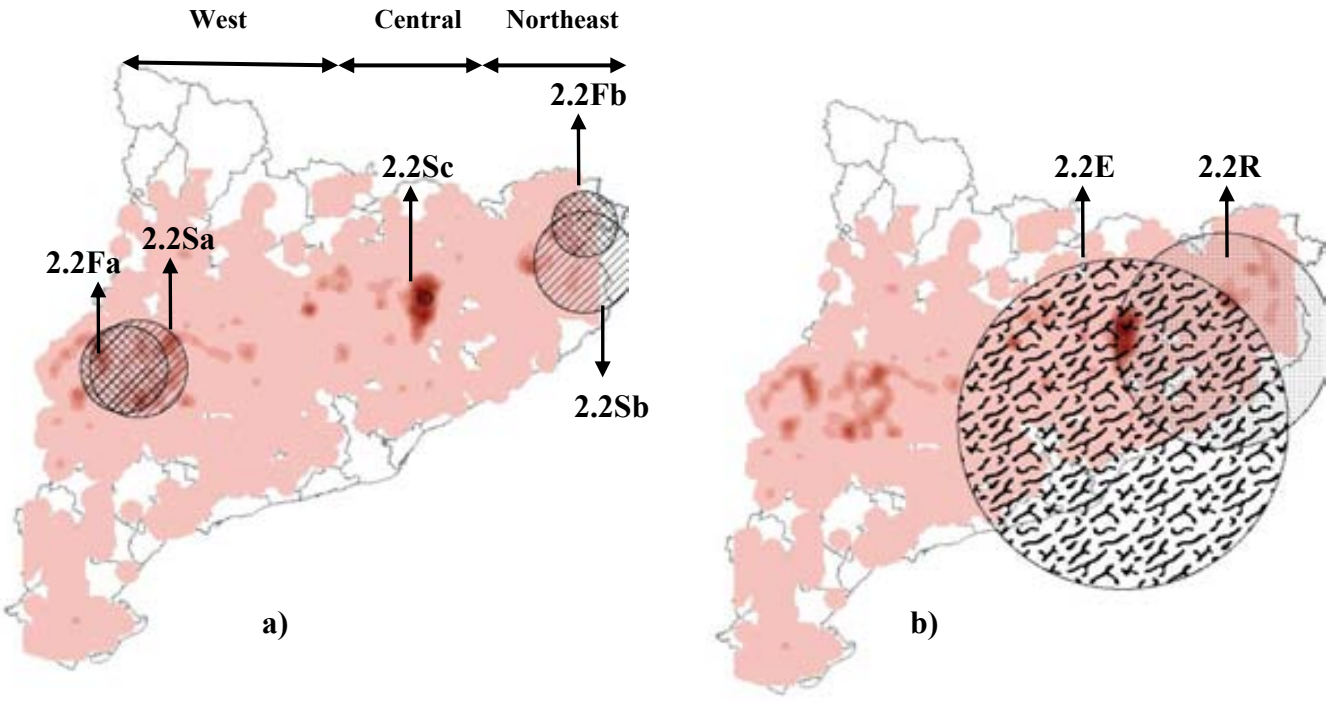


Figure 10. Clusters identified in the third period (June 2005 – May 2006) with the spatial scan statistic (Bernoulli model). a) clusters of positive sow farms (S) and fattening farms (F), b) clusters of elimintion (E) and re-infection (R). Clusters are represented over a Kernell density surface of pig farms.

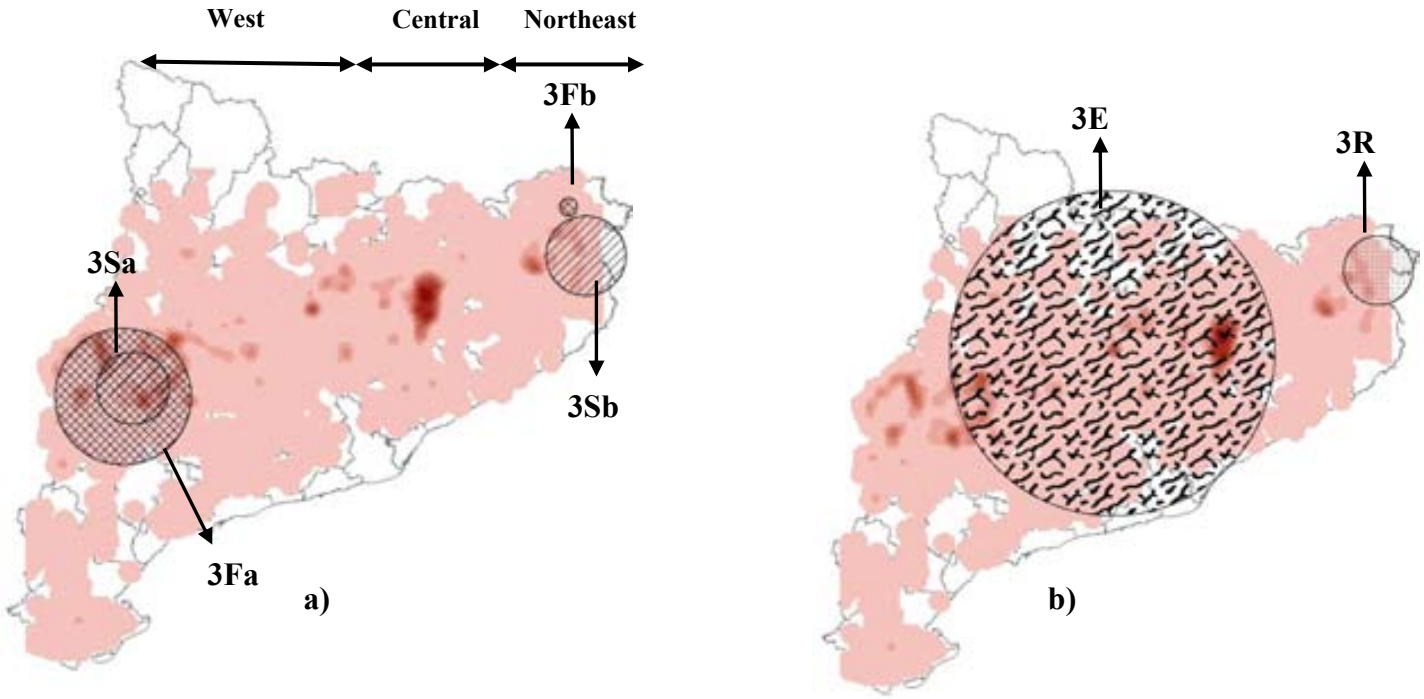


Figure 11. Clusters identified in the fourth period (June 2006 – May 2007) with the spatial scan statistic (Bernoulli model). a) clusters of positive sow farms (S) and fattening farms (F), b) clusters of elimination (E) and re-infection (R). Clusters are represented over a Kernell density surface of pig farms.

In central, western and north eastern Catalonia there was a geographical association between the clusters of positive sow farms, positive fattening farms and re-infected sow farms, as indicated in figure 12. In areas where clusters of positive sow or fattening farms were detected, we identified clusters of sow farms that became positive in the same or the following period.

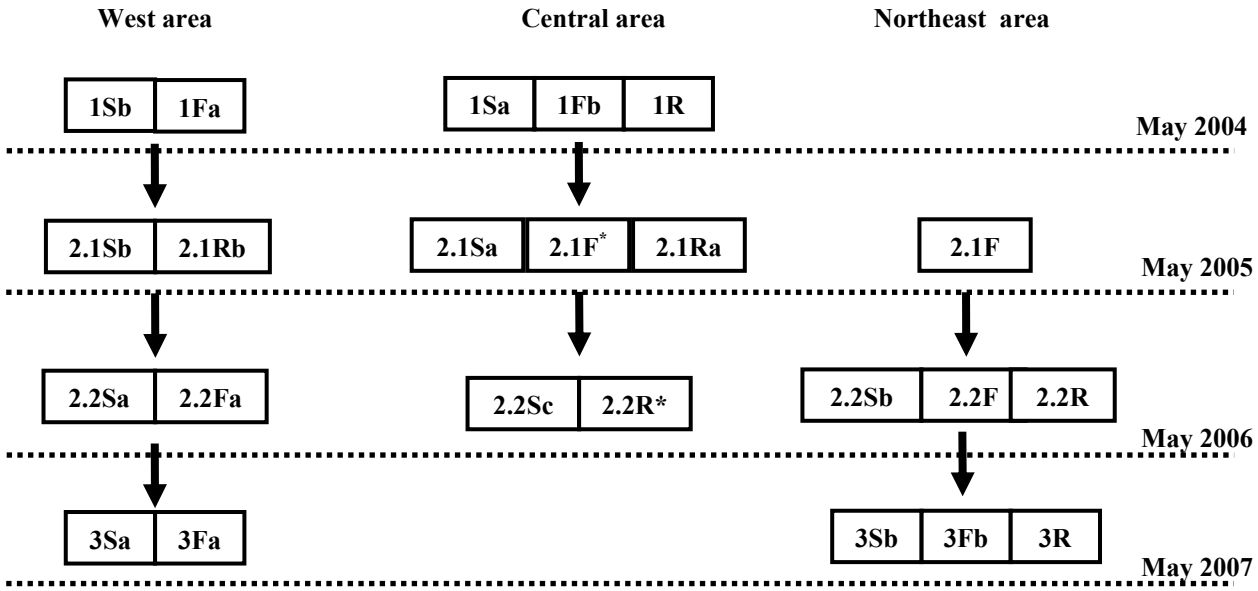


Figure 12. Geograhical overlap of clusters identified during the four periods of study (see the text to interpret the clusters notation).*These clusters covered an area of central and northeastern part of the region.

## 4.2.3. Pig farm density and proportion of serologically positive fattening farms in reinfection and elimination clusters

The mean density of pig farms was 0.40 farms per square Km (median of 0.29 and standard deviation of 0.33) in clusters where more sow farms became negative (elimination) and 1.51 (median of 0.70 and standard deviation of 1.62) in clusters where more sow farms became positive (reinfection). There was a statistically significant difference between them (p value<0.05). The proportion of positive fattening farms was higher in those areas where more sow farms became positive (Table 10).

| | Clusters of reinfection | | Clusters of elimination | | OR (CI: 95%) |
|---|---|---|---|---|---|
| | Positive fattening farms | Fattening farms | Positive fattening farms | Fattening farms | |
| **April 2003 May 2004** | 67 | 348 | 40 | 303 | 1.57 (1.02 – 2.4)* |
| **June 2004 May 2005** | 9 | 154 | 2 | 170 | 5.28 (1.1 – 24.8)* |
| **June 2005 May 2006** | 61 | 1011 | 46 | 1333 | 1.86 (1.3 – 2.8)* |
| **June 2006 May 2007** | 29 | 162 | 103 | 1559 | 3.08 (1.9-4.8)* |

Table 10. Percentage of fattening farms with AD serological positive animals in clusters of reinfection and eradication. (*) indicates that differences are statistically significant (p-value<0.05).

## 4.3. Modelling

### 4.3.1. Parameter results

The hierarchical bayesian binomial model, suggested that the number of positive fattening animals in the neighbourhood of each sow farm, increases the probability of being AD positive. On the other hand, the number of positive sows in the neighbourhood did not influence significantly on the likelihood of being AD positive in the different periods (i.e. the bayesian credible interval of the regression parameter for this covariate conatined the zero). The type of farm (farrow to weaning or farrow to finish) also did not influence the probability of being AD positive.

In period 1, positive fattening pigs in the neighbourhood (750 meters radius) influenced positively the probability of being AD positive: the 95% bayesian credible interval of the regression parameter for this variable is 0.005 - 0.31. i.e.: for 1,000 increases in number of positive fattening animals in the neighbourhood of each sow farm the odds of being AD positive increases by a factor between 1.005 and 1.36 (logarithmic inverse of 0.005 and 0.31).

In period 2.2, having positive fattening animals in the neighbourhood increased the likelihood of each sow farm to be AD positive between 1.84 and 3.22 (logarithmic inverse of 0.61 – 1.17). In period 2.1 and period 3, any variable had a positive relation with the probability of being positive.

For each period, the estimate values, monte carlo error (MC error) and percentile distribution for the bayesian credible intervals of the intercept ($\beta_0$) and the fixed regression parameters of the type of farm ($\beta_1$), positive sows in the neighbourhood ($\beta_2$) and positive fattening in the neighbourhood ($\beta_3$), are presented in table 11.

| Model ( logit ($p_i$) ) | Estimate | OR | MC error | Percentile distribution | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | **5%** | **10%** | **median** | **90%** | **95%** |
| **Period 1 (May 2004)** | $\beta_0$: -8.45 | | 0.0041 | -8.61 | -8.58 | -8.46 | -8.32 | -8.28 |
| | $\beta_1$: -0.03 | | 0.0012 | -0.13 | -0.11 | -0.02 | 0.03 | 0.05 |
| | $\beta_2$: 0.03 | | 0.0046 | -0.35 | -0.21 | 0.01 | 0.32 | 0.48 |
| | **$\beta_3$: 0.15** | **1.61*** | 0.0027 | **0.005** | 0.03 | 0.15 | 0.27 | **0.31** |
| **Period 2.1 (May 2005)** | $\beta_0$: -9.95 | | 0.0121 | -10.38 | -10.32 | -9.92 | -9.61 | -9.55 |
| | $\beta_1$: 0.02 | | 0.0012 | -0.07 | -0.05 | 0.02 | 0.11 | 0.15 |
| | $\beta_2$ : -0.15 | | 0.0031 | -0.34 | -0.29 | -0.13 | -0.02 | 0.01 |
| | $\beta_3$: 0.008 | | 0.0005 | -0.02 | -0.01 | 0.01 | 0.03 | 0.04 |
| **Period 2.2 (May 2006)** | $\beta_0$: -14.56 | | 0.0177 | -15.14 | 15.05 | -14.55 | -14.07 | -13.99 |
| | $\beta_1$: 0.14 | | 0.0033 | -0.03 | -0.01 | 0.12 | 0.32 | 0.37 |
| | $\beta_2$: -0.45 | | 0.0179 | -2.15 | -1.61 | -0.11 | 0.11 | 0.22 |
| | **$\beta_3$: 0.89** | **2.43*** | 0.0057 | **0.61** | 0.67 | 0.89 | 1.11 | **1.17** |
| **Period 3 (May 2007)** | $\beta_0$: -33.1 | | 0.6165 | -48.72 | -48.0 | -28.95 | -16.62 | -16.09 |
| | $\beta_1$: -0.01 | | 0.0175 | -0.55 | -0.26 | -0.0003 | 0.04 | 0.52 |
| | $\beta_2$: -0.35 | | 0.0869 | -3.42 | -1.29 | -0.004 | 0.38 | 1.13 |
| | $\beta_3$: 0.51 | | 0.0825 | -1.01 | -0.36 | 0.007 | 1.96 | 4.47 |

Table 11. Estimate value, odds ratio (OR), monte carlo error (MC error) and percentile distribution for the 95% bayesian credible intervals of the intercept ($\beta_0$) and the fixed regression parameters of the type of farm ($\beta_1$), positive sows in the neighbourhood ($\beta_2$) and positive fattening in the neighbourhood ($\beta_3$) of the binomial model (logit ($p_i$)). In bold are represented those parameters which influence the probability of being AD positive. *Interpretation: for 1,000 increases in number of positive fattening animals in the neighbourhood of each sow farm the odds of being AD positive increases by that factor.

As mentioned before, we assumed that the different slaughterhouses and roads influence '$p_i$' differently (the number of transported and slaughtered pigs is not homogenous on the study region) and therefore the regression parameters for these covariates were individualised. They were introduced in the model as random effects.
All the bayesian credible intervals estimated by the binomial model, of these individualised regression parameters, contained the value zero. Therefore, the distance to the nearest slaughterhose and conventional road did not significantly influenced the probability of being AD positive in any period.

The estimate of the standard deviation, MC error and percentile distribution of the bayesian credible intervals of the regression parameters of the explanatory variables included as random effects (distance to nearest slaughterhouse ($\beta_4$) and distance to nearest road ($\beta_5$) and the random effects for the uncorrelated heterogeneity (a) and correlated heterogeneity (s) are shown in table 12.

| Model ( logit ($p_i$) ) | Estimate | MC error | Percentile distribution | | |
|---|---|---|---|---|---|
| | | | **5%** | **median** | **95%** |
| **Period 1 (May 2004)** | $\beta_4$: 0.11 | 0.0030 | 0.02 | 0.11 | 0.23 |
| | $\beta_5$: 0.12 | 0.0040 | 0.03 | 0.11 | 0.28 |
| | $a$: 0.07 | 0.0016 | 0.02 | 0.06 | 0.13 |
| | $s$: 0.06 | 0.0015 | 0.03 | 0.06 | 0.13 |
| **Period 2.1 (May 2005)** | $\beta_4$: 0.03 | 0.0005 | 0.01 | 0.03 | 0.05 |
| | $\beta_5$: 0.09 | 0.0027 | 0.02 | 0.07 | 0.21 |
| | $a$: 0.06 | 0.0017 | 0.02 | 0.05 | 0.13 |
| | $s$: 0.16 | 0.0052 | 0.03 | 0.14 | 0.38 |
| **Period 2.2 (May 2006)** | $\beta_4$: 0.05 | 0.0014 | 0.02 | 0.04 | 0.12 |
| | $\beta_5$: 0.15 | 0.0083 | 0.03 | 0.10 | 0.45 |
| | $a$: 0.10 | 0.0031 | 0.03 | 0.08 | 0.45 |
| | $s$: 0.11 | 0.0043 | 0.02 | 0.08 | 0.31 |
| **Period 3 (May 2007)** | $\beta_4$: 0.38 | 0.0384 | 0.01 | 0.13 | 1.66 |
| | $\beta_5$: 0.21 | 0.0146 | 0.01 | 0.06 | 0.92 |
| | $a$: 12.29 | 0.2928 | 4.02 | 11.18 | 19.96 |
| | $s$: 0.12 | 0.0082 | 0.01 | 0.06 | 0.40 |

Table 12. Estimate of the standard deviation, monte carlo error (MC error) and percentile distribution of the bayesian credible intervals of the regression parameters of the explanatory variables included as random effects (distance to nearest slaughterhouse ($\beta_4$) and distance to nearest road ($\beta_5$)) and the random effects for the uncorrelated heterogeneity (a) and correlated heterogeneity (s) for each period.

The residuals were useful to evidence those areas where the probability of each sow farm to be infected was less explained by the covariates included in the model. In figures 13 to 17, the predicted probability of infection by the model, the observed infected sow farms and the residuals of the model for each period are represented. Most of the areas were the probability of being infected was high was not predicted by the model. Because of that, the geographic distribution of the observed infected farms is quite similar to the distribution of the residuals of the model. The predictive value of the model was quite higher in the west part of the region as compare to the rest of the region.

