HYBRID MODEL OF CLIMATE EFFECTS ON DENGUE FORECAST AND CONTROL IN SELANGOR, MALAYSIA

LU XINYI

UNIVERSITI SAINS MALAYSIA

2025

HYBRID MODEL OF CLIMATE EFFECTS ON DENGUE FORECAST AND CONTROL IN SELANGOR, MALAYSIA

by

LU XINYI

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

July 2025

ACKNOWLEDGEMENT

Completing the PhD program has been an invaluable experience, enriched by the support and assistance of many individuals. First and foremost, I express my heartfelt appreciation to my main supervisor, Associate Professor Dr. Teh Su Yean. Her dedicated guidance, inspiration, and unwavering support have been instrumental throughout the completion of this study. Her insightful observations and constructive ideas helped shape the overall direction of the research and refine the detailed methods. My sincere thanks also go to my co-supervisors, Dr. Fam Pei Shan and Dr. Tay Chai Jian, who provided their expertise in statistical analysis and the modelling of infectious diseases. I would like to extend my gratitude to Professor Koh Hock Lye for his valuable and concise comments on some of the work included in this thesis. To all of you, thank you for investing your valuable time to review my work and offer insightful feedback. I am also grateful to the School of Mathematical Sciences and the Institute of Postgraduate Studies at Universiti Sains Malaysia for providing a supportive working environment. I am also thankful for the financial support provided by the Fundamental Research Grant Scheme (FRGS/1/2024/STG06/USM/02/5), Ministry of Higher Education, Malaysia. Additionally, I extend my special thanks to the administrative officers for their assistance and kindness. Last but not least, my deepest gratitude goes to my best friend, Sharon Lin, for her continuous understanding and encouragement throughout these years, as well as to my parents, for their warm care and support. Thank you once again to everyone who has helped and supported me. In the future, I believe that I can continue to be a better me.

TABLE OF CONTENTS

ACKNOWLEDGEMENT .... **ii**

TABLE OF CONTENTS .... **iii**

LIST OF TABLES .... **vii**

LIST OF FIGURES .... **ix**

LIST OF SYMBOLS .... **xii**

LIST OF ABBREVIATIONS .... **xiv**

ABSTRAK .... **xvi**

ABSTRACT .... **xvii**

CHAPTER 1 INTRODUCTION **1**

1.1 Introduction 1

1.2 Problem Statement 2

1.3 Significance of the Study 4

1.4 Research Questions 5

1.5 Objectives of Thesis 5

1.6 Scope and Organization of Thesis 6

  1.6.1 Scope of the Research 6

  1.6.2 Organization of the Thesis 6

CHAPTER 2 LITERATURE REVIEW **8**

2.1 Introduction 8

2.2 Deterministic Dengue Models 8

  2.2.1 Single-serotype models 8

  2.2.2 Multi-serotype models 10

2.3 Statistical Dengue Models 11

2.3.1 Time series analysis 11

2.3.2 Spatial statistics 14

2.3.3 Machine learning and deep learning approaches 15

2.4 Climate Effects on Dengue Transmission 17

2.5 Vector Control 19

2.5.1 Mosquito fogging 20

2.5.2 *Wolbachia*-based intervention 21

CHAPTER 3 COUPLED SI-SIR AND ARIMAX MODEL 24

3.1 Introduction 24

3.2 Materials and Methods 25

3.2.1 Dengue and meteorological data 25

3.2.2 Stationary test 28

3.2.3 Granger causality test 28

3.2.4 Coupled SI-SIR and ARIMAX model 30

3.2.4(a) SI-SIR model 30

3.2.4(b) ARIMAX 33

3.3 Results 35

3.3.1 Affinity between dengue and climate variables 35

3.3.1(a) Average biting rate 36

3.3.1(b) Stationary test 37

3.3.1(c) Granger causality test 37

3.3.2 Performance of the coupled SI-SIR and ARIMAX model 38

3.4 Conclusion 45

CHAPTER 4 SI-SIR WITH MLR-LSTM 46

4.1 Introduction 46

4.2 Materials and Methods 47

4.2.1 Data 47

4.2.2 Pearson correlation coefficient and multiple linear regression (MLR) model 48

4.2.3 Multiresolution analysis (MRA) 50

4.2.4 Long Short-Term Memory (LSTM) 50

4.2.5 Evaluation criteria 53

4.3 Results 54

4.3.1 Multiple linear regression model with interaction effect 54

4.3.2 Analysis and forecast of the residuals 57

4.3.3 Forecasted dengue cases for the prediction window 59

4.3.4 Comparison with alternative approaches 61

4.3.5 Forecasting competence for various prediction periods 62

4.4 Conclusion 64

CHAPTER 5 VECTOR CONTROL EFFECTIVENESS 66

5.1 Introduction 66

5.2 Conventional Vector Control Methods 66

5.2.1 Mosquito fogging 67

5.2.1(a) Continuous mosquito fogging 67

5.2.1(b) Intermittent mosquito fogging 68

5.2.2 Elimination of mosquito breeding grounds and larvicide application 70

5.3 Emerging Vector Control Method: *Wolbachia*-based Intervention 72

5.4 Cost Estimation for the *Wolbachia* Release Program 76

5.4.1 Strategies for *Wolbachia* deployment 76

5.4.2 Detailed cost components.. 77

5.4.3 Comparison with dengue costs 79

5.5 Effect of *Wolbachia*-based Intervention in Changing Climate 79

5.6 Conclusion 84

CHAPTER 6 CONCLUDING REMARKS 86

6.1 Conclusions 86

6.2 Contributions of the Research 87

6.3 Limitations and Recommendations for Future Research 88

REFERENCES 90

LIST OF PUBLICATIONS

LIST OF TABLES

Page

Table 2.1 Comparison of time series forecasting models 13

Table 3.1 Definition, unit, and value of the parameters in the SI-SIR model used in this thesis 32

Table 3.2 ADF test for average biting rate and climate factors 37

Table 3.3 Granger causality test between average biting rate and climate factors 38

Table 3.4 Forecasted number of dengue cases for 10 rounds with MRE in brackets 41

Table 3.5 Comparison of MMRE for various forecast periods using coupled SI-SIR and ARIMAX model 42

Table 3.6 MMRE of forecasted values by coupled model (SI-SIR + ARIMAX) and ARIMAX denoting a lower average MMRE for the coupled model compared to ARIMAX based on ten rounds of forecast 42

Table 3.7 Comparison of the values of averaged MMRE for the three outbreak periods indicated in Figure 3.9 and Figure 3.10 44

Table 4.1 Results of correlation tests among the dengue cases and climate variables 54

Table 4.2 Results of correlation tests among the average biting rate and climate variables 55

Table 4.3 Results of correlation tests among the climate variables 55

Table 4.4 Evaluation criteria (MAE and MAPE) for the prediction window (from week 262 to week 365) 60

Table 4.5 Evaluation criteria (MAE and MAPE) of four approaches for the prediction window (from Weeks 262-365) 62

Table 4.6 Evaluation criteria (MAE and MAPE) using SI-SIR with MLR-LSTM for Weeks 276-365 63

Table 4.7 Evaluation criteria (MAE and MAPE) using SI-SIR with MLR-LSTM for Weeks 290-365 63

Table 4.8 Evaluation criteria (MAE and MAPE) using SI-SIR with MLR-LSTM for Weeks 314-365 64

Table 5.1 Definitions, values, and sources for parameters used in the augmented SI-SIR model (Equation((5.3)) 74

Table 5.2 Ratios of death rate and transmission probability between *Wolbachia*-carrying and *Wolbachia*-free mosquitoes............. 75

Table 5.3 Cost components of the *Wolbachia* release program 78

Table 5.4 Estimated dengue costs per case in Malaysia 79

Table 5.5 Simulation for the scenarios for the change of temperature and humidity 81

Table 5.6 Labels and explanations used in Figure 5.7 82

LIST OF FIGURES

Page

Figure 1.1 Global map of areas with dengue risk (CDC, 2024) 2

Figure 1.2 Annual dengue cases in Malaysia from 1995 to 2023 (figure plotted by the author using data provided by MOH) 2

Figure 2.1 Illustration of how cytoplasmic incompatibility (CI) gives *Wolbachia*-infected females a reproductive advantage by making the progeny of infected males and wild-type females non-viable (Ross, 2021) 22

