Faculty of Science and Engineering
School of Earth and Planetary Sciences

Development of a Web-enabled Spatial Decision Support System (SDSS) for Prevention of Tick Borne Disease in Kuantan, Malaysia

Kasmah Hamid

This thesis is presented for the Degree of
Master of Philosophy
of
Curtin University

March 2022

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:

Date: 23 March 2022

ABSTRACT

Ticks are the second most common vectors of human disease after mosquitoes. They are found on many small mammal hosts and also blood-feed on humans with the risk of transmitting diseases. In Malaysia, they are known to cause human otoacariasis disease along the east coast area as well as being the vector for other diseases such as Lyme disease in other countries. Considering the impact of this disease, this study will investigate the potential for a web-enabled spatial decision support system (SDSS) to improve the prevention of tick borne diseases specifically in Kuantan, Malaysia. In order to achieve this goal, four specific objectives have been defined, which are 1) to understand the decision making process for managing tick borne disease by understanding the different needs of decision makers and users, 2) to conduct spatial analysis and modelling of the contributing factors of tick-borne disease and development of future risk maps, 3) to investigate the major components of an SDSS system consisting of disease, users, Web GIS technology and decision support system (DSS), and 4) to explore the potential of interactive Web 2.0 technologies for user participation. This study will show how an SDSS can give benefits to the users and stakeholders by allowing different levels of user to access and share the information via the World Wide Web (WWW). The Web-enabled SDSS can also provide a platform for the users to perform spatial analysis and disease modelling for prediction of disease in the future. Furthermore, the system can be an effective tool for decision makers to get timely and accurate information in order to perform disease surveillance, policy making and prevention of disease.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful.

Alhamdulillah, praise to Allah for giving me the opportunity to pursue and successfully complete my research.

My special acknowledgement goes to my supervisor A/Prof Dr. Cecilia Xia for her guidance and assistance even before I start my journey at Curtin University. Thank you so much for your support and confidence in me. Thank you to Dr Robert (Rob) Corner for being a good mentor and for his valuable contributions of ideas to my studies. Thank you too to Professor Geoffrey (Geoff) West for his contribution and for the sharing of knowledge and expertise. Thank you to the people at the Department of Spatial Science, and the people in the Spatial Sciences office: Caroline Rockliff, Lori Patterson and Pam Kapica, who had been very accommodating and helpful throughout the time I was in Curtin University from 2011 to 2014.

Thank you to my sponsor, the Ministry of Science, Technology and Innovation (MOSTI) especially to the people in the HRD Unit who took care of everything with regards to the financial affairs. Thank you to the Director General of MYSA, Tuan Haji Azlikamil Napiah and my immediate supervisor Mrs. Jamilah Ismail who have given me their full support and cooperation.

I would also like to thank all my friends at room 310A and in building 207, especially my friends Charity Mundava, Aik Asmiwyati, Qian Sun, Chun Mei and many others. Thank you very much for the wonderful friendship and especially for the good times that we shared.

My special thanks to my dearest husband Che Abd Aziz Che Hussin who stands by me all the time and supports me during the ups and downs throughout my journey. Thank you to my sons Che Hamidi Husaini and Che Muhammad Khuzairi - together as a family we have experienced life in the beautiful city of Perth and have made good friends along the way. Last but not least, thank you to our families in Malaysia who have been really supportive and understanding, especially my mother: Hjh Kalsom Hashim. To my late father Hj. Hamid bin Ali, you have taught me well that education is very important. To my late parents-in-law, we could never have made it without your support, understanding and love.

To the people who are reading my thesis, I hope that you would be able to gain some useful knowledge from my research.

TABLE OF CONTENTS

ABSTRACT ii
ACKNOWLEDGEMENTS iii
TABLE OF CONTENTS v
LIST OF FIGURES viii
LIST OF TABLES x
TERMS AND ACRONYMS xi
1. Introduction 1
1. 1 Problem formulation 1
1. 2 Background 3
1.2.1 Tick borne disease and *human otoacariasis* 4
1. 3 Aim and Objectives of Research 5
1.3.1 Aim 5
1.3.2 Objectives of Research 5
1.3.3 Benefits of Research 6
1. 4 Research Framework and Methodologies 7
1. 5 Thesis overview 9
1. 6 Chapter summary 11
2 Vector borne disease 12
2. 1 Public health and disease management 12
2. 2 History of vector borne disease 14
2. 3 Vectors of vector borne disease 16
2.3.1 Mosquitoes 16
2.3.2 Ticks 17
2. 4 Ticks as a vector of vector borne disease 18
2.4.1 Ticks species 19
2.4.2 Ticks lifecycle 20
2. 5 Ticks species in South East Asia 22
2.5.1 The *Dermacentor* family 22
2.5.1.1 *Dermacentor variabilis* 24
2.5.1.2 *Dermacentor (Indocentor) atrosignatus* (Neumann, 1906) 26
2.5.1.3 *Dermacentor (Indocentor) steini* (Schulze, 1933) 26
2. 6 Tick borne disease in Malaysia 27

2.6.1 Control and management of tick borne disease in Malaysia 29

2.6.2 The lifecycle of species found in Malaysia…………………………… 30

2.6.3 Tick seasons 32

2.6.4 Habitats and preferred environments 33

2. 7 Chapter summary 34

3 Geospatial modelling of tick borne disease 35

3. 1 Remote sensing and public health.. 35

3.1.1 Vector borne diseases and remote sensing.. 39

3.1.2 Types of remote sensing data for disease monitoring and surveillance. 41

3.1.3 Future applications of remote sensing satellite images.......................... 48

3. 2 Spatial analysis and modelling for vector borne disease 49

3. 3 Spatial analysis and modelling for tick borne disease 50

3.3.1 Statistical approach to disease modelling 51

3.3.1.1 Species Distribution Modelling (SDM) and pseudo-absences 52

3.3.2 Biological approach to disease modelling 54

3. 4 Web GIS and Web-enabled SDSS 56

3.4.1 Web GIS applications in public health and disease management.......... 58

3.4.2 Decision Support Systems (DSS) 59

3.4.2.1 DSS architecture and framework 59

3.4.2.2 Spatial Decision Support System (SDSS) for tick borne disease 61

3. 5 Chapter summary 63

4 Study area and research methodology 64

4. 1 Study area... 64

4.1.1 The topography of Kuantan 65

4.1.2 The climate of Kuantan... 66

4.1.3 Vegetation, land use and economic activities 68

4.1.4 Factors influencing risk.. 68

4. 2 Data requirement for spatial data analysis 69

4.2.1 Defining variables for tick borne disease modelling 70

4.2.2 Possible factors and data for analysis in study area 70

4. 3 Spatial analysis and modelling methods 85

4.3.1 Exploratory Data Analysis and Visualization....................................... 85

4.3.2 Data analysis 86

4.3.3 Species distribution modelling (SDM) and pseudo-absence analysis ... 86
4. 4 Software ... 90
4.4.1 GIS software ... 90
4.4.2 Statistical software ... 90
4. 5 Chapter summary ... 91

5 Analytical results ... 92
5. 1 Spatial data analysis for tick borne disease ... 92
5. 2 Visualization and statistical analysis ... 95
5.2.1 Meteorological data ... 95
5.2.2 Remote sensing and GIS ... 101
5. 3 Developing the CART Model ... 111
5. 4 Chapter summary ... 120
6 Development of a prototype for web-enabled SDSS for tick borne disease.... 122
6. 1 Significance of Web-enabled SDSS for Tick Borne Disease ... 122
6. 2 The Components of Web-enabled SDSS for Tick Borne Disease ... 123
6. 3 Web-enabled Spatial Decision Support System (SDSS) for tick borne disease ... 132
6. 4 The Prototype Development of Web-enabled SDSS for tick borne disease 134
6.4.1 Features and functions of dashboard ... 135
6.4.2 Dashboard Conceptual architecture ... 137
6.4.3 Sample interface of prototype system ... 141
6. 5 The potential of Web 2.0 Technologies for user participation ... 145
6. 6 Chapter summary ... 148
7 Conclusions and Recommendations ... 149
7. 1 Research achievement ... 149
7. 2 Problems encountered and setback to research ... 151
7. 3 Opportunity for improvement and further research ... 154
7. 4 Chapter summary ... 158
REFERENCES ... 160
APPENDICES ... 204

LIST OF FIGURES

Figure 1.1 Research design framework 8
Figure 2.1 Map showing deaths from vector borne disease (WHO 2005) 15
Figure 2.2 The anatomy of hard tick (left) and soft tick (right) (Richards, Rawlings, and Hollan 2000) 18
Figure 2.3 Relative sizes of several ticks at different life stages. (CDC 2017)……………………………………………………………………20
Figure 2.4 Lifecycle of a three-host hard tick based on *Dermacentor variabilis* (American dog tick) Illustration by Scott Charlesworth, Purdue University (2008)……………………………………………………………………21
Figure 2.5 The life stages of the American dog tick, *Dermacentor variabilis* (Chan and Kaufman 2008) 25
Figure 2.6 The Generalized Lifecycle of *Dermacentor variabilis* and *Dermacentor andersoni* ticks (family Ixodidae) (Mayer 2015) 32
Figure 3.1 The map by John Snow showing the area of cholera outbreak in London epidemic of 1854 (Koch 2004) 36
Figure 3.2 The minimum architecture of a Web GIS system (Piarsa, Sudana, and Gunadi 2012) 57
Figure 3.3 Schematic Representation of Web SDSS Components (Sugumaran and Sugumaran 2007) 62
Figure 4.1 The Kuantan district and its sub districts 65
Figure 4.2 Map showing the locations of meteorological stations in Peninsular Malaysia……………………………………………………………………….. 71
Figure 4.3 Graph showing the monthly mean maximum temperature from 2002 to 2007 72
Figure 4.4 Land use map showing different types of land use categorized in code as described in Table 4.1………………………………………………………78
Figure 4.5 Map showing cases distribution within two-kilometre buffers of river .. 80
Figure 4.6 The CART diagram for modelling of tick borne disease 88
Figure 4.7 A basic ROC graph showing five discreet classifiers (Fawcett 2006)……………………………………………………………………90

Figure 5.1 Temporal distributions of human otoacariasis cases from 2002 to 2007..93
Figure 5.2 Map showing the spatial distribution of human otoacariasis cases from year 2002 to 2007 94
Figure 5.3 Graphs showing linear correlation between monthly mean temperature against number of cases for each year 99
Figure 5.4 Pie charts showing the total percentage of land use type within a two kilometre radius from each point of infections for every year from the year 2002 to 2007..........105
Figure 5.5 Pie chart showing the accumulated total percentage of land use type within a two-kilometre radius from each point of infections from the year 2002 to 2007 107
Figure 5.6 Graphs showing Mean EVI values for each human otoacariasis cases in the year 2004.......... 109
Figure 5.7 Graph showing comparison between Mean EVI for presence and absence data for the year 2004 110
Figure 5.8. Linear regression between oil palm, urban and swamp for 2004.......... 113
Figure 5.9 The ROC plot for presence data (actual cases of human otoacariasis) .. 116
Figure 5.10 ROC plot for first set of random points (pseudo-absence data)........... 117
Figure 5.11 The ROC plot for the second set of random points (pseudo-absence data) 118
Figure 5.12 Distribution of cases against significant factors of tick borne disease. 119
Figure 5.13 The summarised concept of modelling for prediction of tick borne disease 120
Figure 6.1 The three components of a web-enabled SDSS for tick borne disease .. 123
Figure 6.2 DSS Users.......... 131
Figure 6.3 Schematic Representation of Web enabled SDSS for tick borne disease 132
Figure 6.4 Dashboard design for Web-enabled SDSS for tick borne disease 136
Figure 6.5 Conceptual Dashboard Architecture.......... 138
Figure 6.6 Main Interface of the Web-enabled SDSS.. 142
Figure 6.7 Interface for Health Workers.......... 143
Figure 6.8 Interface for Researchers.......... 144
Figure 6.9 Interface for the Public..........145

LIST OF TABLES

Table 1.1 Vectors and the diseases that they can transmit (WHO 2014) 2

Table 2.1 Ticks, hosts and their virus as found in Malaysia according to research by Nadchatram (2008)……………………………………………………………31

Table 3.1 The use of remote sensing in various disciplines (Chandra and Ghosh 2006) 38

Table 3.2 Characteristics of retired (a) and active (b) Landsat Satellites and Sensors as compiled from NASA (Irons 2011; USGS 2013; USGS 2014; Irons 2015) 43

Table 3.3 Specifications of retired (a) and active (b) SPOT satellites. Table updated from CNES (2010) and Airbus Defence and Space (n.d.) 45

Table 3.4 MODIS applications according to bands by NASA (n.d.) 47

Table 4.1 Land use type code and its description 79

Table 4.2 Result of Chi square test for number of cases within specified distances from the river for the year 2002 until 2007…………………………………………..82

Table 4.3 Total of population by gender for selected area, Pahang, Malaysia. Source: Population and Housing Census of Malaysia 2010 84

Table 5.1 Monthly records of human otoacariasis cases from year 2002 to 2007..... 93

Table 5.2 Pearson Correlation between case numbers and monthly meteorological variables 95

Table 5.3 The Pearson correlation value (R^2) between the volume of rain and the number of rainy days against number of cases for each month 101

Table 5.4 Percentage of land use types within a two-kilometre radius of an infection point 106

Table 5.5 The result of chi square test between land use type and cases 108

Table 5.6 The result of linear regression analysis for the three dominant land use 114

Table 6.1 The different types of databases…………………………………………125

Table 6.2 List of databases supported by ArcGIS…………………………………..127

Table 6.3 Source Data System for Web-enabled SDSS……………………………139

Table 6.4 Web 2.0 applications. Examples and open source solutions …………….146

TERMS AND ACRONYMS

Below is a list of terms and acronyms used in the thesis:

AUC	Area Under The Curve
AVHRR	Advanced Very High Resolution Radiometer
AWT	Australian Wet Tropics
BRT	Boosted Regression Tree
CART	Classification And Regression Tree
CCHF	Crimean-Congo Haemorrhagic Fever
CDC	Centre For Disease Control (USA)
CDCIS	Communicable Diseases Control Information System
CHA	Community Health Assessment
CNES	Centre National d'Etudes Spatiales
CSF	Classical Swine Fever
DENV	Dengue Virus
ECER	East Coast Economic Region
ECERDC	East Coast Economic Region Development Council
ECF	East Coast Fever
EHEC	Enterohemorrhagic E. Coli
ESRI	Environmental Systems Research Institute
ETM+	Enhanced Thematic Mapper
EVI	Enhanced Vegetation Index
EWS	Early Warning Systems
GAM	Generalized Additive Models
GLM	Generalized Linear Models
GHS	Ghana Health Service

DBMS	Database Management System
DOSM	Department Of Statistics Malaysia
DSS	Decision Support System
FP	False Positive
FN	False Negative
FPR	False Positive Rate
GIS	Geographical Information System
H1N1	Influenza A
H5N1	Avian Influenza
HTTP	Hyper Text Transfer Protocol
IMR	Institute Of Medical Research
IOM	United States Institute Of Medicine
IS	Information Systems
IT	Information Technology
JEV	Japanese Encephalitis Virus
JPNP	Jabatan Perhutanan Negeri Pahang
LB	Lyme Borreliosis
LD	Lyme Disease
LPDAAC	Land Processes Distributed Active Archive Center
LST	Land Surface Temperature
MERS	Middle East Respiratory Syndrome
MMD	Malaysian Meteorological Department
MOD13Q1	MODIS/Terra Vegetation Indices 16-Day L3 Global 250m
MODIS	Moderate Resolution Imaging Spectroradiometer
MOH	Ministry Of Health (Malaysia)

MOHCW	Ministry Of Health And Child Welfare (Zimbabwe)
MRSA	Malaysian Remote Sensing Agency
MSS	Multispectral Scanner
MySED	Malaysia Strategic Workplan For Emerging Disease
NASA	National Aeronautics And Space Administration
NDVI	Normalized Difference Vegetation Index
NE	Nephropathia Epidemica
NIR	Near Infrared
OGC	Open Geospatial Consortium
OLAP	On-Line Analytical Processing
OLI	Operational Land Imager
PID	Parasitic And Infectious Diseases
PAN	Panchromatic
RBV	Return Beam Vidicon
RCA	Radio Corporation Of America
RH	Relative Humidity
RMSF	Rocky Mountain Spotted Fever
ROC	Receiver Operating Characteristics
RRV	Ross River Virus
RSSEV	Russian Spring Summer Encephalitis Virus
RVF	Rift Valley Fever
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
SARS	Severe Acute Respiratory Syndrome
SDM	Species Distribution Modelling
SDSS	Spatial Decision Support System

SEPPSF	South-East Pahang Peat Swamp Forest
SEZ	Special Economic Zone
SNR	Signal-To-Noise Ratio
SPSS	Statistical Package For The Social Sciences
SQL	Structured Query Language
SWIR	Short Wave Infrared
SOP	Standard Operating Procedure
SOVAT	Spatial OLAP Visualization And Analysis Tool
SPOT	Satellite Pour l'Observation De La Terre
TBD	Tick-Borne Disease
TBE	Tick-Borne Encephalitis
TBEV	Tick-Borne Encephalitis Virus
TIRS	Thermal Infrared Sensor
TM	Thematic Mapper
TP	True Positive
TN	True Negative
TPR	True Positive Rate
UMMC	University Of Missisippi Medical Centre
USA	United States Of America
USGS	United States Geological Survey
VBD	Vector Borne Disease
VBDC	Vector Borne Disease Control
VI	Vegetation Index / Indices
Web-enabled SDSS	Web-Enabled Spatial Decision Support System
WHO	World Health Organization

WFS	Web Feature Service
WMS	Web Map Service
WNV	West Nile Virus
WWW	World Wide Web
YFV	Yellow Fever Virus

1 INTRODUCTION

This chapter presents the problem statement regarding tick borne disease in general with specific focus on the study area of Kuantan, Pahang in Malaysia. The objective is to identify the significant factors leading to tick bite incidents and to predict the risk of incidents in the future based on the presence of the significant factors. Meanwhile, current practice on decision making during infectious and non-infectious diseases outbreak is also addressed in finding solutions and improving the decision making process through the implementation of a Web-enabled Spatial Decision System (SDSS). On the other hand, the proposed solutions can only be described in theory, as an actual system development would require more time for research and testing. Nevertheless, this research is able to provide basic methods and guidelines for development and implementation on a web-enabled SDSS for tick borne disease and a web-enabled prototype system has been developed to demonstrate basic functionalities of the system. Four research objectives have been identified and will be described further in this chapter.

1. 1 Problem formulation

Vector borne diseases (VBD) has become one of a number of global health concerns as well as "an increasing cause of death and suffering worldwide" (Roberts and Andre 1994). These diseases have affected the human population as well as animals, causing symptoms of mild pain or discomfort, to paralysis and in some severe cases resulting in casualty. When animals contracted a disease, it may cause health problems and deterioration that eventually contributed to economic loss. For human, diseases such as VBD may not just decrease the quality of public health in general, they also have the potential to spread and trigger an outbreak if not controlled and contained at an early stage.

To control and prevent VBD outbreaks, the vectors need to be identified earlier prior to formulation of any control and preventive actions. These vectors are carriers of disease and the most frequent and common vectors for VBD are small insects such as mosquitoes and ticks. When they bite, these vectors are able to transmit harmful viruses or bacteria to their hosts. For instance, mosquitoes are the most common vectors for malaria, dengue and yellow fever. Heinz and Stiasny

(2012) classified the most important human pathogenic flaviviruses regarding to their impact on humans as: yellow fever virus (YFV), dengue virus (DENV), Japanese encephalitis virus (JEV), West Nile virus (WNV) and tick-borne encephalitis virus (TBEV).

The World Health Organization (WHO 2014) have listed the most important vectors causing different type of serious vector borne diseases as mosquitoes, sandflies, triatomine bugs, ticks, fleas and various species of flies. The details of vectors and their respective diseases are shown in Table 1.1.

Vector		**Diseases**
Mosquitoes:	*Aedes aegypti*	Dengue, Yellow fever, Chikungunya, Zika virus
	Aedes albopictus	Chikungunya, Dengue, West Nile Virus
	Culex quinquefasciatus	Lymphatic filariasis
	Anopheles (more than 60 known species can transmit diseases)	Malaria, Lymphatic filariasis (in Africa)
	Haemagogus	Yellow fever
Sandflies		Leishmaniasis
Triatomine bugs		Chagas disease
Ticks		Crimean-Congo haemorrhagic fever, Tick-borne encephalitis, Typhus, Lyme disease
Fleas		Plague, Murine typhus
Flies (various species)		Human African trypanosomiasis, Onchocerciasis

Table 1.1 Vectors and the diseases that they can transmit (WHO 2014)

Ticks are one of the important vectors of vector borne disease. Diseases related to ticks are referred to as tick borne disease. According to Hönig et al. (2011), “tick-borne diseases (TBD) belong to the so-called emerging diseases, caused by new or altered disease agents, diseases occurring in a new context or with different intensity”. Ticks are haematophagous ectoparasites of vertebrates that derive nutrition through blood feeding and are efficient vectors of disease (de la Fuente et al., 2017). Pathogens transmitted by ticks are responsible for the majority of the vector-borne diseases in temperate North America, Europe and Asia (Rochlin & Toledo, 2020). Some tick-borne viruses pose a significant threat to the health of humans (tick-borne encephalitis virus, Crimeean-Congo haemorrhagic fever virus) or livestock (African swine fever virus, Nairobi sheep disease virus) (Labuda & Nuttall, 2008).

Realizing the threats of these tick borne diseases to humans, this research has been undertaken to help the related stakeholders, particularly the government decision makers to understand the potential threat of tick borne disease and thus be able to control and prevent the diseases from becoming widespread in Malaysia. However, before any action and control can be taken, the Malaysian government's health authorities first need to understand the nature of tick borne disease, ticks' potential as vectors and how ticks transmit bacteria and viruses to humans. Background research has been carried out through literature review before identifying potential factors leading to tick infestation and tick bites. The potential factors are then investigated to identify significant factors, which are then incorporated in the model to predict the risk of tick bites in the future.

Data and information can be shared with other users using a platform, the web-enabled Spatial Decision Support System (SDSS). By having the system available through the web, data can be shared among the users in an efficient manner, with quick access and can be distributed to other users who have access to the system. Data can be presented in spatial maps with functions to perform basic queries, allowing users to perform visual analysis. The information shared in the system will be useful in assisting the health authorities to plan, manage their resources and take preventive actions. At the same time, by enabling quick access to data as well as information sharing, it can help the respective stakeholders and decision makers to make informed decisions during, before and after an outbreak.

1. 2 Background

There is no geographical boundary that can limit the spread of a disease including infectious diseases. Emerging diseases such as SARS, MERS and Influenza A (H1N1) viruses are among the threats faced by the every country in the world including Malaysia. Therefore it is important to understand the causes and threat that can be brought by these viruses therefore preventing the consequences of outbreak. The Malaysian government and health authorities have realized the importance of controlling and preventing vector borne disease outbreaks in Malaysia. Therefore actions are being taken to control the vectors or carriers of the harmful viruses and bacteria to prevent the spread of disease. Vectors such as mosquitoes, ticks, mites and flies are among important vectors of vector borne disease. However for this study, ticks will be the main focus of research.

For this study, ticks have been identified as causing human otoacariasis, or infestation in the human ear canal, in particular in the Kuantan, Pahang in Malaysia. In this case, ticks came into contact with the patients and entered their ear canal, thus causing pain. If ticks are detected early during their infestation, they can be removed easily and the pain can be treated by administration of antibiotics. As for human otoacariasis, it is not easy to extract ticks and patients may not realize that they have been infested until they experience pain in the ears, thus have to seek treatment. However, if human otoacarisis patients are not treated at an early stage, tick bites may cause further pain and even facial paralysis. Therefore the potential of ticks to cause further harm is a serious matter and combined with their capabilities as vectors of harmful bacteria and viruses, ticks can definitely pose as a threat to public health.

1.2.1 Tick borne disease (TBD) and human otoacariasis in Malaysia

Ticks are vectors of bacteria and virus, where some of them may cause harmful diseases to human such as Tick-borne Encephalitis (TBE) and Lyme disease. Tick-borne encephalitis (TBE) is also the most common viral tick-borne disease in Europe causing thousands of human infections every year (Walter et al., 2020). Meanwhile in the Kuantan area even though there were high numbers of human otoacariasis cases recorded every year, there were no reports on associated disease found on the patients. However, the potential of ticks as vectors of diseases cannot be simply dismissed since there are records of tick borne disease which have occurred earlier in other places in the country. Therefore this study was focussed towards identifying factors that may cause high incidences of tick bites, using the factors in a model to predict future incidence and risks, as well as channelling the information to the government decision makers, health workers, researchers and the public with better understanding and timely information on ticks as vectors of vector borne disease.

This study has also identified the research gaps on the previous and current studies of tick borne disease in Malaysia. Based on the literature review, ticks have been an important subject of studies in Malaysia as they are efficient vectors of multiple pathogens due to their potential interactions with several different vertebrate hosts during their life cycle (Science., 2011). However, previous research have only looked into the biological aspects and investigation on the possible transmission of disease from ticks to human and animals, without specific research on the contributing factors of infection, specifically the environmental factors. Therefore this research is dedicated to investigate the potential environmental factors by analysing meteorological parameters and vegetation data derived from remote sensing satellite in order to identify significant factors of tick borne disease.

1.3 Aim and objectives of research

1.3.1 Aim

The aim of this research is to investigate the impact of risk factors of tick-borne diseases and develop a web-enabled spatial decision support system to effectively manage the diseases.

1.3.2 Objectives of research

The objectives of this research are:

- To get a better understanding of ticks and their role as vectors of vector-borne disease
- To identify the factors of tick-borne disease or human otoacariasis in Kuantan, Pahang

- To conduct spatial analysis and modelling of tick-borne disease for risk management
- To develop a Web-enabled Spatial Decision Support System (SDSS) to facilitate decision-making and sharing of information between different types of stakeholders
- To identify the potential of Web 2.0 technologies for user participation

1.3.3 Benefits of research

Tick borne disease is not as common as mosquito borne disease in Malaysia. However the threat of the disease is not something that can be overlooked since serious incidents of tick borne disease have been recorded in the past, which is described further in Chapter 2.

Since currently tick borne disease does not have the same impact as mosquito borne disease, the issues and problems with regards to the disease are not being regarded as a priority. However ticks as a vector have the potential to transmit dangerous bacteria and virus to human, therefore it is still a threat and cannot be regarded as unimportant. Hence, fewer studies have been done on ticks as vectors as compared to mosquitoes. Therefore it is important that the government and the health authorities need to understand the threat of tick borne disease before they are able to take the right action to manage and prevent the disease from spreading.

Through this research, data and information on ticks can be compiled and made available to four identified users; the government decision makers, health workers, researchers and the general public. The decision makers consist of three levels of health authorities, which are the district health office at the lowest, state office in the middle and national level health office, which is the highest level. The lower level of health office reports to the upper level while orders and instructions flow from top to bottom. Each level of health decision makers can make use of new and latest information on ticks and events of disease occurrence so that they are able to make plans and take action at their own level. The second category of users is health workers who need information to make correct diagnosis and treatment of patients. The third category is the researchers who require more information on ticks and the threat of diseases, for example to come out with cures and methods to control the tick population. The fourth category of users is the general public, who need to understand the threat of tick borne disease, what should be done if they have been bitten, when to get treatment and how to avoid themselves from being bitten amongst others. All of the users can benefit from the data and information on ticks, made available to them through access to the system known as a web-enabled SDSS.

At present, there are online systems developed for reporting of infectious diseases including a dedicated system for dengue. However the online system for tick borne disease is more towards reporting of cases and locations of patients and is oriented more towards a traditional database. Therefore the web-enabled SDSS, which is supported with GIS capabilities, may assist better decision-making and conveyance of information with added features such as spatial analysis and maps for improved visual analysis. Instead of text and statistics, users are able to view spatial maps, able to perform queries and view statistics information and have access to more presentable reports. At the same time different categories of users can use the system as a platform for exchange of information. The health decision makers can convey their orders and plans to the health workers and researchers while informing the public with the health policy and community program. At the same time the health workers and the researchers are able to submit their reports and findings to the health decision makers, which will assist the decision makers for their strategic planning. The general public can also contribute to the system by offering information on the ground. Therefore, using the information, analysis can be made to discover factors leading to an increase of infection, identify pattern and trends of infection and helping everybody to be better prepared and sufficiently equipped during events of emergencies.

1.4 Research Framework and Methodologies

The research framework for this study has followed the design framework as described in Figure 1.1.

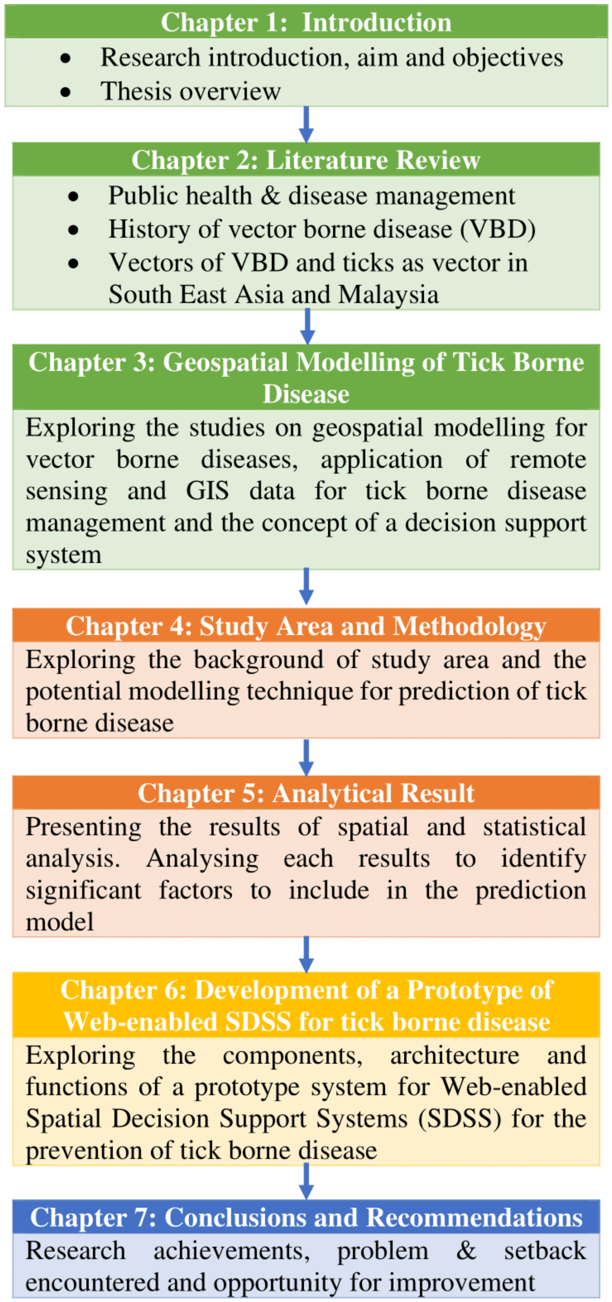


Figure 1.1 Research design framework

Research methodologies involved reviewing the literature to get some background information on tick species found in Kuantan study area to understand tick biology and their habitat. Even though research concentrates on the study area, there are other species of ticks found elsewhere in Malaysia based on the literature.

A further step is identifying potential factors of tick infestation such as temperature, climate, wind, humidity, the type of land use cover, vegetation, other factors such as vicinity to road or river networks. Data obtained consist of records of human otoacariasis patients from the year 2002 until 2007, remote sensing derived data such as land use data and vegetation, administrative boundary map, river and road network as well as population data.

The potential factors are then analysed using statistical methods and spatial analysis to identify significant factors of tick infection. Based on the type of significant factors, a modelling method has been selected to model and predict the risk of tick infection. The result of the analysis and modelling can become important information for decision-making. These data and information as a result of analysis and modelling can be incorporated into the web-enabled Spatial Decision Support System, which will facilitate quick flow of accurate and timely information between users of the system which includes the government decision makers, health workers, researchers and the general public.

1.5 Thesis overview

This thesis comprises a total of seven chapters. Chapter 1 covers the background of vector borne disease (VBD) in general with specific focus on tick borne disease problems at a global level, the effects and threats of VBD to public health and the potential of disease to spread and become an outbreak. The objectives of research are identified and provide the guidelines towards achieving the goal of research. This chapter also gives insight into each individual chapter.

Chapter 2 provides the literature review on ticks as vectors of tick borne disease. Ticks species, their lifecycle as well as habitats and preferred environments are described further in this chapter. This includes bacteria and viruses that can be transmitted by ticks to their hosts as well as symptoms and diseases caused by tick bites. The chapter also addresses the potential diseases that may threat the human population if the diseases are not controlled at an early stage, as well as prevention measures that can be taken to avoid an outbreak.

Chapter 3 describes the analysis process and literature review on the geo-spatial modelling of tick borne disease. A few selected methods for modelling of vector borne disease including modelling applications in tick borne disease performed by other researchers are described in this chapter as options for choosing the best method for application in the study area. The identified modeling method can be expanded in the future for other areas in Malaysia and countries of the South East Asia Region, which have similar climatic and weather conditions as well as vegetation types.

Chapter 4 describes the background of the study area including the climate of Kuantan, its population as well as the economic activities of its population. Data selected for the study area has been obtained based on the best available data during the six years of the study period. The methodology is selected based on the literature review of previous research covered in Chapter 3. The best method has been identified as being able to produce good results based on the criteria of study area and type of available data, which also depends on data availability and the quality of available data. The research summary right from data selection to methods are explained in this chapter.

Chapter 5 discusses the results of statistical analysis of possible factors of tick borne disease. The factors with statistically significant results have been identified as factors that contributed to the occurrence of tick borne disease. These significant factors or predictors are then integrated into a model to predict the potential risk of disease occurrence in the future. This model can then be applied as a prediction model to other areas in the country.

Chapter 6 describes an integration of analysis results into web-enabled SDSS. The result of analysis can be displayed in the form of spatial maps of tick borne distribution and as well as disease risk map. The chapter also explores the potential of Web 2.0 technologies, particularly the application of social media to encourage user participation and improve real time reporting of any event of outbreak as well as disseminating information to the public. A web-enabled prototype system has been developed and made accesssible through the Internet to show some basic functionalities of the proposed actual system.

Chapter 7 presents the problems and setbacks encountered during research, which has an impact on the research direction. The conclusions of the research are also summarized in this chapter based on the outcomes and findings during analysis. Finally recommendation for further research is included for the purpose of improving the research in the future.

1.6 Chapter summary

This chapter describes the rationale and motives of this research to identify the significant factors of tick borne disease in the study area. The factors can then be integrated into a model for prediction of disease risk in the future. At the same time data and information can be incorporated into a Web-enabled Spatial Decision Support Systems (SDSS), which can facilitate quick data flow and dissemination of information to the users. The decision makers will benefit from the system, as they are able to access accurate data to make critical decisions before, during and after an outbreak. This will facilitate the planning of resource management during emergency situation and for the authorities to coordinate inter-agency with other agencies such as the National Security Council, local councils, fire department and the police department. At the same time other users such as health workers; researchers and the general public will also be able to benefit from the system. Through the web-enabled SDSS, not only can users have access to the information, but they can also provide their input and share the information with other users.

2 VECTOR BORNE DISEASE

This chapter covers the history and background of vector borne diseases at a global level and their threat to the human population. There are also descriptions on types of vectors of vector borne disease with ticks as the main vector in focus. Then the chapter will focus on the tick species found in Kuantan District study area, located in the state of Pahang, Malaysia. This background information is able to provide a better understanding of the current and possible threat that can be brought by tick bites. It will also be able to give some insight or information that can assist the health authorities to mitigate actions on vectors control, identify risk factors and eventually be able prevent disease outbreak in the future.

2.1 Public health and disease management

Public health is one of the utmost priorities of all governments in the world. Based on the definition from the United States Institute of Medicine (IOM), public health is what we do collectively to assure the conditions for people to be healthy (Hodge 2005). Also according to Winslow (cited in Koplan et al. 2009), "public health is the science and art of preventing disease, prolonging life and promoting physical health and efficacy through organized community efforts for the sanitation of the environment, the control of communicable infections, the education of the individual in personal hygiene, the organization of medical and nursing services for the early diagnosis and preventive treatment of disease and the development of social machinery which will ensure every individual in the community a standard of living adequate for the maintenance of health; so organizing these benefits in such a fashion as to enable every citizen to realize his birthright and longevity". Every good government wants to ensure that its citizens' health and well-being are looked after in the best possible way to ensure a healthy population and prosperous nation. Therefore, to make good decisions on what is best for public health, government decision makers require quick, reliable and accurate information at the right time and this can be achieved by utilizing the best available tools for decision making.

Decision makers may rely on different tools to assist them in the decision-making process. One of the best decision tools available is known as a Decision Support System (DSS). A Decision Support System (DSS) is defined by Alter (2002) as "an interactive information system that provides information, models and data manipulation tools to help the decision maker make decisions in semi-structured and unstructured situations, where the decision process is both subjective and objective”. While Shim et al. (2002) has described it as “a computer technology solution that can be used to support complex decision making and problem solving”.

DSSs have been used to assist decision makers in public health to make informed and better decisions with regards to health matters. DSS has, for example, been applied as a framework known as RealOpt for modelling and optimizing the public-health infrastructure for all hazard emergency responses such as bioterrorist attacks or pandemics (Lee *et.al* 2009). By having such systems in place, it may allow public health emergency coordinators to allocate and despatch resources quickly as well as make performance analyses and investigate alternative strategies during such emergencies.

On-Line Analytical Processing (OLAP) is a multidimensional data warehouse technique that is commonly used as a decision support system which is not sufficient for solving numerical–spatial problems that frequently occur in community health assessment (CHA) research; however, coupling OLAP with GIS technology offers the potential for a very powerful and useful system (Scotch & Parmanto, 2006). The system is known as Spatial OLAP Visualization and Analysis Tool (SOVAT), an OLAP-GIS decision support systems to access health community data in order to identify health priorities (Scotch, Parmanto and Monaco 2008).

The extensive use of DSSs in public health has been acknowledged and proven as an important tool for public health management. The application of DSS in public health and other field of applications to support decision making will be described further in Chapter 3.

Therefore the focus in this study is to investigate the use of DSS for management of disease to maintain a good public health system. The identified disease for this study will focus on management and predicting risks of vector borne diseases caused by ticks.

2.2 History of vector borne disease

Vector borne diseases (VBDs) have been one of the most important worldwide health problems for many years and are still a serious risk to a large part of the world's population (Andrianasolo et al. 1999). Two of the most well-known vectors of vector borne diseases are mosquitoes and ticks. Even though many studies have been performed on mosquito borne diseases such as dengue and malaria, very little work has been done on ticks, which are the second most common vector for VBD. In the United States of America (USA), ticks have become the most common vectors of VBD (Edlow 2008) with a total of 248,074 cases of Lyme disease caused by tick infestations reported to the Centre for Disease Control (CDC) by the country's health departments during the period of 1992-2006 (Bacon, Kugeler and Mead 2008). At the same time in 2009 alone, health departments had reported a total of 29,959 of confirmed cases and 8,509 probable cases of Lyme disease to the CDC (Hall-Baker et al. 2009).

In this world, there are many countries with responsible and concerned governments, which have made a lot of effort to provide quality health services for their people. These governments are aware of the increase of emerging diseases such as infectious diseases, which are threatening the human population. Therefore, realizing the significant effects that can be brought by VBDs, much research has been conducted to investigate the cause of vector borne diseases by identifying the vectors, finding cures and taking control measures to prevent the disease from the early stages or when threat is detected. Efforts are also being made to educate the public on the disease to prevent people from getting infected as well as how to detect symptoms of disease, right until getting the treatment or cure.

Besides at country level, VBDs are also being closely monitored by the World Health Organization (WHO), which publishes their findings and reports at world level. Figure 2.1 below shows the death per million populations caused by vector borne disease in 2002, which was included as an estimate for the World Health Report in 2004. Africa, Asia and South America have been observed as regions with very high number of vector borne disease related deaths.

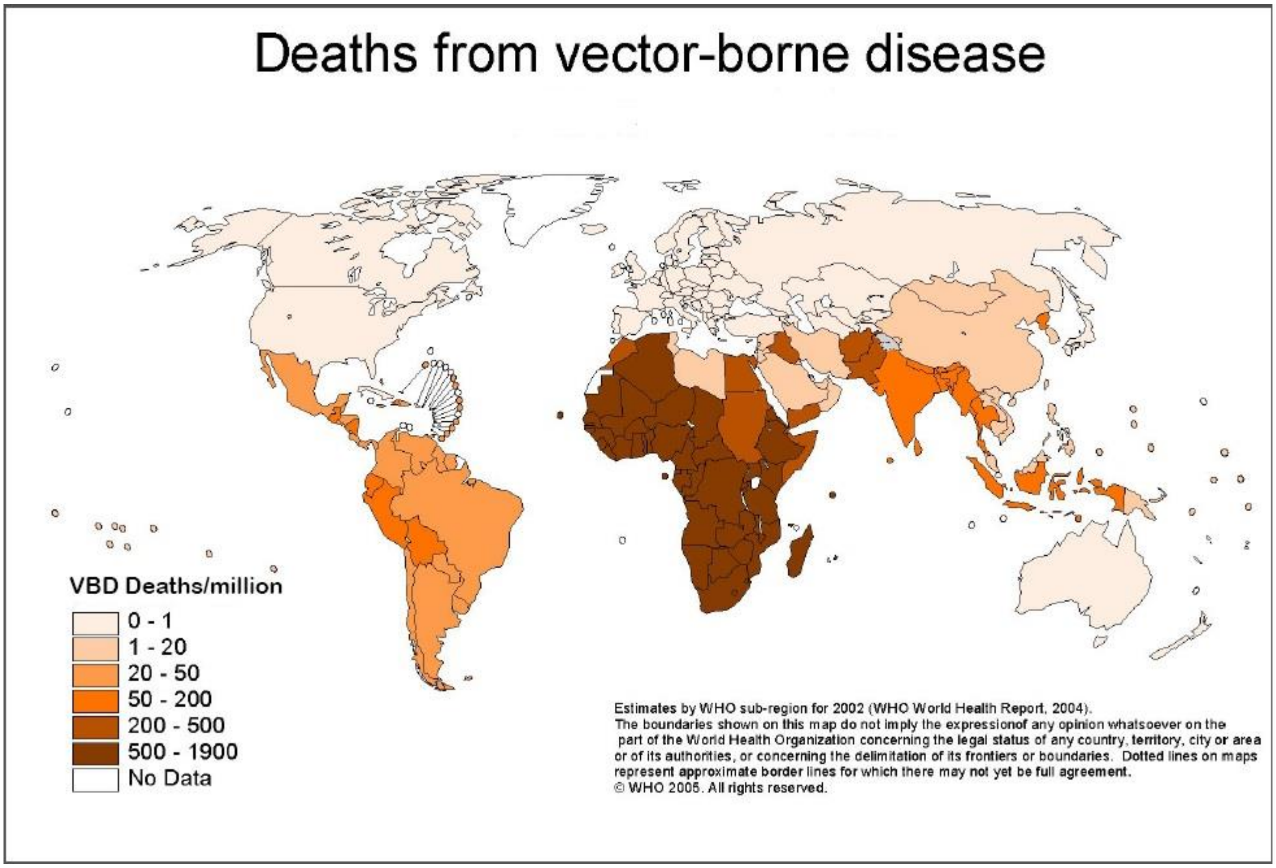


Figure 2.1 Map showing deaths from vector borne disease (WHO, 2005)

In Malaysia, the mortality rate for vector-borne disease is very low or non-existent, ranging from 0 to 0.45 per 100,000 population as recorded in 2020 and mosquito-borne dengue fever contributed to the highest number of incidences at 276.11 per 100,000 population (MOH, 2021). In the latest complete 6-year cycle, 2007–2012, dengue cases peaked from 2007 to 2010 with an annual average of 46,460 reported dengue cases (an incidence of 167 reported cases per 100,000 population) and 104 dengue deaths (Packierisamy et al., 2015). However, there are no specific details, especially on the location or map of vector-borne disease death,available for reference.

Different kinds of vectors may carry different types of bacteria, viruses and parasites that can be transmitted to other animals and also humans. For instance, diseases caused by mosquito bites are malaria, yellow fever and dengue fever while the mosquito’s togaviruses known to infect humans are Ross River and Chikungunya virus. These togaviruses includes others such as O’nyong-nyong, Mayaro and Barmah Forest virus belongs to the genus Alphavirus and Chikungunya virus itself is responsible for extensive urban disease in Africa and Southeast Asia (Lam et al., 2001). Ross River virus is the most prevalent mosquito borne virus in Australia, followed by Barmah Forest virus while malaria and dengue may also become a threat in Australia, with higher threat of dengue if there is an increase in distribution of the vector (Russell, 2010).

Meanwhile, ticks are also the carriers of Lyme disease in the USA, United Kingdom (Marcu et al., 2013) and in Canada (Ogden et al., 2010), as well as causing Tularemia (Brown et al. 2011) and Rocky Mountain Spotted Fever (RMSF) in the USA. In Asia, ticks were discovered to be the cause of the Kyasanur Forest disease outbreak, which was first detected in the spring of 1957 near Kyasanur Forest in Karnataka State in India (Holbrook 2012). These are among a few known life-threatening tick borne diseases, which if left untreated, can be fatal to humans. In Malaysia besides tick infestations such as human otoacariasis, a few deadly viruses such as Langat, Lanjan and Selatar viruses have been reported to be transmitted by ticks. These two important vectors, mosquitoes and ticks are explained further below to gain a better understanding of the species and emphasize how important it is to control, manage and prevent vector borne diseases from becoming an outbreak in the future.

2.3 Vectors of vector borne disease

Vectors of VBD have become one of the most important subjects of research in disease epidemiology. Studies have been carried out to understand vector borne disease transmission and its distribution, identifying species of vectors that can transmit disease and modelling disease based on suspected environmental factors and if certain condition such as climate change have an influence on the disease. Two of the most important vectors are mosquitoes and ticks, which are being described in the sections below.

2.3.1 Mosquitoes

Mosquitoes are the principal vectors of vector borne disease in the world. One of the most deadly mosquito-borne diseases is malaria, which is transmitted by bites from female *Anopheles* mosquitoes. A study in Ghana undertaken by Appiah, Mueller and Cross (2011) has identified that the entire nation's population is at risk of the disease, with the most vulnerable being children under five years of age and pregnant women. Malaria is also the leading cause of morbidity in Ghana where between 3.1 to 3.5 million clinical cases of the disease (38% of all out-patient illness) are reported in public health facilities each year (GHS 2007; MOHCW 2008).

The two mosquito species *Aedes aegypti* and A*edes albopictus* can transmit viruses causing dengue fever (CDC 2012). Currently there is no cure for dengue fever and there are better chances for recovery if patients are treated at the early stages of fever, as the consequences for failure to do so can be fatal.

2.3.2 Ticks

Ticks are arthropods and members of the Class *Arachnida*. They are blood feeders and can transmit diseases to the host, thereby affecting the host's blood and/or lymphatic system and later on causing fever and anaemia. Ticks affect people in two ways, which is by toxic response to their saliva and by transmission of other pathogens. According to Somayaji and Rajeshwari (2007), "when a tick feeds on human blood, the saliva containing toxins is secreted which can cause facial paralysis or respiratory paralysis to the human". The capability of ticks to cause severe toxic conditions such as paralysis and severe allergic reactions (Gauci et al. 1989) as well as their ability to transmit bacteria and viruses to human can become a threat and a major health concern. Most importantly a single tick bite can transmit multiple pathogens, a phenomenon that has led to typical presentations of some classic tick-borne diseases. According to Edlow (2008), ticks can carry and transmit a remarkable array of pathogens, such as bacteria, spirochetes, rickettsiae, protozoa, viruses, nematodes and toxins which are of importance in both human and veterinary medicine (Lindgren 1998a). The rickettsiae are a diverse collection of obligately intracellular Gram-negative bacteria found in ticks, lice, fleas, mites, chiggers and mammals, which may cause tick borne rickettsioses such as Rocky Mountain spotted fever, rickettsial pox, other spotted fevers, epidemic typhus and murine typhus (Walker, 1996). Meanwhile, other important tick borne diseases are human babesiosis, ehrlicihiosis, anaplasmosis, tick-borne relapsing fever, Colorado tick fever and tick paralysis (Gaff and Schaefer 2010). The list of diseases above proves the significant roles of ticks as vectors of vector borne diseases.

2.4 Ticks as a vector for vector borne disease

There are two well-established families of ticks: the *Ixodidae* (hard ticks) and *Argasidae* (soft ticks) and both are important vectors of disease causing agents to humans and animals throughout the world (Vredevoe 2014). The anatomy of both the hard and soft tick families is shown in Figure 2.2 below.

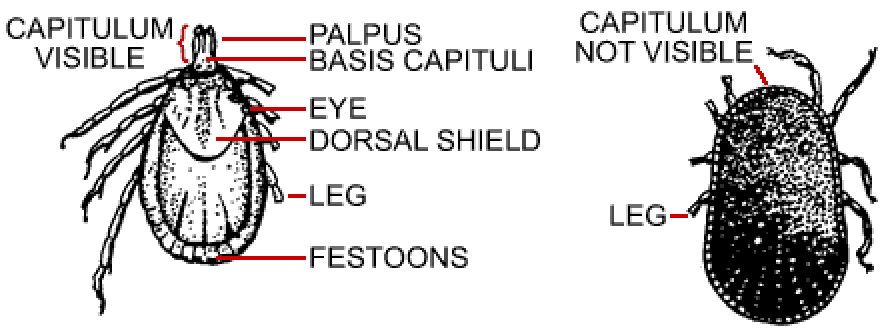


Figure 2.2 The anatomy of hard tick (left) and soft tick (right) (Richards, Rawlings and Hollan 2000)

According to Hay, Packer and Rogers (1997) the biology of ticks is quite different from that of insects: "ticks of the family *Ixodidae* take a blood meal only once per life-cycle stage, as a larva, nymph and adult (Sonenshine 1991). To act as vectors, therefore, ticks must transmit the pathogen between stages within the same generation (trans-stadially) and sometimes from female to the larvae of the next generation via the eggs (trans-ovarially). Between meals ticks drop from their hosts into the vegetation where they undergo long periods of development, lasting between one and twelve months depending on ambient temperature. The survival and development rates of ticks and therefore the transmission dynamics of tick-borne pathogens, are thus directly determined by environmental conditions (Sonenshine 1993)".

The preferred host, or hosts, are also species-dependent and can be different for each life stage, adding complexity to the study of tick-borne diseases (Gaff and Gross 2007).

2.4.1 Tick species

There are many species of ticks found all over the world. However not all species are fully discovered and documented. Due to their very small size, their existence may not be noticed unless one comes in contact with or is being bitten by ticks. For example a common tick called *Dermacentor variabilis* or also known as American dog ticks may easily infect dogs, which is one of the reason the ticks are discovered. There are also many species of ticks that have been identified, however not all of them are identified as vectors of VBDs. Anderson and Magnarelli (2008) have reported in their research that of 878 identified tick species, 222 are known to feed on humans; 28 of which are known to transmit human pathogens (cited in Brown et al. 2011).

Dermacentor variabilis is a good example of tick species that may transmit disease. This is shown in a study by Goethert and Telford (2009) who has identified the American dog ticks (*Dermacentor variabilis*) as an important if not the major factor in perpetuation of *F. tularensis tularensis* on Martha's Vineyard, which is the cause of tularemia disease (Goethert and Telford 2009). Meanwhile in South East Asia, a study by (Hoogstraal et al., 1965), has discovered that the *H. semermis* and *H. papuana nadchatrami ssp. n.* ticks are widely distributed through the forests of certain islands of Indonesia, Borneo, Peninsular Malaya and parts of Thailand, where adult ticks parasitize carnivores, pigs, tapirs, ungulates and also attack humans and domestic dogs.

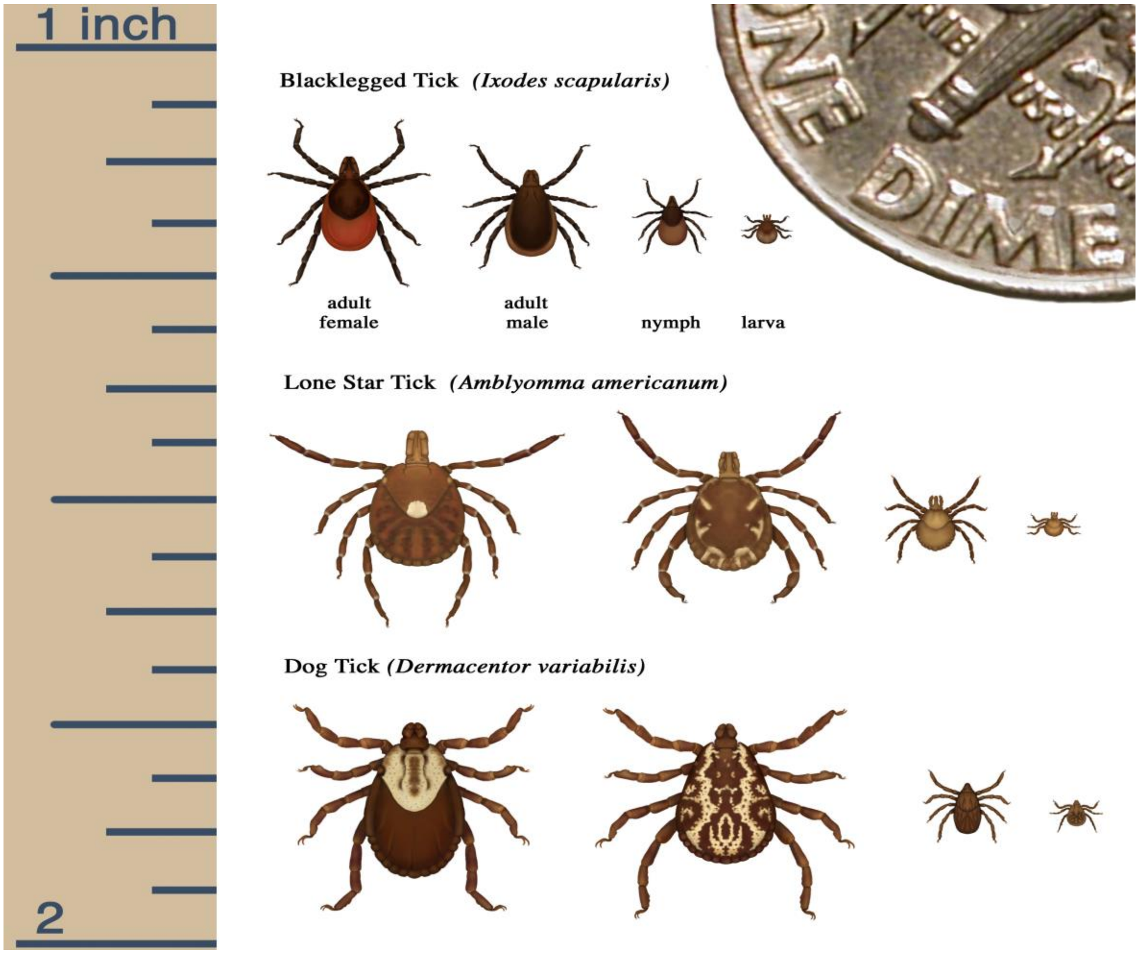


Figure 2.3 Relative sizes of several ticks at different life stages (CDC 2017)

2.4.2 Ticks lifecycle

Ticks may need from one to three types of hosts to complete their lifecycle. When only one host is needed to complete a full lifecycle from the stage of egg until they become adults, then it is known as one host tick. When it requires two or three hosts to complete a lifecycle then they are known as two and three-hosts ticks respectively. The lifecycle of a three-host tick, *Dermacentor variabilis* is shown in Figure 2.3. Meanwhile, it is difficult to identify when exactly a tick would complete its lifecycle as ticks depends on when they feed before they can advance to the next stage. Once they have fed and are replete, the tick detaches and, after dropping from the host, finds a resting place where it can digest its blood meal and moult to the next feeding stage, or enter diapause, a state characterized by reduced metabolism and delayed development (Parola and

Raoult 2001). Then it begins to quest for another host, feed on blood and moult until they turn into adult to lay eggs, reproduce and die.

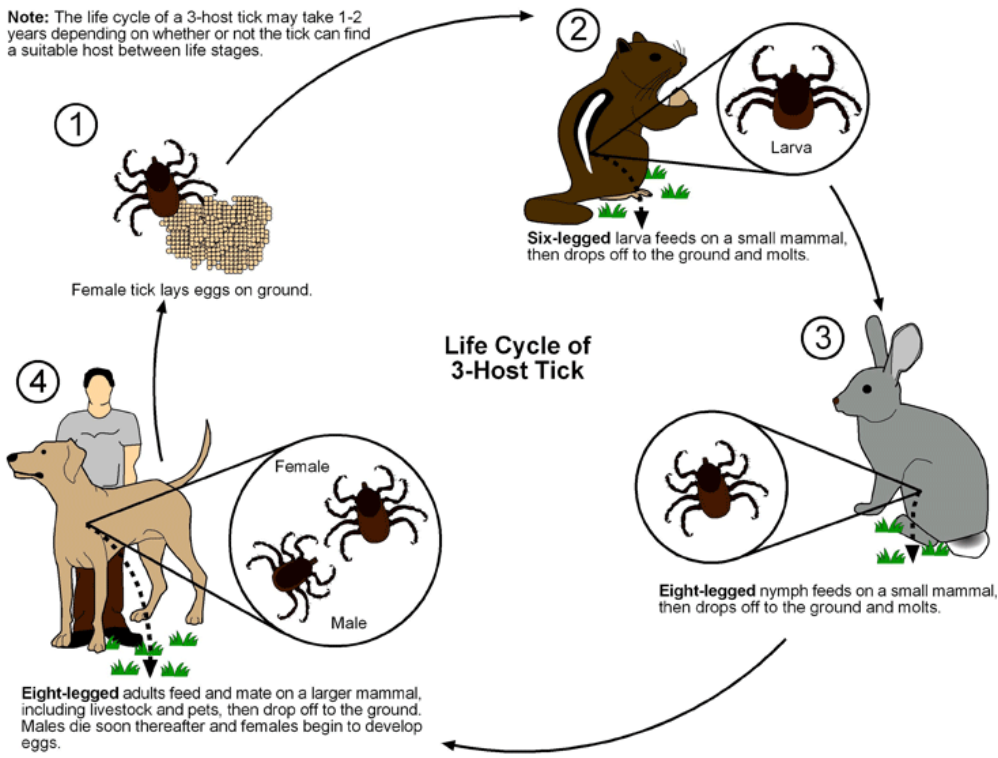


Figure 2.4 Lifecycle of a three-host hard tick based on *Dermacentor variabilis* (American dog tick) Illustration by Scott Charlesworth, Purdue University (2008)

Research conducted by Suss et al. (2008) has summarised that as all parasites do, ticks as *Hematophagous ectoparasites* strive 1) for the best conditions to start and finish their lifecycle; 2) to find a host and take a blood meal; and 3) to transmit pathogens without killing the host. Ticks detect the approach of a host using sensory receptors in Haller's organ, which is located on the tarsus of the first pair of legs and they seek a suitable host by waving those legs in the air ('questing') (Bates, 2012). These hosts are detected based on their shadow, body heat, odour and vibrations caused by their movement (Suss et al. 2008).

2.5 Tick species in South East Asia

Even though ticks can be found almost anywhere in the world, their species may differ among certain geographical locations and environmental conditions. For example, in the United States of America, common species of ticks are from the genus *Ixodide* known to cause Lyme disease. Certain species of ticks have only been found in tropical countries of South East Asia such as Malaysia. There are also species of ticks found in Malaysia, which are unique and can only be found in countries within the South East Asian region. According to studies by Nadchatram (2008), among the species found in Malaysia are those of genus *Dermacentor* and *Amblyomma*. These samples of tick species have been retrieved from the ears of human otoacariasis patients in Kuantan, Pahang and through fieldwork sampling in Peninsular Malaysia.

Local researchers have also performed studies on the ticks, including on identifying how ticks may come into contact with humans. As ticks are blood feeders, they would feed on animals, usually small animals or pigs. They may have accidentally infected humans when the human become contact with them, either in their natural habitats or close by. Ticks usually lay waiting in bush to jump and get attached to a host, which may be small animals or even humans. Samples taken from fieldwork in the Malaysian forests have shown that ticks infest small animals like rats, birds, fruit bats and wild pigs.

Meanwhile, the species of ticks extracted from human otoacariasis patients were mostly of *Dermacentor* species. *Dermacentor* is a three-host tick along with other three-host species ticks such as *Haemaphysalis and Amblyomma.* The three-host ticks drop off following feeding to repletion in each stage of its parasitic life and attach to another host after moulting had taken place away from the host in suitable soil or vegetation. Since *Dermacentor* is the main species found in this study, it will be described further in the sections below.

2.5.1 The *Dermacentor* family

There are a few species of ticks in the *Dermacentor* family, however, not all of the *Dermacentor* species are known to transmit diseases to human. Since there is no complete documentation or background on the species of *Dermacentor* which includes *Dermacentor atrosignatus*, *Dermacentor compactus* and *Dermacentor steini* that are found in Malaysia, the search is for the species with the most similar characteristics. The closest similarity to the Malaysian dermacentor species, in term of characteristics, is found to be with the American dog tick (*Dermacentor variabilis).* These species are of the same *Ixodidae* family or hard ticks and also in the three-host tick category. *Dermacentor variabilis* ticks lay their eggs on the ground, which moult into larvae, after which they latch onto hosts like mice, voles and similar small mammals. As a nymph, *Dermacentor variabilis* feeds on small animals like cats, dogs, oppossums, rabbits and raccoons. Adult ticks of *Dermacentor variabilis* feed on cats, dogs, coyotes, raccoons, horses, cattle, humans and other large mammals while the Malaysian *Dermacentor atrosignatus* has been found to feed on tree shrews, wild boar and fruit bats.

Both American and Asian *Dermacentors* are also known as carrier or vector of bacteria and viruses. The *Rickettsiae* bacteria has been found on *Dermacentor variabilis* ticks (Socolovschi et al. 2009) while Lanjan virus is found on the Malaysian *Dermacentor* ticks. Azad and Beard (1998) have reported that Rocky Mountain Spotted Fever disease is carried by *Dermacentor variabilis.* At the same time, no information on disease is found in the literature for Malaysian *dermacentor*, which can be investigated further.

Environmental conditions preferred by both the American and Asian species are almost the same. *Dermacentor variabilis* prefers grassy meadows and young forests and along roadways and trails, which includes areas with high relative humidity such as dense grassy or herbaceous vegetation with a leaf litter under story (Berrada and Telford 2009). Based on a research by McEnroe (1971) and Carroll et al. (1991), there is convincing evidence that *Dermacentor variabilis* adults can move considerable distances independently and that their movement is orientated toward areas (trails) utilized by potential vertebrate hosts, including humans (Burg 2001).

The *dermacentor* species in Malaysia has been found in areas near orchards, secondary forest, plantations and other types of vegetation. Forest areas are a habitat where ticks are abundant and as proof, thousands of questing male and female ticks and occasionally nymphs were collected from vegetation by Nadchatram (2008). Based on the collection, ticks found were of the following species: *Dermacentor steini, Dermacentor atrosignatus, Dermacentor compactus, Dermacentor auratus, Haemaphysalis semermis, Haemaphysalis nadchatrami* and *Haemaphysalis* *koningsberger*. Individuals of both sexes were collected where they wait on the tips of vegetation for a suitable host to come along to attach to, as they respond to carbon dioxide and movement of the prospective host.

Dermacentor variabilis and *dermacentor* ticks have similarly been found attached to the ear particularly the tympanic membrane and this is known as human otoacariasis (Grady et al. 2011). According to Strickland et al. (1976) and Allan (2001), as cited in Monello and Gompper (2007) *Dermacentor spp.* are known to cause tissue damage, anaemia and paralysis in domestic animals. However, currently there is no information in the literature on effects of tick bites on animals in Malaysia.

Besides D*ermacentor*, other Ixodid ticks (*Acarina*, *Ixodidae*) are also known as transmitters of vectors of many dangerous diseases of humans and domestic animals and among species in Asia found on pigs (*suidae*) are *Haemaphysalis papuana, H. susphilippensis, H. formosensis, Amblyommacyprium*, *Dermacentor auratus*, *Dermacentor atrosignatus, Dermacentor compactus, Dermacentor steini* and *Dermacentor taiwanensis* (Kolonin 2007).

2.5.1.1 *Dermacentor variabilis*

Since the actual *ecology* of the *Dermacentor* species found in Kuantan is unknown, this research will focus at the nearest species of *Dermacentor* with similar characteristics, which is *Dermacentor variabilis* or also called the American dog tick. *Dermacentor variabilis* develops from the egg stage, to the six-legged larva, to the eight-legged nymph and finally to the adult and each cycle requires the ticks to feed on blood as shown in Figure 2.4. Each of the cycle involves ticks feeding on different hosts which requires at least 54 days to complete, but it can take up to two years for the lifecycle to complete depending on the host availability, host location and the temperature (Chan & Kaufman, 2008). However, according to Parola and Raoult (2001), “even though the life cycle of *ixodid* ticks is usually completed in two to three years, it may take from six months to six years to complete depending on environmental conditions, including temperature, relative humidity and photoperiod”.

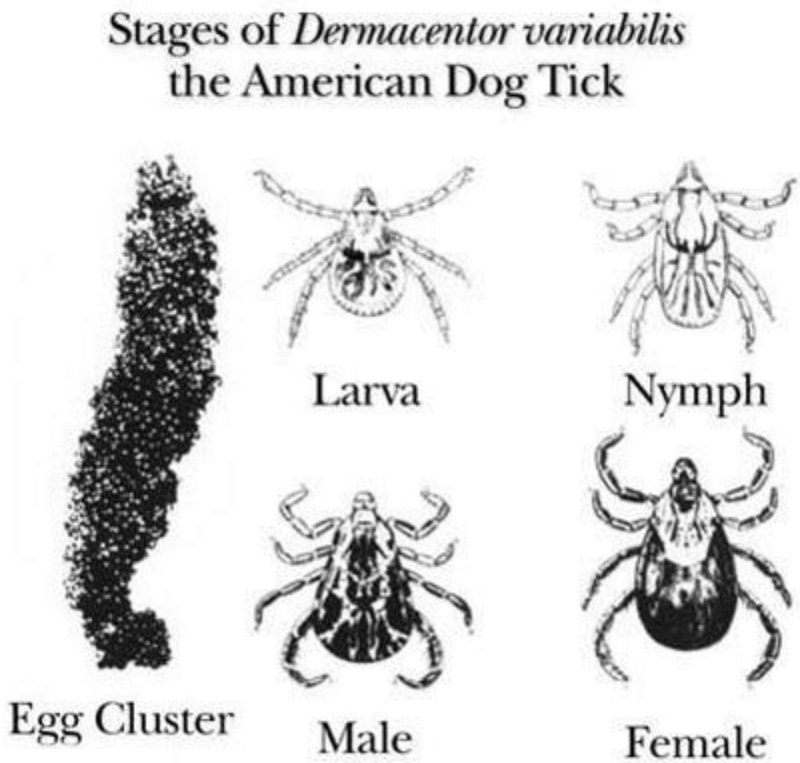


Figure 2.5 The life stages of the American dog tick, *Dermacentor variabilis* (Chan & Kaufman, 2008).

According to Matheson (1950), an female *dermacentor variabilis* takes from five to 14 days of blood feeding before becoming fully engorged and dropping from the hosts, where she then digests her meal, develops her egg clutch over the next four to 10 days before laying from 4,000 to 6,500 eggs on the ground, where they hatch into larvae about 26 to 40 days later, depending on temperature (James et al. 1969), as cited in Chan and Kaufman (2008).

A study by Adams et al. (2002) has stated that non-nidicolous ticks (species like *Ixodes* and *Dermacentor* that occupy open, exposed habitats) will remain attached to vegetation in wait (questing) for a host. According to them, the questing and subsequent orientation to the host for attachment may involve movement, odour, sweat, colour, size, carbon dioxide and other factors.

2.5.1.2 *Dermacentor (Indocentor) atrosignatus* (Neumann, 1906)

This tick species was discovered by Neumann in the year 1906, however neither the nymph nor the larva *of dermacentor atrosignatus* has been described in the findings (Camicas et al. 1998). Based on data collection in Sulawesi, *Dermacentor atrosignatus* has been found on house shrew, lesser tube-nosed fruit bat, Sulawesi rousette, humans, Celebes warty pig, pigs, water buffalo and rats. Elsewhere in Southeast Asia, the adults of *Dermacentor atrosignatus* mainly parasitize wild pigs (*Suidae*) although they also occasionally parasitize domestic pigs, water buffalo, Malayan sun bears, dogs and pangolins (Hoogstraal and Wassef 1985; Petney and Keirans 1996).

Dermacentor atrosignatus is also the probable vector of Lanjan virus, which was discovered in Peninsular Malaysia where it mainly causes pathology in rodents (Karabatsos 1985). Petney and Keirans (1996) have also reported how earlier confusion regarding *dermacentor* species identification in South East Asia has led to uncertainty regarding which species of *dermacentor* is or are vectors of this Lanjan virus.

Other ticks that may transmit pathogens to wild vertebrates in Sulawesi include *Dermacentor spp., Rhipicephalus spp., Haemaphysalis hystricis, H. wellingtoni* and *I. granulatus*. A *Rickettsia* have been detected in pools of *Dermacentor sp*. larvae collected from a wild pig nest in Thailand (Parola et al. 2003) and some *Rhipicephalus spp.* are vectors of the agents of *equine babesiosis* in Asia (Petney and Keirans 1996). Based on data collection in Central Sulawesi (Sulawesi Tengah), van Peenan et al. (1974) has recorded their findings on ticks from several small mammals (mainly bats and rodents) and added their finding of *Dermacentor (Indocentor) atrosignatus* to the Sulawesi faunal list.

2.5.1.3 *Dermacentor (Indocentor) steini* (Schulze, 1933)

Dermacentor steini is one of the *dermacentor* species, which was discovered in the northern Indonesian province of Sulawesi. It is widely distributed in the Indo-Australasian region with records from Vietnam, Thailand, Malaysia, Borneo, several Indonesian islands, the Philippines and New Guinea (Wassef and Hoogstraal 1988; Petney and Keirans 1996).

The only documented collections for Sulawesi are from vegetation in North Sulawesi (Durden and Watts 1988). Durden (1986) and Ibrahim et al. (1999) have both documented the ectoparasites, including ticks, that were parasitizing rodents in North Sulawesi (Sulawesi Utara). Later on Durden and Watts (1988) have also documented ticks from the same Province, including the first Sulawesi records for *dermacentor (Indocentor) steini*. More recent research by Durden, Merker and Beati (2008), has also identified *Dermacentor atrosignatus* and *Dermacentor steini* as two *Ixodidae* (hard ticks) species found in their study.

2.6 Tick borne disease in Malaysia

In Malaysia, ticks have been discovered from the ears of human otoacariasis patients. Human otoacariasis is not a natural happening and human otoacariasis where ticks are involved is not a natural occurrence, but rather an accidental phenomenon when humans living in close contact with tick-infested animals, or living close to the environment where ticks naturally occur, get accidentally infested (Nadchatram 2008). Based on a study by Indudharan et al. (1999) on aural foreign bodies or human otoacariasis, almost half of the 348 cases investigated involved cockroach as the most common arthropod, followed by a cattle tick known as *Boopilus microplus.* However this species is not being the centre of study since *Boopilus microplus* usually infest cattle and only the dermacentor species have been extracted from the ears human otoacariasis patients in the study area.

Based on a report by Lazim et al. (2012), most human otoacariasis patients have complained of experiencing otalgia or ear pain, which is likely caused by the enzymes secreted by the ticks during their attachment to the external auditory canal. They have also documented a case of a woman experiencing facial paralysis due to tick bite. Therefore tick bites do not just cause pain, bleeding and vertigo but may also cause temporary facial paralysis until the tick is removed. Paralysis caused by a tick is usually resolved in 24 hours after its removal where the afflicted person may experience jerky body movements and muscle weakness beginning in the lower extremities and progressing upwards. Breathing difficulties often occur. Typically, the sites of attachment of the ticks are the hairline, the neck region below the hairline, or the back of the body. Human otoacariasis patients may cause intense ear pain, decreased hearing and giddiness and inflammatory reaction of the tympanic membrane after tick removal. Treatment for human otoacariasis is done by removing the tick from patient's ear and the patients will have to take antibiotics.

All genera of ticks incriminated in causing tick paralysis in other countries also occur in Malaysia. For instance, Lam and Chua (2002) were able to determine that fruit bats (*Pteropus hypomelanus*) can serve as a natural reservoir and may be responsible for the Nipah virus outbreak at pig farms. Between September 1998 and June 1999, there was a serious outbreak of viral encephalitis among pig farmers in Malaysia (Lam 2002). The outbreak resulted in 265 cases of acute encephalitis with 105 deaths and leading to a near collapse of the billion-dollar pig-farming industry (Looi and Chua 2007).

