

Use of high spatial resolution satellite imagery to characterize landscapes at risk for bluetongue

Hélène GUI^{a,b,c}, Annelise TRAN^{b,c,*}, Stéphane de LA ROCQUE^b, Thierry BALDET^b, Guillaume GERBIER^b, Bruno BARRAGUÉ^b, Fabienne BITEAU-COROLLER^b, François ROGER^b, Jean-François VIEL^a, Frédéric MAUNY^a

^a Department of Public Health, Biostatistics and Epidemiology Unit, Faculty of Medicine, St. Jacques University Medical Center, 2 place St. Jacques, 25030 Besançon Cedex, France

^b French Agricultural Research Center for International Development (CIRAD), Epidemiology and Ecology of Animal Diseases Unit, Baillarguet Campus, TA 30 B, 34398 Montpellier Cedex 5, France

^c Territories, Environment, Remote Sensing and Spatial Information Joint Research Unit (UMR TETIS), Maison de la Télédétection, 500 rue J.-F. Breton, 34093 Montpellier Cedex 5, France

(Received 18 October 2006; accepted 27 March 2007)

Abstract – The recent and rapid spread in the Mediterranean Basin of bluetongue, a viral disease of ruminants transmitted by some species of *Culicoides* (biting midges), highlights the necessity of determining the conditions of its emergence. This study uses high spatial resolution satellite imagery and methods from landscape ecology science to identify environmental parameters related to bluetongue occurrence in Corsica, a French Mediterranean island where the disease occurred for the first time in 2000. A set of environmental variables recorded in the neighborhood of 80 sheep farms were related to case occurrence through a logistic regression model computed within three subsequent buffer distances of 0.5, 1 and 2 km. The results reveal the role of landscape metrics, particularly those characterizing land-use units such as prairies and woodlands, as well as farm type, latitude and sunshine to explain the presence of bluetongue. Internal and external validation both indicate that the best results are obtained with the 1 km buffer size model (area under Receiver Operating Characteristic curve = 0.9 for internal validation and 0.81 for external validation). The results show that high spatial resolution remote sensing (i.e. 10 m pixels) and landscape ecology approaches contribute to improving the understanding of bluetongue epidemiology.

epidemiology / bluetongue / remote sensing / geographic information systems / landscape ecology

1. INTRODUCTION

Bluetongue (BT) is a vector-borne disease which affects ruminants, with high severity for sheep. BT is recorded in the list of diseases with mandatory notification to the World Organisation for Ani-

mal Health (Office International des Epizooties) because of its major consequences on animal health, economies and international trade¹. Bluetongue virus (BTV) is transmitted by some species of biting

¹ OIE, Diseases notifiable to the OIE [on line] (2004) http://www.oie.int/eng/maladies/en_classification.htm [consulted 16 October 2006].

* Corresponding author: annelise.tran@cirad.fr

midges of the *Culicoides* genus (Diptera: Ceratopogonidae). Until recently Europe was BT-free, although incursions did occur at intervals. Since 1998, in Europe, the circulation of six different BTV serotypes have caused the most severe outbreaks of BT on record [25]. In Western Europe, the first cases of BT were detected in August 2000 in the Italian island of Sardinia [12]. The virus was then rapidly detected in the nearby island of Sicily, the mainland of Italy, the French island of Corsica, and the Spanish islands of Menorca and Mallorca (October 2000) [25]. In the western part of the Mediterranean basin, climate changes could have contributed to this spread through the northward expansion of the main Afro-Asian BT vector, *Culicoides imicola* [39]. In Corsica, *C. imicola* was first detected in October 2000 [20] a few weeks before BT outbreaks occurred. An entomological surveillance network implemented from 2002 onwards has confirmed that this vector is now well established throughout the island [2]. BTV circulation has been detected every year since its introduction in 2000. More knowledge is needed to understand the conditions of BT emergence in Corsica, the first step being to better define the risk zones.

To identify the areas at risk for bluetongue, several models predicting BT vector distribution have been developed using climatic data [14, 50] and low spatial resolution remote sensing data [3–7, 38, 46]. Low spatial resolution satellite sensors can provide surrogates of meteorological data through specific channel values or the calculation of indices related to temperature, humidity and vegetation. These methods have been developed for different vector-borne diseases [8, 41, 42] and take benefit from the area-wide availability of the data and their repetitiveness over time. The accuracy of these predictions on the presence/absence or abundance of *C. imicola* is variable (from $R^2 = 0.53$ [3, 4] to $\kappa = 0.91$ [46]): understandably,

they are more accurate in the regions from which entomological training data sets come from. Predictions from the only model taking into account Corsican entomological data are in good agreement with the observed vector distribution at the low spatial resolution of the study². Nevertheless, to reveal local contrasts in distribution patterns of BT, additional fine scale analyses are necessary [37].

High spatial resolution approaches can both take benefit from knowledge on vector and disease ecology and landscape biometrics [19], and help acquire additional knowledge on vector ecology when only limited field information is available, as is still the case for *C. imicola*. Using such approaches, the environmental landscapes favorable or unfavorable to vectors and to the diseases they carry can be identified. To characterize these landscapes, a wide variety of metrics referring to complementary aspects of vector or host ecology are available. These metrics quantify the composition of the mosaic (patch type, abundance, diversity, etc.) or its structure (shape, edge, neighbors, etc.). Applications of landscape ecology include different diseases such as trypanosomiasis [10, 18] and Lyme disease [22, 27, 28, 40]. Such applications to vectors and vector-borne diseases have been reviewed by Kitron [29].

The aim of this study was to characterize favorable landscapes for BT occurrence in southern Corsica. We test the hypothesis that BT outbreaks in southern Corsica are not distributed randomly but can be explained by landscape elements detected at fine spatial remote sensing resolutions and by farm characteristics. To identify potential landscape risk factors, the environment surrounding BT-free and

² Roger F.L., Emergence of bluetongue disease in the mediterranean basin: Modelling locations at risk for potential vectors (*Culicoides* spp.) using Satellite Imagery, MSc, University of London, 2002.