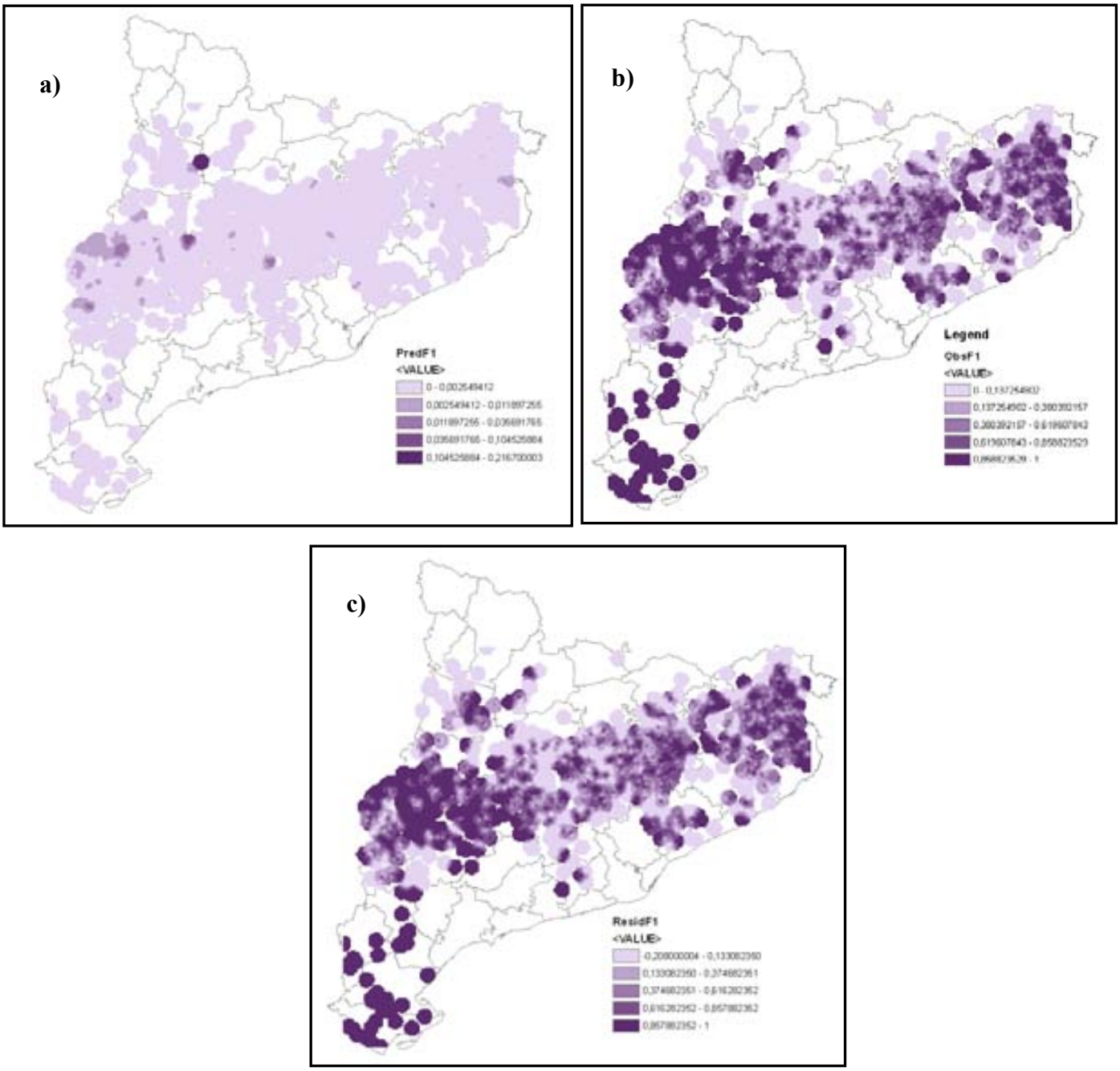


Figure 13. Predicted probability of AD infection in sow farms by the hierarchical Bayesian binomial model (a), observed infected sow farms (b) and residuals of the model (c); represented with the inverse distance weighting function (fixed bandwidth of 4km) for period 1 (May 2004).

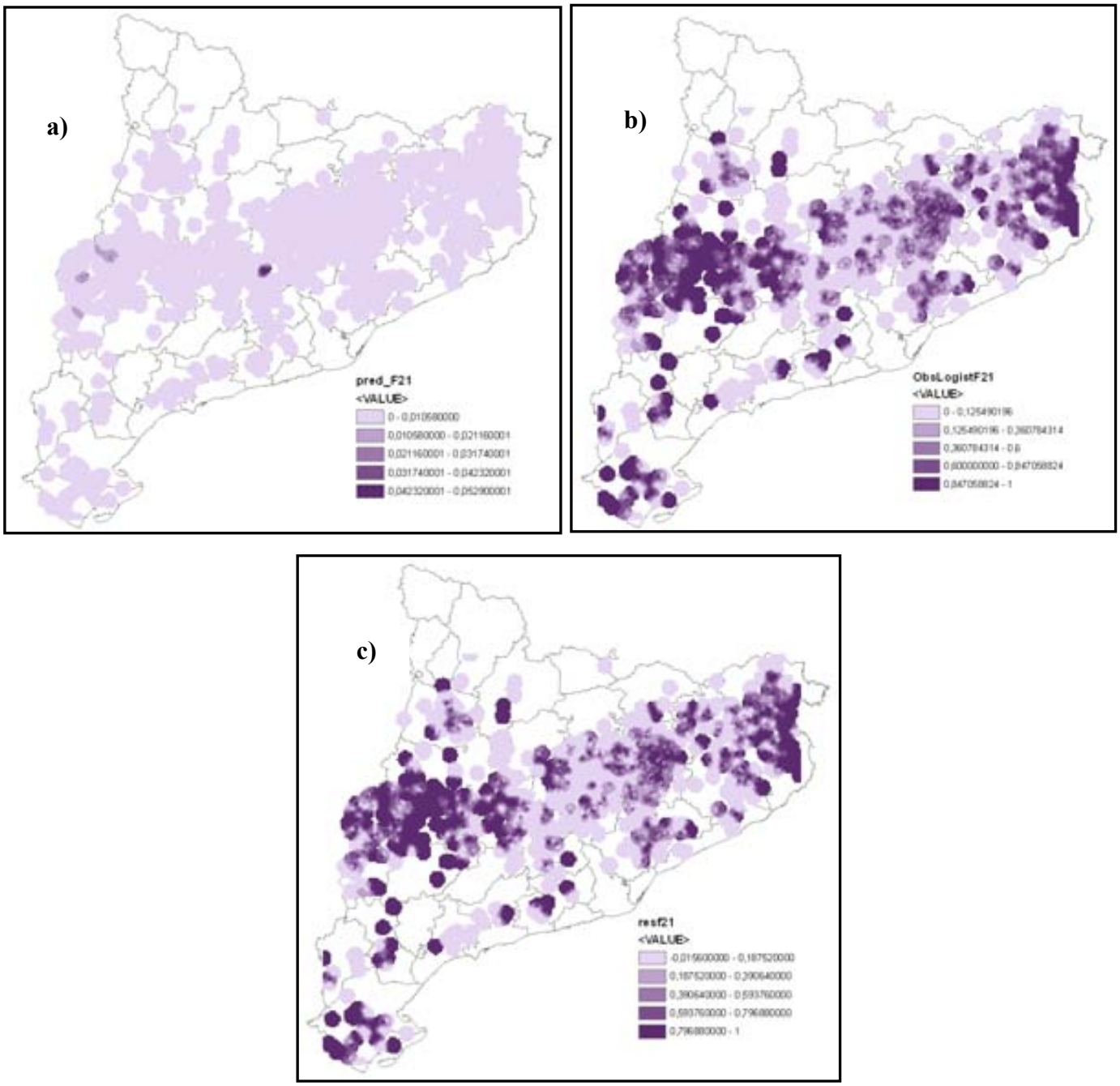


Figure 14. Predicted probability of AD infection in sow farms by the hierarchical Bayesian binomial model (a), observed infected sow farms (b) and residuals of the model (c); represented with the inverse distance weighting function (fixed bandwidth of 4km) for period 2.1 (May 2005).

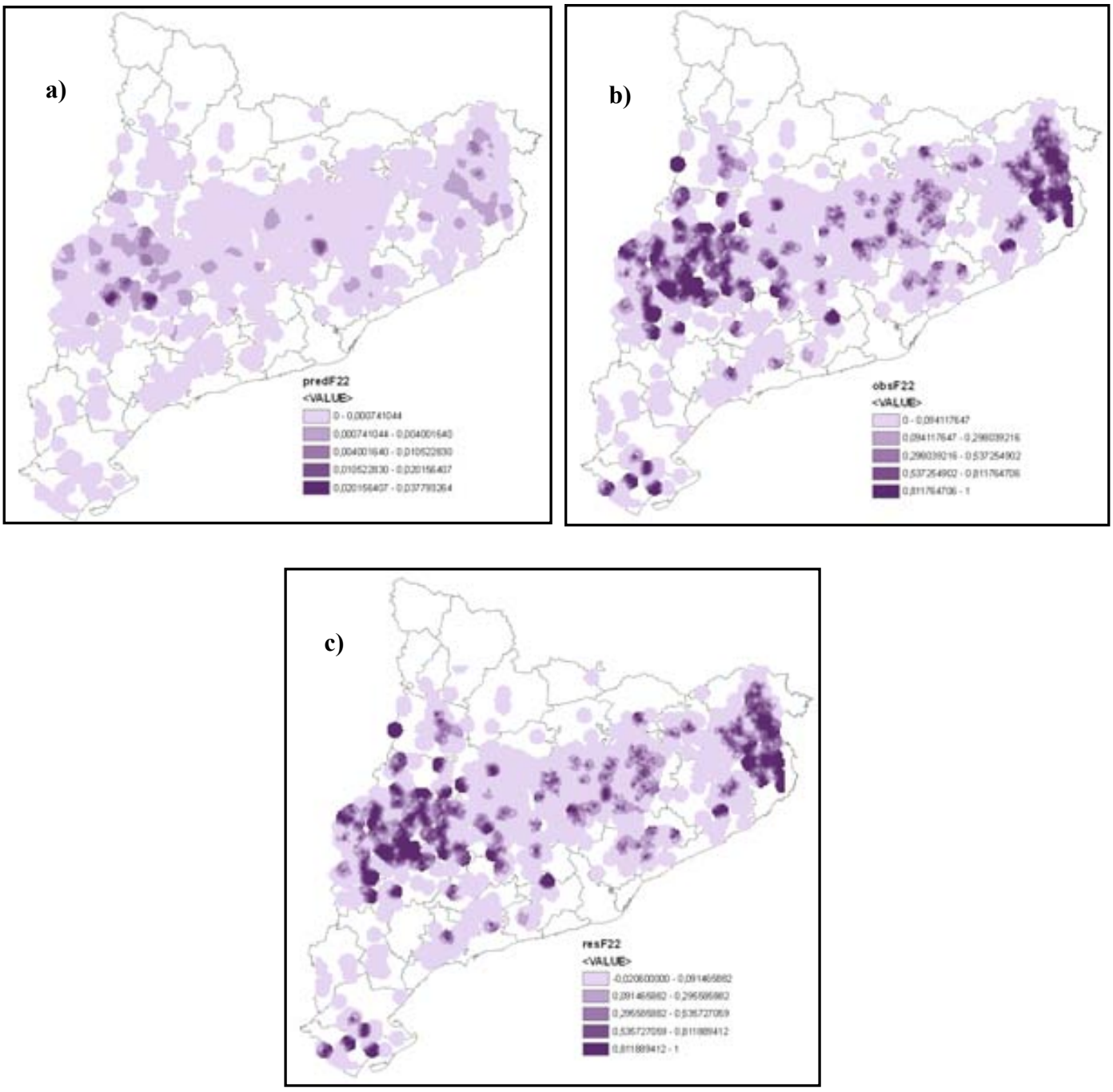


Figure 15. Predicted probability of AD infection in sow farms by the hierarchical Bayesian binomial model (a), observed infected sow farms (b) and residuals of the model (c); represented with the inverse distance weighting function (fixed bandwidth of 4km) for period 2.2 (May 2006).

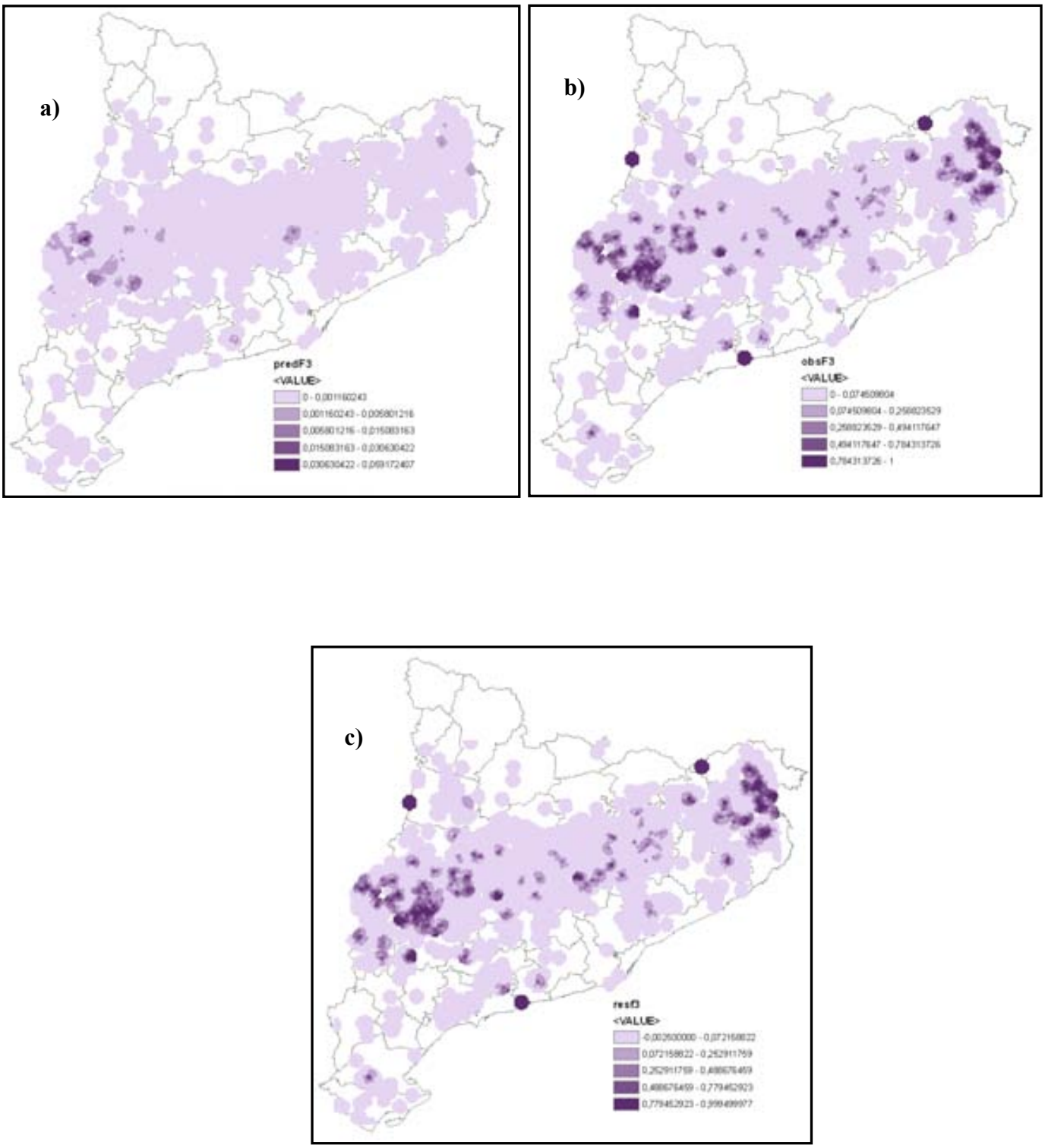


Figure 16. Predicted probability of AD infection in sow farms by the hierarchical Bayesian binomial model (a), observed infected sow farms (b) and residuals of the model (c); represented with the inverse distance weighting function (fixed bandwidth of 4km) for period 3 (May 2007).

## 4.3.2. Convergence diagnostic

As mentioned in the introduction chapter, when Bayesian methods are employed, the parameter or parameters of interest are described by the posterior distribution and therefore, this posterior distribution must converge to a stationary distribution. Once convergence reached, samples should look like a random scatter about a stable mean value. It is also possible to check the convergence using the Gelman-Rubin convergence diagnostic implemented in WinBUGS; at convergence is close to 1.

In figures 17 to 20, trace plots for the deviance of the models and the Gelman-Rubin convergence diagnostic for each period are represented. Convergence has been reached in all the priods; as trace plots look like a random scatter about a stable mean value and the Gelman-Rubin convergence diagnostic is close to 1.

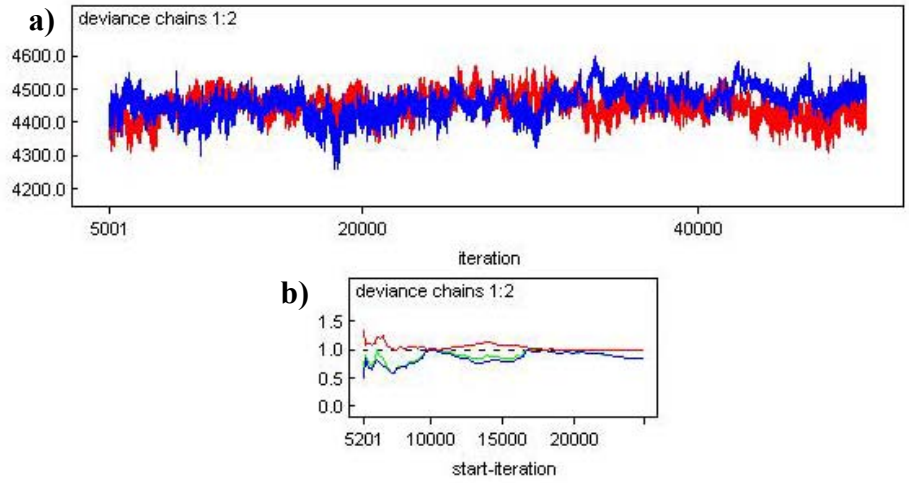


Figure 17. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 1 (May 2004).

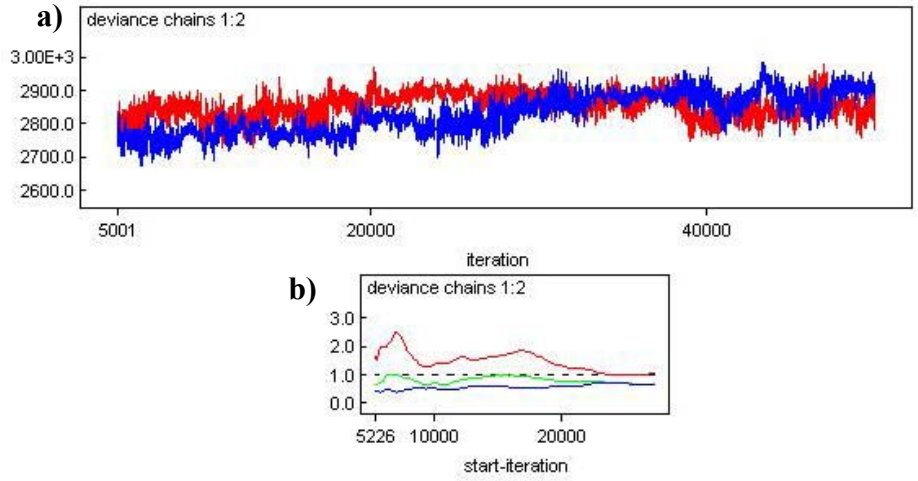


Figure 18. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 2.1 (May 2005).

