
Mapping and Modeling Neglected Tropical Diseases and Poverty in Brazil, Bolivia and Colombia

JB Malone, P Nieto, P Mischler, M Martins, JC McCarroll

Louisiana State University, USA

Penelope Vounatsou, Ronaldo Scholte

Swiss TPH, Switzerland

ME Bavia

Universidade Federal da Bahia, Brazil

HysPIRI Symposium, NASA Goddard Space Flight Center , May 29-30, 2013



Objectives

- Data Portal – A resource data base accessible by FTP was developed for 6 NTD in Brazil, Bolivia and Colombia (Chagas disease, Leishmaniasis, Schistosomiasis, Leprosy, Lymphatic Filariasis and Soil-Transmitted Helminths), with relevant climatic, environmental, population and poverty data
- Risk Modeling – Maximum Entropy, Bayesian and GIS methodologies were used to map and model environmental and socioeconomic risk of 6 NTD
- Course Development – A 4-day short course was developed for training use by PAHO on data portal access and geospatial analysis using ArcGIS 9.3.1, Maximum Entropy (Maxent) and Bayesian modeling

Data Portal

All data clipped to the country boundary; WGS84 projection, 1 km spatial resolution; in ASCII format for Maxent or Bayesian modeling

This example shows the data available for Colombia

Worldclim (global coverage, 1km resolution) used for ecological Niche modeling and by the climate change community

MODIS EVI, LST annual composites for 2005-2009

Socioeconomic Data at the Municipality level

PAHO Data Portal

COLOMBIA



Disease Mapping and Modeling for Neglected and Other Poverty-Related Diseases in Latin America and the Caribbean

Minimum Medical Database	WorldClim Data
The following data are from the South America MMDb and have been clipped to Colombia.	The following data has been prepared to use in Maxent.
Images	Bioclim Variables
MODIS 2003 Composites	Altitude
EVI image - ascii	Min Temperature
Temp Day image - ascii	Mean Temperature
Temp Day image - ascii	Mean Temperature
Temp Night	Max Temperature

[Return to PAHO Project Wikipage](#)

La
in
D
in
S
C
E
H
I
P
P

PAHO Data Portal

Disease Mapping and Modeling for Neglected and Other Poverty-Related Diseases in Latin America and the Caribbean

Click on the country of interest to see data available for download.



Return to [PAHO Project Wikipage](#)

Worldclim Global Climate Data

Tmin, Tmax, Precip, SRTM, Bioclim – 1 km resolution

Bioclimatic variables are derived from the monthly temperature and rainfall values in order to generate more biologically meaningful variables. These are often used in ecological niche modeling (e.g., BIOCLIM, GARP).

BIO1 = Annual Mean Temperature

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3 = Isothermality (P2/P7) (* 100)

BIO4 = Temperature Seasonality (standard deviation *100)

BIO5 = Max Temperature of Warmest Month

BIO6 = Min Temperature of Coldest Month

BIO7 = Temperature Annual Range (P5-P6)

BIO8 = Mean Temperature of Wettest Quarter

BIO9 = Mean Temperature of Driest Quarter

BIO10 = Mean Temperature of Warmest Quarter

BIO11 = Mean Temperature of Coldest Quarter

BIO12 = Annual Precipitation

BIO13 = Precipitation of Wettest Month

BIO14 = Precipitation of Driest Month

BIO15 = Precipitation Seasonality (Coefficient of Variation)

BIO16 = Precipitation of Wettest Quarter

BIO17 = Precipitation of Driest Quarter

BIO18 = Precipitation of Warmest Quarter

BIO19 = Precipitation of Coldest Quarter

Contents of Data Portal/FTP Site

MODIS

Mean annual composites for 2005-2009:

Enhanced Vegetation index (EVI), Normalized difference Vegetation Index (NDVI)

Land surface temperature (LST) day and night and dT

Climate GRID

Long term normal (LTN) climate grid (18x18 km cell size) – Precip, Tmax, Tmin, PET, PPE

Environmental

World Wildlife Fund Ecoregions

Locations of springs, dams, rivers, small streams

Health Data

Bolivia: Ministerio de Salud y Deportes/ Sistema Nacional de información en Salud

Brazil: Ministerio da Saude, SINAN

Colombia: Instituto Nacional de salud/Estadísticas de la Vigilancia en Salud Pública

Ministerios de la protección Social (SIVIGILA) , literature reports.

Infrastructure

Roads, airfields/airports, rail road lines layer, utility lines

Political Boundaries

Counties, major cities, States/Departments, Municipalities

Socioeconomical Variables at Municipality Level Used for Risk Analysis of NTDs in Colombia

Area of municipality	Floors: carpet, marmol, hardwood, tablet	Garbage: in the river, stream, lake, lagoon
Displacement (just COL)	Floors: carpet, brick , vinyl,	Garbage: in another way
Population	Floors: cement	Drinking water from: running water service
Extension Km2	Floors: tough wood, other vegetal material	Drinking water from: well, pump
Human development index	Floors: soil, sand	Drinking water: rain fall
Unsatisfied basic needs * UBN	Walls: block, brick, stones, hardwood	Drinking water: public tank
Miseria (2 or more *UBN)	Walls: adobe, bahareque	Drinking water: car-tank
Un adequate housing * UBN	Walls: rough wood	Drinking water from: river, stream, lake , lagoon
Unsatisfied services* UBN	Wall: pre fabricated walls	Drinking water from: bottles, bag
Overcrowding * UBN	Walls: cane, bamboo, vegetal material	Infant mortality
Educational needs* UBN	Walls: zinc, fabric, cardboard, plastic	Life expectancy
Economical dependency*UBN	No walls	Attendance educational institution YES
Sewage	Electricity: yes	Attendance /educational institution NO
Running water	Electricity: no	
Toilet connected to sewage	Garbage collection services	
Toilet connected to septic tank	Burrow the garbage	
Latrine	Burn the garbage	
No sanitary service	Garbage: patio, back yard, ditch	

Table 1. Socioeconomical variables (47) selected for risk analysis of NTDs in Colombia

*UBN: http://www.dane.gov.co/files/investigaciones/boletines/censo/Bol_nbi_censo_2005.pdf



Opennlp.maxent package is a mature Java package for training and using maximum entropy models.

