



Faculty of Resource Science and Technology

**Assessment of Environmental Variables as Potential Risk Factors for the
Transmission of *Leptospira* spp.**

Rosdi Anak Kira

**Master of Science
2020**

Assessment of Environmental Variables as Potential Risk Factors for the
Transmission of *Leptospira* spp.

Rosdi Anak Kira

A thesis submitted

In fulfillment of the requirements for the degree of Master of Science

(Environmental Microbiology)

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK

2020

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.



.....

Signature

Name: Rosdi Anak Kira

Matric No.: 17020184

Faculty of Resource Science & Technology

Universiti Malaysia Sarawak

Date: 4th November 2020

ACKNOWLEDGEMENT

First and foremost, I want to praise God the Almighty, for giving me the spiritual strength throughout my postgraduate journey. His grace and mercies had given me strength, patience and determination in overcoming the challenges that had come upon me throughout this project.

I wish to express my sincere appreciation to my respected supervisor, Associate Professor Dr Lesley Maurice Bilung and co-supervisors, Professor Kasing Apun and Professor Lela Su'ut, for their full support and guidance in completing this project. Their guidance, patience, advices and supports enabled me to complete this research. It was a privilege and a great pleasure for me to be able to complete this project under their supervisions.

I would also like to extend my deepest gratitude to my other co-supervisor and mentor, Dr Romano Ngui for sharing his expertise and knowledge with me especially in the analytical phase of this study. He was the substance of a genius: he convincingly guided and encouraged me to be professional and keep on moving forward even when the journey got tough. Without his persistent motivation, I would not have been able to finish this project.

Last but not least, I wish to acknowledge the full support and great love from my family, especially my mother, Liah Anak Elaw; my sister and her husband: Meliny Anak Kira and Mohd. Hafiz Bin Jumaat. They supported me financially and kept me going on this project. This work would not have been possible without their input.

ABSTRACT

A diverse *Leptospira* serovars and reservoir hosts have been reported to circulate in various environmental settings in Sarawak. The risk of infection has been linked to socio-economic and environmental risk factors. This underscored the potential for a shift in the transmission dynamics of leptospirosis in Sarawak. It is imperative to consider the broad socio-demographical and environmental determinants such as climate and agricultural intensity, which may act as modifiable predictors for a disease like leptospirosis. These factors may be unevenly distributed, and their spatial distribution may influence the dynamics of leptospirosis. Hence, this study aims to determine the demographical and spatial distribution of leptospirosis using confirmed cases from 2011-2018 at a district level. This study also aims to investigate the global and localised relationship between leptospirosis with climate and agricultural variables. The demographical distribution of leptospirosis was analysed based on descriptive and inferential statistics by aggregating the data according to demographical groups. Spatial statistics including global and local Moran's *I* was employed to study the spatial autocorrelation of leptospirosis incidence rate by districts. The potential clustering of hot spot and cold spot regions were investigated using Getis-Ord-Gi* statistics. The relationship between environmental factors and spatial density of geocoded leptospirosis cases was explored by using global Poisson regression (GPR) and geographically weighted Poisson regression (GWPR). A total of 4,648 cases were reported from 2011-2018 with the highest incidence in 2018, with 31 cases per 100,000 population. One hundred and thirty-nine fatal cases were reported, with the overall case fatality rate of 2.99%. Higher incidence rate per 100,000 population was observed in males (35.79; 95% CI: 22.92- 48.67), individuals aged between 50-59 years (27.22; 95% CI: 15.91- 38.53), other Bumiputera ethnic group (32.21; 95% CI: 16.80-47.82) and elementary occupational

type (56.93; 95% CI: 26.49-87.37). Overall, the leptospirosis incidence in Sarawak demonstrated a positive spatial clustering (Moran's I : 0.28) with significant positive local clustering observed in 2012 (Moran's I : 0.23), 2013 (Moran's I : 0.33) and the highest Moran's I in 2014 (0.45). An increasing incidence trend was observed towards the southeast region with a standardised incidence ratio (SIR) between 0.39 and 4.54. The southeast region was identified as a significant hot spot region (z-scores: 1.58 to 2.83). Leptospirosis in the hot spot region was concentrated in areas with higher water vapour pressure ($\exp(\beta)$: 1.12; 95% CI: 1.02 - 1.25) and annual precipitation ($\exp(\beta)$: 1.15; 95% CI: 1.02 - 1.31), with lower precipitation in the driest month ($\exp(\beta)$: 0.85; 95% CI: 0.75 – 0.96) and the wettest quarter ($\exp(\beta)$: 0.88; 95% CI: 0.77 – 1.00). Water vapour pressure (WVP) has the most variation effect across the hot spot region with a standard deviation of 0.62 (LQ: 0.15; UQ: 0.99) while least variation in effect was observed in annual precipitation (ANNP) with a standard deviation of 0.14 (LQ: 0.11; UQ: 0.30). The GWPR model explained better relationships between leptospirosis and environmental covariates as denoted by the reduction in AICc value from 519.73 to 443.49. The results of localised relationships presented in this study could be used to formulate spatially targeted interventions. This would be useful particularly in localities with strong environmental or socio-demographical determinants for the transmission leptospirosis.

Keywords: Environmental variables, geographically weighted Poisson regression, *Leptospira*, Moran's I , spatial autocorrelation

Penilaian Pembolehubah Alam Sebagai Faktor Risiko untuk Penularan Leptospira spp.

ABSTRAK

Pelbagai serovar dan haiwan pembawa telah dilaporkan beredar dipelbagai tetapan alam di Sarawak. Risiko jangkitan telah dikaitkan dengan faktor sosio-ekonomi dan alam sekitar. Hal ini menekankan potensi kepada perubahan dinamik jangkitan leptospirosis di Sarawak. Adalah amat penting untuk mengambil kira kepelbagaian faktor sosio-demografik dan alam sekitar seperti iklim and keamatan pertanian, yang boleh bertindak sebagai pembolehubah kepada penyakit seperti leptospirosis. Faktor-faktor tersebut juga mungkin tidak teragih secara sekata dan beredarannya yang tidak sekata mengubah dinamik leptospirosis. Oleh itu, kajian ini bertujuan untuk menentukan agihan demografik dan ruangan leptospirosis dengan menggunakan kes yang telah disahkan dari 2011-2018 pada peringkat mukim. Kajian ini juga berniat untuk mengkaji hubungan global dan tempatan diantara leptospirosis dengan pembolehubah iklim dan pertanian. Analisis mengenai agihan demografik telah ditentukan dengan statistic deskriptif dan inferensi, di mana data telah dikumpulkan mengikut kumpulan demografik. Statistik ruangan seperti Moran's I global dan tempatan telah digunakan untuk mengkaji autokorelasi ruangan kadar insiden leptospirosis mengikut daerah. Kewujudan kluster tompok panas dan tompok sejuk telah dikaji menggunakan statistik Getis-Ord-Gi. Hubungan diantara faktor alam dan kepadatan kes leptospirosis telah diteroka menggunakan regresi Poisson global (GPR) dan regresi Poisson berwajaran geografi (GWPR). Sebanyak 4,648 kes telah dilaporkan daripada tahun 2011-2018 dengan insiden tertinggi pada 2018, dengan 31 kes per 100,000 populasi. Seratus tiga puluh sembilan kes maut telah dilaporkan dengan kadar kematian sebanyak 2.99%. Insiden yang lebih tinggi telah dicatatkan oleh kaum lelaki 35.79 (95% CI 22.92- 48.67), kumpulan usia di antara 50-59 tahun (27.22; 95% CI: 15.91- 38.53), kumpulan etnik Bumiputera lain-*

lain (32.21; 95% CI: 16.80-47.82) dan kumpulan pekerjaan asas (56.93; 95% CI: 26.49-87.37). Keseluruhannya, insiden leptospirosis di Sarawak telah menunjukan kluster ruangan global yang positif (*Moran's I*: 0.28) dengan kluster tempatan signifikan dicatatkan pada 2012 (*Moran's I*: 0.23), 2013 (*Moran's I*: 0.33) dan yang paling tinggi pada 2014 (0.45). Kenaikan aliran insiden telah dilaporkan ke arah tenggara dengan nisbah selaras insiden di antara 0.39 dan 4.54. Bahagian tenggara Sarawak telah dikenalpasti sebagai tompok panas yang signifikan (z : 1.58 to 2.83). Leptospirosis di tompok panas tertumpu pada kawasan yang mempunyai tekanan wap air ($\exp(\beta)$: 1.12; 95% CI: 1.02 - 1.25) dan jumlah hujan tahunan yang lebih tinggi ($\exp(\beta)$: 1.15; 95% CI: 1.02 - 1.31), dengan jumlah hujan pada bulan yang paling kering ($\exp(\beta)$: 0.85; 95% CI: 0.75 – 0.96) dan hujan pada suku tahun yang paling basah ($\exp(\beta)$: 0.88; 95% CI: 0.77 – 1.00) yang lebih rendah. Tekanan wap air mempunyai variasi kesan yang paling tinggi dengan sisihan piawai berjumlah 0.62 (LQ: 0.15; UQ: 0.99) manakala jumlah hujan tahunan mencatatkan variasi paling rendah dengan sisihan piawai 0.14 (LQ: 0.11; UQ: 0.30). Model GWPR telah mengesan hubungan di antara leptospirosis dan faktor-faktor alam dengan lebih terperinci berdasarkan penurunan angka AICc daripada 519.73 to 443.49. Hasil daripada hubungan setempat yang ditunjukkan dalam kajian ini dapat diguna untuk merumuskan strategi yang lebih terperinci. Ini amat penting terutamanya di kawasan yang mempunyai penentu alam atau sosio-demografik yang kuat.

Kata kunci: Faktor alam, regresi Poisson berwajaran geografi, Leptospira, *Moran's I*, autokorelasi ruangan

TABLE OF CONTENTS

	Page
DECLARATION	i
ACKNOWLEDGEMENT	ii
ABSTRACT	iii
ABSTRAK	v
TABLE OF CONTENTS	vii
LIST OF TABLES	xi
LIST OF FIGURES	xiii
LIST OF ABBREVIATIONS	xvi
CHAPTER 1: INTRODUCTION	1
1.1 Study background	1
1.2 Problem statement	4
1.3 Research questions	5
1.4 Objectives	5
1.5 Hypotheses	6
1.6 Significance of the study	6
1.7 Chapter summary	7
CHAPTER 2: LITERATURE REVIEW	8
2.1 The bacteria <i>Leptospira</i>	8

2.2	Leptospirosis	9
2.3	Ecology of leptospirosis	11
2.4	Environmental drivers of leptospirosis	13
2.4.1	<i>Leptospira</i> distribution and climate variability	13
2.4.2	Agricultural activities and human leptospirosis	14
2.5	Circulation of <i>Leptospira</i> spp. in Sarawak	15
2.5.1	<i>Leptospira</i> spp. from the wildlife of Sarawak	15
2.5.2	<i>Leptospira</i> in various ecological settings in Sarawak	15
2.5.3	Leptospire seroprevalence among communities in Sarawak	17
2.6	Spatial analysis in epidemiology of vector-borne diseases	18
2.7	Spatial and temporal characteristics of human leptospirosis	20
2.8	Chapter summary	21
	CHAPTER 3: MATERIALS AND METHODOLOGIES	22
3.1	Overview	22
3.2	Study area	23
3.3	Data collection & management	25
3.3.1	Leptospirosis data	25
3.3.2	Socio-demographic data	25
3.3.3	Climate dataset	26
3.3.4	Agricultural dataset	28

3.4	Epidemiological study of leptospirosis (Objective No. 1)	30
3.4.1	Descriptive and inferential statistics	30
3.5	Spatial analysis of leptospirosis (Objective No. 2)	30
3.5.1	The spatial trend map of leptospirosis	30
3.5.2	Spatial autocorrelation analysis of leptospirosis incidence	33
3.5.3	Hot spot analysis (Getis Ord Gi*)	36
3.6	Assessment of potential factors associated with the spatial distribution of leptospirosis (Objective No. 3)	38
3.6.1	Response and explanatory variables	38
3.6.2	Global Poisson regression (GPR)	39
3.6.3	Local Poisson regression	41
3.6.4	Mapping odds ratios and t-values	46
3.7	Chapter summary	47
	CHAPTER 4: RESULTS	48
4.1	Epidemiological study of leptospirosis (Objective No. 1)	48
4.1.1	Prevalence of leptospirosis based on the gender	49
4.1.2	Prevalence of leptospirosis based on the age group & ethnicity	50
4.1.3	Prevalence of leptospirosis based on occupational types	52
4.1.4	Spatial distribution of leptospirosis	54
4.1.5	Temporal distribution of leptospirosis	55

4.2	Spatial analysis of leptospirosis (Objective No. 2)	56
4.2.1	Mapping the spatial trend of leptospirosis	56
4.2.2	Spatial autocorrelation leptospirosis incidence	59
4.2.3	Hot spot analysis (Getis Ord Gi*)	63
4.3	Assessment of potential factors associated with the spatial distribution of leptospirosis (Objective No. 3)	67
4.3.1	Response and explanatory variables	67
4.3.2	Global Poisson regression (GPR)	70
4.3.3	Local Poisson regression	74
4.3.4	Mapping odds ratios and t-values	77
	CHAPTER 5: DISCUSSION	87
	CHAPTER 6: CONCLUSION AND RECOMMENDATIONS	97
6.1	Conclusion	97
6.2	Recommendations	98
	REFERENCES	99
	APPENDICES	120

LIST OF TABLES

	Page
Table 3.1 List of geospatial climate data collected for this study.	27
Table 3.2 List of geospatial agriculture data collected for this study.	29
Table 3.3 The cross-validation measurements to validate the EBK interpolation model.	33
Table 4.1 The prevalence of leptospirosis based on gender.	50
Table 4.2 Prevalence of leptospirosis based on age groups.	51
Table 4.3 Prevalence of leptospirosis based on ethnicity.	52
Table 4.4 Prevalence of leptospirosis in Sarawak based on occupational types.	53
Table 4.5 Results of cross-validation for evaluating the performance of EBK interpolation.	59
Table 4.6 Summary of global Moran's I analysis for leptospirosis incidence in Sarawak.	60
Table 4.7 Number of spatial clusters identified based on LISA analysis.	61
Table 4.8 Summary of hot spot analysis using Getis-Ord Gi* statistics.	63
Table 4.9 Descriptive statistics summarising the variables in the hot spot region (HH).	69

Table 4.10	Summary of univariate assessment by fitting single variable using global Poisson model	72
Table 4.11	Summary of parameter estimates for multivariate GPR model.	73
Table 4.12	Comparison of global (GPR) and local (GWPR) models statistics.	76
Table 4.13	Summary statistics of variable coefficients based on adaptive Gaussian kernel model.	77
Table 4.14	Prediction error statistics for the parameters interpolated by EBK interpolation.	80

LIST OF FIGURES

	Page
Figure 2.1 Darkfield photomicrographs of <i>Leptospira</i> spp. (Source: Adler & de la Peña Moctezuma, 2010).	9
Figure 2.2 An illustration depicting the ecology of leptospirosis. (Source: Lau et al., 2010).	11
Figure 3.1 Flowchart of research activities.	22
Figure 3.2 Sarawak state administrative boundaries based on Global Administrative Areas database (GADM, 2018).	24
Figure 3.3 Schematic representation of geographically weighted regression and its spatial parameters (Source: Feuillet et al., 2015).	45
Figure 4.1 Comparison of annual mean incidence rate of leptospirosis in Sarawak.	48
Figure 4.2 Case fatality rate of leptospirosis in Sarawak from 2011 to 2018.	49
Figure 4.3 A choropleth map illustrating the mean incidence rate among districts in Sarawak during the investigation period, 2011-2018.	54
Figure 4.4 The temporal distribution of leptospirosis incidence in Sarawak over the study period.	55
Figure 4.5 The number of reported leptospirosis cases in Sarawak during the study period, 2011-2018.	56

Figure 4.6	Map showing the standardised incidence ratio (SIR) of leptospirosis during the study period, 2011-2018.	57
Figure 4.7	Leptospirosis incidence clusters and outliers identified based on LISA analysis from 2011 to 2018.	62
Figure 4.8	Map showing the occurrence of hot spot and cold spot of leptospirosis infection over the study period, 2011-2018.	66
Figure 4.9	The geocoded leptospirosis cases from 2011 to 2018 in the hot spot region (HH), located in the southeast of Sarawak.	67
Figure 4.10	The geographical distribution of odds ratios and pseudo-t-values for water vapour pressure (WVP).	81
Figure 4.11	Spatially heterogenous odds ratios and pseudo-t-values for the variable precipitation in the driest month (PDM).	82
Figure 4.12	The geographical distribution of odds ratios and pseudo-t-values for the variable precipitation in the wettest quarter (PWQ).	83
Figure 4.13	Spatially varying odds ratios and pseudo-t-values for the variable annual precipitation (ANNP).	84
Figure 4.14	The smoothed odd ratios for the variable precipitation in the driest month (PDM; top figure) and water vapour pressure (WVP; lower figure).	85

Figure 4.15 Isopleth map illustrating the smoothed odd ratios for the variable annual precipitation (ANNP; top figure) and precipitation in the wettest quarter (PWQ; lower figure).

86

LIST OF ABBREVIATIONS

AI	Annual Incidence
AICc	Corrected Akaike Information Criterion
ANNP	Annual Precipitation
arcmin	arcminute
arcsec	arcsecond
BIC	Bayesian Information Criterion
CHELSA	Climatologies at High Resolution for the Earth's Land Surface Areas
CI	Confidence Interval
CFR	Case Fatality Rate
CLD	Cropland Density
DIFF	Difference
EBK	Empirical Bayesian Kriging
ERDAS	Earth Resource Development Assessment System
ESRI	Environmental Systems Research Institute
<i>exp</i>	exponential
FAO	Food and Agricultural Organisation
GADM	Database of Global Administrative Areas
GIS	Geographical Information System
GLM	Generalised Linear Model
GLW	Gridded Livestock of the World
GPR	Global Poisson Regression
GRASS	Geographic Resources Analysis Support System

GWR	Geographically Weighted Regression
GWPR	Geographically Weighted Poisson Regression
GIS	Geographical Information System
HH	High-High type
HL	High-Low type
ILWIS	Integrated Land and Water Information System
km	kilometer
kPa	kilopascal
LH	Low-High type
LISA	Local Indicator of Spatial Autocorrelation
LL	Low-Low type
LPS	Lipopolysaccharide
LQ	Lower Quartile
LSD	Livestock Density
MAICE	Minimum AIC Estimator
MAT	Microscopic Agglutination Test
Max	Maximum
MDL	Minimum Description Length
ME	Mean Error
Min	Minimum
MSDE	Mean Standardised Error
MSE	Mean Standard Error
mm	milometer
MOH	Ministry of Health
NA	Not Available

OR	Odds Ratios
PDM	Precipitation in the Driest Month
PDQ	Precipitation in the Driest Quarter
Pr	Probability
PS	Precipitation Seasonality
PWM	Precipitation in the Wettest Month
PWQ	Precipitation in the Wettest Quarter
RMSE	Root Mean Square Error
RMSSDE	Root Mean Square Standardised Error
SD	Standard Deviation
SIR	Standardised Incidence Ratio
STH	Soil-Transmitted Helminthiasis
SPSS	Statistical Package for Social Sciences
TS	Temperature Seasonality
UQ	Upper Quartile
VIF	Variance Inflation Factor
WGS	World Geodetic System
WHO	World Health Organisation
WVP	Water Vapour Pressure

CHAPTER 1

INTRODUCTION

1.1 Study background

Leptospirosis is regarded as one of the most important emerging zoonotic diseases with worldwide distribution (Haake & Levett, 2015). The World Health Organization (WHO) estimates the global burden of leptospirosis at over one million severe human cases per year, with a growing number of countries reporting leptospirosis outbreaks (Abela-Ridder et al., 2010; Hartskeerl et al., 2011). The causative organism is a spirochete bacterium in the genus *Leptospira*. Pathogenic *Leptospira* sp. is shed in bodily fluids of reservoir hosts, and infection occurs when the pathogen is transmitted directly and penetrates abraded skins or mucosal membranes, or through transmission via contaminated water bodies or environment. The sources of contamination are rodents, but a broad spectrum of domestic or wild animals can also be carriers of the pathogenic bacteria (Lau et al., 2010). Despite its apparent burden on human health and livestock productivity, the disease is one of the severely neglected tropical diseases (Dupouey et al., 2014). The varying clinical presentation of leptospirosis which often mimics other infectious diseases such as bacterial and viral hepatitis, influenza and malaria, may often lead to its under-reporting and misdiagnosis (Plank & Dean, 2000; Bharti et al., 2003; WHO, 2011; Crump et al., 2013). Outbreaks of leptospirosis have been reported from many regions around the world, but human infection is centred mainly in the tropical and subtropical regions with prominently high incidences in southeast Asia (Bahaman & Ibrahim, 1988; Kawaguchi et al., 2008; Al-shere et al., 2012; Hinjoy, 2016).

Leptospirosis is endemic in Malaysia, and it was gazetted as a notifiable disease since December 2010 under the Prevention and Control of Infectious Diseases Act 1988 (Ministry of Health Malaysia, 2011). The disease has continued to be a burden that significantly affects the Malaysian economy as well as the livelihood of citizens due to losses incurred from poor productivity in livestock, loss of jobs due to a reduction in the number of tourists visiting certain hot spots as well as costs of prevention and control (Garba et al., 2018). In the previous decade, there has been an upsurge of reported cases, with outbreaks primarily related to recreational activities and natural disasters (Hisham et al., 2009; Sopian et al., 2012). Risk factors including the proliferation of reservoir hosts, occupational and recreational exposure and occurrence of natural disasters have compounded the predisposition of humans to the risk of contracting leptospirosis in Malaysia (Garba et al., 2018). Climatic conditions are amongst the foremost factors for the endemicity and the persistence of leptospires in the natural environment across different geographical and ecological regions (Victoriano et al., 2009; Lau et al., 2010; Mwachui et al., 2015; Garba et al., 2018). Single outbreaks following the occurrence of extreme weather events in tropical countries such as Philippines, Laos, India and Malaysia have been deliberated in previous studies (Pappachan et al., 2004; Kawaguchi et al., 2008; Al-shere et al., 2012; Radi et al., 2018).

As urban and industrial development in Sarawak intensifies, new challenges may arise as an implication of induced ecological fragmentation and environmental changes. Ecological changes play a significant role in the emergence of infectious diseases, and urbanisation has been associated with outbreaks and the emergence of numerous infections around the world (Lau et al., 2010). Repercussions from land and natural forest clearing for rural development or agricultural expansion may have modified the distribution of zoonosis

carriers such as rodents and mosquitoes. Rodent populations may expand to where people and settlements are clustered, so population density may be a proxy for leptospirosis risk in some areas. For instance, the establishment of large-scale oil palm estates in Sarawak has provided a feeding ground for rodents and brought them into closer proximity to human settlements. Such circumstance suggested an increase chance for the transmission of *Leptospira* spp. to humans.

In recent years, the use of Geographical Information Systems (GIS) has provided a robust and rapid ability to examine spatial and temporal patterns and processes by incorporating metadata into its analysis. In turn, it has fostered discussions on integrating GIS analytics and geospatial statistics for environmental analyses of infectious diseases (Moore & Carpenter 1999). Previous studies have adopted GIS analysis tools in the development of ecological models to explore and analyse spatial variations between local environmental factors and occurrences of leptospirosis (Mohammadinia et al., 2017; Mayfield et al., 2018; Radi et al., 2018). Local modelling approaches, such as Geographically Weighted Regression (GWR) and Geographically Weighted Poisson Regression (GWPR) enable the exploration of local associations of variables (Fotheringham et al., 2003; Nakaya et al., 2005). GWPR has been used to study spatial variations in associations between health access, disease counts and risks, incidence rates, mortality risks and sets of socioeconomic characteristics (Nakaya et al., 2005; Cheng et al., 2011; Comber et al., 2011; Weisent et al., 2012; Odoi & Busingye, 2014). These studies explored the spatial pattern in the associations between a health outcome with a set of predictor variables and increase our understanding of how spatially varying risk factors can modulate health issues.

1.2 Problem statement

Sarawak Epid News reported that leptospirosis has escalated from 49 cases in 2010 to 616 cases in 2014, with a total of 53 deaths during that period (Sarawak Health Department, 2012). In this regard, the establishment of public health surveillance systems remain essential in achieving better control of this disease and should be parallel with the rapid industrialisation of Sarawak. Such manoeuvre would require an integration of multiple research disciplines including a comprehensive epidemiological study, sophisticated space-time monitoring of disease occurrence, and state of the art computer modelling to identify its underlying determinants. As an attempt to invest in disease prevention and control measures, it is imperative to consider the broad socio-demographical and environmental determinants, which may act as modifiable predictors for a disease like leptospirosis. These factors may be unevenly distributed, and their spatial distribution may influence the dynamics of leptospirosis. The ecological associations of diseases may vary geographically as covariates may change depending on different geographical contexts (Nakaya, 2005). Hence, these knowledge gaps hindered the formulation of reliable control measures and underscored the need for the establishment of site-specific intervention and emergency response.

When the understanding of leptospirosis transmission risk is uncertain, GIS technologies and spatial statistics offer tools to circumvent problems about spatial and temporal relationships in diseases with complex transmission cycle. GIS technologies can combine multiple site-specific metadata from humans, animals and environment for spatially rigorous analysis with an added advantage of interactive visual representation. Examples include the epidemiological application of data obtained from climate-based forecast systems like the observation of land cover, land use, elevation, oceans, surface temperatures

and rainfall, for disease surveillance and development of early-warning systems (Bergquist & Rinaldi, 2010). In Malaysia, studies incorporating GIS-technologies and spatial statistics were primarily focused on dengue (Azhar et al., 2016; Hazrin et al., 2016; Yusof et al., 2017; Hasim et al., 2018) and malaria (Alias et al., 2014; Yoep et al., 2015; Ahmad et al., 2018; Pahrol et al., 2018) with limited studies on leptospirosis (Hassan et al., 2016; Zulkifli et al., 2016; Radi et al., 2018). However, most of these studies were concentrated in west Malaysia. To best of our knowledge, this is the first study in Sarawak to incorporate the element of spatial statistics and GIS technologies in understanding the epidemiology of leptospirosis.

1.3 Research questions

The following research questions motivate this research:

- i. Does the prevalence of leptospirosis infections in Sarawak demonstrate distinctive socio-demographical disparities?
- ii. Which district(s) in Sarawak is most affected by leptospirosis during the study period?
- iii. Does the relationship between leptospirosis and environmental risk factors remain constant throughout the study region?

1.4 Objectives

The objectives of this research were:

- i. To analyse the epidemiological distribution of leptospirosis in Sarawak for over eight years (2011-2018).
- ii. To map and identify the occurrence of spatial trend and clusters of leptospirosis incidence at a district level in Sarawak.

- iii. To evaluate the local and global relationships between leptospirosis incidence with meteorological and agricultural variables using statistical modelling.

1.5 Hypotheses

The null and alternative hypotheses of this research are set up as the followings:

- i. There is significant spatial clustering of leptospirosis at the district level in Sarawak over the study period.
- ii. There is an association between leptospirosis incidence with the meteorological and agricultural variables.
- iii. The relationships that exist between explanatory variables and the incidence of leptospirosis are spatially heterogenous across the high-risk region.

1.6 Significance of the study

While GIS is routinely used in many countries as part of routine public health management and services, its diffusion across developing countries is not uniform (Fletcher-Lartey & Caprarelli, 2016). By mapping leptospirosis incidence, the current estimate of the disease will be apparent, and it is useful for future predictive modelling and assist in the formulation of intervention strategies. Previous studies have emphasised that predictive modelling utilizes current estimates of disease burden to predict future outbreaks based on expected changes in socio-demographical and environmental factors. Moreover, the spatial modelling from this study could provide a foundation for the development of local leptospirosis early-warning systems in Sarawak. The ability to accurately access, predict and pinpoints environmental drivers for leptospires transmission would enable the Sarawak

government to shift from disease surveillance to environmental health hazard surveillance as a more cost-effective approach in directing public health interventions.

1.7 Chapter summary

This chapter has provided a general background of this research regarding the epidemiology of leptospirosis in Sarawak and the application of spatial-temporal statistics with GIS technology. This chapter also presented the problem statement and research questions which aid in the formulation of a framework for this research.

CHAPTER 2

LITERATURE REVIEW

2.1 The bacteria *Leptospira*

Leptospire are spirochetes, about 0.1 mm in diameter by 6–20 mm in length comprising the genus *Leptospira*, which belongs to the family *Leptospiraceae* and order Spirochaetales (Faine et al., 1999). Leptospire have distinctive hooked ends with two periplasmic flagella located in the periplasmic space responsible for motility (Adler & de la Peña Moctezuma, 2010). According to Adler & de la Peña Moctezuma (2010), the family *Leptospiraceae* comprises of 13 pathogenic (i.e. having potential to cause disease in animals and humans) species: *Leptospira alexanderi*, *Leptospira alstonii*, *Leptospira borgpetersenii*, *Leptospira inadai*, *Leptospira interrogans*, *Leptospira fainei*, *Leptospira kirschneri*, *Leptospira licerasiae*, *Leptospira noguchi*, *Leptospira santarosai*, *Leptospira terpstrae*, *Leptospira weilii*, *Leptospira wolffii*, with more than 260 serovars. It also includes saprophytic (i.e. free living and considered not to cause disease) species such as *Leptospira biflexa*, *Leptospira meyeri*, *Leptospira yanagawae*, *Leptospira kmetyi*, *Leptospira vanthielii* and *Leptospira wolbachii*, with more than 60 serovars (Adler & de la Peña Moctezuma, 2010). The serovar classification designated to *Leptospira* was established based on the expression of surface-exposed epitopes of the lipopolysaccharide (LPS) antigens, while the specificity of epitopes is dependent on the sugar composition and orientation of the bacteria. *Leptospira* are obligate aerobic bacteria, and the optimum temperature for their growth is 28-30 °C in laboratory conditions. The doubling-time for pathogenic strains is approximately 6-8 hours and reaching maximum growth in approximately 4-7 days. Meanwhile, non-pathogenic leptospire have a doubling time of 3.5-4.5 hours and reaching maximum growth

in 2-3 days (Faine et al., 1999). Survival and persistence of *Leptospira* bacteria in the environment are dependent on the interaction of multiple factors, including abiotic soil and water conditions. They can live in soil with a moisture content of >20%, water and soil with pHs around 5.5 and 7.6 and temperatures ranging from 4 to 40 °C (Barragan et al., 2017). As such, they are able to survive in alkaline soil, mud, swamps, streams, rivers, organs and tissues of either living or dead animals (Mohammed et al., 2011).

