January 2015

Evaluation of Soil as a Risk Indicator for Human Leptospirosis in Rural, Coastal Ecuador

Chad Allen Weddell
University of South Florida, cweddell@mail.usf.edu

Evaluation of Soil as a Risk Indicator for Human Leptospirosis in Coastal, Rural Ecuador

by

Chad A. Weddell

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Public Health
Department of Global Health
College of Public Health
University of South Florida

Major Professor: Ricardo Izurieta, MD, DrPH, DTM&H
Boo Kwa, Ph.D.
Benjamin Jacob, Ph.D.
Eknath Naik, MD, Ph.D.

Date of Approval:
July 8, 2015

Keywords: tropical diseases, environment, remote sensing, geographic approach

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Leptospirosis, a zoonotic disease caused by pathogenic spirochete bacteria (family *Leptospiraceae*, genus *Leptospira*), is endemic in developing tropical regions of the world. It occurs in epidemics and is endemic in Ecuador where environmental conditions are ideal for maintenance. The role of soil as a long term reservoir has been previously been documented. Geographic Information System (GIS) and Remote Sensing (RS) technology was used in our study to further explore the role of soil as an environmental reservoir and its potential use as a static risk indicator for disease. Red, Green, Blue (RGB) spectral band data from known leptospire positive soil sites were extracted from high resolution satellite images and used to construct the first ever remotely dependent soil-based model. The soil co-variates failed to demonstrate statistical significance; however, elevation was found to be statistically significant. The soil type most associated with soil samples where leptospire DNA was detected using real-time PCR analysis was cambisol, a soil type with a common distribution in Ecuador and Africa. This exploratory analysis presented a novel idea of combining environmental microbiological sampling and GIS/RS technology to better examine static risk indicators such as soil. Further analysis is warranted based on spatial relationships noted.

CHAPTER 1: INTRODUCTION

The causative agents of human leptospirosis are spirochete leptospire bacteria (family *Leptospiraceae*, genus *Leptospira*). Over 200 pathogenic serovars, the basic unit of classification based on serology, have been identified [1]. DNA-DNA hybridization methods classify 20 species of leptospires including six saprophytic (*L. biflexa*), five intermediate, and nine pathogenic species (*L. interrogans*) [2]. Saprophytic leptospires are benign environmental bacteria which subsist on nutrients in surface water and mud [3]. Leptospires of intermediate pathogenicity are opportunistic organisms that have been isolated from humans and animals [2].

Pathogenic leptospires maintain a complex lifecycle between animals and the environment where humans are accidental hosts. The bacteria are maintained in asymptomatic chronic renal carrier animal hosts [4]. Most mammalians are susceptible to infection; however, rats and mice are considered the primary animal host. They maintain lifelong asymptomatic infections and excrete large amounts of infective leptospires into the environment [5]. Livestock, swine, and dogs have been linked with disease transmission which has been attributed to their close proximity to humans. Pathogenic leptospires have been detected in a wide variety of wildlife including bats in the Peruvian rainforest. This reflects the ubiquitous distribution of disease-causing leptospires in domestic and wild animal populations [6,7].

Water and soil environments play a key role in disease transmission. They serve as long term reservoirs for pathogenic leptospires and a continual source of infection. Pathogenic leptospires are associated with stagnant bodies of water and muddy rivers which provide optimal environments for maintenance and viability in the absence of a host [3].

Pathogenic leptospires can survive in soil for long periods of time outside a host. Survival times in soil have been reported from many weeks to months. Leptospires remain viable in acidic soils with a pH of 5.5 for 42 days and in more neutral soil for up to 74 days [8]. The organism can form biofilms in both biotic and abiotic materials (glass, plastic) as an adaptive survival strategy [9]. Additional factors that support environmental maintenance include an alkaline pH, low levels of heterotrophic bacteria, high oxygen and low salt concentration [3]. Moisture is a universal requirement for leptospires. The bacterium is killed by dehydration or temperatures greater than 50 ^{o}C [10].

High humidity, heavy rainfall, and warm temperatures have been identified as environmental drivers of leptospirosis [11]. A link has long been established between heavy rainfall and epidemics [12]. Epidemics have been linked to periods of heavy rainfall and flooding events throughout the world including Argentina, Bangladesh, Brazil, Costa Rica, Cuba, Ecuador, India, Korea, Malaysia, Mexico, Nicaragua, Philippines, Portugal, Puerto Rico, Russia, and the United States [13]. Climate change has been cited as an emerging future driver for disease. Links between global warning and increases in the frequency and intensity of natural disasters and flooding events associated with climate change continue to be made. Global warming due to climate change could potentially increase disease burden by causing carrier animals to migrate into higher elevations or helping to increase survival time of leptospires in the environment [14].

Humans contract leptospirosis when the organism enters the bloodstream through any compromised skin or mucous membranes [15]. This can occur through direct contact with infective animal products or waste (blood, urine, pelts, wool, meat, reproductive materials, or milk). Leptospirosis can be related to occupations (slaughterhouse workers, farmers), food and water consumption, and recreational habits (hunters/trappers). Ingestion of contaminated food

and water and inhalation of contaminated aerosol droplets following the spraying of dairy cow udders during milking have been documented [4].

Leptospirosis is also frequently contracted indirectly through contact with water, soil, or vegetation contaminated with animal urine containing leptospires [15,16]. A widely accepted water-borne transmission model elucidates how indirect leptospirosis transmission likely occurs through environmental exposure. This model asserts that rain water carries pathogenic leptospires in soil contaminated with the urine of infected animals to surface water where new hosts are infected. Specifically, rat nests (habitats) in the ground are posited to be flooded during heavy rains, and pathogenic leptospires may be driven to the surface where humans come into contact with them. This model may not be complete, however, since many cases actually do not occur until months after the onset of heavy rain. Furthermore, cases do not always correlate with years of heavier than normal rainfall [3,17].

Diseases symptoms vary greatly from mild, unapparent to severe, fatal disease. The disease presents as a fever of unknown origin with muscle pains and non-specific complaints. Between 20-40% of the cases involving a fever of unknown origin may be due to leptospirosis [18]. Non-specific signs present similarly in a host of other diseases including dengue fever, malaria, and influenza often leading to a delayed or missed diagnosis [15]. Co-infections with dengue fever occur which further complicates diagnosis and treatment efforts [19,20]). Clinical signs seen in the most severe manifestation of the disease, known as Weil's disease, are related to renal dysfunction, hepatic dysfunction, pneumonia, hemorrhagic disease, and inflammation to the central nervous system [4]. Average reported case fatality rates vary from 5-20% [21]. Vaccination as a preventive method is very limited in both humans and animals. A human vaccine, Vax-Spiral ®, was produced and used in Cuba and demonstrated immunogenicity against all included serovars with a 78% efficacy [22].

Application of highly diluted potentised bacteria in oral vaccines prior to epidemics has been evaluated as a potential preventive measure in Cuba [23]. To date, however, acceptable primary prevention measures are avoidance of stagnant bodies of waters and flood waters, wearing protective gear when engaging in recreational water activities or in areas where contact could occur with animal excreta, and covering open wounds and abrasions when around potential sources of contamination [24]. Secondary prevention measures include early detection and initiation of antibiotic therapy, both of which improve health outcomes.

Direct and indirect methods are used to detect leptospires in various types of human, anima, and environmental samples. Direct detection methods involve antigen detection through direct visualization with microscopy, DNA detection, or culture methods. Direct visualization of leptospires in blood, urine, and rarely cerebrospinal fluid has been conducted using dark-field microscopy. This technique, however, is not sensitive or specific with approximately 10^4 leptospires per ml required for one cell per field for visualization [25]. Culture is a confirmatory method for diagnosis but has limited utility in clinical settings. It requires special, expensive media. Additionally, leptospires are fastidious organisms that often take months to grow [26]. The use of culture is largely limited to reference laboratories for these reasons.

The importance of molecular methods for rapid detection of leptospires is increasing as technologies improve. Polymerase Chain Reaction (PCR) methods including conventional and real-time PCR methods have been described in the literature for detection of leptospires in human, animal and environmental samples [25,27]. Typical processing time is between 2-5 hours. In general terms, conventional PCR techniques involve repeated cycling of DNA denaturing, primer annealing, DNA duplication by a DNA polymerase enzyme, and electrophoresis in agarose gel [28]. This process allows the amplification of small numbers of leptospires that may be present early on in the disease process and increases the ability to obtain definitive diagnosis before antibodies are detectable.

Real-time PCR assays are faster and less prone to contamination issues. Limitations of PCR technologies include expense and requirements for specialized equipment such as thermocyclers [25]. Additionally, samples must be taken prior to antibiotic therapy since antibiotics rapidly clear the organism from blood and urine [29].

Isothermal methods for DNA amplification such as Loop-mediated Isothermal Amplification (LAMP) have been studied for use in various infectious diseases. The use of loop primers have demonstrated ability of amplify small amounts of leptospiral DNA to 10^9 copies within 60-90 minutes [30]. Reported equipment requirements include a heat block or water bath maintained at a single temperature which makes this technology potentially useful in the field or clinical setting as well as for use in low resource countries.

The gold standard for definitive diagnosis is serologic testing for antibodies via the Microscopic Agglutination Test (MAT). Paired serum samples taken at least two weeks apart are tested for the presence antibodies with a fourfold or greater increase in titer levels is considered confirmatory [12]. Widespread use in low resource settings is limited due to cost of testing and specialized laboratory requirements. A reference library of geographically appropriate live leptospire strains is required for testing. Additionally, the MAT shows limited sensitivity in early stages of illness (29.0 -48.7%). The interpretation of test results involves visual inspection for agglutination. The subjective nature of this interpretation has been cited as a diagnostic limitation of the test [12, 31-33].

Estimates of global disease burden range between 500,000 to 1,000,000 cases per year occur; however, true disease burden is unknown due to limitations in diagnostics, inconsistent surveillance, and a low index of suspicion among healthcare workers [18]. The World Health Organization (WHO) prioritizes the disease as the most underreported neglected tropical zoonosis worldwide [12,33]. Leptospirosis occurs with greatest frequency in tropical and sub-tropical regions where year round transmission and maintenance is supported by moist

environments. The estimated incidence of human leptospirosis is 0.1 to 1 per 100,000 population in temperate regions compared with an incidence of 10 or more per 100,000 population in tropical regions [34]. The highest median annual incidences per 100 000 population are as follows: Africa (95.5), Western Pacific (66.4), the Americas (12.5), SE Asia (4.8), and Europe (.05) [12].

In Ecuador, human leptospirosis occurs as an endemic disease and in epidemics following heavy rainfall and flooding. Between 1994 and 2012, 3,606 cases were detected and reported throughout all provinces through the national surveillance operated by Ecuador's Ministry of Health Epidemiological Surveillance System (Figure 1) [35]. A significant spike occurred in 1998 where reported cases increased from 3 in 1997 to 398 in 1998. Heavy flooding occurring in 1997-1998 due to the effects of El Nino was likely a major contributing factor. Similar flooding events associated with the El Nino warm water current typically occur every three to four years in Ecuador making the country prone to leptospirosis epidemics. El Nino has been associated with increases in frequency and intensity in coastal regions in Ecuador as well increases in landslides and storm surges [36]. In 2012, 95.7% (1224/1279) of human leptospirosis cases were from the coastal region. The greatest incidence in 2012 occurred in the coastal province of Manabí 63.71% (975/1279) (Figure 2). The major burden of leptospirosis in Ecuador occurs in Manabí province in rural parishes. Case prevalence (seroprevalence) and incidence data from Ecuador's Ministry of Health Epidemiological Surveillance provided by Veronica Barragan from Northern Arizona University demonstrates that the majority of leptospirosis in Manabí occurs in the rural parishes of Portoviejo and Santa Ana (Table 1 and 2). Figure 5 depicts a 2011 leptospirosis incidence map per 100 000 population by province in Continental Ecuador.

GIS was developed in the 1960s by geographers and mathematicians for the purpose of mapping and management of large amounts of geographically referenced data such as national

census data [37]. Advances in information technology and increased availability of spatial data have contributed to increased use of GIS in both public health practice and research [38]. In public health practice GIS is increasingly being used as an evidence-based decision support tool for guiding policy, resource allocation, and program management. The technology has also been incorporated into disease surveillance and control programs at local, national, and international levels. For example, the World Health Organization (WHO) initiated an ongoing Public Health Mapping Program and GIS program in 1993 designed to strengthen surveillance, prevention and control programs and improve management and knowledge of priority diseases [38].

Two GIS-enabling technologies are Global Position System (GPS) and Remote Sensing (RS). GPS allows manual geo-referencing of important data points that can later be uploaded into the GIS software platform and mapped [39]. RS refers to use of sensor technology on satellites to collect images of the Earths' surface which contains geographical, environmental, and meteorological data. Examples of satellites include Landsat, World-View2, SPOT, GeoEye-1, QuickBird, and IKONOS [39]. Images collected vary in resolution, and higher resolution images are significantly more expensive to purchase. Currently, GeoEye-1 images are the highest commercially available images with a ground resolution of .41 meters [40]. This level of resolution is referred to as sub-meter resolution. Use of RS permits identification of site-specific predictor variables related to geography or calculation of environmental proxy variables from the information captured on the images. It is widely co-applied with GIS in public health studies.

Advances in GIS and RS technologies now offer new public health research tools to explore temporal and spatial relationships in diseases with complex transmission lifecycles like leptospirosis. It allows spatial and temporal analysis among multiple inputs from human, animal, and environmental data that can be applied in assessing risk for disease. Disease risk models created through use of GIS analysis can serve as powerful planning tools for disease

prevention programs in evidence-based public health practice. They can enable more targeted allocation of limited resources for control and prevention, especially important for neglected diseases in resource poor settings.

Problem Statement

In light of its complex ecology, lack of geo-location of representative sites, and lack of availability of mapping data in the literature, identification of a static indicator for creation of a robust infectious disease leptospirosis predictive model is needed.

Study Purpose

The purpose of our study was to explore the potential utility of soil type as a static indicator of risk for leptospirosis using Geographic Information System (GIS) and Remote Sensing (RS) technologies. An additional aim of the study was to analyze potential environmental drivers of leptospirosis in rural, coastal Ecuador within a GIS environment.

Definitions and Terms

Risk models: Quantitative estimations of the probability of specified adverse events from defined hazards [41].

Geographic Information System (GIS): people, data, computer software, hardware, and procedures and methods [42].

Remote Sensing Model: Model based on the use of sensor technology on satellites to collect images of the Earths' surface which contains geographical, environmental, and meteorological data [39].

Spectral Signature: Spectral signatures are radiation signals collected at different spectral bands and are the basis for land surface classification and evaluation of geophysical properties. The spectral signature represents the connected points of solar radiation emitted from a material plotted over a range of wavelengths. The spectral signature can be measured with a

task spectrometer or separation of red, green, blue, and Near Infrared portion of the electromagnetic spectrum in remotely sensed images [43].

Endmember spectra: Pixel spectra that lie at the vertices of the image simplex in n-dimensional space [44].

Research Objectives

The primary objective of our study was to determine if a soil-based remote sensing model could be constructed within a GIS environment from known infected sites for pathogenic or opportunistically pathogenic leptospires in Ecuador that can accurately forecast unknown, un-sampled positive infective sites based on soil type. Our additional research aim was to evaluate potential statistically significant drivers for leptospirosis in rural, coastal Ecuador.

Delimitations

In this study, individual risk factors were not evaluated due to the lack of individual case data and patient confidentiality concerns. The literature review for the study is highly focused on research pertaining to soil, GIS-based studies relative to risk mapping and remote sensing disease risk models, and leptospirosis research conducted in the Americas and Ecuador. Chemical analysis of soils, testing on humans and animals was not performed due to cost.

Limitations

Study limitations are inherent to an ecological study design and ecological fallacy. Population level findings were not able to be applied to individual. The collection of samples was limited temporally to one season due to time constraints of the student's program.

CHAPTER 2: LITERATURE REVIEW

The most comprehensive analysis of leptospirosis incidence in the Americas to date was research published by authors Costa et al. in 2012 [45]. In this original research, the authors' main objective was to evaluate leptospirosis reporting practices throughout the Americas. To achieve this, the authors queried several sources to include the Ministry of Health official websites, international organizations, personal communications, and international morbidity databases. The following three countries represented greater than 80% of the cases reported from the period of 1996-2005: Brazil, Costa Rica, and Cuba. The median annual cases reported by Ministry of Health official websites in these three countries were 3,165.5, 196, and 558.5 respectively. The median reported number of annual human leptospirosis cases in Ecuador was reported at 61 cases per 100 000 [45].

Of the ten South American countries and territories reporting data, three (Argentina, Brazil, and Ecuador reported leptospirosis cases every year from 1996-2005. Case fatalities were only reported for three countries which may highlight an inherent problem with the methodology of this study [45]. This study highlighted some overall weaknesses in leptospirosis surveillance and reporting systems in the Americas. The authors discussed the fact that there is a lack of standardization, regionally and internationally, in reporting requirements, and there is a disparity in information availability for the years 1996-2005 reported in this study. Also, the authors point to the fact that there may be variability in case definitions and laboratory confirmation methodology which may contribute to erroneous numbers of cases reported [45]. The study highlights the fact that the true incidence is most likely grossly under reported due to the variability in surveillance and reporting systems and lack of sensitive and specific tests.

Several studies examine the association of leptospires with specific soil types. Early research of the geographical distribution of pathogenic leptospires conducted by Kingscote (1970), demonstrated a correlation of leptospiral habitats with bedrock type [46]. In this study, Paleozoic bedrock composed of limestone and dolomite was indicated as a reliable marker for enzootic disease with colloidal clay common to positive habitats [46]. The link between specific soil types and increased risk for disease is further supported by an ecological study conducted by Schneider et al (2012) where potential drivers for leptospirosis transmission in Nicaragua were evaluated [47]. In this study increased risk of disease was associated with the following two types of soil combination: cambisol over pyroclastic larval bedrock and andosol over volcanic ash [47]. Andosols are young soils formed in volcanic ash and contain large amounts of glass and colloidal materials commonly found in leptospirosis-prone regions [48,49].

Environmental data, geographic information systems (GIS), spatial statistical analysis, and predictive risk maps have been used for the investigation and management of a range of infectious diseases including malaria, onchocerciasis, West Nile Virus (WNV), Eastern Equine Encephalitis (EEE), schistosomiasis, and Lyme disease [50-55]. GIS-based studies in this area are generally focused on analysis of vector ecology, exploration of environmental drivers in disease transmission, and the development of risk models and early warning systems. These studies identify geographic areas with high disease prevalence and/or risk of outbreaks, and are useful for guiding allocation of scarce public health resources. GIS studies in vector-borne diseases highlight areas where the technology has provided greater insight into transmission patterns and environmental drivers. Our study is the first to use GIS and RS technology to evaluate the feasibility of a soil-based model for human leptospirosis. Remote Sensing (RS) technology has proven highly useful tool for identifying high risk areas to improve allocation of public health resources for many vector-borne diseases including malaria, Eastern Equine Encephalitis, and West Nile Virus. Jacob et al. (2005) in Kenya, Africa examined the effect of

changes in land cover due to rice husbandry practices over time on abundance of anopheline larval habitats [51]. Spatial datasets were constructed from entomological larval sampling data, demographic, hydrological, and agricultural data in order to evaluate potential associations. Findings illuminated a correlation between larval habitat and land use changes and further identified the areas of higher land use changes which could have implications for targeted pest control applications in resource poor settings.

A study was conducted in 1991 to evaluate the distribution of the deer tick which transmits Lyme disease in Northwest Illinois [50]. Geo-referenced data retrieved from tick infested deer were added into the state database which contained additional environmental, geographical, and biological data. Tick distribution on deer was found to be clustered on foci near rivers. The presence of wooded vegetation, sandy soil, and river location were detected as possible risk factors [50]. The study offered insight into the dispersal pattern of the vector and authors planned further studies to continue to document vector dispersal patterns over time. Studying the geographical distribution of vectors and combining this data with human health-related data was referenced as a future initiative in this study. This type of surveillance is highly relevant over twenty years later in light of re-emerging diseases.

A GIS-based study of the ecology of triatomines, the vector for Chagas disease was conducted in a Brazilian village in order to assess risk of transmission [56]. Triatomines were sampled at both inhabited and uninhabited dwellings. Additionally, samples were taken from wood heaps in the village. A kernel density estimation map was generated in a Geographic Information System, and subsequent analysis identified the wood heaps as the greatest source of risk for transmission. Additional local risk areas noted were identified in uninhabited dwellings where animals were kept. From this study, environmental risk factors related to local vector and human behavior were identified. Human practices such as storage of animals in wood buildings were identified as contributing factors to the spread of disease [56].

