Towards improved, cost-effective surveillance of Ixodes ricinus ticks and associated pathogens using species distribution modelling

Signorini, Manuela; Stensgaard, Anna-Sofie; Drigo, Michele; Simonato, Giulia; Mercer, Federica; Montarsi, Fabrizio; Martini, Marco; Cassini, Rudi

Published in:
Geospatial Health

DOI:
10.4081/gh.2019.745

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NG

Citation for published version (APA):
Towards improved, cost-effective surveillance of *Ixodes ricinus* ticks and associated pathogens using species distribution modelling

Manuela Signorini, Anna-Sofie Stensgaard, Michele Drigo, Giulia Simonato, Federica Marcer, Fabrizio Montarsi, Marco Martini, Rudi Cassini

1Department of Animal Medicine, Production and Health, University of Padua, Legnaro (PD), Italy; 2Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark; 3Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro (PD), Italy

Correspondence: Rudi Cassini, Department of Animal Medicine, Production and Health, University of Padua, viale dell’Università 16, 35020 Legnaro (PD), Italy. Tel.: +39.049.8272777. E-mail: rudi.cassini@unipd.it

Key words: *Ixodes ricinus*; Italy; Species distribution models; Remote sensing; Ticks.

Abstract

Various ticks exist in the temperate hilly and pre-alpine areas of Northern Italy, where *Ixodes ricinus* is the more important. In this area different tick-borne pathogen monitoring projects have recently been implemented; we present here the results of a two-year field survey of ticks and associated pathogens, conducted 2009-2010 in North-eastern Italy. The cost-effectiveness of different sampling strategies, hypothesized *a posteriori* based on two sub-sets of data, were compared and analysed. The same two sub-sets were also used to develop models of habitat suitability, using a maximum entropy algorithm based on remotely sensed data. Comparison of the two strategies (in terms of number of ticks collected, rates of pathogen detection and model accuracy) indicated that monitoring at many temporary sites was more cost-effective than monthly samplings at a few permanent sites. The two model predictions were similar and provided a greater understanding of ecological requirements of *I. ricinus* in the study area. Dense vegetation cover, as measured by the normalized difference vegetation index, was identified as a good predictor of tick presence, whereas high summer temperatures appeared to be a limiting factor. The study suggests that it is possible to obtain realistic results (in terms of pathogens detection and development of habitat suitability maps) with a relatively limited sampling effort and a well-planned monitoring strategy.

Introduction

In Europe, the tick *Ixodes ricinus* (L.) is one of the most important vectors of viruses, bacteria and parasites, responsible for the spread of many zoonotic diseases (e.g., tick-borne encephalitis, Lyme disease or borreliosis, rickettsiosis, human babesiosis and human granulocytic anaplasmosis). In recent years a general increase in the geographical range and incidence of these zoonotic diseases has been observed in Europe (Gray et al., 2013; Bergquist et al., 2018). In Italy, *I. ricinus* is reported throughout the peninsula, particularly in temperate hilly and pre-alpine northern areas, where some ticks and tick-borne pathogens (TBPs) monitoring projects have been implemented recently (Piccoli et al., 2006; Nazzi et al., 2010; Capelli et al., 2012; Ceballos et al., 2014).
Due to the complexity related to the emergence of vector-borne diseases, an integrated approach to surveillance is encouraged (Braks et al., 2011). A prime objective of surveillance of TBPs is to assess the panel of pathogens occurring in a given area, estimating their prevalence both in tick populations and in reservoir hosts. Secondly, information on tick distribution, density and seasonal dynamics should be acquired (Capelli et al., 2012). Ticks are strongly influenced by climatic and environmental variables, such as temperature, relative humidity, rainfall and vegetation cover (Ruiz-Fons et al., 2012; Estrada-Peña et al., 2013), something that contributes to the changing geographical patterns of the TBPs (Dantas-Torres et al., 2012).

