South American Journal of Herpetology, 4(2), 2009, 103-116

PREDICTING *LEPTODACTYLUS* (AMPHIBIA, ANURA, LEPTODACTYLIDAE) DISTRIBUTIONS: BROAD-RANGING VERSUS PATCHILY DISTRIBUTED SPECIES USING A PRESENCE-ONLY ENVIRONMENTAL NICHE MODELING TECHNIQUE

Miguel Fernández[1,2], Daniel Cole[3], W. Ronald Heyer[4,7], Steffen Reichle[5], and Rafael O. de Sá[6]

[1] *Center for Biodiversity and Research, California Academy of Sciences, 55 Music Concourse Drive, Golden Gate Park, San Francisco CA 94118, USA.*
[2] *Environmental Systems Graduate Group, School of Engineering, University of California, Merced, PO Box 2039, Merced, CA 95344, USA.*
[3] *Information Technology Office, National Museum of Natural History, MRC 136, PO Box 37012, Smithsonian Institution, Washington, DC 20013-7012, USA.*
[4] *Amphibians and Reptiles, National Museum of Natural History, MRC 162, PO Box 37012, Smithsonian Institution, Washington, DC 20013-7012, USA.*
[5] *The Nature Conservancy, Southern Andes Conservation Program, Avenida 2 de Agosto/4. Anillo, Bolfor Office, PO Box 6204, Santa Cruz, Bolivia.*
[6] *Rafael O. de Sá, Department of Biology, University of Richmond, Richmond, Virginia 23173, USA.*
[7] *Corresponding Author: heyerr@si.edu*

Abstract. Locality data available for many, if not most, species of Neotropical frogs are based on written descriptions of the collecting sites, not on GPS device determined coordinate data. The pre-GPS device data are imprecise relative to GPS data. Niche modeling is a powerful technique for predicting geographic distributions that provides the best results when the locality data are precise. The purpose of this study is to determine whether imprecise historical locality data are sufficient such that niche modeling techniques can yield realistic new insights to species-level distributions. Two sets of frogs of the genus *Leptodactylus* that have known different kinds of distributions are evaluated: two species with broad, presumably continuous distributions, and four species known to occur in patchy, disjunct habitats in South America. BIOCLIM, a presence-only environmental niche modeling algorithm,was used to define suitable occupancy areas based on multiple sets of environmental parameters that include: monthly mean, max, and min temperatures, and monthly precipitation. A Nature Conservancy – Natureserve ecoregion layer and a high resolution elevation layer were also included in the analyses. Our analyses yield new realistic insights and questions regarding distributions of the *Leptodactylus* species we evaluated. We recommend incorporation of the Nature Conservancy – Natureserve layer to evaluate Neotropical distributions, as the layer gave much more robust results than use of only the climatic variable analyses.

Keywords. Environmental niche modeling, locality data precision, *Leptodactylus.*

Introduction

Until recently, distributions of most organisms, including frogs of the genus *Leptodactylus,* were determined by plotting and connecting peripheral localities on a map, tacitly assuming that species are uniformly distributed in space (*e.g.,* Berra, 1981; Hall, 1981; Howell and Webb, 1995; Erwin, 1996). The locality data were drawn from publications and museum specimens. These data typically were in the form of written descriptions rather than precise geographical coordinates. Even when coordinate data were associated with locality data, the coordinates were often determined *"a posteriori"; i.e.,* in the laboratory from maps at hand and guesstimating the coordinates based on the written description of the locality. At broad scales, such as a South American map of 1:10,700,000, plotting coordinates that were retrospectively georeferenced would almost always cover the locality involved, because the dot on the map would cover a rather extensive area. Even when using maps at a scale of 1:1,000,000 to determine coordinates from written locality data, such as *"Brasil: Pará; Altamira, Usina Kararahô, km 10 do acesso do acampamento Juruá,"* would likely result in an error of a minute or two if the map being used to determine the locality did not have *Usina Kararahô* on the map and one would have to make a decision whether the 10 km distance was straight line, 10 km of road, or 10 km of river distance. Historical gazetteers often had different minutes assigned to the same localities. Thus, the precision of historical locality data pre-global positioning devices (GPS), is questionable to inappropriate for fine-scaled understandings of distributions.

A second problem in evaluating distributions based on published data and museum specimen data is determining the distribution boundary/range.

Plotting and connecting peripheral localities as a source of baseline distributional information, clearly defines limits of species distributions, although it likely overestimates the interior area occupied by a species and may underestimate areas inhabited outside known localities. What are the criteria used by a researcher to determine whether a species is distributed continuously or not between two known localities? Such decisions would likely result in different evaluations based on whether the individual doing the evaluation actually collected the specimens involved and travelled between the localities or by someone who had never been in the geographic area involved and was using available maps and gazetteers.

Modern advances in the field of computer science, remote sensing, geographic information systems (GIS) and bioinformatics now permit previously unimaginable levels of analyses for biological collection data. One of the most promising analytical tools emerging from the technological advances are ecological niche models, algorithms embedded in a GIS framework that use the taxonomic and geographic data associated with specimens and fine scale environmental data to characterize suitable habitat for individual species and produce repeatable inferential maps of species distribution, as well as highlight distributional anomalies (Fernández, 2007).

Researchers interested in biogeographic analysis, such as niche modeling, face a heritage of data without coordinates. Most of the coordinate data available for members of the frog genus *Leptodactylus* are not derived from GPS readings taken at the actual collection sites. In other words, most of the available locality data are at least an order of magnitude worse than GPS based data. The question we pursue herein is whether the available coordinate data for species of *Leptodactylus,* although generally imprecise, are accurate enough for predictive distributional software to provide realistic new insights to species-level distributions.

Leptodactylus species in general are lowland species, mostly occurring at elevations below 1500 m, with a few species reaching 2000 m, and a single species reported to reach 3800 m (unpublished data).

We evaluate species with two known contrasting distribution patterns to assess the effectiveness of this computer modeling approach. *Leptodactylus knudseni* and *L. mystacinus* have relatively extensive distributions in South America and have been assumed to have continuous distributions (Heyer *et al.* 2003, Heyer *et al.* 2006). In contrast, *Leptodactylus lithonaetes, L. myersi, L. rugosus,* and *L. syphax* predominantly or exclusively occur on rocky outcrops that are discontinuously scattered across the South American landscape. Three of the latter four species have much smaller distributional extents than either *L. knudseni* or *L. mystacinus.*

Materials and Methods

Locality data and identifications

Species locality data were obtained from publications and museum specimens examined by SR and WRH (unpublished data).

For localities which SR and WRH have examined specimens, the species identifications should be close to 100% correct. Possible identification errors might have been made when small juvenile specimens were the only specimens available from a locality. There is a broad range of difficulty in correctly identifying species of *Leptodactylus* when only preserved specimens are available – and most specimen identifications are based on preserved specimens. Life colors and advertisement calls are often informative for separating closely related species, but these data are usually unavailable. The ease or difficulty in correctly identifying species evaluated in this paper ranges from easy to moderately difficult. An easy identification or diagnosis is when a non-herpetologist can correctly identify a species using available published keys and descriptions. A difficult identification requires expert knowledge of the species complex involved.

Based on our experience, the ranking of possible identification errors due to difficulty of identifications from easiest to most difficult is *L. mystacinus, L. syphax, L. rugosus, L. lithonaetes, L. knudseni, L. myersi.* We have no reason to question the 72% of the *L. mystacinus* localities for which we have not examined specimens as *L. mystacinus* is an easy species to correctly identify. *Leptodactylus lithonaetes* and *L. rugosus* have similar morphologies that taken together readily distinguish them from all other species of *Leptodactylus.* The only morphological features that can be used to separate preserved specimens of these two species are secondary male characteristics, which makes diagnosis of individual specimens based on morphological characters challenging. However, there is a gap between the distributions of these two species that aids in identification (see Results). The most difficult species to identify for the species evaluated in this analysis are *L. knudseni* and *L. myersi.* We have examined specimens of *L. myersi*

from 14 of the 15 known localities; consequently we feel confident and have no concerns about the distributional data for this species as analyzed in this study. However, it is virtually certain that some of the 32% published *L. knudseni* localities for which we have not seen specimens are incorrectly identified.

