

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, entre ellas las 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 que tienen la 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 Gourbiere 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; Carbajal 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 specie 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 Felicísimo, 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.

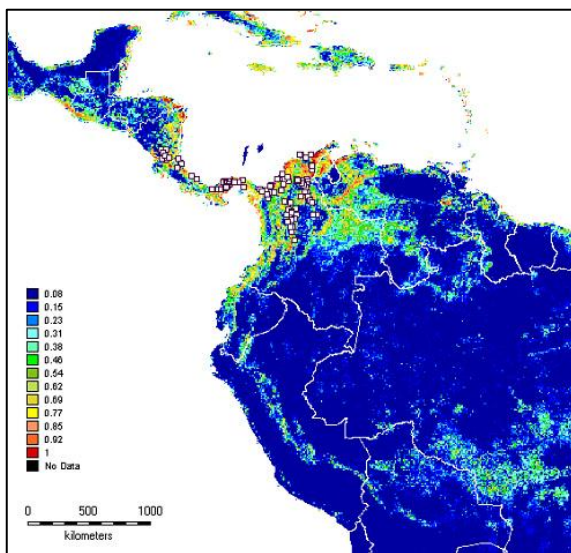


Figure 1. Distribution model of *R. pallescens* using the Maxent method. The white boxes indicate the sites where the species was collected. Values higher than 0.5 (green to red colors) indicate a high probability of *R. pallescens* presence.

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 help 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.

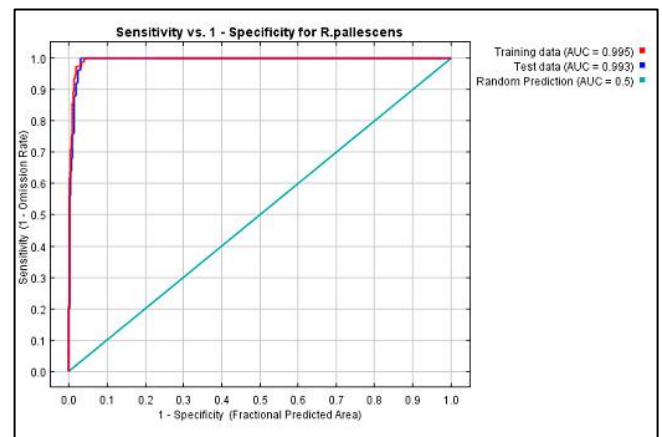


Figure 2. Area under the ROC curve (AUC). The red line shows the model fit to training data and the blue line to test data. The turquoise line shows what would be expected if the model was not better than chance. The closer the blue line is above the top left, the model better predict presence sites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

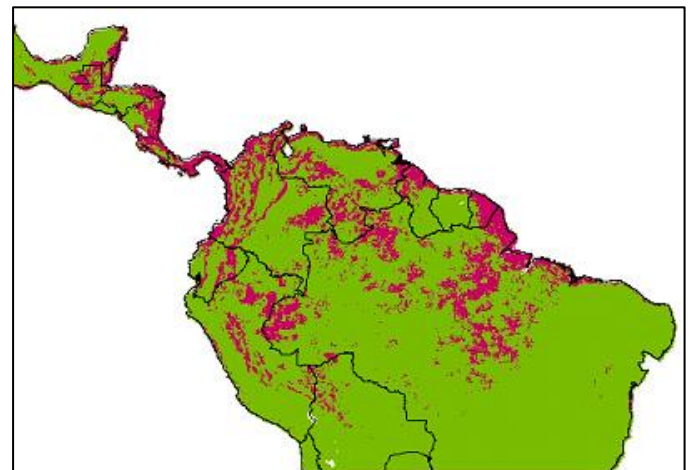


Figure 3. Correspondence between areas predicted as suitable for *R. pallescens* occurrence. Coincidence between Maxent and DA (from Arboleda *et al.* 2009) results are showed in red color.

