



E-ISSN: 2320-7078
P-ISSN: 2349-6800
JEZS 2015; 3(4): 91-97
© 2015 JEZS
Received: 14-06-2015
Accepted: 15-07-2015

Jose Lopez-Collado
Colegio de Postgraduados,
Campus Veracruz, Km 88.5
Carretera Xalapa-Veracruz,
Veracruz, Mexico, Phone and
Fax: 011 52 (229) 2010770.

Catalino Jorge Lopez-Collado
Colegio de Postgraduados,
Campus Veracruz, Km 88.5
Carretera Xalapa-Veracruz,
Veracruz, Mexico, Phone and
Fax: 011 52 (229) 2010770.

Correspondence:
Jose Lopez-Collado
Colegio de Postgraduados,
Campus Veracruz, Km 88.5
Carretera Xalapa-Veracruz,
Veracruz, Mexico, Phone and
Fax: 011 52 (229) 2010770.

Multi-criteria geographic distribution approach: A case study of *Rhyssomatus nigerrimus* Fhåraeus (Coleoptera: Curculionidae)

Jose Lopez-Collado, Catalino Jorge Lopez-Collado

Abstract

The objective of this study was to estimate the potential distribution of *Rhyssomatus nigerrimus* in Mexico. Point presence data and bioclimatic data were used to construct the model. Kernel density distributions were estimated for the bioclimatic layers to represent the response function and to combine them. The vulnerability of municipalities planted with soybean was computed as a function of suitability and revenue. Results showed a medium to high potential occurrence in the northern states of Mexico: Tamaulipas, North of Veracruz in the Gulf of Mexico and Sonora and Sinaloa in the Pacific Ocean. Low suitability occurred in the northern and central areas of the country. The Area under the Curve was 0.805 and the True Skill Statistics was 0.625, both indices supporting the validity of the model. Soybean-planted municipalities in Tamaulipas, San Luis Potosi, Chiapas and Campeche were estimated to have a medium to high vulnerability to pest occurrence.

Keywords: Soybean pests, analytic hierarchic process, species distribution modeling, Curculionidae.

1. Introduction

The Mexican Soybean Weevil (*Rhyssomatus nigerrimus*) Fhåraeus (Coleoptera: Curculionidae) is native to America and is found in Mexico, Barbados, Martinique, St. Vincent, Belize, Guatemala, Panama and South America [1]. It was first reported in Mexico from collection trips in the states of Veracruz and Guerrero [2]. Later, it was recorded in Guanajuato in wild hosts and in Nayarit and Tabasco [3, 4]. Adults of the Mexican Soybean Weevil have different host plants, including *Acacia* spp., *Amaranthus* spp., *Ficus* spp., and *Ipomoea* spp. [4]. Recently, it was found as a pest of soybeans in San Luis Potosi, Tamaulipas and Chiapas [5, 6]. In soybeans, it feeds on the vegetative and reproductive structures. Adults feed on plants and seedlings, attacking stems, branches, sprouts, and buds. The larvae feed inside the pods; later, it drops to the soil, transforms to pupa and overwinters as adult. The insect shows one generation per year [5]. Soybean is planted in Mexico in about 178,530 ha in the states of Campeche, Chiapas, Chihuahua, Jalisco, Nuevo Leon, Quintana Roo, San Luis Potosi, Sonora, Tamaulipas, Veracruz and Yucatan, generating revenue near US\$ 110 million [7].

The geographic distribution of organisms is a key factor in risk analysis and pest control. Habitat distribution is estimated by means of correlative, deductive and hybrid methods. Inductive or correlative methods rely on georeferenced point data, collected by sampling or by specimens stored in museums and bioclimatic layers [8]. Deductive methods use the physiological response of organisms to bioclimatic variables such as temperature and precipitation, which are known to influence the geographic distribution of species; hybrid techniques combine both methods [9].

The diverse results obtained from different modeling tools suggest combining them [10]. Alternatively, input layers can be combined based on their similarity to conditions where the species are known to thrive, using ad-hoc procedures [9]. This approach requires processing different data layers, which are combined by a simple or weighted sum [11, 12]. The analytic hierarchic process AHP combines data from different sources in such a way that it uses heuristics to rank the relative importance of the data by making pairwise comparisons [12, 13]. The ranks are converted to weights and assigned to the data layers which are grouped in categories for combination. This method has been applied to estimate suitable land areas for rice production [14], determine the potential distribution of forest pests [15] and to select landfill

zones [16]. Because *R. nigerrimus* is a pest in its initial colonization stage of soybean crops in Mexico, it is important to know its potential geographic range and which municipalities with soybeans may be affected. Therefore the goal of this study was to apply the AHP to find the potential distribution of *R. nigerrimus* and estimate soybean regions susceptible to potential damage in Mexico.

