Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Grete Wilt Date

Spatial Temporal Incidence and Environmental Determinants of Leptospirosis in Brazil

By

Grete E. Wilt
Master of Public Health

Global Environmental Health

_____

Yang Liu PhD
Committee Chair

_____

Paige Tolbert PhD
Committee Member

_____

Carlos Zambrana-Torrellio MsC
Committee Member

Spatial Temporal Incidence and Environmental Determinants of Leptospirosis in Brazil

By

Grete E. Wilt

B.A.
Colorado College
2014

Thesis Committee Chair: Yang Liu, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2016

Abstract

Background: Leptospirosis is a neglected tropical disease with 1.0 million cases and 59,000 deaths occurring each year. The major disease burden is placed on socio-economically depressed populations. Leptospirosis persists in an endemic state throughout Brazil, where ~3,000 cases are reported per year. Climate change in this country affects the ecology of numerous zoonotic diseases, leading to the increased risk of an outbreak.

Methods: In this study, we investigated the spatial temporal variations of leptospirosis incidence to enhance the understanding of leptospirosis patterns at a municipal scale. Additionally a general linear mixed model controlling for municipality and time was constructed to identify environmental predictors of leptospirosis.

Results: Results suggested high incidence rates of leptospirosis are spatially clustered in Southeastern Brazil. Temporal analyses indicated high seasonality, peaking from December-March, with overall trend of total cases and average cases reported increasing overtime as well. Our results suggest that increases in leptospirosis incidence were significantly associated with increases in mean monthly precipitation over time, mean monthly vegetation index score over time and year. Municipalities with lower urban populations and increased mean soil water pH and isothermality, interpreted as higher temperature evenness over the course of a year, were also significantly associated with higher leptospirosis incidence at baseline. Mean monthly temperature was not significantly associated with an increase in leptospirosis incidence in the final model.

Discussion: Our results clearly show the seasonal temporality of leptospirosis, with increases in leptospirosis observed over time. Our model illustrates an increase in leptospirosis in rural regions with high vegetation in time-periods of elevated rainfall. This approach assists in identifying spatial regions and time-periods of high potential infection risk that may lead to the development of strategies to improve targeted prevention and response.

Spatial Temporal Incidence and Environmental Determinants of Leptospirosis in Brazil

By

Grete E. Wilt

B.A.
Colorado College
2016

Thesis Committee Chair: Yang Liu PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in
2016

Acknowledgements

Firstly, I would like to thank my advisor, Carlos Zambrana-Torrellio, and the modeling and analytics team at EcoHealth Alliance, especially Allison White and Erica Johnson for providing support from project design to (hopeful) publication. Their willingness to discuss complex statistical processes, often over the phone and guidance of my development as a research was invaluable. Additionally I would like to thank Catherine Machalaba at EcoHealth Alliance for consulting on the policy aspects of a One Health zoonosis.

To the professors at the Rollins School of Public Health, especially my advisor, Yang Liu and the Spatial Analysis professors, thank you. You provided assurance, proposal guidance and methodological support.

Lastly to my family: Thank you for your constant support. Always.

Table of Contents

Introduction

Background

Leptospirosis, a tropical zoonotic spirochete bacterial disease of increasing worldwide importance, results in an estimated 1.0 million cases and 59,000 deaths on an annual basis (1). Despite this health burden, leptospirosis remains a relatively "neglected tropical disease". Furthermore, a growing body of evidence suggests associations between environmental factors (including environmental changes) and leptospirosis. For example, higher leptospirosis incidence rates have been associated with greater precipitation and flooding, inadequate floodwater drainage and increasing temperatures (2-5). These findings have major pubic health implications as the number of severe climatic events continues to rise.

Leptospira spp. range from 6-20 micrometers in length and prefer a consistent temperature of 28-30 °C (6). There are 13 pathogenic and 6 saprophytic species of *Leptospira*. Pathogenic strains, such as *L. interrogans* and *L. Icterohamorrhagiae,* are very adaptable and can grow in low nutrient environments, including moist soil and freshwater, surviving outside a host for weeks to months (7). *Leptospira spp.* have been identified in most mammals. Rodents, livestock and domesticated mammals serve as the predominant hosts of pathogenic species (8). The exact infectious load of leptospirosis is unknown, but estimates range from 10^3 -10^8 organisms/mL for all pathogenic *Leptospira spp.* (9). Shedding of the bacteria occurs through urine as hosts carry leptospirosis in the proximal renal tubules of the kidney (10). Exposure transpires through contact with mammalian reservoir host urine or environmental persistence(6). The bacteria enter the host through a small cut, abrasion or the mucous membrane (11).

Human exposure and subsequent cases have been linked to environmental, social, occupational and pathogenic risk factors. Human-to-human transmission is extremely rare (12). Pathogenic species are responsible for a myriad of signs and symptoms that range from mild to severe (11-13). While most infections result in nonspecific flu-like symptoms that resolve without treatment, leptospirosis can cause kidney damage, meningitis, liver failure, respiratory distress and death (14, 15). Severe manifestations of the disease present with organ failure and internal bleeding (Weil's disease) and associated severe hemorrhagic pulmonary syndrome. These severe diagnoses have 10% and 75% case fatality rates, respectively (16). With severe infections, morbidity and mortality can occur despite generic antibiotic treatment; these complications are at least in part attributed to poor patient immunity or delayed treatment initiation (17).

In Brazil, a large country with a tropical climate and heavy rainy season, there are approximately three to five thousand cases of leptospirosis confirmed annually through Sistema de Informacaco de Agravos de Noti Cacao (SINAN) (17). These case numbers are estimates and may underreport true disease burden due to lack of reporting in many municipalities throughout Brazil. In addition, the number of cases continues to rise, placing a significant burden on the public health and economy of the nation. Using SINAN leptospirosis data for 2007, it has been estimated that leptospirosis resulted in the loss of 6, 940 potential life years in Brazil with 75% of these lost years occurring in the 20-49 age group range. Approximately 11.85 million USD from years of minimum wage work were lost to the disease in 2007 (18). The estimated total economic cost of a one specific leptospirosis outbreak in Nova Friburgo, Brazil was estimated to range between 21,500 and 100,800 USD in 2007 (19). Given the inadequate appreciation of the adverse

health and socioeconomic consequences of leptospirosis, there is a lack of effective implementation of disease control measures. Research examining spatial temporal incidence and environmental determinants of leptospirosis in Brazil is needed to advance knowledge of the disease ecology and improve prevention, surveillance and response to disease outbreaks.

Epidemiology, Environmental Predictors and Prior Research of Leptospirosis

Numerous small-scale epidemiologic studies conducted throughout tropical regions of the globe have examined cross-sectional associations between Leptospirosis prevalence or aggregated incidence over time and a multitude of environmental factors including increased precipitation and flooding (2, 9-11, 20-26), greater temperature, low elevation, presence of agricultural lands and alkaline soils (2, 9-11, 20-34). Early epidemiologic studies suggested that leptospirosis predominantly manifests as a rural disease. However, due to massive urban migration and poor urban slum conditions, there has been an increase of leptospirosis cases in urban areas (22, 35-37). Sporadic outbreaks of leptospirosis are more frequent in urban settings (35). The Leptospirosis Burden Epidemiology Reference Group (LERG), supported by the World Health Organization, cites both environmental and social variables as important potential risk factors for leptospirosis that need further scientific support (4). While previous investigations have examined spatial prediction of hotspots and clustering of outbreaks, they have failed to examine the temporality of the associations of potential environment predictors with incidence rates. Climate change events have led to shifting environmental conditions throughout Brazil (38). Furthermore, habitat loss and increasing urbanization have

resulted in vegetation loss, placing large populations in leptospirosis suitable habitats at increased risk for outbreaks.

Since 2001, leptospirosis has been on the nationally notifiable disease registry of the Brazil Ministry of Health. However, many municipalities still do not report cases due to diagnostic challenges and limited funding (17). Previous local epidemiological studies examining leptospirosis in Brazil indicate the endemic nature of the disease. Dias et al. (37), reported that overall seroprevalence in Salvador, Brazil was 12.4%, indicating an endemic state of leptospirosis in Brazil with 61% of the seroprevalent individuals showing high titers for the highly infectious Icterohamorrhagiae serogroup. Work of Goncalves et al. in 2006 suggested that occupational exposures, such as farming or other jobs with proximity to mammalian reservoir hosts, are risk factors for leptospirosis in Brazil. For example, a study that examined illness in 150 slaughterhouse workers in Parana State, Brazil reported increased prevalence in workers with contact to slaughtered animals (39). Similarly, exposure to livestock in rural areas of Parana State, Brazil was associated with a higher likelihood of Leptospirosis, whereas control of rodents on the property and higher education level had protective effects (40). Lacerda et al. (41) found that *Leptospira spp.* infections occurred primarily in rice fields and persisted even in times of low rainfall in rural communities, perhaps because of prolonged exposure to reservoir hosts. These prior studies illustrate that restricted by the lack of spatial and temporal data and analyses. Highlighting regions of the country with elevated incidence rates remains a gap in the literature. In addition, prior research has not adequately addressed the changes in incidence over time across Brazil or comprehensively evaluated associations of environmental factors with the disease, at a regional scale.

A Spatial Temporal Analysis

Complex interactions in disease ecology require expanding methodological capabilities in statistics, geographic information systems and remote sensing (42). These tools improve the ability to analyze spatial temporal relationships of environmental risk factors of infectious diseases. Examining disease trends through a spatial temporal lens will improve predictions of future disease trends (43). In turn, these applications may enhance delivery of public health and clinical interventions to most efficiently and effectively reduce disease burden. Dynamic time series maps, as well as models accounting for space and time, can characterize and quantify how the spatial distribution of disease changes over time, especially in regards to a changing climate. Several studies have examined the relationships of cases and incidence of a number of infectious diseases with environmental factors including land classification, vegetation, temperature and precipitation using general linear mixed modeling (GLMM) (44-46). However, no GLMM models exist for leptospirosis. Allowing for changing environmental factors across the dimensions of time and space provides an improved model fit with the opportunity to account for random effects and adjust for correlations within time and space. Understanding the association between incidence rates, geographical location and time is crucial to exploring the seasonality of leptospirosis as well as forecasting future estimated burdens of disease in regards to the changing environment and climate.

Purpose of Study

Our study examined the associations between several environmental factors and incidence of leptospirosis in cases/ 100,000 person-months, through a spatial temporal

lens. The principal aims of this study were to: **1)** examine the clustering of total leptospirosis incidence in Brazil on a spatial scale to identify regions with high incidence, **2)** evaluate the trend and seasonality of leptospirosis in Brazil and **3)** identify potential environmental drivers of leptospirosis incidence accounting for variations and clustering in time and space.

This analysis characterized the temporal trends of leptospirosis in Brazil as well as identified high incidence both spatially and seasonally. Through GLMMs, we model the variation in the true rates and highlight systemic from random variability and develop a spatial temporal ecological model for leptospirosis at a municipal level in Brazil. Additionally, we identified significant environmental predictors associated with leptospirosis incidence, allowing for an improved understanding of the eco-epidemiology of leptospirosis in Brazil.