Figure 3.1 Reported dengue cases from Jan 2012 to Dec 2020. The data in the shaded periods (from Jan 2012 to Dec 2013 and from Jan 2020 to Dec 2020) were excluded from the analysis to minimize the anomalous factors affecting the reported number of dengue cases 27

Figure 3.2 Actual (dot) and smoothed (line) dengue data for Selangor from Jan 2014 to Dec 2019 27

Figure 3.3 Average temperature (a), relative humidity (b), and accumulated precipitation (c) for Selangor from years 2014 to 2019 28

Figure 3.4 Flow chart denoting the sequence of analyses performed in Chapter 3 35

Figure 3.5 Estimated average biting rate by curve fitting using (a) unsmoothed dengue data and (b) smoothed dengue data 37

Figure 3.6 The forecasted number of dengue cases for up to four weeks for ten rounds of forecast (forecast starts from Week 283) using the coupled SI-SIR and ARIMAX model 39

Figure 3.7 The forecasted number of dengue cases for up to four weeks for ten rounds of forecast (forecast starts from Week 283) using ARIMAX 43

Figure 3.8 The forecasted number of dengue cases for up to four weeks for ten rounds of forecast (forecast starts from Week 204) using the coupled SI-SIR and ARIMAX model 43

Figure 3.9 Real-time prediction of dengue cases for three outbreak periods using ARIMAX 44

Figure 3.10 Real-time prediction of dengue cases for three outbreak periods using the coupled SI-SIR and ARIMAX model 44

Figure 4.1 Time series of average temperature (a), relative humidity (c), accumulated precipitation (e), and the histogram of the data of average temperature (b), relative humidity (d), and accumulated precipitation (f) for Selangor from years 2014 to 2020........... 48

Figure 4.2 The architecture of a typical LSTM block (Van Houdt et al. (2020)) 52

Figure 4.3 Flow chart denoting the sequence of analyses performed in Chapter 4 53

Figure 4.4 Plotted data (a) when the value of accumulated precipitation is nil and (b) when the value of accumulated precipitation is not nil using the multiple linear regression model with interaction effect 58

Figure 4.5 Decomposition and denoising of residual components derived from the multiple linear regression model incorporating interaction effects.. 58

Figure 4.6 Forecasted residuals for the prediction window (from Weeks 262-365) of the multiple linear regression model with interaction using LSTM.. 59

Figure 4.7 Forecasted dengue cases for the prediction window (from week 262 to week 365) 60

Figure 4.8 Comparison of the forecasting performance of four approaches: (a) using only MLR to forecast dengue cases subject to climate variables (Approach A), (b) using only LSTM to forecast dengue cases subject to climate variables (Approach B), and (c) using both MLR and LSTM without associating mosquito biting rate with climate variables (Approach C) and (d) using both MLR and LSTM and associating mosquito biting rate with climate variables (Approach D) 61

Figure 4.9 Forecasted dengue cases using SI-SIR with MLR-LSTM for Weeks 276-365 63

Figure 4.10 Forecasted dengue cases using SI-SIR with MLR-LSTM for Weeks 290-365 64

Figure 4.11 Forecasted dengue cases using SI-SIR with MLR-LSTM for Weeks 314-365 64

Figure 5.1 Effect of various mosquito removal rate η due to continuous fogging on dengue incidence 68

Figure 5.2 Effect of fogging frequency, namely once every 2 weeks, 3 weeks, and 1 month, on dengue incidence 69

Figure 5.3 Effect of once-per-month fogging at various intensities η, on dengue incidence 70

Figure 5.4 Effect of various breeding ground elimination and larvicide application intensity θ due to mosquito breeding grounds elimination and larvicide application............................... 71

Figure 5.5 Simulated impact of *Wolbachia* on dengue incidence for Scenario 1 ($q_1 = 1$), Scenario 2 ($q_1 = 1.1$) and Scenario 3 ($q_1 = 1.25$).. 75

Figure 5.6 Simulated dengue cases considering (a) no change in temperature and humidity; and (b) increasing average temperature by 5% and decreasing relative humidity by 4.5%, following patterns observed from 2014 to 2019 81

Figure 5.7 Effect of *Wolbachia*-based intervention with (Simulation 3, blue line) and without considering mosquito migration (Simulation 4, red line) when (a) $q_1 = 1$, (b) $q_1 = 1.1$, and (c) $q_1 = 1.25$...... 83

LIST OF SYMBOLS

α_V^c	Birth rate of *Wolbachia*-carrying mosquitoes
α_V^f	Birth rate of *Wolbachia*-free mosquitoes
β_{HV}	Probability of human-to-mosquito transmission
β_{VH}	Probability of mosquito-to-human transmission
$\widetilde{\beta_{VH}}$	Probability of mosquito-to-human transmission by *Wolbachia*-carrying mosquitoes
B	Average biting rate
$C_{m.i}$	Maintenance cost of *Wolbachia*-infected mosquito colony
$C_{release}$	Cost of identifying and selecting the appropriate *Wolbachia* strain
$C_{r,i}$	Cost of rearing *Wolbachia*-infected mosquitoes
C_s	Cost of identifying and selecting the appropriate *Wolbachia* strain
C_t	Total cost for the artificial introduction of *Wolbachia*
$L_{r,i}$	Loss rate of mosquitoes during rearing
γ_H	Recovery rate of infected human
H_I	Infected human population
H_R	Recovered human population
H_S	Susceptible human population
μ_V	Natural mortality rate of mosquito
μ_H	Natural birth/mortality rate of human
$\widetilde{\mu_v}$	Mortality of *Wolbachia*-carrying mosquitoes
N_H	Human population
N_V	Mosquito population
N_V^c	Total *Wolbachia*-carrying mosquito population

N_V^f	Total *Wolbachia*-free mosquito population
q_1	Ratio of death rate *Wolbachia*-carrying cf. *Wolbachia*-free
q_2	Ratio of transmission probability *Wolbachia*-carrying cf. *Wolbachia*-free
Q_i	Number of mosquito embryos used for transinfection
V_I	Infected mosquito population
V_I^c	Infected *Wolbachia*-carrying mosquito population
V_I^f	Infected *Wolbachia*-free mosquito population
V_S	Susceptible mosquito population
V_S^c	Susceptible *Wolbachia*-carrying mosquito population
V_S^f	Susceptible *Wolbachia*-free mosquito population

LIST OF ABBREVIATIONS

ADE	Antibody-Dependent Enhancement
ADF	Augmented Dickey-Fuller
AI	Artificial Intelligence
AIC	Akaike Information Criterion
AR	Autoregressive
AR6	Sixth Assessment Report
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average Exogenous Variable
ARMA	Autoregressive Moving Average
AUC	Area Under the Curve
BIC	Bayesian Information Criterion
CI	Cytoplasmic Incompatibility
DENV	Dengue Virus
DL	Deep Learning
EIP	Extrinsic Incubation Periods
ETS	Error, Trend, Seasonal
GIS	Geographic Information Systems
GWR	Geographically Weighted Regression
IGRs	Insect Growth Regulators
IPCC	Intergovernmental Panel on Climate Change
LSTM	Long Short-Term Memory
MA	Moving Average

MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCO	Movement Control Order
MetMalaysia	Malaysian Meteorological Department
ML	Machine Learning
MLR	Multiple Linear Regression model
MMRE	Mean Magnitude of Relative Error
MOH	Malaysian Ministry of Health
MRA	Multiresolution Analysis
MRE	Magnitude of Relative Error
OLS	Ordinary Least Squares
ODEs	Ordinary Differential Equations
RNN	Recurrent Neural Network
SARIMA	Seasonal Autoregressive Integrated Moving Average
SI-SIR	Susceptible-Infective for vector populations; Susceptible Infective-Recovered for human populations
SIR	Susceptible Infective-Recovered
SSA-LSTM	S-LSTM With Spatial Attention
SVR	Support Vector Regression
TBATS	Trigonometric seasonality, Box-Cox transformation, ARMA errors
TCI	Temporary Cross-Immunity
VAR	Vector Autoregressive
WHO	World Health Organization