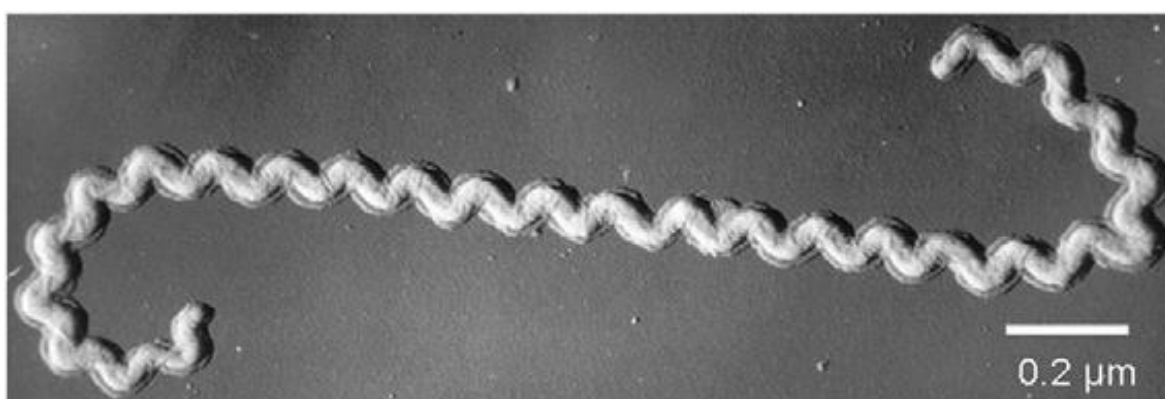


Figure 2.1: Darkfield photomicrographs of *Leptospira* spp. (Source: Adler & de la Peña Moctezuma, 2010).

2.2 Leptospirosis

Leptospira can enter the host via small abrasions, breaches of the surface integument, conjunctiva, mucous membrane and genital tract. They may settle in the convoluted tubules of the kidneys and be shed in the urine for a period of a few weeks to several months and occasionally even longer (Mohammed et al., 2011). Toxic cellular components and consequent symptoms appear once the number of leptospires in blood and tissues reaches a critical level. Leptospiral LPS (leptospiral lipopolysaccharide) preparations exhibit activity in biological assays for endotoxin like other Gram-negative bacteria (Mohammed et al., 2011). The clinical manifestation may include fever, renal and hepatic insufficiency, pulmonary manifestation, and reproductive failure. However, the severity of leptospirosis in

humans may differ depending on the infecting serovar, age, health and immunological competence of the patient (Adler & de la Peña Moctezuma, 2010). As these clinical symptoms mimic several other common diseases such as dengue, typhoid fever and hantavirus infection, incorrect diagnoses occur frequently and consequently lead to the underestimation of the number of individuals with leptospirosis. Leptospirosis may develop into a much severe form of infection (Weil's disease) in 10% of patients and can be lethal. The clinical manifestations of this form include fever, jaundice, renal failure, myocarditis, uveitis, meningitis and lethal pulmonary haemorrhage, culminating with multiorgan failure (Faine et al., 1999; Seguro & Andrade 2013). This severe form of the pulmonary disease may be responsible for 50% of the lethal cases of leptospirosis (Fraga et al., 2015).

Besides human infections, leptospirosis has also been reported in animals such as dogs, horses, pigs and cattle (Bernard, 1993; Andre-Fontaine, 2006; Grooms, 2006). Clinical manifestations in animals may also vary due to host-adapted serovars, such as Canicola in dogs, Bratislava in horses and pigs, Hardjo in cattle and Australis and Pomona in pigs. According to Levett (2001), animals recovering from leptospirosis may become asymptomatic carriers harbouring virulent leptospires in the renal tubules for extended periods and shedding infectious leptospires into the environment. Meanwhile, other species such as mice (*Mus musculus* and other *Mus* species) and rats (mainly *Rattus norvegicus* and *R. rattus*) serve as reservoirs for their host-related serovars (mice for Ballum, Icterohaemorrhagiae and rats for Copenhageni) (Bharti et al., 2003). Hence, these asymptomatic carriers and reservoir animals become a valuable source of infection for susceptible humans.

2.3 Ecology of leptospirosis

The ecology of leptospirosis involves the complex interaction between humans, animal reservoirs, leptospires and the environment in which they coexist (Lau et al., 2010). *Leptospira* bacteria are maintained by a broad spectrum of reservoir animals serving as the primary contamination source. As leptospires are disseminated into the environment via the urine of infected or reservoir animals, their survival is subjected to environmental drivers such as climate and land use as illustrated in Figure 2.2 (Lau et al., 2010). Sociodemographic factors, including occupational types and the involvement of water-based activities, would increase the risk of transmission to susceptible humans. This elucidates the potential of epidemiological discrepancy as transmission cycles, and covariates vary depending on ecological settings.

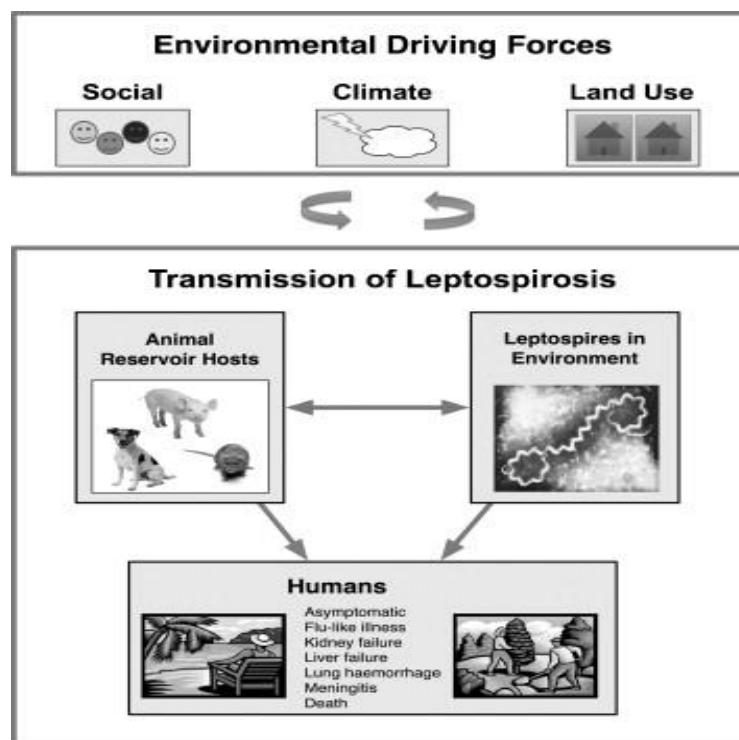


Figure 2.2: An illustration depicting the ecology of leptospirosis. (Source: Lau et al., 2010).

According to Lau et al. (2010) transmission in rural areas are linked to farming and livestock with increased risk during warm and rainy months (Lau et al., 2010). Individuals with an occupational type that requires continuous contacts with contaminated soil and water have a higher risk of contracting leptospirosis. Previous reports describe leptospirosis as an occupational disease, due to higher prevalence among specific high-risk group individuals which include farmers, abattoir workers, miners, sewers and town cleaners among others (Cook et al., 2016; Saha & Thakurta, 2016). In Malaysia, high leptospiral seroprevalence was reported among individuals working in oil palm plantation (Tan, 1970; Ridzuan et al., 2016; Tan et al., 2016). *Leptospira* have also been proven to circulate among domestic and wildlife animals including dogs, cattle, primates, rats and bats (Bahaman et al., 1988; Thayaparan et al., 2013; Lau et al., 2016). Such circumstances highlight the relevance of animal contacts and host abundance as risk factors contributing to the transmission of leptospirosis in rural areas.

Meanwhile, infections in urban areas are usually associated with overcrowding, poor hygiene standards and inadequate sanitation. The presence of garbage and other wastes fosters the proliferation of reservoir animals such as rats. In Malaysia, diverse *Leptospira* serovars have been documented to circulate among rats captured from various urban settings including villages, residential areas, hawker centres, wet and night markets (Benacer et al., 2013; Pui et al., 2017a). A similar observation has also been reported in other countries such as India, Brazil and Tokyo (Bhardwaj et al., 2008; Reis et al., 2008; Koizumi et al., 2009).

2.4 Environmental drivers of leptospirosis

2.4.1 *Leptospira* distribution and climate variability

Climatic conditions are at the forefront among the common factors for the endemicity and the persistence of leptospirosis (Victoriano et al., 2009; Lau et al., 2010). Tropical countries have been challenged by various endemic infectious diseases, including leptospirosis, due to their favourable climate (Gallana et al., 2013). In Thailand, prominent occurrences of leptospirosis infections in different geographical regions were reported during flooding and rainy seasons (Chadsuthi et al., 2012; Suwanpakdee et al., 2015). Comparable observations were noticeable in the Philippines where leptospirosis epidemics have demonstrated seasonal cycles associated with temperature, rainfall and humidity (Al-shere et al., 2012; Sumi et al., 2017; Matsushita et al., 2018). The intricate associations between climatic factors and leptospirosis were also reported in various geographical settings and climate zones including Sri Lanka, New Caledonia, Brazil, Reunion Island and Columbia with distinctive lagged effects (Desvars et al., 2011; Weinberger et al., 2014; Ghizzo-Filho et al., 2018; Guetierrez & Martinez-Vega, 2018). For instance, rainfall at a lag of 8-months was significantly associated with leptospirosis in New Caledonia (Weinberger et al., 2014). In the Philippines, Matsushita et al. reported a substantial increase in leptospirosis cases associated with rainfall at a lag of two weeks (Matsushita et al., 2018). In Malaysia, flooding is a recurrent phenomenon during the monsoon season (Soo et al., 2019). Outbreaks of leptospirosis infections following extreme weather events like flooding have been reported in Peninsular Malaysia, including Kelantan, Terengganu and Pahang (Yaakob et al., 2015; Radi et al., 2018). Benacer et al. deliberated the disparity in seasonal cycles of leptospirosis in Malaysia (Benacer et al., 2016). A higher incidence in peninsular Malaysia was reported to occur between October to March (Benacer et al., 2016). These studies were

primarily focused in peninsular Malaysia, while the relationship between climatic factors and leptospirosis in Sabah and Sarawak, particularly in the state of Sarawak remains unclear.

2.4.2 Agricultural activities and human leptospirosis

The changes in human behaviours as driven by population growth, economic and technological advancement, along with the spatial expansion of agriculture are creating more intensive interactions between humans, livestock and wildlife (Jones et al., 2013). Encroachment of human settlements and agriculture on natural ecosystems results in expansion of transition zones between adjacent ecological systems known as ecotones (Jones et al., 2013). Such circumstance provides new opportunities for pathogen spillover, genetic diversification, and adaptation. Associations between disease emergence and ecotones have been suggested for several diseases, including leptospirosis (Jones et al., 2013). A study in northeastern Malaysia by Rafizah et al. revealed a remarkably higher frequency of leptospirosis infection among agricultural workers (Rafizah et al., 2013). A comparable observation was also reported by Tan et al. where individuals involved in agriculture-based or plantation sector contributed 14.7% of the leptospirosis cases reported during 2012 and 2013 (Tan et al., 2016). A cross-sectional study by Ridzuan et al. in selected oil palm plantations in southern states of Malaysia revealed an exceptionally higher seroprevalence among fruit collectors (Ridzuan et al., 2016). According to Ridzuan et al. (2016) fruit collectors in plantations have a higher risk for leptospirosis infection due to their continuous contact with environmental soil and water and prone to obtain wounds (Ridzuan et al., 2016). Besides plantation workers, animal rearing could also heighten the risks of contracting leptospirosis. Daud et al. (2018) delineated the presence of garbage dumping areas in cattle farms as a contributing factor for the high leptospirosis seroprevalence amongst the workers (Daud et al., 2018).

2.5 Circulation of *Leptospira* spp. in Sarawak

2.5.1 *Leptospira* spp. from the wildlife of Sarawak

The abundance of reservoir animals and asymptomatic carriers have a direct bearing on the dissemination of *Leptospira* in the natural environment. A wide spectrum of domestic and peridomestic animals have been documented to harbour pathogenic *Leptospira* with host-adapted serovars. This includes Canicola in dogs, Bratislava in horses and pigs, Hardjo in cattle, Australis and Pomona in pigs and Icterohaemorrhagiae and Copenhageni in rats (Bernard, 1993; Bharti et al., 2003; Andre-Fontaine, 2006; Grooms, 2006). Serovars such as Australis, Autumnalis, Bataviae, Canicola, Copenhageni, Celledoni, Shermani, Djasiman, Grippotyphosa, Hebdomadis, Icterohaemorrhagiae, Javanica, Patoc, Pomona, Pyrogenes, Sejroe, Lai and Lepto 175 (Sarawak) are commonly found in Malaysia (Thayaparan et al., 2013b). In Sarawak, *Leptospira* serovars like Lai, Lepto 175, Pomona, Copenhageni, Pyrogenes, Icterohaemorrhagiae and Noguchii were reported being circulated amongst wildlife captured from national parks that are frequently visited by tourists (Thayaparan et al., 2013; Thayaparan et al., 2014; Thayaparan et al., 2015; Pui et al., 2017b). *Leptospira* spp. also isolated from non-human primates, rats, squirrels and mongoose (Thayaparan et al., 2013a; Thayaparan et al., 2013b; Thayaparan et al., 2015; Pui et al., 2017a). Such circumstances evidently underscored the substantial role of wildlife as the source of infection and amplifies the magnitude of leptospirosis transmission in Sarawak.

2.5.2 *Leptospira* in various ecological settings in Sarawak

Leptospirosis infections in human usually involved direct and indirect contact with infected animal urine and contaminated environmental surroundings. This includes environmental soil, water from various ecological settings in which humans and reservoir

animals coexist. A study by Pui et al. (2015) deliberated the prevalence of pathogenic, intermediate and saprophytic *Leptospira* strains in environmental soil and water samples collected from national parks in Sarawak (Pui et al., 2015). *L. borgpetersenii* serovar Mini was detected in soil samples collected from Bako National Park, a prominent tourist attraction in Sarawak. Intermediate strain, *L. wolffii* serovar Khorat were also detected in water and soil samples from Tanjung Datu and Bako National Parks (Pui et al., 2015). According to Pui et al. (2015), it is a possibility that urine from an infected animal was the source of *L. borgpetersenii* serovar Mini, as it is the main serogroup circulating in animals (Pui et al., 2015). Besides national parks, Pui and colleagues also documented the prevalence of *Leptospira* spp. from urban areas including a public university, villages, residential areas, commercial and hawker centres as well as markets (Pui et al., 2017a). The study revealed *L. noguchii*, as the dominant pathogenic strain in rats, soil and water samples and emphasized the role of reservoir host as the disseminating agents. It was highlighted that indiscriminate disposal of food waste, presence of domestic animals, stagnant pond water and poor hygiene are the contributing factors for the presence of *Leptospira* spp. detected in the study (Pui et al., 2017a). Additionally, *Leptospira* spp. have also been documented to circulate in national service centres and paddy fields of Sarawak. Pui and colleagues revealed 2.9% each of soil and water samples collected from national service training centres and paddy fields harboured pathogenic *Leptospira* strains (Pui et al., 2017b). Interestingly, intermediate strains, *L. wolffii* serovar Khorat were also detected in the study with a prevalence of 2.7% and 1.9% in soil and water, respectively (Pui et al., 2017b). Since little is known about the pathogenicity of intermediate *Leptospira* group (Slack et al., 2006), it is postulated that it has the potential to induce life-threatening complications and that precautionary action should be taken in handling soil and water samples (Pui et al., 2015). The presence of

Leptospira spp. from paddy fields and various urban settings in Sarawak poses potential risk of acquiring leptospirosis by engaging in agricultural and water-based activities.

2.5.3 Leptospire seroprevalence among communities in Sarawak

From 2010-2014, 53 victims have succumbed to leptospirosis in Sarawak (Sarawak Health Department, 2012). However, as the number of documented cases were established based on clinical records, the actual burden of leptospirosis in Sarawak is underestimated. A serological prevalence study was done by Thayaparan et al. (2015) to assess the extent of leptospire exposure among the local communities in Sarawak (Thayaparan et al., 2015). It was revealed that 35.9% of the sampled population demonstrated seropositivity towards serovar Lepto 175 and detected as the dominant strain affecting the population in the study area (Thayaparan et al., 2015). Males demonstrated higher seroprevalence in comparison to female, and those between the ages of 36-45 years and 18-25 years were at higher risk of being seropositive. Thayaparan et al. (2015) underscored the significance of living near forested areas, having skin wounds, contact with farm animals and the presence of rats as contributing factors to the seroprevalence among the studied group (Thayaparan et al., 2015). A comparable finding was deliberated by Su'ut et al. (2016) in Rejang Basin, Sarawak, where an overall seroprevalence of 37.4% was reported. Interestingly, 53.2% of the respondents showed antibodies against 2 to 5 serovars, an indication of repeated infections to different serovars under various circumstances (Su'ut et al., 2016). *L. interrogans* serovars Djasiman, Shermani, and Pomona were the main circulating pathogenic strain detected in the study. Residential areas near forest fringes and far upriver, agro-based and subsistence economic activities, high annual rainfall and high density of wild animals were regarded as factors associated with high seropositivity.

2.6 Spatial analysis in epidemiology of vector-borne diseases

Shaw et al. (2012) deliberated the significance of Geographic Information Systems (GIS) in health care, surveillance of infectious diseases, mapping and monitoring of the spatio-temporal vector distributions (Shaw, 2012). It incorporates sophisticated algorithms, spatial analysis, geo-statistics and modelling, hence, making GIS technology a powerful tool for the prediction of disease patterns and parasite ecology associations (Higgs, 2004; Guo et al. 2005; García-Rangel & Pettorelli, 2013). According to Fletcher-Lartey & Caprarelli (2016), every GIS is structured around five fundamental components: (i) spatially referenced data, collected and stored in a relational geodatabase (ii) the hardware physically storing data and processing tools; (iii) the software assembling the user-interface algorithms by which users access the database, query and analyse the data; (iv) the algorithms and data management procedures; and (v) the people, both producers and consumers of spatial data (Fletcher-Lartey & Caprarelli, 2016). Basic operations include maps production by overlaying data stored as tables comprising details of geographic features symbolized by points, lines or polygons, or raster datasets, and their geographic coordinates (Fletcher-Lartey & Caprarelli, 2016). Once the features have been mapped, geostatistical analysis, important disease monitoring and investigation analysis such as cluster detection can be carried out using the tools included in the GIS software package (Bergquist & Rinaldi, 2010). This approach can be followed by more sophisticated modelling by linking disease processes with explanatory spatial variables to comprehend the mechanisms of disease spreading. Such circumstances have increased the accessibility and reliability of integration between health data and mapping processes which enable researchers to study the relationship between spatial and temporal trends of environmental factors and diseases. Geographic Resources Analysis Support System (GRASS), Earth Resource Development Assessment System

(ERDAS), MapInfo, Integrated Land and Water Information System (ILWIS) and ArcGIS (Environmental Systems Research Institute ESRI, Redlands, California) are examples of GIS software available (Sharma et al., 2015).

The hybrid of remote sensing and GIS technologies have allowed rapid epidemiological mapping to comprehend the spatial variation of vector diversity and abundance. Such application has facilitated the control of vector-borne diseases such as tickborne-diseases (Ozdenerol, 2015; Ben et al., 2018), rabies (Saputra et al., 2015; Castillo-Neyra et al., 2017), malaria (Gwitira et al., 2018) and dengue (Dom et al., 2017; Sirisena et al., 2017). In Malaysia, GIS-based technologies were applied in the analysis of vector ecology, exploration of environmental drivers in facilitating disease transmission and the development of early warning systems and risk models. A study by Saffawati et al. (2019) in the southwest of Penang Island (Malaysia), mapped the distribution of *Aedes albopictus* and evaluated the association between larval abundance with temperature, humidity and light intensity. The study revealed that the spatial interpolation between parameters and mean larvae density demonstrated distinct patterns, with concentrated population density at a temperature range of 32.25 to 34.07 °C, relative humidity 61.97 to 65.13% and light intensity at 2.49 to 2.80 kLux (Saffawati et al., 2019). In Sabah, Pahrol et al. (2018) investigated the spatial distribution of malaria from 2012 to 2016. The results indicated that the incidence of malaria in Sabah was clustered for all years ($p\text{-value} < 0.05$), within central and western Sabah, and recorded the highest incidence rate. Beside dengue and malaria, GIS-based techniques were also incorporated into the mapping and modelling of soil-transmitted helminthiasis (STH). A study in peninsular Malaysia by Ngui et al. (2014) demonstrated no clear pattern for the geographical distribution of STH across 99 surveyed locations. These studies identify geographic areas with high disease prevalence, and the outcomes are useful

for guiding the allocation of scarce public health resources. Integration of GIS-techniques in vector-borne diseases underscored areas where the technology has provided greater insight into transmission patterns and environmental drivers.

2.7 Spatial and temporal characteristics of human leptospirosis

Over the years, the combination of GIS technologies and spatial statistics have improved our understandings on the spread of leptospirosis and its underlying social and environmental determinants. In Gilan, Iran, a study by Nia et al. (2015) utilized satellite images of vegetation data, digital elevation model and meteorological data points to evaluate the spatial heterogeneity of leptospirosis (Nia et al., 2015). Results of the study revealed that temperature, precipitation, humidity and vapour were positively associated with the disease (Nia et al., 2015). Rood and colleagues identified spatial clusters of high and low leptospirosis incidence rate in the Netherlands using Moran's *I* and Local Indicator of Spatial Association (LISA) (Rood et al., 2017). The results identified seven spatially distinct hot spots (high incidence rate) with municipality incidence to be correlated up to the approximately 12.8km distance (Rood et al., 2017). There were significant associations between leptospirosis with built-up areas, grassland, heavy sabulous clay in soil and the proportion of the population between 45-64 years of age (Rood et al., 2017).

In China, Dhewantara et al. (2018) documented the demographical, ecological and socioeconomic characteristics between high and low leptospirosis infection risk counties. In their study, occurrences of high-risk clusters were situated at higher elevation (576.01 m; 95% CI: 451.17–700.25, $p < 0.001$), higher precipitation rate (136.86 mm per month; 95% CI: 123.61–150.12, $p < 0.001$), lower farmland production (2,949.67 kg/ha; 95% CI: 1,953.41–3,945.93 $p < 0.001$) and lower GDP (440.80 Yuan, 95% CI: 236.61–644.98, $p <$

0.001) (Dhewantara et al., 2018). According to Dhewantara et al. (2018), the changes in leptospirosis distribution in the study may be explained by effective control efforts, ecological and social changes that have occurred in China. For instance, alteration in landscapes due to industrial and urban development changed vegetation structure and population dynamics of host animals such as rodents, leading to changes in leptospirosis distribution (Dhewanara et al., 2018). In Malaysia, only a few studies have outlined the characteristics of the occurrence of high leptospirosis infection clusters and their underlying determinants. A study in Selangor has identified high-risk areas around a temperature of 20 to 30 °C, with peculiar occurrences at agricultural, forested and recreational areas (Hassan et al., 2016). In Kelantan, Radi and colleagues revealed a significant clustering of leptospirosis incidence in post-flood period with a Moran's I value of 0.19 and a z-score of 9.74, the highest during the study course (Radi et al., 2018). This highlights extreme weather events such as flooding as a significant risk factor for the transmission of leptospirosis. The study also reported higher case density in urbanized settings with proximity to garbage areas, presence of vegetation i.e., crops and horticultural lands, warmer temperature and high rainfall in the preflood period (Radi et al., 2018). These studies demonstrated the spatially heterogeneous relationship between leptospirosis with ecological and socio-economic determinants.

2.8 Chapter summary

This chapter summarises the ecology and circulation of *Leptospira* spp. as delineated in previous studies. The findings from previous studies which have incorporated GIS-technology and spatial modelling in studying the epidemiology of leptospirosis were also summarised in this chapter.

CHAPTER 3

MATERIALS AND METHODOLOGIES

3.1 Overview

The methodology of this research is simplified in Figure 3.1.

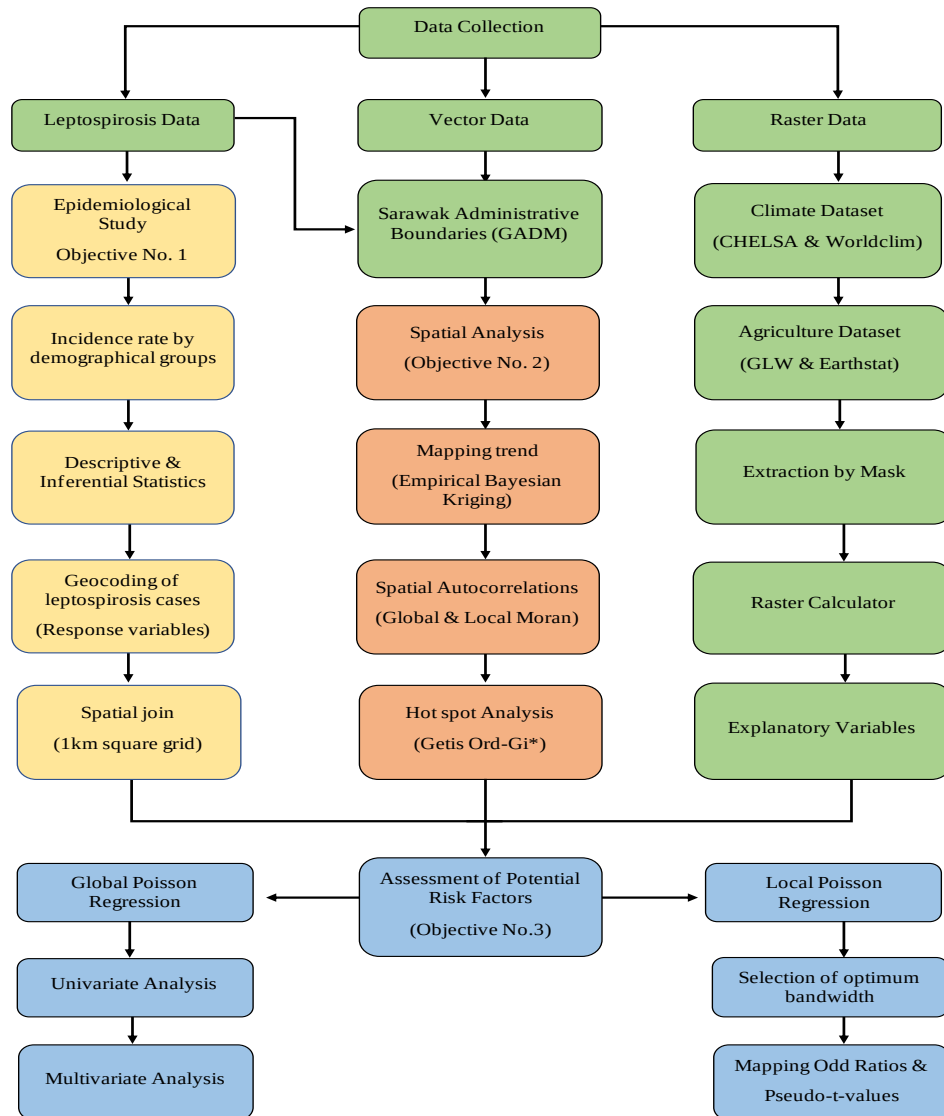


Figure 3.1: Flowchart of research activities. The boxes are color coded to represent different components in this study: green: data collection and management; yellow: first objective; red: second objective; blue: third objective.

3.2 Study area

The state of Sarawak is located between latitude $0^{\circ} 50' 5^{\circ}$ N and longitude $109^{\circ} 36' 115^{\circ} 40'$ E on the island of Borneo (Figure 3.2). Sarawak is the largest state in Malaysia with an area of $124,451 \text{ km}^2$ and a total population of 2,830,000 (Department of Statistics Malaysia, 2019). Based on the local government classification, there are 12 divisions with 57 districts and sub-districts in Sarawak. The population is concentrated in districts such as Kuching (693,800) and Miri (347,600), while the southeast districts e.g Song and Belaga have a lower population with 23,700 and 43,200 inhabitants (Department of Statistics Malaysia, 2019). The Sarawak natives (Bumiputera Sarawak) made up the largest portion of the population with a total of 1,147,634 followed by Chinese and Malays with 560,150 and 551,567, respectively. Natural vegetation covered most of the land in Sarawak with a total area of $106,115.8 \text{ km}^2$ followed by cropland with an area of $15,147 \text{ km}^2$ and urban associated areas make up 830.1 km^2 of the land use. Agriculture is a crucial industry in Sarawak with main crops such as oil palm, rubber, paddy, pepper, cocoa and coconut (Department of Statistics Malaysia, 2019). Oil palm has the largest planted areas in Sarawak with a total of 1,572,477 hectares followed by rubber with 166,636 hectares. Besides crops, domestication of livestock including pig, cattle, goat, rabbit, deer, chicken, duck and quail has also contributed towards the economy of the state. Sarawak has an equatorial climate and experiences two monsoon seasons; northeast and southwest monsoon. The average minimum and maximum temperatures are 23.8°C and 32.1°C , respectively with a mean humidity of 84.40%. On average, the amount of rainfall in Sarawak is approximately 3,696.29 mm (Department of Statistics Malaysia, 2019). Vector data for the study area was retrieved from the Database of Global Administrative Areas, GADM (2018).



Figure 3.2: Sarawak state administrative boundaries based on Global Administrative Areas database (GADM, 2018).

3.3 Data collection & management

3.3.1 Leptospirosis data

Ethical approval for the study was obtained from the National Medical Research (NMRR, Ethical Number:18-2713-42365), Ministry of Health Malaysia (MOH) and Sarawak State Planning Unit, SPU (Appendix A). Reported leptospirosis cases in Sarawak from the year 2011 to 2018 were retrieved from the Infectious Disease Unit, Sarawak State Health Department. All data obtained were anonymised to ensure confidentiality. Cases were confirmed through clinical evaluation and laboratory diagnostic tests. Under the guideline published in 2011, confirmed cases of leptospirosis must be notified within one week upon diagnosis to the district health office. A probable case of leptospirosis is defined as a subject with positive ELISA or any other rapid tests. Confirmed leptospirosis case is defined as a suspected or probable case with confirmatory microscopic agglutination test (MAT) with single serum specimen - titre $\geq 1:400$ and a four-fold or higher rise in titre for paired sera (Ministry of Health Malaysia, 2011). The dataset comprises of confirmed cases in each district along with addresses, month, year, gender, occupational types and age. It reflects the occurrence of the disease in different districts in the study region.