Methods

Study Design

We conducted a retrospective analysis of leptospirosis incidence using confirmed cases from SINAN aggregated by month and municipality collected and reported by the Brazil Ministry of Health from 2001 to 2014 (17). We report monthly incidence in cases/100,000 person-months for the 13-year period. Analyses were conducted with the statistical software R3.2 (47), ArcMap 10.3 by ESRI (48), Cluster Seer 2.0 (49), GeoDa (50).

Data Collection, Matching, and Merging Methods

Individual cases meeting the serological definition of leptospirosis (i.e positive microscopic agglutination test (MAT) or the enzyme-linked immunosorbent assay (ELISA) were collected by SINAN and identified at the municipality of residence (17). We used Microsoft Excel to gather reported cases of leptospirosis from January 2001 to December 2014. Data from 2007 was not available for analysis. We calculated incidence rates using municipality population from 2010 the last population census in Brazil. Incidence was reported as number of cases per 100,000 person months. We assumed that municipality population size remained constant over the analysis period (51).

The global administrative areas database reports a total of 5,504 municipalities within Brazil's 26 states. Despite leptospirosis being a nationally mandated reported disease in the country, not all municipalities report due to personnel and financial constraints. Over the 13 year surveillance period 45.2% of all municipalities reported incidence rates (2,487 out of 5,504) (52). The data was log-transformed to adjust for the non-normal, strongly right skewed data. Potential environmental predictors, identified from a review of relevant literature, were acquired from different sources in raster form (Table 1). For each covariate, we used zonal statistics to calculate the mean value by municipality. Remotely sensed environmental data were grouped into three time scales: time invariant, monthly time variant or monthly time variant over the entire study period (Table 1). All analyses were performed using the statistical software R v3.2 (47) and the following packages: sp, maptools, raster, dplyr, lattice, ggplot2 and nlme.

Descriptive Analysis

We performed a series of descriptive analyses of both the count and incidence case data as well as the environmental independent variables of interest. We preliminarily investigated the potential predictor variables of interest with summary statistics and histogram plots. None of the environmental predictors were severely skewed and thus were not transformed. Mean, standard deviation, median, and percent of data missing in the potential predictor variables were reported (Table 2).

General correlation matrices were used to explore the covariate data for potential multicollinearity as well as associations with the outcome of interest: leptospirosis incidence. The ranges of the continuous environmental variables were explored. We reported the mean, standard deviation and percent of data missing for each potential dependent and independent variable (Table 2).

Spatial Data Statistics

Spatial regions were defined in this analysis as GADM standardized Brazil municipalities with reported incidence. We performed all spatial statistical analyses using ArcGIS 10.1 by ESRI (48), Cluster Seer 2.0 (49), and GeoDa (50). We spatially explored the distribution of leptospirosis cases and incidence across the municipalities. We projected the Brazil shapefile in the WGS 1984 World Mercator Coordinate system. Spatial statistics were performed only on municipalities with reported cases and incidence data. We calculated the geographic distribution of mean incidence and the standard deviational ellipse to assess the incidence distribution and concentration.

Choropleth maps were used to represent variations in incidence, local autocorrelation, and hotspots across municipalities.

Global Autocorrelation

Global indices of spatial autocorrelation compared incidence rate distribution within municipalities across Brazil. The Global Moran's *I* test is a well-recognized method with which to test the significance of the degree to which similar incidence rates of leptospirosis occurred in neighboring municipalities (near in space). The null hypothesis of this test assumed leptospirosis incidence is spatially independent of its neighbors with the disease occurring randomly across regions and the population at risk has an even distribution. Using Clusterseer software (49), we assed significance of global autocorrelation using queens contiguity to build the spatial weights at a 95% significance level ($p<0.05$) with 999 repeated Monte Carlo Simulations.

Local Spatial Autocorrelation

Local indicators of spatial autocorrelation (LISA) highlight where localized clusters of disease occur in space. The Local Moran's *I* is similar to the Global Moran's *I*, with the *I* statistic calculated for every municipality instead of the entire county. The sum of the Local *I* values is proportional to the Global Moran's *I* statistic. We performed the Local Moran's I using a queen's contiguity formation to build the spatial weights matrix in GeoDa (50). The Global and Local Moran's *I* detect the presence of similar value clustering. Moran's *I* values close to +1 indicate that the municipality has a high incidence rate of leptospirosis and is surrounded by other municipalities with high

incidence rates (i.e. high-high clustering). For both the Global and Local Moran's *I*, Z-scores and p-values assessed the significance of the overall and local incidence clustering.

Hot-Spot Analysis

We conducted a hotspot analysis using the Getis-Ord Gi* statistic with a queens contiguity weighting, that allowed each municipality to have one or more neighbor. This method identifies municipalities with higher and lower incidence than predicted. The Getis-Ord Gi* statistic significantly separates clusters of high values from clusters of low values, which are defined as hotspots and cold spots, respectively as compared to neighboring values at varying degrees of significance.

Temporal Analysis

Temporal variation of leptospirosis cases were assessed using a decomposed seasonal, trend line analysis (STL). This procedure filters time series data into trend, seasonal and remainder components (53). This analysis included 13 time points with 12 observations per unit time due to the monthly nature of the data. From these models and graphs, we assessed the variation and magnitude in total and average leptospirosis cases over time of overall temporal trend and seasonality. The STL allows for removal of seasonal peaks to report on the secular trend of the data.

Additionally, we constructed both an unconditional means model and an unconditional growth model. These are the first two linear models in the multilevel hierarchical model taxonomy and assess variation in leptospirosis incidence due to the

inter- and intra-municipality differences, respectively. The intraclass correlation coefficient (ICC), $\rho = \sigma^2_0 / (\sigma^2_0 + \sigma^2_\varepsilon)$ (54) was calculated using the variances of the unconditional means model. The ICC illustrates the variation in incidence rates attributable to differences among municipalities rather than time. These models were the first in the taxonomic multilevel model hierarchy.

Spatial Temporal Environmental Model

We acquired potential predictor variable data from Worldclim (55), Soilgrids (56), NASAs near earth observatory program (NEO) (57) and the 2010 Brazil census (51). To estimate vegetation by municipality, the normalized difference vegetation index (NDVI) calculated a greenness measure of vegetation from 0.0 to 0.8 and was used to quantify vegetation. Remote sensing data was also used to report temperature, precipitation,, soil water pH and elevation. Isothermality, the quantification of temperature evenness throughout the year, and urban population by municipality were reported as percentages (Table 1). Missing data trends were explored. Due to a large amount of missing data (44%) in the NDVI dataset, vegetation index values for municipalities were imputed using the median value. General correlation matrices were used to explore the covariate data for potential multicollinearity, as other methodologies such as variance inflation factors are not possible to calculate with multilevel models. Pearson correlation values for all covariates were <0.4. We explored ranges of the continuous environmental variables and joined case data to the environmental data.

Due to the temporal nature of the data, a general linear model would fail to account for spatial and temporal variability that is factored into potential independent

variables. Thus we apply a general linear mixed modeling framework, as these models can account for temporal and spatial clustering of the data by structured random effects in the linear predictor. The error covariance structures of mixed models are allowed to vary at both the municipal and time levels which improve model fit and factors temporal and spatial autocorrelation into the random effects. The Gaussian multi-level mixed model used in this analysis assumes the taxonomy of statistical models, with a systematic sequence of models that extend a prior model. With this complex multilevel model, both the level 1 (municipal) and level 2 (time) change trajectories are assumed to be linear. These models assume univariate normality at level 1 and bivariate normality at level 2. These assumptions were checked and validated in the exploratory analysis (54). We included an autoregressive time series component and utilized the continually Gaussian autoregressive structure to account for spatial dependence. We used the taxonomic hierarchy of the model framework, to compare independent variables, identify significant predictors (at $p<0.05$) and select an appropriate covariance structure for the best fitting model according to lowest values for deviance, AIC and BIC parameters and significance of each independent variable. Models for consideration appear in the model selection framework (Table 3). The final model was back transformed for interpretability. To assess the model fit, fixed and random residuals for the selected model were interpreted and graphed.

Results

Descriptive Analysis

A total of 110,430 cases of leptospirosis were documented from 2001-2014 by the Brazil Ministry of Health. For the 2,417 municipalities reporting any cases (43% of all municipalities), the total cases reported by municipalities over the time-period ranged from 1 to 228. The mean number of leptospirosis cases per month in municipalities was 2.16 cases respectively, with an SD of 4.57. Population size in these municipalities ranged from 1,216 to 11,250,000 with a mean of 368,000 with a SD of 114,000.

There was a wide range of incidence rates across municipalities reporting cases with a minimum incidence of 0.055 to a maximum incidence of 200.80 cases/100,000 person-months. Mean incidence of leptospirosis among municipalities reporting cases in Brazil was 10.82 cases/100,000 person-months with an SD of 13.52. The data were strongly right skewed, according to the histogram. Incidence was selected as the outcome of interest for spatial and modeling methodologies.

Spatial Data Statistics

Descriptive

We observed cases of leptospirosis across all regions of Brazil (though in less than 50% of municipalities due to absence of reporting by 57% of municipalities). High case totals (Figure 1) were clustered in the Southeast region. Additionally, the mean center and the directional distributions showed incidence clustering in the Southeast regions of the country. The directional distribution, summarized by the standard deviation

ellipse, depicted the spread of the data longitudinally from north to south in the eastern region of the country (Figure 2).

Global Autocorrelation Analysis

The Global Moran's *I* value under Monte Carlo simulation was 0.099 with a p-value of 0.002. The magnitude of the *I* value indicated that this clustering was modest in magnitude, yet statistically significant.

Local Autocorrelation Analysis

Results of the Local Moran's I are displayed in a LISA cluster map (Figure 3). Relationships are categorized as high leptospirosis incidence next to high leptospirosis incidence (high-high), and low incidence next to low incidence (low-low). Outliers are identified as high-low and low-high regions. All clusters noted in the cluster map are significant at $p = 0.05$.

We identified High-High clustering in 76 municipalities, primarily located in the South. low-low clustering was identified in 177 municipalities scattered around Brazil, primarily in the East. We highlighted 97 low-high and 31 high-low outliers, where municipalities with low and high incidence occurred next to municipalities with high and low incidence respectively. Municipalities without significant clustering (n=2016) and neighborless municipalities (n=150) were also observed, due to the missing data.

Hotspot Analysis

The incidence hotspot analysis, using the Getis-Ord Gi* statistic, isolated significant hotspots at 90%, 95% and 99% confidence levels (Figure 4). Similarly to the local autocorrelation analysis, hotspots were scattered throughout the country, but most concentrated in the south. Ninety-two municipalities had a 99% ($p < 0.01$) confidence of hotspots, 77 municipalities for 95% ($p < 0.05$) confidence of hotspots and 51 with 90% ($p < 0.10$) confidence of hotspot activity. These hotspots indicated municipalities where higher incidence than expected was occurring. Five coldspots, regions with lower incidence than expected, were detected at a 90% confidence level ($p < 0.10$). No additional significant coldspots were identified at $p<0.05$ and $p<0.01$.

