APPLICATION OF SATELLITE IMAGES TO STUDY THE DISTRIBUTION OF Rhodnius pallescens BARBER 1932, VECTOR OF CHAGAS DISEASE

APLICACIÓN DE IMÁGENES SATELITALES PARA EL ESTUDIOS DE LA DISTRIBUCIÓN DE Rhodnius pallescens BARBER 1932, VECTOR DE LA ENFERMEDAD DE CHAGAS

Sair Arboleda* and Nicolás Jaramillo

Grupo Biología y Control de Enfermedades Infecciosas (BCEI), Instituto de Biología, Universidad de Antioquia, Sede de Investigaciones Universitarias – SIU, Calle 62 # 52–59 Laboratory 620, P.O. Box: 1226, Medellín, Colombia. *sairorieta@yahoo.es.

ABSTRACT

Chagas disease is the most disabling disease in Latin America. Because the current increase of global human migration threatens to widen the disease range, including vector borne diseases, to study the distribution of insect vectors becomes an important task. Rhodnius pallescens is the main Chagas disease vector in Panama, and a secondary one in Colombia, Nicaragua and Costa Rica. Given their wild habits, their distribution could be highly influenced by climatic factors. In this study we built a geographical distribution model of this vector using the method of maximum entropy (Maxent) to identify sites having the highest occurrence probability of finding the vector species. The results obtained demonstrate that the sites predicted by the model, as those with the greatest occurrence probability, fitted with those already recorded in the field as vector presence sites. The high sensitivity of the prediction was evidenced by the value of the area under the ROC curve (AUC=0.995).

RESUMEN

La enfermedad de Chagas es la más incapacitante en América Latina. Debido a que el actual aumento de la migración humana mundial amenaza con ampliar la gama de enfermedades, incluyendo enfermedades transmitidas por vectores, el estudio de la distribución de los insectos responsables se convierte en una tarea importante. Rhodnius pallescens es el principal vector de la enfermedad de Chagas en Panamá, y es secundario en Colombia, Nicaragua y Costa Rica. Dados sus hábitos silvestres, su distribución está muy influenciada por factores climáticos. En este estudio se construyó un modelo de distribución geográfica de este vector utilizando el método de máxima entropía (Maxent) para identificar los sitios donde se produce una mayor probabilidad de ocurrencia de esta especie. Los resultados obtenidos demuestran que los sitios predichos por el modelo con mayor probabilidad de ocurrencia coinciden con los datos registrados en campo como sitios de la presencia del vector. La alta sensibilidad de la predicción se evidencia por el valor del área bajo la curva ROC (AUC = 0.954).

Keywords: Rhodnius pallescens, Chagas disease, Geographic Information Systems, Satellite images, AVHRR, maximum entropy.

Palabras clave: Rhodnius pallescens, enfermedad de Chagas, Sistemas de Información Geográfica, imágenes satelitales, AVHRR, máxima entropía.
INTRODUCTION

Chagas disease is considered an emerging disease in Spain and the United States, and is re-emerging in many countries of the Americas. According to estimates of WHO (2007) the disease affects approximately 15 million people in Central and South America. The etiologic agent is the protozoan Trypanosoma cruzi, which is transmitted by blood-sucking insects of Triatominae subfamily. Since massive treatment or vaccination against T. cruzi is not feasible due to the lack of suitable drugs or vaccines, the best approach to Chagas disease control concerns transmission mechanisms.

It is estimated that Chagas disease causes two million DALYs (Disability Adjusted Life Years) annually, making it the most disabling disease in Latin America (Tarleton et al., 2007). Additionally, the growing trend of global migration threatens to expand exponentially Chagas disease from rural to urban areas and from endemic to non-endemic regions (Franco-Paredes et al., 2007), so the distribution study of the insect vectors becomes an important task in this epidemiological scenario.

The Geographic Information systems (GIS) had been used to study the Chagas disease vectors in different ways, namely to stratify the risk of infection with the parasite (Ramsey et al., 2000; Dumonteil and Goubriere 2004; López-Cárdenas et al., 2005), to associate host and vector species (Peterson et al., 2002), and to identify sources and patterns of Triatoma infestans infestation after chemical control (Cecere et al., 2004; Porcasi et al., 2006).

Satellite images integrated into a GIS have allowed to model the potential distribution of both wild and domestic triatomines (Beard et al., 2003; Carabajal de la Fuente et al., 2009; Arboleda et al., 2009) and to identify areas with favorable environmental profiles for its occurrence. These areas could be potential markers for epidemiological risk (Arboleda et al., 2009).