2. Materials and Methods

This work applied a composite hierarchic approach to compute the potential distribution of *R. nigerrimus* [17], and it was realized during 2013-2015 from data compiled at different dates. The first step was extracting the bioclimatic data layer values from the points where the specimens have been reported, these values were then transformed using a response function. In the second step, weights associated to each layer were computed based on pairwise comparisons; weights corresponded to the normalized eigenvalues of a pairwise comparison matrix representing the variables grouped by categories. Weights derived from the pairwise comparisons allow ranking the relative importance of each layer. The third step involved combining the layers by using map algebra operations. The model was evaluated by splitting the occurrence data in two sets: one used for training and another for testing. Crop vulnerability was computed by multiplying the potential distribution index by revenue.

2.1. Specimen data

To estimate the potential distribution of *R. nigerrimus*, geo-referenced point data were pooled from collection reports [2, 4, 6], the data set was split in $n=25$ (70 %) for training and $n=10$ (30 %) for testing.

2.2. Selection of bioclimatic layer values and response function

Values were extracted from numeric [18] and nominal bioclimatic layers at the points where the species was present. The numeric variables were: mean annual temperature, mean temperature of coldest quarter and mean annual precipitation. Nominal values for soil and climate types were also extracted [19, 20]. These variables are thought to affect some biological traits of the species, such as overwintering and survival [6]. To combine the layers it is necessary to standardize the raster values between zero and one [12], zero meaning no suitability and one indicates maximum suitability. Different transformation functions have been proposed depending of the problem domain. For numeric variables in the sample, the response was modeled by a kernel density distribution, while for categorical variables each value was ranked according to its frequency to obtain a discrete density estimate. The layers were then transformed to their corresponding density estimates. In both cases, the kernel distributions reflected the information contained in the observed values so that the most frequent received the highest response.

2.3. Assessment and pooling of bioclimatic layers

In the second step, the input variables were ranked by pairs, using a nine points scale [21]; rankings are subjective measures of how important a variable is related to another one, and the weights were obtained by computing the normalized eigenvalues of the square matrix, where rows and columns corresponded to the input variables and the matrix cells to the pairwise rank [17]. The suitability map (Y_f) was created by combining the layers with their assigned weights:

$$X_c(ij) = \sum_i w_i X_{ij}$$

, where $X_c(ij)$ is a particular cell value of the composite layer, w and X are the weight and input

layer respectively, while j refers to any cell value of the i -th raster layer. This procedure was applied at each step depicted in Fig. 1. The suitability map was aggregated using the k-mean clustering algorithm to obtain four classes indicating different degrees of geographic suitability: unsuitable, low, medium and high suitability.

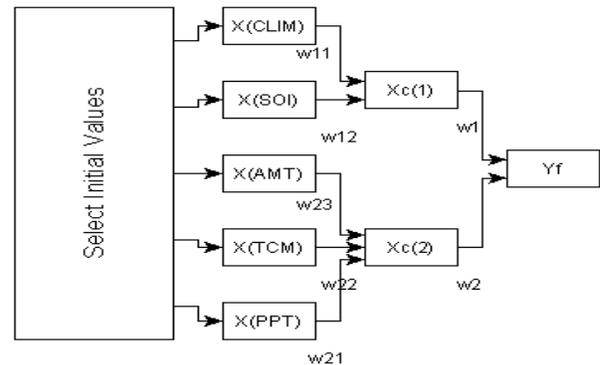


Fig 1: Hierarchic approach applied to compute the distribution model for *R. nigerrimus*. The $X(i)$'s are the input variables or data layers, the $X_c(i)$'s are the composite layers; the w 's are weights used for combining the layers to obtain Y_f , the final distribution map.

2.4 Model evaluation

Model performance was evaluated by using the independent test data to estimate two different metrics: the Receiver Operating Characteristic (ROC) curve and the corresponding Area under the Curve (AUC), and by computing the True Skill Statistics (TSS). These metrics were computed using the test data and 10,000 background point data extracted randomly from the suitability map [22]. The AUC is a threshold independent index that tends to one when the model is appropriate and to 0.5 when it cannot be distinguished from a random model. The TSS range from -1 to 1, the closer to 1 the better the model [23]. The TSS requires converting the suitability map to a presence (1)-absence (0) map U by applying a threshold value th , so that $U = 1$ if $Y_f \geq th$, 0 otherwise. TSS value was computed from observed data point at the lower suitability value. For the AUC, test data was resampled to obtain an estimate of the confidence interval [24].