MODEL HIBRID KESAN IKLIM TERHADAP RAMALAN DAN KAWALAN DENGGI DI SELANGOR, MALAYSIA

ABSTRAK

Pendekatan ramalan denggi yang sedia ada di Malaysia terhad oleh dua kelemahan utama: model deterministik lazimnya mengandaikan kadar gigitan nyamuk yang malar, dengan mengabaikan variasinya yang dipengaruhi oleh iklim, manakala model siri masa statistik pula biasanya hanya berkesan untuk ramalan jangka pendek. Kekangan ini menyukarkan ramalan tepat dinamik wabak dan mengehadkan keberkesanan sistem amaran awal. Kajian ini membangunkan dua pendekatan bersepadu: (i) model gandingan Susceptible-Infective bagi populasi vektor dan Susceptible-Infective-Recovered bagi populasi manusia (SI-SIR) dengan model Autoregressive Integrated Moving Average dengan pemboleh ubah eXogenous (ARIMAX), dan (ii) model SI-SIR dengan Regresi Linear Berbilang dan Long Short-Term Memory (MLR-LSTM). Negeri Selangor dipilih sebagai lokasi kajian. Model SI-SIR terlebih dahulu menganggarkan kadar gigitan nyamuk daripada data denggi. Kemudian, model gabungan SI-SIR dan ARIMAX menghubungkan kadar gigitan dengan suhu, kelembapan, dan taburan hujan, serta menggunakan ramalan kadar gigitan ini untuk meramalkan kes denggi secara mingguan. Model gabungan ini mencapai ramalan 4 minggu dengan purata Mean Magnitude of Relative Error (MMRE) sebanyak 8.01%. Model SI-SIR dengan MLR-LSTM seterusnya melanjutkan tempoh ramalan kepada 20–60 minggu dengan menggabungkan ramalan MLR yang dipacu oleh iklim dan LSTM untuk baki ramalan, sekali gus meningkatkan ketepatan ramalan semasa tempoh yang stabil. Strategi kawalan vektor turut dinilai dalam rangka kerja ini. Kaedah konvensional, seperti semburan nyamuk (fogging), penghapusan tempat pembiakan nyamuk, dan penggunaan larvasid, memerlukan intensiti tinggi yang berterusan, yang sukar dilaksanakan dalam praktik. Kaedah baharu seperti intervensi berasaskan *Wolbachia* menunjukkan potensi, namun migrasi nyamuk mungkin menjejaskan keberkesanannya.

HYBRID MODEL OF CLIMATE EFFECTS ON DENGUE FORECAST AND CONTROL IN SELANGOR, MALAYSIA

ABSTRACT

Existing dengue forecasting approaches in Malaysia are constrained by two main limitations: deterministic models typically assume a constant mosquito biting rate, neglecting its climate-driven variability, while statistical time series models are usually effective only for short-term forecasts. These constraints hinder the accurate prediction of outbreak dynamics and limit the utility of early warning systems. This study develops two integrated approaches: (i) coupled Susceptible-Infective for vector populations and Susceptible-Infective-Recovered for human populations (SI-SIR) and Autoregressive Integrated Moving Average with eXogenous variables (ARIMAX) model, and (ii) SI-SIR with Multiple Linear Regression and Long Short-Term Memory (MLR-LSTM) model. Selangor was selected as the study site. The SI-SIR model first estimates the mosquito biting rate from dengue data. Then the coupled SI-SIR and ARIMAX model links biting rate to temperature, humidity, and precipitation, and uses predicted biting rates to forecast dengue cases with weekly updates. This coupled model achieved a 4-week forecast with an average Mean Magnitude of Relative Error (MMRE) of 8.01%. The SI-SIR with MLR-LSTM model further extends prediction horizons to 20–60 weeks by combining climate-driven MLR predictions and LSTM for residuals, improving accuracy during stable periods. Vector control strategies were evaluated within this framework. Conventional methods, namely mosquito fogging, elimination of mosquito breeding grounds and larvicide application, require impractical sustained high intensity for effectiveness. Emerging method like *Wolbachia*-based intervention showed promise, but mosquito migration might compromise its effectiveness.

CHAPTER 1
INTRODUCTION

1.1 Introduction

Dengue is a mosquito-borne viral infection, which is caused by a virus called dengue virus (DENV). The dengue virus consists of four distinct serotypes, which are classifications of viruses within the same species based on their unique surface proteins. These serotypes are known as DENV-1, DENV-2, DENV-3, and DENV-4. Each serotype can cause dengue fever, and infection with one serotype provides lifelong immunity to that specific serotype but not to the others (Guzman & Harris, 2015; Messina et al., 2019). Transmission primarily occurs through two routes: mosquito-to-human and human-to-mosquito transmission. Mosquito-to-human transmission entails the transmission of the virus to humans via bites from infected female mosquitoes. Though several species can act as vectors during this process, the *Aedes aegypti* mosquito is considered the primary vector of DENV. Human-to-mosquito transmission arises when mosquitoes become infected by feeding on individuals who are viremic with DENV.

Dengue disease is often reported in tropical and sub-tropical climates worldwide, namely the regions of Africa, the Americas, the Eastern Mediterranean, Southeast Asia, and the Western Pacific (Figure 1.1). Compared to the years before 1970, the incidence of dengue has dramatically increased around the world in recent decades and it is now endemic in more than 100 countries (Ilic & Ilic, 2024). Globally, the World Health Organization (WHO) reported an eight-fold rise in dengue cases over the last two decades (2000 - 2019). It remains a major public health concern, particularly in Asia where 70% of the global burden of the disease is accounted for (Bhatt et al., 2013). The modelling results provided by Bhatt et al. (2013) indicated that 390 million dengue infections occur every year mostly in Asia while Brady et al. (2012) estimated that approximately 3.9 billion people are at risk of being infected with dengue viruses.

Similar to the global situation, Malaysia also exhibited a rising trend with a surge in dengue cases every four to five years. Shahin Mia et al. (2013) suggested that from 2000 to 2010 the number of dengue cases and the number of deaths increased by 14% and 8% per year respectively and the number of infection cases has incrementally risen since

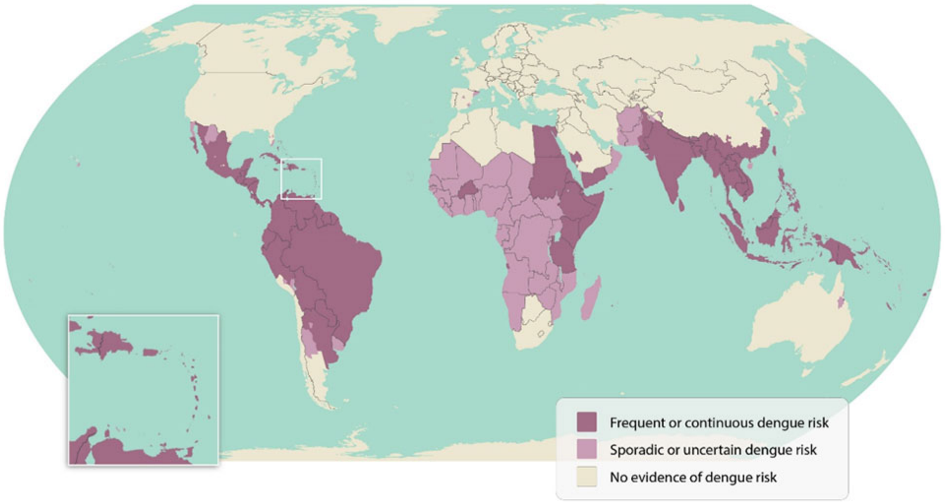

Figure 1.1: Global map of areas with dengue risk (CDC, 2024)

2014 (Pang & Loh, 2016). Figure 1.2 presents the trend of dengue cases in Malaysia from 1995 to 2023, based on data obtained from the Malaysian Ministry of Health (MOH) after the application to the National Medical Research Register. Although the number of dengue cases seems to decrease in 2020 and 2021, probably due to the pandemic lockdown effect, the number of dengue cases surged again in 2022, following the pre-lockdown trend.