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

- Arboleda S., D.E. Gorla, X. Porcasi, A. Saldaña, J. Calzada, O.N. Jaramillo. 2009. Development of a geographical distribution model of *Rhodnius pallescens* Barber, 1932 using environmental data recorded by remote sensing. *Infect. Genet. Evol.* 9:441-448.
- Beard, C.B., G. Pye, F.J. Steurer, R. Rodríguez, R. Campean, A.T. Peterson, J. Ramsey, R.A. Wirtz, L.E. Robinson. 2003. Chagas disease in a domestic transmission cycle in Southern Texas, USA. *Emerg. Inf. Dis.* 9:103-105.
- Carbajal de la Fuente, A.L., X. Porcasi, F. Noireau, L. Diotaiuti, D.E. Gorla. 2009. The association between the geographic distribution of *Triatoma pseudomaculata* and *Triatoma wygodzinskyi* (Hemiptera:Reduviidae) with environmental variables recorded by remote sensors. *Infect. Genet. Evol.* 9:54-61.
- Cecere, M.C., G.M. Vazquez-Prokopec, R. Gürtler, U. Kitron. 2004. Spatio-temporal analysis of reinfestation by *Triatoma infestans* (Hemiptera: Reduviidae) following insecticide spraying in a rural community in northwestern Argentina. *Am. J. Trop. Med. Hyg.* 71:803-810.
- Dumonteil E., S. Gourbiere. 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.
- Elith J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J.M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, N.E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129-151.
- Fielding, A.H., J.F. Bell. 2007. A review of methods for the assessment of prediction errors in conservation presence / absence models. *Environ. Conservat.* 24:38-49.
- Franco-Paredes, C., A. Von, A. Hidron, A.J. Rodríguez-Morales, I. Tellez, M. Barragán, D. Jones, C.G. Náquira, J. Mendez. 2007. Chagas disease: an impediment in achieving the Millennium Development Goals in Latin America. *BMC Int. Health Hum. Rights* 7:7.
- Hanley, J.A., B.J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29-36.
- Jurberg, J., E.F. Rangel. 1984. Ciclo biológico de *Rhodnius pallescens* Barber, 1932 (Hemiptera, Reduviidae, Triatominae). *Revista Brasileira Biol.* 40:569-577.
- López-Cárdenas, J., F.E. González-Bravo, P.M. Salazar-Schettino, J.C. Gallaga-Solórzano, E. Ramírez-Barba, J. Martínez-Méndez, V. Sánchez-Cordero, T. Peterson, J.M. Ramsey. 2005. Fine-scale predictions of distribution of Chagas disease vectors in the state of Guanajuato, Mexico. *J. Med. Entomol.* 42:1068-1081.
- Muñoz, J., A.M. Felicísimo. 2004. Comparison of statistical methods commonly used in predictive modelling. *J. Veg. Sci.* 15:285-292.
- Phillips, S.J., R.P. Anderson, R.E. Schapire. 2006. A maximum entropy modelling of species geographic distributions. *Ecol. Model.* 190:231-259.
- Pizarro, J.C.N., C. Romaña. 1998. Variación estacional de una población silvestre de *Rhodnius pallescens* Barber 1932 (Heteroptera: Triatominae) en la costa Caribe colombiana. *Bull. Inst. Fr. études andines* 27:309-325.
- Porcasi, X., S.S. Catalá, H. Hrellac, M.C. Scavuzzo, D.E. Gorla. 2006. Infestation of rural houses by *Triatoma infestans* (Hemiptera: Reduviidae) in southern area of Gran Chaco in Argentina. *J. Med. Entomol.* 43:1060-1067.
- Ramsey, J.M., R. Ordóñez, A. Cruz-Celis, A.L. Alvear, V. Chávez, R. López. 2000. Distribution of domestic Triatominae and stratification of Chagas disease transmission in Oaxaca, Mexico. *Med. Vet. Entomol.* 14:1-12.
- Segurado, P., M.B. Araújo. 2004. An evaluation of methods for modelling species distributions. *J. Biogeography* 31:1555-1568.
- Tarleton, R.L., R. Reithinger, J.A. Urbina, U. Kitron, R.E. Gürtler. 2007. The challenges of Chagas Disease-grim outlook or glimmer of hope. *PLoS Med.* 4(12):e332.
- WHO. 2007. Nuevas perspectivas de investigación aplicadas a las actuales estrategias de control vectorial. Reporte del grupo de trabajo científico sobre la enfermedad de Chagas, 2005. Guhl F., Lazdins-Helds eds. Buenos Aires, Argentina. pp 31-42.
- Zeledón, R., F. Marín, N. Calvo, E. Lugo, S. Valle. 2006. Distribution and ecological aspects of *Rhodnius pallescens* in Costa Rica and Nicaragua and their epidemiological implications. *Mem. Inst. Oswaldo Cruz* 101:75-79.