3.3.2 Socio-demographic data

Population data during the study period were obtained from the Department of Statistics Malaysia. Demographical factors analysed in this study include age group, gender, ethnicity and occupation. For this study, age was aggregated into seven groups: <11, 11–20, 21–30, 31–40, 41–50, 51–60, and > 60 years. The ranges for the age groups were based on the grouping from the Department of Statistics Malaysia. Occupational data was aggregated into seven major groups based on Malaysian Standard Classification of Occupations

(MASCO); clerical, craft and support workers, elementary occupations, plant and machine operators and assembler, professionals, service and sales workers, skilled agricultural, forestry, livestock and fishery and technician & associate professionals (Ministry of Human Resource Malaysia, 2013). Children and students were classified as not working. In terms of ethnicity, the data was aggregated into seven major groups based on the ethnic classification by the Department of Statistics Malaysia: Bidayuh, Chinese, Iban, Malay, Melanau, Other Bumiputera and others. The case profiles with no occupational information were denoted as Not Available (NA).

3.3.3 Climate dataset

Spatially interpolated gridded climate data were used in many areas of work, particularly in the environmental, agricultural and biological sciences (Fick & Hijmans, 2017). As the evaluation of spatial heterogeneity of climate factors with the probable transmission of leptospires is part of the study objective, eight geospatial climatic variables were included in this study. The variables were retrieved from CHELSA (Karger et al., 2016) and WorldClim (Fick & Hijmans, 2017). The climate data included in this study are listed in Table 3.1.

Table 3.1: List of geospatial climate data collected for this study.

Variables (unit)	Types	Resolution	Source
Water Vapour Pressure, WVP (kPa)	Raster	30 arcsec, ~1 km	Worldclim (Fick & Hijman, 2017)
Temperature Seasonality, TS (SD)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)
Precipitation Seasonality, PS (Coefficient of variation)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)
Annual Precipitation, ANNP (mm/year)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)
Precipitation in Driest Quarter, PDQ (mm/quarter)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)
Precipitation in Driest Month, PDM (mm/month)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)
Precipitation in Wettest Month, PWM (mm/month)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)
Precipitation in Wettest Quarter, PWQ (mm/quarter)	Raster	30 arcsec, ~1 km	CHELSA (Karger et al., 2016)

3.3.4 Agricultural dataset

This study also aims to explore the relationship between leptospirosis and agricultural factors. Hence, livestock density and cropland density were included as independent variables. Livestock densities were obtained from the FAO Gridded Livestock of the World (GLW) dataset (Wint & Robinson, 2007; Robinson et al., 2014). The GLW provided modelled livestock densities of the world at a spatial resolution of 5 arcmins (~10 km). The estimated densities were based on statistical relationships between observed densities within administrative units ('training data') derived from survey and census data, and several explanatory variables, including a time-series of remotely-sensed satellite data relating to climate and the environment, and other spatial data relating to demography, land cover and terrain (Robinson et al., 2014). EarthStat dataset of crop harvest area, which provides global crop harvest areas at 5 arcmins (~10 km) resolution was also incorporated in this study (Ramankutty et al., 2008). The list of variables in the agricultural dataset is presented in Table 3.2.

Table 3.2: List of geospatial agriculture data collected for this study.

Variables (unit)	Types	Resolution	Source
Chicken Density (total number per pixel)	Raster	5 arcmin, ~ 10 km	GLW (Wint & Robinson, 2007)
Duck Density (total number per pixel)	Raster	5 arcmin, ~ 10 km	GLW (Wint & Robinson, 2007)
Pig Density (total number per pixel)	Raster	5 arcmin, ~ 10 km	GLW (Wint & Robinson, 2007)
Cattle Density (total number per pixel)	Raster	5 arcmin, ~ 10 km	GLW (Wint & Robinson, 2007)
Cropland Density (%)	Raster	5 arcmin, ~ 10 km	EARTHSTAT (Ramankutty et al., 2008)

All products are in a geographic coordinate system referenced to the *World Geodetic System* (WGS 84) horizontal datum, with the horizontal coordinates expressed in decimal degrees. Variables processing and management were conducted using ArcGIS Desktop 10.3 (Environmental Systems Research Institute, 2015). The geospatial meteorological and agricultural variables are presented in Appendices B and C, respectively.

3.4 Epidemiological study of leptospirosis (Objective No. 1)

3.4.1 Descriptive and inferential statistics

Data curation were conducted in Microsoft Excel, and statistical analyses were performed using Statistical Package for Social Sciences, (SPSS, 2011). The data was aggregated by year from 2011 to 2018. An independent-samples t-test was conducted to compare the incidence rate between males and females. A One-way Analysis of Variance (ANOVA) was employed to investigate whether leptospirosis incidence rate differ with respect to demographical factors (age groups, occupational types, ethnic groups) and spatial-temporal factors (districts and months). The independent variables were the different groups in demographical, spatial and temporal factors while the dependent variable was the incidence rate per 100,000 population over the study period. Shapiro-Wilks test was conducted prior to any analysis to examine the distribution of the data. When the homogeneity of variance assumption is violated (Levene's test, $p < 0.05$), Welch's F test was opted instead of classic ANOVA. Annual incidence (AI) rate per 100,000 population for each district was calculated by dividing the number of cases for each year by the corresponding district population obtained from the Department of Statistics Malaysia (Department of Statistics Malaysia, 2019). Case fatality rates (CFRs) were calculated by dividing the number of fatal cases by the number of reported cases of leptospirosis.

3.5 Spatial analysis of leptospirosis (Objective No. 2)

3.5.1 The spatial trend map of leptospirosis

Disease mapping is used for reporting the results of geographical correlation study and pinpoint areas of high and low incidence. The distribution of disease cases whether in the form of counts or case address locations, can be thought to follow an underlying model,

and the observed data may contain extra noise. In this context, the noise refers to the random variation around the model of interest (Lawson, 2013). For instance, the observed disease pattern could arise due to the unspecified spatial distribution which includes the population spatial distribution. As such, disease map should be based on smoothed estimates, clean of noise and adjusted for variation in ‘at-risk’ population (Lawson, 2013). In this study, the standardised incidence ratio (known hereafter as SIR) for each district was calculated in order to compare the incidence rates among districts. The SIR was calculated using the following formula:

$$SIR_i = O_i/E_i \quad \text{Equation 3.1}$$

Where the expected number of cases (E) by district (i) was calculated by:

$$E_i = \left(\frac{\sum_{i=1}^n O_i}{\sum_{i=1}^n P_i} \right) P_i \quad \text{Equation 3.2}$$

In Equation 3.2, P_i denotes the population of each n total communities. If the calculated SIR of a district equals to one, it indicates that the number of observed cases is equal to the expected number of cases. SIR value greater than one denotes higher observed cases than expected, while SIR less than one indicates less observed cases than the expected. Mapping of SIR provides an invaluable starting point for epidemiologic enquiry as it would reveal the spatial trend of leptospirosis infection across the study area. It is useful in identifying the underlying environmental risk factors as these factors are spatially continuous. In order to visualise the overall trend of SIR across the study area, an interpolation (smoothing) is required to produce a spatially continuous surface.

Empirical Bayesian kriging (EBK) was applied to generate a continuous surface map. In common kriging method, the prediction in unknown locations is made through calculation of semi-variogram concerning known data location, resulting in the underestimation of the standard error of the prediction. EBK model uses an intrinsic random function for spatial interpolation and does not assume a tendency toward an overall mean, leading to the same chance for large deviations to get larger or smaller (Mirzaei & Sakizadeh, 2016). The automated simulation-based optimization of EBK modelling technique facilitates a quick fitting of many theoretical semi-variogram models by estimating the model parameters instead of adjusting the suitability of a single semi-variogram (Krivoruchko, 2012). The use of geostatistical modelling approach in generating spatial distribution map requires the accuracy of the estimated values to be assessed. Hence, the cross-validation technique was employed to validate the spatial map generated by evaluating the accuracy of critical parameters that may affect the interpolation process. In this study, five evaluation criteria, namely mean error (ME), mean standardized error (MSDE), mean standard error (MSE), root-mean-square standardized error (RMSSDE), and root mean square error (RMSE) were employed to evaluate model performance (Gupta et al., 2017). The details of cross-validation measurements and evaluation criteria are given in Table 3.3.

Table 3.3: The cross-validation measurements to validate the EBK interpolation model.

Cross Validation Parameters	Descriptions
Mean Error (ME)	ME value should be close to zero for a model that provides accurate predictions.
Mean Standardised Error (MSE)	MSDE value should be as close to zero as possible.
Mean Standard Error (MSE)	The value of MSE should be as close to RMSE as possible.
Root-Mean-Square-Standardised Error (RMSSDE)	The value should be close to 1 as possible. RMSSDE value >1 signifies underestimation, while value <1 signifies overestimation.
Root-Mean-Square Error (RMSE)	RMSE value should be as small as possible to indicate how closely the model predicts the measured values.

3.5.2 Spatial autocorrelation analysis of leptospirosis incidence

The ability to visualize spatial data allows for the quick identification of any obvious patterns, and in general, spatial patterns can be classified as regular, random, or clustered. The term ‘clustering’ is used to describe the spatial aggregation of disease events. There are several methods for analysing clusters as specific or non-specific or commonly referred to as ‘local’ and ‘global’. Global (non-specific) clustering methods are used to assess whether clustering is apparent throughout the study region but do not identify the location of clusters (Anselin, 1995). Local (specific) methods of cluster detection define the locations and extent

of clusters (Anselin, 1995). In other words, the term ‘clustering’ applies to global methods of cluster analysis, while ‘cluster detection’ refers to methods of analysis.

3.5.2.1 Global Moran’s I (global clusters)

Global Moran’s I statistics were calculated in order to determine the presence of spatial dependence in age-standardized incidence rate (SIR). Moran's Index (I) is one of the most preferred and used measures of spatial autocorrelation (Moran, 1950). It assesses the pattern of a data set spatially and determines if it is dispersed, clustered, or random based on the locations and values of the given feature. The tool computes the Moran's I Index value along with the z-score and p-value to further validate the Index statistically. It is calculated based on the following formula:

$$I = \frac{n}{S_0} \frac{\sum_{x=1}^n \sum_{y=1}^n w_{xy} z_x z_y}{\sum_{x=1}^n z_x^2} \quad \text{Equation 3.3}$$

Where z_x stands for deviation of an attribute from its mean ($x_i - \bar{X}$) for feature X , w_{xy} is the spatial weight among feature X and Y , n is the total number of features and s_0 is the summation of all spatial weights (Cliff & Ord, 1981). Equation 3.3 standardised value x for observation i to determine if it is high or low relative to the mean and standardised values of x for j to determine if the neighbourhood exhibit high or low relative to the mean. In this study, the incidence rate for each district was standardised in a similar manner as a statistical z-score that compares observations to the mean in order to determine the observations’ relative position within a distribution. The range of the Moran's Index lies between +1 and -1. A positive Moran’s I indicate positive autocorrelation (i.e., clustering of areas of similar attribute values or clustered), while a negative coefficient indicates negative spatial autocorrelation (i.e., that neighbouring areas tend to have dissimilar attribute values or

scattering). Moran's, I value close or equal to zero indicates the null hypothesis of no clustering or absence of autocorrelation.

The conclusions about the observed pattern are drawn only after the interpretation of the z-score and the p-value of the index. The basic assumption for the Global Moran's I statistic is that the data values are independent and randomly arranged in the geographical space. When the p-value is greater than 0.05, the basic assumption is accepted indicating that the data values are spatially random. When the p-value is less than 0.05 and the z-score is negative, the underlying assumption of randomness is rejected, inferring that the values in the dataset are dispersed spatially. Similarly, when the p-value is less than 0.05 with a positive z-score, the assumption of randomness is again ruled out and the inference drawn is that data values are spatially clustered in the geographical space (Moran, 1950).

3.5.2.2 Local Indicator of Spatial Autocorrelation (LISA)

Locality with high incidence rate (hot spots) can be spatially clustered or exist individually as outliers. In this study, spatial clusters of leptospirosis prevalence would be districts with high incidence rate surrounded by other districts with high incidence. In contrast, spatial outliers of leptospirosis prevalence would be districts with high incidence rate surrounded by districts with low incidence. Spatial clusters and outliers can be identified using local Moran's I index by using the following formula:

$$I_i = \frac{z_i - \bar{z}}{\sigma^2} \sum_{j=1, j \neq i}^n [w_{ij} (z_j - \bar{z})] \quad \text{Equation 3.4}$$

where z_i is the value of the variable z at location i ; $\bar{z}$ is the average value of z with the sample number of n ; z_j is the value of the variable z at all the other locations (where $j \neq i$); σ^2 is the variance of variable z ; and w_{ij} is a weight which can be defined as the inverse of the

distance d_{ij} among locations i and j (Anselin, 1995; Getis & Ord, 1996). The outcomes of LISA analysis can be classified into four categories:

- High-High: high incidence value in a district with neighbouring districts having high values of incidence (positive association)
- Low-Low: low incidence value in a district with neighbouring districts having low incidence values (positive association)
- High-Low: high incidence value in a district with neighbouring districts having low incidence values (negative association)
- Low-High: low incidence value in a district with neighbouring districts having high incidence values (negative association)

The significance of global and local Moran analysis can be tested using conditional randomization approach based on Monte Carlo method. Complex permutation methods are used to conduct the tests of significance by randomly permuting the observed values over all the locations and calculating the local Moran statistic for each new permutation (Anselin, 1995). The significance of the local Moran statistic is determined by generating a reference distribution using 999 random permutations. All Moran's I analyses were conducted using first-order queen contiguity spatial weights, using software program GeoDA (Anselin, 2003).

3.5.3 Hot spot analysis (Getis Ord G_i^*)

In order to test for statistically significant higher and lower incidence rate among districts, a hot spot analysis was done using Getis -Ord G_i^* statistics. Getis-Ord G_i^* statistics

measure the clustering intensity of high and low incidence values among features about its neighbouring features. The local sum for a feature and its neighbours is compared proportionally to the sum of all features; when the local sum is much different from the expected local sum, and that difference is too large to be the result of random chance, it will result in a statistically significant z-score. Getis-Ord G_i^* statistic generates z-score and p-values for each feature to indicate whether the standardized leptospirosis incidence rate at a given district is statistically clustered. The statistical equation of Getis-Ord G_i^* can be expressed as:

$$G_i^*(d) = \frac{\sum j W_{ij}(d) x_j - W_i \bar{x}^*}{s^* \{[(n S_{ii}^* - W_i^{*2}) / (n-1)]^{1/2}\}} \quad \text{Equation 3.5}$$

where $W_{ij}(d)$ is the spatial weight with values of all features 'j' within distance d of targeted feature i. W_i is the sum of weights, S_{ii}^* is the sum of squared weights and s^* is the standard deviation of the data in the features. A z-score above 1.96 or below -1.96 denotes a statistically significant hot spot or cold spot at $p < 0.05$. If a z-score is statistically significant, the larger the z-score, the more intense the clustering of high values (hot spot). In contrast, a significant z-score, with negative value denotes clustering of low values (cold spot). In this analysis, the queen case contiguity method was selected to define the neighbourhood as districts in the study area sharing edges and corner. In this study, the cumulative incidence rate for each district was standardised in a similar manner as a statistical z-score that compares observations to the mean in order to determine the observations' relative position within a distribution. Hot spot analysis was conducted using spatial analyst toolbox in ArcGIS 10.3 (Environmental System Research Institute, 2011). The hot spot region

identified in this analysis was used in modelling the relationship between leptospirosis and environmental variables in objective number three.

3.6 Assessment of potential factors associated with the spatial distribution of leptospirosis (Objective No. 3)

The concepts and techniques presented so far have dealt with describing, visualizing, and exploring spatial data of leptospirosis. In this section, the analytical techniques of regression and discrimination are used as a mean of quantifying the effect of a set of explanatory variables on the spatial distribution of leptospirosis.

3.6.1 Response and explanatory variables

Individual leptospirosis case in the hot spot region (known hereafter as hot spot region) was geocoded using latitude and longitude of residential address which were derived using Google Earth Pro (version 7.3.2). The effect of the geographical inaccuracy of geocoded data was reduced by aggregating spatial points into 1×1 km square grid (Rood et al., 2017). A grid size of 1 km was chosen as the analytical unit because it retains more details of geospatial environmental data with 1 km resolution (exp: meteorological data). The grids were generated using ArcMap ‘create fishnet function’ in the data management toolbox. The grids encompass the area over which the spatial join is to be performed. Geocoded leptospirosis cases from 2011 to 2018 were spatially joined into the grid (Appendix D). The number of events within each square was used as attribute values representing the case count and was used as the response variable in the regression analysis.

The explanatory variables used in this study were presented in Section 3.2.3 (climate dataset) and Section 3.2.4 (agriculture dataset). Sampling block (1 × 1 km) were intersected with the environmental data layers to extract data in each cell grid. Descriptive statistics

were used to summarise the environmental variables in the high-risk region. All environmental variables are standardized to have a mean of zero and a standard deviation of 1, enabling the parameter estimates to be directly comparable (Vega-Corredor & Opadeyi, 2014). Spatially joined leptospirosis cases and the extracted geospatial environmental data was managed in excel table and was saved in comma-separated values (.csv) file format.

3.6.2 Global Poisson regression (GPR)

3.6.2.1 Univariate Analysis

Evaluation of potential risk factors was done by using global Poisson regression (GPR). The Poisson model was adopted in accordance with the characteristic of the leptospirosis data which, due to low numbers of observed counts, were expected to follow Poisson distribution (Lovett & Flowerdew 1989). The global Poisson regression model can be express as written as:

$$O_i \sim \text{Poisson}[E_i \exp(\beta_0 + \sum_k \beta_k x_{ki})] \quad \text{Equation 3.6}$$

Where β_0 denotes the model intercept and β_k refers to the parameters for k explanatory variables, x_k in locations i . (Nakaya, 2005). Parameters of Poisson regression are estimated by maximum likelihood. By allowing l_i to be the log-likelihood of the sample in place (i), the estimation problem can be defined as:

$$\begin{aligned} \max_{\beta} L &= \sum_i l_i(\beta | O_i) \\ &= \sum_i (-\hat{O}_i(\beta) + O_i \log \hat{O}_i(\beta)) \end{aligned} \quad \text{Equation 3.7}$$

Where $\beta = (\beta_0, \beta_1, \dots)$ is the vector of the regression parameters. The first part of the analysis involved the evaluation of a single explanatory variable (univariate). The data frame was imported into R-studio using the “read” function. Prior to any analysis, the structure and summary of each variable in the data frame were explored using “summary” and “str” command. Poisson regression model for each variable was estimated using the “glm” function and 95% confidence interval for each parameter estimates was generated using “confint” function. The parameter estimate from the Poisson model is interpreted as the following:

- $\exp(\alpha)$ = effect on the mean Y, when $X = 0$
- $\exp(\beta)$ = with every unit increase in X, it has a multiplicative effect of $\exp(\beta)$ on the mean of Y
- If $\beta = 0$, then $\exp(\beta) = 1$, and the expected count is $\exp(\alpha)$ and, Y and X are not related.
- If $\beta > 0$, then $\exp(\beta) > 1$, and the expected count is $\exp(\beta)$ times larger than when $X = 0$
- If $\beta < 0$, then $\exp(\beta) < 1$, and the expected count is $\exp(\beta)$ times smaller than when $X = 0$

Variance inflation factors (VIF) were estimated using “vif” function from the car package to examine the relations among all the independent variables and checking for multicollinearity. In this study, VIF value greater than 4 ($VIF > 4$) was eliminated and the data fitting process was reiterated until only variables with $VIF < 4$ were left. Variables with $VIF < 4$ were included in the multivariate analysis by employing stepwise regression using

“step” function with “both” directions. The best predictors based on the minimisation of *AIC* in stepwise regression will be used in the formulation of multivariate global Poisson regression (GPR) and geographically weighted Poisson regression (GWPR).

3.6.2.2 Multivariate Analysis

The fitting process for the multivariate model was done in similar steps as the univariate model, with the addition of two or more variables. The goodness of fit test was performed on the residual deviance to assess the model fit. The residual deviance is the difference between the deviance of the current model and the maximum deviance of the ideal model in which the predicted values are identical to the observed. Hence, if the residual difference is small, the goodness of fit test will not be significant ($p > 0.05$), an indication that the model fits the data well. However, if the test had been statistically significant, it is an indication of biasness. The goodness of fit test was performed using “pchisq” command. Univariate and multivariate analysis was performed in the R-statistical program (R core team, 2013). The R-codes and packages used in this study are presented in Appendix E.

3.6.3 Local Poisson regression

3.6.3.1 Geographically weighted Poisson regression (GWPR)

While a global regression model is a significant approach in studying diseases, it assumes that the relationships are spatially homogeneous may hide the underlying local relationship that may exist. Such instance highlights the need to consider the underlying spatial variability in order to avoid bias in the model outcomes. This problem can be circumvented by integrating Geographically Weighted Regression (GWR) into the analysis. GWR allows estimated regression associations to vary smoothly across a study area by using weighted estimation where nearby observations are weighted more heavily than distant ones.

This is motivated by Tobler's First Law of Geography – everything is related to everything else, but closer things are more related (Tobler, 1970). GWR can be extended to a generalised linear model (GLM) framework as described in previous studies (Fortheringham et al., 2003; Nakaya et al., 2005). Poisson regression models are generalized linear models with the Poisson distribution function as the assumed probability distribution of the response (Alves et al., 2016). Poisson models may be appropriate when the response variable is a count, especially in the current study case where observed counts include low numbers, such as the number of leptospirosis cases in each cell grid.

In general, O_i denotes the number of leptospirosis cases observed in each cell grid i in the high-risk region, $i = 1, \dots, n$ (total number of grids). In this study, the numbers of leptospirosis cases in each cell grid were assumed to follow independent Poisson distributions. Geographically Weighted Poisson Regression (GWPR) allows the examination of varying spatial coefficients over space. The model can be expressed as:

$$O_i \sim \text{Poisson} \left(E_i \exp \left(\sum_k \beta_k(u_i, v_i) x_{ki} \right) \right) \quad \text{Equation 3.8}$$

where O_i denote the dependent variable (leptospirosis case counts) and x_{ki} is the k th independent variable including the constant term, while E_i is the offset variable corresponding to population size at risk at grid i . (u_i, v_i) is the geographic coordinate of the centroid of the i th grid (the location of i). The coefficients $\beta_k(u_i, v_i)$ are assumed to be smoothly varying conditional on their location (Alves et al., 2016). Following the local

likelihood methodology, the parameters at location i can be estimated by calculating the following maximisation problem of geographically weighted log-likelihood:

$$\max L(u_i) = \sum_j^N (-\hat{\theta}_j(\beta(u_i)) + \theta_j \log \hat{\theta}_j(\beta(u_i))) \cdot w_{ij}(\|u_i - u_j\|) \quad \text{Equation 3.9}$$

where $\hat{\theta}_j(\beta(u_i))$ is the predicted number of leptospirosis cases at location j with the parameter regression point i ,

$$\hat{\theta}_j(\beta(u_i)) = E_j \exp \left(\sum_k \hat{\beta}_k(u_i) x_{k,j} \right) \quad \text{Equation 3.10}$$

and w_{ij} is the geographical weight of the j th observation at the i th regression point. As the distance between the regression point i and location at which the j th observation increase, the weights of the observations also gradually decreases.

In GWR, geographic weights are assigned using a smoothly decaying kernel function of distance such as a Gaussian kernel or a Bi-square kernel. The Gaussian kernel weight function continuously and gradually decreases from the centre of the kernel, but never reaches zero. The bi-square kernel has a clear-cut range where kernel weights are non-zero. Gaussian kernel and bi-square Kernel can be expressed using the following equations:

$$w_{ij} = \exp \left(-\frac{d_{ij}^2}{G^2} \right) \quad \text{Equation 3.11}$$

$$w_{ij} = \begin{cases} \left(1 - \frac{d_{ij}^2}{G^2}\right)^2 & d_{ij} \leq G \\ 0 & d_{ij} > G \end{cases} \quad \text{Equation 3.12}$$

Where w_{ij} is the geographical weight assigned to the i th observation at a regression based at point j in the study area (referred to as the j th regression location), d_{ij} is the Euclidean distance between observation i and regression location j , and G is a fixed bandwidth size defined by a distance metric measure (Fotheringham et al., 2003; Alves et al., 2016). The bandwidth controls the rate at which weights decay with increasing distance. When the bandwidth is large, the weights decay slowly and vice versa.

In terms of influencing the fit of the model, the choice of bandwidth is more important than the shape of the kernel (Fotheringham et al., 2003). As an alternative to the weighting scheme is presented in Equation 3.8 and Equation 3.9, an adaptive kernel weighting function can also be used, where the bandwidth is not fixed, and the number of areas included in the kernel is kept constant (Figure 3.3). This weighting scheme is applicable when there is a large variation in the geographical density of the observed data (Fotheringham et al., 2003). In the present study, the calibration of the best kernel and bandwidth type was done by comparing four different kernel types available in GWPR; fixed Gaussian, fixed bi-square, adaptive Gaussian and adaptive bi-square. In order to avoid large

standard errors and biasness in local parameter estimates, optimal bandwidth size is necessary (Cheng et al., 2011).

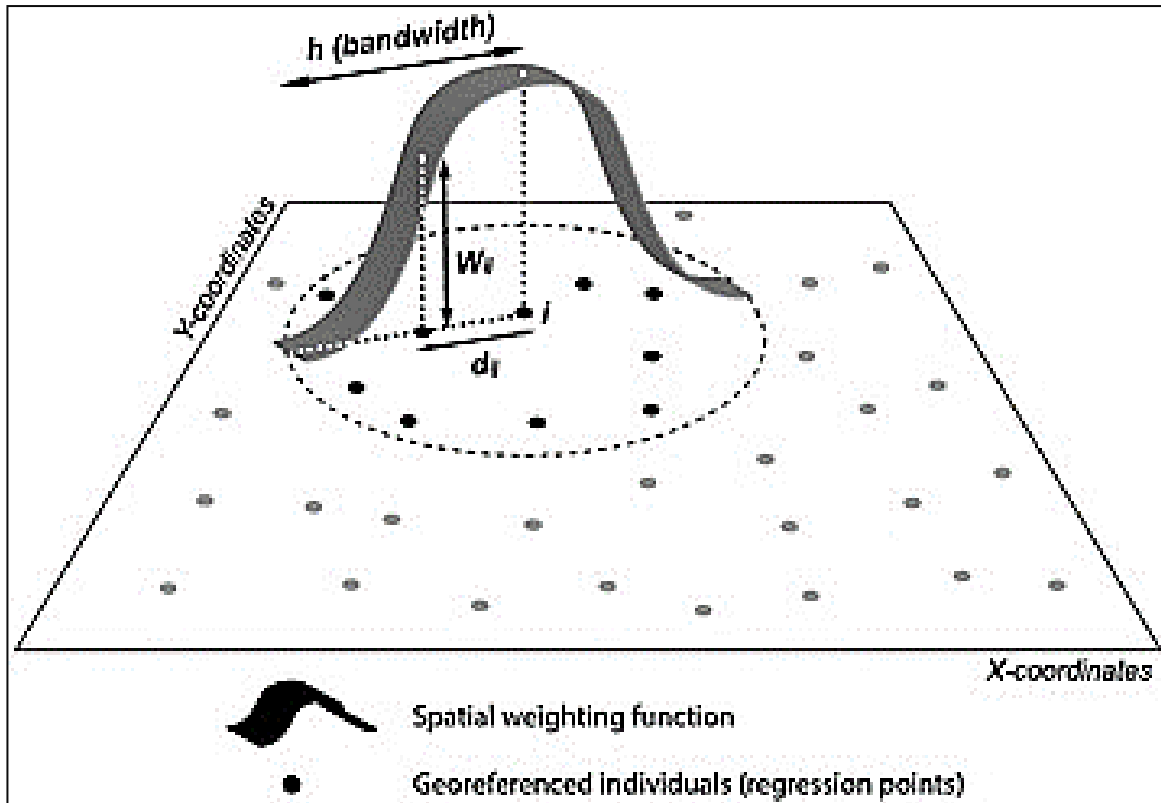


Figure 3.3: Schematic representation of geographically weighted regression and its spatial parameters (Source: Feuillet et al., 2015).

The Akaike Information Criterion with a correction for finite sample sizes (AICc) was used as an indicator for model performance and optimum bandwidth (Fotheringham et al 2003). The AICc, together with golden selection bandwidth were employed as GWPR indicators to complete model calibration. The geographical variability of each varying coefficient was also tested by comparing the fitted GWPR model and a model which only the k th coefficient is fixed. If the fitted GWPR model attains a statistically better fit, the value of model comparison indicator will be smaller than the original GWPR, thereby, the “Diff of Criterion” becomes a positive value, suggesting no spatial variability. Global

Moran's I was employed to examine the spatial autocorrelations of residuals of GWPR model to ensure that non-stationary effects have been accounted for. The GWPR analysis was conducted using the software GWR version 4.0 (GWR4, 2014).

3.6.4 Mapping odds ratios and t-values

After GWPR model calibration, pseudo-t-statistic was calculated to assess the local significance of the parameters. The pseudo-t-statistic was calculated as:

$$t_k(u_i) = \beta_k(u_i)/Se(\beta_k(u_i)) \quad \text{Equation 3.13}$$

Where Se denotes the local standard error of the k th parameter estimate, considering the variation in the data (Vega-Corredor & Opadeyi, 2014). The pseudo- t -values follow a standard normal distribution if the true regression parameter is zero and mapping the values would be beneficial for identifying spatial variations in the relationship between leptospirosis and its explanatory covariates (Vega-Corredor & Opadeyi, 2014). Standard odds ratio will show the sensitivity of leptospirosis cases to a change of one standard deviation in each covariate as:

$$Odds_k(u_i) = \exp(\beta_k(u_i)SD(x_k)) \quad \text{Equation 3.14}$$

Where $SD(x_k)$ is the standard deviation of the k th variable (Vega-Corredor & Opadeyi, 2014). Spatial interpolation was employed to produce a continuous map of odds ratios and pseudo t values across the high incidence region. In this study, Empirical Bayesian kriging (EBK) was employed to smooth the odds ratio and pseudo- t -values across the study area. The details of the EBK interpolation and cross-validation method was explained in section

3.4.1. Spatial data processing and management were conducted in ArcMap and, geostatistical modelling was conducted using geostatistical analyst tool in ArcGIS 10.3 software (Environmental System Research Institute, 2011).