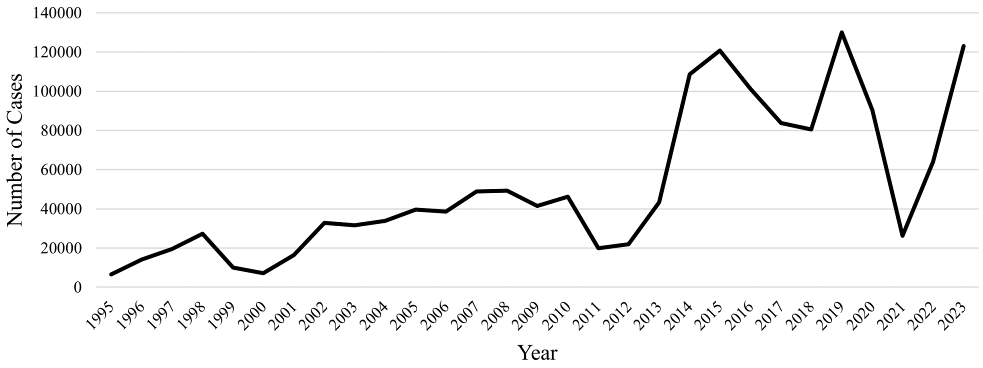

Figure 1.2: Annual dengue cases in Malaysia from 1995 to 2023 (figure plotted by the author using data provided by MOH)

1.2 Problem Statement

Climate variables such as temperature, precipitation and humidity can affect dengue transmission dynamics through their effect on vector populations. As the rising temperature and frequency of extreme weather intensifies due to climate change, dengue incidence could be further elevated. As one of Malaysia's most urbanized and populous states, Selangor consistently reports the highest number of dengue cases in the country, making it a critical hotspot for studying climate-driven dengue transmission. Dengue incurs a substantial burden on healthcare. Economic burden of dengue is the cost associated with treatment, diagnosis, outcome and prevention of dengue disease that is imposed on society (Castro et al., 2017). It is estimated that dengue causes an economic burden of USD 2.1 billion (2010 USD) annually in the Americas and USD 950 million (2010 USD) annually in 12 countries in Southeast Asia (Pitisuttithum & Bouckenooghe, 2016). In Malaysia, the total annual cost of dengue prevention and illness in 2009/2010 was USD 175.7 million (Packierisamy et al., 2015). There is a pressing need to contain dengue surges, particularly now that emerging diseases such as COVID-19, are parallel threats with extensive health care, societal, and economic burdens (Olive et al., 2020). Currently, there is no specific treatment for dengue. Although the dengue tetravalent vaccine (live attenuated), Qdenga, was officially launched in Malaysia on June 11, 2024, the cost, estimated to be MYR 500 for two doses, is deemed unaffordable to the government and general public. Further, the vaccine does not stop the transmission of dengue virus; it will only reduce the severity of the disease. Hence, early detection and effective vector control measures remain pivotal in curbing the spread of the disease and mitigating its impact on communities. Most of the studies on dengue prediction used standalone deterministic, statistical, and machine learning models, which has its limitation.

However, the existing prediction approaches mainly use standalone deterministic model or statistical model, which has its limitation. For example, statistical models like time series models are used to project dengue incidence subject to climate variability but such approach does not allow further analysis of the intervention measures such as vector control. The rate constants in deterministic models such as SIR models are often set as constant so further analysis of the effect of changing climate on dengue incidence using such models may not be effective. Further, real-time dengue forecasting has its challenges. Ideally, an abundance of reliable, refined, and continuously updated epidemiological, entomological, and environmental data is required for accurate prediction. Yet, obtaining such data poses difficulties for researchers in Malaysia due to data collection costs and data sharing policies. Though various forecasting methods have been put forward, many practical applications have shown that epidemic outbreak forecasts often fall far short of the mark, especially for long-term prediction. Hence, it is imperative to develop a reliable model for forecasting dengue cases that is practical in terms of data requirement and allows prediction over an extended period. Given Selangor's vulnerability to the dengue burden, it serves as an ideal study site to develop and validate a robust dengue forecasting model. Such model would enable early detection of evolving dengue trends and facilitate the planning and implementation of mitigation measures, especially in a changing climate.

1.3 Significance of the Study

Dengue prediction models provide a means for stakeholders in the public health system to detect potential outbreaks in advance for timely and strategic implementation of vector control measures. However, most dengue prediction studies used standalone deterministic, statistical, and machine learning models, which may either require an extensive amount of data or be less robust in the sense that it cannot be further adapted for investigating the effectiveness of various dengue mitigation measures subject to a changing climate. To address such model limitations, this study first developed a coupled deterministic SI-SIR (Susceptible-Infective for vector populations and Susceptible-Infective-Recovered for human populations) and statistical ARIMAX (Autoregressive Integrated Moving Average with eXogenous variables) model that can forecast dengue cases based on temperature, humidity, and precipitation. Further, most of the studies on dengue prediction demonstrated the capacity to forecast dengue incidence spanning from 15 days to 4 months. It is desirable to have a credible forecast of dengue cases for an extended period of time to allow early detection of changing dengue trend and to allow appropriate mitigation strategies to be planned. To extend the temporal horizon of dengue forecast, the model was revised into an approach that coupled the deterministic SI-SIR model with the statistical Multiple Linear Regression model (MLR) with an element of deep learning (Long Short-Term Memory (LSTM)). With heavy reliance on vector eradication for curbing dengue outbreaks in Malaysia, the models developed in this study allowed the assessment of intervention strategies and their feasibility under the changing conditions of climate variables. The models developed in this study were calibrated and validated using the meteorological and dengue data for Selangor, a state in Malaysia with the highest number of dengue cases. At the time of writing, there is no published research that used such coupled models. It is envisaged that the models developed in this study can be extended to other study sites and to other vector-borne diseases.

1.4 Research Questions

The following research questions were addressed in this study.

1. How statistical and deterministic models can be coupled to project dengue cases subject to climate variables?
2. How the approach can be improved to provide a longer-term forecast of dengue incidence?
3. What is the level of vector control measures required for lower dengue incidence?

1.5 Objectives of Thesis

The objectives of this thesis are as follows:

1. To formulate a coupled deterministic SI-SIR and statistical ARIMAX model for forecasting dengue incidence in Selangor subject to climate variables.
2. To develop an approach integrating SI-SIR, MLR, and LSTM that extends the temporal horizon of dengue forecast.
3. To analyze potential intervention strategies, including mosquito fogging, elimination of mosquito breeding grounds and larvicide application, and *Wolbachia*-based intervention for dengue control by means of model simulations.

1.6 Scope and Organization of Thesis

In this section, the scope of the research and the organization of the thesis are elaborated.

1.6.1 Scope of the Research

This study focuses on the state of Selangor, a densely populated and highly urbanized region located on the west coast of Peninsular Malaysia, which consistently records the highest number of dengue cases nationwide. Covering the period from January 2014 to December 2020, this research investigates the relationship between climate variables (specifically temperature, humidity, and precipitation) and dengue transmission mediated through mosquito biting rate. For this purpose, two hybrid modeling approaches were developed, namely: a coupled SI-SIR and ARIMAX model designed for real-time prediction of dengue cases, and a SI-SIR model integrated with multiple linear regression and long short-term memory (MLR-LSTM) to enable extended forecasts ranging from 20 to 60 weeks. Additionally, the study analyzes both conventional (mosquito fogging, elimination of mosquito breeding grounds and larvicide application) and novel (*Wolbachia*-based intervention) vector control strategies by simulating their effects within these modeling frameworks.

1.6.2 Organization of the Thesis

This thesis is organized as follows:

Chapter 1 makes an introduction to dengue and the climate effects on dengue transmission, drawing insights into the important roles played by temperature, humidity, and precipitation. This is followed by the problem statement, which emphasizes that reliable, adaptable forecasting models that can integrate various data sources and support extended predictions are crucial for early detection and effective interventions of dengue disease, reducing public health and economic burdens. The significance, research questions, and objective of the study are also stated.

Chapter 2 provides a comprehensive review of the related literature, beginning with two main types of models used in dengue research: deterministic models and statistical models. Representative applications of single-serotype and multi-serotype models within the deterministic framework is introduced and commonly used statistical methods for dengue research, including time series analysis, spatial statistics, machine learning, and deep learning approaches are compared in the chapter. Climate effects on dengue transmission are also reviewed to underscore the connections between climate and dengue. Given the ongoing challenges brought by climate influences, the chapter also makes a closer examination of some existing vector control strategies.

