

---

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

JB Malone, P Nieto, P Mischler JC McCarroll– Louisiana State University, USA

Penelope Vounatsou – Swiss TPH, Switzerland

Mara Bavia – Universidade Federal da Bahia, Brazil

Thomas Kristensen – University of Copenhagen, Denmark

*ISPRS-wg8 , Santa Fe  
September 12, 2011*

Swiss TPH



Swiss Tropical and Public Health Institute  
Schweizerisches Tropen- und Public Health-Institut  
Institut Tropical et de Santé Publique Suisse

**LSU**



# Objectives

---

---

- Data Portal – A municipality level 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 the 6 NTD at municipality scales
- Course Development – A 4-day short course was developed for 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) is used for ecological Niche modeling and by climate change community


MODIS EVI, LST annual composites for 2005-2009

Socioeconomic Data at the Municipality level

### PAHO Data Portal

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

**COLOMBIA**



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.
<b>Images</b>	<a href="#">Bioclim Variables</a>
MODIS 2003 Composites	<a href="#">Altitude</a>
EVI <a href="#">image</a> - <a href="#">ascii</a>	<a href="#">Min Temperature</a>
Temp Day image - ascii	<a href="#">Mean Temperature</a>
Temp Day image - ascii	<a href="#">Mean Temperature</a>
Temp Night	<a href="#">Max Temperature</a>


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.

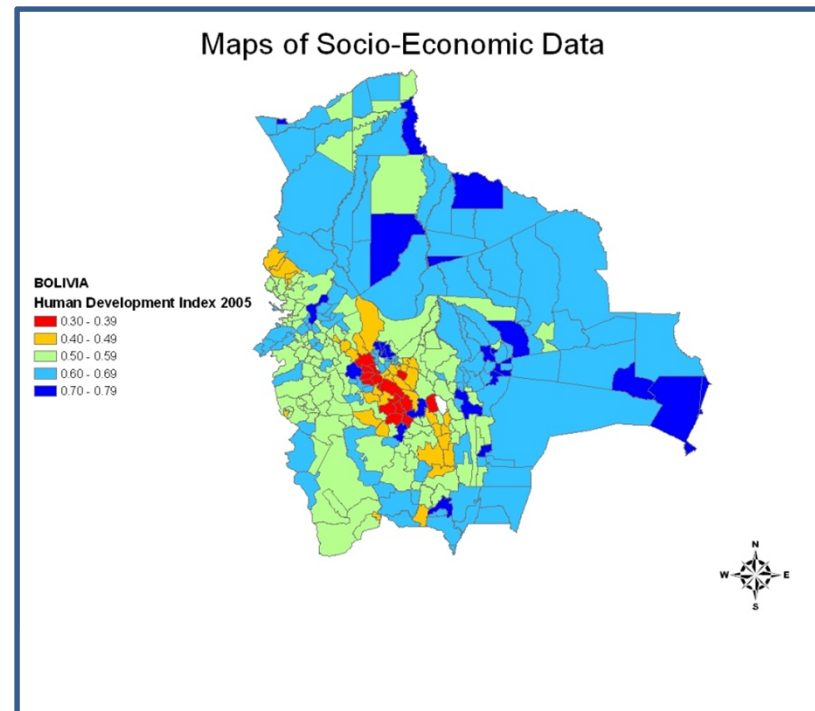
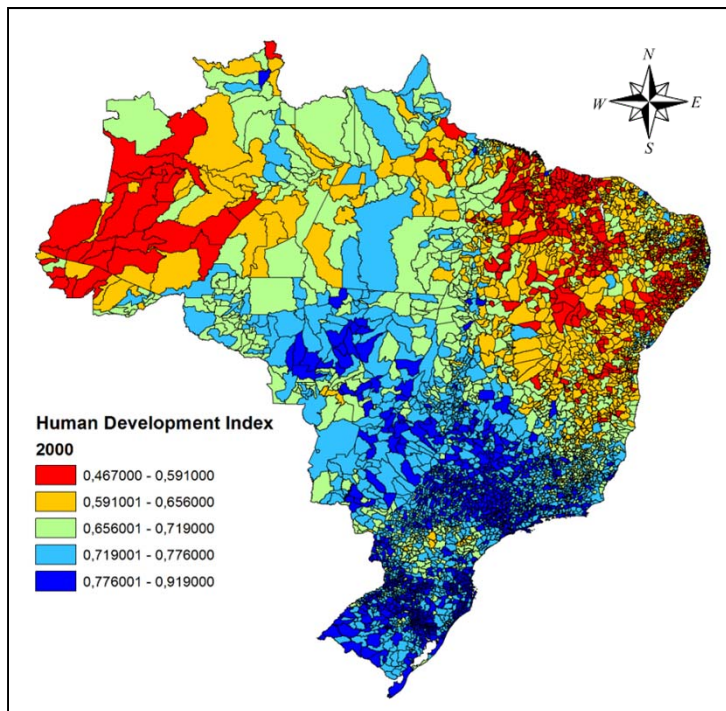
## **Infrastructure**

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

## **Political Boundaries**

Counties, major cities, States/Departments, Municipalities

## Maps of socio-economic data



Unsatisfied Basic Needs for Brazil and Bolivia, via PAHO

# 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](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.**

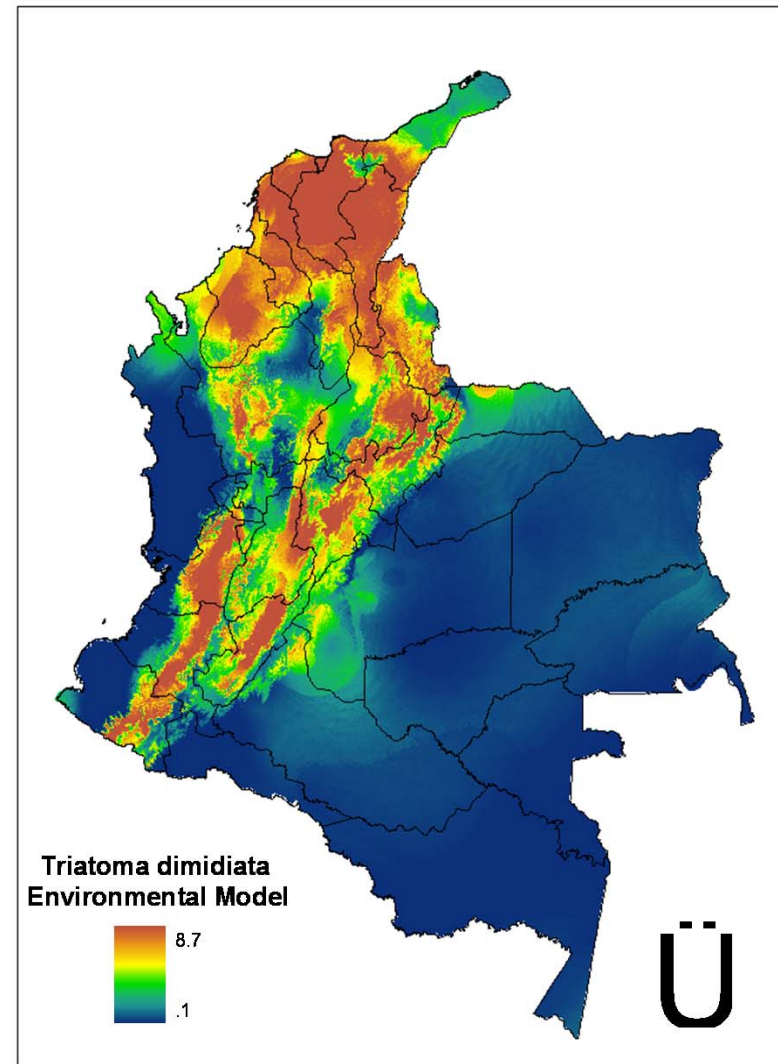
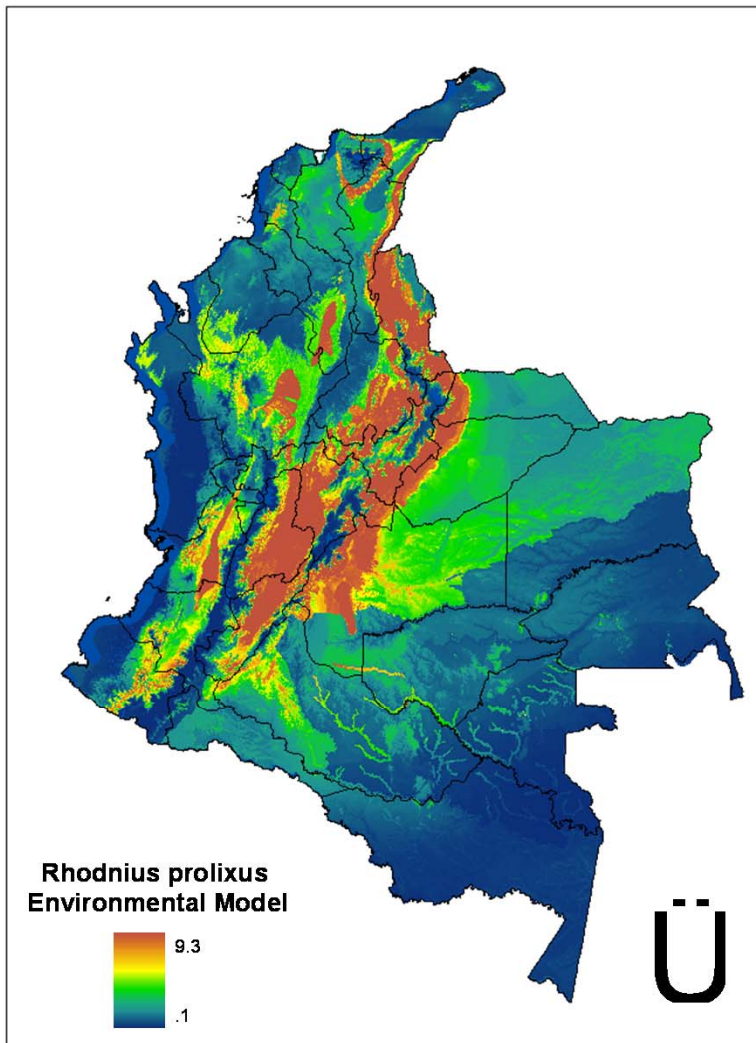
*Ecological Modelling*, 190:231-259, 2006.

