Geographical distribution modelling for *Neospora caninum* and *Coxiella burnetii* infections in dairy cattle farms in northeastern Spain

C. NOGAREDA¹, A. JUBERT², V. KANTZOURA³, M. K. KOUAM³, H. FEIDAS⁴ AND G. THEODOROPOULOS²*¹

¹ Department of Animal Production, Centre for Research in Animal Production (CIPA) University of Lleida, Lleida, Spain
² Interprofessional Dairy Laboratory of Catalonia (ALLIC), Cabrils, Spain
³ Department of Anatomy and Physiology of Farm Animals, Faculty of Animal Science and Hydrobiology, Agricultural University of Athens, Votanikos, Athens, Greece
⁴ Department of Meteorology and Climatology, School of Geology, Aristotle University of Thessaloniki, Thessaloniki, Greece

Received 2 October 2011; Final revision 10 January 2012; Accepted 2 February 2012; first published online 28 February 2012

SUMMARY

Maximum entropy ecological niche modelling and spatial scan statistic were utilized to predict the geographical range and to investigate clusters of infections for *Neospora caninum* and *Coxiella burnetii* in dairy cattle farms in Catalonia, northeastern Spain, using the Maxent and SaTScan programs, respectively. The geographical distribution of *Neospora* and *Coxiella* with the highest level of probability ($P > 0.60$) covers central Catalonia and spreads towards northeastern Catalonia which contains a high concentration of dairy cattle farms. The most important environmental factor that contributed to the ecological niche modelling was precipitation of driest month followed by elevation. Significant clusters ($P < 0.001$) were detected for *Neospora* and *Coxiella* infections in the western and eastern regions of Catalonia, respectively.

Key words: Cattle, *Coxiella burnetii*, ecological niche modelling, dairy farms, *Neospora caninum*.

INTRODUCTION

*Neospora caninum* is an obligate intracellular protozoan and *Coxiella burnetii* is an intracellular Gram-negative bacterium. Both pathogens cause abortions in cattle worldwide [1–3]. Cattle infection with *N. caninum* is facilitated through the ingestion of sporulated oocysts in the environment passed in the faeces of the definitive carnivore host [4]. Cattle infection with *C. burnetii* is facilitated through contact with birth products (birth fluids, placenta), vaginal mucus, milk, faeces, urine and semen of *C. burnetii*-shedding cattle [1, 5]. A common characteristic of *Neospora* and *Coxiella* infections is that in both cases cattle are infected by the respective pathogens present in the environment. Therefore, the presence of these pathogens in the environment makes their spread and geographical distribution susceptible to environmental conditions.

Ecological niche modelling offers the opportunity to derive predictive distribution maps from species occurrence and environmental data [6, 7]. In the present study a cross-sectional survey was conducted...
in dairy cattle farms in Catalonia, northeastern Spain to collect information on the spatial distribution of *N. caninum*-infected and *C. burnetii*-infected dairy farms in order to construct predictive maps of Catalonia for the two pathogens, based on ecological niche modelling, and to investigate clusters of infections.

**MATERIALS AND METHODS**

**Farms and sampled areas**

Bulk tank milk (BTM) samples were collected on the same day in October 2010 from dairy cattle farms in Catalonia, northeastern Spain during daily quality BTM control of all Catalan dairy farms. Catalonia is divided into 41 regions and at the time of the study there were 814 dairy farms which were located in 15 regions [8]. Farms were randomly sampled within each region (10–12%) in order to be representative of all areas and the coordinates of the location of each farm were recorded in a database. The association between the prevalence of *Neospora*-positive and *Coxiella*-positive farms was estimated using Fisher’s exact test (*P* < 0.05) and Cramer’s *V*² test. Cramer’s *V*² test measures the association between two categorical variables. A value of zero indicates that there is no association. A value of 1 indicates that there is a perfect association.

