Spatiotemporal Analysis of Dengue Hemorrhagic Fever and Dengue Shock Syndrome Incidence within Trinidad, West Indies

Sarah Hosein
Department of Geomatics Engineering and Land Management,
The University of the West Indies,
St. Augustine
Trinidad and Tobago
1(868)6622002 ext 83705
hosein.sarah@gmail.com

Raid Al-Tahir
Department of Geomatics Engineering and Land Management,
The University of the West Indies,
St. Augustine
Trinidad and Tobago
1(868)6622002
raid.altahir@sta.uwi.edu

Bheshem Ramlal
Department of Geomatics Engineering and Land Management,
The University of the West Indies,
St. Augustine
Trinidad and Tobago
1(868)6622002 ext 82108
bheshem.ramlal@sta.uwi.edu

ABSTRACT

During the 17th to early 20th centuries, malaria, yellow fever, dengue and other vector-borne diseases were responsible for more human disease and mortality than all other causes combined. Transmitted by the Aedes mosquito, particularly type *Aedes albopictus* and *Aedes aegypti*, dengue has become one of the world's major infectious diseases. Within the Caribbean, individuals can be especially susceptible to dengue due to the tropical climate; such cases can be mapped using spatial analytical tools.

This study focuses on the spatiotemporal behaviour of dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS), for the years 1998-2007, for the island of Trinidad. Geostatistical techniques as well as correlations with environmental and spatial factors were performed. The DHF and DSS datasets were normalized against population, followed by the exploration of geographic distributions and patterns for each year. Lastly, local and global statistical type analyses were tested for spatial dependencies and autocorrelations in order to determine the statistical significance of existing patterns in the data. The years that displayed statistically significant clustering, spatial stationarity and spatial dependency for the island of Trinidad were found to be: 1998, 1999, 2001, 2002 and 2003.

Categories and Subject Descriptors

I.5.3 [Pattern Recognition]: Clustering- *algorithms, similarity measures*

General Terms

Reliability, Theory, Verification

Keywords

Dengue, Geostatistics, Clustering, Spatial Stationarity, Spatial Dependency

Conference HEALTH GIS '13, November 5–8, 2013, Orlando, Florida USA.
Copyright: 2013 ACM ISBN 978-1-4503-2529-5…$15.00.

1. INTRODUCTION

Dengue, a vector-borne disease, has been ranked as the most significant arboviral disease worldwide. Contributing to 50 to 100 million cases and 12000 to 24000 deaths annually (World Health Organization 2002; Gibbons and Vaughn 2002), the virus has now been reported as an endemic in excess of 100 countries worldwide in locations lying within 35°N-35°S latitude, an area known as the 'dengue belt' (Chadee, Shivnauth, et al. 2007; Gubler 1998).

Spatial analytical approaches to health concerns are becoming increasingly more acceptable due to the spatiotemporal nature of health outbreaks. As a result, these outbreaks allows for an integration of visual and analytical applications in the realm of Geographic Information Systems (GIS) (Bertazzon, Olson and Knudtson 2009).

Contracted by humans and transmitted via the bite of an infected female *Aedes aegypti* or *Aedes albopictus* mosquito, the virus belongs to the *Flavivirus*, family *Flaviviridae* where there are 4 serotypes; DEN-1, DEN-2, DEN-3 and DEN-4 (Foster et al. 2003).

All serotypes can cause the different types of dengue viruses: dengue fever (DF) classed as a mild feverish illness, dengue hemorrhagic fever (DHF) or dengue shock syndrome (DSS). DF, despite a very serious illness, records support that the latter two types are significantly more serious and can be potentially fatal (Foster et al. 2003).

2. STUDY AREA

Dengue incidence and associated factors were analysed for the island of Trinidad (Figure 1), located between 10° 02' to 10° 50' N latitude and 60° 55' to 61° 56' W longitude. With an area of 5,128 sq km, it lies approximately 15 km off the north eastern coast of Venezuela, between the Caribbean Sea and the North Atlantic Ocean (Central Intelligence Agency 2012). The climate of the island is tropical with two seasons- a rainy season from June to December and a dry season from January to May. The temperature ranges from approximately 22°C to 30.5°C but can get hotter during the dry season and cooler during the rainy season (Hemme et al. 2009).

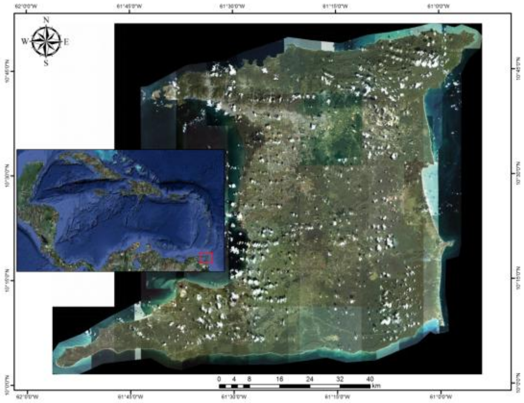

Figure 1: Location of Study Area: Trinidad, West Indies.

3. DATA ACQUISITION

This paper analysed the spatial patterns of DHF and DSS within Trinidad for the years 1998-2007 as shown in Figure 2. The datasets for the years 1998-2006 was obtained from The Ministry of Health, Trinidad and Tobago and provided to the University of the West Indies in a GIS ready shapefile format. The 2007 data was also obtained from the Ministry of Health however; this was in a tabular format. In order to convert the data to a GIS ready format, geocoding was carried out where the street address of each incidence was the parameter to which the geocoding was performed. Using the IKONOS 2007 satellite imagery for Trinidad, the buildings dataset and the road network for the island, a point dataset was created. This was done at a scale of 1:1,000. Figure 2 illustrates the distribution of all geocoded points of DHF and DSS used for this study.

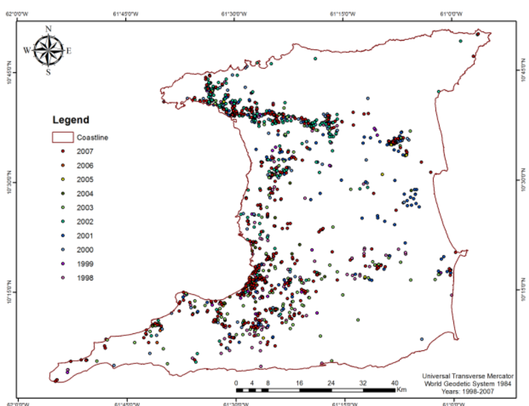

Figure 2: Distribution of DHF/DSS Point Locations

In addition to the datasets previously mentioned, the following are other datasets used in this study confined for the island of Trinidad: Coastline, Major Highways and Wards (Figure 3) that were all obtained from the Department of Geomatics Engineering and Land Management, the University of the West Indies; as well as Land Cover (Figure 4) (Jehu 2011).