2.5 Vulnerability of soybean crop

The vulnerability of soybean to this species was estimated as the product of mean potential occurrence and total revenue at the municipality level [11], using data for year 2013 [7], revenue values were normalized between zero and one [12]. Operations were performed with R 3.1.2 and Mathematica 8.01 [25, 26].

3. Results and Discussion

3.1. Kernel distributions and layer combination

The kernel density estimates are presented in Fig. 2 and 3 for the numeric and nominal sample points. The estimates represented the expected suitability values for the bioclimatic layers derived from the train presence data. Suitable temperatures ranged from 10 to 15 °C up to near 30 °C while precipitation increased from zero up to a maximum near 1,000 mm and then decreased in an irregular shape to about 4,000 mm. In the case of climate and soil types, which are nominal variables, the response values corresponded to the normalized frequency values. The most suitable climate was Aw, corresponding to a tropical wet type. The most common soil types where the insect was found were Pellic Vertisol (Vp) and Chromic Luvisol (Lc) (Fig. 3).

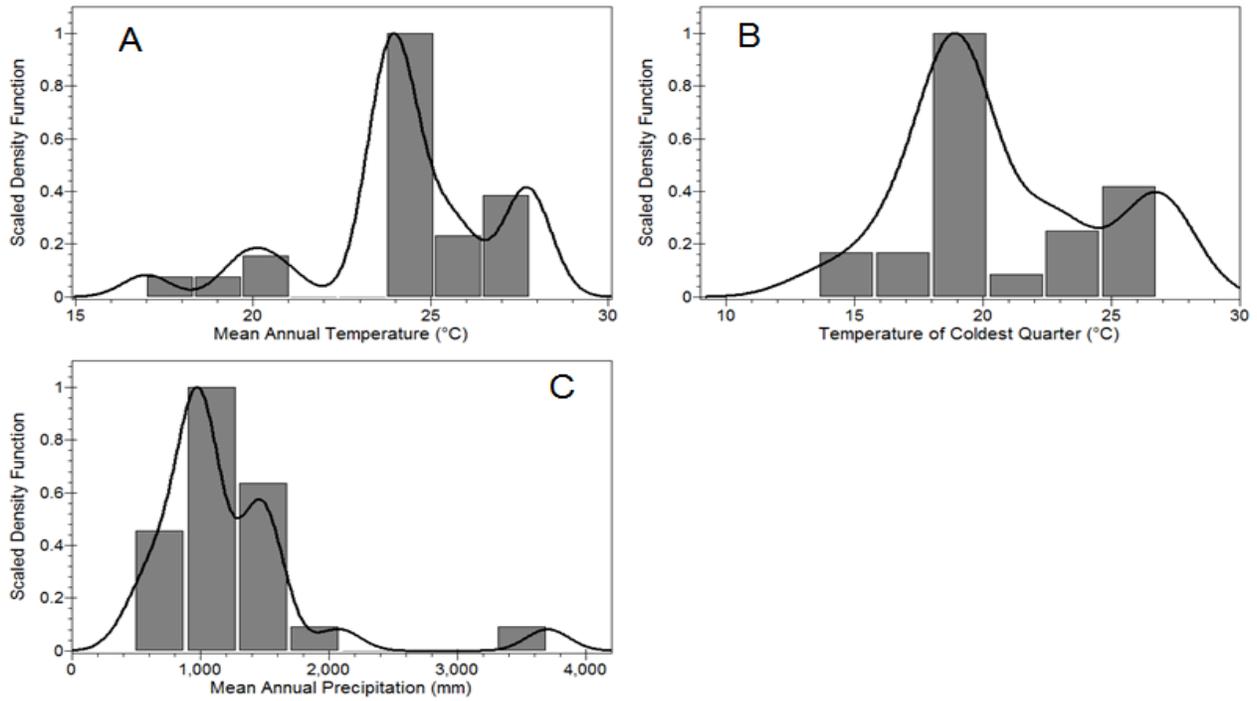


Fig 2: Kernel density estimates of mean annual temperature (A), temperature of coldest quarter (B) and mean annual precipitation (C). The lines represent the kernel density distribution and the bars are the normalized histogram values obtained from the train data.