3.7 Chapter summary

This chapter has extensively described the methodology and concepts needed to achieve the objectives of this study. This includes the collection and processing of geospatial data, conceptualising the spatial relationships of georeferenced data, model optimization and cross-validations of estimated odds ratios.

CHAPTER 4

RESULTS

4.1 Epidemiological study of leptospirosis (Objective No. 1)

From January 2011 to December 2018, there were 4,648 cases reported in Sarawak with mean annual cases of 48.35 ± 2.51 cases (95% CI: 43.38- 53.33) and a standard deviation of 24.54. The mean annual incidence rate in Sarawak is 1.73 ± 0.88 (95% CI: 1.56- 1.91) cases per 100,000 population. A higher mean incidence rate was observed in 2018 with 2.57 ± 0.19 (95% CI: 2.15-2.99) per 100,00 population while the lowest mean for annual incidence rate was observed in 2011 with 0.53 ± 0.10 (95% CI: 0.31- 0.76) per 100,000 population. The trend for the incidence rate is presented in Figure 4.1.

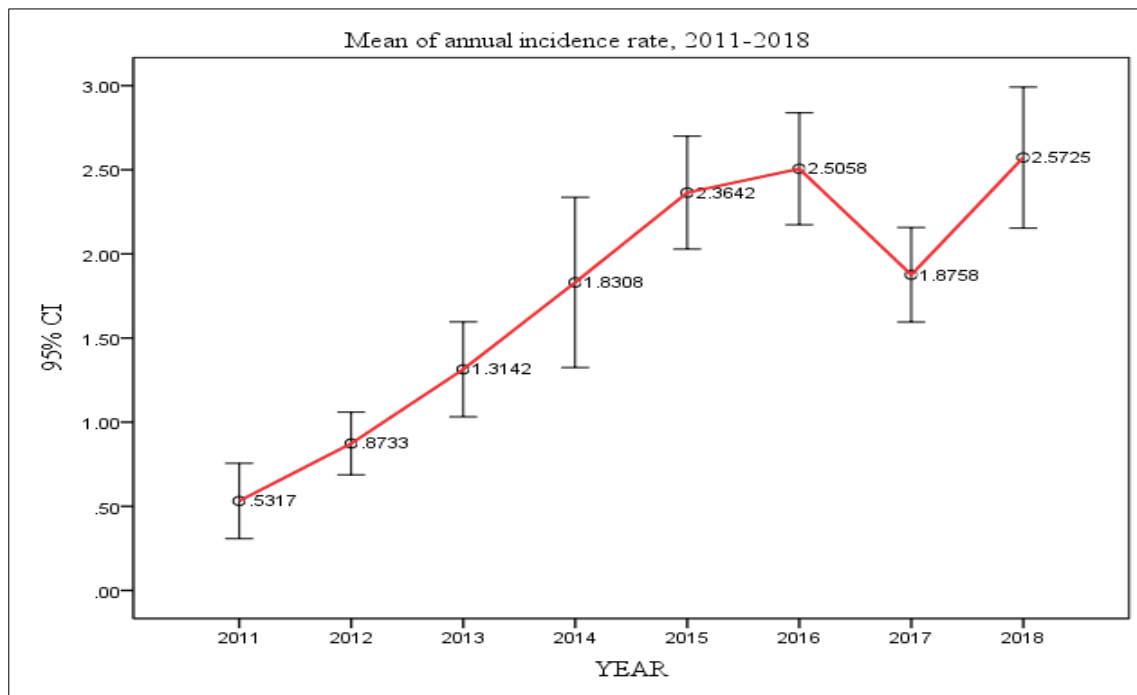


Figure 4.1: Comparison of annual mean incidence rate of leptospirosis in Sarawak. The circles denote the mean values of reported cases while the bar represent the 95% confidence interval. The trend of annual mean incidence rate over the study period is depicted by the red line.

A significant increase in the annual incidence (AI) was noticeable with 6 cases in 2011 to 31 cases per 100,000 population in 2018. The number of fatalities reported during the study period was 139 cases, with an overall case fatality rate (CFR) of 2.99%. There was a reduction in CFR from 2011 (5.62%) to 1.51 % in 2018. Comparisons of CFR among divisions revealed a significantly higher rate in Mukah division with 6.75% ($X^2(11) = 31.68$, $p < 0.01$), followed by Limbang Division with 5.91%. Figure 4.2 demonstrates the decreasing trend of CFR of leptospirosis in Sarawak during the study period.

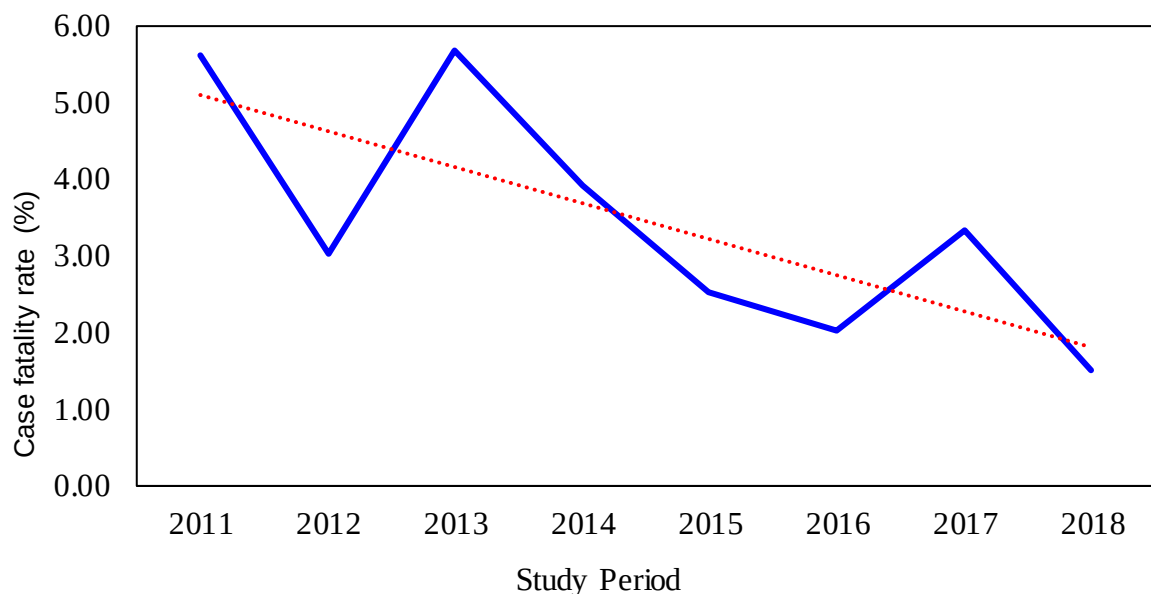


Figure 4.2: Case fatality rate of leptospirosis in Sarawak from 2011 to 2018. The red dotted line denotes the decreasing trend in the case fatality rate (CFR).

4.1.1 Prevalence of leptospirosis based on the gender

There was a significant difference in the incidence rate for male (mean: 35.79, SD:15.40), and female (mean:13.97, SD: 6.71); $t(9.56) = 3.68$, $p < 0.05$. The overall male to female ratio was 3:1, with incidence of 35.79 ± 5.45 (95% CI 22.92- 48.67) and 13.97 ± 2.37 (95% CI: 8.36-19.57) cases per 100,000 population, respectively. Male accounted for 3394 (73.13%) cases as compared to 1254 (26.87%) cases in the female as presented in Table 4.1.

Table 4.1: The prevalence of leptospirosis based on gender.

Group	Total	Mean incidence rate	95% CI	
		($\pm$ std.error)	Lower	Upper
Male	3394	35.79 $\pm$ 5.45	22.92	48.67
Female	1254	13.97 $\pm$ 2.37	8.36	19.57

4.1.2 Prevalence of leptospirosis based on the age group & ethnicity

The result of ANOVA analysis showed a statistically significant difference in the incidence rate among age groups (*Welch F* (8, 25.86) = 2.74, *p* = 0.025). Higher incidence rate was observed among those aged between 50-59 years old with a mean of 27.22 $\pm$ 4.78 (95% CI: 15.91- 38.53) cases per 100,000 population. In contrast, the lowest incidence was observed in the group aged 0-9, with a mean of 5.10 $\pm$ 1.09 (95% CI: 2.52-7.68) cases per 100,000 population. The overall mean age for leptospirosis infection in Sarawak was 37.53 $\pm$ 0.26 (95% CI: 37.01-38.04) years old with a median age of 36 and a standard deviation of 17.94. The summary for the prevalence of leptospirosis based on age groups is presented in Table 4.2.

Table 4.2: Prevalence of leptospirosis based on age groups.

Group	Frequency	Mean incidence rate	95% CI	
		($\pm$ std.error)	Lower	Upper
0 - 9	185	5.10 $\pm$ 1.09	2.52	7.68
10 -19	592	15.20 $\pm$ 2.44	9.44	20.96
20 - 29	954	23.99 $\pm$ 3.38	16.00	31.99
30 - 39	929	25.80 $\pm$ 3.67	17.12	34.48
40 - 49	758	27.14 $\pm$ 4.35	16.85	37.44
50 - 59	618	27.22 $\pm$ 4.78	15.91	38.53
60 - 69	370	25.08 $\pm$ 4.91	13.49	36.67
>70	232	23.24 $\pm$ 5.27	10.78	35.70
NA	10			

NA: Data not available

Prevalence of leptospirosis incidence based on ethnic groups in Sarawak showed that other Bumiputera (Kayan, Kenyah, Orang Ulu, Penan etc) accounted for the highest incidence with an average of 32.21 $\pm$ 6.56 (95% CI: 16.80-47.82) cases per 100,000 population. (*Welch F* (6, 19.47) = 18.58, $p < 0.05$). The lowest incidence was observed in other ethnic group (Indonesian, Indian and non-Malaysians) with 2.46 $\pm$ 0.411 (95% CI: 1.49-3.43) cases per 100,000 population. Table 4.3 summarised the prevalence of leptospirosis in Sarawak based on ethnicity.

Table 4.3: Prevalence of leptospirosis based on ethnicity.

Group	Frequency	Mean incidence	95% CI	
		rate ($\pm$ std.error)	Lower bound	Upper bound
Bidayuh	372	21.09 $\pm$ 3.35	13.16	29.02
Chinese	374	7.43 $\pm$ 1.50	3.88	10.98
Iban	1950	30.11 $\pm$ 5.24	17.71	42.51
Malay	860	16.54 $\pm$ 2.32	11.05	22.03
Melanau	154	13.31 $\pm$ 2.59	7.19	19.44
Other Bumiputera	471	32.21 $\pm$ 6.56	16.8	47.82
Others	37	2.46 $\pm$ 0.411	1.49	3.43
NA	430			

NA: Data not Available

4.1.3 Prevalence of leptospirosis based on occupational types

. Comparison of mean incidence among occupational groups revealed a statistically significant difference (*Welch F* (8, 25.86) = 2.74, $p < 0.025$) among the studied groups. This suggests that occupational type is a significant factor in explaining the demographical discrepancy of leptospirosis infection in Sarawak. Result in the present study identified higher occurrences among people engaged in elementary occupations with a total of 859 cases and a mean incidence rate of 56.93 ± 12.87 (95% CI:26.49-87.37) per 100,000 population. Skilled agricultural, forestry, livestock and fishery reported the lowest case frequency among the occupational types with a total of 110 cases and a mean incidence of

7.54 $\pm$ 2.27 (95% CI: 2.18-12.91) per 100,000 population. The comparison of mean incidence rate for each occupational group during the study period is presented in Table 4.4.

Table 4.4: Prevalence of leptospirosis in Sarawak based on occupational types.

Group	Frequency	Mean incidence rate	95% CI	
		($\pm$ std.error)	Lower bound	Upper bound
Clerical, Craft and support workers	256	13.36 $\pm$ 3.83	4.29	22.42
Elementary occupations	859	56.93 $\pm$ 12.87	26.49	87.37
Plant and machine operators and assembler	155	12.99 $\pm$ 3.03	5.83	20.16
Professionals	176	14.51 $\pm$ 3.97	5.12	23.89
Service and sales workers	210	9.49 $\pm$ 1.76	5.34	13.64
Skilled agricultural, forestry, livestock and fishery	110	7.54 $\pm$ 2.27	2.18	12.91
Technicians & associate professionals	168	24.23 $\pm$ 5.94	10.2	38.27
Not working	262	14.19 $\pm$ 3.62	5.62	22.75
Unemployed	84	25.12 $\pm$ 8.07	6.04	44.19
NA	2362			

NA: Data not available

4.1.4 Spatial distribution of leptospirosis

The mean incidence rate by district varies between 7.92 and 96.63 cases per 100,000 population with the overall mean of 27.45 ± 1.84 (95% CI: 23.82-31.08). There was a statistically significant difference in the mean incidence rate among districts (*Welch F* (29, 69.96) = 4.08, $p < 0.05$). Kapit district demonstrated the highest mean incidence rate of 96.62 ± 27.41 (95% CI: 31.81-161.44) while the lowest incidence was reported by Dalat district with 7.91 ± 9.60 (95% CI: 0.10- 15.94). The choropleth map in Figure 4.3 illustrates the spatial variation in the mean incidence rate of leptospirosis across the study region.

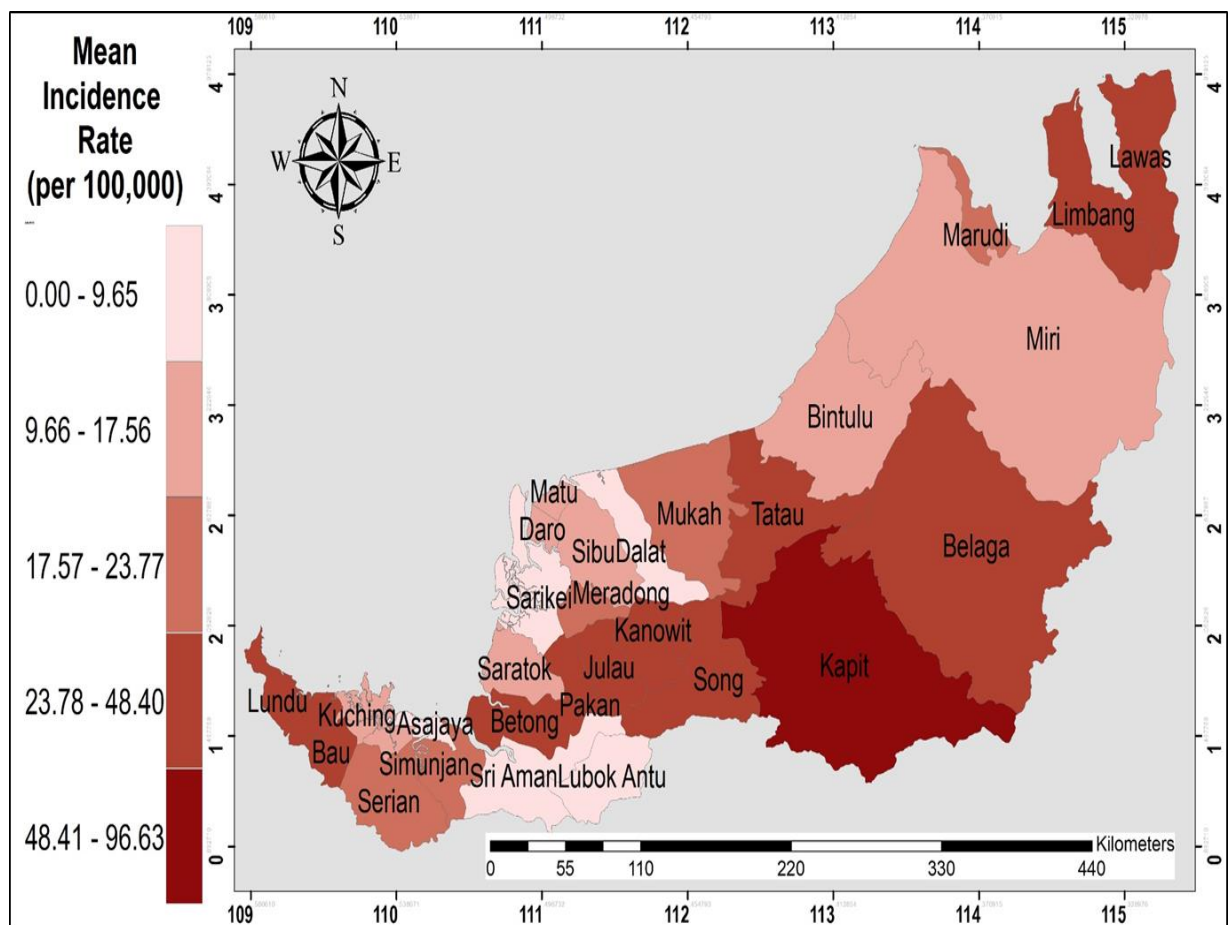


Figure 4.3: A choropleth map illustrating the mean incidence rate among districts in Sarawak during the investigation period, 2011-2018.

4.1.5 Temporal distribution of leptospirosis

The mean value for leptospirosis cases by month oscillates between 41 cases and 56 cases over the study period. January has the highest mean with 55.75 ± 11.24 (95% CI: 29.16-82.34) while April reported the lowest mean with 41.00 ± 8.12 (95% CI: 21.26-60.21). Interestingly, result of the ANOVA analysis revealed that there was no statistically significant difference in the mean value of the monthly cases ($F(11, 84) = 0.31, p=0.98$). This suggests that temporal factor does not influence the annual distribution of leptospirosis incidence in Sarawak. January reported the highest variation during the study period (SD= 31.80) followed by March (SD= 30.70), while December demonstrated the lowest variation with SD= 15.30. The comparison of monthly leptospirosis cases over the study period is presented in Figure 4.4.

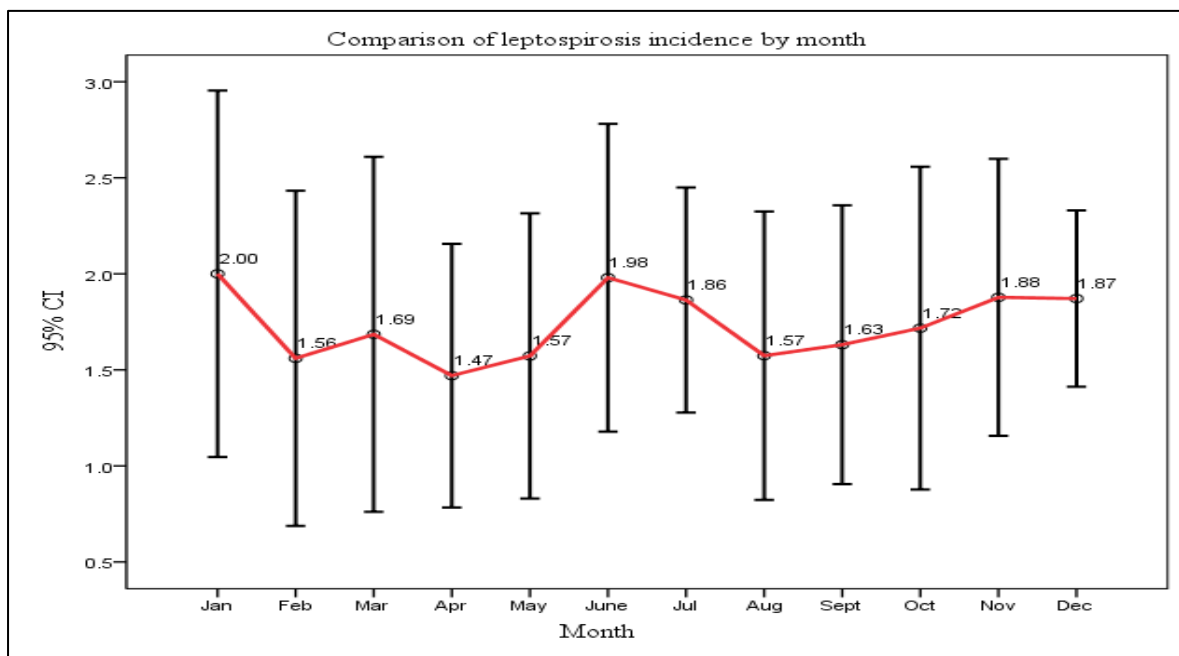


Figure 4.4: The temporal distribution of leptospirosis incidence in Sarawak over they study period. The circles denote the mean values while the bar represent the 95% confidence interval. The red line depicts the pattern of mean incidence rate by month.

4.2 Spatial analysis of leptospirosis (Objective No. 2)

4.2.1 Mapping the spatial trend of leptospirosis

The total number of leptospirosis cases among districts varies between 14 to 648 cases with a mean of 155 cases per district (± 152.02) and a median of 108 cases. The highest number of cases was observed in Kuching district with 648 cases followed by Kapit with 487 cases (Figure 4.5).

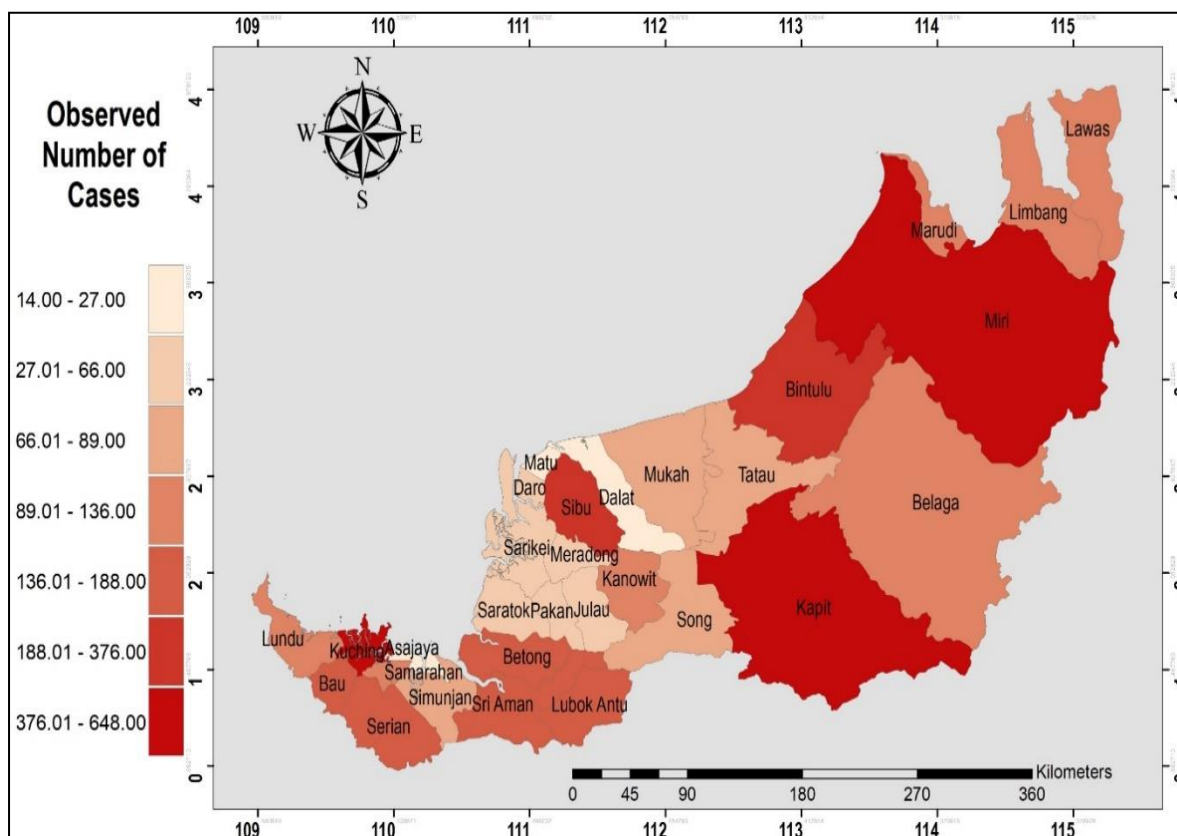


Figure 4.5: The number of reported leptospirosis cases in Sarawak during the study period, 2011-2018.

Dalat district demonstrated the lowest cases among districts with 14 cases and an average of 2 cases per year. However, these values do not represent the true incidence as the number of populations at risk was not taken into consideration. After accounting for the

number of populations at risk, the calculated standardised incidence ratio (SIR), varies between 0.39 and 4.54 with a mean of 1.31 ($\pm$ 1.32). Higher SIR was observed in Kapit district with 4.54, while Dalat district reported the lowest SIR value of 0.39 (Figure 4.6).

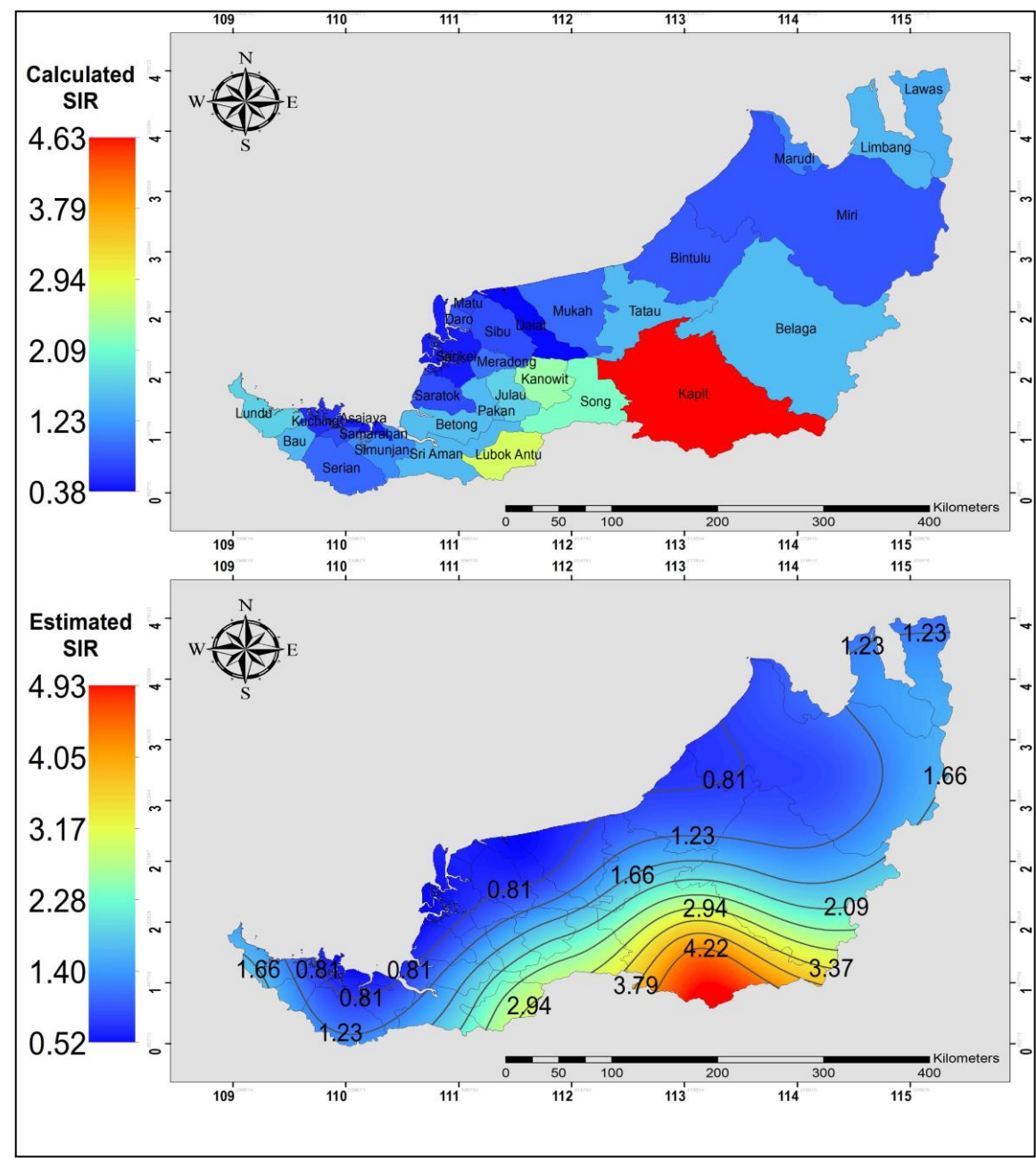


Figure 4.6: Map showing the standardised incidence ratio (SIR) of leptospirosis during the study period, 2011-2018.

While there appears to be a general trend towards the southeast region, as shown in Figure 4.6, the overall geographical pattern of the calculated SIR values is not clear. The isopleth map shows a smoothed geographical distribution of leptospirosis SIR values and facilitates the visualization of patterns across the study area. Lower incidence ratio ($SIR < 1$) is observed in the north, the central and western region of Sarawak, particularly along the coastal region. In contrast, high incidence ($SIR > 1$) were noticeable in the northeast, southwest and central districts. Overall, leptospirosis incidence demonstrated an increasing pattern towards the south east region of Sarawak as denoted by the change in colour from blue to red (Figure 4.6). Based on the isopleth in Figure 4.6, the noticeable increase in SIR values towards the southeast region suggested the presence of spatial disease clusters.

The estimated SIR map was produced using EBK interpolation with K-Bessel detrended semi-variogram model. The isopleth map is presented with an overlaid isolines separated with equal intervals of estimated SIR values. Results of cross-validation reported mean error (ME) and mean standardised error (MSDE) values to be close to zero (-0.02). The root-mean-squared standardised error (RMSSDE) value is comparatively close to 1 (0.99) while the value of mean standard error (MSE) is comparatively closer to the value of root mean square (RMSE) (0.50:0.57). These values suggested that the estimated SIR values based on EBK interpolation were reliable and unbiased. Cross-validation results are summarised in Table 4.5.

Table 4.5: Results of cross-validation for evaluating the performance of EBK interpolation.

Parameters	Values
Mean Error (ME)	-0.02
Mean Standardised Error (MSDE)	-0.02
Mean Standard Error (MSE)	0.50
Root-Mean-Square-Standardised Error (RMSSDE)	0.99
Root-Mean-Square Error (RMSE)	0.57
Mean Error (ME)	-0.02

4.2.2 Spatial autocorrelation leptospirosis incidence

4.2.2.1 Global Moran's *I*

The global Moran's Index for the standardised incidence rate (SIR) revealed a variation in the spatial distribution of leptospirosis from 2011 to 2018 (Table 4.6). Overall, the leptospirosis incidence in Sarawak demonstrated a positive spatial clustering (Moran's *I*: 0.28). Significant positive clustering was observed in 2012 and 2013 with the highest Moran's *I* in 2014 (0.45). Random spatial patterns were reported in the year 2011, 2015, 2016, 2017 and 2018 ($p > 0.05$).

Table 4.6: Summary of global Moran's I analysis for leptospirosis incidence in Sarawak.

