Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Large-scale mapping of tick communities (Ixodidae) uncovers the tick dark diversity

Agustín Estrada-Peña

antricola@me.com

University of Zaragoza

Research Article

Keywords: ticks, co-occurring communites, spatial modelling, environmenatl filtering

Posted Date: April 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9418011/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Experimental and Applied Acarology on August 27th, 2026. See the published version at https://doi.org/10.1007/s10493-026-01172-5.

Abstract

Tick-borne disease systems are inherently multi-species networks, yet most predictive frameworks model vector distributions independently, overlooking community structure. Here, we develop a scalable approach to infer and map tick communities from species distribution models (SDMs), using continuous habitat suitability to reconstruct patterns of co-occurrence across large spatial extents.

We applied this framework to seven widespread tick species across the Iberian Peninsula using high-resolution climatic and vegetation data (1 km; years 2010–2025). Species-level predictions were integrated and clustered using Gaussian mixture models to identify statistically supported communities, followed by analyses of dominance and specialization.

Climatic variables alone achieved high predictive performance, with limited gains from vegetation descriptors. Species distributions showed extensive spatial overlap, revealing that most of the territory supports multi-species assemblages rather than isolated populations. Community reconstruction identified a structured gradient from Atlantic systems dominated by *Ixodes ricinus* to Mediterranean assemblages characterized by *Hyalomma lusitanicum* and *Rhipicephalus bursa*, with transitional communities forming a continuous ecological spectrum, demonstrating a strong environmental filtering in Mediterranean assemblages.

Our results show that SDM-derived suitability surfaces can be leveraged to infer biologically meaningful community structure, providing a practical alternative to data-intensive joint modelling approaches. This framework offers a robust and scalable tool to integrate community ecology into vector surveillance and to improve spatial risk assessment of tick-borne diseases.

Introduction

The concept of parasite community ecology emerged in the late twentieth century, when it became evident that parasite assemblages behave analogously to free-living ecological communities, exhibiting processes such as competition, facilitation, niche partitioning, and environmental filtering (e.g. Poulin, 2007; Yoder et al., 2005; Williams et al., 2013; Dallas et al., 2019). These concepts are particularly well suited for ticks, since they represent one of the most intensively studied parasite groups; multiple species f ticks frequently coexist within the same habitats and exploit overlapping host communities (Sonenshine & Roe, 2014). This is of importance when addressing issue related to human and animal health, and the circulation of tick-bone pathogens (e.g. Aardema and Von Loewenich, 2015; Rataud et al., 2024; Levi et al., 2016)

Ticks rarely occur in complete isolation across broad geographic regions. Even in areas commonly described as dominated by a single species, such as large parts of Central Europe with *Ixodes ricinus*, or extensive regions of northern Eurasia with *Ixodes persulcatus*, this perception often reflects a sampling bias driven by medical relevance rather than true ecological dominance (Estrada-Peña et al., 2015; Randolph, 2009). Standard field surveys tend to emphasize species of epidemiological importance, while overlooking the full assemblage of co-occurring ticks, including both Ixodidae and Argasidae, distributed along climatic and physiographic gradients. Only in low-diversity systems, such as isolated islands or Arctic–Antarctic

environments, do tick assemblages tend to be species-poor or even monospecific (Guglielmone et al., 2014). In contrast, most terrestrial biotopes harbor multi-species tick assemblages, together with diverse pathogen communities and vertebrate hosts that sustain complex transmission cycles (Ostfeld & Keesing, 2012). These assemblages form tightly coupled ecological systems structured by host–parasite interactions, environmental constraints, and the population dynamics of both ticks and vertebrates.

Despite this ecological complexity, the distribution of ticks has historically been modeled on a species-by-species basis. Even when multiple species were analyzed within the same study, modelling frameworks typically treated them independently, ignoring potential co-occurrence patterns and interspecific dependencies. As a result, our understanding of tick communities as integrated ecological units remains limited. Early efforts to characterize tick assemblages across environmental gradients (e.g. Estrada-Peña et al., 2015) demonstrated that climate and vegetation structure are primary determinants of community composition, but also revealed non-random patterns of co-occurrence suggestive of ecological interactions among species occupying similar niches. However, such studies have largely been restricted to local or regional scales, often constrained by targeted sampling designs.

Parallel research has highlighted that interactions among ticks, vertebrate hosts, and pathogens form multi-layered transmission networks, where ignoring one or more components can lead to biased epidemiological inference (Ostfeld et al., 2014; Levi et al., 2016). A simplified focus on single vectors or pathogens may fail to capture the emergent properties of these systems, including amplification or dilution effects driven by community composition. This gap is particularly critical in the context of species distribution modelling (SDM), where the question of “which species occur when the focal species is absent” remains largely unaddressed. Neglecting the spatial structure of co-occurring species results in an incomplete representation of ecological gradients and undermines the use of tick distributions as indicators of disease risk.

An emerging direction in vector ecology is the explicit modelling of multi-trophic communities, integrating hosts, vectors, and pathogens within a unified framework. Joint Species Distribution Models (JSDMs; Warton et al., 2015; Tous et al., 2025) and related approaches provide a statistical basis to capture shared environmental responses and residual correlations among species, offering a pathway to infer community structure from observational data (Iter et al., 2024). Such approaches have been successfully applied to other vector systems, including mosquito assemblages, where both environmental filtering and interspecific interactions contribute to community organization (e.g. Carlson et al., 2004; Golding et al., 2015; Tikhonov et al., 2017;. In these systems, climatic gradients explain a substantial fraction of community composition, while species associations point to processes such as competition and habitat partitioning. These approaches have been never focused on the analysis of tick communities, and modelling work is still focused on the mapping of single species (Zannou et al., 2021; Kopsco et al., 2022) with few efforts towards the understanding of the gradients of tick co-occurrence (e.g. Estrada-Peña et al., 2025)

In vector-borne disease ecology, this perspective aligns with a community-based understanding of transmission, where outcomes depend on the joint effects of host diversity, vector assemblages, pathogen

communities, and environmental drivers (Ostfeld & Keesing, 2012). Concepts such as the dilution effect illustrate how increased host diversity can reduce pathogen transmission by redistributing vector feeding toward less competent hosts (Keesing et al., 2006). In tick systems, these relationships are further complicated by interactions across multiple spatial and temporal scales, including land use, landscape configuration, and climatic variability, which jointly shape both tick and host communities. Although these factors have been explored individually, an integrated, harmonized framework capturing their combined effects on tick community structure remains lacking.

A further challenge arises from detection bias. Tick species differ markedly in detectability due to variation in sampling methods, life stages, and activity patterns. Failure to account for imperfect detection can distort estimates of species associations and community composition. Hierarchical occupancy models, often embedded within JSDM frameworks, offer a partial solution by separating detection processes from ecological processes (Dorazio & Royle, 2005; Ovaskainen et al., 2017). However, these approaches primarily quantify correlations rather than explicitly delineating community assemblages from raw occurrence data.

Here, we address these limitations by developing a high-resolution, multi-species tick distribution prediction framework aimed at mapping communities of co-occurring ticks. The approach is designed to identify epidemiologically meaningful assemblages across environmental gradients, rather than focusing on individual species. We apply this framework to the Iberian Peninsula, a region of high tick diversity, using climate and vegetation data at 1-km spatial resolution for the period 2010–2025, as a proof of concept. The framework is aimed at the detection of co-occurring ticks in large geographical regions, which has an impact on the circulation of tick-borne pathogens. This allows us to assess the implications for disease risk in both human and animal populations, and how the joint modelling strategy improves the knowledge of the ticks at the level of countries. Ultimately, this approach provides a scalable basis for integrating community ecology into the surveillance and modelling of tick-borne diseases.

Material and Methods

Aim and modelling background.

We aimed to reconstruct spatially explicit tick communities from individual species models (SDMs) while preserving the probabilistic information generated by the SDMs. We applied the method to map the distribution of eight species of ticks that have been reported in the Iberian Peninsula. This is a heterogeneous territory, the second higher in altitude in Europe, with link to the wet and cold Atlantic climate in North-West and the warm and dry Mediterranean climate in South-East. The territory is thus a gradient of climate and vegetation features, as well as for vertebrates and ticks. Figure 1 summarizes the raw details of the Iberian Peninsula in elevation, biogeographical, and climate regions. Figure 2 is an appraisal of the hypotheses and the modelling steps of this study.

Figure 1.

Figure 1. The flow of work followed in this study. After collecting and cleaning the records of ticks in the Iberian Peninsula (A), we integrated climate and vegetation data (B) in a hexagonal grid of 1 km. resolution, on which models were carried out separately for each species (C). Ecah hexagon carries the vector of probabilities of the occurrence of each species (C), that is examined by Principal Components Analysis to remove spurious values. Species are clustered in communities using a Correspondence Analysis, together with other methods to correlate the spatial range of communities with ecological or climate regions.

Hexagonal grids are commonly used in spatial ecology because they minimize directional bias and provide uniform neighborhood relationships among cells. We integrated the complete modelling framework into a set of 770,762 hexagons, 1 km. resolution each, covering the complete territory in the Iberian Peninsula, and clipped in the periphery to accommodate the coastline. All the records of ticks were annotated into each corresponding hexagon, as present-absent. The climate and vegetation data (see below) were transferred to hexagons by simple GIS algebra operations. All the modelling steps, the projection into the complete area of IP, calculation of communities, and further estimation of entropy (see below) were referred to the hexagonal grid.

The flow of work in modelling consists of (a) modelling the occurrence of ticks in Iberian Peninsula, estimating species-specific habitat suitability using Random Forest models (after testing other algorithms, see below), (b) transformation of species suitability predictions into a community-level suitability index using a geometric mean approach, (c) derivation of community composition matrices at the hexagonal spatial unit, representing the predicted probability of occurrence of each species within each spatial cell, using Principal Components Analysis and Gaussian Mixture Models, (d) further calculations to capture the dominant species, the species diversity of each community, and the links of each tick community with standard definitions of climate or biogeographical regions.