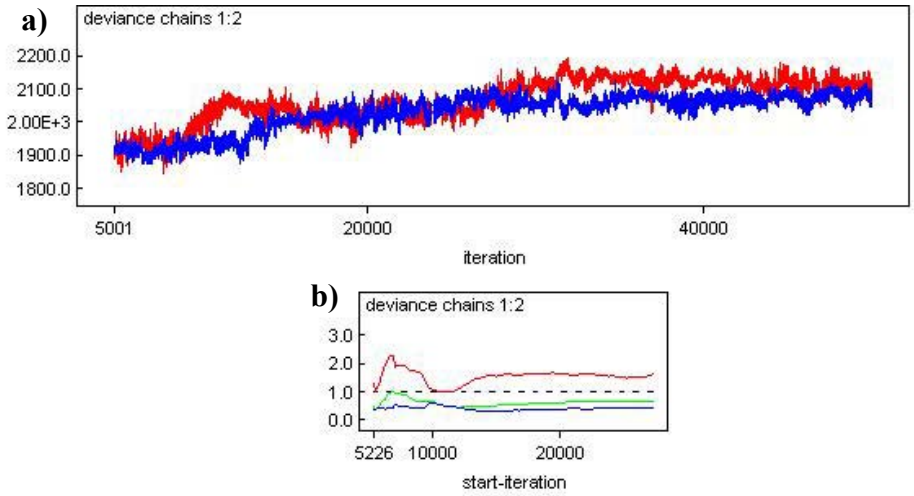


Figure 19. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 2.2 (May 2006).

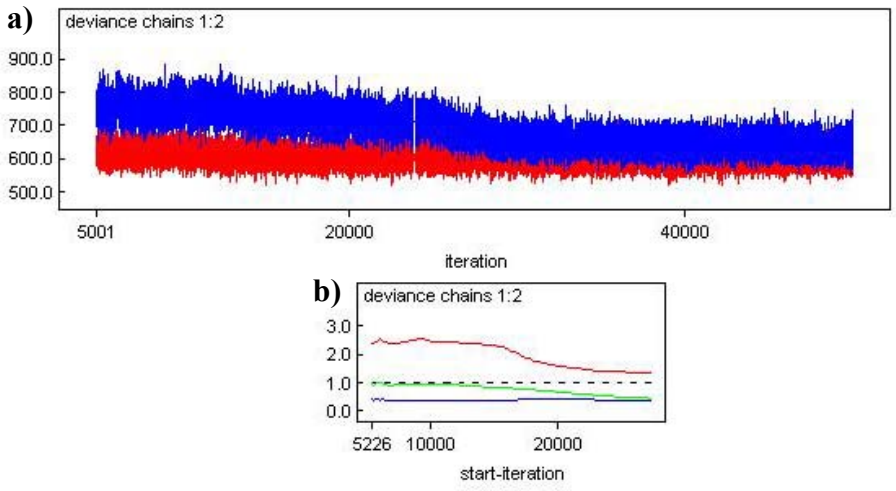


Figure 20. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 3 (May 2007).

# 5. DISCUSSION

AD eradication programmes started in the UE in the 80`s. Since then, the number of countries free of the disease has been increasing. In the European Union, Finland, Austria, Cyprus, Czech Republic, Germany, Denmark, Luxembourg, Slovakia, England, Scotland, Wales, Sweden and France are free of AD and vaccination is prohibited. There is an approved eradication programme in Belgium, Netherlands, Bolzano province in Italy and some provinces of Spain (Anonymous, 2008).

Despite the increase in the use of spatial epidemiological analysis methods in the last years, a spatial analysis of the evolution and progress of Aujeszky's Disease eradication programme during different years, to our knowledge, has not yet been carried out. Most of the papers reporting the progress in the eradication and control programmes are mainly descriptive and aimed at documenting the evolution of AD in different countries, the key points of the eradication programmes and to discuss the effectiveness of the different measures applied such as mass vaccination, stamping out, test and removal and continuous serological monitoring (Auvigne and Bougeard, 2000; Elbers et al., 2000; Martini et al., 2003; Müller at al., 2003; Robertson and Wierup 2000; Taft, 2000; Vannier et al., 2000).

The main goal of this study is to conduct a spatial analysis of the AD eradication programme in Catalonia from 2003 to 2007. In the first part of the research, we conducted an **exploratory analysis** in order to explore the spatial distribution of the disease, and then we applied **spatial modelling** in order to test the role of different geographical factors in the success of the AD eradication programme.

We used data from the different samples obtained in the catalan farms from 2003 to 2007. In most of the cases, the sampling method used in the eradication campaign allowed to detect the disease given that the within-herd prevalence was higher than 5% in sow farms and 20% in fattening farms (with a 95% confidence level). Even though the sample size and the sensitivity and specificity of the ELISA could generate false positive or false negative farms, we consider that this possible misclassification would only affect a small proportion of farms and these would most probably be homogenously distributed. Therefore, this possible inaccuracy would not significantly affect the final results.

In the **exploratory analysis** we applied the spatial scan statistic to observational data of AD in consecutive periods to determine the presence of clusters of infected sow farms and infected fattening farms, as well as clusters representing the elimination of the infection and reinfection. In this part of the study, we identified areas with a significantly higher proportion of sow farms that became negative in each of the four study periods, which indicates that the eradication of the disease had a spatial component (Figures 8, 9, 10 and 11). In Italy, Martini et al., (2003) evaluated the progress of the control programme for Aujeszky's disease using a survival approach applied to gE-seropositive swine herds (farrow to weaning and farrow to finish) at the beginning of the programme. They also found significant differences in the probability of herds from different areas becoming gE-seronegative.

Moreover, the cluster analysis of positive sow and fattening farms evidenced that the geographical distribution of ADV and its evolution on Catalonia is not homogeneous. There are areas where the proportion of ADV-seropositive sow farms and fattening

farms, is higher as compared to the whole region. In the west part of the study area, clusters of ADV-seropositive sow or fattening farms were detected during the four periods of study. In the central part, clusters were detected in the three first period of study and in the northeast area in the three last periods of study (figures 8, 9, 10 and 11). The prevalence ratio values within these clusters increased throughout the study period (Table 7). This is due to the high proportion of positive farms in the whole region at the beginning of the study compared with the final period: as the number of serologically positive farms out of the clusters decreases (the denominator), the prevalence ratio value of the cluster increases.

In this exploratory study, areas with a high proportion of reinfections were located roughly in the same areas where more sow and fattening farms were positive. These zones had a higher pig farm density than areas of elimination. High pig density has been previously considered as a risk factor for AD-seropositivity by different authors (Marsh et al., 1991; Weigel et al., 1992; Stegeman et al., 1995a; Boelaert et al., 1999; Heliez et al., 2000; Zanardi et al., 2000; Tamba et al., 2002). These results suggest a potential role for the local spread of infection, and that pig density and the proportion of positive fattening farms in the area may affect the success of AD eradication programmes.

The questions that arise from these results are: Is the l**ocal spread of ADV responsible for the persistence of AD in certain areas?** If so, **which underlying factors are implicated in this local spread, and why does it occur in these areas?** These factors may be of paramount importance, and may imply the need to apply specific control measures in these areas in order to eradicate AD faster.

The spread of ADV between neighbouring farms could be attributed in some cases to airborne transmission of the virus (Christensen et al., 1990, 1993; Casal et al., 1997), but this probably only occurs under favourable circumstances (Pejsak and Truszcynski, 2006). Other factors such as indirect contact: vehicles, equipment or personnel (Tamba et al., 2002; Solymosi et al., 2004) might also play a role in local spread. It has also been suspected that wildlife may contribute to between-herd transmission. The role of wild mammals was suggested by Kirkpatrick et al., (1980) but supporting empirical evidence has not been available (Austin and Weigel, 1992).

In the **second part** of the study we tried to elucidate the role of some factors that could be implicated in local ADV transmission. We studied several geographic variables describing the possible risk factors associated to proximity: Distance to the nearest slaughterhouse, distance to conventional roads, mean number of AD serological positive sows in the neighbourhood (750 meters radius) of each sow farm and mean number of AD serological positive fattening pigs in the neighbourhood (750 meters radius) of each sow farm. The role of those factors in the local spread of ADV is reasonable both under airborne virus spread and under virus transmission by indirect contacts (carrier animals, mechanical vectors, equipment, personnel, etc). These variables were included in a hierarchical Bayesian model.

The use of hierarchical Bayesian models, allowed us to include these factors as covariates in the model. The first part of the study was mainly **exploratory** and these types of studies only involve examining the data statistically for the presence of patterns, but they cannot be used to test causal hypothesis (Pfeiffer and Jones, 2002). Moreover, the use of these models allowed us to take into account the spatial

dependence (autocorrelation) among the data; which can be included in the model as a random effect. Bayesian modelling approaches allow these random effects to be analysed exactly (Lawson and Zhou, 2005).

As the AD eradication programme progresses an increased number of sow farms and areas are being classified as AD officially free. The purpose of this measure is to reduce the transmission of ADV between farms. As the numbers of ADV-free sow farms increased, the spread of ADV by purchased pigs would decrease (Elbers et al., 2000). In such situation, topographical factors like proximity to a slaughterhouse or a road with frequent animal transport may be an uncontrolled risk factor because animals from infected areas can be slaughtered in free areas. In some areas, especially in the central part of Catalonia, there are big slaughterhouses. In the last years these facilities have been a concern for farmers and veterinarians of these areas and the role of the distance to slaughterhouses and roads in the persistence of the disease has been frequently speculated. The relationship between topographical factors such as distance to roads, rivers, lakes or forest, and the probability of AD infection has been studied by some authors. Marsh et al., (1991) did not found association between the likelihood of becoming infected and distance to nearest road. The presence of a river or lake within a 1km distance had a protective effect. However, their data covered just 41% of the pig units in the study area. More recently, Solymosi et al., (2004) found that the presence of a lake or a highway were positively associated with AD seropositivity whereas the presence of a forest had a protective effect. The authors speculate that the effect of the lake could be due to the existence of more abundant vegetation around the lake, which increases the risk through wildlife or wildlife vectors (e.g. rodents, foxes, birds), or to an increase in the possibility of airborne virus transmission because of the fog and higher humidity around the lake. On the other hand, the forest could reduce the risk of infection by decreasing airborne transmission. In our study, we did not evaluate the effect of topographical factors such as the presence of a lake, a river or a forest in the neighbourhood of each sow farm. However, distances to the nearest conventional road and to the nearest slaughterhouse were evaluated. According to our results, these variables did not increase the probability of a sow farm becoming infected in any of the study periods. This result might be influenced by the overall AD prevalence in the study region. As the number of AD positive animals in the region decreases, the possible influence of the proximity to a slaughterhouse or a road also decreases (as the number of infected slaughtered animals and infected transported pigs decreases). Our study began in 2003, when the 60% of the sow farms had serological positive animals and the mean prevalence of animals within positive farms was over 20% (table 2). In such situation, we believe that the number of transported and slaughtered positive animals was large enough in order to have detected a relation, if there was any, between these factors and the probability of a sow farm to be infected. The analysis of the data of previous years, when the disease was even more distributed over the region, could have been useful in order to analyze the effect of these geographical factors on the probability of being AD positive. However, data was not available for all the region of study before 2003, so we decided to begin the analysis on that year.

From the four geographical variables included in the model, only positive fattening animals in the neighbourhood of sow farms increased the probability of being AD positive**.** The results of the model suggested that this probability was influenced by the presence of fattening positive animals in the neighbourhood, except in period 2.1 and period 3. In period 2.1, this was more likely due to the fact that in this period we did not

have data on the serological results of many fattening farms (table 1), than to a lack of effect of positive fattening animals in the neighbourhood. Despite that in Catalonia a serological testing was implemented on purely fattening units in 2003 (beginning of period 1) it was not until 2005 that national legislation (206/2005 decree (RD)) made compulsory the serological controls on fattening units.

The lack of association found in period 3, could be attributable to the low number of positive sow farms (only 180) and to the low proportion of them that had positive fattening animals in their neighbourhood (52). These values were higher in previous periods: 1,194 positive sow farms, 720 of them with positive fattening farms in the neighbourhood in period 1, and 377 positive sow farms from which 200 had infected fattening farms in the neighbourhood in period 2.2 (table 6). However, the proportion of positive fattening units in period 2.2 and period 3 were quite similar: 9.4% during period 2.2 and 8.5% during period 3. Results from period 1 and 2.2 suggest that positive fattening animals in the neighbourhood can difficult the eradication of the disease in sow farms, as they increase the probability of AD infection. However, the fact that in period 3 the number of positive sow farms with positive fattening animals in their neighbourhood is low, as compared with previous periods, and the lack of association found in this period, suggest that local spread from neighbouring fattening farms in the neighbourhood has not a high effect in order to eradicate AD from sow farms.

On the other hand, the number of positive sows in the neighbourhood of each sow farm did not influence the probability of being AD positive in any of the periods of study. The probability of a fattening farm infecting a neighbouring sow farm seems to be higher than the probability of a sow farm infecting a neighbouring sow farm. This result could be related to herd size and density within finishing herds. In the model developed by Van Nes et al., (1998) herd size appeared to be an important risk factor in between-herd transmission of ADV. However, Bouma et al., (1995), concluded that ADV transmission was not influenced by the size of the population, but was more likely related to the density of pigs. Density of pigs within the herd, determines that the reproduction ratio (the number of units infected by one infectious unit) within a finishing herd is higher than in a sow herd. From field observations, the reproduction ratio within a finishing herd was estimated for groups of twice-vaccinated pigs as 1.5 and for pigs vaccinated once as 3.4 (Stegeman et al., 1995b). For vaccinated sows it was estimated to be 0.7 (Van Nes et al., 1996). When major outbreaks occur in finishing herds they become highly infectious to other herds by non-animal contacts, e.g., airborne transmission and people visiting herds (Van Nes et al., 1998).

In our study, the type of farm (farrow to weaning or farrow to finish) did not have an influence on the probability of a sow farm becoming infected. The relation between the type of farm and the probability of being infected has produced contradictory results in the literature. Martini et al., (2003) did not found significant differences on the probability of becoming gE-seronegative between herds of different type (farrow to weaning versus farrow to finish). However, Marsh et al., (1991) and Norman et al., (1996), among others, found that being a farrow to finish herd increased the probability of becoming AD infected. These results might be linked with herd immunity as a consequence of vaccination. Stegeman et al., (1995b) concluded that double vaccination of finishing pigs, abolished the risk of new infections of sows by the presence of finishing pigs in the farm. In Spain, since 2003, fatteners have to be vaccinated at least

two times; the first application between 10 and 12 weeks of age and the second one three or four weeks after (Anonymous 2003a).

The spatial locations of the residuals of the model were useful to evidence the geographical areas where the probability of infection was less explained by the explanatory variables included in the model. As shown in figures 13 to 16, the probability of infection predicted by the model was much lower than the obtained from the real data. Therefore, the pattern of the residuals of the model (observed versus predicted) was very similar to the observed infection in sow farms, showing that neighbourhood transmission might not be the main factor related to the success of the eradication of Aujeszky`s disease in sow farms.
In a study conducted in Brittany by Heliez et al., (2000) the authors suggest that the two main risk factors for new infections by Aujeszky`s disease virus are airborne transmission and purchase of infected piglets. Austin and Weigel (1992) suggested that high density of AD infection was the primary risk factor for the spread of ADV between herds, and Norman et al., (1996) found that AD risk decreased with high density of non infected herds and low density of infected herds. If there is a high geographic density of ADV infected herds, there is a reservoir of virus for potential spread and transmission risks are increased (Austin and Weigel 1992). From the results of our study, it is concluded that neighbourhood transmission might not be the main factor related to the probability of infection of a sow farm. However, the spread of AD within the area was first suggested by the exploratory analysis; where a geographical association was found between clusters of positive sow farms, fattening farms and sow farms that became positive. Area includes many aspects influencing the epidemiology of ADV infection, some of them are related to local AD spread such as contacts between herds or pig trade pattern, and others are not related to local spread, such as: presence of related industries, farmers' attitude and compliance or efficiency of the veterinary services. We agree with Martini et al., (2003) that the existence of disease clusters is possible without farm-to-farm spread. However, as it was not possible to obtain reliable information on these factors, they were not included in the analysis.
The results concerning to ADV local spread from neighbouring farms might be influenced by the definition of neighbourhood used in the analysis. The decision to use a buffer radius of 750 meters as neighbourhood definition, to calculate positive density of animals around each sow farm, had the objective of testing neighbourhood spread and avoiding the introduction of more variability in the model. At such spatial scale, the relationship between the probability of being AD positive and the number of positive animals in the neighbourhood might have been undetected in some occasions. Marsh et al., (1991) did not found association between probability of becoming infected by AD and distance to nearest infected herd, however density of swine herds within a 5 km radius, was considered a risk factor. The results of Norman et al., (1996) suggesting the importance of local spread were obtained at a 3.22 km and a 6.44 km buffers, whereas the data of the 1.6 km buffer radius were too sparse to evaluate. Also, Tamba et al., (2002) in northern Italy, found that the number of pigs in a 6 km radius was a significant risk factor for sow farms (farrow to finish and farrow to weaning).

Risk mapping studies are based on the idea that, aside of herd-specific risk factors there might be unobserved geographical varying risk factors present (Berke and Beilage 2003). However, from the results of our study we conclude that geographical factors have had moderate importance and that herd-specific risk factors such as variables describing biosecurity measures, vaccination schemes or the health status of the

purchased pigs might be much more related to the probability of AD infection, and therefore to the success of the AD eradication programme, than the variables included in this study.

In the last years, spatial epidemiological analysis has become more easily accessible for epidemiologists. However, while developments in spatial statistics within the human health field have advanced considerably, there has been less evidence of a corresponding interest in such methodology within veterinary medicine (Lawson and Zhou, 2005). The development of models to explain the spatial evolution of the disease taking into account the location of the herds (considering the spatial autocorrelation) at point level has been a rarely studied topic in veterinary medicine. Other studies have attempted to develop models in order to analyze observational point data at some point time. Biggeri et al., (2006) used Bayesian Gaussian spatial exponential models and Bayesian kriging to study the risk of dog parasite infection in the city of Naples. Stevenson et al., (2005b) used a mixed-effects geostatistical model, accounting for extra-binomial variation and spatial dependence in Varroa prevalence on honey bees in New Zealand. The effect of distance from the likely site of incursion of Varroa into New Zealand on the probability of an apiary being Varroa-positive was introduced in the model as a covariate. In both studies, the spatial autocorrelation was parameterised as a continuous process; it was not based on a finite set of locations. In the study of Stevenson et al., (2005b), apiaries i and j were correlated to an extent determined by the distance separating them ($d_{ij}$). Also, Biggeri et al., (2006) used exponential decaying function to parameterise the spatial correlation. In our study, at first, we also tried to parameterise the spatial autocorrelation as a function of distance between sow herds (as an exponential decaying function). However, the simulation was computing not possible and we decided to change the approach. In our opinion, the size of the study population have make not possible to use this approach. The study population of Stevenson et al., (2005b) was comprised of 641 apriaries and Biggeri et al., (2006) collected faeces from 415 different points. In our study, the period with less sow farms was period 1 with 1,485 sow farms.