In Chapter 3, the relationship between climate and dengue transmission is examined through the mosquito biting rate. The development of the coupled model, combining the statistical ARIMAX model and the deterministic SI-SIR model, to enable real-time forecasting of dengue cases using climate variables including average temperature, relative humidity, and accumulated precipitation was detailed in the chapter. Further, the model forecasting performance is assessed and compared to that of the stand-alone ARIMAX model.

Chapter 4 presents the SI-SIR with MLR-LSTM model developed to extend the temporal horizon of dengue prediction. The forecasting performance of model is evaluated and compared to three alternative methods in the chapter. Forecasting is also performed and assessed for various prediction periods.

Chapter 5 explores both conventional and emerging vector control methods in mitigating dengue outbreaks using the model developed in Chapter 4. Conventional methods discussed include mosquito fogging, and the elimination of mosquito breeding grounds and the application of larvicides while emerging vector control method involves the use of *Wolbachia*-infected mosquitoes.

Chapter 6 summarizes the work performed to address the research questions, which led to the achievement of the research objectives. Some recommendations for future research are also provided.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

This chapter provides a review of various models for dengue research. These models can be broadly categorized into two main genres: deterministic models and statistical models. Deterministic models are further divided into single-serotype models and multi-serotype models, each offering unique insights and applications in understanding dengue dynamics. Statistical models contain three prevalent methodologies: time series analysis, spatial statistics, and machine learning (ML) and deep learning (DL) approaches, each offering distinctive advantages in analyzing dengue-related data. Additionally, this chapter conducts a comprehensive review of two significant vector control measures: mosquito fogging and *Wolbachia*-based intervention. By examining these intervention strategies, insights into effective approaches for managing dengue diseases are provided.

2.2 Deterministic Dengue Models

Numerous deterministic models have been published to investigate dengue transmission dynamics. They serve as vital tools for understanding dengue transmission dynamics and facilitating the exploration of intervention strategies and vaccination policies to combat this infectious disease effectively. Generally, these deterministic models can be categorized into two genres: single-serotype and multi-serotype models.

2.2.1 Single-serotype models

As a simple vector-host transmission model, Equation (2.1) (Bailey, 1975) often serves as the basis for dengue models addressing a single serotype. In this existing model, H_S, H_I, and H_R represent the numbers of susceptible, infected, and recovered humans while V_S and V_I represent the numbers of susceptible and infected mosquitoes, respectively. B is the mosquito biting rate, β_{HV} represents the probability of human-to-mosquito transmission and β_{VH} represents the probability of mosquito-to-human transmission. The natural mortality rate of mosquitoes, the natural mortality rate of humans, and the recovery rate of infected humans are notated as μ_V, μ_H, and γ_H, respectively. The vector recruitment rate A was assumed to be constant.

$$
\begin{aligned}
\frac{dH_S}{dt} &= \mu_H N_H - B\beta_{VH}\frac{V_I}{N_H}H_S - \mu_H H_S,\\
\frac{dH_I}{dt} &= B\beta_{VH}\frac{V_I}{N_H}H_S - \gamma_H H_I - \mu_H H_I,\\
\frac{dH_R}{dt} &= \gamma_H H_I - \mu_H H_R,\\
\frac{dV_S}{dt} &= A - B\beta_{HV}\frac{H_I}{N_H}V_S - \mu_V V_S,\\
\frac{dV_I}{dt} &= B\beta_{HV}\frac{H_I}{N_H}V_S - \mu_V V_I.
\end{aligned}
\tag{2.1}
$$

Based on the model shown in Equation (2.1), various derived models have been developed to study dengue transmission with assumptions related to population representations and vaccinations. Population representation enhancements include the incorporation of subpopulations such as travelers, as demonstrated by Pongsumpun et al. (2004) and Polwiang (2016), who observed correlations between infection risk and duration of stay in endemic areas. Age-structured models, integrating different age classes within the host population, have also been developed (Pongsumpun & Tang, 2003; Supriatna et al., 2008; Kristiani et al., 2017).

In addition to human populations, vector populations can be described across various stages of development, allowing for nuanced analyses of vector control strategies. Yang and Ferreira (2008) considered mosquito maturation stages (eggs, larvae, pupae, adults), relaxing the assumption of a constant recruitment rate. Erickson et al. (2010) and Luz et al. (2011) further refined these models by dividing the mosquito population into pre-adult and adult stages, considering factors like insecticide resistance and fitness cost. Since the introduction of the first dengue vaccine, Dengvaxia, vaccination effects have also been integrated into mathematical models. Chanprasopchai et al. (2018) incorporated vaccination effects into a SIR (Susceptible-Infected-Recovered) model, demonstrating significant reductions in hospitalization time with Dengvaxia vaccination. Furthermore, Tay et al. (2021) developed a pricing model to estimate the economic burden of dengue and establish an appropriate threshold price for the vaccine in Malaysia.

In addition to single-serotype models, multi-serotype models, which are reviewed in the following section, represent another important category of deterministic models.

2.2.2 Multi-serotype models

Multi-serotype models are expansions of single-serotype models, given the co-circulation of four serotypes and the lack of long-term cross-immunity. These models, focusing on phenomena like Antibody-Dependent Enhancement (ADE) and Temporary Cross-Immunity (TCI) (Andraud et al., 2012; Aguiar et al., 2022), are broadly categorized into vector-host transmission and host-to-host transmission models.

Vector-host transmission models aim to elucidate interactions between vectors, hosts, and pathogens, thereby informing disease control strategies including vector control and vaccination. Rashkov and Kooi (2021) proposed a model incorporating TCI for hosts and secondary infections, evaluating endemic equilibria and conducting bifurcation analyses. Similarly, Lourenço and Recker (2013) developed a multi-strain vector-host model considering explicit mosquito vector dynamics, TCI, and seasonal mosquito biting patterns. Knerer et al. (2015) extended this approach, incorporating seasonality, age structure, and sequential infections by all four serotypes, successfully capturing national dengue data from Thailand. Their work also assessed the combined impact of vector control and vaccination, demonstrating the effectiveness of vaccination even with low efficacy. Subsequently, Knerer et al. (2020) explored the cost-effectiveness of various dengue control measures using the same model. Knipl and Moghadas (2015) investigated a two-serotype vector-host model incorporating ADE, cross-protection, and seasonality, highlighting challenges in eradicating dengue with vaccines of varying efficacy. Maier et al. (2017) considered ADE in determining optimal vaccination ages for each dengue serotype, revealing variations based on serotype circulation patterns in Brazil.

Host-to-host transmission models seek to understand dengue virus spread among individuals. Erickson et al. (2010) developed a two-serotype model assuming rapid transmission and dense mosquito populations, aiming to elucidate cross-reactive antibody effects on secondary infections. Building upon Ferguson's model, Billings et al. (2008) incorporated two vaccination schemes, analyzing the impact of single serotype versus dual-strain vaccination strategies. Moreover, studies like those by Recker et al. (2009), Wikramaratna et al. (2010), and Bianco and Shaw (2011) expanded these models to accommodate more than two serotypes, assessing the influence of four-serotype co-circulation on infection dynamics.

2.3 Statistical Dengue Models

Statistical models for dengue transmission contain a diverse range of methodologies, including time series analysis, spatial statistics, and ML and DL techniques. These models use available data on dengue cases, climate variables, population demographics, and vector abundance to quantify the relationships between these factors and predict disease dynamics. While traditional epidemiological models, such as deterministic models, provide insights into the mechanisms of disease transmission, statistical models offer complementary approaches by capturing complex nonlinear relationships and heterogeneities in data.

2.3.1 Time series analysis

Time series analysis is a widely used statistical approach for modelling the temporal dynamics of dengue transmission and it helps uncover patterns in data by its very nature, which can then be used to predict future data points. It contains a suite of powerful techniques designed to analyze and forecast sequential data points, often encountered in fields such as economics, finance, meteorology, and beyond. At the core of this analysis approach mainly lie Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), and Seasonal Autoregressive Integrated Moving Average (SARIMA) models.