Besides Pahang, the neighbouring state of Kelantan has also recorded several cases of the disease (Indudharan et al. 1999 and Ahamad et al. 1996). A single case was reported in the state of Kelantan where a tick has entered the right ear of an elderly lady and has caused a right-sided, isolated, complete, lower-motor-neuron facial palsy (facial paralysis) and this can happen if tick has fed for a few days and if unrecognised or undetected, the paralysis can lead to respiratory failure (Indudharan et al., 1996). According to studies and fieldwork by local researchers, the most common tick species found in Malaysia are *Dermacentor, Haemaphysalis, Ixodes* (Ahamad et al. 2008) and *Amblyomma* ticks (Nadchatram 2008). The most common genus found is *dermacentor* at 99.7%, specifically of *dermacentor atrosignatus, dermacentor compactus, dermacentor steini* and there was only one tick of genus *haemaphysal* found (Ahamad and Nordin 2008) in the study, which also concludes that different ticks are found on different types of hosts and environment. A number of different hosts are involved in tick infections and based on research work in Malaysia, it has shown that wild pigs are suspected to be the main host for intra aural ticks indicated by signs of their frequent visits to patients' houses (Ahamad and Nordin 2008). Later on, Ahamad et al. (2010) has also studied the movement of shrews, which were identified as potential carriers of ticks from the wild into houses and associated compounds.

Based on the report by Nadchatram (2008), during the five-year period of 2002 to 2006 a total of 329 ticks from 318 cases were recovered from human otoacariasis patients seeking treatment at the Tengku Ampuan Afzan Hospital in Kuantan, Pahang. It was mentioned that all active stages from larva, nymph and adult were represented, while 84.4% were nymph and usually one tick was found per case. Meanwhile for seven of the cases, two or three ticks were extracted from a single ear. The ticks were mostly found in the bony part of the auditory canal, followed by tympanic membrane and cartilage part of the ear canal. A few ticks were attached to the outer pinna (outer ear). Though the attachment of ticks caused pain, no facial paralysis was recorded perhaps due to the short period of attachment. Throughout the study there were six repeat cases, where two cases were presented 3 times in a year. The afflicted individuals were mostly children (76.4%), nearly 69.0% were females with the Malay as the dominant ethnic group with 96.6% cases while it was claimed that the states where human otoacariasis were reported were not well developed economically and socially and the afflicted individuals may have lived in environmental conditions close to the habitat of the ticks.

2.6.1 Control and management of tick borne disease in Malaysia

Even though human otoacariasis does not cause death, the Malaysian government is still very concerned about the potential harm from possible diseases that may be transmitted by tick bites. Therefore the government has formed the necessary guidelines to monitor and perform surveillance of communicable and non-communicable diseases in the country. The surveillance of disease is handled at three different levels from the highest level of National Level down to the middle level of State Level to the lowest at the District Office Level. The flow of information between levels is made easier using a web-based reporting system, which currently reports on details of patients. However there is still undue delay in the analysis and interpretation of the data, which hence may cause a delay in response during critical times. With regard to spatial analysis and interpretation, the opportunity exists in the ability to use the latest in Geographic Information Systems (GISs) to ensure that the flow is seamless and the correct information is relayed in real time.

The control and management for tick borne disease in Malaysia is under the jurisdiction of the Disease Control Division, Ministry of Health Malaysia (MOH). The governing law for management of infectious disease is under Act 342, Prevention and Control of Infectious Diseases Act 1988, which incorporates all amendments up to 1 January 2006. At the same time, research on tick epidemiology is being undertaken by the Acarology Unit of the Institute of Medical Research (IMR), in collaboration with the Malaysian Remote Sensing Agency (MRSA) to study tick borne disease and to map the distribution of tick infection cases and hence be able to predict the risk of disease outbreak.

In Malaysia and other Southeast Asian countries, the awareness of the potential of tick borne disease needs to be increased. The public awareness need to be instilled in order for them to be prepared in case of any tick bites incidents, knowing which area should they avoid and how to prevent themselves from being bitten. When the effects of tick borne disease is recognized by the government, the medical practitioners and researchers will also be able to focus more on treatment and research. At the same time the public may be able to provide information to the authorities on the location of suspected tick bites in order for the authorities can take some actions or make the necessary precautions.

2.6.2 The lifecycle of species found in Malaysia

Predicting tick borne disease is more complicated than predicting mosquito-borne disease since tick lifecycle is not fully known as compared to mosquito's lifecycle. As mentioned earlier, it may take up to two years for ticks to complete their lifecycle from eggs to adults, depending on when they are able to feed. In contrast to ticks, mosquitoes have shorter life span and their seasons can also be predicted using climate-based prediction model (Reiter, 2001). Table 2.1 describes the type of ticks, their hosts and types of viruses transmitted.

Species of ticks	Hosts	Virus
Ixodes granulatus	Rodents	Langat
Dermacentor auratus, I. granulatus and H. semermis	Rodents, shrews, large mammals and even reptiles	Lanjan
B. microplus	Cattle	Selatar
Argas pusillis	Bat	Keterah
A. Cordiferum	Not known	Unnamed virus

Table 2.1. Ticks, hosts and their virus as found in Malaysia according to research by Nadchatram (2008)

According to Hoogstraal (1966), the Langat virus may cause human encephalitis, be able to invade other species of ticks and the infection is trans-stadially transmitted from larvae to nymphs and from nymphs to adults while the infected nymphs may also transmit the virus to mice, rats and chicks. As a three-host tick, the female tick lays eggs on the ground and the eggs will hatch into larva. The six-legged larva will feed on small mammal then drops off to the ground and moult into a nymph. The eight-legged nymph will feed on a small mammal before dropping off to the ground and moult into an adult tick. The eight-legged adults then feed and mate on a larger mammal, including livestock and pets, then drop off to the ground. Male tick will die soon afterwards while the female will begin to develop and lay eggs.

Figure 2.5 shows the generalized life cycle of *dermacentor variabilis* and *dermacentor andersoni*, two species of ticks in the same *dermacentor* family found in Malaysia. The characteristics of the Malaysian *dermacentor* ticks follow the same lifecycle as *dermacentor variabilis* and *dermacentor andersoni*. However the difference is only on the type of small animals as the hosts as well as the type of bacteria or viruses that they transmit.

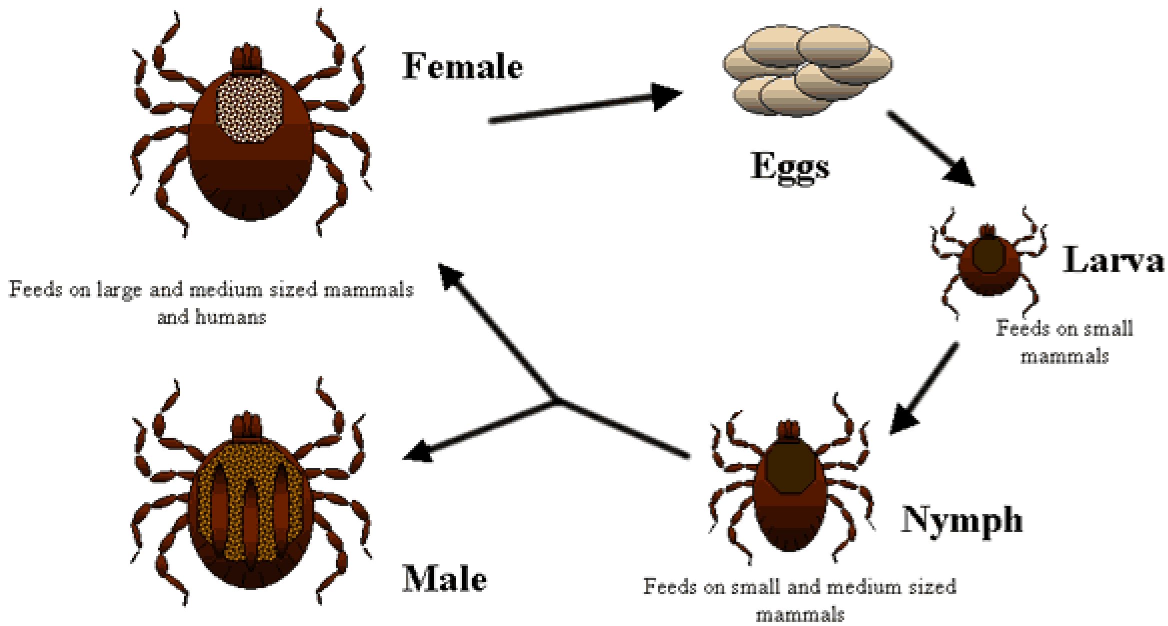


Figure 2.6 The Generalized Lifecycle of *Dermacentor variabilis* and *Dermacentor andersoni* ticks (family Ixodidae) (Mayer 2015)

2.6.3 Tick seasons

There is very little information found in the literature regarding the tick season or period when ticks are found in abundance. A study was conducted by Lim (1972) on the seasonal distribution of immature *Dermacentor* and *Haemaphysalis* spp. ticks for a two-year period in primary and mixed secondary rainforest in Sungei Buloh Forest Reserve, about 26 km northwest of Kuala Lumpur. According to the study, a seasonal peak occurred in July for *Dermacentor spp.* within the mixed-secondary forest while in the primary forest a seasonal peak occurred in August for *Haemaphysalis spp.* and an apparent peak occurred in November for *Dermacentor* species. Meanwhile, a trend of tick infestation has also been detected by analysing patients' records, with a significant peak of number of tick infection cases in January, December and a noticeable peak during the mid-year.

2.6.4 Habitats and preferred environments

Ticks natural habitat is in the forest or areas with high vegetation. According to Civitello, Flory and Clay (2008), there are certain types of habitats preferred by ticks. The habitats favoured by ticks are areas with exotic plants with high human activity, such as along trails, roads and forest edges and in disturbed riparian areas.

Based on field observations, the agricultural activities in the study area are focused on oil palm plantation, fruit orchards and also cattle rearing. There are many activities being carried out, particularly in the oil palm plantation area before, during and after the planting season, as well as during harvesting season. At the same time cattle and sheep are also being reared to control the amount of weed in the plantations as well as for their commercial values, which is to meet the local demand of fresh milk and meat.

In the oil palm plantations, leguminous cover plants are planted following land clearing, which is prior to planting of oil palm itself. Cover plants are important because they help to prevent soil erosion and surface run-off, improve soil structure and palm root development, increase the response to mineral fertilizer in later years and reduce the danger of micronutrient deficiencies. Leguminous cover plants also help prevent outbreaks of *Oryctes* beetles, which nest in exposed decomposing vegetation. The *Oryctes* beetles also known as Coconut Rhinoceros beetles are pest that attack crops such as oil palm and coconut trees.

Three years after oil palm is planted, side crops such as peanuts, corn and vegetables, are planted. Other normal agricultural activities involve pruning of leaves, weeding, fertilizing, spraying of insecticide, assisted pollination and finally, harvesting. All of these agricultural activities have a very high chance to expose humans to tick.

Meanwhile, settlements in the rural areas consist of small villages where some of the houses are located within a short distance from each other. Different types of vegetation, such as fruit trees and shrubs, usually surround these typical village houses. Some villages are located close to the forest fringe or at the foot of a hill and some are located in the coastal areas where they are mostly surrounded by coconut trees. A lot of vegetation is also observed along the riverbanks and there are also swamp reserve areas as well. All of the areas mentioned above are rich in flora and fauna, where there are a lot of human activities and animals alike, which provides a suitable ground for accidental infection since ticks prefer habitats with high moisture and vegetation.

2.7 Chapter summary

When a tick bites, it may not only bring pain and discomfort to humans and animals but most importantly ticks have high potential for transmitting deadly bacteria and viruses to their hosts leading to a few type of tick borne diseases. These tick-borne diseases can be treated if detected early, however it may take years for the victims to recover or in severe cases may result in deaths. Therefore considering their significance role in transmitting diseases, ticks have been recognised as the second most important vectors for vector borne disease.

Based on information found in the literature, a few *dermacentor* species have been extracted and identified from the human otoacariasis patients in the study area. The species are also found in another South East Asian country, which is Indonesia. However, to date there is no complete documentation or descriptions for the tick species that has been discovered in the literature. Since detailed information such as when do the ticks complete their lifecycle is not available, further literature review has been made to find the nearest *dermacentor* species with the most similarities to get better understanding of the ticks background and behaviour. During this review, the American Dog tick or *dermacentor variabilis* has been identified as having similar characteristics as the *dermacentor* species found in the study area. Even though the related information is based on the closest *dermacentor* species, but it has provided some guidance and a better understanding on the species and on how disease transmission have occurred. This information can help identify the significant factors affecting tick-borne disease to incorporate into a spatial model for producing a disease risk map in the Chapter 3.. This map will be most beneficial to the decision-makers in understanding and managing the risk of disease outbreak. The next chapter will explore the geospatial modelling of tick-borne disease and investigate research work carried out previously for tick-borne disease risk modelling and mapping.

3 GEOSPATIAL MODELLING OF TICK BORNE DISEASE

This chapter describes the background to studies on geospatial modelling for vector borne diseases in general and the application of remote sensing and GIS data for tick borne disease management in particular. It also explores the concept of a decision support system, which when combined with the web GIS technology can be used as an important tool for decision makers to make well-informed decisions. This would allow them to manage their resources and make control and prevention of tick borne disease outbreak in the future.

3.1 Remote sensing and public health

Remote sensing data have been used extensively for a wide variety of applications, from monitoring of environmental resources such as land use changes and forest logging, to applications in public health for disease epidemiology, such as mapping disease distribution and prediction of future risks areas. According to Last (2001), disease epidemiology is defined as “the study of the distribution and determinants of health-related states or events in specified populations and the application of this study to the control of health problems”. By visualizing disease distribution on a map, it may help people to see the pattern of disease distribution and the surrounding areas and provide some insight on what could be the suspected factors leading to an infection. Mapping disease distribution is not new since it first occurred about 160 years ago. The first known disease map was produced by the father of modern spatial epidemiology, Dr. John Snow for mapping a cholera outbreak in London in 1854. His theory was published as an essay in 1849, where he described cholera as a communicable disease spread by contamination of the water supply and by acting on his theory of water transmissions; he successfully expedited the end of the Broad Street London epidemic by removing the handle of the local water pump (Ramsay 2006).

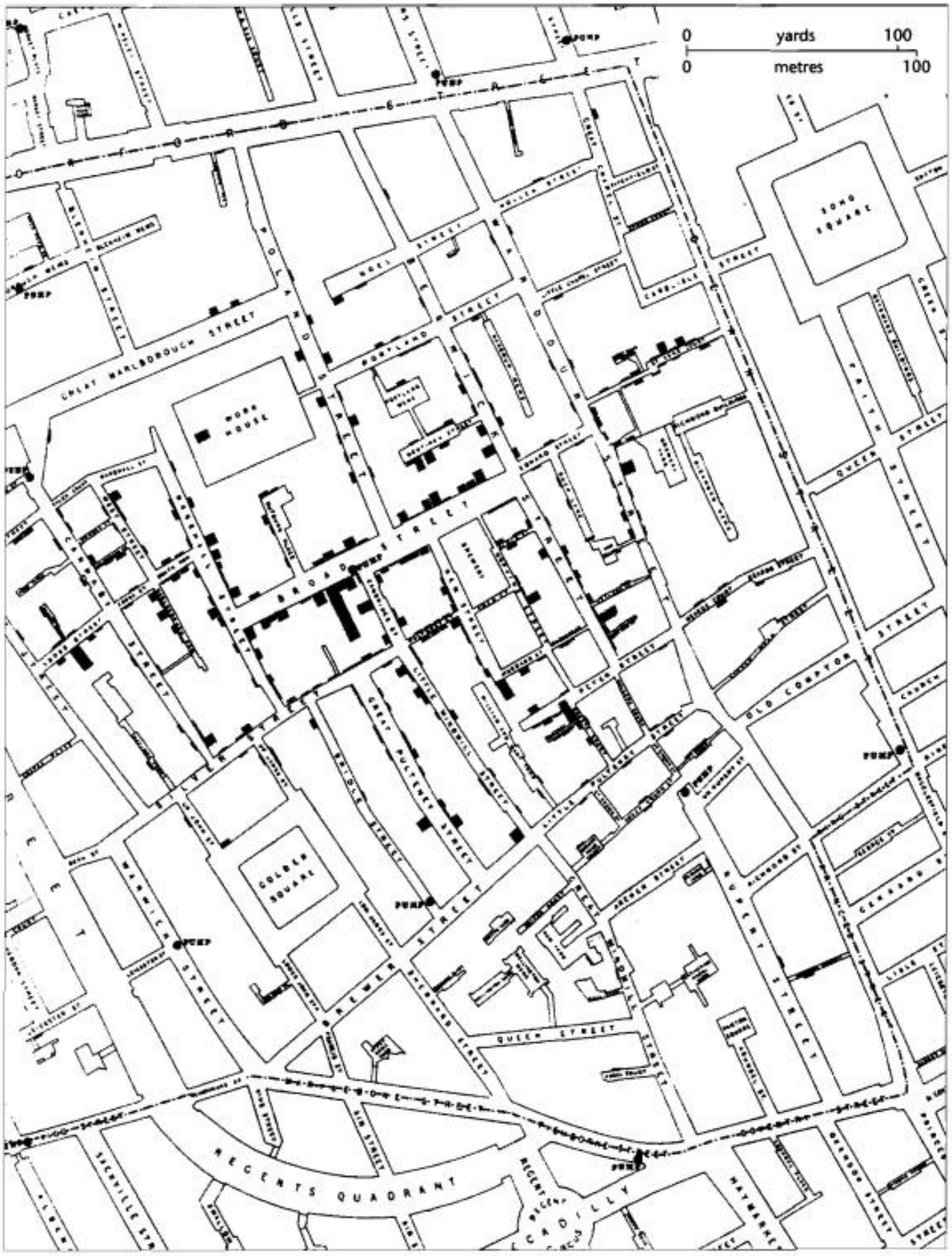


Figure 3.1 The map by John Snow showing the area of cholera outbreak in London epidemic of 1854 (Koch 2004)

Remote sensing (RS) is the process of acquiring information about an object, area or phenomenon from a distance (Hay 2000) while according to Campbell and Wynne (2011), remote sensing is the practice of deriving information about the earth's land and water surfaces using images acquired from an overhead perspective, using electromagnetic radiation in one or more regions of the electromagnetic spectrum, reflected or emitted from the earth's surface. According to Hugh-Jones (1991), “satellite imagery is best considered as a spatially orientated digital dataset obtained from orbiting satellites that can be used to display and compare vegetation classes and biomass, ground temperature and moisture and since each scene or image is acquired at a known time and date, different datasets can be compared to demonstrate changes over time”. Remote sensing satellite data range from high, through medium to lower resolution images. Different type of image resolution can be used for a wide range of applications for earth and natural resources monitoring purposes. The application areas of remote sensing in various disciplines has been broadly summarized by Chandra and Ghosh (2006) in Table 3.1.

Agriculture and forestry	**Land use mapping**	**Geology**	**Water resources**	**Oceanography and marine**	**Environment**
Discrimination of vegetation types	Classification of land use	Recognition of rock type	Determination of water boundaries	Detection of living marine organism	Monitoring of surface mining and reclamation
Measurement of crop species and acreage	Cartographic mapping and map updation	Mapping of major geological units	Mapping of gloods and flood plains	Determination of turbidity patterns and circulation	Mapping and monitoring of water pollution
Measurement of timber acreage and volume by species	Categorization of land suitability	Revising geological maps	Determination of aerial extent of snow and snow boundaries	Mapping of shoreline changes	Detection of air pollution and its effects
Determination of range radiness and biomass	Distinguishing Urban and rural areas	Delineation of unconsolidated rocks and soils	Measurement of glacial fetures	Mapping of shoals and shallow areas	Determination of effects of natural disaster

Agriculture and forestry	Land use mapping	Geology	Water resources	Oceanography and marine	Environment
Determination of vegetation vigour	Regional planning	Mapping igneous intrusions	Measurement of sediment and turbidity patterns	Mapping of ice	Monitoring environmental effects of human activities
Determination of vegetation stress	Management of transport networks	Mapping recent volcanic surface deposits	Inventory of lakes	Study of waves and eddies	
Determination of soil condition	Mapping of land water boundaries	Mapping of landforms	Delineation of irrigated fields		
Determination of soil association		Determination of regional structures			
Assessment of forest fire		Mapping of lineaments etc.			

Table 3.1 Some application areas of remote sensing
Source: Table adapted from Chandra and Ghosh (2006, 8)

Each type of remote sensing satellite sensor has its own spatial, spectral, radiometric and temporal resolution characteristics, which may benefit different types of studies. Spatial resolution means the fineness of the spatial detail visible in an image or the size of the smallest objects that can be identified on an image, radiometric resolution can be defined as the ability for an imaging system to record many levels of brightness, spectral resolution denotes the ability of a sensor to define fine wavelengths and finally temporal resolution which is an important consideration in many application where remote sensing has the ability to record sequences of images, thereby representing changes in landscape pattern over time (Campbell and Wynne 2011). The popularity of remote sensing satellite data as one of the best sources of data for research in public health is due to its capability to cover a much wider area at different spectral (from low to medium to high spectral resolution) and different dates (multi-temporal) at much lower cost as compared to airborne data or ground data. According to de Castro (2007) although available imagery at all levels of resolution may be suitable for public health studies, the spatial resolution should be chosen based on the area being studied, the purpose of the research, the future use of the results and the geographical coverage of the study.

3.1.1 Vector borne diseases and remote sensing

The principal goal of remote sensing in the epidemiology of vector borne diseases is to map the distribution of a disease (often by mapping the distribution of its intermediate host) so that control efforts in endemic situations and intervention strategies in epidemic situations may be most efficiently directed (Hay et al., 1997). The application of remote sensing data in disease epidemiology is not something new as it has been applied effectively for many years in research. Furthermore the emergence and re-emergence of deadly diseases has proven that disease surveillance and monitoring is crucial for control and prevention of outbreaks. To control and prevent outbreaks, it is important to understand the transmission of diseases and their possible vectors. In recent years, cases of deadly and infectious disease such as avian influenza (H5N1) as well as respiratory diseases such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome coronavirus (MERS) have been reported. At the same time the emergence of vector borne diseases caused by mosquito bites such as malaria, yellow fever and dengue as well as tick borne diseases like Lyme disease, Tick Borne Encephalitis (TBE) and Crimean-Congo Hemorrhagic Fever (CCHF) has become a public health concern due to their potential of causing illness to humans and in the worst case scenario, death. As mentioned by Wormser et al. (2006), Lyme disease in the United States is caused by *Borrelia burgdorferi*, which is transmitted by the bite of the tick species *Ixodes scapularis* and *Ixodes pacificus*. Based on information provided by the CDC (2014), TBE is caused by the tick-borne encephalitis virus (TBEV), a member of the family *Flaviviridae* and there are three virus sub-types: the European or Western tick-borne encephalitis virus, the Siberian tick-borne encephalitis virus and the Far eastern Tick-borne encephalitis virus (formerly known as Russian Spring Summer encephalitis virus, RSSEV). Meanwhile the CCHF is a zoonotic viral disease caused by tick-borne virus *Nairovirus* of the family *Bunyaviridae* (Appannanavar and Mishra 2011).

Remote sensing satellites can be used to detect changes in the environment, which can be used to map vectors' species and disease distribution. Vector borne disease research usually involves identifying vectors for disease, their habitat, environmental factors and understanding how, when and where disease transmission may take place. These data can be visualized on a map using specialised software, which can convey meaningful information for useful interpretation of data. The availability of remote sensing data over the years and the ability to provide good quality data with high accuracy has helped to ensure that researchers can depend on the data to do spatial and temporal analysis for prediction of disease by helping people identify possible risk areas. The information will then be able to assist government agencies to plan their resources in case of a disease outbreak. The authorities can use this information to develop a standard operating procedure (SOP) in case of an outbreak, which will enable them to manage their resources as well as to control and prevent disease from spreading.

Remote sensing satellite images, when combined with Geographical Information Systems (GIS) capabilities, can provide public health officials with vital information needed to detect and manage certain disease outbreaks as it is very important for decision makers at all levels throughout all regions of the world to have up to date and relevant information to properly plan, manage and monitor any public health system. According to Daniel, Kolar and Zeman (2004) “remote sensing can be used to describe landscape elements that influence the patterns and prevalence of disease and as an addition, GIS provide tools for modelling spatially their occurrence in space and time”.

To monitor and predict the human health consequences of environmental change, a growing segment of the public health community is advocating the use of satellite imagery to monitor environmental parameters that influence the spatial and temporal patterns of vector-borne diseases and it is anticipated that soon a dedicated remote sensing surveillance capability will lead to the development of a ‘disease early warning system’ for identifying areas of high disease transmission risk and directing control measures (Xue et al., 2008). Consequently GIS applications lend themselves to the development of Early Warning Systems (EWS) needed to permit pre-emptive planning to limit risk and the impact of emerging new diseases (Bergquist, 2011). Early identification is an important first step towards implementing effective interventions to control epidemics and reduce the impact on humans and/or animals (Estrada-Peña et al., 2007).

Remote sensing and GIS techniques offer significant potential for disease detection, monitoring and prevention by predicting areas at risk which depend on the combination of remotely sensed images with appropriate spatial, spectral, radiometric and temporal resolutions and ground truth information analysed with GIS to match information concerning specific patterns in the images with areas of known habitat and/or disease infestation in the ground truth data (Kazmi & Usery, 2001).

3.1.2 Types of remote sensing data for disease monitoring and surveillance

There are several types of remote sensing satellites data, which have been applied extensively in the field of public health, specifically in disease monitoring and surveillance. The popularity of remote sensing data is due to their spatial coverage of the area of interest and the temporal availability of data for studies that would require historical data. The reliability and accuracy of data are also two important criteria for their selection in research. Three remote sensing satellites suitable for disease monitoring and surveillance are:

(a) Landsat

Landsat 1, the first of the Landsat series of satellites was launched into orbit by the United States of America's (USA) National Aeronautics and Space Administration (NASA) in 1972 with the purpose of monitoring the earth natural resources through remote sensing (NASA 2014). To date, a total of eight Landsat satellites have been launched, but only Landsat 7 and Landsat 8 are still operational.

Both Landsat 1 and 2 satellites carried two types of sensors: the Return Beam Vidicon (RBV) and Multispectral Scanner (MSS) with seven spectral bands. The Landsat 3 satellite had the same sensors as both Landsat 1 and 2 but the RBV instrument on-board Landsat 3 had an improved 38 metres ground resolution and used two cameras built by the Radio Corporation of America (RCA), which both imaged in one broad spectral band (green to near-infrared; 0.505–0.750 µm) instead of three separate bands (green, red, infrared) like its predecessors.

Landsat 4 and 5 was equipped with the MSS and a new instrument called Thematic Mapper (TM) whose sensors had improved spectral and spatial resolution. However, the next Landsat series, the Landsat 6 was lost during launch on 5th October 1993. The next satellite, Landsat 7 was launched in April 1999, carrying the ETM+ sensor, which is the enhanced TM. The latest satellite Landsat 8 is equipped with two science instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), providing seasonal coverage of the global landmass at a spatial resolution of 30 meters (visible, Near Infrared or NIR, Short Wave Infrared or SWIR); 100 meters (thermal); and 15 meters (panchromatic). The summarized technical information for each series of Landsat is as described in Table 3.2 below:

Retired sensors (LANDSAT 1,2,3,4 & 5)				
Series	**Sensor**	**Bandwidths**	**Spectral resolutions (μm)**	**Resolution**
Landsat 1 & 2	RBV	Band 1,2,3	3 visible bands	80 metres
	MSS	Band 4, 5, 6 & 7	2 visible bands (4&5) and 2 Near-Infrared bands (6&7)	60 metres
Landsat 3	RBV	Band 1	Green to near-infrared	40 metres
	MSS	Band 4, 5, 6 & 7	2 visible bands (4&5) and 2 Near-Infrared bands (6&7)	60 metres
Landsat 4 & 5	MSS	Band 4, 5, 6 & 7	2 visible bands (4&5) and 2 Near-Infrared bands (6&7)	60 metres
	TM	Band 1,2,3,4,5,6&7	3 visible bands (1,2&3), 1 Near-Infrared (band 4), 2 Short-wave Infrared (band 5 & 7), 1 Thermal Infrared (band 6)	30 metres

a) List and characteristics of retired Landsat satellites

Active sensors (LANDSAT 7&8)				
Series	**Sensor**	**Bandwidths**	**Spectral resolutions (µm)**	**Resolution**
Landsat 7	ETM+	Band 1	Visible blue (0.441 – 0.514 µm)	30 metres
		Band 2	Visible green (0.519- 0.601 µm)	
		Band 3	Visible red (0.631 - 0.692 µm)	
		Band 4	Near Infrared (0.772- 0.898 µm)	
		Band 5	Short-wave Infrared-l (1.547 - 1.749 µm)	
		Band 6	Thermal Infrared (10.31 - 12.36 µm)	60 metres
		Band 7	Short-wave Infrared-2 (2.064- 2.345 µm)	30 metres
		Band 8	Panchromatic (0.515-0.896 µm)	15 metres
Landsat 8	OLI and TIRS	Band 1	Coastal / Aerosol (0.435 - 0.451 µm)	30 metres
		Band 2	Visible blue (0.452 - 0.512 µm)	
		Band 3	Visible green (0.533 - 0.590 µm)	
		Band 4	Visible red (0.636- 0.673 µm)	
		Band 5	Near Infrared (0.851 - 0.879 µm)	
		Band 6	Short-wave Infrared-l (1.566 - 1.651 µm)	
		Band 10	Thermal Infrared-1 (10.60-11.19 µm)	100 metres
		Band 11	Thermal Infrared-2 (11.50-12.51 µm)	
		Band 7	Short-wave Infrared-2 (2.107 - 2.294 µm)	30 metres
		Band 8	Panchromatic (0.503 - 0.676 µm)	15 metres
		Band 9	Cirrus cloud detection (1.363 - 1.384 µm)	30 metres

b) List and characteristics of active Landsat satellites

Table 3.2 Characteristics of retired (a) and active (b) Landsat Satellites and Sensors as compiled from NASA (Irons 2011; USGS 2013; USGS 2014; Irons 2015)

Landsat satellite data have previously been applied for studies on ticks. Hugh-Jones and O'Neal (1986) carried out a study to investigate the potential of Landsat-1 MultiSpectral Scanner (MSS) imagery to identify tick habitats, specifically for tick species of *Amblyomma variegatum* in St Lucia, Caribbean. Through the research, it was discovered that variables such as plant composition, grazing cover, soil type and depth, slope and rainfall, were able to distinguish four main tick habitats which were: 1) lightly infested dry meadows; 2) moderately infested foothills; 3) heavily infested dry scrub; and 4) rocky grasslands. These four habitats could be discriminated remotely and were resolved by unsupervised classification of a 1986 Landsat-TM scene of Grand Terre, Guadeloupe (Hugh-Jones et al. 1988). Later on, Hugh-Jones (1991) extended that study using

Landsat-TM imagery to infer the habitat quality for the same *Amblyomma variegatum* species in Guadeloupe by measuring tick density in 103 cattle herds and applying discriminant analysis.

Features such as temporal, spatial and spectral resolutions are all extremely important factors in determining the choice of a satellite-based sensor for epidemiological and public health studies. Relatively high spatial resolution sensors such as the Landsat Thematic Mapper with 16 days re-visit time have a low frequency of image capture for any point on the Earth. This means that few cloud-free images can be obtained, especially over the equatorial tropics, limiting the possibilities of building complete temporal records of environmental variables (Tatem, Goetz, and Hay 2004). However, when there is such a limitation, other sensors with moderate or coarse resolution such as AVHRR can be used (Hay 2000).

(b) SPOT

The SPOT (*Satellite Pour l'Observation de la Terre*, French for "earth observation satellite") satellite is a high-resolution optical imaging system operating from space, being part of the earth observation strategy of the *Centre National d'Etudes Spatiales* (CNES) (CNES 2009). A programme called tele-epidemiology was carried out under the auspices of CNES to monitor and study the propagation of human and animal diseases (water, air and vector borne diseases), which are closely linked to climate and environmental changes. This concept is currently being applied by CNES and its partners to different infectious diseases including: 1) malaria in urban areas: Puerto Iguazu (Argentina), Dakar (Senegal), Bamako (Mali), Ndjamena (Tchad); 2) malaria in rural areas in Burkina Faso; 3) bilharzia in China; 4) vibrio in the Mediterranean basin; 5) dengue in Argentina and in Martinique Island; and 6) Malaria and Rift Valley Fever (RVF) in Senegal (Lafaye 2014).

To date, there have been a total of 7 SPOT satellites launched to the orbit. However Spot 1, 2, 3, 4 and 5 have been retired and only Spot 6 and 7 are operational and actively collecting data. The SPOT satellites specification is summarized in Table 3.3.

Retired sensors (SPOT 1, 2, 3, 4 & 5)			
SPOT satellite	**Spectral bands**	**Ground pixel size**	**Spectral resolutions**
1, 2 & 3	P: Panchromatic	10 metres	0.51 – 0.73 µm
	B1: green		0.50 – 0.59 µm
	B2: red	20 metres	0.61 – 0.68 µm
	B3: near-infrared		0.78 – 0.89 µm
4	M: Monospectral	10 metres	0.61 – 0.68 µm
	B1: green		0.50 – 0.59 µm
	B2: red		0.61 – 0.68 µm
	B3: near-infrared	20 metres	0.78 – 0.89 µm
	B4: short-wave (SWIR)		1.58 – 1.75 µm
5	P: Panchromatic	2.5 or 5 metres	0.48 – 0.71 µm
	B1: green		0.50 – 0.59 µm
	B2: red	10 metres	0.61 – 0.68 µm
	B3: near-infrared		0.78 – 0.89 µm
	B4: short-wave (SWIR)	20 metres	1.58 – 1.75 µm

a) List and characteristics of retired Spot satellites

Active sensors (SPOT 6 & 7)			
SPOT satellite	**Spectral bands**	**Ground pixel size**	**Spectral resolutions**
6 & 7	P: Panchromatic	1.5 metres	0.450-0.745 µm
	B1: Blue		0.450-0.520 µm
	B2: green		0.530-0.590 µm
	B3: red	6 metres	0.625-0.695 µm
	B4: Near-infrared		0.760-0.890 µm

b) List and characteristics of active Spot satellites

Table 3.3 Specifications of retired (a) and active (b) SPOT satellites. Table updated from CNES (2010) and Airbus Defence and Space (n.d.)

SPOT data has also been used for mapping of Q fever incidence in French Guiana (Tran et al. 2002). Meanwhile Dambach et al. (2009) mapped the risk of malaria in Burkina Faso using a SPOT 2008 image and later used another SPOT 5 satellite image, taken during the rainy season in 2009 for calculating indices by combining the image's spectral bands. Another study on malaria in

Northern Morocco by Adlaoui et al. (2011) has applied SPOT 10m data to produce a land use map to quantify and map malaria transmission risk in the region using biological and environmental data based on a risk model developed by Tran et al. (2008).

(c) MODIS

The Moderate Resolution Imaging Spectroradiometer or MODIS, which was launched by NASA is a key instrument aboard the Terra and Aqua satellites where Terra's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon and in the process, the satellites are able to view the entire Earth's surface at least daily, acquiring data in 36 spectral bands, or groups of wavelengths (Maccherone, 2014). There are three spatial resolutions with their respective bands: the 240 meter for bands one to two, the 500 meter for bands three to seven and 1000m for bands eight to 36. The types of applications and the technical specifications for different bands are summarized in Table 3.4 below:

Primary Use	Band	Bandwidth[1]	Spectral Radiance[2]	Required SNR[3]
Land/Cloud/Aerosols Boundaries	1	620 - 670	21.8	128
	2	841 - 876	24.7	201
Land/Cloud/Aerosols Properties	3	459 - 479	35.3	243
	4	545 - 565	29.0	228
	5	1230 - 1250	5.4	74
	6	1628 - 1652	7.3	275
	7	2105 - 2155	1.0	110
Ocean Colour/ Phytoplankton/ Biogeochemistry	8	405 - 420	44.9	880
	9	438 - 448	41.9	838
	10	483 - 493	32.1	802
	11	526 - 536	27.9	754
	12	546 - 556	21.0	750
	13	662 - 672	9.5	910
	14	673 - 683	8.7	1087
	15	743 - 753	10.2	586
	16	862 - 877	6.2	516

Primary Use	Band	Bandwidth[1]	Spectral Radiance[2]	Required SNR[3]
Atmospheric Water Vapour	17	890 - 920	10.0	167
	18	931 - 941	3.6	57
	19	915 - 965	15.0	250
Primary Use	**Band**	**Bandwidth[1]**	**Spectral Radiance[2]**	**RequiredNE[delta]T(K)[4]**
Surface/Cloud Temperature	20	3.660 - 3.840	0.45(300K)	0.05
	21	3.929 - 3.989	2.38(335K)	2.00
	22	3.929 - 3.989	0.67(300K)	0.07
	23	4.020 - 4.080	0.79(300K)	0.07
Atmospheric Temperature	24	4.433 - 4.498	0.17(250K)	0.25
	25	4.482 - 4.549	0.59(275K)	0.25
Cirrus Clouds Water Vapour	26	1.360 - 1.390	6.00	150(SNR)
	27	6.535 - 6.895	1.16(240K)	0.25
	28	7.175 - 7.475	2.18(250K)	0.25
Cloud Properties	29	8.400 - 8.700	9.58(300K)	0.05
Ozone	30	9.580 - 9.880	3.69(250K)	0.25
Surface/Cloud Temperature	31	10.780 - 11.280	9.55(300K)	0.05
	32	11.770 - 12.270	8.94(300K)	0.05
Cloud Top Altitude	33	13.185 - 13.485	4.52(260K)	0.25
	34	13.485 - 13.785	3.76(250K)	0.25
	35	13.785 - 14.085	3.11(240K)	0.25
	36	14.085 - 14.385	2.08(220K)	0.35

[1] Bands 1 to 19 are in nm; Bands 20 to 36 are in µm
[2] Spectral Radiance values are (W/m^2 -µm-sr)
[3] SNR = Signal-to-noise ratio
[4] NE(delta)T = Noise-equivalent temperature difference
Note: Performance goal is 30-40% better than required

Table 3.4 MODIS applications according to bands by NASA (n.d.)

MODIS data have been very useful in detecting changes on land and in the atmosphere. There are several types of MODIS data products that can be used to retrieve data such as Land Surface Temperature (LST), Vegetation Indices (VI), land cover changes as well as monitoring atmospheric aerosol and water vapour amounts.

MODIS data products have also been used for vector borne disease modeling. For example, LST and NDVI have been derived from MODIS as two bioclimatic variables in developing a spatio-temporal prediction models for arbovirus activity in Northern Australia (Klingseisen et al., 2011). MODIS NDVI and LST data have also been applied by Anyamba et al. (Anyamba et al., 2014) in their study of weather extremes and their impact on crop production in large agricultural regions and how these extremes contributed to vector borne disease outbreaks in diverse regions including the continental United States, Russia, East Africa, Southern Africa and Australia. Meanwhile, MODIS NDVI has also being used in a study to estimate the spatial spread of nephropathia epidemica (NE) transmitted by bank voles (*Myodes glareolus*), a native rodent species and Lyme borreliosis (LB) caused by bites by ticks of the genus *Ixodes* in Belgium (Barrios et al., 2012).

After investigating these remote sensing data such as SPOT 5 data and Landsat data, MODIS data have been identified for this study. The main reason for MODIS data selection is due to the unavailability of SPOT 5 coverage of the area during the study dates while Landsat data are affected by cloud cover and it was almost impossible to find any cloud free data. The MODIS data obtained are from Vegetation Indices 16-Day L3 Global 250m MOD13Q1. However, the coarser resolution MODIS data has its own limitation as MODIS is an optical sensor and cannot observe the surface when cloud cover is present (National Snow & Ice Data Center, 2021). There is also inconsistency in MODIS LST temperatures, therefore difficult to reconcile with post-hoc methods, whereby correcting MODIS data requires a dense network of ground recording stations to obtain accurate temperature estimates (Alonso-Carné et al., 2013).

3.1.3 Future applications of remote sensing satellite images

There are more applications of remote sensing data use in the field of public health, which are yet to be discovered. Remote sensing satellites are able to provide data of increasingly high accuracy and quality and if the data can be analysed quickly, it may be able to provide real time information to the decision makers. By having such critical information at the right time, it will allow the decision makers and health authorities to locate the disease affected as well as manage and control the disease from spreading. The information will also assist them in managing resources, making decision in case of emergency and to predict areas with high risk in the future and make necessary prevention action.

3.2 Spatial analysis and modelling for vector borne disease

Research on vector borne disease involves studies aimed at investigating how and why disease transmission occurs, such as the significant factors that may trigger an infection or outbreak and how to predict disease risk and prevent an outbreak. Every possible factor needs to be investigated such as environmental variables, number of cases, geographical and historical records of when any disease incidence occurs and trying to determine connections or correlations between the suspected factors and disease incidence using statistical analysis. If there is a significant relationship between any of the factors and disease incidence, then the factors can be used in a model to predict the probability of disease occurring in the future. However there is a difference between causation and correlation. Certain factors may show some kind of relationship or pattern with disease incidences, however this ‘relationship’ might just be accidental and may not be the real cause for the disease. Therefore, every pattern or correlation needs to be statistically proven as well as fit in with medical understanding of the diseases to identify the real underlying causes.

Spatial analysis and modelling is an important process for discovering significant and proven factors for disease and hence for monitoring and predicting vector borne disease. Prediction of disease will help the authorities and researchers understand how the presence of certain factors will increase disease risk. It would give them a head start on preparing and managing their resources such as hospitals and health practitioners and apply preventive action such as spraying of pesticides, which will help to control and prevent disease outbreaks and save lives. Prediction may enable the government, especially the health authorities, to perform disease surveillance.

Disease surveillance can be divided into three parts: detecting, understanding and responding to the spread of disease (Blatt 2013).

Spatial analysis components have been successfully added to environmental public health surveillance to enhance the disease prediction and surveillance capabilities for Lyme disease in the USA (Aviña et al., 2011). Time series analyses have been increasingly used in epidemiologic research such as to forecast epidemics of Ross River virus (RRV) disease using the multivariate seasonal auto-regressive integrated moving average (SARIMA) technique for Brisbane, Australia (Hu et al. 2004). A climate-based prediction model has been applied for malaria since analysis has found that the increase in temperature accelerates the rate of mosquito larval development, the frequency of blood feeding by adult females on humans and reduces the time it takes the malaria parasites to mature in female mosquitoes. Increased rainfall creates additional breeding sites for mosquitoes, thus increasing their numbers (Githeko and Ndegwa 2001).

3.3 Spatial analysis and modelling for tick borne disease

Over the past 15 years, there has been a significant increase in the number of attempts to produce vector-borne disease "risk maps" based on the various environmental factors identified from statistical or biological modelling (Kuhn, Campbell-Lendrum and Davies 2004). According to Osnas et al. (2009), "understanding the spatial and temporal distributions of a disease often requires the application of statistical methods to surveillance data to generate a map that describes the variations in risk". The availability of the risk maps will be very useful in assisting health authorities and decision makers to plan their resources, controlling the spread of disease, mitigating relief during the outbreak and getting prepared for any potential of outbreak. However there are different modelling techniques or algorithms that can be used, depending on a number of factors.

According to Myers et al. (2000) there are two types of disease modelling: 1) the statistical approach which requires samples from as wide a range of environmental conditions as possible; and 2) the biological approach which requires details on all the parameters and variables considered to be important in transmission (these may sometimes be estimated by post hoc analysis of disease data sets).

3.3.1 Statistical approach to disease modelling

Modelling methods such as Bayesian modelling, Logistic regression and Decision Tree classification are among the statistical methods applied to tick borne disease studies. At the same time, even if the same factors are found, different models may generate different outcomes since there are different situations in study areas such as different environmental conditions, species of ticks and even the intermediate hosts.

In the USA, the most commonly found tick-borne disease is Lyme disease and both linear regression and discriminant analyses were used to evaluate associations between tick abundance and climate variables in Rhode Island (Brewer, Mather and Mather 2002). Another study by Kalluri et al. (2007) performed a logistic regression analysis resulting in 83.9% classification accuracy while discriminant analysis achieved a 85.7% accuracy to prove that soil order and land cover are the two dominant factors for *Ixodes scapularis* tick presence. These two methods of analyses have also being proposed by Rogers (2006) to model disease distribution. According to Randolph (2000), different types of ticks have different geographical patterns, therefore different modelling techniques are required and the climate-based approach might not be accurate since other non-climatic factors might play a role. Another study by Kiffner et al. (2010) has applied logistic regression analysis for Tick Borne Encephalitis (TBE) in Germany and found that a combination of landscape and climatic variables as well as host species dynamics influence TBE infection risk in humans.

Bayesian networks are also a popular technique and have been applied for modelling of bluetongue virus distribution in Northern Australia (Klingseisen 2000), soil-transmitted helminth infections in Kenya (Pullan et al. 2011) and human schistosomiasis in Africa (Brooker 2007).

Decision trees have been used to predict diseases in slum areas in India (Smitha and Sundaram 2012), for the study of clinical disease most importantly for the prediction of heart disease (Amin, Agarwal and Beg 2013) and for other diseases such as liver disorders (Sug 2012). The same method has also been applied in a study on a decision support system for a web-based heart disease decision support system (Palaniappan and Awang 2007). One type of decision tree is Classification and Regression Trees, or better known as CART. According to Loh (2014), CART is popular among practitioners due to strong reasons such as the interpretability of the tree structures, reasonably good prediction accuracy, fast computation speed and wide availability of software. According to Jones, Conner and Song (2012), one of the major benefits of a decision tree is its ability to use missing data (where part of the data values is missing or incomplete), which can often be as informative as known data, unlike regression techniques, which cannot process this information directly. Therefore it has been applied for their research in describing the occurrence of Lyme disease in Tennesse, USA. The CART method has been used for analysis and control of malaria in Arunachal Pradesh, India (Murty and Arora 2007) using epidemiological and meteorological data. CART has also been used for the analysis and control of the mosquito vector from the species *Culex quinquefasciatus* for Banchroftian filariasis in India (Kumar et al. 2005).

Meanwhile, another popular modelling approach is for predicting species distribution using the techniques of Species Distribution Modelling (SDM).

3.3.1.1 Species Distribution Modelling (SDM) and pseudo-absences

Species distribution models are empirical models relating field observations to environmental predictor variables, based on statistically or theoretically derived response surfaces (Guisan & Zimmermann 2000). According to Stokland, Halvorsen and Støa (2011), predictive species distribution models are commonly based on presence-only observations of species occurrences and a common procedure to overcome the lack of true absence data is to generate pseudo-absence data to facilitate the use of group discrimination methods. According to Franklin (2010), SDM extrapolates species distribution data in space and time, usually based on a statistical model where the model can be a quantitative or rule-based model and, if the fit is good between the species’ distribution and the predictors that are examined, this can provide insight into a species environmental tolerances or habitat preferences.

Miller (2010) has also stated that beyond predicting species distributions, SDM models have become an important and widely used decision-making tool for a variety of biogeographical applications, such as studying the effects of climate change, identifying potential protected areas, determining locations potentially susceptible to invasion and mapping vector-borne disease spread and risk.

According to Senay, Worner and Ikeda (2013), there are three main methods that are currently used to generate pseudo-absence points, which are: 1) randomly generated pseudo-absence locations from background data; 2) pseudo-absence locations generated within a delimited geographical distance from recorded presence points; and 3) pseudo-absence locations selected in areas that are environmentally dissimilar from presence points.

There are instances when it is not easy to differentiate between areas with infection cases and areas with no infection cases. Therefore a pseudo absences method is selected to observe the difference between areas where infection cases is detected (present) and areas where no infection has been detected (absent). This technique is described as one of the techniques for mapping of species distribution to understand if the species prefer a certain type of condition or area. VanDerWal et al. (2009) has suggested in their study of 12 rainforest vertebrates from the Australian Wet Tropics (AWT), that pseudo-absences are meant to provide a comparative data set to enable the conditions under which a species occurs to be contrasted against those where it is absent and the selection of pseudo-absences at large distances from known occurrences may be problematic. Barbet-Massin et al. (2012) created two geographically distinct virtual species in their study to make sure that the results of study is not influenced by the choice of a species and the peculiarities thereof and suggested using the same number of pseudo-absences as available presences for classification techniques such as boosted regression trees, classification trees and random forests. There are different types of modelling approach, which are able to utilize the pseudo-absence technique. For example, the pseudo-absences technique has been applied for species distribution modelling of native New Zealand fern by Zaniewski, Lehmann and Overton (2002) using a GAM Model. Another application for modelling species distribution of a threatened endemic Iberian moth species (*Graellsia isabelae)* has been applied by Chefaoui and Lobo (2008) using a GLM-logistic regression model in mainland Spain and the Balearic islands. Stokland, Halvorsen and Støa (2011) have also used Boosted Regression Tree (BRT) using pseudo-absences to model the distribution of two species of wood-decaying fungi belonging to the group polypores (*Fomitopsis rosea* and *Xylobolus frustulatus*) and two species of longhorn beetles with larval development in decaying wood (*Leptura maculata* and *Anoplodera sexguttata*) that are associated with forests and dead wood of particular host tree species in Norway.

For the purpose of this study, presence data, which are also known as occurrence data are obtained from the data on the location of infection. The location of infection is observed from records of human otoacariasis patients who were bitten by tick(s) and was obtained when they were getting treatment at the hospitals. This data have been converted into 'points' data in GIS. While the pseudo-absence data will be areas where there are no record of infection. Since the study involves analysing two types of data, which are categorical data for example land use data, and continuous data such as EVI, which are derived from remote sensing data, the best modelling technique that can be applied is the Classification and Regression Tree (CART). The technique for selection of pseudo absence data will be described further in the research methodology section of Chapter 4.

3.3.2 Biological approach to disease modelling

The biological approach for disease modelling is the best method to apply when the necessary data for significant variables is available. For tick borne disease, one of the most consistently significant variables for predicting tick distributions, the remote sensing Normalized Difference Vegetation Index (NDVI), has a sound biological basis in that it is related to moisture availability to free-living ticks and is correlated with tick mortality rates (Randolph 2000).

The biological modelling approach has been used to study tickborne disease transmission between cattle and ticks in Africa which have been identified as hosts for tick borne diseases such as East Coast fever (ECF) or theileriosis caused by the protozoan *Theileria parva* and transmitted by the ixodid tick *Rhipicephalus appendiculatus* (also commonly known as the brown ear tick), heartwater caused by the protozoan *Cowdria ruminantium* and transmitted by the ixodid tick *Amblyomma hebraeum* and babesiosis caused by the protozoan *Babesia bigemia* (Mwambi, 2002).

Climate-based prediction has also been used to predict tick-borne encephalitis (TBE) in Sweden where analysis has detected possible relationships between seasonal variation in temperature, precipitation and changes in TBE incidence in a highly endemic region in the country (Lindgren 1998a). Therefore a prediction model was developed for projections of possible changes in the incidence of tick-borne encephalitis (TBE), including subsequent changes in vaccination needs during the next half-century in Sweden (Lindgren 1998b). The ideal climate-driven model of a disease system would involve a biological approach where the dynamics of both vector and pathogen are modelled explicitly (Brownstein et al., 2003). The effect of climate change has also been shown to influence the distribution of *Dermacentor reticulatus* (Gray et al., 2009).

Other potential modelling techniques include Generalized Linear Models (GLM), Generalized Additive Models (GAM) and Fuzzy Rule-Based Methods. The Generalized Linear Models (GLM) has been applied for tick borne study in Prague, Czech Republic by(Jiruše et al., 2004) to find covariates which influence the risk of tick borne infection while a Bayesian approach was used to estimate the parameters of the risk model. (Guernier et al., 2004), have also used generalized linear multivariate models and Monte Carlo simulations by conducting a series of comparative analyses to test the hypothesis that human parasitic and infectious diseases (PIDs) exhibit the same global patterns of distribution as other taxonomic groups. The GLM method has also being used by (Das et al., 2002) to model a discrete spatial response using generalized linear mixed models for application to Lyme disease vectors in the USA.

The Generalized Additive Models (GAM), or in this case Generalized Additive Mixed-effect Models were used by (Wyk et al., 2014) to model the infectious status of tick borne pathogens against each haematological parameter, including significant interactions between pathogens of East African Short-horn Zebu calves. A modelling technique called Fuzzy logic modelling has been used by Estrada-Peña and Venzal (Estrada-Peña & Venzal, 2007) in a framework to evaluate the impact of ticks on human health under various scenarios of climate change, such as when the mean temperature in north-western Spain has increased or if the Mediterranean region temperature has decreased.

For this study, a suitable model for prediction of tick borne disease can be selected when significant factors of tick borne disease have been identified. The model will be able to produce a disease map which then can be an important feature for the web-enabled spatial decision support system (SDSS) which will be described in the next section.

3.4 Web GIS and Web-enabled SDSS

Technological advance in computing has also promoted the use of research applications for vector borne disease. One of the most popular technologies is known as Geographical Information Systems (GIS). GISs are computer systems for acquiring, storing, interpreting and displaying spatially organized information (Green and Bossomaier 2002) as well as for interpreting and visualizing geographically related data as maps, reports and charts (Florea et al. 2009). GIS can assist in this study as it has strong capabilities for mapping and analysing not only spatial data, but also non-spatial data and can integrate many kinds of data to greatly enhance disease surveillance (Gao et al. 2008a). GIS helps decision making by allowing data to be organised and viewed efficiently, by integrating it with other data, by analysis and by the creation of new data that can then be operated on in turn (Heywood, Cornelius and Carver 2006). It is also used by public health organizations such as the Centers for Disease Control and Prevention (CDC), USA to study the spread of disease and by health planners, business managers, large health organizations and insurance companies (Lang 2002). At the same time, the continual advance of Internet technology makes it the chosen medium for publishing and distributing GIS data (Lang 2002). Recent rapid developments of the Internet such as the integration of World Wide Web (WWW) and telecommunications technologies into one system has permitted instant and worldwide availability to individuals of both personal and public information, as well as their own contribution to both types of information (Kellerman 2014). This promotes the popularity of Web-based GIS, which itself shows great potential for the sharing of health information through distributed networks (Gao et al. 2008b). Web-based GISs can have several advantages including: worldwide access, a standard interface, cost effectiveness and a structure that doesn't require central data collation (Green and Bossomaier 2002).

Early web GIS technologies were not integrated and were inflexible to change but nowadays they are more open and scalable, using independent programming languages and operating systems (Fu 2010). Web GIS can also provide a robust infrastructure to support collaboration across multiple spatial and time scales (Mason and Dragićević 2006).

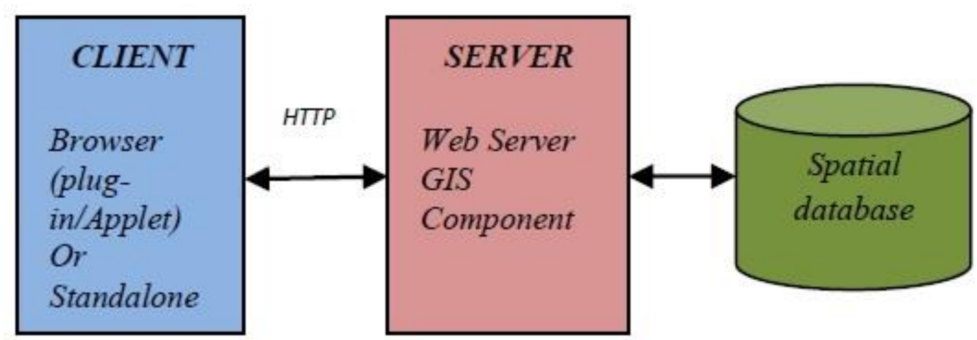


Figure 3.2 The minimum architecture of a Web GIS system (Piarsa et al., 2012)

The minimum architecture of a Web GIS system is shown in Figure 3.2. The concept as explained by Piarsa, Sudana and Gunadi (2012), is described below:

1. The application is client-sided, communicating with the data-providing server through web protocols such as Hyper Text Transfer Protocol (HTTP). To display and interact with GIS data, a browser requires a Plug-In or Java Applet or even both to enhance the functionality of the browser.
2. The Web Server is responsible for processing requests from the client and sending the response back. In the web architecture, a web server also manages the communication with server side GIS Components.
3. The server side GIS Components are responsible for the connection to spatial databases and translating queries into SQL, for example, for extracting the appropriate information from the databases.

There are two main types of Web GIS architecture, namely thin client and thick client. The difference between the architectures is determined by whether the processing is being performed by the client or the server. The benefits of a thin client side architecture are that processing is done by the Server where it is easier to update, cheaper, easier to deploy to many users and easy to start. However it can reduce response time, doesn’t take into account local needs and requires advance data formats. In thick client architecture, data and logic reside locally on the client side, resulting in a lower frequency of data transmissions, as data can be stored at the client. However high internet speed is needed as the server sends raw data which must be processed by the client itself (Agrawal & Gupta, 2014). The thin client architecture is highly recommended for Web GIS users with roles of decision makers and users with less technical skill as they do not need to perform heavy or technical processing work while having the ability to do simple analysis and query. It also allows analysis to be carried out on resource poor mobile devices such as tablets and smart phones.

3.4.1 Web GIS applications in public health and disease management

GIS has become an important tool in public health for disease management. GIS applications such as Geolocation can be used to identify the location of a disease, tracking of disease movement and spread, visualization and mapping of a disease. Meanwhile, the rapid development of the Internet influences the popularity of Web-based GIS, which itself shows great potential for the sharing of disease information through distributed networks and through distributing and sharing disease maps via the Web. It has the potential to help decision makers across health jurisdictions and authorities collaborate in preventing, controlling and responding to a specific disease outbreak (Gao et al. 2008a). Recently GIS and web-based/web-enabled GIS have been widely used for mapping and publishing maps online. By getting maps published online, information can be made available for access to anyone who has access to the Internet. However there are still some restrictions on maps, which need to be considered including issues related to access control and security.

The terms online GIS, web GIS and web enabled GIS have been widely used to describe a system able to publish online maps as well as having the capabilities for data overlay, querying and statistical analysis for generating maps and reports. Web based GIS has been successfully implemented for systems such as the EpiScanGIS for meningiococcal disease surveillance in Germany and has the potential to be extended to an international level and used for other infectious diseases (Reinhardt et al., 2008). The system uses GIS to present maps showing the distribution of all cases of meningococcal disease and acts as an early warning system that provides user with easy to access and timely information on the diseases (Reinhardt et al. 2008). The University of Mississippi Medical Center (UMMC) has also developed a web-based real-time syndromic surveillance system called GeoMedstat which has geographic information system disease mapping capabilities as a control measure against emerging disease outbreaks or bioterrorism attacks (Li et al., 2005).

3.4.2 Decision Support Systems (DSS)

Decision Support Systems (DSS) are interactive, computer-based systems that aid users in making judgment and choice (Druzdzel & Flynn, 2002). The availability of DSS to decision makers will provide them with the capabilities to make informed decisions based on a collection of information. Decision support systems have been used to help decision makers make informed decisions using accurate and timely information.

3.4.2.1 DSS architecture and framework

A DSS is generally composed of: 1) a database which contains relevant data and is managed by a database management system (DBMS); 2) models including statistical, economic, spatial, management, science or other quantitative models that provide the system with analytical capabilities; and 3) a user interface through which the users can communicate with each other and share information within the DSS (Crauwels et al. 2001). DSS is the area of the information systems (IS) discipline that is focused on supporting and improving managerial decision-making. Essentially, DSS is about developing and deploying IT based systems to support decision processes (Arnott and Pervan 2007).

DSS consists of a user interface, database and model base, which effectively supports users in making decisions about semi-structured spatial problems and is characterized by an interactive interface between GIS and users with the role of decision makers (Rinner and Jankowski 2002), which helps them to make decisions in semi-structured and unstructured situations where the decision process is both subjective and objective (Alter, 2002). The integration of both GIS and DSS technologies has resulted in Spatial Decision Support Systems (SDSSs), which harness the decision and analysis power of DSSs and the spatial capabilities of GISs (Sugumaran and Sugumaran 2005). At an organizational or strategic level the most important aspect of GIS output is its use in spatial decision support, whereby it provides a much needed framework for addressing, supporting and making spatial decisions (Heywood, Cornelius and Carver 2006).

DSS has been found to be useful in assisting the decision making process in a wide range of areas such as:

1. *Public health including disease surveillance*

   DSS for monitoring of animal disease, specifically classical swine fever (CSF) in the Netherlands to improve the quality of decision making by modelling all alternatives and by forecasting their contributions to the goal (Crauwels et al. 2001).

2. *Agriculture*

   DSS has been used for the control of apple scab disease and management of fungicide usage (Carisse et al. 2010) and also fish disease and health management (FIDSS), an expert system for fish disease diagnosis in Beijing (Xiaoshuan et al. 2009).

3. *Medical emergency*

   Application of DSS in medical emergency includes enabling stakeholders such as hospitals and firefighters to create Standard Operating Procedures (SOPs), which can be applied during exceptional events such as medical disasters (Herrmann, Kothari and Shaikh 2007).

4. *Environmental monitoring*

   DSS has also been applied for the purpose of managing, analysing and optimizing forest fire fighting strategies in Southern Europe (Bonazountas et al. 2007). While at the same time it has helped in the evaluation of urban air pollution control options in Thessaloniki, Greece (Vlachokostas et al. 2009).

3.4.2.2 Spatial Decision Support System (SDSS) for tick borne disease

The integration of both GIS and DSS technologies has resulted in SDSS, which harnesses the decision analytic power of DSS and the spatial capabilities of GIS (Sugumaran and Sugumaran 2005). SDSSs are integrated computer systems that support decision makers in addressing semi-structured or unstructured spatial problems in an interactive and iterative way with functionality for handling spatial and non-spatial databases, analytical modelling capabilities, decision support utilities such as scenario analysis and effective data and information presentation utilities (Sugumaran and DeGroote 2010). SDSSs are integrated frameworks designed to help explore weakly structured or unstructured problems characterised by many actors, many possibilities, and high uncertainty (Rutledge et al. 2007). It is also able to help health authorities and decision makers at all levels to make informed decisions.

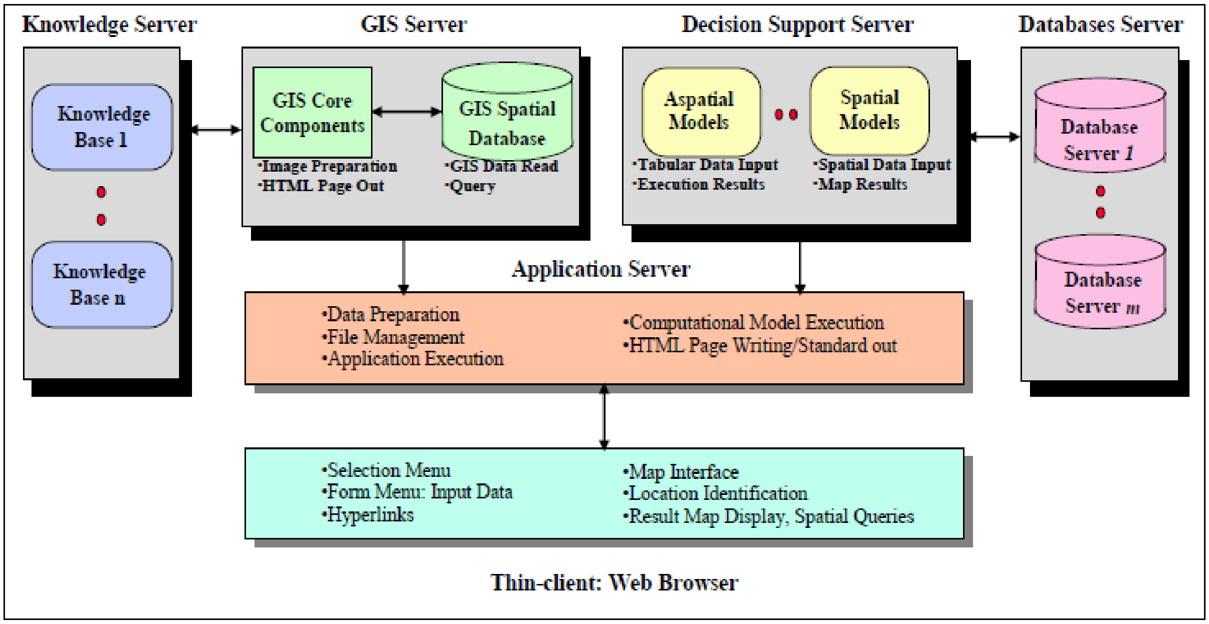


Figure 3.3 Schematic Representation of Web SDSS Components (Sugumaran and Sugumaran 2007)

Figure 3.3 shows the basic components in a Web SDSS architecture consisting of Knowledge Server, GIS Server, Decision Support Server, Database Server and Application Server. Users may access the system using a web-browser user interface where they will be able to make selections from a menu, input data and perform queries or requests. These requests or queries will be submitted via HTTP to the Application server, which will process the request accordingly from the four types of servers: the Knowledge Server, GIS server, Decision Support Server and the Database Server. The Knowledge Server may contain rules that enable the user to select the appropriate type of model to use for a particular task and perform sensitivity analysis, organizational policies, procedures, business rules and constraints that may be relevant for the problem at hand (Sugumaran and Sugumaran 2007). The GIS Server consists of a GIS Spatial database and GIS Core Components, which will be able to handle GIS data query and generating output as a map. The Decision Support Server's function is to execute the modelling request by users and perform processing of spatial model and be able to produce the output in the form of a table or map. Meanwhile, the Database Server stores and manages data for spatial models. This architecture can serve as a basis for developing a Web enabled SDSS for tick borne disease and will be described further in Chapter 6.

3.5 Chapter summary

Remote sensing satellite data have proven to be one of the most popular and best sources of data for applications in tick borne disease monitoring and surveillance due to its capability to provide good quality spatial and temporal data. Both remote sensing and GIS technology can provide the necessary tools to analyse possible factors that may cause disease incidence, which if not controlled effectively may lead to an outbreak. These identified factors can then be used in modelling for prediction of disease risk. Thus, the authorities will be able to perform control measures and prevention of disease in the future.

At this moment, there are two types of modelling approach, the biological approach and statistical approach, which have been described earlier in this chapter. The biological approach is best suited for studies where all or almost all of the biological background and data on disease and species are known. For the statistical approach, there are choices that can be made regarding the modelling techniques or algorithms, depending on the type of data and the suitability for the certain type of modelling. In this case, the best modelling techniques to choose will be the technique that is able to produce the most accurate result to predict tick borne disease occurrence based on significant factors. Consequently, such information can be integrated into a decision support system, which would help decision makers to make informed decisions on the possible risk of an outbreak in the future. The Decision Support System (DSS) may help managers and health authorities in managing and controlling disease outbreak and also predicting risk so that preventive measures can be taken.

At the same time, by enabling viewing of spatial maps and the decision support system to be accessible through the World Wide Web (WWW) will enhance the system as a Web-enabled Spatial Decision Support System (SDSS). Besides providing reliable, accurate and crucial information to the decision makers, the Web-enabled SDSS will also be able to function as a medium for interaction between different levels of users with different level of needs for information such as public users. It can also be used as a medium to disseminate information to the public and may also invite feedback and allow public participation as an input to the system.

4 STUDY AREA AND RESEARCH METHODOLOGY

The previous chapters have covered the background information on ticks, disease modelling techniques and applications of web GIS and decision support systems for the management of vector borne disease. Subsequently, this chapter will continue to describe the background of the study area including the significant reason it was selected for the study. Related data on the study area have also been identified and prepared based on literature review on possible factors of tick borne disease based on previous research. In addition, this chapter will discuss the selection of a modelling technique for prediction of tick borne disease. The results from disease modelling will be able to assist people, especially researchers and government decision makers, to understand how the significant factors may help to predict tick borne disease, allowing them to manage and control any incidence, thus be able to prevent any possible outbreak in the future.

4.1 Study area

This research is focussed on Kuantan District as the study area. This study area covers an area of 2453 km^2 and it is one of the administrative districts of the state of Pahang, the largest state in Peninsular Malaysia. This district in particular has been selected as a study area since it has recorded a very high number of human otoacariasis cases, with a total of 395 cases from January 2002 until December 2007. There are six sub districts in the Kuantan district namely Penor, Sungai Karang, Hulu Kuantan, Beserah, Hulu Lepar and Kuala Kuantan which can be seen in Figure 4.1. The total population in the district according to year 2010 census is 427,515 (Department of Statistics 2011a), consisting of Malay, Chinese and Indians as the three biggest ethnic groups (Department of Statistics, 2011).

Kuantan town itself is the capital city of Pahang and also the state’s centre of administration. In the year 2009, Kuantan town and Pahang state in general were included in the Malaysian government initiative to develop an East Coast Economic Region (ECER) which also comprises of the states of Kelantan, Terengganu and the Mersing District in Johor as a gateway for trade and industry in the Asia Pacific Region by the year 2020 (ECERDC 2010). Kuantan has also been selected as a Special Economic Zone (SEZ) to be developed as an urban region under the National Regional Growth Conurbation in the National Physical Plan (John 2011).

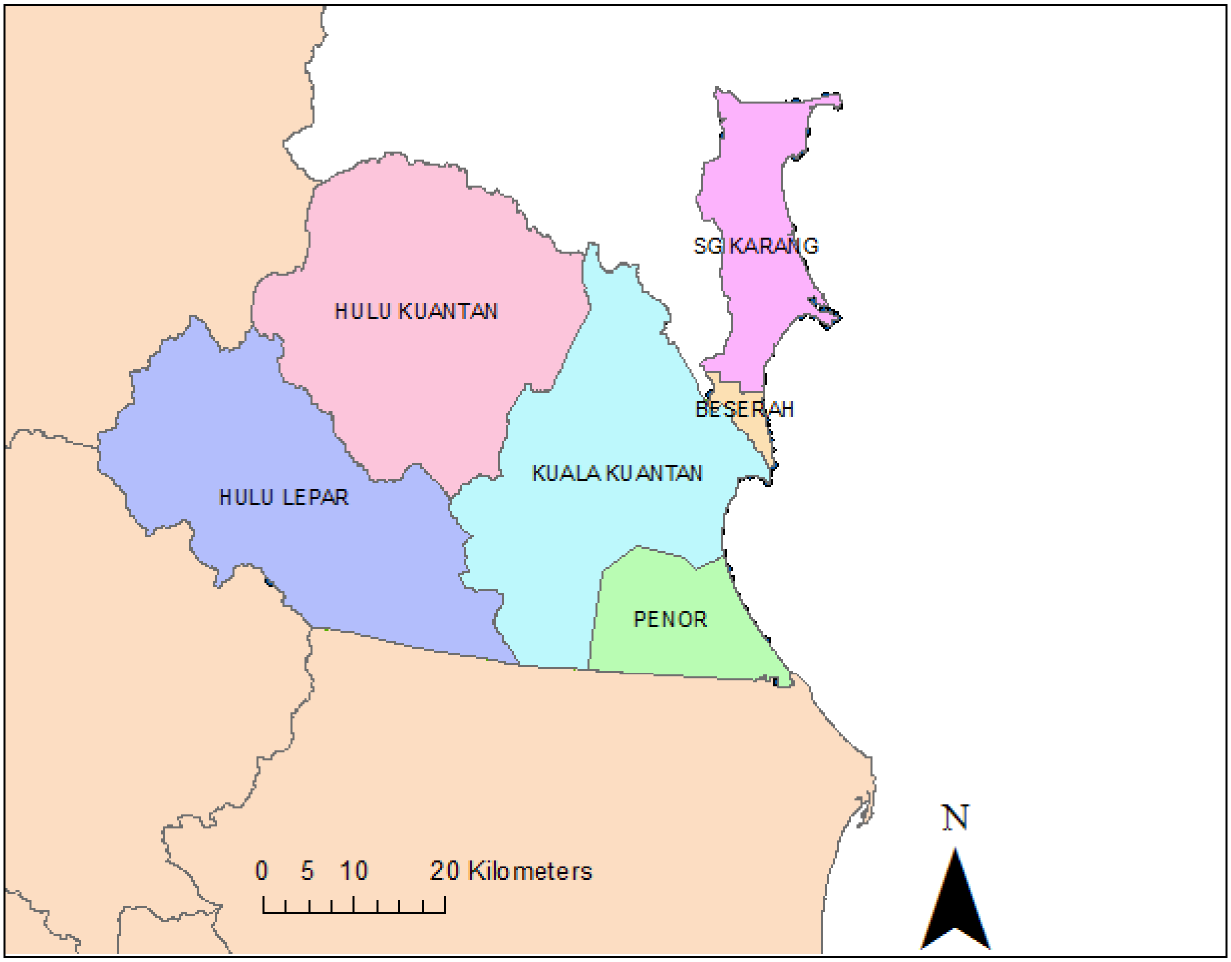


Figure 4.1 The Kuantan district and its sub districts

4.1.1 The topography of Kuantan

Kuantan is located about 250 kilometres from Kuala Lumpur, the capital of Malaysia. It is a coastal town facing the South China Sea on the east, with the state of Terengganu bordering at the north while its south borders other districts in Pahang. A river network runs through the district and is known as the Kuantan River Basin. The Kuantan River Basin is one of the most important river basins in Pahang and has a catchment area of 1630 km^2, which starts from the forest reserve area in Mukim Ulu Kuantan, flowing through the agricultural areas, Kuantan town and towards the South China Sea (Nasir et al. 2012). According to Ghani, Othman and Baharudin (2013), the Kuantan River Basin is largely rural and the upper catchments are largely dominated by natural forest while the middle and lower catchments are dominated by oil palms and slight amount of rubber plantations.

