

Spatial patterns and eco-epidemiological systems–part II: characterising spatial patterns of the occurrence of the insect vectors of Chagas disease based on remote sensing and field data.

Emmanuel Roux, Annamaria de Fátima Venâncio, Jean-François Girres, Christine A Romaña

▶ **To cite this version:**

Emmanuel Roux, Annamaria de Fátima Venâncio, Jean-François Girres, Christine A Romaña. Spatial patterns and eco-epidemiological systems–part II: characterising spatial patterns of the occurrence of the insect vectors of Chagas disease based on remote sensing and field data.. Geospatial Health, 2011, 6 (1), pp.53-64. 10.4081/gh.2011.157 . ird-01369930

HAL Id: ird-01369930
https://ird.hal.science/ird-01369930

Submitted on 21 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Geospatial Health 6(1), 2011, pp. 53-64

Spatial patterns and eco-epidemiological systems – part II: characterising spatial patterns of the occurrence of the insect vectors of Chagas disease based on remote sensing and field data

Emmanuel Roux[1], Annamaria de Fátima Venâncio[2], Jean-François Girres[3], Christine A. Romaña[4]

[1]ESPACE-DEV, UMR228 IRD/UMII/UR/UAG, Institut de Recherche pour le Développement, Cayenne, French Guiana; [2]Instituto do Meio Ambiente e Recursos Hídricos do Estado da Bahia e Centro de Desenvolvimento Sustentável da Universidade de Brasília, Brazil; [3]IGN, COGIT, Saint Mandé, France/IRD, ESPACE-DEV (UMR228), Cayenne, French Guiana; [4]Université Paris Descartes, Paris, France/IRD, ESPACE-DEV (UMR228), France

Abstract. While the former part of this back-to-back paper dealt with the identification of multi-scale spatial patterns associated with the presence, abundance and dispersion of the insect vectors (Triatominae) of Chagas disease, this latter part examines the need for pattern characterisation by means of detailed data on environmental, residential, peri-domiciliary and human behaviour. The study site was, in both cases, a single village situated in Bahia, Brazil, wherefrom the data were collected through field observation and a standardised questionnaire, while the environmental characteristics were derived from satellite images and landscape characterisation. Following this, factorial analysis of mixed group (FAMG), an exploratory data analysis method, was applied to "mine" the huge dataset in a hierarchical way and to evaluate the relative impact of different factors such as the surrounding environment, the domiciliary/peri-domiciliary space properties and the presence of domestic animals. In the study village, five principal "districts" associated with different possible causes of infestation were identified. The results favour the role of depressions of the ground surface due to collapse of karstic subsoil (dolines) and open rock faces as infestation sources, vector attraction by outdoor lighting, risk of insect domiciliation in dwellings constructed without finishing materials and associated with apparent disorder. Ultimately, this study not only provides the basic information needed for decision-making and specification of vector control in the study village, but offers also a knowledge-base for more general control strategies in the region.

Keywords: principal coordinate of neighbourhood matrices, factorial analysis of mixed groups, remote sensing, landscape characterisation, Chagas disease insect vectors, Triatomines, Brazil.

Introduction

In part I of this back-to-back paper, we suggested that modelling of the spatial and/or temporal properties of the entomological and/or epidemiological data should be carried out before relating them with explanatory variables. We showed that spatial modelling of Chagas disease insect vectors provides substantial information on presence, abundance and dispersal behaviour of the triatomine vectors at different scales. We also mentioned the need for a deeper insight into which parameter(s), from land cover/use, intra- and peri-domiciliary spaces and inhabitant behaviour characteristics, could be related to these spatial patterns.

The association of the spatial distribution of Chagas disease insect vectors with environmental variables has been shown previously (Costa et al., 2002; Peterson et al., 2002. Dumonteil and Gourbière, 2004; Walter et al., 2005; Batista and Gurgel-Gonçalves, 2009). From the methodological point of view, Vazquez-Prokopec (2009) and Khan (2010) went one step further by applying geo-statistic analysis of entomological data for the identification of hotspots of domestic *Triatoma infestans* (Klug, 1834) infestations. Thus, multivariate regression analysis based on environmental and demographic factors (elevation, percentage of land covered by secondary forest or deforested areas, land surface temperature and rural houses density) could help

Corresponding author:
Emmanuel Roux
ESPACE-DEV, UMR228 IRD/UMII/UR/UAG
Institut de Recherche pour le Développement
Cayenne, French Guiana
Tel. +594 (0)5 94 29 92 77 Fax +594 (0)5 94 31 98 55
E-mail : emmanuel.roux@ird.fr

deciding whether or not a community would be at risk.

In eco-epidemiological system studies, data mapping is usually used to establish direct links between entomological and/or epidemiological data with environmental explanatory variables. For example, Dumonteil and Gourbière (2004) modelled the relationships between the domestic abundance of *Triatoma dimidiata* (Latreille, 1811), the vector transmitting *Trypansoma cruzi*, and the bioclimatic factors by means of an analysis of covariance (ANCOVA) logistic regression. They produced a natural transmission risk map that fitted well with the locations of seropositive cases including people with chronic Chagas disease. Likewise, Costa et al. (2002), Peterson et al. (2002), and Batista and Gurgel-Gonçalves (2009) characterised and discriminated ecological niches suitable for *Triatoma brasiliensis* (Neiva, 1911) and *Rhodnius* species in north-eastern Brazil based on the genetic algorithm for rule-set prediction (GARP). These examples focus on environmental conditions defining ecological niches, obtained with different ways of modelling and data analysis, but all based on the construction and combination of two tables: one related to species presence or abundance, and the other describing the environmental conditions (Dray et al., 2003; Tsoar et al., 2007). However, if the goals are to map the transmission risk and support its control by the health autorities, many dimensions have to be taken into account, in particular societal factors related to population vulnerability and risk exposure. Walter et al. (2005) studied the abundance of *T. brasiliensis* and *T. pseudomaculata* (Correa and Espinola, 1964) in relation to environmental, demographical and sociological data using a variable selection procedure and a multivariate regression analysis to show that *T. pseudomaculata* is associated with environmental characteristics, while *T. brasiliensis* depend on peri-domiciliary structure and human activities.