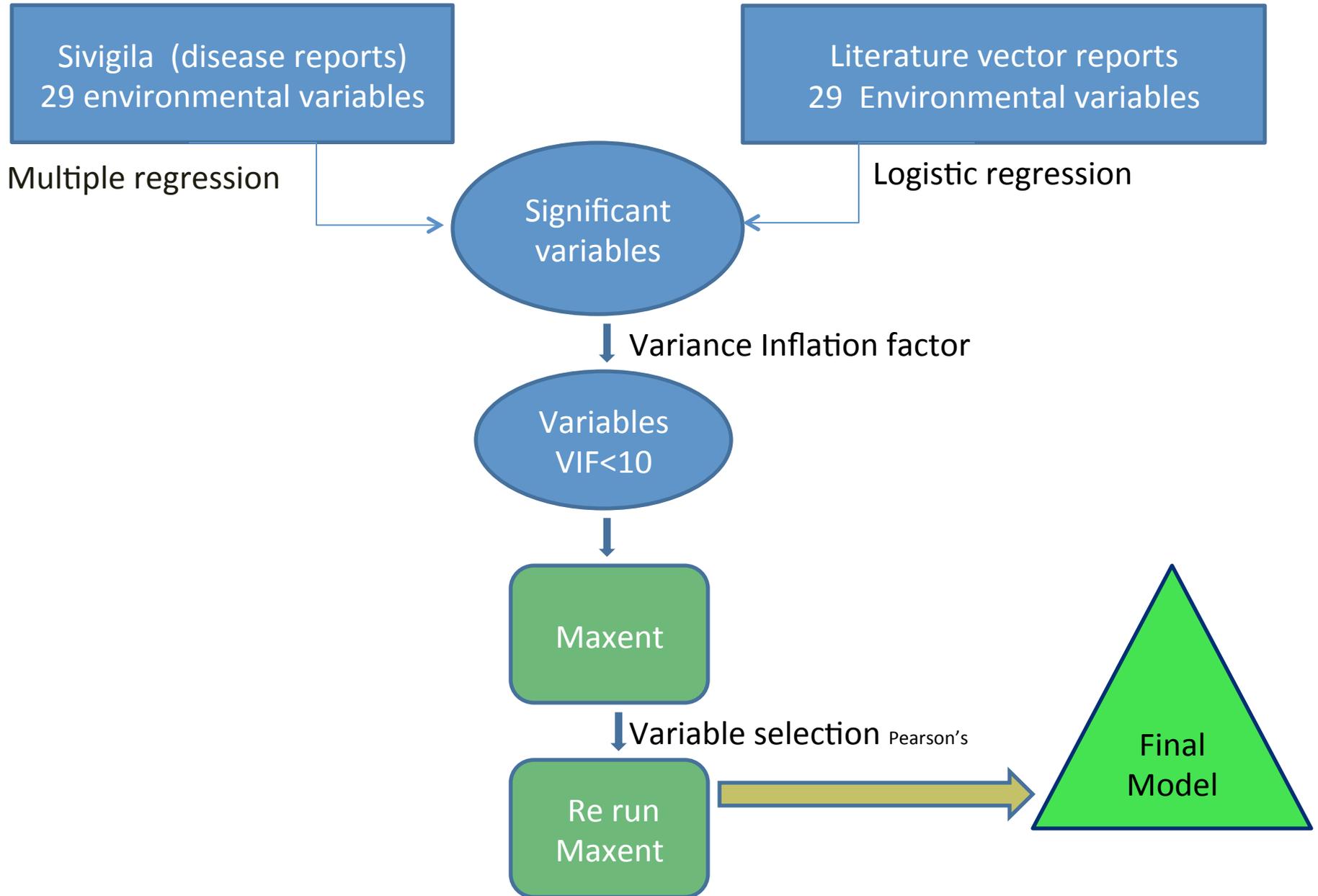
Check out the [Sourceforge page for Maxent](#) for the latest news. You can also ask questions and join in discussions on the [forums](#).
[Download](#) the latest version of Maxent.

[Steven J. Phillips](#), [Robert P. Anderson](#), [Robert E. Schapire](#).

Maximum entropy modeling of species geographic distributions.

Ecological Modelling, 190:231-259, 2006.

Environmental Models



Chagas Disease

Trypanosoma cruzi - 20 million infected in the Americas - Chronic Cardiomyopathy

Circulating Trypomastigote and Tissue Amastigote forms in mammals

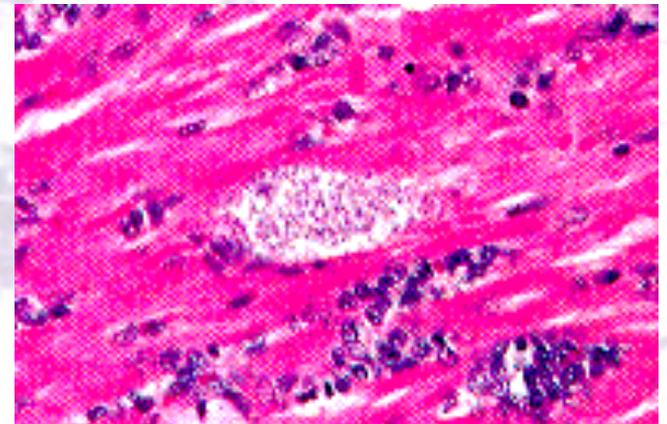
Triatomid 'kissing' bug vectors



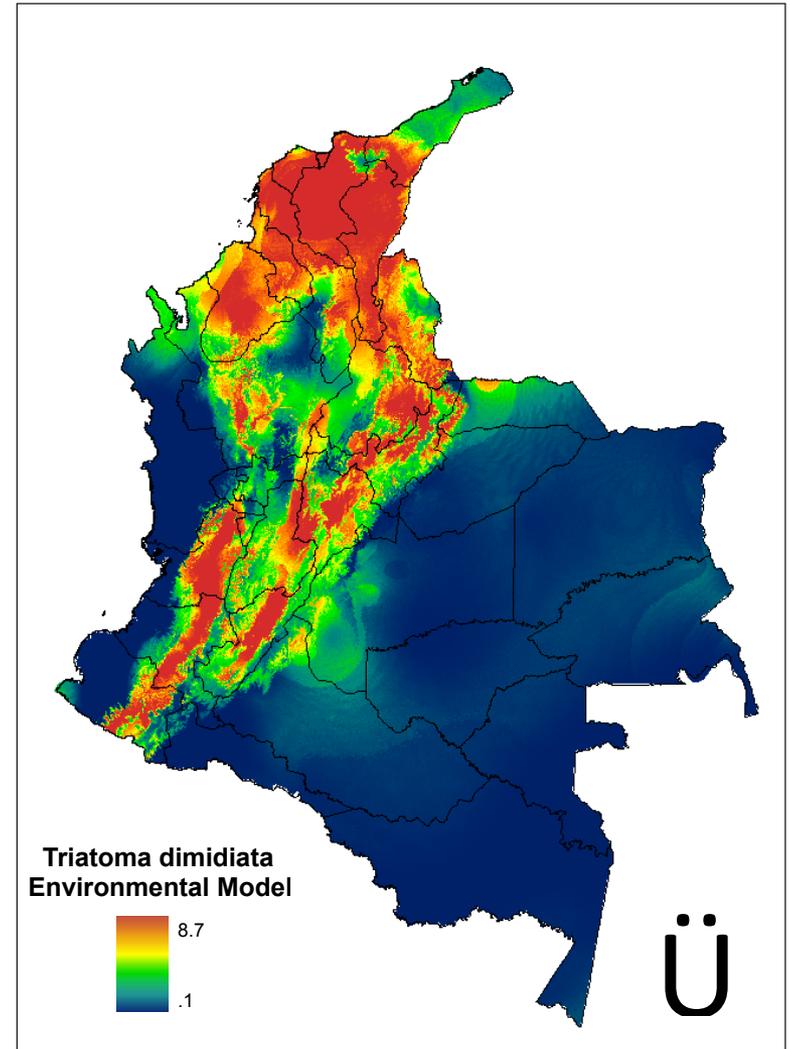
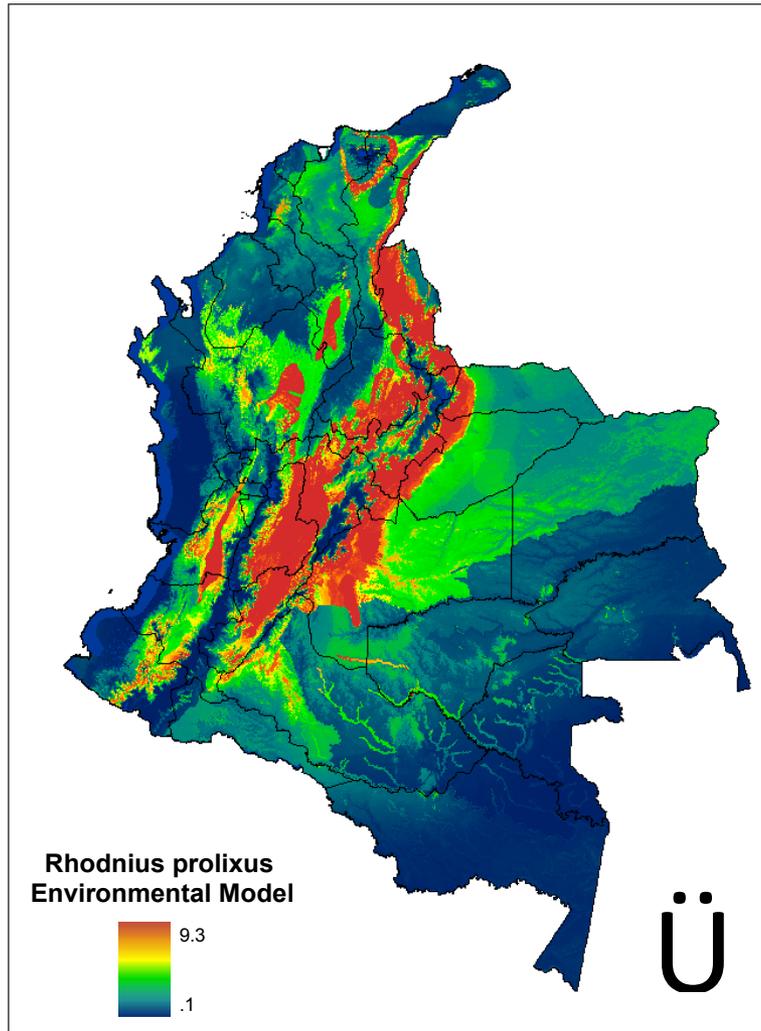
Romana's Sign