Similarly another study design utilized remotely sensed data in combination with intensive field sampling of anopheles mosquito habitat to study human malaria and rice field anopheline population dynamics [57]. The study presented findings from a 1987 background study conducted by NASA in Sacramento California in 1985 designed to identify environmental parameters influencing mosquito production using remotely-sensed data. The original 1985 study found that rice fields with high larval populations were associated with fields with early developing vegetation canopies. The 1987 study expanded the original study and focused on 104 irrigated rice fields in the same area in California. Spectral values, by channel, were extracted from LANDSAT remote satellite images, and a digital land use coverage was created in ARC GIS software. Land use coverage maps were used to located sampled fields and calculate distances between each rice field and nearest livestock pastures. High producing fields were found to be clustered in areas adjacent to livestock pastures. The study demonstrated practical applications of RS technology to improve mosquito abatement efforts [57].

The focus of a research article published in 2012 on malaria in the Amhara region of Ethiopia was on the development of a computerized early warning system capable of forecasting epidemic risk [52]. The methodology centered on the use of RS to extract environmental, climate, and land surface variables to use in modelling risk. Disease specific data were incorporated in a time-series analysis to calculate risk indices. The reported predictability of the model through a cross-validation study was 50% [52]. Several limitations to the study were discussed.

Data related to human behaviors and other potential environmental drivers were not included in the model. The use of district level data limited the detection of finer-scale heterogeneity. Limited availability of malaria surveillance data was also mentioned. The system demonstrated promise for monitoring and surveillance activities but was not fully developed as a predictive tool.

Similar focus on developing predictive models is found in GIS-based dengue fever and West Nile Virus research. A study on dengue fever in the Philippines outlined a prediction method for detecting dengue incidence levels four weeks in advance. Retrospective validation was performed with a reported prediction accuracy of 78% [53]. Authors caution that any predictive model developed will be highly dependent on accuracy of data [53]. This is of concern in many parts of the world where surveillance and reporting methods lack standardization and underreporting is commonplace.

A study performed in the United States on West Nile Virus evaluates a national-level predictive model generated using GIS and RS technology [58]. The model was found retrospectively to predict national-level WNV incidence with a resulting correlation coefficient of .86 [58]. The authors, however, noted potential incongruities with regional model predictability values much lower than national predictability which compromised overall model validity. Limitations of small numbers of human cases were also discussed. A WNV study utilizing GIS and RS conducted in Iowa reported a significant novel finding of an association between rural agricultural settings and human WNV disease [54].

A study was conducted on the encephalitic viral disease Equine Eastern Encephalitis (EEE), a disease which maintains a lifecycle between mosquito vectors and birds with humans and horses as dead end hosts [55]. The study conducted in Tuskegee, Alabama in 2005 utilized GIS and RS with field and remote sampled mosquito and bird data to explore spatial relationships between vector and host [55]. Through use of various functions in GIS,

regression and spatial liner models were generated. Significant environmental and terrain variables were identified. The findings demonstrated the utility of GIS and RS to better understand spatial relationships between vector and host with the potential for future use in surveillance programs and prediction of EEE transmission risk [55]. Additionally, a highly sensitive and specific remote sensing model was developed and validated to predict *Simulium damnosum S.l.* breeding sites in West Africa which is the black fly vector for onchocerciasis, a highly debilitating disease and major cause of blindness worldwide. The model utilized a spectral signature of known positive sites associated with fast flowing water passing over pre-cambrian rock. The spectral signature for bedrock was highly predictive for breeding sites of the black fly vector of onchocerciasis in Sub-Saharan Africa [59]. The model was generated through use of remotely sensed QuickBird images at .61 meter resolution containing geo-referenced validated vector aquatic habitats and the Li-Strahler model.

Class estimates were obtained using ENVI, a commercial GIS software product that utilizes an object-based classification algorithm. Non-parametric estimators from the endmember spectra and the geometric-optical model were then used to construct a Boolean model that generated a robust spectral signature reference in an ArcGIS database specific for the verified *Simulium damnosum s.l.* habitats. These unique signatures were then used to predict larval habitats along rivers in unvisited, untested sites in Togo and Uganda [59].

In a study conducted by Jacob et al. the following steps are described for developing a robust model for forecasting canopied *Simulium damnosum s.l.* larval habitats in Burkina Faso: habitat mapping, generation of remote sensing models, object-based image classification, successive progressive algorithm, 3D radiative transfer equation/Li-Strahler Geometric Optical Model, and Validation [60]. These steps were outlined for using an unbiased stochastic spectral endmember interpolator of a potential vector/disease agent habitat which could be useful for soil-based diseases if a static indicator is discovered.

Image classification can be achieved by the following three methods: pixel-wise image classification, sub-pixel wise classification, and object-based image classification. Pixel-wise image classification methods work under the assumption that each pixel is pure and is labelled as a single land use land cover type. Given the previous discussion of heterogeneity in images, it is not ideal for classification for lower resolution images. Using sub-pixel wise classification a more accurate estimation of the proportion of each land cover type can be attained. Fuzzy classification, artificial neural networks, regression modeling, spectral mixture analysis, and multiple endmember spectral analysis are examples of sub-pixel wise classification [67]. Higher accuracy has been demonstrated when object-based classification is utilized. Geographical objects are the basic unit of analysis in object-based models.

Numerous endmember extraction techniques and models are described in the literature. A 3D radiative transfer model and L-Strahler geometric optical model were used to extract endmembers in the study conducted by Jacob et al. [60]. The 3D radiative transfer model calculated adsorption of seasonally oriented canopy radiance measurement and infused Eddington approximations. Eddington approximation is a special case of the two stream approximation applied in spectral unmixing approximations with the underlying assumption that intensity is a linear function where $\mu=\cos 0$ [60].

The Li-Strahler geometric optical model which is an invertible model commonly used in remote sensing. It allows calculation of spectral properties given surface conditions and extraction of surface structures from remotely sensed signals over vegetation canopies [68]. The model is based on the assumption that Bi-directional Reflectance Distribution Function (BRDF) is a purely geometric phenomenon resulting from a scene of discrete 3-dimensional objects being illuminated and viewed from different positions in the hemisphere. The reflectance of a single pixel can be modeled by linear combination of the following three components: sunlit canopy (C), sunlit background (G) and shadow (T) [69].

Many spectral unmixing algorithms are described in literature. When using remotely sensed images for analysis the issue of spectral mixing must be addressed. Spectral mixing refers to the fact that within any remotely sensed image few image pixel spectra are homogenous [44]. Image endmembers, commonly referred to as endmembers, are pixel spectra that lie at the vertices of the image simplex in n-dimensional space [44]. Endmember variability related to spectral mixing can severely affect the accuracy of subpixel land cover fractions [70]. Spectral Mixture Analysis (SMA) is used to account for mixture of pixels in remotely sensed images. The two categories that broadly define SMA techniques are those that reduce within class variation or enhance between class variation and those that test all endmember combinations that are potentially feasible and selection of the best fit model [70]. Techniques in the former category include automated short-wave form infrared (SWIR) unmixing, stable zone unmixing (SZU), and various spectral weighting techniques. Methods described for the latter category include Multiple Endmember Spectral Mixture Analysis (MESMA) and endmember bundling [70].

The broad steps in conducting MESMA include development of a spectral library followed by image unmixing using every possible combination of two, three, and four endmembers applied to each pixel [60]. In a study conducted by Myint and Okin, a MESMA technique was applied to characterize urban land cover in a growing desert environment in Arizona [71]. The technique was employed over conventional SMA because it allows endmembers to vary on a per pixel basis, and SMA is unable to account for non-linear mixing [71]. The study utilized a combination of three land cover types including impervious surfaces, soil, and vegetation. Findings suggest that MESMA is highly effective in mapping urban land covers. Some signature confusion was noted between bright soils and impervious surfaces. Also confusion was noted between dry/exposed soil and bright impervious surfaces. To address the latter it was recommended to separate urban and non-urban land covers using

different models and different endmembers separately. Using two endmembers for one land type can be employed to improve model performance [71]. General limitations cited for MESMA include the requirement use of hundreds of endmembers which can complicate interpretation and computation [60].

Conventional mathematical algorithms for spectral image classification do not quantify the dependence between a pixel and its neighbors known as spatial autocorrelation [66]. A robust Successive Progressive Algorithm used within a Boolean {false, true}domain is described by Jacob et al. to spectrally separate with-in class feature attributes in n-dimensional space and quantitate sub-pixel endmember heterogeneity [60]. Decomposed endmembers can be used to generate a robust spectral signature in a GIS environment that can then be kriged in order to identify unknown, unsampled productive habitat sites in a blind study fashion [60]. Ordinary kriging is a geostatistical method which utilizes spatial correlation of data to determine weighting values and models correlation between data points to determine the estimate value at an unsampled point [66].

The final step in the described process for generating a robust spectral signature using endmember spectra is the validation process. In the referenced study, this was accomplished through a process referred to as ground truthing where field verification acquired through sampling of all forecasted habitat sites is used to generate a validation model. The validation model generated by Jacob et al. demonstrated a 100% correlation among predictive geo-referenced productive vector habitats based on seasonal-sampled larval density count values [60]. A robust remote sensing model known as the Black Rock-Rapid (BRR) model was similarly developed to predict larval habitats demonstrated a sensitivity of 80% and a specificity of 92% for predicting productive larval habitats [59].

CHAPTER 3: METHODS

The Republic of Ecuador is located in Western South America with a western border of the Pacific Ocean at the Equator with Colombia to the north and Peru to the south. It represents a total area of land and water of 283, 561 kilometers and includes the Galapagos Islands. Four distinct regions are recognized: Costa, Sierra, Oriente, Galapagos Islands [72]. Figure 3 is a population map representing the provinces of continental Ecuador, which does not include the Galapagos Islands. The islands were not evaluated as a part of this study. The climate is highly variable across the regions ranging from tropical along the coast and in the Amazonian jungle lowlands to more temperate conditions inland and at higher elevations. The terrain is also variable from the coastal plain, inter-Andean central highlands (sierra), and the flat and rolling eastern jungle terrain. Ecuador is prone to frequent earthquakes, landslides, volcanic activity, periodic droughts, and floods [72]. Furthermore, the largest coastal and mountain cities are in earthquake-prone areas where populations are particularly vulnerable to leptospirosis outbreaks following natural disasters [36].

Sites targeted for testing were adjacent to areas where productive water sampling is being conducted in riverine bodies in eastern, rural communities in Manabí, Ecuador. Manabí province is a coastal province bordered to the north by the province of Esmeraldas, to the south by Guayas and Santa Elena, and to the east by Guayas, Los Rios, Santo Domingo de los Tsachilas, and west by the Pacific Ocean. Hydrology and elevation are significant factors to flooding and epidemics in this region. The climate is subtropical dry to moist tropical. The winter season is from early December to the end of May which warms due to El Nino. Summer season extends from June to December and is mildly influenced by the cold Humboldt Current.

The temperature is not uniform across the province, the average temperature in Portoviejo, thecapital, is 25 degrees Celsius and in the port city, Manta where average temperatures are 23 degrees Celsius. The province is marked by rapidly urbanizing and rural areas where access to basic services is limited [73]. Figure 4 depicts a 2010 population map of the cantons of Manabí province. We conducted our sampling in the cantons of Portoviejo and Santa Ana. Portoviejo has a greater population which is highly concentrated in urban areas; however, our sample sites were outside the major urbanized areas in less populated locations in both Portoviejo and Santa Ana.

In the first step of our methodology we identified known positive sites for disease-causing leptospires (pathogen mapping) in rural parishes in Ecuador where leptospirosis is endemic. Our study was conducted with conjunction with an ongoing leptospirosis collaborative study between the Northern Arizona University (NAU) and the University of San Francisco-Quito (USF-Q). In their year-long study, water samples from river sites in the rural parishes of Calderon and Santa Ana were being tested for the presence of pathogenic leptospires in addition to human and animal samples. The overall goal of the study was to link species between humans, animals, and the environment to better understand the risk factors associated with disease transmission of leptospirosis in Ecuador. Investigators developed real-time quantitative polymerase chain reaction assays with specificity to discriminate among leptospiral species.

We coordinated our soil sampling with their ongoing scheduled water sampling and collected 64 field soil samples adjacent the river sites where water was being tested. We collected the samples in two iterations in July and August three weeks apart. Sampling dates were July 24-25, 2014 and August 13-14, 2014. We collected samples in the following parishes: Portoviejo, Abdon Calderon, Rico Chico, Santa Ana De Vuelta Larga (Santa Ana), La Union, and San Pueblo. Active human cases were present in La Union at the time of sampling

(August 14, 2014). Soil sites positive for leptospires in the communities sampled are depicted in Figure 6 and 7. All samples were collected as aseptically as possible in a field setting using separate collection devices (plastic spoons) and double sealed in plastic bags to minimize risk of contamination. Samples were collected within five meters of the river edge no more than 3 cm in depth per recommended guidance in previous studies [75]. Samples were kept in coolers while being transported back to the laboratory of Dr. Gabriel Trueba at the University of San Francisco Quito (USF-Q) where we processed them. DNA extraction using a commercial DNA extraction kit was performed on each sample. Extraction was conducted in accordance with the manufacturer's protocol, and no modifications were necessary.

Liquid DNA extracted samples were frozen and shipped by FedEx® to collaborators Veronica Barragan and Dr. Talima Pearson at Northern Arizona University (NAU) laboratory. Real-time PCR techniques were used to amplify extracted DNA, if present, in samples. Two proprietary assays developed by Veronica Barragan were used to discriminate between pathogenic, intermediate, and saprophytic species.

A handheld commercial Global Positioning System (GPS) unit, Garmin GPSMAP ® 64, was utilized to capture geographic coordinates (latitude, longitude) of all locations soil samples were collected. Elevation levels were also recorded at each sample location. High resolution IKONOS 3.2 meter and sub-meter QuickBird .61 meter satellite remotely sensed imagery was provided through a digital imagery grant from the commercial producer Digital Globe Inc., Longmont, CO, USA. Positive geo-referenced sites for leptospires were geo-referenced and incorporated into a GIS environment using the commercial GIS software ARC GIS 10.0.

The Harmonized World Soil Database (HWSD) and associated viewer developed and managed by the Food and Agriculture Organization of the United Nations (FAO) was used to determine soil types associated with infective sites. The HWSD contains greater than 15,000 different soil mapping units and combines existing regional and national soil information updates

globally [61]. General soil type data were available and recorded for all georeferenced sample locations based on latitude, longitude coordinates.

Next, a dataset of RGB values, remotely dependent explanatory environmental covariates, and GIS-based indices such as the Normalized Difference Vegetation Index (NDVI) was constructed. Unfortunately, many of the archived images were either not usable due to cloud cover, or the sites did not fall within the images. Data were able to be extracted from 12 positive sites in total. RGB values were extracted from 11 sites on IKONOS and 1 site on a QuickBird image. These values were then log transformed, and a single composite RGB value was calculated for each site. These represented spectral signatures associated with positive infective sites for leptospires. The composite RGB value was used as the dependent variable in regression analysis.

Normalized Difference Vegetation Index (NDVI) values were recorded for each positive site detected on the IKONOS and QuickBird image used in the analysis. The NDVI is the ratio between the red and near-infrared bands useful to quantitatively and qualitatively assess vegetation cover based on spectral data [62]. Based on leptospire moisture requirements and association with hydrology, other important inputs were Digital Elevation Models (DEM) and Soil-Adjusted Vegetation Index (SAVI) where vegetation can be associated with soil type. A DEM represents ground surface topography and is useful for generation of highly accurate predictor variables associated with vector habitats based on spatiotemporal field sampled count data [66]. Several SAVI, NDVI, and DEM maps were constructed using IKONOS and QuickBird images within ARC GIS 10.0 (Figures 8-15)

When developing a soil-based remote sensing model the following factors influencing soil reflectance must be considered: mineral composition, soil moisture, organic matter content, and soil surface texture. Absorption bands ranging from 1.4-1.9 μm relate to soil moisture content. The soil line of reflectance spectra can be calculated using the least squares

regression method where NIR (soil) is equal to “a”, red (soil) is equal to “b” where Red (soil) is the soil reflectance in the red band. NIR (soil) is equal to the soil reflection in the near-infrared [63]. The soil line is the graphical relationship between red and near-infrared (NIR) bands [62,63]. The a,b parameters of the soil line are estimated by the least squares regression method. High organic matter and rough texture in a soil can produce spectral interference for band characteristics [63].

Individual soil sample analysis for mineral composition, organic matter levels, or texture information was not performed due to cost; however, soil moisture data were available and acquired from satellite-derived ASCAT soil moisture. The ASCAT soil moisture product is produced by EUMETSAT using WARP NRT software originally developed by IPF/TU Wien (Institute of Photogrammetry and Remote Sensing, Vienna University of Technology) [64,65]. Ground surface topography has been shown useful for generation of highly accurate predictor variables associated with vector habitats based on spatiotemporal field sampled count data [66].

Generalized linear regressions models based on remotely sensed and geographic data were constructed where the RGB value was the dependent variable. Independent variables evaluated in the model were NDVI and SAVI values in addition to elevation and Euclidean distances of sample sites from the river edge.

The Poissonized model using endmember spectral values was based on the Poisson distribution, a discrete probability distribution with the following formula:

$P(X)= \lambda^{x}e^{-\lambda}/X!$

Where P(X) is the probability of exactly X occurrences;

λ (lambda) equals the number of occurrences per unit of time (mean occurrence rate);

e=2.718, the base of the natural logarithms

X equals specific values (0,1, 2, 3, etc…) of the random variable;

Where the mean and variance of the Poisson distribution are equal and computed as follows:

expected value = λ

variance = λ [75]

A negative binomial regression model was constructed to model count variables and to account for over-dispersion (excess Poisson variability) which was tested for using a likelihood ratio test. A directory of variables was created and analyzed using a commercial statistical analysis software (SAS 9.4) to generate pseudo R-squared values. This procedure was duplicated for an additional regression model where case prevalence rates from 2009-2013 (Table 1) represented the response variable. In this non-ordinal regression model, the independent variables were as follows: soil moisture (upper soil layer < 2 cm expressed as % of saturation (%)), average annual rainfall (mm), evaporation potential (average annual percentage %), humidity (average annual % relative humidity), temperature (average annual temperature (^{o}C), river flow (average annual river level (m)), land surface temperature (average annual surface temperature (^{o}C).

CHAPTER 4: RESULTS

A total of 64 soil samples were collected and analyzed by real-time PCR in order to detect leptospiral DNA and classify pathogenic status of all positive samples. Out of the 64 soil samples, 21/64 (32.8%) were positive for the presence of leptospiral DNA and 43/64 (67.2%) were negative. 14/21 (66.7%) of the leptospire positive soil samples were further determined to be of intermediate pathogenicity. 8/21 (38.1%) of the leptospire positive soil samples were determined to be saprophytic species. 1/21 (4.76%) of positive samples were positive for both intermediate and saprophytic leptospire species. No positive findings for pathogenic (*L. Interrogans*) species were reported.

Table 3 shows the dominant soil types associated with all samples. Cambisol was found in soils testing positive for intermediate leptospires 11/64 (17.2%) and in negative soil samples 22/64 (34.4%). Phaeozems soil type was found in benign, intermediate, and negative samples. Luvisols soil type was found in benign and negative samples. Regression outputs for the response variable RGB composite variable from the commercial Statistic Analysis Software 9.4 are provided in Figures 27-34. An alpha level of .05 was used for all tests. Of all variables analyzed, only elevation demonstrated statistical significance in the Generalized Linear Model were the p value was .0343 (Figure 32). The results of the non-ordinal regression model using prevalence as the response variable were negative for both Portoviejo and Santa Ana. Ordinary krig maps using composite RGB and NDVI values associated with positive leptospire soil sites were constructed (Figures 16-18).

CHAPTER 5: CONCLUSION

Discussion

Soil-related remote co-variates failed to show significance in constructed regression models when evaluated using composite RGB values as the response variable; however, this may be due to the small sample size of 12 that used in the analysis. Elevation demonstrated significance in the Generalized Linear Model regression; however, a negative confidence interval was reported likely attributable to the small sample size. It is biologically plausible for the statistical significance not to be artifact as lower elevations are more prone to flooding which is a well-established geographic risk factor for leptospirosis. A sample size of at least 31 known infective soil site is needed to fully evaluate statistical significance (B. Jacob, personal communication, April 10, 2013). The study would have been strengthened by acquisition of remotely sensed imagery of the entire study area in order to have the ability to extract all RGB values. High- and sub-meter resolution RS imagery is expensive, and to acquire the entire study site was cost-prohibitive. The study area is not routinely imaged with commercial satellites and would require a request for satellite images to be taken which significantly adds to the cost. Archived images are often available at a significantly reduced cost. Fortunately, we were able to acquire IKONOS and QuickBird images which were kindly provided through a free student digital grant from Digital Globe, Inc. Additionally, several important soil-related factors were unable to be directly measured including mineral composition, organic matter content, and soil surface texture which affect soil reflectance and would be important in the soil remote sensing model. Ideally, analysis at a soil laboratory, of individual soil sample properties would be performed and would provide more fidelity to any soil-based model. This is especially

important considering the complexity of soil science at the microscopic level. Dominant soil types found in sample sites included cambisol, phaeozems, and luvisols. Cambisol has previously been associated as a potential driver of leptospirosis [47]. It was observed that, except for one sample, intermediate and benign species were detected independent of each other. This trend may be due the relative low tolerance of pathogenic leptospires to compete with environmental bacteria. This trend should be explored further to determine if the presence of high bacterial loads in environmental samples (soil, water) could be an indicator for the presence or absence of pathogens. In other words, it may be easier to measure risk in this manner given the difficulty in detecting *L. interrogans* in the environment. Also, as an observation of this study, it was noted that cambisol soil distribution is common to both Latin America and Africa where leptospirosis is problematic (Figure 24). This soil type was associated with disease risk in Nicaragua [47].