A better understanding of the climatic and habitat characteristics that determine the patterns of tick distribution can be obtained through ecological studies and modelling, and can subsequently be used to predict their current distributions over continuous surfaces. Species distribution models (SDMs), widely used in biodiversity research for modelling species geographic distributions based on correlations between known occurrence records and the environmental conditions at those localities (Elith and Leathwick, 2009), have also proven to be a useful tool in disease epidemiology. In fact, SDMs have been widely used in biogeographical and ecological studies of vectors and vector-borne diseases (Illoldi-Rangel et al., 2012; Mughini-Gras et al., 2013; Mwase et al., 2014; Signorini et al., 2014).

In the present study, using data from a 2-years field monitoring activity of ticks and tick-associated pathogens conducted in a limited area of north-eastern Italy (Drigo et al., 2011), we set out to determine the optimal approach to the surveillance of ticks and associated pathogens; we did so by: i) comparing the cost-effectiveness of different sampling strategies for detection of ticks and associated pathogens using two different sub-sets of field data; ii) developing and comparing the outputs of two species distribution models (MaxEnt) based on remotely sensed environmental variables and the two field data-subsets.

Materials and Methods

Study area

The study area comprised the Colli Euganei Regional Park (45°13’-45°24’ N; 11°13’-11°48’ E), a hilly area of volcanic origin, located in the central part of Veneto Region, North-eastern Italy (Figure 1). The Park occupies a small area of approximately 180 km². The area has near-Mediterranean environmental and climatic characteristics and human activities in the park include agriculture, tourism and recreation. Due to these characteristics, there is a potential risk for *I. ricinus* presence and infections associated with this tick. The altitudinal ranges from 0 to 601 m above the mean sea level. The climate is characterized by mild winters (minimum -6°C) and moderately hot summers (maximum 36°C). The annual mean amount of precipitation is 850 mm characterized by two peaks, one in spring (April) and one in autumn (between October and November) and two minima in summer (July) and winter (between January and February). On average, there are 80 days with precipitation per year.

Figure 1. Study area and tick dragging sites.
The territory spans a wide range of vegetation types \(\text{e.g., Mediterranean scrub, arid meadows, woodlands mainly consisting of chestnut, oak and Robinia}\) harbouring a diverse wildlife \(\text{e.g., fox, weasel, badger, small insectivorous mammals, rodents, fallow deer and wild boar}\) that may act as hosts and/or provide blood meals for \(I.\ricinus\) ticks.

**Tick collection and molecular analyses**

The study is based on field data collected during the tick activity seasons in 2009 and 2010. Collections were performed by dragging a 1 m white flannel cloth along a transect of 100 m and examining it at 5-m intervals. Collected specimens were pooled together in vials keeping them alive, identified according to morphological features in the laboratory \(\text{Manilla and Iori, 1992; Cringoli et al., 2005}\) and then stored at -20°C, for further molecular detection of TBPs. Overall, 52 sites were monitored during the two years, with a total of 188 sessions of dragging (Figure 1). During the first year \(\text{2009}\), 52 sites were identified and sampled between April and October to assess the tick distribution in the study area: four out of the 52 sites, were sampled on a monthly basis for 7 months, whereas all other 48 sites were visited only once that year. Then, in 2010, to better examine ticks seasonality, 16 of the original 52 sites were randomly selected and monitored monthly from April to October \(\text{112 drags}\) each site was geo-referenced using GPS \(\text{Juno SB Trimble, USA}\).

Molecular analyses were performed on single adult ticks and on pooled sample of nymphs \(\text{1 to 10}\) and larvae \(\text{1 to 20}\), following the methodology described in detail in Drigo et al. \(\text{2011}\). Briefly, a multiplex real-time polymerase chain reaction \(\text{PCR}\) test was performed with the aim to detect \(Anaplasma\phagocytophilum\) and \(Borrelia\burstdorferi\) s.l., using primers and probes described by Courtney et al. \(\text{2004}\) in the Roche LightCycler. \(Rickettsia\ ssp\). detection was conducted with a conventional PCR targeted at gI\(A\) gene, as described by Regnery et al. \(\text{1991}\).