Occurrence data

We used a set of georeferenced collections of 8 species of ticks in the Iberian Peninsula, namely *Ixodes ricinus*, *Dermacentor marginatus*, *Dermacentor reticulatus*, *Haemaphysalis punctata*, *Haemaphysalis sulcata*, *Hyalomma marginatum*, *Hyalomma lusitanicum*, and *Rhipicephalus bursa*. *Ixodes hexagonus*, *Ixodes trianguliceps*, *Hyalomma scupense* and *Haemaphysalis hispanica* were omitted because the low number of records and their endophilic behaviour, therefore rendering unsuitable the modelling of their niches using only climate features. *Rhipicephalus sanguineus* s.l. was omitted after initial tests in which issues attributed to the identification of the ticks were noticed. Of interest, the group *R. sanguineus* shows a predicted range incompatible with its known distribution, with low sensitivity and specificity, and high threshold. We consider that these results derive of two important facts, namely (1) *R. sanguineus* s.s. is an endophilic species, therefore difficult to collect with the classic flagging/dragging methods, and thus hard to model with variables explaining the exophilic environment, and (2) the existing confusion between the nominal species and other species in the group, that have been named either *Rhipicephalus turanicus* or *Rhipicephalus hibericus* (see Nava et al., 2018, Millán et al., 2024). Since it is not possible to check the reliability of identifications from the dataset of records, and only approximately 20% of records were

available for re-examination, we preferred not to include the group of species in the study use. The initial results of modelling as "*R. sanguineus* s.l." are reported for completeness.

We used records from a previous compilation of 1,636 unique records from 2,285 surveys and more than 30,000 tick specimens as surveyed in the period 1995–2020 (Estrada-Peña et al., 2025), and data available in GBIF (2025). Data from GBIF were used if coordinates had a resolution of +/- 300 meters. We explored and cleaned the records for overlay on land masses and for duplications. Ultimately, we compiled and analysed a dataset of 2,302 unique tick records, with the 100% having a spatial resolution below 300 m.

Explanatory variables

Environmental variables consisted of a set of climate-derived predictors representing temperature and atmospheric moisture conditions, plus a set of vegetation features. These explanatory variables were extracted from raster layers and associated with occurrence locations to construct the modelling dataset.

Climate data used to model environmental suitability for ticks were obtained from the CHELSA repository (https://www.chelsa-climate.org). The choice of this dataset is because is a highly resolution one, compatible with the modelling framework at 1 km. and because it includes measurements of water in the air, that results in better environmental proxies for ticks than the precipitation (Estrada-Peña et al., 2016). Monthly data for maximum temperature (TMAX), minimum temperature (TMIN) and vapor pressure deficit (VPD) were compiled across the target area for the period 1995–2020, at monthly intervals. VPD is a measure of the deficit of water in the air and is related with relative humidity. Harmonic regression (Scharlemann et al. 2008) was performed on the monthly averages for the complete period, independently for each variable, and the first three coefficients of each regression were retained as explanatory variables. This yielded nine climatic predictors: three coefficients each for TMAX, TMIN and VPD. All the calculations were done in R (v. 4.0.2, R Core Team 2025) using the package terra (v 1.8–60, Hijmans et al., 2025).

Vegetation data were derived from the CORINE dataset of European vegetation types of the year 2018, at 100 meters resolution, available at https://land.copernicus.eu/en/products/corine-land-cover. CORINE is an harmonized classification system of vegetation for the complete Europe. We derived a set of explanatory variables from the CORINE raw raster, transferred to the 1 km. hexagons covering the target territory. The explanatory variables obtained from the vegetation were (a) the majority of the vegetation category (i.e. the category of vegetation of the largest patch in each hexagon), (b) the number of different categories of vegetation (i.e. the fragmentation of the territory), (c) the category of vegetation of the largest patch, and (d) the minority of the vegetation category (i.e. the less represented vegetation). All these variables were obtained by basic routines in qGIS software (QGIS Development Team, 2024).

Random Forest modelling

Species distributions were modelled independently for each species of tick using the Random Forest (RF) algorithm, an ensemble machine learning method widely used in ecological niche modelling due to its high predictive performance and ability to capture complex nonlinear relationships between species occurrence

and environmental predictors (Zhang et al., 2019; Rongcheng et al., 2025). Random Forest provides a flexible, non-parametric framework capable of capturing complex, non-linear environmental relationships using only presence data, making it particularly suitable when absences are unreliable or sampling is sparse and no abundance counts are available.

Random Forest constructs numerous classification trees using bootstrap samples of the training data. At each node, a random subset of predictors is evaluated to determine the best split, thereby reducing overfitting and improving model generalization. For each species, the model output consisted of a continuous probability of occurrence, representing the potential suitability for the species across the study area. In all cases, 70% of the records were used for model training, while 30% were used for model testing. The algorithm generated pseudo-absences in a proportion 1:1 to the actual records of the organism being modelled. To select the best models we considered several metrics: sensitivity (the proportion of correctly predicted presences), specificity (the proportion of correctly predicted absences), the threshold value (the value that maximises specificity and sensitivity), the omission rate (the proportion of missing records with the best model), and the proportion of correctly detected records (overall accuracy). The confusion matrix was computed by converting habitat suitability maps into binary presence/absence maps obtained from the optimal threshold of sensitivity and specificity for each species. Additionally, we included the Area Under the Curve (AUC) for comparison purposes among the different candidate models, although its usefulness as a metric for final model performance is limited to comparisons within the same species. Model predictions were generated for all hexagonal cells covering the study region, producing a spatially explicit probability surface for each species.

Prediction of species probabilities across spatial units

To quantify the suitability of each spatial unit for the tick community as a whole, we developed and calculated a *community* suitability index using the geometric mean of species potential suitability in each hexagon. The geometric mean is particularly appropriate when combining probabilities because it preserves the multiplicative structure of independent probabilities and reduces the influence of extreme values.

For each hexagon h, the community suitability index C_h was calculated as:

$$C_h = \left(\prod_{i=1}^{S} P_{ih} \right)^{1/S}$$

where P_{ih} is the predicted probability of occurrence of species i in hexagon h, and S is the total number of species included in the analysis.

This metric provides an estimate of the overall potential suitability for the assemblage of tick species, emphasising areas where multiple species are simultaneously predicted to occur. The final output of this step in the modelling framework consisted of a community composition matrix where each hexagon is characterized by a vector of predicted probabilities for all species.

Formally, for a hexagon h:

$$\mathbf{P}_h = (P_{1h}, P_{2h}, \dots, P_{Sh})$$

where P_h represents the predicted probability of occurrence of species i. These vectors describe the expected composition of tick communities across space, allowing spatial comparisons among hexagons in terms of species assemblages. The vector is then scaled

$$P'_{h,s} = \frac{P_{h,s}}{\sum_s P_{h,s}}$$

and normalized, because the subsequent steps are sensitive to scale:

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

Therefore, each hexagon could host a community of up to eight tick species; however, it is highly improbable that a single hexagon contains all eight species. It is thus necessary to remove noise (i.e., low-probability signals) and infer the effective number of communities. This was achieved using a Principal Component Analysis (PCA). Probabilities within each vector are expected to be correlated but structured along gradients of occurrence. PCA therefore orders and filters these gradients, retaining the dominant structure while discarding non-informative variation.

We retained the PCA axes explaining 85% of the variance, corresponding in our case to four components. Consequently, each hexagon (the minimal mapping unit) is represented by four coordinates describing its position in a four-dimensions ecological space. These coordinates were then used in a Gaussian Mixture Model (GMM) to estimate the number of communities across the > 770,000 hexagons covering the study area. Formally:

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x \mid \mu_k, \Sigma_k)$$

The GMM detects Gaussian clusters, that are the separated communities. These communities result from the ecological mixtures detected in the points above: each hexagon can be classified as a community, considering the scaled and normalized vector of probability of co-occurrence of each species. Moreover, each community can be handled to calculate the dominating species (the maximum scaled and normalized probability) and the entropy, or the incertitude of the assignation to a given community. If a hexagon has low values for every species probability, it belongs to an "inter-community" territory, like an ecological gradient among two or more communities.

Correlation with ecological regions

We wanted to leverage existing official descriptors of the climate and the vegetal cover and aimed to correlated the spatial range of each tick community with these descriptors. This correspondence is not

aimed to use broad vegetation definitions (e.g. forest, sclerophilous vegetation, etc.) that are too broad as to have a clear correlation with tick communities. To evaluate the relationship between tick community composition and broad-scale environmental structure, we compared the spatial distribution of communities with two independent categorical frameworks. Climatic regions for the target territory were obtained from the LANMAP2 initiative (Mücher et al., 2006) that has been used in previous estudies on ticks in Europe (e.g. Estrada-Peña and Fernández-Ruiz, 2020). Six categories of climate (Fig. 1B) cover the Iberian Peninsula, namely Alpine South, Atlantic, Lusitanian, Mediterranean Mountains, Mediterranean North and Mediterranean South (Mücher et al., 2006). Biogeographic regions followed Olson et al. (2001). This is a harmonized and standard classification of the ecosystems of the World at different hierarchical levels, of which we used those at highest resolution, resulting in eight regions for the Iberian Peninsula (see Fig. 1C).

To associate each hexagon with an environmental category, we overlaid the hexagonal grid with the corresponding categorical layer (climatic or biogeographic regions). In cases where hexagons intersected multiple polygons, a dominant category was assigned based on the largest area of overlap, ensuring a one-to-one correspondence between each hexagon and a single environmental class. We constructed contingency tables describing the frequency of each community across environmental categories. Row-wise proportions were calculated to estimate the conditional probability of environmental categories given community identity.

We analyzed the structure of the associations using the Cramer's V, and a correspondence analysis (CA). Cramer's V provides a standardized measure of association ranging from 0 (no association) to 1 (perfect association), allowing comparison across different classification schemes. In example, Cramer's V explores if a tick community X is associated with any of the categorical descriptors (e.g., a specialisation). Then, the CA measures the structure of the associations between communities and categorical descriptors.

CA is an ordination technique specifically suited for categorical data, enabling visualization of relationships between rows (communities) and columns (environmental classes) in a reduced-dimensional space. Community membership was treated as a categorical response variable, and environmental class (climate or ecoregion) as the predictor. The model estimates the log-odds of assignment to each community relative to a reference category as a function of environmental class. From this model, we derived predicted probabilities of community occurrence within each environmental category. These probabilities provide an interpretable measure of how strongly each community is associated with specific environmental conditions. The complete flow of work is included in Figure XXX; a script of the R programming environment for the complete flow of work is provided as supplementary material 1.