To our knowledge, so far no study has proposed the exploration of the numerous factors involved in eco-epidemiological system functioning in a hierarchical way. Beyond the identification of the most explanatory variables in a non-structured set of environmental and societal variables, it would be of interest to identify the relative impact of *a priori* factors, i.e. complex elements particularly associated with environmental, demographical and behavioural variables. However, this demands the *a priori* equilibration of the relative weights of these factors in the analysis. Here, we apply the factorial analysis of mixed groups (FAMG) (Pagès, 2002, 2004), an approach that has contributed to the explanation of the presence of Chagas disease insect vectors *T. sordida* (Stål, 1859) and *P. geniculatus* (Latreille, 1811) regarding the surrounding environment, human behaviour, dwellings and peri-domiciliary features (Roux et al., 2009). However, in the present study, we did not consider raw capture data but the components of the spatial models for the distribution of adults and juveniles obtained with the methodology described in part I of this paper. Consequently, the focus here is on the characteristics of insect presence and abundance that present strong spatial autocorrelations at different scales.

Materials and methods

Study area

As in part I, this study concerns Santa Rita, the rural locality of Iraquara municipality in Bahia, Brazil. The Santa Rita geological relief is slightly undulating with depressions of the ground surface due to collapse of the karstic subsoil (dolines). Twenty-four such dolines are situated near the village, and some of them present rock faces and caves that have become the habitat for many animal species such as birds, rodents, marsupials, monkeys, reptiles and chiropterans. These places also provide resting sites (crevices in rock faces) and trophic resources for triatomine species.

The land cover consists of a mosaic of caatinga (the dry shrub land that constitutes the main vegetation type in this area), cerrado (tropical Brazilian savanna characterised by strong plant and animal biodiversity) and semi-deciduous seasonal forest. These floral systems are associated with different geological and geomorphologic units, while human settlement has contributed to a high degree of landscape fragmentation. Agriculture based on the cultivation of beans and corn is the traditional and principal mean of subsistence. Intensive agriculture is developing but is still limited to beans, corn and sugarcane. Irrigation and pesticides are essentially only used for tomatoes, coffee and pepper.

Data collection

Entomological data

The domiciliary unit (DU), i.e. the human intra and peri-domiciliary spaces with annexes as defined by Silveira and Rezende (1994), constituted the sampling unit used. For each of the 132 DUs of the Santa Rita

village, insect eggs, juveniles, adults and exuvii were collected from October to December 2007. Details concerning collection methodology and results are presented in part I of this paper.

Application of principal coordinate analysis of neighbourhood matrices (PCNM) provided two spatial models of these entomological data (see part I): one for juvenile abundance and the other for the presence of adults. The juvenile and adult models were composed of 28 and 32 orthonormal eigenvectors representing 82.3% and 79.9% of the total variance of the data, respectively. In the following, these two sets of eigenvectors, denoted $\{JVi\}_{i\,[1,28]}$ and $\{AVi\}_{i\,[1,32]}$, respectively, define the two sets of quantitative variables used in the data analysis.

Household data

Data related to humans were obtained by means of field observations and an extensive household survey. This work was done from October to December 2007. Table 2 lists the variables and the survey items. The DUs were geo-localised by means of a GARMIN eTrex legend global positioning system (GPS) instrument.

Remote sensing and landscape characterisation

Two concomitant SPOT-5 images (http://www.seas-guyane.org) acquired on 15 January 2007: one panchromatic image with 2.5 m spatial resolution and one 4-band colour image (red, green and near infra-red at 10 m spatial resolution + middle infra-red at 20 m resolution) provided the geographical base for the work. These images were produced with radiometric and geometric corrections in the UTM WGS84 map projection without ground control points (the 2A production level), then registered (bilinear model) by using GPS-based DU geo-localisations. A pan-sharpened 2.5 m resolution image, obtained by combining the colour and the panchromatic images, was exploited for a photo-interpretation of the scene and a manual delimitation of the dolines and rock faces surrounding the village (Fig. 1). Segmentation, based on the free remote sensing image processing software Spring (Camara et al., 1996), of this image, followed by a supervised classification, was performed in order to provide a land cover map. Eight classes of land cover were identified: dense forest, low-density forest, bare soil and five types of exploited soils (Fig. 1). The 4th type of exploited soil corresponds essentially to

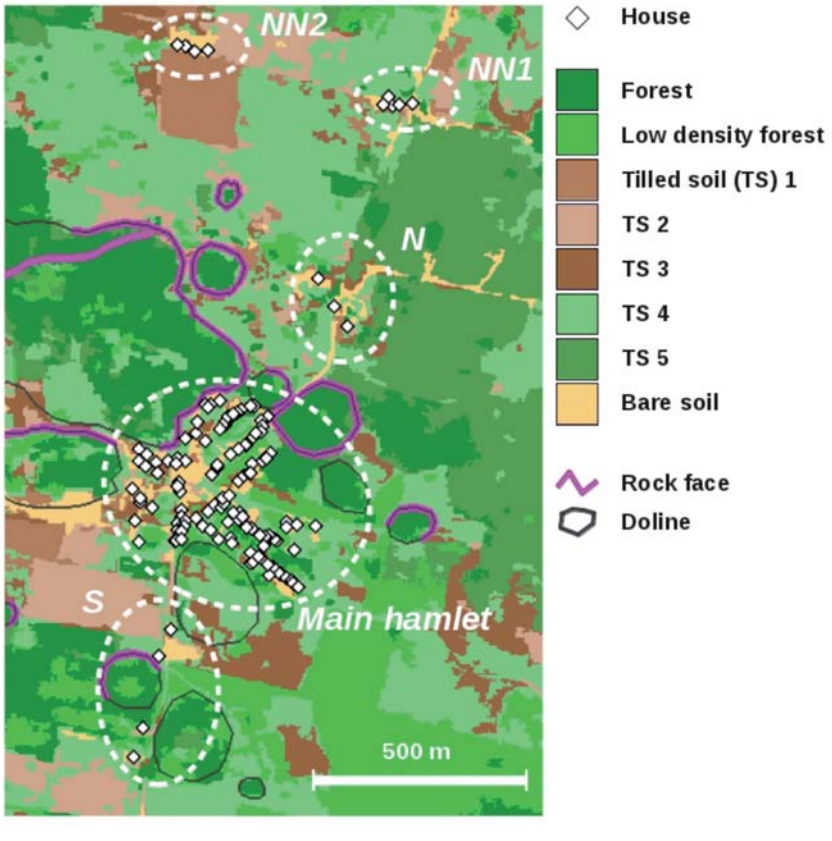


Fig. 1. Results of the supervised land cover classification. Dolines and rock faces were manually digitalized. Dotted white lines delimit dwelling groups.

Table 1. Variables defined for each DU from image photo-interpretation and land cover characterisation.

