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
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
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

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

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

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.

Chagas Vector Distribution

Rhodnius prolixus Environmental Model

Triatoma dimidiata Environmental Model
Chagas vector - Environmental Maxent model
Chagas Environmental Niche Model
Environmental Models

Sivigila (disease reports) 29 environmental variables

Multiple regression

Significant variables

Variables VIF<10

Variance Inflation factor

Literature vector reports 29 Environmental variables

Logistic regression

Maxent

Variables selection Pearson's

Re run Maxent

Final Model
Socioeconomic Factors – Municipality level
Socio-Economical Model

Multiple Regression and VIF

Choose variables for weighted models

Weighted model:

Combined (Socio economical – environmental) final model
Combined model of the potential distribution of Chagas in Brazil
<table>
<thead>
<tr>
<th>Variable</th>
<th>Percent contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>prec02_brazil</td>
<td>75.3</td>
</tr>
<tr>
<td>bio14_brazil</td>
<td>13.1</td>
</tr>
<tr>
<td>alt01_brazil</td>
<td>5.4</td>
</tr>
<tr>
<td>lstnight_2008_brazil</td>
<td>4.5</td>
</tr>
<tr>
<td>brazil_ubn24</td>
<td>1.1</td>
</tr>
<tr>
<td>brazil_gdp1</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Combined model of the potential distribution of Leprosy in Brazil
<table>
<thead>
<tr>
<th>Variable</th>
<th>Percent contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>bio17_brazil</td>
<td>44.3</td>
</tr>
<tr>
<td>lstday_2009_brazil</td>
<td>28.8</td>
</tr>
<tr>
<td>bio11_brazil</td>
<td>17</td>
</tr>
<tr>
<td>prec07_brazil</td>
<td>7.2</td>
</tr>
<tr>
<td>tmin06_brazil</td>
<td>1.8</td>
</tr>
<tr>
<td>brazil_ubn16</td>
<td>0.8</td>
</tr>
</tbody>
</table>
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

(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

HW Transmission cycles
joinSTHclim
prev_anc
- 0.000000000 - 5.000000000
- 5.000000001 - 20.0000000
- 20.00000001 - 70.2211990

Climate
TranCycHW
- 0.00 - 4.00
- 4.01 - 8.00
- 8.01 - 8.00
- 8.01 - 10.00
- 10.01 - 12.00
- 12.01 - 17.24
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.