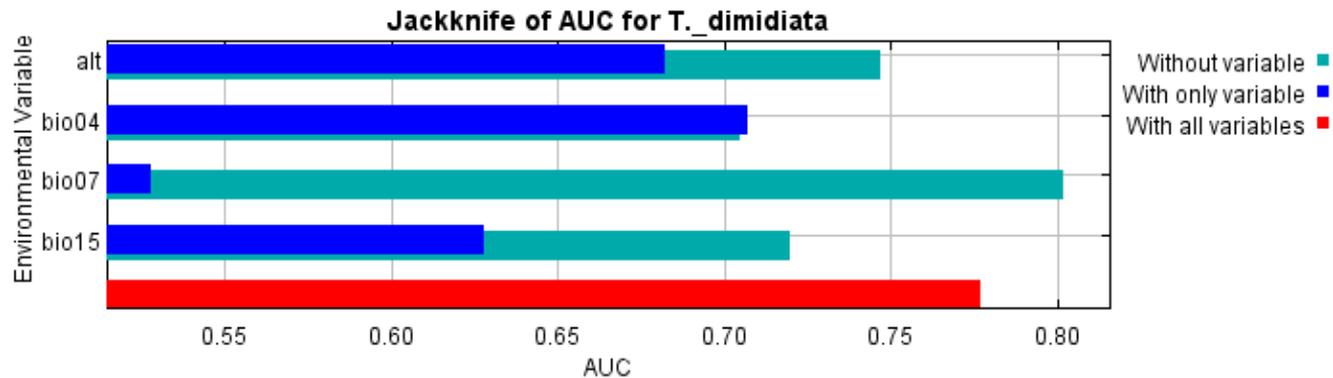
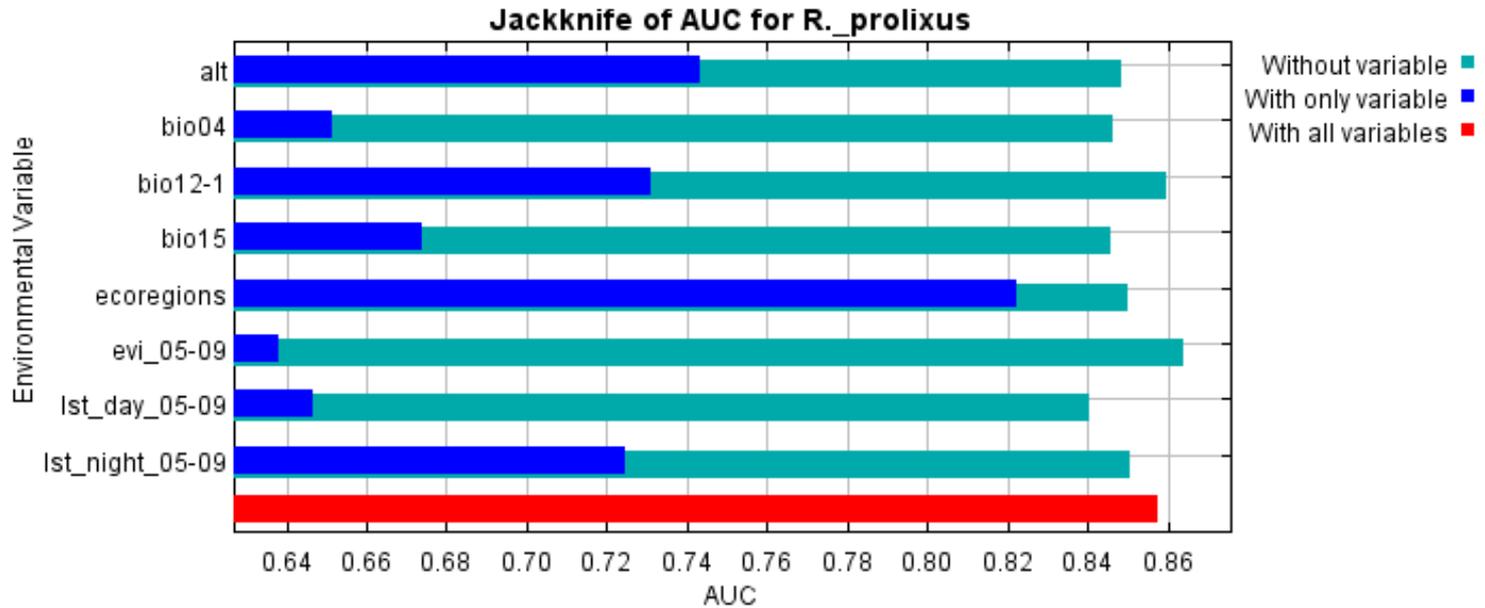
Tissue amastigote form



