

Article

Exploratory Analysis of Dengue Fever Niche Variables within the Río Magdalena Watershed

Austin Stanforth ^{1,2}, Max J. Moreno-Madriñán ^{1,*} and Jeffrey Ashby ¹

¹ Department of Environmental Health Science, Fairbanks School of Public Health, Indiana University, IUPUI, Indianapolis, IN 46202, USA; Austin.Stanforth@gmail.com (A.S.); jlashby@imail.iu.edu (J.A.)

² Department of Geography, Indiana University, IUPUI, Indianapolis, IN 46202, USA

* Correspondence: mmorenom@iu.edu; Tel.: +1-317-274-3170

Academic Editors: Naser El-Sheimy, Zahra Lari, Adel Moussa, Deepak R. Mishra, Douglas G. Goodin, Xiaofeng Li and Prasad S. Thenkabail

Received: 1 July 2016; Accepted: 9 September 2016; Published: 19 September 2016

Abstract: Previous research on Dengue Fever have involved laboratory tests or study areas with less diverse temperature and elevation ranges than is found in Colombia; therefore, preliminary research was needed to identify location specific attributes of Dengue Fever transmission. Environmental variables derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) and Tropical Rainfall Measuring Mission (TRMM) satellites were combined with population variables to be statistically compared against reported cases of Dengue Fever in the Río Magdalena watershed, Colombia. Three-factor analysis models were investigated to analyze variable patterns, including a population, population density, and empirical Bayesian estimation model. Results identified varying levels of Dengue Fever transmission risk, and environmental characteristics which support, and advance, the research literature. Multiple temperature metrics, elevation, and vegetation composition were among the more contributory variables found to identify future potential outbreak locations.

Keywords: Dengue Fever; *Aedes aegypti*; remote sensing; GIS; factor analysis; vector modeling

1. Introduction

The resurgence of tropical diseases in the Americas, including Dengue Fever (DF), may be related to the suspension of vector eradication programs, such as those conducted by the Pan American Health Organization (PAHO) programs until the 1970s [1]. Implementation of new control programs require ongoing data collection. Cost limitation of in situ data collection has increased the use of remote sensing (RS) and Geographic Information Science (GIS) applications to estimate disease spread and population vulnerability [2,3]. Public health agencies can use such tools to implement monitoring and surveillance programs for multiple purposes, including the estimation of vector abundance, implement prevention strategies, or to control for further dispersion of the disease and vector [4,5]. Additionally, local metrics and research are needed to improve current vector disease monitoring practices before cost effective predictive or forewarning systems can be studied or implemented [6–9]. Therefore, the goal of this study was to investigate the variables which would provide the most benefit to future investigations of predictive modeling. The outputs of this study identified variables capable of inclusion in such studies, but the statistical methods used here are not designed for predictive purposes.

One of the more commonly implemented remote sensing (RS) tools for vector research has been the Moderate Resolution Imaging Spectroradiometer (MODIS). This dual satellite platform, Aqua and Terra, of NASA's Earth Observation System maintains a sun-synchronous orbit and is capable of collecting data over the entire Earth multiple times a day [10]. The high temporal resolution reduces the sensor's spatial capabilities, 250 m for multispectral data and 1 km for land surface temperature (LST), but its ability to collect a variety of spectral wavelengths daily has proven useful for

a variety of environmental assessments [11]. The satellite has already been implemented for a variety of studies. Fuller et al. [12] demonstrated how predicative modeling of DF can be achieved through temperature records and remote sensing products. Vector modeling could therefore be implemented into current autonomous monitoring systems, such as MODIS's Famine Early Warning Systems Network (FEWS NET [13]) [2,12,14–16]. NASA scientists have developed such a Malaria transmission model through the Goddard Space Flight Center's Health Planet program, which incorporates input variables for parasites, hosts, vectors, and environmental and human factors [4]. Similar vector warning products could be developed to improve disease dispersion monitoring practices for additional diseases [2,12].

The *Aedes aegypti* mosquito vector is capable of spreading a variety of diseases, including: Yellow Fever, Chikungunya, Zika, and DF. Disease transmission rates of this vector are considered to be influenced by both cultural and climatic conditions. Although commonly associated with tropical and subtropical climates, due to the habitable zone of the mosquito, changing climate conditions may expand the dispersion potential of these diseases [1,17–19]. This demonstrates why improved methods to identify habitable zones of this vector will become increasingly important to improve future disease mitigation and prevention practices.

Since no vaccine for DF is currently available, mosquito control practices are critical to reduce infection [1]. Identification of geographical areas to reduce mosquito populations and contact between the vector and people are essential to attenuate potential outbreaks. Previous research has identified flower pots, water storage containers, impervious surface catchments, and other anthropogenic water sources as opportunities for *Aedes aegypti* larval development [1,6]. These features are common to domesticated or developed areas, perpetuating the likelihood of the *Aedes aegypti* mosquito to be found in urban/populated environments, and increasing the opportunity for mature specimens to feed on human populations to disperse the virus [4,6,20].

The vector's habitat and disease transmission potential are largely set by given meteorological and elevation boundaries [3,21]. Identification of these vector boundaries, and the pathogen it carries, is known as detection of the zoonotic niche [22–25]. The variables defining this niche can be estimated through RS data to derive the geographic extent of the mosquito's habitat [3], and could become increasingly important to accurately monitor future transmission risk of DF. Similar habitat and dispersion studies are already being conducted for this and other vector diseases, including West Nile Virus and Malaria [2,4,7]. The focus of this manuscript is the DF virus, although some of the incorporated variables and discussion directly address the *Aedes aegypti* mosquito vector, which we assume the vector dispersion to be associated to the disease spread.

A study of DF in Singapore described an increase of 22%–184% in reported cases given a 2–10 °C rise in maximum or minimum temperature [8]. Additional laboratory and field research in Cali, Colombia have also demonstrated a strong relationship between DF dispersion and temperature range [26,27]. Under lab conditions, Brady et al. [28] reported low survival of the mosquito vector below 14–15 °C or above 35 °C [29]. Rueda et al. [30] similarly saw survival decline outside the 15–34 °C temperature range, but also documented an inverse relationship between body size and increasing temperature [30]. These temperature limits may also impact the latitude and altitude of the mosquito's habitat [21]. In Mexico and Colombia, the mosquito has been reported to be abundant up to 1700 [18,31] and 1800 m [32], respectively. However, more recently it was reported present (although not abundant) at 2130 m in Central Mexico [18] and 2300 m in Colombia [33]. Cultural parameters can involve the presence of anthropogenic water storing activities, which may be responsible for up to 79% of the reported cases [3,8]. Water storage tanks, potted plants, and similar peridomestic related containers can provide larval rearing sites proximal to populated areas.

Exploratory factor analysis has previously been used to investigate the relationship between a dependent variable and potential contributing, independent variables [34,35]. Principal component analysis (PCA) is a factor analysis method used to identify trends within a multivariate dataset to reduce or derive new, efficient, component variables [5,36]. Independent variables are plotted against the dependent variable during the PCA to identify linear relationships, or clusters, to build the component variable relationships. The amount of the original dataset variability a component variable encompasses is demonstrated by its eigenvalue; the larger the eigenvalues, the stronger the composite component. Composite variables with an eigenvalue of one (1), or greater, are classified as principal components according to the Kaiser Criteria as they describe at least as much variability as any one of the original variables [35,37]. Components which do not reach the Kaiser Criteria can be considered 'noise' and not implemented in future stages; therefore, fewer variables are used to describe the majority of the data [37,38]. These methods have been shown to reduce the amount of data used and identify trends in the dataset [8,36]. Other studies have used regression techniques, including the geographically weighted regression and auto-regression, methods for virus modeling [8,21,39], and boosted regression trees [22–25]. These methods can be computationally intense and difficult to implement. Although they were chosen for their respective studies due to their predictive ability, they sometimes incur complications with non-normalized data, such as that used in this study. Since the main goal of the present study was not prediction, but data reduction to identify strong determinant variables, the PCA methodology was identified to be more appropriate for this purpose [38,40].

This project was designed as an exploratory analysis to identify variables which may explain the annual incidence of DF within Colombia's Magdalena River watershed during the 2012–2014 study period. This study site is characterized by climate conditions ranging from sea level to permanent snow, and from desert to tropical rainforest. The wide range of climate conditions, which exceeds the presence of the upper and lower climate limits of the *Aedes aegypti*, makes this an isolated and ideal study site for this type of exploratory analysis. This methodology was intended to investigate whether the study site supports habitat and transmission information obtained about the vector during previous laboratory and other geographical area studies. Although this PCA analysis has a descriptive, rather than predictive, connotation, its results are aimed to guide future research projects intended to develop predictive models or early warning systems for the disease.