**Comparison of monitoring strategies**

Two different monitoring strategies were hypothesized a posteriori, based on two sub-sets of data. Strategy 1 included the 16 permanent sites sampled monthly from April to October \(\text{2010}\) \(\text{112 dragsessions}\), while Strategy 2 included only one annual sampling of the 52 sites monitored \(\text{52 dragsessions}\) in \(\text{2009}\). For each strategy, the sampling effort \(\text{number of sites; number of dragging; and the relative output (number of positive sites; number of ticks collected; prevalence of pathogens) were calculated and compared}\).

**Modelling tick distribution and habitat suitability**

To investigate environmental and/or climatic predictors of tick occurrence and to produce maps of tick habitat suitability, one of the most commonly used programmes for modelling species distributions from presence data only, MaxEnt, was chosen \(\text{Phillips et al., 2006}\). This approach is based on a machine-learning algorithm that estimates the distribution across geographic space combining a set of environmental data with presence localities and background records \(\text{pseudo absences}\) sampled from the overall study area \(\text{Phillips and Dudik, 2008}\).

Two different models were fitted, using the sub-set of occurrence data of strategies 1 \(\text{Model 1}\) and 2 \(\text{Model 2}\), to compare the impact of the different sampling strategies on the model output in terms of predicted tick distributions. We used MaxEnt software, version 3.3.3k \(\text{http://www.cs.princeton.edu/schapire/maxent/}\) with its default settings, but only the linear and quadratic features were employed \(\text{Phillips et al., 2004}\). The MaxEnt model outputs were set to logistic, which returns an estimated relative probability of species presence at a given location ranging from \(0\) \(\text{low probability}\) up to \(1\) \(\text{very high relative probability}\). We used a 10-fold cross-validation and assessed predictive performance using the area under the receiver operating characteristic curve, to measure the ability of the model predictions to separate presence from the background \(\text{pseudo- absences}\). For model validation, the leave-one-out or jackknife procedure, proposed by Pearson et al. \(\text{2007}\), was used as this approach has been demonstrated to be effective when only small numbers of occurrence localities are available. In this method, each observed locality is excluded once from the dataset used to build the model, and its predictive performance is assessed based on the ability of the model to predict each single locality excluded in turn. It requires the application of a threshold, for dichotomizing the continuous map output surface into presence and absence values, to assess the model performance by predicting the single excluded presence point from the training dataset. The 10th percentile training presence threshold was used. It sets as threshold the value that excludes the 10\% of the predicted locations having the lowest predicted values, as these may represent recording errors and/or ephemeral populations \(\text{Morueta-Holme et al., 2010; Radosavljevic and Anderson, 2014}\). The predictive ability of models was compared using the test suggested by Pearson et al. \(\text{2007}\). The functional relationship between \(I.\ricinus\) presence and the input environmental variables was evaluated and compared through inspection of the MaxEnt response graphs. The relative importance of each environmental variable in the models was evaluated by jackknife testing of the variable importance procedure, as implemented in MaxEnt. Finally, differences between Model 1 and Model 2 outputs were assessed by quantifying niche overlap, using the Ecological Niche Modelling Tools \(\text{ENMTools}\) \(\text{Warren et al., 2010}\). Niche overlap was calculated using Schoener’s D statistic index. The metric ranges from \(0\) \(\text{species having completely discordant predicted distribution}\) to \(1\) \(\text{species with identical predicted distribution}\) \(\text{Warren et al., 2008}\).