Variable	Name	Source	Interpretation
Distance between DU and nearest doline	Dist_doline	Manual digitalisation of dolines	Interaction degree between DU and dolines
Length of doline perimeter within 150 m radius from each DU	L_doline		
Distance between DU and nearest rock face	Dist_rock_face	Manual digitalisation of rock faces	Interaction degree between DU and rock faces
Length of rock face within a 150 m radius	L_rock_face		
Proportion of each land cover class within 150 m radius	Modality names = class names	Supervised characterisation of land cover	Landscape composition
Jaeger landscape division (Jaeger, 2000) within a 150 m radius	Landscape_div	Supervised characterisation of land cover	Landscape structure/ division

pastures and fallows with sparse trees (notably palm trees) and shrubs. The other types of exploited soils correspond to various types of tilled soils.

Even if the landscape structures (spatial distribution of patches, connectivity, etc.) are rarely considered by spatial epidemiologists, they have, according to Ostfeld et al. (2005), "a strong potential to influence disease dynamics through impacts on both abiotic conditions and species interactions that are important to disease spread and prevalence". Based on this reasoning, we defined two kinds of variables: (i) distance-based variables and (ii) variables associated with landscape composition and structure within a given radius. The radius was set to 150 m since it has been demonstrated that *T. sordida* is capable of flight exceeding 100 m (Schofield et al., 1991). Moreover, this radius allowed an effective discrimination of the DUs and a consistent characterisation of the surrounding landscape. The landscape description variables are presented in Table 1. They also appear in Table 2 gathered in the variable group "Environment". All feature extractions were performed with GRASS GIS, a freely available geographical information system (GIS) (GRASS Development Team, 2010).

Table 2. Variables, questions and groups of variables/questions.

Group (GROUP NAME)	Variable/questions
Spatial model for juveniles (SPAT STRUC JUV)	Eigenvectors defining the spatial model for juvenile data
Spatial model for adults (SPAT STRUC ADULTS)	Eigenvectors defining the spatial model for adult data
Land cover/use, landscape characteristics (ENVIRONMENT)	Percentage of each land cover/use type within a 150 m radius around each DU, distances to the nearest doline and rock faces, lengths of doline perimeter and of rock faces, landscape division according to Jaeger (2000).
Dwelling (DWELLING)	Wall building materials, plastering inside/outside of the house, plastering missing, roof building materials, floor materials, presence of cracks, use of insecticide, presence of animals, disorder, or litter in the house, overall hygiene
Peri-domiciliary (PERI-DOM)	Presence of lighting at the front/rear of the house, presence of stacks of materials, use of insecticides, peri-domiciliary parcelling out, peri-domiciliary structure (opened, closed, semi-closed), enclosure building materials, tree species (mango tree, jackfruit, avocado, Syagrus coronata-licuri palm)
Domestic animals and animal breeding (ANIMALS)	Presence of hens, dogs, cats, pigs, birds, other animals, presence of henhouse, animal(s) sleeping in the house and/or in the garden

Data analysis

The groups of variables investigated were defined according to the objects, or the situations, they characterise and consequently define *a priori* factors (see Table 2). We did not include raw capture data but the components of the spatial models for adults and juveniles, obtained with the methodology described in part I. The FAMG (Pagès, 2002, 2004) was used based on: (i) a weighted principal component analysis (PCA) applied to each group, within which the inertia of the quantitative and qualitative variables were balanced, and (ii) a weighted PCA applied to the whole table with the inverse of the maximum eigenvalue of each separated PCA (for each group) as weight to ensure the balance of the group contributions in the global analysis. With these weights, the maximum inertia of each group could be set to 1, i.e. each group was considered as equivalent with no *a priori* knowledge of the relative importance of the groups and no weighting bias due to differences of variable numbers in the groups. This method was implemented in the R environment for statistical computing (R Development Core Team, 2010) with modified FactomineR library functions (Lê et al., 2008). Here, the terms "variables" and "modalities" refer to quantitative variables and modalities of the categorical variables, respectively.

In the results section, we show only: (i) the variables and modalities that contributed the most to the factorial axis and (ii) the better represented variables and modalities as it facilitates interpretation and guarantees statistical significance of the results. These two criteria are expressed in mathematical terms as follows:

(i) for a given axis, after ranking variables and modalities in descending order of their contributions to the axis, variable and modality cumulative contributions must be at most equal to 75%, ensuring that the main variables/modalities that explain the axis are considered; and

(ii) for the j^{th} variable or modality and the p^{th} factorial axis, $cos^2_{j,p} \geq 0.1$, indicating that the p^{th} axis explains at least 10% of the total variance of the j^{th} variable or modality and ensuring that the variable or modality is well represented on the factorial plane.

The contribution of the j^{th} variable (or modality) to the p^{th} axis, $Crt_{j,p}$, is the ratio of the linear correlation of the variable (or of the correlation ratio of the modality) with the factorial axis p, over the sum of the correlations and correlation ratios of all the variables and modalities with the same axis. Thus, we have:

$$\sum_{j=1}^{J} Ctr_{j,p} = 100 ,$$

where J represents the total number of variables and modalities. $cos^2_{j,p}$ in the square of the cosine of the angle between the p^{th} axis, and the segment linking the origin of the plane and the j^{th} variable/modality location. Thus, it represents the quality of the representation onto the factorial plane. We then have:

$$\sum_{p=1}^{P} \cos^2_{j,p} = 1 ,$$

where P represents the number of factorial axes.

Results

Analysis of the whole dataset

Fig. 2 shows the quantitative variables projected on the first FAMG factorial plane as well as the spatial representations of the spatial model components, while Fig. 3 depicts the DUs and the modalities of the categorical variables jointly projected on the first factorial plane. Fig. 4 represents variable groups by means of the average correlation (for quantitative variables) and/or correlation ratio (for categorical variables) with the first two axes, which represent 3.5% and 3.3% of the total variance, respectively. These low values are explained by the high number (n = 165) of columns of the analysis table. Eighty percent of the total variance is represented by 44 factorial axes, but we only consider the first two axes here.

The first axis is clearly associated with the first model component for adults (contribution to the axis = 25.6% with a linear correlation coefficient of -0.86) that is monotonously linked to the distance to the village centre and clearly dissociates DUs of the groups *N*, *NN1, NN2* and *S* (Fig. 1). It is also positively correlated with long distances to dolines and rock faces (Fig. 2), null doline lengths, presence of pigs and hens, open peri-domiciliary spaces (no paddock), presence of jackfruits, tilled soils and low proportions of dense forest (Fig. 3). It is negatively correlated with the presence of forest and bare soil, a closed peri-domiciliary space and the absence of animals, hens in particular. The strongest contributors to this axis are the DUs belonging to the zone denoted *NN* in Fig. 3, corresponding, in the geographical space, to the DUs of the groups denoted *NN1* and *NN2* in Fig. 1. The great majority (4 out of 5) of the *P. geniculatus* were found in these DUs (Fig. 3 and Fig. 3b in Part I).