Environmental predictors

The potential distribution of each species was analyzed for the South American region.

1) *Climatic layers (CLIM):* Working under the assumption that at the scale of our research, real climatic conditions are well represented by environmental layers, data obtained from Worldclim 1.4 (Hijmans *et al.,* 2005) were consistently utilized in the analysis. Environmental layers for monthly mean, maximum and minimum temperature and precipitation were utilized in the modeling process. The 48 layers at the native resolution of 1 km^2 were resampled, at a higher resolution (smaller cell size: 450 m^2), to match the resolution of the rest of the layers (*e.g.,* elevation and ecological systems). The resample process performs a bilinear interpolation, and determines the new value of a cell based on a weighted distance average of the four nearest input cell centers. This technique is useful for continuous data, and will cause some smoothing on the resulting surface (Kidner, 2003).

2) *The Nature Conservancy-Natureserve ecological systems layer (TNC-NAT):* The Nature Conservancy South America Ecoregional Assessment Team developed a GIS modeling approach, using a variety of abiotic and biotic data sets (*e.g.* topography, geology, bioclimate, landcover, and landform), to produce regional maps of the distribution of Latin American and Caribbean (LAC) 492 ecological systems (Sayre *et al.,* 2005). Ecological systems are defined as spatially co-occurring assemblages of vegetation types sharing a common underlying substrate, ecological process, or gradient (Josse *et al.,* 2003). We combined these regional shape files (Amazonian, Atlantic forest, Central and South America Savannas, Northern South America, Southern South America East and Southern South America West) into one single shape file. Since the niche modeling algorithm that we chose cannot be fed with categorical data (Nix, 1986; Busby, 1991), we integrated the localities with the categorical ecological systems layer to produce a non-categorical binary output where a value of 1 was assigned to the areas (all the polygons that represent a classification unit) where the species had at least one locality. A value of 0 was assigned to the areas where the species had no localities. Finally, we rasterized the output at the recommended resolution of the product: 450 m^2 (Sayre *et al.,* 2005).

3) *Elevation (E):* Elevation at 450 meter resolution was also included in the analysis. The original data at 90 meter resolution was obtained from the Shuttle Radar Topography Mission and distributed by the U. S. Geological Survey. The data tiles were cleaned from anomalies and void areas also using a resample process (bilinear interpolation), at a lower resolution (450 m^2). Finally all the files were combined (mosaic) into a single elevation file for South America.

Bioclimatic-envelope modeling approach

Using the biogeographic data residing in natural history museum collections, several algorithms have been developed for modeling species-environmental relationships and generating inferential biodiversity maps (Phillips *et al.,* 2004; Segurado and Araújo, 2004; Guisan and Thuiller, 2005; Penman *et al.,* 2005; Elith *et al.,* 2006). We selected BIOCLIM (Nix, 1986; Busby, 1991) for the following reasons:

1) BIOCLIM is a deterministic non-black box modeling approach (Nix, 1986; Busby, 1991). Compared to other niche modeling algorithms (*e.g.,* Domain, MaxEnt or GARP), BIOCLIM is conceptually straightforward, transparent and well suited to model presence-only data in huge environmental datasets. This method relates occurrence localities to climatic conditions (environmental space), and produces a single rule that identifies all areas with a similar climate (suitable habitat), to the locations of the species, within a minimal rectilinear "climatic envelope". This environmental space can be projected onto geographic space to identify appropriate conditions for the species to occur (Guisan and Zimmerman, 2000).

2) BIOCLIM was coded inside DIVA GIS (Hijmans *et al.,* 2004), a freeware developed by the International Potato Center, the International Plant

Genetic Resources Institute and the Museum of Vertebrate Zoology at UC Berkeley. This software provides the capability of: a) building cumulative frequency distribution plots for the selected environmental variables; b) detecting multiple-variable outliers in environmental space; and c) environmental envelope visualization (Hijmans *et al.,* 2005).

3) BIOCLIM analyses are appropriate to answer our research question: Can generally imprecise locality data for *Leptodactylus* species provide new distributional insights into *Leptodactylus* distributions?

Sequential modeling

Using BIOCLIM, we model the distribution of each species in three steps. First, each species was modeled using the 48 climatic layers as predictors (CLIM). As a second step, the ecological systems layer (TNC-NAT) was included. Finally in the third round of models we included elevation (E) as a predictor.

Evaluation

A common process for evaluating model quality has been to divide known localities randomly into two groups: 1) the training data or data used to create the model, and 2) the test data or data to evaluate the model quality (Fielding and Bell, 2002). However, excluding part of the dataset, especially when not many localities are available for the species, will have a profound effect in the model quality no matter which algorithm is being used (Hernandez *et al.,* 2006). As a consequence the best model that can be created would include all the available records for the species (Anderson *et al.,* 2003).

A different question in the evaluation process is how well the localities represent the environmental space occupied by the species. We explore this issue by 10-folding the localities for each species, randomly sampling with replacement for nine groups of localities each time and modeling the distribution of the species. We exported the raw outputs into Idrisi format and used the Map Comparison Kit, software developed by the Netherlands Environmental Assessment Agency (Visser and Nijs, 2006) to compare the outputs using a fuzzy kappa (Hagen-Zanker *et al.,* 2005). The differences among the resulting preliminary models were on average 67% similar on a pairwise comparison suggesting that all the available localities were providing new information to the model.

In addition, our final models including all localities were evaluated by specialists (SR, WRH and RdS) based on understanding of the species' ecology and knowledge of the characteristics of the area climatic regime (Townsend Peterson *et al.,* 1999).

Results

Adequacy of environmental parameter data

There is a correlation between robustness of environmental distribution models and number of localities available for analysis; this is particularly important at the lower end of the correlation where the models are weaker because there are few known GIS localities. Examination of the environmental data for our data set suggests that for two of the species we may not have enough known localities available for the environmental parameter modeling results to be robust.

Twenty five localities for *L. lithonaetes* were available for analysis. Analysis of the annual precipitation data (Fig. 1) suggests that 36% of the available locality data samples are outliers with respect to at least three user-defined environmental layers. There are three plausible explanations for the high incidence of annual precipitation outliers: (1) the available data are inadequate to characterize the role that annual precipitation plays in the distribution of *L. lithonaetes;* (2) *L. lithonaetes* consists of 2 or 3 populations adapted to different climatic environments; or (3) the characteristics of the rocky habitats in which *L. lithonaetes* occurs may be quite similar amongst all known localities in terms of how annual precipitation impacts activity patterns even though the habitats surrounding the rocky outcrops show marked variation in annual precipitation. In other words, the adaptations of *L. lithonaetes* to rocky habitat environments override the significance of annual precipitation. The annual precipitation data for the morphologically similar *L. rugosus* (Fig. 2) has fewer outliers (14%) than *L. lithonaetes,* suggesting that the number of localities involved provides better environmental modeling results.

The 15 localities known for *L. myersi* show contrasting patterns for the annual mean temperature and annual precipitation data (Fig. 3). The annual

precipitation data demonstrate a pattern that allows characterization of those data for modeling purposes. In contrast, the annual mean temperature data indicate that there are two temperature patterns involved and combining the data from the two patterns might be inappropriate. Interestingly, *L. myersi* demonstrates geographically based morphological variation indicating that there are at least two groups of differentiated population systems or perhaps two distinct species involved (Heyer, 1995, Fig. 1). The correlation of differentiated morphological populations with the mean annual temperature data is weak, as there are 4 and 11 localities respectively represented in the morphologically differentiated units, whereas there are 7 and 8 localities respectively represented in the annual mean temperature differentiated units.