Temporal Models

Temporal variation in the data was assessed via seasonal trend line (STL) decomposition. Seasonality accounted for a large magnitude of total cases and average cases by municipality over time. Case totals over the 13 years of data peaked from January to March, while the lowest case numbers of leptospirosis were observed in September (Figure 5). Variation occurred in total cases per year (Figure 6). Our STL extracted seasonality, overall trend and the remainder from the data. The seasonal STL value was 210.52 with a trend line STL of 80.79 (Figure 7). This illustrates that much of the variation is occurring seasonally. However once the seasonality was removed, a general positive increase in the secular trend was observed. Figure 7 illustrates a general increasing trend in total cases over time, accounting for seasonality

Results of the unconditional means model (Model A, Table 4) and unconditional growth model (Model B) are shown in Table 4. The intracorrelation coefficient, ρ was 0.829 indicating that 82.9% of the total variation in the incidence of leptospirosis was attributable to differences among municipalities. Despite the evidence that much of variation in incidence was to difference in municipalities, the estimated geometric mean incidence of leptospirosis (after back transformation of model B) was 5.62 cases/100,000 person-months at baseline (Table 4). When fit in a linear model, incidence significantly decreased by 0.02% per month, controlling for municipalities. The unconditional growth model (Model B, Table 4) allowed for random effects within the model and estimated the within municipality variation and variation of incidence at baseline (Table 4). These variance components assess the amount of outcome variability remaining after fitting the multilevel model. As we lack comparison points for the random effect estimates, these estimates simply quantify the amount of variation remaining in the model against the null benchmark of 0.

Spatial Temporal Environmental Models

Multi-level mixed models were fitted to leptospirosis incidence data to identify candidate environmental variables that were significantly associated with incidence over time. Models C-L (Table 4) were constructed with the independent environmental variables univariately, to identify variables for further evaluation (using a $p < 0.2$). Examination of partial plots and spaghetti plots revealed a roughly linear relationship between time and log incidence of leptospirosis. Multicollinearity did not present as an issue to this model. Level 1 and level 2 variables included in each hierarchical model

evaluated are outlined in table 3. Model N was selected as the best model and the conditional autoregressive error covariance structure was the best fit (model N*, Table 4). This structure is a spatial statistics function used to define the statistical measurements in neighboring municipalities and used when significant local autocorrelation is present, as defined by a LISA statistic (Table 5)

In the final back-transformed model (N*, Table 4) there were six significant independent variables; two were time variant and four were time invariant (baseline) variables. Precipitation and vegetation were significant predictors of leptospirosis incidence over time. Each 10 mm increase in precipitation over time was significantly associated with a 0.3% monthly geometric mean increase in incidence. Each unit increase in the NDVI vegetation index was significantly associated with a 1.03% monthly geometric mean increase in incidence. For every year increase, the geometric mean incidence of leptospirosis significantly increased by 8.42%. Time invariant variables at baseline that significantly predicted leptospirosis incidence included urban population percentage, soil water pH, and isothermality. Each 1% decrease in urban population percentage was associated with a 2.70% geometric mean increase in leptospirosis, while each 10 ℃ decrease in mean monthly temperature was associated with a 0.01% geometric mean increase in incidence. This association was not significant at the 5% level. Each 0.10 point increase in soil water pH was significantly ($p<0.10$) associated with a 1.3% geometric mean increase in incidence Isothermality, (58), was significantly associated with leptospirosis incidence at baseline but not over time. Each 1% decrease in isothermality at baseline was associated with a 3.24% geometric mean increase in leptospirosis incidence. However, every 10% increase in isothermality over time was

associated with a 0.02% increase in geometric mean incidence. This suggests the strength the association between temperature evenness associated with increased incidence, decreased over time. All associations were significant at $p<0.05$ without an error covariance structure specified. In the final complete autoregressive correlation structure: mean temperature was not significant at the 95% confidence level.

Variances of the intercept, slope and residuals were significantly different than zero ($p<0.05$), suggesting other variables not included in the final model (N*) may contribute to the variation (Table 4). These values also stress the importance of temporality in the model as it provides a better fit to explain the variation in the data.

To examine the model fit, we plotted residuals of the data set (Figure 8). Figure 8 shows the errors between the observed and expected incidence values as a result of the model. The residuals from the data set had a median of -0.772 with an interquartile range (IQR) of 1.73. The small IQR suggests that the expected and observed values are relatively close to each other in value. Additionally the fixed and random residuals were plotted against each other (Figure 9). The random residual median was -0.16 with an IQR of 0.60. This depicts the improved fit when the model is allowed to vary on the intercept according to the spatial location of the municipalities.

Discussion

This analysis presents the first spatial temporal investigation of leptospirosis. We identified significant clustering of incidence rates of leptospirosis in space. In addition, cases of leptospirosis showed strong seasonality, however the secular trend of cases increased overtime as well. Linearly fitted model illustrating incidence rates over time

did not show a strong magnitude of change. Multilevel modeling highlighted five environmental variables significantly associated with increases in leptospirosis incidence in the final model with seven covariates.

We observed cases of leptospirosis in all 21 states of Brazil, with the case and incidence distributions in addition to the spatial mean and standard deviational ellipse suggesting a higher case presence in Southeast municipalities in Brazil. This significant clustering is also supported by findings of the global and local autocorrelation and hotspot analyses. Global and local autocorrelation indicated similar incidence rates are significantly clustered in space. Providing evidence that the distribution of leptospirosis incidence across municipalities of Brazil is not random. The results of the hot spot analysis indicated spatial pinpointing of high incidence clusters in the South occurred. The intraclass correlation coefficient additionally suggested spatial significance, attributing 82.9% of the total variation in the incidence of leptospirosis to differences among municipalities.

The results of the unconditional growth model and time series analysis demonstrated significant temporality in leptospirosis incidence. The seasonal decompression analysis highlighted a strong seasonality in cases, with a peak from December-March. These findings suggested the presence of associations of ecological factors with incidence that were later explored in the mixed model. Secular trend lines indicate a general increase in leptospirosis in Brazil over time. However, the unconditional growth model does not highlight a strong linear magnitude in leptospirosis incidence decrease by month. Controlling for municipality, in every month increase, there is a significant 0.02% decrease in leptospirosis incidence. However the unconditional

growth model, and the overall hierarchical framework does not factor in seasonality to the model (Model B, Table 4). The inability to separate seasonal and overall trends make this model less capable of elucidating temporal trends. Thus, we conclude that leptospirosis in Brazil municipalities is increasing over time, regardless of a strong case seasonality.

The hierarchical mixed model framework allowed for exploration of potential associations of environmental factors with incidence that may further explain the variation in incidence between municipalities across time. Six independent variables were significantly associated with leptospirosis incidence in the final model consisting of seven covaraites. Increases in precipitation, vegetation, soil water pH and time were each associated with increases in leptospirosis incidence, while decreases in isothermality and percent of urban population were associated with increases in leptospirosis incidence. With these findings, we conclude that numerous environmental factors affect the distribution of leptospirosis across Brazil spatially and temporally.

While mean temperature was significantly associated in the final model with no error covariance structure specified, it was not significant upon applying the conditional autoregressive formula that accounts for spatial dependency. This is potentially due to significant spatial autocorrelation among the mean temperature variable, reducing the variable's importance in the model, once the covariance structure was specified. The conditional autoregressive error covariance structure allowed for random effects correlations by neighboring municipalities. There was no significant univariate association with elevation. While other literature supports a negative association between elevation and leptospirosis incidence (46), this independent variable was excluded due to

lack of significance in the hierarchical model. The lack of significance could be related to a modifiable area unit problem as elevation significant at a finer scale. Alternatively the lack of significance could be a result of elevation acting as a distal predictor. Localized epidemiology studies by Lau suggest soil acidity may affect distribution of leptospirosis cases (23). Our model suggested an increase in incidence in basic pH soils. Increasingly basic soils may suggest increased water content in the soil as the median soil water pH was acidic. An analysis of leptospirosis and potential predictors at different geographic scales in urban Brazil indicated that the strongest correlations at the regional level were population density while at the local level, flood risk was strongest (59). This could potentially explain why altitude was not significant in this municipal scale model. Municipalities with a higher proportion of rural residents were associated with increased leptospirosis incidence. Despite outbreaks in slums becoming increasingly frequent, our research suggests they do not account for a significant disease burden (22). Our model adds support to the literature base that increases in precipitation, isothermality and vegetation over time and rural regions are significantly associated with leptospirosis. However, we explored the disease through a spatial temporal lens as that remained a gap in the literature and from this analysis identified regions of Brazil in space and time where a high risk of leptospirosis exists.

Limitations

The spatial and temporal resolution of both the dependent and independent variables lead to limitations in this analysis, due to the ecological fallacy and modifiable area unit problem where conclusions based on data at a particular scale may not translate

to different scales. Case aggregation at the municipal level dilutes the ability of spatial statistics to identify specific data clusters and hotspots. Municipalities lacking case reports (n= 3,017), may lead to underestimations in true clustering and hotspots of leptospirosis in Brazil municipalities. Additionally, we calculated incidence rates and urban population percent under the assumption of a stagnant population. Spatial statistics were calculated ignoring the municipalities without incidence data. Future analysis could apply Empirical Bayesian spatial smoothing statistics to obtain incidence estimates for unreported regions and adjust rates by nearest neighbor population averages. Missing data was also apparent in ecological variables, most notably vegetation index (table 2), however this was attributed to random cloud cover and not thought to have a large affect on the model. The missing vegetation data was imputed using the median NDVI value, however in the future, interpolation models could be applied to strengthen our estimations. The case data of leptospirosis in Brazil municipalities was not reported consistently across months for regions under surveillance (51), limiting the temporal validity of the model. Due to unavailable case information data such as sex, race or age, we did not examine population demographic characteristics of the disease. As with all ecological epidemiological analyses, we cannot attribute causality to the significant associations observed in the mixed model. Land use change was not incorporated into the model and differentiations between climate change and anthropogenic events are not distinguishable in the model. We assume our data met the assumptions of the multilevel mixed model approach. Currently, there is no prominent method to assess multicollinearity or variance inflation factors in multilevel models and we used correlation matrices to assess potential multicollinearity. This model considered change

in incidence over time but did not incorporate seasonality as a unique variable. To our knowledge, this analysis was the first spatial temporal mixed model analysis of leptospirosis.

Despite limitations stemming from the resolution of both the incidence and independent variable data, this study explored the dynamics of leptospirosis and associated environmental factors across space and time in a new lens. Additionally, the rational quadratic covariance correlation structure led to an improved model fit by allowing for municipalities close in space and time to have similar error covariance structures. Testing of reservoir hosts for seroprevalence rates in regions with high incidence could improve future models. Limited funding and time are severe restrictions on the collection of this spatial and temporal data.