2. Materials and Methods

2.1. Study Site

Analysis was conducted in the Río Magdalena watershed in the South American country of Colombia. The approximately 273,000 km² area is bounded by a split in the Andes Mountain range on three sides and the Caribbean Sea coast to the north, as can be seen in the site map in Figure 1. The 2014 projected population of the watershed, including the capital city of Bogota, was calculated to be over thirty-six million. The watershed is in close proximity to the equator, indicating a warmer climate, and well within the currently-documented latitude range of the *Aedes aegypti* mosquito [8]. The elevation of the watershed's mountain boundaries should naturally isolate the *Aedes aegypti* vector from external areas, providing a study area with limited external influence [41,42]. Due to its location on the western side of South America, it is also influenced by El Niño/La Niña seasonal variations. DF has been present in this region for some time. Reports indicate the disease will continue to be an issue for the region, and may potentially reach epidemic proportions [32].

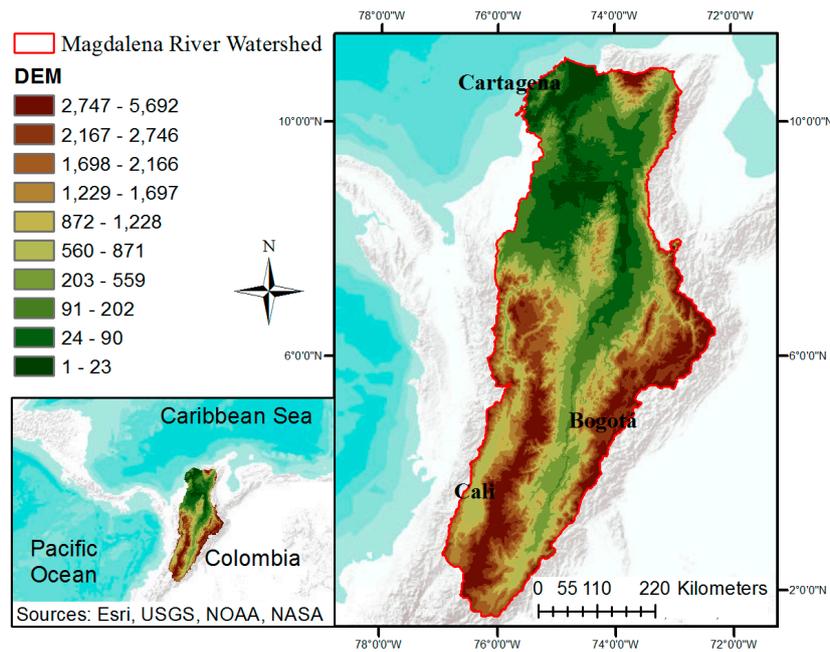


Figure 1. Study site identification of the Magdalena Watershed in Colombia, demonstrating its geographical location, elevation diversity, and presence of larger populated urban environments.

2.2. Population Data

Population variables incorporated into the analysis included socioeconomic data collected during the 2005 General Census administered by El Departamento Administrativo Nacional de Estadística (The National Administrative Department of Statistics). Projected populations for 2012–2014 were created by the data's proprietors, a sample of the data can be viewed in Table 1, and consisted of tabulated total population, rural and urban populations distinguished by gender [43]. Population variables were collected at the municipality resolution, which became the assessment size for this study. Colombian municipality boundaries were distinguished through an open access Esri country shapefile [44], which was clipped to the watershed boundary. Projected population levels and incidence rates for the three study periods can be identified in the following graph. Since the disease is well established, these numbers do not reflect prevalence, but basic annual infection rates based on the population size.

Table 1. Projected population and Dengue Fever Case records.

	Projected Population	New Reported Cases	New Cases Rate
2012	35,375,287	15,811	0.45/1000 people
2013	35,751,512	36,977	1.03/1000 people
2014	36,127,443	24,949	0.69/1000 people

2.3. Remote Sensing Data

Environmental variables were obtained from the U.S. Geological Survey's (USGS) Earth Explorer [45] and NASA [46] open access databases [47]. Daily MODIS data was collected using an automated MS-DOS batch file operation covering 2012–2014, independently. Due to the high volume of files, a Python 2.7 [48] script was composed for ArcGIS 10.3's [44] model builder to automatically process the satellite imagery to use a consistent projection file (WGS 1984 UTM Zone 18N), mosaic, and apply correction algorithms as the products required. Rather than use the eight-day MODIS composite product available through the database, the daily imagery was composited into seven day composite

raster files to coincide with the weekly calendar year. The composite data was necessary to reduce the impact of cloud cover [11]. All imagery was clipped to the watershed shapefile.

ArcMap's raster calculator tool [44] was used to convert MODIS MYD11A1 composited data to LST through additional scripting. Both daytime and nighttime thermal wavelength raster cells were multiplied by a gain factor of 0.02, then converted to Celsius by subtracting 273.15 [49]. Both day and night temperatures were included in this study to investigate previously reported low temperature [3,21] and high/low temperature range [26–28] impact on vector survivability. Another MODIS product, the MYD09GQ, was used to calculate the two-band enhanced vegetation index (EVI) ($2.5 \times [(NIR - Red)/(NIR + Red + 1)]$) [49]. The two-band EVI was used due to the MYD09GQ's 250 m spatial resolution, the prevalence of high biomass/cloud cover in the study region, the index's ability to differentiate vegetation canopy from background noise, and its proven usefulness in previous DF research [12,50].

Precipitation data was collected through the Tropical Rainfall Measuring Mission (TRMM); a joint venture between NASA and the Japanese Aerospace Exploration Agency launched in 1997 [51]. TRMM's daily temporal resolution has proved useful for vector research [8]. One of TRMM's three onboard sensors, the Visible Infrared Scanner (VIRS), provided the 3B42 V7 dataset, which reports global precipitation estimates at a resolution of 5 km. Although the TRMM was decommissioned in 2015, its flight time sufficiently covers the 2012–2014 study period.

Additional annual products for land use land cover (LULC) and elevation were incorporated into the analysis. The pre-processed MODIS MCD12Q1 product identified patterns of LULC at a resolution of 500 m. The data identifies major landscape elements, including: barren, cropland, forest, savanna, shrub land, snow and ice, urban, and wetland features [11]. The Shuttle Radar Topography Mission (SRTM) provided elevation data at a spatial resolution of 30 m. These resolutions were considered to be sufficient to quantify variations across the watershed's contour.

This project consisted of three stages for data acquisition and analysis, which is graphically represented in Figure 2. The acquisition and development of the weekly composite satellite data, previously described, represent the preprocessing part of Stage 1. Composited data were further compounded into annual values using Esri ArcMap's Cell Statistics tool. Cell statistics calculated an input variable's maximum, minimum, and average value at a pixel's location over the calendar year. The final part of Stage 1 consisted of zonal statistics being applied to the yearly composited cell values, using Colombian municipality boundaries. Variables acquired as yearly variables, elevation and LULC, did not require a cell statistics step, but were still processed through zonal statistics. The processing order demonstrates the naming structure of the variables used throughout the rest of this project. Nomenclature begins with the variable type (example LST or VEG), followed by the cell statistic (annual longitudinal: min, max, mean), and concludes with the zonal statistic calculation (cross-sectional per municipality: min, max, range, mean, standard deviation, sum). The zonal statistic step also provided a unique identification code per municipality to improve database organization and allow for the municipality population data to be joined to the environmental variables.

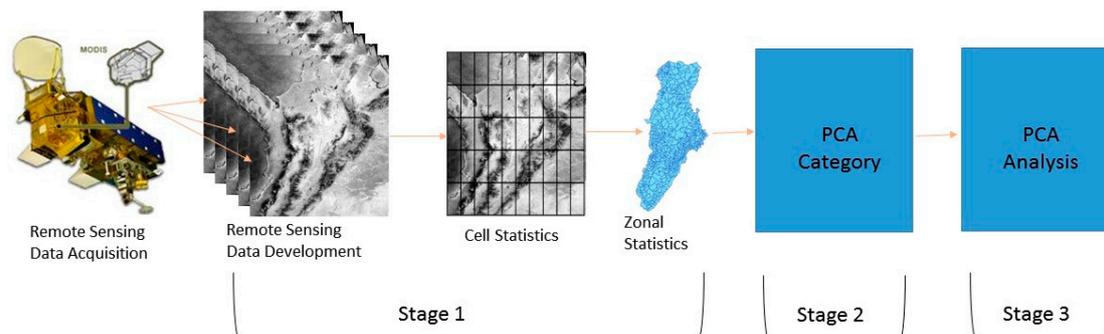


Figure 2. Graphic depicting stages of analysis.