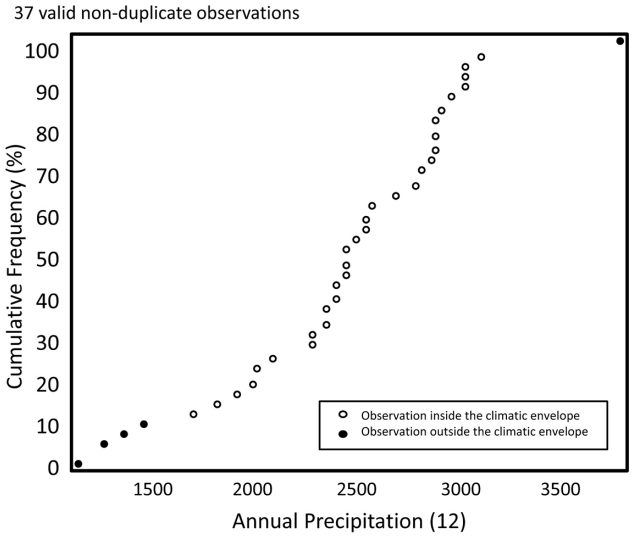


FIGURE 1. Cumulative frequencies for annual precipitation at known localities of *Leptodactylus lithonaetes.* Circles represent localities within the general environmental envelope defined by all the variables. Black dots represent outlier localities for at least three environmental layers (user defined)..

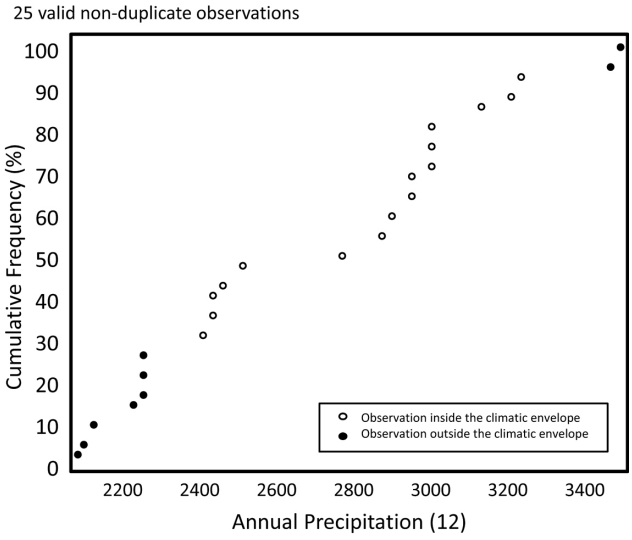


FIGURE 2. Cumulative frequencies for annual precipitation at known localities of *Leptodactylus rugosus.* Circles represent localities within the general environmental envelope defined by all the variables. Black dots represent outlier localities for at least three environmental layers (user defined)..

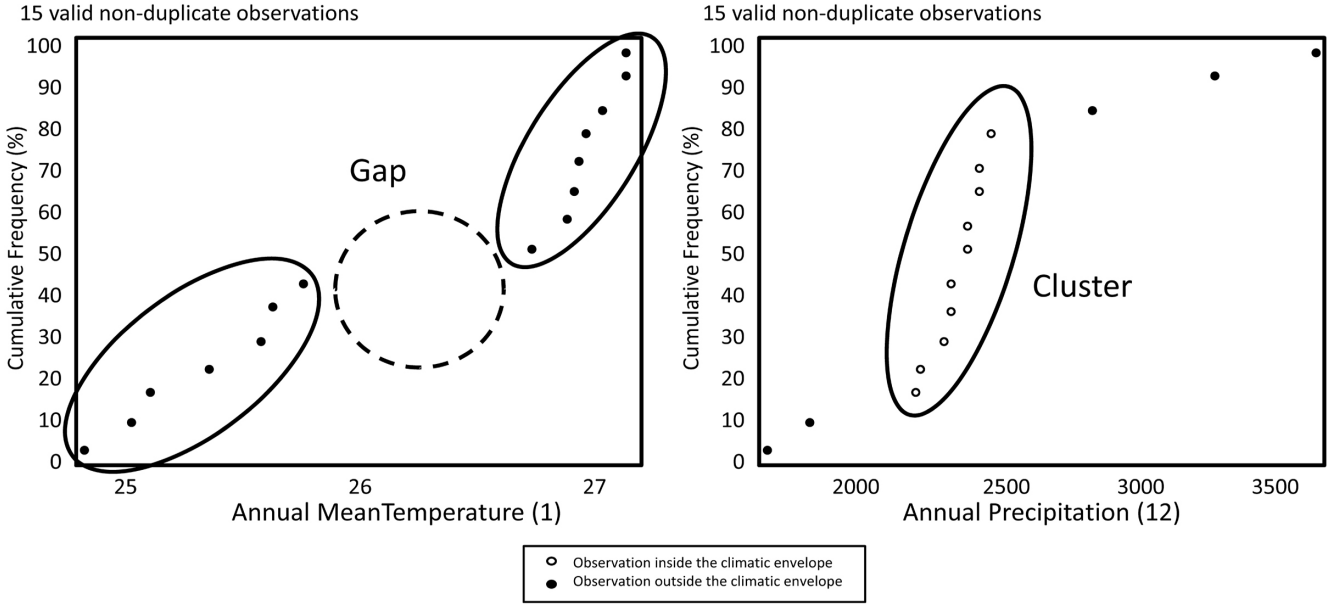


FIGURE 3. Cumulative frequencies for mean annual temperature (left) and annual precipitation (right) for known localities of *Leptodactylus myersi.* The left figure illustrates a gap in the mean annual temperature data, the right figure illustrates a tight cluster of annual precipitation data. Circles represent localities within the general environmental envelope defined by all the variables. Black dots represent outlier localities for at least three environmental layers (user defined).

Verified versus unverified identification locality data

As indicated in Materials and Methods, it is probable that some specimens from localities that we have not checked their identifications are misidentified and represent some other species of *Leptodactylus.* We examine this potential problem by comparing BIOCLIM output for all localities identified as *L. knudseni* (Fig. 4, upper left) versus localities whose identifications were verified by us (Fig. 4, lower left). Comparison of the two BIOCLIM maps indicates that there are only minor differences between the two, suggesting that most of the unverified identifications were in fact correct or from suitable habitats for *L. knudseni,* and therefore use of either data set is appropriate for predictive occurrence purposes.

Broadly ranging species distributions

Predicted distributions of *Leptodactylus knudseni* show similarities among the CLIM (Fig. 4, upper left), CLIM+TNC-NAT (Fig. 4, upper right), and CLIM+TNC-NAT+E (Fig. 4, lower right) models. However, the three models differ from each other markedly in how accurately the known locality data