---

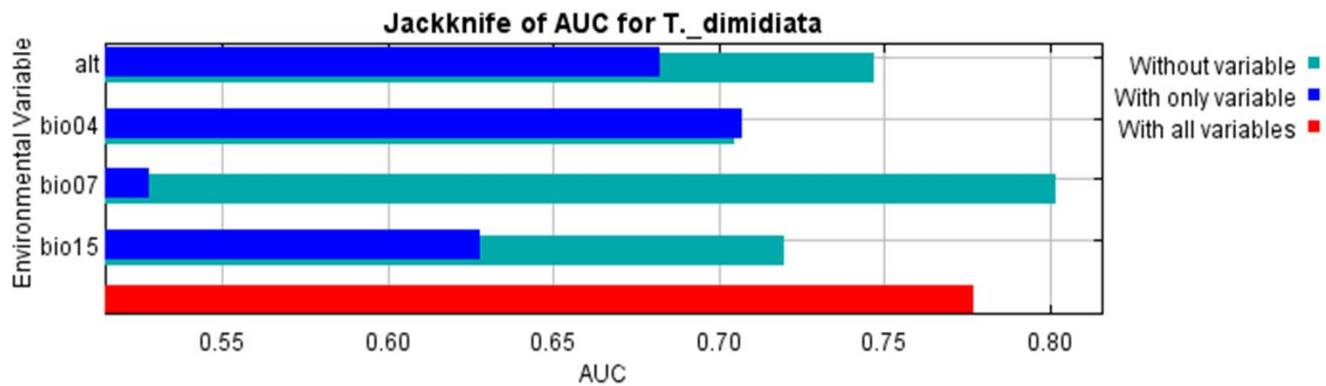
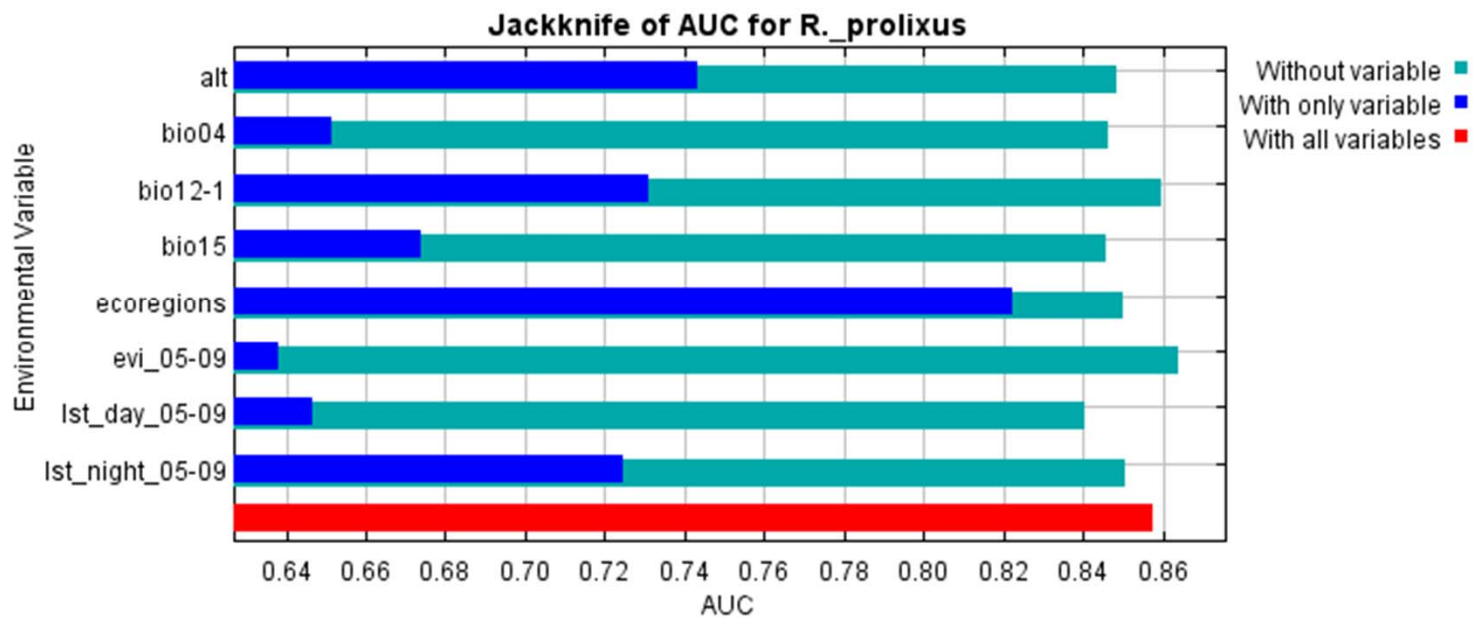
---