Public Health Significance

Leptospirosis is a neglected tropical zoonosis of increasing importance as case fatalities rise to nearly 60,000 annually (1). Leptospirosis incidence had significant high-high local autocorrelation and identified hotspots predominately in Southern Brazil, with a significant increase in incidence temporally. Low-low clustering was present in the North. Numerous environmental factors, including precipitation and temperature, were shown to be significantly associated with increased incidence. These findings will allow for enhanced predictive models and can open discussions on improved surveillance and disease prevention in the high-risk regions of the South while directing fewer resources to the low-risk regions in Northern Brazil.

The World Health Organization launched the multi-sectoral Global Leptospirosis Environmental Action Network (GLEAN) in 2010. GLEAN was developed to improve global and local strategies for prediction, prevention, detection and interventions of leptospirosis outbreaks in high-risk regions (60). Combining GLEAN's action-oriented platform with results from our spatial temporal analysis, allow for targeted policy suggestions. Leptospirosis is an ideal zoonosis to apply a "One Health" research and prevention framework. "One Health" focuses on a collaborative, multidisciplinary approach to the critical interactions between humans, wildlife, and ecosystems. Disease endemicity, animal infections, and circulating serovars are crucial in their contributions to human outbreaks (60). Our analysis identified municipalities with high and increasing incidence rates, with seasonal peaks occurring yearly from December through March. Significant environmental factors associated with these incidence increases including precipitation and mean temperature were highlighted. Additionally leptospirosis cases are increasing over time in Brazil. The number of severe climate-related events are expected to double in the next twenty-five years (61). Coupled with land-use change projections and humans migrating into previously heavily vegetated areas, climate change presents the potential for increased zoonotic outbreaks. Additionally land-use change can lead to a change in the soil water pH, potentially causing more widespread environmental persistence of the bacteria. If these trends continue as projected, leptospirosis outbreaks are anticipated to become more frequent as exposure risk increases.

Drawing from our findings, we recommend an improved "One Health" active surveillance plan in several municipalities in the South, to enhance knowledge on serovar prevalence in animals and humans in these endemic regions. Targeted public health

messaging during seasons of high infection (December-March) in rural regions should focus on education to health care providers concerning signs and symptoms of leptospirosis and stress proper wound bandaging. Individuals should limit exposure in vegetated regions during the rainy season. Basic occupational interventions including proper workplace attire in the fields (long sleeves, proper footware, gloves, etc.), improvements to informal housing and cross-ministry involvement have the potential to reduce infection in males ages 20-49. This is both the largest demographic at risk for infection and the primary income earners in this society. We suggest collaboration between the Ministries of Health, Housing and Agriculture to target drivers of leptospirosis across organizations. A cost-benefits analysis of targeted interventions does not yet exist, however de Souza et al. stress the high social cost of leptospirosis including years of life lost and hospital costs. An estimated 11,847,151.32 USD were lost to the disease over years of work lost in cases that progressed to death from a 2007 study of Brazil leptospirosis (18). Applying these focused directives in Brazil to individuals and regions at highest risk, predicted by social and environmental drivers, will improve prevention response initiatives in a leptospirosis endemic country.

Conclusions

The dynamics of disease ecology between *leptospira spp.*, their reservoir, dead-end hosts and the environment are complex. By investigating environmental drivers of leptospirosis through a spatial temporal lens, we obtain an improved understanding of leptospirosis incidence across Brazil. Our findings support previous studies, which suggest leptospirosis incidence is associated with increases in precipitation and

vegetation, while stressing the large impact of leptospirosis on rural municipalities with soils prone to flooding. While environmental determinates play a crucial role in the distribution of *leptospira spp.*, examining spatial temporal interactions at a smaller scale will further elucidate this potentially causal relationship. With these advancements come increasingly sensitive and accurate models for predicting regions at high risk for epidemic or endemic leptospirosis.

Works Cited

1. Torgerson PR, Hagan JE, Costa F, Calcagno J, Kane M, Martinez-Silveira MS, et al. Global Burden of Leptospirosis: Estimated in Terms of Disability Adjusted Life Years. PLoS Neglected Tropical Diseases. 2015;9(10).
2. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. Nature. 2008;451(7181):990-U4.
3. Lubchenco J. Entering the century of the environment: A new social contract for science. Science. 1998;279(5350):491-7.
4. Organization WH. Report of the Second Meeting of the Leptospirosis Burden Epidemiology Reference Group. Geneva, Switzerland: World Health Organization, 2011.
5. Global administrative areas (boundaries) [Internet]. University of Berkeley, Museum of Vertebrate Zoology and the International Rice Research Institute 2012.
6. Adler B, de la Pena Moctezuma A. Leprospira and leptospirosis. Veterinary Microbiology. 2010;140(3-4):287-96.
7. Pertuiset E, Chong MF, Duval G, Genin R. CLINICAL-FEATURES AND PROGNOSIS OF LEPTOSPIROSIS (WEILS DISEASE) IN ADULTS - A STUDY OF 249 CASES IN LA REUNION. Revue De Medecine Interne. 1988;9(5):487-93.
8. Ivanova S, Herbreteau V, Blasdell K, Chaval Y, Buchy P, Guillard B, et al. Leptospira and Rodents in Cambodia: Environmental Determinants of Infection. American Journal of Tropical Medicine and Hygiene. 2012;86(6):1032-8.
9. Agampodi SB. Spatial epidemiology of leptospirosis in Sri Lanka. Epidemiology and Infection. 2012;140(8):1530-1.
10. Evangelista KV, Coburn J. Leptospira as an emerging pathogen: a review of its biology, pathogenesis and host immune responses. Future Microbiology. 2010;5(9):1413-25.
11. Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, Lovett MA, et al. Leptospirosis: a zoonotic disease of global importance. Lancet Infectious Diseases. 2003;3(12):757-71.
12. Levett PN. Leptospirosis. Clinical Microbiology Reviews. 2001;14(2):296-+.
13. CDC. Leptospirosis CDC2014 [cited 2015]. Available from: http://www.cdc.gov/leptospirosis/.
14. Ullmann LS, Langoni H. Interactions between environment, wild animals and human leptospirosis. Journal of Venomous Animals and Toxins Including Tropical Diseases. 2011;17(2):119-29.
15. Hartskeerl RA, Collares-Pereira M, Ellis WA. Emergence, control and re-emerging leptospirosis: dynamics of infection in the changing world. Clinical Microbiology and Infection. 2011;17(4):494-501.
16. Gouvela EL, Metcalfe J, de Carvalho ALF, Aires TSF, Villasboas-Bisneto JC, Queirroz A, et al. Leptospirosis-associated severe pulmonary hemorrhagic syndrome, Salvador, Brazil. Emerging Infectious Diseases. 2008;14(3):505-8.

17. Health BMo. Leptospirosis Surveillience Database. 2015.
18. de Souza VMM, Arsky M, de Castro APB, de Araujo WN. Years of potential life lost and hospitalization costs associated with leptospirosis in Brazil. Revista De Saude Publica. 2011;45(6).
19. Pereira C, Barata M, Trigo A. Social Cost of Leptospirosis Cases Attributed to the 2011 Disaster Striking Nova Friburgo, Brazil. International Journal of Environmental Research and Public Health. 2014;11(4):4140-57.
20. Cann KF, Thomas DR, Salmon RL, Wyn-Jones AP, Kay D. Extreme water-related weather events and waterborne disease. Epidemiology and Infection. 2013;141(4):671-86.
21. Deng Z, Xun H, Zhou M, Jiang B, Wang S, Guo Q, et al. Impacts of Tropical Cyclones and Accompanying Precipitation on Infectious Diarrhea in Cyclone Landing Areas of Zhejiang Province, China. International Journal of Environmental Research and Public Health. 2015;12(2):1054-68.
22. Felzemburgh RDM, Ribeiro GS, Costa F, Reis RB, Hagan JE, Melendez AXTO, et al. Prospective Study of Leptospirosis Transmission in an Urban Slum Community: Role of Poor Environment in Repeated Exposures to the Leptospira Agent. Plos Neglected Tropical Diseases. 2014;8(5).
23. Lau CL, Clements ACA, Skelly C, Dobson AJ, Smythe LD, Weinstein P. Leptospirosis in American Samoa - Estimating and Mapping Risk Using Environmental Data. Plos Neglected Tropical Diseases. 2012;6(5).
24. Leal-Castellanos CB, Garcia-Suarez R, Gonzalez-Figueroa E, Fuentes-Allen JL, Escobedo-De La Pena J. Risk factors and the prevalence of leptospirosis infection in a rural community of Chiapas, Mexico. Epidemiology and Infection. 2003;131(3):1149-56.
25. Sakundarno M, Bertolatti D, Maycock B, Spickett J, Dhaliwal S. Risk Factors for Leptospirosis Infection in Humans and Implications for Public Health Intervention in Indonesia and the Asia-Pacific Region. Asia-Pacific Journal of Public Health. 2014;26(1):15-32.
26. Schneider MC, Najera P, Aldighieri S, Bacallao J, Soto A, Marquino W, et al. Leptospirosis Outbreaks in Nicaragua: Identifying Critical Areas and Exploring Drivers for Evidence-Based Planning. International Journal of Environmental Research and Public Health. 2012;9(11):3883-910.
27. Wasinski B, Dutkiewicz J. Leptospirosis - current risk factors connected with human activity and the environment. Annals of Agricultural and Environmental Medicine. 2013;20(2):239-44.
28. Wiwanitkit V. Climate change and leptospirosis. Ciencia & Saude Coletiva. 2012;17(12):3451-.
29. Victoriano AFB, Smythe LD, Gloriani-Barzaga N, Cavinta LL, Kasai T, Limpakarnjanarat K, et al. Leptospirosis in the Asia Pacific region. Bmc Infectious Diseases. 2009;9.
30. Schneider MC, Jancloes M, Buss DF, Aldighieri S, Bertherat E, Najera P, et al. Leptospirosis: A Silent Epidemic Disease. International Journal of Environmental Research and Public Health. 2013;10(12):7229-34.