Compounded statistical analysis on the environmental variables, both cell and zonal statistics, was used to investigate the complexity of environmental influences on transmission rates. Previous 'non-laboratory' studies often incorporated simple statistics into their models, such as the mean temperature, which may not quantify the most useful variable to study [6,22,27,52]. The complexity added by the two-stage statistics in this project allows for the assessment of variable diversity to be expanded from what is available in the literature [21,52]. For example, the application of the municipality zonal statistic on the night LST Min cell values explored the minimum, mean, and maximum zone values of each minimum pixel value with the zone. This methodology was anticipated to expand the individual values previously used and further the literature's understanding on environmental influencers. For example, the maximum zone value may better demonstrate whether the minimum survival temperature of the mosquito is reached within a municipality, which would be an improvement over the amount of information available through the mean value. Similar examples to this use of extended variables include the use of day and nighttime mean and temperature ranges for zoonotic niche studies in Africa [23], or max and minimum temperatures for DF in Singapore [8]. The involvement of additional data compounds the complexity of the dataset, further demonstrating the usefulness of a data reduction tool, such as PCA. This statistical tool is able to identify local attributes of infection and could be reproduced in additional study areas/time periods by investigators with limited programming or GIS experience.

2.4. Health Data

Municipality records of DF were generated and made available by the National Institute of Health (Instituto Nacional de Salud) of Colombia for the three study years [53]. Confirmed cases were identified through laboratory tests of patients exhibiting symptoms of: headache, ocular pain, myalgia, arthralgia, and/or rash. Specific virus specimen antibodies were identified by polymerase chain reaction (PCR) tests for fever patients who had symptoms for less than five days, while an ELISA (enzyme-linked immunosorbent assay) test was used for patients who had symptoms for more than five days [17]. The municipality records were compounded into annual records and merged to the shapefile database to be used as the dependent variable in the statistical analysis. Since this methodology incorporated an annual aggregation, it extended beyond the normal gestation or gonotrophic cycle of the virus and vector. Therefore, no temporal lag was needed to be incorporated into the methodology.

2.5. Principal Component Analysis

Following the creation of the complete database incorporating health data, population demographics, and yearly zonal environmental variables, Stage 2 conducted a PCA independently on each variable category using the IBM SPSS 23 statistical software's dimension reduction tool [54]. A category consisted of the zonal statistics derived from a single data source, such as the nighttime LST or precipitation data. The dependent variable was comprised of the number of confirmed cases of DF per municipality. This stage was used to identify the strongest unique variables within each variable category to include in Stage 3 [40]. A variable was considered for the subsequent stage if it strongly loaded in component 1 or 2 in both 2012 and 2014. The El Niño Southern Oscillation (ENSO) conditions between 2012 and 2014 were weak, thus, no important influence from El Niño or La Niña should have affected the studied relationship between meteorological variables and DF [55]. Stage 2 was implemented to remove variables which were highly correlated to one another and eliminate noise within the dataset [36,38]. One variable from component 1 and component 2 was added to Stage 3 to represent the largest diversity of information provided by the category, as multiple variables from a single component are commonly highly correlated to one another and, therefore, contribute little more than noise to the results. LULC and population variables were the exception to this rule. Four LULC variables were included in Stage 3 due to the literature's consistent documentation of their value during this type of analysis. The four variables included urban (1.8% of the study site's landmass), forest (39.8%), cropland (29.9%), and barren land (0.03%). Due to the similar results of population

category, total population and population density were investigated independently in different models in Stage 3. The purpose of Stage 2 was to reduce the amount of input variables, from the original 150 zonal statistic derived variables, to the strongest variables for Stage 3. This step allowed for a more robust relationship to be built from the derived Stage 3 models.

Stage 3 also incorporated a PCA, which identified the relationship between the top variables identified in Stage 2. Three models were analyzed over each of the three years, 2012–2014. Stage 2 results identified total population as a superior variable over gender or urban vs. rural divisions; the models, therefore, incorporated variables of the total municipality population count, population density, and an empirical Bayesian estimation (EBE) of DF counts. The population model included the yearly projected population per municipality as one of the 14 independent variables, while the population density model used the municipality's projected population divided by its area (population/km²). The municipalities are not consistent in size, but it was anticipated the higher populated urban areas would increase the density value of the municipality enough to distinguish them from less developed areas, regardless of the municipality's size. Both population models used DF cases as the dependent variable.

The EBE variable was developed following the procedures of Brandel et al. [56] using a standardized event ratio for each municipality and an adjacency matrix to produce an estimate of vector presence without political boundary influence [56,57]. The adjacency matrix was created through the ArcGIS plug-in 'Adjacency for WinBUG Tool' (available through USGS [58]). Since the population of the municipality was incorporated into the EBE dependent variable, no population independent variable was required for the model, so only 13 variables were used.

Normalization of all variables was conducted for Stage 3 to reduce distribution or range bias in the results. Independent variables were normalized to Z-scores using SPSS's descriptive analysis tool. Similarly, the Poisson distribution of the dependent variable, reported DF cases, was transformed through a natural log calculation (Ln [cases + 1]). The EBE dependent variable included a normalizing step in its development, so no further manipulation was required. The Stage 3 PCA was processed by the SPSS dimension reduction tool [54], using a varimax rotation for an oblique transformation method, and included a regression variable output [59]. As with previous vulnerability studies, these component values can be added together to derive a sum factorial output for the municipality to provide a comparable value and output for mapping [34,60–65].

Comparative assessment of the Stage 3 models was conducted for the three models through a Receiver Operating Characteristic (ROC) Area under the Curve (AUC). The AUC demonstrates the positive reporting of DF cases in high sum factorial areas compared to modeled false positives in areas without reported DF cases. This can demonstrate the ability of the PCA regression output to correctly identify areas of increased risk of transmission, whether the variables are appropriate for identifying transmission risk, and provided a comparison between the results across years [22].

3. Results

The Stage 2 categorical PCA reduced the list of 150 independent variables to 14 variables which were able to more efficiently explain the variance in the dataset without extraneous variable noise. Among the more prominent variables found from Stage 3 were nighttime LST minimum cell-max zone value, elevation minimum value, vegetation min cell-max zone, and daytime LST max cell-mean zone. Four LULC variables were used in the Stage 3 assessment based on Stage 2 results and a priori knowledge from the literature, such as the relationship between the mosquito and urban populated areas, but all four variables demonstrated limited results [1,6]. A complete list of the Stage 3 independent variables can be viewed, along with their component scores, in Figure 3; negative numbers indicate an inverse relationship, and the strongest component loading was bolded for each variable. The variables which load heavier in the earlier components (example 1 or 2) demonstrate a stronger relationship to the dependent variable.