AR models represent a fundamental concept in time series analysis, capturing the relationship between an observation and a linear combination of past observations. The order of an AR model, denoted by p, specifies the number of lagged observations considered in predicting the current value. While MA models emphasize the relationship between an observation and a linear combination of past error terms or shocks to the system. The order of an MA model, denoted by q, determines the number of lagged error terms included in the model. Combining aspects of both AR and MA models, the ARMA model blends these two approaches to provide a comprehensive framework for modelling time series data. ARMA models are adept at capturing both short-term dependencies represented by autoregressive components and the influence of random shocks represented by moving average terms. For non-stationary time series data, where the mean and variance exhibit trends or shifts over time, the ARIMA model proves invaluable. By incorporating differencing operations to induce stationarity, ARIMA models can effectively model and forecast a wide range of time series data. ARIMA models can be viewed as a generalization of ARMA models. Specifically, when the order of differencing d in an ARIMA model is zero (i.e., no differencing is applied), the ARIMA model reduces to an ARMA model. Furthermore, in situations where time series data exhibit seasonal patterns or trends, SARIMA models offer an extension to the ARIMA framework. SARIMA models introduce seasonal differencing along with seasonal autoregressive and moving average terms, enabling the modelling of complex seasonal dynamics present in the data. These techniques of time series analysis occur frequently in literature among different fields, including studies of dengue disease.

For instance, Cheng et al. (2020) utilized an AR model to assess the association between daily weather variability and daily dengue case numbers in Kaohsiung, Taiwan, during the years 2014 and 2015. Their results indicated an association between dengue cases and humidity, emphasizing the simplicity and utility of the developed model for real-time prediction and public health decision-making. Dom et al. (2013) employed ARIMA models based on data collected from 2005 to 2010, utilizing weekly variation to predict dengue incidence in Malaysia for 2010. Their findings underscored the efficacy of ARIMA models in disease control and prevention programs. Similarly, Naher et al. (2022) found ARIMA models outperforming other techniques such as Error, Trend, Seasonal (ETS), and Trigonometric seasonality, Box-Cox transformation, ARMA errors (TBATS) models in forecasting dengue incidences in Bangladesh. In Malaysia, the multivariate Poisson regression models, SARIMA models, and SARIMA models with external regressors have been employed by Jayaraj et al. (2019) to analyze the relationship between weather predictors and dengue incidence in Tawau. The SARIMA model was found to be the most effective in their study for forecasting future outbreaks in Tawau, which would enhance decision-making in public health settings. To sum up, Table 2.1 compares key characteristics of different time series models.

Table 2.1: Comparison of time series forecasting models

Model	Advantages	Disadvantages
ARIMA	• Implement easily • Handle non-stationary data effectively • Require low computational cost	• Fail to incorporate external variables • Capture only temporal patterns • Perform poorly during sudden outbreaks
SARIMA	• Capture seasonal patterns • Model periodic disease cycles effectively • Extend ARIMA capabilities	• Require complex parameter tuning • Exclude climate-related factors • Depend on long historical data
ARIMAX	• Incorporate climate covariates • Enable causal analysis • Combine temporal and external influences	• Depend on high-quality external data • Involve more parameters to estimate • Demand higher computational resources

In this study, the ARIMAX model was selected because it can uniquely integrate temperature, humidity, and precipitation as direct inputs while retaining the temporal forecasting strengths of the ARIMA model. Additionally, by including climate covariates, the ARIMAX model allows the identification and quantification of the influence of specific environmental factors on dengue incidence, thereby offering insights into potential causal relationships between climate conditions and disease transmission.

2.3.2 Spatial statistics

Spatial statistical models are instrumental in elucidating the spatial distribution of dengue cases and identifying environmental risk factors associated with disease transmission. These models utilize geostatistical techniques, such as spatial autocorrelation analysis, spatial regression models, and cluster detection methods, to discern spatial patterns of dengue incidence and pinpoint high-risk areas. By integrating geographic information systems (GIS) data with epidemiological and environmental variables, spatial statistical models identify hotspots of dengue transmission and associated risk factors and contribute to targeted interventions and resource allocation for vector control and public health interventions.

For example, to investigate the spatial and temporal trends of dengue outbreaks in Curaçao, Roelofs et al. (2024) utilized geocoded address information alongside dengue infection data. They applied spatial autocorrelation analysis (Moran's I index) to identify spatial autocorrelation among incident locations, revealing a concentration of cases in densely populated areas and highlighting a correlation between population density and dengue incidence. In a study conducted in the Dhaka metropolitan city in Bangladesh (Sharif et al., 2022), spatial regression models including ordinary least squares (OLS) regression, geographically weighted regression (GWR), and count data regression methods were employed to analyze dengue case locations. Their findings suggested no direct association between dengue case locations and vector distribution or density, prompting further exploration of other transmission risk factors. Together, these studies highlighted different dimensions of dengue risk factors across various settings. The studies by Sharif et al. (2022) and Roelofs et al. (2024) underscored the importance of spatial trends and risk factors in densely populated areas while Abd Naeeim and Abdul Rahman (2022) added to this understanding by emphasizing the seasonal influence on dengue clustering, particularly during monsoon periods in Malaysia. To identify spatial, temporal, and spatiotemporal clusters of dengue in Peninsular Malaysia from 2015 to 2017, Abd Naeeim and Abdul Rahman (2022) utilized two space-time scan statistics, Kulldorff's cylindrical space-time scan statistic and Takahashi et al.'s prismatic space-time scan statistic, to detect clustering in the data. Their results revealed spatio-temporal clusters of dengue occurring in densely populated areas throughout the year, particularly during the monsoon season, suggesting a plausible association between meteorological factors and dengue incidence. This identification of high-risk locations and time periods enables Malaysian public health authorities to make informed decisions regarding vector control measures and dengue awareness campaigns.

Compared to time series analysis and spatial statistics, machine learning and deep learning have experienced rapid development and widespread application in recent years. Relevant literature concerning its use is examined in the following section.

2.3.3 Machine learning and deep learning approaches

In recent decades, there has been a remarkable surge in the utilization of ML and DL techniques, both of which are types of artificial intelligence (AI) that enable computers to learn essential features from datasets and predict outcomes of specific events (Balakumar et al., 2022). Consequently, ML and DL approaches offer promising avenues for modelling the complex relationships underlying dengue transmission dynamics and multiple predictor variables.

In ML algorithms, supervised learning techniques, including decision trees, random forests, support vector machines, and neural networks, are increasingly employed to develop predictive models for dengue incidence. These models utilize large datasets containing dengue cases, climate data, land use patterns, and socioeconomic indicators to forecast disease outbreaks and evaluate the effectiveness of interventions. Furthermore, unsupervised learning methods, such as clustering algorithms, aid in identifying subgroups within dengue data, facilitating the characterization of high-risk populations and targeted intervention strategies. These algorithms have been successfully applied in dengue prediction (Gangula et al., 2023). Among these research, Hoyos et al. (2021) also extensively discussed and compared the advantages and limitations of each algorithm in terms of accuracy, sensitivity, specificity, and area under the curve (AUC).

DL, compared to traditional ML techniques, is gaining popularity due to its superior ability to handle complex patterns and capture non-linear relationships for predictive forecasting (Shrestha & Mahmood, 2019). In recent years, researchers have increasingly employed deep DL techniques, such as long short-term memory (LSTM), for dengue prediction. For example, Doni and Sasipraba (2020) developed a dengue prediction model for India utilizing climatic conditions, and their LSTM-based model achieved forecasting accuracies of over 89% for infection and 81% for mortality. Saleh and Baiwei (2021) compared the performance of LSTM and Support Vector Regression (SVR) on dengue prediction using weather and climate data and found LSTM to be superior in capturing the rise and fall trends of dengue. Similarly, Lestari et al. (2022) compared LSTM with six other methods and found LSTM to have superior forecasting accuracy in predicting the number of dengue cases in Malang Regency, Indonesia. Majeed et al. (2023) developed and compared six different LSTM models for dengue prediction based on climate, topographic, demographic, and land-use variables, with their study suggesting that the S-LSTM with spatial attention (SSA-LSTM) model exhibited the best forecasting performance, underscoring the effectiveness of LSTM models in capturing complex patterns and making accurate predictions.