Results

High-resolution vegetation features do not improve the predictive performance

Models trained using climatic variables alone showed consistently high predictive performance across all species but for *Rhipicephalus sanguineus s.l.* (Table 1). The most informative predictors included the nine variables describing annual averages and seasonal climate dynamics, specifically the spring rate of increase of maximum (TMAX) and minimum (TMIN) temperatures and the autumn rate of decrease of TMAX and TMIN, as well as vapour pressure deficit (VPD). As mentioned above, models of the group of species *R. sanguineus* s.l. underperformed. We consider that this resulted from the inclusion of records of a mix of species that have been unreliably reported. We show here the indexes of performance of the models for *R. sanguineus* s.l., but the group of species is removed from further results.

Table 1

The indexes of performance of models for individual tick species carried out using Random Forests and nine climate variables in Europe. Values were derived from CORINE to obtain the dominant vegetation, the fragmentation, and the less represented vegetation in each hexagon of the grid. Indexes included are the Threshold, AUC, Omission rate, Sensitivity, Specificity, and Accuracy. The Threshold (cut-off value) is a scalar t [0,1] used to convert continuous model outputs (e.g., predicted probabilities) into binary predictions: the choice of t determines the trade-off between different error types (false positives vs. false negatives). In our study, it is the value that maximises the sum of sensitivity and specificity. The Threshold is the value that maximises sensitivity and specificity. The AUC is the area under the Receiver Operating Characteristic (ROC) curve, which plots Sensitivity vs. 1 - Specificity across all possible thresholds. Formally, AUC is the probability that a randomly chosen positive instance has a higher predicted score than a randomly chosen negative instance. AUC is threshold-independent and ranges from 0.5 (random) to 1 (perfect discrimination). The omission rate is the proportion of observed presences incorrectly predicted as absences (false negatives: omission rate = 1 - Specificity). Specificity represents the proportion of correctly predicted absences and thus the quantification of commission errors; It quantifies the model's ability to detect presences. Sensitivity represents the proportion of correctly predicted presences and thus the quantification of omission errors; It quantifies the model's ability to correctly reject absences. Accuracy is the proportion of correctly classified instances (both positives and negatives).

CLIMATE ALONE	**Threshold**	**AUC**	**Omission rate**	**Sensitivity**	**Specificity**	**Accuracy**
Dermacentor marginatus	0.4480	0.9129	0.1569	0.8421	0.8438	0.8431
Dermacentor reticulatus	0.5050	0.8903	0.1733	0.8262	0.8272	0.8267
Haemaphysalis punctata	0.4797	0.9038	0.1680	0.8315	0.8325	0.8320
Haemaphysalis sulcata	0.3731	0.8201	0.1980	0.8765	0.7912	0.7791
Hyalomma lusitanicum	0.4360	0.8623	0.2151	0.7853	0.7847	0.7849
Hyalomma marginatum	0.4350	0.8492	0.2152	0.7853	0.7843	0.7848
Ixodes ricinus	0.4227	0.8625	0.2114	0.7886	0.7886	0.7886
Rhipicephalus bursa	0.4900	0.8700	0.2108	0.7899	0.7887	0.7892
Rhipicephalus sanguineus s.l.	0.5080	0.6669	0.1986	0.7028	0.8000	0.8014

Table 2

The indexes of performance of models for individual tick species carried out using Random Forests and nine climate variables and high-resolution vegetation descriptors in Europe. Values were derived from CORINE to obtain the dominant vegetation, the fragmentation, and the less represented vegetation in each hexagon of the grid. Indexes included are the Threshold, AUC, Omission rate, Sensitivity, Specificity, and Accuracy. The Threshold (cut-off value) is a scalar t [0,1] used to convert continuous model outputs (e.g., predicted probabilities) into binary predictions: the choice of t determines the trade-off between different error types (false positives vs. false negatives). In our study, it is the value that maximises the sum of sensitivity and specificity. The Threshold is the value that maximises sensitivity and specificity. The AUC is the area under the Receiver Operating Characteristic (ROC) curve, which plots Sensitivity vs. 1 - Specificity across all possible thresholds. Formally, AUC is the probability that a randomly chosen positive instance has a higher predicted score than a randomly chosen negative instance. AUC is threshold-independent and ranges from 0.5 (random) to 1 (perfect discrimination). The omission rate is the proportion of observed presences incorrectly predicted as absences (false negatives: omission rate = 1 - Specificity). Specificity represents the proportion of correctly predicted absences and thus the quantification of commission errors; It quantifies the model's ability to detect presences. Sensitivity represents the proportion of correctly predicted presences and thus the quantification of omission errors; It quantifies the model's ability to correctly reject absences. Accuracy is the proportion of correctly classified instances (both positives and negatives).

CLIMATE AND VEGEGATION	Threshold	AUC	Omisson rate	Sensitivity	Specificity	Accuracy
Dermacentor marginatus	0.2805	0.9578	0.4677	0.9012	0.5298	0.5322
Dermacentor reticulatus	0.5974	0.9639	0.0866	0.9136	0.9130	0.9133
Haemaphysalis punctata	0.6305	0.9322	0.0981	0.9034	0.9049	0.9082
Haemaphysalis sulcata	0.7736	0.9476	0.1008	0.9001	0.8982	0.8991
Hyalomma lusitanicum	0.4360	0.8623	0.2151	0.7853	0.7847	0.7849
Hyalomma marginatum	0.3970	0.8758	0.2039	0.7954	0.7967	0.7961
Ixodes ricinus	0.4350	0.8492	0.2152	0.7853	0.7843	0.7848
Rhipicephalus bursa	0.4207	0.9467	0.2249	0.8652	0.7699	0.8750
Rhipicephalus sanguineus s.l.	0.5207	0.6467	0.5201	0.6652	0.5899	0.4150

Incorporating high-resolution vegetation descriptors (including dominant vegetation types and fragmentation metrics) did not result in a substantial improvement in model performance (Table 2). While some species exhibited slight increases in AUC when vegetation variables were included, these gains were often accompanied by reduced specificity or inflated omission rates, indicating limited overall improvement. Despite this, models including the full set of predictors (climate + vegetation) were retained for subsequent analyses to preserve potential fine-scale ecological information.

Dominant tick species emerge from overlapping species distributions

The predicted distributions of individual tick species showed extensive spatial overlap across the Iberian Peninsula. *Rhipicephalus bursa* had the largest predicted range covering 80.7% of the study area, followed

by *Hyalomma marginatum* (48.7%), *Hyalomma lusitanicum* (48.5%) and *Ixodes ricinus* (32.3%). Other species had more restricted distributions: *Haemaphysalis sulcata* (16.3%), *Haemaphysalis punctata* (15.3%), *Dermacentor marginatus* (16.2%) and *Dermacentor reticulatus* (16.1%). The cumulative occupancy exceeded 100% because multiple species were predicted to co-occur within the same areas reflecting overlapping ecological niches rather than mutually exclusive distributions.

By analysing the relative probabilities of occurrence within each grid cell, we identified the dominant species across the study area (Fig. 3). Dominance was defined as the species with the highest predicted suitability in each cell. *Hyalomma lusitanicum* was the dominant species in the largest proportion of the Iberian Peninsula (46.5%) followed by *R. bursa* (19.5%) and *I. ricinus* (16.8%). Other species were dominant in smaller fractions: *D. marginatus* (5.8%), *H. punctata* (4.4%), *H. marginatum* (4.1%), *D. reticulatus* (2.3%) and *H. sulcata* (0.4%). Approximately 3.2% of grid cells were classified as uncertain because two or more species had similar potential suitability and similar probabilities of occurrence. Notably *H. sulcata* showed very low dominance over the entire target territory.

A clear geographic gradient in tick dominance was observed, primarily along the Atlantic–Mediterranean axis from northwest to southeast. This gradient was further influenced by altitude and latitude. *Ixodes ricinus* dominates most Atlantic regions but is expected to spread more widely at lower relative abundance. *Dermacentor reticulatus* and *Haemaphysalis punctata* show patchy dominance in western regions and mountainous areas like the Pyrenees and parts of the central-northern Iberian Peninsula. Conversely, southern and southeastern regions are mainly dominated by *H. lusitanicum*, *R. bursa* and, to a lesser extent, *H. marginatum*. *Dermacentor marginatus* is mainly found at intermediate altitudes where it alternates with other dominant species.

Integrating species-level suitability predictions into a multivariate framework and using Gaussian mixture modelling (GMM) revealed 12 communities of co-occurring ticks. These communities are statistically robust clusters in the multidimensional space of species composition and provide a spatially explicit representation of tick assemblages. Some communities share the same species but differ in relative abundance. We compacted similar groups based on species composition but in different proportions, resulting in ten communities with different species combinations. The complete set of 12 communities (without compacting the similar ones) is shown in Fig. 4; Table 3 includes the final set of tick communities, after compacting the similar ones. Notably, *H. sulcata* and *D. reticulatus* were excluded from every community due to the greater probability of other communities in the territory.

Table 3
The communities of ticks and its species composition. Each community is numbered consecutively as detected by the modelling framework, for simplification. The communities were obtained from a vector with the suitability of the eight species at each hexagon of the grid covering the territory. Noise was removed using a Principal Components Analysis and communities were recognized after a Gaussian Multilineal Model. The dominant species is the tick that displayed the maximum suitability in the community, the secondary species are those that ranked second or third in the building of communities. It can be observed that some combinations of species may appear twice; this is because the relative composition of suitability for each taxon is different according to the community.