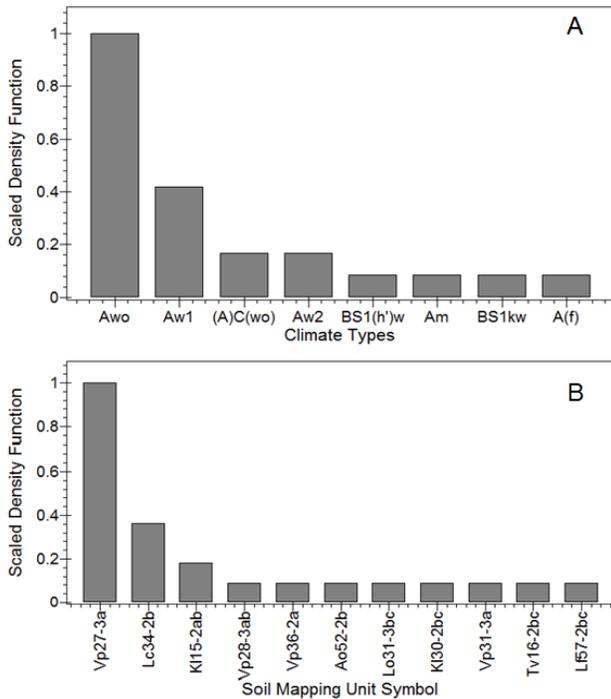


Fig 3: Discrete density estimates for climate (A) and soil layers (B). The bars are the standardized histogram values obtained from the train data.

Most of the acceptable values matched those of tropical climates; for example, temperature of coldest month is higher than 18 °C for A type climates [27] and fell within the observed range. Mean annual temperature synthesized temperature variations and reflected average conditions. Though the minimum threshold temperature is unknown for this species, many of the Coleopterans have a minimum value above 4 °C [28], and the estimated range pointed to a species with low tolerance to cold and, again, described an insect adapted to

tropical regions. Climate types where the insect was found corresponded mostly to tropical ones, as reported previously [1]. Few observations fell within a BS climate which is a dry, semi-arid type and may correspond to marginal areas of the species distribution. Biologically, soil is important because the adults overwinter in it [6], vertisol and luvisol types are mostly found in tropical grasslands or turned to agricultural fields [29]. Recently, kernel density models have been proposed to estimate species distributions [30], and previously, this modeling approach has been applied to extrapolate from point presence data to geographic aerial occurrence. In the current approach, kernel density estimation provided an alternative to other modeling techniques that use distributions such as those based on quantiles like the BIOCLIM or DOMAIN models [31]. The combination of layers required assigning values representing their importance. The weights were 0.636, 0.104 and 0.258 for mean annual temperature, precipitation and temperature of coldest quarter, respectively. Therefore, mean annual temperature was considered the most important because it comprised the whole range of values the insect experiment during its life cycle and because it is known that temperatures have a direct influence on insect development and survival [32]; temperature of coldest quarter was ranked second because it affects adult overwintering and survival [5]. Precipitation had the lower value as an indirect factor related to the development of the host plants. Climate and soils layers had weights of 0.75 and 0.25 respectively, climate was considered more important because it integrates different environmental factors. Climate and soil type formed the first composite category with weight 0.25, while mean annual temperature, precipitation and temperature of coldest quarter were grouped in a second category with weight 0.75. Therefore, the second composite category was considered more important than the first one because it contained more specific variables.

3.2. Potential distribution map

The final potential distribution map is presented in Fig. 4. The most suitable regions included areas of the northwestern states

like Sonora and Sinaloa and northeastern states like Tamaulipas, San Luis Potosi and northern Veracruz. Medium suitable regions scattered throughout the previous states, the Pacific coastal regions of Nayarit, Jalisco, up to Oaxaca, and parts of the Yucatan Peninsula. Low suitability regions spread from Yucatan peninsula to central Veracruz in the Gulf of

Mexico and from Chiapas up to Nayarit in the Pacific States. Unsuitable regions comprised most of the Baja California peninsula, and the central and northern states, like Chihuahua, Coahuila, Durango, Zacatecas, and Hidalgo. These areas corresponded mostly to dry, temperate climates.