**Anti-*N. caninum*** antibody detection in BTM

A commercial indirect enzyme-linked immunosorbent assay (ELISA) kit (Civtest Bovis Neospora, Spain), based on the whole tachyzoite lysate of *N. caninum* NC-1 was used to determine antibodies against *N. caninum*. The skimmed milk samples were centrifuged at 1000 *g* for 15 min at 4 °C. Absorbance was measured as optical density (OD) values at 405 nm using a microplate reader (Multiskan FC, Finland). A relative index percent (RIPC) per BTM sample was obtained according to the manufacturer’s procedure:

\[
\text{RIPC} = \frac{\text{OD sample} - \text{OD negative control}}{\text{OD positive control} - \text{OD negative control}} \times 100.
\]

The kit manufacturer considers individual cow prevalence of antibodies against *N. caninum* to be less than 5–10% when BTM RIPC is ≤ 3·0 and greater than 10% when BTM RIPC is > 3·0. Comparing results of BTM with individual serology, it was found that at the cut-off point of 3·0 the sensitivity and specificity of the test were 60% and 91%, respectively. For analysis of data, BTM was considered negative when RIPC ≤ 3·0 and positive when RIPC > 3·0.

**Anti-*C. burnetii*** antibody detection in BTM

A commercial indirect ELISA LSIVET Ruminant Milk/Serum Q Fever (CoxLS kit; Laboratoire Service International, France) was used to determine antibodies to *C. burnetii* in the BTM samples. The milk samples were skimmed and the test was performed according to the manufacturer’s instructions. The antigen of the ELISA CoxLS kit has been isolated from domestic ruminants by INRA (France). Absorbance was measured as OD values at 450 nm using a microplate reader (Multiskan FC). The results were expressed as a S/P ratio calculated as follows:

\[
\text{S/P} = \frac{\text{OD sample} - \text{OD negative control}}{\text{OD positive control} - \text{OD negative control}}
\]

The BTM results were expressed as titres (titre = S/P × 100). BTM was considered negative when the titre was ≤ 30 and positive when the titre was > 30 [9].

**Environmental data**

The environmental datasets obtained for this study were climate, elevation, and land cover data. These datasets were converted to a common projection, map extent and resolution prior to use in the modelling program. WorldClim version 1·4 climate data [10] was obtained from the WorldClim website (http://www.worldclim.org). WorldClim is a set of global climate layers (climate grids) with a spatial resolution of 1 km². They can be used for mapping and spatial modelling in a geographical information system (GIS) or other computer programs. The data layers were generated through interpolation of average monthly climate data from weather stations on a 30 arc-second (1 km²) resolution grid. Variables included are monthly total precipitation, and monthly mean, minimum and maximum temperature. The data are further processed into a series of bioclimatic variables (Table 1). Input data were gathered from a variety of sources and, where possible, were restricted to records from the period 1950–2000. The thin-plate smoothing spline algorithm implemented in the ANUSPLIN package (Centre for Resources and Environmental Studies at the Australian National University,
Australia) was used for interpolation, using latitude, longitude, and elevation as independent variables. The database is documented in detail in Hijmans et al. [10]. The bioclimatic variables with an approximate spatial resolution of 1 km² were used for this project.

Elevation data were also downloaded from the WorldClim website. WorldClim processed this dataset from NASA Shuttle Radar Topography Mission (SRTM) data to have the same projection and resolution as the other WorldClim layers. SRTM obtained elevation data on a near-global scale to generate the most complete high-resolution digital topographic database of Earth. SRTM consisted of a specially modified radar system that flew onboard the Space Shuttle Endeavour during an 11-day mission in February 2000. Land-cover data were downloaded from the United States Geological Survey’s (USGS) Global Land Cover Characteristics (GLCC) database, version 2 Global. It is a 1 km resolution global land-cover characteristics database generated for use in a wide range of environmental research and modelling applications.

The land-cover data were developed from 1 km advanced very high-resolution radiometer (AVHRR) satellite data from April 1992 to March 1993. More precisely, the 1 km AVHRR normalized differential vegetation index (NDVI) composites are the core dataset used in land-cover characterization. In addition, other key geographical data include digital elevation data, ecoregion interpretations and country-level or regional-level vegetation and land-cover maps. Land-cover data are available in multiple classification schemes; the Global Ecosystems land cover classification, which contains 100 classes, was used for the modelling work. This global database was used because it is not geographically limited and is available on the USGS data portal for most of the countries in the world. In this way, the presented application can be easily reproduced worldwide, producing comparable results. Moreover, the GLCC database is compatible with the WorldClim global climate layers used in this study having the same spatial resolution (1 km²).