Community	Dominant species	Secondary species
1	*Hlusitanicum*	*Hmarginatum + Rbursa*
2	*Hlusitanicum*	
3	*Rbursa*	*Hmarginatum*
4	*Iricinus*	*Rbursa*
5	*Iricinus*	*Hpunctata*
7	*Dmarginatus*	*Rbursa + Hlusitanicum*
9	*Hlusitanicum*	*Rbursa*
10	*Iricinus*	

The spatial distribution of communities follows a structured pattern across the Iberian Peninsula (Fig. 5). Northern regions are largely characterized by communities dominated by *I. ricinus*, particularly in coastal and humid environments. The species is predicted either alone or in combination with *R. bursa* and *H. punctata*, as the physiogeographic gradient enter into the land from the coast and becomes colder and with higher elevation. These assemblages represent the most common community types in Atlantic and northwestern regions. The community of *I. ricinus* and *H. punctata* is also predicted to occur in large areas in the Pyrenees. Toward the south and southeast, communities shift progressively to Mediterranean assemblages. A recurrent community composed of *H. lusitanicum* and *R. bursa* replaces the northern *I. ricinus*-dominated systems. In some areas, this association shows variation in dominance, with either species prevailing locally. Additional configurations of the communities involve combinations of *H. marginatum*, *H. lusitanicum*, *R. bursa*, and *D. marginatus*, forming a mosaic of assemblages across Mediterranean environments. Notably, the dry and warm areas of the Iberian peninsula are foci of biodiversity of ticks, in terms of the number of species and relative composition of communities.

Communities of ticks exhibit associations with standard descriptions of environmental features. Given the clear gradient in tick occurrence patterns across the Iberian Peninsula, we investigated whether an index of specialisation to environmental features could reveal the links between communities and ecological features. This would help identify potential areas of endemism and determine if communities are linked to

climate or specific ecological features. Cramer's V indicates a strong link between communities and climate, but for groups dominated by *I. ricinus*, along with *H. punctata* or *R. bursa*, which are not predicted to coexist (Table 4).

Table 4
The link in the range 0–1 of each community of ticks with a standard category of climate. The communities have been consecutively numbered as follows: 2. *H. lusitanicum*, 4. *I. ricinus* + *R. bursa*, 5. *I ricinus* + *H. punctata*, 6. *H. lusitanicum* + *R. bursa* + *H. marginatum*, 8. *R. bursa* + *H. marginatum*, 9. *H. lusitanicum* + *R. bursa*, 10. *I. ricinus*, 12. *D. marginatus* + *R. bursa* + *H. lusitanicum*.

H. lusitanicum	***I. ricinus* + *R. bursa***	***I. ricinus* + *H. punctata***	***H. lusitanicum* + *R. bursa***	***R. bursa* + *H. marginatum***	***H. lusitanicum* + *R. bursa***	***I. ricinus***	***D. marginatus* + *R. bursa* + *H. lusitanicum***
0.73	0.40	0.38	0.67	0.79	0.95	0.87	0.45

The resulting communities generally exhibit a clear ecological structure along major climatic gradients, particularly the transition from Atlantic to Mediterranean conditions (Fig. 5). Some communities are highly specialised, confined to specific climatic regions, while others have broader distributions and likely represent transitional assemblages (Table 4). *Ixodes ricinus*, either alone or in combination, is strongly associated with the Lusitanian-type climate (see Fig. 1 for a climate region outline). Interestingly, communities of *I. ricinus* with either *Haemaphysalis punctata* or *Rhipicephalus bursa* can colonise other climates. Five tick communities show varying degrees of link with Mediterranean-type climates: the monospecific *H. lusitanicum*, the combination of *H. lusitanicum* (or *H. marginatum* or both species) and *R. bursa* and/or *D. marginatus*. This pattern aligns with the concept of environmental filtering, where climatic conditions restrict species co-occurrence while local factors and ecological interactions introduce additional variability. However, a one-to-one correspondence between a climate descriptor and a tick community does not exist with statistical support (Fig. 5).

Table 5
The link in the range 0–1 of each community of ticks with standard definitions of biogeographic regions. The communities have been consecutively numbered as follows: 1. *H. lusitanicum* + *R.bursa* + *D. marginatus*, 2. *H. lusitanicum*, 3. *R. bursa* + *D. marginatus*, 4. *I. ricinus* + *R. bursa*, 5. *I ricinus* + *H. punctata*, 6. *H. lusitanicum* + *R. bursa* + *H. marginatum*, 7. *D. marginatus* + *R. bursa* + *H. lusitanicum*, 8. *R. bursa* + *H. marginatum*, 9. *H. lusitanicum* + *R. bursa*, 10. *I. ricinus*, 11. *H. lusitanicum* + *R. bursa*, 12. *D. marginatus* + *R. bursa* + *H. lusitanicum*. Results in bold typeface

H. lusitanicum	***I. ricinus* + *R. bursa***	***I. ricinus* + *H. punctata***	***H. lusitanicum* + *R. bursa***	***R. bursa* + *H. marginatum***	***H. lusitanicum* + *R. bursa***	***I. ricinus***	***D. marginatus* + *R. bursa* + *H. lusitanicum***
0.89	0.39	0.30	0.72	0.74	0.90	0.95	0.41

Cramer's V highlights a strong link between communities and ecological regions. However, as before, this is particularly evident for groups formed by *I. ricinus* as the dominant species alongside *H. punctata* or *R. bursa* (Table 5) and the community comprising *D. marginatus*, *R. bursa* and *H. lusitanicum*. Figure 6 illustrates the cross-relationships between communities and ecological regions, which are clearer than those derived from raw climate data. The Cantabrian mixed forest (see Figs. 1 and 6) is characterised by *I. ricinus* as the sole or dominant species. It's also found in the Northwest Iberian montane forests (NIMF) (Fig. 6) and the Pyrenees conifer and mixed forests (PCM). The Iberian sclerophyllous and semi-deciduous forests (ISSF) ecological descriptor hosts the greatest species diversity and thus the combinations of modelled tick communities, including *I. ricinus*. Interestingly, both R. bursa and H. marginatum are projected to coexist in the Northwest Iberian montane forests (NIMF) but not in areas where *I. ricinus* and *H. punctata dominate*.

Discussion

From tick species maps to community structure.

Species Distribution Models (SDMs) have become crucial tools for predicting the geographic distribution of vectors, hosts and pathogens (e.g. Eisen and Eisen, 2014). By estimating relationships between species occurrences and environmental variables, SDMs help identify suitable habitats and assess potential range shifts due to environmental change. In vector-borne disease systems, these approaches have been extensively applied to taxa like ticks, mosquitoes and sandflies (Reisen, 2010) consistently emphasising the significance of climatic drivers, particularly temperature and atmospheric moisture, in shaping species distributions.

This study focused on modelling the spatial distribution of seven tick species across the Iberian Peninsula. This allowed the reconstruction of spatially coherent tick communities. This demonstrates that SDM-derived probabilities of habitat suitability can capture not only individual species niches (Broennimann et al., 2012) but also higher-order community organisation patterns. Importantly, the resulting communities showed moderate to strong correspondence with independent environmental classifications, including climatic regions (Mücher et al., 2010) and biogeographic regions (Olson et al., 2021).

These findings are epidemiologically significant. Tick-borne pathogens can be transmitted by multiple vector species, and differences in vector competence, phenology and host associations strongly influence transmission dynamics (Dallas and Park, 2017). Community-level approaches help identify regions where multiple vectors might co-occur, potentially increasing pathogen circulation (Estrada-Peña et al., 2016). While classical spatially explicit models (SDMs) are widely used, they often treat each species independently, overlooking the ecological dependencies necessary for pathogen transmission. Modelling transmission processes separately can miss the system's emergent properties. Community-based approaches offer a more realistic framework by explicitly considering co-occurrence patterns and shared environmental constraints.

An epidemiological indicator of this framework's usefulness is its ability to investigate areas where tick-borne pathogens circulate through multiple vectors. For example, the etiological agent of Crimean-Congo

haemorrhagic fever, *Orthonairovirus haemorrhagiae*, is transmitted by *H. marginatum* and *H. lusitanicum* (Sánchez-Mora et al., 2026). Similarly, vectors of *Babesia* spp., *Theileria* app. and *Anaplasma phagocytophilum* include Defaye et al. (2002), EFSA panel AHAW (2010) and Dreghiciu et al. (2025).

Methodological robustness

Tick communities consist of co-occurring species with varying compositions and dominances, derived from continuous suitability outputs. This study doesn't compare modelling algorithms or optimise tick range predictions, as these have been extensively covered elsewhere (e.g. Alkise and Peterson, 2021; Westcott et al., 2025). The resulting communities are statistically robust, showing that SDM-derived habitat suitability probabilities capture not only individual species niches (Broennimann et al., 2012) but also higher-order community organisation.

Stacked SDMs (S-SDMs), proposed by Ferrier and Guisan (2006), combine individual species predictions to estimate species richness and composition. However, these methods don't explicitly model community structure and may produce assemblages that don't match naturally occurring communities due to statistical noise in probability estimates. More recent advances, like joint species distribution models (JSDMs), allow simultaneous modelling of multiple species and estimating residual correlations between them. Frameworks developed by Warton et al. (2015), Ovaskainen et al. (2017) or Tikhonov et al. (2020) integrate environmental drivers, species traits and hierarchical structures to explain community patterns across spatial scales. However, these approaches typically require large datasets and are designed to analyse existing communities rather than infer them from sparse or uneven data.

Our approach is conceptually simpler and particularly well-suited for situations with limited or spatially biased occurrence data, as is often the case in vector surveillance (Maxwell et al., 2023). It achieves this by integrating species-level ecological niches into a unified community-level representation. Furthermore, it avoids the information loss associated with threshold-based presence-absence transformations by retaining continuous suitability values. Finally, it's inherently scalable, allowing for the incorporation of additional species, new geographical regions or projections onto a different time period without altering the modelling structure.

The Iberian Peninsula and the biodiversity of ticks

Our findings confirm a clear swap of tick distribution in the Iberian Peninsula, supported by previous field studies (). A northern group, thriving in colder and wetter climates, is dominated by *I. ricinus*, which often co-occurs locally with *H. punctata*. In contrast, Mediterranean-type habitats host communities of four species – *R. bursa*, *H. marginatum*, *H. lusitanicum* and *D. marginatus* – that co-occur at varying densities. These co-occurring groups are well-documented in field surveys in the Mediterranean basin (Dantas-Torres and Otranto, 2013; Barandika et al., 2014; Grench-Angelini et al., 2016; Tsatsaris et al., 2016).

