Does Land Cover Influence the Spatial Distribution of Reservoir Rodent Necromys Lasiurus?

Janduhy Pereira dos Santos1, Stefan Vilges de Oliveira2*, Marco Tulio Antônio Garcia-Zapata3 and Valdir Adilson Steinke1

1Department of Geography of the University of Brasilia, Brasilia, Distrito Federal, Brazil
2Postgraduate Program in Tropical Medicine, University of Brasilia, Distrito Federal, Brazil
3Institute of Tropical Pathology and Public Health, Federal University of Goias, Goiania, Brazil

Received: 01 July, 2017; Accepted: 08 August, 2017; Published: 14 August, 2017
*Corresponding author: Janduhy Pereira dos Santos, Department of Geography of the University of Brasilia, Brasilia, Distrito Federal, Brazil, Email: janduhy.santos@gmail.com

Abstract
The rodent Necromys lasiurus (Rodentia, Sigmodontinae) is considered to be the main hantavirus reservoir in the Brazilian Cerrado biome. Significant changes to the landscape of this biome have occurred in recent decades. Considering the fact that change in land use patterns can influence species distribution and, as a result, the emergence of diseases, this study sought to analyze the potential geographic distribution of the species N. lasiurus and to understand how land use and occupation contribute to the distribution of this species. Using Ecological Niche Modeling (ENM), environmental suitability models were constructed for N. lasiurus. The Maxent algorithm was used with predictor variables indicated by the results of the jackknife technique. Test accuracy was evaluated using the area under the curve (AUC). The differences in the potential distribution models resulting from the inclusion or exclusion of land use as a variable were not random (p < 0.0001). The jackknife method indicated that the variables referred to as percentage of vegetation cover and land use are the strongest predictive models (38.9% and 37.9%, respectively). The AUC evaluation indicated good performance (0.93 and 0.92). When inserted into the model, the variables land use and vegetation cover were found to significantly influence potential N. lasiurus distribution. These variables should be considered when predicting the occurrence of this medically important rodent.

Keywords: Ecological niche modeling; Ecoepidemiology; Hantavirus; Rodent-borne disease; Maxent

Introduction
The genus hantavirus persists in the wild because rodents act as reservoirs for this zoonosis. In Brazil, human infections by etiological agents of this genus are also referred to as hantaviruses [1]. Human infections caused by hantaviruses occur mainly through the inhalation of airborne viral particles, which are formed from excreta of infected rodents. The infection appears in two different clinical forms: Hemorrhagic Fever with Renal Syndrome (HFRS) and Hantavirus Pulmonary Syndrome (HPS) [2].

HPS is an emerging disease which has thus far been the most prevalent form in the Americas [3]. In Brazil, the first occurrence of HPS was reported in 1993 in the city of Júquituba, in the state of São Paulo [4]. Since then, knowledge of the disease has been growing parallel to its range in Brazil; the prevalence of the disease has been established in all regions of the country [5,6]. HPS is a notifiable disease for the Brazilian Ministry of Health and its importance for public health is due to the high case-fatality rates (~ 40%) and limited knowledge about its epidemiology [1,2]. The main risk factors for illness are the exercise of agricultural activities, domestic activities associated with exposure to rodents and houses near forest remnants where wild rodents occur [2].

Necromys lasiurus (Lund 1840) is reservoirs of hantavirus in Brazil. Is an omnivorous and terrestrial species which feeds mainly on seeds [6]. Is widely distributed within the Brazilian Cerrado and Caatinga biomes, areas where human land use and modification of the natural environment are widespread. Most HPS cases in Brazil occurred in areas with the presence of this rodent. Can be found in many habitats, but prefer open and dry areas, being absent or infrequent in moist forest environments [6]. A number of studies have indicated that regional distributions of host and pathogens are associated with climatic and environmental factors [6-10].