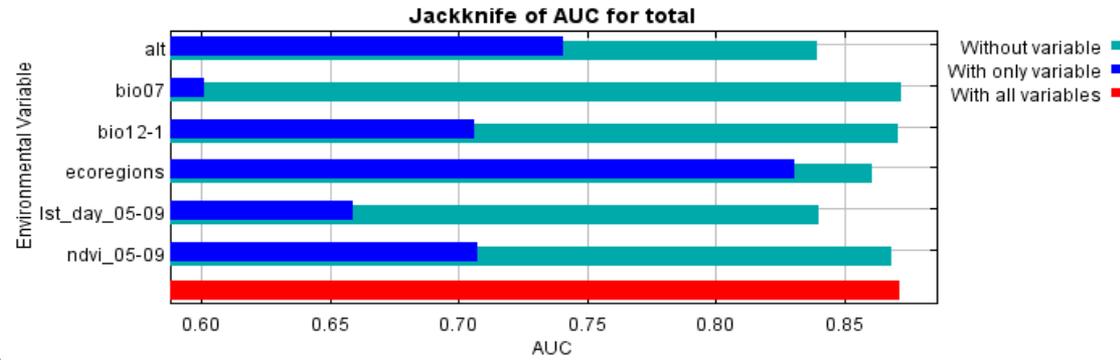
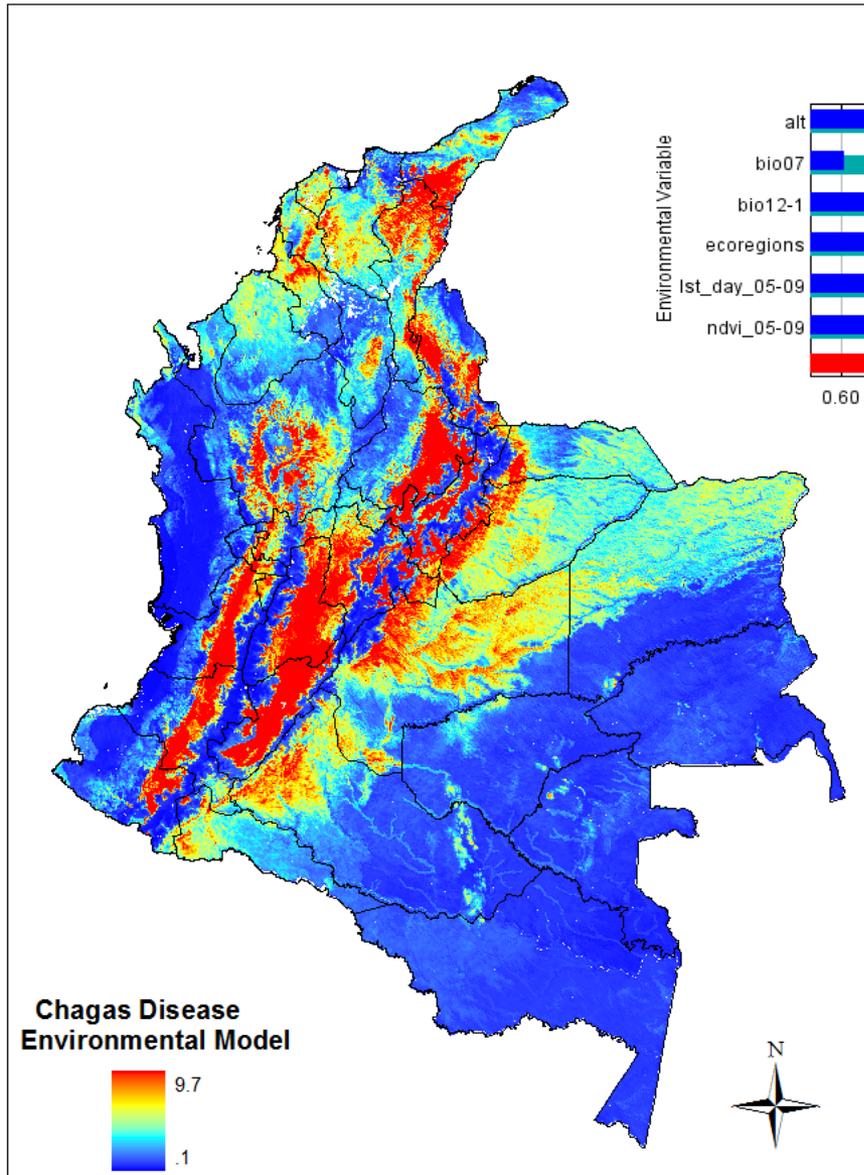
Chagas Vector Distribution



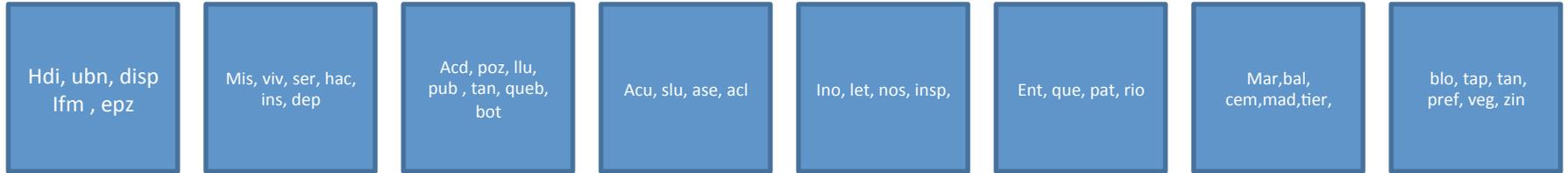
Chagas vectors - Environmental Niche model