In other research, Mohd Azmi et al. (2009) reported that the state of Pahang is also the host of the South-East Pahang Peat Swamp Forest (SEPPSF), comprising of about 200,000 hectares or about 66% of the surviving peat swamp cover in Peninsular Malaysia and believed to be mainland Asia's largest and most intact peat swamp forest. The estuary is composed of interwoven ecosystem structures including water column, scrub-shrub, forested wetland, emergent marsh, mud/sand flat, shallow slope and deep tidal channel and these habitats contribute to maintaining biodiversity and making available wetland bioresources such as fisheries to the local people (Jalal et al., 2012). Within the Kuantan District itself there are two mangrove forest reserves: the Tanjung Lumpur mangrove forest reserve located adjacent to the Kuantan River estuary near the town of Kuantan (Rahman et al. 2013) and Balok Mangrove Forest reserve which covers an area of 121 hectares at Balok Beach near Kuantan town (Rozainah and Mohamad 2006). The topography of Kuantan is composed of low lying coastal areas and hilly hinterland and the city is recognised as a growth centre for the eastern region of West Malaysia (Salleh 2003).

4.1.2 The Climate of Kuantan

The characteristic features of the climate of Malaysia are uniform temperature, high humidity, copious rainfall and generally light winds and since Malaysia is situated in the equatorial doldrum area, it is extremely rare to have a full day with completely clear sky even during periods of severe drought (MMD 2013). According to the Malaysian Meteorological Department (MMD), the wind flow in Malaysia is generally light and variable however there are some uniform periodic changes in the wind flow patterns during the southwest monsoon, northeast monsoon and two shorter periods of inter-monsoon seasons. The southwest monsoon season is usually established in the latter half of May or early June and ends in September while the Northeast monsoon usually commences in early November and ends in March.

Based on the information provided by the Malaysian Meteorological Department (MMD 2013), “the seasonal variation of rainfall in Peninsular Malaysia consists of three main types, which differ according to the locations below:

(a) Over the east coast states, November, December and January are the months with maximum rainfall, while June and July are the driest months in most districts. This condition applies to Kuantan.

(b) Over the rest of the Peninsula with the exception of the southwest coastal area, the monthly rainfall pattern shows two periods of maximum rainfall separated by two periods of minimum rainfall. The primary maximum generally occurs in October - November while the secondary maximum generally occurs in April - May. Over the northwestern region, the primary minimum occurs in January - February with the secondary minimum in June - July while elsewhere the primary minimum occurs in June - July with the secondary minimum in February.

(c) The rainfall pattern over the southwest coastal area is much affected by early morning "Sumatras" from May to August with the result that the double maxima and minima pattern is no longer distinguishable. October and November are the months with maximum rainfalls and February the month with the minimum rainfall. The March - April - May maximum and the June -July minimum rainfalls are absent or indistinct. “

A significantly high number of human octocariasis cases have been observed from December to March for every year during the six-year period. The number of cases increased in December, peaked in January before decreasing in February and March. This coincides with the northeast monsoon season, which brings heavy rains to the East Coast region. Heavy and continuous rain for 24 hours or more may bring flash floods to low lying areas and areas along the rivers. Kuantan has also experienced floods every year during the monsoon season (Win and Win 2014). These flood disasters have resulted in damages to the infrastructure, economic loss and have also claimed lives. During floods, people are evacuated to higher ground and have to seek shelter at temporary place such as schools that have been set up as evacuation centres. Since the number of human otoacariasis patients has increased in January every year, it may seem that the amount of water or floods may have a possible connection with the increase in tick infection in humans. This possible connection will be further investigated in the analysis section.

4.1.3 Vegetation, land use and economic activities

Over the years, Kuantan has been fast developing into an urban city. As an important city in the East Coast of Peninsular Malaysia, its economic sectors consist of services, manufacturing, wholesale and tourism, agriculture, transport and communication, construction and mining. The services sector such as government, education and banking contributed 32% of economic activity, manufacturing contributed 28%, wholesale and tourism made up 11.5%, agriculture 11.5%, transport and communication at 9% while construction is at 6% and mining is only 2% (Hamzah 2013). In the agriculture sector, the main produce from the district comes from palm oil, rubber, rice, fruit orchards and also cattle rearing. At the same time the manufacturing industries are developing and Kuantan port has been an important hub for trade in the East Coast Region. The tourism industry is also one of the main sources of income for the state due to its beautiful beaches and local attractions and many hotels and resorts are located along the coast. In the coastal areas, there are also small fishing villages where the locals make a living by fishing, making homemade seafood products and also crafts such as batik painting. There is also an interstate road network along the coastline connecting Pahang to Terengganu and the recently completed East Coast Highway has also become an important road network connecting the states in the East Coast Region.

4.1.4 Factors influencing risk

The state of Pahang covers an area of 3,596,500.00 hectares and more than half of it or 2,068,605.18 hectares of the state are covered with three forest categories: 1) natural forest; 2) wild life park; and 3) government's reserve land **(**JPNP 2014**).** Pahang is also a host and gateway to the world-renowned Taman Negara National Park. The park is estimated to be 130 million years old and with an area of 4,343 square kilometres, it is a combination of three protected areas in three states: Taman Negara Pahang National Park, Taman Negara Kelantan National Park and Taman Negara Terengganu National Park with the highest mountain in the peninsular, Gunung Tahan (2,187 meter) is also located in the area (Pakhriazad et al. 2009). As ticks natural habitat is in the forest, it is unsurprising that the forest provides a suitable place for their ecosystem. Therefore it has attracted the interest of the local researchers to investigate how the transmission occurs from ticks to their hosts and how ticks ended up biting and transmitting viruses and bacteria to humans. In doing so, field work such as vector trapping activities have been carried out by the researchers at suspected areas to identify the species of ticks, their lifecycle, the type of hosts and how vector transmission occurs.

These studies have been carried out in the ticks' natural habitat in the forest and also near forest fringe areas where human settlements are found nearby. Besides forest areas, the study area also included villages surrounded by fruit trees and orchards or settlements located near oil palm plantations. It was observed by the researchers that animals that have been infected with ticks have come out at night to look for food and even stray into these villages since traces of their activities were found in the area. As a result of these studies, the researchers were able to identify a number of main vectors for ticks and how their activities near human settlements may have provided an insight into how tick encounters with humans will take place. The identified vectors found with ticks on their body are normally small animals like birds, tree shrews and fruit bats. Wild pigs are also common hosts of ticks as traces of their activities have been found in villages located near forest areas. At the same time there are groups of indigenous people, also known as 'orang asli', who live in the forest and may accidentally get exposed to ticks. These people make a living by hunting wild animals for food and looking for forest products such as tropical fruits and rattan to be sold in the market.

4.2 Data requirement for spatial data analysis

A review of tick borne disease studies by Hay et al. (1997) has suggested that factors such as vegetation, moisture (humidity) and temperature have an influence on tick borne disease incidence. Therefore, these factors have been included in this study to identify if they have any influence on tick borne disease incidence in the Kuantan District study area.

4.2.1 Defining variables for tick borne disease modelling

Suitable data for analysis need to be identified before modelling can be carried out. The data identified will include at least all factors that had been suggested based from other tick borne disease studies. In this case as much useful data as possible on environmental factors, meteorology, disease data and the tick ecosystem will be beneficial to the study. However since not every data source is available and suitable for this study, a selection needs to be made based on the best quality data that is available.

4.2.2 Possible factors and data for analysis in study area

Selection of data is made after considering the factors identified in other tick borne disease studies, data availability within the specified time period as well as conditions of the study area. The following data and factors have been selected for analysis:

(a) Meteorological data

Meteorological data are one of the most important sources of data for this study. Meteorological data were obtained from the Malaysian Meteorological Department (MMD) for a period of six years from the year 2002 to 2007. The meteorological data consist of rainfall data (number of rainy days and volume of rain), wind speed, humidity level and temperature data, which can be referred in the Appendix. Based on literature review, weather and climate have an influence on the spread of tick borne disease. Therefore this study will be investigating if the same factors have the same effect in the study area and to identify if there are any specific weather conditions that can become one of the factors for tick borne disease to spread.

The meteorological data during the study period of 2002 to 2007 for the study area were obtained from the Malaysian Meteorological Department (MMD). The data were recorded at the Kuantan meteorological station, which is located at the latitude of 3° 47' N longitude 103° 13' E as in Figure 4.2 and this data can be referred to in Appendix A.

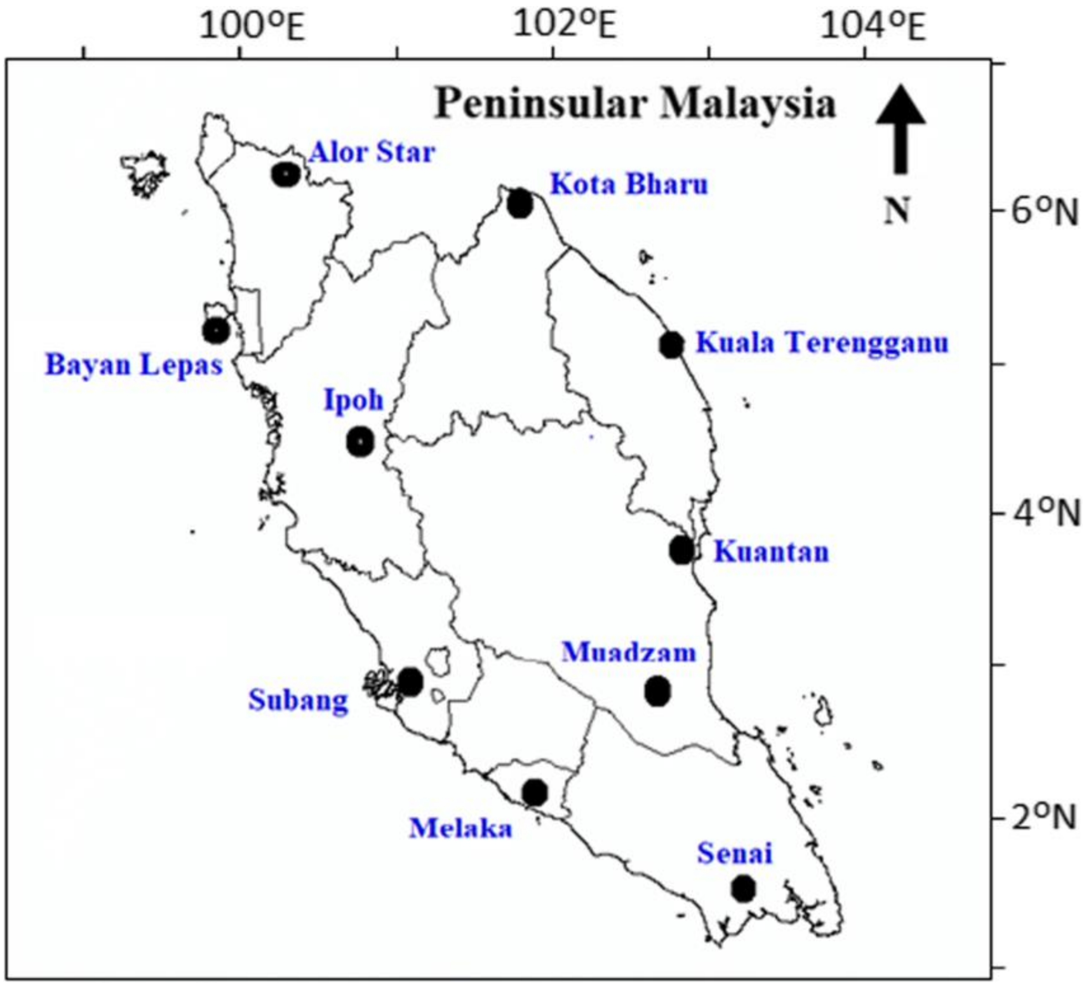


Figure 4.2 Map showing the locations of meteorological stations in Peninsular Malaysia (Muhammad et al., 2021)

Every parameters of the meteorological data are compared against the number of cases for each month to identify any pattern or significant relationship with tick borne disease incidence. The parameters included in this study are:

(i) *Temperature*

The temperature data consist of monthly mean minimum and maximum temperature measured in degrees Celsius recorded for the Kuantan District. Comparison is to be made between temperature and the number of cases for each month to detect if there is any visible pattern or relationship between the parameters during high or low temperature. If there is any pattern, then chi square statistical analysis can be performed to prove if the relationship is significant and is not just a coincidence. The Chi-square test of independence (also known as the Pearson Chi-square test, or simply the Chi-square) is one of the most useful statistics for testing hypotheses when the variables are nominal and unlike most statistics, the Chi-square ($\chi2$) can provide information not only on the significance of any observed differences, but also provides detailed information on exactly which categories account for any differences found (McHugh, 2013). A graph showing Mean monthly maximum temperature for every year from 2002 until 2007 is shown in Figure 4.3 below.

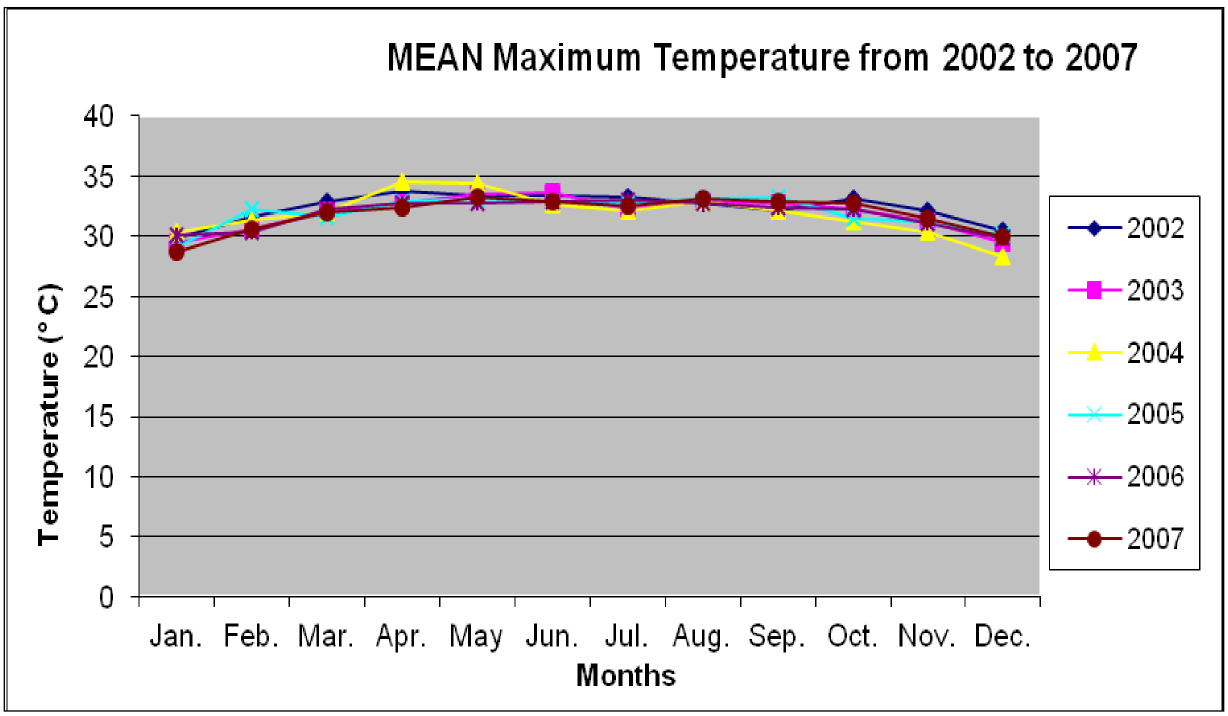


Figure 4.3 Graph showing the monthly mean maximum temperature from 2002 to 2007

(ii) *Wind speed*

Wind speed is one of the parameters included with the meteorological data. It is measured in metres per second. Even though there was no indication of a relationship between wind speed and tick borne disease in previous research, the analysis is still performed to see if any relationship exists.

If there is no relationship or pattern detected, then this parameter will not be considered as a factor and can be ruled out from the study.

(iii) Relative humidity

The Relative humidity (RH), which is expressed as a percentage, measures the actual amount of water vapour in the air compared to the total amount of vapour that can exist in the air at its current temperature, resulting in cooler air for higher relative humidity and warmer air for lower relative humidity (Louisville 2004). The analysis performed will investigate if there is any visible pattern of the level of relative humidity during the period of time when tick infections occur.

(iv) Rainfall data

There are two parameters for rainfall data, which are analysed in this study: the monthly rainfall and number of rain days:

1) Monthly rainfall

Rainfall data are measured in millimetres (mm) and the daily amount of rain is totalled to become monthly rainfall for every month. This monthly amount of rain will be compared against the number of cases for each month to see whether any pattern exists between the number of cases and wet weather conditions. Ticks are known to prefer to live in environments with high moisture conditions. Therefore the study will investigate if the amount of rain has an influence on their activities and lifecycle by analysing the number of tick infections during the wet seasons.

2) Number of rain days

The number of rain days is defined as the number of rainy days for every month during the period of six years from 2002 to 2007. An investigation will be made to identify if more cases were being recorded during the months with the most rainy days, without considering the amount of rain. If a pattern exists when comparing the number of cases and the number of rain days, then further investigation can be made to understand and prove that there is a relationship between the parameters.

(b) Remote sensing and GIS data

According to Young et al. (2004) "remote sensing and GIS can be used to monitor weather and climate changes, generate precision forecasts, classify land use, examine associations between location and adverse outcomes, identify and measure community exposure to public health threats and improve survey sampling methods". Furthermore, the researchers have stated, "by using these technologies to refine the methods by which information is gathered and analysed after a disaster, it will effectively contributes towards prevention and mitigation". Useful information that can be extracted from remote sensing data include vegetation and land use type, identifying build-up areas, soil types and also identifying water body such as rivers and lakes. Remote sensing derived data that are being applied in this study are:

(i) Vegetation index (VI)

Ticks natural habitat is in the forest and forest fringe areas where there are canopy and high vegetation area. Therefore, it is important to analyse areas with high vegetation to find tick habitats. There are two measurements of vegetation density; namely the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI). The Vegetation Index was first discovered by Rouse et al. (1974) when applying Landsat data to study and measure the amount of green biomass of vegetation on earth. The vegetation indices are used for global monitoring of vegetation conditions and are used in products displaying land cover and land cover changes (LPDAAC 2014). As tick habitats are in the forest, areas with high vegetation index may be able to pinpoint forest and high vegetation density areas, which are the natural habitats of ticks. The abundance of ticks in their habitats may lead to higher probability for the ticks to bite when they encounter an unsuspecting host.

In Malaysia, the rapid change in land use, including the conversion of forests to farmlands or plantations and the spread of urbanization to forest-fringe areas, resulted in increased encroachment into an animal, especially wildlife habitats which lead to greater opportunities for humans-animals interactions contributing to possible zoonotic disease transmission (Jing et al., 2017). However more studies are carried out on tick species (Ernieenor et al., 2017), tick-borne pathogens (Sharifah et al., 2020), personal prevention against tick bites (Ghane Kisomi et al., 2016) and tick borne disease than the study on the risk factors associated with tick infections, which includes vegetation.

Since bio-climatic conditions are key determinants of arthropod vector distribution and abundance and consequently affect transmission rates of any diseases they may carry, NDVI is commonly used in models of vector suitability, abundance and disease transmission and early warning while Multi-temporal NDVI, when combined with land surface temperature measures across seasons, has proven to be a potent combination for vector-borne disease prediction (Kelly et al. 2011).

The formula for NDVI is:

$$NDVI = (R_{NIR} - R_{RED})/(R_{NIR} + R_{RED}) \tag{4.1}$$

where R_{NIR} and R_{RED} indicate reflectance in the near infrared and red wavebands, respectively (Gamon et al. 1995).

These NDVI and EVI data need to be derived from satellite imagery such as SPOT, Landsat and MODIS. However, the selection of satellite imagery depends on the availability of the satellite data during the study period, the coverage of the study area and the quality of data. Landsat satellite data are not available to be used for this study since there is no cloud-free image that can be found of the study area. It is also difficult to remove the cloud cover from each of the Landsat images, as there are also elements involved such as cloud shadow and haze. Since Landsat data are not suitable, the next step is investigating the possibility of selecting a higher resolution SPOT images for deriving vegetation indices. However SPOT data are also found to be having the same problems of data with cloud cover and additionally some parts of the study area were not covered during the study period of between 2002 to 2007, with only recent data from the year 2008 being available.

The next step is investigating the possibility of using a coarser resolution MODIS data. MODIS data for this study has been obtained from the MODIS Vegetation Indices 16-Day L3 Global 500m data set and are readily available for a period of six years from the year 2002 until 2007. These images are made available free for download from the MODIS Land Processes Distributed Active Archive Centre (LP DAAC) website: https://lpdaac.usgs.gov/.

According to Huete et al. (2002), “the Enhanced Vegetation Index (EVI) was developed to optimise the vegetation signal with improved sensitivity in high biomass regions and improved vegetation monitoring through a de-coupling of the canopy background signal and a reduction in atmosphere influences.” Therefore, EVI has been chosen for analysis in this study since it is more suitable for the study are which has high vegetation cover as well as problem with cloud cover. Furthermore, the EVI is not just “able to reduce both atmospheric and soil background noise simultaneously, but also its better performance in many applications as compared to NDVI (Guoyu et al. 2007)”. In this study, areas with high EVI values can be identified as areas with high vegetation density, which will also point as areas of tick’s natural habitat and may contribute to higher chances of tick bites.

(ii) *Land use type*

There are different types of land use in the study area, including areas with high vegetation cover such as forests, plantations and fruit trees. There are also human settlements such as small villages in the rural areas with houses that are surrounded with shrubs and different types of vegetation such as fruit trees that are normally being planted nearby the houses. These areas with certain types of vegetation can become a good habitat or questing area for ticks. Regarding this information, remote sensing satellites are known with the ability to detect natural resources on earth, which can then be used to generate more specific information on land use types. This information can be presented on a map and can be further categorized into different types of land use such as agricultural areas, urban areas and forest areas. A few types of land use can also be categorized and divided into their own smaller groups such as different types of agriculture or forest. Figure 4.4 shows the categories of land use in the study area, which are common land use type in Malaysia. According to Mat Akhir, S. A. and A. Mohammad (2007), the land use categories or land classes considered in their study are derived from the Anderson et al. (1976) classification level I and level II, which is a base for classification of remote sensing derived land use. Another study has also applied the similar land cover classification for land use and cover mapping with airborne hyperspectral imager in Setiu, Malaysia (Jusoff 2009).

Further description of land use type available in the area is described in Table 4.1. To investigate the influence of land use on tick borne disease incidence, spatial analysis can be performed using the GIS software. The analysis will concentrate on the areas within a two kilometres radius from the point of infections. Land use type in the areas is to be identified to determine the land use type with the largest number of patients' records. The result of analysis is presented in Chapter 5.

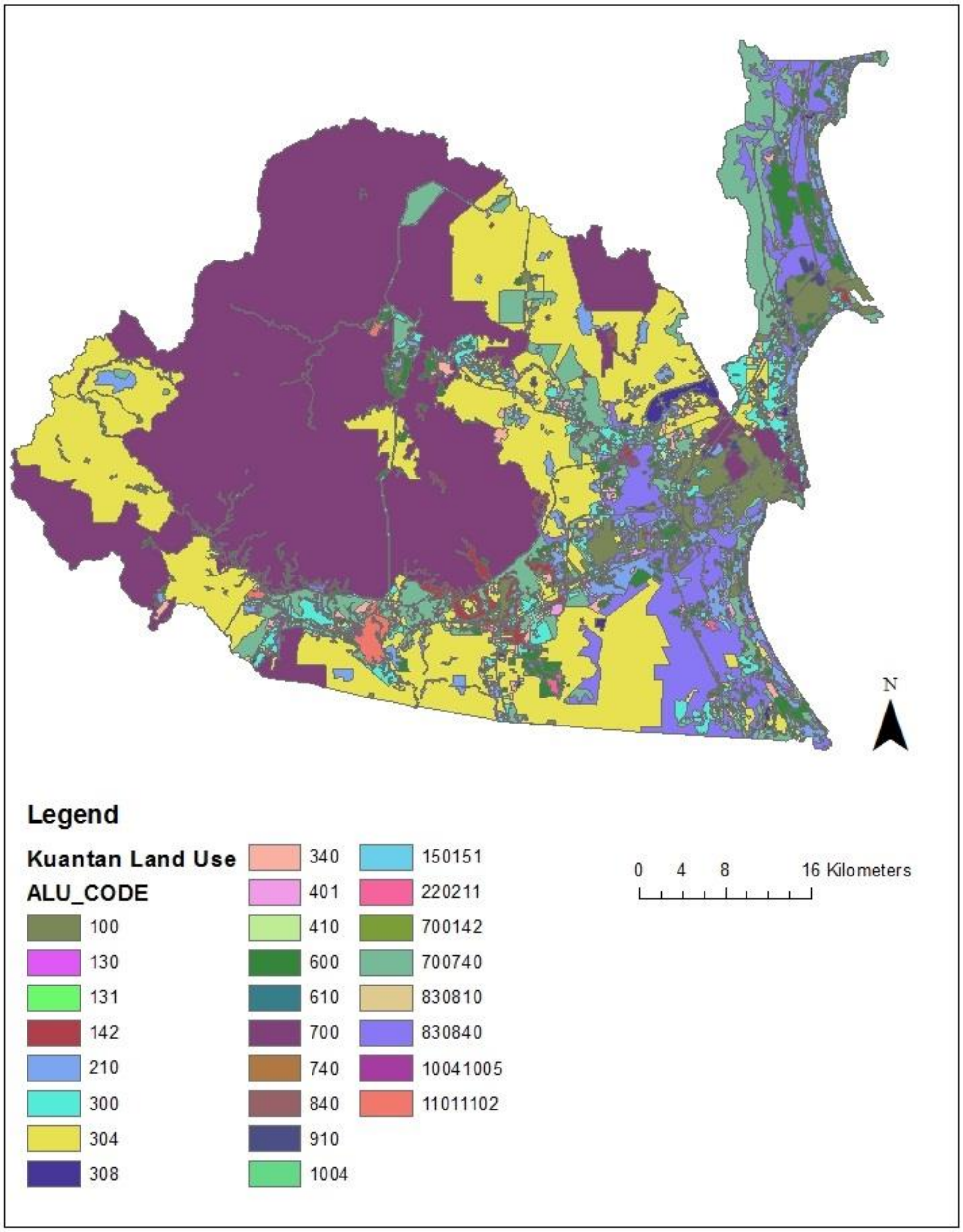


Figure 4.4 Land use map showing different types of land use categorized in code, which is described in Table 4.1.

Land use code (ALU_CODE)	Description
100	Urban and associated areas
130	Power line
131	Gas pipe line
142	Other mining areas
210	Mixed horticulture
300	Rubber
304	Oil palm
308	Coconut
340	Orchard
401	Diversified crops
410	Wet paddy
600	Grassland/ *lallang*
610	Scrub
700	Forest
740	Secondary forest
840	Marshland
910	Cleared land
1004	River/canal/waterway/drain
150151	Main highway/ main road
220211	Vegetables/ floriculture
700142	Forest/ other mining areas
700740	Forest/ secondary forest
830810	Swamp forest/ mangrove
830840	Swamp forest/ marshland
10041005	River/canal/waterway/ponds and lake
11011102	Non ruminant/ ruminant

Table 4.1 Land use type code and its description

(iii) River and road network

Rivers and road are both mediums of communication and transport. Besides connecting towns, they can also be possible mediums of the transfer for disease. Therefore it is interesting to investigate if the river or road network has any influence to the increase in infection cases. The river and road networks are two important networks to be included in the analysis. The rivers that run through the area may become one of the factors since high-density vegetation is found along the river as well and also swamp areas, as ticks are found to prefer areas with high vegetation and moisture. Buffer areas were created within two-kilometre distances from the rivers as in Figure 4.5 below.

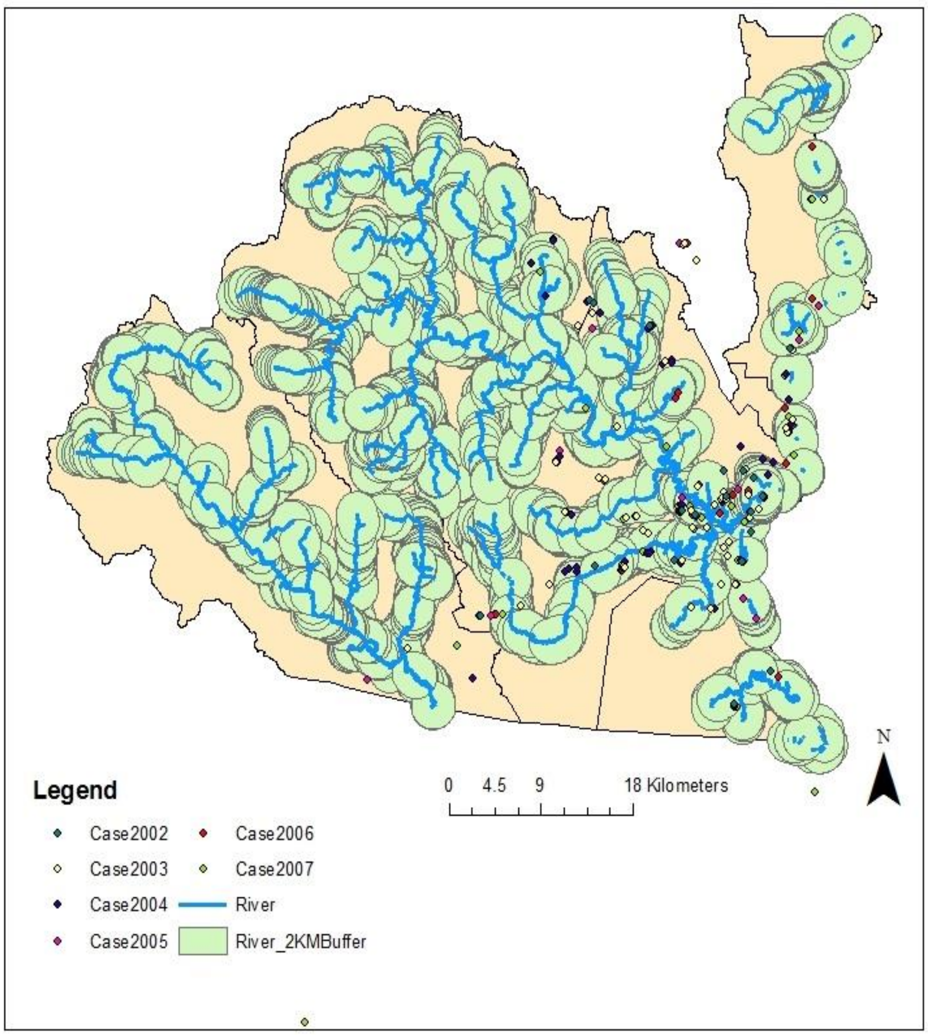


Figure 4.5 Map showing cases distribution within two-kilometre buffers of river

The points on the map represent individual cases for each year from 2002 until 2007. Based on visual observation, significant numbers of cases are found to be located close along the river while fewer cases are found further from the rivers.

Since visual observation shows a relationship between distance from the river and the number of cases, the chi square analysis has been chosen to further analyse and investigate whether the pattern does signify the relationship between case occurrence and distance from the river. The chi square statistical analysis is calculated according to the chi square formula as described in 4.2. The observed frequency represents the actual cases that have been reported while the expected frequency is the cases that are expected to occur under complete spatial randomness.

Based on the calculation produced in Microsoft Excel, p value is found to be less than 0.5 as shown in Table 4.2. Consequently, when the result of $p < 0.5$, then the pattern of observed cases is not due to randomness and therefore represents the real pattern of infections. This analysis concludes there are more cases that can be found near the river, which suggests that the river does have an influence on disease occurrences.

The Chi square formula is as the form:

$$x2 = \sum \frac{(Observed\ Frequency - Expected\ Frequency)^2}{Expected\ Frequency} \quad (4.2)$$

$$x2 = \sum \frac{(O - E)^2}{E} \quad (4.3)$$

Where O is the frequencies observed and E is the frequencies expected.

Distance from river	Expected/ Observed	Year						
		2002	2003	2004	2005	2006	2007	Total
<1km	Expected	4.6	12.4	9	4.2	7	11	48.2
	Observed	16	32	26	14	17	40	145
1-2 km	Expected	4.6	12.4	9	4.2	7	11	48.2
	Observed	3	17	12	4	15	11	62
2-3 km	Expected	4.6	12.4	9	4.2	7	11	48.2
	Observed	4	7	6	2	1	0	20
3-4 km	Expected	4.6	12.4	9	4.2	7	11	48.2
	Observed	0	3	0	0	0	1	4
>4 km	Expected	4.6	12.4	9	4.2	7	11	48.2
	Observed	0	3	1	1	2	3	10
Total	Expected	23	62	45	21	35	55	241
	Observed	23	62	45	21	35	55	241

Distance from river	Observed	Expected	p
<1km	145	48.2	1.34E-60
1-2 km	62	48.2	
2-3 km	20	48.2	
3-4 km	4	48.2	
>4 km	10	48.2	
total	241	241	

Table 4.2 Result of Chi square test for number of cases within specified distances from the river for the year 2002 until 2007

(iv) *Administrative boundary*

The administrative boundary is the virtual boundary on the map separating Kuantan District from other districts in the state of Pahang. Even though there is no limit or geographical boundary to the spread of any diseases, this administrative boundary is required to focus specifically on Kuantan District as the study area. Hence, other cases and parameters outside of the boundary can be eliminated or ignored in this study to conduct research only within the specified study area.

(v) Land Surface Temperature (LST)

Land surface temperature (LST) is one potential factor or influence that can be included in this study. Daily LST data can be compared with individual case's dates and analysed to find if there is any significant relationship between the two. For this research, MODIS LST data from the year 2002 to 2007 was obtained for each respective case. However, heavy cloud cover during the wet Monsoon season in the East Coast from November to March each year has affected the availability of MODIS data resulting in no data values in the area. This has posed a setback for further analysis since the highest number of cases were recorded during the period.

(c) Hospital records of patients

The original records of hospital patients who were treated for human otoacariasis contain the information on their location. However, these individual records have been mapped digitally as points and saved as vector data. Based on these points, analysis can be carried out on the point of infection and within a specified distance surrounding the point of infection to understand the type of environment the patient was living in at the time of infection. By having information on patients' locations, investigation can be made into the conditions of area that may be suitable for tick bites to happen.

(d) Population data

Information about the human population and their activities may provide some insight into how people get infected at the first place. It is important to find out if there is any link that exists between areas or towns with high population density, their economic and social activities with high number of tick infection cases. Factors that can be investigated are the background of patients from the infected area including to identify their age and gender to understand if there is any specific category of people that are most likely to be infected.

For this study, the population data from the year 2010 was obtained from the Department of Statistics Malaysia (DOSM). There were no population data collected during the study period from the year 2002 to 2007 as the department only performed the population census once in every ten years (Talha, Nair and Ismail 2009). Therefore it is noted that the population data from the year

2010 may not accurately represent the population during 2002 until 2007. At the same time, the data only represent a few selected areas from the Kuantan District. As a result, it is not possible to use these data in the analysis since data are incomplete and may not generate the correct result to portray the number of population in the study area. Meanwhile, the population data of the few selected study areas is shown as in Table 4.3.

Area	Population		
	Total	**Male**	**Female**
Bandar Baru Chendor	3,424	1,940	1,484
Bukit Kuantan	2,885	1,492	1,393
Felda Lepar Utara	4,030	2,208	1,822
Kawasan Kelab Golf	205,975	105,813	100,162
Kg. Balok Baru & Kg. Cengal Lempung	23,353	12,321	11,032
Kg. Batu Lima	692	423	269
Kg. Berahi	14,582	8,132	6,450
Kg. Geliga Besar & Kg. Paya Berenjut & Kg. Fikri	18,631	9,787	8,844
Kg. Panching	1,485	771	714
Kg. Padang & Kg. Padang Perdana	10,745	5,540	5,205
Kg. Pelindung & Kg. Ceti & Kg. Baharu	14,860	7,581	7,279
Kg. Pertul & Kg. Hijrah	53,403	27,449	25,954
Sg. Isap	75,223	38,483	36,740
Kg. Kempadang	1,824	903	921

Table 4.3: Total of population by gender for selected area, Pahang, Malaysia.
Source: Population and Housing Census of Malaysia 2010

4.3 Spatial analysis and modelling methods

Spatial analysis and modelling is an important method that can help to identify factors of tick borne disease related with high risk of disease outbreak by analysing trends or patterns of disease occurrence. These patterns of disease occurrence can be identified by investigating the historical records, such as when the disease was most detected and what are the specific environmental factors that were present during these occurrences. These factors can be identified as suspected factors and their significance can be tested by statistical analysis to prove their significant relationship. Therefore, temporal records of disease occurrence are important for the study. Meanwhile, investigation will be made to identify if spatial patterns such as temperature, rain and vegetation may have contributed to disease occurrence. These spatial patterns may also be observed to see if the same pattern occurs over a period of time. Therefore this study poses a spatial and temporal problem and further analysis must be made to investigate if spatial distributions of disease vary through time and if temporal distributions vary through space.

When all data and suspected factors have been identified and obtained, the next step is to perform spatial analysis using GIS software. Spatial data analysis include visualization, exploration of possible patterns and modelling (Pfeiffer 1996). In this study, the first stage of data analysis involves performing exploratory data analysis.

4.3.1 Exploratory Data Analysis and Visualization

Data analysis is one of the most important tasks in research. Data need to be identified, gathered and analysed for the purpose of identifying their suitability and significance for further analysis and modelling. There are two types of data analysis; either quantitative data analysis or qualitative data analysis, which depends on their respective type of data.

The initial step to data analysis is to perform exploratory data analysis. Exploratory data analysis is a method of summarizing the data, usually in graphics for visual observation and examinations, which can provide an overall picture of the situation before getting into further investigation. This graphics visualization involves the display of data in a graph or map format and can be achieved by using software such as statistical software. For instance, data will be able to be presented as a disease distribution map to assist visual analysis. On such a disease distribution map, each individual case can be represented as a point, along with other data such as land use, river and road networks that may help people to interpret and transform their data into useful information as well as providing some insight or direction for further analysis.

4.3.2 Data analysis

When possible factors have been identified during exploratory data analysis, further analysis will involve performing spatial data analysis, queries and developing hypotheses to determine the pattern of tick infection by analysing the significance effect of each possible factor. This includes identifying any relationship between the factors. The results of analysis can be presented either in a map, graph or as results of statistical analysis. Factors proven to have significant relationship with tick borne disease are called significant factors, which can then be used in a model to predict disease risk in the future.

For this study, comparisons have been made between the infection cases and spatial patterns such as temperature, rain, wind, distance to the rivers and land use. The analysis results are presented in Chapter 5. During analysis, it is shown that there is no significant relationship between ticks with vegetation or EVI. However, since ticks habitats are all in areas with high vegetation, further analysis is needed, which is by performing the species distribution modelling and adopting the pseudo-absences strategy.

4.3.3 Species distribution modelling (SDM) and pseudo-absence analysis

Since there is limited information in the literature on the biological aspects of ticks, this research has been focussed on the area where ticks are abundant and likely to bite, in other words, to study the tick's species distribution. The pseudo-absence analysis strategy being adopted involves comparing between presence and absence, where presence means the environment or area where disease or cases are found while absence means the environment or area where diseases or cases are not found or reported. To ensure that the analysis is performed without any bias or conflict, it is important to select pseudo-absences data from random locations where there are no diseases or cases reported. Therefore, two-kilometre buffer areas were created around each infection points to signify presence data, which will be excluded from the selection of pseudo-absences. By utilising the Random Points tool in the ArcMap software, in this case, two sets of random points or pseudo-absence points are collected from the areas outside of the presence data buffer selection. Then a Zonal Statistics as Table analysis was performed for the pseudo-absences data to obtain the Mean EVI that corresponds with the same date of the presence data.

Next, Chi square analysis was performed for dominant land use against Mean EVI and also non-dominant land use consisting of other land use types. The analysis will be able to show whether the result the analysis is significant ($p < 0.5$) for further analysis. Since the result for dominant land use is significant, further analysis of logistic regression was performed between the EVI of dominant land use and the EVI of pseudo-absences data. The purpose of logistic regression is to model the relationship between a dependent and one or more independent variables. In this study, the objective is to build a model to predict the probability of disease occurrence using the Mean EVI as the predictor. The dependent variable is the positive (pseudo-positive where disease occurs) and negative (pseudo-negative where there is no disease). The result of the analysis will be able to show how good the predictive capacity of the model is.

When EVI has proven to be a factor for disease modelling, further analysis in logistic regression is required to prove that the significant land use type such as oil palm, swamp and urban are the factors of tick infestation, which leads to an increase of tick bites. Since decision needs to be made involved modelling of categorical and continuous data, the best modelling technique that can be applied is Classification and Regression Tree (CART) with logistic regression. According to Pfeiffer et al. (2008), the objective of logistic regression is to identify factors that influence the risk of disease being present or absent at specific locations (e.g. farm or household) using the binary labels ‘positive’ (i.e. disease present) or ‘negative’ (i.e. disease absent). While the Classification and Regression Tree (CART) can be utilised as an analytical tool that can be used to explore public health data which are often complex, unbalanced and contain missing values, where the relationships between a health outcome and its determinants may not be linear and necessitate higher order interactions (Speybroeck 2012). The CART diagram and logistic regression method for this analysis is shown in Figure 4.6.

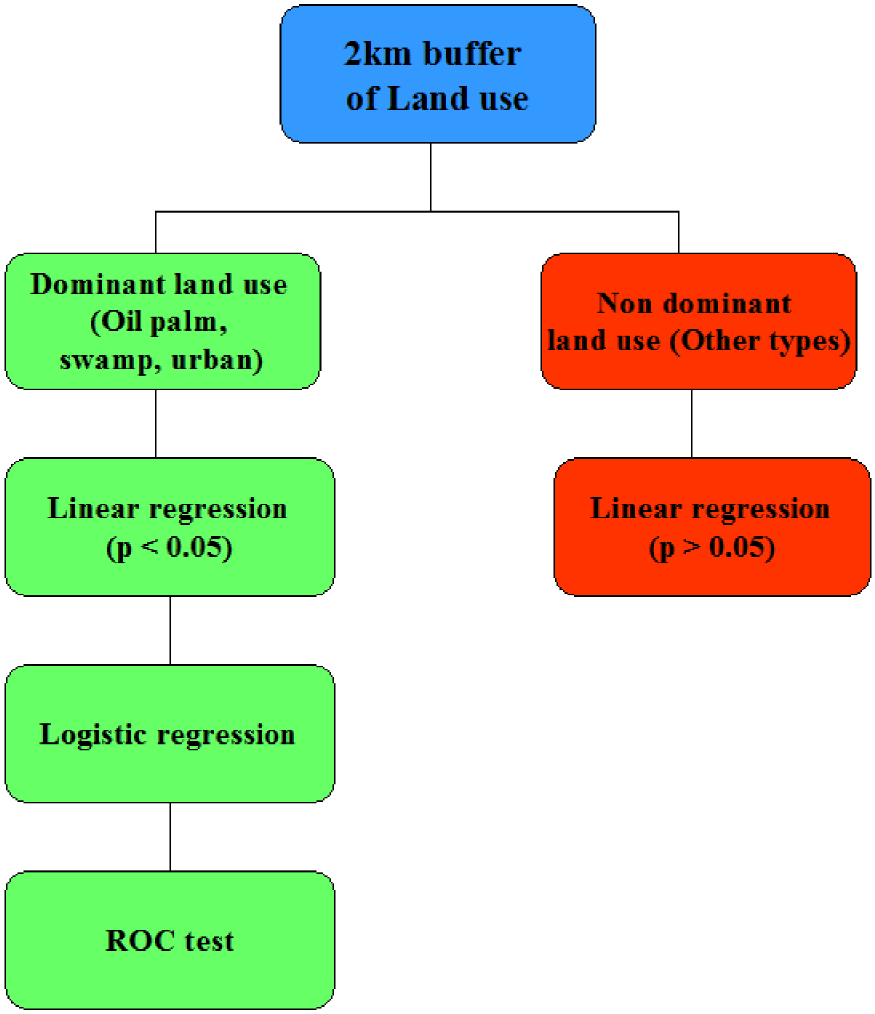


Figure 4.6 The CART diagram for modelling of tick borne disease

For this CART analysis, the inclusion and exclusion strategy of factors to engage in CART analysis have been considered from the study sample. The inclusion criteria identify the study population in a consistent, reliable, uniform and objective manner while the exclusion criteria include factors or characteristics that make the recruited population ineligible for the study (Garg, 2016).

To prove that the models can be used, the two sets of pseudo-absences or random points need to be tested using the Receiver Operating Characteristics (ROC) test to determine their accuracy and capability for modelling. ROC analysis is a useful tool for evaluating the performance of diagnostic tests and more generally for evaluating the accuracy of a statistical model (eg, logistic regression, linear discriminant analysis) that classifies subjects into one of two categories, diseased or non-diseased (Zou, O'Malley and Mauri 2007). An ROC curve shows the trade-off between sensitivity and specificity, where any increase in sensitivity (true positive rate) is accompanied by a decrease in specificity (false positive rate). According to (Zhu, Zeng and Wang 2010) for a given diagnostic test, the true positive rate (TPR) against false positive rate (FPR) can be measured, where:

$$TPR = TP/(TP+FN) \tag{4.4}$$

And

$$FPR = FP/(FP+TN) \tag{4.5}$$

As can be seen from the above equations, TPR is equivalent to sensitivity and FPR is equivalent to (1 – specificity) and all possible combinations of TPR and FPR compose a ROC space. One TPR and one FPR together determine a single point in the ROC space and the position of a point in the ROC space shows the trade-off between sensitivity and specificity.

Meanwhile according to Fawcett (2006) as shown in Figure 4.7, informally one point in ROC space is better than another if it is to the northwest where true positive rate is higher and false positive rate is lower or both, which means A is better than B. C's performance is random while D is the perfect classification. Meanwhile E performs the worst.

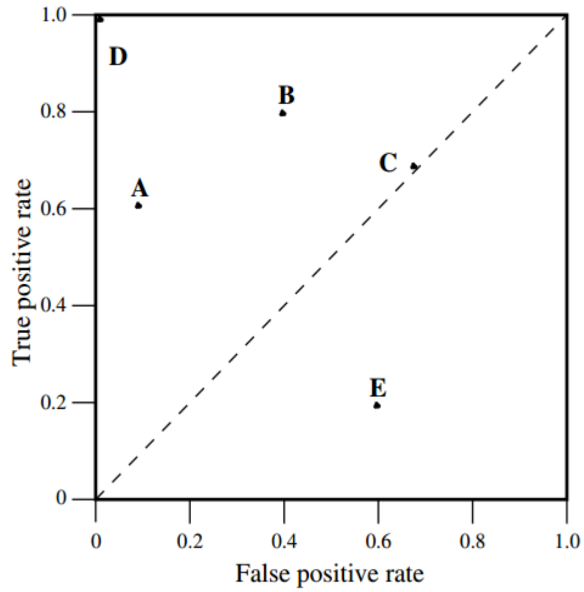


Figure 4.7 A basic ROC graph showing five discreet classifers (Fawcett 2006)

Based on analysis, three types of land use, namely oil palm, swamp and urban have been identified to be present and the most dominant type of land use during tick infections. Then the logistic regression statistical tests for each year from 2002 to 2007 have been carried out for the three dominant land use types to prove their significance as the factors for tick borne disease and also presented in Chapter 5.

4.4 Software

Software being used for this study consisted of the following:

4.4.1 GIS software

Software used is ESRI's ArcGIS version 10.0 for spatial data analysis and map visualization.

4.4.2 Statistical software

The Statistical Package for the Social Sciences (SPSS) software and Microsoft Excel were used throughout the analysis, which is described in Chapter 5.

4.5 Chapter summary

Different species of ticks are prevalent in certain types of geographical location and climates. Specific environmental conditions of the study area may contribute to the suitable conditions that may result in increase in tick bites. Therefore getting more background information of the study area is important to understand the type of suitable data that can be used for analysis such as climate, vegetation, human activities and understanding suitable habitats for ticks. The data can then be analysed to identify the parameters or suspected factors of tick borne disease.

The suspected factors need to be tested to identify if they are the significant factors before modelling can be done. Hence, the type of significant factors may be able to point to a suitable modelling method since the techniques applied for tick borne disease studies in other regions may not necessarily be suitable for modelling for the study area. A correct selection of modelling technique will be able to generate an accurate result and therefore be able to give prediction of disease risk in the future.

5 ANALYTICAL RESULTS

This chapter presents the results of spatial analysis and statistical tests performed based on the methodology described in Chapter 4. The results of each test will be discussed here to identify the significant factors of tick borne disease. The factors can then be integrated in a model to predict the future risk of tick borne disease.

5.1 Spatial data analysis for tick borne disease

Prior to modelling, the most important task is to identify the significant factors of tick borne disease. Therefore spatial data analysis must be performed on each suspected factor to identify the real factors and to prove the result statistically. During spatial data analysis, data for these factors will be closely analysed by mapping disease location, investigating and analysing the surrounding areas as well as performing statistical tests to prove if there is any connection. The statistical test results will be able to highlight the influence of the data and prove the real factors of tick borne infections of human in the future.

For testing and analysis, a selection of spatial data has been made and the motivations behind those data selections are explained in Chapter 4. These data consist of the district level administrative boundary with information on names of locality, land use type, meteorological data, forest type, flood data, road and river networks as well as records of human otoacariasis patients.

During the study period of six years from the year 2002 until 2007, a total of 395 human otoacariasis cases or tick infestations of the ears were recorded as shown in Table 5.1. These records are kept manually as hardcopy at the Tunku Ampuan Afzan hospital in Kuantan, where the patients were treated. Based on these records, there were no cases observed from the month of April to December 2005 however this is due to records being missing and not because there were no infection cases. A graph has been plotted in Figure 5.1 to better visualize the disease distribution across the six-year period. The high number of cases every year were detected from January until March and also in December during the Northeast monsoon season. The monsoon season, which lasts from November to March, brings heavy rain and flooding in the Kuantan area and may pose one of the significant factors for tick infection in humans. The distribution map of human otoacariasis from 2002 to 2007 is shown in Figure 5.2.

Year	Month												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2002	27	24	8	4	3	3	8	5	3	0	0	18	103
2003	17	17	7	3	4	2	3	4	3	2	0	12	74
2004	25	9	4	3	4	8	3	3	0	1	0	3	63
2005	21	9	2	0	0	0	0	0	0	0	0	1	33
2006	10	8	7	1	1	1	5	3	1	0	1	7	45
2007	14	9	10	6	1	2	10	3	6	3	3	10	77
Total	114	76	38	17	13	16	29	18	13	6	4	51	395

Table 5.1 Monthly records of human otoacariasis cases from year 2002 to 2007

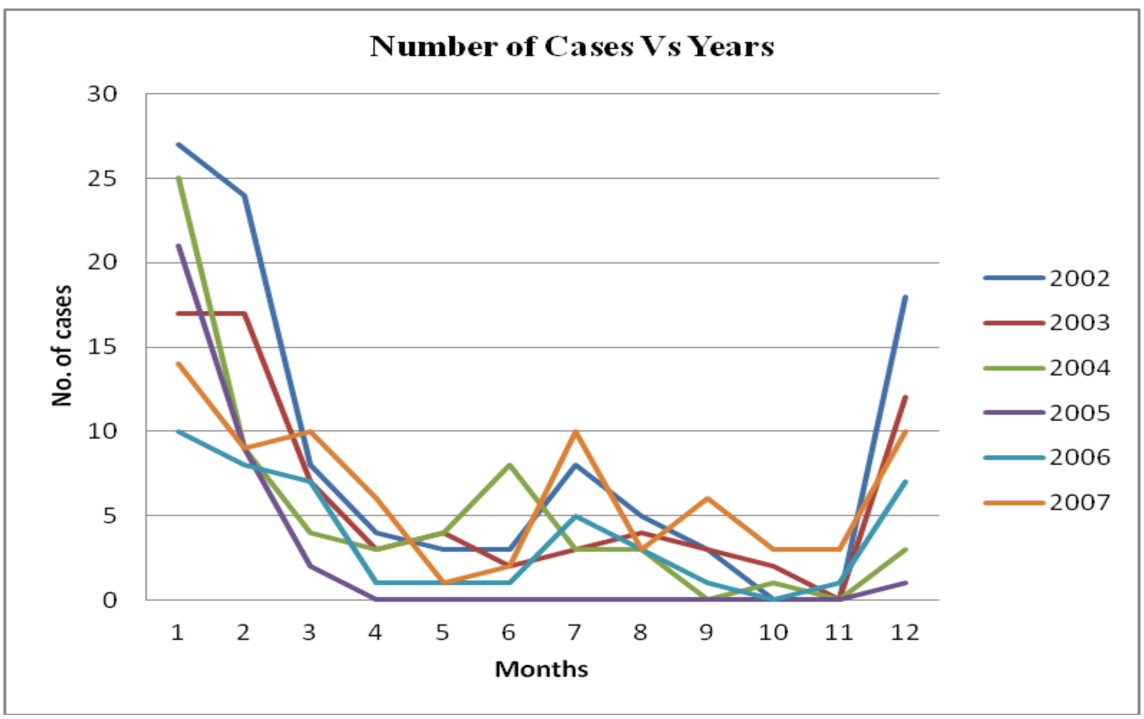


Figure 5.1 Temporal distributions of human otoacariasis cases from 2002 to 2007

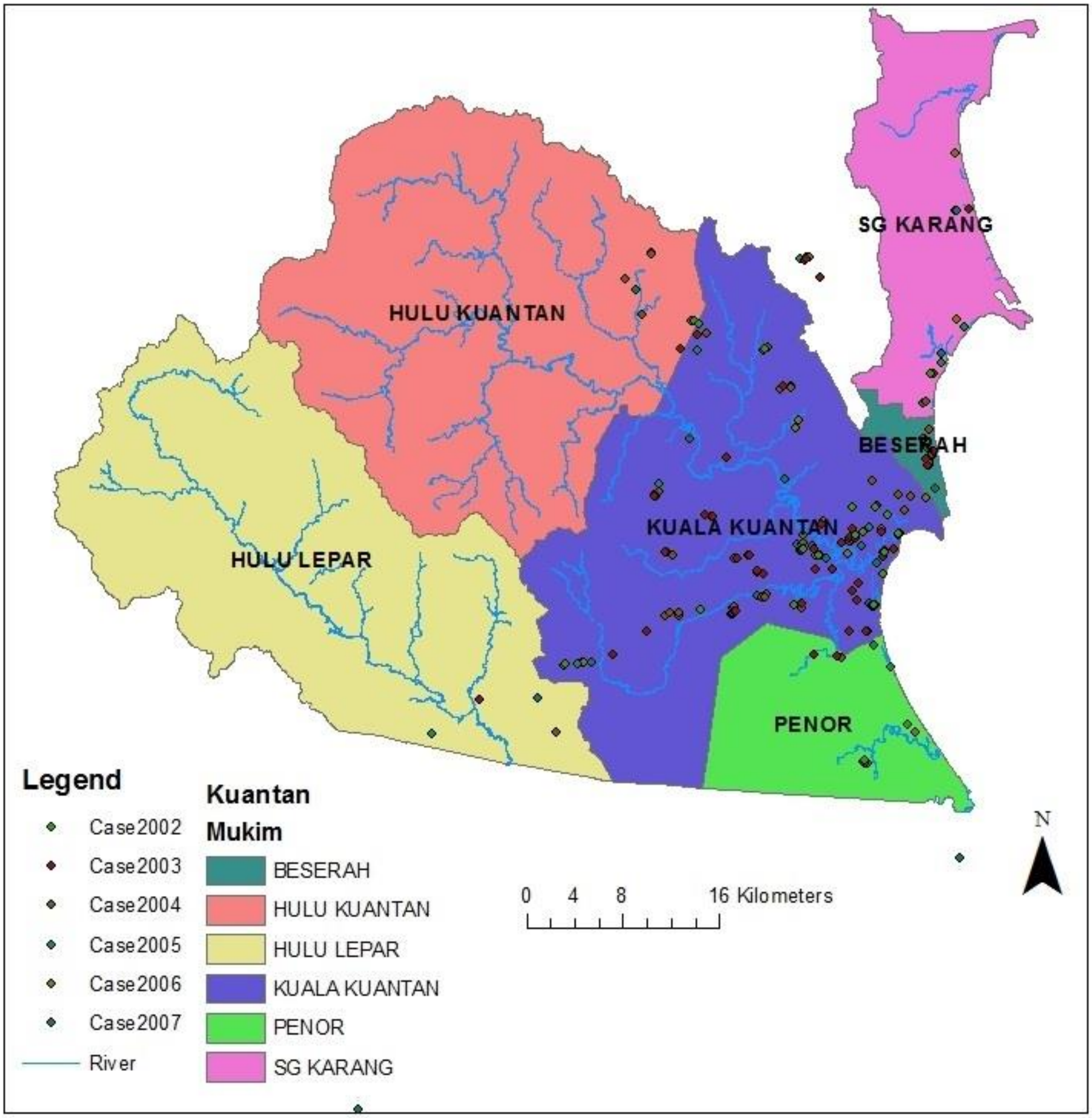


Figure 5.2 Map showing the spatial distribution of human otoacariasis cases from year 2002 to 2007

5.2 Visualization and statistical analysis

The goal for data analysis is to identify factors of tick borne disease by analysing each suspected factors as described in Chapter 3.

5.2.1 Meteorological data

Comparisons were made between tick infections cases against each of the meteorological data, both monthly and yearly. However, these comparisons did not show any significant relationship between these meteorological data and tick infections cases. The results for each comparison are shown and discussed below:

(a) Temperature against number of cases for each year

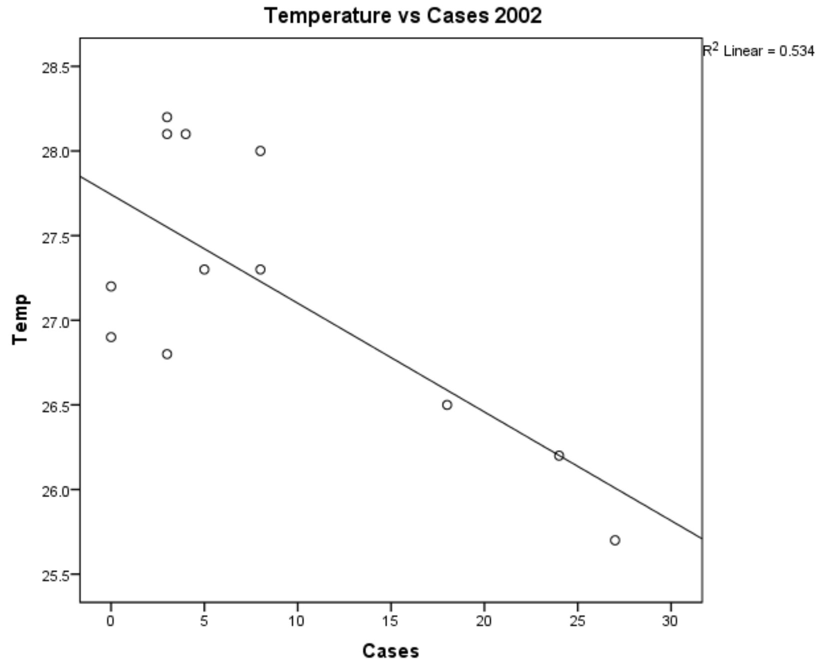
The result of correlations in Table 5.2 below shows the relationship between the numbers of tick infections cases against monthly temperature for each year, from 2002 to 2007. For most of the years, the pattern shows that more cases were present when temperatures were getting lower, except for in 2004. In 2005, there was a significant high value of correlation between temperature and cases, followed by 2002 and 2007. While in 2002, the correlation value was at its lowest. However there is very little difference in temperature since the study area experiences a uniform equatorial climate with no distinct or significant changes in temperature recorded within one year, which is unlike countries experiencing different seasons where there are drastic changes in weather and temperature.

Comparison	**Year**					
	2002	**2003**	**2004**	**2005**	**2006**	**2007**
Temperature versus cases	-0.534	-0.362	-0.002	-0.629	-0.208	-0.499
Wind speed versus cases	0.673	0.212	0.101	0.054	0.408	0.197
Relative humidity versus cases	0.017	0.042	0.005	0.002	0.085	0.054

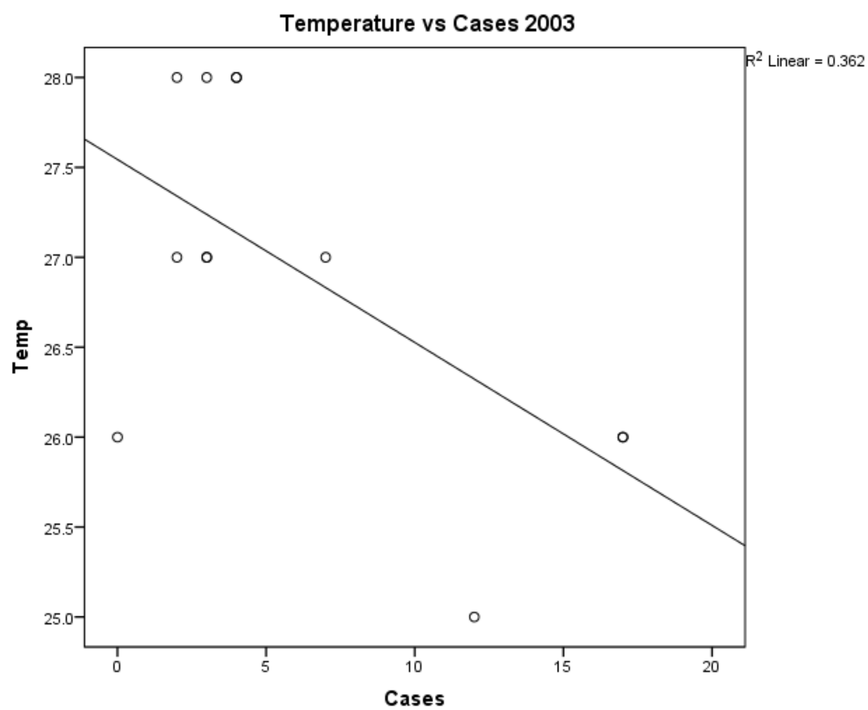
Table 5.2 Pearson Correlation between case numbers and monthly meteorological variables

Table 5.2 shows the result of linear correlation between two variables. A positive correlation means the value of the first variable increase will result in an increase in the second variable. A negative correlation means the increase of one variable will decrease the value of the other variable while no correlation means there is no relationship between the two variables where the values is toward 0. A perfect correlation will produce an R^2 value of 1.0, which is a perfect linear relationship.

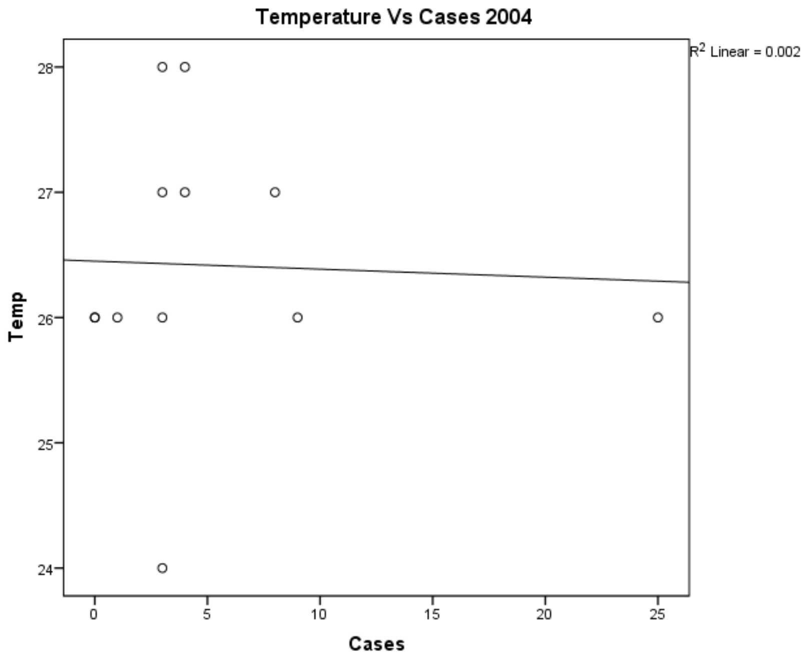
Figure 5.3 shows negative correlation between temperature and disease incidences. This signifies that less cases are recorded when temperature gets lower. However, as mentioned above, using temperature as one of the parameter may not be a good choice due to the study area having an equatorial climate where temperature is relatively uniform throughout the year.



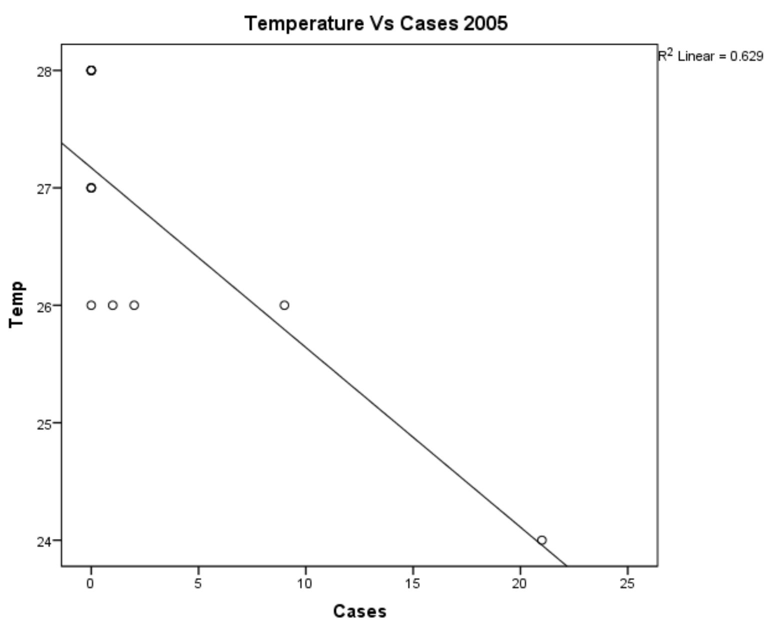
a) Temperature against cases for the year 2002



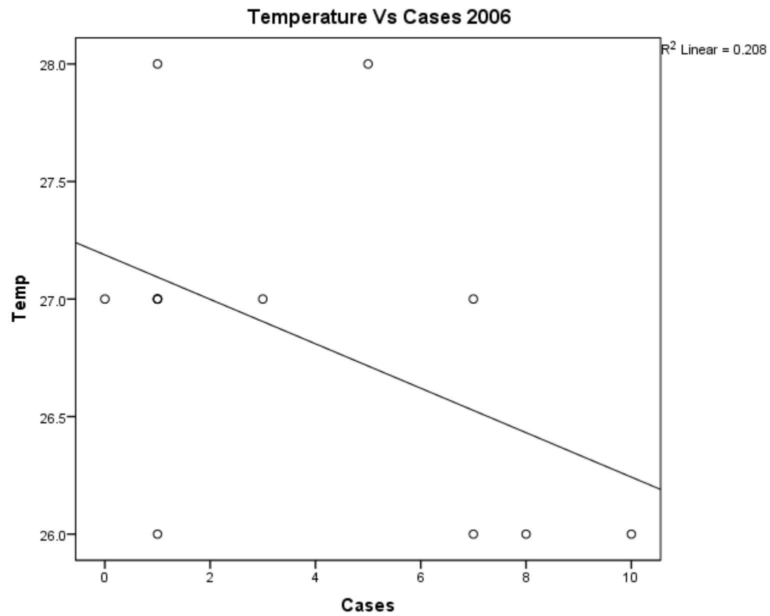
b) Temperature against cases for the year 2003



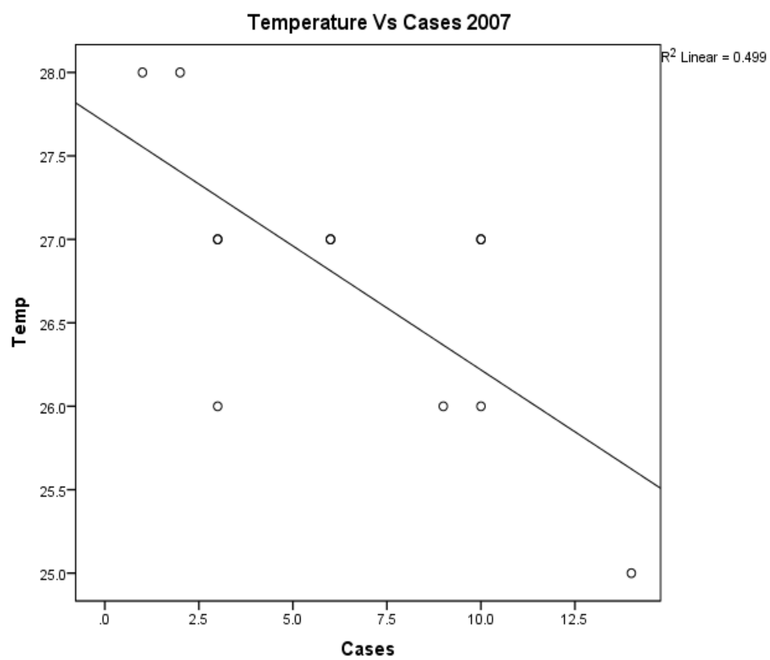
c) Temperature against cases for the year 2004



d) Temperature against cases for the year 2005



e) Temperature against cases for the year 2006



f) Temperature against cases for the year 2007

Figure 5.3 Graphs showing linear correlation between monthly mean temperature against number of cases for each year *(b) Wind speed*

Wind speed is measured in metre per second and even though there was no literature review or previous findings on wind speed relationship or windy conditions with tick borne disease, it is still interesting to investigate if there is any chance at all of relationship that has not been found. The value of wind speed for each month was plotted against the number of cases in each year as shown in the Appendix. The R^2 values for the relationship between wind speed and number of cases are shown in Table 5.2. In 2002 and 2006, the correlation values were significantly higher compared to the lowest value recorded in 2005. Even though the graphs showed similar pattern, the actual distributions are very scattered and not distributed closely along the lines.

(c) Relative humidity (RH)

Ticks are known to thrive in environments with high moisture levels. Therefore high humidity conditions may be one of the indicators that signal high abundance of ticks, which in turn may lead to tick infestation and infections on human. Based on the graphs plotted between relative humidity and number of cases (in Appendix), the result shows that there is a positive relationship between the two variables. However, based on the R^2 values shown in Table 5.2, the correlation value was the highest in the year 2006 and the lowest in 2005, but the relationship is not significant enough to require further investigation.

(d) Monthly rainfall

Based on previous research on ticks' survival, they are known to be able to survive wet and rainy seasons as well as being submerged in water for a few days. Therefore the occurrence of heavy rainfalls may not have an effect to them but may possibly slow down their activity especially during host-questing activities. The total amount of rain and number of rain days for each month are summed up and compared against the total number of cases for the respective months. However, graphs plotted between the amount of rain and cases show no significant relationship or visible pattern as shown in Table 5.3 below. Graphs showing correlation between monthly rainfall volume and the number of cases for a period of six years from 2002 until 2007 are included in the Appendix.

Month	Rainfall data from 2002-2007	Number of rain days from 2002-2007
January	-0.290	-0.001
February	-0.098	-0.012
March	-0.004	0.034
April	0.010	-0.020
May	-0.063	-0.309
June	0.053	0.239
July	-0.042	0.005
August	0.0503	0.323
September	-0.303	-0.210
October	0.495	0.222
November	-0.167	-0.240
December	0.066	0.481

Table 5.3 The Pearson correlation value (R^2) between the volume of rain and the number of rainy days against number of cases for each month.

(e) Number of rain days

The number of rain days means total number of days for each month on which any rain falls. This analysis investigates whether high numbers of rainy days or if more wet days for the month have any influence on an increase in tick infections. Based on the graph plotted between the total numbers of rainy days against cases for each month in the Appendix, it is observed that there is no pattern or significant relationship between these two variables.