BT-infected sheep farms is described using a high spatial resolution SPOT (Satellite pour l'Observation de la Terre) image and a digital elevation model (DEM). The satellite image of southern Corsica was classified to obtain a land cover map that was used in a landscape analysis to characterize the spatial structure of the mosaic of vegetation patches. Statistical models of BT occurrence were created and their accuracy was assessed.

2. MATERIALS AND METHODS

2.1. Localization of farms and disease outbreaks

The study was restricted to the farms under the charge of one of the governmental veterinary assistants of southern Corsica. With his help, the farms under his responsibility were georeferenced in a Geographical Information System (GIS) using ESRI ArcGis™ (Redlands, CA, USA), their official identification number recorded and species bred listed. Amongst these livestock farms, only those including sheep holdings were selected ($n = 80$, referred to as *sheep farms*). Two types of farming systems were recorded in the sheep farms: one with only sheep holdings (referred to as *monospecific sheep farms*), and the second with sheep and other livestock (cattle and/or goats, referred to as *mixed farms*). The official declaration of BT occurrence by French veterinary services was based on detection of BTV by RT-PCR in clinically suspected sheep by the national reference laboratory. BT infected farms were defined as farms in which at least one BT outbreak (whatever the serotype) had been officially declared between 2000 and 2003 ($n = 46$). All other sheep farms for which no BT outbreaks had been officially declared were considered BT-free ($n = 34$).

2.2. Remote sensing derived environmental data

To characterize the environment in the neighborhood of the 80 sheep farms, buffer zones were defined around each farm. The active flight range of *C. imicola* was not precisely described, but referring to other *Culicoides* species [30], it should range between a few hundred meters and a few kilometers. Therefore, three sizes of buffer zones were tested: 0.5, 1 and 2 km in radius. Since knowledge on the bio-ecology of *C. imicola* is limited, a wide range of environmental data, derived from a DEM and from a high spatial resolution SPOT satellite image, were tested.

2.2.1. Environmental data derived from the DEM

The DEM of Corsica was provided by the IGN (Institut Géographique National); its pixel size is 50 m. Five variables were extracted from the DEM, namely altitude (mean in the buffer), slope (mean in the buffer), a sunshine index (mean and standard deviation in the buffer), and aspect (modal class). The sunshine index was calculated according to the sun's position on the 30th of June 2001 at 16 h GMT (ENVI® software, ITT, CO, USA). This index varied between 0 and 1 and can be considered as an indicator of temperature and humidity. The aspect was measured clockwise in degrees with 0 = 360° facing north and 180° facing south. It was classified in three equal classes: one with a southern exposure facing Sardinia (130–250°), and two others facing north-east (10–130°) and north-west (250–10°), to test if areas facing Sardinia (the closest infected area) could be more at risk than the others. Finally, hydrographical data were used to quantify the length of the rivers flowing through the buffer zones.

2.2.2. Environmental data derived from the SPOT image

Data on vegetation coverage were extracted from a SPOT 4 satellite image (pixel of 10 m) acquired on the 6th of June 2001. SPOT 4 images include 4 channels in the green, red, near infrared (NIR) and middle infrared (MIR) portions of the spectrum.

2.2.2.1. NDVI

The NDVI, a standardized vegetation index [49], was calculated for this image. Minimum, maximum, mean, and standard deviation values in the buffer were extracted. The image was also used to produce a land-cover classification.

2.2.2.2. Land-cover map

A land-cover map was obtained by carrying out a supervised object-oriented classification (eCognitionTM software, Definiens-imaging, Munich, Germany). In object-oriented approaches, the image is first segmented into homogeneous polygons, so-called objects. A supervised classification then allocates the objects to pre-defined land-cover classes in three key stages [16]. The first stage was the acquisition and characterization of training sites; more than 180 sites of known land cover, identified in the field with a GPS (Global Positioning System) receiver, are geolocated on the satellite image and the features of the respective objects were used as a training set. At the second stage, a standard nearest neighbor classification algorithm was used to allocate each object, according to its features, to the most similar land-cover class. In this way, a thematic classification was produced. A major advantage of classifying objects rather than pixels is the fact that features such as topology (i.e. relations to the neighboring objects), shape, texture, and a wider range of statistics on spectral signatures can be used to characterize the

objects [9]. In all, 49 features were used to classify the image. During the last stage, the accuracy of the obtained classification was evaluated using a new set of training data, i.e. ground-truth data, coming from more than 210 sites. In a confusion matrix, the allocated land-cover class of the new set of objects is compared to the observed land-cover class.

2.2.2.3. Landscape metrics

A landscape analysis was conducted on the land-cover map to characterize its spatial structure. In this study, being the first in the field to link landscape metrics and *Culicoides*, a wide range of landscape metrics were chosen, including area-density-edge, diversity, isolation-proximity and connectivity. Using the Fragstats freeware³, metrics were calculated at two levels: one for each land-cover class found in a buffer (class-level metrics) and the other for the whole landscape, i.e. for the entire buffer regardless of the class (buffer-level metrics). The description of the selected metrics is provided in Table I.

2.3. Statistical analysis

2.3.1. Univariate analysis

As a first step, a descriptive analysis (including a Pearson correlation matrix) was performed for all the variables. The quantitative variables were tested both as quantitative and, after coding in classes, as qualitative variables. They were kept as quantitative variables whenever possible, and were coded in terciles (or in presence/absence if more relevant) when necessary to accommodate for non-monotonic

³McGarigal K., Cushman S.A., Neel M.C., Ene E., FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps, version 3.3, University of Massachusetts, Amherst, USA, [on line] 2002 www.umass.edu/landeco/research/fragstats/fragstats.html [consulted 16 October 2006].

Table I. Landscape metrics extracted from Fragstats.