Chagas Environmental Niche Model



Socio-Economical Model

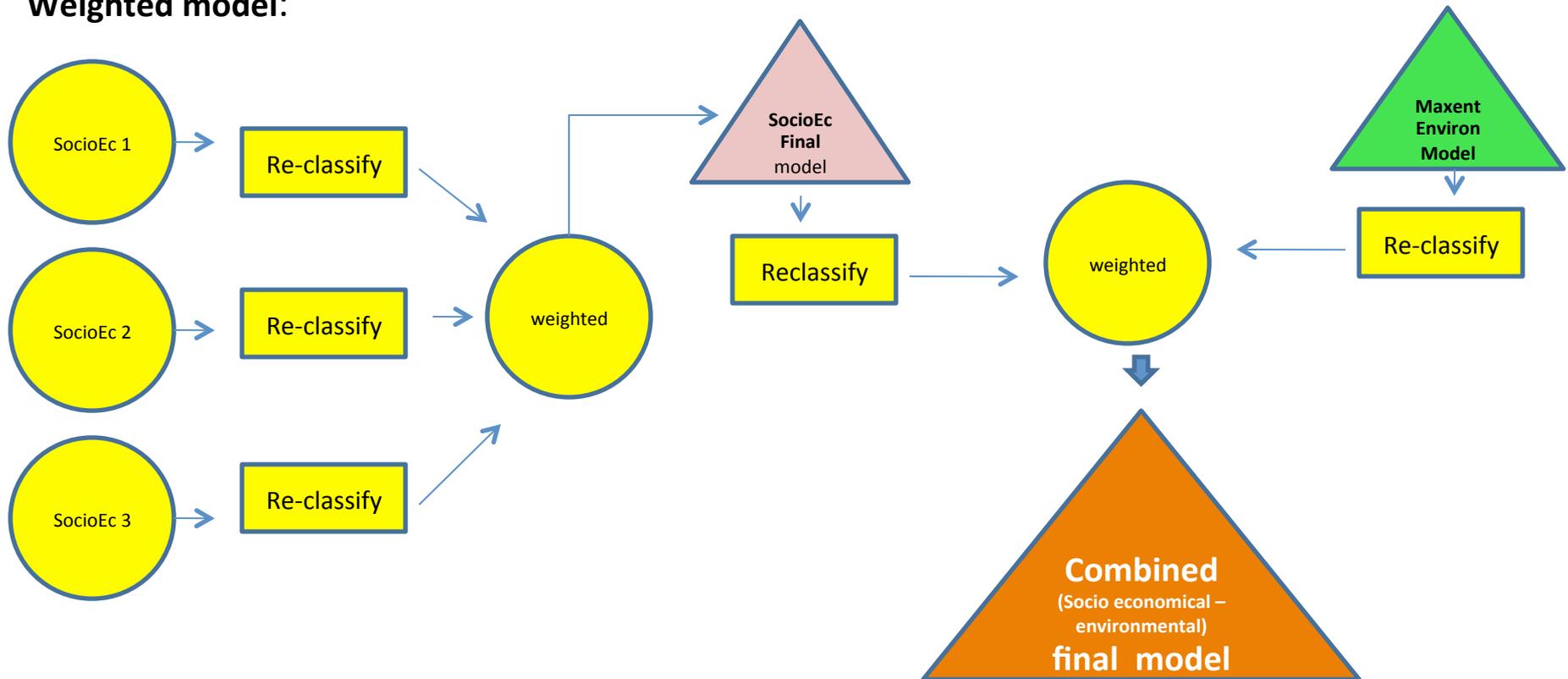


41 socio economical variables divided in 8 groups

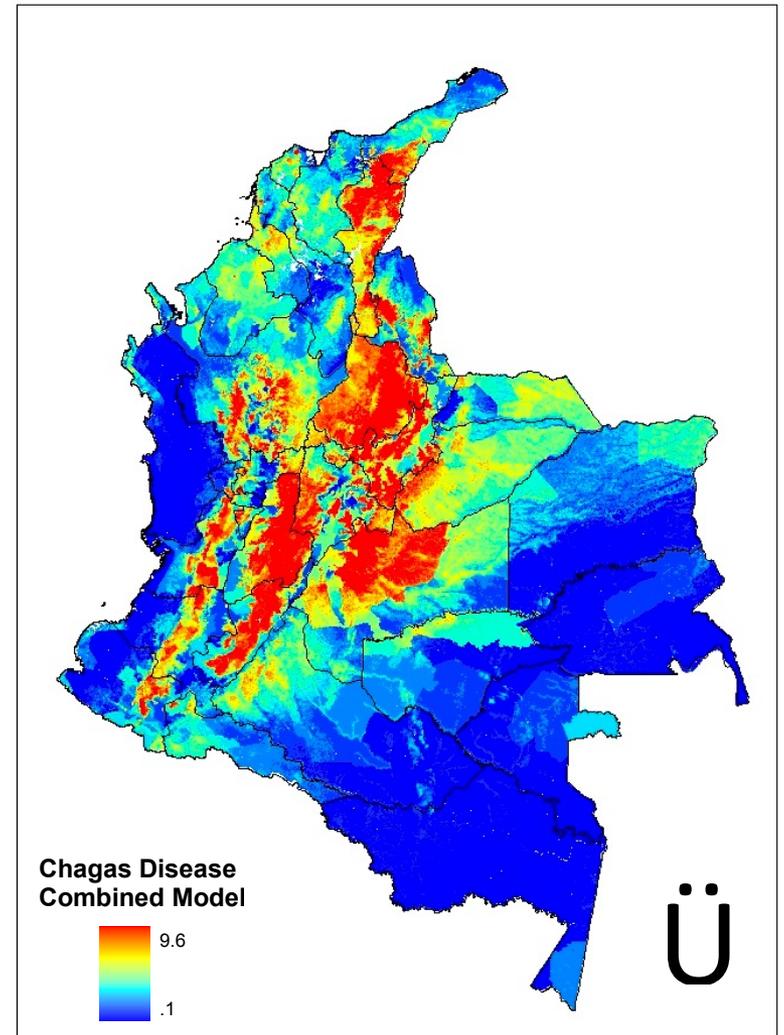
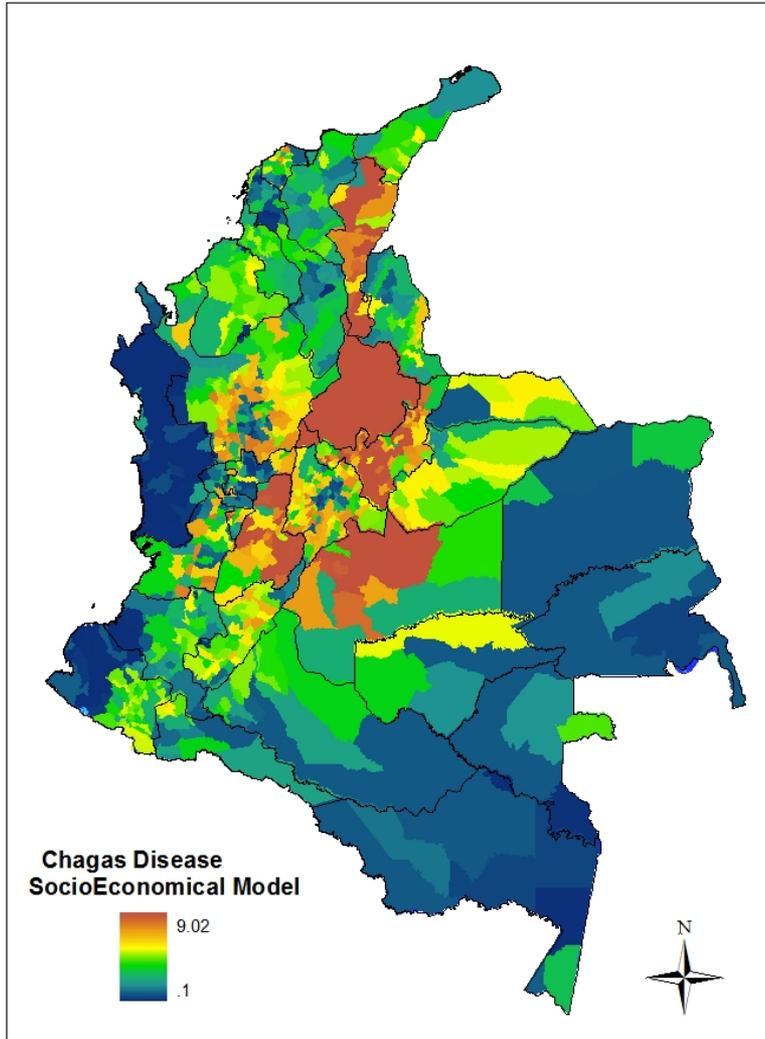
Multiple Regression and VIF

Choose variables for weighted models

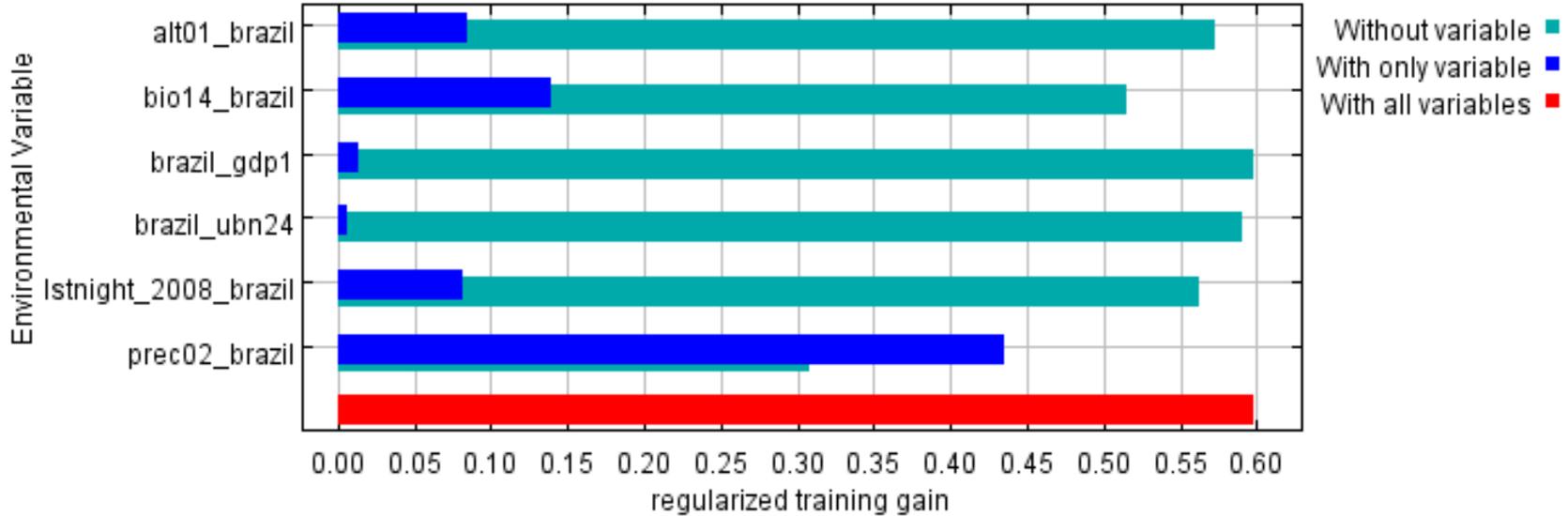
Weighted model:



Socioeconomic Factors – Municipality level



Jackknife of regularized training gain for Trypanosoma_cruzi

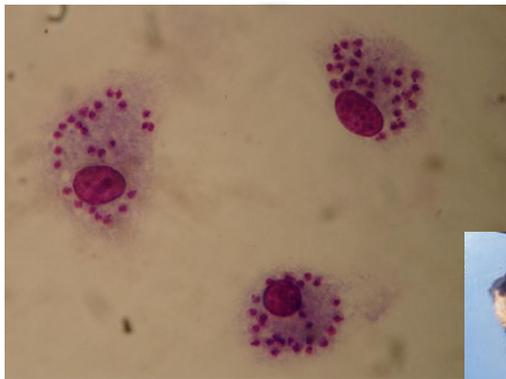


Variable	Percent contribution
prec02_brazil	75.3
bio14_brazil	13.1
alt01_brazil	5.4
Istnight_2008_brazil	4.5
brazil_ubn24	1.1
brazil_gdp1	0.7