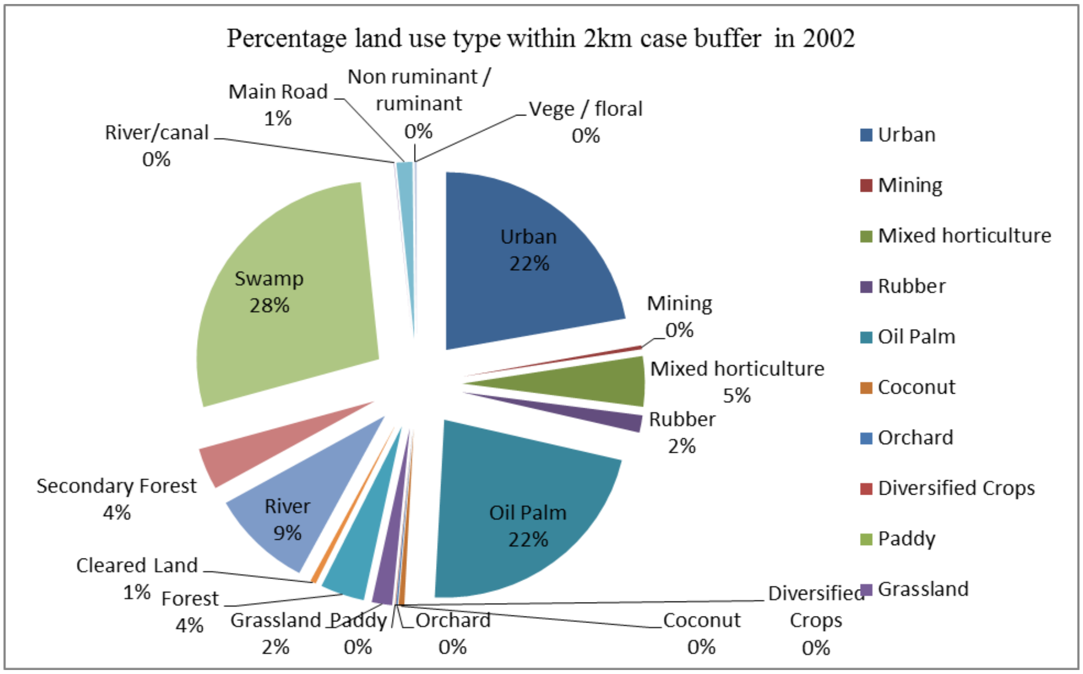
5.2.2 Remote sensing and GIS

Remote sensing satellites are able to observe the earth's natural resources and therefore able to detect land use changes based on temporal images. This technology has helped researchers to understand specific environmental conditions suitable for their subjects' habitats, therefore providing a better understanding on ticks infestations and infections. At the same time GIS data may come from many sources, including remote sensing data. By using specialised GIS software, researchers are able to integrate all data including remote sensing data into a virtual map and perform spatial analysis such as query and analysis, data overlay, buffer and perform statistical analysis to produce disease distribution and risk maps.

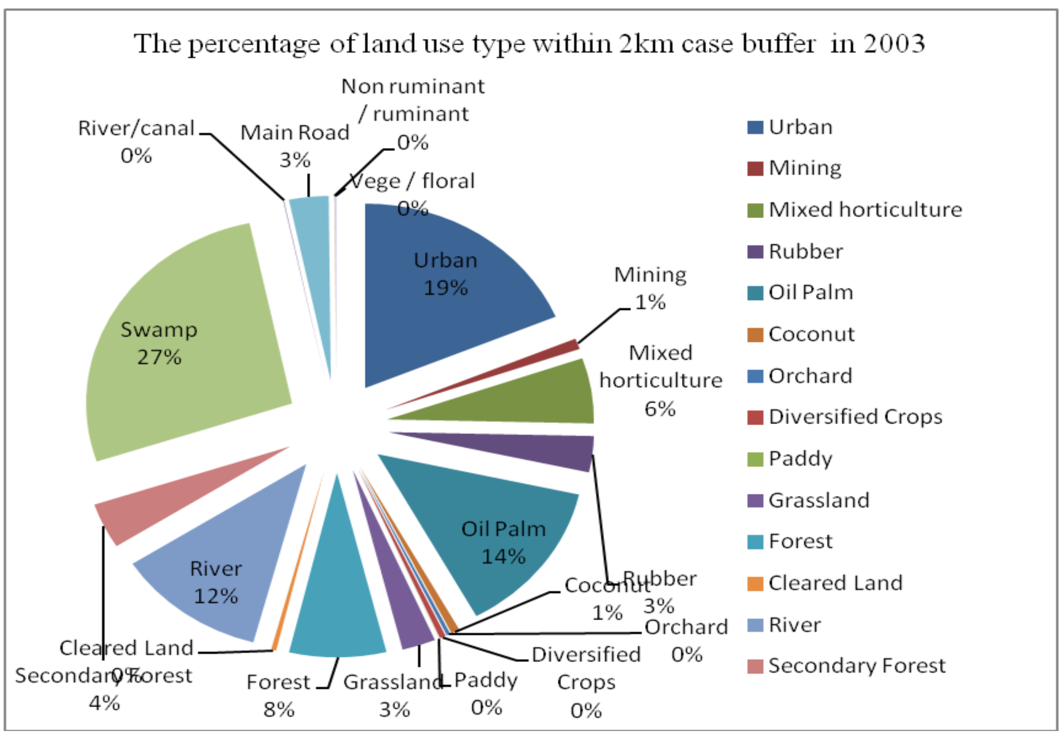
a) Land use type

An analysis was carried out to identify the type of land use within a two-kilometre buffer of an infection point or location of patient. The two-kilometre buffer was selected since humans activities in the area were carried out close to the area where they live, which exposed them to contact with ticks.

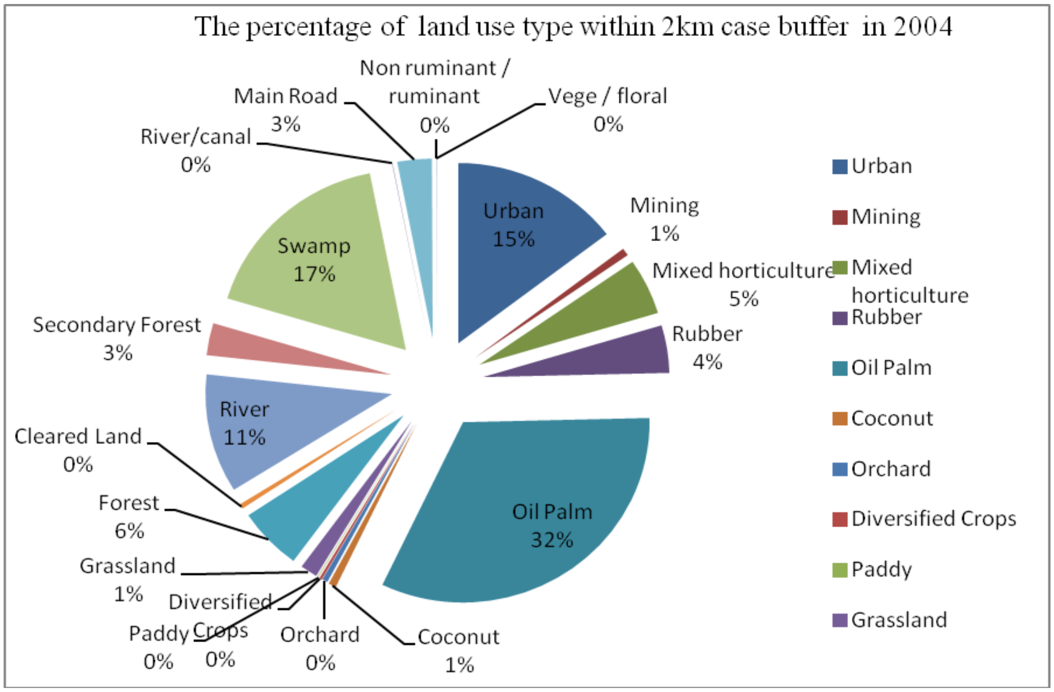
The result of analysis is presented in pie charts as in Figure 5.4 below. According to the charts, three types of land use were identified as the three most common land use types found within the two-kilometre radius of infection point every year. The land use types are namely swamp, urban and oil palm plantation. The detailed percentage for each type of land use for each year is presented in Table 5.4. Figure 5.5 summarizes the total percentage of land use types for a period of six years from the year 2002 until 2007.



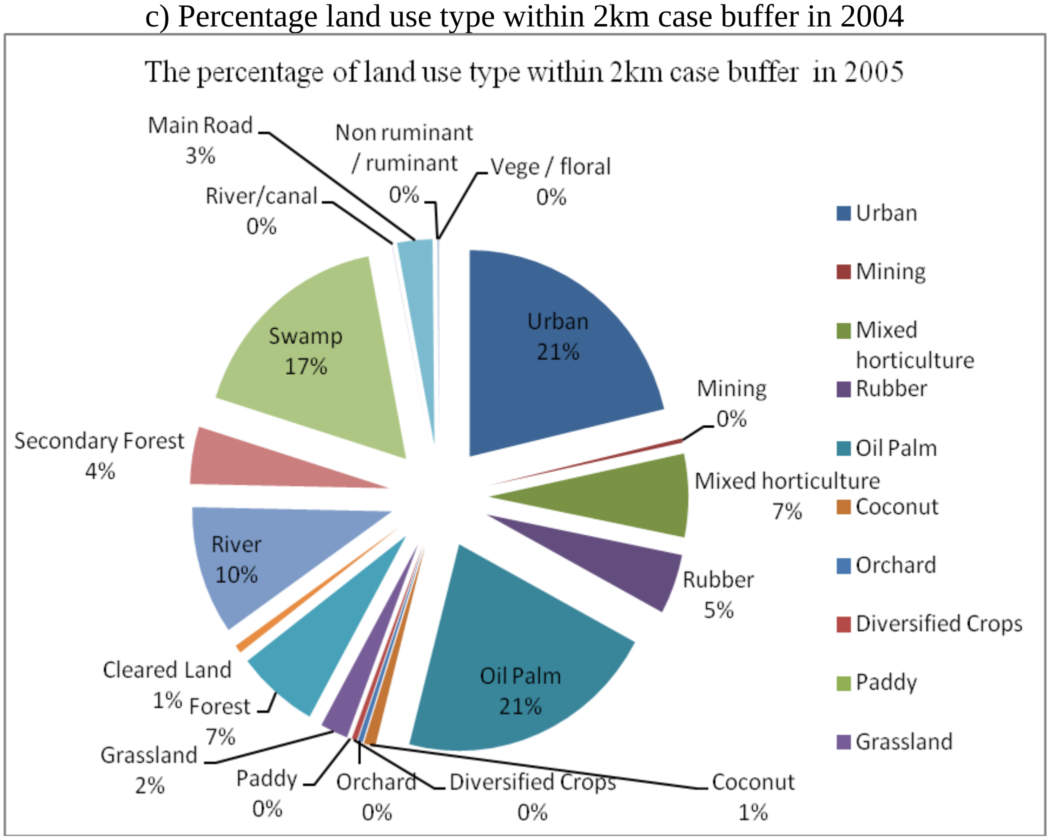
a) Percentage land use type within 2km case buffer in 2002



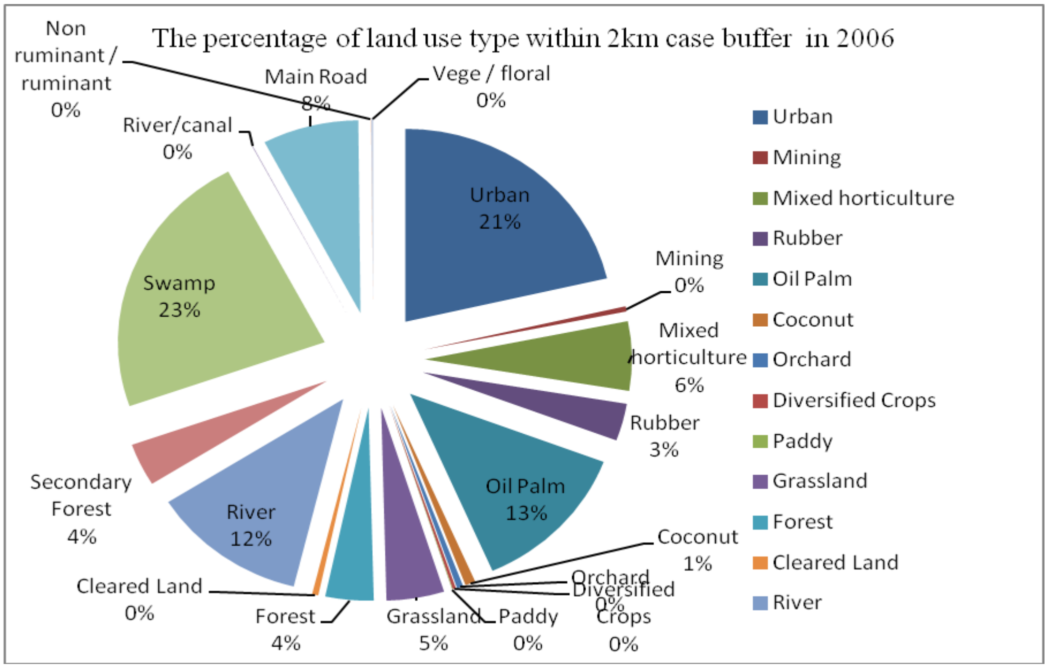
b) Percentage land use type within 2km case buffer in 2003



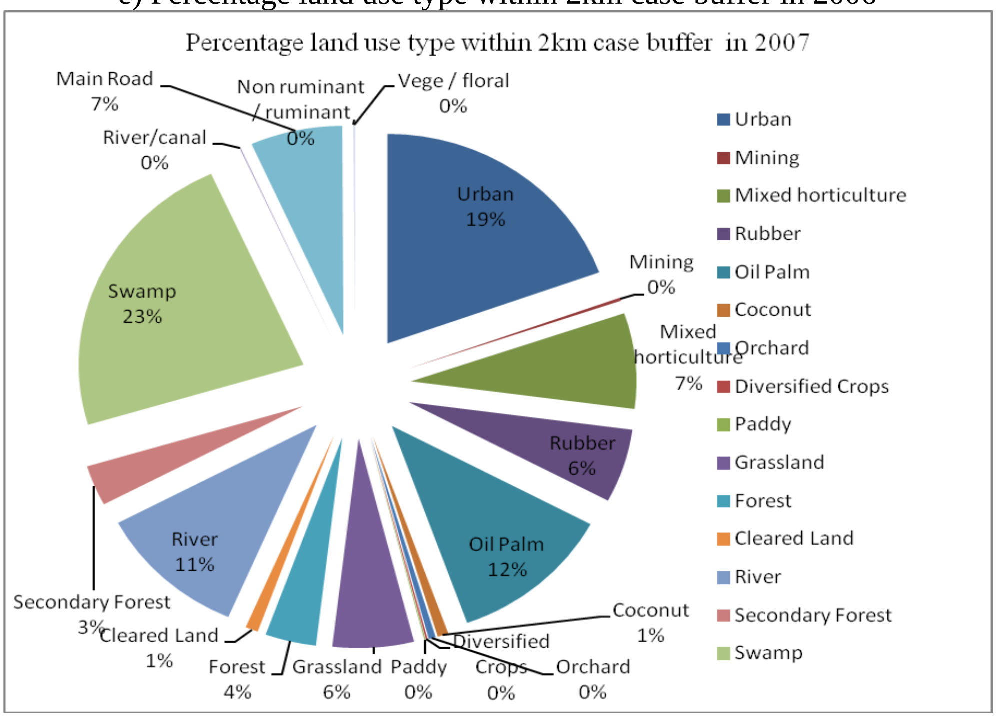
c) Percentage land use type within 2km case buffer in 2004



d) Percentage land use type within 2km case buffer in 2005



e) Percentage land use type within 2km case buffer in 2006



f) Percentage land use type within 2km case buffer in 2007

Figure 5.4 Pie charts showing the total percentage of land use type within a two-kilometre radius from each point of infections for every year from the year 2002 to 2007

Land use type	Year						
	2002	**2003**	**2004**	**2005**	**2006**	**2007**	**Total**
Urban and associated areas	22.17046	18.68793	14.64331	21.30086	21.38355	19.48132	117.6674
Mining	0.346173	0.94298	0.686812	0.333588	0.432044	0.241499	2.983096
Mixed horticulture	4.572695	5.70689	5.025634	6.519235	5.799226	7.382931	35.00661
Rubber	1.566532	3.169736	4.268175	4.698934	3.12784	5.679328	22.51054
Oil Palm	22.1757	13.52783	32.49607	21.17225	12.73625	11.78306	113.8912
Coconut	0.496969	0.678016	0.667203	0.934838	0.80405	0.779668	4.360745
Orchard	0.145412	0.299331	0.48088	0.388697	0.492176	0.492018	2.298515
Diversified Crops	0.056841	0.472115	0.236917	0.434324	0.258407	0.130779	1.589384
Paddy	0.015449	0.018641	0.059095	0.006944	0.052897	0.079255	0.232282
Grassland	1.804497	2.628694	1.493198	2.246043	4.513743	5.819651	18.50583
Forest	3.913132	7.609651	5.59679	6.595639	3.763581	3.65319	31.13198
Cleared Land	0.515689	0.396031	0.433429	0.633697	0.484106	0.90864	3.371592
River	9.093803	11.79881	10.66236	10.0689	12.20358	10.80761	64.63506
Secondary Forest	3.775888	3.970325	2.882864	4.485345	3.575408	3.124678	21.81451
Swamp	27.73611	26.76788	17.26016	17.15172	22.64709	22.89464	134.4576
River/canal	0.040686	0.065211	0.054295	0.024066	0.056402	0.064591	0.30525
Main Road	1.411349	3.098426	2.981048	2.856593	7.502736	6.594716	24.44487
Non ruminant / ruminant	0	0.045593	0.002487	0	0.031905	0	0.079984
Vege / floral	0.162616	0.115925	0.069275	0.148331	0.135002	0.082425	0.713573
Total percentage	100	100	100	100	100	100	600

Table 5.4 Percentage of land use types within a two-kilometre radius of an infection point

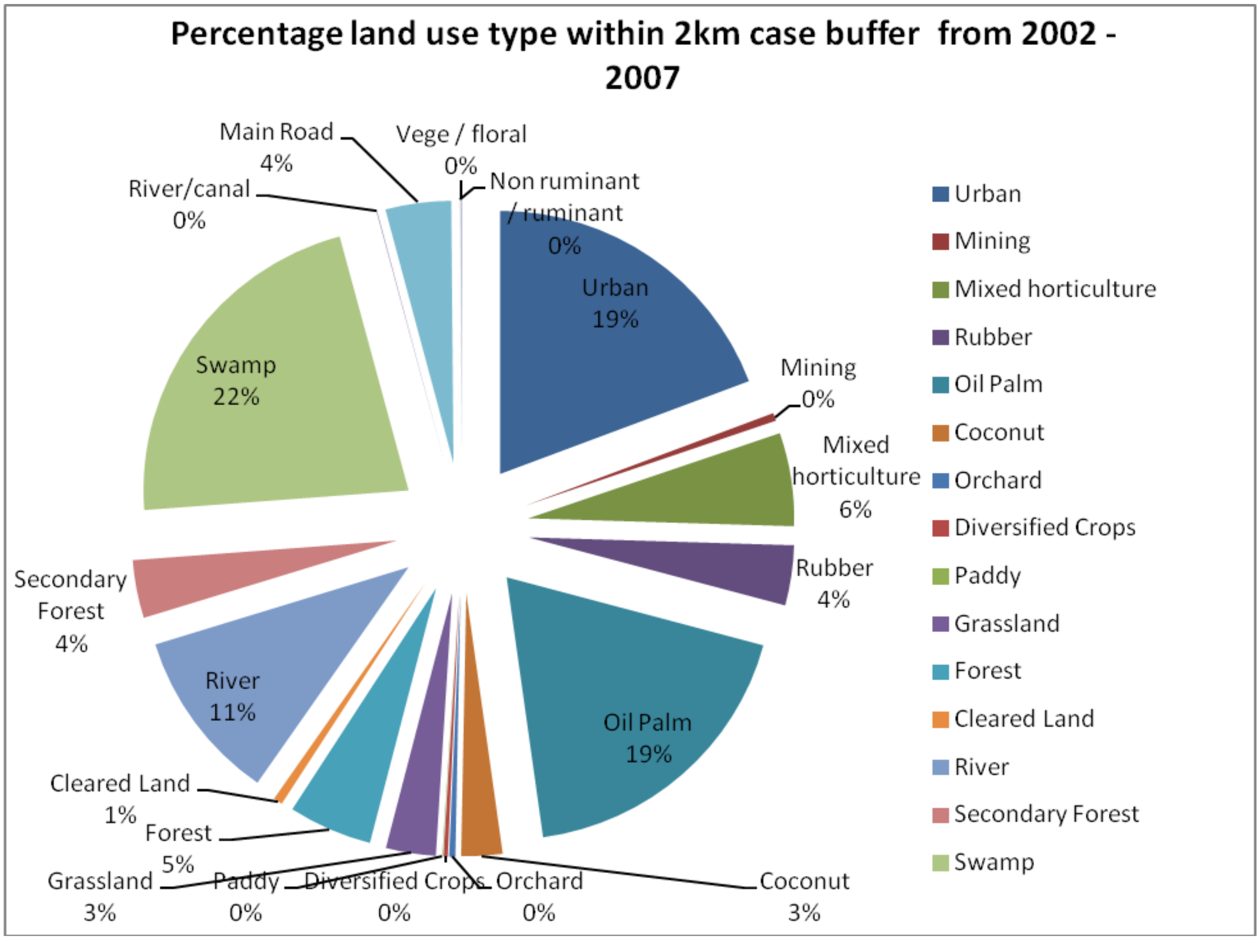


Figure 5.5 Pie chart showing the accumulated total percentage of land use type within a two kilometre radius from each point of infections from the year 2002 to 2007

To see whether land use has a significant relationship with disease occurrence, a Chi-square statistical analysis was performed. Based on this test, analysis shows a positive relationship between land use type and disease occurrence cases in the area. The Chi-square test of independence (also known as the Pearson Chi-square test, or simply the Chi-square) is one of the most useful statistics for testing hypotheses when the variables are nominal, as often happens in clinical research (McHugh 2013). The result of the Chi-square test, $p = 0.00$ is significant since $p \leq 0.05$ and is shown in Table 5.5 below.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	784.637[a]	456	.000
Likelihood Ratio	337.673	456	1.000
Linear-by-Linear Association	19.986	1	.000
N of Valid Cases	157		

a. 507 cells (100.0%) have expected count less than 5. The minimum expected count is .01.

Table 5.5 The result of chi square test between land use type and cases

b) Vegetation index (VI)

According to Gao (1996), "the Normalized difference vegetation index (NDVI) is the most widely used index for remote sensing of vegetation". The Vegetation index (VI) data can be derived from remote sensing satellite images and represents the vigour and density of vegetation on earth. To obtain the vegetation index for the study area within the study time frame, the first step is to find a good quality remote sensing data that fulfil the criteria.

Previous studies on tick borne disease have indicated that high vegetation index (VI) has significant relationship with the high number of tick infestations, which may lead to infections to human and animals. Based on a study on Lyme disease in the USA, human case distribution by county of exposure was significantly correlated with tick distribution; both were positively correlated with high NDVI values in spring and fall, when wooded vegetation could be distinguished from agricultural crops in the satellite image (Kitron and Kazmierczak 1997). Agoulon et al. (2012) have noted in their study that tick density was related significantly to a Vegetation Index (V.I., ranging from 1 to 5) that took into account the abundance of trees and bushes on the edge of pastures: most ticks (57%) were found in transects with the highest V.I. (covering 15% of the explored surface in the study area). While Estrada-Peña et al. (2014) have inferred the abiotic niches of four species of Neotropical ticks, *Amblyomma mixtum*, *Amblyomma cajennense*, *Amblyomma tonelliae* and *Amblyomma sculptum* in their studies using coefficients of a harmonic regression of the temperature and the Normalized Difference Vegetation Index (NDVI, reflecting plant stress) from remotely sensed data from MODIS satellites with 0.05° spatial resolution. For this study, vegetation index (VI) for infected areas were gathered and analysed to find out the values of vegetation index. Since the natural habitats of tick are in the forest with high vegetation density, it is assumed that the values of VI in the infected areas will be high.

As mentioned in Chapter 4, even though there are two types vegetation index, the NDVI and EVI, the EVI has been the chosen for this study since it suits areas with high vegetation cover as well as cloud cover.

Graphs were plotted for visual analysis to determine if there is any relationship or pattern between EVI values and cases as in Figure 5.6 below. However, there is no visible pattern even though EVI or high vegetation has been proven in previous literature as one of the factors for predicting tick abundance. However as shown in Figure 5.6 below, there is no significant pattern and it has shown an inverse relationship for certain cases, which are also detected in areas with low EVI values. The graphs for the year 2003, 2005, 2006 and 2007 can be referred in the Appendix. The EVI analysis for the year 2002 is unavailable since there are missing records on certain dates of patients; therefore this research is unable to identify the corresponding EVI.

Through this investigation, it was found that the types of land use are very mixed even within such small areas.

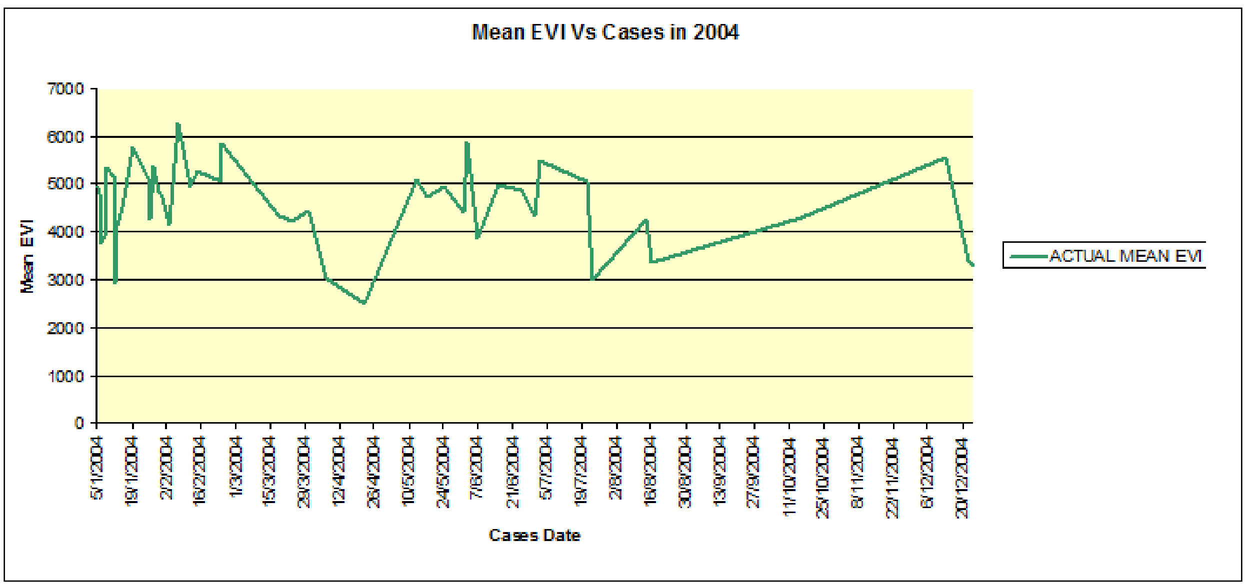


Figure 5.6 Graph showing Mean EVI values for each human otoacariasis cases in the year 2004

Since the value of VI cannot be used directly to determine a vegetation connection to tick infections, species distribution modelling with pseudo absence techniques was used. In this technique, the goal is to compare the difference between areas with presence or actual infected areas with absence or areas with no infections. For better and unbiased results, two sets of random points or absence data are created and compared to presence data. These random points are generated automatically using the random point generator tool in ESRI's ArcMap software. The results of the analysis of EVI for actual (presence data) and pseudo absences are shown as below.

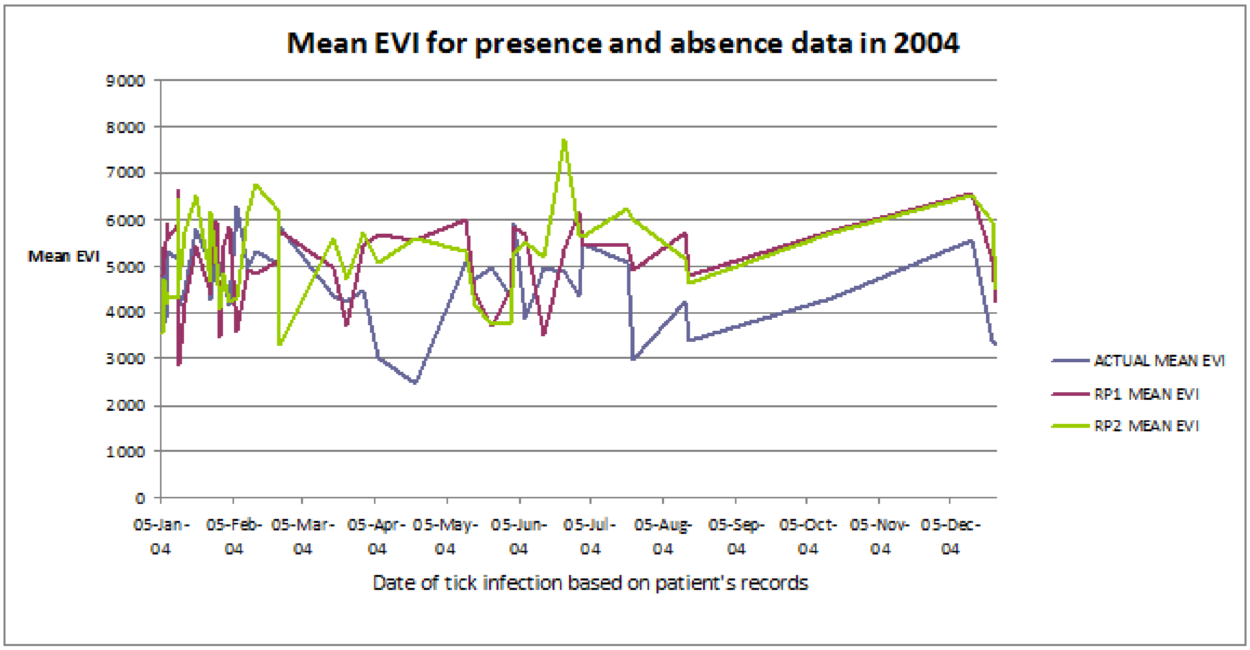


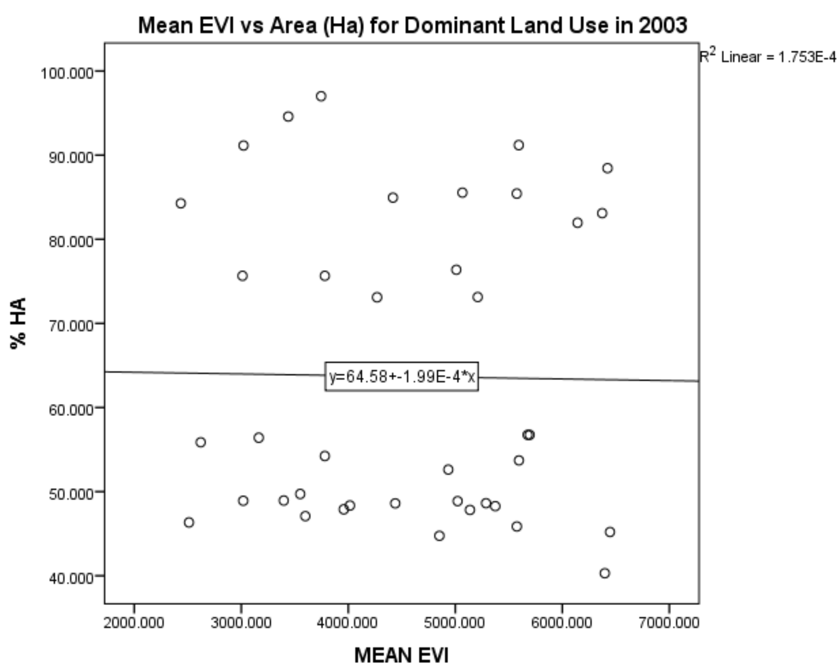
Figure 5.7 Graph showing comparison between Mean EVI for presence and absence data for the year 2004

As Vegetation Indices (VI) has been identified as one of the factors of tick borne disease in the literature, therefore, the EVI values for the study area have been extracted from MODIS satellite data. Statistically Independent T-Tests was carried out for the EVI data from 2003-2007. EVI data from the year 2002 was not available for analysis since there were only five cases with date's information recorded. Therefore the number of cases for 2002 is not sufficient enough for testing and comparison.

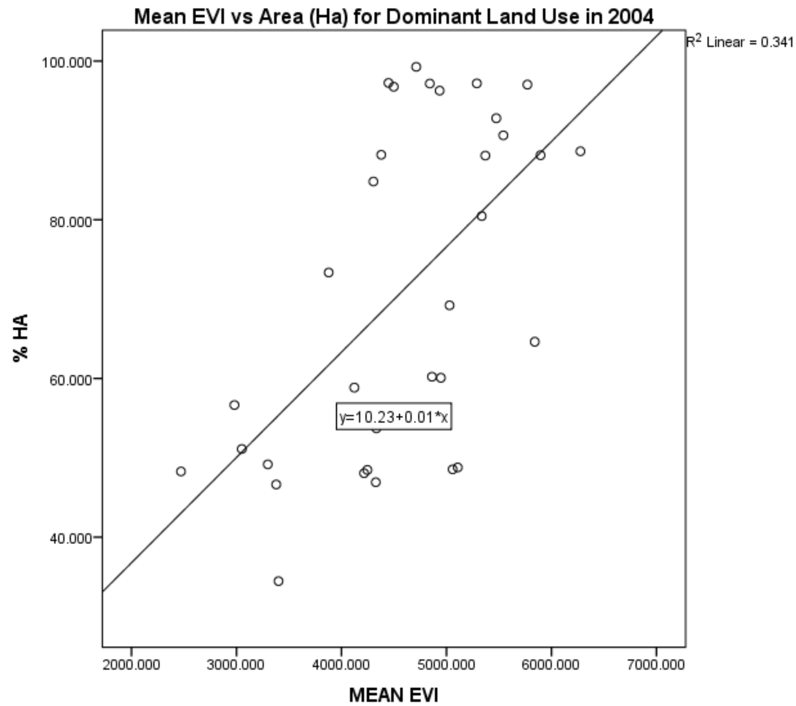
5.3 Developing the CART model

Three factors have been identified: 1) distance from the river which has been described on Chapter 4; 2) land use; and 3) three types of land use: oil palm, urban and swamp areas. As described in Chapter 4, since the significant factors involved both categorical and continuous factors, therefore the most suitable technique is to apply the CART model with logistic regression.

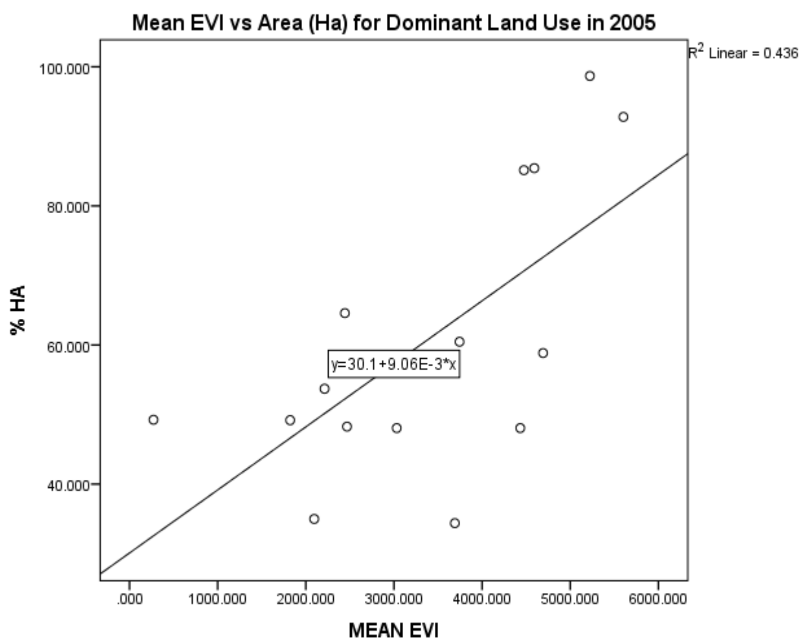
The first analysis performed was linear regression to determine the significance of the three dominant land use type as factors for tick borne disease. Analysis is performed for data from the year 2003 until 2007, except for the year 2002 due to missing records on dates of infection. The relationship graph plotted for each year is shown in Figure 5.8.



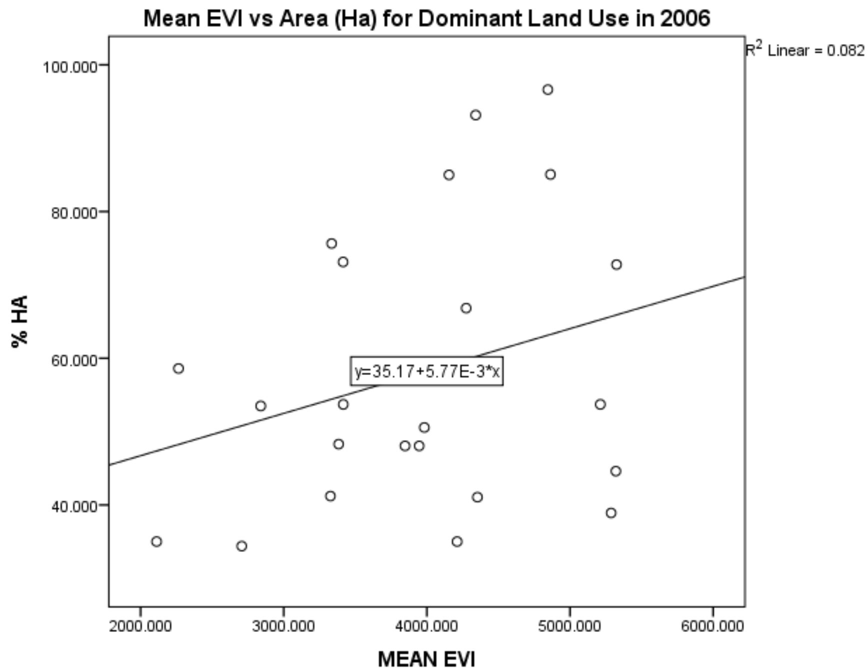
a) Mean EVI vs area (Ha) for dominant land use in 2003



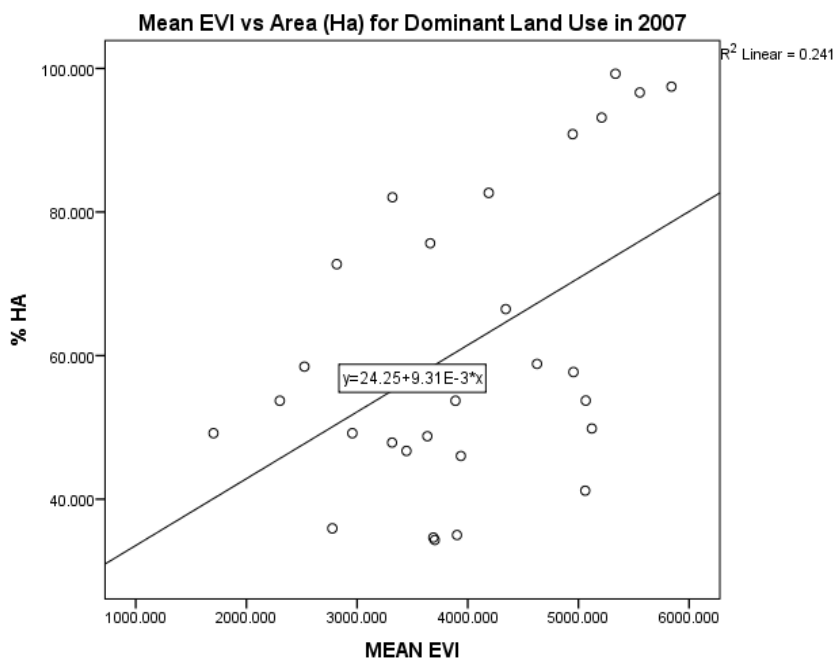
b) Mean EVI vs area (Ha) for dominant land use in 2004



c) Mean EVI vs area (Ha) for dominant land use in 2005



d) Mean EVI vs area (Ha) for dominant land use in 2006



e) Mean EVI vs area (Ha) for dominant land use in 2007

Figure 5.8 Linear regression between oil palm, urban and swamp for 2004

Table 5.6 shows the result of linear regression analysis between swamp, oil palm and urban against the Mean EVI for the year 2003 until 2007. The values of p shows significant relationship between EVI and the three dominant land use types, but only for three years, which are in 2004, 2005 and 2007.

Year	p value
2003	0.936
2004	0.000*
2005	0.007*
2006	0.186
2007	0.006*

*: The variable is statistically significant since $p \leq 0.05$

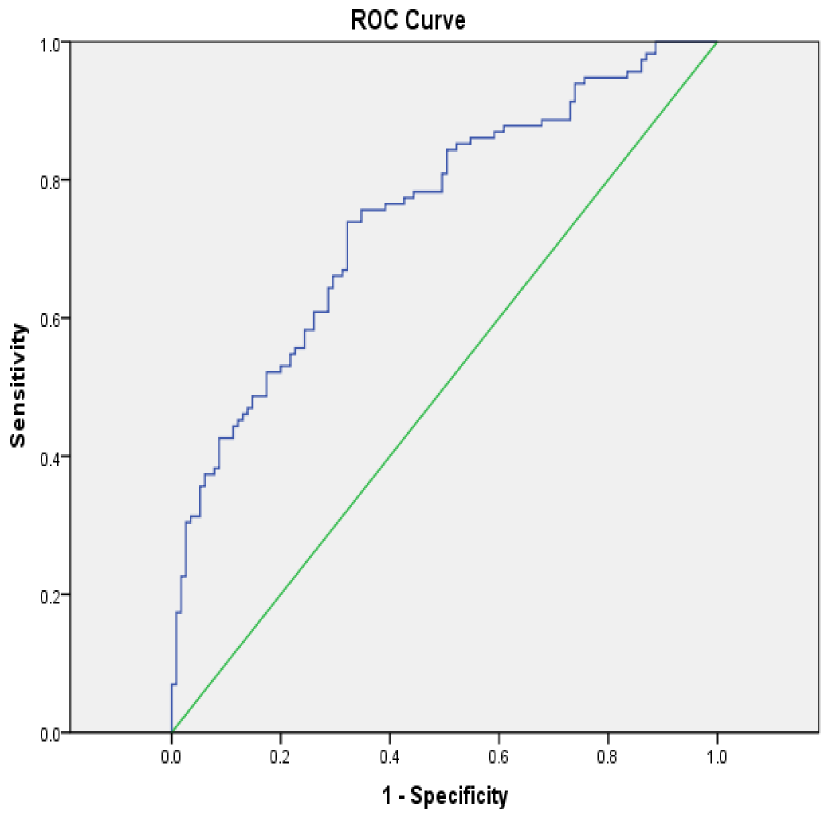
Table 5.6 The result of linear regression analysis for the three dominant land use types

Logistic regression is widely used in the medical literature for analyzing binary outcome data and has many similarities to linear regression, but it is more complex and harder to evaluate graphically (Sainani, 2014). Logistic regression blends two very different statistical traditions: one is the analysis of contingency tables (cross-tabulations or crosstabs) and secondly is ordinary least squares (OLS) multiple regression analysis (Menard, 2010). According to Debanne and Rowland (2002), an important difference between this technique and linear regression is that the latter requires that the outcome, or dependent variable be continuous - normally distributed, actually - whereas a logistic regression requires that the outcome variable be dichotomous.

Thus, to prove that the EVI for the three land use types are significant factors for tick infection and suitable for modelling, a logistic regression statistical analysis was performed. The result of analysis is found to be statistically significant.

Meanwhile, an ROC test was performed to prove that the factors are good enough for modelling. The graph in Figure 5.9 shows the ROC curve plotted for presence data (actual cases of human otoacariasis) with Area Under the Curve (AUC) of 0.75. Meanwhile the graphs in Figure 5.9 and Figure 5.10 shows the ROC curve plotted for two sets of pseudo absences or random points. The figure shows the relationship of sensitivity (true positive rate) versus 1-specificity (false positive rate) for two sets of random points. The maximum value for the AUC is 1.0, thereby indicating a (theoretically) perfect test (i.e., 100% sensitive and 100% specific). An AUC value of 0.5 indicates no discriminative value (i.e., 50% sensitive and 50% specific) and is represented by a straight, diagonal line extending from the lower left corner to the upper right (Fan, Upadhye and Worster 2006).

The area under the curve (AUC) is a measure of the model's ability to correctly distinguish between presence and absence and can be interpreted in a straightforward manner. According to Marzban (2004), an AUC of 0.5 reflects random forecasts, while AUC of 1 implies perfect forecasts. Meanwhile, Hosmer, Lemeshow and Sturdivant (2013) have stated that the AUC value of 0.5 means no discrimination, the value between 0.5 and 0.7 means poor discrimination, the value between 0.7 and 0.8 are acceptable and the value between 0.8 and 0.9 are excellent while the value above 0.9 signifies outstanding discrimination.



Area Under the Curve

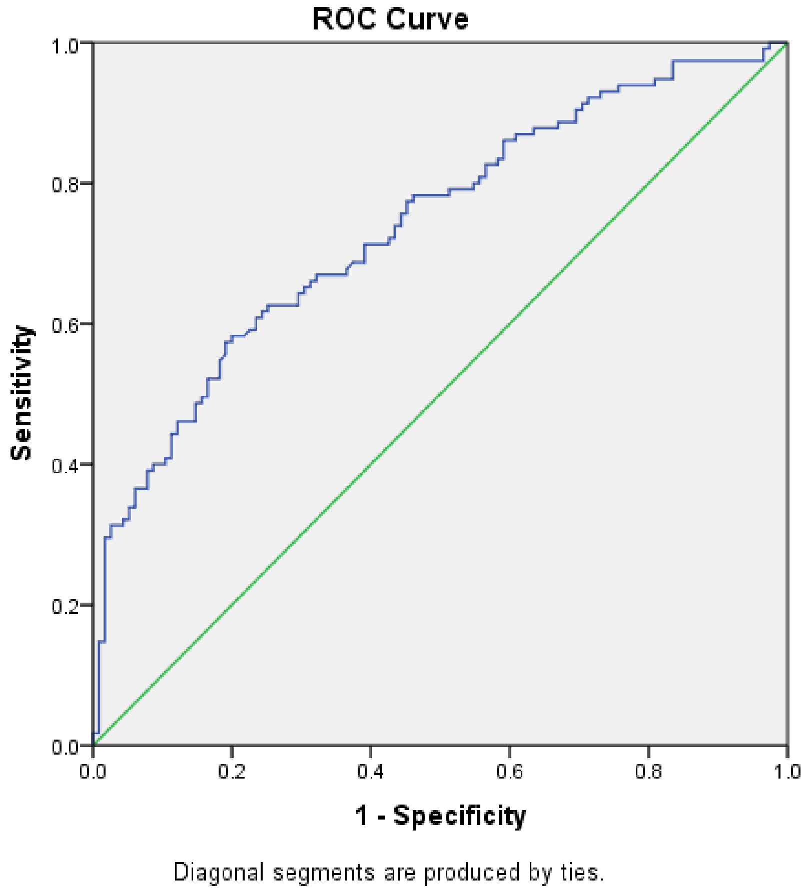
Test Result Variable(s):Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.750	.032	.000	.688	.812

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Figure 5.9 The ROC plot for presence data (actual cases of human otoacariasis)



Area Under the Curve

Test Result Variable(s): EVI

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.739	.032	.000	.675	.802

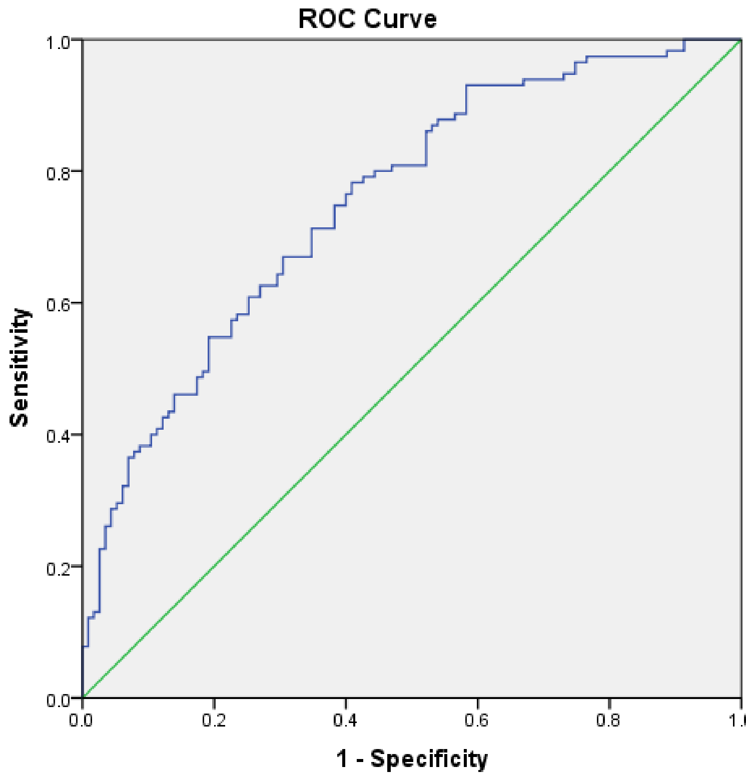
The test result variable(s): EVI has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Figure 5.10 ROC plot for first set of random points (pseudo-absence data)

Based on Figure on the first ROC analysis for first set random points, value of Area Under the Curve (AUC) is 0.739 and for the second set is 0.754, while the AUC for presence data is 0.75. These values indicate that the diagnostic test is fair and acceptable since they fall between 0.7 and 0.8. Therefore the shorter distance from the river, the three types of land use (oil palm, swamp and urban area) as well as the vegetation index or EVI are significant predictors and can be used for the modelling and prediction of tick borne disease.



Area Under the Curve

Test Result Variable(s): EVI

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.754	.031	.000	.692	.815

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Figure 5.11 The ROC plot for the second set of random points (pseudo-absence data)

The map in Figure 5.12 shows the distribution of cases from the year 2002 until 2007 along with the significant factors. As can be seen from the map, the cases are distributed within the areas where the factors are present.

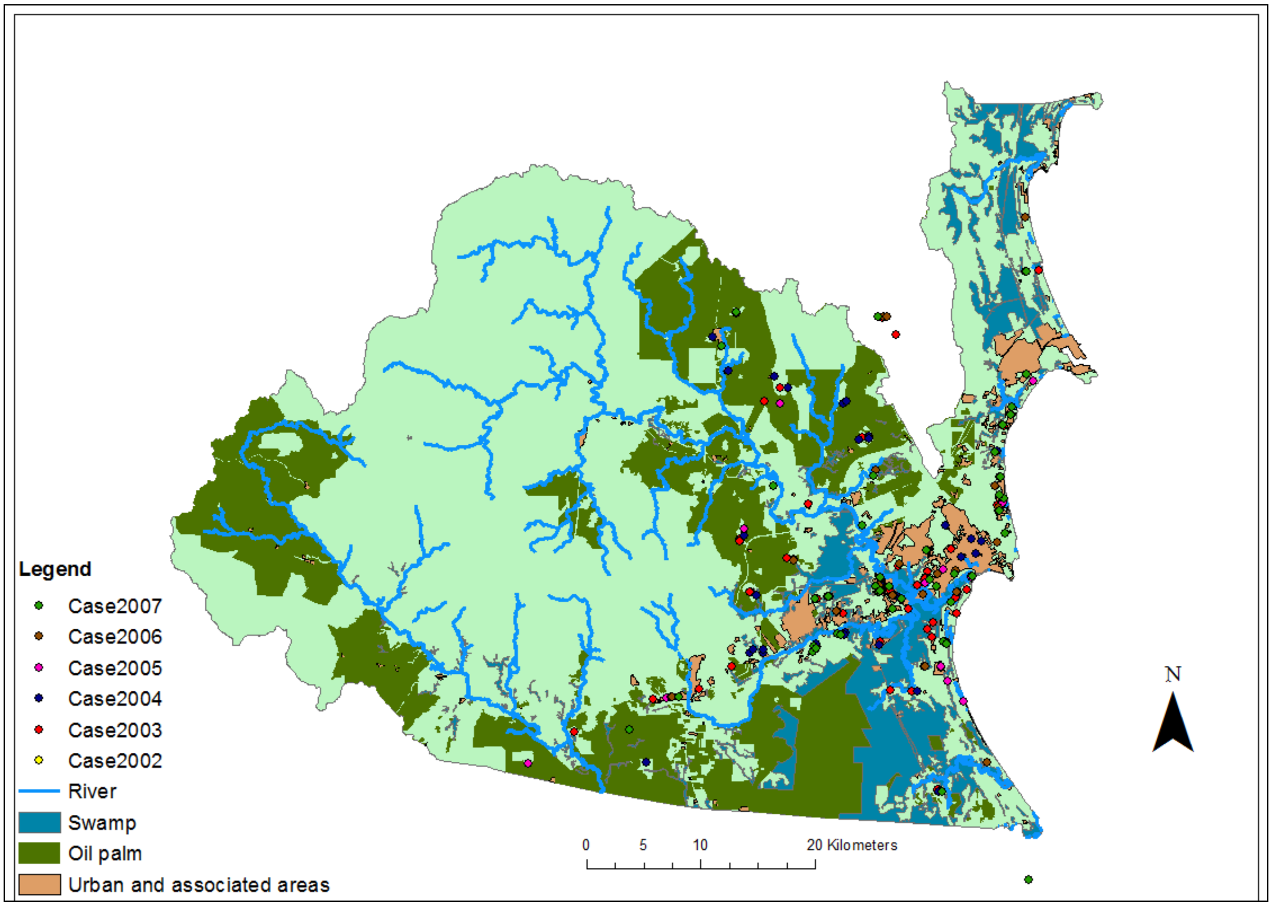


Figure 5.12 Distribution of cases against significant factors of tick borne disease

Therefore to describe the relationship, the conceptual modelling for prediction of tick borne disease can summarised as in Figure 5.13.

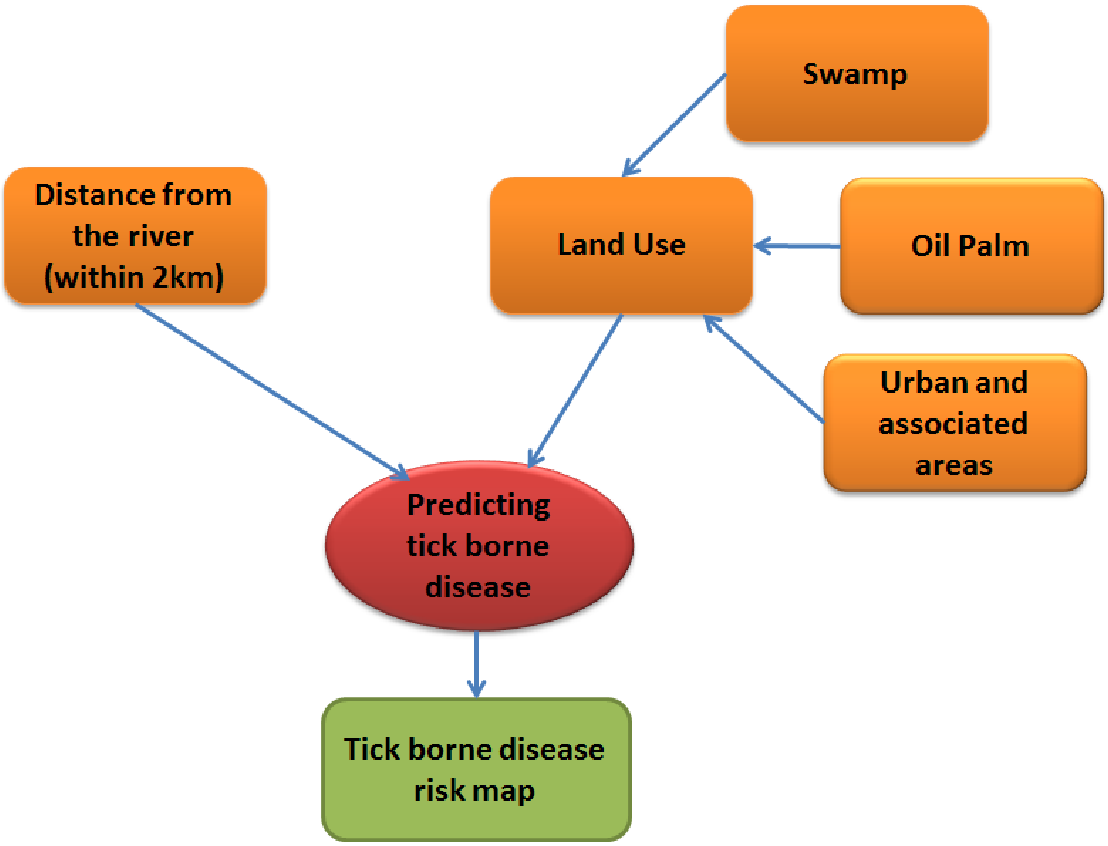


Figure 5.13 The summarised concept of modelling for prediction of tick borne disease

5.4 Chapter summary

In this chapter, the suspected factors of tick infection have been identified and analysed based on the best available data for this research. Several of the factors included for analysis have been mentioned in the literature as significant factors of tick borne disease. However, these parameters need to be analysed statistically since the conditions of the study area differ from the study areas in the literature. The parameters identified for this study consist of environmental factors such as temperature, rain, wind and relative humidity as well as analysis on land use type and river.

The result of statistical analysis has shown that there are three significant factors for tick infections which are: distance from the river, land use and, in particular, three types of land use which are oil palm plantation, urban areas and swamp areas. These three factors can be integrated into modelling to predict future risk of tick borne disease in the future. Since significant factors involve categorical and quantitative data, the most suitable modelling technique suggested for the combination of data is a Classification and Regression Tree (CART) with logistic regression. To test the suitability of this modelling, a diagnostic test has also been performed. Therefore, by using the mentioned method with the presence of significant factors, we are able to predict the risk of disease in the future. A suitable modelling method has been identified is CART and logistic regression which is able to predict the occurrence of tick borne disease based on the presence of the significant factors consisting of 3 land use type (urban, swamp and oil palm areas), shorter distance to the river and vegetation (EVI). This information will be useful to the decision makers and health authorities to make necessary plans and precautions to manage and prevent disease outbreak in the future.

6 DEVELOPMENT OF A PROTOTYPE FOR WEB-ENABLED SDSS FOR TICK BORNE DISEASE

This chapter presents the development of a prototype sytem for a Web-enabled Spatial Decision Support Systems (SDSS) for the prevention of tick borne disease in Kuantan District of Pahang, Malaysia. The development of a Web-enabled prototype of the system, its components and architecture with features and functions that can be implemented and developed further as a full system.

6.1 Significance of Web-enabled SDSS for Tick Borne Disease

Web-enabled SDSS has been applied in many fields of applications including in disease control and prevention of disease outbreak such as controlling animal diseases outbreak in Egypt ((Bakr et al., 2015). In Malaysia, development of web-enabled SDSS for disease has been attempted such as prototype development of MyGeoTBIS for tuberculosis (TB) (Rasam et al., 2021). Another system for the disease is more focused towards web-GIS for plotting of tuberculosis cases in the Gombak District in Selangor (Mohidem et al., 2021). Mosquito born disease such as dengue has also a online systems such as iDengue for community and Online Dengue Outbreak Management System (DOMS) specifically for Ministry of Health (MoH) Malaysia. However to date, especially in Malaysia there is no GIS or web-enabled SDSS particularly for tick borne disease has been developed.

Currently, the method of decision making by the government decision makers is done manually which is by analysing report and status of disease. The report comes from district and state level health authorities and monitored at the national level by the Ministry of Health (MoH). At present, there is no GIS based system being implemented, for example no representation of location of tick borne disease in a map which will help decision makers to make better visual interpretation and overview of the situation. Thus, the study will propose a method to improve the current method of decision making by development of a web-enabled SDSS for tick borne disease based on the GIS technology.

The government decision makers will be able to make better and informed decisions when more data is provided such as more information about disease and how to control and prevent any disease. Besides report and statistics, by applying the current GIS technology, tick infections cases can be visualised in a GIS map to show location and time of infection such as cases reported in months or years which is more user friendly instead of presented in texts and numbers. GIS map is not just able to present data better than text and statistics, decision makers can also view analysis of trend of infection, hotspot analysis as well as factors contributing to infection which can assist in prevention of tick borne disease outbreak in the future.

The web-enabled SDSS will improve decision making as it is accessible to the users through the internet. Information can be represented in a better way to suit the different types of users which may require complex to simple information that can help them make decisions. To develop the system, three major components have been identified.

6.2 The Components of Web-enabled SDSS for Tick Borne Disease

There are three integrated components that have been identified in development of the web-enabled Spatial Decision Support System (SDSS) for tick borne disease. The components are Spatial Database Management System, Decision Support System and Tick borne Disease Management System as illustrated in Figure 6.1.

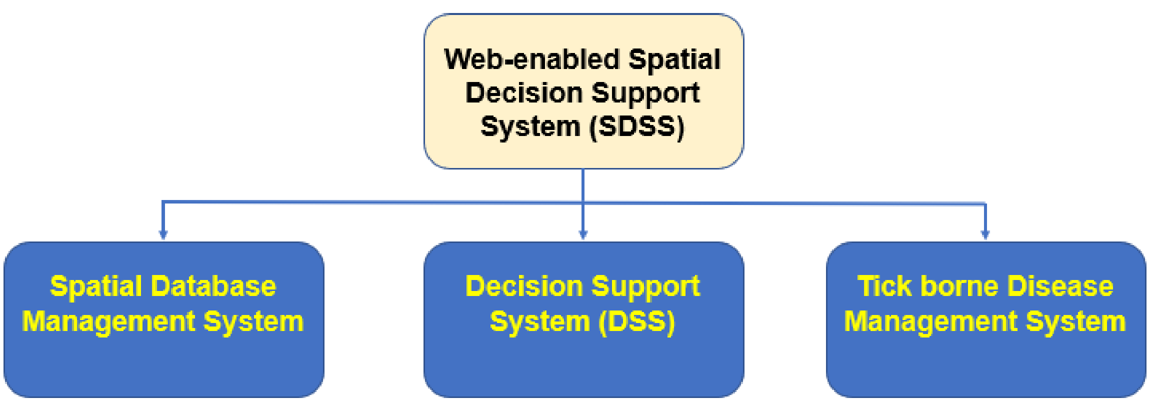


Figure 6.1 The three components of a web-enabled SDSS for tick borne disease

a) Spatial Database Management System

The Spatial Database Management System consists of a GIS Spatial database and GIS Core Components, which handle GIS data query and generate spatial maps as output. A GIS can be built as the front-end for any Spatial Database Management System and its applications consists of tools that allow users to create interactive queries (user-created searches), analysing spatial information, editing data of maps, and presenting the results of all these operations (K. and Wadhwa 2016).

The spatial database management system for tick borne disease stores and manages spatial data such as remote sensing satellite, satellite-derived EVI from the MODIS satellite, land use maps, river and road networks, forest maps, administrative boundaries; and point data representing the location of tick infection cases.

There are two types of spatial data: the geometry and geography data type. The geometry data type supports planar, or Euclidean (flat-earth), data and conforms to both Open Geospatial Consortium (OGC) Simple Features for SQL Specification version 1.1.0 and is compliant with SQL MM (ISO standard). In addition, SQL Server supports the geography data type, which stores ellipsoidal (round-earth) data, such as GPS latitude and longitude coordinates (Microsoft, 2021).

The difference between normal database and spatial database management system is, unlike classical relational database management system which store data as numbers, alphabets, alphanumeric or even symbols, spatial databases store abstract data types (ADTs) such as points, lines, polygons, coordinates, topology or other data types that can be mapped (Samson et.al 2017). The difference between the types of databases is described in the table below:

Database Type	Description	Examples
Database Management System (DBMS)	A database management system (DBMS) is a software package with computer programs that control the creation, maintenance, and use of a database.	File systems, xml
Relational DBMS (RDBMS)	is a DBMS that is based on Relational model that stores data in tabular form	SQL Server, Sybase, Oracle, MySQL, IBM DB2, MS Access
Spatial DBMS (SDBMS)	store abstract data types (ADTs) such as points, lines, polygons, coordinates, topology or other data types that can be mapped	ArcGIS Server
Geodatabase	a collection of geographic datasets of various types held in a common file system folder, a Microsoft Access database, or a multiuser relational DBMS (such as Oracle, Microsoft SQL Server, PostgreSQL, Informix, or IBM DB2)	ArcGIS Server

Table 6.1 The different types of databases

Spatial data types store the spatial attributes that allow data visualisation on a map. Many databases automatically include spatial data types but others may require some configuration or an installation to use a spatial data type. Table 6.2 shows the list of databases supported by ArcGIS, spatial data types supported in each, and what type of configuration is needed to use a spatial data type in each database (ESRI).

Database	Spatial data type	Configuration
Altibase	Geometry	User have to create two system tables in ALTIBASE
Dameng	ST_Geometry	User have to initialize it
IBM DB2	ST_Geometry	No configuration is required
IBM Informix	ST_Geometry	Require registration of database with Informix Spatial DataBlade.
IBM Netezza Data Warehouse Appliance	Netezza Spatial Package	ArcGIS requires that the spatial column in the table be named shape. If the name of the column is not shape, create a view on the table and set the alias of the spatial column to shape.
	Netezza Spatial Esri Package	spatial package used affects the configuration of the Netezza ODBC driver to connect to Netezza from client applications.
Microsoft SQL Server and Microsoft Azure SQL	Geometry and Geography	No configuration is required.

Database	Spatial data type	Configuration
Oracle	ST_Geometry and SDO_Geometry	Configuration is required by using the Create Spatial Type geoprocessing tool to install the ST_Geometry spatial type in Oracle database.
PostgreSQL	ST_Geometry, PostGIS geometry, and PostGIS geography	Installation is required in the database
SAP Hana	ST_Geometry	Included in default data warehouse installation
SQLite	ST_Geometry	Either install SpatiaLite or use the Create SQLite Database geoprocessing tool to create a GeoPackage or SQLite database that uses ST_Geometry or SpatiaLite storage.
Teradata Data Warehouse Appliance	Included	No installation is required but ArcGIS requires that the spatial column in a Teradata feature class be named shape**.**

Table 6.2 List of databases supported by ArcGIS

Spatial database allows spatial data queries such as spatial range queries, nearest-neighbour queries and spatial join queries and the applications of spatial data is not limited

Geographics Information Systems (GIS), but also includes Computer-Aided Design/Manufacturing and Multimedia Databases ((Ramakrishnan & Gehrke, 2003).

All the databases mentioned can be deployed as standalone in a desktop computer, or installed and accessible remotely through a local network or made available through the internet (web-enabled database).

Web-enabled spatial database systems are database systems designed especially for access over the Internet using the World Wide Web (Web) protocol (Yeung and Hall 2007). In this case, the web-enabled SDSS for tick borne disease prototype is built on ArcGIS server technology and hosted in ArcGIS Online, utilizing the web-enabled spatial database system and therefore is accessible through the internet.

According to Farooq, a good spatial database management software package should be able to:

1. Scale and rotate coordinate values for "best fit" projection overlays and changes.
2. Convert (interchange) between polygon and grid formats.
3. Permit rapid updating, allowing data changes with relative ease.
4. Allow for multiple users and multiple interactions between compatible databases.
5. Retrieve, transform, and combine data elements efficiently.
6. Search, identify, and route a variety of different data items and score these values with assigned weighted values, to facilitate proximity and routing analysis.
7. Perform statistical analysis, such as multivariate regression, correlations, etc.
8. Overlay one file variable onto another, i.e., map superpositioning.
9. Measure area, distance, and association between points and fields.
10. Model and simulate, and formulate predictive scenarios, in a fashion that allows for direct interactions between the user group and the computer program.

b) Decision Support System (DSS)

The Decision Support System (DSS) is an important component in the prototype system. The design for the system is made specifically according to the needs of three (3) users category or stakeholders specifically the government decision makers, the health workers and the researchers which is summarised in Figure 6.2.

Each stakeholders have their own respective and important roles. The government decision maker, the Ministry of Health (MoH) is the regulator and policy maker in Malaysia. The ministry is the main provider of health care services to the public, whose main objective is to provide a greater network of physical facilities, equity, accessibility and utilization of health care resources (Thomas et al., 2011). The health workers role is in providing health care services to the public. The health workers consists of the doctors, nurses and health practioners from government and private-owned hospitals, clinics and pharmacies who provide consultation and treatment to patients. They are frontliners who attend to patients and will be the first to detect the occurrence of disease. These workers may report if there is any alarming increase of a certain disease to the higher authority, in this case to the district level health authority which will forward the report to the Ministry of Health (MOH). The MOH will gather all inputs and to assess the situation whether a certain reported disease may become a concern in the future or may lead to outbreak before they can decide on taking necessary actions as control measures. On the other hand, researchers play important roles in terms of investigating from the tick species as vector itself, existing and the potential of new tick borne disease, pathogens, treatment and prevention of tick borne disease. The scope of research is not limited to biological and medical aspects, but will be broader to cover all aspects including improvement of the system and optimisation of decision making process. Hence, there are six research institutes established by the government under the National Institutes of Health consisting of Institute for Medical Research, Institute for Clinical Research, Institute for Public Health, Institute for Health Systems Research, Institute for Health Management and Institute for Health Behavioural Research. At the same time there are other research bodies undertaking studies on medical microbiology such as the Tropical Infectious Diseases Research and Education Centre

(TIDREC), a department under the University of Malaya and the Malaysian Space Agency (MYSA) under the Ministry of Science, Technology and Innovation (MOSTI) which specialises in GIS research and mapping technology using data derived from remote sensing satellite data.

In the Decision Support System (DSS), the government decision makers require strategic information for the control and prevention of disease outbreak. The DSS component for the system will be able to fulfill the users requirement in accessing information on the current status of tick related incidence, hotspots.

The health workers need to access information on ticks, species that may cause disease and type of disease, view reported cases in the district or nearby states or at national level.

The researchers provide further information on tick species and their disease. The prototype system will enable the researchers to download data from system (historical records eg. patients records), meteorological data including temperature, rainfall, population data, land use data, statistics, and reports from other researchers. The data from the system may assist in performing trend analysis and modelling. Submit their report and findings or new method or discovery on tick borne disease.

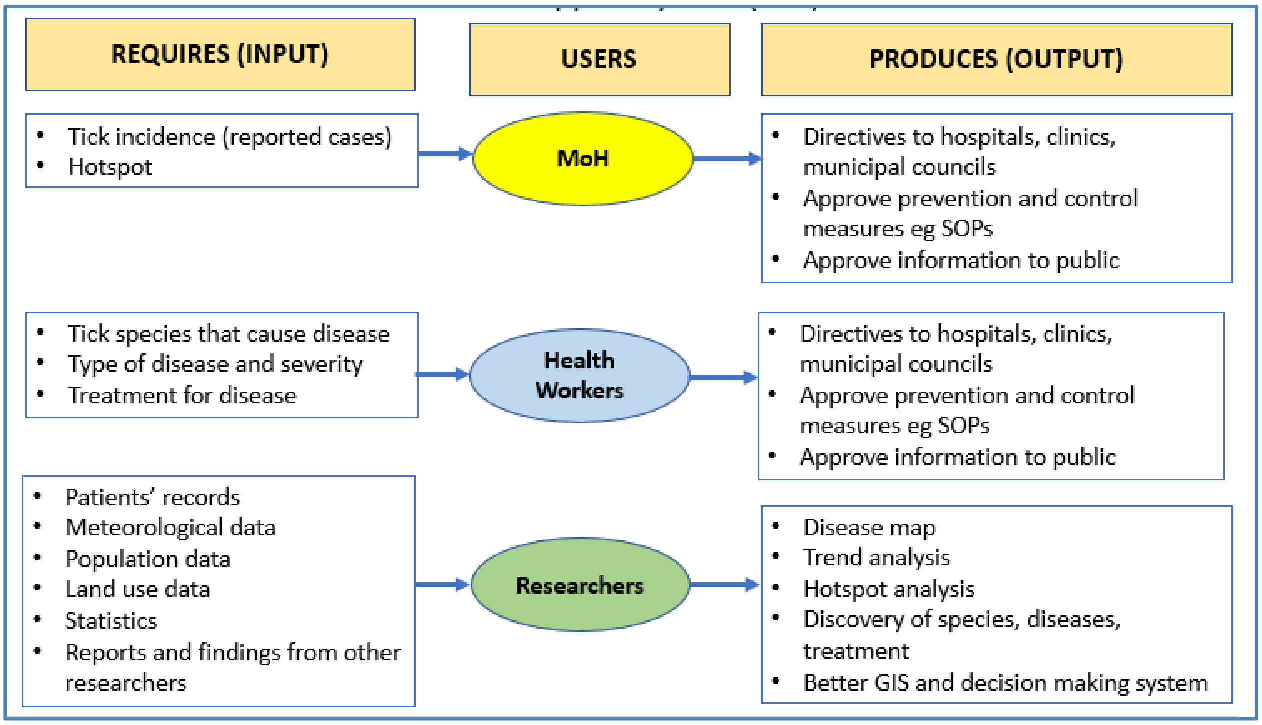


Figure 6.2 DSS Users

Tick borne Disease Management System manages and store related data on tick borne disease. Data includes background data on ticks and its diseases, how tick bite occurs, types of disease it may cause and the suspected factors of tick borne disease. In this case existing records and data relating to the spatial and temporal distribution of outbreaks and also further information found in the literature review can be included in the system. This data can be stored as records and also in the form of spatial maps to visualize disease pattern and distribution. The web-enabled SDSS system can then provide the facilities for users to retrieve, provide, as well as share these data and information with other users within the system. Users can submit information, retrieve and share information about tick and diseases, symptoms, risk level and actions to be taken to prevent the disease from spreading. Data and information can be shared in spatial map and visualize pattern and distribution of disease. Three types of users are able to benefit from the system by providing information to the government authorities to devise better control and prevention strategies of tick borne disease. At the same time health workers need the information to provide the correct treatment while the researchers may utilise the existing data to do more studies on ticks and tick borne diseases, find cure or develop better treatment and prevention methods.