Descriptor	Name (<i>Metric level; Pattern component</i>)	Definition
PD	Patch Density (<i>Buffer-level; Area-density-edge</i>)	Number of patches in the buffer (of all patch types), divided by the buffer area, and then converted to 100 hectares.
LSI	Landscape Shape Index (<i>Buffer-level; Area-density-edge</i>)	Total length of edge divided by the minimum length of class edge possible for a maximally aggregated class (achieved when the class is maximally clumped into a single, compact patch).
ENN_MN	Mean Euclidian Nearest Neighbour Distance (<i>Buffer-level; Isolation-proximity</i>)	Mean distance (m) to the nearest neighboring patch (regardless of the patch type).
CONNECT	Connectance index (<i>Buffer-level; Connectivity</i>)	Number of joinings between all patches of the same type within a 50 meter range (user-defined threshold), divided by the total number of possible joinings between all patches of the same type (whatever the range), and then converted to a percentage.
PRD	Patch Richness Density (<i>Buffer-level; Diversity</i>)	Number of different patch types present within the buffer divided by buffer area, and then converted to 100 hectares.
SIDI	Simpson's Diversity Index (<i>Buffer-level; Diversity</i>)	SIDI equals 1 minus the sum, across all patch types, of the proportional abundance of each patch type squared. It represents the probability that any 2 pixels selected at random would be different patch types.
PLAND_X*	Percentage of Landscape covered by patches of class X (<i>Class-level; Area-density-edge</i>)	PLAND equals the sum of the areas (m ²) of all patches of the corresponding patch type, divided by total landscape area (m ²) and then converted to a percentage.
NP_X*	Number of patches of class X (<i>Class-level; Area-density-edge</i>)	Number of patches of each of the 9 types of land-cover classes.
LSI_X*	Landscape shape index of class X (<i>Class-level; Area-density-edge</i>)	LSI calculated for each of the 9 types of land-cover classes.

* X represents each of the 9 types of land-cover classes.

and nonlinear responses (e.g. little risk associated with both small and high values of the variable, and high risk associated with a mean value of the variable). Mann-Whitney, Pearson chi-square or Fisher exact tests [1] were used to test associations between BT status and environmental data. Significant variables in preliminary univariate screening analysis at a 0.15 *p*-value for at least one size of buffer were then

tested in a multivariate analysis with a logistic regression stepwise procedure.

2.3.2. Multivariate analysis

One of the main purposes of multivariate logistic regression is to examine the role of explanatory variables, having adjusted for other variables [17]. Much

caution was taken to examine the effects of correlations on variable selection at all steps of the process. The variables were first gathered in five thematic groups (topography, land-use units, buffer-level landscape metrics, class-level number of patches and class-level landscape shape indexes). Variables of each theme that were significant in the multivariate analysis with a 0.10 p -value were next tested with the significant variables of the other themes. The entire process was repeated for each buffer size, thus producing one logistic model for each buffer size (Systat[®] software, SSI, CA, USA).

2.3.3. Autocorrelation

In spatial data sets, observations that are close in space tend to be more similar than observations that are far apart, i.e. data are spatially correlated. Underlying logistic regression techniques are the two basic assumptions that the residuals of the model are independently and identically distributed. If the data are spatially correlated and the predictor variables do not take this autocorrelation into account, the residuals of the model will be spatially autocorrelated, and the statistical link between the predictor variables and the predicted variable may be over-estimated [48]. A Monte Carlo test based on the variogram was used to check for evidence of spatial correlation of residuals⁴. This test consists in comparing the observed variogram with variogram “envelopes” computed by simulating permutations of the data values

⁴Ribeiro P.J. Jr., Christensen O.F., Diggle P.J., geoR and geoRglm: Software for Model-Based Geostatistics. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20–22, Vienna, Austria, [online] 2003 <http://www.ci.tuwien.ac.at> in Conferences - DSC-2003 – Proceedings [consulted 9 January 2007].

across locations. This test was carried out with the R freeware⁵ (geoR and geoRglm packages).

2.4. Internal accuracy assessment

The models were first validated internally, i.e. on the same data set of 80 farms of southern Corsica. The predictive accuracy of the models was assessed using the ROC (Receiver Operating Characteristic) curve method [32, 36]. The area under curve (AUC) of the ROC curve, its 95% confidence interval and the sensitivity and specificity (cut-off point 0.5) of the models were calculated to evaluate the quality of the models. The greater the AUC, the more discriminating the model is and the closer the predictions are to the observed data.

2.5. External accuracy assessment

The models were validated externally, i.e. on new environmental and epidemiological data sets. A SPOT image of the Ajaccio region (Corsica, France), situated north of the first study area, was classified using the same methods and land-cover classes. The variables selected in the models were extracted and the risk probability was calculated and compared to the official disease status of sheep farms located in the Ajaccio region. When more than 25% of the buffer zone surrounding a farm fell outside the satellite image, the farm was excluded, thus the number of farms included varied from 130 farms with the 2 km buffer to 151 farms with the 0.5 km buffer. Similar to the internal validation, external validation relied on the AUC of the ROC curve.

⁵R Development Core Team, R: a language and environment for statistical computing, version 2.3.1, R Foundation for statistical computing, Vienna, Austria, [on line] 2005 <http://www.r-project.org/> [consulted 16 October 2006].