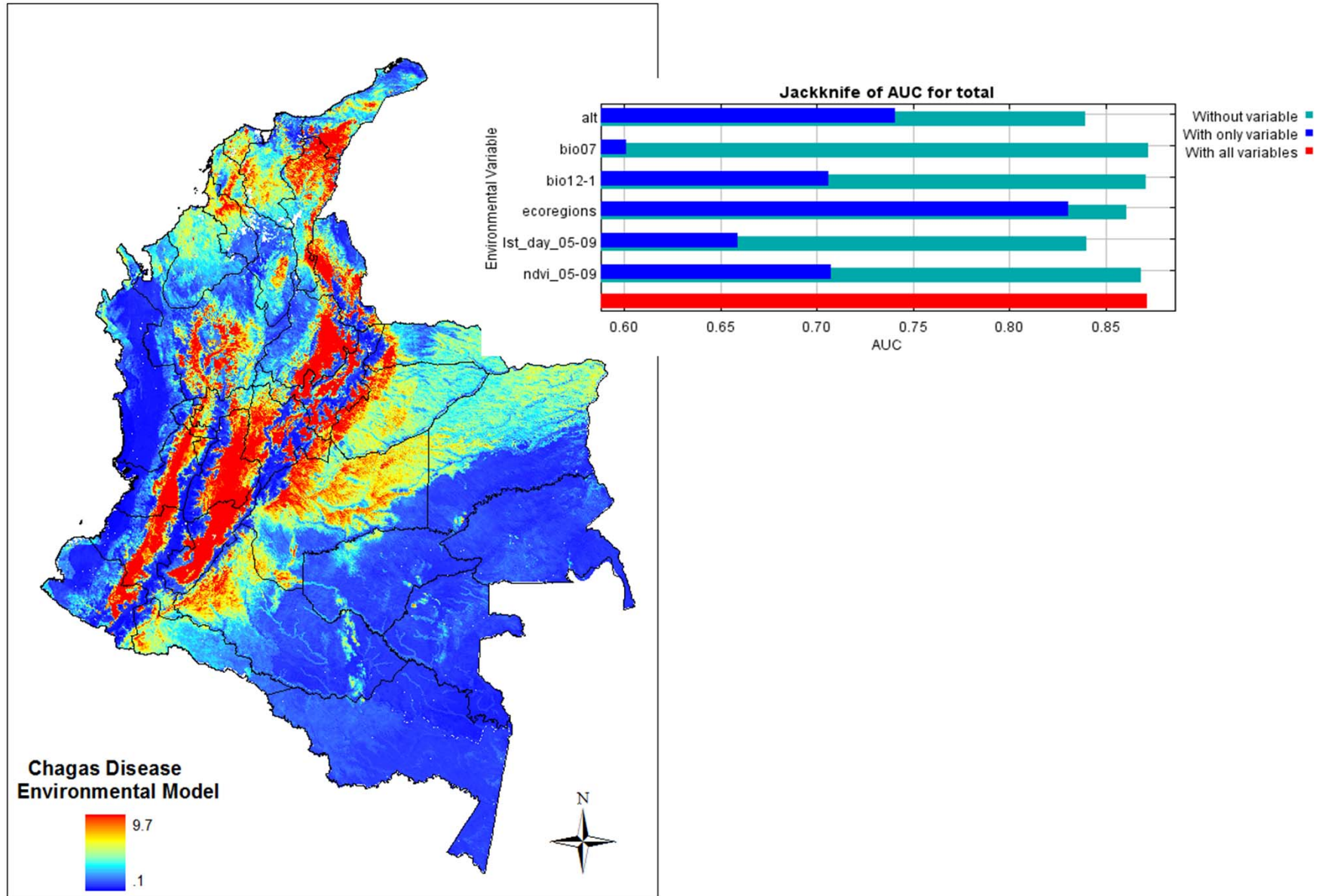
# Chagas Vector Distribution



# Chagas vector - Environmental Maxent model



# Chagas Environmental Niche Model



# Environmental Models

Sivigila (disease reports )  
29 environmental variables

Literature vector reports  
29 Environmental variables

Multiple regression

Logistic regression

Significant variables

↓ Variance Inflation factor

Variables  
VIF < 10

↓

Maxent

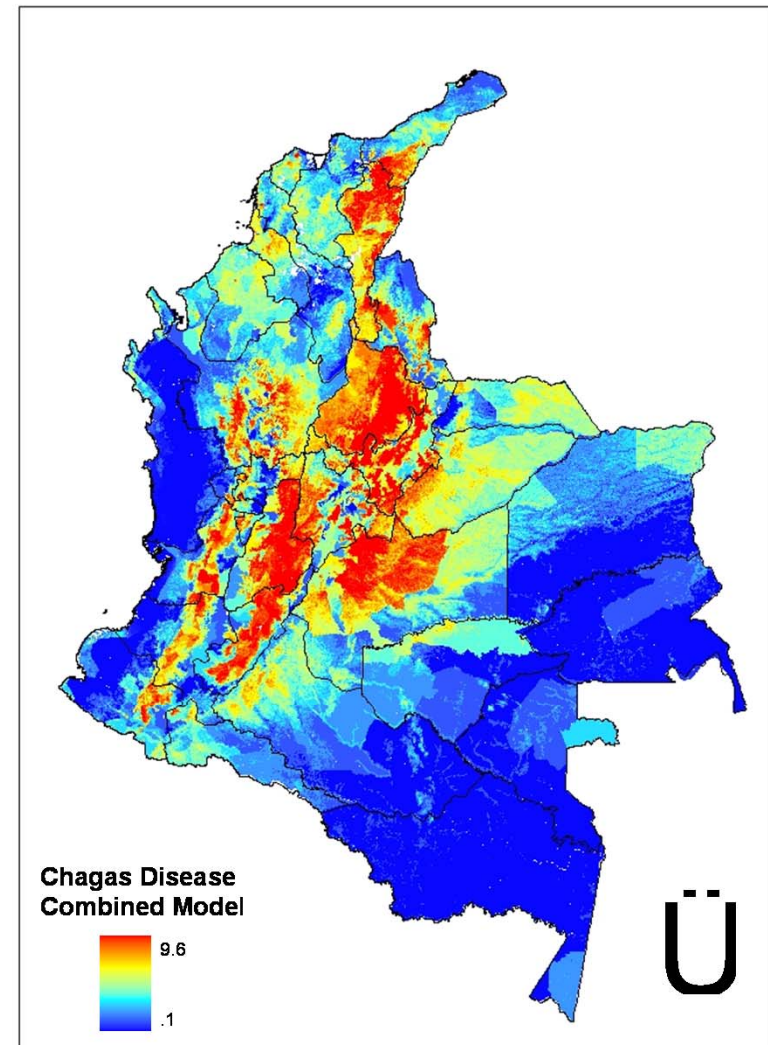
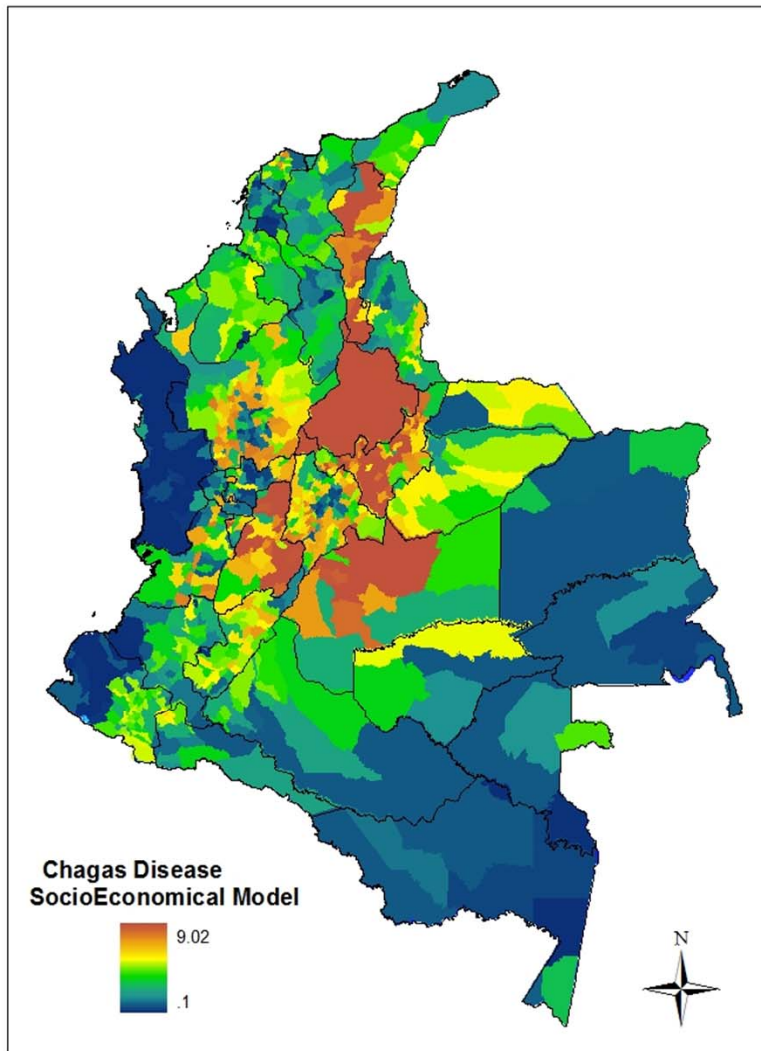
↓ Variables selection Pearson's