Period	Moran's Index	Z	p-value	Pattern
2011	-0.01	0.33	0.30	Random
2012	0.23	2.10	<0.01*	Clustered
2013	0.33	2.61	<0.01*	Clustered
2014	0.45	3.60	<0.01*	Clustered
2015	0.09	1.38	0.10	Random
2016	0.09	1.15	0.13	Random
2017	0.13	1.19	0.12	Random
2018	0.04	0.68	0.23	Random
2011-2018 (Overall)	0.28	2.97	<0.01*	Clustered

*significant at $p < 0.05$

4.2.2.2 Local Indicator of Spatial Autocorrelation (LISA)

Leptospirosis incidence rate was found to be heterogeneous across all 30 districts during the study period (2011-2018). High-High (HH) cluster was not reported in the year 2011, but significant Low-Low (LL) clusters were noticeable in the western districts involving Matu, Daro and Dalat. Song district was identified as a significant high incidence cluster (HH) from 2015 to 2018. Sarikei district was revealed as a significantly low cluster (LL) from 2014 to 2018. Interestingly, a higher number of districts showed significant clustering in 2014 with 4 districts identified as HH clusters and 8 districts as LL clusters.

Besides high and low clusters, spatial outliers of high (HL) and low (LH) incidence rate were also observed in this study. Outliers of high incidence rate (HL) of leptospirosis incidence was reported in 2015 and 2016, which involved Betong and Samarahan district. Tatau district was a significantly low incidence outlier in 2012 and 2017, while in 2013, Kapit was the significant outlier. Overall, HH clusters are predominantly localised in the southeastern districts while clusters of low incidence rate (LL) were identified in the western region involving districts such as Matu and Sarikei. The total number of significant clusters by year is summarised in Table 4.7 and the map of spatial clusters is presented in Figure 4.7.

Table 4.7: Number of spatial clusters identified based on LISA analysis.

Year	HH	LL	HL	LH	Not Significant
2011	0	3	0	0	27
2012	1	5	0	1	23
2013	2	3	0	1	24
2014	4	8	0	0	18
2015	1	1	2	0	26
2016	1	1	1	0	27
2017	1	2	0	1	26
2018	1	2	0	0	27

*HH: High-High type; LL: Low-Low type; HL: High-Low type; LH: Low-High type

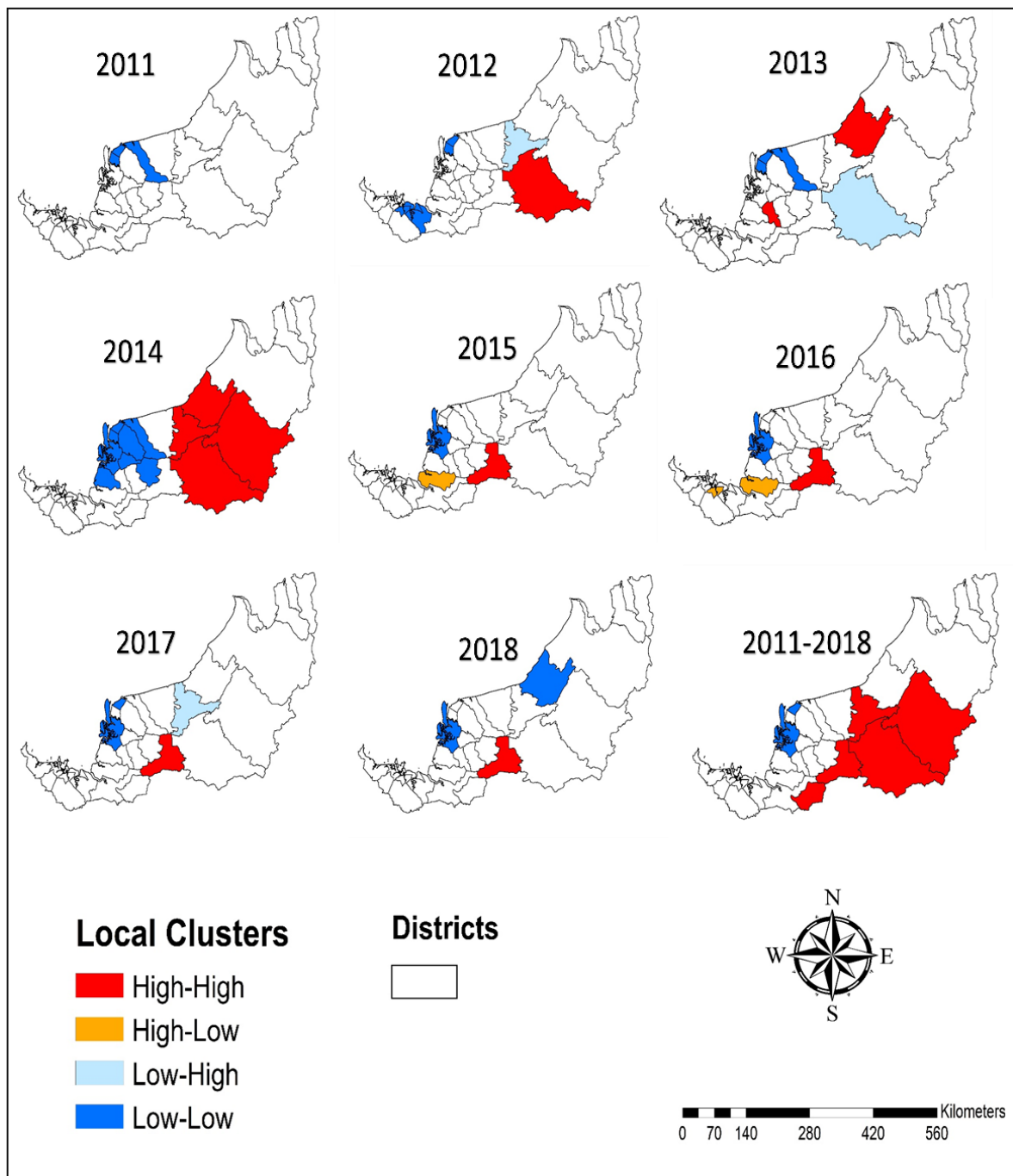


Figure 4.7: Leptospirosis incidence clusters and outliers identified based on LISA analysis from 2011 to 2018. Colour-coded polygons indicated districts with significant spatial clustering, while the blank polygons indicated areas with insignificant spatial clustering.

4.2.3 Hot spot analysis (Getis Ord Gi*)

The result of hot spot analysis delineated a clear spatial pattern of high incidence clusters, concentrated in the southeast of Sarawak. Five districts were identified as spatial hot spot namely, Kapit, Belaga, Song, Lubok Antu and Tatau. These districts demonstrated extreme p-values ($p < 0.01$) with z-scores varies between 1.58 to 2.83. Kapit has the most intense clustering as denoted by the highest z-score of 2.83 with a p-value less than 0.01. Song and Tatau have the second-highest z-scores of 2.53 and 1.97, respectively. The lowest positive z-score among the districts in the hot spot region were observed in Belaga (1.55) and Lubok Antu (1.58). Hot spot analysis also revealed clustering of low incidence denoted as a cold spot for leptospirosis incidence. Samarahan district in the western region exhibited a negative z-score of -1.83 with $p < 0.10$. Other districts in the study area were found to be not statistically significant, with z-scores varies between -1.51 and 0.96. The result of hot spot analysis is presented in Table 4.8 and Figure 4.8.

Table 4.8: Summary of hot spot analysis using Getis-Ord Gi* statistics.

Districts	Z-scores	p-value	Pattern
ASAJAYA	-1.30	0.19	Random
BAU	-0.58	0.56	Random
BELAGA	1.86	0.07*	Clustered
BETONG	0.61	0.54	Random
BINTULU	-0.37	0.71	Random
DALAT	-0.56	0.58	Random

Table 4.8 continued

DARO	-1.47	0.14	Random
JULAU	0.96	0.34	Random
KANOWIT	0.00	1.00	Random
KAPIT	2.83	<0.01***	Clustered
KUCHING	-0.90	0.37	Random
LAWAS	-0.27	0.79	Random
LIMBANG	-0.27	0.79	Random
LUBOK ANTU	1.88	0.06*	Clustered
LUNDU	-0.29	0.77	Random
MARUDI	-0.74	0.46	Random
MATU	-1.51	0.13	Random
MERADONG	-0.38	0.71	Random
MIRI	-0.43	0.66	Random
MUKAH	-0.27	0.78	Random
PAKAN	-0.02	0.98	Random
SAMARAHAN	-1.83	0.07*	Clustered
SARATOK	-0.04	0.97	Random
SARIKEI	-1.11	0.27	Random

Table 4.8 continued

SERIAN	-1.26	0.21	Random
SIBU	-1.44	0.15	Random
SIMUNJAN	-1.29	0.20	Random
SONG	2.53	0.01**	Clustered
SRI AMAN	1.64	0.10	Random

***significant at $p < 0.01$

**significant at $p < 0.05$

*significant at $p < 0.10$

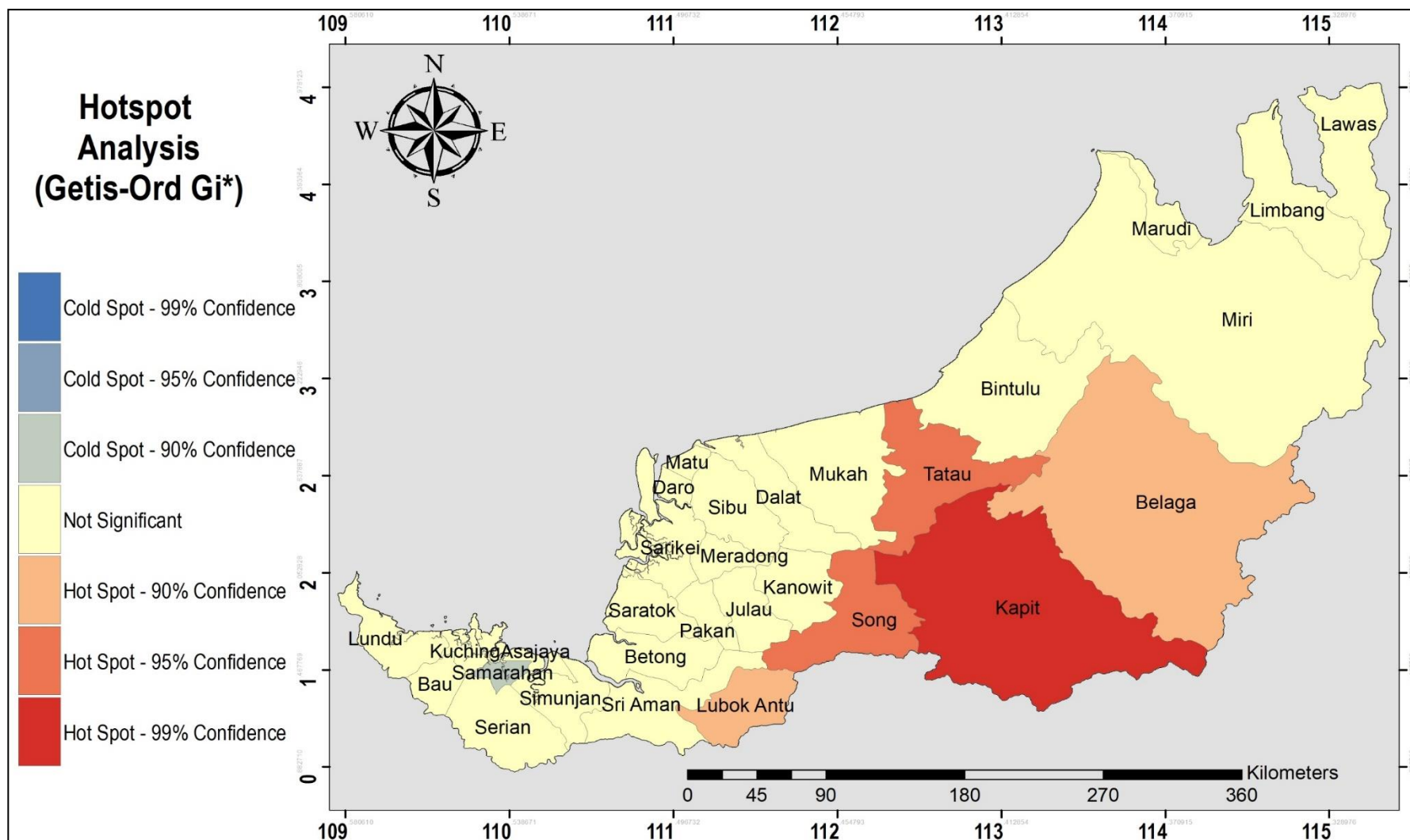


Figure 4.8: Map showing the occurrence of hot spot and cold spot of leptospirosis infection over the study period, 2011-2018.

4.3 Assessment of potential factors associated with the spatial distribution of leptospirosis (Objective No. 3)

4.3.1 Response and explanatory variables

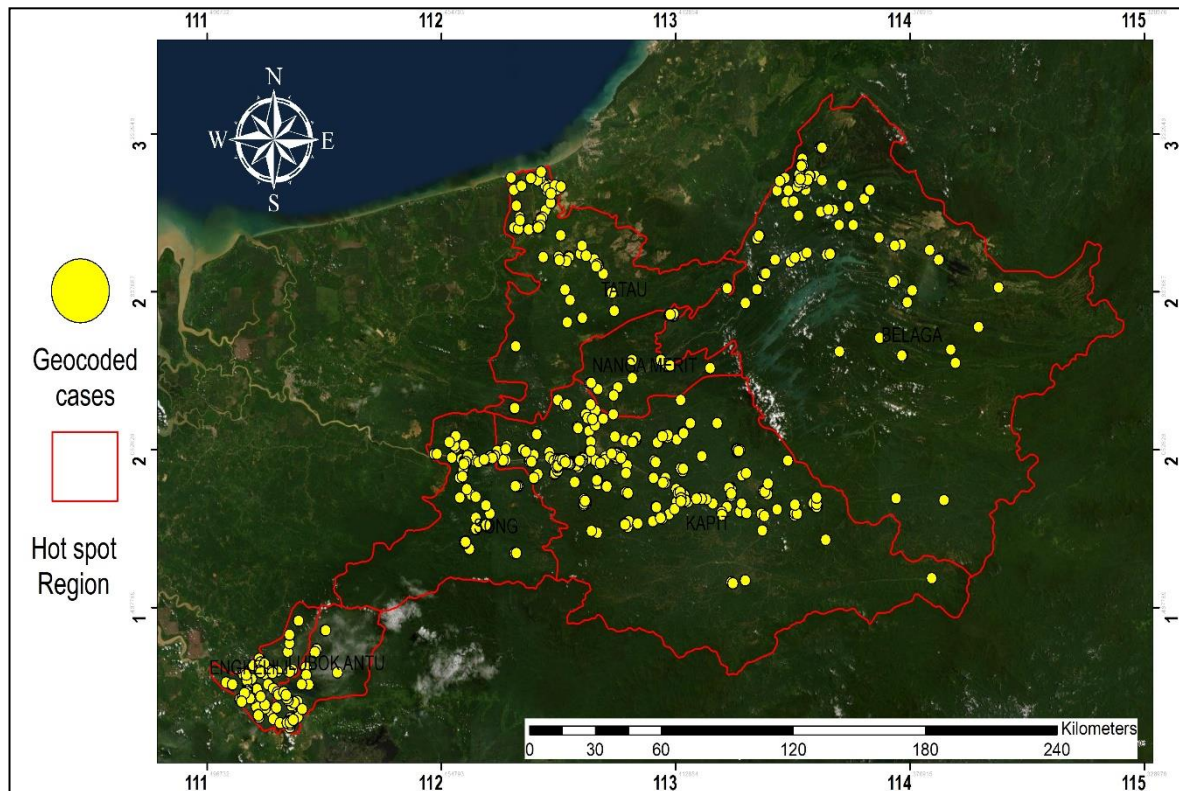


Figure 4.9: The geocoded leptospirosis cases from 2011 to 2018 in the hot spot region (HH), located in the southeast of Sarawak.

A total of 916 cases were reported in the hot spot region (known hereafter as HH region). Figure 4.19 showed the geocoded leptospirosis cases in the hotspot region during the study period (2011-2018). In general, the HH region has an average incidence of 58 cases (per 100,000 population) per year. Males contributed 70% of the total cases with an average incidence of 92 cases (per 100,000) per year. Leptospirosis infection in the HH region was predominantly among people aged between 40-49 years, with an average incidence of 93 cases (per 100,000) per year. The lowest incidence rate was observed in the group aged between 0-9, with an average of 10 cases per year. When comparing the incidence by

ethnicity, the Melanau reported the highest incidence in the HH region with 95 cases (per 100,000) per year. The majority of the cases involved people engaged in elementary occupations, making up 42% of the overall cases. When comparing the frequency of leptospirosis infection by month, the month of June revealed the highest number with an average of 12 cases per year. This was followed by April with 6 cases per year. The lowest number of monthly cases was observed in February with an average of 4 cases per year. In this study, the transmission of leptospires was investigated by modelling their relationship with selected environmental variables. The descriptive statistics of each environmental variable is presented in Table 4.9.

Table 4.9: Descriptive statistics summarising the variables in the hot spot region (HH).

Category	Variables (unit)	Min	Max	Mean	Std.D
Agriculture	LSD (no.)	1.00	83179.60	14684.98	23715.56
	CLD (%)	0.01	0.67	0.24	0.14
Meteorology	WVP (kPa)	2.03	3.00	2.77	0.13
	TS (%)	260.00	478.00	347.99	39.28
	PWQ (mm)	749.00	1785.00	1211.10	142.59
	PWM (mm)	256.00	649.00	428.35	51.91
	PS (%)	10.00	33.00	19.91	3.21
	PDQ (mm)	476.00	1054.00	646.76	81.89
	PDM (mm)	155.00	342.00	205.15	26.11
	ANNP (mm)	215.00	439.00	305.98	31.43

LSD: Livestock Density; CLD: Cropland Density; WVP: Water Vapour Pressure; TS: Temperature Seasonality; PS: Precipitation Seasonality; ANNP: Annual Precipitation; PDQ: Precipitation in Driest Quarter; PDM: Precipitation in Driest Month; PWM: Precipitation in Wettest Month; PWQ: Precipitation in Wettest Quarter

4.3.2 Global Poisson regression (GPR)

4.3.2.1 Univariate Analysis

- i. Association between the spatial distribution of leptospirosis and agriculture variables

A fitting of univariate Poisson regression model with agriculture variables revealed that the distribution of leptospirosis in the high incidence region (HH) was significantly associated with cropland density ($p < 0.05$). Based on Table 4.10, one unit increase in cropland density corresponds to approximately 1-fold ($\exp(\beta) = 1.09$) increase in the frequency of leptospirosis. Fitting livestock density to the univariate model did not show significant association with leptospirosis as indicated by $p > 0.05$. Hence, it suggests that livestock density does not explain the distribution of leptospirosis in the HH region.

- ii. Association between the spatial distribution of leptospirosis and meteorological variables

Regarding meteorological variables, water vapour pressure (WVP), annual precipitation (ANNP), precipitation in the driest month (PDM), precipitation in the driest quarter (PDQ) and precipitation in the wettest quarter (PWQ) were found to be significant factors in affecting the distribution of leptospirosis in the HH region (Table 4.10). WVP has a positive influence in comparison with the other meteorological covariates based on the parameter estimate of 0.12. This indicates that one unit increase in WVP will correspond to 1-fold ($\exp(0.12) = 1.11$) increase in the number of leptospirosis case count. In contrast, ANNP, PDQ, PDM and PWQ has a negative association with leptospirosis as denoted by the negative parameter estimates of -0.08, -0.13, -0.17, and -0.09, respectively. Such values indicate that as these variables increase by one unit, the distribution of leptospirosis would decrease by approximately 1-fold. Three meteorological variables did not reveal statistical

significance to leptospirosis in this study: temperature seasonality (TS), precipitation seasonality (PS), and precipitation in the wettest month (PWM).

All significant variables in univariate analysis (6 variables) were included in the formulation multivariate global Poisson regression model. Multicollinearity among the predictor variables was checked based on variance inflation factor (VIF), before the analysis. Variables with VIF values greater than 4 were eliminated, and the best predictor combinations were selected based on stepwise regression. The final variables (i.e., with $VIF < 4$ and low AIC values) such as vapour pressure (WVP), annual precipitation (ANNP), precipitation in the driest month (PDM), and precipitation in the wettest quarter (PWQ) were included in the analysis. The summary of univariate assessment is presented in Table 4.10.

Table 4.10: Summary of univariate assessment by fitting single variable using global Poisson model.

Variables	Variable types (unit)	β	Std. error	$\exp(\beta)$ (95%)	$\Pr(> z)$
	Intercept (α)	0.56	0.034	1.74 (1.63 – 1.86)	<0.01*
Agriculture	LSD (no.)	0.06	0.03	1.06 (0.99 - 1.13)	0.07
	CLD (%)	0.09	0.03	1.09 (1.02 - 1.16)	<0.01*
Meteorology	WVP (kPa)	0.12	0.04	1.12 (1.04 - 1.21)	<0.01*
	TS (%)	-0.01	0.03	0.99 (0.93 - 1.06)	0.86
	ANNP (mm)	-0.08	0.03	0.92 (0.86 - 0.99)	0.02*
	PS (%)	0.05	0.03	1.05 (0.98 - 1.12)	0.145
	PDQ (mm)	-0.13	0.04	0.88 (0.81 - 0.94)	<0.01*
	PDM (mm)	-0.17	0.04	0.85 (0.79 - 0.91)	<0.01*
	PWM (mm)	-0.03	0.03	0.97 (0.91 - 1.04)	0.42
	PWQ (mm)	-0.09	0.03	0.91 (0.85 - 0.98)	<0.01*

*significant at $p < 0.05$; α : intercept; β : parameter estimates; exp: exponential function; LSD: Livestock Density; CLD: Cropland Density; WVP: Water Vapour Pressure; TS: Temperature Seasonality; PS: Precipitation Seasonality; ANNP: Annual Precipitation; PDQ: Precipitation in Driest Quarter; PDM: Precipitation in Driest Month; PWM: Precipitation in Wettest Month; PWQ: Precipitation in Wettest Quarter.

4.3.2.2 Multivariate analysis

The results of global Poisson regression indicated that leptospirosis in the HH region was significantly associated with higher water vapour pressure ($\beta = 0.12$), higher annual precipitation ($\beta = 0.14$), lower precipitation in the driest month ($\beta = -0.16$) and lower precipitation in the wettest quarter ($\beta = -0.12$) (Table 4.11).

Table 4.11: Summary of parameter estimates for multivariate GPR model.

Variables	β	Std. error	Exp(β) (95% CI)	Pr(> z)	VIF
intercept	0.54	0.03	1.72 (1.61 - 1.84)	<0.01*	
WVP (kPa)	0.12	0.05	1.12 (1.02 - 1.25)	0.02*	1.95
ANNP (mm)	0.14	0.06	1.15 (1.02 - 1.31)	0.02*	3.05
PDM (mm)	-0.16	0.06	0.85 (0.75 - 0.96)	0.01*	3.36
PWQ (mm)	-0.12	0.07	0.88 (0.77 - 1.00)	0.05*	2.84

*significant at $p < 0.05$; β : variable coefficients; Exp: exponential function; WVP: Water Vapour Pressure; ANNP: Annual Precipitation; PDM: Precipitation in Driest Month; PWQ: Precipitation in the wettest quarter.

Following the model fitting, the multivariate Poisson model can be expressed as follows:

$$\hat{O}_i = \exp (0.54 + 0.12 [\text{WVP}] + 0.14 [\text{ANNP}] - 0.16 [\text{PDM}] - 0.12 [\text{PWQ}])$$

Where the predicted leptospirosis cases are denoted by ($\hat{O}_i$). From Table 4.11, all parameter estimates displayed statistical significance ($|z| > 1.96$ at the 95% confidence level) in the multivariate GPR model. When all variables are held constant (intercept only), two

leptospirosis case would still be expected ($\exp(0.54) = 1.72$). The positive parameter estimates observed in WVP and ANNP indicated that as water vapour pressure increases, leptospirosis occurrence also increases. Similarly, as annual precipitation increases in the high incidence region, so does the occurrence leptospirosis. In contrast, PDM and PWQ show a negative association with leptospirosis as denoted by the negative parameter estimates. This suggests that as PDM decreases, the occurrence of leptospirosis in the high incidence region is expected to increase. Likewise, an increase in leptospirosis infection is also anticipated as PWQ in the HH region decreases.

Factors which were found to be significantly associated with leptospirosis in the univariate models but not included in the final multivariate model are cropland density (CLD) and precipitation in the driest quarter (PDQ). Inclusion of these variables did not lead to a significant improvement of the model fit suggesting strong correlations with other factors included in the final model. The global Poisson model reported a residual deviance of 509.61 at $df = 490$, which was substantially lower than the null deviance with 539.58 at $df = 494$. Goodness of fit for the final model was tested using deviance chi-squared test and was found not to be significant ($p = 0.26$), an indication that the model fit the data well.

4.3.3 Local Poisson regression

4.3.3.1 Geographically Weighted Poisson Regression, GWPR

GWPR model was develop using the variables WVP, PDM, PWQ and ANNP to explore their spatial variabilities with leptospirosis. There was a significant improvement in the model fitting using GWPR in comparison to the global model, as indicated by the reduced AICc value. The AICc value decreases from 519.73 in the global model (GPR) to 506.32 (fixed Gaussian), 484.05 (fixed Bi-square), 443.49 (adaptive Gaussian) and 456.03 (adaptive

Bi-square). The smallest decrease in AICc value was observed when a fixed kernel with Gaussian function was employed with 13.41 difference from the AICc of the global model. In contrast, the highest difference in AICc was noticeable when adaptive kernel with Gaussian function as used with 76.04. In practice, if the difference between AICc values of the two models is less or equal to two, there is no significant difference in the performance of them (Cheng et al. 2011). The selection of the best-fitted model was completed by following the criteria of the minimum AIC estimator (MAICE) (Nakaya et al., 2005). The model with adaptive kernel and Gaussian bandwidth function showed the lowest AICc with a bandwidth size of 43 km, indicating a better fit to the observed data. The percentage of deviance explained also revealed a substantial difference from 0.05 in the global model, to 0.32 in the local model. These values indicated that the GWPR model explains better the relationship between environmental factors and leptospirosis by allowing these relationships to varying spatially. By allowing the parameter estimates to vary spatially in the GWPR model, the model residuals (observed – predicted) did not demonstrate spatial autocorrelations (global Moran's $I = -0.003$, $p = 0.48$). This indicated that local model fits the data well. The comparison of the global model with local model is summarised in Table 4.12.

Table 4.12: Comparison of global (GPR) and local (GWPR) models statistics.

Model	Diagnostics			Optimum		Deviance
	AICc	AIC	BIC/MDL	Bandwidth (km)	Deviance	Explained (%)
Global	519.73	519.61	640.63	NA	509.61	0.06
Fixed Gaussian*	506.32	505.98	541.96	1.00	488.86	0.09
Fixed Bi-square*	484.05	481.72	578.30	1.03	484.06	0.19
Adaptive Gaussian*	443.49	437.48	591.23	43.00	364.35	0.32
Adaptive Bi-square*	456.03	444.40	655.62	98.00	343.93	0.36

*local model (GWPR)

AICc; corrected Akaike Information Criterion; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion, MDL: Minimum Description Length

4.3.4 Mapping odds ratios and t-values

The sensitivity in the frequency of leptospirosis cases in response to changes of one standard deviation of each environmental covariates is presented as odds ratios (OR). Among the variables fitted to the GWPR model, water vapour pressure (WVP) has the most variation in effect across the hot spot region with a standard deviation of 0.62 (LQ: 0.15; UQ: 0.99) (Table 4.13). The negative value of -4.14 in the difference of criterion for WVP confirmed that the variable has spatial variability. In contrast, the least variation in effect was observed in annual precipitation (ANNP) with a standard deviation of 0.14 (LQ: 0.11; UQ: 0.30) (Table 4.13). The spatially heterogeneous relationships between leptospirosis and environmental covariates were visualised by mapping the odds ratios and pseudo-t-values.

Table 4.13: Summary statistics of variable coefficients based on adaptive Gaussian kernel model.

Variable	Odds Ratio			Pseudo-t-values			DIFF of
	Robust SD	LQ	UQ	Min	Max	SD	Criterion
WVP	0.62	0.15	0.99	-1.25	5.85	1.64	-4.14
PWQ	0.28	-0.23	0.15	-2.28	3.44	1.44	3.57
PDM	0.35	-0.51	-0.04	-5.01	1.07	1.65	0.93
ANNP	0.14	0.11	0.3	-0.63	2.49	0.71	7.06

WVP: Water Vapour Pressure; PWQ: Precipitation in the Wettest Quarter; PDM: Precipitation in the Driest Month; ANNP: Annual Precipitation; SD: Standard Deviation; LQ: Lower Quartile; UQ: Upper Quartile; Min: Minimum; Max: Maximum

Water vapour pressure has the most variation in the odds ratios as indicated by higher values in the central region ($OR > 1$) and lower values in the northern and northeast region ($OR < 1$). Based on Figures 4.10, odds ratio higher than 1 ($OR > 1$) were concentrated in the central part of the hot spot region, particularly in urbanised areas of Song and Kapit district. High pseudo-t-values (4.18-5.85) observed in the similar area suggested that the association between WVP and leptospirosis in this area is strong. The north and northeastern part of the hot spot region is least sensitive to the changes in WVP as denoted by the low odds ratio ($OR < 1$). The relationship was also very weak as low pseudo-t-values were observed in the region (-1.25- 1.66).

Regarding the relationship between leptospirosis with precipitation in the driest month (PDM), the northeast region was more sensitive to the changes in PDM as denoted by higher odds ratios (0.16-0.38), particularly in Belaga district (Figure 4.11). The pseudo-t-values between 0.40-1.07 suggested a more substantial relationship in this region as compared to other areas. Interestingly, the central region of Kapit district also showed a high pseudo-t-values (0.40-1.07), indicating that the region has a significant association with PDM. The northern and central regions were least affected by the changes in PDM. These regions showed lower odds ratios between -1.21 to -0.21 with small pseudo-t-values between -5.01- -1.54 as compared to other localities (Figure 4.11). Overall, the effect of PDM on the spatial distribution of leptospirosis is very low as denoted by odd ratios lower than 1 ($OR < 1$) across the study area.