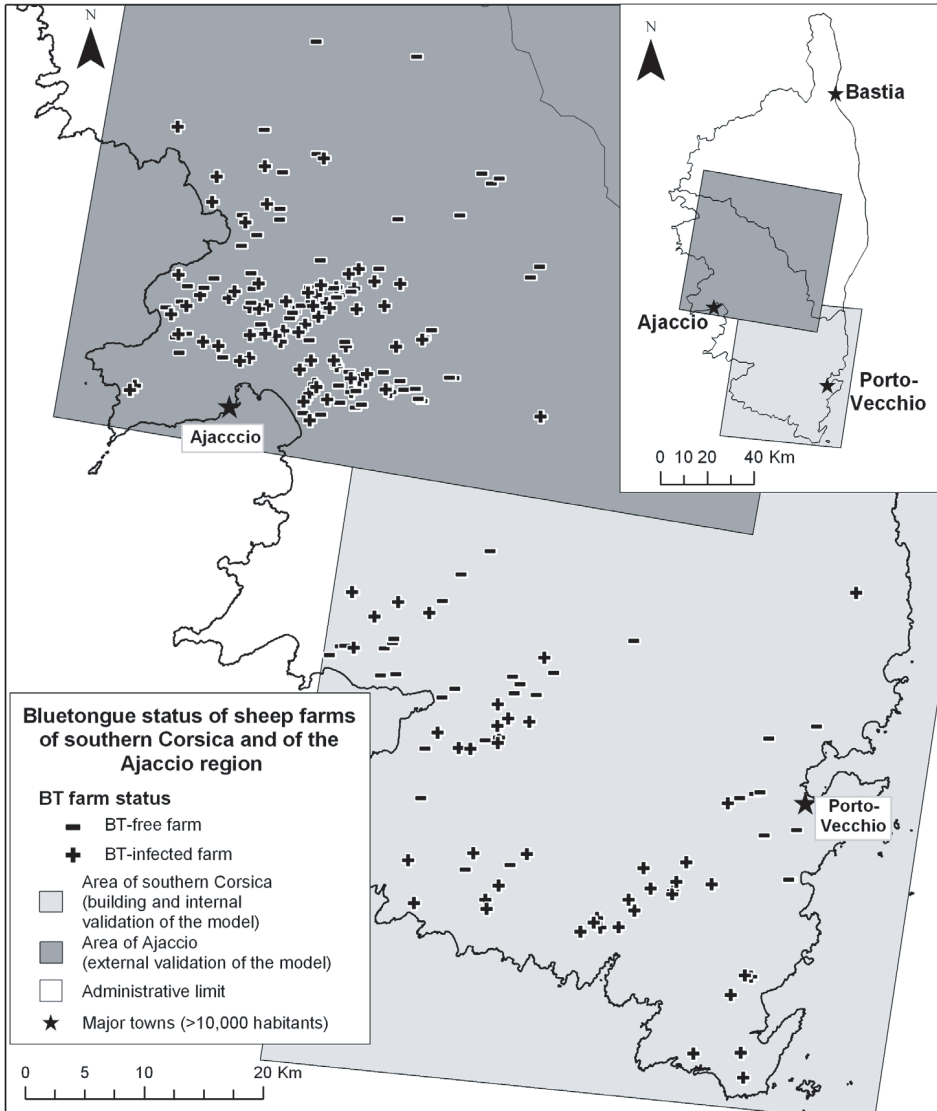


Figure 1. Bluetongue status of sheep farms of southern Corsica and of the Ajaccio region.

3. RESULTS

3.1. Farm characteristics

The locations of the BT-infected ($n = 46$) and BT-free ($n = 34$) farms are presented in Figure 1. The absence of

farms in the central part of the island is due to a high mountain range (1400 m above sea level), with all farms being located under 800 m. Seven farms were monospecific sheep farms and 73 were mixed farms.

Table II. Variables selected in the BT models for the three sizes of buffers.

Variables* classes	Models					
	0.5 km buffer		1 km buffer		2 km buffer	
	OR	(90% CI)	OR	(90% CI)	OR	(90% CI)
Farm type						
Monospecific sheep farm	1		1		1	
Mixed farm	57.83	(3.97–842.76)	133.29	(10.74–1653.83)	37.16	(4.93–280.08)
Latitude (km)	0.88	(0.83–0.94)	–		0.86	(0.81–0.92)
Mean sunshine	–		1E-4	(3E-8–0.56)	2E-6	(4E-11–0.08)
Patch richness density						
Low values (T1)	1		–		–	
Medium values (T2)	4.86	(1.54–15.33)	–		–	
High values (T3)	0.81	(0.21–3.09)	–		–	
Low shrublands						
Absence	–		1		–	
Presence	–		5.2	(1.43–18.92)	–	
LSI of impervious surface	–		1.53	(1.05–2.22)	–	
LSI of open prairies						
Low values (T1)	–		1		–	
Medium values (T2)	–		0.13	(0.03–0.66)	–	
High values (T3)	–		1.31	(0.28–6.20)	–	
LSI of woodlands						
Low values (T1)	–		1		–	
Medium values (T2)	–		12.18	(2.27–65.35)	–	
High values (T3)	–		10.07	(1.98–51.07)	–	
Nb of patches of open prairies	–		–		1.24	(1.10–1.40)

Intercept: 513.06 for the 0.5 km buffer, –3.37 for the 1 km buffer and 612.86 for the 2 km buffer.

* Variables selected with $p < 0.1$. Variables were kept as continuous variables unless it was more relevant to recode them in terciles (T) or to dichotomize (presence/absence) them.

OR: odds ratio; CI: confidence interval.

3.2. Land-cover maps

Nine land-cover classes were defined, namely low shrublands, woodlands (including forests and *maquis*, a typical Mediterranean association of dense shrubs and trees), pine trees, open prairies, prairies with tree cover, cultivated land, marshes, impervious surface and water. The confusion matrixes revealed that the overall accuracies of the classifications were fair (71%) for southern Corsica and very good (90%) for the Ajaccio region.

3.3. Selected variables in the models

The variables included in each of the three buffer-specific models as well as their contribution are detailed in Table II. All three models include landscape metrics, three environmental variables are common to at least two models and six others are included in only one model.

– The three models include the type of farm as a predictive variable. The high odds ratio values and large confidence intervals reflect the sampling imbalance of

Table III. Internal and external evaluation of the BT models for the three sizes of buffers.

Model evaluation	0.5 km buffer	1 km buffer	2 km buffer
Internal Validation ($n = 80$)			
AUC ROC curve (95% CI)	0.85 (0.77–0.93)	0.90 (0.83–0.97)	0.88 (0.81–0.96)
Sensitivity (cut-off: 0.5)	87%	85%	87%
Specificity (cut-off: 0.5)	62%	85%	68%
External Validation			
Number of farms	151	134	130
AUC ROC curve (95% CI)	0.73 (0.65–0.81)	0.81 (0.74–0.88)	0.77 (0.69–0.85)

CI: confidence interval; n : number of farms.