Although ML and DL methods have shown considerable promise in capturing relationships between dengue incidence and diverse predictor variables, some limitations remain. Traditional ML models often rely on feature engineering and may struggle with the complex temporal dependencies of disease spread, whereas DL models can be computationally demanding and less interpretable despite their superior accuracy with time-series data (Makridakis et al., 2018; Sharma et al., 2020). To address these gaps, an integrated approach combining deterministic, statistical, and deep learning models is proposed in this thesis. Tested with dengue data from Selangor, Malaysia, this method uses multiple linear regression (MLR) for initial predictions of mosquito biting rates based on climate factors, enhances accuracy by fitting MLR residuals with LSTM, and employs an SI-SIR model to predict dengue cases. By taking advantage of the strengths of each method, this approach aims to improve both accuracy and interpretability in dengue forecasting, providing a more robust tool for public health applications.

While statistical, spatial, and machine learning models have proven invaluable for understanding dengue transmission and predicting outbreaks, they often rely on a range of climate and environmental factors that play critical roles in the disease's dynamics. Among these factors, temperature, relative humidity, and precipitation emerge as some of the most influential, directly and indirectly affecting mosquito development, survival, and transmission potential. Recognizing these factors' unique contributions can enhance model precision and aid in designing targeted interventions. Therefore, it becomes essential to understand how specific meteorological variables shape the patterns of dengue transmission in different regions. The following Section 2.4 reviews the multiple effects of climate on dengue dynamics, highlighting recent findings that underscore the connections between climate and dengue.

2.4 Climate Effects on Dengue Transmission

Meteorological variables have long been recognized as critical factors driving dengue outbreaks (Alkhaldy, 2017; Nosrat et al., 2021). Among these variables, temperature, humidity, and precipitation directly influence all phases of a mosquito's life cycle and can indirectly impact the dengue transmission cycle (Nuraini et al., 2021).

Temperature plays a pivotal role in the development and survival of *Aedes* mosquitoes, influencing essential biological processes such as reproduction, biting rates, and developmental rates (Liu et al., 2023). The optimal temperature range for their development spans from 25°C to 30°C; however, temperatures exceeding 40°C result in adult mosquito mortality, with eggs and larvae failing to develop (Reiskind & Zarrabi, 2012). While rising temperatures generally coincide with increased *Aedes* mosquito activity, excessively high temperatures can shorten their lifespan and reduce population size (Myer et al., 2020). *Aedes aegypti* faces challenges in extreme temperature conditions, exhibiting reduced mobility and feeding capability below 11°C or above 36°C, ultimately leading to mortality. Additionally, prolonged survival requires regular blood meals, with survival rates plummeting after 2-3 days without feeding, especially in tropical temperatures (Alto & Bettinardi, 2013). During the developmental stage, temperature variations manifest in external traits such as wing length and adult size. Higher temperatures (24°C-29°C) and adequate nutrition foster mosquito ingestion, resulting in shorter wings and heavier weight in *Aedes* mosquitoes, whereas lower temperatures (14°C-19°C) and inadequate food yield mosquitoes with longer wings and lower weight (Reiskind & Zarrabi, 2012; Alto & Bettinardi, 2013). Moreover, in regions with an average annual temperature of 22°C, the weekly production of *Aedes* mosquito eggs peaks when the Daily Mean Temperature Range falls between 12°C and 18°C, declining thereafter (Betanzos-Reyes et al., 2018). Extreme temperatures exceeding 36°C can significantly diminish egg production (Marinho et al., 2016).

Relative humidity also plays a significant role in the mosquito life cycle. The combination of heat and moisture profoundly influences mosquito feeding patterns and interactions, with increasing relative humidity prolonging vector survival rates and feeding duration on infective hosts (Rowley & Graham, 1968; McMichael et al., 1996; Thu et al., 1998). Conversely, reduced relative humidity negatively impacts adult survival, potentially decreasing the proportion of the vector population surviving extrinsic incubation periods (EIP) (Christophers, 1960). Relative humidity influences various aspects of mosquito behaviour, including longevity, mating, dispersal, feeding, and egg-laying behaviour, thereby impacting dengue transmission rates (McMichael et al., 1996; Hales et al., 1999; Mellor & Leake, 2000). Mosquitoes exhibit longer lifespans and disperse further under high relative humidity conditions, increasing the likelihood of infecting multiple hosts before succumbing to mortality. Additionally, relative humidity affects the evaporation rates of vector breeding sites, contributing to geographical variations in dengue transmission (Promprou et al., 2005).

Precipitation emerges as another critical factor affecting mosquito populations and dengue transmission patterns. While abundant rainfall negatively correlates with dengue incidence by flushing away mosquito eggs, it can also create suitable breeding environments for mosquitoes (de Melo et al., 2012; Méndez-Lázaro et al., 2014). In some instances, increased rainfall may bolster larval habitats and vector populations by creating new breeding sites (Gubler et al., 2001; Woodruff et al., 2002; Kelly-Hope et al., 2004). The intensity and pattern of rainfall are crucial, as heavy rains can displace mosquito larvae from breeding sites or destroy them, while lighter rain replenishes existing breeding sites and maintains higher levels of relative humidity (McMichael et al., 1996). In dry seasons with limited rainfall, mosquito habitats form when rivers dry up, leaving behind pools of water conducive to breeding, thus facilitating dengue transmission (Gubler et al., 2001). Additionally, while manmade water containers like abandoned tires and trash receptacles may not consistently affect dengue transmission, domestic water storage practices directly influence the availability of larval-rearing sites (Wu et al., 2009; Jansen & Beebe, 2010).

Knowing the association between climate and mosquito is essential for controlling dengue outbreak because it facilitates the development of management strategies for vector populations. Vector control measures specifically aim to reduce mosquito abundance and interrupt transmission cycles, addressing both environmental and biological factors that sustain dengue spread. Given the ongoing challenges posed by climatic influences, a closer examination of effective vector control strategies is warranted, the findings of which is summarized in the following Section 2.5.

2.5 Vector Control

Vector control plays a pivotal role in mitigating the transmission of dengue. Effective vector control strategies aim to reduce mosquito populations, thereby minimizing the risk of dengue transmission to humans. Conventional vector control methods encompass a variety of approaches, including the use of insecticides, plant derivatives, insect growth regulators (IGRs), and pheromones (van den Berg et al., 2021). Insecticides and plant derivatives are widely used to reduce mosquito populations by directly killing adult mosquitoes or larvae. These substances target the insects at different stages of their life cycle, significantly lowering their numbers (Onen et al., 2023). However, the development of resistance in mosquito populations to these chemicals is a growing concern, as it diminishes the effectiveness of these methods over time. Additionally, the environmental impact of using chemical insecticides raises ecological concerns, prompting the need for more sustainable solutions. IGRs are another critical tool in mosquito control. IGRs work by disrupting the development of mosquito larvae, preventing them from maturing into adults capable of reproduction and biting. This method, while effective, requires consistent and repeated applications to maintain control over mosquito populations, as it only targets the immature stages of the insects. Chemical signals, such as pheromones, offer innovative ways to manage mosquito populations. Pheromones can be used to attract mosquitoes to traps or to interfere with their mating behaviours, thus reducing their ability to reproduce. This method can be highly specific and environmentally friendly, as it targets mosquito behaviours rather than killing them outright. In addition to these conventional methods, several innovative approaches have emerged. One promising strategy involves the introduction of *Wolbachia*-infected mosquitoes (Flores & O'Neill, 2018). *Wolbachia* is a type of bacteria that can inhibit mosquitoes' ability to transmit dengue virus. By releasing these infected mosquitoes into the wild, the overall infection rate can be reduced. Another novel approach includes the use of larvivorous fish and crustaceans, which naturally prey on mosquito larvae. Introducing these predators into mosquito breeding sites can significantly reduce larval populations, offering a biological control method that is both effective and environmentally sustainable (Rather et al., 2017). Collectively, these methods highlight the multifaceted approach needed to combat dengue and other vector-borne diseases. Effective and sustainable mosquito control can reduce dengue incidence and ease its impact on public health. In this thesis, two vector control methods, one conventional (mosquito fogging) and another emerging (*Wolbachia*-based interventions) were assessed.