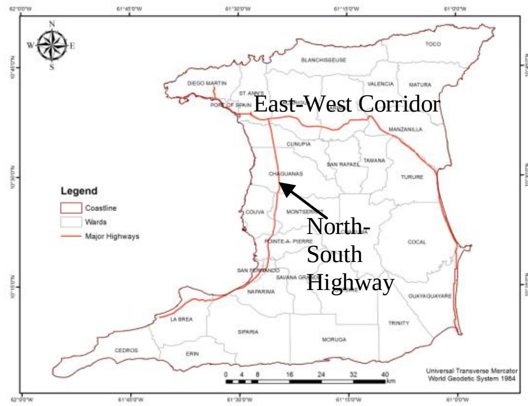

Figure 3: Datasets used: Major Highways, Coastline, Wards

3.1 Environmental Parameters

According to Arcari, Tapper and Pfueller (2007), rainfall is one of the principal climatic parameters to impact on the spatiotemporal patterns of dengue incidence. They also reported that temperature critically affects the intensity of dengue outbreak intensities based on studies carried out on dengue and DHF cases in Indonesia (Arcari, Tapper and Pfueller 2007).

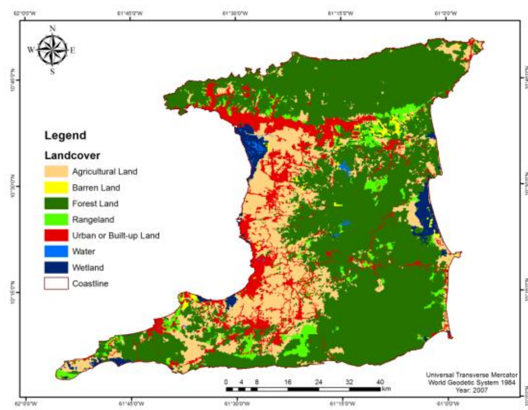

Figure 4: Trinidad Land Cover, 2007, Trinidad (Jehu 2011)

Epidemics of dengue in Caribbean regions have been thought to be in some correlation to similar climatic anomalies (Kovats, Bouma and Haines 1999) such as the El Nino Southern Oscillation (ENSO). It comprises of a warm El Nino phase where there is a warming of sea surface temperatures (SST), and a cold La Nina phase where there is a cooling effect that brings about anomalies in global temperature and precipitation (Stern and Easterling 1999). Since dengue is transmitted via the *Aedes aegypti* mosquito, it is likely to be impacted on by changes in climatic conditions and hence climatic anomalies related to the ENSO (Seghal 1997).

The El Nino and La Nina years were identified utilising the Oceanic Nino Index (ONI) (Kousky and Higgins 2007) from the Climate Prediction Centre website (NOAA 2012). The recorded El Nino events that encompass this study are as follows- 1997-1998, 2002-2003, 2004-2005 (weak) and 2006-2007 (weak). La Nina episodes were- latter part of 1998-2003, latter part of 2005-2006, latter part of 2007-2008 (NOAA 2012).

3.2 Rainfall

Rainfall data was obtained from the Water Resources Agency (Trinidad and Tobago) in a tabular format. The data was investigated and different rainfall stations were chosen such that they were located befittingly about the island and also provided that they had data for the entirety of the temporal span of the study.

For each year, the data was categorized by annual rainfall, rainy season and dry season. A surface was then created using the Inverse Distance Weighted tool (IDW). IDW was chosen since the surface interpolated does not generate values that exceed the highest input or is lower than the lowest value input (Fischer and Getis 2009).

3.3 Temperature

Temperature data for Trinidad via ground temperature stations was found to be quite limited with respect to the study area: the only available data accessible was retrieved from one temperature gauge at the Piarco International Airport, Piarco, Trinidad. Since the dataset focuses on the entire island of Trinidad, one temperature station was deemed insufficient and therefore based on a study conducted for Brazil by Roseghini et al. (2011) the Moderate Resolution Imaging Spectro radiometer (MODIS) - Land Surface Temperature (LST) was used at a 1km resolution. The data was obtained from NASA's Earth Observing System Data and Information System.

3.4 Software used

In order to carry out the necessary GIS spatial analysis, ESRI ArcGIS (version 10.0) and Crimestat (version 3.3) were utilized. All statistical representations and analyses were done in Minitab 16 Statistical Software.

4. PREVIOUS RESEARCH CONDUCTED

From a previous study conducted (Hosein and Al-Tahir 2011), geographic patters were investigated. The central features, directional distribution ellipses and kernel densities for each epoch of data were obtained (Figure 5) and the following conclusions can be made based on their annual behaviour.

For the years 1998- 2007, the mean centres demonstrates spatial pattern follows closely to the North-South Highway and mostly to the central and south western part of the island. This held similar for the annual analysis, the rainy season and for the dry season. It appears that the patterns begins closer to the south west in 1998 near the Ward of Pointe-a-Pierre, travel northwards to Chaguanas, plunge deep to the south as far as La Brea and end back near to the starting position in the general area of or next to Pointe-a-Pierre in 2007.

Kernel densities were generated from the point incidence data at a cell size of 1km for consistencies. It is very effective in visualising trends of spatial clustering over large areas. The kernel density analysis for the island displayed that the clusters of incidence followed closely to the East-West Corridor to the north. There were almost always clusters present in the Wards of Port-of-Spain and St. Ann's to the north west, Manzanilla to the east and in Naparima and San Fernando to the south (Hosein and Al-Tahir 2011) .

The mean centres were chosen to be mapped against each other to observe their directionality per year since according to Fischer and Getis (2009); this tool is useful in tracking distributional changes. The directional distribution ellipses were mapped in order to observe possible trends in the data by summarising the spatial characteristics of each dataset.

5. Spatiotemporal Analysis of Dengue Incidence in Trinidad

For this study, in order to analyse spatial and temporal behaviour of dengue incidence, a multitude of analyses was carried out. The DHF and DSS datasets were initially normalized against population obtained from the Wards dataset in order to get a true representation of annual outbreaks for the island of Trinidad.