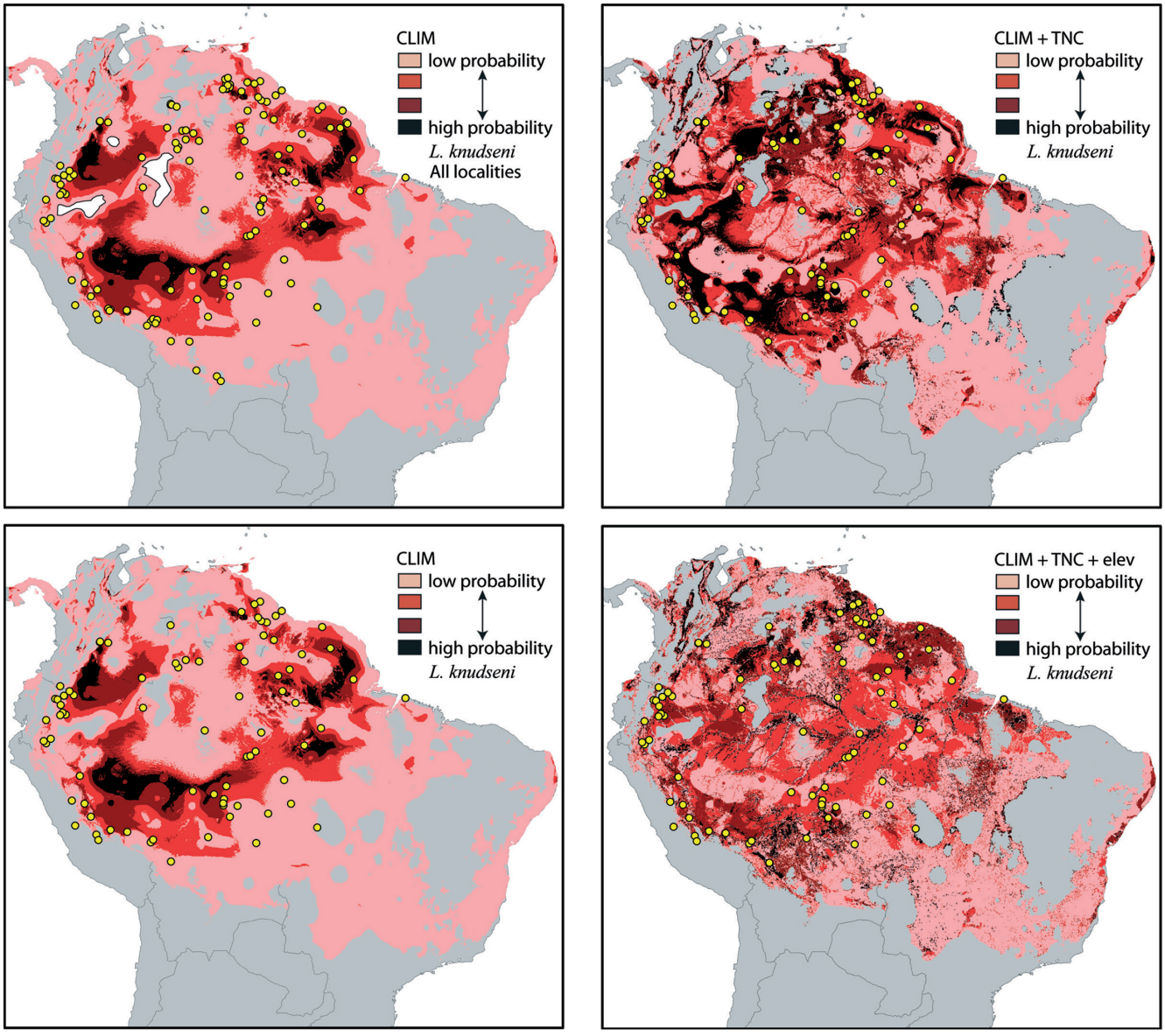


FIGURE 4. Predicted distributions for *Leptodactylus knudseni.* Upper left – BIOCLIM model for all locality data using the CLIM environmental variables, white areas highlight areas predicted to not be suitable for *L. knudseni;* lower left – BIOCLIM model based on CLIM environmental variables for locality data where authors examined and confirmed the species identifications; upper right – BIOCLIM model based on CLIM+TNC-NAT variables; lower right – BIOCLIM model based on CLIM+TNC-NAT+E variables.

Table 1. Known locality data occurrences within predicted area categories.

Species	Environmental variables	# localities in highest predicted areas (%)	# localities outside top 3 areas (%)
L. knudseni	CLIM	0 (0)	51 (47)
L. knudseni	CLIM+TNC	55 (50)	3 (3)
L. knudseni	CLIM+TNC+ELEV	63 (58)	1 (1)
L. mystacinus	CLIM	32 (14)	88 (37)
L. mystacinus	CLIM+TNC	116 (49)	3 (1)
L. mystacinus	CLIM+TNC+ELEV	139 (59)	9 (4)
L. lithonaetes	CLIM	2 (7)	9 (33)
L. lithonaetes	CLIM+TNC	14 (52)	1 (4)
L. lithonaetes	CLIM+TNC+ELEV	14 (52)	2 (7)
L. rugosus	CLIM	1 (2)	9 (21)
L. rugosus	CLIM+TNC	19 (44)	0 (0)
L. rugosus	CLIM+TNC+ELEV	24 (56)	0 (0)
L. myersi	CLIM	0 (0)	9 (38)
L. myersi	CLIM+TNC	3 (19)	7 (44)
L. myersi	CLIM+TNC+ELEV	9 (56)	2 (12)
L. syphax	CLIM	1 (4)	22 (42)
L. syphax	CLIM+TNC	40 (77)	1 (2)
L. syphax	CLIM+TNC+ELEV	34 (65)	2 (4)

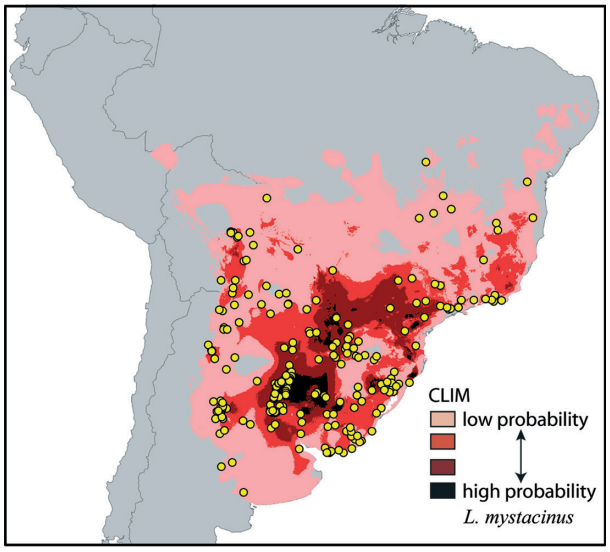
fit the models (Table 1). The CLIM+TNC-NAT and CLIM+TNC-NAT+E models demonstrate a better fit of the data than the CLIM model in that many more known localities coincide with the best category of predicted suitable environment/ecoregion. Furthermore, the CLIM+TNC-NAT+E model is marginally better than the CLIM+TNC-NAT model.

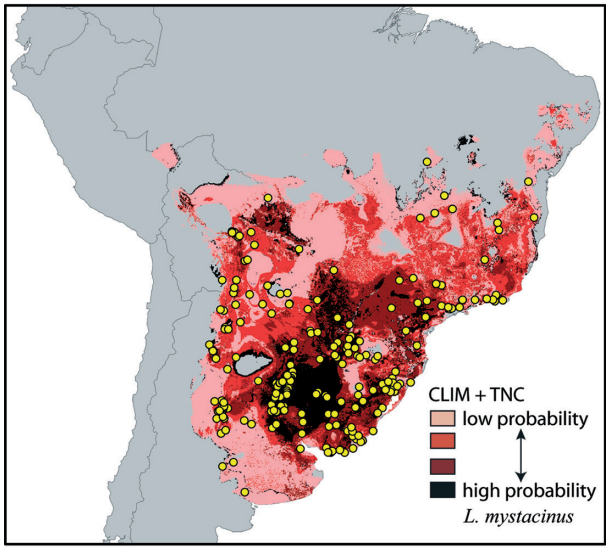
Previous range maps for *L. knudseni* have assumed a continuous distribution throughout Amazonia (*e.g.,* Heyer *et al.* 2006). The CLIM map for *L. knudseni* indicates three extensive areas within Amazonia that are environmentally unsuitable (Fig. 4 upper left, unsuitable areas in white). Each of these three areas would require additional fieldwork to determine whether, in fact, *L. knudseni* does not occur in these areas. The CLIM+TNC-NAT and CLIM+TNC-NAT+E models indicate that *L. knudseni* would be able to occur in the Atlantic Forest regions of northeast Brasil. Interestingly, there is no large species of the *Leptodactylus pentadactylus* clade that is known to occur in the northern Atlantic Forests. *Leptodactylus flavopictus* is the only species of the *L. pentadactylus* clade that occurs in the Atlantic Forests and is known only from localities in the states of Espírito Santo and São Paulo. The lack of a large species of the *L. pentadactylus* group in the northern Atlantic Forests is puzzling. There are other taxa known from the suitable regions identified in Fig. 4 (upper and lower right), such as *Leptodactylus mystaceus,* that are known from both Amazonian and Atlantic Forest ecoregions (Heyer, 1978). Amazonian and Atlantic Forest populations of species such as *L. mystaceus* were doubtless in genetic contact during the wettest periods of the Pleistocene. It is possible that *L. knudseni* formerly occurred in the northern Atlantic Forest region and went extinct there sometime in the not too distant past, possibly due to environmental changes, habitat degradation and loss, or human activities.

Leptodactylus mystacinus is another broadly distributed species, presumed to have a continuous distribution (Heyer *et al.* 2003) extending from southern tropical regions of Brasil to temperate regions of Argentina. Geographic sampling for this species is uneven, with a much more intensive sampling along the Río Paraná in the State of Entre Ríos. We did not assess how this biased sampling may have impacted the overall distributional model. In terms of models fitting the available locality data, the CLIM+TNC-NAT (Fig. 5, middle) model and CLIM+TNC-NAT+E (Fig. 5, lower) model provide much better fits than the CLIM model (Fig. 5, upper). The CLIM+TNC-NAT+E model places more known localities in the predicted area of best quality occupancy conditions than the CLIM+TNC-NAT model; the CLIM+TNC-NAT model places fewer known localities in poor quality occupancy conditions than the CLIM+TNC-NAT+E model (Table 1).