The study of potential Rhodnius pallescens distribution through satellite imagery is attractive because, despite being strictly wild, Rhodnius pallescens is the principal vector in Panama and a secondary one in Nicaragua, Costa Rica and Colombia. Its life characteristics, such as its close association with palm trees, mainly Attalea butyracea (Pizarro and Romaña 1998) and its stenohydric behavior (Zeledón et al., 2006; Jurberg and Rangel, 1984) indicate that their distribution could be highly influenced by climatic factors.

Using a discriminant analysis method (DA), Arboleda and collaborators (Arboleda et al., 2009) recently built a geographical model which identified potential sites of R. pallescens presence and absence. In this study we aimed to build a new model using the maximum entropy method (Maxent) to evaluate the degree of concurrence of the distribution profiles predicted by both techniques to validate the vector presence areas as markers of epidemiological risk.

MATERIALS Y METHODS

Rhodnius pallescens potential distribution was estimated with the software MaxEnt version 3.2.1 (Phillips et al., 2006), using 25% of the data as training sites and the parameters defined by default. Data corresponded to 101 presence sites (determinate by field sampling) covering the entire area of registered occurrence of the species (Arboleda et al., 2009) and also nine environmental variables obtained from temporal series of 20 years of the AVHRR sensor. These variables were those with the highest discriminator power between presence and absence sites in the specific distribution study (Arboleda et al., 2009). According with that study, the variables here considered are: minimum vapor pressure deficit, amplitude tri-annual of air temperature, variance of the complete series for land surface temperature, phase tri-annual of air temperature, phase tri-annual of middle infra-red, percentage of the total variance of the three first Fourier components for air temperature, amplitude bi-annual of land surface temperature, average of land surface temperature, maximum air temperature.

To determine the discriminatory capacity of the model, the area under the ROC curve (Receiver Operating Characteristic) was used (Hanley and McNeil, 1982). This criterion has been successfully applied to evaluate distribution models based on presence-absence methods (Fielding and Bell, 1997; Segurado and Araújo, 2004; Muñoz and Feliciísimo, 2004) and presence only methods. The area under the ROC curve (AUC) indicates the probability that the value provided by the model for the presence site is greater than that expected randomly. Values close to 1 indicate a good fit to the model and those close to 0.5 a fit not better than that obtained by chance.
RESULTS AND DISCUSSION

Methods that use vector presence and absence data are adequate for modeling potential vector distribution but the lack of absence data in most of the studies (Phillips et al., 2006) is a constraint that demands the use of techniques, such as Maxent, using presence data only.

Contrary to DA method, Maxent allows identifying sites with the highest occurrence probability in the greatest area of distribution of a species, based only on its recorded distribution. Maxent uses entropy as a mean to generalize specific observations of presence of a species and does not need incorporate absence points within its theoretical framework. Recent comparative analysis classifies Maxent as a successful technique for modeling distribution using occurrence data only (Elith et al., 2006).

The presence sites used to build the model of *R. pallescens* potential distribution greatly coincides with the occurrence sites of increased probability (Figure 1). Moreover, the area under the ROC curve shows high sensitivity of the method (AUC = 0.954; Figure 2). The predicted *R. pallescens* distribution showed by Maxent has a good correspondence with distribution predicted by DA method (Figure 3), which shows the usefulness of the Maxent.

Results suggest that this application of spatial analysis to disease-vector ecology can identify geographic areas with environmental characteristics suitable for the species occurrence, and could helps to identify at a species distribution level, the epidemiological risk areas. This is an important tool for surveillance regarding the effort required for field sampling (in terms of time and cost) of Chagas’ vector species broadly distributed. Maxent modeling excels over other alternatives because the software is freely distributed, is of easy handled, and only requires data about specie presence.

![Figure 1](image1.png)

**Figure 1.** Distribution model of *R. pallescens* using the Maxent method. The white boxes indicate the sites where the species was collected. Values higher than 0.5 (green to red colors) indicate a high probability of *R. pallescens* presence.

![Figure 2](image2.png)

**Figure 2.** Area under the ROC curve (AUC). The red line shows the model fit to training data and the blue line to test data. The turquoise line shows what would be expected if the model was not better than chance. The closer the blue line is above the top left, the model better predict presence sites (Fielding and Bell, 2007).

![Figure 3](image3.png)

**Figure 3.** Correspondence between areas predicted as suitable for *R. pallescens* occurrence. Coincidence between Maxent and DA (from Arboleda et al. 2009) results are showed in red color.
CONCLUSIONS

In a broad scale of analysis, the use of environmental variables related with temperature and humidity, which are considered with its temporal variation, are useful inputs (and predictors) for distribution models. DA and Maxent methods produced similar results, but the latter method only requires presence data. The geo-spatial analysis of satellite imagery linked with field information and those related to ecology and biology of wild triatomines can help in assessing their potential distribution and delimiting the areas for surveillance programs.

LITERATURE CITED