2.5.1 Mosquito fogging

Mosquito fogging, also known as space spraying, involves the use of specialized equipment to generate a fine mist or fog of insecticide droplets. The fogging machines disperse the insecticide solution into the air in the form of tiny droplets, which remain suspended and can penetrate various areas, including vegetation and indoor spaces. Fogging is typically used for wide-area mosquito control, covering large outdoor spaces such as neighborhoods, parks, or agricultural fields. Usually, it serves as an emergency response measure during dengue outbreaks or to target adult mosquito populations in high-risk areas (Arham et al., 2023). Mosquito fogging has demonstrated short-term effectiveness in reducing mosquito densities and suppressing dengue transmission (Esu et al., 2010; Thammapalo et al., 2012; Karunaratne et al., 2013). However, the sustainability and long-term impact of fogging remains debated, with concerns about insecticide resistance, environmental impact, and cost-effectiveness. Research by Abeyasuriya et al. (2017) indicates that insecticide fogging may adversely affect non-target insects, such as pollinators, suggesting the need for controlled fogging operations and caution against indiscriminate fogging. Moreover, misconceptions about fogging may lead to complacency in practicing routine mosquito breeding site eradication Wijayanti et al. (2020).

2.5.2 *Wolbachia*-based intervention

Wolbachia is a naturally occurring bacterium that can be introduced into *Aedes* mosquitoes to reduce their ability to transmit dengue virus. When introduced into the *Aedes aegypti* mosquito species, *Wolbachia* can help reduce the transmission of the dengue disease. The mechanism through which *Wolbachia* controls dengue disease involves several factors that disrupt the virus's ability to thrive and spread within *Aedes aegypti* mosquitoes, ultimately lowering transmission to humans. This section outlines these key mechanisms:

(a) Cytoplasmic Incompatibility (CI)

Wolbachia induces a phenomenon called cytoplasmic incompatibility, which occurs when infected male mosquitoes mate with uninfected females (as shown in Figure 2.1). The embryos produced from such matings either fail to develop or die, reducing the overall mosquito population. However, when *Wolbachia*-infected males mate with *Wolbachia*-infected females, the embryos develop normally. This leads to the spread of *Wolbachia* throughout mosquito populations over time (Zabalou et al., 2004; Dorigatti et al., 2018).

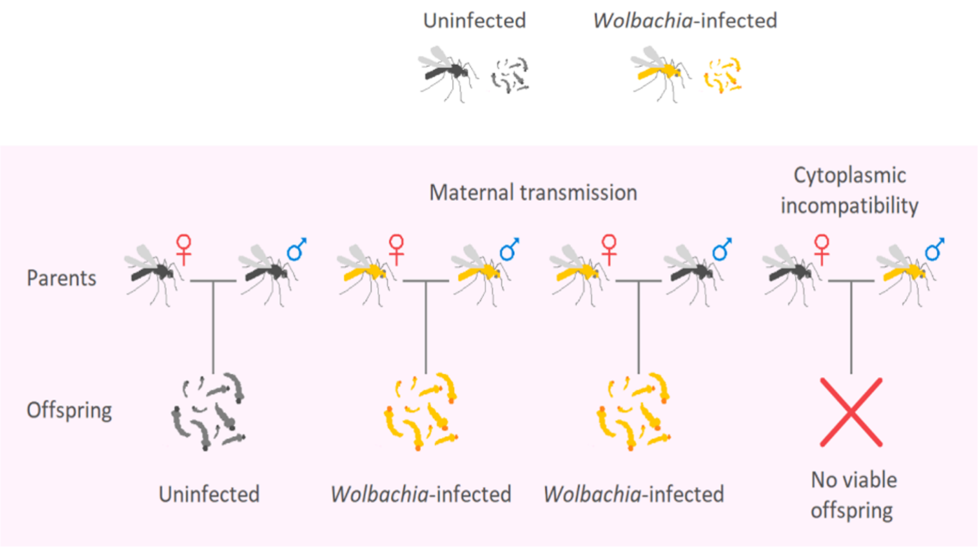

Figure 2.1: Illustration of how cytoplasmic incompatibility (CI) gives *Wolbachia*-infected females a reproductive advantage by making the progeny of infected males and wild-type females non-viable (Ross, 2021)

(b) Reduced vector competence

Studies have shown that *Aedes aegypti* mosquitoes carrying *Wolbachia* have reduced vector competence for dengue viruses. This means that they are less likely to become infected with dengue viruses when they feed on blood from infected individuals. As a result, *Wolbachia*-infected mosquitoes are less likely to transmit dengue viruses to humans when they subsequently feed on human blood (Bian et al., 2010; Majeed et al., 2023).

(c) Competition with pathogens

Wolbachia establishes a competitive relationship within the mosquito's body, occupying niches that would otherwise be exploited by dengue viruses. This competition can limit the replication and transmission of dengue viruses within the mosquito, further reducing the risk of dengue transmission to humans (Bian et al., 2013).

(d) Activation of mosquito immune responses

Wolbachia infection can also stimulate the mosquito's immune responses, enhancing its ability to fight off dengue virus infections. This activation of immune pathways within the mosquito can contribute to the reduction of dengue transmission (Kambris et al., 2009; Pan et al., 2018).

Additionally, *Wolbachia*-based interventions offer a sustainable and environmentally friendly alternative to insecticide-based vector control methods, as *Wolbachia* does not pose harm to humans or the environment. Successful invasions of two natural *Aedes aegypti* populations in Australia have demonstrated the feasibility of *Wolbachia*-based strategies for dengue suppression with potential for area-wide implementation (Hoffmann et al., 2011). Trial field releases of *Wolbachia*-carrying mosquitoes in Malaysia reported a significant reduction in dengue incidence in high-rise residential areas, indicating the feasibility and promising impact of wider-area applications (Hoffmann et al., 2024).

In addition to field trials, researchers have explored the impact of *Wolbachia* on dengue transmission dynamics through mathematical modelling. For instance, Ndii et al. (2015) developed a compartment-based mathematical model to quantify the effect of introducing *Wolbachia* into mosquito populations on human dengue cases, finding a substantial reduction in cases provided *Wolbachia*-carrying mosquitoes persist. Similarly, Li and Liu (2021) established an SEI-SIS model to study dengue transmission between humans and mosquitoes, concluding that *Wolbachia* can significantly reduce dengue fever transmission depending on its fitness effect on mosquitoes and persistence.

Conventional dengue forecasting models often suffer from limitations such as fixed rate constants in deterministic frameworks and restricted capacities for evaluating intervention strategies under changing climate conditions in purely statistical methods. To address these shortcomings, this study proposes a novel coupled modeling framework that integrates deterministic and statistical approaches for real-time predictions and combines deterministic, deep learning, and statistical methods for extended-period forecasting. By integrating these diverse modeling approaches, the proposed framework also enhances the capacity to assess and explore effective intervention strategies. These modeling components and their application in intervention analysis will be systematically introduced in Chapters 3, 4, and 5, respectively.

CHAPTER 3

COUPLED SI-SIR AND ARIMAX MODEL

3.1 Introduction

The ability to detect an outbreak of dengue in advance would assist in the timely and cost-effective deployment of vector control measures. Climate or meteorological variables, especially temperature, humidity, and precipitation, have long been recognized as critical factors driving dengue outbreaks. In time series models, a link between climate variables and dengue cases is often established so that future dengue cases can be predicted based on the forecasted climate. However, the effect of planned intervention measures cannot be predicted using such time series models. To overcome this, a dengue forecast model that incorporates the elements of a deterministic model and a statistical model is proposed. In this chapter, the deterministic SI-SIR model is coupled with the statistical ARIMAX model to enable the prediction of dengue cases based on forecasted temperature, humidity, and precipitation. The novelty of this approach lies in the combination of a deterministic epidemic model (SI-SIR) with a statistical time series model (ARIMAX), enabling the integration of climate-driven dynamics and transmission mechanisms within a unified dengue forecasting framework. The biting rate is chosen as the coupling parameter because it directly reflects mosquito activity and transmission potential, while also being sensitive to climate variations. By dynamically linking climate forecasts to the biting rate, the coupled model captures both environmental drivers and transmission dynamics, bridging the gap between statistical correlations and mechanistic processes. The model was calibrated and validated using the dengue and meteorological data for the state of Selangor in Malaysia. Selangor was chosen as the study site because this state has the highest number of dengue cases and its meteorological data such as average temperature, relative humidity, and accumulated precipitation were readily available. Notably, the work in this chapter has been published in Lu et al. (2024).