Re run  
Maxent

→

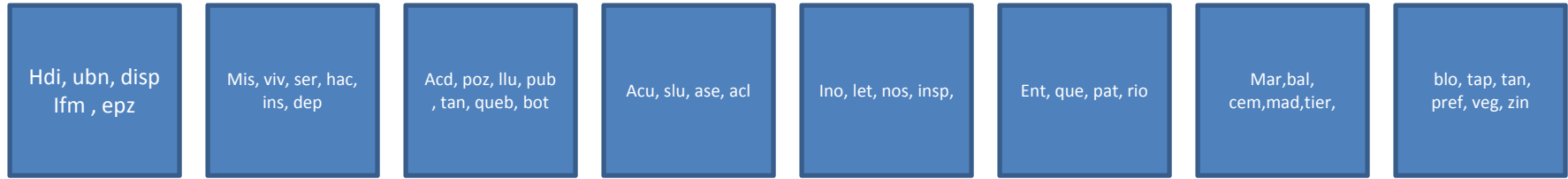
Final Model

# Socioeconomic Factors – Municipality level





# Socio-Economical Model

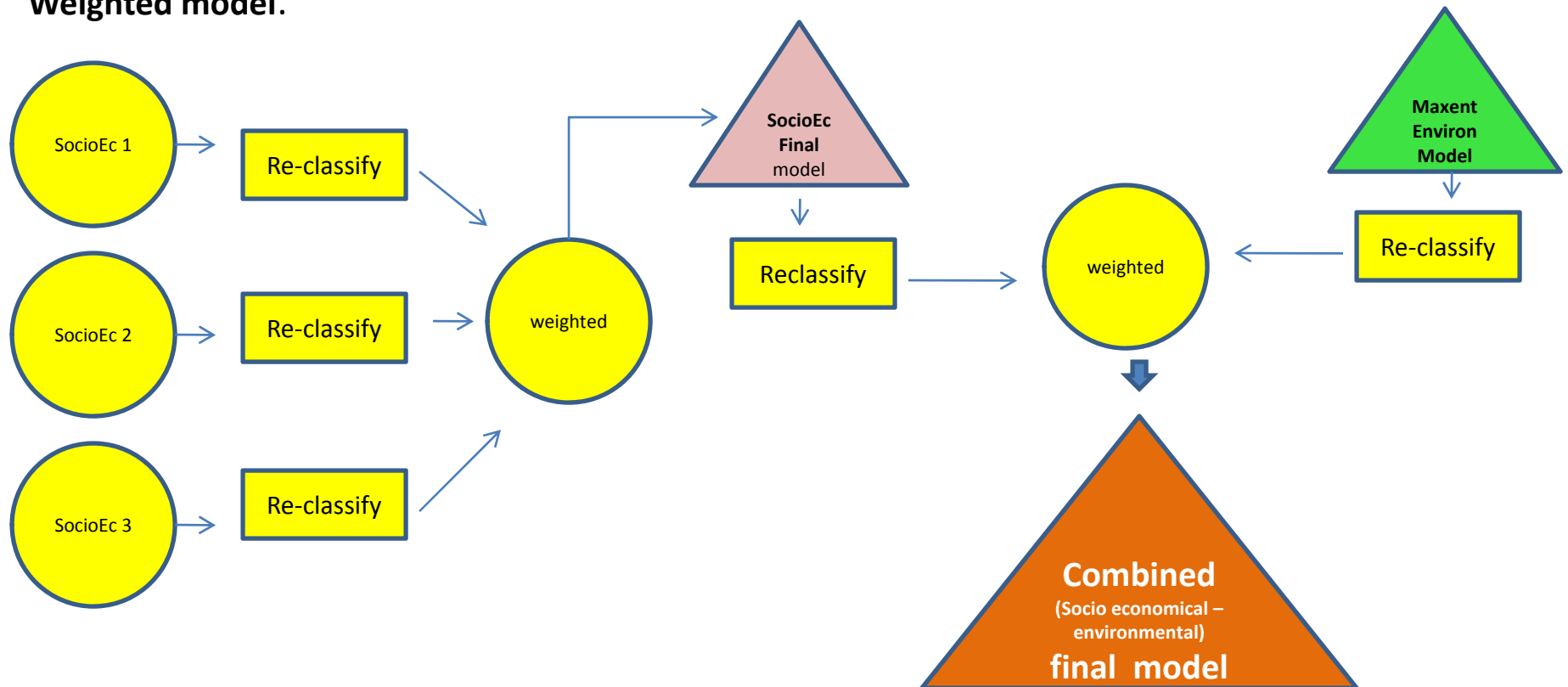


41 socio economical variables divided in 8 groups

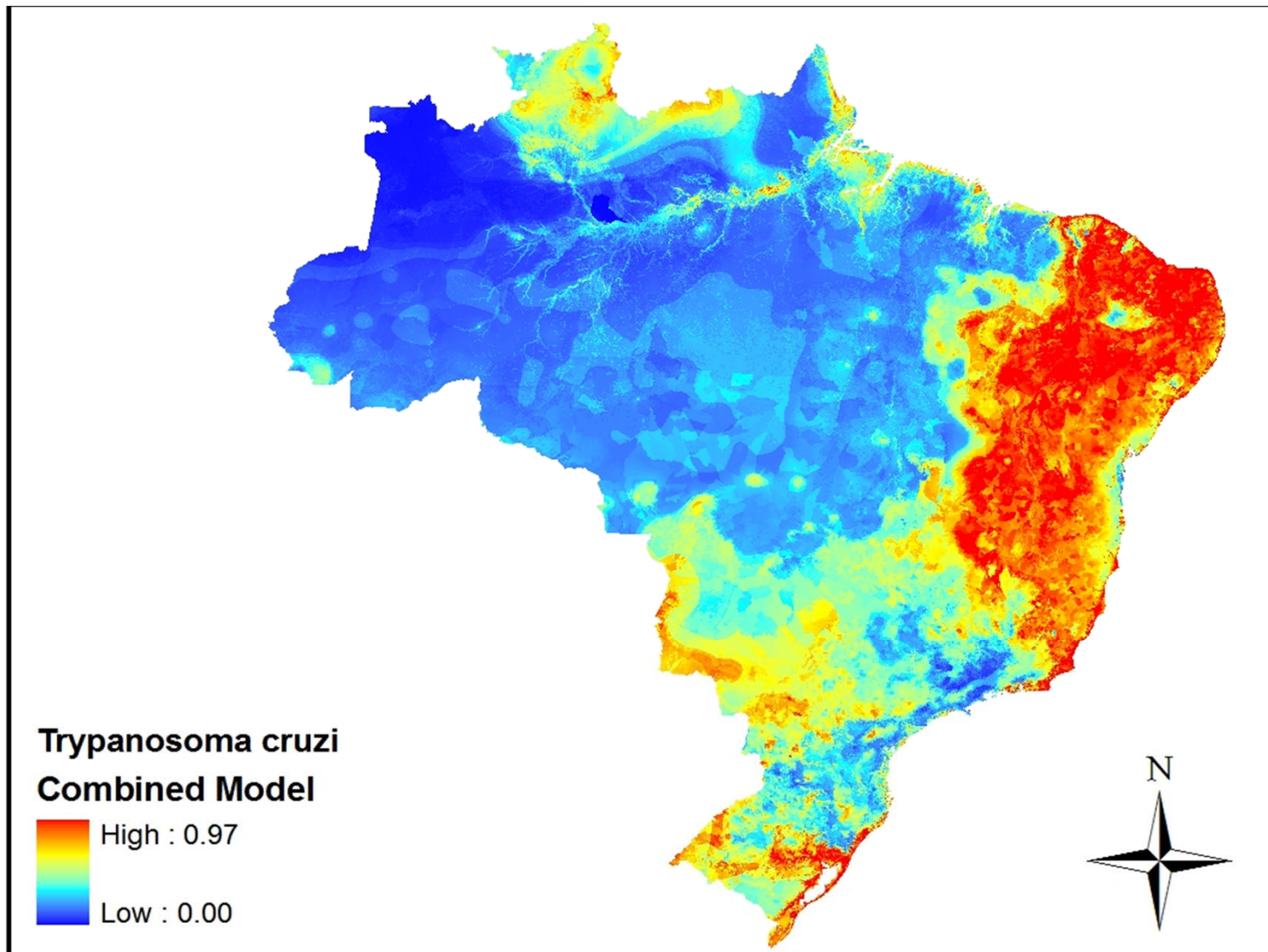
Multiple Regression and VIF

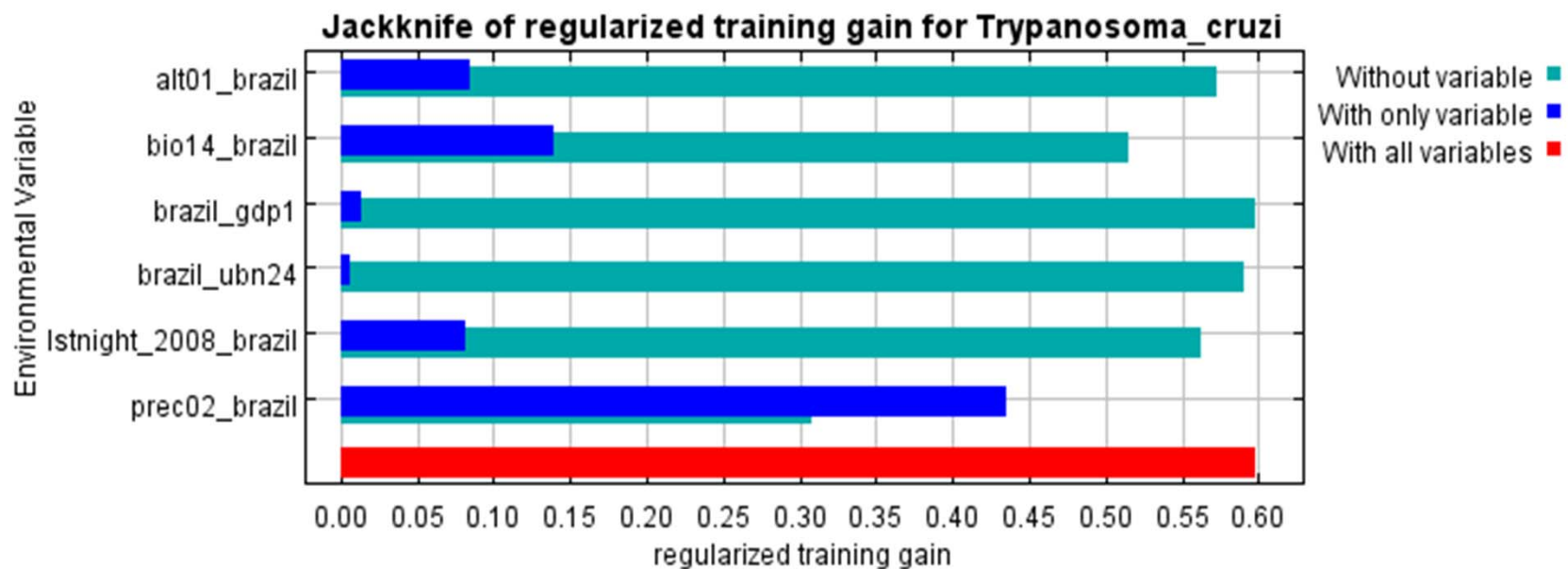