		Population vs. Ln Cases					Population Density vs. Ln Cases					EBE					
		Component					Component					Component					
		1	2	3	4	5	1	2	3	4	5	1	2	3	4		
2012	Elevation Min	-0.95	-0.09	0.07	0.06	0.00	LST Min Max	0.95	0.10	-0.04	-0.07	-0.03	LST Min Max	0.95	0.10	0.01	-0.04
	LST Min Max	0.95	0.09	-0.07	-0.04	-0.03	Elevation Min	-0.95	-0.09	0.07	0.08	0.07	Elevation Min	-0.95	-0.09	-0.01	0.07
	Vegetation Min Max	0.84	-0.05	0.18	0.20	-0.07	Vegetation Min Max	0.84	-0.07	0.17	0.20	-0.11	Vegetation Min Max	0.82	-0.06	0.24	0.21
	Daytime LST Max Mean	0.62	-0.27	-0.39	0.47	0.06	Daytime LST Max Mean	0.63	-0.25	0.45	-0.34	0.03	Daytime LST Max Mean	0.65	-0.26	-0.38	0.44
	LST Mean Range	0.00	0.84	0.18	-0.34	-0.06	Daytime LST Max Range	0.15	0.86	-0.03	-0.06	-0.11	Daytime LST Max Range	0.18	0.84	-0.09	-0.11
	Daytime LST Max Range	0.17	0.83	-0.15	-0.13	0.04	LST Mean Range	-0.02	0.83	-0.31	0.24	-0.13	LST Mean Range	-0.02	0.84	0.25	-0.31
	Vegetation Mean Min	0.09	-0.65	0.48	0.26	-0.10	Vegetation Mean Min	0.09	-0.73	0.14	0.43	-0.07	Vegetation Mean Min	0.05	-0.68	0.45	0.27
	Precipitation Min Max	-0.15	0.56	0.43	0.32	0.02	Forest Area	0.00	-0.06	0.85	-0.33	0.01	Precipitation Min Max	-0.19	0.53	0.40	0.36
	Precipitation Sum Min	0.39	-0.14	0.79	-0.09	-0.05	Crop Area	0.09	-0.22	0.78	0.25	0.08	Precipitation Sum Min	0.32	-0.17	0.79	-0.05
	Barren Area	0.17	-0.01	-0.44	0.02	0.00	Precipitation Sum Min	0.39	-0.20	-0.19	0.74	0.02	Barren Area	0.21	0.01	-0.41	0.00
	Crop Area	0.08	-0.20	0.13	0.77	-0.05	Precipitation Min Max	-0.11	0.45	0.10	0.53	0.09	Urban Area	-0.08	-0.05	-0.37	0.01
	Forest Area	-0.04	0.19	0.54	-0.74	-0.12	Barren Area	0.20	0.03	0.03	-0.42	0.16	Crop Area	0.07	-0.22	0.09	0.77
	Population	0.01	0.12	0.05	0.00	0.83	Urban Area	-0.01	0.04	0.00	-0.21	0.77	Forest Area	-0.10	0.18	0.59	-0.70
	Urban Area	-0.06	-0.09	-0.11	0.01	0.82	Population Density	-0.13	-0.14	0.07	0.11	0.74					
2013	LST Min Max	0.95	0.07	-0.09	-0.03	-0.10	LST Min Max	0.95	0.07	-0.09	-0.01	-0.09	LST Min Max	0.95	0.07	0.01	-0.08
	Elevation Min	-0.94	-0.09	-0.11	0.00	0.13	Elevation Min	-0.94	-0.09	0.11	0.03	0.13	Elevation Min	-0.95	-0.08	0.01	0.11
	Vegetation Min Max	0.80	-0.21	0.01	-0.07	0.22	Vegetation Min Max	0.80	-0.21	0.01	-0.10	0.21	Vegetation Min Max	0.79	-0.21	0.15	0.22
	Daytime LST Max Range	0.05	0.85	-0.03	0.04	0.20	Daytime LST Max Range	0.05	0.85	-0.04	-0.01	0.20	Daytime LST Max Range	0.06	0.84	-0.08	0.21
	LST Mean Range	0.06	0.82	0.36	-0.08	0.21	LST Mean Range	0.06	0.82	0.36	-0.10	0.21	LST Mean Range	0.03	0.84	0.33	0.19
	Vegetation Mean Min	0.13	-0.79	0.08	-0.11	0.11	Vegetation Mean Min	0.13	-0.80	0.08	-0.09	0.10	Vegetation Mean Min	0.11	-0.79	0.18	0.09
	Crop Area	0.08	-0.51	-0.26	-0.04	0.48	Crop Area	0.08	-0.52	-0.26	-0.04	0.47	Crop Area	0.10	-0.54	-0.16	0.49
	Forest Area	0.03	0.29	0.87	-0.10	-0.21	Forest Area	0.03	0.29	0.87	-0.09	-0.21	Forest Area	-0.07	0.36	0.83	-0.26
	Daytime LST Max Mean	0.43	-0.31	-0.75	0.04	0.00	Daytime LST Max Mean	0.43	-0.31	-0.75	0.05	0.00	Precipitation Sum Min	0.36	-0.26	0.74	0.15
	Precipitation Sum Min	0.46	-0.33	0.68	0.01	0.20	Precipitation Sum Min	0.46	-0.32	0.69	0.02	0.20	Daytime LST Max Mean	0.52	-0.37	-0.66	0.05
	Barren Area	0.16	0.08	-0.38	0.02	-0.21	Barren Area	0.16	0.08	-0.38	0.03	-0.19	Barren Area	0.21	0.05	-0.37	-0.16
	Population	0.02	0.11	0.02	0.83	0.06	Population Density	-0.04	0.03	-0.01	0.98	0.01	Urban Area	-0.11	-0.02	-0.33	-0.03
	Urban Area	-0.08	-0.03	-0.11	0.83	-0.05	Urban Area	-0.06	0.00	-0.10	0.98	-0.01	Precipitation Min Max	-0.05	0.25	-0.14	0.76
	Precipitation Min Max	-0.03	0.24	0.13	0.03	0.76	Precipitation Min Max	-0.03	0.24	0.13	0.02	0.77					
2014	LST Min Max	0.96	0.06	-0.02	-0.07	-0.02	LST Min Max	0.96	0.08	-0.01	-0.05	-0.07	LST Min Max	0.96	0.69	-0.01	-0.06
	Elevation Min	-0.95	-0.05	0.05	0.07	-0.01	Elevation Min	-0.95	-0.08	0.03	0.06	0.11	Elevation Min	-0.95	-0.07	-0.01	0.08
	Vegetation Min Max	0.83	-0.08	0.18	0.19	-0.08	Vegetation Min Max	0.84	-0.14	-0.09	0.13	0.15	Vegetation Min Max	0.83	-0.13	-0.08	0.22
	Daytime LST Max Mean	0.57	-0.29	0.47	-0.47	0.04	LST Mean Range	0.01	0.86	-0.08	0.33	0.03	Daytime LST Max Range	0.14	-0.85	0.12	-0.01
	Daytime LST Max Range	0.14	0.83	-0.20	-0.08	0.05	Daytime LST Max Range	0.13	0.86	-0.01	-0.02	0.03	LST Mean Range	0.03	0.81	0.40	0.20
	LST Mean Range	0.03	0.82	-0.36	0.22	-0.05	Vegetation Mean Min	0.15	-0.77	-0.10	0.22	0.17	Vegetation Mean Min	0.14	-0.75	-0.12	0.34
	Vegetation Mean Min	0.13	-0.67	0.32	0.34	-0.13	Population Density	-0.03	0.02	0.99	0.00	-0.01	Precipitation Min Max	-0.15	0.54	-0.23	0.38
	Precipitation Min Max	-0.15	0.62	0.35	0.17	-0.04	Urban Area	-0.05	-0.01	0.98	-0.09	0.00	Forest Area	-0.02	0.13	0.91	0.23
	Forest Area	-0.03	0.14	-0.74	0.57	-0.12	Precipitation Sum Min	0.37	-0.20	-0.01	0.80	0.06	Daytime LST Max Mean	0.56	-0.26	-0.63	-0.27
	Crop Area	0.08	-0.24	0.72	0.17	-0.02	Forest Area	-0.06	0.29	-0.11	0.76	-0.47	Crop Area	0.08	-0.34	-0.56	0.38
	Precipitation Sum Min	0.36	-0.17	-0.10	0.80	-0.04	Daytime LST Max Mean	0.59	-0.35	0.06	-0.59	0.16	Precipitation Sum Min	0.36	-0.29	0.47	0.55
	Barren Area	0.19	-0.08	-0.10	-0.50	0.00	Barren Area	0.18	0.01	0.01	-0.45	-0.25	Urban Area	-0.10	-0.07	-0.02	-0.54
	Population	0.01	0.11	0.01	0.07	0.84	Crop Area	0.12	-0.45	-0.02	-0.05	0.63	Barren Area	0.19	0.02	-0.14	-0.45
	Urban Area	-0.07	-0.07	0.02	-0.12	0.82	Precipitation Min Max	-0.12	0.48	-0.02	0.09	0.58					

Figure 3. Model output of principal component loadings. The strongest component loading was bolded for each variable.

The Stage 3 PCA descriptive results in Table 2 indicate the methods were fairly consistent between the three models. The results show both population-based models contain five components, and explain over 73% of the variance within the dataset. The EBE dependent model contained one fewer component, with four, and a lower explanation of variance compared to the other models. This may be due to the reduced number of input variables, aka population, but further analysis would be needed to confirm such a theory. The three models also had similar results when compared through the AUC, as is seen in Table 3. These results suggest there were limited statistical differences between them. The ROC AUC also demonstrated the exploratory methods used in this study were able to provide results better than statistical chance, but could be improved. The ROC AUC represent a ranking of “Poor–Fair” for the statistical method results, but are acceptable for an exploratory study [66]. With the similarity in model PCA results, similar component loadings between the three models will support the incorporation of those strongly-loading variables in future projects.

Table 2. Principal components analysis results.

		2012	2013	2014
Population	# Components	5	5	5
	Cumulative Variance	74.224	73.403	74.07
Population Density	# Components	5	5	5
	Cumulative Variance	72.283	77.347	77.966
EBE	# Components	4	4	4
	Cumulative Variance	69.622	68.674	69.599

Table 3. ROC AUC result. There was no difference between the inputs in the population and population density models for 2012–2013 when only the top two component variables were used.