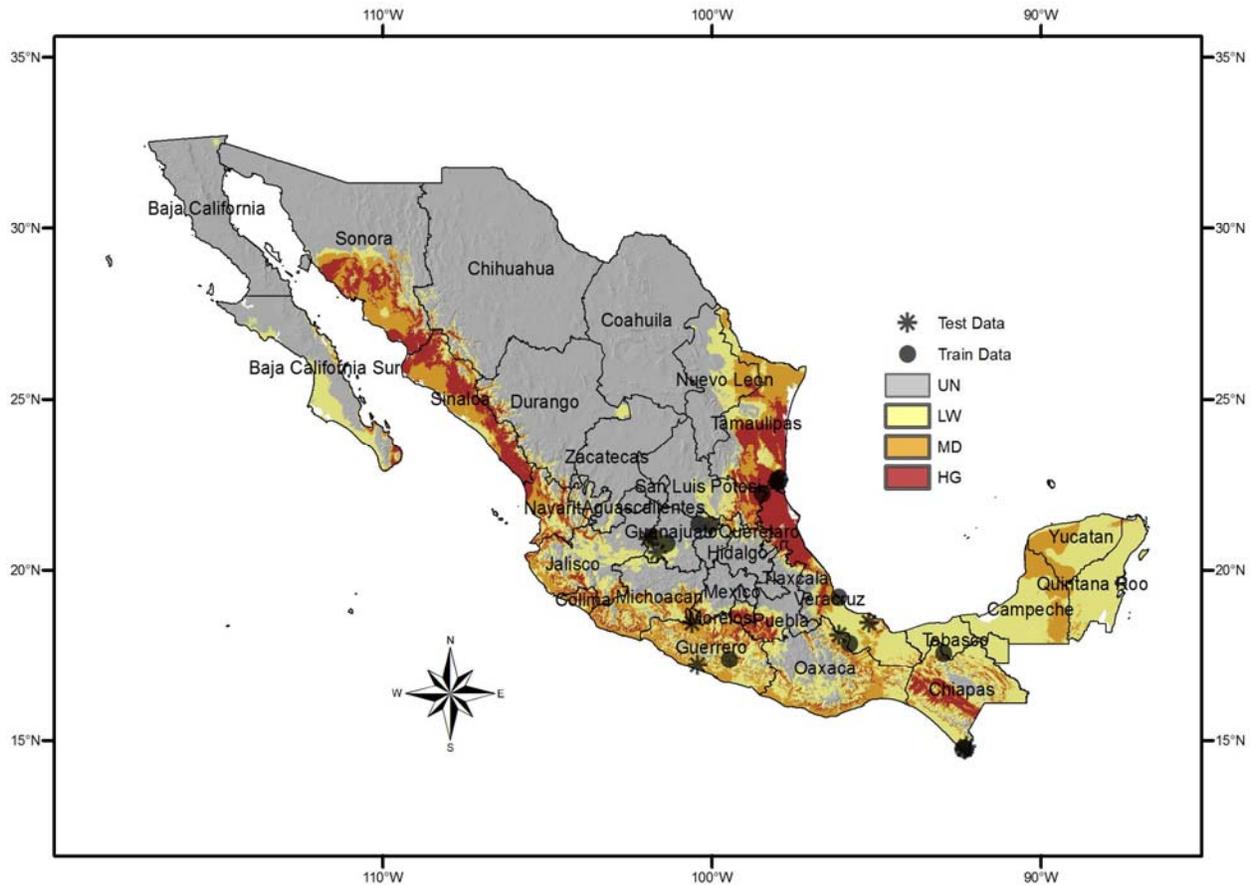


Fig 4: Potential distribution of *R. nigerrimus* in Mexico grouped in four classes. Lines correspond to administrative state boundaries. Circles represent train data and stars correspond to test data. Suitability classes are: unsuitable UN, low LW, medium MD and High HG.

The highly suitable regions corresponded mostly to states in the northern coastal states of Mexico and parts of Chiapas and Puebla and reflected the sampling effort of the species as a pest of soybeans. On the other hand, the previous reports of this species in wild hosts [2, 3], located individuals in the low and medium suitable regions, the same happened with insects reported in Guanajuato [4], which can be considered as a marginal region because of the low suitability values (Fig. 4). These results indicated that *R. nigerrimus* is mostly adapted to semitropical conditions but may reach arid areas which translated to the low suitable areas of the model. Worldwide, the findings agreed with reports relating this species to tropical locations in the Americas [1, 2].

3.3. Model evaluation

The ROC curve is presented in Fig. 5. The bootstrap estimation yielded an AUC value of 0.805 with a 95 % confidence interval that ranged from 0.625 to 0.875. The TSS index had the highest value of 0.625 at the lowest observed test point; this index was used instead of the kappa index because is considered to have better statistical properties [23]. Both indices showed a departure from randomness and implied a significant model [23, 33]. The indices supported the validity of the full model, which presented a medium suitability covering

most of the states of Tabasco and Nayarit, reported as regions for the presence of *R. nigerrimus* [5].

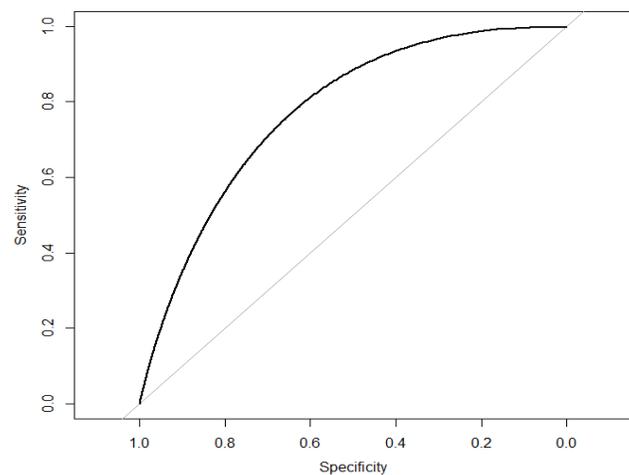


Fig 5: Bootstrap Receiver Operating Curve (ROC) for the test data of *R. nigerrimus*.