Changes in the landscape of the Cerrado have largely been caused by the newly permitted use of land for grazing beef cattle and, more recently, for large-scale agriculture (the soybean in particular). Current estimates indicate that, every year, approximately 3 million hectares of natural habitat are lost from an original area of 2.045 million square kilometers [11]. These patterns of land use have been identified as key drivers of habitat loss, and it is essential to consider such factors when conducting predictive modeling of the distribution of zoonotic reservoir species [12]. However, there are currently few modeling studies that have explicitly emphasized changes in land use and their impacts on habitat and emerging diseases [13]. Given these factors, the current study sought to analyze the potential geographical distribution of N. lasiurus while trying to understand how land use and occupation influence the predictive distribution models of this hantavirus reservoir species in the Brazilian Cerrado.

Materials and Methods

Data on the Occurrence of N. Lasiurus
Data on the occurrence of N. lasiurus was obtained from SpeciesLink, a system that integrates data from biological collections a variety of Brazilian museums and research centers (Appendix 1). These institutions provide free access to the biological information from their scientific collections. In addition, a search of the literature on N. lasiurus was performed [6,14-16].

Distributional data was collected and converted into geographic coordinates (latitude / longitude) in Decimal Degrees format (DD) using the WGS-84 datum surface. Occurrence points for N. lasiurus were
Does Land Cover Influence the Spatial Distribution of Reservoir Rodent Necromys Lasius?

The Maxent software was used because it is considered the most appropriate when a data set is comprised of only occurrence data [21]. When Maxent is used to model species distribution, it estimates the probability of species occurrence and the distribution of probability of maximum entropy. The results are subject to a set of constraints that represent the incompleteness of the information on the target distribution. This process is known as the principle of maximum entropy [20].

The principle of maximum entropy is expressed as a finite set X (which will later be interpreted as a set of pixels in the study area). Individual elements are referred to as X points. The distribution π assigns a non-negative probability π(x) for each point x, and the sum of these probabilities is 1. Approximation to π is also a probability distribution, which is denoted with π. The entropy of π is defined as (Figure 1):

\[ H(\hat{\pi}) = -\sum_{x \in X} \hat{\pi}(x) \ln \hat{\pi}(x) \]

Where ln is a natural logarithm. The entropy is not negative and is not at maximum, and the natural logarithm is the number of elements in X. The Maxent algorithm then computes the probability distribution over the study area (background) based on the distribution of the environmental variables. For each pixel, Maxent indicates a numerical value ranging from 0-1. This indicates environmental suitability rather than probability of occurrence [12]. In this study, the Maxent software version 3.3.3 was used and operated under the default program settings.

Environmental and Anthropogenic Data

The environmental layers were selected according to a Pearson correlation analysis. The data used in the modeling of potential N. lasiurus distributions, which was obtained from the Brazilian National Institute for Space Research, the Food and Agriculture Organization of the United Nations and the Oak Ridge National Laboratory, is provided in Table 1 [17-19].

<table>
<thead>
<tr>
<th>Type</th>
<th>Spatial resolution</th>
<th>Temporal resolution</th>
<th>Fonte</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topography</td>
<td>Elevation</td>
<td>90 m</td>
<td>*INPE</td>
</tr>
<tr>
<td>Bio1</td>
<td>Average annual temperature</td>
<td>1 km 1950-2000</td>
<td>INPE</td>
</tr>
<tr>
<td>Bio2</td>
<td>Diurnal variation average temperature</td>
<td>1 km 1950-2000</td>
<td>INPE</td>
</tr>
<tr>
<td>Bio3</td>
<td>Isoterm</td>
<td>1 km 1950-2000</td>
<td>INPE</td>
</tr>
<tr>
<td>Bio12</td>
<td>Annual rainfall</td>
<td>1 km 1950-2000</td>
<td>INPE</td>
</tr>
<tr>
<td>Bio14</td>
<td>Driest month precipitation</td>
<td>1 km 1950-2000</td>
<td>INPE</td>
</tr>
<tr>
<td>Percentage of herbaceous cover</td>
<td>Vegetation</td>
<td>500 m 2001</td>
<td>**ORNL</td>
</tr>
<tr>
<td>Land use</td>
<td>Use and land cover</td>
<td>1:5,000,000 2008</td>
<td>***FAO</td>
</tr>
</tbody>
</table>