In order to fit the models we decided to parameterise the spatial dependence based on a finite set of locations. This set of locations was defined as those farms located in a 500 meters buffer radius around each sow farm; this decision was arbitrary and based on not including more variability in the model. To our knowledge, such approach to take into account the spatial autocorrelation has never been used, but we think it can be a good approach to take into account spatial dependence when the most common method, distance decaying based on an exponential function, is not possible to apply. Another choice could have been to create a lattice surface by aggregating data at some area level, creating a grid of a specific size and overlaying layers, in order to apply the CAR model in its classical framework (lattice data). However, we believe that as data was available at point level, it was not worthy to aggregate data as we would lose precision. On the other hand, the relation with covariates would have been very difficult to assess.

Alternatively to the study of the relation between topographical factors and the presence (probability of infection) of ADV in the herd during the different phases of the AD eradication programme, we tried to study the relation of those factors and the within-herd seroprevalence in sow herds. For that purpose, we also developed a hierarchical Bayesian model. In this model, we assumed that the number of serological positive sows in each sow farm follow a Poisson distribution. Also, because count data

contained an excess of zeros relative to the Poisson distribution (more than 90% in period 3) we decided to use a mixture model. In medical applications, a popular approach to analyse such data is to use a zero-inflated Poisson (ZIP) regression model. Also, in practice, data count is often overdispersed so that alternative distributions such as the zero-inflated binomial (ZINB) may be more appropriate than the ZIP (Xiang et al., 2007). In appendix 2, the specification, syntax and results of the ZINB model for each period, have been included.

We believe that the results from these models are not valid. The ZINB model used in our study nearly predicted all the observed cases in each of the sow farms. Because of that, as shown in figures 25 to 28, the residuals are very small. We believe that is not possible for those variables to predict almost all the observed cases in each sow farm, and therefore we have considered that the results were not valid. However, as the methodology and specification of the model might be of interest, we decided to include them in an appendix (appendix 2). ZINB models, to our knowledge, have not been used in veterinary medicine. Examples of such applications can be found in caries prevention in dental epidemiology (Böhning et al., 1999), fisheries research (Minami et al., 2007) or accident analysis (Li et al., 2008) among other disciplines.
A possible explanation for the lack of validity of the results of this model was the assumption that the observed number of cases in each sow farm to followed a Poisson distribution. The distribution of the counts is typically assumed to come from a Poisson distribution when the diseases considered are rare, so the distribution gives a good approximation of the underlying binomial distribution that would hold for each risk stratum. In a Poisson distribution, the probability of the counts must be independent for i=1,2,…, n (Richardson et al., 2004). AD is not a rare disease (at least in the first periods), and besides is an infectious disease, so the assumption that the number of positive sows present in the different sow farms can be modelled by the Poisson distribution might not be correct. In the bayesian framework, to our knowledge, there are not adjustment criteria for the models. If convergence is achieved, is believed that the results from the model are valid. DIC value can be used to compare model adjustment with different combinations of variables. In our models, convergence of the deviance (overall parameter of the model) was achieved in all the periods; however, we consider the results can not be valid, as only geographical factors are not likely to predict all the observed cases.
In veterinary medicine, hierarchical Bayesian models based on Poisson distributions have been used to study the geographical distribution of Bovine Spongiform Encephalopathy (BSE) risk and the relation with covariates such as pig population (Abrial et al 2005b; Ducrot et al., 2005; Stevenson et al., 2005a; Allepuz et al., 2007). However, BSE is a rare and non infectious disease, and might match better the Poisson assumption than AD. Different approaches might be necessary in order to study the spatial distribution of the AD risk and its relation with different covariates.

Another option for the lack of validity of the ZINB model in our study is the bias in the observed number of cases within each farm. Within sow prevalence was estimated from the number of positive samples obtained in the farm. In most of the cases, the AD serological screening carried out during the different eradication periods was based on a sample of animals that made it possible to state that the within-herd prevalence was lower than 5% in sow farms with a 95% confidence level. Because of that, samples obtained in this screening is not a good estimate of the within herd prevalence and might altered the results of the model.

# 6. CONCLUSIONS

**1. -** The distribution of Aujeszky's disease in Catalonia during the period 2003-2007 was not homogeneous. Clusters of positive sow and fattening farms were detected in all the study periods in the same areas. Clusters of negative sow farms that became positive were also detected in these areas. This geographical relationship suggests a possible role of the local spread in AD.

**2. –** The proportion of positive farms that eliminated the infection was higher in some areas. Therefore, the elimination of the disease had a spatial component.

**3. -** High number of fatteners AD serologically positive in the proximity of a sow farm increased its probability of becoming infected.

**4. –** The distance to the nearest slaughterhouse and road did not influence the probability of being AD positive in any of the periods of study.

**5. -** The studied geographical factors do not seem to have played an important role on the success of the AD eradication programme. Herd-specific risk factors such as biosecurity measures, vaccination schemes or the health status of the purchased pigs might be much more related to the probability of AD infection, and therefore to the success of the AD eradication programme, than the geographical variables included in this study.

# REFERENCES

Aamodt G., Samuelsen S.O., Skrondal A. (2006). A simulation study of three methods for detecting disease clusters. Intern Journal Health Geograph. Apr 12, 5-15.

Abrial D., Calavas D., Lauvergne N., Morignat E., Ducrot C. (2003). Descriptive spatial analysis of BSE in western France. Vet Res. 34, 749-760.

Abrial D., Calavas D., Jarrige N., Ducrot C. (2005a). Spatial heterogeneity of the risk of BSE in France following the ban of meat and bone meal in cattle feed. Prev Vet Med. 67, 69-82.

Abrial D., Calavas D., Jarrige N., Ducrot C. (2005b). Poultry, pig and the risk of BSE following the feed ban in France – A spatial analysis. Vet Res. 36, 615-628.

Alba, A., Allepuz, A., Serrano, E., Casal, J. (2008). Seroprevalence and Spatial Distribution of Maedi-Visna Virus and Pestiviruses in Catalonia (Spain). Small Rumin Res (in press).

Albina E., Mespleáde A., Chenut G., Le Potier M.F., Bourbao G., Le Gal S., Leforban Y. (2000).A serological survey on classical swine fever (CSF), Aujeszky's disease (AD) and porcine reproductive and respiratory syndrome (PRRS) virus infections in French wild boars from 1991 to 1998. Vet Microbiol. 77, 43-57.

Allepuz A., Lopez-Quilez A., Forte A., Fernandez G., Casal J. (2007). Spatial analysis of bovine spongiform encephalopathy in Galicia, Spain (2000-2005). Prev Vet Med. 79, 174 - 185.

Anonymous (1964). Council Directive 64/432/EEC of 26 June 1964 on animal health problems affecting intra-community trade in bovine animals and swine. Official Journal P 121, 29/07/1964, 1977–2012.

Anonymous (1993a). Commission Decision 93/24/EEC of 11 December 1992 concerning additional guarantees relating to Aujeszky's disease for pigs destined to Member States or regions free of the disease. Official Journal L 016, 25/01/1993, 0018–0021.

Anonymous (1993b). Commission Decision 93/244/EEC of 2 April 1993 concerning additional guarantees relating to Aujeszky's disease for pigs destined for certain parts of the territory of the Community. Official Journal L 185, 28/07/1993, 0054.

Anonymous (1995). BOE. Real decreto 245/1995.

Anonymous (2001). Commission decision 2001/618/EC of 23 July 2001 on additional guarantees in intra-community trade of pigs relating to Aujeszky's disease, criteria to provide information on this disease and repealing Decisions 93/24/EEC and 93/244/EEC. Official Journal L 215, 09/ 08/2001, 0048–0054.

Anonymous (2003a). BOE. Real decreto 427/2003.

Anonymous (2003b). ARP/254/2003 de 27 de Maig.

Anonymous (2005a). BOE. Real decreto 206/2005.

Anonymous (2005b). Joint Research Centre. Institute for Environment & Sustainability, Rem Group. URL: http://www.ai-geostats.org/index.php?id=faq

Anonymous (2006a). BOE. Real decreto 636/2006.

Anonymous (2006b). Programa de lluita, control i erradicació de la malaltia d'Aujeszky. Fase III. Document B. 29 de juny de 2006, pp: 66-67.
URL: http://www.gencat.net/darp/c/ramader/aujeszky/doc/documentb.pdf

Anonymous (2007a). Aujeszky's disease. Centre for Food security & Public Health. Iowa State University. URL: http://www.cfsph.iastate.edu/Factsheets/pdfs/aujeszkys_disease.pdf

Anonymous (2007b). Censo agrario 2007. http://www.idescat.es.

Anonymous (2008). Commision decision 2008/269/CE

Austin C.C. and Weigel R.M. (1992). Factors affecting the geographic distribution of pseudorabies (Aujeszky's disease) virus infection among swine herds in Illinois. Prev Vet Med. 13, 239-250.

Auvigne V. and Hery D. (1997). Analysis of the relationship between seroprevalence of Aujeszky´s disease and pig density within the different areas of Brittany. Vet Microbiol. 55, 153-158.

Auvigne V. and Bougeard C. (2000). Progress of the Aujeszky`s disease virus eradication programme in Brittany (France). Vet Res. 31, 155-156.

Barceló M.A., Saez M., Cano-Serral G., Martíneza-Beneito M.A., Martínez J.M., Borrel C., Ocaña-Riola R., Montoya I., Calvo M., López-Abente G., Rodrígyez Sanz M., Toro S., Alcalá J.T., Saurina C., Sánchez-Villegas P., Figueiras A. (2008). Métodos para la suavización de indicadores de mortalidad en ciudades del Estado Español (Proyecto MEDEA). Gaceta Sanitaria (in press).

Benschop J., Hazelton M.L., Stevenson M.A., Dahl J., Morris R.S., French N.P. (2007) Application of a novel method of spatially adaptive smoothing. GisVet 2007 proceedings. URL: http://www.gisvet.org/Conferences_2007.aspx

Berke P. and Beilage G.E. (2003). Spatial Relative Risk Mapping of Pseudorabies-Seropositive pig herds in an animal dense region. J. Vet. Med. B. 50. 322 -325.

Berke O. (2004). Exploratory disease mapping: kriging the spatial risk function from regional count data. Intern Journal Health Geograph. 3:18.

Berke O. (2005). Exploratory spatial relative risk mapping. Prev Vet Med. 71, 173-82.

Besag J., York J., Mollie A. (1991). Bayesian image restoration with applications in spatial statistics. Ann Inst Math Stat. 43, 1-20.

Best N. and Abellan J. (2006). Bayesian Analysis, Spatial Statistics and WinBUGS. Notes of statistical course held on Girona.

Best N., Richardson S., Thomson A. (2005). A comparison of Bayesian spatial models for disease mapping. Stat. Methods in Medical Research 14, 35-39.

Biggeri A., Dreassi E., Catelan D., Rinaldi L., Lagazio C., Cringoli G. (2006). Disease mapping in veterinary epidemiology: a Bayesian geostatistical approach. Stat Meth Med Res. 15, 1-16.

Boelaert F., Deluyker H., Maes D., Godfroid J., Raskin A., Varewijck H., Pensaert M., Nauwynck H., Castrick F., Miry C., Robijns J.M., Hoet B., Segers E. Van Vlaenderen I., Robert A., Koenen F. (1999). Prevalence of herds with young sows seropositive to pseudorabies (Aujeszky's disease) in northern Belgium. Prev Vet Med. 41, 239-255.

Boender G.J., Hagenaars T.J., Bouma A., Nodelijk G., Elbers A.R.W., de Jong M.C.M., van Boven M. (2007). Risk Maps for the Spread of Highly Pathogenic Avian Influenza in Poultry. PLOS Comput Biol. 3, 71.

Böhning D., Dietz E., Sclatmann P., Mendonça L., Kirchner U. (1999). The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. J Roy Stat Soc. 162, 195-209.

Bouma A., de Jong M.C.M., Kimman T.G. (1995). Transmission of pseudorabies virus within pig population is independent of the size of the population. Prev Vet Med. 23, 163-172.

Calistri P., Conte A., Freier J., Ward M. (2007a). Geographic information systems: introduction. Vet Ital. 43, 379

Calistri P., Savini L., Ippoliti C., Conte A. (2007b). Web-based geographic information system tools and international surveillance network for bluetongue in the Balkans and south-east Europe (East-BTNet). Vet Ital. 43, 739-744.

Calvete C., Miranda M.A., Estrada R., Borras D., Sarto i Monteys V., Collantes F.,Garcia-de-Francisco J.M., Moreno N., Lucientes J. (2006). Spatial distribution of Culicoides imicola, the main vector of bluetongue virus, in Spain. Vet Rec. 158, 130-131.

Capela R., Purse B.V., Pena I., Wittman E.J., Margarita Y., Capela M., Romão L., Mellor P.S., Baylis M. (2003). Spatial distribution of Culicoides species in Portugal in relation to the transmission of African horse sickness and bluetongue viruses. Med Vet Entomol. 2, 165-77.

Carpenter T.E. (2001). Methods to investigate spatial and temporal clustering in veterinary epidemiology. Prev Vet Med. 48, 303-320.

Casal J., Moreso J.M., Planas-Cuchí E., Casal J. (1997). Simulated airborne spread of Aujeszky's disease and foot and mouth disease. Vet Rec. 140, 672 – 676.

Christensen L.S., Mousing J., Mortensen S., Soerensen K.J., Strandbygaard S.B., Henriksen C.A., Andersen J.B. (1990). Evidence of long distance airborne transmission of Aujeszky's disease (pseudorabies) virus. Vet Rec. 127, 471-474.

Christensen L.S., Mortensen S., Botner A., Strandbygaard B.S., Ronsholt L., Henrikesn C.A. Andersen J.B. (1993). Further evidence of long distance airborne transmission of Aukeszky's disease (pseudorabies) virus. Vet Rec. 132, 317 – 321.

Clark P.J. adn Evans F.C. (1954). Distance to nearest neighbour as a measure of spatial relationships in populations. Ecology 35, 73-104.

Clayton D. and Bernardinelli L. (1992) Bayesian methods for mapping disease risk. In: Geographical and environment epidemiology : methods for small area studies. eds Elliott, P., Cuzick, J., English, D. and Stern, R. Oxford University Press, Oxford.

Congdon P. (2001). Bayesian Statistical Modelling. Wiley publishing. Chapter 1: Introduction; the Bayesian Method, its Benefits and Implementation.

Cuzick J. and Edwards R. (1990). Spatial clustering for inhomogeneous populations. *J* Roy Stat Soc B Stat Meth. 52, 73-104.

De la Rocque S., Michel V., Plazanet D., Pin R. (2004). Remote Sensing and epidemiology : examples of applications for two vector-borne diseases. Comp Immunol Microbiol Infect Dis. 27, 331-341.

Diggle, P.J., Chetwynd A.D., Haggkvist R. Morris S.E. (1995). Second-order analysis of space-time clustering. Stat Meth Med Res. 4, 124-136.

Doherr, M.G., Carpenter, T.E., Wilson, W.D., Gardner, I.A. (1999). Evaluation of temporal and spatial clustering of horses with Corynebacterium pseudotuberculosis infection in horses. Am. J. Vet. Res. 60, 284-291.

Doherr M.G., Hett A.R., Rüfenacht J., Zurbrigen A., Heim D. (2002). Geographical clustering of cases of bovine spongiform encephalopathy (BSE) born in Switzerland after the feed ban. Vet Rec. 151, 467-472.

Ducrot C., Abrial D., Calavas D., Carpenter T. (2005). A spatio-temporal analysis of BSE cases born after the reinforced feed ban in France. Vet Res. 36, 839-853.

Durr P.A., Argyraki A., Ramsey M., Clifton-Hadley R.S. (2000). Agro-ecological databases for spatial correlation studies: methodological issues. Thrusfield, M.V. and Goodall, E.A. (Ed) Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine, 225-235.

Durr P.A. and Gatrell A.C. (2004). GIS and Spatial Analysis in Veterinary Science. CABI Publishing.

Elbers A.R.W., Braamskamp J., Dekkers L.J.M., Voets R., Duinhof T., Hunneman W.A., Stegeman J.A. (2000). Aujeszky's disease virus eradication campaign successfully heading for last stage in the Netherlands. Vet. Quart. 22, 103-107.

Elliot P. and Wartenberg D. (2004). Spatial Epidemiology: Current Approaches and Future Challenges. Environ Health Perspect. 112, 998–1006.

French N.P., Berriatua E., Wall R., Smith K., Morgan KL. (1999) Sheep scab outbreaks in Great Britain between 1973 and 1992: spatial and temporal patterns. Vet Parasitol. 83, 187-200.

French N.P. and White P.C.L. (2004). GIS and Spatial Analysis in Veterinary Science. CABI Publishing. Chapter 7: The use of GIS in modelling the spatial and temporal spread of animal diseases, 117-204.

Gangnon R.E., Clayton M.K. (2000). Bayesian detection and modelling of spatial disease clustering. Biometrics. 56, 922-935.

Gatrell A.C. (2004). GIS and Spatial Analysis in Veterinary Science. CABI Publishing. Chapter 3: Geographical Information Science and Spatial Analysis in Human Health: Parallels and Issues for Animal Health Research.

Guis H, Tran A, de La Rocque S, Baldet T, Gerbier G, Barragua B, Biteau-Coroller F, Roger F, Viel J.F, Mauny F. (2007). Use of high spatial resolution satellite imagery to characterize landscapes at risk for bluetongue. Vet Res. 38, 669-683.

Hars J., Rossi S., (2005). Actualites dans le domaine de la surveillance des maladies transmissibles en France (peste porcine classique, maladie d´Aujeszky, tuberculose, brucellose, leptospirose, trichinellose, influenza aviaire, virus West Nile). Proceedings of the 23èmes Rencontres du GEEFSM, Andorre.

Heliez S., Auvigne V., Fourichon C. (2000). Risk factors of new Aujeszky's disease virus infection in swine herds in Brittany (France). Vet. Res. 31, 146-147.

Keeling M.J., Woolhouse M.E., Shaw D.J., Matthews L., Chase-Topping M., Haydon D.T., Cornell S.J., Kappey J., Wilesmith J., Grenfell B.T. (2001). Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science. 294, 813-817.

Kirkpatrick C.M., Kanitz C.L., McCrocklin S.M. (1980). Possible role of wild mammals in transmission of pseudorabies to swine. J Wildl Dis. 4, 601-614.

Kulldorf M. (1997) A spatial scan statistic. Comm Stat Theor Meth. 26, 1481–1496.

Kulldorf M. (2006). SaTScan™ User Guide for version 7.0. URL: http://www.SaTScan.org

Lawson A.B. and Williams F.L.R. (2001). An Introductory Guide to Disease Mapping. John Wiley & Sons.

Lawson A.B. Browne W.J. and Rodeiro C.L.V. (2003). Disease mapping with WinBUGS and MLwiN. John Wiley & Sons.

Lawson A.B., Zhou H. (2005).Spatial statistical modeling of disease outbreaks with particular reference to the UK foot and mouth disease (FMD) epidemic of 2001. Prev Vet Med. 71, 141-156.

Leblond A, Sandoz A, Lefebvre G, Zeller H, Bicout D.J. (2007). Remote sensing based identification of environmental risk factors associated with West Nile disease in horses in Camargue, France. Prev Vet Med. 16, 20-31.

Lertxundi A. (2006). Phd thesis: 'Métodos de estadística espacial para evaluar la influencia de factores medioambientales sobre la incidencia y mortalidad por cáncer'. URL: http://www.tesisenxarxa.net/TDX-0313107-114758/index_cs.html

Leuenberger R., Boujon P., Thür B., Miserez R., Garin-Bastuji B., Rüfenacht J., Stärk K. D. C. (2007). Prevalence of classical swine fever, Aujeszky's disease and brucellosis in a population of wild boar in Switzerland. Vet Rec. 160, 362-368.