6.3 Web-enabled Spatial Decision Support System (SDSS) for tick borne disease

The integration of both GIS and DSS technologies has resulted in SDSS, which harnesses the decision analytic power of DSS and the spatial capabilities of GIS (Sugumaran and Sugumaran 2005). SDSSs are integrated computer systems that support decision makers in addressing semi-structured or unstructured spatial problems in an interactive and iterative way with functionality for handling spatial and non-spatial databases, analytical modelling capabilities, decision support utilities such as scenario analysis and effective data and information presentation utilities (Sugumaran and DeGroote 2010). SDSSs are integrated frameworks designed to help explore weakly structured or unstructured problems characterised by many actors, many possibilities, and high uncertainty (Rutledge et al. 2007).

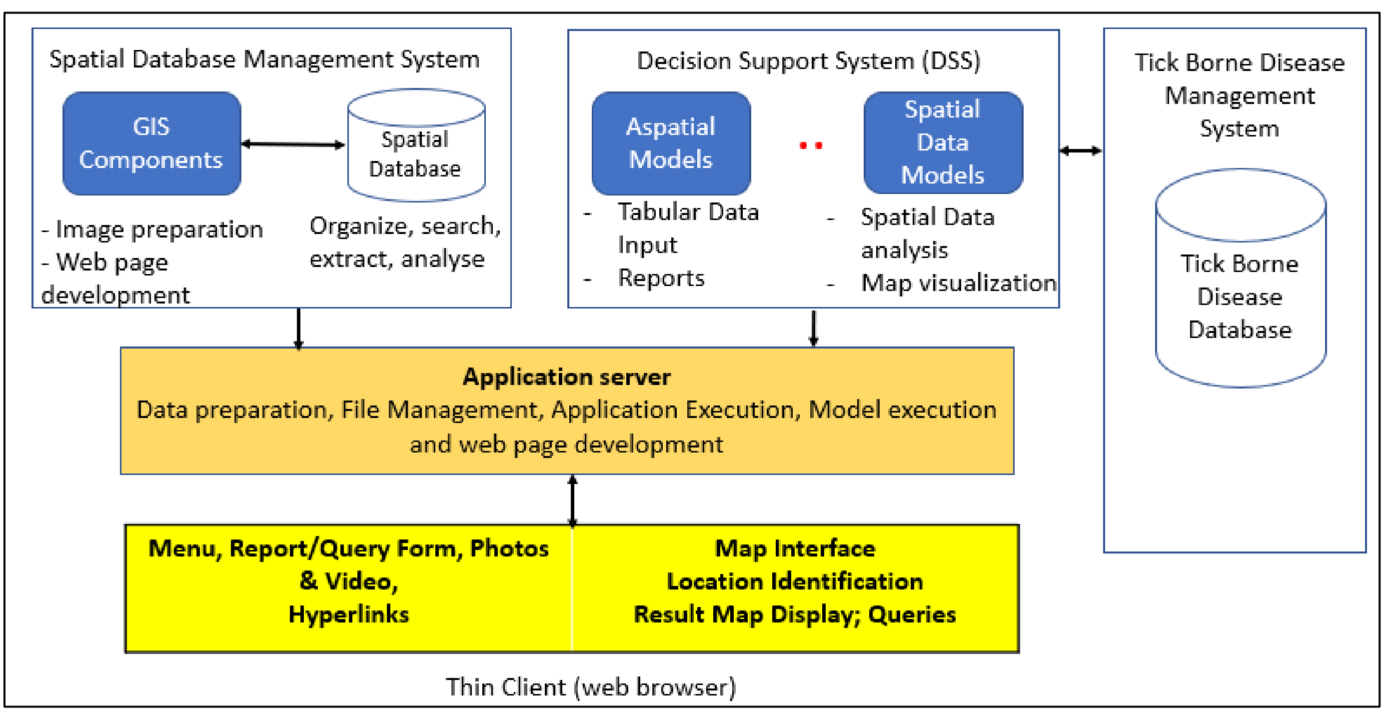


Figure 6.3 Schematic Representation of Web enabled SDSS for tick borne disease

Figure 6.3 shows the three components of Web-enabled SDSS for tick borne disease consisting of the Spatial Database Management System, Decision Support System and Tick Borne Disease Management System. The Spatial Database Management System comprises of spatial database and GIS components. The spatial database allows spatial data manipulation such as organize, search, extract and analyse spatial data. The GIS components utilise the spatial data for image preparation and web page development as an output.

The second component, Decision Support System (DSS) consist of aspatial models and spatial data models. The aspatial models manage tabular data input as well as generate text based reports which may include graphics and statistics. The spatial data models provides the medium for spatial data analysis, modelling and map visualisation. The third component, tick borne disease management system supports the DSS with a database which manage data on tick borne disease.

To access the system, users may use a web-browser interface and select menu, view and download report or query, photos, videos and hyperlinks. For spatial map display, a user friendly map interface allow users to perform functions such as location identification, including query and display the resulting map. All the requests or queries made by users are submitted via HTTP to the Application Server, which process the request accordingly from the three components: Spatial Database Management System, Decision Support System and Tick borne Disease Management System. The Applications Server accepts request from users and serves as interface between the three components and users requests. It performs task such as data preparation, file management, application execution, model execution as well as web page development.

The strength of the DSS framework proposed is its capability to assist decision-makers in making efficient and quick decision, which can beachieved when necessary data and information are made available to thedecision-makers at any time required. The DSS is also designed to provide services to different types of users and may provide a platform for the exchange of information among them. However, the weakness of the framework is the management of the system will depend on the administrator who must ensure the integrity, accuracy and security of data in the system. There is also the danger of having limited data and information, which may lead to biased analysis and bad decision-making.

6.4 The Prototype Development of Web-enabled SDSS for tick borne disease

The development of the protoype system for the Web-enabled SDSS for tick borne disease is based on the ArcGIS Online platform. ArcGIS Online is a cloud-based mapping and analysis solution which allows users to to make maps, analyze data, and to share and collaborate (ESRI). The platform allows users to create web maps, applications and visualise data which can be published online. Two of the ArcGIS Online applications being utilised for the development of the prototype system are ArcGIS StoryMaps and ArcGIS Dashboards.

ArcGIS StoryMaps is a story authoring web-based application that enables people to share maps in the context of narrative text and other multimedia content. ArcGIS StoryMaps allows author stories with story builder to include maps, narrative text, lists, images, videos, embedded items and other media. These stories can be published and shared using designated URL with specific groups or just anyone. For this study, a storymap is developed and intended for the public to get the information on tick borne disease which is very informative and presented in a very interactive and user friendly manner.

On the other hand, ArcGIS Dashboards is an information management system that supports the access to and visualization of information for management of tick borne disease. It serves as a central information resource and support decision-making during response to possible tick borne disease outbreak. The dashboard functions as interface to other information system consisting of public health surveillance, land use change and weather conditions monitoring system. It supports decision-making during possible incidents or outbreaks and have functions for data access and data visualisation which benefits specific categories of users such as the government (Ministry of Health), health workers and the researchers.

6.4.1 Features and functions of dashboard

The dashboard for the prototype system is designed specifically for four (4) types of user categories: the government, health workers, researchers and the public users. Two specific features made available for these users are navigation through different data representation and geospatial presentation.

a) Navigation through different data representations

The dashboard allow users to view data which can be visualisaed using visualisation tools such as maps, tables and graphs. Tick infection cases from year 2002 to 2007 can be presented as tabular data and as well as graphs showing trend of tick infections throughout the years.

b) Geospatial presentation

The geospatial presentation refers to the visualisation of tick borne disease map in a spatial context relative to points of infection, type of land use in the areas and other spatially referenced data. The common underlying layer is referred to as a basemap, which is usually a map that contains non-utility features such as streets, highways, and administrative boundaries. Additional information is displayed in layers overlying the basemap. The basic functions for the geospatial presentations are:

- Pan and zoom to enable users to scroll across the map and zoom in or out to specific areas
- Layer control to allow users to select the layers to display and to set display attributes, such as transparency, for each layer (see below for a more detailed description of layers)
- Distance measurement between points or along a path
- The ability to identify or view GIS attribute data of a specific feature on the screen (for example, street address, parcel ID, work order number, or political boundary)
- The ability to search by GIS attribute data
- The ability to add annotations and sketches as a new layer
- Bookmarking of views to enable them to be retrieved later; and
- The ability to display detailed data by clicking on the related icon on the display

The GIS layers utilize symbols to convey additional details about the data represented in the layer. Tick infection layer includes map symbols as points to denote the location infection or address of patients. A layer can also include symbols or icons that change depending on the status of the element represented by the icon to convey additional information. Referred to as thematic mapping, this visualization can be as simple as icons that change depending on the current value of the parameter represented by the icon. The layout of the dashboard is as shown in Figure 6.4 below:

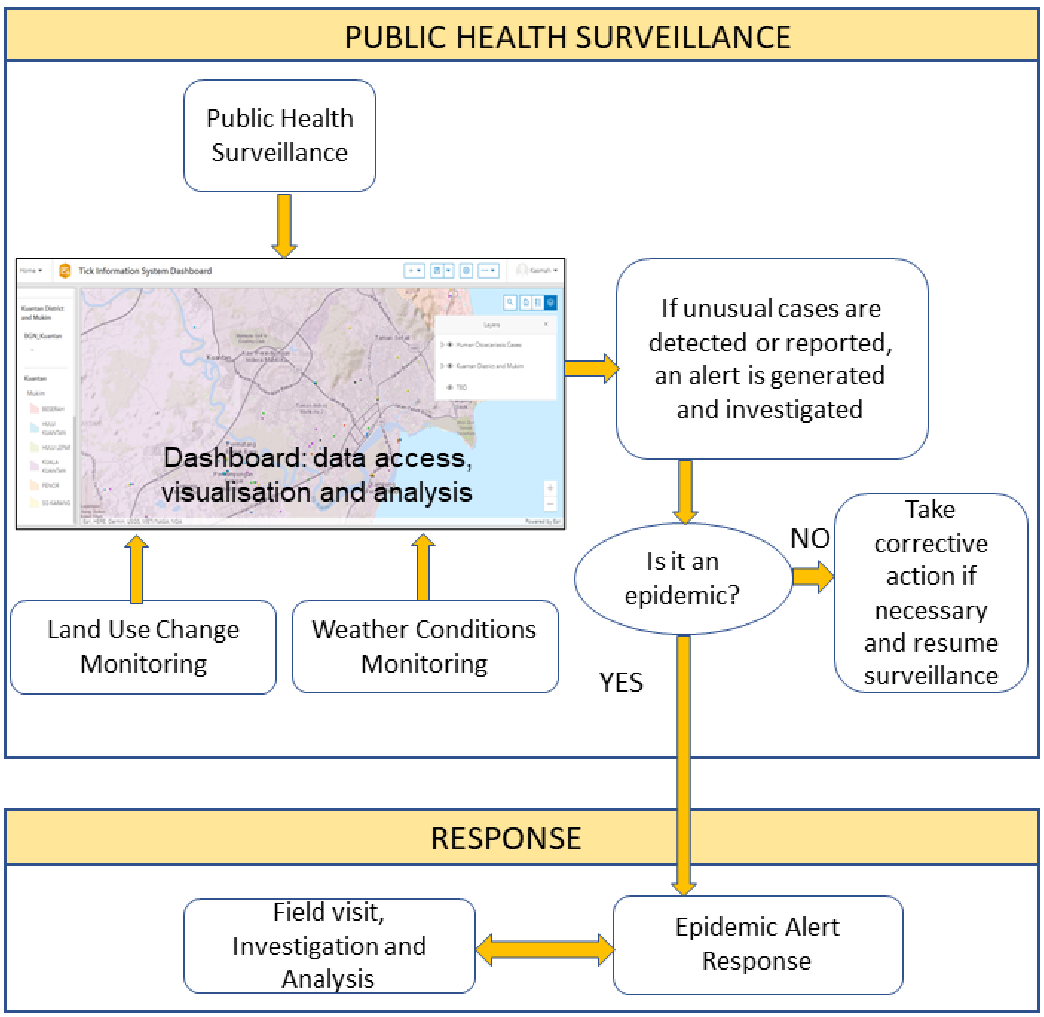


Figure 6.4 Dashboard design for Web-enabled SDSS for tick borne disease

Specific dashboard functions for dashboard (for each user category)

1. Government (Ministry of Health)

   The government users can view the location or points of infection on a map and identify tick infection hotspots. They are able to see the surrounding areas, whether infection occurs in an urban area and the location of health facilities nearby.

2. Health workers

   The health workers in the hospitals and clinics may also view the location of infections near their facilities. When more cases are detected they are able to notify the higher authority to request for further actions

3. Researchers

   Researchers may be able to view patients' records, meteorological data, population data, land use data, statistics, reports and findings from other researchers. These information may guide them in finding the possible factors of tick infection, which may due to type of vegetation, weather conditions etc

4. Public

   Public users may also want to have information on ticks and what will happen in case they get bitten, or if the ticks get in their ears (human otoacariases).

6.4.2 Dashboard Conceptual architecture

To achieve the functionality described previously, the dashboard will need to interface with the other parts of the information management system, in particular the source data systems utilized by the dashboard. The dashboard conceptual architecture describes the high-level functional elements required to meet the requirements and the ways in which these various elements must be integrated to achieve a functional system.

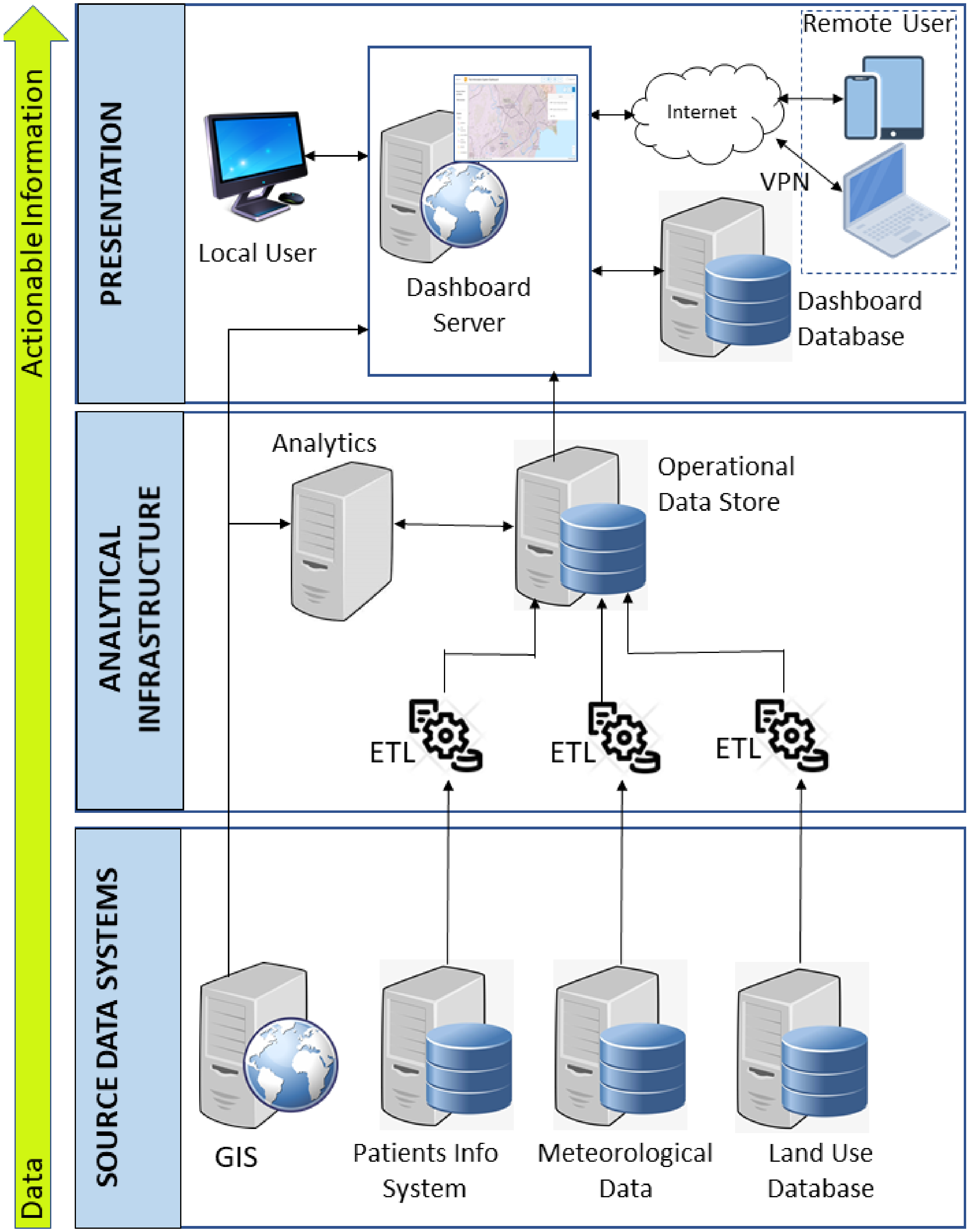


Figure 6.5 Conceptual Dashboard Architecture

The conceptual dashboard architecture illustrated in Figure 6.5 is modeled after a standard architecture that extracts basic data from multiple source data systems, transforms the source data into a format that can be more readily displayed and understood, and consolidates the data into a single repository. A conceptual architecture does not represent all of the physical elements necessary to build the system, but does identify the major elements of the system. This conceptual Web-enabled SDSS for tick borne disease dashboard architecture generally includes the following three tiers: source data systems, analytical infrastructure, and presentation. The Figure 6.5 also shows how these tiers relate and the type of source data systems that may be available to support the dashboard.

i) Source Data Systems

Each data source within the system needs a system of record (such as an application or database) that houses and manages the source data for the component. These source data systems provide data to be presented by the dashboard. Examples of possible source data systems are provided below.

Source Data System	**Data Provided to the Dashboard**
GIS	Map services that provide geospatial context or topographic features such as administrative boundaries, street addresses, roads, utility assets, etc.
Patient's Information System	Basic information on patients who seek treatment at the health facilities. (age, gender, address, date of treatment).
Meteorological data	Meteorological data including temperature, rainfall, humidity, wind speed, current weather which are collected from meteorological station in Kuantan.
Land Use Database	Land use information derived from remote sensing satellite data which consists of type of land use in the area such as urban, vegetation, forest etc.

Table 6.3 Source Data System for Web-enabled SDSS

ii) Analytical Infrastructure

The analytical infrastructure serves as the interface between the source data systems and the presentation tier. Without this tier, the dashboard would need to retrieve data directly from a source data system in response to every user request, and the resulting traffic could impair the performance of the source data systems or result in data transfer delays.

The analytical infrastructure includes an intermediate repository known as the operational data store, which provides a short-term repository that stores only recent data that would frequently be displayed on the dashboard. The operational data store allows for the data to be extracted once from the source data system and then transformed into a format that allows for efficient display before it is loaded into the operational data store through a process termed Extract-Transform-Load (ETL).

The ETL converts source data into a more efficient data store for use within the dashboard. This is achieved by extracting the necessary data from the source data systems, transforming the data into a format that is more closely aligned to the intended display or report structure, and loading the data into the operational data store for use by the dashboard. The efficiency gained in performing the transformation once during the ETL, as opposed to doing it on-the-fly for every user interaction, improves the performance of the dashboard. Data validation and data quality checks are also performed during the ETL. The dashboard displays this transformed data within the presentation tier.

Some data does not require transformation or temporary storage in the operational data store and thus bypasses this process in the analytical infrastructure tier. In the example architecture shown in Figure 2-13, both GIS and the security video server data bypass this tier. The data in both of these systems does not require transformation prior to display in the presentation tier. In the case of GIS, the map services are designed to provide other systems with GIS data on demand and the dashboard is just another system that can connect to the GIS resource. In the case of security video, the large video files are provided to the presentation tier on demand, and software running on the client-side browser allows the video to play without transformation.

iii) Presentation Tier

The presentation tier provides the interface to the user. The results from the analytical infrastructure are made available to the presentation tier where they are combined with the GIS map services to provide geospatial context. This creates the display layers described earlier in this section, which can be manipulated by the user through an interface. On-screen interactions with the dashboard may include requests for different data presentations and user input to the alert management toolbox. This tier includes user access through devices such as work stations and mobile devices.

The dashboard may generate some data for alert management and may require some configuration information to be stored in a database that can be accessed by the dashboard as needed. This dashboard database also allows users to access data created through interactions with the dashboard, such as alert investigation records.

6.4.3 Sample interface of prototype system

The prototype for web-enabled SDSS for tick borne disease is developed on the ArcGIS Server and hosted through ArcGIS Online. The ArcGIS server is built on geospatial database and stores spatial data from the study including point data (location of tick infection cases) from 2002 to 2007, polygons (land use map, administrative boundary) and lines (river and road networks). The server allows upload of spatial data and preparation of map layout with additional information such as descriptions and photo images that can be published in the internet. The prototype system is in the form of a ArcGIS Story Map and can be accessed at the following link: https://arcg.is/1zayfD.

The prototype system is named as Web-enabled SDSS for Tick Infections in Kuantan, Pahang, has four different categories based on four types of users: government decision makers, health workers, researchers and public users. The main page of the system is as shown in Figure 6.6.

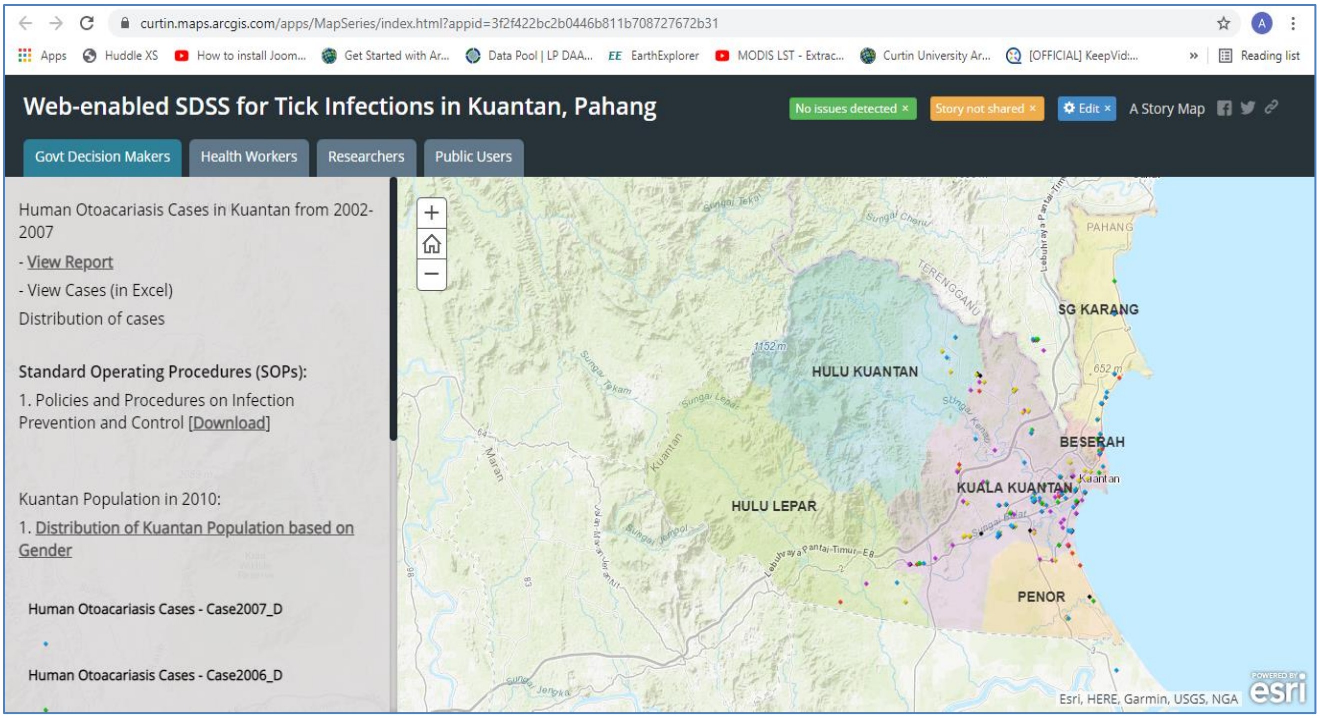


Figure 6.6 Main Interface of the Web-enabled SDSS

a) Government decision makers

The government decision makers consists of management level officers in the Ministy of Health (MoH). In order to formulate policy and decision making, they would need quick and accurate information whenever needed. For analysis purposes, useful data can be made available such as reports on human otoacariasis cases over a period of time, demographics of the area and existing policies and standard operating procedures (SOPs) on infection prevention and control. This category of user can also be able to explore the map and have a better understanding of the situation on the ground by looking at the pattern and distribution of cases in the area. By having such information, it will help them to identify risk areas, manage resources and formulate plans to control any outbreak.

b) Health workers

Health workers are users working in health facilities like government owned hospitals and clinics. This category of users are the frontliners or the people responsible to attend to human otoacariasis patients who seek treatment at their health facilities. As in normal operation procedures, the health workers keep records on patients on their own database system. However, upon detection of any infection that may cause concern or outbreak, they are able to submit an incident report on tick borne disease through an online form which will be submitted to the attention of the Ministy of Health (MoH). This online form is made accessible for collection of data from all health facilities. Data collected can then further be analysed by the Ministry of Health in order to produce report and statistics, and are also able to be accessed by the researchers.

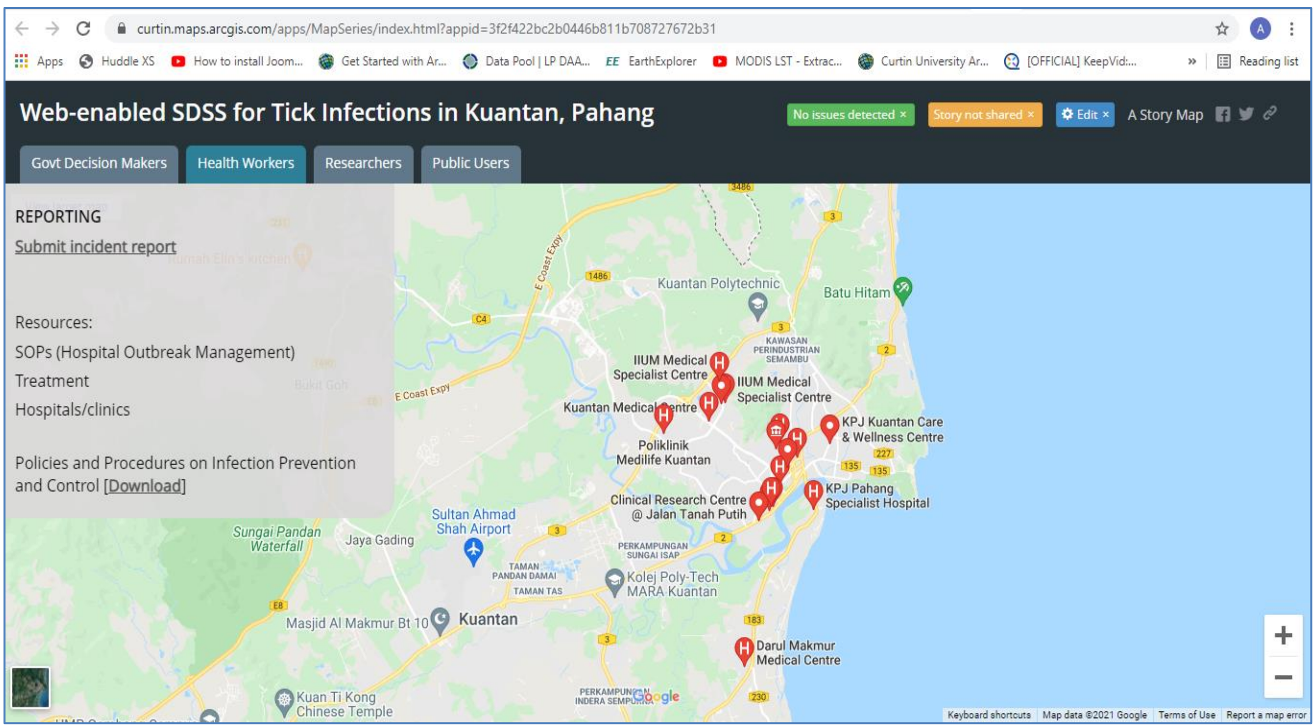


Figure 6.7 Interface for Health Workers

c) Researchers

Researchers are able to access and utilise data made available for them including human otoacariasis data, meteorology data, population data and related publications and related research. The data are downloadable from the system. At the same time they are also able to provide input to the system or contribute their product of analysis and research to the government such as on improvement of the web-enabled system, discovery of new disease information and treatment methods.

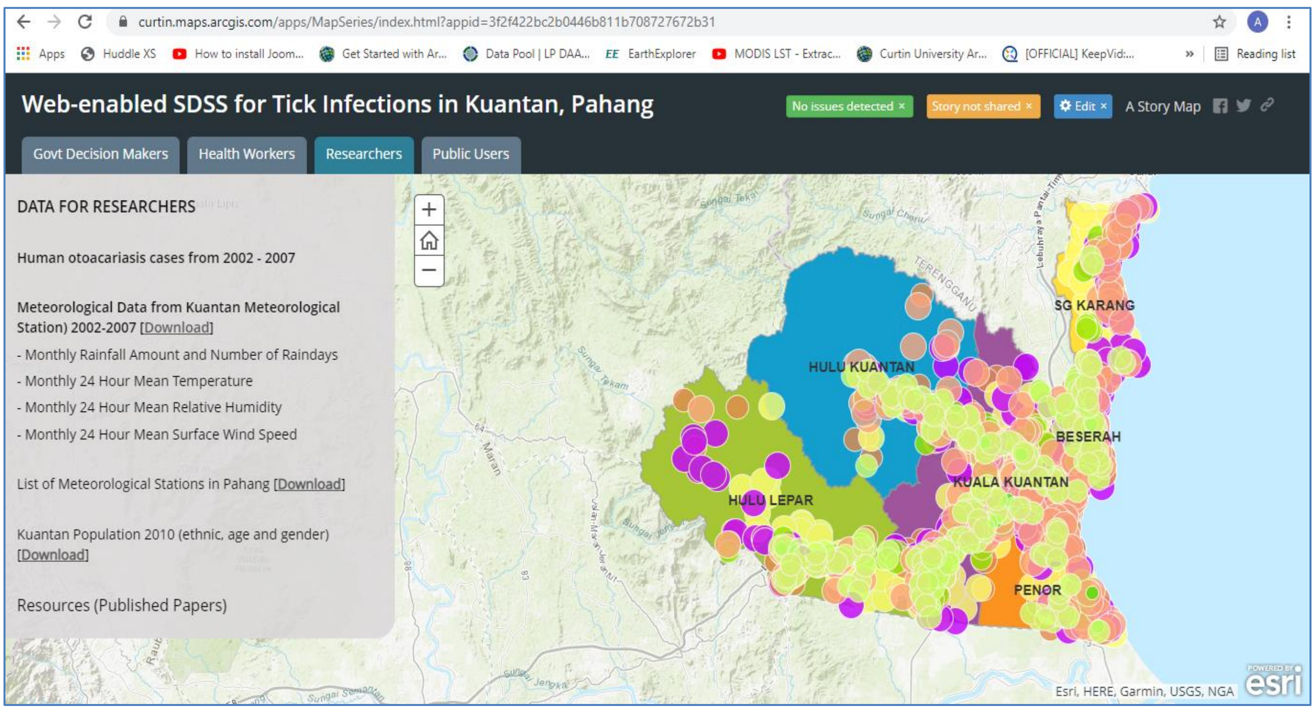


Figure 6.8 Interface for Researchers

d) Public users

The interface for the public users as in Figure 6.9 is designed to be more interactive, informative and user friendly in order to attract the attention of the public. There are general information on Kuantan as the hub of economic activity, transportation and tourism in Pahang and in the east coast of Malaysia. Information such as land use, agricultural types and distribution of human otoacariasis cases are visualised in maps. Photos and videos are also included in the system. The public may also learn more about ticks, diseases it may cause and symptoms related to tick bites. In case of any encounter with tick or tick bites, they may seek medical treatment from nearby health facilities based on the locations of hospitals and clinics visualised in the map.

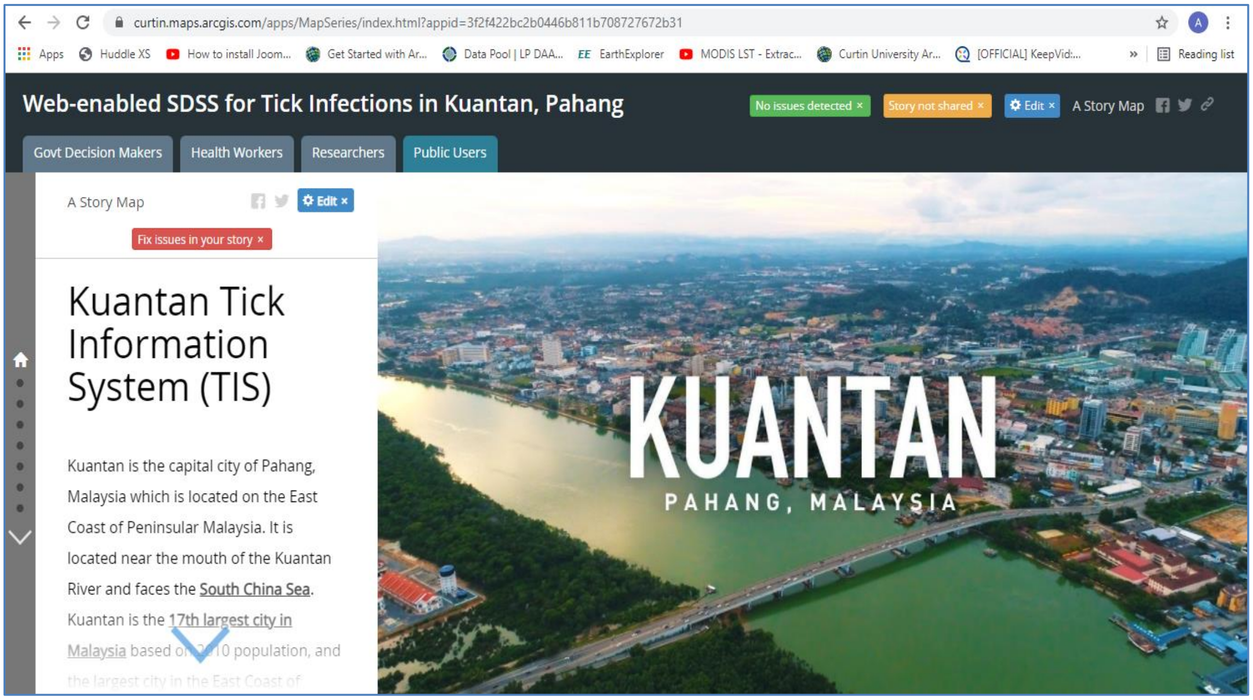


Figure 6.9 Interface for the Public

6.5 The potential of Web 2.0 Technologies for user participation

Web 2.0 technologies provide members of the health community - health professionals, health consumers, health carers and medical and health science students - with new and innovative ways to create, disseminate and share information both individually and collaboratively (Doherty, 2008). Web 2.0 is the next generation of Internet applications that depend on interaction with users: in other words, Web 2.0 aims to build collaborative frameworks that considering the user as an active entity, one that participates in the framework by providing experience, knowledge, information and feedback through applications such as healthcare blogs, wikis, social networking and podcasting, (Saini et al., 2010). Table 6.4 shows some examples of the Web 2.0 applications specifically related for medical applications.

Web 2.0 apps	Purpose	Online examples/tools
Blogs, photo blogs	Provide medical consultation, news, announcements, photos, allow comments	www.docnotes.net http://casesblog.blogspot.com
		www.wordpress.com www.flickr.com
RSS feeds and news syndication	Instantly receive medical information right after it is published	http://www.doctorslounge.com/rss http://www.rss4medics.com www.medicalnewstoday.com
		http://www.feedforall.com
Podcast and Vidcast	Provide consults, courses and information in audio and video stream format	http://conversations.acc.org/ http://www.annals.org/podcast/index.shtml http://www.clevelandclinic.org
		http://video.google.com/ http://www.archive.org/details/movies
Wiki	Collaboratively construct an archive of medical knowledge	http://askdrwiki.com/mediawiki/ http://www.radiopaedia.org
		http://www.mediawiki.org/ http://www.splitbrain.org/go/dokuwiki
Collaborative Tagging and Social bookmarking	Link to informative content, evaluate sources and organize knowledge	http://www.bibsonomy.org/ http://www.citeulike.org/
		http://www.flickr.com/ http://www.connotea.org/
Cyberspaces	Provide a virtual and interactive learning environment	http://www.secondlife.com http://opensimulator.org/

Table 6.4 Web 2.0 applications. Examples and open source solutions
Source: Table adapted from Varlamis & Apostolakis (2010, 319)

Web 2.0 has intervened in the public health sector, with its main aim to guard and enhance the health of public through the promotion and propagation of healthy lifestyles, disease prevention, detection and control of outbreaking viral infections (Shah, 2018). Using the applications, the different types of stakeholders can communicate and collaborate among each other and share information which is useful to other types of users.

For the Web-enabled SDSS, all of the Web 2.0 applications listed above can be applied to invite collaboration from the users, especially the public. Some applications that can be useful are described in Table 6.5 below.

Web 2.0 applications	**Functions**
Blogs, photo blogs	Users can share post information, photos, and links on any topics regarding tick-borne disease and invite comments from readers.
RSS feeds and news syndication	Users can share content such as research articles or news and allow other users to subscribe automatically and get feeds whenever new content is published
Podcast and Vidcast	Users such as health workers, researchers, and patients can broadcast audio or video, which can be shared with others
Wiki	Users can create content on a topic and invite other users to collaborate and edit the content
Social networking	A platform or site where people can connect and network with one another, either between patients, health workers, researchers and decision-makers

Table 6.5 Web 2.0 applications for Web-SDSS for tick borne disease

Besides promoting participation and networking among government decision-makers, health workers, researchers and the public, Web 2.0 applications can also be a platform for risk management. The government decision makers may utilise an application to request a detailed report from the health workers, assess the situation on the ground, and decide should an increase of cases pose a health concern or risk to the community. The detailed report, along with information on the locations of cases, will be useful to the researchers to visualise in the form of disease risk maps. The disease risk maps can in turn, be shared among selected users, especially health workers in hotspot areas which will allow them to be able to implement necessary control measures such as for allocation of resources of manpower, medical supplies and more. The patients or the public, may also contribute by providing the location of the infection and the symptoms that they experienced, which will be helpful to the health workers to identify the disease and for the researchers to study.

6.6 Chapter Summary

The rapid advancement in the World Wide Web (WWW) technology has enabled many applications to be widely accessible to Internet users. This includes applications for Web GIS where spatial maps can be accessed, shared and downloaded onto computers and mobile devices. Based on this technology, a prototype for web-enabled Spatial Decision Support System (SDSS) for tick borne disease has been successfully developed. The system is able to cater for four categories of users consisting of government decision makers, researchers, health workers and the general public. It can serve as a platform for communication and information sharing between each category of users. At the same time different access levels can be assigned for each user category for protection of classified data and information, which can only be accessed by authorized users. For prototype development, access level has not been set in order to show the capabilities and functions for each user category.

At the same time, Web 2.0 technologies offer web-based applications that are able to include participation among the stakeholders and the public. Using the readily available applications, administrators can choose the different platform for information sharing, dissemination and invite contributions and participations from many categories of users.

7 CONCLUSIONS AND RECOMMENDATIONS

This final chapter describes the findings of this research including the problems encountered, which have an influence on the method of analysis and eventually the final outcomes of this study. At the same time, there are a few proposed suggestions that can be undertaken to solve the respective problems. Finally, recommendations are given on the potential to improve and extend the research further in the future.

7.1 Research achievements

Based on the literature review, the species of ticks, their habitats and transmission to humans have been explored. Extensive studies have been found to be more on the biological aspects of tick and their pathogens, with the most species of ticks belonging to the genus Dermacentor (99.7%). There is no study found on the environmental factors that may lead to the increase of infection. Therefore this research has focused on the potential factors of tick infection based on the meteorological data as well as land use and vegetation data derived from remote sensing satellite images. As the meteorological data was obtained from only one station, the analysis has not successfully found any significant relationship with the infection cases.

Studies have found that tick habitat is in the vegetation area and the transmission to human and animals occur when human and animals venture into their area. Therefore this study has investigated the potential transmission of ticks to humans by identifying the surrounding enviroment of each patients' location. Based on the findings, there are three significant predictive factors of tick bite incidents have been successfully identified and incorporated in a model for prediction of incidents in the future. The presence of these significant factors can provide a clue as to when and where tick bites may occur, which may pose a risk of future disease outbreaks if it is not contained at an early stage. Three significant factors have been identified consisting of land use, distance from the river (with more cases detected closer to the rivers) and in the presence of three types of land use: swamp, urban and oil palm areas. The three identified factors have been proven to be significant predictors of tick infestation leading to tick bite incidents when incorporated in the CART modelling algorithm as presented in Chapter 5.

Consequently, the results of modelling, in the form of spatial maps can be incorporated into a Web-enabled Spatial Decision Support System (SDSS). The web-enabled SDSS will then be able to provide data and information for decision makers, which can improve the current practice and process of decision-making, which is currently directed more towards reporting and storing of patients' records. The system can also be a platform for distribution and dissemination of information to other users including the general public. Through a web-enabled SDSS, data and information can be shared in real time and the availability of spatial data may assist the decision maker to have better evaluation capability on the current situation on the ground. Four categories of users, namely decision makers, researchers, health workers and the general public have been identified for the system. Each category of users has its own specific requirement for the type of information as well as a different access level to the system.
The Web-SDSS can be improved with the integration of Web 2.0 technologies, which is able to initiate participation from users especially the general public. However in this case further study should be made to investigate the suitability of implementation of Web 2.0 technologies at this stage, as there are still issues arising in terms of limited Internet access and service coverage in the study area.

Finally there are potential improvements to this research in the future, for example higher chances of probability in identifying more factors of tick borne disease, investigating the effects of climate change, as well as the impact of deforestation and urbanisation, which may affect ticks and their natural habitat. When more time is available to conduct further investigations, researchers may learn more about ticks and at the same time new data may assist the authorities to review their current and future development plan, which will provide for systematic urban development and sustainable forest management while at the same time be able to prevent the occurrence of a disease outbreak.

7.2 Problems encountered and setback to research

During research, there are a few problems and setbacks encountered resulting in certain limitation and difficulties faced before and during data analysis. Therefore these issues have an influence on the research methodology, thus affecting the end result. The problems encountered among others, are due to:

a) Insufficient data

Currently data on the three Malaysian *Dermacentor* tick species such as *Dermacentor atrosignatus*, *Dermacentor compactus* and *Dermacentor steini* found in the study area is very limited as there is no complete documentation on the species available in the literature. Even though these tick species are discovered from human otoacariasis patients, the actual biology and lifecycle of the species are still unknown. Through personal communication, the lead researcher on ticks from the Acarology Unit, Institute of Medical Research (IMR) has informed this researcher that their documentation on ticks is yet to be completed. Therefore at present more detailed documented information is not yet available for reference, apart from general information published by local researchers. Hence, to better understand the species, reference and comparison has been made with the *Dermacentor variabilis* or also known as the American Dog tick, which has similar characteristics to the species found in Malaysia. However the purpose of the comparison is towards having better understanding of the species and cannot be applied directly in the study. Thus, this research could be improved if more data and information about the local species was available, including their lifecycle, habitats, their preferred hosts and how do they come in contact with human and the types of bacteria and virus that ticks can transmit and diseases caused by the transmission. Furthermore, the information about the tick lifecycle itself will be very helpful in predicting when tick borne disease may occur. However at this time it is difficult to identify the specific time when a tick lifecycle will complete since ticks can only moult from one stage to another after they fed, therefore it may take a few years for a lifecycle to complete. Tick's lifecycle completion also does not depend on other factors such as weather or climate as is the case with mosquitoes, which have a shorter life span. Therefore modelling methods such as climate-based prediction models which can be applied for prediction of mosquito borne disease is not suitable for predicting tick borne disease.

b) The actual events of tick bite incidence

For the purpose of this study, the locations of tick bite incidents have been spatially geo-coded according to the address of patients. However, the patients' address may not be the actual location where the incidents took place. Therefore to determine the incidents location and how it happens, it will be very helpful if the patients can notify the authorities about where they were bitten or at least provide some insights or clues on where it took place. At the same time patients can also provide information on the symptoms that they experienced, so that health workers can make a correct diagnosis and give an early treatment to prevent it from getting worse. However, some patients may only seek treatment when the pain has become unbearable to them, after a few days of being in contact or bitten by ticks. Therefore it will be difficult to identify the exact location at which the patient came into contact with ticks.

At the same time, even though there are patients' records for a period of six years from 2002 to 2007, there are still some gaps or missing data due to incomplete records especially from the year 2005. Therefore analysis of the area may not be a hundred percent accurate since only geo-coded locations can be included in the spatial analysis. Besides missing locations, there are also missing records on patients' admission date, which is the date when they seek treatment. At the same time the date of hospital treatment may not signify the date when tick bites happened as patients may have been bitten a few days before.

Considering the issues mentioned above, it is important to get more details about patients to understand the pattern of tick infestation and contact with human. Such details may help the investigation to further understand if tick bite is present and related to a certain age group, gender, or due to activities engaged by the patients that may expose them to ticks in the first place. At the same time more information from the patients will be helpful to find out if they have contracted any disease transmitted by ticks.

c) Satellite remote sensing data on study area

Since the study area is located in a region with an equatorial climate, it is almost impossible to get cloud-free data. Data coverage is also an issue since there may be no coverage for specific remote sensing satellites on the study area during the study period. Since the study area is small, higher resolution satellites such as SPOT have been considered for this research. However, upon investigation it was found that there was no coverage of SPOT satellite on the area during the period of 2002 until 2007. The next available option is the Landsat remote sensing satellite, for which data are available during the study period. However the Landsat data are not free from cloud cover, which would require more processing works to remove the cloud cover as well as shadows, with a risk of losing valuable data beneath the cloud. Thus, to overcome both problems of cloud cover and coverage, the best available option is to apply the coarser resolution MODIS remote sensing satellite data. The MODIS satellite data come with composite images of 8-day or 16-day satellite images, which are able to solve the problem of cloud cover in the area. MODIS also has its own limitations. Its optical sensor is unable to observe the earth surface during cloud cover. The coarse resolution data can also affect the quality of analysis including in this study. However, recently current higher resolution SPOT data are now available for use for future research.

d) Mixed vegetation types in land use data

At present, the land use map of the study area is not being updated frequently. The updating work is done once in every few years since there are not many changes that can be detected within the study area in a short period of time such as in between one or two years. However, there are different types of land use or vegetation detected within the radius of two kilometres used in this study. The types of vegetation can be classified using the Vegetation Index (VI), however, it is difficult to classify or identify the EVI values that correspond with a certain type of vegetation, due to a mixture of vegetation within such a small area. Therefore it is quite a challenge to identify which type of vegetation is present at the location of tick bites. Thus a different analysis technique, in this case, pseudo absence analysis is being applied to prove that the area where ticks are found (present) is completely different from the area where ticks are not found (absent).

7.3 Opportunity for improvement and further research

There are several opportunities that can be considered to improve this research in the future. These suggestions may be able to solve the problems and setbacks while taking into consideration the potential of technological advance. In general, the availability of good quality data such as tick background, remote sensing data and human population data as well as their economic activities will be able to extend this research further in the future. The suggestions are:

1. *Getting more data on tick species background and biology*

Since data for this research were limited, it has reduced the chances to investigate other potential factors of tick borne disease. Having access to more and complete data, would further improve the chances of identifying more significant factors, which can be incorporated into a model for prediction of tick borne disease. For example, more data on ticks biology will be helpful to identify ticks habitat preferences, their lifecycle and if certain conditions have an influence on their population, questing activities and survival. According to Knülle and Devine (1972), "ticks absorb water vapour from the atmosphere". Later on Needham and Teel (1991) discovered in their study that "between meals, the off-host stages ticks can survive for several months, even years without feeding". Therefore the availability of existing data combined with more recent and complete data on tick infestation and tick background will be able to help further expand this research. These data can be compared and overlayed with land use data and other auxiliary data in GIS to detect spatial patterns of tick borne disease.

2. *Good quality and real-time remote sensing satellite data*

Remote sensing satellite data have been used in a wide range of applications including in public health, and specifically in disease management. The satellite data are able to provide multi-resolution and multi-temporal data of high quality and accuracy. However there are situations when good quality data are not available especially when conducting research based on old or historical data. For future research, if real-time data are available and can be processed quickly, it will allow researchers to perform quicker analysis and modelling. Information can be extracted from the satellite data and decision makers and health authorities can make use of the current information to assess the situation when disease outbreaks occur. Therefore, they will be able to decide possible areas for quarantine or containment, as well as improve the management and allocation of resources during disease outbreak emergencies. However, even though getting real-time remote sensing satellite data is not impossible, the challenge is more towards acquisition and processing time, which has restricted its use specifically for identified emergency related applications only such as during events of national security such as forest fire and flood.

3. *Develop and test the system for implementation*

Since the actual system development and testing is not available at this stage of research, the capability of the system can only be identified if the system is developed and tested. Therefore an assessment needs to be made on a working system to check the effectiveness and to detect any weakness before suggestions can be made for improvement and system enhancement. In addition, prior to development, a feasibility study should be carried out to identify the physical location of hardware, an actual system design framework, as well as the installation and configuration of hardware and software including system development and maintenance.

4. *Investigate other possible factors based on more data such as elevation and soil type*

During data analysis it was discovered that high number of cases came from areas where land use types present are of urban, oil palm and swamp areas. These three areas are significant with more human activities and high vegetation where ticks are abundant resulting in highly potential transmission and infection to humans. A study by Morand & Lajaunie (2021) has found that there is an association between the increasing number of outbreaks of vector-borne diseases and the increase of oil palm plantations. There are also very high human activities involved in the swamp areas of Kuantan. Coastal development has impact on mangrove area as it is being cleared for tourism such as building jetties, used for constructions and alsor for commercial and residential (Saad et al., 2009). Green localities inside cities present vegetation cover suitable for tick life where large and middle-sized animals which are good hosts for adult ticks provide means for tick transport from wilderness into such localities and support the establishment of tick populations there (Uspensky, 2015). Ticks are an example of species that adapt to new conditions; hence, an increasing number of tick bites in urban and suburban areas, such as city parks or suburban forests, are reported based on studies in Europe by Grochowska et al. (2020). The combination of urbanization, climate change, and alterations in land-use patterns along with socio-economic factors (outdoor sports and leisure-time activities, gardening, an increased density of pets, and companion animals near human settlements) act in creating favorable conditions for increasing the exposure of humans to ticks, thus favoring the transmission of tick-borne pathogens in urban and peri-urban areas (Rizzoli et al., 2014).

At the same time, there may be more significant factors that have not been discovered due to data limitations. Therefore when more data are available, further study can be done to identify more factors of tick borne disease to complement the current findings. Factors that can be considered, for instance is to identify whether climate change has an influence on ticks and their lifecycle in the study area. This follows from information found in the literature which suggest that ticks are influenced and affected by climate change and water moisture as well as being able to survive underwater for a period of time. A study by Sutherst (1971) on the flooding of the Ixodid ticks *Boophilus microplus (Canestrini)* has discovered that "the survival of engorged female ticks was influenced by their age and the temperature of the water where it was found that eggs and larvae were more resistant to submersion than were engorged females and their survival was increased at low temperatures and in water with high oxygen content". That researcher also noted that "the persistence of this tick in areas prone to flooding was attributed to the survival of parasitic stages and to larvae on the pasture and heavy rains produce favourable pasture conditions for tick reproduction so that a large increase in population size may be expected to follow such heavy rain".

Climate change and the different seasons have an influence on ticks and tick-borne diseases in Europe. According to Daniel et al. (2008) who performed a study in the influence of meteorological conditions of the preceding winter on the incidences of Lyme Borreliosis (LB) and Tick-borne Encephalitis (TBE) in the Czech Republic, "the severity of winter influences TBE incidence in the following season while the occurrence of LB is correlated in similar, but not fully significant manner". Later on a study was performed by Gray et al. (2009) on the effects of climate change on ticks and tick-borne diseases in Europe has found that "climate change has indirect effects on tick-borne pathogen transmission by affecting the survival and abundance of tick maintenance hosts such as deer and pathogen-reservoir hosts such as rodents and birds and may also influence disease risk by affecting the long-term use of land (e.g., farming, tourism, etc.) while weather patterns may influence short-term human behaviour such as picnics and mushroom picking". Even though the study area does not experience dramatic changes of seasons, there are still weather conditions involved such as rainy and dry seasons, which possibly may have some effects on ticks and their habitats.

Meanwhile Fielden et al. (2011) discovered in their study that "one aspect of off-host tick survival is their ability to survive immersion in water after heavy rainfall events or temporary flooding due to the respiratory system of adult and nymphal ticks being similar to that of insects, with a tracheal system to supply oxygen throughout their body". This has served their purpose of their study to demonstrate that unfed adult *Ixodid* ticks can survive water immersion for a considerable period of time, specifically up to 15 days in the case of female *Dermacentor variabilis*. This discovery can be investigated further in the future to study the effects of floods on ticks.

5. *Exploration of different modelling techniques*

Since data are limited especially on tick biology, only certain factors can be identified, thus limiting the choice of modelling technique for this research. However when more data are made available, there are better chances to identify more possible and significant factors of tick borne disease. Therefore the choice of modelling technique will not be limited to only one method. For example when more biological data such as the background and lifecycle of tick is known, biological modelling methods can be applied. For instance, since climate has been identified as a significant factor, the climate-based suitability model was used to model the dynamics of both vector and pathogen of *Ixodes scapularis,* the main vector of Lyme disease in North America by Brownstein, Holford and Fish (2005). Olwoch, Reyers and Jaarsveld (2007) have also applied the climate suitability model in their study to explore effect of climate change on the distribution of *Rhipicephalus appendiculatus* tick in sub-Saharan Africa. Hönig et al. (2011) has used two factors, namely the vegetation cover and elevation level in a habitat suitability model to predict the tick borne disease risk in the Czech Republic. Jones, Conner and Song (2012) have tested and compared four separate modelling techniques (stepwise logistic regression, classification decision tree, gradient boosted tree and neural network) using the same dataset to determine which model type performs best in describing the occurrence of two tick-borne diseases: Lyme disease (LD) and Rocky Mountain spotted fever (RMSF), which are commonly detected in Tennessee, USA. Later on Porretta et al. (2013) have also applied methods of Species Distribution Modelling (SDM) to incorporate the effects of climate change on the *Ixodes ricinus* vector in the Europe. Meanwhile, statistical modelling techniques have also been applied in predicting tick borne disease. Hierarchical Bayesian modelling was applied by Wimberly, Baer and Yabsley (2008) in predicting the geographic distributions of two tick-borne pathogens: *Ehrlichia chaffeensis*, the causative agent of human monocytotropic ehrlichiosis and *Anaplasma phagocytophilum*, the causative agent of human granulocytotropic anaplasmosis in the United States using serology and environmental data. The examples mentioned above have shown that different modelling techniques can be applied for tick borne disease using similar or different types of factors.

6. *Input from related government agencies and other users*

The decision makers for the web-enabled SDSS consist of three levels of government health departments from the district, state and up to the national level. The proposed system design is customized towards preparing a platform to facilitate an efficient flow of data and information among the three levels of decision makers. However, in the future the decision makers might wish to include more features or make specific requests on the system based on their current needs. Therefore, input from the decision makers will be very helpful to understand their actual requirements as the priority users of the system. The users' feedback on the system as well as their expectation can be used as a reference for enhancement of system to improve it further. At the same time, access to the system can also be extended to other government agencies that may gain benefit from information sharing, such as the National Security Council and the Fire Department.

7.4 Chapter summary

This chapter summarizes the achievements of this research, which has successfully managed to identify the significant factors of tick borne disease and incorporates them into a model for prediction of disease risk in the future. However, during research there were also problems and setbacks encountered, which have influenced the methods of analysis and eventually affecting the final result of study. The major setback is related to the unavailability of good quality data including the lack of data on tick biology and background in the study area. Finally suggestions and recommendations are provided to solve the problems, as well as suggestions to improve this study further in the future.

REFERENCES

Adams, D. R., B. E. Anderson, C. T. Ammirati, and K. F. Helm. 2002. Identification And Diseases Of Common U.S. Ticks. *The Internet Journal of Dermatology* 2 (1). https://ispub.com/IJD/2/1/7515 (accessed 1 April 2015).

Adlaoui, E., C. Faraj, M. E. Bouhmi, A. E. Aboudi, S. Ouahabi, A. Tran, D. Fontenille, and R. E. Aouad. 2011. MappingMalaria Transmission Risk in Northern Morocco Using Entomological and Environmental Data. *Malaria Research and Treatment* 2011: 9. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3265283/pdf/MRT2011-391463.pdf (accessed 22 November 2012).

Ahamad, M., T. M. Ho, I. Salleh, and R. Indhudaran. 1996. Species distribution of ticks in two localities in Kelantan. (Research Note). *Tropical Biomedicine* 13: 185-88.

Ahamad, M., Z. Z, H. T.M., M. K. B., S. I., S. M.N., and S.-A. M.S. 2008. Ticks (Ixodidae) and other ectoparasites in Ulu Muda Forest Reserve, Kedah, Malaysia. *Southeast Asian Journal of Tropical Medicine and Public Health* 39 (3): 496-506. http://proquest.umi.com.dbgw.lis.curtin.edu.au/pqdweb?index=0&did=1502995271&SrchMode=1&sid=1&Fmt=6&VInst=PROD&VType=PQD&RQT=309&VName=PQD&TS=1315448862&clientId=22212 (accessed 7th September 2011).

Ahamad, M., and S. Nordin. 2008. Intra-aural ticks (Metastigmata: Ixodidae) from human otoacariasis cases in Pahang, Malaysia. *Asian Pacific Journal of Tropical Medicine 1* 1 (3): 20-24.

Ahamad, M., S. M.N., M. N. H., I. N. B., and H. T.M. 2010. Movements and home range of a common species of tree-shrew, Tupaia glis, surrounding houses of otoacariasis cases in Kuantan, Pahang, Malaysia. *Asian Pacific Journal of Tropical Medicine* 3 (6): 427-434. http://www.sciencedirect.com/science/article/pii/S1995764510601034 (accessed 7th September 2011).

Agoulon, A., L. Malandrin, F. Lepigeon, M. Vénisse, S. Bonnet, C. A. M. Becker, T. Hoch, S. Bastian, O. Plantard, and F. Beaudeau. 2012. A Vegetation Index qualifying pasture edges is related to Ixodes ricinus density and to Babesia divergens seroprevalence in dairy cattle herds. *Veterinary Parasitology* 185 (2–4): 101-109. http://www.sciencedirect.com/science/article/pii/S0304401711006819 (accessed 4 April 2017).

Agrawal, S., and R. D. Gupta. 2014. *ISPRS Technical Commission IV Symposium, Development and comparison of open source based Web GIS frameworks on Wamp and Apache tomcat web servers*. Suzhou, China. http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-4/1/2014/isprsarchives-XL-4-1-2014.pdf (accessed 23 July 2014).

Airbus Defence and Space. n.d. *SPOT 6 | SPOT 7 Technical Sheet*. http://www.geo-airbusds.com/sg/3215-spot-6-7 (accessed 30 March 2015).

Alonso-Carné, J., García-Martín, A., & Estrada-Peña, A. (2013). Systematic errors in temperature estimates from MODIS data covering the western palearctic and their impact on a parasite development model. *Geospat Health*, *8*(1), 1-12. https://doi.org/10.4081/gh.2013.49

Alter, S. 2002. *Information systems : foundation of e-business / Steven Alter*. 4th ed: Prentice Hall.

Amin, S. U., K. Agarwal, and R. Beg. 2013. Data Mining in Clinical Decision Support Systems for Diagnosis, Prediction and Treatment of Heart Disease. *International Journal of Advanced Research in Computer Engineering & Technology* 2 (1): 6. http://ijarcet.org/wp-content/uploads/IJARCET-VOL-2-ISSUE-1-218-223.pdf (accessed 26 July 2014).

Anderson, J. F., and L. A. Magnarelli. 2008. Biology of Ticks. *Infectious Disease Clinics of North America* 22 (2): 195-215. http://www.sciencedirect.com/science/article/pii/S0891552007001237 (accessed 12 June 2014).

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer. 1976. *A Land use and land cover classification system for use with remote sensor data*. Edited by J. R. Anderson: Washington: U.S. Govt. Print. Off.

Andrianasolo, H., D. Fages, J.-P. Gonzalez, P. Barbazan, and K. Nakhapakorn. 1999. A Methodology in Detailed Environment Mapping for Viral Disease Survey. In *Asian Conference on Remote Sensing (ACRS)*. Hong Kong. GISdevelopment.net

Angelaccio, M., A. Krek, and A. D'Ambrogio. 2009. A Model-driven Approach for Designing Adaptive Web GIS Interfaces. In *Information Fusion and Geographic Information Systems*, ed. M. S. Vasily V. Popovich, Christophe Claramunt, Kyrill V. Korolenko, 137-148. Springer Berlin Heidelberg.

Anyamba, A., J. L. Small, S. C. Britch, C. J. Tucker, E. W. Pak, C. A. Reynolds, J. Crutchfield, and K. J. Linthicum. 2014. Recent Weather Extremes and Impacts on Agricultural Production and Vector-Borne Disease Outbreak Patterns: e92538. *PLoS ONE* 9 (3): Biological Sciences; Computer and Information Systems Abstracts; Materials Research Database; ProQuest SciTech Collection. http://search.proquest.com.dbgw.lis.curtin.edu.au/docview/1509205458/fulltextPDF/47C97FFDC142499EPQ/2?accountid=10382 (accessed 11 November 2014).

Appannanavar, S. B., and B. Mishra. 2011. An Update on Crimean Congo Hemorrhagic Fever. *Journal of Global Infectious Diseases* 3 (3): 285-292. Pmc. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3162818/ (accessed 26 March 2015).

Appiah, S. K., U. Mueller, and J. Cross. 2011. Spatio-temporal modelling of malaria incidence for evaluation of public health policy interventions in Ghana, West Africa. In *19th International Congress on Modelling and Simulation (MODSIM)*. Perth, Australia.

Arnott, D., and G. Pervan. 2008. Eight key issues for the decision support systems discipline.(Report). *Decision support systems* 44 (3): 657. http://www.sciencedirect.com.dbgw.lis.curtin.edu.au/science/article/pii/S0167923607001698 (accessed 16th June 2011)

Aviña, A., C. Tiwari, P. Williamson, J. Oppong, and S. Atkinson. 2011. A Spatially Explicit Environmental Health Surveillance Framework for Tick-Borne Diseases. In *Geospatial Analysis of Environmental Health*, ed. J. A. Maantay and S. McLafferty: Springer Netherlands.

Azad, A. F., and C. B. Beard. 1998. Rickettsial pathogens and their arthropod vectors. *Emerging infectious diseases* 4 (2): 8. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2640117/ (accessed 23 May 2012).

Bacon, R. M., K. J. Kugeler, and P. S. Mead. 2008. Surveillance for Lyme Disease --- United States, 1992--2006. *Morbidity and Mortality Weekly Report (MMWR)*, http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5710a1.htm.

Bakr, M., et al. (2015). "Cooperative Spatial Decision Support System for Controlling Animal Diseases Outbreaks in Egypt." International Journal of Advanced Networking and Applications **6**(6): 2533-2533.

Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3 (2): 327-338. http://dx.doi.org/10.1111/j.2041-210X.2011.00172.x (accessed 23 March 2015).

Barrios, J. M., W. W. Verstraeten, P. Maes, J.-M. Aerts, J. Farifteh, and P. Coppin. 2012. Using the Gravity Model to Estimate the Spatial Spread of Vector-Borne Diseases. *International Journal of Environmental Research and Public Health* 9 (12): 4346-64. ProQuest Central;

ProQuest Public Health; ProQuest SciTech Collection. http://search.proquest.com.dbgw.lis.curtin.edu.au/docview/1316623058/fulltextPDF?accountid=10382# (accessed 11 November 2014).

Bates, P. 2012. *External Parasites of Small Ruminants : A Practical Guide to their Prevention and Control*. Wallingford: CABI. http://CURTIN.eblib.com.au/patron/FullRecord.aspx?p=862275 (accessed 31 October 2014).

Benelli, E., M. Innorta, R. Satolli, and S. Salmaso. 2013. *The A/H1N1 Pandemic in Italy: Surveillance, Communication and Formation through the Internet* Edited by E. Mordini and M. Green, *Internet-Based Intelligence in Public Health Emergencies*. Amsterdam: Amsterdam : IOS Press.

Bergquist, R. 2011. New tools for epidemiology: a space odyssey. *Memórias do Instituto Oswaldo Cruz* 106: 892-900. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762011000700016&nrm=iso (accessed 25 November 2012).

Berrada, Z. L., and S. R. Telford III. 2009. Burden of Tick-borne Infections on American Companion Animals. *Topics in Companion Animal Medicine* 24 (4): 175-181. http://www.sciencedirect.com/science/article/pii/S1938973609000737 (accessed 23 May 2012).

Blatt, A. J. 2013. Geospatial Applications in Disease Surveillance: Solutions for the Future. *International Journal of Applied Geospatial Research (IJAGR)* 4 (2): 1-8. http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/jagr.2013040101 (accessed 16 June 2014).

Bonazountas, M., D. Kallidromitou, P. Kassomenos, and N. Passas. 2007. A decision support system for managing forest fire casualties. *Journal of Environmental Management* 84 (4): 412-418. http://www.sciencedirect.com/science/article/pii/S0301479706001769 (accessed 26 July 2014).

Brewer, M. J., T. N. Mather, and J. R. Mather. 2002. *Using climate predictions in tick-borne disease risk modeling*. http://vintage.joss.ucar.edu/joss_psg/meetings/archived/cpa-wkshp/brewer.pdf (accessed 26 July 2014).

Brooker, S. 2007. Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 101 (1): 1-8. http://www.sciencedirect.com/science/article/pii/S0035920306002574 (accessed 27 August 2011).

Brown, H., K. F. Yates, G. Dietrich, K. MacMillan, C. B. Graham, S. M. Reese, W. S. Helterbrand, W. L. Nicholson, K. Blount, P. S. Mead, S. L. Patrick, and R. J. Eisen. 2011. An acarologic survey and Amblyomma americanum distribution map with implications for tularemia risk in Missouri. *The American journal of tropical medicine and hygiene* 84 (3): 411. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3042817/ (accessed 12th June 2012).

Brownstein, J. S., T. R. Holford, and D. Fish. 2003. A Climate-Based Model Predicts the Spatial Distribution of the Lyme Disease Vector *Ixodes scapularis* in the United States. *Environ Health Perspect* 111 (9). http://dx.doi.org/10.1289%2Fehp.6052 (accessed 13 April 2012).

Brownstein, J. S., T. R. Holford, and D. Fish. 2005. Effect of Climate Change on Lyme Disease Risk in North America. *Ecohealth* 2 (1): 38-46. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2582486/ (accessed 10 April 2014).

Burg, J. G. 2001. Seasonal activity and spatial distribution of host-seeking adults of the tick Dermacentor variabilis. *Medical and Veterinary Entomology* 15 (4): 413-421. http://dx.doi.org/10.1046/j.0269-283x.2001.00329.x (accessed 23 May 2012).

Burkhardt, P. 2009. Social Software Trends in Business: Introduction. In *Social software and Web 2.0 technology trends*, edited by P. C. Deans.

Camicas, J. L., J. P. Hervy, F. Adam, and P. C. Morel. 1998. Les Tiques du Monde. Nomenclature, Stades Decrits, Hoˆtes, Repartition. The Ticks of the World. Nomenclature, Described Stages, Hosts, Distribution (Acarida, Ixodida). http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers11-05/010014377.pdf (accessed 15 July 2015).

Campbell, J. B., and R. H. Wynne. 2011. *Introduction to Remote Sensing, Fifth Edition*. New York: Guilford Publications. http://CURTIN.eblib.com.au/patron/FullRecord.aspx?p=843851 (accessed 17 June 2014).

Carisse, O., D.-M. Tremblay, T. Jobin, and A. S. Walker. 2010. Disease Decision Support Systems: Their Impact on Disease Management and Durability of Fungicide Effectiveness edited by O. Carisse. InTech.

Carroll, J. F., E. Russek-Cohen, J. D. Nichols, and J. E. Hines. 1991. Population dynamics of American dog ticks (Acari: Ixodidae) along park trails. *Population dynamics of American dog ticks (Acari: Ixodidae) along park trails* 20 (3): 922-929. http://ee.oxfordjournals.org.dbgw.lis.curtin.edu.au/content/20/3/922 (accessed 13 July 2015).

CDC. 2012. Dengue Entomology & Ecology. http://www.cdc.gov/dengue/entomologyEcology/.

CDC. 2017. Life cycle of Hard Ticks That Spread Diseases. https://www.cdc.gov/ticks/life_cycle_and_hosts.html (accessed 10 October 2017)

CDC, C. f. D. C. a. P. 2014. *Tick-borne Encephalitis (TBE)*. http://www.cdc.gov/vhf/tbe/ (accessed 30 March 2015).

Chan, W.-H., and P. E. Kaufman. 2008. *Common name: American dog tick Scientific name: Dermacentor variabilis (Say) (Arachnida: Ixodida: Ixodidae)*. http://entnemdept.ufl.edu/creatures/urban/medical/american_dog_tick.htm (accessed 29 October 2014).

Chandra, A. M. and S. K. Ghosh (2006). Remote Sensing and Geographical Information System, Alpha Science.

Charlesworth, S. 2008. Lifecycle of a three-host hard tick based on Dermacentor variabilis (American dog tick) Purdue. Purdue University.

Chefaoui, R. M., and J. M. Lobo. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling* 210 (4): 478-486. http://www.sciencedirect.com.dbgw.lis.curtin.edu.au/science/article/pii/S030438000700419X (accessed 23 March 2015).

Civitello, D. J., S. L. Flory, and K. Clay. 2008. Exotic grass invasion reduces survival of Amblyomma americanum and Dermacentor variabilis ticks (Acari: Ixodidae). *J Med Entomol* 45 (5): 867-72. http://www.florylab.com/wp-content/uploads/2012/05/Civitello-et-al-2008-JME.pdf (accessed 12th June 2012).

CNES. 2009. *SPOT Looking down on earth*. http://www.cnes.fr/web/CNES-en/1415-spot.php (accessed 16 July 2014).

CNES. 2010. *Resolutions and spectral modes*. http://www2.geo-airbusds.com/files/pmedia/public/r451_9_resolutionspectralmodes_uk_sept2010.pdf (accessed 30 March 2015).

Constantinides, E., and S. J. Fountain. 2008. Web 2.0: Conceptual foundations and marketing issues. *J Direct Data Digit Mark Pract* 9 (3): 231-244. http://dx.doi.org/10.1057/palgrave.dddmp.4350098 (accessed 5 June 2015).

Contino, G., and G. Virgilio. 2002. *5th AGILE Conference on GIScience, A Spatial Decision Support System for Situational Crime Prevention in Urban Environment*. Palma, Balearic Islands http://plone.itc.nl/agile_old/Conference/mallorca2002/proceedings/dia26/Session_8/s8_Virgilio.pdf (accessed 4th June 2011)

Crauwels, A. P., R. de Koning, M. Nielen, A. Elbers, A. Dijkhuizen, and M. Tielen. 2001. A Concept for a Decision Support System Based on Practical Experiences from a National Disease Emergency : The Dutch Experience. In *Acta Veterinaria Scandinavica Proceedings of the 13th Internordic Symposium of the Nordic Committee for Veterinary Scientific Cooperation (NKVet)*.