BT infection in monospecific sheep farms (1 out of 7, 14%) and in mixed farms (45/73, 62%).

– Latitude and mean sunshine are selected in two models, the southernmost farms or those with the least sunshine being at higher risk.

– The 1 km buffer size model includes the landscape shape index (LSI) of 3 land-cover classes (impervious surfaces, open prairies and woodlands). The LSI is related to the length of the edges.

– Finally, other metrics such as the patch richness density, the presence of low shrublands and the number of patches of open prairies are selected in the 0.5, 1 and 2 km buffer size models respectively.

– High LSI (important length of edges) and high numbers of patches of open prairies are both associated with higher risk in the 1 and 2 km buffer size models, respectively.

– Neither isolation-proximity nor connectivity metrics was selected by the models.

3.4. Validation results

The results of the internal and external validation are presented in Table III. The three models have a good ($0.80 < \text{AUC} < 0.9$) or a high ($\text{AUC} > 0.9$)

accuracy (1 km buffer size model). The 1 km buffer size model has the best compromise between sensitivity and specificity (85% in both cases), while the other two models have slightly better values of sensitivity (87%) but far worse values of specificity (62 and 68%).

The 1 km buffer size model was also the best model when tested with a new data set (external validation) since it has a good accuracy (AUC of the ROC curve = 0.81) whereas the 0.5 and 2 km buffer size models have a fair accuracy ($0.7 < \text{AUC} < 0.8$).

Figure 2 shows the distribution of BT-infected and BT-free farms as predicted by the 1 km buffer size model for southern Corsica. True positives ($n = 39$) and true negatives ($n = 29$) are farms whose status was correctly predicted (BT-infected and BT-free respectively). False positives ($n = 5$) are farms which are BT-free but were predicted BT-infected and false negatives ($n = 7$) are farms which are BT-infected but were predicted BT-free.

The residuals of the 1 km buffer size model were not spatially autocorrelated, reflecting the fact that the spatial interactions (if any) had already been taken into account in the model. The map (Fig. 2) shows that the farms that were wrongly predicted as infected (false positives) or disease-free (false negatives) do not cluster in particular areas of southern Corsica.

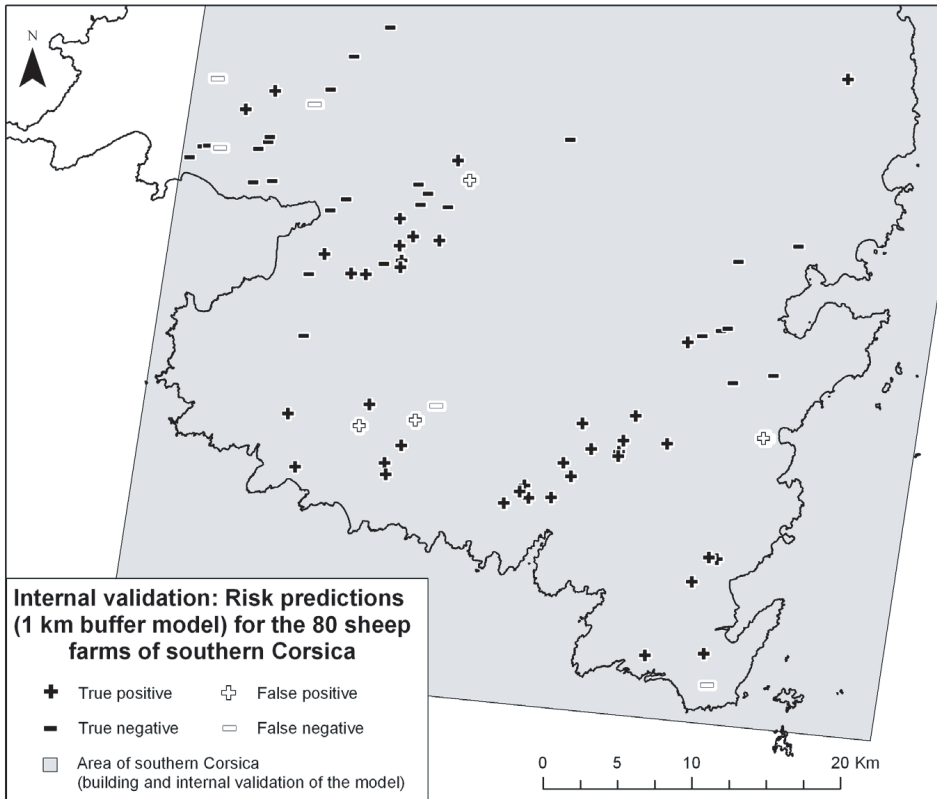


Figure 2. Internal validation: Risk predictions (1 km buffer model) for the 80 sheep farms of southern Corsica.

4. DISCUSSION

4.1. Landscape metrics

Our results suggest that landscape components may be important in defining favorable environments for BT since there are significant statistical associations between landscape metrics and BT risk, regardless of the buffer scale considered. With a few exceptions [22, 26], in most landscape ecology studies one type of metrics, namely the percentage of land surface, is explored since it is the simplest metric and can easily be extracted from GIS software (such as the Spatial Analyst extension of ArcView, for example).

Although Fragstats is a free and simple software, more complex landscape metrics (such as LSI, PRD, connectance, etc.) are still rarely used even if they have proven to be very useful in characterizing subtle ecological aspects of transmission, as in the case of Lyme disease [22]. This study shows that landscape approaches can also be useful when little is known about the vector's ecology.

4.2. Buffer scales

Three sizes of buffer zones were tested since active flight range of *C. imicola* is not precisely known. The results show that

different variables are selected according to the buffer scale chosen, although these variations may be in part amplified by the small number of farms included. Since environmental factors seem to have an influence at the 3 buffer scales chosen, testing a wider range of buffer scales would be necessary to determine within what range of buffer sizes there is an influence of the neighbouring environment on BT risk. Since the 1 km buffer size model appeared to be the most discriminating, we chose to focus the discussion on the significant ($p < 0.1$) variables of this model.