Li Z., Knight S., Cook L.J., Hyde L.K., Holubkov R., Olson L.M. (2008). Modelling motor vehicle crashes for street racers using zero-inflated models. Accident Analysis and Prevention 40, 835-839.

Mainar J. R.C., Litgh-Pereira P.L., Epp T., Waldner C. (2005). The Application of Spatial Analysis Tools in Small-Ruminant Brucellosis Eradication Programs in Northern Spain. Intern. J Res Vet Med. 3, 179-188.

Maroney S.A., McCool M.J., Geter K.D., James A.M. (2007). The evolution of internet-based map server applications in the United Stated Department of Agriculture, Veterinary Services. Vet. Ital. 43, 723-730.

Marsh W.E., Damrongwatanapokin T., Larntz K., Morrison R.B. (1991). The use of a geographic information system in an epidemiological study of pseudorabies (Aujeszky's disease) in Minessota swine herís. Prev Vet Med. 11, 249-254.

Martin V., Simone L.D., Lubroth J. (2007). Geographic information systems applied to the international surveillance and control of transboundary animal diseases, a focus on highly pathogenic avian influenza. Vet. Ital. 43, 437-450.

Martini M., Drigo M., Dalla Poza M., Ferrari G., Martinello F., Sona B. (2003). A study of the progress of the Aujeszky's Disease Control Programme in Italy Using Survival Analysis. J. Vet. Med. B. 50, 191-195.

Mettenleiter T.C. (2000) Aujeszky's disease (pseudorabies) virus: the virus and molecular pathogenesis – State of the art, June 1999. Vet Res. 31, 99-115.

Miller R., Kaneene J.B., Schmitt S.M., Lusch D.P., Fitzgerald S.D. (2007). Spatial analysis of *Mycobacterium bovis* infection in white-tailed deer (*Odocoileus virginianus*) in Michigan, USA. Prev Vet Med. 82, 111-122.

Minami M., Lennert-Cody C.E., Gao W., Román-Verdesoto M. (2007). Modelling shark bycatch : The zero-inflated negative binomial regression model with smoothing. Fish Res. 84, 210-221.

Morris R.S., Wilesmith J.W., Stern M.W., Sanson R.L., Stevenson M.A. (2001). Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain, 2001. Vet Rec. 149, 137-44.

Mostashari F., Kulldorf M., Hartman J.J., Milerr J.R., Kulasekera V. (2003). Dead Bird Clusters as an Early Warning System for West Nile Virus Activity. Emer. Infec. Dis. 9, 6.

Müller T., Teuffert J., Zielder K., Possardt C., Kramer M., Staubach C., Conraths F.J. (1998). Pseudorabies in the European Wild Boar from Eastern Germany. J Wildl Dis. 34, 251-258.

Müller T., Bätza H.-J., Schlüter H., Conraths F.J., Mettenleiter T.C. (2003). Eradication of Aujeszky's Disease in Germany. J Vet Med B. 50, 207-213.

Mwalili S.M., Lesaffre E., Declerck D. (2004). The Zero-Inflated Negative Binomial Regresión Model with correction for misclassification: An example in Caries Research, Technical report. URL: http://www.stat.ucl.ac.be/ISpub/tr/2004/TR0462.pdf

Norman H.S., Sischo W.N., Pitcher P., Nesselrodt A., Day L. (1996). Spatial and temporal epidemiology of pseudorabies virus infection. Am J Vet Res. 57, 1563–1568.

O'Brien D.J., Kaneene J.B., Getis A., Lloyd J.W., Swanson G.M., Leader R.W. (2000). Spatial and temporal comparison of selected cancers in dogs and humans, Michigan, USA, 1964-1994. Prev Vet Med. 47, 187-204.

OIE (2004). World Organisation for Animal Health. Manual of Diagnostic Tests and Vaccines for terrestrial Animals, 2004. URL: http://www.oie.int/eng/normes/mmanual/A_00041.htm

Olea-Popelka FJ, Griffin JM, Collins JD, McGrath G, Martin SW. (2003). Bovine tuberculosis in badgers in four areas in Ireland: does tuberculosis cluster? Prev Vet Med. 59, 103-111.

Olea-Popelka FJ, Flynn O, Costello E, McGrath G, Collins JD, O'Keeffe JO, Kelton DF, Berke O, Martin SW. (2005). Spatial relationship between Mycobacterium bovis strains in cattle and badgers in four areas in Ireland. Prev Vet Med. 71, 57-70.

Pejsak Z.K. and Truszcynski J. (2006). Aujeszky's disease (pseudorabies). In Diseases of Swine (9th edition). Straw B.E., Zimmerman J.J., Allaire S.D., Taylor D.J. Ed: Blackwell publishing.

Perruci S., Pinello E., Fichi G., Ciardi E., Bàrberi P., Moonen C., Ragaglini G., Bibbiani C. (2007). Parasitic infections in an organic grazing cattle herd in Tuscany using geographic infromation systems to determine risk factors. Vet. Ital. 43, 415-424.

Perez A.M., Ward M.P., Torres P., Ritacco V. (2002). Use of spatial statistics and monitoring data to identify clustering of bovine tuberculosis in Argentina. Prev Vet Med. 29, 63-74.

Picado A., Guitian F.J., Pfeiffer D.U. (2007). Space–time interaction as an indicator of local spread during the 2001 FMD outbreak in the UK. Prev Vet Med. 79, 3-19.

Pfeiffer D.U. (2000). Spatial analysis – A new challenge for veterinary epidemiologists. SVEPM proceedings, 87-105.

Pfeiffer D.U. (2004). GIS and Spatial Analysis in Veterinary Science. CABI Publishing. Chapter 5: Geographical Information Science and Spatial Analysis in Animal Health.

Pfeiffer D.U. and Jones M.H. (2002). Geographical information systems as a tool in epidemiological assessment and wildlife disease management. Rev. Sci. tech. Off. Int. Epiz. 21, 91-102.

Pfeiffer D.U., Minh P.Q., Martin V., Epprecht M., Otte M.J. (2007). An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data. Vet J. 174, 302-309.

Plummer M., Best N., Cowles K., Vines K. (2006). CODA: Concergence Diagnosis and Output Analysis for MCMC. The Newsletter of the R Project. Volume 6/1, March. URL: http://cran.r-project.org/doc/Rnews/Rnews_2006-1.pdf

Poljak Z., Dewey C.E., Martin S.W., Christensen J.C., Carman S., Friendship R.M. (2007). Spatial clustering of swine influenza in Ontario on the basis of herd-level disease status with different misclassification errors. Prev Vet Med. 81, 236-249.

Poljak Z., Friendship R.M., Carman S., McNab W.B., Dewey C.E. (2008). Investigation of exposure to swine influenza viruses in Ontario (Canada) finisher herds in 2004 and 2005. Prev Vet Med. 83, 24- 40.

Porphyre T., McKenzie J., Stevenson M. (2007). A descriptive spatial analysis of bovine tuberculosis in intensively controlled cattle farms in New Zealand. Vet. Res. 38, 465-479.

Purse B.V., Nedelchev N., Georgiev G., Veleva E., Boorman J., Denison E., Veronesi E., Carpenter S., Baylis M., Mellor P.S. (2006). Spatial and temporal distribution of bluetongue and its Culicoides vectors in Bulgaria. Med. Vet. Entomol. 20, 335-344.

Richardson S., Thomson A., Best N., Elliot P. (2004). Interpreting Posterior Relative Risk Estimates in Disease-Mapping Studies. Environ Health Perspect. 112, 9.

Robertson J.A. and Wierup M. (2000). The eradication of Aujeszky's disease from pig production in Sweden. Vet Res. 31, 153-154.

Robson J., Yeoman G.H. Ross J.P.J. (1961). *Rhipicephalus appendiculatus* and East Coast Fever in Tanganyika. East Afr Med J. 38, 206-214.

Rogerson P.A. (2006). Statistical methods for the detection of spatial clustering in case-control data. Stat. Med. 25, 811-823.

Romero C. H., Meade P. N., Shultz J. E., Chung H. Y., Gibbs E. P., Hahn E. C., Lollis G. (2001). Venereal transmission of pseudorabies viruses indigenous to feral swine. J Wildl Dis. 37, 289-296.

Romero C. H., Meade P. N., Homer B.L., Shultz J. E., Lollis G. (2003). Potential sites of virus latency associated with indigenous pseudorabies viruses in feral swine. J Wildl Dis. 39, 567-575.

Ruiz Fons F. (2006). Interacciones epidemiológicas sobre la enfermedad de Aujeszky entre el porcino doméstico y el jabalí en Castilla-La Mancha, España. Riesgos saniatarios asociados a la producción cinegética del jabalí: la enfermedad de Aujeszky, phd thesis.

Saez M. and Saurina C. (2007). Estadística y epidemiología espacial. Girona: Edicions a petició. Documenta Universitaria. Capítulo 1: Introducción a la estadística especial.

Schmidt J., Gerdts V., Beyer J., Klupp B.G., Mettenleiter T.C. (2001). Glycoprotein D-Independent Infectivity of Pseudorabies Virus Results in an Alteration of In Vivo Host Range and Correlates with Mutations in Glycoproteins B and H. J Virol. 1001, 10054-10064

Segalés J, Chianini F, Cortes R, Domingo M. (2003). Neurological disorders associated with Aujeszky's disease virus infection in fattening pigs. Vet Rec. 23, 240-241.

Sheridan H.A., MacGrath G., Fallon R., Shoukri M.S., Martin S. W. (2005). A temporal-spatial analysis of bovine spongiform encephalopathy in Irish cattle herds, from 1996 to 2000. Can. J. Vet. Res. 69, 19-25.

Solymosi N., Reiczigel J., Berke O., Harnos A., Szigeti S., Fodor L., Szigeti G., Bodis K. (2004). Spatial risk assessment of herd sero-status of Aujeszky's disease in a county in Hungary. Prev Vet Med. 65, 9-16.

Solymosi N., Wagner S. E., Speiser F., Allepuz A. P., Valcz G., Lakos A. maps2WinBUGS: a tool to prepare maps and data for Bayesian mapping. Environmental Modelling & Software *SUBMITTED*

Song C., Kulldorf M. (2003). Power evaluation of disease clustering tests. Intern. Journal. Health. Geograph. 2, 9.

Spiegelhalter D. (2002). Bayesian measures of model complexity and fit. J.R. Statistic. Soc. B. 64, 583-639.

Stegeman A., Elbers A.R.W., Van Oirschot J.T., Hunneman W.A., Kimman T.G., Tielen M.J.M. (1995a). A retrospective study into characteristics associated with the seroprevalence of pseudorabies virus-infected breeding pigs in vaccinated herds in the southern Netherlands. Prev Vet Med. 22, 273-283.

Stegeman A., Van Nes A., De Jong M.C.M., Bolder F.W.M.M. (1995b). Assessment of the effectiveness of vaccination against pseudorabies virus in finishing pigs. Am J Vet Res. 56, 573-578.

Stevenson M.A. (2003). Phd thesis. The Spatio-temporal Epidemiology of Bovine Spongiform Encephalopathy and Foot-and-Mouth Disease in Great Britain. URL: http://epicentre.massey.ac.nz/Portals/0/EpiCentre/Downloads/Publications/Thesis/MarkStevensonPhD.pdf

Stevenson M.A., Wilesmith J.W., Ryan J.B.M., Morris R.S., Lawson A.B., Pfeiffer D.U., Lin D. (2000). Descriptive spatial analysis of the epidemic of bovine spongiform encephalopathy in Great Britain to June 1997. Vet Rec 147, 371-384.

Stevenson M.A., Morris R.S., Lawson A.B., Wilesmith J.W., Ryan J.B.M., Jackson R., (2005a). Area-level risks for BSE in British cattle before and after July 1988 meat and bone meal feed ban. Prev Vet Med. 69, 129-144.

Stevenson M.A., Benard H., Bolger P., Morris R.S. (2005b). Spatial epidemiology of the Asian honey bee mite (*Varroa destructor*) in the North Island of New Zealand. Prev Vet Med. 71, 241-252.

Taft A.C. (2000). The eradication of Aujeszky's disease in the United States. Vet Res. 31, 158.

Tamba M., Calabrese R., Finelli E., Cordioli P. (2002). Risk factors for Aujeszky`s-disease seropositivity of swine herds of a region of northern Italy. Prev Vet Med. 54, 203-212.

Tatem A.J., Baylis M., Mellor P.S., Purse B.V., Capela R., Pena I., Rogers D.J.(2003). Prediction of bluetongue vector distribution in Europe and north Africa using satellite imagery. Vet Microbiol. 97, 13-29.

Torina A., Caracappa S., Mellor P.S., Baylis M., Purse B.V. (2004). Spatial distribution of bluetongue virus and its Culicoides vectors in Sicily. Med Vet Entomol. 18, 81-89.

Turner C., Williams S.M., Cumby T.R. (2000). The inactivation of foot and mouth disease, Aujeszky's disease and classical swine fever viruses in pig slurry. J Appl Microbiol. 89, 760-767.

Ugarte M.D., Ibáñez B., Militino A.F. (2005). Detection of spatial variation in risk using CAR models for smoothing relative risks. Stoch Environ Res Risk Assess. 19: 33-40.

Van Nes A., Stegeman J.A., De Jong M.C.M., Loeffen W.L.A., Kimman T.G., Verheijden J.H.M. (1996). No major outbreaks of pseudorabies virus in well-vaccinated sow herds. Vaccine. 14, 1042-1044.

Van Nes A., De Jong M.C.M., Buitjels J.A.A.M., Verheijden J.H.M. (1998). Implications derived from a mathematical model for eradication of pseudorabies virus. Prev Vet Med. 33, 39-58.

Vannier P., Strugar S., Pansart J.F., Le Gal M. (2000). Progress in eradication and control programmes against Aujeszky's disease (pseudorabies) in France. Vet. Res. 31, 141-145.

Van Oirschot JT, Rziha HJ, Moonen PJ, Pol JM, van Zaane D., (1986). Differentiation of serum antibodies from pigs vaccinated or infected with Aujeszky's disease virus by a competitive enzyme immunoassay. J Gen Virol. 67, 1179-1182.

Van Oirschot J.T., Daus F., Kimman T.G., van Zaane D., (1991). Antibody response to glycoprotein I in maternally immune pigs exposed to a mildly virulent strain of Pseudorabies virus. Am J Vet Res. 52, 1788-1793.

Vengust G., Valencak Z., Bidovec A. (2006). A Serological Survey of Selected Pathogens in Wild Boar in Slovenia. J Vet Med B. 53, 24-27.

Vicente J., Ruiz-Fons F., Vidal D., Höfle U., Acedevedo P., Villanúa D., Fernández de Mera I.G., Martín M.P., Gortazar C. (2005). Serosurvey of Aujeszky´s disease virus infection in European wild boar in Spain. Vet Rec. 156, 408-412.

Vigre H., Baekbo P., Jorsal S.E., Billi-Hansen V., Hassing A.G., Enoe C., Botner A. (2005). Spatial and temporal patterns of pig herds diagnosed with Postweaning Multisystemic Wasting Syndrome (PMWS) during the first two years of its occurrence in Denmark. Vet Microbiol. 110, 17-26.

Ward M.P. (2007). Spatio-temporal analysis of infectious disease outbreaks in veterinary medicine: clusters, hotspots and foci. Vet Ital. 43, 559-570.

Ward M.P. and Carpenter, T.E. (1995). Infection of cattle in Queensland with bluetongue viruses. II. Distribution of antibodies. Aust Vet J. 72, 217-220.

Ward M.P. and Carpenter T.E. (2000). Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. Prev Vet Med. 45, 257–284.

Weigel R.M., Austin C.C., Siegel A.M., Biehl L.G., Taft A.C. (1992). Risk factors associated with the seroprevalence of pseudorabies virus in Illinois swine herds. Prev Vet Med. 12, 1-13.

Westergaard J.M. (2000). Aujeszky's disease: the position of the European Commission and regulation perspectives. Vet Res. 31, 160-161.

Wint W. (2007). Veterinary Web-geographic information systems: what's the point and what's involved?. Vet Ital. 43, 709-721.

Xiang L., Lee A.H., Yau K.K.W., McLachlan J.M. (2007). A score test for overdispersed in zero-inflated poisson mixed regression model. Stat. Med. 26, 1608-1622.

Yeoman G.H. (1966). Field vector studies of epizootic East Coast Fever. II. Seasonal studies of R. appendiculatus on bovine and non-bovine hosts in East Coast Fever enzootic, epizootic and free areas. Bull Epizoot Dis Afr. 14, 113-140.

Zanardi G., Tamba M., Macchi C., Alborali L., Guadagnini P. (2000). Evaluation of Italian Aujeszky's disease control programme in two regions of northern Italy in 1997. Vet Res. 31, 154-155.

Zuckermann F.A. (2000). Aujeszky´s disease virus: opportunities and challenges. Vet Res. 31, 121-131.

Zupancic Z., Jukic B., Lojkic M., Cac Z., Jemersic L., Staresina V. (2002) Prevalence of Antibodies to Classical Swine Fever, Aujeszky's Disease, Porcine Reproductive and Respiratory Syndrome, and Bovine Viral Diarrhoea Viruses in Wild Boars in Croatia. J. Vet. Med. B. 49, 253-256.

# Summary

Aujeszky's disease (AD) eradication programme started in Spain in 1995, but it was not until 2003, due to the additional guarantees in intra-community trade relating to Aujeszky's, that AD eradication programme was adapted and ensured. The aim of this study is to conduct a spatial analysis of the Aujeszky's disease (AD) eradication programme in Catalonia, Spain, from 2003 to 2007. The study has been divided in four periods, based on the phases designed in the AD eradication programme in Catalonia.

In the first part of the study, we explore for *high risk areas* (clusters) in order to test whether the spatial distribution of AD in the region during the consecutive eradication periods was homogeneously distributed over the territory or clustered in space. Different purely spatial analyses, based on the Bernoulli model, were run with SaTScan® v6.1 in each period. Clusters of positive sow farms (farrow to weaning and farrow to finish) and/or fattening farms were identified in the four study periods in the western, central and north eastern part of the region. The prevalence ratio values of these clusters increased throughout the study period due to the fact that the risk of disease decreased faster outside the clusters than inside the clusters. In order to study the evolution of the disease, we explored for areas where more negative sow farms became infected and areas where more sow farms eliminated the infection. These analyses demonstrated areas with significantly higher proportions of sow farms that became negative, which indicates that the eradication of the disease has a spatial component.
Clusters of negative sow farms that were infected again (reinfections) were also detected in the four study periods. The relative risk values of these clusters were much higher compared to the other cluster analyses. There was a geographical association between the clusters of positive sow farms, positive fattening farms and re-infected sow farms. This association could be attributable to the local spread of Aujeszky´s disease virus. Pig farm density could be a factor influencing the local spread of infection and was therefore evaluated for clusters of re-infected sow farms and clusters of sow farms that eliminated the infection. The mean density of pig farms was 0.40 farms per square Km (median of 0.28 and standard deviation of 0.33) in clusters of sow farms that became negative and 1.51 (median of 0.70 and standard deviation of 1.61) in clusters where more sow farms became positive (p-value<0.05).