### Modelling algorithm

Model building was performed using the maximum entropy method implemented in the Maxent program [6]. The Maxent program develops models of species’ distributions using species’ presence data and environmental data [6, 11, 12]. Maxent is based on the idea of estimating a target probability distribution by finding the probability distribution of maximum entropy (i.e. the most spread out, or closest to uniform), subject to a set of constraints that represent the incomplete presence-only data information about the target distribution. The constraints are that the expected value of each environmental variable should match its empirical average. A detailed description of the Maxent algorithm is given in Philips et al. [6]. Maxent has been used for a number of studies and has been shown to be a high performing modelling program [7, 13].

### Model performance evaluation and contribution of environmental factors to the models

The model performance was evaluated using the threshold independent method based on the area under the curve (AUC) of receiver-operating characteristics curve (ROC). The AUC is computed through a ROC plot of sensitivity (the true-positive fraction) against 1 – specificity (the false-positive fraction) [14]. The AUC value ranges from 0.5 (random accuracy) to a maximum value of 1 (perfect discrimination). In order to determine which variables contribute most to the model’s development, the Maxent program was

<table>
<thead>
<tr>
<th>Bioclimatic variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIO1</td>
<td>Annual mean temperature</td>
</tr>
<tr>
<td>BIO2</td>
<td>Mean diurnal range (mean of monthly (max temp – min temp))</td>
</tr>
<tr>
<td>BIO3</td>
<td>Isothermality (BIO2/BIO7) (× 100)</td>
</tr>
<tr>
<td>BIO4</td>
<td>Temperature seasonality (standard deviation × 100)</td>
</tr>
<tr>
<td>BIO5</td>
<td>Max temperature of warmest month</td>
</tr>
<tr>
<td>BIO6</td>
<td>Min temperature of coldest month</td>
</tr>
<tr>
<td>BIO7</td>
<td>Temperature annual range (BIO5-BIO6)</td>
</tr>
<tr>
<td>BIO8</td>
<td>Mean temperature of wettest quarter</td>
</tr>
<tr>
<td>BIO9</td>
<td>Mean temperature of driest quarter</td>
</tr>
<tr>
<td>BIO10</td>
<td>Mean temperature of warmest quarter</td>
</tr>
<tr>
<td>BIO11</td>
<td>Mean temperature of coldest quarter</td>
</tr>
<tr>
<td>BIO12</td>
<td>Annual precipitation</td>
</tr>
<tr>
<td>BIO13</td>
<td>Precipitation of wettest month</td>
</tr>
<tr>
<td>BIO14</td>
<td>Precipitation of driest month</td>
</tr>
<tr>
<td>BIO15</td>
<td>Precipitation seasonality (coefficient of variation)</td>
</tr>
<tr>
<td>BIO16</td>
<td>Precipitation of wettest quarter</td>
</tr>
<tr>
<td>BIO17</td>
<td>Precipitation of driest quarter</td>
</tr>
<tr>
<td>BIO18</td>
<td>Precipitation of warmest quarter</td>
</tr>
<tr>
<td>BIO19</td>
<td>Precipitation of coldest quarter</td>
</tr>
</tbody>
</table>
set to perform the jackknife tests in which the method is run multiple times: (i) using all variables, (ii) dropping one variable at a time, and (iii) running the model using only one variable. Variables which produced the highest training gains or reduced the training gain when omitted from the model are considered to be the most important variables. Strong correlation between variables can be located in curves showing how each environmental variable affects the Maxent prediction disregarding all other variables. The GIS software ArcGIS v. 9.2 (ESRI Inc., USA) was used to display the prediction results and represent the sampled localities.