While some taxa are predicted to occur alone, the framework supports a community structure, particularly in typical Mediterranean-type vegetation associations like sclerophylous, conifer, deciduous and mixed forests, shrubs and woodlands. These areas experience long dry periods but are also subject to severe

precipitation and high winter temperatures (Suc, 1988). However, two species of ticks, already reported in the target territory and detected by our modelling framework, have been removed from the communities, namely *H. sulcata* and *D. reticulatus*. A closer examination of the maps confirmed that both species have been predicted to occur along a patchy distribution, and therefore were removed from communities, which are larger spatial units. In other words, the presence of this two species in isolated cells of our modelling framework favoured its removal in subsequent steps of the modelling (*viz*. the communities assembling). It is ecologically interesting to conciliate modelled results and field findings: the framework correctly detected the missing species, but the statistical filtering did not consider those as belonging to a community because their patchy occurrence patterns. The occurrence of ticks in largely isolate patches and how they interact with the climate and vegetation has not been addressed. We foresee a field for further research.

We linked these resulting communities to standard climate or ecological region classifications. Human-made classifications of either climate or ecological regions can be useful for a generalist definition of co-occurring tick groups, but a one-to-one correspondence does not always exist. The same community composition can occur under different ecological conditions, and *vice versa*. We observed that tick communities form gradients of relative abundance radiating across ecological regions. Our results are highly suggestive of a difference of interpretations of the limits of a an area with a given ecological conditions that are human-made, and therefore, only a qualitative assessment based on large scale features, *versus* an explicit definition of dynamic communities build upon a layer of climate and vegetation features, without human intervention.

Gaps and caveats

This study has some limitations. We demonstrated the existence of communities using standard SDM testing methods where the model's performance was well-established. However, we did not investigate further the model's transferability to Europe, as suggested by Linner et al. (2024). Furthermore, models trained on tick records from the Iberian Peninsula cannot not be directly projected onto other regions even under the same constraints. This is because the models were trained on a portion of their niche rather than the entire range. Developing a model for the entire Palearctic region would require significant computing power. The framework ideally needs to be at 1 km resolution, resulting in hundreds of millions of cells that need to be individually computed for several species. This could be achieved without changing the concept, as the same clustering and categorisation methods would work on larger areas.

Before including *R. sanguineus* s.l. in modelling developments, we need to address the concerns in its identity and distribution in Europe. Our study showed that its inclusion in the current framework, based on its known distribution, does not represent a single species. This led to poorly performing models, attributed to training data representing different species within the genus, resulting in unreliable outcomes. Previous research has demonstrated that *R. turanicus* doesn't exist in the Iberian Peninsula (Dantas-Torres et al., 2017; Bakkes et al., 2020; Mumcuoglu et al., 2022). Furthermore, *Rhipicephalus hibericus* was recently added to the rhipicephalid fauna (Millán et al., 2024) after laboratory crosses between the neotype of *R. sanguineus* and field-collected specimens historically identified as *Rhipicephalus turanicus* in Spain and Portugal. The lack of fertile progeny from these crosses indicated two species were present in the

laboratory colonies. We believe including "*R. turanicus*" in any modelling effort without reassessing its taxonomic identity is futile. Additionally, *R. sanguineus* s.s. is an endophilous species, relatively independent of environmental climate conditions. We are concerned that these restrictions to indoor habitats could negatively impact model performance due to the limited descriptive power of commonly used climate variables.

Conclusions

Independent predictions of disease systems can be ecologically inconsistent when vectors, hosts and pathogens are modelled separately. This often leads to distributions that do not fit the real world. For example, tick populations live within vector communities and different vector species may compete for hosts or occupy different ecological niches. Similarly, hosts can harbour multiple pathogens whose interactions influence transmission. As a result, disease systems often function like ecological networks rather than simple pairs. Future research will likely focus on combining mechanistic epidemiological models, trait-based approaches and high-resolution environmental data into community frameworks. This will be particularly useful for predicting how tick communities respond to climate change. Traditional species distribution models have offered valuable insights into the environmental factors influencing individual species distributions but are limited in their ability to capture the complexity of disease transmission. By incorporating environmental drivers, host communities, vector assemblages and pathogen distributions, community modelling approaches provide a more realistic representation of disease ecology and improve our ability to predict disease emergence and assess future risks under environmental change. Independent predictions can be ecologically inconsistent when vectors, hosts and pathogens are modelled separately, leading to distributions that don't fit reality. Future research will likely focus on integrating mechanistic epidemiological models, trait-based approaches and high-resolution environmental data into community frameworks, particularly for predicting tick community responses to climate change.

Declarations

Author contributions The author of the study designed the research, did the programming and run the scripts to obtain the data, wrote the paper, and prepared the figures.

Data availability No new datasets were generated or analysed during the current study.

Competing interests The author declares no competing interests.

References

1. Alkishe A, Peterson AT (2021) Potential geographic distribution of *Ixodes cookei*, the vector of Powassan virus. J Vector Ecol 46:155–162

2. Barandika JF, Berriatua E, Barral M, Juste RA, Anda P, García-Pérez AL (2006) Risk factors associated with ixodid tick species distributions in the Basque region in Spain. Med Vet Entomol 20:177–188
3. Barandika JF, Olmeda SA, Casado-Nistal MA, Hurtado A, Juste RA, Valcárcel F, García-Pérez AL (2014) Differences in questing tick species distribution between Atlantic and continental climate regions in Spain. J Med Entomol 48:13–19
4. Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Guisan A (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. Glob Ecol Biogeogr 21:481–497
5. Dallas T, Park AW, Drake JM (2019) Predicting cryptic links in host–parasite networks. Ecography
6. Dantas-Torres F, Otranto D (2013) Species diversity and abundance of ticks in three habitats in southern Italy. Ticks Tick Borne Dis 4:251–255
7. Defaye B, Moutailler S, Pasqualini V, Quilichini Y (2022) Distribution of tick-borne pathogens in domestic animals and their ticks in the countries of the Mediterranean basin between 2000 and 2021: a systematic review. Microorganisms 10:1236
8. Dreghiciu IC, Hoffman D, Florea T, Oprescu I, Dumitru S, Imre M, Ilie MS (2025) A systematic review on the occurrence of *Babesia* spp. and *Anaplasma* spp. in ticks and wild boar from Europe: a 15-year retrospective study. Pathogens 14:612
9. Dorazio RM, Royle JA (2005) Estimating size and composition of biological communities by modelling the occurrence of species. J Am Stat Assoc 100:389–398
10. EFSA Panel on Animal Health and Welfare (2010) Scientific opinion on geographic distribution of tick-borne infections and their vectors in Europe and the other regions of the Mediterranean basin. EFSA J 8:1723
11. Eisen RJ, Eisen L (2014) Spatial modeling of human risk of exposure to vector-borne pathogens based on epidemiological versus arthropod vector data. J Med Entomol 45(2):181–192
12. Esser HJ, Mögling R, Cleton NB, Van Der Jeugd H, Sprong H, Stroo A, Reusken CB (2019) Risk factors associated with sustained circulation of six zoonotic arboviruses: a systematic review for selection of surveillance sites in non-endemic areas. Parasit Vectors 12:265
13. Estrada-Peña A, Gray JS, Kahl O, Lane RS, Nijhof AM (2013) Research on the ecology of ticks and tick-borne pathogens: methodological principles and caveats. Front Cell Infect Microbiol 3:29
14. Estrada-Peña A, Farkas R, Jaenson TGT, Koenen F, Madder M, Pascucci I, Jongejan F (2013) Association of environmental traits with the geographic ranges of ticks (Acari: Ixodidae) of medical and veterinary importance in the western Palearctic: a digital data set. Exp Appl Acarol 59:351–366
15. Estrada-Peña A, Estrada-Sánchez A, de la Fuente J (2014) A global set of Fourier-transformed remotely sensed covariates for the description of abiotic niche in epidemiological studies of tick vector species. Parasit Vectors 7:302
16. Estrada-Peña A, Sprong H, Cabezas-Cruz A, de la Fuente J, Ramo A, Coipan EC (2016) Nested coevolutionary networks shape the ecological relationships of ticks, hosts, and the Lyme disease bacteria of the *Borrelia burgdorferi* (sl) complex. Parasit Vectors 9:517

17. Estrada-Peña A, Martínez-Estéllez MAH, Pradera C, Castellà J (2025) The historical baseline of hard tick records in Spain (1985–2024). Pathogens 14:173
18. Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. J Appl Ecol 43:393–404
19. GBIF (2025) Download DOI10.15468/dl.qspbnr
20. Golding N, Nunn MA, Purse BV (2015) Identifying biotic interactions which drive the spatial distribution of a mosquito community. Parasit Vectors 8:367
21. Grech-Angelini S, Lancelot R, Ferraris O, Peyrefitte CN, Vachiery N, Pédarrieu A, Peyraud A, Rodrigues V, Bastron D, Libeau G, Fernandez B, Holzmuller P, Servan de Almeida R, Michaud V, Tordo N, Comtet L, Métras R, Casabianca F, Vial L (2020) Crimean-Congo hemorrhagic fever virus antibodies among livestock on Corsica, France, 2014–2016. Emerg Infect Dis 26:1041–1044
22. Guglielmone AA, Robbins RG, Apanaskevich DA, Petney TN, Estrada-Peña A, Horak IG (2014) The hard ticks of the world. Springer, Dordrecht
23. Hijmans R (2025) terra: spatial data analysis. R package version 1.8–55
24. Keesing F, Holt RD, Ostfeld RS (2006) Effects of species diversity on disease risk. Ecol Lett 9:485–498
25. Levi T, Keesing F, Holt RD, Barfield M, Ostfeld RS (2016) Quantifying dilution and amplification in a community of hosts for tick-borne pathogens. Ecol Appl 26:484–498
26. Maxwell SP, Brooks C, Kim D, McNeely CL, Cho S, Thomas KC (2023) Improving surveillance of human tick-borne disease risks: spatial analysis using multimodal databases. JMIR Public Health Surveill 9:e43790
27. Millán J, Rodríguez-Pastor R, Estrada-Peña A (2024) Description of *Rhipicephalus hibericus* sp. nov. Ticks Tick Borne Dis 15:102340
28. Moerbeck L, Parreira R, Seixas G, Velez R, Domingos A, Antunes S (2024) Update on tick-borne rickettsioses in mainland Portugal: emerging threats and potential vectors. Parasit Vectors 17:538
29. Mücher CA, Klijn JA, Wascher DM, Schaminée JHJ (2010) A new European landscape classification (LANMAP): a transparent, flexible and user-oriented methodology to distinguish landscapes. Ecol Indic 10:87–103
30. Nava S, Beati L, Venzal JM, Labruna MB, Szabó MPJ, Petney T, Estrada-Peña A (2018) *Rhipicephalus sanguineus*: neotype designation, morphological re-description of all parasitic stages and molecular characterization. Ticks Tick Borne Dis 9:1573–1585
31. Norberg A, Abrego N, Blanchet FG, Adler FR, Anderson BJ, Anttila J, Ovaskainen O (2019) A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecol Monogr 89:e01370
32. Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on earth. Bioscience 51:933–938