4.3. Reservoir species

The results show that whatever the buffer scale considered, mixed farms seem to be more at risk than monospecific sheep farms (where only sheep are bred). Some authors have suggested that cattle may be more attractive to *Culicoides* than sheep and thus may serve to protect sheep from infection by drawing the insects to themselves [21, 31]. Our results seem to indicate the opposite, although, as previously stated, because of the imbalance between mixed and monospecific sheep farm groups, the results should be interpreted with much caution. If these results are confirmed, one hypothesis may be that cattle and goats, known reservoirs of BT virus [21, 31], may have a more important role as reservoirs rather than just as feeding hosts in the epidemiology of the disease.

4.4. Low shrublands and latitude: a source-point effect

Latitude was a key variable selected in the 0.5 and 2 km buffer size models. While not selected in the 1 km buffer size model, latitude may be indirectly present since the correlated low shrublands unit (correlation coefficient: -0.6) that is preferentially distributed in the southernmost part

of Corsica, is selected in the model. This importance of latitude probably reflects a distance to a disease source-point, i.e. the distance to Sardinia. Since introduction of ruminants from BT-infected areas like Sardinia was prohibited two months before first Corsican outbreaks, the risk of incoming viremic animals can be considered minor [24]. This finding supports the hypothesis that there are regular invasions of infected midges from Sardinia to Corsica. Indeed, the French and Italian BT-surveillance networks have shown that the first outbreaks due to BTV2 in 2000 and BTV4 in 2003 in Corsica were preceded by outbreaks of the same serotype in Sardinia⁶ [23]. Since *Culicoides* can be passively transported by winds over hundreds of kilometers [11, 44, 45], the 12 km of sea between Corsica and Sardinia could have been easily crossed. Furthermore, weather conditions and wind patterns were shown to be favorable for transportation of BTV4 infected midges in 2003 [24].

4.5. Sunshine and lack of humidity

Mean sunshine was negatively associated with BT risk. This may seem rather surprising since *C. imicola* is a tropical midge whose northern limit in cooler European regions is linked to temperature [39]. Since sunshine was estimated for a summer afternoon, the minimal temperatures required for *C. imicola* are surpassed (the mean monthly temperature in June in southern Corsica is 19 °C). One hypothesis could be that the limiting factor is the lack of humidity rather than temperature.

⁶ Gerbier G., Baldet T., Cêtre-Sossah C., et al., Emergence of Bluetongue in France 2000-2004, in: SciQuest (Ed.), Proc. 11th ISVEE, Cairns, Australia, [on line] (2006) <http://www.sciquest.org.nz/default.asp?pageid=69&pub=10&vol=11> [consulted 16 October 2006].

4.6. The importance of edge metrics

High values of lengths of edges (LSI) of woodlands, open prairies and impervious surfaces are associated with high BT risk. This shows that it may be more important to consider the fragmentation of the environment rather than the percentage of surface covered by a land-cover class. It is interesting to note that the percentage of surface and the LSI are not always positively correlated and may even be negatively correlated, as in the case of woodlands. Edges and borders may be important for disease transmission by several means such as by providing meeting points between hosts and vectors, or by providing landscapes where essential elements of a habitat are closely intricate.

The results suggest that fragmented landscapes containing woodlands and prairies are linked with BT risk. More thorough knowledge on the life history traits of *C. imicola* is necessary to determine whether edges of woodlands are important because they offer resting sites for *C. imicola*, or whether landscapes containing large compact surfaces of woodlands with small lengths of edges characterize farms which are more isolated, with big patches of woodlands acting like a barrier for *C. imicola*. These results show that until then, the links with environmental characteristics must be explored with caution, using simple metrics, assessing their robustness in different study areas and must be completed with complementary entomological field studies.

4.7. Combining high and low spatial resolution approaches

This study shows that the scope of satellite imagery to model vector-borne diseases is very wide; at low spatial and high temporal resolutions, climatic surrogates can be extracted [3–7, 38, 46], and at finer spatial resolutions, landscape approaches

can help untangle the interplay between vector-borne diseases and their environment and support entomological and epidemiological field studies. While the use of local scale approaches is still limited, it will hopefully be promoted in the future, taking benefit of the increasing development of sensors. The aims of both approaches are then different but highly complementary; a key issue now is to establish a link between the two and combine them to describe both climatic and landscape conditions of occurrence of diseases.

4.8. Conclusion

An innovative ecological high spatial resolution approach was carried out to model the occurrence of BT in southern Corsica. Since little is known on the ecology of *C. imicola*, many precautions were taken; several sizes of neighborhoods were tested, the best model being chosen according to internal and external validation criteria. The results highlight the importance of landscape features (especially edge metrics), farm types, and reveal the ‘source point effect’ of Sardinia. The method presented here is relatively flexible, and new environmental characteristics (such as pedological information [15]) can easily be taken into account if necessary. Before these results can be used to help identify future hotspots outside Corsica which could aid the management of the disease, it is necessary to locally adapt the model to take into consideration prevailing ecological and environmental conditions, especially if different *Culicoides* species are involved, as in southern [13, 43], eastern [35] and recent outbreaks of northern Europe [47]. Such an approach will be adapted to the Var, a French region where *C. imicola* is present but not the virus, to help focus the trapping sites of the entomological surveillance system in the areas where BT risk is the greatest.

As a first approach, the models presented here were based on easily accessible epidemiological data, i.e. official BT outbreaks. Since disease-based models suffer from several limitations (a major one being a looser link with environmental conditions), a similar analysis (same spatial resolution, buffer scales, variables and study area) will be developed using entomological data coming from an extensive *Culicoides* trapping campaign. This will help to determine whether models based on more costly and labor intensive *Culicoides* collections [33, 34] outperform disease-based models.