The second axis is defined by the spatial model components (Fig. 4), more precisely by the ninth eigenvector of the adult model (contribution 21.5% and correlation 0.76) and the 23th eigenvector associated with

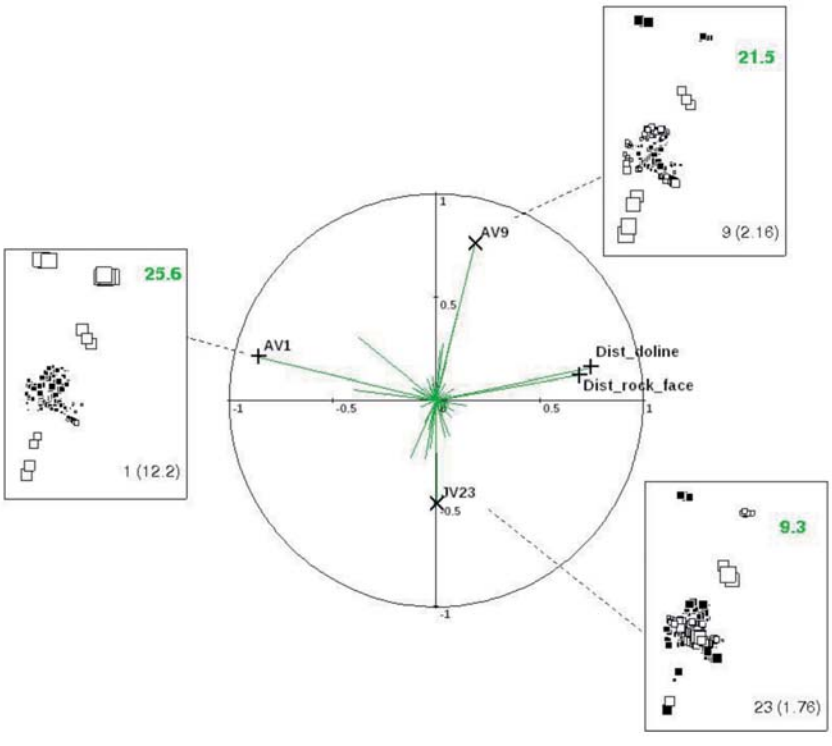


Fig. 2. Quantitative variables that contribute the most to the first two factorial axes, projected on the first factorial plane (first two axes). Symbols + and x represent variables that contribute to the first and the second factorial axis, respectively. The circle represent the correlation circle (radius = 1). For each component of the spatial models (eigenvectors), the spatial representation is shown: black and white squares respectively correspond to positive and negative values of the eigenvector components and the square size is proportional to the absolute value of the eigenvector components. At bottom right, numbers correspond to: the rank of the eigenvector in the spatial model; the eigenvalue (in brackets). At top right, numbers correspond to the contribution percentages to the definition of the first or second axis.

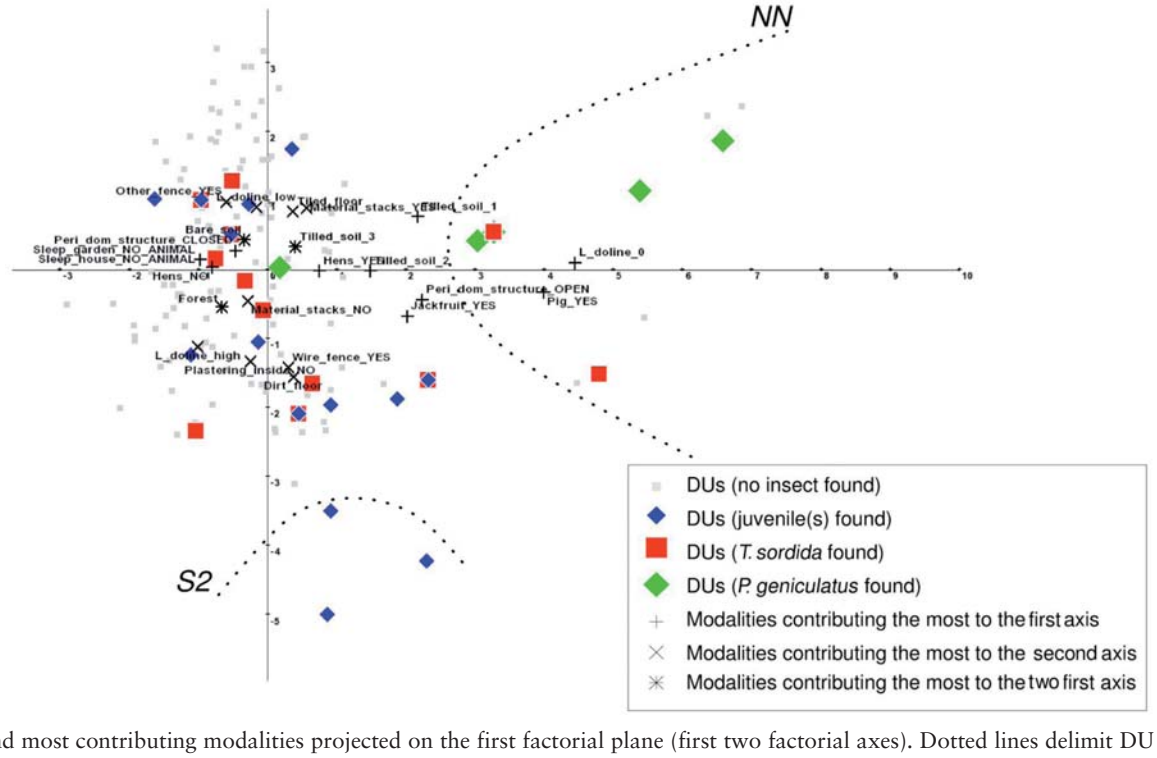


Fig. 3. DU and most contributing modalities projected on the first factorial plane (first two factorial axes). Dotted lines delimit DU groups corresponding to the isolated DUs at the extreme north ($NN=\{NN_1, NN_2\}$) and south ($S_2 \in S$) of the village.