Source:
*National Institute for Space Research– INPE - Brazil
**United Nations Food and Agriculture Organization– FAO - Italy
***Oak Ridge National Laboratory– USA

The selected bioclimatic variables in Table 1 (Bio 1, Bio 2, Bio 3, Bio 12 and Bio 14) are parameters corresponding to observation data, cover the period 1950 to 2000, and were given the operational resolution of 30 arc-seconds (1km). Temperature values are given in °C, and precipitation values in mm (millimeters). The information on topography (elevation) comes from SRTM data (Shuttle Radar Topographic Mission) at a horizontal resolution (spatial resolution) of 1km, and a vertical resolution (height) of 1 m. The layer regarding the percentage of herbaceous cover is from the MODIS sensor containing the estimate of herbaceous vegetation coverage. The raster has a spatial resolution of 500 meters with a time resolution of 2000-2010. Raster use and land cover was designed in 2010 with scale of 1: 5,000,000 with Datum WGS-84. Data were resampled (nearest neighbor) to allow maximum base sample compatibility. This allowed the Maxent algorithm to be used.

The layer land use has a spatial resolution of 5 arc minutes or 0.083333 decimal degrees with scale of 1: 5,000,000 with Datum WGS-84. Land use classes present in the modeling are: urban, bare, water, cropland, grassland, rangeland and forestry land.

Ecological Niche Modeling

The Ecological Niche Model (ENM) was used was generated using the Maxent algorithm (maximum entropy). The objective was to model potential geographic species distribution. To develop the ecological niche model algorithms and to find nonrandom relations, N. lasiurus occurrence data and environmental and anthropogenic data mentioned in the previous steps of the methodology were entered into the model [20].

The Maxent software was used because it is considered the most appropriate when a data set is comprised of only occurrence data [21]. When Maxent is used to model species distribution, it estimates the probability of species occurrence and the distribution of probability of maximum entropy. The results are subject to a set of constraints that represent the incompleteness of the information on the target distribution. This process is known as the principle of maximum entropy [20].

The principle of maximum entropy is expressed as a finite set X (which will later be interpreted as a set of pixels in the study area). Individual elements are referred to as X points. The distribution π assigns a non-negative probability π(x) for each point x, and the sum of these probabilities is 1. Approximation to π is also a probability distribution, which is denoted with π. The entropy of π is defined as (Figure 1):

\[ H(\hat{\pi}) = -\sum_{x \in X} \hat{\pi}(x) \ln \hat{\pi}(x) \]

Where ln is a natural logarithm. The entropy is not negative and is not at maximum, and the natural logarithm is the number of elements in X. The Maxent algorithm then computes the probability distribution over the study area (background) based on the distribution of the environmental variables. For each pixel, Maxent indicates a numerical value ranging from 0-1. This indicates environmental suitability rather than probability of occurrence [12]. In this study, the Maxent software version 3.3.3 was used and operated under the default program settings.

Assessment of the Models Generated

Receiver Operating Characteristic (ROC) curves were used for the statistical assessment of the species distribution models [20]. Calculating the Area Under the Curve (AUC) provides a measure of model performance that is independent of the previous choice of the specified threshold, because the curve is constructed from a series of thresholds (with each threshold representing one point on the curve). Calculating the area under the curve therefore eliminates the need to choose a specific threshold.

An area of 1 would represent the “perfect” model; an area of 0.5 indicates that the model was selected at random. From a practical point of view, a validation test can adopt the following AUC values as indicators of model quality: excellent (0.90 - 1.0); good (0.80 - 0.90); average (0.70 - 0.80); poor (0.60 - 0.70); or very poor (0.50 - 0.60) [22].

To identify which variables had the greatest influence on the distribution of the rodent species, a jackknife test was run in Maxent [20]. This test measures the predictive effects of each variable in the model, and, in doing so, determines the quality of the models produced only with the variables being tested; the models themselves are omitted.