33. Ostfeld RS, Keesing F (2012) Effects of host diversity on infectious disease. Annu Rev Ecol Evol Syst 43:157–182
34. Ostfeld RS, Levi T, Jolles AE, Martin LB, Hosseini PR, Keesing F (2014) Life history and demographic drivers of reservoir competence for three tick-borne zoonotic pathogens. PLoS ONE 9:e107387
35. Ovaskainen O, Tikhonov G, Norberg A, Blanchet FG, Duan L, Dunson D, Abrego N (2017) How to make more out of community data? a conceptual framework and its implementation as models and software. Ecol Lett 20:561–576
36. Pfäffle MP, Santos-Silva MM, Jaenson TGT, Vatansever Z, Petney TN (2018) Haemaphysalis punctata. Ticks of Europe and North Africa: a guide to species identification. Springer
37. Poulin R (2007) Are there general laws in parasite ecology? Parasitology 134:763–776
38. Randolph SE (2009) Tick-borne disease systems emerge from the shadows: the beauty lies in molecular detail, the message in epidemiology. Parasitology 136:1403–1413
39. QGIS Development Team (2024) QGIS geographic information system. Open Source Geospatial Foundation
40. Reisen WK (2010) Landscape epidemiology of vector-borne diseases. Ann Rev Entomol 55(1):461–483
41. Rongcheng R, Yi H, Jialing M, Ying Y, Fanning L, Xiya W, Jianglong Q (2025) The species distribution model based on the random forest algorithm reveals the distribution patterns of *Neophocaena asiaeorientalis*. Sci Rep 15:10037
42. Sánchez-Mora P, Habela MA, del Peso T, Grande Ávila AC, García López AM, Mata García Soldado J, Tapia MM, Molero F, Herrero L, Olmeda AS, Valcárcel F, Estrada-Peña A, Negredo A, Sánchez-Seco MP (2026) Detection and genetic characterization of Crimean-Congo hemorrhagic fever virus in ticks from western Spain (2017, 2020–2024). Front Vet Sci 13:1789622
43. Santos-Silva MM, Beati L, Santos AS, De Sousa R, Núncio MS, Melo P, Bacellar F (2011) The hard-tick fauna of mainland Portugal (Acari: Ixodidae): an update on geographical distribution and known associations with hosts and pathogens. Exp Appl Acarol 55:85–121
44. Scharlemann JP, Benz D, Hay SI, Purse BV, Tatem AJ, Wint GW, Rogers DJ (2008) Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. PLoS ONE 3:e1408
45. Sonenshine DE, Roe RM (2014) Biology of ticks, vol 2. Oxford University Press
46. Suc JP (1984) Origin and evolution of the Mediterranean vegetation and climate in Europe. Nature 307:429–432
47. Tikhonov G, Opedal ØH, Abrego N, Lehikoinen A, De Jonge MM, Oksanen J, Ovaskainen O (2020) Joint species distribution modelling with the R-package Hmsc. Methods Ecol Evol 11:442–447
48. Tsatsaris A, Chochlakis D, Papadopoulos B, Petsa A, Georgalis L, Angelakis E, Ioannou I, Tselentis Y, Psaroulaki A (2016) Species composition, distribution, ecological preference and host association of ticks in Cyprus. Exp Appl Acarol 70:523–542
49. Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2021) Modelling species presence-only data with random forests. Ecography 44:1731–1742

50. Warton DI, Blanchet FG, O'Hara RB, Ovaskainen O, Taskinen S, Walker SC, Hui FKC (2015) So many variables: joint modelling in community ecology. Trends Ecol Evol 30:766–779
51. Westcott JR, Bowden JJ, Savage J, Doody KM (2025) Rapid northward expansion of the blacklegged tick, *Ixodes scapularis*, in response to climate change. Glob Change Biol 31:e70591
52. Yang MS, Lai CY, Lin CY (2012) A robust EM clustering algorithm for Gaussian mixture models. Pattern Recognit 45:3950–3961
53. Zhang L, Huettmann F, Liu S, Sun P, Yu Z, Zhang X, Mi C (2019) Classification and regression with random forests as a standard method for presence-only data SDMs. Ecol Inf 52:46–56

Figures

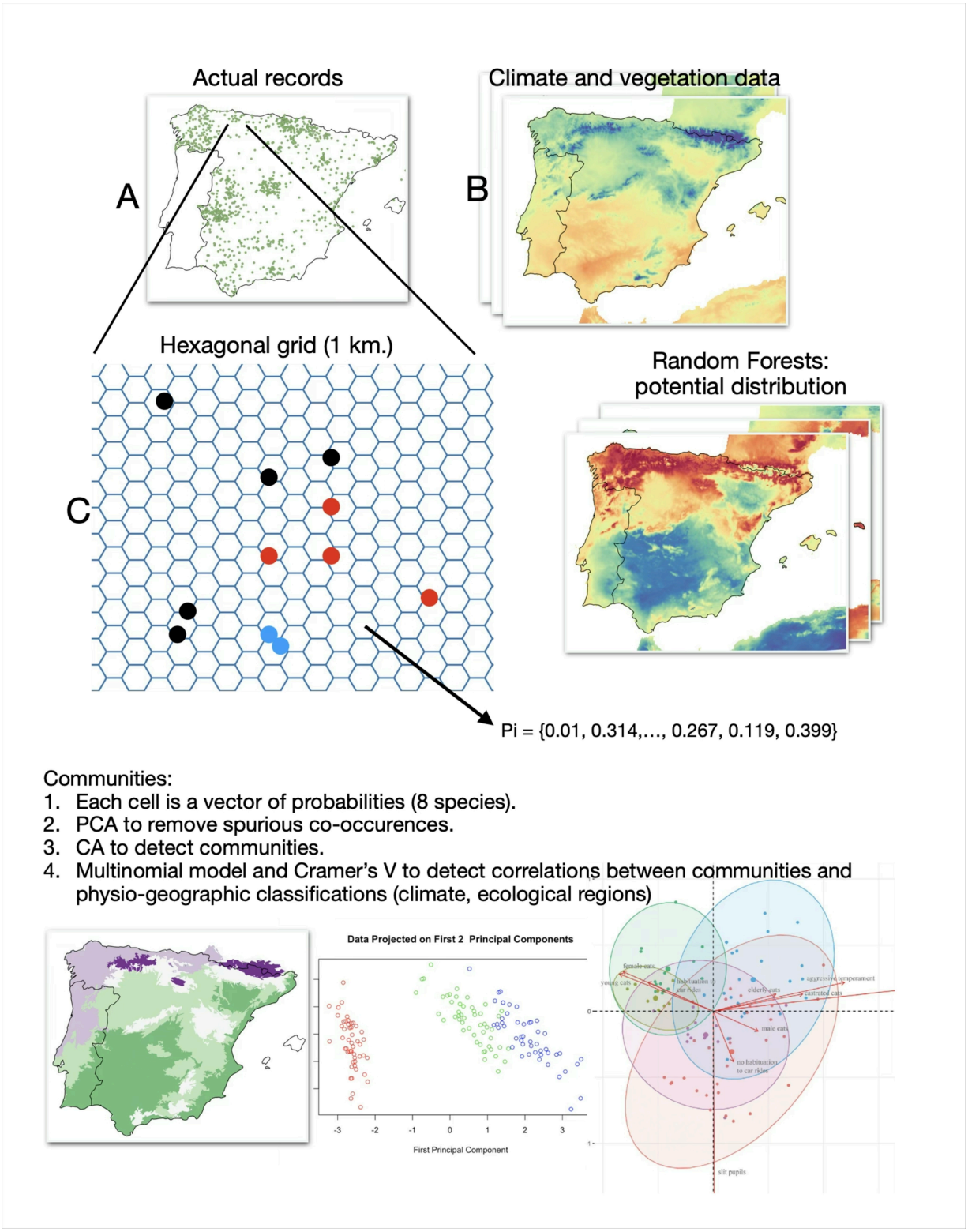

Figure 1

The flow of work followed in this study. After collecting and cleaning the records of ticks in the Iberian Peninsula (A), we integrated climate and vegetation data (B) in a hexagonal grid of 1 km. resolution, on which models were carried out separately for each species (C). Ecah hexagon carries the vector of probabilities of the occurrence of each species (C), that is examined by Principal Components Analysis to

remove spurious values. Species are clustered in communities using a Correspondence Analysis, together with other methods to correlate the spatial range of communities with ecological or climate regions