Patchily distributed species

Four species of *Leptodactylus* occur exclusively or primarily on sandstone and/or granitic rock outcrops. These rocky outcrops are disjunctly scattered





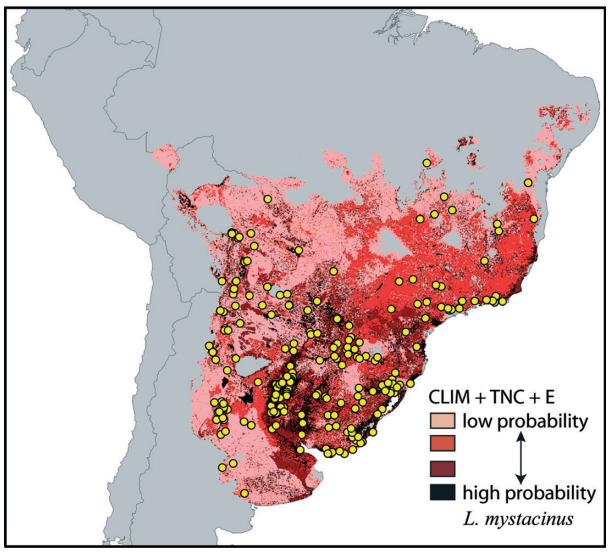
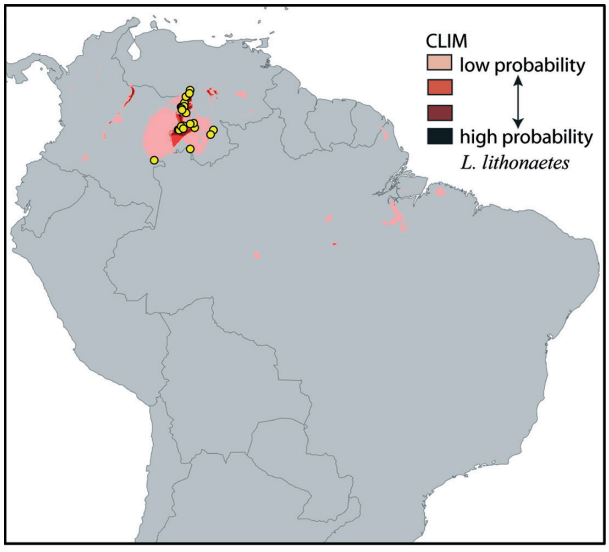
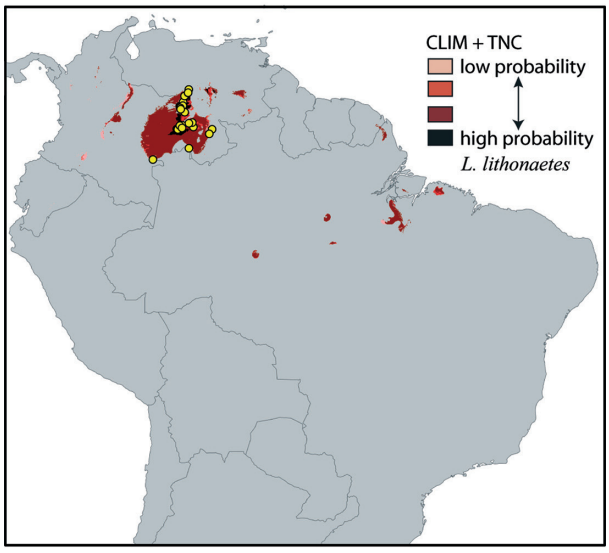


FIGURE 5. Predicted distributions for *Leptodactylus mystacinus*. Upper figure – BIOCLIM model using the CLIM environmental variables; middle figure – BIOCLIM model using the CLIM+TNC-NAT variables; lower figure – BIOCLIM model using the CLIM+TNC-NAT+E variables.





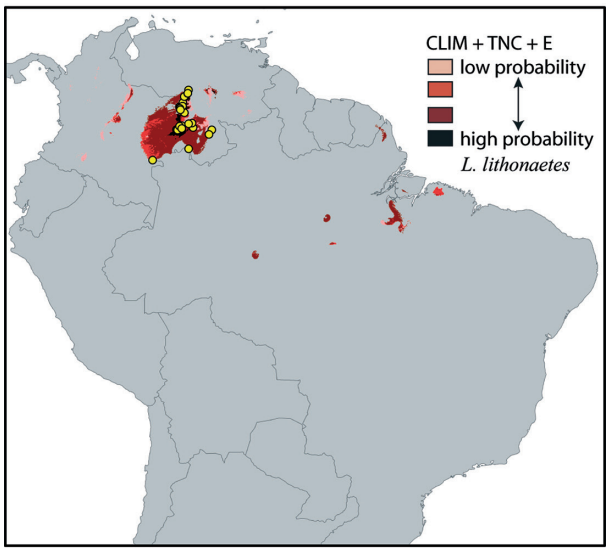


FIGURE 6. Predicted distributions for *Leptodactylus lithonaetes*. Upper figure – BIOCLIM model based on CLIM environmental variables; middle figure – BIOCLIM model based on CLIM+TNC-NAT environmental variables; lower figure – BIOCLIM model based on CLIM+TNC-NAT+E variables.

across the South American landscape where *Leptodactylus* occur. The rocky outcrops vary considerably in size, from small local areas measured in terms of fractions of hectares to extensive formations such as the campos rupestres of the Serra do Espinhaço and Chapada Diamantina in Brasil.

Leptodactylus lithonaetes and *L. rugosus* are closely related rocky outcrop associated species, which can be told apart morphologically from each other only on the basis of male secondary sexual characteristics.

The CLIM+TNC-NAT and CLIM+TNC-NAT+E models for *L. lithonaetes* perform similarly in accounting for known locality data and both perform much better than the CLIM model (Fig. 6, Table 1).

Of the three predictive occurrence models for *L. rugosus* (Fig. 7), the CLIM+TNC-NAT+E model is the most robust in terms of fitting of known locality data, with the CLIM+TNC-NAT model clearly superior to the CLIM model (Table 1).

The rather extensive disjunct areas in the Brasilian states of Amazonas and Pará as particularly shown in the CLIM+TNC-NAT model, indicate suitable macrohabitat conditions for *L. rugosus* (Fig. 7 middle). If there are rocky outcrops in these areas of suitable environmental conditions, there could be a rocky outcrop associated species of *Leptodactylus* occurring on them. Given the extensive geographical distances involved from known *L. rugosus* localities, a rocky outcrop species in the Amazonas and Pará regions of suitable environmental conditions would most likely represent undescribed species of *Leptodactylus*.

Leptodactylus myersi occurs on granitic and sandstone outcrops in a disjunct pattern from the Brasilian states of Pará and Roraima to Suriname and French Guiana. One specimen (from Roraima) has been collected from forested habitat somewhat distant from the nearest rocky outcrop. The CLIM+TNC-NAT+E model far outperforms the other two models in terms of accounting for the known locality data (Fig. 8, Table 1).

Leptodactylus syphax occurs primarily in drier ecoregions, usually associated with rocky outcrops. The association with rocky outcrops is true for almost all localities, but is not absolute. Specimens have been found in ant colonies of the genus *Atta* nowhere near rocky outcrops and specimens have been collected from one restinga (sandy scrub coastal habitat) site far removed from any rocky outcrops (Heyer *et al.,* in press). The CLIM+TNC-NAT model best fits the known locality data for *L. syphax,* closely followed by the CLIM+TNC-NAT+E model with the CLIM model performing rather poorly (Fig. 9, Table 1).

Table 2. Extracted model values for localities of Cerro Guanay, Bolívar, Venezuela and Maigualida, Amazonas, Venezuela from three sets of GIS models.