Results

After the unreliable records, the duplications, and the points in areas with multiple samplings were all excluded, 64 spatially unique records remained for use in the modeling process. ENM revealed the areas that possess suitable environments for N. lasiurus both with and without the inclusion of the land use variable (Figures 2 and 3). Figure 2 shows the model for N. lasiurus with the inclusion of the land use variable. The results show potential distribution in areas of slope forest and semi-deciduous forest on the Atlantic plateau, as well as in the transition zone between the Atlantic forest-savanna-Cerrado and in some areas of the Amazon (Central-West Region and Southeast Region).

Figure 2 shows areas with high suitability, beginning northeast of Roraima State where savanna-type vegetation exists. The map also shows suitability in the region surrounding the city of Novo Airão in the state of Roraima State where savanna-type vegetation exists. The map also shows suitability in the region surrounding the city of Novo Airão in the state of Roraima State where savanna-type vegetation exists.
grassy vegetation with shrubs and small trees. There are additional areas of very suitable habitat in southwestern Rondônia State and in central Mato Grosso State, specifically on the Parecis Plateau and on the outskirts of the city of Cuiabá.

In Northeastern Brazil, there is a strong potential for occupancy at higher altitudes areas, including the mountains of Ibiapaba and Baturité and the Cariri Plateau on the border between the states of Ceará, Pernambuco, and Paraiba, as well as in certain areas of Rio Grande do Norte State, mainly in the regions of Seridó and Borborema. In addition, areas favorable to the species were found near the São Francisco River and in the highland Cerrado regions in the states of Goiás, Minas Gerais and São Paulo.

The model also infers a strong likelihood of the species being present in the Mantiqueira Mountains and the Serra do Mar mountain range in Southeastern and Southern Brazil, as well as in areas originally covered by the Atlantic Forest in the coastal states of Parana, Santa Catarina, and Rio Grande do Sul.

As an overview, Figure 3 shows the model without the insertion of the land use and land cover layers. The model indicates that areas with strong environmental suitability occur in the lower Amazon basin near Santarem, Para. In the Brazilian Northeast, potential areas were found to be located in the mountains of Ibiapaba, Baturité and Arraipre in Ceara State, as well as in the region known as Gariri Paraibano in rural Pernambuco State, on the coast of Sergipe, and also in the northern-central region of Bahia State. In South-Central Brazil, areas with high suitability for *N. lasiurus* were detected in the area surrounding Brazil’s capital district (the Distrito Federal), in the city of Goiania, Goias, in the planning regions known as the Triângulo Mineiro and Alto Paraíba, and also in areas of the Espinhaco Mountains to the Sao Francisco River valley in Minas Gerais State.

There is also significant potential for *N. lasiurus* occupancy in northwestern São Paulo State, where the Cerrado biome has been substantially modified by sugarcane plantations. There is also potential for distribution in the coastal regions of the states of Espírito Santo, Rio de Janeiro, Santa Catarina, and Rio Grande do Sul, where there is a mixed vegetation complex that includes areas of Atlantic Forest, mangroves, and salt marsh vegetation.

Another area highlighted and indicated by ENM as having a very strong probability of *N. lasiurus* occurrence (Figures 2 and 3) was the northwestern region of Roraima State, which is best known for its unique Lavrado ecosystem: a savanna area characterized by open vegetation covering very poor white sand soil. Large variations in altitude and rainfall result in a mosaic of vegetation types, including hills covered by extensive grasslands, narrow gallery forests, and rocky areas dominated by cactus [23]. Given this combination of floral characteristics and edaphic and pedological features, this unique ecosystem is very similar to the savanna of the Cerrado biome.

Table 1 shows the statistical results of the Maxent algorithm, including the Minimum Training Presence (MPT) threshold values, which represent the cut-off points chosen to examine areas with good environmental suitability for *N. lasiurus*. The table also presents the results of the inclusion and exclusion of the land use variable in the habitat quality modeling process.

The jackknife test revealed that percentage of herbaceous cover accounted for 36.9% of modeling variation, while the land use variable accounted for 37.9% (Table 3). Both had the same weight in the model. When the model was created without the land use variable, the jackknife result for the percentage of herbaceous cover layer increased to over 60%. Additionally, applying the jackknife technique and omitting the percentage of occurrences of the species that lays outside the predicted area also indicated this model. The results found in the absence of the land use layer were much more reliable than those produced when this variable was included.