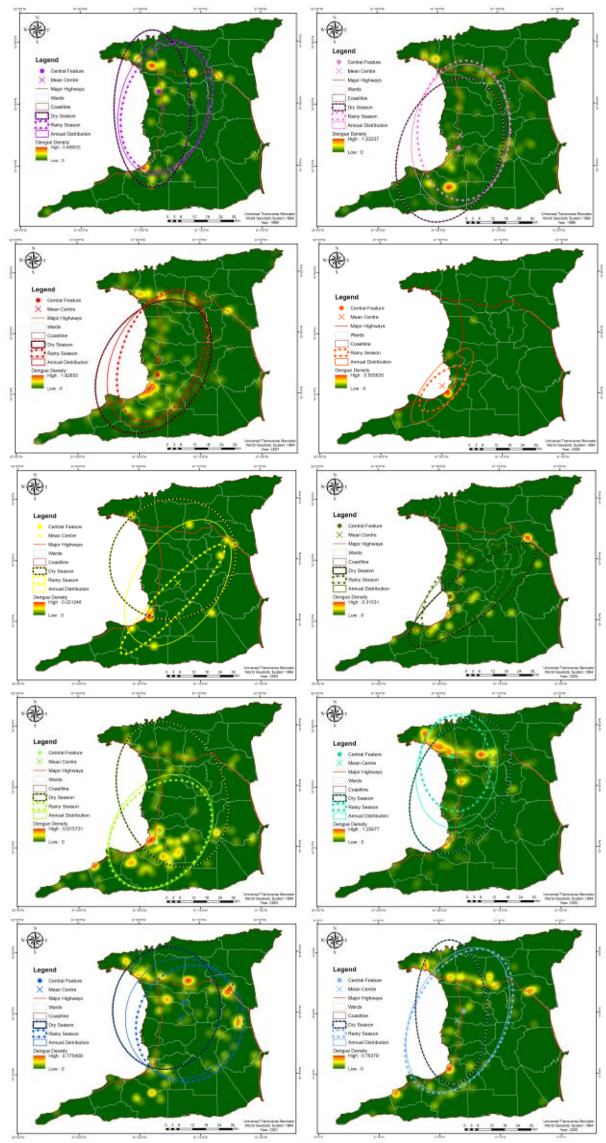

Figure 5: Temporal locations of Central Features, Directional Distribution Ellipses, Kernel Densities for DHF and DSS Cases in Trinidad

Local and global statistical type analysis that tested for spatial dependencies and autocorrelations for each epoch of data was performed. The directional distribution ellipses, done in 3 categories: annual, rainy season and dry season ellipses for each year were compared the north/south and east/west highway dataset (Figure 3) and to land cover derived from a 2007 dataset (Figure 4).

Statistical analyses were then performed. The Average Nearest Neighbour, High/Low Clustering (Getis-Ord General G), Multi-Distance Spatial Cluster Analysis (Ripley K), Spatial Autocorrelation (Moran's I), Cluster and Outlier Analysis (Aselin Local Moran's I), Hot Spot Analysis (Getis-Ord Gi*), and the Knox test were all performed on the datasets.

For some processes, the raw point data was sufficient, in other cases there was a requirement for an input numeric field to be evaluated statistically. In order to suffice, a 1km fishnet grid for Trinidad was generated and each year of dengue incidence data was spatially joined to give a count per km of dengue incidence. The grid is broken up into 5105 1km² grid blocks. When investigating the resolutions of all datasets used in this study, the resolution of the temperature data was the coarsest at 1km, therefore this was the chosen resolution for all datasets.

6. DENGUE INCIDENCE DATA

Table 1: Annual Number of Cases of DHF and DSS

Year	Number of Cases
1998	156
1999	48
2000	129
2001	158
2002	212
2003	135
2004	27
2005	9
2006	8
2007	382

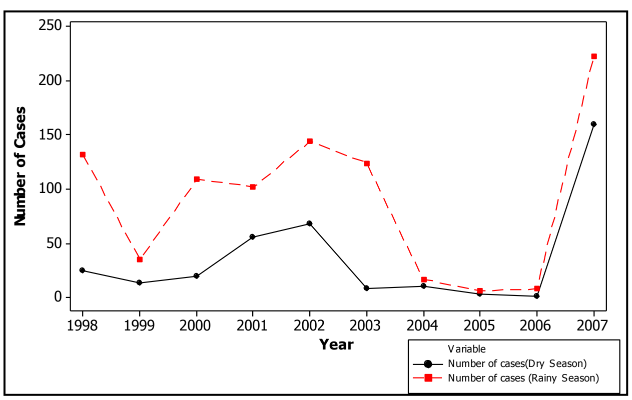

Figure 6: Seasonal DHF and DSS Cases

Table 1 shows the annual number of cases of DHF and DSS that were obtained; the data seems to show almost a cyclic pattern from peak to peak. The data starts out in 1998, an epidemic year, and illustrates the changes in the number of cases of DHF and DSS for all years until 2007 seasonally.

According to Hemme et al. (2009) the years 1998, 2002 and 2003 have been reported as epidemic years of DHF in Trinidad. This holds for the datasets utilized as seen in Table 1. The year 2007 is an obvious epidemic year also since it has the highest recorded number of cases for all datasets used in this study. Figure 6 shows the variation of the number of incidences per year seasonally and it is seen that the rainy season accounts for the greater number of cases annually. Most apparent is that as the number of cases increase for the rainy season, so does the cases for the dry season. The epidemic years corresponds with the rainy season well but also shows that the year 2000 had a high number of cases. When compared to the dry season it is observed that the epidemic years did hold as well. There are a high number of cases for the dry season for the years 1998, 2001, 2002 and 2007 but a very low number of cases for the year 2003.

7. CLUSTERING

Testing for clustering statistically involves the disproving of the null hypothesis which states that there is spatial randomness in the distribution of a given dataset. The presence of similar values found together is termed positive spatial autocorrelation whereas negative spatial autocorrelation is where dissimilar values are found being clustered. This can be explained most simply in the case of categorical data where one area is coloured white (W) and the other is coloured blue (B). Positive spatial autocorrelation would exist in areas where adjacent colours are similar i.e. BB or WW and negative spatial autocorrelation would be those areas of differing colours BW or WB (Haining 2004).

Statistical processes like Getis-Ord and the Moran's index generate a cross-product statistic where there is a measure of spatial proximity of locations with similarity of data values. The output is a statistical measure of spatial dependency within data (Haining 2004).