Precipitation in the wettest quarter has little effect on the spatial distribution of leptospirosis as indicated by odds ratios less than 1 ($OR < 1$). The central region, particularly in Kapit district, showed relatively higher odds ratios (0.27-0.94) as compared to other areas

(Figure 4.12). The association in this region was found to be stronger with higher pseudo-t-values between 1.29-3.44. Lower odd ratios were identified in the southwest region of Song district, but the low pseudo-t-values suggested that the relationship was very weak. Among the variables fitted to the GWPR model, annual precipitation (ANNP) has the least variation in spatial effect. This is shown by the low odd ratio between -0.15 to 0.60 ($OR < 1$) with the lowest standard deviation (SD: 0.14) as compared with other covariates. The southwest region has substantially higher odd ratios ($OR: 0.47-0.60$), particularly in Lubok Antu district (Figure 4.13). However, the association between leptospirosis and ANNP was found to be moderate in this region (pseudo-t: 0.41-1.18). Higher association were observed in the central and southeast region as indicated by high pseudo-t-values between 1.19-2.49. The odd ratios and pseudo-t-values for each covariate were interpolated by using Empirical Bayesian kriging. The resulting isopleth maps are presented in Figures 4.14 and 4.15, respectively. The mean prediction error for all parameters was close to 0 with the root mean-standardized-prediction error close to 1. The values for root mean square error and average standard error were small. The prediction error statistics for the interpolated parameters are presented in Table 4.14.

Table 4.14: Prediction error statistics for the parameters interpolated by EBK interpolation.

Variables	Parameters	Mean	Root Mean Square	Mean std	Root Mean Square Std	Average Standard Error
WVP	Odds Ratio	0.01	0.03	-0.01	1.06	0.03
	t-values	-0.01	0.17	-0.03	0.98	0.12
PWQ	Odds Ratio	-0.01	0.03	-0.01	0.96	0.04
	t-values	-0.01	0.08	-0.01	0.96	0.13
PDM	Odds Ratio	0.01	0.04	-0.1	1.03	0.03
	t-values	0.01	0.15	0.01	1.06	0.16
ANNP	Odds Ratio	-0.01	0.08	-0.02	1.04	0.09
	t-values	-0.01	0.04	-0.01	0.96	0.06

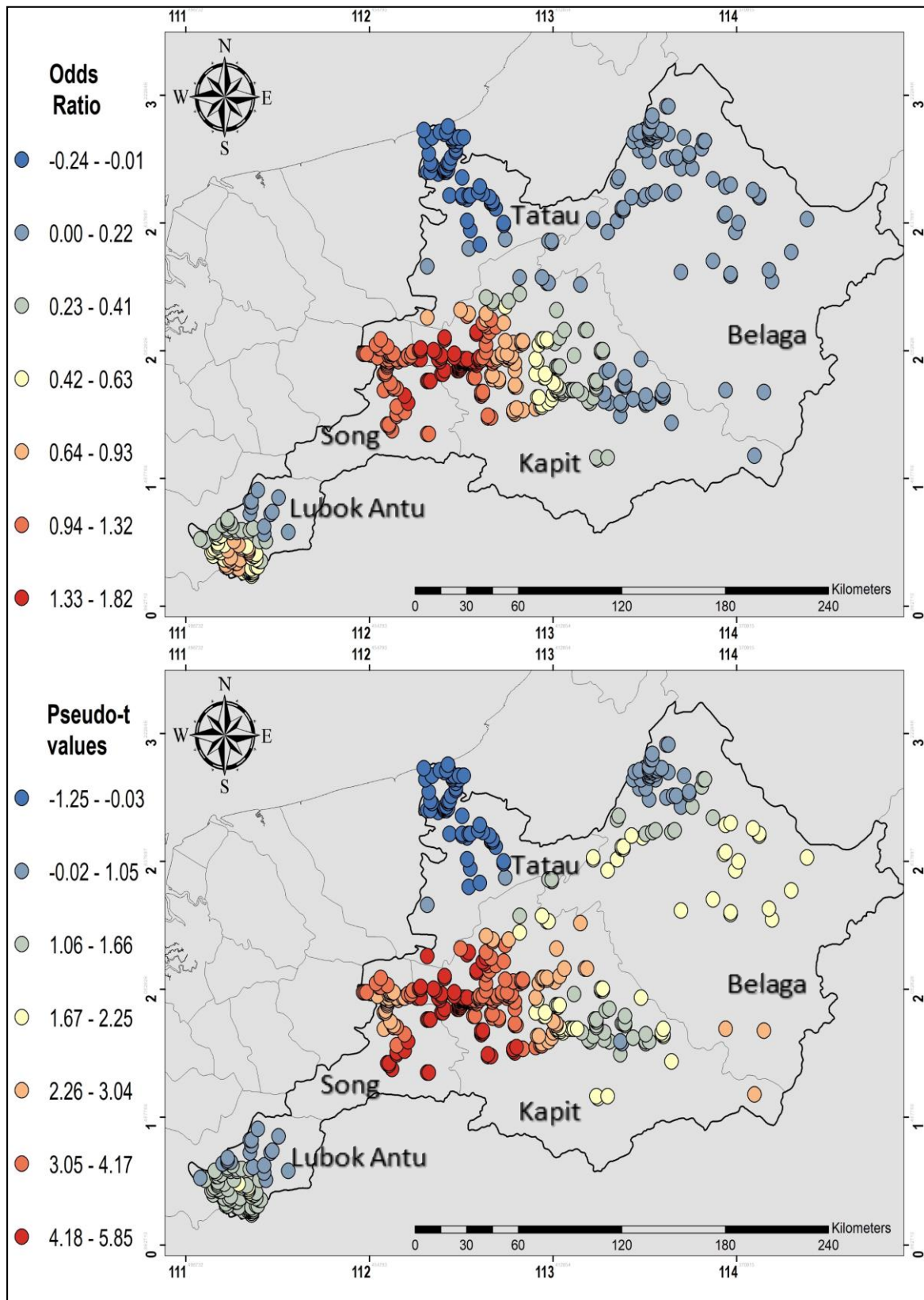


Figure 4.10: The geographical distribution of odds ratios and pseudo-t-values for water vapour pressure (WVP).

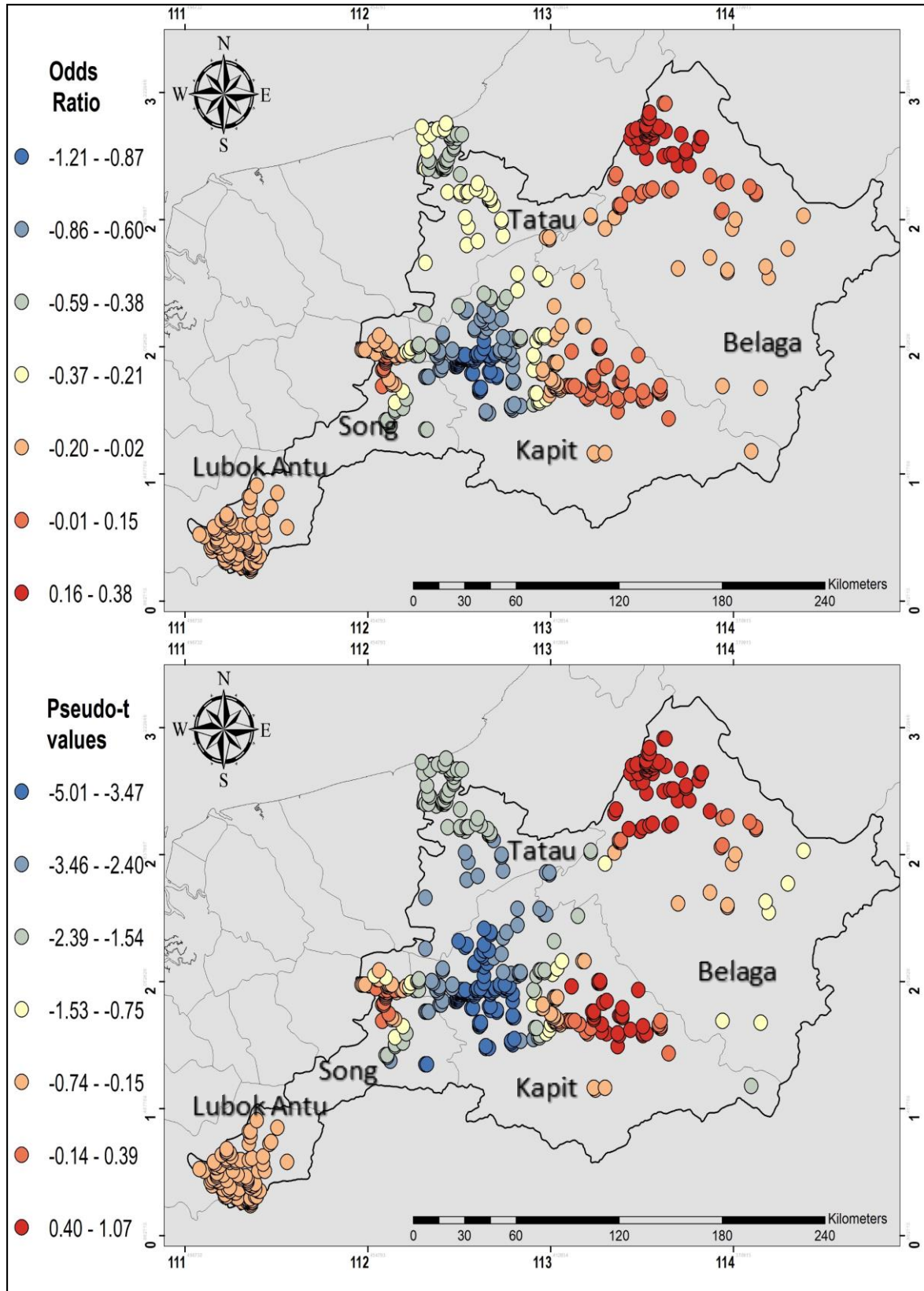


Figure 4.11: Spatially heterogeneous odds ratios and pseudo-t-values for the variable precipitation in the driest month (PDM).

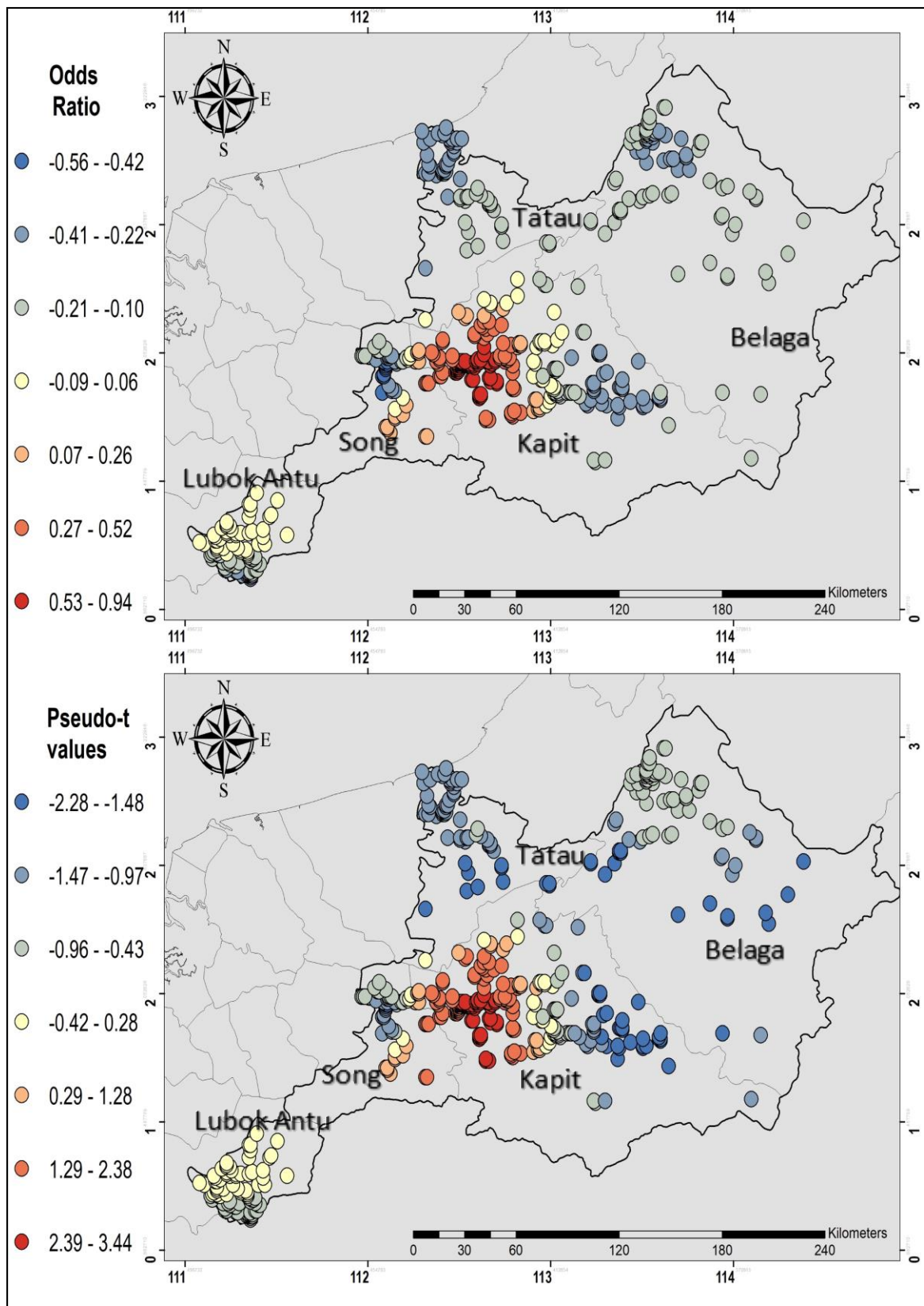


Figure 4.12: The geographical distribution of odds ratios and pseudo-t-values for the variable precipitation in the wettest quarter (PWQ).

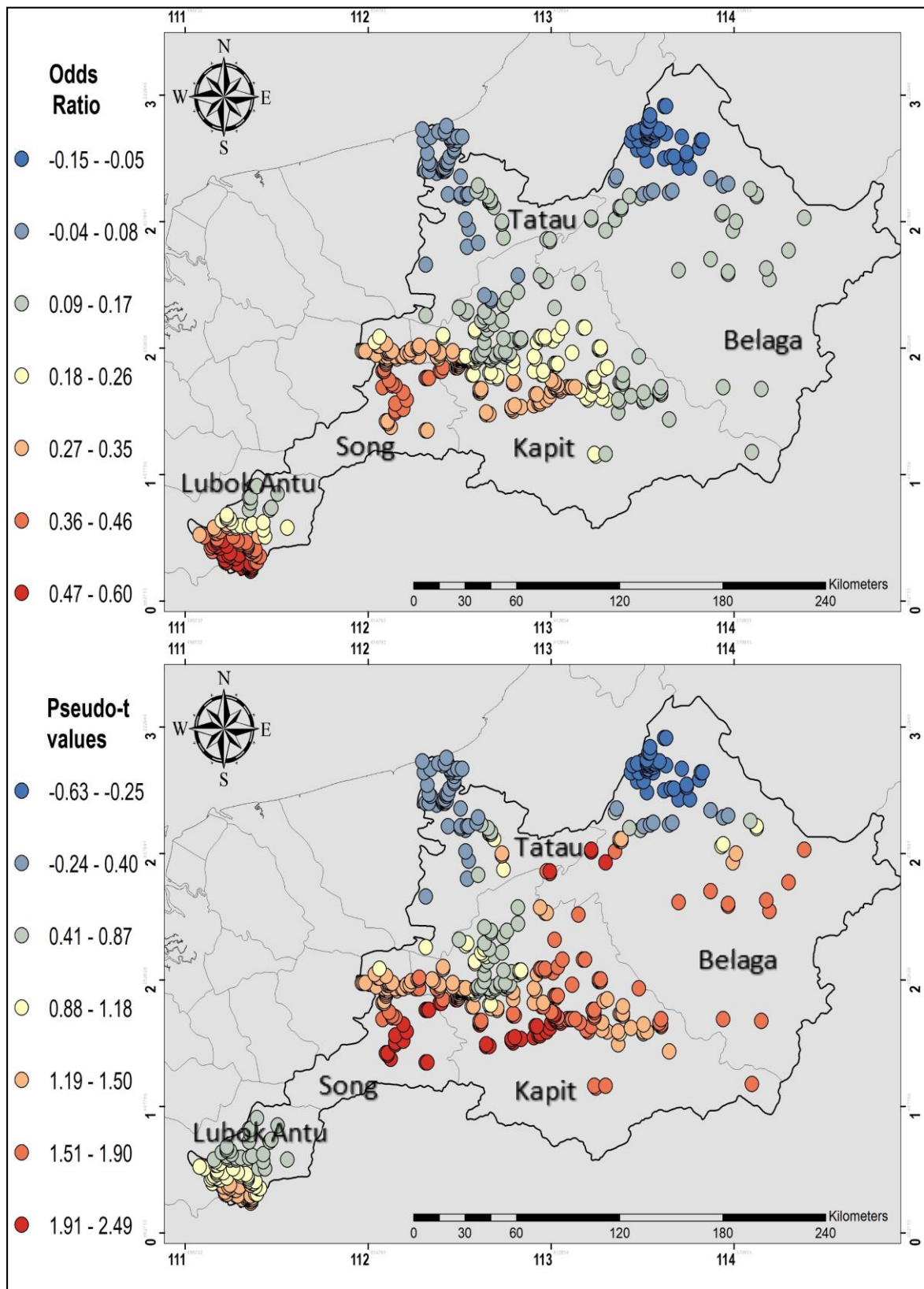


Figure 4.13: Spatially varying odds ratios and pseudo-t-values for the variable annual precipitation (ANNP).

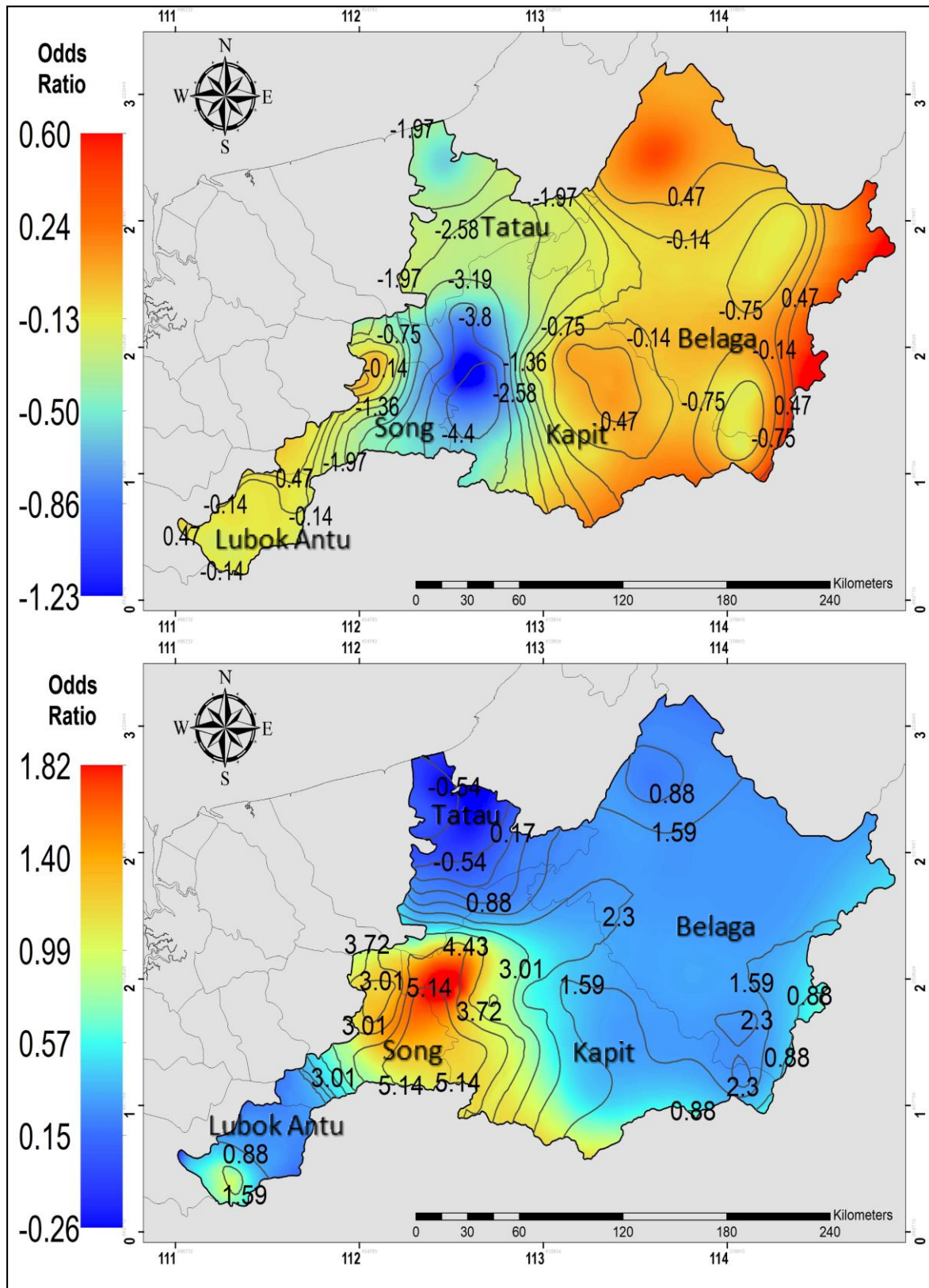


Figure 4.14: The smoothed odd ratios for the variable precipitation in the driest month (PDM; top figure) and water vapour pressure (WVP; lower figure). The isopleth maps were overlaid with contour lines representing pseudo-t-values.

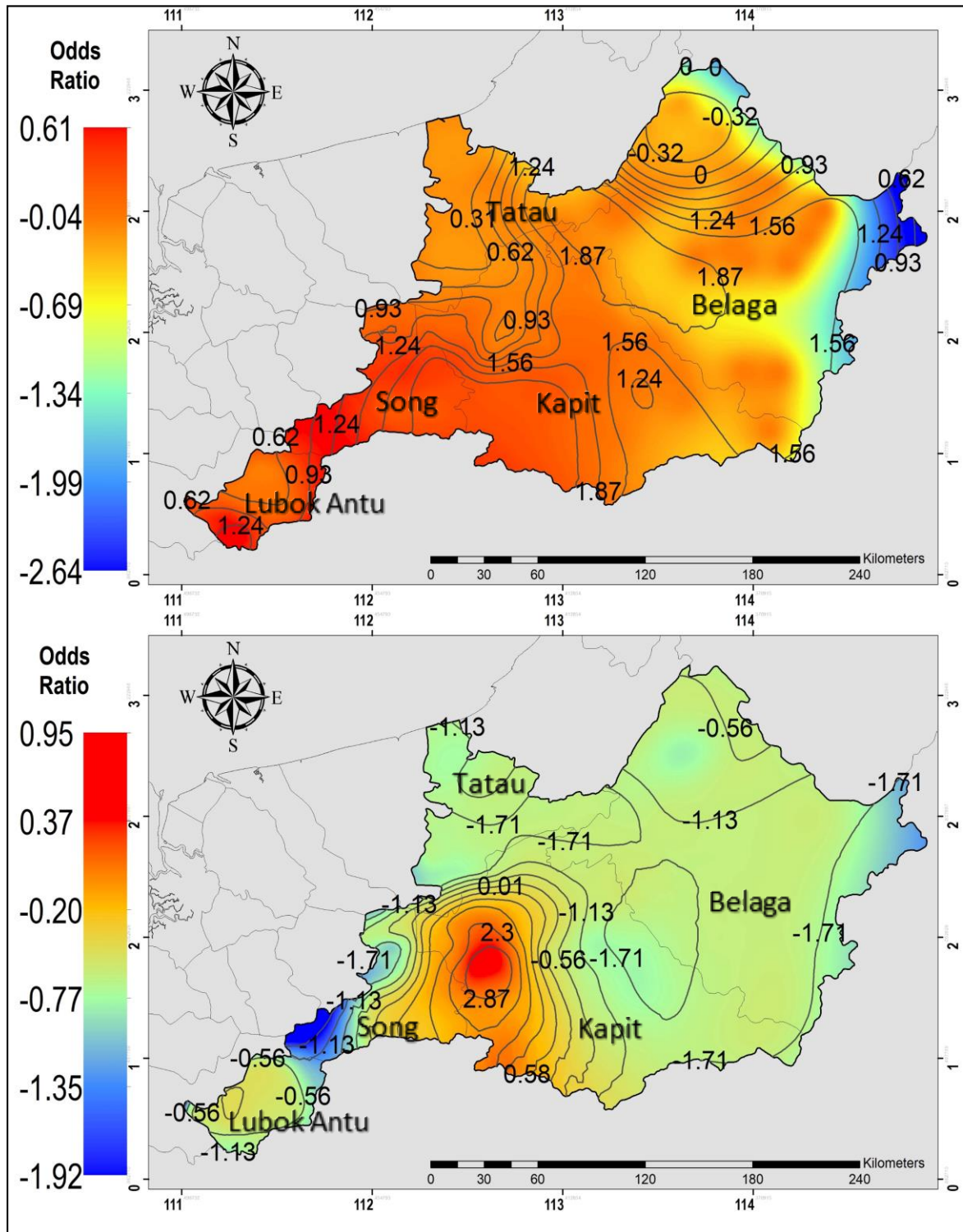


Figure 4.15: Isopleth map illustrating the smoothed odd ratios for the variable annual precipitation (ANNP; top figure) and precipitation in the wettest quarter (PWQ; lower figure). Pseudo-t-values were overlaid as contour lines denoting the strength of association between leptospirosis and the environmental covariates.

CHAPTER 5

DISCUSSION

The present study revealed a substantial increase in the average incidence of leptospirosis from 2011 to 2018. This indicated that leptospirosis remains a public health challenge in the state of Sarawak. The average incidence rate reported in this study was relatively higher as compared to previous studies by Benacer et al. (2016) and Tan et al. (2016) with an average of 3.61 and 15.44 cases per 100,000 population, respectively. Such differences could be due to study period variability, in which Benacer et al. (2016) analysed data from 2004 to 2012, whereas Tan et al. (2016) included data for 2012 and 2013. Furthermore, the actual figure before 2010 could be subjected to underreporting as leptospirosis was only gazetted as a notifiable disease in Malaysia in 2010 (Ministry of Health Malaysia, 2011).

Reduction in case fatality rate reported in the current study can be attributed to the progressive efforts made by local health authorities in the surveillance and management of this disease. The demographical discrepancy in the prevalence of leptospirosis by gender, age and ethnic group can be associated with the occupational risk factor. Based on the result of the epidemiological study, elementary occupational type reported the highest incidence of leptospirosis in Sarawak during the study period. This includes agricultural, forestry, farming and fishery labourers, cleaners, helpers and street related sales and service workers. The majority of the cases in the high incidence demographical groups: males (37.64%), age group 50-59 years old (47.08%) and other Bumiputera ethnic group (37.96%), were associated with an elementary occupation type. Individuals engaged in elementary occupational type are prone to *Leptospira* infection mainly because protective clothing and

equipment are not being worn. For instance, the use of rubber boots and gloves by farmers in a wet paddy field could reduce the likelihood of *Leptospira* spp. to enter open wounds and establish infection. A study by Ridzuan et al. (2016) reported a high prevalence among plantation workers and deliberated the presence of rats, dampness of plantation areas and the frequent contact with soil and polluted water as contributing factors. According to Ridzuan et al. (2016), the risk of transmission is very high, especially when protective clothing is not worn. Other studies in Malaysia have also reported high leptospirosis incidence especially among male-dominated occupations such as paddy planters, rubber tappers, and palm oil workers, labourers dealing with sewage, draining, cleaner, forestry, those involved in antimalarial programs, and military personnel (Tan, 1973; Chou et al., 2008; Shafei et al., 2012; Goris et al., 2013; Benacer et al., 2016; Tan et al., 2016).

Another key question investigated in this study was about the spatio-temporal patterning of leptospirosis in Sarawak over the study period. The occurrence of a high number of cases are expected to be clustered in the area surrounding pathogen sources, and when a disease aggregates in a cluster, then one or more of its causes or sources of infection must aggregate as well (Hassan et al., 2018). Overall, the southeast region of Sarawak is at the highest risk concerning leptospirosis as this region was identified as a significant hot spot region (Figure 4.8). This finding corresponds to Su'ut et al. (2016) where high seroprevalence was reported in the Rejang Basin, located in the central and southeast of Sarawak (Su'ut et al., 2016). According to Su'ut et al. the rapid development in the Rejang Basin through dam construction, the opening of the vast plantation, as well as agrarian environment and other subsistence activities performed by the local communities could have facilitated the interaction between human and the rodents' host and leptospires-contaminated environment (Su'ut et al., 2018).

The Rajang river which flows from northeast to southwest of Sarawak passing Belaga, Kapit, Song and Sibu and eventually to the South China Sea, has played a significant role in sustaining the life of the communities along the river. It serves as an important source of fish protein, freshwater, food and transportation (Tawan et al., 2019). There are numerous longhouses, agricultural sites, forest plantations and logging camps along the river. Deforestation and timber harvesting have contributed to the deterioration of the surface water quality especially during high river discharge associated with rainfall event (Tawan et al., 2019). As this happens, the loose soil from the forest floor can easily be transported by rainwater and run-off into the nearby river, potentially carrying *Leptospira* spp. and contaminating the water bodies. In addition, the hydropower dam construction also contributed to several environmental problems which include flooding of large areas for the formation of reservoirs (Tawan et al., 2019). These circumstances may contribute towards the profound leptospirosis incidence rate observed in this area as compared to other regions in Sarawak.

The complex spatial-temporal pattern in this study can also be linked to the existence of local characteristics that are commonly shared by the clustered region. For instance, behavioural risk factors of the local population may have promoted disparities in the distribution of leptospirosis detected in the current study. An example of such behavioural risk factor is the involvement in water-based recreational and agricultural activities, which gives rise to frequent contact with water bodies such as lakes or rivers. These circumstances facilitate the transmission of pathogenic *Leptospira* spp. from the environment and elevate the risk of infection among susceptible hosts. In 2016, a study by Su'ut and colleagues revealed a significantly higher *Leptospira* seropositivity among respondents in Kapit as compared to those living in Sibu and Sarikei (Su'ut et al., 2016). According to Su'ut et al.

(2016), the demographics of Kapit residents who live in settlements along the Rejang River, and their involvement in agro-based and subsistence economic activities like fishing and farming are likely to contribute to this result.