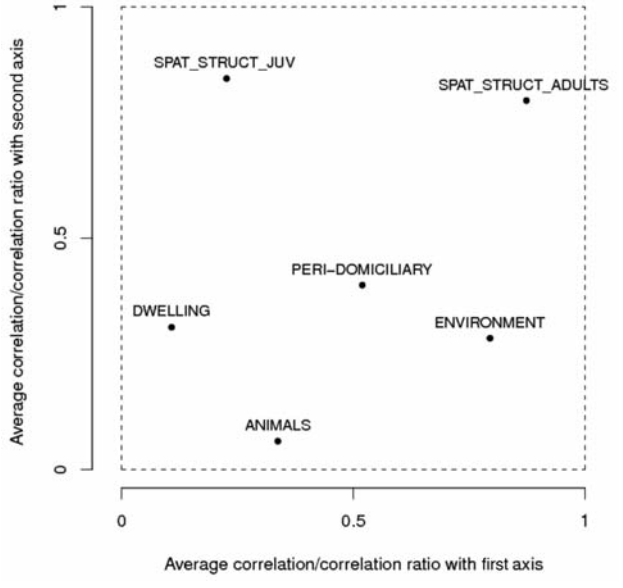


Fig. 4. Representation of the relations between the variable groups and the first two factorial axis.

the juveniles (9.3% and -0.50, respectively) (Fig. 2). These two spatial patterns are associated with the distance to the village centre, both distinguishing the DUs of the centre of the main hamlet from the ones in the peripheral zone. This second factorial axis is associated with the presence of juveniles in the southern part of the village. In fact, DUs belonging to *S2* in Fig. 3 were those located in the extreme south of the *S* group in Fig. 1, and they were all infested by juveniles. DU features associated with the positive part of this axis (centre of the main hamlet) consisted of tiled floors, closed peri-domiciliary spaces, loose materials stacked near the house, and the surrounding environment characterised by weak interactions with dolines and relative high proportions of bare or tilled soil (Fig. 3). Dirt floors and no inside plastering were DU features associated with the negative part of the axis (the periphery of the main hamlet) for the "Dwelling" variable group. For the "Environment" variable group, the features were wire fences and strewn, rather neatly piled, loose materials plus a relatively high proportion of forest and strong interactions with dolines. Figure 4 substantiates the results: variable groups associated with the environment and the spatial model for adults especially linked with the first factorial axis, spatial models for adults and juveniles linked with the second axis.

This first analysis emphasizes differences between the DUs located in the main hamlet and isolated ones. It also contributes to the discrimination of DUs located in the very heart of the village and the ones located in its immediate periphery. However, the fact that all DUs, including isolated ones, participated in the analysis can mask local spatial characteristics associated with the main hamlet. Even if such local patterns could be observed by investigating the next factorial axes, we preferred performing a second analysis by removing DUs belonging to the groups *N*, *NN*1, *NN*2 and *S* (Fig. 1) in order to focus on the main hamlet characteristics.

Analysis of DUs in the main hamlet

Figures 5, 6 and 7 are related to DUs belonging to the main hamlet only and have the same symbols as those in Figs. 2, 3 and 4. The first two factorial axes of this analysis represent 5.6% and 4.7% of the total variance, while 39 of them were needed to cover 80% of the total variance.

The first axis is associated with spatial patterns separating the centre and the peripheral DUs of the main hamlet: eigenvectors JV23 (contribution 12.0% and correlation -0.74); AV24 (7.4%, -0.75); AV4 (3.0%, -0.48); AV9 (7.1%, 0.73); and AV1 (7.3%, 0.75). Fig. 8a, representing the DU coordinates on the first factorial axis, underlines this result depicting information similar to the second axis of the first analysis. The dwellings of the main hamlet centre (the positive coordinates of the first factorial axis) had tiled floors, lacked cracks in the walls and were also characterised by generally good hygiene and no apparent disorder (Fig. 6). Mango trees in the peri-domiciliary spaces and stacked materials were observed and the surrounding environment was associated with a relative high proportion of bare or tilled land and a relatively high level of landscape fragmentation (Fig. 5).

The peripheral DUs were associated with non-plastered walls and a general apparent disorder inside the houses as well as in the peri-domiciliary spaces and there was no mango tree. In addition, a relative high proportion of forest and long stretches of rock faces characterised the surrounding environment. The fact that the spatial model components that contributed most to the axis were provided by the adult model (Fig. 7) associates this factorial axis with the presence of adult insects in the main hamlet periphery. These results are confirmed by (i) the fact that 64% of the DUs were infested by adult specimens (7 out of 11) and had negative coordinates on the factorial axis (Fig. 6), a sign particularly associated with the peripheral DUs (white squares in Fig. 8a) and (ii) by the spatial distribution of the adults (see Fig. 3b in Part I).

The second axis was found to be especially associated with the spatial model of juvenile abundance (Figs.

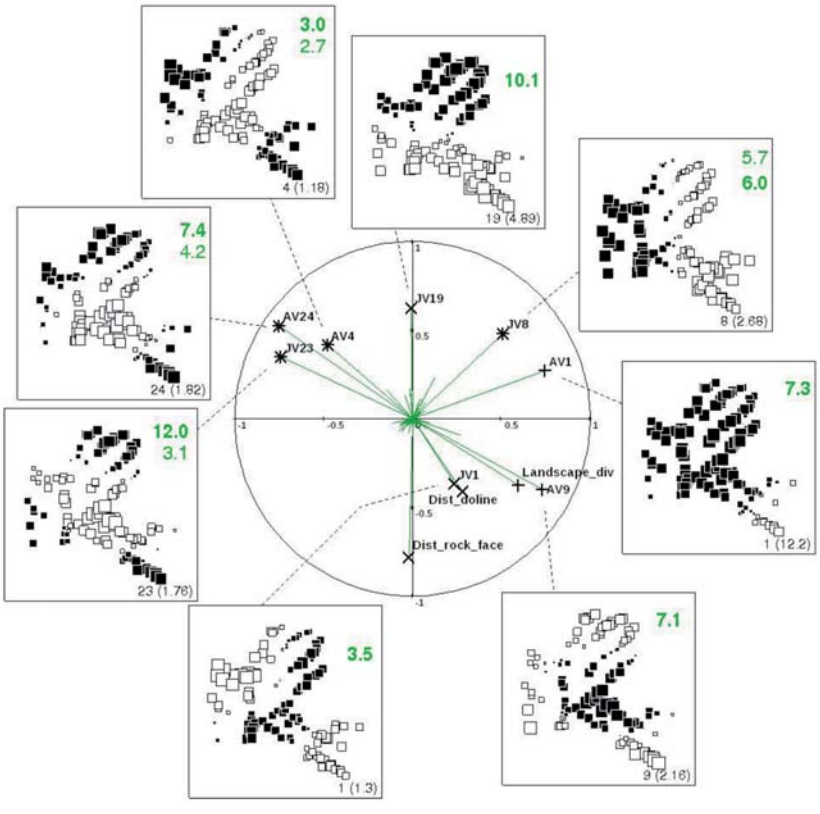


Fig. 5. Quantitative variables that contribute the most to the first two factorial axes, projected on the first factorial plane (first two axes), after the exclusion of the isolated DUs. See the caption of the Fig. 2 for details on symbols.