Dambach, P., A. Sié, J.-P. Lacaux, C. Vignolles, V. MAchault, and R. Sauerborn. 2009. *Using high spatial resolution remote sensing for risk mapping of malaria occurrence in the Nouna district Burkina Faso*.

Daniel, M., B. Křížˇ, J. Valter, I. Kott, and V. Danielová. 2008. The influence of meteorological conditions of the preceding winter on the incidences of tick-borne encephalitis and Lyme borreliosis in the Czech Republic. *International Journal of Medical Microbiology* 298, Supplement 1 (0): 60-67. http://www.sciencedirect.com/science/article/pii/S1438422108000659 (accessed 9 April 2012).

Daniel, M., J. Kolar, and P. Zeman. 2004. GIS tools for tick and tick-borne disease occurrence. *Parasitology* 129 (SupplementS1): S329-S352. http://dx.doi.org/10.1017/S0031182004006080 (accessed 1st November 2011).

Darulmakmur, T. P. 2013. *Kuantan River Cruise*. http://www.pahangtourism.org.my/index.php/destinations/islands-beaches/kuantan/kuantan-river-cruise (accessed 7 July 2014).

Das, A., S. R. Lele, G. E. Glass, T. Shields, and J. Patz. 2002. Modelling a discrete spatial response using generalized linear mixed models: application to Lyme disease vectors. *International Journal of Geographical Information Science* 16 (2): 151-166. http://dx.doi.org/10.1080/13658810110099134 (accessed 8 November 2014).

Debanne, S. M., and D. Y. Rowland. 2002. Logistic regression. *Gastrointestinal Endoscopy* 55 (1): 142-143. http://www.sciencedirect.com/science/article/pii/S001651070268647X (accessed 5 April 2017).

de Castro, M. C. 2007. Scale of questions & scale of answers: the role of remote sensing for public health policy. In *Annual Meeting of the Population Association of America*. New York, USA.

de la Fuente, J., Antunes, S., Bonnet, S., Cabezas-Cruz, A., Domingos, A. G., Estrada-Peña, A., Johnson, N., Kocan, K. M., Mansfield, K. L., Nijhof, A. M., Papa, A., Rudenko, N., Villar, M., Alberdi, P., Torina, A., Ayllón, N., Vancova, M., Golovchenko, M., Grubhoffer, L., Caracappa, S., Fooks, A. R., Gortazar, C., & Rego, R. O. M. (2017). Tick-Pathogen Interactions and Vector Competence: Identification of Molecular Drivers for Tick-Borne

Diseases. *Frontiers in cellular and infection microbiology*, *7*, 114-114. https://doi.org/10.3389/fcimb.2017.00114

Department of Statistics, M. 2011a. Population Distribution by Local Authority Areas and Mukims, 2010, edited by M. Department of Statistics. Kuala Lumpur.

Department of Statistics, M. 2011b. Population Distribution and Basic Demographic Characteristics, edited by M. Department of Statistics. Kuala Lumpur.

Disease Control Division, M. o. H. M. 2003. Infectious Diseases Outbreak Rapid Response Manual, edited by P. D. V. L. Dr. Shafie Ooyub, Dr. Lee Cheow Pheng, Dr. Cheng Hoei Hon, Dr. Zainudin Abdul Wahab, Dr. K. Devan, Dr. Amal Nasir Mustafa, Dr. Fadzilah Kamaludin, Datin Dr. Salbiah Hj Nawi, Dr. Leela Anthony, Dr. Rohani Jahis.

Doherty, I. (2008). Web 2.0: A Movement Within The Health Community. *Health Care & Informatics Review Online, 12*.

Druzdzel, M. J., and R. R. Flynn. 2002. Decision Support Systems. In *Encyclopedia of Library and Information Science,*, ed. A. Kent, 15. New York: Marcel Dekker Inc.

Durden,L.A., 1986. Ectoparasites and other arthropod associates of tropical rain forest mammals in Sulawesi Utara, Indonesia. Nat. Geogr. Res., 2: 320.

Durden, L. A., and C. H. S. Watts. 1989. A Collection of Ticks (Ixodidae) From Sulawesi Utara, Indonesia. *BIOTROPIA : the Southeast Asian Journal of Tropical Biology* 2: 32-37. http://journal.biotrop.org/index.php/biotropia/article/view/94/70 (accessed 20 April 2014).

Durden, L., S. Merker, and L. Beati. 2008. The tick fauna of Sulawesi, Indonesia (Acari: Ixodoidea: Argasidae and Ixodidae). *Experimental and applied acarology* 45 (1): 85-110. http://search.proquest.com.dbgw.lis.curtin.edu.au/docview/225532584?OpenUrlRefId=info:xri/sid:primo&accountid=10382 (accessed 15 September 2013).

ECERDC, E. C. E. R. D. C. 2010. Kuantan Port City Asia Pacific's emerging trade and industrial hub, edited by E. C. E. R. D. C. ECERDC. Kuala Lumpur.

Edlow, J. A. 2008. Tick-Borne Diseases Introduction to Tick-Borne Diseases. http://emedicine.medscape.com/article/786652-overview.

Ernieenor, F. C. L., Ernna, G., & Mariana, A. (2017). Phenotypic and genotypic identification of hard ticks of the genus Haemaphysalis (Acari: Ixodidae) in Peninsular Malaysia. *Experimental & applied acarology*, *71*(4), 387-400. https://doi.org/10.1007/s10493-017-0120-3

Estrada-Peña, A., E. L. Tarragona, U. Vesco, D. d. Meneghi, M. Mastropaolo, A. J. Mangold, A. A. Guglielmone, and S. Nava. 2014. Divergent environmental preferences and areas of sympatry of tick species in the Amblyomma cajennense complex (Ixodidae). *International Journal for Parasitology* 44 (14): 1081-1089. http://www.sciencedirect.com/science/article/pii/S002075191400229X (accessed 4 April 2017).

Estrada-Peña, A., and J. M. Venzal. 2007. A GIS framework for the assessment of tick impact on human health in a changing climate. *Geospatial Health* 1 (2): 12. http://www.geospatialhealth.unina.it/articles/V1I2/estrada.pdf (accessed 9 November 2014).

Estrada-Peña, A., Z. Vatansever, A. Gargili, and T. Buzgan. 2007. An early warning system for Crimean-Congo haemorrhagic fever seasonality in Turkey based on remote sensing technology. *Geospatial Health* 2 (1): 127-35. http://www.geospatialhealth.unina.it/articles/v2i1/gh_v2i1_13_estrada_pena.pdf (accessed 25 November 2012).

Fan, J., S. Upadhye, and A. Worster. 2006. Understanding receiver operating characteristic (ROC) curves. *CJEM* 8 (1): 19. http://www.cjem-online.ca/sites/cjem-online.ca/files/pg19.pdf (accessed 13 February 2015).

Farooq, S. *Spatial Databases Concept, Design and Management*. http://www.geol-amu.org/notes/m14b-4-4.htm

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27 (8): 861-874. https://ccrma.stanford.edu/workshops/mir2009/references/ROCintro.pdf (accessed 13 February 2015).

Fielden, L. J., L. M. Knolhoff, S. M. Villarreal, and P. Ryan. 2011. Underwater survival in the dog tick Dermacentor variabilis (Acari:Ixodidae). *Journal of Insect Physiology* 57 (1): 21-26. http://www.sciencedirect.com/science/article/pii/S0022191010002544 (accessed 14th May 2012).

Florea, M. C., N. Duclos-Hindie, E. Bosse, and P. Valin. 2009. A Web–Service Approach for Multi–Sensor Data Fusion and Geographic Information Systems. In *Information Fusion and Geographic Information Systems*, ed. M. S. Vasily V. Popovich, Christophe Claramunt, Kyrill V. Korolenko, 209-223. Springer Berlin Heidelberg.

Franklin, J. 2010. *Mapping Species Distributions : Spatial Inference and Prediction*. Cambridge: Cambridge University Press. http://CURTIN.eblib.com.au/patron/FullRecord.aspx?p=542846 (accessed 27 June 2013).

Fu, P. 2015. *Getting to know web GIS / Pinde Fu*: Redlands, California : Esri Press.

Fu, P., and J. Sun. 2010. *Web GIS Principles And Applications* 1st ed ed. Redlands, California: ESRI Press.

Gaff, H. D., and L. J. Gross. 2007. Modeling Tick-Borne Disease: A Metapopulation Model.(Author abstract). *Bulletin of mathematical biology* 69 (1): 265. http://search.proquest.com.dbgw.lis.curtin.edu.au/docview/209598976/fulltextPDF?accountid=10382 (accessed 30 July 2012).

Gaff, H., and E. Schaefer. 2010. Metapopulation Models in Tick-Borne Disease Transmission Modelling. In *Modelling Parasite Transmission and Control*, ed. E. Michael and R. Spear, 51-65. Springer New York.

Gamon, J. A., C. B. Field, M. L. Goulden, K. L. Griffin, A. E. Hartley, G. Joel, J. Penuelas, and R. Valentini. 1995. Relationships between NDVL canopy structure, and photosynthesis in three California vegetation types. *Relationships between NDVL canopy structure, and photosynthesis in three California vegetation types* 5 (1): 28-41. http://www.jstor.org.dbgw.lis.curtin.edu.au/stable/1942049?origin=crossref&seq=1#page_scan_tab_contents (accessed 13 February 2015).

Gao, B.-C. 1996. NDWI - A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space. *Remote sensing of environment* 58: 10. ftp://ftp.eri.ucsb.edu/pub/org/swim/WTLDS/Encylopedia_Wetlands/papers/Gao_1996_NDWI.pdf (accessed 29 September 2014).

Gao, S., D. Mioc, F. Anton, X. Yi, and D. J. Coleman. 2008a. Online GIS Service for mapping and sharing disease information. *International Journal of Health Geographics.*: 12. http://www.ij-healthgeographics.com/content/7/1/8 (accessed 4th June 2011).

Gao, S., D. Mioc, X. Yi, F. Anton, and E. Oldfield. 2008b. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *Geospatial Services*

For Decision Support On Public Health. Beijing. http://www.isprs.org/proceedings/XXXVII/congress/8_pdf/14_ThS-20/02.pdf (accessed 4th June 2011).

Gao, Sheng. 2010. Advanced Health Information Sharing With Web-based GIS. PhD Dissertation, Department of Geodesy and Geomatics Engineering, University of New Brunswick Fredericton, New Brunswick,Canada.

Garg, R. (2016). Methodology for research I. *Indian J Anaesth*, *60*(9), 640-645. https://doi.org/10.4103/0019-5049.190619

Gauci, M., R. K. S. Loh, B. F. Stone, and Y. H. Thong. 1989. Allergic reactions to the Australian paralysis tick, Ixodes holocyclus: diagnostic evaluation by skin test and radioimmunoassay. *Clinical & Experimental Allergy* 19 (3): 279-283. http://dx.doi.org/10.1111/j.1365-2222.1989.tb02384.x (accessed 28 October 2014).

Ghane Kisomi, M., Wong, L. P., Tay, S. T., Bulgiba, A., Zandi, K., Kho, K. L., Koh, F. X., Ong, B. L., Jaafar, T., & Hassan Nizam, Q. N. (2016). Factors Associated with Tick Bite Preventive Practices among Farmworkers in Malaysia. *PLoS One*, *11*(6), e0157987. https://doi.org/10.1371/journal.pone.0157987

Ghani, N. A. A. A., N. Othman, and M. K. H. Baharudin. 2013. Study on Characteristics of Sediment and Sedimentation Rate at Sungai Lembing, Kuantan, Pahang. *Procedia Engineering* 53: 12. http://www.sciencedirect.com/science/article/pii/S187770581300132X (accessed 16 February 2015).

GHS, G. H. S. 2007. The Health Sector in Ghana Facts and Figures 2007, edited by G. H. Service.

Githeko, A. K., and W. Ndegwa. 2001. Predicting malaria epidemics in the Kenya highlands using climate data: a tool for decision makers. *Global Change and Human Health* 2 (1): 9.

http://download.springer.com.dbgw.lis.curtin.edu.au/static/pdf/102/art%253A10.1023%252FA%253A1011943131643.pdf?auth66=1406559124_9847658eb52d39213189599d9ec65107&ext=.pdf (accessed 26 July 2014).

Goethert, H. K., and S. R. Telford. 2009. Nonrandom Distribution of Vector Ticks (*Dermacentor variabilis*) Infected by *Francisella tularensis*. *PLoS Pathog* 5 (2): e1000319. http://dx.doi.org/10.1371%2Fjournal.ppat.1000319 (accessed 12th June 2012).

Grady, C. A., D. R. McDonald, C. F. Poppen, and B. S. Pritt. 2011. Dermacentor tick attached to tympanic membrane. *The Lancet* 378 (9788): 347. http://www.sciencedirect.com/science/article/pii/S0140673611609811 (accessed 23 May 2012).

Gray, J. S., H. Dautel, A. Estrada-Pen, O. Kahl, and E. Lindgren. 2009. Effects of Climate Change on Ticks and Tick-Borne Diseases in Europe. *Interdisciplinary Perspectives on Infectious Diseases* 2009. http://dx.doi.org/10.1155/2009/593232 (accessed 9 April 2012).

Green, D. G., and T. J. Bossomaier. 2002. *Online GIS and Spatial Metadata. London*: Taylor & Francis.

Grochowska, A., Milewski, R., Pancewicz, S., Dunaj, J., Czupryna, P., Milewska, A. J., Róg-Makal, M., Grygorczuk, S., & Moniuszko-Malinowska, A. (2020). Comparison of tick-borne pathogen prevalence in Ixodes ricinus ticks collected in urban areas of Europe. *Scientific Reports*, *10*(1), 6975. https://doi.org/10.1038/s41598-020-63883-y

Guernier, V., M. E. Hochberg, and J.-F. Guégan. 2004. Ecology Drives the Worldwide Distribution of Human Diseases. *PLoS Biol* 2 (6): e141. http://dx.doi.org/10.1371%2Fjournal.pbio.0020141 (accessed 9 November 2014).

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135 (2): 147-186.2 April 2015).

Guoyu, Q., O. Yuyichi, C. Jin, Y. Wei, and M. Bunkei. 2007. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-density Cypress Forest. *Sensors* 7 (11): 2636-2651. http://www.mdpi.com/1424-8220/7/11/2636 (accessed 11 November 2015).

Hall-Baker, P. A., S. L. Groseclose, R. A. Jajosky, D. A. Adams, P. Sharp, W. J. Anderson, J. P. Abellera, A. E. Aranas, M. Mayes, M. S. Wodajo, D. H. Onweh, M. Park, and J. Ward. 2009. Summary of Notifiable Diseases --- United States, 2009. *Morbidity and Mortality Weekly Report (MMWR)*, http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5853a1.htm?s_cid=mm5853a1_w.

Hamzah, H. H. 2013. The Implementation of National Physical Plan 2: Kuantan, edited by K. M. Council. Kuantan.

Hay, S. I. 2000. An Overview of Remote Sensing and Geodesy for Epidemiology and Public Health Application. *Adv. Parasitol* 47: 35. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3164799/ (accessed 17 June 2014).

Hay, S. I., M. J. Packer, and D. J. Rogers. 1997. Review article The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *International journal of remote sensing* 18 (14): 2899-2930. http://www.tandfonline.com.dbgw.lis.curtin.edu.au/doi/pdf/10.1080/014311697217125 (accessed 22 November 2012).

Heinz, F. X., and K. Stiasny. 2012. Flaviviruses and flavivirus vaccines. *Vaccine* 30 (29): 4301-4306.

http://www.sciencedirect.com.dbgw.lis.curtin.edu.au/science/article/pii/S0264410X11015568 (accessed 16 February 2015).

Herrmann, J. W., A. Kothari, and S. Shaikh. 2007. *Industrial Engineering Research Conference, Designing a Decision-Making System for Exceptional Events: the Case of Medical Disaster Mutual Aid*. Nashville, TN. http://www.isr.umd.edu/Labs/CIM/projects/mocep/Hospital-decision-making.pdf (accessed 24th May 2011).

Heywood, D. I., S. Cornelius, and S. Carver. 2006. *An introduction to geographical information systems*. Edited by r. ed.: Harlow, England ; New York : Pearson Prentice Hall.

Hodge, J. G. 2005. An Enhanced Approach to Distinguishing Public Health Practice and Human Subjects Research. *The Journal of Law, Medicine & Ethics* 33 (1): 125-141. http://dx.doi.org/10.1111/j.1748-720X.2005.tb00215.x (accessed 19th July 2011).

Holbrook, M. R. 2012. Kyasanur forest disease. *Antiviral Research* 96 (3): 353-362. http://www.sciencedirect.com/science/article/pii/S0166354212002355 (accessed 28 October 2014).

Hönig, V., P. Švec, O. Masař, and L. Grubhoffer. 2011. *GIS Ostrava 2011, Eight International Symposium, Tick-borne diseases risk model for South Bohemia (Czech Republic)*. Ostrava, Czech Republic: VŠB - TU Ostrava. gis.vsb.cz/GIS_Ostrava/GIS_Ova_2011/sbornik/papers/Svec.pdf (accessed 16 September 2013).

Hoogstraal, H. 1966. Ticks in Relation to Human Diseases Caused by Viruses. *Annu. Rev. Entomol.* 11: 261-308.

http://www.annualreviews.org.dbgw.lis.curtin.edu.au/doi/abs/10.1146/annurev.en.11.010166.001401 (accessed 13 January 2013).

Hoogstraal, H., H. Trapido, and G. M. Kohls. 1965. Southeast Asian Haemaphysalis Ticks (Ixodoidea, Ixodidae). H. (Kaiseriana) papuana nadchatrami ssp. n. and Redescription of H. (K.) semermis Neumann. *The Journal of Parasitology* 51 (3): 433-451. http://www.jstor.org/stable/3275971 (accessed 30 October 2014).

Hoogstraal, H., and H. Y. Wassef. 1985. Dermacentor (Indocentor) atrosignatus (Acari: Ixodoidea: Ixodidae): hosts and distribution in the Malay Peninsula, Indonesia, Borneo, and southern Philippines. *Journal of Medical Entomology* 22 (6): 644. http://jme.oxfordjournals.org.dbgw.lis.curtin.edu.au/content/22/6/644.full-text.pdf (accessed 7 September 2014).

Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant. 2013. *Applied Logistic Regression*. Edited by S. Lemeshow and R. X. Sturdivant. 3rd ed.. ed, *Wiley Series in Probability and Statistics*. Hoboken: Hoboken : Wiley.

Hu, W., N. Nicholls, M. Lindsay, P. Dale, A. J. Mcmichael, J. S. Mackenzie, and S. Tong. 2004. Development of a predictive model for ross river virus disease in Brisbane, Australia. *The American journal of tropical medicine and hygiene* 71 (2): 9. http://www.ajtmh.org/content/71/2/129.full.pdf+html (accessed 26 July 2014).

Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83 (1): 195-213. https://cimss.ssec.wisc.edu/dbs/China2011/Day2/Lectures/MODIS_MOD13_NDVI_referenc.pdf (accessed 11 November 2015).

Hugh-Jones, M. 1991. The remote recognition of tick habitats. *The remote recognition of tick habitats* 8 (4): 309-315. http://scentsoc.org/Volumes/JAE/v8/4/00084309.pdf (accessed 15 February 2014).

Hugh-Jones, M. 1991. Satellite imaging as a technique for obtaining disease-related data. *Rev Sci Tech* 10 (1): 8. http://www.oie.int/doc/ged/D8227.PDF (accessed 26 November 2012).

Hugh-Jones, M. E., N. Bnrre, G. Nelson, C. Wheyness, J. Warner, G. Garris, and W. Hubbert. 1988. Remote recognition of Amblyomma variegatum habitats in Guadeloupe using LANDSAT-TM imagery. Acta Vet. Scand. Suppl. 84: 259-261.

Huiji Gao, G., R. Barbier, and R. Goolsby. 2011. Harnessing the Crowdsourcing Power of Social Media for Disaster Relief. *Intelligent Systems, IEEE* 26 (3): 10-14. http://star-tides.net/sites/default/files/documents/files/Harnessing-the-Crowdsourcins-Power-of-Social-Media-for-Disaster-Relief.pdf (accessed 24 June 2015).

Ibrahim, I. N., T. Okabayashi, E. Ristiyanto, T. Lestari, Y. Yanase, H. Muramatsu, C. Ueno, and C. Morita. 1999. Serosurvey of wild rodents for Rickettsioses (spotted fever, murine typhus and Q fever) in Java Island, Indonesia. *European Journal of Epidemiology* 15 (1): 89-93. http://link.springer.com.dbgw.lis.curtin.edu.au/article/10.1023/A%3A1007547721171 (accessed 21 September 2013).

Indudharan, R., M. Ahmad, T. M. Ho, R. Salim, and Y. N. Htun. 1999. Human otoacariasis. *Annals of Tropical Medicine & Parasitology* 93 (2): 163-167.

Indudharan, R., A. S. Dharap, and T. M. Ho. 1996. Intra-aural tick causing facial palsy. *The Lancet* 348 (9027): 613. http://linkinghub.elsevier.com/retrieve/pii/S0140673605648271 (accessed 1 November 2014).

Irons, J. 2011. *Landsat 7 Science Data Users Handbook*. http://landsathandbook.gsfc.nasa.gov/program/ (accessed 30 March 2015).

Irons, J. R. 2015. *Landsat 8*. http://landsat.gsfc.nasa.gov/?p=3186 (accessed 4 April 2015).

Jalal, K. C. A., Kamaruzzaman, B. Y., Arshad, A., Ara, R., & Rahman, M. F. (2012). Diversity and distribution of fishes in tropical estuary Kuantan, Pahang, Malaysia. *Pak J Biol Sci*, *15*(12), 576-582. https://doi.org/10.3923/pjbs.2012.576.582

Jankowski, P., and T. Nyerges. 2001. *Geographic Information Systems For Group Decision Making Towards A Participatory, Geographic Information Science*: Taylor & Francis.

Jing, K. J., Shiang, L. F., Sieng, K. C., Bolongon, G. G., Kulaimi, N. A. M., & Bakar, S. A. (2017). Mini-review: Zoonotic agents from ticks in Malaysia. *Journal of Wildlife and Parks*, *32*, 67-77. https://jwp.wildlife.gov.my/index.php/jwp/article/view/51

Jiruše, M., J. Machek, V. Beneš, and P. Zeman. 2004. A Bayesian Estimate of the Risk of Tick-Borne Diseases. *Applications of Mathematics* 49 (5): 389-404. http://dx.doi.org/10.1023/B%3AAPOM.0000048119.55855.65 (accessed 9 November 2014).

John, D. J. I. 2011. Kuantan conurbation. PowerPoint slides. Economic Planning Unit, Prime Minister's Department.

Jones, S., W. Conner, and B. Song. 2012. Spatially Explicit Nonlinear Models for Explaining the Occurrence of Infectious Zoonotic Diseases. *ISRN Biomathematics* 2012: 12. http://dx.doi.org/10.5402/2012/132342 (accessed 14 November 2014).

JPNP, P. S. F. D. 2014. Info Hutan. http://forestry.pahang.gov.my/index.php/component/content/article/19-artikel-menu/info-hutan/306-info-hutan-2.html.

Jusoff, K. 2009. Land Use and Cover Mapping with Airborne Hyperspectral Imager in Setiu, Malaysia. *Journal of Agricultural Science* 1 (2): 120-131. http://search.proquest.com.dbgw.lis.curtin.edu.au/docview/840388014?rfr_id=info%3Axri%2Fsid%3Aprimo (accessed 11 November 2015).

Kalluri, S., P. Gilruth, D. Rogers, and M. Szczur. 2007. Surveillance of Arthropod Vector-Borne Infectious Diseases Using Remote Sensing Techniques: A Review. *PLoS Pathogens* 3 (10): e116. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2042005/?tool=pmcentrez (accessed 4th May 2011).

Karabatsos, N. (ed.) (1985) International Catalogue of Arboviruses including Certain Other Viruses of Vertebrates, 3rd edn, American Society of Tropical Medicine and Hygiene, San Antonio.

Kazmi, S. J. H., and E. L. Usery. 2001. Application of remote sensing and gis for the monitoring of diseases: A unique research agenda for geographers. *Remote Sensing Reviews* 20 (1): 45-70. http://dx.doi.org/10.1080/02757250109532427 (accessed 22 November 2012).

Kellerman, A. 2014. *The Internet as Second Action Space*. Hoboken: Hoboken : Taylor and Francis.

Kelly, M., S. D. Blanchard , E. Kersten, and K. Koy. 2011. Terrestrial Remotely Sensed Imagery in Support of Public Health: New Avenues of Research Using Object-Based Image Analysis. *Remote sensing* 3 (11): 2321. http://www.mdpi.com/2072-4292/3/11/2321 (accessed 26 November 2012).

Khorram, S., F. H. Koch, C. F. van der Wiele, and S. A. C. Nelson. 2012. *Remote Sensing / by Siamak Khorram, Stacy A.C. Nelson, Frank H. Koch, Cynthia F. van der Wiele*. Edited by S. A. C. Nelson, F. H. Koch, C. F. van der Wiele and SpringerLink. Boston, MA: Boston, MA : Springer US.

Kiffner, C., W. Zucchini, P. Schomaker, T. Vor, P. Hagedorn, M. Niedrig, and F. Rühe. 2010. Determinants of tick-borne encephalitis in counties of southern Germany, 2001-2008. *International Journal of Health Geographics* 9 (42): 10. http://www.ij-healthgeographics.com/content/9/1/42 (accessed 21st September 2011).

Kitron, U., and J. J. Kazmierczak. 1997. Spatial analysis of the distribution of Lyme disease in Wisconsin. *American Journal of Epidemiology* 145 (0002-9262 (Print)). http://aje.oxfordjournals.org/content/145/6/558.long (accessed 20 March 2014).

Klingseisen, B. 2010. *Spatio-temporal modelling of bluetongue virus distribution in Northern Australia based on remotely sensed bioclimatic variables*.

Klingseisen, B., R. J. Corner, and M. Stevenson. 2011. *Surveying and Spatial Sciences Biennial Conference, Developing spatio-temporal prediction models for arbovirus activity in nothern Australia based on remotely sensed bioclimatic variables*. Wellington, New Zealand: New Zealand Institute of Surveyors and the Surveying and Spatial Sciences Institute. . http://espace.library.curtin.edu.au/webclient/StreamGate?folder_id=0&dvs=1415604527018~36&usePid1=true&usePid2=true (accessed 10 November 2014).

Knülle, W., and T. L. Devine. 1972. Evidence for active and passive components of sorption of atmospheric water vapour by larvae of the tick Dermacentor variabilis. *Journal of Insect Physiology* 18 (9): 1653-1664. http://www.sciencedirect.com/science/article/pii/0022191072900959 (accessed 14th May 2012).

Koch, T. 2004. The Map as Intent: Variations on the Theme of John Snow. *Cartographica* 39 (4): 14. http://www.ph.ucla.edu/epi/snow/cartographica39%284%291_14_2004.pdf (accessed 13 January 2015).

Kolonin, G. V. 2007. Mammals as hosts of Ixodid ticks (Acarina, Ixodidae). *Entomological Review* 87 (4): 401-412. http://dx.doi.org/10.1134/S0013873807040033 (accessed 15 June 2012).

Koplan, J. P., T. C. Bond, M. H. Merson, K. S. Reddy, M. H. Rodriguez, N. K. Sewankambo, and J. N. Wasserheit. 2009. Towards a common definition of global health. *The Lancet* 373 (9679) 1993-1995. http://www.sciencedirect.com/science/article/pii/S0140673609603329 (accessed 27th July 2011).

K., T. and D. M. Wadhwa (2016). "Spatial Data System: Architecture and Applications." International Journal of Computer Science Trends and Technology (IJCST) **4**(5): 7.

Kuhn, K.G.; Campbell-Lendrum, D.H.; Davies, C.R.: Tropical Diseases in Europe? How we can learn from the past to predict the future. *EpiNorth*, 5: 6-13; 2004

Kumar, D. V. R. S., K. Sriram, K. M. Rao, and U. S. Murty. 2005. Management of filariasis using prediction rules derived from data mining. *Bioinformation* 1 (1): 4. http://www.bioinformation.net/001/000400012005.pdf (accessed 15 February 2015).

Labuda, M., & Nuttall, P. A. (2008). *Viruses transmitted by ticks*. Cambridge University Press.

Lafaye, M., C. Vignolles, J. Haynes, and S. Estes. 2014. CNES strategy: satellite data and modelling for public health: towards a cooperation with NASA. *Geocarto International* 29 (6):663-670.http://www-tandfonline-com.dbgw.lis.curtin.edu.au/doi/abs/10.1080/10106049.2013.827751 (accessed 11 September 2015).

Lam, S. K. 2002. Emerging and Re-Emerging Diseases in Malaysia. *Asia-Pacific Journal of Public Health* 14 (1): 6-8. http://aph.sagepub.com/content/14/1/6.abstract (accessed 23 May 2011).

Lam, S. K., and K. B. Chua. 2002. Nipah Virus Encephalitis Outbreak in Malaysia. *Clinical Infectious Diseases* 34 (Supplement 2): S48-S51. http://cid.oxfordjournals.org/content/34/Supplement_2/S48.abstract (accessed 4 September 2011).

Lam, S., K. Chua, P. Hooi, R. MA, S. Kumari, M. Tharmaratnam, S. Chuah, D. Smith, and I. Sampson. 2001. Chikungunya infection- an emerging disease in Malaysia. *The Southeast Asian journal of tropical medicine and public health* 32 (3): 5. http://www.thaiscience.info/journals/Article/Chikungunya%20infection-an%20emerging%20disease%20in%20malaysia.pdf (accessed 24 October 2014).

Lang, L. 2002. *GIS for Health Organizations*. Redland, CA: ESRI Press.

Last JM, editor. Dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001. p. 61

Lazim, N. M., I. Mohamad, M. K. Mohammed Daud, and R. Salim. 2012. The Many Faces of Intra-Aural Tick Clinical Presentation. *Journal of Pioneering Medical Sciences* 3 (1): 4. http://www.jpmsonline.com/wp-content/uploads/2012/12/JPMS-VOL3-ISSUE1-PAGES34-37-CR.pdf (accessed 7 August 2014).

Lee, E. K., C. Chien-Hung, F. Pietz, and B. Benecke. 2009. Modeling and Optimizing the Public-Health Infrastructure for Emergency Response. *Interfaces* 39 (5): 476-490. bth. http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=47181241&site=ehost-live (accessed 27th July 2011).

Li, H., F. Faruque, W. Williams, and R. Finley. 2005. Application of Geographic Information Systems to Syndromic Surveillance. In *Proceedings of ESRI 2005 International User Conference*. San Diego, USA.

Lim, B.L. (1972). Host relationships and seasonal abundance of immature ticks (Haemaphysalis spp. and Dermacentor spp.) in primary and mixed-secondary rainforests in West Malaysia. Southeast Asian Journal of Tropical Medicine and Public Health 3: 605-12.

Lindgren, E. 1998a. Climate and tickborne encephalitis. *Conservation Ecology [online]* 2 (1). http://www.ecologyandsociety.org/vol2/iss1/art5/ (accessed 26 July 2014).

Lindgren, E. 1998b. Climate change, tick-borne encephalitis and vaccination needs in Sweden—a prediction model. *Ecological Modelling* 110: 9. http://www.uco.com.es/servicios/informatica/windows/filemgr/download/stella/articulos/Lindgren_1998.pdf (accessed 26 July 2014).

Loh, W.-Y. 2014. Fifty Years of Classification and Regression Trees. *International Statistical Review* 82 (3): 329-348. http://onlinelibrary.wiley.com.dbgw.lis.curtin.edu.au/doi/10.1111/insr.12016/abstract (accessed 25 March 2015).

Looi, L. M., and K. B. Chua. 2007. Lessons from the Nipah virus outbreak in Malaysia. *Malays J Pathol* 29 (2): 63-7. http://www.mjpath.org.my/2007.2/02Nipah_Virus_lessons.pdf (accessed 16 March 2015).

Louisville, N. W. S. 2004. *A discussion on water vapour, humidity, dewpoint, and relationship to precipitation*. http://www.crh.noaa.gov/lmk/soo/docu/humidity.php (accessed 8 July 2014).

LPDAAC, L. P. D. A. A. C. 2014. *Vegetation Indices 16-Day L3 Global 250m MOD13Q1*. https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1 (accessed 20 March 2015).

Maccherone, B. 2014. *MODIS Data Products*. http://modis.gsfc.nasa.gov/data/dataprod/ (accessed 6 November 2014).

Mason, B. C., and S. Dragićević. 2006. Web GIS and Knowledge Management Systems: An Integrated Design for Collaborative Community Planning. In *Collaborative Geographic Information Systems*, ed. S. Balram and S. Dragićević.

Marcu, A., J. Barnett, D. Uzzell, K. Vasileiou, and S. O'Connell. 2013. Experience of Lyme disease and preferences for precautions: a cross-sectional survey of UK patients. *BMC Public Health* 13. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3681680/ (accessed 36 May 2014).

Marzban, C. 2004. The ROC Curve and the Area under It as Performance Measures. *Weather and Forecasting* 19 (6): 1106-1114. http://journals.ametsoc.org/doi/abs/10.1175/825.1 (accessed 16 April 2015).

Mat Akhir, J., S. M. S. A., and N. A. Mohammad. 2007. Land use evaluation for Kuala Selangor, Malaysia using remote sensing and GIS technologies. *Geografia : Malaysian Journal of Society and Space* 3 (1): 1-19. Land use evaluation for Kuala Selangor, Malaysia using remote sensing and GIS technologies (accessed 11 November 2015).

Matheson R. 1950. Medical Entomology, 2nd Edition. Comstock Publishing Company, Inc. Ithaca, NY.

Mathiyalagan, V., S. Grunwald, K. R. Reddy, and S. A. Bloomb. 2003. *ESRI User Conference, A WebGIS And Geodatabase For Florida's Wetlands*. San Diego, CA. http://www.cp-idea.org/documentos/tecnologia/10_webgisdatabase.pdf (accessed 6th June 2011).

Mayer, G. 2015. *Bacteriology - chapter twenty one Rickettsia, Orientia, Ehrlichia, Anaplasma, Coxiella and Bartonella*. http://www.microbiologybook.org/mayer/ricketsia.htm (accessed 1 April 2015).

McEnroe, W. D. 1971. The effect of automobile traffic on american dog tick distribution (Dermacentor variabilis, Say: Acarina, Ixodidae). *Environmental Pollution (1970)* 2 (2): 135-143. http://www.sciencedirect.com.dbgw.lis.curtin.edu.au/science/article/pii/0013932771900164 (accessed 25 July 2014).

McHugh, M. L. 2013. The Chi-square test of independence. *Biochemia Medica* 23 (2): 143-149. Pmc. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900058/ (accessed 16 April 2015).

Menard, S. 2010. Logistic Regression: From Introductory to Advanced Concepts and ApplicationsThousand Oaks, California: SAGE Publications, Inc.

MMD, M. M. D. 2013. *General Climate of Malaysia*. http://www.met.gov.my/index.php?option=com_content&task=view&id=75&Itemid=1089&limit=1&limitstart=0 (accessed 10 October 2014).

Miller, J. 2010. Species Distribution Modeling. *Geography Compass* 4 (6): 20. https://noppa.oulu.fi/noppa/kurssi/791626s/materiaali/791626S_artikkeli3_2.pdf (accessed 24 March 2015).

Milosavljevic, A., S. Dordevic-Kajan, and L. Stoimenov. 2008. An application framework for rapid development for web based GIS : Ginis Web. In *Geospatial Services and*

Applications for the Internet, ed. J. T. Sample, K. Shaw, S. Tu and M. Abdelguerfi, 48-71. Springer US.

MOH. 2012. Malaysia Strategic Workplan for Emerging Diseases (MySED Workplan) 2012-2015, edited by M. o. H. Malaysia. Kuala Lumpur. Ministry of Health Malaysia.

MOH, C. D. S. S. 2004. Standard Operating Procedure For Potential Infectious Disease, edited by D. C. D. Communicable Disease Surveillance Section. Communicable Disease Surveillance Section.

MOH, M. o. H. M. (2021). Health facts 2021. https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/HEALTH%20FACTS/Health_Facts_2021.pdf

MOHCW, M. o. H. a. C. W. Z. 2008. National Malaria Control Programme Strategy (NMCP) 2008-2013. Ministry of Health and Child Welfare Zimbabwe.

Mohd Azmi, M. I., R. Cullen, H. R. Bigsby, and A. G. Awang Noor. 2009. *NZARES Conference, The Existence Value of Peat Swamp Forest in Peninsular Malaysia*. Nelson, New Zealand. http://purl.umn.edu/97133 (accessed 16 February 2015).

Monello, R. J., and M. E. Gompper. 2007. Biotic and abiotic predictors of tick (Dermacentor variabilis) abundance and engorgement on free-ranging raccoons (Procyon lotor). *Parasitology* 134 (14): 2053-2062. http://dx.doi.org/10.1017/S0031182007003423 (accessed 23 May 2012).

Morand, S., & Lajaunie, C. (2021). Outbreaks of Vector-Borne and Zoonotic Diseases Are Associated With Changes in Forest Cover and Oil Palm Expansion at Global Scale [Original Research]. *Frontiers in Veterinary Science, 8*(230). https://doi.org/10.3389/fvets.2021.661063

Muhammad, M. K. I., Shahid, S., Ismail, T., Harun, S., Kisi, O., & Yaseen, Z. M. (2021). The development of evolutionary computing model for simulating reference evapotranspiration over Peninsular Malaysia. *Theoretical and applied climatology*, *144*(3-4), 1419-1434. https://doi.org/10.1007/s00704-021-03606-z

Murty, U. S. N., and N. Arora. 2007. Application of Correlation and Regression Tree (CART) for management of Malaria in Arunachal Pradesh, India. *The Internet Journal of Tropical Medicine* 5 (1). https://ispub.com/IJTM/5/1/10515 (accessed 15 February 2015).

Mwambi, H. G. 2002. Ticks and tick-borne diseases in Africa: a disease transmission model. *Mathematical Medicine and Biology* 19 (4): 275. ProQuest Central; ProQuest SciTech Collection. http://imammb.oxfordjournals.org.dbgw.lis.curtin.edu.au/content/19/4/275.full.pdf+html (accessed 6 November 2014).

Myers, M. F., D. J. Rogers, J. Cox, A. Flahault, and S. I. Hay. 2000. Forecasting disease risk for increased epidemic preparedness in public health. In *Advances in Parasitology*, ed. S. E. R. D. J. R. S.I. Hay, 309-330. Academic Press.

Nadchatram, M. 2008. The beneficial rainforest ecosystem with environmental effects on zoonoses involving ticks and mites (Acari). A Malaysian perspective and review. *Tropical Biomedicine (Supplement)* **25**: 1–92.

NASA, N. A. a. S. A. n.d. *MODIS Specifications*. http://modis.gsfc.nasa.gov/about/specifications.php (accessed 30 March 2015).

Nasir, M. F. M., M. Abdul Zali, H. Juahir, H. Hussain, S. M. Zain, and N. Ramli. 2012. Application of receptor models on water quality data in source apportionment in Kuantan River Basin.

Iranian Journal of Environmental Health Science and Engineering 9 (18): 12. http://www.ijehse.com/content/9/1/18 (accessed 15 January 2015).

National Snow & Ice Data Center, N. (2021). *What are the limitations of MODIS data?* https://nsidc.org/support/21553418-What-are-the-limitations-of-MODIS-data-

Needham, G. R., and P. D. Teel. 1991. Off-Host Physiological Ecology of Ixodid Ticks. In *Annu. Rev. Entomol.*

Ogden, N. H., C. Bouchard, K. Kurtenbach, G. Margos, L. R. Lindsay, L. Trudel, S. Nguon, and F. Milord. 2010. Active and passive surveillance and phylogenetic analysis of Borrelia burgdorferi elucidate the process of Lyme disease risk emergence in Canada. *Environmental Health Perspectives* 118 (7): 909-914. rzh. http://search.ebscohost.com/login.aspx?direct=true&db=rzh&AN=2010731816&site=ehost-live Publisher URL: www.cinahl.com/cgi-bin/refsvc?jid=2714&accno=2010731816 (accessed 20 March 2012).

Olwoch, J. M., B. Reyers, and A. S. Van Jaarsveld. 2009. Host–parasite distribution patterns under simulated climate: implications for tick-borne diseases. *International Journal of Climatology* 29 (7): 993-1000. http://onlinelibrary.wiley.com.dbgw.lis.curtin.edu.au/doi/10.1002/joc.1801/epdf (accessed 15 June 2013).

Osnas, E. E., D. M. Heisey, R. E. Rolley, and M. D. Samuel. 2009. Spatial and temporal patterns of chronic wasting disease: fine-scale mapping of a wildlife epidemic in Wisconsin. *Ecological Applications* 19 (5): 1311-1322. http://dx.doi.org/10.1890/08-0578.1 (accessed 26 July 2014).

Packierisamy, P. R., Ng, C. W., Dahlui, M., Inbaraj, J., Balan, V. K., Halasa, Y. A., & Shepard, D. S. (2015). Cost of Dengue Vector Control Activities in Malaysia. *Am J Trop Med Hyg*, *93*(5), 1020-1027. https://doi.org/10.4269/ajtmh.14-0667

Painho, M., M. Peixoto, P. Cabral, and R. Sena. 2001. *ESRI User Conference, 9-13 July 2001: WebGIS as a teaching tool*. San Diego, CA. http://www.isegi.unl.pt/unigis/papers/ESRI_2001.pdf (accessed 6th June 2011).

Pakhriazad, H. Z., I. Mohd Hasmadi, H. M. K. Aida, and K. Jusoff. 2009. Historical and Current Legislations of Taman Negara National Park Peninsular Malaysia. *Journal of Politics and Law* 2 (1): 6. http://ccsenet.org/journal/index.php/jpl/article/view/384/344 (accessed 15 January 2015).

Palaniappan, S., and R. Awang. 2007. *9th International Conference on Information Integration and Web-based Applications & Services, Web-Based Heart Disease Decision Support System using Data Mining Classification Modeling Techniques*. Jakarta, Indonesia. http://pdf.aminer.org/000/248/170/using_neural_networks_for_modeling_the_input_requirements_of_electronic.pdf (accessed 26 July 2014).

Parola, P., and D. Raoult. 2001. Ticks and Tickborne Bacterial Diseases in Humans: An Emerging Infectious Threat. *Clinical Infectious Diseases* 32 (6): 32. http://cid.oxfordjournals.org/content/32/6/897.long (accessed 9 March 2015).

Parola, P., J.-P. Cornet, Y. O. Sanogo, R. S. Miller, H. V. Thien, J.-P. Gonzalez, D. Raoult, S. R. Telford Iii, and C. Wongsrichanalai. 2003. Detection of Ehrlichia spp., Anaplasma spp., Rickettsia spp., and Other Eubacteria in Ticks from the Thai-Myanmar Border and Vietnam. *Journal of Clinical Microbiology* 41 (4): 1600. http://jcm.asm.org.dbgw.lis.curtin.edu.au/content/41/4/1600 (accessed 2 October 2014).

Peng, Z.-R., and M.-H. Tsou. 2003. *Internet GIS : Distributed Geographic Information Services For The Internet And Wireless Networks*: John Wiley & Sons Inc.

Petney, N., Keirans, J.E., 1996, Ticks of the genera Boophilus, Dermacentor, Nosomma and Rhipicephalus (Acari: Ixodidae) in South-east Asia, Tropical Biomedicine, 13, 73-84.

Petney, T., R. Robbins, A. Guglielmone, D. Apanaskevich, A. Estrada-Peña, I. Horak, and R. Shao. 2011. A Look at the World of Ticks. In *Progress in Parasitology*, ed. H. Mehlhorn, 283-296. Springer Berlin Heidelberg.

Pfeiffer, D. U. 1996. *Proceedings of the epidemiology and state veterinary programmes New Zealand Veterinary Association / Australian Veterinary Association Second Pan Pacific Veterinary Conference, 23-28 June 1996: Issue related to handling of spatial data*. Christchurch. ftp://131.252.97.79/Transfer/ES_Pubs/ESVal/spatial_statistics/issues_related_to_spatial_data.pdf (accessed 9 July 2014).

Pfeiffer, D. U., T. P. Robinson, M. Stevenson, K. B. Stevens, D. J. Rogers, and A. C. A. Clements. 2008. *Spatial Analysis in Epidemiology*. Oxford: Oxford University Press.

Piarsa, I. N., A. A. K. O. Sudana, and G. W. M. Gunadi. 2012. Web-based GIS by using Spatial Decision Support System (SDSS) Concept for Searching Commercial Marketplace - using Google MAP API. *International Journal of Computer Applications* 50 (7): 5. http://research.ijcaonline.org/volume50/number7/pxc3880867.pdf (accessed 24 July 2014).

Porretta, D., V. Mastrantonio, S. Mona, S. Epis, M. Montagna, D. Sassera, C. Bandi, and S. Urbanelli. 2013. The integration of multiple independent data reveals an unusual response to P leistocene climatic changes in the hard tick I xodes ricinus. *Molecular Ecology* 22 (6): 1666-1682.

http://onlinelibrary.wiley.com.dbgw.lis.curtin.edu.au/doi/10.1111/mec.12203/epdf (accessed 15 January 2014).

Pullan, R., P. W. Gething, J. L. Smith, C. S. Mwandawiro, H. J. W. Sturrock, C. W. Gitonga, S. I. Hay, and S. Brooker. 2011. Spatial modelling of soil-transmitted helminth infections in Kenya: a disease control planning tool. *PLoS Neglected Tropical Diseases* 5 (2): e958. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3035671/?tool=pmcentrez (accessed 22 August 2011).

Rahman, M. M., M. B. Awang, K. C. A. Jalal, S. Aisha, and B. Y. Kamaruzzaman. 2013. Study on Toxic Chemicals in Kuantan River During Pre and Post Monsoon Season. *Australian Journal of Basic and Applied Sciences* 7 (4): 6. http://ajbasweb.com/old/ajbas/2013/March/24-30.pdf (accessed 16 February 2015).

Ramakrishnan, R. and J. Gehrke (2003). Database management systems / Raghu Ramakrishnan, Johannes Gehrke. Boston, Boston : McGraw-Hill.

Ramsay, M. A. E. 2006. John Snow, MD: anaesthetist to the Queen of England and pioneer epidemiologist. *Baylor University Medical Center* 19 (1): 5. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1325279/pdf/bumc0019-0024.pdf (accessed 22 July 2014).

Randolph, S. E. 2000. Ticks and tick-borne disease systems in space and from space. In *Advances in Parasitology*, ed. S. I. Hay, S. E. Randolph and D. J. Rogers, 217-243. Academic Press.

Reinhardt, M., J. Elias, J. Albert, M. Frosch, D. Harmsen, and U. Vogel. 2008. EpiScanGIS: an online geographic surveillance system for meningococcal disease. *International Journal of Health Geographics* 7 (1): 33-33. http://www.ij-healthgeographics.com/content/7/1/33 (accessed 27th May 2011).

Reiter, P. 2001. Climate Change and Mosquito-Borne Disease. *Environmental Health Perspectives* 109 (ArticleType: research-article / Issue Title: Supplement 1: Reviews in Environmental Health, 2001 / Full publication date: Mar., 2001 / Copyright © 2001 The National Institute of Environmental Health Sciences (NIEHS)): 141-161. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1240549/pdf/ehp109s-000141.pdf (accessed 29 October 2014).

Richards, T., C. Croner, G. Rushton, C. Brown, and L. Fowler. 1999. Geographic information systems and public health: mapping the future. *Public Health Rep* 114: 359 - 373.

Richards, B., J. Rawlings, and H. Hollan. 2000. Tick Texas: Ticks 101-anatomy. http://www.ticktexas.org/ticks/ticks101_anatomy.htm.

Rinner, C., C. Keßler, and S. Andrulis. 2008. The use of Web 2.0 concepts to support deliberation in spatial decision-making. *Computers, Environment and Urban Systems* 32 (5): 386-395. http://www.sciencedirect.com/science/article/pii/S0198971508000574 (accessed 10th October 2011).

Rinner, C., and P. Jankowski. 2002. Web-Based Spatial Decision Support – Technical Foundations and Applications. In *The Encyclopedia of Life Support Systems (EOLSS), Theme 1.9 – Advanced Geographic Information Systems edited by Claudia Bauzer Medeiros*. UNESCO / Eolss Publishers.

Rizzoli, A., Silaghi, C., Obiegala, A., Rudolf, I., Hubálek, Z., Földvári, G., Plantard, O., Vayssier-Taussat, M., Bonnet, S., Spitalská, E., & Kazimírová, M. (2014). Ixodes ricinus and Its Transmitted Pathogens in Urban and Peri-Urban Areas in Europe: New Hazards and Relevance for Public Health. *Frontiers in public health*, *2*, 251-251. https://doi.org/10.3389/fpubh.2014.00251

Roberts, D. R., and R. G. Andre. 1994. Insecticide resistance issues in vector-borne disease control. *The American journal of tropical medicine and hygiene* 50 (6 Suppl): 21. http://www.ajtmh.org.dbgw.lis.curtin.edu.au/content/50/6_Suppl/21.full.pdf+html (accessed 13 January 2015).

Robinson, T. P. 2000. Spatial statistics and geographical information systems in epidemiology and public health. In *Advances in Parasitology*, ed. S. E. R. D. J. R. S.I. Hay, 81-128. Academic Press.

Rochlin, I., & Toledo, A. (2020). Emerging tick-borne pathogens of public health importance: A mini-review. *Journal of medical microbiology*, *69*(6), 781-791. https://doi.org/10.1099/jmm.0.001206

Rogers, D. J. 2006. Models for Vectors and Vector-Borne Diseases. In *Advances in Parasitology*, ed. A. G. Simon I. Hay and J. R. David, 1-35. Academic Press.

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W. (1973) Monitoring Vegetation Systems in the Great Plains with ERTS. Proceedings of the Third ERTS Symposium, Washington DC, 10-14 December 1973, 309-317.

Rozainah, M. Z., and M. R. Mohamad. 2006. Mangrove Forest Species Composition and Density in Balok River, Pahang, Malaysia. *Ecoprint* 13: 6. http://nepjol.info/index.php/ECO/article/download/1624/1814 (accessed 16 February 2015).

Russell, R. C. 2010. “Dramas” Down-Under: Changes and Challenges in Australia. In *Vector Biology, Ecology and Control*, ed. P. Atkinson, 81-96. Springer Netherlands.

Rutledge, D., G. McDonald, M. Cameron, G. McBride, J. Poot, F. Scrimgeour, R. Price, D. Phyn, H. van Delden, B. Huser, B. Small, L. Wedderburn, and T. Fenton. 2007. Development of Spatial Decision Support Systems to Support Long-term, Integrated Planning. In *MODSIM*

2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, edited by D. K. Les Oxley.

Saad, S., Ahmad, Z., Yunus, K., & Khan Chowdhury, A. J. (2009). Impact of coastal development on mangrove distribution in Kuantan, Pahang.

Sainani, K. L. 2014. Logistic Regression. *PM&R* 6 (12): 1157-1162. http://www.sciencedirect.com/science/article/pii/S1934148214014439 (accessed 5 April 2017).

Saini, D., Al Maskari, S., & Omar, W. (2010). *Web 2.0 and Healthcare Collaborative Framework for Health Awareness and Education*.

Salleh, A. M. 2003. *Report on AUICK Associate Cities Kuantan, Malaysia*. Asian Urban Information Centre of Kobe. http://www.auick.org/database/apc/apc043/apc0430305.html (accessed 7 July 2014).

Samson, G. L., et al. (2017). Spatial Databases: An Overview. Ontologies and Big Data Considerations for Effective Intelligence. J. Lu and Q. Xu. United States of America, IGI Global.

Science., I. o. M. U. C. o. L. D. a. O. T.-B. D. T. S. o. t. (2011). *Critical needs and gaps in understanding prevention, amelioration, and resolution of lyme and other tick-borne diseases the short-term and long-term outcomes, workshop report*. National Academies Press.

Scotch, M., and B. Parmanto. 2006. Development of SOVAT: A numerical-spatial decision support system for community health assessment research. *International Journal of Medical Informatics* 75 (10-11): 771-784. http://www.sciencedirect.com/science/article/pii/S1386505605002133 (accessed 15th July 2011).

Scotch, M., B. Parmanto, and V. Monaco. 2008. Evaluation of SOVAT: An OLAP-GIS decision support system for community health assessment data analysis. *BMC Medical Informatics and Decision Making* 8 (1): 22. http://www.biomedcentral.com/1472-6947/8/22 (accessed 23 November 2011).

Senay, S. D., S. P. Worner, and T. Ikeda. 2013. Novel Three-Step Pseudo-Absence Selection Technique for Improved Species Distribution Modelling. *PLoS ONE* 8 (8): e71218. Pmc. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3742778/ (accessed 23 March 2015).

Shah, U. U. (2018). Web 2.0 & Public Health. *Library philosophy and practice*, 1-10.

Sharifah, N., Heo, C. C., Ehlers, J., Houssaini, J., & Tappe, D. (2020). Ticks and tick-borne pathogens in animals and humans in the island nations of Southeast Asia: A review. *Acta tropica*, *209*, 105527-105527. https://doi.org/10.1016/j.actatropica.2020.105527

Shim, J. P., M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson. 2002. Past, present, and future of decision support technology. *Decision support systems* 33 (2): 111-126. http://www.sciencedirect.com.dbgw.lis.curtin.edu.au/science/article/pii/S0167923601001397 (accessed 5 July 2011).

Smitha.T, and V.Sundaram. 2012. Classification Rules by Decision Tree for Disease Prediction. *International Journal of Computer Applications* 43 (8): 3. http://research.ijcaonline.org/volume43/number8/pxc3878323.pdf (accessed 26 July 2014).

Socolovschi, C., O. Mediannikov, D. Raoult, and P. Parola. 2009. The relationship between spotted fever group Rickettsiae and Ixodid ticks. *Veterinary Research* 40 (2): 20. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2695030/pdf/vetres-40-34.pdf (accessed 24 May 2012).

Somayaji, K. S. G., and A. Rajeshwari. 2007. Human otoacariasis. *Indian Journal of Otolaryngology and Head & Neck Surgery* 59 (3): 237-239. Pmc. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3452122/ (accessed 13 January 2013).

Sonenshine, D. 1991. “Biology of Ticks”. Volume 1. Oxford University Press, New York.

Sonenshine, D. 1993. “Biology of Ticks”. Volume 2. Oxford University Press, New York.

Spatial Data Types Overview. (2021). Microsoft. https://docs.microsoft.com/en-us/sql/relational-databases/spatial/spatial-data-types-overview?view=sql-server-2017

Speybroeck, N. 2012. Classification and regression trees. *International Journal of Public Health* 57 (1): 243-246. http://dx.doi.org/10.1007/s00038-011-0315-z (accessed 10 April 2015).

Stokland, J. N., R. Halvorsen, and B. Støa. 2011. Species distribution modelling—Effect of design and sample size of pseudo-absence observations. *Ecological Modelling* 222 (11): 1800-1809. http://www.sciencedirect.com/science/article/pii/S0304380011001062 (accessed 11 July 2013).

Sug, H. 2012. *Better Decision Tree Induction for Limited Data Sets of Liver Disease*. Edited by T.-h. Kim, J.-J. Kang, W. I. Grosky, T. Arslan and N. Pissinou. Vol. 353, *Communications in Computer and Information Science*.

Sugumaran, R., and J. Degroote. 2010. *Spatial decision support systems : principles and practice*: Boca Raton : Taylor & Francis.

Sugumaran, V., and R. Sugumaran. 2005. Web-based Spatial Decision Support Systems (WebSDSS) : Evolution, Architecture, and Challenges. In *Third Annual SIGDSS Pre-ICIS Workshop Designing Complex Decision Support: Discovery and Presentation of Information and Knowledge.* Las Vegas, Nevada.

Sugumaran, V., and R. Sugumaran. 2007. Web-Based Spatial Decision Support Systems (WebSDSS): Evolution, Architecture, Examples And Challenges. *Communications of the Association for Information Systems* 19: 34. http://aisel.aisnet.org/cgi/viewcontent.cgi?article=2683&context=cais (accessed 25 July 2014).

Sultana, S. R., A. Ali, A. Ahmad, M. Mubeen, M. Zia-Ul-Haq, S. Ahmad, S. Ercisli, and H. Z. E. Jaafar. 2014. Normalized Difference Vegetation Index as a Tool for Wheat Yield Estimation: A Case Study from Faisalabad, Pakistan. *The Scientific World Journal* 2014: 8. http://dx.doi.org/10.1155/2014/725326 (accessed 14 April 2015).

Suss, J., C. Klaus, F.-W. Gerstengarbe, and P. C. Werner. 2008. What Makes Ticks Tick? Climate Change, Ticks, and Tick-Borne Diseases. *Journal of Travel Medicine* 15 (1): 6. http://onlinelibrary.wiley.com/doi/10.1111/j.1708-8305.2007.00176.x/full (accessed 9 April 2012).

Sutherst, R. W. 1971. An Experimental Investigation into the Effects of Flooding on the Ixodid Tick Boophilus microplus (Canestrini). *Oecologia* 6 (3): 208-222. http://www.jstor.org/stable/4214641 (accessed 13 January 2014).

Talha, R., P. L. Nair, and M. F. Ismail. 2009. Population and Housing Census 2010 Malaysia 2010: Development in Technology and Approach. *Journal of the Department of Statistics, Malaysia* 2: 13. http://www.statistics.gov.my/portal/download_journals/files/2009/Volume2/contents_Population_and_Housing_Census2010.pdf (accessed 23 January 2015).

Tatem, A. J., S. J. Goetz, and S. I. Hay. 2004. Terra and Aqua: new data for epidemiology and public health. *International Journal of Appled Earth Observation and Geoinformation* 6

(1): 33-46. http://www.ncbi.nlm.nih.gov/pmc/articles/pmid/22545030/ (accessed 15 October 2014).

Thomas, S., Beh, L., & Nordin, R. B. (2011). Health care delivery in Malaysia: changes, challenges and champions. *J Public Health Afr*, *2*(2), e23. https://doi.org/10.4081/jphia.2011.e23

Tran, A., J. Gardon, S. Weber, and L. Polidori. 2002. Mapping Disease Incidence in Suburban Areas using Remotely Sensed Data. *American Journal of Epidemiology* 156 (7): 662-668. http://aje.oxfordjournals.org/content/156/7/662.abstract (accessed 22 November 2012).

Tran, A., N. Ponçon, C. Toty, C. Linard, H. Guis, J.-B. Ferré, D. L. Seen, F. Roger, S. de la Rocque, D. Fontenille, and T. Baldet. 2008. Using remote sensing to map larval and adult populations of Anopheles hyrcanus (Diptera: Culicidae) a potential malaria vector in Southern France. *International Journal of Health Geographics* 7: 12. http://www.ij-healthgeographics.com/content/7/1/9 (accessed 26 July 2014).

USGS, U. S. G. S. 2013. *Frequently Asked Questions about the Landsat Missions*. https://landsat.usgs.gov/best_spectral_bands_to_use.php (accessed 30 March 2015).

USGS, U. S. G. S. 2014. *Frequently Asked Questions about the Landsat Missions*. http://landsat.usgs.gov/band_designations_landsat_satellites.php (accessed 30 March 2015).

Uspensky, I. (2015). *Urbanization, ticks and tick-borne diseases*.

VanDerWal, J., L. P. Shoo, C. Graham, and S. E. Williams. 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling* 220 (4): 589-594. http://www.sciencedirect.com/science/article/pii/S0304380008005486 (accessed 11 July 2013).

Van Peenan, P.F.D.; Carney, W.P.; Sudomo, M.; & SUlianti Saroso,J. 1974. Parasites of mammals in Gumbasa valley, central Sulawesi, Indonesia. Tropical and Geographical Medicine 26:352-358.

van Velsen, L., J. E. van Gemert-Pijnen, D. J. Beaujean, J. Wentzel, and J. E. van Steenbergen. 2012. Should health organizations use web 2.0 media in times of an infectious disease crisis? An in-depth qualitative study of citizens' information behavior during an EHEC outbreak. *J Med Internet Res* 14 (6). http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3799610/ (accessed 8 June 2015).

Varlamis, I., & Apostolakis, I. (2010). The Evolution of Healthcare Applications in the Web 2.0 Era. In (Vol. 309, pp. 315-328). https://doi.org/10.1007/978-3-642-14464-6_14

Vlachokostas, C., C. Achillas, N. Moussiopoulos, E. Hourdakis, G. Tsilingiridis, L. Ntziachristos, G. Banias, N. Stavrakakis, and C. Sidiropoulos. 2009. Decision support system for the evaluation of urban air pollution control options: Application for particulate pollution in Thessaloniki, Greece. *Science of The Total Environment* 407 (23): 5937-5948. http://www.sciencedirect.com/science/article/pii/S0048969709007414 (accessed 26 July 2014).

Vredevoe, L. 2014. Background Information on the Biology of Ticks. *Tick Biology*, http://entomology.ucdavis.edu/Faculty/Robert_B_Kimsey/Kimsey_Research/Tick_Biology/.

Walker, D. H. 1996. *Medical Microbiology*, ed S. Baron. Galveston, Texas: University of Texas Medical Branch at Galveston. http://www.ncbi.nlm.nih.gov/books/NBK7624/ (accessed 31 October 2014).

Walter, M., Vogelgesang, J. R., Rubel, F., & Brugger, K. (2020). Tick-Borne Encephalitis Virus and Its European Distribution in Ticks and Endothermic Mammals. *Microorganisms, 8*(7). https://doi.org/10.3390/microorganisms8071065

Wasseef, H. Y., and H. Hoogstraal. 1988. Dermacentor (Indocentor) steini (Acari: Ixodoidea: Ixodidae): hosts, distribution in the Malay Peninsula, Indonesia, Borneo, Thailand, the Philippines, and New Guinea. *Dermacentor (Indocentor) steini (Acari: Ixodoidea: Ixodidae): hosts, distribution in the Malay Peninsula, Indonesia, Borneo, Thailand, the Philippines, and New Guinea* 25 (5): 315-320. http://jme.oxfordjournals.org.dbgw.lis.curtin.edu.au/content/25/5/315 (accessed 15 March 2014).

Wimberly, M. C., A. D. Baer, and M. J. Yabsley. 2008. Enhanced spatial models for predicting the geographic distributions of tick-borne pathogens. *International Journal of Health Geographics* 7 (1): 15. http://www.ij-healthgeographics.com/content/7/1/15 (accessed 22 August 2011).

WHO. 2005. *Deaths from vector borne disease*. http://www.who.int/heli/risks/vectors/en/vbdmap.pdf (accessed 31 October 2014).

WHO. 2014. A global brief on vector-borne diseases. http://apps.who.int/iris/bitstream/10665/111008/1/WHO_DCO_WHD_2014.1_eng.pdf

Win, N. L., and K. M. Win. 2014. The Probability Distributions of Daily Rainfall for Kuantan River Basin in Malaysia. *International Journal of Science and Research (IJSR)* 3 (8): 7. http://www.ijsr.net/archive/v3i8/MDIwMTUzODQ%3D.pdf (accessed 19 January 2015).

Wormser, G. P., R. J. Dattwyler, E. D. Shapiro, J. J. Halperin, A. C. Steere, M. S. Klempner, P. J. Krause, J. S. Bakken, F. Strle, G. Stanek, L. Bockenstedt, D. Fish, J. S. Dumler, and R. B. Nadelman. 2006. The Clinical Assessment, Treatment, and Prevention of Lyme Disease, Human Granulocytic Anaplasmosis, and Babesiosis: Clinical Practice Guidelines by the

Infectious Diseases Society of America. *Clinical Infectious Diseases* 43 (9): 47. http://cid.oxfordjournals.org/content/43/9/1089.full (accessed 26 March 2015).

Wyk, I. C. V., A. Goddard, B. M. De C Bronsvoort, J. A. W. Coetzer, I. G. Handel, O. Hanotte, A. M. Y. Jennings, M. Lesosky, H. Kiara, S. M. Thumbi, P. Toye, M. W. Woolhouse, and B. L. Penzhorn. 2014. The impact of co-infections on the haematological profile of East African Short-horn Zebu calves. *Parasitology* 141 (3): 374-88. ProQuest Central; ProQuest Public Health; ProQuest SciTech Collection. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4021814/pdf/S0031182013001625a.pdf (accessed 9 November 2014).

Xiaoshuan, Z., F. Zetian, C. Wengui, T. Dong, and Z. Jian. 2009. Applying evolutionary prototyping model in developing FIDSS: An intelligent decision support system for fish disease/health management. *Expert Systems with Applications* 36 (2, Part 2): 3901-3913. http://www.sciencedirect.com/science/article/pii/S095741740800170X (accessed 26 July 2014).

Young, S., J. Malilay, C. Sanchez, and D. Batts-Osborne. 2004. *UJNR 36th Joint Meeting Panel on Wind and Seismic Effects, Integrating Public Health, Earth Observation Sciences, and Engineering Information Systems for Critical Disaster Reduction*. Gaithersburg. http://www.pwri.go.jp/eng/ujnr/joint/36/paper/73young.pdf (accessed 7 December 2012).

Zou, K. H., A. J. O'Malley, and L. Mauri. 2007. Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models. *Statistical Primer for Cardiovascular Research* 115: 4. http://circ.ahajournals.org/content/115/5/654.full (accessed 22 April 2015).

APPENDICES

METEOROLOGICAL DATA

(TEMPERATURE, RAINFALL AMOUNT, NUMBER OF RAINDAYS, RELATIVE HUMIDITY AND WIND SPEED)

Records of Mean Maximum Temperature

Unit :° C

Month / Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	29.9	31.6	32.9	33.8	33.4	33.4	33.3	32.8	32.1	33.1	32.1	30.5	32.4
2003	29.4	30.6	32.1	32.8	33.5	33.6	32.3	32.9	32.7	32.3	31.2	29.4	31.9
2004	30.4	31.4	32.0	34.5	34.4	32.6	32.1	32.9	32.1	31.2	30.4	28.3	31.9
2005	29.1	32.3	31.6	32.9	33.1	32.9	32.7	33.1	33.3	31.5	31.2	30.1	32.0
2006	30.1	30.3	32.2	32.8	32.7	32.9	33.0	32.8	32.4	32.3	31.1	29.9	31.9
2007	28.7	30.6	32.0	32.4	33.2	32.9	32.5	33.1	32.9	32.7	31.5	30.0	31.9

Records of Mean Minimum Temperature

Unit :° C

Month / Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	22.8	22.1	23.1	24.0	24.7	24.2	24.2	23.8	23.7	23.7	23.8	24.0	23.7
2003	23.2	23.2	24.0	24.1	24.9	24.1	23.8	24.0	23.7	23.6	23.6	23.0	23.8
2004	23.2	22.8	24.2	24.0	24.3	23.2	22.7	23.3	22.4	22.5	22.7	22.2	23.1
2005	21.3	22.3	23.0	24.4	24.2	24.3	23.8	24.0	23.6	23.6	23.3	23.4	23.4
2006	22.9	23.5	23.3	24.0	24.0	24.1	24.3	23.7	23.6	23.9	23.4	23.6	23.7
2007	23.3	23.0	23.7	24.1	24.2	24.4	23.9	23.7	23.9	23.5	23.6	23.1	23.7

JABATAN METEOROLOGI MALAYSIA

Station : Kuantan

Lat. : 03° 47' N

Long. : 103° 13' E

Ht. above M.S.L.: 15.3 m

Records of Monthly Rainfall Amount

Unit : mm

Month Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	61.2	10.0	158.3	118.4	346.6	144.4	147.4	344.5	169.9	173.4	218.3	644.4	2536.8
2003	521.4	161.8	248.4	254.8	56.8	90.2	312.4	166.2	209.6	314.8	431.3	1031.9	3799.6
2004	208.8	22.2	130.8	37.8	274.8	208.0	193.6	172.4	204.0	591.2	343.2	928.2	3315.0
2005	85.8	57.0	124.0	138.2	293.6	127.8	188.8	141.2	199.0	179.2	757.8	420.6	2713.0
2006	222.8	470.4	108.4	135.2	152.2	194.0	119.0	191.2	302.4	147.8	276.6	657.8	2977.8
2007	414.4	35.2	79.0	167.2	270.4	296.2	214.2	189.0	150.0	520.2	248.2	889.6	3473.6

Records of Number of Raindays

Month Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	17	7	4	10	18	12	10	15	19	14	22	25	173
2003	22	13	16	16	10	14	20	10	17	25	26	22	211
2004	19	6	13	7	15	15	15	13	18	21	20	16	178
2005	13	5	10	11	19	11	13	9	17	23	23	21	175
2006	13	18	10	17	16	15	13	16	13	15	19	16	181
2007	24	8	17	13	14	11	19	13	12	20	20	21	192

Ht. above M.S.L. : 15.3 m

Records of 24 Hour Mean Temperature

Unit :° C

Month / Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	25.7	26.2	27.3	28.1	28.2	28.1	28.0	27.3	26.8	27.2	26.9	26.5	27.2
2003	25.6	26.2	27.2	27.6	28.4	28.0	27.0	27.7	27.2	26.8	26.3	25.3	26.9
2004	25.9	26.3	27.2	28.4	28.1	27.0	26.3	27.0	25.9	25.6	25.5	24.5	26.5
2005	24.5	26.5	26.4	27.9	27.7	27.8	27.4	27.7	27.3	26.6	26.0	25.9	26.8
2006	25.7	26.1	26.9	27.3	27.3	27.5	27.6	27.3	27.1	27.3	26.3	26.1	26.9
2007	25.3	26.1	27.1	27.4	27.6	27.8	27.3	27.3	27.4	27.0	26.4	25.6	26.9

Records of 24 hour Mean Relative Humidity

Unit : %

Month / Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	84.4	77.4	77.5	78.7	82.1	79.6	79.5	80.7	82.7	82.2	84.1	87.8	81.4
2003	86.3	82.7	82.7	82.9	80.4	80.4	83.6	80.3	81.1	84.8	89.9	89.1	83.7
2004	87.0	83.0	85.6	80.2	83.3	84.6	83.7	82.8	85.4	88.2	88.5	89.2	85.1
2005	85.6	80.4	82.9	81.8	83.9	83.1	81.8	81.5	82.9	86.1	88.5	89.3	84.0
2006	86.0	85.7	84.0	84.9	86.1	84.3	83.7	82.7	83.5	84.3	87.3	86.8	84.9
2007	88.7	82.6	83.6	84.5	84.7	85.2	85.0	83.6	85.0	87.2	89.4	90.8	85.9

Ht. of anemometer head above ground: 14.0 m

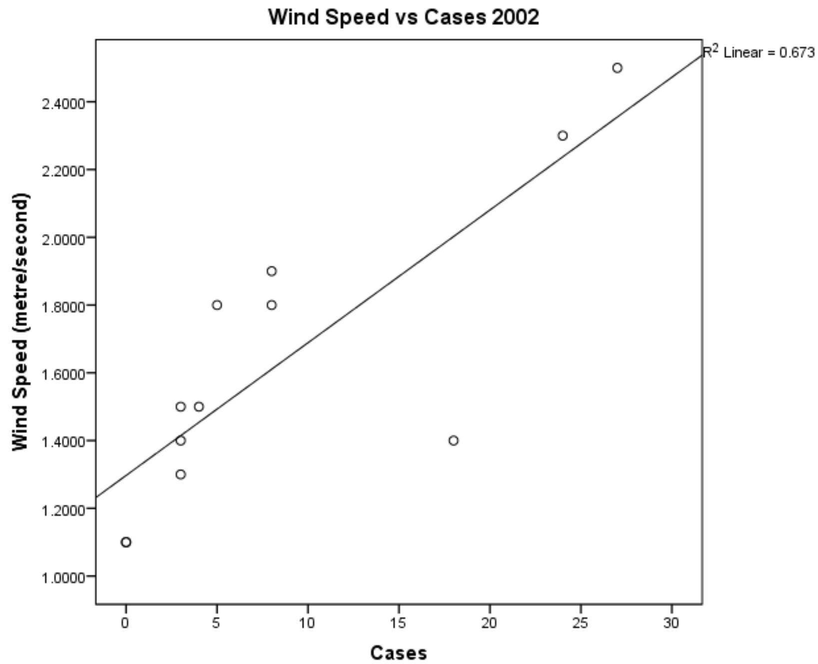
w.e.f. Apr 2005, Wind Sensor been installed.