A higher risk of *Leptospira* infection has been shown in communities that have frequent contact with water bodies (Philip et al., 2013). Leptospirosis outbreaks have been reported following exposure to contaminated water and soil during sporting activities like jungle water rafting (Haake et al., 2002) and swimming (Sejvar et al., 2003;) and hiking (Wynwood et al., 2014). A study in China reported that counties with a higher risk of leptospirosis have a more significant proportion of younger population, males and farmers as compared to low-risk counties (Dhewantara et al., 2018). In their study, Dhewantara et al. (2018) revealed that high-risk counties were also economically less developed and more rural with higher precipitation (Dhewantara et al., 2018). These findings underscored the significance of demographical and socio-behavioural disparity as modulators to the dynamics of leptospirosis.

A localised influence of environmental factors could also explain the geographically inconsistent patterns of leptospirosis observed in this study. The global Poisson model revealed that the geographical distribution of leptospirosis in the hot spot region was concentrated in areas with higher water vapour pressure (WVP) and annual precipitation (ANNP), with lower precipitation in the driest month (PDM) and the wettest quarter (PWQ). A localised relationship of the covariates on the spatial density of leptospirosis was also captured by the local model (GWPR). It is showed that the relationship of WVP and ANNP with leptospirosis were stronger in the central (for WVP) and southwest (for ANNP) as indicated by odd ratios (Figure 4.14 & 4.15). The variable WVP was found to exhibit spatial

variability (-4.14 difference of criterion), thereby suggesting that its influence towards leptospirosis is more localised. The positive relationship indicates that as water vapour pressure and annual rainfall increases, the spatial density of leptospirosis cases also increases. This is in agreement with the previous study in Sarawak, where high annual rainfall in the Rejang Basin and seasonal flooding during monsoon season have contributed towards the high seroprevalence in these areas (Su'ut et al., 2016).

By definition, water vapour pressure is the pressure exerted by water vapour molecules in the air (Bagirov et al., 2017). As the pressure increases, water vapour condenses into water droplets, thereby promoting more precipitation. A study by Wang et al. (2016) has demonstrated a positive correlation between annual mean precipitation with mean water vapour pressure (Wang et al., 2016). As this happened, it encourages the formation of small lakes, muddy ponds, and streams, which can prolong the survival of *Leptospira* in the natural environment. In addition, continuous rainfall also unclogs drains and flush out *Leptospira* in the soil and becomes a source of contamination especially when the runoff flows into the river or water bodies near human settlements. This may result in higher concentrations of *Leptospira* in the media to which humans are exposed to like mud, river or sewer water. Furthermore, continuous rainfall often leads to a natural disaster such as flooding. Flooding has been extensively deliberated as a profound risk factor for the transmission of *Leptospira* (Matsushita et al., 2018; Radi et al., 2018; Togami et al., 2018). Other countries like the Reunion Island (Desvars et al., 2011), Republic of Korea (Joshi et al., 2017), the Philippines (Sumi et al., 2017), Brazil (Cunha et al., 2019) and China (Dhewantara et al., 2019) have also reported a similar positive relationship between leptospirosis with humidity and precipitation.

Conventionally, the amount of precipitation has a positive relationship with leptospirosis (Victoriano et al., 2009; Chadsuthi et al., 2012; Benacer et al., 2016). However, the result of this study found that the spatial density of leptospirosis cases to be concentrated in areas with lower precipitation in the driest month (PDM) and the wettest quarter (PWQ). Although this is counter-intuitive, the relationships were very weak, as indicated by the small pseudo-t-values (Figure 4.14 & Figure 4.15), thereby suggesting that the relationship vary geographically. A decrease in precipitation may have influenced human activities and animal mobility during the dry and wet season. The dry period in Sarawak is associated with the southwest monsoon, which extends from May to September while the wet season is between November to February (Dindang et al., 2013). A decrease in the amount of precipitation reduces surface water availability and may encourage water-based activities for humans and animals like swimming, bathing and drinking (Lau et al., 2010). This would promote contact between humans, livestock and wildlife through more intense sharing of the depleting surface water sources. For instance, in some remote areas of Sarawak, the access to clean water is limited and the local community often relies on rivers and streams as their primary source of water. Such practice increases the risk of acquiring *Leptospira* spp. especially when the water bodies have been contaminated by the excreta of infected animals.

Lower precipitation could also lead to the intensification of agricultural activities. For example, in Sarawak, the month of July and August are often associated with land preparation phase for rice cultivation (Hoki, 1977). This may facilitate the interaction between humans, rodents, and the leptospires-contaminated environment. A study in Central Java, Indonesia has reported the significance of harvesting season as a critical period associated with leptospirosis transmission (Widiastuti et al., 2019). Jones et al. (2013) have also emphasized that encroachment of human settlements and agriculture on natural

ecosystems has led to the expansion of ecotones (transition zones between adjacent ecological systems), where species assemblages from different habitats mix. This provides new opportunities for pathogen spillover, genetic diversification, and adaptation. For instance, the increase in human population along with economic and technological advancement has led to an anthropogenic modification to the natural environment. This includes the opening of vast plantations and livestock farming for food security and urbanisation to support population expansion. The dynamics associated with these anthropogenic changes have led to the alteration of land and water interfaces, which may act as vector breeding sites. As this happens, it encourages more intensive interactions among humans, livestock, and wildlife. Livestock could potentially become an intermediate host for pathogens to evolve and spill over into humans, or humans can be infected directly from wildlife or vectors. In each host species, there is a vast array of symbiotic microorganisms which may be pathogenic to a new host species or evolve to become pathogenic. A new transmission cycle may be established if the newly evolved pathogen is transmissible into another new host species. Associations between disease emergence and ecotones have been suggested for several diseases, including leptospirosis (Jones et al., 2013). Human settlements and opening of vast agricultural farms in ecotonal areas provide abundant nesting, feeding, and breeding sites for the vectors of leptospirosis like rodents. For instance, when livestock rearing is juxtaposed next to a forest, disease like leptospirosis may spillover to domesticated livestock and be amplified before additional spillover into human populations.

The local model (GWPR) has explained a better relationship between environmental factors and leptospirosis by allowing the relationships to vary spatially. This is supported by the reduction in the AICc value of the local model in comparison to its global counterpart.

A similar observation was reported by Vega-Corredor & Opadeyi (2014), whereby the GWPR model explained a better relationship between hydrological factors and leptospirosis on the island of Trinidad. In their study, a stronger underlying local relationship was observed when rainfall was excluded from the model as evidence from the change in bandwidth size (Vega-Corredor & Opadeyi, 2014). Standard global modelling techniques, such as ordinary least square (OLS) linear regression cannot detect nonstationarity and may obscure regional variation in the relationships between predictors and response variables (Matthews & Yang, 2012).

In GWR, the spatial structure is accounted for and estimates a discrete model and local parameter estimates for a unique geographical location using a different weighting scheme (Matthews & Yang, 2012). The use of GWR in the spatial modelling of leptospirosis have also been deliberated in other studies in countries like Indonesia (Widayani et al., 2016), Fiji (Mayfield et al., 2016)), India (Pawar et al., 2018), and Iran (Mohammadinia et al., 2019). The model with adaptive Gaussian kernel function exhibits a more localised effect as evidence seen in the reduction in bandwidth size. As bandwidth increases, a GWPR model is likely to tend towards the GPR model due to the increased smoothing of weights given to the data points. Also, estimates may become biased due to the increasing distances of data from the regression point (Fotheringham et al., 2003; Vega-Corredor & Opadeyi, 2014).

Several policy implications could be drawn from the findings of this study. In high incidence region (hot spot), intervention strategy should be focused on investing in public health education coverage. This way, the local community will have a better understanding of the risks of leptospirosis and the necessary actions needed to reduce transmission. As high incidence clusters were localised in remote areas of Sarawak, efforts can be directed in

promoting more competent care by making healthcare services more accessible to the remote community with low socioeconomic status. This includes the establishment of efficient transportation options and locating healthcare facilities in areas where services are lacking.

The spatially varying relationships between climate factors and leptospirosis implied that each covariate should be examined separately when tailoring healthcare delivery and programs to a particular area. For instance, monitoring and surveillance of leptospirosis should be prioritised in areas where water vapour pressure and annual precipitation has the strongest influence, as a high volume of rainfall is expected. Such circumstance suggested the risk of flooding especially in areas where the community are living near a large water body like a river. As for precipitation in the driest month and wettest quarter, a negative association with leptospirosis was revealed from the findings. A reduced in the amount of precipitation would lead to a shortage of water supplies especially in a remote region of Sarawak, where rivers and ponds were the main sources of water for the locals. This calls upon a need for district water board and stakeholders to collaborate in establishing a clean water system to minimise the reliance of local community on natural water sources for daily use.

While the data presented in this study can be beneficial in identifying high-risk areas for more directed surveillance and interventions, some study limitations need to be considered. The present study only incorporated confirmed leptospirosis cases in the analyses. However, there could be unreported cases who experienced subclinical infection, and undiagnosed, especially in resource-limited rural areas of Sarawak. Also, the missing data in the dataset, particularly on occupational data (Not Available: 2033 cases) would contribute to biasesness in the statistical analysis and impedes the identification of high-risk

occupational groups. This could underestimate the role of occupation on the actual incidence of leptospirosis in Sarawak. While this study acknowledged the use of GWPR to study the non-stationary relationships, it is important to consider that ecological study such as this was not able to identify the causation. Hence, future research is required to investigate the risk factors identified as they relate to the ecology of *Leptospira* and the pathology of leptospirosis.

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

This study has integrated multidisciplinary approaches to study the spatial characteristics of leptospirosis and its underlying environmental determinants. The epidemiological analysis has revealed significant disparities in the demographical distribution of leptospirosis throughout the districts of Sarawak. A higher incidence rate was observed among males, age groups of 50-59 years, those engaged in elementary occupations and the other Bumiputera ethnic group. These groups were found to be significant in the inferential statistics, thereby confirming that human leptospirosis in Sarawak demonstrated distinctive socio-demographical disparities.

The results of global and local Moran's *I* analysis have found that leptospirosis in Sarawak was clustered and exhibit a spatio-temporal patterning throughout the study period. Overall, the southeast region of Sarawak is at the highest risk concerning leptospirosis as evidence from the trend surface map, and the hot spot analysis confirmed this region as part of the hot spot cluster. Hence, this leads to the conclusion that there are significant spatial clustering of leptospirosis at the district level in Sarawak, confirming the first hypothesis of this study.

The present study also has identified significant associations between leptospirosis and meteorological variables, thereby supporting the second hypothesis of this study. This is evidently showed by the result of regression modelling where the geographical distribution of leptospirosis in the hot spot region was concentrated in areas with higher water vapour

pressure and annual precipitation, with lower precipitation in the driest month and the wettest quarter. Interestingly, the use of geographically weighted Poisson regression (GWPR) model have revealed that the relationships were spatially heterogeneous. The spatially varying odds ratios and pseudo-t-values supported the final hypothesis of this study, whereby the relationships that exist between explanatory variables and the incidence of leptospirosis are spatially heterogeneous across the high-risk region.

6.2 Recommendations

In the present study, the spatial modelling of leptospirosis with environmental variables were focused in the hot spot region as it was part of the study objective. However, future works should also model the relationship in other districts to provide a more comprehensive understanding of the influence of environmental variables on the distribution of leptospirosis in Sarawak. As leptospirosis is a multifactorial disease, it is imperative to include other variables such as hydrology and topography into the analysis. A study incorporating time-series analysis could provide more robust analysis, especially when the seasonality of environmental variables is considered. The techniques presented in this study could also be applied to future works regarding public health and microbial ecology.

REFERENCES

- Abela-Ridder, B., Sikkema, R., & Hartskeerl, R. A. (2010). Estimating the burden of human leptospirosis. *International Journal of Antimicrobial Agents*, 36(11), 5-7.
- Adler, B., & de la Peña Moctezuma, A. (2010). *Leptospira* and leptospirosis. *Veterinary Microbiology*, 140(4), 287-296.
- Ahmad, I., Arsin, A. A., & Thaha, R. M. (2018). A Spatial Analysis of Malaria Cases Using Geographic Information System (GIS) Application in Gorontalo Regency in 2016. In *Proceedings of the 2nd International Conference on Medical and Health Informatics*, Tsukuba, Japan, June. Association for Computing Machinery.
- Alias, H., Surin, J., Mahmud, R., Shafie, A., Zin, J. M., Nor, M. M., & Rundi, C. (2014). Spatial distribution of malaria in Peninsular Malaysia from 2000 to 2009. *Parasites & Vectors*, 7(1), 186-192. <https://doi.org/10.1186/1756-3305-7-186>.
- Al-shere, T. A., Ujiie, M., Suzuki, M., Salva, E., Belo, M. C. P., Koizumi, N., & Villarama, J. B. (2012). Outbreak of leptospirosis after flood, the Philippines, 2009. *Emerging Infectious Diseases*, 18(1), 91-94.
- Alves, A. T., Nobre, F. F., & Waller, L. A. (2016). Exploring spatial patterns in the associations between local AIDS incidence and socioeconomic and demographic variables in the state of Rio de Janeiro, Brazil. *Spatial and Spatio-temporal Epidemiology*, 17(5), 85-93.
- André-Fontaine, G. (2006). Canine leptospirosis—do we have a problem? *Veterinary Microbiology*, 117(1), 19-24.

- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2), 93-115.
- Anselin, L. (2003). GeoDa 0.9 user's guide. *Spatial Analysis Laboratory (SAL). Department of Agricultural and Consumer Economics, University of Illinois, Urbana Champaign, IL.*
- Azhar, Z. I., Jusoh, A., Abdul Rahim, S. S. S., Hassan, M. R., Safian, N., & Shah, S. A. (2016). Temporal spatial distribution of dengue and implications on control in Hulu Langat, Selangor, Malaysia. *Dengue Bulletin*, 39(16), 19-31.
- Bahaman, A. R., & Ibrahim, A. L. (1988). A review of leptospirosis in Malaysia. *Veterinary Research Communications*, 12(2), 179-189.
- Barragan, V., Olivas, S., Keim, P., & Pearson, T. (2017). Critical Knowledge Gaps in Our Understanding of Environmental Cycling and Transmission of *Leptospira* spp. *Applied and Environmental Microbiology*, 83(19), 1-10. <https://doi.org/10.1128/AEM.01190-17>.
- Bergquist, R., & Rinaldi, L. (2010). Health research based on geospatial tools: a timely approach in a changing environment. *Journal of Helminthology*, 84(1), 1-11.
- Ben, I., Lozynskyi, I., & Semenyshyn, O. (2018). Experience of GIS Technology Application in the Surveillance of Tick-Borne Infections. *Online Journal of Public Health Informatics*, 10(1), 1-2. <http://doi.org/10.5210/ojphi.v10i1.8898>.
- Benacer, D., Thong, K. L., Min, N. C., Verasahib, K. B., Galloway, R. L., Hartskeerl, R. A., & Zain, S. N. M. (2016). Epidemiology of human leptospirosis in Malaysia, 2004-2012. *Acta Tropica*, 157(5), 162-168.

- Bernard, W. V., Bolin, C., Riddle, T., Durando, M., Smith, B. J., & Tramontin, R. R. (1993). Leptospiral abortion and leptospiruria in horses from the same farm. *Journal of the American Veterinary Medical Association*, 202(8), 1285-1286.
- Bharti, A. R., Nally, J. E., Ricaldi, J. N., Matthias, M. A., Diaz, M. M., Lovett, M. A., & Vinetz, J. M. (2003). Peru-United States Leptospirosis Consortium. Leptospirosis: a zoonotic disease of global importance. *Lancet Infectious Disease*, 3(12), 757-771.
- Bhardwaj, P., Kosambiya, J. K., & Desai, V. K. (2008). A case control study to explore the riskfactors for acquisition of leptospirosis in Surat city, after flood. *Indian Journal of Medical Sciences*, 62(11), 431-438.
- Castillo-Neyra, R., Zegarra, E., Monroy, Y., Bernedo, R. F., Cornejo-Rosello, I., Paz Soldan, V. A., & Levy, M. Z. (2017). Spatial association of canine rabies outbreak and ecological urban corridors, Arequipa, Peru. *Tropical Medicine and Infectious Disease*, 2(3), 38-46.
- Caruso, C., Accardi, G., Virruso, C., & Candore, G. (2013). Sex, gender and immunosenescence: a key to understand the different lifespan between men and women? *Immunity and Ageing*, 10(1), 20-22.
- Chadsuthi, S., Modchang, C., Lenbury, Y., Iamsirithaworn, S., & Triampo, W. (2012). Modeling seasonal leptospirosis transmission and its association with rainfall and temperature in Thailand using time-series and ARIMAX analyses. *Asian Pacific Journal of Tropical Medicine*, 5(7), 539-546.
- CHELSEA. (2018). Climatologies at high resolution for earth's land surface areas, version 1.2. [online] Available at < <http://chelsea-climate.org/> > [Assessed on 25th October 2018].

- Cheng, E. M., Atkinson, P. M., & Shahani, A. K. (2011). Elucidating the spatially varying relation between cervical cancer and socio-economic conditions in England. *International Journal of Health Geographics*, 10(1), 51-67.
- Chou, Y. L., Chen, C. S., & Liu, C. C. (2008). Leptospirosis in Taiwan, 2001-2006. *Emerging Infectious Diseases*, 14(5), 856-857.
- Cliff, A.D. & Ord, J.K. (1981) *Spatial processes: Models and applications*, London: Pion Limited.
- Comber, A. J., Brunsdon, C., & Radburn, R. (2011). A spatial analysis of variations in health access: linking geography, socio-economic status and access perceptions *International Journal of Health Geographics*, 10(1), 44-54.
- Cook, E. A., de Glanville, W. A., Thomas, L. F., Kariuki, S., Bronsvoort, B. M., & Fèvre, E. M. (2017). Risk factors for leptospirosis seropositivity in slaughterhouse workers in western Kenya. *Occupational and Environmental Medicine*, 74(5), 357–365. <https://doi.org/10.1136/oemed-2016-103895>.
- Crump, J. A., Morrissey, A. B., Nicholson, W. L., Massung, R. F., Stoddard, R. A., Galloway, R. L., & Muiruri, C. (2013). Etiology of severe non-malaria febrile illness in Northern Tanzania: a prospective cohort study. *PLoS Neglected Tropical Diseases*, 7(7), 1-8. <https://doi.org/10.1371/journal.pntd.0002324>.
- Cunha, M., Costa, F., Ribeiro, G. S., Carvalho, M. S., Reis, R. B., Júnior, N. N., & Wunder, E. A. (2019). Rainfall and other meteorological factors as drivers of urban transmission of leptospirosis. *BioRxiv*, 14(6), 1-32. <https://doi.org/10.1101/658872>.

- Daud, A., Fuzi, N. M. H. M., Arshad, M. M., Kamarudin, S., Mohammad, W. M. Z. W., Amran, F., & Ismail, N. (2018). Leptospirosis seropositivity and its serovars among cattle in Northeastern Malaysia. *Veterinary World*, 11(6), 840-844.
- Department of Statistics Malaysia. (2019). *Statistics Yearbook Sarawak*, 2018. [e-book] Kuching: Department of Statistics Malaysia, Sarawak. Available through: Department of Statistics Malaysia Official Website. <<https://newss.statistics.gov.my>> [Accessed on 3rd January 2020].
- Desvars, A., Jégo, S., Chiroleu, F., Bourhy, P., Cardinale, E., & Michault, A. (2011). Seasonality of human leptospirosis in Reunion Island (Indian Ocean) and its association with meteorological data. *PloS One*, 6(5), 1-10. <https://doi.org/10.1371/journal.pone.0020377>.
- Dhewantara, P. W., Al Mamun, A., Zhang, W. Y., Yin, W. W., Ding, F., Guo, D., & Magalhães, R. J. S. (2018). Geographical and temporal distribution of the residual clusters of human leptospirosis in China, 2005–2016. *Scientific Reports*, 8(1), 1-12.
- Dhewantara, P. W., Hu, W., Zhang, W., Yin, W. W., Ding, F., Al Mamun, A., & Magalhães, R. J. S. (2019). Climate variability, satellite-derived physical environmental data and human leptospirosis: A retrospective ecological study in China. *Environmental Research*, 176(9), 1-10.
- Dindang, A., Taat, A., Phuah, E. B., Alwi, A. M., Mandai, A. A., Adam, S. F., & Lah, D. (2013). Statistical and trend analysis of rainfall data in Kuching, Sarawak from 1968 2010. *Jabatan Meteorologi Malaysia Research Publication*, 6(13), 1-17.

- Dom, N. C., Ahmad, A. H., Latif, Z. A., & Ismail, R. (2017). Integration of GIS-Based Model with Epidemiological data as a Tool for Dengue Surveillance. *Environment Asia*, 10(2), 135-146.
- Dupouey, J., Faucher, B., Edouard, S., Richet, H., Kodjo, A., Drancourt, M., & Davoust, B. (2014). Human leptospirosis: an emerging risk in Europe. *Comparative Immunology, Microbiology and Infectious Diseases*, 37(2), 77-83.
- Earthstat. (2000). GIS data for agriculture and environment. [online] Available at < <http://www.earthstat.org>> [Assessed on 25th October 2018].
- Environmental Systems Research Institute. (2015). ArcGIS 10. 3. 1.
- Faine, S. (1994). *Leptospira* and leptospirosis, Boca Raton: CRC Press Inc.
- Feuillet, T., Charreire, H., Menai, M., Salze, P., Simon, C., Dugas, J., & Oppert, J. M. (2015). Spatial heterogeneity of the relationships between environmental characteristics and active commuting: towards a locally varying social ecological model. *International Journal of Health Geographics*, 14(3), 12-25.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315.
- Fletcher-Lartey, S. M., & Caprarelli, G. (2016). Application of GIS technology in public health: successes and challenges. *Parasitology*, 143(4), 401-415.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*, Chichester: John Wiley and Sons, Ltd.

- Fraga, T. R., Carvalho, E., Isaac, L., & Barbosa, A. S. (2015). *Leptospira* and leptospirosis. In *Molecular medical microbiology* (pp. 1973-1990). Academic Press.
- GADM. (2018). GADM database of global administrative areas. <https://gadm.org/index.html>. Accessed 7 September 2018.
- Gallana, M., Ryser-Degiorgis, M. P., Wahli, T., & Segner, H. (2013). Climate change and infectious diseases of wildlife: altered interactions between pathogens, vectors and hosts. *Current Zoology*, 59(3), 427-437.
- Garba, B., Bahaman, A. R., Bejo, S. K., Zakaria, Z., Mutalib, A. R., & Bande, F. (2018). Major epidemiological factors associated with leptospirosis in Malaysia. *Acta Tropica*, 178(2), 242-247.
- García-Rangel, S., & Pettorelli, N. (2013). Thinking spatially: the importance of geospatial techniques for carnivore conservation. *Ecological Informatics*, 14(3), 84-89.
- Getis, A., & Ord, J. K. (1996). Spatial analysis and modeling in a GIS environment. In *A Research Agenda for Geographic Information Science* (McMaster, R. B., & User, E.L). pp157-196. Boca Raton: CRC Press.
- Ghizzo Filho, J., Nazário, N. O., Freitas, P. F., Pinto, G. D. A., & Schlindwein, A. D. (2018). Temporal analysis of the relationship between leptospirosis, rainfall levels and seasonality, Santa Catarina, Brazil, 2005-2015. *Revista do Instituto de Medicina Tropical de São Paulo*, 60(7), 1-9. <https://doi.org/10.1590/S1678-9946201860039>.
- GLW. (2014). Gridded livestock of the world, version 2.1. [online] Available at <<http://www.fao.org/geonetwork/srv/en/main.home>> [Assessed 25th October 2018].

- Goris, M. G., Boer, K. R., Duarte, T. A., Kliffen, S. J., & Hartskeerl, R. A. (2013). Human leptospirosis trends, the Netherlands, 1925–2008. *Emerging Infectious Diseases*, 19(3), 371-378.
- Grooms, D. L. (2006). Reproductive losses caused by bovine viral diarrhea virus and leptospirosis. *Theriogenology*, 66(3), 624-628.
- Guerra-Silveira, F., & Abad-Franch, F. (2013). Sex bias in infectious disease epidemiology: patterns and processes. *PloS One*, 8(4), 113. <https://doi.org/10.1371/journal.pone.0062390>.
- Guo, J. G., Penelope, V., Cao, C. L., Jürg, U., Zhu, H. Q., Daniel, A., & Chen, M. G. (2005). A geographic information and remote sensing-based model for prediction of *Oncomelania hupensis* habitats in the Poyang Lake area, China. *Acta Tropica*, 96(23), 213-222.
- Gupta, A., Kamble, T., & Machiwal, D. (2017). Comparison of ordinary and Bayesian kriging techniques in depicting rainfall variability in arid and semi-arid regions of north-west India. *Environmental Earth Sciences*, 76(15), 512-527.
- Gutiérrez, J. D., & Martínez-Vega, R. A. (2018). Spatio-temporal dynamics of human leptospirosis and its relationship with rainfall anomalies in Colombia. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 112(3), 115-123.
- Gwitira, I., Murwira, A., Zengeya, F. M., & Shekede, M. D. (2018). Application of GIS to predict malaria hot spots based on *Anopheles arabiensis* habitat suitability in Southern Africa. *International Journal of Applied Earth Observation and Geoinformation*, 64(2), 12-21.

- Haake, D. A., Dundoo, M., Cader, R., Kubak, B. M., Hartskeerl, R. A., Sejvar, J. J., & Ashford, D. A. (2002). Leptospirosis, water sports, and chemoprophylaxis. *Clinical Infectious Diseases*, 34(9), 40-43. <https://doi.org/10.1086/339942>.
- Haake D.A., Levett P.N. (2015) Leptospirosis in Humans. In *Leptospira and Leptospirosis* (Adler B), pp 65-97, Heidelberg: Springer.
- Hartskeerl, R. A., Collares-Pereira, M., & Ellis, W. A. (2011). Emergence, control and re-emerging leptospirosis: dynamics of infection in the changing world. *Clinical Microbiology and Infection*, 17(4), 494-501.
- Hasim, H., Hiong, T. G., Hatta, M., Amierul, F. M., Aliza, L., Yoep, N., & Paiwai, F. (2018). Spatial density of dengue incidence: A case study of a dengue outbreak in Seksyen 7, Shah Alam. *International Journal of Mosquito Research*, 5(2), 9-14.
- Hassan, A. A., & Tahar, K. N. (2016). Evaluation of spatial risk factor for leptospirosis outbreak using GIS application. *International Journal of Advanced and Applied Sciences*, 3(7), 60-68.
- Hazrin, M., Hiong, H. G., Jai, N., Yeop, N., Hatta, M., Paiwai, F., & Othman, W. (2016). Spatial distribution of dengue incidence: A case study in Putrajaya. *Journal of Geographic Information System*, 8(1), 89-97.
- Higgs, G. (2004). A literature review of the use of GIS-based measures of access to health care services. *Health Services and Outcomes Research Methodology*, 5(2), 119-139.
- Hinjoy, S. (2016). Epidemiology of leptospirosis from Thai national disease surveillance system, 2003-2012. *Outbreak, Surveillance, Investigation and Response Journal*, 7(2), 1-5.

- Hisham, B., Marzukhi, M., & Daud, A. R. (2009). Spectrum of flood-related diseases encountered during flood disaster in Johore, Malaysia. *Malaysian Journal of Community Health*, 15(9), 15-23.
- Hoki, M. (1977). Farming operations and labor requirements for paddy cultivation in Sarawak, East Malaysia. *Japanese Journal of Southeast Asian Studies*, 15(3), 457-471.
- Jones, B. A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M. Y., & Pfeiffer, D. U. (2013). Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences*, 110(21), 8399-8404.
- Joshi, Y. P., Kim, E. H., & Cheong, H. K. (2017). The influence of climatic factors on the development of hemorrhagic fever with renal syndrome and leptospirosis during the peak season in Korea: an ecologic study. *BioMed Central Infectious Diseases*, 17(1), 406-416.
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4(9), 1-20.
- Kawaguchi, L., Sengkeoprasedh, B., Tsuyuoka, R., Koizumi, N., Akashi, H., Vongphrachanh, P., & Aoyama, A. (2008). Seroprevalence of leptospirosis and risk factor analysis in flood-prone rural areas in Lao PDR. *The American Journal of Tropical Medicine and Hygiene*, 78(6), 957-961.
- Koizumi, N., Muto, M., Tanikawa, T., Mizutani, H., Sohmura, Y., Hayashi, E., & Watanabe, H. (2009). Human leptospirosis cases and the prevalence of rats

- harbouring *Leptospira interrogans* in urban areas of Tokyo, Japan. *Journal of Medical Microbiology*, 58(9), 1227-1230.
- Krivoruchko, K. (2012). Empirical bayesian kriging. *ESRI: Redlands, CA*. [Online] Available at <<http://www.esri.com/news/arcuser/1012/empirical-byesian-kriging.html>> [Assessed on 3rd January 2020].
- Lau, C. L., Smythe, L. D., Craig, S. B., & Weinstein, P. (2010). Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104(10), 631-638.
- Lawson, A. B. (2013). *Statistical methods in spatial epidemiology*, 2nd ed., Chichester: John Wiley & Sons, Ltd.
- Levett, P. N., Branch, S. L., Whittington, C. U., Edwards, C. N., & Paxton, H. (2001). Two methods for rapid serological diagnosis of acute leptospirosis. *Clinical and Diagnostics Laboratory Immunology*, 8(2), 349-351.
- Lovett, A., & Flowerdew, R. (1989). Analysis of count data using Poisson regression. *The Professional Geographer*, 41(2), 190-198.
- Mayfield, H. J., Lowry, J. H., Watson, C. H., Kama, M., Nilles, E. J., & Lau, C. L. (2018). Use of geographically weighted logistic regression to quantify spatial variation in the environmental and sociodemographic drivers of leptospirosis in Fiji: a modelling study. *The Lancet Planetary Health*, 2(5), 223-232. [https://doi.org/10.1016/S2542-5196\(18\)30066-4](https://doi.org/10.1016/S2542-5196(18)30066-4).
- Matsushita, N., Ng, C. F. S., Kim, Y., Suzuki, M., Saito, N., Ariyoshi, K., & Hashizume, M. (2018). The non-linear and lagged short-term relationship between rainfall and

- leptospirosis and the intermediate role of floods in the Philippines. *PLoS Neglected Tropical Diseases*, 12(4), 1-13. <https://doi.org/10.1371/journal.pntd.0006331>.
- Matthews, S. A., & Yang, T. C. (2012). Mapping the results of local statistics: Using geographically weighted regression. *Demographic Research*, 26, 151-166.
- Ministry of Health Malaysia. (2011). Guidelines for the Diagnosis, Management, Prevention and Control of Leptospirosis in Malaysia. Retrived from <https://www.moh.gov.my/moh/resources/auto%20download%20images/589d71cb77d8.pdf>.
- Ministry of Human Resource Malaysia. (2013). Malaysian Standard Classification of Occupations (MASCO). Retrieved from <http://jtksm.mohr.gov.my/en/sumber-sumber/penerbitan/garis-panduan/814-malaysia-standard-classification-occupations-masco-2013.pdf>.
- Mirzaei, R., & Sakizadeh, M. (2016). Comparison of interpolation methods for the estimation of groundwater contamination in Andimeshk-Shush Plain, Southwest of Iran. *Environmental Science and Pollution Research*, 23(3), 2758-2769.
- Mohammadinia, A., Alimohammadi, A., & Saeidian, B. (2017). Efficiency of geographically weighted regression in modeling human leptospirosis based on environmental factors in Gilan Province, Iran. *Geosciences*, 7(4), 136-177.
- Mohammadinia, A., Saeidian, B., Pradhan, B., & Ghaemi, Z. (2019). Prediction mapping of human leptospirosis using ANN, GWR, SVM and GLM approaches. *BioMed Central Infectious Diseases*, 19(1), 1-18.