ACKNOWLEDGEMENTS

The authors would like to thank the French Veterinary Services of southern Corsica (DDSV 2A), especially G. Bousquet and J. Parodi, for their collaboration during data collection. They would also like to thank D. Pleydell for his help with R software and V. Soti for her help with the field work. The authors thank the reviewers for their very constructive comments on the manuscript. The SPOT image was obtained with financial support of the ISIS (Incitation à l'utilisation scientifique d'images SPOT) program of the CNES (Centre National d'Etudes Spatiales). This work was funded by a grant ACI "écologie quantitative" from the French Ministry of Research. H. Guis received a doctoral fellowship from the University of Franche-Comté and the French Ministry of Research.

REFERENCES

- [1] Armitage P., Colton T., Encyclopedia of biostatistics, John Wiley and Sons, New-York, USA, 1998.
- [2] Baldet T., Delécolle J.-C., Mathieu B., De La Rocque S., Roger F., Entomological surveillance of bluetongue in France in 2002, *Vet. Ital.* (2004) 40(3):226–231.
- [3] Baylis M., Bouayoune H., Touti J., El Hasnaoui H., Use of climatic data and satellite imagery to model the abundance of *Culicoides imicola*, the vector of African horse sickness virus, in Morocco, *Med. Vet. Entomol.* (1998) 12:255–266.
- [4] Baylis M., Rawlings P., Modelling the distribution and abundance of *Culicoides imicola* in Morocco and Iberia using climatic data and satellite imagery, *Arch. Virol. Suppl.* (1998) 14:137–153.
- [5] Baylis M., Meiswinkel R., Venter G.J., A preliminary attempt to use climate data and satellite imagery to model the abundance and distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in southern Africa, *J. S. Afr. Vet. Assoc.* (1999) 70:80–89.
- [6] Baylis M., Mellor P.S., Wittmann E.J., Rogers D.J., Prediction of areas around the Mediterranean at risk of bluetongue by modelling the distributions of its vector using satellite imaging, *Vet. Rec.* (2001) 149:639–643 (erratum *Vet. Rec.* (2002) 150:404).
- [7] Baylis M., O'Connell L., Purse B.V., Modelling the distribution of bluetongue vectors, *Vet. Ital.* (2004) 40(3):176–181.
- [8] Beck L.R., Lobitz B.M., Wood B.L., Remote sensing and human health: new sensors and new opportunities, *Emerg. Infect. Dis.* (2000) 6:217–227.
- [9] Benz U.C., Hofmann P., Willhauck G., Lingenfelder I., Heynen M., Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, *ISPRS J. Photogramm.* (2004) 58:239–258.
- [10] Bouyer J., Guerrini L., Desquesnes M., de la Rocque S., Cuisance D., Mapping African Animal Trypanosomosis risk from the sky, *Vet. Res.* (2006) 37:633–645.
- [11] Braverman Y., Chechik F., Air streams and the introduction of animal diseases borne on *Culicoides* (Diptera, Ceratopogonidae) into Israel, *Rev. Sci. Tech.* (1996) 15:1037–1052.
- [12] Calistri P., Giovannini A., Conte A., Nannini D., Santucci U., Patta C., Rolesu S., Caporale V., Bluetongue in Italy: Part I, *Vet. Ital.* (2004) 40(3):243–251.
- [13] Caracappa S., Torina A., Guercio A., Vitale F., Calabro A., Purpari G., Ferrantelli V., Vitale M., Mellor P.S., Identification of a novel bluetongue virus vector species of *Culicoides* in Sicily, *Vet. Rec.* (2003) 153:71–74.
- [14] Conte A., Giovannini A., Savini L., Goffredo M., Calistri P., Meiswinkel R., The effect of