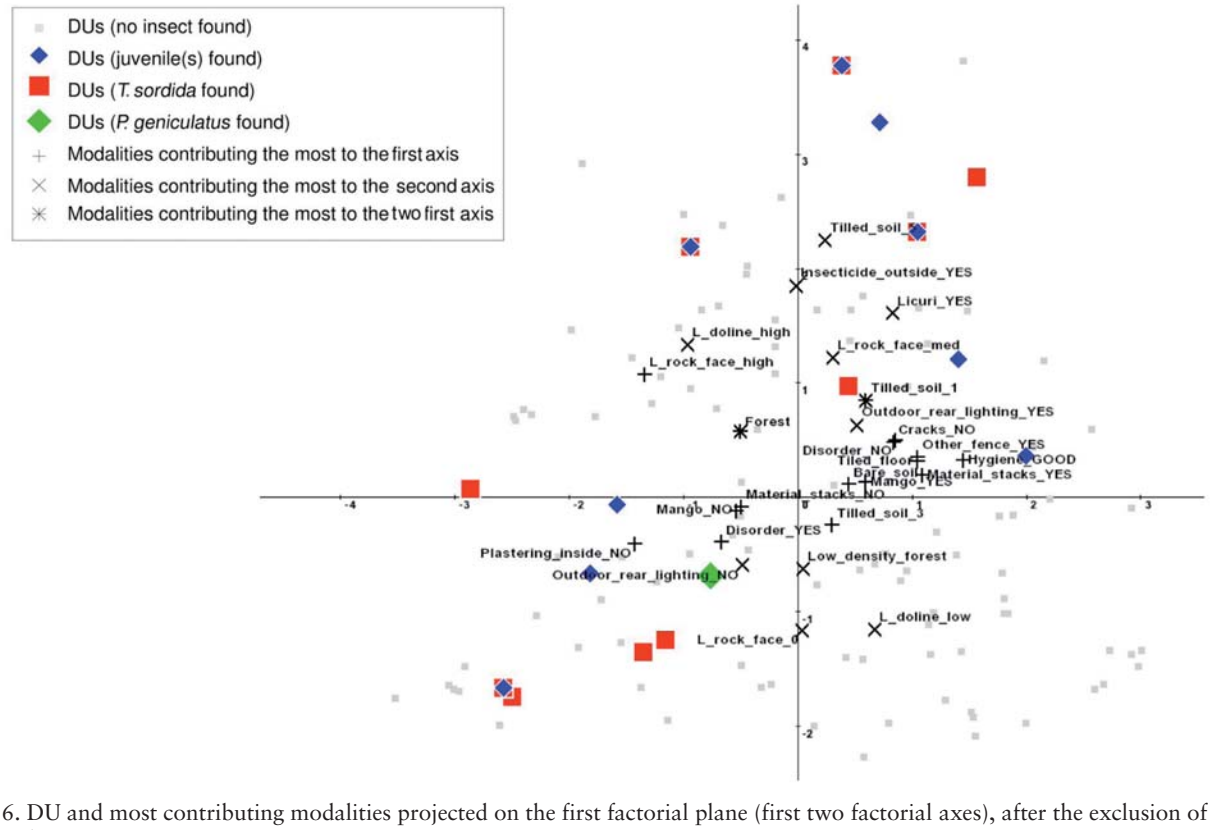


Fig. 6. DU and most contributing modalities projected on the first factorial plane (first two factorial axes), after the exclusion of the isolated DUs.

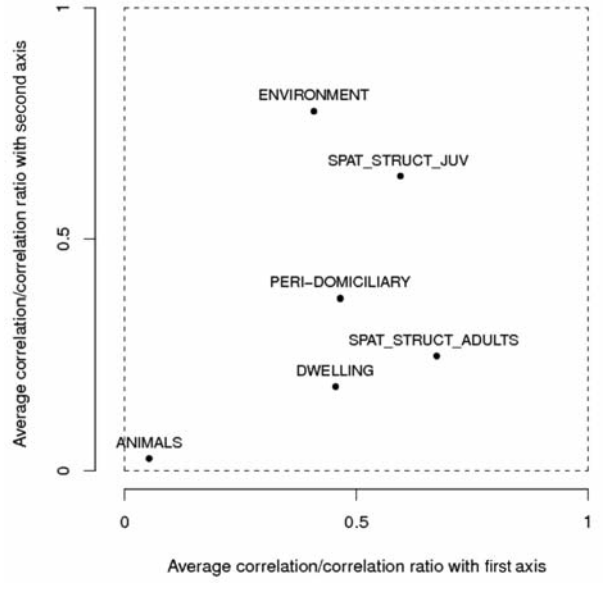


Fig. 7. Representation of the relations between the variable groups and the first two factorial axis, after the exclusion of the isolated DUs.

5 and 6). This result contributes to the tendency of differentiation between the south-eastern and the north-western parts of the village. In fact, the dwellings of the latter were associated with positive values of the following eigenvectors: JV19 (contribution 10.1%, correlation 0.62), JV8 (5.7%, 0.51), AV24 (4.2%, 0.52), AV4 (2.7%, 0.41), and JV23 (3.1%, 0.35) and with negative values of JV1 (3.5%, -0.37). The south-eastern part of the village had the opposite sign of the same eigenvectors. Again, the DU coordinates shown in Fig. 8a underline the result.

The dwellings in the south-eastern part of the main hamlet are characterized by being far from rock faces (4.6%, -0.77) and dolines, which are also not prominent structures (Fig. 5). In addition, there is a high proportion of low-density forest and no outdoor lighting at the rear of the houses. The DUs of the north-western part of the village, on the other hand, are characterised by the presence of outdoor lighting, the use of insecticides outside, a relatively high proportion of tilled soils, long lengths of doline perimeter and rock faces plus presence of licuri palms (*Syagrus coronata*).

The fact that the spatial model components that contribute the most to the axis is provided by the juvenile model associates this factorial axis with juvenile presence (Fig. 7). This result is confirmed by (i) the fact that 67% of the DUs infested by juvenile insects (6 out of 9) had positive coordinates on the factorial axis (Fig. 6), a sign also associated with the peripheral DUs (black squares in Fig. 8b) and (ii) by the spatial distribution of the juveniles (see Fig. 3a in Part I).