7.1 Global Calculators

A global calculation identifies broad overall patterns or trends. They work by comparing feature locations and/or attributes to a theoretical random distribution in order to determine if there is statistically significant clustering or dispersion (Haining 2004) and focuses on spatial dependency while disregarding non-stationarity (Bertazzon, Olson and Knudtson 2009).

7.1.1 Average Nearest Neighbour

The purpose of the nearest neighbour index is to compare distances between incidences that would be expected as the basis of chance (Fischer and Getis 2009).

Table 2: Annual Average Nearest Neighbour Summary

Year	Nearest Neighbour Ratio	z-score	p-value	Clustered
1998	0.565645	-10.411783	0.000000	Yes
1999	0.740687	-3.436967	0.000588	Yes
2000	0.619621	-8.232902	0.000000	Yes
2001	0.493007	-12.191619	0.000000	Yes
2002	0.531608	-13.046933	0.000000	Yes
2003	0.605424	-8.803006	0.000000	Yes
2004	0.661646	-3.363439	0.000770	Yes
2005	0.981087	-0.120004	0.904480	Random
2006	0.406426	-3.766190	0.000166	Yes
2007	0.430623	-21.289335	0.000000	Yes

The z-score gives an indication of the strength of clustering being displayed between the incidences. The negativity of the z-score indicates the strength of clustering whereas the positivity of the z-score indicates the dispersion of the data therefore if there is a z-score tending to zero; the data exhibits no spatial pattern and is random. A small p-value indicates clustering while a p-value tending to 1 indicates randomness (Fischer and Getis 2009).

From the data as seen in Table 2, the years 1998, 2002 and 2007 exhibits the strongest clustering since the z-scores and their corresponding p-values are < 0.01 for all years except 2005 which falls within the 99% confidence interval and is therefore statistically significant. For the year 2005, the data is random since the z-score is -0.120004 which is very close to zero and the p-value of 0.904480 is very close to 1.

However, the Nearest Neighbour Index is a ratio of the observed mean distance to the expected mean distance of the cases. An index less than 1 means that the pattern exhibits clustering whereas an index greater than 1 indicates a trend of dispersion (Fischer and Getis 2009). When observing the nearest neighbour indices for all the years, they are all positive which therefore means that the trend in the data is not clustered.

7.1.2 High/Low Clustering (Getis-Ord General G)

The Getis ord general G statistic for a specified study area gives a measure of the concentration of high or low data values. Positivity of the z-score indicates clustering of high values whereas negative z-scores are indicative of clustering of low values (Fischer and Getis 2009).

Table 3: Annual Getis-Ord General G Summary

Year	Observed General G	z-score	p-value	Clustered
1998	0.000010	18.225253	0.000000	High Clusters
1999	0.000021	8.498823	0.000000	High Clusters
2000	0.000009	12.739287	0.000000	High Clusters
2001	0.001149	9.568638	0.000000	High Clusters
2002	0.000008	18.164190	0.000000	High Clusters
2003	0.000006	10.097104	0.000000	High Clusters
2004	0.000021	5.548298	5.548298	High Clusters
2005	0.000026	1.817846	0.069088	High Clusters
2006	0.000158	16.578630	0.000000	High Clusters
2007	0.000008	26.920080	0.000000	High Clusters

From Table 3, the data for all years exhibits high clustering with the years 1998, 2002 and 2007 having the most clustering of high values. The associated p-values for all years except 2005 are statistically significant (>0.01) therefore the null hypothesis can be rejected for the General G process for the years 1998-2004 and 2006-2007.

7.1.3 Multi-Distance Spatial Cluster Analysis (Ripley's K Function)

Based on the Ripley's K-function, the multi-distance spatial cluster analysis allows for summarizing spatial dependencies on a dataset for a specified range of distances that is controlled by the user (ESRI 2012).

The distance range and the increments are Expected Ks. The average number of neighbouring features that are associated with each feature is computed. As distances change, the associated neighbouring features associated with each feature are likely to increase. For clustering to be present, the average concentration of features must be lower than the average neighbours for a specific evaluation distance (ESRI 2012).

The output of this tool gives the Expected K, Observed K, Lower Confidence Envelope and the Higher Confidence Envelope. For spatial clustering to be present, the observed K must be larger than the expected K. For statistically significant spatial clustering to take place, the observed K must be larger than the high confidence envelope. For spatial dispersion to take place, the converse of what is expected for clustering is to be observed i.e. the observed K will be less than the expected K and the observed K will be less than the low confidence.

The 2005 dataset illustrates that the observed K was greater than the expected K for each specified distance however only for the distances of 3500m, 4000m and 8500m was there statistically significant spatial clustering occurring.

Therefore, for the years 1998-2004 and 2006-2007 for the distance threshold specified there is statistically significant spatial clustering taking place and this leads to the fact that as distance continues to increase, there will be more spatially significant clustering taking place for the specified study area.

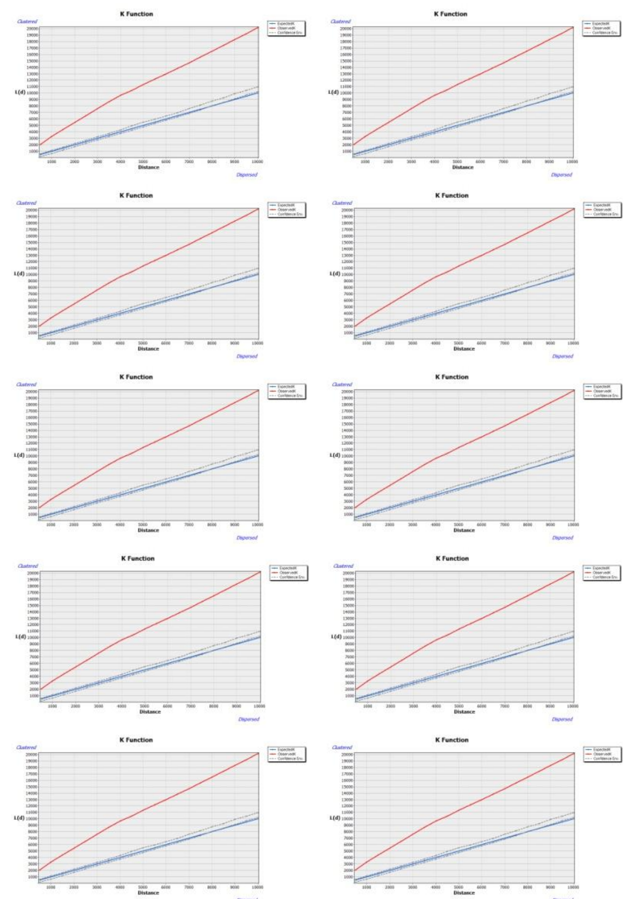

Figure 7: Multi-Distance Spatial Cluster Analysis Ripley's K Function

7.1.4 *Spatial Autocorrelation (Global Moran's I)*

When investigating the spatial autocorrelation process, its concept should be first understood. If a set of incident locations are arranged such that there is no spatial relationship between any of the incidence then this is termed spatial independency. Spatial autocorrelation is the opposite of spatial independence. It is such that the arrangements of the location of incident points are related to each other and spatial independency has been voided.