Species/Environmental layers	Cerro Guanay	Maigualida
Leptodactylus lithonaetes CLIM	0	0
Leptodactylus lithonaetes CLIM+TNC-NAT	0	0
Leptodactylus lithonaetes CLIM+TNC-NAT+E	45.12	45.11
Leptodactylus rugosus CLIM	0	0
Leptodactylus rugosus CLIM+TNC-NAT	31.23	29.7
Leptodactylus rugosus CLIM+TNC-NAT+E	24.81	24.83

Can modeling resolve species identifications?

At the time the computer modeling was done for this project, there were two localities for which only juvenile specimens were available that could not be distinguished between *L. lithonaetes* or *L. rugosus.* Both localities occur between the known localities for *L. lithonaetes* and *L. rugosus.* An obvious question was whether the GIS models could assist on deciphering the species identifications of the juveniles from these two localities. Neither locality falls within the predicted areas of suitable ecoregions for *L. lithonaetes* or *L. rugosus* (Figs. 6, 7). The extracted data for the two localities of Cerro Guanay, Bolívar, Venezuela and Maigualida, Amazonas, Venezuela were examined to determine whether the model parameters would indicate which of the two species occur at each site (Table 2). The *Leptodactylus lithonaetes* CLIM+TNC-NAT+E model values indicate that those model data are most robust and thus *Leptodacytlus lithonaetes* is predicted to occur at both localities. As this paper was being written, Myers and Donnelly (2008:87) reported that they examined adult male specimens from Cerro Guanay (104 km distance from the nearest *L. lithonaetes* locality) and identified them as *L. lithonaetes.* These results are most encouraging: individual locality data can be used in the GIS modeling framework to evaluate species identifications of problematic specimens for certain classes of identification problems.

Discussion, Conclusions, and Recommendations

Incorporation of the TNC-NAT vegetation classification layer consistently provided a distributional model with a much better fit of known locality data than those based solely on CLIM environmental

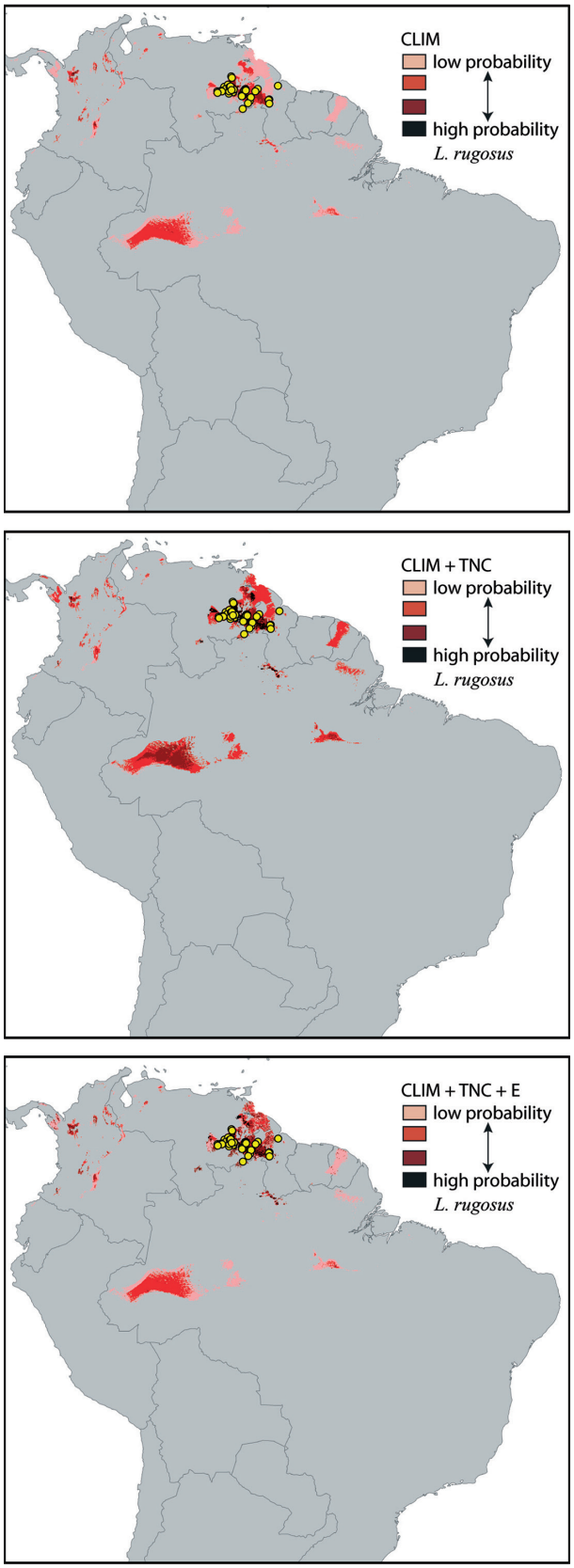


FIGURE 7. Predicted distributions for *Leptodactylus rugosus*. Upper figure – BIOCLIM model based on CLIM variables; middle figure – BIOCLIM model based on CLIM+TNC-NAT variables; lower figure – BIOCLIM model based on CLIM+TNC-NAT+E model.

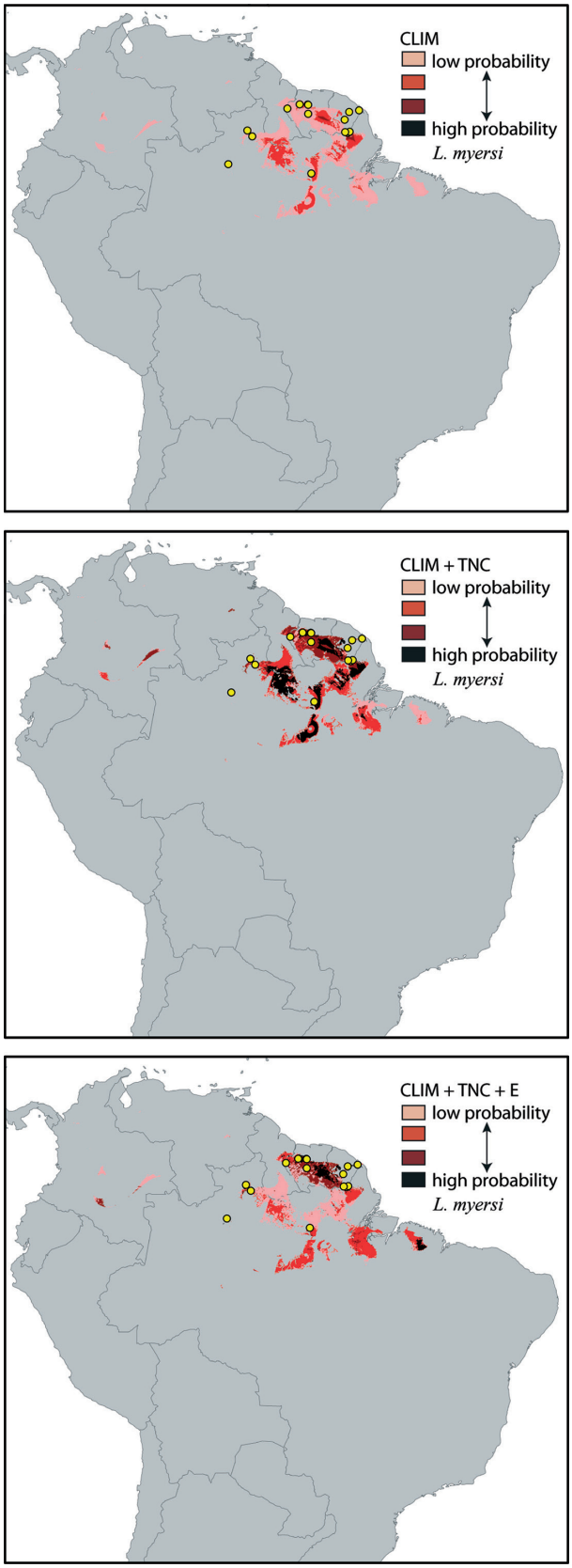


FIGURE 8. Predicted distributions for *Leptodactylus myersi*. Upper figure – BIOCLIM model based on CLIM variables; middle figure – BIOCLIM model based on CLIM+TNC-NAT variables; lower figure – BIOCLIM model based on CLIM+TNC-NAT+E variables.