Conclusions

Soil-based co-variates were used to determine its potential as a potential stationary indicator with predictive power for identification of high risk transmission. Geo-spatial trends were noted that point to the potential use of soil as a stationary indicator. Soil holds predictive potential for identifying high risk transmission areas for leptospirosis that warrants further analysis. Leveraging GIS and RS technology with soil sampling is a novel idea that could be developed further to create a robust model for predicting leptospirosis risks areas. If soil could be linked directly as a static indicator it would offer tremendous power to public health planners in resource poor settings for focal targeting of limited resources and public health education programs where risk is greatest. A targeted control program where prophylactic antibiotics could be dispensed prior to seasonal occurrences could be feasible if risk areas were able to be accurately delineated. Environmental sampling, especially in highly impacted, endemic areas could be incorporated into ongoing surveillance programs if predictive patterns of disease can

be associated with soil type. Molecular analysis of strains found in soil need to be further linked to human or animal strains to better establish the role of leptospires of intermediate pathogenicity.

GIS and RS public health research has significantly advanced the general understanding of disease transmission patterns in many human and animal diseases by allowing maximum and real-time exploration of the total environment. This technology offers researchers an enhanced tool for conducting spatial epidemiology where multiple environmental, geographical, social, and demographic factors can be simultaneously explored to uncover patterns in transmission and characterize risk.

Recommendations

Public health policy and planning for a disease like leptospirosis with a complex transmission cycle involving humans and animals bridged by environments such as soil and water requires active, ongoing surveillance in all three domains. Establishment of national, regional, and local real-time geodatabases that allow submission, management, and synchronous analysis of health and environmental sampling data is highly recommended. Ecuador's Ministry of Health has established GeoSalud 2.0 (GeoHealth) national database which maps health information and data from the 9 health zones within Ecuador [76]. Establishment of these systems allow for continuous refinement of disease risk assessment and offers a highly efficient public health planning tool and are crucial in order to be able to focus limited resources and public health programs where they are needed the most.

TABLES AND FIGURES

Table 1: Human leptospirosis area prevalence, Manabí, Ecuador, 2009-2013. Color strata from light to dark represents low to high quantile.

Areas	2009	2010	2011	2012	2013
	Cases	Cases	Cases	Cases	Cases
PORTOVIEJO	245	252	289	521	114
MANTA	1	8	12	30	52
CHONE	5	9	9	57	25
JIPIJAPA	1	1	2	12	4
BAHIA	2	6	8	26	3
CALCETA	13	19	25	71	23
ROCAFUERTE	8	26	15	32	28
EL CARMEN	0	0	0	16	0
PAJAN	16	5	12	7	1
SANTA ANA	29	71	81	110	46
PICHINCHA	0	2	2	5	3
PEDERNALES	0	2	1	6	0
TOTAL	320	401	456	893	299

Table 2: Human leptospirosis area incidence per 10 000, Manabí, Ecuador, 2009-2013. Color strata from light to dark represents low to high quantile.

Areas	2009	2010	2011	2012	2013
	Incidence	Incidence	Incidence	Incidence	Incidence
PORTOVIEJO	9.0	9.2	10.6	19.0	4.2
MANTA	0	0.3	0.4	1.0	1.8
CHONE	0.3	0.5	0.5	3.5	1.5
JIPIJAPA	0.1	0.1	0.2	1.3	0.4
BAHIA	0.2	0.7	0.9	3.0	0.3
CALCETA	1.3	1.9	2.5	7.0	2.3
ROCAFUERTE	1.5	5.0	2.9	6.1	5.3
EL CARMEN	0	0	0.0	1.5	0.0
PAJAN	3.9	1.2	2.9	1.7	0.2
SANTA ANA	3.1	7.4	8.5	11.5	4.8
PICHINCHA	0	0.6	0.6	1.5	0.9
PEDERNALES	0	0.4	0.2	1.1	0.0
TOTAL	23.3	2.9	3.3	6.4	2.2

Table 3: Dominant soil type associated with each sampled soil site

Leptospire species	Cambisol	Phaeozems	Luvisols	Total
Benign	0	4	3	7
Intermediate	11	4	0	15
Not detected	22	13	8	43
Total	33	21	11	65*

*64 samples were collected. One sample tested positive for both benign and intermediate leptospire specie

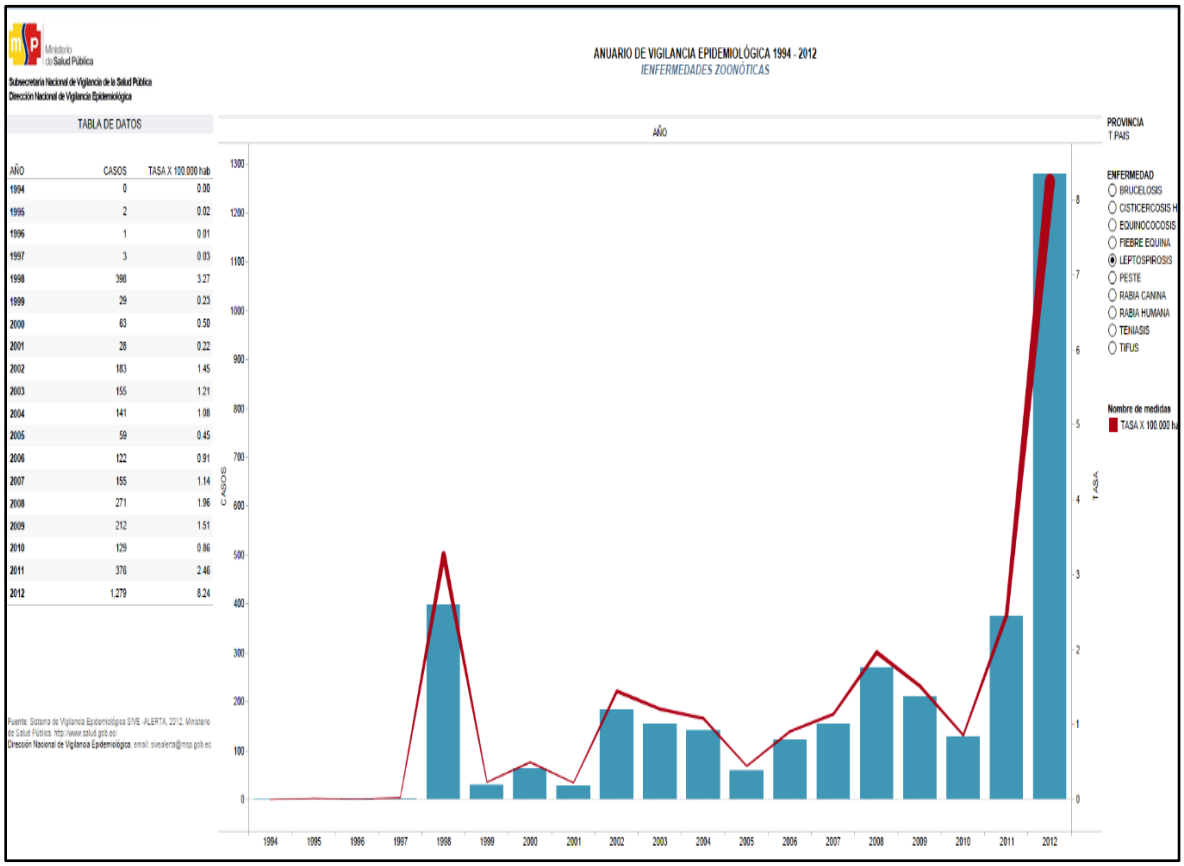


Figure 1: Annual human leptospirosis prevalence rates in Ecuador, 1994-2012

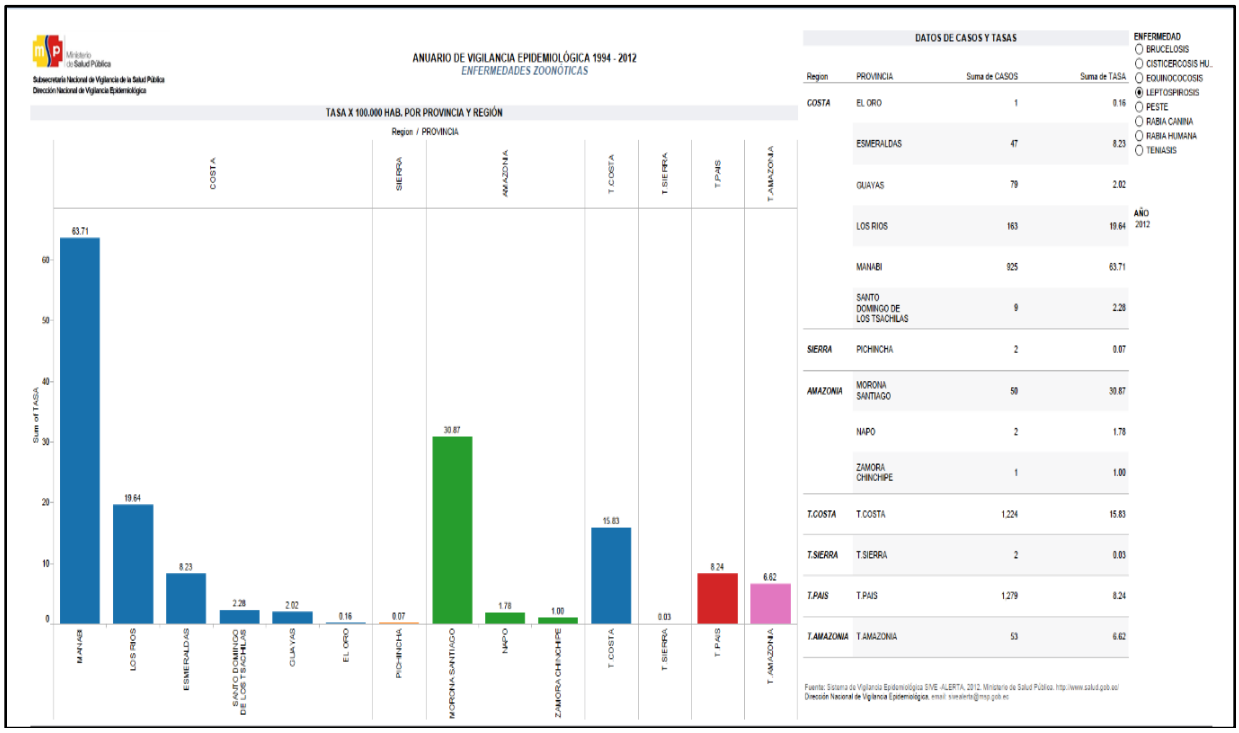


Figure 2: Human leptospirosis prevalence and incidence by region and province, Ecuador, 1994-2012

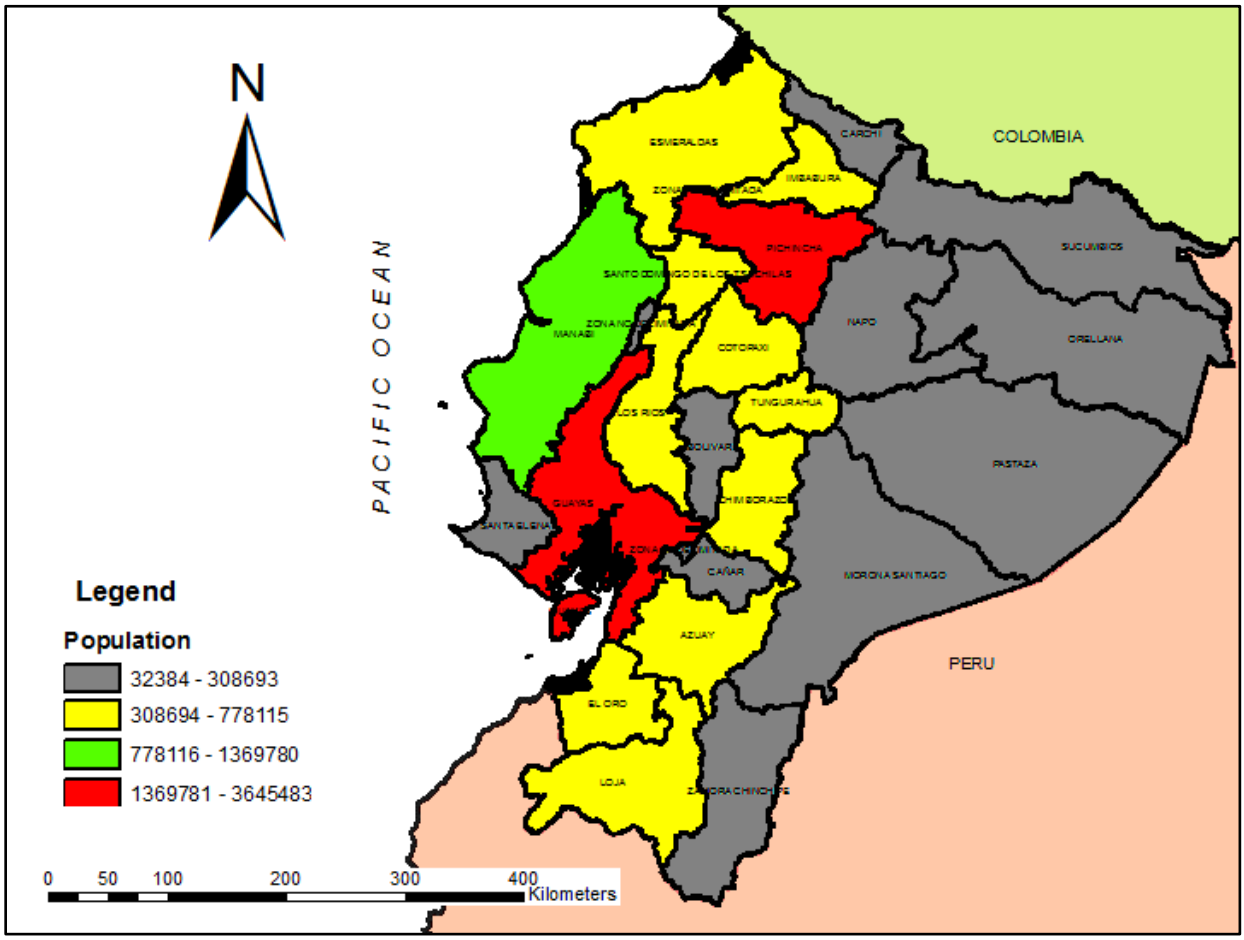


Figure 3: 2010 population map of the Republic of Ecuador (Continental)

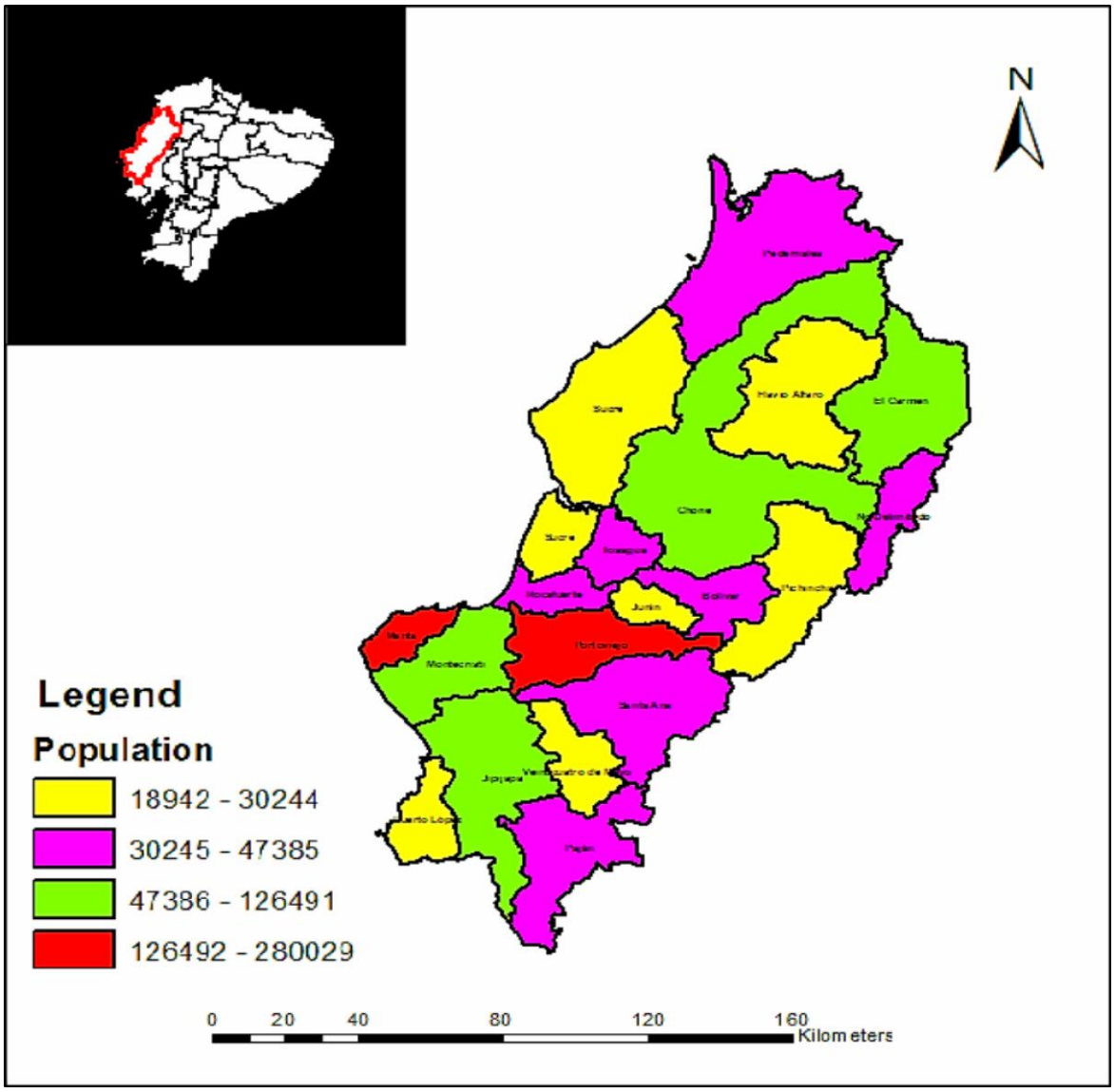


Figure 4: 2010 population map of the cantons of Manabí Province, Ecuador

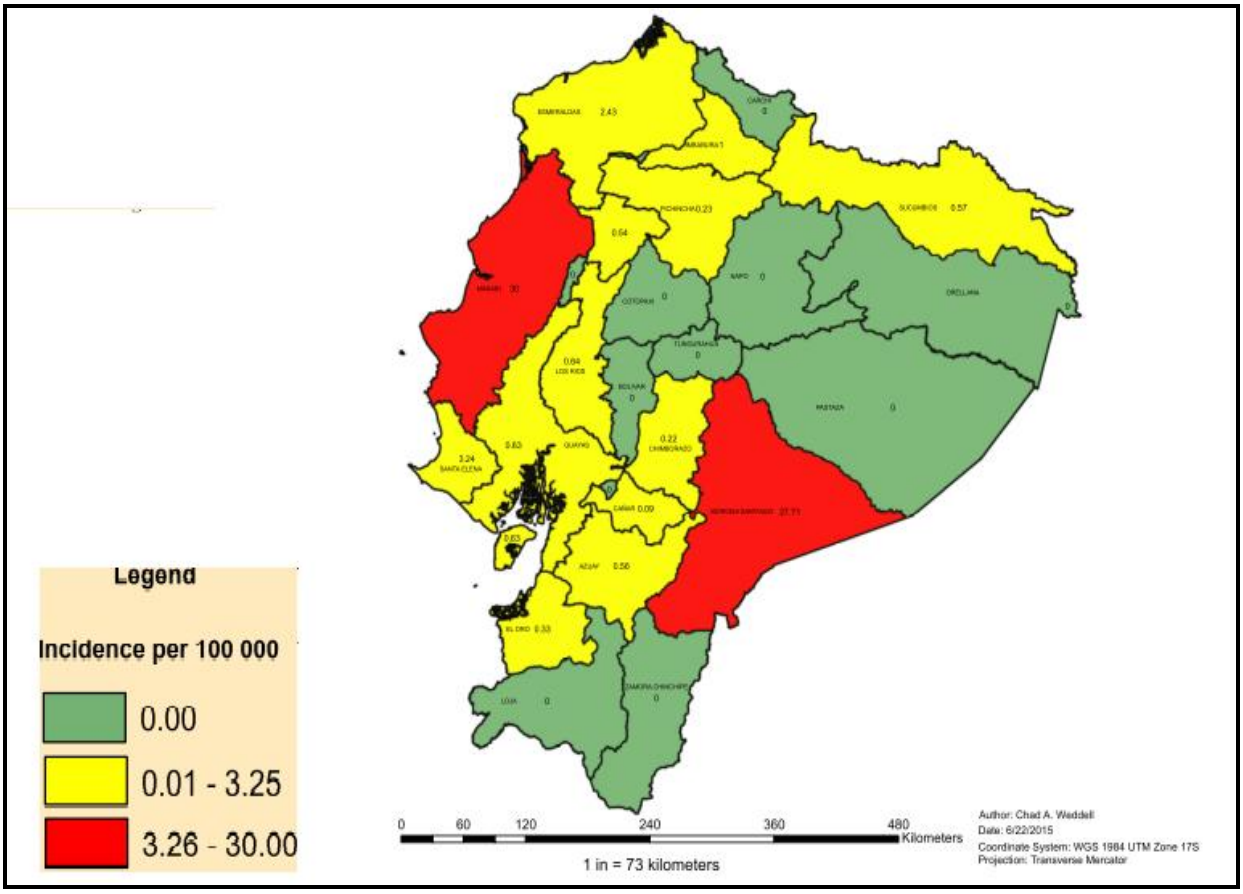


Figure 5: 2011 leptospirosis incidence map per 100 000 population by province, Continental Ecuador. The actual incidence values are reflected next to province names.