- climate on the presence of *Culicoides imicola* in Italy, *J. Vet. Med. B Infect. Dis. Vet. Public Health* (2003) 50:139–147.
- [15] Conte A., Ippooliti C., Calistri P., Pelini S., Savini L., Salini R., Goffredo M., Meiswinkel R., Towards the identification of potential infectious sites for bluetongue in Italy: a spatial analysis approach based on the distribution of *Culicoides imicola*, *Vet. Ital.* (2004) 40(3):311–315.
- [16] Curran P.J., Atkinson P.M., Foody G.M., Milton E.J., Linking remote sensing, land cover and disease, *Adv. Parasitol.* (2000) 47:37–80.
- [17] Darlington G.A., Collinearity, in: Armitage P., Colton T. (Eds.), *Encyclopedia of biostatistics*, John Wiley and Sons, New-York, USA, 1998, pp. 788–789.
- [18] De La Rocque S., Augusseau X., Guillobez S., Michel V., De Wispelaere G., Bauer B., Cuisance D., The changing distribution of two riverine tsetse flies over 15 years in an increasingly cultivated area of Burkina Faso, *Bull. Entomol. Res.* (2001) 91:157–166.
- [19] De La Rocque S., Michel V., Plazanet D., Pin R., Remote sensing and epidemiology: examples of applications for two vector-borne diseases, *Comp. Immunol. Microbiol. Infect. Dis.* (2004) 27:331–341.
- [20] Delécolle J.-C., de La Rocque S., Contribution à l'étude des *Culicoides* de Corse. Liste des espèces recensées en 2000/2001 et redescription du principal vecteur de la fièvre catarrhale ovine : *C. imicola* Kieffer, 1913 (Diptera, Ceratopogonidae), *Bull. Soc. Entomol. Fr.* (2002) 107:371–379.
- [21] Du Toit R.M., The role played by bovines in the transmission of bluetongue in sheep, *J. S. Afr. Vet. Med. Assoc.* (1962) 33:483–490.
- [22] Estrada-Pena A., Understanding the relationships between landscape connectivity and abundance of *Ixodes ricinus* ticks, *Exp. Appl. Acarol.* (2002) 28:239–248.
- [23] Gerbier G., Parodi J., Biteau-Coroller F., Baldet T., Mathieu B., Zientara S., Cêtre-Sossah C., Roger F., Surveillance de la fièvre catarrhale ovine (bluetongue) en France et dans l'ouest méditerranéen : bilan et perspectives, *Epidemiol. et Santé Anim.* (2006) 37–44.
- [24] Gerbier G., Biteau-Coroller F., Grillet C., Parodi J., Zientara S., Baldet T., Guis H., Roger F., A description of the outbreak of Bluetongue (BTV 4) in Corsica in 2003: lessons for surveillance, *Vet. Rec.* (2007) (in press).
- [25] Gomez-Tejedor C., Brief overview of the bluetongue situation in Mediterranean Europe, 1998–2004, *Vet. Ital.* (2004) 40(3):57–60.
- [26] Graham A.J., Danson F.M., Giraudoux P., Craig P.S., Ecological epidemiology: landscape metrics and human alveolar echinococcosis, *Acta Trop.* (2004) 91:267–278.
- [27] Guerra M., Walker E., Jones C., Paskewitz S., Cortinas M.R., Stancil A., Beck L., Bobo M., Kitron U., Predicting the risk of Lyme disease: habitat suitability for *Ixodes scapularis* in the north central United States, *Emerg. Infect. Dis.* (2002) 8:289–297.
- [28] Jackson L.E., Hilborn E.D., Thomas J.C., Towards landscape design guidelines for reducing Lyme disease risk, *Int. J. Epidemiol.* (2006) 35:315–322.
- [29] Kitron U., Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis, *J. Med. Entomol.* (1998) 35:435–445.
- [30] Lillie T.H., Marquard W.C., Jones R.H., The flight range of *Culicoides variipennis* (Diptera: Ceratopogonidae), *Can. Entomol.* (1981) 133:419–426.
- [31] Luedke A.J., Anakwenze E.I., Bluetongue virus in goats, *Am. J. Vet. Res.* (1972) 33:1739–1745.
- [32] McNeil B.J., Keller E., Adelstein S.J., Primer on certain elements of medical decision making, *N. Engl. J. Med.* (1975) 293:211–215.
- [33] Meiswinkel R., Gomulski L.M., Delécolle J.-C., Goffredo M., Gasperi G., The taxonomy of *Culicoides* vector complexes – unfinished business, *Vet. Ital.* (2004) 40(3):151–159.
- [34] Mellor P.S., Boorman J., Baylis M., *Culicoides* biting midges: their role as arbovirus vectors, *Annu. Rev. Entomol.* (2000) 45:307–340.
- [35] Panagiotatos D.E., Regional overview of bluetongue viruses, vectors, surveillance and unique features in Eastern Europe between 1998 and 2003, *Vet. Ital.* (2004) 40(3):61–72.
- [36] Park S.H., Goo J.M., Jo C.H., Receiver operating characteristic (ROC) curve: practical review for radiologists, *Korean J. Radiol.* (2004) 5:11–18.

- [37] Pili E., Ciucce S., Culurgioni J., Figus V., Pinna G., Marchi A., Distribution and abundance of bluetongue vectors in Sardinia: comparison of field data with prediction maps, *J. Vet. Med. B Infect. Dis. Vet. Public Health* (2006) 53:312–316.
- [38] Purse B.V., Tatem A.J., Caracappa S., Rogers D.J., Mellor P.S., Baylis M., Torina A., Modelling the distributions of *Culicoides* bluetongue virus vectors in Sicily in relation to satellite-derived climate variables, *Med. Vet. Entomol.* (2004) 18:90–101.
- [39] Purse B.V., Mellor P.S., Rogers D.J., Samuel A.R., Mertens P.P., Baylis M., Climate change and the recent emergence of bluetongue in Europe, *Nat. Rev. Microbiol.* (2005) 3:171–181.
- [40] Rizzoli A., Merler S., Furlanello C., Genchi C., Geographical information systems and bootstrap aggregation (bagging) of tree-based classifiers for Lyme disease risk prediction in Trentino, Italian Alps, *J. Med. Entomol.* (2002) 39:485–492.
- [41] Rogers D.J., Randolph S.E., Snow R.W., Hay S.I., Satellite imagery in the study and forecast of malaria, *Nature* (2002) 415:710–715.
- [42] Rogers D.J., Randolph S.E., Studying the global distribution of infectious diseases using GIS and RS, *Nat. Rev. Microbiol.* (2003) 1:231–237.
- [43] Savini G., Goffredo M., Monaco F., Di Gennaro A., de Santis P., Meiswinkel R., Caporale V., The isolation of bluetongue virus from field populations of the *Obsoletus* Complex in central Italy, *Vet. Ital.* (2004) 40(3):286–291.
- [44] Sellers R.F., Pedgley D.E., Tucker M.R., Possible windborne spread of bluetongue to Portugal, June–July 1956, *J. Hyg. (Lond)* (1978) 81:189–196.
- [45] Sellers R.F., Gibbs E.P., Herniman K.A., Pedgley D.E., Tucker M.R., Possible origin of the bluetongue epidemic in Cyprus, August 1977, *J. Hyg. (Lond)* (1979) 83:547–555.
- [46] Tatem A.J., Baylis M., Mellor P.S., Purse B.V., Capela R., Pena I., Rogers D.J., Prediction of bluetongue vector distribution in Europe and north Africa using satellite imagery, *Vet. Microbiol.* (2003) 97:13–29.
- [47] Thiry E., Saegerman C., Guyot H., Kirten P., Losson B., Rollin F., Bodmer M., Czaplicki G., Toussaint J.F., De Clercq K., Dochy J.M., Dufey J., Gillemann J.L., Messemann K., Bluetongue in northern Europe, *Vet. Rec.* (2006) 159:327.
- [48] Thomson M.C., Connor S.J., Environmental information systems for the control of arthropod vectors of disease, *Med. Vet. Entomol.* (2000) 14:227–244.
- [49] Tucker C.J., Red and Photographic Infrared Linear Combinations for Monitoring Vegetation, *Remote Sens. Environ.* (1979) 8:127–150.
- [50] Wittmann E.J., Mellor P.S., Baylis M., Using climate data to map the potential distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in Europe, *Rev. Sci. Tech.* (2001) 20:731–740.