**Spatial cluster analysis**

The spatial scan statistic implemented in SaTScan software (version 8.0; SaTScan™, USA) was used to investigate geographical clusters of infection. The concept of a spatial scan statistic is based on the generalization of a test probability [15–17]. The spatial scan statistic uses a circular window of variable radius that moves across the map to represent potential geographical clusters. The radius of the cluster varies from zero up to a specified maximum value. By gradually changing the circle centre and radius, the window scans the geographical areas for potential localized clusters without incorporating prior assumptions about their size and location and notes the number of observed and expected observations inside the window at each location. The test of significance is based on the likelihood ratio test for which the window with the maximum likelihood is the most likely cluster [16]. The assessment of a cluster is performed by comparing the number of cases (infection) within the circle with the number of expected cases under the assumption that cases are randomly distributed in the space. The $P$ value is obtained through Monte Carlo hypothesis testing [18]. The spatial scan statistic adjusts for spatial variations in the density of the population in the study area [17]. The detection of clusters in the present study was performed under the Bernoulli probability model using the maximum cluster size of 50% of the total population for *Neospora* and *Coxiella* infections separately. Test-positive farms were considered as cases while test-negative farms were regarded as controls. The number of simulations for Monte Carlo testing was set to 9999. For each window of varying position and size, the SaTScan program tested the risk of *Neospora* and *Coxiella* infection each within and outside the window, with the null hypothesis of equal risk.

**RESULTS**

A total of 90 (11%) dairy cattle farms were sampled. The prevalence of *Neospora*-positive and *Coxiella*-positive farms in the examined farms was 90% and 74%, respectively. No association between the prevalence of *Neospora*-positive and *Coxiella*-positive farms was found (Fisher’s exact test: $P=0.22$, Cramer’s $V=0.02$). Figures 1 and 2 display the potential geographical distribution of *Neospora*-positive and *Coxiella*-positive farms predicted by the Maxent program, respectively. The geographical distribution of *Neospora*-positive and *Coxiella*-positive farms with the highest level of probability ($P>0.60$) covers central Catalonia and spreads towards northeastern Catalonia.

**Model validation**

The AUC values for the models range from 0.886 to 0.926, indicating that the models are good to very good (Table 2). Both models have low $P$ values (Table 2) using minimum training presence as the threshold with *Neospora*-positive farms having the lowest ($P=0.0001$) indicating the best prediction.

**Effect of environmental variables on the model**

For *Neospora*-positive and *Coxiella*-positive farms, ‘precipitation of driest month’ (BIO14) followed by elevation were the variables that achieved the highest training gain when used to build a model with no other variables (Fig. 3). Other environmental variables that show training gain when modelling with only a single variable included: ‘precipitation seasonality’ (BIO15), land cover, ‘mean temperature of warmest quarter’ (BIO10), ‘annual mean temperature’ (BIO1), ‘precipitation of wettest quarter’ (BIO16), ‘mean temperature of wettest quarter’ (BIO8), and ‘mean temperature of coldest quarter’ (BIO11), for *Neospora*-positive farms, and ‘precipitation of driest month’ (BIO14), elevation, ‘mean temperature of warmest quarter’ (BIO10), ‘precipitation seasonality’ (BIO15), ‘precipitation of wettest quarter’ (BIO16), ‘mean temperature of driest quarter’ (BIO9), ‘precipitation of coldest quarter’ (BIO19), ‘mean temperature of wettest quarter’ (BIO8), and land cover for *Coxiella*-positive farms (Fig. 3).

Clear or almost clear linear (positive or negative) or no linear correlation between predicted suitability and environmental variable are indicated in Figure 3.
Land-cover types that occurred at the species sampling areas are shown in Table 3. The primary land covers are similar for both *Neospora*-positive and *Coxiella*-positive farms and include mostly Mediterranean scrub followed by cool crops and towns.

**Fig. 1.** [colour online]. Predicted geographical distribution ranges for *Neospora caninum*. The circle encloses the location identified within the most likely cluster.

**Fig. 2.** [colour online]. Predicted geographical distribution ranges for *Coxiella burnetii*. The circle encloses the location identified within the most likely cluster.

Read the full text at [https://doi.org/10.1017/S0950268812000271](https://doi.org/10.1017/S0950268812000271)
The results of the spatial scan statistic analyses showed one most likely cluster of infected farms ($P < 0.05$) for each pathogen. The relative risk (RR) of the most likely cluster was 1.25 and 1.57 for Neospora-positive and Coxiella-positive farms, respectively (Table 4).