Visceral Leishmaniasis

Caused by protozoans of the genus *Leishmania*

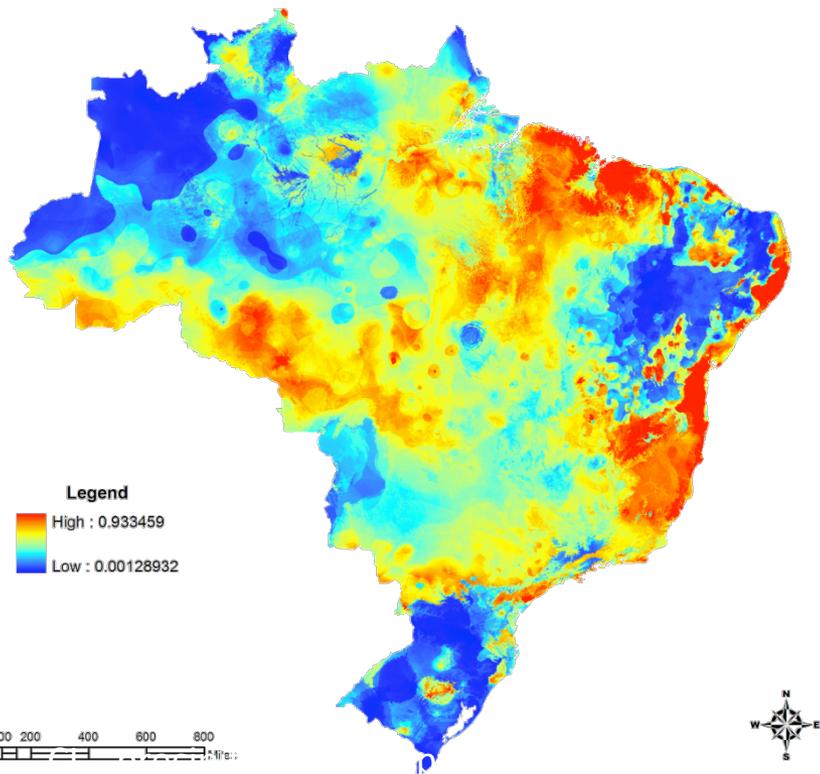
- Amastigote form – mammals
- Promastigote form – arthropod vector



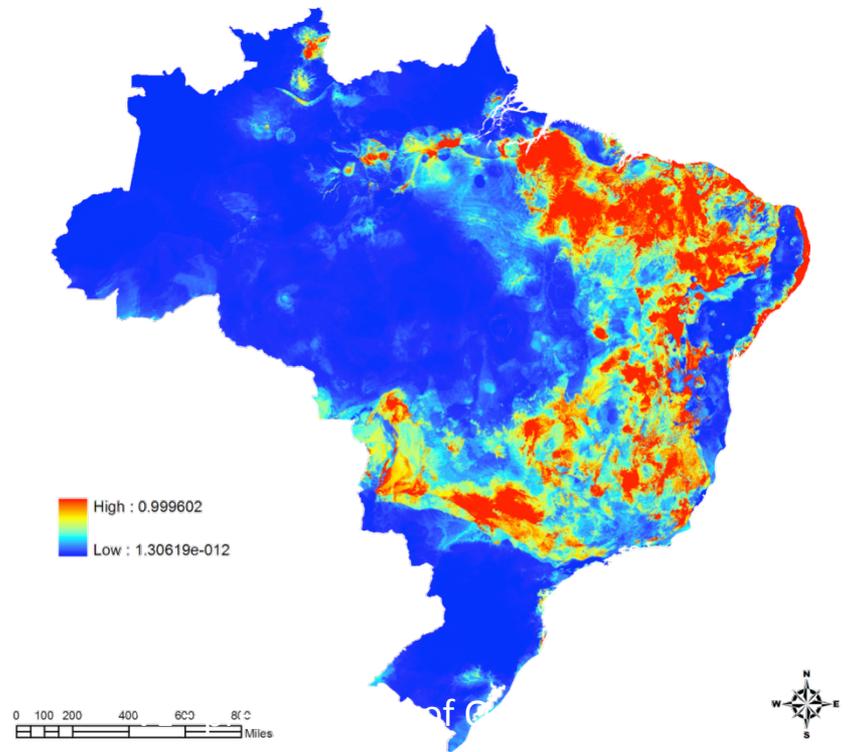
Sandfly vector (*Lutzomyia*)

Leishmaniasis – Visceral and Cutaneous

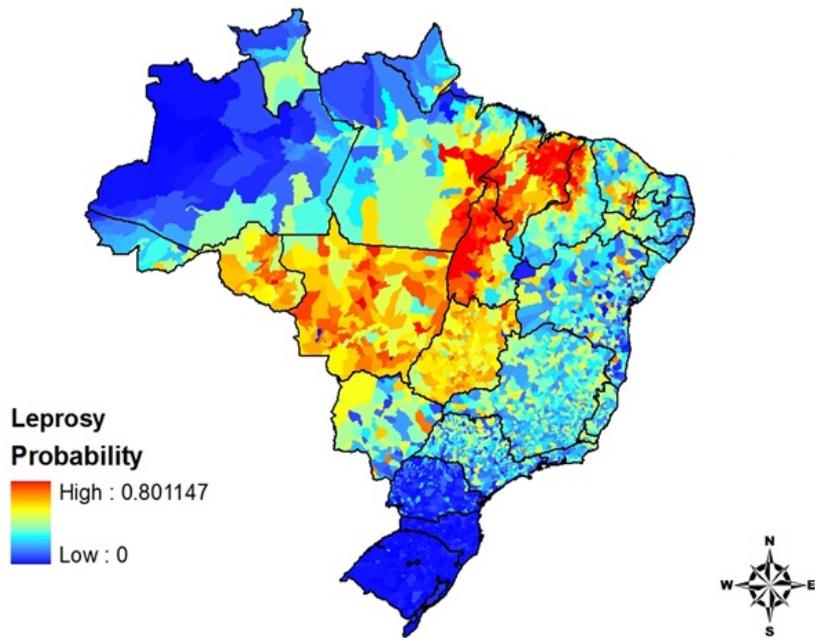
Maxent Environmental Model using Worldclim data
Cutaneous Leishmaniasis



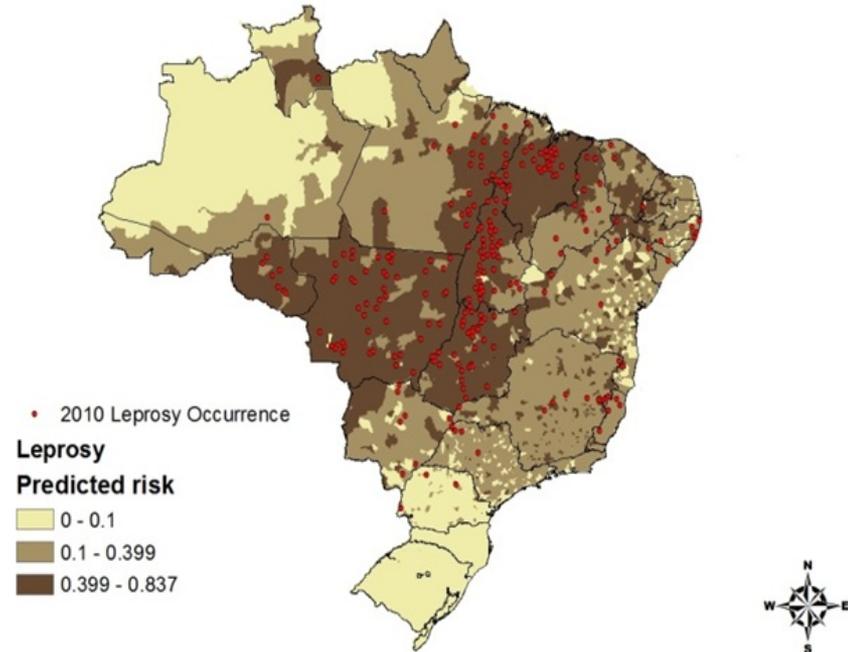
Maxent Environmental Model using Worldclim data
Visceral Leishmaniasis