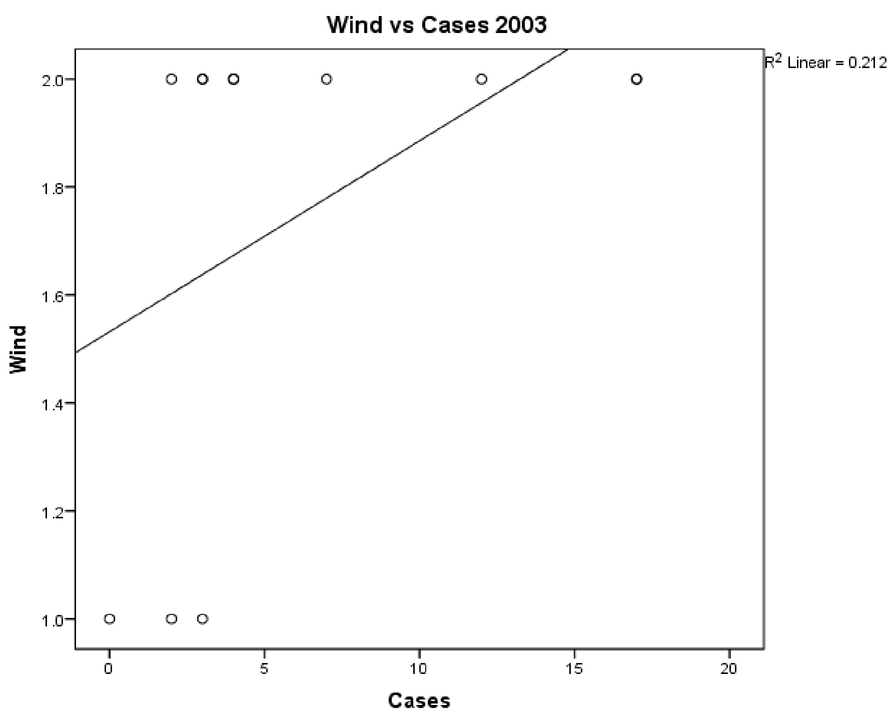
Ht. of Wind Sensor above ground: 10.0 m

Month / Year	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Annual
2002	2.5	2.3	1.9	1.5	1.3	1.4	1.8	1.8	1.5	1.1	1.1	1.4	1.6
2003	2.1	2.3	1.9	1.4	1.6	1.5	1.6	1.7	1.5	1.4	1.4	2.1	1.7
2004	2.3	2.1	1.8	1.6	1.4	1.5	1.7	1.9	1.4	1.4	1.6	2.0	1.7

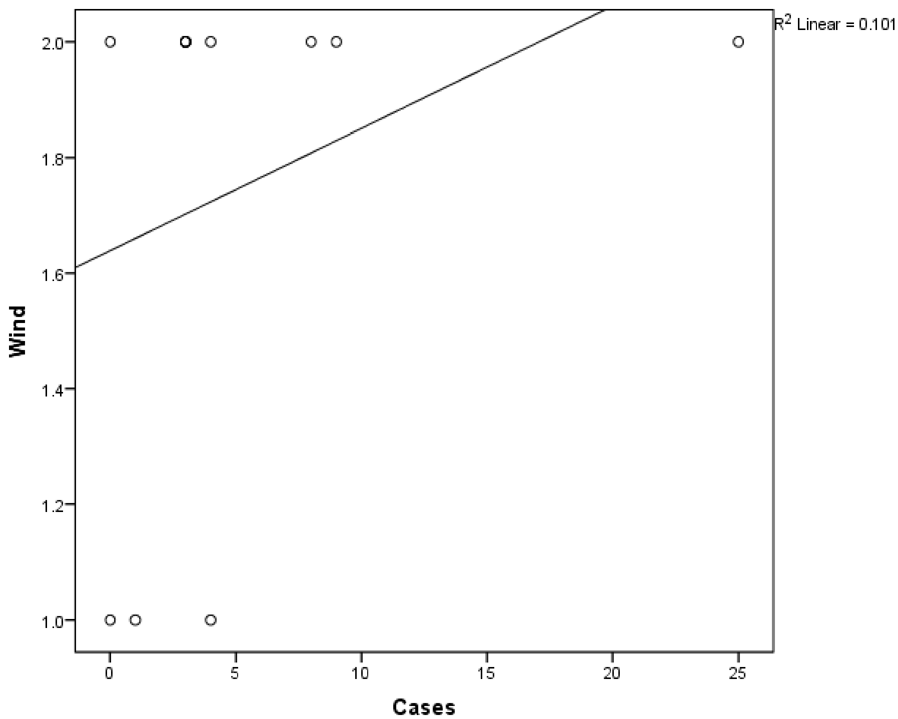
2005	2.1	2.1	2.1	1.6	1.5	1.5	1.7	1.7	1.5	0.8	0.8	0.9	1.5
2006	1.6	2.2	1.7	1.4	1.4	1.8	1.8	1.8	1.7	1.4	1.4	1.9	1.7
2007	1.9	1.9	1.6	1.5	1.5	1.4	1.5	1.8	1.6	1.5	1.2	1.7	1.6

CORRELATION BETWEEN WIND SPEEDS AGAINST NUMBER OF CASES FOR EACH YEAR FROM 2002 TO 2007

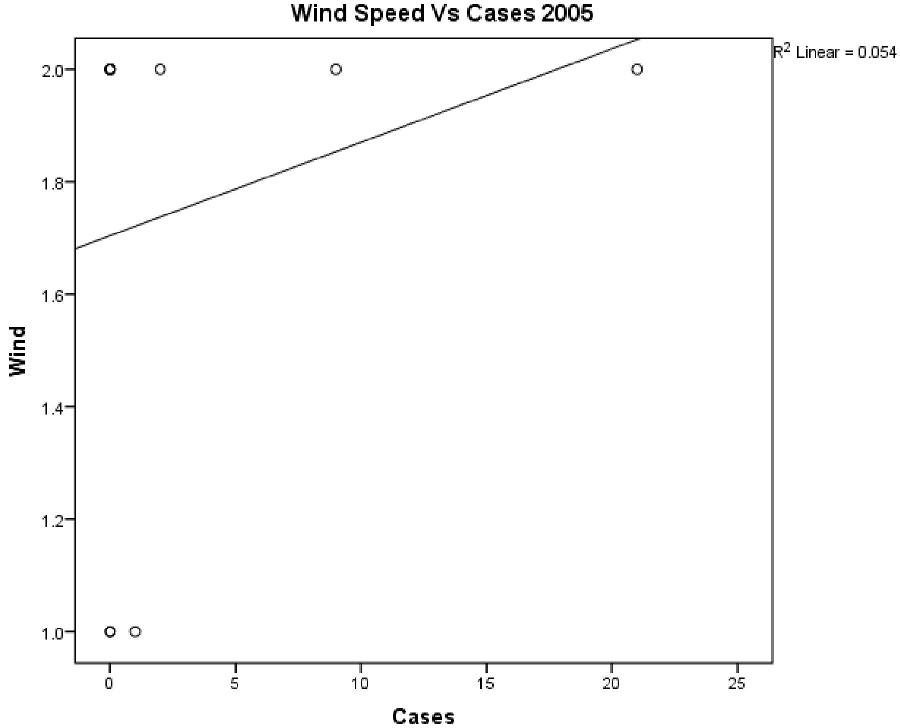


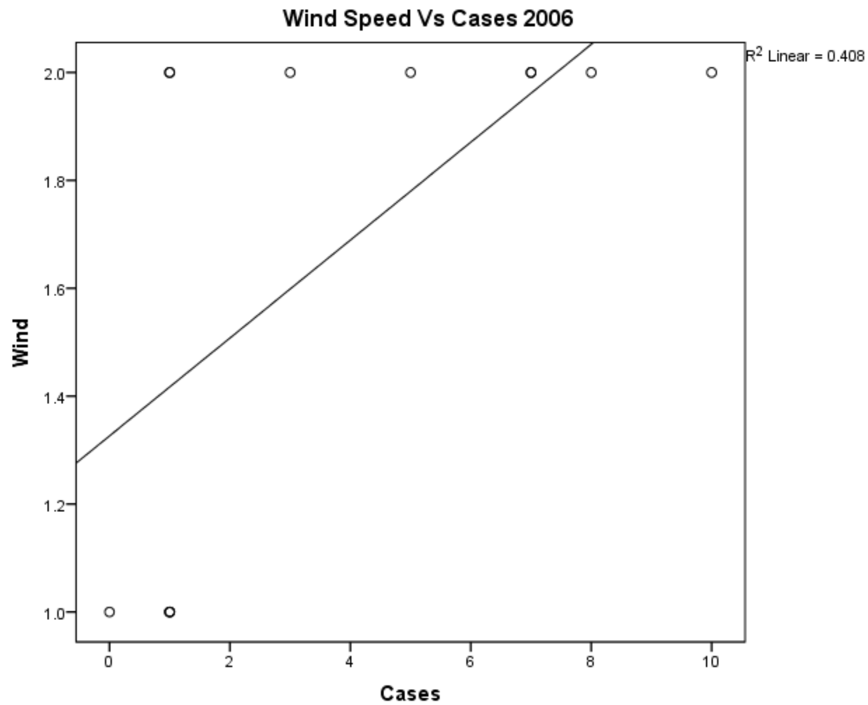


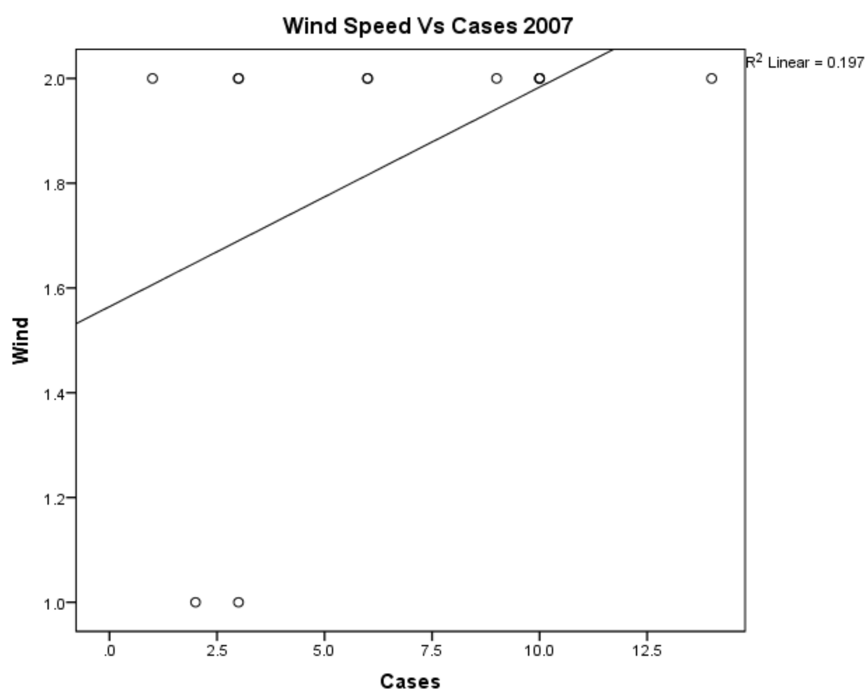
Wind Speed Vs Cases 2004



Wind Speed Vs Cases 2005

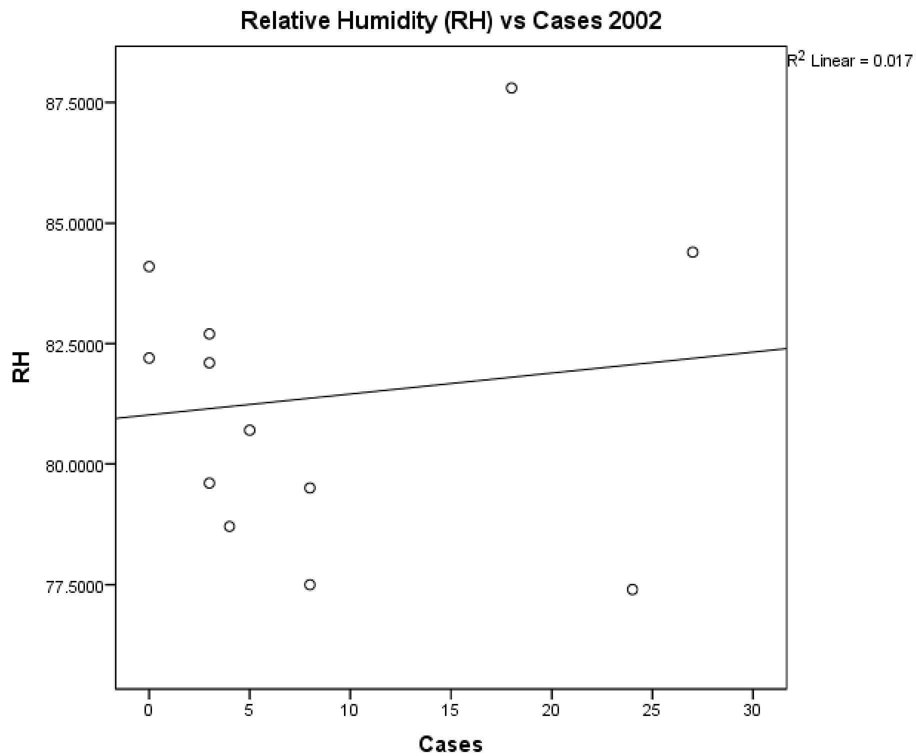


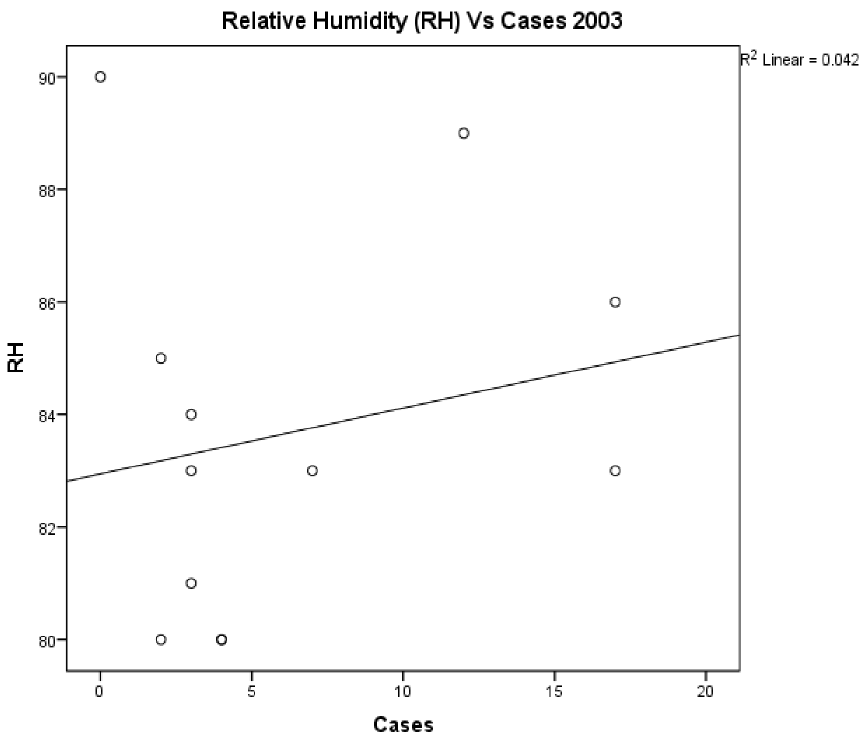


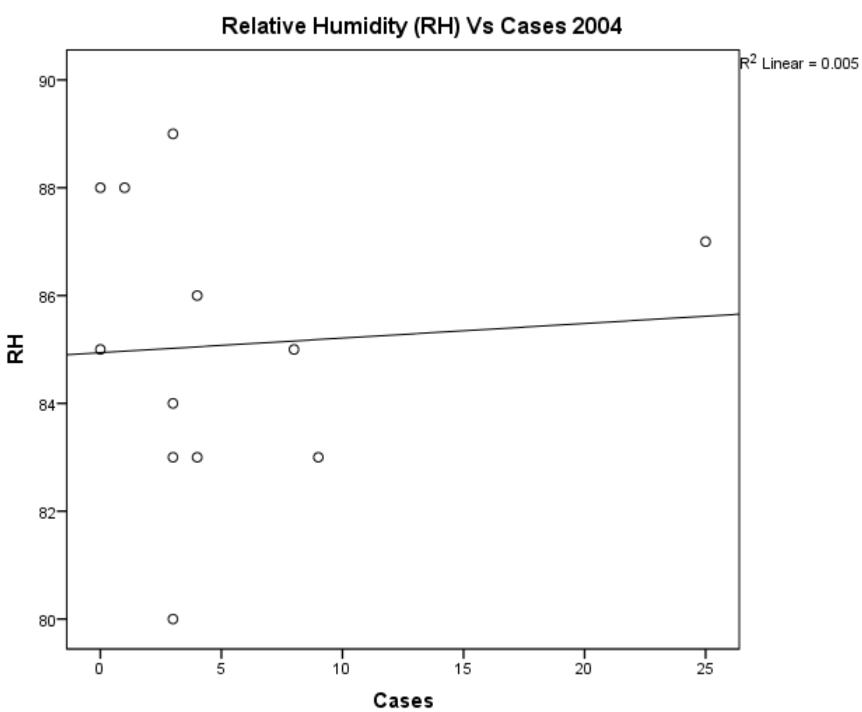


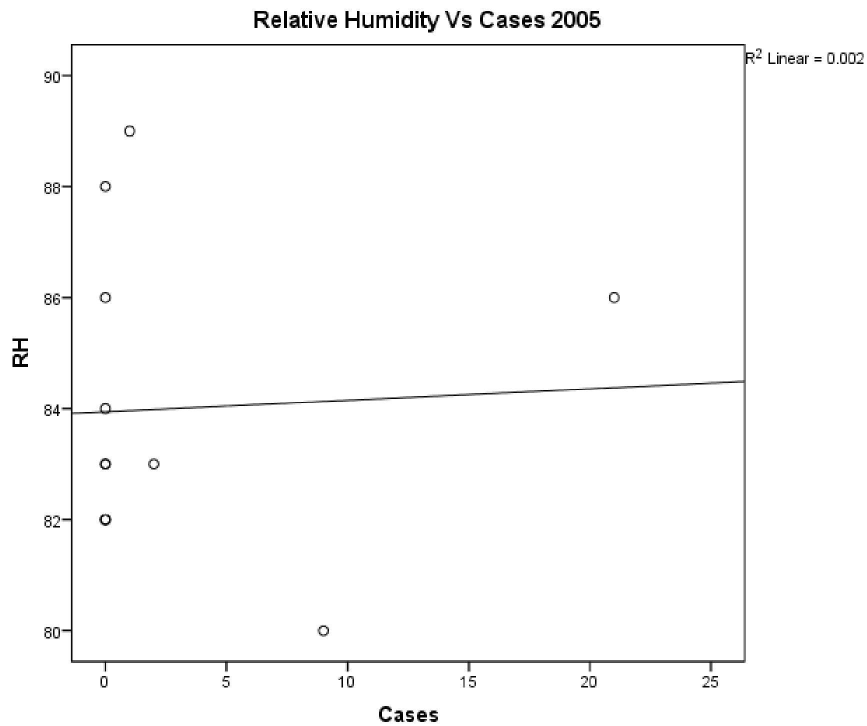
Graphs showing correlation between wind speed against number of cases for each year

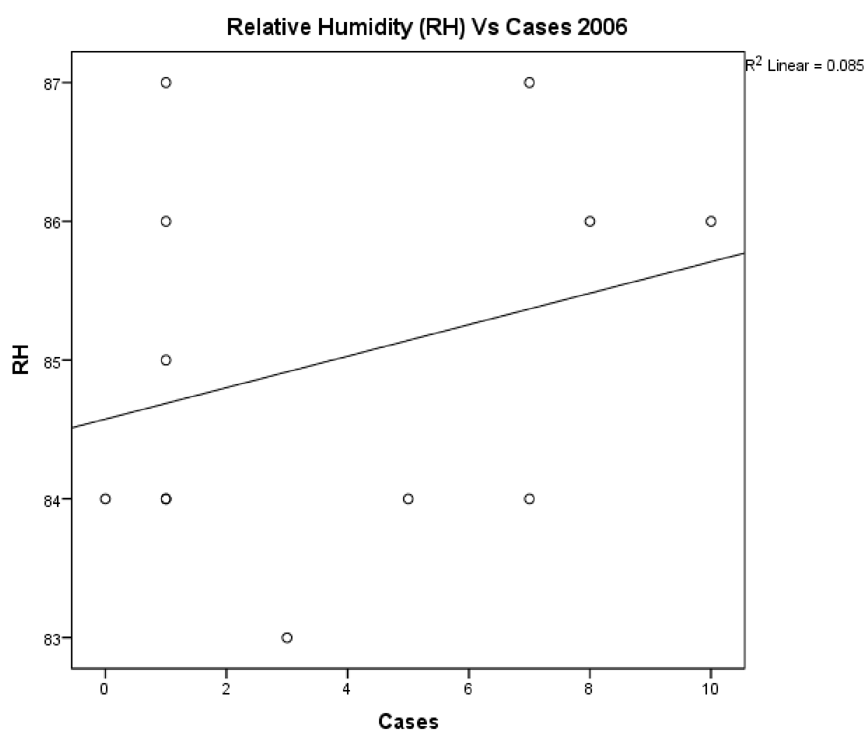
CORRELATION BETWEEN RELATIVE HUMIDITY AGAINST NUMBER OF CASES FOR EACH YEAR FROM 2002 TO 2007

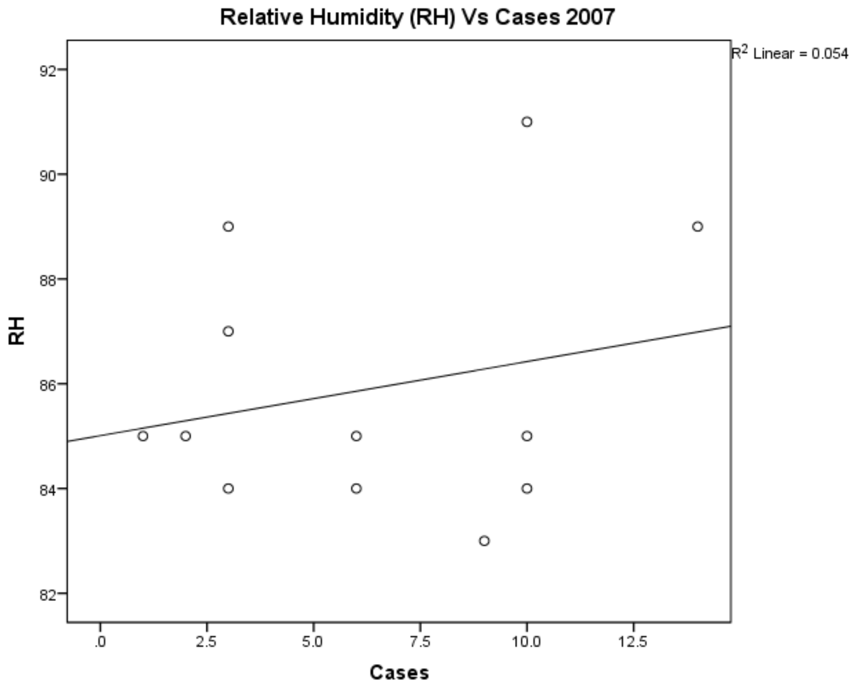



Relative Humidity (RH) Vs Cases 2003
R^2 Linear = 0.042
RH
90
88
86
84
82
80
0
5
10
15
20
Cases

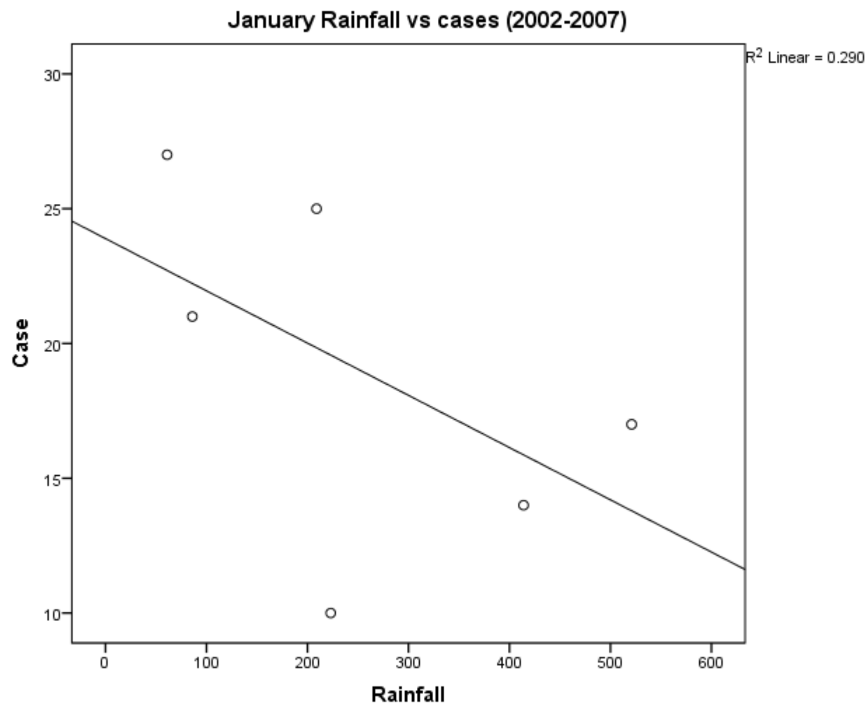

Relative Humidity (RH) Vs Cases 2004
R^2 Linear = 0.005
RH
90
88
86
84
82
80
0
5
10
15
20
25
Cases

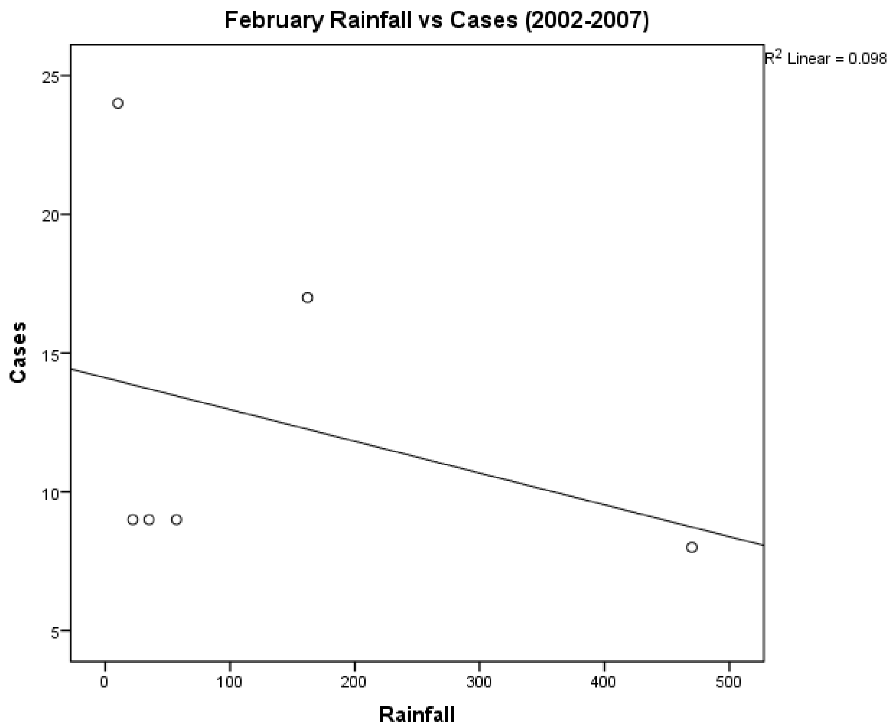

Relative Humidity Vs Cases 2005
R2 Linear = 0.002
RH
90
88
86
84
82
80
0
5
10
15
20
25
Cases

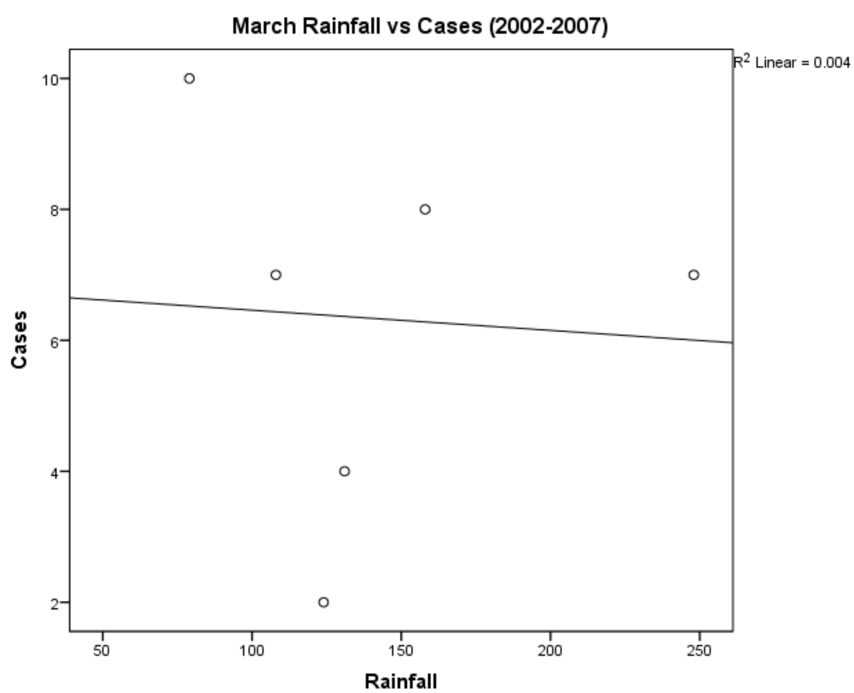

Relative Humidity (RH) Vs Cases 2006
R2 Linear = 0.085
RH
87
86
85
84
83
0
2
4
6
8
10
Cases

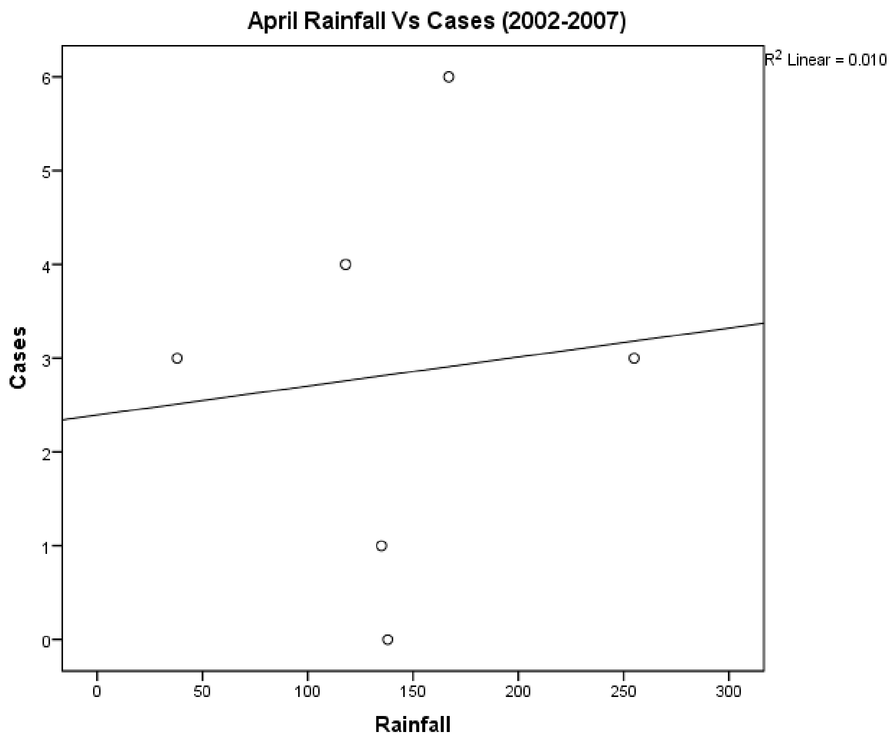

Relative Humidity (RH) Vs Cases 2007
R2 Linear = 0.054
92
90
88
86
84
82
RH
.0
2.5
5.0
7.5
10.0
12.5
Cases

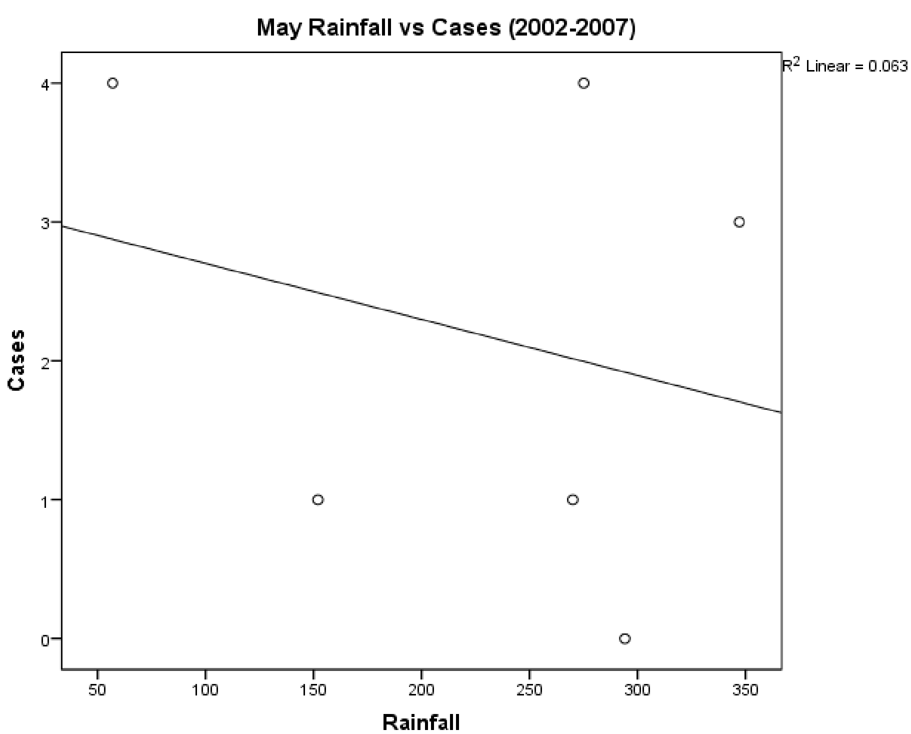
CORRELATION BETWEEN MONTHLY RAINFALL VOLUMES AGAINST NUMBER OF CASES FOR A PERIOD OF SIX YEARS FROM 2002 UNTIL 2007

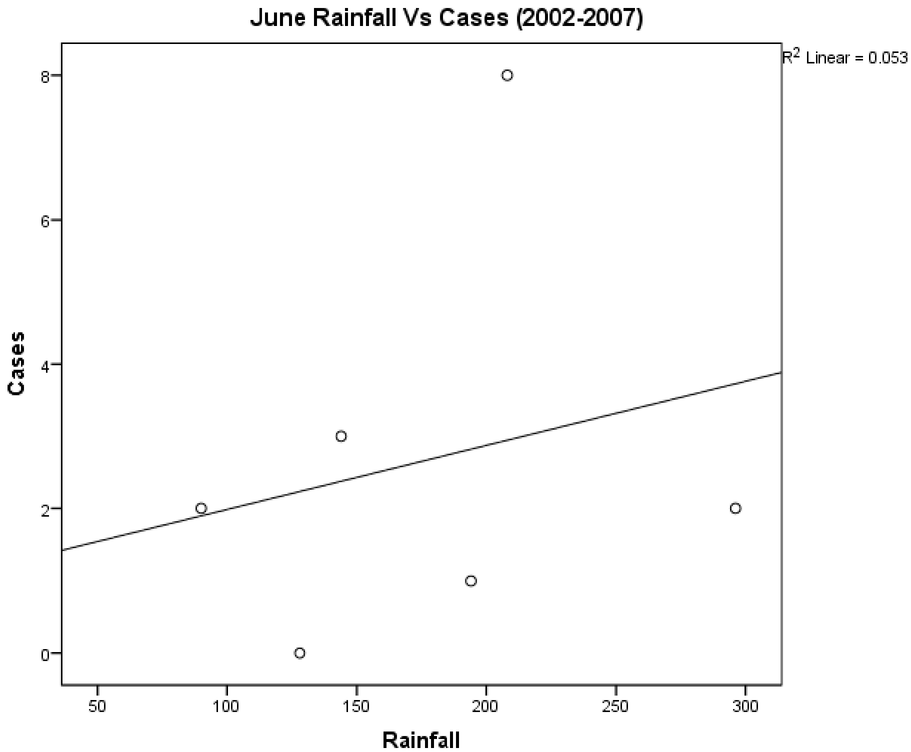


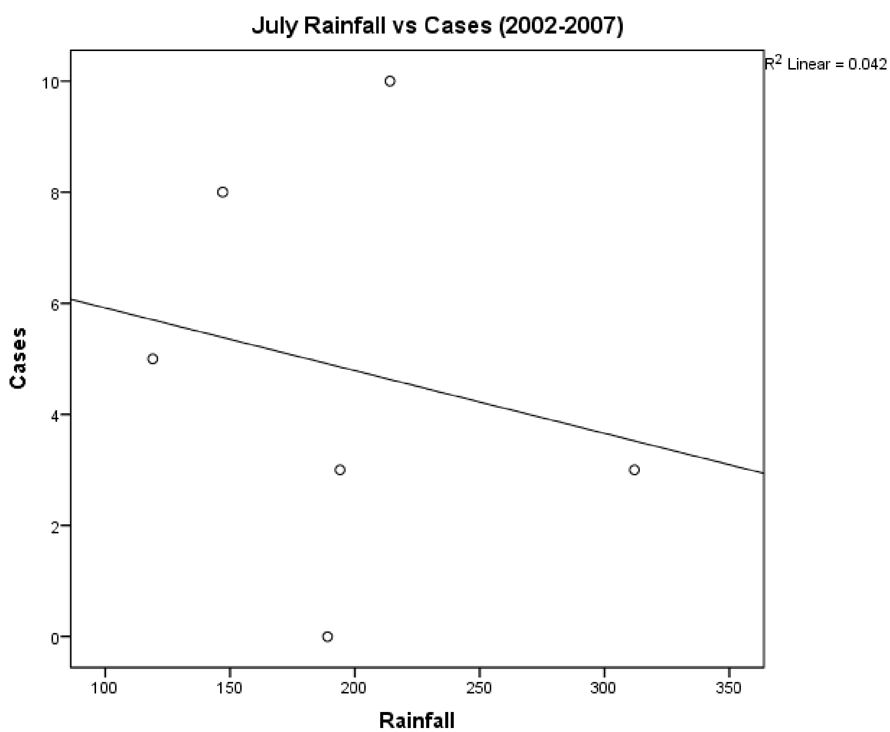

February Rainfall vs Cases (2002-2007)
R2 Linear = 0.098
Cases
Rainfall
25
20
15
10
5
0
100
200
300
400
500

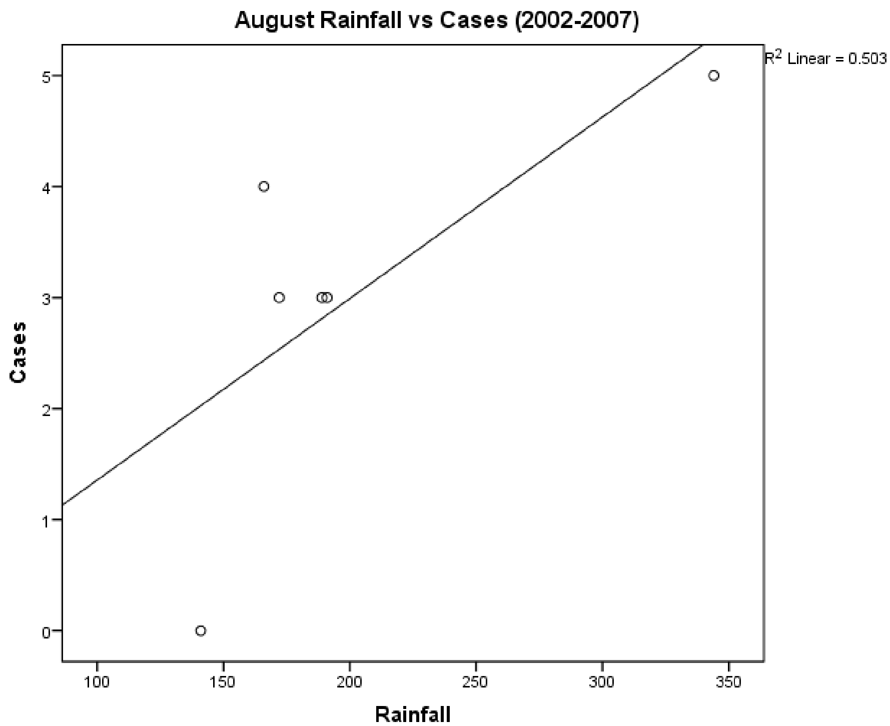

March Rainfall vs Cases (2002-2007)
R2 Linear = 0.004
Cases
Rainfall
10
8
6
4
2
50
100
150
200
250

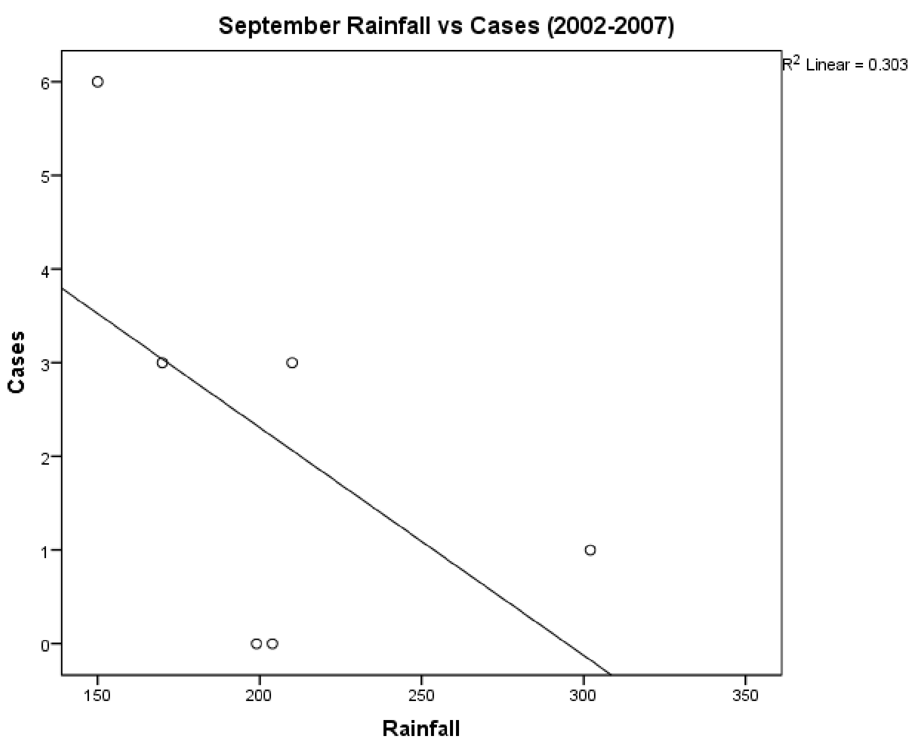

April Rainfall Vs Cases (2002-2007)
R2 Linear = 0.010
Cases
6
5
4
3
2
1
0
0
50
100
150
200
250
300
Rainfall

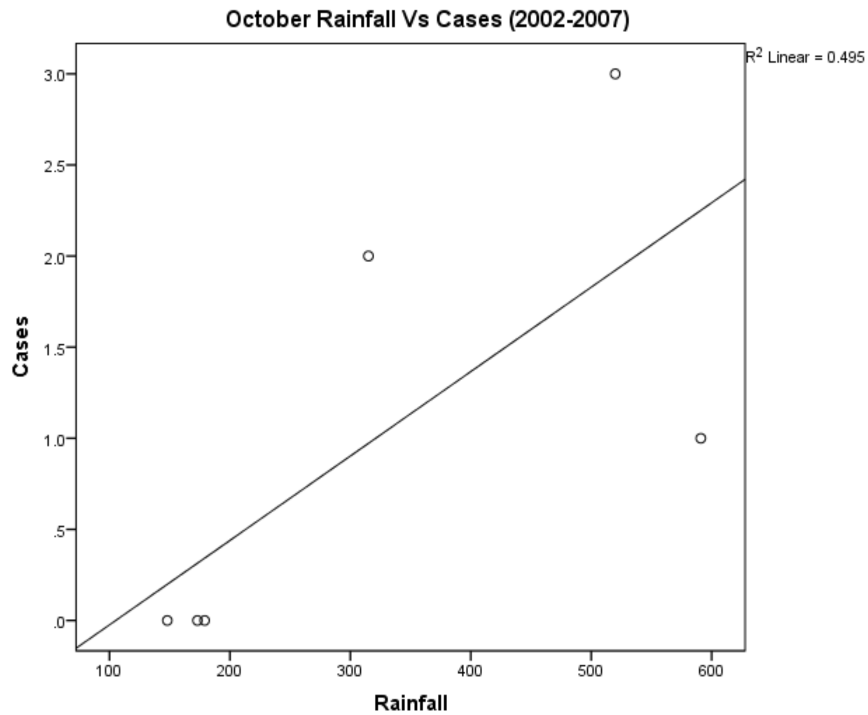

May Rainfall vs Cases (2002-2007)
R2 Linear = 0.063
Cases
4
3
2
1
0
50
100
150
200
250
300
350
Rainfall

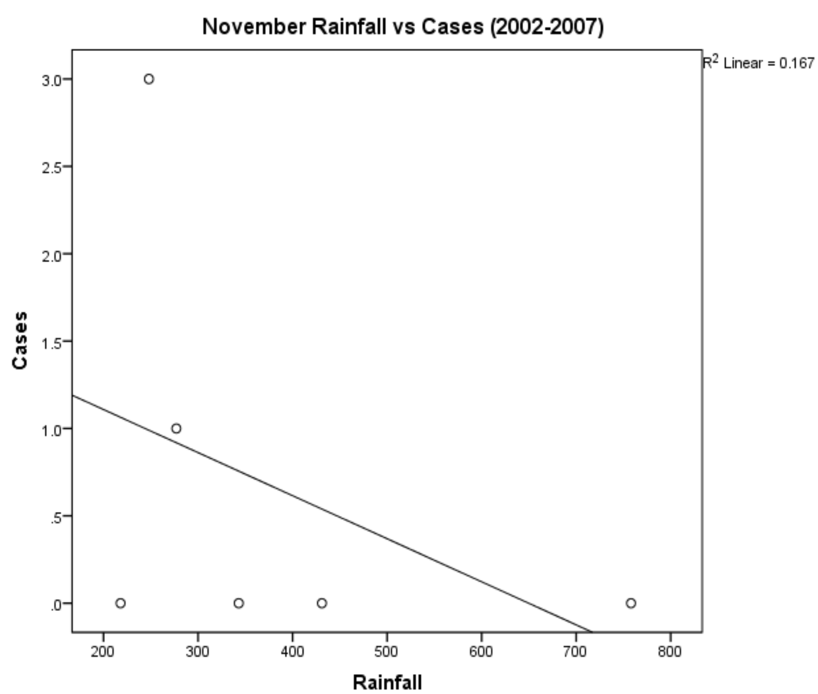

June Rainfall Vs Cases (2002-2007)
R² Linear = 0.053
Cases
Rainfall
0
2
4
6
8
50
100
150
200
250
300

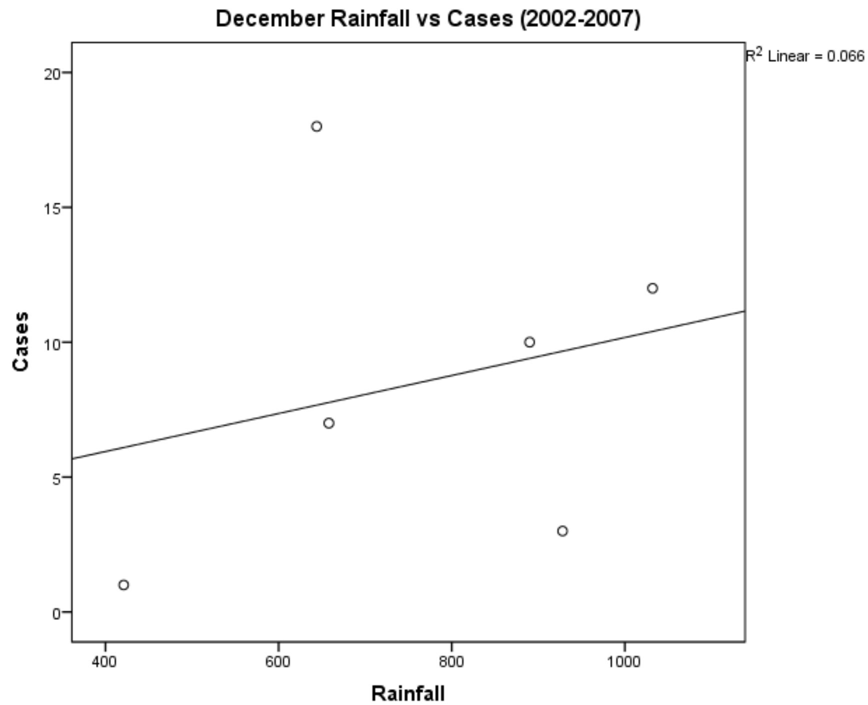

July Rainfall vs Cases (2002-2007)
R² Linear = 0.042
Cases
Rainfall
0
2
4
6
8
10
100
150
200
250
300
350


August Rainfall vs Cases (2002-2007)
R2 Linear = 0.503
Cases
Rainfall
0
1
2
3
4
5
100
150
200
250
300
350

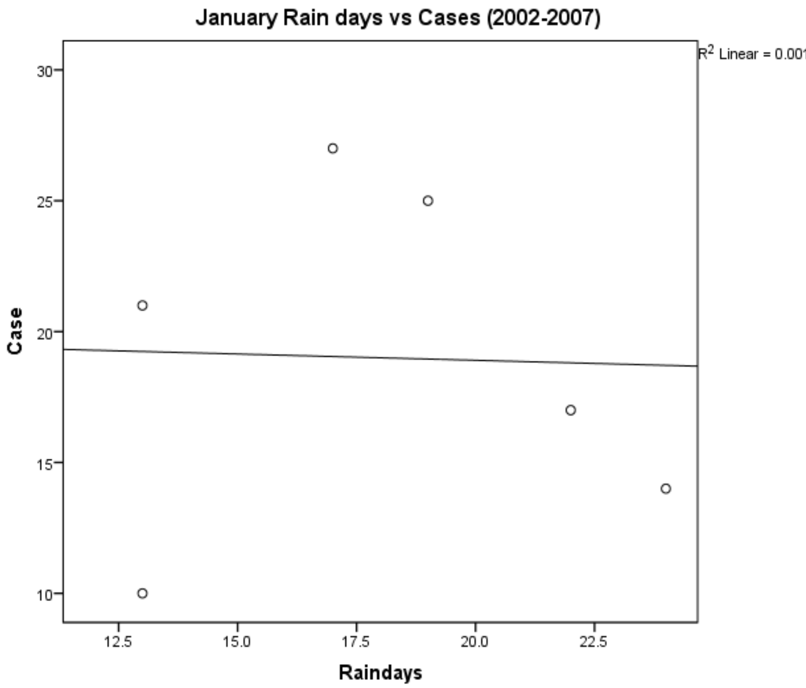

September Rainfall vs Cases (2002-2007)
R2 Linear = 0.303
Cases
Rainfall
0
1
2
3
4
5
6
150
200
250
300
350

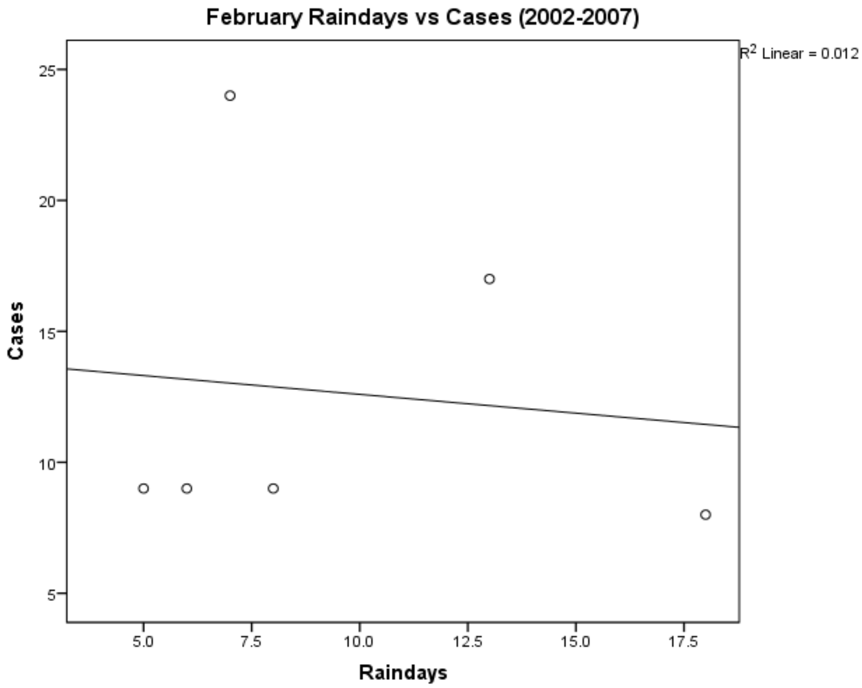

October Rainfall Vs Cases (2002-2007)
R2 Linear = 0.495
3.0
2.5
2.0
1.5
1.0
.5
.0
Cases
100
200
300
400
500
600
Rainfall

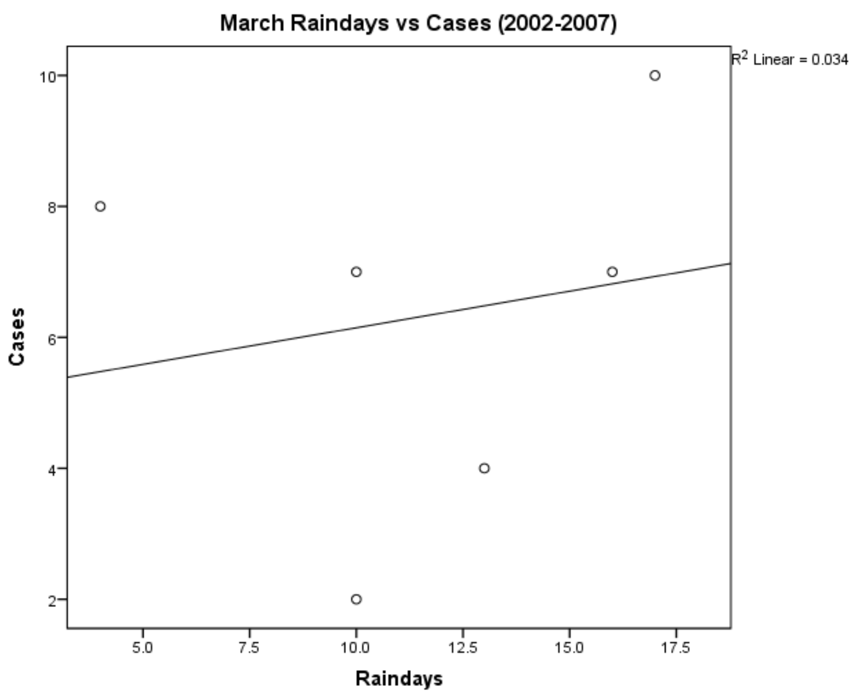

November Rainfall vs Cases (2002-2007)
R2 Linear = 0.167
3.0
2.5
2.0
1.5
1.0
.5
.0
Cases
200
300
400
500
600
700
800
Rainfall

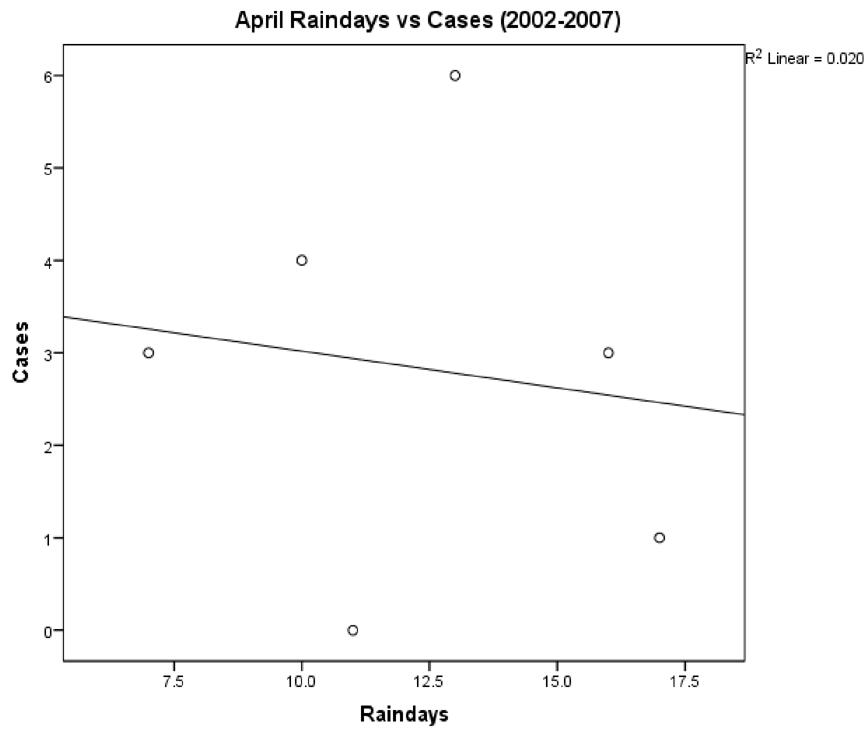

December Rainfall vs Cases (2002-2007)
R2 Linear = 0.066
Cases
20
15
10
5
0
400
600
800
1000
Rainfall

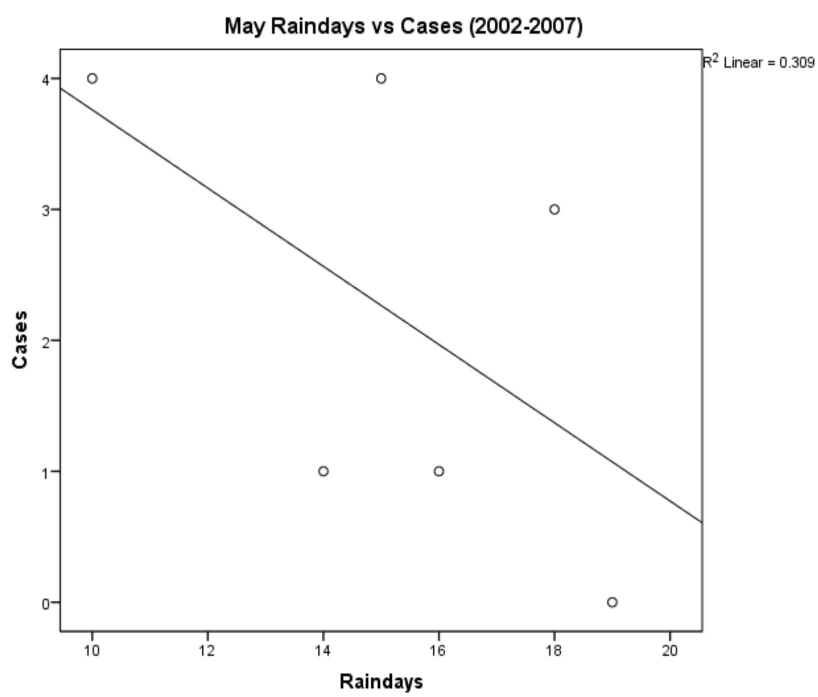
CORRELATION BETWEEN MONTHLY NUMBERS OF RAINY DAYS AGAINST NUMBER OF CASES FOR A PERIOD OF SIX YEARS FROM 2002 UNTIL 2007

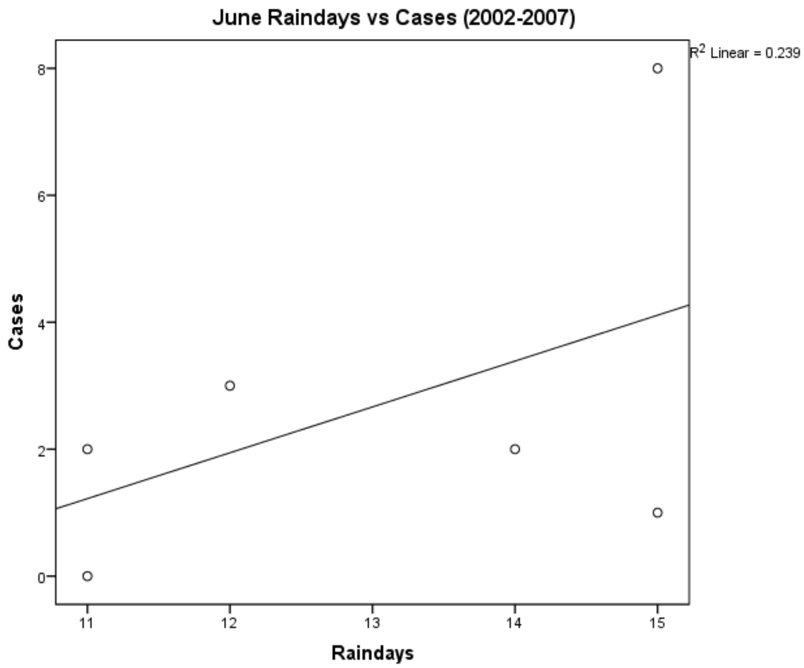


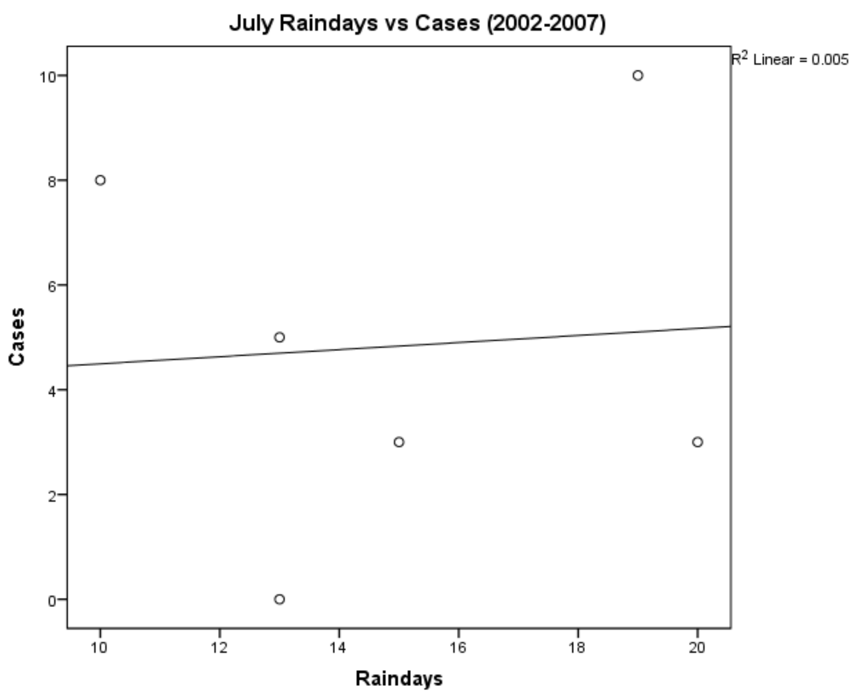

February Raindays vs Cases (2002-2007)
R2 Linear = 0.012
Cases
25
20
15
10
5
5.0
7.5
10.0
12.5
15.0
17.5
Raindays

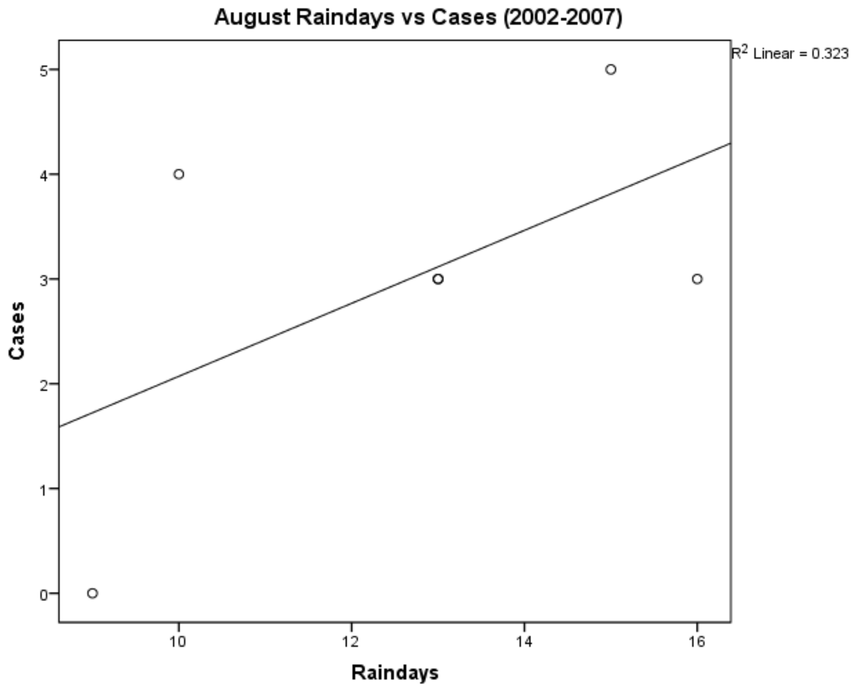

March Raindays vs Cases (2002-2007)
R2 Linear = 0.034
Cases
10
8
6
4
2
5.0
7.5
10.0
12.5
15.0
17.5
Raindays

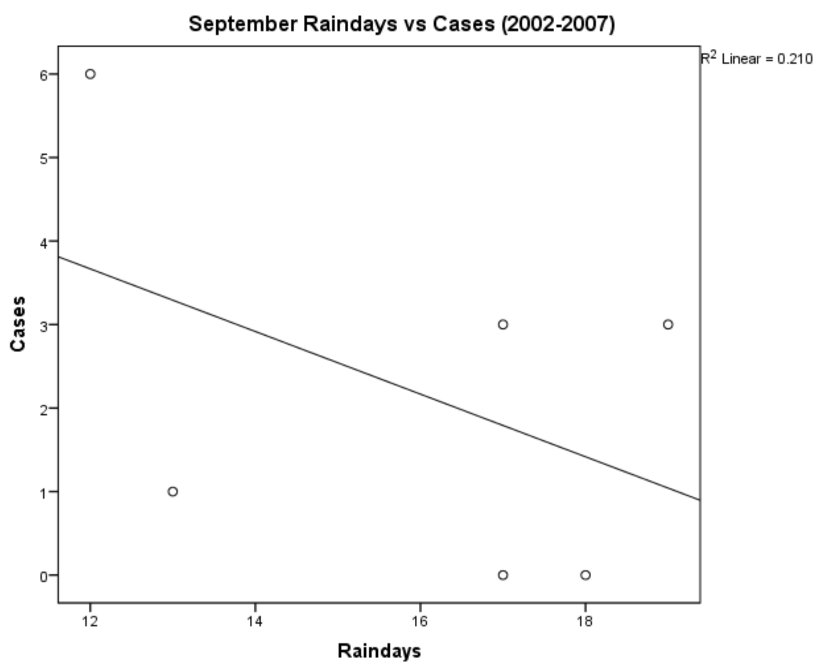

April Raindays vs Cases (2002-2007)
R^2 Linear = 0.020
Cases
0
1
2
3
4
5
6
Raindays
7.5
10.0
12.5
15.0
17.5

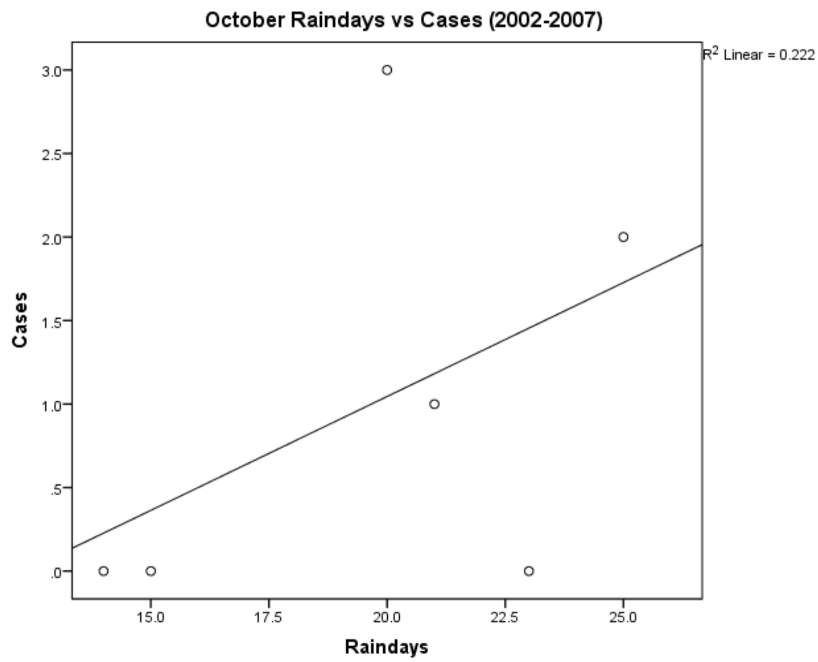

May Raindays vs Cases (2002-2007)
R^2 Linear = 0.309
Cases
0
1
2
3
4
Raindays
10
12
14
16
18
20

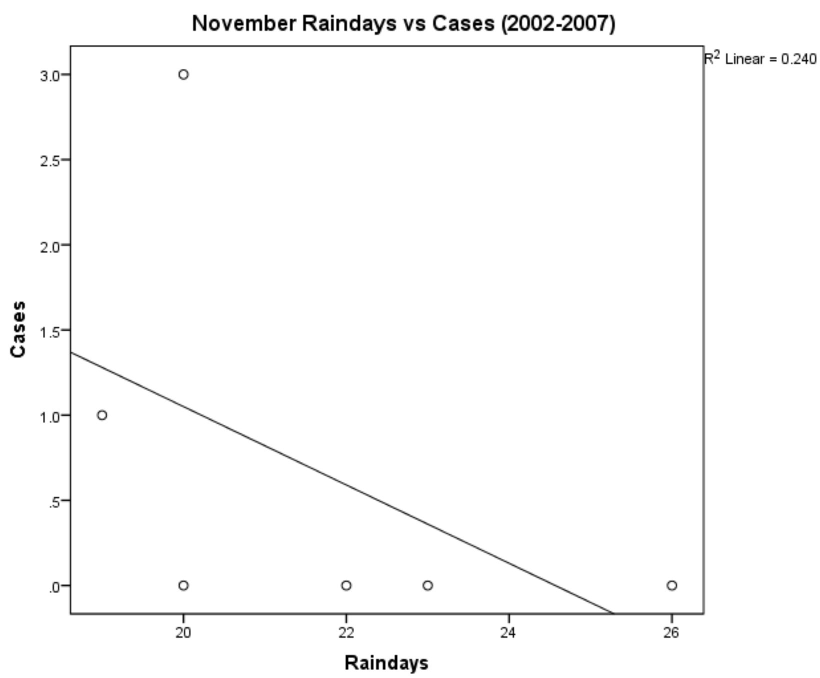

June Raindays vs Cases (2002-2007)
R2 Linear = 0.239
Cases
Raindays
0
2
4
6
8
11
12
13
14
15

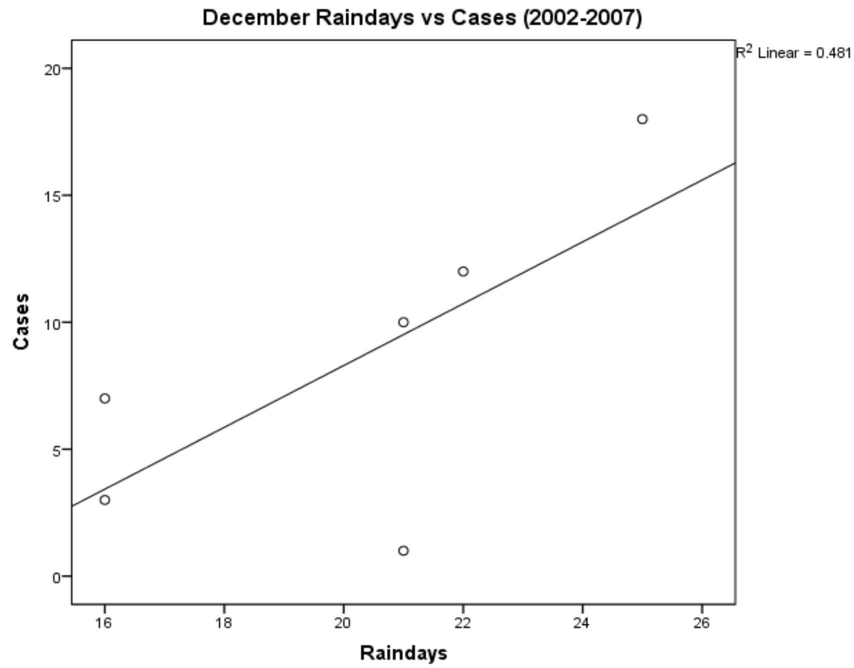

July Raindays vs Cases (2002-2007)
R2 Linear = 0.005
Cases
Raindays
0
2
4
6
8
10
10
12
14
16
18
20


August Raindays vs Cases (2002-2007)
R2 Linear = 0.323
Cases
Raindays
0
1
2
3
4
5
10
12
14
16


September Raindays vs Cases (2002-2007)
R2 Linear = 0.210
Cases
Raindays
0
1
2
3
4
5
6
12
14
16
18