Clustering of events is referred to as positive spatial autocorrelation and the dispersion of events is referred to as negative spatial autocorrelation.

Spatial autocorrelation is based on Tobler's first law (Fischer and Getis 2009). It can be measured via a number of indices. This report investigates the Global Moran's I index on the dengue incidence. For the software used, the global Moran's I index is obtained per dataset via calculating the mean and variance. The means are then subtracted from each value to obtain a deviation from the mean. The deviations are then multiplied together to give what is known as a cross-product. A positive cross-product would be indicative of data whose neighbours are both larger than or both smaller than the mean. A negative cross-product would be obtained where one value is larger and the other is smaller than the mean. Therefore, spatial clusters that give rise to a positive Moran's I index occur where high cross-product values cluster near each other or low cross-product values cluster near each other. However, if the index is near zero, this is indicative of a positive value balancing out a negative value.

The output of the process for the datasets used in this study (Table 4) is consistent with the results of the average nearest neighbour and the Getis-Ord General G statistics. The p-values are <0.01 and the z-scores are all positive for all years except for 2005 where the p-value is greater than 0.01 and the z-score value is closest to 1 indicating clustering in all years with the least in 2005 and the strongest clustering for the years 1998, 2002 and 2007. The positivity of the Moran's I index indicates clustering of high cross-product values however since they are very close to zero, this can also indicate that there are clustering of low values near clusters of high values and they may have been some balancing out.

Table 4: Spatial Autocorrelation (Morans I) Summary

Year	Moran's Index	z-score	p-value	Clustered
1998	0.132346	18.234136	0.000000	Yes
999	0.027948	8.386368	0.000000	Yes
2000	0.090333	12.646987	0.000000	Yes
2001	0.092002	9.569654	0.000000	Yes
2002	0.182276	18.175151	0.000000	Yes
2003	0.072289	10.000999	0.000000	Yes
2004	0.016163	5.586030	0.000000	Yes
2005	0.002215	1.761637	0.078131	Yes
2006	0.041702	16.649707	0.000000	Yes
2007	0.195410	26.920750	0.000000	Yes

7.2 Local Calculators

Local calculations identify the extents and locations of clustering and indicate where spatial clustering occurs. It works by processing every feature with respect to neighbouring features in order to determine whether or not it represents a spatial outlier, or if it is part of a statistically significant cluster (Fischer and Getis 2009). Here, the focus is on non-stationarity while spatial dependency is disregarded (Bertazzon, Olson and Knudtson 2009).

7.2.1 *Cluster and Outlier Analysis (Anselin Local Moran's I)*

The cluster and outlier process discovers spatial clusters where the attributes of the data being investigated has values that are similar in magnitude. It is done via calculating a local Moran's I index value for clusters (Fischer and Getis 2009). The output gives a representation of the statistical significance of the index values in the z-scores and p-values (ESRI 2012).

For an output of a cluster to be obtained, there must be a positive value of the Moran's I index that signifies that a feature is surrounded by similar features of high or low values. Negative values of the index indicate that features are surrounded by features with dissimilar values and are termed an outlier. For statistically significant clusters or outliers, the p-values must be small enough (Fischer and Getis 2009).

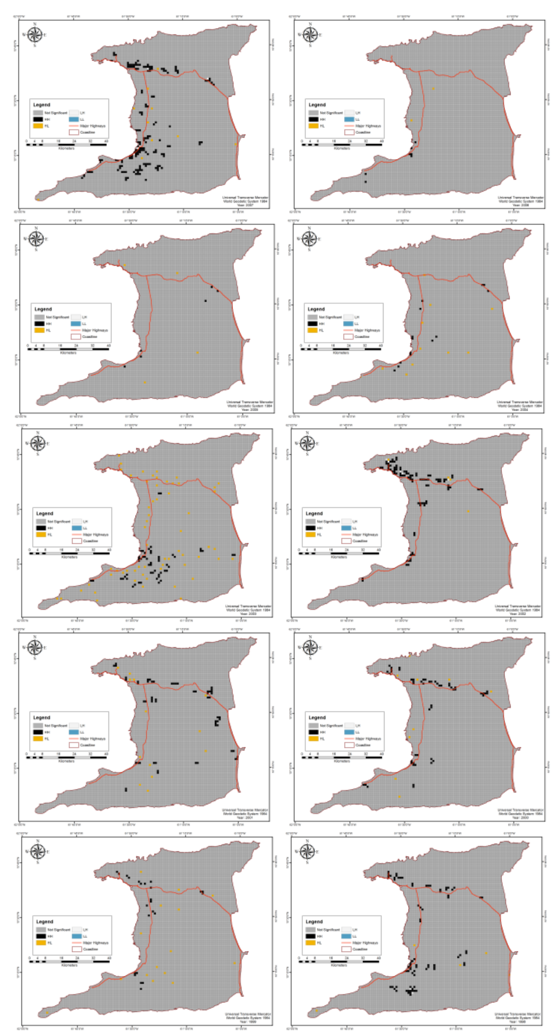

Figure 8: Cluster and Outlier Analysis (Anselin Local Moran's I)

The generated output is a map that shows statistically significant to the 0.05 confidence level areas of clusters that are of high values HH, low values LL, an outlier of a low value surrounded by high values LH and an outlier of a high value surrounded by low values HL (ESRI 2012).

When the Cluster and Outlier Analysis (Anselin Local Moran's I) statistic was performed, it was seen in Figure 8, that for the majority of the years, there were HH clusters occurring along the East-West Corridor and along the North-South Highway. Most interesting was the presence of HH cluster in the Ward of Naparima in every year. Also, the Ward of Sipara had HH clusters for the years 1998, 1999, 2000, 2001, 2003 and 2007. The HH clusters were concentrated mostly to the north west of the island near the capital city of Port-of-Spain and in the south in the Wards of Naparima and San Fernando.