**Environmental variables**

Based on the biological and ecological knowledge on \(I.\ricinus\) requirements, an initial set of environmental factors was chosen as potential explanatory variables of tick distribution. The variables initially included were altitude \(\text{obtained from the United States Geological Survey};\) slope, aspect \(\text{derived from ArcGIS software};\) land cover \(\text{extracted from the Corine Land Cover 2006 raster map};\) and rainfall \(\text{downloaded from the WorldClim database that reports average monthly rainfall data at a \(1\) km\(^2\) resolution, based on interpolated climate data from weather stations}\). Finally, normalized difference vegetation index \(\text{NDVI}\), and day and night time temperature were derived from the Moderate Resolution Imaging Spectroradiometer \(\text{Table 1}\). To reduce issues related to over-fitting and collinearity in the environmental variables, only a subset of potential predictors was chosen from the full set to reduce the number of variables \(\text{Elith et al., 2010}\). An initial explorative analysis was performed to exclude the variables with the lowest variability in the study area. The correlation among the variables was tested using Pearson’s correlation analysis \(\text{software IBM SPSS statistics 20}\). Among those highly correlated \(\text{R}^{2} \geq 0.80\), only the most meaningful ones from a biological point of view were selected for inclusion in the models.
Table 1. Preliminary set of environmental variables selected for developing the models.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
<td>Digital Elevation Model (DEM) GTOPO30, obtained from United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Data Center (<a href="http://tds.ucar.edu/datasets/ds758.0/">http://tds.ucar.edu/datasets/ds758.0/</a>)</td>
</tr>
<tr>
<td>Aspect, Slope, Hill shade</td>
<td>DEM GTOPO30, obtained using the spatial analyses toolset of ArcGIS® software version 10.2 (ESRI, Redlands, CA, USA)</td>
</tr>
<tr>
<td>NDVI*</td>
<td>Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra satellite (<a href="https://lpdaac.usgs.gov/">https://lpdaac.usgs.gov/</a>)</td>
</tr>
<tr>
<td>LSTday°</td>
<td>Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra satellite (<a href="https://lpdaac.usgs.gov/">https://lpdaac.usgs.gov/</a>)</td>
</tr>
<tr>
<td>LSTnight°</td>
<td>Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra satellite (<a href="https://lpdaac.usgs.gov/">https://lpdaac.usgs.gov/</a>)</td>
</tr>
<tr>
<td>BIO</td>
<td>WorldClim database (<a href="http://www.worldclim.org">http://www.worldclim.org</a>)</td>
</tr>
</tbody>
</table>

*The average seasonal normalized difference vegetation index (NDVI) for 2009-2010 (MOD13A2); °LST, land surface temperature. The average seasonal day and night land surface temperature for 2009-2010 (MOD11A2); #The average seasonal day and night land surface temperature for 2009-2010 (MOD11A2); §BIO 12 = Annual precipitation; BIO 13 = Precipitation of the wettest month; BIO 14 = Precipitation of the driest month; BIO 16 = Precipitation of the wettest quarter; BIO 17 = Precipitation of the coldest quarter; BIO 18 = Precipitation of the warmest quarter; BIO 19 = Precipitation of the driest quarter.

Table 2. Number of positive sites, ticks, pools and pathogens prevalence according to different monitoring strategies.

<table>
<thead>
<tr>
<th>Site/Collection/Outcome</th>
<th>Complete data 2009-2010</th>
<th>Strategy 1* 2010</th>
<th>Strategy 2° 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampled sites</td>
<td>52</td>
<td>16</td>
<td>52</td>
</tr>
<tr>
<td>Number of dragging sessions</td>
<td>188</td>
<td>112</td>
<td>52</td>
</tr>
<tr>
<td>Positive sites</td>
<td>28</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Number of collected I. ricinus (adults/ nymphs/larvae)</td>
<td>341 (54/95/192)</td>
<td>200 (15/25/160)</td>
<td>96 (30/57/9)</td>
</tr>
<tr>
<td>Number of pools analysed</td>
<td>118</td>
<td>46</td>
<td>50</td>
</tr>
<tr>
<td>Borrelia burgdorferi s.l. Positive/analysed pools (%) (95% CI)</td>
<td>22/118 (18.6) (CI: 11.6-25.7)</td>
<td>11/46 (23.9) (CI: 11.6-36.2)</td>
<td>5/50 (10.0) (CI: 1.7-18.3)</td>
</tr>
<tr>
<td>Rickettsia spp. Positive/analysed pools (%) (95% CI)</td>
<td>30/118 (25.4) (CI: 17.6-33.3)</td>
<td>19/46 (41.3) (CI: 27.0-55.5)</td>
<td>8/50 (16.0) (CI: 5.8-26.7)</td>
</tr>
</tbody>
</table>

*Model 1: permanent sites - April to October; °Model 2: temporary sites - April to October.