Leprosy in Brazil

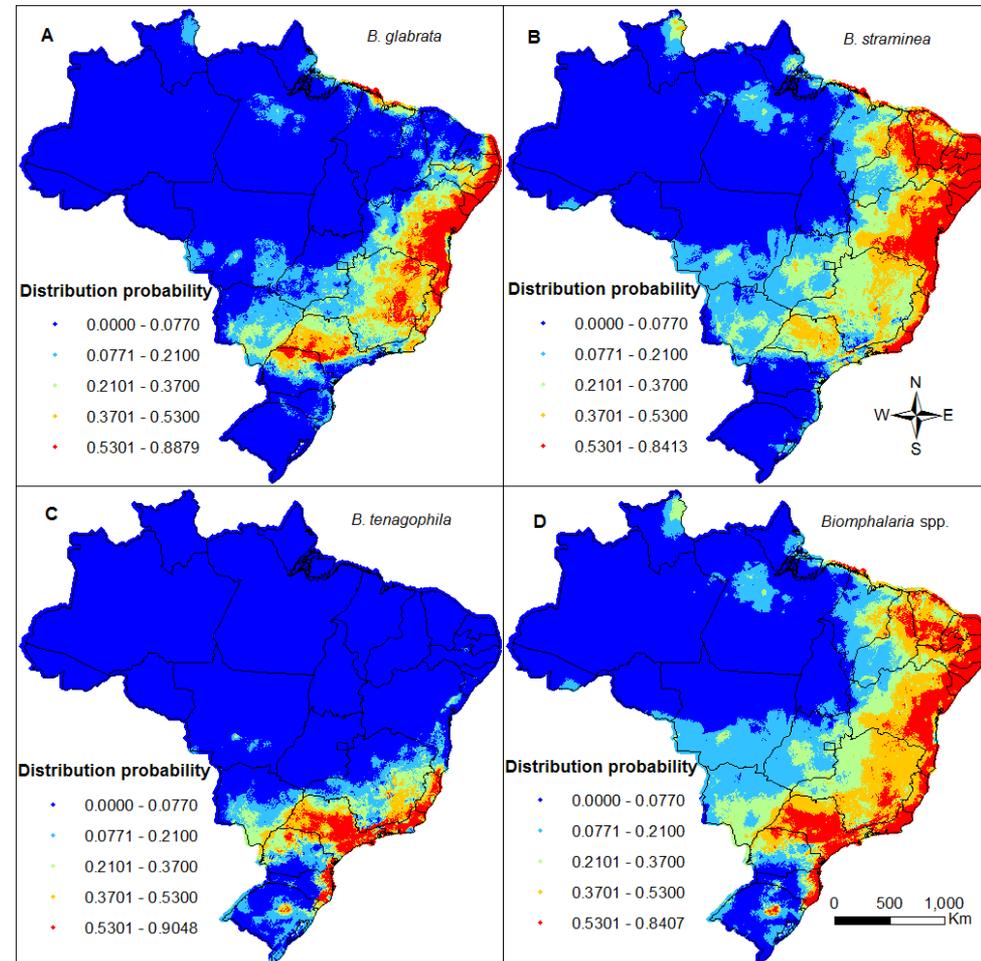
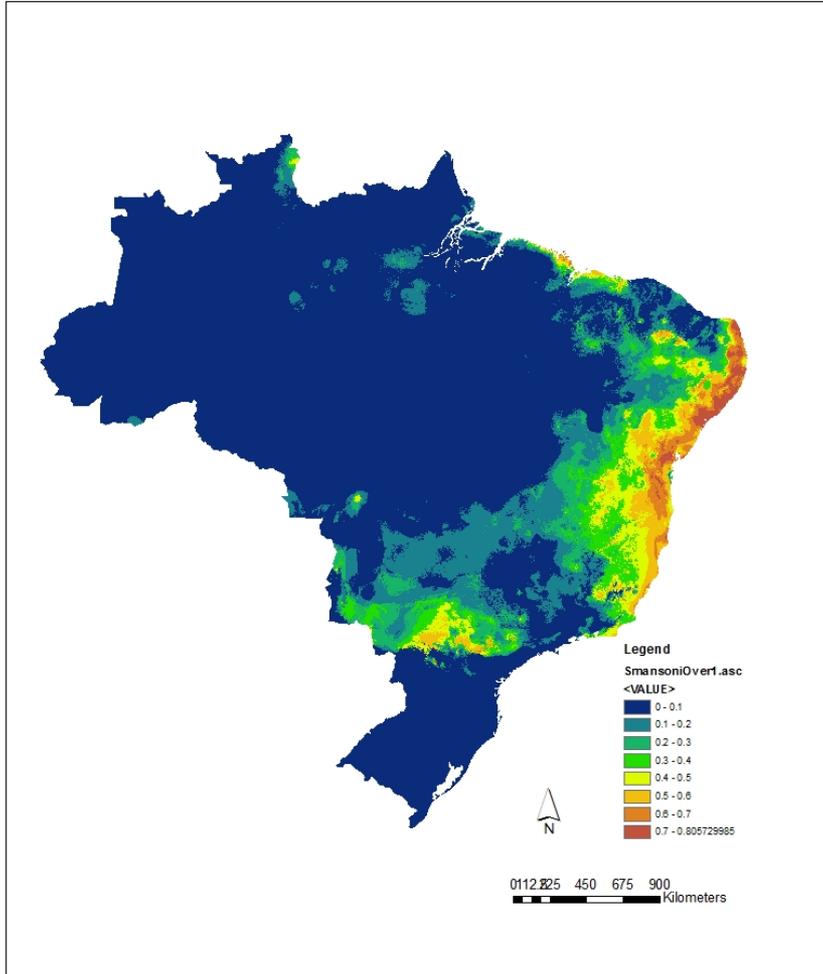


Maxent predictive model showing the distribution probability of leprosy occurrence. Red indicates a higher probability of occurrence, while blue indicates a low probability of occurrence.



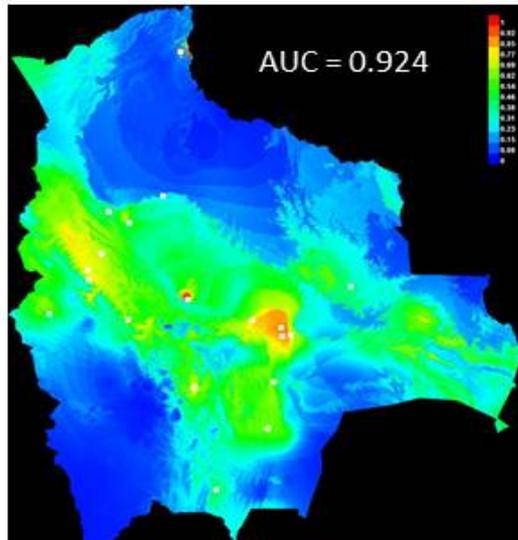
The predicted risk map of leprosy overlaid with 2010 leprosy occurrence data.