	Model	Primary Results AUC	TOP 2 Component AUC
2012	Population	0.6907276	N/A
	Density	0.6856312	0.6733368
	EBE	0.6656901	0.6733368
2013	Population	0.6612869	N/A
	Density	0.6623720	0.6981112
	EBE	0.6656901	0.6981112
2014	Population	0.6612869	0.7389620
	Density	0.7122281	0.7341314
	EBE	0.6587307	0.7341314

Night LST min cell-max zone consistently ranked high across all three models. Night LST mean cell-min zone and daytime LST max cell-range zone were also highly loaded in the first or second component, supporting previous documentation on the influence of temperature range, or diurnal temperature, on mosquito survivability and their use for DF modeling [26–28,52]. Since both warmer daytime and cooler nighttime temperatures were found to be highly influential in the results, it demonstrates the variability of temperature within an area may influence DF transmission. This is particularly relevant since the range zonal statistic was determinant in both day and night temperature variables. Although daily diurnal temperatures has previously been demonstrated to influence the development of the mosquito, it should be noted these results are representative of data aggregated to an annual value which may not directly compare [27,52,67,68]. The general term ‘temperature range’ may be used hereafter to refer to the general temperature variability observed on the dependent variable. LST data documented an annual mean nighttime pixel range of -8.218 – 21.88 °C, and 8.1 – 42.73 °C during the day between both 2012 and 2014. Although large ranges are detected in this study, they may be influenced by either the weather or the diverse landscapes within a municipality, and may not provide direct support for the previous laboratory identified vector temperature limits [27,28,30]. The importance of the temperature range on DF is further supported by the consistently strong inverse loading of elevation in the first component across all three models; high elevations contain colder temperatures and have been documented to impede mosquito habitation [3,28,32]. Another influential explanatory variable was identified as vegetation min cell-max zone; which loaded positive in the first component for all models and could represent areas capable of supporting flora and fauna [1].

DF has been widely recognized as a mainly urban or peri-urban disease. The mosquito vector prefers populated spaces, due to the availability of anthropogenic water breeding sites and human food sources [1,6,17,19,22,26,52]. However, these results showed Urban LCLU as one of the lowest ranked variables in all three of the models, and across all three of the study years. This unexpected result might have been due to the low annual temporal or spatial resolutions used. Similarly, the population and population density variables ranked low across all years, not listing higher than the fourth component. Conversely, the vegetation index results were superior to those of the LULC variables. However, these results do not necessarily contradict previous work, as they still loaded within the principal components of Stage 2. The lower than expected importance of urban LCLU and population in this study suggests the need for future studies to investigate contributory urban attributes, particularly if using different spatial or temporal resolutions [1,3,17,41]. For example measures of local density, urban heat island, or water storage potential variables could be investigated rather than an individual LULC urban variable.

The SPSS regression variable output from Stage 3 was mapped by municipality to demonstrate the geographical distribution of the project’s results, following the procedures of previous vulnerability studies [34,62–64,69]. Figure 4 demonstrates the visual comparison of the population density and

EBE models to the distribution of reported DF cases per 10,000. The population and population density models were similar in component and mapped results. All three of the models follow a similar spatial pattern across all years by indicating low risk in higher elevated municipalities and increased risk in the more moderate temperature valleys, which coincide with areas more suitable for vector survival [4,18,19,28]. The maps suggest the population based models have a propensity to indicate higher risk in more densely populated areas, even if they exist above the elevation limit of the mosquito, such as Bogota. The EBE model does not indicate as high of a risk for densely populated areas, even those within the mosquito’s habitable elevation. Therefore, it is anticipated that the population models added weight to municipalities with larger urban areas, potentially due to the inclusion of the population independent variable, which the EBE model did not have. These mapped results show similar distribution of the vector count, supporting the models were able to appropriately identify the contributing independent variables, for example, the temperature range and elevation.

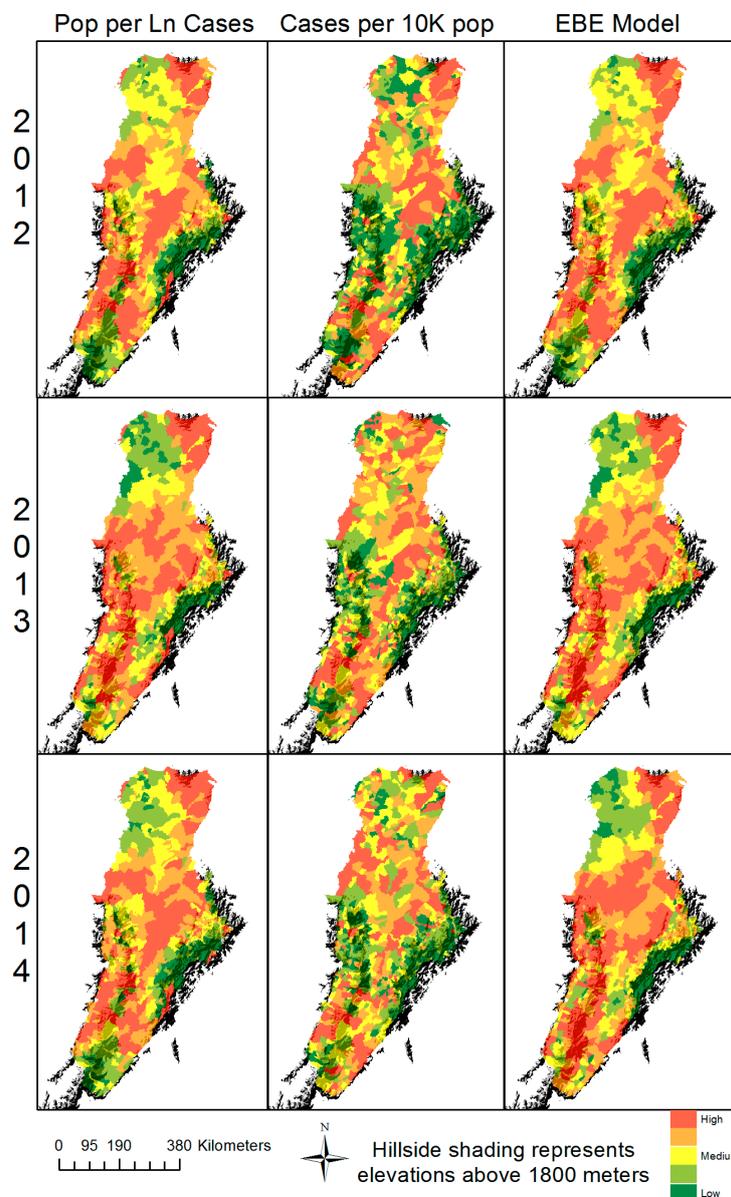


Figure 4. Population density and EBE model graphical comparisons to the reported cases per 10,000 population.

4. Discussion

Both nighttime LST min cell-max zone and daytime LST max cell-mean zone, were consistently ranked high in the top two components (Figure 3) of this study. These results are similar to previous laboratory studies on the *Aedes aegypti* mosquito which indicated temperature range can impact the vector's ability to thrive [28,52,67,68], and field studies identifying higher low temperatures improve vector resilience [3,8]. Due predominantly to the mountainous terrain, the Magdalena River watershed has a large temperature range which exceeds both the minimum and maximum temperature thresholds (which has been found to extend from 14 °C and 24 °C, respectively, depending on location or study parameters [3,28]). This provides more influence on the mosquito's ability to thrive [3,18,28], and impacts the extrinsic incubation period and duration of the gonotrophic cycle [30,41], than has been witnessed in previous studies. As noted, the temperature range variables were based off three annual assessments, so the results do not necessarily document the actual temperature limits of the *Aedes aegypti*. Additionally, since the temperature range is identified in this study as indicative of DF, it supports the theory of vector habitat re-distribution is impacted by climate change [4,6,21,70].

Although the LULC forest classification often contained an inverse relationship in Stage 3's results, the EVI variables (vegetation) had markedly better results in the top component of Stages 2 and 3 (Figure 3). It seems logical to expect a lower prevalence of DF transmission in dense forest, since there are lower population rates. However, the vegetation index will have a more inclusive classification of features (grasslands, agriculture, etc.), which may allow for more anthropogenic uses, compared to the individual dense forested space variable. The positive relationship between vegetation and dengue cases might be explained by the increased vegetation in wetter areas, which in turn are more likely to provide larval rearing sites. This is particularly relevant near populated areas where surrounding vegetation (for example, gardens, potted plants or even subsistence agriculture) may be more likely to be irrigated and to have a greater human exposure potential as compared to dense forest. Furthermore, greener areas are often associated with increased moisture and biodiversity, leading to more fauna and flora, which may also provide opportunities for suitable breeding sites [1]. These water sources are commonly located in more populated or urbanized areas, which may have lower forest LULC-classified space. The lower resolution of the forest LULC product used, compared to EVI vegetation index, could also explain the difference.