31. Richard S, Oppliger A. Zoonotic occupational diseases in forestry workers - Lyme borreliosis, tularemia and leptospirosis in Europe. Annals of Agricultural and Environmental Medicine. 2015;22(1):43-50.
32. Lau CL, Smythe LD, Craig SB, Weinstein P. Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? Transactions of the Royal Society of Tropical Medicine and Hygiene. 2010;104(10):631-8.
33. Kucerova P, Cermakova Z. Leptospirosis: a neglected zoonosis of global distribution. Reviews in Medical Microbiology. 2013;24(3):63-9.
34. Batchelor TWK, Stephenson TS, Brown PD, Amarakoon D, Taylor MA. Influence of climate variability on human leptospirosis cases in Jamaica. Climate Research. 2012;55(1):79-90.
35. Forouhar E, Mitsani D. Sporadic urban leptospirosis. Journal of community hospital internal medicine perspectives. 2011;1(1).
36. de Araujo WN, Finkmoore B, Ribeiro GS, Reis RB, Felzemburgh RDM, Hagan JE, et al. Knowledge, Attitudes, and Practices Related to Leptospirosis among Urban Slum Residents in Brazil. American Journal of Tropical Medicine and Hygiene. 2013;88(2):359-63.
37. Dias JP, Teixeira MG, Costa MCN, Mendes CMC, Guimaraes P, Reis MG, et al. Factors associated with Leptospira sp infection in a large urban center in northeastern Brazil. Revista Da Sociedade Brasileira De Medicina Tropical. 2007;40(5):499-504.
38. Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li WH, Nobre CA. Climate change, deforestation, and the fate of the Amazon. Science. 2008;319(5860):169-72.
39. Goncalves DD, Teles PS, Rosimarie dos Reis C, Ruiz Lopes FM, Freire RL, Navarro IT, et al. Seroepidemiology and occupational and environmental variables for leptospirosis, brucellosis and toxoplasmosis in slaughterhouse workers in the Parana State, Brazil. Revista do Instituto de Medicina Tropical de Sao Paulo. 2006;48(3):135-40.
40. Goncalves DD, Benitez A, Ruiz Lopes-Mori FM, Alves LA, Freire RL, Navarro IT, et al. Zoonoses in humans from small rural properties in Jataizinho, Parana, Brazil. Brazilian Journal of Microbiology. 2013;44(1):125-31.
41. Lacerda HG, Monteiro GR, Oliveira CCG, Suassuna FB, Queiroz JW, Barbosa JDA, et al. Leptospirosis in a subsistence farming community in Brazil. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2008;102(12):1233-8.
42. Kalluri S, Gilruth P, Rogers D, Szczur M. Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: A review. Plos Pathogens. 2007;3(10):1361-71.
43. McMichael AJ, Woodruff RE, Hales S. Climate change and human health: present and future risks. Lancet. 2006;367(9513):859-69.
44. Eisen L, Lozano-Fuentes S. Use of Mapping and Spatial and Space-Time Modeling Approaches in Operational Control of Aedes aegypti and Dengue. Plos Neglected Tropical Diseases. 2009;3(4).
45. Nobre AA, Schmidt AM, Lopes HF. Spatio-temporal models for mapping the incidence of malaria in Para. Environmetrics. 2005;16(3):291-304.

46. Hu Y, Li R, Bergquist R, Lynn H, Gao F, Wang Q, et al. Spatio-temporal Transmission and Environmental Determinants of Schistosomiasis Japonica in Anhui Province, China. Plos Neglected Tropical Diseases. 2015;9(2).
47. Team RC. R: A language and environment for statistical computing. R: Foundation for Statistical Computing. Vienna, Austria2015.
48. ESRI. ArcGIS Desktop. Release 10 Redlands, CA: Environmental Systems Research Institute; 2011.
49. BioMedware. ClusterSEER 2.0. 2003.
50. Anselin LI, S.; and Youngihn, K. . GeoDa: An Introduction to Spatial Data Analysis. Geographical Analysis. 2006;38(1):5-22.
51. Brazil Census [Internet]. 2010. Available from: http://www.ibge.gov.br/english/.
52. GADM database of Global Administrative Areas [Internet]. 2012. Available from: http://www.gadm.org.
53. Cleveland RB, Cleveland, W.S., McRae, J.E., Terpennin, I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics. 1990;6(1):3.
54. Fitzmaurice GL, N.; Ware, J. . Allied Longitudinal Analysis. 2 ed: John Wiley & Sons; 2012.
55. Hijmans RJC, S.E.; Parra, J.L.; Jones, P.G. and Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 2005;25:1965-78.
56. Hengl TdJ, JM.; MacMillan, RA; Batjes, NH. . SoilGrids: an automated system for global soil mapping. World Soil Information; 2015.
57. Administration NAaS. Near Earth Observations. In: NASA, editor. http://neo.sci.gsfc.nasa.gov/about/2016.
58. Ignizio MSODaDA. Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. Geological Survey Data Series US Geological Survey. 2012;691.
59. Gracie R, Barcellos C, Magalhaes M, Souza-Santos R, Guimaraes Barrocas PR. Geographical Scale Effects on the Analysis of Leptospirosis Determinants. International Journal of Environmental Research and Public Health. 2014;11(10):10366-83.
60. Durski KN, Jancloes M, Chowdhary T, Bertherat E. A Global, Multi-Disciplinary, Multi-Sectorial Initiative to Combat Leptospirosis: Global Leptospirosis Environmental Action Network (GLEAN). International Journal of Environmental Research and Public Health. 2014;11(6):6000-8.
61. Cutter SL, Ismail-Zadeh A, Alcantara-Ayala I, Altan O, Baker DN, Briceno S, et al. Pool knowledge to stem losses from disasters. Nature. 2015;522(7556):277-9.

Appendix: Tables and Figures

Table 1. Source of Independent Variable Data

Data	Source	Time Scale	Spatial Scale*	Description of Data
Total Precipitation	NEO (57)	Monthly over years	0.1 degree	Total precipitation in mm x 10 for each month from NASA remote sensing data
Mean Temperature	WorldClim (55)	Monthly	2.5 km	Mean temperature in ℃ by month x 10
Minimum Temperature	WorldClim (55)	Monthly	2.5 km	Minimum temperature in ℃ by month x 10
Maximum Temperature	WorldClim (55)	Monthly	2.5 km	Maximum temperature in ℃ by month x 10
Isothermality	WorldClim (55)	Time invariant	2.5 km	Percentile (oscillation in day to night temperatures relative to annual oscillations of summer to winter)
Altitude	NEO (57)	Time invariant	90 m	Altitude at 90 m resolution from NASA remote sensing
Soil water pH	Soilgrids (56)	Time invariant	1 km	Soil pH measured through water, x 10
Population	Brazil Institute of Geography and Statistics (51)	Time invariant	*Municipality*	2010 Brazil Census
% Population: Urban	Brazil Institute of Geography and Statistics (51)	Time invariant	*Municipality*	2010 Brazil Census
Vegetation	NEO (57)	Monthly over years	0.1 degree	Normalized difference vegetation index (NDVI) with an index of 0.0 to 0.8.

*These values were averaged by municipality

Table 2: Characteristics of Sample

Characteristic	Mean ±SD	Median	IQR	% Missing
Cases of Leptospirosis	2.16 ± 4.58	1	1-2	0
Log Incidence of Leptospirosis	0.609 ± 0.00105	0.54	-0.599-1.796	0
Total Precipitation (mm)	188.89 ± 31.35	194.45	169.0-212.5	6.7
Mean Temperature (°□) x 10	217.51 ± 34.89	223.60	196.10-244.10	$1.8x10^{-4}$
Minimum Temperature(°□) x 10	164.66 ± 37.98	171.00	140.80-192.90	$1.8x10^{-4}$
Maximum Temperature (°□)x 10	270.09 ± 34.30	276.12	250.90-297.40	$1.8x10^{-4}$
Isothermality	62.62 ± 8.81	63.00	55.71-67.86	$1.8x10^{-4}$
Altitude (m)	370.56 ± 316.08	267.23	78.320-633.60	$3.6x10^{-4}$
Soil water pH x 10	51.56 ± 3.53	51.61	49.57-53.40	0
Population	$3.68 x10^5 \pm 1.14x10^6$	$7.89x10^4$	21,400-294,600	0
% Population Urban	81.7 ± 21.50	91.55	68.60-98.50	$5.4x10^{-4}$
Vegetation (NDVI Index)	164.26 ± 31.93	171.00	151.0-186.0	44.3 (0 after imputation)