Complementary results can be derived from the joint interpretation of the first two factorial axes and the representation of the infested DUs in the factorial plane (Fig. 6). In fact, infestation seems to be associated with: i) the bottom left quarter of the factorial plane (Fig. 6), i.e. the peripheral DUs located in the south-eastern part of the main hamlet and essentially characterized by dwellings with neither outdoor rear lighting nor inside plastering and with apparent disorder; ii) the top right quarter of the factorial plane (Fig. 6), i.e. the north-western part of the main hamlet centre, characterised by dwellings with outdoor lighting, no cracks in the house walls, good hygiene and no disorder as well as by peri-domiciliary and environmen-

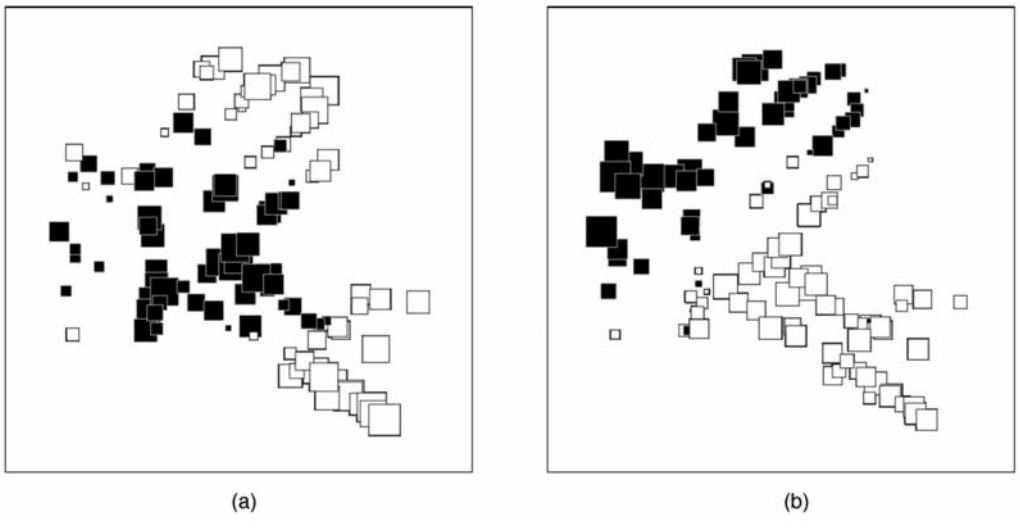


Fig. 8. Spatial representation of DU coordinates on (a) the first and (b) the second factorial axis of the FAMG (with DUs of the main hamlet only). Black and white squares correspond to positive and negative values of the DU coordinates on the axis, respectively. The square size is proportional to the absolute value of the coordinates.

tal variables: use of insecticides, medium to long rock faces, materials collected in neat piles, presence of licuri palms plus a high proportion of tilled soil.

Discussion

As any exploratory factorial analysis, FAMG makes it possible to rapidly isolate relevant information in spite of many variables when no *a priori* knowledge is available, neither of data nor process. In addition, it permits joint analysis of categorical and continuous variables, while interpretation is straightforward as the PCA rules for quantitative variables, and the multiple correspondence analysis ones for categorical variables, can be applied. Moreover, the results can be interpreted at the level of variable groups, allowing the comparison of relative influences of factors that are *a priori* equilibrated, also at the variable/modality level. On the other hand, although the approach permits the reduction of dimensionality and facilitates specifying further investigations that can be performed with more classical statistical methods, the interpretation of the FAMG results remains mainly qualitative.

In this study, FAMG provided many factorial axes, each of them gathering a relatively low portion of the total variance. However, method permitted us to hierarchically explore numerous and heterogeneous variables and, as discussed below, produce substantial information based on relatively small entomological datasets.

The presence of *T. sordida* juveniles in the southern, isolated DUs was associated with simple dwellings (dirt floors and no plastering). However, these residential shortcomings cannot fully explain the insect presence, as 19 juvenile specimens out of 20 were found in the peri-domiciliary space. On the other hand, the strong association with dolines and the high proportion of dense forest in the surrounding areas support the findings. Moreover, hen breeding in the peri-domiciliary space indicates juvenile presence, even if this variable does not explicitly appear in the FAMG results.

The same critical analysis has to be conducted concerning the peripheral DUs of the main hamlet. These DUs were associated with dwellings without inside plastering and with apparent disorder. However, as mentioned in Part I and in previous studies (Slimi et al., 2009; Barbu et al., 2010; Ramirez-Sierra et al., 2010), infestation can proceed from the surrounding area towards the village centre and with small dispersal distances. Thus, the explanatory factor for insect presence seems to be the spatial structure of the village, which is displayed in Fig. 7, where the spatial model components were shown to be the most correlated groups with the fisrt axis.

The presence and abundance of the insects in the north-western part of the main hamlet did not seem to be associated with house characteristics but rather with behavioural parameters such as outdoor lighting at the rear, stacked loose materials and presence of licuri palms in the peri-domiciliary areas as well as proximity to dolines and rock faces in the surrounding environment. Moreover, the fact that insects, essentially only juveniles, were found in the peri-domiciliary spaces, confirms the predominant role of this area in the infestation of this part of the village. Outdoor lighting has previously been identified as an attractive factor for adult triatomines (Walter et al., 2005; Roux et al., 2009; Barbu et al., 2010), and can thus be considered a facilitating factor for insect settlement. On the other hand, the presence of adult specimens in the south-eastern part of the village seems to be associated with features such as lack of plastering inside the houses and apparent disorder, a scene that corresponds to suitable conditions for *T. sordida* domiciliation. This thinking is confirmed by the fact that all insects were found inside the houses in this part of the village.

Although DU features and the occurrence of insects are not necessarily causally related, our results point to some particular residential parameters found in the five main village areas:

(i) the northern isolated DUs, associated with the presence of *P. geniculatus* adults, are situated far from dolines and rock faces, characterised by pig and hen breeding and surrounded by tilled areas;
(ii) the southern isolated DUs, associated with the presence of *T. sordida* juveniles, are situated near dolines and have a high proportion of dense forest in the surrounding areas;
(iii) the peripheral DUs of the main hamlet, associated with the presence of *T. sordida* adults, were characterised by dwellings in apparent disorder and without plastering inside. However, the insect infestation there tended to be due to the dispersal from the "natural" environment towards the village centre;
(iv) the north-western DUs of the main hamlet centre, associated with *T. sordida* juveniles in the peri-domiciliary space, are situated near dolines and rock faces. They were further characterised by outdoor rear lighting, presence of material stacks and licuri palms. In this area, infestation seemed to come from a highly infested site, where 43% of

the exuvii, 16% of the eggs and 63% of the juveniles were found (see Part I); and

(v) the south-eastern part of the main hamlet periphery, where *T. sordida* adults were found inside the houses, displayed factors promoting triatomine domiciliation such as dwellings in apparent disorder and without plastering inside.