The validation process showed independent support for the full

model which used only five data layers. Other models use most detailed information, for example NAPPFAST takes into account threshold temperatures for insect development and computes the number of generations per year [34] while CLIMEX [35] uses weekly variables. However, previous work has suggested that bioclimatic variables like the ones used in this research, provide enough insight into the distribution of a species [36]. In the current model, soil and extreme cold temperatures accounted for basic biological aspects influenced by these factors, such as overwintering survival and location [6]. For a novel pest species like *R. nigerrimus*, it might be appropriate to establish the quantitative nature of temperature and development or to estimate stress thresholds to apply other detailed models. Because the suitability map correctly predicted most of the test data, it suggested a simple way to compute the potential geographic distribution of insect pests for which detailed quantitative information is lacking.

Regarding the sample size used to estimate the distribution of *R. nigerrimus*, previous work found that sizes as small as ten can yield valid estimates of the geographic distribution but it also depended on different factors such as species type, climatic requirements, sampling bias or modeling technique [37]. The current estimation was computed to obtain regions where the species may thrive given its new pest status and provides information for better planning, for example detection sampling [38]. Better estimation can be computed

with more point data; however, this may happen when the pest invades new areas, thus reducing the advantage of forecasting its distribution.

3.4. Economic impact on soybeans

Based on the total revenue and the potential occurrence, the vulnerability map is presented in Fig. 6. Because it weights revenue by suitability, high vulnerable areas reflected high production and high occurrence values, such as those municipalities of Tamaulipas, San Luis Potosí and Campeche. Intermediate vulnerable areas occurred across Sonora, Chiapas and Campeche. Low vulnerable regions spread from Chihuahua, northern Tamaulipas, Jalisco and Quintana Roo.

The economic impact was based on a simple product of suitability and revenue, however, more detailed analysis is required to better know the potential economic impact of this native species. Other factors affecting the economic impact are alternative crop hosts, spread rate and the dynamics of damage through time [39] and may need to be modeled by other methods. Because *R. nigerrimus* is native to America and Mexico, control measures such as eradication may not work; a related species, *R. subtilis* Fiedler is a pest of soybeans and beans in Argentina [40], therefore, it is apparent that the genus has the potential of causing damage to soybeans and related crops.