Choose variables for weighted models

## Weighted model:



## Combined model of the potential distribution of Chagas in Brazil

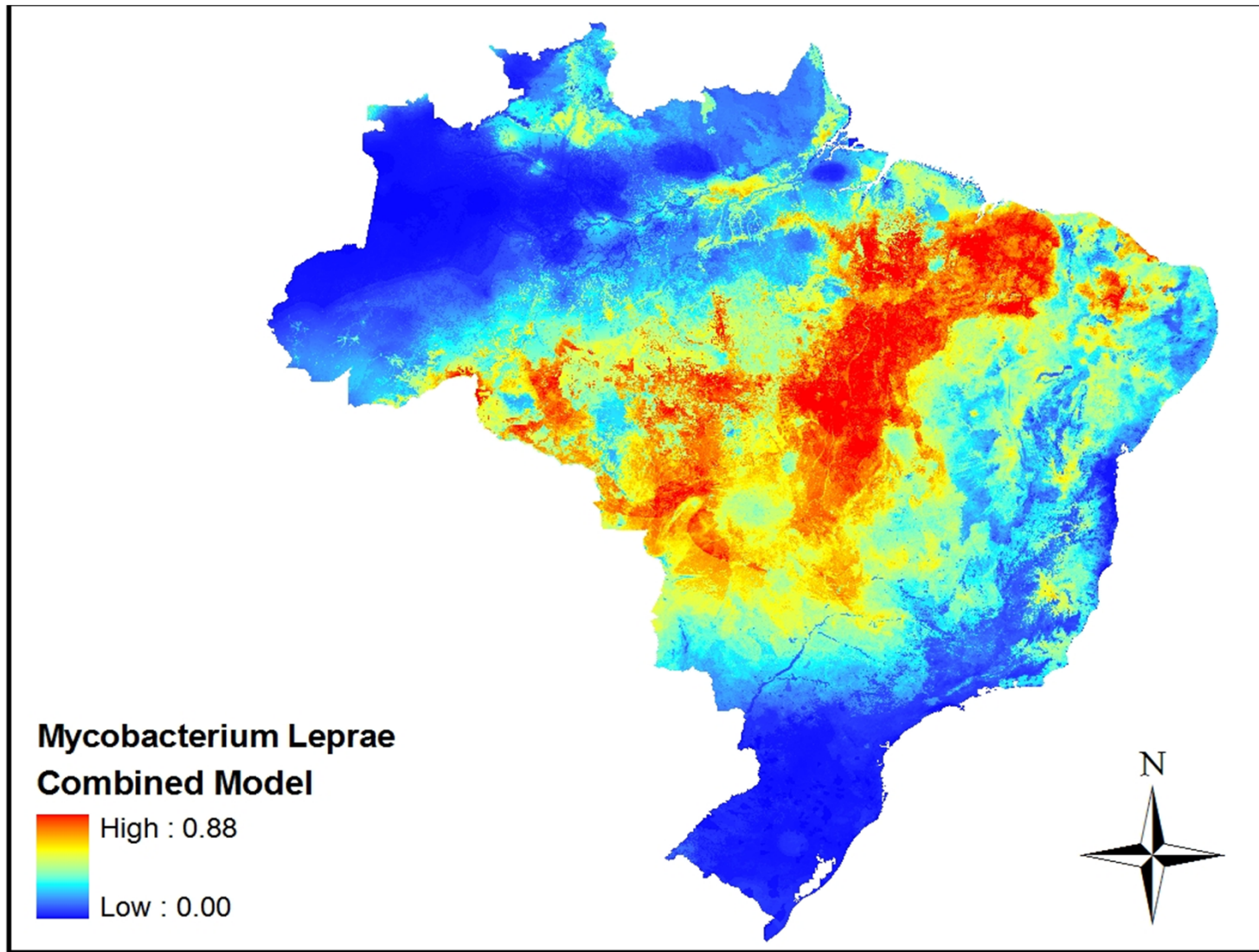




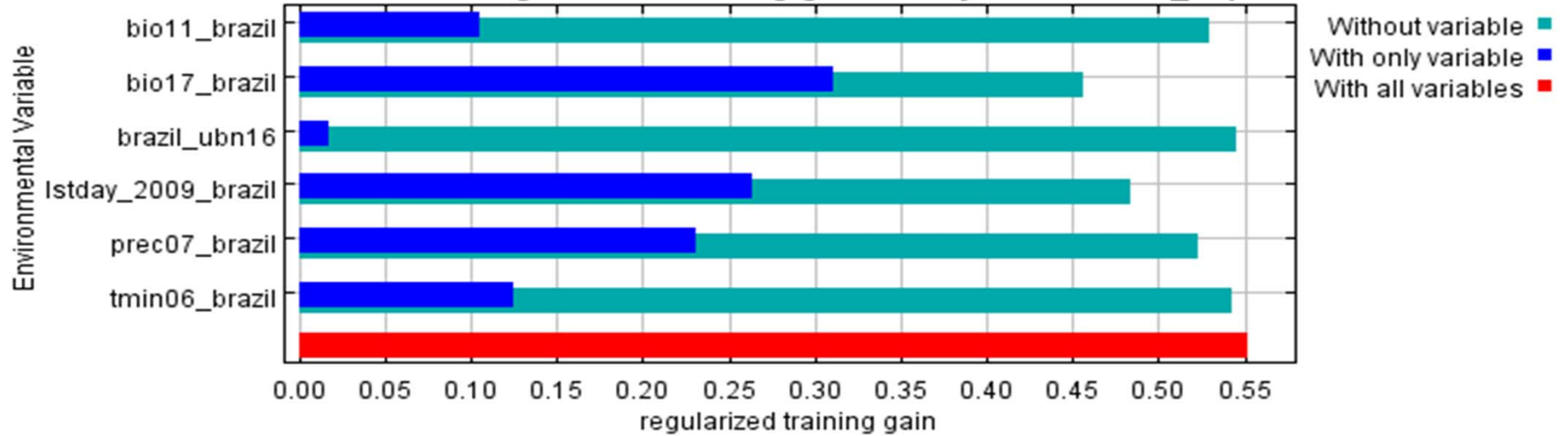
Variable	Percent contribution
prec02_brazil	75.3
bio14_brazil	13.1
alt01_brazil	5.4
Istnight_2008_brazil	4.5
brazil_ubn24	1.1
brazil_gdp1	0.7