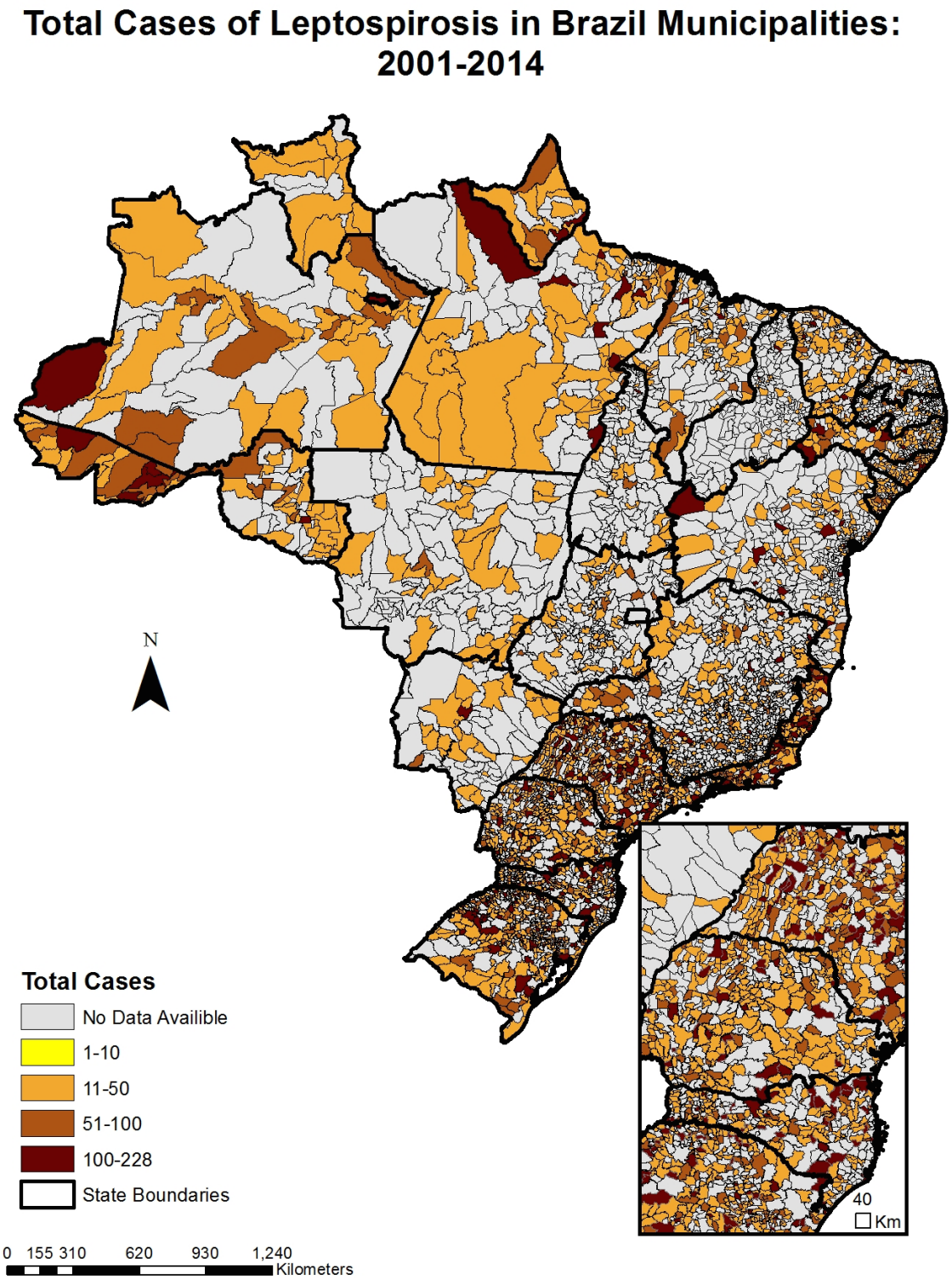

Figure 1: Total leptospirosis cases over study years presented by municipality

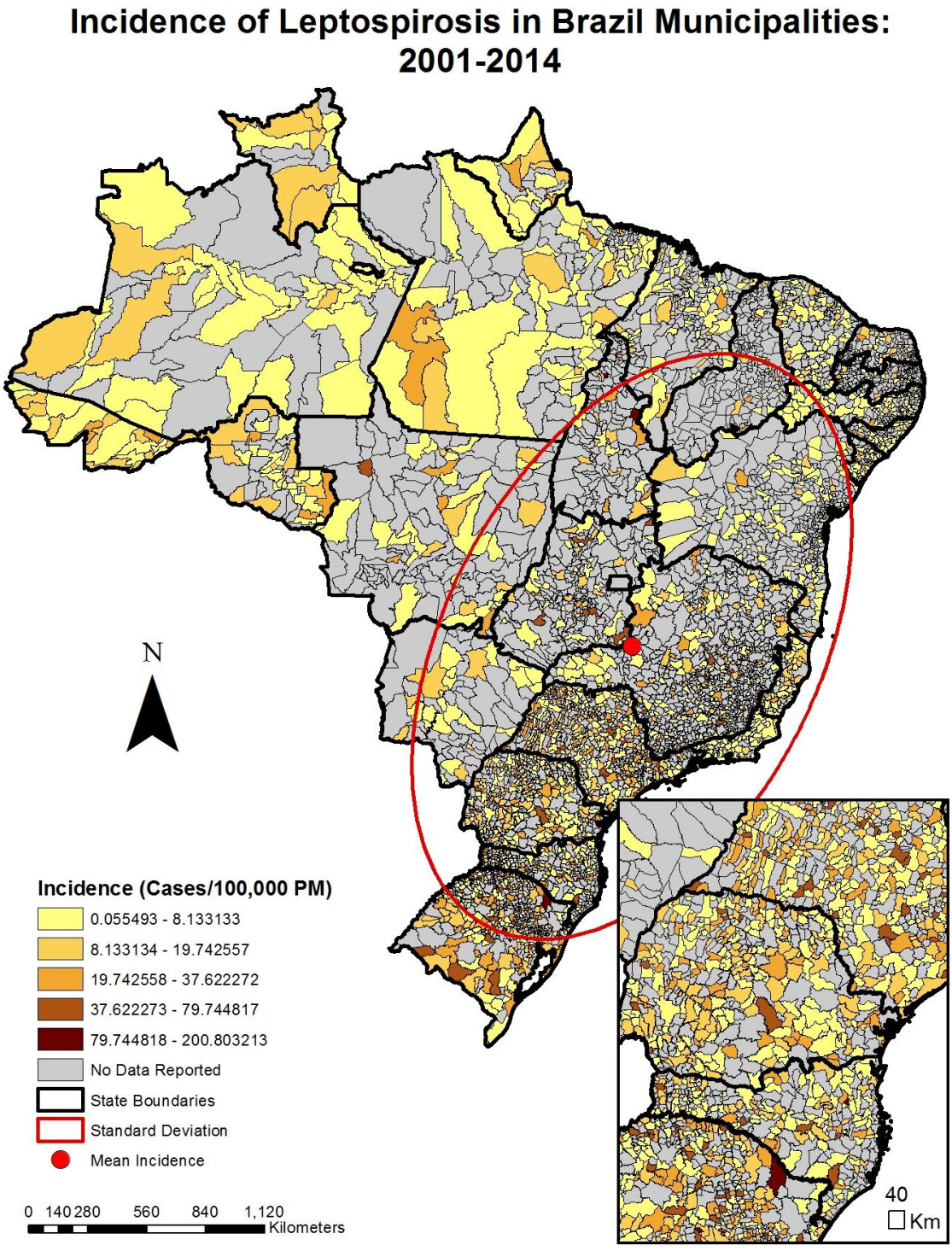

Figure 2: Total leptospirosis incidence over study years with median and standard deviational ellipse values displayed to show incidence distribution

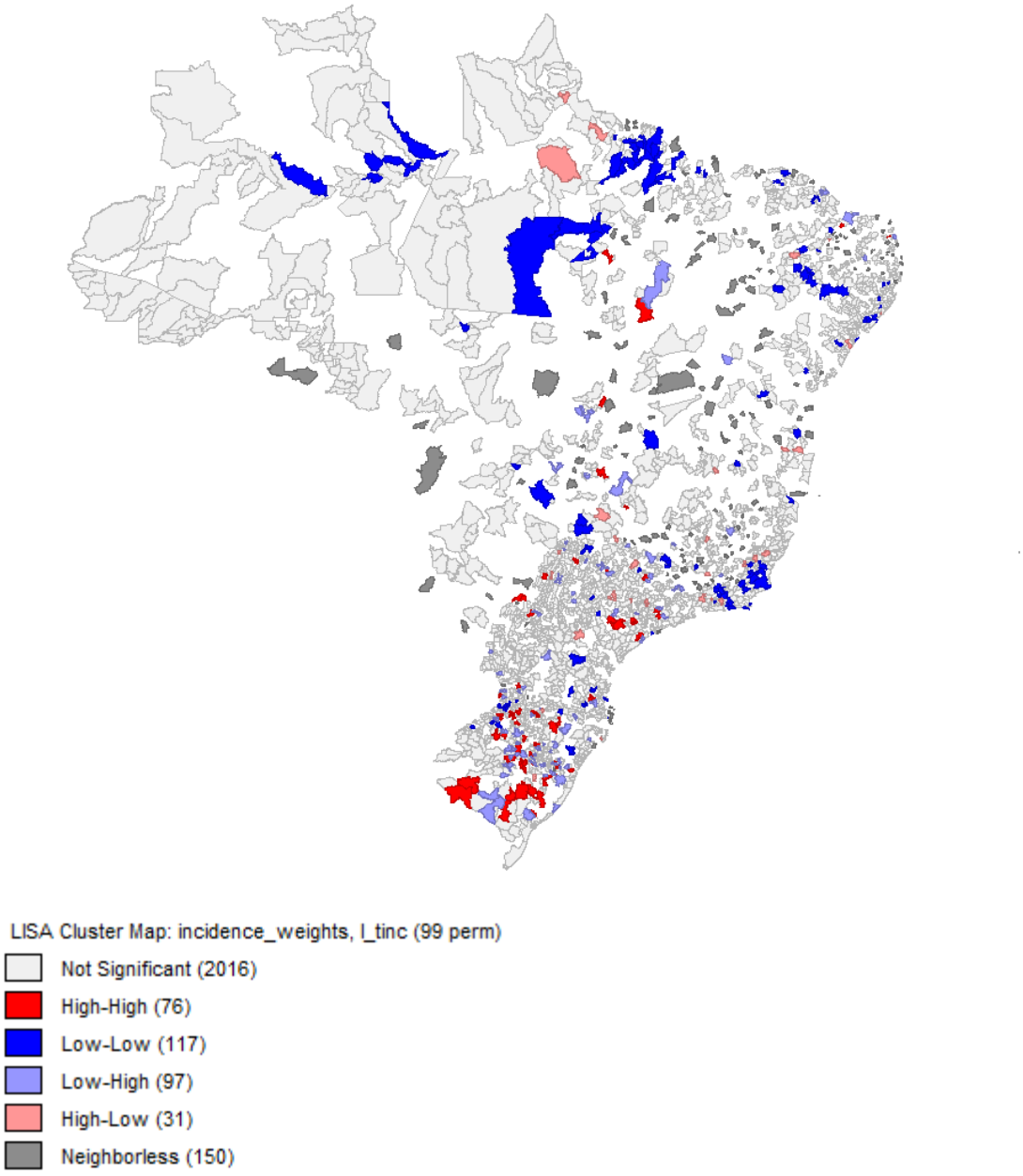

Figure 3: Local indicators of spatial association (LISA) of total incidence of leptospirosis/100,000 person months using the Local Moran I statistic, with a queens contiguity weighting. High-High clusters indicate municipalities with high incidence rates that border other high incidence municipalities. Low-Low clusters highlight the opposite. Low-High and High-Low clusters are evidence of outliers, with municipalities of high and low incidence neighboring each other.

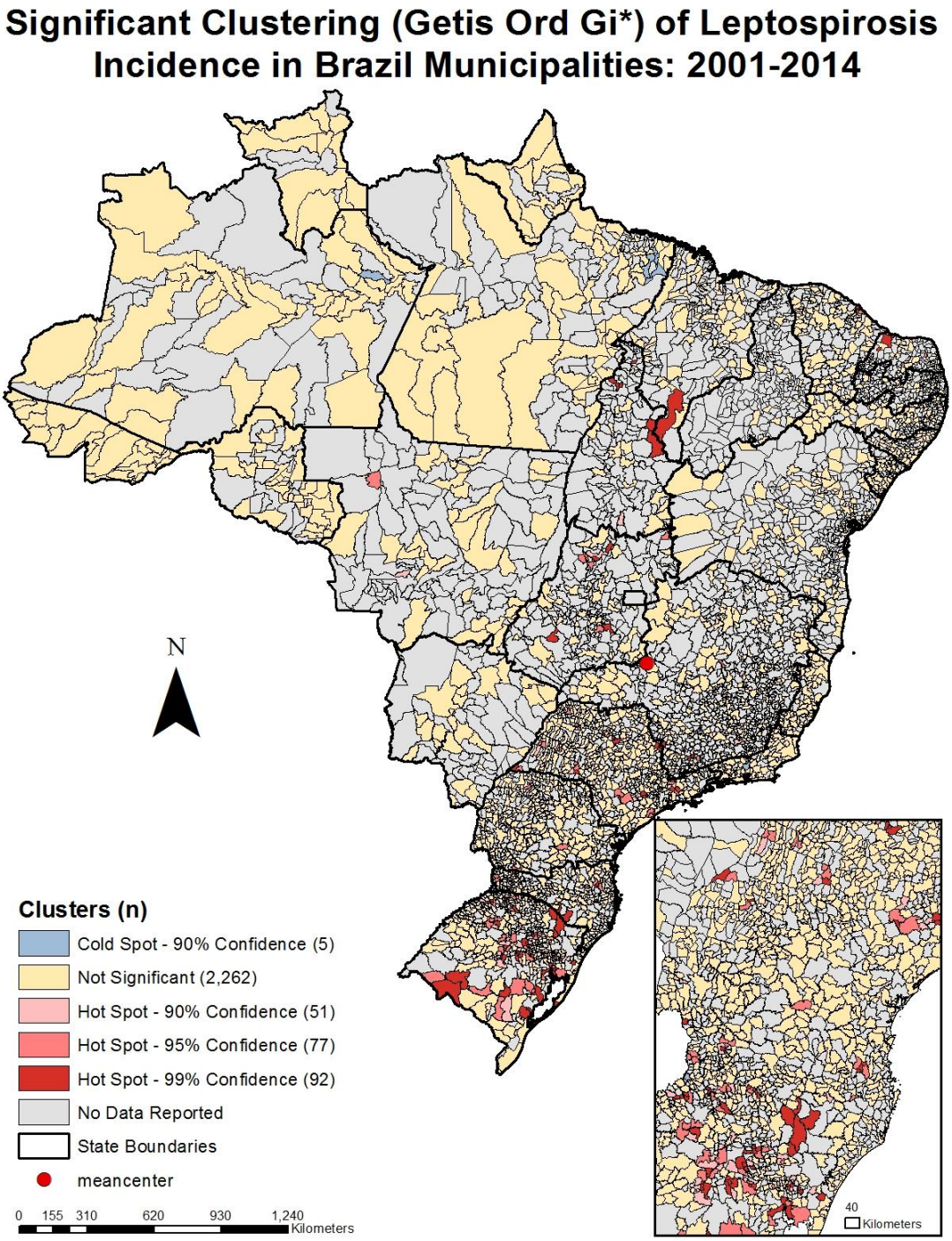

Figure 4: Hotspot and Cold Spot analysis of total incidence of leptospirosis/100,000 person months using the Getis-Ord Gi* statistic with a **euclidean inverse distance measurement**.