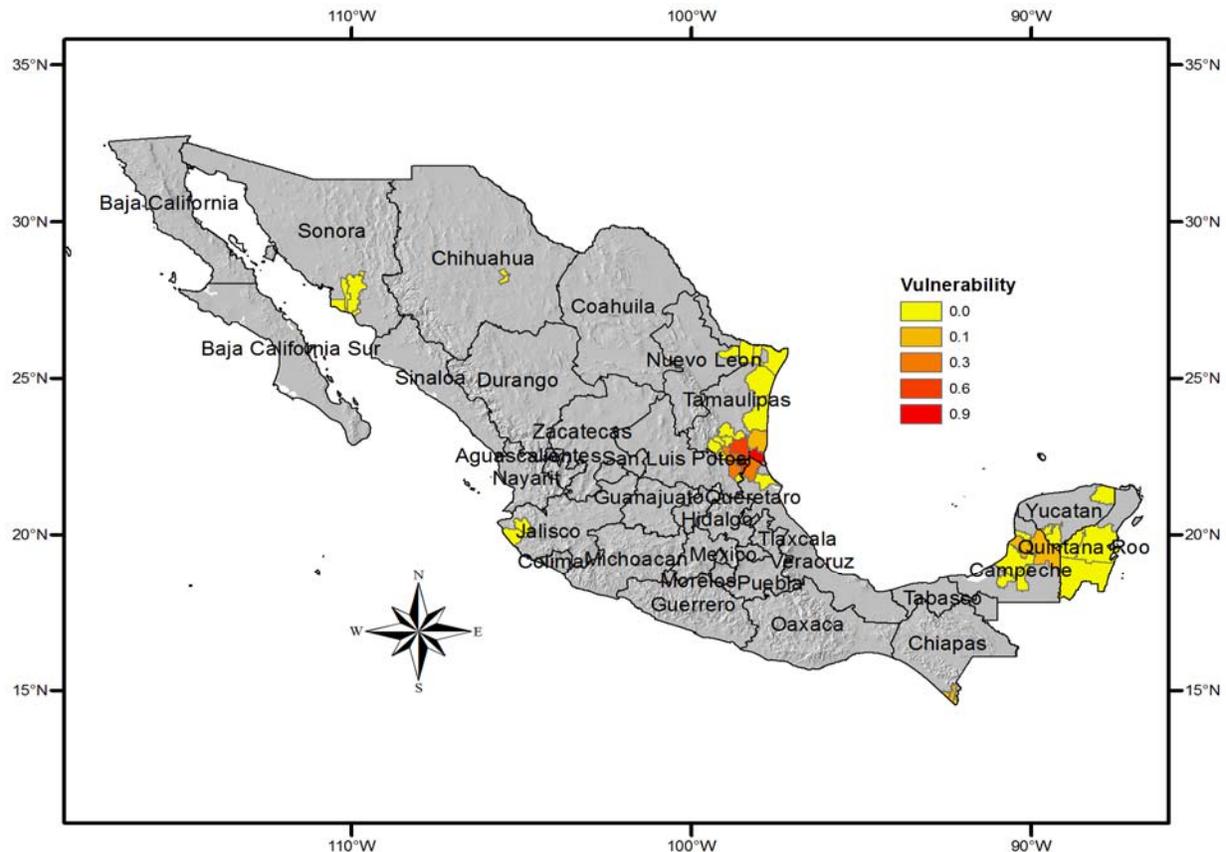


Fig 6: Expected vulnerability of soybean-planted municipalities to *R. nigerrimus*.

4. Conclusion

A hierarchic approach was used to combine data layers to estimate the geographic distribution of *R. nigerrimus* and its potential economic impact on soybeans in Mexico. It was found that *R. nigerrimus* had a medium to high potential distribution in some northern states of both Gulf of Mexico

and the Pacific Ocean. Different validation indices supported the distribution model capacity to represent the potential geographic distribution of this species. Soybean areas with medium to high risk of economic impact comprised some municipalities in the states of Tamaulipas, San Luis Potosí, Veracruz and Campeche.