## Combined model of the potential distribution of Leprosy in Brazil

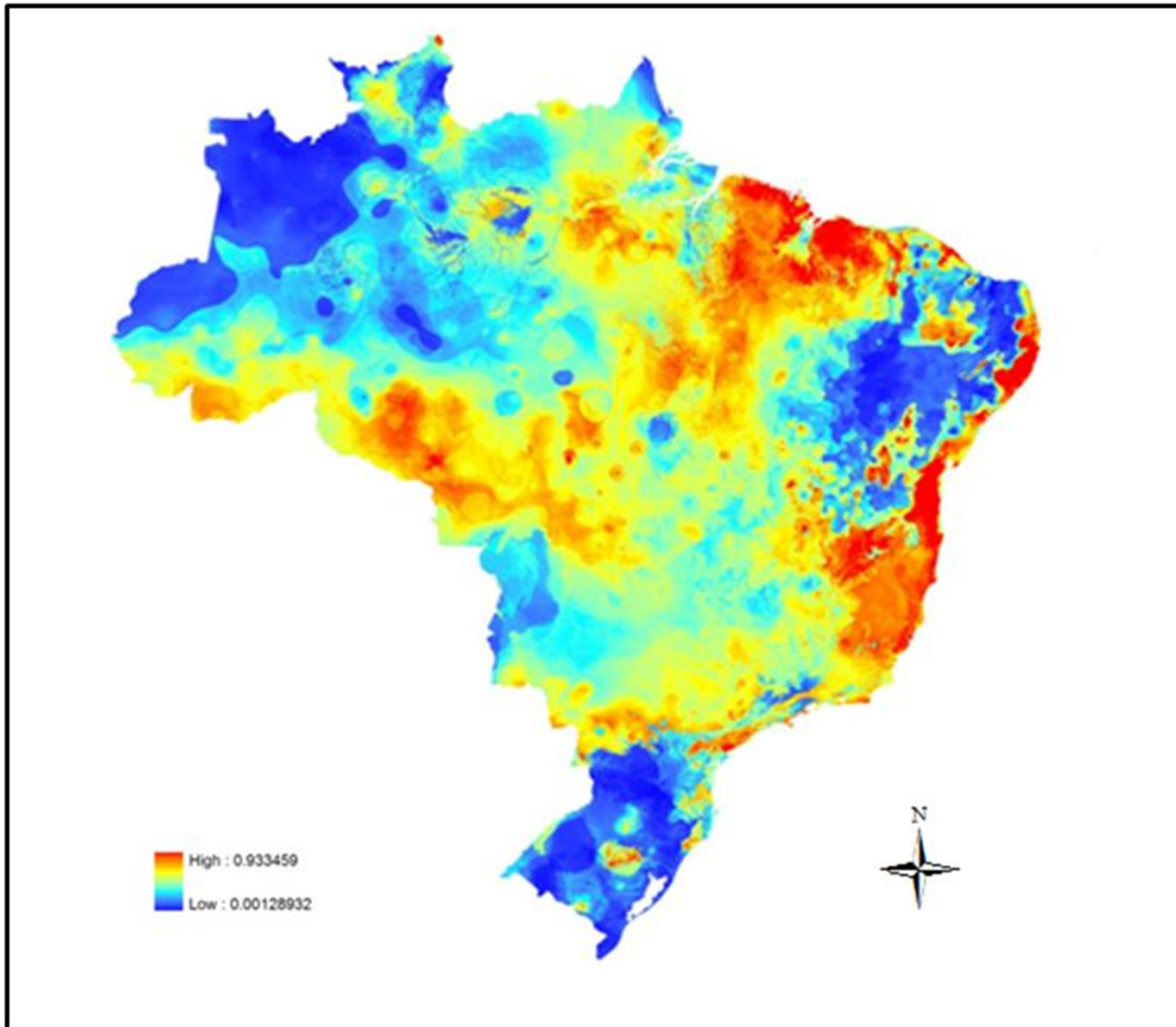


### Jackknife of regularized training gain for *Mycobacterium\_Leprae*

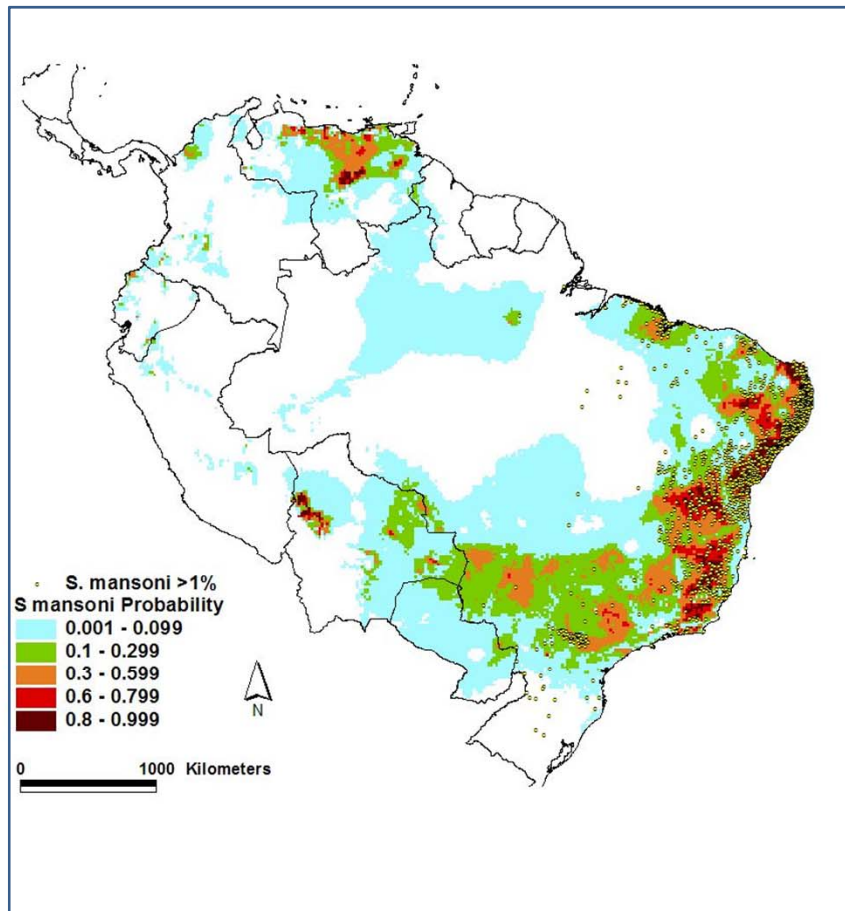


Variable	Percent contribution
bio17_brazil	44.3
lstday_2009_brazil	28.8
bio11_brazil	17
prec07_brazil	7.2
tmin06_brazil	1.8
brazil_ubn16	0.8