Based on these results, in the second part of the study, we tested the role of geographical factors that could be implicated in local spread and persistence of AD in certain areas. Several geographic variables describing the possible risk factors associated to neighbourhood transmission: Distance to the nearest slaughterhouse, distance to conventional roads, mean number of AD serological positive sows and serological positive fattening pigs in the neighbourhood (750 meters radius) of each sow farm were included in a hierarchical Bayesian binomial model. A non geographic variable; type of farm (farrow to weaning *versus* farrow to finish) was also included. The use of Bayesian models allowed us to take into account the spatial dependence (autocorrelation) among the data; included in the model as a random effect. Spatial dependence was parameterised with a conditional autoregressive distribution (CAR) based on a set of neighbours. The set of neighbours was defined as those farms located in a 500 meters buffer radius around each sow farm.

From the four geographical variables included in the model, only positive fattening animals in the neighbourhood of sow farms increased the probability of being AD positive**.** In the first period, 1,000 positive fattening pigs in the neighbourhood (750 meters buffer radius) increase the odds of each sow farm being AD positive by a factor between 1.005 and 1.36. In period 2.2, having positive fattening animals in the neighbourhood increased the likelihood of each sow farm to be AD positive between 1.84 and 3.22. In period 2.1 and period 3, none of the variables had a positive relation with the probability of being positive. The type of farm (farrow to weaning or farrow to finish) also did not influence the probability of being AD positive in any period.
The geographical pattern of the residuals of the hierarchical bayesian binomial model (observed versus predicted) was very similar to the observed infection in sow farms in all the eradication periods, showing that neighbourhood transmission might not be the main factor related to the eradication of Aujeszky`s disease in sow farms. Other herd-specific risk factors might be much more related to the probability of AD infection than the geographical variables included in this study.

# Resum

El programa d'eradicació de la malaltia d'Aujeszky va començar a Espanya l'any 1995, però no va ser fins el 2003, quan a causa de les garanties suplementàries establertes en els intercanvis intracomunitaris de l'espècie porcina amb relació a la malaltia d'Aujeszky, que aquest programa es va reforçar i es varen establir les bases del programa coordinat de lluita, control i eradicació de la malaltia. L'objectiu d'aquest estudi és realitzar una anàlisi espacial de l'eradicació de la malaltia d'Aujeszky a Catalunya (Espanya) des del 2003 fins el 2007. L'estudi s'ha dividit en quatre períodes, en base a les diferents etapes establertes al programa d'eradicació a Catalunya.

A la primera part de l'estudi, varem analitzar si la distribució espacial de la malaltia d'Aujeszky a Catalunya ha estat homogènia o hi han hagut zones d'alt risc (conglomerats) durant les diferents etapes del programa d'eradicació. Per fer-ho, en cada període varem realitzar diferents anàlisis espacials amb el programa SaTScan v6.1 basats en el model de Bernoulli. En els quatre períodes d'estudi, varem identificar conglomerats de granges positives de truges (cicle obert i cicle tancat) i/o de granges positives d'engreixos, tant a l'oest com al centre com a l'est de Catalunya. Com que el risc d'infecció va disminuir més ràpidament fora dels conglomerats que dintre, els valors del ràtio de prevalença d'aquests conglomerats augmenten al llarg del temps. Per analitzar l'evolució de la malaltia, varem estudiar si hi havia àrees en les que la proporció de granges que s'havien reinfectat o que havien eliminat la infecció era més gran. Aquestes anàlisis van demostrar que hi havia zones en les que la proporció de granges que havien eliminat la infecció era més alta, i per tant que l'eradicació de la malaltia té també un component espacial.

En els quatre períodes d'estudi, també es van detectar àrees en les que la proporció de granges reinfectades havia estat més alta. El risc relatiu d'aquests conglomerats era més gran que el dels conglomerats descrits abans. D'altra banda, existeix una associació geogràfica entre els conglomerats de granges de mares positives, granges d'engreix positives i granges de mares reinfectades. Aquesta associació podria ser deguda a la transmissió a nivell local del virus d'Aujeszky. Ja que la densitat de granges a una zona podria ser un factor relacionat amb aquesta transmissió local, varem analitzar aquesta variable en conglomerats d'eradicació i de reinfeccions. La densitat mitjana de granges de porc als conglomerats d'eradicació és de 0.4 granges per quilòmetre quadrat (mitjana de 0.28 i desviació estàndard de 0.33) i de 1.51 (mitjana de 0.7 i desviació estàndard de 1.61) als conglomerats on més granges de truges s'han reinfectat (valor de $p<0.05$).

En base a aquests resultats, a la segona part de l'estudi varem analitzar el paper que podien exercir factors geogràfics en la transmissió a nivell local del virus i en la persistència de la malaltia d'Aujeszky a determinades zones. Per fer-ho, varem usar un model jeràrquic bayesià, en el que varem incloure diferents variables geogràfiques que podien estar implicades en la transmissió a nivell local del virus; com són la distància a l'escorxador més proper, distància a la carretera més pròxima, nombre d'animals d'engreix positius pròxims a la granja (radi de 750 metres) i nombre de truges positives pròximes a la granja (radi de 750 metres). Al model també varem incloure una altra variable no geogràfica: tipus de granja (cicle obert o cicle tancat). L'ús d'aquests models jeràrquics bayesians permet d'incorporar un terme que té en compte la dependència espacial (autocorrelació) existent a les dades. La dependència espacial va

ser inclosa al model mitjançant una distribució normal condicionalment autoregressiva (CAR) basada en un nombre de veïns. Aquests veïns van ser definits com aquelles granges localitzades en un radi de 500 metres de cada granja de truges.

De les quatre variables geogràfiques incloses al model, només la presència d'animals d'engreix positius presents a la proximitat d'una granja de truges incrementava la probabilitat d'infecció pel virus d'Aujeszky. Al primer període, per cada 1000 porcs d'engreix al voltant de cada granja de mares, l'odds (raó de probabilitats) de cada granja d'ésser positiva s'incrementava per un factor entre 1.005 i 1.36. En el període 2.2, tenir porcs d'engreix al voltant augmentava la raó de probabilitats d'infecció per un valor d'entre 1.84 i 3.22. En el període 2.1 i en el període 3, cap de les variables va influir de forma significativa en la probabilitat de ser una granja positiva. El tipus de granja (cicle obert o cicle tancat) tampoc es va relacionar amb la probabilitat de ser una granja positiva en cap dels períodes de l'estudi.
El patró geogràfic dels residus (observats *versus* predits) del model binomial jeràrquic bayesià va ser molt similar al dels observats, en tots els períodes de l'estudi. Aquest resultat evidencia que la transmissió a nivell local del virus d'Aujeszky probablement no hagi estat el principal factor relacionat amb la persistència del virus en granges de truges. Altres factors, específics de cada granja, probablement han tingut una relació més alta en la probabilitat d'infecció que les variables geogràfiques incloses en aquesta anàlisi.

# Resumen

El programa de erradicación de la enfermedad de Aujeszky comenzó en España en 1995, pero no fue hasta el 2003, cuando debido a las garantías suplementarias establecidas en los intercambios intracomunitarios de la especie porcina en relación a la enfermedad de Aujeszky, que dicho programa se reforzó y se establecieron las bases del programa coordinado de lucha, control y erradicación de la enfermedad. El objetivo de este estudio es realizar un análisis espacial de la erradicación de la enfermedad de Aujeszky en Cataluña (España) desde el 2003 hasta el 2007. El estudio se ha dividido en cuatro periodos, en base a las diferentes etapas establecidas en el programa de erradicación en Cataluña.

En la primera parte del estudio, analizamos si la distribución espacial de la enfermedad de Aujeszky en Cataluña ha sido homogénea o han existido *zonas de alto riesgo* (conglomerados) durante las distintas etapas del programa de erradicación. Para ello, en cada periodo realizamos diferentes análisis espaciales con el programa SaTScan® v6.1 basados en el modelo de Bernoulli. En los cuatro periodos de estudio, identificamos conglomerados de granjas positivas de cerdas (ciclo abierto y ciclo cerrado) y/o de granjas positivas de engordes, tanto en la parte oeste como en el centro y este de Cataluña. Debido a que el riesgo de infección disminuyó más rápido fuera de los conglomerados que dentro, los valores del ratio de prevalencia de estos conglomerados aumentaron a lo largo del tiempo. Para analizar la evolución de la enfermedad, estudiamos si había áreas en las que la proporción de granjas que se habían reinfectado o que habían eliminado la infección era mayor. Estos análisis demostraron que había zonas en las que la proporción de granjas que habían eliminado la infección era más alta, y por lo tanto que la erradicación de la enfermedad tiene también un componente espacial.

En los cuatro periodos de estudio, también se detectaron áreas en las que la proporción de granjas reinfectadas fue más alta. El riesgo relativo de estos conglomerados era mayor que el de los otros análisis de conglomerados. Por otro lado, existía una asociación geográfica entre los conglomerados de granjas de madres positivas, granjas de engorde positivas y granjas de madres reinfectadas. Esta asociación podría ser debida a la transmisión a nivel local del virus de Aujeszky. Ya que la densidad de granjas en una zona podría ser un factor relacionado con esta transmisión local, analizamos esta variable en conglomerados de erradicación y de reinfecciones. La densidad media de granjas de porcino en los conglomerados de erradicación fue de 0.4 granjas por kilómetro cuadrado (mediana de 0.28 y desviación estándar de 0.33) y de 1.51 (mediana de 0.7 y desviación estándar de 1.61) en los conglomerados donde más granjas de cerdas se habían reinfectado (valor de $p<0.05$).

En base a estos resultados, en la segunda parte del estudio analizamos el papel que podían desempeñar factores geográficos en la transmisión a nivel local del virus y en la persistencia de la enfermedad de Aujeszky en determinadas zonas. Para ello, usamos un modelo jerárquico bayesiano y en él incluimos diferentes variables geográficas que podían estar implicadas en la transmisión a nivel local del virus; como son la distancia al matadero más cercano, distancia a la carretera más próxima, número de animals de engorde positivos próximos a la granja (radio de 750 metros) y número de cerdas positivas próximas a la granja (radio de 750 metros). En el modelo también incluimos otra variable no geográfica: tipo de granja (ciclo abierto o ciclo cerrado). El uso de estos

modelos jerárquicos bayesianos permite incorporar un término que tiene en cuenta la dependencia espacial (autocorrelación) existente en los datos. La dependencia espacial fue incluida en el modelo mediante una distribución normal condicionalmente autoregresiva (CAR) basada en un número de vecinos. Dichos vecinos fueron definidos como aquellas granjas localizadas en un radio de 500 metros de cada granja de cerdas.

De las cuatro variables geográficas incluidas en el modelo, sólo la presencia de animales de engorde positivos presentes en la proximidad de una granja de cerdas incrementaba la probabilidad de infección por el virus de Aujeszky. En el primer periodo, por cada 1000 cerdos de engorde en la vecindad de cada granja de madres, el odds (razón de probabilidades) de ser positiva de cada granja se incrementaba por un factor entre 1.005 y 1.36. En el periodo 2.2, tener cerdos de engorde en la vecindad aumentaba la razón de probabilidades de infección por un valor entre 1.84 y 3.22. En el periodo 2.1 y en el periodo 3, ninguna de las variables influyó de forma significativa en la probabilidad de ser una granja positiva. El tipo de granja (ciclo abierto o ciclo cerrado) tampoco se relacionó con la probabilidad de ser una granja positiva en ninguno de los periodos del estudio.
El patrón geográfico de los residuos (observados *versus* predichos) del modelo binomial jerárquico bayesiano fue muy similar al de los observados, en todos los periodos del estudio. Este resultado evidencia que la transmisión a nivel local del virus de Aujeszky probablemente no haya sido el principal factor relacionado con la persistencia del virus en granjas de cerdas. Otros factores, específicos de cada granja, probablemente tengan una relación más alta en la probabilidad de infección que las variables geográficas incluidas en este análisis.

# APPENDIXES

# Appendix 1. Sintaxis of the binomial model

```
model {
  for (i in 1 : N) {
     O[i]  ~ dbin(p[i],n[i])
logit(p[i]) <- alpha+b1*x0[i]+c11[i]*x11[i]+c31[i]*x31[i]+c35*x35[i]+c36*x36[i]+ S[i]+eta[i]
eta[i]~dnorm(0, tau.eta)}

      for (i in 1 : N) {
      c11[i]~ dnorm(0, tau.c11)
      c31[i]~ dnorm(0, tau.c31)}

  # CAR prior distribution for spatial random effects:
  S[1:N] ~ car.normal(adj[], weights[], num[], tau.S)
  for(k in 1:sumNumNeigh) {
     weights[k] <- 1
  }
  # Other priors:
  alpha  ~ dflat()

# Prior on regression coefficients

  b1 ~ dnorm(0,tau.b1)
  c35 ~ dnorm(0,tau.c35)
  c36 ~ dnorm(0,tau.c36)

#Hyperpriors

  tau.S   ~ dgamma(0.1,0.001)
  sdS <- sqrt(1 / tau.S)
  tau.eta   ~ dgamma(0.1,0.001)
  sdeta <- sqrt(1 / tau.eta)
  tau.c11  ~ dgamma(0.1, 0.001)
  sdc11 <- sqrt(1 / tau.c11)
  tau.c31  ~ dgamma(0.1, 0.001)
  sdc31 <- sqrt(1 / tau.c31)
  tau.b1  ~ dgamma(0.1, 0.001)
  sdb1 <- sqrt(1 / tau.b1)
  tau.c35  ~ dgamma(0.1, 0.001)
  sdc35 <- sqrt(1 / tau.c35)
  tau.c36  ~ dgamma(0.1, 0.001)
  sdc36 <- sqrt(1 / tau.c36)
  }
```

**Interpretation*:***

**alpha:** constant (intercept)

**b1, c11, c31, c35, c36:** Parameters of type of farm, distance to slaughterhouse, distance to roads, mean number of positive sows and mean number of positive fatteners covariates.

**x0, x11, x31, x35, x36:** Type of farm, distance to slaughterhouse, distance to roads, mean number of positive sows and mean number of fatteners covariates.

**S, eta:** random effects for the correlated and uncorrelated heterogeneity.

# APPENDIX 2: ZINB model

***1) Specification of the model:***

For each sow farm we estimated the relative risk ($RR_i$), which represents an increase/decrease in the risk of infection compared to a standard risk evaluated in the whole region:

$$RR_i = \frac{O_i}{E_i}$$

Being Oi and Ei the observed and expected positive sows, respectively, in each farm. The observed positive sows in each farm were computed as follows:

$$O_i = \frac{P_i}{T_i} * N_i$$

Where i=1,…, n indexes the farm, '$N_i$' is the number of sows present and '$P_i$' and '$T_i$' are the number of positive sows to the gE official enzyme-linked immunosorbent assay method and the number of sows tested, respectively, at the end of each eradication period.
The expected positive sows in each farm were estimated by applying the overall AD ratio in the whole of Catalonia to the sow population in each farm:

$E_i = R*N_i$

Where 'R' and '$N_i$' are the overall ratio of positive sows in the whole region and the number of sows in each 'i' farm respectively.

Due to the counts of cases observed in a particular location ($O_i$) is a counter variable, the number of AD positive sows in each farm can be modelled as a Poisson distribution centred on $\mu$:

$\mu_i$ ($\mu_i = RR_i E_i$)

$O_i \sim$Poisson ($\mu_i$)

However, because the excess of zeros in the data count (i.e. high proportion of farms with no positive animals, especially in the last periods) we decided to use a mixture distribution rather than a Poisson distribution for the number of AD positive sows in

each farm. This model assumes an underlying partition of the population into two homogeneous components, one for the zero counts (farms without seropositive animals) and the other for the non-zero counts (farms with seropositive animals). Therefore, the data likelihood is a mixture of two distinct processes:

$$L(Y_i) = \pi_i f_1 + (1-\pi_i) f_2$$

where $i=1,...,n$ indexes farm; '$\pi'_i$ is the probability of being a zero count and $1-\pi_i$ the probability of a non-zero count; $f_1$ is discrete with mass point at zero and $f_2$ is a negative binomial count distribution (Poisson with extra variation) (Greene, 1994) with mean $\mu_i$. Leading to a Zero Inflated Negative Binomial model (ZINB).

The probability of having positive animals in the sow farm is given by:

$$\log(\mu_i) = \beta_0 + a_i + s_i$$

Where '$a_i$' and '$s_i$' indicates the uncorrelated heterogeneity and spatially-correlated heterogeneity random effects respectively.

***2) Incorporating covariates in the models:***

The number of slaughtered animals is not homogeneous over the different slaughterhouses of the region. Also, the number of transported pigs is not homogenous in the different roads. Because of that, we assumed that the different slaughterhouses and roads influence '$RR_i$' differently so the regression parameters were individualised. The prior distribution for these parameters was assumed to follow a normal distribution:

$$\beta_1 \sim N(0,\tau_1)$$
$$\beta_2 \sim N(0,\tau_2)$$

Where '$\tau_1$' and '$\tau_2$' are the precision of the prior distribution of '$\beta_1$' and '$\beta_2$' respectively.

The four random effects (*uncorrelated and correlated heterogeneity, distance to nearest slaughterhouse and conventional road*) and three covariates (*type of farm, mean number of positive sows and positive fattening animals)* were linearly added to the prior distribution of '$\mu_i$'.

$$\log(\mu_i) = \beta_0 + \beta_1 type_i + \beta_2 sows_i + \beta_3 fattening_i + \beta_{4i} slaughterhouse_i + \beta_{5i} road_i + a_i + b_i$$

The residuals of the model were computed as:

$$resid_i = O_i - p_i$$

Prior distributions were established for the parameters. Non-informative prior knowledge with a flat distribution was considered for the intercept ($\beta_0$) and normal distributions of mean 0 and variance 100,000 for covariate parameters (*mean number of positive sows and fattening animals and sows:fattening*). For the precision of the random covariates (*distance to nearest slaughterhouse and conventional road*) a gamma

distribution of mean 0.5 and variance 2,000 was used and a power distribution with mean 0 and variance -0.5 for the precision of the two random effects (*uncorrelated and correlated heterogeneity)* (Barceló et al., 2008).

### 3) Sintaxis of the ZINB model:

```
model {
  for (i in 1 : N) {
     mu[i,1]<-0
     mu[i,2]<-lambda[i]
     mu.i[i]<-mu[i,index[i]]
     index[i] ~ dcat(theta[1:2])
     # mixture
     O[i]  ~ dpois(mu.i[i])
     log(lambda[i])                                                                    <-
log(E[i])+alpha+b1*x0[i]+c11[i]*x11[i]+c31[i]*x31[i]+c35*x35[i]+c36*x36[i]+S[i]+eta[i]
     RR[i] <- mu.i[i]/E[i]
     PRP[i]<-step(RR[i]-1)
     resid[i]<-O[i]-mu.i[i]
     eta[i] ~ dnorm(0, tau.eta)}
     p ~ dbeta(1,1)
     theta[1]<-p
     theta[2]<-1-p

      for (i in 1 : N) {
      c11[i]~ dnorm(0, tau.c11)
      c31[i]~ dnorm(0, tau.c31)}

  # CAR prior distribution for spatial random effects:
  S[1:N] ~ car.normal(adj[], weights[], num[], tau.S)
  for(k in 1:sumNumNeigh) {
     weights[k] <- 1
  }
  # Other priors:
  alpha  ~dflat()
  b1 ~ dnorm(0,1.0E-6)
  tau.S <- pow(sdS,-2)
  sdS ~ dunif(0,5)
  tau.eta <- pow(sdeta,-2)
  sdeta ~ dunif(0,5)
  tau.c11  ~ dgamma(0.5, 0.0005)
  sdc11 <- sqrt(1 / tau.c11)
  tau.c31  ~ dgamma(0.5, 0.0005)
  sdc31 <- sqrt(1 / tau.c31)
  c35  ~ dnorm(0,1.0E-6)
  c36 ~ dnorm(0,1.0E-6)
}
```

**Interpretation:**

**alpha:** constant (intercept)

**b1, c11, c31, c35, c36:** Parameters of type of farm, distance to slaughterhouse, distance to roads, mean number of positive sows and mean number of positive fatteners covariates.