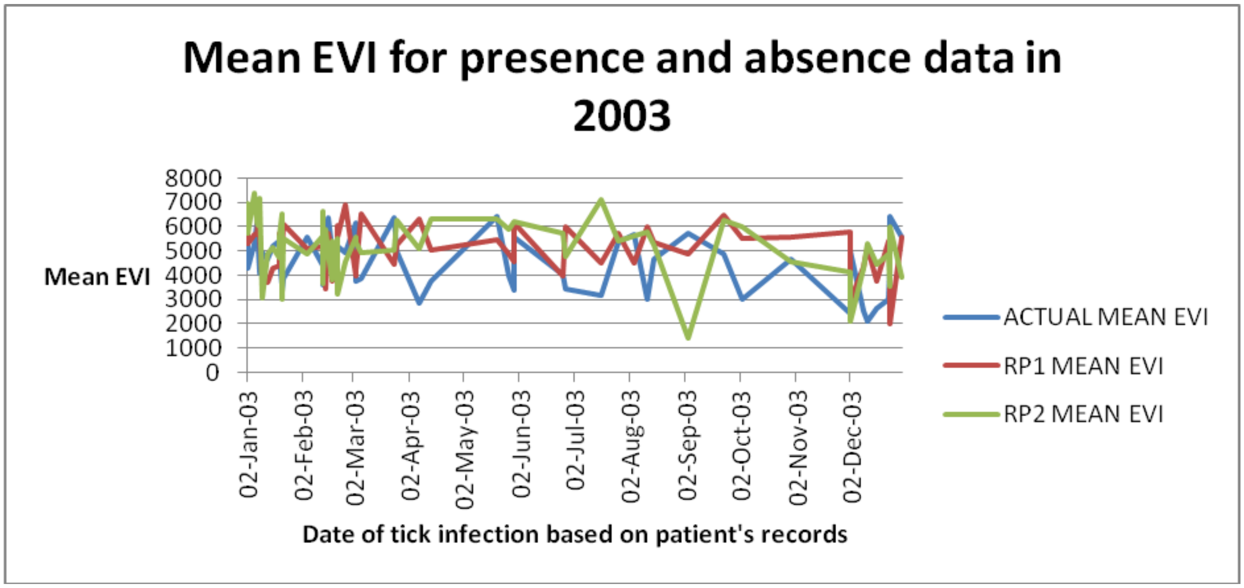

October Raindays vs Cases (2002-2007)
R^2 Linear = 0.222
Cases
Raindays
3.0
2.5
2.0
1.5
1.0
.5
.0
15.0
17.5
20.0
22.5
25.0

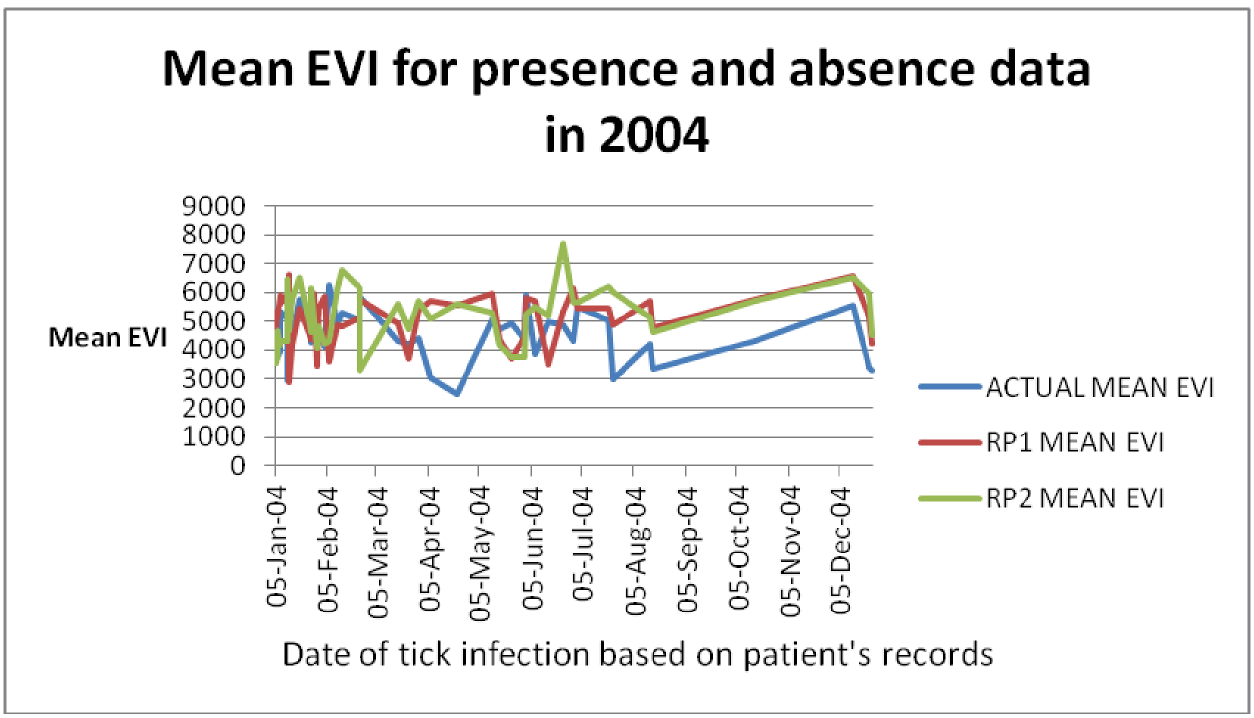

November Raindays vs Cases (2002-2007)
R^2 Linear = 0.240
Cases
Raindays
3.0
2.5
2.0
1.5
1.0
.5
.0
20
22
24
26

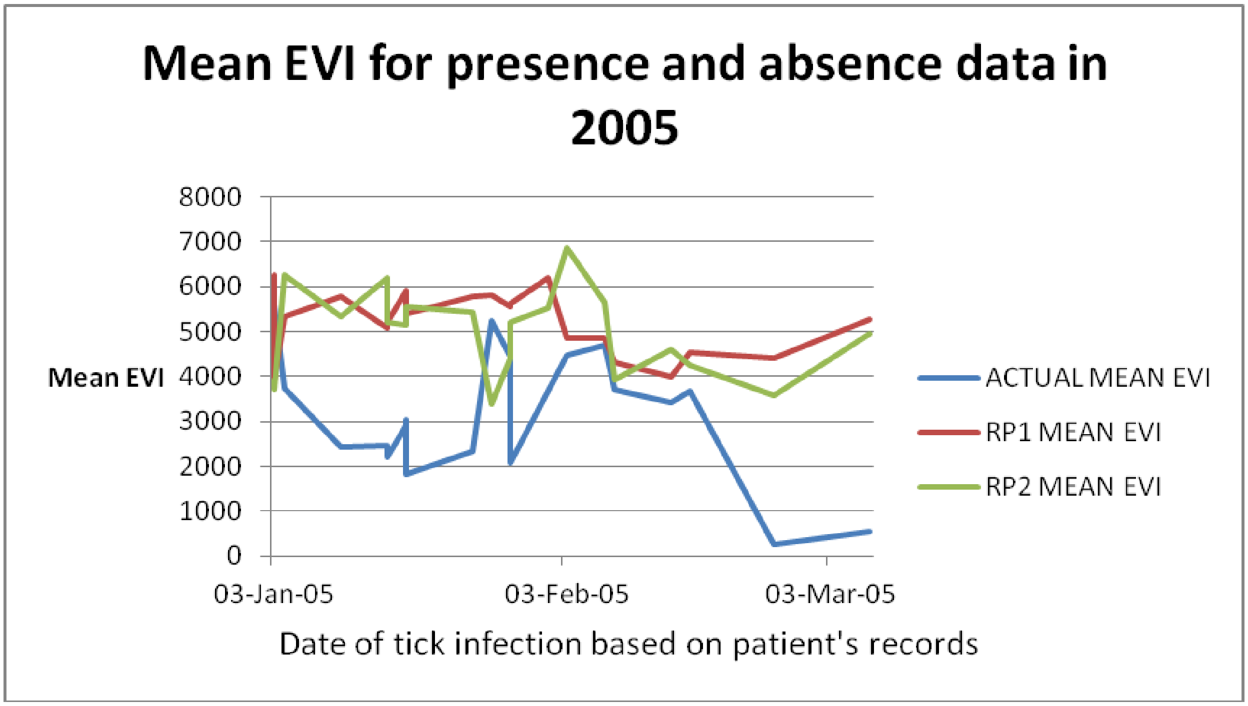


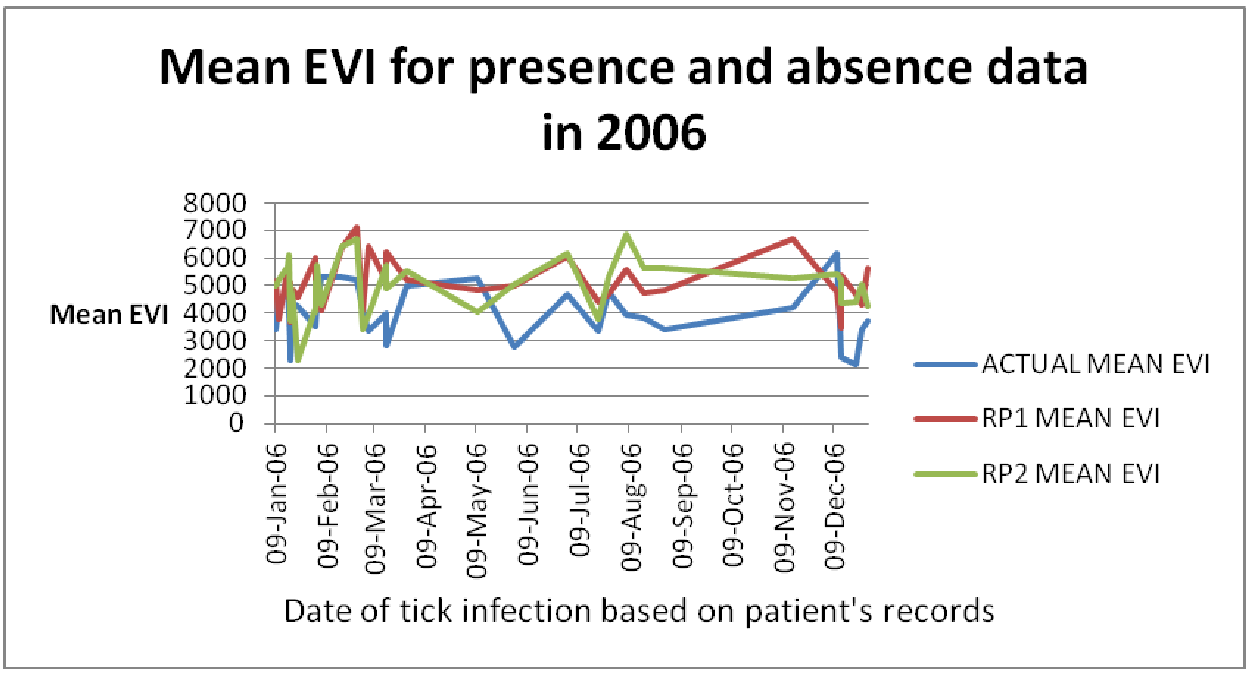
Graphs showing correlation between monthly numbers of rainy days against number of cases for a period of six years from 2002 until 2007

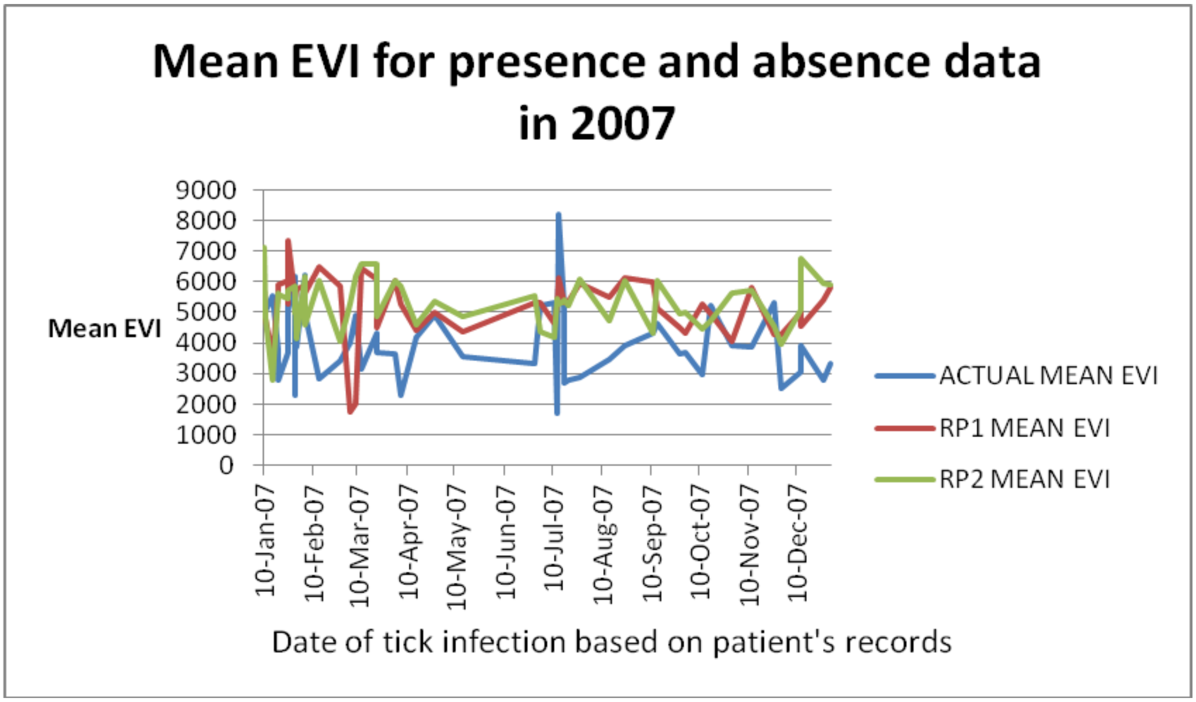
GRAPHS SHOWING THE CORRELATION BETWEEN MEAN EVI AGAINST PRESENCE (ACTUAL REPORTED CASES) AND ABSENCE (2 SETS OF RANDOM POINTS) FOR THE YEAR 2003 UNTIL 2007





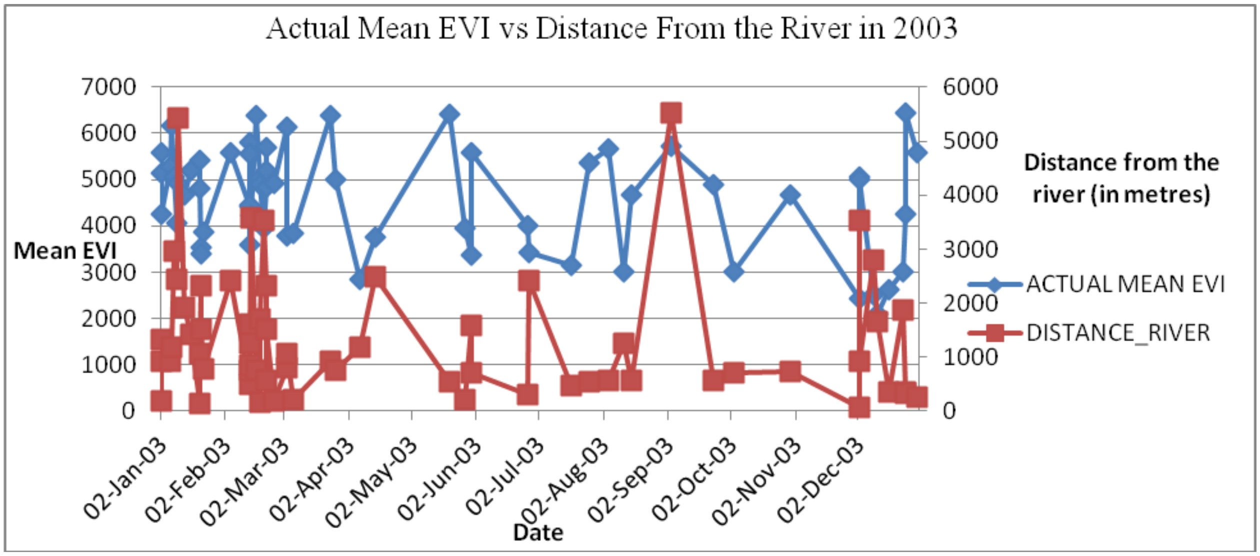

Mean EVI for presence and absence data in 2005
8000
7000
6000
5000
4000
3000
2000
1000
0
Mean EVI
03-Jan-05
03-Feb-05
03-Mar-05
Date of tick infection based on patient's records
ACTUAL MEAN EVI
RP1 MEAN EVI
RP2 MEAN EVI

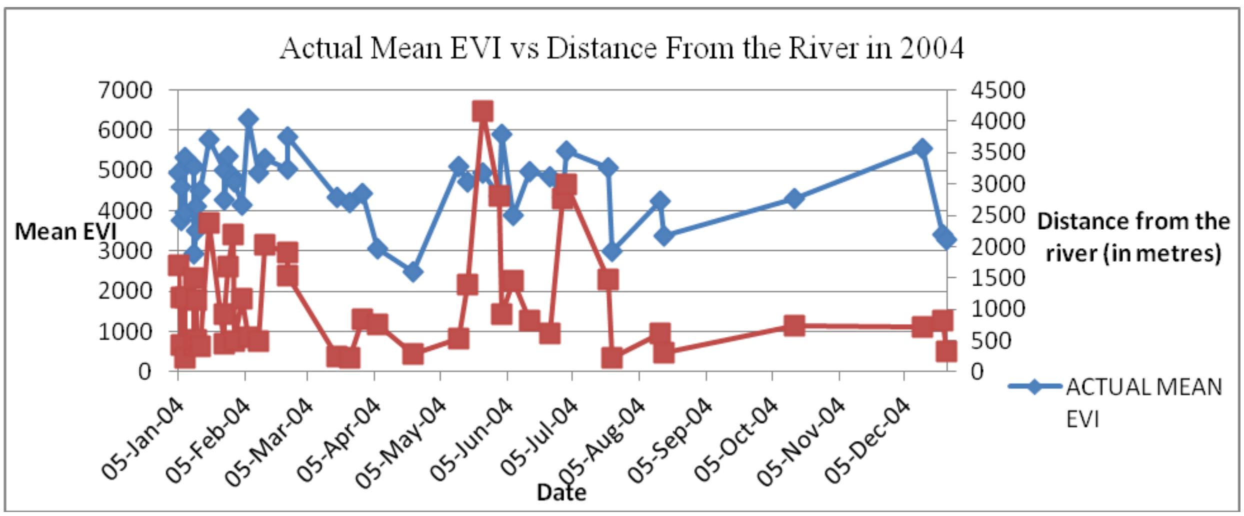

Mean EVI for presence and absence data in 2006
8000
7000
6000
5000
4000
3000
2000
1000
0
Mean EVI
09-Jan-06
09-Feb-06
09-Mar-06
09-Apr-06
09-May-06
09-Jun-06
09-Jul-06
09-Aug-06
09-Sep-06
09-Oct-06
09-Nov-06
09-Dec-06
Date of tick infection based on patient's records
ACTUAL MEAN EVI
RP1 MEAN EVI
RP2 MEAN EVI

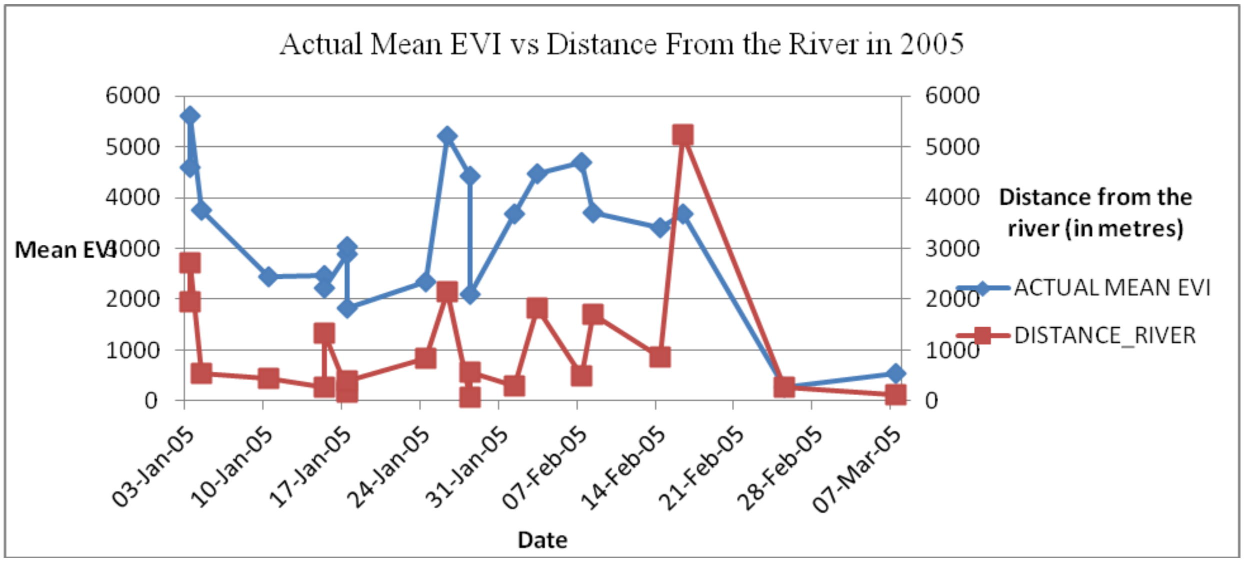


Mean EVI for presence and absence data in 2007
Mean EVI
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
10-Jan-07
10-Feb-07
10-Mar-07
10-Apr-07
10-May-07
10-Jun-07
10-Jul-07
10-Aug-07
10-Sep-07
10-Oct-07
10-Nov-07
10-Dec-07
Date of tick infection based on patient's records
ACTUAL MEAN EVI
RP1 MEAN EVI
RP2 MEAN EVI

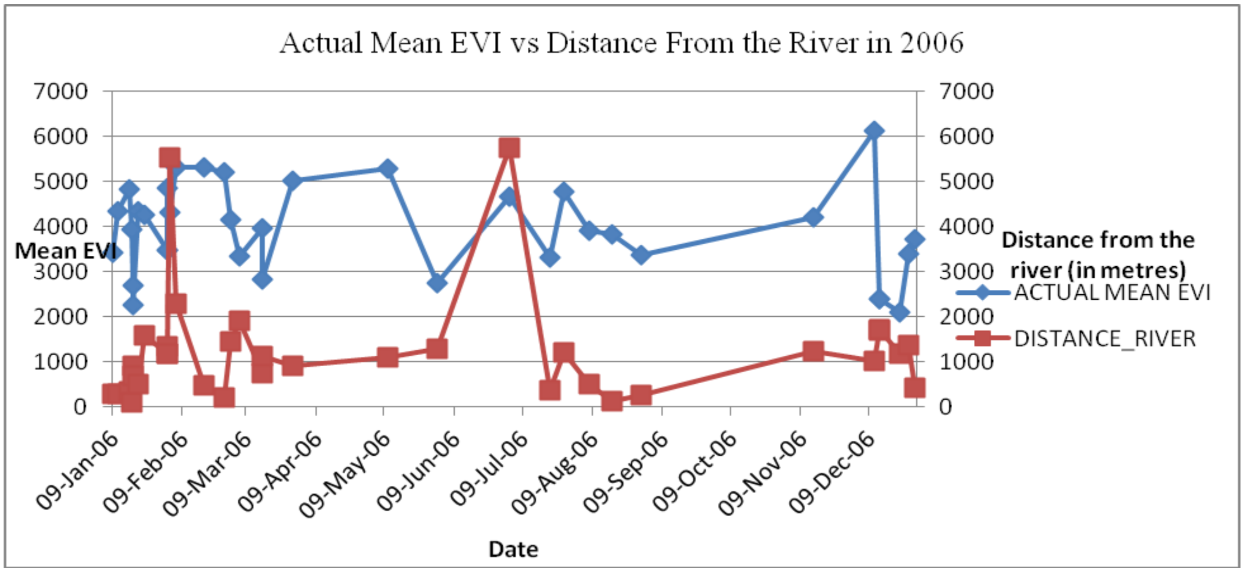
GRAPHS SHOWING RELATIONSHIP BETWEEN MEAN EVI AND DISTANCE FROM THE RIVER FOR THE YEAR 2003 TO 2007

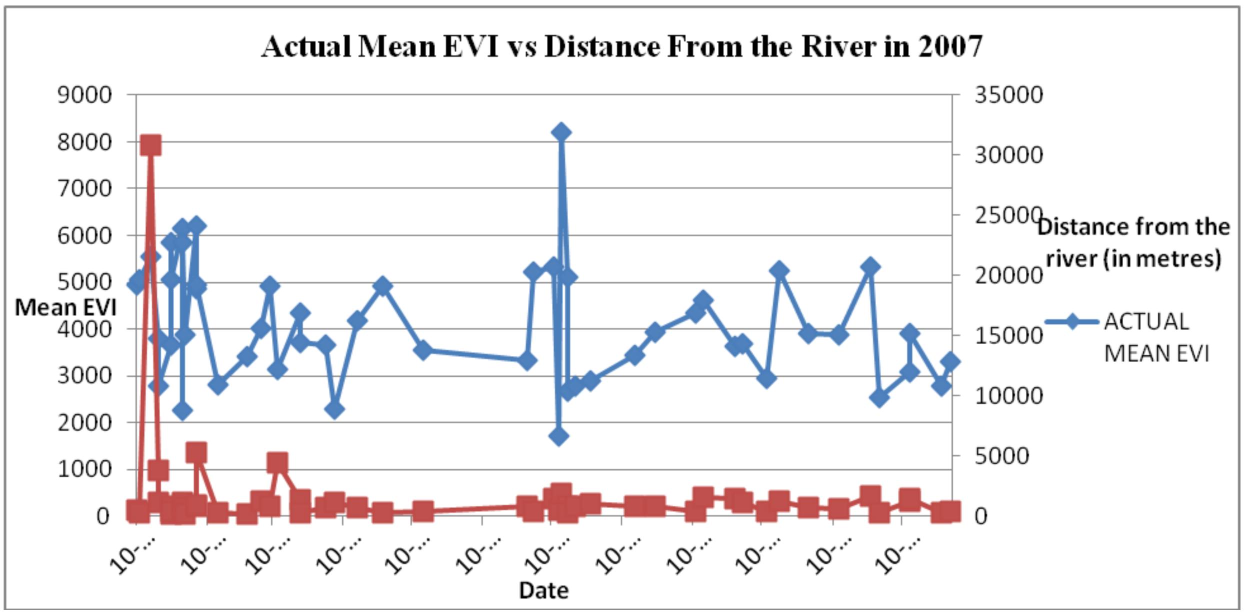






Actual Mean EVI vs Distance From the River in 2005
6000
5000
4000
3000
2000
1000
0
Mean EVI
Distance from the river (in metres)
ACTUAL MEAN EVI
DISTANCE_RIVER
03-Jan-05
10-Jan-05
17-Jan-05
24-Jan-05
31-Jan-05
07-Feb-05
14-Feb-05
21-Feb-05
28-Feb-05
07-Mar-05
Date





Graphs showing MEAN EVI between presence (actual reported cases) and distance from the rivers

Based from the graphs above, there seems to be a kind of pattern between mean EVI and distance from the river. To prove the relationship, regression analysis is being done in SPSS.

LOGISTIC REGRESSION ANALYSIS TO IDENTIFY RELATIONSHIP BETWEEN MEAN EVI AND DISTANCE FROM THE RIVER FOR THE YEAR 2003 TO 2007

The result of linear regression analysis is summarised in table below. The null hypothesis is to prove that there is no relationship between mean EVI and distance from the river.

Regression Analysis

H_0 = no relationship between Mean EVI and distance from the river

Year	P value
2003	0.748 (>0.05)
2004	0.012(<0.05)
2005	0.040 (<0.05)
2006	0.668 (>0.05)
2007	0.173 (>0.05)

However, only for 2004 and 2005 shows significant result, where the p value is less than 0.05.

Detail of regression analysis is as below:

Regression

[DataSet2] C:\Users\14645822\Documents\MEETING 42\FINAL FOR MEETING\2003.sav

Descriptive Statistics

	Mean	Std. Deviation	N
ACTUAL MEAN EVI	4576.9266	1105.34879	62
DISTANCE_RIVER	1374.2846	1181.27157	62

Correlations

		ACTUAL MEAN EVI	DISTANCE_RIVER
Pearson Correlation	ACTUAL MEAN EVI	1.000	.042
	DISTANCE_RIVER	.042	1.000
Sig. (1-tailed)	ACTUAL MEAN EVI	.	.374
	DISTANCE_RIVER	.374	.
N	ACTUAL MEAN EVI	62	62
	DISTANCE_RIVER	62	62

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	DISTANCE_RIVER[b]	.	Enter

a. Dependent Variable: ACTUAL MEAN EVI

b. All requested variables entered.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.042[a]	.002	-.015	1113.55643

a)

a. Predictors: (Constant), DISTANCE_RIVER

b. Dependent Variable: ACTUAL MEAN EVI

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	129077.871	1	129077.871	.104	.748[b]
	Residual	74400475.209	60	1240007.920		
	Total	74529553.080	61			

a. Dependent Variable: ACTUAL MEAN EVI

b. Predictors: (Constant), DISTANCE_RIVER

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4523.410	217.977		20.752	.000
	DISTANCE_RIVER	.039	.121	.042	.323	.748

b)

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	4087.392	4959.428
	DISTANCE_RIVER	-.202	.280

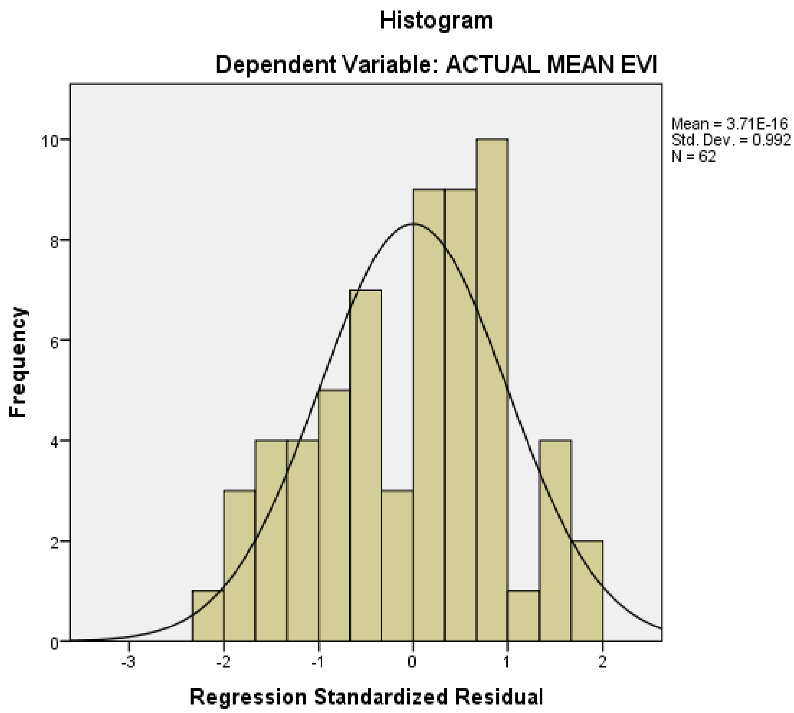
a. Dependent Variable: ACTUAL MEAN EVI

Residuals Statistics[a]

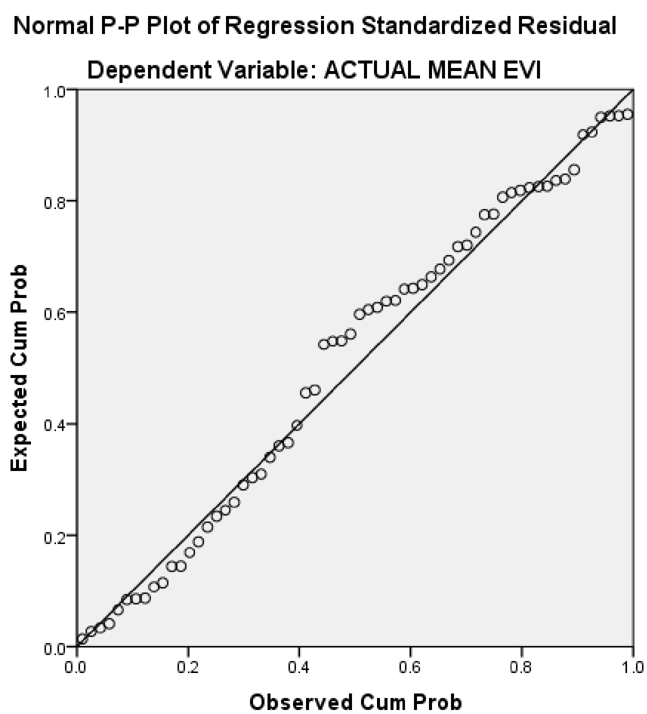
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4526.0605	4738.4009	4576.9266	46.00033	62
Residual	-2458.03003	1886.34534	.00000	1104.39120	62
Std. Predicted Value	-1.106	3.510	.000	1.000	62
Std. Residual	-2.207	1.694	.000	.992	62

a. Dependent Variable: ACTUAL MEAN EVI

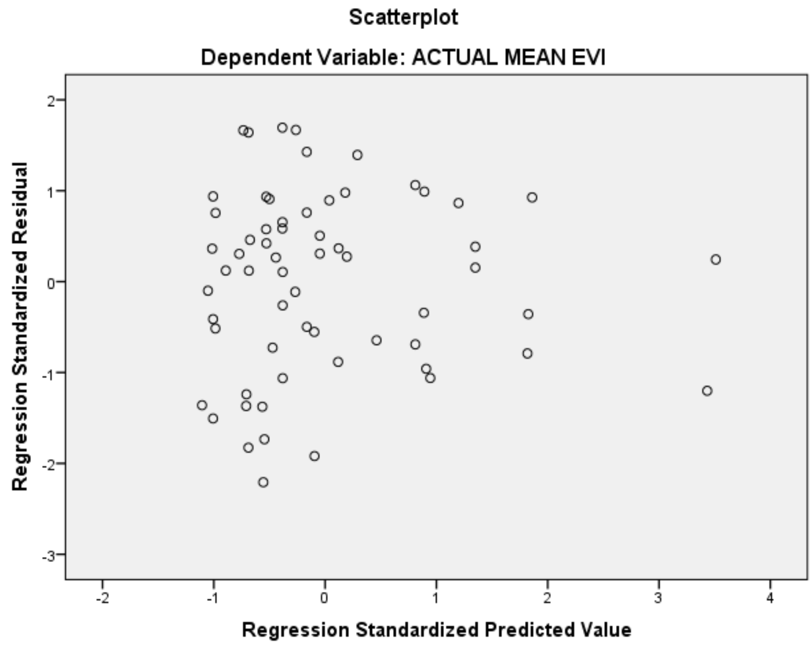
Charts



Histogram
Dependent Variable: ACTUAL MEAN EVI
Mean = 3.71E-16
Std. Dev. = 0.992
N = 62
Frequency
0
2
4
6
8
10
-3
-2
-1
0
1
2
Regression Standardized Residual



Normal P-P Plot of Regression Standardized Residual
Dependent Variable: ACTUAL MEAN EVI
Expected Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0
Observed Cum Prob



Regression

[DataSet3] C:\Users\14645822\Documents\MEETING 42\FINAL FOR MEETING\2004.sav

Descriptive Statistics

	Mean	Std. Deviation	N
ACTUAL	4523.3235	866.42404	45
V5	1123.1564	879.28224	45

Correlations

		ACTUAL	V5
Pearson Correlation	ACTUAL	1.000	.371
	V5	.371	1.000
Sig. (1-tailed)	ACTUAL	.	.006
	V5	.006	.
N	ACTUAL	45	45
	V5	45	45

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	V5[b]	.	Enter

a. Dependent Variable: ACTUAL

b. All requested variables entered.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.371[a]	.138	.117	813.93265

a. Predictors: (Constant), V5

b. Dependent Variable: ACTUAL

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4543473.383	1	4543473.383	6.858	.012[b]
	Residual	28486913.581	43	662486.362		
	Total	33030386.964	44			

a. Dependent Variable: ACTUAL

b. Predictors: (Constant), V5

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4112.855	198.214		20.750	.000
	V5	.365	.140	.371	2.619	.012

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	3713.119	4512.591
	V5	.084	.647

a. Dependent Variable: ACTUAL

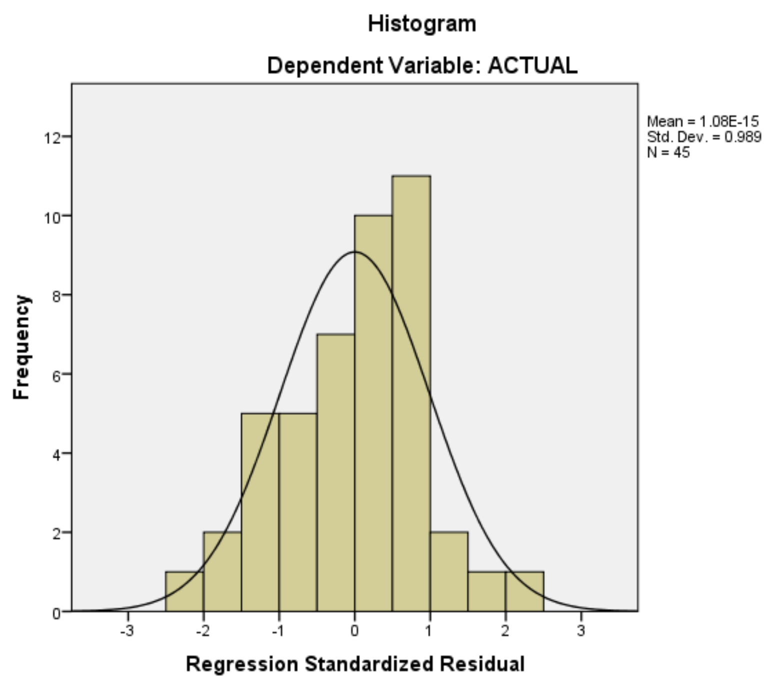
Residuals Statistics[a]

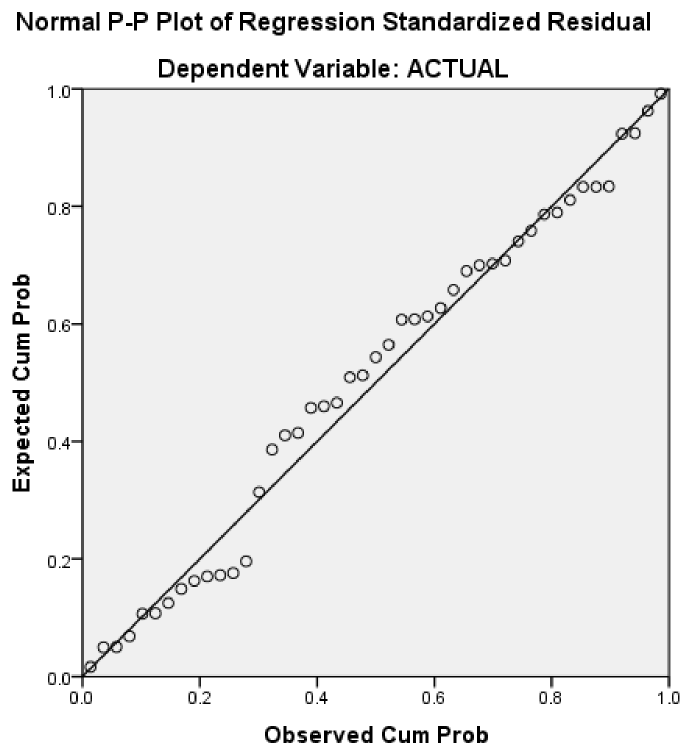
	Minimum	Maximum	Mean	Std. Deviation	N

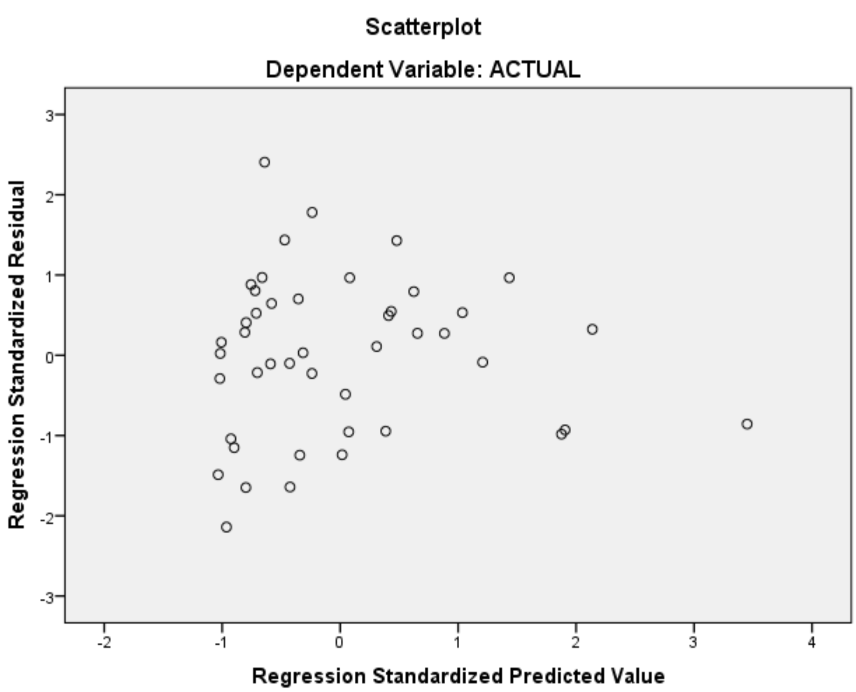
Predicted Value	4191.0527	5631.9673	4523.3235	321.34212	45
Residual	-1742.32544	1958.77979	.00000	804.63026	45
Std. Predicted Value	-1.034	3.450	.000	1.000	45
Std. Residual	-2.141	2.407	.000	.989	45

a. Dependent Variable: ACTUAL

Charts




Normal P-P Plot of Regression Standardized Residual
Dependent Variable: ACTUAL
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0


Scatterplot
Dependent Variable: ACTUAL
Regression Standardized Residual
Regression Standardized Predicted Value
-3
-2
-1
0
1
2
3
4

Regression

[DataSet4] C:\Users\14645822\Documents\MEETING 42\FINAL FOR MEETING\2005.sav

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	V4[b]	.	Enter

a. Dependent Variable: ACTUAL

b. All requested variables entered.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.451[a]	.203	.161	1293.27019	.203	4.847

Model Summary[b]

Model	Change Statistics		
	df1	df2	Sig. F Change
1	1[a]	19	.040

a. Predictors: (Constant), V4

b. Dependent Variable: ACTUAL

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8106496.648	1	8106496.648	4.847	.040[b]
	Residual	31778408.024	19	1672547.791		
	Total	39884904.672	20			

a. Dependent Variable: ACTUAL

b. Predictors: (Constant), V4

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2654.438	378.187		7.019	.000
	V4	.518	.235	.451	2.202	.040

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	1862.884	3445.992
	V4	.026	1.011

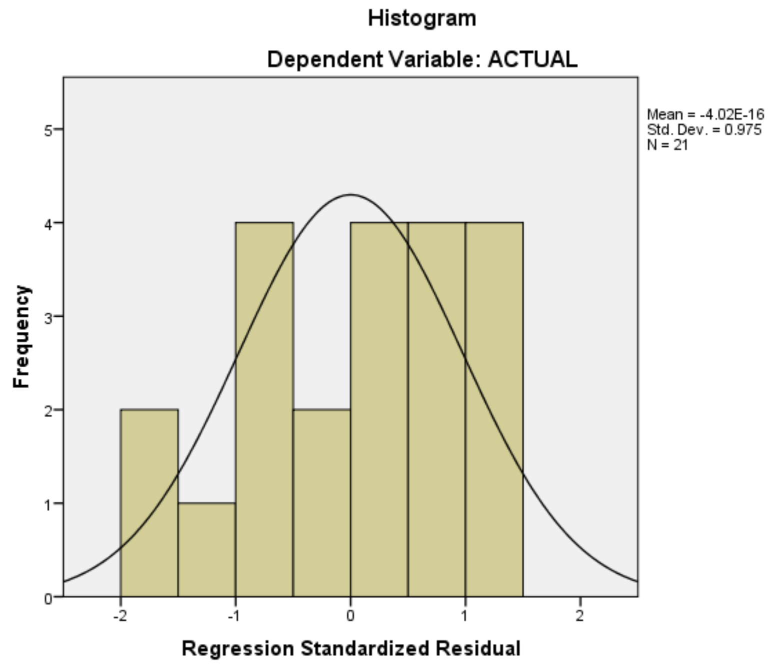
a. Dependent Variable: ACTUAL

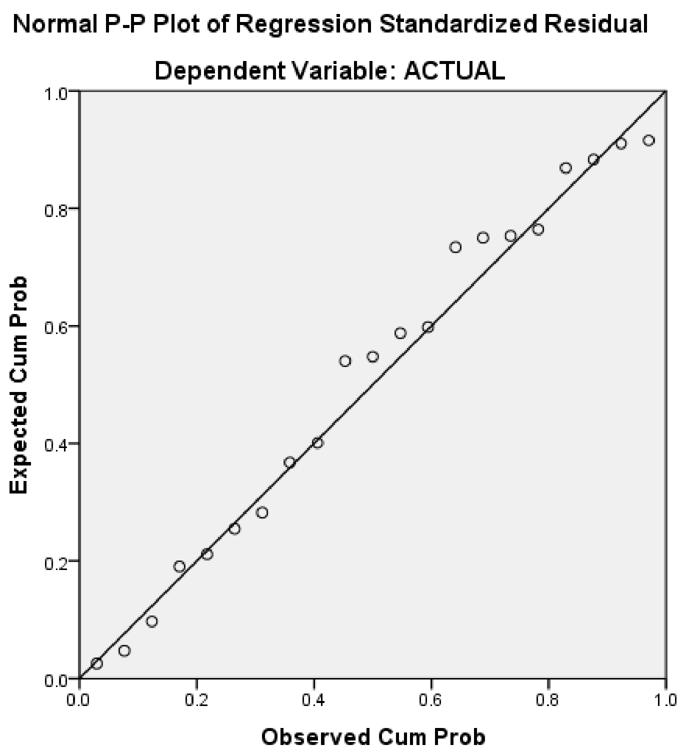
Residuals Statistics[a]

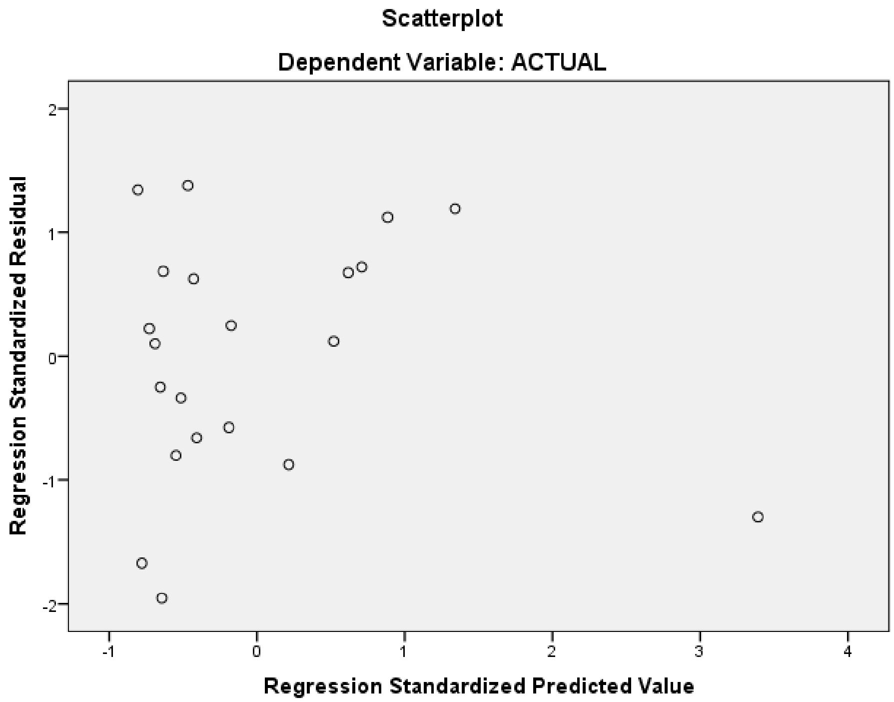
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2694.7637	5367.7246	3208.6856	636.65126	21
Residual	-2526.27661	1781.18030	.00000	1260.52386	21
Std. Predicted Value	-.807	3.391	.000	1.000	21
Std. Residual	-1.953	1.377	.000	.975	21

a. Dependent Variable: ACTUAL

Charts


Histogram
Dependent Variable: ACTUAL
Mean = -4.02E-16
Std. Dev. = 0.975
N = 21
Frequency
0
1
2
3
4
5
-2
-1
0
1
2
Regression Standardized Residual


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: ACTUAL
Expected Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0
Observed Cum Prob



Regression

[DataSet5] C:\Users\14645822\Documents\MEETING 42\FINAL FOR MEETING\2006.sav

Descriptive Statistics

	Mean	Std. Deviation	N
ACTUAL	3939.7422	998.58912	35
V4	1212.6553	1237.05818	35

Correlations

		ACTUAL	V4
Pearson Correlation	ACTUAL	1.000	.075
	V4	.075	1.000
Sig. (1-tailed)	ACTUAL	.	.334
	V4	.334	.
N	ACTUAL	35	35
	V4	35	35

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method

1	V4[b]	.	Enter

a. Dependent Variable: ACTUAL

b. All requested variables entered.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.075[a]	.006	-.024	1010.73725

a. Predictors: (Constant), V4

b. Dependent Variable: ACTUAL

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	191665.434	1	191665.434	.188	.668[b]
	Residual	33712462.702	33	1021589.779		
	Total	33904128.136	34			

a. Dependent Variable: ACTUAL

b. Predictors: (Constant), V4

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3866.142	240.959		16.045	.000
	V4	.061	.140	.075	.433	.668

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	3375.907	4356.377
	V4	-.224	.346

a. Dependent Variable: ACTUAL

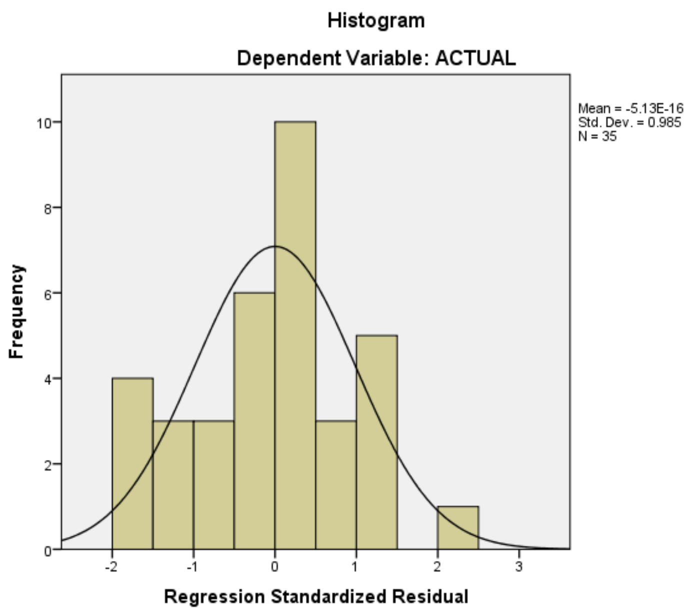
Residuals Statistics[a]

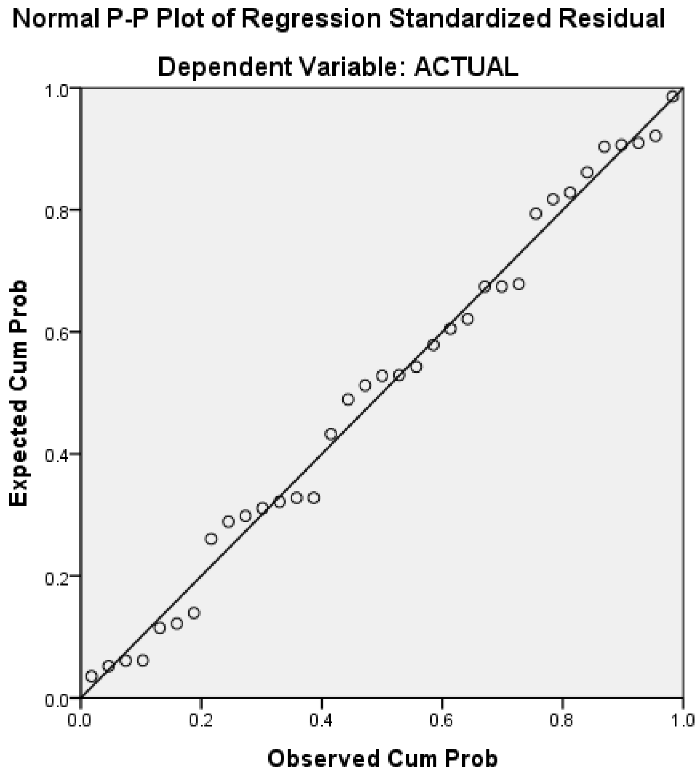
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3872.3508	4215.2769	3939.7422	75.08141	35

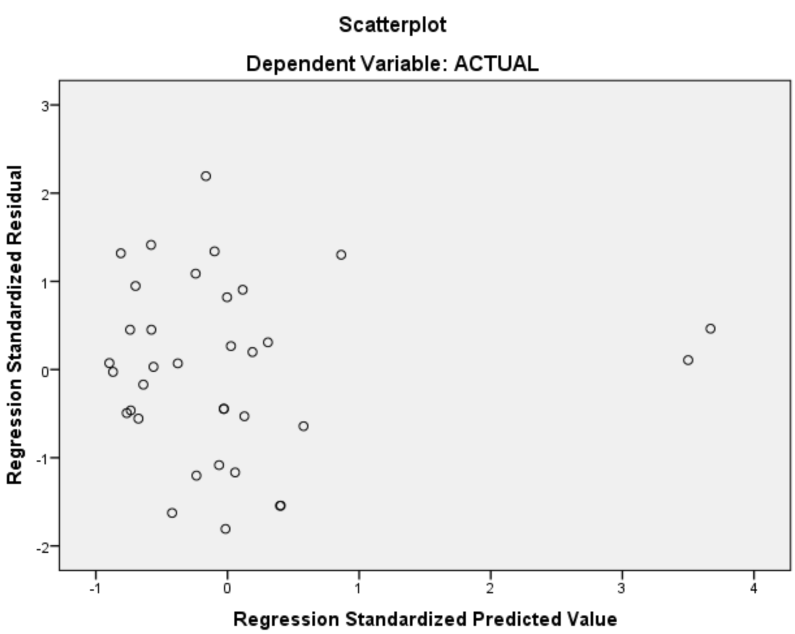
Residual	-1826.60815	2216.28638	.00000	995.76253	35
Std. Predicted Value	-.898	3.670	.000	1.000	35
Std. Residual	-1.807	2.193	.000	.985	35

a. Dependent Variable: ACTUAL

Charts




Normal P-P Plot of Regression Standardized Residual
Dependent Variable: ACTUAL
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0


Scatterplot
Dependent Variable: ACTUAL
Regression Standardized Residual
Regression Standardized Predicted Value
-2
-1
0
1
2
3
4

Regression

[DataSet1] C:\Users\14645822\Documents\MEETING 42\FINAL FOR MEETING\2007.sav

Descriptive Statistics

	Mean	Std. Deviation	N
ACTUAL	4109.2959	1237.46828	52
V4	1562.4253	4260.52455	52

Correlations

		ACTUAL	V4
Pearson Correlation	ACTUAL	1.000	.192
	V4	.192	1.000
Sig. (1-tailed)	ACTUAL	.	.086
	V4	.086	.
N	ACTUAL	52	52
	V4	52	52

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	V4[b]	.	Enter

a. Dependent Variable: ACTUAL

b. All requested variables entered.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.192[a]	.037	.018	1226.54073

a. Predictors: (Constant), V4

b. Dependent Variable: ACTUAL

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2877607.291	1	2877607.291	1.913	.173[b]
	Residual	75220108.168	50	1504402.163		
	Total	78097715.459	51			

a. Dependent Variable: ACTUAL

b. Predictors: (Constant), V4

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4022.186	181.378		22.176	.000
	V4	.056	.040	.192	1.383	.173

Coefficients[a]

Model		95.0% Confidence Interval for B	
		Lower Bound	Upper Bound
1	(Constant)	3657.878	4386.494
	V4	-.025	.137

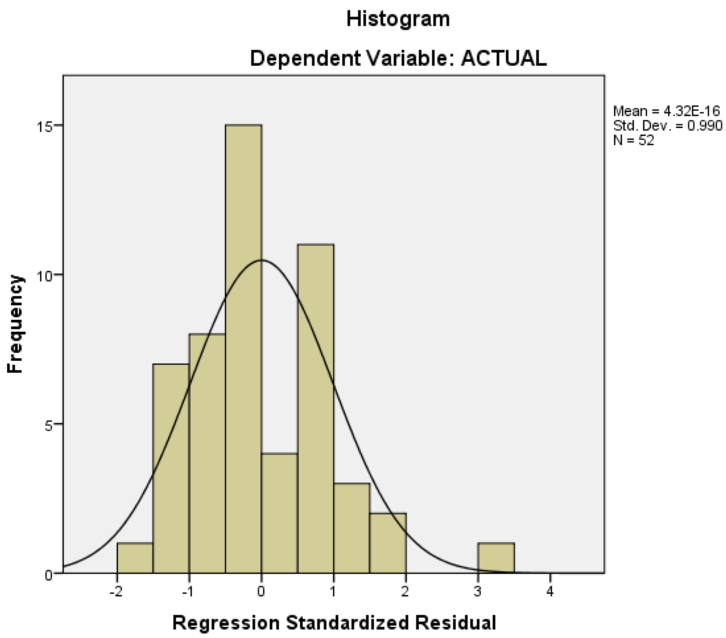
a. Dependent Variable: ACTUAL

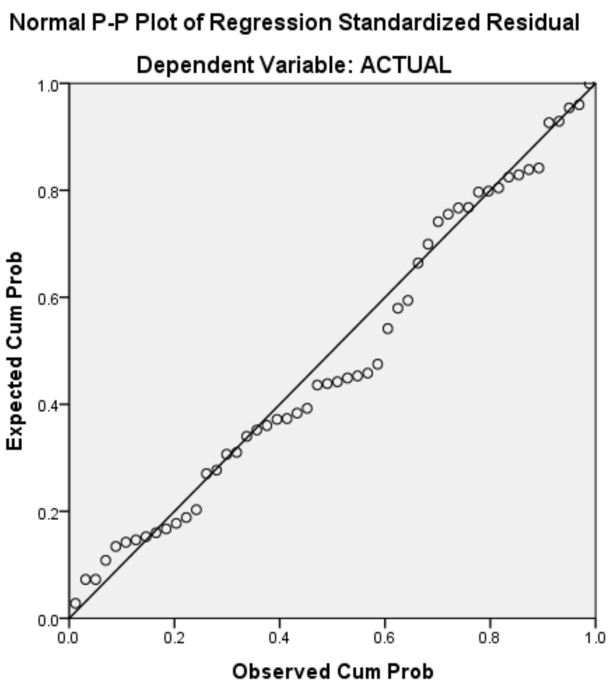
Residuals Statistics[a]

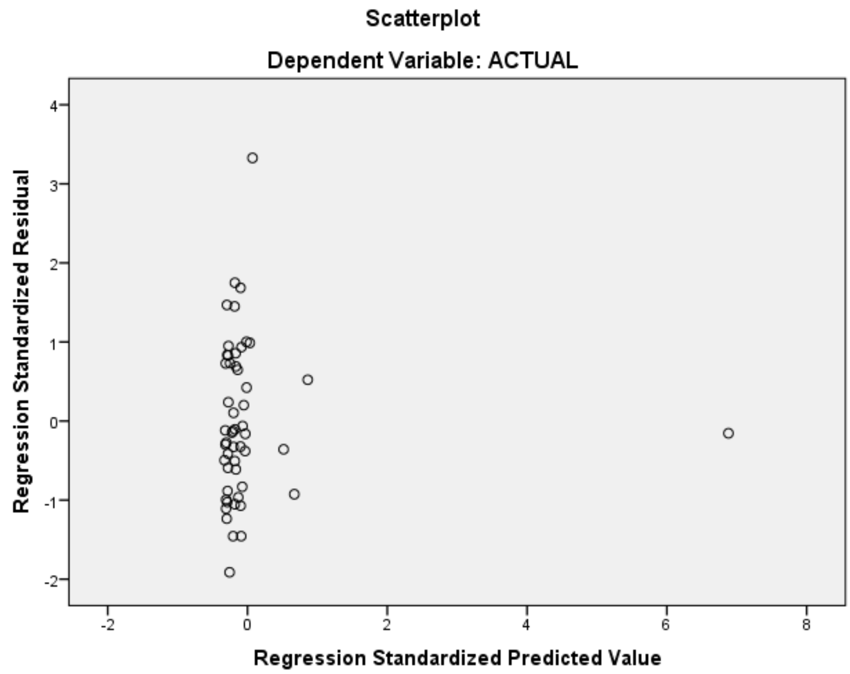
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4031.0769	5743.7681	4109.2959	237.53668	52
Residual	-2345.94336	4081.11646	.00000	1214.45629	52
Std. Predicted Value	-.329	6.881	.000	1.000	52
Std. Residual	-1.913	3.327	.000	.990	52

a. Dependent Variable: ACTUAL

Charts


Histogram
Dependent Variable: ACTUAL
Mean = 4.32E-16
Std. Dev. = 0.990
N = 52
Frequency
15
10
5
0
-2
-1
0
1
2
3
4
Regression Standardized Residual


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: ACTUAL
Expected Cum Prob
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Observed Cum Prob


Scatterplot
Dependent Variable: ACTUAL
Regression Standardized Residual
Regression Standardized Predicted Value
4
3
2
1
0
-1
-2
-2
0
2
4
6
8

LOGISTIC REGRESSION ANALYSIS FOR PSEUDO POSITIVE AND PSEUDO NEGATIVE DATA

CASES	POSITIVE / NEGATIVE	EVI
1	1	3454.655
2	1	3638.082
3	1	3016.373
4	1	4287.811
5	1	5336.754
6	1	3026.8
7	1	4864.241
8	1	5552.158
9	1	2032.842
10	1	5039.533
11	1	3283.104
12	1	4684.414
13	1	4297.386
14	1	4130.732
15	0	4801.065
16	0	5513.483
17	0	4831.484
18	0	5909
19	0	6072.379
20	0	6083.631
21	0	4685.458
22	0	5032.569
23	0	5571.586
24	0	4845.589
25	0	5717
26	0	4577.098
27	0	4929.39
28	0	4856.081

1=pseudo positive

0=pseudo negative

[DataSet4] C:\Users\14645822\Documents\MEETING 52\LOGISTIC REGRESSION 3.sav

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	230	100.0
	Missing Cases	0	.0
	Total	230	100.0
Unselected Cases		0	.0
Total		230	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
Negative	0
Positive	1

Block 0: Beginning Block

Classification Table[a,b]

	Observed		Predicted		
			POSITIVE NEGATIVE		Percentage Correct
			Negative	Positive	
Step 0	POSITIVE NEGATIVE	Negative	0	115	.0
		Positive	0	115	100.0
	Overall Percentage				50.0

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.000	.132	.000	1	1.000	1.000

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	EVI	40.582	1	.000
	Overall Statistics		40.582	1	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	45.036	1	.000
	Block	45.036	1	.000
	Model	45.036	1	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	273.812[a]	.178	.237

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.351	8	.968

Contingency Table for Hosmer and Lemeshow Test

		POSITIVE NEGATIVE = Negative		POSITIVE NEGATIVE = Positive		Total
		Observed	Expected	Observed	Expected	
Step 1	1	19	18.620	4	4.380	23
	2	16	16.770	7	6.230	23
	3	15	16.208	9	7.792	24
	4	14	14.239	9	8.761	23
	5	14	12.857	9	10.143	23
	6	13	11.132	10	11.868	23
	7	9	9.598	14	13.402	23
	8	9	7.614	14	15.386	23
	9	4	5.395	19	17.605	23
	10	2	2.566	20	19.434	22

Classification Table[a]

	Observed	Predicted

		POSITIVE NEGATIVE		Percentage Correct
		Negative	Positive	
Step 1	POSITIVE NEGATIVE Negative	81	34	70.4
	POSITIVE NEGATIVE Positive	42	73	63.5
	Overall Percentage			67.0

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	EVI	-.001	.000	34.041	1	.000	.999
	Constant	4.240	.750	31.991	1	.000	69.389

a. Variable(s) entered on step 1: EVI.

LOGISTIC REGRESSION FOR SECOND SET OF PSEUDO ABSENCE DATA

```
[DataSet3] C:\Users\14645822\Documents\MEETING 52\LOGISTIC REGRESSION 4.sav
```

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	230	100.0
	Missing Cases	0	.0
	Total	230	100.0
Unselected Cases		0	.0
Total		230	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 0: Beginning Block

Classification Table[a,b]

	Observed		Predicted		
			POSITIVE NEGATIVE		Percentage Correct
			0	1	
Step 0	POSITIVE NEGATIVE	0	0	115	.0
		1	0	115	100.0
	Overall Percentage				50.0

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.000	.132	.000	1	1.000	1.000

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	EVI	44.103	1	.000
	Overall Statistics		44.103	1	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	48.915	1	.000
	Block	48.915	1	.000
	Model	48.915	1	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	269.933[a]	.192	.255

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.702	8	.952

Contingency Table for Hosmer and Lemeshow Test

		POSITIVE NEGATIVE = 0		POSITIVE NEGATIVE = 1		Total
		Observed	Expected	Observed	Expected	
Step 1	1	20	19.130	3	3.870	23
	2	18	17.013	5	5.987	23
	3	16	15.850	7	7.150	23
	4	13	14.458	10	8.542	23
	5	10	12.878	13	10.122	23
	6	12	10.923	11	12.077	23
	7	10	9.277	13	13.723	23
	8	8	7.471	15	15.529	23
	9	5	5.223	18	17.777	23
	10	3	2.775	20	20.225	23

Classification Table[a]

	Observed		Predicted		
			POSITIVE NEGATIVE		Percentage Correct
			0	1	
Step 1	POSITIVE NEGATIVE	0	81	34	70.4
		1	41	74	64.3
	Overall Percentage				67.4

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	EVI	-.001	.000	36.796	1	.000	.999
	Constant	4.139	.704	34.564	1	.000	62.771

a. Variable(s) entered on step 1: EVI.

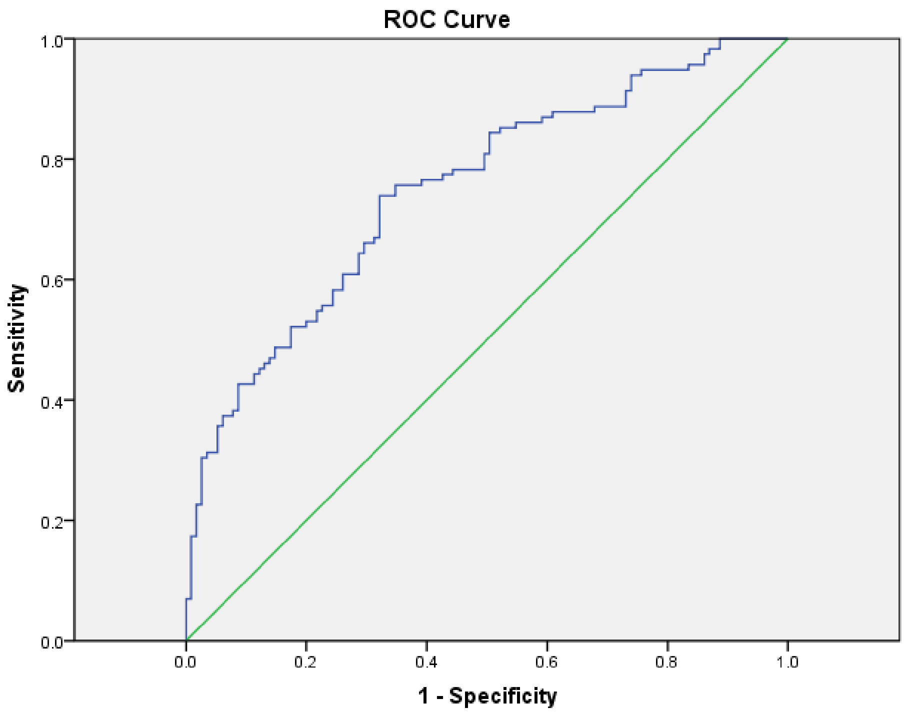
RESULT OF ROC CURVE FOR PRESENCE DATA (ACTUAL CASES)

Case Processing Summary

POSITIVE NEGATIVE	Valid N (listwise)
Positive[a]	115
Negative	115

Larger values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is 1.



Area Under the Curve

Test Result Variable(s):Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.750	.032	.000	.688	.812

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Area Under the Curve

Test Result Variable(s):Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.750	.032	.000	.688	.812

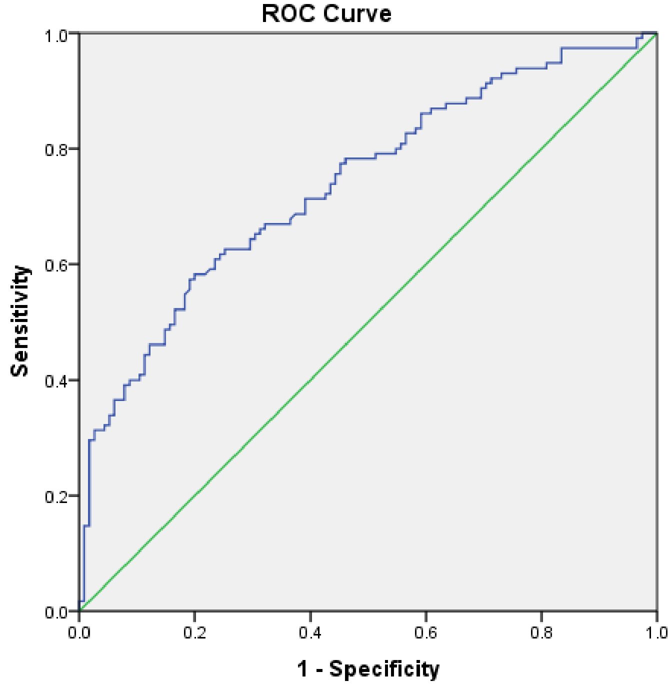
RESULT OF ROC CURVE FOR ABSENCE DATA (FOR THE FIRST SET OF RANDOM POINTS)

Case Processing Summary

POSITIVE NEGATIVE	Valid N (listwise)
Positive[a]	115
Negative	115

Smaller values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is Positive.



Diagonal segments are produced by ties.

Area Under the Curve

Test Result Variable(s): EVI

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.739	.032	.000	.675	.802

The test result variable(s): EVI has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

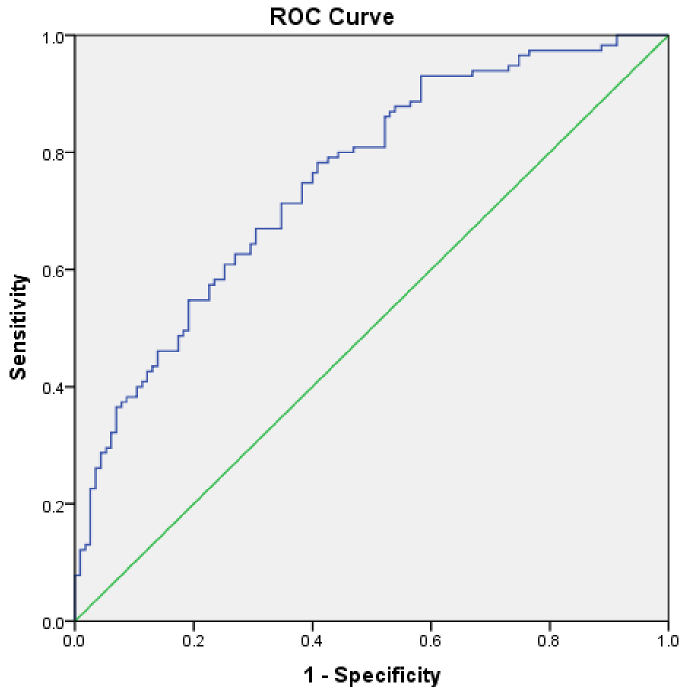
RESULT OF ROC CURVE FOR ABSENCE DATA (FOR THE SECOND SET OF RANDOM POINTS)

Case Processing Summary

POSITIVE NEGATIVE	Valid N (listwise)
Positive[a]	115
Negative	115

Smaller values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is 1.



Area Under the Curve

Test Result Variable(s): EVI

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.754	.031	.000	.692	.815

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5