7.2.2 *Hot Spot Analysis (Getis-Ord Gi*)*

The Getis-Ord Gi* statistic calculates hot spots within a dataset. High clusters or low clusters of features are produced based on resultant z-scores and p-values. The desired output is statistically significant hot spot of spatially clustered features with a statistically significant z-score. The tool obtains the z-scores for a dataset first by identifying features of high values surrounded by other features with high values.

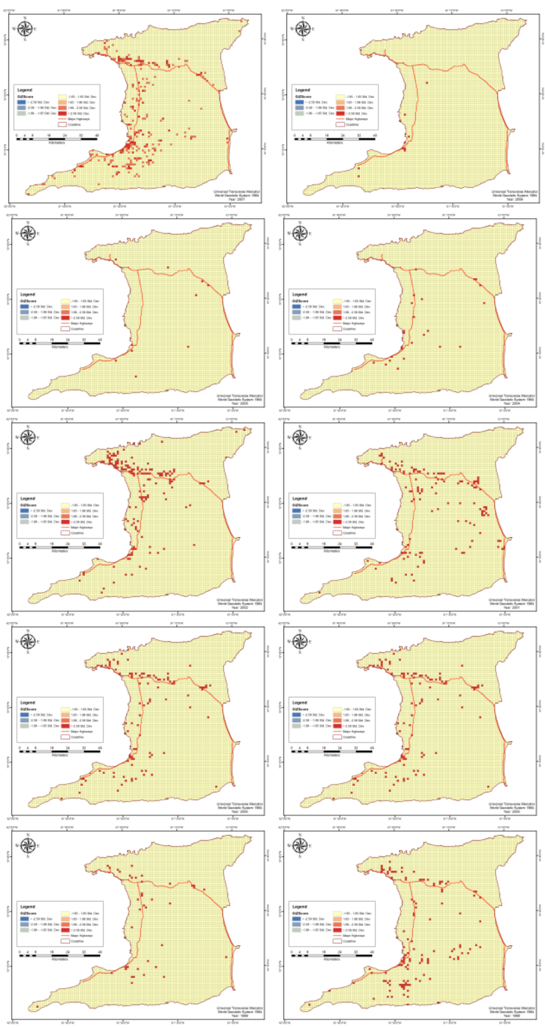

Figure 9: Hot Spot Analysis (Getis-Ord Gi*)

The larger the Gi* value i.e. a positive z-score, the tenser the clustering of high values which in turn gives a statistically significant hot spot. Conversely, a cold spot arises where statistically significant negative z-scores cluster (Fischer and Getis 2009).

When the Hot Spot Analysis (Getis-Ord Gi*) statistic was performed, the results (Figure 9) were similar to the Cluster and Outlier analysis such that the locations of the statistically significant hot spots followed closely the major highways for the island. For all years, there were hotspots located in the Wards of Naparima and Manzanilla. Also, all years illustrated that there was only statistically significant clusters of hotspots that were greater than 2.58 standard deviations except for 2007 that also had clusters of 1.96-2.58 standard deviations. There were no cold spots present.

7.3 The Knox Test

One of the most commonly used statistical space-time interaction techniques is that proposed by Knox (1964). For a dataset, the time and geographic location is noted for each event and also the distances between each pair of cases are calculated with respect to space and time. To have spatiotemporal interaction, cases that are 'close' in space must also be 'close' in time similarly; those that are 'close' in time must be 'close' in space. The process is straightforward to perform the statistic since there is no need for controls, only knowledge of the incidence (Kulldorff and Hjalmars 1999).

Since the Knox test is a measure of the closeness of space and time, the test is performed by comparing each pair of points to therefore give N*(N-1)/2 pairs. Once this is done, the distances between points are divided into 2 groups- close in distance and not close in distance; and the time intervals between points are also divided up into close in time as well as not close in time.

The Knox Index is calculated under a Monte Carlo simulation of the Chi-square statistic. If the user defines a simulation, the tool selects random pairs and at a time interval in the dataset, calculates the Knox index and the chi-square test.

The output contains the sample size, number of pairs, chi-square index, ten percentiles from the simulation, the maximum and minimum chi-square value from the Knox Index.

Table 5: The Knox Test Annual Summary

Year	Knox Chi-Square	95 Percentile Chi-Square	p-value
1998	8.96205	7.94922	0.0001
1999	8.04424	4.71932	0.0001
2000	3.7098	7.31943	0.0001
2001	24.3209	7.36545	0.0001
2002	315.30446	9.31139	0.0001
2003	288.40984	7.17649	0.0001
2004	1.17301	3.93689	1
2005	0.99417	4.05	1
2006	2.30604	3.45812	0.05
2007	0.86819	12.92022	1

To produce a good test for significance, 1000 simulations were chosen. In analysing the results, it is crucial to select an appropriate significance level. Since the Knox test is a one-tailed test it means that only high values for the chi-square would indicate spatial interaction therefore an upper cut off of the 95 percentile is chosen therefore if values of the Knox chi-square index is greater than the 95 percentile value, the null hypothesis can be rejected.

When investigating the statistic that attribute to both space and time clustering, the Knox test was utilised since the time and geographic location is taken into account for each event as well as the distances between the pairs of cases are calculated with respect to space and time. Therefore, a measure of closeness in space and closeness in time is obtained (Kulldorff and Hjalmars 1999). The output Knox indices (Table 5) indicated that only the years 1998, 1999, 2001, 2002 and 2003 were close in both space and time.

8. DENGUE AND LAND COVER

Land use change is a major component of global environmental change. It can be a potentially significant factor when investigating how mosquitoes-borne diseases are dispersed (Vanwambeke, et al. 2007). During the immature stages of the vector, a freestanding water habitat is crucial for their development and survival (Service and Townson 2002). Breeding sites can either be natural or artificial containers and/or bodies of water. Changes in land use could result in colonization of new breeding sites or the extension or reduction of the range of the mosquito. Also, there is the probability that there could be modifications of the composition of the vector community since different species of vectors have varying preferences in habitat conditions (Patz and Norris 2004; Vanwambeke et al. 2007).

The results for the association of DHF and DSS cases with land cover are shown in Table 6.