Results

Tick collection and detection of pathogens

I. ricinus was observed at 28 (53.8%) out of 52 monitored sites. Besides, 63 (33.5%) out of 188 dragging sessions were positive for this species. Overall, 341 specimens were collected: 54 adults, 95 nymphs and 192 larvae, with an overall average density of 2.05 ticks/100m². Molecular analysis showed a high prevalence of Borrelia burgdorferi s.l. (15/54 positive adults, 5/35 positive pools of nymphs and 2/29 positive pools of larvae) and of Rickettsia spp. (11/54 positive adults, 10/35 positive pools of nymphs and 9/29 positive pools of larvae).

Comparison of tick monitoring strategies

The results of the comparison among the different monitoring strategies and the results obtained from the complete dataset are shown in Table 2. Strategy 1 permitted collection of more specimens than Strategy 2, but a lower number of pools were analysed at the end with this strategy. This was due to the occasional recovery of many larvae (n=120) in two samplings (July and August 2010) at one single site, and all these larvae were grouped in a few pools according to the defined methodology. The occurrence of the two pathogens (Borrelia burgdorferi s.l. and Rickettsia spp.) was detected using both strategies and their prevalence values were not significantly different from those obtained using the complete dataset. However, Strategy 1 estimated a higher prevalence than Strategy 2; for Rickettsia spp. this difference was significant.

Tick habitat suitability model

Based on the exploratory and correlation analyses (Pearson correlation test was used to exclude highly correlated environmental predictors, retaining only those with the perceived highest biological meaningfulness for ticks) the final set of predictor variables included the following variables for the two models: the NDVI seasonal averages for spring and autumn and the land surface temperature during the light hours (LSTday) in the summer.

The MaxEnt produced tick habitat suitability maps (binary and continuous outputs) for Model 1 and 2 are shown in Figure 2. Both models performed well as measured by the leave-one-out validation approach, i.e. 10 out of 12 excluded points predicted for Model 1 (P=0.005) and 21 out of 24 for Model 2 (P<0.001). The overlap among the model continuous outputs was high (Schoener’s D=0.93) indicating that the model predictions did not change significantly with respect to the input training dataset (12 and 24
The jackknife test of variable importance showed that NDVI_spring was the most influential variable, in both models, indicating that NDVI is the most important predictor for tick distribution in the current study. It was also the environmental variable that decreased the gain (a measure of goodness of fit closely related to deviance) the most when omitted, which therefore is a reflection that this variable in isolation is the strongest contributor (Table 3).

The set of response graphs produced by generating a model using only the corresponding variable, were similar for the two models and showed that the highest relative probability of *I. ricinus* presence was predicted in areas with more dense vegetation cover during autumn and spring resulting in higher NDVI values (>0.85 and >0.78, respectively) and lower LST_day values (T°<24°C).

### Table 3. The results of jackknife test of variable importance as measured by the regularized training gain for *I. ricinus* for the two MaxEnt models.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training gain* Without variable</th>
<th>Training gain* With only variable</th>
<th>Training gain° Without variable</th>
<th>Training gain° With only variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>LST_day in summer</td>
<td>0.42</td>
<td>0.04</td>
<td>0.27</td>
<td>0.05</td>
</tr>
<tr>
<td>NDVI in spring</td>
<td>0.30</td>
<td>0.40</td>
<td>0.12</td>
<td>0.27</td>
</tr>
<tr>
<td>NDVI in autumn</td>
<td>0.40</td>
<td>0.30</td>
<td>0.27</td>
<td>0.12</td>
</tr>
</tbody>
</table>

LST_day, Land Surface Temperature during light hours; NDVI, normalized difference vegetation index. *Model 1: 12 sites, area under the curve = 0.78 (averaged over the replicate runs); °Model 2: 24 sites, area under the curve = 0.73 (averaged over the replicate runs).
Discussion

*Ixodes ricinus* was found to be the most prevalent tick species in the study area, but at low density (about 2 ticks/100m²) compared to pre-alpine areas of North-eastern Italy (Nazzi et al., 2010; Tagliapietra et al., 2011). This may be due to environmental characteristics that are not the optimal for this tick species, and to limited presence of wild ruminants.