Schistosomiasis



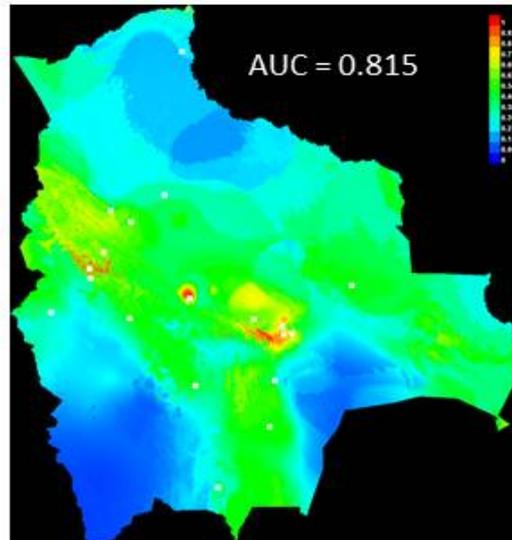
Hookworm in Bolivia

Worldclim Model



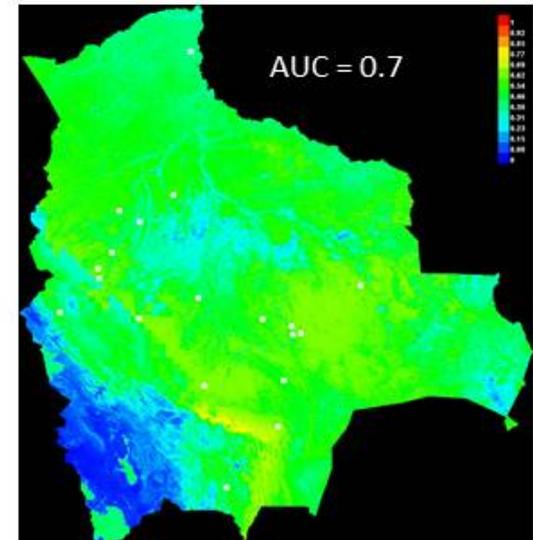
Variable	% Contribution
tmax06	28.2
prec01	18.4
prec08	13.3
tmax04	9.5
tmin05	8.9
tmin04	7.1
tmin03	3.9
prec03	3.0
prec04	2.0
prec05	1.8
prec09	1.7

Bioclim Model



Variable	% Contribution
Bio13 - Precipitation wettest month	37.2
Bio15 - Precipitation Seasonality	29.4
Bio03 - Isothermality	13.0
Bio09 - Mean temperature driest quarter	11.5
Bio01 - Annual mean temperature	6.3
Bio02 - Mean diurnal range	2.1
Bio07 - Temperature annual range	0.3
Bio12 - Annual precipitation	0.2

MODIS Model



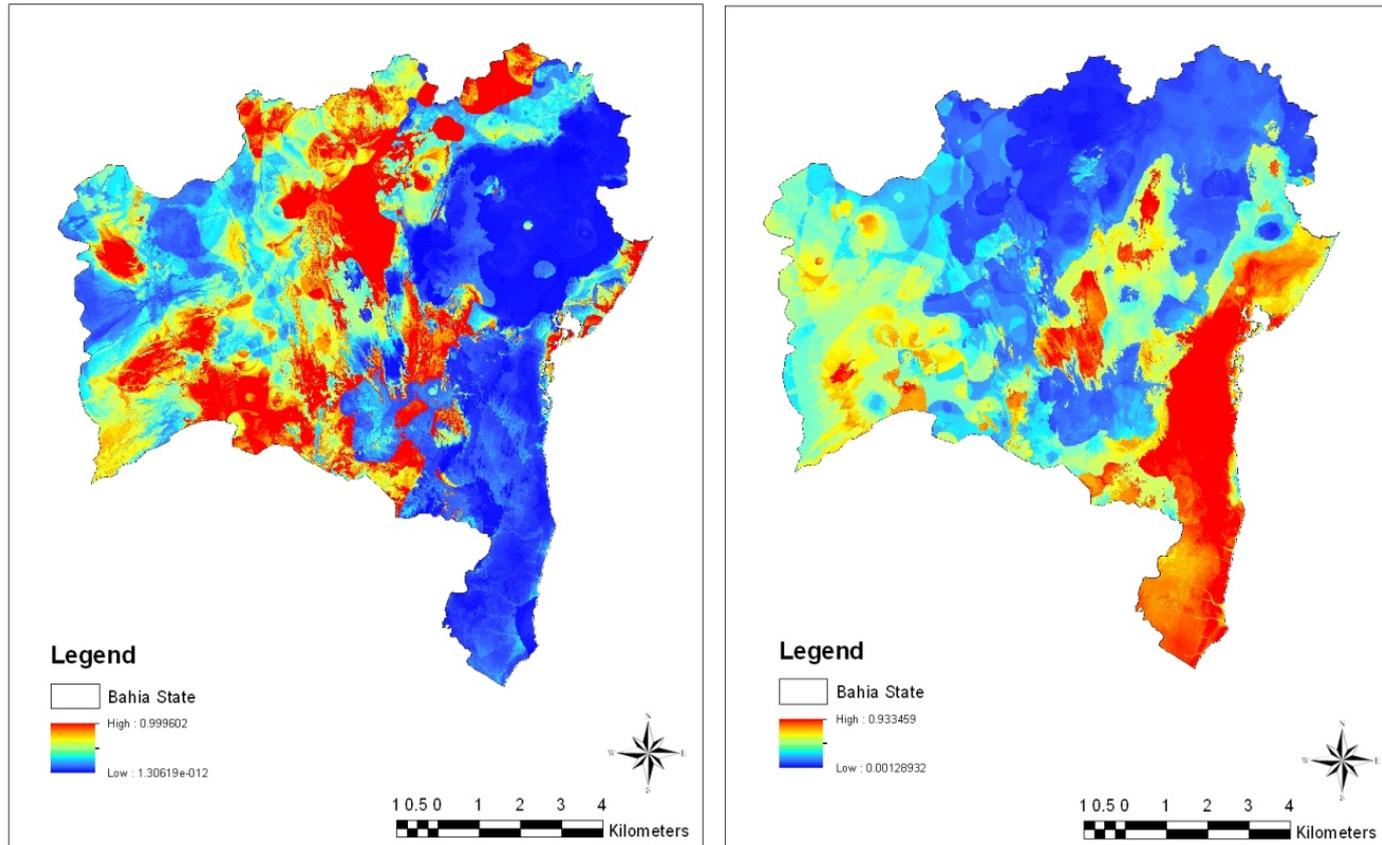
Variable	% Contribution
GDD/WB Gradient	44.3
LST day	21.8
LST night	20.3
Temperature Difference (dT)	12.8
Enhanced Vegetation Index	0.7

Conclusions and Recommendations

1. Maxent Ecological Niche Modeling is a useful tool to guide surveillance and control programs for NTD, particularly where health surveillance data are scarce
2. Extrapolation of risk surfaces is of limited validity where representative survey data are absent in a given ecosystem
3. Socioeconomic data or poverty indicators should be at the census tract level; Municipality level data is typically too heterogeneous
4. Results of Maxent ecologic niche mapping and modeling should be validated by alternative methods eg. biology based GDDxWB climate models

Future Work

Maxent distribution probability models for Leishmaniasis in Bahia



Maxent generated risk surfaces extracted for Bahia from national scale maps on visceral leishmaniasis (a) and cutaneous leishmaniasis (b) using MODIS environmental satellite annual composite data on vegetation index (EVI) and land surface temperature (LST).

Select High, Medium, Low Risk municipalities (5 each) using SINAN case reports, vector records



Community Profile modeling System

Environmental
(15-30 m²)

Vulnerability
(census block)

Climate

Hydrology

Landuse

Poverty

Population
#/Density/
migration

Exposure/
occupation

Reservoir
Hosts

Vectors



Local Intervention Scenarios

Vector Control

Reservoir control

Surveillance Planning