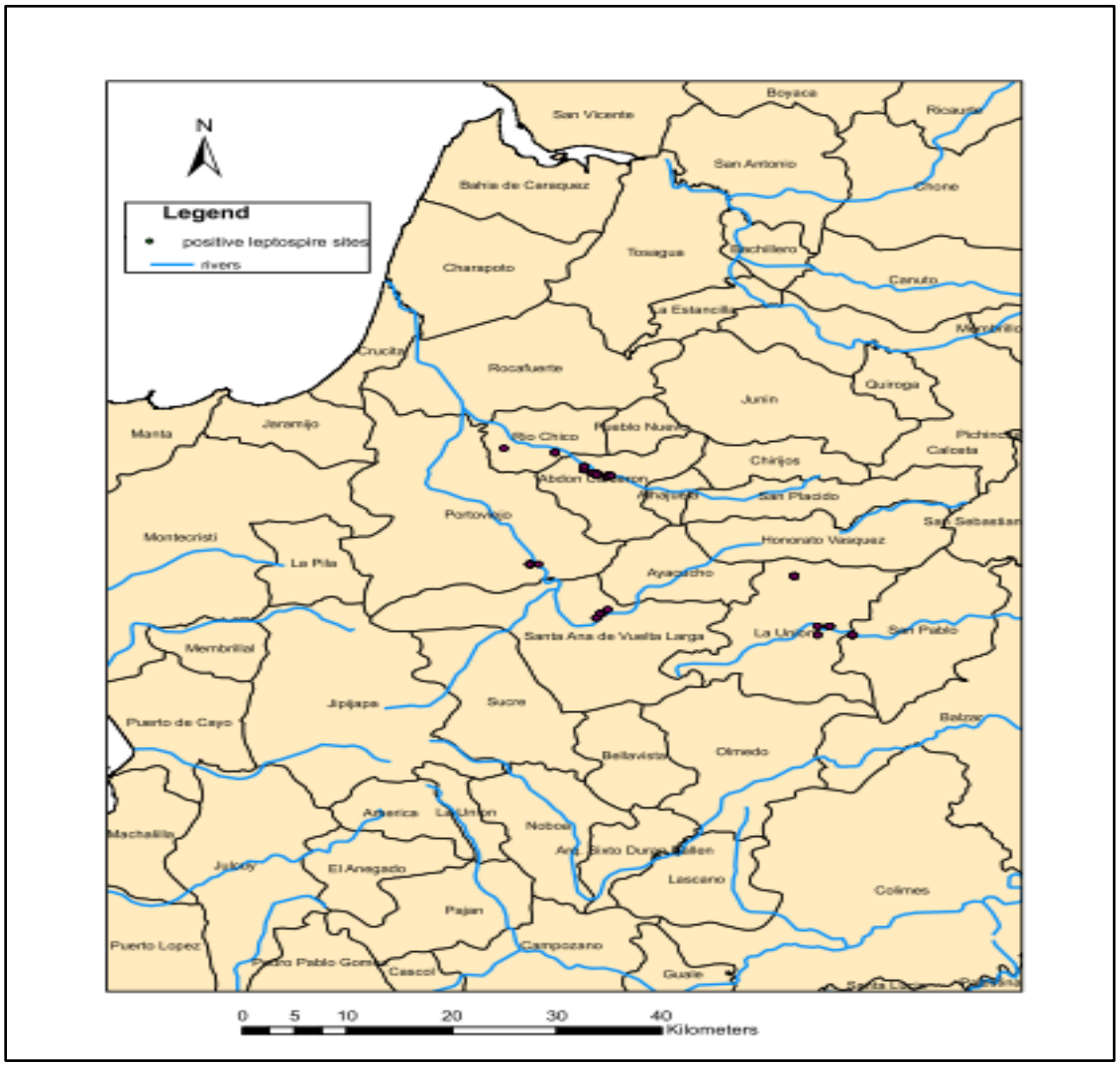


Figure 6: Positive leptospire soil sites in sampled parishes, Manabí, Ecuador

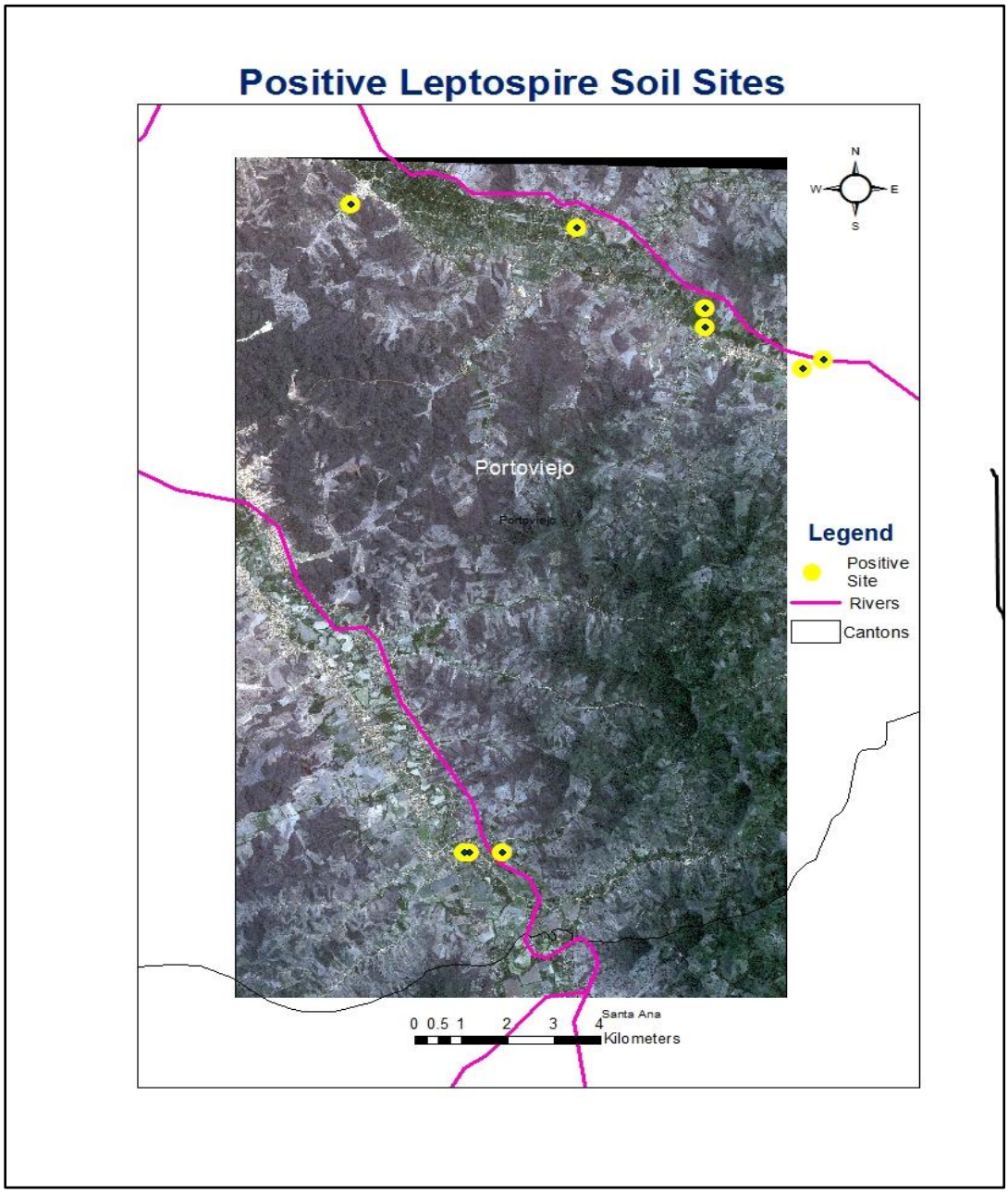


Figure 7: Positive leptospire soil sites displayed on high resolution (3.2 m) IKONOS satellite imagery.

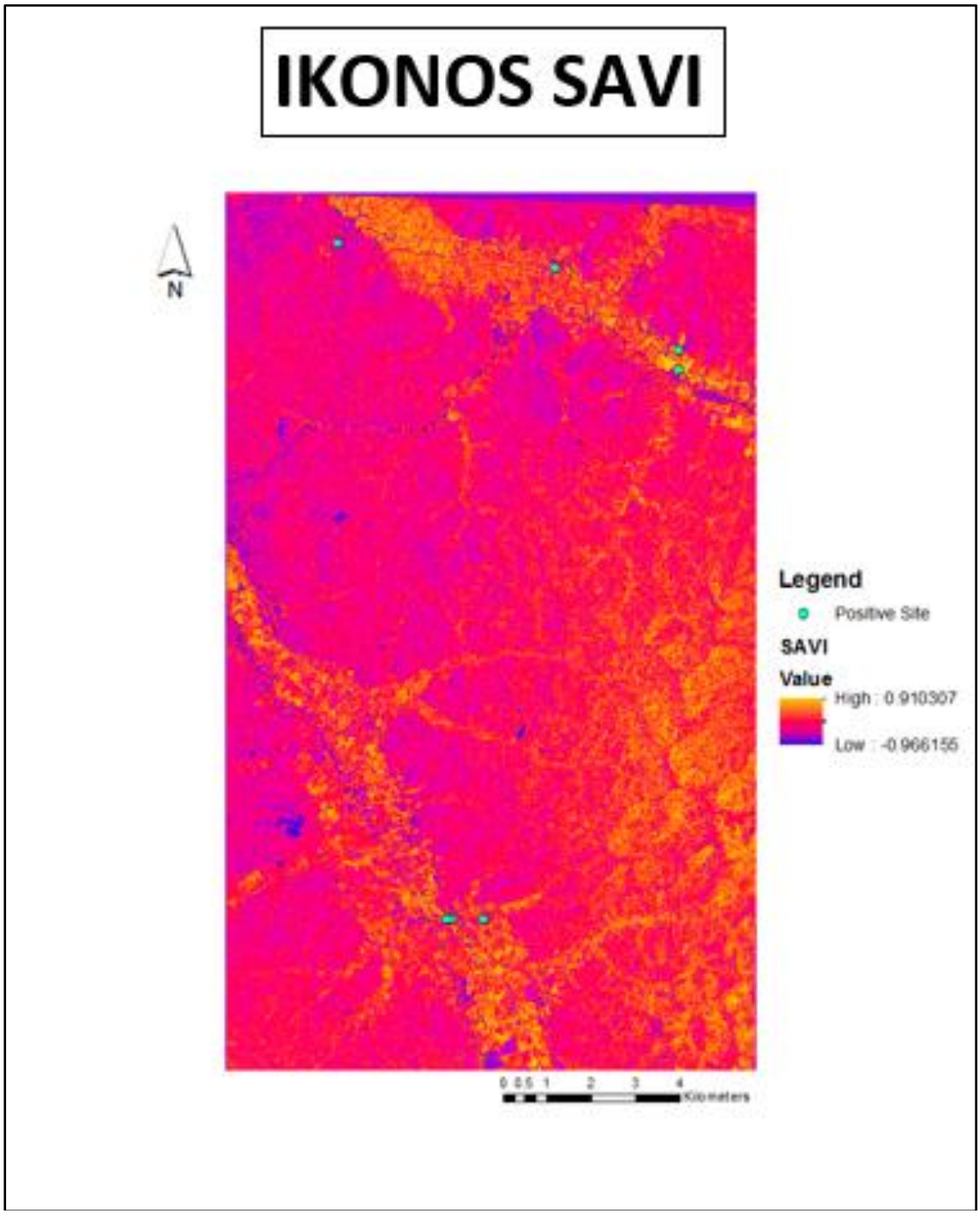


Figure 8: Soil Adjusted Vegetation Index (SAVI) associated with positive leptospire soil sites in Portoviejo, Manabí, Ecuador constructed using IKONOS (3.2m)

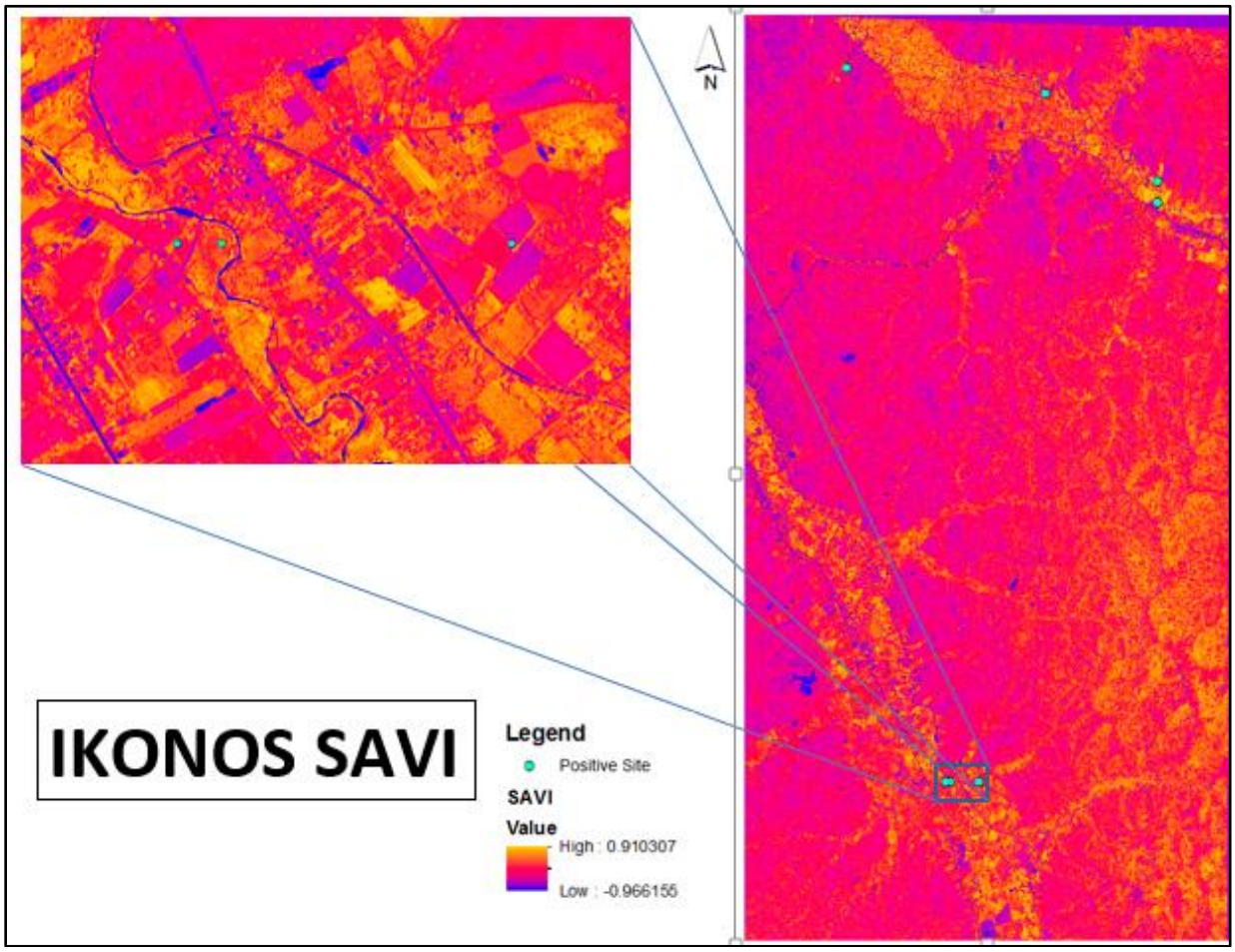


Figure 9: Zoomed view of SAVI associated with positive leptospire soil sites in Portoviejo, Manabí, Ecuador on IKONOS image (3.2 m)

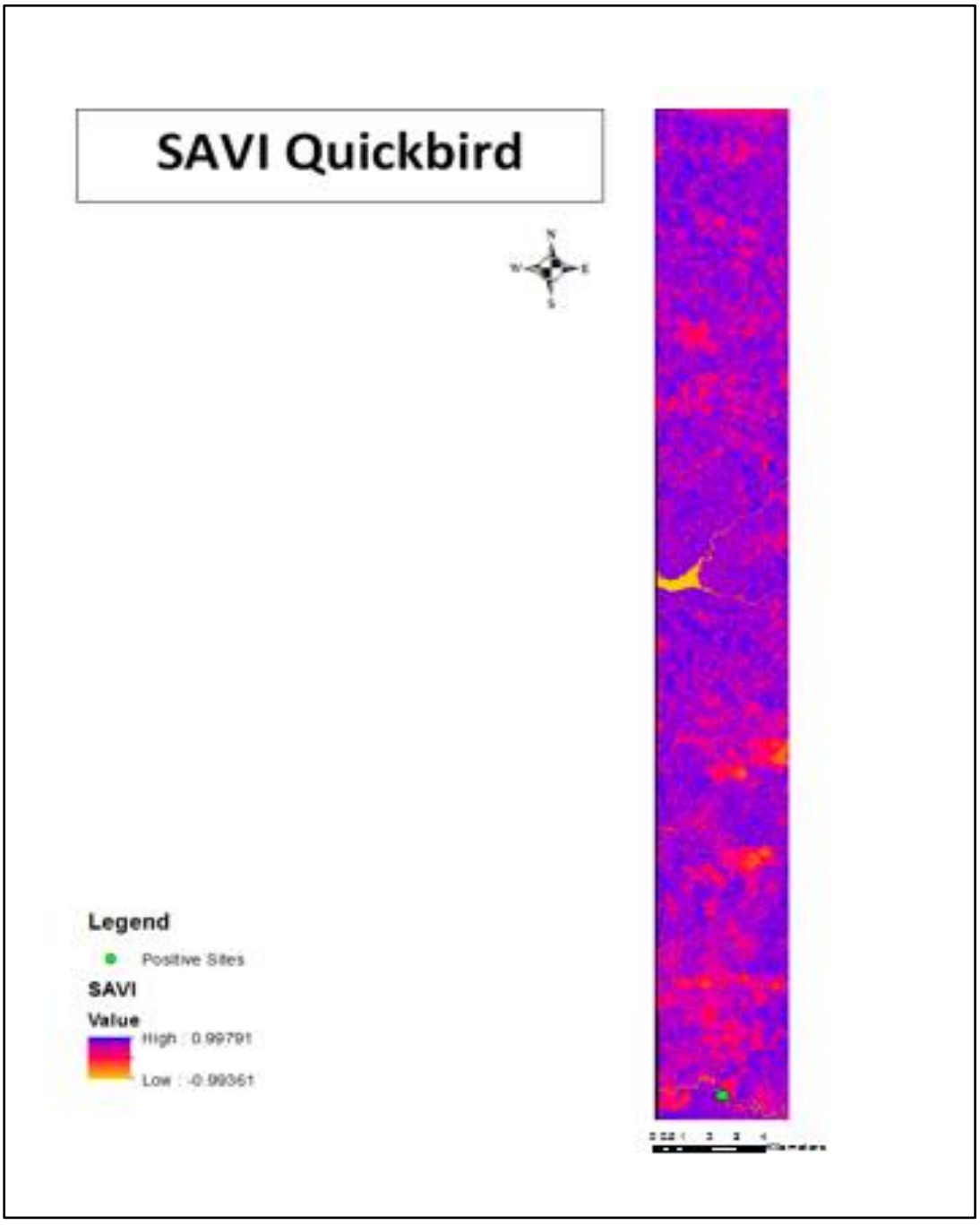


Figure 10: SAVI constructed from QuickBird (.61m) image showing positive leptospire soil site in Santa Ana, Manabí, Ecuador