**x0, x11, x31, x35, x36:** Type of farm, distance to slaughterhouse, distance to roads, mean number of positive sows and mean number of fatteners covariates.

**S, eta:** random effects for the correlated and uncorrelated heterogeneity.

### *4) Parameter results*

For each period, the estimate values by the ZINB model, monte carlo error (MC error) and percentile distribution of the bayesian credible intervals of the intercept ($\beta_0$) and the fixed regression parameters of the type of farm ($\beta_1$), positive sows in the neighbourhood ($\beta_2$) and positive fattening in the neighbourhood ($\beta_3$) are presented in table 13.

| model ($\log(\mu_i)$) | Estimate | MC error | Percentile distribution | | | | |
|---|---|---|---|---|---|---|---|
| | | | **5%** | **10%** | **median** | **90%** | **95%** |
| **Period 1 (May 2004)** | **$\beta_0$: 10.47** | 0.1361 | **5.76** | 6.44 | 10.66 | 14.13 | **14.51** |
| | **$\beta_1$: 5.22** | 0.0791 | **1.59** | 2.43 | 5.23 | 7.05 | **7.24** |
| | $\beta_2$ : -0.32 | 0.0155 | -0.85 | -0.75 | -0.24 | 0.01 | 0.03 |
| | $\beta_3$ : 0.21 | 0.0106 | -0.07 | -0.05 | 0.19 | 0.53 | 0.61 |
| **Period 2.1 (May 2005)** | **$\beta_0$: 12.67** | 0.1714 | **6.66** | 7.54 | 13.37 | 16.99 | **17.26** |
| | **$\beta_1$: 6.52** | 0.0975 | **2.51** | 3.67 | 7.07 | 8.79 | **9.08** |
| | $\beta_2$ : 0.72 | 0.0242 | -0.05 | 0.01 | 0.75 | 1.35 | 1.39 |
| | **$\beta_3$ : 1.22** | 0.0242 | **0.41** | 0.58 | 1.26 | 1.87 | **1.95** |
| **Period 2.2 (May 2006)** | **$\beta_0$: 10.77** | 0.1417 | **5.31** | 6.41 | 11.03 | 14.17 | **15.12** |
| | **$\beta_1$: 7.48** | 0.1076 | **3.32** | 4.07 | 7.63 | 10.37 | **10.76** |
| | **$\beta_2$ : 3.34** | 0.0635 | **1.37** | 1.59 | 3.39 | 5.08 | **5.41** |
| | $\beta_3$ : -0.001 | 0.0216 | -0.74 | -0.67 | 0.11 | 0.51 | 0.60 |
| **Period 3 (May 2006)** | *$\beta_0$: -0.03* | 0.0313 | -1.04 | -0.86 | -0.14 | 0.96 | 1.09 |
| | $\beta_1$: 0.82 | 0.0547 | -0.56 | -0.46 | 0.44 | 2.37 | 2.46 |
| | **$\beta_2$ : 0.37** | 0.0099 | **0.12** | 0.15 | 0.32 | 0.68 | **0.76** |
| | $\beta_3$ : -0.26 | 0.0089 | -0.56 | -0.51 | -0.24 | -0.02 | 0.03 |

Table 13. Estimate values, monte carlo error (MC error) and percentile distribution for the bayesian credible intervals of the intercept ($\beta_0$) and the fixed regression parameters of the type of farm ($\beta_1$), positive sows in the neighbourhood ($\beta_2$) and positive fattening in the neighbourhood ($\beta_3$) of the ZINB model ($\log(\mu_i)$). In bold are represented those parameters which influence the AD relative risk of sow farms.

In the first period, 258 out of 2,027 (12.7%) of the individualised regression parameters of the slaughterhouse covariate were related to AD risk. In period 2.1, this proportion was of 8.5% (127 out of 1,485), in period 2.2 it fell to 5.8% (94 out of 1,612) and in period 3 all the individualised bayesian credible intervals of the regression coeffcients contained the zero value. In this period, there was any relationship between the distance from each farm to the nearest slaughterhouse and AD risk.

In relation to the distance to the nearest conventional road, in the first period 87 (4.3%) sow farms presented an AD risk related with this covariate. This proportion was 3.5% (52 farms) in period 2.1; 1.5% (25 farms) in period 2.2 and zero in period 3.

Figures 21 to 24 show the density and location of sow farms where the distance to the nearest slaughterhouse or conventional road was related to the AD risk

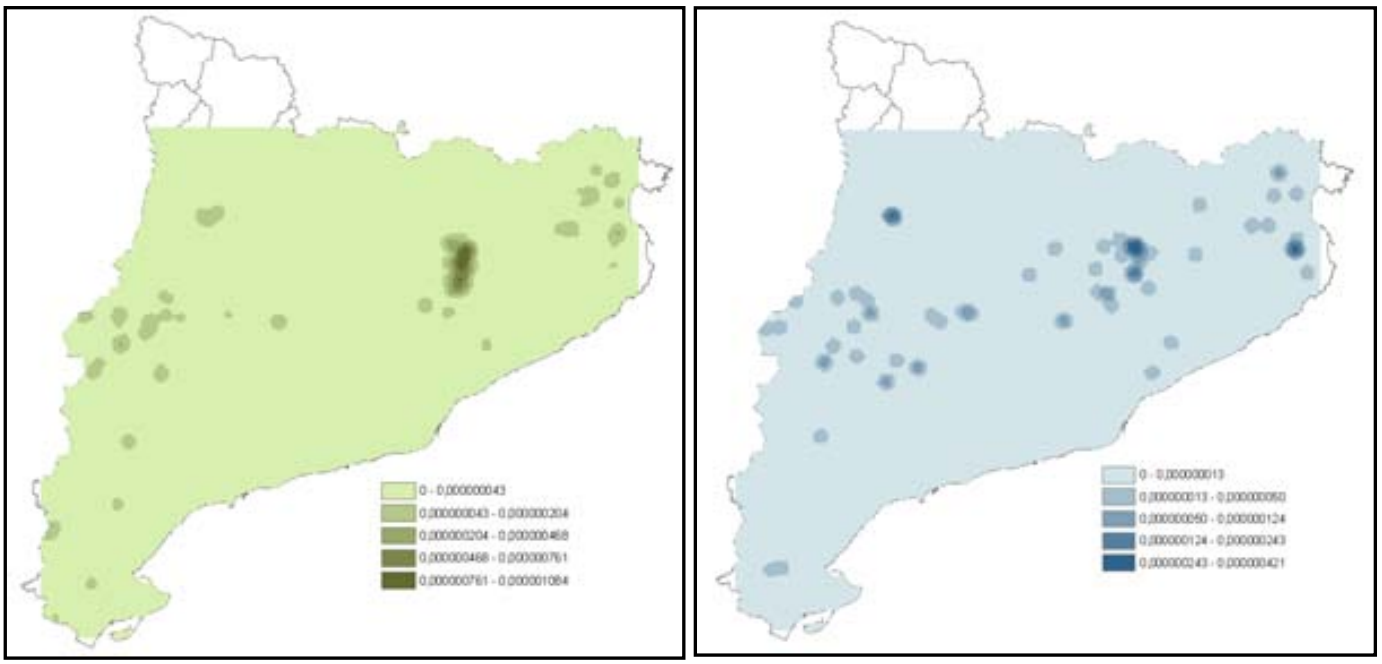

Figure 21. Kernell density surface of sow farms (calculated with a bandwidth of 5km) where the distance to the nearest slaughterhouse (left) or conventional road (right) was related to the AD risk in period 1 (May 2004).

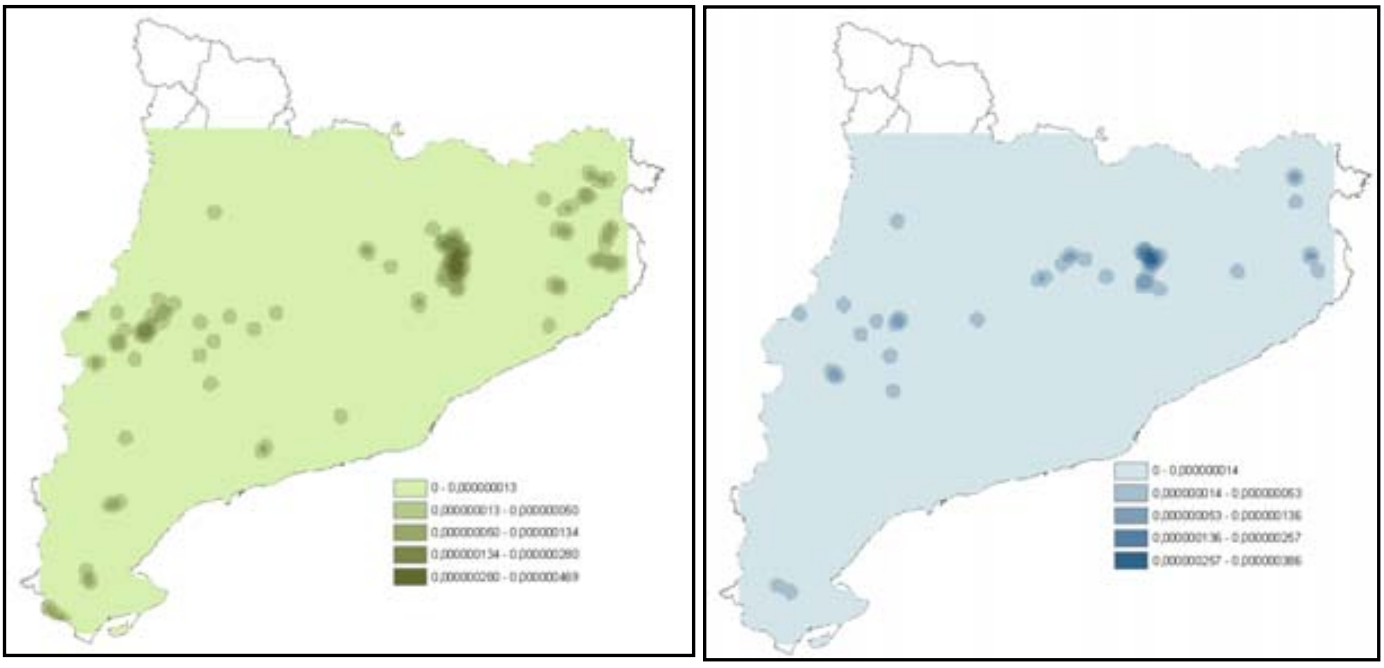

Figure 22 Kernell density surface of sow farms (calculated with a bandwidth of 5km) where the distance to the nearest slaughterhouse (left) or conventional road (right) was related to the AD risk in period 2.1 (May 2005).

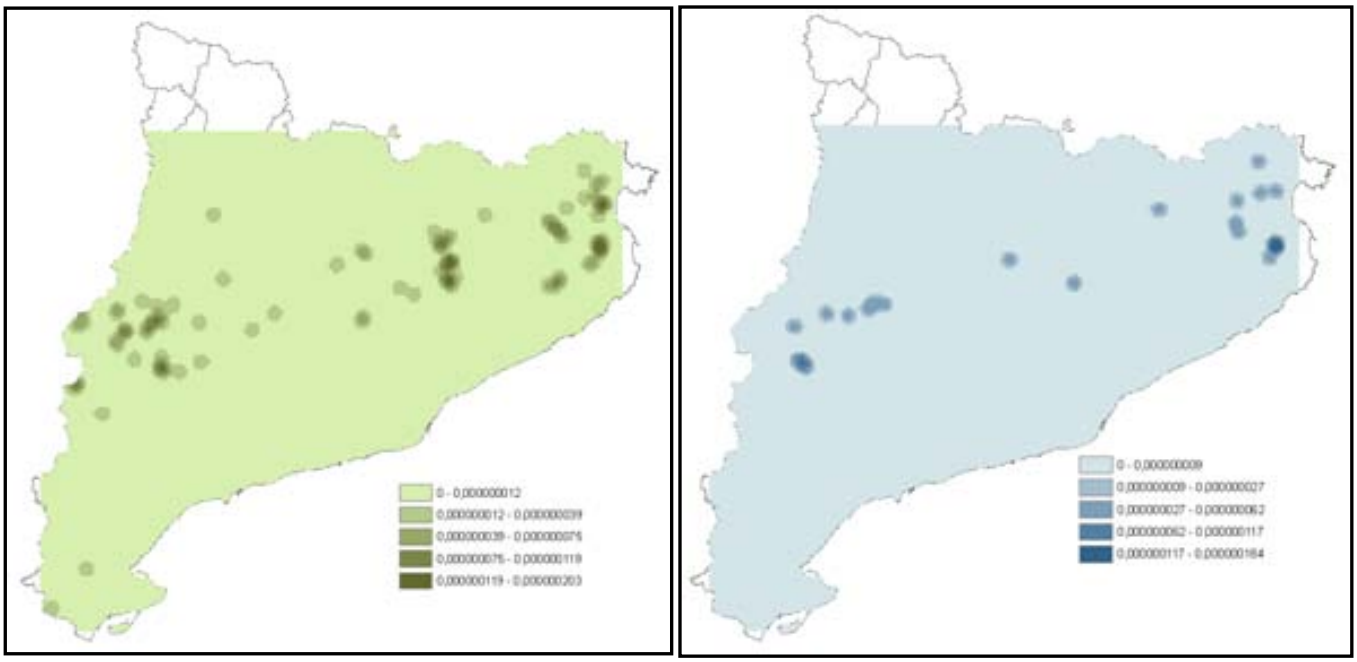

Figure 23. Kernell density surface of sow farms (calculated with a bandwidth of 5km) where the distance to the nearest slaughterhouse (left) or conventional road (right) was related to the AD risk in period 2.2 (May 2006).

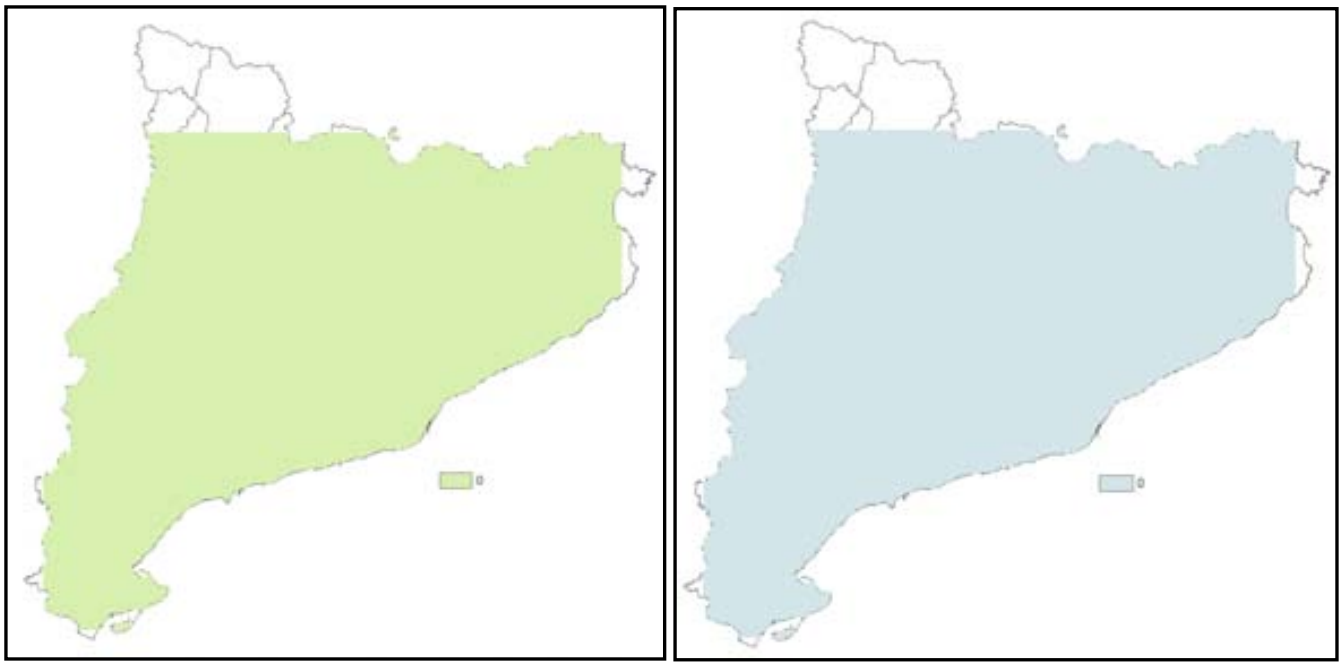

Figure 24. Kernell density surface of sow farms (calculated with a bandwidth of 5km) where the distance to the nearest slaughterhouse (left) or conventional road (right) was related to the AD risk in period 3 (May 2007).

The estimate of the standard deviation, MC error and percentile distribution of the bayesian credible intervals of the regression parameters of the explanatory variables included as random effects (distance to nearest slaughterhouse ($\beta_4$) and distance to nearest road ($\beta_5$) and the random effects for the uncorrelated heterogeneity (a) and correlated heterogeneity (s) are shown in table 14.

| Model ($\log(\mu_i)$) | Estimate (sd) | MC error | Percentile distribution | | |
|---|---|---|---|---|---|
| | | | **5%** | **median** | **95%** |
| **Period 1 (May 2004)** | $\beta_4$: 27.9 | 0.5634 | 10.48 | 27.8 | 47.24 |
| | $\beta_5$: 19.19 | 0.3433 | 8.12 | 19.55 | 31.08 |
| | *a*: 4.98 | 0.0002 | 4.95 | 4.99 | 5 |
| | *s*: 4.95 | 0.0007 | 4.86 | 4.97 | 4.99 |
| **Period 2.1 (May 2005)** | $\beta_4$: 28.11 | 0.4426 | 13.47 | 27.28 | 41.69 |
| | $\beta_5$: 19.16 | 0.2893 | 9.07 | 18.65 | 28.1 |
| | *a*: 4.97 | 0.0006 | 4.92 | 4.98 | 4.99 |
| | *s*: 4.92 | 0.0016 | 4.76 | 4.94 | 4.99 |
| **Period 2.2 (May 2006)** | $\beta_4$: 43.14 | 0.7521 | 15.16 | 45.47 | 66.65 |
| | $\beta_5$: 27.71 | 0.4187 | 12.68 | 28.62 | 42.14 |
| | *a*: 4.96 | 0.0011 | 4.88 | 4.97 | 42.14 |
| | *s*: 4.82 | 0.0067 | 4.46 | 4.88 | 4.99 |
| **Period 3 (May 2007)** | $\beta_4$: 2.31 | 0.0644 | 0.77 | 1.83 | 4.93 |
| | $\beta_5$: 1.32 | 0.0483 | 0.26 | 1.12 | 3.24 |
| | *a*: 2.37 | 0.0334 | 1.31 | 2.33 | 3.46 |
| | *s*: 0.90 | 0.0323 | 0.20 | 0.71 | 2.15 |

Table 14. Estimate of the standard deviation, MC error and percentile distribution of the regression parameters of the explanatory variables included as random effects (distance to nearest slaughterhouse ($\beta_4$) and distance to nearest road ($\beta_5$)) and the random effects for the uncorrelated heterogeneity (a) and correlated heterogeneity (s) of the ZINB model for the different periods of study.