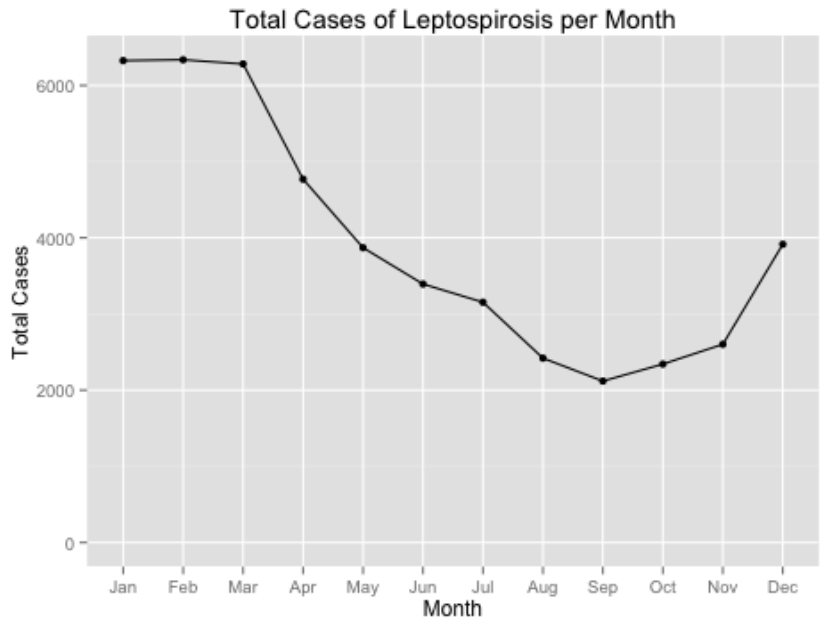

Figure 5: Total cases of leptospirosis by month. Cases peak from January-March.

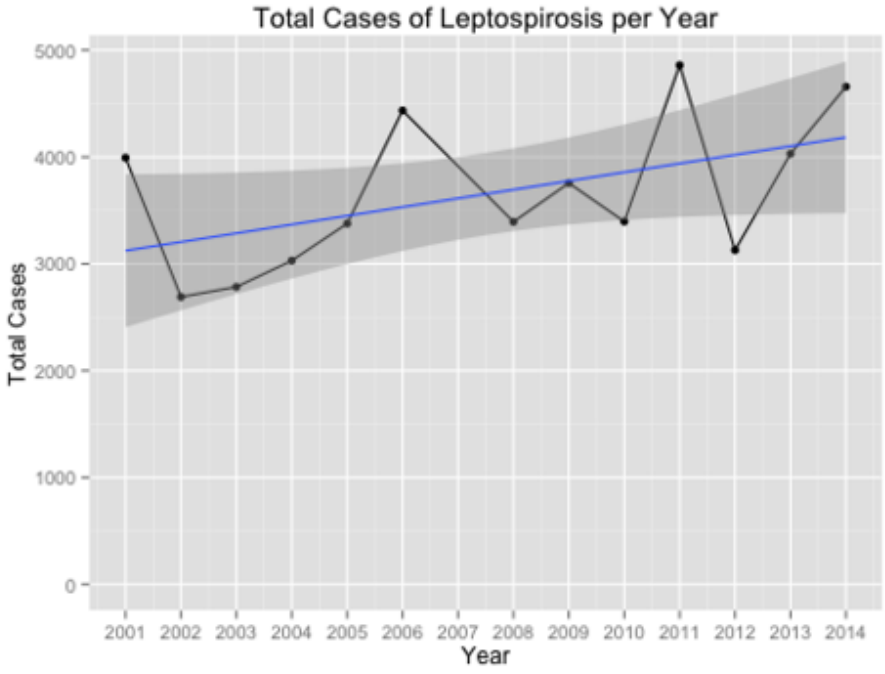

Figure 6: Total cases of leptospirosis by year with a fitted linear regression at 95% confidence intervals. Case totals vary by year, with a slight positive trend illustrated by the regression line.

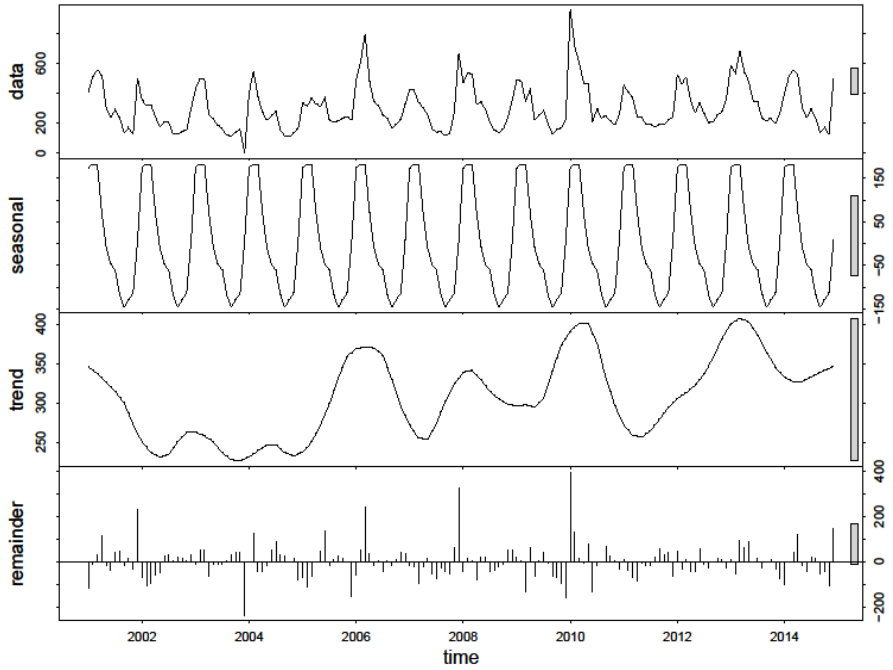

Figure 7a: Time series analysis of leptospirosis cases in Brazil over time using a seasonal trend decomposition.

Table 3: Taxonomy of multilevel models for change fitted to log transformed leptospirosis incidence

	Level-1/Level-2 specification		
Model	Level-1 model	Level-2 model	Composite Model
A	$Y_{ij} = \pi_{0i} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$	$Y_{ij} = \gamma_{00} + (\varepsilon_{ij} + \zeta_{0i})$
B	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{10}TIME_{ij} + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
C	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}PRECIP_{ij}\ \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{10}TIME_{ij} + \gamma_{20}PRECIP_{ij} + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij} + PRECIP_{ij})$
D	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{02}URBAN_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{12}URBAN_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{02}URBAN_i + \gamma_{10}TIME_{ij} + \gamma_{11}TIME_{ij}URBAN_i\ (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
E	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{03}TEMP(mean)_i + \gamma_{04}YEAR_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{13}TEMP(mean)_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{03}TEMP(mean)_i + \gamma_{04}YEAR_i + \gamma_{10}TIME_{ij} + \gamma_{13}TIME_{ij}TEMP(mean)_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
F	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{07}ISOTHERM_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{17}ISOTHERM_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{07}ISOTHERM_i + \gamma_{10}TIME_{ij} + \gamma_{17}TIME_{ij}ISOTHERM_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
G	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{08}ELEVATION_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{18}ELEVATION_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{08}ELEVATION_i + \gamma_{10}TIME_{ij} + \gamma_{18}TIME_{ij}ELEVATION_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
H	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{09}SOILPH_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{19}SOILPH_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{09}SOILPH_i + \gamma_{10}TIME_{ij} + \gamma_{19}TIME_{ij}SOILPH_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
I	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}VEG_{ij}\ \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{10}TIME_{ij} + \gamma_{30}VEG_{ij} + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij} + VEG_{ij})$
J	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{05}TEMP(min)_i + \gamma_{04}YEAR_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{15}TEMP(min)_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{05}TEMP(min)_i + \gamma_{04}YEAR_i + \gamma_{10}TIME_{ij} + \gamma_{15}TIME_{ij}TEMP(min)_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
K	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{06}TEMP(max)_i + \gamma_{04}YEAR_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{16}TEMP(max)_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{06}TEMP(max)_i + \gamma_{04}YEAR_i + \gamma_{10}TIME_{ij} + \gamma_{16}TIME_{ij}TEMP(max)_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
L	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{010}TEMP(range)_i + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{110}TEMP(range)_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{010}TEMP(range)_i + \gamma_{10}TIME_{ij} + \gamma_{110}TIME_{ij}TEMP(range)_i + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij})$
M	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}PRECIP_{ij} + \pi_{2i}VEG_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{02}URBAN_i + \gamma_{03}TEMP(mean)_i + \gamma_{04}YEAR_i + \gamma_{07}ISOTHERM_i + \gamma_{09}SOILPH + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{03}TEMP(mean)_i + \gamma_{07}ISOTHERM_i + \gamma_{09}SOILPH + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{02}URBAN_i + \gamma_{03}TEMP(mean)_i + \gamma_{04}YEAR_i + \gamma_{13}TIME_{ij}TEMP(mean)_i + \gamma_{07}ISOTHERM_{ij} + \gamma_{17}TIME_{ij}ISOTHERM_i + \gamma_{09}SOILPH_i + \gamma_{19}TIME_{ij}SOILPH_i + \gamma_{10}TIME_{ij} + \gamma_{20}PRECIP_{ij} + \gamma_{30}VEG_{ij} + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij} + PRECIP_{ij})$
N	$Y_{ij} = \pi_{0i} + \pi_{1i}TIME_{ij} + \pi_{2i}PRECIP_{ij} + \pi_{2i}VEG_{ij} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{02}URBAN_i + \gamma_{03}TEMP(mean)_i + \gamma_{04}YEAR_i + \gamma_{07}ISOTHERM_i + \gamma_{09}SOILPH + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{07}ISOTHERM_i + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{02}URBAN_i + \gamma_{03}TEMP(mean)_i + \gamma_{04}YEAR_i + \gamma_{07}ISOTHERM_{ij} + \gamma_{17}TIME_{ij}ISOTHERM_i + \gamma_{09}SOILPH_i + \gamma_{10}TIME_{ij} + \gamma_{20}PRECIP_{ij} + \gamma_{30}VEG_{ij} + (\varepsilon_{ij} + \zeta_{0i} + \zeta_{1i}TIME_{ij} + PRECIP_{ij})$