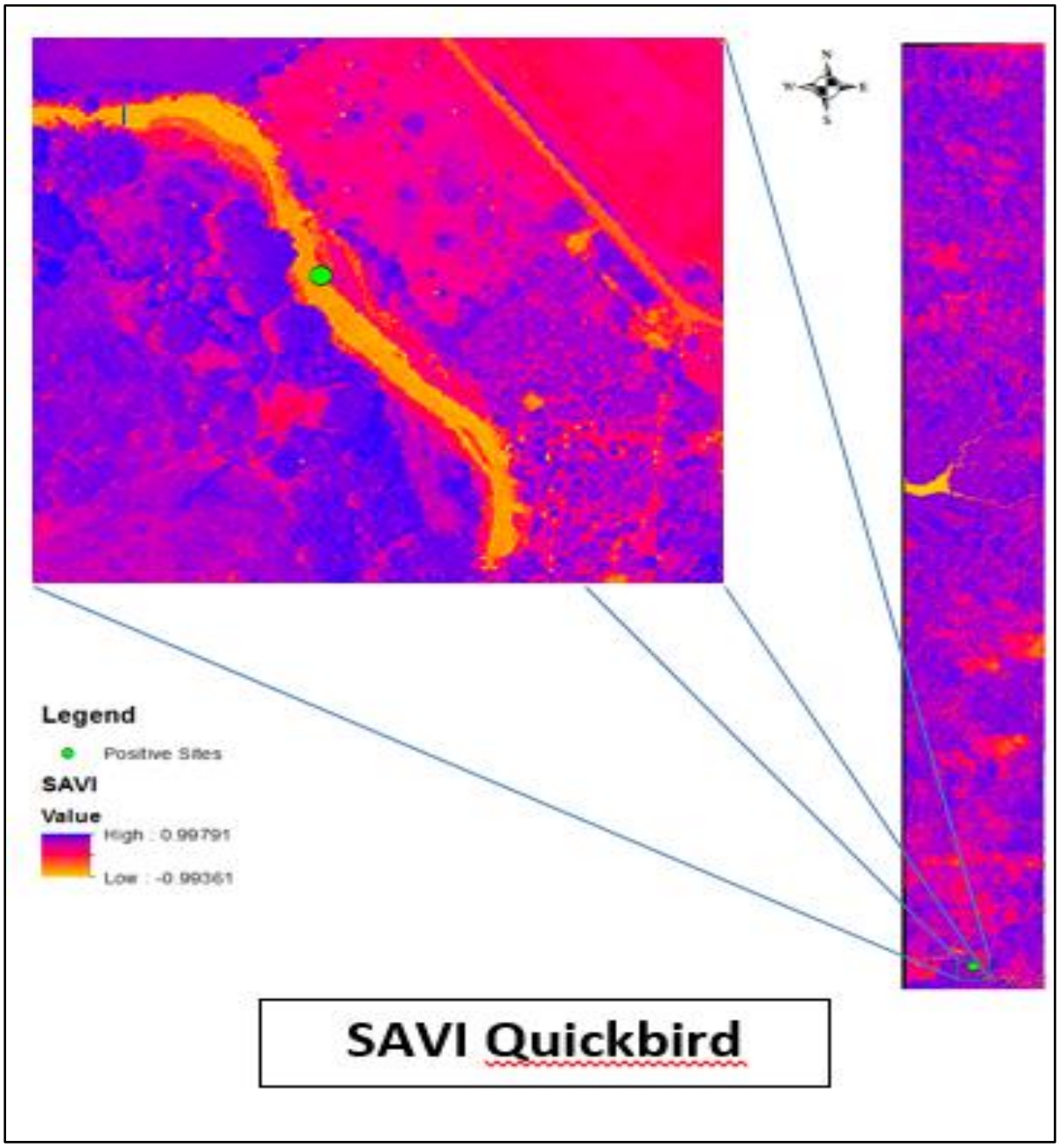


Figure 11: Zoomed view of a SAVI constructed from QuickBird (.61m) imagery showing positive leptospire soil site in Santa Ana, Manabí, Ecuador

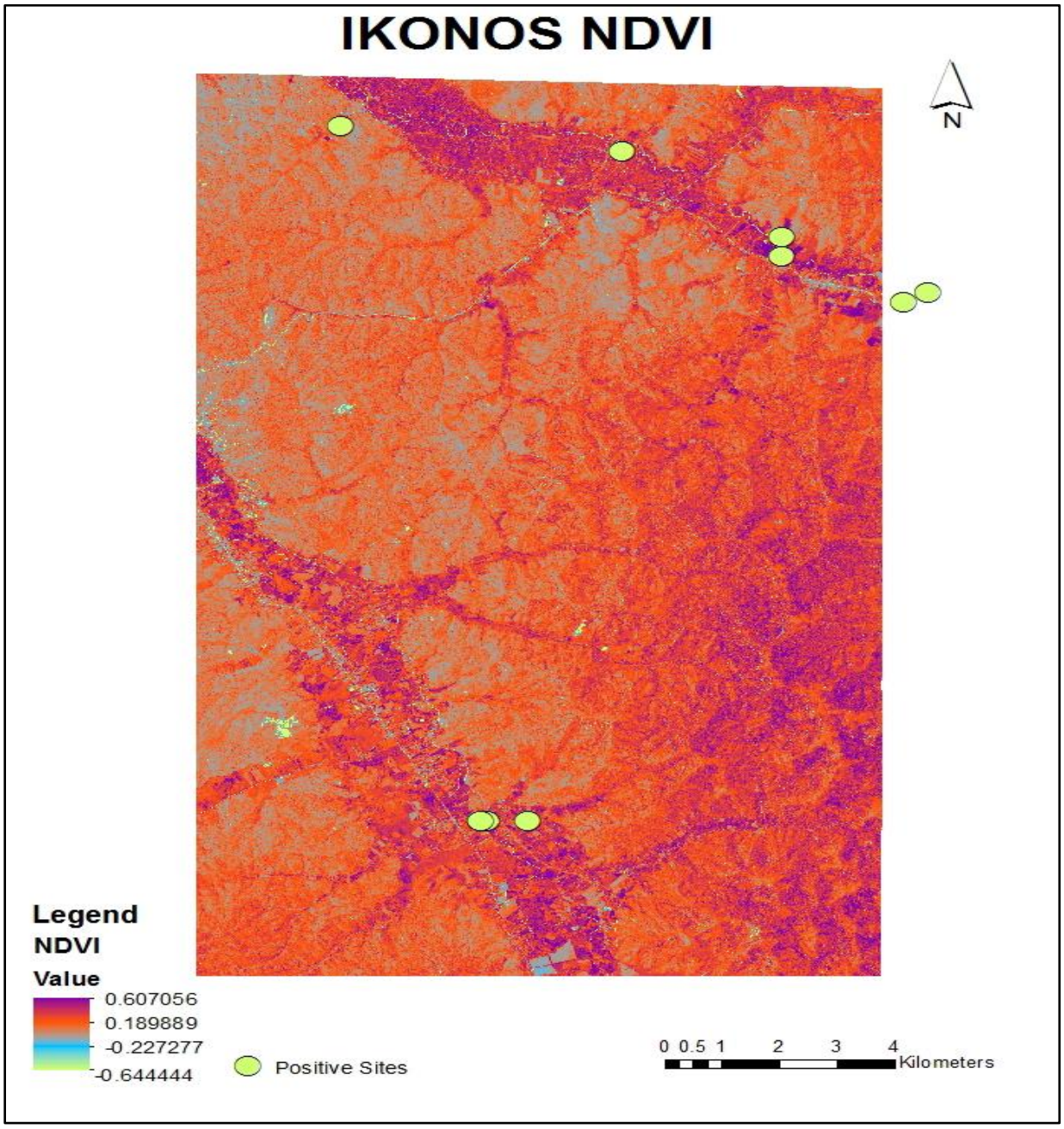


Figure 12: Normalized Difference Vegetation Index (NDVI) on IKONOS (3.2 m) image showing positive leptospire soil sites in Portoviejo, Manabí, Ecuador

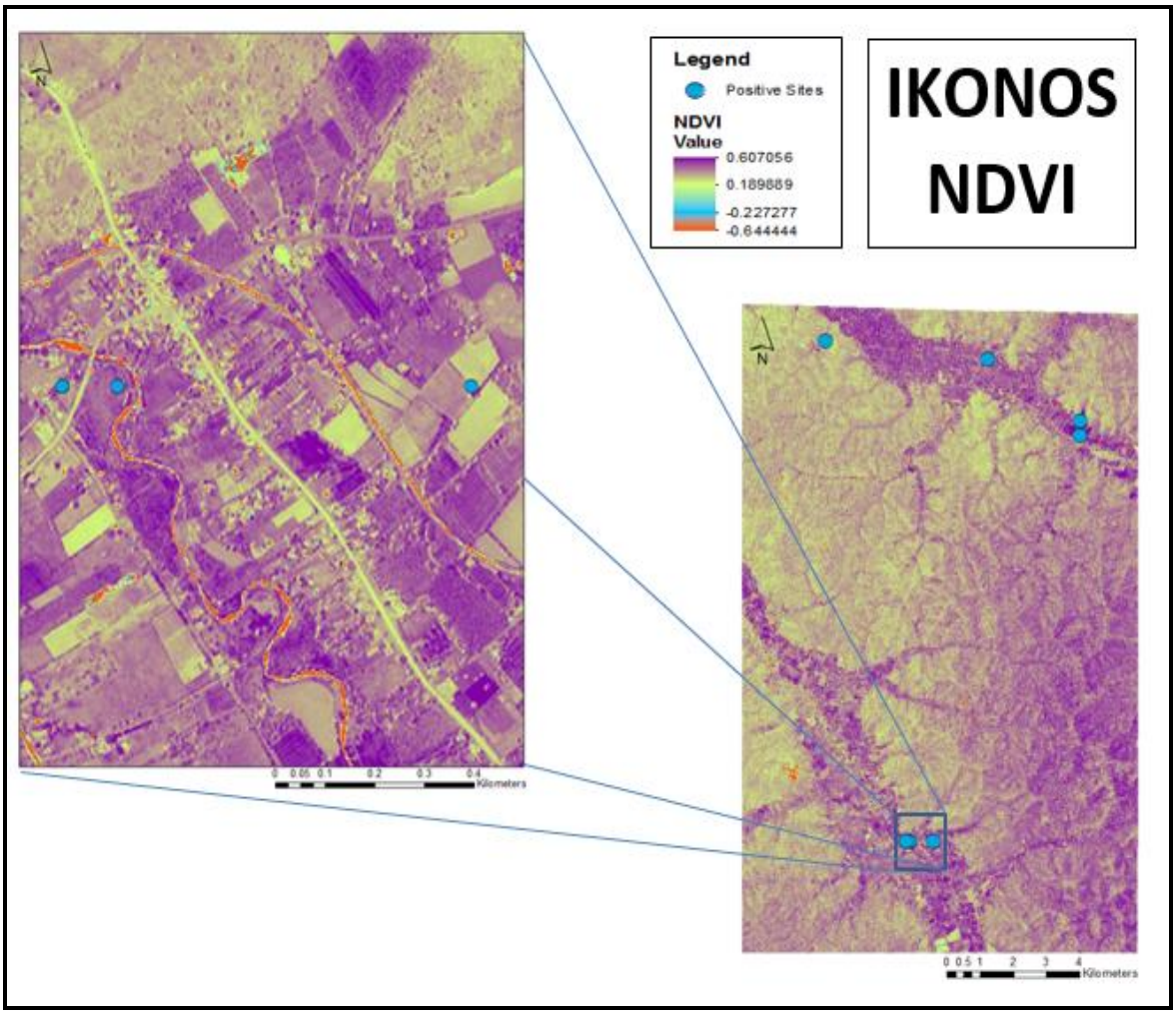


Figure 13: Zoomed view of NDVI constructed from IKONOS (3.2 m) image showing positive leptospire soil sites in Portoviejo, Manabí, Ecuador

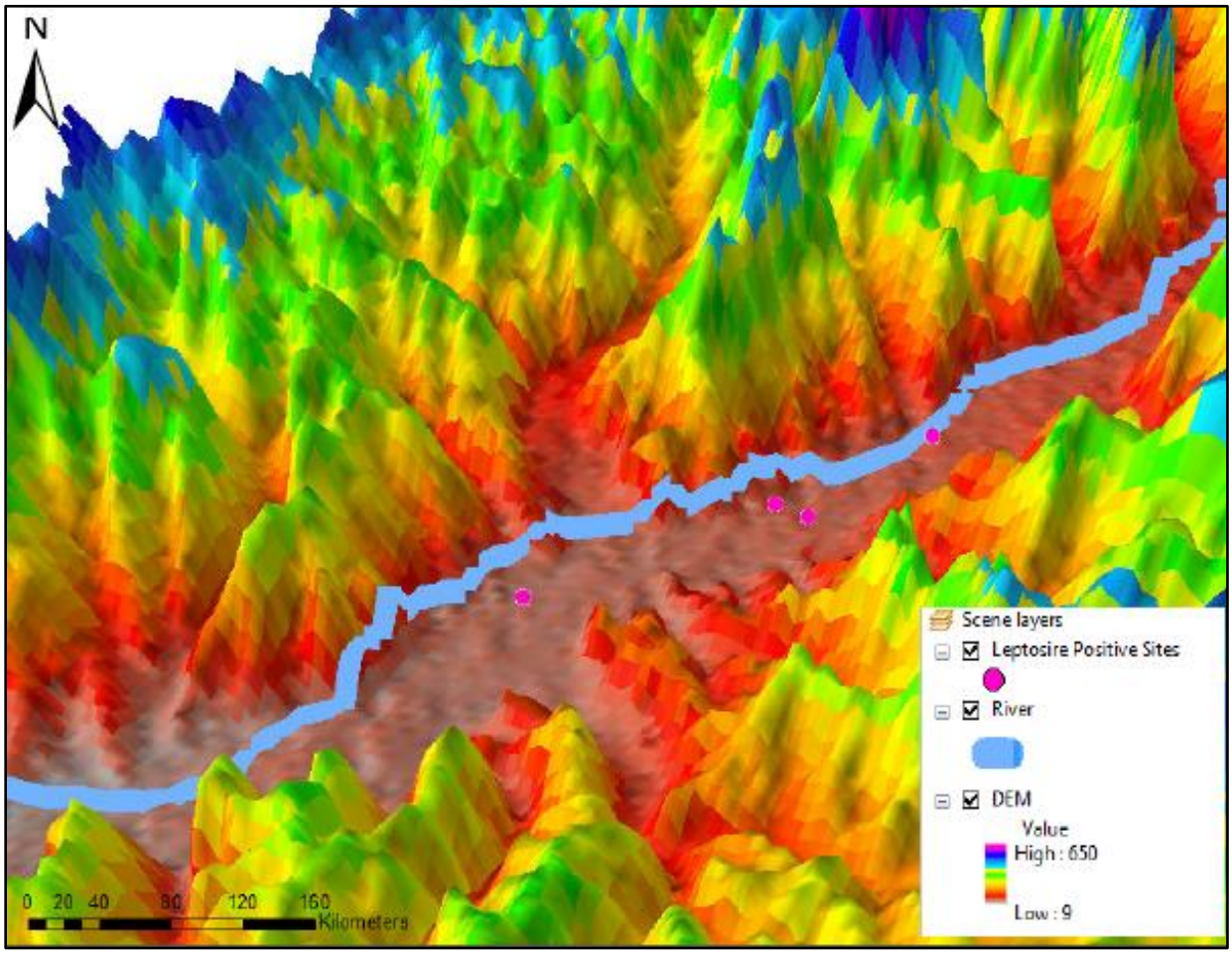


Figure 14: Digital Elevation Model (DEM) with georeferenced positive leptospire soil sites, Portoviejo, Manabí, Ecuador

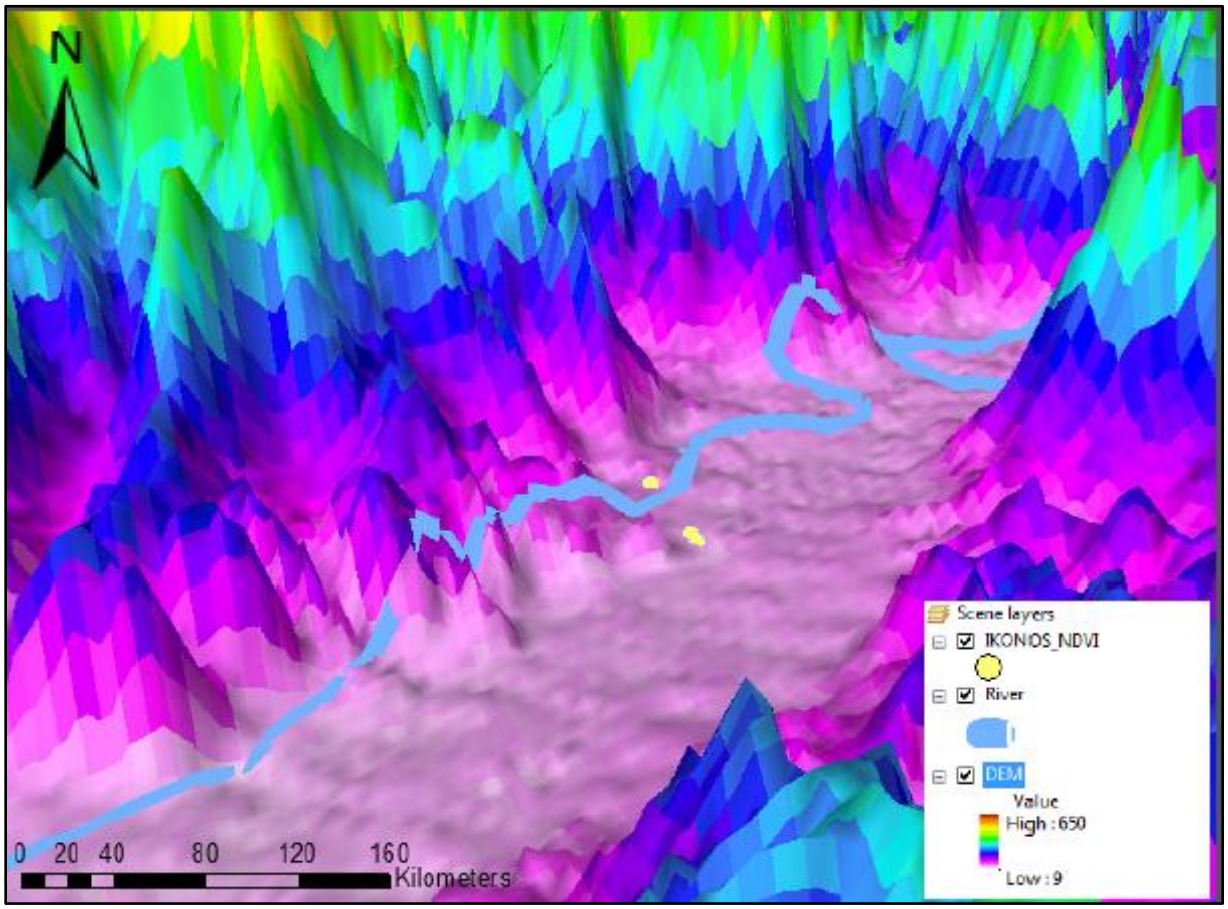


Figure 15: DEM on IKONOS image (3.2 m) showing NDVI positive leptospire soil sites, Portoviejo, Manabí, Ecuador