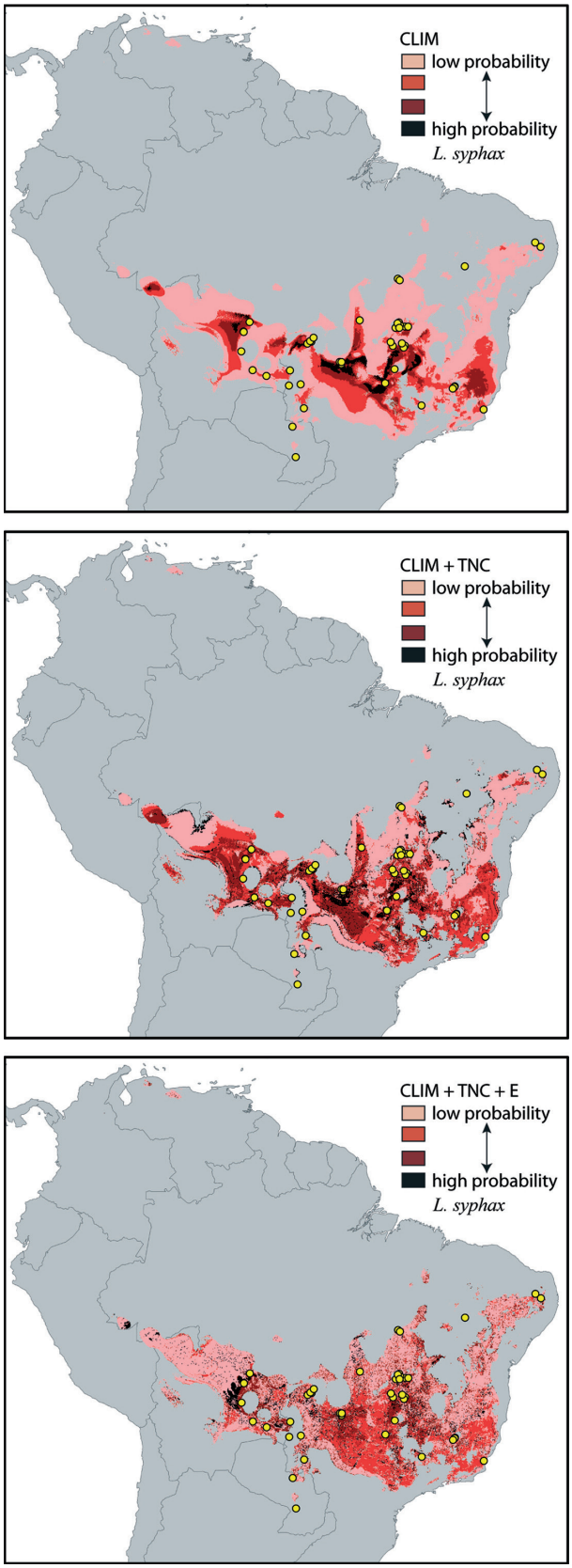


FIGURE 9. Predicted distributions for *Leptodactylus syphax*. Upper figure – BIOCLIM model based on CLIM variables; middle figure – BIOCLIM model based on CLIM+TNC-NAT variables; lower figure – BIOCLIM model based on CLIM+TNC-NAT+E variables.

variables analyses. We recommend use of this layer for all GIS studies involving the Neotropical fauna.

The results were variable regarding performance of the CLIM+TNC-NAT versus the CLIM+TNC-NAT+E models. In four cases, the CLIM+TNC-NAT+E model yielded more robust results, in one case the CLIM+TNC-NAT model was the most robust, and in one case the two models were equivalent. The known elevations for the six species included in our analyses range from 50-1000 to 0-1800 m, with most localities occurring below 1500 m. One would expect that elevation would not be a consistently critical parameter for lowland frogs. However, our results indicate that the high resolution elevation layer is certainly worth evaluating for Neotropical lowland frogs such as *Leptodactylus*.

Niche modeling for broadly and narrowly distributed species yielded different insights than distribution interpretations based on dot maps with experienced investigator delineation of distributional limits, which was the only option available prior to GIS analyses. The GIS modeling of the six *Leptodactylus* species provides new insights and questions regarding their distributions.

Leptodactylus knudseni and *L. mystacinus* clearly are not continuously distributed within the region bounded by the most peripheral localities. We are impressed by the sizeable areas predicted to be environmentally unsuitable within the overall range of the species. At least some of these predicted unsuitable environments should be ground-truthed to verify or refute the postulate that the species really does not occur there.

Given the strikingly bleak rocky habitats where some species of *Leptodactylus* occur, some of the authors assumed that the environmental niche modeling techniques would produce either unrealistically large or small geographic distributions, because the rocky outcrops themselves are small and locally distributed in most cases. *Leptodactylus lithonaetes* and *L. rugosus* occur on granitic outcrops, where there is no buffering vegetation layer that ameliorates the effect of sun on the rocky surface where these frogs live. These rocky outcrops are similar to desert conditions for the limited biota that occurs on them (*e.g.,* cacti occur on the outcrops). The environmental niche modeling results that *L. lithonaetes* and *L. rugosus* could be adapted to general environmental conditions surrounding the rocky outcrops was unexpected. The insight that the environmental niche models provide is that the general environmental conditions in which these rocky outcrops occur differ for *L. lithonaetes* and

L. rugosus, and these differences do an unexpectedly good job of modeling the distributions of these two species. The analyses suggest that *L. lithonaetes* and *L. rugosus* have adapted to the unique environmental conditions found in the areas where the two species occur, such as differences in the precipitation of the warmest quarter. Adaptation to these environmental differences may have been the primary selective factors leading to differentiation of these two species. Barrio-Amoros and Brewer-Carias (2008, pages 27-28) included *L. lithonaetes* and *L. rugosus* as examples of a distribution pattern of "vicariance associated with an axis following the Maigualida-Parima Mountains, in which the headwaters of the upper Orinoco and Ventuari rivers are separated from the Caura, Paragua, and Caroní drainages." The proposed vicariant pattern is supported by the distinctive environments associated with the two areas, as demonstrated in our analyses.

Under certain classes of species identification problems, GIS analyses of the relatively imprecise *Leptodactylus* data can provide inferences as to which species occurs at localities where species identifications are problematic.

Our results provide blueprints for fieldwork to refine the accuracy of our understanding for the *Leptodactylus* species we have analyzed. The following examples are only a sample of the field work that should be undertaken to refine our understanding of the species involved.

- *Leptodactylus rugosus* – The high probability area indicated in Fig. 7 on the Venezuela-Guyana border that is significantly farther north than any known *L. rugosus* locality should be worked to see if the species occurs there. The very distant area of suitable distributional conditions for *L. rugosus* in Amazonas, Brasil (between northern Bolivia and Peru) is begging for ground truthing. If there are rocky outcrops and frogs on the outcrops in this area of high probability environmental conditions in Amazonas, Brasil, it is most likely that they would represent a new species of *Leptodactylus,* rather than representing a distant population of *L. rugosus.*

- *Leptodactylus syphax* – There are many higher probability areas for *L. syphax* that need to be ground truthed. Field work in the higher probability areas would certainly refine our understanding of the distribution of *L. syphax.* Of greatest priority, field work needs to be undertaken in the tantalizing areas bordering Peru, Brazil, and Bolivia (Fig. 9).

Our results indicate that at least one species, *Leptodactylus myersi,* may actually contain two or more species, as there is a gap in distribution of at least one environmental parameter, annual mean temperature (Fig. 3). There are several, disjunct areas of high probability that should be explored to determine whether there are *L. myersi*-like frogs, and if so, whether there are two or more species involved (Fig. 8).

Our results also have conservation implications. The environmental niche model maps for all the species we analyzed have high probability areas well outside the distributions based only on locality records. For example, our results indicate that the environment in the northern Atlantic Forests of Brasil (Fig. 4) is suitable for *L. knudseni.* Transporting and releasing *L. knudseni* into the northern Atlantic Forests could result in a fundamental restructuring of the ecological community where the *L. knudseni* were released. This type of environmental modeling can assist in conservation efforts by identifying clusters of populations that vary in their environmental requirements and are currently placed and considered a single and widely distributed taxon. Further analyses of those populations may reveal morphological or genetic differentiation to consider them separate evolutionary lineages; maintaining genetic diversity is a guiding principle for long term viability of biodiversity. We should review the "least concern" conservation status of widely distributed taxa currently considered to be single species.