Table 4: Results of fitting taxonomic multilevel models for change to log transformed leptospirosis incidence data

		Parameter	Model A	Model B	Model C	Model D	Model E	Model F	Model G	Model H
Fixed Effects										
Initial status, π_{0i}	Intercept	γ_{00}	1.705***	1.728***	1.095***	3.487***	1.264***	3.820***	1.670***	0.409
	Precipitation	γ_{01}	--	--	--	--	--	--	--	--
	Urban	γ_{02}	--	--	--	-0.025***	--	--	--	--
	Mean Temperature	γ_{03}	--	--	--	--	0.0014***	--	--	--
	Year		--	--	--	--	0.238***	--	--	--
	Isothermality	γ_{04}	--	--	--	--	--	-0.033***	--	--
	Altitude	γ_{05}	--	--	--	--	--	--	0.00012	--
	Soil Water pH	γ_{06}	--	--	--	--	--	--	--	0.0252***
	Vegetation	γ_{07}	--	--	--	--	--	--	--	--
	Min Temperature	γ_{08}	--	--	--	--	--	--	--	--
	Max Temperature	γ_{09}	--	--	--	--	--	--	--	--
	Temperature Range	γ_{010}	--	--	--	--	--	--	--	--
Rate of change, π_{1i}	Intercept	γ_{10}	--	-0.0002***	-0.0001	-0.00001	-0.0213***	-0.0003***	-0.00064***	0.00372***
	Precipitation	γ_{11}			0.00314***	--	--	--	--	--
	Urban	γ_{12}	--	--	--	-0.000004	--	--	--	--
	Mean Temperature	γ_{13}	--	--	--	--	0.000075*	--	--	--
	Isothermality	γ_{14}	--	--	--	--	--	0.000047***	--	--
	Altitude	γ_{15}	--	--	--	--	--	--	0.0000010***	--
	Soil Water pH	γ_{16}	--	--	--	--	--	--	--	-0.00008***
	Vegetation	γ_{17}	--	--	--	--	--	--	--	--
	Min Temperature	γ_{18}	--	--	--	--	--	--	--	--
	Max Temperature	γ_{19}	--	--	--	--	--	--	--	--
	Temperature Range	γ_{110}	--	--	--	--	--	--	--	--
Variance components										
Level-1:	Within-person	σ^2_ε	0.283***	0.524***	0.524***	0.514***	0.514***	0.524***	0.524***	0.523
Level-2:	In initial status	σ^2_0	1.269***	1.185***	1.162***	1.034***	1.191***	1.151***	1.182***	1.180
	In rate of change	σ^2_1	--	0.00164***	0.00167***	0.00126***	0.00167***	0.00161***	0.00160***	0.00163
	Covariance	σ_{01}	--	-0.156**	-0.177**	-0.182**	-0.141**	-0.102**	-0.145**	-0.142
Goodness-of-fit										
	Deviance		--	-20863.50	-19465.99	-20554.4	-20520.85	-20816.08	-20870.12	-20869.18
	AIC		42028.55	41738.99	38945.98	41124.80	41059.70	41648.16	41756.24	41754.37
	BIC		42045.90	41787.02	39001.53	41188.84	41131.74	41712.20	41820.27	41818.41

Table 4 Cont.: Results of fitting taxonomic multilevel models for change to log transformed leptospirosis incidence data

		Parameter	Model I	Model J	Model K	Model L	Model M	Model N	Model N*
Fixed Effects									
Initial status, π_{0i}	Intercept	γ_{00}	1.055***	1.247***	1.138***	1.770***	3.961***	4.062***	4.720***
	Precipitation	γ_{01}	--	--	--	--	--	--	
	Urban	γ_{02}	--	--	--	--	-0.0264***	-0.0265***	-0.0271***
	Mean Temperature	γ_{03}	--	--	--	--	0.00154***	0.00180***	-0.00002
	Year		--	0.233***	0.234***	--	0.2124***	0.2199***	0.0809***
	Isothermality	γ_{04}	--	--	--	--	-0.0382***	-0.0399***	-0.0330***
	Altitude	γ_{05}	--	--	--	--	--	--	--
	Soil Water pH	γ_{06}	--	--	--	--	0.0192***	0.0171***	0.0103***
	Vegetation	γ_{07}	--	--	--	--	--	--	--
	Min Temperature	γ_{08}	--	0.0020***	--	--	--	--	--
	Max Temperature	γ_{09}	--	--	0.0006***	--	--	--	--
	Temperature Range	γ_{010}	--	--	--	-0.00040	--	--	--
Rate of change, π_{1i}	Intercept	γ_{10}	-0.00013	-0.020***	-0.0227***	-0.0026***	-0.0196***	-0.0203***	-0.008***
	Precipitation	γ_{11}	--	--	--	--	0.0011***	0.0011***	0.0003***
	Urban	γ_{12}	--	--	--	--	--	--	--
	Mean Temperature	γ_{13}	--	--	--	--	0.00003	--	--
	Isothermality	γ_{14}	--	--	--	--	0.000041***	0.00005***	0.000017***
	Altitude	γ_{15}	--	--	--	--	--	--	--
	Soil Water pH	γ_{16}	--	--	--	--	-0.000022	--	--
	Vegetation	γ_{17}	0.0028***	--	--	--	0.00188***	0.00190***	0.00103***
	Min Temperature	γ_{18}	--	0.000004	--	--	--	--	--
	Max Temperature	γ_{19}	--	--	0.00012***	--	--	--	--
	Temperature Range	γ_{110}	--	--	--	0.000022***	--	--	--
Variance components									
Level-1:	Within-person	σ^2_ε	0.518***	0.513***	0.514***	0.524***	0.514***	0.514***	0.609***
Level-2:	In initial status	σ^2_0	1.113***	1.200***	1.190***	1.184**	0.967***	0.968***	0.914***
	In rate of change	σ^2_1	0.0015	0.00168***	0.00165***	0.0016***	0.0017***	0.0017***	0.0013***
	Covariance	σ_{01}	-0.205	-0.149*	-0.130*	-0.150*	-0.148*	-0.150*	-0.174*
Goodness-of-fit									
	Deviance		-8479.57	-20491.23	-20549.75	-20871.31	-18795.92	-18777.85	-1321.05
	AIC		16973.14	41000.46	41117.50	41758.62	37623.83	37583.71	26452.11
	BIC		17022.98	41072.50	41189.54	41822.66	37750.77	37694.78	26571.12

***p<0.05, *p<0.10

Table 5: Error Covariance Structure for Selected Model

	SYMM[a]	EXP	GAUS	LIN	CS	RATIO	AR(1)	*CAR(1)**
Parameter	120	15	15	15	2	2	2	*2*
Deviance	--	-13211.05	-15736.33	-15301.42	-18777.85	-14190.69	-13211.05	*-13211.05*
AIC	--	26452.11	31502.67	30632.84	37585.71	28411.38	26452.11	*26452.11*
AICC	--	--	--	--	--	--	--	--
BIC	--	26571.12	31621.68	30751.85	37704.71	28530.39	26571.12	*26571.12*

SYMM= General correlation matrix EXP=Exponential; GAUS=Gaussian; LIN=linear; CS=compound symmetry; RATIO = Rational quadratics; AR(1) = Autoregressive; CAR(1) = Continuous autoregressive process.

*Chosen

[a] Did not converge

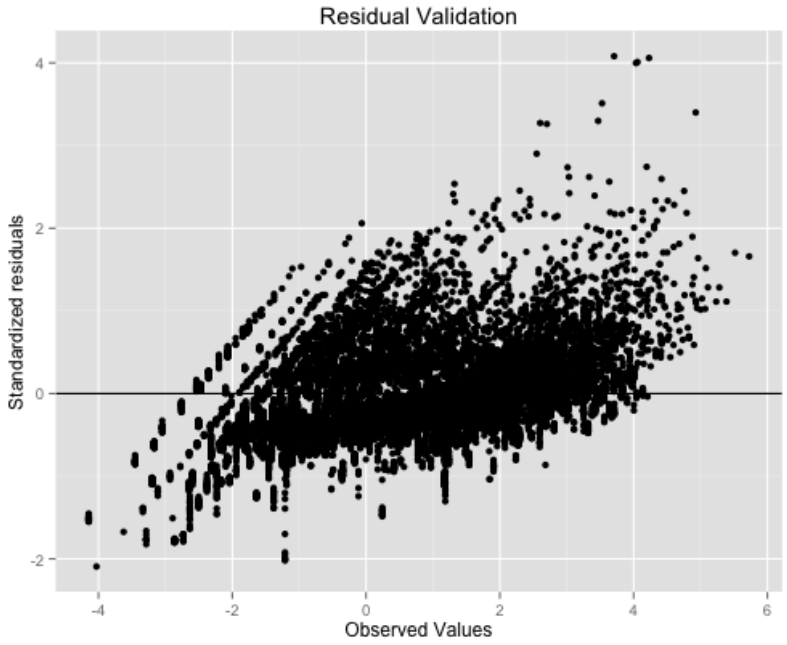

Figure 8. Observed values of leptospirosis rates and their corresponding standardized residuals. Clustering around the x-axis indicates a small residual value and suggest improved model fit.

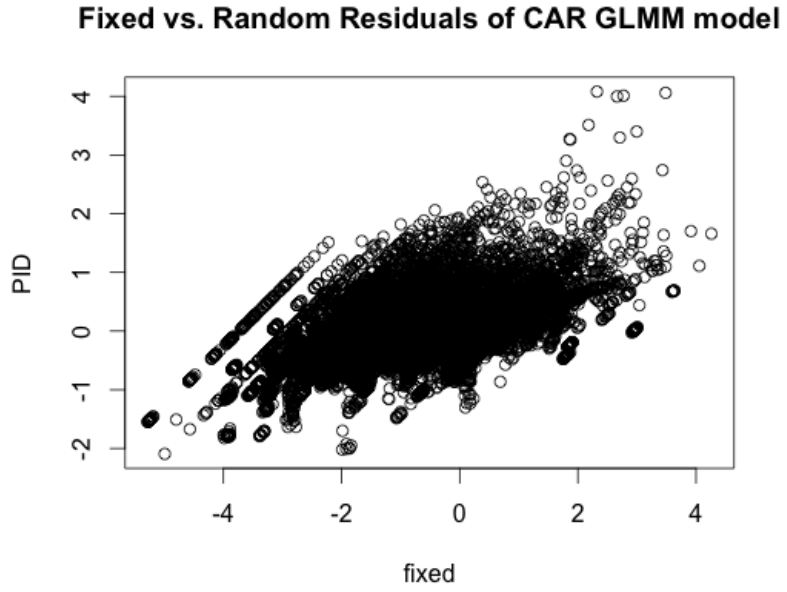

Figure 9. Fixed vs. Random residuals in the conditional autoregressive general linear mixed model. Clustering around the x and y-axis indicate small residual values and suggest improved model fit.