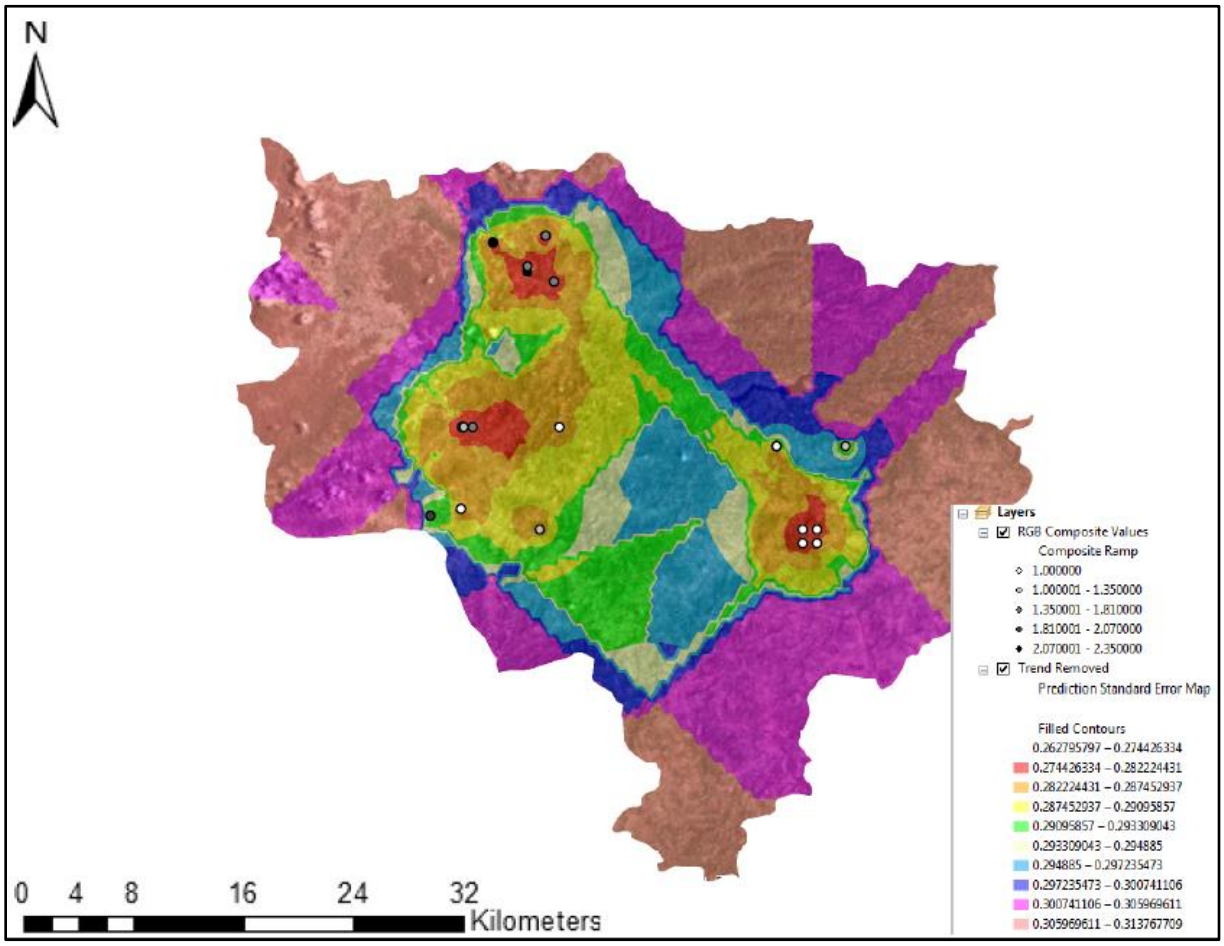


Figure 16: RGB composite value krig, Portoviejo, Manabí, Ecuador

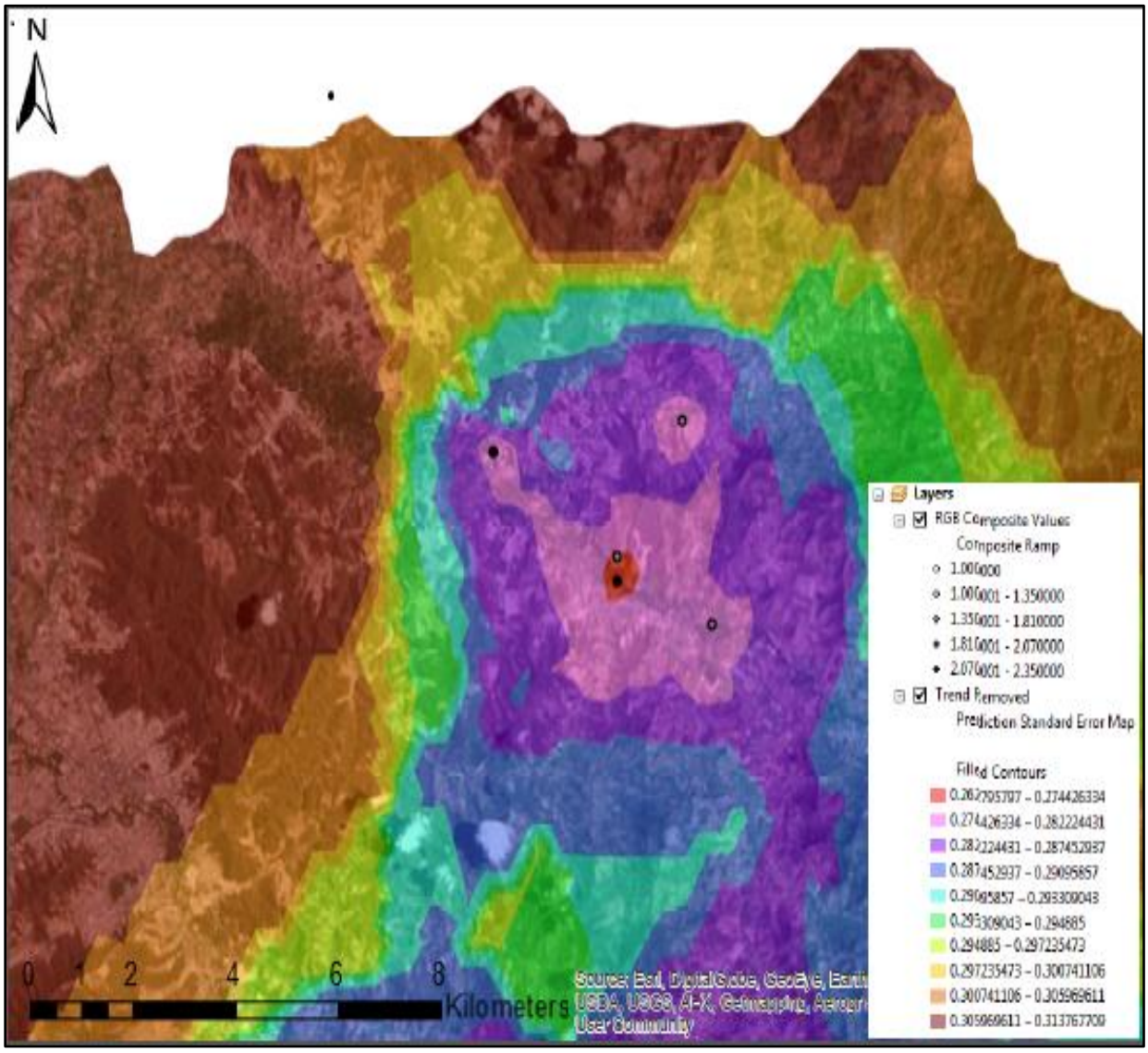


Figure 17: Zoomed RGB composite value krig, Portoviejo, Manabí, Ecuador

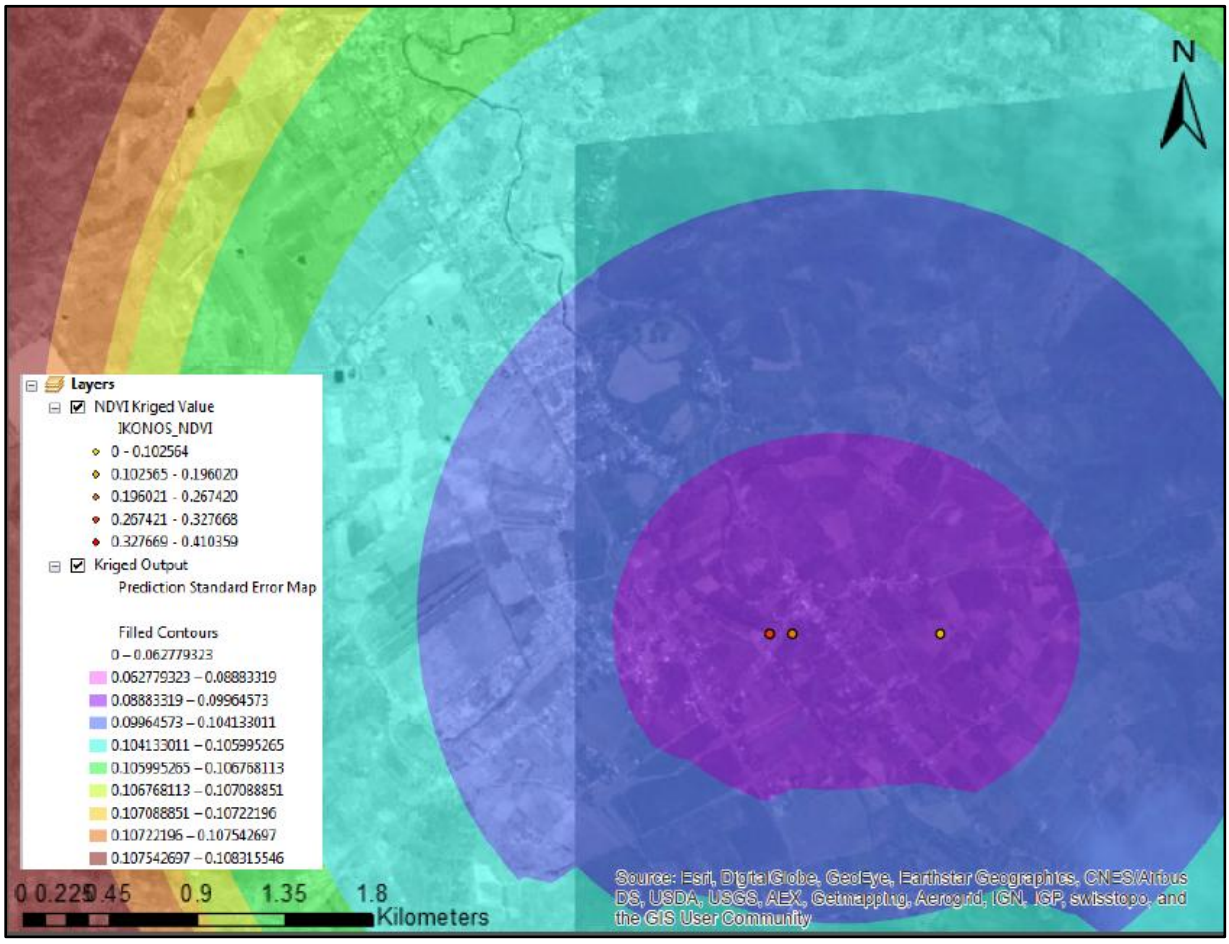


Figure 18: NDVI value krig, Portoviejo, Manabí, Ecuador

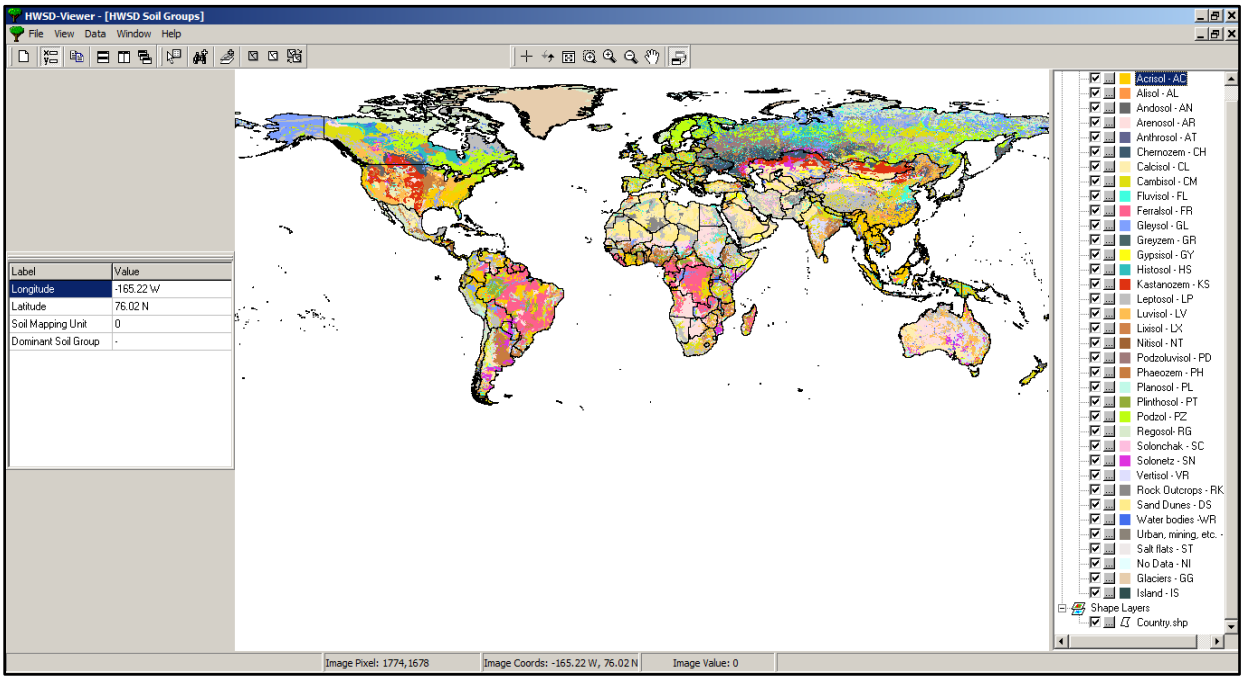


Figure 19: Screenshot of the Harmonized World Soil Database (HWSD) viewer depicting the dominant soil types of the world

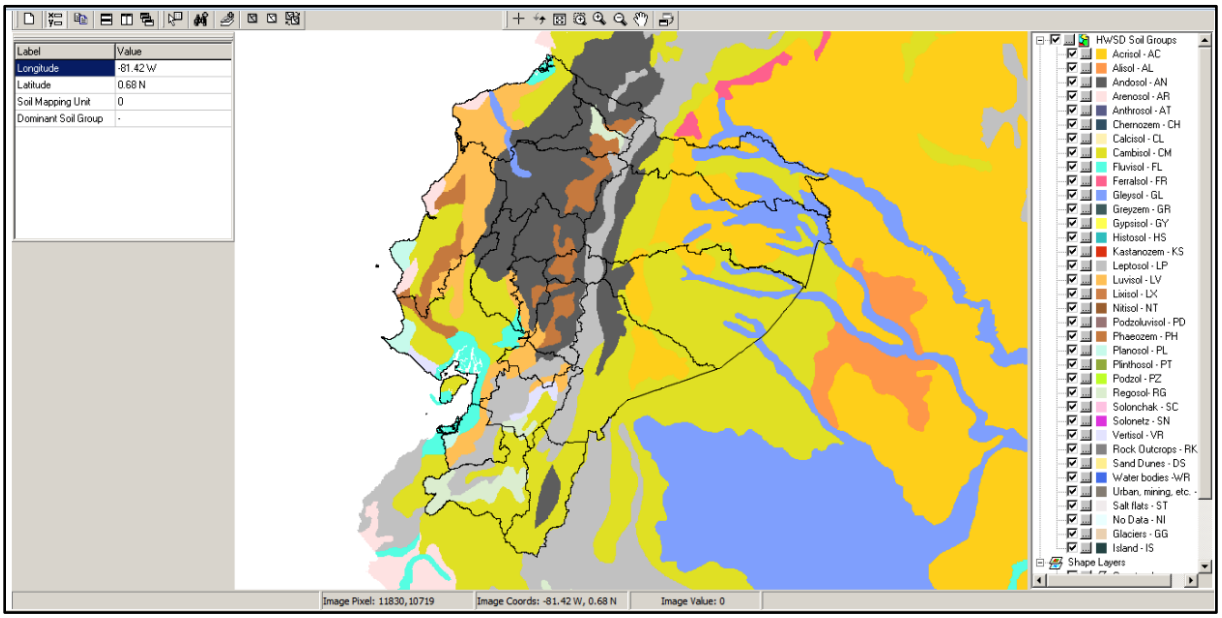


Figure 20: Screenshot of the HWSD viewer depicting the dominant soil types of the Republic of Ecuador (Continental)

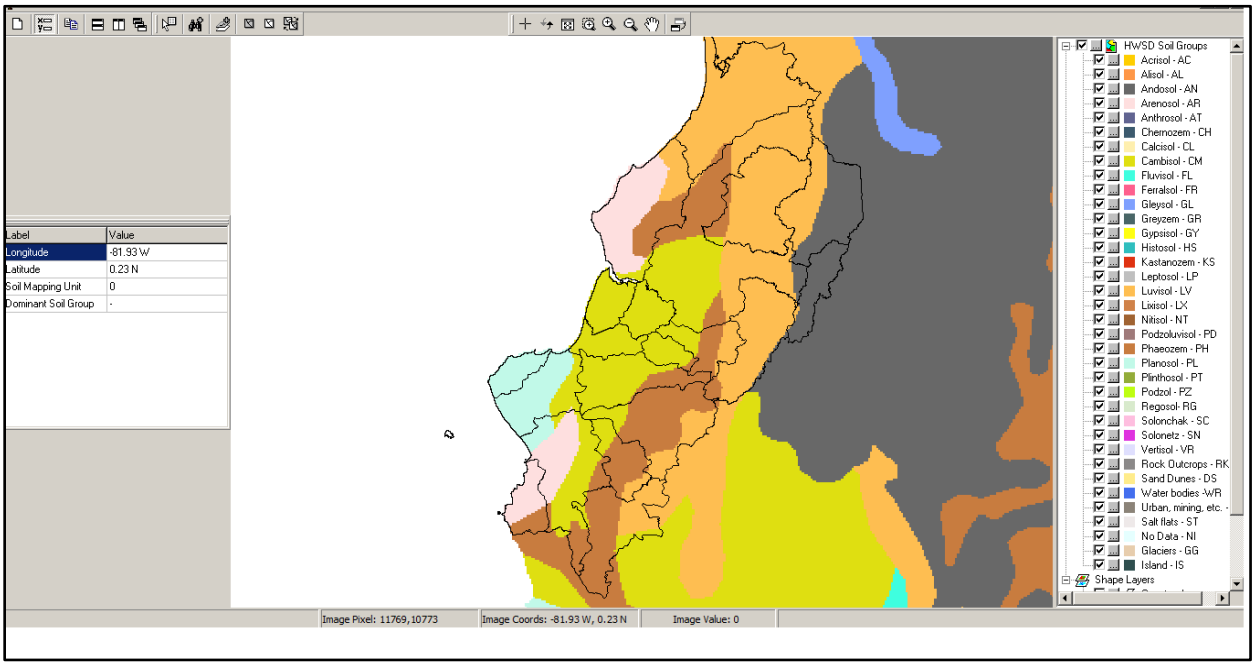


Figure 21: Screenshot of the HWSD viewer depicting the dominant soil types of Manabí, Ecuador

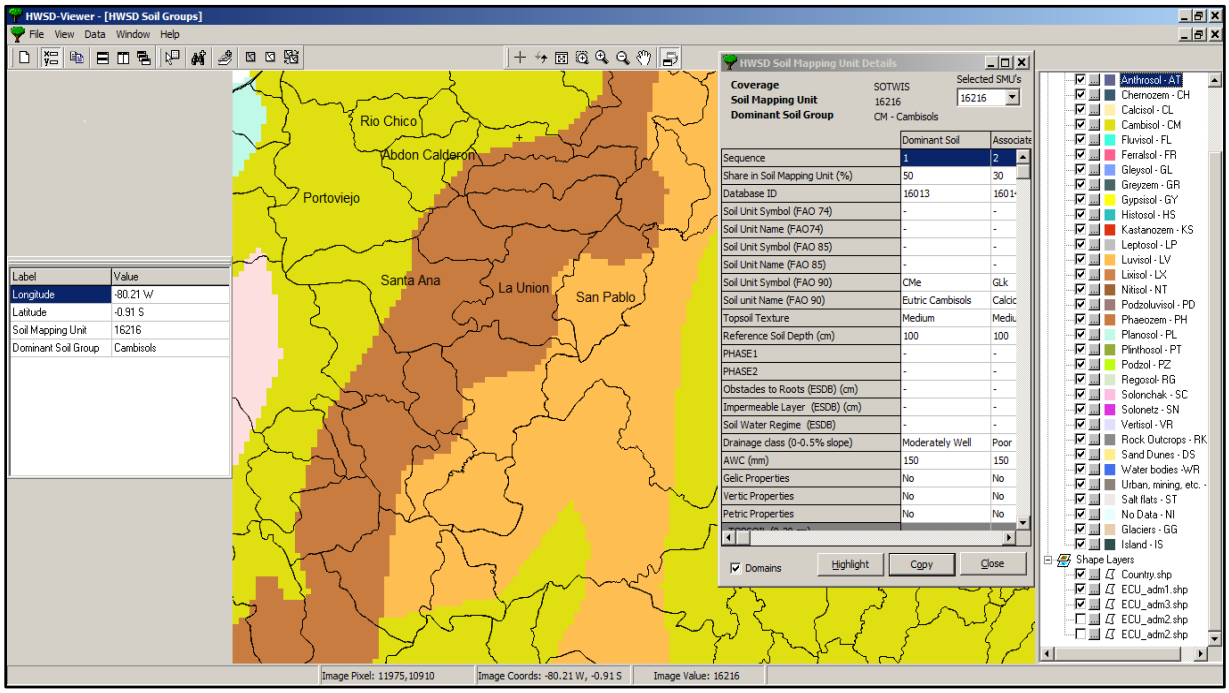


Figure 22: Screenshot of the HWSD viewer depicting the dominant soil types of parishes of Manabí, Ecuador