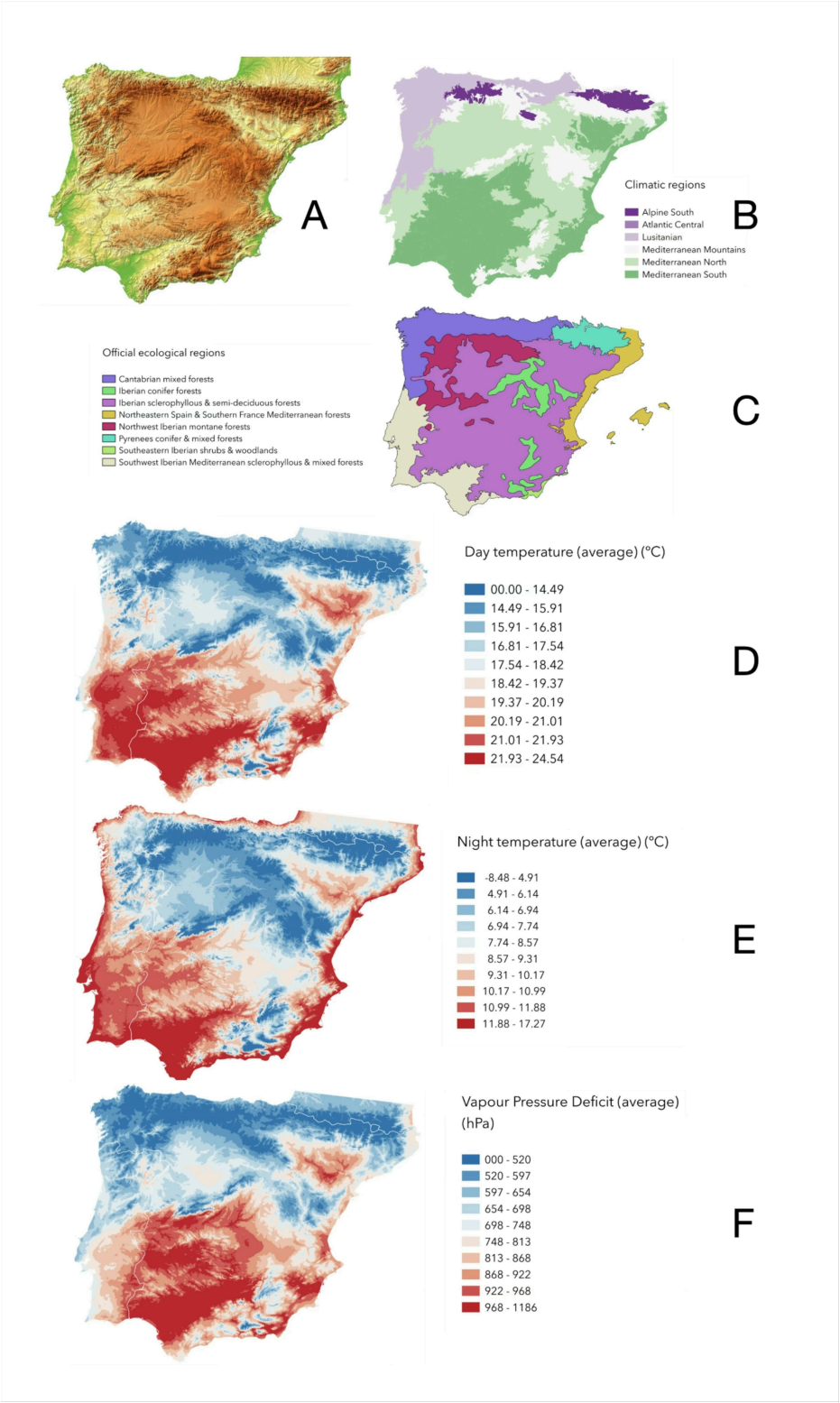

Figure 2

An overview of the physical, ecological features and climate values of the Iberian Peninsula. A: the elevation of the territory (maximum, 3.500 meters a.s.l.). B: the standard categories of climate.C: the

spatial distribution of the official ecological regions. D: annual averaged maximum temperature (ºC). E: annual averaged minimum temperature (ºC). F: annual averaged vapour pressure deficit (in hPa).

Ecoregions were abbreviated as follows: CMF = Cantabrian mixed forests; ICF = Iberian conifer forests; ISSF = Iberian sclerophyllous and semi-deciduous forests; NSF-Med = Northeastern Spain and Southern France Mediterranean forests; NIMF = Northwest Iberian montane forests; PCM = Pyrenees conifer and mixed forests; SIW = Southeastern Iberian shrubs and woodlands; SWI-Med = Southwest Iberian Mediterranean sclerophyllous and mixed forests.

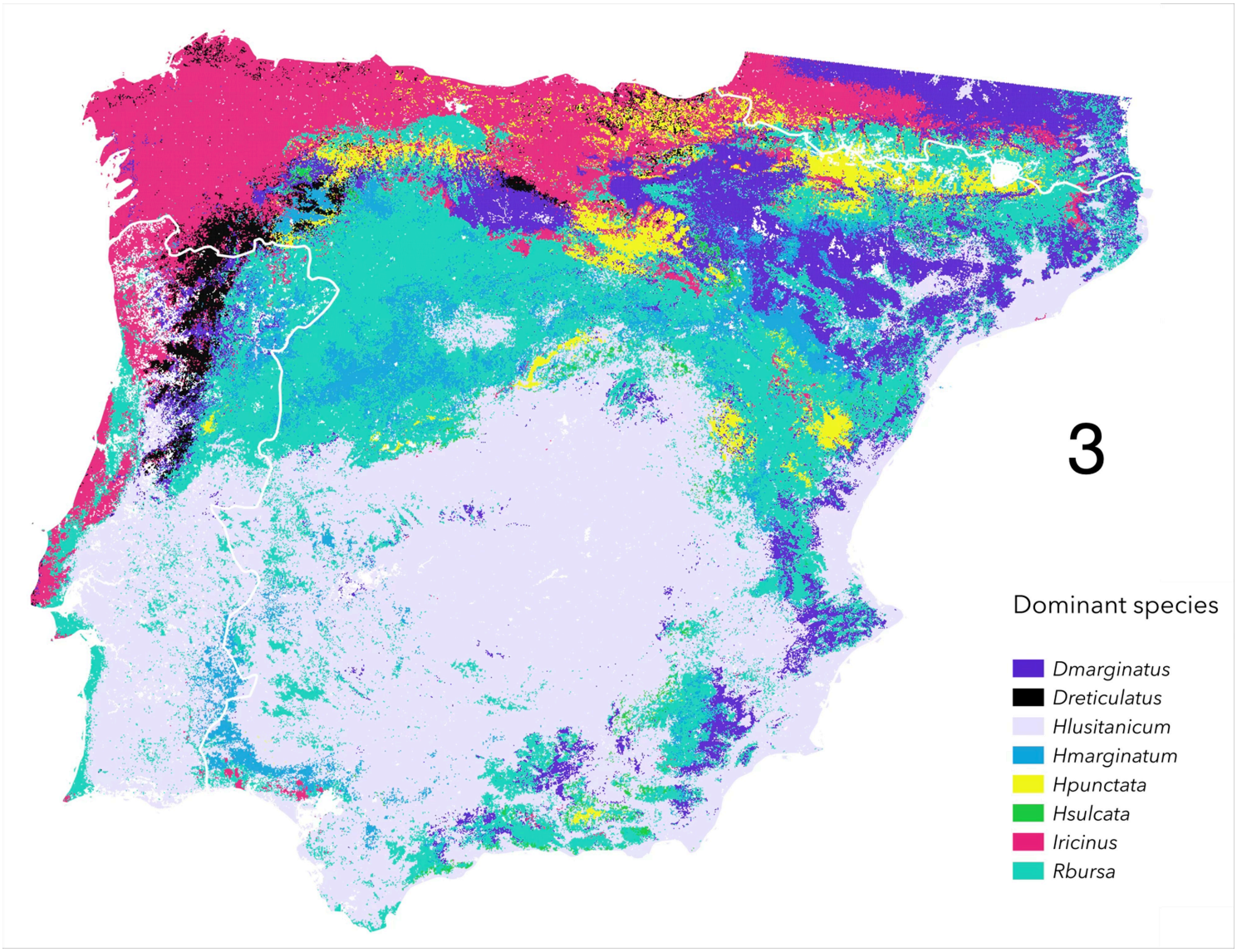

Figure 3

The spatial distribution of the dominant specie sof ticks, at 1 km. resolution, in the Iberian Peninsula.

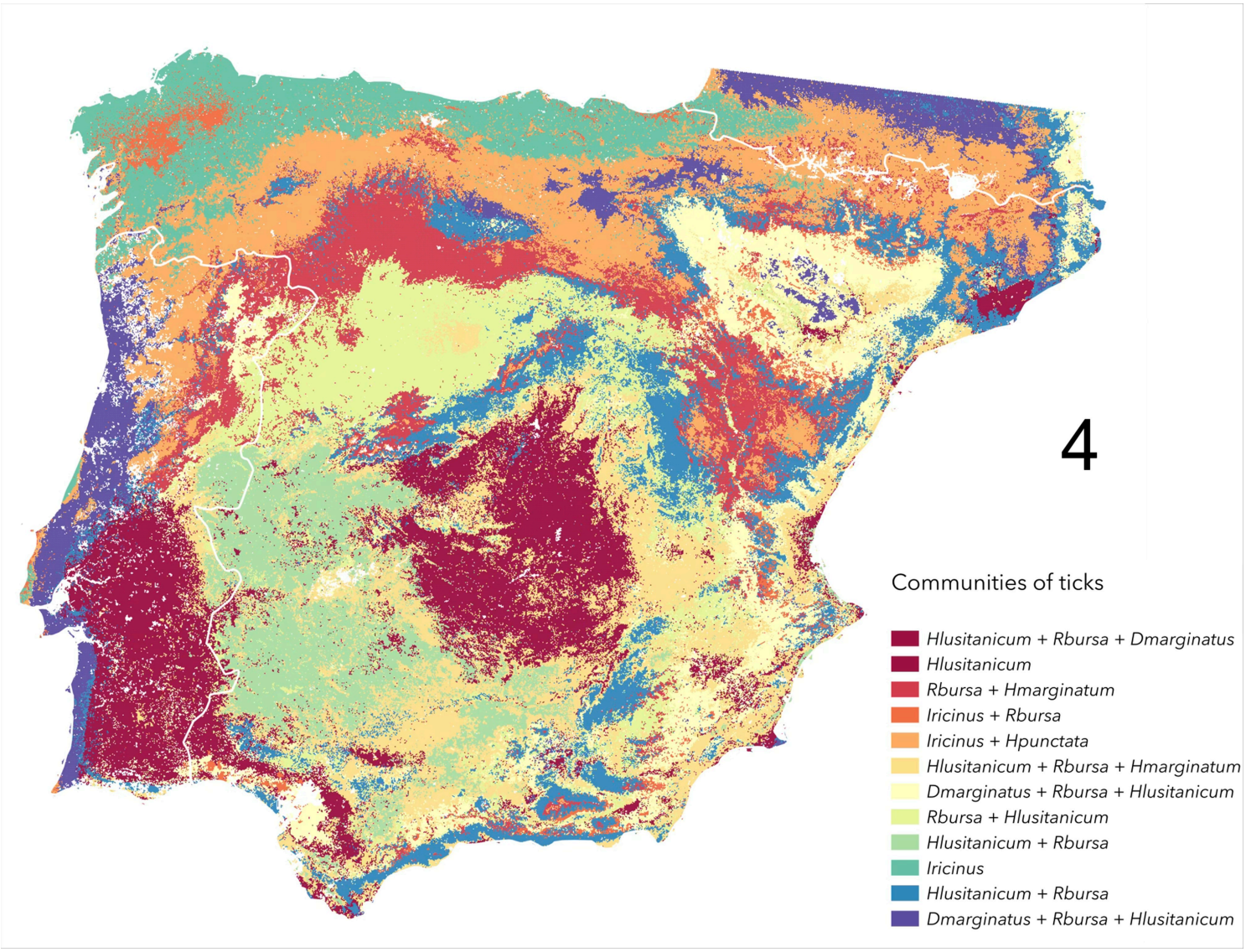

Figure 4

The spatial distribution of the detected 12 communities of ticks, before compacting the similar ones.