Table 6: Annual Summary of Dengue Incidence Correlated with Land Cover

Year	Agri-cultural Land	Barren Land	Forest Land	Range Land	Urban or Built-up Land	Wetland
1998	29	0	14	6	107	1
1999	9	0	3	5	31	0
2000	22	2	13	5	84	2
2001	23	1	34	3	95	1
2002	26	0	17	4	163	2
2003	58	3	52	12	241	0
2004	6	0	3	1	16	1
2005	1	0	2	0	8	0
2006	0	0	1	0	10	0
2007	26	0	9	1	346	0

There are 2 epochs of land cover datasets for the island of Trinidad: 1994 (based on 1994 aerial orthophotomosaic) and 2007 (based on the Ikonos 2007 satellite imagery). The Ikonos 2007 imagery is a compilation of 8 years of satellite data: 2000-2007. The temporal span of that imagery was to accommodate the specification of less than ten percent of cloud cover or less within a scene that is appropriate for satellite imagery purchasing (Jehu 2011). Since the temporal span of the imagery encompasses 80% of the temporal span of the dengue data used in this study, it was deemed more appropriate.

As the incidence per year was correlated with the land cover of the island, it was found that the majority of incident data was found to be located on Urban or Built-up Land followed by Agricultural Land. Since the mapping scale of the land cover was at a larger scale when comparing to the scale of geocoding of the DHF/DSS, this can be one of the reasons why there are incident cases located in land cover types such as Barren Land, Forest Land, Rangeland and Wetland. The scale of geocoding for the years 1998-2006 are unclear however, that for the year 2007 was done at 1:1,000. At a scale of 1:25,000 features tend to be more generalised as compared to a scale of 1:1,000 therefore there could have been houses within other land cover classes than Urban or Built-up Land but at the scale of digitizing, it may not have been identified. Also, due to the temporal span of the satellite imagery, some of the tiles may have been in years previous to the ones used for dengue comparison indicating that construction of buildings may not have taken place yet.

9. CLIMATIC CONDITIONS AND DENGUE

The El Nino events were in the years: 1997-1998, 2002-2003, 2004-2005 (weak) and 2006-2007 (weak) and the La Nina episodes were: latter part of 1998-2003, latter part of 2005-2006, latter part of 2007-2008 (NOAA 2012). However, the rainfall and temperature data was not at a sufficient resolution to identify proper correlations with these events and subsequently, the dengue incidence and therefore the data is omitted from this study. However, the ENSO events do coincide with the outbreak periods of 1998, 2002 and 2007 such that where there were strong periods of El Nino and La Nina, there was an epidemic for the island. This can be attributed to the research of Proveda, et al. (1999) that stated the presence of a drought as in an El Nino event, this may force persons to store water and as a result, create new habitats for mosquito breeding.

10. LIMITATIONS

The limitations for this study are: it is assumed that the addresses of the cases i.e. the addresses that have been used in geocoding are the locations at which the cases were infected and the actual geocoding was done at an appropriate scale so as to identify individual house locations. Secondly, it is assumed that the datasets of cases were comprehensive such that all cases have been reported and collected to be utilised in this study. Also, it is assumed that the tests for dengue detection were accurate and all cases used in this study are valid DHF and DSS cases. Lastly, limitations in resolutions of data such as the temperature and rainfall data have prevented a more robust study from being conducted.

11. CONCLUSION

It was found that the rainy season had the most impact on the overall distributions of dengue incidence to indicate that the majority of dengue cases presented themselves during this season and also the clustering present during this season was most tense and therefore very influential on the overall geographic distributions performed.

For the temporal span of this study, the mean centres led to a spatial pattern that holds closely the North-South Highway and mostly to the central and south western part of the island. This remained consistent for the annual investigation, the rainy season and for the dry season as well. This establishment of the cases following the major road network leads to the question - what is the dominating land cover that cases have been located on? This study suggests that most of the incidence correlated with Urban or

Built-up Land as well as Agricultural land as expected given that the factors that are contribution to the global incline of dengue incidence are land use change and increased and unplanned urbanisation (Gubler 1989; Vanwambeke et al. 2007).

For all the statistical tools run, all years have z-scores that are within the 99% such that they are either <-2.58 or >2.58 standard deviations with a p-value being <0.01 confidence level except for 2005 that was within the 90% confidence level with a p value falling within the <0.10 probability limit and z-scores of either <-1.65 or >1.65 standard deviations.

To be able to reject the null hypothesis for this study, the confidence level should be within 99% since that would indicate that the probability of the dengue spatial pattern per year is random is less than a 1% probability (Haining 2004).

The Average Nearest Neigbor statistic illustrated that for all the years, 1998, 2002 and 2007 exhibited the strongest clusters while the remaining years displayed clustering except for 2005 that showed no clustering. This was the similar results for the High/Low Clustering (Getis-Ord General G) as well as the Spatial Autocorrelation (Global Moran's I) statistics. When the Multi-Distance Spatial Cluster Analysis (Ripley's K Function) was run, it was seen that for the years 1998-2004 and 2006-2007 there is statistically significant spatial clustering taking place as distances continue to increase since for the distance threshold specified, only values where the observed K is greater than the expected K was obtained. However, for 2005 only at distances of 3500-4000m and 8500m were statistically significant clusters.

The Cluster and Outlier Analysis (Anselin Local Moran's I) and the Hot Spot Analysis (Getis-Ord Gi*) when compared to the Kernel Density analysis show similar results to the Kernel Density process. It can be deduced from this study that St. Anns, Port-of-Spain (the capital of the island), Naparima, San Fernando, Siparia and Manzanilla are the most common areas of dengue incidence in Trinidad.

For the island of Trinidad, the epidemic years were 1998, 2002, 2003 (Hemme et al. 2009) and 2007. These years coincided with the rainy season with high numbers of cases but also showed that 2000 was a year of interest. The dry season also corresponded with the epidemic years but with the year 2001 having a high number of incidence as well. Despite being an epidemic year, the year 2003 showed a low number of cases for the dry season.

The clustering tools used in this study all played different roles which were alluded to from their names- the global calculators gave overall statistically significant distributions per year whereas the local calculators gave statistically significant areas for different epochs in time and in space. When taking the Knox index into account, both space and time are of importance and hence, for those years of data that were of importance to this index, we can obtain truly statistically significant clustering.

The global statistical calculators identified that the years that displayed clustering was present in all years except for 2005 and the strongest clustering was present in 1998, 2002 and 2007. Therefore, the null hypothesis was accepted only for the year 2005.

The local statistical calculators deduced that the areas of most spatial clustering annually were in the Wards of Port-of-Spain and St. Ann's to the North West, Manzanilla to the east and in Naparima and San Fernando to the south west while Siparia, also in the south, displayed significant clustering as well.

When the Knox index was generated, it indicated that years of 1998, 1999, 2001, 2002 and 2003 displayed statistically significant clustering.