The widely accepted view that *Aedes aegypti* preferentially breeds in peridomestic water sources, such as storage tanks or flower pots [18], which are more abundant in urban areas, plus the higher density of people around these sites makes transmission of DF easier [71]. Indeed, urban spaces may even be oasis breeding sites, due to impervious surfaces collecting water and increasing the UHI, in unexpected habitat ranges (temp or elevation) [72]. The unexpected weak relationship detected here between urban LULC and DF in the Stage 3 models could be due to a low proportion of urbanized space within the municipalities, misclassification of suburban/rural communities in the LULC dataset, and/or the spatial resolution used in this study. Conversely, it may suggest future investigations as to whether higher contact with the vector is made outside, or on the fringe of, urban space, but reported in urban areas. The LULC might be identifying here primarily developed space, such as infrastructure, rather than metropolitan characteristics. Populated areas can be very green, which would increase misclassification of the urban LULC [50,73], while high vegetation index records would coordinate with the populated spaces.

Contrary to what was expected, precipitation, population, and LULC variables were less determinant within the Stage 3 results. Precipitation frequently loaded within the lowest 3–5 components across all models. This could represent a limitation of the data resolution or methodology used to measure rainfall's influence. Another potential limitation is the use of an annual aggregation of RS variables, which does not require the use of a temporal lag. Future studies incorporating higher temporal resolution should incorporate a time lag, to correspond with the gonotrophic cycle, to better identify precipitation trends. The low precipitation results could alternatively be explained by the dependence of this mosquito on peridomestic water sources for larval

rearing sites. Use of containers filled with water by humans (i.e., water storage tanks, flower pots), may allow the mosquito to be less dependent on rainfall [74], particularly during the moderate ENSO seasons experienced during the study period [55]. LULC overall loaded in the weakest 4–5 components during this project, particularly the EBE model, even though the literature indicated landscape composition to be important in modeling DF [3,6,18,19]. This may indicate the 500 m low resolution MODIS LULC product is less capable of documenting environmental discontinuity for this type of study [3,34,50,73]. This may impact the ability to identify urban LULC due to the large pixel resolutions and municipality aggregation used in this study [75,76]. It should be noted that barren LULC frequently contained stronger results than urban, which might be due to a misclassification of urban, agriculture, or periphery/transitional landscapes [50]. The yearly assessment of this project may also fail to identify seasonal transitional landscapes, land with intermittent use, which may impact the rate of transmission by bringing people to areas with a higher mosquito breeding potential [12,39]. Seasonally transitional land could include agriculture or seasonal markets. Additionally, the majority of urbanized features in Colombia are built higher in elevation, where the mosquito has not been shown to naturally thrive [3,32]. This cultural dynamic may reduce the effect of urban or similar LULC results within this study area. Future studies should continue to explore varying LULC variables, even though they were less influential during this project.

All three of the mapped models locate higher risk areas within the mountain valleys, which is supported by the knowledge that the *Aedes aegypti* does not naturally thrive in high altitudes [3,32]. Although the population variables suggest increased identification of disease transmission risk in highly populated areas. Since impervious surfaces increase local thermal conditions in a phenomenon known as the urban heat island, it should be considered that urban environments hold the potential to be oasis breeding habitats. The increased temperature and water storage capacities of urban impervious features may provide a haven for mosquito habitats in higher elevations [3,72]. Methods of transportation, particularly for goods, could bridge the gap between natural vector breeding sites in agricultural fields and municipalities outside the habitation zone. Human activities can also increase the promotion of vector breeding or host-vector contact through the establishment of gathering spaces in high vector breeding sites, such as seasonal markets, who then return infected to their urban residences where the disease is reported [39]. Additionally, the EBE had a lower explanation of variance than the other models across all years, which may be attributed to the lack of a population independent variable. This suggests a need to strongly consider the use of population in future vector modeling or zoonotic niche studies. However, it seems appropriate to mention a priori knowledge of the areas would be required to take advantage of this type of information.

A supplementary study was conducted to ascertain the strength of the variables identified from the top 2 components in the Stage 3 analysis. A Stage 3 type PCA was conducted on the 6–7 variables which weighed heavier (0.7 rounded or higher) in the top two components of the original Stage 3 (see Figure 3). The variables were predominantly from the temperature, elevation, and vegetation categories. These results were very similar to those of the original 14 variable PCA. Although the top two studies had fewer components and a larger explanation of variance, which could be explained by the fewer input variables. The AUC results also demonstrate the similarity between the results, so the component results of the top two component models were not provided. These results simply supported the previously demonstrated influence of temperature, elevation, and vegetation.

5. Conclusions

This analysis was able to identify variables which are statistically useful to estimate the likelihood of DF transmission in the Río Magdalena watershed in Colombia, as is demonstrated by Figure 3. Stage 2 and 3 results indicate temperature is an important modeling variable within this study area. The results of nighttime min cell and daytime max cell variables demonstrate that temperature range is particularly relevant, presumably on how it relates to the habitation zone for the vector and virus replication. Elevation min also loaded highly through all of the results as an inverse value, supporting

previously-identified altitude limitations to the vector. Vegetation was, similarly, an important variable to explain dengue cases, presumably due to a greater likelihood of larval rearing sites in wetter areas, which could also support increased quantities of vegetation. These results consistently suggest temperature and elevation as strongly related to the mosquito vector's ability to thrive and transmit the disease.

The PCA, AUC, and mapped results suggest the variables used in this exploratory analysis should be considered in future DF monitoring or predicting tools. Whether to adopt a population or EBE-based model would be dependent on the individual's prior knowledge and focus. The focus of population models appear to identify risk in higher populated areas, sometimes despite being located above the mosquito's expected habitable elevation. Such methodology could provide future studies with a population weight to identify spaces where there is a higher rate of transmission, rather than simply the presence of the mosquito vector.

As an exploratory study, the conditions and methods depicted within this report were able to appropriately quantify and model DF records through the use of environmental and population data. The results were able to explain approximately 70% of the dataset variability, across all of the models used. Future studies should include the variables identified in Stage 2 of this study, but consider modifying them in an attempt to further improve the model strength. Although elevation min proved to be a strong variable in this study, additional variations on elevation may prove beneficial in future studies. Future exploratory analysis could investigate whether a more computationally heavy variable could improve analysis, such as by measuring the amount of municipality area which is within the vector's habitable elevation limits or surface slope, either would better quantify the amount of potential breeding space over the simpler minimum elevation variable used in this study. Additionally, this analysis was focused on variables estimated through RS, however, it is imperative for future studies to incorporate a more direct variable of socioeconomic status. Other statistical methods, such as geographically weighted regression or boosted regression trees, should also be explored to interpret the relationship between similar independent variables and DF. Still, this project was able to appropriately and accurately accomplish its goal of identifying niche variables which identify increased *Aedes aegypti*, through DF cases, within the Magdalena Watershed. The results described by this exploratory analysis could be used to aid mitigation practices or develop an early warning system for the transmission of DF by the *Aedes aegypti* mosquito in future studies.

Acknowledgments: This study was supported by the IUPUI Chancellor for Research—Research Support Funds Grant. We thank the National Institute of Health (Instituto Nacional de Salud, INS) of Colombia for making available the public health data used in this study. The authors would like to thank Constantine Yiannoutsos, PhD for his consultation on statistical methods, Mary Lefevers for her assistance in the data preprocessing and Victor E. Casallas-Bedoya from INS for his assistance with ancillary information.