The relative risk values increased during the different periods of study. In the first period, these values increased until 6.24 whereas in period 3 reach the value of 29.82. As mentioned before, the residuals were used to evidence those areas where the AD risk of each sow farm was less explained by the covariates included in the model.
The distribution of the relative risk values and the residuals of the model are shown in figures 25 to 28.

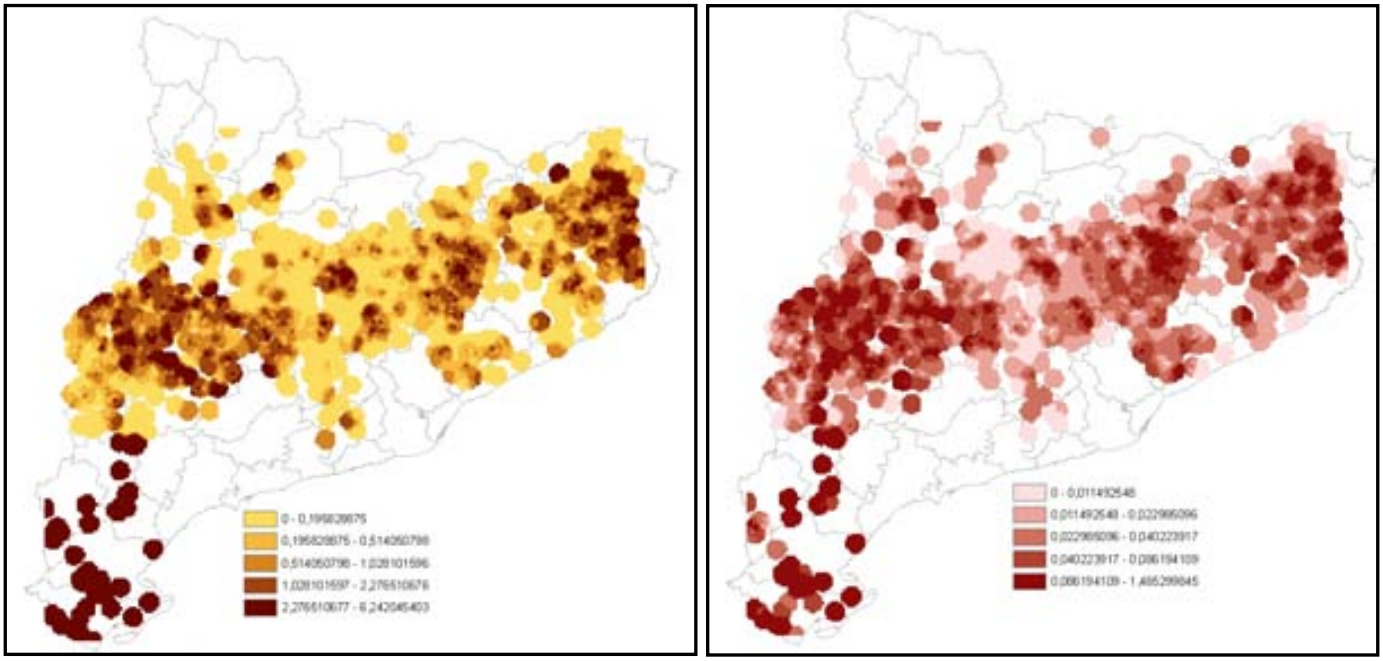

Figure 25. Relative risk values (left) and residuals (right) for period 1 (May 2004).

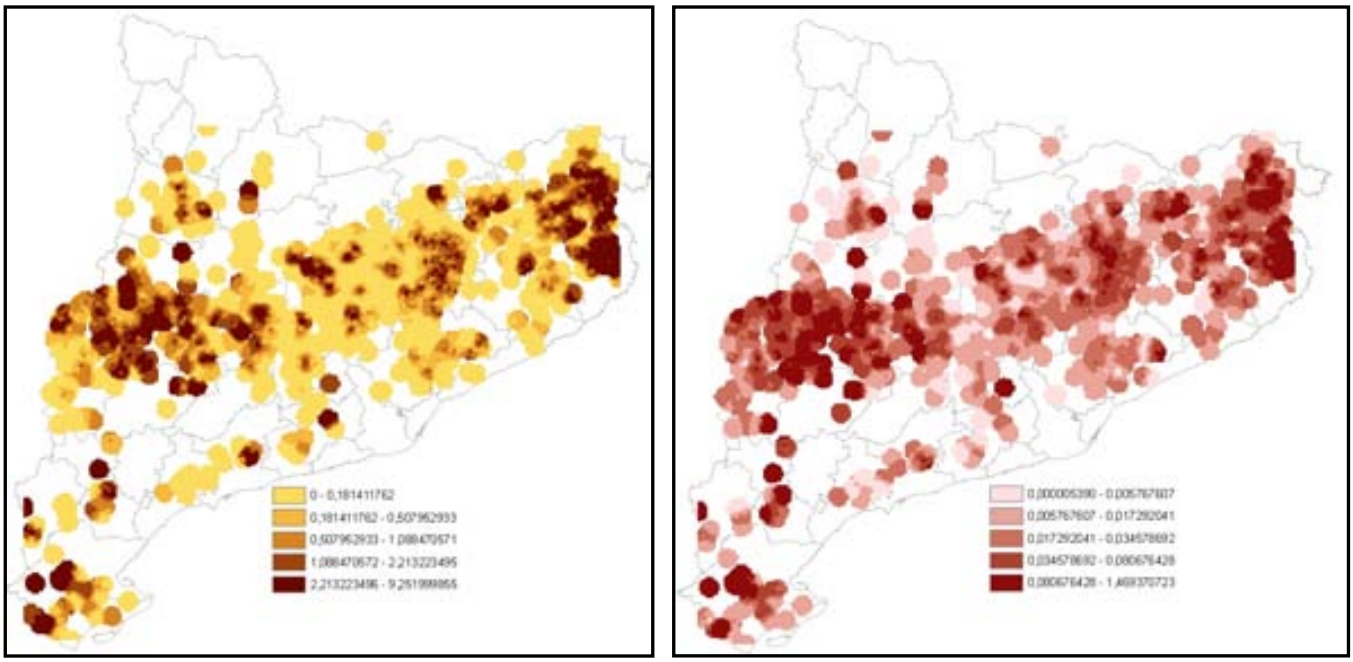

Figure 26. Relative risk values (left) and residuals (right) for period 2.1 (May 2005).

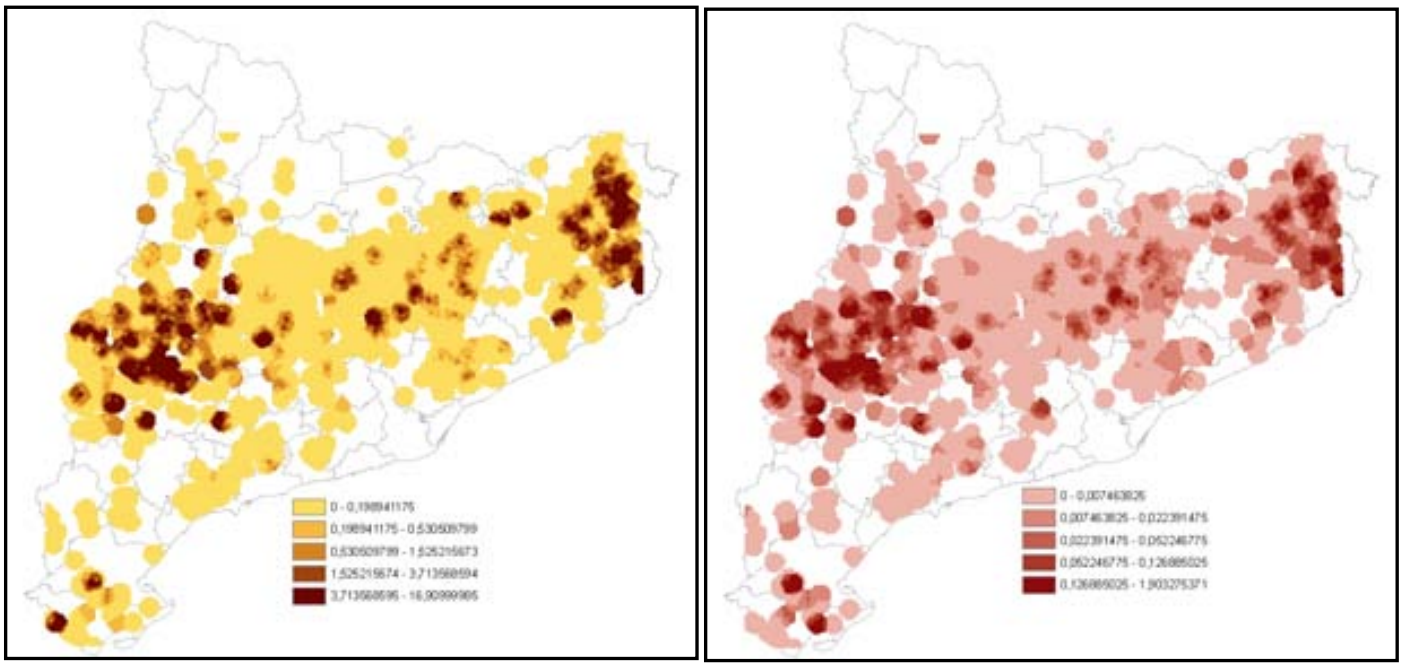


Figure 27. Relative risk values (left) and residuals (right) for period 2.2 (May 2006).

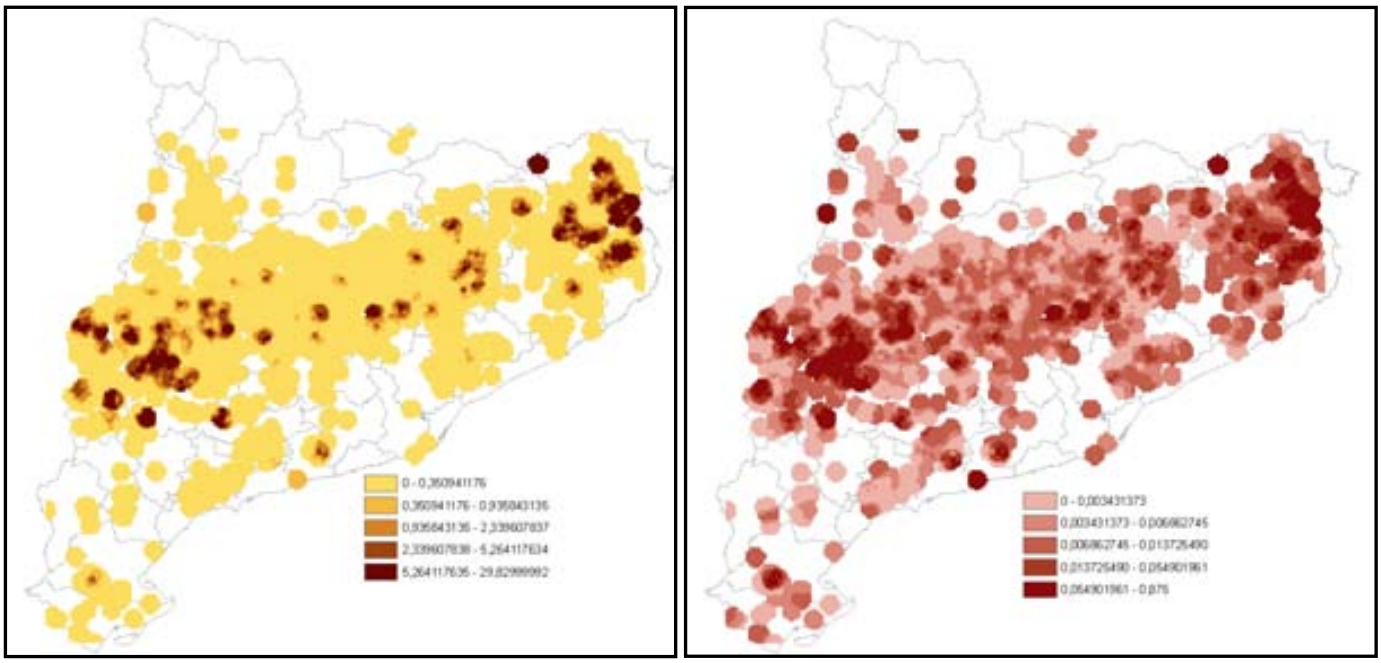


Figure 28. Relative risk values (left) and residuals (right) for period 3 (May 2007).

***5) Convergence diagnostic:***

In figures 29 to 32, trace plots for the deviance of the models and the Gelman-Rubin convergence diagnostic for each period are represented. Convergence has been reached; as trace plots look like a random scatter about a stable mean value and the Gelman-Rubin convergence diagnostic is close to 1.

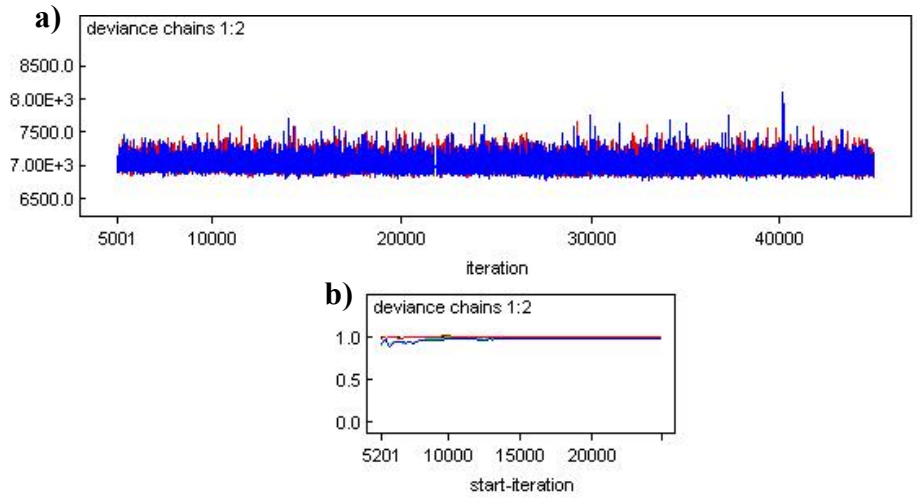


Figure 29 Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 1 (May 2004).

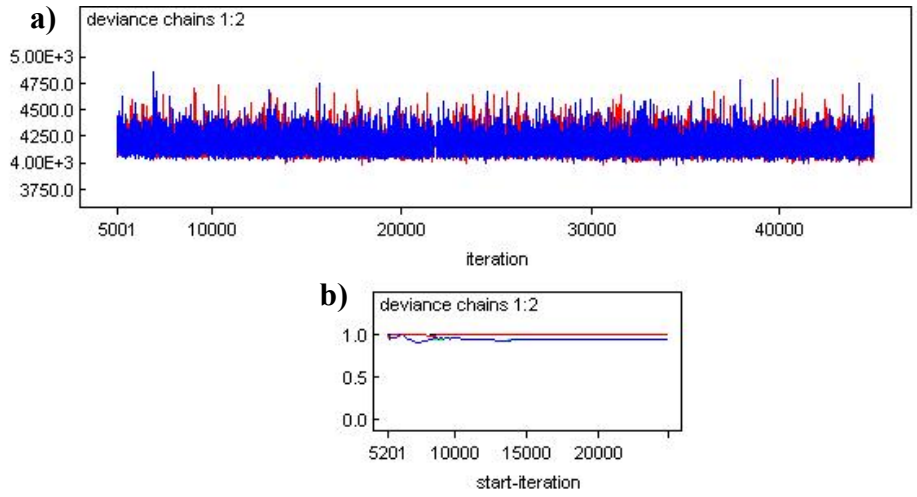


Figure 30. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 2.1 (May 2005).

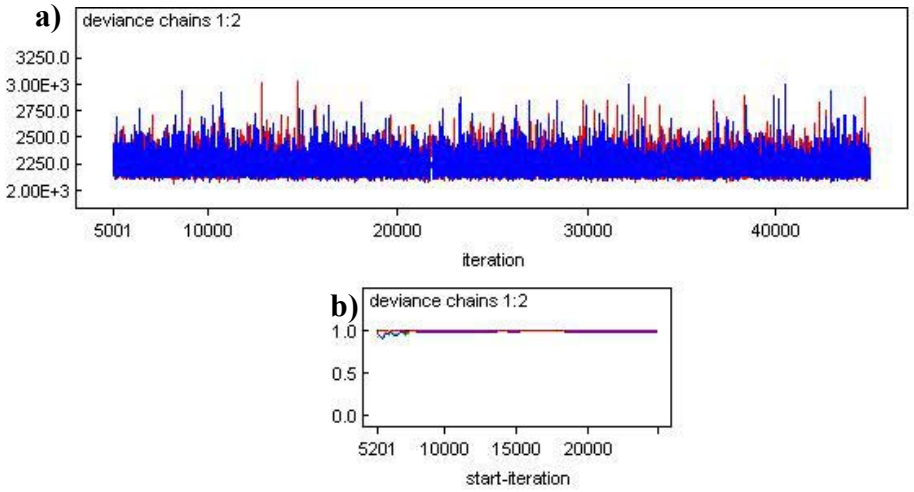


Figure 31. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 2.2 (May 2006).

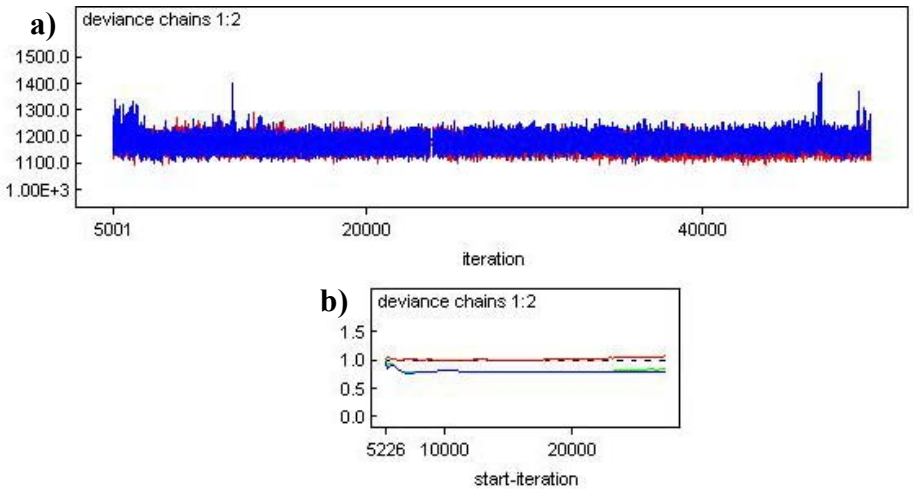


Figure 32. Trace plots (a) and Gelman-Rubin convergence diagnostic (b) of the deviance for period 3 (May 2007).