The performance evaluation, which used the Area Under the Curve (AUC) for the occurrence points used to perform a trial with the data, found that the values derived from the ROC curves indicated that modeling with and without the variable land use both exhibited both a good performance (Table 2). Table 3 shows estimates of the contributions of the variables used in each model. These estimates were calculated using the jackknife technique on the points used in predictive analysis for *N. lasiurus* distribution.

The jackknife test revealed that percentage of herbaceous cover accounted for 36.9% of modeling variation, while the land use variable accounted for 37.9% (Table 3). Both had the same weight in the model. When the model was created without the land use variable, the jackknife result for the percentage of herbaceous cover layer increased to over 60%. Additionally, applying the jackknife technique and omitting the percentage of herbaceous cover variable was found to cause a loss of gain in both models.

**Discussion**

This study adds information about the ecoepidemiology of the main reservoir of hantavirus in the Brazilian Cerrado biome. Knowledge of the areas of occurrence as well as the factors that determine their distribution...
areas may be useful for the epidemiological surveillance services of the disease in Brazil. Based on the descriptive and quantitative information in the predictive model, the insertion of the land use layer into the model suggests that this variable has a significant influence on potential distribution of *N. lasiurus*. Though the jackknife analysis showed that the percent vegetation cover layer made a greater contribution to the models, and though the anthropogenic layer, land use, was found to exert an influence on the prediction of species occurrence, it is possible to see very marked visual differences (Figures 2 and 3) when the latter layer was included or excluded.

When the land use variable was included, a decrease in predictive pixels was observed. It can be inferred that this decrease in predictive modeling was a result of environmental changes caused by human activity, such as farming and urban expansion into abandoned agricultural regions or even into regions which still harbor native forest (Table 2). This finding differed from when this variable was excluded, in which case an increase in predictive pixels was observed.

<table>
<thead>
<tr>
<th>Variable (Land use)</th>
<th>Cut limit value</th>
<th>Predicted Area fraction</th>
<th>Omission rate</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>0.075</td>
<td>0.504</td>
<td>0.1</td>
<td>0.93</td>
</tr>
<tr>
<td>Absent</td>
<td>0.049</td>
<td>0.678</td>
<td>0.033</td>
<td>0.92</td>
</tr>
</tbody>
</table>

These results are similar to the analyses conducted in China by Wei, et al. and Yan, et al. [24,25]. This consistency indicates that the anthropogenic land use variable has a significant impact on the distribution of the target species. Indeed, environmental impacts of anthropogenic origin such as land use changes are well known for leading to biodiversity loss. *N. lasiurus* is a generalist species in terms of habitat use, but it appears to prefer open areas since it is absent or infrequent in humid forest environments. This factor suggests that the expansion of deforested areas over the past few years could explain the higher occurrence of this species in the Atlantic Forest region, a conclusion which reinforces the importance of analyzing the human-created habitats accepted by this rodent.

The distribution of *N. lasiurus* throughout Brazil that was predicted by the two models is compatible with the distribution modeled by Oliveira, et al. [6]. These authors also found areas of potential distribution in South-Central Brazil and in regions further north and northeast. Their results suggest that there is also the potential for hantavirus transmission in other parts of the country.

These conclusions are consistent with those reported by Lima, et al. who found hantavirus antibodies in asymptomatic patients from Gurai State, even though there were no records of the disease there at the time [26]. The authors also mention the capture of *N. lasiurus* with the same seroreactive in the Ibiapaba Mountains, located in the same state. This site is highlighted in both cartograms as a predictive area for the rodent (Figures 2 and 3).

In the dry season, the seeds of the grass genus *Brachiaria* are important sources of both energy and water for this species, such that the presence of *N. lasiurus* has been linked to that of this plant. Studies of hantavirus infection in humans living in the Brazilian Cerrado have found that most cases occurred in houses close to *Brachiaria* fields, a result which suggests a strong association between *N. lasiurus* and this type of vegetation this open areas [27].