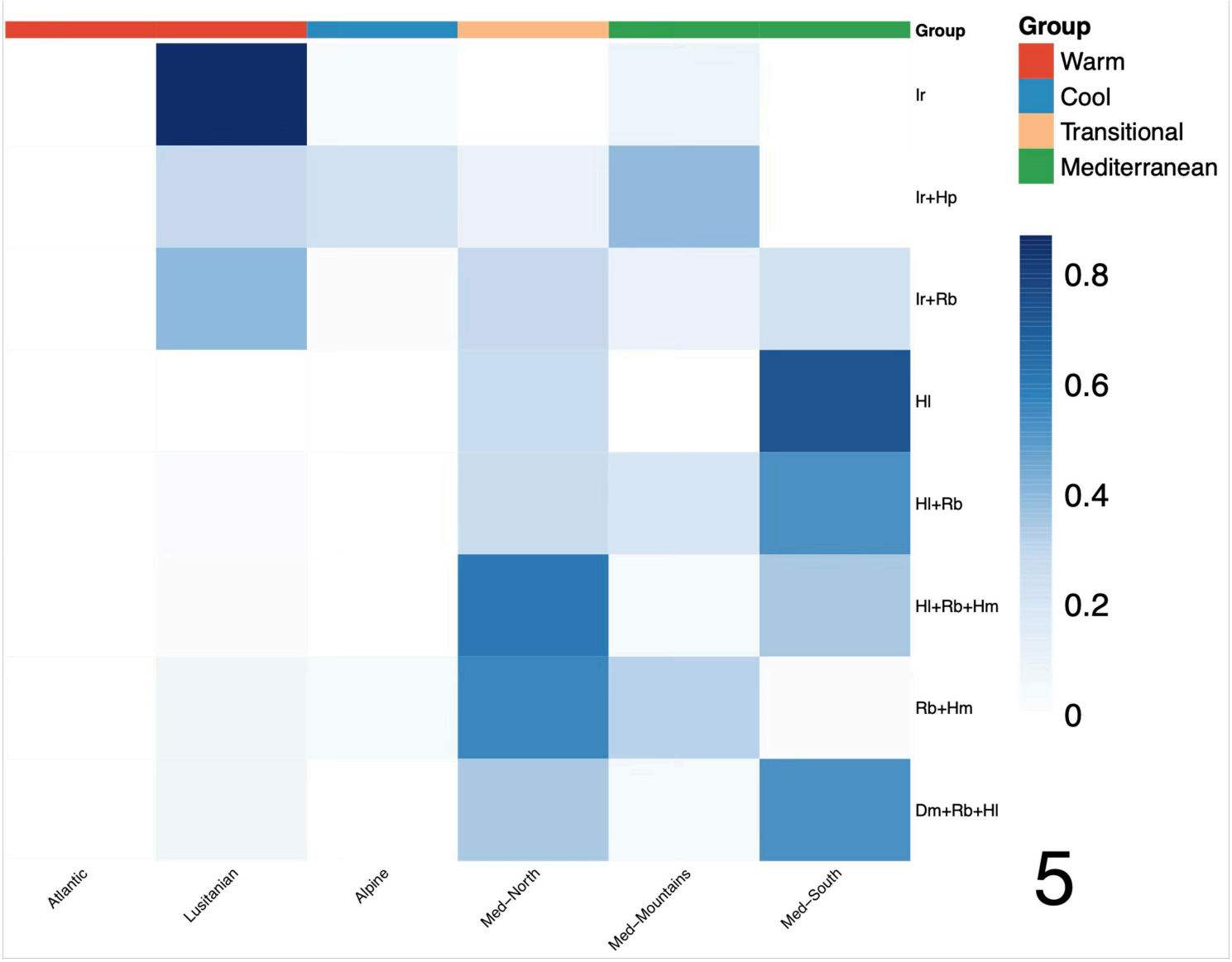

Figure 5

A heatmap displaying the relationships between the communities of ticks (rows) and the standard classifications of climate (columns) in Iberian Peninsula. The scale is the relationship value of a Correspondence Analysis based on the spatial distribution of the communities and the climatic regions. Communities are named along the first letter of the genus and species of each taxon in the community. Ir: *Ixodes ricinus*, Hp: *Haemaphysalis punctata*, Rb: *Rhipicephalus bursa*, Hl: *Hyalomma lusitanicum*, Hm: *Hyalomma marginatum*, Dm: *Dermacentor marginatus*.

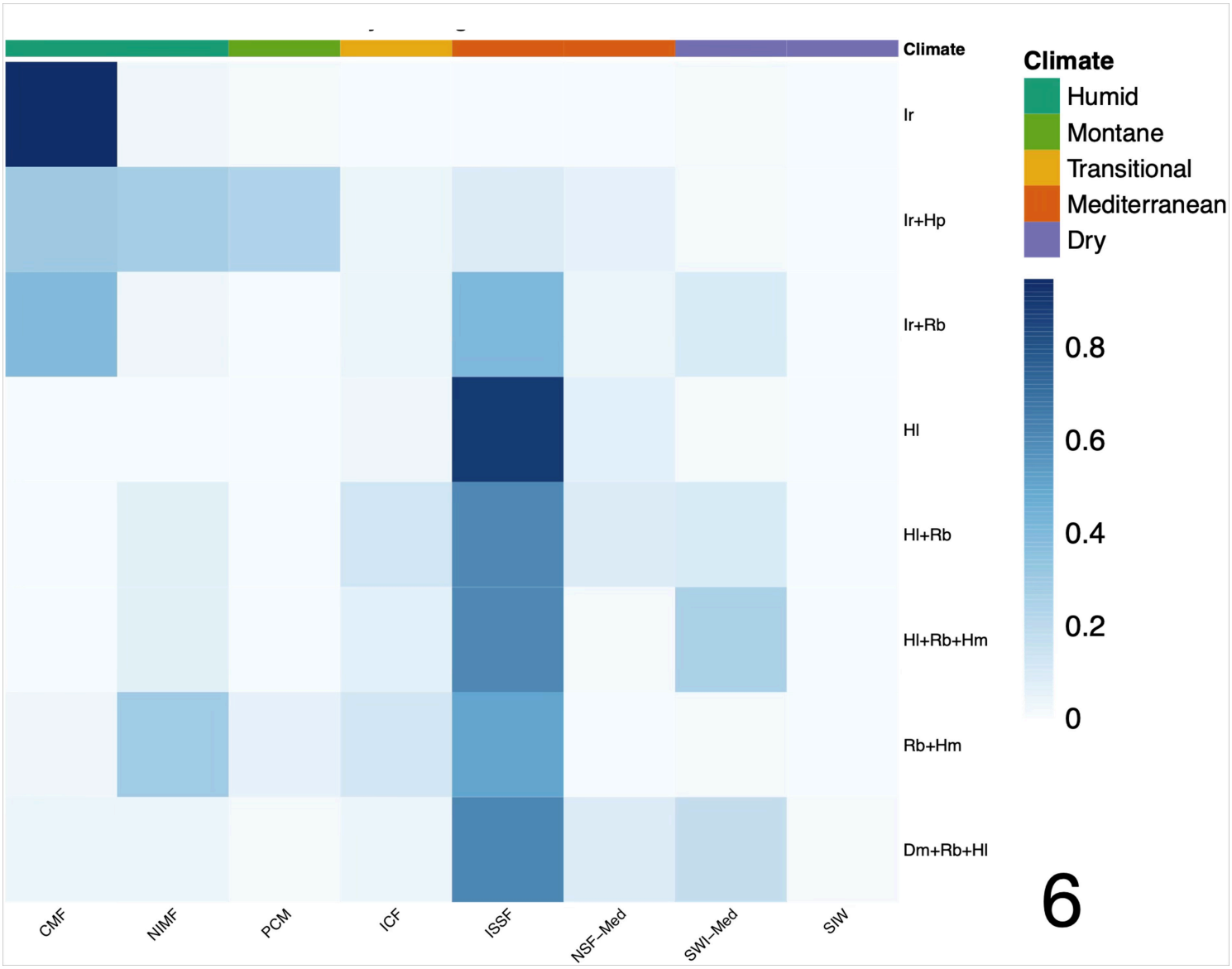

Figure 6

A heatmap displaying the relationships between the communities of ticks (rows) and the standard ecological regions (columns) in Iberian Peninsula. The scale is the relationship value of a Correspondence Analysis based on the spatial distribution of the communities and the climatic regions. Communities are named along the first letter of the genus and species of each taxon in the community. Ir: *Ixodes ricinus*, Hp: *Haemaphysalis punctata*, Rb: *Rhipicephalus bursa*, Hl: *Hyalomma lusitanicum*, Hm: *Hyalomma marginatum*, Dm: *Dermacentor marginatus*. Ecological regions are abbreviated as CMF: Cantabrian Mixed Forests, ICF: Iberian conifer forests, ISSF: Iberian sclerophyllous and semi-deciduous forests, NSF-Med: Northeastern Spain and Southern France Mediterranean forests, NIMF: Northwest Iberian montane forests, PCM: Pyrenees conifer and mixed forests, SIW: Southeastern Iberian shrubs and woodlands, SWI-Med: Southwest Iberian Mediterranean sclerophyllous and mixed forests