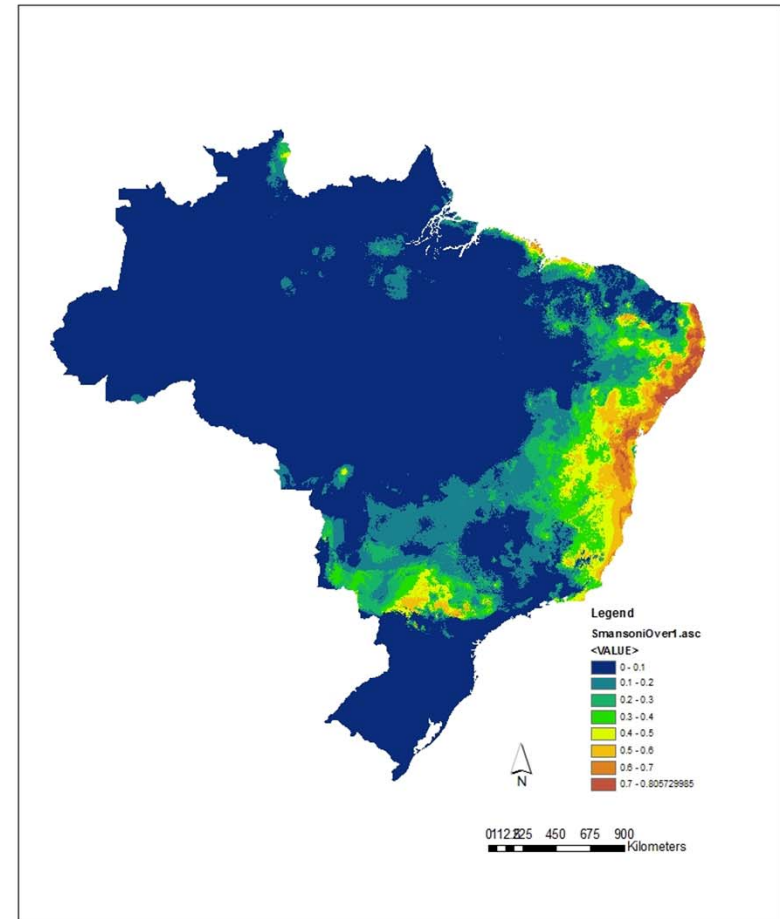
## Environmental model of the potential distribution of Cutaneous Leishmaniasis.(CL) in Brazil



# Optimize methods: Statistical Niche Modeling vs. Climate Based Probability Surfaces for *S. mansoni*



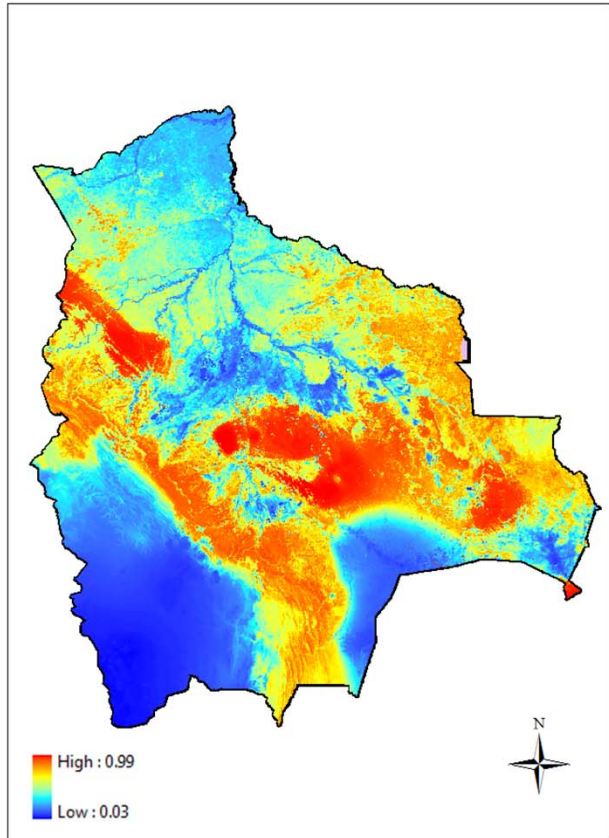
Floramap Model (Malone et.al., 2005)