The potential coastal distribution of *N. lasiurus* demonstrates that the possibility of the rodent’s occurrence is due to the presence of *Brachiaria* spp. and other exotic grasses grown to establish pastures in the region, which provide shelter and food for open-area rodents in these formerly forested areas.

The use of the land use variable in the models shows that anthropogenic layers used in models of potential distribution place significant weight on model outcome. It is clear that the use of this variable is key in understanding the current distribution of the species investigated herein, since environmental changes represented by the land use variable are responsible for recent changes in floral and faunal biodiversity.

The results of the current study indicate that the ecological model of the potential distribution of wild animals involved in infectious disease transmission cycles provide supplemental information for studies on disease transmission. Together, these studies can generate information on hotspots that may eventually be the center of future epidemics.

It is important to recognize that there are challenges to obtaining appropriate data for environmental and anthropic variables to perform these predictive studies. The scale of the information used must be compatible with the species distribution. Therefore, it is necessary to monitor the Receiver Operating Characteristic (ROC) curves in order to avoid potential overestimation of the distribution of the species under study. This step is vital because, if overestimation occurs, it can cause a series of uncertainties and may invalidate the entire model.

We also emphasize the importance of knowing the extent to which these rodents participate in the epidemiological cycle of this disease. The models created herein indicate a strong potential of *N. lasiurus* occurrence in the Northern and Northeastern regions of Brazil. These areas are poorly explored and lack ecogeographical studies to know what species of rodents occur and the prevalence of infection in rodents [28].

It is also important to exercise caution when interpreting the results. More eco-physiological studies on *N. lasiurus* are needed to develop an accurate understanding of which of the potential range areas this rodent actually inhabits. Furthermore, a species movement analysis focusing on the dispersibility of the species would provide information on which locations are accessible to it. It is need expand our knowledge on enzootic cycles of the hantavirus strains present in Brazil. To do so, more eco-epidemiological and spatial analyses are needed. They will provide information that will aid in the understanding the complexity of hantavirus and ultimately, in the development of local epidemiological surveillance systems.

This study showed that the variables land use and vegetation cover were found to significantly influence potential *N. lasiurus* distribution. These variables should be considered when predicting the occurrence of this medically important rodent.

**Appendix 1**

Biological collections and research centers where the information of rodents *N. lasiurus* was accessed. The Mammal Collection from Mato Grosso State University’s Novo Xavantina Campus (UNEMAT) (UNEMAT-CM); The Mammal Collection from the Federal University of Mato Grosso
Does Land Cover Influence the Spatial Distribution of Reservoir Rodent Necromys Lasiurus?

(UFMT); the Mammmology Collection from the Federal University of Parana (UFPR) (DZUP-Mamillia); the Reference Section of the Zoological Collection for Arthropod-Transmitted Viruses from the Aldo Lutz Institute (IAL) in Sao Paulo (IAL-Roedores), the Mammal Database from the Federal University of Espirito Santo (UFES) (Mamferos-ES); The Mammal Collection from the Capao do Imbiua Natural History Museum in Curitiuba, Parana (MHNCI); the Mammal Collection from the Zoology Program of the Federal University of Parana (UFPR) (MZUEL); the Biota Program Information System from the Sao Paulo Research Foundation (FAPESP) (SinBiota); and the Mammal Collection from the State University of Campinas (UNICAMP) Zoology Museum (ZUEC-MEM).

Acknowledgments

We would like to thank the staff of the Health Surveillance Secretariat of the Ministry of Health the Brazil. We would also like to thank Adrian Paul Ashton Barnett for reviewing the scientific content of the article and Danielle Jacqueline Deremo Cosimo for revising the English.

Author Contributions:

J.P.S., V.A.S. and S.V.O designed the experiments; J.P.S. performed the experiments; J.P.S., M.T.A.G.Z., V.A.S and S.V.O analysed the data; J.P.S. and S.V.O wrote the paper and all authors reviewed the final draft.

References