- Mohammed, H., Nozha, C., Hakim, K., Abdelaziz, F., & Rehia, B. (2011). *Leptospira*: morphology, classification and pathogenesis. *Journal of Bacteriology and Parasitology*, 2(6), 120-123.
- Moore, D. A., & Carpenter, T. E. (1999). Spatial analytical methods and geographic information systems: use in health research and epidemiology. *Epidemiologic Reviews*, 21(2), 143-161.
- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(2), 17-23.
- Mwachui, M. A., Crump, L., Hartskeerl, R., Zinsstag, J., & Hattendorf, J. (2015). Environmental and behavioural determinants of leptospirosis transmission: a systematic review. *PLoS Neglected Tropical Diseases*, 9(9), 1-15. <https://doi.org/10.1371/journal.pntd.0003843>.
- Nakaya, T., Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, 24(17), 2695-2717.
- Nia, A. M., Alimohammadi, A., Habibi, R., & Shirzadi, M. R. (2015). Spatial and statistical analysis of leptospirosis in Guilan province, Iran. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(1), 497-502.
- Ngui, R., Shafie, A., Chua, K. H., Mistam, M. S., Al-Mekhlafi, H. M., Sulaiman, W. W., & Lim, Y. A. (2014). Mapping and modelling the geographical distribution of soil transmitted helminthiases in Peninsular Malaysia: implications for control approaches. *Geospatial Health*, 8(12), 365-376.

- Odoi, A., & Busingye, D. (2014). Neighborhood geographic disparities in heart attack and stroke mortality: comparison of global and local modeling approaches. *Spatial and Spatio-Temporal Epidemiology*, 11(10), 109-123.
- Ozdenerol, E. (2015). GIS and remote sensing use in the exploration of lyme disease epidemiology. *International Journal of Environmental Research and Public Health*, 12(12), 15182-15203.
- Pahrol, M., Noraishah, M. S., & Nasir, R. A. (2018). Spatial Distribution of malaria incidence in Sabah from 2012 to 2016. *Geoinformatics & Geostatistics: An Overview*, 6(3), 1-5.
- Pappachan, M. J., Mathew, S., Aravindan, K. P., Khader, A., Bharghavan, P. V., Kareem, M. M., Tuteja, U., Shukla, J., & Batra, H. V. (2004). Risk factors for mortality in patients with leptospirosis during an epidemic in northern Kerala. *The National Medical Journal of India*, 17(5), 240–242.
- Pawar, S. D., Kore, M., Athalye, A., & Thombre, P. S. (2018). Seasonality of leptospirosis and its association with rainfall and humidity in Ratnagiri, Maharashtra. *International Journal of Health & Allied Sciences*, 7(1), 37-40.
- Philip, R. R., Indu, P. S., Anish, T. S., Shrinivasa, B. M., Joy, T. M., & Sujina, C. M. (2013). Occupational and environmental risk factors of leptospirosis: a case control study in a tertiary care setting in Kerala, India. *Health Sciences*, 2(1), 1-12.
- Plank, R., & Dean, D. (2000). Overview of the epidemiology, microbiology, and pathogenesis of *Leptospira* spp. in humans. *Microbes and Infection*, 2(10), 1265-1276.

- Pui, C. F., Bilung, L. M., Su'ut, L., & Apun, K. (2015). Prevalence of *Leptospira* species in environmental soil and water from national parks in Sarawak, Malaysia. *Borneo Journal of Resource Science and Technology*, 5(1), 49-57.
- Pui, C. F., Bilung, L. M., Su'ut, L., Chong, Y. L., & Apun, K. (2017a). Detection of *Leptospira* spp. in Selected National Service Training Centres and Paddy Fields of Sarawak, Malaysia using Polymerase Chain Reaction Technique. *Pertanika Journal of Tropical Agricultural Science*, 40(1), 99-109.
- Pui, C. F., Bilung, L. M., Apun, K., & Su'ut, L. (2017b). Diversity of *Leptospira* spp. in Rats and Environment from Urban Areas of Sarawak, Malaysia. *Journal of Tropical Medicine*, 2017, 3760674. <https://doi.org/10.1155/2017/3760674>.
- Radi, M. F. M., Hashim, J. H., Jaafar, M. H., Hod, R., Ahmad, N., Naw, A. M., & Ayub, N. I. F. (2018). Leptospirosis outbreak after the 2014 major flooding event in Kelantan, Malaysia: a spatial-temporal analysis. *The American Journal of Tropical Medicine and Hygiene*, 98(5), 1281-1295.
- Rafizah, A. N., Aziah, B. D., Azwany, Y. N., Imran, M. K., Rusli, A. M., Nazri, S. M., & Zaliha, I. (2013). Risk factors of leptospirosis among febrile hospital admissions in northeastern Malaysia. *Preventive Medicine*, 57(1), 11-13.
- Ramankutty, N., Evan, A. T., Monfreda, C., & Foley, J. A. 1. (2008). Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemistry Cycles*, 22(1), 1-19. <https://doi.org/10.1029/2007GB002952>.
- R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available on <URL <http://www.R-project.org/>> [Assessed on 7th August 2018].

- Reis, R. B., Ribeiro, G. S., Felzemburgh, R. D., Santana, F. S., Mohr, S., Melendez, A. X., & Carvalho, M. S. (2008). Impact of environment and social gradient on *Leptospira* infection in urban slums. *PLoS Neglected Tropical Diseases*, 2(4), 228-237. <https://doi.org/10.1371/journal.pntd.0000228>.
- Ridzuan, B. J. M., Aziah, B. D., & Zahiruddin, W. M. (2016). Study on seroprevalence and leptospiral antibody distribution among high-risk planters in Malaysia. *Osong Public Health and Research Perspectives*, 7(3), 168-171.
- Robinson, T. P., Franceschini, G., & Wint, W. (2007). The Food and Agriculture Organization's gridded livestock of the world. *Veterinaria Italiana*, 43(3), 745-751.
- Robinson, T. P., Wint, G. W., Conchedda, G., Van Boeckel, T. P., Ercoli, V., Palamara, E., & Gilbert, M. (2014). Mapping the global distribution of livestock. *PloS One*, 9(5), 1-13. <https://doi.org/10.1371/journal.pone.0096084>.
- Rood, E., Goris, M., Pijnacker, R., Bakker, M. I., & Hartskeerl, R. A. (2017). Environmental risk of leptospirosis infections in the Netherlands: Spatial modelling of environmental risk factors of leptospirosis in the Netherlands. *PloS One*, 12(10), 1-11. <https://doi.org/10.1371/journal.pone.0186987>.
- Saffawati, T. N., Ismail, T., Kassim, N. F. A., Rahman, A. A., Hamid, S. A., Yahya, K., & Webb, C. E. (2019). The application of geographic information system (GIS) to assess the population abundance of *Aedes albopictus* (Skuse) in mangrove forests of Penang, Malaysia. *International Journal of Mosquito Research*, 6(1), 50-54.
- Saha, R., Sil, S., & Thakurta, R. G. (2016). Incidence of leptospirosis in India: a cross sectional study. *European Journal of Pharmaceutical and Medical Research*, 3(1), 502-504.

- Sapian, M., Khairi, M. T., How, S. H., Rajalingam, R., Sahhir, K., Norazah, A., & Jamalludin, A. R. (2012). Outbreak of melioidosis and leptospirosis co-infection following a rescue operation. *Medical Journal of Malaysia*, 67(3), 293-297.
- Saputra, I. W. A., Dibia, I. N., & Puja, I. K. (2015). Risk Factors and Spatial Distribution of Rabies in Bali in the Year of 2014. *Veterinary Science and Medicine Journal*, 3(2), 69-72. <https://ojs.unud.ac.id/index.php/jikh/article/view/20533>.
- Sarawak Health Department. (2012). *Sarawak Weekly Epid. News*. December 24-30.
- Seguro, A. C., & Andrade, L. (2013). Pathophysiology of leptospirosis. *Shock: Injury, Inflammation, and Sepsis: Laboratory and Clinical Approaches*, 39(5), 17-23.
- Sejvar, J., Bancroft, E., Winthrop, K., Bettinger, J., Bajani, M., Bragg, S., & Tappero, J. (2003). Leptospirosis in “eco-challenge” athletes, Malaysian Borneo, 2000. *Emerging Infectious Diseases*, 9(6), 702-707.
- Shafei, M. N., Sulong, M. R., Yaacob, N. A., Hassan, H., Mohamad, W. M. Z. W., Daud, A., & Abdullah, M. R. (2012). Seroprevalence of leptospirosis among town service workers in Northeastern state of Malaysia. *International Journal of Collaborative Research on Internal Medicine & Public Health*, 4(4), 395-403.
- Sharma, B., Tyagi, S., Singh, P., Dobhal, R., & Jaiswal, V. (2015). Application of remote sensing and GIS in hydrological studies in India: An overview. *National Academy Science Letters*, 38(1), 1-8.
- Shaw, N. T. (2012). Geographical information systems and health: current state and future directions. *Healthcare Informatics Research*, 18(2), 88-96.
- Sirisena, P., Noordeen, F., Kurukulasuriya, H., Romesh, T. A., & Fernando, L. (2017). Effect of Climatic Factors and Population Density on the Distribution of Dengue in Sri

- Lanka: A GIS-based Evaluation for Prediction of Outbreaks. *PloS One*, 12(1), 1-14. <https://doi.org/10.1371/journal.pone.0166806>.
- SPSS, I.I.B.M. (2011). IBM SPSS statistics for Windows, version 20.0. *New York: IBM Corp*, 440.
- Soo, Z. M. P., Khan, N. A., & Siddiqui, R. (2019). Leptospirosis: Increasing importance in developing countries. *Acta Tropica*, 201(1), 1-9.
- Sumi, A., Telan, E. F. O., Chagan-Yasutan, H., Piolo, M. B., Hattori, T., & Kobayashi, N. (2017). Effect of temperature, relative humidity and rainfall on dengue fever and leptospirosis infections in Manila, the Philippines. *Epidemiology & Infection*, 145(1), 78-86.
- Su'ut, L., Mazlan, M. N. A., Arif, M. T., Yusoff, H., Abdul Rahim, N. A., Safii, R., & Suhaili, M. R. (2016). Serological prevalence of leptospirosis among rural communities in the Rejang basin, Sarawak, Malaysia. *Asia Pacific Journal of Public Health*, 28(5), 450-457.
- Su'ut, L., Mazlan, A., Arif, M. T., Katip, T., Nor Aliza, A. R., & Haironi, Y. (2018). Serovar diversity of *Leptospira* sp. infecting wild rodents in Sarawak, Malaysia. *Tropical Biomedicine*, 35(1), 252-258.
- Suwanpakdee, S., Kaewkungwal, J., White, L. J., Asensio, N., Ratanakorn, P., Singhasivanon, P., & Pan-Ngum, W. (2015). Spatio-temporal patterns of leptospirosis in Thailand: is flooding a risk factor? *Epidemiology and Infection*, 143(10), 2106-2115.
- Tawan, A. S., Ling, T. Y., Nyanti, L., Sim, S. F., Grinang, J., Soo, C. L., & Ganyai, T. (2020). Assessment of water quality and pollutant loading of the Rajang River and

- its tributaries at Pelagus area subjected to seasonal variation and river regulation. *Environment, Development and Sustainability*, 22(5), 4101-4124.
- Tan, D. S. (1970). Leptospirosis in rural West Malaysia. *Medical Journal of Malaya*, 24(4), 261-266.
- Tan, W. L., Soelar, S. A., Mohd Suan, M. A., Hussin, N., Cheah, W. K., Verasahib, K., & Goh, P. P. (2016). Leptospirosis incidence and mortality in Malaysia. *Southeast Asian Journal of Tropical Medicine and Public Health*, 47(3), 434-440.
- Thayaparan, S., Robertson, I. A. N., Amraan, F., Su'ut, L., & Abdullah, M. T. (2013a). Serological prevalence of leptospiral infection in wildlife in Sarawak, Malaysia. *Borneo Journal of Resource of Science and Technology*, 2(2), 79-82.
- Thayaparan, S., Robertson, I. D., Fairuz, A., Suut, L., & Abdullah, M. T. (2013b). Leptospirosis, an emerging zoonotic disease in Malaysia. *Malaysian Journal of Pathology*, 35(2), 123-132.
- Thayaparan, S., Robertson, I. D., Fairuz, A., Suut, L., Gunasekera, U. C., & Abdullah, M. T. (2015). Seroepidemiological study of leptospirosis among the communities living in periurban areas of Sarawak, Malaysia. *The Medical Journal of Malaysia*, 70(5), 288-294.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(6), 234-240. <https://www.jstor.org/stable/143141>.
- Togami, E., Kama, M., Goarant, C., Craig, S. B., Lau, C., Ritter, J. M., & Nilles, E. J. (2018). A large leptospirosis outbreak following successive severe floods in Fiji, 2012. *The American Journal of Tropical Medicine and Hygiene*, 99(4), 849-851.

- Vega-Corredor, M. C., & Opadeyi, J. (2014). Hydrology and public health: linking human leptospirosis and local hydrological dynamics in Trinidad, West Indies. *Earth Perspectives*, 1(3), 1-14. <https://doi.org/10.1186/2194-6434-1-3>.
- Victoriano, A. F. B., Smythe, L. D., Gloriani-Barzaga, N., Cavinta, L. L., Kasai, T., Limpakarnjanarat, K., & Yanagihara, Y. (2009). Leptospirosis in the Asia Pacific region. *BioMed Central Infectious Diseases*, 9(1), 147-155.
- Wang, C., Zhong, S., Zhang, F., & Huang, Q. (2016). Precipitation interpolation by multivariate bayesian maximum entropy based on meteorological data in Yun-Gui Guang region, mainland china. In *IOP Conference Series: 6th Digital Earth Summit, Beijing, China, 7-8 July 2019*. IOP Publishing.
- Weinberger, D., Baroux, N., Grangeon, J. P., Ko, A. I., & Goarant, C. (2014). El Niño southern oscillation and leptospirosis outbreaks in New Caledonia. *PLoS Neglected Tropical Diseases*, 8(4), 2798-2804.
- Weisent, J., Rohrbach, B., & Dunn, J. R. (2012). Socioeconomic determinants of geographic disparities in campylobacteriosis risk: a comparison of global and local modelling approaches. *International Journal of Health Geographics*, 11(1), 45-60.
- Widayani, P., Gunawan, T., Danoedoro, P., & Mardihusodo, S. J. (2016). Application of geographically weighted regression for vulnerable area mapping of leptospirosis in Bantul District. *The Indonesian Journal of Geography*, 48(2), 168-177.
- Widiastuti, D., Puspitaningsih, D., & Wijayanti, S. M. P. (2019). Leptospirosis Outbreak during Rice Harvesting Season in Kebumen, Central Java Indonesia (The First Case Report in Kebumen). *Indian Journal of Public Health Research & Development*, 10(3), 654-658.

- WorldClim. (2016). Global climate data, version 2. [online] Available at < <http://worldclim.org/version2>> [Assessed on 25th October 2018].
- World Health Organization (2011). Report of the second meeting of the leptospirosis burden epidemiology reference group. World Health Organization, Geneva.
- Wynwood, S. J., Graham, G. C., Weier, S. L., Collet, T. A., McKay, D. B., & Craig, S. B. (2014). Leptospirosis from water sources. *Pathogens and Global Health*, 108(7), 334-338.
- Yaakob, Y., Rodrigues, K. F., & John, D. V. (2015). Leptospirosis: recent incidents and available diagnostics—a review. *The Medical Journal of Malaysia*, 70(6), 351-355.
- Yoep, N., Hasim, H., Yusoff, U. N., Yusoff, M., & Mahpot, N. R. (2015). Spatio-temporal distribution of malaria in Perak, Malaysia. *Advances in Infectious Diseases*, 5(4), 154-161.
- Yusof, M. M., Dom, N. C., & Zainuddin, A. (2017). Spatial pattern distribution of dengue fever in sub-urban area using GIS tools. *Serangga*, 21(2), 127-148.
- Zulkifli, Z., Shariff, A. R. M., & Tarmidi, Z. M. (2016). Human leptospirosis distribution pattern analysis in Hulu Langat, Selangor. In *IOP Conference Series: Earth and Environmental Science*, 8th IGRSM International Conference Exhibition on Geospatial & Remote Sensing, Kuala Lumpur ,13-14 April 2016. IOP Publishing

APPENDICES

Appendix A: Permission to conduct research from Sarawak Planning Unit.



STATE PLANNING UNIT
CHIEF MINISTER'S DEPARTMENT
6th & 7th Floor, Wisma Bapa Malaysia
93502 Petra Jaya, Kuching
Sarawak, Malaysia

Tel : (6)082 – 319387/440857

Fax : (6)082 – 449481/442536

Website : <http://www.jkm.sarawak.gov.my>

Our Ref. : (27-) JKM/SPU/608-8/2/1 VOL.2
Date : 6 June 2018

Mr. Rosdi anak Kira
No.4, Taman Lee Cheng,
97200, Tatau, Bintulu,
Sarawak.

Dear Sir,

APPLICATION TO CONDUCT RESEARCH IN SARAWAK

I am pleased to inform you that approval is hereby given to enable you to conduct a research in Sarawak titled '**Assessment of Environmental Variables as Potential Risk Factors for Leptospira Transmission**' from 7.06.2018 to 1.03.2020 (20 months).

This approval is subject to the following terms and conditions :-

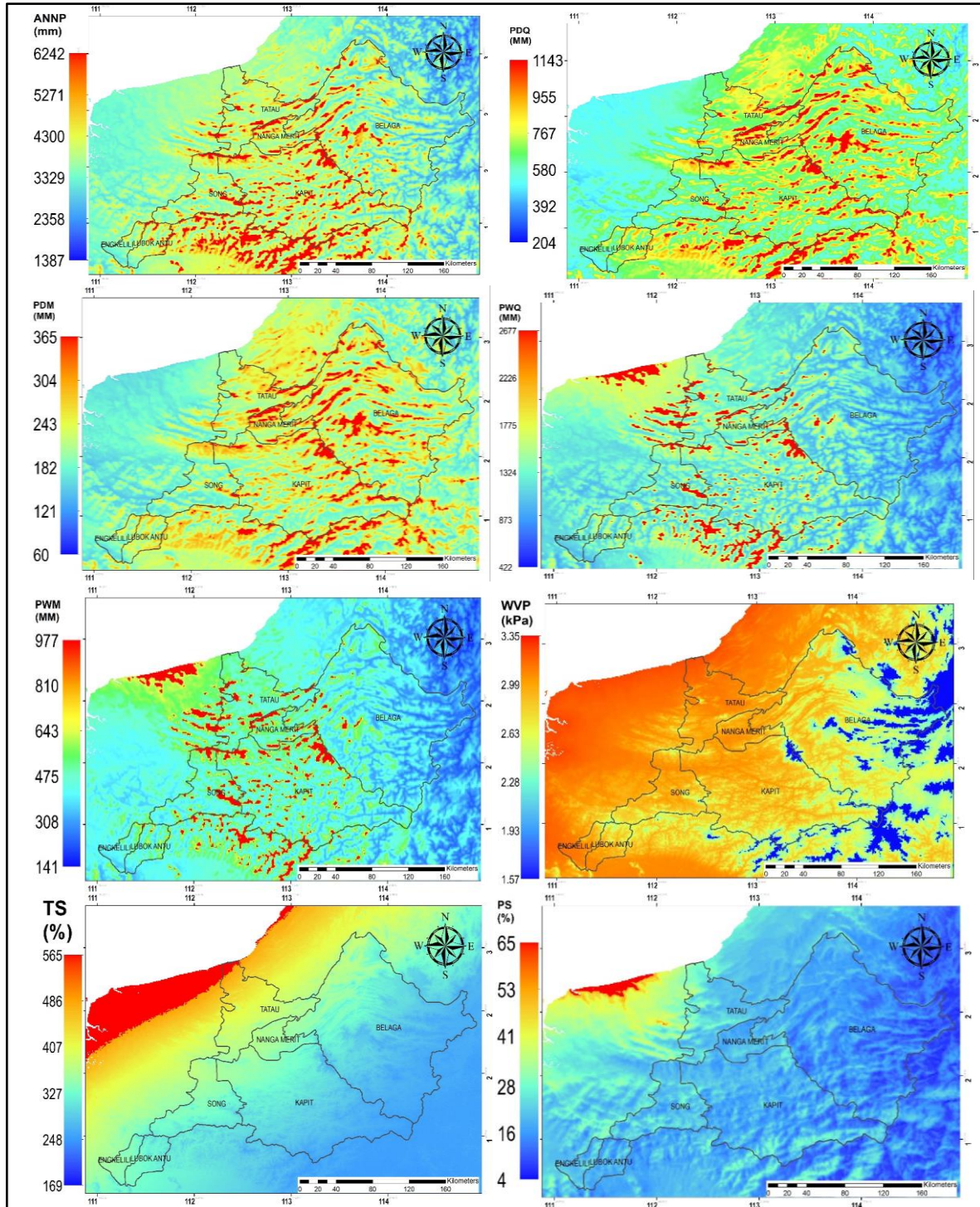
- (i) The research will not involve any financial aid from the State Government of Sarawak;
- (ii) You are to complete your research within the stipulated time as indicated in this letter of approval;
- (iii) Your study should be limited to your research area i.e. Kuching;
- (iv) The Government has the right to withdraw this approval should you deviate from the original scope of your research based on the application form submitted to us; and
- (v) Upon completion of your research, printed copies of the final report of your study written in the English Language are to be extended free to the Sarawak State Library (that is designated as the State Depository Centre) (1 copy); the State Planning Unit (1 copy) and other departments/ agencies connected with your study at no cost.

Thank you.

"BERSATU BERUSAHA BERBAKTI"
"AN HONOUR TO SERVE"

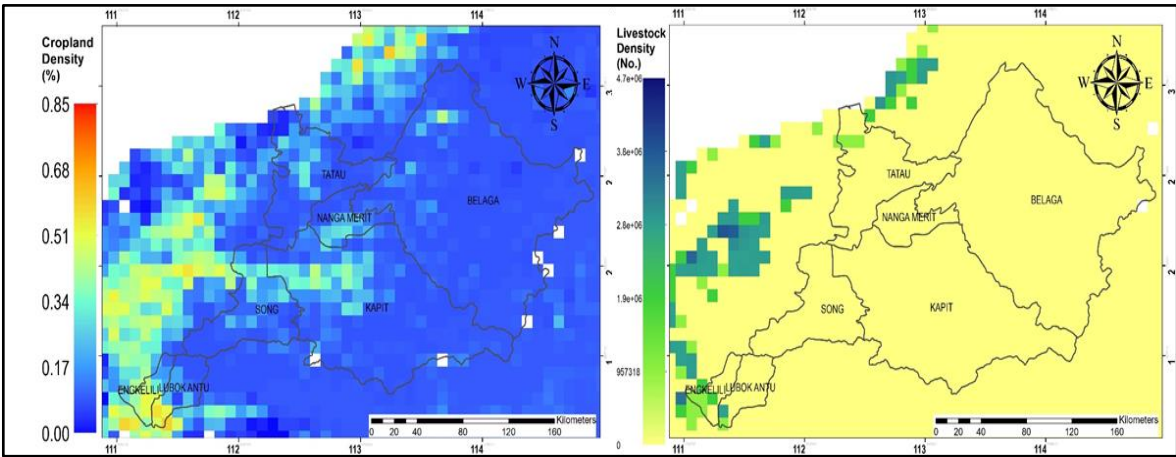
(BUCKLAND BANGIK)
Director
State Planning Unit
for State Secretary Sarawak

Appendix B: Geospatial climate variables from CHELSA and Worldclim.



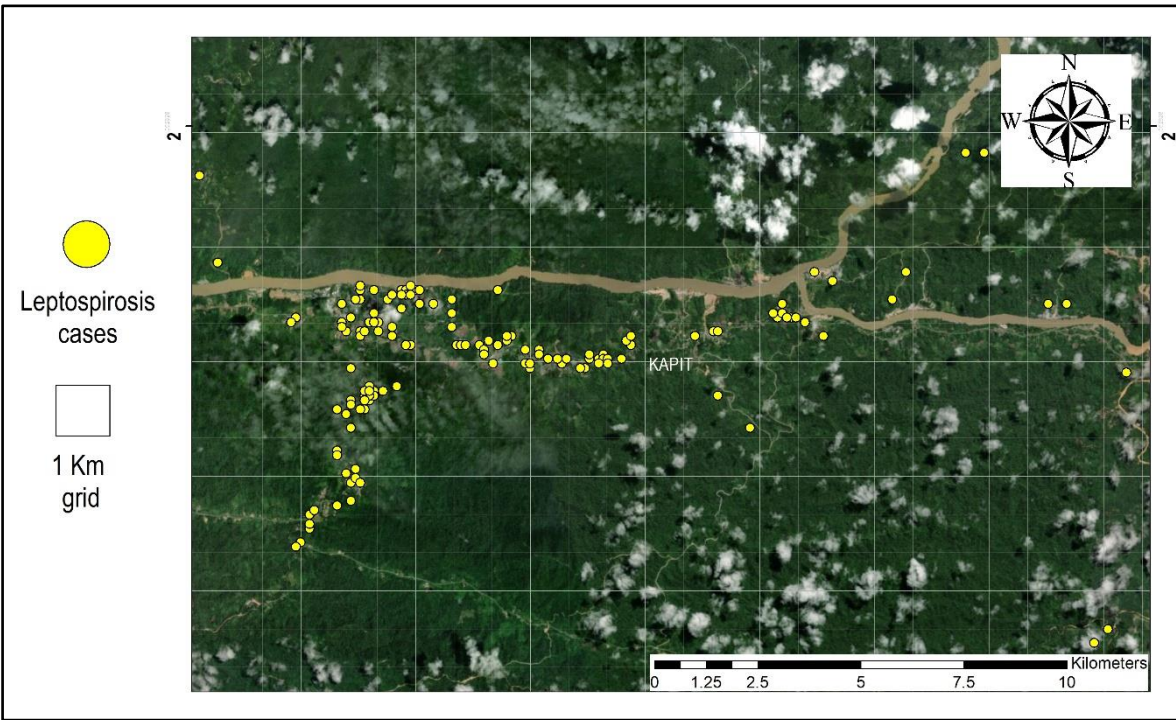
ANNP: annual precipitation; PDQ: precipitation in the driest quarter; PDM: Precipitation in the driest month; PWQ: precipitation in the wettest quarter; PWM: precipitation in the wettest month; WVP: water vapour pressure; TS: temperature seasonality; PS: precipitation seasonality.

Appendix C: Geospatial agricultural variables from Gridded Livestock of the World (GLW) and Earthstat



CLD: Cropland Density; LSD: livestock density

Appendix D: Spatially joining the geocoded leptospirosis cases in hot spot region with 1×1 km square grid.



Appendix E: The R-scripts used in the development of global Poisson regression.

```
#variables
Count_~std.lsd+std.crp+std.wvp+std.ts+std.annp+std.ps+std.pdq+std.pdm+std.pwm+std.pwq
#univariate analysis
univar<-glm(Count_~std.lsd, data=df, family=poisson(link="log"))
summary(univar)
exp(coef(univar))
exp(confint(univar))
#goodness of fit test univariate
with(univar,cbind(res.deviance= deviance, df= df.residual,
                  p= pchisq(deviance,df.residual, lower.tail = FALSE)))
#robust stadard error for parameter estimates
cov.univar<-vcovHC(univar,type = "HC0")
std.err<-sqrt(diag(cov.univar))
r.rest<-cbind(Estimate=coef(univar),"Robust SE"=std.err,
              "Pr(>|z|)"=2*pnorm(abs(coef(univar)/std.err),lower.tail = FALSE),
              LL=coef(univar)-1.96*std.err,UL=coef(univar)+1.96*std.err)
r.rest
m2<-update(univar,~.-std.lsd)
anova(m2,univar, test="Chisq")
#variable elimination (vif>4) (only significant variables in univariate analysis)
multivar<-glm(Count_~std.crp+std.wvp+std.annp+std.pdq+std.pwq+std.pdm, data = df,
family=poisson(link="log"))
summary(multivar)
vif(multivar)
#MULTIVARIATE MODEL AFTER VIF>4 ELIMINATION
multivar2<-glm(Count_~std.crp+std.wvp+std.annp+std.pwq+std.pdm, data = df,
family=poisson(link="log"))
summary(multivar2)
vif(multivar2)
#model selection stepwise regression with backward elimination
step(multivar2,direction = "both")
#FINAL MULTIVARIATE MODEL
```

```

finalmodel<-glm(Count_ ~ std.wvp + std.annp + std.pdm + std.pwq, data = df,
family=poisson(link="log"))
summary(finalmodel)
vif(finalmodel)
#robust stadard error for parameter estimates
cov.final<-vcovHC(finalmodel,type = "HC0")
std.err<-sqrt(diag(cov.final))
r.rest<-cbind(Estimate=coef(finalmodel),"Robust SE"=std.err,
              "Pr(>|z|)"=2*pnorm(abs(coef(finalmodel)/std.err),lower.tail = FALSE),
              LL=coef(finalmodel)-1.96*std.err,UL=coef(finalmodel)+1.96*std.err)
r.rest
exp(coef(finalmodel))
exp(confint(finalmodel))
#residual test
with(finalmodel,cbind(res.deviance= deviance, df= df.residual,
                      p= pchisq(deviance, df.residual, lower.tail = FALSE)))

```