The years that displayed statistically significant clustering, spatial stationarity and spatial dependency for the island of Trinidad were: 1998, 1999, 2001, 2002 and 2003. Hence, these are the years that the null hypothesis has been rejected.

Therefore in conclusion, taking into account the epidemic years and the statistically significant years, it can be observed that even though an epoch of data may be not be from an epidemic year, there still can be statistically significant clustering occurring.

12. REFERENCES

[1] Arcari, Paula, Nigel Tapper, and Sharron Pfueller. 2007. Regional variability in relationships between climate and dengue/DHF in Indonesia. *Singapore Journal of Tropical Geography*: 251-272.

[2] Bertazzon, Stefania, Scott Olson, and Merril Knudtson. 2009. A Spatial Analysis of the Demographic and Socio-economic Variables Associated with Cardiocascular Disease in Calgary (Canada). *Springer Sciene and Business Media*.

[3] Bhandari, K P, P L N Raju, and B S Sokhi. 2008. Application of GIS Modelling for Dengue Fever Prone Area Based on Socio-Cultural and Environmental Factors- A Case Study of Delhi City Zone. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.*

[4] Central Intelligence Agency. *Central Intelligence Agency.* 15 08 2012. https://www.cia.gov/library/publications/the-world-factbook/geos/td.html (accessed 08 16, 2012).

[5] Chadee, D D, B Shivnauth, S C Rawlins, and A A Chen. 2007. Climate, mosquito indices and the epidemiology of dengue fever in Trinidad (2002-2004). *Annals of Tropical Medicine & Parasitology*: 69-77.

[6] Chadee, D D, R Doon, and D W Severson. 2007.Surveillance of dengue fever cases using a novel Aedes aegypti population sampling method in Trinidad,West Indies: the cardinal points approach.*ACTA TROPICA*: 1-7.

[7] Corwin, AL, et al. 2001.Epidemic dengue transmission in southern Sumatra, Indonesia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*: 257-265.

[8] ESRI. *ArcGIS Resource Centre.* 3 7 2012. (accessed 7 20, 2012).

[9] Fischer, Manfred M., and Arthur Getis. 2009. *Handbook of Applied Spatial Analysis.* Springer 2009.

[10] Foster, J E, S N Bennett, H Vaughan, V Vorndam, W O McMillan, and C V F Carrington. 2003. Molecular Evolution and Phylogeny of Dengue Type 4 Virus in the Caribbean. *Elsevier Science*: 126-134.

[11] Gagnon, Alexandre S., Andrew B. G. Bush, and Karen E. Smoyer-Tomic. 2001.Dengue epidemics and the El Niño Southern Oscillation. *Climate Research*: 35-43.

[12] Gibbons, R.V., and D.W. Vaughn. 2002. Dengue: An Escalating Problem. *British Medical Journal*: 324.

[13] Gubler, D J. 1998. Epidemic dengue and dengue hemorrhagic fever: a global public health problem in the 21st century. In *Emerging Infections*, by W. Michael Scheld, Donald Armstrong and James M Hughes, 1-14. Washington: ASM Press.

[14] Gubler, D. J. 2004.The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? *Comparative Immunology, Microbiology & Infectious Diseases 27*: 319-330.

[15] Gubler, D., and G. Clark. 1995. Dengue/Dengue Hemorrhagic Fever: The Emergence of a Global Health Problem. *Emerging Infectious Diseases*.

[16] Gubler, DJ, W Suharyono, I Lubis, S Eram, and SJ Sulianti. 1978. Epidemic DHF in rural Indonesia. *American Journal of Tropical Medicine and Hygiene*: 701-710.

[17] Haining, Robert. 2004. *Spatial Data Analysis Theory and Practice.* Cambridge: Cambridge University Press.

[18] Hemme, Ryan R., Jennifer L. Tank, Dave D. Chadee, and DavidW. Severson. 2009.Environmental conditions inwater storage drums and influences on Aedes aegypti in Trinidad,West Indies. *Acta Tropica*: 59-66.

[19] Hosein, Sarah, and Raid Al-Tahir. 2011. Spatio-Temporal Analysis of Dengue Incidence in Trinidad using Directional Distribution Ellipses. West Indian Medical Journal.

[20] Jehu, Adam. 2011. *Use of Principal Component Analysis and Geostatistics for the Derivation of a Forest Fragmentation Model.* St Augustine : The University of the West Indies.

[21] Kousky, V E, and R W Higgins. 2007. An Alert Classifications System for monitoring. *Weather and Forecasting*: 353–371.

[22] Kovat, S. 2000. El Niño and human health. *Bulletin of the World Health Organization* (Bulletin of the World Health Organization): 1127-1138.

[23] Kovats, RS, MJ Bouma, and A Haines. 1999. *El Niño and Health (WHO/SDE/PHE/99.4).* Geneva: Task Force on Climate and Health.

[24] Kulldorff, Martin, and Ulf Hjalmars. 1999. The Knox Method and Other Tests for Space-Time Interaction. *BIOMETRIC*: 544-552.

[25]Nathin, MA, SR Harun, and Sumarmo. 1998. Dengue haemorrhagic fever and Japanese B encephalitis in Indonesia. *Southeast Asian Journal of Tropical Medicine and Public Health*: 475-481.

[26] NOAA. *NOAA/ National Weather Service Climate Prediction Centre.* 8 6 2012. http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml (accessed 8 7, 2012).

[27] Patz, J A, and D E Norris. 2004. Land use change and human health. In *Ecosystems and Land Use Change, Geophysical Monograph*, by R DeFries, G Asner and R Houghton, 159–167. Washington, DC: American Geophysical Union.

[28] Seghal, R. 1997. Dengue fever and El Niño. *Lancet*: 729–730.

[29] Service, M W, and H Townson. 2002. The Anopheles vector. In *Essential Malariology*, by D A Warrel and H M Gilles, 59-84. London: Arnold.

[30] Stern, PC, and WE Easterling. 1999. *Making climate forecasts matter.* Washington, DC: National Academy Press.

[31] van Peenen, F F D, S J Atmosoedjono, J C Lien, and S J Saroso. 1972. Seasonal abundance of adult Aedes aegypti in Djakarta, Indonesia. *Journal of the American Mosquito Control Association*: 176-179.

[32] Vanwambeke, Sophie O., et al. 2007. Impact of Land-use Change on Dengue and Malaria in Northern Thailand. *EcoHealth*: 37-51.

[33] World Health Organization. 2002. Dengue prevention and control. *Fifty-fifth World Health Assembly*.