5. References

- Peck S. The beetles of Barbados, West Indies (Insecta: Coleoptera): Diversity, distribution and faunal structure. *Insecta Mundi* 2009; 0073:1-51.
- Champion GC. Insecta. Coleoptera. Rhynchophora. Biologia Centrali-Americana, R.H. Porter Publisher, London 1902-1906; 4(IV):750.
- Morrone JJ, Muñiz R, Asiain J, Marquez J. Lista de las especies de Curculionioidea (Insecta: Coleoptera) depositadas en la colección del museo de zoología "Alfonso L. Herrera", Facultad de Ciencias, UNAM (MZFC). *Acta Zoologica Mexicana* 2002; 87:147-165.
- Salas-Araiza MD, O'Brien CW, Romero-Napoles J. Curculionioidea (Insecta: Coleoptera) from the state of Guanajuato, Mexico. *Insecta Mundi* 2001; 15:45-57.
- Lopez-Guillen G, Teran-Vargas AP, Gomez Ruiz J, San-Juan Lara J, Rosado-Neto GH, O'Brien CW *et al.* First Record of *Rhyssomatus nigerrimus* (Curculionidae: Molytinae: Cleogonini) infestations in soybeans in Mexico. *Florida Entomologist* 2012; 95:524-528.
- Lopez-Guillen G, Gomez-Ruiz J, Cruz-López L, Teran-Vargas AP. El picudo mexicano de la soya (*Rhyssomatus nigerrimus*) una plaga nueva del tropico. *Agroproductividad* 2012; 5:9-14.
- SIAP (Servicio de Información Agroalimentaria y Pesquera). Cierre de la producción agrícola por estado 2013. <http://www.siap.gob.mx/cierre-de-la-produccion-agricola-por-estado/>. December, 2014.
- Elith J, Leathwick JR. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics* 2009; 40:677-697.
- Franklin J, Miller JA. Mapping species distributions, Cambridge University Press, New York, 2009, 320.
- Araujo M, New M. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution* 2006; 22:42-47.
- Dupin M, Brunel S, Baker R, Eyre D, Makoski D. A comparison of methods for combining maps in pest risk assessment: application to *Diabrotica virgifera virgifera*. *EPPO Bulletin* 2011; 41:217-225.
- Eastman JR, Jin W, Kyem PAK, Toledano J. Raster procedures for multi-criteria/multi-objective decisions. *Photogrammetric Engineering and Remote Sensing* 1995; 61:539-547.
- Saaty T. How to make a decision: The analytic hierarchy process. *European Journal of Operation Research*. 1990; 48:9-26.
- Kihoro J, Bosco NJ, Murage H. Suitability analysis for rice growing sites using a multicriteria evaluation and GIS approach in great Mwea region, Kenya. *SpringerPlus* 2013; 2:265.
- Krist FJ Jr, Sapio FJ, Tkacz BM. A multicriteria framework for producing local, regional, and national insect and disease risk maps. In: Pye, Rauscher JMHM, Sands Y, Lee DC, Beatty JS. (eds.). *Advances in threat assessment and their application to forest and rangeland management*, General Technical Report PNW-GTR-802, U.S. Forest Service, Pacific Northwest and Southern Research Stations, Portland, 2010, 621-636.
- Vasiljevic TZ, Srdjevic Z, Bajcetic R, Miloradov MV. GIS and the analytical hierarchy process for regional landfill site selection in transitional countries: A case study from Serbia. *Environmental Management* 2012; 49:445-458.
- Hossain MS, Das NG. GIS-based multi-criteria evaluation to land suitability modelling for giant prawn (*Macrobrachium rosenbergii*) farming in Companigonj Upazila of Noakhali, Bangladesh. *Computers and Electronics in Agriculture* 2010; 70:172-186.
- Fernandez-Eguiarte A, Zavala-Hidalgo J, Romero-Centeno R. Atlas Climatico Digital de Mexico. Centro de Ciencias de la Atmosfera. UNAM. <http://www.atmosfera.unam.mx/uniatmos/atlas/kml/>, August, 2013.
- CONABIO (Consejo Nacional de Biodiversidad). Portal de GeoInformacion. Sistema Nacional de Informacion sobre Biodiversidad. <http://www.conabio.gob.mx/informacion/gis>, October, 2013.
- FAO (Food and Agriculture Organization). Digital soil map of the world. <http://www.fao.org/geonetwork/srv/en/metadata.show?id=14116>, September, 2013.
- Saaty TL. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*. 1977; 15:234-281.
- Elith, J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 2006; 2:129-151.
- Allouche O, Tsoar A, Kadmon R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*. 2006; 43:1223-1232.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacekm F, Sanchez JC, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; 12:77.
- R Core Team. R: A language and environment for statistical computing, Vienna, Austria, 2014.
- Wolfram Research. Mathematica 8.01, Wolfram Research, Champaign, Illinois, 2011.
- Peel MC, Finlayson BL, McMahon TA. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* 2007; 11:1633-1644.
- Nietschke BS, Magarey RD, Borchert DM, Calvin DD, Jones E. A development database to support insect phenology models. *Crop Protection* 2007; 26:1444-1448.
- Geissen V, Sanchez-Hernandez R, Kampichler C, Ramos-Reyes R, Sepulveda-Lozada A, Ochoa-Goana S, *et al.* Effects of land-use change on some properties of tropical soils- An example from southeast Mexico. *Geoderma* 2009; 151:87-97.
- Blonder B, Lamanna C, Violle C, Enquist BJ. The n-dimensional hypervolume. *Global Ecology and Biogeography*, 2014. DOI: 10.1111/geb.12146.
- Carpenter G, Gillison AN, Winter J. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation* 1993; 2:667-680.
- Shi P, Ge F. A comparison of different thermal performance functions describing temperature-dependent development rates. *Journal of Thermal Biology*. 2010; 35:225-231.
- Landis JR, Koch GG. The measurement of observed agreement for categorical data. *Biometrics* 1977; 33:159-174.
- Nietschke BS, Borchert DM, Magarey R, Ciomperlik MA. Climatological potential for *Scirtothrips dorsalis* (Thysanoptera: Thripidae) establishment in the United

- States. Florida Entomologist 2008; 91:79-86.
35. Sutherst RW, Maywald GF. A computerized system for matching climates in ecology. Agriculture, Ecosystems and Environment 1985; 13:281-299.
 36. Watling J, Romañach SS, Bucklin DN, Speroterra C, Brandt LA, Pearlstine LG, *et al.* Do bioclimatic variables improve performance of climate envelope models? Ecological Modelling 2012; 246:79-85.
 37. Hernandez PA, Graham CH, Master LL, Albert DL. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 2006; 29:773-785.
 38. Berc L, Kean JM, Epanchin-Niell R, Liebhold AM, Haight RG. Designing efficient surveys: spatial arrangement of sample points for detection of invasive species. Biological Invasions 2015; 17:445-459.
 39. Cook DC, Liu S, Edwards J, Villalta ON, Aurambout JP, Kriticos D, *et al.* Predicted economic impact of black sigatoka on the Australian banana industry. Crop Protection 2013; 51:48-56.
 40. Cazado LE, Murua MG, Casmuz A, Socias MG, Vera MT, O'Brien CW, *et al.* Geographic distribution and new host associations of *Rhyssomatus subtilis* (Coleoptera: Curculionidae) northwestern in Argentina. Florida Entomologist 2013; 96:663-669.