**DISCUSSION**

Dairy cattle farms are an important source of income for the economy of Catalonia region in northeastern Spain. Despite the low number of farms, Catalonia is one of the most dynamic dairy regions of Spain [19]. The samples for testing were chosen randomly within each region of Catalonia (10–12%) in order to be representative of the overall area. Abortions due to *N. caninum* and *C. burnetii* have a grave impact on the reproductive performance of dairy cattle with major consequences on the profitability of cow production [20, 21]. The seroprevalence of infections due to *N. caninum* and *C. burnetii* has been reported to be 83% in Spanish farms [22] and 61% in Catalan farms [23], respectively. In Catalonia, the risk of abortion was found to be 12–19 times higher in Neospora-seropositive dairy cows than in their seronegative partners [21, 24]. No similar data is available for Coxiella abortions.

Even though *N. caninum* can be transmitted vertically when tachyzoites are transmitted from an infected dam to her foetus during pregnancy [4, 25], the ingestion of oocysts is the only demonstrable mode for horizontal transmission in herbivores [4] and the mode of first introduction of the disease to the farm.

*N. caninum* [26, 27] and *C. burnetii* [9] analysis of antibodies using ELISA in BTM has been demonstrated to be very useful in estimating the prevalence...
of these pathogens in dairy farms. However, as the volume of milk that each lactating cow provides to the tank is variable throughout the year, the prevalence value obtained from BTM has to be considered as a guiding result. The BTM samples used in the present study were the same routinely obtained by the Catalan dairy industry for milk quality testing and enabled detection of Neospora and Coxiella infection at an affordable price. This available, reliable, and simple to use diagnostic technique for screening and monitoring large number of cows can be very useful for bovine neosporosporosis and coxiellosis implementation control programmes [26, 27], but individual milk or serum analysis should be performed in order to better control the positive farms.

The present study reports the potential geographical range of N. caninum and C. burnetii in northeastern Spain using the Maxent program. Maxent is a mathematical framework that can be applied for several species at once and/or for spatially structured communities [28]. There have been no previous attempts to model Neospora and Coxiella distribution using GIS.

Papes & Baubert [29] found that the minimum number of records needed for accurate modelling with Maxent is around 15. Pearson et al. [30] found that models built with as few as five records showed some statistical significance using Maxent. It is important to mention that by modelling the ecological niche for the two pathogens in the present study, it does not mean that the models represent their actual or realized distribution. These models should be considered as an initial effort to approximate their potential geographical range.

The results of the present study indicate that various bioclimatic variables, and most importantly, precipitation of driest month (BIO14), precipitation seasonality (BIO15) and elevation correlate similarly with the geographical distribution of Neospora and Coxiella. In two previous studies on climate effects on the risk of Neospora in cattle in Germany [31] and Italy [32] it was found that ‘mean temperature in July in the municipality where the herd is localized’ and ‘mean temperature in spring in a buffer zone around farm location’ were identified as risk factors, respectively. In addition, two previous studies on climate effects on abortions due to Neospora in The Netherlands [33] and California [34] indicated that mild temperatures and humidity increase the risk of infection. Other studies on climate effects on the risk of Coxiella in cattle showed that the number of antibody-positive cows and their antibody titre in Japan were significantly elevated in winter and decreased in summer [35], while in Italy all the C. burnetii-related abortions were recorded between October and April [36]. These observations for Neospora and Coxiella may be attributed to the effect of climate on the rate of sporulation or survival of Neospora oocysts and the preservation of Coxiella bacteria in the environment.

Another finding of the present study is that Mediterranean scrub followed by cool crops and towns are the primary land covers for the geographical distribution of both Neospora-infected and Coxiella-infected farms. Farm location or proximity to a city, town, or village has been found to be a risk factor for N. caninum-associated abortion in Switzerland [37] and N. caninum positivity in bulk milk ELISA [31]. The observed land covers that favour the geographical distribution of Neospora may be associated with high human population density which is correlated with a high density of dogs [32] that serve as reservoirs of Neospora. In the case of Coxiella, the observed favourable land covers may be associated with the presence of ticks that have an active role in maintaining the disease in wild and peridomestic cycles [38] or other wild ungulates [39].