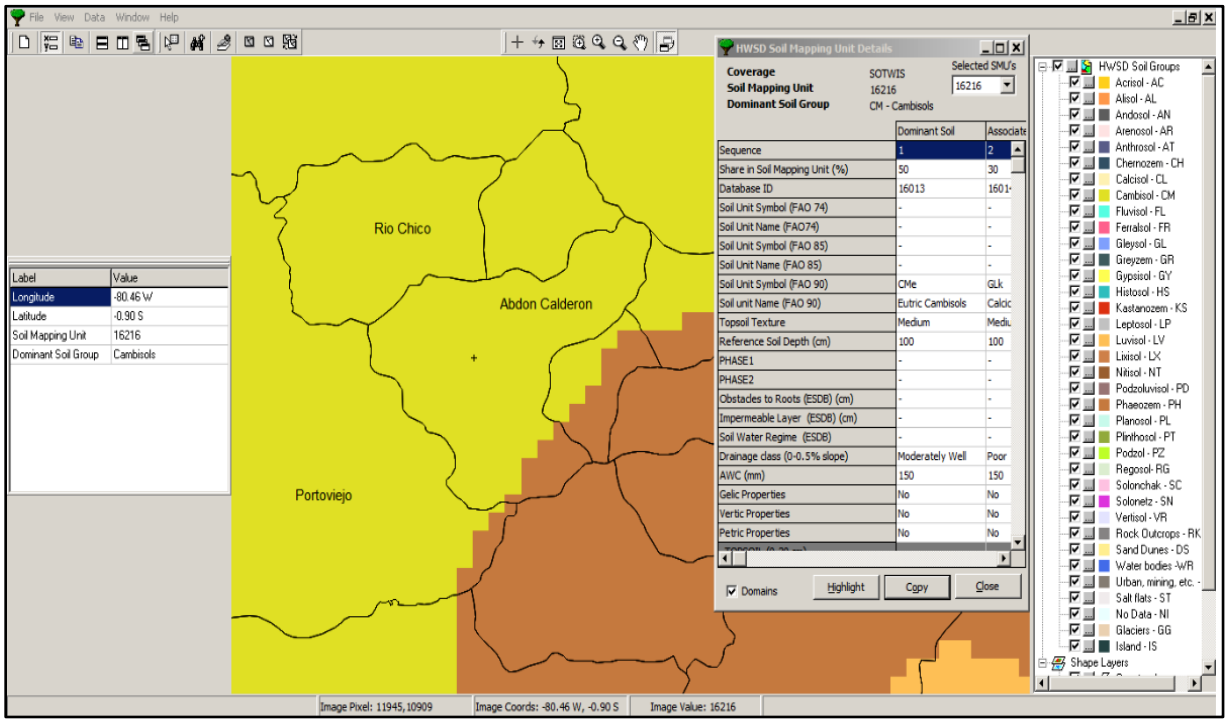


Figure 23: Screenshot of the HWSD viewer depicting the dominant soil type associated (cambisols) with a georeferenced positive leptospire site in Abdon Calderon, Portoviejo, Manabí, Ecuador

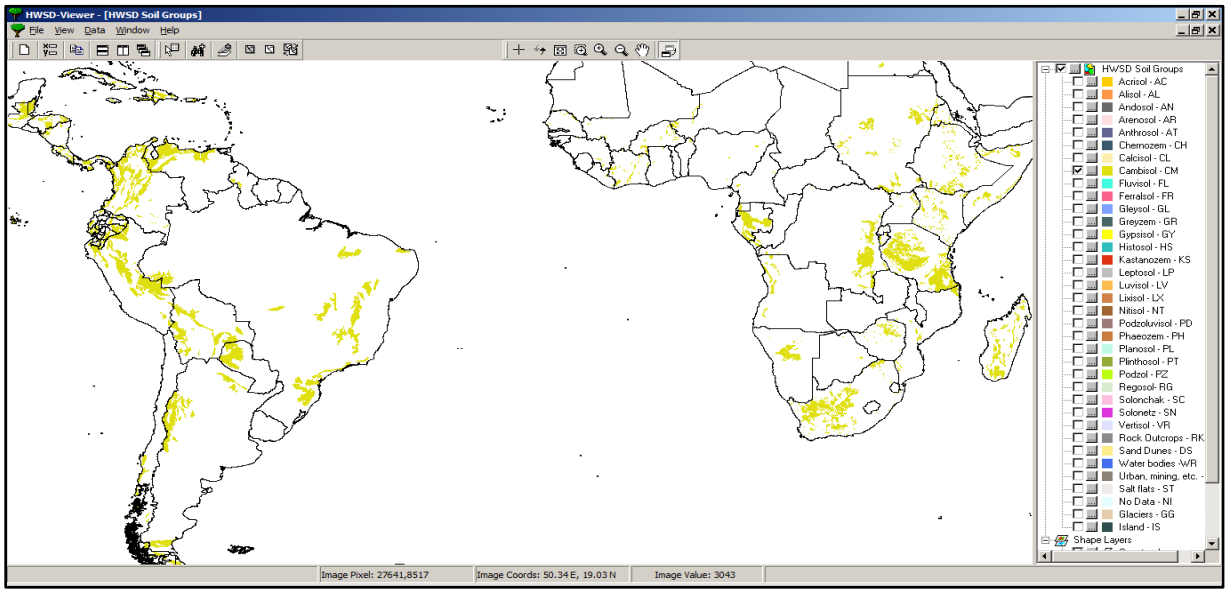
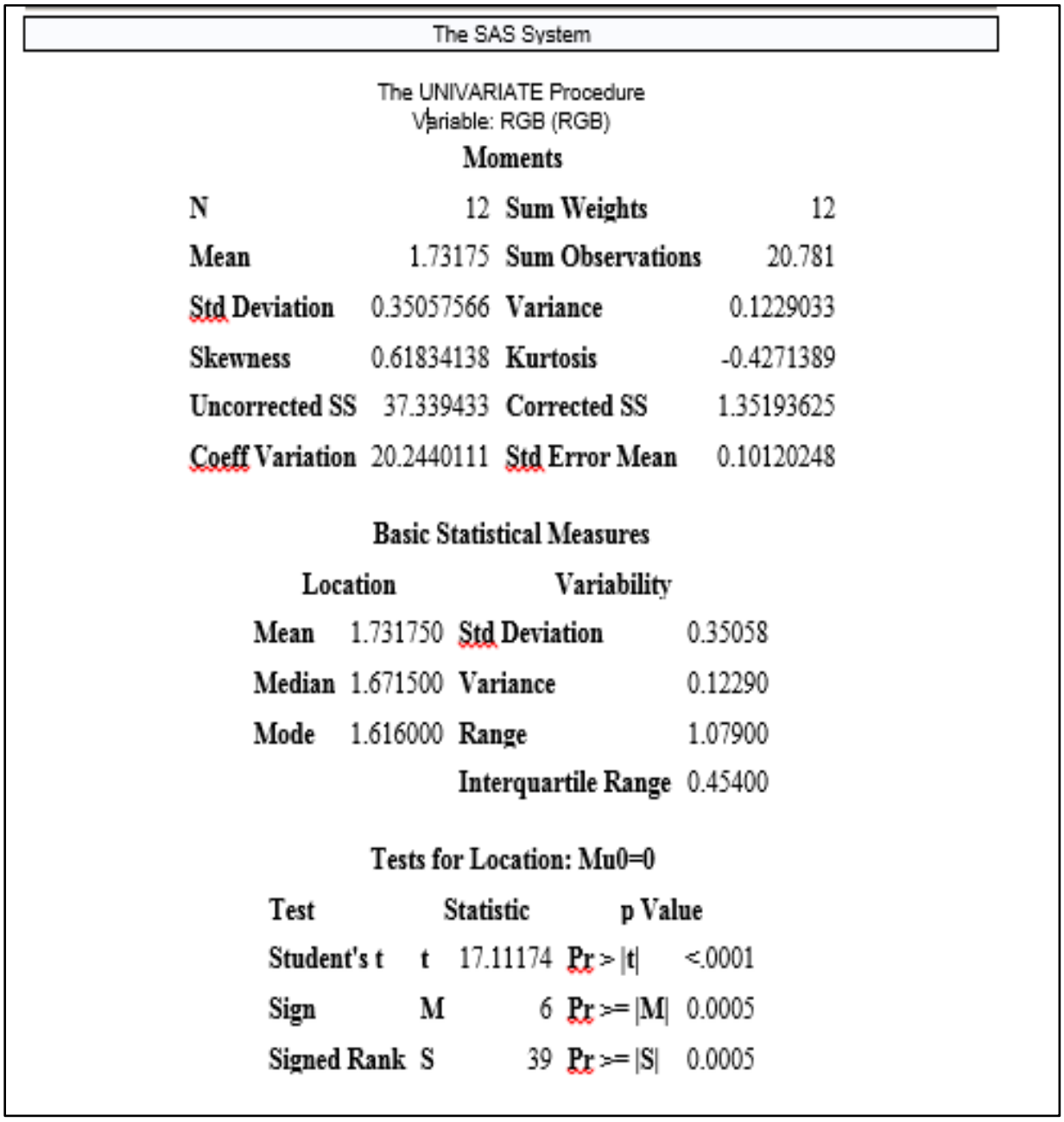


Figure 24: Screenshot of the HWSD viewer depicting distribution of cambisol soil group in Latin America and Africa

Figure 25: ARC GIS Online map of intermediate leptospire positive soil sites in Portoviejo, Manabí, Ecuador. Positive soil samples are labeled Q5 and Q6 on map. The inlay photos show the presence of cattle at the site, and the sandy/clay texture of the soil is classified by the HWSD as cambisol.

Figure 26: ARC GIS Online map of site of positive intermediate leptospire soil sample, Portoviejo, Manabí, Ecuador. The inlay photos depict the site, and the sandy/clay texture of the soil is classified by the HWSD as cambisol.



The SAS System

The UNIVARIATE Procedure
Variable: RGB (RGB)

Moments

N	12	Sum Weights	12
Mean	1.73175	Sum Observations	20.781
Std Deviation	0.35057566	Variance	0.1229033
Skewness	0.61834138	Kurtosis	-0.4271389
Uncorrected SS	37.339433	Corrected SS	1.35193625
Coeff Variation	20.2440111	Std Error Mean	0.10120248

Basic Statistical Measures

Location		Variability	
Mean	1.731750	Std Deviation	0.35058
Median	1.671500	Variance	0.12290
Mode	1.616000	Range	1.07900
		Interquartile Range	0.45400

Tests for Location: Mu0=0

Test		Statistic	p Value	
Student's t	t	17.11174	Pr > |t|	<.0001
Sign	M	6	Pr >= |M|	0.0005
Signed Rank	S	39	Pr >= |S|	0.0005

Figure 27: Univariate analysis for dependent variable RGB (SAS 9.4)

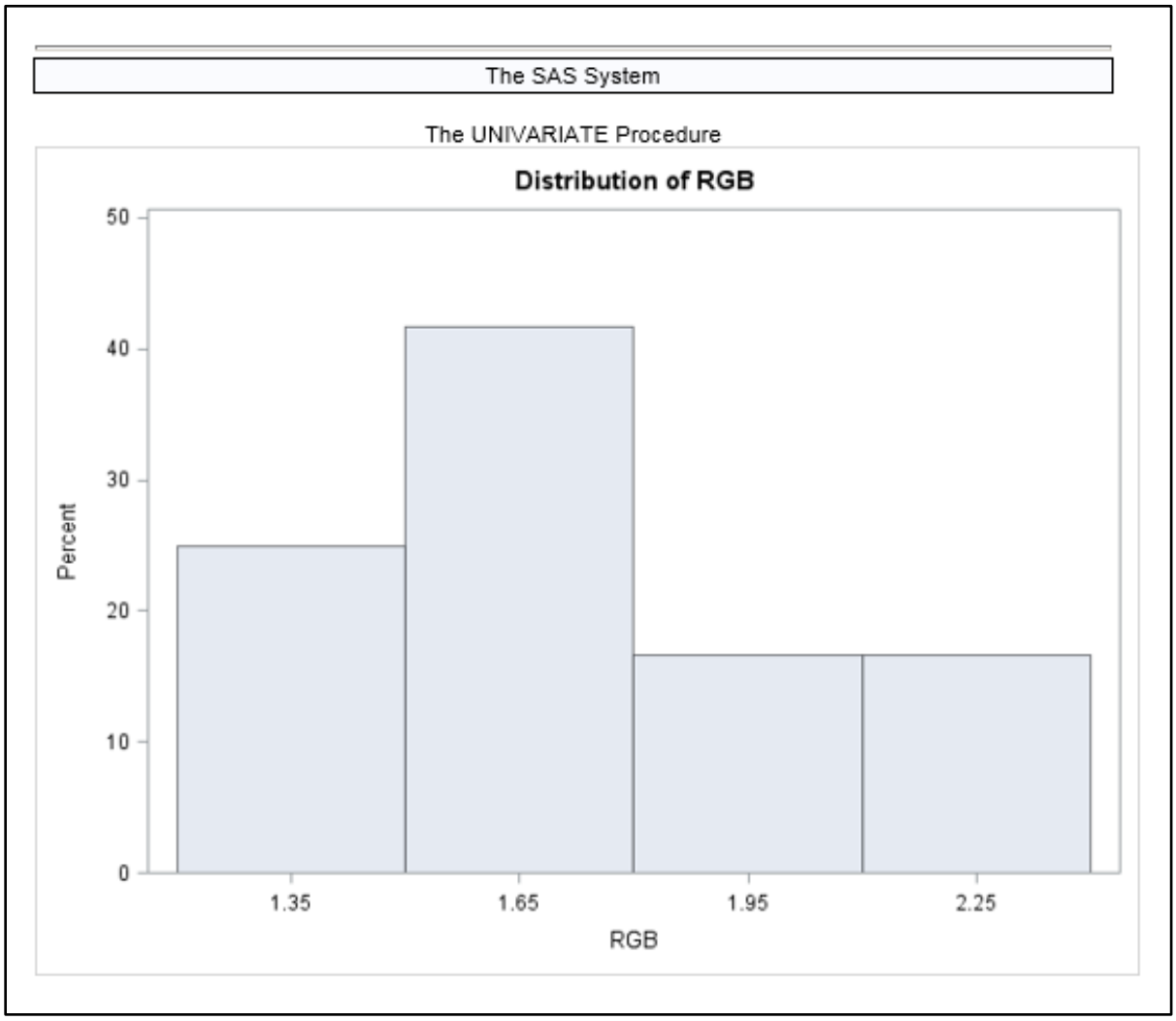


Figure 28: Histogram showing distribution dependent variable RGB (SAS 9.4)

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: RGB RGB

Number of Observations Read 12

Number of Observations Used 12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.56814	0.14203	1.27	0.3666
Error	7	0.78380	0.11197		
Corrected Total	11	1.35194			

Root MSE	0.33462	R-Square	0.4202
Dependent Mean	1.73175	Adj R-Sq	0.0890
Coeff Var	19.32268		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	2.85508	0.80512	3.55	0.0094
SAVI	SAVI	1	-0.18613	0.41889	-0.44	0.6702
NDVI	NDVI	1	0.70943	0.96230	0.74	0.4850
Elevation	Elevation	1	-0.02213	0.01369	-1.62	0.1500
riverdistance	riverdistance	1	-0.27588	0.26529	-1.04	0.3330

Figure 29: Simple linear regression results for dependent variable RGB (SAS 9.4)

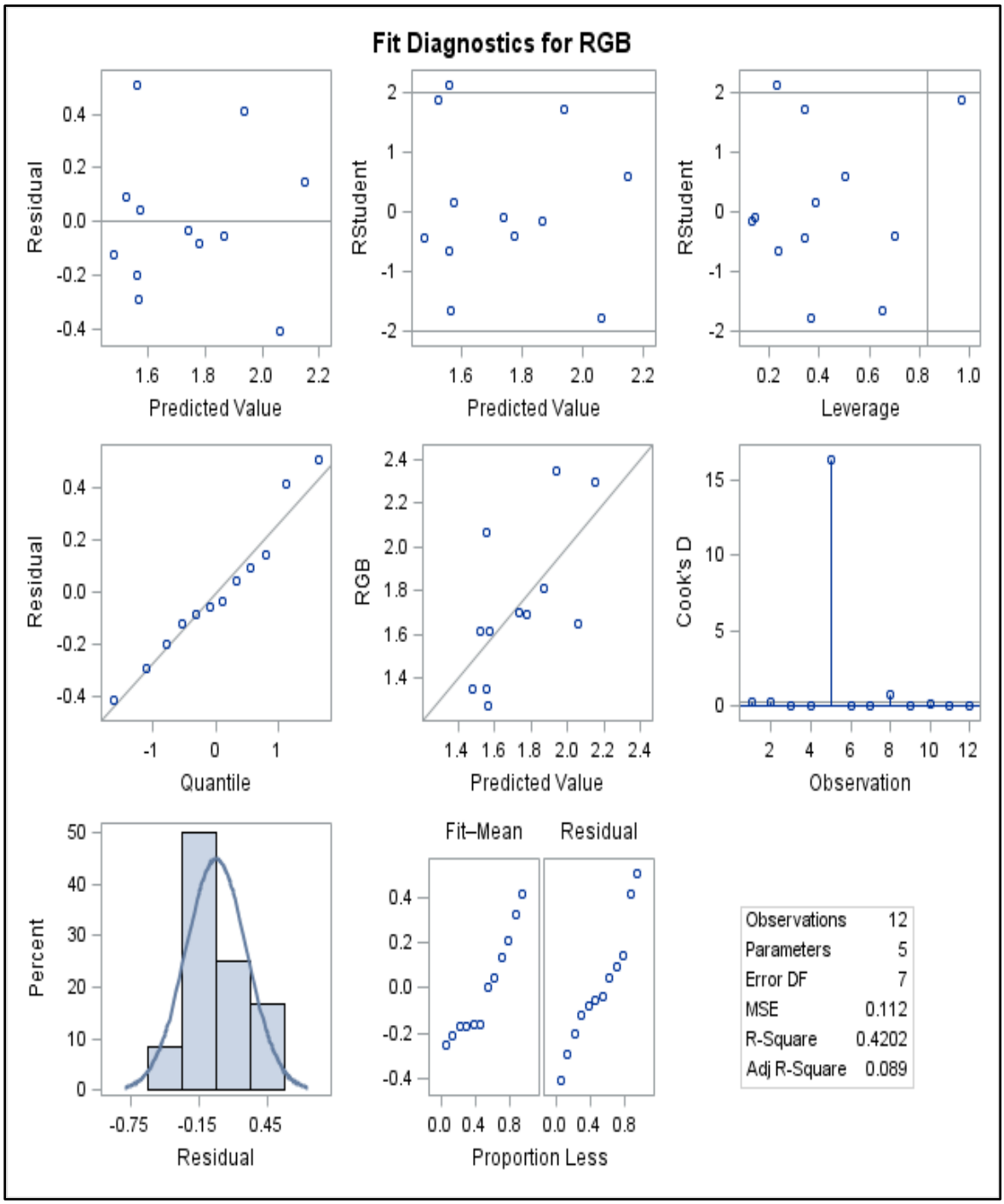


Figure 30: Simple linear regression fit diagnostics for dependent variable RGB (SAS 9.4)

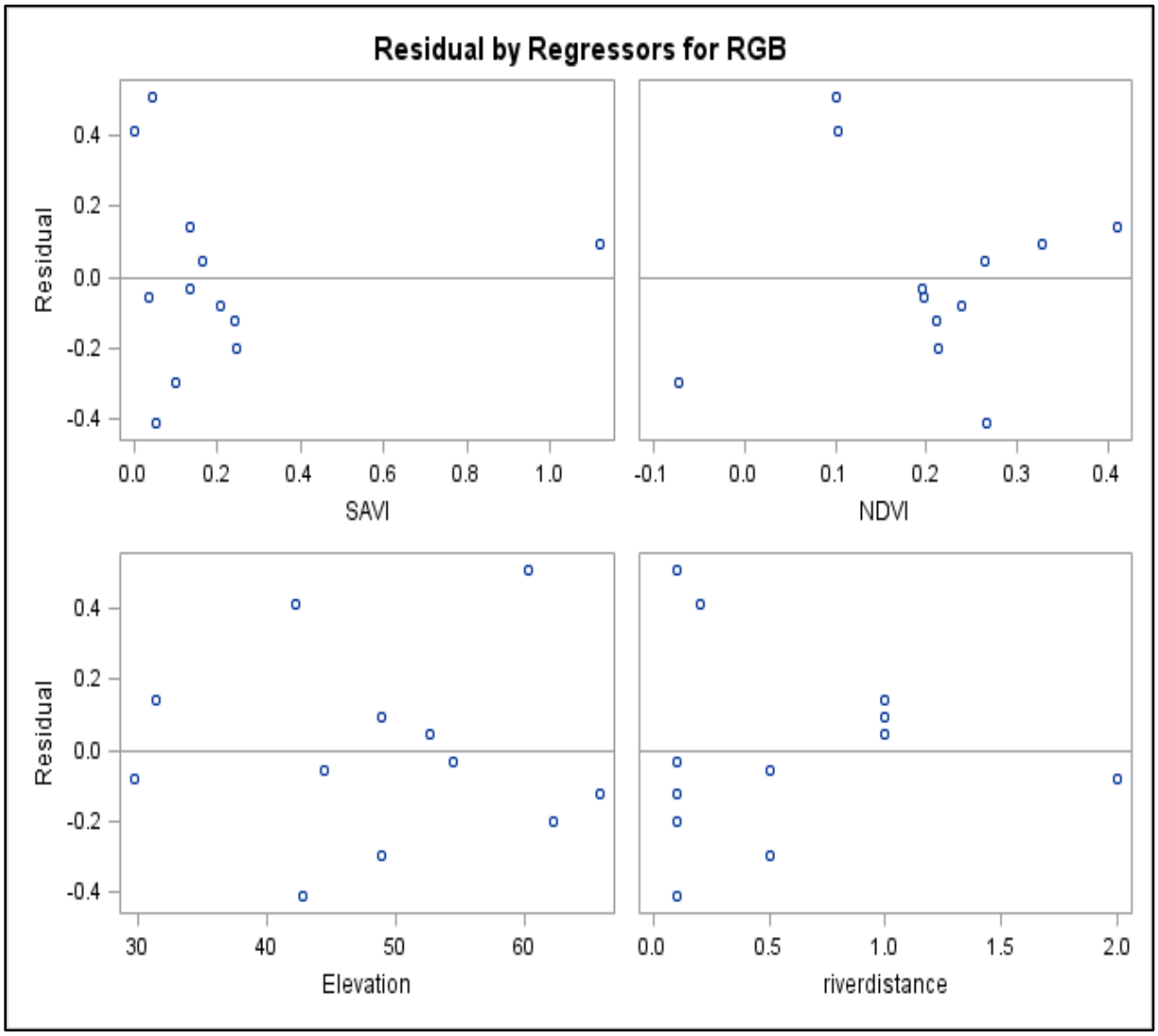


Figure 31: Simple linear regression residual by regressors for dependent variable RGB (SAS 9.4)

The SAS System

The GENMOD Procedure

Model Information

Data Set	WORK.RGB
Distribution	Normal
Link Function	Identity
Dependent Variable	RGB
Number of Observations Read	12
Number of Observations Used	12

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	7	0.7838	0.1120
Scaled Deviance	7	12.0000	1.7143
Pearson Chi-Square	7	0.7838	0.1120
Scaled Pearson X2	7	12.0000	1.7143
AIC (smaller is better)		13.3124	
AICC (smaller is better)		30.1124	
BIC (smaller is better)		16.2218	

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	2.8551	0.6149	1.6499	4.0603	21.56	<.0001
SAVI	1	-0.1861	0.3199	-0.8132	0.4409	0.34	0.5607
NDVI	1	0.7094	0.7350	-0.7311	2.1499	0.93	0.3344
Elevation	1	-0.0221	0.0105	-0.0426	-0.0016	4.48	0.0343
riverdistance	1	-0.2759	0.2026	-0.6730	0.1212	1.85	0.1733
Scale	1	0.2556	0.0522	0.1713	0.3813		

Note: The scale parameter was estimated by maximum likelihood.