Author Contributions: Austin Stanforth participated in the research design, data collection, analysis and interpretation, and prepared the manuscript. Max J. Moreno-Madriñán conceived the project idea, participated in the research design, data collection, analysis and interpretation, and reviewed the manuscript. Jeffrey Ashby participated in the research design, data collection, analysis and interpretation, and reviewed the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gubler, D.J. Dengue and dengue hemorrhagic fever. *Clin. Microbiol. Rev.* **1998**, *11*, 480–496. [[PubMed](#)]
2. Ceccato, P.; Connor, S.J.; Jeanne, I.; Thomson, M.C. Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. *Parassitologia* **2005**, *47*, 81–96. [[PubMed](#)]
3. Moreno-Madriñán, M.J.; Crosson, W.L.; Eisen, L.; Estes, S.M.; Estes, M.G., Jr.; Hayden, M.; Hemmings, S.N.; Irwin, D.E.; Lozano-Fuentes, S.; Monaghan, A.J.; et al. Correlating remote sensing data with the abundance of pupae of the dengue virus mosquito vector, *Aedes aegypti*, in Central Mexico. *Int. J. Geo Inf.* **2014**, *3*, 732–749. [[CrossRef](#)]

4. Cleckner, H.; Allen, T.R. Dasymetric mapping and spatial modeling of mosquito vector exposure, Chesapeake, Virginia, USA. *Int. J. Geo Inf.* **2014**, *3*, 891–913. [[CrossRef](#)]
5. Li, Z.; Roux, E.; Dessay, N.; Girod, R.; Stefani, A.; Nacher, M.; Moiret, A.; Seyler, F. Mapping a knowledge-based malaria hazard index related to landscape using remote sensing: Application to the cross-border area between French Guiana and Brazil. *Remote Sens.* **2016**, *8*, 319. [[CrossRef](#)]
6. Beebe, N.W.; Cooper, R.D.; Mottram, P.; Sweeney, A.W. Australia's dengue risk driven by human adaptation to climate change. *PLoS Negl. Trop. Dis.* **2009**, *3*, e429. [[CrossRef](#)] [[PubMed](#)]
7. Liu, H.; Weng, Q.; Gaines, D. Spatio-temporal analysis of the relationship between WNV dissemination and environmental variables in Indianapolis, USA. *Int. J. Health Geogr.* **2008**, *7*, 66. [[CrossRef](#)] [[PubMed](#)]
8. Pinto, E.; Coelho, M.; Oliver, L.; Massad, E. The influence of climate variables on dengue in Singapore. *Int. J. Environ. Health Res.* **2011**, *21*, 415–426. [[CrossRef](#)] [[PubMed](#)]
9. Zhang, S.; Zhao, J. Spatio-temporal epidemiology of hand, foot and mouth disease in Liaocheng City, North China. *Exp. Ther. Med.* **2015**, *9*, 811–816. [[CrossRef](#)] [[PubMed](#)]
10. Harlan, S.L.; Brazela, A.J.; Prashada, L.; Stefanovb, W.L.; Larsenc, L. Neighborhood microclimates and vulnerability to heat stress. *Soc. Sci. Med.* **2006**, *63*, 2847–2863. [[CrossRef](#)] [[PubMed](#)]
11. NASA. MODIS. 2016; Available online: <http://www.modis.gsfc.nasa.gov/> (accessed on 4 September 2016).
12. Fuller, D.; Troyo, A.; Beier, J. El nino southern oscillation and vegetation dynamics as predictors of Dengue Fever cases in Costa Rica. *Environ. Res. Lett.* **2009**, *4*, 140111–140118. [[CrossRef](#)] [[PubMed](#)]
13. Famine Early Warning Sysems Network - FEWS NET. Available online: <http://www.fews.net> (accessed on 14 September 2016).
14. Thomson, M.C.; Conn, S.J. *Malaria Early Warning Systems—A Framework for Field Research in Africa*; World Health Organization: Geneva, Switzerland, 2001.
15. Brown, M.E.; Brickley, E.B. Evaluating the use of remote sensing data in the US agency for international development famine early warning systems network. *J. Appl. Remote Sens.* **2012**, *6*, 063511.
16. NASA Applied Sciences Projects Improve Famine Predictions Worldwide. Available online: <http://www.nasa.gov/topics/earth/features/famine.html> (accessed on 14 September 2016).
17. CDC. Dengue. 19; January; 2016; Available online: <http://www.cdc.gov/dengue/> (accessed on 2 February 2016).
18. Lozano-Fuentes, S.; Hayden, M.H.; Welsh-Rodriguez, C.; Ochoa-Martinez, C.; Tapia-Santos, B.; Kobylinski, K.C.; Uejio, C.K.; Zielinski-Gutierrez, E.; Monache, L.D.; Monaghan, A.J.; et al. The dengue virus mosquito vector *Aedes aegypti* at high elevation in Mexico. *Am. J. Trop. Med. Hyg.* **2012**, *87*, 902–909. [[CrossRef](#)] [[PubMed](#)]
19. Dengue and Severe Dengue. Available online: <http://www.who.int/mediacentre/factsheets/fs117/en/> (accessed on 2 February 2016).
20. Delmelle, E.; Dony, C.; Casas, I.; Jia, M.; Tang, W. Visualizing the impact of space-time uncertainties on Dengue Fever patterns. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1107–1127. [[CrossRef](#)]
21. Brunkard, J.M.; Cifuentes, E.; Rothenberg, S.J. Assessing the roles of temperature, precipitation, and ENSO in dengue re-emergence on the Texas-Mexico border region. *Salud Pública Méx.* **2008**, *50*, 227–234. [[CrossRef](#)] [[PubMed](#)]
22. Kraemer, M.U.; Sinka, M.E.; Duda, K.A.; Mylne, A.Q.N.; Shearer, F.M.; Barker, C.M.; Moore, C.G.; Carvalho, R.G.; Coelho, G.E.; van Bortel, W.; et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *Elife* **2015**, *4*, e08347. [[CrossRef](#)] [[PubMed](#)]
23. Mylne, A.Q.; Pigott, D.M.; Longbottom, J.; Shearer, F.; Duda, K.A.; Messina, J.P.; Weiss, D.J.; Moyes, C.L.; Golding, N.; Hay, S.I. Mapping the zoonotic niche of lassa fever in Africa. *Trans. R. Soc. Trop. Med. Hyg.* **2015**, *109*, 483–492. [[CrossRef](#)] [[PubMed](#)]
24. Pigott, D.M.; Golding, N.; Mylne, A.; Huang, Z.; Henry, A.J.; Weiss, D.J.; Brady, O.J.; Kraemer, M.U.G.; Smith, D.L.; Moyes, C.L.; et al. Mapping the zoonotic niche of Ebola virus disease in Africa. *Elife* **2014**, *3*, e04395. [[CrossRef](#)] [[PubMed](#)]

25. Pigott, D.M.; Golding, N.; Mylne, A.; Huang, Z.; Weiss, D.J.; Brady, O.J.; Kraemer, M.U.; Hay, S.I. Mapping the zoonotic niche of Marburg virus disease in Africa. *Trans. R. Soc. Trop. Med. Hyg.* **2015**, *109*, 366–378. [[CrossRef](#)] [[PubMed](#)]
26. Eastin, M.D.; Delmelle, E.; Casas, I.; Wexler, J.; Self, C. Intra-and interseasonal autoregressive prediction of dengue outbreaks using local weather and regional climate for a tropical environment in Colombia. *Am. J. Trop. Med. Hyg.* **2014**, *91*, 598–610. [[CrossRef](#)] [[PubMed](#)]
27. Lambrechts, L.; Paaijmans, K.P.; Fansiri, T.; Carrington, L.B.; Kramer, L.D.; Thomas, M.B.; Scott, T.W. Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7460–7465. [[CrossRef](#)] [[PubMed](#)]
28. Brady, O.J.; Johansson, M.A.; Guerra, C.A.; Bhatt, S.; Golding, N.; Pigott, D.M.; Delatte, H.; Grech, M.G.; Leisnham, P.T.; Maciel-de-Freitas, R.; et al. Modelling adult *Aedes aegypti* and *Aedes albopictus* survival at different temperatures in laboratory and field settings. *Parasites Vectors* **2013**, *6*, 1–12. [[CrossRef](#)] [[PubMed](#)]
29. Miron, I.J.; Montero, J.C.; Criado-Alvarez, J.J.; Díaz, J.; Linares, C. Effects of temperature extremes on daily mortality in Castile-La Mancha (Spain): Trends from 1975 to 2003. *Gac. Sanit.* **2010**, *24*, 117–122. [[PubMed](#)]
30. Rueda, L.; Patel, K.J.; Axtell, R.C.; Stinner, R.E. Temperature-dependent development and survival rates of *Culex quinquefasciatus* and *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **1990**, *27*, 892–898. [[CrossRef](#)] [[PubMed](#)]
31. Ibanez-Bernal, S. New altitudinal record of *Aedes* (*Stegomyia*) *aegypti* (Linnaeus, 1762) (Diptera: Culicidae) in Mexico. *Folia Entomol. Mex.* **1987**, *72*, 163–164.
32. Kanaroglou, P.; Delmelle, E. *Spatial Analysis in Health Geography*; Routledge: New York, NY, USA, 2015; p. 344.
33. Suárez, M.F.; Nelson, M.J. Registro de altitud del *Aedes aegypti* en Colombia. *Biomédica* **1981**, *1*, 225. [[CrossRef](#)]
34. Stanforth, A.C. *Identifying Variations of Socio-Spatial Vulnerability to Heat-Related Mortality during the 1995 Extreme Heat Event in Chicago, IL, USA*; Indiana University: Bloomington, IN, USA, 2011.
35. StatSoft. *Electronic Statistics Textbook*; StatSoft: Tulsa, OK, USA, 2015.
36. DeCoster, J. Data Analysis in SPSS. 21 February 2004. Available online: <https://msu.edu/course/psy/818/snapshot.afs/deshon/Projects/Misc%20files/spss%20overview.pdf> (accessed on 18 September 2016).
37. Kaiser, H.F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–150. [[CrossRef](#)]
38. Idre. Annotated SPSS Output—Principal Components Analysis. 2016. Available online: http://www.ats.ucla.edu/stat/spss/output/principal_components.htm (accessed on 19 September 2016).
39. Lin, C.-H.; Wen, T.-H. Using geographically weighted regression (GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence of dengue. *Int. J. Environ. Res. Public Health* **2011**, *8*, 2798–2815. [[CrossRef](#)] [[PubMed](#)]
40. UCLA, Introduction to SAS. UCLA: Statistical Consulting Group. Available online: http://www.ats.ucla.edu/stat/SPSS/output/principal_components.htm (accessed on 14 September 2016).
41. Morin, C.W.; Comrie, A.C.; Ernst, K. Climate and dengue transmission: Evidence and implications. *Environ. Health Perspect.* **2013**, *121*, 1264–1272. [[CrossRef](#)] [[PubMed](#)]
42. Venkatesan, M.; Rasgon, J.L. Population genetic data suggest a role for mosquito-mediated dispersal of west Nile virus across the Western United States. *Mol. Ecol.* **2010**, *19*, 1573–1584. [[CrossRef](#)] [[PubMed](#)]
43. DANE, Colombia. Proyecciones De Población Por Área 2005–2020. 2010; El Departamento Administrativo Nacional de Estadística; Available online: <http://www.dane.gov.co/index.php/poblacion-y-demografia/proyecciones-de-poblacion> (accessed on 10 September 2016).
44. ArcGIS, version 10.3. Esri: Redlands, CA, USA, 2014.
45. Earth Explorer. Available online: <http://earthexplorer.usgs.org> (accessed on 14 September 2016).
46. NASA. Satellite Data was obtained from NASA. Available online: <http://disc2.nascom.nasa.gov> (accessed on 1 November 2015).
47. U.S.G. Survey (USGS). *Remote Sensing Data*; U.S.G. Survey: Reston, VA, USA, 2016.
48. Python, version 2.7. Python Software Foundation: Wilmington, DE, USA, 2010.