The comparison between the two sampling strategies in the present study suggests that monitoring temporary sites only once/year (Strategy 2) allows the discovery of more positive sites and seems to ensure more cost-effectiveness in term of number of pools analysed: 50 pools with 52 dragging sessions in Strategy 2 vs 46 pools with 112 sessions in Strategy 1. The more pools available for pathogens detection, the more significant the results of molecular analyses. Strategy 1 seems to over-estimate the true prevalence and this may be due to sites with many positive ticks that are repeatedly sampled. However, both strategies detected the pathogens circulating in the area and the estimated prevalence was not significantly different from what was estimated using the complete dataset. Of note, Strategy 2 obtained this result with about one-fourth of the dragging sessions of the whole survey. Clearly, a reduction in sampling effort corresponds to an abatement of sampling-costs (reduction of travel and operators costs), which are one of the more relevant expenditures of monitoring, as previously demonstrated (Capelli et al., 2012). Monitoring fewer permanent sites (Strategy 1) may help in a sound description of tick species seasonality, but it represents an additional cost when the main objectives of surveillance activity are the development of a risk map (habitat suitability) and the detection of TBPs.

Overall, the models demonstrated a good predictive ability, which increased with the number of sites used to train the model, confirming the outcome of previous studies (Hernandez et al., 2006; Pearson et al., 2007; Wisz et al., 2008). However, the MaxEnt predictions based on the two sub-sets of data obtained a high niche overlap index and both produced realistic distribution maps, even the one with a low number of presence sites (n=12) as input.

The SDM approach identified a number climatic and other environmental conditions suitable for *I. ricinus* and, as such, reflects the known ecological requirements of the *I. ricinus* species. The most important variable for development of the models was NDVI$_{spring}$ followed by the NDVI$_{autumn}$. Spring corresponds to the period of peak tick activity, whereas autumn represents both the period of the second peak and the beginning of diapause, according to climatic condition of northern Italy (Piccolin et al., 2006; Ceballos et al., 2014). These results are consistent with the biological and ecological requirements of ticks, which are sensitive to desiccation during the environmental phases of their life cycle and therefore can survive only in areas with good vegetation cover, which provides relatively high moisture levels (Gray, 1998). Furthermore, the results are consistent with previous studies demonstrating that NDVI is a good predictor for *I. ricinus* ticks, being a measure of photosynthetic activity of vegetation and therefore a good proxy for soil moisture (Wang et al., 2007; Ruiz-Fons et al., 2012; Estrada-Peña et al., 2013; Ceballos et al., 2014). We also found that that low day-time summer temperatures were associated with higher relative probability of *I. ricinus* presence. This finding confirms the hypothesis generated by our previous preliminary ecological analysis, based only on field-collected environmental data (Drigo et al., 2011), that a significant association between the presence of adult ticks and relatively low temperatures exist. These results suggest that the main limiting factor in this area is represented by the degree of heat during summer and dryness during spring and autumn, in place of the cold winter of the pre-alpine environment.

Conclusions

The implementation of cost-effective sampling strategies may serve to optimize available resources, achieving an efficient and cost-effective tick and TBP surveillance. In this study the analyses of tick data collected during a 2-year field investigation demonstrated that TBP detection and map development can be obtained in the study area with a reduced sampling effort. Based on our findings, we suggest that the most cost-effective approach is to monitor tick presence only once, at an appropriate number of sites relative to the study area, and then use the positive sites to develop tick habitat suitability maps, by means of an SDM-remote sensing approach. A wider application of this kind of integrated approach is foreseen, but needs to be tested in other areas at risk of ticks’ infestation, to prove its effectiveness in large-scale survey plans.

References


Moderate Resolution Imaging Spectroradiometer (MODIS), onboard the Terra satellite. Available from: https://lpdaac.usgs.gov/