---

### Table 4. Significant clusters of Neospora caninum-infected and Coxiella burnetii-infected dairy farms

<table>
<thead>
<tr>
<th>Species</th>
<th>Cluster</th>
<th>Location (coordinates) [radius]</th>
<th>No. of cases</th>
<th>Expected cases</th>
<th>Relative risk</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neospora caninum</td>
<td>Most likely</td>
<td>Almacelles (41°738712′ N, 0°424497′ E) [158.08 km]</td>
<td>45</td>
<td>40-50</td>
<td>1.25</td>
<td>0.01</td>
</tr>
<tr>
<td>Coxiella burnetii</td>
<td>Most likely</td>
<td>Caldes de Montbui (41°636959′ N, 2°158792′ E) [64-16 km]</td>
<td>35</td>
<td>27.54</td>
<td>1.57</td>
<td>0.01</td>
</tr>
</tbody>
</table>
The areas where significant clusters of *Neospora*-positive and *Coxiella*-positive farms were observed are centred on Almacelles and Caldes de Montbuy, respectively. Almacelles is an irrigated flat area with fruit trees and forage crop production with a medium density of big dairy farms and a high density of beef feedlots. Dogs and other small wild animals are present in this area facilitating the spread of neosporosis to cattle [40, 41]. The area of Caldes de Montbuy is surrounded by mountains and is very close to two natural parks (Parc natural del Montseny and Parc natural de Sant Llorenç del Munt i l’Obac). This area is associated with wild animals and ticks that can carry coxiellosis [38].

The most likely cluster for *Neospora*-positive farms did not coincide with that for *Coxiella*-positive farms, suggesting that different risk factors might be at play for clusters. Since both *Neospora* and *Coxiella* occur in the same hosts, a possible discriminating factor might be the presence of reservoir hosts such as wild animals and ticks. The distribution range of suitable reservoir hosts may vary between high-risk areas associated with both significant clusters as a result of the heterogeneity of climates and landscapes. For example it is known that the habitat of ticks is more set by abiotic factors such as vegetation and climate (which determine tick development and survival rates on the ground) rather than by host-related factors [42].

Another interesting observation in the present study is that the identified clusters encompass areas with high and low probability of presence for both *Neospora*-positive and *Coxiella*-positive farms predicted by Maxent program. In addition, even though the bioclimatic variables were found in the present study to affect similarly the geographical distribution of both *Neospora*-positive and *Coxiella*-positive farms, no association between the prevalence of these pathogens was found. These observations suggest that in addition to environmental factors, other factors such as animal and farmer as well as farm and pasture management status affect infection with these two pathogens.

The environmental factors identified in the present study are expected to be valid to model the geographical distribution of *Neospora*-positive and *Coxiella*-positive farms in other regions of the world, but other enzootic areas in the world should be located to test this assumption. In addition, it seems that the distribution of *Neospora*-positive and *Coxiella*-positive farms is also influenced by other variables not analysed in our study.

In conclusion, the highest level of probability for the geographical distribution of *Neospora*-positive and *Coxiella*-positive farms covers central Catalonia and spreads towards northeastern Catalonia which contains a high concentration of farms with potential dairy cattle hosts. Maximum entropy ecological niche modelling has proved to be a useful tool in mapping *Neospora* and *Coxiella* in northeastern Spain, while the significant clusters detected might be helpful in investigating the underlying causes of increased risk in the identified areas.

ACKNOWLEDGEMENTS

The authors thank the students of the School of Agricultural Engineering of the University of Lleida Rosa Colome and Montse Navarro and the staff members of the Interprofessional Dairy Laboratory of Catalonia (ALLIC) for their help with the collection of samples for this study. This work was supported by Spanish CICYT grant GL2010-21273-C03-01/GAN.

DECLARATION OF INTEREST

None.

REFERENCES

milk and individual samples by PCR and ELISA. 
Veterinary Record 2012; 168: 79


42. **Randolph SE.** Tick and tick-borne disease systems in space and from space. *Advances in Parasitology* 2000; **47**: 217–243.