Overall, we are very encouraged by our results. In answering our original question, we conclude that the imprecise locality data predominantly available for *Leptodactylus* species localities are in fact adequate to provide new insights into *Leptodactylus* distributions, both for broadly and narrowly distributed species. Environmental niche modeling should be one of the tools used to understand *Leptodactylus* species distributions.

Resumo

Dados de localidades disponíveis para muitas, senão a maioria, das espécies de anfíbios Neotropicais são baseadas em descrições das áreas de coleta, não de coordenadas obtidas através do dispositivo GPS. Os dados pré-GPS são imprecisos se comparados aos dados obtidos através deste dispositivo. A modelagem de nichos consiste em uma poderosa técnica na previsão da distribuição geográfica que proporciona os melhores resultados quando a localidade é precisa. O objetivo deste

estudo é de determinar se as localidades imprecisas são suficientes para que as técnicas de modelagem de nichos promovam perspectivas mais realistas acerca da distribuição ao nível de espécie. Dois conjuntos de anfíbios do gênero *Leptodactylus* que apresentam diferentes padrões de distribuições foram avaliados: duas espécies com ampla, presumivelmente distribuição contínua e quatro espécies conhecidas por apresentarem distribuição mais restrita, em corredores ou habitats disjuntos na América do Sul. BIOCLIM, um algoritmo de modelagem de nicho que se baseia somente em dados de presença foi utilizado para definir as prováveis áreas de ocorrência baseada em múltiplos conjuntos de parâmetros ambientais que incluem: média e precipitação mensal e temperatura mínima e máxima. Foram incluídas nas análises dois *layers* da Nature Conservancy Natureserve, um de ecoregião e outro de altitude. Nossa análise fornece uma nova visão acerca da distribuição das espécies de *Leptodactylus* analisadas neste estudo. Nós recomendamos a incorporação dos *layers* da Nature Conservancy Natureserve para avaliar a distribuição Neotropical uma vez que estes forneceram resultados mais robustos que simplesmente o uso de análises baseadas em variáveis climáticas.

Acknowledgments

We thank Dr. Peltzer and Dr. Lajmanovich for distributional data for samples of *L. mystacinus*. This work was supported through NSF Award 0342918 to RdS and WRH. MF received support from the Lakeside Foundation, California Academy of Sciences and the WWF Russell E. Train Education for Nature Program. We thank The Nature Conservancy for allowing us early access to The Nature Conservancy – Natureserve ecological systems layer.

Literature Cited

Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecological Modelling, 162(3):211-232.

Barrio-Amoros, C. L. and C. Brewer-Carias. 2008. Herpetological results of the 2002 expedition to Sarisariñama, a tepui in Venezuelan Guayana, with the description of five new species. Zootaxa, (1942):1-68.

Berra, T. M. 1981. An atlas of distribution of the freshwater fish families of the world. University of Nebraska Press, Lincoln, xxix+197.

Busby, J. R. 1991. BIOCLIM–A bioclimatic analysis and prediction system; pp. 64-68. In: C. R. Margules and M. P. Austin (Eds.), Nature Conservation: Cost Effective Biological Surveys and Data Analysis. CSIRO, Melbourne.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, 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, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecogeography, 29:129-151.

Erwin, T. L. 1996. Arboreal Beetles of Neotropical Forests: *Agra* Fabricius, the *cayennensis* complex (Coleoptera: Carabidae: Lebiini: Calleidina). Annales Zoologici Fennici, 33(1):17-21.

Fernández, M. 2007. Assessing the effect of locality accuracy in predicting species distribution. Geography. Master of Arts thesis, San Francisco State University, San Francisco, 75 pp.

Fielding, A. H. and J. F. Bell. 2002. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation, 24(1):38-49.

Guisan, A. and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8(9):993-1009.

Guisan, A. and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling, 135(2):147-186.

Hagen-Zanker, A., B. Straatman, and I. Uljee. 2005. Further developments of a fuzzy set map comparison approach. International Journal of Geographical Information Science, 19(7):769-785.

Hall, E. R. 1981. Mammals of North America. Second Edition. Volume 1. John Wiley and Sons Publisher, New York, 600+90 pp.

Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography, 29(5):773-785.

Heyer, M. M., W. R. Heyer, S. Spear, and R. O. de Sá. 2003. *Leptodactylus mystacinus*. Catalogue of American Amphibians and Reptiles, (767):1-11.

Heyer, W. R. 1978. Systematics of the *fuscus* group of the frog genus *Leptodactylus* (Amphibia, Leptodactylidae). Natural History Museum Los Angeles County, Science Bulletin, (29):1-85.

Heyer, W. R. 1995. South American rocky habitat *Leptodactylus* (Amphibia: Anura: Leptodactylidae) with description of two new species. Proceedings of the Biological Society of Washington, 108:695-716.

Heyer, W. R. and M. M. Heyer. 2006. *Leptodactylus knudseni*. Catalogue of American Amphibians and Reptiles, (807):1-12.

Heyer, W. R., M. M. Heyer, and R. O. de Sá. In Press. *Leptodactylus syphax*. Catalogue of American Amphibians and Reptiles.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25(15):1965-1978

Hijmans, R. J., L. Guarino, C. Bussink, P. Mathur, M. Cruz, I. Barrantes, and E. Rojas. 2004. DIVA-GIS, version 4. A geographic information system for the analysis of biodiversity data. [Online] Available at www.diva-gis.org.

Hijmans, R. J., L. Guarino, A. Jarvis, R. O'Brien, P. Mathur, C. Bussink, M. Cruz, I. Barrantes, and E. Rojas. 2005. DIVA-GIS, Version 5.2. Manual. Available at www.diva-gis.org.

Howell, S. N. G. and S. Webb. 1995. A guide to the birds of Mexico and northern Central America. Oxford University Press, Oxford, 851 pp.

Josse, C., G. Navarro, P. Comer, R. Evans, D. Faber-Langendoen, M. Fellows, G. Kittel, S. Menard, M. Pyne, and M. Reid. 2003. Ecological Systems of Latin America and the Caribbean: A Working Classification of Terrestrial Systems. NatureServe, Arlington, Virginia, USA.

Kidner, D. B. 2003. Higher-order interpolation of regular grid digital elevation models. International Journal of Remote Sensing, 24(14):2981-2987

Myers, C. W. and M. A. Donnelly. 2008. The summit herpetofauna of Auyantepui, Venezuela: Report from the Robert G. Goelet American Museum-Terramar Expedition. Bulletin of the American Museum of Natural History, (308):1-147.

Nix, H. A. 1986. A biogeographic analysis of Australian elapid snakes; pp. 4-15. In: Bureau of Flora and Fauna (Ed.), Atlas of Australian Elapid Snakes, Australian Government Publishing Service.

Penman, T. D., M. J. Mahony, A. L. Towerton, and F. L. Lemckert. 2005. Bioclimatic analysis of disjunct populations of the giant burrowing frog, *Heleioporus australiacus*. Journal of Biogeography, 32(3):397-405.

Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Ecological modeling. Proceedings of the 21st International Conference on Machine Learning, Banff, Canada, 69:83-90.

Sayre, R., J. Bow, and C. Josse. 2005. Modeling NatureServe's Latin America and Caribbean Ecological Systems throughout South America. The Nature Conservancy, 25 pp.

Segurado, P. and M. B. Araújo. 2004. An evaluation of methods for modelling species distributions. Journal of Biogeography, 31:1555-1568.

Townsend Peterson, A. and K. P. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modelling, 117(1):159-164.

Visser, H. and T. de Nijs. 2006. The map comparison kit. Environmental Modelling and Software, 21(3):346-358.

Wieczorek, J., Q. Gu., and R. Hijmans. 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. International Journal of Geographical Information Science, 18(8):745-767.

Submitted 22 September 2008
Accepted 18 June 2009