Although environmental parameters alone cannot explain spatial distribution repartition of the insects, they play a key role in the explanation of spatial structures. PCNM (see Part I) is a multi-scale approach permitting the specification of the relevant spatial scale for environmental characterisation. In this study, the environmental characteristics were especially linked with large-scale patterns. The landscape characterisation was performed within a 150 m radius around the houses and an environmental characterisation bias can proceed from this radius choice. In fact, a smaller radius would permit underlining environmental differences at a more local scale and to characterise more local patterns. Without precise biological knowledge of insect flight capacities and behaviour, further investigations should consider various buffer sizes or a specific methodology to choose the most relevant buffer size(s). However, investigating the next factorial axes of the FAMG should permit the characterisation of more local phenomena.

Conclusions

The methodology proposed in this two-part paper constitutes a unified and intuitive approach in terms of modelling and analysis methods and criteria, which should be suitable for eco-epidemiological system studies. The finding of an association of certain village "districts" with different types and causes of infestation, can contribute to the formulation of an improved control strategy for Chagas disease insect vector populations.

Acknowledgements

The field work, realised by A. F. Venâncio during her PhD studies, benefitted from collaboration and methodological support elaborated during the CNPq-IRD programme "Ecologie du paysage, dynamique des agroécosystèmes et complexes pathogènes: la définition du risque éco-épidémiologique dans le cas de la trypanosomose américaine, EDCTA" (2002-2004).

We are thankful for the authorisation for field work in the APA by the Brazilian institutions (IBAMA, CECAV, Ministério da Saúde, Comitê de Ética em Pesquisa and Secretaria de Meio Ambiente e Recursos Hídricos do Estado da Bahia - SEMARH, Superintendência de Biodiversidade, Florestas e Unidades de Conservação - SFC). All authors thank the inhabitants of Santa Rita for their participation and the SEAS-Guyane project, which provided the SPOT5 image.

References

Barbu C, Dumonteil E, Gourbière S, 2010. Characterization of the dispersal of non-domiciliated *Triatoma dimidiata* through the selection of spatially explicit models. PLoS Negl Trop Dis 4, e777.

Batista TA, Gurgel-Gonçalves R, 2009. Ecological niche modelling and differentiation between *Rhodnius neglectus* Lent, 1954 and *Rhodnius nasutus* Stål, 1859 (Hemiptera: Reduviidae: Triatominae) in Brazil. Mem Inst Oswaldo Cruz 104, 1165-1170.

Camara G, Souza RCM, Freitas UM, Garrido J, 1996. SPRING: integrating remote sensing and GIS by object-oriented data modelling. Computers & Graphics 20, 395-403.

Costa J, Peterson AT, Beard CB, 2002. Ecologic niche modeling and differentiation of populations of *Triatoma brasiliensis* Neiva, 1911, the most important Chagas' disease vector in northeastern Brazil (Hemiptera, Reduviidae, Triatominae). Am J Trop Med Hyg 67, 516-520.

Dray S, Chessel D, Thioulouse J, 2003. Co-inertia analysis and linking of ecological data tables. Ecology 84, 3078-3089.

Dumonteil E, Gourbiére S, 2004. Predicting *Triatoma dimidiata* abundance and infection rate: a risk map for natural transmission of Chagas disease in the Yucatan peninsula of Mexico. Am J Trop Med Hyg 70, 514-519.

GRASS Development Team, 2010. Geographic Resources Analysis Support System (GRASS GIS) Software. Open Source Geospatial Foundation, USA. http://grass.osgeo.org (accessed on December 2010).

Khan OA, Davenhall W, Ali M, Castillo-Salgado C, Vazquez-Prokopec G, Kitron U, Soares Magalhaes RJ, Clements A, 2010. Geographical information systems and tropical medicine. Ann Trop Med Parasitol 104, 303-318.

Lê S, Josse J, Husson F, 2008. FactoMineR: an R package for multivariate analysis. Department of Statistics, University of California at Los Angeles, USA. http://www.jstatsoft.org/v25/i01/paper (accessed on December 2010).

Ostfeld RS, Glass GE, Keesing F, 2005. Spatial epidemiology: an emerging (or re-emerging) discipline. Trends Ecol Evol 20, 328-336.

Pagès J, 2002. Analyse factorielle multiple appliquée aux variables qualitatives et aux données mixtes. Rev Stat Appl 52, 5-37.

Pagès J, 2004. Analyse factorielle de données mixtes. Rev Stat Appl 52, 93-111.

Peterson AT, Sánchez-Cordero V, Beard CB, Ramsey JM, 2002. Ecologic niche modeling and potential reservoirs for Chagas

disease, Mexico. Emerg Infect Dis 8, 662-667.

R Development Core Team, 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Ramirez-Sierra MJ, Herrera-Aguilar M, Gourbière S, Dumonteil E, 2010. Patterns of house infestation dynamics by non-domiciliated *Triatoma dimidiata* reveal a spatial gradient of infestation in rural villages and potential insect manipulation by *Trypanosoma cruzi*. Trop Med Int Health 15, 77-86.

Roux E, de Fátima Venâncio A, Girres J-F, Romaña C, 2009. High resolution remote sensing and heterogeneous data analysis for local scale characterization of environmental risk: an application to Chagas disease in endemic areas. In XIV Simpósio Brasileiro de Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais (INPE), Natal, Brazil, pp. 7577-7586.

Schofield CJ, Lehane MJ, McEwan P, Catalá SS, Gorla DE, 1991. Dispersive flight by *Triatoma sordida*. Trans R Soc Trop Med Hyg 85, 676-678.

Silveira AC, Rezende DF. 1994. Epidemiologia e controle da transmissão vetorial da doença de Chagas. Rev Soc Bras Med Trop 27, 11-22.

Slimi R, Elyacoubi S, Dumonteil E, Gourbiere S, 2009. A cellular automata model for Chagas disease. Appl Math Model 33, 1072-1085.

Tsoar A, Allouche O, Steinitz O, Rotem D, Kadmon R, 2007. A comparative evaluation of presence-only methods for modelling species distribution. Diversity and Distributions 13, 397-405.

Vazquez-Prokopec GM, Spillmann C, Zaidenberg M, Kitron U, Gürtler RE, 2009. Cost-effectiveness of Chagas disease vector control strategies in north-western Argentina. PLoS Negl Trop Dis 3, e363.

Walter A, Pojo do Rego I, Ferreira AJ, Rogier C, 2005. Risk factors for reinvasion of human dwellings by sylvatic triatomines in northern Bahia State, Brazil. Cad Saude Publica 21, 974-978.