49. Salano, R.; Didan, K.; Jacobson, A.; Huete, A. *MODIS Vegetation Index User's Guide (MOD13 Series)*; The University of Arizona: Tucson, AZ, USA, 2010; p. 42.
50. Jensen, J. *Introductory Remote Sensing—A Remote Sensing Perspective*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2005.
51. NASA. TRMM—Tropical Rainfall Measuring Mission. 18; May; 2016; Available online: <http://www.trmm.gsfc.nasa.gov/> (accessed on 23 May 2016).
52. Liu-Helmersson, J.; Stenlund, H.; Wilder-Smith, A.; Rocklöv, J. Vectorial capacity of *Aedes aegypti*: Effects of temperature and implications for global dengue epidemic potential. *PLoS ONE* **2014**, *9*, e89783. [[CrossRef](#)] [[PubMed](#)]
53. Vigilancia Rutinaria. Available online: <http://www.ins.gov.co/lineas-de-accion/Subdireccion-Vigilancia/sivigila/Paginas/vigilancia-rutinaria.aspx> (accessed on 14 September 2016).
54. *SPSS Statistical Software*, version 23. IBM Corp.: New Castle, NY, USA, 2016.
55. Null, J. El Niño and La Niña Years and Intensities. Available online: <http://www.ggweather.com/enso/oni.htm> (accessed on 4 August 2016).
56. Brandel, J. *Empirical Bayes Methods for Missing Data Analysis*; Department of Mathematics, Uppsala University: Uppsala, Sweden, 2004.
57. Woodward, M. *Epidemiology: Study Design and Data Analysis*; CRC Press: Boca Raton, FL, USA, 2013.
58. *Adjacency For WinBUGS Tool*; USGS: Menlo Park, CA, USA, 2012.
59. Costello, A.B.; Osborne, J.W. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* **2005**, *10*, 1.
60. Cutter, S.L.; Barnes, L.; Berry, M.; Burton, C.; Evans, E.; Tate, E.; Webb, J. A place-based model for understanding community resilience to natural disasters. *Glob. Environ. Chang. Hum. Policy Dimens.* **2008**, *18*, 598–606. [[CrossRef](#)]
61. Cutter, S.L.; Boruff, B.J.; Shirley, W. Social vulnerability to environmental hazards. *Soc. Sci. Q.* **2003**, *84*, 242–261. [[CrossRef](#)]
62. Cutter, S.L.; Burton, C.G.; Emrich, C.T. Disaster resilience indicators for benchmarking baseline conditions. *J. Homel. Secur. Emerg. Manag.* **2010**, *7*, 1547–7355. [[CrossRef](#)]
63. Cutter, S.L.; Finch, C. Temporal and spatial changes in social vulnerability to natural hazards. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2301–2306. [[CrossRef](#)] [[PubMed](#)]
64. Johnson, D.P.; Stanforth, A.; Lulla, V.; Lubert, G. Developing an applied extreme heat vulnerability index utilizing socioeconomic and environmental data. *Appl. Geogr.* **2012**, *35*, 23–31. [[CrossRef](#)]
65. O'Neill, M.S.; Zanobetti, A.; Schwartz, J. Disparities by race in heat-related mortality in four US cities: The role of air conditioning prevalence. *J. Urban Health* **2005**, *82*, 191–197. [[CrossRef](#)] [[PubMed](#)]
66. Tape, T.G. *Interpreting Diagnostic Tests*; University of Nebraska Medical Center: Omaha, NE, USA, 2006.
67. Carrington, L.B.; Seifert, S.N.; Willits, N.H.; Lambrecht, L.; Scott, T.W. Large diurnal temperature fluctuations negatively influence *Aedes aegypti* (Diptera: Culicidae) life-history traits. *J. Med. Entomol.* **2013**, *50*, 43–51. [[CrossRef](#)] [[PubMed](#)]
68. Mohammed, A.; Chadee, D.D. Effects of different temperature regimens on the development of *Aedes aegypti* (L.) (Diptera: Culicidae) mosquitoes. *Acta Trop.* **2011**, *119*, 38–43. [[CrossRef](#)] [[PubMed](#)]
69. Reid, C.E.; O'Neill, M.S.; Gronlund, C.J.; Brines, S.J.; Brown, D.G.; Diez-Roux, A.V.; Schwartz, J. Mapping community determinants of heat vulnerability. *Environ. Health Perspect.* **2009**, *117*, 1730–1736. [[CrossRef](#)] [[PubMed](#)]
70. Curtis, A.; Ye, X.; Heob, E.; Targhetta, J.; Salvato, V.; Reyna, M.; Bueno, R., Jr.; Holmes, L. A comparison of three approaches to identify west nile virus mosquito space-time hotspots in the houston vicinity for the period 2002–2011. *Appl. Geogr.* **2014**, *51*, 58–64. [[CrossRef](#)]
71. Lounibos, L.P. Invasions by insect vectors of human disease. *Annu. Rev. Entomol.* **2002**, *47*, 233–266. [[CrossRef](#)] [[PubMed](#)]
72. Ruiz-López, F.; González-Mazo, A.; Vélez-Mira, A.; Gómez, G.F.; Zuleta, L.; Uribe, S.; Vélez-Bernal, I.D. Presence of *Aedes (Stegomyia) aegypti* (Linnaeus, 1762) and its natural infection with dengue virus at unrecorded heights in Colombia. *Biomédica* **2016**, *36*, 303–308. [[CrossRef](#)] [[PubMed](#)]

73. Vogelmann, J.E.; Howard, S.M.; Yang, L.; Larson, C.R.; Wylie, B.K.; van Driel, J.N. Completion of the 1990's national land cover data set for the conterminous United States. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 650–662.
74. Moreno-Madriñán, M.J.; Turell, M.J. Factors of concern regarding Aedes-transmitted viruses in the U.S. *J. Med. Entomology* **2016**, in press.
75. Kwan, M.-P. The uncertain geographic context problem. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 958–968. [[CrossRef](#)]
76. Kwan, M.-P. How GIS can help address the uncertain geographic context problem in social science research. *Ann. GIS* **2012**, *18*, 245–255. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).