Figure 32: Generalized Linear Model (GLM) regression analysis for variable RGB (SAS 9.4)

The GENMOD Procedure

Model Information	
Data Set	WORK.RGB
Distribution	Poisson
Link Function	Log
Dependent Variable	rgb

Number of Observations Read	12
Number of Observations Used	12

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance	7	0.2845	0.0406
Scaled Deviance	7	0.2845	0.0406
Pearson Chi-Square	7	0.2911	0.0416
Scaled Pearson X2	7	0.2911	0.0416
Log Likelihood		-9.1298	
Full Log Likelihood		-14.9465	
AIC (smaller is better)		39.8930	
AICC (smaller is better)		49.8930	
BIC (smaller is better)		42.3175	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	1.6175	1.7522	-1.8168	5.0518	0.85	0.3560
savi	1	-0.4030	1.0646	-2.4895	1.6836	0.14	0.7050
ndvi	1	-0.0176	2.3866	-4.6953	4.6600	0.00	0.9941
elev	1	-0.0163	0.0276	-0.0705	0.0378	0.35	0.5546
river	1	-0.2079	0.5504	-1.2867	0.8709	0.14	0.7057
Scale	0	1.0000	0.0000	1.0000	1.0000		

Figure 33: Poisson regression analysis for dependent variable RGB (SAS 9.4)

The SAS System

The GENMOD Procedure

Model Information

Data Set	WORK.RGB
Distribution	Negative Binomial
Link Function	Log
Dependent Variable	RGB RGB

Number of Observations Read	12
Number of Observations Used	12

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	7	0.4429	0.0633
Scaled Deviance	7	0.4429	0.0633
Pearson Chi-Square	7	0.4522	0.0646
Scaled Pearson X2	7	0.4522	0.0646
Log Likelihood		-9.2091	
Full Log Likelihood		-15.0257	
AIC (smaller is better)		42.0514	
AICC (smaller is better)		58.8514	
BIC (smaller is better)		44.9609	

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	1.1771	1.8338	-2.4171	4.7712	0.41	0.5210
SAVI	1	-0.1076	1.0034	-2.0742	1.8590	0.01	0.9146
NDVI	1	0.3718	2.1902	-3.9210	4.6646	0.03	0.8652
Elevation	1	-0.0124	0.0313	-0.0739	0.0490	0.16	0.6915
riverdistance	1	-0.1514	0.5933	-1.3143	1.0114	0.07	0.7985
Dispersion	0	0.0000	0.0000	0.0000	0.0000		

Note: The negative binomial dispersion parameter was estimated by maximum likelihood.

Figure 34: Negative binomial regression analysis for dependent variable RGB (SAS 9.4)

REFERENCES

1. Smythe LD et al. (2002). A Quantitative PCR (TaqMan) Assay for Pathogenic *Leptospira spp. BMC Infectious Diseases*, 2:13.

2. Cerqueira GM and M Picardeau. (2009). A century of *Leptospira* strain typing, *Infection*, *Genetics and Evolution*, 9, 760-768.

3. Barragan et al. (2011). Interactions of *Leptospira* with Environmental Bacteria from Surface Water. *Curr Microbiol*, 62, 1802–1806.

4. Guerrant, RL, Walker DH, and PF Weller. (2011). Tropical Infectious Diseases: Principles, Pathogens, and Practice-3rd Edition. China: Saunders.

5. Perez JBF, Becam J, Mauron C, and C Goarant. (2011). Rodent Abundance Dynamics and Leptospirosis Carriage in an Area of Hyper-Endemnicity in New Caldonia. *PLoS Negl Trop Dis*, 5(10): e1361.

6. Matthias et al. (2005). Diversity of Bat-Associated *Leptospira* in the Peruvian Amazon inferred by Bayesian Phylogenetic Analysis of 16S Ribosomal DNA Sequencing. *Am. J. Trop. Med. Hyg*., 73(5), 964-974.

7. Ullmann LS and H Langoni. (2011). Interactions between environment, wild animals and human leptospirosis. *The Journal of Venomous Animals and Toxins including Tropical Diseases*, 17(2), 119-129.

8. Jeffery S and WH Van de Putten. (2011). Soil Borne Human Diseases. Retrieved from the following website on March 3, 2014: http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/22432/2/lbna24893enn.pdf

9. Ristow et al. (2008). Biofilm formation by saprophytic and pathogenic leptospires. M*icrobiology*, 154, 1309-1317.

10. Tilahun, Z., Reta D., and K. Simenew. (2013). Global Epidemiological Overview of Leptospirosis*. International Journal of Microbiological Research*, 4(1), 9-15.

11. Coelho Micheline SZS and E Massad. (2012). The impact of climate on leptospirosis in Sao Paulo, Brazil. *Int J Biometerol*, 56, 233-241.

12. World Health Organization (2010). Report of the First Meeting of the Leptospirosis Burden Epidemiology Reference Group. Retrieved from the following website: http://whqlibdoc.who.int/publications/2010/9789241599894_eng.pdf

13. Few R, Ahern M, Matthies F, and S Kovats. (2004). Floods, health and climate change: a strategic review. Tyndall Centre for Climate Change Research, Working Paper 63.

14. Guerra MA. (2013). Leptospirosis: Public health perspectives. *Biologicals*, 41, 295-297.
15. Izurieta R, Galwankar S, and A Clem. (2008). Leptospirosis: The "mysterious" mimic. *J Emerg Trauma Shock*, 1(1), 21-33.

16. Watson JT, Gayer M, and MA Connolly. (2007). Epidemics after Natural Disasters. *Emerging Infectious Diseases*, 13(1).
17. Trueba et al. (2004). Cell aggregation: a mechanism of pathogenic Leptospira to survive in fresh water. *International Microbiology*, 35–40.

18. Desvars et al. (2011). Seasonality of Human Leptospirosis in Reunion Island (Indian Ocean) and Its Association with Meterological Data. *PLoS ONE*, 6(5):e20377.

19. Gamboa AA, Vasco L, Espinel M, Coloma J, and Gabriel Trueba. (2013). Difficulties in the differential diagnosis of dengue and leptospirosis in Guayaquil. *Avances*, 5(1), B10-B13.

20. Levett PN, Branch SL, and CN Edwards. (2000). Detection of Dengue Infection in Patients Investigated for Leptospirosis in Barbados. Am J Trop Med Hyg, 62(1), 112-114.

21. Marinho, Marcia. (2012). Leptospirosis: Epidemiologic Factors, Pathophysiological and Immunopathogenic. Current Topics in Tropical Medicine, Dr. Alfonso Rodriguez-Morales (Ed.). Retrieved from: http://www.intechopen.com/books/current-topics-in-tropicalmedicine/leptospirosisepidemiologic-factors-pathophysiological-and-immunopathogenic

22. González, Marta, et al. (2004). Vax-Spiral ®. Trivalent Antileptospirosis Vaccine For Human Use; Research, Development And Impact On The Disease In Cuba. http://www.medicc.org/publications/medicc_review/1004/pages/cuban_medical_literature2.html

23. Bracho et al. (2010). Large-scale application of highly-diluted bacteria for Leptospirosis epidemic control. *Homeopathy*, 99, 156-166.

24. Hartskeerl RA, Collares-Pereira M, and WA Ellis. (2011). Emergence, control and re-emerging leptospirosis: dynamics of infection in the changing world. *Clinical Microbiology and Infection*, 17(4).

25. Picardeau et al. (2014). Rapid tests for diagnosis of leptospirosis: Current tools and emerging technologies. *Diagnostic Microbiology and Infectious Disease*, 78, 1-8.

26. Wuthiekanun V et al. (2013). Rapid Isolation and Susceptibility Testing of Leptospira spp. Using New Solid Medium, LVW Agar. *Antimicrobial Agents and Chemotherapy*, 57(1), 297-302.

27. Benacer D, Who PY, Zain SNM, Amran F, and KL Thong. (2013). Pathogenic and Saprophytic *Leptospira* Species in Water and Soils from Selected Urban Sites in Peninsular Malaysia. *Microbe Environ*., 28(1), 135-140.

28. Dongyou L. (2011). Molecular Detection of Human Bacterial Pathogens. *Leptospira*. Retrieved from the following website: http://books.google.com/books?id=nnGhc44bypAC&pg=PA1187&lpg=PA1187&dq=Molecular+Detection+of+Human+Bacterial+Pathogens+leptospirosis+pcr&source=bl&ots=gL6XEhyYf9&sig=v5l2jHIBzGH56zKA0Hlna0CD3os&hl=en&sa=X&ei=a1uGU4aSBYKMqgaBoDgCw&ved=0CFEQ6AEwBA#v=onepage&q=Molecular%20Detection%20of%20Human%20Bacterial%20Pathogens%20leptospirosis%20pcr&f=false

29. Villumsen et al. (2012). Novel TaqMan® PCR for detection of *Leptospira* species in urine and blood: Pit-falls of in silico validation. *Journal of Microbiological Methods,* 91:184-190.

30. Chirathaworn C, Inwattana R, Poovorawan Y, and D Suwancharaoen. (2014). Interpretation of microscopic agglutination test for leptospirosis diagnosis and seroprevalence. *Asia Pacific Journal of Tropical Biomedicine*, 4(1):S162-S164.

31. Fonseca CA et al. (2006). *Lepospira* DNA detection for the diagnosis of human leptospirosis. *Journal of Infection,* 52, 15-22.

32. Goris et al. (2013). Prospective Evaluation of Three Rapid Diagnostic Tests for Diagnosis of Human Leptospirosis. *PLoS Negl Trop Dis*., 7(7): e2290.

33. World Health Organization (WHO). (2003). Human Leptospirosis: Guidance for Diagnosis, Surveillance, and Control. Retrieved from the following website: http://whqlibdoc.who.int/hq/2003/WHO_CDS_CSR_EPH_2002.23.pdf

34. Florence et al. (2009). Leptospirosis in the Asia Pacific region. *BMC Infectious Diseases*, 9:147.

35. Ecuador Ministry of Health (2014). Retrieved from the following website: http://www.salud.gob.ec/

36. World Bank GFDRR. (2010). Disaster Risk Management in Latin America and the Caribbean Region: global facility for disaster reduction and recovery (GFDRR). Country Notes Ecuador. Retrieved from the following website: http://www.gfdrr.org/sites/gfdrr.org/files/Ecuador-2010.pdf

37. Coppock JT and DW Rhind. nd. The History of GIS. Retrieved from the following website: http://www.grossmont.edu/judd.curran/History_of_GIS.pdf

38. YangT-C, Shoff C, and AJ Noah. (2013). Spatialising health research: what we know and where we are heading. *Geospatial Health*, 7(2), 161-168.

39. Clarke KC, McLafferty SL, and BJ Tempalski. (2010). On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Diseases*, Vol. 2, No.2, 85-92.

40. LandInfo. (2015). GeoEye-1 High-Resolution Satellite Imagery. Retrieved from the following website: http://www.landinfo.com/geo.htm

41. Woolhouse M. (2011). How to make predictions about future infectious disease risks. *Phil. Trans. R. Soc. B*, 366, 2045-2054.

42. Murayama Y and RC Estoque. Fundamentals of Geographic Information System. University of Tsukuba. 28 June 2010. Lecture.

43. Liang S, Schaepman M, and M Kneubuhler. (2008). Advances in Photogrammerty,Remote Sensing, and Spatial Information Sciences. Li, Chen, & Baltsavias (Ed.). London: Taylor and Francis Group.

44. Rogge et al. (2007). Integration of spatial–spectral information for the improved extraction of endmembers. *Remote Sensing of Environment,* 110, 287–303.

45. Costa et al. (2012). Surveillance for leptospirosis in the Americas, 1996–2005: a review of data from ministries of health *Rev Panam Salud Publica,* 32(3):169–77.

46. Kingscote BF. (1970). Correlation of Bedrock Type with the Geography of Leptospirosis. *Can. J. comp. Med*., Volume 34, 31-37.

47. Schneider et al. (2012). Leptospirosis Outbreaks in Nicaragua: Identifying Critical Areas and Exploring Drivers for Evidence-Based Planning. *Int. J. Environ. Res. Public Health*, 9, 3383-3910.

48. Hartsig JJ. (2011). The Effects of Land-Use Change on the Hydrological Properties of Andisols in the Ecuadorian Paramo. Master's Theses University of Tennessee. Retrieved from the following website: http://trace.tennessee.edu/utk_gradthes/978

49. Pappas, et al. (2008). The globalization of leptospirosis: worldwide incidence trends. *International Journal of Infectious Diseases*, 12:351—357.

50. Kitron U, Bouseman JK, and CJ Jones. (1991). Use of the ARCINFO GIS to study the distribution of Lyme disease ticks in an Illinois county. *Preventive Veterinary Medicine,* 11, 243-248.

51. Jacob et al. (2005). Evaluation of Environmental Data for Identification of Anopheles (Diptera: Culicidae) Aquatic Larval Habitats in Kisumu and Malindi, Kenya.

52. Wimberly MC, Chuag T-M, Henebry GM, Liu Y, Midekisa A, Semuniguse P, and G Senay. (2012). A Computer System for Forecasting Malaria Epidemic Risk Using Remotely-Sensed Environmental Data. International Congress on Environmental Modelling and Software. Proceedings.

53. Buczak AL, Baugher B, Babin SM, Ramac-Thomas LC, Guven E, et al. (2014). Prediction of High Incidence of Dengue in the Philippines. *PLoS Negl Trop Dis* 8(4):e2771.

54. DeGroote JP, Sugumaran R, Brend SM, Tucker BJ, and LC Bartholomay. (2008). Landscape, demographic, entomological, and climatic associations with human disease incidence of West Nile virus in the state of Iowa, USA. *International Journal of Health Geographics* 7:19.

55. Jacob et al. (2010). Developing GIS-based eastern equine encephalitis vector-host models in Tuskegee Alabama. *International Journal of Health Geographics* 9:12.

56. Coutinho CFS, Souza-Santos R, and MM Lima. (2012). Combining geospatial analysis exploratory study of triatome ecology to evaluate the risk of Chagas disease in a rural locality. *Acta Tropica* 121, 30-33.

57. Wood BL, Washino RK, Hibbard KA and JS Salvte. (1992). Estimating high mosquito-producing rice fields using spectral and spatial data. *International Journal of Remote Sensing*, 13:15, 2813-2826.

58. Young SG, Tullis JA and Jackson Cothruen. (2013). A remote sensing and GIS-assisted landscape epidemiology approach to West Nile Virus. *Applied Geography*. 45, 241-249.

59. Jacob et al. (2013). Validation of a Remote Sensing Model to identify *Simulium damnosum s.l*. Breeding Sites in Sub-Saharan Africa. *PLoS Negl Trop Dis*., 7(7): e2342

60. Jacob et al. (2013b). Unbiasing a Stochastic Endmember Interpolator Using ENVI Object-Based Classifiers and Boolean Statistics for Forecasting Canopied *Simulium damnosum s.l*. Larval Habitats in Burkina Faso. *J Geophys Remote Sensing*, 2:109.

61. Food and Agriculture Organization of the United Nations (FAO) Harmonized World Soil Database (HWSD). (2015). Retrieved from the following website: http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/

62. Genu AM, Roberts D, and JAM Dematto. (2013). The use of multiple endmember spectral mixture analysis (MESMA) for the mapping of soil attributes using Aster Imagery. *Acta Scientiarium Agronomy*, 35(3), 377-386.

63. Remote Sensing and Soil. (nd). University of Calgary, California. Course material. Retrieved from: http://www.ucalgary.ca/GEOG/Virtual/Remote%20Sensing/rssoils.htmlv

64. European Organization for the Exploitation of Meteorological Satellites (EUMOSAT). (2015). Retrieved from the following website: http://www.eumetsat.int/website/home/Data/Products/Land/index.html

65. Bartalis, Z., Naeimi, V., Hasenauer, S., Wagner, W. (2008). ASCAT Soil Moisture Product Handbook. ASCAT Soil Moisture Report Series, No. 15, Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Austria.

66. Jacob BG, Mwangangi JM, Mbogo CM, and RJ Novak. (2011). A Taxonomy of Unmixing Algorithms Using Li-Strahler Geometric Optical and Other Spectral EndmemberExtraction Techniques for Decomposing a QuickBird Visible and Near Infra-Red Pixel of an Anopheles arabiensis Habitat. *The Open Remote Sensing Journal*, 4, 1-25.

67. Li M, Zang S, Zhang B, Shanshan L, and C Wu. (2014). A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information. *European Journal of Remote Sensing*, 47: 389-411.

68. University of North Carolina. (nd). Li-Strahler Model. Retrieved from http://webcache.googleusercontent.com/search?q=cache:G6FnezxRR54J:www.unc.edu/courses/2008spring/geog/577/001/www/Li-Strahler.ppt+&cd=1&hl=en&ct=clnk&gl=us

69. Zeng et al. (2008). Comparison of the inversion of two canopy reflectance models for mapping forest crown closure using imaging spectroscopy. *Can. J. Remote Sensing,* 34(3), 235-244.

70. Deng C and C Wu. (2008). A spatially adaptive spectral mixture analysis for mapping subpixel urban impervious surface distribution. *Remote Sensing of Environment*, 133, 62–70.

71. Myint SW and GS Okin. (2010). Modeling Land-Cover Types Using Multiple Endmember Spectral Mixture Analysis in a Desert City. Arizona State University. Working Paper Number 2010-06.

72. CIA World Factbook. (2014). Country Profile: Ecuador. Retrieved from the following website: http://www.ciaworldfactbook.us/south-america/ecuador.html

73. Manabi (2014). Retrieved from the following website: http://www.manabi.gob.ec

74. Henry RA and RC Johnson. (1978). Distribution of the Genus *Leptospira* in Soil and Water. *Applied and Environmental Microbiology*, 35(3), 492-499.

75. Render B, Stair RM, and ME Hanna. (2006). Quantitative Analysis for Management. M. Pfaltzgraff (Ed.). Upper Saddle River, NJ: Pearson.

76. Ecuador Ministry of Health GeolSalud 2.0. (2015). Retrieved from the following website: https://aplicaciones.msp.gob.ec/salud/publico/dniscg/geosalud/gui/