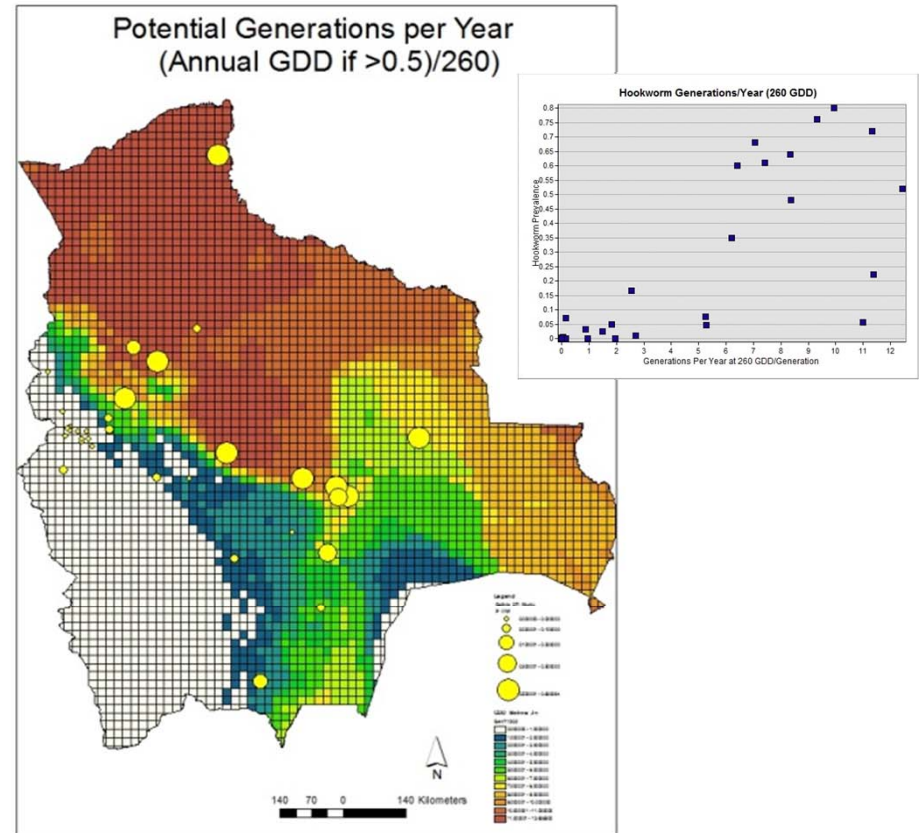
Maxent - 19 Variable Bioclim Model



# Maxent vs Biology Based GDD/WB Models for Hookworm in Bolivia

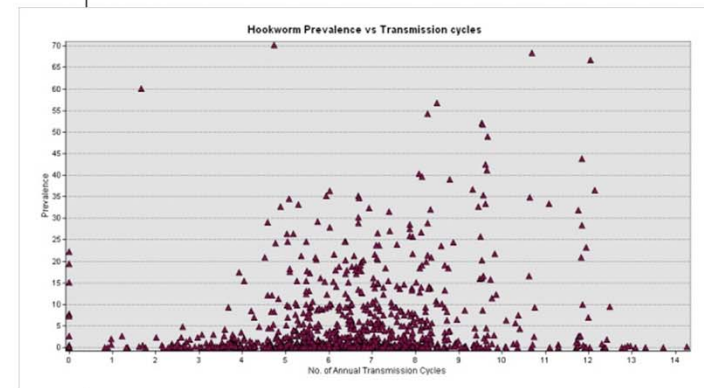
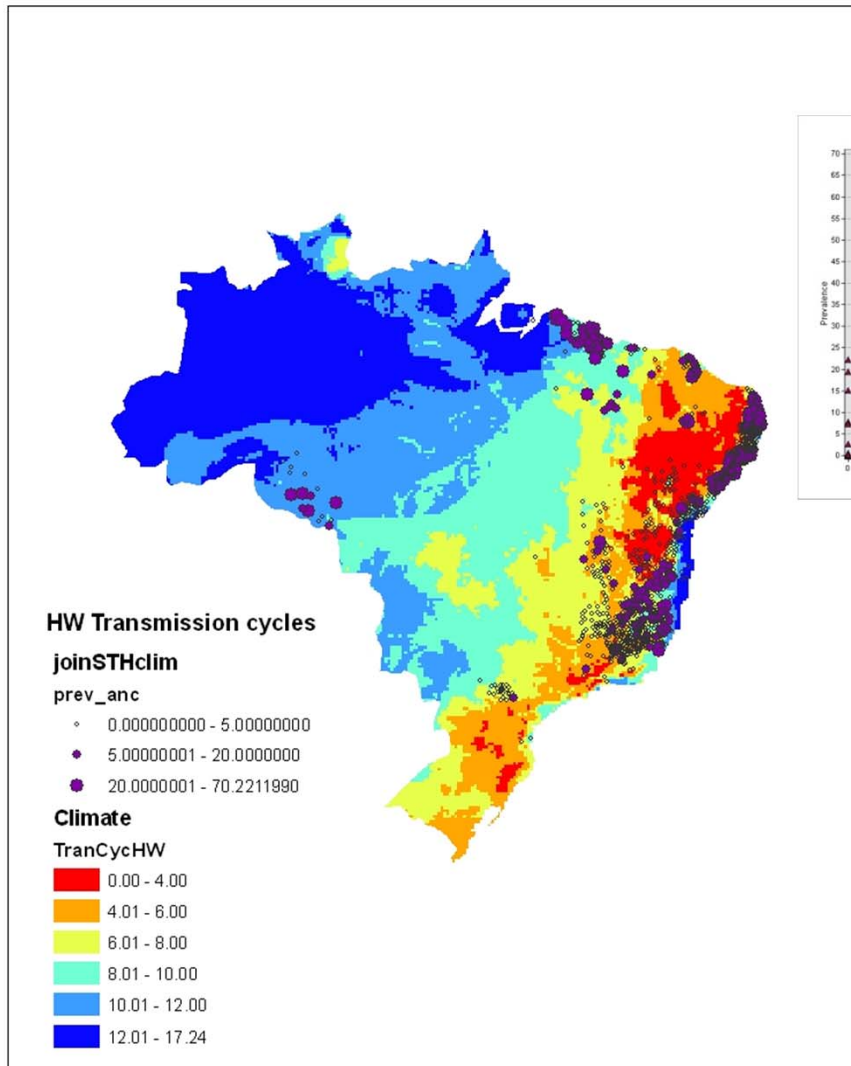


Hook Worms Environmental Model Based on Available Studies

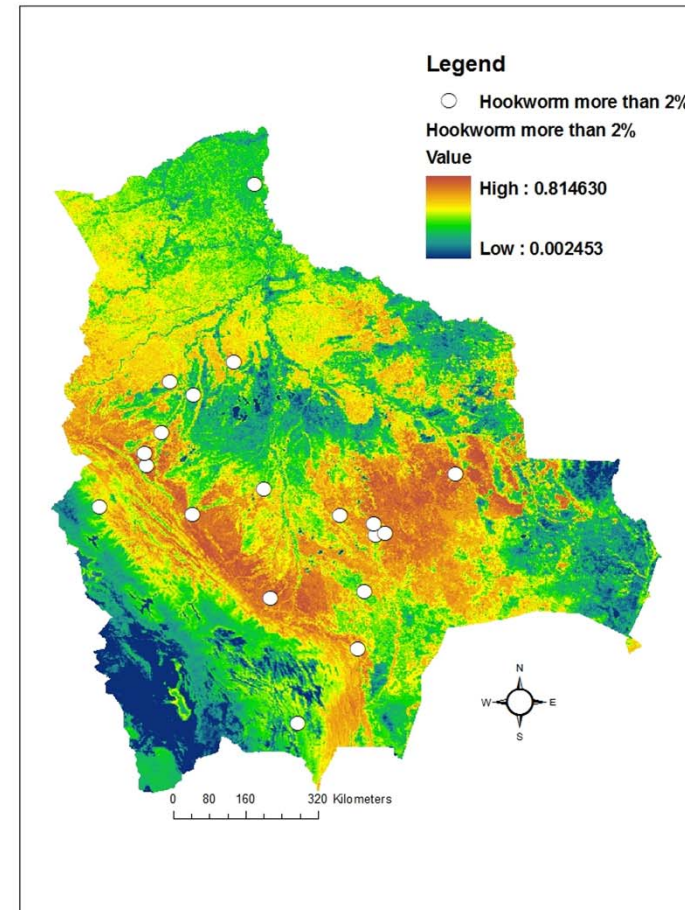
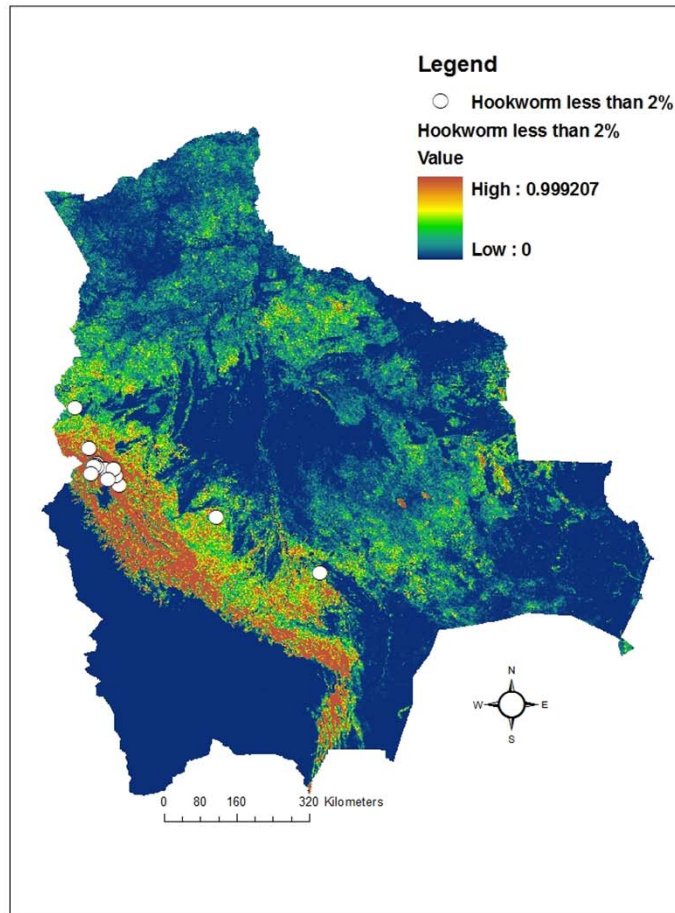


(Mean monthly temperature – 15)\*Number of days in a month; Months where Rain/PET < 0.5 were not counted; One generation requires 260 GDD = cycles of hookworm development and transmission in soil (Egg to L3 + mean L3 survival time)

# Biology Based GDD/Water Budget Analysis for Hookworms in Brazil



# Maxent Threshold Analysis vs MODIS LST, NDVI & dT – Hookworm Prevalence <2% and >2% in Bolivia



Control program thresholds for intervention decisions can be set using Maxent

# Future Work

---

---

1. Implement NTD modeling system capabilities by health agencies in Brazil, Bolivia and Colombia, including short courses on application of modeling methods and use of the data portal to test and refine use in current control programs
2. Develop similar NTD databases and risk modeling systems for additional countries in Latin America
3. Continue research on development of maximally effective mapping methods and risk modeling systems for use in surveillance and control of NTD in Latin America and the Caribbean



# 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 heterogeneous
4. Results of Maxent ecologic niche mapping and modeling should be validated by alternative methods eg. biology based GDDxWB climate models